Encyclopedia of Arabic Language and Linguistics

General Editor
Kees Versteegh

Associate Editors
Mushira Eid
Alaa Elgibali
Manfred Woidich
Andrzej Zaborski

Volume V
Index

BRILL

Encyclopedia of Arabic Language and Linguistics

Volume I

A-Ed

GENERAL EDITOR
Kees Versteegh
(University of Nijmegen)

ASSOCIATE EDITORS
Mushira Eid
(University of Utah)
Alaa Elgibali
(University of Maryland)
Manfred Woidich
(University of Amsterdam)
Andrzej Zaborski
(University of Cracow)

ADVISORY BOARD
Ramzi Baalbaki *(American University in Beirut)*
Elsaid Badawi *(American University in Cairo)*
Dominique Caubet *(INALCO, France)*
Clive Holes *(University of Oxford)*
Manfred Kropp *(Orient-Institut Der Deutschen Morgenländischen Gesellschaft)*
Jérôme Lentin *(Institut National des Langues et Civilisations Orientales)*
John McCarthy *(University of Massachussetts)*
Jamal Ouhalla *(University College Dublin)*
Jan Retsö *(Göteborg University)*
Sabah Safi *(King Abdulaziz University)*

COPY EDITOR
Margaret Owen

ENCYCLOPEDIA OF ARABIC LANGUAGE AND LINGUISTICS

VOLUME I
A-Ed

General Editor
Kees Versteegh

Associate Editors
Mushira Eid
Alaa Elgibali
Manfred Woidich
Andrzej Zaborski

BRILL
Leiden – Boston
2006

Illustrations to 'Arabic Alphabet: Origin'
The President and Fellows of Harvard College:
Ia - Ib: from HOLLIS # 001008865, IIa - IIb: from HOLLIS # 000605524, IIIa - IIIb: from HOLLIS # 002238303, IVa - IVb: from HOLLIS # 005292920, V: from HOLLIS # 003062799, VIa - VIb: from HOLLIS # 001350425

Brill Academic Publishers has done its best to establish rights for the use of the illustrations printed on/in this volume. Should any other party feel that its rights have been infringed, we would be glad to hear from them.

Copyright © 2006 by Koninklijke Brill NV, Leiden, The Netherlands
Koninklijke Brill NV incorporates the imprints Brill Academic Publishers, Martinus Nijhoff Publishers and VSP.

www.brill.nl

All rights reserved. No part of this publication may be reproduced, translated, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without prior written permission from the publisher.

Authorization to photocopy items for internal or personal use is granted by Brill provided that the appropriate fees are paid directly to The Copyright Clearance Center, 222 Rosewood Drive, Suite 910, Danvers MA 01923, USA.
Fees are subject to change.

ISBN 90 04 14976 2 (Set)
ISBN 90 04 14473 0 (Volume 1)

Printed in The Netherlands on acid-free paper.

Introduction

1. Reference tools for Arabic linguistics

For many disciplines within the field of Arabic studies major reference tools exist. The *Encyclopaedia of Islam*, especially useful for historical matters, with an emphasis on persons and places, has now embarked on its third edition. The *Encyclopaedia of the Qurʾān* covers the entire domain of Qurʾānic studies and has only one more volume to go to completion. For Arabic literature there is the *Encyclopedia of Arabic literature*, as well as the *Cambridge history of Arabic literature*. For written production in Classical Arabic Brockelmann's *Geschichte der arabischen Literatur* has been superseded for the period up until 430 A.H. by Sezgin's *Geschichte des arabischen Schrifttums* (1967–2000). For Islamic history the *Cambridge history of Islam* is a comprehensive source. There are bibliographical and biographical manuals, such as the *Index Islamicus*.

Yet, for Arabic linguistics comparable reference tools are lacking. The literature before 1983 has been recorded in Bakalla's bibliography (1983), but there has been no follow-up for the literature since then, although the general *Bibliographie linguistique* makes up for this to some extent. The standard reference grammars of → Classical Arabic (such as Howell 1883–1911; Wright 1859–1862; Reckendorff 1898–1898; Blachère and Gaudefroy-Demombynes 1952; and Fleisch 1961, 1979) are in need of revision because they are outdated; Fischer's (2002) more recent grammar is not meant to be a complete reference grammar but rather a textbook for students. For the standard handbooks on varieties of Middle Arabic see the entries on → Middle Arabic, → Christian Arabic, and → Judaeo-Arabic. For → Modern Standard Arabic the situation has improved now that the survey in three volumes by Cantarino (1974–1975) has been replaced by the reference grammar by Badawi, Carter, and Gully (2004), and by the large-scale syntax of Modern Standard Arabic by El-Ayoubi, Fischer, and Langer (1st vol. 2001, 2003), which is in the process of being published.

In the field of → lexicography the situation has improved as well, although the great project of an etymological dictionary of Arabic (→ language academies) has never materialized. The dictionary of the Deutsche Morgenländische Gesellschaft is moving slowly from the letter *kāf* toward the end of the alphabet, and there is still a need for an authoritative dictionary of Classical Arabic. For Qurʾānic Arabic the old dictionary of Penrice (1873) has finally found a successor in the form of a new dictionary by Badawi and Abdelhalim (to be published in 2006); a concise dictionary was published by Procházka and Ambros (2004). Arabic/Arabic dictionaries are being published in the Arab world, for instance the *Wasīṭ* of the Arabic Language Academy in Cairo. Large-scale dictionaries of Modern Standard Arabic now exist for all major Western languages.

For dialectology there is the *Handbuch der arabischen Dialekte* by Fischer and Jastrow (1980), which however does not deal with sociolinguistic topics, nor with the external history of the language, while the coverage of the internal development of the language and the peripheral dialects is not comprehensive. Dialect atlases exist for some of the major areas (Egypt by Behnstedt and Woidich 1985–1999; Yemen by Behnstedt 1985–1987; and Syria by Behnstedt 1997), and the introduction to Arabic dialect geography by Behnstedt and Woidich (2005) has just appeared (→ dialect geography). For individual dialects the situation varies considerably.

For a long time Bateson (1967) was the only handbook that could serve as an introduction to the entire field; it was republished in 2003 as a classic reference work but is obviously outdated. A small number of handbooks fill part of the gap, such as the *Grundriß der arabischen Philologie*, whose first volume, edited by Fischer (1982), focuses on the history and the structure of the language and on the philological study of the written documents. More recent textbooks, like the ones by Anghelescu (1995), Versteegh

(2001), Ferrando (2001), and Holes (2004), are intended for use by students, they are not exhaustive, and they deal only with selected perspectives on the Arabic language.

In short, there is no major reference tool to represent the state of the art in all aspects of Arabic linguistics. Not all relevant linguistic topics are included in the *Encyclopaedia of Islam*, and for those that are, the treatments vary in depth. Many aspects of the history and structure of Arabic are not covered in either the first or the second edition. Given the progress in several fields of Arabic linguistics (in particular in sociolinguistics, structural analysis of Standard Arabic, and dialect studies) and the proliferation of publications, a new comprehensive reference tool is needed. More than other disciplines, the study of the Arabic language is characterized by a fragmentation of the efforts of scholars, both between the Arab and the Western countries and between Europe and the United States. As a result, scholars often are not aware of the work done by colleagues elsewhere.

The *Encyclopedia of Arabic language and linguistics*, the first volume of which is presented here, fills the gaps. It is intended in the first place as a reference tool for linguists working with Arabic, but also for scholars from many other disciplines (Islamic studies, Arabic literature, social sciences), whose fields of research frequently intersect with that of linguistics, if only because Arabic as the language of the *Qurʾān* plays such a pervasive role in the entire Arab and Islamic world. Beyond this, the *EALL* will also be a reference tool for general linguists. In an article on the importance of Arabic for → general linguistics, Comrie (1991:29) points out that Arabic incorporates "a wealth of fascinating data relating to the variation among the many vernaculars". Yet, this material is underrepresented in general linguistics because of the lack of authoritative and accessible sources. By bringing together data on all varieties of Arabic, the *EALL* contributes to the dissemination of knowledge about one of the world's major languages.

Perhaps the most important category of users targeted by the *EALL* is that of students, especially those at the graduate and the postgraduate levels. For this reason, when presented with a choice between succinctness and a surplus of information, the editors have accepted the resulting overlaps. Another aspect of the didactic side of this policy decision is that the use of abbreviations has been avoided as much as possible; names of varieties of Arabic, for instance, are always written out in full, and even in grammatical contexts the use of abbreviations has been restricted, except in morph-by-morph translations, where standard coding has been used.

2. Scope and approach of the *Encyclopedia of Arabic language and linguistics*

The *EALL* is an encyclopedic handbook covering all relevant aspects of the study of Arabic and dealing with all levels of the language (pre-Classical Arabic, Classical Arabic, Modern Standard Arabic, Arabic vernaculars, mixed varieties of Arabic), both synchronically and diachronically. It will be published in four volumes with a total of two million words, distributed over approximately 500 entries. The treatment includes both the external and the internal history of the language, as well as the structural analysis of the different varieties of the language, the interaction between varieties in mixed levels (such as Middle Arabic), the linguistic contacts between Arabic and other languages, and the place of Arabic within larger language groups.

The terminology with which the varieties of Arabic and the various stages in its history are indicated is notoriously complicated (→ history of Arabic). In principle, the term 'Old Arabic' is used to cover the pre-Islamic period for which the sources are the earliest inscriptions in a language that is recognizably Arabic, the language of the *Qurʾān*, pre-Islamic poetry, and information culled from the dialects of the Bedouin tribes by the Arab grammarians. Earlier forms of a language closely resembling Arabic are called 'Proto-Arabic', while 'Ancient North Arabian' is used as the term for the languages of the North Arabian inscriptions that were related to Arabic. 'Classical Arabic' is used for the language as it was codified by the Arab grammarians. 'Modern Standard Arabic' is the modern form of this language. The term 'Arabic dialects' is used freely without any connotation of lesser status; it is the common phrase for the vernacular varieties of Arabic. Terms like 'Colloquial Arabic' and 'Educated Arabic' (and the many variants of these terms), on the other hand, indicate sociolinguistic levels of the language. 'Middle Arabic', finally, is used to indicate a category of texts written in mixed varieties, rather than a stage in the history of the language. As terminology is usually connected closely with the theoretical views of individual researchers, authors have been given considerable latitude in their use of terms. In some cases the editors have refrained from interfering, even if an author's terminology did not agree exactly with the terms preferred by the editors.

With regard to the relationship between dialectal forms and standard forms, any terminology presupposes a theoretical stance on the development of the language. In order to avoid as much as possible taking a position in the debate about this relationship, the editors have chosen to refer to the dialectal forms as reflexes of either the codified form of Arabic, i.e. Classical Arabic, or the historical predecessors of this codified form, i.e. Old Arabic. This is meant as a neutral way of referring to related forms, which does not suggest any genetic or evolutionary relationship between the two varieties.

The *EALL* was set up as a meeting place for a wide variety of theoretical approaches, and the editors have made no effort at all to harmonize these approaches. Rather than selecting one descriptive model, they believe that alternative analyses, whether traditional, functionalist, generativist, minimalist, or reflective of any other linguistic school, should all be represented. The indigenous Arabic tradition, too, is covered extensively, primarily in the entries with an Arabic title (e.g. → *ʾiʿrāb*, → *ism*, → *fāʿil*).

For each topic a synthesis of the most recent research is given, with the emphasis on adequate bibliographical coverage. This applies in particular to what the editors regard as the central articles (e.g. → syntax, → morphology, → diglossia, → multilingualism), which are intended as general introductions to the field. Other entries are more in the nature of essays (e.g. → language and culture, → language and ethnicity, → language and nationalism). Still others are more technical (e.g. → prosody, → X-bar syntax).

All entries are written from an encyclopedic point of view, which means that authors were requested to 'objectify' their views. This is not to say that they were discouraged from presenting novel ideas. In fact, for some entries authors were asked and permitted to give their own theories, even when these were not universally accepted (e.g. → ʿArab, → matrix and etymon theory, → poetic koine). As long as these views are clearly presented as such, the readers of the encyclopedia deserve a chance to become acquainted with innovatory, sometimes even controversial, views.

In such an approach overlap is unavoidable and perhaps even desirable: on basic topics of Arabic structure the reader will find side by side in the *EALL* traditional philological treatment and modern syntactic analysis of the same linguistic facts. As a result, the same data are sometimes repeated in several entries, but from different angles. Unfortunately, it turned out to be impossible to cover all topics originally envisaged, simply because authors could not be found for all topics. The extensive index that will be published in Volume IV will make it possible to find information about those items which do not have their own entry.

Cross-references in the entries have been used sparingly. A few terms without their own entry are cross-referred to a more general entry (e.g., 'plural' and 'singular' are cross-referred to 'number'). The index at the end of Volume IV will cover the entire encyclopedia and allow readers to find all relevant loci.

Some special features of the *EALL* deserve to be mentioned. In the first place, the indigenous linguistic tradition, not always represented in analyses of Arabic, has been included here. Another feature is the inclusion of dialect sketches of more than 40 dialects, described according to a predetermined format, which allows the user to make quick cross-dialectal comparisons. For all Arabic-speaking countries a linguistic profile has been included that sketches the sociolinguistic and dialectological make-up of these countries. The relations between Arabic and the other Semitic (and Afro-Asiatic) languages are dealt with in separate entries (→ South Semitic languages, → Northwest Semitic languages, → Semitic languages, → Afro-Asiatic languages). The relations between Arabic and other languages in the Islamic world (e.g. → Persian, → Indonesian/Malay, → Swahili, → Hausa) are dealt with in two ways: entries with the name of a language as their title deal with the Arabic influence in these languages, whereas entries with the term 'loanwords' in their title deal with the influence of these languages on Arabic (where 'loanwords' has been chosen as a blanket term, covering all levels of interference).

3. TRANSCRIPTION

Transcription is always a problem, especially in the case of an encyclopedia that brings together data from Standard Arabic and dialects. The transcription follows in the main the one adopted by Fischer and Jastrow in the *Handbuch der arabischen Dialekte* (1980:11–14), but with some adaptations. The editors have decided to use *ḏ̣* (not *ẓ*) throughout, except in proper names (thus *ḏ̣uhr*, but Ibn Manẓūr).

Table 1: Transcription of the Arabic alphabet

أ	*ʾ*
ب	*b*
ت	*t*
ث	*ṯ*
ج	*j*
ح	*ḥ*
خ	*x*
د	*d*
ذ	*ḏ*
ر	*r*
ز	*z*
س	*s*
ش	*š*
ص	*ṣ*
ض	*ḍ*
ط	*ṭ*
ظ	*ḏ̣*
ع	*ʿ*
غ	*ġ*
ف	*f*
ق	*q*
ك	*k*
ل	*l*
م	*m*
ن	*n*
ه	*h*
و	*w*
ي	*y*

In the transcription of Modern Standard Arabic and Classical Arabic, the following rules have been followed:

- *hamza* at the beginning of the word is always transcribed
- the article is transcribed in its assimilated form (*as-sikka*, etc.) and written with a hyphen; the *waṣla* is not transcribed, thus *wa-l-faras*
- suffixes *-hu, -hā*, etc. and clitics *bi-, li-, wa-* etc. are written attached to the word to which they belong, optionally with hyphen
- endings are written when relevant; otherwise, nouns are given in their pausal form without case endings and connecting vowels, thus *min al-bayt, hum al-muʾminūna*, but verbs retain the last short vowel, thus *kataba*, *yaktubu*; suffixes always retain their connected form, thus *kitābu-hu*, not *kitābuh*
- the feminine ending is transcribed *–a* (not *–ah*), except in genitive constructions, thus *al-madīna*, but *madīnat an-nabī*

For the Arabic dialects a standardized phonological transcription is used (without slashes and in italics), unless phonetic detail is at issue. In the standardized transcription articles are always assimilated and written without hyphen, likewise the pronominal suffixes; words like *wi, bi*, and *li* are written as separate words. When phonetic transcription is needed, this is given in square brackets using IPA signs (see www2.arts.gla.ac.uk/IPA). When it is necessary to indicate phonological transcription explicitly, this is given between slashes using the standard transcription signs rather than IPA signs. For morphological notation straight lines are sometimes used.

The following signs are used in standard phonological transcription:

obstruents: *p, b, ḅ, t, ṭ, d, ḍ, k, ḳ, g, g̣, q, ʾ*
affricates: t^s, *ć, ǵ, č, j*
stridents: *f, f̣, v, ṿ, ṯ, ṯ̣, ḏ, ḏ̣, χ, ġ, ḥ, ʿ, h*
sibilants: *s, ṣ, z, ẓ, ś, ź, š, ž*
laterals and vibrants: *l, ḷ, r, ṛ*
nasals: *m, ṃ, n, ṇ, ŋ, ɲ, ñ*
semivowels/glides: *w, ẉ, y*
vowels: *i, ī, e, ē, a, ā, ä, ǟ, å, å̄, u, ū, o, ō, ö*
diphthongs: *ay, āy, aw, āw*

To indicate short vowels the vowel sign + breve is used (*ă, ĕ*, etc.); superscript vowel signs are used to indicate ultrashort/epenthetic vowels; subscript dot (*ạ, ẹ*, etc.) is used to indicate open vowels outside phonetic notation, rather than the IPA signs; to indicate nasalized vowels tilde is used (*ã, ẽ*, etc.).

For primary stress the vowel sign + acute is used (*á, é*, etc.); for secondary stress the vowel sign + gravis is used (*à*, *è*, etc.); alternatively, and always so in phonetic transcription, stress is indicated by an apostrophe ' before the tonic syllable. To indicate palatalization superscript *y* is used (t^y, d^y, etc.); to indicate labialization superscript *w* is used (m^w, etc.).

For the transcription of Persian and Ottoman Turkish a phonological transcription has been the preferred option (thus, for instance, in Persian *vaġt* is written rather than *waqt*); for Modern Turkish the standard orthography has been used. For other languages standard orthography is used when this exists, for instance for Indonesian, Hausa (including ejectives such as *ɓ*, *ɠ*, *ɗ*, etc.), and Swahili. Otherwise, scholarly practice in the field is followed, for instance in the transcription of glottalized consonants in Ethiopian languages as *k*ʿ, *t*ʿ, etc.

4. ACKNOWLEDGMENTS

The idea for an encyclopedia of Arabic was first suggested to Brill by Andrzej Zaborski in the early 1990s, and was later developed by a group of linguists during a meeting of the Association Internationale de Dialectologie Arabe in Cambridge in 1995. After the initial stages, Brill enthusiastically accepted the idea, under the responsibility first of Albert Hoffstädt, then of Olaf Köndgen. In the final stages of Volume I, the project was taken over by Joed Elich. The editorial committee convened for the first time in 2000 and set about to establish a list of entries. A list of more than 2,000 terms was compiled covering all subdomains (morphology, phonology, semantics, historical linguistics, syntax, language contacts, indigenous grammar, sociolinguistics, psycholinguistics, and dialectology). Subsequently, 500 terms were selected as individual entries to represent the major topics. About 35 entries were chosen for general survey articles.

Once the project was underway, an Advisory Committee was appointed to assist the editors in their work, consisting of Ramzi Baalbaki (Beirut), Elsaid Badawi (Cairo), Dominique Caubet (Paris), Clive Holes (Oxford), Manfred Kropp (Beirut), Jérôme Lentin (Paris), John McCarthy (Amherst), Jamal Ouhalla (Dublin), Jan Retsö (Göteborg), and Sabah Safi (Jedda). The editors wish to thank the members of the Advisory Committee for their advice and support, especially in the formative stages of the project, and for agreeing to write some of the articles. The first invitations to authors were sent out in 2003.

The entire project would have been impossible without the support of Brill. Special mention must be made of Ingrid Heijckers, without whom the project would literally have been impossible. Thanks to her meticulous management of the *EALL* database and her unfailing common sense and good humor, this first volume could be realized according to schedule. The tireless copy editors Margaret Owen (Volume I) and Carolyn Russ (later volumes) did a wonderful job of harmonizing the text and improving the style.

Finally, of course, the authors of the entries are to be thanked for their enthusiastic cooperation, especially at a time when many research assessment organizations refuse to take into account articles written for encyclopedias and to recognize the value such articles have for the scholarly community. The editors hope to bring out the remaining three volumes in the years 2006 and 2007.

October 2005

Mushira Eid
Alaa Elgibali
Kees Versteegh
Manfred Woidich
Andrzej Zaborski

BIBLIOGRAPHICAL REFERENCES

Anghelescu, Nadia. 1995. *Langage et culture dans la civilisation arabe*. Paris: L'Harmattan.

Badawi, Elsaid, Michael G. Carter, and Adrian Gully. 2004. *Modern written Arabic: A comprehensive grammar*. London and New York: Routledge.

Bakalla, Muhammad. 1983. *Arabic linguistics: An introduction and bibliography*. London: Mansell.

Bateson, Mary Catherine. 1967. *Arabic language handbook*. Washington, D.C.: Center for Applied Linguistics. (Repr., Washington, D.C.: Georgetown University Press, 2003.)

Behnstedt, Peter. 1985–1987. *Die nordjemenitischen Dialekte*. 2 vols. Wiesbaden: L. Reichert.

——. 1997. *Sprachatlas von Syrien. Kartenband + Beiheft*. Wiesbaden: O. Harrassowitz.

Behnstedt, Peter and Manfred Woidich. 1985–1999. *Die ägyptisch-arabischen Dialekte*. 5 vols. Wiesbaden: L. Reichert.

——. 2005. *Einleitung in die arabische Dialektgeographie*. Leiden: E.J. Brill.

Blachère, Régis and Maurice Gaudefroy-Demombynes. 1952. *Grammaire de l'arabe classique (Morphologie et syntaxe)*. 3rd ed. Paris: G.-P. Maisonneuve.

Cantarino, Vicente. 1974–1975. *Syntax of modern Arabic prose*. 3 vols. Bloomington: Indiana University Press.

Comrie, Bernard. 1991. "On the importance of Arabic for general linguistic theory". *Perspectives on Arabic linguistics,* III, ed. Bernard Comrie and Mushira Eid, 3–30. Amsterdam and Philadelphia: J. Benjamins.

El-Ayoubi, Hashem, Wolfdietrich Fischer, and Michael Langer. 2001, 2003. *Syntax der arabischen Schriftsprache der Gegenwart*. I/1. *Das Nomen und sein Umfeld*. I/2. *Die konnektiven Wortarten des Nomens: Pronomina, Adverbien, Präpositionen*. Wiesbaden: L. Reichert.

Ferrando, Ignacio. 2001. *Introducción a la historia de la lengua árabe: Nuevas perspectivas*. Saragossa: Navarro y Navarro.

Fischer, Wolfdietrich. 1982. *Grundriß der arabischen Philologie*. I. *Sprachwissenschaft*. Wiesbaden: L. Reichert.

——. 2002. *A grammar of Classical Arabic*. Transl. Jonathan Rodgers. 3rd ed. New Haven, Conn.: Yale University Press.

Fischer, Wolfdietrich and Otto Jastrow. 1980. *Handbuch der arabischen Dialekte*. Wiesbaden: O. Harrassowitz.

Fleisch, Henri. 1961, 1979. *Traité de philologie arabe*. 2 vols. Beirut: Imprimerie Catholique.

Holes, Clive. 2004. *Modern Arabic: Structures, functions and varieties*. 2nd rev. ed. Washington, D.C.: Georgetown University Press. (1st ed., London and New York: Longman.)

Howell, Mortimer Sloper. 1883–1911. *A grammar of the Classical Arabic language*.

Penrice, John. 1873. *A dictionary and glossary of the Korān with copious grammatical references and explanations of the text*. (Repr., New York: Praeger, 1971.)

Procházka, Stephan and Arne Ambros. 2004. *A concise dictionary of Koranic Arabic*. Wiesbaden: L. Reichert.

Reckendorff, Hermann. 1895–1898. *Die syntaktischen Verhältnisse des Arabischen*. 2 vols. Leiden: E.J. Brill. (Repr., 1967.)

Sezgin, Fuat. 1967–2000. *Geschichte des arabischen Schrifttums*. 12 vols. Leiden: E.J. Brill.

Versteegh, Kees. 2001. *The Arabic language*. 2nd ed. Edinburgh: Edinburgh University Press.

Wright, William. 1859–1862. *A grammar of the Arabic language*. London. (3rd ed., rev. W. Robertson Smith and M.J. de Goeje. 2 vols. Cambridge: Cambridge University Press, 1896–1898.)

List of Contributors

Soha Abboud-Haggar La Complutense University, Madrid, Spain
Hussein Abdul-Raof University of Leeds, U.K.
Farida Abu-Haidar London, U.K.
Jordi Aguadé University of Cádiz, Spain
Abdul Sahib Mehdi Ali University of Sharjah, United Arab Emirates
Khawla Aljenaie Kuwait University, Kuwait
Muhammed Al-Sharkawi American University in Cairo, Egypt
Mohamadou Aminou University of Yaounde I, Cameroon
Joseph Aoun University of Southern California, Los Angeles, U.S.A.
Werner Arnold University of Heidelberg, Germany
Ramzi Baalbaki American University of Beirut, Lebanon
Maher Bahloul American University of Sharjah, United Arab Emirates
Peter Behnstedt Chipiona, Spain
Sabrina Bendjaballah Centre National de la Recherche Scientifique, France
Elabbas Benmamoun University of Illinois at Urbanan-Champaign, U.S.A.
Lidia Bettini University of Florence, Italy
Alexander Borg Ben-Gurion University of the Negev, Israel
Aziza Boucherit Université René Descartes – Paris 5, France
Sami Boudelaa MRC-Cognition and Brain Sciences Unit, U.K.
Naima Boussofara-Omar University of Kansas, U.S.A.
Dinie Bouwman Groningen, The Netherlands
Ahmed Chanfi Centre for Modern Oriental Studies in Berlin, Germany
Leila Chérif-Chebbi Paris, France
Noureddine Chenfour Dhar El-Mehraz, Fes, Morocco
Lina Choueiri American University of Beirut, Lebanon
Federico Corriente University of Zaragoza, Spain
Mihai Dat University of Bourgogne, France
Humphrey Davies American University in Cairo, Egypt
Everhard Ditters University of Nijmegen, The Netherlands
George Echu University of Yaounde I, Cameroon
Lutz Edzard University of Oslo, Norway
John C. Eisele College of William and Mary, U.S.A.
Kerstin Eksell University of Copenhagen, Denmark
Abderrahman El Aissati Tilburg University, The Netherlands
Dalal Mahmoud El-Gemei al-Azhar University Women's Branch, Egypt
Mohssen Esseesy George Washington University, U.S.A.
Said Faiq American University of Sharjah, United Arab Emirates
Ahmed Fakhri West Virginia University, U.S.A.
Mohammed Farghal Kuwait University, Kuwait
Ignacio Ferrando University of Cádiz, Spain
Wolfdietrich Fischer Erlangen, Germany
Adam Gacek McGill University, Canada
Alain Girod Bordeaux, France
JAcques Grand'Henry Catholic University of Louvain, Belgium
Beatrice Gruendler Yale University, U.S.A.
Martine Haak Diemen, The Netherlands
Niloofar Haeri Johns Hopkins University, U.S.A.

ALASTAIR HAMILTON School of Advanced Studies, London University attached to the Warburg Institute, U.K.
CLIVE HOLES University of Oxford, U.K.
JAN HOOGLAND University of Nijmegen, The Netherlands
BRUCE INGHAM School of Oriental and African Studies, London, U.K.
TAMÁS IVÁNYI Budapest, Hungary
OTTO JASTROW University of Erlangen-Nuremberg, Germany
RUDOLF DE JONG University of Amsterdam, The Netherlands
PATRICE JULLIEN DE POMMEROL Ankara, Turkey
HELÉNE KAMMENSJÖ Göteborg, Sweden
ALAN S. KAYE California State University, U.S.A.
ESAM KHALIL University of Nijmegen, The Netherlands
GHADA KHATTAB University of Newcastle, U.K.
ALAIN KIHM CNRS – UMR 7110 Paris, France
MICHAEL LANGER Schlettau/Erzgebirge, Germany
PIERRE LARCHER University of the Provence, France
JÉRÔME LENTIN Institut National des Langues et Civilisations Orientales, Paris, France
MARK S. LETOURNEAU Weber State University, U.S.A.
ABDUL AZIZ LODHI Uppsala, Sweden
UTZ MAAS University of Osnabrück, Germany
ADRIAN MĂCELARU University of Bucharest, Romania
JACOB MANSOUR University of Haifa, Israel
GUNVOR MEJDELL University of Oslo, Norway
CATHERINE MILLER Aix-en-Provence, France
SAMIA NAÏM CNRS-LACITO, France
JONATHAN OWENS University of Bayreuth, Germany & College Park Maryland, U.S.A.
HEIKKI PALVA University of Helsinki, Finland
YISHAI PELED Tel Aviv University, Israel
STEPHAN PROCHÁZKA University of Vienna, Austria
ROBERT RATCLIFFE Tokyo University of Foreign Studies, Japan
JOHANNES REESE University of Zürich, Switzerland
JAN RETSÖ Göteborg University, Sweden
ALI RIAZ Illinois State University, U.S.A.
TONIO SEBASTIAN RICHTER University of Leipzig, Germany
ANDRÉ ROMAN Université Lumière – Lyon 2 – CRTT, France
JUDITH ROSENHOUSE Technion I.I.T. Haifa, Israel
VALERIY RYBALKIN Institute of Oriental Studies, National Academy of Sciences of Ukraine, Kiev, Ukraine
KAREN C. RYDING Georgetown University, U.S.A.
MARIE-CLAUDE SIMEONE-SENELLE LLACAN-CNRS, INALCO, University of Paris 7, France
YASIR SULEIMAN University of Edinburgh, U.K.
DAVID TESTEN Reston, VI, U.S.A.
KEES VERSTEEGH University of Nijmegen, The Netherlands
ÁNGELES VICENTE University of Zaragoza, Spain
RAINER VOIGT Free University of Berlin, Germany
MARY-ANN WALTER Massachusetts Institute of Technology, U.S.A.
KEITH WALTERS University of Texas at Austin, U.S.A.
MICHAEL WALTISBERG Philipps University Marburg, Germany
JANET C.E. WATSON University of Salford, U.K.
EDWIN WIERINGA University of Cologne, Germany
JAMES M. WILCE Northern Arizona University, U.S.A.
MANFRED WOIDICH University of Amsterdam, The Netherlands

ANDRZEJ ZABORSKI University of Cracow, Poland
PETR ZEMÁNEK Charles University, Czech Republic
CHAKIR ZEROUAL University Sidi Mohamed Ben-Abdellah, Morocco & CNRS – UMR 7018, Paris
TAMAR ZEWI University of Haifa, Israel
OTTO ZWARTJES University of Amsterdam, The Netherlands

A

Abbreviations

Just as in the Western tradition, so also in the Arabic context, frequent repetition of the same word or phrase in the text leads the writer or scribe to use abbreviations, acronyms, and initialisms. The difference between acronyms and initialisms is that the former are pronounced as words while the latter are spelled out letter by letter. As far as is known, initialisms were not used in the manuscript age but became a common feature in modern Arabic. Abbreviations are usually designated in various sources as: *ʿalāmāt, rumūz, muṣṭalaḥāt* (*iṣṭilāḥāt*), and *muxtaṣarāt*. Although two important lists of abbreviations were recently published (al-Māmaqānī 1992; aẓ-Ẓufayrī 2002), there is to date no complete study of their usage in Arabic, whether in the past or in modern times.

Generally speaking, there are four main categories of abbreviations encountered in Arabic texts:

i. Suspensions: abbreviation by truncation of the letters at the end of the word, e.g. المصـ = *al-muṣannif*. Perhaps the most interesting here is the case of suspensions that look like, or were considered by some to be, numerals. To this category belong the signs that resemble the numerals ٢ and ٣, but which may represent the unpointed *tāʾ* and *šīn* (for *tamām* and *šarḥ*) when used in conjunction with marginal glosses.
ii. Contractions: abbreviating by means of omitting some letters in the middle of the word, but not the beginning or the ending, e.g. قه (*qawlu-hu*).
iii. Sigla: using one letter to represent the whole word, e.g. م (*matn*).
iv. Abbreviation symbols: symbols in the form of logographs used for whole words. A typical abbreviation symbol is the horizontal stroke (sometimes hooked at the end) which represents the word *sana* 'year'. Another example is the 'two teeth stroke' (which looks like two unpointed *bā*'s), which represents the word قف 'stop', or the suspension فته (for *fa-taʾammal-hu/hā* 'reflect on it'), used in manuscripts for notabilia or side-heads.

Closely connected with these abbreviations is a contraction of a group of words into one 'portmanteau' word (*naḥt*; → compounds), for instance *basmala* (*bi-sm Allāh*) *ḥamdala* (*al-ḥamdu li-llāh*) and *ṣalwala* (*ṣallā llāh ʿalayhi*). To all intents and purposes, the word *naḥt* corresponds to an acronym, i.e. a word formed from the abbreviation of, in most cases, the initial letters of each word in the construct. Most of these constructs are textual and pious formulae. Apart from *basmala*, *ḥamdala*, and *ṣalwala*, we encounter *ṭalbaqa* (*ṭāla llāh baqāʾa-hu*), *ḥawqala* or *ḥawlaqa* (*lā ḥawla wa-lā quwwata ʾillā bi-llāh*), *ṣalʿama* (a synonym of *ṣalwala*), *ḥasbala* (*ḥasbunā allāh*), *mašʾala* (*mā šāʾa llāh*), *sabḥala* (*subḥāna llāh*), and *ḥayʿala* (*ḥayya ʿalā ṣ-ṣalāt*) (as-Samarrāʾī 1987; Gacek 2001).

Abbreviations, especially contractions and sigla, may be (and often are) accompanied by a horizontal stroke (*tilde*) placed above them. This mark may resemble the *madda* but has nothing to do with the latter's function in Arabic script. Suspensions, on the other hand, were indicated by a long downward stroke, a mark that is very likely to have been borrowed from Greek and Latin paleographic practice.

The use of abbreviations was quite popular among Muslim scholars, although originally some of the abbreviations, such as those relating to the prayer for the Prophet (*taṣliya, ṣalwala*), were disapproved of. In the manuscript age, abbreviations were extensively used, not only in the body of the text but also in marginalia, ownership statements, and in the primitive critical apparatus (Ben Cheneb 1920; Maḥfūẓ 1964).

Medieval scholars could not always agree on the meaning of some of the abbreviations used in manuscripts. The letter ح, for instance, which is used to separate one *ʾisnād* from another, was thought by some to stand for *ḥāʾil* or *ḥaylūla* ‘separation’ and by others for *ḥadīṯ* and even *ṣaḥḥa*. Some scholars even thought that the letter *ḥāʾ* should be pointed (خ – *xāʾ muʿjama*) to stand for *ʾisnād ʾāxar* ‘another *ʾisnād*’. The contemporary scholar may face a similar dilemma (see e.g. Alič 1976).

Abbreviations in manuscripts are often unpointed and appear sometimes in the form of word-symbols (logographs). Here, the context, whether textual or geographical, is of great importance. Thus, for instance, what appears to be the letter ط may in fact be a ظ, and what appears to be an *ʿayn* or *ġayn*, in its initial (عـ) or isolated form (ع), may actually be an unpointed *nūn* or *xāʾ* (for *nusxa ʾuxrā* ‘another copy’). Similarly, the same word or abbreviation can have two different functions and/or meanings. For example, the words *ḥāšiya* and *fāʾida* can stand for a gloss or a side-head (‘nota bene’), while the ص or صـ can be an abbreviation of *ṣaḥḥa* (when used for an omission/ insertion or evident correction) or *ʾaṣl* (the body of the text), or it can stand for *ḍabba* ‘door-bolt’, a mark indicating an uncertain reading and having, for all intents and purposes, the function of a question mark or ‘sic’. Also, the abbreviation ن may stand for *bayān* ‘explanation’ or *nusxa ʾuxrā*; the latter is often found in manuscripts of Persian/ Indian provenance.

The earliest use of abbreviations in the Arabic language is probably connected with its orthography and possibly the ‘mysterious letters’ (*al-ḥurūf al-muqaṭṭaʿa*) at the beginning of some chapters of the *Qurʾān* (Bellamy 1973). In terms of orthography, for instance, the initial form of *jīm* (جـ) or *mīm* (مـ) was regarded by some scholars as an abbreviation of *jazma*. Furthermore, the unpointed initial form of *šīn* (سـ) was used for *tašdīd* (or *šadda*), and the initial form of *ṣād* (صـ) was thought to represent *waṣla* (or *ṣila*) (Wright 1967:13–14, 19; Gacek 2001:23).

Most of the abbreviations are found in the body of the text. They were introduced in order to speed up the process of transcription and their usage varied according to the subject or type of a given work. Abbreviations can be found in almost all types of works, but especially in compositions on the recitation of the *Qurʾān*, compilation and criticism of *Ḥadīṯ*, philosophy, lexicography, poetry, genealogy, biography, and astronomy. The lists of these are often included in prefaces and frequently concern either the names of authors or titles of compositions. In addition, we find didactic poems that were composed specifically in order to help memorize given sets of abbreviations (see, e.g., ʿAlawān 1972). They are especially common in works on *Ḥadīṯ* and jurisprudence (both Sunni and Shiʿi) (al-Māmaqānī 1992; aẓ-Zufayrī 2002), and although some abbreviations were standardized, most were specific to a given work. Among the commonly used abbreviations for major *Ḥadīṭ* compilations are: خ (al-Buxārī), م (Muslim or Mālik), د (ʾAbū Dāʾūd), ت (at-Tirmiḏī), ك (Mālik), ه (ʾAbū Ḏarr or Ibn Māja), ن or س (an-Nasāʾī), and the like (Gacek 1989:56).

Specific to *Ḥadīṯ* literature are other abbreviations connected with the frequent repetitions of such expressions as *ḥaddaṯanā, ʾaxbaranā*, and *ʾanbaʾanā*, which were commonly abbreviated as: نا, ثنا, دثنا (*ḥaddaṯanā*); انا, ارنا, ابنا (*ʾaxbaranā*); قثنا, ق ثنا (*qāla ḥaddaṯanā*). The transition from one *ʾisnād* to another, as mentioned above, was marked with ح (*ḥāʾil*, *taḥwīl*, *ḥaylula*, *ḥadīṯ* or *ṣaḥḥa*) (Gacek 1989:56), and for the evaluation of *ḥadīṯ*s the following abbreviations were used: ض (*ḍaʿīf*), صح (*ṣaḥīḥ*), ح (*ḥasan*); م (*majhūl*), مو (*muwāfiq* or *mawqūf*), قف (*mawqūf*); ق (*muwaṭṭaq* or *muttafaq ʿalayhi*), ل (*mursal*) (e.g. Gacek 1985:xiv, 96).

With the gradual development of various Qurʾānic disciplines, and a growing body of literature, various *Qurʾān*-specific abbreviations were introduced. They relate to either the evaluation of pauses (*waqf*) or variant readings (*qirāʾāt*). These abbreviations are found written in the text in red ink above the relevant word. For pauses in the text the following are encountered: م (*lāzim*), ط (*muṭlaq*), ج (*jāʾiz*), ص or ض (*muraxxaṣ ḍarūratan*, *ḍarūrī*), ز (*mujawwaz li-*

wajh), ق (*qīla ʿalayhi l-waqf* or *qad qīla*), قف (*yūqaf ʿalayhi*), صل (*qad yūṣal*), صلى (*al-waṣl ʾawlā*), قلى (*al-waqf ʾawlā*), س (*sakt*), مع (*muʿānaqa*) (Ahmad 1984:104–108). As for variant readings, the most common are the abbreviations introduced in the *Šāṭibiyya* of Ibn Firruh aš-Šāṭibī (d.590/1194), as well as the *Ṭayyiba* and *Ġāyat an-nihāya fī ṭabaqāt al-qurrāʾ* by Ibn al-Jazarī (d.833/1429). They use either the *ʾabjad* sequence (e.g. ا = Nāfiʿ, ب = Qālūn, ج = Warš, and د = Ibn Kaṯīr), or the sigla which employ one letter from the author's name (e.g. ج = ʾAbū Jaʿfar Yazīd al-Maxzūmī, ى = Yaʿqūb al-Ḥaḍramī, and خ = Xalaf al-Bazzār).

Another type of literature which uses specific abbreviations includes language and subject dictionaries. Here we may mention the *Muqaddimat al-ʾadab* by Maḥmūd az-Zamaxšarī (d. 538/1144), *ʿUmdat aṭ-ṭabīb fī maʿrifat an-nabāt* by ʾAbū l-Xayr al-ʾIšbīlī (6/12th century), *al-Qāmūs al-muḥīṭ* by Muḥammad ibn Yaʿqūb al-Fīrūzābādī (d. 817/1415), and *ar-Rāmuz fī l-luġa* by Muḥammad ibn Ḥasan ibn ʿAlī (d. 860/1455). Al-Fīrūzābādī, for instance, gives the following list: ع *(mawḍiʿ*), د (*balad*), ة (*qarya*), ج (*jamʿ*), and م (*maʿrūf*).

A similar picture is seen in works on jurisprudence, grammar, and theology. Here one encounters numerous abbreviations of such frequently repeated words as: *ʾaḥaduhumā* (اح), *ḏāhir* (ظ), *bāṭil* (بط), *kaḏālika* (كك), *fa-kaḏālika* (فكك), *ḥīnaʾiḏin* (ح), *hāhunā* (هن), *mawḍūʿ* (ضع), *mamnūʿ* (مم), *hāḏā xalaf* (هف), etc. This type of abbreviation became very common from about the 10th/16th century onward and was used extensively in manuscripts of Persian provenance (Gacek 1985, xiii–xiv; Heer 1969).

Apart from the subject-specific works, another abbreviation-rich category of literature, commentary (*šarḥ*) and gloss (*ḥāšiya*), should also be mentioned here. In the manuscript age, there were several ways of distinguishing between the original text (*matn*), i.e. the text commented upon, and the commentary or gloss. In comment-text books (*šarḥ mamzūj*), the *matn* was either written in red and the *šarḥ* in black, or the following abbreviations were used: م (*matn*), ص (*ʾaṣl*), and ش (*šarḥ*), ق (*qāla*), اق (*ʾaqūlu*), قه (*qawlu-hu*), المص (*al-muṣannif*), الش (*aš-šāriḥ)*, and المح (*al-muḥaššī*).

The passage commented upon or glossed was quoted either in full or in a shortened form. Thus, the end of a quotation was indicated by اه , ه (various forms) and هى (all abbreviated forms of the verb *intahā*) or الخ (= *ʾilā ʾāxirih*), whereas the gloss could be introduced by the word *ʾay* 'that is'.

Another large category of abbreviations concerns eulogies and prayers. Although their usage was looked upon with disapproval, particularly when applied to Allāh (*ism al-jalāla*) and the Prophet Muḥammad, with time they became prevalent in most manuscripts and printed books. In this category one may encounter the following: عج / عج (*ʿazza wa-jalla)*; تع / تعـ (*taʿālā*); جش (*jalla šaʾnu-hu*); ص / صم / صله / صلم / صلعم / صلع / صع (*ṣallā llāh ʿalayhi wa-sallam*); عه / عسـ / عم / عم / ع / ع (*ʿalayhi as-salām*); قس / قس / قده / قد / ق / س (*qaddasa llāh sirra-hu* or *rūḥa-hu* or *quddisa sirru-hu*); ض / رضه / رض / ر (*raḍiya llāh ʿan-hu*); ره / رم / رحه (*raḥima-hu llāh*); طه / طر (*ṭāba llāh ṯarā-hu*); and طلع / طع / طع / طره (*ʾaṭāla llāh ʿumra-hu*) (Gacek 2001:174–175).

Specific to manuscripts are also abbreviations of months, closing formulae in the colophon, quire signatures, notabilia (side-heads), and the primitive critical apparatus.

The abbreviation of the months of the year seems to have been a common Ottoman practice. Here we encounter: م = Muḥarram, ص = Ṣafar, را or ع ا or ع ل = Rabīʿ al-ʾawwal, ر or ر٢ or ع٢ = Rabīʿ aṯ-ṯānī (ʾāxir), جـ or ج ا or جا = Jumādā l-ʾūlā, ج or ج٢ = Jumādā al-ʾāxira, ب or ر = Rajab, ش or شع = Šaʿbān, ن or مض = Ramaḍān, ل or ش = Šawwāl, ذا or قع = Ḏū l-Qaʿda, ذ = Ḏū l-Ḥijja (Gacek 2001:174–175).

The end of the colophon in manuscripts is usually indicated by any of the following formulae: *tatmīm*, *taʾmīn*, *tafqīṭ*, *intihāʾ*, which most often appear in their abbreviated forms: م , ام , فقط or ط, and ه or هى . The *tafqīṭ* is almost exclusively used in the Persian/Indian context and the هى is often seen in Maghrebi manuscripts.

For quire signatures (or the numbering of quires) normally the words *juzʾ* or *kurrās* or *kurrāsa* were used in full or in their abbreviated forms namely: ج or جـ , and ك or ک . This practice is very similar to the one encountered in Latin manuscripts where the letter *Q* (for *quaternio*) was used.

The notabilia or side-heads were most commonly introduced by the word *qif* or *qif hunā, qif wa-taʾammal* (very often employed in the Maghreb) from which developed a logograph used as an overlining (*tawqīf*) (see earlier). Among other expressions used for this purpose were *unḏur* (abbrev. ظ – usually in the Maghreb), and *fāʾida*. The word *fāʾida* was mostly used in the Iranian and Indian context and may be seen in its abbreviated forms: ف (with a long horizontal stroke), or فيه, فصل or even ص (= *fāʾidat al-ʾaṣl*).

The primitive critical apparatus is a result of collation with the exemplar and other extant manuscripts. The collation (*muqābala, muʿāraḍa*), sometimes marked by the letter ع (*ʿūriḍa*), attracted a great variety of abbreviations and symbols, which differed from region to region and from period to period.

The ـصـ, originally a mark for *ḍabba* (also referred to as *ʿalāmat at-taḍbīb* or *at-tamrīḍ* or *taškīk*) was used for uncertain (doubtful) readings and resembled the initial form of the letter *ṣād*. Later, the ـصـ came to be regarded as an abbreviation of *ṣaḥḥa* and scholars were instructed to add to it (i.e. the *ṣād*) the letter ح (*ḥāʾ*), if the reading was confirmed, or the correct version was to be inscribed in the margin. Other scholars used ـضـ (*ḍād*) as an abbreviation of *ḍabbabtu-hu*. The ـصـ was also used in its suspended form for omissions as an abbreviation of *ṣaḥḥa*.

The letters ب and ن (above the restored word in the margin) (*bayān* or *bayānu-hu*) were employed for cacographic errors. However, in Shiʿi Imami manuscripts these errors were often indicated by the word *badal* or its abbreviation ل.

For metathesis or word transpositions the following abbreviations (placed above the relevant words) were used: خ = *muʾaxxar*; ق = *muqaddam* or *qabla*; ق خ = *muʾaxxar muqaddam*; ق م = *muʾaxxar muqaddam*; م = *muqaddam*; مم = *muqaddam muʾaxxar*; خم = *muʾaxxar muqaddam*; ب= *baʿda*. On the other hand, lacunae in the text were indicated by ض = *bayāḍ* or *hunā bayāḍ* and the ک often stood for *kaḏā* (or *hākaḏā*), i.e. sic/thus.

In manuscripts of Persian/Indian provenance a word to be replaced (substituted) by another word was marked by ل or د (*badal* or *baddil-hu*). It often appears with the خ (for *nusxa ʾuxrā*), i.e. خ ل (sometimes ل خ) or ن ل . The combination خ٢ may stand either for a simple variant (٢ = *ʾuxrā*) or a variant, which is deemed to be more correct. In this case, the ٢ can be read as an unpointed and suspended *bāʾ* (= *badal*). Similarly, ل٢ might represent the word *badal* itself (with an unpointed *bāʾ*) or the reference mark (*signe de renvoi*) ٢ (*bāʾ hindiyya*) and ل .

Furthermore, conjectures were marked by ـعـ (*raʾs al-ʿayn*) or somtimes as عه (*laʿalla-hu*), and ظ (*ʾaḏunnu-hu* or *ḏāhir*). The latter meaning is almost always encountered in the Persian/Indian context.

For textual variants it was common to employ sigla corresponding to the names of various transmitters (*rāwī*). Thus, for instance, in connection with the *Ṣaḥīḥ* of al-Buxārī (d.256/870) we may find the following: سـ = as-Saraxsī, حـ = al-Ḥamawī, هـ = al-Kušmihānī, سـ and حـ (sometimes suprascript) for both as-Saraxsī and al-Kušmihānī. Other abbreviations are: هـ = ʾAbū Ḏarr al-Harawī, ص = al-ʾAṣīlī, ش = Ibn ʿAsākir ad-Dimašqī, ظ = ʾAbū l-Waqt, ست – al-Mustamlī, حس = al-Ḥamawī and al-Mustamlī, حه = al-Ḥamawī and al-Kušmihānī (Quiring-Zoche 1998).

Non-specific variants (usually from another copy) were marked by a variety of abbreviations (often unpointed), e.g. خـ , خ , خه , نخـ , نخ , نخه, and ن (mainly India). Both نخـ and نخ often look like the letter ـعـ / ع or ـغـ / غ while the initial form of *xāʾ*, when unpointed, may look like *dāl* (د). They all represent the word *nusxa* (*ʾuxrā*), or the expressions *fī nusxa*, *fī nusxa ʾuxrā*, *fī ʾuxrā*.

For glosses and scholia the most often encountered words were *ḥāšiya* (in the Eastern Islamic lands, Mashriq) and *ṭurra* (Maghreb). The *ḥāšiya* was often abbreviated by scribes as: حـ, حه, حشـ, حيه, ح, and the *ṭurra* as ط. Among other words we find: *taʿlīq(a)* (abbrev. تـ or ـعـ), *tafsīr*, *šarḥ* (abbreviated سـ often unpointed and suspended or logographed), the above-mentioned *fāʾida* (abbreviated فـ, فيه, فصـ, فصل – the last two for *fāʾidat al-ʾaṣl*), and *hāmiš* (abbrev. هـ or ه).

For other types of glosses either a short title (including the word *šarḥ*, often in the form of a logograph) or short name (or sigla for these) were used. Other common references were given to well-known dictionaries, for example, ق for *al-Qāmūs al-muḥīṭ* of al-Fīrūzābādī, ص for *aṣ-Siḥāḥ* of al-Jawharī.

The end of the gloss was usually indicated by words or abbreviations of words which carry the meaning of an end or finish. Here we encounter the following: ه (in the form of a circle) or ٥ (in the form of an inverted heart), هـ (*hāʾ mašqūqa*), هی (*hāʾ* and *yāʾ*) or هـ ا, all representing the word *intahā* ‘it is finished’; number-like marks ٢ or ١٢ or ١٣ (١٤) (= *tamām*, *tamām šud*, Persian/Indian context only, see above); فقط (Iran and India only), often as a logograph (Gacek 1984:88);

and يه (*nihāya*, seen mostly in manuscripts of Indian provenance).

With the arrival of printing and the development of modern Arabic many new abbreviations came to be created. Prominent among them are acronyms and initialisms (Oman 1961 and Bosworth 1981). The component individual letters of initialisms are often but not always separated by full stops. In the field of bibliography, for instance, one encounters the following: ص (*ṣafḥa*) = page, ط (*ṭabʿa*) = printing or edition, دت (*dūna taʾrīx*) = n.d., دن (*dūna nāšir*) = s.n., دم (*dūna makān*) = s.l., سم (*santīmatr*) = cm., ج (*juzʾ*) = section, volume, مج (*mujallad*) = volume, بم (*baʿda l-mīlād*) = A.D., and ت (*tuwuffiya*, *al-mutawaffā*) = died. Among other common initialisms and acronyms are: ببس (BBC), جعم (al-Jumhūriyya al-ʿArabiyya al-Muttaḥida), شمم (*šarika ḏāt masʾūliyya maḥdūda*) = Co. Ltd., صب (*ṣundūq barīd*) = P.O.B., al-Yūniskū (= UNESCO), al-Yūnisīf (UNICEF), Tadmak or Radmak (at-Tarqīm ad-Duwalī al-Miʿyārī li-l-Kutub) = ISBN, Tadmad (at-Tarqīm ad-Duwalī al-Miʿyārī li-d-Dawriyyāt) = ISSN (Repp 2001:31–33).

Bibliographical references

Ahmad, Syed Barakat. 1984. *Introduction to Qur'anic script*. London: Curzon Press.

ʿAlawān, Muḥammad Bāqir. 1972. "Ṯalāṯ ʿarājīz fī rumūz 'al-Jāmiʿ aṣ-Ṣagīr'". *Majallat Maʿhad al-Maxṭūṭāt al-ʿArabiyya* 18.151–158.

Alič, Salih H. 1976. "Problem kratica u arapskim rukopisima (sa spiskom arapskih kratica iz 16. vjeka)" [The problem of abbreviations in Arabic manuscripts (with a list of Arabic abbreviations from the 16th century onward)]. *Prilozi za Orientalnu Filologiju* 26.199–212.

Bellamy, James A. 1973. "The mysterious letters of the Koran: Old abbreviations of the Basmalah". *Journal of the American Oriental Society* 93.267–285.

Ben Cheneb, Mohammed. 1920. "Liste des abréviations employées par les auteurs arabes". *Revue Africaine* 302–303, 134–138.

Bosworth, C. Edmund a.o. (eds.). 1981. "Abbreviations". *Encyclopaedia of Islam*, Supplement, fasc. 1–2. 2nd ed. Leiden: E.J. Brill.

Gacek, Adam. 1984. *Catalogue of Arabic manuscripts in the library of the Institute of Ismaili Studies*. I. London: Islamic Publications.

——. 1985. *Catalogue of Arabic manuscripts in the library of the Institute of Ismaili Studies*. II. London: Islamic Publications.

——. 1989. "Technical practices and recommendations recorded by classical and post-classical Arabic scholars concerning the copying and correction of manuscripts". *Les manuscrits du Moyen-Orient: Essais de codicologie et paléographie*, ed. François Déroche, 51–60, pl. XXB–XXXIIB. Istanbul and Paris: Institut Français d'Etudes Anatoliennes and Bibliothèque Nationale.

——. 2001. *The Arabic manuscript tradition: A glossary of technical terms and bibliography*. Leiden: E.J. Brill.

Heer, Nicholas. 1969. "Some common abbreviations used in Arabic philological and theological texts". *An-Nashrah* 3:2.17–22.

Maḥfūẓ, Ḥusayn ʿAlī. 1964. *al-ʿAlāmāt wa-r-rumūz ʿinda l-muʾallifīn al-ʿArab qadīman wa-ḥadīṯan*. Baghdad: Maṭbaʿat al-Maʿārif.

Māmaqānī, Muḥammad Riḍā al. 1992. *Muʿjam ar-rumūz wa-l-ʾišārāt*. Beirut: Dār al-Muʾarrix al-ʿArabī.

Oman, Giovanni. 1961. "Abbreviature e sigle nell'arabo moderno". *Oriente Moderno* 41.800–802.

Quiring-Zoche, Rosemarie. 1998. "How al-Buḫārī's '*Saḥīḥ*' was edited in the Middle Ages: ʿAlī al-Yūnīnī and his '*Rumūz*'". *Bulletin d'Etudes Orientales* 50.191–222.

Repp, Hanna. 2001. *Glossar bibiothekarischer Fachtermini: Arabisch–Deutsch*. Wiesbaden: O. Harrassowitz.

Samarrāʾī, ʾIbrāhīm as-. 1987. "al-Muxtaṣarāt wa-r-rumūz fī t-turāṯ al-ʿArabī". *Majalla Majmaʿ al-Luġa al-ʿArabiyya al-ʾUrdunnī* 32.105–114.

Wright, William. 1967. *A grammar of the Arabic language*. 3rd ed. Cambridge: Cambridge University Press.

Ẓufayrī, Maryam Muḥammad Ṣāliḥ aẓ-. 2002. *Muṣṭalaḥāt al-maḏāhib al-fiqhiyya wa-ʾasrār al-fiqh al-marmūz fī l-ʾaʿlām wa-l-kutub wa-l-ʾārāʾ wa-t-tarjīḥāt*. Beirut: Dār Ibn Ḥazm.

Adam Gacek (McGill University)

Ablaut → Apophony

Academies → Language Academies

Accommodation → Speech Accommodation

Acehnese

1. Acehnese as an Islamic language

The population of the Indonesian province of Aceh, located on the northernmost tip of the island of Sumatra, is estimated at 4.7 million (2005). Aceh was one of the first areas in Indonesia to become converted to Islam. By the end of the 13th century, Islam was established in North Sumatra, gradually spreading to other parts of Aceh. The first Sultan of Aceh appeared in the 16th century, and in the early 17th century Aceh was unified by Sultan Iskandar Muda (r. 1607–1636), inaugurating Aceh's golden age, which spanned almost the entire 17th century (Hadi 2004). Fiercely resisting Dutch attempts at colonization, a long and bitter struggle was fought out, known as the Aceh War, which officially lasted from 1873 until 1903, but in fact a 'pacification' never really succeeded. After the Second World War, Aceh became in theory a province of the Republic of Indonesia, which declared its independence on 17 August 1945, but the central government has never been able to crush guerrilla separatist movements operating in the name of Islam. In 1959, the Indonesian government accepted the creation of what was virtually an Islamic state within the nation by giving Aceh the status of a 'Special District'. In 2002, the formal name of the province was officially changed into Nanggroe Aceh Darussalam: *nanggroe* is Acehnese for 'state', whereas the honorific *darussalam* comes from Arabic *dār as-salām* 'abode of peace'. In the same year, as part of a special autonomy package that the Indonesian government hoped would appease separatists, Aceh was granted the right to implement Islamic law (*šarīʿa*).

Adherence to Islam and fluency in the Acehnese language are generally mentioned as the two most defining characteristics of 'Acehneseness' (Wessing 1984:3; Durie 1985:6). Geographically the closest part of the Indonesian archipelago to the Arab world, Aceh is popularly known as *s(eu)ramoe/seurambi Makah* 'front porch of Mecca'. According to Djajadiningrat (1934:II, 900), Aceh and its capital Banda Aceh owed this old epithet to its function as a transit place for pilgrims from the archipelago en route to and from Mecca. The term was also used, however, to designate the strongly Islamic character of the region as is indicated by an episode in the 17th-century Malay work *Bustān as-Salāṭīn* 'Garden of the Sultans', in which admiring visitors of the Acehnese court exclaim: 'Verily, Aceh, the Abode of Peace, is Mecca's front porch of God the Exalted' (Iskandar 1966:68).

At the end of the 19th century, Snouck Hurgronje (1906:I, 18) noted that according to the Acehnese themselves their origins went back to the Arabs, the Persians, and the Turks. The popularity of this myth of pure Islamic roots, which in Snouck Hurgronje's eyes was an 'absurd idea' (1906:I, 18), should be understood, however, against the background of the Aceh War which constituted nothing less than a holy war for the Acehnese. Decades of sanguinary warfare have meanwhile resulted in the distorted view (fostered by Snouck Hurgronje and his contemporaries) that the Acehnese, in contrast to other Indonesian peoples, are 'fanatics'. This image still persists today, but as Wessing (1984:7) has pointed out, the 'fanaticism' is mostly a matter of emphasis on the necessity of being a Muslim rather than a faithful observance of all the tenets of the faith.

The impact of the Aceh War on Acehnese language and literature can hardly be overestimated. Perhaps the most popular work in Acehnese literature of the 19th and 20th centuries is the *Hikayat Prang Sabi* 'Song of the Holy War', which (in different versions) contains exhortations to wage a holy war against the Dutch infidels (Damsté 1928; Hasjmy 1977; Alfian 1992). Branded by the Dutch as subversive literature, a great number of manuscripts of this *hikayat* were confiscated and burnt, but its circulation could not be stopped. In the 1980s, an anthropologist observed that older Gayo men and women (a neighbouring Muslim people living in the highlands of northern Sumatra) could still remember hearing or reading the poem in Acehnese (Bowen 1991:67). For other texts belonging to this specific genre of belligerent literature, see Snouck Hurgronje (1906:II, 100–120), Iskandar (1986:94–120) and Wieringa (1998:298–308).

Even a scanty examination of Djajadiningrat's 1934 two-volume dictionary of nearly 2,400 pages will yield many examples of what Snouck Hurgronje (1906:I, 172) once called "the hatred or at least the contempt felt for all others than Mohammedans in Acheh". For example, whereas the Arabic term for 'uncircumcised' (*ġulf*, plural of *ʾaġlaf*) is used in Javanese as a common term of address for a young boy or son (*kulup*), Djajadiningrat (1934:I, 792) explains that *kulôb* has a very negative connotation in Acehnese (nowadays spelled *kulôp* according to Aboe Bakar a.o. 2001: 466), being used as an opprobrious epithet applied to infidels and a dog's name. Conversely, the word *moseulimin* (Arabic *muslimīn* [plural]; Malay *muslimin* 'Muslims'), which normally means 'Muslim(s)', became synonymous during the Aceh War with 'fighter in the Holy War' and 'waging the Holy War' (Djajadiningrat 1934:II, 120). In the same vein, again according to Djajadiningrat (1934:II. 118), the term *mokmin* 'believer(s)' (Arabic *muʾmin*; Malay *mukmin*) may also have the meaning of 'kind (-hearted), mild, benevolent'.

In Kreemer's (1931) concise dictionary of 367 pages, however, no mention is made of any such connotations of the examples mentioned above. As Djajadiningrat's dictionary was based mainly on written manuscript sources, and many quotations were drawn from contemporaneous literature on the 'holy war', the seemingly pro-Islamic bias in the Acehnese lexicon may be attributed to special wartime circumstances and should not be regarded as yet another proof of Acehnese 'fanaticism'. Incidentally, despite the high status of the Arabic language, the quotations in Djajadiningrat (1934:I, 72) under the entry for *arab* show that the Acehnese shared the common repertoire of dirty jokes featuring sex-crazed Arabs, known throughout Indonesia. Other irreverent and bawdy jokes in Djajadiningrat's dictionary (omitted in the recent dictionary of Aboe Bakar a.o. 2001) concerning mystics/hypocrites (e.g. 1934:I, 75 under *aréh* < Arabic *ʿārif*) or even alluding to verses of the *Qurʾān* in a smutty context (e.g. 1934:I, 31 under *alamtarakòy* < Arabic *ʾa-lam tara kay*[*fa*] . . ., the beginning of *sūra* 105), equally question the received image of the Acehnese as stern zealots

2. Script and literature

The Acehnese language, which has affiliations with languages of the Southeast Asian mainland, belongs to the Aceh-Chamic subgroup within the Austronesian language family (Cowan 1981: 523, 1991:53–83; Durie 1995:407). Acehnese has many dialects and each dialect again many variants (see Durie 1985:4–5, 1995:410–411 for a discussion of linguistic studies of Acehnese with bibliographical references). The Acehnese dialect described in Djajadiningrat's dictionary represents the variety spoken in the neighborhood of Banda Aceh around the turn of the century. In the course of time, however, considerable linguistic change has taken place, while the so-called *banda* dialect has not retained its former prestige (Voorhoeve 1994:20).

Traditionally, Acehnese manuscript literature was written in an Arabic-derived script. At least from the late 18th century until the end of the 19th century a distinctive school of Acehnese manuscript illumination was flourishing, deriving firmly from the broader Islamic tradition (Gallop 2004). During the Aceh War, large numbers of manuscripts in Malay, Arabic, and Acehnese were seized by the Dutch as war booty and subsequently entered Dutch collections. As a result of economic factors and political turmoil, a more recent outpouring of manuscripts from Aceh has taken place over the past two decades, finding their way to libraries in neighboring Malaysia and Brunei (Gallop 2004:194–195).

The traditional Arabic-based script of the manuscripts can give little indication of variation: for example, the Arabic loanword *manfaʿa* 'profit' (Malay *manfaat*), which is spelled <m-n-f-ʿ-tāʾ marbūṭa>, is (erroneously) transliterated as *mena-faât* in Van Langen's dictionary (1889: 266), but Djajadiningrat (1934:II, 67; 70) gives the possibilities *meunapaʿat* and *meunepeuʿat* for the written language, noting the more popular forms of *mupaʿat* or *mupeuʿat* for the spoken language (taken over in Aboe Bakar a.o. 2001:594 as *meunapaʾat/meneupeuʾat* and *mupaʾat/meupeuʾat* respectively). In a more recent dictionary we find the spelling *munap(h)aʿat* (Basry 1994:243; 492). Snouck Hurgronje (1893) developed a Latin orthography, which was adopted by Djajadiningrat (1934). New insights in Acehnese phonology have necessitated a modernization of Acehnese spelling, in keeping with present-day conventions for Bahasa Indonesia (applied in Aboe Bakar a.o. 2001, which for the rest closely follows Djajadiningrat's dictionary).

Short mnemonic descriptions of the letters helped facilitate the learning of the Arabic script, e.g. *aléh meutungkat beusoe* 'the *ʾalif* looks like an iron rod'. Just as in other parts of the Muslim world, in Aceh, too, speculations about letter symbolism of the Arabic alphabet were well-known. For example, in one text the name of the Prophet Muḥammad is symbolically explained as follows: *ban haraih Muḥammad nyata, dum angèeta meukheuluk Rabi; ban ulèe mim ulèe énsan, nyata badan misé haʾi; ban mim akhé pruet nyataʾan, misé day kan dua gaki* 'the body parts of the creatures of our Lord are like all the letters of [the word] Muḥammad; the head of a human being is like the initial *mīm*, the body is like the *ḥāʾ*; the belly is like the final *mīm* and the two legs are like the *dāl*' (Djajadiningrat 1934:II, 89 under *mim*). Letter symbolism, however, also gave rise to such risqué sayings as *aléh lam ha* 'the *alif* in the *hāʾ*', in which these two letters because of their shape are metaphors for the male and female genitals.

Acehnese has a rich oral and written literature, for which the reader is referred to Snouck

Hurgronje (1906:II, 66–189). Voorhoeve's (1994) catalogue can be read as a bibliographical complement. Important text editions are those of Drewes (1979, 1980) and Abdullah (1991), which not only contain transliterations, translations, and commentaries of some classical 'epic' poems, but also discuss Acehnese literature in general and have valuable appendices and bibliographies covering nearly all publications on Acehnese literature.

To the orally transmitted literature belong the *miseue* 'proverb, saying' (< Arabic *miṯāl*; Malay *misal*) and the *haba* 'story' (< Arabic *xabar*; Malay *kabar*; Minangkabau *kaba*), among others the *haba jameun* 'story of bygone days' (< Arabic *zamān*; Malay *jaman*) and the *hadih maja* 'tales or traditions of female ancestors' (< Arabic *ḥadīṯ*; Malay *hadis*). The most important genre of written Acehnese literature is the *hikayat* which, unlike the Arabic *ḥikāya* and the Malay *hikayat*, is always in poetical form, written in *sanjak* (< Arabic *sajʿ*; Malay *sajak*). This truly Acehnese poetic meter finds its parallel in a similar meter in Cham (Cowan 1933:149–155). Every line of poetry, called *ayat* (< Arabic *ʾāya*; Malay *ayat*), numbers four times four units or syllables. *Hikayat* literature is governed by a strict set of conventions. One of the recognized characteristics is that it should commence with a *kōteubah* (< Arabic *xuṭba* 'sermon, lecture'), i.e. an introduction with "certain formulas in praise of Allah and his Apostle, to which are sometimes appended other general views or reflections of the author's own, till finally the actual subject is reached" (Snouck Hurgronje 1906:II, 77; cf. Abdullah 1991:17, 30–31; Wieringa 1998: 298–308).

In *hikayat* literature the phrase *ajayéb sōbeuhanalah* (*taleungo lōn kisah saboh calitra*) 'Wonderful! Astonishing! (Listen, I am going to tell a story)' is typically used to announce the introduction of a story or a chapter (< Arabic *ʿajāʾib* and *subḥāna llāh*; Malay *ajaib* and *subhanallah*). The original meaning of *sōbeuhanalah* as 'Praise be to God' is not felt any more: the syllables are usually divided into *ajayéb sō / beuhan alah //* and the fact that *sōbeuhan* is all one word is obliterated. The equally formulaic expression *ama ba'du* 'further' (Arabic *ʾammā baʿdu*; Malay *ammaba'du*), often followed by the Acehnese translation *dudoe nibak nyan* 'after this', serves to introduce a new subject.

3. The important role of Malay

The oldest borrowings into the Acehnese lexicon are from Mon-Khmer and Sanskrit, whereas for centuries Malay has continued to be the most important source of borrowings. An old inscription, dated 1380 C.E., from Pasai (Aceh), comprises a Malay poem in Indian script and meter with various Arabic words and Muslim imagery. This text forms important proof that already at an early stage Malay was apparently considered the appropriate literary language for official (commemorative) purposes (Stutterheim 1936: 268–271; Marrison 1951:162–165; Teeuw 1959:149). Malay was not only the trading language, but also the prestige language of the royal courts and Islamic scholarship. During its golden age, in the 17th century, Aceh was one of the most powerful centers in which Malay literature flourished, resulting in the creation under royal patronage of such important early Malay works as a panegyric of Sultan Iskandar Muda (Iskandar 1958; Penth 1969), and two encyclopedic 'Mirrors for Princes', viz. *Tāj as-Salāṭīn* 'Crown of the Sultans' by a certain Buxārī al-Jawhārī in 1603 and the voluminous *Bustān as-Salāṭīn* (mentioned above) by Nūr ad-Dīn ar-Rānīrī in 1638 (on the rise of a Malay Islamic literature in Aceh, see Andaya 2001:45–50). Furthermore, some early 17th-century letters from the Sultan of Aceh belong to the oldest extant manuscripts in Malay (Shellabear 1897: 107–151).

In fact, the first Western beginnings of the study of Malay are language descriptions of its regional variety in Aceh. Frederick de Houtman acquired his knowledge of spoken Malay in an Acehnese prison and published his lexicographical work in 1603 (Lombard 1970). A much shorter Dutch-Malay vocabulary, compiled by van Elbinck in Aceh, is dated 1 June 1604 (van Ronkel 1896:13–18). It has been suggested that all Malay manuscripts in the collection of Thomas Erpenius (1584–1624), among others a late 16th-century Malay translation of al-Būṣīrī's *al-Burda* (Drewes 1955), may originally have been acquired by van Elbinck in Aceh around 1604, but more study is needed to confirm this hypothesis. Because of Aceh's important position, replacing Malacca (after the Portuguese capture in 1511) as the spiritual and intellectual capital in the Malay world, al-Attas

(1988) has claimed that the 'oldest known Malay manuscript', a Malay translation of the *ʿAqāʾid*, a popular catechism of an-Nasafī (d. 537/1142), dating from the latter half of the 16th century, must have been written in Aceh. The earliest extant Acehnese adaptations of Arabic and Malay works date from the 17th century (Voorhoeve 1952:335–345).

It remains a moot point through which pathways Arabic words entered the Acehnese lexicon, but in view of such circumstances as the profound Malay character of the premodern Acehnese Sultanate and the intensive interaction between Malay and Acehnese literature (cf. Snouck Hurgronje 1906:II, 121–123; Andaya 2001:46), Durie's (1995:410) bold statement that most Arabic borrowings were derived directly from Arabic, while only some came via literary Malay, would seem to be much too strong.

Both Kreemer (1931) and Djajadiningrat (1934) have indicated the sources of borrowings into Acehnese, but are of little help in establishing the 'donor' language. Kreemer did not always recognize the exact derivation of words, and as Cowan (1981:522) puts it, his dictionary is "handy but not always altogether reliable, [it] mentions many doubtful Malayisms as if they were true Acehnese". As a rule Djajadiningrat (1934) always gives the ultimate source, rarely mentioning the possibility of indirect borrowing through Malay. One of the very few examples in which Djajadiningrat (1934:II, 83) explicitly mentions the intermediary role of Malay is the word *meuseutay* 'ruling-board' which in his opinion comes from Arabic *misṭara* through Malay *mistar*.

It is impossible on the basis of Djajadiningrat's dictionary to draw conclusions on the size and currency of the Arabic loan stock. Djajadiningrat included, for example, such words as *aphiet* 'good health' (< Arabic *ʿāfiya*; Malay *afiat*) and *amba* 'ambergris' (< Arabic *ʿanbar*; Malay *ambar*), about which he remarked that they were little known and only rarely used. It should be noted that a considerable part of the Arabic loan stock is restricted to *hikayat* literature, which is very much indebted to Malay examples. An example of the influence of Malay *hikayat* literature upon Acehnese *hikayat* literature is the specific meaning of the Arabic loanword *bidʿa* (Malay *bidaʾah*). The Acehnese word *bédeuʿah* or *beudeuʿah* has the same meaning as in Arabic and Malay, viz. 'innovation', more often with the negative connotation of 'heresy', but according to Djajadiningrat (1934:II, 154) it is used in Acehnese *hikayat* literature as a term of abuse ('accursed; wretched'). Djadiningrat does not explain this development, but in all probability this semantic shift was influenced by Malay *bedebah* (< Persian *badbaxt*, cf. Bausani 1974:353), which is a stock term of abuse in Malay literature and phonetically closely matches *bédeuʿah*/*beudeuʿah*.

4. The preponderance of Arabic loanwords

The large amount of Arabic loanwords in Djajadiningrat's dictionary may easily create the false impression that the presence of Arabic in Acehnese is all-pervasive and much more pronounced than in other Indonesian regional languages. It should be remembered, however, that Djajadiningrat simply tried to register as many words as possible, and as he also included many religious works among his main sources, the size of the Arabic loan stock inevitably became extensive.

Generally speaking, loans exclusively belonging to the literary and theological domain are not normally used in daily communication and are to a large extent even unknown to most native speakers (cf. al-Harbi 1991:102), but on the other hand some technical terms may have had a wider currency than perhaps expected. For example, a highly specialized word like *rakibah* 'guardian of the page', i.e. 'catchword' (cf. Latin *custos*), from Arabic *raqībah* (feminine of *raqīb*) is unknown in Malay lexica and is not even listed as such in (European) dictionaries of Arabic or even in the specialized glossary of Gacek (2001). Judging from Djajadiningrat's quotation (1934:II, 472), however, *rakibah* seems to have been a fairly common word, but it is not included in other Acehnese dictionaries (apart from Aboe Bakar a.o. 2001, of course, which is based on Djajadiningrat 1934).

Understandably, religious terms are very numerous among the Arabic loanwords. Djajadiningrat lists such Islamic technical terms as *amè* 'collector of *zakāt* (Acehnese *jakeuet*)' (< Arabic *ʿāmil*; Malay *amil*); *peureulèe kipayah* 'collective obligation' (< Arabic *farḍ al-kifāya*; Malay *fardul-kifayah*), and *pasah* 'to annul a marriage' (< Arabic *fasx*; Malay *pasah*). A special subsection are the appellations of the chapters of the *Qurʾān*, e.g. *aleuham* (after the opening *al-ḥam*[*du li-llāhi*]) for *sūra* 1 and *alépeulam* for *sūra* 2 (< *ʾalif-lām-mīm*).

The Islamization process resulted in a radical transformation of the rhythm of daily and annual life: not only the names of the times of prayer, days of the week, months, and the annual festivals are Arabic loanwords, but also the very word for 'time' itself, viz. *wa(k)tèe* (Arabic *waqt*; Malay *waktu*), is of Arabic origin (cf. Meuleman 1994:16 for exactly the same situation in Malay). Personal names, too, were Islamized: Abaih (< ʿAbbās); Usén (< Ḥusayn); Uma (< ʿUmar); Usu(i)h (< Yūsuf). The popular name Muḥammad is rendered as Muhamat, Mamat, or Mat. Such formulaic expressions as *aseuta(g)pirōlah* 'God forbid!' (< Arabic *ʾastaġfiru llāh*; Malay *astagfirullah*) and *aleuhamdulélah* (Arabic *al-ḥamdu li-llāh*; Malay *alhamdulillah*) spice everyday conversation.

The high status of Arabic is indicated by the use of Arabic loanwords functioning as euphemistic or respectful terms alongside words from other origins. The examples for Malay given by Jones (1984:14) concerning pregnancy and death are exactly the same as in Acehnese: *hamè* (Arabic *ḥāmil*; Malay *hamil*) for 'pregnant' is considered to be more refined than *bunténg* (Malay *bunting*), while *ma(n)yèt* 'corpse' (Arabic *mayyit*; Malay *mayat*) is used of human beings (especially Muslims) and *bangké* (Malay *bangkai*) can only be used in a derogatory way for human beings (sometimes of unbelievers). A recent loan in this field seems to be *jimak*, i.e. a more refined term for 'sexual intercourse' (< Arabic *jimāʿ*; probably via Indonesian *jimak*), which is included in the newer dictionaries of Basry (1994) and Aboe Bakar a.o. (2001), but which was omitted by the older lexicographers Kreemer and Djajadiningrat, who give only the coarse terms.

In general it may be said that the same semantic fields which have been proposed for Malay are equally applicable in the case of Arabic loanwords in Acehnese. Many semantic shifts are also the same: for example *mudim/mudém* (Arabic *muʾaḏḏin*; Malay *modin/mudin*) does not have the Arabic meaning of 'muezzin' but, exactly as in Malay, denotes a 'circumciser'.

5. ASSIMILATION

Loans are assimilated to the Acehnese phonological patterns, which may occasionally result in homonyms: for example, *peutua* (also spelled *peutuha*) can mean 'headman, village elder' (root *tu(h)a* 'old'), but it can also mean 'considered legal opinion; ruling' (< Arabic *fatwā*). According to Versteegh (2001:500; 2003), the occurrence of /l/ as the reflex of Arabic /ḍ/ may be indicative of loanwords belonging to the earliest period. The example of *halé* 'to be present' (< Arabic *ḥāḍir*) is used in daily conversation when offering a meal: *ka halé béseumélah* 'it is ready, begin please' (*béseumélah* is from the common invocation *bi-smi llāh* 'in God's name'). A more detailed analysis of the assimilation process, concentrating on phonetical phenomena, was published by al-Harbi (1991). The special case of words ending in *-u/-i*, to which Versteegh (2003) has drawn attention for Malay, also exists in Acehnese: *napeusu* 'lust, passion' (= Malay *napsu/nafsu*, idem; Arabic *nafs* 'mind, soul'), *wahi* 'revelation' (= Malay *wahi/ wahyu*; Arabic *waḥy*), and *nèseupu* 'half' (= Malay *nisfu*; Arabic *niṣf*). This remarkable phenomenon of identical sound changes in Malay and Acehnese would further point to the importance of Malay as donor language (cf. Cowan 1981:547 for a discussion with different arguments as to why *nafs* probably entered Acehnese via Malay).

Many Arabic loanwords not only exist as independent words, but may also undergo modification as roots with affixation. For example, *abeudi* 'slave' (< Arabic *ʿabd*; Malay *abdi*) is a nominal, appearing as a root in the derived verbal forms *meuabeudi* and *peuabeudi*. The prefix *meu-* with a nominal as root derives a verb with the general meaning of 'the Agent acts as X, makes himself to be X', thus *meuabeudi* means 'to act as a slave' as in the expression *meuabeudi akayji* 'his nature (mind) is (as despicable) as of a slave' (*akay* < Arabic *ʿaql*; Malay *akal*; *-ji* denotes familiar 3rd person). The prefix *peu-* with a nominal as root derives a verb with the general meaning of 'to treat the undergoer as X', thus *peuabeudi* means 'to treat as a slave'.

Loanwords beginning with *mu-* are often (erroneously) regarded as derived verbal forms with prefix *mu/meu-*: for example, the root of *mupakat* (Malay idem), properly speaking borrowed from Arabic *muwāfaqa* (root *w-f-q*), is believed to be *pakat* 'agreement; discussion' and so we also find such forms as *sipakat* 'agreed; unanimously' (Malay *sepakat*). The word *mumida* (also pronounced as *mumanyèh* and *mumeunyèt*, tentatively [and erroneously] romanized in van Langen 1889:264 as *mumiza* [?], spelled <m-m-y-ḏ-a>), which is borrowed from *mumayyiz*, i.e. a technical term in Islamic law denoting 'intelligent, discriminating' minors (Malay *muma(y)iz*), is regarded as a derivation with prefix *m(e)u-* and the root *mida*, which in turn is seen as a variant of *bida* 'difference' (< Malay *beda*), so *mumida* is a variant of *mubida* 'come to the age of discretion, come of age'.

Bibliographical references

Abdullah, Imran Teuku. 1991. Hikayat meukuta alam*: Suntingan teks dan terjemahan beserta telaah struktur dan resepsi* [*Hikayat meukuta alam*: Text edition and translation with an analysis of the structure and reception]. Jakarta: Intermasa.

Aboe Bakar a.o. 2001. *Kamus bahasa Aceh-Indonesia* [Acehnese-Indonesian dictionary]. Jakarta: Balai Pustaka.

Al-Attas, Syed Muhammad Naquib. 1988. *The oldest known Malay manuscript: A 16th century translation of the* ʿAqāʾid *of al-Nasafī*. Kuala Lumpur: University of Malaya.

Alfian, Teuku Ibrahim. 1992. *Sastra perang: Sebuah pembicaraan mengenai* Hikayat Perang Sabil [War literature: A discussion of the *Hikayat Perang Sabil*]. Jakarta: Balai Pustaka.

Al-Harbi, Awwad Ahmad al-Ahmadi. 1991. "Arabic loanwords in Acehnese". *Perspectives on Arabic linguistics*, III, ed. Bernard Comrie and Mushira Eid, 93–117. Amsterdam and Philadelphia: J. Benjamins.

Andaya, Leonard Y. 2001. "Aceh's contribution to standards of Malayness". *Archipel* 61.29–68.

Basry, M. Hasan. 1994. *Kamus umum Indonesia-Aceh* [General Indonesian-Acehnese dictionary]. Jakarta: Yayasan Cakra Daru.

Bausani, Alessandro. 1974. "Notes sur les mots persans en malayo-indonésien". *Acta Iranica* 2.347–379.

Bowen, John. 1991. *Sumatran politics and poetics: Gayo history, 1900–1989*. New Haven and London: Yale University Press.

Cowan, H.K.J. 1933. "Het Atjèh'sch metrum 'sandja' in verband met een Tjamsch gedicht" [The Acehnese metre 'sandja' in connection with a Cham poem]. *Bijdragen tot de Taal-, Land- en Volkenkunde* 90.149–155.

——. 1981. "An outline of Acehnese phonology and morphology". *Bulletin of the School of Oriental and African Studies* 44.522–549.

——. 1991. "Achehnese dialects in connection with Chamic migrations". *Vical 2. Western Austronesian and contact languages: Papers from the fifth international conference on Austronesian linguistics, Auckland, New Zealand 1991*, I, II, ed. Ray Harlow, 53–83. Auckland: Linguistic Society of New Zealand, University of Auckland.

Damsté, H.T. 1928. "Hikajat Prang Sabi". *Bijdragen tot de Taal-, Land- en Volkenkunde* 84.545–609.

Djajadiningrat, Hoesein. 1934. *Atjèhsch-Nederlandsch woordenboek* [Acehnese–Dutch dictionary]. 2 vols. Batavia: Landsdrukkerij.

Drewes, G.W.J. 1955. *Een 16de eeuwse Maleise vertaling van de* Burda *van al-Būṣīrī (Arabisch lofdicht op Mohammad)* [A 16th century Malay translation of the *Burda* by al-Būṣīrī (Arabic panegyric to Muḥammad)]. The Hague: Nijhoff.

——. 1979. Hikajat Potjut Muhamat: *An Achehnese epic*. The Hague: Nijhoff.

——. 1980. *Two Achehnese poems:* Hikajat Ranto *and* Hikajat Teungku di Meuké. The Hague: Nijhoff.

Durie, Mark. 1985. *A grammar of Acehnese on the basis of a dialect of North Aceh*. Dordrecht and Cinnaminson: Foris Publications.

——. 1995. "Acehnese". *Comparative Austronesian dictionary: An introduction to Austronesian studies*, I:1, ed. Darrell T. Tryon, 407–420. Berlin and New York: Mouton de Gruyter.

Gacek, Adam. 2001. *The Arabic manuscript tradition: A glossary of technical terms and bibliography*. Leiden: E.J. Brill.

Gallop, Annabel Teh. 2004. "An Acehnese style of manuscript illumination". *Archipel* 68.193–240.

Hadi, Amirul. 2004. *Islam and state in Sumatra: A study of seventeenth-century Aceh*. Leiden: E.J. Brill.

Hasjmy, A. 1977. *Apa sebab rakyat Aceh sanggup berperang puluhan tahun melawan agressi Belanda* [What is the reason why the Acehnese people were able to fight for decades against Dutch aggression]. Jakarta: Bulan Bintang.

Iskandar, Teuku. 1958. *De* Hikajat Atjéh [The *Hikayat Aceh*]. The Hague: Nijhoff.

——. 1966. *Nuru'd-din ar-Raniri Bustanu's-Salatin bab II, fasal 13*. Kuala Lumpur: Dewan Bahasa dan Pustaka.

——. 1986. "The Hikayat Prang Geudōng". *Cultural contact and textual interpretation: Papers from the fourth European colloquium on Malay and Indonesian studies, held in Leiden in 1983*, ed. C.D. Grijns and S.O. Robson, 94–120. Dordrecht and Cinnaminson: Foris Publications.

Jones, Russell. 1984. "Loan-words in contemporary Indonesian". *Towards a description of contemporary Indonesian: Preliminary studies*, II, ed. John W.M. Verhaar, 1–38. Jakarta: Universitas Atma Jaya.

Kreemer, J. 1931. *Atjèhsch handwoordenboek (Atjèhsch-Nederlandsch)* [Concise Acehnese dictionary (Acehnese-Dutch)]. Leiden: E.J. Brill.

Langen, K.F.H. van. 1889. *Woordenboek der Atjehsche taal* [Dictionary of the Acehnese language]. The Hague: Nijhoff.

Lombard, Denys (with Winarsih Arifin and Minnie Wibisono). 1970. *Le* "Spraeck ende woordboek" *de Frederick de Houtman: Première méthode de malais parlé (fin du XVI*^{e} *s.)*. Paris: Ecole Française d'Extrême-Orient.

Marrison, G.E. 1951. "A Malay poem in Old Sumatran characters". *Journal of the Malayan Branch of the Royal Asiatic Society* 24.162–165.

Meuleman, Johan H. 1994. "Arabic in Indonesia". *Indian Journal of Applied Linguistics* 20.11–34.

Penth, Hans. 1969. Hikajat Atjeh: *Die Erzählung von der Abkunft und den Jugendjahren des Sultan Iskandar Muda von Atjeh (Sumatra)*. Wiesbaden: O. Harrassowitz.

Ronkel, Philippus Samuel van. 1896. "Account of six Malay manuscripts of the Cambridge University Library". *Bijdragen tot de Taal-, Land- en Volkenkunde* 46.1–53.

Shellabear, W.G. 1897. "An account of some of the oldest Malay MSS now extant". *Journal of the*

Straits Branch of the Royal Asiatic Society 31.107–151.

Snouck Hurgronje, C. 1893. "Studiën over Atjèhsche klank- en schriftleer" [Studies on Acehnese phonology and orthography]. *Tijdschrift voor Indische Taal-, Land- en Volkenkunde* 35:346–442. (Repr. *Verspreide Geschriften*, V, 169–247. Bonn and Leipzig: Schroeder, 1925.)

——. 1906. *The Achehnese*. 2 vols. Leiden: E.J. Brill.

Stutterheim, W.F. 1936. "A Malay sha'ir in Old-Sumatran characters of 1380 A.D. (with 4 plates)". *Acta Orientalia* 14.268–279.

Teeuw, A. 1959. "The history of the Malay language". *Bijdragen tot de Taal-, Land- en Volkenkunde* 115.138–156.

Versteegh, Kees, 2001. "Linguistic contacts between Arabic and other languages". *Arabica* 48.470–508.

——. 2003. "The Arabic component of the Indonesian lexicon", *Rintisan kajian leksikologi dan leksikografi* [Preliminary studies of lexicology and lexicography], ed. Lilie Suratminto and Munawar Holil, 216–229. Jakarta: Fakultas Ilmu Pengetahuan Budaya Universitas Indonesia.

Voorhoeve, P. 1952. "Three old Acehnese manuscripts". *Bulletin of the School of Oriental and African Studies* 14.335–345.

——. 1994. *Catalogue of Acehnese manuscripts in the library of Leiden University and other collections outside Aceh*. Leiden: Leiden University Library (Legatum Warnerianum) in co-operation with Indonesian Linguistics Development Project (ILDEP).

Wessing, Robert. 1984. "Acehnese". *Muslim peoples: A world ethnographic survey*, ed. Richard V. Weekes, 3–10. 2nd ed. London: Aldwych Press.

Wieringa, Edwin. 1998. "The dream of the king and the holy war against the Dutch: The kôteubah of the Acehnese epic, *Hikayat Prang Gômpeuni*". *Bulletin of the School of Oriental and African Studies* 61.298–308.

Edwin Wieringa (University of Cologne)

Acronyms → Abbreviations

Address, Terms of → Terms of Address

Adjective Phrase

1. Definition

In contrast with nouns and verbs, → adjectives are a disputed category. Arabic has figured prominently in discussions about whether they constitute a special word class, or just special cases of nouns or verbs (see, e.g., Hengeveld 1992; Bhat 1994; Baker 2003). Here, we adopt the minimal theoretical requisite in the matter of word classes, to be found in the so-called 'functionalist' perspective of Croft (1991) or in the 'paraminimalist' approach of Baker (2003): the *unmarked* function of nouns is reference; the *unmarked* function of verbs is predication; a third function is property assignment. A word-form having this as its unmarked function may be called an adjective. The interesting question, therefore, is not whether a language 'has' adjectives or not, but which word-forms assign properties, and whether their morphosyntactic identity is that of nouns, verbs, neither, or both.

The question of the lexical category of property words in Arabic is addressed first. Then the various shapes and morphological structures of such words are examined. Finally agreement and concord phenomena are dealt with.

2. The lexical category of property words

The main problem posed by Arabic in this matter is the concurrence of two kinds of property words. On the one hand, we find lexemes with various templates, including the 'minimal' *faʿl* template, which seem to behave just like attributive or predicative adjectives, e.g. *sahl* 'easy' (see list in Wright 1991:I, 133 and → adjective). On the other hand, there are so-called 'stative' verbs having /u/ or /i/ as the middle vowel of the Perfect in Form I, usually translated as 'to be *p*', e.g. *sahula* 'to be easy', *fariḥa* 'to be glad' next to *fariḥ* 'glad'. The question then is what is the difference between examples such as *hāḏā l-ʾamr sahl* and *sahula hāḏā l-ʾamr*, both apparently translatable as 'this business [is] easy'.

In fact, the translation of 'stative' *faʿula* or *faʿila* 'statives' as 'to be *p*' is misleading. As shown in Cohen (1984:137ff.; see also Fassi Fehri 1993:175–190; Goldenberg 1995), such lexemes always retain processual force and they contrast in this respect with the predicative adjectives sharing the same root, e.g. *mā kāna karīm fa-karuma* 'he was not generous, but he became generous' (Cohen 1984:140). In other cases, the subjective process of perceiving the property is what is implied, as in *ṯaqula ʿalay-hi l-ʾamr* [lit. 'became-heavy on-him the-business'] 'the business weighed on him' (i.e. he found it heavy). It would be unacceptable to say *?*ṯaqula hāḏā* *l-lawḥ* to mean 'this board is heavy'; only *hāḏā l-lawḥ ṯaqīl* will do (Jean-Patrick Guillaume, p.c.). Stative verbs are thus 'real' verbs.

Adjectives such as *sahl* or *karīm*, in contrast, assign stable and/or 'objective' properties and they enter the same types of construction as do nouns. Thus, there is no formal difference (except for agreement) between *hāḏihi l-marʾa karīma* 'this woman [is] generous' and *hāḏihi l-marʾa ʾuxtuhu* 'this woman [is] his sister'. In attributive function, adjectives follow (immediately or not) the noun they modify and they share a referential index, e.g. *al-marʾa al-karīma* 'the generous woman', *imraʾa(tun) karīma(tun)* 'a generous woman', *ʾabū l-marʾa(ti) l-karīm(u)* 'the generous father of the woman'. Such phrases bear a striking similarity to relative clause constructions. There may indeed be a historical connection, insofar as the relative pronoun *allaḏī/allatī* clearly involves the definite determiner *l-*, and it is not inserted when the noun phrase is indefinite. Yet, claiming that [*al-karīma*] in *al-marʾa al-karīma* is a relative clause would amount to imparting verbhood to the adjective, and it would blur the formal and semantic distinction between such noun phrases and *al-marʾa allatī karumat* 'the woman who became generous'. One should therefore side with the more conservative account, according to which *(al-)karīma* is a noun phrase – quite commonly used by itself to mean 'the/a generous one' – which stands in apposition (possibly multiple as in *al-kawkabu n-nayyiru l-ʾaḥmar* 'the bright red star') to the noun phrase projected from the 'head' noun. No attempt will be made to be more specific about the syntax of → appositions.

Property items like *sahl* or *karīm* are therefore nouns. They differ from other nouns, however, in having no autonomous value for definiteness, gender, number, and case. Correspondingly, predicative adjectives such as *karīma* in *hāḏihi l-marʾa karīma* 'this woman [is] generous' should be analyzed as noun predicates involving a phonologically unrealized predicator node, probably an entity different from a 'zero copula' (Cohen 1984; Baker 2003).

The nounhood of Arabic property words is further enhanced by the fact that a Construct State Nominal can be used as the equivalent of a noun-adjective phrase, e.g. *ṣāḥibu ṣidq(in)* 'a true friend' [lit. 'a friend of truth' = *ṣāḥib(un) ṣadīq(un)*]. Moreover, adjectives can head Construct State Nominals as in *rijāl(un) ḥisān(u) l-wujūh(i)* 'handsome men' [lit. 'men handsome of the faces' = 'handsome-faced men']. Nouns not in a Construct State Nominal can also be used adjectivally, as in Wright's (1991:II, 274) example *imraʾa(tun) ʿadl(un)* 'a just woman' [lit. 'a woman justice'].

Finally, active and passive participles (*fāʿil* and *mafʿūl*) can be used as property words, the former retaining the verbal capacity to govern accusative complements.

3. Forms and morphological structures of property words

A root-and-site approach is used, according to which (consonantal) roots are uncategorized, and verbs and nouns differ depending on whether the V or N verbalizing or nominalizing functional element identifies a dedicated derivational site internal to the root (Guerssel and Lowenstamm 1990; Kihm 2003, 2004).

Simple adjectives consist in a root plus one or two vowels devoid of morphological function and often variable, e.g. *ṣaʿb* 'difficult', *ḥasan* 'handsome', *faṭun* ~ *faṭin* 'clever'. Their morphological structure therefore contains the root and N. As already mentioned, the difference between, for instance, *ṣaʿb* 'difficult' and *šayx* 'old man' is that the latter (call it a substantive noun) bears an inherent value for class and gender (human, masculine) and is given definiteness, number, and case values according to the speaker's choice and its syntactic position, whereas the former (the adjective noun) receives all these values from the substantive noun it modifies. This difference can be formalized by assuming that N in Arabic is a set comprising two members, one unmarked and one marked: N_{-F} as in *šayx* and N_{F} as in *madīna* 'city' or *šams* 'sun' (see Kihm, 2005, for the view of → gender as a value of N). In adjective nouns, N has no inherent value. It must obtain one, however, because bare N is uninterpretable (in Arabic), hence the gender concord (see next section). The uninterpretable absence of a value for N entails that adjective nouns are unable to refer unless they get a referential index from a substantive noun (possibly implicit). Definiteness, number, and case concord follows from this. Perhaps one may see a connection between this approach and the traditional grammarians' notion that "every adjective contains a pronominal agent within itself" (Wright 1991:II, 284; see also Goldenberg 1995) insofar as the complex [Definiteness$_x$, Gender$_y$, Number$_z$, Case$_w$] with variables as values may indeed be regarded as an abstract pronoun.

Other adjective templates manifest morphological activity in the root site, e.g. *jabān* 'cowardly', *kurām* 'generous', *karīm* 'generous', etc. It consists in inserting a glide in the Nominal Derivational Site between C$_2$ and C$_3$, which is also used for broken plural and *maṣdar* formation (Asfour 2001; Kihm 2003). For instance, *jabān* = {J.B.{C$_A$.}N.}, where the low glide /A/ identifying a C position surfaces as the long vowel [ā]. Such adjective nouns are supposed to add an intensive touch to their basic meanings. Intensification is certainly implied by the *faʿʿāl* template, e.g. *ʾakkāl* 'gluttonous, a glutton'. Less common templates involve it as well, e.g. *kubbār* 'very large', *fārūq* 'very timorous'.

Special mention must be made of the *ʾafʿal(u)* diptotic template. It is the template of the so-called → 'elative' as in *ʾashal* 'easier, easiest', as well as of those adjective nouns that denote conspicuous properties, traditionally known as adjectives of color and defect, e.g. *ʾaḥmar* 'red', *ʾaṭraš* 'deaf'.

The last adjective noun type covered here is the so-called 'relative adjective' or *nisba* 'relation', derived from noun bases with the ending *-iyy*, as in *ʾarḍī* 'earthly', *ʿilmī* 'scientific', etc. (Fleisch 1961:434 ff.). Note that, although this is frowned upon in Classical Arabic, the derivation base may be a broken plural, as in *kutubī* 'bookseller' < *kutub* 'books' (sg. *kitāb*). As this example shows, relative adjectives can be substantivized, their feminine singular being a common source of abstract nouns, e.g. *māhiyya* 'substance, quiddity' < *mā* 'what?'.

There are also a number of adjective-like items such as → demonstratives, → numerals, etc. that must be ignored due to space limitations, just like the → participles.

4. The syntax of adjective nouns: concord and agreement

Concord, following Wechsler and Zlatić (2003), will be used here to refer to feature value sharing within the noun phrase as in *al-marʾa(tu) l-karīma(tu)* 'the generous woman', and agreement will be used to refer to the corresponding process involving a subject and its predicate as in *ʾinna l-marʾa(ta) karīma(tun)* 'the woman is generous'. Agreement differs from concord in Arabic in that neither definiteness nor case need be shared. In the above example, *al-marʾa(ta)* is definite and accusative because of the discourse particle *ʾinna* vs. *karīmatun*, indefinite and nominative. Only gender must be shared. (Note the feminine endings: *-a(tun)*, the most common, *-ā* as in *kubrā* 'bigger [fem.]', and *-āʾ(u)* as in *ḥamrāʾ(u)* 'red [fem.]'). Except for a few epicene adjectives, always appearing in the masculine, e.g. *imraʾa ṣabūr wa-šakūr* 'a patient and grateful woman', *imraʾa ḥāmilun* 'a pregnant woman', all features must be shared in the concord relation. A difficulty arises with the so-called 'indirect modification' as in *raʾaytu mraʾatan ḥasanan wajhuhā* 'I saw a woman with a handsome face', in which *ḥasanan* 'handsome' shares accusative case 'by attraction' with preceding *mraʾatan* 'a woman', but has the gender and number features of *wajhuhā* 'her face', which it modifies. Such constructions should not be confused with the adjective-headed Construct State Nominals mentioned above (cf. *imraʾa ḥasana wajh(in)* 'a handsome-faced woman'). This issue will be left aside here.

Number introduces a complication. Adjective nouns modifying or predicated of broken plurals appear in the singular, and they are feminine, whatever the gender of the substantive noun, e.g. *ʾumūr sahla* 'easy affairs', *nisāʾ karīma* 'generous women', unless the latter denotes a male human being, in which case the adjective noun is also pluralized, in the broken mode if possible, e.g. *rijāl ṭiwāl* 'tall men', *rijāl ṣāliḥūn* 'virtuous men'. The first type is called 'deflected', whereas the second is said to be 'strict'. Note that in Modern Standard Arabic, substantive nouns denoting female human beings regularly entail strict concord and agreement, so that *nisāʾ karīmāt* is used instead of *nisāʾ karīma* and *an-nisāʾ(u) karīmāt(un)* 'the women are generous', instead of *an-nisāʾ(u) karīma(tun)*. Note further that sound plurals of substantive nouns denoting non-human entities entail deflected concord and agreement just like their broken counterparts (e.g. *ijtimāʿāt musʾima* 'boring meetings', not **musʾimāt*, at least not in Modern Standard Arabic). The relevant feature for the division is therefore the human or non-human reference of the substantive noun – in the sense of noun class – rather than pluralization type. This is only a sketch; a full picture would lead to many more complications than can be dealt with here. In the singular, concord and agreement are always strict, and likewise in the dual, despite rare examples of deflection, as in *mīlān ʾifranjiyya* 'two European miles' (instead of *mīlān ʾifranjiyyān*).

A fully formalized treatment of deflected concord-agreement cannot be attempted here. The generalization that emerges is that the feature set {non-human, plural} receives a non-rigid, collective interpretation in the substantive noun phrase (Kwon and Zribi-Hertz 2004; see also Link 1983). This explains why the concording or agreeing adjective turns up in the singular. The feminine ending could be explained by the fact that one function of the feminine in Arabic is to extract individuals from collections, as in *baqar* 'cows' vs. *baqara* 'a cow'. This mental operation has a symmetric counterpart, namely converting collections into second-order, abstract individuals fit for referring to undifferentiated masses or to kinds (as in 'The cat was first domesticated in Egypt'). Putting both observations together leads us to assume that the feminine in Arabic is the gender of *derived individuals*, with the derivation proceeding 'downward', to form real individuals, or 'upward', to form abstract individuals. Compare Chierchia's (1998) operators: 'up', from property to kind, and 'down', from kind to property. This also explains why femininization derives abstract nouns such as *māhiyya* (see above). The effect is clearly seen in the Syrian dialect of Damascus, where *il-kutub mā bihimm-hū* with the verb in the 3rd person plural means 'The books [in question] don't interest him' and contrasts with *il-kutub mā bithimm-hū*, with the verb in the 3rd person feminine singular, meaning 'Books [in general] don't interest him' (see Holes 1995:166). Only the latter agreement is grammatical in Modern Standard Arabic.

5. Conclusion

Conventional wisdom has it that adjectives belong to the verbal paradigm, they are 'like the verb' (*šibhu l-fiʿl*). In other words, *sahl* 'easy' is to *sahula* 'to be(come) easy' what the verbal noun *ʿilm* 'knowledge' is to *ʿalima* 'to know'. There is much truth in this, provided 'paradigm' is understood in a broad sense, i.e. in the sense that English *knowledge* is a member of the extended paradigm of KNOW, defined as a root or an abstract lexeme, rather than in the way *knows* or *knowing* are forms of the verb *(to) know*. The morphological foundation of Arabic adjective nouns is thus the root itself, not a categorized representative of the root. There is one exception to this claim, namely relational adjectives, such as *ʿilmī* 'scientific', which are indeed derived from fully specified bases, agglutinatively rather than according to the dominant, inner site mode.

Bibliographical references

Baker, Mark C. 2003. *Lexical categories: Verbs, nouns, and adjectives.* Cambridge: Cambridge University Press.

Bhat, D.N.S. 1994. *The adjectival category: Criteria for differentiation and identification.* Amsterdam and Philadelphia: J. Benjamins.

Chierchia, Gennaro. 1998. "Reference to kind across languages". *Natural Language Semantics* 6.339–405.

Cohen, David. 1984. *La phrase nominale et l'évolution du système verbal en sémitique: Etudes de syntaxe historique.* Paris: Société de Linguistique de Paris.

Croft, William. 1991. *Syntactic categories and grammatical relations: The cognitive organization of information.* Chicago: University of Chicago Press.

Fassi Fehri, Abdelkader. 1993. *Issues in the structure of Arabic clauses and words.* Dordrecht: Kluwers.

Fleisch, Henri. 1961. *Traité de philologie arabe.* I. *Préliminaires, phonétique, morphologie nominale.* Beirut: Imprimerie Catholique.

Goldenberg, Gideon. 1995. "Attribution in Semitic languages". *Langues orientales anciennes, philologie et linguistique* 5–6.1–20. (Repr. *Studies in Semitic linguistics, selected writings*, by Gideon Goldenberg, 46–65. Jerusalem: Magnes Press, 1998.)

Guerssel, Mohand and Jean Lowenstamm. 1990. "The derivational morphology of the Classical Arabic verb". Ms. University of Quebec at Montreal/ University of Paris 7.

Hengeveld, Kees. 1992. *Non-verbal predication: Theory, typology, diachrony.* Berlin: Mouton de Gruyter.

Holes, Clive. 1995. *Modern Arabic: Structures, functions and varieties.* London: Longman.

Kihm, Alain. 2003. "Les pluriels internes de l'arabe: systèmes et conséquences pour l'architecture de la grammaire". *Recherches Linguistiques de Vincennes* 32.109–155.

——. 2004. "Verbal noun (*maṣdar*) formation in Classical Arabic: A root-and-site account". Ms. Centre National de la Recherche Scientifique (Laboratoire de linguistique formelle)/University of Paris 7.

——. 2005. "Noun class, gender, and the lexicon-syntax-morphology interfaces: A comparative study of Niger-Congo and Romance languages". *Handbook of comparative syntax*, ed. Guglielmo Cinque and Richard S. Kayne 459–512. Oxford: Oxford University Press.

Kwon, Song-Nim and Anne Zribi-Hertz. 2004. “Number from a syntactic perspective: Why plural marking looks ‘truer’ in French than in Korean”. *Empirical issues in formal syntax and semantics*, V, ed. Olivier Bonami and Patricia Cabredo-Hofherr, 133–158. Paris: Presses de la Sorbonne.

Link, Godehard. 1983. “The logical analysis of plural and mass nouns: A lattice theoretic approach”. *Meaning, use, and interpretation of language*, ed. Rainer Bäuerle, Christoph Schwarze, and Arnim von Stechow, 302–323. Berlin: de Gruyter.

Wechsler, Stephen and Larisa Zlatić. 2003. *The many faces of agreement*. Stanford: Center for the Study of Language Information.

Wright, William. 1991. *A grammar of the Arabic language*. 2 vols. Cambridge: Cambridge University Press. (1st ed., 1896–1898.)

Alain Kihm (CNRS – UMR 7110, Paris)

Adjectives

1. General definitions

Adjectives refer to a noun (substantive) which they qualify. They may take the position of an attribute, a predicate, or a predicative (circumstantial predicative). Usually attributes, predicates, and circumstantial predicatives come after the reference noun (Fassi Fehri 1999:107–112). Their reference to the noun is marked by agreement in gender and number and also, if they function as an attribute, in case and definiteness/indefiniteness. There are no special morphological signs to mark a noun as an adjective because, in principle, any adjective in Arabic is capable of functioning as a noun. Therefore, syntactic function of an attribute and morphological distinction of the two genders, masculine and feminine, are the only criteria to establish the adjective as a specific part of speech in Arabic. A small number of adjectives which quantify the reference noun form a special group insofar as they function partly as adjectives and partly as substantives.

2. Morphology

The adjective comprises the following morphological categories:

2.1 *Adjectives which agree with their reference noun*

2.1.1 Patterns which take the feminine ending *-at-u(n)* for agreement with a feminine noun, e.g. masc. *murr-un* ‘bitter’, fem. *murr-at-un*

i. Verbal adjectives of the patterns:
faʿl-un like *ṣaʿb-un* ‘difficult’ (from *ṣaʿuba* ‘to become difficult’), *sahl-un* ‘easy’ (from *sahula* ‘to become easy’);
fiʿl-un like *diqq-un* ‘fine, thin’ (from *daqqa* ‘to become thin’), *rixw-un* ‘loose’ (from *raxiya* ‘to become loose’);
fuʿl-un like *ḥulw-un* ‘sweet’ (from *ḥaluwa* ‘to become sweet’), *ṣulb-un* ‘hard, firm’ (from *ṣaluba* ‘to become hard, firm’);
faʿal-un like *ḥasan-un* ‘nice’ (from *ḥasuna* ‘to become nice’), *baṭal-un* ‘brave’ (from *baṭula* ‘to be brave’);
faʿil-un like *xašin-un* ‘rough’ (from *xašuna* ‘to become rough’), *fariḥ-un* ‘glad’ (from *fariḥa* ‘to become glad’);
faʿul-un like *yaqud-un* ‘awake’ (from *yaquda* ‘to be awake’); in Modern Standard Arabic most of the adjectives of this pattern turn into the pattern *faʿil-un*;
faʿīl-un like *qadīm-un* ‘ancient’, *jadīd-un* ‘new’, *ṭawīl-un* ‘long’, *qaṣīr-un* ‘short’, *kabīr-un* ‘big, large’, *ṣaġīr-un* ‘small’, *kaṯīr-un* ‘many’, *qalīl-un* ‘few’;
faʿūl-un like *ṭamūḥ-un* ‘ambitious’ (from *ṭamaḥa* ‘to aspire’), *ʿaṭūf-un* ‘compassionate’ (from *ʿaṭafa* ‘to feel compassion’);
fuʿlān-un like *ʿuryān-un* ‘naked’ (from *ʿariya* ‘to be naked’).

ii. Intensive adjectives of the patterns:
faʿʿāl-un like *bakkāʾ-un* ‘weepy’, *bassām-un* ‘smiling much’ (this pattern often indicates nouns of profession: *ʿaṭṭār-un* ‘druggist’, *ṣarrāf-un* ‘money-changer’);
fiʿʿīl-un like *ṣiddīq-un* ‘very truthful’, *sikkīt-un* ‘habitually silent’.

iii. Active and passive (→) participles are often used not only in their verbal function, but also as adjectives describing qualities, e.g.:
fāʿil-un like *kātib-un* ‘writing’, but also ‘able to write’ (from *kataba* ‘to write’); *bārid-un* ‘cold’ (from *barada* ‘to feel cold’), *wāḍiḥ-un* ‘clear’ (from *waḍuḥa* ‘to become clear’);
mufʿil-un like *muṣīb-un* ‘pertinent’ (from *ʾaṣāba* ‘to hit, to strike’), *muqmir-un* ‘moonlit’ (from *ʾaqmara* ‘to be shining [moon]’);
munfaʿil-un like *munfaʿil-un* ‘excited’ (from *infaʿala* ‘to become affected’);
muftaʿil-un like *muttaṣil-un* ‘continuous’ (from *ittaṣala* ‘to be connected’);
mafʿūl-un like *masʾūl-un* ‘responsible’ (from *saʾala* ‘to ask’), *maʿrūf-un* ‘well-known’ (from *ʿarafa* ‘to get to know’);

mufaʿʿal-un like *murabbaʿ-un* 'quadrangular' (from *rabbaʿa* 'to quadruple');
mufʿal-un like *muṭlaq-un* 'unlimited, absolute' (from *ʾaṭlaqa* 'to set free').

iv. Adjectives of relationship formed with the so-called *nisba*-ending *-iyy-un* indicate the belonging or relationship to something or somebody. In pre-Classical Arabic the *nisba*-ending is normally suffixed to names of tribes and places: *qays-iyy-un* 'belonging to the tribe of Qays', *ḥijāz-iyy-un* 'belonging to Hijaz', but *nisba*-adjectives derived from other nouns like *durr-iyy-un* 'being like pearls, brilliant' (Q. 24/35) (from *durr* 'pearls'), *ʾaʿjam-iyyun* 'belonging to the people who do not speak Arabic' (Q. 16/103) (from *ʾaʿjam-u* 'stammering, dumb') emerged very early; on peculiarities in pre-Classical Arabic see Fischer (1965:193–196, 383–417). In Classical Arabic the number of adjectives of relationship increased, because every noun could become the derivational base, e.g. *ʾinsān-iyy-un* 'human' (from *ʾinsān* 'human being'), *šams-iyy-un* 'belonging to the sun' (from *šams* 'sun'), *salb-iyy-un* 'negative' (from *salb* 'deprivation'). In Modern Standard Arabic the *nisba*-ending is extremely productive and may even replace the genitive, e.g. *aṯ-ṯawb-u n-nawm-iyy-u* or *ṯawb-u n-nawm-i* 'the nightdress'. The feminine ending *-at-* is usually dropped before adding *-iyy-*, e.g. *makk-iyy-un* 'belonging to Mecca (*Makk-at-u*)', *ḥukūm-iyy-un* 'belonging to the government' (from *ḥukūm-at-un* 'government') but is added to the endings *-ā* and *-āʾ*, with insertion of *w* between *-ā* and *-iyy*, e.g. *dunyaw-iyy-un* 'belonging to this world' (from *ad-dunyā* 'this world'), *ṣaḥrāw-iyy-un* 'belonging to the desert' (from *ṣaḥrāʾ-u* 'desert'). In the syllable before the *nisba*-ending the vowels *i*, *ī,* and *ay* change to *a*, e.g. *namar-iyy-un* 'belonging to the tribe of Namir', *nabaw-iyy-un* 'belonging to the Prophet (*nabīy*), *quraš-iyy-un* 'belonging to the tribe of Qurayš', but this phonetic rule is often neglected, e.g. *rudayn-iyy-un* 'a spear made by Rudayna', *ḥaqīq-iyy-un* 'real' (from *ḥaqīq-at-un* 'reality'). A *nisba*-ending *-ān-iyy-*, e.g. *fawq-ān-iyy-un* 'located above' (from *fawqu* 'above'), *ʿaql-ān-iyy-un* 'rational' (from *ʿaql* 'reason') appeared in post-Classical Arabic. For further details see Wright (1933:I, 149–165); El-Ayoubi a.o. (2002:137–143).

2.1.2. Patterns with supplementary feminine and plural forms; all of them follow the diptotic inflexion (→ diptosis):

i. *faʿlān-u*, fem. *faʿlā*, pl. com. *faʿālā* or *fuʿālā* like *sakrān-u*, fem. *sakrā*, pl. *sakārā, sukāra* 'drunk', *kaslān-u*, fem. *kaslā*, pl. *kasālā, kusālā* 'lazy'. Sometimes adjectives of this pattern shift to the triptotic inflection; in this case they take *-at-un* as feminine marker, e.g. *nadmān-u*, fem. *nadmā* or *nadmān-un*, fem. *nadmān-at-un* 'repentant'.

ii. *ʾafʿal-u*, fem. *faʿlāʾ-u*, pl. com. *fuʿl-un* indicates colors and striking qualities, e.g. *ʾaḥmar-u*, fem. *ḥamrāʾ-u*, pl. *ḥumr-un* 'red', *ʾabyaḍ-u*, fem. *bayḍāʾ-u*, pl. *bīḍ-un* (< **buyḍ-*) 'white', *ʾaʿraj-u*, fem. *ʿarjāʾ-u*, pl. *ʿurj-un* 'lame', *ʾaʿwaj-u*, fem. *ʿawjāʾ-u*, pl. *ʿūj-un* 'crooked'; Fischer (1965) deals with the adjectives of this pattern.

iii. *ʾafʿal-u*, fem. *fuʿlā*, pl. masc. *ʾafʿal-ūna* or *ʾafāʿil-u*, pl. fem. *fuʿlayāt-un* or *fuʿal-u* are (→) elatives expressing relative proportions, e.g. *al-ʾax-u l-ʾaṣġaru* 'the younger brother', *al-ʾuxt-u ṣ-ṣuġrā* 'the younger sister'; *ʾawwal-u* 'first' and *ʾāxar-u* 'another' belong to this class: sg. fem. *ʾūlā*, pl. masc. *ʾawwal-ūna*, *ʾawāʾil-u*, pl. fem. *ʾūlayāt-un*, *ʾuwal-u* and sg. fem. *ʾuxrā*, pl. masc. *ʾāxar-ūna*, *ʾawāxiru*, pl. fem. *ʾuxrayāt-un, ʾuxar-u.*

iv. adjectives formed by *ḏū*, fem. *ḏāt-u*, pl. masc. *ḏaw-ū* or *ʾul-ū*, pl. fem. *ḏawāt-u*, *ʾulāt-u* 'the one with . . ., endowed with . . .' followed by a noun in the genitive (on the inflection see Fischer 2002:§ 283, for further details El-Ayoubi a.o. 2002:143–146) like *rajul-un ḏū xams-īna ʿām-an* 'a man [one] of fifty years', i.e. 'a fifty-year-old man', *ad-dirās-at-u ḏāt-u l-bāl-i* 'the study endowed with attention' i.e. 'the notable study'.

2.2 *Patterns which do not agree with the reference noun*

faʿāl-un like *sawāʾ-un* 'equal', *jabān-un* 'cowardly', *ṣaḥāḥ-un* 'sound';
fiʿāl-un like *mirāḥ-un* 'lively', *liyāḥ-un* 'bright';
fuʿāl-un like *kurām-un* 'noble', *huḏām-un* 'keen'.

Most of the adjectives of these patterns are found only in old Arabic poetry, and no longer in Modern Standard Arabic.

ʾafʿal-u, the so-called (→) elative that describes something or somebody as pre-eminently characterized by some quality. It serves to form expressions corresponding in European languages to the comparative and superlative; it differs, however, in its syntactic structures from the adjective. In Modern Standard Arabic the elative is partly assimilated to the adjective (El-Ayoubi a.o. 2002:279– 282), but it does not form feminine and plural, with the exception of the elatives which have the supplementary feminine and plural forms mentioned above.

mifʿāl-un and *mifʿal-un* like *mixṣāb-un* 'very fertile', *miḏʿān-un* 'obedient'. They are originally nomina instrumenti, but used metaphorically as adjectives to mean 'doing something like a machine' (Wright 1933:I, 138).

The Arab lexicographers quote many adjectives of various other patterns (cf. Wright 1933:I. 137–140), but most of them cannot be verified in texts.

2.3 *Dual and plural*

All adjectives form the dual like the noun with nom. *-āni*, gen./acc. *-ayni*, and, in principle, they are able to form the plural externally with the markers masc. nom. *-ūna*, gen./acc. *-īna*, fem. nom. *-āt-un*, gen./acc. *-āt-in*. Many verbal adjectives have in addition internal plural patterns, which mostly function for the masc. pl., but sometimes for both genders. These patterns are:
faʿlā: *marḍā* to *marīḍ-un* 'ill', *qatlā* to *qatīl-un* 'killed';
faʿalat-un: maharat-un to *māhir-un* 'skilful';
fuʿul-un: judud-un to *jadīd-un* 'new', *ġuyur-un* to *ġayūr-un* 'jealous';
fiʿāl-un: kibār-un to *kabīr-un* 'large', *niyām-un* to *nāʾim-un* 'sleeping';
faʿālā or *fuʿālā: samājā*, *sumājā* to *samij-un* 'ugly';
fuʿūl-un: quʿūd-un to *qāʿid-un* 'sitting';
fawāʿil-u: kawāfir-u to *kāfir-un* 'unbelieving', *qalāʾil-u* to *qalīl-un* 'few';
fuʿalāʾ-u: fuḍalāʾ-u to *fāḍil-un* 'excellent', *fuqarāʾ-u* to *faqīr-un* 'poor';
fuʿʿal-un: sujjad-un to *sājid-un* 'prostrate in adoration';
fuʿʿāl-un: subbāq-un to *sābiq-un* 'former', *juhhāl-un* to *jāhil-un* 'ignorant';
ʾafʿāl-un: ʾaḥrār-un to *ḥurr-un* 'free'; *ʾamwāt-un* to *mayyit-un* 'dead';
ʾafʿilāʾ-u: ʾaġniyāʾ-u to *ġanīy-un* 'wealthy'.

3. Substantivization

Every adjective may be employed as a substantive; in this case, it includes the sense of 'person' or 'thing', e.g. *faqīr-un* 'poor' or 'poor man', *baṭal-un* 'brave' or 'hero', *jadīd-un* 'new' or 'anything new'. On the other hand, there are substantives which may be used for qualification of nouns, e.g. *ar-raʾy-u l-xaṭaʾ-u* 'the opinion, the error', i.e. 'the wrong opinion', *ʾarḍ-un qafr-un* 'land, desert', i.e. 'wasteland'. Therefore, one cannot be quite certain whether a noun was originally an adjective or a substantive; furthermore, in pre-Classical Arabic, many adjectives do not agree in gender with their reference noun and may well be classified as substantives. For this reason the Arab grammarians did not distinguish within the noun between substantive and adjective (cf. Diem 1971).

To create abstract nouns the so-called feminine ending *-at* is added to the adjective, e.g. *ḥasan-at-un* 'good deed', *xāṭiʾ-at-un* 'mistake', *faḍīl-at-un* 'virtue', *qawm-iyy-at-un* 'nationalism'.

Adjectives of estimating an action may be employed like the elative, i.e. they are used as substantives annexed to the noun they qualify. The annexion structure expresses a higher intensity than the attributive one does (El-Ayoubi a.o. 2002:157; Fassi Fehri 1999:115–117), e.g. *fāʿiqu t-taqdīr-i* 'the excellent of esteem', i.e. 'the most excellent esteem' vs. *at-taqdīr-u l-fāʾiq-u* 'the excellent esteem', *ṣādiq-u tamanniyāt-ī* 'the sincere of my wishes', i.e. 'my most sincere wishes' vs. *tamanniyāt-ī ṣ-ṣādiq-at-u* 'my sincere wishes'. In a similar way, the quantitative adjectives *kaṯīr* 'many, much', *ʿadīd* 'numerous', and *qalīl* 'few, little' may, in connection with the preposition *min* 'of', be employed as substantives, e.g. *kaṯīr-un min aṭ-ṭullāb-i* 'many of the students', *qalīl-un min aṭ-ṭullāb-i* 'a few of the students', as equivalents of *ṭullāb-un kaṯīr-ūna* 'many students', *ṭullāb-un qalīl-ūna* 'few students'.

4. Agreement

Adjectives take with reference to the noun they qualify the syntactic positions of attribute, predicate, or circumstantial predicative. The reference is marked by agreement in gender and number with the reference noun. In addition, when they are used as attributes they adopt the case and definite/indefinite state of the reference noun. As for the agreement in number, Arabic distinguishes between persons and non-persons; in agreement with the plural of persons the adjective takes the plural form, but in agreement with the plural of non-persons it is in the feminine singular.

Table 1. The adjective as attribute

	Masculine: *bayt-un jamīlun* ‘a beautiful house’ / *al-bayt-u l-jamīl-u* ‘the beautiful house’	
	singular	dual
nom.	*bayt-un jamīl-un /al-bayt-u l-jamīl-u*	*bayt-āni jamīl-āni /al-bayt-āni l-jamīl-āni*
gen.	*bayt-in jamīl-in /al-bayt-i l-jamīl-i*	*bayt-ayni jamīl-ayni /al-bayt-ayni l-jamīl-ayni*
acc.	*bayt-an jamīl-an /al-bayt-a l-jamīl-a*	*bayt-ayni jamīl-ayni /al-bayt-ayni l-jamīl-ayni*
	plural of non-persons	plural of persons: *rijāl-un jamīl-ūna* ‘beautiful men’
nom.	*buyūt-un jamīl-at-un /al-buyūt-u l-jamīl-at-u*	*rijāl-un jamīl-ūna /ar-rijāl-u l-jamīl-ūna*
gen.	*buyūt-in jamīl-at-in /al-buyūt-i l-jamīl-at-i*	*rijāl-in jamīl-īna /ar-rijāl-i l-jamīl-īna*
acc.	*buyūt-an jamīl-at-an /al-buyūt-a l-jamīl-at-a*	*rijāl-an jamīl-īna /ar-rijāl-a l-jamīl-īna*
	Feminine: *ṣūr-at-un jamīl-at-un* ‘a beautiful picture’ /*aṣ-ṣūr-at-u l-jamīlat-u* ‘the beautiful picture’	
	singular	dual
nom.	*ṣūr-at-un jamīl-at-un /aṣ-ṣūr-at-u l-jamīl-at-u*	*ṣūr-at-āni jamīl-at-āni /aṣ-ṣūr-at-āni l-jamīl-at-āni*
gen.	*ṣūr-at-in jamīl-at-in /aṣ-ṣūr-at-i l-jamīl-at-i*	*ṣūr-at-ayni jamīl-at-ayni /aṣ-ṣūr-at-ayni l-jamīl-at-ayni*
acc.	*ṣūr-at-un jamīl-at-an / ʾaṣ-ṣūr-at-a l-jamīl-at-a*	*ṣūr-at-ayni jamīl-at-ayni /aṣ-ṣūr-at-ayni l-jamīl-at-ayni*
	plural of non-persons	plural of persons: *ban-āt-un jamīl-āt-un* ‘beautiful girls’
nom.	*ṣuwar-un jamīl-at-un /aṣ-ṣuwar-u l-jamīl-at-u*	*ban-āt-un jamīl-āt-un /al-ban-āt-u l-jamīl-āt-u*
gen.	*ṣuwar-in jamīl-at-in /aṣ-ṣuwar-i l-jamīl-at-i*	*ban-āt-in jamīl-āt-in /al-ban-āt-i l-jamīl-āt-i*
acc.	*ṣuwar-an jamīl-at-an /aṣ-ṣuwar-a l-jamīl-at-a*	*ban-āt-in jamīl-āt-in /al-ban-āt-i l-jamīl-āt-i*

When the adjective is used as a predicate, it appears in the nominative and agrees with the subject in gender and number. The subject is normally definite, whereas the predicate is indefinite. Since Arabic has no verbal copula, indefiniteness of the predicate in contrast to definiteness of the subject marks the difference between the predicative and attributive nominal phrase: *aṣ-ṣūrat-u jamīlat-un* ‘The picture is beautiful’ vs. *aṣ-ṣūr-at-u l-jamīl-at-u* ‘the beautiful picture’ or *ṣūr-at-un jamīl-at-un* ‘a beautiful picture’. Predicative function neutralizes the opposition between substantive and adjective, e.g. *al-bayt-u xarāb-un* ‘The house is ruined’ or ‘the house is a ruin’. When the nominal phrase is verbalized by *kāna* ‘to be’ or another modifying verb, the predicate changes into the accusative: *kānat-i ṣ-ṣūr-at-u jamīl-at-an* ‘The picture was beautiful’.

When the adjective is used beside the verbal predicate as a second predicate usually called circumstantial predicative (→ *ḥāl*), it is in the accusative and indefinite, but agrees with its reference noun or pronoun in gender and number, in accordance with the rules given for the predicate, e.g. *xarajat-i l-bintu bākiy-at-an* ‘The girl went out [and was while doing so] weeping’, *daxalnā l-ġurf-at-a fariḥ-īna* ‘We went into the room [and were while doing so] cheerful’. After verbs of perception and others the object may be the reference noun, e.g. *raʾaytu n-nisāʾa ḥāʾir-āt-in* ‘I saw the women [and they were at the same time] confused’.

Adjectives which denote qualities specific for females are excluded from the gender agreement and do not take the feminine marker *-at-*, e.g. *imraʾ-at-un ḥāmil-un* ‘a pregnant woman’. Adjectives of the patterns *faʿīl-un* and *faʿūl-un* do not always show gender agreement. The Arab grammarians give as a basic rule that *faʿīl-un* with an active sense and *faʿūl-un* with a passive sense do not agree, e.g. *qanāt-un rafīd-un* ‘broken lance’, *ruʾyā kaḏūb-un* ‘false vision’. However, in the course of time, usage deviated more and more from this rule.

The rules of gender and number agreement given above apply without reservation only to Modern Standard Arabic and to a certain extent to Classical Arabic as well. In the pre-Classical language, however, the agreement in gender is more limited; only participles, adjectives of relationship, and adjectives with a supplementary feminine show full agreement, while many patterns of verbal adjectives are absolutely or partially not capable of agreement (Kahle 1975:77–78), e.g. *bi-maḥall-at-in šaks-in* ‘on an inaccessible place’, *hiya rixw-un* ‘she is loose’ (Kahle 1975:89, 97). With regard to the number agreement, there is no distinction between persons and non persons in pre-Classical Arabic: most adjectives show full agreement with nouns in the plural (Kahle 1975:111–134), e.g. *buyūt-un wāḍiʿ-āt-un* ‘shabby cottages’, *ar-rimāḥ-u ṭ-ṭiwāl-u* ‘the long lances’, *al-qibāb-u l-ḥumr-u* ‘the red tents’. On the other hand, adjectives of quantity remain in the masculine singular (Kahle

Table 2. The adjective as predicate

sg. masc.	*al-bayt-u jamīl-un*	‘The house is beautiful’
	ʾinna l-bayt-a jamīl-un	‘Verily, the house is beautiful’
	kāna l-baytu jamīl-an	‘The house was beautiful’
sg. fem.	*aṣ-ṣūr-at-u jamīl-at-un*	‘The picture is beautiful’
	ʾinna ṣ-ṣūr-at-a jamīl-at-un	‘Verily, the picture is beautiful’
	kānat-i ṣ-ṣūr-at-u jamīl-at-an	‘The picture was beautiful’
du. masc.	*al-bayt-āni jamīl-āni*	‘The two houses are beautiful’
	ʾinna l-bayt-ayni jamīl-āni	‘Verily, the two houses are beautiful’
	kāna l-bayt-āni jamīl-ayni	‘The two houses were beautiful’
du. fem.	*aṣ-ṣūr-at-āni jamīl-at-āni*	‘The two pictures are beautiful’
	ʾinna ṣ-ṣūr-at-ayni jamīl-at-āni	‘Verily, the two pictures are beautiful’
	kānat-i ṣ-ṣūr-at-āni jamīl-at-ayni	‘The two pictures were beautiful’
pl. of non-persons	*al-buyūt-u jamīl-at-un*	‘The houses are beautiful’
	ʾinna l-buyūt-a jamīl-at-un	‘Verily, the houses are beautiful’
	kānat-i l-buyūt-u jamīl-at-an	‘The houses were beautiful’
pl. of persons	*al-ban-āt-u jamīl-āt-un*	‘The girls are beautiful’
	ʾinna l-ban-āt-i jamīl-āt-un	‘Verily, the girls are beautiful’
	kānat-i l-ban-āt-u jamīl-āt-in	‘The girls were beautiful’

1975:117), e.g. *rijāl-un kaṯīr-un* ‘many men’. The language of the *Qurʾān* follows partly the pre-Classical, partly the Classical rules of agreement; e.g. for gender agreement *bald-at-an mayyit-an* ‘an inanimate place’ (Q. 50/11), but *al-ʾarḍ-u* [fem.] *l-mayyit-at-u* ‘the inanimate earth’ (Q. 36/33), and for number agreement *judad-un bīḍ-un wa-ḥumr-un* ‘white and red stripes’ (Q. 35/27), *ʾayyām-an maʿdūd-āt-in* ‘for some countable days’ (Q. 2/184), but *ʾayyām-an maʿdūd-at-an* ‘for some countable days’ (Q. 2/80). In some phrases and with the adjectives of color the old usage is preserved in Modern Standard Arabic, e.g. *saʿāt-un ṭiwāl-un* ‘long lasting hours’, *šaʿar-āt-un bīḍ-un* ‘white hairs’.

5. Extended adjectives

Besides the extension of (→) participles and verbal adjectives with direct or indirect objects and other complements like *hāḏā l-ʿamal-u ṣaʿb-un ʿalay-nā* ‘this work is hard for us’, *al-qary-at-u baʿīd-at-un min al-ʿāṣim-at-i* ‘the village is far away from the capital’, there are two kinds of extension specific to adjectives, the annexion of a limiting genitive and the adjectival clause.

The adjective may annex a limiting term in the genitive, which is always marked as definite, but does not make the adjective phrase definite, so that it takes the definite article for agreement with a definite reference noun, e.g. *imraʾ-at-un bayḍāʾ-u š-šaʿr-i* ‘a woman white of [the] hair’, i.e. ‘a white-haired woman’, *al-marʾ-at-u l-bayḍāʾ-u š-šaʿr-i* ‘the woman white of [the] hair’, i.e. ‘the white-haired woman’.

The (→) adjective clause has its own subject which contains a pronoun referring to the reference noun of the adjective, e.g. *imraʾ-at-un ʾabyaḍ-u šaʿr-u-hā* ‘a woman white [is] her hair’, i.e. ‘a white-haired woman’, *rajul-un bayḍāʾ-u liḥy-at-u-hu* ‘the man white [is] his beard’. i.e. ‘a white-bearded man’. The adjectival clause has a two-fold reference: on the one hand it agrees in gender with its subject, on the other hand in case and definiteness/indefiniteness with its reference noun. It is used like other adjectives as an attribute, predicate, or circumstantial predicative, e.g. *raʾaytu l-marʾ-at-a l-ʾabyaḍ-a šaʿr-u-hā* ‘I saw the white-haired woman’, *ar-rajul-u bayḍāʾ-u liḥy-at-u-hu* ‘the man is white-bearded’, *raʾaytu r-rajul-a bayḍāʾ-a liḥy-at-u-hu* ‘I saw the man [and he was at the same time] white-bearded’. The adjectival clause is a transformation of a qualifying relative clause (*šaʿr-u-hā ʾabyaḍ-u* ‘her hair is white’, *liḥy-at-u-hu bayḍāʾ-u* ‘his beard is white’), its predicate is put before the subject (i.e. *ʾabyaḍ-u šaʿr-u-hā*, *bayḍāʾ-u liḥy-at-u-hu*) and then attached by agreement to the noun it qualifies (cf. Diem 1998).

Bibliographical references

El-Ayoubi, Hashem, Wolfdietrich Fischer and Michael Langer. 2001. *Syntax der Arabischen Schriftsprache der Gegenwart*. I/1. *Das Nomen und sein Umfeld*. Wiesbaden: L. Reichert.

Diem, Werner. 1971. “Nomen, Substantiv und Adjektiv bei den arabischen Grammatikern”. *Oriens* 23–24, 312–332.

——. 1998. fa-waylun li-l-qāsiyati qulūbuhum*: Studien zum arabischen adjektivischen Satz*. Wiesbaden: O. Harrassowitz.

Fassi Fehri, Abdelkader. 1999. “Arabic modifying adjectives and DP structures”. *Studia Linguistica* 53:105–154.

Fischer, Wolfdietrich. 1965. *Farb- und Formbezeichnungen in der Sprache der altarabischen Dichtung.* Wiesbaden: O. Harrassowitz.
——. 2002. *A grammar of Classical Arabic*, trans. Jonathan Rodgers. New Haven and London: Yale University Press.
Kahle, Erhard. 1975. *Studien zur Syntax des Adjektivs im vorklassischen Arabisch*. Erlangen: H. Lüling.
Wright, William. 1933. *A grammar of the Arabic language*, ed. W. Robertson Smith and M.J. de Goeje. 2 vols. 3rd rev. Cambridge: Cambridge University Press.

Wolfdietrich Fischer (Erlangen, Germany)

Adjunction → X-bar-Syntax

Adverbial Clause → Mafʿūl fīhi; X-bar Syntax

Adverbs

Classical Arabic has few words that function solely as adverbs. More often, a word with a basic nominal or adjectival function may be used as an adverbial in certain syntactic contexts. Prepositional phrases typically function as adverbials.

The accusative is the fundamental marker of adverbiality in Classical Arabic. (The few exceptions to this rule will be dealt with below in sections 1.1 and 1.2.) This general pattern is most obviously apparent in forms such as *dāʾim-an* ‘always’ derived from the adjective *dāʾim* ‘lasting’ or *dawām-an* ‘permanently’ derived from the noun *dawām* ‘permanence’; however, it also manifests itself in genitive phrases such as *ḥīna wuṣūlihi* ‘at the time of his arrival’, where the Western grammatical tradition would incline the reader to regard *ḥīn-a* as a preposition. Even in cases where Arabic has true → prepositions (i.e. words for which there is no corresponding noun or adjective, such as *fī* ‘in’), the prepositional phrase can be regarded as having an internal genitive structure, hence the genitive ending on the noun modifier, the whole phrase functioning as an adverbial. In certain syntactic contexts, the distinction between adverbial and object is blurred: for example, someone with a Western grammatical background would be inclined to analyze *ramaḍān-a* in *ṣāma ramaḍān-a* ‘he fasted Ramaḍān’ as an adverbial; however, the fact that this sentence is passivizable as *ṣīma ramaḍān-u* ‘Ramaḍān was fasted’, suggests that *ramaḍān-a* in *ṣīma ramaḍān-a* is to be indeterminately analyzed as both object and adverbial. This entry does not investigate the syntactic structure of prepositional phrases or details of the relationship between object and adverbial. Rather, the focus is on semantic categories of adverbs in Classical Arabic under the headings of temporal, local, direction, degree, manner, and interrogative adverbials. Within each section adverb types are considered in the following categories, where applicable: words which function solely as adverbs; words which function mainly, but not solely, as adverbs; words which function both as adverbs and as nominals; and words which function to varying degrees as adverbs and adjectives. In the discussion of adverbs in modern Arabic dialects, the principal deviations from Classical Arabic adverbs, loan adverbs, and innovative forms are examined.

1. Adverbs in Classical Arabic

1.1 *Temporal adverbs*

In Classical Arabic, the temporal adverbs include nouns which have an almost entirely adverbial function, such as *ʾamsi* ‘yesterday’ (cf. *ʾamsu* ‘the day before’), *ġad-an* ‘tomorrow’, *al-ʾāna* ‘now’, and nouns, noun phrases, and genitive phrases in the accusative which function mainly in an adverbial sense, such as: *taww-an* ‘immediately’; *marrat-an* ‘once’; *marrat-an ʾuxrā* ‘once more’; *mubāšarat-an* ‘straight away, immediately’; *ʾaḥyān-an* ‘sometimes’; *ʾabad-an* + negation ‘never’; *baʿḍ-a l-ʾaḥyāni* ‘sometimes’; and *ʿādatan* ‘usually, habitually’.

Terms used wholly adverbially in Classical Arabic include temporal circumstants when they take the archaic ending *-u*: *qabl-u* ‘previously, formerly’ and *baʿd-u* ‘then, afterwards, later, still, yet’; these may also follow the preposition *min* ‘from’ without changing either the meaning or the ending *-u*. These circumstants may also take as suffix the temporal particle *ʾiḏ(in)*, maintaining their adverbial function though taking the accusative case ending, viz. *qablaʾiḏin*, *baʿdaʾiḏin*. This particle may be suffixed to other temporal terms, as in *yawmaʾiḏin* ‘on that day’, *ʿāmaʾiḏin* ‘in that year’, *waqtaʾiḏin* ‘then, at that time, by then’ and *waqtaḏāka* ‘then, at that time, then’, used independently, as in *ʾiḏ* and *ʾiḏā* ‘at that moment, then’ and *ʾiḏan* ‘hence, then’, or prefixed to *ḏāka*, as in *ʾiḏ-ḏāka* ‘at that time’ (cf. Fischer 1997:201).

Terms which are almost as likely to function adverbially as nominally (in the latter case with case endings as appropriate to syntactic status) include *ḥāl-an* ‘straight away’, *waqt-an* ‘at once, one day’, *yawm-an* ‘one day’, *al-yawm-a* ‘today’, and demonstrative-noun phrases such as *hāḏā l-ʾusbūʿ-a* ‘this week’, *hāḏā š-šahr-a* ‘this month’, *hāḏihi s-sanat-a* ‘this year’.

A number of words whose stem is on the typical adjectival pattern *faʿīl* or *fāʿil* function as temporal adverbs in adverbial contexts, taking the accusative ending *-an*. These include: *dāʾim-an* ‘always’, *ṭawīl-an* ‘a long time’, *nādir-an* ‘rarely’, *kaṯīr-an* ‘a lot, often’, *qadīm-an* ‘in the past’, *qalīl-an* ‘rarely’. Of these, *kaṯīr-an, qalīl-an*, and *nādir-an* may also take following *mā* while maintaining the sense of a temporal adverb.

1.2 *Local adverbs*

The locative demonstratives are pure adverbs. Distance is denoted by the suffixal element *-ka*, as in the pronominal demonstratives. The local adverbs include *hunā* ‘here’ and *hunāka/hunālika*, both of which can also occur with prefixed *hā-*, as *hāhunā* and *hāhunāka*. The demonstrative phrases *ṯamma, ṯammata* and *min ṯamma* ‘there’ also have a purely adverbial function. Words on nominal patterns which frequently function as local adverbs include *barr-an* ‘out, outside’. Words on adjectival patterns which typically, but not exclusively, function as local adverbs include *baʿīd-an* ‘far away’ and *qarīb-an* ‘near’. Local circumstants which take the archaic ending *-u* (cf. 1.1) are also used adverbially, as in: *taḥt-u* ‘underneath’, *fawq-u* ‘up, upstairs, on top, above’, *xalf-u* and *min xalf-u* ‘behind’, *ʿal-u* and *min ʿal-u* ‘above’.

1.3 *Direction*

The majority of nominal forms used as adverbs of direction may also function predicatively. These include *yasār-an* ‘left’, *yamīn-an* ‘right’, *ġarb-an* ‘west’, *šarq-an* ‘north’, *janūb-an* ‘south’, *šimāl-an/šamāl-an* ‘north, left’. Direction adverbs also include a number of words built on adjectival patterns which are typically, but not exclusively, used adverbially. These include *jānib-an* ‘to the side’, *dāxil-an* ‘inside’, and *xārij-an* ‘outside’.

1.4 *Degree*

The principal degree adverbs in Classical Arabic are *jidd-an* ‘very’, and words on adjectival patterns *kaṯīr-an* ‘a lot, much’ and *qalīl-an* ‘a little, somewhat’, all of which are mainly, but not exclusively, used in an adverbial sense.

1.5 *Manner and modal*

The manner demonstratives all involve the original preposition **ka-* ‘like, as’. As for the local demonstratives (cf. above), distance is denoted by suffixal *-ka*. Examples include: *ka-ḏā* ‘so, thus, that way’, *ka-ḏāli-ka* ‘so, like this, thus, equally, likewise’. *Ka-ḏā* can be preceded by the demonstrative element *hā-*, as in: *hākaḏā* ‘so, thus’.

The word *ʾayḍ-an* ‘also’, from the root *ʾ-y-ḍ* ‘to return’, only functions adverbially. Words on typical nominal patterns which frequently function as manner or modal adverbs include *sawiyyat-an* ‘together, jointly’, *mahl-an* ‘slowly, leisurely’, *karh-an/kurh-an* ‘unwillingly, under duress’, *fawr-an* ‘immediately, directly’, *maṯal-an* ‘for example’, *ṭabʿ-an* ‘of course’. Adjectives which may function as manner adverbs include *sarīʿ-an* ‘quickly’, *baṭīʾ-an* ‘slowly’, *sawiyy-an* ‘in common, jointly’, *qalīl-an qalīl-an* ‘by and by, slowly, gradually’, and *jayyid-an* ‘well, excellently, thoroughly’.

1.6 *Interrogative adverbials*

The interrogative adverbials include *ʾayna* ‘where?’, *ʾannā* ‘where . . . from?’, *li-māḏā* ‘why?’, *matā* ‘when?’, *ʾayyāna* (from **ʾayya-ʾāna*) ‘what time?’, *kam* ‘how many; how much?’, and *kayfa* (derived, according to Fischer 1997:202, from **ka-ʾayyin fa-*) ‘how?’

2. Adverbs in the Arabic dialects

In contrast to Classical Arabic, the modern Arabic dialects have a large range and number of pure adverbs. On the whole, these have developed from Arabic nouns, noun phrases, adjectives, or prepositional phrases which function adverbially in Classical Arabic in certain contexts. In some cases, adverbs have been derived from a noun plus a suffix; in others, emphatic suffixes have led to further sets of demonstrative adverbs. Adverbs in modern Arabic dialects will be examined in terms of demonstrative adverbials, the treatment of the Arabic indefinite accusative ending *-an*, the innovative development of adverbs though grammaticalization of content or function words, adverbs formed by suffixation of a (probably substrate) suffix, loan adverbs, and interrogative adverbials.

2.1 *Demonstrative adverbials*

In the dialects final short vowels are either lost or lengthened, and in most dialects final long vowels are reduced to short vowels – thus, Classical Arabic *hunā* 'here' is realized in Cairene as *hina*, and *hunāka* 'there' as *hināk*. In a number of different dialect areas, the local demonstratives are derived ultimately from *hāhunā* and *hāhunāka* with loss of the intervocalic /h/ and accompanying sound changes, as in Ṣanʿānī *hānā* 'here' and *hānāk* 'there', Tunisian *hūnī* 'here', Lebanese *hawn* and *hawne* (Fischer 1969:119–20), Mardin *hawne* 'here' and *hawnak/e* 'there' (Jastrow 1978). In Tunis and Rabat, alongside *hnāk* 'there' are forms which can be traced back to Classical *ṯamma*, viz: *ṯemma* or *femma* in Tunis, and *temma* or *temmāk* in Rabat. Various dialects have developed a third distance level in the local demonstratives with the sense of 'over there'. These are derived respectively from *ġādī* and *li-hunā*: Tunisia and Cherchell *ġādi*; Palmyra *ġād*; Algeria and Morocco *lihen*, *lhī* (Fischer and Jastrow 1980; Grand'Henry 1972 for Cherchell).

In all dialect regions, the manner demonstratives go back to *kaḏā* and *hākaḏā* (Fischer and Jastrow 1980:83), with the dialects of the Peninsula most closely resembling the mother forms: Ṣanʿānī *hākaḏā, kaḏā, kaḏayya* 'like this'; Adeni *hākida* (Fischer 1969:135); in a number of dialects outside the Peninsula, including Palestinian rural dialects and some Mesopotamian dialects, the long vowel is raised through → *ʾimāla*, as in: Mardin *hēkiḏ* (Fischer 1969:135); Jewish Baghdadi *hēkid*; Muslim Baghdadi, Kwayriš *hīči*, *hīč* (Fischer and Jastrow 1980:151); Palestine *hēkiḏ* (Fischer 1969:135). In North Africa, the middle vowel usually disappears, as in Sfax *häkḏa* and Algiers *hākda* (Fischer 1969:135). Dialects of the Nile Valley and eastern Sudan only construct manner demonstratives without the *hā-* prefix (Fischer 1969:132).

A number of dialects have two sets of adverbial demonstratives, with the second set taking emphatic endings (Fischer 1969:98–9): e.g. Ṣanʿānī *-ayyih*, *hinayyih* 'here', *hinayyik* 'there', *kaḏayyih* 'like that, this'; Manāxa, Yemen *-eyya* (Werbeck 2001), *hineyya* 'here', *hineyyik* 'there', *kaḏeyya* 'in the same way, likewise'; Muslim Baghdadi *hnāya* 'here'; and Tangiers *hnāya* (Werbeck 2001).

There are a number of adverbs in modern Arabic dialects that take the accusative *-an* ending of Classical/Modern Ṣtandard Arabic. These are particularly frequent in higher registers of speech, often where more obviously colloquial alternatives exist: *mubāšaratan* 'immediately' (as against Ṣanʿānī *bisāʿathā*), *dāʾiman/dāyman* 'always' (as against Khartoum *dīma*), *ʾaḥyānan* 'sometimes' (as against Ṣanʿānī *zārаṯḥīn*), *ʿādatan* 'usually', *maṯalan* 'for example', *jiddan* 'very' (as against Khartoum *jadd*), *ʾasāsan* 'basically'.

2.2 *Adverbs derived through grammaticalization*

While Classical Arabic has only a few pure adverbs and these occurring in a small number of semantic classes (temporal, local, and manner), → grammaticalization and concomitant → semantic bleaching of content words in the neo-Arabic dialects has led to the development of a large range of adverbs. Words expressing the adverbial concept 'now' in the different dialects have developed from the temporal nouns **sāʿat-un* 'hour', **waqt-un* 'time', **ḥīn-un* 'time', from a grammaticalization of **ʾiḏā bi-* (Fischer 1969), as well as from the adverb, **taww-an*. Illustrative examples are given in Table 1 (adapted from Durand 1995:96; Algiers data checked with Aziza Boucherit, p.c.).

Other temporal adverbs are derived from grammaticalization of prepositional phrases. These include Ṣanʿānī *baḥīn* 'early', from the prepositional phrase **bi-ḥīn-in*, which has since developed a comparative adverb, *abḥan*

Table 1. 'Now' in Arabic dialects

Dialect	Dialect form	Classical etymology
Baghdad	*hassa*	**hāḏihi s-sāʿa*
Khartoum	*hassi/hassaʿ*	**hāḏihi s-sāʿa*
Damascus	*hallaʾ*	**hāḏā l-waqt*
Jerusalem	*halʾēt*	**hā-l-wuqayt*
Ṣanʿāʾ	*ḏalḥīn*	**hāḏā l-ḥīn*
Cairo	*dilwaʾti*	**hāḏā l-waqt*
Algiers	*drūk/derwaq*	**hāḏā l-waqt*
Rabat	*dāba*	**ʾiḏā bi-*
Tunis	*tawwa*	**taww-an*

'earlier', and a verb, _baḥḥan_ 'to be early'; _fīsāʿ_ 'quickly'; _bi-sāʿat_ + pronoun 'immediately', probably related to Classical Arabic _min sāʿatihi_ 'the same hour, immediately'; and Sudanese _lissa/lissaʿ_, Egyptian _lissa_ 'not yet, just now', derived from *_li-s-sāʿat-i_. Procházka (2000) also suggests that Egyptian _barḍu_ 'also' results from the grammaticalization of the prepositional phrase, *_bi-ʾarḍihi_, rather than, as commonly thought, from the Turkish word _bir de_. Various grammaticalized combinations involving *_ka_- or *_kayf_ and *_ġayr_, the presentative _rāh_ or _bāqe_ in North Africa, convey simultaneity or immediate precedence: the sense of 'at the moment that, as soon as' is expressed in Fes as _ki_, Djidjelli as _kīma_, Rabat as _ġer kīf_ (Taine-Cheikh 2004: 323); that of 'just' as _kīv änn_ (+ pronoun suffix) in Ḥassāniyya, _ki_ in northern Tunisia, _bāqe kī_ in Morocco, and _ġīr kī_ or _rāh kī_ in Tlemcen (Taine Cheikh 2004:324).

Adverbs in other semantic fields result from grammaticalization of nouns or adjectives. Thus, the diminutive noun _šuwayya_ 'small thing' in most non-peripheral dialects has now developed the adverbial sense 'a little'; in Cairene, _aḥsan_ has the adverbial sense of 'rather; better' in some contexts, as in _ikkallimu f-ḥāga tanya aḥsan_ 'rather/better talk about something else' (Woidich 1995); Cairene _ʾawi_, Yemeni _gawī/qawī_ (*_qawīy_ 'strong') has the sense of 'very' following an adjective, 'much, a lot' following a verb; in the Omani dialect of Khābūra, _yōm_ (*_yawm_ 'day') means 'when' and _il-ʿām_ (*_al-ʿām_ 'the year') 'last year' in adverbial contexts (Brockett 1985:225, 164); Khartoum _gawām_, Damascene _ʾawām_ (*_qawām_ 'support') has the adverbial sense of 'immediately'.

2.3 _Adverbs with suffixal elements_

A number of dialects have adverbs which involve suffixal _-n_ elements, most notably the now almost pan-Arabic _baʿdēn_ 'then, afterwards' and _kamān_ 'also'. Procházka (2000) analyses _-ēn_ in _baʿdēn_ as a demonstrative element, possibly influenced by the Aramaic substrate in Syrian and Lebanese dialects where it originated, with the word then spreading to other dialect areas which enjoyed contact with Syrian/Lebanese speakers. Further support for this theory is seen in _mitēn_ 'when' and _waktēn_ 'when' in Sudanese dialects, which can also be analyzed as [time word] + _-ēn_ (J. Dickins, p.c.). In some dialects in Upper Egypt and Sudan, _baʿdēn_ can be further extended through the demonstrative suffixes _-ak/-ik_, as in Egyptian _baʿdēnak_ and Šukriyya _baʿdēnlik_ (Procházka 2000). In various dialects, adverbs take suffixes with a resulting emphasis in meaning, as in Šukriyya _hassaʿ/hassiʿ_ 'now' vs. _hassaʿtiyya/ hassaʿtīn_ 'right now' (Reichmuth 1983), and _essaʿ_ 'now' vs. _essaʿni_ 'now, right now' and _hássaʿni_ 'just now' in Mardin (Procházka 2000:100).

A few adverbs are formed through suffixation of a demonstrative in some dialects in which the demonstrative pronoun follows the noun, as in Cairene _innaharda_ 'today', or Khartoum _al-lēladi_ 'today'. In Khartoum, a demonstrative _da_ or _di_ following _hina_ 'here' and _hassi_ and _hassaʿ_ 'now' emphasizes immediacy, as in _hina da_ 'right here', _hassi di_ and _hassaʿ da_ 'right now'.

2.4 _Loan adverbs_

A few dialects have adverbs borrowed from superstrate languages or derived from substrate languages. These include pan-Arabic _bass_ 'only' (Persian _bas_), Cairene _duġri_ 'immediately; straight on' (Turkish _doğru_), Khartoum _šārb_ 'sharp' (English _sharp_), as in _taji as-sāʿa t-talāta šārb_ 'come at three o'clock sharp!' (J. Dickins, p.c.), _aččax/aščax_ 'when' (*_ayš_ + Kurdish _çax_ 'time') in some Mardin dialects (Jastrow 1978), Jewish Baghdadi _gárag_ 'probably' (Turkish _gerek_) (Mansour 1991), and Khābūra _hest_ 'very' (Persian _hest_) and _fūl_ 'at full throttle' (English _full_) (Brockett 1985).

2.5 _Interrogative adverbials_

The interrogative adverbials are derived in varying degrees through grammaticalization of other parts of speech and merging of Classical Arabic interrogatives with prepositions. In the peninsular dialects, some of the interrogatives are little changed from the Classical Arabic originals. Terms for 'when?' are mainly reflexes of either *_ʾayya ḥīn_ 'which time?', *_ʾayya matā_ 'which when?', *_ʾayya waqt_ 'which time?', or *_matā_ (when?). 'Why?' is generally derived from *_li-māḏā_ 'why?' or *_li-ʾayyi šayʾin_ 'for which thing?', in some dialects from *_ʿalā_ + _mā/ʾayyi šayʾin_. 'How?' is derived from *_kayfa_ 'how?', *_kayfa_ + _ʾayy šayʾin_ 'which thing?', *_ʾayya lawnin_ 'which type?', *_ʾayya_ + _šayʾin_ + _lawn_ 'which color?', or *_ʾayya ziyyin_ 'which guise?' 'Where?' is almost invariably derived directly

Table 2. Interrogative pronouns in Arabic dialects

	*matā	_*ʾayna_	_*li-māḏā_	_*kayf_	_*kam_	how much?
Ṣanʿānī	_ʾayyaḥīn_	_ʾayn_	_lilmā_	_kayf_	_kam_	
Cairene	_ʾimta_	_fēn_	_lēh_	_ʾizzāy_	_kām_	_ʾaddi ʾē_
Damascus	_ʾēmta_	_wēn/fēn_	_lēš_	_kīf/šlōn_	_kamm_	_ʾaddēš_
Muslim Baghdad	_yemte/(i)šwakit_	_wayn_	_layš/luwayš_	_šlōn_	_bayš/šged_	_čem/škem/šged_
Mardin	_aymat(e)_	_ayn_	_layš_	_ʾašwan_		
Cherchell, Algeria	_ḏīwqāš_	_fāyen_	_ʿalāš/lāš_	_kifāš/kīš_	_šḥāl_	
Khartoum	_mitēn_	_wēn_	_lē šnu/lēh_	_kēf_	_kam_	

from _*ʾayna_ 'where?' with prefixation of _*fa-_ 'then' or _*wa-_ 'and' in some dialects; some Sudanese dialects, however, have _šiggēš_ (Hillelson 1930). 'How many, much?' is derived either from _*kam_ 'how many/much?' or from a prepositional phrase, _*bi-ʾayyi šayʾin_ 'of which thing?', or a genitive phrase, _*qadda ʾayyi šayʾin_ 'the size/quantity of which thing?'. Some dialects, including Damascene and Muslim Baghdadi, distinguish lexically the notions 'how much?' and 'how many?'.

Sound changes in the different dialects, including vowel deletion or reduction, palatalization, and monophthongization, often heavily disguise the etymological origin of the interrogative. Illustrative examples of reflexes of the interrogative adverbials _*matā_ 'when?', _*ʾayna_ 'where?', _*li-māḏā_ 'why?', _*kayfa_ 'how?' and _*kam_ 'how many, much?' from seven dialects are given in Table 2. Where the dialect in question makes a lexical distinction between the notions 'how much?' and 'how many?', the term for 'how much?' is given in the far right-hand column and the form/s in the _*kam_ column is/are the 'how many?' form/s. Illustrative data for Ṣanʿānī is taken from Watson (1993), for Cairene from Woidich (2002), for Damascus from Cowell (1964), for Muslim Baghdadi from McCarthy and Raffouli (1964), for Mardin from Jastrow (1978), for Cherchell from Grand'Henry (1972), and for Khartoum from J. Dickins (p.c.).

BIBLIOGRAPHICAL REFERENCES

Brockett, Adrian A. 1985. _The spoken Arabic of Khābūra on the Bāṭina of Oman._ Manchester: Journal of Semitic Studies, University of Manchester.

Cowell, Mark W. 1964. _A reference grammar of Syrian Arabic (based on the dialect of Damascus)._ Washington, D.C.: Georgetown University Press.

Durand, Olivier. 1995. _Introduzione ai dialetti arabi._ Milan: Centro Studi Camito-Semitici.

Fischer, Wolfdietrich. 1969. _Die demonstrativen Bildungen in den neuarabischen Dialekten._ The Hague: Mouton.

——. 1997. "Classical Arabic". _The Semitic languages_, ed. Robert Hetzron, 187–219. London: Routledge.

—— and Otto Jastrow (eds.). 1980. _Handbuch der arabischen Dialekte._ Wiesbaden: O. Harrassowitz.

Grand'Henry, Jacques. 1972. _Le parler arabe de Cherchell, Algérie._ Louvain-la-Neuve: Université Catholique de Louvain, Institut Orientaliste.

Hillelson, Ṣ. 1930. _Sudan Arabic: English-Arabic vocabulary._ London: Sudan Government.

Jastrow, Otto. 1978. _Die mesopotamisch-arabischen_ qəltu-_Dialekte._ I. _Phonologie und Morphologie._ Wiesbaden: F. Steiner.

Mansour, Jacob. 1991. _The Jewish Baghdadi dialect: studies and texts in the Judaeo-Arabic dialect of Baghdad._ Or-Yehuda: The Babylonian Jewry Heritage Center.

McCarthy, R.J. and Faraj Raffouli. 1964. _Spoken Arabic of Baghdad._ 2 vols. Beirut: Librairie Orientale.

Procházka, Stefan. 2000. "Über einige rätselhalfte Formen auf _-(ē)n_ in den arabischen Dialekten". _Oriente Moderno_ 19.99–109.

Reichmuth, Stefan. 1983. _Der arabische Dialekt der Šukriyya im Ostsudan._ Hildesheim: G. Olms.

Taine-Cheikh, Catherine. 2004. "De la grammaticalisation de 'comme' (comparatif) en arabe". _Approaches to Arabic dialects: A collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday_, ed. Martine Haak, Rudolf de Jong, and Kees Versteegh, 309–328. Leiden: E.J. Brill.

Watson, Janet C.E. 1993. _A syntax of Ṣanʿānī Arabic._ Wiesbaden: O. Harrassowitz.

Werbeck, Wolfgang. 2001. _Laut und Formenlehre des nord-jemenitischen Dialekts von Manāxa._ Münster: Rhema.

Woidich, Manfred. 1996. "Some cases of grammaticalization in Egyptian Arabic". _Proceedings of the 2nd international conference of l'Association Internationale pour la Dialectologie Arabe_, ed. Joe Cremona, Clive Holes and Geoffrey Khan, 259–268. Cambridge: Cambridge University.

——, 2002. _Ahlan wa Sahlan: Eine Einführung in die Kairoer Umgangssprache._ Wiesbaden: L. Reichert.

JANET C.E. WATSON (University of Salford)

Affixation → Derivation; Inflection

Affrication

1. Introduction

Affrication is a process creating what are commonly called affricated stops or affricates. These sounds consist phonetically of a cluster formed by a plosive + homorganic fricative, but behave phonologically as one segment (Jakobson a.o. 1952; Sagey 1986; Rubach 1994; Clements 1999). They frequently arise as allophonic variants of non-affricated stops before high vowels and glides (e.g. [t] > [tˢ, tᶴ] / __ [i]/[j], [t] > [tˢ] / __ [u]/[w] in Korean, Japanese, Danish, Romance [Clements 1999]; and /k/ > [tᶴ] in Slavic, Arabic), but constitute contrastive phonemes in other languages (e.g. Athabaskan) where they are not bound to specific contexts.

2. Acoustic correlates of affrication

The acoustic events constituting affricates are similar to events that are usually present even in non-affricated stops. Every stop consonant has, among others, a complete constriction generally in the oral cavity, called 'closure', which induces silence at the acoustic level. This component is usually followed by a second phase, called 'release', created when the articulators come apart, and can contain up to three acoustic events (Fant 1960; Cho and Ladefoged 1999):

i. Transient: on spectrograms, this appears as vertical striations of very short duration. This noise occurs systematically at the release of voiceless plosives, and non-systematically at the release of voiced ones.
ii. Frication: this is turbulence noise generated at the narrow passage created during the release, which excites the front cavity and produces a sound similar to the homorganic fricative. This noise has peaks of energy in restricted frequency ranges that vary with the place of articulation of the stop consonant and the following vowel. Its duration is inversely proportional to intra-oral air pressure, the articulatory velocity of the release and the aperture of the following vowel. This frication, too, appears systematically during voiceless released plosives and non-systematically during voiced ones.
iii. Aspiration: this occurs at the point of the release where the supralaryngeal constriction becomes larger than the glottalic one; it is a turbulence noise, less intense than frication, generated in the glottis. This turbulence has the acoustic properties of /h/, and is concentrated at the level of the following vowel formants (mainly F2, F3, and F4). Such aspirated plosives are generally voiceless, and produced with a very large glottal opening (Dixit 1989). The aspiration duration seems, then, to be inversely proportional to the velocity of the glottal closing gesture.

It is often difficult to find sharp acoustic boundaries between these three events (Hanson and Stevens 2003). 'Transient' and 'frication' are often grouped together to form a single event called the 'burst' (Klatt 1975).

Affricates differ from non-affricated stops in that their release is dominated by the frication phase, which constitutes the main acoustic correlate of affrication. Shifts such as *t > tˢ, tᶴ* before high vocoids are generally attributed to phonetic parameters because in this context, the narrow constriction created during the release lasts longer (prolonging the duration of the frication phase) than before low vowels.

3. Affrication in Arabic

Based on Arab grammarians' phonetic descriptions (mainly Sībawayhi), the majority of modern linguists claim that Classical Arabic *jīm* was pronounced as a palato-alveolar affricate [dᶾ] (Mitchell 1993; Moscati 1980) or palatal affricate (al-Nassir 1993). [dᶾ] is also attested in several modern Arabic dialects, generally as a reflex of Classical [dᶾ] (for example in Jordan, Iraq, Algeria, Kuwait) or an allophone of /g/, the reflex of Classical Arabic /q/, when /g/ is followed by a front vowel (as in Baghdad, Kuwait). The Arab grammarians mentioned the presence, in some dialects of their time, of [tᶴ] as an allophonic pronunciation of /k/ followed by a front vowel (→ *kaškaša*). This alternation is also attested nowadays (for example in Jordan, Kuwait, and Iraq).

It is widely accepted that in Classical Arabic the noise release is much longer following /t k/

than following /ṭ q/. This asymmetry might explain why Sībawayhi groups /t k/ together with the voiceless (→ *mahmūsa*) consonants, and /ṭ q/ (most probably voiceless) with the voiced (*majhūra*) ones (Blanc 1967; for review see Ghazeli 1977). This situation also exists in several modern Arabic regions (Blanc 1967; Mitchell 1993; Odisho 1987), such as Moroccan Arabic (Heath 1987; Zeroual 2000). /t k/ are generally regarded as aspirated or slightly aspirated, and /ṭ q/ as non-aspirated and even glottalized (Marçais 1948; Odisho 1987). Heath (1987) considers that, in the central area of Morocco, "/q/ is consistently glottalized", /k/ "usually aspirated", and /t/ has a "slightly affricated release". In eastern Morocco, /t k ṭ q/ are voiceless non-aspirated stops, /ṭ q/ non-glottalized, and /t/ always produced with affrication as in [t^{s}] (Zeroual 2000).

4. Phonological analyses of affrication

Following Sagey (1986), many phonologists have claimed that an affricate is a 'contour segment', having the two phonologically ordered values [-continuant, +continuant]. In this model, alternations between [t] and [t^{s}, $t^{ʃ}$] are generally analyzed as the spreading of the feature [+continuant] from the following vowel. However, this model predicts that affricate and fricative may form a natural class with [+continuant] sounds, which seems not to be attested (Rubach 1994; Clements 1999).

For other phonologists (Jakobson a.o. 1952; Rubach 1994; Shaw 1991; Clements 1999), affricates are simple stops and therefore only possess the value [-continuant], as well as [+strident] to explain the presence of its affrication. In this model, it is at the level of phonetic implementation that the combination [-continuant, +strident], which cannot be produced simultaneously, is temporally ordered. Alternations between simple plosives and affricates are analyzed as the insertion of the feature [+strident] (Clements 1999; Kim 2001).

Notice that Classical Arabic *jīm* is counted by the Arab grammarians as one of the 'sun letters' since it is not involved in the assimilation of the definite article prefix /l/ with a following coronal (called 'moon letter') radical consonant. According to many phonologists (Clements 1976; Lahiri and Evers 1991), palato-alveolar and even palatal consonants are [+coronal]. These two observations support Cantineau's (1960) analysis according to which Classical Arabic *jīm* was not /$d^{ʒ}$/ but /g^{j}/: a 'palatalized dorsal-palatal plosive'. This analysis is, however, isolated, and it seems that Classical Arabic *jīm* was phonetically coronal, but phonologically continued to behave as its proto-Semitic cognate /g/.

Bibliographical references

Al-Nassir, Abdulmunim Abdulamir. 1993. *Sibawayh the phonologist*. London and New York: Kegan Paul International.

Blanc, Haim. 1967. "The 'sonorous' vs. 'muffled' distinction in old Arabic phonology". *To honor Roman Jakobson: Essays on the occasion of his seventieth birthday*, 295–308. The Hague and Paris: Mouton.

Cantineau, Jean. 1960. *Cours de phonétique arabe*. Paris: Klincksieck.

Cho, Taehong and Peter Ladefoged. 1999. "Variation and universals in VOT: Evidence from 18 languages". *Journal of Phonetics* 27.207–229.

Clements, Georges N. 1999. "Affricates as noncontoured stops". *Proceedings of LP '98: Item order in language and speech*, ed. Osamu Fujimura, Brian D. Joseph, and Bohumil Palek, 271–299. Prague: Karolinum Press. 271–299.

—— and Elizabeth Hume. 1995. "The internal structure of speech sounds". *Handbook of phonological theory*, ed. John Goldsmith, 245–306. Oxford and Cambridge, Mass.: Blackwell.

Dixit, R. Prakash. 1989. "Glottal gestures in Hindi plosives". *Journal of Phonetics* 17.213–237.

Fant, Gunnar. 1960. *Acoustic theory of speech production*. The Hague and Paris: Mouton.

Ghazeli, Salem. 1977. "On the controversy of the Arabic '*qaf*'". *Texas Linguistic Forum* 6.40–50.

Hanson, Helen M. and Kenneth N. Stevens. 2003. "Models of aspirated stops in English". *XVth International Congress of Phonetic Sciences*, 783–786. Barcelona.

Heath, Jeffrey. 1987. *Ablaut and ambiguity: Phonology of a Moroccan Arabic dialect*. Albany, N.Y.: State University of New York Press.

Jakobson, Roman, Gunnar Fant, and Morris Halle. 1952. *Preliminaries to speech analysis: The distinctive features and their correlates*. Cambridge, Mass.: M.I.T. Press.

Kim, Hyunsoon. 1997. *The phonological representation of affricates: Evidence from Korean and other languages*. Ph.D. diss., Cornell University, Ithaca, N.Y.

Klatt, Dennis H. 1975. "Voice onset time, frication, and aspiration in word-initial consonant clusters". *Journal of Speech and Hearing Researches* 18.686–706.

Lahiri, Aditi and Vincent Evers. 1991. "Palatalization and coronality". Paradis and Prunet (1991: 79–100).

Marçais, Philippe. 1948. "L'articulation de l'emphase dans un parler arabe maghrébin". *Annales de l'Institut d'Etudes Orientales (Alger)* 7.5–28.

Mitchell, T.F. 1993. *Pronouncing Arabic.* Oxford: Clarendon Press.
Moscati, Sabatino (ed.). 1980. *An introduction to the comparative grammar of the Semitic languages: Phonology and morphology.* 3rd ed. Wiesbaden: O. Harrassowitz.
Odisho, Edward Y. 1987. "Comments on Arabic *mufaxxama* and pharyngeal sounds". *Journal of the Syrian Academy* 4.1–11.
Paradis, Carole and Jean-François Prunet (eds.). 1991. *Phonetics and phonology.* II. *The special status of coronals: Internal and external evidence*, 125–157. San Diego: Academic Press.
Rubach, J. 1994. "Affricates as strident stops in Polish". *Linguistic Inquiry* 25.119–144.
Sagey, Elizabeth. 1986. *The representation of features and relations in nonlinear phonology.* Ph.D. diss., Massachusetts Institute of Technology.
Shaw, Patricia. 1991. "Consonant harmony systems: The special status of coronal harmony". Paradis and Prunet (1991:125–157).
Zeroual, Chakir. 2000. *Propos controversés sur la phonétique et la phonologie de l'arabe marocain.* Thèse de Doctorat, Université de Paris 8.

CHAKIR ZEROUAL (University Sidi Mohamed Ben-Abdellah & CNRS-UMR 7018, Paris)

Afghanistan Arabic

1. GENERAL

1.1 *Area*

The Arabic dialect of Afghanistan is an offshoot of the better described dialects of Central Asia, which became known initially through the work of Ceret'eli (1956). The tradition of the speakers is that they arrived in the area in the time of Tamerlane, Amīr Taymūr Kūraghān as they called him, i.e. in the 14th century C.E. There may be some truth in this as Tamerlane is known to have deported Arab populations from Syria to Central Asia. They also claim to be of the Qurayš coming from Yaman. This is less easy to substantiate. However, linguistic evidence suggests that they are an outlying section of the Bukhara Arabs who were, up till the 19th century, involved in nomadic sheep husbandry on behalf of the Bukhara Sultanate. They lived at that time in yurts and were, according to a contemporary report (Barfield 1981:8) quite prosperous, being the main suppliers of sheep for the area. These would seem to be a recent splinter group from them, having come over in the 1870s (Barfield 1981:15).

At the time of the collection of this data in 1969, the Arabs lived in four villages According to an informant, there were two villages in the region of Mazār Sharīf (Xošḥālābād, 102 families, and Yakhdān, 16 families) and two in Shibarghān (Sulflān Ariq, near Akhche, and Ḥasanābād, no details given). This entry is based on material obtained from the village of Xošḥālābād.

1.2 *Linguistic type*

The dialect is also undoubtedly related to some dialects of Khurasan in Persia. Material recently published by Seeger (2002) shows a type of Arabic resembling the Central Asian type in some ways, but differing in others and showing certain East Arabian features typical of southern Iraq *ʿarab* dialects, such as velar fronting of /k/ and /g/ as in *čibīr* 'big' and *jidīm* 'old' (Seeger 2002: 633); resyllabication of syllables with the guttural group as in *yaḥalib* 'he milks' and *yiʿarif* 'he knows' (Seeger 2002:634); syllabication of certain verbal forms such as *čitibat* 'she wrote' (Seeger 2002:635); and the form of the plural suffixes *-ow* [masc. pl.], *-an* [3rd fem. pl.] and, *-tan* [2nd fem. pl.]. This is strange, since from a purely geographical point of view, northeast Persia, northern Afghanistan, and Uzbekistan are all part of that area of Central Asia that lies along the valley of the Amu Darya. It may therefore be that the Arabic dialects of eastern Persia separated from the main block more recently than those of Afghanistan and Central Asia.

The dialect shows the influence of the neighboring languages of Dari (Afghan Persian), Uzbek (Turkic), and Pashto and has many syntactic, morphological, and phonological features not found in other Arabic dialects. It would seem on this basis to have been in the area for a considerable length of time confirming the tradition of the speakers themselves.

This area is one of considerable plurilingualism with a very uncentralized and scattered population and would also seem to be one which has historically experienced continuous linguistic change and instability right up till the present era. It is on the border of the Indo-European and Altaic language areas, with three politically important languages spoken in the immediate area, namely Persian (locally called Dari), Turkic (the local form being Uzbek), and Pashto, the language of the Afghans proper. Other groups include Dardic, Nuristani (Kafiri), and Pamir speech from Kabul northeastward to the Chinese frontier and southeast, and a dialect of Mongolian is spoken south of Herat. In addition, language, racial origin, and political grouping do not always correspond in this area. The Mongolian-descended Hazaras speak Persian, much of the 'Arab' population speaks Tajik or Uzbek, and many Pathans are of Dardic ethnic origin.

Many people in the area are at least partly bilingual, the main lingua franca being Persian. Tsereteli (1970:169) also mentions this bilingual or trilingual situation as normal for the area where the Central Asian Arabs live.

The factors contributing to the survival of the dialect so far from the Arab homeland seem to be first, the very decentralized nature of society in the area mentioned above, which contributes to the prevailing plurilingualism and second, a conscious effort made by the speakers to preserve it by the expedient of not allowing the girls of the community to marry out, hence preserving a reservoir of Arabic-speaking mothers. They also have strategies for keeping their 'foreignness' in terms of language a secret and avoid speaking Arabic in public if anyone is close by. In their own words *nās ḥayarōn tālīn ki had iš raqam ilsōn ikūn* 'people would be surprised at what type of language this was'.

1.3 *State of research*

Sources on the dialect are, in Persian Sīrat (1961), in French Kieffer (1980), and in English Sīrat and Knudsen (1973) and Ingham (1994; 2003). Sources from the neighboring dialects of Central Asia and Khurasan can also be used profitably.

2. LINGUISTIC DESCRIPTION

The main characteristics of the dialect are described below.

2.1 *Phonology*

Both the vowel and the consonant system show new developments. The quality of vowels is very much influenced by neighboring languages and the consonant system has shown reduction in the disappearance of the emphatic (pharyngealized) and the interdental set of consonants. These have merged with their non-pharyngealized alveolar correlates in the following manner:

_*ṯ > s_ _*ḏ > d_
_*ṭ > t_ _*ṣ > s_ _*ḍ > z_ _*ḏ̣ > z_

Axvlediani (1985:99) regards /ṭ/ as still being present in the Bukhara dialect. Occasional occurrences of a *ts* variant may be noted, as in *tsēr* 'bird', but in the main /ṭ/ seems to have been replaced by /t/ everywhere in the Afghan variety.

2.1.1 Inventory

i. Consonants

b	voiced bilabial plosive
v/w	alternating bilabial voiced continuant and voiced labiodental fricative
f	voiceless labiodental fricative
d	voiced dental plosive
t	voiceless dental plosive
z	voiced alveolar fricative
s	voiceless alveolar fricative
j	voiced palato-alveolar affricate
č	voiceless palato-alveolar affricate
y	voiced palatal continuant
š	voiceless palato-alveolar fricative
g	voiced velar plosive
k	voiceless velar plosive
ġ	voiced uvular fricative, sometimes pronounced plosive when initial
q	voiced uvular plosive
x	voiceless uvular fricative
ʿ	voiced pharyngeal continuant
ḥ	voiceless pharyngeal fricative
h	voiceless glottal fricative

ii. Vowels

ī ū
i u
ē (or ay) ō (or aw)
a
ā

Of these /ū/ is mid and sometimes fronted like Turkish /u/. In many cases /ō/ also occurs in places where in Arabic /ā/ would be expected, due to its having arrived via Afghan Persian which has a rounded variant of this, as in *kōn* (< *kān*) 'he was', *lōkin* 'but', *salōt* 'prayer', *xōtir* 'emotion', and in other places because of the environment of a back consonant either extant, such as *qōl* 'he said' or reduced, such as *sōr* (< *ṣār*) 'he became', *ōrd* (< *ʾarḍ*) 'earth'. Sometimes also /ū/ occurs where /ō/ is expected as in *yūm* 'day'. Final /ʿ/ is often replaced by /ḥ/ as in *šibiḥ* 'to be satisfied', *waqaḥ* 'to fall'.

2.1.2 Stress

Some unusual stress patterns occur in association with verbal suffixes. See in particular final weak verbs below.

2.2 *Morphology*

The dialect shows interesting developments in morphology. Certain developments show the influence of neighboring languages and others may purely represent the result of internal structural drift.

2.2.1 Pronouns and similar elements

i. Personal pronouns

The usual Arabic 3rd person pronouns do not occur and are replaced by forms of the demonstrative *duk* 'that'.

ana	'I'	*niḥna*	'we'
hint	'you [masc.sg.]'	*hintu*	'you [masc.pl.]'
hinti	'you [fem.sg.]'	*hintin*	'you [fem.pl.]'
duk	'he'	*duklaw*	'they [masc.]'
duki	'she'	*duklan*	'they [fem.]'

ii. Object pronoun suffixes

The forms of the suffixes are as follows: *-ni* 'me', *-(a)k* 'you', *-ki* 'you [fem. sg.]', *-u* 'him', *-(h)a* 'her', *-na* 'us', *-kum* 'you [masc.pl.]', *-kin* 'you [fem.pl.]', *-(h)um* 'they', *-(h)in* 'they [fem.]'

The 3rd person object suffixes *-ha*, *-hum*, *-hin* in many cases lose the initial *h-*, with attendant morphophonemic alternations in the verb. This is particularly important with the final weak verb type, of which many forms end in a vowel. Stress therefore becomes important in distinguishing the object function from the subject function of the 3rd pers. fem. pl. suffix *-in* as in *xassálin* (< *xassal-hin*) 'he washed them [fem.]', *xassalín* 'they [fem.] washed' (< *xassal-in*), *ana xassáltin* 'I washed them [fem.]' (< *xassalt-hin*), *hintin xassaltín* 'you [fem. pl.] washed' (< *xassaltin*), *raššūn* 'they [fem.] threw them [fem.]' (< *raššū-hin*), *libsínin* (< *libsin-hin*) 'they [fem.] wore them [fem.]', *libsítin* (< *libsit-hin*) 'she wore them [fem.]', *nintīn* (< *nintī-hin*) 'we give them [fem.]'. This stress reflects an older *-inn* form, which is still occasionally seen in the dialects of northern Arabia.

iii. Demonstrative pronouns

A demonstrative *had* [masc.], *hadi*, *hay* [fem.] 'this' also occurs in combinations as follows: *hamad* 'this very one, just this one' (< *ham hāḏa*), *hanqas* 'this many' (< *hāḏā l-qadr*), *hankīt*, *ankīt* 'here', *hallič* 'this way' (< *hāḏā l-wajh*), *hallaw* 'thus, like this' (< *hallōn*), *mimmāy* 'from this very one' (< *min ham hāḏi*), *ʿalhadyam* 'in this direction' (< *ʿalā hāḏā l-janb*), *ʿalhamadyam* 'in just this direction' (< *ʿalā ham hāḏā l-janb*), *hamallūč* 'in this very direction' (< *ham hāḏā l-wajh*). In one word, *alhawa* 'weather', the Classical Arabic definite article has been fossilized. Others demonstratives include *(h)ankīt*, *fi kīt* 'there', *min kī(t)* 'from there', *fi hān* 'here'.

iv. Interrogative pronouns

The following WH- elements occur:

iššay, iš-	'what?' (< *ʾayyu šay*ʾ)
išqar, išqas, išqa	'how many?' (< *iš qadr*)
ištūn, ištū	'how, what type?' (< *iš ṭōr*, *iš lōn*)
šahid	'which?' (< *iš wāḥid*?)
kē, ki	'why?' (< *kēf*)
kō	'where?'
ēsō	'where, in which direction?' (< *ʾayy ṣawb*)
ʿališam	'in which direction?' (< *ʿalā iš yamm*)
mata	'when?'
min	'who?'
limin	'whose?'

Examples: *iššay mitrīd, iš-mitrīd* 'what do you want?', *ištūnak* 'how are you?', *iš-missi* 'what are you doing?', *haššayāt limin inin* [sic] 'to whom do these things belong?', *hint ki matnām* 'why don't you sleep?', *mata tuġdi* 'when will you go?', *šahid bittīxa tōxida* 'which melon will you take?', *alhawa ištuwa* 'how is the weather?', *ʿumrak išqa* 'how old are you?', *bittīxa išqas mitbīʿa* 'for how much do you sell melons?', *hint ams ēso kō ġādinnak* 'in which direction were you going yesterday?', *hintu mininkum* 'who are you [masc. pl.]?', *duklan mininin* 'who are they [fem.]?', *walad mininnak* 'whose son are you?'.

Occasionally, the interrogative suffix *-mi* can also occur with the above, as in *iššay tišrab-mi* 'what are you drinking?', *iš tiktib-*

mi ‘what are you writing?’ Note also *iškinki* (< *iš yikūn ki*) ‘whatever’, showing the Persian relative pronoun *ki* ‘which’.

2.2.2 Particles

Syntactic and grammatical elements, usually undeclinable, characteristic of the dialect, which occur both in the nominal and verbal phrase, are shown below.

i. *fad* indefinite article: *fad maktab* ‘an office’.
ii. *ham* ‘also’ (Persian and Turkic): *ḥōlō ham miḥvīni* ‘he is also now protecting me’.
iii. *hast* ‘there is’ (Persian existential particle): *min kīˁud qarītin rōxir fi šibarġōn hast* ‘then in that direction there is another village in Shibarghan’. Negative existence is shown by *mōlin*: *fi bēt pašša mōlin* ‘there are no mosquitoes in the house’, *fi pulixumri kīt ˁarabi mōlīn* ‘there in Pul-i-Khumri, there are no Arabs’.
iv. *ki* ‘which, who, that, when’ (Persian): *duk parvardigōrin ki fi raḥim ummi ḥāvīni, ḥōlō ham miḥvīni* ‘that Protector, who protected me in my mother’s womb, is also protecting me now’, *nās ḥayarōn tālīn ki had iš raqam ilsōn ikūn* ‘people would be surprised at what type of language this was’, *fils ki antu xila* ‘after when they have given him money’.
v. equational particles *-inn-*, *-wa* and *-ya*. The 3rd pers. sg. shows *-wa* and *-ya* (< Classical Arabic *huwa*, *hiya*). Others show *-inn-* plus object pronoun suffixes as with the participle form: *fi masjid maktab-wa* ‘there is an office in the mosque’, *ismak iš-wa* ‘what is your name?’, *alhawa ištu-wa* ‘what is the weather like?’, *duk min-wa* ‘who is he?’, *duki min-ya* ‘who is she?’, *anāy-inni* ‘it is me’, *ana afġōn-inni* ‘I am an Afghan’, *duklaw ˁarabīn-innum* ‘they [masc.] are Arabs’, *duklan ˁarabiyāt-inin* ‘they [fem.] are Arabs’, *hint ˁarab-innak* ‘you are an Arab’, *hintin ˁarabiyāt-inkin* ‘you [fem. pl.] are Arabs’, *niḥna ˁarab-inna* ‘we are Arabs’. In the negative the forms *mōnn-* occurs: *ana afġōni mōnni* ‘I am not an Afghan’, *hintu afġōni mōnkum* ‘you [masc. pl.] are not Afghans’.
vi. *yō* ‘or’: *abu bint ilēna bintu li’awlōdna yō li’axunna mintiya* ‘the father of the girl gives his daughter to our sons or to our brothers’.

2.2.3 The noun

Nominal morphology does not, in general, differ from the general Arabic pattern except that the verbal noun has been regularized to the form *faˁalān* or *faˁlān* as in *savyān* ‘doing’, *ayfān* ‘seeing’, *nayamān*, *nīmān* ‘sleep’.

2.2.4 The verb

The basic Arabic verbal morphology system has been maintained. However, the effect of language drift has produced a quite distinct and symmetrical system, where stress operates on an underlying structure and the elision in some environments of the *-h-* in *-ha, -uh, -hum*, and *-hin* has led to other developments.

The imperfect shows a prefix *mi-* in some forms. In forms where the resulting stem begins with – CC-, i.e. the strong and Form I final weak types, this occurs in the 1st pers. sg. and the forms which would show *ya-* prefix in Classical Arabic. In forms where the resulting stem begins with Cv-, i.e. Form I medial weak and initial weak and Form II and III of all types, it occurs on all imperfect forms except the 1st pers. pl.

This is shown under the individual verb classes below.

i. Form I

This includes the types *faˁal* and *fiˁil*, the latter also including a type *fiˁl*. The majority type is *faˁal*. The *fiˁil/fiˁl* verbs, which are often intran-sitive or stative, include *ḥilf* ‘to swear’, *ḥilim* ‘to dream’, *himid* ‘to die down’, *hirb* ‘to flee’, *libis* ‘to wear’, *luḥq* ‘to follow’, *nizil* ‘to descend’, *rikb* ‘to ride’, *šibiḥ* ‘to be satisfied’, *širib* ‘to drink’, and *ˁibir* ‘to cross’. A prefix *mi-, m-* occurs with some persons in the imperfect. The basic paradigm of the strong verb is as follows:

kátab	he wrote	*míktib*	he writes
kátabit	she wrote	*tíktib*	she writes
kátabu	they [masc.] wrote	*miktibū́n*	they [masc.] write
katabín	they [fem.] wrote	*miktibín*	they [fem.] write
katábt	I wrote	*máktib*	I write
katábna	we wrote	*níktib*	we write
katábt	you [masc.sg.] wrote	*tíktib*	you [masc. sg.] write
katábti	you [fem.sg.] wrote	*tiktibī́n*	you fem.sg.] write
katábtu	you [masc.pl.] wrote	*tiktibū́n*	you [masc. pl.] write
katabtín	you [fem.pl.] wrotewrite	*tiktibín*	you [fem.pl.] wrote

Participle form

duk kātib	he has written
duki kātba	she has written
duklaw kātbīn	they [masc.] have written
duklan kātbāt	they [fem.] have written
ana kātbanni	I have written
niḥna kātbīnna	we have written
hint kātbinnak	you [masc.sg.] have written
hinti kātbanki	you [fem.sg.] have written
hintu kātbīnkum	you [masc.sg.] have written
hintin kātbankin	you [fem.pl.] have written

Where a further object suffix is attached, it is either added directly or, with the 1st and 2nd person forms, by means of the particle *-ya-*, *-yō-* (Classical Arabic *-iyyā*) as in *hinti xādankiya* 'you [fem.sg.] have taken it [fem.]', *hint xādinnakiha* 'you [masc.sg.] have taken it [fem.]', *hintin xādankinyaha* 'you [fem.pl.] have taken it [fem.]', *duk xādīha* 'he has taken it [fem.]', *zīˁānkinyōha* (< *zayyaˁankinyōha*) 'you [fem.pl.] have lost it [fem.]', *zīˁinnakyōhin* 'you [masc.sg.] have lost them [fem.]', *zīˁinkumyāha* 'you [masc.pl.] have lost it [fem.]', *zīˁankinyaha* 'you [fem.pl.] have lost it [fem.]'.

Note that the fem. pl. suffix *-in* is stressed. This reflects an underlying form *-inn*, although it is nowhere realized as double any longer. This occurs stressed also before suffixes as in *raššinin* (< *raššin-hin*) 'they [fem.] threw them [fem.]', *raššinna* (< *raššin-ha*) 'they threw it [fem.]', *tizrubinnu* (*tizru-bin-u*) 'you [fem.pl.] hit him'

The forms *fiˁil*, *fiˁl* and *fuˁl* differ only in the 3rd person of the perfect showing *libis, libsit, libsu, libsin, ḥilf, ḥilfit, ḥilfin, ḥilim, ḥilmit, ḥilim, luḥqit, luḥqin, hirb, hirbit.*

ii. Derived patterns
The dialect shows Form II: *ˁaddal/miˁiddil* 'to make, fix', *tallaq/mitilliq* 'to divorce', *qassam/ miqissim* 'to divide', III: *yāmar/ myāmir* 'to order', IV: *awra* 'to show', V: *itxarrab* 'to spoil', VI: *ityābas* 'to become dry', *izzārab* 'to fight', VII: *intabax* 'to be cooked', *insanad* 'to lean on', VIII: *ixtalat* 'to be mixed'. Note also Form IV initial weak *wijiˁ/tūjiˁ* 'to hurt'.

iii. Doubled verbs
These are now incorporated under Form II final weak. Examples are *daqqa/miduqqi* 'to beat', *farra/mifirri* 'to fly', *ḥalla/miḥilli* 'to open', *ḥamma/miḥimmi* 'to become ill', *ḥatta/miḥutti* 'to put', *lamma/milummi* 'to collect', *laffa/ miluffi* 'to twist', *qassa/miqussi* 'to cut', *šadda/mišiddi* 'to tie', *zarra/mizirri* 'to reap'.

iv. Weak verbs
a. Initial weak verbs. The weakness of these verbs is apparent in the imperfect and imperative, showing initial /ʾ/ (*hamza*) and *w*-types. Originally, initial /ʾ/ shows two verbs, which are now final weak in the perfect and initial weak with *-w-* in the imperfect: *xada* 'he took', *mōxid* 'I take', *myōxid* 'he takes', *myōxdūn* 'they [masc.] take', etc., and *kala* 'he ate', *mōkil* 'I eat', *myōkil* 'he eats', *myōklūn* 'they [masc.] eat', etc. If the initial consonant is *w* it has no exponent in the imperfect as in *waqaḥ* 'he fell', *miqaḥ* or *miqaˁ* 'he falls', *yiqaˁ* 'fall!'. Other examples are *waqaf/ miqaf*, 'to stand', *wuldit/tūld* 'to give birth', *daˁa/ midaˁi* 'to put'. The 1st pers. sg., and those which would have *ya-* in Classical Arabic in the imperfect show the *m-*, *mi-* prefix as in *maqaˁ* 'I fall', *miqaˁ* 'he falls', *miqafūn* 'they [masc.] stand', *miqaˁin* 'they [fem.] fall'.

The verb *daˁa/midaˁi* 'to put, allow' is irregular, occurring as a final weak verb in the perfect and initial and final weak in the imperfect, giving *daˁa* 'he put', *daˁin* 'they [fem.] put', *daˁīt* 'I put', *daˁīna* 'we put', but *midaˁi* 'he puts', *madaˁi* 'I put', *nidaˁi* 'we put', *middaˁi* 'you put', and in the imperative *daˁ*.
b. Medial weak verbs. Here the medial *w* or *y* is realized vocalically or as zero, giving such forms as *nām* 'he slept', *nāmat* 'she slept', *numt* 'I slept', *numna* 'we slept', *minām* 'he sleeps', *mināmūn* 'they [masc.] sleep', *nām* 'sleep!', etc. Other types show a stem in /ū/ as in *šāf/mišūf* 'to see' and a stem in /ī/, as in *šāl/mišīl* 'to carry'.

The 1st sg., 3rd fem.sg., and 2nd sg. imperfect forms show the *mi-* prefix as in *manām* 'I sleep', *minām* 'he sleeps', *mitnām* 'you [masc.sg.] sleep', *mitnāmin* 'you [fem.pl.] sleep', *miššūf* 'you [masc.sg.] see', and *miššūfin* 'you [fem.pl.] see'.
c. Final weak verbs. These only include final *-y* types. They follow the usual pattern as exemplified by *qiri/miqri* 'to read':

qíri	he read	*míqri*	he reads
qírit	she read	*tíqri*	she reads
qíru	they [masc.] read	*miqrū́n*	they [masc.] read
qirín	they [fem.] read	*miqrín*	they [fem.] read
qirī́t	I read	*máqri*	I read

qirī́na	we wrote	*níqri*	we read
qirī́t	you [masc.sg.] read	*tíqri*	you [masc. sg.] read
qirī́ti	you [fem.sg.] read	*tiqrī́n*	you [fem.sg.] read
qirī́tu	you [masc.pl.] read	*tiqrū́n*	you [masc. pl.] read
qirītín	you [fem.pl.] read	*tiqrín*	you [fem.pl.] read

Here the distinction between 3rd masc.sg. and 3rd fem.sg. object suffix is partly marked by the junction between verb and suffix as in *nintáh* 'we give him', *nintíya* 'we give her', *liqáh* 'he found him', *liqíyah* 'he found her', *nisíyah* 'he forgot her', *raššū* 'they [masc.] threw it [masc.]', *raššúwa* 'they [masc.] threw it [fem.]', *saváh* 'he did it [masc.]', *savíya* 'he did it [fem.]', *liqītī* 'you [fem. sg.] found it [masc.]', *liqītáh* 'you [fem.sg.] found it [fem.]', *šāfō* 'they [masc.] saw it [masc.]', *šāfúwa* 'they [masc.] saw it [fem.]', *šuftúh* 'you [masc.pl.] saw it [masc.]', *šuftúwa* 'you [masc.pl.] saw it [fem.]', *lá-ssah* 'do not make it [masc.]!', and *hint tintáh mi* 'will you give it to him'. The verb *sava/misi* 'to do' shows considerable reduction as follows:

sáva	he did	*mísi*	he does
sávit	she did	*míssi*	she does
sávu	they [masc.] did	*misū́n*	they [masc.] do
savín	they [fem.] did	*misín*	they [fem.] do
sīt	I did	*mási*	I do
sī́na	we did	*nísi*	we do
sīt	you [masc.sg.]did	*míssi*	you [masc.sg.] do
sī́ti	you [fem.sg.] did	*missī́n*	you [fem.sg.] do
sī́tu	you [masc.pl.] did	*missū́n*	you [masc.pl.] do
sītín	you [fem.pl.] did	*missín*	you [fem.pl.] do

Imperative

The imperative resembles that of the Arabian dialects in having no characteristic final vowel as in *irm* 'throw!', *ḥass* 'wake up!'. Note also *su* 'do!', *sō* 'do it [masc.]!', and *suwa* 'do it [fem.]!'

2.2.5 Preverbal particles

These mark negation, tense, and mode. They are often reduced forms of verbs or other elements and include the following:

i. Negators
These are *mā* and *lā*. *Mā* occurs before verbs in the indicative as in *matnām* 'will you not sleep?', *jōʿān ma talīt* 'haven't you become hungry?', *hinti ma nāymanki* 'you [fem. sg.] are not sleeping', *ana mā nāymanni* 'I am not sleeping'. However, in the 1st pers. sg., which already has the prefix *ma-*, *la* occurs as in *la manām* 'I do not sleep', *la maktib* 'I do not write'. The item *la* occurs in imperatives and wishes as in *la ssi* 'don't do [it]!', *la tintiya* 'don't give her [it]!', *la ysiya* 'let him not do it [fem.]', *la ttaʿʿibinum* 'do not [fem.pl.] make them [masc.] tired!'.

ii. Verbal prefixes
These include *mi-* and *ta-* . Of these *mi-* marks the indicative in those persons with which it occurs, as shown above, while *ta-* marks the subjunctive indicating purpose or wish, and possibly conditions as in *daʿu t-ēqaḥ* (< *ta-yiqaḥ*) 'let him fall', *kūšiš ta-nsi* 'let us make an effort', *agar xāhiš ta-nsi fad bint min fad aḥid ta-nōxid* 'if we ask to take a girl from someone [in marriage]', *kul ši ta-tizraʿ handūk tuḥsud* 'what you sow, that shall you reap'.

iii. Verbal suffixes
The suffix *-k* or *-kin* occurs on the 3rd person of some verbs. This is a type of ethic dative as in *jak(k)* 'he came', *ġadakk*, *ġadakki(n)* 'he went', *jattakin* 'she came', *ġattakin* 'she went', *hastakin* 'he is'.

iv. Modal particles
The only item extant in the available data is *daʿ-* from *daʿa/midaʿ* 'to put, allow' marking the jussive as in *daʿu tēqaḥ* 'let him fall'.

2.3 *Syntax*

2.3.1 Noun phrase structure

The definite article *al-* is not used, with some exceptions (see 2.2.1.3 above). Thus, as in neighboring Persian and Turkic, the plain noun signifies the definite. Indefiniteness is shown by the particle *fad* (see 2.2.2 i) above) and also by nunation in the form *-in*, especially between noun and adjective. This often shows assimilation of the *-n* to the following consonant as in *bētik kabīr* (< *bētin kabīr*) 'a big house', *byūtik kibōr* (*byūtin kibōr*) 'big houses'. Examples in context include *fad gappin maḥqūl* 'reasonable words', *ʿirsin kasīr* 'an abundant wedding', *min ahlin zīn* 'from a good family', *ḥintatin ḥamra* 'a red wheat', *min qōminʿarab* 'from an Arab people', *faz zamōnin* 'one time', *xītin* 'a thread', *wazīrin kōʿindu* ‘he had a wazīr’, *darwīšin šuft* ‘I saw a dervish’.

2.3.2 Sentence structure

Syntax has been deeply influenced by the neighboring languages. Most markedly, the dialect is verb final like Persian and Turkic as in *šīrviyya li-xōja šāfu* ‘Shīrwiyya saw Khāja’. Verbal objects also are frequently marked by suffixes in the verb especially if they are definite. Animate verbal objects may also be marked by the preposition *li-* or *ila-*. Also remarkable is the development of postpositions in addition to prepositions. Interrogatives are marked by a suffix *-mi* as in Turkic. This comes in sentence final position: *hint tintahmi* ‘will you give it to him?’

Inanimate object marking

bēt kōn ʿidla ‘he had made a house’, *mū nijību* ‘we bring water’, *had xyūt rūdōba ḥavītin* ‘Rūdāba kept these threads’.

Animate object marking

sīmurġ li-dūk zaġīr šāftu ‘the phoenix saw that child’, *šīrviyya li-xōja šāfu* ‘Shīrwiyya saw Khāja’, *illay giddām rūdāba waddinni* ‘take me to Rūdāba’, *ana li-duk zīn mašūfu* ‘I like him’.

2.3.3 Verbal compounds and impersonal verbs

These occur on the Persian model often with the verb *sava* ‘to do’, as in *kūšiš sava* ‘to make an effort’, *amir sava* ‘to order’. In these the object is not marked in the verb: *taʿajjuw savin* ‘they [fem.] were surprised’, *ġīm xada* ‘it clouded over’, *ana rādyō izin marmi* ‘I am listening to the radio’. Impersonal constructions include *nūmna jakkin* ‘we fell asleep’, *ġīzu jakkin* ‘he became angry’, *xōšu jak* ‘he liked it’.

2.3.4 Postpositional usage

Postpositions have grown up, possibly under the influence of Turkic. These include: *xila(f)* ‘after’, *jimīʿ* ‘with’, *giddām* ‘before’: *yōmīn xila* ‘after two days’, *min had xilaf* ‘after this’, *faras jimīʿ* ‘with a horse’, *min nayamān giddām* ‘before sleeping’.

2.3.5 Comparatives

The comparative form *afʿal* does not occur regularly. The plain adjective or the adjective with the Persian comparative suffix *-tar* is used: *had bēt min duk bēt kabīr(tar)-wa* ‘this house is bigger than that house’, *had bēt min kul byut zaġīr-wa* ‘this house is smaller than all the houses’.

3. Lexicon

Certain characteristic lexical items are shown here. The lexicon shows a perceptible resemblance to that of Iraq and eastern Arabia, though not exclusively to either: *agar* ‘if’ (Persian), *anta/minti* ‘to give’, *awta* ‘under’, *baliġa* ‘fish’ (Turkic), *ganda* ‘bad’ (Persian), *ġada(k)/muġdi* ‘to go’, *ḥidir* ‘under’, *irḥa(t)* ‘mill’, *ja(k)/miji* ‘to come’, *jōndōr* ‘wolf’ (Persian), *mu* ‘water’, *našad/minšid* ‘to ask’, *ōrd* ‘earth’, *sōr/misōr* ‘to become’, *tili/mitli* ‘to remain’, *‘ūd* ‘then’, *watar* ‘wet’, *xašim* ‘mouth’ [sic], *zaww* ‘fire’, *zīn* ‘good’, *ʿāyan/miʿāyin* ‘to look at’.

A number of lexical items connected with agriculture are shared with neighboring languages. They are listed here as evidence of the marked linguistic convergence of the area: *mōldōri* ‘animal husbandry’ (Tajik, Pashto), *mōl* ‘cattle’ (Uzbek), *qūš* ‘to plough, migrate’ (Tajik, Uzbek, Pashto), *bēda* ‘rick of twisted hay’ (Dari, Tajik, Pashto; also Uzbek ‘clover, lucerne’), *māyda* ‘small, ground up’ (Pashto, Dari, Tajik, Uzbek ‘flower’), *čiġil* ‘sieve’ (Tajik), *māysa* ‘feed, grass’ (Tajik; also Uzbek ‘young grass’), *kallakbōd* ‘pruning’ (Tajik, Uzbek *kallak* ‘to prune’).

Bibliographical references

Axvlediani, Vladimir. 1985. *Buxarskij arabskij dialekt* [The Arabic dialect of Bukhara]. Tbilisi: Mecniereba.

Barfield, T.J. 1981. *The Central Asian Arabs of Afghanistan*. Austin: University of Texas Press.

Ceret’eli, Giorgi V. 1956. *Arabskie dialekty Srednej Azii*. I. *Buxarskij Arabskij dialekt* [Arabic dialects of Central Asia. I. The Bukharan Arabic dialect]. Tiflis: Izdatel’stvo Akademii Nauk Gruzinskoj SSSR.

Ingham, Bruce. 1994. “The effect of language contact on the Arabic dialect of Afghanistan”. *Actas del congreso internacional sobre interferencias linguïsticas Árabo-Romances y paralelos extra-Iberos*, ed. Jordi Aguadé, Federico Corriente, and Marina Marugan, 105–117. Saragossa: Navarro y Navarro.

——. 2003. “Language survival in isolation: The Arabic dialect of Afghanistan”. *Proceedings of the 5th AIDA conference in Cadiz, September 2002*, ed. Ignacio Ferrando and Juan José Sánchez Sandoval, 21–37. Saragossa: Universidad de Cádiz.

Kieffer, Charles M. 1980. “L’arabe et les arabophones de Bactriane (Afghanistan). I. Situation ethnique et linguistique”. *Welt des Islams* 20.178–96.

Seeger, U. (2002), “Zwei Texte im Dialekt der Araber von Chorasan”. *“Sprich doch mit deinen Knechten aramäisch, wir verstehen es!”: 60 Beiträge zur Semitistik. Festschrift für Otto Jastrow zum 60. Geburtstag*, ed. Werner Arnold and Hartmut Bobzin, 629–646. Wiesbaden: O. Harrassowitz.

Sīrat, Abdul-Sattār. 1961. “Naxostīn taḥqīq-e ʿarabi-ye ʿāmmiyāne dar ḥavāli-ye Balkh”. *Majalle-ye Adab-e Kābul* 1.

—— and Ebbe Egede Knudsen. 1973. “Notes on the Arabic dialect spoken in the Balkh region of Afghanistan”. *Acta Orientalia* 35.89–101.

Tsereteli, George V. [Ceret’eli, Giorgi V.]. 1970. “The influence of the Tajik language on the vocalism of Central Asian Arabic dialects”. *Bulletin of the School of Oriental and African Studies* 33.169–170.

Bruce Ingham (School of Oriental and African Sudies, London)

Afro-Asiatic Languages

Together with other Semitic languages, Arabic belongs to the greater language family (or ‘phylum’, as some linguists prefer) called Afro-Asiatic (some scholars prefer the spelling Afroasiatic, since Semitic languages are spoken both in Africa and in Asia; ‘Afro-Asiatic’ was contracted by Diakonoff to ‘Afrasian’). This name has largely, although not completely, supplanted the older Hamito-Semitic (with variants Hamitosemitic and Semitohamitic/Semito-Hamitic), which has been criticized for its inadmissible reference to non-existing ‘Hamitic’. As a matter of fact, the concept of ‘Hamitic’ languages and of ‘Hamitic’ peoples has been rejected decisively by both linguists (first of all Marcel Cohen) and physical as well as cultural anthropologists. Other names, such as ‘Erythraic’ and ‘Lisramic’, have eventually been rejected.

There is general agreement that the Afro-Asiatic language family consists of the following branches or subfamilies: Semitic, Berber, Cushitic, Egyptian, and Chadic. Berber is spoken in north and northwest Africa from the Siwa oasis in Egypt to Morocco and Mauritania and in Mali and Niger further south; up to the 16th century it was spoken by at least a part of the Guanche population of the Canary islands. The name Libyco-Berber found especially in French publications is misleading since the so-called ‘Libyan’ inscriptions (actually mostly from today’s Tunisia) are in older Berber which, although known to a limited extent since most of the inscriptions are epitaphs, cannot be opposed to the rest of Berber as a separate branch. Tuareg Berber in particular has preserved many old features.

Cushitic languages are spoken in northeast Africa from Upper Egypt in the north through Sudan (east of the Nile), Ethiopia, Djibouti, Somalia, the northern half of Kenya, and even in Tanzania. The most important Cushitic languages both from the practical and the comparative linguistic point of view are: Oromo, spoken by well over 32 million native speakers mainly in Ethiopia but also in Kenya; → Somali, spoken in Somalia, Djibouti, Ethiopia, and northern Kenya by some 12.5 million; Beja, spoken by around 1.2 million people in Sudan, Eritrea, and Egypt; and ʿAfar, spoken by 1.5 million speakers in Djibouti and in Ethiopia. Beja (very tentatively classified by Hetzron as a separate branch but finally reintroduced into Cushitic, see Zaborski 1987), ʿAfar, and closely related Saho as well as Rendille (spoken in northern Kenya by over 30,000 people and classified either as a separate language or as an idiosyncratic dialect of Somali) have preserved the largest number of old linguistic features. The West Cushitic group from southwest Ethiopia was very tentatively reclassified as a separate Afro-Asiatic subfamily called Omotic, but according to Zaborski (2004b) a part of the alleged ‘Omotic’, viz. Ari, Hamer, the Banna group, and the ‘Mao’ group are not Afro-Asiatic at all, while the rest (e.g. Wollamo, Yem/Janjero, Kafa) should still be classified provisionally as West Cushitic in spite of the very strong influence of the neighboring Nilo-Saharan languages.

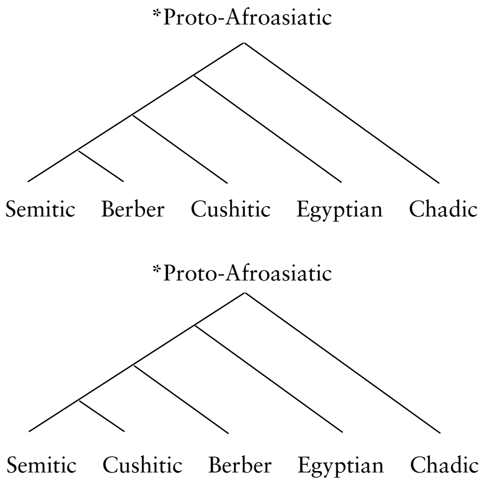
Egyptian, recorded since ca. 3000 B.C.E., is an extinct language since its last historical stage, → Coptic, died out as a spoken language under the impact of Arabic, ultimately in the period between the 17th and 19th centuries. Contrary to a widespread opinion, Arabic rather than Coptic has been the liturgical language of the Coptic church since the 1850s.

The Chadic subfamily consists of some 150 languages spoken in northern Nigeria, Niger, Chad, and northern Cameroon. → Hausa has at least some 25 million native speakers, while most Chadic languages are spoken by small groups of people, many of the Chadic languages being on the verge of extinction. Hausa is usually quoted as an example of a Chadic language since it has been described extensively, although from a linguistic point of view it displays many innovations. East Chadic languages, which are still little known, such as Dangaleat, Migama, and Mubi, preserve a number of older elements.

So far no other languages have been convincingly shown to be genetically related to Afro-Asiatic. While attempts to find Afro-Asiatic morphological cognates in some Saharan languages (e.g. Tubu/ Teda) deserve attention, the alleged genetic Afro-Asiatic links of Songhay, Nubian, not to mention Meroitic, Ful, the Mande group, and Basque cannot be accepted. The inclusion of Afro-Asiatic within the alleged Nostratic super-family is rather tentative.

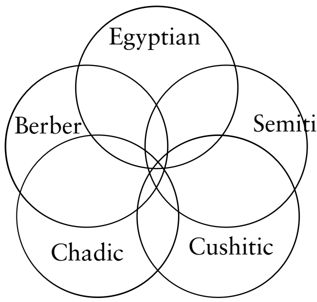
The internal relations between the main branches of Afro-Asiatic still need further elucidation, but there is no doubt that Semitic, Berber, and Cushitic, which have the so-called prefix conjugations, must be separated from Egyptian, which has only suffix conjugations. The position of Chadic in relation to other branches (although Chadic languages do belong to Afro-Asiatic) is still unclear, since most of the Chadic languages remain practically unknown. Within the Semito-Berbero-Cushitic subgroup it is possible that Berber and Cushitic are more closely related, but there are also strong isoglosses showing a special genetic relation between Semitic and Berber. As a matter of fact, isoglosses connecting and disconnecting relatively closely related languages and dialects are always more or less contradictory, so that genealogical trees as in Fig. 1 actually distort the relationship to some extent.

Figure 1. Alternative genealogical trees of Afro-Asiatic languages



The diagram in Fig. 2 may be a better graphic presentation of the complicated, partially overlapping and partially exclusive sets of features.

Figure 2. Relationships between the Afro-Asiatic languages



Personal pronouns provide the most obvious proof of the genetic unity of the Afro-Asiatic languages (Table 1).

Personal morphemes of the conjugation also show relatively close relationship (Table 2).

Afro-Asiatic languages have a considerably large Proto-Afro-Asiatic lexicon in common. Comparative lexical-etymological studies are still at an initial stage, since many indispensable preparatory studies (good dictionaries of many languages, comparative dictionaries of smaller groups, etc.) are still to be made. Consequently, regular sound correspondences or sound laws have only partially been discovered. A comparative dictionary of the Semitic roots (D. Cohen 1970–1976) is slowly proceeding; Militarev and Kogan (2000) is a good dictionary but concentrates on particular semantic fields; Dolgopolskij (1973) is a very good example of combined comparative-historical phonology and etymology with many references to Arabic. A dictionary of Berber roots has been started by Kamal Naït Zerrad (1998). For Egyptian, Takacs (1999) has launched a large-scale etymological project. There have been initial reconstructions of Proto-Chadic (mainly Jungraith-mayr and Ibriszimow 1994; Newman 1977) and of Proto-Cushitic (Dolgopolskij 1973; Sasse 1979). On the other hand, preliminary comparative dictionaries of the whole of Afro-Asiatic, starting with Marcel Cohen's pioneer study (1947), are either controversial (Orel and Stolbova 1995) and unfinished (e.g. Diakonoff 1993–1997), or not acceptable to many scholars (Ehret 1995). In any case, there is no doubt that regular sound correspondences will be discovered, eventually.

Table 1. Pronouns in the Afro-Asiatic languages

		Arabic	Akkadian	Tuareg	ʿAfar	Rendille	Coptic	Hausa	Kulere
sing.									
	1	*ʾana*	*ana*, *anāku*	*näk*	*anu*	*ani*	*ani*	*ni*	*nì*
	2m	*ʾanta*	*attā*	*kay*	*atu*	*ati*	*ento-k*	*kay*	*yá*
	2f	*ʾanti*	*attī*	*käm*	*atu*	*ati*	*ento*	*kē*	*kí*
	3m	*huwa*	*šū*	*ənta*	*usu-k*	*usu*	*ento-f*	*šī*	*ší*
	3f	*hiya*	*šī*	*ənta*	*is*	*iče*	*ento-s*	*ita*	*tí*
plur.									
	1	*naḥnu*	*nīnu*	*näkkan(t)i-ḍ*	*nanu*	*naḥ*, *inno*	*ano-n*	*mū*	*ní*
	2m	*ʾantum*	*attunu*	*kawani-ḍ*	*isin <*itin*	*atin*	*ento-tn*	*kū*	*kú*
	2f	*ʾantunna*	*attina*	*kämäti-ḍ*	*isin*	*atin*	*ento-tn*	*kū*	*kú*
	3m	*hum*	*šunu*	*əntani-ḍ*	*oson*	*ičo*	*ento-w*	*sū*	*sí*
	3f	*hunna*	*šina*	*əntanäti-ḍ*	*oson*	*ičo*	*ento-w*	*sū*	*sí*

Table 2. Personal morphemes of the conjugation in Afro-Asiatic languages

		Kabyle Berber	Tuareg Berber		Saho		
		preterite	preterite	present	perfect	imperfect	subjunct.
		‘to send’	‘to take’		‘to know’		
sing.	1.	*u-zne-ġ*	*e-rmese-ġ*	*râmmes-e-ġ*	*ē-ḍeg-e*	*ā-ḍig-e*	*ā-ḍag-o*
	2m	*tu-zne-ḍ*	*te-rmese-d*	*te-râmmese-d*	*tē-ḍeg-e*	*tā-ḍig-e*	*tā-ḍag-o*
	2f	*yu-zen*	*te-rmese-d*	*te-râmmese-d*	*tē-ḍeg-e*	*tā-ḍig-e*	*tā-ḍag-o*
	3m	*yu-zen*	*i-rmes*	*i-râmmes*	*yē-ḍeg-e*	*yā-ḍig-e*	*yā-ḍag-o*
	3f	*tu-zen*	*te-rmes*	*te-râmmes*	*tē-ḍeg-e*	*tā-ḍig-e*	*tā-ḍag-o*
plur.	1.	*nu-zen*	*ne-rmes*	*ne-râmmes*	*nē-ḍeg-e*	*nā-ḍig-e*	*nā-ḍag-o*
	2m	*tu-zne-m*	*te-rmese-m*	*te-râmmese-m*	*tē-ḍeg-in*	*tā-ḍig-in*	*tā-ḍag-ō-na*
	2f	*tu-zne-m-t*	*te-rmes-m-et*	*te-râmmes-m-et*	*tē-ḍeg-in*	*tā-ḍig-in*	*tā-ḍag-ō-na*
	3m	*u-zne-n*	*e-rmese-n*	*râmmese-n*	*yē-ḍeg-in*	*yā-ḍig-in*	*yā-ḍag-ō-na*
	3f	*u-zne-n-t*	*ermes-n-et*	*râmmes-n-et*	*yē-ḍeg-in*	*yā-ḍig-in*	*yā-ḍag-ō-na*

Apart from genetic relationship, Afro-Asiatic languages have influenced each other through contact, so that some areal features have developed. Arabic has been in contact with → Berber for some 13 centuries, and Berber has greatly influenced Western, i.e. Maghrebi, dialects of Arabic in the fields of phonetics (reduction of vowels and introduction of many consonant groups), vocabulary, and syntax (→ Berber loanwords). Coptic, which has totally receded under the pressure of Arabic, has left some loanwords in Egyptian dialects (a few even in Literary Arabic) and very few, if any at all, traces in the morphology and syntax (→ Coptic loanwords). Since Arabic as spoken in Ethiopia and in the eastern Sudan is still little known, it is difficult to say to what extent it has interfered with the local Afro-Asiatic languages there, although it is known that in the spoken Arabic of → Ethiopia the syntactic order SOV typical of Cushitic (in Ethio-Semitic it is due to Cushitic influence) has been adopted at least in some varieties, resulting in sentences like *huwa l-bayt jāʾa* ‘he came home’ instead of Standard Arabic *huwa jāʾa l-bayt*. Chadian and Nigerian Arabic must have been influenced to some extent by Chadic (not to mention Nilo-Saharan) languages, but details require further investigation. Pidgin and Creole Arabic (→ Ki-Nubi of Uganda and Nairobi, → Juba Arabic, and a Chadian Arabic pidgin, → Chad) appeared due to special contact with non-Afro-Asiatic languages (→ creolization, → pidginization).

As far as typological features are concerned, internal inflexion, also called ablaut or → apophony is typical not only of Semitic, but of all other ‘old’ branches of Afro-Asiatic/Hamito-Semitic, although it is receding in more innovating languages. Berber, Cushitic, and Egyptian have highly regular sets of separate negative conjugations parallel to positive ones (Zaborski 2003). This coincides with the fact that in Arabic the old preterite, viz. ‘apocopate’, *yaqtul*, and in later Akkadian its cognate *iprus* survive mainly in negative sentences. Older Afro-Asiatic languages do not seem to have periphrastic verbal constructions, although in all probability such constructions can be reconstructed for Proto-Afro-Asiatic (Zaborski 2005). Attempts at reconstructing ergative constructions for Proto-Afro-Asiatic have as yet failed to convince (see Zaborski 1999b).

For Arabists it is important to know that other Afro-Asiatic languages provide evidence that some features of Arabic go back not only to Proto-Semitic but also to Proto-Afro-Asiatic, thus proving that Classical Arabic, contrary to the opinion current mainly among Assyriologists, is not among the most innovating Semitic languages at all, but has preserved many ancient elements, so that in relative terms it is as ‘ancient’ as Akkadian, which had first been recorded over 2,500 years earlier. Feminine *-(a)t* is found everywhere, e.g. Egyptian *sn* ‘brother’ and *sn-t* ‘sister’, Tuareg *a-mîdi* ‘companion, friend’, fem. *ta-mîdi-t*. Singulatives or *nomina unitatis* are formed with *-Vt/tV*, e.g. ‘Afar *adám* ‘man’, singulative *adām-to*, pl. *adām-um*, *dummu* ‘cat(s)’, singulative *dummū-ta*. On the other hand, *-u/w* is the morpheme of masculine, e.g. in the Beja definite article ʾ*un* < **w-un*; Kabyle Berber absolute state *a-rgaz* ‘man’, annexed state *wə-rgaz*, *a-meksa/u-meksa* ‘shepherd’.

Internal plurals, which have been considered by many Semitists to be a ‘Southern Semitic’ innovation, have cognates (not only in the case of forms with *-ā-*), first of all in Berber, e.g. Tuareg *a-fus* ‘hand’, pl. *i-fass-en*, *akâl* ‘country’, pl. *i-kall-en*; Cushitic, e.g. Beja *banūn* ‘eyebrow’, pl. *banín*, *hūm* ‘brain’, pl. *him*, *ōr* ‘son’, pl. *ar*; in Chadic, and even in Egyptian. The masculine plural ending *-ū* (see Zaborski 1976) has a good cognate in Egyptian *-w* (**-ow* or **-aw*), e.g. *sn-w* ‘brothers’; Berber *-aw*, e.g. Tuareg *ul* ‘heart’, pl. *ul-aw-en*, *e-ġef* ‘head’, pl. *i-ġaf-aw-an*; Cushitic, e.g. ‘Afar *bus* ‘vagina’, pl. *bus-wa*, *lubak* ‘lion’, pl. *lubak-wa*; and Chadic. Plural and abstract nouns with plural in *-ān* (e.g. Arabic *furs-ān* ‘riders’) occur also at least in Berber, e.g. Tuareg *ânu* ‘well, spring’, pl. *ûn-ân*, *a-kli* ‘slave, serf’, pl. *i-kl-ân*, and in Cushitic and Chadic. The dual is found in Egyptian, e.g. *sn-w-y* ‘two brothers’, *sn-t-y* ‘two sisters’.

The construct state is common and the genitive ending *-i* has a good cognate at least in Beja and in ‘Afar-Saho, e.g. *awk-a* ‘boy’ and *awk-i migaʿ* ‘a boy’s name’. In Cushitic there is also good evidence of the diptotic declension (cf. Sasse 1984; → diptosis) with absolutive/oblique case in *-a* used not only as accusative but also as citation form and after a copula. *Nisba* ending *-ī* is found practically everywhere, e.g. Egyptian *nwt* ‘town’, *nwt-y* ‘from the town, urban, citizen’; *kmt* ‘Egypt’, *kmt-y* ‘Egyptian’.

Nouns of place and nouns of instrument have cognates, e.g. in Egyptian *m-sḫn* ‘rest place’, *sḫny* ‘to stay, to settle’; *m-nḫ-t* ‘clothes, dress’, *wnḫ* ‘to dress, to put on clothes’; Hausa *ma-karant-ā* ‘school’, *karant-ā* ‘to read’; *ma-dūb-ī* ‘mirror’, *dūb-ā* ‘to look at’; *ma-kull-ī* ‘key’, *kullē* ‘to lock’.

Especially Cushitic ‘Afar-Saho shows that the Arabic → energicus (which has good cognates, for instance, in Ugaritic and in Modern South Arabian, while being only residual in Akkadian in the form of the so-called ‘ventive’) goes back to Proto-Semito-Berbero-Cushitic, as well, e.g. ‘Afar *ta-dūr-en keh* ‘that you return’, *ta-dūr-em takkeh* ‘you may return’ (see Zaborski 2004c).

The Arabic tense/aspect/mood endings, ‘imperfect’ *-u* and ‘subjunctive’ *-a* have cognates, too, with rather common functional and phonological shifts, in ‘Afar endings, e.g. *ta-dūr-u* ‘that you return’, *tu-dūr-e* ‘you returned’, negative imperfect *ma ta-dūr-a* ‘you do not return’ (cf. East Chadic Birgit perfect *-i*, imperfect *-a*, subjunctive *-o*), and thus go back to the prehistoric period, while in Akkadian *-u* has survived mainly in non-final position, where it has been reinterpreted as a morpheme of subordination (wrongly labeled ‘subjunctive’), e.g. subordinate present *iparras-u*, independent (final) present *iparras* (Zaborski *forthcoming*). Berber ‘present’ (usually called ‘habitual’ or ‘intensive aorist’) is formed not only with gemination of the second root consonant (e.g. Tuareg *i-kârras* ‘he makes a knot’), which has a good cognate in Akkadian and in Ge‘ez, but also with prefixed *t-* (e.g. Tuareg *i-tâkär* < **yi-tä-hkär* ‘he steals’, *yu-kär* ‘he stole’), which may be accompanied either by gemination of C2 (e.g. Tuareg *i-tâ-mättât* ‘he dies’, *yä-mmut* ‘he died) or by vowel lengthening (e.g. Tuareg *i-tâfâdây* ‘he is skinned’, *yä-fidäy* ‘he was skinned’). This demonstrates that not only intensive *yuqattilu/***yaqattalu* but also *yataqattalu* and *yataqātalu* and **yat(a)qatilu* >

yaqtatilu/*yaqqatilu* forms were already used for the renewal of the 'present' in Proto-Semito-Berbero-Cushitic. As a matter of fact, many Arabic verbs of the Forms II, V, VI, and VIII have exactly the same meaning as Form I verbs (e.g., they are not intensive, causative, reflexive, mediopassive, or passive), which means that they are remnants from the period when Proto-Arabic, like Berber, had 'present' conjugations of the **yaqattal*, *yataqattal*, *yataqātal*, and *yaqqatal* (cf. also *yanqatil* > *yaqqatil*, e.g. in Akkadian and in Hebrew) type. The derived form *qātala* of Form III is not an innovation of 'Southern Semitic' preserved mainly in Arabic (residual in Geʿez) since it is found also in Cushitic Beja (Zaborski 1994), where it functions as an intensive form as well as the non-intensive present of a group of verbs. This explains why so many *qātala* forms in Arabic are actually intensive/iterative. The internal passive of the *qutila* type has also frequently been regarded as an innovation, but it has cognates in Egyptian (Zaborski *forthcoming*). Both Beja and ʿAfar-Saho have the auxiliary *an* 'to be', which most probably has a cognate in the Semitic *-n-* prefix of mediopassive verbs of Form VII. Berber has both *n-* and *m-*, which make mediopassive and passive forms, Cushitic has *m-* and *n-*, which may explain, together with Egyptian (e.g. *m-ḥnk* 'endowed' from *ḥnk* 'to give a gift, to endow'; *m-nhzy-w* 'watching, being on a guard' from *n-hzy* 'to be vigilant') and Chadic (e.g. Hausa *má-kaḍ-í* 'taylor') the origin of Arabic *m-a-*, i.e. the prefix of the passive participles, which can hardly be identified with the interrogative *ma* (Zaborski 1999a). It may also explain the origin of *m-u* as prefix of both active and passive participles in derived verbs. It is also possible that the *-t-* prefix of Form VIII goes back to an original auxiliary, found in Egyptian as a suffix. Cushitic and partially Berber confirm that all verbal derivational prefixes *t-*, *n-*, *š/s/h/ʾ-* of Semitic including Arabic were originally free standing morphemes, since in Cushitic they appear not only as prefixes but also as suffixes. The active participle *qātil*, which is related to the originally intensive *qātala* has a cognate in Egyptian (Zaborski 1999a). The Arabic Form III *yuqātilu* (in which the lengthening of the *-a-* preceding C2 was only a phonetic variant of the gemination of C2) was originally intensive > durative. The allegedly principal 'conative' function was only a semantic variant of intensive, especially in resultative verbs: *yuqātilu* with singular object can mean only 'to repeat efforts to kill', since a single object can be killed only once. It was used for the renewal of the 'present' in Proto-Arabic and is still used as present/'imperfect' in Modern South Arabian, e.g. Mehri *ya-rákb-en* 'he rides' with originally 'energetic' *-en*. It also occurs in conditional sentences (subjunctive *ya-rōkəb*). In Beja past < old present, e.g. *te-kātim-a* 'you used to come repeatedly to the same place'. In this respect, Classical Arabic is actually more ancient than Akkadian since it had several 'presents', both with gemination and long *-ā-* (and both sometimes accompanied by *ta-*), while Akkadian seems to have preserved the *iparras(-u)* present only.

Bibliographical references

Cohen, David (ed.). 1988. *Les langues dans le monde ancien et moderne*. III. *Les langues chamito-sémitiques*. The Hague and Paris: Mouton.

—— a.o. 1970–1976. *Dictionnaire des racines sémitiques ou attestées dans les langues sémitiques*. Paris: Mouton.

Cohen, Marcel. 1947. *Essai comparatif sur le vocabulaire et la phonétique du chamito-sémitique*. Paris: Champion.

Diakonoff, Igor Mikhailovich. 1988. *Afrasian languages*. Moscow: Nauka.

—— a.o. 1993–1997. "Historical comparative vocabulary of Afrasian". *Sankt-Peterburgskij Žurnal Afrikanskix Issledovanij* [St. Petersburg Journal of African Studies] 1.7–15; 2.5–28; 3.5–26; 4.7–38; 5.5–32; 6.12–35.

Dolgopolskij, Aron Borisovič. 1973. *Sravnitel'no-istoričeskaja fonetika kušitskix jazykov* [The comparative-historical phonetics of the Cushitic languages]. Moscow: Akademija Nauk SSSR.

——. 1999. *From Proto-Semitic to Hebrew: Phonology: Etymological approach in a Hamito-Semitic perspective*. Milan: Centro Studi Camito-Semitici.

Ehret, Christopher. 1995. *Reconstructing Proto-Afroasiatic (Proto-Afrasian): Vowels, tone, consonants and vocabulary*. Berkeley: University of California Press.

Hetzron, Robert. 1980. "The limits of Cushitic". *Sprache und Geschichte in Afrika* 2.7–126.

Jungraithmayr, Hermann and Dymitr Ibriszimow. 1994. *Chadic lexical roots*. I. *Tentative reconstruction, grading, distribution*. II. *Documentation*. Berlin: D. Reimer.

Kienast, Burkhart. 2001. *Historische Semitische Sprachwissenschaft*. Wiesbaden: O. Harrassowitz. [With contributions by Erhart Graefe (Old Egyptian) and Gene B. Gragg (Cushitic).]

Militarev, Alexander and Leonid Kogan. 2000. *Semitic etymological dictionary*. I. *Anatomy of man and animals*. Münster: Ugarit-Verlag.

Nait-Zerrad, Kamal. 1998–. *Dictionnaire des racines berbères (formes attestées)*. Paris and Louvain: Peeters.

Newman, Paul. 1977. “Chadic classification and reconstructions”. *Afroasiatic Linguistics* 5.1–42.

Orel, Vladimir E. and Olga V. Stolbova. 1995. *Hamito-Semitic etymological dictionary: Materials for a reconstruction*. Leiden: E.J. Brill.

Rössler, Otto. 1950. “Verbalbau und Verbalflexion in den semitochamitischen Sprachen”. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 100. 461–514. (English trans., *Essays in historical linguistics in memory of J.A. Kerns*, ed. Yoël L. Arbeitman and Allan R. Bomhard, 679–748. Amsterdam: J. Benjamins, 1981.)

Sasse, Hans-Jürgen. 1979. “The consonant phonemes of Proto-Cushitic”. *Afroasiatic Linguistics* 7.1-67.

——. 1984. “Case in Cushitic, Semitic and Berber”. *Current Progress in Afro-Asiatic Linguistics*, ed. James Bynon, 111–126. Amsterdam: J. Benjamins.

Satzinger, Helmut. 1997. “Egyptian in the Afroasiatic frame”. *Afroasiatica Neapolitana*, ed. Alessandro Bausi and Mauro Tosco, 27–48. Naples: Istituto Universitario Orientale.

Takács, Gabor. 1999ff. *Etymological dictionary of Egyptian*. Leiden: E.J. Brill.

Voigt, Rainer M. 2002. “The Hamitic connection: Semitic and Semitohamitic”. *Israel Oriental Studies* 20.265–290.

Zaborski, Andrzej. 1987. “Reinisch and some problems of the study of Beja today”. *Leo Reinisch: Werk und Erbe*, ed. Hans G. Mukarovsky, 125–139. Vienna: Verlag der Akademie.

——. 1994. “Archaic Semitic in the light of Hamito-Semitic”. *Zeitschrift für Althebraistik* 7.234–244.

——. 1999a. “On Hamitosemitic participles”. *Afroasiatica Tergestina*, ed. Marcello Lamberti and Livia Tonelli, 35–39. Padua: Unipress.

——. 1999b. “On the alleged ergativity in Hamitosemitic/Afroasiatic languages”. *Collectanea linguistica in honorem Casimiri Polanski*, ed. M. Brzezina and H. Kurek, 309–317. Cracow: Ksiegarnia Akademicka.

——. 2003. “Negative conjugations in Hamitosemitic”. *Das alte Ägypten und seine Nachbbarn: Festschrift zum 65. Geburtstag von Helmut Satzinger*, ed. Monika R.M. Hasitzka, Johannes Diethart, and Günther Dembski, 353–359. Krems: Österreichisches Literatur-forum.

——. 2004a. “Non-causative verbs of the ‘causative’ *ʾaqtala* class in Arabic and **yuqtilu* conjugation in Arabic”. *Festschrift Gideon Goldenberg*, 000. Jerusalem: 000.

——. 2004b. West Cushitic: A genetic reality”. *Lingua Posnaniensis* 46.173–186.

——. 2004c. “Energicus and other modals in Cushitic”. *Verbum et calamus: Semitic studies in honour of the sixtieth birthday of professor Tapani Harviainen*, 435–439. Helsinki: Academia Fennica.

——. 2005. “Proto-Hamitosemitic periphrastic conjugations”. *Proceedings of the 10th meeting of Hamito-Semitic/Afroasiatic Linguistics*, ed. P. Fronzaroli and P. Marrasini, 85–94. Città di Castello: Università di Firenze.

——. Forthcoming. “Verbal categories of Proto-Hamitosemitic”. *Proceedings of the Conference on Semitic Linguistics, Oslo 2004*.

Andrzej Zaborski (University of Cracow)

Agent

There has been a terminological mix-up and indeterminacy over the syntactic and the semantic notion of *fāʿil* by Arab grammarians, who distinguished between transitive and intransitive verbs, but did not focus on inherent componential features of the verb or the noun (phrase) when investigating the semantic role of *fāʿil* in a given proposition. Arab grammarians did not provide clear-cut semantic criteria for the potentially agentive nominal. Their characterization, therefore, was syntactic, rather than semantic. Arab grammarians assigned the role of Agent (→ *fāʿil*) to post-verbal nominal constituents regardless of their inherent semantic properties and without taking into consideration the semantic properties of the verb. Thus, the role of Agent is designated by its position: postverbal for the Baṣra school of grammar and post- or preverbal for the Kūfa school of grammar. The quasi-semantic label *fāʿil* was applied to a general syntactic category, namely, Subject, and lost its semantic thrust. Although case markers, such as nominative and accusative, are surface structure grammatical indicators, they dominated the characterization of the underlying grammatical functions, such as *fāʿil* and *mafʿūl*, of nominal constituents of a given proposition. Case markers in Arabic, however, cannot consistently account for the semantic roles of nominals in a proposition. The traditional Arab grammarians’ account of *fāʿil* could not capture the intuitive judgment that *al-waladu* ‘the boy’ in sentences like *ġafā l-waladu* ‘the boy fell asleep’ and *rabiḥa l-waladu jāʾizatan* ‘the boy won a prize’ is not the *fāʿil* ‘Agent’ of any action. It is rather an Experiencer and Benefactive, respectively.

The inadequacy of the traditional account of the nominal Subject has led modern Arab linguists such as Ḥasan (1963) and al-Maxzūmī (1964) to distinguish between *al-fāʿil al-muxtār* ‘the volitional doer’ and *al-fāʿil ġayr al-muxtār* ‘the non-volitional doer’. The impact of European linguistics on the Arab linguistic tradition has now started to take shape. Agent, in the view of Kearns (2000:188 ff.), is a thematic role in terms of thematic roles theory (→ case theory, → theta roles), which was already proposed by the Sanskrit grammarian Pāṇini in 500–400 B.C.E. who established classes of noun phrases according to the broad interpretation of their grammatical form. However, thematic roles have mainly been syntactic, and the thematic role of Agent overlaps with other causative nominals such as *ar-riyāḥu* 'the wind' in *kasarat ar-riyāḥu s-siyāja* 'the wind smashed the fence'. Traditional thematic roles theory was concerned with the interpretation of nominals according to their grammatical characteristics, in particular, position in the sentence, case marking, and prepositions, if any, preceding the noun phrase (Kearns 2000:191).

We view Agent as a semantic role assigned to nominal constituents in a given proposition. It is a semantic notion that denotes the actual doer of an action embodied in an action verb. Thus, before embarking on the syntactic characterization of a given noun (phrase), one needs to distinguish between an action verb and a non-action verb, and 'an actual doer' *fāʕil ḥaqīqī* and 'a non-actual doer' *fāʕil* in the traditional sense, i.e. the Subject which can also be 'a fronted Subject' *mubtadaʾ*, (→ *ibtidāʾ*) regardless of its semantic role.

Agent is a case notion in terms of case grammar theory developed by Charles Fillmore in 1968. Case grammar, however, does not provide a solution to the problem of which nominal constituents should be assigned the role of Agent and why. Linguists failed to pinpoint clear-cut semantic criteria for the characterization of agentivity, though they are in agreement about animacy as the criterial feature of agentivity, i.e. only an animate noun (phrase) can undertake the semantic role of Agent. Proponents of this view include Gruber (1967), Fillmore (1968), Lakoff (1968), Lyons (1977), and Jackendoff (1972). Other linguists, such as Chafe (1970) and Platt (1971), disagreed, arguing that the role of Agent could be assigned to both animate and inanimate constituents. The inherent componential features of either the noun or the verb may determine what constitutes an Agent. Two opposing views emerged regarding which is the central and which the peripheral constituent, the noun or the verb, when agentivity is investigated. Linguists such as Fillmore (1968) and Lakoff (1977) accounted for agenthood according to the inherent features of the noun (phrase), while linguists such as Chafe (1970) and Fillmore (who made a u-turn in 1970) claimed that agentivity can only be determined by the inherent features of the verb. The latter view of the centrality of the verb in a given proposition gained momentum and was supported by other linguists such as Cook (1971) and Andrews (1985). It was held earlier by the French linguist Tesnière (1965).

Here, the view will be taken that Animacy is the archetypal feature of agentivity, while inanimate subject nominals cannot be assigned the semantic role of Agent. Similarly, the semantic nature of the verb should be the criterial ingredient of agenthood. A refined blend of the noun-based and verb-based accounts needs therefore to be introduced for the characterization of the case role Agent. In other words, agentivity cannot be established by a one-sided analysis; the semantic features of both the noun (phrase) and the verb should be considered. Thus, the criterial features of both the agentive noun (phrase), i.e. agentive Subjects, and the agentive verb need to be specified. First of all, potentially agentive nominals should enjoy a number of archetypal inherent agentive componential features such as: [+ Human], to be distinguished from non-humans; [+ Higher Animate], to be distinguished from lower animates like animals, which lack the mental capability to perceive commands; [+ Volition], to be distinguished from nominals which lack the will in making a decision such as lower animates; [+ Potent], to be distinguished from impotent animate or inanimate nominals; [+ Control], to be distinguished from nominals that cannot control their actions; [+ Intentionality], to be distinguished from nominals that do not express an intention while carrying out an action; [+ Imperativeness], to be distinguished from nominals that cannot receive an order; and [+ Active], to be distinguished from non-active nominals such as Objects.

However, some of these relational features are shared by non-agentive nominals which co-occur with agentive verbs, as in *dammara l-maddu l-baḥriyyu l-manāzila wa-qatala n-nāsa* 'the tsunami destroyed the houses and killed the people'. Although the Subject nominal *al-maddu l-baḥriyyu* 'the tsunami' is [- Human], [- Animate], [- Volition], [- Control], [- Intentionality], and [- Imperativeness], this Subject nominal enjoys other agentivity features such as [+ Potent] and [+ Active]. In other words, it is a causer and can bring about some event or state. Yet, it cannot be assigned the case role Agent, since only [+ Higher Animate] can be commanded because they are potentially [+ Human], which explains why sentences like *yā maddu dammir al-manāzil wa-qtul an-nāsa* 'oh tsunami, destroy the houses and kill the people' are not acceptable. Similarly, [- Higher Animate] and [- Human] Subject nominals such as dogs or chimpanzees cannot undertake the role of Agent since they do not enjoy sophisticated mental systems and decision-making which are characteristic of [+ Human] nominals. Although lower animates can receive commands, they cannot, for instance, solve mathematical problems, invent, write articles, report, be witnesses in court, perform actions purposively and with a presupposed will, or give orders or advice. Thus, the relational feature [+ Human] is a necessary corollary of the case role Agent. The [+ Human] feature, however, has not been accounted for by Wierzbicka (2002), who links the notion of Agent directly with the universal concept 'Do' *yaʿmal* as in *šaxṣun māʿamila šayʾan mā* 'someone did something' without linking the fundamental agentive feature [+ Human] to the agentive nominal. For Givón (2001:52), volition is the prototypical agentive feature. Kearns (2000:192, 237), however, assigns the case role Agent to [- Human] inanimate nominals, but links agentivity with the relational features of conscious control, volition, and intention.

After the investigation of the criterial features of the agentive nominal, those of the agentive verb have to be specified as well. A syntactically motivated analysis is irrelevant here since it accounts only for surface structure syntactic functions of nominal constituents in a given construction. Arabic derivative verb forms are vital in the assignment of agentivity. Morphological changes in a given verb lead to a change in the case frame of the verb. A verb like *kataba* 'to write' is an agentive verb, and so is its derivative form *kattaba* 'to make someone write'. Both verb forms allow a potential agentive nominal. However, sentences like *kattaba l-mudarrisu ṭ-ṭullāba d-darsa* 'the teacher made the students write the lesson' can be regarded as having two Agents. The first is the primary Agent *al-mudarrisu* 'the teacher', who initiates the action of writing, and the second, although the Object, is the secondary Agent *aṭ-ṭullāba* 'the students', which is the performer of the action of writing. The verb form, therefore, is crucial in the assignment of the role of Agent to a nominal. Two verb types can be differentiated in terms of their selectional semantic features: dynamic (action) verbs and stative verbs. In the present account, we are concerned only with dynamic verb forms, which are agentive verbs that allow a potential agentive nominal. Stative verbs, however, are nonagentive and co-occur with case roles other than Agent, which are outside the scope of the present investigation. Agents, therefore, do not co-occur with stative verbs, such as *māta* 'to die', *ʿalima* 'to know', *nāma* 'to fall asleep', *xasira* 'to lose'. It should be clear by now why a verb-based account is paramount in the analysis of agentivity in Arabic. The relational feature [+ Human] can co-occur with both dynamic and stative verbs. A Subject nominal like *al-fatātu* 'the girl' can occur with both verb types, but the case role it is assigned differs depending on the type of verb associated with it. Thus, the case role Agent is contingent on the semantic nature of the verb in a given proposition. Relational features like [+ Volition], [+ Control], and [+ Active] can also be imputed to the agentive nominal. These relational components allow the dynamic verb to occur in the progressive aspect, as in constructions like *mā zāla r-rajulu yukassiru l-ʾaxšāba* 'the man is still breaking the logs into pieces' and *kānat xadījatu tasīru bi-surʿatin* 'Xadīja was walking fast'. The semantic properties of dynamic verbs like *kassara* 'to break into pieces' and *sāra* 'to walk' illustrate that such verbs can occur in the progressive aspect, which pictures the action as still being carried out and as being performed with volition by a wilful [+ Human] Subject nominal in full control over what is being done. Dynamic verbs like the above can also take the relational feature [+ Imperativeness], as illustrated by imperatives like *iksir* 'break', *kassir* 'break into pieces', or *sir* 'walk'. Stative verbs, however, like *šakka* 'to doubt' and *raʾā* 'to see' cannot; they are semantically [- Imperative] due to their semantic nature. As a result, Agents can occur with dynamic verbs only.

An exhaustive characterization of Agent in Arabic should take into account both the lexical features of the Subject noun (phrase) and the semantic properties of the verb. Neither an agentive verb nor agentive lexical features, such as [+ Human], alone can predict the occurrence of Agent. The Subject nominal *aṭ-ṭabību* 'the doctor' in *istalama ṭ-ṭabību jāʾizatan* 'the doctor received a prize' is a [+ Human], but it cannot be assigned the case role Agent, because semantically the verb *istalama* 'to receive' is characterized by the componential feature [+ Benefactive]. The verb *kasara* 'to break', for instance, co-occurs with [+ Human] nouns and similarly with non-human, inanimate, and natural force nouns, as in *kasara l-kalbu l-lawḥata* 'the dog broke the painting', *kasarat al-ḥijāratu š-šubbāka* 'the stone broke the window', and *kasarat ar-rīḥu s-siyāja* 'the wind broke the fence'. Yet, none of the Subject nominals *al-kalbu* 'the dog', *al-ḥijāratu* 'the stone', and *ar-rīḥu* 'the wind' is an Agent.

Bibliographical references

Andrews, Avery. 1985. "The major functions of the noun phrase". *Language typology and syntactic description: Clause structure*, ed. Timothy Shopen, 62–154. Cambridge: Cambridge University Press.
Chafe, Wallace L. 1970. *Meaning and the structure of language*. Chicago: University of Chicago Press.
Cook, Walter A. 1971. "Improvements in case grammar". *Language and linguistics: Working papers*, II, 10–22. Washington, D.C.: Georgetown University.
Fillmore, Charles. 1968. "The case for case". *Universals in linguistic theory*, ed. Emmon W. Bach and Robert T. Harms, 1–88. New York: Holt, Rinehart and Winston.
——. 1970. "Subjects, speakers and roles". *Ohio State University Working Papers in Linguistics* 4.31–63.
Givón, Talmy. 2001. *Syntax: An introduction*. 2 vols. Amsterdam and Philadelphia: J. Benjamins.
Gruber, Jeffrey S. 1967. "Look and see". *Language* 43. 937–947.
Ḥasan, A. 1963. *an-Naḥw al-wāfī*. 2 vols. Cairo: Dār al-Maʿrifa.
Jackendoff, Ray S. 1972. *Semantic interpretation in generative grammar*. Cambridge, Mass.: MIT Press.
Kearns, Kate. 2000. *Semantics*. New York: Palgrave.
Lakoff, George. 1968. "Instrumental adverbs and the concept of deep structure". *Foundations of Language* 4.4–29.
——. 1977. "Linguistic gestalts". *Papers from the Thirteenth Regional Meeting, Chicago Linguistics Society*, 236–287. Chicago: Chicago Linguistic Society.
Lyons, J. 1977. *Semantics*. 2 vols. Cambridge: Cambridge University Press.
Maxzūmī, Mahdī al-. 1964. *Fī n-naḥw al-ʿarabī: Naqdun wa-tawjīh*. Beirut: al-Maktaba al-ʿAṣriyya.
Platt, John T. 1971. *Grammatical form and grammatical meaning: A tagmemic view of Fillmore's deep structure case concepts*. Amsterdam: North-Holland.
Tesnière, Lucien. 1965. *Eléments de syntaxe structurale*. Paris: Klincksieck.
Wierzbicka, Anna. 2002. "Semantic primes and linguistic typology". *Meaning and universal grammar: Theory and empirical findings*, ed. Cliff Goddard and Anna Wierzbicka, 257–300. Amsterdam and Philadelphia: J. Benjamins.

Hussein Abdul-Raof (University of Leeds)

Agreement

Agreement is a relational feature obtaining between members of different phrases and clauses. Within the Standard Arabic verbal paradigm, for example, verbs agree with subjects in three features: gender, person, and number. This is illustrated in sentences (1), (2), and (3).

(1) a. *nām-a* *al-walad-u*
slept-3.s.m the-boy-nom
'The boy slept'

b. *nām-**at*** *al-bint-u*
slept-3.s.f the-girl-nom
'The girl slept'

(2) a. *ʾana* *nim-tu*
I.s.m/f slept-1s.m/f
'I slept'

b. *ʾan**ta*** *nim-**ta***
you.s.m slept-2.s.m
'You slept'

c. *huwa* *nām-a*
he.s.m slept-3s.m
'He slept'

(3) a. *ʾan**ta*** *nim-**ta***
you.s.m slept-2.s.m
'You slept'

b. *ʾan**tumā*** *nim-**tumā***
you.d.m/f slept-2.d.m/f
'You slept'

c. *ʾan**tum*** *nim-**tum***
you.p.m slept-2.p.m
'You slept'

The example in (1) shows different agreement morphology on the verb according to whether the subject is masculine or feminine. While the verb in (1a) shows a masculine suffix /-a/, (1b) exhibits a feminine suffix /-at/. In (2), the different suffixal morphology on the verbs is due to the change of person: 1st person in (2a), 2nd person in (2b), and 3rd person in (2c). As for the verbs in (3a), (3b), and (3c), they clearly show different suffixes that are sensitive to the number feature. Thus, the suffix /-ta/ in (3a) marks the singular feature, /-tumā/ in (3b) marks the dual feature, and /tum/ in (3c) marks the plural feature.

The Arabic verbal agreement paradigms in perfect and imperfect forms are summarized in Tables 1 and 2.

Table 1. Morphophonemic shapes of the Perfect agreement morphemes

	Singular		Dual		Plural	
	Masculine	Feminine	Masculine	Feminine	Masculine	Feminine
3rd pers.	/-a/	/-at/	/-ā/	/-atā/	/-u/	/-na/
2nd pers.	/-ta/	/-ti/	/-tumā/	/-tumā/	/-tum/	/-tunna/
1st pers.	/-tu/	/-tu/	/-nā/	/-nā/	/-nā/	/-nā/

Table 2. Morphophonemic shapes of the Imperfect agreement morphemes

	Singular		Dual		Plural	
	Masculine	Feminine	Masculine	Feminine	Masculine	Feminine
3rd pers.	/y-/_/-u/	/t-/_/u/	/y-/_/-āni/	/t-/_/-āni/	/y-/_/-ūna/	/t-/_/-na/
2nd pers.	/t-/_/-u/	/t-/_/ina/	/t-/_/- āni/	/t-/_/-āni/	/t-/_/-ūna /	/t-/_/-na /
1st pers.	/ʾ-/_/-u/	/ʔ-/_/-u/	/n-/_/-u/	/n-/_/-u/	/n-/_/-u/	/n-/_/-u/

While agreement morphology within the perfect paradigm is prefixal, the imperfect paradigm combines prefixes and suffixes resulting in discontinuous morphemes, as shown above and illustrated below in (4), (5), and (6) for person, gender, and number respectively.

(4) a. ***ʾana*** ***ʾ-adrus-u***
I.s.m/f 1.m/f-study-s
'I am studying'

b. *ʾanta* ***t****-adrus-****u***
you.s.m 2.m-study-s
'You are studying'

c. *huwa* ***y****-adrus-****u***
he.s.m 3.m-study-s
'He is studying'

(5) a. *ʾanta* ***t****-adrus-****u***
you.s.m 2.m-study-s
'You are studying [masc.]'

b. *ʾanti* ***t****-adrus-****īna***
you.s.f 2.m-study-f
'You are studying [fem.]'

(6) a. *huwa* *y-adrus-****u***
he.s.m 3.m-study-**s**
'He is studying'

b. *humā* *y-adrus-****āni***
they.d.m 3.m-study-**d**
'They are studying'

c. *hum* *y-adrus-****ūna***
they.p.m 3.m-study-**p**
'They are studying'

The generalizations in (7a), (7b), and (7c) below clearly summarize the distribution of nominal agreement morphology on the verb.

(7) a. person morphology (1st, 2nd, 3rd) is constantly encoded in the prefix;
b. number morphology (singular, dual, plural) is encoded in the suffix, except for the 1st person;
c. gender morphology (masculine, feminine) appears on the suffix in the plural and on the prefix in the singular, except for the 1st person;

Following this brief summary of subject agreement morphology as it is spelled out on verbs, the remaining part of this entry focuses on a number of agreement discrepancies. First, it examines the contexts under which these agreement inconsistencies are achieved; second, it highlights the major analyses; finally, it provides an analysis which accounts for the observed phenomena and extends to similar agreement discrepancies within the Arabic morphological system.

Subject agreement morphology on the verb is sensitive to the subject position in the sentence. As such, if the subject precedes the verb, all agreement morphemes (person, gender, and number) are realized on the verb. If the subject follows the verb, person and gender are realized, while number agreement is not observed. This asymmetry between preverbal and postverbal subjects in Arabic is illustrated by the contrast between (8a) and (8b).

(8) a. _nām-a_ (/*_nām-ū_) _l-ʾawlādu_ (VS)
slept-3.**m.s** (slept-3.**m.p**) the-boy-nom
'The boys slept'

b. _al-ʾawlād-u_ _nām-**ū**_ (*_nām-**a**_) (SV)
the-boys-nom slept-3.**m.p** (slept-3.**m.s**)
'The boys slept'

In (8a) the verb _nām-a_ 'slept' carries singular morphology, as indicated by the suffix /-a/, while the subject _al-ʾawlādu_ 'the boys' is plural as indicated by the plural form of the noun (_walad_ [sg.], _ʾawlād_ [pl.]). In (8b), however, the verb _nām-**ū**_ 'slept' carries plural morphology, as indicated by the suffix /-ū/, whereas the subject _al-ʾawlādu_ 'the boys' is in the same plural form. This asymmetry is further supported by the ungrammaticality of the plural form of the verb in (8a) and singular form of the verb in (8b). In short, Arabic verbs agree in number with preverbal but not postverbal subjects.

The examples in (8) with a masculine plural subject and the ones in (9) with a feminine dual show that postverbal subjects, as well as preverbal subjects, control agreement on the verb with respect to the features of gender and person.

(9) a. _nām-**at**_ (*_nām-atā_) _al-bint-āni_ (VS)
slept-3.**f.s** (slept-3.**f.d**) the-girl(**f**)-3.d
'The two girls slept'

b. _al-bint-āni_ _nām-atā_ (*_nām-**at**_) (SV)
the-girl(**f**)-3.d slept-3.**f.d** (slept-3.**f.s**)
'The two girls slept'

The above agreement asymmetry whereby subject verb agreement obtains in the SV order with all pronominal features of gender, number, and person, while partial agreement, i.e. gender and person only, obtains in VS order has been discussed and analyzed within the investigation of the syntax of clauses and functional categories in Universal Grammar (UG) in an extensive literature which cannot be reviewed here; see, for example, Fassi Fehri (1982, 1988, 1993); Abd El Moneim (1989); Mohammad (1990, 1999); Benmamoun (1990, 1992, 2000); Eid (1991); Bahloul and Harbert (1993); Bahloul (1994); Harbert and Bahloul (2002); Aoun, Benmamoun, and Sportiche (1994); and Parkinson (1995). Agreement asymmetry is quite common in a number of languages and language groups. For Celtic languages, see McCloskey (1986), Rouveret (1991), and Bahloul and Harbert (1993); for North Italian Dialects, see Brandi and Cordin (1989); for Russian and other Slavic languages, see Corbett (1983). In Arabic and in other languages with similar verbal agreement alternations, the essential question is how to account for such asymmetry. In other words, each analysis has to explain how each agreement pattern is licensed. The following section briefly highlights the contributions of major analyses which have provided an answer to this question.

1. THE EXPLETIVE ANALYSIS

Fassi Fehri (1982, 1988) and Mohammad (1990, 1999) provide an answer to the question of asymmetry by contending that agreement patterns differently in VSO and SVO sentences because the latter is licensed by a different element in each case. While in SVO cases, full agreement obtains through specifier-head agreement with the thematic subject, partial agreement in VSO cases is the result of an agreement with a preverbal null expletive pronoun. This analysis is motivated by the observation that Arabic VSO structures may be preceded by expletive topics as shown in (10).

(10) _ʾinna-**hu**_ _zāra-nī_ _ṯalāṯu_ _šāʿirāt-in_
that-**it** visited-me three-nom poets.f.-gen
(Fassi Fehri 1993:39)
'It visited me three poets' = 'Three poets visited me'

Under the Expletive Analysis, the singular agreement feature on the verb _zāranī_ is licensed by the 3rd person singular pronominal form /_-hu_/ on the complementizer _ʾinna_ 'that'. Fassi Fehri (1988) postulates an empty expletive pronoun (_pro_) in order to license the 3rd person singular agreement features on the verb; Mohammad (1990:98) claims that "VSO sentences in Arabic contain two subjects: the 'real' subject and an expletive subject . . . that dictates the agreement features on the verb". It is later observed that while the expletive has inherent singular number, its gender features are variable, and it can "change . . . to feminine if followed by a feminine subject" (Mohammad 1999:144). It is not clear, however, that this analysis is descriptively adequate. In a much more recent work, Fassi Fehri (1993:39) observes that this analysis "arbitrarily limits the list of expletives to the one needed". Benmamoun (2000:125) further notes the lack of "any independent motivation for such an expletive". Harbert and Bahloul (2002: 65–66) question the validity of this approach on the basis of its theoretical and practical limitations, since it fails to answer a number of fundamental questions.

2. Incorporation and merger analysis

Another account which has been proposed to explain the Arabic agreement asymmetry claims that the differences between full and partial agreement are the products of fundamentally different processes. On the basis of examples such as (11) and (12), Fassi Fehri (1993:111) contends that full agreement is achieved through pronoun incorporation.

(11) a. *jāʾ-**ū***
came.3.**pl**.m
'They [masc.] came'

b. *jiʾ-**na***
came.3.**pl**.f
'They [fem.] came'

(12) a. *jāʾ-**a*** *l-ʾawlad-u*
came-**s.m** the-boys-nom
'The boys came'

b. *jāʾ-at* *al-banāt-u*
came-f the-girls-nom
'The girls came'

As the examples in (11a) and (11b) suggest, full agreement obtains with null pronominals, while partial agreement is obtained with postverbal full lexical subjects. The pronouns are null because they have been incorporated into the verb. The gist of the merger analysis, on the other hand, is that it is partial agreement which involves incorporation, not of pronouns, but of postverbal lexically headed subject noun phrases (NPs). To provide further support, Benmamoun observes that it is not always the case that morphosyntactic features are realized through affixes by showing that such features may be spelled out periphrastically through independent words, as in the case of tense being realized through auxiliaries or modals. Along the same line of argument, Benmamoun (2000:130) claims that in the VS cases where only partial agreement is observed, the number feature on the verb is not spelled out by an affix, but rather by the lexical subject which merges with the verb. In other words, "since the subject is inherently specified for number features, its merger with the verb amounts to spelling-out those features on the latter, thus making the number suffix redundant". In short, Benmamoun's approach offers an alternative analysis by claiming that number agreement in Arabic is spelled out in two different ways: as a single word through affixation or periphrastically through the merger of the verb and the postverbal subject.

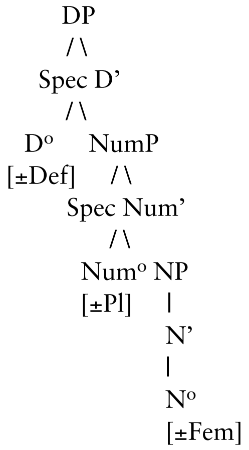
Different as the incorporation account of Fassi Fehri and the merger account of Benmamoun may be, they converge on one crucial prediction: there should never be full agreement with overt postverbal subjects. In Fassi Fehri's account, this follows because full agreement is an incorporated subject, and therefore should never cooccur with an overt subject. In Benmamoun's account, it follows because postverbal subjects are obligatorily merged with the verb, thus precluding the independent realization of number agreement morphology.

3. The government-agreement analysis

A more promising analysis has been advanced in Bahloul and Habert (1993) and Harbert and Bahloul (2002). Under this analysis, labeled the Government-Agreement (GA) approach, it is argued that SVO and VSO sentences exhibit different agreement patterns because agreement is effected under different syntactic conditions. In the former, it encodes a Spec-Head relationship between the subject and the agreeing verb, while in the latter it is obtained under a relationship of government between those elements. Thus, the difference in agreement morphology results because different feature sets are accessible under these two types of agreement. Under this account, Determiner Phrases (DPs) headed by lexical nouns are represented as in (13).

Agreement features may be inherent or derived. Gender features are inherent features of lexical nouns since they are associated with invariant gender distinctions. Definiteness, however, is not an inherent feature of lexical head nouns, but rather of the functional category 'Determiner' and therefore of DPs. In (13) gender features, being inherent features of nouns, are associated with the lowest layer of projection

(13)



(the Base of the projection system, in terms of Bahloul and Harbert 1993). Number and Definiteness features, on the other hand, are associated with higher level functional heads in the projection system. Accordingly, only features of the Base of a projection system are visible for purposes of agreement under government, as specified by the principle in (14), while all features are visible for Spec-Head agreement.

(14) The Government-Agreement constraint: only those features that originate on the Base of a Projection System are evaluated for agreement under government.

This yields the observed asymmetry between gender and number agreement in Arabic. Gender features originate on lexical Ns, hence on the Bases of Projection Systems. Thus, they are visible for agreement under government. Number features are features of the functional head Num°. They are not features of the Base node in Projection Systems terminating in lexical nouns and are therefore not visible to such agreement.

4. AGREEMENT IN CONJOINED SUBJECTS

Another advantage of the Government-Agreement analysis is that it readily explains the agreement contrast between postverbal and preverbal conjoined subjects as illustrated in (15).

(15) a. *[al-walad-u wa l-bint-u] xaraj-ā*
the-boy-nom and the-girl-nom left-**d.m**
'The boy and the girl left'
b. *[al-bint-u wa l-walad-u] xaraj-ā*
the-girl-nom and the-boy-nom left-**d.m**
c. *xaraj-**at** [al-bint-u wa l-walad-u]*
left-**3.s.f** the-girl-nom and the-boy-nom
d. *xaraj-**a** [al-walad-u wa l-bint-u]*
left-**3.s.m** the-boy-nom and the-girl-nom

In (15a), we see that a conjoined subject consisting of a masculine conjunct and a feminine conjunct triggers (dual) masculine agreement on a verb in SV order. A masculine NP conjoined with a feminine NP yields a coordinate structure with the composite gender feature [+Masc], and this feature controls agreement on the verb. The example in (15b) shows that the order of the conjuncts does not affect the outcome of agreement. For the purpose of agreement, it does not matter whether the first conjunct is masculine and the second feminine, as in (15a), or that the first is feminine and the second masculine, as in (15b). In both cases the controlling gender agreement remains the same, namely, masculine. The contrast between (15a) and (15b) shows that gender agreement is not affected by the relative order of the conjuncts. The verb in SV order agrees with the composite masculine gender feature of the topmost node.

In VS order, however, a very different pattern emerges. The verb is singular with postverbal subjects, as expected, but its gender inflection is determined by the closest conjunct. In (15c) it is feminine since the feminine conjunct comes first. In (15d) it is masculine since the masculine conjunct is first. The same pattern obtains with person agreement. The examples in (16) illustrate agreement with the first conjunct in postverbal conjoined subject constructions.

(16) a. *xaraj-**ti*** *[ʾanti* *wa* *huwa]*
left-2.**s.f** you.**s.f** and he.3.s.m
'You and he left'
b. *xaraj-**ta*** *[ʾanta* *wa* *hiya]*
left-2.**s.m** you.**s.m** and she.3.s.f
'You and he left'

The example in (16) is given the representation in (17) whereby the conjunction heads a Conjunction Phrase (ConjP), and the first conjunct is its Specifier.

As Benmamoun has noted, under the assumption of an appropriate definition of minimality government, INFL in configuration (17) governs the Specifier of the Conjunction Phrase, since that Spec is not more locally o-commanded by any other head. Accordingly, INFL may agree in

```
(17)      INFL'
            /    \
        INFL          PredP
        V                /    \
             Spec Pred           Pred'
        ConjP[+MASC]          /  \
              /     \            Pred  VP
      SpecConj   Conj'
              /     |     \
           DP   Conjº DP
             |        |      |
           NP      |    NP
          [+Fem]  and [+Masc]
```

gender with the DP in Spec Conj position, through agreement under government. It may not agree with the second conjunct, the complement of Conjº, since the complement is in the domain of a more local head. To ensure the impossibility of feminine agreement in (15d), certain other well-motivated assumptions are required. First, if conjunctions are heads, it appears that they must be functional, rather than lexical heads. They are closed class elements, not belonging to a major word class, and they lack descriptive content. If so, then the NP complement of Conjº in (17) constitutes the Base of a Projection System extending upward to Conjunction Phrase. Given that features of Bases of Projection Systems are projected freely upward through those systems, a question is raised as to why the feminine feature of the noun in the complement phrase in (15d), for example, is not available for agreement. Harbert and Bahloul (2002) propose that the projection of these features is disrupted in such cases because the head of the Conjunction Phrase is simultaneously assigned a second abstract gender feature through Spec-Head agreement with the DP in the SpecConj position. Because of these competing features, one from the Base of the Projection System and one from the Specifier of the phrase, the Conjunction Phrase receives a composite gender feature, counting for purposes of agreement as Masculine. This composite gender feature is what determines masculine agreement in the Spec-Head agreement cases in (15a–b).

The three analyses of agreement presented above – the Expletive Analysis, the Incorporation and Merger Analysis, and the Government-Agreement Analysis – provide different accounts for Arabic agreement asymmetry. The Government-Agreement Analysis, however, provides a unified account for the agreement asymmetry and related issues such as agreement with conjoined subjects.

BIBLIOGRAPHICAL REFERENCES

Abd El-Moneim, Aliaa. 1989. *The role of INFL.* Ph.D. diss., University of Connecticut.

Aoun, Joseph, Elabbas Benmamoun, and Dominique Sportiche. 1994. "Agreement, conjunction and word order in some varieties of Arabic". *Linguistic Inquiry* 25.195–220.

Bahloul, Maher. 1994. *The syntax and semantics of taxis, aspect, tense and modality in Standard Arabic.* Cornell University: Department of Modern Languages and Linguistics Publications.

—— and Wayne Harbert. 1993. "Agreement asymmetries in Arabic". *The eleventh West Coast conference on formal linguistics*, ed. Jonathan Mead, 15–31. Stanford: Center for the Study of Language and Information.

Benmamoun, Elabbas. 1990. "Inflectional morphology: Problems of derivation and projection". Ms., University of Southern California.

——. 1992. *Functional categories: Problems of projection, representation and derivation.* Ph.D. diss., University of Southern California.

——. 2000. *The feature structure of functional categories: A comparative study of Arabic dialects.* New York: Oxford University Press.

Eid, Mushira. 1991. "Verbless sentences in Arabic and Hebrew". *Perspectives on Arabic linguistics*, III, ed. Bernard Comrie and Mushira Eid, 31–61. Amsterdam and Philadelphia: J. Benjamins.

Fassi Fehri, Abdelkader. 1982. *Linguistique arabe: Forme et interprétation.* Rabat, Morocco: Publications de la Faculté des Lettres et Sciences Humaines.

——. 1988. "Agreement, binding and coherence". *Agreement in natural language: Approaches, theories, descriptions*, ed. Michael Barlow and Charles A. Ferguson, 107–158. Stanford: Center for the Study of Language and Information.

——. 1993. *Issues in the structure of Arabic clauses and words.* Dordrecht: Kluwer.

Harbert, Wayne and Maher Bahloul. 2002. "Postverbal subjects and the theory of agreement". *Themes in Arabic and Hebrew syntax*, ed. Jamal Ouhalla and Ur Shlonsky, 45–70. Dordrecht: Kluwer Academic Publishers.

Mohammad, Mohammad A. 1990. "The problem of subject-verb agreement in Arabic: Towards a solution". *Perspectives on Arabic linguistics*, I, ed. Mushira Eid, 95–125. Amsterdam and Philadelphia: J. Benjamins.

——. 1999. *Word order, agreement and pronominalization in standard and Palestinian Arabic.* Amsterdam and Philadelphia: J. Benjamins.

Parkinson, David. 1995. "Iterated subject agreement in standard Arabic". *Cornell Working Papers in Linguistics* 13.121–150.

MAHER BAHLOUL (American University of Sharjah)

ʿĀʾid

When dealing with anaphoric reference, the Medieval Arab grammarians normally use the verb *ʿāda* (*yaʿūdu ʾilā* 'referring to'). The term *ʿāʾid* (occasionally *rābiṭ*, *rājiʿ*) is typically used to denote the resumptive pronoun in either a clausal predicate (*jumla xabar*) or a relative clause (*jumla ṣila/ṣifa*). The underlying principle in both cases is that the basic form of the predicate and the attribute is a phrase referring back to the *mubtadaʾ* (the subject in a non-VS sentence, sometimes translated as 'topic', a pragmatic term that does not cover all instances of *mubtadaʾ*) or head (*manʿūt*) respectively. In other words, there must be a semantic relation of identity between the *xabar/ṣifa* (→ *sabab*) and the *mubtadaʾ/manʿūt* respectively. A clause, in contrast to a phrase, is by definition a non-referential element. Thus, in order to establish a predicative or an attributive relationship between a clause and a *mubtadaʾ/* head, some referential linking element should be entered into the clause, usually a personal pronoun referring back to the *mubtadaʾ*/head. Ibn Yaʿīš (d. 642/1245) explains, when dealing with the types of clausal *xabar* (*Šarḥ* I, 88–89), that the *jumla* has the value of an independent sentence (*kalām mustaqill*) and therefore cannot qualify as a *xabar*, unless it contains some referential element (*ḏikr*) to link it to the *mubtadaʾ*. Absence of such an element would render the clause syntactically disconnected (*ʾajnabiyya*) from its *mubtadaʾ*, and the sentence as a whole unacceptable. Thus, while *zaydun qāma ʾabūhu* 'Zayd, his father stood' is a perfectly grammatical sentence, with *-hu* in *ʾabūhu* functioning as *ʿāʾid*, *zaydun qāma ʿamrun* 'Zayd, ʿAmr stood' is disallowed because of the lack of an *ʿāʾid*. The principle requiring an *ʿāʾid* in a clausal *xabar* applies not only to verbal and nominal clauses but also to cases where the *xabar* position is filled by a conditional sentence or by what some of the later grammarians refer to as *jumla ḏ̣arfiyya*. Regarding the former, Ibn Yaʿīš (*Šarḥ* I, 89) indicates that the *ʿāʾid* may occur either in the protasis (*šarṭ*), as in *zaydun ʾin tukrim-hu yaškur-ka ʿamrun* 'Zayd, if you honor him, ʿAmr will thank you', or both in the protasis and the apodosis (*jazāʾ*), as in *zaydun ʾin yaqum ʾukrim-hu* 'Zayd, if he stands up, I will honor him'. As for cases such as *zaydun fī d-dār* 'Zayd is in the house' (*jumla ḏ̣arfiyya*), Ibn Yaʿīš (*Šarḥ* I, 91) indicates that this sentence is paraphrasable by *zaydun yastaqirru fī d-dār*, where the deleted *yastaqirru* 'he is located' contains the *ʿāʾid*.

The grammarians cite other cases where the *ʿāʾid* is absent from the clausal *xabar*. They emphasize that such a deletion occurs only in cases where it is easily retrievable from the context. One typical example is the sentence *as-samnu manawāni bi-dirhamin* 'the butter, two *manan* for a dirham'. Ibn Yaʿīš (*Šarḥ* I, 91) explains that in such cases the *mubtadaʾ* is a generic noun designating a kind of commodity, and that the sentence as a whole is de-signed for pricing that commodity. The addressee can thus easily relate this sentence to the underlying (*muqaddar*) full version: *as-samnu manawāni* ***min-hu*** *bi-dirhamin*, where *-hu* in *min-hu* is the *ʿāʾid* referring to *as-samn* (for further examples and details, see, e.g., Ibn Yaʿīš, *Šarḥ* I, 91–92).

A well-known Kufan claim is that the *ʿāʾid* in the clausal *xabar* is the assigner of the *rafʿ* case to the *mubtadaʾ*. Dealing with the verse *kullun qad ʿalima ṣalāta-hu wa-tasbīḥa-hu* 'He knows the prayers and praise of each' (Q. 24/41), al-Farrāʾ (*Maʿānī* II, 255) states that *kullun* is assigned the *rafʿ* case by the pronoun *-hu* referring back to it in *ṣalāta-hu wa-tasbīḥa-hu*. For extensive citations of passages from al-Farrāʾ's *Maʿānī l-Qurʾān*, dealing with anaphoric reference, see Kinberg (1996:529–532).

Ibn Hišām (d. 761/1360) (*Muġnī* II, 647–663) provides an extensive discussion of various kinds of linking elements (*rawābiṭ*) in various kinds of constructions. He starts with cases where the *rābiṭ* occurs in a clause standing in predicative relationship to some nominal. He points out that the personal pronoun (*ḍamīr*) is the basic but not the only form of the *rābiṭ*. Among the other devices he outlines, the most common are:

i. A demonstrative pronoun, as in *wa-llaḏīna kaḏḏabū bi-ʾāyāti-nā wa-stakbarū ʿan-hā ʾulāʾika ʾaṣḥābu n-nār* 'those who deny and display arrogance toward Our signs will be in Hell' (Q. 7/36);
ii. Repetition of the actual nominal functioning as *mubtadaʾ* as in *wa-ʾaṣḥābu l-yamīn mā ʾaṣḥābu l-yamīn* 'those on the right hand: what [awaits] those on the right hand?' (Q. 56/27). Ibn Hišām (*Muġnī* II, 553) indicates

that this strategy is normally used for frightening and aggrandizement (*at-tahwīl wa-t-tafxīm*);

iii. An agnomen (*kunya*) referring to the *mubtada*ʾ, as in *zaydun jāʾa-nī ʾabū ʿabdi llāhi* ‘Zayd, ʾAbū ʿAbdallāh came to me’, where ʾAbū ʿAbdallāh is the *kunya* of Zayd; and

iv. The *rābiṭ* may be realized as a general term inclusive of the referent of the *mubtada*ʾ, as in *zaydun niʿma r-rajul* ‘Zayd, what a nice person’.

Among other constructions requiring a *rābiṭ*, Ibn Hišām enumerates the following:

i. The relative clause, whether syndetic (*ṣila*) or asyndetic (*ṣifa*). This is exemplified by the following Qurʾānic verses respectively: [*ḏālika l-kitābu* . . .] *hudan li-l-muttaqīna llaḏīna yuʾminūna bi-l-ġayb* ‘[this Book . . .] is a guide for the godfearing who believe in the Unseen’ (Q. 2/3), *ḥattā tunazzila ʿalaynā kitāban naqraʾu-hu* ‘till you send down for us a book which we can read’ (Q. 17/93). Ibn Hišām (*Muġnī* II, 653–656) indicates that the *rābiṭ* in relative clauses is normally a personal pronoun which may, however, be deleted. Deletion in these cases is more likely in a syndetic than in an asyndetic clause, and it is least likely in a *xabar* clause.

ii. The circumstantial clause (*jumla ḥāl*). In this case the function of the *rābiṭ* may be implemented either by the particle *wa-* or by a personal pronoun, or by both. The latter is exemplified by: *lā taqrubū ṣ-ṣalāta wa-ʾantum sukārā* ‘do not get near to prayer when you are drunk’ (Q. 4/43).

iii. The *ištiġāl* construction (a term denoting an accusatival topic, for which an underlying verbal operator must be posited), as in *zaydan ḍarabtu-hu* ‘Zayd, I hit him’. For further details, see Ibn Hišām (*Muġnī* II, 657–658).

iv. *Maʿmūl aṣ-ṣifa l-mušabbaha*. Here Ibn Hišām (*Muġnī* II, 658–659) deals with constructions analogous to *naʿt sababī*. The *ʿāʾid* in these cases is either morphologically realized or, otherwise, assumed (*muqaddar*). Ibn Hišām cites the verse *ʾinna li-l-muttaqīna la-ḥusna maʾābin jannāti ʿadnin mufattaḥatan la-humu l-ʾabwābu* ‘for the righteous there is a blessed resort, the gardens of Eden, where the gates will be open for them’ (Q 38/49–50). He argues that underlying (*al-ʾaṣl*) *al-ʾabwābu* is either *al-ʾabwābu min-hā* or *ʾabwābu-hā*. In other words, the definite article in this case substitutes for the *ʿāʾid*.

Bibliographical references

Primary sources

Farrāʾ, *Maʿānī* = ʾAbū Zakariyyāʾ Yaḥyā ibn Ziyād al-Farrāʾ, *Maʿānī l-Qurʾān*. Ed. ʾAḥmad Yūsuf Najātī and Muḥammad ʿAlī an-Najjār. 3 vols. Cairo, 1955–1972.

Ibn Hišām, *Muġnī* = Jamāl ad-Dīn ʾAbū Muḥammad ʿAbdallāh ibn Yūsuf Ibn Hišām al-ʾAnṣārī, *Muġnī l-labīb ʿan kutub al-ʾaʿārīb*. Ed. Māzin al-Mubārak and Muḥammad ʿAlī Ḥamdallāh. Beirut: Dār al-Fikr, 1985.

Ibn Yaʿīš, *Šarḥ* = Muwaffaq ad-Dīn Yaʿīš ibn ʿAlī Ibn Yaʿīš, *Šarḥ al-Mufaṣṣal*. 10 vols. Cairo: Maktabat al-Mutanabbī, n.d.

Secondary source

Kinberg, Naphtali. 1996. *A lexicon of al-Farrāʾ's terminology in his* Qurʾān *commentary with full definitions, English summaries, and extensive citations*. Leiden: E.J. Brill.

Yishai Peled (Tel Aviv University)

ʿAjamī → Faṣīḥ

Aktionsart

The notion of Aktionsart has arisen as a counter-notion to the concept of → aspect. Aspect refers to temporal structures. This notion has been developed in studies on Arabic (Maas p.c.), where a dichotomy exists reflecting a boundedness distinction (a state of affairs is considered ‘bounded’ if it is terminated). Structures that are not connected to this distinction are described with the term Aktionsart. However, there has been a lot of confusion about the use of the notion of aspect. Other languages, too, have dichotomous or partly dichotomous verb systems, but some of them convey distinctions different from the Arabic system (e.g. the Slavic). The notion of aspect has often been applied to these systems as well; the Slavic tradition has been the most influential. Hence, the delimitation of Aktionsart against aspect is often obscured, leading to different definitions of Aktionsart. Their common ground is the differentiation of the verbal lexicon and the opposition to aspect.

Many scholars have done without the distinction between aspect and Aktionsart, preferring to subsume both concepts under either aspect or aspectuality (Sasse 2002).

Nowadays, Aktionsart is generally regarded as acting on whole sentences (see, e.g., Verkuyl 1972). However, this hinders cross-linguistic comparison. As a notion referring to verbs, there are two approaches for defining Aktionsart (Steinitz 1981), an abstract one and a morphology-based one. Under the latter assumption, all verbs count as Aktionsart that are morphologically derived from a certain basic verb. Russian, for instance, has an elaborate derivational system. Isačenko (1968) gives a list of possible derivations, contrasting a number of simple verbs with their derivations: *pet'* 'to sing' vs. *zapet'* 'to begin a song', *sidet'* 'to sit' vs. *posidet'* 'to sit for a while' vs. *siživat'* 'to sit now and then', *kusit'* 'to bite' vs. *perekusit'* 'to bite each member of a set (in turn)', etc. Many scholars have this system in mind when they investigate Aktionsart. As a result, interactions with arguments, intensifiers, etc. are often subsumed under Aktionsart.

The verbal systems of the Arabic languages – except for Standard Arabic, these are usually called 'dialects' – are relatively uniform with respect to Aktionsart, so that examples will be taken indifferently from any Arabic language. Unless indicated otherwise, examples are from Standard Arabic. Minor differences exist, though (see below).

Blohm (1990) follows the morphology-based approach to Aktionsart in Arabic, like most of the authors he cites. He defines the so-called verbal forms or stems as the derivational system. Unlike Russian, however, in Arabic, the semantic connections between the basic and the derived forms are seldom transparent. For this reason, it seems unwise to view these derivations as an Aktionsart system. The verbal forms are morphologically connected, but can no longer be viewed as being semantically connected. As these forms are the only derivational means for verbs in Arabic, a morphology-based approach to Aktionsart cannot be applied. In Arabic, there is only one verb for each contrast in Isačenko's list.

Whereas some scholars define Aktionsart as adding meaning to a basic concept represented by a verb (mostly proponents of a morphology-based approach), others ascribe any verb to a certain Aktionsart (predominantly proponents of the more abstract approach). Aktionsart can thus be described either as a classification of verbs or as a specification task for states of affairs. Under the latter view, the most prominent temporal task of Aktionsart is to pick out subintervals of states of affairs (including such that are longer than the interval itself). Arabic specifies these subintervals mostly by analytical constructions or outside the predicate. The constructions may involve complex predicates, in Arabic called coverbs (Maas 1995) → or auxiliaries. Alternatively, a state of affairs may be expressed by a noun, which then functions as the object to a verb whose only task is to provide the information that a subinterval of the state of affairs is in the focus of attention:

(1) Tunisian *bdī-t* *na-fham*
start:PFV-1S 1S-understand
'j'ai commencé à comprendre' (Simeone-Senelle 1985)

(2) Tunisian *bēdi* *fi l-xidma*
start:AP in ART-work
'je commence le travail' (Simeone-Senelle 1985)

(3) Moroccan *kemmel* *l-makla*
complete:PFV:3Sm ART-meal
'he finished his meal'

The main verbs in coverb constructions are not always verbs which only serve for specifying Aktionsart. Compare the following durative construction:

(4) Tunisian *gāˁad* *ye-rǧaˁ*
sit:AP 3Sm-return:IMPV
'il est en train de revenir (on le voit en train d'arriver)' (Simeone-Senelle 1985)

Arabic varieties differ in the extent to which they make use of complex predicates (cf. Youssi 1992; Mitchell and al-Hassan 1994); other kinds of predicates containing more than one verb exist, too. Standard Arabic is at the lower end of this 'verb serialization scale' (→ serial verbs). Examples like (4) are therefore impossible in Standard Arabic.

Normally, an Arabic verb comprises more subintervals of a state of affairs, compared to European languages. The famous two-phase verbs of Arabic are a good example of this tendency: *jalasa* 'to sit/to sit down', *waqaḍa* 'to wake up/to be awake', *waqafa* 'to stand up/to stand'. In two-phase verbs, two different stages of a state of affairs may be referred to by the same aspectual form.

Many modern investigations viewing Aktionsart as a classification of verbs are based on Vendler (1967). He distinguishes between stative verbs like English *to know*, activity verbs like *to play*, (punctual) achievements like *to explode*, and accomplishments like *to open*. Achievements and accomplishments share the feature of being oriented towards a goal; this feature is called 'telicity'.

States are often expressed by nominal predicates in Arabic, more often than in English:

(5) *al-kursī* *wazn-u-hu*
ART-chair weight-NOM-3Sm
ʿišrūn kīlo
twenty kilogram
'the chair weighs twenty kilograms'

(6) *ʿumr-u-hu* *ṯalaṯūna* *sana*
age-NOM-3Sm thirty:NOM year
'he's thirty years old'

In Standard Arabic, such nominal predicates may be → participles. In some modern dialects, the participles have the same function, but have ceased to be a nominal form; they may be regarded as stative verb forms (Reese *forthcoming*), which, however, are used less often than stative verbs in European languages.

Arabic lacks the feature known from English that some verbs may be viewed as either atelic or telic (*active accomplishments* like eating, running in Van Valin and LaPolla 1997).

(7) *ʾakala* *l-ʾakl*
eat:PFV:3Sm ART-meal
'he ate the meal'

(8) *ʾakala* *l-ʾakl* *kull-a-hu*
eat:PFV:3Sm ART-meal all-ACC-3Sm
'he ate the meal up'

(9) *ʾakala* *l-ʾakl,* *wa-lākin*
eat:PFV:3Sm ART-meal but
lā *kull-a-hu*
not all-ACC-3Sm
'he ate the meal, but not all of it'

Russian as the classic 'Aktionsart language' distinguishes two verbs for the two readings of the English *active accomplishments*.

In the examples, telicity is specified, if at all, not in the verbs, but outside the predicate or even the clause (though the telic interpretation may be regarded as default). The lack of telicity may be specified, too, again outside the predicate:

(10) *ʾakala* *min* *al-xubz*
eat:PFV:3Sm from ART-bread
'he ate some of the bread'

Unmarked telic markers for *active accomplishments* are found in nomadic (Caubet 1991) dialects; for a marked example, see (3) above.

(11) Mauritanian *vet-t* *tʿaššey-t*
complete:PFV-1S have.meal:PFV-1S
'j'ai déjà mangé' (Tauzin 1985)

On the other hand, Arabic has 'purely telic verbs', that is, verbs that only focus on the goal of an action, like *waṣala* 'to arrive'. Most motion verbs behave similarly (Arabic being a verb-oriented language; see Talmy 1985): they focus on the goal or path of a motion, its manner remaining unspecified, e.g. *daxala* 'to enter, to go in'.

This means that Arabic varieties add few Aktionsart specifications to verbs. In most cases, these are omitted. If they need to be specified, this specification is achieved either by additional verbs (pure telic verbs or special Aktionsart verbs like 'to begin') or by additional information outside the predicate.

Bibliographical references

Blohm, Dieter. 1990. "Gibt es Aktionsarten im Arabischen?". *Zeitschrift für Arabische Linguistik* 22.7–28.

Caubet, Dominique. 1991. "The active participle as a means to renew the aspectual system: A comparative study in several dialects of Arabic". *Semitic studies in honor of Wolf Leslau on the occasion of his eighty-fifth birthday, November 14th, 1991*, ed. Alan Kaye, I, 209–224. Wiesbaden: O. Harrassowitz.

Isačenko, Aleksandr. 1968. *Die russische Sprache der Gegenwart*, I, *Formenlehre*. 2nd ed. München: Hueber.

Maas, Utz. 1995. "Koverben im marokkanischen Arabisch in typologischer Perspektive". *Sprachtypologie und Universalienforschung* 48.239–274.

Mitchell, Terence F. and Shahir A. El-Hassan. 1994. *Modality, mood and aspect in spoken Arabic, with special reference to Egypt and the Levant*. London and New York: Kegan Paul International.

Reese, Johannes. Forthcoming. *Das Aktivpartizip als Prädikat im Marokkanischen*. Ph.D. diss., University of Osnabrück.

Sasse, Hans-Jürgen. 2002. "Recent activity in the theory of aspect: Accomplishments, achievements, or just non-progressive state?" *Linguistic Typology* 6.199–271.

Simeone-Senelle, Marie-Claude. 1985–1986. "Systèmes aspecto-temporels en arabe maghrébin (Tunisie)". *Matériaux Arabes et Sudarabiques* 63–81.

Steinitz, Renate. 1981. *Der Status der Kategorie "Aktionsart" in der Grammatik (oder: Gibt es Aktionsarten im Deutschen?).* (= *Linguistische Studien*: Reihe A, *Arbeitsberichte*, 76). Berlin: Akademie der Wissenschaften der DDR, Zentralinstitut für Sprachwissenschaft.

Talmy, Leonard. 1985. "Lexicalization patterns: Semantic structure in lexical forms". *Language typology and syntactic description.* III. *Grammatical categories and the lexicon*, ed. Timothy Shopen, 57–149. Cambridge: Cambridge University Press.

Tauzin, Aline. 1985–1986. "Systèmes aspecto-temporels en arabe maghrébin (Mauritanie)". *Matériaux Arabes et Sudarabiques* 81–97.

Van Valin, Robert and Randy LaPolla. 1997. *Syntax: Structure, meaning and function.* Cambridge: Cambridge University Press.

Vendler, Zeno. 1967. *Linguistics in philosophy.* Ithaca, N.Y.: Cornell University Press.

Verkuyl, Henk. 1972. *On the compositional nature of the aspects.* Dordrecht: Reidel.

Youssi, Abderrahim. 1992. *Grammaire et lexique de l'arabe marocain moderne.* Casablanca: Wallada.

Johannes Reese (University of Zürich)

Algeria

Research on Algerian Arabic may be divided into two main periods: the first began during the first years of the 20th century and ended some time after Algerian independance (1962), i.e. during the 1970s. The basic research on dialectal geography (Cantineau 1937, 1938, 1940, 1941) and the most comprehensive monographs on Algerian Arabic, particularly on some representative dialects (W. Marçais 1902, 1908; Ph. Marçais 1945, 1956, 1960a, 1960b, 1977; Grand'Henry 1972, 1976a) were produced during this first period. The second period began during the 1980s and has continued up to the present: it is mainly devoted to the study of modern Arabization, Algerian sociolinguistics, and to different aspects of the dialect of Algiers (→ Algiers Arabic).

1. Algerian Arabic dialect geography and some important Algerian dialects

It should be mentioned at the outset that the most comprehensive bibliography on Algerian Arabic is now found in Maïri (1987:96–107).

It is well known that Arabization of Algeria was achieved after two conquests. The first was a military conquest during the 7th century C.E. At this time, the cities (Tlemcen, Constantine) and their rural surroundings (villages and mountains) were partially Arabized and it is quite probable that many Algerians still preserved their mother language, i.e. Berber. Most were sedentary. The second conquest, by contrast, was a Bedouin one which penetrated deeply inside the country and resulted in the growing process of the Arabization of Algeria. This produced an intermixing of ethnic groups and brought about broad changes in the Arabic dialects spoken everywhere outside the northern areas. So we may divide the Algerian dialects into two main groups: the pre-Hilālī and the Hilālī dialects (from the name of Banū Hilāl, the tribe which settled in the center of the Maghreb, near the Banū Sulaym in Tunisia and the Banū Maʿqil in Morocco).

1.1 *Pre-Hilālī Dialects*

Pre-Hilālī dialects may be divided into village (or mountain) dialects and urban dialects.

1.1.1 Village dialects

Village dialects are located in the Oran area, between the Trara mountains and the sea. Nedroma is the main town center of this area and was probably Arabized at the time of the Idrisids. There is also a village dialect area located in eastern Algeria: it is in the eastern Kabylia between Djidjelli, Mila, and Collo. This area was occupied by the Aghlabid garrisons.

1.1.1.1 Phonetics

**q* is pronounced /k/: kalb 'heart'; **k* on the other hand is prepalatalized and pronounced [k^{y}, $\mathrm{k}^{ʃ}$, $\mathrm{t}^{ʃ}$, ʃ]. **t* is affricated and becomes [t^{s}]; the interdentals **ṯ*, **ḏ*, **ḏ̣* pass into the dentals /t/, /d/, /ḍ/. The pronunciation [ŧ] often appears for **ḍ*. **j* appears as [ʒ] when it is single and [dʒ] when it is doubled. The diphthongs **ay* and **aw* pass into /ī/ and /ū/; /m/, /b/, /q/ assimilate the article.

1.1.1.2 Morphology

The verbs III*w*/*y* have been totally rebuilt: *nsā*, *nsāt*, *yensā*, *yensāw* 'to forget'; *bkā*, *bkāt*, *bkāw*, *yebkī*, *yebkīw* 'to cry'. The verbs Iʾ have been similarly rebuilt: *klā*, *klīt*, *klāw*, *yākel*, *kūl* 'to eat'. The dual ending is *-āyen*: *yūmāyen* 'two days'. Broken plurals of the quadriliterals have a short vowel in the last syllable: *mfīteḥ* 'small keys'. Diminutives have the pattern: *ṭfeyyel* 'little child'. Both genders have merged for the 2nd person singular in the verbs and pronouns. The syllable structure of verbs and nouns + pronominal suffixes is: *ḍrab* 'he struck' + *u* = *ḍarbu* 'he struck him'; *ūled* 'child' + *u* = *weldu* 'his child'. So far, the village dialects of the Trāra and eastern Kabylia are alike. But they diverge on other points: the Trāra has *yeḍḍarbu*, eastern Kabylia *iḍarbu* 'they strike'; Trāra *reqqebtek*, eastern Kabylia *rqebtek* 'your neck'; Trāra *beʿt*, eastern Kabylia *bīʿt* 'I sold'; for the expression of durative and/or usual action, Trāras *yekteb*, eastern Kabylia *ka-yekteb* 'he is writing', or 'he usually writes'. Some remarks on syntax and vocabulary: the direct connexion (*ʾiḍāfa*) disappeared in these village dialects and was superseded by an indirect connexion through the particles *di*, *eddi*, *dyāl* + *elli* (Collo). Some words of the Arabic lexicon have the gender and/or the number of their Berber equivalent word: *ṣōf* is feminine like Berber *taḍuf* 'wool'; *ma* 'water' is considered a plural like Berber *aman*.

1.1.2 Urban dialects

The traditional distinction made between Jewish and Muslim dialects is no longer valid since most of the Algerian Jews settled either in Israel or in other countries just before or after independence (see on these dialects Ph. Marçais 1960b:376; Cohen 1981:91–105). Therefore only Muslim dialects are dealt with here.

There are urban dialects in which linguistic features from the first Arabization are still preserved: Tlemcen, Nedroma, Cherchell, Dellys, Djidjelli, and Collo. There are others in which external influences from rural and Bedouin neighbors became dominant: Tenes, Miliana, Medea, Blida, Algiers, Bejāya, Mila, Skikda, and Constantine. In some of them, the ancient urban dialects were partially superseded by the neighboring Bedouin dialects: Tenes, Blida, Miliana, Medea, Mila, Skikda, and Constantine. In some towns the dialect is completely Bedouin: Oran, Mostaganem, Mascara, Mazouna, and Annaba.

1.1.2.1 Phonetics

On the whole, village dialects, and urban dialects show the same phonetics. But only old Tenes, Cherchell, Dellys, and Constantine preserved the interdental consonants. In Medea, Blida, Algiers, occlusive and fricative interdentals may be heard. /j/ for **j* appears in the dialects of Tlemcen, Tenes, Cherchell, Miliana, Medea, Blida, Algiers, Dellys, Mila, and Constantine. Everywhere else in Algeria, this consonant is pronounced /ž/. In all urban dialects, **q* is voiceless /q/ except in Tlemcen where it is a glottal stop /ʾ/. It is a well-known feature that Maghrebi sedentary people pronounce **q* voiceless as /q/, whereas the Bedouin have voiced /g/, but the Bedouin pronunciation has also spread in some towns, at least partially.

1.1.2.2 Morphology

All urban dialects have rebuilt forms of the verbs III*w*/*y* (the same for *klā* 'to eat' and *xdā* 'to take'), broken plurals of the quadriliterals of the type *mfīteḥ* 'small keys', triliteral diminutives of the type *ṭfeyyel* 'little child', personal pronoun suffix *-u*/*-o* after a consonant, personal feminine pronoun suffix *-ah* in Cherchell instead of *-ha* elsewhere, *entūmān*, *hūmān* in Cherchell only where an ancient Andalusian influence is to be observed (particularly in lexical entries, see Grand'Henry 1972:165–166). The merger of both genders into one for the 2nd person in verbs and nouns appears in western and eastern Algeria in the urban dialects, but not in the center. The Classical Arabic *yaḍribūna* 'they strike' may be *yeḍḍarbu*, *iḍarbu*, or *yeḍarbu* according to the Algerian towns; *raqqebti*, *rqebti*, *raqebti* 'my neck' appear symmetrically. *ḍarbet* + *-u* 'she struck him' is *ḍarbātu* in the west and the center, *ḍarbettu* in the east. Indirect connexion is marked through particles: *di*, *dyāl*, *mtāʿ*, *ntāʿ*.

1.2 *Bedouin dialects*

1.2.1 Phonetics

Interdentals are preserved in most Bedouin dialects; on the whole, there is no affrication of /t/ in [t^s]; *qāf* is pronounced /g/ except in some words borrowed from the religious and/or juridical domains. Some short vowels are preserved. There is no reduction of the diphthongs /ey/, /ow/.

1.2.2 Morphology

2nd persons masculine and feminine singular never merge, the dual is fairly widespread compared with its use in sedentary dialects and it is not restricted to nouns of measure and nouns designating paired parts of the human body.

1.2.3 Syntax and lexicon

Indetermination is often expressed through the basic form of the word, without addition of *waḥd-el*. *ʾIḍāfa* is expressed the way it is in Classical Arabic, rather than through particles of indirect connexion. Plurals of the quadriliteral nouns are of the type *ṣnādīg* 'chests' instead of *ṣnādeg*. Diminutives are of the type *mfītīḥ* 'small keys' instead of *mfīteḥ*, and *ṭfeyl* 'little child' instead of *ṭfeyyel* in towns and villages. Some plurals are of the *mfaᶜᶜla* type from singular *mafᶜūl*: *maġbūn* 'deceived': plural *mġabbna*. The original pronunciation of *ᶜayn* is preserved in some numerals between 11 and 19: *tlātāᶜš* instead of *tlāṭāš* 'thirteen'. Taken as a whole, Bedouin dialects may be differentiated through phonetic and morphological/syntactical discriminating features: the pronunciation of /ġ/ like /q/; the use of a 3rd person masculine singular suffix pronoun in *-ah* instead of *-o*/*-u*; specific patterns bringing about changes in the syllable structure: *ḏ̣arbātek*, *ḏ̣arbettek*, *ḏ̣arebte k* 'she struck you'; *yeḏ̣arbu*, *raqebti*, and *yeḏ̣ḏ̣arbu*, *raqqebti* 'they strike', 'my neck'. The conjugation of verbs III*y* has the pattern: *mšet*, *mšu*, *yemšu or mšīt*, *mšāw*, *yemšīw* 'I went, they went, they go'. On the basis of these discriminating features, Ph. Marçais (1960b:378 f.) divides the Bedouin dialects into five groups:

i. Eastern Bedouin dialects have /j/, /ġ/, *ḏ̣arbātek*, *yeḏ̣arbu*, *raqebti*, *mšet*/ *mšu*/*yemšu*, diphthongs reduced to /ē/, /ō/ (this is the *E* area of Cantineau);
ii. Center and west Oran dialects have /j/, /ġ/, *-ah*, *ḏ̣arebtek*, *yeḏ̣ḏ̣arbu*, *raqqebti*, *yemšu*, diphthongs /ey/, /ow/ or /ē/, /ō/ (this is the *D* area of Cantineau);
iii. Bedouin dialects of central and Saharan Algeria have /j/, /q/ for /ġ/, *ḏ̣arebtek*, *yeḏ̣arbu*, *raqebti*, diphthongs *ey*/*ow* or *ē*/*ō*, *yemšu* (this is the *A* area of Cantineau; see also Grand'Henry 1976a:1–6);
iv. Bedouin dialects of the Tell and Sahel between Algiers and Oran have /j/ [dʒ], /ġ/, *-u*/-o, *ḏ̣arbātek*, *yeḏ̣ḏ̣arbu*, *raqqebti*. Diphthongs are /ey/, /ow/ or /ī/, /ū/. They have *mšāw* (not *mšu*) but *yemšu* (not *yemšīw*) features, which seem to be typical for a transitional area (this is the *B* area of Cantineau);
v. Bedouin dialects of the Constantine high plains have /j/ [dʒ], /ġ/, *-u*, *ḏ̣arbettek*, *yeḏ̣arbu*, *raqebti*, diphthongs developed to /ī/, /ū/ and *mšīt*, *mšāw*, *yemšīw* as in dialects of sedentary population.

Ph. Marçais (1960b:379) believes that group i may be named 'sulaymite' because it has connections with Tunisian Bedouin dialects; group ii may be named *maᶜqilian* because it has connections with Moroccan Bedouin dialects; groups iii, iv, v may be named *hilālī* with reference to the Banū Hilāl.

2. Modern Arabization of Algeria and Algerian sociolinguistics

After independence in 1962, language remained a focus of conflict in Algeria and its relationship with the state and nation is still disputed. The massive influence of French language and culture on Algerian Arabic does not result from French colonization only, but was also stimulated through emigration. Over a 40-year period, 1914–1954, two million Algerians had lived in France. During the first years of post-independent Algeria, because a large number of people found themselves unable to understand radio and television broadcasts in Standard Arabic, many advocated the use of dialectal Arabic, but the *ᶜulamāʾ* were given considerable influence in language policy. They saw Arabization primarily in religious terms and inseparable from Islamization (or re-Islamization). A large part of cultural life was Arabized (schools, a part of the university, broadcasting on radio and televison), but the centers of power and administration remained dominated by French. After the student strikes and demonstrations and the 1980 Berber Spring, a new ideology came to the fore, represented by the *Front Islamique du Salut*, which claimed legitimacy and authenticity. A new law on the generalization of (Modern and Classical) Arabic was promulgated in 1991: it committed Algeria to total Arabization by 1997 (Holt 1994:25–41). Nevertheless, Arabization has been a particularly difficult task in Algeria because of the complexity of the basic choice as to which is the most legitimate language:

Modern/Classical Arabic, considered to be the national language, and a symbol of religious identification; dialectal Arabic and Berber, marks of connection with the native soil; or French, which was long considered a symbol of modernity and change. From the beginning of the 1980s, the term 'Arabization policy' seems to have balanced between two poles: it essentially means 'Islamization' for some officials, but it does not exclude a bilingualism policy (Arabic-French) for some others. As it appears that Arabization policy did not succeed in creating one unified cultural reference for everybody in Algeria (i.e. Classical Arabic/Modern Standard Arabic), there are sociolinguists who are in favor of an official vivification of the Algerian dialects because they are potentially able to assume the three values mentioned above: identification with Islam, connection with the native soil, and modernization (Grandguillaume 1991:45–54; Redjala 1969–1970: 109–123). It has been also stressed that the dialectal Arabic variety spoken by the educated is markedly different from the one spoken by the illiterate (50 percent in Algeria). As to Berber (the mother tongue of approximately 25% of the population in Algeria, about 3 million people), there is a considerable similarity between all Berber non-Tuareg varieties. People whose mother tongue is Berber are generally bilingual and Berber is submerged by Arabic loans. In Algeria, there is a radio channel that exclusively broadcasts programs in the Kabylian language but there is a decline of Berber, which is not taught in schools. Finally, in Algeria and Tunisia, research has shown that Arabic-French bilingualism is perceived to be the best language policy (Ennaji 1991:7–25). A good way to observe the evolution of Arabic in Algeria is to check which kind of Arabic is used by theater companies: from this viewpoint, Siagh's inquiry (1991) came up with the following two results. In various theater companies from eastern, central, and western Algeria, a 'Middle dialectal Algerian Arabic' is used, i.e., a variety which tends to be as free as possible from any social or regional mark. Lexical borrowings from Classical Arabic are made, but without case endings and with a dialectal phonology. Kabylian Berber is sometimes used, especially for prologues and songs. The actors who play the roles of directors, persons in charge, or officials speak Classical Arabic only (Siagh 1991:71–86).

3. RECENT STUDIES ON PHONOLOGY, LEXICON, SYNTAX, AND POPULAR LITERATURE IN VARIOUS ALGERIAN DIALECTS

A phonological generative study of the Arabic dialect of Mila (Constantine area) resulted in the proposal of a basic phonological group CCVC which phonetically alternates with CVCC and -CCC-. If the last syllable is 'overheavy', it has the main stress. The general rule which characterizes Maghrebi dialects, prohibiting short vowels in open syllables, is confirmed (Lechheb 1986: 325–351). A study of the lexical variation of some Arabic dialects in Algeria (urban dialects of Tlemcen, Nedroma; Bedouin dialects of Tiaret, Oran, Bechar, Aïn Temouchent, Sidi Bel Abbes, El Bayad, and Mascara, with reference to Classical Arabic) reveals that only four varieties (Bechar, Nedroma, Tlemcen, and El Bayad) are "above the 70% requirement for the two (i.e. these varieties and Classical Arabic) to be considered varieties of the same language. The other five varieties can be considered separate languages from Classical Arabic". There is a greater cognation between Bedouin dialects than between urban ones, but there is a leveling of linguistic differences within the urban dialects and between the latter and Bedouin dialects, resulting from numerous migrations from rural to urban areas and from progress in the Arabization policy (Bouamrane 1994:52–79).

A syntax study based on a Tlemcen tale indicates the limits of a functionalist theory when it tries to describe the Algerian narrative clauses and the textual structure. In contrast, the theory of ancient Arab grammarians, which makes a distinction between a logic-semantic level on the one hand and a morphosyntactical level on the other hand, allows an account of clause and text unity (Mered 1992:75–102).

Verbal auxiliaries *rāḥ*, *irōḥ*, *ja*, *iji*, *mša*, *imši*, *gˁad*, *igˁod*, *bqa*, *ibqa*, *bda*, *yibda*, *wəlla*, *iwəlli*, *zād*, *izīd*, *ˁāwəd*, *iˁāwəd*, *ġādi* in the Arabic dialect of Sidi-Bel-Abbes have been studied by Madouni. He makes a distinction between verbs that preserve their original meaning (often verbs of movement) and verbs mainly used as auxiliaries (inchoatives, continuatives, repetitives) (Madouni 1994:127–139; on auxiliary verbs in Maghrebi Arabic dialects, see also Grand' Henry 1976a:457–475; 1977a:237–258; 1977b: 439–456; 1978:211–224).

A treasure of popular literature in Algerian Arabic dialect is that of riddles: a corpus of 95 riddles has been carefully transcribed (with a French translation) and studied by Bensalah (1991). The corpus is classified according to the number and length of 'séquences rythmiques', including short and long pauses, and according to rhyme. Riddles are a form of ritual game which presupposes a participation of people from various social layers and of different ages. This is why they constitute a fascinating field of study in sociolinguistics (Bensalah 1991:229–263).

The linguistic analysis of Bedouin poems in Algerian dialectal Arabic reveals remarkable similarities with old Andalusian Arabic poems, e.g. those written by Ibn Quzmān (el Cancionero): both reflect simultaneously conservative features (they use *leys* = Classical Arabic *laysa*, as a simple negative auxiliary, a typical → Middle Arabic feature), prevailing concord of feminine singular adjectives with plural nouns, and linguistic changes, such as a gradual shift from the two-term noun phrase (subject and predicate) to a three-term noun phrase (subject + *rā-* + predicate). Compare Algerian Bedouin Arabic: *eš-šowq rā-h ḥālu ṣoʿba* 'the desire, its case is a serious matter' with Andalusian Arabic of Ibn Quzmān: *kull ʿāšiq fīka hū mawlūʿ* 'everyone who is in love with you is passionate' (Grand'Henry 1995:51–57).

Algerian Judaeo-Arabic is no longer alive in Algeria, and is thus not commented on here, but it remains an attractive field of research, mainly in Israel (Bar-Asher 1993:135–191; 1992:184; 1996:167–177; Chetrit 1980:125–159; 1993: 169–204; Cohen, 1988:569; 1981:91–105).

Bibliographical references

Bar-Asher, Moshe. 1992. *La composante hébraïque du judéo-arabe algérien: Communautés de Tlemcen et Aïn-Témouchent.* Jerusalem: Magnes Press.

——. 1993. "The Hebrew component in the Judeo-Arabic dialect of Algeria (the communities of Tlemcen and Ain Temouchent)". *Hebrew and Arabic Studies in honour of Joshua Blau*, ed. Hagai Ben Shammai, 135–191. Jerusalem: Hebrew University.

——. 1996. "La recherche sur les parlers judéo-arabes modernes du Maghreb: Etat de la question". *Histoire, Epistémologie, Langage* 18:1.167–177.

Bensalah, Yasmina. 1991. "Approche rythmique de quelques devinettes d'ed-Dis et de Khermam (Bousaada, Algérie)". *Matériaux Arabes et Sudarabiques* N.S. 3.229–263.

Bouamrane, Ali. 1994. "Lexical variation among Arabic dialects in Algeria". *Proceedings of the International Congress of Dialectologists. Bamberg 29.7–4.8.1990*, ed. Wolfgang Viereck, III, 52–79. Stuttgart: F. Steiner.

Cantineau, Jean. 1937. "Les parlers arabes du département d'Alger. *Revue Africaine* 81.703–711.

——. 1938. "Les parlers arabes du département de Constantine". *Actes du IV. Congrès de la Fédération des Sociétés Savantes de l'Afrique du Nord* II, 849–863.

——. 1949. "Les parlers arabes du département d'Oran". *Revue Africaine* 84.22–231.

——. 1941. "Les parlers arabes des territoires du Sud". *Revue Africaine* 85.72–77.

Chetrit, Joseph. 1981. "L'influence du français dans les langues judéo-arabes d'Afrique du Nord". *Judaïsme d'Afrique du Nord aux XIX^{e}–XX^{e} siècles*, ed. Michel Abitbol, 125–159. Jerusalem: Institute Ben Zvi.

——. 1993. "A sociolinguistic and pragmatic study of the Hebrew component in the Judeo-Arabic poetry of North-Africa: Theoretical aspects". *Hebrew elements in Jewish languages*, ed. Yitzak Avishur and Shlomo Morag, 169–204. Jerusalem: Hebrew University.

Cohen, David. 1981. "Remarques historiques et sociolinguistiques sur les parlers arabes des Juifs maghrébins". *International Journal of the Sociology of Language* 30.91–105.

——. 1988. "Chuintement et zézaiement dans les parlers juifs du Maghreb". *Etudes Chamito-sémitiques du Groupe Linguistique d'Etudes Chamito-sémitiques* 24–28:3.569.

Ennaji, Moha. 1991. "Aspects of multilingualism in the Maghreb". *International Journal of the Sociology of Language* 87.7–25.

Grandguillaume, Gilbert. 1991. "Arabisation et langues maternelles dans le contexte national au Maghreb". *International Journal of the Sociology of Language* 87.45–54.

Grand'Henry, Jacques. 1972. *Le parler arabe de Cherchell (Algérie).* Louvain-la-Neuve: Université Catholique de Louvain (Institut Orientaliste).

——. 1976a. *Les parlers arabes de la région du Mzâb: Sahara algérien.* Leiden: E.J. Brill.

——. 1976b. "La syntaxe du verbe en arabe parlé maghrébin, I". *Le Muséon* 89.457–475.

——. 1977a. "La syntaxe du verbe en arabe parlé maghrébin, II". *Le Muséon* 90.237–258.

——. 1977b. "La syntaxe du verbe en arabe parlé maghrébin, III". *Le Muséon* 90.439–456.

——. 1978. "La syntaxe du verbe en arabe parlé maghrébin, IV". *Le Muséon* 91.211–224.

——. 1995. "Conservation and change in Algerian Bedouin Arabs". *Proceedings of the 2nd International Conference of AIDA, Cambridge 11–14.9. 1955*, ed. Joe Cremona, Clive Holes, and Geoffrey Khan, 51–57. Cambridge: Cambridge University.

Holt, Mike. 1994. "Algeria: Language, nation and state". *Arabic sociolinguistics: Issues and Perspectives*, ed. Yasir Suleiman, 25–40. London: Curzon Press.

Lechheb, Samya. 1986. "Structure syllabique et représentation phonologique dans le parler arabe de Mila". *Arabica* 33.325–351.

Madouni, Jihane. 1994. "De l'auxiliarité dans un parler de Sidi-bel-Abbes". *Matériaux Arabes et Sudarabiques* N.S. 6.127–139.

Maïri, Lies. 1987. “A bibliography of Algerian Arabic linguistics”. *Zeitschrift für Arabische Linguistik* 17.96–107.

Marçais, Philippe. 1945. “Contribution l’étude du parler arabe de Bou-Saâda”. *Bulletin de l’Institut Français d’Archéologie Orientale* 44.21–88.

——. 1956. *Le parler arabe de Djidjelli: Nord Constantinois, Algérie*. Paris: A. Maisonneuve.

——. 1960a. “*ʿArabiyya* (Western dialects)”. *Encyclopaedia of Islam* I, 578–583. 2nd ed. Leiden: E.J. Brill.

——. 1960b. “Algeria”. *Encyclopaedia of Islam* I, 374–379. 2nd ed. Leiden: E.J. Brill.

——. 1977. *Esquisse grammaticale de l’arabe maghrébin*. Paris: A. Maisonneuve.

Marçais, William. 1902. *Le dialecte arabe parlé à Tlemcen*. Paris: E. Leroux.

——. 1908. *Le dialecte arabe des ūlâd Brâhîm de Saïda*. Paris: Champion.

Mered, Zoulikha. 1992. “La notion de prédicat dans la théorie d’André Martinet et dans celle des anciens grammairiens arabes: Réflexion sur les résultats d’une analyse du discours narratif algérien”. *Cahiers de l’Institut de Linguistique de Louvain* 18:1–2.75–102.

Redjala, Mbarek. 1969–1970. “Remarques sur les problèmes linguistiques en Algérie”. *Etudes Chamito-sémitiques du Groupe Linguistique d’Etudes Chamito-sémitiques* 14.109–123.

Siagh, Zohra. 1991. “Le théâtre amateur en Algérie: Choix et usage de langues”. *International Journal of the Sociology of Language* 87.71–86.

Jacques Grand’Henry
(Catholic University of Louvain)

Algiers Arabic

1. GENERALITIES

1.1 Located in the center of the Algerian coast, Algiers, al-Jazāʾir, Algeria’s capital, chief town of a *wilāya*, faces both the sea and its hinterland. It is organized into 15 communes and has 1,483,000 inhabitants. The incorporation of 28 suburban communes into Greater Algiers makes that agglomeration the second in the Maghreb after Casablanca; it is the first city in the country with 2,562,428 inhabitants out of an estimated population of 29,100,867 inhabitants and represents 13.69 percent of the urban population. Only a small number of inhabitants of Algiers originate from families settled there before independence in 1962; like other Algerian cities, Algiers, a political, administrative, economical, and cultural capital, has been subject to strong migratory movements of several (urban or rural) populations. The geographical and social origins of these groups are different, and they have their own linguistic varieties. The contacts between these ‘new’ Algiers inhabitants and the older ones have generated a common variety of dialectal Arabic that coexists with other Arabic varieties (dialectal or non-dialectal) as well as non-Arabic idioms (Berber, French). As a meeting place, the city reinforces the role played by this common use, while maintaining the other varieties. Other factors contribute to give Algiers Arabic its special tinge; among these, Standard Arabic plays an important role.

1.2 Dialectal Arabic is the first language of about 80 percent of the population. The number of users of the other Arabic varieties (‘arabe médian’, Standard, Classical) is difficult to estimate; it is related to educational level and socio-cultural and professional activity. The use of these varieties depends on the communicational situation (functional distribution) and the speaker’s skills (passive/active users).

Dialectal Arabic, code-switching dialectal Arabic/Standard Arabic or code-switching dialectal Arabic/French have no codification, or official norms (Academy, official instructions). Classical and Standard Arabic are highly codified and submitted to official norms.

Dialectal and Middle Arabic, contrary to Standard Arabic, do not obey an official spelling code. Private (personal correspondance) or literary (theater plays, press, comics) texts in dialectal Arabic are regularly written in Arabic characters.

1.3 According to a historical classification used in Arabic dialectology (Marçais 1938), Algiers Arabic is connected with the sedentary Western dialect group. Algiers Arabic marks the boundary between the Eastern and Western Maghrebi dialects; it is classified in the urban pre-Hilālī dialectal group, which itself includes two subgroups: the Jewish dialects (that have disappeared today from Algeria) and the Muslim dialects. Thus, among the latter, strictly pre-Hilālī dialects are separated from the rest. Algiers, where the social and geographical mobilities encourage linguistic mixing, belongs to the second category.

This classification is based on the history of Arab conquests and on the contributions of the population connected to them. The distinction between urban and nomadic dialects could be explained by the fact that Arabization started from two distinct strata, corresponding to two different periods. The urban dialect stratum goes back to the first wave of Arab invasions (7th–10th centuries) and to the foundation of Arab cities. In the wake of an old urban koine the urban dialects would be the following step. As for the basis of the Bedouin and rural dialects, it could be the kind of Arabic brought by the nomadic invaders (11th century) including the Banū Hilāl. However, if the dialects were constituted according to koines on their basis (D. Cohen 1969, 1988) this scenario is not very likely. Some pre-Hilālī dialects must have existed, and the urban dialects cannot be the heir of only one kind. So if this classification, on the one hand, provides a good working hypothesis for studies, it should, on the other hand, not obscure the facts; the dialectal diversity in one single place, and the features that form it, must be examined according to each situation.

1.4 The lack of documentation does not permit a historical approach. Apart from the conversational handbooks that followed the beginning of the French conquest (Delaporte 1839, Bellemare 1850), there is little more material. During the 20th century four works are based on Algiers Arabic as their only source. The oldest one (M. Cohen 1912), devoted to the Jewish dialect, provides us with precious information about the Muslim variety. The more recent ones each deal with one field: nominal determination (Georgin 1980), phonology and morphology (Maïri 1981), verbal system (Boucherit 2002); these contain some references to more general works.

There is no dictionary of Algiers Arabic. Beaussier (1887) covers a vaster area; it was improved by Marçais (1905), updated by Ben Cheneb (1931), and completed by Lentin (1959). To this may be added Cherbonneau (1869) and his introduction, “Différences essentielles entre l’arabe littéral et l’arabe vulgaire”, and Ben Cheneb (1922) on loans in Algerian Arabic from Turkish and Persian.

With respect to texts, concerning different genres, we can refer to Bencheneb (1943) and Boucherit (2002:206–319) in phonetic transcription, with notes and translation.

2. Linguistic description

Some characteristics are specific to the Maghrebi or Algerian area, but most of them are also found in other sedentary dialects. In fact, the retained typological criteria, which are often historical, do not suggest an absolute division between the dialects that are actually in use. Those that assign Algiers to the urban pre-Hilālī dialects are present; but, more than their presence/absence, what becomes significant is the coexistence of the speakers’ different variants. When the facts are examined within this framework, we realize that, from a phonological point of view, variation predominates even in some essential features (voiceless realization of /q/, interdental treatment). Speech is more homogeneous on a morphological and syntactic level, and the tendencies found in other Arabic dialects are observed here, too. This applies to the analytic tendency with the development of nominal and verbal particles, for example the linking particle to express the dependence relationship between two nominals: *mtaʿ* or *djal* (Boucherit 1997:63–67) or, for the verb, the concomitant particle *ra* + suffix (Boucherit 2002). As for the vocabulary (see below, 3), it is the object of a double movement: it maintains the dialectal basis and the regional differentiations and it renews several loans from Standard Arabic and French, with the development of hybrid terms.

2.1 *Phonology*

2.1.1 Consonants

2.1.1.1 Two systems are represented in Table 1. In (i) the system is organized into triads (t/d/ṭ, ṯ/ḏ/ḏ̣, s/z/ṣ) and the voiceless/voiced opposition is not relevant for the velarized, although it is in (ii) for the dental velarized (ṭ/ḍ) and non-velarized (t/d) but not for the sibilant. In (i), the existence of [ƶ], which is a variant of /ḏ̣/, might lead to the combining of the phonologization of this variant and result into near-balance of the velarized subsystem, just as in the other Maghrebi dialects (D. Cohen 1970:161).

In (i) and (ii), /ʾ/ is not shared by all speakers, and has a weak contrastive output, only one pair: [saʔəl] ‘to ask [act. part.]’ – [sahəl] ‘easy’. The phoneme is only attested in older speakers; its phonic restitution by young people reflects Standard Arabic influence.

The existence of doublets: [rukba], [ᵊɍːukba] ‘knee’ and some forms like [rɛfrɛf] ‘[he] has quivered’ – [ɍafɍaf] ‘[he) has floated’ has suggested (Georgin 1980:32) the existence of an opposition /r/ – /ṛ/, but the collected data do not attest it.

Table 1. Inventory of consonants in Algiers Arabic

i. with interdental

	labial	dental	inter-dental	sibilant	palato-alveolar	affricate	velar	uvular	pharyngeal	laryngeal
Non velarized										
Voiceless	f	t	ṯ	s	š	–	k	x	ḥ	h (’)
Voiced	b	d	ḏ	z	–	j	–	ġ	ʿ	–
Nasal	m	n								
Lateral		l								
Vibrant		r								
Velarized		ṭ	ḏ̣	ṣ			q			

ii. without interdental

	labial	dental	sibilant	palato-alveolar	affricate	velar	uvular	pharyngeal	laryngeal
Non velarized									
Voiceless	f	t	s	š	–	k	x	ḥ	h (’)
Voiced	b	d	z	–	j	–	ġ	ʿ	–
Nasal	m	n							
Lateral		l							
Vibrant		r							
Velarized									
Voiceless		ṭ	ṣ			q			
Voiced		ḍ							

[p] and [f] are common in loans from French: [plɑsɑ] ‘square, place’, [fakɑ̃s] ‘holidays’; just as [tʃ] in words borrowed, a long time ago, from Turkish: [tʃərək] ‘pastry’< Turkish *çörek*; [tʃabtʃaq] ‘white iron pot’< Turkish *çömlek*.

Standard Arabic influence appears mostly in hypercorrections and fluctuations involving the interdental (Boucherit 1991).

2.1.1.2 Historically, the treatment of *qāf* differentiated between dialects. In a city like Algiers where the voiced and voiceless realizations are present, it is no longer an absolute distinctive feature. It seems that besides this historical distinction there is another one, which reconstitutes it at an another level. The voiceless realization is primarily found among middle-class and privileged people (more likely city-born); the voiced realization is found among poor people (more likely of rural origin). Thus, using one or the other variant tends to become a social marker and leads to a switch of the discriminant value.

For the dental and the interdental, the grammar handbooks of the 19th century note t^s for the Arabic consonants {t} and {ṯ}, and *d* and *ḍ* for [ḏ] and [ḏ̣]. This suggests that during this era, spirants were confused with stops and *t* was regularly realized as [t^s]. In our days, the affricate is found only sporadically among older women and some Algiers speakers use interdentals phonologically. There again, the value of the feature presence/absence of interdentals has been modified, probably because of Standard Arabic influence. It has become a ‘good language’ marker, hence its frequent hypercorrective restitution.

/’/ is not very frequent and its realization is often a literary variant that reappears by hypercorrection in initial and final position.

2.1.1.3 Phonetic realization

The above mentioned phonetic phenomena are common to several dialects, especially with respect to velarization, maintenance/loss of interdental features, and assimilation or devoicing. Velarized phonemes and /q/ have a strong effect on the other consonants and surrounding vowels. From this point of view, Algiers Arabic is not exceptional. Thus, velarization can be achieved by assimilation: [ᵴfaᵲ] 'yellow'; by expressiveness: [ˀᵯːoᵯwaᵰe] 'granny'; or by naturalization as far as French loans are concerned: [ᵬɫɑᵴɑ] 'square, place'. When close to /ḥ/ and /ʿ/, it can weaken: [ħaᵵː] 'he has put' or even disappear: [təlːaħ] 'he climbed up' (with devoicing of /ʿ/, which is common in final position). Unlike other Algerian dialects, the velarization is not striking, but even when it is almost inaudible, it maintains its influence on the quality of the adjacent vowel; this indicates, *a contrario*, that the consonant is a velarized phoneme.

/q/ has three variants. [q]: [ħqarha] 'he has oppressed her'; [qˀ] (less frequent): [ð̴a:qˀ] 'he has tasted' and [g]: in the rural based vocabulary: [gɔrn] 'a horn' but also elsewhere: [nəggəz] 'I jump'. The distribution of [q] and [qˀ] does not seem to respond to one conditioning; the alternation between [q ~ qˀ] and [g]: [gəᵵ:ɑħa ~ qta:ħa] 'he has cut' belongs especially to younger speakers.

The realizations of /ṯ/, /ḏ/, /ḏ̣/, and especially /ɖ/, are particularly unstable. /ṯ/, /ḏ/ alternate regularly with the corresponding stops [θlæθæ ~ tlætæ] 'three'; [dɔrk ~ ðɔrk] 'now'. /ḏ̣/ becomes [ð]: [inoːð] 'he stands up' – [na:ð̴] 'he stood up'; or [ᵭ] : [ʕɑᵭ: – ʕɑð̴ː] 'he has bitten, bite!' and [d] : [æðːif] 'guest' – [ð̴ijɛf] 'guests'.

The affricate realization of /t/ is sporadically observed among older women: palatalized [sobtˢ'o] 'I have found him' or not palatalized [kuntˢum] 'you have been'.

The assimilation of consonants that are in contact is regular: [ʕɑᵵ:ejjib] for *ʿad-t n-ṭeyyib* 'I will start cooking', but the insertion of a disjunctive vowel is regular too: [ʕauᵭtᵊð̴rət ~ ʕauᵭtᵊð̴rəbt] 'I have called again'.

Two further cases are also frequent: on the one hand, devoicing due to pause anticipation in final position [beːt] 'eggs', but [beːdːjalək] 'your eggs'; on the other hand, assimilation of /r/ to /ḥ/ and /ʿ/: [lfrħa] 'the joy', [ifəħ:o] 'they will be delighted'. Finally, /ʿ/ weakens at the initial: [ʕadəs ~ adəs] 'lentil' and the final: [smɑʕ ~ sma] 'he has heard' and gets devoiced at the final: [dʒuħ]'hungry'.

2.1.2 Vowels

The many vowel types, the broad dispersion of phoneme realization fields, their overlapping (as a result of vowel exposure to surrounding consonants), the role of the accent and the syllabic structure on quantity, and the abundance of neutralizations are the main obstacles to vowel system determination. The action of analogical regularization, the fate of the feminine marker *-a*, the variation due to the diverse communicational situations, and the traces of older states that interact with phonic facts are additional factors. This may explain why the following presentation is a simplified one. It will also make the passage on variable-triggering factors clearer.

2.1.2.1 The system presented in Table 2 (Boucherit 1994) consists of four phonemes, organized in two classes: front vowel /i/ – back vowel /u/ and three degrees of aperture: /i/ ~ /u/ – /ə/ – /a/. Vowel quantity is not taken into account because it is not distinctive of vowels with the same quality (see 2.1.2.3).

Georgin (1980) and Mairi (1981) suggest a second system, with similar quality of the vowels, but in which /u/ and /ū/ are opposed.

2.1.2.2 Historically, this situation extends the evolution of Maghrebi dialects (D. Cohen 1970). If we take Classical Arabic as a hypothetical representative, modern dialects exhibit a tendency, with respect to short vowels, towards a binary system: nomads = /ə/ (Classical *i*, *u*) – /a/, sedentary = /ə/ (Classical *i*, *a*) – /u/. In both cases, the length contrast is relevant only for /a/ – /ā/ or /u/ – /ū/. According to Georgin (1980) and Mairi (1981), Algiers represents the latter case. According to Boucherit's hypothesis, the situation here is similar to the case observed by Marcel Cohen for Algiers Jewish Arabic at the

Table 2. Vowel contrasts in Algiers Arabic

/i/ – /u/ :	/ʿid/	'party'	–	/ʿud/	'piece of wood'
/i/ – /a/ :	/snin/	'tooth'	–	/snan/	'teeth'
/u/ – /a/:	/šuf/	'see, look at!'	–	/šaf/	'he has seen'
/ə/ – /i/ :	/smən/	'kind of butter'	–	/smin/	'fat'
/ə/ – /u/:	/ḥəbb/	'buttons'	–	/ḥubb/	'love'
/ə/ – /a/ :	/rjəl/	'foot'	–	/rjal/	'men'

beginning of the 20th century: its evolution resulted in the confusion of the three short vowels of Classical Arabic. At the beginning of the 21st century, length correlation is not relevant, but this does not rule out the realization of phonetically long variants.

2.1.2.3 Phonetic realization: the conditioning factors

2.1.2.3.1 Consonantal environment (cf. 2.1.1.3)

This does not really influence quantity, but it modifies vowel quality. It does not totally change the fundamental quality of /i/, /u/, /a/ but determines the color of /ə/:

velarized consonants and /q/ have an opening effect or back the vowel: /i/ = [e̞, e], /u/ = [u̞, o̞] and sometimes [ɔ], /a/ = [ɑ], /ə/ = [ɔ] or [ɑ];

/q/, /x/ and /ʿ/, /ḥ/ back /a/ = [ɑ] and /ə/ = [o̞, ɔ, ɑ]; /i/ and /u/ are in that case less perceptible;

front consonants and liquids reinforce the anteriority of /i/, and orient /ə/ toward frontness; they have little effect on /u/ and /a/;

/k/, /g/, /ʾ/, and /ḥ/ do not act on the vowel's quality;

[w] makes the preceding or following vowel round or back;

[j] anteriorizes.

In addition to the conditioned variants, we have individual and situational variants and cases in which the consonant's environment is neutral, for example the feminine marker *-a* is generally realized as [a].

2.1.2.3.2 Quantity, accent, syllable

In general, vowels are long when the syllable is stressed and, in that case, they are less exposed to the consonant's environment. In an unstressed syllable, they are realized as middle or short, except in absolute open final syllable, most of which are middle. In the absolute word-initial, vowels are short but, since the syllables V# : [u] 'and' or VC- : [aʕ'ma] 'blind' are not very frequent, these cases are rare and result from the elision /ʾ/. Thus, we note, [u'dᵊn ~ ʔu'dᵊn] 'ear', [ʔimala ~ imala ~ mala ~ ᵊmmala] 'therefore'. Ultra-short vowels, mostly [ə] colored, serve as disjunction elements in the consonant cluster or as an onglide for the initial consonant, see above [ᵊmmala].

Whatever its pronunciation, the phoneme /ə/ is realized as a short one, except when in contact with pharyngeal fricatives where it becomes longer.

When close to [j] and [w], the variants of /i/ and /u/ are longer and the diphthongs have a tendency to be reduced to long sounds: [wajin] but also [weːn] and [wiːn] 'where'.

In some cases, the deletion of the pronominal suffix sg.masc. *-h* is compensated for by stressing and lengthening the final vowels. Thus [ikət'buː] 'they write it' and [i'kətbu] 'they write' are either distinguished by their final vowel and by stress position, or by the context or the situation. Boucherit (2003) considers that there is no reason to point out that stress is distinctive since the opposition would be effective only in that case and neutralized everywhere else, and because *-h* is always capable of appearing.

A syntagmatic compensation phenomenon br ings about the move, in the chain, of syllabic quantity as in [haːdak] and [hadːɛk] 'that-one [masc.]' where long vowel and geminate consonant (quantitatively long) alternate. Based on ten monosyllabic pairs ([ŧo̞ːƀ] 'brick' / [ŧɔƀː] 'medicine'), Georgin (1980:72) postulates length for the back vowel. But, for Boucherit, these examples are phonologically difficult to accept, unless we oppose consonants and vowels.

2.1.2.3.3 The main variants

A few examples will show the phonetics of the dialect without listing all cases according to consonantal environment and syllabic type:

/i/	
[iː]	['biːt] 'room', [ɍɔˑħiː] 'go! [fem.]', [l'ɣaːʃi] 'the crowd', [kiˑfɛʃ] 'how'
[eː]	['kbeːrɑ] 'big' [fem.], [ŧːwɑˑ'qeː] 'the windows', [ħeːŧ] 'wall', ['rɔƀːe] 'my God!'
[i] ~ [e̞]	['ʃwiˑjːa ~ 'ʃwe̞jːa] 'little'.
/a/	
[ɑː]	[l'mɑː] 'the water', ['qɑːʕdɑ] 'seated [fem.]', ['wɑːħd] 'one', [ɣɑ:ləb] 'winner'.
[aː]	[l'ʔaːn] 'now', [ᵊsːxaːna] 'the heat'.
[ɛ] ~ [a] ~ [ɑ]	[baˑ'lɛːk ~ ba'laːk] 'attention', [aħ'wɛːjdʒi ~ aħ'waːjdʒi] 'my business', ['ħɛtta ~ 'ħɑtta] 'until'.
/u/	
[uː]	[dʒuˑ'nuːd] 'soldiers', [i'quːlo̞] 'they say'.
[oː]	[l'foːq] 'above', [nəssɔknoˑ] 'we live'

[u] ~ [ọ] ~ [ɔ] [l'kulː ~ l'kọlː ~ l'kɔlː] 'the totality, the whole'

/ə/ : all the listed vowels appear in close relationship with the consonantal environment from [i] to [ɑ] passing through [e, ø, ɔ].

2.1.2.4 Semivowels

[w] and [j] are variants of /u/ and /i/ but in terms of syllables, they have a consonantal role: [huwa] 'he, him', and [huma] 'they, them'.

2.1.2.5 Syllabic structure

The syllabic structure of Algiers Arabic is governed by a general law that tends to avoid the appearance of short vowels in open non-final syllables (Cohen 1912:141). This law explains the syllabic economy of the dialect, which is characterized by the abundance of consonantal clusters and by the predominance of closed over open syllables. This gives the Algiers dialect, just like all the other Maghrebi dialects, a syllabic structure and, therefore, an auditory impression that is very different from those known in Oriental dialects. Word-initially, the syllable begins either with a simple consonant CV, CVC(CC), or with a group of two CCV, CCVC(CC) or of three consonants CCCV, CCCVC(C), more rarely with a vowel V, V-, VC. Word-finally, the syllable can be opened or closed by one, two, or three consonants. Finally, consonants can be long word-initially and word-finally and geminated medially (the above mentioned consonant and vowel forms illustrate some of these types).

2.1.2.6 Stress

It is not distinctive, but its contrastive value is increased because, when stressed, a vowel is long and its position predictable. Usually, the stress falls on the last syllable if the latter is closed: [u'dən] 'ear' or on the penultimate (open or closed): ['nqɑɍːe] 'I teach', [qɑːʕdɑ] 'seated [fem.]'. These rules do not apply if the last closed syllable is a 2nd/ 3rd person suffixed pronoun: *-kum*, *-hum*, *-k*, *-h*: ['ꞩɑbkum] and not [ꞩɑb'kum] 'he has found you'.

2.2 *Morphology and syntax*

2.2.1 Personal pronouns

The subject and object personal pronouns are set out in Table 3.

2.2.2 Particles

2.2.2.1 *ha-*

This is a demonstrative interjection found in demonstrative pronouns (*had*, *hada*, *hadi* 'this, this one [masc./fem.]'), is constructed with the help of suffixed pronouns, and is compatible with nominal predicates: *hāhu ṛājel* 'Here is a

Table 3. Personal pronouns in Algiers Arabic

	Subject personal pronouns				Object personal pronouns		
	Affixes		Autonomous		Suffixes		Autonomous
	Prefix Conjugation	Suffix Conjugation	Independents	Coordinated[1]	Direct[4]	Indirect	Indirect
3rd sg. masc.	*i- ~ y-*	*-Ø*	*huwa*	*iyya-h*	*-u ~ -h*[3]	*-lu*	*lu(h)*
3rd sg. fem.	*t-*	*-ət*	*hiya*	*iyya-ha*	*-(h)a*	*-lha*	*liha*
3rd pl.	*i- ~ y-…-w*	*-w*	*huma*	*iyya-hum*	*-hum*	*-lhum*	*lihum*
2nd sg. masc.	*t-*	*-t*	*nta ~* ***ntaya***[2]	*iyya-k*	*-k*	*-lək*	*lik*
2nd sg. fem.	*t-…-i*	*ti*	*nti ~* ***ntiyya***[2]	*iyya-k*	*-ki*	*-lki*	*liki*
2nd pl.	*t-…-w*	*-tu*	*ntuma*	*iyya-kum*	*-kum*	*-lkim*	*likum*
1st sg.	*-n*	*-t*	*ana ~* ***anaya***[2]		*-ni*	*-li*	*liyya*
1st pl.	*n-…-u*	*-na*	*ḥna*		*-na*	*-l-na*	*lina*

[1] Used when two pronouns are co-ordinated: *ana u ijjak* 'me and you'.

[2] In bold: emphasis forms : *škun* ᵊ***ntijja*** *–* ***anaja*** *ʿammtik* . . . '**You,** who are you? – **Me**? I am your aunt . . .' [= I am not just anybody].

[3] Realized as *-u* after a consonant: *qtəlt-u* 'I killed him' and *-h* after vowel *qtəltu-h* 'You [pl.] have killed him'.

[4] These are combined with the verb: *ḥagro-ni* 'They have oppressed me', the demonstrative: *ha- ha-ni nrūḥ* 'I am going away', the particle *ra-* (copula or presentative): *rā-hi ttmənja w ərbəʿ* 'it is a quarter past eight', and the negative copula: *had əṣṣbiḥa ma-niš mliḥ* 'This morning, I am not feeling well'.

man [= that, this is . . .]' or verbal predicates: *ai* (< *hahi*) *tmesxret bīk* 'There she is, making fun of you'. In negative sentences, the attraction of the negative copula regularly results in confusing *ha*-, *ra*-, and *ma*-. *ha*- is maintained only exceptionally; *ra*- is maintained more frequently (*ma-ṛāni ma ṭwīl ma qṣīr* 'I am neither tall nor short') but not systematically (*rani maṭlob maniš mgaže* 'I was called up into the army, I didn't join up') and should be considered sometimes as a mark of instance.

2.2.2.2 *ra-*

Very common in Algeria, and to a smaller extent in the rest of the Maghreb, this particle, which has some nominal and verbal uses, is conjugated with the series of the suffixed personal pronouns: *ra-ni*, *rā-nā* 1st person sg., pl., etc. When used in a nominal sentence, it functions as a copula or a presentative: *rani fi lkuzina* 'I am in the kitchen'. When used in a verbal sentence, it is placed before a verb in the prefixed or suffixed conjugation. It then qualifies the aspectual value: *ra-h ixaus aᶜli* 'he looks for you ~ he is busy looking for you'; *ma-rāk-š ᵊtšəmm* . . . 'don't you feel ~ aren't you . . .'.

In a nominal sentence, these particles are used with a noun class which refers to actual situations or states. But *ha*-, which has conserved its nature of interjectional particle connected with speech acts, mostly denotes situations or states viewed as momentary. In this respect, the durational criterion, which is not distinctive in the case of copula sentences (temporary duration) and two-term nominal sentences (permanence), seems more operational to differentiate *ha*- and *ra*-, even if in sentences like *ha-ni mrēḍ*, *ra-ni mrēḍ*, *ana mrēḍ* 'I am sick', it is difficult to use. Only the communicational situation shows whether the situation in question is temporary or permanent.

2.2.2.3 *mtaᶜ* and *dyal*

Both of these particles are used to express the dependence relationship between two nominals: *lḥɛlfa djali* ~ *mtaᶜi* 'my espadrille'. Both of them may coexist in the same sentence: *ḥɛbbɛt ṭəṣnaᶜ roppa kīma mtaᶜ hadīk əlmaḍāma dyal lfilm* 'She wanted to make herself a dress, just like the lady's in the film'. This construction, although it is generalized, has not completely superseded the synthetic construction. The latter continues to be used for frequent terms and kinship terms, or for terms related to intimate and everyday life: *qəlbi* 'my heart'.

2.2.3 Negation

Nominal: *maši: maši ana* 'it is not me' [neg. me].

Verbal: *ma-...-š: ma-txəl-š* 'he didn't come in' or *ma-: ma-nᶜarf* 'I don't know' (the use of *ma*- helps to avoid heavy consonant clusters). When the verb is preceded by *ra*- (proverb) or an auxiliary, they carry the negation marker: *ma-rākš ᵊtšəmm* 'Don't you smell?', *ma-iᶜawwəd yəddi* 'He won't take again' ('[neg.] he starts again he takes').

ᶜand 'at someone's', *fi* 'in' are treated and negated as verbs: *ma-ᶜandiš sdiq, gutlək ma-ᶜandiš saḥbi* 'I have no friend, I tell you: "I have no friend"', *ma-fihaš taᶜb* 'There is no trouble' (= 'It is not difficult'). This is more rare with *ᶜla* 'on' except in frozen sentences *ma-ᶜliš* 'It doesn't matter'.

Negative copula: *ma* + suffix: *had əṣṣbīḥa maniš mliḥ* 'This morning, I am not feeling well'.

2.2.4 Diminutive

There is a diminutive form with a geminated [j]: *ṭfəjjəl* 'young kid' and another one of a *mfītəḥ* 'small key' type. A diminutive type, currently obsolete, was noted as regular (Delaporte 1939:41) in the last century for aCCaC template adjectives referring to a color or a deformity: *aḥmar* ⇒ *ḥmimar* 'red'; *aṣfar* ⇒ *ṣfifar* 'yellow'; *ᶜawar* ⇒ *ᶜwiwar* 'one eyed'.

2.2.5 Verbal paradigms

The paradigm of the imperfect and the perfect verb is set out in Tables 4 and 5.

2.2.6 Auxiliary elements

2.2.6.1 *kan*

It is normally inflected with affixed subject pronouns. In the prefixed conjugation, the stem is *kun* and in the 1st and 2nd person singular, masculine and feminine are not distinguished. In the suffixed conjugation, the stem is *kun* in 1st and 2nd person and *kan* in 3rd person.

kan is both a verb of existence and an auxiliary verbal which supports temporal and modal determination. It refers to the past when used in the suffixed conjugation: *kant matət yimmaha* 'his/her mother was dead'. In the prefixed conjugation it refers to the future; it adds a modal shade and is sometimes interpretable as a present tense: *tkun zina u šəbba* 'She is pretty and beautiful'.

Table 4. Prefixed conjugation: Imperfect

	‘to write’	‘to eat’	‘to say’	‘to come’	‘to take’	‘to bite’
1st sg.	*nə-ktəb*	*n-akul*	*n-qul*	*n-ži*	*nə-ddi*	*n-ʿaṭṭ*
1st pl.	*nə-ktb-u*	*n-akl-u*	*n-qul-u*	*n-ži-w*	*nə-ddi-w*	*n-ʿaṭṭ-u*
2nd sg. masc.	*tə-ktəb*	*t-akul*	*t-qul*	*t-ži*	*tə-ddi*	*t-ʿaṭṭ*
2nd sg. f.	*tə-ktb-i*	*t-akl-i*	*t-qul-i*	*t-ži*	*tə-ddi*	*t-ʿaṭṭ-i*
2nd pl.	*tə-ktb-u*	*t-akl-u*	*t-qul-u*	*t-ži-w*	*tə-ddi-w*	*t-ʿaṭṭ-u*
3rd sg. masc	*yi-ktəb*	*y-akul*	*i-qul*	*i-ži*	*yə-ddi*	*i-ʿaṭṭ*
3rd sg. fem	*tə-ktəb*	*t-akul*	*t-qul*	*t-ži*	*tə-ddi*	*t-ʿaṭṭ*
3. pl.	*yi-ktb-u*	*y-akl-u*	*i-qul-u*	*i-ži-w*	*yə-ddi-w*	*i-ʿaṭṭ-u*

Table 5. Suffixed conjugation: Perfect

	‘to write’	‘to eat’	‘to say’	‘to come’	‘to take’	‘to bite’
1st sg.	*ktəb-t*	*kli-t*	*qul-t*	*ži-t*	*ddi-t*	*ʿaṭṭi-t*
1st pl.	*ktəb-na*	*kli-na*	*qul-na*	*ži-na*	*ddi-na*	*ʿaṭṭi-na*
2nd sg. masc.	*ktəb-t*	*kli-t*	*qul-t*	*ži-t*	*ddi-t*	*ʿaṭṭi-t*
2nd sg. fem.	*ktəb-ti*	*kli-ti*	*qul-ti*	*ži-ti*	*ddi-ti*	*ʿaṭṭi-ti*
2nd pl.	*ktəb-tu*	*kli-tu*	*qul-tu*	*ži-tu*	*ddi-tu*	*ʿaṭṭi-tu*
3rd sg. masc.	*ktəb*	*kla*	*qal*	*ža*	*dda*	*ʿaṭṭ*
3rd sg. fem.	*ktəb-ət*	*kla-t*	*qal-at*	*ža-t*	*dda-t*	*ʿaṭṭ-ət*
3rd pl.	*ktəb-u*	*kla-w*	*qal-u*	*ža-w*	*dda-w*	*ʿaṭṭ-u*

Its active participle *kayin* yields an existential predicate: *waš kayin (l-)makla* ‘What is there for eating?’.

2.2.6.2 *raḥ*

For the immediate future the active participle of the verb *rāḥ* ‘to go’ is used. It agrees in gender and in number and is followed by the prefixed conjugation. The verb’s meaning is lost and the auxiliary is also used with verbs that do not imply any movement: *rāḥ iṭēb* ‘it will cook’.

2.2.6.3 *wāsā*: ‘to do, to act, to put, to place’

This serves to express the inchoative: *naḍətsbaḥ ~ wasit xəmməlt darħa* ‘One morning she woke up and started doing housework’.

2.2.6.4 *ʿadd*: ‘to turn back, to come back, to restart’

This expresses reiteration: *fħam ħadak ššīx ʿad ᵊrjaʿ ᵊrjaʿ məlbɛb* ‘the old man understood [and] moved back from the door’.

2.2.6.5 Active participle

In its predicative uses the active participle (inflected for masc., fem., and pl.) expresses an ongoing process: *qaʿdat . . . ʿiša mʿa ixwaħa* ‘she used to stay . . . living with her brothers’. When combined with *ra-* it keeps the same function and meaning and the particle plays the role of a copula: *ra-ni xaddəm* ‘I am busy working’.

2.2.7 Modes

Only the imperative has a specific paradigm (*qul*, *quli*, *qulu* ‘to say’, respectively 2nd person sg. masc., fem. and pl.). To express potentiality, unreality, wishes or injunctive meanings, the pre-fixed/suffixed conjugation is used without *ra-* : *ifrəj rəbbe* ‘with God’s help’, *ma-troḥoš* ‘do not leave’, *lukan-tlagga (< ntallaga) biħa nqulħa* ‘if I meet her, I will tell her’, *ila kamәlt kulši nroḥo nsəmʿoh* ‘if I have finished everything, we will go and listen to him [= the musician]’. In these conditional sentences, the verb in the protasis can be preceded by *ila* or *lukān* ‘if’, but this is not obligatory.

2.2.8 Derived forms

There are about ten such forms which are more or less productive and have a more or less stable value. Thus, for verbs with a three-consonant stem, where Form I is $C_1C_2vC_3$ (type *ktəb* ‘he wrote’), the following derivatives may be mentioned:

2.2.8.1 1st group

gemination of the second radical: $C_1vC_2C_2vC_3$ (= Form II): *rqəd* ⇒ *rəqqəd* ‘to sleep ⇒ to send to sleep, to make [someone sleep]’;

infixation of an *-a-* vowel between the 1st and the 2nd radical: $C_1aC_2vC_3$ (= Form III): *ktəb* ⇒ *katəb* 'to write ⇒ to write to someone';

prefixation of *t-* to a 2nd form: t-$C_1vC_2C_2vC_3$ (= Form V): *ʿalləm* ⇒ *tʿalləm* 'to learn ⇒ to learn, to learn for oneself';

prefixation of *t-* to a 3rd form: t-$C_1aC_2vC_3$ (= Form VI): *samaḥ* ⇒ *tsamaḥ* 'to forgive someone ⇒ to forgive each other'.

2.2.8.2 2nd group

prefixation of *t-* to a Form I: t-$C_1C_2vC_3$. It is likely that this form, which does not have any Classical Arabic equivalent but is known in Maghreb, was constructed by analogy, on the model of Forms V and VI. This process has regularized the mode of formation of the derived forms where a derived form with a prefix *t-* corresponds to each of the first three forms: *bna* ⇒ *tbna* 'to build, to construct ⇒ to build oneself, to construct oneself';

prefixation of *n-* to a Form I: n-$C_1C_2vC_3$ (= Form VII): *ḍrəb* ⇒ *nḍrəb* 'to hit ⇒ to be hit', or 'to build, to construct ⇒ to build oneself, to construct oneself', to be brought closer to *bna* ⇒ *tbna* mentioned above.

2.2.8.3 3rd group

infixation of *-t-* between the 1st and the 2nd radical: C_1-t-C_2vC_3: (= Form VIII): *štrək* 'to associate with';

prefixation *st-* to a Form I: st- $C_1C_2vC_3$: (= Form X): *staʿgeb* 'to be amazed';

infixation of an *-a-* or *-ə-* vowel between the 2nd and the 3rd radical: $C_1C_2aC_3$ (= dialectal form): *zraq* 'to become blue', *smən* 'to put on weight'.

3. Lexicon

In Algiers, as in the dialects of other cities, → diminutives are frequent and their formation is quite productive (cf. 2.2.4). There is a certain regional differentiation (e.g. 'pumpkin' = *qarʿa* in Algiers, *dziriwɛt* in Constantine), but since there is no systematic study, it is difficult to be precise on this point.

Bibliographical references

Beaussier, Marcelin. 1958. *Dictionnaire pratique arabe – français*. 2nd ed. Rev. Mohamed Ben Cheneb. Algiers: J. Carbonel.

Bellemare, Alexandre. 1850. *Grammaire arabe (idiome d'Algérie) à l'usage de l'armée et des employés civils de l'Algérie*. Paris and Algiers: Hachette et Cie and Dubos Frères.

Ben Cheneb, Mohamed. 1922. *Mots turcs et persans conservés dans le parler algérien*. Algiers: J. Carbonel.

Bencheneb, Rachid. 1943–1944. "Textes arabes d'Alger". *Revue Africaine* 87.396–397, 221–243; 88.398–399, 123–140.

Boucherit, Aziza. 1991. "Convergence et résistance des hommes et des langues". *International Journal of the Sociology of Language* 87.55–69.

——. 1994. "Phonétique ou phonologie: Les voyelles d'Alger". *Matériaux Arabes et Sudarabiques N.S.* 6.103–125.

——. 1997. "L'expression du rapport de dépendance: Constructions synthétique et analytique en arabe algérois". *Afroasiatica Neapolitana*, ed. Alessandro Bausi and Mauro Tosco, 63–67. Naples: Istituto Universitario Orientale.

——. 2002. *L'arabe parlé à Alger: Aspects sociolinguistiques et énonciatifs*. Paris and Louvain: Peeters.

——. 2003. "Le phonème 'virtuel' ou Comment le passé fait partie du présent". *Mélanges David Cohen*, ed. Jérôme Lentin and Antoine Lonnet, 119–127. Paris: Maisonneuve et Larose.

—— and Jérôme Lentin. 1989. "Les dialectes féminins dans le monde arabe: Des dialectes minoritaires et leur évolution". *Linx* 21.17–37.

Cherbonneau, Auguste, 1986. *Dictionnaire français – arabe pour la conversation en Algérie*. Beirut: Librairie du Liban. 1st ed. 1869.

Cohen, David. 1970. *Études de linguistique sémitique et arabe*. The Hague and Paris: Mouton.

——. 1973. "Variantes, variétés dialectales et contacts linguistiques en domaine arabe". *Bulletin de la Société de Linguistique de Paris* 68.215–248.

Cohen, Marcel. 1912. *Le parler arabe des Juifs d'Alger*. Paris: Champion.

Delaporte, J.-H. 1839. *Principes de l'idiome arabe en usage à Alger*. Algiers and Paris: Brachet et Bastide and Charles Hingray.

Georgin, Pierre. 1980. *Esquisse phonologique et détermination nominale du parler arabe d'Alger*. Ph.D. diss., René Descartes University-Paris 5.

Lentin, Albert, 1959. *Supplément au Dictionnaire pratique arabe – français de Marcelin Beaussier.* Algiers: J. Carbonel.

Mairi, Lies, 1981, *The phonology and morphology of Algiers Arabic*. Ph.D. diss., Exeter University.

Marçais, William. 1905. "Quelques observations sur le Dictionnaire pratique arabe – français de Beaussier". *Recueil de mémoires et de textes*, Algiers University, École Supérieure des Lettres, 409–504. Algiers: P. Fontana.

——. 1938. 'Comment l'Afrique du Nord a été arabisée'. *Annales de l'Institut des Etudes Orientales* IV, 1938, 1–12; XIV, 1956, 5–17. (Repr. William Marçais, *Articles et conférences*, 171–192. Paris: Adrien Maisonneuve 1961.)

Aziza Boucherit
(Université René Descartes – Paris 5)

ʿAmal

1. GOVERNMENT AND GOVERNORS

The syntactic term *ʿamal* ‘action, performance’ denotes ‘governance’, i.e. the grammatical effect of one word of a sentence on another. All constituents of a sentence are either *ʿawāmil* (sg. *ʿāmil*) ‘governors’ or *maʿmūlāt* (sg. *maʿmūl*) ‘governed’. The effect of this government is a case ending (→ *ʾiʿrab* ‘declension’). For the noun these endings are:

-u nominative (*rafʿ*): *rajul-**un*** ‘a man’;
-a accusative (*naṣb*): *rajul-**an***;
-i genitive (*jarr* or *xafḍ*): *rajul-**in***.

In the verb only the imperfect has declined forms:

-u indicative (*rafʿ*): *yaḏhab-**u*** ‘he goes’;
-a subjunctive (*naṣb*): *yaḏhab-**a***;
-ø jussive (*jazm*): *yaḏhab*.

Since no categorical distinction is made between verbal and nominal endings, *rafʿ* can mean either ‘nominative’ or ‘indicative’, depending on context (Owens 1988:39). All parts of speech (nouns, verbs, and particles) can operate as governors, while only nouns and imperfect form of verbs can be governed. Particles are indeclinable.

An early classification of all types of governors is found in the *Kitāb al-jumal fī n-naḥw*, ascribed to al-Xalīl ibn ʾAḥmad (d. 791; cf. Owens 1990:189–193). A summary of the theory of governance is given by ʿAbd al-Qāhir al-Jurjānī (d. 1078) in *ʿAwāmil miʾa* (cf. Jirjāwī, *Šarḥ*) and *Jumal*.

Governors are divided into overt (*ʿawāmil lafdiyya*, lit. ‘verbal’) and virtual (*ʿawāmil maʿnawiyya*, lit. ‘notional’) (ʿAbd al-Laṭīf n.d.: 168; ʿAmāyira 1987:56); the existence of the latter is assumed when there is no overt governor. Overt governors are divided into regular (*ʿawā-mil qiyāsiyya*, lit. ‘analogical’) and those that govern on the basis of usage (*ʿawāmil samāʿiyya*, lit. ‘aural’).

2. REGULAR GOVERNORS

Regular governors are represented by the following categories of words (ʿAbd al-Laṭīf n.d.: 168):

i. Verbs, which are either transitive (*mutaʿaddin*) or intransitive (*ġayr mutaʿaddin*, or *lāzim*). Transitive verbs govern an agent (noun or personal pronoun), in the nominative, and all direct objects in the accusative, e.g. *ḍaraba zaydun ʿamran* ‘Zayd struck ʿAmr’. They may accept up to three direct objects, e.g. *ʾarā l-lāhu zaydan ʿamran xayra n-nāsi* ‘God showed Zayd that ʿAmr is the best of the people’.

Verbs accepting one direct object in Form I, become ditransitive in Forms II and IV, as in *ʾaḍrabtu zaydan ʿamran* ‘I forced Zayd to strike ʿAmr’; verbs accepting two direct objects in Form I, become tritransitive in Forms II and IV, e.g. *ʾaʿlama l-lāhu zaydan ʿamran ʾaxā-ka* ‘God revealed to Zayd that ʿAmr is your brother’.

In the passive, transitive verbs govern a direct object in the nominative; this object is said to replace an agent (*al-mafʿūl al-qāʾim maqāma l-fāʿil*), the other objects taking the accusative, e.g. *ḍuriba zaydun ḍarban šadīdan* ‘Zayd was hit strongly’. Every object, except objects of reason and partnership, can replace an agent in the passive, e.g. *sīra farsaxāni* ‘two parasangs were traveled’.

The Basran grammarians claim that if there is an overt direct object, this must replace the agent in the passive, while the Kufan grammarians accept such replacement by all objects, even if there is a direct object. Therefore, they allow expressions like *ḍuriba ḍarbun šadīdun zaydan* ‘a strong blow struck Zayd’, where the direct object *zaydan* remains in the accusative, while the absolute object *ḍarbun šadīdun* replaces the agent (Ibn ʿAqīl, *Šarḥ* II, 121).

Intransitive verbs govern their agent in the nominative, and all objects, except the direct object, in the accusative, e.g. *qāma zaydun ʾikrāman li-s-sulṭāni* ‘Zayd rose to honor the sultan’. They become transitive in Forms II and IV or govern by means of particles. Thus, in the phrase *marra zaydun bi-ʿamrin* ‘Zayd passed ʿAmr’, the action of the verb *marra* affects the object *ʿamrin* through the particle *bi-*. Verbs governing their object by means of particles, sometimes become transitive, when the particle is omitted, and govern a direct object in the accusative. Thus, in the phrase *daxaltu l-bayta* ‘I entered the house’, the intransitive verb *daxala* has become transitive after deletion of the particle *ʾilā* ‘into’.

In the passive, an object with a particle replaces the agent of the intransitive verb, e.g. *murra bi-ʿamrin* ‘ʿAmr was passed’, where *bi-ʿamrin* operates as the object replacing the agent of *murra*.

ii. The verbal noun (→ *maṣdar*), when it is used independently with *tanwīn*, may operate as a verb, governing the agent in the nominative and the object in the accusative, e.g. *ʾaʿjaba-nī ḍarbun zaydun ʿamran* ‘I was astonished that Zayd beat ʿAmr’. More often, though, it governs its agent in the genitive and its object in the accusative, e.g. *ʾa ʿjaba-nī ḍarbu zaydin ʿamran*. Alternatively, the object may be governed in the genitive and the agent in the nominative, as in *ʾaʿjaba-nī ḍarbu ʿamrin zaydun*.

In the example *ʾaʿjaba-nī ḍarbu zaydin ʿamran* the agent has a genitive case ending, but is considered to have a nominative position in the sentence. Attributes may agree with it either in the genitive, e.g. *ʾajibtu min šurbi zaydin ad-darīfi* ‘I was surprised by the drinking of lovely Zayd’, or in the nominative, *ad-darīfu* (Ibn ʿAqīl, *Šarḥ* III, 104).

Substantive verbal nouns (*ʾasmāʾ al-maṣdar*), such as *kalām* ‘speech’, *ʿaṭāʾ* ‘gift’ etc. sometimes govern like nouns of action, e.g. *min qublati r-rajuli imraʾata-hu l-wuḍūʾu* ‘ablution [is made necessary] by a husband kissing his wife’. Here *qublatun*, being used in the meaning of ‘kissing’, governs the agent *ar-rajuli* in the genitive and the object *imraʾata-hu* in the accusative.

iii. Active participles (*ism al-fāʿil*), when indefinite and expressing present or future tense, govern similarly to a transitive verb, with the agent in the nominative, and the object in the accusative, e.g. *zaydun ḍāribun ġulāmu-hu ʿamran* ‘the slave of Zayd is beating ʿAmr’.

This is permitted only if the participle is a) a predicate of a topic (*mubtadaʾ*), as in the example; b) an attribute of a preceding noun, e.g. *jāʾa-nī rajulun ḍāribun ʿabdu-hu ʿamran* ‘a man came to me, whose slave beats ʿAmr’; c) an adverbial modifier (→ *ḥāl*) of a preceding noun, e.g. *jāʾa zaydun rākiban farasan* ‘Zayd arrived riding a horse’; or d) preceded by interrogative or negative particles, e.g. *ʾa-ḍāribun zaydun ʿamran?* ‘is Zayd beating ʿAmr?’ (Zamaxšarī, *Mufaṣṣal* 226–229; Jirjāwī, *Šarḥ* 294–295).

If these conditions are not met, or when a participle is used with a past reference, its object is governed in the genitive, as in *hāḏā ḍāribu zaydin ʾamsi* ‘this [is who was] beating Zayd yesterday’. However, if a participle is used with the definite article *al-*, it governs an object in the accusative, irrespective of time, e.g. *hāḏā ḍ-ḍāribu zaydan* ‘this [is who was/is/will be] beating Zayd’.

Nouns coordinated with the object of a participle in the genitive can be coordinated with it either in the genitive, e.g. *hāḏā ḍāribu zaydin wa-ʿamrin* ‘this [is who is] beating Zayd and ʿAmr’, or in the accusative, e.g. *hāḏā ḍāribu zaydin wa-ʿamran*, since the participle may govern the object in the accusative (Ibn ʿAqīl, *Šarḥ* III, 119).

According to the grammarians, adjectives of the pattern *faʿʿāl, mifʿāl*, *faʿūl*, *fāʿil*, and *faʿīl*, called intensive (*ʾamṯilat al-mubālaġa*), are regular governors, because their use resembles that of a participle, the intensive meaning expressing permanent occupation with something, e.g. *ʾammā l-ʿasala fa-ʾana šarrābun* ‘as for honey, I am [permanently] drinking [it]’. Basran grammarians assume that intensive adjectives of the first three patterns govern like a participle, but they disagree about the governance of adjectives of the last two patterns. Kufans do not recognize governance of these adjectives. They assume that the noun functioning as object of these adjectives is governed in the accusative as the result of the action of a deleted verb.

iv. Passive participles (*ism al-mafʿūl*) govern under the same conditions as a passive verb, that is, they put the object replacing the agent in the nominative, e.g. *zaydun maḍrūbun ʿabdu-hu* ‘Zayd, his slave [is] beaten’, ‘Zayd’s slave is beaten’ (Ibn ʿAqīl, *Šarḥ* III, 122).

However, the relation between a passive participle and an object replacing an agent may be expressed by annexing the latter to the former, e.g. *jāʾa-nī rajulun maḍrūbu ʿabdin* ‘a man came to me whose slave was beaten’, or by placing the object in the accusative of specification, e.g. *maḍrūbun ʿabdan* (Zamaxšarī, *Mufaṣṣal* 229; Jirjāwī, *Šarḥ* 299).

v. Adjectives assimilated to the participles (*ṣifāt mušabbaha bi-ʾasmāʾ al-fāʿil*) govern as participles by putting the agent in the nominative and the object in the accusative. Thus, in *zaydun ḥasanun al-wajha* ‘Zayd [is] lovely by [his] face’ the adjective *ḥasanun* acts as if there is an implied personal pronoun *huwa* ‘he’ with its object *al-wajha*. The object of assimilated adjectives, which always comes after its governor, can also stand in the nominative or the genitive, e.g. *zaydun ḥasanun al-wajhu* and *ḥasanu l-wajhi* (Zamaxšarī, *Mufaṣṣal* 230–231; Ibn ʿAqīl, *Šarḥ* III, 140–146; Jirjāwī, *Šarḥ* 300–302).

Among the regular governors are the adjectives expressing comparative and superlative degree (*ʾafʿāl at-tafḍīl*). They govern their object in the genitive with the particle *min*, e.g. *zaydun ʾafḍalu min ʿamrin* 'Zayd is better than ʿAmr'.

When these adjectives remain in the constant form, without the definite article, they govern their agent in the nominative only if they are preceded by an interrogative or negative particle, and when they can be replaced by a verb, e.g. *mā raʾaytu rajulan ʾaḥsana fī ʿayni-hi l-kuḥlu min-hu fī ʿayni zaydin* 'I did not see a man in whose eyes the antimony would be more perfect than in Zayd's eye'. Here, *al-kuḥlu* is put in the nominative by *ʾaḥsana*, which can be replaced by the verb *ḥasuna* 'was lovely'.

When they express the superlative degree, these adjectives govern their objects in the genitive. They remain either in the constant form, e.g. *az-zaydāni ʾafḍalu l-qawmi* 'two Zayds [are] the best [man] from the entire tribe', or agree in gender and number, e.g. *az-zaydāni ʾafḍalā l-qawmi*. When used attributively, they agree in gender and number with a definite noun, e.g. *zaydun al-ʾafḍalu* 'the best Zayd', *hindun al-fuḍlā* 'the best Hind' (Zamaxšarī, *Mufaṣṣal* 232–237; Ibn ʿAqīl, *Šarḥ* III, 181).

vi. The annexed noun (*ism muḍāf*) is treated by some grammarians as a regular governor, since it governs its object in the genitive, whereas others assume that the second noun is governed in the genitive by a deleted particle, *li-*, *min*, or *fī* (Ibn ʿAqīl, *Šarḥ* III, 43; cf. Owens 1988: 153–154).

vii. Complete nouns (*asmāʾ tāmma*) are called thus because while governing they lose neither *tanwīn* nor the final *nūn* of dual and plural. These are nouns indicating a measure or weight and the cardinal numerals of tens, which govern their objects in the accusative of specification, e.g. *raṭlun zaytan* 'a pound of olive oil', *ʿišrūna dirhaman* 'twenty dirhams', etc. (ʿAbd al-Laṭīf n.d.: 168; Carter 1972).

3. ABROGATING GOVERNORS

An initial word (*mubtadaʾ*) or topic is used in the nominative by the initial place it occupies in the sentence, being independent from any preceding governor. But frequently topics are preceded by various grammatical governors affecting their declension. These are called → *nawāsix al-ibtidāʾ* 'abrogators of initiality', because they cancel the effect of the initial position (Junaydī 1981:992). They are overt governors that govern an initial word in the accusative, and its predicate in the nominative; govern a predicate in the accusative, and a topic in the nominative; or govern both of them in the accusative (→ *ibtidaʾ*).

3.1 *Overt governors governing a topic in the accusative and its predicate in the nominative*

i. The particle *ʾinna* and its 'sisters' (→ *ʾinna wa-ʾaxawātu-hā*): *ʾanna* 'that', *ka-ʾanna* 'as if; as though', *lākinna* 'but', *layta* 'if only', and *laʿalla* 'perhaps'; these are referred to as particles resembling verbs (*al-ḥurūf al-mušabbaha bi-l-ʾafʿāl*) because they have the same meaning as the verbs *ʾakkada* 'to assure', *tamannā* 'to wish', etc.; like them, they have a final vowel *-a* and require a noun after them (Ibn Mālik, *ʾAlfiyya* 13–14).

In a nominal sentence, these particles govern a topic in the accusative and its predicate in the nominative, e.g. *ʾinna zaydan ʾaxū-ka* 'verily Zayd is your brother'. The topic is referred to as the noun of *ʾinna* (*ism ʾinna*) and its predicate as the predicate of *ʾinna* (*xabar ʾinna*). Basran grammarians compare the phrase *ʾinna zaydan ʾaxū-ka* with *ḍaraba zaydan ʾaxū-ka* 'your brother struck Zayd'. Kufan grammarians claim that these particles do not affect the declension of the predicate, since it remains in the nominative for the same reason as in the phrase *zaydun ʾaxū-ka* 'Zayd [is] your brother' (Ibn al-ʾAnbārī, *ʾInṣāf* 81–84; Zamaxšarī, *Mufaṣṣal* 27).

The noun of *ʾinna* should always precede its predicate, except when this is an adverbial modifier of place or time or a locative adverbial, as in *ʾinna fī d-dāri zaydan* 'truly Zayd [is] in the house'. Therefore, Basran grammarians stipulate that nouns coordinated to the noun of *ʾinna*, when they are mentioned before the predicate, should be governed in the accusative, e.g. *ʾinna zaydan wa- ʿamran qāʾimāni* 'verily, Zayd and ʿAmr [are] standing'. If the coordinated noun is mentioned after the predicate, it can agree with the noun of *ʾinna* in the accusative, e.g. *ʾinna zaydan qāʾimun wa-ʿamran*, or in the nominative *wa-ʿamrun*, since the noun of *ʾinna* has the same case in the underlying structure. The Kufans allow agreement of coordinated nouns in the nominative, before the predicate is mentioned, *ʾinna zaydan wa-ʿamrun qāʾimāni* (Ibn al-ʾAnbārī, *ʾInṣāf* 85–87).

The Basran grammarians believe that the particles *ʾinna* and *ʾanna* in the 'light' form (*muxaffafa*), i.e. with deletion of one *nūn*, in some cases keep their governance; according to the Kufans, they lose their governing force, and therefore do not place their noun in the accusative case (Ibn al-ʾAnbārī, *ʾInṣāf* 88–91).

Kisāʾī and Farrāʾ allow the use of *layta* 'if only' with the sense of the verb *tamannā* 'to wish' with two accusatives, e.g. *layta zaydan qāʾiman* 'if only Zayd [was] standing!' (Girgas 1873:113).

ii. The negative particle *lā*, which is called *lā li-n-nafy* 'the *lā* of negation', governs an indefinite noun in the accusative without *tanwīn*, and its predicate in the nominative case, e.g. *lā rajula fī d-dari* 'there is no man in the house'. The predicate is frequently omitted, e.g. *lā baʾsa* 'there is nothing bad'. The Basrans assume that such indefinite words are indeclinable, with a final vowel *-a*; the Kufans regard them as declinable, governed in the accusative (Ibn al-ʾAnbārī, *ʾInṣāf* 161–163).

Attributes of the noun governed by *lā* agree with it either in the accusative without *tanwīn*, e.g. *lā rajula darīfa fī d-dāri* 'there is no lovely man in the house', or with *tanwīn*, *lā rajula darīfan*, or in the nominative, *lā rajula darifun*, since in the underlying structure the noun is governed in this case. Coordinated nouns, being indefinite, are governed in the accusative or in the nominative with *tanwīn*, e.g. *lā ʾabā wa-bnan miṯla marwāna wa-bni-hi* 'there is no father and son like Marwān and his son'. But when the coordinated noun is definite, it is governed in the nominative, e.g. *lā ġulāma la-ka wa-lā l-ʿabbāsu* 'you have neither slave, nor ʿAbbās'.

If the negative *lā* is repeated before a coordinated noun, as in *lā ḥawla wa-lā quwwata ʾillā bi-l-lāhi* 'there is no power and no strength, except with God', when the first noun is in the accusative without *tanwīn*, the coordinated noun can also be governed in the accusative without *tanwīn* or with *tanwīn*, *lā quwwatan*, or in the nominative, *lā quwwatun*. If the first noun is in the nominative with *tanwīn*, the coordinated noun either agrees with it in the same case, *lā ḥawlun wa-lā quwwatun*, or in the accusative without *tanwīn*, *lā quwwata*.

3.2 *Overt governors governing the topic in the nominative and the predicate in the accusative*

i. Verbs similar to *kāna* (→ *kāna wa-ʾaxawātu-hā* '*kāna* and its sisters'): *ṣāra* 'to become'; *ʾaṣbaḥa* 'to be in the morning'; *ʾamsā* 'to be in the evening'; *ʾaḍḥā* 'to be before noon'; *dalla* 'to be by day'; *bāta* 'to stay overnight'; *mā zāla, mā bariḥa, mā nfakka* and *mā fatiʾa* 'to continue to be'; *mā dāma* 'as long as'; and *laysa* 'not to be'. These are referred to as defective verbs (*ʾafʿāl nāqiṣa*), because, unlike other verbs, which need a noun in the nominative to form a complete sentence, these verbs require for completeness of sense a topic in the nominative and a predicate in the accusative, e.g. *kāna zaydun qāʾiman* 'Zayd was standing'. The noun governed by *kāna* is called *ism kāna* 'the noun of *kāna*', and the predicate is referred to as *xabar kāna* 'predicate of *kāna*'.

Predicates of these verbs may be placed before their noun and even precede the verb, e.g. *qāʾiman kāna zaydun*. However, the Basrans believe that verbs with the negative particle *mā* cannot be preceded by their predicate, while the Kufans allow this. Thus, in their opinion, one can say: *qāʾiman mā zāla zaydun* 'Zayd did not cease to be standing'. They reject this possibility for the predicate of *laysa*, e.g. **qāʾiman laysa zaydun* 'Zayd is not standing', which is accepted by the Basrans (Ibn ʿAqīl, *Šarḥ* I, 278).

ii. Verbs similar to *kāda* 'to be almost' (*kāda wa-axawātu-hā* '*kāda* and its sisters'): *ʿasā* 'it could be that'; *kāda, ʾawšaka*, and *karaba* 'to be on the point [of doing something]'; and *ʾaxaḏa, jaʾala*, and *ṭafiqa* 'to begin'. These are referred to as *ʾafʿāl al-muqāraba* 'verbs of proximity', since some of them express the fact that the predicate is close to accomplishment. They govern the topic in the nominative, and the predicate in the accusative, e.g. *fa-ʾubtu ʾilā fahmin wa-mā kidtu ʾāʾiban* 'and so I returned to [the tribe] Fahm, but I was very near not returning' (Wright 1986:II, 106). Examples of this are extremely rare; more often the predicate is expressed by an imperfect verb, e.g. *kāda zaydun yamūtu* 'Zayd almost died', or by the particle *ʾan* 'that' with a subjunctive, e.g. *ʿasā zaydun ʾan yaxruja* 'perhaps Zayd will leave'.

iii. According to the Basrans, the negative particles *mā* and *lā* in the Ḥijāzī dialect govern a topic in the nominative and the predicate in the accusative, since they resemble the negative verb *laysa*, e.g. *mā hāḏā bašaran* 'this [is] not a man', *lā šayʾun ʿalā l-ʾarḍi bāqiyan* 'nothing on earth is eternal'. The Kufans assert that the particles *mā* and *lā* do not govern a predicate; they explain the accusative of the predicate by deletion of a particle (Ibn al-ʾAnbārī, *ʾInṣāf* 76–79).

For this governance the predicate must follow the topic and cannot be separated from it by *ʾillā* 'except for' or any other particle. Furthermore, for the governance of *lā* both topic and predicate must be indefinite. In the Tamīmī dialect the predicate is governed in the nominative, *mā hāḏā bašarun*, which according to Sībawayhi, is more correct, since these particles are not verbs (Girgas 1873:116).

The particle *lāta* 'not' governs in the accusative only nouns of time, e.g. *lāta ḥīna manāṣin* '[there is] no moment of escape'. The Basrans assert that the particle *ʾin* 'not' does not govern, while the Kufans admit its governance in some cases, e.g. *ʾin huwa mustawliyan ʿalā ʾaḥadin* 'he is not dominating anybody' (Girgas 1873: 116).

3.3 *Overt governors governing both the topic and the predicate in the accusative*

These are the verbs similar to *ḏ̣anna* 'to think' (*ḏ̣anna wa-ʾaxawātu-hā*), such as *ḥasiba* 'to consider'; *xāla* 'to imagine'; *darā* and *ʿalima* 'to know'; *raʾā* 'to see, to consider'; *zaʿama* 'to assert', etc. These are called judgment verbs (*ʾafʿāl al-qalb*, lit. 'verbs of the heart'), because they express intellectual actions. Thus, in the phrase *ḏ̣anantu zaydan jāhilan* 'I thought that Zayd was ignorant' *zaydan* is the first object (*al-mafʿūl al-ʾawwal*) of the verb *ḏ̣anna*, and *jāhilan* acts as its second object (*al-mafʿūl aṯ-ṯānī*).

4. GOVERNORS OF THE VERB

According to the grammarians, only the imperfect verb can be declined. The Basrans claim it is used in *rafʿ* (*marfūʿ*) because it replaces a noun, e.g. *zaydun yaktubu* 'Zayd writes', which is equivalent to *zaydun kātibun* 'Zayd [is] writing'. The Kufans believe that the verb is used in *rafʿ* because it does not depend on overt governors requiring *naṣb* or *jazm*. This opinion was shared by later grammarians, such as Ibn Mālik, Ibn al-Ḥājib and others (Girgas 1873:117).

A verb is governed in *naṣb* (*manṣūb*) by *nawāṣib*, i.e. overt governors requiring subjunctive mood. These are: a) *ʾan* and *kay* 'in order to', the negation *lan*, and *ʾiḏan* 'in that case', which govern directly in *naṣb*, e.g. *ʾurīdu ʾan taqūma* 'I wish you to rise'; *lan yaḍriba* 'he will not strike', *ʾiḏan ʾukrima-ka* 'then, I will respect you'; and b) *ḥattā*, *li-*, *ʾaw* 'that; so that'; *fa-*, *wa-* 'and', e.g. *ʾaslamtu ḥattā ʾadxula l-jannata* 'I embraced Islam so as to enter Paradise'; *jiʾtu-ka li-tukrima-nī* 'I have come to you, so that you respect me'; *la-ʾalzamanna-ka ʾaw tuʿṭiya-nī ḥaqqī* 'I will not leave you, until you give me my due' (Zamaxšarī, *Mufaṣṣal* 246–252).

The particle *ʾan* does not govern *naṣb* after verbs expressing certainty (*yaqīn*), e.g. *ʿalimtu ʾan taqūmu* 'I knew that you would rise', since here *ʾan* is considered to be derived from *ʾanna-ka*. After verbs expressing probability (*rujḥān*), it can operate both the subjunctive, e.g. *ḏ̣anantu ʾan taqūma* 'I thought that you would rise', and the indicative, *ʾan taqūmu* (Girgas 1873:118).

According to the Kufans, the particles *kay* and *ḥattā* may be reinforced by *ʾan*, without affecting the following verb, e.g. *jiʾtu li-kay ʾan ʾukrima-ka* 'I came so as to respect you'. The Basrans believe that *kay* is sometimes used as a particle and does not accept *ʾan* (Ibn al-ʾAnbārī, *ʾInṣāf* 230–232, 235–238).

The particle *ḥattā* governs the subjunctive, when the following verb is used with the sense of a future tense, otherwise it loses its governance, e.g. *mariḍa ḥattā lā yarjūna-hu* 'he became so ill that they cannot hope for his [recovery]'. The Kufans assert that *ḥattā* governs the subjunctive directly. The Basrans, however, believe that *ḥattā* governs nouns in the genitive directly, but verbs in the subjunctive through an implied *ʾan*. Likewise, the Kufans assert that *li-* 'in order to' directly governs the verb in the subjunctive, but that it may accept *ʾan* for emphasis, e.g. *mā kāna zaydun li-ʾan yadxula dāra-ka* 'Zayd is unable to enter your house'. Furthermore, they allow a direct object preceding the verb in the subjunctive, e.g. *mā kāna zaydun dāra-ka li-yadxula*. This is unacceptable to the Basrans (Ibn al-ʾAnbārī, *ʾInṣāf* 241–243).

The subjunctive particles *fa-*, expressing a consequence, and *wa-*, expressing simultaneity of action, govern a following verb in the subjunctive only when the preceding verb expresses an order, prohibition, negation, question, desire, or hope, e.g. *zur-nī fa-ʾukrima-ka* 'visit me so that I respect you'; *lā tanhā ʿan xuluqin wa-taʾtiya miṯla-hu* 'do not keep [someone else] from any act while you are doing the same'. The Basrans believe that these particles govern through an implied *ʾan*, while the Kufans explain the subjunctive by the disagreement between the two verbs.

A verb is governed in *jazm* (*majzūm*) by *jawāzim*, i.e. overt governors implying an imperative. There are two categories: a) those that govern one verb in *jazm: lam, lammā, li-* (*lām al-ʾamr*), and *lā*, e.g. *lam yaqum* 'he has not risen yet', *li-yaktub* 'let him write', *lā taḍrib* 'don't hit!'; and b) those that govern two verbs in *jazm*: *ʾin* 'if'; *man* 'the one who'; *mā* 'what'; *mahmā* 'whatever'; *ʾayyun* 'whoever'; *matā, ʾayyāna* and *ʾiḏmā* 'whenever'; *ʾayna*, *ʾaynamā* and *ḥayṯumā* 'wherever'; and *ʾannā* 'in whatever way' (Wright 1986:II, 14). The first verb governed should express a condition (*šarṭ*), the second one the consequence of that condition (*jazāʾ aš-šarṭ* or *jawāb aš-šarṭ*), e.g. *ʾin tukrim-ni ʾukrim-ka* 'if you respect me, I respect you'; *man yaʿmal sūʾan yujzā bi-hi* 'whosoever commits evil, will be punished for it'; *mā tafʿalū min xayrin ya ʿlam-hu l-lāhu* 'whatever good you [pl.] do, God will know about it'.

The second verb also has *jazm*, when the first one expresses a prohibition, negation, question, desire, or hope, e.g. *uṭlub tajid* 'search and you will find'; *lā takfur tadxul al-janna* 'don't be irreligious, and you will enter Paradise'. These verbs are governed in *jazm* by an implied *ʾin* 'if', since the underlying sentence is *uṭlub fa-ʾin taṭlub tajid* 'search, and if you search, you will find'.

Yet, grammarians disagree about the governor causing *jazm* in the second verb of conditional clauses. Thus, some Basrans claim that these governors affect both verbs, others believe that the first verb governs the second one; still others assert that the governing word places the first verb in *jazm*, and this verb, in its turn, governs the second one. According to the Kufans, the verb expressing the consequence is governed in *jazm* by its proximity (*majzūm bi-l-jiwār*) to the first verb expressing the condition (cf. Dévényi 1988). Hence, they believe that if the agent of a verb expressing consequence precedes its verb, the latter should be used in the indicative, e.g. *ʾin taʾtī-nī zaydun yukrimu-ka* 'if you come to me, Zayd will respect you'. Unlike them, the Basrans believe that this does not interrupt governance, so that the verb should be used in *jazm* (Ibn al-ʾAnbārī, *ʾInṣāf* 250–254).

The Kufans claim that *kayfa* 'how' governs a verb in *jazm* just as *ḥayṯumā* and *ʾaynamā* 'wherever', whereas the Basrans reject this (Ibn al-ʾAnbārī, *ʾInṣāf* 262–264).

5. Governors based on usage

The second category of overt governors, based on *samāʿ* are the following:

i. Words called 'prepositions' in the Western tradition are called by the gramarians *ḥurūf al-jarr* or *ḥurūf al-xafḍ* 'particles [governing their objects] in the genitive'. Az-Zamaxšarī calls these *ḥurūf al-ʾiḍāfa* 'particles of connection', because they join verbs with the nouns to which the action passes, e.g. *min* 'from'; *ʾilā* 'to'; *fī* 'in'; *ḥattā* 'up to'; *bi-* 'in, with'; *li-* expressing 'belonging (to)'; *rubba* 'many'; and *wa-* and *ta-* 'by!' (particles that introduce oath). Other 'prepositions' are regarded as nouns, e.g. *ʿalā* 'on, above'; *ʿan* 'from'; *ka-* 'like'; *muḏ* and *munḏu* 'since'; or as verbs, e.g. *ḥāšā, xāla*, and *ʿadā* 'except' (Zamaxšarī, *Mufaṣṣal* 283).

   According to the grammarians, some particles are pleonastic (*zāʾida*), like *bi-* in the phrase *kafā bi-l-lāhi šāhidan* (Q. 48/28) 'God suffices as a witness' (Arberry II, 229). The Basrans believe that the particle *min* 'from' is pleonastic in interrogative and negative sentences, when its object is an indefinite word, e.g. *mā jāʾa-nī min ʾaḥadin* 'nobody came to me'; the Kufans allow the use of *min* in affirmative sentences, acknowledging phrases like *qad kāna min maṭarin* 'it has already rained' (Girgas 1873:122).

   Sometimes, particles governing in the genitive are omitted; this frequently happens with *rubba* 'how many!' and *bi-* in oaths. The Basrans assert that the deleted particle should be replaced by another particle, the *ʿamal* remaining with the deleted preposition, for instance, *wa-*, replacing *rubba*, does not govern by itself the following word in the genitive. The Kufans disagree with this; they also assert that an oath can be used in the genitive governed by the implied particle, even when it is not replaced. In such cases, the Basrans insist that the deleted particle is replaced with an interrogative or another particle, e.g. *hā-l-lāhi* 'by God!' (Ibn al-ʾAnbārī, *ʾInṣāf* 167–171).

   The Kufans assert that if the object of *muḏ* and *munḏu* 'since' is used in the nominative, e.g. *mā raʾaytu-hu muḏ yawmu l-jumʿati* 'I haven't seen him since Friday', it is governed by an underlying verb. The Basrans regard *muḏ* and *munḏu* as topics, the noun following them being their attribute. When they are used as particles, their objects require the genitive (Ibn al-ʾAnbārī, *ʾInṣāf* 165–167).

ii. Particles of exception (*ḥurūf al-* → *istiṯnāʾ*): *ʾillā* 'except for'; *xalā, ʿadā, ḥāšā, laysa, lā yakūnu* 'excluding'; and *ġayrun, siwā, sawāʾun* 'except for'. The grammarians disagree about the government of the excluded noun (*al-mustaṯnā*) after *ʾillā*. The Basrans assume that the noun is governed in the accusative by an underlying verb *ʾastaṯnī* ‘I exclude’, governing through *ʾillā*. Some Kufans assert that *ʾillā* governs the noun by itself, while others, like al-Farrāʾ, suppose that *ʾillā* is formed from *ʾin* (short form of *ʾinna* ‘verily’) and *lā* ‘not’, hence the accusative of the excluded noun in affirmative sentences (in other sentences it agrees with *al-mustaṯnā min-hu* ‘that from which the exception is made’ as a conjunctional apposition) (Ibn al-ʾAnbārī, *ʾInṣāf* 116–118).

The verbs *xalā, ʿadā, ḥāšā, laysa*, and *lā yakūnu* ‘excluding’ govern an excluded noun in the accusative, e.g. *qāma l-qawmu xalā zaydan* ‘all the tribe stood up, excluding Zayd’; the first three govern an excluded noun also in the genitive, *xalā zaydin*. The Basrans believe that *ḥāšā*, expressing exception, is a particle, whereas the Kufans consider it an imperfect verb.

The other particles, being actually nouns, always govern an excluded noun in the genitive and have themselves the same case as an excluded noun after *ʾillā*, e.g. *jāʾa-nī l-qawmu ġayra zaydin* ‘the tribe came to me, except for Zayd’; *mā qāma ġayru zaydin* ‘nobody stood up, except for Zayd’.

iii. Particles of appeal like *yā*, *ʾayā*, and *hayā*, according to az-Zamaxšarī, are used when the person is far from the speaker; whereas *ʾay* and *ʾa* are used when the addressee is close to the speaker (*Mufaṣṣal* 309); the other grammarians, like Ibn Mālik, disagree with this (Girgas 1873:123).

The nominative – in the singular always without *tanwīn* – is used when the addressee (*al-munādā*) is addressed directly by the speaker, no explanatory term of any description being appended to it, e.g. *yā muḥammadu* ‘o Muḥammad’, *yā sayyidu* ‘o sir’, etc. The accusative is used: a) when the addressee is indefinite and not directly addressed by the speaker, e.g. when a blind man says *yā rajulan xuḏ bi-yadī* ‘some man, take my hand; and b) when it is directly addressed by the speaker, but has an explanatory term appended to it, e.g. *yā ʿabda l-lāhi* ‘o ʿAbdallah’, *yā xayran min zaydin* ‘o you that are better than Zayd’, etc. (Wright 1986:II, 85–86).

The particle of appeal is frequently omitted, except in lamentations for the dead (*an-nudba*), e.g. *wa-zaydāh* ‘alas Zayd!’ and in calling for help (*al-istiġāṯa*), e.g. *yā la-zaydin* ‘o Zayd, help [me]!’ (Girgas 1873:123).

iv. Indefinite pronouns (*al-ʾasmāʾ al-mubhama*), i.e. quantitative numerals from 11 up to 99 (except for the tens, which are ‘complete nouns’), interrogative and exclamatory particles *kam* and *kaʾayyin* ‘how much?’, *kaḏā* ‘so-and-so much’, expressing an uncertain number, govern the estimated subjects in the accusative of specification. When expressing a question, *kam* governs the accusative, e.g. *kam dīnāran ʿinda-ka* ‘how many dinars do you have?’ In other cases it governs its object in the genitive, either in the singular or in the plural, e.g. *lā naʿrifu kam rajulin* (or *rijālin*) *ʿinda-ka* ‘we do not know, how many men you have’. If *kam* is separated from its object by other words, this object is governed, in the Basrans’ opinion, in the accusative, e.g. *kam fī d-dāri rajulan* ‘how many men [are] in the house!’ whereas the Kufans believe that the object should be governed in the genitive, e.g. *kam fī d-dāri rajulin*.

Ka-ʾayyin and *kaḏā* govern similarly to *kam*: they require their objects in the accusative; but *ka-ʾayyin* is more often used with the preposition *min*, e.g. *ka-ʾayyin min qaryatin ʾahlaknā-hā* (Q. 7/4) ‘how many a city We have destroyed!’ (Arberry 1996: I, 171), and *kaḏā* is repeated, e.g. *malaktu kaḏā kaḏā* (or *kaḏā wa-kaḏā*) *dirhaman* ‘I had so-and-so many dirhams’. The Kufans allow the government of the object of single *kaḏā* in the genitive, singular, or plural, e.g. *kaḏā ṯawbin* or *ʾaṯwābin* ‘so-and-so many dresses’ (Girgas 1873:124).

v. Interjections (→ *ism al-fiʿl*) govern like verbs, e.g. *hayhāta Zaydun* ‘Zayd is far’, *zaydun* having nominative as the agent of *hayhāta*, which is equivalent to the verb *baʿuda* ‘to be distant’. If a verb governs nominative and accusative, an interjection with its meaning also governs two cases. Thus, in *ḍarābi zaydan* ‘beat Zayd!’, the agent of *ḍarābi* is the implicit personal pronoun in the nominative, whereas *zaydan* is used in the accusative as the direct object of the interjection (Ibn ʿAqīl, *Šarḥ* III, 305).

vi. Verbs of praise and blame (*ʾafʿāl al-madḥ wa-ḏ-ḏamm*) like *niʿma* ‘to be good’, *biʾsa* ‘to be bad’, etc. govern both the first noun expressing the quality and the second one denoting the praised or blamed person in the nominative, e.g. *niʿma r-rajulu zaydun* ‘how excellent is Zayd, as a man!’. *Zaydun* may be analyzed as the topic placed at the end of the sentence, whereas its attribute is the preceding verbal sentence consisting of the verb and the agent; alternatively, *zaydun* may be the predicate of an underlying topic *huwa* ‘he’. One may also say *niʿma rajulan zaydun*, *zaydun* being the agent of the verb *niʿma*, whereas *rajulan* is used in the accusative of specification (Ibn ʿAqīl, *Šarḥ* III, 165).

Most grammarians consider *niʿma* and *biʾsa* verbs, but some Kufans, like al-Farrāʾ, treat them as nouns, since they are used with prepositions, e.g. *niʿma s-sayru ʿalā biʾsa l-ʿayru* ‘what a beautiful trip on such a miserable donkey!’ The Basrans explain this use of the preposition by an underlying attribute after the verb: *niʿma s-sayru ʿalā ʿayrin maqūlin fī-hi biʾsa l-ʿayru* ‘what a beautiful trip on a donkey about which it is said: such a miserable donkey!’ (Ibn ʿAqīl, *Šarḥ* III, 160–161).

Bibliographical references

Primary Sources

Ibn al-ʾAnbārī, *ʾInṣāf* = ʾAbū l-Barakāt ʿAbd ar-Raḥmān ibn Muḥammad Ibn al-ʾAnbārī, *Kitāb al-ʾinṣāf fī masāʾil al-xilāf bayna n-naḥwiyyīn al-baṣriyyīn wa-l-kūfiyyīn*. Ed. Gotthold Weil. Leiden: E.J. Brill, 1913.

Ibn ʿAqīl, *Šarḥ* = Bahāʾ ad-Dīn ʿAbdallāh Ibn ʿAqīl, *Šarḥ ʿalā ʾAlfiyyat Ibn Mālik*. 20th ed. 4 vols. Cairo: Dār at-Turāṯ, 1980.

Ibn al-Ḥājib, *Kāfiya* = Jamāl ad-Dīn ʿUṯmān ibn ʿUmar Ibn al-Ḥājib, *al-Kāfiya*. Būlāq, 1266 A.H.

Ibn Mālik, *ʾAlfiyya* = Jamāl ad-Dīn ʾAbū ʿAbdallāh Muḥammad ibn ʿAbdallāh Ibn Mālik, *Matn al-ʾAlfiyya*. Beirut: al-Maktaba aš-Šaʿbiyya, 1970.

Jirjāwī, *Šarḥ* = Xālid al-ʾAzharī al-Jirjāwī, *Šarḥ al-ʿawāmil al-miʾa fī ʾuṣūl ʿilm al-ʿarabiyya li-ʿAbd al-Qāhir al-Jurjānī*. Ed. Zahrān al-Badrāwī. 2nd ed. Cairo: Dār al-Maʿārif, 1988.

Jurjānī, *Jumal* = ʿAbd al-Qāhir al-Jurjānī, *al-Jumal*. Ed. ʿAlī Ḥaydar. Damascus: Majmaʿ al-Luġa al-ʿArabiyya, 1972.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. 3rd ed. 2 vols. Beirut: Muʾassasat al-ʾAʿlamī li-l-Maṭbūʿāt, 1990.

[Xalīl], *Jumal* = [ʾAbū ʿAbd ar-Raḥmān al-Xalīl ibn ʾAḥmad al-Farāhīdī], *Kitāb al-jumal fī n-nahw*. Ed. F. Qabāwa. Beirut: Muʾassasat ar-Risāla, 1985.

Zamaxšarī, *Mufaṣṣal* = ʾAbū l-Qāsim Maḥmūd ibn ʿUmar az-Zamaxšarī, *Kitāb al-Mufaṣṣal fī ʿilm al-ʿarabiyya*. [Together with: Muḥammad an-Naʿasānī, *Kitāb al-Mufaḍḍal fī šarḥ ʾabyāt al-Mufaṣṣal*.] 1st ed. Cairo: Maṭbaʿat at-Taqaddum, 1323 A.H.

Secondary Sources

ʿAbd al-Laṭīf, Muḥammad Ḥamāsa. n.d. *al-ʿAlāma al-ʾiʿrābiyya fī l-jumla: Bayna l-qadīm wa-l-ḥadīṯ*. Cairo: Dār al-Fikr al-ʿArabī.

ʿAmāyira, Xalīl Aḥmad. 1987. *Fī t-taḥlīl al-luġawī*. Zarqa: Maktabat al-Manār.

Arberry, Arthur J. 1996. *The Koran interpreted*. New York: Touchstone.

Carter, Michael G. 1972. “‘Twenty dirhams’ in the *Kitāb* of Sībawaihi”. *Bulletin of the School of Oriental and African Studies* 35.485–490.

Dévényi, Kinga. 1988. “The treatment of conditional sentences by the mediaeval Arabic grammarians”. *Budapest Studies in Arabic* 1.11–42.

Girgas, Viktor. 1873. *Očerk grammatičeskoj sistemy arabov* [Survey of the grammatical system of the Arabs]. Saint Petersburg: Tipografija Imperatorskoj Akademii Nauk.

Junaydī, Muḥammad. 1981. *aš-Šāmil: Muʿjam fī ʿulūm al-luġa al-ʿarabiyya wa-muṣṭalaḥāti-hā*. Beirut: Dār al- ʿAwda.

Owens, Jonathan. 1988. *The foundations of grammar: An introduction to medieval Arabic grammatical theory*. Amsterdam and Philadelphia: J. Benjamins.

——. 1990. *Early Arabic grammatical theory: Heterogeneity and standardization*. Amsterdam: J. Benjamins.

Wright, William. 1986. *A grammar of the Arabic language*. 3rd rev. ed. Cambridge: Cambridge University Press.

Valeriy Rybalkin (Institute of Oriental Studies, National Academy of Science of Ukraine)

Amharic → Ethiopia

Amman Arabic → Jordanian Arabic

ʿĀmmiyya → Diglossia; Colloquial

Analogy

1. Analogy in affixational morphology

In historical linguistics, analogy is generally defined as a process by which words or morphemes are created or recreated on the model of existing linguistic patterns. Analogy operates independently of sound change and often regularizes irregularities brought about by sound change. The most widely discussed form of analogy is proportional analogy: A is to A’ as B is to B’. It has been suggested that analogy operates in the direction of maximizing ‘semantic iconicity’, i.e. the one-to-one match of meaning (or function) and form (Bynon 1977:35; Anttila 1989:88–108). Generative linguists have analyzed analogic change as the extension of a rule (Anderson 1988). These notions of analogy are complementary and overlapping. All can be helpful to understanding morphological changes

Table 1. Assimilation of the article

	Classical	Egyptian	
labial	*al-bint*	*il-bint*	'the girl'
dental	*aṭ-ṭaalib*	*iṭṭaalib*	'the student'
dental	*ad-dars*	*id-dars*	'the lesson'
alveolar	*as-suuq*	*is-suuʾ*	'the market'
palatal	*aš-šams*	*iš-šams*	'the sun'
velar	*al-jaamiʿa*	*ig-gamʿa*	'the university'
velar	*al-kitaab*	*ik-kitaab*	'the book'

in the affixational morphology of Arabic, although, as will be seen, none can fully capture the possibilities of analogic change within a fixed-pattern, or fixed-output morphological system typical of Arabic and Semitic languages.

The assimilation of the definite article before velars in Egyptian Arabic is the result of an analogic change that can be readily analyzed as rule extension (Table 1). In Classical Arabic the /l/ of the definite article /al-/ assimilates to a following consonant which is articulated between the teeth and the soft palate. (The consonant *jīm* is a possible exception to this rule insofar as it is pronounced as a palatal affricate [dʒ], as is the case in most pronunciations of Modern Standard Arabic. However, it reflects proto-Semitic /g/ and its earlier Arabic pronunciation is a matter of debate). In the speech of some Egyptian Arabic speakers the rule assimilating the definite article has been extended to the velar consonants /k/ and /g/.

Another example, which can be readily analyzed as proportional analogy, involves the reshaping of the passive/reflexive verb derivation in several dialects across North Africa from Egypt to Morocco (Table 2). Classical Arabic has several verb patterns which generally indicate a passive or reflexive of a transitive verb, all of which involve an affix in some position. For the basic, so-called Form I verb, which has the shape CvCvC (assuming no glides or 'weak' consonants are present), there are two reflexive/passive forms. One is formed with prefix /n-/, another with infixation of /-t-/ after the second consonant of the stem. By contrast, the passive/reflexive of so-called Form II and III verbs (basic pattern CvCCvC and CvvCvC, with medial geminate and long vowel respectively) is derived with prefix /ta-/. Many modern dialects have created a new medio-passive of the Form I verb with prefix /t-/, as the Egyptian examples below indicate. (Roman numbers in parentheses indicate the conventional Form numbers of the active.)

The position of the *t-* affix in the intransitive/passive of the Form I verb has been moved from infix to prefix on the analogy of the Form II and III verbs.

A slightly more complicated but nonetheless straightforward proportional analogy affects the development of the imperfect verb conjugation in Maghrebi dialects (Table 3). In Classical Arabic and the Eastern dialects, the singular-plural distinction in the 1st person is marked solely by a difference in prefixes *ʾa-ktubu* 'I write'/*na-ktubu* 'we write'. This contrasts with the situation in 2nd and 3rd person masculine, where the prefix remains the same in the plural, and plurality is indicated by a suffix *-uu(na)*: *ya-ktub-u* 'he writes' *ya-ktub-uu(na)* 'they write', *ta-ktub-u* 'you [sg.] write' *ta-ktub-uu(na)* 'you [pl.] write'. In Maghrebi dialects the 1st person forms have been reshaped on the analogy of the 2nd and 3rd persons masculine. (Feminine forms have been lost except in the 3rd person singular.) This leads to a closer form:function match. Prefixes exclusively indicate person, while the suffix indicates number.

2. ANALOGY IN ROOT AND PATTERN MORPHOLOGY

All of these examples are easily integrated into traditional approaches to analogy because all

Table 2. Reshaping of the passive/reflexive verb derivation

	Classical	Egyptian		
active	passive/reflexive	active	passive/reflexive	
kattaba	*takattaba*	*kattib*	*itkattib*	'to cause to write (II)
kaataba	*takaataba*	*kaatib*	*itkaatib*	'to write to' (III)
kataba	*iktataba*	*katab*	*itkatab*	'to write' (I)

Table 3. Development of the imperfect verb in Maghrebi dialects

Classical sg.		pl.	Moroccan sg.		pl.
yaktubu	::	***yaktubuu(na)***	***yəktəb***	::	***yəktəbu***
			‘he/they write(s)’		
taktubu	::	***taktubuu(na)***	***təktəb***	::	***təktəbu***
			‘you [sg./pl.] write’		
ʾaktubu	::	***naktubu***	***nəktəb***	::	***nəktəbu***
			‘I/we write’		

involve affixation of some kind. Interesting problems emerge, however, when one turns to the so-called root-and-pattern morphology of Arabic (→ derivation; → root). The type of proportion generally assumed in proportion-based analogy is one in which derivatives of the same word are related in a consistent way, like English *big* :: *bigger*. Such a proportion can be stated as a rule (approximately ‘add *-er* to the positive to form the comparative’). But the defining feature of fixed-pattern morphology is precisely that consistency is found not in a proportion or relationship between a base and a derivative, but in a consistent pattern imposed on all forms representing a particular grammatical category regardless of the form of the source word. The Classical Arabic → elative (comparative/superlative) form of the adjective provides a clear illustration. The elative consistently has the syllable and vowel pattern ʾaCCaC(u) regardless of the syllable structure or vowelling of the positive, as the following examples illustrate.

positive		elative	
sahl	>>	*ʾashal*	‘easy’
kabiir	>>	*ʾakbar*	‘big’
ṣabuur	>>	*ʾaṣbar*	‘patient’
jaahil	>>	*ʾajhal*	‘ignorant’

The elative respects the principle of ‘one meaning, one form’ (semantic iconicity), since a single pattern ʾaCCaC conveys the single meaning ‘more/most’, but it violates the principle of proportional relationship between base (positive) and derivative (comparative). Thus, proportion and semantic iconicity are fundamentally at odds in systems of this type. Speakers’ attempts to resolve this tension in one direction or the other have led to various types of morphological reshaping.

The development of the → diminutive is a case in point (Ratcliffe 2001). Unlike the elative, the diminutive in Classical Arabic does not strictly adhere to the principle of the fixed pattern (Table 4). The diminutive has a fixed vowel sequence *u-ay-(i)*. But there are three distinct syllabic shapes CuCayC, CuCayCiC, and CuCayCiiC, depending upon the syllabic shape of the base from which they are derived. Stems with the shapes CvCC, CvCvC, CvvC (triliteral in Arabic script, tri-moraic in phonemic terms) have a diminutive CuCayC. Quadriliteral (quadrimoraic) stems CvCCvC, CvvCvC, and CvCvvC have a diminutive CuCayCiC. Quinquiliteral (quinquimoraic) CvCCvvC stems have the diminutive CuCayCiiC.

If the data offered only these possibilities, we would have to conclude that the principle of proportionality between input and output has triumphed here over the principle of having a strictly fixed pattern for each grammatical category. These diminutives can be described in terms of the proportion

Cvx[..] :: CuCay[..]

(where Cvx indicates the first heavy syllable of the word, and [..] the remainder of the word). This proportion could be stated in terms of a rule infixing the sequence *-ay-* after the first heavy syllable of the input with change of the vowels of the first and last syllable.

The situation is complicated, however, by the fact that some nouns of the shape CvvCvC and CvCCvC, which should form a trisyllabic diminutive (CuCayCiC), also have a bisyllabic diminutive alternant (CuCayC) in Early Arabic. The CuCayCiC pattern is more common in Classical Arabic and is the only productive pattern in modern written Arabic. However, the alternant CuCayC is also cited in medieval grammatical sources, as reported by Wright (1896:282–3); see Table 5.

This variation indicates that two strategies for diminutive formation were in competition in

Table 4. Classical Arabic diminutives

base		diminutive	examples		
CvCC	>>	CuCayC	*kalb*	>>	*kulayb* ‘dog’
CvCCvC	>>	CuCayCiC	*masjid*	>>	*musayjid* ‘mosque’
CvvCvC	>>	CuwayCiC	*ṭaalib*	>>	*ṭuwaylib* ‘student’
CvCvvC	>>	CuCayyiC	*fanaar*	>>	*funayyir* ‘lighthouse’
CvCCvvC	>>	CuCayCiiC	*sulṭaan*	>>	*sulayṭiin* ‘sultan’

Table 5. Residual diminutive patterns (in parentheses) in Classical Arabic

		CuCayC	CuCayCiC	
ḥaariṯ-	>>	*(ḥurayṯ)*	*ḥuwayriṯ-*	‘plowman’
ḥaamid-	>>	*(ḥumayd-)*	*ḥuwaymid-*	‘sour’
miʿṭaf-	>>	*(ʿuṭayf-)*	*muʿayṭif-*	‘coat’
ʾaswad-	>>	*(suwayd-)*	*ʾusaywid-*	‘black’

early Arabic, one proportional, based on maintaining a consistent relationship across the row (between base and derived form), the other semantically iconic, based on maintaining consistency down the column (a consistent diminutive pattern for all nouns regardless of the shape of the base).

consistency ‘down the column’

CvCC	CvvCvC	CvCCvC	>>	CuCayC
kalb			>>	*kulayb*
	ḥaariṯ		>>	*ḥurayṯ*
		miʿṭaf	>>	*ʿuṭayf*

consistency ‘across the row’

CvCC	CvvCvC	CvCCvC	
kalb			>>
	ḥaariṯ		>>
		miʿṭaf	>>
CuCayC	CuwayCiC	CuCaaCiC	
kulayb			
	ḥuwayriṯ		
		ʿuṭayf	

The system has undergone a further development in Moroccan Arabic (Harrell 1962:81–84, Heath 1987:113–133). The residual CuCayC diminutives of quadriliteral nouns have completely disappeared. More surprisingly, the CuCayC diminutive of triliteral nouns has also become extremely rare. By regular sound changes (loss of unstressed short vowels in many environments, reduction of the diphthong /ay/ to /i/) the Classical Arabic diminutive pattern CuCayC should appear in Moroccan Arabic as *CCiC. Yet the reflexes of Classical Arabic CvCC stems (CvCC or CCvC where V is /o/ or /ə/) and CvvC stems (CVC, where V is one of the stable vowels /a/, /i/, or /u/) in fact form a diminutive CCiyəC or CCiCa, as illustrated below.

Moroccan Arabic diminutives

	bab	>>	*bwiyəb* *bwiba*	‘door’
	kəlb	>>	*kliyəb*	‘dog’
	bġəl	>>	*bġiyəl*	‘donkey’
cf.	*bəgra*	>>	*bgira*	‘cow’
	fnar	>>	*fniyər*	‘lighthouse’
	bakit	>>	*bwikət*	‘packet’
	məktub	>>	*mkitəb*	‘pocket’

Undoing the effects of sound change reveals the following subsystem.

Moroccan Arabic diminutive system reconstructed

*CvvC		>>	*CuwayCaa
‘ ‘		>>	*CuwayyiC
*CvCC		>>	*CuCayyiC
*CvCCa		>>	*CuCayCaa
	*CvCvvC	>>	*CuCayyiC
	*CvvCvC	>>	*CuwayCiC
	*CvCCvC	>>	*CuCayCiC

If one contrasts this with the corresponding sub-system in Classical Arabic, the nature of the analogic reshaping which has taken place becomes clear:

Classical Arabic diminutive system
(Wright 1896:166–175)

			CuCayC	CuCayCiC
CvvC		>>	CuwayC	
CvCC		>>	CuCayC	
CvCCa		>>	CuCayCa	
	CvCCvv	>>		CuCayCaa
	CvCvvC	>>		CuCayyiC
	CvvCvC	>>	(CuCayC)	CuwayCiC
	CvCCvC	>>	(CuCayC)	CuCayCiC

It appears that at some point in the development of Moroccan Arabic speakers interpreted the diminutive as a strict case of fixed-output morphology, thus giving rise to the appearance of allomorphy and pressure for eliminating one of the two 'patterns'. Instead of trying to extend the CuCayC pattern by deleting stem material (as had been possible at an earlier stage of the language), they extended the CuCayCiC pattern by supplying a default consonant (/y/ or a copy of C2) or a default stable vowel (/a/) to fill out the final syllable of the pattern.

A parallel process has affected the broken plural of feminine nouns in Moroccan Arabic (as well as other dialects, to some extent) (Ratcliffe 2002, 2003). In parallel with the diminutive, triliteral (trimoraic) singulars generally have bisyllabic plurals (the most frequent plural patterns for feminine CvCCa singulars being CvCaC and CiCaaC, and for masculine CvCC singulars CuCuuC, ʾaCCaaC, and CiCaaC), and quadriliteral singulars (CvCCvC) have trisyllabic plurals (CaCaaCiC). Just as most etymological CvCC singulars have come to take a trisyllabic diminutive in Moroccan Arabic, so some feminine CvCCa singulars have come to take a trisyllabic plural. Two new plural patterns have developed. One has the shape CCayəC (*CaCaayiC, if the effects of sound changes are undone).

Classical

fitna(t)	>>	*fitan*
šafa(t)	>>	*šifaah*, *šafawaat*

Moroccan

fətna	>>	*ftayən*	'riot'
šəffa	>>	*šfayəf*	'lip'

The source of analogy for these form are feminines with a long vowel in the second syllable, which regularly have the trisyllabic plural CaCaaiC in Classical Arabic. This pattern becomes CCayəC in Moroccan Arabic by regular sound change:

Classical

xizaana >> *xazaaʾin*

Moroccan

xzana	>>	*xzayən*	'cupboard'

The second and more frequent new plural pattern is CCaCi (*CaCaaCiy, if sound changes are undone). The pattern CaCaaCiy of course exists in Classical Arabic as the regular CaCaaCiC plural of four-consonant nouns whose last consonant happens to be /y/ (*kursiiy*>> *karaasiiy* 'chair'). But this is pattern CaCaaCiC with the final /y/ reflecting the fourth consonant of the singular. In the Moroccan Arabic feminine case there is no fourth consonant in the singular, and the /y/ of the plural is in effect a kind of suffix or pseudo-suffix supplied simply to fill out the pattern. The irregular development of the feminine suffix *tāʾ marbūṭa* plays a central role in this development. The suffix is preserved as stable /a/, not deleted or reduced to /ə/ as expected by regular sound change. Singular feminine nouns with the shape CvCCa thus acquire the same phonological shape as nouns with an etymological long vowel or glide in the final syllable: CvCCvv or CvCCvG (glide). These latter provide the likely source of analogy, although they are notably less frequent in Classical Arabic than CvCCa nouns which undergo the analogical change.

Classical

rukba	>>	*rukab*
kulwa	>>	*kula(y)*
qamara	>>	*qamaraat*

Moroccan

rokba	>>	*rkabi*	'knee'
kəlwa	>>	*klawi*	'kidney'
gəmra	>>	*gmari*	'moonlight'

The likely source of analogy:

Classical

malqa(y)a	>>	*malaaqi(y)*
miqla(y)a	>>	*maqaali(y)*

Moroccan

məlqa	>>	*mlaqi*	'junction'
məqla	>>	*mqali*	'frying pan'

If the relevant subsystems in Classical and Moroccan are compared, it becomes clear that the principle of imposing a fixed pattern CCaCəC has begun to take precedence over the principle of maintaining a proportional relationship between input and output.

Classical

sg.		pl.
CvCCa	>>	CvCaC
	>>	CiCaaC
CvCvvCa	>>	CaCaayiC
CvCCa(y)a	>>	CaCaaCi(y)
CvCCvC	>>	CaCaaCiC

Moroccan

sg.		pl.	
CvCCa	>>	*CvCaayiC	(CCayəC)
	>>	*CvCaaCiy	(CCaCi)
CvCvvCa	>>	*CaCaayiC	(CCayəC)
CvCCa(y)	>>	*CaCaaCiy	(CCaCi)
CvCCvC	>>	*CaCaaCiC	(CCaCəC)

Thus in both the plural and the diminutive in Moroccan Arabic the fixed pattern has triumphed over proportion. While traditional proportional analogy works 'across the row', extending a proportional relationship between words which share a feature of specific semantic reference, these examples reveal a type of analogy in Arabic which works 'down the column', imposing a similar shape on words which share a feature of grammatical function or an abstract semantic feature.

In other systems and other dialects, however, proportion triumphs over the fixed pattern to give an 'across the row' analogy. One example comes from Iraqi (Muslim Baghdadi) Arabic (Ratcliffe 2003). In Classical Arabic, nouns with more than four consonants must either take an external ('sound') plural or be brought into conformity with the CaCaaCiC pattern by loss of a high sonority consonant or of the final consonant (e.g. *ʿankabuut* >> *ʿanaakib* 'spider'; *zanbarak* >> *zanaabik* 'metal spring'). In Iraqi Arabic five-consonant nouns CvCCaCaaCa have the newly innovated plural CvCCaCaayiC. Furthermore, four consonant nouns CvC(a)CaaCa ending in -aaCa have a plural CvC(a)CaayiC, rather than the expected Classical Arabic CvCaaCiiC. It is clear that the source for this analogy are forms like CvCaaCa, which in Classical Arabic normally have the plural CaCaaʾiC, regularly becoming CaCaayiC in Iraqi.

Iraqi

CvCaaCa	CaCaayiC	
zabaana	*zabaayin*	'casing'
sidaara	*sidaayir*	'cap'
sidaana	*sidaayin*	'clay container'
CvCCaaCa	CvCCaayiC	
karxaana	*karxaayin*	'factory'
CvCaCaaCa	CvCaCaayiC	
qarawaana	*qarawaayin*	'metal serving bowl'
CvCCaCaaCa	CvCCaayiC	
tallaxaana	*tallaxaayin*	'casino'
xastaxaana	*xastaxaayin*	'hospital'

These forms are all related by a proportion that can be stated as [..]CaaCa :: [..]CaayiC. For words which have a final sequence -CaaCa, this is converted to -CaayiC in the plural, and the rest of the word is ignored.

A somewhat similar phenomenon occurs in → Maltese (Mifsud 1994). Nouns ending in a geminate followed by /a/, regardless of length, form a plural ending in a -CeC sequence.

fidda	*fided*	'silver'
mħadda	*mħaded*	'pillow'
qawsalla	*qawsalel*	'rainbow'
kaxxa	*kaxex*	'box'
vlegga	*vlegeg*	'arrow'
skwerra	*skwerer*	'set-square'
biccilla	*biccilel*	'lace'
umbrella	*umbrelel*	'umbrella'
pozambrella	*pozambrelel*	'umbrella stand'

The proportion relating these forms can be stated as [..]CCa :: [..]CeC. It is clear that the basis for the analogy here is the regular plural for feminine triconsonantal nouns CvCCa >> CvCaC, like Classical Arabic *ḥujja* >> *ḥujaj* 'proof'.

While each of the analogic changes affecting plurals discussed above leads to the regularization of a particular subsystem, the overall, pan-systemic, pan-dialectal effect is greater variation. It has been suggested that this type of give-and-take between pattern-based and proportional analogy, possibly operating in different prehistoric dialects, may have given rise to the great variety of 'broken' plural formations found in Classical Arabic and related languages (Ratcliffe 1998). More generally it has been suggested that the richness of the Classical Arabic morphology vis-à-vis other Semitic languages may be due to creative analogic processes rather than to conservative retentions (Carter 1996).

In addition to being a source for the creation of new morphological patterns and redistribution of old ones, pattern-based analogy may also be the basis for the creation of new words and new roots. Since fixed pattern morphology works by imposing a fixed syllabic shape on words representing a particular grammatical category, it also incidentally imposes the requirement that those words have at least three consonants. The famous triconsonantalism of Arabic thus follows from the nature of the morphological processes in the language. For words which are 'defective' – lacking a third consonant in part of their paradigm – or which have become defective as a result of sound change, there is a strong diachronic tendency to generalize a three-consonant root throughout the paradigm. (This applies only to nouns and verbs. Parts of speech which do not enter into derivation, such as prepositions, conjunctions, pronouns, etc., are not affected.) 'Primitive' biconsonantal nouns in Classical Arabic are supplied with filler consonants in their derived forms in a somewhat idiosyncratic way. (→ biradicals)

sg.	pl.	relational adjective	
fam	*ʾafwaah*		'mouth'
dam	*dimaaʾ*	*damawiy*	'blood'
šifa	*šifaah*	*šafawiy*	'lip'

In the modern dialects many of these biconsonantal nouns have been reshaped and regularized, usually on the analogy of stems with a final geminate, like *muxx*, pl. *mixaax*, *ʾamxaax* 'brain'.

Moroccan

fomm	>>	*fmam*, also *fwam* (Harrell 1962:117)
dəmm	>>	*dmayat*
šəffa	>>	*šfayəf*

Iraqi

damm	>>	*dammaat, dumuum*
šiffa	>>	*šifaaf*, *šifaayif*

Egyptian

fumm	>>	*ʿifmaam*, *fimaam*
šiffa	>>	*šifaayif, šifaf*

Sometimes this analogical reshaping produces variant roots of what were originally the same word. Thus, the Moroccan plural alternants for mouth appear to be based on two different roots: *f-m-m* (found also in other dialects) and *f-w-m*, likewise the word for blood with *d-m-m* and *d-m-y*. Two-consonant stems which have emerged as the result of sound change are also subject to this type of analogical reshaping. In Moroccan Arabic a regular sound change deletes word initial syllables beginning with glottal stop. Thus the Classical Arabic *ʾanf* 'nose', pl. *ʾaanaaf* or *ʾunuuf*, root *ʾ-n-f* emerges in Moroccan Arabic as *nif*, pl. *nyuf*, with a root *n-y-f*.

The situation in the verb is more complex. One small class of verbs has a simple biconsonantal -CvC- stem in the imperfect, but acquires an initial *w-* in the perfect and in the derived forms.

Classical

ya-ṣil-u	*waṣal-a*	*yu-waṣṣil-u*
'he arrives'	'he arrived'	'he connects'
waṣṣal-a		
'he connected'		

In most dialects these verbs have been regularized on the analogy of regular three-consonant verbs so that the /w/ is maintained throughout the paradigm:

Moroccan	
yəwṣəl	*wṣəl*
Egyptian	
yiwṣal	*wiṣil*
Iraqi	
yooṣal	*wuṣal*

Another class of verbs has the stem shape CvvC, with two consonants and a long vowel in both the imperfect and the perfect. These are generally analyzed as containing a medial glide consonant (/w/ or /y/) which has been deleted by (morpho)phonological rules.

Classical

ṭaal-a	*ya- ṭuul-u*	'become long'
daax-a	*ya-duux-u*	'be dizzy'

In Egyptian Arabic many of these verbs have been regularized, with a glide in second position, functioning as a stable consonant (Carter 1996). Sometimes, the older form is maintained with a different meaning.

ṭiwil	*yiṭwal*	'become long'
ṭaal	*yiṭuul*	[idiomatic and auxiliary uses]
dawax	*yidwix*	'make dizzy'
daax	*yiduux*	'be dizzy'

Another well-known case, which leads to a word fracturing into two roots through analogic reshaping, is the development of geminated Form I verbs in all dialects outside the Arabian Peninsula (Ferguson 1959). In Classical Arabic the geminate in such forms remains intact where syllable structure constraints allow it, but is split in other environments, notably when the stem is followed by a suffix beginning with a consonant.

Classical	3rd sg.	1 sg.	
	ḥabba	*ḥababtu*	'he/I loved'

In the dialects the geminate resists splitting, and there develops a stem variant with a diphthong /ay/ (which is etymological) between the stem and the suffix.

Egyptian	3rd sg.	1 sg.	
	ḥabb	*ḥabbayt*	'he/I loved'

What has happened here is that geminate Form I verbs have been reanalyzed and partly recategorized on the analogy of Form II verbs (with derived gemination) of roots with a final glide like *ṣallaa* 'pray'.

Classical

root	*ḥ-b-b*	*ṣ-l-y*
pattern	CaCvCa (I)	CaCCaCa (II)
	ḥabba	*ṣallaa*
	ḥababtu	*ṣallaytu*
	'to love'	'to pray'

Dialect

root	*ḥ-b-b*	*ḥ-b-y*
pattern	CaCvC (I)	CaCCaC (II)
	ḥabb	*ḥabbayt*
	'to love'	'to love'

At the same time that biconsonantal and defective verbs are being reshaped on a triconsonantal pattern, new classes of biconsonantal words emerge in some dialects as a result of sound change. The loss of glottal stop in initial syllables in Moroccan Arabic should have given rise to a new set of two-consonant verbs from Classical Arabic forms like *ʾaxaḏa* 'to take', *ʾakala* 'to eat'. In fact, these verbs have been reshaped in Moroccan Arabic as follows: in the perfect the stems are reshaped on the model of verbs with a stem final glide such as Moroccan Arabic *rma/rmit* 'he threw'/'I threw' (= Classical Arabic *rama(y)a/ramaytu*) and, in effect, acquire a new third consonant. In the participle and the Form II verb, a new first consonant /w/ has been supplied (possibly on the analogy of *w-* initial verbs, discussed above).

Two new Form II verbs have developed in Moroccan Arabic (Heath 1987)

Classical

imperfect	*yaʾkulu*	*yaʾxuḏu*
perfect	*ʾakala*	*ʾaxaḏa*
participle	*ʾaakil*	*ʾaaxiḏ*
Form II perfect	*ʾakkala*	*ʾaxxaḏa*

Moroccan

imperfect	*yakul*	*yaxud*
perfect	***kla***	***xda***
participle	***wakəl***	***waxəd***
Form II perfect	***wakkəl***	***waxxəd***

In effect two new 'roots' of the shape *C-C-y* and *w-C-C* have developed by analogy from stems with initial glottal stop, in response to a sound change deleting /ʾ/. These are *k-l-y* and *w-k-l* from *(ʾ)-k-l* and *x-d-y* and *w-x-d* from *(ʾ)-x-ḏ*.

In the prehistory of Arabic it is possible that the variety of roots associated with the meaning 'one' in Classical Arabic may have come about through a comparable process of forcing an original bi-consonantal stem into triconsonantal patterns (Ratcliffe 2001):

root	word	
ʾ-ḥ-d	*ʾaḥad*	'one' [noun]
w-ḥ-d	*waaḥid*	'one' [adjective, formally participle CaaCiC]
ḥ-d-y	*ḥaadiy ʿašara*	'eleventh' [also participle CaaCiC]

Indeed much of the phenomenon of 'root variation' across Semitic (pairs like *d-w-k* and *d-k-k* both meaning 'to crush') discussed by Voigt (1988) and Zaborski (1991) may owe its development to analogic processes of the type discussed here.

(Except where otherwise noted, the sources for the dialect forms cited in this article are: Egyptian, Hinds and Badawi 1986; Iraqi, Woodhead and Beene 1967; Moroccan, Harrell and Sobelman 1966.)

BIBLIOGRAPHICAL REFERENCES

Anderson, Stephen. 1988. "Morphological change". *Linguistics: The Cambridge survey*. I. *Linguistic theory: The foundations*, ed. Frederick Newmeyer, 146–191. Cambridge: Cambridge University Press.

Anttila, Raimo. 1989. *Historical and comparative linguistics*. Amsterdam and Philadelphia: J. Benjamins.

Bynon, Theodora. 1977. *Historical linguistics*. Cambridge: Cambridge University Press.

Carter, Michael. 1996. "Signs of change in Egyptian Arabic". *Understanding Arabic: Essays in honor of El-Said Badawi*, ed. Alaa Elgibali, 137–143. Cairo: The American University in Cairo Press.

Ferguson, Charles A. 1959. "The Arabic koine". *Language* 35.616–630.

Harrell, Richard S. 1962. *A short reference grammar of Moroccan Arabic*. Washington, D.C.: Georgetown University Press.

—— and Harvey Sobelman (eds.). 1966. *A dictionary of Moroccan Arabic: Arabic-English*. Washington, D.C.: Georgetown University Press.

Heath, Jeffrey. 1987. *Ablaut and ambiguity: Phonology of a Moroccan Arabic dialect*. Albany, N.Y.: State University of New York Press.

Hinds, Martin and El-Said Badawi. 1986. *A dictionary of Egyptian Arabic*. Beirut: Librarie du Liban.

Mifsud, Manwel. 1994. "Internal pluralization in Maltese: Continuity and innovation". *Actes des premières journées internationales de dialectologie arabe de Paris*, ed. Dominique Caubet and Marine Vanhove, 90–105. Paris: INALCO.

Ratcliffe. Robert. 1998. *The 'broken' plural problem in Arabic and comparative Semitic: Allomorphy and analogy in non-concatenative morphology*. Amsterdam and Philadelphia: J. Benjamins.

——. 2001. "Analogy in Semitic morphology: Where do new roots and new patterns come from?" *New data and new methods in Afroasiatic linguistics: Robert Hetzron in memoriam*, ed. Andrzej Zaborski, 153– 162. Wiesbaden: O. Harrassowitz.

——. 2002. "The broken plural system of Moroccan Arabic: Diachronic and cognitive perspectives". *Perspectives on Arabic linguistics*, XIII–XIV, ed. Dilworth Parkinson and Abbas Benmamoun, 87–110. Amsterdam and Philadelphia: J. Benjamins.

——. 2003. "The historical dynamics of the Arabic plural system: Im1plications for the theory of morphology". *Research in Afroasiatic grammar*. II. *Selected papers from the fifth conference on Afroasiatic languages, Paris, 2000*, ed. Jacqueline Lecarme, 339–362. Amsterdam and Philadelphia: J. Benjamins.

Voigt, Rainer. 1988. *Die infirmen Verbaltypen des Arabischen und das Biradikalismus-Problem*. Stuttgart: F. Steiner.

Woodhead, Daniel R. and Wayne Beene. 1967. *A dictionary of Iraqi Arabic: Arabic-English*. Washington, D.C.: Georgetown University Press.

Zaborski, Andrzej. 1991. "Biconsonantal roots and triconsonantal root variation in Semitic: Solutions and prospects". *Semitic studies in honor of Wolf Leslau on the occasion of his eighty-fifth birthday*, ed. Alan S. Kaye, 1675–1703. Wiesbaden: O. Harrassowitz.

ROBERT RATCLIFFE
(Tokyo University of Foreign Studies)

Analytic Genitive

In the analytic or periphrastic genitive, the genitive relation is expressed by means of a particular genitive exponent placed between the noun and its genitival modifier, e.g. *d-dukkān* ***tabaʿ*** *t-tājir* (cf. English 'the shop of the merchant'). The analytic genitive is a dialectal innovation. In *fuṣḥā*, the genitive is of a synthetic kind, hereafter called the synthetic genitive (→ construct state), formed by the juxtaposing of the two terms (and by inflecting the second term in the genitive case), e.g. *dukkān*^{u} *t-tājir*^{i} (cf. English 'the merchant's shop'). Although the synthetic genitive can be resolved by prepositions functioning in a way similar to the analytic genitive exponents, the analytic genitive is not formalized in the written language.

The analytic genitive is a typological phenomenon that reflects internal and external influences on the language structure. Since the synthetic genitive also exists in the same dialects, we are confronted with a choice between two systems. The development of the analytic genitive is a complex process operating on several levels of a speech or a text, depending on phonetic, syntactic, and stylistic (rhetoric) factors, affected by the socio-geographical and social surrounding as well as by the situation of the actual speech act. It is restricted by a teleological ambition toward economy, and a structuring of semantic categories, the principles of which are not yet fully understood. It has been referred to as a phenomenon of general drift (Ferguson 1989:7).

Most dialects possess a genitive exponent. A great number of exponents occur, each one typical of a particular geographical area. The variations in the choice of genitive exponent reflect different stages of language development, in addition to being criteria for distinguishing the dialects from each other. They may be roughly defined as belonging to one of two groups:

i. etymological substantives denoting 'possession' or 'property'. These exponents may or may not agree in number and gender with the preceding noun, as in the following examples: *žna* (Jordan); *gey(y)* (Jordan, Algeria), *gī* (Algeria), *gīt* (Egypt); *ḥāl* (Chad); *hana*, *hen*, *hine*, *hine*; *hila* (Chad), *hīl* (Sudan); *mtāʿ* (*matāʿ*, *mataʿ*, *matā*, *mata*, *mita*, *mtaʿ*) (North Africa); *ntāʿ* (*ntaʿ*), *tāʿ*, *taʿ*, *tīʿ* (North Africa); *tabaʿ* (Greater Syria); *btāʿ* (Egypt); *māl* (Iraq); *šuġl* (Negev); *ḥaqq* (Arabian Peninsula, Sudan); *hnīn* (Upper Egypt); *hūl* (Sudan);
ii. etymological relative pronouns or particles built up from relative and/or demonstrative elements. These exponents are invariable in number and gender, e.g.: *d-; di; dēl ḏēl, ḏēla, ḏīl, ḏīla, dyāl, dyal; eddi, elli; allīl; l-, lē, līt, lēl, līl; šat, šayt, šēt, šīt; t-, tā, ta, te, tel.*

Exponents of the *d-* and *dyāl* types and the *eddi* and *elli* types are typical of Morocco and Algeria. Other exponents in this group are used, more or less frequently, in the eastern Arabic countries of Syria, Palestine, and Iraq, including the isolated dialects of Cypriot Maronite Arabic (*te, tel, šayt, šat*) and the Arabic of Daragözü, Turkey (*lē, līl*).

The analytic genitive syntagm consists minimally of a noun + exponent + modifier. The genitive relation is denoted by the exponent. The noun is usually a substantive. The phrase may be extended by the addition of several modifying or coordinated components. The noun and the modifier are extended independently of one another: any component related to the noun immediately precedes or follows the noun, while any component related to the modifier immediately precedes or follows it. Examples include *l-bēt tabaʿ l-malik* 'the king's house'; *l-bēt l-kbīr tabaʿ l-malik* 'the king's big house', *l-bēt tabaʿ l-malik l-kbīr* 'the great king's house'. The synthetic genitive syntagm, on the other hand, is based on the juxtaposition of the noun and the genitival modifier. The noun may be any nominal concept, a substantive or an adjective or (more rarely) an infinitive. The noun and the modifier may not be separated by anything other than the definite article, which defines the whole syntagm, and all modifications except genitival ones must be placed after the modifier. Constructions of the extended synthetic genitive phrase in the dialects, where case endings no longer exist, may thus be ambiguous in meaning, e.g. *dukkān t-tājir l-kbīr* may mean both 'the merchant's big shop' and 'the big merchant's shop'. As compared with the synthetic genitive, the analytic genitive is additively constructed, restricted to express modifications of substantive nouns, and potentially emphatic (being a larger and heavier entity than the synthetic genitive). The synthetic genitive, on the other hand, is synthetically constructed, more economical, has the noun and the modifier more intimately connected, and is better able to express a greater number of semantic categories.

The analytic genitive may be preferred for *formal* reasons: to avoid the complexity or ambiguity of the synthetic genitive syntagm, for example, if the noun has an attribute; if there is more than one coordinated noun or modifier; or if the syntagm contains one or more multiple annexions. Quite often, the analytic genitive is chosen for *stylistic* reasons, the potential emphasis of the heavier construction being used. Munzel (1949) was the first to identify this type of influencing factors in the choice of the analytic construction in the Cairo dialect. Situations in which the analytic genitive is chosen for emphasis may be, for instance, a contrast between two concepts; the introduction of a new theme; a standstill or climax in a chain of events; the tendency in an enumeration to pass from a shorter to a longer phrase; and the use of a (heavier) analytic genitive phrase in final position to counteract the falling intonation.

The centers of the semantic spheres of the analytic genitive and the synthetic genitive are remarkably stable. The center of the analytic genitive sphere consists of loose (often alienable) (non-intimate) relations, such as concrete possession, authority, or other hierarchical relationships between two concrete entities. In the category of qualification, the analytic genitive is used optionally for a diversity of combinations. The center of the synthetic genitive sphere consists of intimate, mutual relations (often inalienable) between concrete or abstract concepts, typically the categories of kinship, body parts, mutual relations such as friendship, and partitive relations. The productiveness of the synthetic genitive sphere is manifested outside the semantic center, in its ability to move freely between the abstract and the concrete, between hierarchical and mutual relationships.

When the frequency of the analytic genitive increases, formal and stylistic factors remain the same, and the expansion is achieved by a gradual exhaustion of the semantic possibilities of the analytic genitive. Thus, the analytic genitive is extended from the category of concrete possession to the more indefinite relations of place and even as far as the category of abstract possession.

There are two general criteria that combine to define the relative status of the analytic genitive. The first is geographical: there is a major dividing line between the areas of Morocco/Algeria, on the one hand, and the areas of eastern North Africa, Greater Syria, Iraq, and the Arabian Peninsula, on the other. In the western region, the analytic genitive tends to be the ordinary way of expressing the genitive. In the eastern region, the analytic genitive is a complement to the synthetic genitive, more or less extensively used.

The historically isolated dialects tend to occupy either of the two extremes. In Malta and Daragözü, the analytic genitive is the ordinary way of expressing the genitive. In most Anatolian dialects and in Uzbekistan, the analytic genitive either does not exist, or is not a productive alternative.

The second criterion is a sociocultural one: in all dialects sufficiently accounted for, a differentiation is made between *madani, qurawi*, and *badawi* dialects. The analytic genitive is most extensively used in the *madani* dialects, less extensively in the *qurawi* dialects, and not used at all, or very rarely, in the *badawi* dialects.

In the majority of dialects, the analytic genitive serves as a complement to the synthetic genitive. It is chosen on the basis of its being additively composed and potentially emphatic. The synthetic genitive is still a fully productive construction.

In the *madani* dialects of Morocco and Algeria, where the analytic genitive tends to be the ordinary way of expressing the genitive, the extension follows the same semantic pattern as in the eastern dialects but has proceeded much further than in the east. The analytic genitive has been transferred to all semantic categories and even introduced into the center of the synthetic genitive sphere, since it may be used for partitive relations. The synthetic genitive is considerably reduced and restricted to a non-productive center of relationships expressing definiteness, adverbs, and lexicalized combinations.

The historical development is imperfectly known, but there are a few facts that may throw light upon the process. Both *mtāʿ* and *btāʿ* are known as exponents from → Middle Arabic texts. *Mtāʿ* may in fact be quite old: derivations of it are used in Malta, which was separated from the Arabic mainland in the 11th century. It may reasonably be assumed to have been brought to Malta with the Tunisian conquerors, perhaps as early as the 9th century, giving the 11th century as a *terminus post quem*, and possibly the 9th century as a *terminus ante quem*.

Accepting the Syro-Lebanese origin of the Cypriot Maronite Arabic exponents *tel* and *šayt*, the corresponding Syro-Palestinian exponents *del* and *šīt/šet* must have existed before the Maronite emigration from Lebanon, giving the 13th century as a *terminus ante quem* (the end of the epoch of the Crusades).

The *d-/dyāl* exponents in North Africa probably appeared much later than the *mtāʿ* exponents, since they are not known from Andalusian Arabic. We know of a *addi yali* from the early 16th century. However, the use of the (longer forms of the) relative pronoun as genitive exponents does occur in Andalusian Arabic. It can be argued that the modern *d-* and *dyāl* forms go back to intermediate forms such as *aldi* and *addi*, and that these were relative pronouns related to *allaḏi* and similar forms, and even that short forms like *di-* were used very early as relative pronouns. The Jews of Fes use *di-* both as a genitive exponent and as a relative pronoun; in Djidjelli, the process has gone so far as to use the *d*-particle to connect the noun with any kind of modifier, "une modalité variable" (Marcais 1956:506, cited in Eksell 1984). The structural prerequisite for the development is a basic similarity between the two syntacto-semantic patterns for expressing nominal modifications, especially between relatival and genitival qualifications. In the eastern region of Syria, Palestine, and Iraq, the historical development is clouded. Late and Modern Aramaic, substrate languages for New Arabic, use *d*-particles, originally relative pronouns, as markers of any modifier of the noun, but it is uncertain how or whether this phenomenon was taken over by Arabic, because *d*-particles (including *tel, dēl*) are normally not productive in the eastern Arabic mainland, and there is no documentation to show that they were ever very frequent.

It is probable that the principles governing the choice of genitive came into existence at an early date, too, stabilizing the analytic genitive as a complement to the synthetic genitive, operating with semantic constraints. There is no sign of a pulsatory process, nor is there evidence that the analytic genitive should once have been more frequent than today.

The sociolinguistic setting obviously plays, or has played, a role in the development of the analytic genitive. The higher pace of change in urban societies and its diversity of linguistic groups may stimulate the phenomenon of periphrastic genitive constructions. Analogically, the analytic genitive may be favored in periods of change. Thus, Versteegh (1984) includes it in the type of phenomena appearing in pidgin languages, in Arabic corresponding to the formative period of New Arabic. The *mtāʿ* phrase, for example, would meet with the demands prescribed for a pidgin feature. The *d*-particles, on the other hand, would be mainly an internal linguistic phenomenon, developing independently of the social setting.

An internal linguistic factor favoring the choice of the synthetic genitive may be its economical construction, and its ability to express any type of genitival relation. As such, it is a highly functional device. It has also proved strong in all other Semitic languages, where it belongs to the basic syntactic features and has survived, in varying degrees, in modern Semitic variants.

Analytic genitive constructions occur in most Semitic languages (see Lipiński 1997). They appear as complementary to the synthetic genitive, sometimes rather rarely, as in Biblical Aramaic *d̠ī*, Mishnaic Hebrew *šel*, and Akkadian *ša*, sometimes more frequently, as in Geʿez *za*, or even as the ordinary way of expressing genitival and other subordinated relations as in Late and Modern Aramaic *d*. The genitival exponents in Semitic in general tend to be formed from demonstrative-relative particles.

BIBLIOGRAPHICAL REFERENCES

Diem, Werner. 1986. "Alienable und inalienable Possession im Semitischen". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 136.227–291.
Eksell Harning, Kerstin. 1980. *The analytic genitive in the modern Arabic dialects.* (= *Orientalia Gothoburgensia*, 5). Ph.D. diss., University of Göteborg.
——. 1984. "On the development of *d*-particles as genitive exponents in Arabic dialects". *Acta Orientalia* 45. 21–42.
——. 1995. "Complexity of linguistic change as reflected in Arabic dialects". *Dialectologia Arabica: A collection of articles in honour of the sixtieth birthday of Professor Heikki Palva*, 63–73. Helsinki: Finnish Oriental Society.
Ferguson, Charles. 1959. "The Arabic koine". *Language* 35.616–630.
Lipiński, Edward. 1997. *Semitic languages: Outline of a comparative grammar*. Leuven: Peeters.
Munzel, Kurt. 1949. *Der Gebrauch des Genitivexponenten im arabischen Dialekt von Ägypten*. Ph.D. diss., University of Erlangen.
Versteegh, Kees. 1984. *Pidginization and creolization: The case of Arabic*. Amsterdam and Philadelphia: J. Benjamins.

KERSTIN EKSELL
(University of Copenhagen)

Anaphora → Pronominalization; Deixis

Anaptyxis

Anaptyxis is defined as the insertion of a short/extra short (non-etymological) vowel between consonants in order to make a word more easily pronounceable. In Arabic, it is also employed to resolve consonantal clusters prohibited by the syllable structure rules, which generally leads to a creation of a new syllable. It is also called → epenthesis; a similar phenomenon in nouns may be called 'nomina segolata' (especially employed with regard to Hebrew).

This phenomenon takes place in various positions; at the beginning of a word, it is usually called → prothesis. It may also take place across morpheme and word boundaries. The most frequent position where such a process is employed is in patterns such as CVCC, and also at the juncture of two morphemes, where there are consonants in contact position.

In Classical and Modern Standard Arabic, anaptyxis is found in weak verbs of the type *marra* 'to pass by': in a paradigm, when a suffix beginning with a consonant is attached, the verb changes to *marar-ta* 'you passed by'. This process is also associated with stress, as the newly created syllable receives stress.

This feature is mainly present at the juncture of two morphemes, such as **katab-tum-hā* > *katab-tum-ūhā* 'you wrote it/these'. Another typical usage is the application to loanwords that do not fit within the syllable structure of Arabic, to resolve a prohibited initial consonantal cluster #CC-, such as **fransā* > *faransā* 'France', etc. (but also *brāġ/brāhā* 'Prague': such exceptions are allowed especially in roots with R2 = r/l; for the role of sonorants cf. also below).

In a similar way, this process manifests itself also as the insertion of a prothetic vowel in a prohibited #CC- cluster (as a rule in the derived verbal Forms VII, VIII, and X, formed by prefixation of a consonant, e.g. **nkataba* > *inkataba* 'to subscribe', etc.).

Other instances of this phenomenon in Classical and Modern Standard Arabic are rare. The opposition of CVCC and CVCVC formations cannot be viewed as anaptyxis in Classical or Modern Standard Arabic, as these formations are in a great majority of cases semantically distinct (e.g. *labbun* 'remaining, staying' vs. *lababun* 'upper part of the chest').

In older Arabic (like Early Arabic), this phenomenon may have been widespread, but the graphemic notation does not allow us to study it (see Hopkins 1984:8).

In the dialects of Arabic, this phenomenon is very frequent with both nouns and verbs. The nature and extent of the process sometimes even serves to classify Arabic dialects, as in the case of the *qəltu* vs. *gilit* dialects (cf. Fischer and Jastrow 1980:26 and passim). With nouns, it usually appears at the end of a word in order to resolve a cluster of two consonants, the most common pattern change being CVCC > CVCVC (**tamr* > *tamer* 'dates'). Such formations seem to operate at the surface level; underlyingly the form can remain without anaptyxis (cf. *isim* 'name' vs. *ism-i* 'my name'; see Abu-Mansour 1991:139).

In this context, a rule of sonority hierarchy is being applied, according to which the amount of sonority has to decrease from left to right in order for anaptyxis to be applied (cf. especially Palva 1965:35; Taine-Cheikh 1988:217–218). One may compare this with the structuring of the consonantal system by Petráček (1971), based on combinations of ±*cons* and ±*voc* features. This rule can be formalized as follows:

$$C_1VC_2C_3 > C_1VC_2VC_3 \quad \text{if} \quad \text{Sonority}(C_2) < \text{Sonority}(C_3)$$

This rule reflects the fact that in one and the same dialect one may find a number of anaptyctic processes, but also a number of constructions that do not undergo such changes (e.g. in Meccan Arabic: **šukr* > *šukur* 'thanks' vs. *kanz* 'treasure', etc.; Abu-Mansour 1991). However, in the relatively rare cases of non-resolved initial clusters (such as *brāġ/brāhā* 'Prague'), the rule seems to be working in the opposite direction.

In verbs, this process is applied to resolve the final CC# cluster (such as **katabt* > *katabit* 'I wrote'), but both types of form can be also found within the verbal paradigm of one dialect (Baghdad dialect: *kitbaw* 'they wrote' vs. *kitabna* 'we wrote').

Sequences of three consonants are prohibited in most dialects of Arabic. Such sequences are usually not found in the lexicon, but may result from a juxtaposition (mostly with a suffixed pronoun). In such an instance, an anaptyctic vowel is inserted, as in the examples from Egyptian Arabic: *kull* + *hum* > *kullu-hum* 'all of them', etc. (Watson 2002:64).

The quality (color) of the anaptyctic vowel is usually governed by vowel harmony (progressive assimilation at CVCC and – C#C types of clusters: **šukr* > *šukur*; regressive at CC- types of clusters: **fransā* > *faransā*).

It is relatively difficult to specify a clear geographical area of Arabic dialects in which this process takes place. One might say that this rule is applied throughout the Arabic world, even in Nubi Arabic creole (Owens 1985:248). For the distribution of this feature in the Arabic dialects, see Fischer and Jastrow (1980).

Bibliographical references

Abu-Mansour, Mahasen H. 1990. "Epenthesis, gemination and syllable structure". *Perspectives on Arabic linguistics*, II, ed. Mushira Eid and John McCarthy, 167–191. Amsterdam and Philadelphia: J. Benjamins.

Fischer, Wolfdietrich and Otto Jastrow. 1980. *Handbuch der arabischen Dialekte*. Wiesbaden: O. Harrassowitz.

Hopkins, Simon. 1984. *Studies in the grammar of Early Arabic based upon papyri datable to before 300 A.H./912 A.D.* Oxford and New York: Oxford University Press.

Owens, Jonathan. 1985. "Arabic dialects of Chad and Nigeria". *Zeitschrift für Arabische Linguistik* 14.45–61.

Palva, Heikki. 1965. *Lower Galilean Arabic: An analysis of its anaptyctic and prothetic vowels with sample texts*. Helsinki: Societas Orientalis Fennica.

Petráček, Karel. 1971. "Die innere Strukturation des phonologischen Systems im Schriftarabischen". *Travaux Linguistiques de Prague* 4.37–40.

Taine-Cheikh, Catherine. 1988. "Métathèse, syncope, épenthèse: À propos de la structure prosodique du ḥassāniyya". *Bulletin de la Société Linguistique de Paris* 83.213–252.

Petr Zemánek (Charles University)

Anatolian Arabic

1. GENERAL

There are three distinct areas in Turkey where Arabic dialects are spoken:

i. The coastal region of the Eastern Mediterranean from Hatay (Antakya) to Mersin and Adana; all the Arabic dialects spoken in this region are linguistically part of the Syrian Arabic dialect area (→ Cilician Arabic).
ii. Parts of Urfa province which are close to the Syrian border; the dialects spoken in this region are a continuation of the Bedouin dialects of the Syrian desert.
iii. Eastern Anatolia, an area comprising the Turkish provinces of Mardin, Siirt, and Diyarbakır. Only the dialects spoken in this easternmost area are called Anatolian Arabic and form the subject of the present entry. Unlike the dialects mentioned under (i) and (ii) above, they are part of the larger Mesopotamian dialect area. In other words, they can be considered as a continuation of the Iraqi Arabic dialects (→ Iraq). All Anatolian dialects are *qəltu* dialects, according to Blanc's (1964) classification of Mesopotamian Arabic. As implied by the epithet *qəltu*, all Anatolian dialects are characterized by a voiceless reflex of the Old Arabic uvular stop *q* and the inflectional suffix *-tu* of the 1st pers. sg. perfect.

The following is a short dialect classification of Anatolian Arabic:

i. Mardin group
    Mardin town (Muslims; Christians, emigrated)
    Mardin villages (Muslims; Christians, emigrated)
    Plain of Mardin (Muslims; Christians, extinct)
    Kōsa and Mḥallami dialects (Muslims)
    Āzəx (Christians, now emigrated)
    Nusaybin and Cizre (Jews, emigrated)
ii. Siirt group
    Siirt town (Muslims; Christians, extinct)
    Siirt villages (Muslims)
iii. Diyarbakır group
    Diyarbakır town (Christians, extinct; Jews, emigrated)
    Diyarbakır villages (Christians, extinct)
    Diyarbakır, Siverek, Çermik, Urfa (Jews, emigrated)
iv. Kozluk-Sason-Muş group
    Kozluk (Muslims; Christians extinct?)
    Sason (Muslims; Christians extinct?)
    Muş (Muslims; Christians extinct?)

Until the beginning of the 20th century, the various *qəltu* dialects in Anatolia were spoken by Christians, Jews, and Muslims, respectively. As can be inferred from the above list, however, not all of the dialects listed can still be found *in situ*, and some, in fact, must now be considered extinct. The latter category comprises dialects which were spoken exclusively by Christians and came close to extinction as early as during the First World War, as a result of the genocide of the Armenians and other Christian groups. The majority of these idioms are irretrievably lost, and only a few could be partly salvaged with the help of survivors. A second category comprises dialects which remained *in situ* until the middle of the 20th century and then gradually disappeared due to the emigration of the speakers. These dialects were spoken by those Christians who had survived the First World War massacres, and by Jews. The emigration of the Christians was triggered by a resumed, or in fact never interrupted, hostility toward the Christian minorities. The speakers who now live in exile in Europe, North America, and other parts of the world are easily accessible to research but they are not likely to retain their native dialects for more than a few decades. The Jews were not persecuted to the same extent as the Christians but they also suffered some measure of discrimination. Most of them emigrated to Israel after its establishment in 1948. Linguistically, their fate is similar to that of the Christians: in other words, in a few decades from now their specific idioms will be irretrievably lost. The group which fared best were the Arabic-speaking Muslims. Most of them are still to be found *in situ* but they too are subjected to an ever increasing pressure for assimilation and therefore their Arabic dialects are not likely to survive the next one or two generations (→ Turkey).

Although one particular dialect of Anatolia, Mardin Arabic, has been known for over 120 years, having first been described in Socin (1882–1883), the vast majority of these dialects have been discovered and first published by

Jastrow and some of his former students (Jastrow 1973, 1978, 1981; Wittrich 2001; Talay 2001). Some areas are still awaiting further fieldwork, especially the mountainous area between Kozluk and the plain of Muş.

All Anatolian Arabic dialects are minority idioms spoken in small linguistic islands. Most of the speakers also know Kurdish (the regional trade language) and Turkish (the official language of the state). The phenomenon of Arabic → diglossia does not exist in Anatolia, because the knowledge of Literary Arabic is restricted to the clergy, and school instruction is in Turkish only. Therefore the position of the 'High variety' is occupied by Turkish, and the 'Low variety', Anatolian Arabic, has remained purely dialectal. The speakers of Anatolian Arabic do not attach any prestige to their own mother tongue and do not make any noticeable efforts to preserve it.

2. Linguistic description

2.1 *Phonology*

2.1.1 Consonants

Table 1 shows the inventory of consonant phonemes in Kinderib (Kəndērīb) Arabic, a conservative dialect belonging to the Mardin group, Mḥallami subgroup, which has been the subject of two recent monographs (Jastrow 2003, Jastrow forthcoming). The variations of this basic system found in other Anatolian dialects are discussed subsequently.

This inventory calls for the following remarks:

2.1.1.1 A number of new phonemes have been introduced into Anatolian Arabic via loanwords from Kurdish, Turkish, and Aramaic: the voiceless bilabial stop /p/, the voiced labiodental fricative /v/, the voiceless affricate /č/, the voiced platal fricative /ž/, and the voiced velar stop /g/. Examples from Kinderib are:

parčāye 'piece' [< Turkish *parça*]
pūš 'dry grass, hay' [< Kurdish *pûš*]
davare 'ramp' [< Kurdish *dever* fem. 'place']
čəqmāq 'lighter' [< Turkish *çakmak*]
çāx 'time, moment' [< Kurdish *çax*]
ṭāži 'greyhound' [< Kurdish *tajî*]
bažž 'non-irrigated land' [< Kurdish *bej* 'land']
gōmlak 'shirt [modern]' [< Turkish *gömlek*]
magzūn 'large sickle [< Aramaic *magzūnā*, cf.
Ṭuroyo *magzūno*]

2.1.1.2 The interdental fricatives /ṯ/, /ḏ/ and /ḏ̣/ (the latter being the joint reflex of Old Arabic *ḍād* and *ḏ̣ā*) have been retained in the vast majority of the Mardin group dialects, e.g. Kinderib *ṯaqīl* 'heavy', *ḏahab* 'gold', *bayḏ̣* 'eggs', *ḏ̣əhər* 'noon'. In the Diyarbakır group dialects they have shifted to the dental stops /t/, /d/ and /ḍ/, e.g. *tnayn* 'two', *axad* 'he took', *abyaḍ* 'white'. In the Kozluk-Sason-Muş group dialects and in the dialect of Āzəx (Mardin group) the interdental fricatives have shifted to the sibilants /s/, /z/ and /ẓ/, e.g. Āzəx *saʿlab* 'fox', *axaz* 'he took', *bayẓ* 'eggs', whereas in the Siirt group dialects they have yielded the labiodental spirants /f/, /v/ and /ṿ/, e.g. *baʿaf* 'he sent', *vahab* 'gold', *bīṿ* 'white ones'.

2.1.1.3 The voiceless uvular stop /q/ has been preserved in most Anatolian Arabic dialects except the Siirt group dialects where it alternates with a glottal stop /ʾ/ and sometimes [Ø] under conditions which have not yet been completely clarified, e.g. Mardin *qāl* 'he said' vs. Siirt *ʾāl* ~ *āl*.

Table 1. Inventory of consonant phonemes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
stop	p b		t d ṭ		k g	q		(ʾ)
affricate				č j				
fricative		f v	ṯ ḏ ḏ̣	š ž	x ġ		ḥ ʿ	h
			s z ṣ					
nasal	m		n					
lateral			l ḷ					
vibrant			r ṛ					
semi-vowel	w			y				

(1) bilabial, (2) labio-dental, (3) apical, (4) palatal, (5) velar, (6) uvular, (7) pharyngeal, (8) glottal

2.1.1.4 The glottal stop /ʾ/ is a marginal phoneme, e.g. Mardin *saʾal* 'he asked'. Word initial [ʾ] is not interpreted as a phoneme but as a phonetic vowel onset; in open word juncture it is sometimes retained or assimilated to the preceding consonant, e.g. Kinderib *əlʾarḏ̣* ~ *əllarḏ̣* 'the ground'.

2.1.1.5 Alongside the lateral /l/ and the vibrant /r/ there exist the emphatic (velarized) counterparts /ḷ/ and /ṛ/, which have a acquired a marginal phonemic status, e.g. Mardin *kara* 'he rented' vs. *kaṛa* (< **kəl-aṛa*) 'he has seen'. Also /b/ and /m/ can have velarized variants. In some Kurdish and Turkish loans with emphatic pronunciation it is debatable whether they should be analyzed with an emphatic consonant or an emphatic *ạ̄* vowel [ɑː], e.g. *ḅāš* or *bạ̄š* 'good' (< Kurdish *baş*).

2.1.2 Vowels

2.1.2.1 Long vowels and diphthongs
Table 2 is typical for all Anatolian Arabic dialects:

Table 2. Inventory of vowels

ī				ū		
	ē		ō		ay	aw
		ā				

The Old Arabic diphthongs *ay* and *aw* have been preserved by and large, although in a few lexical items they may have been monophthongized to /ē/ and /ō/ respectively, e.g. Mardin *bayt* 'house', *mawt* 'death', but *fōq* 'above'. The mid long vowels /ē/ and /ō/ have entered the inventory mainly by the following processes:

i via loanwords from Turkish and Kurdish, e.g. widespread items like *čōl* 'desert', *xōrt* 'young man', *ṭrəmbēl* 'car', *tēl* 'wire'.
ii by lowering of /ū/ and /ī/ in contact with emphatic and back consonants, e.g. Kinderib *ṛōḥ* 'soul; go!', *daqēq* 'flour'.
iii by the so-called → *ʾimāla*, i.e. the conditioned shift of Old Arabic /ā/ to /ē/ when the preceding or following syllable contained an i or ī vowel, e.g. Āzəx *jēməʿ* 'mosque', *jawēməʿ* 'mosques', *dakēkīn* 'shops'.

2.1.2.2 Short vowels
Anatolian Arabic has a system of two short vowels – /ə/ and /a/ – resulting from the unconditioned merger of the Old Arabic short high vowels /i/ and /u/ into /ə/, thus *bənt* 'daughter', *əxt* 'sister' < Old Arabic *bint*, *uxt*. In the Diyarbakır group /ə/ in word final unstressed syllables has the phonetic value [e], e.g. *mayyet* 'dead', *awnek* 'there' (cf. Mardin *mayyət*, *hawnək*). In Siirt, /ə/ in the same position is split into the two allophones [e] and [o], depending on the consonantal environment, e.g. *yāxev* 'he takes', *ṣārot* 'she became', *ləḥoq* 'he reached' (cf. Mardin *yāxəḏ*, *ṣārət*, *ləḥəq*).

In open unstressed syllables the Old Arabic short high vowels /i/ and /u/ or their merged reflex /ə/ have been elided, e.g. *byūt* 'houses' < Old Arabic *buyūt*, *nēšfe* [fem.] 'dry' < Old Arabic *nāšifa*. In loanwords /ə/ may be retained, e.g. *nəkāḥ* 'marriage ceremony'. In the imperfect of the verb, /ə/ in open syllables is always retained, e.g. Kinderib *yəmsək* 'he seizes' > *yəmsəkūn* 'they seize'.

Old Arabic /a/ in open unstressed syllables has in general been preserved, e.g. *fataḥ* 'he opened', *fataḥət* 'she opened', *fataḥu* 'they opened'. In Daragözü (Kozluk-Sason-Muş group), /a/ has been elided in open unstressed syllables and raised to /ə/ in closed unstressed syllables, e.g. *fatḥət* 'she opened', *fətḥə́tu* 'she opened him/it', see Jastrow (1973).

In certain nominal forms /a/ has been elided, probably after an intermediate assimilation to /ī/ of the following syllable, e.g. Old Arabic *kaṯīr* > **kiṯīr* > Mardin, Kinderib *kṯīr* 'much'. In the broken plural forms KaKēKəK and KaKēKīK (with *ʾimāla* vowel /ē/) the /a/ of the first syllable has been preserved in Āzəx and the Siirt dialects but shifted to /ə/ in the dialects of the Mardin group without, however, being elided, thus, e.g., Āzəx *jawēməʿ* 'mosques', *dakēkīn* 'shops' but Mardin *jəwēməʿ*, *dəkēkīn*.

2.1.3 Suprasegmental features

2.1.3.1 Stress
In all Anatolian Arabic dialects stress is usually determined by syllable structure, according to the rule of thumb that stress will fall on vKK or v̄K closest to the end of the word, otherwise on the first v. In addition, there is a morphological rule which requires stress always to be on the last syllable of a nominal or verbal form if a pronominal suffix is added, e.g. Mardin *axaḏ* 'he took' > *axáḏu* 'he took him', *axaḏət* 'she took' > *axaḏə́tu* 'she took him'.

Certain word classes, in particular numerals (including *kəll* 'all'), negations, and interrogatives have a tendency to form a stress unit with the following word, the main stress being on the first element. This is expressed by the symbol = joining the two elements, e.g. *ṯmən=təšhəṛ* 'eight months', *mō=tākəl* 'she does not eat'.

2.1.3.2 Consonant clusters and anaptyxis

Final clusters of two consonants are in general not separated by an anaptyctic vowel, e.g. Kinderib *bəxš* 'hole', *bard* 'cold [noun]', *ḏ̣ayyaˁt* 'you lost'. Only when the last consonant is *l*, *r*, *m*, or *n* an anaptyctic [ə] is inserted, e.g. *baġəl* 'mule', *šahəṛ* 'month', *baṭən* 'belly, stomach'. Although strictly speaking this vowel is not a phoneme in most Anatolian dialects as it cannot be stressed, e.g. *báġəlna* 'our mule', *báṭənki* 'your [fem.] stomach', it is written as a full vowel /ə/ by the present author.

Word initial clusters of two consonants may be preceded by an anaptyctic [ə] vowel (written as a raised ᵊ), e.g. Kinderib *ᵊṣṣēr* 'she/it becomes', *kūjjāb ᵊḥmāṛu* 'he has brought his donkey'. Similarly word internal clusters of three consonants may have an anaptyctic vowel between the first and second consonant; in this case a full vowel is written, e.g. *kūt* + *nqatal* > *kūtənqátal* 'he has been killed'.

A different system of syllabication obtains in Daragözü (Kozluk group) where a word initial KK cluster is realized with an anaptyctic vowel between the two consonants, e.g. *ftaḥ* [fə'taħ] 'open!'

2.1.3.3 Word final devoicing

Voiced consonants in word final position have a tendency to become unvoiced: in the case of stops there can be additional aspiration, e.g. Mardin *axaḏ* [θ] 'he took', *katab* [p'] 'he wrote'. The sonants *l*, *r*, *m*, *n* are not subject to final devoicing. In the dialects of the Mardin group word final /ˁ/ is not subject to devoicing; it is, however, devoiced in the remaining Anatolian groups, e.g. Mardin *yəqṭaˁ* 'he cuts', Siirt *yəqṭaḥ*.

2.2 *Morphology*

The gender distinction in the 2nd and 3rd pers. pl. in verbs and pronouns has not been preserved in Anatolian Arabic, as in all *qəltu* dialects. The former masculine forms have been generalized as the new *communis* forms.

2.2.1 Pronouns

2.2.1.1 Independent pronouns

Table 3 shows the independent personal pronouns in two dialects, Mardin and Daragözü (Kozluk group).

The form *ənta* for the 2nd pers. sg. masc. is found only in Mardin and surroundings, the majority of Anatolian Arabic dialects has *ənt*. The initial *h-* of the 3rd person forms has been elided in the Siirt dialects (*ūwe*, *īye*, *ənne*) and in the Sason dialects which, however, follow the Daragözü model (*īyu*, *īya*, *īyən*); see next paragraph. In the Siirt group *anā* with word final stress is used for the 1st pers. sg.; this explains Daragözü *nā*.

In Daragözü the expected form *hīye* has become *hīya* by analogy to the pronominal suffix 3rd pers. sg. fem. *-a*. The forms *hīyu* and *hīyən* in turn are back formations from *hīya*, by attaching to a basis *hīy-* the respective pronominal suffixes *-u* and *-ən*. The 2nd person forms *ənte* and *ənto* owe their final vowel to the analogy with the inflected verb (see 2.2.6.2.2).

2.2.1.2 Copula

In Anatolian Arabic a copula is used regularly in nominal sentences. It consists of the unstressed and sometimes shortened forms of the independent pronouns which follow the predicate enclitically; in the Siirt group they precede the predicate. In some dialects the 3rd person forms have different allomorphs after vowels and consonants, as shown in the Āzəx forms. The paradigms in Table 4 show the copula with two predicates, one with final consonant (*fə-lbayt* 'in the house') and one with final vowel (*hawne* ~ *awne* 'here').

Table 3. Independent personal pronouns

	Mardin	Daragözü
3rd sg. masc.	*hūwe*	*hīyu*
3rd sg. fem.	*hīye*	*hīya*
3rd pl.	*hənne*	*hīyən*
2nd sg. masc.	*ənta*	*ənt*
2nd sg. fem.	*ənti*	*ənte*
2nd pl.	*əntən*	*ənto*
1st sg.	*ana*	*nā*
1st pl.	*nəḥne*	*naḥne*

Table 4. Copula with predicate

	Mardin	Āzəx	Siirt
3rd sg. masc.	*fə-lbayt-we*	*fə-lbayt-u*	*ūwe fə-lbayt*
	hawne-we	*hawne-we*	*ūwe awne*
3rd sg. fem.	*fə-lbayt-ye*	*fə-lbayt-i*	*īye fə-lbayt*
	hawne-ye	*hawne-ye*	*īye awne*
3rd pl.	*fə-lbayt-ənne*	*fə-lbayt-ən*	*ənne fə-lbayt*
	hawne-nne	*hawne-nən*	*ənne awne*
2nd sg. masc.	*fə-lbayt-ənta*	*fə-lbayt-ənt*	*ənt fə-lbayt*
2nd sg. fem.	*fə-lbayt-ənti*	*fə-lbayt-ənti*	*ənti fə-lbayt*
2nd pl.	*fə-lbayt-əntən*	*fə-lbayt-əntən*	*əntən fə-lbayt*
1st sg.	*fə-lbayt-ana*	*fə-lbayt-ana*	*anā fə-lbayt*
1st pl.	*fə-lbayt-nəḥne*	*fə-lbayt-nəḥne*	*nəḥne fə-lbayt*

In negative sentences a copula formed from a negation **mā* + copula precedes the predicate, e.g. Kinderib *mawwe fə-lbayt* 'he is not at home'. The forms are for Kinderib: 3rd pers. sg. masc. *mawwe*, 3rd pers. sg. fem. *mayye*, 3rd pers. pl. *manne*; 2nd pers. sg. masc. *mant*, 2nd pers. sg. fem. *manti*, 2nd pers. pl. *mantən*; 1st pers. sg. *mana*, 1st pers. pl. *mánəḥne*.

2.2.1.3 Pronominal suffixes

The pronominal suffixes are attached to nouns (to express possession), to verbs (to express a direct verbal object), and to prepositions. They have different allomorphs after vowels and consonants; in some dialects (e.g. Āzəx, Siirt group) the nature of the vowel also matters.

Table 5 shows the pronominal suffixes for Mardin after bases ending in a consonant, in *-ā* or *-ū* (*bayt* 'house', *waddā-* 'he took away', *abū-* 'father'). The forms for Siirt differ slightly from those for Mardin.

In Fəsken (a dialect of the Siirt group) and in the Diyarbakır dialects we find the 3rd pers. sg. masc. suffix *-nu* after *-u* (in Diyarbakır also after *-a*), e.g. Fəsken *abūnu*, *qatalūnu* 'they killed him', but *waddāhu*; Diyarbakır *abūnu*, *qatalūnu*, *waddānu*. The suffix *-nu*, which is also found in the Tigris group of → Iraqi Arabic can be explained as a reanalysis of forms like Fəsken *yəqtəlūnu* (< *yəqtəlūn* + *-u*) 'they kill him' > *yəqtəlū-nu*.

2.2.1.4 Relative pronouns

The relative pronoun is *la-* in Mardin, *lə-* in the majority of the Mardin group dialects. In the Diyarbakır, Siirt, and Kozluk groups we find a form *lē*.

2.2.2 Adverbs

2.2.2.1 Demonstrative adverbs

For 'thus' there are forms harking back to Old Arabic *hā-kaḏā*, e.g. Mḥallami *hāgge* ~ *hāg*, Diyar-bakır *āge* ~ *āg*, and forms harking back to simple **kiḏā*, e.g. Mardin *kəḏe*, Daragözü *kəze* ~ *kəz*.

Table 5. Pronominal suffixes (Mardin)

3rd sg. masc.	*bayt-u*	*waddā-hu*	*abū-hu*, Siirt *abū́*,
3rd sg. fem.	*bayt-a*	*waddā-ha*	*abū-wa*
3rd pl.	*bayt-ən*, Siirt *-en*	*waddā-hən*, Siirt *-hen*	*abū-wən*, Siirt *-wen*
2nd sg. masc.	*bayt-ək*, Siirt *-ok*	*waddā-k*	*abū-k*
2nd sg. fem.	*bayt-ki*	*waddā-ki*	*abū-ki*
2nd pl.	*bayt-kən*, Siirt *-ken*	*waddā-kən*, Siirt *-ken*	*abū-kən*, Siirt *-ken*
1st sg.	*bayt-i*	*waddā-ni*	*abū-yi*
1st pl.	*bayt-na*	*waddā-na*	*abū-na*

Most forms for 'here' hark back to Old Arabic *hā-hunā*, e.g. Mardin *hawne* ~ *hawn*, Siirt and Diyarbakır *awne*. Āzəx has a form harking back to simple *hunā*, i.e. *nna*, with a longer variant *nnane*.

The forms for 'there' hark back to a common Anatolian **hawnak*, cf. Mardin *hawnak* ~ *hawnake*, Diyarbakır *awnek*, Siirt *ạwnok*.

2.2.2.2 Interrogative adverbs

A reflex of Old Arabic *kayfa* 'how' is preserved in the Mḥallami dialects, e.g. Kinderib *kēf*. Most Anatolian words for 'how', however, hark back to **ayš-lawn* 'what color, what kind', e.g. Kōsa dialects *áššōn*, Mardin *ašwan*, Siirt *ayšám* ~ *ašám*. Diyarbakır has *əšṭōr* (< **ayš-ṭawr*, cf. Jastrow 1997), Āzəx has *áštawf*, probably a contamination of **ayš-ṭawr* and *kayfa*.

The word for 'where' is *ayn* in Mardin and Āzəx, a direct reflex of Old Arabic *ʾayna*. In most Anatolian dialects, however, the words used can be tracked back to compound forms like **ayna mōḏ̣aʿ* 'which place' (Diyarbakır *əndaḥ* ~ *ənda*, Daragözü *əmmaḥ* ~ *əmma*), **ayna kēs* 'which direction' (Kōsa, Mḥallami *angəs*) or **ayna ṣawb* 'which direction' (Siirt *aysáb*, with de-emphatization).

The Old Arabic form for 'when', *matā*, survives in Āzəx *mate*, but forms harking back to a compound form **ayy matā* are more frequent, e.g. Mardin *áymate*. Most Anatolian dialects, however, have reflexes of two compound forms, **ayš-waqt* 'which time' and *ayš-čāx* 'which time' (< Kurdish *çax* 'time'), e.g. *áyšwaxt*, *ášwaxt*, *aščax*, *aččax*.

2.2.3 Particles

2.2.3.1 The definite article is *əl-*, the demonstrative article *hal-*, thus *əlbayt* 'the house', *halbayt* 'this house'. The /l/ is usually assimilated to preceding 'sun letters', e.g. *əddəkkān* 'the shop'; outside the Mardin group, however, there is a tendency to keep the /l/ unassimilated.

2.2.3.2 Indefiniteness is expressed in the Mardin group by *wēḥəd*, fem. *wəḥde* 'one'. In Kinderib, e.g., *wēḥəd xōrt* means 'a [some] young man', whereas *əlwēḥəd*, fem. *əlwəḥde* following a noun means 'a certain', e.g. *yawm əlwəḥəd* '[on] a certain day, one day'. Thus *wēḥəd* functions like Iraqi Arabic *fad* (→ article, indefinite). In one part of the Anatolian Arabic area, i.e. the Diyarbakır group, there is a lexeme *faqet* which seems to be a cognate of Iraqi Arabic *fad* (< *fard*) and functions like it, e.g. Kaʿbīye *ləhu faqed tawṛ* 'he has an ox', *faqed mōḍaḥ* 'a certain place'. In the Sason-Muş dialects, enclitic *-ma* functions as an indefiniteness marker, e.g. Hasköy *yōm-ma* '[on] a certain day, one day'.

2.2.3.3 There are two main forms of a genitive marker: *ḏīl(a)* ~ *ḏēl(a)* in most of the Mardin group, with a variant *dēl* in the Diyarbakır group, and *lēl* in Āzəx, *lē* in the Siirt group and Daragözü.

2.2.3.4 Negations

Anatolian Arabic has two different negations for the present and the past tense: *mō* is used with the present tense and in nominal sentences, *mā* with the past tense, e.g. Mardin *mō yəji* 'he does not come', *mā jā* 'he did not come'. In a nominal sentence: Mardin *mō fə-lbayt-we* 'he is not at home' (but Kinderib *mawwe fə-lbayt*, see 2.1.2). Optative and imperative are negated by *lā*, e.g. *lā yəji* 'may he not come', *lā təjawn* 'don't come [pl.]'.

Negations usually form a stress unit with the following noun (2.1.3.1), thus *mā=jā*, *mō=təji*, *lā=yəji*, *lā=təjawn*.

2.2.4 Noun

2.2.4.1 Feminine forms. The nominal feminine ending has two allomorphs: *-a* after emphatic and back consonants and and *-e* otherwise, e.g. *ḥənṭa* 'wheat', *waraqa* 'a leaf; one Turkish pound', but *mədde* 'period of time', *jəbne* 'cheese'.

2.2.5 Numerals

Table 6 shows the numerals from 1–10 and 11–20 in the dialect of Kinderib.

Table 6. Numerals (Kinderib)

1	*wēḥəd*, fem. *wəḥde*	11	*ḥḍaʿš*
2	*ṯnayn*, fem. *ṯəntayn*	12	*ṯṇaʿš*
3	*ṯāṯe* ~ *ṯəṯ*	13	*ṯələṭṭaʿš*
4	*aṛbʿa* ~ *aṛbaʿ*	14	*arbīṭaʿš*
5	*xamse* ~ *xams*	15	*xamṣṭaʿš*
6	*sətte* ~ *sətt*	16	*səṭṭaʿš*
7	*sabʿa* ~ *sabəʿ*	17	*sabīṭaʿš*
8	*ṯmēnye* ~ *ṯmən*	18	*ṯmənṭaʿš*
9	*təsʿa* ~ *təsəʿ*	19	*sāṭaʿš* [sic]
10	*ʿašara* ~ *ʿaš*	20	*ʿəšrin*

wēḥəd is constructed as an adjective but can either precede or follow a noun (see 2.2.3.2). The number two is frequently expressed by the dual which is not restricted to a few time units, e.g. *šahrayn* ‘two months’, *sāˁtayn* ‘two hours’, but can be used with ordinary nouns as well, e.g. *waladayn* ‘two boys’, *bəntayn* ‘two girls’. When combined with a counted noun (always in the plural) the numerals 3–10 have shorter forms without the final vowel. They frequently form a stress unit with the counted noun, the main stress being on the numeral. This is expressed by the symbol =, e.g. *sətt=banāt* ‘six girls’, *ˁaš=ġalamāt* ‘ten goats’.

A small list of nouns which originally began with ʾV in the plural have special count plurals with initial *t-*; they are used after the numerals 3–10, e.g. Kinderib *xams=tālāf* ‘five thousand’, *ṯmən=tīyēm* ‘eight days’, *təsˁ təšhəṛ* ‘nine months’, *xams tərġafe* ‘five loaves’.

The numerals 11–19 have a single form, regardless of whether they are used independently or in connection with a following noun.

2.2.6 Verb

2.2.6.1 Derivation

2.2.6.1.1 Form I

The Anatolian dialects preserve two different vowel patterns in the perfect, reflecting Old Arabic ‘transitive’ CaCaCa and ‘intransitive’ CaCiCa/ CaCuCa patterns, respectively. In the imperfect, the stem vowel is *ə* (< Old Arabic *i*, *u*) or a. Thus, e.g., Kinderib *ḏ̣arab*, *yəḏ̣rəb* ‘to hit, shoot’, *šəṛəb*, *yəšṛab* ‘to drink’.

2.2.6.1.2 Derived forms

Form IV survives only in some rare fossilized expressions, e.g. Mardin *awdaˁnākən* ‘goodbye’, lit. ‘we commend you [to God]’. The internal passive has disappeared.

In the Siirt and Diyarbakır group Forms II, III, V, VI, and X have identical inflectional bases for perfect and imperfect, the last syllable being always vocalized with *ə* (for allophones of /ə/ in these dialects see 2.1.2.2), e.g. Siirt *ˁallem*, *yəˁallem* ‘to teach’, *čāḷoš*, *yčāḷoš* ‘to work’, *tˁawwoq*, *yətˁawwoq* ‘to be late’, *stanṿor*, *yəstanṿor* ‘to wait’; Diyarbakır *ḥaddet*, *yḥaddet* ‘to speak’, *zzawwej*, *yəzzawwej* ‘to get married’, *staxber*, *yəstaxber* ‘to ask’.

Form III in Anatolian Arabic has a long *ē* vowel in the first syllable of both the perfect and the imperfect, e.g. Kinderib *ˁēwan*, *yˁēwən* ‘to help’. The *ʾimāla* was triggered by the vowel *i* in the Old Arabic imperfect form. i.e. **yuˁāwin* > *yˁēwən*. The *ē* was then extended by analogy to the perfect. In a few cases *ē* was even taken over into Form VI (both perfect and imperfect), e.g. Kinderib *tsēwa*, *yətsēwa* ‘to become even, flat’.

Characteristically, Form III in Anatolian Arabic forms causatives for a few verbs of motion, e.g. *dēxal* ‘to make come in, introduce’ to *daxal* ‘to come in’, *ṭēlaˁ* ‘to make come out, take out’ to *ṭalaˁ* ~ *ṭələˁ* ‘to come out’, *qēˁad* ‘to put’ to *qaˁad* ~ *qəˁəd* ‘to sit down’.

2.2.6.2 Inflection

2.2.6.2.1

Table 7 shows the inflection of Form I of the strong verb in Mardin Arabic, ‘transitive’ *katab* ‘to write’ and ‘intransitive’ *šəṛəb* ‘to drink’.

These forms call for several remarks:

i. Note that in the perfect of the ‘intransitive’ verb /ə/ in the first syllable is never elided but /ə/ in the second syllable is. This reflects the fact that /ə/ of the first syllable harks back to **a*, e.g. **šaribtu* > *šəṛəbtu*.
ii. The inflectional morpheme *-tu* of the 1st pers. sg. perfect is an important hallmark of the *qəltu* dialects, both Anatolian and Iraqi.
iii. The inflectional morpheme *-tən* of the 2nd pers. pl. perfect, on the other hand, is an important isogloss distinguishing between the Anatolian and Iraqi branches of *qəltu* dialects. Altogether there are five morphemes which in Anatolian Arabic end in *-n* but in Iraqi *qəltu* Arabic in *-m* (Table 8)
iv. The retention of final *-n* in the imperfect forms 2nd pers. sg. fem., 2nd pers. pl., and 3rd pers. pl. is common in both Anatolian and Iraqi Arabic (in other words, in Iraq it is found in both *qəltu* and *gələt* dialects). The *-n* is dropped when pronominal object suffixes are added, e.g. Mardin *yjībūn* ‘they bring’, *yjībū-hu* ‘they bring him’. The dialect of Āzəx (Mardin group) and the Siirt group dialects drop the final *-n* of the free forms, e.g. *yjībū* ‘they bring’; since the pronominal suffix 3rd pers. sg. masc. in these dialects is Ø after *-ū*, *yjībū* may also mean ‘they bring him’.

Table 7. Inflection Form I (Mardin)

	perfect	imperfect	perfect	imperfect
3rd sg. masc.	*katab*	*yəktəb*	*šəṛəb*	*yəšṛab*
3rd sg. fem.	*katabət*	*təktəb*	*šəṛbət*	*təšṛab*
3rd pl.	*katabu*	*yəktəbūn*	*šəṛbu*	*yəšṛabūn*
2nd sg.masc.	*katabt*	*təktəb*	*šəṛəbt*	*təšṛab*
2nd sg. f.	*katabti*	*təktəbīn*	*šəṛəbti*	*təšṛabīn*
2nd pl.	*katabtən*	*təktəbūn*	*šəṛəbtən*	*təšṛabūn*
1st sg.	*katabtu*	*aktəb*	*šəṛəbtu*	*ašṛab*
1st pl.	*katabna*	*nəktəb*	*šəṛəbna*	*nəšṛab*

Table 8. Differences between Anatolian and Iraqi

Anatolia (Mardin)	Iraq (Christian Baghdad)	gloss
jəbtən	*jəbtəm*	you [pl.] brought
əntən	*əntəm*	you [pl.] – independ. pronoun
baytkən	*bētkəm*	your [pl.] house
hənne	*həmma*	they – independent pronoun
baytən	*bētəm*	their house

2.2.6.2.2 Treatment of final weak verbs

In Anatolian Arabic a distinction is made in the inflection of strong and final weak (III*y*) verbs, as in Table 9.

In a deviation from the Old Arabic pattern, however, the endings *-ayn*, *-awn* have been extended by analogy to inflectional bases of the imperfect ending in *-ī*, e.g. Mardin *təbnayn* 'you [sg. fem.] build', *təbnawn*, *yəbnawn* 'you [pl.], they build', cf. Old Arabic *tabnīna*, *tabnūna*, *yabnūna*.

The dialect of Daragözü has generalized the endings of the final weak verbs for strong verbs as well, as shown in forms like *ẓarbo* 'they hit', *ẓərbawni* 'they hit me' (cf. Mardin *ḏ̣arabūni*). In the imperfect the final *-n* was dropped and, subsequently, the endings were subjected to the analogy of the perfect, thus *īẓrəbo* 'they hit', *īẓrəbawni* 'they hit me'. In a final step, the inflectional suffixes of the 2nd pers. sg. fem. and 2nd pers. pl. perfect were reshuffled according to the same pattern, thus *ẓərabte* 'you [sg. fem.] hit', *ẓərəbtayni* 'you [sg. fem.] hit me', *ẓərabto* 'you [pl.] hit', *ẓərəbtawni* 'you [pl.] hit me'.

2.3 *Syntax*

2.3.1 Noun phrase

2.3.1.1 Definiteness and indefiniteness

In the dialects of the Kozluk-Sason-Muş group there is a tendency to drop the definite article while retaining it before a preposition, e.g. Daragözü *baqər ẓāʿo* 'the cows got lost', *kalb jā* 'the dog came' but *ṭalaʿ mə ddaḥle* 'he came out of the wood'. Corresponding to the loss of the definite article there is an increasing use of postpositional *waʿd*, fem. *waʿde* 'one' to express indefiniteness, e.g. *f-daḥle waʿde* 'in a wood'.

Table 9. Inflection of strong and weak verb

			Old Arabic	Mardin
imperfect	3rd pl. masc.	strong verb	*yaktubūna*	*yəktəbūn*
		final weak verb	*yansawna*	*yənsawn*
	2nd sg. fem.	strong verb	*taktubīna*	*təktəbīn*
		final weak verb	*tansayna*	*tənsayn*
perfect	3rd pl. masc.	strong verb	*katabū*	*katabu*
		final weak verb	*banaw*	*banaw*

2.3.1.2 Numeral phrase

As described in 2.2.5, the numerals 3–10 have a longer form used independently (i.e. in counting) and a shorter form used before a noun, e.g. Kinderib *xamse* 'five' vs. *xams nəswān* 'five women'. The numerals 11–19 have only a single form, the noun follows in the singular; likewise after tens, hundreds, and thousands, e.g. Kinderib *xamṣṭaʕš bayḏ̣a, mīt bayḏ̣a* 'fifteen, a hundred eggs'. There are no different forms for use with masculine or feminine nouns.

2.3.2 Verbal phrase

2.3.2.1 Object marking

A noun functioning as a definite object usually follows the verb; unlike in Iraqi Arabic it is not normally marked by an anticipatory object suffix on the preceding verb, e.g. Kinderib *šaʕaltu ṣṣōba* 'I lit the oven', *jaṛṛaytu əlxanjaṛ* 'I drew the dagger'. In the Kozluk-Sason-Muş dialects, however, the object noun usually precedes the verb which takes a referential object suffix, e.g. Daragözü *čəftᵊwātna nšilən* 'we take our rifles [lit. our rifles we take them]'.

2.3.2.2 Expression of tense and aspect

2.3.2.2.1 Present tense. The dialect of Mardin does not distinguish between general present and present continuous, e.g. *təṛāni ana aqčəm maʕək ʕaṛabi* 'you see, I am talking Arabic to you'. The majority of the Anatolian dialects, however, mark the present continuous with a verb modifier *kū-* prefixed to the imperfect, e.g. Kinderib *kū-təqšaʕūn ṣīyād-ana* 'you see [lit. are seeing] that I am a hunter'.

2.3.2.2.2 Future and intent. Future and intent are expressed by *tə-*, in the Siirt group *də-*, prefixed to the imperfect, e.g. Kinderib *tə-nə́ji* 'we shall come', Siirt *də-nəzzawwej* 'we shall get married'. With the inflectional prefix *y(ə)-*, *tə-* and *də-* coalesce to *tī-*, *dī-*, e.g. Kinderib *tīrōḥūn* 'they shall go'; with the inflectional prefix *a-* they coalesce to *ta-*, *da-*, e.g. Mardin *taġanni* 'I shall sing'.

2.3.2.2.3 Habitual past. In Mardin this tense is expressed by the particle *kān*, in the remaining dialects by a prefix *ka-*, with the imperfect, e.g. Mardin *kān yjībūn*, Kinderib *kayjībūn* 'they used to bring'.

2.3.2.2.4 Perfect. The perfective aspect is expressed by the morphological perfect with a prefix *kəl-* in Mardin and Siirt, *kū* ~ *kūt-* in the majority of the other dialects, e.g., Mardin *kəlmāt* 'he has died', Kinderib *kūtmāt*.

2.3.2.2.5 Pluperfect. In Mardin *kān kəl-* precedes the morphological perfect to express the remote past, e.g. *abūhu kān kəlmāt* 'his father had died'; in Kinderib the prefix is *kāt-*, e.g. *əlhabwe kāssaddət* (< *kāt-saddət*) *əššəbēbīk* 'the snow had covered [lit. closed] the windows'.

3. LEXICON

3.1 *Borrowings*

There are two main sources for lexical borrowing into Anatolian Arabic, viz. Turkish (both Ottoman and Modern Turkish) and Kurdish (Kurmancî). To a lesser degree Aramaic words survive in Anatolian Arabic (Jastrow 2001). While in the larger cities (Mardin, Diyarbakır, Siirt) more Turkish than Kurdish borrowings are found, the opposite is true for rural dialects.

In → Turkish loanwords which are originally Arabic borrowings into Turkish, Arabic phonemes such as the emphatics, the pharyngeals, and the interdentals are frequently restituted, e.g. *ḥ̣ālbūki* 'however' (< Turkish *halbuki*). Turkish *k* after back vowels is rendered by *q*, e.g. *čəqmāq* 'lighter' (< Turkish *čakmak*), *bāšqa* 'different' (< Turkish *başka*), but *balki* 'perhaps' (< Turkish *belki*). Similarly, Turkish *t* and *d* in nouns with back vowels are rendered by *ṭ* and *ḍ* respectively, e.g. *ḍolma* 'stuffed zucchini' (< Turkish *dolma*), *māzōṭ* 'heavy oil' (< Turkish *mazot*).

Kurdish feminine nouns receive the Arabic feminine morpheme *-e*/*-a*, e.g. *tūre* 'shoulder bag' (< Kurdish *tûr* fem.), *pēžne* 'sound' (< Kurdish *pêjin* fem. 'echo').

3.2 *Lexical variation*

Lexical variation in Anatolian Arabic is quite considerable. Thus the word for 'to talk, to speak' has the following renderings: (a) Mardin group: Mardin *qačam*, Kosa *nēdam*, Mḥallami *ḥaka*, Qarṭmīn *twannas*, Āzəx *štaġal*; (b) Siirt *gara*; and (c) Diyarbakır *ḥaddet*.

BIBLIOGRAPHICAL REFERENCES

Blanc, Haim. 1964. *Communal dialects in Baghdad.* Cambridge, Mass.: Harvard University.

Jastrow, Otto. 1973. *Daragözü: Eine arabische Mundart der Kozluk-Sason-Gruppe (Südostanatolien).* Nuremberg: H. Carl.

——. 1978. *Die mesopotamisch-arabischen* qəltu-*Dialekte. I. Phonologie und Morphologie.* Wiesbaden: O. Harrassowitz.

——. 1981. *Die mesopotamisch-arabischen* qəltu-*Dialekte*. II. *Volkskundliche Texte in elf Dialekten*. Wiesbaden: O. Harrassowitz.

——. 1997. "Wie arabisch ist Uzbekistan-Arabisch?". *Built on solid rock: Studies in honour of Professor Ebbe Egede Knudsen on the occasion of his 65th birthday April 11th 1997*, ed. Elie Wardini, 141–153. Oslo: Novus.

——. 2001. "Aramäische Lehnwörter in den arabischen Dialekten der Südost-Türkei". *Akten des 27. Deutschen Orientalistentages*, ed. Stefan Wild and Hartmut Schild, 615–621. Würzburg: Ergon.

——. 2003. *Arabische Texte aus Kinderib*. Wiesbaden: O. Harrassowitz.

——. Forthcoming. *Glossar zu Kinderib (Anatolisches Arabisch)*. Wiesbaden: O. Harrassowitz.

Sasse, Hans-Jürgen. 1971. *Linguistische Analyse des arabischen Dialekts der Mḥallamiye in der Provinz Mardin (Südosttürkei)*. Ph.D. diss., University of Munich.

Socin, Albert. 1882–1883. "Der arabische Dialekt von Mosul und Märdin". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 36.1–53, 238–277; 37.293–318.

Talay, Shabo. 2001. "Der arabische Dialekt von Hasköy (Dēr-Khāṣ), Ostanatolien. I. Grammatikalische Skizze". *Zeitschrift für Arabische Linguistik* 40.71–89.

——. 2002. "Der arabische Dialekt von Hasköy (Dēr-Khāṣ), Ostanatolien. II. Texte und Glossar". *Zeitschrift für Arabische Linguistik* 41.46–86.

Wittrich, Michaela. 2001. *Der arabische Dialekt von Āzəx*. Wiesbaden: O. Harrassowitz.

Otto Jastrow
(University of Erlangen-Nuremberg)

Andalus

1. Introduction

After the initial conquest of al-Andalus by the Muslim armies, a process of Arabization both linguistic and cultural started (→ Andalusi Arabic). Arabic culture remained a crucial factor on the Iberian Peninsula in the period between the invasion of the Muslims in 711 and the expulsion of the Moriscos at the beginning of the 17th century. The fundamental question is to what degree Arabic was used, or not used, among the inhabitants of the Iberian Peninsula (Christians, Jews, and Muslims), in different periods and different regions. Sources are relatively sparse and are interpreted differently by different scholars.

Some theories describe al-Andalus as a completely monolingual Arabic-speaking society: the Romance vernacular is then assumed to have disappeared completely in the regions under Arab control. Other theories maintain that the use of Arabic was restricted to a relatively small part of the population, whereas the Romance dialects served as the colloquial language for the majority of the inhabitants of al-Andalus. According to yet other theories, Romance and Arabic coexisted, each in its own domain. Apart from diachronic and geographic differences, social stratification is another complicating factor.

2. Evidence of the replacement of Romance by Arabic

Many studies, recent and not so recent, quote the well-known passage of Alvarus of Cordoba from his *Indiculus luminosus* (854) as evidence of the almost total disappearance of the Romance dialects among the Christian citizens of al-Andalus, the Mozarabs. The text complains that Christians "have forgotten their own language, and there is hardly one among a thousand to be found who can write to a friend a decent greeting letter in Latin. But there is a countless multitude who express themselves most eloquently in Arabic and make poetry in this language with more beauty and more art than the Arabs themselves" (Simonet 1888:I, x; von Schack 1877:278; Roth 1994:54–55; Wright 1982:157). Abbot Samson, who translated letters from Arabic into *latinum eloquium* in the year 863 (Wright 1982:159) was proud of his erudition and complained about the lack of *latinitas* of others, which supports the theories of those scholars who maintain that Latin was not used on a large scale.

After the Reconquista we still find various remarks about the linguistic situation in al-Andalus. The most conspicuous views are those of the Jesuit Juan de Mariana (1535–1624) who wrote – following the text of Alvarus – that the use of Latin had almost disappeared in al-Andalus and that Christians had a thorough command of the Arabic language. The fact that Juan [Hispalense] translated the Bible into Arabic in order to facilitate access to Christian sources for Muslim and Mozarabs was adduced by Juan de Mariana in support of his theory (Mariana [1601]1950:195). The predominance of Arabic is also observed by other Jesuit scholars, such as Bernardo de Aldrete ([1606]1972: 141) and Andrés Marcos Burriel (1719–1762). The latter quotes the works of the archbishop of Toledo San Eulogio who said that Christians had a perfect command of Arabic, sometimes even better than native Arabs and stressed that they had almost lost the command of what he calls

"Latin language", although he made a clear distinction between the "colloquial", spoken form of Latin on the one hand which had almost disappeared, and the "erudite language necessary for the faith" on the other, which survived as the language of the Church (Burriel 1755:207–208; Thompson 1971:22). In the 10th century nothing remained of the Romance dialects. Romance had become obsolete among all the inhabitants of the peninsula under Muslim rule and the process of Arabization was already complete in that period. Other authors situated the gradual disappearance in later periods, particularly the age of the Almoravids (1086–1147) and the Almohads (1171–1223) who tried to forbid the use of the Romance language. However, Hanssen (1913:8) demonstrated that the Mozarabs of Toledo had completely forgotten the Romance language when the armies of King Alphonse VI reconquered the city in the year 1085, which means that the process of Arabization was already completed before the Almoravids conquered the Taifa kingdoms. The Christians who lived there even had Arabic names. Bernardo de Aldrete and Martin de Viciana also attempted to demonstrate the importance of Arabic, stressing the fact that it was still spoken in their age in Valencia and Granada.

The Arabization process was realized by two forces. In any society, social or ethnic groups have their own attitudes toward each other. In a society dominated by Arabic-speaking Muslims, Christians began to form a non-prestige social group, which may have led to a negative attitude toward their own language. If they wished to play a more prominent part in Muslim society, they had to speak the Arabic language. The second factor was that Muslims attempted to prohibit the use of Latin and during the rule of Caliph Hišām I (788–796) Christians were forced to attend Arabic schools (Thompson 1971:69). It is impossible to reconstruct the linguistic situation of this early period in detail, but one may assume that the number of Arabic-speaking monolinguals must have been very limited during the first generations. One important argument supporting such a theory is the numismatic evidence. Bilingual Arabic–Latin coins have been discovered from the first generation living under Muslim rule, which implies that Muslims made certain administrative concessions with regard to the language (Thompson 1971:68; Amador de los Ríos 1862:II, 581).

Castro's theory relates to this discussion. He assumes (1956:6) that al-Andalus was a direct spiritual and linguistic continuation of the Islamic East. Although he admits that Romance was spoken on a modest scale, he emphasizes that there is no evidence of a widespread use of this language. Recently, Wasserstein (1991:2) has come to the same conclusion:

> The variety is too great. What is said of the northern part of the country is not necessarily correct of other areas; cities may well offer patterns different from those of the countryside; the earlier periods differ very much from later ones; class and education introduce other variables; and sex probably does so too. Categorization, again, is very difficult: in terms of language itself, there is the difference between spoken and written forms of any specific language; for their users we have to organize a mixed bag of ethnic and religious boundary markers between and within groups; and in using terms like bilingualism, multilingualism, diglossia, and so on there is the risk of appearing to give more exactness to the situation than the facts themselves often warrant.

Wasserstein concludes that the Andalusis were already monolingual Arabic speaking in the 11th century when the Romance language had disappeared almost totally. This view is not compatible with the fact that captive Romance-speaking Mozarabs were integrated in North-African Almoravid armies. In this context, it may be relevant to cite a statement by al-ʾIdrīsī who mentions the fact that in the year 1154, the native Romance language of North Africa was still spoken in the cities (Lewicki 1951–1952: 418, 430).

The 17th-century author al-Maqqarī mentions the fact that the Mozarabs spoke Arabic very well. He states that the spoken dialect of Arabic in al-Andalus was corrupt and that the Andalusis were very competent in writing classical Arabic texts, and could compete with the Eastern sources (Gayangos 1840:II, 142–143).

3. Evidence of the survival of Romance

This alternative view of the linguistic situation in al-Andalus has achieved less success. Eguílaz y Yanguas (1886:viii–ix) states that the Andalusis never succeeded in influencing the Christian civilization. He even states that the Andalusi Arabic culture was not the product of Arabs, but of renegade Christians, *muwalladūn*, and Jews. His main argument for his theory is the fact that Arab authors mention the use of Romance in several regions, such as Aragón, Zaragoza, and Valencia. The Hispano-Romans had been a civilized and literate people before the Muslim conquest and this situation did not change during the age of Muslim rule in al-Andalus. Cejador y Frauca (1932:I, 107–108) assumes that the Mozarabic dialect was not very different from the northern dialects. He even states that Andalusi authors such as Ibn Ḥazm wrote "in the tradition of Isidorus of Seville".

According to Sánchez-Albornoz the process of Arabization and Islamization proceeded very slowly (1946:I, 356). The Christians showed themselves averse to accepting Islam and Arabic language and manners. Romance was even spoken by the caliph himself (Thompson 1971:68): a minority would never have been able to impose their language on seven or eight million Christians. He also supposed that the Berbers only spoke their own language and that they were unable to speak Arabic correctly. Sánchez-Albornoz took his information from an episode from the *Kitāb al-quḍāt bi-Qurṭuba* written by the historian and jurist al-Xušanī (10th century). The fragment in question describes a prestigious Muslim individual who had testified in Romance in a court in Cordoba. This person lived in the reign of ʿAbd ar-Raḥmān II (822–852), which led Sánchez-Albornoz to the conclusion that Romance was still current in these days. As Thompson (1971:78) observes, "the period with noticeable shifts among the youth Christians to an interest in and mastery of Arabic is the mid-ninth century". Most scholars do not make a clear distinction between 'Latin' and 'Romance', on the one hand (von Schack 1988:278), and between 'Classical Arabic' and 'colloquial', on the other. Sánchez-Albornoz probably alluded to Classical Arabic. It seems premature to draw the conclusion that the use of Arabic was not very common. Of course, a preacher in a mosque had to be a learned Muslim, familiar with classical Arabic. Such learned people may have been scarce, but this does not mean that colloquial Arabic was not spoken on a large scale. Probably, well-educated Andalusis with a perfect knowledge of Classical Arabic constituted a minority at this early stage.

4. Coexistence of Romance and Arabic

The theories sketched above represent opposite views. Since the 19th century, more moderate theories have been formulated. Von Schack (1988:278) describes the linguistic situation in his study of Arabic poetry in detail. According to him, most Mozarabs were Arabized quickly after the Reconquista, but continued to speak 'Latin' or 'Romance', which survived as the language of the people. There were even many Andalusis who spoke the Romance language. Menéndez Pidal (1904:22) states that the manifold mutual political, commercial, and social contacts, such as marriages between Andalusis and Christians, had resulted in a large number of Arabic elements in the Romance languages of the Iberian Peninsula. He mentions the so-called *enaciados*, individuals who operated between the two parties as spies, intermediaries, messengers, or couriers, and who were completely bilingual. He also mentions the so-called *moros latinados* or *ladinos* in the Arabic-speaking community who spoke Romance (cf. the French term *latinier*, a term used for a Muslim who has command of the language of the Christians). The word is also used for 'any interpreter', whereas Roth (1994:54, 254, n.55) observes that the term *ladino* also means 'astute'. On the other hand, there were also the *cristianos algarabiados* who spoke Arabic fluently, and *dragomanes*, also called *trujamanes* or in Catalan *torsimanys* (Burns 1984:186). The existence of these individuals does not constitute evidence that widespread bilingualism really existed. The fact that these *dragomanes* were needed rather demonstrates the lack of bilinguals, apparently because many people were monolingual. We know nothing about the number of these *moros latinados* and *cristianos algarabiados*. For Menéndez Pidal (1973), the fact that Rodrigo Díaz de Vivar, the Cid, when he entered Valencia, found Christians there who spoke only Arabic, is clear evidence of the continuation of Christian culture. Recently, Epalza and Llobregat (1982:27) refuted the thesis of Menéndez Pidal; according to them, most of these Christians immigrated to the Taifa kingdoms in the 11th century. They even state that already in the 8th century on the eastern coast (*Šarq al-ʾAndalus*) and in the Baleares, no indigenous Christians were left (1982:8). Those Christians who were persecuted by the Almoravids and Almohads were mainly 'foreign' Christians from the north, rather than 'indigenous'. The most important argument for such a thesis is the disappearance of organized Christians in al-Andalus and the non-existence of episcopates. The process of Arabization and

Islamization in the eastern part of al-Andalus was completed in the 13th century (Epalza and Llobregat 1981:31). Another argument used by Menéndez Pidal is the fact that Jacobo de Vitry states that Latin was still used by the Mozarabs in the 13th century. Lapesa (1983:129–130) only points to the existence of the two languages side by side. He does not mention which specific groups used what language and when.

A fragment written by the famous linguist and lexicographer Ibn Sīda from Murcia (1007–1066), author of the two dictionaries *Kitāb al-muḥkam* and *Kitāb al-muxaṣṣaṣ* (I, 14) has been interpreted as evidence of the existence of a Romance language, spoken in Murcia and coexisting with Arabic. In this fragment, Ibn Sīda complains about the difficult circumstances in which he works as a purist, living among *ʿajam* people (non-Arabic). Ribera translates this word as "personas que hablan romance", but recently Bramon (1977:20) has pointed out that this translation must be corrected. According to her thesis, *ʿajam* must be interpreted in ethnic terms rather than linguistic. *ʿAjam* means 'non-Arabic people' and not necessarily 'non-Arabic-speaking people' (cf. Epalza 1981:168; Barçeló 1979). For Roth (1994:54), the fragment shows the existence of a Romance-speaking population, but the evidence is restricted since it reflects the situation in Murcia, "from which we cannot generalize for all of Muslim Spain".

5. EVIDENCE OF ARABIC/ROMANCE BILINGUALISM

Simonet (1988:xxvi; cf. Anssens-Lestienne 1983:12) maintains that the Mozarabs never forgot the religious and literary language of their ancestors. The existence of codices written in Latin until the end of the Reconquista proves this, as does the occurrence of many Romance loanwords in the Hispano-Arabic language, especially as recorded in the *Vocabulista in Arábigo* by Pedro de Alcalá (1505). This shows that Romance was spoken on a large scale throughout the country by a major part of its inhabitants. The idea of a totally successful integration and a complete bilingualism in al-Andalus was sustained by Steiger (1967:96). In his view, al-Andalus was a bilingual society at least until the 12th century when the Mozarabs became gradually monolingual Arabic-speaking individuals and were forced to move to the northern kingdoms.

Lévi-Provençal (1953:76) posited a linguistic opposition between the cities, where the Arabic language dominated as the vehicle of literate society, and the countryside, where Romance was prevalent. Madariaga was the first to consider the difference in terms of the social stratification of al-Andalus. Entwistle (1936:106) – without specifying geographically or chronologically – maintains that "Romance was the language of the marketplace, of all women and of unofficial intercourse". Arabic was the language of "administration, literature and highclass families claiming Arabian descent".

A fragment frequently quoted by some scholars as evidence of the fact that bilingualism occurred on a large scale is from al-Muqaddasī's *Kitāb ʾaḥsan at-taqāsīm* (985). According to this fragment, the variety of spoken Arabic in al-Andalus was difficult to understand for an Arab from the East, and a form of Romance, similar or related to the Romance or Latin language (*rūmī*) was current in al-Andalus. He even states that it was unusual for certain Arabic families of high society not to know Romance. Other Arabic sources, too, mention the fact that Romance was spoken in al-Andalus. Ibn Ḥazm (994–1063) states that not knowing Latin was unusual for the Muslims who lived in Aguilar and Morón in his time (Burns 1984:174; Castro 1956:8).

It is well known that many learned Jews were fluent in both Arabic and Romance. Apparently, Moše ibn Ezra (1055–ca.1138) studied Christian commentaries on the Bible in Latin (Díez Macho 1953:15), and in his *Kitāb al-muḥāḍara wa-l-mudākara* (chapter 3, 24a; cf. Brann 1991:196) we read that when he was young, he asked an Islamic scholar to recite the first *sūra* of the *Qurʾān* – the *fātiḥa* – in Romance. This scholar did what he asked, although the result sounded ugly and the language distorted. There must therefore have been learned Muslims who were able to speak Romance, but Roth (1994:53) states that this "certainly does not prove that this was common. On the contrary, his mention of it indicates that it was unusual". In recent years more chronological distinctions have been made.

6. CONCLUSION

Perhaps the best way to analyze the linguistic situation in al-Andalus is by combining elements of each theory. This is what Galmés de Fuentes (1983:14–17) does when he combines various theories and distinguishes three periods in the linguistic development in Muslim Spain. In the first period, the Mozarabs not only preserved their Romance language for domestic purposes, but Romance was also the dominant language of all inhabitants of al-Andalus, even of Muslims. There were even Muslims who were not proficient in the Arabic language and only spoke *aljamiado* or Romance. Galmés also makes a clear distinction between the main cities, such as Toledo and Seville, inhabited almost exclusively by *hispano-godos*, and the countryside, where the Muslims preferred to live. During the second period, the impact of the Mozarabs diminished steadily up till the year 1099, when the first persecution of the Mozarabs took place. After the year 1102, the majority of the Mozarabs emigrated to the north. The third period is dominated by the two invasions from North Africa, the Almoravid and the Almohad. The number of Mozarabs decreased dramatically because of execution or forced emigration, not only to the north, but also to North Africa. Yet, the culture of the Mozarabs did not disappear. The Romance language still had a considerable social and even literary power. Galmés de Fuentes adduces the example of the frequently quoted botanists Ibn Buklāriš, who wrote shortly before the reconquest of Alfonso el Batallador in the year 1118, and Ibn al-Bayṭār, from Málaga who died in 1248. These authors used Hispano-Romance loanwords, and this demonstrates that bilingualism existed. He also adduces as evidence the presence of many Romance items in the Arabic vocabulary of the Granadan Muslims in the *Vocabulista* of Pedro de Alcalá (1505). Galmés even speaks about a "mixed language". In his recent study of the *xarjas* (1994:81–88), Galmés comes to the conclusion that there is enough evidence for the existence of bilingualism in al-Andalus until the 13th century. Yet, many of the sources adduced by Galmés are connected with the presence of Christians or Mozarabs. Few arguments are related exclusively to the linguistic situation in general and bilingualism in particular.

Corriente (1991:66) and Wright (1994:265) agree with Galmés insofar as the first period is concerned. They agree that the use of Romance and Latin was diminishing rapidly. According to Corriente, Romance soon became the "prestigeless language of women, peasants, and slaves. Most people still understood it to some degree but very few cared to speak it any more, although it remained in use in images of domestic scenes". In the 13th century, Andalusis were completely monolingual and bilingualism was only marginally recorded in the 11th and 12th centuries. Corriente (1994:448) refutes the theory of some authors who overrate the persistence of bilingualism and the use of Mozarabic. Wright (1994:265) concludes that the Christian community in al-Andalus was mostly bilingual in speech in the period after the Muslim conquest, using *ladino*, according to his terminology, and Arabic. By the end of the 9th century, they became literate in Arabic alone. After the year 860, there was no Christian left in al-Andalus who was able to write Latin, which explains the lack of documents in Latin from that period. An important source of bilingual utterances is the collection of bilingual Romance-Arabic closing lines, the *xarajāt*, appended to the Hispano-Arabic and Hispano-Hebrew strophic poems called the *muwaššaḥāt*. Some interpret these as authentic Romance songs which survived until the 13th century, in which case they could be proof of the persistence of Romance and the existence of bilingualism. Since many Romance or partly Romance *xarajāt* are love songs sung by women, one is inclined to see these texts as evidence of the fact that Romance was indeed used for the 'lower registers' of language use, i.e. for unofficial, often humorous and witty love songs. More recently, there is a tendency among scholars to mitigate their theories concerning the importance and persistence of 'Mozarabic poetry' since many readings and interpretations of the texts are no longer supported (see also Zwartjes 1997).

Bibliographical references

Primary sources

Alcalá, Pedro de. *Vocabulista arauigo en letra castellana*. Dictionary appended to the *Arte para ligeramente sauer la lengua arauiga* (1505). Ed. P. de Lagarde, *Petri Hispani de lingua Arabica libri duo*. Göttingen: Arnold Hoyer, 1883. Repr. Osnabrück: Otto Zeller.

Aldrete, Bernardo de. 1606. *Del origen y principio de la lengua castellana romance que oi se usa en España*. Rome: Carlo Willetto. Ed. Lidio Nieto Jiménez. Madrid: Consejo Superior de Investigaciones Científicas, 1972.

Alvarus of Cordoba. *Indiculus luminosus*. Ed. J. Gil, *Corpus scriptorum Muzarabicorum*. 2 vols. Madrid: Consejo Superior de Investigaciones Científicas, 1974.

Burriel, Andrés Marcos. 1755. "Paleographia española". *Espectáculo de la naturaleza o conversaciones* *a cerca de la particularidades de la historia natural*, XIII. Ed. Abad M. Pluché, trans. Estevan de Terreros y Pando (1707–1782), 201–359. Madrid: Gabriel Ramírez; microfilm: Oviedo: Pentalfa Microediciones, 1989.

Ibn Sīda. *Kitāb al-muxaṣṣaṣ*. 3rd ed. 17 vols. Cairo: Dār al-Fikr, 1970.

Maqqarī, al-. See Gayangos 1843.

Mariana, Juan de. 1601. *Historia de España*. Ed. Madrid: Real Academia Española, Biblioteca de Autores Españoles, 1950.

Muqaddasī, al- (d. after 375/985). *ʾAḥsan at-taqāsīm fī maʿrifat al-ʾaqālīm*, III. Ed. M.J. de Goeje. 2nd ed. Leiden: E.J. Brill, 1906.

Viciana, Martin de. *Libro de alabanças de las lenguas Hebrea, Griega, Latina, Castellana y Valenciana*. Valencia: Joan Navarro, 1574.

Secondary sources

Amador de los Ríos, José. 1862. *Historia crítica de la literatura española*. 8 vols. Madrid: Imprenta de José Rodríguez.

Anssens-Lestienne, Yannick. 1983. "L'arabe andalou. Sources et bibliographie". *Matériaux Arabes et Sudarabiques* 11–59.

Barçeló Torres, María del Carmen. 1979. "La llengua àrab al País Valencià". *Arguments* 4.123–149.

Bramon, Dolors. 1977. "Una llengua, dues llengües, tres llengües". *Raons d'identitat del país Valencià*, 17–47. Valencia: Pere Sisè.

Brann, Ross. 1991. *The compunctious poet: Culture ambiguity and Hebrew poetry in Muslim Spain*. Baltimore and London: Johns Hopkins University Press.

Burns, I. Robert. 1984. *Muslims, Christians and Jews in the crusader kingdom of Valencia: Societies in symbiosis*. Cambridge: Cambridge University Press.

Cejador y Frauca, Julio. 1932. *Historia de la lengua y literatura castellana*. 3rd ed. Madrid: Casa Editorial Hernando.

Corriente, Federico. 1992. *Árabe andalusí y lenguas romances*. Madrid: Mapfre.

——. 1994. "Linguistic interference between Arabic and the Romance languages of the Iberian Peninsula". *The legacy of Muslim Spain*, ed. S. Kh. Jayyusi, 443–451. Leiden: E.J. Brill.

Díez Macho, Alejandro. 1953. *Mošé Ibn ʿEzra como poeta y preceptista*. Madrid and Barcelona: Consejo Superior de Investigaciones Científicas.

Eguílaz y Yanguas, Leopoldo. 1886. *Glosario etimológico de palabras españolas (castellanas, catalanas, gallegas, mallorquinas, portuguesas, valencianas y vascas) de origen oriental (árabe, hebreo, malayo, persa y turco)*. Granada: Imprenta de Lealtad.

Entwistle, William, J. 1936. *The Spanish language*. London: Faber and Faber.

Epalza, Míkel de. 1981. "Notas sobre el lingüista Ibn Sidah y la historia de Denia y su región en el siglo XI". *Revista del Instituto de Estudios Alicantinos* 33.164–172.

—— and Enrique Llobregat. 1982. "¿Hubo mozárabes en tierras valencianas? Proceso de islamización del Levante de la Península (Sharq al-ʾAndalus)." *Revista de Investigación y Ensayo* 36.7–31.

Galmés de Fuentes, Álvaro. 1983. *Dialectología mozárabe*. Madrid: Gredos.

Gayangos, Pascual de (trans). 1843. *The history of the Mohammedan dynasties in Spain*. London. (Repr. New York: Johnson Reprint Company, 1964.)

Hanssen, Federico [Friedrich]. 1913. *Gramática histórica de la lengua castellana*. Halle: Max Niemeyer. (Repr. Buenos Aires: El Ateneo, 1945.)

Lapesa, Rafael. 1983. *Historia de la lengua española*. 9th ed. Madrid: Gredos.

Lévi-Provençal, Emile. 1953. *Histoire de l'Espagne musulmane*. Paris: G.P. Maisonneuve.

Lewicki, Tadeusz. 1951–1952. "Une langue romane oubliée de l'Afrique du Nord". *Rocznik Orientalistyczny* 17.415–480.

Menéndez Pidal, Ramón. 1973. *Manual de gramática histórica española*. 14th ed. Madrid: Espasa-Calpe. (First published 1904.)

Roth, Norman. 1994. *Jews, Visigoths and Muslims in medieval Spain: Cooperation and conflict*. Leiden: E.J. Brill.

Sánchez-Albornoz, Claudio. 1946. *La España musulmana según los autores islamitas y cristianos medievales*. Buenos Aires: El Ateneo. (3rd ed., 2 vols., Madrid: Espasa-Calpe, 1973.)

Schack, Adolf Friedrich von. 1877. *Poesie und Kunst der Araber in Spanien und Sicilien*. 2nd ed., 2 vols. Stuttgart: Verlag der Cotta'schen Buchhandlung.

Simonet, Francisco Javier. 1888. *Glosario de voces ibéricas y latinas usadas entre los mozárabes precedido de un estudio sobre el dialecto hispano-mozárabe*. 2 vols. Madrid: Fortanet.

Steiger, Arnald. 1967. "Arabismos". *Enciclopedia lingüística hispánica*, ed. M. Alvar, II, 93–126. Madrid: Consejo Superior de Investigaciones Científicas.

Thompson, Billy Bussell. 1971. *Bilingualism in Moorish Spain*. Ph.D., diss. University of Virginia.

Wasserstein, David. 1991. "The language situation in al-Andalus". *Studies on the* muwaššaḥ *and the* kharja. *Proceedings of the Exeter international colloquium*, ed. Alan Jones and Richard Hitchcock, 1–15. Reading, U.K.: Ithaca Press.

Wright, Roger. 1982. *Late Latin and early Romance in Spain and Carolingian France*. Liverpool: F. Cairns.

——. 1994. "La muerte del ladino escrito en Al-Andalus". *Euphrosyne* 22.255–268.

Zwartjes, Otto. 1997. *Love songs from al-ʾAndalus: History, structure and meaning of the* kharja. Leiden: E.J. Brill.

Otto Zwartjes (University of Amsterdam)

Andalusi Arabic

1. Andalusi Arabic

Andalusi Arabic is a dialect bundle, constituted by scarcely differentiated members and generated by the occupation of the Iberian Peninsula at the beginning of the 8th century by armies of Muslim Arabs and (partially) Arabicized Berbers. It appears to have spread rapidly and been in general oral use in most parts of the geopolitical entity resulting from those events, called al-Andalus by its native population, between the 9th and 15th centuries. It reached its highest peak of users, which can be roughly estimated at 5–7 million, during the 11th and 12th centuries. It then dwindled as a consequence of the gradual but relentless takeover by the Christian northern states of all lands held by the Muslims, although it remained in use in certain areas already under Christian political control until the final expulsion of the Muslims at the beginning of the 17th century. It was in all likelihood also spoken by Andalusi immigrants in North Africa, at least for a few generations, as hinted by its pervasive influence on many North African Arabic dialects. It may also be easily surmised that Andalusi Arabic played an important role in the Arabicization of the countries in that region and the gradual disappearance of Berber dialects from urban milieus, although not so in rural and, above all, mountainous areas, where they remain to this day alive and healthy.

Andalusi Arabic soon became the main linguistic link between all the inhabitants of al-Andalus, if due allowance is made for the initial stages of that historical entity and for remote areas where Romance monolingualism might have lasted longer.

As in every other Arabic-speaking land, the Andalusi people were diglottic, i.e. spoke their local dialect in all low-register situations, but only Classical Arabic was resorted to when a high register was required, as well as for written purposes (→ Andalus).

Andalusi Arabic clearly belongs to Early Western Neo-Arabic, which does not allow for any separation between Bedouin, urban, or rural types or dialects, nor does it show any detectable difference between communal dialects, such as Muslim, Christian, and Jewish, beyond the well-known features of → Middle Arabic written documents.

The oldest evidence of Andalusi Arabic utterances can be dated from the 10th and mostly 11th century, in isolated quotes, both in prose and stanzaic Classical poems *(muwaššaḥāt)*, and then, from the 11th century on, in stanzaic dialectal poems *(ʾazjāl)* and dialectal proverb collections, while its last documents are a few business records and one letter written at the beginning of the 17th century in Valencia.

The Andalusi Arabic dialect bundle reflects a local evolution in an uneven melting-pot of Classical Arabic dialects, quite akin by drift or selection to the Neo-Arabic type in matters such as loss of *ʾiʿrāb*, with other dominant creole traits, likely of Nabaṭī or Yemenite ancestry, some characteristic substratal Romance, and fewer adstratal Berber features, above all in its lexicon.

Since the Middle Ages, Andalusi Arabic texts have been now and then transmitted by Western language sources, in brief quotes, glossaries, loanwords and place-names, even grammatical sketches, but there was no comprehensive description of them before that of Colin (1960), nor any detailed account of Andalusi Arabic grammar until Corriente (1977). Its main sources were critically edited only after that and the first attempt at a full account of its lexicon was that of Corriente (1997).

2. Linguistic description

2.1 *Phonology*

2.1.1 Inventory

The phonemic inventory of the Andalusi Arabic dialect bundle counts 27 consonantal and 3 vocalic phonemes.

2.1.1.1 Consonants

The consonantal phonemes are the same and with identical realizations as those of the received pronunciation *(tajwīd)* of Classical Arabic, but for the merger of /ḍ/ into /ḏ̣/. There are also 3 marginal consonants, /p/, /č/ and /g/, found in Romance and Berber borrowings. The status of /ʾ/ is quasi-marginal, as it was realized only occasionally in intervocalic position.

Interdentals, pharyngeals, sibilants, and liquidae are generally preserved except for isolated instances of substitution of dentals, alternation or even loss of pharyngeals in prejunctural position, hesitation of voice in sibilants, and alternation of liquidae respectively.

The phonemes /q/, /ʤ/, /j/ and /k/ had standard realizations, with only some hints of idiolectal voicing of /q/ into /g/, early instances of Yemenite occlusive *gīm*, and loss of affrication of /j/ (= /ž/).

A genuine lateral /ḍ/ is reflected by early loanwords in Romance (e.g. Castilian *alcalde* ‘mayor’ < *alqáḍi*, *aldea* ‘village’ < *aḍḍáyʿa*), but later on and as a consequence of the standardization of Andalusi Arabic, the merger of /ḍ/ with /ḏ̣/ became complete (→ *ḍād*).

As in other Arabic dialects, it appears that velarized /r/ may have acquired full-fledged phonemic status in cases such as *baṛṛád* 'it hailed' vs. *barrád* 'it cooled'.

Andalusi Arabic emphasis may have belonged to either the velarization or pharyngealization types, as proven by its effects upon the vocalic environment, e.g. in the Arabic loanwords of Romance, although the glottalization type cannot be altogether excluded in its earliest stages, before standardization, when Yemenite traits were pervasive. While there are no hints of affrication of dentals, /b/ is sometimes spirantized and even vocalized with graphemic results ranging from /f/ to /w/, e.g. *yafqá* 'it remains' (< *yabqà*) and *qáwqab* 'clog' (< *qabqāb*).

Distribution of phonemes within syllable boundaries: Andalusi Arabic does not allow postjunctural clusters (CCv-), requires a vowel as center of any syllable, and rejects clusters of more than two consonants, except in some prejunctural biconsonantal sequences containing liquidae and sibilants (e.g. *kálb* 'dog', *bánt* 'daughter', *búrj* 'tower', *áṣṭ* 'arse').

2.1.1.2 Vowels

The vocalic inventory of Andalusi Arabic consists of three phonemes, /a/, /i/, and /u/, with environment-conditioned allophones (more open in contact with pharyngeal and velarized phonemes, as well as in closed syllables), while quantity distinctions, so characteristic of Classical Arabic and most Neo-Arabic dialects, appear to have been eliminated. There is a possibility, though not positively clear, that /a/ may have split into /a/ and /e/ as a result of palatalization of Classical Arabic /ā/, a phenomenon called → *ʾimāla* by native grammarians.

Andalusi Arabic reflexes of Classical Arabic vowels are sometimes irregular for reasons other than combinatory phonetics. In some instances, where /i/ has developed into /a/ in stressed syllables, one suspects the effect of Philippi's law (e.g. *záyy* 'clothing' < *ziyy*, *sákka* 'ploughshare' < *sikkah*), which probably also triggered hypercorrect forms, such as *zínd* 'arm' < *zand*, *fírq* 'difference' < *farq*, etc.

Final /ā/ and /āʾ/ not only are not distinguished in Andalusi Arabic, but have even merged with the feminine morpheme *-a(t)*, regardless of their origin, as proven by instances like *dawátu* 'its medicine' < *dawāʾuhū*, *iḥdátha* 'one of them' < *ʾiḥdāhā*, *ʿaṣátak* 'your stick' < *ʿaṣāka*, *arriḥatáyn* 'both mills' < *arraḥawāni*, etc.

Second degree, i.e., intense *ʾimāla* eventually caused the merger of /a/ and /i/ into the archiphoneme /Í/, almost regularly in late stages (e.g. Granadan *bíb* 'door' < *bāb*, *tíj* 'crown' < *tāj*, but also in earlier sporadic cases from every area, like the widespread *wíld* 'father' < *wālid* and *jímiʿ* 'mosque' < *jāmiʿ*.

Accommodation to the consonantal environment triggers the development of vocalic allophones, labialized, palatalized, or velarized, merely detectable in borrowings by Romance languages, which identified them with Castilian *aceite* < *azzáyt* 'oil', *faneca* 'land measure' < *faníqa*), less often patent in written records (e.g. *yašrúb* 'he drinks' < *yašrabu*, *dijája* 'hen' < *dajāja*, *ḍúrr* 'damage' < *ḍarr*).

Vowel distribution in Andalusi Arabic syllables and words is not entirely free. Although inheriting certain preferences and constraints from Classical Arabic, such as partial vocalic harmony and avoidance of sequences where /i/ or /y / would abut on /u/ or /w/, it shares with Neo-Arabic the preference for CaCCūC ~ CaCCīC over its harmonized Classical counterpart (e.g. *ʿaṣfúr* 'bird', *xanzír* 'pig'), and shares with Western Arabic the occasional allowance of a diphthong /iw/ (e.g. *istiwbár* 'hair-raising', *iwrá* 'showing').

Andalusi Arabic vowels are generally stable, whether historically long or short, but can disappear in post-tonic syllables, in cases like *wíld* 'father' < *wālid*, *ṣáḥb* 'owner' < *ṣāḥib*, *xábya* 'jar' < *xābiʾa*, within a trend prevailing in Neo-Arabic dialects.

2.1.1.3 Diphthongs

Andalusi Arabic is one of the most conservative branches of the whole Neo-Arabic type regarding preservation of Classical Arabic diphthongs /aw/ and /ay/.

However, there are several cases of monopthongization (e.g. *lís* 'not' < *laysa*, *ṣunúbra* 'pine-tree' < *ṣanawbara*, *qíḥ* 'pus' < *qayḥ*), as well as others that could be considered hypercorrect diphthongs generated by a reaction to that infracorrect trend (e.g. *ṯáwm* 'garlic' < *ṯūm*, *ṣáwf* 'wool' < *ṣōf*). This can be construed as evidence that monophthogizing dialects were extant, but in a minority, among those brought to the Iberian Peninsula by the Arabs.

As for the realization of diphthongs, whether preserved or reduced, it must be supposed that their vocalic onset should suffer some degree of assimilation both to the preceding consonant and to the next glide, as hinted by Romance transcriptions (e.g. Portuguese *açoute* < *assáwṭ* ‘scourge’, Castilian *aceite* < *azzáyt* ‘oil’, *aljéun* < **aljawn* ‘the gulf’). However, as these often reflect the strong monophthongizing trends of Romance, they are not absolutely reliable as a guide to the actual pronunciation by native Andalusi people, which does not allow us to posit */e/ and */o/ phonemes with any certainty in those cases, instead of a mere reduction to /i/ and /u/, as in other Neo-Arabic areas.

2.1.1.4 Syllable

The inventory of possible syllable types in Andalusi Arabic includes Cv and CvC without any positional constraint, and CvCC, only allowed in prejunctural position. An additional syllable type v(C) should be added to that list, if /ʾ/ is excluded from the phonemic inventory, which could be a legitimate analysis of the situation.

Consonant clusters do not call for additional comments beyond the remarks above and in 2.1.1.

2.1.1.5 Stress

Stress in Andalusi Arabic, quite exceptionally for any kind of Arabic, is phonemic, as a consequence of the loss of phonemic quantity of vowels and syllables. This is the only possible inference that can be derived from the scribes’ consistent habit of marking stressed syllables with matres lectionis (i.e. long vowel graphemes) and dispensing with them in the case of unstressed syllables, in both cases often in open disagreement with Classical Arabic customary spelling rules (e.g. *māʿi* < = *máʿi* ‘with me’ < *maʿī*, *ḏanābu* < = *ḏanábu* ‘its tail’ < *ḏanabuh*, *ẓahārat* < = *ẓahárat* ‘she appeared’ < *ẓaharat*).

In most cases, but not always, stress patterns are predictable in terms of syllable structure, such as CvCv́C(v), Cv́CC(v), Cv́CCvCa, mvCCv́Ca, and Cv́CvC when that first vowel was historically long, regardless of the condition of the second one, CvCCv́C(a) when the second vowel was historically long and the first one short, as well as CvCCv́C(a), with some degree of hesitation, perhaps idiolectic, with Cv́CCvC(a), e.g. *katáb* ‘he wrote’, *katábu* ‘they wrote’, *qaṣába* ‘castle’, *qalámi* ‘my pen’, *kálbi* ‘my dog’, *mánxar* ‘nose’, *qánṭara* ‘bridge’, *maḍrába* ‘tuna fishery’, *kátib* ‘writer’ < *kātib*, *ṣáḥ(i)ba* ‘female friend’ < *ṣāḥiba*, *dínar* ‘dinar’ < *dīnār*, *xanázir* ‘pigs’ < *xanāzīr*, *xanzīr* ‘pig’ < *xinzīr*, *silbáḥa* ‘eel’, *muqaddám* ‘commander’ (cf. Castilian *almocadén*), but *muwáddan* ‘muezzin’ (cf. Castilian *almuédano*), it being remarkable that there is some hesitation only when the last two syllables were closed. Such a system appears to inherit another where stress was weak and entirely predictable in terms of syllable length and structure, as was probably the case of the Yemenite dialects brought by a majority of the Arab invaders of the Iberian Peninsula. But here it became stronger and occasionally unpredictable upon the loss of quantitative rhythm, in agreement with the Hispanic substratum.

The stress position is fixed, except for its shift to the *-u* pl. suffix of verbs when a pronominal suffix is added, e.g. *yaktúbu* ‘they write’, but *yaktubú+h* ‘they write it’, and to pronominal suffixes ending in vowel when clitic indirect objects are added (e.g. *ʿaṭahá+li* ‘he gave them to me’). The same capacity to attract the stress is regular in the sound plural and dual morphemes, and often observed in the *nisba* suffix.

2.1.2 Phonotactics

Assimilation between consonants or consonants and vowels in contact, less often distant, simple, or reciprocal, can take place, for example, with respect to voice (e.g. *maḥfún* ‘rotten’ < *maʿfūn*), and point or manner of articulation (e.g. *ẓahál-lak* ‘it seemed to you’ < *ẓahar(a) lak*, *nallás* ‘I sit down’ < *najlis*, *ṣaqṣá* ‘he asked’ < *istaqṣā*, *naččakí* ‘I complain’ < *naštakī*, *úččak* ‘your face’ < *wajhak*, *jizzár* ‘butcher’ < *jazzār*, *kiṯír* ‘much’ < *kaṯīr*, *xarínj* ‘heath tree’ < *xalanj*). There are also cases of ultracorrect dissimilative reactions, such as *ísd* ‘arse’ < *ist*, *ṣifráwi* ‘bilious’ < *ṣafrāwī*, *qíwwa* ‘strength’ < *quwwa*.

Dissimilation in a sequence of identical vowels or consonants happens sometimes, e.g. in cases like *niṣráni* ‘Christian’ < *naṣrānī*, *ṣirṣáf* ‘willow’ < *ṣafṣāf*, *kurnása* ‘writing pad’ < *kunnāsa*, *kaymún* ‘cumin’ < *kammūn*.

Metathesis is sometimes found in cases like *yazhú* ‘he mocks’ < *yahzaʾu*, *lutáyra* ‘spider’ < *rutaylāʿ*, *raʿʿáda* ‘catapult’ < *ʿarrāda*.

Suprasegmental spread of emphasis, evidence of its being velar or pharyngeal, is met with in cases like *infiṭáq* ‘ripping’ < *infitāq*, *ṣaqṣá* ‘he asked’ < *istaqṣà*, *ṣáṭl* ‘bucket’ < *saṭl*.

2.1.3 Morphophonology

Some Andalusi Arabic nouns exhibit elision of vowels when compared with Classical Arabic or Neo-Arabic, e.g. *jáml* ‘camel’, *ṭárf* ‘point’, *jábl* ‘mountain’, *báqra* ‘cow’, *wázġa* ‘gecko’ vs. *jamal*, *ṭaraf*, *jabal*, *baqara*, *wazaġa*. It cannot presently be ascertained whether such instances are an inheritance from old dialects already having those shapes brought along by some Arab tribesmen, or just reflect the effects of standardization of stress patterns upon Cv́CvC(a) structures, similar to those prevailing in Eastern Arabic dialects and Maltese.

As for the insertion of vowels, it is characteristic of Andalusi Arabic that it allows only certain prejunctural biconsonantal clusters while others are eased by means of a non-phonemic disjunctive [a], e.g. *nám(a)l* ‘ants’, *báṭ(a)n* ‘belly’, etc. Those disjunctive vowels are often (optionally) phonemicized (e.g. *naʿáš* ‘stretcher’ < *naʿš*, *zanjafúr* ‘cinnabar’ < *zunjufr*); otherwise they are dropped when the cluster disappears, e.g. through the agency of a suffix with a vocalic onset (e.g. *báṭn+ak* ‘your belly’), or are functionally replaced in open junctures by connective /i/ (or *a* in pharyngeal and laryngeal settings) (e.g. *kúlli yáwm* ‘every day’, *sábʿa míyya* ‘seven hundred’).

Following a Pan-Semitic trend, but not always in agreement with Classical Arabic, Andalusi Arabic mostly goes along with Neo-Arabic in the treatment of biconsonantal and monoconsonantal root morphemes, which are lengthened by gemination of its last consonant (e.g. *dámm* ‘blood’, *fúmm* ‘mouth’, *ḥírr* ‘vulva’, *šúffa* ‘lip’ vs. *dam*, *fam*, *ḥir*, *šafa*), or generalization of construct state morphemes (e.g. *axú* ‘brother’, *ḥamú* ‘father-in-law’, which reflect the Classical Arabic construct shapes *axū* and *ḥamū* of *ʾax* and *ḥam*). Conversely, there are cases of proper names in which *ʾab* ‘father’ does not exhibit the usual construct state morpheme, e.g. *ab jáʿfar* and *ab ʿámir*.

Among other clitics, the monoconsonantal prepositions *bi*, *fi*, and *li* exhibit the unusual feature among Neo-Arabic dialects of losing their final vowel when abutting upon the definite article *al-*, e.g. *b+al+muftáḥ* ‘with the key’, *f+al+báyt* ‘in the house’, *la+r+rajúl* ‘to the man’.

In the case of verbal complements, not only pronominal direct objects are clitic, but also the indirect objects introduced with the preposition *li-*, e.g. *naqúl+lak* ‘I tell you’, *naʿmál+lak* ‘I make for you’, even after a pronominal direct object, e.g. *naxrij+á+lu* ‘I will put them out for him’, *yaʿṭi+há+li* ‘he gives her to me’.

As usual in Neo-Arabic, the feminine sg. morpheme *-a(t)* exhibits its complete shape only and always in the construct state (e.g. *midínat assulṭán* ‘the sultan’s city’), it being possible, however, that exceptions in both ways could take place in certain low registers (cf. Castilian *batafalúa* < **ḥábbat ḥalúwwa* ‘aniseed’, lit. ‘sweet grain’, where the insertion of /t/ is irregular, or the place-name Cantaralcadi, which reflects **qánṭara alqáḍi* ‘the judge’s bridge’, where the construct state would require *qánṭarat*).

Neither does Andalusi Arabic differ from Neo-Arabic in maintaining the constant shape of the dual and masculine sound plural morphemes, *-áyn* and *-ín*, without loss of their final consonant in the construct state (e.g. *ʿaynáyn al+ʿúmš* ‘the eyes of the blind’, *mudalliyín al+uḏnáyn* ‘crestfallen’), except in the case of dual parts of the body with possessive suffixes, e.g. *rijláy+na* ‘our feet’, *ʿaynáy+k* ‘your eyes’.

2.2 *Morphology*

2.2.1 Pronouns

2.2.1.1 The personal independent pronouns are: *aná* ~ *aní* ‘I’, *ánt(a)* or *át(ta)* ‘thou’, *hú* or *húwwa(t)* ‘he’, *hí* or *híyya(t)* ‘she’, *niḥín(at)*, *nuḥún*, *naḥán*, *(n)íḥna*, *aḥín*, *ḥínat* or *náḥnu* ‘we’, *ántum* ‘you’, and *húm* or *húma(t)* ‘they’. Gender distinction exists only in the 3rd pers. sg., although some lexical sources posit a 3rd pers. fem. pl. *húnna(t)*, which is not registered in the texts.

2.2.1.2 The matching possessive pronominal suffixes are: *-i* or *-y(a)*, *-(a)k*, *-u* or *-h*, *-a* or *-ha*, *-(i/a)na*, *-(u)kum*, *-(u)hum*. The same forms are used in the case of object suffixes, except for *-(a)ni* in the 1st. pers. sg. The optional forms with vocalic onset are used after consonant clusters or in order to avoid them, and full *-ya* only after *-ay* or *-i(y)*, e.g. *fíyya* ‘in me’, *bíyya* ‘with me’, but *ʾaxúy* ‘my brother’, *liwáy* ‘my banner’.

2.2.1.3 The demonstratives are, for the near deixis, *(há)ḏá* or *ḏí* ‘this’, *háwl(ay)* or *hawlín* ‘these’ and, for the remote deixis, *(há)ḏák* or *ḏík* 'that', *ḥáwlak* or *ḥáwlink* 'those'. There is no gender distinction, and the sg. can be substituted for the pl., it being questionable whether the prefixation of *há* introduces a third degree of deixis, as extant in substratal Romance and adstratal Berber. As adjectives, they always precede the qualified noun with the definite article.

2.2.1.4 The standard Andalusi Arabic relative has the invariable shape *allaḏí*, with the low-register allomorphs *allí*, *addí*, and *aḏḏí*. Standard Arabic or hypercorrect inflected shapes (e.g. *allatína* and *alliyát* for the fem. pl.) merely reflect interference by high registers, while the vernacular syntax occasionally allows the substitution of personal independent pronouns for the relative (e.g. *qíṭʿat árḍ híyya laddáyr* 'a plot which belongs to the monastery').

2.2.1.5 The interrogative pronouns are: *mán* 'who?', *má* or *áš(šu)* 'what?' and *áy (min)* 'which?', all of which, together with *ášma* 'whatever', may be used as (cor)relatives (e.g. *mará man tukún qaríbatak* 'a woman who is your relative', *kábš ma niḍaḥḥí* 'a ram that I can slaughter', *ášma yuqúl* 'whatever he says'). When used as interrogatives, they are stressed and open the sentence, while as (cor)relatives they are clitic and connect the antecedent with the consecutive phrases. By Romance interference, Andalusi Arabic allows relatives to be ruled by prepositions, instead of being represented by a recalling pronoun *(ḍamīr ʿāʾid)* at the end of the sentence (e.g. *aṭlúb šurráfa ʿalaš taʿtalí* 'look for a merlon from which you would throw yourself').

2.2.2 Adverbs

Andalusi Arabic adverbs include some innovations, like *dába* 'now', *makkár* 'at least' and *yáḏḏa* 'too'. It is noteworthy that adverbial *tanwīn* always has pausal reflexes in the vernacular (e.g. *ġadá ~ ġadÍ* 'tomorrow', *ḥáqqa* 'truly', *xáṣṣa* 'particularly'), and that preservation of /n/ is merely found in borrowings from the high register (e.g. *áwilan* 'firstly'). Among the interrogatives, *ašḥál* 'how much?' competes with *kám*.

2.2.3 Particles

2.2.3.1 The definite article has the shape *al*+ with an invariable and stable first /a/, as shown by examples like *naʿṭí alxúbz* 'I give the bread', and place-names such as Benialfaquí < *baní alfaqí* 'sons of the faqih'.

There was also an indefinite invariable article *wáḥ(i)d al*+ (e.g. *wáḥd alfarás* 'a horse', *wáḥd aṣṣabíyya* 'a girl'), probably due to interference by substratal Romance and adstratal Berber.

2.2.3.2 The invariable genitive marker *matá(ʿ)* was commonly used in Andalusi Arabic (e.g. *arrajúl matáʿha* 'her husband', *alqulúb matá nuḍḍáru* 'the hearts of those who see him').

2.2.3.3 Andalusi Arabic negative markers are manifold, with functional, diachronic, and diatopic distributions. *las* ~ *lis*, multifunctional and optionally incorporating pronominal suffixes, was in general use except in late periods, when it was replaced by *iš* (e.g. *las nuḥún ṣubyán* 'we are not children', *las nisammí aḥád* 'I mention nobody', *las akfá* 'it was not enough', *iš aní ṣáliḥ* 'I am not a saint', *iš ʿaṭáytu* 'I did not give it', *iš taʿmálu* 'you do not do it'). *ma*, a negative of verbal predicates above all, appears to have increased its frequency in later times, while *la*, except as the absolute negative adverb 'no' and in negative imperative or jussive nuances, is less common than in Standard Arabic.

2.2.3.4 As in other kinds of Arabic, *(ya) tará ~ turá* is used in order to introduce questions, e.g. *ya tará ikkín hu líyya*? 'could it be mine by any chance?', *tará lbašár yaltahám* 'maybe people will remember'.

2.2.3.5 The existential marker of Andalusi Arabic is *ṯám(ma)* 'there is' (e.g. *ṯám ʿilál* 'there are reasons', *ḥáddi ṯámma siwák* 'is there anybody but you?', or sometimes *báh* (e.g. *áy ʿúqla báh* 'which obstacle is there?'). With an exclamative nuance are found *ráni* 'lo, I am', *ráhi* 'lo, she is', *wará ḏá faxx* 'and lo, this is a trap', *taráh ahník* 'there he is', *awwaḏáni* 'here I am!'.

2.2.3.6 The functionals (prepositions and conjunctions) call for little comment, apart from a few innovations such as the preposition *biḥal* ~ *baḥál* 'like', the final conjunction *baš* 'in order to', or causal *ḥúrma fi* 'because of'.

2.2.3.7 The vocative markers are *ya*, *a*, and *al*+ (e.g. *ya rább* 'oh God!', *a ṭífli* 'oh my child', *alqamḥ aljidíd* 'oh fresh harvested wheat!'), which may be sometimes dispensed with.

2.2.3.8 Exclamations, oaths and curses offer little novelty, e.g. *áy xáḏla fíhum* 'how disappointing they are!', *áš yurá min manáḥis* 'how many calamities are seen!', *wáš qadár qálbi yahwák* 'how much my heart loves you!', *kam ḏá ṣudúd* 'how much scorn!', *bayáḏÍ* 'how lucky I am!', *a sawádi* 'poor me!', *ayyák tasál* 'beware of asking!', *ya ʿaláy taʿníq* 'if only I could hug him!', *háyya* 'come on!', *(w)alláh* 'by God!', *la kánu min ṣibyán* 'accursed boys!'.

2.2.4 Noun

2.2.4.1 Besides the leveling of feminine markers (2.1.1.2.2), Andalusi Arabic morphological gender rules differ little from other kinds of Arabic. There are unmarked feminine nouns most often continuing the situation in Classical Arabic, but these have been leveled in some cases (e.g. *ʿajúza* ‘old woman’, *ʿarúsa* ‘bride’), as generally in Neo-Arabic. Some feminine nouns have become masculine, most likely under the impact of substratal Romance, e.g. *alʿáyn alakḥál* ‘the black eye’, *aššáms ṭaláʿ* ‘the sun rose’ (cf. Castilian *ojo*, *sol*), as shown by converse cases in which the Romance feminine gender prevails, like *báyt* ‘house’ and *mawt* ‘death’ (Castilian *casa* and *muerte*).

2.2.4.2 The Classical Arabic derivational noun pattern system has considerably shrunk in Neo-Arabic, and above all in Western Arabic, which has forsaken many templates. Andalusi Arabic has retained reflexes of only CvCC, CaCvC, CiCaC ~ CuCaC, CāCaC ~ CāCiC, CāCūC, CayCaC ~ CaCCaC, CaCūC ~ CaCīC, CaCāC, CuCāC ~ CuCūC, CaCCāC, ʾaCCaC, maCCaC ~ maCCiC, maCCūC, miCCāC, CvCCā, CvCCān, CvCCūt, CaCCaC, CiCCiC, CaCCūC ~ CaCCīC, and CaCCaCa in more or less frequent use.

2.2.4.3 The sound plural morphemes have gained some ground at the expense of broken plural patterns (above all in adjectives, e.g. *ṣaʿbín* ‘difficult ones’, *aṣammín* ‘deaf ones’, *ʿazizín* ‘glorious ones’), where in the adjectives the masculine morpheme tends to supersede the feminine, e.g. *suqiyyín* ‘market women’, *xamriyyín* ‘brunettes’). However, both the sound feminine plural (e.g. *jarḥát* ‘wounds’, *baġlát* ‘mules’, mostly without anaptyctic vowel and exclusive for diminutives in Granadan, e.g. *rujaylít* ‘little men’, *uxayyít* ‘little brothers’) and broken plural patterns remain alive and productive in spite of abandoning certain templates and nuances, such as the so-called *pluralis paucitatis*.

Following are some examples of the most frequent broken plural patterns:

ʾaCCáC: *ajnáḥ* ‘wings’
ʾaCCúC: *arjúl* ‘feet’
ʾaCCíCa: *anʿíša* ‘stretchers’ (with an alternative ʾaCCáCa: *asárra* ‘beds’)
ʾaCCiCá: *aġniyá* ‘rich ones’
CúCC: *rúxṣ* ‘tender ones’, *ṭúrq* ‘roads’
CuCúC: *quṣúr* ‘castles’, *kutúb* ‘books’
CuCúCa: *ḏukúra* ‘penises’
CiCáC ~ CiCÍC: *ṯiyáb* ‘clothes’, *ibár* ‘needles’, *rimíl* ‘sands’
CaCíC: *ḥamír* ‘donkeys’
CaCáC: *xadám* ‘female slaves’, *daráq* ‘shields’
CuCáC: *luqám* ‘morsels’
CuCaCá ~ CuCaCí or CuCáCa: *fuqahá* ‘faqihs’, *ʿuzazí* ‘glorious ones’, *usára* ‘captives’.
CuCCáC ~ CuCCíC: *fussáq* ‘perverts’, *xunníṯ* ‘hermaphrodites’, with an alternative CuCáC in defective roots, e.g. *rumá* ‘bowmen’.
CaCCá ~ CaCCí: *marḍá* ‘sick ones’, *ḍaʿfí* ‘weak ones’.
CaCáCa: *waráṯa* ‘heirs’, *labába* ‘wolves’, *ʿaḏára* ‘maidens’.
CiCCán ~ CuCCín: *ʿiqbán* ‘eagles’, *fursín* ‘knights’.
CaCáCiC(a): *fanádiq* ‘inns’, *falásifa* ‘philosophers’.

2.2.4.4 The dual, marked with the endings *-áy(n)* or *-ín*, is nearly restricted to certain nouns, double by nature or countable, such as the names of body parts, weights, and measures, e.g. *saqáyn* ‘legs’, *ʿaynín* ‘eyes’, *uḏnáy* ‘ears’, *martáy* ‘twice’, with some exceptions, like *rajuláy* ‘two men’. Otherwise, its expression is analytical, e.g. *záwj rixáx* ‘two castles’, *záwj min alqurún* ‘two horns’. However, in the case of the double parts of the body, such forms are in fact a → pseudo-dual, i.e. they can mean the plural as well, e.g. *taftáḥu ʿayníkum* ‘you [pl.] open your eyes’, *alḥiṭán laha uḏnáy* ‘walls have ears’, *yijú ʿala saqáyhum* ‘they return on foot [on their feet]’.

2.2.4.5 The diminutive is very productive in Andalusi Arabic, with the templates CuCáyyaC, fem. CuCáyCa, for triconsonantal forms and CuCáyCaC(a) for the quadriconsonantic, e.g. *kuláyyab* ‘little dog’, *buġáyla* ‘little mule’, *muráykab* ‘little ship’, *suráysala* ‘little chain’. Defective nouns have CuCáy (e.g. *ṣubáy* ‘little boy’), some adjectives have a special template CuCáyCaC (e.g. *kubáybar* ‘slightly big’, and Romance diminutive suffixes are also used (e.g. *ʿarus*+ÉLLA ‘weasel’, literally, a euphemistic ‘little bride’).

2.2.4.6 Among the adjectival patterns remaining in use in Andalusi Arabic some call for special comments, like ʾaCCáC, characteristic of colors and defects, which preserves inflections very close to those of Classical Arabic, e.g.

aḥmár ‘red’, fem. *ḥámra* and pl. *ḥúmr*, but it is noteworthy that this fem. suffix behaves exactly like *-a(t)* (e.g. *zárqat alyamáma* ‘the blue [woman] from Alyamāmaʾ), and that the disjunctive vowel which eventually developed in the plural template is often phonemicized (e.g. *kuḥál* ‘black ones’, *buláh* ‘stupid ones’).

2.2.4.7 The elative pattern ʾaCCáC is frequently used in Andalusi Arabic, e.g. *aṯqál* ‘heavier’, *aqdám* ‘older’, occasionally also in exclamative utterances, e.g. *má ashál* ‘how easy!’. This happens, as is common in Neo-Arabic, even with adjectives of this same template, e.g. *má abyáḍu* ‘how white he is!’. In later stages, however, analytical idioms become frequent, e.g. *ġáli akṯár* ‘more expensive’, *akṯár akḥál* ‘darker’.

2.2.5 Numerals

From 3 to 10 Andalusi Arabic has two series of numerals, the first one short, used in the construct state, *ṯaláṯ*, *arbáˁ*, *xáms*, *sítt*, *sábˁ*, *ṯamán*, *tísˁ*, *ˁášr*, and the second one long, used in the absolute state, *ṯaláṯa*, *arbáˁa*, *xámsa*, *sítta*, *sábˁa*, *ṯamánya* ~ *ṯamínya*, *tísˁa* and *ˁášra*, although an *-at* interfix appears at times in front of numbered items beginning with a vowel (e.g. *arbáˁ ayyám* ‘four days’, but *taláṯat ašyá* ‘three things’). As for ‘one’, *wáḥid* and fem. *wáḥida* are used, as in Standard Arabic, only to emphasize oneness, occasionally replaced by a preceding invariable *fárd*, e.g. *fárdi márra* ‘once’. ‘Two’, *iṯnáy(n)*, which does not distinguish gender either, is mostly restricted to compound numerals, otherwise being replaced by *záw(j) (min)* ‘a pair of’.

From 11 to 19, there is no gender distinction in the series *ḥidáˁšar*, *iṯnáˁšar*, *ṯalattáˁšar*, *arbaˁtáˁšar*, *xamistáˁšar*, *sittáˁšar*, *sabˁatáˁšar*, *ṯamantáˁšar*, *tisˁatáˁšar*, with hints of loss of /ˁ/ in the second constituent and compensatory velarization of the preceding dental, as well as of decay of final *-ar*, like in North African Arabic. Higher numerals are also invariable (e.g. the tens, *ˁišrín*, *ṯalaṯín* ~ *ṯaliṯín*, *arbaˁín*, etc., the hundreds, *míyya*, *mitáy(n)*, *ṯalaṯmíyya*, *arbaˁmíyya*, *xamsumíyya*, *sittumíyya*, *sabˁamíyya*, *ṯaminmíyya*, *tisˁamíyya*, and the thousands, *álf*, *alfáy(n)*, *ṯalát álaf*, *arbáˁ álaf*, etc.).

As for other series of numerals, in the case of ordinals it is noteworthy that ‘first’ has been adapted to the template CāCiC, i.e. *áwil*, *ṯáni*, *ṯáliṯ*, etc., for which the dialect of Valencia exhibits the allomorphs *awilí*, *ṯaliṯí*, *arbaˁí*, *xamsí*, etc.

2.2.6 Verb

2.2.6.1 The Andalusi Arabic triconsonantal verb preserved the Forms I (simple) and derived II, III, IV, V, VI, VII, VIII, IX–XI, and X, from which every verb has full paradigms for the perfective and imperfective aspects and the imperative mood, in addition to the nominal derivates known as *maṣdar* and participles.

The system of characteristic vowels expressing semantic values in Form I of the Classical Arabic verb, never altogether effective because of phonetic affinities between the vowels and certain consonants, has left only some traces in Andalusi Arabic, such as a considerable number of stative verbs with *-u-* as characteristic vowel of both perfect and imperfect (e.g. *qarúb*, *yaqrúb* ‘to be near’, *ḍaˁúf*, *yaḍˁúf* ‘to be weak’, *samúj*, *yasmúj* ‘to be ugly’) and the preservation of a host of active verbs with the alternation *-a-* ~ *-u-* in the perfective and imperfective themes. But every alternation based upon the contrast between /a/ and /i/ has been leveled into steady /a/, and other phonetic and analogical phenomena have altered the old situation considerably.

The derived forms of the Andalusi Arabic verb are the same in the themes of perfect and imperfect, with the partial exception of Form IV, which is no longer productive, and probably was felt as a mere freak variant of Form I, but included such commonly used verbs as *axráj*, *yaxríj* ‘to put out’ and *adxál*, *yadxíl* ‘to put in’. The remaining forms had the paradigms: II CaCCáC, III CáCaC, V atCaCCáC, VI atCáCaC, VII anCaCáC, VIII aCtaCáC, XI–XI aCCáCC, X astaCCáC. Their semantic functions are standard, but only Forms II and VII were really productive. There were substitutions, like Form II for IV or III) and combinations, like Form II + X.

The non-agentive voice or internal passive, with its characteristic vocalic marking CuCíC, has survived in Andalusi Arabic better than in the rest of Neo-Arabic. However, it appears that Form VII tends to take its place, as in other dialects, and that certain non-agentive expressions were converted into agentive, e.g. *atwaffá* ‘to pass away’.

2.2.6.2 Inflectional paradigms

The perfect expresses person, number, and gender with the following suffixes: sg. 3rd masc. =

Ø, 3rd fem. *-at*, 2nd *-t*, 1st *-t*; pl. 3rd *-u*, 2nd *-tum*, 1st *-na*. When a pronominal suffix is added, the 2nd pl. lengthens its suffix with a stressed *ú*, e.g. *katabtumúh* 'you wrote it'.

The imperfect expresses person, number, and gender with the following prefixes: sg., 3rd masc. = *y*V-, 3rd fem. *t*V-, 2nd *t*V-, 1st *n*V-, to which a suffix *-u* is attached in the pl., except for the 3rd fem., which is the same as the masc. The prefix vowel is /a/, except in Forms II and III, vocalized with /i/ more often than with /a/ or /u/. A prefix *sa-* indicates the future, while *ki(n)-*, with assimilation of /n/ to the person prefixes, implies eventuality.

The imperative is expressed with the simple stem, having only a sg., and a pl. with the suffix *-u*. When the stem begins with two consonants, this is avoided with a prefixed *a-*, e.g. *aktúb* 'write! [sg.]', *aktúbu* 'write! [pl.]'.

The *maṣdar*s or verbal nouns of Andalusi Arabic differ only a little from their Classical Arabic counterparts, since they belong mostly to high registers.

The same applies to participles, both agentive and non-agentive. However, in the derived forms the voice opposition tends to disappear, e.g. *muʿallám* 'teacher', *muslamín* 'Muslims'.

2.2.7 Weak verbs

Geminate verbs follow the Standard Arabic rules, but cases of uncontracted or haplological forms are more frequent (e.g. *nastaḥbáb* 'I make myself loved', *astaḥast* 'I spied'). Unusual for Neo-Arabic, there is no substitution of -Cay- for -CaC- in front of the perfect consonantal suffixes, e.g. *ḥabábt* 'I loved'.

Hamzate verbs often depart from the standard forms by (a) decay of /ʾ/ without compensation (e.g. *kálna* 'we ate', *yaḏḏán* 'he calls to prayer', *tasál* 'you ask', *ráyt* 'I saw'; (b) decay of /ʾ/ with compensation by stress shift or consonantal lengthening (e.g. *nákul* 'I eat', *naxxúḏ* 'I take'); or (c) substitution of /w/ or /y/ for /ʾ/ (e.g. *waxádt* 'I took', *badáyt* 'I began', *nirayyás* 'I start').

I*w* verbs often exhibit forms departing from the standard rules either by loss without compensation of radical /w/ (e.g. *nastaṯáq* 'I trust', *nastaqáʿ* 'I covet') or its aberrant preservation (e.g. *nawṯáb* 'I jump', *yawtaqád* 'it burns').

II*w*/*y* verbs follow the standard rules most of the time, although the prefix vowel of Form I is idiolectically variable (e.g. *tizúr* 'you visit', *tuqúl* 'you say'), and non-agentive perfects may exhibit a prefixed *u-* (e.g. *uqíl* 'it was said').

III*y* verbs, which have absorbed the roots IIIʾ, appear almost exclusively in two types, CaCá ~ yaCCí and CaCá ~ yaCCá, as is characteristic of Western Arabic. There is hesitation in the treatment of /i-u/ sequences (e.g. *tamšú* 'you walk' ~ *yabníw* 'they build', as well as in the preservation of thematic alternation in Forms VII, VIII, and X (e.g. *yaxtabí* 'he hides himself' vs. *yamtaḥá* 'it is erased'). Non-agentive participles of Form I always have a *mu-* prefix (e.g. *murmí* 'thrown').

2.2.8 Quadriconsonantal verbs

These do not call for a specific treatment, as they follow the pattern of Forms II and V of the triconsonantal, mostly in agreement with standard rules. Many of them are innovations, obtained from loanwords, repetition of biconsonantal roots or of the last radical consonant in the triconsonantal, dissimilation of an /r/ in Form II, infixation of /w/, or suffixation of /n/, etc.

2.3 *Syntax*

2.3.1 Noun phrase

Annexation in Andalusi Arabic has yielded some ground to the analytical genitive. As for the use of the definite article in annexation and qualifying syntagms, it is noteworthy that the standard rules are often infringed, for instance in compound substantives (e.g. *alḥabb almulúk* 'the cherries', *alwalád ziná* 'the son of a whore'), or by omission in the head of a qualifying syntagm (e.g. *riḥá aljidíd* 'the new mill', *masjíd alaxḍár* 'the Green Mosque').

Early Andalusi Arabic used indefinite nouns connected to a following qualifier by means of an interfix *-an*, a reflex of older *tanwīn* (e.g. *zamán+an áxar* 'another time', *qálban jáyd* 'good will', also introducing relative clauses, e.g. *zamán+an qad bád* 'a time which is already bygone', *dár+an fíh zawáj* 'a house in which there is a couple'). It disappeared in later epochs, apart from a few idioms.

2.3.2 Verbal phrase

Verbal phrases are integrated by a finite verb and optional extensions such as a direct or indirect object, other complements, and adverbs. The object is sometimes introduced with the preposition *li-* (e.g. *aštúm li-wíldi* ‘curse my father’, *rá li-ʕáynak* ‘he saw your eyes’).

The process of abandonment of aspectual connotations in favor of tense, already begun in Classical Arabic and advanced in Neo-Arabic, is nearly concluded in Andalusi Arabic, but for certain optative idioms (e.g. *alqá lláh fi rásu ḍárbat šuqúr* ‘may God send a blow with an ax on his head!’, *la nasáytu* ‘may I never forget it!’). Perhaps for this very reason, the aspect booster *qad* has evolved into a mere adverb of manner (e.g. *qad tadrí* ‘you certainly know’, *qad tammámt* ‘I have really finished’, *qad aná maḏ̣lúm* ‘I am wronged indeed’).

There are some instances of narrative imperative, like *bítna fi riḍá qabbál aw ʕannáq* ‘we spent the night in a friendly manner, kissing and embracing’.

2.3.3 Word order

Word order in Andalusi Arabic, like the rest of Neo-Arabic, is more fixed and linear than in Classical Arabic, with general precedence of subjects over predicates, substantives over adjectives, verbs over complements, etc., but there are cases of emphatic inversion (e.g. *záwj úmmak nakún* ‘I am your mother’s husband’, *ḥalál hú* ‘it is lawful’.

2.3.4 Agreement

Agreement in Andalusi Arabic tends to be natural (e.g. *assaḥáb yusáqu* ‘the clouds are carried’, *xudaydát ḥumár* ‘red cheeks’), but the standard concord of inanimate pl. with sg. fem. is still often observed (e.g. *ḏukírat almudún* ‘the cities were mentioned’, *mirár kiṯíra* ‘many times’). Duals can be treated in either way, e.g. *ʕaynáyn súd* ‘black eyes’, *xubzatáyn takfíni* ‘two loaves are enough for me’.

2.3.5 Marginal phrases

Marginal phrases, functionally equivalent of adverbs and integrated by a preposition and the noun governed by it, may be used as extensions of a nominal or verbal phrase or as predicate of the former (e.g. *aní falbáyt* ‘I am in the house’). Such syntagms can convey alienable or inalienable possession, obligation, etc. (e.g. *máʕi maʕšúq* ‘I have a love’, *áš ʕaláyna min ḏák* ‘what does it matter to us?’, *lú an yumút* ‘he must die’, cf. Castilian *tiene que morir*).

2.3.6 Complex sentences

Nominalized sentences can be obtained by mere juxtaposition in the case of certain verbs (e.g. *tiríd tará* ‘you want to see’, *dáʕni nafráḥ* ‘let me rejoice’), through an appropriate marker (e.g. *niríd an niqabbál* ‘I want to kiss’, *abát ma tudúr* ‘she refused to turn’) or, less often, by the use of the *maṣdar* (e.g. *qad ḥán inqiṭáʕak* ‘the time for your departure has arrived’).

Causal, final, temporal, and modal sentences offer little novelty, except for some innovated markers (e.g. *sídi mašġúl kamá ṭaláʕ larruqád* ‘my master is busy, because he just went up to sleep’, *faríḥat kayf yaḏháb* ‘she was glad that he left’, *fi ḥáqqat yuqál liḏá ʕanbarí* ‘so that this be called amber-like’, *rabbá janáḥ baš yiṭír* ‘he grew wings to fly’, *jáni an yaftaqád ḥáli* ‘he came to ask for my news’, *kamá aṣbáḥ labás ṯiyábu* ‘when he got up in the morning, he put on his clothes’, *biḥál iḏ bátat fi ṣahríj* ‘as if they had spent the night in a pool’).

Conditional sentences are introduced by the conjunctions *in (kán)* (also *ikkán ~ ikkín*), *id(á)*, and *law* in ways similar to Standard Arabic (*in katáb naḏ̣ám fi kútbu aljáwhar* ‘if he writes, he strings gems in his books’, *iḏá ḥább qatál* ‘if he loves, he kills’, *law kán biwúddi ma maššáyt ʕannu aḥmáq* ‘had I had any choice, I would not have gone crazy over him’), apart from some deviations, like the frequent abandonment of the main clause markers *fa-* and *la-*, and the substitution of the imperfect, often with prefixation of *kán ~ kín* as a mark of eventuality, for the perfect forms in either or both clauses (*iḏa ráytu nabhát* ‘when I see him, I am confused’, *iḏa turá alákl aqrúb* ‘when you are shown food, come close’, *law kán falbúm xáyr ma kiyyaslám ḥala ṣṣayyád* ‘if there were anything good in owls, they would not be safe from hunters’).

3. LEXICON

The Andalusi Arabic lexicon is integrated by a main core of Arabic items and about 3 percent of mostly Romance loanwords, including some dozens of Berber items. Needless to say, those Arabic, Romance, and Berber elements could be inherited or borrowed from other languages, such as Latin, Greek, Persian, Aramaic, Hebrew, Coptic, etc. (→ Ibero-Romance loanwords).

With respect to Classical Arabic items, there has been a considerable evolution due to (a) replacement or abandonment of lexemes; (b) semantic shifts; and (c) morphological reshaping.

Bibliographical references

Alcalá, Pedro de. 1505. *Arte para ligera mente saber la lengua araviga y vocabulista arauigo en letra castellana*. Granada. Ed. by Paul de Lagarde, Göttingen 1993. (Repr. Osnabrück 1971.)
Ayala, Martín de. 1566. *Doctrina Christiana en lengua arabiga y castellana*. Valencia. (Repr. 1911.)
Binšarīfa, Muḥammad. 1971–1975. *ʾAmṯāl al-ʿawāmm fī l-ʾandalus*. Fez: Maṭbaʿa Muḥammad al-Xāmis.
Colin, Georges. 1960. "Al-Andalus: L'Arabe hispanique". *Encyclopédie de l'Islam*, I, 516–19. 2nd ed. Leiden: E.J. Brill.
Corriente, Federico. 1977. *A grammatical sketch of the Spanish Arabic dialect bundle*. Madrid: Instituto Hispano-Árabe de Cultura.
——. 1988. *El léxico árabe andalusí según P. de Alcalá*. Madrid: Universidad Complutense.
——. 1989. *El léxico árabe andalusí según el Vocabulista in arabico*. Madrid: Universidad Complutense.
——. 1992. *Árabe andalusí y lenguas romances*. Madrid: MAPFRE.
——. 1994. "Textos andalusíes de céjeles no quzmanianos en Alḫillī, Ibn Saʿīd Almaġribī, Ibn Xaldūn y en la Genizah". *Foro Hispánico* 7.61–104.
——. 1997. *A dictionary of Andalusí Arabic*. Leiden: E.J. Brill.
Ferrando, Ignacio. 2000. "The Arabic language among the Mozarabs of Toledo during the 12th and 13th centuries". *Arabic as a minority language*, ed. Jonathan Owens, 45–63. Berlin and New York: Mouton de Gruyter.
Harvey, L.P. 1971. "The Arabic dialect of Valencia". *al-Andalus* 36.81–115.
Ould Mohamed Baba, Ahmed Salem. 1999. *Estudio dialectológico del refranero andalusí de Abū Yaḥyā Azzajjālī*. Saragossa: Universidad de Zaragoza.
Singer, Hans Rudolf. 1981. "Zum arabischen Dialekt von Valencia". Orbis 27–28.316–323.

Federico Corriente (University of Zaragoza)

Antiochia Arabic

1. General

1.1 Geographical area

Arabic in Antiochia is spoken by Sunnis, Alawis (Nuṣayris), Christians, and by the Jewish community of the city of Antakya. Until 1999 a small Arabic-speaking community also existed in the city of Iskenderun. In the second half of the 20th century, many Arabs left Antiochia and settled in Europe, so that, for example, the Christian community of Yayladağ no longer exists.

All of the Alawites live in the Western part of the province of Antioch, west of a line from Iskenderun to Kılıçtutan. The Sunnis live east of this line except for the Arabic-speaking Sunnis in the city of Antakya. Bedouin have settled in the town of Reyhanlı and in the area of the former lake of ʿAmq. Their dialects, as is generally known, differ greatly from the dialects of the sedentary populations. The Christians live in the two big cities of the Antioch region, Antakya and Iskenderun, as well as in a few smaller places.

The total number of Arabic speakers in the region is about 200,000, the majority of them (170, 000–180,000) are Alawis. The estimated number of Arabic-speaking Sunnis is 20,000. The Christian and Jewish communities today are very small; probably fewer than 5,000 Christians and no more than 50 Jews live in the region today.

1.2 *Speakers' lifestyle*

The inhabitants of the villages are farmers. Along the coast there are also fishermen. The inhabitants of the cities are tradesmen, craftsmen, employees in the administration of the city, and unskilled workers. The Bedouin of the region probably gave up their nomadic way of life two or three generations ago but they continue to call themselves 'Bedouin' (*badu*).

1.3 *Position within the dialects of the area and within society*

Linguistically the Arabic dialects of the province of Antioch lie on the northwestern periphery of the Syro-Palestinian dialect area. An offshoot of these dialects also exists in the Çukurova region, where Arabic-speaking Alawites settled mainly in the 20th century (Procházka 2002). The dialect of the Jews is very similar to the dialect of Aleppo, in particular to the dialect of the Jews of Aleppo (Nevo 1991) to which rabbinate the Jewish community of the Antioch region belonged until 1938.

The Bedouin dialects in the Antioch region are closely connected with the Bedouin dialects in northern and northeastern Syria, spoken by the Šāwi Bedouin.

Though the Arabic dialects of the Antioch region are linguistically connected with the dialects spoken in neighboring Syria, there is little or no influence from Modern Standard

Arabic, as these dialects are spoken outside the Arabic world.

Since Antiochia became a province of Turkey, Arabic has been completely banished from public life and therefore the teaching of Arabic, the singing of Arabic songs in public, and the use of Arabic personal names are forbidden (→ Turkey). A few people have learned with difficulty through a sheikh to decipher the Arabic letters. Others have tried to learn Standard Arabic by themselves, with little success. Most of the Arabs in Antioch today are illiterate in Arabic.

1.4 *Historical evidence*

Antiochia was conquered by the Arabs in 637. It was part of Syria and under French mandate until 1938 when it was given to Turkey. It became a Turkish province called Hatay with the capital city Antakya, the former Antiochia. The dialect of the village of Samandağ was the first dialect of this region to be described by Bergsträsser (1915).

1.5 *State of research*

The dialect of the village of Samandağ is the only one in the region to have found its way into Bergsträsser's *Sprachatlas* (1915). A text from an Alawi village was published by Stroomer (1991). More detailed information about the linguistic situation can be found in Behnstedt (1996), where the first information about the dialects of the Bedouin can also be found. The main source is Arnold (1998).

2. LINGUISTIC DESCRIPTION

2.1 *Phonology*

2.1.1 Inventory

2.1.1.1. Consonants

The majority of the sedentary dialects have consonantal inventory in Table 1.

Table 1. Inventory of consonants in sedentary dialects

b	t	ṭ	d	ḍ			k		q			
f	s	ṣ	z		š	j	x	ġ		ḥ	ʿ	h
m	n	l	r									
w					y							

Marginal consonants are /g/, /ẓ/, /ʾ/ and also, in the dialects of the Christians, /ž/. The city dialect of the Alawites and Sunnis of Antakya has a postvelar /ḳ/ instead of /q/. Two Sunni villages near the Syrian border and the Jews have /ʾ/ instead of /q/. The old interdentals shifted to the corresponding plosives. Especially in the Alawi dialects **ḏ* in some words is replaced by /z/.

The Bedouin have mainly the system in Table 2.

Table 2. Inventory of consonants in Bedouin dialects

b	t	d	ṭ				k	g			
f	ṯ	ḏ	ḏ̣								
	s	z	ṣ	š	j	č	x	ġ	ḥ	ʿ	h
m	n	l	r								
w				y							

In all Bedouin dialects new velarized allophones occur (*gāḷ* 'he said', *ṃayy* 'water'). In two Bedouin villages the interdentals are replaced by the corresponding dentals. In many Bedouin dialects /ġ/ has the allophone /q/ when followed by a vowel (*qēm* 'cloud', *muġrib* 'evening'). Old **q* is replaced by /g/, and in some dialects in the vicinity of front vowels also by /j/ (*rifīj* 'friend'). Old **k* is replaced by /č/ in contact with front vowels and always as suffix of the 2 sg. fem.: *abūk* 'your [masc.] father', *abūč* 'your [fem.] father'.

2.1.1.2. Vowels

The sedentary dialects of the Alawites, Sunnis, and Christians have the two short vowels /a/ and /i/ (< **i* and **u*), the Jews have /a/ and /ə/ (< /i/ and /u/). In word final syllables of the structure -CvC in all Sunni dialects, in the Christian dialects of Antakya and the northern part of Antioch, and in a few village dialects of the Alawites the vowel /u/ is preserved. This is also the case in the dialects of the Jews but the vowels /i/ and /u/ in this position are lowered to /e/ and /o/. In many dialects of the Christians and Alawites an allophone /u/ also occurs after the labial /ṃ. In the majority of the village dialects of the Alawites, in closed syllables the vowel /a/ in word final position is realized as /o/ in the vicinity of emphatic consonants (*abyoḍ* 'white'), as /a/ in the vicinity of back consonants (*azraq* 'blue'), and as /i/ in the vicinity of front consonants (*aswid* 'black'). In all dialects of the Alawites and the Christians *a* shifts to /i/ in pre-stressed syllables if the vowel of the following syllable is /a/ (*ṣálla* 'he prayed', *ṣilláyt* 'I prayed').

In Antioch different inventories of long vowels can be found. A system with the three long vowels /ā/, /ī/, and /ū/ exists in some dialects of Christians and Alawites in which the old diphthongs are preserved and in which either no → *ʾimāla* or an *ʾimāla* /ā/ > / ī / occurs. A system of the four vowels /ā/, / ī /, / ū /, and /ō/ is found in those dialects of the Alawites in which monophthongization of **ay* > /ā/ and **aw* > / ō / took place. The inventory /ā/, /ī/, /ū/, and /ē/ exists in dialects in which the diphthongs are preserved and an *ʾimāla* /ā/ > /ē/ occurs. This is the case in all city dialects (except that of the Jews), in most dialects of the Sunnis, and in some dialects of the Alawites and the Christians. A system with the five long vowels /ā/, /ä/-, /ī/, /ō/, and /ū/ is found in two Alawi village dialects in which the diphthongs **aw* and **ay* shifted to /ō/ and /ā/, through which the allophone /ä/ of **ā* became a phoneme (*bǟt* < *bāt* he stayed overnight, *bāt* < **bayt* house). The Alawi dialect of Harbiye has a similar system but with /ɔ̄/ instead of /ā/. Finally all dialects, in which the diphthongs **aw* and **ay* are monophthongized to /ō/ and /ē/, have an inventory with /ā/, /ē/, /ī/, /ō/, and / ū /. This system is found in the Jewish dialects, in the Sunni dialect of Reyhanlı, and in two Alawi villages, where /ē/ is also a result of the *ʾimāla*. In many Alawi village dialects, the vowel /ā/ is split into /ā/ and /ō/ in word-final syllables, where /ō/ occurs in the vicinity of emphatic consonants (*fillāḥ* 'farmer', but *xiyyōṭ* 'tailor').

In word-final position, the Alawis and Christians have three vowels /a/, /i/, /u/, the Sunnis have four /a/, /i/, /u/, /e/), and the Jews five /a/, /i/, /u/, /o/, /e/.

The Bedouin dialects have the short vowels /a/ and /i/. In the vicinity of labials, back consonants and emphatics (/u/-coloring environments) /u/ also occurs (*gumar* 'moon', *uxt* 'sister'). Old **a* is often replaced by /i/, /u/ as in *rguba* < **ragaba* 'neck', *sina* < **sana* 'year'. Unstressed and even stressed short vowels in open syllables are often elided (*šrab* 'he drank').

The Bedouin dialects have in addition to the old long vowels /ā/, /ī/, /ū/, the new vowels /ē/ and /ō/ resulting from monophthongization of **ay* and **aw*.

2.1.1.3 Diphthongs

The diphthongs /aw/ and /ay/ are preserved in open and closed syllables in all dialects of the Christians, with two exceptions, also in all Sunni sedentary dialects, in the city dialects of the Alawites, and in some Alawi village dialects. In a few Alawi villages in the North of Antioch the diphthongs are realized as [ow] and [ey]. Most of the Alawi village dialects have the diphthongs /aw/ and /ay/ only in open syllables but /ō/ and /ā/ in closed syllables (*yawmān* 'two days'; *baytān* 'two houses'). In two Alewi villages the monophthongization of /ay/ in closed syllables /ē/ is found, as in *baytēn*, whereas the little town of Harbiye has /ɔ̄/ as in *baytɔ̄n*. In some Alawi village dialects monophthongization in closed syllables is limited to /aw/, while /ay/ is preserved (*yōm* 'day', *bayt* 'house'). In the dialects of the Bedouin, of the Jews, of one Alawi village in the south, and in the Sunni sedentary dialect of the town of Reyhanlı and a neighboring village, the diphthongs in all syllables are monophthongized to /ō/ and /ē/. In these villages, monophthongization was probably influenced by the Bedouin dialects.

2.1.1.4 Syllable

All dialects of the region have the following syllable types:

Cv, Cv̄, CvC, Cv̄C, CvCC, CCv, CCv̄, CCvC, CCv̄C, CCvCC. In the dialects of the Christians and Alawis CCCvCC is also attested (*štrayt* 'I bought').

2.1.1.5 Consonant clusters

An anaptyctic vowel [ə] may be inserted between two consonants in word-final position or after the first consonant in a cluster of three consonants. In Bedouin dialects, the anaptyctic vowel is [u] after /u/ in the preceding syllable, otherwise [i] (*jubun* 'cheese', *baṭin* 'belly').

2.1.1.6 Stress

In all dialects of the region, stress lies on the last syllable if it is long (-Cv̄C and -CvCC). If the last syllable is short, the stress lies in most dialects of the Christians and Alawites on the penultima, but in the dialects of Sunnis with some exceptions on the first long syllable (*fáttašitu* 'she inspected him'). Stress is significant in the dialects of the Sunnis.

Christians/Alawites	*yiktíbu*	‘they write/ he writes it’
Sunnis (Kamberli)	*yíktibu*	‘they write’
	yiktíbu	‘he writes it’

2.1.2 Phonotactics

The → *gahawa* syndrome is attested in all Bedouin dialects of Antioch (*aḥámar*, in some dialects *ḥamar* ‘red’).

The overwhelming majority of the sedentary dialects have an *i*-conditioned *ʾimāla* up to /ē/. In some Christian and Alawi villages south of the capital city of the province an *ʾimāla* up to /ī/ is found. Only a few dialects at the periphery of the region such as the communal dialects of Iskenderun in the north, the Sunni dialect of Reyhanlı in the east and the dialect of the Christians in Yayladağ in the south have no *ʾimāla*.

In some Christian and Alawi villages lowering or diphthongization of vowels in pause is attested:

Limanköyü:	*inti* > *inta*	‘you [fem.]’
Altinözü:	*ana* > *ano*	‘I’
Yaylıca:	*inti* > *intey*	‘you [fem.]’
Mağaracık:	*nsīt* > *nsēyt*	‘I forgot’

In a few villages glottalization and/or devoicing of /ʿ/ in pause is attested: *simiʿ* ‘he heard’ > [simiʔħ] (Yayladağ), [simiʔ] (Altınözü), *simiħ* (Koyunoğlu). Glottalization of /ʿ/ in pause also occurs in some Bedouin dialects. In word-final position devoicing of /b/, /d/, and /g/ is attested in the Bedouin dialect of Han Muratpaşa (*nišat* ‘he asked’, but *nišadum* ‘they asked’).

2.1.3 Morphophonology

The short vowel /i/ is elided in all sedentary dialects in unstressed open syllables. This is also the case for the vowel /a/ in all dialects of Christians and Alawites, while in Sunni and Jewish dialects /a/ is normally preserved:

‘I heard/ I wrote’

differential:	Sunnis/Jews:	*smiʿt/katabt*
non-differential:	Alawites/ Christians:	*smiʿt/ktabt*

Suffixation:

3rd sg. fem. perf.+ suffix:
‘she hit him’

Christians	*ḍarbit* + *-u* > *ḍarbítu*
Sunnis	*ḍarabit* + *-u* > *ḍarabítu*
Jews	*ḍarbet* + *-o* > *ḍarbəto*
Alawis (Antakya)	*ḍarbit* + *-u* > *ḍaribtu*
Alawis (villages)	*ḍarbit* + *-u* > *ḍarbītu*
Bedouin	*ḏ̣irbat* + *-u* > *ḏ̣irbattu*

3rd pl. masc. perf. + suffix:
‘they forgot him’

Sedentaries	*nisu* + *-u* > *nisū*
Bedouin	*nisu* + *-u* > *nisō*

3rd pl. fem. perf. + suffix:
‘they [fem.] forgot him’

Only Bedouin	*nisan* + *-u* > *nisanno*

Double suffixation is attested in the dialect of the Jews: *bilabbsawe* ‘he dresses her in it’.

2.2 *Morphology*

Gender distinctions in 3rd person pl. and 2nd person pl. is attested only in the Bedouin dialects of Antioch. The Christians in the city of Antakya and in the villages in the south of the region have *inti* for 2 sg. masc. and 2 sg. fem. of the independent personal pronoun.

2.2.1 Pronouns

2.2.1.1 Independent

3rd sg. masc.: *hūwi* (Christians, most Alawis), *hūwe* (Sunnis, Jews), *hūwa* (Sunnis in Babatorun), *hūwit* (some Alawi dialects in the southwest of the region) *huwwa* ~ *huwwe* (Bedouin); the short form *hū* is also used by many speakers of all dialects.

3rd sg. fem.: *hīyi* (Christians, most Alawis), *hīye* (Sunnis, Jews), *hīya* (Sunnis in Çetenli, and Alawis in the southernmost part of the region), *hīyit* (some Alawi dialects in the southwest of the region), *hiyya* ~ *hiyye* (Bedouin); the short form *hī* can alternatively be used in all dialects.

3rd pl. com.: *hinni* (Christians, most Alawis), *hinne* (most Sunni dialects), *hənnen* (Jews), *hinnin* (Sunnis in Reyhanlı and Alawis in Koyunoğlu), *hinn* (Alawis in Kuzeytepe), *hinnit* (two Alawi dialects in the southwest of the region). Bedouin: *humma* (pl. masc.), *hinna* ~ *hinne* (pl. fem.).

2nd sg. masc.: *int* (most Alawi dialects, Christians in the south), *inti* (Christians in the north and in the city of Antakya), *ənte* (Jews), *inte* (Sunnis), *inta* (Alawis in the city of Iskenderun), *hint* (some Alawi villages in the southernmost part of the region). Bedouin: *inta* ~ *inte*.

2nd sg. fem.: *inti* (most dialects of the Alawis, all Christians and Sunnis), *ənti* (Jews), *hint* (some Alawi villages in the southernmost part of the region). Bedouin: *inti*.

2nd pl.: *intu* (most dialects of the Alawis, all Christians and Sunnis), *əntu* (Jews), *hintu* (some Alawi villages in the southernmost part of the region). Bedouin: masc. *intum* ~ *intam*, fem. *intan*.

1st sg.: *ana* (all sedentary dialects). Bedouin: *āni*.

1st pl.: *niḥna* (most dialects of the Alawis and Sunnis, all Christians), *nəḥna* (Jews), *niḥni* (some Alawi villages in the central part of the region), *niḥne* (southeastern Sunni villages), *niḥᵊn* (in the Alawi village of Kuzeytepe), *naḥni* (in some Alawi villages), *naḥne* (in the Sunni village of Kavalcık). Bedouin: *iḥna*; the Bedouin in Arpahan and Paşaköy have adopted the sedentary form *niḥna*.

2.2.1.2 Possessive/Object suffixes

Most sedentary dialects have the suffixes in Table 3.

The rule given for short vowels in 2.1.1.2 accounts for variation *-in* ~ *-un*, etc. In some Alawi villages in the north and along the coast the suffix 3rd person sg. masc. after vowel is *–h*; some Christian dialects have *-hni* after vowels (*ʿaṭāhni* 'he gave him').

Table 3. Possessive/object Suffixes in sedentary dialects

	After consonants	After vowels (final vowel lengthened)
3rd sg. masc.	*-u*	-ø
3rd sg. fem.	*-a*	*(ā)-ha*, *(ī)-ya*, *(ū)-wa*
3rd pl.	*-in* ~ *-un*	*(ā)-hin* ~ *-hun*, *(ī)-yin* ~ *-yun*, *(ū)-win* ~ *-wun*
2nd sg. masc.	*-ak*	*-k*
2nd sg. fem.	*-ik*	*-ki*
2nd pl.	*-kin* ~ *-kun*	*-kin* ~ *-kun*
1st sg.	*-i* (object *-ni*)	*-y* (object *-ni*)
1st pl.	*-na*	*-na*

The Bedouin have the suffixes in Table 4:

Table 4. Possessive/object suffixes in Bedouin dialects

3rd sg. masc.	*-u*
3rd sg. fem.	*-ha*
3rd pl. m	*-hum*
3rd pl. fem.	*-hin* ~ *-han*
2nd sg. masc.	*-ak*
2nd sg. fem.	*-ič*
2nd pl. masc.	*-kum* ~ *-kam*
2nd pl. fem.	*-čin* ~ *-čan*
1st sg.	*-i* (object *-ni*)
1st pl.	*-na*

2.2.1.3 Indirect object suffixes

Indirect objects are expressed by inserting *-l-* (after -CC *-ill*) between the verb and pronominal suffixes. In the Bedouin dialects the final *n* of the feminine plural is assimilated: /čitaban/ + /lu/ > *čitaballu* 'they [fem.] wrote it'.

2.2.1.4 Demonstratives

Near deixis

sg. masc.: *hāda* ~ *hǟda* (most sedentaries), *hǟd* (Sunnis in Antakya), *hēda* (Christians in Altınözü), *hada* (Alawis in Madenli and Üçgüllük); Bedouin *hāḏa* (in Alaattin *hāḏ̣a*, in Paşaköy and Arpahan *hāda*).

sg. fem.: Alawis *hāya* (except Kavalcık *hāydi*, Iskenderun *hīya*, and Ötençay *hāyi*); Sunnis: *hayye* (except Antakya *hayy* and Babatorun *hǟdi*); Christians: *hǟdi* (Altınözü *hēdi*); Jews *hǟdi* ~ *hayye*; Bedouin *hāya*.

pl.: Alawis *hawdi* (except some villages near the coast *hawdin* and Iskenderun *hǟdōl*); Sunnis *hawwe* (except Antakya *haww* and Babatorun *hǟdu*); Christians *hǟdōl* (Altınözü *hanni*); Jews *hadōl(e)*; Bedouin masc.: *hāḏōla* (Paşaköy and Arpahan *hawdōle*), fem.: *hāḏanna* (Reyhanlı *hāḏinna*, Paşaköy and Arpahan *hawdinne*).

Far deixis

sg. masc.: most sedentary dialects *hǟka*; Jews and some rural dialects *hadǟk*; Alawis in Madenli and Üçgüllük *hada*; Christians in Altınözü *hēka*; Bedouin *hāḏāk*, Alaattin *hāḏ̣āk*, Gölbaşı *hāḏūk*; Arpahan and Paşaköy have adopted the sedentary form *hāka*.

sg. fem.: most sedentary dialects *hǟki*, Jews and some rural dialects *hadīk*; Christians in Altınözü *hēki*; Bedouin *hāḏīče*; Alaattin and

Gölbaşı *hāḏīč*; Arpahan and Paşaköy *hādīče*.

pl.: Alawis and Christians mostly *hawki* but *hǟkōl*, *hädōlīk*, and *hǟdawki* are also attested; Christians of Altınözü *hanni*; Sunnis *hawke* (except Antakya *hawk* and Babatorun *hǟku*); Jews *hadōlīk(e)*; Bedouin: masc. *hāḏōlāk* (Paşköy and Arpahan *hawdōlāk*), fem. *hāḏannīš* (Gölbaşı *hāḏannāk*, Paşaköy and Arpahan *hādannīč*).

2.2.1.5 Presentatives

For the Sunni village of Keskinci *kōha* ‘here he is’ (fem. *kēha*, pl. *kēnna*) is attested.

2.2.1.6 Relative Pronoun

With the exception of the Jews the sedentaries have *il*, which is identical to the definite article. The Jews have *əlli* and the Bedouin *hal*.

2.2.1.7 Interrogative Pronouns

All dialects have *mīn* ‘who?’, ‘what?’ is *ayš* (Christians, most of the Sunni dialects and the city dialects of the Alawites), *šū* or *šō* (most of the Alawi villages), *ašu* (Sunnis of Antakya and two other villages), *aš* (Tavla), *əš* (Jews), *iššu* (Reyhanlı), *wēš* (Üçırmak), *šikᵊl* (some Alawi villages), *šinu* or *šunu* (Bedouin). For ‘which?’ the Christians, Sunnis, and the Alawites in the cities have *ayna*, the Alawi villages have *nǟ* and the Jews *ēni*.

2.2.2 Adverbs

2.2.2.1 Temporal

Sedentary *ḥallaq* ‘now’ (city of Antakya *ḥallaḳ*, Jews and some Sunni villages *ḥallaʾ*), Bedouin *alḥaz*.

2.2.2.2 Local

‘here’: *hawn* or *hōn* (most sedentaries); *hān* ~ *hām* and *hawm* (some Christian and Sunni dialects); *hōnit* (in one Alevi village); *hēna* or *hēne* (Bedouin).

‘there’: *hawnīk*, *hūnik*, *hōnīk* (most sedentaries); *hūnāk*, *hnīk*, *hamīk* (some Sunni dialects); *hawmīk* (in a few Christian and Sunni villages); *hnāk* (Bedouin).

2.2.2.3 Manner

Most Alawites: *hīk* ‘so’; Christians, most Sunnis, and the Alawites in Iskenderun *hayk*; Jews and the Sunnis in Reyhanlı *hēk*; Bedouin *hīč*.

2.2.2.4 Interrogatives

‘where?’: *ayn* (many sedentary dialects), *aynaḥall* (Christians and Alawites), *ayna* (Christians), *aynaḥᵊl*, *aynaḥᵊn*, *waynaḥᵊn*, *naḥᵊn*, *ayni*, *ən*, *ān*, *ōn*, *wǟn*, and *fayn* (only Alawites), *wayn* and *wēn* (Alawites and Sunnis including the Bedouin).

‘when?’: *aymat*, *aymāt*, *ēmat*, *īmat*, *aymti*, *ǟmti*, *amtik*, *īmtan*, *īmtān*, *aymtin*, *amayt*, and *aynti* are attested, rare forms in only three villages are *ayšwaqt* and *aywaq*.

‘why?’: Christians, Sunnis, and the Alawi city dialects have *layš*, the Alawi villages usually *lašo*, *lašu*, *lašū*, *lašō*, *lišō*. The Jews and some other dialects have *lēš*. Rare forms are *la ayš*, *minšān ayš*, *minšō*, *liš*, *lǟš*, and *ʿalayš*. The Bedouin have *ʿalēš* and *ʿáleš*.

‘how?’: Most sedentary dialects have *kīf*, rural also *škīf*. Some Sunni dialects have *šlawn* or *šlōn*, the Jews *əšlōn*. As well as *əšlōn* the Bedouin also have *šnōn*.

‘how many?’: beside *kam* (Bedouin *čam*) the forms *kan*, *kǟn*, *kayn*, *kām*, and *kawm* are also attested.

‘how much?’: *ašqadd*, *šqadd*; Bedouin *šgadd*.

2.2.3 Particles

The article: is *il-* (Jews *əl-*) in the sedentary dialects. The Bedouin have *al-*. Besides the negations *mā* and *lā* the Alawi villages have *ǟ* (*hūwi ǟ b-il-bēt* ‘he is not at home’) and the Jews *lam* (*lam ambiṣīr* ‘it is not possible’).

2.2.4 Noun

Expressions of paired parts of the body (e.g. *ijər* ‘foot’, *īd* ‘hand’) and beings of female sex (*imm* ‘mother’) are feminine without a feminine marker. In many dialects, the dual forms of these nouns take a *t*-suffix in the construct state and in the dual (*ijərtayn* ‘two feet’) and the old dual is used as a plural (*ijrayn* ‘feet’).

Nouns denoting family members are often replaced by the diminutive forms in the dialects of the Christians and Alawis (*bayy* ‘father’, *xayy* ‘brother’, *xayt* ‘sister’), but in the construct state the old forms are sometimes preserved (*xūk* ‘your brother’).

In many dialects, old internal plurals are replaced by the external plural *-āt* (*rāsāt* ‘heads’, *qalbāt* ‘hearts’). By contrast, the internal plural of the CCv̄C type is very productive in both sedentary and Bedouin dialects in the whole area: *jbūl* 'mountains', *ḥyūṭ* 'walls', *ṣbūʿ* 'fingers'.

2.2.6 Verb

2.2.6.1 Forms of the verb

2.2.6.1.1 Form I

Most of the dialects have perfect bases with the vowel *a* or the vowel *i*: *katab* 'he wrote', *širib* 'he drank'. In the dialects of the Christians of Altınözü and the Alawis of Toygarlı the vowel of the second syllable is always /i/ as in *katib*, *širib*. The Christians of Samandağ have only one base with the vowel /i/ as in *kitib*, *širib*. The Bedouin dialects have bases with /i/ in the first syllable and /a/ in the second syllable (*čitab* 'he wrote'). If the first radical is a back consonant the radical of the first syllable is /a/ (*ḥamal* 'he carried'). In bases with /i/ in both syllables, the first vowel is elided in open syllables with a short vowel (*šrib* 'he drank', *širbat* 'she drank').

In the imperfect, the Jews and the Sunnis, but also some Christian and Alawi dialects, have three bases: *yišrab*, *yiktub*, *yilbis* (Jews: *yəšrab*, *yəktob*, *yəlbes*). Most of the Christian and Alawi dialects have only two bases: *yišrab*, *yilbis*. Verbs with an original vowel **u* shifted in some dialects to the /a/-group (*yiktab*) in others to the /i/-group (*yiktib*).

The Bedouin have imperfect bases with the vowel /a/ or /i/: *yiḏ̣rib*, *yismaʿ*. In u-coloring environments the vowel /u/ occurs: (*yunṭur* ~ *yinṭur* 'he waits'). Verbs influenced by the → *gahawa* syndrome are of the *yfaʿil* type: *yḥamil* 'he carries'.

2.2.6.1.2 Derived forms

Form II is in all dialects of the *faʿʿal/yfaʿʿil* type; only the Christians of Samandağ have *faʿʿil/yfaʿʿil*. Form V is similar except for the added *t*-prefix.

Forms III and VI are without *ʾimāla* in all dialects of the Christians and Jews and in most of the Alawi dialects too: *fāʿal/yfāʿil*. Among the Sunni dialects only the dialect of Reyhanlı has no *ʾimāla (fāʿal/ yfāʿal)*, all the other Sunni dialects and some Alawi dialects have forms with *ʾimāla* (*fīʿil/yfīʿil*), some Alewi dialects in the north even have *fēʿal/yfēʿil*.

Form IV is not attested.

Forms VII and VIII are attested with different vowel distribution: *nfaʿal/yinfaʿal* : *ftaʿal/yiftiʿil*. The Christians of Samandağ have *nfiʿil/yinfiʿil*, for Form VIII also *ftaʿil/yiftiʿil* and *ftaʿal/yiftaʿil* are attested. The Bedouin have *nfiʿal/yinfaʿil* and *ftiʿal/ yiftaʿil*.

Form IX is preserved in the sedentary dialects (*ḥmarr/yiḥmarr*); the forms of the 1st and 2nd person have a base vowel /i/ in the dialects of the Christians and Alawites: *ḥmirrayt* (Jews *ḥmarrēt*, Sunnis *ḥmarrayt*). In the dialects of the Bedouin Form IX is replaced by Form I: *ḥamar/yiḥmar*.

Form X is found only in the dialects of the Bedouin (*stafʿal/yistafʿil*) and of the Jews (*stafʿa/ ystafʿel*). In the other sedentary dialects, some originally Form X verbs have shifted to Form VIII, for example: *staraḥ/yistiriḥ* 'to rest', *stafad/yistifid* 'to profit'.

2.2.6.2 Inflection of imperfect and perfect

2.2.6.2.1 Imperfect: paradigm 'to drink' (Tables 5 and 6)

Table 5. Imperfect in sedentary dialects

	Jews	Kamberli (Sunnis) Yayladağ (Christians)	Other sedentary dialects
3rd sg. masc.	*yəšrab*	*yišrab*	*yišrab*
3rd sg. fem.	*təšrab*	*tišrab*	*tišrab*
2nd sg. masc.	*təšrab*	*tišrab*	*tišrab*
2nd sg. fem.	*təšrabi*	*tíšrabi*	*tišrábi*
1st sg. com.	*ašrab*	*išrab*	*išrab*
3rd pl. com.	*yəšrabu*	*yíšrabu*	*yišrábu*
2nd pl. com.	*təšrabu*	*tíšrabu*	*tišrábu*
1st pl. com.	*nəšrab*	*nišrab*	*nišrab*

Table 6. Imperfect in Bedouin dialects

3rd sg. masc.	*yišrab*	3rd pl. masc.	*yišrabūn*
3rd sg. fem.	*tišrab*	3rd pl. fem.	*yišraban*
2nd sg. masc.	*tišrab*	2nd pl. masc.	*tišrabūn*
2nd sg. fem.	*tišrabīn*	2nd pl. fem.	*tišraban*
1st sg. com.	*ašrab*	1st pl. com.	*nišrab*

2.2.6.2.2 Perfect: paradigm 'to write' (Table 7)

The Sunnis of Kavalcık have for the 3rd person sg. fem. *kátabit*. In the dialect of the Christians in Samandağ the base vowel is /i/ *(kitib*, *ktibt*, *kitbu)* and the inflection suffix of the 3rd person sg. fem. is *-at*: *kitbat*.

Table 7. Perfect in sedentary dialects

	Sunnis	Jews	Alawites/ Christians
3rd sg. masc.	*katab*	*katab*	*katab*
3rd sg. fem.	*katbit*	*katbet*	*katbit*
2nd sg. masc.	*katabt*	*katabt*	*ktabt*
2nd sg. fem.	*katabti*	*katabti*	*ktabti*
1st sg. com.	*katabt*	*katabt*	*ktabt*
3rd pl. com.	*kátabu*	*kátabu*	*katbu*
2nd pl. com.	*katabtu*	*katabtu*	*ktabtu*
1st pl. com.	*katabna*	*katabna*	*ktabna*

Bedouin:
Two different types of inflectional suffixes are attested. (Table 8)

Table 8. Perfect in Bedouin dialects

	Type 1	Type 2
3rd sg. masc.	*-ø*	*-ø*
3rd sg. fem.	*-at*	*-at*
2nd sg. masc.	*-t*	*-t*
2nd sg. fem.	*-ti*	*-ti*
1st sg. com.	*-t*	*-t*
3rd pl. masc.	*-um*	*-am*
3rd pl. fem.	*-an*	*-an*
2nd pl. masc.	*-tum*	*-tam*
2nd pl. fem.	*-tan*	*-tan*
1st pl. com.	*-na*	*-na*

2.2.7 Weak verbs

2.2.7.1 Geminated verbs

All dialects have verbs with the vowel /a/ or /i/ in the imperfect: *ḍall/yḍall*, *ḥabb/yḥibb*. In the dialects of the Alawites and the Christians, the base vowel is /i/ in all forms of the 1st and 2nd person of the perfect *(ḍillayt)*.

2.2.7.2 Verbs Iʾ

The Sunnis have *akal/yākul*, the Alawites and most of the Christians have *akal/yākil*. The form *akil*, *yākil* is attested in the dialect of the Alawites of Samandağ. Forms with inflectional suffixes beginning with a consonant have lost their first syllable in the dialects of the Alawites and the Christians (*kalt* 'I ate'; some Christian dialects have *kilt*). The imperative sg. masc. is *kōl* in the dialects of the Sunnis and the Jews, but *kēl* or *kīl* in the dialects of the Alawites and the Christians. Most of the Bedouin have *čila/yākul*, but other forms such as *ačal/yākul* are also attested.

2.2.7.3 Verbs I*w/y*

All sedentary dialects have forms with a long vowel in the imperfect: *wirit/yūrat*, *yibis/yības*. In some Alawi and Christian village dialects, the verbs I*y* shifted to the II*y*: *ybās/ybīs*.

Some Bedouin dialects have also forms with long vowel: *yibas/yības*, *wiṣal/yōṣal*, but forms with loss of the first radical are also attested: *ybis/yibas*, *wiṣil/yiṣil*.

2.2.7.4 Verbs II*w/y*

In all dialects of the region the forms are *nām*, *ynām*, *rāḥ/yrūḥ* and *bāˁ/ybīˁ*.

2.2.7.5 Verbs III*w/y*

In this verbal group, many different forms have developed. In some dialects all types have collapsed into one form like *nisi*, *biki, qiri* or *nisa, bika, qira*; other dialects have different forms such as *nisi, biki, qara* (many Sunni dialects), or *nisa, bika, qara* (some Alawi dialects). The forms *nisi, baka, qiri* are also attested in an Alawi dialect. In the sedentary dialect of Reyhanlı, the original verbs of the *i*-type (**nasiya*) have preserved the third radical when inflectional endings beginning with a vowel are added (*nisyit*, but *bikit* and *qarit*). In most of the Alawi and Sunni dialects, the third radical is lost *(nisit)*. In many dialects of the Christians, the /y/ of the *i*-type was transferred by analogy to the other verbs III*w/y* (*nisyit*, *bikyit*, *qiryit*). This is also true for the dialect of the Jews (*nəsyet*, *bəkyet*, *ʾəryet*).

A similar development can be observed in the Bedouin dialects. They all have 3rd person sg. masc. *nisa*, *biča*, *gira*, but the Bedouin in Reyhanlı have 3rd person sg. fem. *nisyat*, *bičyat*, *giryat* whereas the other Bedouin have *nisat*, *bičat*, *girat*.

3. LEXICON

The vocabulary of village life and traditional agriculture is almost without → Turkish loanwords, but the terms of administration and modern technology are all borrowed from Turkish, for example *emekli* 'pensioner', *süč* 'debt', *hafta* 'week'.

The different religious communities have their own terminology which is not known in the other communities. The Christians have, for example, *ˁirrāb* 'godfather' and the Alawites *naqfe* 'ritual drink'.

BIBLIOGRAPHICAL REFERENCES

Arnold, Werner 1998. *Die arabischen Dialekte Antiochiens*. Wiesbaden: O. Harrassowitz.
——. 2000. "The Arabic dialects in the Turkish province of Hatay and the Aramaic dialects in the Syrian mountains of Qalamūn: Two minority languages compared". *Arabic as a minority language*, ed. Jonathan Owens, 347–370. Berlin and New York: W. de Gruyter.
——. 2002. "Code switching and code mixing in the Arabic dialects of Antioch". *Aspects of the dialects of Arabic today: Proceedings of the 4th Conference of the International Arabic Dialectology Association (AIDA), Marrakesh, Apr. 1–4. 2000*, ed. Abderrahim Youssi, Fouzia Benjelloun, Mohamed Dahbi, and Zakia Iraqui-Sinaceur, 163–168. Rabat: AMAPATRIL.
——. 2004. "Homonymenfurcht in den arabischen Dialekten Antiochiens". *Jerusalem Studies in Arabic and Islam* 29.136–144.
Behnstedt, Peter 1997. *Sprachatlas von Syrien*. Wiesbaden: O. Harrassowitz.
Bergsträsser, Gotthelf 1915. "Sprachatlas von Syrien und Palästina". *Zeitschrift des Deutschen Palästina-Vereins* 38.169–222.
Nevo, Moshe 1991. *The Arabic dialect of the Jews of Aleppo: Phonology and morphology*. Ph.D. diss., Hebrew University of Jerusalem.
Procházka, Stephan 2002. *Die arabischen Dialekte der Çukurova (Südtürkei)*. Wiesbaden: O. Harrassowitz.
Stroomer, Harry 1991. "The Arabic dialect of Antakiya". *Semitic studies in honor of Wolf Leslau*, ed. Alan S. Kaye, II, 1515–1519. Wiesbaden: O. Harrassowitz.

WERNER ARNOLD (University of Heidelberg)

Apocopate → Mood

Apophony

The fact that Classical Arabic uses vowel changes to indicate grammatical information such as aspect, voice, and number was observed very early by the Arabic scholars of the 'Classical' tradition (10th–15th centuries); it is discussed in the so-called *taṣrīf* (→ *ṣarf*) part of Arabic grammars (Ibn Jinnī, *Munṣif*, Ibn al-Ḥājib, Ibn as-Sarrāj, Ibn ʿUṣfūr, *Mumtiʿ*, Ibn Yaʿīš, *Šarḥ al-Mulūkī*). The word 'apophony' (or *Ablaut*) was first used by Grimm (1819) to refer to context-free, morphologically meaningful vowel alternations in some verbal paradigms of German and its ancestors, back to Indo-European. Vowel alternations in Classical Arabic are discussed in standard studies such as Barth (1889), Brockelmann (1908–1913), Nöldeke (1897), Philippi (1894), and Wright (1859), but the first comprehensive theory of apophony in Semitic is Kuryłowicz's (1957–1958, 1961, 1972). Kuryłowicz set out the questions that still form the basis of current research (Chekayri and Scheer 1996, forthcoming; Guerssel 2003; Guerssel and Lowenstamm 1990, 1996; McCarthy 1990; Ségéral 1995, 1996, 2000).

1. FACTS

The apophonic alternations discussed in the literature belong mainly to two domains: the expression of voice and aspect in the verbal system, and plural formation in the nominal system. The first ten Forms of a Classical Arabic sound triliteral root are given in Table 1.

The verb Forms in Table 1 can be divided into two groups, depending on the stem vocalization. The first group consists of Active Forms I, the second group of Passive Forms I, and all Forms II–X. In the first group, the thematic vowel (i.e. the vocalization of C_2) is lexically conditioned: it may be *a, i* or *u*, depending on the root. In the second group, the vocalization is the same for all roots, e.g. *kutib*, *ḍurib*, *lubis*.

For both groups, Perfect and Imperfect forms differ in two respects: their stem vocalization, and the fact that Imperfect forms are prefixed, whereas Perfect forms are not. The first point is of interest to us here, and the different patterns are summarized in Table 2.

As Table 2 shows, voice opposition is also expressed by vowel alternation.

In nominal morphology, singular and plural in some classes of internal or so-called → broken plurals differ only with respect to their vocalic melody: the number opposition is expressed by vowel alternation. Some examples are given in Table 3.

Plural formation in Classical Arabic is particularly intricate since the shape of the plural depends on many factors, such as gender, the syllabic structure of the singular, and the semantics of the root.

Moreover, for a given singular type, different plural types may be attested, as shown by the last example in Table 3. A discussion of these alternations is beyond the scope of this entry (see Barth 1889; Brame 1970; Brockelmann 1908; Fleisch 1961; Idrissi 1997; Kuryłowicz 1961; McCarthy 1979, 1983, 1990; Murtonen 1964; Petráček 1960; Philippi 1894; Ratcliffe 1998; Wright 1859, for analysis and references).

In the *taṣrīf* part of traditional Arabic grammars, definitions of the basic structures of the

Table 1. The verbal system

	Perfect		Imperfect		Gloss
	Active	Passive	Active	Passive	
I	*kata̱b*	*kutib*	*ya-ktu̱b*	*yu-ktab*	‘to write’
	ḍara̱b	*ḍurib*	*ya-ḍri̱b*	*yu-ḍrab*	‘to hit’
	labi̱s	*lubis*	*ya-lba̱s*	*yu-lbas*	‘to wear’
	kabu̱r	–	*ya-kbu̱r*	–	‘to be tall’
II	*kattab*	*kuttib*	*yu-kattib*	*yu-kattab*	‘to cause to write’
III	*ka:tab*	*ku:tib*	*yu-ka:tib*	*yu-ka:tab*	‘to correspond’
IV	*ʾaktab*	*ʾuktib*	*yu-ʾaktib*	*yu-ʾaktab*	‘to cause to write’
V	*takattab*	*tukuttib*	*ya-takattab*	*yu-takattab*	‘to be caused to write’
VI	*taka:tab*	*tuku:tib*	*ya-taka:tab*	*yu-taka:tab*	‘to write to each other’
VII	*nkatab*	*nkutib*	*ya-nkatib*	*yu-nkatab*	‘to subscribe’
VIII	*ktatab*	*ktutib*	*ya-ktatib*	*yu-ktatab*	‘to write, be registered’
IX	*ktabab*	–	*ya-ktabib*	–	
X	*staktab*	*stuktib*	*ya-staktib*	*yu-staktab*	‘to write, make write’

Table 2. Apophony in the verbal system

		Perfect	Imperfect
a. Form I Active (thematic vowel):		*a*	*u*
		a	*i*
		i	*a*
		u	*u*
b. Other Forms (except V, VI):	Active:	*a…a*	*a…i*
	Passive:	*u…i*	*(a)…a*
c. Forms V, VI:	Active	*a…a…a*	*a…a…a*
	Passive:	*u…u…i*	

Table 3. Apophony in the nominal system

Singular	Plural	Gloss
ki̱ta̱:b	*ku̱tu̱b*	‘book’
qa̱ṭi̱:ʿ	*qi̱ṭa̱:ʿ*	‘group’
qi̱ndi̱:l	*qa̱na̱:di̱l*	‘lamp’
bu̱ndu̱q	*ba̱na̱:di̱q*	‘hazelnut’
ʾi̱ta̱:n	*ʾu̱tn*, *ʾu̱tu̱n*, *ʾa̱:tu̱n*, *maʾtu̱:na̱:ʾ*	‘she-donkey’

verbal system are given, as well as a description of the morphological processes yielding the derived forms. Two types of processes apply to a basic form: processes without augment, and processes with augment. Augmentation is understood as a modification of the consonantal structure of the base. Apophony is a process without augment.

2. Apophony theory

The discussion of apophony focuses on the relationship between the vowels involved in the alternation. Do both terms of an apophonic alternation have to be lexicalized? Or is there a basic vocalization and a derived one? If yes, can the derived vocalization be predicted from the basic one? The strongest hypothesis is that apophonic alternations are predictable without ambiguity from a basic vocalization.

The first comprehensive theory of apophony in Semitic is by Kuryłowicz (1957–1958, 1961, 1972). The basic principles of this theory are very roughly summarized below.

i. Origin of apophony:
“Historically the various kinds of apophony go as a rule back to purely phonemic alternations” (Kuryłowicz 1972:32). These alternations spread by analogy to forms where they were not conditioned by the context, and then acquired a morphological function.

ii. Apophony may have three values:
It may be “(1) part of a discontinuous morpheme [. . .] (2) full morph(on)eme [. . .] (3) morphoneme with semantic zero value” (Kuryłowicz 1972:36).

iii. Apophony is oriented, that is, there is a basic form and a derived form. The derived form is called 'founded form': "Tout comme n'importe quel autre morphème de dérivation ou de flexion, le degré vocalique doit être référé à une forme de fondation, que ce soit le mot-base (lorsqu'il s'agit de dérivation) ou une forme flexionnelle" (Kuryłowicz 1957–1958:3).

iv. An apophonic relation obeys two principles:
a. Proportionality: for each pair of forms related by apophony, the semantic or grammatical relation between the two forms is constant.
b. "Law of polarization, i.e. [the] distance between the basic and the founded form [is maximal]" (Kuryłowicz 1972:40)

v. "Reverse relations […] are also liable to take place" (Kuryłowicz 1972:35). Thus, for instance, both *i* > *a* and *a* > *i* can co-exist in a given apophonic system.

vi. There are four types of apophonic alternations in the Semitic verbal system (Kuryłowicz 1972:35): (a) vowel/zero alternation; (b) *i*, *u* > *a*; (c) *u* > *i*; (d) short/long vowel.

The principle under (iii) defines the relation between two terms involved in an apophonic alternation as a derivation. The grammatical relation between the basic and the derived form is constrained by the principle of proportionality. Vowel quality is constrained by the law of polarization. However, this law does not allow one to define the vowels involved in an apophonic relation unambiguously: *u* and *i*, for instance, are equally distant from *a*, so both alternations *u~a* and *i~a* are equally possible. Finally, length alternations are classified as apophonic too (vi).

Applying the theory of Autosegmental Phonology to Semitic morphology, McCarthy (1979, 1981) gives an account of the distribution of consonantal patterns in the Classical Arabic verbal system. Some configurations are predicted to be absent from the system because they violate basic principles of the theory. However, this approach does not predict the quality of vocalic melodies. Most linguists consider apophonic alternations only partially predictable, and analyze the vocalic melodies as part of the lexical entry of the verb.

Guerssel and Lowenstamm (1990, 1996) propose a formula, the 'Apophonic Path', which unambiguously predicts the identity of the vowels involved in an apophonic alternation. This analysis of the Classical Arabic verbal system reduces the role of the lexicon in the vocalization of the verbal root. Guerssel and Lowenstamm (1996) focus on the alternations of the thematic vowel in Form I active of sound triliterals. This analysis is sketched below. The reader is referred to Guerssel and Lowenstamm (1990) for other Forms.

The relevant data are repeated in (1).

(1)	Perfect	Imperfect	
1.	*labi̲s*	*ya-lba̲s*	'to wear'
2.	*ḍara̲b*	*ya-ḍri̲b*	'to hit'
3.	*kata̲b*	*ya-ktu̲b*	'to write'
4.	*kabu̲r*	*ya-kbu̲r*	'to be tall'

Consider this system of alternations as summarized in (2).

(2)				
Perfect	*i*	*a*	*a*	*u*
	| 1	| 2	| 3	| 4
Imperfect	*a*	*i*	*u*	*u*

The hypothesis that apophony is a derivation seems to be weakened by two properties of (2), which obtain whether one assumes the directionality to be Perfect ⇒ Imperfect, or Imperfect ⇒ Perfect:

i. Opacity
One input has two different outputs: Perfect *a* ⇒ Imperfect *i, u*, or if the other directionality is chosen: Imperfect *u* ⇒ Perfect *a*, *u*. The derivation is ambiguous.

ii. Polarity
Both derivations *i* ⇒ *a* and *a* ⇒ *i* have to be postulated. Guerssel and Lowenstamm (1996:129) adopt a position which clearly differs from that of Kuryłowicz: "Since derivational operations are directional, it is difficult to believe that if X ⇒ Y is the manifestation of a process unidirectionally relating A and B, X Y obtains regardless of whether A ⇒ B or B ⇒ A!". On this question, see also Brame (1970).

In order to reduce the opacity, Guerssel and Lowenstamm apply a method established in phonology: since *a* in *ḍara̲b* behaves systematically differently from *a* in *kata̲b*, these segments must have two different phonological identities, which they note *a* and *x*. The same reasoning applies to *u* if the directionality Imperfect ⇒ Perfect is chosen. Guerssel and

Lowenstamm (1996:130) argue that a single non-opaque and non-polar system underlies the alternations in (2), as in (3):

$$
\text{(3) Perfect} \quad
\begin{array}{lcccc}
 & i & x & a & u \\
 & | & | & | & | \\
 & 1 & 2 & 3 & 4 \\
 & \downarrow & \downarrow & \downarrow & \downarrow \\
\text{Imperfect} & a & i & u & u
\end{array}
$$

They analyze *x* as an underlying empty vowel position: the vocalization of *ḍarab* is *a*...*Ø*. The empty position between C_2 and C_3 is filled phonologically by propagation from the first vowel, yielding *ḍara̱b*. Under this hypothesis, the derivation of the Imperfect can have one of the following four forms:

(4) Perfect ⇒ Imperfect Verb classes

Perfect	⇒ Imperfect	Verb classes		
Ø	⇒ *i*	1. *ḍarØ̱b*	⇒	*ya-ḍri̱b*
i	⇒ *a*	2. *labi̱s*	⇒	*ya-lba̱s*
a	⇒ *u*	3. *kata̱b*	⇒	*ya-ktu̱b*
u	⇒ *u*	4. *kabu̱r*	⇒	*ya-kbu̱r*

They propose that the system does not involve four independent derivations, but four steps of a fixed sequence of derivations, the so-called 'Apophonic Path' given in (5).

(5) Ø ⇒ i ⇒ a ⇒ u ⇒ u

Given (5), apophonic alternations are unambiguous derivations. In particular, the hypothesis of a fourth element, Ø, yields a system without polarity. Note also that their analysis excludes short vs. long vowel alternations: the Apophonic Path affects the melodic level only. Length alternations are conditioned by another level of representation, the skeletal level.

In establishing hierarchical relations between vowels, the Apophonic Path recalls a principle of the Classical Arabic tradition called *xiffa* 'lightness', which is given in the phonology section of traditional grammars, the second part of the *taṣrīf*. According to this principle, phonological rules are motivated by the avoidance of heavy sequences, where *a* is lighter than *i*, which is lighter than *u*.

The analysis of apophony as a principled phenomenon concentrates on the regularity of the alternations. Kuryłowicz defines a derivational link between two forms involved in an apophonic alternation (law of polarization, proportionality). Guerssel and Lowenstamm propose a restrictive analysis of apophony: apophonic alternations are unambiguous sequences of four derivations. This theory provides insight in other parts of Classical Arabic morphology (see Chekayri and Scheer 1996, [forthcoming] for an analysis of so-called → 'weak' verbs). Subsequent work has shown that the sequence of derivations postulated by the Apophonic Path is not specific to Classical Arabic (see for Akkadian and Geʿez: Ségéral 1995, 1996, 2000; Kabyle Berber: Bendjaballah 2001; German: Ségéral and Scheer 1998; Spanish: Boyé 2000).

Bibliographical references

Primary sources

Ibn al-Ḥājib = Jamāl ad-Dīn ʾAbū ʿAmr ʿUṯmān ibn ʿUmar Ibn al-Ḥājib, *aš-Šāfiya fī ʿilm at-taṣrīf*. Ed. Hasan ʾAḥmad al-ʿUṯmān. Mekka: al-Maktaba al-Makkiyya, 1995.

Ibn Jinnī, *Munṣif* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *Al-Munṣif šarḥ li-kitāb at-taṣrīf li-ʾAbī ʿUṯmān al-Māzinī*. Ed. ʾIbrāhīm Muṣṭafā and ʿAbdallāh ʾAmīn. Cairo, 1954.

Ibn ʿUṣfūr, *Mumtiʿ* = ʾAbū l-Ḥasan ʿAlī ibn Muʾmin Ibn ʿUṣfūr, *al-Mumtiʿ fī t-taṣrīf*. Ed. Faxr ad-Dīn Qabāwa. Beirut: Dār al-ʾĀfāq al-Jadīda, 1970.

Ibn Yaʿīš, *Šarḥ al-Mulūkī* = Muwaffaq ad-Dīn Yaʿīš ibn ʿAlī Ibn Yaʿīš, *Šarḥ al-Mulūkī fī t-taṣrīf*. Ed. Faxr ad-Dīn Qabāwa. Aleppo: al-Maktaba al-ʿArabiyya, 1973.

Secondary sources

Aro, Jussi. 1964. *Die Vokalisierung des Grundstammes im semitischen Verbum*. Helsinki: Societas Orientalis Fennica.

Barth, Jacob. 1889. *Die Nominalbildung in den semitischen Sprachen*. Leipzig: Hinrichs.

Bendjaballah, Sabrina. 2001. "The negative preterite in Kabyle Berber". *Folia Linguistica* 34.185–223.

Boyé, Gilles. 2000. *Problèmes de morpho-phonologie verbale en français, en espagnol et en italien*. Ph.D. diss., University of Paris 7.

Brame, Michael. 1970. *Arabic phonology: Implications for phonological theory and historical Semitic*. Ph.D. diss., Massachusetts Institute of Technology.

Brockelmann, Carl. 1908–1913. *Grundriss der vergleichenden Grammatik der semitischen Sprachen*. Berlin: Reuther und Reichard. (Repr. Hildesheim: G. Olms, 1961.)

Chekayri, Abdellah and Tobias Scheer. 1996. "The apophonic origin of glides in the verbal system of Classical Arabic". Lecarme a.o. 1996: 62–76.

——. Forthcoming. "The appearance of glides in Classical Arabic defective verbs". *Folia Orientalia*. <http://www.unice.fr/dsl/tobweb/papers/ChekScheerAppGlides.pdf>.

Fleisch, Henri. 1961. *Traité de philologie arabe*. Beirut: Imprimerie Catholique.

——. 1968. *L'arabe classique: Esquisse d'une structure linguistique*. Beirut: Imprimerie Catholique.

Greenberg, Joseph H. 1950. "The patterning of root morphemes in Semitic". *Word* 6.162–181.

Grimm, Jacob. 1819. *Deutsche Grammatik*. I. *Formenlehre*. Göttingen: Dieterich. 2nd ed. 1822–1837. (Repr. Hildesheim: G. Olms, 1967.)

Guerssel, Mohamed. 2003. "Why Arabic guttural assimilation is not a phonological process". *Living on the edge: 28 papers in honour of Jonathan Kaye*, ed. Stefan Ploch, 581–598. Berlin and New York: Mouton de Gruyter.

—— and Jean Lowenstamm. 1990. *Classical Arabic apophony*. Ms. University of Quebec at Montreal and University of Paris 7.

——. 1996. "Ablaut in Classical Arabic measure I active verbal forms". Lecarme a.o. 1996:123–134.

Idrissi, Ali. 1997. "Plural formation in Arabic". *Perspectives on Arabic linguistics*, X, ed. Mushira Eid and Robert Ratcliffe, 123–145. Amsterdam: J. Benjamins.

Kuryłowicz, Jerzy. 1957–1958. "Esquisse d'une théorie de l'apophonie en sémitique". *Bulletin de la Société Linguistique de Paris* 53.1–38.

——. 1961. *L'apophonie en sémitique*. Wrocław and Warsaw: Polska Akademia Nauk. Komitet Językoznawstwa.

——. 1972. *Studies in Semitic grammar and metrics*. Wrocław and Warsaw: Polska Akademia Nauk. Komitet Językoznawstwa.

Larcher, Pierre. 2003. *Le système verbal de l'arabe classique*. Aix-en-Provence: Publications de l'Université de Provence.

Lecarme, Jacqueline, Jean Lowenstamm and Ur Shlonsky (eds.). 1996. *Studies in Afroasiatic grammar*. The Hague: Holland Academic Graphics.

McCarthy, John. 1979. *Formal problems in Semitic phonology and morphology*. Ph.D. diss., Massachusetts Institute of Technology.

——. 1981. "A prosodic theory of nonconcatenative morphology". *Linguistic Inquiry* 12.373–418.

——. 1983. "A prosodic account of Arabic broken plurals". *Current trends in African linguistics*, I, ed. Ivan R. Dihoff, 289–320. Dordrecht: Foris.

—— and Alan Prince. 1990. "Foot and word in pro-sodic morphology: The Arabic broken plural". *Natural Language and Linguistic Theory* 8.109–183.

Murtonen, A. 1964. *Broken plurals: Origin and development of the system*. Leiden: E.J. Brill.

Nöldeke, Theodor. 1897. "Zur Grammatik des klassischen Arabisch". *Denkschriften der kaiserlichen Akademie der Wissenschaften, Philosophisch-historische Classe*, Band 45. Wien: Carl Gerolds Sohn.

Petráček, Karel. 1960, 1961, 1962. "Die innere Flexion in den semitischen Sprachen". *Archív Orientální* 28.547–606, 29.513–545, 30.361–408.

Philippi, Friedrich. 1894. "Die semitische Verbal- und Nominalbildung in ihrem Verhältnis zu einander". *Beiträge zur Assyriologie und Vergleichenden Semitischen Sprachwissenschaft* 2.359–389.

Ratcliffe, Robert R. 1998. *The 'broken' plural problem in Arabic and comparative Semitic*. Amsterdam and Philadelphia: J. Benjamins.

Schramm, Gene M. 1962. "An outline of Classical Arabic verb structure". *Language* 38.361–375.

——. 1970. "The Semitic languages: An overview". *Current trends in linguistics*, VI, ed. Thomas A. Sebeok a.o., 257–260. The Hague: Mouton.

Ségéral, Philippe. 1995. *Une théorie généralisée de l'apophonie*. Ph.D. diss., University of Paris 7.

——. 1996. "L'apophonie en ge'ez". Lecarme a.o. 1996:360–391.

——. 2000. "Théorie de l'apophonie et organisation des schèmes en sémitique". *Research in Afroasiatic grammar*, ed. Jacqueline Lecarme, Jean Lowenstamm, and Ur Shlonsky, 263–299. Amsterdam and Philadelphia: J. Benjamins.

—— and Tobias Scheer. 1996. "A generalized theory of Ablaut: The case of Modern German strong verbs". *Models of inflection*, ed. Albert Ortmann, Ray Fabri, and Teresa Parodi, 28–59. Tübingen: Niemeyer.

Silvestre de Sacy, Antoine-Isaac. 1831. *Grammaire arabe à l'usage des élèves de l'Ecole spéciale des langues orientales vivantes*. 2nd ed. Paris: Imprimerie Royale. (Repr. Paris: Institut du Monde Arabe.)

Wright, William. 1859–1862. *A grammar of the Arabic language*. Cambridge: Cambridge University Press. (3rd ed. 1896–1898.)

Zaborski, Andrzej. 1995. "Kuryłowicz and the so-called 'aspect' in Classical and Modern Arabic". *Kuryłowicz Memorial Volume*, I (= *Analecta Indoeuropea Craco-viensia*, 2.), ed. W. Smoczynski, 529–541. Cracow: Universitas.

Sabrina Bendjaballah
(Centre National de la Recherche Scientifique)

Apposition

The term *badal* 'apposition, substitution, permutative, parenthetic' is found in the works of early medieval Arabic grammarians and continues to be used to designate the appositive noun phrase as well as those structures in which typically a pair of noun phrases – the first *mubdal minhu* 'the principal noun'; the second, *badal* 'the apposed noun' – are juxtaposed (asyndetically) without an intervening overt device connecting them. Typically, these two noun phrases fulfill similar syntactic functions in any given structure and are governed equally by the same external element. The semantics of *badal* with regard to *mubdal minhu* ranges from designating an attribute to partial or complete coreferentiality with the principal noun phrase. As such, this definition brings Arabic appositional structures in line with English appositions (Matthews 1981:224–236; Burton-Roberts 1975:391–419), though the latter in many grammar treatises (e.g. Celce-Murcia and Larsen-Freeman 1983) are often found to be nearly synonymous with non-restrictive relative clauses. Nevertheless, the grammatical structures of appositions extend well beyond these confines and include a host of diverse structures in Arabic. Although the concept is as old as the codification of Arabic grammar (for the early treatment see Sībawayhi, *Kitāb* I, 150–158, 439), its continual refinement is demonstrated in the writings of later grammarians with recycled examples from their forerunners, for example Zajjājī, an Arabic grammarian of the 10th century (see the commentaries of Ibn ʿUṣfūr on Zajjājī's *Jumal* I, 250–267).

The three major types of *badal* constructions classified in terms of semantic relations that hold between the noun phrase and its apposed one include:

i. *badal kull min kull* 'full substitution', e.g. *marartu bi-ʾaxīka zaydin* 'I passed your brother, Zayd', where Zayd (Ibn ʿAqīl, *Šarḥ* II, 249), a proper name, is apposed to *ʾaxīka* 'your brother'. In this instance, Zayd fully substitutes for *ʾaxīka* as its semantic equivalent; when the latter is dropped, as in *marartu bi-zaydin* 'I passed Zayd', referentiality is not compromised.

ii. *badal baʿḍ min kull* 'permutative apposition' e.g. *ʾakaltu r-raġīfa ṯuluṯahu* 'I ate the loaf, one third of it' = 'I ate one third of the loaf' (Ibn Hišām, *Qaṭr* 346), where *ṯuluṯahu* 'one third', showing a partial feature of coreference, stands in apposition to the whole, *r-raġīfa* 'the loaf'. Structures of this type have an underlying genitive construction (*ʾiḍāfa* construct) whose first term is a quantifier *ʾakaltu ṯuluṯa r-raġīfi*. This type requires a pronoun suffix *-hu* to carry the referential function to the principal noun; without such a suffix, the structure becomes ungrammatical.

iii. *badal ištimāl* 'substitution of inclusiveness', e.g. *ʾaʿjabanī zaydun ʿilmuhu* 'Zayd, his learning, pleased me' (Ibn Hišām, *Qaṭr* 346), where one of Zayd's included attributes, his learning, substitutes for him. Like the type in (ii) it too requires a pronoun suffix for reference to the preceding noun to which it is apposed.

A coreference hierarchy for the noun phrases in apposition is conceivable, as in Figure 1:

Figure 1: Scale of coreference between the principal noun phrase and its apposed one

badal kull → *badal baʿḍ* → *badal ištimāl*

As Figure 1 shows, the highest level of equivalency to the principal noun is found in *badal kull*. The rightward move shows decline in co-reference since *badal baʿḍ* refers to a portion of the same (principal) noun, whereas *badal ištimāl* refers to a quality that the principal noun possesses.

Three other types of *badal* of lesser frequency and of minor importance, all of which carry a corrective or rectifying interpretation, parallel these. Contrary to the three types above, the following ones seem to violate the coreferentiality clause for apposition structures since the second noun phrase usually nullifies the preceding noun. The following three types are found in the works of medieval Arabic grammarians:

iv. *badal al-ʾiḍrāb* 'permutative of recanting', e.g. *ʾakaltu xubzan laḥman* 'I ate bread, meat' = 'I ate bread, well, meat' (Ibn Mālik, *Šarḥ* II, 249), where after stating the eating of bread, the speaker reflects that it is better to state that it was meat that was eaten, as an amendment to the preceding noun phrase.

v. *badal al-ġalat* 'permutative of error', e.g. *raʾaytu rajulan ḥimāran* 'I saw a man, a donkey' = 'I saw a man, rather, a donkey' (Ibn Mālik, *Šarḥ* II, 249), where the speaker self-corrects what she or he has stated. Like the preceding permutative, this class has an abrogating effect.

vi. *badal an-nisyān* 'permutative of forgetfulness'. To illustrate, *marartu bi-zaydin ḥimārin* 'I passed by Zayd, donkey' = 'I passed by Zayd, [oh! it was] a donkey', where the speaker intended to recall having passed a donkey but instead has forgotten this fact. In turn, the speaker misstated that it was Zayd whom he or she had passed.

Arabic grammarians included *badal* within a subclass of *tawābiʿ* 'modifiers', which include *ʿaṭf* 'coordination', *naʿt* 'adjectival', and *tawkīd* 'corroboration'. The syntactic and semantic boundaries among these subclasses were not always drawn sharply (see Owens 1990:58–63 for details, 74 for summary), leading to instances where syntactic ambiguity becomes inevitable, as in *ḍarabtu ʾabā ʿabdillāhi zaydan* ‘I hit ʾAbū ʿAbdallāh, Zayd’, which is bound to be construed either as *ʿaṭf bayān* ‘explicative coordinating’ or as *badal* (Ibn ʿAqīl, *Šarḥ* II, 221).

Appositional structures and their semantic denotations are by and large heterogeneous. They not only include declarative clauses of the types already presented but also interrogatives like *kayfa zaydun, ʾa-marīḍun ʾam ṣaḥīḥun?* How is Zayd, is he ill or well?’ (Ibn Mālik, *Šarḥ* III, 199).

Although *badal* has received extensive attention in its treatment as a syntactic and semantic phenomenon warranting a special section in grammatical treatises in medieval times, research on the subject with such intensity has not yet been paralleled in modern times. In this regard, similar to many other languages, apposition in Modern Standard Arabic and spoken dialects is less than adequately investigated and remains on the margin of grammatical investigation.

In Modern Standard Arabic, *badal* continues to include usages that have not been attested in the classical period (at least these were not included under such a rubric). Ibn ʿAqīl (*Šarḥ* II, 209) includes *jāʾa r-rakbu kulluhu* ‘the cavalcade has arrived, all of it’ = ‘the cavalcade has arrived in its entirety’ as a case of what is commonly known in Classical Arabic as the ‘corroborative’, where the second term emphasizes its preceding one. In this example, the quantifier *kull* ‘each, every, all, totality’, performing the corroborating function, has a pronoun suffix attached to it that refers back to the preceding noun *ar-rakbu* and copies its case marking. In Modern Standard Arabic there are similar usages of the quantifier *kull*, as in the example cited in Cantarino (1975:II, 73), *wa-qad jamaʿtu l-qurʾāna kullahu ḥifḏan* ‘I had memorized the whole *Qurʾān*’, in which he designates *kullahu* as having appositive function to its preceding noun *al-qurʾān*.

The frequency in usage of *badal* in Modern Standard Arabic manifests itself in the type that precludes interruptions by other elements to the two noun phrases in sequence, namely what Bloomfield (cited in Matthews 1981:227) labels ‘close apposition’. Its widest occurrences are in certain honorific titles followed by proper nouns *as-sulṭān ḥasan* ‘Sultan Ḥasan’. The sequence may not be broken or interrupted **as-sulṭān sumuww ḥasan* ‘the Sultan, Royal Highness, Hasan’; each noun phrase of this type is fully capable of standing alone without loss of referentiality or grammaticality.

Testing for appositive relation has led researchers (e.g. Burton-Roberts 1975) to conceive of criteria by which syntactic relations are to be judged. Two tests include ‘isolation’ of constituents as a token of the two noun phrases’ syntactic independence of each other and ‘reversal’ of their order, which tests their syntactic relations vis-à-vis relations specified as subordination, complementation, or modification. Burton-Roberts also observed that an additional test, insertion of an overt appositive marker, such as ‘that’s to say’ between the two noun phrases in apposition, would contribute to the sufficiency of these two tests in screening for appositive relations, as in “An upholsterer, that’s to say, Mr. Pontefract, called to-day” (Burton-Roberts 1975: 414). Such a marker existed in Classical Arabic and it holds well in some Modern Standard Arabic non-restricted appositives *ʾanna l-ʾuslūba huwa ṯamatu l-ʿaqli wa-l-qalbi ʾayi l-ʾafkāri wa-l-ʿawāṭifi* ‘that style is the fruit of intelligence and heart, that is to say, of thought and feelings’ (Cantarino 1975:II, 70).

When the task is to define appositive relations in anything like rigorous terms by way of other syntactic relations, Matthews concludes that it is “undifferentiated relation” and that the boundaries between appositives, coordination, complementation, and parataxis may not be drawn aptly.

Bibliographical references

Primary sources

Ibn ʿAqīl, *Šarḥ* = Bahāʾ ad-Dīn ʿAbdallāh ibn ʿAqīl, *Šarḥ al-ʾAlfiyya*. Ed. Muḥammad Muḥyī d-Dīn ʿAbd al-Ḥamīd. Beirut, n.d.

Ibn Hišām, *Qaṭr* = Jamāl ad-Dīn ʾAbū Muḥammad ʿAbdallāh ibn Yūsuf Ibn Hišām, *Qaṭr an-nadā wa-ball aṣ-ṣadā*. Ed. Muḥammad Muḥyī d-Dīn ʿAbd al-Ḥamīd. Beirut: al-Maktaba al-ʿAṣriyya, 2000.

Ibn Mālik, *Šarḥ* = Jamāl ad-Dīn Muḥammad ibn ʿAbdallāh ibn ʿAbdallāh Ibn Mālik, *Šarḥ at-tashīl*. Ed. Muḥammad ʿAbd al-Qādir ʿAṭāʾ and Ṭāriq Faṭhī as-Sayyid. Beirut: Dār al-Kutub al-ʿIlmiyyah, 2001.

Ibn ʿUṣfūr, *Šarḥ* = ʾAbū l-Ḥasan ʿAlī ibn Muʾmin ibn Muḥammad ibn ʿAlī Ibn ʿUṣfūr, Š*arḥ Jumal az-Zajjājī*. Ed. ʾImīl Badīʿ Yaʿqūb. Beirut: Dār al-Kutub al-ʿIlmiyya, 1998.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān ibn Qanbar Sībawayhi, *Kitāb Sībawayhi*. Ed. ʿAbd as-Salām Muḥammad Hārūn. Beirut: Dār al-Kutub al-ʿIlmiyya, 1988.

Secondary sources
Burton-Roberts, Noel. 1975. “Nominal apposition”. *Foundations of Language* 13.391–419.
Cantarino, Vicente. 1975. *Syntax of modern Arabic prose: The expanded sentence*. Bloomington: Indiana University Press.
Celce-Murcia, Daniel and Diane Larsen-Freeman. 1983. *The grammar book: An ESL/EFL teacher's course*. Boston: Heinle and Heinle.
Matthews, Peter. 1981. *Syntax*. London and New York: Cambridge University Press.
Owens, Jonathan. 1990. *Early Arabic grammatical theory: Heterogeneity and standardization*. Amsterdam and Philadelphia: J. Benjamins.

Mohssen Esseesy (George Washington University)

ʾAqsām al-kalām → Parts of Speech

ʿArab

The term Arab as a designation for groups of people in Arabia and adjacent areas is documented continuously from the middle of the 9th century B.C.E. until the present day. Its meaning has changed substantially during this period. Its original etymology is not clear but a suggestion will be presented below. This entry is based on an investigation of the documentation of Arabs in all pre-Islamic sources as well as the most important early Islamic ones, documented in Retsö (2003). The textual evidence, an analysis of it, as well as references to sources and secondary literature are found in that work.

1. Pre-Islamic period

We possess rich contemporary source material, dating from the 1,400 years between the 9th century B.C.E. and the rise of Islam, mentioning Arabs (almost 4,000 instances of the word) in Akkadian, Hebrew, Persian, Aramaic/Syriac, Greek, Latin, and South Arabian sources. Only a couple of them are certain to originate directly from the Arabs themselves. A group called *ar-ba-a-a* (= *arbāya*) appears for the first time among the kings of Syria in the monolith inscription of the Assyrian King Salmanasar III (858–824 B.C.E.), describing the battle of Qarqar in 853 B.C.E. In Assyrian inscriptions from the time of Tiglath Pileser III (744–727 B.C.E.) and onwards, people called *arab* are mentioned as dwellers in the Syrian desert in and around the oasis of Dūmah. They were the target of at least two major Assyrian attacks, under Sennacherib and Assurbanipal (Retsö 2003:124–193). The *arab* and the *arbāya* further occur in the inscriptions of the Achaemenid kings (Darius I, Xerxes I, Artaxerxes I) and are identified as *arábioi* in Herodotus (Retsö 2003:235–250). The *arab* in the Syrian desert also appear several times in the Old Testament (e.g. Jeremiah 25:18–26; Retsö 2003:212–228). Herodotus' *arábioi* lived in the area between the eastern Nile Delta, southern Palestine, and the Sinai Peninsula (*Historiae* 2.75, 3.5). One of their deities was *Alilat*, in which we perhaps have the earliest documentation of an Arabic dialect with the definite article *(a)l* (*Historiae* 1.198, 3.8). The Arabs in this area had probably been settled there already by the Assyrians (Retsö 2003:243–250).

From the time of Darius I (522–486 B.C.E.) we discern a political entity between Syria and Egypt, ruled by the Arabs described by Herodotus, which lasted until 106 C.E., when it was incorporated into the Roman provincial system as Provincia Arabia. One of its rulers was Gešem/Gašmū ha-ʿarbī ‘the Arab’, mentioned in Nehemiah (2:19, 4:7, 6:1–2), who lived in the second half of the 5th century B.C.E. (Retsö 2003:250–251). This entity, which during its later history was ruled by the kings of NBṬ, the Nabataeans, residing in present-day Petra, obviously played a central role, functioning as a cultural and economic link between the Fertile Crescent, the Mediterranean world, and the Arabian Peninsula. The language and script they used was Aramaic, but a substantial part of the population spoke dialects classifiable as Arabic. Most of the pre-Islamic documentation of a language called Arabic refers to this area (Retsö 2002:141).

Other areas where we hear of groups called Arabs in pre-Islamic times are Ḥawrān, the Biqāʿ valley, central and northern Syria (around ar-Ristan, around Qinnasrin), Central Mesopotamia, and South Arabia (Retsö 2003:308–358, 440–448, 536–567). In the latter two, Arabs played a crucial role after the turn of the era. In Mesopotamia they were an important factor in the state ruled from Hatra from the 1st to the mid-3rd centuries C.E., serving as a buffer between Parthia and Rome. Arabs are mentioned in Osrhoene and Adiabene as well, probably originating from Hatra. In South Arabia, Arabs are mentioned approximately 40 times in Sabaean and Qatabanian inscriptions from the period between the 1st and 6th centuries C.E. (Retsö 2003:536–566). An isolated early mention (ʿRBM) may, however, be found in the inscription RÉS 3945 by the Sabaean ruler Karibʾil Watar, datable to the 7th or 6th century B.C.E. (Retsö 2003:537–539).

The pre-Islamic evidence does not support the idea that Arab was a general designation for people living in the peninsula. Instead, they appear as groups in the above-mentioned areas living among other peoples. When sometimes used in Greek and Latin texts as a general term, it is derived from the name given to the entire peninsula, Arabia, a designation going back to the Greek discoveries in connection with the campaign of Alexander the Great. The Greek geographers from Eratosthenes onwards named the newly discovered continent Arabia, and consequently its inhabitants were sometimes called Arabs, *árabes*, or *arábioi*. Furthermore, the evidence does not support the idea found in most modern textbooks that Arab at this time designated nomads, Bedouin, or desert-dwellers in general. Apart from the fact that these terms do not necessarily mean the same thing, it has been pointed out that the classical Bedouin culture arose after the turn of the era, especially around the 3rd century C.E., so that the word Arab cannot refer to Bedouin of this kind since it occurs as early as the 9th century B.C.E. (Caskel 1953a, 1953b; Bulliet 1975:28–110; Knauf 1988: 9–15). The evidence shows that the Arabs in antiquity had a special relationship with the camel, documented already in their first appearance at Qarqar. According to the Assyrian evidence, the *arab* in Dumah were ruled by priest-queens, *apkallatu*, which is probably reflected in the legend of the Queen of Sheba in the Old Testament (Retsö 2003:173–176). From Hebrew and Greek sources we hear that they did not till the soil, they did not drink wine, they did not live in houses built from stone, they worshipped only two gods, Ruḍā and Allāt (= Alilat), and they had their hair cropped in a special way (Jeremiah 25:25; Herodotus, *Historiae* 3.8; Diodorus 19.94.2–95.2; Diodorus' original text speaks about Arabs only, not Nabataeans, cf. Retsö 2003:285–289). On the Assyrian reliefs one can see that they were dressed in a short loincloth, like the *ʾizār* of the present-day pilgrims to Mecca. They thus seem to constitute a distinct group of their own among the inhabitants of Arabia.

It is striking that they almost always appear in military contexts even if their weaponry seems to have been quite primitive. They were, however, amply used as auxiliary troops, border guards, and police forces by the empires. The occurrence of two forms of the designation for them may be connected with this. The Assyrians fought against *arab* but their employees are mostly called *arbāya*. Likewise, in South Arabia the enemies of the Sabaean and Himyarite kings are ʿRB, whereas their auxiliaries, allies, and border guards are usually called ʾʿRB. *Arbāya* is a *nisba* adjective from the collective noun *arab* and ʾʿRB is a plural of the Sabaean *nisba* ʿRBY. The Arabs employed by the empires often stood under a special command represented by an officer, a KBR in South Arabia, an *arabárkhes* in Egypt, Nabataea, and Syria and a RBYTʾ or a ŠLṬʾ in Hatra and Osrhoene (Retsö 2003:409–411). Finally, *arábioi* appear in a great mythical scene in Nonnos' *Dionysiaca*, written in the 5th century C.E. but deriving from a Greek text from ca. 400 B.C.E. In this epic the Arabs constitute the entourage of the semi-divine hero Lycurgus fighting against Dionysus, the god of wine. The myth itself originated in southern Syria where both Lycurgus and Arab presence are well documented in texts and inscriptions from late antiquity (Retsö 2003:610–614).

The picture of the pre-Islamic Arabs is thus somewhat reminiscent of that of the Rechabites and the Levites in the Old Testament. Even if the Rechabites in some way lived like nomads they were not typical nomads. The Arabs may originally have been a similar group, for whom the nomadic way of living was due to explicit ruling and ideology, not to economic necessity. There are no traces among them of the elaborate tribal system found in the later Arabo-Islamic sources. As a matter of fact, we do not hear of specific Arab tribes at all.

A special case is the group called Qidru or Qadru in Assyrian texts and Qedar in the Old Testament They seem to have had a unique relationship with the *arab* in Dūmah from the 8th century B.C.E. onwards and their kings are called king of Qedar or king of Arabs. The same double designation is found with the kings of Nabataea who mostly appear as kings of the *árabes* in Greek sources but as MLK NBṬ in the inscriptions and coins from Nabataea. From

Greek, Latin, and Talmudic sources it appears that Arabs and Nabataeans were not identical (Retsö 2003:312–314, 397, 527, 528). There are also hints at a close relationship between Qedar and the rulers of Nabataea: Qedar appears to dwell in the same area from the 5th century B.C.E. (Knauf 1985:96–108). The parallel between Arabs and Qedar/Nabataeans is unique and points to a special social and political structure in the regions east and south of Palestine during a long period.

A remarkable fact is that the term Arab disappears from North Arabia from ca. 300 C.E. onwards. The Namara inscription from 328 C.E. is, in fact, the last major testimony of Arabs in the Syrian desert. The ruler mentioned in it, Imruʾ al-Qays ibn ʿAmr, 'king of all Arabs', seems to have ruled Arabs in southern Syria and along the Roman Limes up to Osrhoene and taken part in a major military operation to South Arabia (see analysis and discussion in Retsö 2003:467–485). In the 4th century Arabs are replaced by Saracens in the Greek–Latin sources and by *ṭayyāyē* in the Syriac ones (Retsö 2003:505–521). There are numerous passages in Greek, Syriac, and Talmudic sources showing that these two were considered separate from the Arabs (e.g. Retsö 2003:487–491, 517–520). The disappearance of the Arabs from the sources could thus indicate the disappearance of the people so named, which supports the assumption that Arabs were a social or religious institution rather than being representative of a way of living or a genealogically defined group. It has been assumed that major changes took place in North Arabia in the 3rd century C.E., caused by the introduction of new weaponry from Iran (Caskel 1953a, 1953b; Bulliet 1975:100–105), and the disappearance of the Arabs in this period may thus be connected with this process. Classical Bedouin culture arose in connection with these changes, which means that it was not Arab in the old sense of the word.

2. The Islamic period

The study of the Arabs in the Islamic period is complicated by the problematic source situation for the most crucial period, viz. between the time of the Prophet and the fall of the Umayyads. Almost all relevant sources are written after the fall of the Umayyad dynasty and the analysis of them and their trustworthiness constitutes a major problem. In the poetry ascribed to the pre-Islamic poets there are astonishingly few mentions of the word Arab. This confirms the picture of the disappearance of the Arabs in non-Arabic sources from late antiquity and is, in fact, an argument in favor of the genuineness of the poetry. In the *Qurʾān*, the word *ʿarab* does not occur. The adjective *ʿarabiyy* is found on eleven occasions in the Meccan suras, always, it seems, referring to the language of the Revelation. The word *ʾaʿrāb* occurs in ten passages in the Medinan suras (Q. 33/20, 48/11, 16; 9/90, 97–101, 120). The *ʾaʿrāb* are criticized for not having performed their duties as participants in military operations undertaken by the *muhājirūn*, the *ʾanṣār*, and 'those who followed them' (Q. 9/100–101) According to Q. 33/20 they live in the *bādiya*, a word which most likely originally means the land outside a main, usually walled city, in this case most probably Yathrib/Medina (cf. Ibn Xaldūn, *Muqaddima* 1,2).

From several statements in the later literature it seems clear that the *ʾaʿrāb* were people of subordinate status to city-dwellers or the Muslim community with a duty to take part in warfare, a picture conforming to the one in the *Qurʾān* (Retsö 2003:82–87, against Marbach 1992). All inhabitants of the *bādiya* were thus not necessarily ˣRB, who look very similar to the ˣRB in the Sabaean texts. It might be surmised that, indeed, *ʾaʿrāb* is a loanword from South Arabia. The parallel function of the *ʾaʿrāb* in the *Qurʾān* and the ˣRB in the South Arabian inscriptions might indicate that they were a similar kind of institution, perhaps originating in South Arabia. Later on, in Islamic literature *ʾaʿrāb* is used as a general designation for people living as shepherds outside the sown area and becomes more a synonym of our term Bedouin.

As far as the word *ʿarab* is concerned, it is used with several different meanings in the Arabo-Islamic literature, especially when referring to the time of the Prophet and the 1st Islamic century. It may be used as a comprehensive term for all the Muslims, i.e. the tribesmen and their associates, in other words the citizens of the Umayyad Empire (Retsö 2003:63ff.). It can designate the full tribesman as opposed to the *mawlā* 'client', who could be a tribesman from another tribe (or a non-tribesman, Retsö 2003:69ff.). On several occasions, it is used for the *muhājirūn*, sometimes including the *ʾanṣār*, i.e. the kernel troops of the Islamic movement, whereas the Qurayš of Mecca are often contrasted with the *ʿarab*. It also seems to have been used by the *ʾanṣār* and their Yemeni allies as a designation for themselves opposed to other groups (Retsö 2003:71–76).

From this usage the idea developed that the Yemenis and the *ʾanṣār* were the 'real Arabs', *al-ʿarab al-ʿāriba*, a concept that can be traced to the beginning of the 8th century C.E., when the influence of these two groups was waning (Retsö 2003:28–40). This idea had a deep impact on historiography and is reproduced by the genealogical system found in most Islamic historical works. This latter case shows that the term *ʿarab* in the 1st Islamic century was a term with political and ideological importance due to some kind of prestige. From the contemporary, non-Arabic evidence from the 1st Islamic century it can be seen that the original designation for the Muslim army and the Islamic state was *muhājirūn*, Syriac *mahgrāyē*, Greek *moagarîtai*, but that the term Arab, Greek *árabes*, was introduced in the latter half of the Umayyad period (Retsö 2003:96–99).

Apart from these usages, which are widely spread in Arabic historiography, there are traces of further employments of the word *ʿarab*. Quite often it seems to designate a section of a tribe or people somehow attached to a tribe. An expression like *ʿarab Kināna* may well mean Arabs coming from, being attached to, or being a part of the Kināna tribe (Retsö 2003:76ff.). A similar usage is found in pre-Islamic South Arabian inscriptions, in which expressions like ʾʿRB KDT 'the *ʾaʿrāb* of Kinda' parallels ʾʿRB SBʾ 'the *ʾaʿrāb* of Saba' or ʾʿRB ḤḌRMWT the *ʾaʿrāb* of Ḥaḍramawt (Jamme 1962:665.2; Iryani 1990: 323; Robin and Gajda 1994:7.13; see Retsö 2003:552–553, 555). In the latter cases, there can be no doubt about the difference between the ʾʿRB and the 'mother nations' like Kinda, and a similar distinction can be claimed for Central and North Arabia. It is, however, to be observed that in this context the Islamic sources usually employ the term *ʿarab* rather than *ʾaʿrāb*, although there clear indications that the distinction between the two was often blurred in the Arabic tradition (Retsö 2003:80, n. 30). It is remarkable that the non-Arabic sources do not know the word *ʾaʿrāb*, a term which seems to have been forgotten early. The difference between ʿRB and ʾʿRB in South Arabia is reminiscent of the one between *arab* and *arbāya* in the Akkadian texts (cf. also Retsö 2003:5 for the difference between *ʿarab* and *ʿurbān* in contemporary Arabia).

Another deviant use of the word Arab is found in connection with the term *al-qurā l-ʿarabiyya* 'the villages (even cities) of the Arabs'. This expression seems above all to be connected with the area around present-day al-ʿUla. One source defines them as the towns of Xaybar, Fadak, as-Suwāriqiyya, and some others, which are sites around that area. Another passage refers to the sites around Dūmah. Several scholars have assumed a connection between this concept and the *ʾahl al-qurrāʾ* who appear in the earliest phase of Islamic history: the Biʾr Maʿūna incident, the *ridda* wars, and some events during the reign of *al-xulafāʾ ar-rāšidūn* (see Retsö 2003: 61, n. 166 for references). These people were known as readers of the Qurʾānic text but apparently also as connected with these *qurā*. The exact meaning of this is not yet clear but it is worthwhile comparing it with the statements in later sources that the *ʿarab* are people living in fortified cities (*ʾamṣār*, Retsö 2003:48–51). Even if this partly reflects the conditions in the conquered lands in the 1st Islamic century, we still have an obscure tradition that Arabs are connected with a special kind of town or settlement, which corresponds well with much of the pre-Islamic evidence. In Assyrian and Achaemenid times several Arab settlements (*dūr*) in Mesopotamia are mentioned. Arabs are also connected with the town of Hatra and, probably, with Hagar in Eastern Arabia, present-day al-Hufūf (Retsö 2003:168, 304–305, 307, 449, 434–453).

The many contradictory usages of the word Arab found in Arabo-Islamic literature are an important fact and may be interpreted plausibly as reflecting a historical development in the period from the time of the Prophet to the fall of the Umayyads. It seems that the comprehensive meanings of the word are later than the more limited references. This is also compatible with the disappearance of the word from the sources of late antiquity and its preservation in South Arabian ones. The settlement of Arabs in special, limited regions, on the one hand, and the function of the *ʾaʿrāb*, manifest in the Arabic sources, on the other, fits well the picture found in the South Arabian texts, which show a similar difference between ʿRB and ʾʿRB.

Their absence from the *Qurʾān* corresponds with the picture gleaned from pre-Islamic poetry that Arabs were not an important group in the environment where these texts were composed. But it seems that the designation received a renewed importance during the early period of the caliphate. It was used both by the Yemenis/*ʾanṣār* and the emerging tribal aristocracy as a self-designation. In the latter half of the Umayyad period, it was used as a term for all warriors and, eventually, for all members of the Islamic state. The classic view of the Arabs as the tribes that constituted the empire of the Umayyads and their descendants, as well as the distinction between the Yemeni branch of these tribes, *al-ʿarab al-ʿāriba*, and the Ishmaelite branch, *al-ʿarab al-mustaʿriba* or *al-mutaʿarriba*, seems to have been established in the latter half of the Umayyad period, perhaps during the reign of ʿUmar ibn ʿAbd al-ʿAzīz (717–720; cf. Dagorn 1981:208–209, 217).

The application of the term *ʿarab* to the Yemenis shows very clearly that it was a term that could be used for political and ideological purposes and that its meaning was not a self-evident fact. In pre-Islamic times, for instance, there was a clear difference between Arabs and the peoples of Yemen, a difference which was thus abolished by this new politically conditioned definition. The main factor is its transformation from a designation of an institution to a characteristic of members of a tribal society during the time of the first Islamic empire. The use of *ʿarab* as a designation for the Muslims lost its meaning after the fall of the Umayyad dynasty. It seems, however, that the term from then on had been adopted definitely by the tribes as a designation for themselves. The extent of its use among the tribes during the Islamic Middle Ages is, however, still uninvestigated. The present-day usage may give a clue (see below).

The multifarious usages of the term Arab in the sources dealing with the early Islamic period is reflected in the entries on *ʿarab* in the great medieval dictionaries, from al-Xalīl's *Kitāb al-ʿayn* via al-ʾAzharī's *at-Tahḏīb fī l-luġa* and Ibn Manẓūr's *Lisān al-ʿarab* to az-Zabīdī's *Tāj al-ʿarūs*. These entries, which largely tend to copy each other, preserve the many contradictory meanings of the word and should be read with this fact in mind. In his *Muqaddima*, Ibn Xaldūn defines the people of the *badw* as those who dwell outside the cities, making their living as peasants or shepherds. The harder the living conditions, the stronger the sense of blood ties. The most extreme of these are the *ʿarab*, who dwell under the most severe conditions of all, and for whom tribal solidarity built on blood ties (*ʿaṣabiyya*) is the main factor of identity. The *ʿarab* are thus a section of the *badw*, the people of the *bādiya* or the Bedouin, characterized by the all-encompassing importance of genealogical relations. This means that not all Bedouin are *ʿarab* (Ibn Xaldūn I, 2/Rosenthal 1958:I, 249 ff.).

3. The modern period

During the last two centuries there have been basically two definitions of Arabs. The first is the one used by many of the tribes in Arabia and North Africa and their settled relatives, distinguishing them from other people, sedentary or nomads. The main criterion for being an Arab according to this definition is the possession of an accepted genealogy, making it possible to trace one's ancestors back to some famous forebear and/or ultimately back to the tribes of Arabia at the time of the Prophet. This definition is widespread and well-documented through modern texts collected from informants (Retsö 2003:1–7). It seems that it is quite old, closely related to Ibn Xaldūn's description and traceable even to the earliest Arabic sources as one of the meanings occurring there. It can be assumed that it originated in the Umayyad period. An *ʿarab* is thus a tribesman with his genealogy in good order, be he a nomad or a sedentary (one might compare the use of the term *arap* in Afghanistan for tribes that do not speak Arabic but trace themselves back genealogically to the peninsula; cf. Kieffer 1980).

The other definition is the one launched by the ideologists behind the modern Arab nationalist movements. The criteria are diverse but as a rule rather vague: common language, common culture, common history, sometimes even a common ancestry. It is ultimately based on the romantic European concepts of nationhood (Choueiri 2000). This definition, developed during the last decades of the 19th century, was proclaimed to a wide public for the first time by N. Azoury in his *Le réveil de la nation arabe* (1905; see Hourani 1970:277–279). At this time the extent of 'the Arab nation' had not yet been precisely defined, and Azoury excluded the whole of North Africa. The concept of the Arab nation turned out to be very useful as a main ideological weapon in the struggle against the European colonial powers after the First World War. The political dominance of Great Britain and France in the Middle East and North Africa is perhaps the main factor behind the creation of the modern Arab nation, which also has transformed the citizens of the modern states into Arabs. Eventually, the pan-Arab ideology was made politically manifest in the creation of the Arab League in 1946. The influence of the new concept of Arabism culminated ideologically in the writings of Sāṭiʿ al-Ḥuṣrī (d. 1949) and politically with the career of Gamāl ʿAbd an-Nāṣir, who tried to implement the idea by taking the first steps toward creating a pan-Arab political structure. The failure of Nāṣir's political project, combined with the repeated defeats of the Arab states by Israel has diminished the practical importance of the Arab ideology, although it still plays an important role in political rhetoric. Its most important result is, apart from the abolition of foreign political control in the Arab world, the revivification of the → *ʿarabiyya* language and its introduction as the official language in 18 states from the Atlantic to Oman. A widespread sense of belonging to an Arab nation, independent of the borders of the modern Arab states, bound together by a common language and cultural heritage, is also a lasting result of the modern Arab nationalist movement.

4. The language of the Arabs

A distinction should be made between the linguistic data from Arabia and adjacent regions, which by modern linguists are classified as Arabic, and the explicit characterization of a language in the ancient sources, pre-Islamic and Islamic, as Arabic. In Greek and Latin pre-Islamic sources there is a handful of references to an Arabic language, *arabikē glōssa*, *arabikē diálektos*, the *phōnē* of the *árabes*, *arabicus sermo*, and *arabica lingua* (Retsö 2003:591). The earliest one is found in Agatharchides' description of the Red Sea, written ca. 140 B.C.E., mentioning the name of a plant, *lárimna*, as *arabistí*. In the Talmud there are approximately 30 words which are said to be used by Arabs or to derive from Arabia (Retsö 2002). Most of the occurrences date from after the turn of the era and refer to a language/languages in and around the Provincia Arabia. The only definite exception is a notice in the *Periplus Maris Erythraei* (ca. 70 C.E.) about 'holy men' on the island of Sarapis, i.e. present-day Masira, using an *arabikē glōssa*. The exact nature of this/these language(s) escapes us. The words in the Talmud are varied and only a few of them are recognizable as Arabic according to later standards.

The Qurʾānic passages referring to an Arabic language all deal with the oral performance of the text, *al-Qurʾān*, and seem to distinguish this from the writing, *al-kitāb*. They occur in contexts where the authority of the revelation is at stake, in which the Arabic character of the recitation is adduced as proof that the text is indeed of divine origin. The *ʿarabī* language thus seems in some way to have been associated with speech from the non-human world. The opposite term for *ʿarabī* in the *Qurʾān* is *ʾaʿjamī*, a term originally meaning 'crooked', 'twisted', in a linguistic sense (Q. 41/44, 26/198–199, 16/103). In later Arabic lexicography the term *ʾaʿjamī* contrasted with → *faṣīḥ*. The lexicographers make a distinction between *ʾaʿjam* 'non-*faṣīḥ*', and *ʿajam*, 'foreigner, non-Arab', especially Iranians, but this distinction is probably secondary (Retsö 2003:24–28). The lexicographers sometimes identify *ʿarab* and *faṣīḥ*, but mostly state that one can be *faṣīḥ* without being *ʿarabī* (*Lisān al-ʿarab*, s.v. *ʿJM*). This would mean that the Arab identity was not dependent on linguistic criteria. This raises the problem why the language was called Arabic in the *Qurʾān*, which is not quite clear yet. The connection between the Arabs as they appear in the 1st Islamic century and the language of the *Qurʾān* thus remains more obscure than usually realized. It is, however, clear that, at least until the beginning of the 9th century C.E., there were people among the tribes who had a competence in the language of the *Qurʾān* through their knowledge of the language of the poetry. The early grammarians often refer to informants called *ʿarab*. The often adduced claim about the linguistic excellence of the Qurayš in Mecca, found in many medieval works, is based on theology and ideology rather than historical and linguistic facts.

5. Etymology

Many etymological explanations of the word *ʿarab* have been suggested, none of which are quite convincing. An explanation found in the Middle Ages is that the name is derived from the word *ʿaraba*, the name of a region (*Lisan al-ʿarab* s.v. ʿRB; Retsö 2003:52–53). There is, however, no immediate connection between the areas called ʿAraba (part of the Jordan valley, possibly an area near Mecca) and the Arabs of the sources. Explanations based on the assumption of metathesis of a root ʿBR 'pass by', sometimes linking it with *ʿibrī* 'Hebrew', ascribing to it the meaning 'passer-by', i.e. 'nomad', can be dismissed.

An explanation connecting *ʿarab* with a root meaning 'to be mixed' is perhaps closer to the truth (Lane 1863–93 s.v. ʿRB). It could explain the opposite meaning found in *muʿrib*, *ʿārib* 'pure, unmixed', which is well documented as a characterization of the *ʿarabiyya* language or of someone speaking it. The polarity of meanings in a root is well known in Semitic languages. It is, however, doubtful whether this is the original meaning of the designation for the groups of people concerned. The concern for genealogical purity is well known among the traditional *ʿarab* today, but since the genealogical definition of an Arab seems to have risen quite late (end of the Umayyad period), genealogical purity cannot have been the original meaning of the term. The identification of the earliest Arabs as a kind of religious community dedicated to a deity or a divine hero might provide an explanation of a different kind. The root ʿRB is well known in the meaning 'enter' (with a variant ĠRB) from which 'mix' can be derived. In Ancient South Arabia ʿRB is well documented in the meaning 'offer', 'give', 'dedicate to a god'. With this is connected the meaning 'give a pledge or security' from which the Arabic nouns *ʿarabūn*, *ʿurbān*, the Hebrew *ʿerabōn*, and the Syriac *ʿrābā* are formed. This word was borrowed at an early stage into Greek as *arrabōn* 'pledge, security' (cf. also Latin *arrabo*). When Herodotus describes the initiation ceremony of the *arábioi* he characterizes it by a Greek word, *pístis*, which may also mean 'pledge', 'guarantee'. To this may be added the many instances in Akkadian where forms of this root are used to designate people and things standing in various kinds of dependency on others: thus *ērib bītim* 'temple official', *ana ilkim erēbum* 'become vassal to someone', *ana libbi adê erēbum* 'to swear an oath, *ana mārūtīšu erēbum* 'to be adopted by someone', etc. (Retsö 2003:597). Both Akkadian and Ugaritic know the noun *erbum*/ʿRBM 'gift to a temple'.

All these meanings can be connected with the concept of 'entering', i.e. 'entering into a dependency upon someone'. It is to be noticed that the use of the root ʿRB in these specialized meanings is found in the earliest Semitic languages, a use which is not alive in the *ʿarabiyya*, showing that it is an archaic survival in that language. The word *ʿarab* would then be a *qatal* formation of a verb originally serving as an abstract or verbal noun and then as an infinitive. Starting from these functions, it could be used as a collective noun designating those who were involved in the process designated by the verb. This structure can still be seen in a few Arabic words of the same pattern such as *ṭalab* 'search'; 'searchers' (from *ṭalaba/yaṭlubu* 'to search'), *ḥaras* 'guarding'; 'guards' (from *ḥarasa/yaḥrusu* 'to guard'); *ġalab*, 'noise'; 'noisy ones' (from *ġalaba/yaġlibu* 'to be noisy').

Bibliographical references

Primary sources

Diodorus of Sicily. Trans. C.H. Oldfather. 12 vols. London, New York and Cambridge, Mass.: Loeb Classical Library, 1933–1967.

Herodotus, *Historiae*. Trans. A.D. Godly. Cambridge, Mass. and London: Loeb Classical Library, 1920–1981.

Ibn Manẓūr, *Lisān al-ʿArab*. 20 vols. Cairo: Dār al-Maʿārif, n.d.

Ibn Xaldūn, *Muqaddima = The Muqaddimah. An introduction to history*. Trans. Franz Rosenthal. New York: Routledge and Kegan Paul, 1958.

RÉS = *Répertoire d'épigraphie sémitique*, vols. V–VII. Paris, 1928–1950.

Secondary sources

Azoury, Negib. 1905. *Le réveil de la nation arabe dans l'Asie turque: En présence des intérêts et des rivalités des puissances étrangéres, de la curie romaine et du patriarcat œcuménique*. Paris: Plon.

Bulliet, Richard. 1975. *The camel and the wheel*. Cambridge, Mass. and London: Harvard University Press.

Caskel, Werner. 1953a. *Die Bedeutung der Beduinen für die Geschichte der Araber*. Opladen and Cologne.

——. 1953b. "Zur Beduinisierung Arabiens". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 28.*28–*36.

Choueiri, Youssef M. 2000. *Arab nationalism: A history*. Oxford and Malden, Mass.: Blackwell.

Dagorn, René. 1981. *La geste d'Ismaël d'après l'onomastique et la tradition arabes*. Geneva and Paris: Librairie Droz and Librairie Champion.

Hourani, Albert. 1967. *Arabic thought in the liberal age 1798–1939*. Oxford: Oxford University Press.

Iryani, Muṭahhar ʿAbd ar-Raḥmān al-. 1990. *Nuqūš musnadiyya wa-taʿlīqāt*. 2nd. ed. Sanaʾ.

Jamme, Albert. 1962. *Sabaean inscriptions from Mahram Bilqîs (Mârib)*. Baltimore: Johns Hopkins Press.

Kieffer, Charles M. 1980. "L'arabe et les arabophones de Bactriane (Afghanistan). I. Situation ethnique et linguistique". *Welt des Islams* 20.178–196.

Knauf, E. A. 1985. *Ismael: Untersuchungen zur Geschichte Palästinas und Nordarabiens im 1. Jahrtausend v. Chr.* Wiesbaden: O. Harrassowitz.

——. 1988. *Midian: Untersuchungen zur Geschichte Palästinas und Nordarabiens am Ende des 2. Jahrtausends v. Chr.* Wiesbaden: O. Harrassowitz.

Lane, Edward William. 1863–1893. *Arabic–English lexicon.* 8 vols. London: Williams and Norgate.

Marbach, Amikam. 1992. "Maʿnā l-muṣṭalaḥ ʿarab ḥasb al-maʿājim wa-l-*Qur'ān* wa-l-*Kitāb* li-Sībawayhi wa-*Muqaddimat* Ibn Xaldūn". *al-Karmil* 13.145–178.

Retsö, Jan. 2002. "Das Arabische der vorislamischen Zeit bei klassischen und orientalischen Autoren". *Neue Beiträge zur Semitistik. Erstes Arbeitstreffen der Arbeitsgemeinschaft Semitistik in der Deutschen Morgenländischen Gesellschaft vom 11. bis zum 13. September 2000 and der Friedrich-Schiller-Universität Jena*, ed. Norbert Nebes, 139–146. Wiesbaden: O. Harrassowitz.

——. 2003. *The Arabs in antiquity: Their history from the Assyrians to the Umayyads.* London: Routledge Curzon.

Robin, Christian and Iwona Gajda. 1994. "L'inscription du Wadi ʿAbadan". *Raydan* 6.113–137.

Jan Retsö (Göteborg University)

Arabic Alphabet for Other Languages

The Arabic alphabet is employed for a large number of languages other than Arabic, the Semitic language for which it was originally designed. After the use of the Latin script for the written manifestation of many of today's languages, the Arabic writing system has spread far and wide, chiefly due to Islam. Consequently, it is the second most widespread segmental script in the world rendering a variety of different languages from different language families. Among the most important languages today using a modified form of the Arabic script are the following (in alphabetical order using the language family as the characteristic designation): the Berber languages of North Africa (with the notable exception of Tuareg, which has developed its own Tifinagh script); the Dravidian language Moplah, a dialect of Malayalam (which is closely related to Tamil); the Indo-Aryan languages of the Indian subcontinent, Urdu, Sindhi, and Kashmiri; the Iranian languages Balochi, Pashto, Persian, and Kurdish; and the Malayo-Polynesian (Austronesian) languages Malagasy, Malay, and Sulu. The Arabic alphabet has also been used to write Turkic and Caucasian languages and is now competing with the Cyrillic script in some of the former parts of the Soviet Union. Chief among these latter cases is Ottoman Turkish (Osmanlı), which was written in Arabic script from approximately 1300 C.E. until 1928.

Besides these languages, several other languages were commonly written in a modified version of the Arabic alphabet, and a few still are, particularly by some Muslim scholars. These are: the Bantu language Swahili in East Africa; Nilo-Saharan Kanuri in eastern Nigeria; Chadic Hausa in northern Nigeria and Niger; West Atlantic Fulani (Fula) in West Africa; Ethio-Semitic Harari in Ethiopia; Indo-European Albanian; Slavic Serbo-Croatian in Bosnia; and others.

It is important to stress that the Arabic script was successfully adopted because of Islam, which spread rapidly from its humble beginnings in the 7th century C.E. in western Arabia (Ḥijāz). Since the *Qur'ān* was to be studied along with the Muslim prayers in the original Arabic, the Arabic script came to have a unifying effect on Muslims everywhere. Soon, it made its calligraphic presence felt over the entire Arab world, Iran and Afghanistan, the countries of the Ottoman Empire, the Mughal Empire, Malaysia and Indonesia, and elsewhere. Quite differently from the Copts in Egypt, the last of whom gave up speaking their language in favor of the ubiquitous Arabic in the 14th century, the peoples of Iran, Afghanistan, Malaysia, and the other countries mentioned never gave up their native languages. Thus, the Iranians embraced Islam, giving up their native religions of Zoroastrianism and Manichaeism, while at the same time adopting the Arabic script to fit their needs. Moreover, they borrowed numerous Arabic vocabulary items. This has been a familiar story with other Muslim peoples.

The Arabic alphabet has proven itself to be adaptable to the phonological structure of the borrowing non-Semitic languages. The usual scenario is that all the Arabic graphemes have been borrowed primarily for the preservation of the original Arabic orthographic representations. Of course, the pronunciation of these Arabic loanwords differs in accordance with the phonologies of the borrowing languages.

Since the Arabic emphatics, interdentals, and pharyngeals are, in fact, rarities in the 7,000 or so languages of the world, these consonants are most often pronounced differently by the borrowing languages (for example, Kurdish has both [ħ] and [ʕ]). Thus the Arabic /ḍ/ ض is pronounced /z/ in Persian and Urdu. Persian /s/ may be rendered in one of three distinct ways: ث, س, and ص. Persian /z/ may be rendered in one of four distinct manners: ز, ذ, ض, and ظ. Naturally, five out of the seven preceding graphemes are pronounced differently in Arabic, although even one of these, ذ, is pronounced /z/ by many Egyptians, for example when speaking or reading Modern Standard Arabic. Similarly, these same Egyptians tend to pronounce the voiceless interdental fricative ث as /s/.

The borrowing non-Semitic languages have borrowed all the Arabic letters including the feminine marker *tāʾ marbūṭa* (with a few exceptions). Since the latter grapheme is often pronounced /t/, Persian has changed the spelling of words containing it to the regular *tāʾ ṭawīla; ḥaqīqatan* 'really' is not spelled in Persian the same way it is spelled in Arabic. To be sure, new graphemes had to be created for non-Arabic phonemes. Special superscript or subscript diacritics were invented for the new graphemes. In Persian, for example, پ <p> is ب <b> but with three dots underneath the basic configuration rather than one; ژ <ž> is ز <z> but with three dots on top. This type of modification is the basic strategy for the creation of the new graphemes, although other languages use devices other than dots placed above or below the basic curvature marking of the grapheme.

Ottoman Turkish (Osmanlı) contained awkward deficiencies in the Arabic script adapted to this Turkic language. For example, Arabic *kāf* renders, in addition to [k], also [g] and [ɲ]. Swahili in Arabic characters has some problems which affect even native speakers. There are no distinct symbols for [p], [g], [č], [v], [ɲ], and [ŋ]. Thus, there is much ambiguity present. Many Persian forms are used when natives are beginning to become literate, but rarely later on. Since Swahili does not have a uvular stop, there is little synchronic justification of spelling any word with a *qāf*, except Arabic orthographic tradition.

1. Persian

Table 1 (typical of all the tables which follow) presents the Persian script. The left-hand column gives the Library of Congress (LC) transliteration, and International Phonetic Alphabet (IPA) transcription, including the major variant pronunciations. The remainder of the table presents the allographs, traditionally arranged according to the isolated, final connected, initial, and medial shapes. There are many calligraphic styles, such as the sometimes difficult-to-decipher *šekaste* (lit. 'broken') writing, which is not treated here.

Due to its un-Semitic phonology, many Arabic graphemes have been revalued, as has been noted. Arabic *tanwīn* and *šadda* are not normally written. There are four new letters all characterized by three dots: <p>, <g>, <č>, and <ž>. Today's گ originally had three superscript dots, which evolved into a straight line stroke – a basic feature of handwritten forms (one reason is that they are faster to produce than the cumbersome dots).

2. Kurdish

Kurdish has taken the approximates [w, j] and the laryngeals [ʔ, h] and used them to represent the vowels of the language. Thus, it has attained the status of a true alphabet. Some items are in the Persian style; e.g., <ž> is ژ. However, <č> uses the Persian چ, which has three dots in the middle and an additional dot as a superscript. Table 2 contains a complete list of consonants and vowels.

3. Pashto

Pashto writing has much in common with its neighbor, Persian. Thus, /g/ is often written as in Persian with *gāf* گ or the pure Pashto *kāf* ګ with a circle added below the upper stroke. Table 3 contains a complete list of consonants and vowels. There are some differences from the forms presented in Table 3 depending on dialect. For example, in Pakistan, Pashto retroflexes can be written in Urdu fashion (see under Urdu). There are a few options in spelling as well; e.g., the *hamza* is occasionally used to represent [e]. Also, *plene* spelling is an option, with ی <y> indicating [i] or [e], and و <w> for short [u] or [Ö]. This is a built-in ambiguity.

4. Kashmiri

One of the most developed vocalic offshoots of the Arabic script is used for Kashmiri. Here 16 vocalic graphemes represent the vowel phonemes plus a palatalization marker written with undotted ى. The aspirated stops are uniformly done as in Urdu, with the medial <h> attached to the non-aspirated version of the stop, e.g. پھ <ph>. Table 4 lists the consonants and vowels.

5. Urdu

Since the Urdu script has been adapted directly from that of Persian (one should keep in mind that Persian was a literary language of India), many features of Persian pronunciation and script also apply to Urdu. One of its basic characteristics is the use of the emphatic <ṭ> to represent the retroflex consonants. In some published works, between two and four dots as superscript are variations. The reason ط was adopted is that it was considered to be an unusual kind of /t/ in Arabic, and thus it resembled a retroflex /ṭ/ (which does not exist in Arabic). There are three aspirated retroflex stops: <ṭh>, <ḍh>, and <ṛh>; however, the last two of these are allophones.

Aspiration is marked with the figure-eight looking allograph of <h>. All the phonemic nasalized vowels are written in final position as a *nūn* without the dot. Elsewhere, they are written with *nūn*. The script distinguishes <ē> from <ī>: *laṛkē* 'boy(s)' vs. *laṛkī* 'girl'. Table 5 presents the consonants and vowels.

6. Sindhi

Sindhi, although an Indo-Aryan language, has phonemes not found in Urdu. It has created a new aspirated stop grapheme, a new retroflex grapheme, and a new imploded grapheme (three dots on top of a <d> is an imploded <ɗ>. Table 6 presents the consonants and vowels.

7. Ottoman Turkish

One of the most basic typological features of Turkish is vowel harmony. The graphemes of the Arabic emphatic consonants are associated with the back vowels, while the corresponding non-emphatics are associated with the front vowels. Thus a word such as Turkish *kara* could be written *qaraa* ~ *qarah* ~ *qaarah*, and the <q> implies the back vowel /a/ in the first two of these. If it is written with an *ʾalif*, there is no ambiguity possible between front and back vowel. If this word were written with a <k>, it would imply that a front vowel would follow. Similarly, *dört* 'four' is written either as <dwrt> or <drt>. If it were written with an emphatic <ḍ>, that would imply that a back vowel would follow. Table 7 lists the consonants and vowels. It is easy to see that <v> has eight different values, thus causing the reader of Ottoman Turkish considerable difficulty. This was one of the reasons Atatürk's spelling reform was so successful.

8. Uyghur

The most interesting thing about Uyghur as a written language is that it does not preserve the spelling of the Arabic loanwords. This in itself is quite unusual. As in Kurdish, the <h> is used to mark [e]. There are many resemblances to Persian and Urdu types, e.g. the letters <p>, <č>, and <ž>. Some possible ambiguities in the vowels are disambiguated, as, for example, [u] is a *wāw* with a small *wāw* on top as a diacritical mark; otherwise it is [o], and [ö] has a haček on top, and [ü] has a small vertical stroke on top of it. The script also quite cleverly differentiates [i] from [e] – something which even unvocalized Arabic does not do (Arabic *mayl* ~ *mēl* 'inclination' vs. *mīl* 'mile' are both written ميل. Table 8 presents the consonants and vowels.

9. Malay (Jawi)

The Arabic (or Jawi) script is still used for Malay on the eastern coast of peninsular Malaysia (there are two daily newspapers printed in it). However, it has lost ground over the past decades to the Latin script as a result of the use of Latin script for Bahasa Indonesia, almost the same language as Malay. However, one can still see Arabic script occasionally on buildings (e.g. banks) in Kuala Lumpur and other cities. Arabic has a long history of many centuries' use in Malaysia, and it has made certain innovations; for example, an <f> with three dots on top is a <p>, and an <f> can be read in one of two fashions – either as an [f] or a [p]. An <ŋ> is written as a *ġayn* with three dots on top. Table 9 paints an accurate picture as to the consonants and vowels.

10. Hausa

Hausa was written in Arabic script (called *ajami* < Arabic *ʿajam* 'Persian; non-Arab') long before the introduction of the Roman alphabet (called *boko*

< ? English *book*). The former is tied to Islam by Hausa scholars, while the latter was introduced by the British and French colonials around 1900. Many traditional Hausa scholars still use *ajami*. However, most written Hausa is in *boko*.

The short vowels are represented as follows: *fatḥa* /a/, *kasra* /i/, a dot below a consonant /a/; *ḍamma* /o/ or /u/; long vowels use the same diacritics plus *ʾalif, yāʾ*, or *wāw* respectively. The diphthongs are written as in Arabic.

11. SWAHILI

Table 11 depicts the Swahili adaptation of the Arabic script. It is based on Allen (1945) and Beech (1918). Apparently, the literacy rate in Swahili in 1945 in Arabic script was quite high. The vowels are usually written, including the short vowel diacritics, except in well-known Arabic phrases.

The use of the Arabic script for Swahili was never standardized. Thus, there is some variation in orthographic practice. Since pronunciation varies, so does the spelling. The educational background of the writer is also of great importance. If a writer has studied Arabic (and almost all have, at least to some extent), many Arabic loanwords are spelled in Arabic fashion. With minimum exposure to Arabic, more phonetic (un-Arabic) spellings prevail. Strange as it may seem, the word *kitabu* 'book' is spelled in un-Arabic fashion without *ʾalif*. The word *bwana* 'Mr., Sir' is written both with and without *ʾalif*: <bwn> and <bw?n.

The short vowels are represented by *fatḥa* /a/, *kasra* /e/ or /i/, *ḍamma* /o/ or /u/; *sukūn* is used for zero vowel. The *šadda* is often used to mark the gemination of a consonant; however, some writers omit it. The *tāʾ marbūṭa* sometimes occurs as in Arabic, as do *ʾalif maqṣūra, hamza, madda*, etc.

BIBLIOGRAPHICAL REFERENCES

Allen, J.W.T. 1945. *Arabic script for students of Swahili*. Dar es Salaam: published for the editorial committee, Tanganyika Notes and Records.

Barry, Randall K. (ed.). 1991. *ALA-LC romanization tables: Transliteration schemes for non-roman scripts*. Washington, D.C.: Library of Congress.

Beek, Mervyn W.H. 1918. *Aids to the study of Ki-Swahili*. London: Kegan Paul, Trench, Trübner.

Bright, William and Saeed A. Khan. 1958. *The Urdu writing system*. Washington, D.C.: American Council of Learned Societies.

Doroszkiewicz, Małgorzata. 1990. "The orthographical rules of ajami script in Hausa". *Hausa Studies* (Warsaw) 3.36–48.

Hahn, Reinhard F. 1991. *Spoken Uyghur*. Seattle: University of Washington Press.

Lewis, M.B. 1958. *A handbook of the Malay script*. London: Macmillan.

MacKenzie, D.N. 1987. "Pashto". *The world's major languages*, ed. Bernard Comrie, 547–563. London: Croom Helm.

McCarus, Ernest N. 1958. *A Kurdish grammar: Descriptive analysis of the Kurdish of Sulaimaniya, Iraq*. New York: American Council of Learned Societies.

Naim, C. Muhammad. 1971. "Arabic orthography and some non-Semitic languages". *Islam and its cultural divergence*, ed. Girdhari L. Tikku, 113–144. Urbana: University of Illinois Press.

Nemeth, J. 1962. *Turkish grammar*, trans. T. Halasi-Kun. The Hague: Mouton.

Newman, Paul. 2000. *The Hausa language: An encyclopedic reference grammar*. New Haven: Yale University Press.

Paper, Herbert H. and Muhammad Ali Jazayeri. 1955. *The writing system of modern Persian*. Washington, D.C.: American Council of Learned Societies.

Yegorova, R.P. 1971. *The Sindhi language*, trans. E.H. Tsipan. Moscow: Nauka.

ALAN S. KAYE
(California State University)

Table 1. The Persian alphabet

LC Transliteration	*Transcription*	*Isolated*	*Final*	*Initial*	*Medial*
–	[ʔ, ɔ, æ, Ø]	ﺍ	ﺎ	–	–
b	[b]	ﺏ	ﺐ	ﺑ	ﺒ
p	[p]	ﭖ	ﭗ	ﭘ	ﭙ
t	[t]	ﺕ	ﺖ	ﺗ	ﺘ
s̱	[s]	ﺙ	ﺚ	ﺛ	ﺜ
j	[dʒ]	ﺝ	ﺞ	ﺟ	ﺠ
ch	[tʃ]	ﭺ	ﭻ	ﭼ	ﭽ
ḥ	[h, Ø]	ﺡ	ﺢ	ﺣ	ﺤ
kh	[x]	ﺥ	ﺦ	ﺧ	ﺨ
d	[d]	ﺩ	ﺪ	–	–
ẕ	[z]	ﺫ	ﺬ	–	–
r	[r]	ﺭ	ﺮ	–	–
z	[z]	ﺯ	ﺰ	–	–
zh	[ʒ]	ﮊ	ﮋ	–	–
s	[s]	ﺱ	ﺲ	ﺳ	ﺴ
sh	[ʃ]	ﺵ	ﺶ	ﺷ	ﺸ
ṣ	[s]	ﺹ	ﺺ	ﺻ	ﺼ
z̤	[z]	ﺽ	ﺾ	ﺿ	ﻀ
ṭ	[t]	ﻁ	ﻂ	ﻃ	ﻄ
ẓ	[z]	ﻅ	ﻆ	ﻇ	ﻈ
ʿ	[ʔ, Ø], preceding V → V:	ﻉ	ﻊ	ﻋ	ﻌ
gh	[ɣ]/V_V; [q, ɢ, x]	ﻍ	ﻎ	ﻏ	ﻐ
f	[f]	ﻑ	ﻒ	ﻓ	ﻔ
[q]	[q, ɢ]	ﻕ	ﻖ	ﻗ	ﻘ
k	[k]	ﻙ	ﻚ	ﮐ	ﮑ
g	[g]	ﮒ	ﮓ	ﮔ	ﮕ
l	[l]	ﻝ	ﻞ	ﻟ	ﻠ
m	[m]	ﻡ	ﻢ	ﻣ	ﻤ
n	[n]	ﻥ	ﻦ	ﻧ	ﻨ
v	[v, u, o, ow]	ﻭ	ﻮ	-	-
h	[h, Ø, ɛ, æ], Arab. fem. [t]	ﻩ, ﺓ	ﻪ, ﺔ	ﻫ	ﻬ
y	[j, i, e]	ﯼ	ﯽ	ﯾ	ﯿ

Table 2. The Kurdish alphabet

LC Transliteration	*Transcription*	*Isolated*	*Final*	*Initial*	*Medial*
–	[ʔ]	ا	ـا	–	–
b	[b]	ب	ـب	بـ	ـبـ
p	[p]	پ	ـپ	پـ	ـپـ
t	[t]	ت	ـت	تـ	ـتـ
j	[ʤ]	ج	ـج	جـ	ـجـ
ch	[tʃ]	چ	ـچ	چـ	ـچـ
ḥ	[ħ]	ح	ـح	حـ	ـحـ
kh	[x]	خ	ـخ	خـ	ـخـ
d	[d]	د	ـد	–	–
r	[ɾ]	ر	ـر	–	–
ṛ	[r̄]	ڕ	ـڕ	–	–
z	[z]	ز	ـز	–	–
zh	[ʒ]	ژ	ـژ	–	–
s	[s]	س	ـس	سـ	ـسـ
sh	[ʃ]	ش	ـش	شـ	ـشـ
ṣ	[s̶]	ص	ـص	صـ	ـصـ
ʻ	[ʕ]	ع	ـع	عـ	ـعـ
gh	[ɣ]	غ	ـغ	غـ	ـغـ
f	[f]	ف	ـف	فـ	ـفـ
v	[v]	ڤ	ـڤ	ڤـ	ـڤـ
q	[q]	ق	ـق	قـ	ـقـ
k	[k]	ك	ـك	کـ	ـکـ
g	[ɡ]	گ	ـگ	گـ	ـکـ
l	[l]	ل	ـل	لـ	ـلـ
ḷ	[ɫ]	ل	ـل	لـ	ـلـ
m	[m]	م	ـم	مـ	ـمـ
n	[n]	ن	ـن	نـ	ـنـ
w	[w]	و	ـو	–	–
h	[h, ə]	ه	ـه	هـ	ـهـ
y	[j, iː, eː]	ی	ـی	یـ	ـیـ
a	[ə]	ه(ه)	ـه	هـ	ـهـ
ā	[ɑ]	ا	ـا	–	–
u	[uː, ʉ, o]	و	ـو	–	–
ū	[uː]	وو	ـوو	–	–
ī	[iː]	یی	ـیی	ییـ	ـییـ
ē	[eː]	ی	ـی	–	ـیـ
o	[o]	ۆ	ـۆ	–	ـۆ

Table 3. The Pashto alphabet

LC Transliteration	*Transcription*	*Isolated*	*Final*	*Initial*	*Medial*
-	[ʔ, Ø, ə, ɑ(:), i, u]	ا	ـا	–	–
b	[b]	ب	ـب	بـ	ـبـ
p	[p]	پ	ـپ	پـ	ـپـ
t	[t]	ت	ـت	تـ	ـتـ
ṭ	[ʈ]	ټ	ـټ	ټـ	ـټـ
s̲	[s]	ث	ـث	ثـ	ـثـ
j	[ʤ]	ج	ـج	جـ	ـجـ
ch	[tʃ]	چ	ـچ	چـ	ـچـ
ḥ	[h, Ø]	ح	ـح	حـ	ـحـ
s	[ts]	څ	ـح	حـ	ـحـ
ż	[dz]	ځ	تح	حـ	ح
kh	[x]	خ	ـخ	خـ	ـخـ
d	[d]	د	ـد	–	–
ḍ	[ɖ]	ډ	ـډ	–	–
z̲	[z]	ذ	ـذ	–	–
r	[r]	ر	ـر	–	–
ṛ	[ɽ]	ړ	ـړ	–	–
z	[z]	ز	ـز	–	–
zh	[ʒ]	ژ	ـز	–	–
ẓh	[ʐ]	ږ	ـږ	–	–
s	[s]	س	ـس	سـ	ـسـ
sh	[ʃ]	ش	ـښ	ښـ	ـښـ
ṣh	[ʂ]	ښ	ـص	صـ	ـصـ
ẓ̈	[z]	ض	ـض	ضـ	ـضـ
ṭ̈	[t]	ط	ـط	طـ	ـطـ
ẓ	[z]	ظ	ـظ	ظـ	ـظـ
ʿ	[ʔ, Ø, ɑ], preceding V→Vː	ع	ـع	عـ	ـعـ
gh	[ɣ]	غ	ـغ	غـ	ـغـ
f	[f]	ف	ـف	فـ	ـفـ
q	[q]	ق	ـق	قـ	ـقـ
k	[k]	ك	ـك	کـ	ـکـ
g	[ɡ]	ګ	ـګ	ګـ	ـګـ
l	[l]	ل	ـل	لـ	ـلـ
m	[m]	م	ـم	مـ	ـمـ
n	[n]	ن	ـن	نـ	ـنـ
ṇ	[ɳ]	ڼ	ـڼ	ڼـ	ـڼـ
w	[w, o, u(ː)]	و	ـو	–	–
h	[h, ɑ, ə, Ø], Arab. fem. [t]	ه, ة	ـه, ـة	هـ	ـهـ
y	[j, e, aj, i(ː)]	ي, ی	ـي, ـی	يـ	ـيـ

Table 4. The Kashmiri alphabet

LC Translit.	*Value*	*Isolated*	*Final*	*Initial*	*Medial*	*LC Translit*	*Value*	*Isolated*	*Final*	*Initial*	*Medial*
b	[b]	ب	ـب	بـ	ـبـ	k	[k]	ك	ـك	کـ	ـکـ
p	[p]	پ	ـپ	پـ	ـپـ	g	[g]	گ	ـگ	گـ	ـگـ
t	[t]	ت	ـت	تـ	ـتـ	l	[l]	ل	ـل	لـ	ـلـ
ṭ	[ʈ]	ٹ	ـٹ	ٹـ	ـٹـ	m	[m]	م	ـم	مـ	ـمـ
s̱	[s]	ث	ـث	ثـ	ـثـ	n	[n]	ن	ـن	نـ	ـنـ
j	[ʤ]	ج	ـج	جـ	ـجـ	v	[w]	و	ـو	-	-
c	[tʃ]	چ	ـچ	چـ	ـچـ	h	[h]	ھ	ه, ـه	هـ	ـهـ
ḥ	[h]	ح	ـح	حـ	ـحـ	y	[j]	ے	ـے	يـ	ـيـ
kh	[k^{h}]	خ	ـخ	خـ	ـخـ	a	[a]	آ	ـَ	اَ	ـَ
d	[d]	د	ـد	-	-	ā	[aː]	آ	ـا	آ	ـأ
ḍ	[ɖ]	ڈ	ـڈ	-	-	ạ	[ə]	أ	ء	أ	ء
ẕ	[z]	ذ	ـذ	-	-	ạ	[əː]	أ	أ	أ	أ
r	[r]	ر	ـر	-	-	i	[iː]	ا	ا	-	
ṛ	[ɽ]	ڑ	ـڑ	-	-	ī	[ɪ]	اى	ى	ايـ	ـيـ
z	[z]	ز	ـز	-	-	u’	[ɨː]	إ	-	إ	-
ts	[ts]	ژ	ـژ	-	-	ū’	[ɨ]	إ	-	إ	-
s	[s]	س	ـس	سـ	ـسـ	u	[uː]	ا	-	ا	-
ś	[ʃ]	ش	ـش	شـ	ـشـ	ū	[ʊ]	اڑ	ـڑ	اڑ	ـڈ
ṣ	[s]	ص	ـص	صـ	ـصـ	o	[oː]	او	ـاو	او	و
z̤	[z]	ض	ـض	ضـ	ـضـ	ō	[o]	او	ـو	او	و
ṯ	[t]	ط	ـط	طـ	ـطـ	ō	[ɔ]	اوٚ	ـاوٚ	اوٚ	وٚ
ẓ	[z]	ظ	ـظ	ظـ	ـظـ	ọ̄	[ɔː]	اوآ	ـوآ	-	-
ʿ	[ʔØ]	ع	ـع	عـ	ـعـ	e	[e]	ـےٚ	ـے	ايـ	ـيـ, ـيـ
gh	[g]	غ	ـغ	غـ	ـغـ	ē	[eː]	ـے	ـے	ايـ	ـيـ, ـيـ
f	[f, p^{h}]	ف	ـف	فـ	ـفـ	y	[^{j}]	ـےٕ	ى	-	ـۑـ ـۑ
q	[k]	ق	ـق	قـ	ـقـ						

Table 5. The Urdu alphabet

LC Transliteration	*Transcription*	*Isolated*	*Final*	*Initial*	*Medial*
-	/C_[a]; [ə]	ا	ـا	–	–
b	[b]	ب	ـب	بـ	ـبـ
p	[p]	پ	ـپ	پـ	ـپـ
t	[t]	ت	ـت	تـ	ـتـ
ṭ	[ʈ]	ٹ	ـٹ	ٹـ	ـٹـ
s̱	[s]	ث	ـث	ثـ	ـثـ
j	[dʒ]	ج	ـج	جـ	ـجـ
c	[tʃ]	چ	ـچ	چـ	ـچـ
ḥ	[h]	ح	ـح	حـ	ـحـ
k̲h̲	[x]	خ	ـخ	خـ	ـخـ
d	[d]	د	ـد	–	–
ḍ	[ɖ]	ڈ	ـڈ	–	–
ẕ	[z]	ذ	ـذ	–	–
r	[r]	ر	ـر	–	–
ṛ	[ɽ]	ڑ	ـڑ	–	–
z	[z]	ز	ـز	–	–
zh	[ʒ]	ژ	ـژ	–	–
s	[s]	س	ـس	سـ	ـسـ
sh	[ʃ]	ش	ـش	شـ	ـشـ
ṣ	[s]	ص	ـص	صـ	ـصـ
z̤	[z]	ض	ـض	ضـ	ـضـ
t̤	[t]	ط	ـط	طـ	ـطـ
ẓ	[z]	ظ	ـظ	ظـ	ـظـ
ʿ	/C_[ɑ]; [Ø, ʔ,ə]	ع	ـع	عـ	ـعـ
g̲h̲	[ɣ]	غ	ـغ	غـ	ـغـ
f	[f]	ف	ـف	فـ	ـفـ
q	[q]	ق	ـق	قـ	ـقـ
k	[k]	ك	ـك	کـ	ـکـ
g	[g]	گ	ـگ	گـ	ـگـ
l	[l]	ل	ـل	لـ	ـلـ
m	[m]	م	ـم	مـ	ـمـ
n	[n]; nasalization	ن	ـن	نـ	ـنـ
n̲	nasalization	ں	ـں	ٮـ	ـٮـ
v	[v, u, ʊ, o, ow]	و	ـو	–	–
h	/_# [ɑ]; [h, Ø]	ہ	ـہ	ھـ	ـھـ
t	[t] (Arabic feminine)	ة	ـة	–	–
y	[j, i, e, ɛ]	ی	ـی	یـ	ـیـ

Table 6. The Sindhi alphabet

LC Translit.	*Value*	*Isolated*	*Final*	*Initial*	*Medial*	*LC Translit*	*Value*	*Isolated*	*Final*	*Initial*	*Medial*
-	/C_[ɑ]	ا	ـا	-	-	z	[z]	ز	ـز	-	-
b	[b]	ب	ـب	بـ	ـبـ	s	[s]	س	ـس	سـ	ـسـ
ḅ	[ɓ]	ٻ	ـٻ	ٻـ	ـٻـ	sh	[ʃ]				
bh	[bʰ]	ڀ	ـڀ	ڀـ	ـڀـ	ṣ	[s]	ص	ـص	صـ	ـصـ
t	[t]	ت	ـت	تـ	ـتـ	ẓ	[z]	ض	ـض	ضـ	ـضـ
th	[tʰ]	ٿ	ـٿ	ٿـ	ـٿـ	ṭ	[t]	ط	ـط	طـ	ـطـ
ṭ	[ʈ]	ٽ	ـٽ	ٽـ	ـٽـ	ẓ	[z]	ظ	ـظ	ظـ	ـظـ
ṭh	[ʈʰ]					ʻ	[Ø]	ع	ـع	عـ	ـعـ
s̲	[tʰ]	ث	ـث	ثـ	ـثـ	gh	[ɡ]	غ	ـغ	غـ	ـغـ
p	[p]	پ	ـپ	پـ	ـپـ	f	[f]	ف	ـف	فـ	ـفـ
j	[ʤ]	ج	ـج	جـ	ـجـ	ph	[pʰ]	ڦ	ـڦ	ڦـ	ـڦـ
j	[ʄ]	ڄ	ـڄ	ڄـ	ـڄـ	q	[k]	ق	ـق	قـ	ـقـ
jh	[ʤʰ]	جھ	ـجھ	جھـ	ـجھـ	k	[k]	ڪ	ـڪ	ڪـ	ـڪـ
ñ	[ɲ]	ڃ	ـڃ	ڃـ	ـڃـ	kh	[kʰ]	ک	ـک	کـ	ـکـ
c	[tʃ]	چ	ـچ	چـ	ـچـ	g	[ɡ]	گ	ـگ	گـ	
ch	[tʃʰ]	ڇ	ـڇ	ڇـ	ـڇـ	g̈	[ɠ]	ڳ	ـڳ	ڳـ	ـڳـ
ḥ	[h]	ح	ـح	حـ	ـحـ	gh	[ɡʰ]	کھ	ـکھ	کھـ	ـکھـ
k̲h̲	[x]	خ	ـخ	خـ	ـخـ	n	[ġ]	ڱ	ـڱ	ڱـ	ـڱـ
d	[d]	د	ـد	-	-	l	[l]	ل	ـل	لـ	ـلـ
dh	[dʰ]	ڌ	ـڌ	-	-	m	[m]	م	ـم	مـ	ـمـ
ḍ	[ɗ]	ڏ	ـڏ	-	-	n	[n]	ن	ـن	نـ	ـنـ
ḍ	[ɖ]	ڊ	ـڊ	-	-	ṇ	[ɳ]	ڻ	ـڻ	ڻـ	ـڻـ
ḍh	[ɖʰ]	ڍ	ـڍ	-	-	v	[v, u, o]	و	ـو	-	-
z̲	[z]	ذ	ـذ	-	-	h	[h]	ھ, ه	ـه	هـ	ـهـ
r	[r]	ر	ـر	-	-	y	[j, i, e]	ي	ـي	يـ	ـيـ
ṛ	[ɽ]	ڙ	ـڙ	-	-						

Table 7. The Ottoman Turkish alphabet

LC Transliteration[a]	*Transcription*	*Isolated*	*Final*	*Initial*	*Medial*
-	[ɑ, Ø]; sometimes [o, e]	ا	ـا	–	–
b	[b, p]	ب	ـب	بـ	ـبـ
p	[p]	پ	ـپ	پـ	ـپـ
t	[t]	ت	ـت	تـ	ـتـ
s̲	[s]	ث	ـث	ثـ	ـثـ
c	[dʒ, tʃ]	ج	ـج	جـ	ـجـ
ç	[tʃ]	چ	ـچ	چـ	ـچـ
ḥ	[h]	ح	ـح	حـ	ـحـ
ḫ	[h]	خ	ـخ	خـ	ـخـ
d	[d, t]	د	ـد	–	–
ẕ	[z]	ذ	ـذ	–	–
r	[r]	ر	ـر	–	–
z	[z]	ز	ـز	–	–
j	[ʒ]	ژ	ـژ	–	–
s	[s]	س	ـس	سـ	ـسـ
ş	[ʃ]	ش	ـش	شـ	ـشـ
ṣ	[s]	ص	ـص	صـ	ـصـ
ż	[z, d]	ض	ـض	ضـ	ـضـ
ṭ	[t, d]	ط	ـط	طـ	ـطـ
ẓ	[z]	ظ	ـظ	ظـ	ـظـ
ʻ	[Ø, a]	ع	ـع	عـ	ـعـ
ġ	[ɣ, ɡ, k, h]	غ	ـغ	غـ	ـغـ
f	[f]	ف	ـف	فـ	ـفـ
ḳ	[k]	ق	ـق	قـ	ـقـ
k	[k, j]	ك	ـك	كـ	ـكـ
g	[ɡ, k]	گ	ـگ	گـ	ـگـ
ñ	[ɲ]	ڭ	ـڭ	-	ـڭـ
l	[l]	ل	ـل	لـ	ـلـ
m	[m]	م	ـم	مـ	ـمـ
n	[ɲ]	ن	ـن	نـ	ـنـ
v	[v, Ø, u, ɯ, i, y, o, œ]	و	ـو	-	-
h	[h, ɑ, i, e], Arab. fem. [t]	ه، ة	ـه، ـة	هـ	ـهـ
y	[j, i, y, e, ej, ɑ, u, ɯ]	ى	ـى	يـ	ـيـ

Table 8. The Uyghur alphabet

LC Transliteration	*Transcription*	*Isolated*	*Final*	*Initial*	*Medial*
a	[ɑ]	ا	ـا	-	-
ă	[ə]	ه	ـه	-	-
b	[b]	ب	ـب	بـ	ـبـ
p	[p]	پ	ـپ	پـ	ـپـ
t	[t]	ت	ـت	تـ	ـتـ
j	[dʒ]	ج	ـج	جـ	ـجـ
ch	[tʃ]	چ	ـچ	چـ	ـچـ
kh	[x]	خ	ـخ	خـ	ـخـ
d	[d]	د	ـد	-	-
r	[r]	ر	ـر	-	-
z	[z]	ز	ـز	-	-
zh	[ʒ]	ژ	ـژ	-	-
s	[s]	س	ـس	سـ	ـسـ
sh	[ʃ]	ش	ـش	شـ	ـشـ
gh	[ɣ]	غ	ـغ	غـ	ـغـ
f	[f]	ف	ـف	فـ	ـفـ
q	[q]	ق	ـق	قـ	ـقـ
k	[k]	ك	ـك	كـ	ـكـ
g	[ɡ]	گ	ـگ	گـ	
ng	[ŋ, n]	ڭ	ـڭ	ڭـ	ـڭـ
l	[l]	ل	ـل	لـ	ـلـ
m	[m]	م	ـم	مـ	ـمـ
n	[n]	ن	ـن	نـ	ـنـ
h	[h]	ھ	ـھ	ھـ	ـھـ
o	[o]	وٝ	ـوٝ	-	-
u	[u]	و	ـو	-	-
ö	[ø]	ۆٖ	ـۆٖ	-	-
ü	[y]	و	ـو	-	-
v	[v]	ۋ	ـۋ	-	-
e	[e]	ې	ـې	ېـ	ـېـ
i	[i]	ى	ـى	ىـ	ـىـ
y	[j]	ي	ـي	يـ	ـيـ
ʾ				ئـ	ـئـ
la	[la]	لا	ـلا	-	-

Table 9. The Malay (Jawi) alphabet

LC Transliteration	*Transcription*	*Isolated*	*Final*	*Initial*	*Medial*
-	[Ø]	ا	ـا	-	-
b	[b]	ب	ـب	بـ	ـبـ
t	[t]	ت	ـت	تـ	ـتـ
th	[s]	ث	ـث	ثـ	ـثـ
j	[dʒ]	ج	ـج	جـ	ـجـ
ch	[tʃ]	چ	ـچ	چـ	ـچـ
ḥ	[h]	ح	ـح	حـ	ـحـ
kh	[x, k]	خ	ـخ	خـ	ـخـ
d	[d]	د	ـد	-	-
dz	[dz]	ذ	ـذ	-	-
r	[r]	ر	ـر	-	-
z	[z, dʒ]	ز	ـز	-	-
s	[s]	س	ـس	سـ	ـسـ
sh	[ʃ, s]	ش	ـش	شـ	ـشـ
ṣ	[s]	ص	ـص	صـ	ـصـ
ḍ	[z, dʒ]	ض	ـض	ضـ	ـضـ
ṭ	[t]	ط	ـط	طـ	ـطـ
ẓ	[z]	ظ	ـظ	ظـ	ـظـ
ʻ	[Ø]	ع	ـع	عـ	ـعـ
gh	[ɣ, r]	غ	ـغ	غـ	ـغـ
ng	[ŋ]	ڠ	ـڠ	ڠـ	ـڠـ
f	[f, p]	ف	ـف	فـ	ـفـ
p	[p]	ڤ	ـڤ	ڤـ	ـڤـ
ḳ	[k]	ق	ـق	قـ	ـقـ
k	[k]	ك	ـك	كـ	ـكـ
g	[g]	ݢ	ـݢ	ݢـ	ـݢـ
l	[l]	ل	ـل	لـ	ـلـ
m	[m]	م	ـم	مـ	ـمـ
n	[n]	ن	ـن	نـ	ـنـ
w	[w]	و	ـو	-	-
h	[h]	ه	ـه	هـ	ـهـ
la	[la]	لا	ـلا	-	-
ʼ	[Ø]	ء	ء	-	ـئـ
y	[j]	ي	ـي	يـ	ـيـ
ny	[ɲ]	ڽ	ـڽ	ڽـ	ـڽـ

Table 10. The Hausa alphabet

Hausa name	*isolated*	*initial*	*medial*	*final*	*Boko equivalents*
alif	ا	–	–	ـا	.
ba	ب	بـ	ـبـ	ـب	b. ɓ
ta	ت	تـ	ـتـ	ـت	t
ca	ث	ثـ	ـثـ	ـث	c. ts
jim	ج	جـ	ـجـ	ـج	j
ha	ح	حـ	ـحـ	ـح	h
ha mai-ruwa	خ	خـ	ـخـ	ـخ	h
dal	د	–	–	ـد	d
zal	ذ	–	–	ـذ	z
ra	ر	–	–	ـر	r
zaira	ز	–	–	ـز	z
tsa mai-hannu	ط	طـ	ـطـ	ـط	ts. ƭ
zadi	ظ	ظـ	ـظـ	ـظ	z
kaf lāsan	ك	كـ	ـكـ	ـك	k
lam	ل	لـ	ـلـ	ـل	l
mim	م	مـ	ـمـ	ـم	m
nun	ن	نـ	ـنـ	ـن	n
sodi	ص	صـ	ـصـ	ـص	ṣ
lodi	ض	ضـ	ـضـ	ـض	l
ain	ع	عـ	ـعـ	ـع	.
angal	غ	غـ	ـغـ	ـغ	g
fa	ف	فـ	ـفـ	ـف	f
kaf wati	ق	قـ	ـقـ	ـق	k
sin	س	سـ	ـسـ	ـس	ṣ
shin	ش	شـ	ـشـ	ـش	š
ha kuri	ه	هـ	ـهـ	ـه	h
Semi-vowels					
wau	و	–	–	ـو	w
ya	ي	يـ	ـيـ	ـي	yʾ, ʼy

Table 11. The Swahili alphabet

Swahili Name	*Isolated*	*Final*	*Medial*	*Initial*	*Roman*	*IPA*
alifu	ا	ـا	–	–		
be	ب	ـب	ـبـ	بـ	B	[p], [pʰ], [ɓ], [ᵐb]
te	ت	ـت	ـتـ	تـ	T	[t], [tʰ]
te kiti						
the	ث	ـث	ـثـ	ثـ	Th	[θ] (only in Arabic loanwords)
jimu	ج	ـج	ـجـ	جـ	J	[ɟ], [ⁿɟ], sometimes [ʄ]
hhe	ح	ـح	ـحـ	حـ	H	[h]
he ngoke						
khe	خ	ـخ	ـخـ	خـ	Kh	[h] (only in Arabic loanwords)
dali	د	ـد	–	–	D	[ɗ], [ⁿd]
dhali	ذ	ـذ	–	–	Dh	[ð] (only in Arabic loanwords)
re	ر	ـر	–	–	R	[r]
re kiti						
re kusuka	ر	ـر	–	–		
ze	ز	ـز	–	–	Z	[z]
sini	س	ـس	ـسـ	سـ	S	[s]
shini	ش	ـش	ـشـ	شـ	Sh	[ʃ], [tʃ]
swadi	ص	ـص	ـصـ	صـ	S, Sw	[s], [sʷ]
dhadi	ض	ـض	ـضـ	ضـ	Dh	[ɟ]
dhe maʾarufu						
twe	ط	ـط	ـطـ	طـ	T, Tw	[t], [tʷ]
dhwe	ظ	ـظ	ـظـ	ظـ	D, Dw, Z	[ɟ]
aini	ع	ـع	ـعـ	عـ	G	[ʔ],[ø]
ghaini	غ	ـغ	ـغـ	غـ	Gh	[ʄ], [ŋ], [ŋ$_g$], [ɣ] ([ɣ] only in Arabic loanwords)
fe	ف	ـف	ـفـ	فـ	F	[f], [v]
kafu	ق	ـق	ـقـ	قـ	Q, K	[k]
kyafu	ک ك	ـك	ـکـ	کـ	K	[k], [kʰ]
lamu	ل	ـل	ـلـ	لـ	L	[l]
mimu	م	ـم	ـمـ	مـ	M	[m]
nuni	ن	ـن	ـنـ	نـ	N	[n]
he	ه	ـه	ـﻬـ ـﮭـ	ﻫ هـ	H	[h]
he mduara						
wau	و	ـو	–	–	W	[w], [oː], [uː]
ye	ي	ـي	ـيـ	يـ	Y	[y], [iː], [e], [ny]

Extra letters sometimes used in Swahili but not in Arabic

pe	پ	ـپ	ـپـ	پـ	P	[p], [pʰ]
che	څ چ	ـچ ـڅ	ـچـ ـڅـ	څـ چـ	Ch	[tʃ], [tʃʰ]
ge	ڠ	ـڠ	ـڨـ	ڠـ	G	[g], [ʄᵑ]
ve	ڤ	ـڤ	ـڤـ	ڤـ	V	[v]

Arabic Alphabet: Origin

The Arabic alphabet, or more precisely *ʾabjad* ‘consonantary’ takes its origin from the Nabataean variant of late Aramaic script, which suits Semitic morphology based on the tri-consonantal root, but records neither short vowels nor most inflectional endings (Daniels 1990:730). In the process of adoption, the letters were graphically homogenized, and subsequently a variety of mostly supralinear signs were devised to optimize the phonetic precision of the script.

The Arabic alphabet most often denotes the formal variant within the Arabic languages (Classical Arabic, Modern Standard Arabic), though before Islam, in Islamic early private writings on papyri (Hopkins 1984:xliv–xlvii), in → Judaeo-Arabic (Blau 2002), and increasingly from the 5th/11th century onwards informal texts were recorded in writing. Since the script omits the inflectional endings, and the orthography of the glottal stop (*hamza*) as an additional sign is treated loosely in manuscripts, the graphic image partially veils unintended nonstandard forms.

Writing was not immediately accepted by Arab scholarship, which first evolved in the transmitting of Prophetic traditions, and book learners were denigrated as *ṣuḥufiyyūn* (Rosenthal 1947:6–18; Schoeler 2002:40, 120–121; al-Xaṭīb al-Baġdādī, *Taqyīd* 29–63). Until the 3rd/9th century at least, writing therefore coexisted with the aural transmission of knowledge ‘heard’ from a teacher (*samāʿ*), even if the claim of such personal contact was occasionally more fiction than fact.

The Arabic alphabet has been appropriated for numerous mostly non-Semitic Islamic languages in the past and present. Today the most important of these are Berber, Persian, Pashto, Kurdish, Urdu, Sindhi, Kashmiri, and Uyghur for which additional signs were introduced. After the Roman alphabet, Arabic is the second most frequent segmental script in the world (→ Alphabet: Use for other language).

In art and architecture, Arabic letters give the Arabic language permanence and ubiquity far beyond the realm of its spoken use; they pervade the entire area of Islam where they constitute “symbols of a true politico-religious unity” (Sourdel-Thomine 1978:1114). Familiar in Europe since the Crusades from imported luxury objects, they were occasionally reproduced (correctly or incorrectly) on Christian and Jewish monuments and artifacts for decorative purposes.

1. Before the Arabic ʾabjad

Long before Arabs had a script of their own they left traces in other writings of the Ancient Near East. Tribal and royal names survive in Neo-Assyrian historical sources about their dealings with (often female) Bedouin tribal chiefs (Ephʿal 1982; → Arabs). Ancient South Arabian script denotes South Semitic languages different from Old → North Arabian, the ancestor of modern Arabic. However, since the latter half of the 1st millennium B.C.E., the northern Arabs used the Old South Arabian alphabet in five variants mainly in the northern Arabian Peninsula and in Jordan. These are the alphabets of Thamūdic (6th century B.C.E. to 4th century C.E. in west and central northern Arabia), Dedānic and Liḥyānic (5th century B.C.E. to 1st century C.E. in northwest Arabia), Ṣafaitic (1st century B.C.E. to 3rd century C.E. in the Ṣafāʾ stone desert southeast of Damascus), and Ḥasaean (also called Ḥasaitic; 5th to 2nd century B.C.E. in the northeast of southern Arabia on the Persian Gulf; Müller 1982). Their order of letters, identical with the South Arabian order beginning with <h l ḥ m> might be as old as the now predominant Canaanite-Phoenician order beginning <ʾ b g d> (Dietrich and Loretz 1988:289, 294) and survives today in the Ethiopic alphabet and perhaps in the etymology of Latin *elementum* pl. *elementa* ‘letters, alphabet; beginnings, primary causes’, which word according to Müller (1994–1996) reflects the beginning of the Ancient South Arabian order of the first letters of the alphabet <*h-l-ḥ-m*>. These Old North Arabian scripts preserve mostly short texts in languages that still differed from Classical Arabic, such as their use of the article *ha-* as opposed to Classical Arabic *al-*. Arabs also sporadically used Imperial Aramaic as early as the 5th century B.C.E.

Four centuries later, the Arab satellite states of the Roman, Seleucid, and Parthian Empires developed their own varieties of Aramaic script (Nabataean, Palmyrenian, Syriac, and Hatran).

Of these, the Nabataean script (attested 100 B.C.E.–350 C.E.) was the genetic ancestor of the current Arabic alphabet (Cantineau 1930–1932; Healey 1993:49–63). This had been asserted by Theodor Nöldeke as early as 1865 and supported by Adolf Grohmann, based on the development of individual graphemes. Thereafter, a competing explanation with Syriac as an ancestor was proposed by Jean Starcky (1966: 932–34) based on the script's *ductus* (general shape and formation of letters and their combinations), and claims of this kind can still be found today. The calligraphic influence of Syriac script on Arabic, once it had come into being, may be argued profitably (Abbott 1939:19–20; Briquel-Chatonnet 1997:143–44; Healey 1990–1991:41–43). But the epigraphic data leave no doubt that the shapes of the letters and the specific set of homographs can only be derived from a Nabataean provenance (Grohmann 1976–71:2.13, 17–21) [See plates I a–b]. Detailed documentations for each grapheme based on dated inscriptions and papyri are now available (Healey 1990–1991:44–45 and tables; Gruendler 1993:123–30 and charts) [See plates II a–b].

The script of the Nabataeans continued to be used after their defeat by the Romans (106 C.E.) by Jews and Arabs in Syro-Palestine until the 4th century C.E. In two Nabataean inscriptions by Arabs (ʿEn Avdat, between 88–89 and 125–126 C.E.; an-Namāra, 328 C.E.) Arabic language is expressed through Nabataean characters, whereas others (e.g. ʾUmm al-Jimāl, c. 250 C.E.; el-ʿUla, 267 C.E.) show an admixture of Arabic (Negev 1986:48; Versteegh 1997:30–36 with further bibliography).

The move from Nabataean to Arabic was complex, for the Nabataean script combined epigraphic, formal, and free cursive variants, which developed at different speeds. Around the turn of the 1st century C.E. the formal cursive of the Engaddi papyri from Naḥal Ḥever (Starcky 1954: 162 and pls. 1–3; Yardeni a.o. 2002:1.169–256 and pls. 15–26 on P. Yadin [5/6 Ḥev] 1–4 all dated within 93–99 C.E.) and the free, or extreme, cursive found in the same group of papyri (Yardeni a.o. 2002:1.257–76 and pls. 55–56 on P. Yadin [5/6 Ḥev] 6 and 9 dated 119–120 and 122–123 C.E. respectively) and the Nessana ostraca (Rosenthal 1962:200, emended by Naveh 1979: 111, n. 4) already show shapes the epigraphic script arrives at two centuries later. The Nabataean cursive thus served as the model for the Arabic script [See plates III a–b]. Nonetheless, few cursive documents have been preserved, and supplementary evidence must be gleaned from late epigraphic Nabataean (Naveh 1982:156; Healey 1990–1991:43–44, 50–52, with further bibliography).

2. The formation of the Arabic ʾabjad

The Arabic *ʾabjad* first materialized in five brief pre-Islamic inscriptions from Syria and northwest Arabia. They display a clearly Arabic *ductus*, though they are linguistically controversial and graphically heterogeneous. Except for the graffito on a Nabataean sanctuary in Jabal Ramm near ʿAqaba, datable to the first half of the 4th century C.E., all inscriptions belong to the 6th century C.E. They include a trilingual inscription in Greek, Syriac, and Arabic on a Christian martyrion in Zabad southeast of Aleppo (512 C.E.), a historical inscription in Jabal ʾUsays (Sēs) on the Syrian-Roman border approximately 100 kilometers southeast of Damascus (528 C.E.), a graffito in the double church of ʾUmm al-Jimāl southwest of Boṣrā (ca. 6th century C.E.), and a Greek and Arabic bilingual text on a martyrion in Ḥarrān in the Lejāʾ (586 C.E.) (Grohmann 1967–1971:2.14–15; Gruendler 1993:13–14).

For paleographic studies up to the 2nd/8th century, the most reliable conclusions can be drawn by limiting the investigation to the preserved dated specimens listed above, even if these are extremely sparse for the pre-Islamic period. Arabic script has been preserved on stone and objects of craft such as mosaics, metal objects, glass weights, earthenware, coins, and cloth. Cursive script mostly survives on papyrus and parchment and, since the 2nd/8th century, on rag paper. Earlier writing materials were stripped stalks of palm branches, wood (both already attested for the South Arabian cursive; see Ryckmans a.o. 1994), and shoulder blades of camels. Medieval accounts about the formative period of the Arabic script derive from later centuries and cannot be taken at face value. Abbott (1939:3–12) has attempted to reconcile the alleged invention of the script in al-ʾAnbār or al-Ḥīra in Iraq, whereas most epigraphic remains hail from Syria. More plausible is that Muslim scholars tried to locate the invention of their script in places where they knew pre-Islamic Arabic expertise of writing to have come into contact with Aramaic-Hellenistic culture, such as in the figure of the poet and scribe ʿAdī ibn Zayd (d. ca. 590–600 C.E.) (Endress 1982:169–170).

The configuration of Arabic script can be characterized by five trends, the antecedents of which can already be observed in Aramaic: (1) in the 4th century B.C.E. positional variants, or allographs, emerged in the form of larger final letters in the Aramaic cursive; (2) in the 1st century C.E. letters became fully connected in cursive Nabataean; at the same time (3) bars of letters were integrated into continuous strokes, and formerly distinct letters merged (*bēṯ/nūn*, *gīmel/ḥēṯ*, *zayin/rēš*, *yōḏ /tāw*, *pēh/qōp̄*) in the cursive of Naḥal Ḥever; (4) in the free Nabataean cursive, the ‘ceiling-line’ limiting the height of most letters was replaced by a baseline (graffiti only show this in the 3rd century C.E.); finally (5) the Namāra inscription (328 C.E.) shows the *lām-ʾalif* ligature for the first time. (Here and in the following a letter’s shape is referred to by its name, e.g. Nabataean *zayin* or Arabic *zāy*, and a letter’s sound by its phonetic symbol, e.g. *z*). The mergers listed under point (3) are specific to Nabataean and only they explain the Arabic homographs *jīm/ḥāʾ*, *rāʾ/zāy*, non-final *bāʾ/nūn*, *yāʾ/tāʾ*, and non-final *fāʾ/qāf*. By themselves, these homographs preclude a provenance from Syriac where all of the graphemes *gāmal̠*, *ḥēṯ*, *rēš*, *zayn*, *bēṯ*, *nūn*, *yūḏ*, *taw*, *pē* and *qōp̄* remain distinct. Among the three varieties of Nabataean, the free cursive most closely approximates pre-Islamic Arabic: straight (Nabataean:) *ʾalep̄* / (Arabic:) *ʾalif*, short hooked *tāw/tāʾ*, three parallel teeth for *šīn/šīn*, integrated *ṭēṯ/ṭāʾ*, hooked *ʿayin/ʿayn*, a closed loop without stem for *pēh/fāʾ*, rounded *mēm/mīm*, looped *hēh/hāʾ*, lowered curved *waw/wāw*, and s-shaped left-turning final *yōḏ/yāʾ*. At the present state of paleographic evidence, the emergence of the Arabic *ʾabjad* must be surmised in the late 2nd or 3rd century C.E., between the latest cursive Nabataean and the earliest attested Arabic script.

In the Arabic *ʾabjad*, the first two formative trends (positional variants and connections between letters) were integrated into a system, each shape being made to correspond to a specific (initial, medial, final, or isolated) position – excepting the six letters *ʾalif*, *dāl/ḏāl*, *rāʾ/zāy*, and *wāw* which developed no connection to the left. In addition to the above-mentioned mergers (third trend), homographs had already been imported to Arabic from the West Semitic *ʾabgad*, because its reduced inventory of 22 Phoenician sounds only required this amount of graphemes. Conversely, in Arabic all Proto-Semitic sounds except s^2 (also *ś*) had remained distinct and needed to be expressed by an extant graphic shape. This explains the presence of multiple letters doing double duty. The Nabataean letters *tāw*, *ḥēṯ*, *dāleṯ*, *ṣādeh*, *ṭēṯ* and *ʿayin* thus served secondarily to indicate the sounds *ṯ*, *x*, *ḏ*, *ḍ*, *ḏ̣*, and *ġ*, and Nabataean *šīn* denoted both Arabic *s* and *š*. Together, the Phoenician sound mergers and the Nabataean graphic mergers led the Arabic graphemes to shrink to 18 (*ʾalif*, *bāʾ*, *jīm*, *dāl*, *rāʾ*, *sīn*, *ṣād*, *ṭāʾ*, *ʿayn*, *fāʾ*, *qāf*, *kāf*, *lām*, *mīm*, *nūn*, *hāʾ*, *wāw*, *yāʾ*), or 15 in non-final position (with *bāʾ/nūn/yāʾ* and *fāʾ/qāf* becoming identical) expressing a sum of 28 sounds [See plate IV]. This homogeneity would be further enhanced in Arabic calligraphy, but it hampered the legibility of difficult texts. No capitals developed in Arabic script, where varieties of one letter depend strictly on its relative position within the word.

The alphabetic order is based essentially on the Aramaic order with the new homographs distinguished by diacritics added after their respective model *ʾalif*, *bāʾ*, *tāʾ*, *ṯāʾ*, *jīm*, *ḥāʾ*, *xāʾ*, *dāl*, *ḏāl*, *rāʾ*, *zāy*, *sīn*, *šīn*, *ṣād*, *ḍād*, *ṭāʾ*, *ḏ̣āʾ*, *ʿayn*, *ġayn*, *fāʾ*, *qāf*, *kāf*, *lām*, *mīm*, *nūn*, *hāʾ*, *wāw*, *yāʾ*. This order was amply used as an ordering principle for medieval language dictionaries (mostly by a word’s last letter), poets’ collected works, or *dīwān*s (by a poem’s final rhyming letter, or *qāfiya*), and biographical dictionaries (by first letter of given name, or *ism*). The Aramaic order survives in the numerical use of the letters which continued to be applied in the sciences *ʾalif*, *bāʾ*, *jīm*, *dāl*, *hāʾ*, *wāw*, *zāy*, *ḥāʾ*, *ṭāʾ* (= 1–9); *yāʾ*, *kāf*, *lām*, *mīm*, *nūn*, *sīn*, *ʿayn*, *fāʾ*, *ṣād* (= 10–90); *qāf*, *rāʾ*, *šīn*, *tāʾ*, *ṯāʾ*, *xāʾ*, *ḏāl*, *ḍād*, *ḏ̣āʾ* (= 100–900); *ġayn* = 1000). A third order is that of letter frequency, which was used for cryptography (see, e.g., ʾIsḥāq ibn ʾIbrāhīm, *Burhān* 354).

3. The Arabic ʾabjad in Islam

For early Islamic → paleography, the literary accounts still prove problematic, as they lack full descriptions and were composed long after the scribal practices they discuss. An exception is Ibn an-Nadīm's definition of one letter (*ʾalif*) in the early Meccan script, allowing its identification in actual specimens, now referred to as *ḥijāzī* script (Abbott 1939:18–19 and pls. 8–13). The terms *māʾil* and *mašq*, often understood as scripts today, may not have meant that originally (Déroche 1980:213–21). Nonetheless, scholars have ventured to identify scripts listed in the sources: J.G.C. Adler first applied the term → 'Kūfic' in 1780 to Qurʾānic material, Josef von Karabacek, *māʿil* and *ʿirāqī* (Déroche 1980: 209–12). Others identified *badīʿ* (Schroeder 1937:234–48), *ġubār* (Abbott 1939:37–38), *musalsal* (Abbott 1941:98–99), *jalīl* (Grohmann 1952:75–77), *ṯuluṯ rayḥān* (ibid. 81), and *qarmaṭa* (Dietrich 1955:46, 67). Some medieval terms grew too vague, so the *kūfī* of early *Qurʾān*s has been replaced by six groups of 'Abbāsid scripts' by Déroche (1992:34–47), and the use of the term *nasx* for early papyri is discouraged as anachronistic by Khan (1992: 45–46; cf. emendations in Diem 1993). However, irrespective of the often dubious factual accuracy for the early period, the rich literary sources underscore the importance Arabic–Islamic culture placed upon the history of its script and its artful execution.

Groups of dated or datable specimens provide a more reliable basis for early paleographic study. Even so, this research remains in a preliminary state with a vast amount of yet-uncharted material. During the first three centuries of Islam, scripts diverged between (more or less homogeneous) groups of texts with distinct functions: memorial and votive inscriptions, *Qurʾān*s, papyrus documents and letters, and scholarly and literary manuscripts. The application to all of them of one script terminology, derived from later secretarial manuals is problematic, and some scholars prefer a careful analysis of all (or a significant sample) of a script's graphemes in order to build a typology, although the conclusions drawn from small collections remain limited (Flury 1920:8–21; Déroche 1980:213). Different concepts have been introduced to grasp the level of execution in a piece of writing. For example, a cluster of scripts can be viewed as a circle, with the example closest to the 'ideal' at its center, and the loosest reproduction at the periphery (Déroche 1992:16). Similarly, Noam Chomsky's syntactic notion of competence vs. performance serves to distinguish a writer's ideal form, 'competence', from the actual result, 'performance' (Khan 1992:39, n. 53).

Arabic script thrived as the medium for recording the *Qurʾān* and as the official script of the Umayyad caliphate from ʿAbd al-Malik's reform of 78/697. It spread from Upper Egypt to Sogdiana within a century (Gruendler 1993:28, 167). By the end of the 1st century A.H., four groups of scripts had already emerged: (1) an angular epigraphic script, first attested in a clumsily carved Egyptian tombstone (31/652), reached a regular *ductus* in milestone inscriptions (64–86/685–705) and the mosaic band and copper plate of the Dome of the Rock (both 71/691). A rounded cursive, first attested in a requisition of sheep on papyrus (22/643), diversified into a (2) routined, ligatured protocol script, (3a) a wide-spaced slender chancellery hand, preserved in the gubernatorial correspondence of Qurra ibn Šarīk (r. 90–96/709–714) from Qōm ʾIšqāwh, including (3b) a denser and squatter variant for bilingual tax notifications, and (4) a slanting script of *Qurʾān* fragments, now referred to as *ḥijāzī* (Gruendler 1993: 131–141) [See plates V and VI a–b].

Diacritical marks (*ʾiʿjām*, *naqṭ*) were possibly inspired by pre-Islamic Nabataean or Syriac examples (Endress 1982:1.175, n. 82). They appear as a complete system, though used selectively on the earliest dated documents, the said requisition and a building inscription on a dam of Muʿāwiya (58/677), as well as *Qurʾān*s in *ḥijāzī* script datable to the 1st/7th century. During the two following centuries, diacritics were increasingly used in *Qurʾān*s and literary texts, but less so in private and business documents, which form the bulk of early papyri (Grohmann 1952:83, n. 289; Abbott 1957–1972:3, documents 4–7; Ferrando 2001:76–77 to be corrected accordingly). The literary discussion over their usefulness and social acceptability continued through the 3rd/9th century (Rosenthal 1948:17, 26; Rāġib 1990:16). Small diacritic strokes predominated in *Qurʾān*s, diacritic points in papyri and manuscripts. They marked either all meanings of a homograph (<*bāʾ/tāʾ/ṯāʾ/nūn/yāʾ*>, <*ǧīm/ḥāʾ/xāʾ*>, <*fāʾ/qāf*>) or only one of a pair (<*dāl/ḏāl*>, <*ṭāʾ/ḏ̣āʾ*>, <*ʿayn/ġayn*>, <*sīn/šīn*>). Initially, only *qāf* was distinguished by a dot above (or below) in the 2nd/8th century, and *fāʾ* received the respective opposite diacritic later. This distribution was kept in *maġribī* script, whereas a single dot on *fāʾ* and a double dot on *qāf* spread in the East in the 3rd/9th century. In its early history, the system was still fluid, as visible in changing diacritics; a *Qurʾān* manuscript (Paris Ar. 376 b) distinguishes *zāy* from *rāʾ* and *ʿayn* from *ġayn* by a dot beneath and *sīn* from *šīn* by three dots beneath. In the 2nd/8th century, the feminine ending written in pausal form as *hāʾ* received two dots, forming the *tāʾ marbūṭa*, and a century later, a miniature *kāf* was placed inside the final *kāf* to prevent confusion with final *lām*. In a reverse manner of distinction, so-called *muhmal* signs were devised to indicate unmarked letters in the form of dots, tilted small *lā*, or miniatures of the letters themselves.

The use of vowel signs in the mid-2nd/8th century can be deduced from both the theological dispute about them in the written sources and extant contemporary *Qurʾān* fragments (Abbott 1939, nos. 9–13, 15). Most Medieval Arabic sources ascribe the invention of Qurʾānic vocalization to ʾAbū l-ʾAswad ad-Duʾalī (d. 69/688) or his disciple Naṣr ibn ʿĀṣim (d. 89/707), but trace the impulse back to an Umayyad governor, whether Ziyād ibn ʾAbīhi (r. 45–53/665–73) or al-Ḥajjāj (r. 75–95/694–714). In *Qurʾān*s an (often red) colored dot above a consonant indicates the following short vowel *a* (*fatḥ*), beneath it *i* (*kasr*), at the letter's base *u* (*ḍamm*), and a double dot in these positions signifies indeterminacy (*tanwīn*). Further orthographic signs – an inverted half-circle or hook for a double consonant and a line above *ʾalif* for its zero-value – were ascribed to al-Xalīl ibn ʾAḥmad (d. ca. 160–75/776–91) though attested only in the 3rd/9th century. The glottal stop (*hamza*), omitted in the consonantal text (unless an otiose *ʾalif* had been kept or a glide had replaced it) was reinstated as a supplemental sign to an existing letter (*ʾalif*, *wāw*, or *yāʾ*) or placed on the line. The marker was variously a colored dot, a semicircle, or a miniature *ʿayn*. In the same century, papyri began to display the use of short strokes for the vowels *a* and *i*, a small *wāw* for *u*, and a double stroke (or a double *wāw*) for indeterminacy. Further miniature letters were introduced: a small *mīm* derived from *jazm* 'apocopation' for vowellessness (*sukūn*), a small *šīn* from *šadd* or *tašdīd* 'strengthening' for a double consonant, a small *ṣād* derived from *waṣl* or *ṣila* 'connection' for *ʾalif* with zero-value, and small *mīm-dāl* derived from *madd* 'extension' for the word-initial *ʾā* or the word-final *āʾ*. These orthographical signs became fully used a century later in *Qurʾān*s and difficult-to-read texts. No punctuation developed, except markers of *Qurʾān* verses (dots, strokes, circles, or rosettes). Nonetheless, modern editors have introduced Western → punctuation in reverse direction. Occasionally, medieval texts other than *Qurʾān*s were subdivided into sections by dotted circles or extended words (*mašq*).

For the period from the 4th/10th century onwards, medieval sources yield ample information on contemporary practices of script and penmanship, among these notably Ibn an-Nadīm's (d. 380/990) *Fihrist* (4–12), and the extensive treatment by al-Qalqašandī (d. 821/1418) (*Ṣubḥ* II, 440–88, III, 1–226/II2, 440–88, III2, 1–222; cf. Endress 1982:190–91; Gacek 1987:129–30). Later, Mamluk secretarial manuals in particular described and even illustrated chancellery scripts (partly used for calligraphy). By the 7th/13th century, five or more frequently six scripts (later called *al-ʾaqlām as-sitta*) had established themselves in chancellery and in popular practice. They fell into a 'moist' (*muraṭṭab*) subgroup, emphasizing the curvilinear elements and consisting of → *ṯuluṯ*, *tawqīʿ*, and *riqāʿ*, and a 'dry' (*yābis*) subgroup tending towards the rectilinear and including *muḥaqqaq*, *rayḥān*, and → *nasx*. Scripts were further classified by size, the extremes being gigantic *ṭumār* and tiny *ġubār* for pigeon posts, or by the presence of serifs (*tarwīs*) or closed loops (*ṭams*) (Gacek 1989:144–45). The literary sources also recorded pioneering calligraphers: Ibn Muqla (d. 328/939), who codified *nasx*, elevating it to a Qurʾānic script; Ibn al-Bawwāb (d. 423/1032) who further refined it; and Yāqūt al-Mustaʿṣimī (d. ca. 697/1298), who invented a new way of trimming the *qalam* and excelled in the six scripts. Ibn al-Bawwāb leaves behind the first *Qurʾān* in *nasx*, dated 391/1000–1001 and Yāqūt's name appears on several (partly forged) *Qurʾān*s.

4. ORTHOGRAPHY

Similar to the Arabic graphemes, the earliest orthographic conventions take their origin in the late Aramaic script (Diem 1976, 1983:395–401). The non-denotation of short vowels, but the denotation of long *ū* and *ī* by the glides *waw* and *yōd*, and the phonetic writing of the *t*-infix in the verbal Form VIII take their beginning here. Arabic names and words that contained letters which had merged in Aramaic were not denoted with the phonetically closest sound, but with their etymological cognate in Imperial Aramaic or Nabataean, e.g. *naḏara* 'to look' was written as <nṭr> according to Aramaic *nṭar* 'to guard' (Diem 1980:73) or *xalāṣun* 'clearance' by <ḥlṣ> (Yardeni a.o. 2002:28–30). Frequent letter correspondences were then generalized within words that lacked Aramaic cognates. This was a straightforward procedure as most writers of Arabic before Islam were also somewhat familiar with Aramaic. Further traces of Aramaic orthography are the otiose nominative *wāw* in the name *ʿAmr*<w> and the open writing of *tāʾ* in *ḏāt*.

In the late Nabataean stage, the final *-ā* came to be rendered by *yāʾ* when this was justified by paradigmatic derivation. The feminine ending *-atun* was recorded in its pausal form *-ah* by the letter *hāʾ* such as <ʾmh> instead of <ʾmt> for *ʾamatun* 'maidservant'. *Ḏāt*, which never occurred in pausal position, retained the open final *tāʾ* unchanged.

In its *ḥijāzī* stage, the orthography reflected a dialect in which the glottal stop had been lost and the long final *ī* had been shortened to *-i*. This is reflected in the reinterpretation of the now otiose internal *ʾalif* as *ā*, beginning with historical spellings, such as <rʾs> 'head', read as *rās*, and then expanded to words for which no etymological *ʾalif* had previously existed, such as *nās* 'people'. The long final *ī* was similarly not denoted, such as in Qurʾānic <dyn> for *dīni* 'my religion'.

When this dialect orthography was used for recording the *Qurʾān* and the → poetic koine, which were closely related, adjustments had to be made. The Qurʾānic orthography of case inflection, such as in nominative <ʾbʾwhm> for *ʾābāʾuhum* versus genitive <ʾbʾyhm> for *ʾābāʾihim* visible in the glides *wāw* and *yāʾ* preceding the suffix pronoun, may be dated either to the *ḥijāzī* or the subsequent Classical period according to Diem (1983:396). Besides the systematization of *ʾalif* for *ā* in all positions of a word (except a few cases, such as *hāḏā* <hḏʾ> or *allāh* <ʾlh>) and the generalization of the article *al-* in the relative pronoun (except in the singular and masculine plural), the most substantial adaptation concerned the orthography of the glottal stop (*hamza*). The glides *wāw* and *yāʾ* were reinterpreted as carriers for the supralinear *hamza* or, when it had dropped out without substitution, it was placed directly on the line.

5. LEGIBILITY AND AMBIGUITY

Arabic script, unlike many other alphabetic scripts, retains a very high phonetic accuracy when it is fully vocalized. Nonetheless, a paradigmatic-etymological counter trend of writing is visible in the non-assimilation of the article in the script, and the segmental writing of assimilated verbal endings and suffixes.

In grammatological parlance, the Arabic script is an *ʾabjad* (or consonantary) with obligatory notation of long vowels, and it abbreviates words by omitting short vowels, doubled consonants, and inflectional endings. An *ʾabjad* can be read faster than alphabetic script denoting both consonants and vowels, because the reduced denotation makes the larger unit of the word more recognizable, which is actually what the reader processes – not single graphemes (Bauer 1994–1996a:1435b after Coulmas 1989:52). The linguistic reconstruction required in this process is not done by a paradigmatic-etymological, or root pattern, analysis of each word, but rather by lexical recognition. Many words, however, are ambiguous; <ktʾb> for instance may mean either *kitāb* 'book' or *kuttāb* 'scribes'. The correct pronunciation of a word depends on the syntactic and semantic context, which even a native speaker might miss at the first reading of a sentence (Biesterfeldt 1994–1996:1300a). Correct reading demonstrates competence in the → *ʿarabiyya*, and Arabic-Islamic society is unique in the precedence it assigns this knowledge as the foundation and emblem of general culture.

The same graphic economy of omitting inflectional endings and short vowels safeguards the inclusiveness of Arabic script. It tends to veil the → Middle Arabic hypercorrections or Neo-

Arabic dialectal slips of careless or uneducated writers. Inversely, Arabic script also allows written text to be read to a certain degree as colloquial, provided one adds lexical substitution for very common dialectal terms (Bauer 1994–1996b:1485a, 1489a).

Although this ambiguity of the script has been criticized by Arab intellectuals (Meynet 1971), it allows "an efficient linguistic communication fulfilling modern needs without requiring the sacrifice of a literary culture and tradition in the same process" (Bauer 1994–1996b:1490b). See also → Epigraphy, → Palaeography, → Script and art, → Alphabet: Use for other languages.

Bibliographical references

Primary sources

Ibn an-Nadīm, *al-Fihrist* = Muḥammad ibn ʾIsḥāq Ibn an-Nadīm, *al-Fihrist*. Ed. Riḍā Tajaddud. Tehran, 1971.
ʾIsḥāq ibn ʾIbrāhīm, *Burhān* = ʾIsḥāq ibn ʾIbrāhīm ibn Sulaymān ibn Wahb al-Kātib, *al-Burhān fī wujūh al-bayān*. Ed. Ḥifnī Muḥammad Šaraf. Cairo: Maktabat aš-Šabāb, 1389/1969.
al-Xaṭīb al-Baġdādī, *Taqyīd* = ʾAbū Bakr ʾAḥmad ibn ʿAlī al-Xaṭīb al-Baġdādī, *Taqyīd al-ʿilm*. Ed. Yūsuf al-ʿIšš. Mecca: Dār ʾIḥyāʾ as-Sunna an-Nabawiyya, 1949.
Qalqašandī, *Ṣubḥ* = Šihāb ad-Dīn ʾAḥmad ibn ʿAlī al-Qalqašandī, *Ṣubḥ al-ʾaʿšā fī ṣināʿat al-ʾinšāʾ*. 14 vols. 2nd ed. Cairo 1913–1920/1357–1938. (Repr. 1383/1963.)

Secondary sources

Abbott, Nabia. 1939. *The rise of the North Arabic script and its Qurʾānic development*. Chicago: University of Chicago Press.
——. "Arabic paleography". *Ars Islamica* 8.65–104.
——. 1957–1972. *Studies in Arabic literary papyri*. 3 vols. Chicago: University of Chicago Press.
ʿAwwād, Kūrkīs. 1948. "al-Waraq ʾaw al-kāhgad: Ṣināʿatuhu fī l-ʿuṣūr al-ʾislāmiyya". *Revue de l'Académie Arabe de Damas* [previously *Majalla Majmaʿ al-ʿIlmī al-ʿArabī*] 23.409–438.
Bauer, Thomas. 1994–1996a. "Das arabische Schriftsystem". Günther and Ludwig (1994–1996:1433–1436).
——. 1994–1996b. "Die schriftliche Sprache im Arabischen". Günther and Ludwig (1994–1996: 1483–490).
Becker, C.H. 1906. *Papyri Schott-Reinhardt*, I. Heidelberg: C. Winter.
Biesterfeldt, Hans Hinrich. 1994–1996. "Lese- und Schreibunterricht im arabischen Sprachraum". Günther and Ludwig (1994–1996:1299–1309).
Blau, Joshua. 2002. *A handbook of early Middle Arabic*. Jerusalem: Max Schloessinger Memorial Foundation.
Briquel-Chatonnet, F. 1997. "De l'araméen à l'arabe: Quelques réflexions sur la genèse de l'écriture arabe". *Scribes et manuscrits du Moyen-Orient*, ed. François Déroche and Francis Richard, 135–149. Paris: Bibliothèque Nationale de France.
Cantineau, Jean. 1930–1932. *Le nabatéen*. 2 vols. Paris: E. Leroux. (Repr. Osnabrück: O. Zeller, 1978.)
Coulmas, Florian. 1989. *The writing systems of the world*. Oxford: Blackwell.
Daniels, Peter T. 1990. "Fundamentals of grammatology". *Journal of the American Oriental Society* 110.727–730.
—— and William Bright (eds.). 1996. *The world's writing systems*. New York and Oxford: Oxford University Press.
Déroche, François. 1980. "Les écritures coraniques anciennes: Bilan et perspectives." *Revue des Etudes Islamiques* 48.207–224.
——. 1992. *The Abbasid tradition: Qurʾāns of the 8th to the 10th centuries AD*. London: Nour Foundation.
Diem, Werner 1976. "Some glimpses at the rise and early development of the Arabic orthography". *Orientalia* 45.251–261.
——. 1979. "Untersuchungen zur frühen Geschichte der arabischen Orthographie. I. Die Schreibung der Vokale". *Orientalia* 48.207–257.
——. 1980. Untersuchungen zur frühen Geschichte der arabischen Orthographie. II. Die Schreibung der Konsonanten". *Orientalia* 49.67–106.
——. 1981. Untersuchungen zur frühen Geschichte der arabischen Orthographie. III. Endungen und End-schreibungen". *Orientalia* 50.332–83.
——. 1983. "Untersuchungen zur frühen Geschichte der arabischen Orthographie. IV. Die Schreibung der zusammenhängenden Rede. Zusammenfassung". *Orientalia* 52.357–404.
——. 1993. "Philologisches zu den Khalili-Papyri". *Wiener Zeitschrift zur Kunde des Morgenlandes* 83.39–81.
Dietrich, Albert. 1955. *Arabische Briefe aus der Papyrussammlung der Hamburger Staats- und Universitäts-Bibliothek*. Hamburg: Augustin.
—— and Oswald Loretz. 1988. *Die Keilalphabete: Die phönizisch-kanaanäischen und altarabischen Alphabete in Ugarit*. Münster: Ugarit-Verlag.
Endress, Gerhard. 1982. "Die arabische Schrift". Fischer (1982:165–183, 190–197).
Ephʾal, Israel. 1982. *The ancient Arabs: Nomads of the Fertile Crescent, 9th–5th centuries B.C.* Jerusalem and Leiden: Magnes Press and E.J. Brill.
Ferrando, Ignacio. 2001. *Introducción a la historia de la lengua árabe: Nuevas perspectivas*. Saragossa: Navarro y Navarro.
Fischer, Wolfdietrich (ed.). 1882. *Grundriss der arabischen Philologie. I. Sprachwissenschaft*. Wiesbaden: L. Reichert.
Flury, Samuel. 1920. *Islamische Schriftbänder: Amida-Diarbekr. XI Jahrhundert*. Basel: Frobenius.
Gacek, Adam. 1987. "Al-Nuwayrī's classification of Arabic scripts". *Manuscripts of the Middle East* 2.126–130.
——. 1989. "Arabic scripts and their characteristics as seen through the eyes of Mamluk authors". *Manuscripts of the Middle East* 4.144–149.

Grohmann, Adolf. 1952. *From the world of Arabic papyri*. Cairo: Al-Maaref Press.

——. 1967–1971. *Arabische Paläographie*. 2 vols. Vienna: H. Böhlau.

Gruendler, Beatrice. 1993. *The development of the Arabic scripts: From the Nabataean era to the first Islamic century according to dated texts*. Atlanta: Scholars Press.

——. 2001. "Arabic Script". *Encyclopedia of the Qurʾān*, ed. Jane Dammen McAuliffe a.o. I, 127–144. Leiden: E.J. Brill.

Günther, Hartmut and Otto Ludwig (eds.). 1994–6. *Schrift und Schriftlichkeit: Writing and its use*. 2 vols. Berlin and New York: de Gruyter.

Healey, John F. 1990–1991. "Nabataean to Arabic: Calligraphy and script development among the pre-Islamic Arabs". *Manuscripts of the Middle East* 5.41–52.

——. 1993. *The Nabataean tomb inscriptions of Mada'in Salih*. Oxford: Oxford University Press.

Hopkins, Simon. 1984. *Studies in the grammar of Early Arabic based upon papyri datable to before 300 A.H./912 A.D*. Oxford: Oxford University Press.

Khan, Geoffrey. 1992. *Arabic papyri: Selected material from the Khalili Collection*. London: Azimuth Editions and Oxford University Press.

Meynet, Roland. 1971. *L'écriture arabe en question: Les projets de l'Académie de langue arabe du Caire de 1938 à 1968*. Beirut: Dar El-Machreq.

Müller, Walter W. 1982 "Das Frühnordarabische". Fischer (1982:I, 17–29).

——. 1994–1996. "Die altsüdarabische Schrift". Günther and Ludwig (1994–1996:307–312).

Naveh, Joshua. 1979. "A Nabataean incantation text." *Israel Exploration Journal* 29.111–119 and pl. 14.

——. 1982. *Early history of the alphabet*. Leiden: E.J. Brill.

Negev, Abraham, 1986. "Obodas the God". *Israel Exploration Journal* 36.56–60.

Rāġib, Yūsuf. 1990. "L'écriture des papyrus arabes aux premiers siècles de l'Islam". *Les premières écritures islamiques*, ed. Alfred-Louis de Prémare, 14–29. Paris: Edisud.

Rosenthal, Franz. 1947. *Technique and approach of Muslim scholarship*. Rome: Pontificium Institutum Biblicum.

——. 1948. "Abū Ḥaiyān al-Tawḥīdī on penmanship". *Ars Islamica* 13–14.1–27.

——. 1962. "Nabataean and related inscriptions". *Excavations at Nessana*, I, ed. H. Dunscombe Colt, 198–210 and pls. 34–35. London: British School of Archaeology in Jerusalem.

Ryckmans, Jacques, Walter W. Müller, and Yusuf M. ʿAbdallah. 1994. *Textes du Yemen antique inscrits sur bois*. Louvain: Université Catholique de Louvain, Institut Orientaliste.

Schoeler, Gregor. 2002. *Écrire et transmettre dans les débuts de l'islam*. Paris: Presses Universitaires de France.

Schroeder, Erich. 1937. "What was the *badīʿ* script?". *Ars Islamica* 4.232–248

Sourdel-Thomine, Janine. 1978. "*K͟haṭṭ*". *Encyclopaedia of Islam* IV, 1113–1122. 2nd ed. Leiden: E.J. Brill.

Starcky, Jean. 1954. "Un contrat nabatéen sur papyrus". *Revue Biblique* 61.161–81.

——. 1966. "Pétra et la Nabatène". *Dictionnaire de la Bible: Supplément* VII, 886–1017. Paris: Letouzey et Ané.

Versteegh, Kees. 1997. *The Arabic language*. Edinburgh: Edinburgh University Press.

Yardeni, Ada, Barukh Levine, a.o. (eds). 2002. *The documents from the Bar Kochba period from the Cave of Letters: Hebrew, Aramaic and Nabataean-Aramaic papyri*. 2 vols. Jerusalem: Israel Exploration Society.

BEATRICE GRUENDLER
(Yale University)

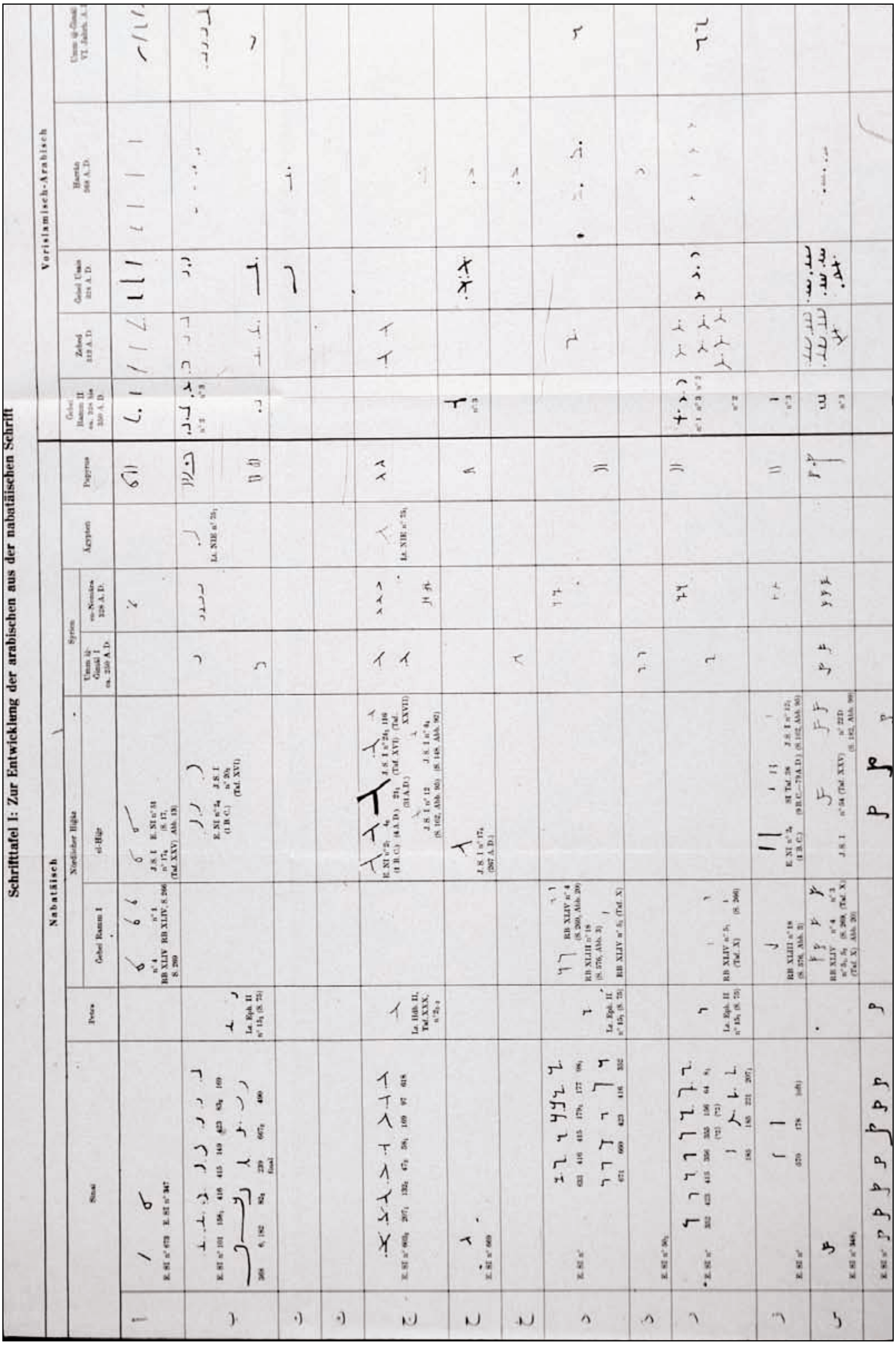

Schrifttafel I: Zur Entwicklung der arabischen aus der nabatäischen Schrift
Nabatäisch
Vorislamisch-Arabisch
Sinai
Petra
Nördlicher Hiǵāz
Gebel Ramm I
el-Hiǵr
Syrien
Umm iǵ-Ǵimāl I ca. 250 A.D.
en-Nemāra 328 A.D.
Aegypten
Papyrus
Gebel Ramm II ca. 328 bis 350 A.D.
Zebed 512 A.D.
Gebel Usais 528 A.D.
Harrān 568 A.D.
Umm iǵ-Ǵimāl VI. Jahrh. A.D.

Plate Ia. Development of pre-Islamic Arabic from Nabataean graphemes (Grohmann 1967–1971:2, folding sheet following 34).

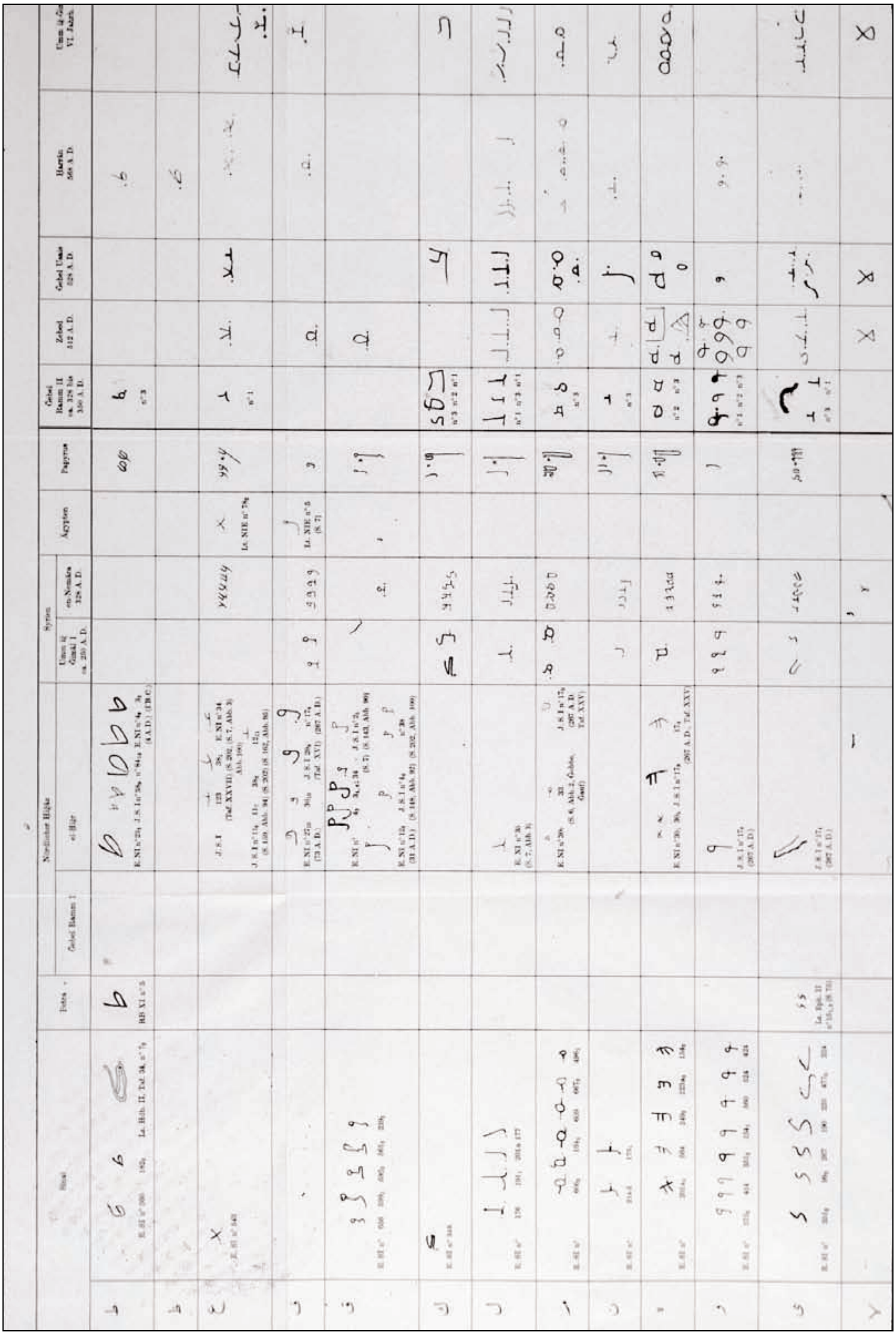

Sinai
Petra
Nördlicher Hiğāz
Gebel Ramm I
el-Hiğr
Syrien
Umm iğ-Ğimāl I ca. 250 A.D.
en-Nemāra 328 A.D.
Ägypten
Papyrus
Gebel Ramm II ca. 328 bis 350 A.D.
Zebed 512 A.D.
Gebel Usais 528 A.D.
Harrān 568 A.D.
Umm iğ-Ğimāl VI. Jahrh.

Plate Ib. Development of pre-Islamic Arabic from Nabataean graphemes (Grohmann 1967–1971:2, folding sheet following 34).

Plate IIa. Chart of early (monumental and cursive) Nabataean graphemes; epitaphs from Madāʾin Ṣāliḥ (cols. A–B), cursive of Naḥal Ḥever papyri and related texts (cols. C–D), cursive of Jabal Ramm graffito and related texts (cols. E) (Healey 1990–1991:50).

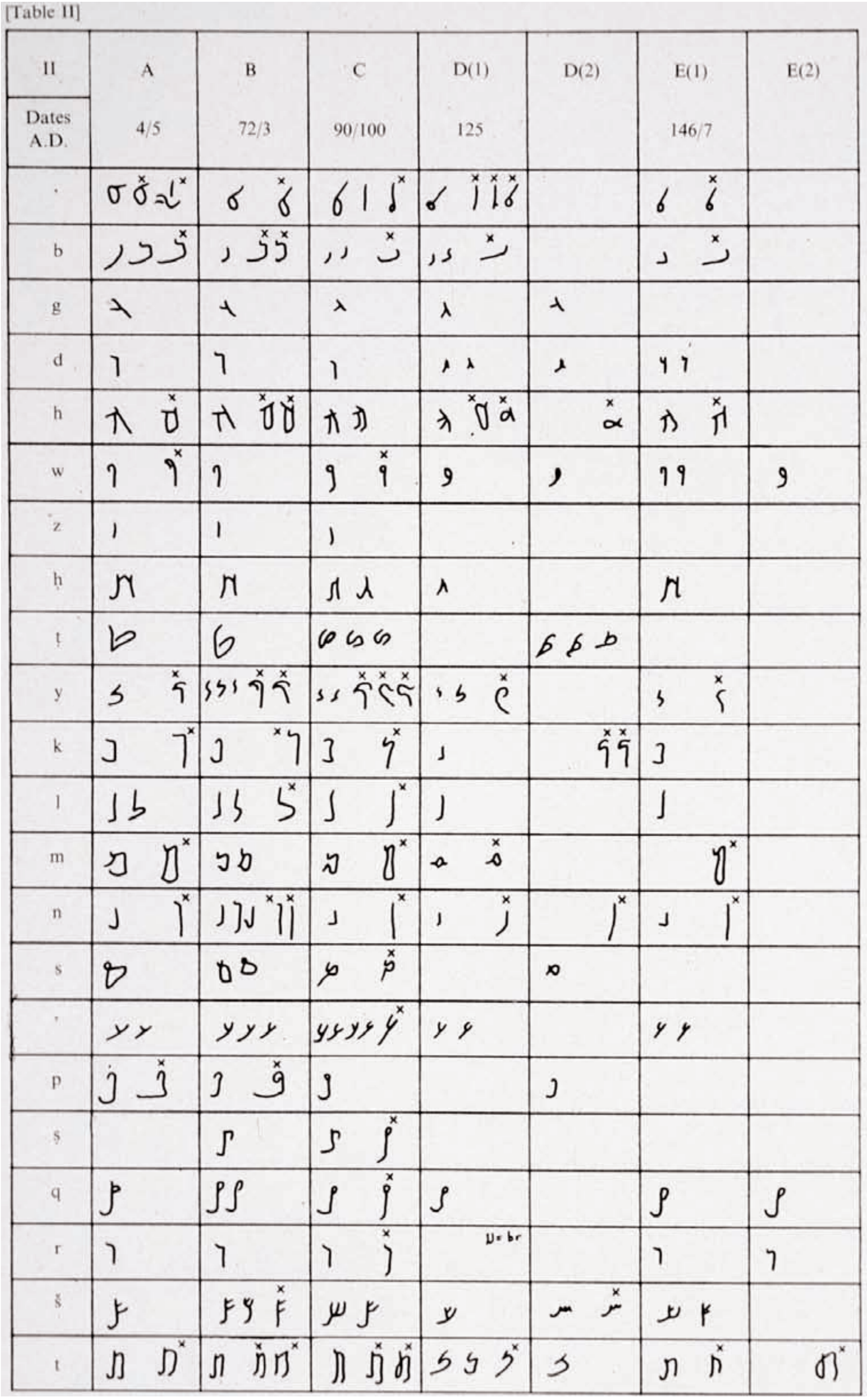

[Table II]
II
A
B
C
D(1)
D(2)
E(1)
E(2)
Dates A.D.
4/5
72/3
90/100
125
146/7
ʾ
b
g
d
h
w
z
ḥ
ṭ
y
k
l
m
n
s
ʿ
p
ṣ
q
r
š
t

Plate IIb. Selective chart of (monumental and cursive) Nabataean graphemes; graffiti from Sinai and Egypt (cols. 11, 16–18), epitaphs from el-ʿUla (cols. 21–22), an-Namāra inscription (col. 23) (for cols. A–B and C–E, see ill. IIa) (Healey 1990–1991:51).

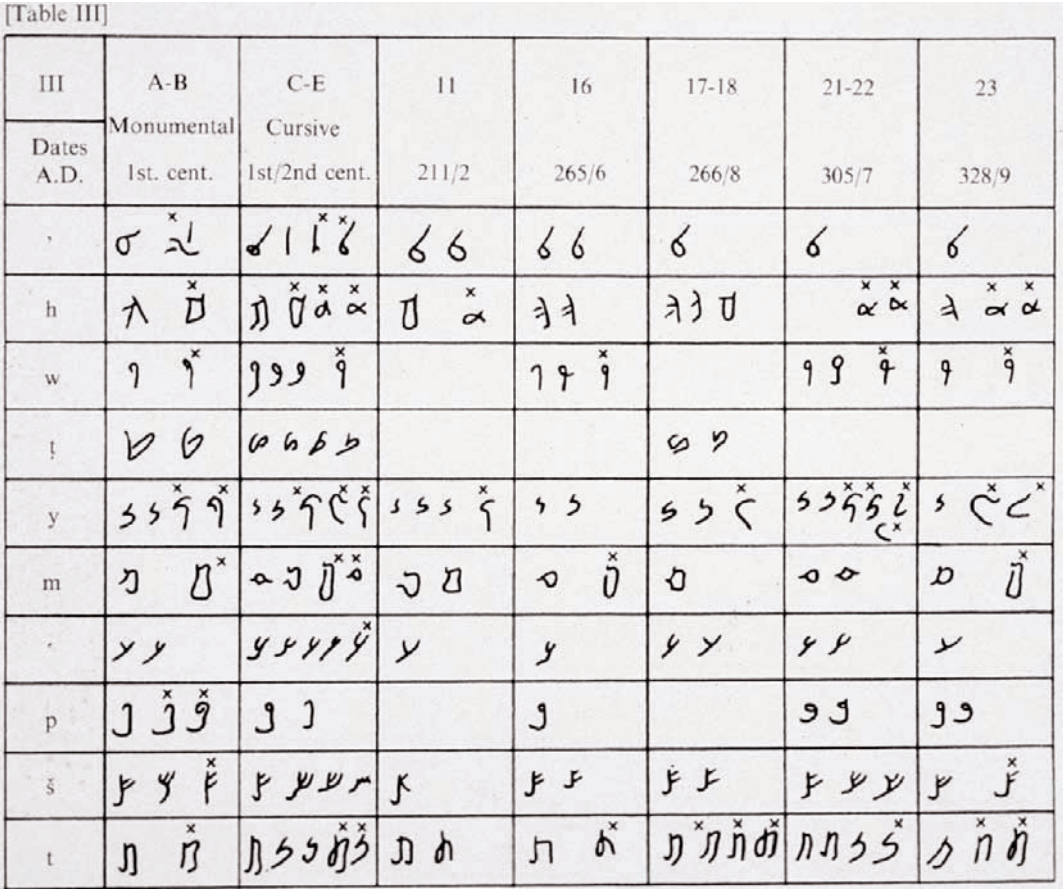

[Table III]
III
Dates A.D.
A-B Monumental 1st. cent.
C-E Cursive 1st/2nd cent.
11 211/2
16 265/6
17-18 266/8
21-22 305/7
23 328/9
ʾ
h
w
ṭ
y
m
ʿ
p
š
t

Plate IIIa. Nabataean tenancy agreement on papyrus P. Yadin (5/6 Ḥev) 6 recto from Naḥal Ḥever dated 119–20 C.E. in extreme cursive; sketch and photograph (Yardeni a.o. 2002:1, 259 and pl. 55).

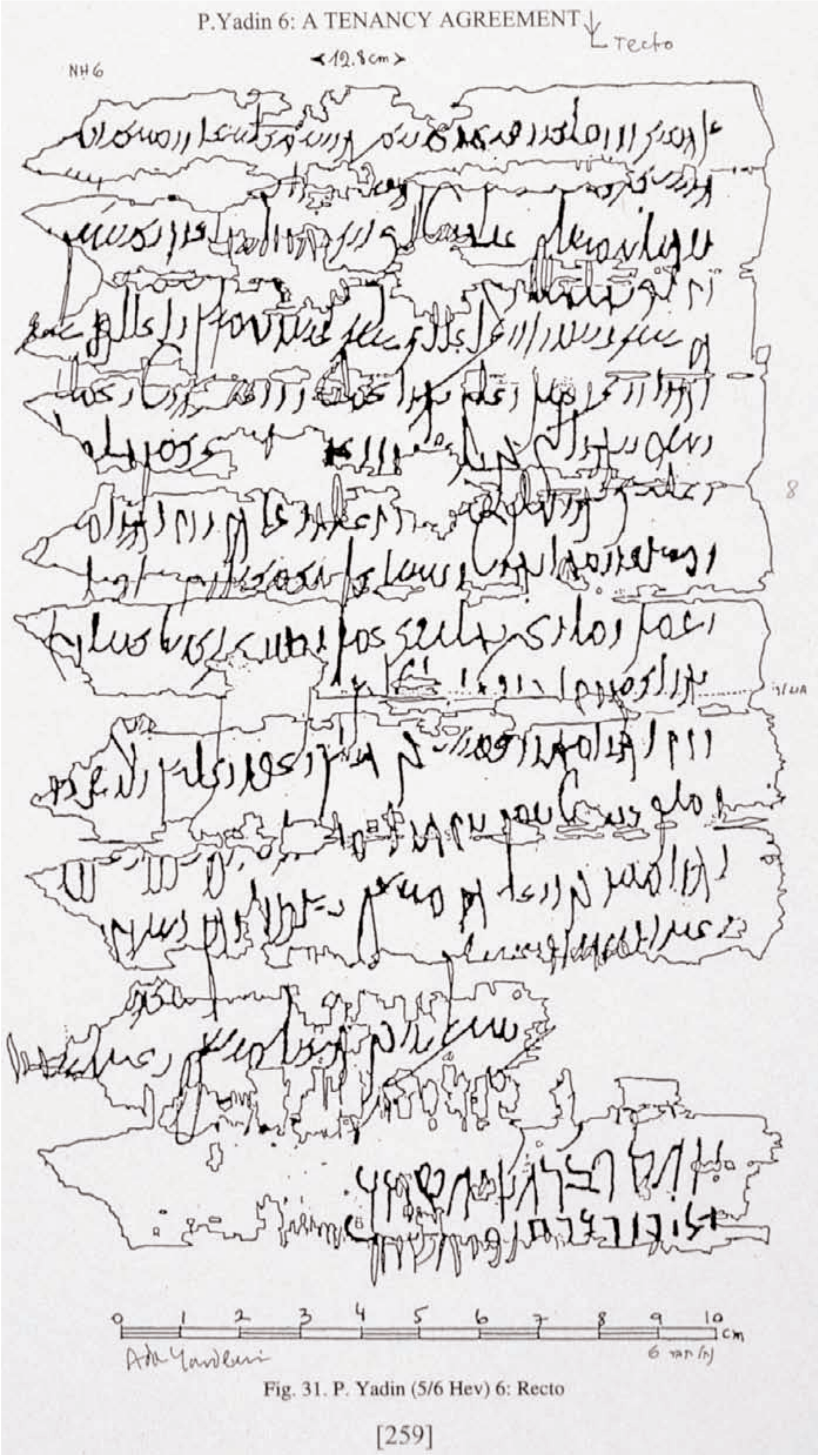

P.Yadin 6: A TENANCY AGREEMENT
Recto
12.8cm
NH6
8
0 1 2 3 4 5 6 7 8 9 10 cm
Ada Yardeni
Fig. 31. P. Yadin (5/6 Hev) 6: Recto
[259]

Plate IIIb. Nabataean tenancy agreement on papyrus P. Yadin (5/6 Ḥev) 6 recto from Naḥal Ḥever dated 119–20 C.E. in extreme cursive; sketch and photograph (Yardeni a.o. 2002:1, 259 and pl. 55).

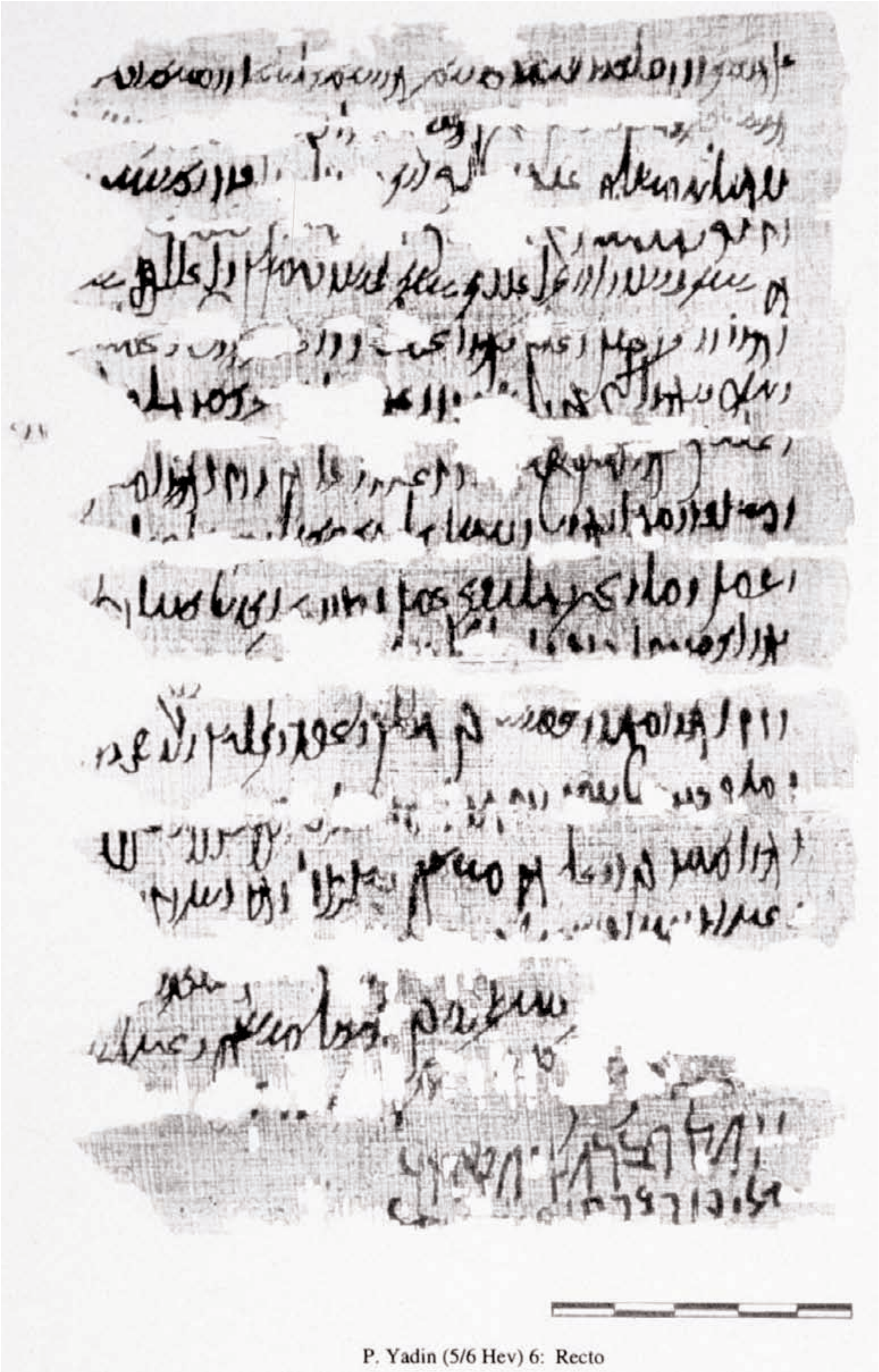

P. Yadin (5/6 Hev) 6: Recto

Plate IV. Chart of Modern Arabic script (Daniels and Bright 1996:560).

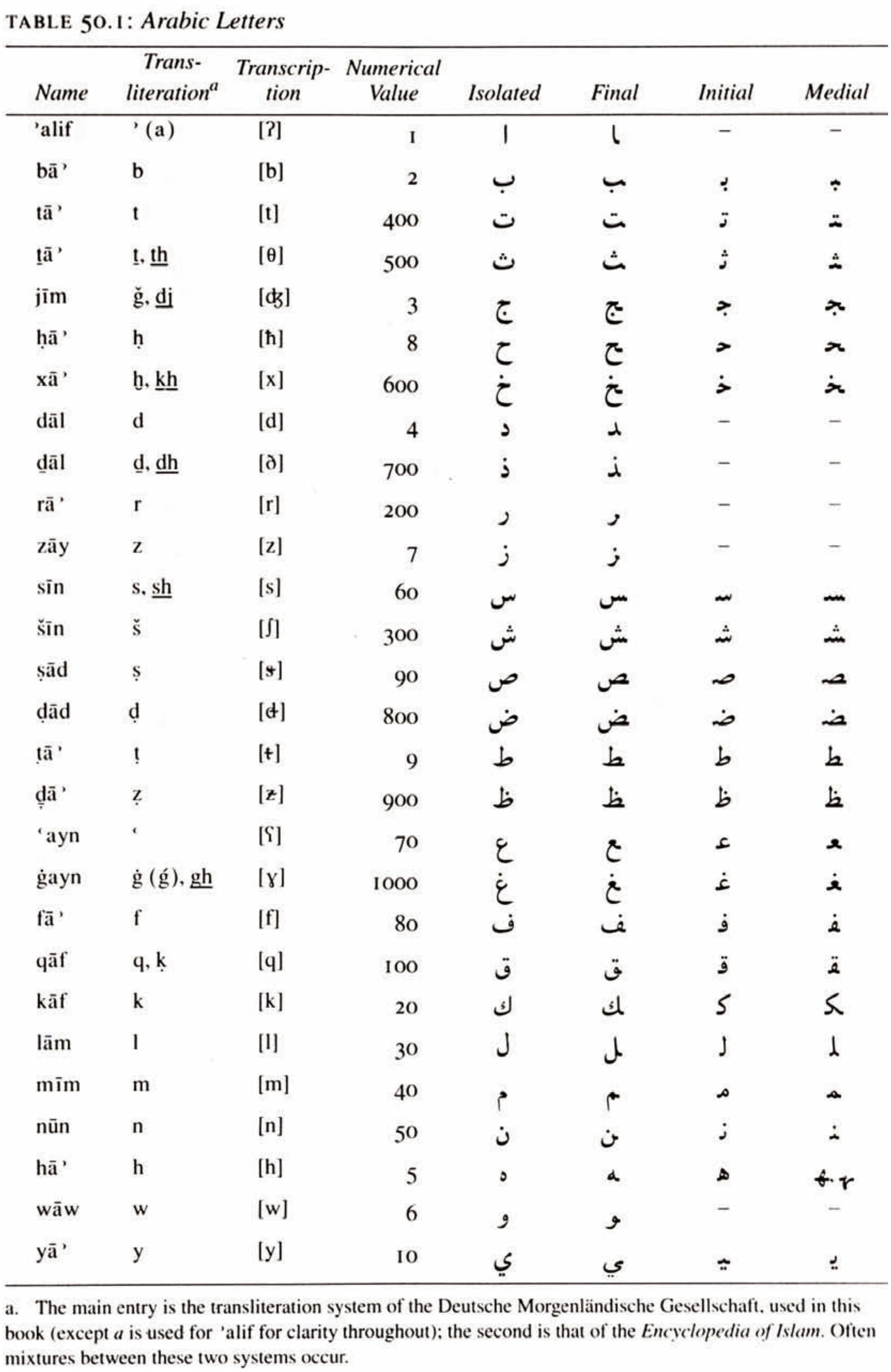


TABLE 50.1: *Arabic Letters*

Name	*Transliteration*[a]	*Transcription*	*Numerical Value*	*Isolated*	*Final*	*Initial*	*Medial*
'alif	' (a)	[ʔ]	1	ا	ـا	–	–
bā'	b	[b]	2	ب	ـب	بـ	ـبـ
tā'	t	[t]	400	ت	ـت	تـ	ـتـ
ṯā'	ṯ, th	[θ]	500	ث	ـث	ثـ	ـثـ
jīm	ǧ, dj	[ʤ]	3	ج	ـج	جـ	ـجـ
ḥā'	ḥ	[ħ]	8	ح	ـح	حـ	ـحـ
xā'	ḫ, kh	[x]	600	خ	ـخ	خـ	ـخـ
dāl	d	[d]	4	د	ـد	–	–
ḏāl	ḏ, dh	[ð]	700	ذ	ـذ	–	–
rā'	r	[r]	200	ر	ـر	–	–
zāy	z	[z]	7	ز	ـز	–	–
sīn	s, sh	[s]	60	س	ـس	سـ	ـسـ
šīn	š	[ʃ]	300	ش	ـش	شـ	ـشـ
ṣād	ṣ	[ᵴ]	90	ص	ـص	صـ	ـصـ
ḍād	ḍ	[ᵭ]	800	ض	ـض	ضـ	ـضـ
ṭā'	ṭ	[ᵵ]	9	ط	ـط	طـ	ـطـ
ḏ̣ā'	ẓ	[ᵶ]	900	ظ	ـظ	ظـ	ـظـ
ʻayn	ʻ	[ʕ]	70	ع	ـع	عـ	ـعـ
ġayn	ġ (ǵ), gh	[ɣ]	1000	غ	ـغ	غـ	ـغـ
fā'	f	[f]	80	ف	ـف	فـ	ـفـ
qāf	q, ḳ	[q]	100	ق	ـق	قـ	ـقـ
kāf	k	[k]	20	ك	ـك	كـ	ـكـ
lām	l	[l]	30	ل	ـل	لـ	ـلـ
mīm	m	[m]	40	م	ـم	مـ	ـمـ
nūn	n	[n]	50	ن	ـن	نـ	ـنـ
hā'	h	[h]	5	ه	ـه	هـ	ـهـ
wāw	w	[w]	6	و	ـو	–	–
yā'	y	[y]	10	ي	ـي	يـ	ـيـ

a. The main entry is the transliteration system of the Deutsche Morgenländische Gesellschaft, used in this book (except *a* is used for 'alif for clarity throughout); the second is that of the *Encyclopedia of Islam*. Often mixtures between these two systems occur.

Plate V. Chart of Arabic scripts from the 1st/7th century; early cursive (ls. 1–2), epigraphic script (l. 3), chancellery cursive (l. 4), entagia cursive (l. 5), protocol cursive (l. 6), cursive of palimpsest PSI 1272v (l. 7); and *ḥijāzī* script of *Qurʾāns* (ls. 8–9) (Gruendler 1993:141).

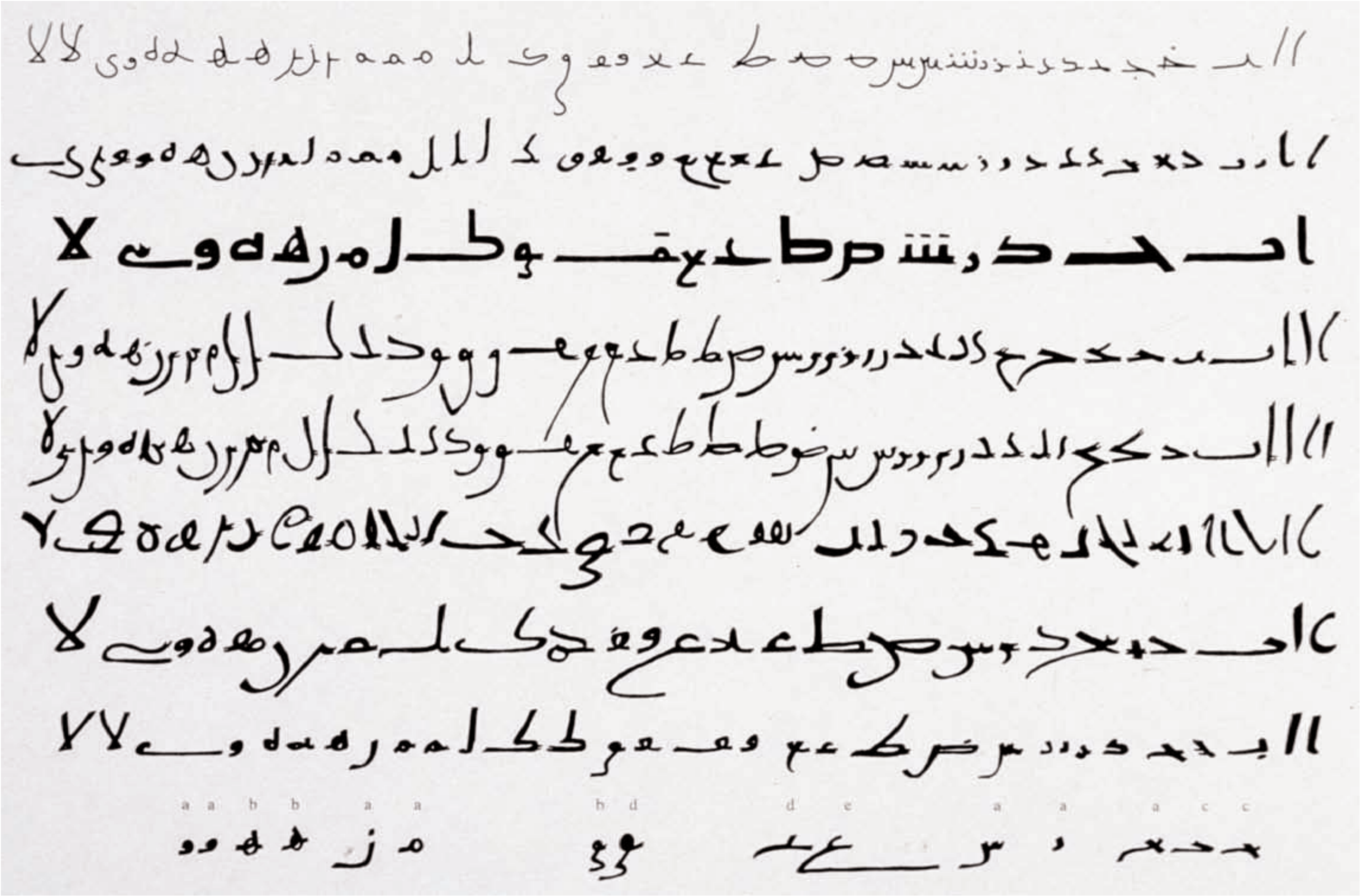

a a b b a a
b d
d e
a a a c c

Plate VIa. Arabic letter on papyrus PSR I, III from the Egyptian governor Qurra ibn Šarīk to Basīl, district head of Qōm ʾIšqawh, instructing him to expedite wheat delivery to Cairo and to prevent abusive tax collection, dated 91/710 (Becker, 1906: pl. 3). *Papyri Schott-Reinhardt* I. Heidelberg: C. Winter. Pl. 3).

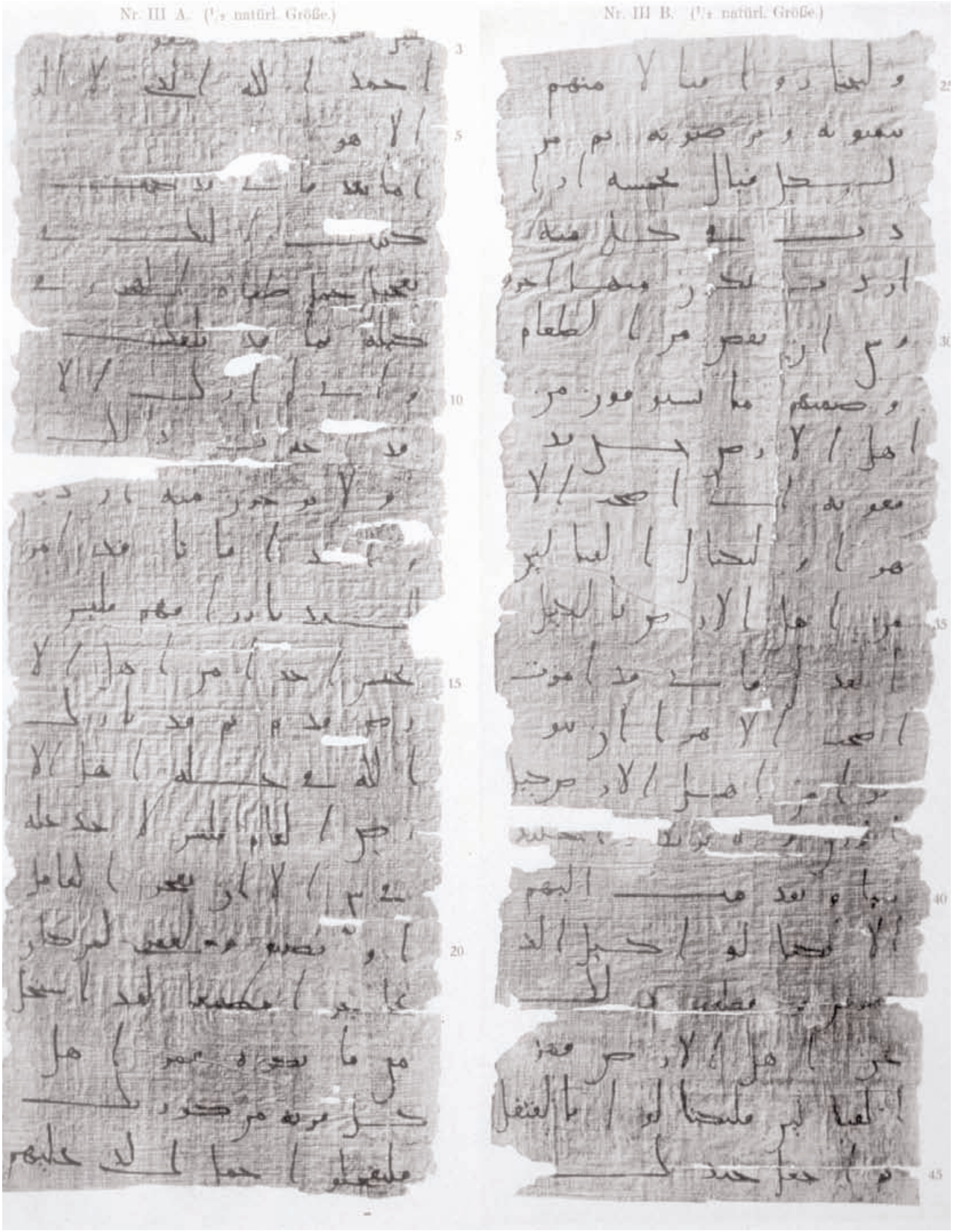
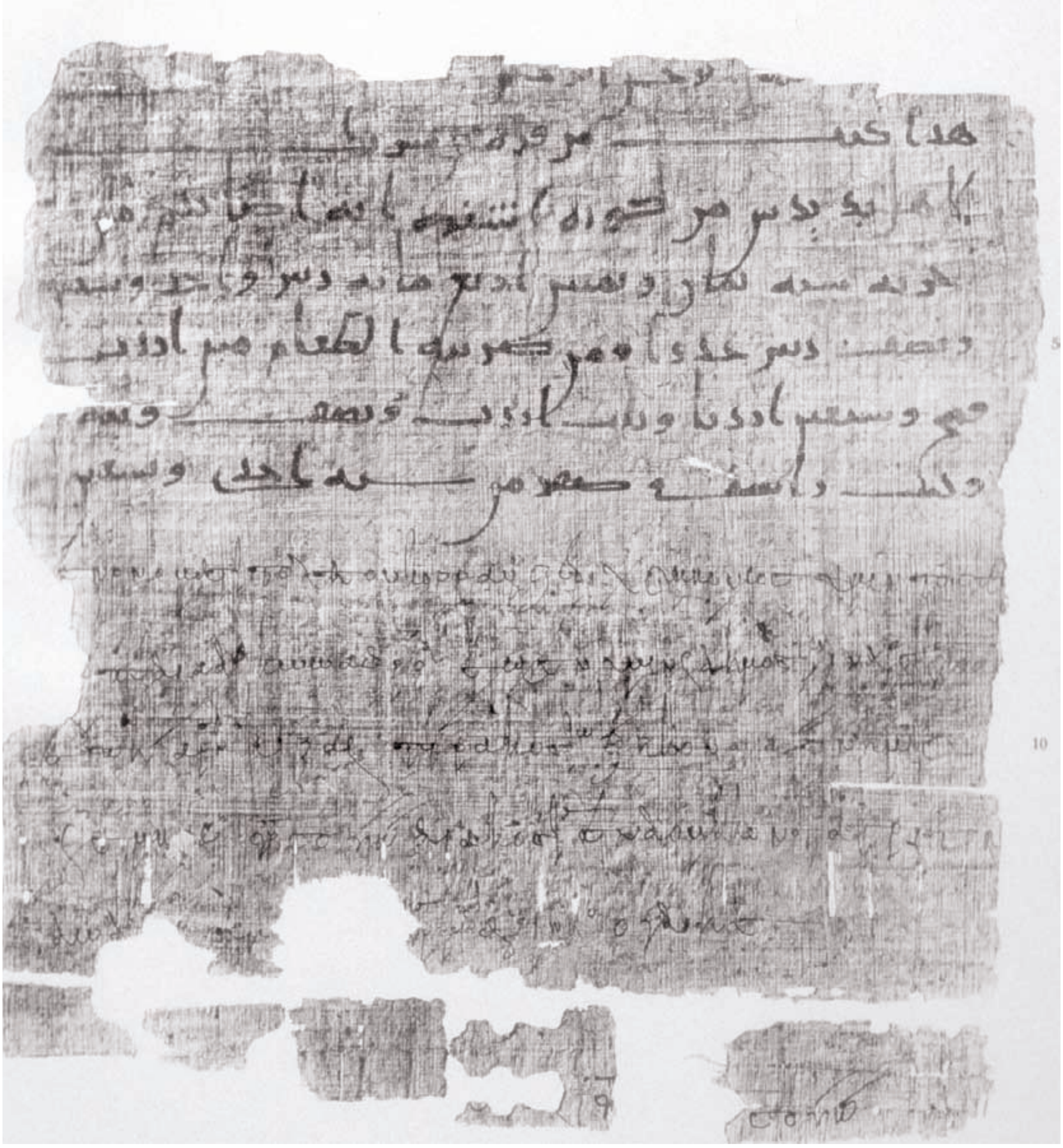

Nr. III A. (1/2 natürl. Größe.)
Nr. III B. (1/2 natürl. Größe.)

Plate VIb. Bilingual Greek-Arabic entagion (demand note) on papyrus PSR I, V from Qurra ibn Šarīk to the inhabitants of Pedias (Badīdas) in the end of ʾIšqawh for poll tax payable in coin and wheat, dated 91/710 (Becker 1906: pl. 6).



Nr. V (natürl. Größe).

Arabic Sign Language → Sign Languages

Arabic Studies in Europe

1. The motives

The study of Arabic in Europe can be traced back to the Middle Ages, and by the 18th century a variety of arguments had been assembled in its support. Frequently intended to attract an uninformed patron, not all of them stand up to critical scrutiny today, but they came to form a standard litany without which no apology of Arabic would be complete (Hamilton 1985: 66–96).

The first reason, which proved remarkably resilient, was the use of Arabic for missionaries (Dannenfeldt 1955). The possibility of converting the Muslims to Christianity and of combating Islam had once raised the highest hopes. These suffered a major setback in the 14th century when the Mongols converted to Islam, yet, if Arabic continued to be studied throughout the later Middle Ages, it was still partly due to the dream of converting Muslims by peaceful methods and partly to pastoral objectives in previously Muslim areas which had been conquered by the Christians. The establishment of chairs in Arabic, as well as in Greek, Hebrew, and Syriac, at the universities of Paris, Oxford, Bologna, Salamanca, and Avignon (the seat of the papacy), was consequently proposed at the Council of Vienna in 1312 and at the Council of Basle in 1434, but it was not carried out. At the same time the determination to win over the Muslims with rational arguments derived from a sound knowledge of Islam induced European scholars to tackle the translation of the *Qur'ān* from the 12th century onwards.

At the Council of Florence, which lasted from 1438 to 1445, a further incentive to teach Arabic to missionaries emerged: the union of the Churches and the wish to convert the Arabic-speaking Christians to Roman Catholicism. In view of the difficulty of converting Muslims this second objective seemed far more practicable. It was to form a significant part of the policy behind the main missionary organization of the 17th century, the Congregation of Propaganda Fide founded by Pope Gregory XV in 1622.

Another reason given for the study of Arabic was the need to read the works of the great Arab scientists. In view of the products of the Arab physicians, astronomers, and mathematicians between the 8th and 12th centuries, and the translations of Greek scientific texts in ʿAbbasid Baghdad, this had once been a valid argument. It was in order to translate from the Arabic that an international group of scholars traveled to Toledo in the course of the 12th century. Robert of Ketton, Hermann of Carinthia, Gerard of Cremona, John of Seville, and many others assembled in the city, which, in 1085, had at last been reconquered from the Muslims by the Christian forces, and settled down to turning some of the main Arabic scientific texts into Latin (Haskins 1924:12–19). In the centuries that followed, their versions came under increasing criticism and later scholars learnt Arabic in order to improve on them. But, although certain texts in the fields of mathematics and astrology which were only available in Arabic continued to fascinate scholars until well into the 18th century, even by the mid-17th century the Arab contribution to science was itself being reassessed. The discovery by the humanists of Greek manuscripts of texts previously only known in Arabic translation had confirmed the suspicion that the Arabic translations were not always reliable, and the growing tendency in the 17th century to question traditional knowledge, to experiment and to base scientific conclusions on personal observation, diminished the demand for the scientific works either of the Ancients or of the Arabs (Klein-Franke 1980).

Then there was the proximity of Arabic to Hebrew and its use for students of the Bible. Throughout the early modern period the majority of academic students of Arabic were theologians, and Hebrew was the first Semitic language they encountered. Hebrew, it was generally believed in Christian Europe, was the original language from which all others descended. In the genealogical trees of languages, Arabic, together with Syriac and Aramaic, occupied a privileged place as its immediate descendants. If students learnt Arabic, it was consequently argued, they would expand their acquaintance with the kindred tongues and gain a deeper knowledge of Hebrew. Many dictionaries of the 17th century, like the Arabic one of William Bedwell in England which remained in manuscript (Hamilton 1985:85–94) and the 'polyglot' *Lexicon pentaglotton* compiled by Valentin Schindler and published in Hanau in 1612 (Hamilton 1989:574), bore out this persuasion and contained countless comparisons between the various Semitic languages which were supposed to be of use to Biblical students (Hamilton 1979:575). They led up to the largest of the polyglot dictionaries, Edmund Castell's *Lexicon heptaglotton* (London 1669), much praised at the time but of little assistance to the progress of Arabic studies (Toomer 1996:255–265).

There also existed a belief among Bible students, which persisted into the 18th century, that the Book of Job, one of the earliest in the Old Testament, had originally been written in Arabic and that a knowledge of that language would be of assistance in solving some of the linguistic obscurities which it contained. Where the study of the New Testament was concerned, on the other hand, scholars all over Europe felt that the Arabic renderings might reflect a far earlier Syriac version and contain interesting variants which would enable them to improve on the standard Latin translation, the Vulgate attributed to St Jerome in the late 4th and early 5th century. This belief brought about the inclusion of Arabic versions of the Scriptures in the two principal polyglot Bibles of the 17th century, the Paris Polyglot (1629–1645) and the London Polyglot (1653–1657).

In fact the advisability of associating Arabic with Hebrew and using it for Biblical studies was soon questioned. In the first years of the 17th century the French scholar Joseph Justus Scaliger expressed his regret that so many people were studying Arabic in conjunction with Hebrew and limiting themselves almost entirely to Biblical texts. Arabic, he pointed out, was above all an Islamic language. It could not be approached profitably without a thorough knowledge of the *Qur'ān*, and should then continue to be studied on the basis of a wide selection of Islamic texts. If anything, he suggested, Arabic should be learnt in association not with Hebrew but with other Islamic languages, notably with Turkish and Persian (Hamilton 1989:576–579).

One of the more practical reasons given for studying Arabic was the need of commercial exchange which had been stressed intermittently ever since the 13th century. This objective grew in significance as the Arab world became of increasing commercial and political importance in the early modern period. The best French Arabist of the 16th century, Guillaume Postel, said that the knowledge of Arabic would enable a traveler to make his way from Morocco to the Moluccas without an interpreter – an attractive prospect for a merchant – and it was very largely for commercial reasons that the curators of Leiden University decided to found a chair of Arabic at the end of the 16th century (Juynboll 1931:10–11).

To the commercial importance of the Arabic-speaking world was added the fascination entailed by its antiquity, its vastness, and its comparative remoteness. In the age of both geographical and intellectual exploration it contained countless facets that intrigued scholars. Cartographers wanted to chart the area and to discover the modern Arabic names of sites only known from the geographers of Antiquity. Physicians, botanists, zoologists, and geologists were eager to assemble material concerning flora, fauna, and geological formations unknown in the West. Historians and chronologists wanted information about the Arabs which would enable them to acquire a more complete idea of the history of the world and its various civilizations. This, in its turn, ultimately led to the revision of the traditional Biblical chronology hallowed in Christian Europe.

The interest in comparative religions which got underway in the 17th century led to a pronounced interest in Islam and an ever greater curiosity about Arabic religious texts, while the collections of Arabic manuscripts, brought back from the East by Jacobus Golius, Edward Pococke, Levinus Warner, and others, stimulated an interest in Arabic literature. It was only well after Antoine Galland had introduced the European reading public to the *Arabian Nights* in the first years of the 18th century and Johann Jakob Reiske had published some of the odes in the *Muʿallaqāt* in 1742, however, that scholars started to learn Arabic in order to gain a more direct acquaintance with literary, rather than scientific, texts. An early example of one who did so, in the 1760s, was the future Sanskrit scholar, William Jones (Fück 1955:130).

2. THE GRAMMARS

The first steps toward compiling an Arabic grammar in Christian Europe depended on the availability of teachers. The Christian Arabists who arrived in Toledo in the 12th century had a wide choice, above all resident Jews (both orthodox and converted) and Mozarabs (Christians once living under Muslim rule), who knew Arabic and Spanish (Burnett 1995). Outside Spain it was considerably less easy to engage an Arabic-speaker as a teacher. In the mid-15th century it took John of Segovia, who had retired to a monastery in Savoy, two years to find a copy of the *Qurʾān* and an Arabic-speaking Muslim ready to help him translate it. He finally found a jurist from Salamanca, but the jurist was only prepared to spend a few months with him and, after his departure, John of Segovia could find nobody else (Southern 1962:86–92).

Spain at first seemed a good place in which to study Arabic, and it was there that the first printed grammar, the *Arte para ligeramente saber la lengua araviga*, appeared in Granada in 1505.

The author, Pedro de Alcalá, the confessor of the archbishop of Granada Fernando de Talavera, was working in Andalusia in the years after the Christian Reconquista of 1492. His grammar, as well as his dictionary, were intended above all for missionaries working in the south of Spain who needed the language of the less educated people in order to converse with them and take their confessions. This required a particular terminology. It also meant that the Arabic he used, transcribed in the Roman alphabet and according to Castilian pronunciation, was the Arabic of al-Andalus. His main models for his grammar were the Latin and Spanish grammars by his fellow-countryman Antonio de Nebrija, and this explains the Latin structure which he imposed. At the same time, however, his Muslim informants, the "learned faqihs" to whom he refers in his dictionary (but whose educated use of the language he rejected), obviously introduced him to the Classical Arabic grammatical tradition, and one of the infelicities of the grammar as a whole is that no distinction is made between Classical Arabic rules and those of the Andalusian dialect (Jones 1988:134–143).

As a result of Christian pressure on the Muslims to convert and the hostility that this aroused, it became increasingly hard to find either a Muslim scholar ready to impart tuition or texts with which to work in Spain. Nicolaus Clenardus, from the southern Netherlands, taught in Salamanca in the 1530s, and there, in the library of Hernán Núñez, the professor of Hebrew, came across the manuscript of a standard Arabic manual on grammar which was to revolutionize the compilation of Arabic grammars in Europe, the 13th-century *ʾĀjurrūmiyya*. Núñez proved unable to teach him how to use it, so Clenardus proceeded to Granada and, thanks to the governor of the city, was provided with a tutor in Arabic. Yet, the difficulty of obtaining Arabic manuscripts induced him to leave for Morocco in 1640 in order to progress with his studies (Jones 1988:144–148).

To begin with, European grammarians had to rely on the combination of Arabic texts such as the *ʾĀjurrūmiyya* and the advice of a native speaker. In the course of the 16th and 17th centuries we find Turkish prisoners-of-war and Arabic-speaking Christians being employed as teachers and copyists. The standard of their knowledge, however, varied greatly and there was seldom a guarantee of quality. By the middle of the 16th century scholars wishing to study Arabic consequently tried to make their way to the Arab or Ottoman worlds.

The French were among the first to avail themselves of permanent diplomatic representation at the Porte, and Guillaume Postel set out for the Levant in 1535 with the king's first ambassador to the sultan. He then devoted himself to the study of Arabic in Istanbul under the tuition of a Turkish Christian (Balagna Coustou 1989). He already knew Hebrew, and his progress in the new language was rapid. He became acquainted with the *ʾĀjurrūmiyya* and another standard grammatical work of the 13th century, az-Zanjānī's *Taṣrīf*. Basing himself on these, he made the first major contribution to the knowledge of Arabic grammar in Europe. His early effort, the *Linguarum characterum differentium alphabetum introductio*, was published on his return to Paris in 1538, and he followed it up with his far more important *Grammatica arabica*, published in 1543. The *Grammatica arabica*, the first of its kind to be printed and to make use of Arabic types, was revolutionary in establishing "the method of incorporating the information contained in Arabic grammatical tracts into Western-style grammar books about Classical Arabic" (Jones 1988:155).

The next step in introducing Europeans to Arabic grammar was the publication of the Arabic grammars themselves, and this was undertaken in the last decade of the 16th century by Giovanni Battista Raimondi and the Tipografia Medicea, the printing press which he ran in Rome. In 1592 he produced the *ʾĀjurrūmiyya* and another work on grammar of the

13th century, Ibn al-Ḥājib's *Kāfiya*, both solely in unvocalized Arabic. Almost 20 years later, in 1610, he added the *Taṣrīf*, this time in vocalized Arabic accompanied by a Latin translation.

Such, together with the far less interesting European Arabic grammars by Jakob Christmann, Rutger Spey, and Bartholomaeus Radtmann, was the printed material at hand when the young Thomas Erpenius decided to study Arabic. Erpenius had proved himself an excellent Hebraist at Leiden University, and one of his professors was Scaliger, who had himself studied Arabic under Postel in Paris. But although Scaliger advised Erpenius on how to tackle the language, he actually started learning it outside the Netherlands. He was taught the rudiments by William Bedwell in London, and received some more, but not very reliable, instruction from the itinerant Egyptian Copt Josephus Abudacnus in Paris. It was also in France that Erpenius met ʾAḥmad ibn Qāsim al-Ḥajarī, an emigrant from Spain in the service of the ruler of Morocco, who was in France on a diplomatic errand in 1611. Despite the uncertain quality of his own Arabic – his first language was Spanish – al-Ḥajarī appears to have provided Erpenius with some of the information which, in combination with the Arabic grammars, enabled him to compile his own *Grammatica arabica*. It was published in Leiden in 1613, the year in which he took up the professorship of Arabic at the university.

Like Postel, Erpenius reconciled the Classical Arabic grammatical tradition with the Latin one, but he did so far more successfully and exhaustively. Postel's grammar is short, and one wonders how far a student would have progressed had he only had that on which to rely. Erpenius's work is far longer and more accurate. He gives numerous examples where Postel simply announces a general rule, and at last produced a grammar from which students could, and did, learn Arabic. Not only did Erpenius's grammar, to which he himself added in the course of his career in Leiden, remain the standard European work on the subject for almost two hundred years, but even when it was at last surpassed by Silvestre de Sacy's *Grammaire arabe* in 1810, it had a profound influence on Sacy himself and continued to affect later grammarians such as Karl Paul Caspari and William Wright (Jones 1988:187–212).

3. THE DICTIONARIES

An early dictionary, which was to be of crucial importance for the compilation of the first Arabic-Latin lexicon to be printed, was the Mozarabic Latin-Arabic glossary acquired by Postel in 1532 and now at the Leiden University Library (Seybold 1900). In a clear hand and with most of the Arabic words vocalized, it was compiled in Toledo shortly before 1175 by an Arabic-speaking Christian who wanted to instruct his countrymen in Latin in the period immediately before the Mozarabic community adopted Castilian as its main language. It is consequently in the Arabic spoken in Spain and had among its sources the Arabic translations of the Scriptures known to the compiler (Koningsveld 1977).

The next dictionary to play an important role in the development of Arabic lexicography in Europe was the *Vocabulista aravigo en letra castellana*, the Spanish-Arabic wordlist compiled by Pedro de Alcalá and published in Granada in 1505 in the same year as his Arabic grammar. Like the grammar it was intended for Spanish missionaries preaching to the converts from Islam in Andalusia – thus for students of Arabic – and it was modeled after Antonio de Nebrija's Spanish-Latin dictionary. Also like the grammar, it was transcribed in the Roman alphabet for Spaniards wishing to speak the dialect of al-Andalus, and not for readers or writers of Classical Arabic.

These two dictionaries were consulted exhaustively by Franciscus Raphelengius, the son-in-law of the printer Christophe Plantin. Just as Scaliger had done, Raphelengius too studied Arabic under Postel in Paris. After working with the team of scholars who produced the Antwerp Polyglot Bible published by Plantin between 1569 and 1572, he moved in 1586 from Antwerp to Leiden and was given the chair of Hebrew at the university. He also ran the printing press established in the university town by his father-in-law. Having had Arabic types cut, he and his sons became the main printers of Arabic in northern Europe.

While he was still in Antwerp Raphelengius had come into possession of the Mozarabic Latin-Arabic glossary once belonging to Postel, as well as of Pedro de Alcalá's *Vocabulista*. It was on these that he drew when he started to compile an Arabic-Latin vocabulary. Encouraged by his colleagues in Leiden, he continued his lexicographical studies, albeit with a relatively limited number of sources. These included the *Pentateuch* printed in Istanbul by Gerson Soncino in 1546 with Saadiya Gaon's Judaeo-Arabic paraphrase, the medieval translation of the *Qur'ān* published by Bibliander in 1543, the few available works printed in Arabic, and a small collection of Arabic manuscripts. After 1593, the year in which Scaliger arrived in Leiden, Raphelengius also used Scaliger's own *Thesaurus linguae arabicae*, the Latin-Arabic glossary which would never be published, but to which Scaliger added in Leiden largely on the basis of the two Spanish dictionaries belonging to Raphelengius. The extent to which the two men used one another's material shows how closely they collaborated (Hamilton 1989: 558–572).

Raphelengius's *Lexicon arabico-latinum* testifies to a considerable ability in turning the contents of the Mozarabic Latin-Arabic glossary and the Spanish-Arabic *Vocabulista* into an Arabic-Latin dictionary, sometimes rectifying mistakes in the original text and generally providing a correct version in Arabic characters. Certainly, the very nature of Raphelengius's two main sources occasionally led to mistakes – to forms which were exclusively Andalusian and to a number of grammatical errors – but, despite the mistakes, the object of his lexicon was strikingly modern. Admittedly, the wordlists in Greek, Hebrew, and Aramaic at the end, as well as the Hebrew equivalents in the text of the dictionary, were for the assistance of theologians and Hebraists. But the Latin wordlist, and the statements in the preface, indicate that he wanted his dictionary to serve scholars working in a variety of fields besides merchants, navigators, and diplomats who required the language for more practical purposes (Hamilton 1989: 572–577).

Raphelengius died in 1597. His dictionary was unfinished and, probably owing to Scaliger who was aware of its imperfections, it remained in manuscript. Only after Scaliger's death in 1609 did Raphelengius's sons, now running the family firm, consider having it completed and published. By 1611 it was in the press, but at the last moment the Raphelengius brothers turned to the young Thomas Erpenius in the hope of adding his Arabic grammar to the dictionary and of having the dictionary corrected. As Erpenius drew up his immense list of corrections he spotted the key to Arabic lexicography.

Just as the key to the European Arabic grammars consisted in the grammars by the Arabs themselves, so the key to Arabic lexicography in Europe was to be found in the monolingual Arabic dictionaries. These Erpenius was unable to obtain, but, when he stopped in Venice in his unsuccessful attempt to sail to the East, he acquired some Arabic-Turkish dictionaries which were based on the monolingual Arabic ones. "As I once started to learn Arabic in order to improve my understanding of Hebrew," he wrote in May 1612, "now I study Turkish in order to know better Arabic" (Hamilton 1989: 581). It was thanks to the Arabic-Turkish lexicon *al-'Axtarī*, compiled in 1545 and based on some of the finest monolingual Arabic dictionaries, including the 10th-century *Ṣaḥāḥ* and *Mujmal* and the 13th-century *Muġrib*, that Erpenius managed to correct so many of Raphelengius's mistakes. His corrections were added in an appendix to the published version of the dictionary which appeared in 1613 with his grammar issued separately by the same publishers (Hamilton 1989:577–584).

Erpenius had only just had time to catch a glimpse of the most important monolingual Arabic dictionary of all, the *Qāmūs* compiled in the late 14th century, before Raphelengius's *Lexicon arabico-latinum* appeared in print. Although he subsequently acquired a copy, he could make little use of it. Yet, it was the *Qāmūs* which, together with the *Ṣaḥāḥ*, would be the main source of the next Arabic-Latin dictionary, the *Thesaurus linguae arabicae* published in Milan in 1632 and compiled in the course of 18 years by Antonio Giggei. Giggei had consulted the very considerable collection of Arabic manuscripts assembled for the Ambrosian Library by his patron, the archbishop of Milan, Cardinal Federico Borromeo, and his vast work in four volumes was indeed an improvement on Raphelengius, lacking comparisons with other Semitic languages or with Greek, and providing a far greater range of words.

Giggei's work, however, was to be superseded by that of Erpenius's successor as professor of Arabic at Leiden, Jacobus Golius. In the introduction to his *Lexicon arabico-latinum* (Leiden 1653), Golius admittted that Giggei had preceded him in the use of the monolingual dictionaries, but he himself consulted a far wider range, which included the main Arabic-Turkish and Arabic-Persian lexicons. If Giggei had freed Arabic of the assocation with Hebrew, Golius, with his Persian and Turkish sources, connected it with the other main Islamic languages (Fück 1955:79–84).

Golius's dictionary, like Erpenius's grammar, remained unsurpassed until the 19th century, when Georg Wilhelm Freytag, and above all Edward Lane, improved on it (Fück 1955:166–170). Yet, in contrast to that of Raphelengius and his predecessors, Golius's dictionary was solely of Classical Arabic, with no attempt to include current or dialectal forms. Invaluable for readers, particularly of poetry and early Islamic texts, it was of little use to merchants or travelers who needed to speak the language of the streets. Not until 1881 were the Mozarabic lexicon and Pedro de Alcalá's *Vocabulista* appreciated as guides to a particular form of spoken Arabic, when R.P.A. Dozy, in his *Supplément aux dictionnaires arabes*, perceived their full value and, implicitly, rendered homage to the work of Raphelengius (Hamilton 1989:577).

4. The schools

Some form of translators' school seems to have existed in Toledo since the mid-12th century, and was followed by other foundations in medieval Spain (Burnett 1995). The growing interest in Arabic in the early modern period, however, is more clearly documented by the creation of academic chairs. The French King François I established a chair in Arabic for Guillaume Postel at his own humanist foundation, the Collège royal in Paris, in 1538. Very gradually other European academies followed suit. Leiden University took the decision to set up a chair in 1599 (Brugman and Schröder 1979:3–4). A chair was endowed at Cambridge in 1632 and at Oxford in 1636 (Toomer 1996:111–114). But while Arabic was being studied at the European universities and academies mainly by theologians, alternative forms of instruction were also being organized for those students who wanted to have a more practical knowledge of the Eastern languages.

Arabic was taught, interruptedly and to varying standards, in the schools and seminaries intended for missionaries. The first was probably the Dominican school in Tunis founded some time before 1250. It was succeeded by schools in Barcelona, Murcia, Valencia, and Játiva, while the Franciscan Raymond Lull set up a school at Miramar in 1276 (Cortabarria Beitia 1970). By the end of the 16th century Arabic could be learnt by prospective missionaries in Rome not only in some of the houses of the various religious orders and at the Collegio Romano, but also at the foundations of Pope Gregory XIII, the Neophyte College (established in 1577 for converts from Judaism and Islam) and the Maronite College (founded in 1584 and run by the Jesuits). It would later be taught, too, at the Collegio Urbano, founded by Pope Urban VIII in 1627. One of the main purposes of the colleges was to attract young Eastern Christians, particularly Maronites, who had been in formal communion with Rome since 1182, as well as Jacobites, Copts, Nestorians, Melkites, Armenians, and, in the 18th century, Greek Catholics and other members of the Uniate churches. Generally aged between 9 and 15, they had received hardly any formal education in the East, and it was in Rome that they obtained intensive instruction, in Italian and Latin, but also in Classical Arabic, Syriac, and Hebrew. Having converted to Roman Catholicism they might either remain in Europe where they often acted as librarians and as language teachers, or return to their home country to propagate the Catholic faith (Heyberger 1994:408–423).

The other institutions for teaching a more practical type of Arabic than was to be obtained at a university were the interpreters' schools. In the early modern period they were originally devised by the Venetians who, in 1551, sent two of their young citizens to Istanbul to study the spoken languages of the Middle East (Palumbo Fossati Casa 1997). Although the project was not immediately successful, it would be imitated. In 1669, thanks to Louis XIV's minister Colbert, the French organized a school in Istanbul for the teaching of Arabic, Turkish, and Persian (Pippidi 1997). Sixty years later the French emulated the Roman system by introducing a school which was to train Eastern Christians as missionaries at the Jesuit Collège de Louis-le-Grand in Paris. Such practical endeavors would lead, after many vicissitudes, to the creation in Paris of the Ecole spéciale des langues orientales in 1795. The Venetians set up a similar institution in Padova in 1699, and other European countries followed them, the Austrians with the Orientalische Akademie in

Vienna in 1745 and the English with the College of Fort William in Calcutta in 1800.

Yet, despite the growing number of institutions at which Arabic was being taught from the early 17th century onwards, most of the best Arabists were largely self-taught. They studied the language independently of any institution and often on the spot. By the first decades of the 17th century both Holland and England had permanent diplomatic representation in Istanbul and consulates scattered over the Arabic-speaking world. The greatest of the Dutch Arabists, Golius, had indeed been a pupil of Erpenius (who had himself learnt his Arabic from Arab travelers in France), but his true progress in the language was made when he accompanied a diplomatic mission to Morocco, and when he spent a far longer period first at the Dutch consulate in Aleppo and then at the embassy in Istanbul (Juynboll 1931:119–183). In England Edward Pococke, who occupied the first chair of Arabic at Oxford in 1636, had had some tuition from William Bedwell in London, but he perfected his knowledge of Arabic when he acted as chaplain to the Levant Company at the English consulate in Aleppo and, later, at the English embassy in the Ottoman capital (Toomer 1996:116–126). And the Italian Lodovico Marracci, who published an edition of the *Qur'ān* in Arabic and Latin in 1698 infinitely superior to all previous efforts, never set foot outside Italy but learnt Arabic on his own in Rome, occasionally asking the advice of a Maronite (Levi della Vida 1959:196). This tendency was to become increasingly marked with the decline of the academic teaching of Arabic in the 18th and 19th centuries.

Bibliographical references

Balagna Coustou, Josée. 1989. *Arabe et humanisme dans la France des derniers Valois*. Paris: Maisonneuve et Larose.

Brugman, Jan and Frank Schröder. 1979. *Arabic studies in the Netherlands*. Leiden: E.J. Brill.

Burnett, Charles. 1995. "The institutional context of Arabic-Latin translations of the Middle Ages: A reassessment of the 'school of Toledo'". *Vocabulary of teaching and research between Middle Ages and Renaissance. Proceedings of the Colloquium London, Warburg Institute, 11–12 March 1994*, ed. Olga Weijers, 214–235. Turnhout, Belgium: Brepols.

Cortabarria Beitia, A. 1970. "L'étude des langues au moyen âge chez les Dominicains: Espagne, Orient, Raymond Martin." *Mélanges de l'Institut Dominicain d'Etudes Orientales du Caire* 10.189–248.

Dannenfeldt, Karl H. 1955. "The Renaissance humanists and the knowledge of Arabic". *Studies in the Renaissance* 2.96–117.

Fück, Johann. 1955. *Die arabischen Studien in Europa bis in den Anfang des 20. Jahrhunderts*. Leipzig: O. Harrassowitz.

Hamilton, Alastair. 1985. *William Bedwell the Arabist 1563–1632*. Leiden: E.J. Brill.

——. 1989. "'Nam tirones sumus': Franciscus Raphelengius' *Lexicon Arabico-Latinum* (Leiden 1613)". *Ex officina Plantiniana. Studia in memoriam Christophori Plantini (ca. 1520–1589)*, ed. Marcus de Schepper and Francine de Nave, 557–589. Antwerp: Vereeniging der Antwerpschen Bibliophielen.

Haskins, C.H. 1924. *Studies in the history of medieval science*. Cambridge, Mass.: Harvard University Press.

Heyberger, Bernard. 1994. *Les chrétiens du Proche-Orient au temps de la réforme catholique (Syrie, Liban, Palestine, XVII^e^–XVIII^e^ siècles)*. Rome: Ecole Française de Rome.

Hitzel, Frédéric (ed.). *Istanbul et les langues orientales. Actes du colloque organisé par l'IFÉA et l'INALCO à l'occasion du bicentenaire de l'Ecole des langues orientales, Istanbul 29–31 mai 1995*. Paris: L'Harmattan.

Jones, John Robert. 1988. *Learning Arabic in Renaissance Europe (1505–1624)*. Ph.D. diss., School of Oriental and African Studies, University of London.

Juynboll, W.M.C. 1931. *Zeventiende-eeuwsche beoefenaars van het Arabisch in Nederland*. Utrecht: Kemink en Zoon.

Klein-Franke, Felix. 1980. *Die klassische Antike in der Tradition des Islam*. Darmstadt: Wissenschaftliche Buchgesellschaft.

Koningsveld, Pieter Sjoerd van. 1977. *The Latin-Arabic glossary of the Leiden University Library*. Leiden: New Rhine Publishers.

Levi della Vida, Giorgio. 1959. *Aneddoti e svaghi arabi e non arabi*. Milan and Naples: Riccardo Ricciardi.

Palumbo Fossati Casa, Isabella. 1997. "L'école vénitienne des 'giovani di lingua'" Hitzel (1997: 109–122).

Pippidi, Andrei. 1997. "Drogmans et enfants de langue: La France de Constantinople au XVII^e^ siècle". Hitzel (1997:131–140).

Seybold, C.F. 1900. *Glossarium Latino–Arabicum*. Berlin: Emil Felber.

Southern, R.W. 1962. *Western views of Islam in the Middle Ages*. Cambridge, Mass.: Harvard Univer-sity Press.

Toomer, G.J. 1996. *Eastern wisdom and learning: The study of Arabic in seventeenth-century England*. Oxford: Oxford University Press.

Alastair Hamilton
(School of Advanced Studies, London University
attached to the Warburg Institute)

ʿArabiyya

For Arabic speakers *al-ʿarabiyya* is a cover term which refers to Arabic in its various forms, both synchronically and diachronically. In particular, it designates what the Arabs call *fuṣḥā* Arabic (lit. 'pure, clear, or universally intelligible', → *faṣīḥ*) and the wide range of dialects – called *ʿāmmiyyāt*, *lahajāti*, or in North Africa the *dārija* – which are the true mother tongues of Arabic speakers. In Arabic folk-linguistics, the dialects are treated as deviations from the norm – or the 'measure/standard' – which the *fuṣḥā* provides in its capacity as the codified language and as the vehicle of religion and high culture. In attitudinal terms, therefore, the 'defective' and, as often mistakenly claimed, 'grammarless' dialects are treated as valid targets for the purifying impulses of verbal hygienists (Cameron 1995). But, paradoxically, these dialects are also socio-psychologically internalized as sites of allegiance and self-definition, through which speakers express their feelings of intimacy as well as personal, local, or regional identities. In an early study on the subject, Ferguson (1959a) suggested that Arabic speakers express loyalty to their dialects, believing them to be the closest to the *fuṣḥā* and the easiest to learn, but that these positive attitudes pale into insignificance when compared with the veneration the Arabs display towards the *fuṣḥā*. This duality (*izdiwājiyya*) of language forms and attitudes in popular conception was the subject of heated debates among Arab intellectuals in the 19th century (Daniels 2002; Suleiman 2003), long before the term → 'diglossia' was used by the French Arabist William Marçais in 1930 and, later, by Charles Ferguson, whose seminal article "Diglossia" (1959b) gave it universal currency.

It is now generally agreed that the *fuṣḥā* and the dialects represent the end points of a variation continuum (Badawī 1973; Holes 1995; Versteegh 1997), but it is worth pointing out that, in the native linguistic-*cum*-intellectual tradition, little recognition is accorded to the taxonomies Western Arabists use to describe the *diachronic* variability of the language. Terms such as → 'Proto-Arabic', → 'Old Arabic', 'Early Arabic', → 'Classical Arabic', 'Early Middle Arabic', 'New Arabic', 'Muslim → Middle Arabic', → 'Christian Middle Arabic' or → 'Judaeo-Arabic', which Western Arabists use as tools of historical categorization, are given short shrift in modern thinking about the language, which prefers to highlight the diachronic continuities and synchronic overlaps in Arabic rather than to dwell on what are regarded as typologies of difference and 'fragmentation'. This unifying orientation in the modern Arab world, as a trope of academic practice and as a truism of popular thinking, is not without socio-political meaning. It is part and parcel of the attitude toward the *fuṣḥā* as the mainstay of pan-Arab unity, at least on the cultural level (Suleiman 1994).

The view that the *fuṣḥā* is superior to other languages has a long and established pedigree in Arab culture. It is found in the writings of the medieval grammarians and rhetoricians, for example ʾAbū Manṣūr aṯ-Ṯaʿālibī (d. 430/1038-9) and Ibn Sinān al-Xafājī (d. 460/1067–1068). With few exceptions, the Arabic grammarians showed little interest in other languages, even when they could speak them – for example, Sībawayhi (d. 180/796), as-Sijistānī (d. 255/869) and ʾAbū ʿAlī al-Fārisī (d. 377/987) who were of Persian origin and were fluent in Persian – believing them, explicitly or implicitly, to be inferior to Arabic. Ibn Jinnī (d. 392/1002), who was not of Arab origin, associated the excellence of Arabic with the principle of *ḥikmat al-ʿArab* 'the wisdom of the Arabs', which, in modern discourses on nationalism, can be equated with ethnolinguistic election as an emblem of distinction of the nation (Suleiman 2002). In the 20th century, the Syrian Zakī al-ʾArsūzī developed this principle into his main dogma of Arab nationalism, declaring time after time in his writings that the "genius of the Arabs resides in/emanates from their language" (1972–1976).

On a different level, this attitude towards the *fuṣḥā* is associated with its exalted position as the language of the Islamic revelation (*tanzīl*) verbatim, expressed through the *Qurʾān* in its capacity as God's inimitable word. It is impossible to exaggerate the role of Islam in the development of the Arabic language and in shaping the attitudes toward it. Addressing the poetically-minded Arabs, God tells them that He revealed the *Qurʾān* in Arabic to challenge them to produce one that can match it in excellence. In Islamic theology and Arabic rhetoric, this challenge was embodied into the principle of → *ʾiʿjāz al-Qurʾān* ‘the inimitability of the *Qurʾān*’, which is an article of faith for the Arab and non-Arab Muslims. The language of the *Qurʾān* is said to be *mubīn* ‘perspicuous’. According to one *ḥadīṯ* (a saying of the Prophet), Arabic is said to be ‘the master of speech’; another *ḥadīṯ* declares that Arabic is the language of Heaven (*janna*). We must, however, not conclude from this and other examples from the same sources that the *fuṣḥā* is a sacred or liturgical language. *Fuṣḥā* Arabic is used for religious purposes and for recording and transmitting sacred material, but it is also deployed in non-religious domains, some of which in fact are the very antithesis of the ‘sacred’ or the ‘liturgical’ (cf. Haeri 2003).

The rise of Islam and the ensuing conquests propelled Arabic to pre-eminence on the world stage. These defining events constitute the ‘golden age’ to which modern-day Arabs return linguistically, for assurance and inspiration, when they feel that their language and what it stands for are under attack (Suleiman 2004). But the connection between Islam and Arabic goes far deeper. Islam expanded the functional domains of Arabic, led to its lexical development through borrowing and semantic expansion, occasioned its orthographic and grammatical codification, turned it into the vehicle of learning and high culture, and made it the medium of government and diplomacy. These gains were consolidated during the early centuries of Islam, but the fortunes of the *fuṣḥā* started to wane when control in the Islamic empire started to pass to the non-Arabs in the 3rd/9th century.

The Islamic conquests were the engine that led to mixing and linguistic leveling among the Arab tribes which, until then, had not pursued common goals on such a spectacular scale. Waves of migration into the conquered territories, coupled with intercommunal life in the newly established settlements, brought the Arabic speakers into sustained contact with speakers of other languages (Versteegh 1984). Mixed marriages intensified this contact. The combination of these factors led to linguistic practices, invariably described as → *laḥn* ‘solecism’ in the Arabic grammatical tradition, in which the mixing of codes ushered in new developments that threatened the presumed ‘pristine purity’ of the *fuṣḥā*. The attitude toward the modern dialects as ‘deviations’ from the norm is no doubt socio-psychologically linked to the attitude towards *laḥn* at this early stage in the expansion of the Arab controlled lands and the attendant spread of the language. It is therefore not surprising that the language guardians use this ideologically loaded term to describe modern ‘deviations’ from the *fuṣḥā*, regardless of their source or provenance. The injection of this term into modern sites of linguistic debate and conflict is another example of the drive to emphasize historical ‘continuity’ in conceptualizations of the *fuṣḥā* in the Arab world.

The position of Arabic as the language of government and diplomacy was dealt a body blow when the hold of the Abbasid caliphs on power started to wane from the 3rd/9th century onward. Under the Samanids in Iran, in the 10th century, Persian replaced Arabic as the language of culture and government. With the rise of the Seljuk Turks to power and their expansion into Anatolia to the west (5th–7th/11th–13th centuries), Turkish replaced Arabic as the official language of the state. The fall of Baghdad at the hands of the Mongols in 657/1258 divested Arabic of any prestige it might have had as the language of administration and high culture. The fall of Granada to the Castilian kings in 897/1492 put an end to the presence of Arabic in the Iberian Peninsula. In Egypt, under the Mamluks (648–923/1250–1517), Turkish was the language of the ruling elite and government. The Ottomans (1517–1918) conducted the affairs of the state and encouraged expressions of high culture in Turkish. And yet, in spite of all the pressure Arabic had come under over the centuries it never lost its pre-eminence as the language of religion and its allied sciences (jurisprudence and theology) in the lands of Islam. *Fuṣḥā* Arabic was used in literary production in these lands, but it never had the prestige it enjoyed in its heyday when the Arabs formed the ruling class.

In the 18th century, attempts were made to breathe a new life into Arabic. The Maronite archbishop of Aleppo, Germanus Farḥāt (d. 1145/ 1732) led the way by, among other things, encouraging education in Arabic and writing books on grammar, lexicography, and literature. However, the real impetus for change in the fortunes of Arabic came as a result of Napoleon’s expedition into Egypt in 1798 and Muḥammad ʿAli’s rise to power in the country in 1805.

As viceroy of Egypt, Muḥammad ʿAli (d. 1849) set out to modernize the country by building a strong administration and army that could withstand the pressure of the European powers and pursue his ambitions against the Ottoman sultans. To achieve this, he dispatched educational missions to France to acquire the new sciences and to transplant them into Egypt (Heyworth-Dunne 1939). He encouraged translation, and, for this purpose, established the famous *Dār al-ʾalsun* (School of Languages), whose task was to transfer into Arabic legal, medical, engineering, military, and other works (aš-Šayyāl 1951; Tājir 1945). Būlāq, the famous printing press, was established to serve this endeavor. The official *al-Waqāʾiʿ al-Miṣriyya* (Egyptian gazette), which was originally published in Turkish only, started to appear in a bilingual edition, with Arabic on the right hand side of the page as if to signal its dominance over Turkish. Later, Turkish was dropped and the gazette started to appear in Arabic only. These developments, coupled with the spread of education, albeit at a modest rate at this stage, invigorated Arabic and jolted it into contact with the modern world. In the course of the 19th century, the expansion of the literary arts, particularly the theater, and the popularity of the press provided enhanced opportunities for the revitalization of Arabic.

The 19th century witnessed similar developments in the Levant. In fact, contact with Europe in this region predated Napoleon's invasion of Egypt. The Maronites of Lebanon, through their centuries-old relations with the Vatican in Rome, played a leading role in invigorating Arabic. In the 19th century, European and American missionaries established schools in the Levant, the most famous of which was the Syrian Protestant College (1863) – the predecessor of the American University of Beirut – which, until 1882, taught all the disciplines, including medicine, through the medium of Arabic. Translations of the Bible as well as a wide range of books and other materials appeared in the language (Zaytūnī 1983). The increasing popularity of the press worked to the advantage of Arabic. More and more people came into its expanding orbits of expression, which had to accommodate readers' expectations for a more accessible style that shifted focus to content and meaning. Books and articles on Arabic grammar, rhetoric, style, and lexicon began to appear. Arabic started to be used in new literary genres, for example the novel and the short story. The net effect of these and other developments was to propel the language into an ever-expanding range of communicative needs. Arabic became a tool of modernization, but it was at the same time made into an object of modernization in its own right.

In addition, Arabic acquired new ideological meanings. In the struggle for autonomy within the Ottoman Empire, Arabic was used as a proxy to express increasingly politicized demands and irredentist tendencies, in some cases bordering on outright independence from the Ottoman Turks (Suleiman 2003). One of the most important achievements of this period was to reconfigure the automatic identification of Arabic with Islam. Arabic was promoted as the language of the Arabs first and foremost, regardless of their religious background. ʾIbrāhīm al-Yāzijī (d. 1906) played a crucial role in promoting this new spirit, using poetry, pamphlets, press articles, books, translations, clandestine activities, public speeches, and other forms of direct interaction with his readers, for example through the 'Letters to the Editor' section of the newspapers he edited (*al-Bayān*, *aḍ-Ḍiyāʾ*, and *aṭ-Ṭabīb*). In the 20th century, this secular conception of Arabic found strong expression in the writings of Sāṭiʿ al-Ḥuṣrī (d. 1968), the most famous ideologue of pan-Arab nationalism in the 20th century (Cleveland 1971).

One of the main issues facing Arabic in the 19th and 20th centuries was the creation of new terms and other lexical resources to express the communicative needs of sociopolitical and economic modernization. Starting in the 19th century (Xūrī 1991), this issue was debated extensively by Arab scholars and linguists, who returned to the 'golden age' and its successes in transferring Greek knowledge into Arabic to assure their audiences that the language can handle the challenges of modernization in the modern period. A typology of methods for creating new vocabularies emerged, and it included the following (Stetkevych 1970): (a) borrowing terms (mainly from English and French) and integrating them into the language phonologically and morphologically; (b) semantic extension of existing words; (c) analogical creations from existing roots; (d) translation of foreign words into Arabic (*calque*); and (e) blending. Not all of these methods, for example blending (→ compounds), have succeeded to the same degree. One area where Arabic word creation still has not been successful is in the provision of acronyms

(→ abbreviations), which are a frequent feature of the two main languages with which Arabic is in constant contact: English and French. Another area where problems have emerged is → technical terminologies. The problem here is not one of generating these terminologies, but their multiplicity and differentiated use, which create problems of comprehension in the scientific exchange among Arabic speakers.

Contact with English and French has been the source of stylistic innovations and code-switching practices in Arabic. The Arab press has acted as the channel through which many stylistic innovations are introduced and popularized in Arab culture. The process started in the 19th century, but it intensified in the 20th century to such an extent that one can confidently talk about the emergence of new styles of writing in Arabic. The same is true of → code-switching, which was present in the dramatic work of the Egyptian Yaʿqūb Ṣannūʿ in the 19th century (Saʿīd 1964; Moosa 1983), and was also the butt of attacks by the Arab nationalists and purists in the Levant in their effort to make Arabic the criterion of their national identity (Suleiman 2003).

In the modern Arab world, code-switching has emerged as a topic of intense debate. As a style of speaking among young Arabic speakers (by no means all young speakers), code-switching as a form of language display is practiced as a sign of sophistication, modernity and, therefore, prestige. It is particularly popular among women who, some argue (aḏ-Ḏwādī 1996), use it to cull a symbolic verbal space in which they can escape from the oppressive 'patriarchal' norms of Arab society. However, on the whole, negative values are read into code-switching, even by the code-switchers themselves (Bentahila 1983). Some Arabs treat code-switching as an expression of inferiority and post-coloniality (aḏ-Ḏwādī 1986). Others think of it as a sign of social affectation or, worse still, as a way through which the speaker 'tarts up' his speech for dubious sociolinguistic gains (al-ʾAbṭaḥ 2001). Yet, these negative attitudes do not seem to depress the appetite of young Arab speakers, particularly women, for code-switching with either English or French.

In spite of the attempts that have been made to promote Arabic in the past century, including the provision of universal education, most Arabs believe that standards in the language have been falling, although the evidence for this tends to be anecdotal (al-Mūsā 1990). Empirical studies reveal, however, that Arab students show little enthusiasm to specialize in the language at university level (al-Frēj 1993; al-ʿAnān 1994), and that they in fact ascribe this to their weak standards in the language. The perception that standards in the language are falling has been attributed by many scholars to, among other things, the complexity of pedagogic grammars and the outdated mode of their presentation and teaching. Even before the discovery of Ibn Maḍāʾ al-Qurṭubī's (593/1196) famous treatise *ar-Radd ʿalā n-nuḥāt* in Cairo in the 1950s, in which he strongly attacked the use of unnecessary causes (*ʿilla*s) in Arabic language teaching (Suleiman 1999a), attempts were made to recast Arabic grammar in new pedagogic moulds to suit the needs of learners (Muṣṭafā 1937/1959). Similar attempts were made by the Lebanese linguist ʾAnīs Frayḥa (1955, 1956, 1959) and the Egyptian Šawqī Ḍayf (1986a, 1986b, 1990). Until recently, the calls for the simplification of Arabic (*tabsīṭ* or *taysīr*) have remained part of the academic 'wish list' for the language. However, important steps have been taken in some Arab Gulf countries recently, notably Qatar, to implement a radical literacy-based reform of Arabic language teaching, applying benchmarking comparisons and a curriculum-standards based approach (al-Majlis al-ʾAʿlā li-t-Taʿlīm 2004), similar to those that have been developed for teaching English in England, the United States, and New Zealand.

As the official language of over 20 countries in the Middle East and North Africa, and as the language of well over 200 million users in these countries and elsewhere in the world, the future of Arabic is assured (Suleiman 1999b). However, in areas where it is a minority language, for example in → Israel and → Turkey, Arabic is much weaker. In Turkey, Arabic is facing linguicide. It is in a more precarious position than Kurdish. In Israel, Arabic is the object of benign neglect. It has been aptly described by Ben-Rafael (2001) as the "non-prestigious language of the underprivileged". The position of Arabic in Israel and Turkey reminds us very strongly that the status of Arabic as a regional and world language is uneven. It has demographic strength and cultural depth, but in some contexts it is barely able to survive. Furthermore, globalization means that Arabic is facing competition from the world languages, mainly English and, to a lesser extent, French, which are promoted as the languages of instruction for medicine and the sciences in an expanding coterie of schools and universities (Abu Absi 1997; Shaaban 1997). The popularity of private education in many Arab countries has given this trend a real boost over the past two decades.

Bibliographical references

ʾAbṭaḥ, Sawsan al-. 2001. “at-Tabarruj al-luġawī: aṭ-Ṭarīq al-ʾasraʿ ʾilā l-hāwiya”. *aš-Šarq al-ʾAwsaṭ* 13 August 2001.
Abu Absi, Samir. 1997. “Innovative second language education in the Middle East and North Afric”. *Second language education*, ed. G. Richard Tucker and David Corson, 196–205. Dordrecht: Kluwer Academic Publishers.
ʿAnān, ʿAbdallāh al-. 1994. “al-Luġa al-ʿarabiyya wa-l-ʾibdāʿ al-fikrī aḏ-ḏātī”. *al-ʾIbdāʿ al-fikrī aḏ-ḏātī fī l-ʿālam al-ʿarabī*, ed. by ʾAnwar ʿAbd al-Malik and Xaldūn an-Naqīb, 62–85. Cairo: Jāmiʿat al-ʾUmam al-Muttaḥida wa-l-Hayʾa al-Miṣriyya al-ʿĀmma li-l-Kitāb.
ʾArsūzī, Zakī al-. 1972–1976. *al-Muʾallafāt al-kāmila*. Damascus: Maṭbaʿat al-ʾIdāra as-Siyāsiyya li-l-Jayš wa-l-Quwwāt al-Musallaḥa.
Badawī, as-Saʿīd Muḥammad. 1973. *Mustawayāt al-ʿarabiyya al-muʿāṣira fī Miṣr*. Cairo: Dār al-Maʿārif.
Ben-Rafael, Eliezer. 2001. “A sociolinguistic paradigm of bilingualism: English, French, Yiddish, and Arabic in Israel”. *Language and communication in Israel*, ed. Hanna Herzog and Eliezer Ben-Rafael, 175–205. New Brunswick and London: Transaction Publications.
Bentahila, Abdelali. 1983. *Language attitudes among Arabic–French bilinguals in Morocco*. Clevedon: Multilingual Matters.
Cameron, Deborah. 1995. *Verbal hygiene*. London and New York: Routledge.
Cleveland, William L. 1971. *The making of an Arab nationalist: Otomanism and Arabism in the life of Satiʿ al-Husri*. Princeton, N.J.: Princeton University Press.
Daniels, Helge. 2002. *Debating variablity in Arabic:* Fuṣḥā *versus* ʿāmmiyya. Ph.D. diss., University of Antwerp.
Ḍayf, Šawqī. 1986a. *Tajdīd an-naḥw*. Cairo: Dār al-Maʿārif.
——. 1986b. *Taysīr an-naḥw at-taʿlīmī qadīman wa-ḥadīṯan maʿa nahj tajdīdihi*. Cairo: Dār al-Maʿārif.
——. 1990. *Taysīrāt luġawiyya*. Cairo: Dār al-Maʿārif.
Ḏwādī, Maḥmūd aḏ-. 1986. “al-Mazj al-luġawī ka-sulūk luġawī li-l-ʾinsan al-maġribī al-maġlūb”. *al-Majalla al-ʿArabiyya li-l-ʿUlūm al-ʾInsāniyya* 6.46–66.
——. 1996. “al-Franco-arabe al-ʾunṯawiyya al-maġāribiyya ka-sulūk iḥtijājī ʿalā l-lāmusāwā maʿa r-rajul wa-ka-ramz li-kasb rihān al-ḥadāṯaʾ. *Dirāsāt ʿArabiyya* 3/4.81–89.
Ferguson, Charles. 1959a. “Myths about Arabic”. *Languages and Linguistic Monograph Series* (Georgetown University) 12.75–82.
——. 1959b. ‘Diglossia’. *Word* 15.325–340.
Frayḥa, ʾAnīs. 1955. *Naḥwa ʿarabiyyatin muyassara*. Beirut: Dār aṯ-Ṯaqāfa.
——. 1956. *Yassirū ʾasālīb taʿlīm al-ʿarabiyya: Hāḏā ʾaysar*. Junyeh: Maṭābiʿ Karīm.
——. 1959. *Tabsīṯ qawāʿid al-luġa al-ʿarabiyya ʿalā ʾusus jadīda: Iqtirāḥ wa-namūḏaj*. Beirut: American University of Beirut.
al-Frēj, Sihām. 1993. *ʿUzūf ṭalaba jāmiʿat al-Kuwayt ʿan at-taxaṣṣuṣ fī l-luġa al-ʿarabiyya*. Kuwait: Jāmiʿat al-Kuwayt.
Haeri, Niloofar. 2003. *Sacred language, ordinary people: Dilemmas of culture and politics in Egypt*. London: Palgrave Macmillan.
Heyworth-Dunne, James. 1939. *An introduction to the history of education in Modern Egypt*. London: Luzac.
Holes, Clive. 1995. *Modern Arabic: Structures, functions and varieties*. London and New York: Longman.
Majlis al-ʾAʿlā li-t-Taʿlīm al-. 2004. *Maʿāyīr al-manāhij at-taʿlīmiyya li-dawla Qaṭar: al-Luġa al-ʿarabiyya min riyāḍ al-ʾaṭfāl ḥattā ṣ-ṣaff aṯ-ṯāniya ʿašar*. Doha, Qatar.
Moosa, Matti. 1983. *The origins of modern Arab fiction*. Washington, D.C.: Three Continents Press.
Mūsā, Nihād al-. 1990. *al-Luġa al-ʿarabiyya wa-ʾabnāʾuhā: ʾAbḥāṯ fī qaḍiyyat al-xaṭaʾ wa-ṣ-ṣawāb wa-ḍaʿf aṭ-ṭalaba fī l-luġa al-ʿarabiyya*. Amman: Maktaba Wisām.
Muṣṭafā, ʾIbrāhīm. 1959. *ʾIḥyāʾ an-naḥw*. Cairo: Lajnat at-Taʾlīf wa-t-Tarjama wa-n-Našr. (1st ed. 1937.)
Saʿīd, Naffūsa Zakariyyā. 1964. *Tārīx ad-daʿwā ʾilā l-ʿāmmiyya wa-ʾāṯāruhā fī Miṣr*. Cairo: Maṭbaʿa Dār Našr aṯ-Ṯaqāfa.
Shaban, Kassim Ali. 1997. “Bilingual education in Lebanon”. *Bilingual education*, ed. Jim Cummins and David Corson, 196–205. Dordrecht: Kluwer Academic Publishers.
Šayyāl, Jamāl ad-Dīn aš-. 1951. *Tārīx at-tarjama wa-l-ḥaraka aṯ-ṯaqāfiyya fi ʿaṣr Muḥammad ʿAlī*. Cairo: Dār al-Fikr al-ʿArabī.
Stetkevych, Jaroslav. 1970. *The modern Arabic literary language: Lexical and stylistic developments*. Chicago and London: University of Chicago Press.
Suleiman, Yasir. 1994. “Nationalism and the Arabic language: An historical overview”. *Arabic sociolinguistics: Issues and perspectives*, ed. Yasir Suleiman, 3–24. Richmond, Surrey: Curzon Press.
——. 1999a. *The Arabic grammatical tradition: A study in* taʿlīl. Edinburgh: Edinburgh University Press.
——. 1999b. “Language education policies: Arabic speaking countries’. *Concise encyclopedia of educational linguistics*, ed. Bernard Spolsky, 106–116. Oxford: Elsevier and Pergamon.
——. 2002. “*Bayān* as a principle of taxonomy: Linguistic elements in Jāḥiz’s thinking”. *Studies on Arabic in honour of G. Rex Smith*, ed. J.F. Healy and Venetia Porter, 273–295. Oxford: Oxford University Press.
——. 2003. *The Arabic language and national identity: A study in ideology*. Edinburgh: Edinburgh University Press.
——. 2004. *A war of words: Language and conflict in the Middle East*. Cambridge: Cambridge University Press.

Tājir, Jāk. 1945. *Ḥarakat at-tarjama bi-Miṣr xilāl al-qarn at-tāsiʿ ʿašar*. Cairo: Dār al-Maʿārif.
Versteegh, Kees. 1984. *Pidginization and creolization: The case of Arabic*. Amsterdam: J. Benjamins.
——. 1997. *The Arabic language*. Edinburgh: Edinburgh University Press.
Xūrī, Yūsuf Qazmā (ed.). 1991. *Najāḥ al-ʾumma al-ʿarabiyya fī luġatihā al-ʾaṣliyya*. Beirut: Dār al-Ḥamrāʾ.
Zaytūnī, Laṭīf. 1983. *Ḥarakat at-tarjama fī ʿaṣr an-nahḍa*. Beirut: Dār an-Nahār.

Yasir Suleiman (University of Edinburgh)

Aramaic/Syriac Loanwords

1. Arabic and Aramaic

During the first half of the 1st millennium B.C.E. Aramaic dialects spread from their original home around the Upper Euphrates (Aram Naharayim) into Syria and Mesopotamia. By the time the Achaemenid Empire was established in the 6th century B.C.E. most areas of Syria-Palestine and Mesopotamia were Aramaic-speaking. Different forms of Aramaic became the dominating spoken language of these areas until the Islamic conquest. It is, however, likely that during this period different forms of Arabic could already be heard in some regions. According to documents, there was a substantial presence of people with Arabic rather than Akkadian names in central and lower Mesopotamia from the late Assyrian period into the Achaemenid times. In Syria we hear about the presence of people called → 'Arabs' not only in the border regions to the desert but also in the Anti-Lebanon, the Biqāʿ valley and around Ḥimṣ/ Emesa in Seleucid and Roman times. Some of the rulers of these Arabs have Arabic-sounding names and it can be assumed that there were speakers of a variety of Arabic among them. Finally, in the Arabo-Nabatean kingdom there was interaction between the users of late Imperial Aramaic as a written language and large groups of speakers of Arabic dialects. The interference between Aramaic and different forms of Arabic is thus most likely to have existed more than one millennium before the Islamic conquest. During the first two centuries of Islam, Aramaic continued to be spoken in Syria and Mesopotamia by the peasantry. They were called *nabaṭ*, a word which in early Arabic sources also means Aramaic-speaker. In the cities, a bilingual situation arose soon after the conquest when Arabic increasingly became the language of the marketplace and public life, whereas Aramaic continued to be spoken at home. With the growing physical separation between the different religious groups in separate quarters within the city walls, from the Crusader period and onwards, Aramaic tended to be limited especially to Christian and Jewish quarters. The religious minorities have tended to preserve the bilingual situation to a larger degree than the Muslim majority, which early on seems to have been Arabicized. Aramaic-speaking Jewish and Christian communities surrounded by Arabic-speakers have existed until the 20th century, especially in the old cities of Iraq. The Aramaic-speaking Jewish communities migrated to Israel in the beginning of the 1950s. In the countryside, Aramaic was preserved in certain areas until quite recently. In Lebanon, Aramaic was generally spoken in the northern Christian mountain villages until the beginning of the 18th century and is still alive in the three villages of Maʿlūla, Baxʿa, and Jubb ʿAdīn in the Anti-Lebanon. In Mesopotamia, Aramaic is still spoken in villages around Mosul and further north. In the south, Aramaic was probably spoken by Mandaeans until fairly recent times. In general, it can be said that the Arabization process has been faster in cities and among Muslims than in the countryside and among religious minorities. A bilingual Arabic-Aramaic situation has probably existed in many areas for a very long time but unfortunately this is poorly documented (Hopkins 1995:37–38).

The interference between Aramaic and Arabic is a much more complex phenomenon than is usually realized. Thus, we have to take into account not only the bilingual situation in many areas during the Islamic period, but also the close contacts between Aramaic- and Arabic-speakers before the Islamic conquests. Further, Aramaic is far from being one unified language. There are substantial differences between the western and eastern dialects, documented as early as the turn of the era, and also within these groups, especially the eastern one. To this is added the preservation among the Aramaic-speaking minorities of traditional literary Aramaic idioms used in religious ceremonies and formal speech. For the Christians, Eastern Aramaic Syriac has been of great importance whereas the Jews have had a mixed Eastern and Western Aramaic literary tradition. The influence of Classical Mandaean upon the spoken language of the Mandaeans is more difficult to grasp. A final problem which should be taken into account is the typological similarity between Aramaic in general and the modern spoken forms of Arabic (Fischer 1984:83), which sometimes makes it difficult to recognize borrowings. It has been suggested (Retsö 2000) that some Arabic dialects, mostly in North Africa, are in fact descendants of dialects spoken in the border regions between Syria and Arabia, originally sharing many morphological and most likely also lexical features with the Aramaic complex.

An important question in connection with the borrowing from Aramaic into Arabic is which kind of Aramaic the borrowing reflects, which also may be an indication about the age of the borrowing. Aramaic has one main phonological shift which distinguishes it from Arabic, viz. the *begadkefat* shift. This means the fricativization of the plosives *b*, *g*, *d*, *k*, *p*, *t* to *ḇ*, *g*, *ḏ*, *ḵ*, *p̱*, *ṯ* when non-geminated in postvocalic position. The *begadkefat* has not affected all consonants in all dialects and its distribution shows a very complicated picture. Two other sound shifts are of importance even if they can also be found in certain forms of modern spoken Arabic: (a) the reduction of short vowels in unstressed open syllables to a murmur vowel or to zero, (b) the change *ā* > *o*(*ō*). Of these (a) is a feature reflected in all Aramaic dialects; (b) has affected the Aramaic dialects spoken in northern-central Syria and Mesopotamia but not those in southern Syria, including Palestine, southern Mesopotamia and the northwestern periphery. These sound shifts took place at different periods in different areas. The earliest traces of (a) are from the 2nd century B.C.E. in Mesopotamia. The *begadkefat* shift presupposes the existence of all short vowels and should thus have started earlier but the exact development of these two changes is difficult to follow.

Another factor to be taken into account is the sound changes in Arabic. Thus, common Semitic *p* is in all forms of Arabic represented by *f*. According to the *begadkefat* rules many varieties of Aramaic have *p* and *f* in complementary distribution. Further, common Semitic s^1 = *š* is in Arabic represented by *s*, whereas Semitic s^3 is *š* in Arabic but *s* in Aramaic. It should also be noticed that both Arabic and Aramaic have *t ṯ* and *dḏ* but in different distribution. All these factors make the tracing of Aramaic words in Arabic a difficult task and the difference between words inherited by Arabic and Aramaic on the one hand, and Aramaic words borrowed into Arabic on the other, is not as simple as has sometimes been assumed. Some of the collections of borrowings (Féghali 1918; Hobaika 1939; Freyha 1973; Nakhla 1973) should be used with caution.

2. ARAMAIC IN THE ʿARABIYYA

The integration of Aramaic loanwords into Arabic is reflected in the literary language, the *ʿarabiyya*, from its earliest stages. The earliest dated text is the *Qurʾān*, but several of the Aramaisms there can also be found in the poetry ascribed to the earliest poets from the 6th century C.E. Many of the most important and frequent words in the *Qurʾān* are clear Aramaic borrowings, which can be shown by a comparison with Syriac: *ʾaslam-* 'to submit [to the new religion]' < *ašlem*; *bāb* 'door', 'gate' < *bāḇā*; *bīʿa* 'church' < *biʿṯā*; *rabb* 'lord', *raḥmān* 'merciful' (most likely via South Arabian); *sabīl* 'way', 'path' < *šḇīlā*; *sabt* 'Sabbath' < *šabṯā*; *sajad-* 'prostrate' < *sgeḏ*; *safīna* 'ship' < *sfi(n)tā*; *tāb-/yatūb-* 'repent' < *tāḇ/yṯūḇ* or *nṯūḇ*; *tatbīr* 'destruction', from Aramaic *tḇar* 'break', cf. Arabic *ṯabar-* 'destroy'; *ʾasbāṭ*, pl. of *sibṭ* < *šiḇṯā* 'tribes'; *ʿālam* 'world' < *ʿālmā*; *ṣalāt* 'religious service, ceremony' < *ṣlūṯā*; *zakāt* 'alms' < *zḵūṯā*; *ʿīd* 'festival' <*ʿīḏā*; *qurbān* 'offering' < *qurbānā*; *furqān* 'salvation', 'redemption' < *purqānā*; *madīna* 'town' < *mḏi(n)tā*; *malakūt* 'kingship' < *malkūṯā*; *masīḥ* 'Christ' < *mšīḥā* (Jeffrey 1938). The Aramaic origin of these words and many others is made likely by the fact that they have no semantic cognates in Arabic from which they can be derived. Thus, for instance, *jannat-* 'garden' has no direct cognate in Arabic where the verb *janna* means 'to cover'. Aramaic *gi(n)tā*, on the other hand, is clearly formed from the root GNN 'to surround, to protect'. In this case the *ʿarabiyya* has the indigenous word *ḥadīqa* from the verb *ḥadaq-* 'to surround, to protect'. When derivations are sometimes found, it can be shown that they are derived from the loanword. Thus, the word *sūq* 'market-place' has many derivations but they are all from the noun, which must be a borrowing from Aramaic *šūqā* and then originally from Akkadian *sūqu* 'street'. In Akkadian, it is obviously connected with *siāqum* 'to be narrow' whereas Arabic *sāq-* has a completely different meaning: 'to lead', 'to conduct'. In the *ʿarabiyya* of the *Qurʾān* we also find several semantic borrowings which give homonyms like *daras-* 'to study' (from Aramaic *draš*) or 'to wipe out' (original Arabic), *katab-* 'to write' (Aramaic/Hebrew) or 'to sew together', 'to put together' (original Arabic), *dīn* 'judgment' (Aramaic *dīnā*) or 'to owe' (original Arabic), *zakā* 'to be pure, innocent' (Aramaic *zḵā*) or to 'be fit, suitable' (original Arabic), *ṣalīb* 'cross' (Aramaic *ṣlaḇ* 'to crucify'), Arabic *šalab-* 'to be hard, stiff', *ṣawm* 'fasting' (Aramaic *ṣawmā*), Arabic 'to stand still', *qaraʾ-* 'to read aloud', 'to recite' (Aramaic *qrā*) or Arabic 'to gather', 'to collect'. This vocabulary is also found in the earliest Islamic prose texts like Ibn ʾIsḥāq's history of the Prophet (Hebbo 1984). Many of them must be very old borrowings as can be seen from the many derivations according to Arabic morphological rules, e.g. of *katab* with the meaning 'to write'. In general it can be said that the Aramaic loans in the *Qurʾān* and the earliest poetry seem to reflect an archaic form of Aramaic. There are no traces of the sound shifts mentioned. The six *begadkefat* consonants are always plosives, although the phonology of the *ʿarabiyya* could have reproduced the fricative articulation of the Aramaic *begadkefat* consonants, including (the fricative) *ḇ* which could be rendered *w*, as in some Mishnaic Hebrew and early Aramaic inscriptions: *gaḇrā* > GWRʾ. Further, all instances of Aramaic *ā* are *ā* in the *ʿarabiyya*, e.g. *furqān*, thus no trace of the shift *ā* > ō. In western Syriac we have *purqōn-* but in eastern Syriac *porqān-*. The Aramaic *š* (= Semitic s^1) is always *s* in these items, which shows that these words were borrowed from Aramaic before the Arabic sound shift *š* > *s* (McDonald 1974), cf. *sabbaḥ-*, Syriac *šabbaḥ* 'to praise'. The same holds for the *p*, which in Aramaic after the *begadkefat* shift has two varieties: *p* and *f*, whereas the *ʿarabiyya* always has *f*. The vowel reduction is also absent in Arabic, cf. *sabīl-*, Syriac *šḇīl*. This does not mean that all borrowings must be from before the 2nd century B.C.E., only that the Aramaic from which the borrowings come had not been seriously affected by the sound shifts. It can be assumed that in certain cases the loanword was transformed when being integrated into the sound and syllable structure of Arabic. This especially holds for verbal borrowings where paradigmatic leveling has been at work. The verb *tāb/yatūb-* 'to repent' is an Aramaic loan, which can be seen from its meaning and from its Arabic counterpart *ṯāb-/yaṯūb* 'to turn back'. But it should be observed that Arabic has *t* in all forms of this verb whereas Syriac, for instance, has the fricative *ṯ* in the imperfect according to the *begadkefat* shift, thus *tāḇ/nṯūḇ*. The *ʿarabiyya* has either introduced the *t* in all forms analogically, or the word was borrowed from an Aramaic dialect which had not yet undergone the *begadkefat* shift. The existence of short vowels in unstressed open syllables as in this example is most likely due to an integration of the borrowing into the verbal paradigm of the *ʿarabiyya*. This does not explain, however, the total absence of traces of the Aramaic sound changes. A noun with the form *šḇīl* could very well have been borrowed into the *ʿarabiyya* as **isbīl* and an Aramaic *zaḵūṯā* should give **zaxūṯ* in Arabic.

In the approach taken here the Aramaic cognates in the *ʿarabiyya* are regarded as borrowings from Aramaic. The much further reaching claim that the *ʿarabiyya* of the *Qurʾān* is in fact a transformation of a text originally written in Aramaic or even Syriac, as claimed by Luxenberg (2000), is most difficult to verify and remains highly unlikely.

3. ARAMAIC IN ARABIC DIALECTS

The Arabic spoken in Syria and Mesopotamia has replaced Aramaic dialects there and it can be assumed that a bilingual situation existed for a long time and that numerous Aramaic lexemes found their way into Arabic during this period. The presence of Aramaic lexemes is well studied in Lebanese Arabic (Féghali 1918; Freyha 1973) and the dialects spoken in the Anti-Lebanon (Arnold and Behnstedt 1993) but can be found in dictionaries from the entire Syro-Palestinian area (cf. Barbot 1961). The material collected by Féghali and Freyha shows that, unlike in the *ʿarabiyya*, most borrowings preserve the Aramaic phonology. Thus *šawb* 'heat', Syriac *šawbā* 'summer heat'; *seġer* 'to be ignited', Syriac *sg̱ar*; *šaleḥ* 'to undress', Syriac *šlaḥ*, cf. Arabic *salax* 'to pull off'; *naṭar* 'to guard', Syriac *nṭar* cf. Arabic *naḏ̣ar-* 'to look at'; 'to observe', *labše* 'to cloth', Syriac *lḇāšā*, cf. Arabic *libs* 'clothes'; *baššaṭ* 'stretch', 'to extend', Syriac *pšaṭ*; *faram* 'to cut', Syriac *pram* (Arnold 2002). Even if most of these words can be found in Syriac, one should not assume that they are borrowed from that variety of Aramaic which is an eastern dialect (Contini 1999:102–103). It is obvious that most of the words designating everyday activities belong to a local western Aramaic dialect originally spoken in Lebanon. It should be observed that in general these words preserve Aramaic *š* and *ṭ* against Arabic *s* and *ḏ*. There are clear traces of the *begadkefat* shift of *g* also in initial position: *ġaddef* ‘to blaspheme’, Syriac *gaddef*, cf. Arabic *jaddaf-* ‘to curse’, ‘to blaspheme’. The Arabic word may in this case be an ancient borrowing from Aramaic. An example of a semantic borrowing is *ṣabeġ* ‘baptize’ which has its meaning from Aramaic *ṣḇeʿ*, but the form is Arabic *ṣabaġ-* ‘to dip’, ‘to dye’. Both these examples belong to the religious semantic field. Anti-Lebanon shows a similar picture (Arnold and Behnstedt 1993:80–92). In this area, Aramaic is still spoken in the three villages and it has been shown that the presence of Aramaic in the Arabic dialects increases the closer one gets to the villages. It is likely that this reflects earlier extension of spoken Aramaic which only quite recently has been reduced to a few places. The Aramaic words in the dialects surrounding Maʿlūla show a reflex of dialects which, unlike the Aramaic of Maʿlūla, did not have the shift *ā* > *ō*, e.g. *tiġār* ‘pot for grape syrup’, Maʿlūla: *tiġōra*, *tuġōra* (originally from Persian), *maššān* ‘extension of plough handle’, Maʿlula *maššōn*. The form *maššōn* is also found in Arabic dialects in the area.

In Mesopotamia/Iraq Aramaic is still spoken in the north around Mosul and the dialects there show many obvious lexical items with an Aramaic origin. Many of these words can probably also be found in other parts of Iraq and in Anatolia, especially in the so-called *qəltu* dialects. Unfortunately, no systematic investigation has as yet been carried out. A comparison between the works of al-Calabī (1935) and Vocke and Waldner (1984) shows only a few common items. The items collected by Calabi from the Mosul area show the preservation of Aramaic sounds, e.g. *šaql* ‘weight’, ‘measure’, Aramaic ŠQL (cf. Vocke and Waldner 1984, s.v.); *daġaš* ‘to show’, ‘to demonstrate’, Syriac *dgaš*. Some lexemes show signs of being older loans like *siḥl* ‘stream of water’, Syriac *šiḥlā*; *ṭamas* ‘to dip’, Syriac *ṭmaš*. Aramaic *ḥ* often appears as *x*, like *fasax* ‘to be wide’, Syriac *pšaḥ*.

The Aramaic vocabulary is likely to be the largest foreign element in the Arabic lexicon even if the exact extent is difficult to define. There has been a tendency to draw the line somewhat too generously (Hopkins 1995:41–43; Contini 1999:112–113). Of the 221 loanwords identified by Hebbo (1984) in the biography of the Prophet 37 percent were Aramaic or have been transmitted into the *ʿarabiyya* via Aramaic. The general picture is that of two main strata of Aramaic loanwords: the old ones in the *ʿarabiyya* and the more recent ones in the dialects. Quite a few of the *ʿarabiyya* words give the impression of being early borrowings from Aramaic dialects, not affected by the characteristic sound shifts. The vocabulary in the *Qurʾān* and in early prose contains words from all aspects of life: religion, agriculture, politics, architecture, administration, and natural phenomena, even if religious terminology dominates, a fact that may be due to the content of the texts. In the dialects, the picture is somewhat different with a predominance of Aramaic words dealing with agriculture and everyday domestic life. To this is added the religious vocabulary among the Christian minorities. The Aramaic element in the Arabic dialects also includes many of the old items in the *ʿarabiyya*, but it is uncertain whether these were borrowed from the *ʿarabiyya* or inherited from a common ancestor. To this is then added a more recent stratum which, unlike the older one, to a large extent reflects the characteristic sound shifts in Aramaic.

Bibliographical references

Arnold, Werner. 2002. “Zur Geschichte der arabischen Lehnwörter im Neuwestaramäischen”. *Neue Beiträge zur Semitistik. Erstes Arbeitstreffen der Arbeitsgemeinschaft Semitistik in der Deutschen Morgenländischen Gesellschaft vom 11 bis 13 September 200 an der Friedrich-Schiller-Universität Jena*, ed. Norbert Nebes, 5–11. Wiesbaden: O. Harrassowitz.

—— and Peter Behnstedt. 1993. *Arabisch-aramäische Sprachbeziehungen in Qalamūn (Syrien): Eine dialektgeographische Untersuchung mit einer wirtschafts- und sozialgeographischen Einführung von Anton Escher*. Wiesbaden: O. Harrassowitz.

Barbot, Michel. 1961. “Emprunts et phonologie dans les dialects citadins syro-libanais”. *Arabica* 8.174–188.

Calabi al-Mawṣilī, Dāwūd al.- 1935. *al-ʾĀṯār al-ʾaramiyya fī luġat al-Mawṣil al-ʿāmmiyya*. Mosul.

Contini, Riccardo. 1999. “Le substrat araméen en neo-libanais”. *Afroasiatica Tergensia: Papers from the 9th Italian Meeting of Afro-Asiatic (Hamito-*

Semitic) linguistics, Trieste, 23–24 April 1998, ed. M. Lamberti and L. Tonelli, 101–128. Padova: Unipress.
Féghali, Michel T. 1918. *Etude sur les emprunts syriaques dans les parlers arabes du Liban*. Paris: E. Champion.
Fischer, Wolfdietrich (ed.). 1985. *Grundriss der arabischen Philologie*, I. Wiesbaden: L. Reichert.
——. 1985. "Frühe Zeugnisse des Neuarabischen". Fischer (1985:83–95).
Fraenkel, Siegmund. 1886. *Die aramäischen Fremdworter im Arabischen*. Leiden: E.J. Brill. (Repr. Hildesheim: G. Olms, 1962.)
Freyha, Anis [ʾAnīs Frayḥa]. 1973. *Muʿjam al-ʾalfāḏ al-ʿāmmiyya (A dictionary of non-classical vocables in the spoken Arabic of Lebanon)*. (2nd ed. Beirut.)
Hebbo, Ahmed Irhayem. 1984. *Die Fremdwörter in der arabischen Prophetenbiographie des Ibn Hischam (gest. 218/834)*. Frankfurt: Lang.
Hobaika [Ḥubayqa), Yūsuf. 1939. "ad-Dawāṯir: Baḥṯ fī baqāyā al-luġa as-suryāniyya fī Lubnān wa-Sūriyā". *al-Mašriq* 37.290–412.
Hopkins, Simon. 1995. "*šarār* 'pebbles': A Canaanite substrate word in Palestinian Arabic". *Zeitschrift für Arabische Linguistik* 30.37–49.
Jeffery, Arthur. 1938. *The foreign vocabulary of the Qurʾān*. Baroda: Oriental Institute.
Luxenberg, Christoph. 2000. *Die syro-aramäische Lesart des Koran: Ein Beitrag zur Entschlüsselung der Koransprache*. Berlin: Das arabische Buch.
McDonald, M.V. 1974. "The order and phonetic value of Arabic sibilants in the ʾabjad". *Journal of Semitic Studies* 19.36–46.
Nakhla, Rafael [Rūfāʾīl Naxla]. 1973. *Ġarāʾib al-lahja al-lubnāniyya as-sūriyya*. Beirut: al-Maṭbaʿa al-Kāṯūlīkiyya.
Retsö, Jan. 2000. "*kaškaša*, t-passives and the dialect geography of Ancient Arabia". *Oriente Moderno* 19/80 N.S.111–118.
Schall, Anton. 1985. "Geschichte des arabischen Wortschatzes: Lehn- und Fremdwörter im klassischen Arabisch". Fischer (1995:142–153).
Vocke, Sibylle and Wolfram Waldner. 1982. *Der Wortschatz des anatolischen Arabisch*. Ph.D. diss., University of Erlangen.

Jan Retsö (Göteborg University)

Argument

An argument is an element of a sentence or predication that enters into the predication as one of the major valencies of the verb or predicate, namely as subject or object. These are also referred to as 'core arguments', while the subject of a sentence usually encompasses what is referred to as the 'external argument'. The notion of argument may be considered from both a syntactic and a semantic perspective, and this janus-faced nature of the notion of argument is clearly reflected in the Arabic tradition where one finds terms referring to arguments from a primarily syntactic perspective (*mubtadaʾ* 'starting place, topic' → *ibtidāʾ*) contrasting with terms of a more semantic cast (→ *fāʿil* 'actor, agent', → *mafʿūl bihi* 'acted-upon, patient'). This distinction in Arabic falls out from a basic distinction in Arabic syntax, namely that between a *jumla ismiyya* 'nominal sentence' and a *jumla fiʿliyya* 'verbal sentence'. The former is simply defined as a sentence which begins with a nominal form, often the subject, as in (2), but not necessarily, while the latter begins with a verbal form as in (1), which is regarded as the unmarked order in Classical and Modern Standard Arabic.

(1) *kataba muḥammad-un risālat-an ʾilā*
wrote.Masc.Sg. Muhammad letter-a to
ʾummi-hi
mother-his
'Muhammad wrote a letter to his mother'
(2) *muḥammad-un kataba risālat-an*
Muhammad wrote.Masc.Sg. letter-a
ʾilā ʾummi-hi
to mother-his
'Muhammad wrote a letter to his mother'

The argument structure of each of these types of sentence has given rise to several different issues of interest to linguists involving the way that arguments (syntactic or semantic) interact with their verb or the predicate in general. The one which has been of primary interest to theoretical linguists involves verb → agreement asymmetries of the sort exemplified in (3) and (4).

(3) *katab-at an-nisāʾ-u risālāt-in ʾilā*
wrote.Fem.Sg. the-women letters to
ʾummahāti-hinna
mothers-their
'The women wrote letters to their mothers'
(4) *an-nisāʾ-u katab-na risālāt-in ʾilā*
the-women wrote.Fem.Pl. letters to
ʾummahāti-hinna
mothers-their
The women wrote letters to their mothers'

That is, while features of gender agreement are copied onto a preceding head verb, features of number agreement are not, while both gender and number features are copied onto a following verb. These facts are handled in a variety of fashions in linguistic analyses (well summarized in Soltan 2004). They are troublesome in large part due to the restrictions which the theory places on the representation of agreement features (gender vs. number), as well as to the indirect fashion in which one must encode the notion of directionality in most theoretical representations. The different analyses handle them according to the different kinds of theoretical machinery which the theory provides.

While these issues of argument structure have long been of concern to theory-driven analyses, other issues involving syntactic/semantic argument mismatches or asymmetries in Arabic have also been of interest in theory-neutral analyses. Prominent among these is the apparent reanalysis of → topics as subjects (i.e., as explicit external arguments) in many Arabic dialects, as detailed for Maltese by Comrie (1982). → Topicalization in general for all forms of Arabic involves placing a nominal (noun or pronoun) from any position in the sentence at the beginning of a sentence, while a trace pronoun is left behind as a place holder. In Modern Standard Arabic this is usually accompanied by an overt marker, namely *ʾammā . . . fa-* as in (5).

(5) *ʾammā muḥammad-un fa-katab-tu lahu*
as for Muhammad I wrote him
risālat-an ʾams
a letter yesterday

The *ʾammā . . . fa-* construction is not available in the dialects, where topicalization may take place with any nominal element of the sentence, and often does so with great frequency, especially for nominals whose reference is human and especially if it is close to the deictic center of the discourse (involving the speaker, or the interlocutor) as in (6).

(6) (Egyptian)
ʾana ʾabū-ya ḥayīgi baʿd^{i} šiwayya
I father-my will-come after a little
'My father will come in a bit'

For certain constructions involving non-verbal predicates, placing a nominal of this sort in a place generally reserved for the external argument (or subject in a language which is primarily SVO, as are most Arabic dialects) conditions a reanalysis of it as a kind of quasi-subject, along with a reanalysis of its non-verbal predicate as a kind of quasi-verb. This is most especially evident in possessive predications involving prepositions (such as *ʿand* 'with, by' in Cairene Arabic), but it may appear with a variety of other similar constructions as well. As noted by Comrie (1982) for → Maltese, certain constructions show a discrepancy between the morphological form and the syntactic function of certain elements involved in those constructions. These include the preposition *għand* 'by, with' (expressing possession) and quasi-auxiliaries such as *għad* 'still', *għodd* 'almost', *qis* 'like', among others (→ pseudo-verb).

Comrie (1982) makes several claims regarding the syntactic-morphological discrepancy of these constructions, each involving, implicitly or explicitly, the definitional criteria for each of the three categories in question: verb, subject (external argument), and topic. First, the non-verbal predicates in these constructions function as verbs; second, the sentence-initial N(oun) P(hrase) in these constructions functions as subject/external argument even though it 'looks like' a topic; third, the reanalyses proposed for these different items in these contexts are 'complete'; fourth, the 'syntactic-morphological discrepancy' is claimed to lie in the fact that the oblique form of the pronoun on the preposition or quasi-auxiliary is actually the subject pronoun.

These claims must be qualified somewhat, however. The claim that these items are 'morphologically irregular verbs' is misleading. This implies that they are morphological verbs, which they are definitely not, since they do not form a morphological word-class along with all the other items that function quasi-verbally, including nouns, pronouns, adverbs, and auxiliaries. Nor are their pronominal markings morphologically obligatory, since they may appear without them with full NPs. Also, if these items were morphological verbs, albeit irregular, then this would resolve the 'syntactic-morphological discrepancy' in favor of a simple morphological irregularity. In other words, if these are morphological verbs, then there is no discrepancy at the syntactic level. It seems better to modify this claim by simply saying that there is a mismatch or asymmetry between the syntax and morphology of these forms, as conditioned by the argument structure at the semantic level: these items function syntactically as verbs in these contexts, despite their morphological form. This modified claim is supported by several different criteria used to define the category of verb: negation, word order, and subject agreement. These items (*għand*, *donn*, etc.) take what is identified as verbal negation, namely *ma—x*, (7), their position in the sentence is that of a 'verb' in an SVO language (8), and they agree with their 'subject' (9).

(7) a. *ma-għand-u-x*
'He does not have . . .'
b. *ma-donnħā-x*
'She is still . . .'
(8) a. *pawlu għandu bejd*
'Paul has an egg'
b. *il-baqra donnħa torqod*
'The cow is still lying'
(9) a. *pawlu*$_i$ *għandu*$_i$ *bejd*
'Paul has an egg'
b. *il-baqra*$_i$ *donnħa*$_i$ *torqod*
'The cow is still lying'

The second claim of Comrie's paper is that the S(entence)-initial NP in both of these constructions is not a topic, but rather a true subject, fully expressive of the external argument of the predication. This claim is supported by using several different criteria to define the categories of subject and, less explicitly, of topic: subject-verb agreement, word order, and the availability or not of a non-topicalized counterpart. First, the S-initial NPs in the above sentences are subjects because they agree (albeit irregularly) with their 'verb': in possessive sentences the preposition *għand* functions as the verb, while in the others one of the quasi-auxiliaries function as the verb. More significant as regards subject-verb agreement, however, is the fact that in sentences with quasi-auxiliaries in the past or future tense, the temporal auxiliary *kien* agrees with this initial NP (10).

(10) *kont għodd-ni waqaj-t*
'I almost fell'

Second, these NPs are 'subject' because they occupy subject position in an SVO language, namely sentence-initial position. Topics are also sentence-initial, but Comrie (1982) claims that subjects differ from topics since subjects trigger 'real' verb agreement. Third, the S-initial NPs in these sentences are subjects according to Comrie because they lack a non-topicalized counterpart (11). Deriving these structures from an obligatory movement rule is ruled out primarily because topicalization would then have lost its 'distinctive function' in the language.

(11) **għand pawlu bajda*

The third claim of Comrie's paper is that all of these reanalyses (of topics as subjects and of non-verbs as verbs) are 'complete', even though he notes (1982:303) traces of an 'ongoing' reanalysis in Maltese. This is supported by the above third point (the non-existence of a non-topicalized counterpart), but this claim is much weaker than it appears, since there are non-topicalized counterparts. The latter group, however, does not have the exact 'possessive/ownership' reading that the topicalized group does. This does not point to a complete reanalysis but rather to the close connection between location and possession. In addition, in Cairene Arabic almost all → locative predicates behave in a similar fashion: they have both topicalized and non-topicalized forms, the first of which has a 'possessive/ownership' reading to it and may be negated with *ma—š*, and the second of which is more purely a locative and does not take *ma—š* negation. Both are grammatical and derivable from the other, and do not involve a complete reanalysis but rather point to the close connection between location and possession.

The final claim is that in these constructions the subject is marked irregularly in the predicate, namely through the use of an oblique pronoun form, hence there is a discrepancy between the syntactic function of these items (viz. subject) and their morphological form (viz. object). This is an important insight, but is less of a discrepancy than it appears at first sight. This is because oblique pronoun forms in Maltese and in other forms of Arabic, and in Semitic in general, are not exclusively restricted to marking non-subjects but often do mark subject. This is so not just in the kind of constructions under consideration here, but in others as well, where the items to which they are conjoined cannot be considered verbal in any way. Rather the oblique form of the ostensible subject seems to be controlled or governed at a superficial level, and is not indicative of a reanalysis of any sort. In Cairene Arabic we have the following:

(12) *zamān-ak guˁt*
'You have (probably) become hungry by now'

(13) *fēn-ak?*
'Where are you?', 'Where have you been?'

(14) *lōlay̱, kān ʾatal-ak*
‘If not for me, he would have killed you’

(15) *(ya) m-aḥlā-ha*
‘How lovely she is!’

Because of these facts, the primary discrepancy in the above constructions is with the ‘prepositional + pronoun’ complex as a whole, and not necessarily with the form of the pronoun marker itself.

The way that the topicalized nominal has been reanalyzed as a kind of quasi-external argument has shed light on a number of criteria for defining both verbhood and subjecthood in Maltese and other Arabic dialects as well. Comrie (1982) revealed that there are at least three main criteria for verbhood in Maltese (the use of *ma—x* negation, agreement with a ‘subject’ as external argument, and word order position), at least two criteria for subjecthood (agreement marking on verb and word order position), and at least two criteria for topic-hood (word order position and availability of a non-topicalized counterpart). Each of these criteria may be applied to similar constructions in almost every Arabic dialect, whereby certain items, ostensibly non-verbal, take on certain verbal features (viz., *ma—š* negation) in contexts involving preposed topics reinterpreted as external arguments. Many of these constructions involve lexical items which are in fact cognate with the items in Maltese identified by Comrie (1982). The criteria supplied by him can therefore be used as a starting point in the analysis of these items in the individual dialects, and as a tool by which the various dialects may be compared in this regard.

For example, in both Maltese and Cairene Arabic almost any construction involving a locative preposition with an oblique pronoun may be reanalyzed as a quasi-verb sufficiently enough to be negated with *ma—š* negation if it is used with a preposed nominal that is co-referent with the oblique pronoun on the preposition, which functions as a quasi-external argument:

(16) *miḥammad ma-ʿand-ū-š ʿarabiyya*
‘Muhammad does not have a car’

(17) *it-ṭarabēẕa ma-taḥta-hā-š siggāda*
‘The table doesn’t have a carpet under it’

(18) *ʾana ma-wara-yā-š šuġl*
‘I don’t have work to do’

In addition, in both Maltese and Cairene Arabic there are constructions involving a limited class of nouns used with a possessive pronoun suffix that may also be reanalyzed as quasi-verbs: for Maltese there is (19).

(19) *jisim-ni pawlu*
‘My name is Paul’

This does not derive from a verb but from the noun *ism* ‘name’. The use of the suffix form *-ni*, generally a verbal feature, indicates that this noun has taken on certain other features of a verb, but agreement features indicate that it retains certain nominal ones. A similar case of a noun taking the *-ni* object suffix exists in Moroccan Arabic, where the noun *ʿemr-* ‘life’, when used adverbially, may take either the *-ī* or the *-ni* form of the oblique 1st person pronoun suffix (Harrell 1962:215): *ʿemmer-ni* or *ʿemr-ī*.

In Cairene Arabic there are two types of constructions involving nominals: (a) those that take a nominal complement only, and (b) those that may take either nominal or sentential complements. Type (a) and type (b) with sentential complements may take discontinuous negation, while type (b) with nominal complements does not take discontinuous negation.

type a: (20) *ism-* ‘name’:
ma-smū-š ṣadīq-ik
‘He is not called “your boyfriend”’

(21) *ḥilt-* ‘resource, (to) have’:
ma-hilt-ū-š
‘He didn’t have a thing’

(22) *lōn-* ‘color’:
ma-lon-ū-š ʾaḥmar
‘It is not red’

type b: (23) *bidd-* ‘desire’:
ma-bidd-ī-š ʾasāfir
‘I don’t want to travel’

(24) *nifs-* ‘desire’:
ma-nifs-ī-š ʾaštaġal
‘I don’t want to work’

(25) *ʾaṣd-* ‘intention’
ma-ʾaṣd-ī-š ʾaʾūl da
‘I don’t mean to say that’

There are also adverbs which may be used in a quasi-verbal manner, although more so in Maltese than in Cairene. These include *għad* 'still', which Comrie (1982) includes among his quasi-auxiliaries and which is not related to the future marker *ġad* (< Classical Arabic, *ġadā*), but from the adverbial usage of *ʿāda* 'to return' (cf. Lane 1984:2189).

There are a number of adverbs in Cairene Arabic that have quasi-verbal features, but in a more limited way than is the case in Maltese or in other dialects. These adverbs, which include *taww-* 'just', *lissā-* 'still', *yadōb-* 'just barely', take pronoun suffixes in agreement with the ostensible subject (or external argument) of the sentence, but unlike the adverb *għad* in Maltese, these adverbs in Cairene Arabic do not take *ma—š* negation.

(26) *taww-u gāyy*
'He has just now come'

In addition to these prepositions, nouns, and adverbs that may function as quasi-verbs in Maltese and Cairene Arabic, there are particles or 'quasi-auxiliaries' probably deriving from aphaeretic verb forms (i.e., verb forms which have lost their initial syllable) which were reanalyzed as adverbs, then further reanalyzed as quasi-verbs. These include for Maltese *donn-*, *qis-*, and *għod-*. Their analysis as aphaeretic verb forms is based on the behavior of the particle *xeṣṣ* in Moroccan Arabic, which patterns as follows (Harrell 1962:185):

present: *xeṣṣ-ha timši le-s-suq l-yum*
'She has to go to the market today'
habitual: *ka-ixeṣṣ-ha temši le-s-suq koll nhar*
'She has to go to the market every day'
past: *kan xeṣṣ-kom teʿṭiw-hom-li*
'You should have given them to me'

Cairene Arabic, on the other hand, does not seem to have the kind of 'quasi-auxiliaries' that Maltese has, although the particle *tann-/tannit-* 'to continue' might qualify as such, although it is of a type quite different from the Maltese ones. It is like the Maltese 'quasi-auxiliaries' in that it takes an oblique pronominal suffix that is coreferent with the subject of the sentence (27a), but unlike them, it does not take *ma—š* negation (27b)

(27) a. *miḥammad tann-u māši*
'Muhammad continues to walk'

b. **miḥammad ma-tann-ū-š māši*

It does, however, have many other features of a verb, including a well-developed imperfect conjugation, as in (28), and a connotation of past time. This latter feature may be due to the fact that it derives from a perfect verb, but it may also derive from a nominal or adverbial source, probably not from Coptic as Hinds and Badawi (1985:139) believe, but related to Kuwaiti Arabic *tamm* or to the Moroccan Arabic full verb form *temm* with meaning and use similar to *tann- (temmina ġadyīn, temmu ġadyīn).*

(28) a. *yitann-u gāyy/yitannit-u gāyy*
'He keeps on coming'

b. *nitann-ina gāyyīn/nitannit-na gāyyīn*
'We keep on coming'

c. *yitann-u-hum gāyy/yitannit-hum gāyyīn*
'They keep on coming'

d. *titann-uku(m)/titannit-ku(m) gāyyīn*
'You [pl.] keep on coming'

This item, whatever its origins are, has become reanalyzed as a verb to the point where it has an imperfect conjugation, yet despite this abundance of verbal and quasi-verbal features, it generally does not take discontinuous negation.

All of the above classes of quasi-verbs are dependent for their reanalysis on the presence of a preposed, topicalized nominal form which is coreferent with the oblique pronoun attached to them. The spread of this kind of structure throughout the wide variety of Arabic dialects is an indication of the degree to which sentence-initial position has become identified as the position for the external argument or subject of the predication, even when the ostensible agreement marker for that external argument is of an oblique (or non-subject) form. There is a great deal more to discover about argument structure in Arabic, and it is hoped that many of the issues mentioned here will be further clarified in the future.

Bibliographical references

Comrie, Bernard. 1982. "Syntactic-morphological discrepancies in Maltese sentence structure". *Communication and Cognition* 15:3–4.281–306.

Harrell, Richard. 1962. *A short reference grammar of Moroccan Arabic.* Washington, D.C.: Georgetown University Press.

Hinds, Martin and El-Said Badawi. 1986. *A dictionary of Egyptian Arabic.* Beirut: Librarie du Liban.

Lane, Edward. 1984. *An Arabic–English lexicon.* Cambridge: Islamic Texts Society. (1st ed. 1877.)

Soltan, Usama. 2004. "An argument for AGREE and

Multiple Spell-out: Standard Arabic agreement asymmetries revisited". Paper presented at the Workshop on Minimalist Theorizing, Indiana University. Also available at <http://www.indiana.edu/~lingdept/syntax/ minimalist/Handouts/Soltan_MT2004.pdf>.

John C. Eisele (College of William and Mary)

Article, Definite

All varieties of Arabic have a prefixed definite article. Its main allomorph is *(ʾ)Vl-*, e.g. *al-bāb* 'the gate', but it is *ʾam-/im-/an-/in-* in some Yemenite dialects and *ʾam-* in Ancient West Arabian and in the dialect of Ṭayyiʾ (Rabin 1951:34–37, 50–51, 205). In some Proto-Arabic dialects (e.g. Liḥyanite) it was *han-*, *hal-* (today found, e.g., in many Syrian, Lebanese, and Tunisian varieties, but usually interpreted as going back to a contraction: *halbēt* 'this/the house' < **hāḏā l-bēt* 'this house') and also *hV-*, most probably with the gemination of the first consonant of the determined word. In Classical and in Modern Literary Arabic it is *al-*, while in other dialects it is usually *il-/əl-*, or even *lə-*, e.g. Damascus and Muslim Tunisian Arabic *lə-kbīr* 'the big one'. In a couple of nouns and in the demonstrative *ʾulāʾi* the definite article *al-* has conditioned the elision of the initial *ʾV-*, e.g. *al-ʾilāhu* > *allāhu* 'the God'; *al-ʾulāʾi* > *allāʾi*. Rarely, the definite article is reduced to *l-* as in *l-aḥmar* which is a variant of common *al-ʾaḥmar* 'the red one'.

From a synchronic point of view (this was the approach of the medieval Arab grammarians, for whose different opinions see Fleisch (1990:II, 56), the definite article has three other allomorphs:

i. *-l-*, viz. with the elision of *(ʾ)V-* in the position after the final vowel of the preceding word, e.g. *al-baytu l-kabīr* 'the big house', usually pronounced, however, *ʾal-bayt ʾal-kabīr* in Modern Literary Arabic. The elision does not take place after the interrogative particle *ʾa*, e.g. *ʾa-ʾal-bint* 'the girl?'. If the preceding word ends in a consonant, a sandhi liaison vowel appears, e.g. *katabat* 'she has written' but *katabat-i l-kitāb* 'she has written a book', although in Modern Arabic this is usually pronounced *katabat ʾal-kitāb*;
ii. *(ʾ)V-C1* after a pause, viz. in an initial position where, according to standard synchronic interpretation *-l-* assimilates to the following apical consonant (here symbolized by C1), i.e., causes its gemination, e.g. *ad-dars* 'the lesson' < **al-dars*, *ar-rajul* 'the man' < **al-rajul*, *as-sana* 'the year' < **al-sana*;
iii. *-C1-* in the position after the final vowel of the preceding word in the same syntactic group (which excludes a pause and necessitates sandhi) or after a sandhi vowel, e.g. *li-r-rajul* 'for the man'. The → assimilation of *-l-* to the following apical consonants (traditionally called 'sun letters' on the basis of *aš-šams* 'the sun'; other, i.e. non-assimilating consonants, are called 'moon letters' following the example of *al-qamar* 'the moon'), viz. /ṯ/, /ḏ/, /t/, /d/, /ṭ/, /ḍ/, /s/ /š/, /ṣ/, /z/, /ḏ̣/ [ẓ], /n/, /r/ (usually grammarians list also /l/ as a 'sun letter'), occurs in virtually all dialects. Outside Classical Arabic and Modern Literary Arabic, *-l-* assimilates also to initial /k/ (e.g. *ik-kitab* 'book' versus *il-kitab*), /g/, /b/, /f/, /m/, /n/, /j/ [ž], and in some West Yemenite dialects (e.g. Jiblah) *-l-* (or -n-/-m-) of the article assimilates to all consonants, e.g. *ab-bēt*/*ib-bēt* 'the house'.

In Arabic writing, the letter *-l-* is always spelled, which results in different Latin, both scientific and non-scientific, transliterations, which are actually a mix of transcription and transliteration, e.g., *ʾal-ṯawra* and *ʾaṯ-ṯawra* 'the revolution', *ʾal-sadāt* and *ʾas-sadāt*. In Modern Literary Arabic, as pronounced even by some radio and television news readers, initial variants *ʾal-* and *ʾaC1-* sometimes occur even in non-initial position after a vowel, e.g. *fī hāḏā ʾal-barnāmij* 'in this program', instead of the standard *fī hāḏā l-barnāmij*. The last example shows that the Arabic definite article must be used with demonstrative pronouns (the phrase amounts to '*in this the program') and in rare cases even with a suffixed pronoun, e.g. Classical Arabic *ʾat-tābiʿī* 'the one who follows me'.

The definite article is used with contextually definite/determined nominals (except as the head/ *possessum* in a genitive construction, e.g. *bāb al-madrasa* 'the gate of the school' and before possessive pronouns) as well as with lexically definite nouns, e.g. *aš-šams* 'sun' (the sun in 'our' solar system). Abstract nouns and generic nouns usually occur with a definite article, e.g., *al-muruwwa* means 'the ideal of manhood', *ar-rajul* means both 'the man' and 'man [in general]' but 'gold [in general]' can be either *aḏ-ḏahab* or *ḏahab*. The definite article is repeated with every attribute in a nominal phrase, e.g. *fī kitāb al-mudarris al-ʿarabī al-jadīd* ‘in the new Arab lecturer’s book/in the new book of the Arab lecturer’.

For the history of research on the origins of the Arabic definite article see Testen (2000:135–182). Initial ʾ- alternates with *h-* and the Arabic definite article is related to the demonstrative *hā-*, which occurs in *hā-ḏ-ā* ‘this [masc.]’, *hā-ḏ-i-hī* ‘this [fem.]’, *hāʾulāʾi* ‘these’, cf. Syriac *hā-n-(ā)* ‘this [masc.]’, *hā-ḏ-(ē)* ‘this [fem.]’, *hā-ll-ēn* ‘these’. It is related to the Northwest Semitic definite article *ha*-C1- and to Modern South Arabian **h/ʾV-* (the latter has been largely lexicalized and occurs only residually). The fact that the definite article shows neither gender nor number, while nouns and adjectives, demonstratives, and other pronouns do show these categories raises the question of whether originally, i.e. in the prehistoric period, the definite article was not inflected. The rule of assimilation of *-l-* to the following consonants articulated with the tip of the tongue is valid synchronically only for the article and it does not work in other circumstances, e.g., the indefinite plural of *lisān* ‘language, tongue’ is *ʾalsun*, not **ʾassun* (the definite plural is *al-ʾalsun* ‘the languages, the tongues’), *ʾalṯaj* ‘having a speech defect’, etc., not to mention many nominal (e.g. *falta*, *jald*, *jild*, *jalsa*, etc.) and verbal forms (e.g. *yalzam*, *yalsan*, *iltajaʾa*, etc.). Therefore, Zaborski (2000) has suggested a hypothesis that the original, i.e. prehistoric, definite article was masculine **ʾan-/han-*, feminine **ʾat/hat-* and plural **ʾal-/hal-*. Both *-n-* and *-t-* assimilate to many following consonants and in this view, in Classical Arabic the plural **ʾal-* variant was chosen out of all four forms as a written symbol because *-l-* was the least assimilating consonant, and **ʾal/hal-* was used in the plural for both feminine and masculine nouns, as well as for collectives (semantic plurals). This made it possible to extend its use to the singular of both genders.

According to Kuryłowicz (1972:131–133), the historically recorded definite article with the main allomorph *al-* is a relative innovation. In the prehistoric period → nunation had the function of the definite article, which it still has with some proper names (e.g. *ḥasan-u-n* vs. later *al-ḥasan-u* ‘Hasan’), and in some semantically determined phrases and words like *buʿayda samāʾi-n* ‘not far from the sky’, *kullu rajul-i-n* ‘the totality of mankind’, *ḏahab-u-n* ‘gold’, etc.

Bibliographical references

Fleisch, Henri. 1990. *Traité de philologie arabe*. 2 vols. 2nd ed. Beirut: Dar el-Machreq.

Gabučan, Gračija M. 1972. *Teorija artiklja i problemy arabskogo sintaksisa*. [Theory of the article and problems of Arabic syntax]. Moscow: Nauka.

Kuryłowicz, Jerzy. 1950. “La mimation et l’article en arabe”. *Archív Orientální* 18.323–328.

——. 1972. *Studies in Semitic grammar and metrics*. Wroclaw: Ossolineum. (1st ed. *Archív Orientální* 18, 1950.)

Rabin, Chaim. 1951. *Ancient West-Arabian*. London: Taylor’s Foreign Press.

Testen, David. 1998. *Parallels in Semitic languages*. Leiden: E.J. Brill.

Zaborski, Andrzej. 2000. “Inflected article in Proto-Arabic and some other West Semitic languages”. *Asian and African Studies* 9.24–35.

Andrzej Zaborski (University of Cracow)

Article, Indefinite

1. Definition

In Classical Arabic no element is found which could be directly compared to the indefinite article (quantifier of existence) in Indo-European languages, e.g. the element *a* in English. Indefiniteness (Arabic *tankīr*) is simply marked by a combination of a Ø-article and → nunation (*tanwīn*) in the case of triptote nouns and by a Ø-article in the case of diptote nouns in the plural (cf. Fleisch 1961:342–345 and Diem 1975 for morphological details and the genesis of mimation and nunation). Furthermore, indefiniteness can be rendered by paronomastic constructions such as *wasīlatun min al-wasāʾili* ‘a medium’ (lit.‘medium of mediums’), or by the quantifiers *baʿḍ* and *biḍʿ* ‘some’. However, at least since the 12th century C.E. the use of the numeral *ʾaḥad/wāḥid*, fem. *ʾiḥdā* ‘one’ with the semantic shade of an indefinite article can be observed (cf. Monteil 1960:234). Thus, constructions are encountered such as *ʾaḥadu l-qabāriṣi* ‘a Cypriot’ (lit. ‘one of the Cypriots’), or *bābu ʾaḥadi l-buyūti* ‘the door of a house’ (lit. ‘the door of one of the houses’ in Modern Standard Arabic. In modern Arabic dialects indefiniteness can be expressed mainly in three ways: (a) by the mere absence of the definite article; (b) by the absence of the definite article in combination with reflexes of the nunation (*tanwīn*); and (c) by a number of quantifiers which are either reflexes of Classical Arabic *wāḥid* or other elements such as *fard* and (with regressive assimilation) *fadd* in Iraqi dialects. These elements are now found in teaching manuals and dictionaries of modern Arabic dialects (e.g. Hinds and Badawi 1986:926–927; Holes 2001: 553–554; Woidich 2002:315; Woodhead and Beane 1967:347, 490; Erwin 1969:91, 194) as well as in the relevant reference works and comparative analyses (e.g. Fischer and Jastrow 1980:88; Versteegh 1984:99; see below). This entry looks first at those elements which can properly be labeled 'indefinite article' in a synchronic perspective. In a further step it deals with modern reflexes of the *tanwīn*.

2. Indefinite article(s)

In connection with the development of deictic elements in Arabic dialects indefinite articles have emerged, either based on the numeral 'one' or on words denoting an 'individual' (cf. Versteegh 1984:99). Most frequently encountered are varieties of *wāḥid* 'one'. In Cairene Arabic (but not necessarily in other Arabic dialects) the use of *wāḥid* as an indefinite article seems to be restricted to persons (cf. Fischer and Jastrow 1980:309; Woidich 2002:315), e.g. *wāḥid maṣri* 'an Egyptian'. The indefinite article in this form is also found in the Western pre-Hilālī group of the Maghreb dialects. According to Marçais (1977:163–164) a morpho-syntactic opposition exists between the use of *wāḥəd*/*waḥd* as indefinite article and as indefinite pronoun. Thus there is *waḥdəṛ-ṛājəl* 'a man' ('a the man' and *waḥd əl-mṛa* 'a woman' ('a the woman)' with gender-invariable *waḥd*, as opposed to *wāḥəd ṛājəl* 'someone (male)' and *waḥda mṛa* 'someone (female)'. In the same dialect group, the indefinite article also surfaces in a shortened form *ḥa-* (as a probable reflex of *ʾaḥad*), e.g. *ḥa-ṛ-ṛājəl* 'a man' and *ḥa-l-mṛa* 'a woman'. Cohen (1975:221–222) and Marçais (1956:400) sketch comparable scenarios for the Jewish vernacular spoken in Tunis and the Algerian Arabic dialect of Djidjelli respectively. In the latter the Arabic indefinite article can also precede nouns of Berber origin with an *a*-prefix. It is noteworthy that the indefinite article in this form occurs always in combination with the definite article *(a)l-*, except in cases where it precedes an *ʾiḍāfa*, as in *ḥa bāb əd-dāṛ* 'a house door' (Marçais 1956:401). Versteegh (2001: 164) suggests that this may have come about in analogy to the construction of the demonstrative with the article (*hāḏā r-rajul* 'this man'). Marçais (1977:163) points out that a comparable construction (indefinite article–definite article–noun) is prominent in many varieties of Berber and may thus have given rise to the analogous construction in Arabic dialects in the Maghreb. However, the definite article is not found after elements expressing indefiniteness other than those deriving from *wāḥid*/*ʾaḥad*. In the Maghreb such elements include *šī* (< *šayʾ*), *fərd* (< *fard*) 'individuum/ unity', and *baʿd (mən)* (< *baʿḍ (min)*), e.g. *šī ṛājəl* 'a man', *šī mṛa* 'a woman', *fərd ktāb* 'a book' (Marçais 1977: 164–165), and *baʿd mən nāṛ* 'a day' (Cohen 1975:222). Caubet (1993:257–260) distinguishes four degrees of determination in Moroccan Arabic, the first of which refers to the Ø-article and the second ('quelque X') comprises the constructions with *wāḥ(ə)d əl-* and *ši*. Outside the Maghreb the particle *éxen*/*éxte* occurs in Cypriot Arabic, e.g. *éxte táka* 'a window' (Borg 1982:218), *éxte mára* 'a (certain) woman' (Borg 1985:146).

The elements in Mesopotamian Arabic deriving from *fard* (*fared* in the Baghdadi Muslim variety, *faġad* in the Baghdadi Jewish variety, *faġed* in the Baghdadi Christian variety, and as syncopated form *fadd* in all varieties) have been labeled "the characteristic Mesopotamian 'indetermination marker(s)'" by Blanc (1964:118) (cf. also Abu Haidar 1991:111). This element also surfaces as *fat* in insular Arabic dialects (largely of the *qəltu* type) in Central Asia (cf. Fischer 1961: 242). Blanc (1964:118) also states that "the degree to which [the indefiniteness marker] contrasts with absence of any mark is yet to be determined". He comes up with a minimal pair of the two answers (in the Muslim variety) *ʿendi fadbēt* 'I have a house' vs. *ʿendi bēt* 'I have a house', the first of which is supposed to respond to the question 'what have you got?', while the second is supposed to respond to the question 'who has a place we can meet at?'. Interestingly, Blanc (1964:119) notes that the indefinite article in this group of Arabic dialects appears to be an areal feature in Mesopotamia, comparable to such elements as Turkish *bir*, Persian *ye(k)*, and Northeastern Neo-Aramaic *xa*.

As in the previously cited cases the distribution of the indefinite article vs. the absence of the same is not easy to determine in the Baghdad varieties of Iraqi Arabic. Neither criteria such as subject position vs. predicate position, nor 'indefinite' vs. 'partitive' fully account for the presence or absence of a dialectal indefiniteness marker. The following short dialogue (Blanc 1964:119) illustrates this situation:

– *tfaḍḍal jigāra*	'have a cigarette'
– *la', aškurak*	'no, thank you'
– *fadgahwa?*	'a cup of coffee?'

Of special interest is the observation by Roth (1979:141), referred to by Versteegh (1984:99), that in some varieties of Chad Arabic the indefiniteness marker *wāḥid* can serve as the *marked* term in the opposition definite vs. indefinite in circumstances where the use of the definite article is reduced.

3. Reflexes of nunation

In some dialects of Arabic traces of the historical nunation can be reasonably labeled 'indefiniteness marker'. A distinction must be made, though, between the mere morphophonological (and prosodic) surface and the actual semantics of such elements. In terms of morphophonology the phenomenon is mentioned, for instance by Holes (1996:14, 143) with reference to Johnstone (1961). Thus, in many Bedouin dialects, the noun in a noun-adjective phrase retains (or rather re-develops) an invariable *-in* or *-an* ending (independently of the Classical Arabic case system), e.g. *bint-in zēna* 'a good girl' in the singular or *fardat-in zēna* 'good dates'. Behnstedt (1987:54) reports a contrast between an *-in* ending for nouns in context and an *-u* ending for nouns in pause in the northernmost dialects of Yemen e.g. *labanin* vs. *labanu* 'milk'. The interesting case here is presented by dialects in which such endings adopt a semantic shade of their own. Such a scenario is drawn by Reichmuth (1983:188–201) for the Arabic dialect of the Šukriyya in eastern Sudan. Here a tripartite system exists, with a third 'state' between indefinitenesss and definiteness, labeled 'declarative' by Reichmuth. The following chart (Reichmuth 1983:188) provides an illustration:

i.	*al-kalām fī bagara*	'the talk about a cow' ('some cow')
ii.	*al-kalām fī bagartan layy*	'the talk about a cow of mine'
iii.	*al-kalām fī l-bagara/ bagarti*	'the talk about the cow/my cow'

Here the *-an* suffix (the use of which is by no means compulsory) conveys the semantic information of a known circumstance which is brought to the attention of the listener (*taʿyīn wa-taxṣīṣ*). This suffix can even be attached to dual and plural endings, e.g. *kurʿēnan* '(two) legs' or *wāḥdīnan* 'some people [masc.]'.

4. Conclusions

The publications in Arabic dialectology that mention the emergence of some form of 'indefinite article' do not allow clearcut distinctions as to when this article is to be used and when not. No indications are found that the non-use of such an article would be ungrammatical under any circumstances. Rather, it seems that we are looking at pragmatic tendencies. What is more, the difference between the use of the element *wāḥid* as a numeral and as an indefinite article (or a partitive element) cannot always be seen in the examples given. Nevertheless the very existence of an indefinite article *per se* seems by now firmly established in a large number of Arabic dialects.

Bibliographical references

Abu Haidar, Farida. 1991. *Christian Arabic of Baghdad*. Wiesbaden: O. Harrassowitz.

Behnstedt, Peter. 1987. *Die Dialekte der Gegend von Ṣaʿdah*. Wiesbaden: O. Harrassowitz.

Blanc, Haim. 1964. *Communal dialects in Baghdad*. Cambridge, Mass.: Harvard University

Borg, Alexander. 1985. *Cypriot Arabic: A historical and comparative investigation into the phonology and morphology of the Arabic vernacular spoken by the Maronites of Kormakiti village in the Kyrenia district of north-western Cyprus*. Stuttgart: F. Steiner. (Based on the author's *Habilitationsschrift*, University of Erlangen, 1982.)

Caubet, Dominique. 1993. *L'arabe marocain*. II. *Syntaxe et catégories grammaticales, textes*. Paris and Louvain: Editions Peeters.

Cohen, David. 1975. *Le parler arabe des juifs de Tunis*. II. *Etude linguistique*. The Hague: Mouton.

Diem, Werner. 1975. "Gedanken zur Frage der Mimation und Nunation in den semitischen Sprachen". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 125.239–258.

Erwin, Wallace. 1969. *A basic course in Iraqi Arabic*. Washington, D.C.: Georgetown University Press.

Fischer, Wolfdietrich. 1961. "Die Sprache der arabischen Sprachinsel in Uzbekistan". *Islam* 36.232–263.

—— and Otto Jastrow (eds.). 1980. *Handbuch der arabischen Dialekte*. Wiesbaden: O. Harrassowitz.

Fleisch, Henri. 1961. *Traité de philologie arabe*. I. *Préliminaires, phonétique, morphologie nominale*. Beirut: Imprimerie Catholique.

Hinds, Martin and El-Said Badawi. 1986. *A dictionary of Egyptian Arabic*. Beirut: Librairie du Liban.

Holes, Clive. 1996. *Modern Arabic: Structures, functions and varieties*. London and New York: Longman.

——. 2001. *Dialect, culture, and society in Eastern Arabia*. I. *Glossary*. Leiden: E.J. Brill.

Johnstone, Thomas. 1961. "Some characteristics of the Dōsiri dialect of Arabic as spoken in Kuwait". *Bulletin of the School of Oriental and African Studies* 24.249–297.

Marçais, Philippe. 1956. *Le parler arabe de Djidjelli: Nord constantinois Algérie*. Paris: Librairie d'Amérique et d'Orient and A. Maisonneuve.

——. 1977. *Esquisse grammaticale de l'arabe maghrébin*. Paris: Librairie d'Amérique et d'Orient and A. Maisonneuve.

Monteil, Vincent. 1960. *L'arabe moderne*. Paris: C. Klincksieck.

Reichmuth, Stefan. 1983. *Der arabische Dialekt der Šukriyya im Ostsudan*. Hildesheim: G. Olms.

Roth, Arlette. 1979. *Esquisse grammaticale du parler arabe d'Abbéché (Tchad)*. Paris: P. Geuthner.

Versteegh, Kees. 1984. *Pidginization and creolization: The case of Arabic*. Amsterdam and Philadelphia: J. Benjamins.

——. 2001. *The Arabic language*. 2nd ed. Edinburgh: Edinburgh University Press.

Woidich, Manfred. 2002. *Ahlan wa-sahlan: Eine Einführung in die Kairener Umgangssprache*. 2nd ed. Wiesbaden: L. Reichert.

Woodhead, Daniel and Wayne Beene. 1967. *A dictionary of Iraqi Arabic: Arabic–English*. Washington, D.C.: Georgetown University Press.

LUTZ EDZARD (University of Oslo)

ʾAṣl

The term *ʾaṣl* is primarily used as one of the major tools of analysis in Arabic grammatical theory. It is first encountered in the *Kitāb* of Sībawayhi (d. 180/796) where it occurs 569 times (Troupeau 1976 *sub ʾṢL*), all of which, with the exception of four instances, indicate a methodological notion. Although the term itself may be generally translated as 'origin', 'principle', or 'base', it is used in a variety of specialized senses, the most basic of which are (for use by Sībawayhi, see Baalbaki 1988:163–164):

i. The form, pattern, case ending, etc. which agrees with the → *qiyās*, that is, with the norm and with the usage which is most frequently attested in accepted dialects. It is, therefore, the *ʾaṣl* in nominal sentences to begin with the definite noun (e.g. *al-ḥamdu li-llāhi* 'praise be to God' and *al-waylu la-ka* 'woe unto you') (Sībawayhi, *Kitāb* II, 165). Usage contrary to this would require justification (e.g. *salāmun ʿalay-ka* lit. 'a greeting to you' and *waylun la-ka*) (*Kitāb* I, 166). In Ibn Mālik's (d. 672/1274) hemistisch *wa-l-ʾaṣlu fī l-mabniyyi ʾan yusakkanā* 'the *aṣl* in indeclinable words is to end in *sukūn*; i.e. absence of vowel'; Ibn ʿAqīl, *Šarḥ* 36), the *sukūn* is perceived as the norm for → *bināʾ* in nouns, verbs, and particles, as in *kam* 'how much?', *iḍrib* 'hit', and *ʾajal* 'yes', and thus a noun like *ʾamsi* 'yesterday' and a particle like *ʾinna* 'indeed', both of which are *mabnī* 'indeclinable', are considered to be contrary to the *ʾaṣl*.

ii. The attested form, pattern, etc. which is assigned as the origin from which a certain usage has developed. Thus, the *ʾaṣl* of *lam yaku* 'he was not', *lā ʾadri* 'I do not know', *ġuzya r-rajulu* 'the man was assaulted', and *bal-ʿAnbar* 'son of ʿAnbar' is said to be *lam yakun, lā ʾadrī, ġuziya r-rajulu*, and *banū l-ʿAnbar* (Sībawayhi, *Kitāb* I, 8; II, 259, 382 and Mubarrad, *Muqtaḍab* I, 251; III, 167).

iii. The supposed, but not attested, origin of a certain form, pattern, etc. For example, the *ʾaṣl* of the word *ʾašyāʾ* 'things' is said to be *_šayʾāʾ_ according to the Basrans and *_ʾašyiʾāʾ_ according to the Kufans, but because of the perceived *ṯiqal* 'heaviness' of both proposed words, they were supposedly changed to a lighter form, hence *ʾašyāʾ* (Ibn al-ʾAnbārī, *ʾInṣāf* II, 812 ff.). Both proposed forms *_šayʾāʾ_ and *_ʾašyiʾāʾ_ are abstract underlying forms which the grammarians do not present as historical origins. Another example is that the negative particle *laysa* – which the grammarians classify as a verb – is said to have developed from the abstract and unattested *ʾaṣl*, *_layisa_, of the pattern *faʿila* which is common in triliteral verbs (Ibn Jinnī, *Munṣif* I, 258–259 and Ibn ʿUṣfūr, *Mumtiʿ* II, 440).

iv. In two related senses, *ʾaṣl* can refer to one of the radicals which form a root, or to the whole root. Sībawayhi, for example, describes triliterals as having the least number of radicals (*ʾaqall al-ʾuṣūl ʿadadan*; *Kitāb* II, 399). In this sense, *ʾaṣl* is contrasted with *zāʾid* 'augmented' to determine the root of a given word (Ibn Jinnī, *Munṣif* I, 11). The second sense may be exemplified by Mubarrad's

d. 285/898) text (_Muqtaḍab_ IV, 180) in which he justifies the impermissibility of using the exclamatory pattern _mā ʾafʿala-hu_ with quadriliterals by saying that if this were to be allowed, one radical would have to be elided from the _ʾaṣl_ (here, 'root').

As an analytical tool, _ʾaṣl_ is closely connected with _qiyās_. Sībawayhi often describes linguistic usage as being both the _ʾaṣl_ and the _qiyās_ (e.g. _Kitāb_ II, 53, 213, 421), or contrasts a usage with both terms (_Kitāb_ II, 214). More specifically, Sība-wayhi's use of the two terms shows that they share several features (Baalbaki 1988: 166–167): Both terms are (a) opposed to _šāḏḏ_ and _šawāḏḏ_ 'anomalous'; (b) linked to descriptions like _ḥasan_ 'well', _jayyid_ 'good', _ʾajwad_ 'better', _ʾakṯar_ 'more frequent', etc., and to the two terms _ḥadd_ and _wajh_ (both of which denote the most appropriate usage); (c) opposed to dialects which are described as _radīʾ_ 'bad', _xabīṯ_ lit. 'evil', _qalīl_ 'infrequent', etc.; (d) used in poetic license as forms that are otherwise unattested; and (e) described as pertaining to forms which have not undergone change.

The contrast between _ʾaṣl_ and change is nowhere more evident than in the study of morphology (_taṣrīf_). Several principles or guidelines related to this contrast are explicitly expressed by the grammarians, but are more often implicit in their discussion of forms which have undergone change. One of the most basic principles is briefly but succinctly formulated by Māzinī (d. 249/863) when he argues that the pattern _fiʿlā_ has to be considered an unchanged use (_ʿalā l-ʾaṣl_) as long as we are not certain that it is the result of change (_Taṣrīf_, in Ibn Jinnī, _Munṣif_ II, 163). In explaining this, Ibn Jinnī (d. 392/1002) establishes another basic principle, namely that there is no need for the grammarian to justify usage which is not the result of change, whereas the reason for change has to be indicated when usage departs from its _ʾaṣl_ (_iʿlam ʾanna mā jāʾa min hāḏā ʿalā ʾaṣli-hi fa-lā kalām fī-hi wa-ʾinnamā sabīl mā xaraja ʿan ʾaṣli-hi ʾan yunḏ̣ara ʾilā ʿillati-hi mā hiya_).

Another general principle in the study of _ʾaṣl_ and change is formulated by Ibn Jinnī in his assertion that forms which are proposed to be the _ʾaṣl_ of attested usage are not meant to represent a historical stage of the language which was later abandoned. Thus, when the grammarians presume that _*qawama, *bayaʿa, ʾaxwafa, *ʾaqwama, *istaʿwana_, and _*istaqwama_ are the origins of _qāma_ 'to rise', _bāʿa_ 'to sell', _ʾaxāfa_ 'to frighten', _ʾaqāma_ 'to set up', _istaʿāna_ 'to ask for help' and _istaqāma_, 'to stand upright' respectively, they do not mean to refer to any prior stage of the language. Rather, they argue that had these forms been in agreement with the norm, then one would have expected them to be used as proposed by the grammarians (Ibn Jinnī, _Munṣif_ I, 190–191; also _Xaṣāʾiṣ_ I, 256ff.). Further to this argument, Ibn Jinnī points out that a proposed form which represents the _ʾaṣl_ is often supported by the existence of anomalous (_šāḏḏ_) examples which do preserve that _ʾaṣl_, such as _istaḥwaḏa_ 'to overwhelm' which corresponds to the _ʾaṣl_ because it was not changed to _*istaḥāḏa_, contrary to other words of its class. He argues that such anomalous words serve as an indication (_manbaha_) of the original forms which preceded the introduced change (_fa-rubba ḥarf yaxruju hākaḏā manbaha ʿalā ʾaṣl bābi-hi_; _Xaṣāʾiṣ_ I, 257; cf. I, 161 and _Munṣif_ I, 191). It can be concluded from this line of thinking, which is also adopted by Ibn ʿUṣfūr (d. 669/1271) in his _Mumtiʿ_ and by ʾAstarābāḏī (ḍ. 686#1287) in his _Šarḥ_, that morphological change of an _ʾaṣl_ is often not exhaustive and that the _ʾaṣl_ can assert itself, so to speak, in rare but significant examples. In the light of this we can understand why the grammarians talk of forms that revert to the _ʾaṣl_, particularly in _nisba_ (gentilic) adjectives, such as _damawiyy_ ('bloody' and _ʾaxawiyy_ 'brotherly' (Sībawayhi, _Kitāb_ II, 79–80; cf. ʾAstarābāḏī, Šarḥ II, 61) and in diminutives, such as _muwayzīn_ and _muwayqīt_ (Sībawayhi, _Kitāb_ II, 125; cf. ʾAstarābāḏī, _Šarḥ_ I, 210), all of which are said to have the radical _wāw_ reinstated. Such forms are often cited as proof of the 'dominance' of the _ʾaṣl_ and of the correctness of the form or pattern which the grammarians postulate to represent each _ʾaṣl_.

Determining the _ʾaṣl_ of a certain usage was not an easy task for the grammarians. Not only did they differ on the proposed _ʾaṣl_, as in whether _*šayʾāʾ_ or _*ašyiʾāʾ_ is the _ʾaṣl_ of _ʾašyāʾ_ (see above), but at times they had to assign one of two actually attested forms as the _ʾaṣl_ of the other. This is especially true in the case of metathesis (_qalb_). The ideal situation, as one may conclude from Ibn Jinnī's discussion of the matter (_Xaṣāʾiṣ_ II, 69, 82), is that the two forms in question, such as _jaḏaba_ and _jabaḏa_ 'to attract', be interpreted as independent of one another since each of them has its own derivatives (e.g. *jāḏib, majḏūb*, and *jaḏb* versus *jābiḏ, majbūḏ*, and *jabḏ*). When this was not feasible, however, the grammarians had to establish several criteria to determine which form is the *ʾaṣl*. A discussion of these criteria may be found in Ibn ʿUṣfūr's *Mumtiʿ* (II, 617–618), but even some of the examples cited there, and hence the criteria they represent, are controversial. The reported difference between Sībawayhi and Jarmī (d. 225/840), for example, over whether *iṭmaʾanna* 'to be reassured' is the *ʾaṣl* of *ṭaʾmana* 'to reassure' or vice versa (*Mumtiʿ* II, 617–618 and Ibn Jinnī, *Xaṣāʾiṣ* II, 74; cf. Sībawayhi, *Kitāb* II, 130, 380) casts doubt on the validity of Ibn ʿUṣfūr's third criterion, which relates to whether the *ʾaṣl* is the form which has no augment as opposed to that in which the augment always appears.

From the general sense of 'origin' or 'principle', the term *ʾaṣl* developed in another direction and became associated with its antonym, *farʿ* 'branch, subsidiary'. Since this association only occurs once in Sībawayhi's *Kitāb*, and not in the sense used by later grammarians, but in the phonetic distinction between *ʾuṣūl* and *furūʿ* (*Kitāb* II, 404), it may be safe to conclude that Māzinī was the first author to have used *ʾuṣūl*, in the plural, as a technical term which refers to the 'fundamental' or main themes related to a certain grammatical topic. These are then opposed to the *furūʿ* or the subsidiary or comparatively secondary questions and problems which came to be known generally as *masāʾil* (see *Taṣrīf* I, 340 where *ʾuṣūl* is opposed to both *furūʿ* and *masāʾil*, and II, 208, 251–252, 340 where the boundaries between *ʾuṣūl* and *masāʾil* are clearly set). Under *ʾuṣūl*, Māzinī mentions the general principles which dominate the issue at hand, or *bāb* lit. 'chapter', whereas he reserves the particulars, especially the complex questions and intricate examples, for discussion under *furuʿ*.

Mubarrad seems to have adopted this distinction from his teacher, Māzinī, and generalized it from the realm of morphology, with which Māzinī was concerned, to the realm of syntax. Mubarrad maintains that to know the *ʾuṣūl* is to achieve perfection and mastery of a *bāb*; the particulars or *masāʾil* embraced within that *bāb* are then to be judged by reference to those *ʾuṣūl* (*Muqtaḍab* IV, 172; cf. Baalbaki 1988:172–173). Accordingly, he often divides his subject into two separate headings, *ʾuṣūl* followed by *masāʾil* (see *Muqtaḍab, bāb al-fāʾ*, II, 14–24; *bāb ḥattā*, II, 38–43; *bāb ʾam wa-ʾaw*, III, 286–300; *bāb al-fiʿl al-mutaʿaddī*, IV, 86–106). Obviously, this distinction between *ʾuṣūl* and *furūʿ* or *masāʾil*, which is absent from Sībawayhi's *Kitāb*, is an early step toward the classification of grammatical questions according to some logical foundation which proceeds from the general to the particular. Furthermore, this classification is basically a didactic technique, and its absence from the *Kitāb* argues against a didactic aim. Sentences of the type *wa-naqūlu fī masāʾil ṭiwāl yumtaḥanu bi-hā l-mutaʿallimūna* 'we pronounce on complex constructions by which learners are examined' are frequent titles in *Muqtaḍab*; e.g. I, 22; II, 62; IV, 59) but are totally alien to Sībawayhi's method.

Mubarrad's student, Ibn as-Sarrāj (d. 316/929), took the distinction between *ʾuṣūl* and *furūʿ* or *masāʾil* one step further and was, as far as we know, the first grammarian who devised his book on the basis of *ʾuṣūl* and thus gave it the title *Kitāb al-ʾuṣūl fī n-naḥw*. He clearly states his plan in a number of places (e.g. *ʾUṣūl* I, 36, 328, 381) and asserts that – if he were to live long enough – he intended to write *Kitāb al-furūʿ* in order to assemble the *furūʿ* which fall under the *ʾuṣūl* he gathered in his first book (*ʾUṣūl* I, 328). The importance which Ibn as-Sarrāj attaches to *ʾuṣūl* is obviously linked to his system of *qiyās* as he believes that what is contrary to *ʾuṣūl* is surely contrary to *qiyās* (*ʾUṣūl* I, 406; cf. I, 56–57). The clear-cut distinction which Ibn as-Sarrāj establishes between *ʾuṣūl* and *furūʿ* throughout his book is probably the main reason for the famous saying that he has, by his *ʾuṣūl* (or perhaps *ʾUṣūl*, i.e. the work itself), rationalized grammar: *mā zāla n-naḥw majnūnan ḥattā ʿaqqalahu Ibn as-Sarrāj bi-ʾuṣūli-hi* (Yāqūt, *Muʿjam* VI, 2535). It is noteworthy, however, that Ibn as-Sarrāj's method of dividing his chapters into *ʾuṣūl* and *masāʾil* was generally not continued by the later grammarians, who seem to have adopted other ways for the internal division of their chapters.

The relationship between *ʾaṣl* and *farʿ* also features in the four elements which, according to the grammarians, constitute *qiyās* (here, 'analogical extension'). These are, in addition to *ʾaṣl* and *farʿ*, *ḥukm* 'rule', and *ʿilla*, the latter being the reason which justifies the application, to a *farʿ*, of the rule which is characteristic of an *ʾaṣl* (Ibn al-ʾAnbārī, *Lumaʿ* 93). According to Suleiman (1999:15). "this constituency relation between *ʿilla* and *qiyās* places the study of *taʿlīl* within the wider framework of *ʾuṣūl an-naḥw* (the fundamental principles of grammar), which, in addition to *qiyās*, is said to include the principles of *samāʿ* (attestation, attested data) and, depending on the orientation of the grammarian, *ʾijmāʿ* (consensus) and/or *ʾistiṣḥāb al-ḥāl* (presumption of continuity)". The interest in the study of *ʾuṣūl an-naḥw* gave rise to a genre of writing which was distinct from the descriptive account of the language, and which formed the explanatory component of the grammatical tradition. Titles like Ibn al-ʾAnbārī's (d. 577/1181) *Lumaʿ al-ʾadilla fī ʾuṣūl an-naḥw* and Suyūṭī's (d. 911/1505) *al-ʾIqtirāḥ fī ʿilm ʾuṣūl an-naḥw* represent this trend which tried to provide explanations and interpretations within the general framework of the grammatical theory and generally to highlight the logic behind linguistic phenomena.

Finally on the relationship between *ʾaṣl* and *farʿ*, it has been noted that the dichotomy *ʾaṣl/farʿ* resembles the dichotomy marked/unmarked in modern linguistics (Owens 1988: 119–226). In particular, Owens (ibid., 220ff.) believes that it is worthwhile to study the distinction *ʾaṣl/farʿ* in the light of modern generative-transformational grammar. The problem with the identification of *ʾaṣl* with the underlying structure of modern linguistics, as Versteegh (1995:238) notes, is that, although the Arab grammarians call the underlying structure *ʾaṣl*, "this does not mean that they regard the surface structure as a derived one".

The term *ʾaṣl* also occurs in other contexts. In a phonetic context, it is used in the expression *ʾaṣl/ʾuṣūl aṯ-ṯanāyā* to refer to the upper part of the central incisors from which /ṭ/, /d/, and /t/ are uttered (Sībawayhi, *Kitāb* II, 405, 419). In etymology, it refers to the origin from which a word is borrowed, hence expressions like *ʾaʿjamiyy al-ʾaṣl* 'of foreign/Persian origin'; *Kitāb* II, 342; Jawālīqī, *Muʿarrab* 5) or *ʾaṣlu-hu bi-l-fārisiyya/bi-n-nabaṭiyya* 'its origin in Persian or Nabatean'; *Muʿarrab* 16). *ʾAṣl* is also connected with the term *waḍʿ* ('positing', 'coining', particularly in expressions such as *ʾaṣl al-waḍʿ* or *mā wuḍiʿa fī l-ʾaṣl* which refer to an original usage or coinage. (For this, and other uses of *ʾaṣl* in sources including Ibn as-Sarrāj's *ʾUṣūl*, Ibn Jinnī's *Xaṣāʾiṣ* and Suyūṭī's *ʾIqtirāḥ*, see an-Nawājī 2001:132–134, 141–154, 209–210, 222, 224.) A special sense of *ʾaṣl* is systematically used by Ibn Fāris (d. 395/1004) in his dictionary *Maqāyīs al-luġa*. In an attempt to demonstrate the semantic relationship between some of the words derived from the same root, and at the same time distinguish semantically between them and other words which are also derived from that root, Ibn Fāris describes roots in term of the number of their *ʾuṣūl*. A root like *ʾ-w-r*, for example, is said to be one *ʾaṣl* because all its derivatives, according to Ibn Fārīs, indicate the general sense of *ḥarr* 'hotness' (*Maqāyīs* I, 155–156). The root *b-ḍ-ʿ* (*Maqāyīs* III, 254–257), on the other hand, is said to comprise three *ʾuṣul*, the first of which indicates a 'part of' , as in *baḍʿa* 'piece of meat', the second indicates a 'spot; location', as in *baḍīʿ* 'sea; island', and the third indicates 'healing', as in *baḍʿ* 'quenching of thirst'. The division of these *ʾuṣūl*, however, is left entirely to the imagination of the author and is often evidently farfetched and unconvincing.

Bibliographical references

Primary sources

ʾAstarābāḏī, *Šarḥ aš-Šāfiya* = Raḍī d-Dīn Muḥammad ibn Ḥasan al-ʾAstarābāḏī, *Šarḥ Šāfiyat Ibn al-Ḥājib*. Ed. Muḥammad Nūr al-Ḥasan, Muḥammad az-Zafzāf, and Muḥammad Muḥyī d-Dīn ʿAbd al-Ḥamīd. 4 vols. Repr. from the Cairo edition, Beirut: Dār al-Kutub al-ʿIlmiyya, 1982.

Ibn al-ʾAnbārī, *ʾInṣāf* = ʾAbū l-Barakāt ʿAbd ar-Raḥmān ibn Muḥammad al-ʾAnbārī, *al-ʾInṣāf fī masāʾil al-xilāf bayna n-naḥwiyyīn al-Baṣriyyīn wa-l-Kūfiyyīn*. Ed. Muḥammad Muḥyi d-Dīn ʿAbd al-Ḥamīd. 2 vols. Cairo: al-Maktaba at-Tijāriyya, 1955.

Ibn al-ʾAnbārī, *Lumaʿ* = ʾAbū l-Barakāt ʿAbd ar-Raḥmān ibn Muḥammad al-ʾAnbārī, *Lumaʿ al-ʾadilla fī ʾuṣūl an-naḥw*. Ed. Saʿīd al-ʾAfġānī. 2nd ed. Beirut: Dār al-Fikr, 1971.

Ibn ʿAqīl, *Šarḥ* = Bahāʾ ad-Dīn ʿAbdallāh Ibn ʿAqīl, *Šarḥ Ibn ʿAqīl ʿalā ʾAlfiyyat Ibn Mālik*. Ed. Ramzī Munīr Baʿalbakī. Beirut: Dār al-ʿIlm li-l-Malāyīn, 1992.

Ibn Fāris, *Maqāyīs* = ʾAbū l-Ḥusayn ʾAḥmad ibn Fāris, *Maqāyīs al-luġa*. Ed. ʿAbd as-Salām Muḥammad Hārūn. 6 vols. Cairo: ʿĪsā al-Bābī al-Ḥalabī, 1946–1952.

Ibn Jinnī, *Munṣif* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Munṣif, šarḥ Kitāb at-taṣrīf li-l-Māzinī*. Ed. ʾIbrāhīm Muṣṭafā and ʿAbdallāh ʾAmīn. 3 vols. Cairo: Muṣṭafā al-Bābī al-Ḥalabī, 1954–1960.

Ibn Jinnī, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. Ed. Muḥammad ʿAlī an-Najjār. 3 vols. Cairo: Dār al-Kutub al-Miṣriyya, 1952–1956.

Ibn as-Sarrāj, *ʾUṣūl* = ʾAbū Bakr Muḥammad ibn Sahl Ibn as-Sarrāj, *al-ʾUṣūl fī n-naḥw*. Ed. ʿAbd al-Ḥusayn al-Fatlī. 3 vols. Beirut: Muʾassasat ar-Risāla, 1985.

Ibn ʿUṣfūr, *Mumtiʿ* = ʾAbū l-Ḥasan ʿAlī ibn Muʾmin Ibn ʿUṣfūr al-ʾIšbīlī, *al-Mumtiʿ fī t-taṣrīf*. Ed. Faxr ad-Dīn Qabāwa. 4th ed. 2 vols. Beirut: Dār al-ʾĀfāq al-Jadīda, 1979.

Jawālīqī, *Muʿarrab* = ʾAbū Manṣūr Mawhūb ibn ʾAḥmad al-Jawālīqī, *al-Muʿarrab min al-kalām al-ʾaʿjamī ʿalā ḥurūf al-muʿjam*. Ed. ʾAḥmad Muḥammad Šākir. Cairo: Dār al-Kutub al-Miṣriyya, 1361 A.H.

Māzinī, *Taṣrīf*. In Ibn Jinnī, *Munṣif*.

Mubarrad, *Muqtaḍab* = ʾAbū l-ʿAbbās Muḥammad ibn Yazīd al-Mubarrad, *al-Muqtaḍab*. Ed. Muḥammad ʿAbd al-Xāliq ʿUḍayma. 4 vols. Cairo: Dār at-Taḥrīr, 1965–1968.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. 2 vols. Būlāq, 1316–1317 A.H.

Suyūṭī, *Iqtirāḥ* = Jalāl ad-Dīn ʾAbū l-Faḍl ʿAbd ar-Raḥmān ibn ʾAbī Bakr as-Suyūṭī, *al-Iqtirāḥ fī ʿilm ʾuṣūl an-naḥw*. Ed. ʾAḥmad Muḥammad Qāsim. Cairo: Maṭbaʿat as-Saʿāda, 1976.

Yāqūt, *Muʿjam* = Šihāb ad-Dīn ʾAbū ʿAbdallāh Yāqūt ibn ʿAbdallāh ar-Rūmī al-Ḥamawī, *Muʿjam al-ʾudabāʾ*. Ed. ʾIḥsān ʿAbbās. 7 vols. Beirut: Dār al-Ġarb al-ʾIslāmī, 1993.

Secondary sources

Baalbaki, Ramzi. 1988. “A contribution to the study of technical terms in early Arabic grammar: The term *aṣl* in Sībawayhi’s *Kitāb*”. *A miscellany of Middle Eastern articles in memoriam Thomas Muir Johnstone*, ed. A.K. Irvine, R.B. Serjeant, and G. Rex Smith, 163–177. Harlow, Essex: Longman.

Nawājī, ʾAšraf Māhir an-. 2001. *Muṣṭalaḥāt ʿilm ʾuṣūl an-naḥw*. Cairo: Dār Ġarīb.

Owens, Jonathan. 1988. *The foundations of grammar: An introduction to medieval Arabic grammatical theory*. Amsterdam and Philadelphia: J. Benjamins.

Suleiman, Yasir. 1999. *The Arabic grammatical tradition: A study in* taʿlīl. Edinburgh: Edinburgh University Press.

Troupeau, Gérard. 1976. *Lexique-index du* Kitāb *de Sībawayhi*. Paris: Klincksieck.

Versteegh, Kees. 1995. *The explanation of linguistic causes: Az-Zaǧǧāǧī’s theory of grammar*. Amsterdam and Philadelphia: J. Benjamins.

Ramzi Baalbaki
(American University of Beirut)

Aspect

1. The definition of ‘aspect’

Aspect is a semantic category of a verb or proposition which describes the internal consistency of a verbal event. That is, it describes the event as either being a complete whole, an incomplete, ongoing process, or a state of some sort. Like time reference with which it interacts in important ways, it may be grammaticalized in morphological verb forms, or through compound verb forms, which may enter into quasi-morphological (or morphosyntactic) paradigms with simple verb forms, which is here termed ‘formal aspect’. It may also enter into an expression through the aspect of a basic lexical item, which is often referred to as → Aktionsart ‘mode of action’ or simply ‘lexical aspect’. In addition to these two formalized avenues of use, the aspect of an expression may also be affected by the types of noun phrases which are part of the predicate (definite singular nouns may often give rise to more ‘perfective’ readings, while indefinite plurals may give rise to more ‘imperfective’ readings), or by the types of adverbials present (the use of an explicitly habitual adverbial such as ‘every day’ may be all that is necessary to provide a habitual reading to a predicate).

While grammaticalized time reference may be designated as ‘tense’, the term ‘aspect’ is used to refer to all of these categories that describe the internal consistency of an event, which often leads to confusion. To avoid this, in this discussion only the semantic category will be referred to as ‘aspect’ while grammaticalized aspect will be referred to as ‘formal aspect’, and lexicalized aspect will be referred to as lexical aspect (Aktionsart). It is important to distinguish each of these categories of aspect one from the other, and to distinguish aspect itself from time reference (to the extent that this is possible) since each of them has been confused or conflated with the others especially in dealing with aspect and tense in Arabic. This is due, in part, to the ways in which tense and aspect interact crucially to impart a particular meaning to a verbal expression, as well as to the peculiarities of the Semitic verbal system which Arabic inherited.

The former is shown most clearly in the contradictory ways in which researchers have viewed the aspectual nature of a present tense or time reference: some (e.g. on German, Koschmieder 1929; on Arabic, Wild 1964; Denz 1971; Woidich 1975) see the ‘present tense’ as necessarily having an ‘interval’ aspect, while others (on English, Dowty 1979; Carlson 1977) view the ‘present’ as necessarily having the aspectual value of a ‘point in time’. Each of these researchers conflated the preferred aspectual value of a particular tense form in a particular language with the semantic category of ‘present time reference’ and generalized it to all types of present time reference. To avoid this pitfall, one must carefully note the contributions of each semantic category to the overall meaning of a phrase, and note as well the different pathways along which these various elements of meaning are introduced into the phrase.

The latter point is shown in the uses of the various forms of the Arabic 'imperfect' verb (*al-muḍāriʿ* or 'the (one) similar [to the noun]': the jussive or apocopated form (*al-muḍāriʿ al-majzūm: yaktub*) is confined to contexts which express past time or to a domain of usage which overlaps with that of the past tense verb, namely conditionals and past time negation after the particle *lam*, while the indicative imperfect (*al-muḍāriʿ al-marfūʿ: yaktubu*) is used to refer by itself to non-past situations (present or future), and the subjunctive imperfect (*al-muḍāriʿ al-manṣūb: yaktuba*) is used in apparently non-temporal contexts following the particle *ʾan* (after certain verbs) as well as after the particle *lan* to express future negation. Compounding the situation was the fact that the earliest analyses of Arabic took place in the context of comparative Semitic studies in which the study of Biblical Hebrew was of primary importance, and the Arabic verbal system was taken to be similar to, if not the same as, the Hebrew verbal system. Given the complications that the '*waw*-conversive' brings to the analysis of Hebrew verbs (a complication which does not exist for Arabic), such a conflation was unfortunate. Thus, owing to the seemingly contradictory semantic domains of these verb forms in Arabic and Hebrew, many Arabists in the 19th and 20th centuries began to apply the notion of 'aspect' to these forms, 'aspect' being the English translation of the Russian term *vid*, which had been newly minted to describe the very highly developed system of aspect-based derivational processes in Russian and other Slavic languages (Binnick 1991:136). The dominant belief since that time is that Arabic verbs do not denote time reference and are therefore not tenses, but rather express aspect (e.g. Ewald 1831; Caspari 1848; Wright 1874; M. Cohen 1924; and more recently Fleisch 1957; D. Cohen 1989; and most if not all of the recent textbooks on Arabic). Despite the dominance of the aspectualist position, however, some Arabists have held that Arabic verbs do express temporal notions. Prominent among them were Bauer (1910), Reckendorf (1895), who recognized that at least the perfect is a tense, Gaudefroy-Demombynes and Blachère (1952), and most especially Aartun (1963), who provided clear evidence that, in most instances, the Arabic perfect verb *does* refer to past time, while the imperfect, in most contexts, *does* refer to a non-past (present or future) time, that is, Arabic verbs prototypically *do* grammaticalize time reference.

However, whether arguing for or against a tense or aspect position, the notion of aspect itself has remained ill-defined in discussions of aspect in Arabic. It is important to note in this regard that the notion of 'aspect' is not found in traditional Arabic discussions of verb functions, and is only implicit in ancient and medieval Western grammatical traditions, which are the source for terms such as perfect(ive), imperfect(ive), aorist, etc. It was explicitly developed initially to describe the system of verbal derivation in Slavic in the early 19th century, and was quickly extended to analyses of Germanic languages and Semitic languages as well.

In the following (based on Eisele 1999), the first two kinds of 'aspect' are described: 'formal aspect', or the aspect associated with a morphological form, and 'lexical aspect', or the aspect associated with the lexical entry of a root or stem. The way that these two categories interact provides important clues both about the meaning of the forms themselves, as well as about the types of lexical classes which exist in the language. The categories of formal and lexical aspect are then related to the temporal schemata developed in Reichenbach (1947) to clarify how aspect interacts with time reference to bring about a particular reading for a verbal expression.

2. Formal aspect classes

In Arabic, verbal forms express rather mild aspectual features, while more detailed aspectual information is provided by the lexical item (through its Aktionsart or lexical aspect), preverbal particles, syntactic context, or the pragmatics of the situation. This varies depending on the type of Arabic being considered: in Classical or Modern Standard Arabic, very little of what is taken to be aspectual is actually due to the verb forms themselves. Most of what is considered aspectual in nature derives from the lexical item itself, or the syntactic or semantic context (the sentence or proposition as a whole.) The verb form in most Arabic dialects (expanded to include the active participle and aspectual verbal prefixes) does more overtly express aspectual notions and there are important aspectual distinctions between forms, although much aspectual information is still delivered through the lexical base as well as the context (syntactic, semantic, and pragmatic).

In general discussions of aspect, there are three general categories of formal aspect: events, processes, and states. A form whose primary aspectual value is 'event' represents the state of affairs predicated of an entity as a single, complete, whole event. Whether it is dynamic or not depends on the lexical item, and whether it has a beginning, middle, or end will depend on the lexical item, not on the verb form. In a sense, it seems to represent the lexical item in the simplest, most transparent fashion. It is not necessarily 'perfective' in the sense of Slavic perfective forms, which often indicate completion or fulfillment of some action; rather an event form portrays an action or activity described by a lexical item as a single whole. In actual fact, an event form may be described as a 'non-process' or 'non-aspectual' form as well.

A 'processive' form, or a form whose primary aspectual value is 'process', represents the state of affairs predicated of an entity as an event occurring over a heterogeneous interval, or one which is true only at intervals larger than a moment or point in time, and not true at all moments within an interval. It will therefore necessarily occupy an interval, without necessarily indicating the end of the process, if there is a specific end associated with the lexical item. In terms of 'beginning-middle-end' it could be described as portraying the 'middle' of a heterogeneous situation. As noted above, the English progressive form is the most commonly cited example of a processive form, while in Arabic one of the two basic meanings associated with the imperfect verb form is processive aspect.

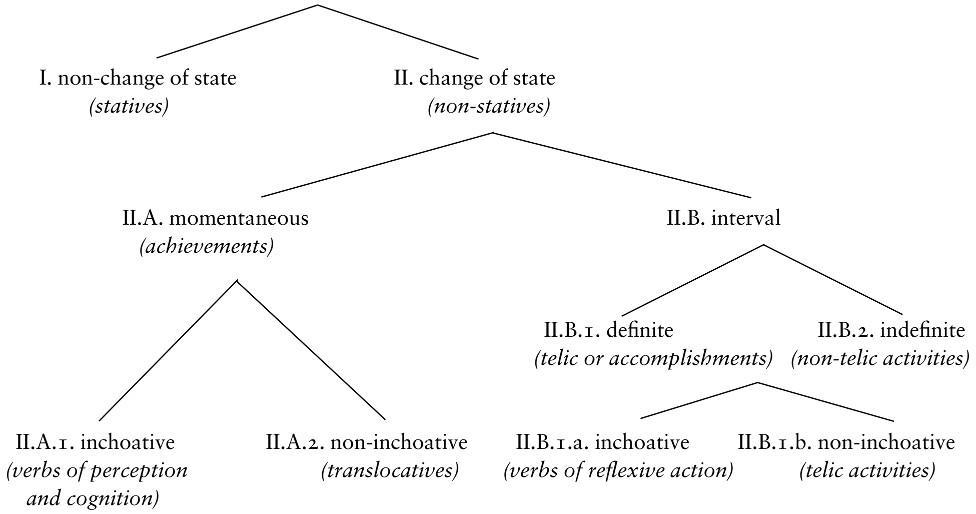
A 'stative' form, or a form with the primary aspectual value of 'state' represents the state of affairs predicated of an entity as homogeneous or true at all points or moments within an interval. There are no finite verb forms in English or in Arabic whose aspectual value is stative. Rather stative forms are verbal adjectives: in English the passive participle and in Arabic the active and passive participles, as well as what are termed 'quasi-active participles' (e.g. *faˁlān* participles). In addition, non-verbal equational sentences in Arabic are syntactically stative, even though there may be no stative verbal form associated with these sentences. It is important to note that the expression of a stative lexical item or predicate does not require a stative form per se. Lexically stative items can be expressed through an event form, as can lexically non-stative items, since there is nothing contained in the definition of event form which would prevent such a usage. As noted above, an event form is the most basic, or unmarked, and aspectually transparent of forms. As such, it can represent static or dynamic lexical items equally well.

3. Lexical aspect classes

Lexicalized verbal expressions are the other primary source for the articulation of aspect in an expression. When this is confined to a single lexical item (or an idiomatized phrase) this is termed Aktionsart or lexical aspect. However, in most cases the expression of lexical aspect is taken to include the verbal phrase as a whole (including direct objects), i.e. the predicate. The categories described in the discussion which follows are meant to subsume not just single verbal lexical items, but predicate expressions as a whole. Also, categories of lexical aspect tend to express much finer nuances of meaning and there are thus more of them, at least for Arabic and English, than categories of formal aspect. The following discussion exemplifies these lexical aspectual categories with reference to Egyptian Arabic and is taken from Eisele (1999), which in turn is derived from Dowty (1979).

Dowty (1979) rephrased Vendler's (1967) Aristotelian based categories of accomplishment, achievement, activity, and state in terms of notions such as change of state, intervals and the internal consistency of these intervals, based upon a series of syntactic and semantic tests. (These categories are summarized in Figure 1.) The first division of categories divides those predicates which involve a change of state (II: Vendler's achievements, activities, and accomplishments) from those which do not (I: statives). In addition to being non-change of state, statives generally denote a homogeneous interval, or one which is true at all moments within it. The test for this distinction in English is the non-use of stative predicates in *do*-constructions (e.g. pseudo-cleft: 'What John did was . . .'), and their non-use in the progressive. One test for this in Egyptian Arabic is similar to the latter test: stative predicates in the *bi*-imperfect form have only

Figure 1. Aspectual Classification of Predicates


I. non-change of state
(statives)
II. change of state
(non-statives)
II.A. momentaneous
(achievements)
II.B. interval
II.B.1. definite
(telic or accomplishments)
II.B.2. indefinite
(non-telic activities)
II.A.1. inchoative
(verbs of perception and cognition)
II.A.2. non-inchoative
(translocatives)
II.B.1.a. inchoative
(verbs of reflexive action)
II.B.1.b. non-inchoative
(telic activities)

a characteristic or habitual reading (or in other words, they may only be used with a non-specific time reference, e.g. *kull*^{i} *yōm* 'every day'). Among the other tests for this class is the reading of the active participle form (AP) when used with the adverb *lissa* 'still, just': for this class of predicates *lissa* + AP has only a 'still' reading, not a 'just' one. (The *bi*-imperfect of non-stative predicates, on the other hand, may have a real present reading or a habitual one, while *lissa* + AP of these predicates may have either a 'just' only reading or both 'still' and 'just').

The second categorial distinction is within the category of change of state predicates, distinguishing between predicates whose associated events take place over an interval (II.B: Vendler's accomplishments and activities) versus those whose event is momentaneous (II.A: Vendler's achievements). Among the tests for momentaneous predicates are the following: their *bi*-imperfect has only a habitual reading, their perfect form used with the adverb *fi sāʿa* 'in an hour' does not entail the *bi*-imperfect during that interval (i.e. 'x V-d in y time' does not entail 'x was V-ing during y time'), and use of their perfect with *sāʿa* 'for an hour' is unacceptable.

Interval predicates, which have the opposite readings and implications from those of momentaneous predicates, are further divided into definite change of state predicates (II.B.1: Vendler's accomplishments, Comrie's telic predicates) versus indefinite change of state predicates (II.B.2: Vendler's activities). (It is important to note that momentaneous predicates [achievements] also indicate a definite change of state, or one which does have an upper bound.) Indefinite change of state predicates or activities on the other hand do not involve an upper bound. There are two important tests for distinguishing definite change of state (or 'telic') interval predicates from indefinite change of state (or 'non-telic') interval predicates. The first involves use of the perfect form of these verbs with a telic adverbial such as *fi sāʿa* 'in an hour': if a predicate may be used in its perfect form with *fi sāʿa* 'in an hour', then it is likely to fall into the category of definite change of state, and if not, then it is likely to be an indefinite change of state predicate. In addition, if the *bi*-imperfect of a predicate entails its corresponding perfect, then it is likely to be a definite change of state predicate ('x was V-ing' [Progressive] entails 'x has V-ed' [Past]), and if not, then it is likely to be an indefinite change of state.

Under both momentaneous predicates (II.A) and definite interval predicates (II.B.1) there is a further subdivision between inchoative and non-inchoative verbs. All inchoatives of both classes are distinguished from non-inchoatives by the reading of their Active Participle forms: the active participles of inchoative verbs are understood as referring to a present state, while those of non-inchoatives are understood as indicating a resultative or perfective state of some sort, i.e. they have a past connotation. In addition, the active participle of inchoative verbs used with the adverbial *lissa* may have both a 'just' or a 'still' reading, while the active participle of non-inchoative verbs may have only a 'just' reading. Inchoatives are distinguished from each other based on their readings in the *bi*-imperfect: the *bi*-imperfect of momentaneous inchoatives may have only a habitual reading, while the *bi*-imperfect of interval inchoatives may have either a habitual or a present progressive reading, indicating that interval inchoatives (II.B.1.a), like other inchoatives, point to entry into a physical or cognitive state, but unlike momentaneous inchoatives the 'entry' or change in state takes place over an interval.

In opposition to inchoatives in both the momentaneous category and the interval definite change of state category, there is a non-inchoative class of predicates. Momentaneous non-inchoatives (II.A.2) include a particularly prominent class of predicates, namely translocatives (verbs which indicate movement from place to place). Like other momentaneous predicates, their *bi*-imperfect has only a habitual reading, while their active participle may have either a past, perfective, present, or a future reading, depending upon the context.

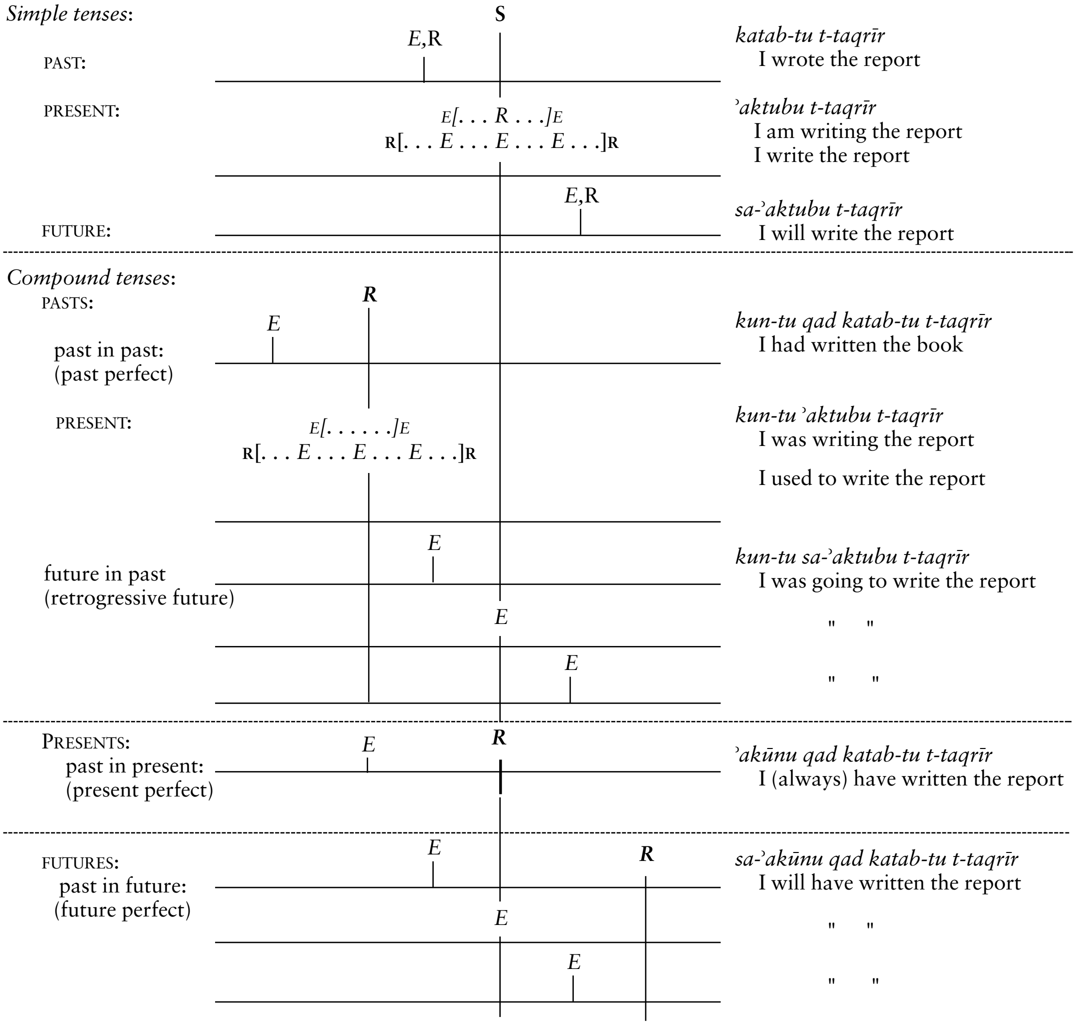
Interval non-inchoatives (II.B.1.b) correspond in large part to Vendler's accomplishments. They are characterized by having *bi*-imperfects that have a present processive reading, and active participles which are understood as past (i.e., as a resultative, meaning that a past event is strongly implied by the active participle), and active participles which have a reading of 'just' only, with *lissa*. The *bi*-imperfect of these predicates, in contrast to indefinite change of state predicates, do *not* entail the corresponding perfect.

The classes denoted by this classification scheme overlap to some extent with classification schemes proposed for Arabic in Cowell (1964:265-276), McCarus (1976), Tonsi (1980), and Woidich (1975), but there are important differences. For example, Cowell's 'developmental' predicates appear to be the same as the inchoative classes (II.A.1 and II.B.1.a), while his 'punctual' class is similar to II.A. momentaneous predicates in this scheme. Cowell's 'inceptive' class is harder to relate to this, but it appears to subsume interval non-inchoatives (II.B.1.b and II.B.2). While each of the classification schemes proposed in the above works delimits somewhat different sets of predicates, each has a certain validity within the framework of the analysis and with regard to the kind of Arabic being described. What the system described here provides is a more precise way of cross linguistic comparison, as well as a way to differentiate issues of aspectual reference more clearly from those of temporal reference.

It should be clear from the preceding discussion that an important facet of the tests used to distinguish classes of predicates according to their 'lexical aspect' is how they behave in particular verb forms, or, in other words, how they interact with the grammaticalized aspect of a particular verb form, as well as with the particular time reference which may be associated with these forms. This may be clarified by relating aspect and time reference explicitly based on Reichenbach (1947) (Figure 2) where the time reference associated with a predicate utterance is defined in terms of three basic times: speech time (S), reference time (R), and event time (E). The reference time is the pivotal time point to which both the speech time and the event time are related: deictic time reference is defined in terms of the relation between speech time and reference time, while non-deictic time reference is defined in terms of the relation between event time and reference time. Speech time (S) is given by the context of the utterance. The reference time (R) is given either explicitly through the use of deictic time adverbials or is derivable from the context. The event time (E) is dependent for its determination upon the reference time. In certain circumstances a time adverbial can make the event time explicit, but this is not usually the case. The event time is, quite simply, the time of application or occurrence of the action/state described by the verb. Time references are defined in terms of the relationships between these three times. For example, *past deictic time reference* is defined as R < S, or 'reference time prior-to speech time', *future deictic time reference* as S < R, or 'reference time subsequent-to speech time', etc. for other time references. Individual verb forms (tenses) may then grammatically encode these different time references.

In addition to the above relationships between E, R, and S, the tense form is further affected by the individual character of the R-time and by the nature of the event associated with the E-time. The latter determines the aspect of the form, while the former determines whether or not it is to be understood with a habitual or characteristic reading. For example, in a processive form (such as the Egyptian Arabic *bi*-imperfect), the

Figure 2. Aspect related to Reichenbach's temporal schemata


Simple tenses:
S
PAST:
E,R
katab-tu t-taqrīr
I wrote the report
PRESENT:
E[. . . R . . .]E
R[. . . E . . . E . . . E . . .]R
ʾaktubu t-taqrīr
I am writing the report
I write the report
FUTURE:
E,R
sa-ʾaktubu t-taqrīr
I will write the report
Compound tenses:
PASTS:
R
E
past in past:
(past perfect)
kun-tu qad katab-tu t-taqrīr
I had written the book
PRESENT:
E[. . . . . .]E
R[. . . E . . . E . . . E . . .]R
kun-tu ʾaktubu t-taqrīr
I was writing the report
I used to write the report
future in past
(retrogressive future)
E
E
E
kun-tu sa-ʾaktubu t-taqrīr
I was going to write the report
" "
" "
PRESENTS:
past in present:
(present perfect)
E
R
ʾakūnu qad katab-tu t-taqrīr
I (always) have written the report
FUTURES:
past in future:
(future perfect)
E
R
E
E
sa-ʾakūnu qad katab-tu t-taqrīr
I will have written the report
" "
" "

event associated with the E-time is a process, meaning that it involves a change of state over at least two different points in time, which means that the E-time in this case will necessarily be an interval. In a form characterized as an event, the event associated with the E-time may be of several different types, depending on the type of lexical aspect involved, which is an indication of the transparency of this form. For example, with lexical items that involve a momentaneous change of state, the event associated with the E-time will necessarily be a momentaneous change of state, in which case the event time will be a moment or point in time. For lexical items that involve a change of state over at least two different points in time, the event will necessarily describe an activity, in which case the E-time will be an interval. If the latter is in addition a telic activity or accomplishment (one that involves a definite end or result such as 'write a book'), then the E-time interval will be a closed one. Finally, if the lexical item is a stative (or non change of state), then what is associated with the E-time is actually a state and not a change of state event.

For truly stative (or non-change of state) forms, what is related to the R-time is not an event conceived of as a change of state, but is rather necessarily a state. The difference between a morphological stative form (i.e. a grammaticalized stative) and a lexical stative expressed through a morphological event form is a formal one – semantically they should express the same idea. The formal difference is that the grammaticalized stative can only express stativity, whereas the final aspect of the event form is dependent on the lexical aspect of the verb involved. In English there is no productive verb form (i.e. one which is a part of the verbal system and enters into temporal oppositions) which can be called purely stative (although perhaps the passive participle might come close), but there are many adjectival forms which can perform this function in the right context (e.g. in copular or equational sentences). In Arabic, however, there are productive forms which are purely stative, namely, the active participle and quasi-participial forms. With these forms it is a state which is associated with the reference time. These forms however in many instances may also have an event (or 'change of state') associated with them in a predictable but non-grammaticalized way, which may give rise for certain lexical classes of predicates to 'resultative' or present perfect reading with the active participle (e.g. with interval non-inchoatives).

Finally, in the analysis presented here, iterative aspect (namely, one which has a habitual or characteristic reading) is not grammaticalized in a verb form in English or in Arabic. The habitual or iterative readings often associated with an imperfect form (in any type of Arabic) are not due to the form itself (or just to the form itself) – rather, they are due to the kind of time reference associated with a predicate in a particular context. In terms of R, S, and E, it has to do with the kind of R, or reference point, associated with the predication: it is a non-specific one, or one which does not tie or anchor the utterance to a unique and delimitable point or interval of the temporal context.

BIBLIOGRAPHICAL REFERENCES

Aartun, Kjell. 1963. *Zur Frage altarabischer Tempora*. Oslo: Universitetsforlaget.

Bauer, Hans. 1910. *Die Tempora im Semitischen: Ihre Entstehung und ihre Ausgestaltung in den Einzelsprachen*. Ph.D. diss., University of Leipzig.

Binnick, Robert. 1991. *Time and the verb*. Oxford: Oxford University Press.

Carlson, Gregory. 1977. *Reference to kinds in English*. Ph.D. diss., University of Massachusetts, Amherst.

Caspari, Carl Paul. 1848. *Grammatica arabica in usum scholarum academicarum*. Leipzig: C.L. Fritzschii.

Cohen, David. 1989. *L'aspect verbal*. Paris: Presses Universitaires de France.

Cohen, Marcel. 1924. *Le système verbal sémitique et l'expression du temps*. Paris: Imprimerie Nationale.

Comrie, Bernard. 1976. *Aspect*. Cambridge: Cambridge University Press.

——. 1985. *Tense*. Cambridge: Cambridge University Press.

Cowell, Mark. 1964. *A reference grammar of Syrian Arabic*. Washington, D.C.: Georgetown University Press.

Denz, Adolf. 1971. *Die Verbalsyntax des neuarabischen Dialektes von Kwayriš (Irak), mit einer einleitenden allgemeinen Tempus- und Aspektlehre*. Wiesbaden: F. Steiner.

Dowty, David. 1979. *Word meaning and Montague grammar*. Dordrecht: D. Reidel.

Eisele, John. 1999. *Arabic verbs in time: Tense and aspect in Cairene Arabic*. Wiesbaden: O. Harrassowitz.

Ewald Heinrich. 1831–1833. *Grammatica critica linguae Arabicae: Cum brevi metrorum doctrina*. Leipzig: Sumptibus Librarie Hahnianae.

Fleisch, Henri. 1957. "Etudes sur le verbe arabe". *Mélanges Louis Massignon* II, 153–181. Damascus: Institut Français de Damas.

Gaudefroy-Demombynes, Maurice and Régis Blachère. 1942. *Grammaire de l'arabe classique*. Paris: G.P. Maisonneuve.

Koschmieder, Edwin. 1929. *Zeitbezug und Sprache: Ein Beitrag zur Aspekt- und Tempusfrage*. Leipzig: B.G. Teubner.

Kuryłowicz, Jerzy. 1973. "Verbal aspect in Semitic". *Orientalia* 42.114–120.

McCarus, Ernest. 1976. "A semantic analysis of Arabic verbs". *Michigan studies in honor of George G. Cameron*, ed. Louis Orlin, 3–28. Ann Arbor: Department of Near Eastern Studies, University of Michigan.

Mitchell, Terence F. 1952. "The active participle in an Arabic dialect of Cyrenaica". *Bulletin of the School of Oriental and African Studies* 14.11–33.

Tonsi, Abbas al-. 1980. *Egyptian colloquial Arabic: A structure review*. Cairo: American University in Cairo.

Reckendorf, Hans. 1895, 1898. *Die syntaktischen Verhältnisse des Arabischen*. 2 vols. Leiden: E.J. Brill.

Reichenbach, Hans. 1947. *Elements of symbolic logic*. New York: Collier-Macmillan. (Repr. New York: Free Press, 1966.)

Vendler, Zeno. 1967. *Linguistics and philosophy*. Ithaca: Cornell University Press.

Wild, Stefan. 1964. "Die resultative Funktion des aktiven Partizips in den syrisch-palästinischen Dialekten des Arabischen". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 114.239–254.

Woidich, Manfred. 1975. "Zur Funktion des aktiven Partizips im Kairenischen-Arabischen". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 125.273–293.

Wright, William. 1874–1875. *A grammar of the Arabic language, translated from the Geman of Caspari, and edited with numerous additions and corrections, by W. Wright*. 2nd ed. London: F. Norgate.

JOHN C. EISELE (College of William and Mary)

Asseverative Particle

The proclitic particle *la-* was used in early literary Arabic as a marker of the category known to Jakobsonian structuralism as 'status', which has been defined as "the subjective evaluation of the narrated event by the speaker" (Aronson 1991:114). The principal role of *la-* seems to have been to underscore the speaker's commitment to the veracity of his or her utterance, and it was thus often encountered in oaths (e.g. *la-'in kašafta ʿannā r-rijza la-nu'minanna laka* '*la*-if you remove the penalty from us, *la*-we shall believe in you', Q 7/134), in evaluations or assessments (e.g. *wa-ma'wāhumu n-nāru wa-la-bi'sa l-maṣīru* 'and their abode is the Fire, and *la*-what an evil refuge it is!' (Q. 24/57), and in adversive or counterintuitive contexts (e.g. *la-'in basaṭta ʾilayya yadaka li-taqtulanī mā ʾana bi-bāsiṭin yadiya ʾilayka li-ʾaqtulaka* '*la*-even if you stretch forth your hand against me to kill me, I do not stretch forth my hand to kill you', Q. 5/28).

La- was known in the traditional grammatical literature under the cover term *lām at-ta'kīd* (or *at-tawkīd*) 'the *lām* of emphasis'. The grammarians distinguished several discrete types of emphatic *la-*, most of which consisted of further subcategories. The taxonomy provided in Ibn Hišām's (d. 761/1359) *Muġnī* (251ff.), for example, lists three principal types of *lām*, each of which in principle was associated with a specific set of syntactic environments.

i. The *lām* of the 'commencement' (*al-ibtidāʾ*) is typically located either at the head of a sentence (e.g. *la-yūsufu wa-'axūhu ʾaḥabbu ʾilā ʾabīnā minnā* '*la*-Joseph and his brother are dearer to our father than we are', Q. 12/8) or before the predicate (*xabar*) of a nominal sentence (*jumla ismiyya*) marked with the presentation particle *ʾinna* (e.g. *ʾinna rabbī la-samīʿu d-duʿāʾiʿ* '*inna* my Lord *la*-(is) the hearer of the call' Q. 14/39).
ii. The *lām* of the 'apodosis' (*al-jawāb*) marks the apodosis of either conditional structures (*wa-law ʾanna-hum ʾāmanū . . . la-maṯūbatun min ʿinda llāhi xayrun* 'and if they had believed . . . *la-* [their] reward before God would be better', Q. 2/103) or oath-formations. The latter include not only cases in which the oath itself is overtly expressed (of the type *wa-llāhi la yaxrujanna zaydun* 'by God, Zayd will surely go out!'), but also cases in which it must be assumed to be underlyingly present on the basis of the presence of certain lexical or morphological features felt to be characteristic of oaths (e.g. *la-yaxrujanna zaydun* '[By God!] Zayd will surely go forth', in which the verb is marked with the → energicus suffix *-anna*).
iii. The *lām* which 'announces' (*al-mu'ḏina*), or which 'paves the way' (*al-muwaṭṭi'a*), is prefixed to the conditional marker *ʾin* 'if', and is said to foreshadow that the following apodosis is governed by an oath (e.g. *la-'in ʾuxrijū lā yaxrujūna maʿahum* '*la*-if they are expelled, [by God!] they will not go out with them', Q. 59/12). The 'announcing' *lām* is often found paired with a following *la-*, the latter being regarded by the grammarians as *lām jawāb al-qasam* (cf. *wa-la-'in naṣarūhum la-yuwallūnna l-'adbāra* 'and *la-* (even) if they do come to their aid, [by God!] *la*-they will surely turn their backs [on them]', Q. 59/12).

The grammarians were aware that the above categories did not exhaust the entire range of attested examples of *la-*, and that, especially in the early poetic corpus, instances of *la-* were encountered which defied ready accommodation into any taxonomy. Ibn Hišām labeled the *la-* in such situations *al-lām az-zā'ida* 'supplementary *lām*'. An example of this may be seen in the following anonymous hemistich, in which the *la-* is uncharacteristically found in the predicate of a sentence headed by *lākinna* 'but': *wa-lākinnī min ḥubbihā la-ʿamīdū* 'but I, on account of loving her, *la-* [am] heartbroken'.

Another instance of the 'supplementary' *lām* which Ibn Hišām cites is found in Q. 22/13 *yadʿū la-man ḍarruhu ʾaqrabu min nafʿihi* 'they call *la-* one who is a likelier source of harm than of benefit', the syntactic complexity of which has given rise to various conflicting analyses.

For the most part, *la-* served as an optional emphasizer, and there was thus little appreciable distinction in meaning between a sentence such as *ʾinna zaydan la-karīmun* 'Zayd *la*-[is] noble' and the simpler *ʾinna zaydan karīmun*. In a few situations, however, the *la-* was said to be obligatory. In certain sentences featuring an initial *ʾin*, the presence of *la-* served to indicate that the *ʾin* was to be read as the 'lightened' (*muxaffafa*) shape of the particle *ʾinna* (e.g. *ʾin wajadnā ʾakṯarahum la-fāsiqīna* '*ʾin*-We found most of them [to be] *la*-corrupt', Q. 7/102) rather than as the synonymous negative particle (the *ʾin an-nāfiya*, as in *ʾin hāḏā ʾillā ʾifkun* 'this is nought but a fabrication', Q. 25/4) or as the conditional *ʾin*.

A second situation in which the presence of *la-* was said to be required was in object-clauses headed by *ʾinna* and preceded by a *verbum sentiendi* (*fiʿl al-qalb*) (e.g. Q. 63/1 *wa-llāhu yaʿlamu ʾinna-ka la-rasūlu-hu* ‘and God knows that you are His messenger’). In the absence of *la-*, the subordinating conjunction was required to take the shape *ʾanna* rather than *ʾinna* – note that Q. 63/1 may be paraphrased using . . . *yaʿlamu ʾanna-ka rasūluhu* – and indeed the conjunction *ʾanna* has largely supplanted *ʾinna* . . . *la-* . . . in all but the oldest stratum of Arabic.

As far as its position within the sentence is concerned, *la-* occurs as a rule either in the sentence-initial slot or at the beginning of the sentence’s last major syntactic constituent. For sentences marked with *ʾinna*, this constituent may be either the predicate – whether clausal (e.g. *ʾinna rabbaka la-yaḥkumu baynahum* ‘ʾ*inna* your Lord *la-* will judge between them’, Q. 16/124) or phrasal (*ʾinna-ka la-ʿalā xuluqin ʿaḏ̣īmin* ‘ʾ*inna-* you *la-* [are] on [have been formed with] a great character’, Q. 68/4) – or a substantival phrase (typically indefinite) which has been shifted to the right (cf. *ʾinna fī ḏālikum la-’āyātin li-qawmin yu’minūna* ‘ʾ*inna* in that *la-* [there are] signs for people who believe’, Q. 6/99), *ʾinna rabbahum bihim yawma’iḏin la-xabīrun* ‘ʾ*inna* their Lord regarding them on that day *la-* [is] fully informed’, Q. 100/11). The position of the *la-* found in *ʾinna-*sentences, known to the grammarians as *al-lām al-muzaḥliqa* ‘the *lām* which slides down’, was said to be the result of a systematic displacement which shifted it from an underlying sentence-initial position to medial position – i.e. an underlying **la-’inna zaydan karīmun* yielded, by virtue of a sort of ‘repulsion’ between the two emphasizing elements, the surface order *ʾinna zaydan la-karīmun* ‘ʾ*inna* Zayd *la-* [is] noble’; as support for this interpretation some grammarians cited the dialectal form *lahinna-* occasionally found in older poetry (e.g. the anonymous hemistich *lahinna-ka min barqin ʿalayya karīmun* ‘What a dear lightning-flash you are to me!’), which was interpreted as the counterpart to the underlying configuration **la-ʾinna-*. For the possible involvement of phonological factors in the distribution of the pre-Arabic ancestor of *la-*, see Testen (1998).

Many modern researchers have come to associate the *lām at-ta’kīd* with the ‘*lām* of the command’ (*lām al-’amr*), the clitic *l(i)-* added to the jussive verb-form to express wishes or commands (e.g. *li-yafʿal* ‘may he do’). Since it is clear that the simple jussive alone (*yafʿal*) was originally sufficient to render such structures, it is quite plausible that the fuller construction *li-yafʿal* could have resulted from a grammaticalization of *yafʿal* with an extra prefixed ‘emphasizer’. The vowel of the *lām al-’amr* is curious in that, although the normative language uses the general shape *li-*, the vowel is absent in the environment of a directly preceding conjunction (*wa-l-yafʿal, fa-l-yafʿal*), and that early dialects also attest *lām al-’amr* in the shape of *la-*.

Several other Semitic languages feature particles which contain an element *l* and which serve roughly the same ‘emphatic’ function as the Arabic *lām at-ta’kīd*. Geʿez, for example, employs a particle *la-* both in wish-formations (*la-yəkun bərhān* ‘let there be light’) and in conditional structures, the resulting *la-’əmma* being in the latter case highly reminiscent of Arabic *la-’in* (see above). Ugaritic seems to have used an emphasizing particle written *l-* (cf. *ʾugr l-rḥq ilm inbb l-rḥq ilnym* ‘[the city of] Ugr *l-* [is] far, O gods, Inbb *l-* [is] far, O divine ones’, *ʿnt* iv, 78-79), and isolated examples of what may be comparable cases have been adduced from Biblical Hebrew (cf. *kî lə-ḵéleḇ ḥay hû ṭôḇ min-hā-’aryeh ham-mēṯ* ‘because *lə-* a live dog is better than the dead lion’, Ecclesiastes 9:4). While many of the Northwest Semitic examples are open to alternative analyses, it is clear that Akkadian made use of a particle *lu*, both as an emphasizer (*lu aprus* ‘I did divide’) and as an element in wish/command constructions (*lu taprus* ‘may you divide’). In a number of Semitic languages we find formations corresponding to the Arabic *lām al-’amr* in which the *l*-particle has merged with the subject-prefix, e.g. Akkadian *luprus* (Babylonian dialect)/*laprus* (Assyrian dialect) ‘may I divide,’ *liprus* ‘may he divide,’ Mehri *lərkēz* ‘may I straighten,’ Socotri *ləʿárəb* ‘may I/he know,’ and Amharic *ləngär* ‘may I tell.’

In addition to Ibn Hišām’s *Muġnī*, *la-* is discussed in two surviving works of Arabic grammar, both of which are entitled *Kitāb al-lāmāt*, the first by az-Zajjājī (d. ca. 337/949) and the second by Ibn Fāris (d. 395/1004). For studies of contextual manifestations of literary Arabic *la-*, see Kinberg (1982, 1988), Nebes (1982, 1985, 1987), and Testen (1999). For discussion of the comparative Semitic dimension of *la-*, see Huehnergard (1983) and Testen (1998).

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Ibn Fāris, *Lāmāt* = ʾAbū l-Ḥusayn ʾAḥmad Ibn Fāris, *Kitāb al-lāmāt*. Ed. Gotthelf Bergsträsser, *Islamica* 1 (1924) 77–99.

Ibn Hišām, *Muġnī* = Jamāl ad-Dīn ʾAbū Muḥammad ʿAbdallāh ibn Yūsuf Ibn Hišām, *Muġnī l-labīb*. Ed. Māzin al-Mubārak and Muḥammad ʿAlī Ḥamdallāh. Damascus: Dār al-Fikr, 1964/1384.

Zajjājī, *Lāmāt* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān ibn ʾIsḥāq az-Zajjājī, *Kitāb al-lāmāt*. Ed. Māzin al-Mubārak. Damascus: al-Maṭbaʿa al-Hāšimiyya, 1969.

SECONDARY SOURCES

Aronson, Howard I. 1991. "Towards a typology of verbal categories". *New vistas in grammar: Invariance and variation*, ed. Linda R. Waugh and Stephen Rudy, 111–131. Amsterdam and Philadelphia: J. Benjamins.

Huehnergard, John. 1983. "Asseverative **la* and hypothetical **lu/law* in Semitic". *Journal of the American Oriental Society* 103.569–593.

Kinberg, Naphtali. 1982. "A study of *la-ʾin* clauses in early literary Arabic". *Jerusalem Studies in Arabic and Islam* 3.203–222.

——. 1988. "Some temporal, aspectual and modal features of the Arabic structure *la-qad* + prefix tense verb". *Journal of the American Oriental Society* 108.291–295.

Nebes, Norbert. 1982. "*ʾin al-muḫaffafa* und *al-lām al-fāriqa*, I". *Zeitschrift für Arabische Linguistik* 7.7–22.

——. 1985. "*ʾin al-muḫaffafa* und *al-lām al-fāriqa*, II". *Zeitschrift für Arabische Linguistik* 14.7–44.

——. 1987. "*ʾin al-muḫaffafa* und *al-lām al-fāriqa*, III". *Zeitschrift für Arabische Linguistik* 16.7–30.

Testen, David. 1998. *Parallels in Semitic linguistics: The development of Arabic* la- *and related Semitic particles*. Leiden: E.J. Brill.

——. 1999. "On *ʾinna, ʾanna* et alia." *Perspectives on Arabic linguistics*, XII, ed. Elabbas Benmamoun, 141–160. Amsterdam and Philadelphia: J. Benjamins.

DAVID TESTEN (Reston, VI, U.S.A.)

Assimilation

Assimilation can be viewed as a process in which the number of common segments of two (usually adjacent) sounds is made higher. In this process, one feature or set of features is dominant. The changes in the character of sounds do not change the meaning of the respective word or phrase. In Arabic, assimilatory processes of various kinds can be found: assimilation between consonants or between vowels and assimilation of a consonant to a vowel or vice versa. It may be progressive, regressive, or reciprocal, and also total or partial. It can happen with adjacent elements or at a distance, within a word or across the word boundaries. The usual position of the less dominant sound is in the syllable coda rather than in the syllable onset.

Assimilation occurs widely in Classical Arabic, Modern Standard Arabic, and Arabic dialects. In Classical and Modern Standard Arabic, assimilation within the root is not very frequent, but the root can become a domain of assimilation in the dialects. The most obvious examples are the assimilation of the definite article and the assimilation of the glottal stop /ʾ/ (*hamza*) and other weak consonants (especially the semivowels /w/, /y/). In the dialects of Arabic, assimilation processes are much more common.

1. TOTAL ASSIMILATION

In standard Arabic, the most common instances of this process are the assimilation of the *l* of the definite article and the assimilation of the glottal stop. The most common type of assimilation is regressive.

The *l* of the definite → article assimilates to the following consonant in case of coronals (dentals, sibilants, and liquids: *t, ṯ, d, ḏ, r, z, s, š, ṣ, ḍ, ṭ, ḏ̣, l, n*) – the so-called 'solar letters' (*ḥurūf šamsiyya*), as opposed to the 'lunar letters' (*ḥurūf qamariyya*), e.g. **al-sayyāratu* > *as-sayyāratu* 'car', etc. These two groups divide the phonemic inventory of Arabic consonants into two equal sets. In some dialects of Arabic, an assimilation of *l* to a velar plosive (*k* and *j/g*) can occur: **il-ginēna* > *ig-ginēna* 'garden', etc. A special case of this type of assimilation is the partial assimilation of *l* of the definite article to *m* under the influence of the following *b* in Modern Standard Arabic: *al-bāriḥata* > dialectal *embāreḥ* 'yesterday'. This expression can be found in many dialects, but the process as such (*l* > *m*) is exceptional and is limited to this example.

The reason for this assimilation can be explained either by violation of the → Obligatory Contour Principle (OCP), which forbids two adjacent coronals, or by historical reasons. The first argument does not correspond with the fact that the OCP acts mainly within the root and does not apply across morpheme boundaries; the second is dependent on the reconstructed form of the definite article in Semitic.

The *hamza* (glottal stop) assimilates very often. It can assimilate to a consonant or to a vowel (sometimes called 'compensatory lengthening’). The two processes can be modeled as follows:

i. i #V’C > #v̄C (’V’C > ’v̄C) (*’a’-mara* > *ʾāmara* ‘to ask advice, to consult’).
ii. vʾC > vCC (**iʾ-ta-xaḏa* > *it-ta-xaḏa* ‘to assume’).

A similar assimilation is also regularly attested with the semivowels (*w*, *y*): *iw-ta-ṣala* > *ittaṣala* ‘to connect’. Also the morphonological changes of *w* and *y* (**-awa-* > *-ā*; **-aya-* > *-ā*; **Cwa* > *Cā*; **Cwu* > *Cū*; **Cyi* > *Cī*, etc.) can be described as assimilation of the two consonants (*w* and *y*) to the neighboring vowels.

Aside from that, there are many instances of total assimilation (usually regressive) across the word boundaries. A common example is composite words consisting of a preposition (*min* ‘from’, *ʿan* ‘from, at, on’) and a relative pronoun *mā* ‘which’ (*min* + *mā* > *mimmā*).

In Classical Arabic, assimilation across word boundaries is very common. It is found in many places in the *Qurʾān*, and Medieval Arabic linguists frequently quote attestations of such assimilation in the dialect of many pre-Islamic Arab tribes.

A regressive assimilation of the final *-n* (mostly the indefinite article) to the following nasal or liquid is frequently attested in the *Qurʾān* (**hudan li-l-muttaqīna* > *hudal-li-l-muttaqīna* ‘a guide for the righteous’, Q. 2/2; **hudan min* > *hudam-min* ‘a path from’, Q. 2/4; *min rabbi-him* > *mir-rabbi-him* ‘from his master’, Q. 2/4, etc.).

Apart from the assimilation of *n*, a number of other instances of regressive assimilation are found in the *Qurʾān* or other medieval sources:

i. regressive spread of voice (e.g. *-td-* > *-dd-*: **ʿitdān* > *ʿiddān* ‘ready, prepared’);
ii. regressive spread of non-voice (e.g. *-b f-* > *-f f-*: *iḏhab fa-man tabiʿa-ka* > *iḏhaffa-man tabiʿa-ka* ‘go, with those who follow you’);
iii. regressive spread of nasalization (e.g. *-b m-* > *-m m-*: *uṭlub muḥammadan* > *utlum-muḥammadan* ‘ask Muḥammad’);
iv. regressive spread of spirantization (e.g. *-ṭ š-* > *-šš-*: *lā tuxāliṭ šarran* > *lā tuxāliššarran* ‘do not associate with evil’);
v. regressive spread of emphasis (e.g. *-d ṭ-* > *-ṭṭ-*: *ʾabʿid ṭāliban* > *ʾabʿiṭṭāliban* ‘chase Ṭālib away’).

For a detailed description see Fleisch (1961: 80–98). One might say that a similar situation can be observed also in the contemporary dialects of Arabic.

All of the above mentioned instances are also found in the Arabic dialects. There is, however, one additional process occurring in these:

i. assimilation of the *t-* detransitivizing prefix (Form V) to a dental or a sibilant (**t-dāxil* >*iddāxil* ‘to interfere’; *t-ṭarrab* > *iṭṭarrab* ‘to be covered’; *t-sābiʾ* > *issābiʾ* ~ *itsābiʾ* ‘to contend with’; *t-šaṭaf* > *iššaṭaf* ~ *itšaṭaf* ‘to be chipped’; *t-ṣabban* > *iṣṣabban* ~ *itṣabban* ‘to be soaped’); an assimilation to a velar plosive may occur as well (**t-kabb* > *ikkabb* ~ *itkabb* ‘to be poured’; **t-gawwiz* > *itgawwiz* ~ *itgawwiz* ‘to get married’) (cf. Watson 2002: 222–224).

2. PARTIAL ASSIMILATION

The most common features playing a role in assimilation are voice and emphasis. Both of them can spread or be neutralized and both types of assimilation are found in Classical Arabic, Modern Standard Arabic, and in all Arabic dialects.

Emphasis (→ *tafxīm*) spreads both regressively and progressively and its spread is so general that it has led some authors to regard emphasis as a prosodic feature (or, in autosegmental theory, to reserve a special tier for emphasis). The emphatic consonant influences its neighborhood (the minimal domain of emphasis is the syllable, but in many cases, especially in the Arabic dialects, its domain can be a whole word (cf., e.g., Hoberman 1989). The spread of emphasis (the usual direction being a left-to-right/progressive, though right-to-left/regressive is attested, too) can be viewed as partial assimilation. For a phonetic description see, e.g., Ghazeli (1981) or Giannini and Pettorino (1982). The spread of emphasis over the whole syllable or word can be observed in most Arabic dialects, with the exception of the peripheral ones, such as the dialects in Central Asia, Cyprus, Malta, or the African Arabic creoles. It is also present in Classical and Modern Standard Arabic.

A common place for partial assimilation in Classical and Modern Standard Arabic is the infix of Form VIII (*iK-ta-TaBa*). Here, the assimilation is progressive (the root consonant is dominant). The most common features are sharing of voice (**iZ-ta-ḤaMa* > *iZ-da-ḤaMa* ‘to be crowded’), emphasis (**iṢ-ta-DaMa* > *iṢ-ṭa-DaMa* ‘to collide’), etc.

In the Arabic dialects, a number of partial assimilation processes of voice are attested. Various rules can be observed, which may differ from one dialect to another. In short, it can be said that clusters of two consonants tend to agree in voice and that most of the assimilation is regressive (Malta: *niktbu* > *nigdbu* ‘they wrote’; Sudan: *yabsim* > *yapsim* ‘he smiles’, etc.). In some dialects, neutralization of voice in final position can take place (Daragözü, Turkey: *bərəd* > *bərət* ‘he got cold’; Malta: *trīd* > *trīt* ‘you/she want(s)’). However, sonorants do not neutralize and can also block neutralization of other sounds when in adjacent position. For details see Abu-Mansour (1996).

Concerning the problem of assimilation in the Arabic dialects, it should also be mentioned that some of the often adduced examples can be interpreted as assimilation only in opposition to Classical/Modern Standard Arabic and not from an internal reconstruction within the dialect itself. This is the case of Egyptian Arabic *zuġayyar* ‘small’ (< Classical/Modern Standard Arabic *ṣaġīrun*).

3. Mutual assimilation

A mutual assimilation (with two dominant features) is not very common, but there are examples of it. In this case, there are in fact two partial assimilations, both progressive and regressive, as in **iḏ-ta-xara* > *id-da-xara* ‘to keep, preserve’ (the *t* of the infix is assimilated to the voiced interdental, the interdental assimilated to the dental). In such cases, the fully assimilated form is allowed as well (*iḏ-ḏa-xara*).

4. Assimilation of vowels

The assimilation processes that involve vowels are usually called vowel harmony. The most common type is the progressive assimilation of *i/y* of the genitive or preposition to the suffixed pronoun, as a rule in the 3rd pers. sg. masc., pl. fem. and pl. (*-hu, -hum, -hunna*), less often in the 2nd pers. (*-kum* ‘your/you [pl.]’: *bi-kum* > *bi-kim* ‘in you [pl.]’, **ʿalay-kum* > *ʿalay-kim*; cf. Fleisch 1961:81–82). For the 3rd pers., the following forms are given:

* *fī bayti-hu* > **fī bayti-hi* ‘in his house’;
* *fī bayti-hum* > **fī bayti-him* ‘in their [masc.] house’;
* *fī bayti-hunna* > **fī bayti-hinna* ‘in their [fem.] house’;
* *ʿalay-hu* > *ʿalay-hi* ‘on him’;
* *ʿalay-hum* > *ʿalay-him* ‘on them [masc.]’;
* *ʿalay-hunna* > *ʿalay-hinna* ‘on them [fem.]’;

Another type of assimilation can be observed in the neighborhood of laryngeals, where the vowels exhibit a strong tendency to accommodate to the laryngeal sound. In most cases, it is /a/ that appears adjacent to laryngeal sounds.

Bibliographical references

Abu-Mansour, Mahasen Hasan. 1996. “Voice as a privative feature: Assimilation in Arabic”. *Perspectives on Arabic linguistics*, VIII, ed. Mushira Eid, 201–231. Amsterdam and Philadelphia: J. Benjamins.

Fleisch, Henri. 1961. *Traité de philologie arabe*. I. *Préliminaires, phonétique, morphologie nominale*. Beirut: Imprimerie Catholique.

Ghazeli, Salman. 1981. “La coarticulation de l’emphase en arabe”. *Arabica* 28.251–277.

Giannini, A. and M. Pettorino. 1982. *The emphatic consonants in Arabic: Speech laboratory report IV*. Naples: Istituto Universitario Orientale.

Hoberman, Robert. 1989. “Parameters of emphasis: Autosegmental analyses of pharyngealization in four languages”. *Journal of Afroasiatic Languages* 2.73–97.

Watson, Janet C.E. 2002. *The phonology and morphology of Arabic*. Oxford and New York: Oxford University Press.

Petr Zemánek
(Charles University)

Asyndetic Construction → Serial Verbs

ʿAṭf → Parataxis

Attitude → Language Attitude

Attrition → Language Attrition

Automatic Language Processing

Because of its morphological, syntactic, phonetic, and phonological properties, the Arabic language may be considered to be one of the most difficult languages for written and spoken language processing. Recent years have shown an increasing interest in Arabic in several fields of natural language processing (NLP). The aim of this entry is to present some of the most recent advances in Arabic language processing, such as morphology, information retrieval, text-to-speech synthesis, and speech recognition.

1. INTRODUCTION

Research on Arabic language processing started in the 1970s, even before the problems of Arabic text editing were completely solved. The first studies focused primarily on lexicon, morphology, and text-to-speech systems.

In the late 1990s, the internationalization of the → Internet and the proliferation of communication tools in Arabic led to the need for a large number of Arabic natural language processing applications. Statistics show that since 1995, when the first Arabic newspapers were launched online, the number of Arabic websites has been growing exponentially: by the year 2000, about 20,000 were counted (Abdelali 2004).

As a result, research activity has extended to address more general areas of Arabic language processing, including syntactic analysis, machine translation, document indexing, information retrieval, Arabic speech recognition and synthesis, speech translation and automatic identification of a speaker, geographic origin discrimination, etc.

2. TEXT PROCESSING

The Arabic alphabet consists of consonants, vowels, digits, and some other diacritics, as well as punctuation marks. Vowels are always preceded by consonants but are marked on the consonants. The particular form of the Arabic word facilitates the definition of the syllables, since there are only three possible forms for syllables: CV, CVC and CVCC, where V includes both short and long vowels.

Arabic glyphs change according to the position in the words: beginning, median, final, or isolated, every consonant having 4 (or fewer) different glyphs. There are 28 consonants and 6 vowels (short and long /i/, /a/, /u/). In addition to this, there are some other diacritics, such as the gemination markers that are used (like the vowel signs) as diacritics on consonants to mark them as geminated. In addition, there are the markers of the *tanwīn* (pronounced /an/, /un/ and /in/), used like the vowels on the consonants. Another marker is the pause or absence of vowels which is marked by the *sukūn* diacritic.

Arabic numbers are represented exactly like those in French and English, but the pronunciation of a number changes according to its grammatical context. The number seven, for instance, has nine possible pronunciations (/sabʿu/, /sabʿi/, /sabʿa/, /sabʿatu/, /sabʿati/, /sabʿata/, /sabʿatun/, /sabʿatan/, /sabʿatin/).

Arabic texts are read and written from right to left, and the vowels are generally not indicated, which means that there are two possible representations, voweled and non-voweled. While reading a non-voweled text, an Arabic speaker implicitly assigns the appropriate vowel to the consonant. Such an operation is very complex, however, to process automatically. The voweling of a word depends not only on its syntactic context, but also on the semantics of the whole sentence. The verb form <ktbt>, for example, can refer to four possible persons: the 1st person singular /katabtu/, the 2nd person singular masculine /katabta/, the 2nd person singular feminine /katabti/ and the 3rd person singular feminine /katabat/. It is impossible to pronounce such a word correctly without context, so for most written texts one must understand the text in order to know how to vowel and pronounce it.

Arabic morphology represents a special kind of morphological system, characterized by the manipulation of two essential factors → root and scheme (Soudi a.o. 2001). The use of these two factors makes the majority of morphological rules perfectly regular. Thus, a morphological parser constitutes the most important component of an Arabic Language Processing system, which is why many researchers have worked in this field (Fassi Fehri 1982; Sadiqi and Ennaji 1992; Jarir 1997; Soudi a.o. 2001; Abuleil and Evens 2004).

Designing a high-quality morphological parser requires both an exhaustive database and a well-adapted organization of this database. Here, the methodology that is presently being used to obtain an optimal morphological database and develop a well-adapted morphological parser for Arabic is presented (Chenfour 2003).

The database is designed with object-oriented modeling. Every morphological primitive is regarded as an object, and objects having the same morphological properties are gathered in the same class. According to the category of the morphological components that each class represents, all classes are arranged in several packages such as verbs, nouns, particles, and affixes.

The linguistic database is implemented using a well-adapted language that has been developed specially for this work: Java-based MOrphology Definition Language (JMODEL; Chenfour 2003). All objects and classes are transcribed to a deterministic finite state automaton, representing the morphological parser itself.

As a first result, this organization made it possible to build a complete morphological database for Arabic with a small number of components. Since most morphological classes are interconnected, either by inheritance relationship or by aggregation, the number of morphological construction rules could be reduced significantly.

Every morphological primitive is treated as an object, called a morphological component (MCM). Every MCM is characterized by a list of morphological properties, called morphological descriptors (MD). The MCMs are gathered into different morphological classes (MCL) according to the following rule: "The components of the same class accept the same prefixes and the same suffixes or they are suffixes or prefixes of the same classes." For example,

```
class OriginSchemeS uses Number {
        faʿala(1) { singular }
        faʿila(2) { singular }
        faʿula(3) { singular }
        . . .
}
```

A morphological class is called 'abstract' when its components are not complete words, but need to be concatenated to some other components. *Final* classes are classes that contain only complete Arabic words.

Morphological descriptors are also gathered into different classes, called morphological properties classes (MPC), e.g. the MPC 'Number' is a class containing all morphological descriptors indicating the property numbers. It is modeled using JMODEL language (Chenfour 2003).

```
property class Number {
singular;
dual;
plural;
}
```

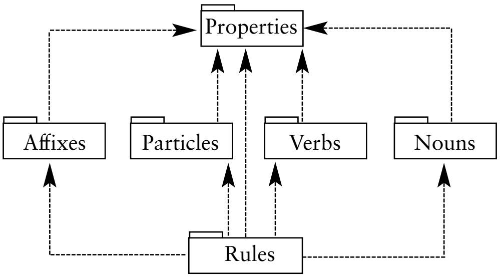
Rules classes (RCL) are defined symbolizing all possible concatenations between the different morphological components defined in the morphological classes. Thus, each class of rules represents a complete Arabic word family.

Finally, the MCL, MPC, and RCL classes are also arranged in different packages according to the kind of morphological components or morphological descriptors that each one contains. There are six main packages, four of which contain all morphological classes, affixes package, particles package, verbs package, and nouns package. These depend on the properties package, containing all morphological properties classes. The last part of the database is the rules package, which contains all rule classes (see Figure 1). This package depends on all the other packages, and every component of this package represents a complete Arabic word.

A complete description of all packages is too long to be included here (Chenfour 2003; Tahir and Chenfour 2004). Therefore, in order to give an idea of the constitution of every package, the Verbs package, which is the most important one of the database, is used as an example.

The morphological system of Arabic verbs is very particular. On the one hand, it is robust and

Figure 1. Morphological database packages



completely regular in the case of sound verbs, using the representation 'root–scheme'. On the other hand, it is generally irregular in the case of the → weak and → defective verbs (a family of verbs being characterized by their particular structure and incomplete conjugation).

A fundamental study of this system makes it possible to establish the representation 'radical–conjugation affixes', which regularize all conjugation rules even in the case of the weak verbs (Jarir 1997). This representation makes it possible to divide the conjugated form of any Arabic verb into two parts: the radical part, which depends just on tense and the infinitive form of the scheme, and the conjugation affixes, which depend on person, number, gender, and tense. It is always represented by a prefix in the case of the perfect and by a pair suffix-prefix in the case of the imperfect.

According to this particular structure of Arabic verbs, the verbs package gathers all verbs' radicals classified in other sub-packages according to the nature and conjugation tense of each kind of verb.

Another important component is the rules package. It gathers all possible concatenation rules that can be applied to all morphological components. Each concatenation rule can generate a new word from a set of morphological components (Figure 2).

Finite-state automata are widely used in language and speech processing (Elgot and Mezei 1965; Kaplan and Kay 1981; Kay 1987; Mohri 1997; Kiraz and Grimley-Evans 1998; Beesley and Karttunen, 2003). They permit a fast processing of input strings and can be easily modified and combined by well-defined operations. A JMODEL compiler is being written that will translate the whole morphological database to a deterministic finite-state automaton, called Deterministic Morphological Automaton (DMA). With this result it will be possible to have a 100 percent rule-based morphological parser, entirely based on automata. Some experiments have already been carried out on many morphological classes and an adequate DMA with very satisfying results has been obtained.

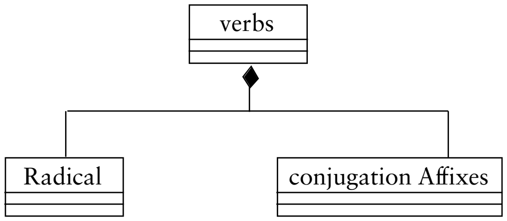
One of the most important fields in Arabic language processing is information retrieval, where a query that is formulated by a user is matched with objects of any media in a database. Arabic information retrieval has become a focus of research and commercial development due to the vital necessity of such tools. But all Arabic texts available online (in newspapers, for example) are non-voweled, which means that the text becomes ambiguous and difficult to process with an Arabic information retrieval system.

Therefore, a deep level of morphological analysis is needed for information retrieval of Arabic. There are two schools of thought about this: according to one, a light stemmer would provide sufficient results, whereas the other school believes that light stemming will stand a greater chance of producing wrong stems. Yet, it is evident that Arabic information retrieval has a particularly acute need for effective normalization and stemming. Both orthography and morphology give rise to a huge amount of lexical form variation. For information retrieval, this abundance of forms means a greater likelihood of mismatch between the form of a word in a query and the forms found in documents relevant to the query. This is what explains the lack of Arabic information retrieval systems. However, a few Arabic information retrieval systems are available such as <http://crl.nmsu.edu/~ahmed/arabic/> or Aramedia (<http://aramedia.com/>), and also some publications (Abdelali 2004; Al-Onaizan a.o. 2003; Larkey and Connell 2002).

3. Speech processing

Research on Arabic speech processing has made significant progress, due to improved signal processing technologies, and to recent advances in

Figure 2. The representation (radical–conjugation affixes) of Arabic verbs



the knowledge of the prosodic and segmental characteristics of Arabic and the acoustic modeling of Arabic schemes. These results should make it possible to progress further in more innovative areas, such as Arabic speech synthesis and recognition.

3.1 *Speech synthesis*

Most of the existing speech synthesis systems can be classified as either formant synthesizers (Klatt 1980, 1987) or concatenative synthesizers (Atal and Schroeder 1967; Chenfour 1997). Formant synthesizers, which are rule-based, have the advantage of a very small database, but the synthesized speech is not very natural. On the other hand, trainable concatenative speech synthesis, using a large speech database, has become popular due to its usual ability to produce a high quality natural speech output. Concatenative synthesis is based on speech signal processing of natural speech databases. The segmental database is built to reflect the major phonological features of a language. For instance, its set of phonemes is described in terms of diphone units, representing the phoneme-to-phoneme junctures. Non-uniform units are also used (diphones, syllables, words, etc.). The synthesizer concatenates speech segments, and performs some signal processing to smooth unit transitions and to match predefined prosodic schemes.

However, achieving high quality text-to-speech synthesis with trainable concatenative methods requires a safe choice of the synthesis method and the speech unit and a conscientious preparation of the speech units database. Many methods and systems have been developed in this field, such as ARABTALK or BrightSpeech. Reference may also be made to the work done by Amr Youssef and Ossama Emam (Youssef and Emam 2004). Their system has been developed at the Human Language Technologies Laboratory of IBM Egypt and is based on the state-of-the-art IBM trainable concatenative speech synthesizer.

There are also other tools, such as Sakhr products, which make it possible to scan the text of an Arabic newspaper and transform it to text, passing the generated text to the automatic corrector to make sure that all known Arabic mistakes are corrected. Afterwards, Sakhr's keyword extractor can identify important concepts in the text and classify the content under one or more categories in a user-defined taxonomy, generate a summary of the article, translate it into English or vowel it, and finally read it with a text-to-speech synthesizer.

In the following sections the general architecture of the PARADIS (Psola & ARAbic DI-Syllable) system is examined; this is based on the concatenation of di-syllables and TD-PSOLA as synthesis method developed by Chenfour and others.

3.1.1 Architecture of the PARADIS system

The text-to-speech system (PARADIS) is composed of two major parts, namely the linguistic processing part and the acoustic processing one. The goal of the first part is to transcribe the input text to a phonetic representation equipped with prosodic markers. The second part consists of translating this output phonetic text to a vocal message of a good quality. The two parts of the system are composed of six independent modules communicating with each other (Figure 3).

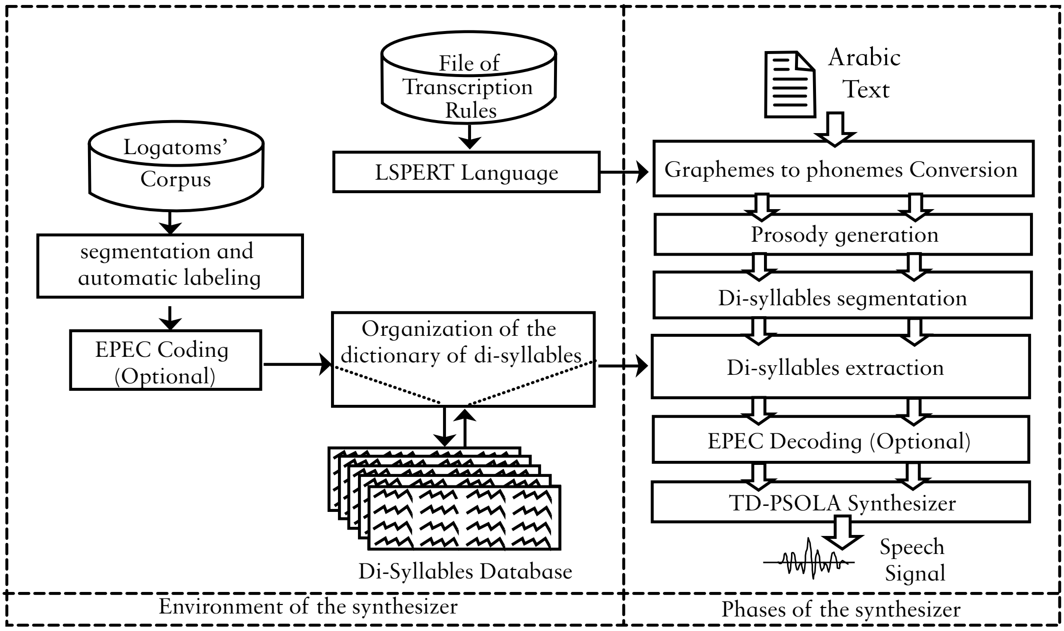
3.1.2 Grapheme to phoneme conversion module

The first module of PARADIS is grapheme-to-phoneme conversion that consists of transcribing an Arabic input text to a corresponding phonetic text. The transcription module was automatically generated by a rules compiler named LSPERT (Language of Specification of the Rules of Transcription) which was developed for this purpose (Chenfour 1997). Input data for LSPERT is a formal specification of all transcription rules (about 150 rules). The syntax used to describe rules is inspired by Chomsky's formalism commonly used in linguistics (Chomsky and Halle 1968). Six classes of rules have been developed: direct conversions (one-to-one mapping); context depending conversions; rules for irregular words; mathematical symbols; abbreviations; and numbers.

3.1.3 Prosody generation

Prosody generation constitutes a very important component of the system. Its aim is to insert prosodic indicators in the phonetic text generated at the previous stage. These indicators determine syllable duration, position and duration of the pauses, stressed syllables, evolution of the melody, etc. The generated text is then used by the synthesis module to ensure the pro-

Figure 3. General architecture of the PARADIS system


File of Transcription Rules
Arabic Text
Logatoms' Corpus
LSPERT Language
Graphemes to phonemes Conversion
segmentation and automatic labeling
Prosody generation
Di-syllables segmentation
EPEC Coding (Optional)
Organization of the dictionary of di-syllables
Di-syllables extraction
EPEC Decoding (Optional)
TD-PSOLA Synthesizer
Speech Signal
Di-Syllables Database
Environment of the synthesizer
Phases of the synthesizer

duction of a high-quality vocal message corresponding to the input Arabic text.

The prosody generation component receives a phonetic text equipped with punctuation marks and has to transform it to a flow of parameters in order to control the synthesizer. The parameters to be computed include three prosodic factors: pitch factor (PF), duration factor (DF), and intensity factor (IF). Therefore, a melodic model and a duration model are needed to implement the prosodic structure. The prosodic module contains five main components: pause generation, stress marker, duration and modality, syllabic lengthening, and finally, the most important component, the melodic model, whose aim is to compute pitch or fundamental frequency curves (see Fig. 4).

3.1.4 Concatenation of di-syllables

The third phase consists of segmenting phonetic text in small units. One of the major problems encountered in concatenative text-to-speech systems is how to make the best selection of units and how to describe their concatenation. Indeed, concatenating units usually creates a problem of distortion because of spectral discontinuity at the connecting points.

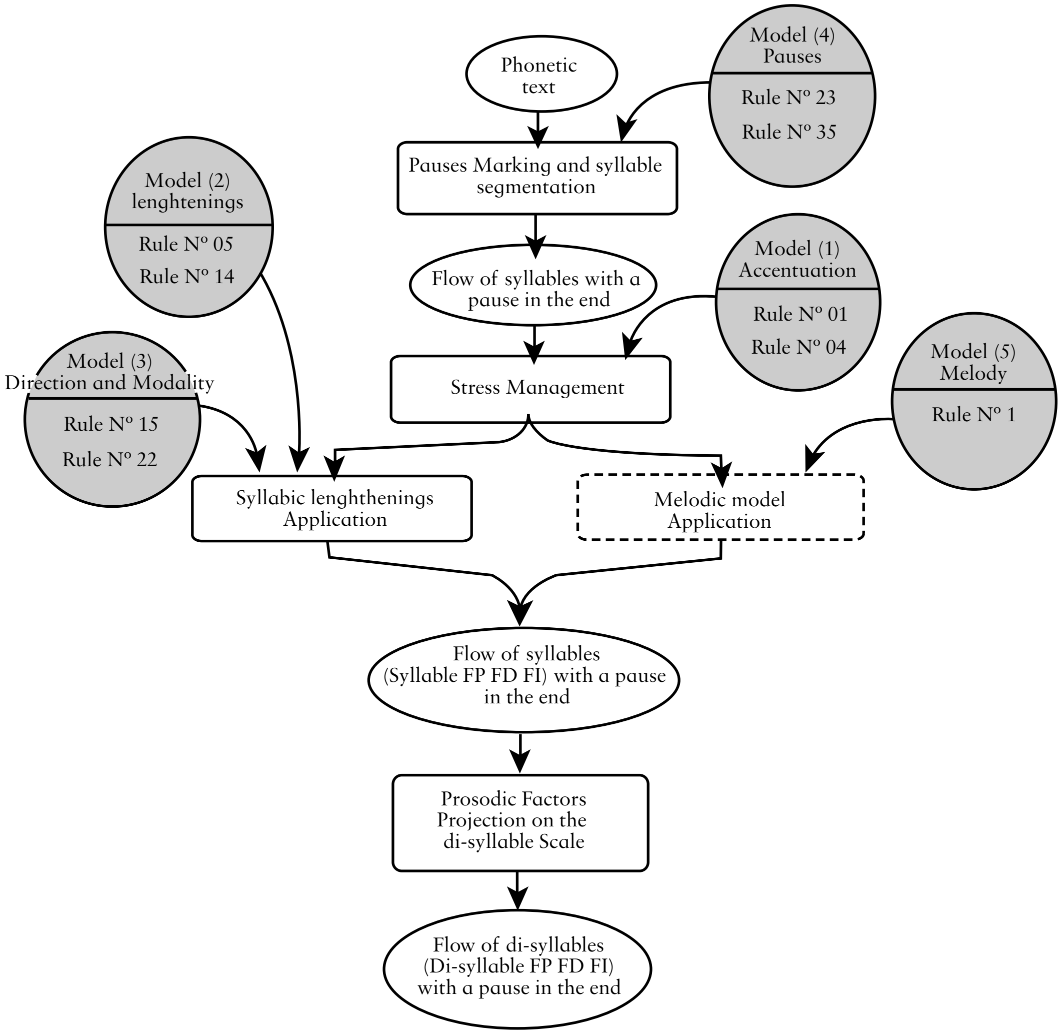
The use of di-syllables (part of signal from vowel to vowel) as concatenation units solved most of this problem, and the generated speech is much smoother at the concatenation points, which are always vowels (Chenfour a.o. 1997). There are only six forms of di-syllables in Arabic: CV at the beginning of words; VC, VCC, and V at the end; and VCV and VCCV in the middle.

By generating automatically long vowels from short ones, and by excluding some impossible shapes it turned out to be possible to decrease the combinatory of di-syllables to about 8,500 units. All these units were recorded, then segmented and labeled automatically. The database of di-syllables was generated afterwards automatically, using a hash-coding method. In this way, the time needed to look up an entry in the database at the synthesis stage was reduced significantly.

3.1.5 EPEC decoding

In order to decrease database size, a speech signal coder had to be used. The most popular coding algorithm is a code-excited linear prediction (CELP) coder. However, it requires digital signal processing (DSP) to be implemented in real time. The speech coding algorithm EPEC (Extensible Pitch period Extraction Coding) has a very low complexity of computation, high reconstruction quality, but reasonable bit-rate compression. It is based on a long-term prediction procedure, which represents a voiced short signal

Figure 4. Prosodic module


Phonetic text
Model (4) Pauses
Rule N° 23
Rule N° 35
Pauses Marking and syllable segmentation
Model (2) lenghtenings
Rule N° 05
Rule N° 14
Flow of syllables with a pause in the end
Model (1) Accentuation
Rule N° 01
Rule N° 04
Model (3) Direction and Modality
Rule N° 15
Rule N° 22
Stress Management
Model (5) Melody
Rule N° 1
Syllabic lenghthenings Application
Melodic model Application
Flow of syllables (Syllable FP FD FI) with a pause in the end
Prosodic Factors Projection on the di-syllable Scale
Flow of di-syllables (Di-syllable FP FD FI) with a pause in the end

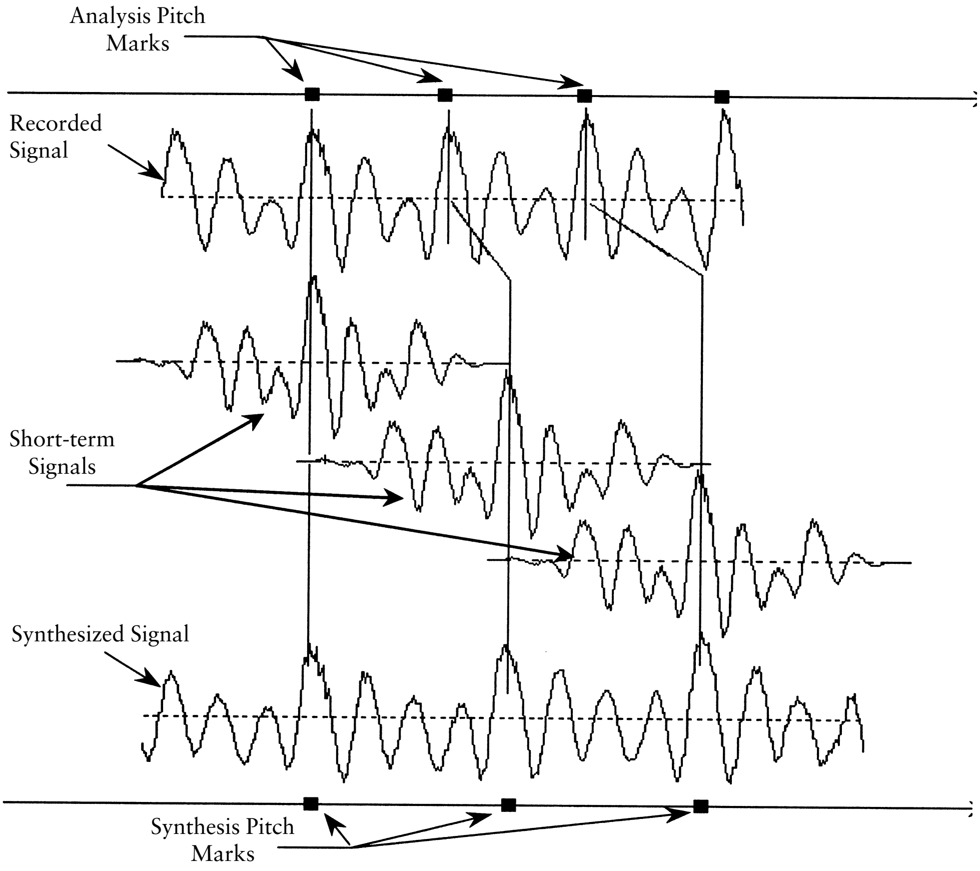
portion (a sequence of 4 or 5 pitch periods) with one chosen extensible pitch period and with the use of the Time-Domain Pitch-Synchronous OverLap-Add (TD-PSOLA) algorithm to overlap short signals and to smooth period transitions (Chenfour a.o. 2000).

3.1.6 TD0-PSOLA synthesizer

The last step for speech synthesis and the most fundamental component of the text-to-speech system is the TD-PSOLA synthesizer that was implemented and adapted for the entire Arabic system. This module must generate synthetic speech according to the segmental and prosodic parameters defined at earlier stages of processing (Fig. 5).

The TD-PSOLA synthesizer is known by its capacity for direct action on the speech signal and the concept of separation between the coding algorithm and the synthesis technique. The basic idea of TD-PSOLA consists of three steps (Mouline and Charpentier 1990): the analysis step where the original speech signal is first divided into separate but overlapping short-term analysis signals; the modification step of each analysis signal to synthesis signal; and the synthesis step, where these segments are recombined by means of overlap adding. TD-PSOLA has significantly improved the synthetic speech quality, as it allows, with great simplicity, the variation of the duration and the fundamental frequency of the synthesized speech signal.

Figure 5. Principle of TD-PSOLA synthesizer



However, it was necessary to determine pitch marks throughout segments. This was achieved automatically, with a labeling system which was developed for this purpose. In this way, a di-syllables database (with adequate pitch marks) was obtained which constitutes the necessary static component of the PARADIS system.

3.2 *Speech recognition*

The speech recognition task, namely the recognition of single words spoken by arbitrary persons, was solved very successfully for Arabic by means of statistical models such as Hidden Markov Models (HMM), where every word has to be recorded from many speakers. The disadvantage of such a speech recognizer is the limitation of its vocabulary. Extending the vocabulary to new words requires collecting the corresponding speech signals, which is a very time-consuming and therefore expensive task.

Significant advances in Arabic speech recognition have been achieved recently by developing speaker-independent word recognizers with unlimited vocabulary. The basic idea is to define and train appropriate subword unit models, such as phone models, rather than whole word models. With the resulting set of subword unit models it is then possible to generate a word model for any desired word by the concatenation of the appropriate subword unit models.

The focus of research then shifted to the transcription of speech such as radio and television broadcast news (Bakis a.o. 1997), and some work on Arabic has become a reality (Zavaliagkos a.o. 1998; Kirchhoff a.o. 2003). Transcription of broadcast news presents technical challenges to Large Vocabulary Continuous Speech Recognition (LVCSR) systems (IBM). Indeed, broadcast news contains speech and non-speech data, and the speech data in broadcast news contains a wide mixture of speaking styles. In addition to that, broadcast news comes from several sources, such as interviews with people under noisy conditions and interviews over the telephone, so that the background may contain music, noise, or other interfering speech.

Recognizing broadcast news data with a system trained on a clean training corpus gives very high error rates. Therefore, the basic task of the speech recognizer is to first try to identify the segments of input data that can belong to several classes and use separate modeling techniques appropriate for each class. For instance, segments detected as music are simply removed, segments identified as telephone quality speech are decoded by a system trained on telephone bandwidth speech, and so on. A typical broadcast news transcription system is thus composed of two main modules:

i. Segmentation and labeling module which is applied to the input data at the first stage, in order to extract and classify segments with the same nature (e.g., speech, noise, etc.). These segments are associated with suitable indicators and all non-speech segments are removed. The result of the segmentation and labeling module is a flow of speech segments with cluster, gender, and wideband or telephone quality indicators.
ii. The second module is the word recognizer which determines for each speech segment the sequence of words that compose it. The speech recognizer makes use of the Hidden Markov Models of the words to be recognized.

Two main classes of Arabic resources must be prepared in order to achieve an Arabic speech recognizer:

i. A speech corpus containing a few dozen hours of recorded audio data from radio and television broadcast news (from different Arabic stations such as Aljazeera or the Moroccan TV RTM).
ii. A text corpus of a few hundreds of millions of words from various newspapers:

Algeria	<http://www.aps.dz/>
Jordan	<http://www.petra.gov.jo/
Kuwait	<http://www. alraialaam.com/>
Lebanon	<http://www.nna-leb.gov.lb/>
London	<http://www.asharqalawsat.com/>
Mauritania	<http://www.rajoul-echaree.com/>
Morocco	<http://www.morocco-today.com/>
Oman	<http://www.alwatan.com/>
Saudi Arabia	<http://www.al-jazirah.com/>
Syria	<http://www.thawra.com/>
Qatar	<http://www.aljazeera.net/>
Qatar	<http://www.raya.com/>

The text obtained usually needs to be corrected with the help of a morphological parser or other automatic text processing systems.

One of the major limitations for developing Arabic speech recognizers is the lack of adequate resources. Only a few standard corpora are available online to users:

i. Agence France Presse (<http://wave.ldc.upenn. edu/Catalog/LDC2001T55>), a collection of 380,000 newswire stories from 1994–2000 available from the Linguistic Data Consortium.
ii. Al-Hayat (<http://www.elda.fr/catalogue/ en/text/W0030.html>), a collection of over 42,000 newspaper stories from 1994 available from the Evaluations and Language resources Distribution Agency (ELDA).

Other resources such as bilingual dictionaries are needed. These can vary from machine translation dictionaries to handcrafted dictionaries for a specific topic or usage (Abdelali a.o. 2004). Available online are the Ajeeb Arabic–English dictionary (<http://dictionary.ajeeb.com/en. htm>) from Sakhr (<http://www.sakhr.com/>) and the Ectaco Bilingual Dictionary (<http://www.ectaco.com/online/diction.php3?refid=0&rfr_id=10370&-rqt_id=81368 &pagelang=23 &lang=3>), which were used in some information retrieval experiments (Larkey a.o. 2002). Other individual efforts were carried out in different applications (Zajac a.o. 2001).

Apparently, then, the availability of adequate Arabic Language Resources (ALR) is recognized as a central component of the linguistic infrastructure, necessary for the development of all Arabic Human Language Technologies (HLT). This issue is now becoming prominent: the lack, on the one hand, of resources and, on the other hand, of real-world applications, highlights the need for improving research in this area and for promoting the use of human language technologies among potential partners, in particular to safeguard some of the cultural heritage of the Arabic-speaking regions.

This is why a new project has been developed, called NEMLAR (Network for Euro-Mediterranean LAnguage Resources: <http://www.nemlar.org>). The NEMLAR consortium consists of 14 partners from European and Arabic countries. NEMLAR's goal is to obtain the necessary funds to produce the required resources and tools, and to make them as widely available as for many other major languages. The European Language Resources Association (ELRA) and the European Network of Excellence in Human Language Technologies (ELSNET) have been promoting the concept of a Basic LAnguage Resources Kit (BLARK), which constitutes a must for each and all languages to allow automatic processing of the language. Hence, one of the goals of the NEMLAR project is to collect information about existing institutions and language resources, and to describe the need for language resources.

4. Conclusion

The aim of this entry was to represent the state of the art of most important aspects of Arabic language processing. Recent contributions to scientific and technical progress in Arabic computer applications have been reviewed, which allow natural speech interaction between human end-users, computer applications, and digital information services. Arabic is at least one order of magnitude more difficult than other common languages due to its particular morphology, the lack of vowels in written texts, and other particularities. An attempt has been made to determine which technologies are suitable for Arabic language processing, whether training data will be necessary, and, if so, what kind of training data will be adequate.

Obviously, it is impossible to review all requirements of all computer applications for Arabic. Thus, for instance, the issue of the different varieties of Arabic language and dialects, the interaction between Arabic and other languages, etc. have not been discussed here. However, the entry offered an opportunity to discuss the most recent advances in both Arabic text and speech processing technologies in different contexts. (See also → computational linguistics, Corpus linguistics)

Bibliographical references

Abdelali, Ahmed, Jim Cowie, and Hamdy Soliman. 2004. "Arabic information retrieval perspectives". *Journée d'étude sur la parole: Traitement automatique du langage naturel*, 191–196. Fes: Sidi Mohamed ben Abdellah University.

Abuleil, Saleem and Martha Evens. 2004. "Arabic noun classifier system (ANCS) to support noun morphology". *Journée d'étude sur la parole: Traitement automatique du langage naturel*, 203–208. Fes: Sidi Mohamed ben Abdellah University.

Al-Onaizan, Y., R. Florian, M. Franz, H. Hassan, Y.S. Lee, S. McCarley, K. Papineni, S. Roukos, J. Sorensen, C. Tillmann, T. Ward, and F. Xia. 2003. "TIPS: A translingual information processing system". *Proceedings of the Human Language Technology conference/North American Chapter of the Association for Computational Linguistics annual meeting*, 1–2. Edmonton.

Atal, B.-S. and M.R. Schroeder. 1967. "Predictive coding of speech signals and subjective error criteria". *Proceedings of the conference on communication and processing*, 360–361. Cambridge, Mass.

Bakis, Raimo, Chen Scott, Ponani Gopalakrishnan, Ramesh Gopinath, Stéphane Maes, and Lazaros Polymenakos. 1997. "Transcription of broadcast news: System robustness issues and adaptation techniques". *Proceedings of the IEEE international conference on acoustics, speech and signal processing*, 711–714. Munich.

Beesley, Kenneth R. and Lauri Karttunen. 2003. *Finite state morphology*. Palo Alto, Calif.: Center for the Study of Language and Information Publications.

Chenfour, Noureddine. 1997. *Réalisation d'un système de synthèse de la parole arabe à partir du texte par concaténation de di-syllabes*. Sc.D. diss., Faculty of Sciences, Mohamed V University, Rabat.

——. 2003. "Réalisation d'un langage de définition de la morphologie JMODEL". *Journée d'étude en réseau et multimédia*, 90–102. Fes: Sidi Mohamed Ben Abdellah University.

——, Abderrahim Benabbou, and Abdelhak Mouradi. 1997. "Synthèse de la parole arabe par concaténation de di-syllabes". *Journées scientifiques et techniques du réseau francophone de l'ingénierie de la langue de l'AUPELF-UREF*, 459–462. Avignon.

——. 2000. "Synthèse de la parole arabe TD-PSOLA, génération et codage automatiques du dictionnaire". *Proceedings of the international symposium on image/video communication*, 112–122. Rabat: Mohamed V University.

Chomsky, Noam and Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row. (Repr., Cambridge, Mass.: MIT Press, 1991.)

Elgot, C.C. and Jorge E. Mezei. 1965. "On relations defined by generalized finite automata". *IBM Journal of Research and Development* 9.47–68.

Fassi Fehri, Abdelkader. 1982. *Linguistique arabe: Forme et interprétation*. Ph.D. diss., University Mohamed V, Rabat.

Jarir, Zahi. 1997. *Mise en place de primitives morphologiques dans un environnement 4GL et application à une base de données terminologiques*. Sc.D. diss, Faculty of Sciences, Mohamed V University, Rabat.

Kaplan, R.M. and M. Kay. 1981. "Phonological rules and finite state transducers". *Linguistic Society of America handbook: Fifty-sixth annual meeting*. New York. [Paper presented to the First Mathematics of Language Conference, University of Michigan. 27–30 December 1981.]

Kay, M. 1987. "Non concatenative finite-state morphology". *Proceedings of the 3rd international conference of the European Chapter of the Association for Computational Linguistics*, 2–10. Copenhagen: University of Copenhagen.

Kiraz, G.A. and E. Grimley-Evans. 1998. "Multi-tape automata for speech and language systems: A prolog implementation". *Automata implementation*, ed. D. Woods and S. Yu. (= *Lecture Notes in Computer Science*, no. 1436.), 87–103. Berlin: Springer Verlag.

Kirchhoff, K., J. Bilmes, S. Das, N. Duta, M. Egan, G. Ji, F. He, J. Henderson, D. Liu, M. Noamany, P. Schone, R. Schwartz, and D. Vergyri. 2003. "Novel approaches to Arabic speech recognition: Report from the 2002 Johns Hopkins Workshop". *Proceedings of the IEEE international conference on acoustics, speech and signal processing*, 1.344–347. Hong Kong.

Klatt, D.H. 1980. "Software for a cascade/parallel Formant synthesizer". *Journal of the Acoustical Society of America* 80.971–995.

——. 1987. "Review of text-to-speech conversion for English". *Journal of the Acoustical Society of America* 87.737–793.

Larkey, Leah S. and Margaret Connell. 2002. "Arabic information retrieval at UMass in TREC-10". *Tenth text retrieval conference*, 562–570. Washington, D.C.: USGPO.

Mohri, M. 1997. "Finite-state transducers in language and speech processing". *Computational Linguistics* 23:2.269–312.

Mouline, E. and F.J. Charpentier. 1990. "Pitch synchronous waveform processing techniques for text-to-speech synthesis using diphones". *Speech Communication* 9:5–6.453–467.

Sadiqi, Fatima and Moha Ennaji. 1992. *Introduction to modern linguistics*. Casablanca: Afrique Orient.

Soudi, Abdelhadi, Abderrahim Jamari, and Violetta Cavali-Sforza. 2001. "A computational lexeme-based treatment of Arabic morphology". *Proceedings of the Arabic NPL workshop at the Association for Computational Linguistics/European Chapter of the Association for Computational Linguistics*. Toulouse.

Tahir, Youssef, Noureddine Chenfour, and Mostafa Harti. 2004. "Modélisation à objets d'une base de données morphologique pour la langue arabe". *Journée d'étude sur la parole: Traitement automatique du langage naturel*, 351–360. Fes: Sidi Mohamed ben Abdellah University.

Youssef, Amr and Ossama Emam. 2004. "An Arabic TTS system based on the IBM trainable speech synthesizer". *Journée d'étude sur la parole: Traitement automatique du langage naturel*, 375–384. Fes: Sidi Mohamed ben Abdellah University.

Zajac, Rémi, Ahmed Malki, Ahmed Abdelali, James Cowie, and William C. Ogden. 2001. "Arabic-English NLP at CRL". *Proceedings of the Arabic NLP workshop at the Association for Computational Linguistics/European Chapter of the Association for Computational Linguistics*. Toulouse.

Zavaliagkos, G., J. McDonough, D. Miller, A. El-Jaroudi, J. Billa, F. Richardson, K. Ma, M. Siu, and H. Gish. 1998. "The BBN Byblos 1997 large vocabulary conversational speech recognition system". *Proceedings of the international conference on acoustics speech, and signal processing, Seattle, Washington*, II, 905–908. Washington, D.C.

Noureddine Chenfour (Dhar El-Mehraz, Fes)

Autosegmental Phonology → Morphology

Auxiliary

The term 'auxiliary' is used as a descriptive label for forms bearing both morphological and positional resemblance to verbs. They occur in combination with a main verb to express tense, aspect, mood, voice, or polarity. Verbal particles with a similar function may also be called 'auxiliaries' (cf. Schachter 1985:41; Steele 1999: 49; Payne 1997:84–85). Both Literary and Colloquial Arabic have elements which fit this rather broad descriptive label. The form of these elements ranges from fully inflected verbs to particles, which may or may not have inflectional properties.

Only verbal auxiliaries are treated here, with special attention for the verb *kāna* 'to be', which can be identified as the most prototypical auxiliary verb in all varieties of the language in view of its functions and distribution. Nonverbal auxiliary devices differ considerably across the various colloquial dialects and are described separately in the dialect sketches.

1. VERBS WITH AUXILIARY FUNCTIONS

Arab grammarians traditionally classify verbs with auxiliary functions in a group, referred to as *al-ʾafʿāl an-nāqiṣa* 'the incomplete verbs'. Apart from the verb *kāna* 'to be' itself, this group contains the verbs *laysa* 'not to be' and a number of phasal-aspect-modifying verbs with meanings like 'to become', 'to remain', 'not to cease', 'to be about', and 'to begin'. Some of these verbs can be used with a nonverbal predicate as well, in which case they have a copular function.

In verbal complexes, both the auxiliary and the main verb are inflected for person, number, and gender. The verb forms are usually coreferential in the sense that they indicate the same syntactic subject, but exceptions are found, in which case the auxiliary is either in the 'neutral' 3rd person masculine singular, or inflected according to the so-called 'logical' subject of the sentence (see, for instance, Bravmann 1953: 81–85; Brustad 2000:145). Verb phrases do not need an overtly expressed subject (e.g., the unmarked expression of 'he writes' is *yaktubu*, and not *huwa yaktubu*). A formal analysis in which the verbal complex is interpreted as a matrix predicate with a complement, or a full verb with an adverbial clause, is possible in most (if not all) cases, since asyndetic clause linkage does occur in all varieties of Arabic. The absence of a convincing syntactic distinction of auxiliary verbs does not preclude a functional analysis of their characteristics (cf. Eisele 1992:157). With this in mind, the following description intends to be practical rather than theoretical.

The verb with auxiliary function precedes the main verb, but not necessarily immediately, as the subject or other parts of the sentence frequently separate them. Negation may be expressed either on the auxiliary or the main verb. Depending on the type of auxiliary, this may lead to a change of meaning, and as such, the effects of negation must be considered for each of the auxiliary verbs separately. Only the verb *kāna* 'to be' is lexically empty, in the sense that it expresses temporal, modal, or aspectual values without modifying the semantic content of the sentence. It occurs, in different phonetic realizations, in all varieties of the language and puts no selection restrictions upon the verbs with which it may combine. It is also used in combination with one of the other verbs with auxiliary function, resulting in a verbal complex of three verb forms.

Verbal complexes with a form of *kāna* are much more frequent than combinations with any other incomplete verb. In her analysis of a text in Modern Standard Arabic, Messaoudi (1985:175) found 360 combinations with a form of *kāna*, whereas 15 other (semi-)auxiliary verbs together accounted for 315 verbal complexes. Although exact counts are not available for other varieties, verbal complexes with a form of (the local equivalent of) *kāna* are estimated to occur frequently in most urban educated colloquials. Holes (1994:192) mentions a far more restricted use of this auxiliary in many rural and Bedouin dialect areas, especially in the speech of the less educated and the elderly. This may be related to the less elaborate tense systems encountered in these varieties.

2. Verbal complexes with a form of the auxiliary verb kāna 'to be'

In all varieties of Arabic, a limited number of temporal, modal, and aspectual distinctions can be expressed by inflection of the verb form. Depending on the type of Arabic and the linguistic context, some of these distinctions must be indicated while others are optional. Whereas simple verb forms can only adopt the morphology of one of the inflectional verb forms (and thus indicate one temporal, modal, or aspectual value), complex verb forms consisting of an inflected form of the auxiliary *kāna* and an inflected form of the main verb combine the expression of two of these values. In fact, the possibility of adding a temporal or modal distinction to a verb which itself is inflected in order to indicate an aspectual value seems to be the main *raison d'être* for the auxiliary form of *kāna* in verbal complexes (see Cuvalay 1994).

In principle, the auxiliary may take any of the verb forms existing in a particular variety. In Classical and Modern Standard Arabic it occurs in the perfect (with or without the preverbal marker *qad*) as well as the imperfect (indicative with and without the prefix *sa-* or one of the preverbal markers *sawfa* or *qad*, subjunctive and jussive, respectively), and even in the special imperative form. In Colloquial Arabic it is also used with the various future and indicative imperfect prefixes, such as *ḥa-* and *bi-* in Egyptian Arabic. For Jordanian Arabic, an auxiliary function of the active participle form *kāyin* is reported (Mitchell and El-Hassan 1994:78).

The form of the main verb is perfect (with or without *qad*) or imperfect (with or without *sa*- or one of the colloquial preverbal markers). It can also be an active participle, and in some colloquial varieties imperative forms have been attested (Bravmann 1953:131–132; Munzel 1982). Negation is usually expressed on the auxiliary, but may also be expressed on the main verb without leading to a different interpretation. It is not feasible to mention all functions ascribed to the possible combinations of auxiliary and main verb forms here. In all varieties, the specific meanings of the verbal inflections and the exact values indicated by the forms of *kāna* and the main verb depend to some extent on the context, which makes it difficult to consider all possible interpretations. Nebes (1982), for example, devotes more than 200 pages to a functional analysis of *kāna yafʿalu* (combinations of a perfect auxiliary with an indicative imperfect main verb) in Classical Arabic only.

Therefore, without claiming to be exhaustive, attested combinations in Classical and Modern Standard Arabic are presented with example sentences, followed by a summary in the form of a table. A similar table with verbal complexes in Egyptian Arabic is included to give an idea of the possibilities in colloquial varieties.

Examples from Classical Arabic (the combinations also occur in Modern Standard Arabic):

perfect auxiliary + perfect main verb (Fischer 1972:95)

(1) *hāḏihī ʾatānu- ki llatī* ***kunti***
this she-ass- your which be:2FSG
xarajti *ʿalay-hā*
go out:2FSG on- her
'This is your she-ass, on which you had gone out (that time)'

perfect auxiliary + indicative imperfect main verb (Fischer 1972:96)

(2) ***kāna*** *l- maliku* ***yamurru***
be:3MSG the- king pass:IND3MSG
bi- hi
by- him
'The king was passing by him'

(3) ***kāna*** ***yakūnu*** *fī l- bayt*
be:3MSG be:IND3MSG in the- house
'He was always [habitually, continuously] in the house'

indicative imperfect auxiliary + *qad* + perfect main verb (Fischer 1972:96)

(4) . . . *fa-* ***nakūnu*** ***qad*** ***ʾaxaḏnā***
then be:1PL ASP MKR take:1PL
ʿiwaḍ- an
substitute- ACC
'. . . then we will have taken a substitute'

subjunctive imperfect auxiliary + perfect main verb (Fischer 1972:97)

(5) *ʿasā* ***ʾan*** ***yakūna*** ***samiʿa***
maybe that be:SUBJ3MSG hear:3MSG
min- nī
of me
'Maybe he has heard of me'

imperative auxiliary + indicative imperfect main verb (Fischer 1972:107)

(6) *fa-* ***kun*** *ʾanta* ***tukallimu-***
so- be:IMPMSG you address:IND2MSG
hum
them
'So you must/should be addressing them'

Examples from Modern Standard Arabic (non-existent or infrequent in Classical Arabic)

perfect auxiliary + *sa*-imperfect main verb (Cantarino 1974:75)

(7) *maʿa ʾanna- hum* ***kānū***
with that- they be:3MPL
sa-yaqtulūna *min al- muslimīna*
FUT- kill:IND3MPL from the- Muslims
man yastaṭīʿūna qatla- hu
whoever can:IND3MPL killing- his
'But they would kill as many Muslims as they could'

sa-imperfect auxiliary + perfect main verb (Moutaouakil 1988:186)

(8) *sa- ʾuʿīru- ka l- kitāba*
FUT- lend:IND1SG- you the- book
ġadan iḏ ***sa- ʾakūnu*** ***ʾanhaytu***
tomorrow as FUT- be:IND1SG finish:1SG
qirāʾata- hu
reading- it
'I will lend you the book tomorrow as I'll have finished reading it by then'

Verbal complexes with a perfect or imperfect form with the preverbal marker *qad* are not represented in the table. *Qad* occurs with the perfect of both the auxiliary (*qad kāna faʿala*) and the main verb (*kāna qad faʿala*), but rarely with both of them in one verbal complex. *Qad* + indicative imperfect has been attested for the auxiliary of verbal complexes, but not for the main verb. So, *qad yakūnu faʿala* and *qad yakūnu yafʿalu* occur, to express the modal meaning of 'he may (well) have done' and 'he may (well) be doing' respectively.

The active participle form is also left out of the table. In verbal complexes only the main verb can take this form, with all possible aspectual values associated with it. Like adjectival and nominal predicates, the active participle gets the accusative case in combination with one of the incomplete verbs, and verbal complexes with an active participle take the form of *kāna/ yakūnu/sayakūnu/ yakūna fāʿilan*.

Some verbal complexes which would be expected to occur, such as ?*(sa-)yakūnu yafʿalu* 'he will be doing' are not mentioned in grammars or attested in descriptive studies. Apart from their use in the negation of the perfect (*lam yakun faʿala* 'he had not done' or *lam yakun yafʿalu* 'he was not doing'), jussive forms must be rare or nonexistent in verbal complexes, as they have not been described so far. The combination of an imperative auxiliary and an imperfect main verb (*kun tafʿalu*, mentioned by Bravmann 1953:124 and Fischer 1972:107) is also rare.

As the colloquial dialects differ with respect to their verbal systems, the combinations and their possible interpretations have to be listed for each of them separately. Most reference grammars pay attention to the use of the local equivalent of *kāna* in verbal complexes, but not all of them do so in a systematic way. The verbal complexes of Egyptian Arabic are reproduced in Table 2 by way of an example.

The translations of these complex forms do not represent all possible interpretations. The combinations with *kān*, for instance, are frequently encountered in conditional phrases, where, depending on the context, the verbal complex is not necessarily translated with a past tense form. Combinations with *yikūn* are often found in the complement phrases of modal expressions, where the occurrence of a bare imperfect (i.e. an imperfect form without *ḥa-* or *bi-*) is obligatory, but in an independent context *yikūn katab* can be translated with 'he may (well) be writing'. For non-dynamic main verbs, combinations of a perfect of *kāna* with a bare imperfect form refer to a past state, which is why a combination such as *kān yiʿraf* is translated as 'he knew', rather than 'he came to know, recognized'. The choice for the complex verb form in these cases is probably motivated by the fact that perfect forms of state verbs generally indicate the inception of a state, and will thus be interpreted as referring to the present.

As in Table 1, verbal complexes with an active participle form are not represented in view of the various functions associated with this form. Active participle forms of *kān* do not occur in Egyptian Arabic verbal complexes, but *kān kātib*, *biykūn kātib, ḥaykūn kātib*, and *yikūn kātib* are all possible, and often preferred to the other verbal complexes which may convey the same meaning. The combination of a *ḥa*-imperfect form of the auxiliary and a *ḥa*-imperfect form of the lexical verb is not included, either, although an example of this verbal complex is given by El-Tonsi (1992:19) in the form of *ḥaykūn ḥaynām* 'he will be about to sleep'. Most speakers, however, seem to prefer the sequence of *ḥaykūn* + *ʾarrab* + imperfect, as a more literal expression of 'to be about', and reject the possibility of a verbal complex consisting of two *ḥa*-imperfects.

In verbal complexes of all varieties we see a tendency for so-called inner aspectual values,

Table 1. Complex verb forms in Classical and Modern Standard Arabic

form of *kāna*	form of main verb	verbal complex	translation
perfect	perfect	*kāna faʿala*	he had done
	indicative imperfect	*kāna yafʿalu*	he was doing/used to do
	sa-imperfect	*kāna sayafʿalu*	he was going to do
indicative imperfect	perfect	*yakūnu faʿala*	he will have done
sa-imperfect	perfect	*sayakūnu faʿala*	he will have done
subjunctive imperfect	perfect	*yakūna faʿala*	he has done [subordinate]
imperative (rare)	indicative imperfect	*kun tafʿalu*	be (one who) does/is doing

Table 2. Verbal complexes in Egyptian Arabic

form of *kān*	form of main verb	verbal complex	translation
perfect	perfect	*kān katab*	he had written
	bi-imperfect	*kān biyiktib*	he was writing/used to write
	ḥa-imperfect	*kān ḥayiktib*	he was going to write
	bare imperfect	*kān yiktib*	he should have written/used to write
	imperative	*kunt iktib*	you should have written
bi-imperfect	perfect	*biykūn katab*	he usually has written
	bi-imperfect	*biykūn biyiktib*	he usually is writing
	ḥa-imperfect	*biykūn ḥayiktib*	he usually is about to write
ḥa-imperfect	perfect	*ḥaykūn katab*	he will have written
	bi-imperfect	*ḥaykūn biyiktib*	he will be writing
bare imperfect	perfect	*yikūn katab*	he has written [subordinate]
	bi-imperfect	*yikūn biyiktib*	he is writing [subordinate]
	ḥa-imperfect	*yikūn ḥayiktib*	he is about to write [subordinate]

such as perfect, progressive, and prospective, to be marked on the main verb, while temporal and modal distinctions are expressed by the auxiliary. Habitual aspect may be marked on the auxiliary or the main verb, depending on whether it combines with an aspectual main verb form, or a temporal or modal form of *kāna*. In combination with an obligatory perfect auxiliary in conditional phrases, the main verb can also express temporal values (see Cuvalay-Haak 1997: 199–234 for a detailed account of the ordering of temporal, aspectual, and modal values in Arabic verbal complexes). For a theoretical discussion concerning the formal distinction between the syntactic categories of the auxiliary (AUX) and main verb (PREDICATE) in Egyptian Arabic, see Jelinek (1981, 1983), who claims that such a distinction is valid, and Eisele (1992), who argues against the validity of this distinction for Egyptian Arabic.

3. Other verbs with auxiliary functions

In Classical and Modern Standard Arabic, the so-called 'verbs of incomplete predication' can be subdivided in groups according to their meaning. Some of these verbs refer to the inception of a state or an activity (*šaraʿa, ṭafiqa, badaʾa* and *jaʿala*, with the sense of 'to start, set in, set out to'; *ʾaṣbaḥa*, *ʾamsā*, *ʾaḍḥā*, etc. with the sense of 'to become'). Others indicate the continuation of a state or activity (*istamarra* 'to continue', *baqiya* 'to remain', *mā zāla* 'not to cease', etc.) or its prospective anticipation (*kāda* 'to be on the point, be about'). They themselves occur most often in the perfect and combine with another verb in the imperfect. As these verbs appear to have a lexical rather than a grammatical function, they are not discussed further here.

The incomplete verb *laysa* expresses negative polarity. It is inflected according to the morphological characteristics of the perfect verb, but does not convey the meaning of completeness or anteriority which is normally associated with this inflection. It is primarily used for the negation of nominal phrases, but can be found with verbal predicates as well. In view of its special function and limited verbal characteristics, *laysa* can best be described as a negative particle which is inflected for person, number, and gender.

In the colloquial varieties we find a number of verbs with auxiliary functions, some of which are grammaticalized to the extent that they may lose part of their original meaning. Ingham (1994:190) reports the use of the verb *baġa* 'to want' in Najdī Arabic for the expression of 'to be about' or (imminent) future. Moroccan Arabic *bġa* 'to want' can be used in a similar way (Maas 1995:257). The loss of the original volitional meaning is demonstrated by the possibility of using these auxiliary verbs with inanimate subjects. For other verbs, the transition from a full lexical verb to an auxiliary is not always easy to prove, especially since the subsequent stages in a diachronic development from lexical verb to auxiliary frequently coexist synchronically.

Apart from modal verbs with a meaning such as 'to want', 'to love', and 'to be able to', colloquial varieties each have a set of verbs similar to the phasal aspect modifying verbs in Classical and Modern Arabic, with meanings like 'to remain', 'to begin' and 'to continue'. Both of these sets usually combine with another verb in the (bare) imperfect. Verbs like 'to go', 'to come' and 'to stand up', roughly described as 'movement' verbs, typically combine with a next verb in the perfect, leading to a sequence of two perfect verbs, which together denote one single action.

Attempts to categorize the different types of verbal complexes are found in Maas (1995, with examples from Moroccan Arabic); Brustad (2000:147ff., with examples from Moroccan, Egyptian, Syrian and Kuwaiti Arabic), and Woidich (2002, with examples from Egyptian Arabic), of which the last mentioned is the most exhaustive, based on a thorough analysis of all occurrences in a number of different text types. In addition to the combinations already mentioned, Woidich (2002:181ff.) describes verbal complexes consisting of two asyndetically linked perfect forms such as *barṭam* 'to grumble' and *ʾāl* 'to say', where the second verb functions as an 'auxiliary' device, in that it provides for the possibility of adding a complement to an intransitive verb (the verb *barṭam* cannot take a complement by itself). Similar examples may well exist in other colloquial dialects, but they have not been described so far (→ serial verbs).

BIBLIOGRAPHICAL REFERENCES

Bravmann, Meir Moshen. 1953. *Studies in Arabic and general syntax*. Cairo: Institut Français d'Archéologie Orientale.

Brustad, Kristen E. 2000. *The syntax of spoken Arabic*. Washington, D.C.: Georgetown University Press.

Cantarino, Vicente. 1974. *Syntax of modern Arabic prose*, I. Bloomington, Ind.: Indiana University Press.

Cuvalay, Martine. 1994. "Auxiliary verbs in Arabic". *Function and expression in functional grammar*, ed. Elisabeth Engberg-Pedersen, Lisbeth Falster Jakobsen, and Lone Schack Rasmussen, 265–283. Berlin and New York: Mouton de Gruyter.

Cuvalay-Haak, Martine. 1997. *The verb in literary and colloquial Arabic*. Berlin and New York: Mouton de Gruyter.

Eisele, John C. 1992. "Egyptian Arabic auxiliaries and the category of AUX". *Perspectives on Arabic linguistics*, IV, ed. Ellen Broselow, Mushira Eid, and John McCarthy, 143–165. Amsterdam and Philadelphia: J. Benjamins.

El-Tonsi, Abbas. 1992. *Egyptian colloquial Arabic: A structure review*, II. Cairo: American University in Cairo, Arabic Language Institute.

Fischer, Wolfdietrich. 1972. *Grammatik des Klassischen Arabisch*. Wiesbaden: O. Harrassowitz.

Holes, Clive. 1995. *Modern Arabic: Structures, functions and varieties*. London and New York: Longman.

Ingham, Bruce. 1994. "Modality in the Arabic dialect of Najd". *Actes des premières journées internationales de dialectologie arabe de Paris*, ed. Dominique Caubet and Martine Vanhove, 185–200. Paris: Institut National des Langues et Cultures Orientales.

Jelinek, M. Eloise. 1981. *On defining categories: AUX and PREDICATE in colloquial Egyptian Arabic*. Ph.D. diss., University of Arizona, Tucson.

——. 1983. "Person-subject marking in AUX in Egyptian Arabic". *Linguistic categories: Auxiliaries and related puzzles*, ed. Frank Heny and Barry Richards, 21–46. Dordrecht: D. Reidel.

Maas, Utz. 1995. "Koverben im marokkanischen Arabisch in typologischer Perspective". *Sprachtypologische Universalienforschung* 48.239–274.

Messaoudi, Leila. 1985. *Temps et aspect: Approche de la phrase simple en arabe écrit*. Paris: Geuthner.

Mitchell, T.F. and Shahir A. El-Hassan. 1994. *Modality, mood and aspect in spoken Arabic, with special reference to Egypt and the Levant*. London and New York: Kegan Paul International.

Moutaouakil, Ahmed. 1988. *Essais en grammaire fonctionelle*. Rabat: Société Marocaine des Éditeurs Réunis.

Munzel, Kurt. 1982. "Die zusammengesetzte Verbalform *kunt iʿmil* im Kairener Arabisch: Versuch einer Erklärung". *Zeitschrift für Arabische Linguistik* 7.76–78.

Nebes, Norbert. 1982. *Funktionsanalyse von* kāna yafʿalu. Hildesheim: G. Olms.

Payne, Thomas E. 1997. *Description of morphosyntax: A guide for field-linguists*. Cambridge: Cambridge University Press.

Schachter, Paul. 1985. "Parts-of-speech systems". *Language typology and syntactic description*. I. *Clause structure*, ed. Timothy Shopen, 3–61. Cambridge: Cambridge University Press.

Steele, Susan. 1999. "Auxiliaries". *Concise encyclopedia of grammatical categories*, ed. Keith Brown and Jim Miller, 49–56. Amsterdam: Elsevier.

Woidich, Manfred. 2002. "Verbalphrasen mit asyndetischem Perfekt im Ägyptisch-Arabischen". *Estudios de Dialectología Norteafricana y Andalusí* 6.121–192.

MARTINE HAAK
(Diemen, The Netherlands)

B

Baby Talk → Caretaker Talk

Backing → Vowel Lowering

Badal → Apposition; ʾIbdāl

Baghdad Arabic

1. Introduction

1.1 Baghdad, the capital and the seat of government and financial operations, is situated on the river Tigris in the center of Iraq. In 2003 it was estimated that its population exceeded five million inhabitants, 60 percent of whom are Shiʿi. The rest are mainly Sunni, with a Christian minority comprising Chaldean Catholics, Assyrians (Orthodox and Catholics), and a small number of Greek Catholics and Protestants. The number of Mandeans, another religious minority, has decreased to fewer than two thousand. Baghdad's once thriving Sephardic Jewish community now numbers fewer than one hundred, mostly elderly individuals. Before the 1950s the number of Jewish residents was more than double that of Christians. They included several families of European Ashkenazi Jews who had arrived during and immediately after the Second World War (→ Baghdad Arabic, Jewish).

1.2 Apart from Arabic, the official language, varieties of Neo-Aramaic, Kurdish, and Armenian are spoken throughout Baghdad. The Arabic dialects comprise what Blanc (1964) has termed *gələt*, the Muslim variety, and *qəltu*, the intra-communal dialect of Christian Iraqis. It is the Muslim dialect (Baghdad Arabic) that has become the lingua franca of Iraq, and the language of commerce and education, co-existing in a diglossic relationship with Modern Standard Arabic, used in the media and in interaction with non-Iraqis. The speech of Baghdadi Christians is constantly shifting in the direction of the *gələt* variety (Abu-Haidar 1992).

1.3 Bagdadi *gələt* Arabic belongs to the 'Mesopotamian' group, and shares many features with Gulf Arabic and with varieties spoken in some parts of Syria near the Iraqi border. *gələt* Arabic is of Bedouin provenance, unlike Christian Baghdadi which is a direct descendant of Medieval (sedentary) Iraqi Arabic (Jastrow 1978). Right up until the late 1950s Baghdad Arabic consisted of a large inventory of borrowings from Turkish, Farsi via Turkish, English, and to a lesser extent French. A number of Turkish and Farsi terms are still used to refer to everyday objects, but are being superseded by Anglo-American loans. During the first decades of the 20th century, when the inhabitants of Baghdad numbered fewer than one million, some inner-city quarters had their own distinctive speech characteristics, maintained for generations. From about the 1960s, with population movement within the city, and the influx of large numbers of people hailing mainly from the south, Baghdad Arabic has become more standardized, and has come to incorporate some rural and Bedouin features.

1.4 There do not appear to exist any studies on the Arabic of Baghdad prior to the 20th century. The earliest known studies in Arabic on vocabulary and popular sayings date from about the 1960s. Works in European languages include a monograph by Malaika (1963) and an updated version of a work by Van Ess (1989). Since the 1960s Georgetown University Press has been publishing manuals and dictionaries, Erwin (1963) being the earliest and Alkalesi (2001) the latest. The data in Alkalesi (2001) is out of date and not based on any scientific methods. Blanc (1964), however, remains the most thorough study of Baghdad Arabic dialects to date.

2. PHONOLOGY

2.1 *Consonants*

The inventory of consonants is set out in Table 1.

/b/, /l/, /m/, /n/ occur as emphatics when they are in the contiguity of /ṭ/, /ḍ/, or /ṣ/, gutturals, /r/ with a back vowel, or if they precede /ā/: *ḅāṣṭərma* 'spicy sausage', *ləḏaṃ* 'to thread', *ṃəṣrān* 'intestine', *ṣaṇṭa* 'quiet', *ḅūri* 'pipe', *ḅāḷōn* 'balloon', *ṃāku* 'there is not'. /l/ is realized as /ḷ/ in *aḷḷa* 'God' and other compound forms like *baḷḷa* 'by the way', *ḥayaḷḷa* 'anyhow'.

The interdentals /ṯ/ and /ḏ/ are retained in borrowings from Standard Arabic, with the exception of *ṯalāṯa* 'three' where initial /ṯ/ > /t/: *tlāṯa*. Older Christian Baghdadi speakers ordinarily realize dentals for interdentals (/t/ < /ṯ/, /d/ < /ḏ/, /ḍ/ < /ḏ̣/), but in more recent loans /ṯ/ is often retained: *ṯəyatar* 'theater', *talawwəṯ* 'pollution'.

/ʿ/ is retained, except in *ʿaṭa* 'to give' which becomes *nəṭa*.

The glottal stop /ʾ/ is rare. Unlike some Levantine Arabic dialects, it is not an allophone of /q/ and occurs in borrowings from Standard Arabic: *səʾal* 'to ask', *tʾaxxar* 'to be delayed'.

Standard /k/ and /q/ are frequently realized as /č/ and /g/ respectively. /k/ and /q/ are sometimes retained. The following examples show how forms with /k/ and /q/ have a non-everyday meaning:

/g/ < /q/	*gaddar* 'to try on [clothes]'	*qaddar* 'to estimate'
	ʿugda 'knot'	*ʿuqda* 'complex'
/č/ < /k/	*čalāwi* 'kidneys [culinary]'	*kalāwi* 'kidneys [anat.]'
	mčattaf 'with arms crossed'	*mkattaf* 'fettered'

/č/ and /q/ occur also in borrowings from non-Arabic languages:

/č/ *čāmərḷuġ* 'mudguard', *čarəx* 'wheel', *ṣaččam* 'to insult', *ṣūč* 'fault';
/q/ *qačaġči* 'smuggler', *qappaṭ* 'to wind down', *qašmar* 'to mock', *šaqa* 'joke'.

Other reflexes of /q/ are /k/, /j/, and /x/:

/q/ > /j/ *jədər* ~ *jədrəyya* 'saucepan', *šarji* 'East wind'.
/q/ > /k/ *wakət* 'time', *kətal* 'to beat, to kill'.
/q/ > /x/ *ṭaxəṃ* 'suit [clothes], set [of items]', *yāxa* 'collar'.

/j/ is invariably realized as a voiced alveolar affricate. Reflex /y/, common in some southern Iraqi Arabic dialects, does not occur in Baghdad Arabic.

Standard /ġ/ is retained, except in the terms *ġasal* 'to wash' and *staġfar* 'to ask forgiveness' where /ġ/ > /x/: *xəsal*, *ṣṭaxfar*.

/r/ is usually realized as /ġ/ [ɣ] in Christian Baghdadi. However, Christian Baghdadi speak-

Table 1. Inventory of consonants

	bilabial	labiodental	dental	interdental	alveolar	palatal	velar	uvular	pharyngeal	glottal
plosive	p b		t d				k g	q		ʾ
emphatic			ṭ							
fricative	w	f	s z	ṯ ḏ	š	y	x ġ		ḥ ʿ	h
emphatic			ṣ	ḏ̣						
affricate					č j					
liquid			l r							
emphatic			ḷ							
nasal	m		n							
emphatic	ṃ		ṇ							

ers retain /r/ in borrowings from standard Arabic and non-Arabic languages: *tarākum* 'accumulation', *kampyūtar* 'computer'.

Phoneme /p/ occurs in a number of loans: *parda* 'curtain', *pulak* 'sequins', *lappa* 'purée of boiled rice'; *pākēt* 'packet, parcel', *pīkāp* 'truck'. In some forms both /p/ and /č/ of the donor language are retained: *pančar* 'flat tyre, puncture', *pačata* 'napkin', *parčam* 'fringe [hair]'

/v/ is not used, except in code-switched forms, like *vīsa* 'visa', *vəlla* 'villa'. The term 'television' occurs as *talfəzyōn.*

2.2 *Vowels*

2.2.1 Short vowels

Table 2. Short vowels

i u

ə o

a

/i/ occurs in unstressed open syllables, in word-final position: *wardi* 'pink', *sarsari* 'layabout, good-for-nothing', *jēti* 'you [sg. fem.] came'.

/ə/ is found in stressed and unstressed, open and closed syllables:

bəčər 'first-born', *čəma* 'truffles', *təmman* 'rice', *bətənjān* 'aubergine'.

/u/ occurs in unstressed open and closed syllables: *ndallētu* 'you pl. found the way', *xəlug* 'patience', *əskut* 'be quiet!'.

/o/ occurs in unstressed, word-final position in loans: *pālṭo* 'overcoat', *rādyo* 'radio', *čmīnto* 'cement'.

/a/ is found in all positions: *agūm* 'I get up', *angas* 'worse', *darbūna* 'alley', *xərga* 'cloth, rag', *qamarəyya* 'trellised vine'.

2.2.2 Long vowels

Table 3. Long vowels

ī ū

ē ō

ā

All long vowels occur in open or closed syllables:

bībi 'grandmother', *pīp* 'large metal container'; *jēbi* 'my pocket', *ləwēš* 'why?'; *bṭūla* 'bottles', *manšūl* 'he has a cold'; *ṃəssāḥa* 'rubber', *jīrān* 'neighbors'; *rōba* 'yoghurt', *lōz* 'almonds'.

In feminine CvCv/CvCCv forms denoting color or defect, Christian Baghdadi speakers invariably realize long /ā/ in open final position:

Baghdad Arabic	Christian Baghdadi
ḥaṃra	*ḥaṃġā* 'red'
ʿamya	*ʿamyā* 'blind'

ʾimāla (*-ā-* > *-ē-*; *-a* > *-i*) does not occur in Baghdad Arabic, although it is a common feature of Christian Baghdadi:

-ā- > *-ē*	Baghdad Arabic	Christian Baghdadi
	člāb	*klēb* 'dogs'
	dkākīn	*dakēkīn* 'shops'
-a > *-i*	*ḥəlwa*	*ḥəlwi* 'pretty'
	bazzūna	*bazzūni* 'cat'

2.2.3 Glide vowel [ie]

The glide vowel [ie] is the reflex of long vowel /ē/. It is more common in the speech of women and men of rural origin. Thus: *bēt* 'house', *hammēna* 'also, again', *nṭēt* 'I gave' are sometimes realized as [biet], [hamːiena], [nɨiet], respectively.

2.2.4 Diphthongs

The diphthongs /aw/ and /ay/, including /ay/ of the dual, are contracted and become long vowels /ō/ and /ē/:

Modern Standard Arabic	Baghdad Arabic	Christian Baghdadi
sawdāʾ 'black'	*sōda*	*sōdā*
baytayn 'two houses [acc.]'	*bētēn*	*bētēn*

Diphthongs /aw/ and /ay/ are retained in some loans: *rawḏ̣a* 'kindergarten' and *ḥaywān* 'animal'. Variants *ḥēwān* and *ḥīwān* are used by older speakers. Diphthongs are retained also in proper names like *xawla* and *maysūn.*

/ay/ occurs in some compound forms with the negative particle *ma*: *mayṣīr* 'definitely not', *mayxāləf* 'never mind'.

It is more usual to find diphthongs where /w/ or /y/ is geminate and the preceding vowel /a/, /u/ or /ə/: *mdawwra* 'round [fem.]', *guwwa* 'strength, defiance', *ʿələwwa* 'his/its height'; *msayyja* 'fenced [fem.]', *həyya* 'she'.

2.2.5 Syllabication

Syllables are of the Cv, Cv̄, CvC, Cv̄C, CCv̄, CvCC and CCv̄C patterns

Cv syllables occur either in initial, medial or final position: *qa-nafa* 'sofa', *mad-ra-sa* 'school', *sayyā-ra* 'car'.

In CvCvC forms the intial vowel of Cv̄ is invariably *ə*: *səmač* 'fish', *gəṣab* 'reeds', *rəsab* 'to fail'.

Cv̄ syllables occur in initial or medial position, or as monosyllabic words: *qū-ṭəyya* 'box', *jā-bō-ha* 'they brought her', *ʿaṭ-šā-na* 'thirsty [sg. fem.]', *mū* 'not, isn't that so?', *šī* 'thing'.

CvC syllables occur in initial position: *fət-ḥīha* 'open it! [sg. fem.]', *rug-ba* 'neck', *man-qala* 'brazier'.

Cv̄C syllables occur in monosyllabic words or in word-final position: *bīr* 'well', *jēb* 'pocket', *šāf* 'to see', *yōm* 'day'; *ba-nāt* 'girls', *ʿan-ka-būt* 'spider'.

Monosyllabic CvCC words end in a geminate consonant: 'to open', *ḥabb* 'watermelon seeds', *ṣəll* 'young snake'.

CCv̄C monosyllabic words are usually plural, where v̄ is either /ū/, /ā/ or /ī/: *ngūˁ* 'dried apricots', *trāb* 'dust', *nhīb* 'theft'.

In reduplicated CvC words the short vowel is /ə/ in nouns and /a/ in verbs: *fəlfəl* 'pepper', *ṣərṣər* 'beetle'; *sarsar* 'to gad about', *pašpaš* 'to whisper'.

2.2.6 Stress assignment

In disyllabic forms stress (´) falls on the second syllable if its vowel is long, otherwise it is the first syllable that is stressed:

jwārī́n 'the neighborhood', *naddā́f* 'carder', *malyā́n* 'full', *mṣārī́n* 'intestines'.

mára 'woman', *bə́ttna* 'our daughter', *ḥā́rəs* 'guard', *ḥjā́ra* 'stone'.

In forms of more than two syllables, stress falls on the penultimate syllable. Trisyllabic: *səndā́na* 'flower pot', *msannā́ya* 'balcony on the river', *sāfárna* 'we traveled', *rattbáthum* 'she tidied them up'. Quadri-syllabic: *ʿarabā́yən* 'carriages', *ṃaṣārī́fkum* 'your expenses', *maḥābə́sha* 'her rings', *nəsēnā́hum* 'we forgot them'.

2.2.7 Phonotactics

2.2.7.1 Assimilation

In Baghdad Arabic, as in Modern Standard Arabic, the /l/ of the definite article is assimilated to the sunletter that follows it: *š-šaməs* 'the sun', *j-jəmmār* 'the edible shoots of the palm tree', *r-rašma* 'the bridle'. But *l-guṃar* 'the moon', *l-kāġad* 'the paper'. In Christian Baghdadi, however, /l/ is often assimilated to the moonletter following it: *əq-qaṃaġ* 'the moon', *əy-yōm* 'the day', *əb-bēʿa* 'the church'.

Modern Standard Arabic /j/ is realized as /š/ when it is followed by /t/ in the same syllable: *j-h-d*: *štəhad* 'he studied hard' but *məjtəhəd* 'hard-working'; *j-m-ˁ: štəmāˁ* 'meeting' but *mujtamaˁ* 'society'.

Where /n/ precedes /t/, /d/ or /r/ assimilation can occur: *ənti* 'you [sg. fem.]' > *ətti*, *bənti* 'my daughter' > *bətti*, *ˁəndi* 'I have' > *ˁəddi*, *yənrād* 'wanted, desirable' > *yərrād*.

2.2.7.2 Metathesis is rare; the most frequently occurring examples are:

Modern Standard Arabic root		Baghdad Arabic
j-n-z-r	>	*mzanjər* 'rusty'; *zanjīl* 'chain'
l-ˁ-n	>	*yənʿal* 'he curses'
séchoir (French)	>	*šəswār* 'hair-dryer'.

2.2.7.3 Voicing of /ṣ/ occurs when it precedes /d/:

qaṣdi > *qaẓdi* 'my intention, I mean', but *qəṣad* 'to mean'; *aṣdəqāʾ* > *aẓdəqāʾ* 'friends', but *ṣadāqa* 'friendship'. Voicing of /s/ and /ṣ/ occurs more frequently in Christian Baghdadi: *masdūd* > *mazdūd* 'closed'; *masbaḥa* > *mazbaḥa* 'rosary'; *sbūˁ* > *zbōˁ* 'week'.

2.2.8 Morphophonology

2.2.8.1 Consonant clusters and anaptyxis

There are no three-consonant clusters in Baghdad Arabic. Two-consonant clusters occur in initial position only, in CCv̄ and CCv̄C: *ḥrāġi* 'screws', *ṇġūla* 'bastards', *zmāl* 'donkey', *gṃāṭ* 'swaddling clothes'.

In word-final position -CC is geminate: *ḍabb* 'to throw away', *ṃurr* 'bitter'. Where the comparable standard form ends in non-geminate -CC an anaptyctic vowel (/ə/ or /u/) is invariably inserted.

Modern Standard Arabic Baghdadi	Baghdad	Christian Arabic
ṇafṭ	*ṇafəṭ*	*ṇafəṭ* 'petrol'
šuġl	*šuġul*	*šəġəl* 'work'
ḍarb	*ḍarub*	*ḍaġəb* 'beating'

2.2.8.2 Elision

In nominal and verbal forms with final -CvC, the vowel of the second syllable is elided on suffixation: *bārəd* [sg.masc.] > *bārdīn* ‘cold [pl.]’; *mkammal* [sg.masc.] > *mkammlīn* ‘perfect [pl.]’; *dəras* ‘he studied’ > *dərsat* ‘she studied’; *rājaʿ* ‘he revised’ > *rājʿaw* ‘they revised’

3. MORPHOLOGY

3.1 *Pronouns*

3.1.1 Independent subject pronouns

1st sg.	*āni*	1st pl.	*əḥna*
2nd sg. masc.	*ənta*	2nd pl.	*əntu*
2nd sg. fem.	*ənti*		
3rd sg. masc.	*huwwa*	3rd pl.	*huṃṃa*
3rd sg. em.	*həyya*		

3.1.2 Suffixed direct object pronouns

1st sg.	*-ni*	1st pl.	*-na*
2nd sg. masc.	*-ak*	2nd pl.	*-kuṃ*
2nd sg. fem.	*-əč*		
3rd sg. masc.	*-a*	3rd pl.	*-huṃ*
3rd sg. fem.	*-ha*		

There is no 3rd pl. gender distinction in Baghdad Arabic proper, only in the speech of Baghdad residents of rural or Bedouin origin.

3.1.3 Suffixed indirect object pronouns

These are formed by the addition of *-l-* to singular and *-əl-* to plural direct object pronouns, with the exception of 1st sg. where *-ni* > *-li*. Thus compare: *jābni* ‘he brought me’ but *jābli* ‘he brought to me’, *wakkala* ‘he deputized him’ but *wakkalla* ‘he deputized for him’.

3.1.4 Suffixed possessive pronouns

These are the same as direct object pronouns, the only exception being 1st sg. which is *-i* , or *-ti* if the noun ends in a vowel: *ktāb* ‘book’ > *ktābi* ‘my book’, *madrasa* ‘school’ > *madrasti* ‘my school’.

3.1.5 Independent demonstrative pronouns

this/these	sg. masc. *hāḏa*
	sg. fem. *hāḏi* ~ *hāy*
	pl. *haḏōla*
that/those	sg. masc. *haḏāk* ~ *haḏāka*
	sg. fem. *haḏīč*
	pl. *haḏōlāk*

3.1.6 Relative pronoun

This is *l* or *(ə)lli*, the latter occurring when emphasis is required.

3.1.7 Interrogative pronouns

The following are the most frequently used: *mənu* ‘who?’, *šənu* ‘what?’, *yāhu* ‘which one?’.

3.2 *Adverbs*

The following are the common adverbs of place, manner, and time:

3.2.1 Place

hnā ‘here’, *hnāk(a)* ‘there’, *lī-gəddām* ‘forwards’, *lī-wara* ‘backwards’, *jawwa* ‘inside’, *ḅarra* ‘outside’, *haṣ-ṣafḥa* ‘this side’, *ḏīč-əṣ-ṣafḥa* ‘that side’.

3.2.2 Manner

hīčī ‘thus’, *kulləč* ‘very’, *hwāya* ‘much, a lot’, *šwayya* ‘a little’, *zēn* ‘well’, *ḥēl* ‘quickly, loudly’, *bəl-ʿajal* ‘without delay, quickly’, *yawāš* ‘slowly’.

3.2.3 Time

l-yōṃ ‘today’, *l-bārḥa* ‘yesterday’, *bāčər* ‘tomorrow’, *mən zamān* ‘a long time ago’, *kull-wakət* ‘always’, *dōrāt* ‘sometimes’, *ṃāṭūl* ‘as long as’, *fad-dōra* ‘directly’.

3.3 *Particles*

3.3.1 Definite article

The definite article in Baghdad Arabic is *l-* preceding moon letters; it is assimilated to the sound it precedes with sun letters (see 2.2.7.1).

3.3.2 Indefiniteness

This is usually expressed by *farəd* ~ *fadd* and *wāḥəd*, both meaning ‘one’.

3.3.3 Particle of possession

This is *māl* ~ *mālat* sg. and *mālāt* pl. ‘of, belonging to’.

3.3.4 Prepositions

b- ‘in’, *l-* ‘to’, *ʿal(a)* ‘on’, *mən* ‘from’, *ʿədd* ‘at’, *fōg* ‘on top of’, *jawwa* ‘under’, *yamm* ‘near’, *gəddām* ‘in front of’, *wara* ‘behind’, *gbāḷ* ‘opposite’, *bēn* ‘between’.

3.3.5 Conjunctions

w ‘and’, *lō* ‘if, or’, *lākən* ~ *lākət* ‘but’, *əlla* ‘except’, *baʿdēn* ‘then’, *ʿala-mūd* ‘so that’.

3.3.6 Interrogative particles

lēš ~ *luwēš* ‘why?’, *bēš* ‘how much?’, *šənu* ‘what?’, *šgadd* ‘how much?’, *šwakət* ‘when?’, *ščamm* ‘how many?’, expressed also as *čamm*.

3.3.7 Negative particles

3.3.7.1 Simple negative particles

ma ~ *mā* ‘not’, *mū* ‘not’, *la* ~ *lā* ‘no’.

3.3.7.2 Compound negative particles

ṃāku ‘there is not’, *mayṣīr* ‘do not’, *malāzəm* ‘must not’.

3.3.8 Vocative particles

Some kinship terms and caritatives occur as vocative particles, the most common being *yāba* ‘father’ and *yuṃṃa* ‘mother’. These are best translated as ‘I say’ or ‘by the way’. Other vocative particles include the terms of endearment *ʿēni* ‘my eye’, *ʿyūni* ‘my eyes’, *gaḷbi* ‘my heart’, and *fədwa* and *ṣadaqa*, which both mean ‘may I be a sacrifice to you’. All these can be translated as ‘my dear’.

3.3.9 Particles that introduce sentences

Some particles introduce both affirmative and interrogative sentences where the verb is perfect or imperfect. These include *ʿād*, *ʿūd*, *ašū*, *bāri*, *baḷḷa*, which can all be translated as ‘incidentally, by the way’, as can *tara* and *waḷḷa*, which introduce affirmative sentences only.

3.4 *The noun*

A noun can be derivative or non-derivative. Derivative nouns are either deverbal or denominal:

deverbal	*najāḥ* ‘success’ < *nəjaḥ* ‘to succeed’
	fərja ‘spectacle’ < *tfarraj* ‘to watch’
denominal	*mwarrad* ‘floral’ < *warəd* ‘flowers’
	mahāra ‘skill’ < *māhər* ‘skillful’

CvCCv̄C nouns where medial CC is geminate denote habit or profession: *ḥayyāl* ‘liar’, *sakkīr* ‘drunkard’, *ballām* ‘boatman’, *sammāč* ‘fisherman’.

Nouns ending in *-i* or *-āwi* often refer to a person’s origin: *baġdādi* ‘Baghdadi’, *īṭāli* ‘Italian’, *namsāwi* ‘Austrian’, *ḅaṣrāwi* ‘Basran’.

Nouns denoting a traditional profession usually take the ending *-či* (< Turkish *-çı*): *pāčači* ‘tripe vendor’, *fītarči* ‘mechanic’, *bəstanči* ‘gardener’

CāCv(C) nouns are usually adjectives: *fāhi* ‘faint, insipid’, *ġāli* ‘expensive’, *šāṭər* ‘clever’, *xāḅəṭ* ‘murky [river]’.

aCCaC nouns are either adjectives of color and defect or elatives: *azrag* ‘blue’, *axras* ‘dumb’; *akḅar* ‘bigger’, *angas* ‘worse’.

3.4.1 Gender

The feminine is usually formed by adding the marker *-a* to sg. masc. nouns or collective substantives: *malək* ‘king’ > *maləka* ‘queen’, *ʿaṭšān* > *ʿaṭšāna* ‘thirsty’, *dəjāj* ‘chicken’ > *dəjāja* ‘hen’, *ṃōz* ‘bananas’ > *ṃōza* ‘a banana’.

There are a number of feminine nouns derived from the base form of Form I verbs: *gaṣṣ* ‘to cut’ > *gaṣṣa* ‘a cut’, *akal* ‘to eat’ > *akla* ‘a dish, type of food’, *nām* ‘to sleep’ > *nōma* ‘a sleep’, *wəgaʿ* ‘to fall’ > *wagʿa* ‘a fall’.

Where the masculine noun ends in *-i* the feminine marker is *-əyya*: *alṃāni* > *alṃānəyya* ‘German [fem.]’, *šurṭi* > *šurṭəyya* ‘policeman > policewoman’.

Substantives that refer to female creatures or double parts of the body are feminine: *uxut* ‘sister’, *bībī* ‘grandmother’, *īd* ‘hand, arm’, *ʿēn* ‘eye’.

3.4.2 Number

The dual is formed by the addition of the suffixes *-ēn* to sg. masc. and *-tēn* to sg. fem. nouns, after the elsion of the final weak vowel: *ktāb* ‘book’ > *ktābēn* ‘two books’, *sana* ‘year’ > *santēn* ‘two years’, *čaləb* ‘dog’ > *čalbən* ‘two dogs’, *qābəla* ‘midwife’ > *qābəltēn* ‘two midwives’.

The plural is formed either by suffixation or by changing the morphemic shape of the singular. Adjectives, participles, and nouns denoting habit or profession usually take the suffix *-īn*: *amīn* > *amīnīn* ‘trustworthy, reliable’, *rāyəḥ* > *rāyḥīn* ‘going’, *manšūl* > *manšūlīn* ‘ill with a cold’, *xaḅḅāz* > *xaḅḅāzīn* ‘bakers’, *čaḏḏāb* > *čaḏḏābīn* ‘liars’.

Nouns referring to origin and ending in *-i* take the suffix *-yyīn*, while those denoting habit take *-əyya*: *sūri* > *sūrəyyīn* ‘Syrians’, *ḥəllāwi* > *ḥəllā-*

wəyyīn 'inhabitants of Ḥilla', *sarsari* > *sarsarəyya* 'lay-abouts', *ḥarāmi* > *ḥarāməyya* 'thieves'.

A number of sg. fem. nouns ending in *-a* take the plural suffix *-āt*: *ḥəlwa* > *ḥəlwāt* 'pretty', *šābba* > *šābbāt* 'young women'.

Plurals that are formed by morphemic change include the following: CCāC *ṭwāl* 'tall', CCūC *byūt* 'houses', CəCaC *xərag* 'rags', CvCĀCəC *qanābəl* 'bombs', CaCāCa *yatāma* 'orphans', CCāCīC *bzāzīn* 'cats'.

3.5 *Numerals*

3.5.1 Cardinal

Numbers 1 to 20 are: 1 *wāḥəd*, 2 *t̠nən*, 3 *tlāt̠a*, 4 *arḅaʿa*, 5 *xamsa*, 6 *sətta*, 7 *sabʿa*, 8 *t̠mānya*, 9 *təsʿa*, 10 *ʿašra*, 11 *daʿaš*, 12 *t̠ṇaʿaš*, 13 *ṭḷaṭṭaʿaš*, 14 *arḅāṭaʿaš*, 15 *xaṃəṣṭaʿaš*, 16 *səṭṭaʿaš*, 17 *ṣḅāṭaʿaš*, 18 *t̠ṃəṇṭaʿaš*, 19 *ṭəṣāṭaʿaš*, 20 *ʿəšrīn*.

Numbers 3 to 10 occur in the construct as: 3 *tlat̠*, 4 *arbaʿ*, 5 *xaməs*, 6 *sətt*, 7 *sabəʿ*, 8 *t̠man*, 9 *təsəʿ*, 10 *ʿašər*.

When a number in the construct, other than *sətt*, is followed by the plural of *yōm* or *šahar*, *t-* is inserted between the two forms: *arḅaʿ t-əyyām* 'four days', *t̠man t-əšhur* 'eight months'.

The numbers 30–100 are: 30 *tlāt̠īn*, 40 *arbaʿīn*, 50 *xaṃsīn*, 60 *səttīn*, 70 *sabʿīn*, 80 *t̠mānīn*, 90 *təsʿīn*, 100 *məyya*. The construct of *məyya* 'hundred' is *mīt*.

The number 200 is formed by adding the dual suffix *-ēn* to the construct of *mīt*: *mītēn*.

From 300 onwards the numbers are formed by adding the prefixes of the construct numbers to *məyya* or to *mīt*: 600 *səttməyya*, *təsəʿmīt dīnār* '900 dinars'.

3.5.2 Ordinal

Numbers 1 to 10:

1st *awwal* 2nd *t̠āni* 3rd *t̠ālət̠* 4th *rābəʿ* 5th *xāməs* 6th *sādəs* 7th *sābəʿ* 8th *t̠āmən* 9th *tāsəʿ* 10th *ʿāšər*

3.6 *The verb*

3.6.1 The perfect

The base form (Form I) of triradical strong verbs is of the pattern CəCaC: *xətal* 'to hide'. In Christian Baghdadi the base form has two patterns: CaCaC and CəCəC. Other Form I verbs are of the patterns aCaC *akal* 'to eat'; CaCC *madd* 'to spread'; CəCa (III*y*) *ḥəča* 'to speak', and CāC (II*w*/*y*) *nām* 'to sleep'. The base form is also the 3rd pers. sg. masc. perfect to which the following suffixes are added, in keeping with the rule of the elision of the unstressed vowel:

3rd sg. masc.	-Ø	3rd sg. fem.	*-at*	3rd pl.	*-aw*
2nd sg. masc.	*-ət*	2nd sg. fem.	*-ti*	2nd pl.	*-tu*
1st sg.	*-ət*			1st pl.	*-na*

aCaC verbs and the medial hamza verb *səʾal* follow the same rules of suffixation in the perfect as CəCaC verbs, as do CāC (II*w*/*y*) verbs, where *ā* > *ə* except in the 3rd persons. Thus compare: *gəmti* 'you [sg. fem.] got up', but *gāṃaw* 'they got up'.

Geminate CaCC, C,Ca (III*y*) and the doubly weak verb *jā* ~ *əja* 'to come' have a long vowel *ē* in 3rd pers. sg. fem., 2nd pers. sg. masc. and 2nd pers. sg. fem., and 1st pers. sg. and pl.:

1st sg., 2 sg. masc.	2nd sg. fem.	1st pl.	2nd pl.
maddēt	*maddēti*	*maddēna*	*maddētu*
ḥəčēt	*ḥəčēti*	*ḥəčēna*	*ḥəčētu*
(ə)jēt	*(ə)jēti*	*(ə)jēna*	*(ə)jētu*

3.6.2 The imperfect

In 1st pers. sg. and pl. of CəCaC verbs the characteristic vowel is /ə/, unless it is in the contiguity of gutturals or /r/ and a back vowel when it is /a/: *anzəl*, *nənzəl* 'I/we go down'; *agdar*, *nəgdar* 'I/we can'. There are exceptions, e.g. *ləbas* > *albas*.

In forms that have /x/ or /ġ/ the vowel is usually /u/: *axṭub*/*nəxṭub* 'I/we propose marriage', *aġlub*, *nəġlub* 'I/we win'.

The vowel of II*w*/*y* verbs in the imperfect can be *ā*, *ū* or *ī*: *ynām* 'he sleeps', *ygūṃ* 'he gets up', *ytīh* 'he gets lost'.

The following are the imperfect affixes of the ə and pers. sg. and pl.:

CəCaC	aCaC	CaCC	CəCa	Cā (jā)	CāC
yəCCəC	yāCuC	yCəCC	yəCCi	yəCi	yCāC ~ yCūC ~ yCīC

The 3rd pers. sg. fem and 2nd pers. sg. masc. have the same shape as the 3rd pers. sg. masc. but take the prefix *t-*; the 3rd pers. pl. has prefix *y-* and suffix *-ūn*; the 2nd pers. sg. fem. has prefix *t-* and suffix *-īn*.

3.6.2.1 A present continuous or habitual action, a future action, and a past continuous or habitual action are expressed by particles *da* and *raḥ*, and auxiliary verb *čān* respectively introducing the imperfect:

da yəthammlūn əhānāt əlʿərāqəyyīn 'The Iraqis are putting up with humiliation'

raḥ yətnāqšūn b ha lmawḏūʿ bāčər 'They will discuss this matter tomorrow'

čənna nəzraʿ lāləngi jawwa ṇṇaxal 'We used to plant mandarin [trees] under the palm trees'

3.6.3 The imperative

The imperative is formed from the imperfect base form:

ənzəl 'go down [masc.]', *nəzli* sg. fem., *nəzlu* pl.; *sədd* 'close [masc.]', *səddi* sg. fem., *səddu* pl.; *gūm* 'get up [masc.]', *gūmi* sg. fem., *gūmu* pl., etc.

3.6.4 Derived forms

Form II verbs, with a geminate medial CC, usually denote an intensive, transitive action: *ləbas* 'to wear' > *labbas* 'to dress someone'; *šagg* 'to tear' > *šaggag* 'to tear to pieces'; *məša* 'to walk' > *mašša* 'to get s.t. moving'.

Form III verbs are formed by lengthening the first vowel of CəCaC verbs: *rəjaʿ* 'to return' > *rājaʿ* 'to review, to revise'; *ḥəkam* 'to judge' > *ḥākam* 'to prosecute'.

Forms V and VI are formed by the addition of *t-* to Forms II and III respectively. Form V conveys an intransitive, reflexive action, while Form VI denotes a reciprocal action: II *ʿallam* 'to teach' > V *tʿallam* 'to learn', *baddal* 'to change [tr.]' > *tbaddal* 'to change [intr.]'; III *wājah* 'to face someone' > VI *twājah* 'to meet someone', *ʿātab* 'to reproach' > *tʿātab* 'to reproach one another'.

Form VII, which frequently conveys the passive, is formed by the addition of prefix *n-* to Form I verbs: *kəsar* 'to break' > *nkəsar* 'to be broken'; *qəra* 'to read' > *nqəra* 'to be read'.

Form VIII is formed by the insertion of *-t-* after the first radical of Form I or II verbs: *šəʿal* 'to light' > *štəʿal* 'to be set alight'; *šəka* 'to complain' > *štəka* 'to lodge a complaint'.

Form IX verbs are restricted to verbs of color and defect, and are of the pattern CCaCC where final -CC is geminate: *ḥṃarr* 'to blush, to go red'; *ṭrašš* 'to become deaf'.

Form X verbs are rare and tend to be denominal Modern Standard Arabic loans: *stəḥaqq* 'to merit' < *ḥaqq* 'right, due'; *stərāḥ* 'to rest' < *rāḥa* 'rest'.

3.6.5 Participles

3.6.5.1 Active participles, which inflect for gender and number, correspond to all the verbal forms. Form I is of the CāCəC ~ māCəC ~ CāCi patterns: *xāyəf* 'afraid', *sādəd* 'having closed'; *mākəl* 'having eaten'; *nāsi* 'having forgotten'. For *əja* it is *jāy* 'coming, having come'. Active participles of derived forms are formed by the addition of prefix *m-* to Forms II and III, and prefix *mə-* to Forms V–X, and by changing the vowel of the second syllable from /a/ to /ə/.

3.6.5.2 Passive participles are of Forms I, II, and III: I maCCūC, II mCaCCaC, III mCāCaC.

3.6.6 Quadriradical verbs

Quadriradical verbs have two Forms, I (tr.) and II (intr.). Most are locally coined and not used outside Iraq:

Form I *xarbaṭ* 'to mix up', *čaqlab* 'to turn upside down [tr.], *qašmar* 'to mock', *hajwal* 'to make homeless', *gaḷwaṭ* 'to put on tenterhooks'.

Form II *txarbaṭ* 'to feel unwell, to be disheveled', *tčaqlab* 'to turn upside down [intr.]', *tqašmar* 'to be fooled', *thajwal* 'to be made homeless', *tgaḷwaṭ* 'to be on tenterhooks'.

Some have reduplicated CaC: *waswas* 'to be anxious', *sarsar* 'to gad about', *dabdab* 'to crawl [child]', *ṭabṭab* to pat'.

3.6.6.1 The active and passive participles are mCCaCCəC and mCCaCCaC respectively.

4. SYNTAX

4.1 *Definiteness*

This is expressed by the definite article introducing the noun: *lxəṭṭār* 'the guest[s]', *ddarūna* 'the alley'. Indefiniteness can be expressed by the use of a noun or a noun phrase without a marker: *fītarči zēn* 'a good mechanic', or by *wāḥəd* 'a, one' (used with masc. nouns only), or *fadd* or *farəd* 'a, one' (used for both genders and numbers): *wāḥəd šāyəb galli* 'an old man told me';

fadd sayyāra mčarqəʿa ‘a battered car’; *fadd ryājīl sarsarəyya* ‘good-for-nothing man’; *fadd banāt ḥəlwāt* ‘pretty girls’. *fadd* sometimes means ‘just one’: *ʿəddhum fadd əbnayya* ‘they have one daughter only’.

4.2 *Possession*

This can be expressed in several ways:

Suffixing possessive pronouns to nouns: *ḥamāta* ‘his sister-in-law’, *ʿagədna* ‘our street’.

māl ~ *mālat* occurring independently or with pronominal suffixes: *lmanādər māl əlqanafa* ‘the cushions of the sofa’, *ššēx mālatna* ‘our religious leader’.

The independent pronoun and the possessive pronominal suffix it refers to: *āni gubbti mū čəbīra* ‘my room isn’t large’; *huwwa bənta da ddārī* ‘his daughter is looking after him’.

Two nominal elements occurring as a construct: *šaʿb əlʿərāq* ‘the people of Iraq’, *murabbəyat əbni* ‘my son’s nanny’.

The terms *abu* ‘father of; he of’ and *umm* ‘mother of; she of’ occur in constructs preceding a defined noun. The compound refers to a person’s profession or some other description, often physical: *abu ləbwāri* ‘the plumber’; *abu lmaṭʿam* ‘the restaurant owner’; *umm əlgēmar* ‘the [sg. fem.] buffalo cream vendor’; *abu ššwārəb* ‘the man with the moustache’; *abu lḥənəč* ‘the man with the [protruding] chin’; *umm əttarāči ləmdandəla* ‘the woman with the dangly earrings’.

4.3 *Concord*

An adjective ordinarily agrees in gender with the substantive it qualifies: *mgaddi ḥayyāl* ‘a deceitful beggar’, *bazzūna sōda* ‘a black cat’. An adjective is definite if it qualifies a definite substantive: *lmaqām əlʿərāqi* ‘the Iraqi Maqam’, *wələdna əššəṭṭar* ‘our clever children’. Pl. masc. and pl. fem. substantives referring to humans are ordinarily qualified by pl. masc. and pl. fem. adjectives respectively: *šabāb tāyhīn* ‘aimless youth’, *ʿammāt ḥanūnāt* ‘loving paternal aunts’. Plural substantives referring to animals or inanimate objects of either gender are qualified by a sg. fem. adjective: *člāb mačlūba* ‘rabid dogs’, *mṣāləx waṣxa* ‘dirty basins’. Duals are ordinarily qualified by plural adjectives: *bētēn zġār* ‘two small houses’, *ġəṣnēn ṭāggīn* ‘two branches in blossom’.

4.4 *The plural of abundance*

This is of the CCāCīn pattern. In Baghdad Arabic the two forms that take a simple plural and a plural of abundance are: *jār* ‘neighbor’ > *jīrān* ‘neighbors [next door]’ > *jwārīn* ‘the neighborhood’; *ḥāyəṭ* ‘wall’ > *ḥīṭān* ‘walls’ > *ḥyāṭīn* ‘[lots of] walls’.

4.6 *Negation*

ma ~ *mā* ‘not’ ordinarily negates verbs, while *mū* ‘not’ negates non-verbs. *la* ~ *lā* ‘no’ is used to negate imperatives. *māku* ‘there is not’, a compound of *ma* + *aku* ‘there is’, ordinarily negates nominal sentences:

ma ygədrūn yʿīšūn bala mukayyəfa bə ṣṣēf ‘They can’t live without an air-conditioner in summer’

ma ṭālʿa mən bēti lyōm ṯman təšhur ‘I haven’t left my house for the past eight months’

əddərāsa baʿd əlməṭawaṣṣəṭa ṃū əlzāməyya ‘After the first three years of secondary school education isn’t compulsory’

ṃū huwwa hāḏa aṯāri ġēra ‘That is not him, apparently it is someone else’

la tətʿāmal wəyyā ‘Don’t have any dealings with him!’

ṃāku aḥad wəyyāhum ‘There isn’t anyone with them’

4.7 *Interrogation*

Affix *-š* ‘what?’ is the hallmark of interrogation. It occurs in the interrogative particles *lēš* ‘why?’, *bēš* ‘how much?’, *šənu* ‘what?’, *šgadd* ‘how much?’, *šwakət* ‘when?’, *ščamm* ‘how many?’. *š-* can also be prefixed to an imperfect or perfect verb or participle in interrogative sentences: *šdatsawwi* ‘what are you doing?’, *šgəttəlla* ‘what did you tell him?’, *šmākəl əlyōm* ‘what did you have to eat today? Unlike some other Arabic dialects, *-š* in Baghdad Arabic has no negative connotation, even in the term *ṃakuš* ‘there is not’, a variant of *ṃāku*.

4.8 *Anticipatory pronominal suffix*

The anticipatory pronominal suffix + the defined object it refers to is frequently used in Baghdad Arabic. This construction has the same semantic value as a subject + verb + direct object with definite marker:

kəsart ərraggəyya	*kəsarətha lərraggəyya* 'I cut open the watermelon'
ʿāzmīn jīrānhum ʿa ssafra	*ʿāzmīhum əljīrānhum ʿa ssafra* 'They have invited their neighbors to the picnic'

The anticipatory pronominal suffix is sometimes used in nominal sentences where emphasis is required:

	gubbt əjjəhāl	'the children's room'
but	*gubbathum lə jjəhāl*	'[it is] the children's room'.

5. Lexicon

Baghdad Arabic has a number of borrowings from Turkish, Farsi via Turkish, English, and French.

5.1 Turkish and Turkish via Farsi terms in Baghdad Arabic constitute older borrowings:

atag 'petticoat', *jaṇṭa* 'handbag', *čāmurluġ* 'mud-guard', *čarəx* 'wheel', *čālġi* 'orchestra, concert', *dugma* 'button', *mēz* 'table', *ṃuṣlux* 'basin, sink', *saxtači* 'cheat', *ṭāwa* 'frying-pan', *ūti* 'iron', *yalag* 'waistcoat', *xāšūga* 'spoon', *xōš* 'good', *zangīn* 'rich'.

5.2 English borrowings refer to mechanical or technological items:

stērən 'steering-wheel', *brēk* 'brake', *hōrən* 'horn', *klāč* 'clutch', *lōri* 'truck', *gḷōb* 'light bulb', *gḷāṣ* 'glass [for drinking]', *pāysəkəl* 'bicycle', *taləfōn* 'telephone', *talfəzyōn* 'television', *kampyūtar* 'computer', *antarnēt* 'internet'.

5.3 French loans refer mainly to clothing and materials:

blūz 'blouse', *sətyān* 'bra', *pāḷṭo* 'overcoat', *rōb ~ rōbdəšām* 'dressing-gown', *tūl* 'tulle', *jōrjēt* 'georgette', *krēša* 'crèpe-de-Chine', *šəswār* 'hair-dryer, *dīkōr* 'décor', *qanafa* 'sofa', *məkyāj* 'make-up'.

5.4 There are a number of terms that have extended or different meanings from standard forms derived from the same root:

Modern Standard Arabic

qaṭṭa 'to sharpen nib'
šawwafa 'to show [in good light]'
tašarraba 'to soak up'
raʾa 'to see'
saffaṭa 'to repair, to put right'
ḥamā 'mother-in-law'
harwala 'to walk fast'
dəʿbəl 'frog spawn'
nəqrəs 'gout'

Baghdad Arabic

məqṭāṭa 'pencil-sharpener'
šāf 'to see'
təšrīb 'pieces of bread in meat broth'
rāwa 'to show'
ṣaffat 'to arrange neatly'
ḥama/ḥamā 'brother/sister-in-law'
harwal 'to jog'
duʿbul '[game of] marbles'
nəjrəs 'small mosquito'

Biographical references

Abu-Haidar, Farida. 1992. "Shifting boundaries: The effect of MSA on dialect convergence in Baghdad". *Perspectives on Arabic linguistics*, IV, ed. Ellen Broselow, Mushira Eid, and John McCarthy, 91–106. Amsterdam and Philadelphia: J. Benjamins.

Alkalesi, Yasin. 2001. *Modern Iraqi Arabic: A textbook*. Washington, D.C.: Georgetown University Press.

Blanc, Haim. 1964. *Communal dialects in Baghdad*. Cambridge, Mass.: Harvard University.

Erwin, Wallace M. 1963. *A short reference grammar of Iraqi Arabic*. Washington, D.C.: Georgetown University Press.

Jastrow, Otto. 1978. *Die mesopotamisch-arabischen* qəltu-*Dialekte*. I. *Phonologie und Morphologie*. Wiesbaden: F. Steiner.

Malaika, Nizar. 1963. *Grundzüge der Grammatik des arabischen Dialektes von Bagdad*. Wiesbaden: O. Harrassowitz.

Van Ess, John. 1989. *The spoken Arabic of Iraq*. 2nd ed. Oxford: Oxford University Press.

Farida Abu-Haidar (London, U.K.)

Baghdad Arabic Jewish

1. General

1.1 *Speakers*

Jewish Baghdad Arabic is the Arabic dialect spoken by the Jews of Baghdad and other towns of Lower Iraq. The dialects spoken by Jews in Upper Iraq, such as in ʿAna and Mosul, differ in some aspects from that of Baghdad. The Jews spoke Jewish Baghdadi at home and with members of their community, but with Muslims they used the Muslim dialect or Classical Arabic. Jewish Baghdadi was also used abroad, in Jewish Baghdadian communities established in the 18th, 19th, and early 20th centuries in Bombay, Calcutta, Rangoon, Singapore, Shanghai, Hongkong, Manchester, etc., but in the course of time the English language replaced Jewish Baghdadi. After the mass emigration of Jews from Iraq in 1949–1952, Israel became the major center of Jewish Baghdadi speakers. However, Jewish Baghdadi nowadays bears traces of the local language. Reliable informants of Jewish Baghdadi are growing more and more scarce. Moreover, it is mainly the older people who still speak Jewish Baghdadi in its original form, and theirs is perhaps the last generation to speak it as a mother tongue.

1.2 *Written language*

The Jews of Baghdad also have a written language, which employs Hebrew characters, and is used for translations (*šarḥ*) of the Old Testament, parts of liturgical literature, such as the Passover Haggada, stories, etc. This language differs from colloquial Jewish Baghdadi, and may be considered as a literary language (Blanc 1964b).

1.3 *Linguistic type*

A most interesting sociolinguistic feature of Baghdad is the existence of three distinct dialects: Muslim, Jewish, and Christian Baghdadi. Muslim Baghdadi belongs to a group called *gilit* dialects and is a more recent type with Bedouin influence, while Jewish Baghdadi (as well as Christian Baghdadi) belongs to the *qəltu* dialects, assumed to be "direct descendants of dialects spoken by the urban population of the Abbasid Iraq", and it is possible that "features that are at present peculiar to Jewish Baghdadi may hark back to the vernacular of medieval Baghdad" (Blanc 1964a: 166–167).

1.4 *State of research*

Phonology and morphology have been studied quite thoroughly, but only very little work has been undertaken on syntax (Blanc 1964a; Mansour 1991; Jastrow 1990). Vocabulary has been worked on only partially (Blanc 1964a:133–159; Mansour 1974–1983:III, 1991:45–50; Avishur 2001) but there are lexicons and collections of proverbs without linguistic treatment (Sassoon 1949:190–199; Ben-Yaacob 1985; Yona-Swery 1995; Me'iri 1997). Additional bibliography can be found in Blanc 1964a:173–180, Mansour 1991:313–316, Avishur 2001:323–344.

1.5 *Recordings*

Jewish Baghdadi recordings can be heard on the Internet at <http://semarch.uni-hd.de/>.

2. PHONOLOGY

2.1 *Consonants*

2.1.1 List of consonants
In Table 1 the unmarked consonants are phonemes, and those in parentheses are borderline cases.

2.1.2 Comparison with Literary Arabic

2.1.2.1 The reflex of Literary Arabic /q/ is /q/ in Jewish Baghdadi. In only a few loanwords Jewish Baghdadi has /g/ (2.1.2.4) and rarely has /j/ for Literary Arabic /q/. The Literary Arabic interdentals /ḏ, ṯ, ḏ̣/ have been retained in Jewish Baghdadi; they are replaced by stops in only a few words: *jġēdi* 'rat' (cf. Literary Arabic *juraḏ*).

Table 1. Inventory of consonants

	bilabial	labio-dental	inter-dental	alveolar	palatal	velar	uvular	pharyngeal	glottal
stops	b p			d t	g k		q		ʾ
fricatives		(v) f	ḏ ṯ	z s	š	ġ x		ʿ ḥ	h
nasals	m			n					
laterals				l					
trills				r					
affricates					j č				
semivowels	w				y				
emphatics	(ḅ ṃ)		ḏ̣	ṭ ṣ (ẓ ḷ ṇ)					

2.1.2.2 Jewish Baghdadi has three emphatics: /ṣ/ (cf. Literary Arabic /ṣ/), /ṭ/ (cf. Literary Arabic /ṭ/) and /ḏ̣/ (cf. Literary Arabic both /ḍ/ and /ḏ̣/). Other emphatics, such as /ḷ, ḅ, ṃ, ṇ, ẓ/, heard in certain words, are better regarded as secondary emphatics (Mansour 1991:57–59). Emphatics are also found in loanwords, possibly due to the effort made by Jewish Baghdadi speakers to articulate foreign sounds: *tōẓ* 'dust, powder' (Turkish *toz*), *gḷāṣ* 'drinking glass' (English *glass*).

2.1.2.3 The realization of Literary Arabic /r/ is one of the chief features distinguishing Jewish Baghdadi and Christian Baghdadi from Muslim Baghdadi. While Muslim Baghdadi has /r/, Jewish Baghdadi (as well as Christian Baghdadi) usually has /ġ/ for Literary Arabic /r/: *ġās* 'head' (Literary Arabic *raʾs*), *ʿášġa* 'ten' (Literary Arabic *ʿašara*). The [r]-sound also exists in Jewish Baghdadi, but generally in loanwords: *brāxa* 'blessing' (Hebrew *bəraḵa*), *náfar* 'person, individual' (Persian *nafar*), *kōndra* 'shoe' (Turkish *kundura*). The Arabic words with /r/ are mostly of recent origin, or are loans from Literary Arabic or from one of the dialects (only a few of them are not obviously loans), e.g. *səyyāra* 'automobile, car', *qiṭār* 'train'. Cf. *ʿáskaġ* 'army' with /ġ/, but *ḥākəm ʿáskari* 'martial law', a modern expression, with /r/.

Moreover, in some cases a change from /ġ/ to /r/ can bring about a change of meaning, e.g. *faġġ* 'to pour, serve food', but *farr* 'to throw'; *qáddaġ* 'to measure', but *qáddar* 'to estimate, value'; *báġġa* 'outside', but *barr* 'desert'. Each of these doublets consists of two forms of the same Arabic root, but the form with /r/ appears to have come into Jewish Baghdadi at a later date and with a different meaning. Or it may be that Jewish Baghdadi has now begun to exploit these two sounds to express different meanings (Mansour 1955–1956, 1991:29–32; Blanc 1964a:20–25).

2.1.2.4 The consonants /p, č, g/, which do not occur in Literary Arabic, came into Jewish Baghdadi from other languages. These consonants have become so integrated with the other Jewish Baghdadi consonants that speakers of Jewish Baghdadi no longer feel them to be foreign. /g/ occurs only rarely in Arabic words, and these are all loans from a dialect in which Literary Arabic /q/ is realized as /g/: *gámaz* 'to jump' (Literary Arabic *qamaz*, Muslim Baghdadi *gumaz*); or from a dialect in which Literary Arabic /j/ is /g/: *dárag* 'drawer (of a table)' (Literary Arabic *durj*).

/v/ occurs in only a very few loanwords, e.g. *vwāl* 'bridal veil or train' (< French *voile*), and its status as a phoneme is therefore questionable.

2.1.3 Assimilations and elisions

2.1.3.1 Assimilation of voiced to voiceless consonants and vice versa occurs in clusters, and is generally regressive: *dfātəġ* > *tfātəġ* 'notebooks'; *qʿádtu* > *qʿáttu* 'I sat down'; *aš-jābak* > *aj-jābak* 'what brought you here?'

Assimilation of stop to fricative and vice versa is relatively rare. They generally occur with other phonetic changes, such as voiced/voiceless or velarized/non-velarized.

2.1.3.2 Velarization. An emphatic consonant, as well as /x, ġ, q/, will generally velarize the adjacent (and sometimes also the non-adjacent) consonant, either regressively or progressively: *ṣlā* > *ṣḷā* 'prayer, synagogue', *stánḏ̣aġ* > *ṣṭáṇḏ̣aġ* 'he waited for'. When /w/ is assimilated to a preceding labial consonant, the resulting geminate may be velarized: *bwáḥdu* > *ḅḅáḥdu* 'by himself', *fwanīṣ* > *ffanīṣ* 'lamps'.

2.1.3.3 /ġ/ in the clusters /ġx/, /xġ/ and /ġq/, /qġ/ tends to assimilate to /x/ and /q/ respectively, either regressively or progressively: *ġxīṣ* > *xxīṣ* 'cheap'; but not in all words, e.g. *áxġas* > *áxxas* 'dumb [sg. masc.]', but *áġxas* 'cheaper'; *qġīb* > *qqīb* 'near', but *qġūn* 'horns'. In some cases /ġ/ is elided: *(l)bōḥi* (Literary Arabic *al-bāriḥa*) 'yesterday'.

2.1.3.4 /l/ may assimilate to /n/ in the cluster /ln/: *qə́lna* > *qə́nna* 'we said'. The second /l/ of *lelt* (construct state) 'night' is heard very weakly or elided altogether: *lēt-əl-ḥə́nni* 'the Henna Night'. Concerning the definite article *l-* see 3.4.3.1.

/n/ and /m/, in certain phonetic contexts, may assimilate to each other: *yənmáġəd* > *yəmmáġəd* 'he will be upset'; *tə́mṭəġ* > *tə́nṭəġ* 'it rains'. The /n/ of *bənt* may assimilate to /t/ in construct state: *bətt-xalə́ti* 'my cousin'.

2.2 *Vowels*

2.2.1 List of vowels

Table 2 lists the phonemic vowels, including borderline cases:

Table 2. Vowels

ī				ū	i				u
	ē		ō			e	ə	o	
		ā					a		
		long vowels					short vowels		

2.2.2 Distribution of Vowels

2.2.2.1 Long vowels

These occur in stressed syllables, both open and closed, both final and non-final. The only exception is /ū/ which does not occur in final open syllables. Long vowels do not, as a rule, occur in unstressed syllables. However, /ī/ < /iy/ and /ū/ < /uw/ occur in unstressed syllables: *msībī́n* < *msiybī́n* 'having abandoned [pl.]', *mūzū́n* < *muwzū́n* 'weighed'.

2.2.2.2 Short vowels

i. Stressed syllables. Short vowels do not occur in stressed final open syllables. In all other stressed syllables, the vowels that regularly occur are /a/ and /ə/. The vowels /i/ and /u/ are heard before /y/ and /w/ respectively: *híyyi* 'she', *húwwi* 'he'. /i/ and /o/ occur in some particles: *íḏa* 'if'.

ii. Post-stress syllables. In open final syllables only /a, i, u/ occur. In closed final syllables, as a rule, only /a/ and /ə/ occur. In words with an antepenultimate stress, the non-final syllable following the stressed syllable is usually an open one: *ṯánawi* 'secondary school'.

iii. Pre-stress syllables. Short vowels are usually elided before the stressed syllable, but they are found in the following categories: (a) a short vowel which was originally a long one (2.2.4.2.i); (b) after /ʾ/ (2.2.4.2.ii); (c) /i/ and /u/ before /y/ and /w/ respectively; (d) loanwords: *jarīda* 'newspaper'; and (e) /ə/ as anaptyctic (2.2.5).

iv. Minimal pairs are found in /a-i-u/ (*bēta* 'her house', *bēti* 'my house', *bētu* 'his house'), and in /a-ə/ (*báqča* 'small flower garden', *bə́qča* 'bundle'). The relation of /ə/ to the other short vowels, with the exception of /a/, is not always clear. In the syllables in which /i/ and /u/ regularly occur, /ə/ does not occur at all.

v. Relation between the long and short vowels. Only /ā/ and /a/ form minimal pairs (*ḥākət* 'she wove, knitted', *ḥákət* 'she spoke'). None of the other short vowels normally occur in those syllables in which their corresponding long vowels occur. Blanc (1964a:30) does not list /o/ and /e/ among the phonemes, and considers them to be merely variants of the corresponding long vowels /ō/ and /ē/.

2.2.3 Comparison with Literary Arabic

2.2.3.1 Literary Arabic /ā/ and the → *ʾimāla*. The *ʾimāla* is a characteristic of the *qəltu* dialects. The most striking feature of Jewish Baghdadi is the treatment of Literary Arabic /ā/. It is often represented in Jewish Baghdadi by /ī/ or /ē/.

In nouns, in which Literary Arabic /ā/ is proximate to /i/, Jewish Baghdadi has /ī/ (while Christian Baghdadi has /ē/, and Muslim Baghdadi has /ā/): Literary Arabic *jāmiʿ* 'mosque', *kilāb* 'dogs' > Jewish Baghdadi *jīməʿ*, *klīb*. However, in verbs, in the active participle of Form I, Literary Arabic /ā/ is represented in Jewish Baghdadi by /ē/ (as in Christian Baghdadi): Literary Arabic *jāmiʿ* 'having gathered' > Jewish Baghdadi *jēməʿ*. These changes have yielded doublets in Jewish Baghdadi: *jīməʿ* 'mosque' vs. *jēməʿ* 'having gathered'.

Not every Literary Arabic /ā/ which is proximate to /i/ has *ʾimāla* in Jewish Baghdadi, e.g. in adjective plurals corresponding to the Literary Arabic pattern *qitāl*, the vowel /ā/ is retained in Jewish Baghdadi: Literary Arabic *kibār* 'big' > Jewish Baghdadi *kbāġ*.

2.2.3.2 Literary Arabic short vowels /i/ and /u/ have merged in Jewish Baghdadi into /ə/ in closed syllables and in open stressed syllables. This merger has resulted in changes of forms and paradigms. For example, Literary Arabic noun patterns CiCC and CuCC are both realized in Jewish Baghdadi as CəCC: Literary Arabic *ʾuxt* 'sister', *bint* 'daughter' > Jewish Baghdadi *əxt*, *bənt*. In pre-stress open syllables Literary Arabic /a, i, u/ are generally elided in Jewish Baghdadi (2.2.2.2.iii).

2.2.4 Position of word stress and its influence on vowels

2.2.4.1 Stress occurs on the final syllable, if this contains a long vowel or ends with a cluster: *yumḗn* 'two days', *afə́rr* 'I throw'. If the last syllable has neither a long vowel nor a cluster, the stress falls on the penultimate, whether the vowel in the penultimate is long or short: *qəbbə́tu* 'his room', *aftáhəm* 'I understand' (an anaptyctic is not included in the vowel count).

In some words, stress is on the antepenultimate. Most of them are either loans in affixless forms: *tánaga* 'tin', *qánafa* 'sofa' (with suffixes they behave according to the word stress proposed above: *tanagā́t* ‘tins’, *qanafə́tu* ‘his sofa’), or compound words and words with preposed particles: *ḥámdəlla* ‘thank God!’, *áš-aku* ‘what’s the matter?’. Blanc (1964a:52) asserts that when the penultimate syllable is open and short, stress occurs on the antepenultimate, but for explaining words like *qəbbə́tu* and *aftáhəm*, he introduces morphological considerations: “certain nominal and verbal base types stress the syllable immediately preceding an object pronoun suffix” (e.g. *qəbbə́tu* and “in the imperfect of Forms VII and VIII verbs” Jewish Baghdadi and Christian Baghdadi “always stress the first syllable of the base” (e.g. *aftáhəm*) (Blanc 1964a:52–53).

2.2.4.2 Vowels from which stress has shifted:

i. A long vowel is shortened and may even undergo a change of quality: /ā, ī, ū/ > /a, i, u/ respectively; and /ē, ō/ > /i, u/ respectively, e.g. *taqā́tal* ‘he quarreled’ > *tqatáltu* ‘I quarreled’, *bēt* ‘house’ > *bitḗn* ‘two houses’, *yōm* ‘day’ > *yumḗn* ‘two days’. In some forms /ē/ and /ō/ may be shortened without change of quality, e.g. in participles with the pattern CēCəC: *kḗtəb* ‘writing, having written [sg.masc.]’ > *ketbī́n* [pl.]; and in some loanwords.

ii. A short vowel from which stress has shifted forward is generally elided: *kátab* ‘he wrote’ > *ktábu* ‘he wrote it [sg.masc.]’. If, as a result, a sequence of three or more word-initial consonants is produced, an anaptyctic vowel intrudes (2.2.5.1). A short vowel preceded by /ʾ/ is not usually elided: *áxaḏ* ‘he took’ > *axáḏu* ‘he took him’. In loanwords the short vowel may be retained: *náfar* ‘person’ > *nafarḗn* ‘two persons’.

2.2.5 Consonant clusters and anaptyxis

2.2.5.1 Initial clusters

i. A sequence of two initial consonants preceding the stressed vowel is always realized as a cluster. It is sometimes necessary, in order to facilitate the realization of the initial consonants as a cluster, to introduce a prosthetic vowel before the two consonants: *[ə]ġkīk* ‘weak, feeble’, *[ə]llēli* ‘tonight’.

ii. Initial $C_1C_2C_3$V yield CəCCV: *ləkbīġ* ‘the big one’, *bəʿṯḗtu* ‘she sent him’. When C_1C_2 are /st/ or /št/, all three initial consonants are articulated as a cluster: *štġáltu* ‘I worked’, *stḥētu* ‘I was ashamed’.

iii. $C_1C_2C_3C_4$V yield generally CCəCCV, e.g. *tkəbōn* ‘you [pl.] write’; but also CəCəCCV, particularly with preposed particles: *mnəlbēt* ~ *mənəlbēt* ‘from the house’. When C_2C_3 are /st/ or /št/, they are not separated: *təstḥōn* ‘you [pl.] are ashamed’.

2.2.5.2 Medial clusters

i. Medial two-consonant clusters occur with any combination of consonants.

ii. Medial $C_1C_2C_3$ clusters occur in certain combinations, e.g. when C_1 is n: *pánjra* ‘window’; when C_1C_2 are /št, st, rt/: *kəštbān* ‘thimble’, *pərtqāl* ~ *pərṭqāl* (also *pəṭqāl*) ‘oranges’. When C_1C_2 are geminate, they are reduced to a single consonant: *fárrni* > *fárni* ‘he threw me’. In many other words speakers unintentionally tend to insert an anaptyctic between C_1 and C_2: *áġbʿa* > *áġəbʿa* ‘four’.

2.2.5.3 Final clusters

Final clusters are limited to two consonants, and usually occur after a short vowel. They are found mostly in noun patterns C_1əC_2C_3 and C_1aC_2C_3, and in verbs in the perfect, 2nd person sg. masc. But while quite a number of them may be realized as a cluster, others require an anaptyctic. The need for an anaptyctic depends on the relative sonority of the consonants. If C_2 has less sonority than C_3 an anaptyctic will intrude between C_2 and C_3. Cf. /lb/ vs. /bl/: *qalb* ‘heart’, *kalb* ‘dog’ vs. *qábəl* ‘before’, *ḥábəl* ‘rope’, *zábəl* ‘garbage’.

2.2.6 Diphthongs

Literary Arabic diphthongs /aw/ and /ay/ have undergone monophthongization in Jewish Baghdadi, as in many Arabic dialects: *aw* > *ō*, *ay* > *ē*. The diphthongs /Vw/ and /Vy/ are retained when /w/ or /y/ is geminate (*qáwwa* ‘he strengthened’, *máyyət* ‘dead [sg. masc.]’), when /aw/ and /ay/ are preceded by /ʾ/ (*áwzən* ‘I weigh’, *áybas* ‘drier; I’ll become dry’), and in loanwords. When unstressed, *ō* (< *aw*) and *ē* (< *ay*) may be changed to /u/ and /i/ respectively (cf. 2.2.4.2.i): Literary Arabic *bayḍāʾ* ‘white [fem.]’, *sawdāʾ* ‘black [fem.]’ > Jewish Baghdadi *biḏ̣ā*, *sudā*.

2.3.7 Hiatus and glide

Two adjacent vowels (a hiatus) occur within a single word when a vowel suffix, or a suffix with an initial vowel, is added to a word ending in a vowel. Jewish Baghdadi tends to avoid a hiatus and does so (a) by inserting a glide between the vowels: *dáwa* ‘medicine’+ *āt* for pl. > **dwaāt* > *dwayāt*; *pālṭu* ‘coat’ + *āt* > **palṭuāt* > *palṭuwāt* (pl.). *dūni* ‘bad [sg. masc.]’ + *i* for fem. > **dunii* > *duníyyi* [sg. fem.]; and (b) by eliding one of the vowels: *šakāsa* ‘cup’+ *āt* > **šakasaāt* > *šakasāt* ‘cups’.

3. MORPHOLOGY

3.1 *The pronoun*

3.1.1 Subject pronouns

3.1.1.1 The free subject pronouns are: *ána* (1st sg.), *ə́nta* (2nd sg. masc.), *ə́nti* (2nd sg. fem.), *húwwi* (3rd sg. masc.), *híyyi* (3rd sg. fem.), *nə́ḥna* (also *əḥna* 1st pl.), *ə́ntəm* (2nd pl.), *hə́mmi* (3rd pl.). There is no gender distinction in the plural.

3.1.1.2 The subject pronoun affixes attached to the perfect are listed in Table 3 (with *kátab* ‘to write’). Suffixes ending in a vowel have stressed alternants, which occur when followed by an additional suffix: *ktəbtōnu* ‘I wrote it [sg. masc.]’. The items in parenthesis occur before 3rd pers. sg. fem. direct pro-nominal suffix: *ktəbtūha* (also *ktəbtúwa*) ‘I wrote it [sg. fem.]’.

Table 3. Subject pronoun affixes (perfect)

	unstressed		stressed
3rd sg. m	*-ø*	*kátab*	*-ø*
3rd sg. f	*-ət*	*kátbət*	*-ət*
3rd pl.	*-u*	*kátbu*	*-ō (ū)*
2nd sg. m	*-t*	*ktábt*	*-t*
2nd sg. f	*-ti*	*ktábti*	*-tē (tī)*
2nd pl.	*-təm*	*ktábtəm*	*-təm*
1st sg.	*-tu*	*ktábtu*	*-tō (tū)*
1st pl.	*-na*	*ktábna*	*-nā*

The suffix *-tu* of the 1st pers. sg. is one of the distinguishing features of the *qəltu*-dialects.

3.1.1.3 The subject pronoun affixes attached to the imperfect and imperative are listed in Table 4.

Table 4. Subject pronoun affixes (imperfect)

	imperfect		imperative	
3rd sg. masc.	*yə-*	*yə́ktəb*		
3rd sg. fem.	*tə-*	*tə́ktəb*		
3rd pl.	*y … ōn*	*ykətbōn*		
2nd sg. masc.	*tə-*	*tə́ktəb*	-ø	*ktəb*
2nd sg. fem.	*t … ēn*	*tkətbēn*	*-i*	*ktə́bi*
2nd pl.	*t … ōn*	*tkətbōn*	*-u*	*ktə́bu*
1sg.	*a-*	*áktəb*		
1st pl.	*nə-*	*nə́ktəb*		

The retention of *-n* of the suffixes is typical of the Mesopotamian area. This *-n* is elided when an object pronominal suffix is added. With the 3rd pers. sg. fem. direct object pronominal suffix the vowels *-ē* and *-ō* change to *-ī-* ~ *-iy-* and *-ū-* ~ *-uw-*, respectively: *tkətbīha* ~ *tkətbiya* ‘you [sg. fem.] write it [sg. fem.]’. When a suffix is added to the imperative the vowels *-i* and *-u* change to *-ē-* and *-ō-* respectively: *kətbēnu* ‘write [sg. fem.] it [sg. masc.]!’

3.1.2 Object pronouns

3.1.2.1 Direct object pronouns
The pronominal suffixes attached to verbs, nouns, and particles are listed in Table 5:

Table 5. Object pronoun suffixes

	after C	after V
3rd sg. masc.	*-u*	*-nu*
3rd sg. fem.	*-a*	*-ha*
3rd pl.	*-əm*	*-həm*
2nd sg. masc.	*-ak*	*-k*
2nd sg. fem.	*-ək*	*-ki*
2nd pl.	*-kəm*	*-kəm*
1sg.	*-i*, *-ni*	*-yi*, *-ni*
1st pl.	*-na*	*-na*

The 1st pers. sg. has two alternants: *-ni* occurs after all verbal forms and after a small number of particles: *jābni* ‘he brought me’, *lə́ni* ‘I have’; *-i* occurs after nouns and particles.

3.1.2.2 Indirect object pronouns
The indirect object pronominal suffixes attached to verbs are listed in Table 6:

Table 6. Indirect object pronominal suffixes

3rd sg. masc.	*-lu*	2nd sg. masc.	*-lak*	1st sg.	*-li*
3rd sg. fem.	*-la*	2nd sg. fem.	*-lək*		
3rd pl.	*-ləm*	2nd pl.	*-lkəm*	1st pl.	*-lna*

3.1.2.3 Double object

The double object pronominal suffix includes both direct and indirect object pronominal suffixes, the latter specified, the former unspecified: *jabəlyāk* 'he brought it/him/her/them to you [sg.masc.]'. The regular Jewish Baghdadi suffixes may be assumed to be as in Table 7.

Table 7. Double object pronominal suffixes

3rd sg. masc.	*-lyānu*	2nd sg. masc.	*-lyāk*	1st sg.	*-lyā*
3rd sg. fem.	*-lyāha*	2nd sg. fem.	*-lyāki*		
3rd pl.	*-lyāhəm*	2nd pl.	*-lyākəm*	1st pl.	*-lnyā*

However, there is some difference of usage among Jewish Baghdadi speakers.

3.2 *The noun*

3.2.1 The feminine suffix

The feminine suffix of the noun is either *-a* or *-i*, depending on the preceding vowel. If the base final syllable contains /y, i, ē/ or /ə/, then the feminine suffix is generally *-i*: *faqāri* 'poor [sg. fem.]'; in all other cases *-a*, e.g. *kálba* 'dog [fem.]', but there are some exceptions. In sandhi the feminine suffix is *-ət*: *qə́bba* 'room' > *qəbbə́tu* 'his room'.

3.2.2 The unit noun

This is formed by adding the feminine suffix to a collective noun, but generally by affixing *-āyi*: *bəṭṭīx* 'melon', *bəṭṭixāyi* 'a melon'.

3.2.3 The plural

3.2.3.1 The broken plural. Jewish Baghdadi also uses it with non-Arabic words: *pánjra* 'window' in plural: *pnājər*.

3.2.3.2 The sound plural. In addition to *-īn* and *-āt*, Jewish Baghdadi also uses two Hebrew suffixes, *-īm* and *-ōṯ*, mainly with Hebrew words: *ḥxamīm* 'rabbis', *braxōṯ* 'blessings'.

3.2.4 The numerals

3.2.4.1 The cardinal numbers. 'one' *wēḥəd* (masc.), *wáḥdi* (fem.). 'two' *ṯnēn* (masc.), *ṯnēn*/*ṯəntēn* (fem.). The dual and plural + *ṯnēn* are also used: *qəbtēn* 'two rooms', *ṣənnāˁ ṯnēn* 'two servants'.

The numerals from 'three' to 'ten' have three shapes: (a) in isolation; (b) when preceding a noun; and (c) when preceding certain nouns (*iyyām* 'days', *ə̂šhəġ* 'months' and *alāf* 'thousands'). There is no difference between masculine and feminine (Table 8).

Table 8. Numerals from 'three' to 'ten'

	A	B	C
3	*tlāṯi*	*tlaṯ-*	*tlattiyām*
4	*áġ(ə)bˁa*	*aġbaˁ-*	*aġbaˁtiyām*
5	*xámsi*	*xams-*	*xamstiyām*
6	*sə́tti*	*sətt-*	*səttiyām*
7	*sábˁa*	*sabˁ-*	*sabəˁtiyām*
8	*ṯmīni*	*ṯmən-*	*ṯməntiyām*
9	*tə́sˁa*	*təsˁ-*	*təsəˁtiyām*
10	*ˁášġa*	*ˁašəġ-*	*ˁašəġtiyām*

Shape C consists of shape B with an added *-t*, and when pronounced, this *-t* is joined to the noun: *ˁášəġ-tiyām*.

11 *idáˁ(ə)š*, 12 *ṯnáʾˁ(ə)š*, 13 *ṭḷəṭṭáˁ(ə)š*, 14 *aġḅaṭáˁ(ə)š*, 15 *xṃəṣṭáˁ(ə)š*, 16 *ṣəṭṭáˁ(ə)š*, 17 *ṣḅaṭáˁ(ə)š*, 18 *ṯṃənṭáˁ(ə)š*, 19 *ṭṣaṭáˁ(ə)š*, 20 *ˁəšġīn*, 30 *tliṯīn*, 40 *ġəbˁīn*, 50 *xəmsīn*, 60 *səttīn*, 70 *səbˁīn*, 80 *ṯminīn*, 90 *təsˁīn*.

100 *mīyi* ~ *míyyi*, 200 *mitēn*, 1,000 *alf*, 2,000 *əlfēn*. 'Hundred' has the variant *mīt* before a numbered noun: *mīt lēra* 'a hundred pounds'. The numbers 300–900 are formed with numerals of shape B: *tlaṯmīyi* ~ *tlaṯmíyyi* '300' The numbers 3,000–9,000 are formed with numerals of shape C: *tlattalāf* '3000'.

3.2.4.2 The ordinal numbers. 1st: *áwwal* and *wlāni*. 2nd–10th: *ṯīni*, *ṯīləṯ*, *ġībəˁ*, *xīməs*, *sīdəs*, *sībəˁ*, *ṯīmən*, *tīsəˁ*, *ˁīšəġ*.

3.3 *The verb*

3.3.1 General

The verbal forms of Classical Arabic, except Form IV, have been preserved in Jewish Baghdadi. The forms CēCaC and CōCaC are listed here as variants of Form III, and their passive forms tCēCaC and tCōCaC as variants of Form VI. Blanc (1964a:110) classified these with quadriradical verbs with /w/ or /y/ as 2nd radical. There is no gender distinction in the plural.

Each of the forms has its base patterns, one for the perfect and one for the imperfect and the imperative. However, as a result of stress shift and other changes, they yield modifications. In the paradigms below, one form of each root pattern is given, as follows: in the perfect: (a) 3rd sg. masc. (*kátab*) for all 3rd persons, but when the root pattern of 3rd pers. sg. fem. and 3rd pers. pl. varies from 3rd pers. sg. masc., the 3rd pers. pl. is also given; (b) 1st pers. sg. (e.g. *ktábtu*) for all 1st and 2nd persons. In the imperfect and imperative: (a) imperfect 2nd pers. sg. masc. (*tə́ktəb*), and when 2nd pers. sg. fem., 2nd pers. pl. and 3rd pers. pl. vary, 2nd pers. sg. fem. is also given; (b) imperative 2nd pers. sg. masc. (*ktəb*), and, if necessary, also 3rd pers. sg. fem. In some cases additional forms are given.

3.3.2 Form I

Strong verbs. *kátab* 'to write'; *fátaḥ* 'to open'. Perfect: *kátab*, *kátbu*, *ktábtu*. Imperfect: *tə́ktəb*, *tkətbēn*; *ktəb*. *tə́ftaḥ*, *tfətḥēn*; *ftaḥ*. The perfect has a single base pattern, with C_2-a, and the imperfect has two, with C_2-*a* and C_2-*ə*.

Geminate. *farr* 'to throw'. Perfect: *farr*, *fərrētu*. Imperfect: *tfərr*; *fərr*.

Iʾ. *ákal* 'to eat'. Perfect: *ákal*, *áklu*, *akáltu*. Imperfect: *ākəl* (1st pers. sg.) *tākəl*, *taklēn*; *kəl*.

I*w*. *wáqaʿ* 'to fall', *wázan* 'to weigh'. Imperfect: *áwqaʿ* (1st pers. sg.), *tə́wqaʿ* ~ *tūqaʿ*, *twəqʿēn*; *wqaʿ*. *áwzən* (1st pers. sg.), *tə́wzən* ~ *tūzən*, *twəznēn*; *wzən*.

I*y*. *yə́bas* 'to dry [intr.]'. Imperfect: *áybas* (1st pers. sg.), *tə́bas*, *tyəbsēn*; *ybas*.

II*w*/*y*. *xāf* 'to be scared, afraid'; *sāġ* 'to become'; *qām* 'to get up'. Perfect: *xāf*, *xə́ftu*. Imperfect: *txāf*, *txafēn*; *xāf*. *tsīġ*, *tsiġēn*; *sīġ*. *tqūm*, *tqumēn*; *qūm*.

III*y*. *báqa* 'to remain'; *qála* 'to fry'. Perfect: *báqa*, *báqu*, *bqētu*. Imperfect: *tə́bqa*, *təbqēn*; *bqā* (2nd pers. sg. masc.), *bqē* (2nd pers. sg. fem.), *bqō* (2nd pers. pl.). *tə́qli*, *təqlēn*; *qlī* (2nd pers. sg. masc.), *qlē* (2nd pers. sg. fem.), *qlō* (2nd pers. pl.).

II*w*/*y* and III*y*. *jā* 'to come'. The whole paradigm: Perfect: *jā*, *jət*, *jō*, *jīt*, *jīti*, *jītəm*, *jītu*, *jīna*. Imperfect: *áji*, *tə́ji*, *təjēn* ~ *djēn*, *yə́ji*, *tə́ji*, *nə́ji*, *təjōn* ~ *djōn*, *yjōn*. The imperative is formed from a different root: *tāl*, *tāli*, *tālu*.

3.3.3 Form II

Strong verbs. *názzal* 'to take down'. Perfect: *názzal*, *názlu*, *nəzzáltu*. Imperfect: *tnázzəl*, *tnəzlēn*, *názzəl*, *názli*.

III*y*. *xálla* 'to put'. Perfect: *xálla*, *xállu*, *xəllētu*. Imperfect: *txálli*, *txəllēn*; *xálli* (2nd pers. sg. masc., 2nd pers. sg. fem.), *xállu* (2nd pers. pl.).

Quadriradical verbs. *tárjam* 'to translate'. Perfect: *tárjam*, *tá(ə)jmu*, *tərjámtu*.

Imperfect: *ttárjəm*, *ttərəjmēn*; *tárjəm*, *tár(ə)jmi*.

3.3.4 Form III

Strong verbs. *ḥārab* 'to fight'. Perfect: *ḥārab*, *ḥārbu*, *ḥarábtu*. Imperfect: *tḥārəb*, *tḥarbēn*; *ḥārəb*, *ḥārbi*.

III*y*. *lāqa* 'to meet'. Perfect: *lāqa*, *lāqu*, *laqētu*. Imperfect: *tlāqi*, *tlaqēn*; *lāqi* (2nd pers. sg. masc., 2nd pers. sg. fem.), *lāqu* (2nd pers. pl.).

Form IIIa *bēġak* 'to bless'. Perfect *bēġak*, *bēġku*, *beġáktu* ~ *biġáktu*. Imperfect: *tbēġək*, *tbiġkēn*; *bēġək*, *bēġki*.

Form IIIb *sōlaf* 'to chat'. Perfect: *sōlaf*, *sōlfu*, *soláftu* ~ *suláftu*. Imperfect: *tsōləf*, *tsolfēn*; *sōləf*, *sōlfi*.

3.3.5 Form V

Strong verbs. *tkámmal* 'to be completed'. Perfect *tkámmal*, *tkámlu*, *tkəmmáltu*. Imperfect: *tətkámmal*, *təkəmlēn*; *tkámmal*, *tkámli*.

III*y*. *tġádda* 'to have a meal, lunch'. Perfect: *tġádda*, *tġáddu*, *tġəddētu*. Imperfect: *tətġádda*, *tətġəddēn*; *tġádda*, *tġáddi*.

Quadriradical verbs. *tšáqlab* 'to tumble, roll over'. Perfect: *tšáqlab*, *tšáq(ə)lbu*, *tšəqlábtu*. Imperfect: *tətšáqlab*, *tətšəqəlbēn*; *tšáqlab*, *tšáq(ə)lbi*.

3.3.6 Form VI

Strong verbs. *tqātal* 'to quarrel'. Perfect: *tqātal*, *tqātlu*, *tqatáltu*. Imperfect: *tətqātal*, *tətqatlēn*; *tqātal*, *tqātli*.

III*y*. *tbāha* 'to brag, pride oneself'. Perfect: *tbāha*, *tbāhu*, *tbahētu*. Imperfect: *tətbāha*, *tətbahēn*; *tbāha*, *tbāhi*.

Form VIa. *tnēšan* 'to be marked, become betrothed'. Perfect: *tnēšan*, *tnēšnu*, *tnišántu* ~ *tnešántu*. Imperfect: *tətnēšan*, *tətnešnēn*; *tnēšan*, *tnēšni*.

Form VIb *tdōxan* > *ddōxan* 'to feel dizzy'. Perfect: *ddōxan*, *ddōxnu*, *ddoxántu*. Imperfect: *təddōxan*, *təddoxnēn*; *ddōxan*, *ddōxni*.

3.3.7 Form VII

Strong verbs. *nkátab* 'to be written down, listed'. Perfect: *nkátab*, *nkátbu*, *nəktábtu*. Imperfect: *tənkátəb*, *tənkətbēn*; *nkátəb*, *nkátbi*.

Geminate. *ndaqq* 'to be ground, knocked'. Perfect: *ndaqq*, *ndəqqētu*. Imperfect: *təndáqq*, *təndəqqēn*, *ndaqq*.

I*w*. *nwálad* 'to be born'. Perfect: *nwálad*, *nwáldu*, *nəwládtu* ~ *nuwládtu* ~ *nūládtu*.

II*w*/*y*. *ndāġ* ‘to turn [intr.]’. Perfect: *ndāġ*, *ndaġētu*. Imperfect: *təndāġ*, *təndaġēn*; *ndāġ*.

III*y* *nqála* ‘to be fried’. Perfect: *nqála*, *nqálu*, *nəqlētu*. Imperfect: *tənqáli*, *tənqlēn*; *nqáli* (2nd pers. sg. masc., 2nd pers. sg. fem.), *nqálu* (2nd pers. pl.).

3.3.8 Form VIII

Strong verbs. *ftáham* ‘to understand’. *štáġal* ‘to work’. Perfect: *ftáham*, *ftáhmu*, *fəthámtu*. *štáġal*, *štáġlu*, *štġáltu*. Imperfect: *təftáhəm*, *təftəhmēn*; *ftáhəm*, *ftáhmi*.

Geminate. *štamm* ‘to smell’. Perfect: *štamm*, *štəmmētu*. Imperfect: *təštámm*, *təštəmmēn*; *štamm*.

I*w*. *ttáfaq* (< **wtafaq*) ‘to agree’. Perfect: *ttáfaq*, *ttáfqu*, *ttəfáqtu*. Imperfect: *təttáfəq*, *təttəfqēn*; *ttáfəq*, *ttáfqi*.

II*w*/*y*. *ʿtāz* ‘to need’. Perfect: *ʿtāz*, *ʿtazētu*. Imperfect: *təʿtāz*, *təʿtazēn*; *ʿtāz*.

III*y*. *štáka* ‘to sue’. Perfect: *štáka*, *štáku*, *štkētu*. Imperfect: *təštáki*, *təštkēn*; *štáki* (2nd pers. sg. masc., 2nd pers. sg. fem.), *štáku* (2nd pers. pl.).

3.3.9 Form IX

Strong verbs. *ṣfaġġ* ‘to turn yellow, become pale’. Perfect: *ṣfaġġ*, *ṣfəġġētu*. Imperfect: *təṣfáġġ*, *təṣfəġġēn*; *ṣfaġġ*.

3.3.10 Form X

Strong verbs. *stáʿjal* ‘to hurry’. Perfect: *stáʿjal*, *stáʿ(ə)jlu*, *stəʿjáltu*. Imperfect: *təstáʿjəl*, *təstəʿəjlēn*; *stáʿjəl*, *stáʿ(ə)jli*.

Geminate. *stʿadd* ‘to get ready, become proficient’. Perfect: *stʿadd*, *stʿəddētu*. Imperfect: *təstʿə́dd*, *təstʿəddēn*; *stʿədd*.

II*w*/*y*. *stġāḥ [ṣṭġāḥ]*. ‘to rest’. Perfect: *stġāḥ*, *stġaḥētu*. Imperfect: *təstġāḥ*, *təstġaḥēn; stġāḥ*.

II*w*/*y* and III*y*. *stáḥa* ‘to be ashamed’. Perfect: *stáḥa*, *stáḥu*, *stḥētu*. Imperfect: *təstáḥi*, *təstḥēn*; *stáḥi* (2nd pers. sg. masc., 2nd pers. sg. fem.) *stáḥu* (2nd pers. pl.).

III*y*. *stáhna* ‘to enjoy’. Perfect: *stáhna*, *stáhnu*, *stəhnētu*. Imperfect: *təstáhni*; *təstəhnēn*; *stáhni* (2nd pers. sg. masc., 2nd pers. sg. fem.), *stáhnu* (2nd pers. pl.).

3.4 *Particles*

3.4.1 Particles preceding the imperfect

qad ~ *qa* preceding the imperfect denotes the present tense. *qad* occurs before the 1st pers. sg., and *qa-* before the other persons: *qad-áktəb* ‘I am writing’, *qa-ykətbōn* ‘they are writing’.

ġaḥ preceding the imperfect denotes the future: *ġaḥ-áji* ‘I’ll come’. The same applies to *ssa* and *sa* (< *hassa*): *ssa-áji*, *sa-áji* ‘I’ll come’.

da preceding the 1st pers. of the imperfect has an optative function. It can convey a wish, a request, an urging: *da-nġūḥ* ‘let’s go!’.

xálli. A similar meaning is produced by the addition of *xalli* before the 1st and 3rd pers. of the imperfect: *xalli(y)ġūḥ* ‘let him go!’.

ma. Unstressed *ma* followed by the 2nd pers. of the imperfect expresses modality; when stressed, *ma* expresses negation. Cf. *ma-téji* ‘do come, come along, will you!’ vs. *má-teji* ‘you’ll not come’.

la. If *la* followed by the imperfect is stressed, the result is a negative imperative: *lá-təḥkīlu* ‘don’t tell him!’; when unstressed, the compound denotes doubt or apprehension: *axāf la-təḥkīlu* ‘I am afraid you may tell him’.

3.4.2 Demonstratives

hāḏa ‘this [masc.]’, *hāyi* ‘this [fem.]’, *haḏōli* ‘these’, *haḏāk* ‘that [masc.]’, *haḏīk* ‘that [fem.]’, *haḏōk* ‘those’, *hal* ‘this, these’. The consonant /h/ of these demonstratives is generally elided when it forms a cluster with a preceding preposition: *bhāḏa* > *bāḏa* ‘in this one [masc.]’, *mən haḏōli* > *mnaḏōli* ‘from these’. /l/ of *hal* is assimilated as the definite article (3.4.3.1): *hal-sana* > *has-sana* ‘this year’.

3.4.3 Definite and indefinite articles

3.4.3.1 The definite article

When the definite article *l-* precedes a dental, interdental, or palatal consonant (excluding the palatal /y/), it will generally assimilate completely to the following consonant. This occurs only before CV: *ṭ-ṭēġ* ‘the bird’, *č-čāy* ‘the tea’. When *l-* precedes CC and an anaptyctic intrudes after *l-*, there is no assimilation: *lə-ṭyūġ* ‘the birds’.

3.4.3.2 The indefinite article

faġad or *fadd* ‘one, some’, when unstressed, expresses indefiniteness, as opposed to the definite article: *fadd-yōm* ‘one day’, *fadd-wēḥəd* ‘a certain person’. When stressed *fadd* means ‘only’: *fádd-wēḥəd* ‘only one’.

3.4.4 Relative pronoun

ə̄lli, *lli*, *l-* ‘who, which’. *l-* is assimilated as the definite article: *j-jā* ‘(he) who came’.

3.4.5 Interrogatives

áškun is a free form for 'what?'. *aš* is a preposed form: *áš-asawwi* 'what shall I do?', *áš-əsmu* 'what's his name?', *ášlon* 'how?'. The variant *-ēš/-eš/-aš* is postposed to prepositions, forming, mostly, an inseparable compound: *bēš* 'how much? with what?', *lēš* 'why?', *ʿáleš* 'on what? what for?', etc. (Mansour 2001). *máni* 'who?', *l-máni* 'to whom?'. *háyyi* 'which [of several]?': *háyyi skámli* 'which chair?', *háyyi l-ákbaġ* 'who is the oldest?'. *wēn* 'where?'. *mnēn* 'whence?'. *ēmta* 'when?'.

3.4.6 Prepositions

mən 'from': *mə́nni* ~ *mmə́nni* 'from me'. *ʿənd* 'at, with'. *qbāl* 'facing'. *b-* 'in': *b-bētu* 'in his house', *bə-mkānu* 'in his place'; with suffix: *bī-*, *bíyyi* ~ *bīyi* 'in me', *bīna* 'in us'. *ʿála* 'on'; *ʿláyyi* 'on me', *ʿlēna* 'on us'. *wáya* 'with', *wəyyāna* 'with us'. *l-* 'to': *l-bētu* 'to his house', *l-wēn* 'where to'. *li-* 'until, up to': *li-wēn* 'up to where? how far?'. *li-qəddām* 'in advance'. The particle *māl* denotes possession, relation, etc.: *lə-ḥkíyyi mal-[ə]s-sáfaġ* 'the story of the journey', *ləmʿálləm māləm* 'their teacher'.

3.4.7 Conjunctions

íḏa 'if', *lō* 'if, or', *ya* 'or', *w* 'and'. *lākən* 'but'. *bass*, when unstressed, means 'but'; stressed *bass* means 'only'; cf. *bass-hṓni* 'but here ...' vs. *báss-hōni* 'only here'. *aṯāġi* 'it turns out, it seems'. For *báqa* as conjunction, see Mansour (1985).

3.4.8 Adverbs

time: *l-yōm* 'today', *(l)bōḥi* 'yesterday', *ġáda* tomorrow', *dāyman* 'always', *ábadan* 'never, not at all'.

place: *hōn*, *hōni* 'here', *wnī́ki* 'there', *fōq* 'above', *jáwwa* 'inside, under', *báġġa* 'outside'.

manner: *šwáyya* 'a little', *kṯīġ* 'a lot, very much', *bə lʿájal* 'quickly'.

4. LEXICON

4.1 *Characteristic words*

Jewish Baghdadi shares certain characteristic words with the other dialects of Baghdad, e.g. nouns: *bəzzūna* 'cat', *ḥə́nṭa* 'wheat' *tə́mman* 'rice [uncooked]', *ḥwās* 'clothes'. Verbs: *čállab* ~ *ččállab* 'to hang on to, cling to', *ṣáffaṭ* 'to arrange in its place', *gádda* 'to beg for alms'. Particles: *áku* 'there is ~ are', *māku* 'there is no', *məšwāġ* 'short time, moment'; *kə́lləš* 'very'.

4.2 *Loanwords from Persian and Turkish*

A common feature of the various dialects of Baghdad is the large number of loanwords from Persian and Turkish, e.g. everyday matters: *azaxāna* 'pharmacy' (Turkish *eczahane*); *čāra* 'cure, remedy' (Turkish *çare*, Persian *čare*); *čarpāya* ~ *čarpāyi* 'bedstead' (Persian *čarpaye*); *čarx* 'wheel' (Persian *čarx*); *čəngāl* 'hook' (Persian *čāngal*, Turkish *çengel*); *čāġak* 'quarter' (Persian *čarāk*, Turkish *çeyrek*); *mēz* 'table' (Persian *miz*, *mēz*); *mīwa* 'fruit' (Persian *mive*, *mēve*). Particles: *xōš* 'good', (Persian *xoš*); *gárag* 'probably' (Turkish *gerek*); *hīč* 'nothing' (Persian *hič*); *ham*, *hámzed* 'also, too' (Persian *ham*, Turkish *hem*). The suffix *-či* (< Turkish) signifies a person's profession or habitual activity: *azāči* 'pharmacist', *bəstánči* 'gardener', *pōsṭači* 'postman', *jāmči* 'glazier', *kundárči* 'shoemaker, cobbler', etc. (also → Persian loanwords; → Turkish loanwords).

4.3 *Loanwords from Hebrew and Judaeo-Aramaic*

Of the non-Arabic lexical items, it is the Hebrew and the Judaeo-Aramaic elements that distinguish Jewish Baghdadi from other Baghdad dialects. Hebrew words mainly occur in connection with Jewish festivals, religious practice and ritual. However, they occur in secular matters, too: *məzzāl* 'luck', *kabōd* 'honor, respect', *səkkāna* 'danger', *afə́llu* 'nevertheless'.

In some cases a new word is coined. From the Judaeo-Aramaic phrase: *haššata haxa* ('this year [we are] here'), in the Passover Haggada, Jewish Baghdadi has derived the noun *šəttāxa* 'the Passover ceremony', and the verb *šáttax* 'to celebrate the Passover'. Compounds of Arabic and Hebrew words are also found, e.g. *ṭġāb w-ʿafār* 'dust, earth' (cf. Literary Arabic *turāb* and Hebrew *ʿafar*).

4.4 *Loanwords from European languages*

Loanwords from French and English were introduced mainly under the influence of modern Jewish schools established in Baghdad since 1864. English also entered through relations with India and more so after the British conquest of Iraq.

BIBLIOGRAPHICAL REFERENCES

Avishur, Yitzhak. 2001. *ha-ʿIḇrīt še-ba-ʿarabit ha-yəhudīt* [Hebrew elements in Judaeo-Arabic]. Tel-Aviv and Jaffa: Archaeological Center.

Ben-Yaacob, Abraham. 1985. *ʿIbrīt wa-ʾaramīt bi-lšōn yəhudey Babel* [Hebrew and Aramaic in the language of the Jews of Iraq]. Jerusalem: Ben-Zvi Institute.

Blanc, Haim. 1964a. *Communal dialects in Baghdad.* Cambridge, Mass.: Harvard University.

——. 1964b. “Notes on the literary idiom of the Baghdadi Jews”. *For Max Weinreich on his seventieth birthday*, ed. L.S. Dawidowicz a.o., 18–30. The Hague: Mouton.

Jastrow, Otto. 1990. *Der arabische Dialect der Juden von ʿAqra und Arbīl.* Wiesbaden: O. Harrassowitz.

Mansour, Jacob. 1955–1956. *Hagiyat ha-reš bə-fī yəhudey Bagdad* [The pronunciation of *r* by the Jews of Baghdad]. *Lešonenu* 20.47–49.

——. 1974–1983. *ha-ʿArabit ha-yəhudit šel Bagdad* [The Judaeo-Arabic dialect of Baghdad]. I. *Studies in phonology and morphology.* II. *Texts.* III. *Dictionary*, part one. Haifa: University of Haifa.

——. 1985. “The particle *baqa* in the Judaeo-Arabic dialect of Baghdad”. *Zeitschrift für Arabische Linguistik* 14.62–75.

——. 1991. *The Jewish Baghdadi dialect: Studies and texts in the Judaeo-Arabic dialect of Baghdad.* Or-Yehuda: The Babylonian Jewry Heritage Center.

——. 2001. “The particles *ʾašľeš* and *ʾašu* in the Jewish-Arabic of Baghdad”. *Zeitschrift für Arabische Linguistik* 39.84-89.

Meʾiri, Yehoshua. 1997. *ʿAl naharōt Babel: Otsar ha-pitgamim šel Yəhudey Babel* [On the rivers of Babylon: Treasury of proverbs of the Iraqi Jews]. Jerusalem: R. Mass.

Sassoon, David Solomon. 1949. *A history of the Jews in Baghdad.* Letchworth: S.D. Sassoon.

Yona-Swery, Gila and Rahamim Rajwan. 1995. *Millōn ʾimrōt wu-mšalīm šel lahag yəhudey Babel* [Dictionary of Iraqi Judaeo-Arabic dialect]. Jerusalem: Association for Jewish Academics from Iraq.

Jacob Mansour (University of Haifa)

Bahraini Arabic

1. GENERAL

1.1 *Area*

Bahrain is an archipelago (552 square kilometers) halfway between the head of the Gulf and the Straits of Hormuz, 20 kilometers off the coast of Saudi Arabia. In 1999 the population was approximately 660,000, of whom 402,000 were Bahrain nationals. Its dialects are of two types: those of the so-called ʿArab (lit. ‘Arabs’, henceforth ‘A’), and those of the Baḥārna (lit. ‘Bahrain-dwellers’, henceforth ‘B’). The A community, which includes the ruling family, and is Sunni, traces its origin to Najd. The B, ‘Twelver’ Shiʿi, and outnumbering the A by about two to one, regards itself as the descendants of the ancient population of the area.

1.2 *Society*

Before the 1970s, there were sharp social divisions between the communities. The A community lived in Muḥarraq, in a few coastal settlements, and near the ruler’s palace; the B in Manāma and about 60 villages. Recently, these divisions have become somewhat blurred.

1.3 *Regional context*

The A dialect is Najdi in origin and Bedouin in type, and similar to that of other Gulf states (Johnstone 1967: *passim*). The rural B dialects bear a strong resemblance to the sedentary dialects of northern Oman (→ Omani Arabic). The B dialect of Manāma and the nearby B villages is similar to that of Baḥārna in eastern Saudi Arabia (Smeaton 1973).

Where Bahraini speech is imitated in the media (plays, newspaper cartoons) it is usually the A dialect which is used. A tradition of dialectal poetry exists. The poets are generally from the A community and use its dialect, or a ‘poeticized’ form thereof.

1.4 *Historical evidence*

There is almost no textual evidence for the Bahraini dialects earlier than the 1930s. Earlier documents in the state archives occasionally betray dialect influences, but inconsistently. Evidence of contact with many languages over centuries, if not millennia, is abundant in the extensive borrowings from Persian, Turkish, Hindi/Urdu, Portuguese, and English in the modern vocabulary, and in toponyms. There is increasing evidence of a substrate vocabulary which is of Semitic but possibly non-Arab origin (Holes 2001:xxix–xlii; 2002).

1.5 *State of research*

In recent years there have been a large number of studies of the Bahraini dialects. A glossary has recently been published (Holes 2001), and a complete description of the dialects will appear soon (Holes *forthcoming*). The studies listed here in the Bibliographical References are concerned wholly with the Bahraini dialects; for those concerned with Gulf dialects more generally but relevant to Bahrain (→ Gulf Arabic).

NB: in what follows, the language level described is that of *uneducated* (often *illiterate*) speakers.

2. Linguistic description

2.1 *Phonology*

2.1.1 Inventory

2.1.1.1 Consonants

Plosives: *(p) b*, *d*, *ṭ*, *ḍ*, *k*, *g*, *q*, *ʾ*
Affricates: *č*, *j*
Fricatives: *f*, *ṯ*, *ḏ*, *ḏ̣*, *x*, *ġ*, *ḥ*, *ʿ*, *h*
Sibilants: *s*, *ṣ*, *z*, *š*
Laterals and vibrants: *l*, *r*
Nasals: *m*, *n*
Semivowels/glides: *w*, *y*

Some speakers have /p/ in borrowings, e.g. *panka* 'ceiling fan', which others pronounce with /b/. /č/ and /g/ are common in borrowings, e.g. *čāra* 'ruse', *gāri* 'donkey cart', and are indistinguishable phonetically from /č/ and /g/ that have resulted from internal phonological changes.

2.1.1.1.1 The A dialects

These have preserved the Classical Arabic interdentals, /ṯ, ḏ, ḏ̣/ and words with etymological /ḍ/ are always pronounced with /ḏ̣/, e.g. *abyaḏ̣* 'white'. Normally /g/ < Classical Arabic /q/, which in front vowel environments was fronted and affricated to /j/, e.g. *jīma* 'value', *mjābil* 'opposite'. In similar environments, Classical Arabic /k/ was fronted and affricated to /č/, e.g. *čibīr* 'great, old' *ḥači* 'talk, gossip'. But these developments did not operate categorically: *jidir* 'cooking pot', but *gidar* 'he was able', *čibd* 'liver', but *kitab* 'he wrote', and occasionally /č/ is heard in back-vowel environments, e.g. *smūč* 'fishes'. In a few cases, the result of the partial operation of the rule was a minimal pair, e.g. *kitab* 'he wrote', but *čitab* 'gold pendants attached to women's plaits'. In a number of verbs, the stop/affricate contrast has been morphologized, e.g. *saggam/yisajjim* 'to give an advance payment [pearling]'. /y/ < Classical Arabic /j/ categorically, but /j/ < /g/ did not usually undergo this change, except in a few items such as *yassam* 'to divide up' < *jassam* < *gassam*. /j/ in foreign borrowings is usually preserved, e.g. *jūti* (Urdu) 'shoes', *jām* (Persian) 'pane of glass'. Classical Arabic /ġ/ > [ɢ] or [q], e.g. [qe:r] for *ġēr* 'other'. Modern Standard Arabic /q/ in neologisms is often pronounced [ɣ] or [ɢ] e.g. [taɣaddum] ~ [taɢaddum] 'progress', its allophones identical with those of dialectal /ġ/. Educated speakers have a few /q/-/j/ minimal pairs within the same root as a result of borrowing from Modern Standard Arabic, e.g. *ytijaddam* 'he comes forward', *yatqaddam* 'he is making progress [in an abstract sense]'. Classical Arabic /ʾ/ disappeared initially, e.g. *kal* 'he ate', or was replaced by /w/, e.g. *winsa* 'fun' or /y/, e.g. *yaddam* 'he presented a guest with food'; finally, it either disappeared, e.g. *ašya* 'things' or was replaced by /w/ e.g. *ḏ̣aww* 'fire' or by /y/, e.g. *simi* 'sky'; medially it was replaced by vowel length, e.g. *yīt* 'I came', *rās* 'head', and after /ā/ by /y/, e.g. *šāyil* 'removing'. /ʾ/ now occurs even in uneducated speech, in a few items which are derived from Modern Standard Arabic, e.g. *yisʾal* 'he asks' (alongside the dialect form *ysāyil*). /l/ and /r/ have velarized allophones in some words, especially when a labial is present, when the whole word may become velarized, e.g. [g̶aƀiƚ] 'before'.

2.1.1.1.2 The B dialects

These all have /f/ for /ṯ/, /d/ for /ḏ/, and /ḍ/ for both Classical Arabic /ḍ/ and /ḏ̣/. Beyond that, they can be divided into three groups. For Group 1, the main rural group, Classical Arabic /q/ in most words is /k/, and Classical Arabic /k/ > /č/ unconditionally e.g. *šōč* 'palm-tree thorns', *ḍuḥč* 'laughter' where the A dialects have /k/. In these dialects, Classical Arabic /j/ is [dʒ], but in a few villages /j/ and /č/ are [dɟ], [tç]. Group 2, the main urban B dialect of Manāma, is like the A dialects in having /g/ not /k/ < Classical Arabic /q/, but does not have /j/ < /g/, and has less affrication of Classical Arabic /k/. It has /j/ for Classical Arabic /j/. Group 3 (three northern villages and the Manāma quarter Rās Rummān) differs from Group 2 in having, like the A dialects, /y/ < Classical Arabic /j/.

2.1.1.2 Vowels and diphthongs

2.1.1.2.1 General

All dialects have three short vowels, /a, i, u/, and most have five long /ā, ē, ī, ō, ū/, but some B Group 2 dialects have preserved the Classical Arabic diphthongs /ay, aw/. Unstressed /i/ in non-final open syllable is often deleted (see 2.1.3). /ē/ and /ō/ are sometimes shortened in particular words, but the resulting short mid-vowels are not phonemic. The distribution and quality of the vowels differ between the A and B dialects.

2.1.1.2.2 A dialects

i. Distribution and quality of short vowels: /i/ occurs to the exclusion of /a/ in open, non-final syllables, except in the contiguity of guttural consonants, /x, ġ, ḥ, ʿ, h/, or where the following consonant is /l/, /n/, or /r/ when at the same time the vowel of the following syllable is /a/ or /ā/ (Johnstone 1967:27). Thus *kitab* 'he wrote', but *barad* 'it got cold', *tkallam* 'he spoke' but *tkallimaw* 'they spoke'. The labials have a rounding and backing effect on /i/ (> /u/) whether /i/ is original or < /a/, especially if a velar, emphatic, or /l/ or /r/ are also present e.g. *mukān* 'place', *buṣal* 'onions', *šrubat* 'she drank', but *xšiba* 'piece of wood', where /b/ is present but none of the other factors. /i/ and /u/ are virtually in complementary distribution in open syllables, but not in closed ones. Final /ā/ and /āʾ/ are shortened, and when not in contiguity with a guttural, emphatic, /l/ or /r/, and when preceded by an open syllable, raised, e.g. *niši* 'starch', *simi* 'sky' (an exception is *māy* 'water') but *ḥamra* 'red [fem.]', *ramḍa* 'hot ground'. Some older A speakers raise final /a/ in these same environments, e.g. *bridi* 'hail' (< *barada*), and also when the preceding syllable is closed, e.g. *gumni* 'we got up'.

/a/ is realized: as [æ] or [ɛ] where gutturals (excluding /h/) and emphatics are absent, e.g. [hæli] 'my family', [dɛʃʃ] 'he entered'; as [a] in guttural environments, e.g. [baʕad] 'after' [xalleːt] 'I/you allowed'; as [ɑ] with an emphatic, and often with labials, e.g. [ŧɑƚƚ] 'mist', [xɑm̶ɑɍ] 'alcohol'. Medial /i/ is retracted, e.g. [bɪnt] 'girl'; in final position it is closer and more front, e.g. [rɪħti] 'you [fem.] went'; with emphatics it is lowered, e.g. [yɪɍŧʌn̶] 'he gabbles'. /u/ is back and rounded, e.g. [ʃrʊbæt] 'she drank'.

ii. Long vowels: generally /ā/ has a very backed and rounded quality in any phonetic environment. This is particularly noticeable among women, e.g. [hɒːði] or even [hɔːði] 'this', [ʔɒːna] 'I'. /ī/ is a close, front vowel, but with the emphatics is more centralized, e.g. [bɨːs̶] 'keel of a boat'. /ē/ and /ō/ correspond to the Classical Arabic diphthongs /ay, aw/ and occur medially, e.g. *sēf* 'sword', *bōg* 'theft'. Where /aw/ or /āw/ occurs medially in multisyllabic forms, it is often reduced to /ā/, e.g. *mithāš* < *mithāwiš* 'arguing', *hān* < *hāwin* 'mortar'. In verb forms, final /ay/ and /aw/ > /ē/ and /ō/ when suffixed, e.g. *gālaw* 'they said', *gālōli* 'they told me', *fitḥay* 'open [fem.]!', *fitḥēh* 'open [fem.] it!'. /ō/ occurs finally in a few words, now obsolete, which appertain to seafaring and traditional culture, e.g. *ḥalwāyō* 'jack pomfret [type of fish]', *ndēndō* 'type of dance with drum accompaniment'. /ō/ is also suffixed to certain personal names as a hypocoristic, e.g. *maryamō*, *xalīlō*, *ḥusnō*.

2.1.1.2.3 The B dialects

They do not have the restrictions on /a/ in open syllable of the A dialects and nor do the labials have the same effects. Final /ā/ and /āʾ/ are similarly shortened and raised, though not to the same height. In other types of form in which Classical Arabic final /āʾ/ occurred, /ʾ/ has been replaced by /y/, e.g. *bannāy* 'builder' (cf. Omani sedentary dialects). Except in emphatic environments, /ā/ in the B dialects is always a front vowel; there is none of the general backing and rounding typical of many A speakers' pronunciation. There is pronounced and widespread → *ʾimāla* in some rural B dialects, e.g. *ktībi* < *kitāba* 'writing', *mi* < *māʾ* 'water' (A dialects have *māy*).

2.1.1.3 Syllable types

2.1.1.3.1 Possible syllable types

Cv:	*kitab* 'he wrote' (Cv-CvC)
CvC:	*čilma* 'word' (CvC-Cv)
CvCC:	*maġarb* (A dialects) 'evening' (Cv-CvCC); *ḍarabk* (B dialects) 'he hit you' (Cv-CvCC)
Cv̄:	*sōlaf* 'he chatted' (Cv̄CvC)
Cv̄C:	*bāg* 'he stole'

The above are the basic types. Cv̄CC also occurs, but only in one type of form (the active participle of geminate verbs), e.g. *ḥāṭṭ* 'putting'. However, the following also arise:

CCv:	*drisat* 'she studied', *ghawa* 'coffee' (CCv-CvC) (A dialects)
CCv̄C:	*smūč* 'fishes', *smīt* 'cement' (borrowing)
CCvCC:	*fhimt* 'I understand', *trinj* 'citron' (borrowing)

2.1.1.3.2 Distribution of syllable types

Excluding borrowings, the last three of the types listed arise via vowel elision and epenthesis rules. CCv (A dialects) is always word-initial: CvCvCv(C) is reduced by elision of the first vowel (see 2.1.3.1). A prosthetic vowel is inserted before the resulting consonant cluster after a word ending in a consonant, e.g. *l-ixšiba* 'the lump of wood'. CvCC is always word-final, and, apart from in monosyllabic words like *šarg* 'east', arises either as a product of the → *gahawa* syndrome (in the A dialects – see 2.1.2.4), or via suffixation (B dialects), as in the examples. With a consonant initial suffix, Cv̄C forms are treated differently in the A dialects depending on whether they are verbs or nouns: *ṣādəna* ‘it hit us’ versus *bētna* ‘our house’. This distinction is absent in the B dialects, which have non-epenthesized forms in both cases.

2.1.1.4 Consonant clusters

The treatment of CCC clusters in the A dialects has also been partly morphologized. Where the cluster is a consequence of the suffixation of doubled verbs, most speakers insert an epenthetic schwa, e.g. *šaggəha* ‘he tore it’ (cf. *ṣādəna* ‘it hit us’ for the same speakers) but reduce the cluster in nouns, e.g. *ḥagna* ‘our right, for us’ (< *ḥagg* + *na*). Otherwise, CCC clusters are stable, e.g. *bintkum* ‘your daughter’, *šiftkum* ‘I saw you’, though some A speakers, unpredictably, have occasional forms of the *binitkum*, *šifitkum* type. The only non-verbal CCC clusters which A speakers normally epenthesize are *ʿind* and *kill*, e.g. *ʿindəkum*, *killəhum*.

In the B dialects, CCC clusters are generally stable, but are reduced, as in the A dialects, in doubled nominal forms: *ḥagg* + *na* → *ḥagna*. Exceptions are again *ʿind* and *kill*, but here the B dialects reduce the cluster (= *ʿidkum*, *kilhum*) rather than epenthesize.

Reduction in one high-frequency phrase is universal for all: *git* < *gilt*, ‘I/you [masc.] said’ in phrases like *git lik/lah* ‘I told you/him’.

Initial CC clusters occur in all dialects but are differently distributed (see above).

The treatment of non-doubled final clusters in words of the structure CvCC in Classical Arabic depends on the preceding vowel and the consonants:

2.1.1.4.1 Classical Arabic CaCC

i. C_2 is /l, n, r/: the form is stable, e.g. *galb* ‘heart’, *ḏanb* (B *danb*) ‘sin’, *warč* ‘thigh’.

ii. C_2 is a guttural: CaCaC is normal, e.g. *baḥar* ‘sea’, *šahar* ‘month’.

iii. in other cases, the A and B dialects differ: the A have CaCvC, in which the v is usually /i/ but in some words /a/; the B retain CaCC, e.g. A *ḥabil* ‘rope’, *ṭabix* ‘cooking’ *xamar* ‘alcohol’, B *ḥabl*, *ṭabx*, *xamr*. There are a few exceptions, e.g. (A) *wagt/wakt* ‘time, weather, climate’.

2.1.1.4.2 Classical Arabic CiCC and CuCC

i. C_2 is /l, n, r/: the form is stable, e.g. *ʿilč* ‘chewing gum’, *bunk* (some B *binč*) ‘essence’, *ʿirs* ‘marriage’.

ii. C_3 is /l, n, r/: CvCvC is normal, with vowel harmony, e.g. *šuġul* ‘work’, *ḏ̣uhur* (B dialects *ḍuhur*) ‘noon’, *dihin* ‘oil, fat’, *čiṯir* (B villages *kufur*) ‘amount’.

iii. in other cases, the A dialects have CvCvC, again with vowel harmony, the B CvCC, e.g. A *xubuz* ‘bread’, *bišit* ‘man’s cloak’, B *xubz*, *bišt*, etc. There are odd exceptions, e.g. (A) *rizg* ‘sustenance’, *ṣibḥ* ‘morning’.

The initial v of CvCvC forms that result from (ii) is deleted by rural B speakers when such forms are preceded by the definite article, and a shwa inserted, e.g. *dihin* ‘oil’ ⇒ *lədhin* ‘the oil’, *ḍuhur* ‘noon [-time]’ ⇒ *ləḍhur* ‘the noon [-time]’ (cf. Omani sedentary dialects). A equivalents are *iddihin*, *iḏ̣ḏ̣uhur*.

In all dialects, any CvCvC derived via the above rules reverts to CvCC- with a vowel-initial suffix.

2.1.1.5 Stress

In the A dialects, the rule is: stress the last syllable of a polysyllabic word if long (Cv̄C, CvCC), the second if there are two long syllables; otherwise stress the penultimate. This applies *after* the application of rules that specify syllable structure (2.1.1.3). In the B dialects, Cv-Cv-Cv forms regularly occur, and in such cases the antepenultimate is stressed. Stress is non-distinctive.

2.1.2 Phonotactics

2.1.2.1 Assimilation

The following are major loci:

the definite article: sun letters and moon letters behave as in Classical Arabic;

the *b-* verb prefix > *m-* before 1st pers. pl. imperfects in the B rural dialects, e.g. *minrūḥ* ‘we’ll go’ (cf. Omani sedentary dialects);

/l/ in the imperative *xall* ‘let’ > /n/ before the *-ni* and *-na* suffixes, e.g. *xanna* ‘let’s . . .’ For some speakers, this assimilation is general, e.g. *šṭaġanna* ‘we worked’;

/ġ/ > /b/ in the imperfect forms of the verb *baġa* ‘to want, need’ e.g. *abbi*, *tabbi*, etc. (A speakers); *abba*, *tubba*, etc. (some B speakers);

/h/ in the 3rd suffixes *ha* and *hum* is assimilated to the *t* of the 3rd pers. sg. fem. perfect verb and other feminine forms, e.g. *šrubatta* 'she drank it [fem.]', *rgubattum* 'their neck' (A speakers);

/t/ in *ti-* verbal prefixes of various kinds is assimilated to /t, ṭ, ṯ, d, ḏ, ḍ, ḏ̣, č, j, s, ṣ, š/ as a consequence of the deletion of unstressed /i/ in open syllable (see 2.1.3), e.g. *(i)ṣṣīr* 'she becomes', *(i)ṭṭawwar* 'it developed', *(i)ččiddūn* 'you work for your money'.

2.1.2.2 Dissimilation

There are a few cases, such as *sajara* < *šajara* 'tree, bush' (B village dialects), *xast* < *xass* 'lettuce' (all speakers). Some B speakers have dissimilated the first /b/ in assimilated forms of the *abba*, *tubba* type (see 2.1.2.1]) to /m/ ⇒ *amba*, *tumba*, *yumba*, etc. for 'I /you/he want(s)'. Certain quadriliteral verbs may have arisen historically via dissimilation, e.g. *karfas* 'to knock to the ground' < *kaffas* 'to crease, bend over'.

2.1.2.3 Metathesis

Examples are *nāwas* and *wānas* ~ *ānas* 'to keep someone company', *jawāz* and *zawāj* 'marriage', *raṣaʿ* and *ʿaṣar* 'to squeeze', *xamaš* and *šamax* 'to scratch, snatch, grab'.

2.1.2.4 *gahawa* syndrome

The → *gahawa* syndrome (A dialects only) is the deletion of /a/ in CaC non-final syllables where C_2 is a guttural and epenthesis of /a/ after C_2, e.g. *nxala* 'palm-tree', *yiʿarf* 'he knows'.

2.1.2.5 Conditioned *ʾimāla*

See 2.1.1.2.3

2.1.2.6 Spread of velarization

Velarization may spread, especially to /l/, /r/, /b/, /m/, e.g. *ṣabi* 'boy' (= [s̴ɒb̴i]), *ṭabil* 'drum' (= [t̴ɒb̴ɪɫ]), *ṣabur* 'patience' (= [s̴ɒb̴ʊr̴]). In a few roots, an originally non-emphatic consonant has become velarized, and velarization has then spread, e.g. (B dialects) *ḍakar* 'to remember, mention' (= [d̴ɑkɑr̴]). Combinations of velars, labials, and /l/ or /r/ may also become velarized when an emphatic consonant is not present, e.g. *txammar* 'to ferment' (= [t̴xɑm̴m̴ɑr]), *xtarab* 'to go rotten' (= [xt̴ɑr̴ɑb̴]) (B examples).

2.1.3 Morphophonology

2.1.3.1 Elision of vowels

/i/ (and /u/ in labial environments) in unstressed non-final open syllable is deleted, and a prosthetic vowel /i/ is inserted if the deleted vowel is in the first syllable and a consonant-final word precedes. Medially doubled consonants are reduced, e.g. *(i)tʿallim* 'you [masc.] teach' but *(i)tʿalmūn* 'you [pl.] teach', *labbisaw* 'they dressed' but, especially in rapid speech, *labsōha* 'they dressed her'. In the A dialects, /a/ is deleted in the initial syllable of CvCvC (v/v̄) strings, e.g. *šarab* 'he drank'→ *išrubah* 'he drank it'. In the B dialects, the equivalent forms are of the CvCvCvC or CvCCvC-type (see 2.2.6.1.)

2.1.3.2 Insertion of vowels

See 2.1.3.1 for prosthetic /i/. See 2.1.1.4 for epenthetic vowels, and 2.1.2.4 for the *gahawa* syndrome.

2.1.3.3 Shortening and lengthening of vowels

Where two or more long vowels occur in a word, the unstressed first (and second, if there is one) long vowel may be shortened, especially in rapid speech, e.g. *jālbūt* 'type of small boat' = [jalbʊːt], *hāḏēlēn* 'these' [haðɛleːn]. Final vowels that are short in unsuffixed form are lengthened and stressed when suffixed, e.g. *yxalli* 'he lets', *yxallīni* 'he lets me'. In some B urban dialects, the *-ah* 3rd pers. sg. masc. suffix is lengthened and stressed when attached to perfect verbs, e.g. *ḍirbāh* 'he hit him'. However, this has also been recorded with the feminine and plural suffixes, e.g. *šaggāha* 'he cut it open', *ṣaffāhum* 'he lined them up', and may be better explained as an analogical extension of the rule which lengthens the final vowel of III*y* perfect verbs on suffixation, cf. *xalla* 'he let', *xallāh*, *xallāha* 'he let him/her'.

2.1.3.4 Clitics

There are a number of clitics:

a question particle *ə*, after vowels *hə* or *yə*, suffixed to the questioned word or phrase, e.g. *awlādiš imšattatīn-ə* 'have your children been split up?' (B dialects only; cf. Omani sedentary dialects).

deictic *ha* prefixed to the definite article, e.g. *halbēt* 'this house' (all dialects).

b- prefix indicating proximate intent, e.g. *b-inrūḥ bāčir* ‘we’ll go tomorrow’ (all dialects, but B dialects *minrūḥ*, as with some Omani sedentaries).

d- prefix with imperatives as an exhortative, e.g. *d-itkallami*, *gūli šay* ‘speak, why don’t you? say something!’ and also for continuous or habitual aspect, e.g. *kil mā bēn ṯalāṯat šuhūr d-yišrabha* ‘he would drink it every three months’ (both uses in some B rural dialects only).

š- ‘what?’, e.g. *š-gilt*? ‘what did you say?’ *š-ḥaggah*? ‘what for?’ (mainly A dialects).

presentative *k(a)-*, e.g. *ka-āna yāya* ‘I’m just coming!’ (A dialects only).

taw(w) ‘just’ (= ‘recently’) is arguably a clitic element, as it only occurs with pronoun suffixes, e.g. *naymīn*, *taw-na naymīn* ‘we’d gone to bed, just gone to bed’ (all dialects).

-in, a vestige of Classical Arabic → *tanwīn* in certain phrase types, e.g. *bint-in zēna* ‘a nice girl’, *arāḍin šādda* ‘soils which hold together’ (see 2.2.3.2) (all dialects).

mā, inserted between the repetition of a word to indicate ‘and suchlike’, e.g. *fār mā-fār* ‘rats and things like that’. A shortened version, in which *m* is substituted for the first consonant of the noun, is also used, e.g. *xalāgīn malāgīn* ‘rags and suchlike’ (all dialects).

2.1.3.5 Construct state

This is less used because of the development of an analytic genitive (2.2.3.3). Examples: *yūniyyat ʿēš* ‘a sack of rice’, *rōḥat issīf* ‘going to the seashore’. Plurals and duals that occur as head noun normally retain their final *n*, e.g. *muwaḏ̣ḏ̣afīn ilḥukūma* ‘government employees’, though some A speakers have dual forms of the type *īḏē lmara* ‘the woman’s hands’. (In suffixed forms, the Manāma A dialects generally have *īdēnik* ~ *yadēnik*-type forms.)

2.1.3.6 Suffixation

i. In the A dialects, the /h/ of the *ha* and *hum* suffixes assimilates to the suffix /t/ of 3rd pers. sg. fem. perfect forms (2.1.2.1). When vowel-initial suffixes are added, they are resyllabified, e.g. *išrúbat* ‘she drank’ + *ah* ⇒ *išrúbtah* ‘she drank it [masc.]’. In all suffixed plural forms, the final *-aw* becomes ō and is stressed, e.g. *šrúbaw* ‘they drank’ + *ah* ⇒ *išrubṓ'h* they drank it’. In the B dialects there is no consonant assimilation or resyllabification, i.e. forms are of the *šarabatha* ~ *širbatha* and *šarabōha* ~ *širbōha*-type

ii. In the A dialects, active partiple + suffix combine as in Table 1:

Table 1. Active participle with suffix (A dialects)

	vowel-initial suffiix *(ah)*	consonant-initial suffix *(ha)*
masc.	*kātbah*	*kātibha* or *kātbaha*
fem.	*kātbitah*	*kātbatta*
pl.	*kātbīnah*	*kātbīnha*

kātibha-type forms are supplanting *kātbaha* type forms. A similar process is occurring with III*y* and II*w/y* verbs, so the older type of masculine forms *bānyah* ‘X [masc.] has built it’ are being replaced by *bānīh*, and *čāyfaha* ‘X [masc.] has seen it’ by *čāyifha*. The feminine forms are as in the strong verb, *bānyitah*, *čāyfitah*, etc.

In the B dialects, there is an obligatory *-inn* infix between the active participle and suffix (Table 2).

Table 2. Active participle with suffix (B dialects)

masc.	*kātbinnah*	*kātbinha*
fem.	*kātbatinnah*	*kātbatinha*
pl.	*kātbīnah*	*kātbīnha*

III*y* and II*w/y* verbs roots follow the same pattern, e.g. *bāninnah*, *bāninha*, etc. Where a participle has nominal meaning, it lacks the infix, so: *huwa mʿalliminnah* ‘he has taught him’ but *huwa mʿallimah* ‘he is his teacher’ (cf. Omani sedentary dialects).

iii. Communal differences in the shape of suffixed pronouns, in combination with differences in preferred syllable structure (2.1.1.3), generate contrasting forms in common types of verb and noun phrase. The B dialects have consonant-initial 2nd pers. sg. pronoun suffixes, the A dialects have vowel-initial. The A dialects elide the first vowel in CvCvCv forms, the B dialects maintain it (Table 3):

Table 3. Pronoun suffixes 2nd pers. sg.

A dialects	B urban	B villages	
ḏ̣rubič	*ḍarabk*	*ḍarabč*	‘he hit you [masc.]
arayḥik	*arayyiḥk*	*arayyiḥč*	‘I give you [masc.] rest’
yiddatič	*jaddatš*	*jiddatš*	‘your [fem.] grandmother’

The treatment of CvCvC nouns in the A dialects is somewhat unpredictable. Some ‘core’ nouns such as *yimal* ‘camel’ become *ymalik* ‘your camel’, but others, seemingly neologisms, do not follow this rule. Contrast *ḥarasah* ‘his defence force’, a neologism, with *ḥrisah* ‘he defended him’, both < *ḥaras* + *ah*.

2.2 *Morphology*

2.2.1 Pronouns

2.2.1.1 Personal independent pronouns

Table 4. Personal independent pronouns

	A dialects	B dialects
1st sg.	*ana/āna*	*ana* (f. *ani* in villages)
2nd sg. m.	*inta*	*inta*
2nd sg. f.	*inti/intay*	*intīn/intīna*
3rd sg. m.	*huwwa/əhwə/əhuwwə*	*hu/huwwa*
3rd sg. f.	*hiyya/əhyə/əhiyyə*	*hi/hiyya*
1st pl.	*iḥna*	*iḥna*
2nd pl.	*intu/intaw*	*intūn/intūna*
3rd pl.	*hum/əhmə/əhummə*	*hum/humma*

The B dialects 1st feminine form *ani* (also used in Iraq) is mainly heard in women’s speech.

2.2.1.2 Possessive/object suffixes

Table 5. Possessive/object suffixes

	A dialects	B city	B villages
1st sg.	*-i* (poss) *-ni* (obj)	*-i*, *-ni*	*-i*, *-ni*
2nd sg. m.	*-ik*	*-k*	*-č*
2nd sg. f.	*-ič*	*-š*	*-š*
3rd sg. m.	*-ah*	*-ah*	*-uh/-ah*
3rd sg. f.	*-ha*	*-ha*	*-ha*
1st pl.	*-na*	*-na*	*-na*
2nd pl.	*-kum*	*-kum*	*-kim/-čim*
3rd pl.	*-hum*	*-hum*	*-hum/-him*

There is much geographically-based allomorphy in the 1st person with certain prepositions: e.g. with *l-*: *lī*, *liyyi*, *liyya*; with *ila*: *ili*, *iliyyi*, *ilayya*; with *fī*: *fiyyi*, *fiyya*, *fīni*

2.2.1.3 Indirect object suffixes

With some verbs, the indirect object is suffixed directly to the verb and the direct object carried by the particle, e.g. *ʿaṭni iyyāh* ‘give it to me’; in most cases of such constructions, it is the direct object which is suffixed to the verb and the indirect object to *iyya*, e.g. *xarribōha iyyāy* ‘they’ve ruined it for me’.

2.2.1.4 Demonstratives

Table 6. Demonstratives

	A dialects	B dialects
Proximal m	*(hā)ḏa*	*(hā)da*
Proximal f	*(hā)ḏi*	*(hā)di*
Proximal pl.	*(hā)ḏēlēn/(hā)ḏēla*	*(hā)dēlēn/(hā)dēla*
Distal m	*(hā)ḏāk*	*(hā)dāk*
Distal f	*(hā)ḏīč*	*(hā)dīk*
Distal pl.	*(hā)ḏēlāk*	*(hā)dēlāk/(hā)dōl/(hā)dōla*

The feminine form of the demonstrative is often used as the form of vague reference, e.g. *šinhu ḏi*? ‘what’s this?’ or with masculine nouns, e.g. *ḏi lbēt* ‘this house’, *ḏīk ilyōm* ‘that day’.

When used as an adjective, the demonstrative precedes the noun; used contrastively, it follows. The proximal set of forms can be reduced to a clitic *ha-* prefixed to the defined noun, e.g. *ha-lbēt* ‘this house’, *ha-rrajājīl* ‘these men’.

2.2.1.5. Presentatives

The A dialects prefix *k(a)-* to any independent pronoun (see 2.1.3.4). The B dialects use a mixed system: *hāk-* + 3rd person independent pronouns only, the pronoun referring forward to a following noun, e.g. *hāk-hu ja, rafīgna* ‘here he comes, our friend!’ and *hāda-* + any independent pronoun, e.g. *hāda-nta taḥči ʿarabi tamām* ‘you speak Arabic perfectly!’. Note also B village *ayya*, e.g. *ayyāhu samād ʿindi mawjūd* ‘look here! I’ve got some manure already’.

2.2.1.6 Relative pronouns

All dialects *illi*. B village variants: *illaḏi*, *illadi*, *illi di*.

2.2.1.7 Interrogative pronouns

All dialects: *šlōn* ‘how?’, ‘what kind?’, *kēf/čēf* ‘how?’, *čam* ‘how much/ many?’, *wēn* ‘where?’, *mita* ‘when?, *min* and *minhu* ~ *minhi* fem. ‘who?’, *šfī* + pron. ‘what’s wrong with . . .?, *š-hast* ‘what’s the matter?’.

A dialects: *šinhu/šinhi* ‘what?’, *šingāyil* ‘what sort?’, *š-* formations: *š-* + verb, ‘what . . .?’; + preposition, *š-ḥaggah*, *š-ʿalēh*, *š-lēh*, *š-minnah* ‘why, because of what?’; + noun: *š-kiṯir š-gadd* ‘how much/many?’, *wara* + pronoun ‘why?’ *yahu/yahi* ‘which one?’

B dialects: *wēš* (urban), *wēšhu/wēšhi*, *wēšin*, *wē*, *ay šo* (various villages) ‘what?’, *wēš-rang* ‘what sort?’ preposition + *wēš* formations, e.g. *ʿala wēš*, *li wēš*, *ḥagg wēš* ‘why?’, *ayhu/ayhi* ‘which one?’, *anu* (villages) ‘who, which?’

Interrogatives have no fixed sentence position: pragmatic factors determine the word order.

2.2.2 Adverbs

Note that only unusual forms or local innovations not common in other dialects are listed in this section.

Some adverbs, e.g. *ʿād* ‘just, so, then, again’, *xōb* ‘then, but, probably’ have very many diverse uses and are not easily classifiable (see Holes 2001:367–368, 163–164 for details). *gad ~ ǧid ~ kid ~ čid* (all < Classical Arabic *qad*), *gaṭ* (< Classical Arabic *qaṭ*), and *čūd* (< Classical Arabic *yakūd*), are all used in partially overlapping senses to indicate possibility, epistemic modality, and the ‘experiential’ perfect (see Holes 2001: 415, 428–429, 467–468 for details).

2.2.2.1 Temporal

All dialects: *ilḥīn ~ alḥīn* ‘now’, *ha-lḥazza ~ ḥazzat ḥāḍir* ‘at this moment’, *ḥadd ilḥīn* ‘up to now’, *bāčir* ‘tomorrow’, *ʿugub bāčir* ‘the day after tomorrow’, *ilbārḥa* ‘yesterday’, *dōm* ‘always’, *min waqt* ‘early’, *(min) iṣṣibḥ* ‘in the early morning’, *iḏ̣ḏ̣iḥa* ‘in the forenoon’, *ilgāyla* ‘in the early afternoon’, *ilʿaṣir* ‘in the late afternoon’, *msayyān* ‘in the evening’, *lawwal* ‘in the old days’, *baʿad* ‘still, yet, also’, *min gabil* ‘beforehand’, *(min) ʿugub* ‘afterwards’, *abdan* ‘totally; ever, never’, *tāli* ‘then, next’, *āxir* ‘finally’, *awqāt* ‘sometimes’, *rāyiḥ* ‘continuously’. A dialects only: *ilǧābla* ‘tomorrow night’, *illābla* ‘the day after tomorrow’, *min wahal* ‘early’. B village dialects only: *baʿadan* ‘afterwards’, *abad* ‘always’, *amsiyya* ‘yesterday’, *ġādi* ‘continuously’, *gaʿadiyya ~ kuʿdiyya* ‘full-time’, *killəh u lā budd* ‘inevitably’.

2.2.2.2 Local

ihni ‘here’, *ihnāk* ‘there’, *minni* ‘over here’, *minnāk* ‘over there’, *minni u minni* ‘here and there, all over the place’, *ḥadir* ‘beneath’, *ḥadir fōg* ‘upside down’, *dāyir madār* ‘around’, *sīda* ‘straight ahead, directly’, *ha-ṣṣōb* ‘over here’, *ḏāk iṣṣōb* ‘over that way’.

2.2.2.3 Manner

ham ‘also’, *bass* ‘only’, *čiḏi* ‘thus’, *ha-ššakil*, *ha-nnamūna*, *ha-llōn* ‘like this’, *wāǧid ~ wāyid* ‘much, very’, *zēn* ‘well’, *killiš* ‘completely’, *iji ~ iyi* ‘approximately’, *mūl ~ mūliyya* and *marra*, *bi lmarra* ‘at all’ (both used only negatively), *zitāt* ‘quickly’, *falla* ‘well, excellently’, *zēn u mā zēn* ‘willy-nilly, come what may’, *marra waḥda* ‘suddenly, just like that’, *wakād* ‘certainly’, *ʿasā* ‘hopefully’, *bi lxašš* ‘surreptitiously’, *ʿala ġafla* ‘suddenly, without warning’; B only: *čidiha* ‘like this’, *ʿafar ~ ʿafarāt* ‘possibly, maybe’, *ḥirwa*, *miḥāri*, ‘approximately’, *dīdīh* ‘quickly, directly’.

2.2.3 Particles

2.2.3.1 Definite article

The article is *(i)l-*, and is assimilated by the sun letters as in Classical Arabic. In a few common phrases it is *al-*, e.g. *alḥīn* ‘now’. When prefixed to nouns whose first syllable is an open, unstressed /i/ or /u/, especially if the following vowel is long, it normally becomes *lə-*, e.g. *ləktāb*.

2.2.3.2 Indefinite article

Some uses of *wāḥid* have this function, when it precedes the noun, e.g. *ʿindəna wāḥid ʿabd yigāl lih* . . . ‘we once had a black man here called . . .’. *šay ~ šī* is used for an unspecified amount, e.g. *šī aṣīdah*, *šī yiṭīr* ‘some [birds] I catch, some fly away’. Dialectal *tanwīn* fulfils the ‘generic’ function of the indefinite article in some phrase-types, e.g. *bintin zēna* ‘a[n unspecified] nice girl’.

2.2.3.3 Genitive markers

māl(at) and *ḥagg* are both used e.g. *ilbēt māl ʿammi* ‘my uncle’s house’, *hāfāt ḥagg iṣbayān* ‘boys’ underpants’. *māl* can be used with inalienable possession, e.g. *irrajāyil mālah* ‘his feet’ (B example). Both markers have a wide range of other uses.

2.2.3.4 Negative particles

With indicative verbs in all dialects: *mā*. *mā* . . . *šay* (variant *ši*) may be used, but only for emphasis, e.g. *mā yiḍirrhum šay* ‘it doesn’t harm them at all’. *mā* is also used with ‘dummy’ verbs expressing possession *(ʿind)* and existence: *mā fīh*, *mā miš*, *mā hast*, *mā šay*, *mā min* and (B dialects) *māku* ‘there isn’t any . . .’, e.g. *čāh mā miš* ‘tea didn’t exist’; *dīra mā šay* ‘there was no town’; *mā min šarika, wi lā min šay* ‘there was no [oil]-company, or anything’; *māku hāda lgarāṭīs, aṣlan mā miš* ‘these banknotes didn’t exist originally at all’. *maḥḥad* or *māḥad* is ‘no one’.

With other parts of speech: *mu*, *muhu* masc., *mi* fem. (all dialects); variants: *mub*, *muhub*, *hub*, *humb* masc. and *mahi*, *hīb*, *mahīb* fem. (A). In nominal sentences, *mā* + independent pronoun is an alternative to pronoun + *mu*, etc., e.g. (A example) *inta mub rayyāl* or *minta (b)rayyāl* ‘you’re not a [real] man’.

lā: for negative imperatives and optatives, e.g. *lā yiṭra lmōt ʿala bālik* ‘let the thought of death not cross your mind!’ and all co-ordinated negatives. *ʿan lā* = ‘lest’, e.g. *yigʿad wiyyāha ʿan lā tistawḥiš* ‘he stays with her so she won’t be lonely’. *wi lā*: categorical negation, e.g. *wi lā xalaka ʿalēha* ‘she hadn’t got a stitch of clothing on her’.

2.2.3.5 Particles to introduce questions

The B dialects attach a clitic *ə* to any word to create a yes–no question (see 2.1.3.4). They also attach *lā* with a rising intonation to the end of sentences as an attention-maintaining device, e.g. *nḥuṭṭ fīh šakkar, lā* ‘we put sugar in it, right?’ The tag *mu čiḏi* ~ *čidi* ‘isn’t that so?’ is used in all dialects for the same purpose.

2.2.3.6 Existentials

fīh and *hast* are common to all dialects. *miš* and *aku* are mainly used by B speakers. All are negated with *mā*: see 2.2.3.4.

2.2.3.7 Prepositions

li ‘to, for’, *bi* (or *əb*) ‘with, by means of’, *fi* (or *əf*) ‘in, on’, *ḥagg* ‘to, for’, *ila* ‘to, towards’, *ʿala* ‘on, against’, *min* ‘from’, *ʿan* ‘away from, instead of’, *fōg* ‘on top of, above’, *taḥt* ‘under, near, compared with’, *ḥadir* ‘under, below’, *ṣōb* ‘towards’, *ḥadd* ‘as far as’, *wara* ‘behind’, *jiddām/giddām* ‘in front of’, *mjābil* ~ *mgābil* ‘opposite’, *bēn* ‘between’, *fi wusṭ* ~ *wasṭ* ‘in the middle of, inside’, *gabil* ‘before’, *baʿad*, *xalf*, *ʿugub* ‘after’, *bilā* and *bilayyā* ‘without’, *ʿind* ‘at, with, according to’, *māl(at)* ‘of, relating to’, *wiyya* ‘with’, *yamm* ‘beside’, *dūn* ‘in contrast with, different from’, *gufa* ‘dependent upon, incumbent on’, *dāyir madār* ‘around’, *miṯil* ~ *miṯlāt* and *šikil* ~ *šiklāt* ‘like’, *min bidd* ~ *biddat* ‘rather than, in preference to’.

2.2.3.8 Conjunctions

wi/u ‘and’, *willa* and *aw* ‘or’, *amma . . . aw* ‘either . . . or’, *lākin* and *bass* ‘but’, *innamā* ‘only, except that’, *inn* ‘that’, *yōm*, *min*, *lamma*, *lamman* ‘when’, *ila min*, *lēn*, *ilēn*, *ilamma* ‘whenever, until’, *lijil*, *liʾann*, *asbāb* (B dialects), *šminnah* (A dialects), *čēfan* ‘because’, *ḥatta*, *ḥakka*, *ḥagg* ‘so that’, *ʿan lā* ‘lest’, *gabil lā*, *gabilmā* ‘before’, *baʿad mā*, *xalf mā*, *ʿugub mā* ‘after’, *miṯil mā*, *šikil mā* ‘like’, *wēn mā* ‘wherever’, *kil mā* ‘the more . . . the more; whenever’, *mā dām* ~ *im dām* ‘as long as’, *činn* ‘as if’, *mā illa* ‘as soon as, no sooner than’. Conditional conjunctions: *in*, *iḏa*, *ila*, *lēn*, *lō*, *yō*, *(in) čān*.

2.2.3.9 Vocative particles

yā is used when addressing people by name.

2.2.3.10 Exclamations

Particles of affirmation and denial are: ‘Yes!’: *ē*, *ē naʿam, ajal, ʿajal, bala*, *mbala*. ‘No!’: *lā*.

yallah ‘come on!’; *yā rēt* ‘would that . . . !’; *bass* ‘stop! that’s enough!’; *ʿajal/ ʿayal* ‘well . . .’; *nzēn* ‘OK, right . . .’; *xōb* (B dialects) ‘very well, . . .’; *hā* ‘well, . . .’; *ila* . . . ‘lo and behold!’; *ʿād*, which has many uses, e.g. mild reproach, e.g. *ismaʿʿād* = ‘just listen, will you!’, cajoling, e.g. *rūḥi šʿād* ‘go, why don’t you?’; *ya* = ‘what?!’ for incredulity, often coupled with dismissal of a proposition; *yū!*, used by women only, to express fear or anxiety; *ax* used to express pain; *wēl* to express woe, sorrow; *bēl* ~ *bwēl* to express surprise; *čabb* ‘shut up!’; *afa* is used to express disgust. *š-* (A) and *wēš* (B) + noun are used to express admiration/ surprise, e.g. *š-ḥalāwatha* ‘how pretty she is!’. Also *amma*, e.g. *amma xōš* ‘how nice!’ and *yā* . . . e.g. *yā zīn ha-ššaʿar* ‘how beautiful your hair is!’, *yā min ʿindəh migaṣṣ al-ḥīn* ‘oh for someone with a pair of scissors right now!’

2.2.4 Nouns

2.2.4.1 Gender

Feminine by usage: double parts of the body; nouns denoting females, e.g. *umm* ‘mother’, *ʿarūs* ‘bride’; names of countries; a few common nouns, e.g. *šams* ‘sun’, *arḍ/arḏ̣* ‘earth, land’, *nār* and *ḍaww* ‘fire’, *rūḥ* ‘soul, spirit’, especially those which are reflexes of Classical Arabic *-āʾ*, e.g. *kahraba* ‘electricity’. *mā/mi/māy* ‘water’ may be of either gender. *nās* ‘people’ and several other human collective nouns may be of either gender, the determining factor being whether the word is used generically sg. fem. (e.g. ‘people have become educated’) or specifically pl. (e.g. ‘the people I told you about’).

2.2.4.1 Productive patterns

Common patterns are as below (unstressed /i/ and /u/ in initial open syllable are normally deleted):

CaCC *ḥalǧ* ‘mouth’; CiCC *bišt* ‘man’s cloak’; CuCC *ṣubḥ* ‘morning’; CaCCa *šagḥa* ‘jump’; CiCCa *dīra* ‘homeland’; CuCCa *ṣubʿa* ‘finger’; CaCaC *baḥar* ‘sea’; CiCiC *ǧidir* ‘cooking pot’; CuCuC *ḍ̣uhur* ‘noon’; CaCiC *ḥači* ‘talk, speech’; CiCaC *risan* ‘halter’; CCaCa *bġala* ‘clay pot’; CCiCa *ḥṭiba* ‘piece of wood’; CāCiC *yāhil* ‘child’; CāCCa *sālfa* ‘matter, business’; CaCāC *ṣarām* ‘harvested fruit’; C(i)CāC *simāṭ* ‘eating mat’; C(u)CāC *turāb* ‘soil’; CaCīC *ʿaṣīd* ‘porridge’; CaCīCa *darīša* ‘window’; CaCūC *xarūf* ‘sheep’; C(i)CīC *jifīr* ‘basket’; CCiCi *ṣfiri* ‘autumn’; CiCCa *dinya* ‘[this] world’; CaCCaC *čanʿad* ‘king mackerel’; CiCCiC *ʿišrig* ‘type of purgative’; CaCCāC *sammāč* ‘fisherman’; CiCCūC *bindūl* ‘area of ship’s deck’; CiCCāC *bindār* ‘storeroom on a ship’; CaCCīC *zanjīl* ‘chain’; CaCCāCa *ṭaffāya* ‘ashtray’; CiCCāCa *dirwāza* ‘gate’; CaCCūC *ṣalbūx* ‘shingle’; CāCāC *sāmān* ‘stuff, gear’; CāCūC *hāmūr* ‘grouper [fish]’; CāCūCa *xāšūga* ‘spoon’; CīCāC *nīšān* ‘target’; CāCīCa *fānīla* ‘undervest’; maCCaC *maʿāš* ‘salary’; maCCaCa *maḏ̣lama* ‘injustice’; maCCiCa *maṣbina* ‘group of men who launch a boat’; miCCaC *mišxal* ‘sieve’; miCCaCa *mirfaʿa* ‘food storage hoist’ (A); miCCiCa *mifliga* ‘oyster-knife’; miCCāC *mīdār* ‘fishhook’; miCCāCa *millāla* ‘food storage hoist’ (B); maCCūC *mamrūs* ‘type of sweet mash’; m(i)CaCCiC *mʿallim* ‘*Qurʾān* teacher’ (B); m(i)CaCCaC *mḥammar* ‘dish of rice and date molasses’; miCCiC *miʿris* bridegroom’; mistaCCaC *mistašfa* ‘hospital’.

A few nouns of very local reference end in the suffix *-ō* e.g. *ḥalwayō* ‘a type of fish’, *čāftō* ‘keelson [in maritime terminology]’, *ndēndō* ‘type of dance with drum accompaniment’.

The *-iyya* suffix is also productive, e.g. *ṭalliyya* ‘mist’, *fahmiyya* ‘understanding’, *ahliyya* ‘family’, *ṣiḥḥiyya* ‘health’ *ʿaṣriyya* ‘afternoon’ (cf. the Omani dialects).

2.2.4.2 External and internal plural

Internal

aCCāC	*aġrāš* < *ǧarša* ‘bottle’
aCCiCa	*amṯila* < *maṯal* ‘proverb’
aCCuC	*ašhur* < *šahar* ‘month’
CuCaC	*čuwal* < *čūla* ‘stove’
C(i)CaC	*rkab* < *rikba* ‘knee’
C(i)CiC	*sṭir* < *saṭir* ‘line’
C(u)CuC	*ḥṣur* < *ḥaṣīr* ‘mat’
CaCāC	*banāt* < *bint* ‘girl’
C(i)CāC	*člāb* < *čalb* ‘dog’
C(u)CūC	*glūb* < *galb* ‘heart’
CaCīC	*ḥarīm* < *ḥurma* ‘woman, wife’
CuCīC	*gufīf* < *guffa* ‘basket’
CuCC	*ġurb* < *ġarīb* stranger’
CiCCiC	*ḥimmil* < *ḥāmil* ‘pregnant’
CCaCān	*ṣbayān* < *ṣabi* ‘boy’
CiCCān	*firgān* < *firīǧ* ‘neighborhood’
CuCCān	*ṣuxlān* < *ṣxala* ‘young goat’
C(i)CāCa	*digāla* < *digla* ‘mast’
CaCāCa	*jamāra* < *jamri* ‘person from Bani Jamra’
CaCāCīC	*galālīf* < *gallāf* ‘shipwright’
CawāCCīC	*jawālbīt* < *jālbūt* ‘small boat’
CiCCāC	*tijjār* < *tājir* ‘merchant’
CīCCāC	*sībān* < *sāb* ‘irrigation channel’
CaCāCCa	*nawāxḏa* < *nōxaḏa* ‘sea-captain’
CaCāCwa	*naxālwa* < *naxlāwi*, *naxxāl* ‘palm-cultivator’
CaCāyCa	*bawāyga* < *bāyig* ‘thief’
CaCāCiC	*gawāṭi* < *gūṭi* ‘tin, packet’
CaCāyiC	*čanāyin* < *čanna* ‘daughter-in-law’
CaCāCīn	*balādīn* < *bilād* ‘village’
CaCāCi	*ʿaṣāri* < *ʿaṣriyya* ‘afternoon’
C(i)CāCi	*jimāʿi* < *jumʿa* ‘Friday’
maCāCiC	*mafāliǧ* < *mifliga* ‘knife for opening clams’
maCāCīC	*magāṣīṣ* < *magṣūṣ* ‘bankrupt person’
maCaCCa	*maṭawʿa* < *mṭawwaʿ* (A) ‘religious teacher’

anomalous or rare patterns:

CiCīC	*sinīn* < *sana* ‘year’
aCāCiC	*arāyil* ~ *arājil* ‘foot’ (alternative to *ryūl* ~ *rjūl*) < *rijl* ~ *rīl*

ubuhāt < *abu* ‘father’
ummahāt < *umm* ‘mother’

External plurals for human nouns are mainly formed with *-īn* masc. and *-āt* fem., the latter also being used for borrowings, e.g. *bāṣāt* < *bāṣ* ‘bus’. *-a* is used for some human nouns of the CaCCāC pattern, e.g. *ḥaddāda* ‘ironmongers’. The *-iyya* suffix is used for some male human nouns, e.g. *drēwiliyya* < *drēwil* ‘driver’, especially those ending in the relational *-i*, e.g. *bnātiyya* < *bnāti* ‘effeminate man’, *ṭirṭangiyya* < *ṭirṭangi* ‘layabout’.

2.2.4.3 Diminutives

Diminutives are much used by women, sometimes with a hypocoristic, sometimes a pejorative sense. Examples: *ṣbay* 'little boy' *bnayya* '[nice] little girl' (these two often used to refer to bride and groom), *ʿuwēnāt* '[nice] little eyes', *wlēdāt* '[nice] little kiddies [at *Qurʾān* school, of either sex]'; but *ḥbayyib ahlik* 'your family's little darling', *umayma* '[your] little mother', and *mraytik* 'your little wife', all sarcastic. In storytelling, diminutives are especially common, e.g. *ḍḥēwa* '[little] forenoon', *ṯwēb* '[little] dress', *mrēgdāt* '[little] beds'. A diminutive in pattern but not meaning is *msayyān* 'in the evening' (cf. Classical Arabic *muġayribān*).

The suffix *-ūna* is used to denote a small example, or a bit, e.g *ḥabbūna* < *ḥabb* in *ḥabbūnat raggi* 'a little watermelon'. Note also *ṣġayrūn* 'little, unimportant'.

CaCCūC is used for the diminutive of names, e.g. *ammūn* < *amīna*, *laṭṭūf* < *laṭīfa*, *xallūd* < *xālid*. The suffix *-ō* or *-aw* is also used for this purpose, e.g. *xalīlō*, *zambaw* (< *zēnabaw*), *fiḍḍaw* (< *fiḍḍa*).

2.2.4.4 Vocatives

There is a complex system of vocatives involving bipolar address forms, e.g. *yābū-k*, *yāxū-k*, *yāxt-ič*, *yumma-k*, *yā nasīb-ik*, etc., as well as *(yā) yubba*, *(yā) yumma*. See Holes (2001:4, 8, 19); Yassin (1977).

2.2.4.5 Adjectives

The following are additional to the nominal patterns:

CaCC, e.g. *ġalğ* 'difficult to understand'

CaCCān, e.g. *ṭabʿān*, pl. *ṭabāʿa* 'shipwrecked'

CaCCāCi, e.g. *waggāfi*, pl. *waggāfiyya* 'in a standing position', *gaʿʿādi* 'in a sitting position'. The pattern is also applied to specific styles of artefact or activity, e.g. *baḥḥāri* 'type of *mačbūs*', *ḥaddāri* and *nawwāri* two types of singing.

CaCCūC ~ CaCCīC, intensive patterns, e.g. *akkūl* ~ *aččūl* and *aččīl*, pl. *ačāčīl* 'greedy'.

CāCūC, intensive pattern, e.g. *šāġūl* 'hardworking'.

CaCCa: the feminine of some CaCīC adjectives, e.g. *xanṯa* 'catamite'.

CuCāC: plural of some adjectives, e.g. *kubār* 'old'

CiCCaC: plural of some adjectives, e.g. *ʿittag* 'old', sg. *ʿatīj*.

2.2.4.6 Color and deficiency adjectives

aCCaC masc., CaCCa fem., e.g. *aswad*, *sōda* 'black'. In the A dialect some of the masculine forms are resyllabified because of the *gahawa* syndrome, e.g. *xaḏar* 'green', *vatam* 'dumb'. The plural is CuCC in the A dialects, CuCCān in the B, e.g. *sumur/simrān* 'brown-skinned', *ḥūl/ḥūlān* 'squint-eyed'.

2.2.4.7 Elatives

These are formed as in Classical Arabic. The aCCaC pattern has been extended, however, e.g. *ašwa* < *šway* 'better', *astad* < *ustād* 'more expert' (B), *ašṯan* < *šēṭān* 'naughtier'. *adna*, elative of *dani* 'inferior' is used in constructs in the form *adnāt* with the sense 'the least . . . /the worst . . .', e.g. *adnāt iš-šay* 'the least little thing'

2.2.5 Numerals

'One': is *wāḥid*, *waḥda* fem. It precedes the noun as an indefinite article (2.2.3.2), and follows it to signify 'one' (as opposed to 'two', etc.), e.g. *b-agūl lik šay wāḥid* 'I'll tell you one thing, . . .', *marra waḥda* 'in one go, all at once'. Note also the phrase *čam wāḥid/waḥda* 'how many?', to which the answer is, e.g., *iṯnaʿšar wāḥid/waḥda* 'twelve'.

'Two': is *ṯnēn*, *ṯintēn* fem. (A variant *hintēn*). The dual or the plural followed by 'two' can be used, with no difference in meaning, e.g. *bintēn* or *banāt ṯintēn* 'two daughters'. Some speakers, A and B, have a feminine form *ṯnēna*, e.g. *kbūr ifnēna* 'two graves' (B village example).

'Three' to 'ten': the gender polarity system is as in Modern Standard Arabic, with no reanalysis of forms with *-t*. Telling the time: the masculine form is normally used (except *sāʿa ṯintēn* 'two o'clock'); in enumerating some currency and similar units, B speakers use the singular noun, e.g. *falāfa dīnār* 'three dinars', *xamsa lakk* 'fifty thousand' (but always *xams rubbiyāt* 'five rupees').

'Eleven' to 'nineteen': when enumerating, all speakers use the long form, e.g *xamstaʿšar sana* 'fifteen years'; in counting, and when no noun is mentioned, the A speakers have the short form without *-ar*, the B the long (data on this last point are at odds with Al-Tajir 1983:97).

'Hundred': *miya* or *imya*, pl. *miyāt*

'Thousand': *alf*, pl. *ulūf*. *ulūf ilulūf* 'thousands upon thousands'

Larger numbers: *lakk*, pl. *lkūk* signifies a large number (cf. English 'zillion'), for some speakers 10,000, for others 100,000.

Ordinal numbers are regular, except that ‘sixth’ has alternatives forms, *sādis* and *sātt*.

Count nouns: some B speakers use the *-āya* suffix, e.g. *ṭamāṭāya* ‘a [single] tomato’; for animals, *rās* is used, e.g. *ifnaʿšar rās bagar* ‘twelve cows’; for any sphere-like object, *ḥabba*, e.g. *ḥabbat raggi* ‘a [single] watermelon’, *arbaʿ xams ḥabbāt* ‘four or five [pearls]’.

2.2.6 Verbs

2.2.6.1 Forms

2.2.6.1.1 Form I

i. Perfect stems
In the A dialects, the factor determining the first vowel in the perfect stem is phonological (see 2.1.1.2). In the B village dialects, it is morphological, corresponding to the Classical Arabic split between roots with v_2 /a/, which have a dialectal *a-a* pattern, and those with /i/, /u/, which have a dialectal *i-i* or *u-u* pattern, depending on the combination of consonants in C_2 and C_3 (cf. the Omani sedentary dialects). The B urban dialects are a compromise: verbs which in Classical Arabic have v_2 /a/ are *a-a*, but Classical Arabic verbs with v_2 /i/ which are *transitive* are reclassified into the *a-a* group (although ‘mixed’ *i-a* forms of the *širab*-type are also heard), while intransitives have an *i-a* pattern, not *i-i* (Table 7).

ii. Imperfect stems
A dialects: if C_2 or C_3 is a guttural, the theme vowel is /a/, and the prefix vowel /i/, e.g. *yigʿad*, *yiṭbax*. A few older speakers have forms of the *yagʿad*-type in this verb class, as in Najd. If C_1 is a guttural, the stem vowel is /i/ and the prefix vowel /a/, the form then being resyllabified according to the rule already given, e.g. *yʿarf* < *yaʿrif*, *yxaṭib* < *yaxṭub*. In non-guttural stems, the A dialect follows the Classical Arabic system: theme vowel /a/ for Classical Arabic /a/, otherwise theme vowel /i/ or /u/ depending on the consonants in the C_2 and C_3 position, with the prefix vowel typically opposite in height from the theme vowel, e.g. *yilbas*, *yišrab* but *yabriz*, *yaḏkir*, *yaṣbur*.

B dialects: the theme vowel mirrors that of Classical Arabic: /a/ where Classical Arabic has /a/, and a high vowel (close front or open back depending on consonant environment) where Classical Arabic has /i/ or /u/. The prefix vowel is always /i/.

2.2.6.1.2 Derived forms

Noteworthy semantic aspects are as follows:

Form II is mainly causative, but also: denominative, e.g. *naggaṣ* ‘to get lumpy’ < *nugṣ* ‘lump’; extensive, e.g. *laggaṭ* ‘to pick up in large quantities’; and estimative, e.g. *bawwag* ‘to declare someone to be a thief [*bāyig*]’. A number of Form II verbs have freely varying Form III forms, e.g. *layyam-lāyam* ‘to go around collecting’, with the same extensive sense.

Form IV exists only in a few fixed expressions, e.g. *aġnāk allāh* ‘May God enrich you!’, *aḥsant* ‘thank you’ (lit. ‘you have done well’), having been replaced by Forms I and (mainly) II.

Forms V and VI: in the perfect the vowel of the *t-* prefix is often elided; in the imperfect the A dialects have *iyt-*, *tti-*, *nti-* whereas the B have *yit-*, *tit-*, *nit-*. Form V is often reflexive or passive, e.g. *tibaṭṭaṭ* ‘to burst open’, *timaššak* ‘to become torn, tattered’. Form VI often implies gradation or repetition, e.g. *tigāṣar* ‘to peter out’, *timāyal* ‘to sway, incline’.

Form VII is the normal means of passivization. There are a few verbs with a Form VIII alternative, e.g. *inbāg*, *ibtāg* both ‘to be stolen’. The *in-* may be prefixed to Forms II, e.g. *inʿawwar* ‘to get hurt’, and V, e.g. *intifarrag* ‘to become dispersed’.

Table 7. Perfect Stems

Classical Arabic	A dialects	B villages	B urban	
šariba [trans.]	*šarab*	*širib*	*šarab*	‘he drank’
zaʿila [intrans.]	*zaʿal*	*ziʿil*	*ziʿal*	‘he got upset’
kaṯura [intrans.]	*kiṯar*	*čufur*	*kufar*	‘it grew more numerous’
ḍaraba	*ḏ̣arab*	*ḍarab*	*ḍarab*	‘he hit’
kataba	*kitab*	*katab*	*katab*	‘he wrote’
dafaʿa	*difaʿ*	*dafaʿ*	*dafaʿ*	‘he pushed’

Form X is very productive, especially of change-of-state verbs: *istamarr* 'to become bitter', *istaḥmag* 'to get angry'. The *ista-* prefix is occasionally prefixed to other forms, e.g. Form III, e.g. *istafāham* 'to come to an understanding'.

2.2.6.1.3 Internal passive

The internal passive occurs commonly in a few fixed phrases, e.g. *wāḥid yigāl lih* 'someone called . . .', *wlidt* (A)/*xlukt* (B villages) 'I was born'; otherwise sporadically e.g. *alḥīn ruwyat, tinǧaḥ* 'now it [crop] has been watered, it'll grow' (B village); *lēn yinzaf u yisikk illōḥ* 'when it has been baled out and the timbers have tightened' (A).

2.2.6.2 Inflection

2.2.6.2.1 Imperfect

Table 8. Inflection imperfect (A dialects)

yaktib 'he writes'	singular	plural
1st	*āktib*	*naktib*
2nd m.	*taktib*	*tikitbūn*
2nd f.	*tikitbīn*	
3rd m.	*yaktib*	*yikitbūn*
3rd f.	*taktib*	

These are the canonical syllable structures for A imperfect forms. Those with the *a*-stem vowel have *yismaˁ*, *yisimˁūn*-type forms. Those with C_1 = guttural have *yˁarf*, *yˁarfūn*-type. See 2.2.6.1.1 (ii) above.

Table 9. Inflection imperfect (B dialects)

Theme vowel /i/, /u/: *yiktib* 'he writes'	singular	plural
1st	*aktib*	*niktib*
2nd m.	*tiktib*	*tikitbūn*
2nd f.	*tikitbīn*	
3rd m.	*yiktib*	*yikitbūn*
3rd f.	*tiktib*	

Theme vowel /a/: *yismaˁ* 'he hears'	singular	plural
1st	*asmaˁ*	*nismaˁ*
2nd m.	*tismaˁ*	*tismaˁūn*
2nd f.	*tismaˁīn*	
3rd m.	*yismaˁ*	*yismaˁūn*
3rd f.	*tismaˁ*	

Syllable structure in B dialects suffixed forms thus depends on whether the theme-vowel is high or low.

Aspect/mood prefixes: all dialects have *b-* for proximate intent. Some B dialects have *d-* for continuous/habitual aspect (see 2.1.3.4). *gāˁid* is used more widely to express continuous or iterative processes.

The imperative: A dialects: *iktib*, *kitbi* (or *kitbay*), *kitbu* (or *kitbaw*); B: *iktib, iktibi, iktibu.* The negative imperative is A: *lā tiktib/tikitbi/tikitbu*; B: *lā tiktib/tiktibi/tiktibu.*

2.2.6.2.2 Perfect

Table 10. Inflection perfect (A dialects)

kitab 'he wrote'	singular	plural
1st	*kitabt*	*kitabna*
2nd m.	*kitabt*	*kitabtaw*
2nd f.	*kitabti/kitabtay*	
3rd m.	*kitab*	*ktibaw*
3rd f.	*ktibat*	

Consonant-initial suffixes can be replaced with *-ēt*-type in any type of stem in the A dialects, e.g. *kitbēt*, *gālēna*, *tzawwajētaw*, but this type of form is now becoming obsolete.

Table 11. Inflection perfect (B dialects)

katab ' he wrote'	singular	plural
1st	*katabt*	*katabna*
2nd m.	*katabt*	*katabtūn(a)*
2nd f.	*katabtīn(a)*	
3rd m.	*katab*	*katabaw*
3rd f.	*katabat*	

Verbs with an *i-i* (in Manāma *i-a*) stem have 3rd person forms of the type *simˁat*, *simˁu*.

2.2.6.3 Participles

The morphology of participles is similar to Classical Arabic. CvCCān is productive, e.g. *jūˁān* (variant *jēˁān*) 'hungry', *ḍarṭān* 'useless [lit 'farty']', *kišḥān* 'cursed with bad luck'.

2.2.6.4 Verbal nouns

Form I. The most common patterns are:

CaCC	*ṭabx/ṭabix* 'cooking'
CiCC	*čiṯr*, *čiṯir*, *kufr* (B) 'amount'

CuCC	*kubr*, *kubur*, *čubr* (B) ‘size’
CiCaC	*rifaʿ* ‘height’
CaCCa	*jēba* ‘bringing’
CuCCa	*buġya* ‘desire’
CaCāC	*ġarās* ‘planting’
C(u)CūC	*blūġ* ‘attaining, puberty’
CaCāCa	*ġašāma* ‘naivety’
C(i)CāCa	*ḥiyāka* ‘weaving’
CaCīC	*barīx* ‘pulling a boat in on a rope’
C(i)CīCa	*nizīfa* ‘baling out’
C(u)CūCa	*ruṭūba* ‘humidity’
CiCCān	*dikrān* (B) ‘remembrance’
CaCaCān	*šayalān* ‘removal’ (II*y*)
maCCaCa	*maḥabba* ‘love’

Derived form verbal nouns that are original to the dialects (and which are attested in Classical Arabic) are:

II:	tvCCūC	*taʿlūm* ‘teaching’
	tiCCāC	*tiswā* ‘doing, deed’
	taCCīCa	*tasgīya* ‘watering’
	CvCCāC	*giffāl* ‘ending of the pearling season’, *yabbāb* ‘ululation’.
III:	m(u)CāCaC	*mrāmaḥ* ‘kicking, struggling’.

Modern Standard Arabic-type verbal nouns Forms V, VI, VII, VIII, and X do occur, but not in uneducated speech.

2.2.7 Weak verbs

2.2.7.1 Geminates

Perfect forms with consonant-initial suffixes are on the usual pattern CaCCēC, e.g. *raddēt* ‘I returned’. There is contraction in the active participle of Form I, e.g. *hābb rīḥ* ‘dextrous, expert’. But in Forms III and VI, there is no vowel contraction, e.g. *ḥājajōni* ‘they argued with me’, *yitḥājaj* ‘he gets into an argument’, *mšādid* ‘resistant’.

2.2.7.2 Iʾ verbs

In the perfect of Form I, some verbs lose the initial /ʾ/, e.g. *akal*, *axaḏ* and behave like strong verbs. In the A dialects, *akal* and *axaḏ* have alternatives *kal* and *xaḏ* which behave like doubled verbs. In other verbs the lost /ʾ/ is replaced by vowel length, e.g. *āman* ‘to be safe’, *āmar* ‘to order’. In all cases, the imperfect is of the form yāCCvC, e.g. *yākil*, *yāmin* and the active participle *māxiḏ*, *māmir* etc. In Forms II and III, /ʾ/ has either been (a) lost, e.g. *aḏḏan* ‘to call to prayer’ or (b) replaced by a semi-vowel, e.g. *waddam* ~ *yaddam* ‘to feed’, *wānas* ‘to keep someone company’. In case (a), in the imperfect, it reappears: *yiʾaḏḏin*, *yiʾakkil*. In Form VII it is preserved, e.g. *inʾaxaḏ*, *inʾakal*, but in Form X it is lost, e.g. *istānas* ‘to be content’.

Form I imperatives: there is much variation, but the main forms are: A dialects *ikil*, *ikli*, *iklaw*; B *kil*, *ikli*, *iklu* (with *č* in villages).

2.2.7.3 I*w* verbs

The /w/ is preserved in the Form I imperfect in all dialects, becoming *yō*- or *yū*-, e.g. *yōguf*, *yōṣal*, *yūʿa*, and in the imperative, e.g. *ōguf* ‘stop!’. One verb of Form VIII fails to assimilate it: *iwtaʿa* ‘to become aware of’.

2.2.7.4 II*w*/*y* verbs

The vowel of the Form I imperative is always long: *gūl* ‘say’, etc. In the passive participle, /y/ and /w/ are treated as strong consonants (with /w/ → /y/), e.g. *madyūs* ‘trodden on’, *mabyūʿ* ‘sold’, *mašyūf* ‘seen’, *mabyūg* ‘stolen’. The Form VIII perfect behaves as Form I, e.g. *iḥtijt* ‘I needed’, *gilt* ‘I said’.

2.2.7.5 III*y* verbs

There are two types: the /a/-type, e.g. *nisa/yinsa* and the /i/-type, e.g. *miša/yimši*. Verbs which in Classical Arabic had final-hamza lost it and were absorbed into the /a/-type, e.g. *gara/yigra*, and Classical Arabic final-/w/ verbs have been absorbed by the /i/-type, e.g. *ḥaba/yaḥbi* ‘to crawl’. One verb is differently classified by the dialects: A *baġa/yabġi* ~ *yabbi*, B *baġa/yibġa* (and variants *yibba*, *yubba*, *yumba*) ‘to want’.

All Bahraini dialects have masculine imperatives which lack a final *-i*, e.g. *ʿaṭ* ‘give!’, *wall* ‘clear off!’, *ištak* ‘complain!’. A further peculiarity of the A dialects only is that (like → Najdi dialects) they allow iCC imperative forms, e.g. *ibg* ‘stay!’, *imš* ‘go!’. Other verbs, somewhat unpredictably, have an epenthetic vowel, e.g. *imiš* (variant of *miš*) ‘go!’, *igir* ‘read!’, *irim* ‘throw!’, *išiw* ‘roast!’. The B speakers have *miš* ‘go!’ but *ibga/igra/irmi*-type forms for the masculine imperative in other verbs.

Peculiarities of the B dialects: in the perfect of the /a/-type, the radical /y/ shows up in the 3rd person forms, e.g. *nisyat*, *nisyaw*. In the imperfect, the endings of the suffixed forms of the imperfect are *-ēn* and *-ōn*, e.g. *tinsēn*, *tinsōn*.

2.2.7.6 Irregular verbs

The verb *ja* (B)/*ya* (A) 'to come' has irregular inflection (Table 12).

Table 12. Inflection of the verb *ja/ya*

singular	A	B	A	B
1st	*yīt*	*jīt*	*ayi*	*aji*
2nd m.	*yīt*	*jīt*	*tyi*	*ijji*
2nd f.	*yīti*	*jītīn*	*tyīn*	*ijjīn*
3rd m.	*ya*	*ja*	*iyi*	*iji*
3rd f.	*yat* ~ *yāt*	*jat*	*tyi*	*ijji*
plural				
1st	*yīna*	*jīna*	*nyi*	*inji*
2nd	*yītaw*	*jītūn*	*tyūn*	*ijjūn*
3rd	*yaw*	*jaw*	*iyūn*	*ijūn*

In the A dialect, the perfect 3rd pers. pl., when suffixed, has forms of the type *yawwōk* 'they came to you' as a variant of *yōk*.

2.2.8 Quadriliterals

These are very common. The main types: reduplicatives, e.g. *lamlam* 'to collect together'; echoic, mimetic in which C_2 = /l, r, n/, e.g. *balbaṣ* 'to dangle', *dandar* 'to bicker'; C_2 = /w/ in color verbs and bodily states (replacing Classical Arabic Form IX), e.g. *bōyaḍ* 'to be whitish, go white', *fōšaḥ* 'to be bow-legged', and denominatives, e.g. *sōlaf* 'to chat' < *sālfa* 'matter, affair', with reciprocal form *tisōlaf* 'to chat to each other'; C_2 = /y/ in denominatives, e.g. *tērab* 'to dance the *trēnbō* [a kind of wedding dance]', and with a *t-* prefix denotes bodily/mental states, e.g. *tkēsal* 'to be lazy, backslide', *tṣēmax* 'to pretend to be deaf'(→ Kuwaiti Arabic, → Omani Arabic); C_2 = /r, n/ inserted into triliteral, e.g. *šarbak* 'to ensnare' < *šabak* 'net', *tḥanda* 'to groan' < *ḥada* 'to urge, goad'. Others fall into no particular form: *laʿwaz* 'to make dirty', *tgašmar* 'to joke, play tricks'. Apart from the above, there are denominatives formed from secondary forms, e.g. *tmalfaʿ* 'to wear a black filigree face-veil (= *malfaʿ*)'.

Inflection, participle formation, and verbal nouns (where they exist) conform to the patterns already described, e.g. verbal nouns: *dandara* 'bickering', *mfōšaḥ* 'state of being bow-legged', *mtērab* 'performance of a type of wedding dance'.

Bibliographical references

Al-Tajir, Mehdi. 1983. *Language and linguistic origins in Bahrain.* London: Kegan Paul International.

Holes, Clive. 1983. "Bahraini dialects: Sectarian dialects and the sedentary/nomadic split". *Zeitschrift für Arabische Linguistik* 10.7–37.

——. 1984. "Bahraini dialects: Sectarian differences exemplified through texts". *Zeitschrift für Arabische Linguistik* 13.27–67.

——. 1987. *Language variation and change in a modernising Arab state: The case of Bahrain.* London and New York: Kegan Paul International.

——. 2001. *Dialect, culture and society in Eastern Arabia.* I. *Glossary.* Leiden: E.J. Brill.

——. 2002. "Non-Semitic elements in the Arabic dialects of Arabia". "*Sprich doch mit deinen Knechten aramäisch, wir verstehen es!": 60 Beiträge zur Semitistik für Otto Jastrow zum 60. Geburtstag*, ed. Werner Arnold and Hartmut Bobzin, 269–280. Wiesbaden: O. Harrassowitz.

——. 2005. *Dialect, culture and society in Eastern Arabia.* II. *Ethnographic texts.* Leiden: E.J. Brill.

——. Forthcoming. *Dialect culture and society in eastern Arabia.* III. *Dialect description.* Leiden: E.J. Brill.

Johnstone T.M. 1967. *Eastern Arabian dialect studies.* London: Oxford University Press.

Prochazka, Theodor. 1981. "The Shiʿi dialects of Bahrain and their relationship to the Eastern Arabian dialect of Muḥarraq and the Omani dialect of al-Ristāq". *Zeitschrift für Arabische Linguistik* 6.16–55.

Smeaton, B. Hunter. 1973. *Lexical expansion due to technical change as illustrated by the Arabic of al-Hasa, Saudi Arabia.* Bloomington, Ind.: Indiana University.

Yassin M.A. 1977. "Bi-polar terms of address in Kuwaiti Arabic". *Bulletin of the School of Oriental and African Studies* 40. 297–330.

Clive Holes (University of Oxford)

Bambara

1. Bambara and Arabic

Bambara is the largest language family of the Mande group (Niger-Congo family). From the 19th century onward, Bambara gained influence in the region as lingua franca of trade and army, and, through Bambara-speakers, in French administration and education. It is now the predominant lingua franca in → Mali and Eastern Senegal, and it is also spoken in Gambia, Burkina Faso, Guinea, and Sierra Leone (Baldi 1992:1–3), by a total of approximately 10 million or more people as first (around 2.8 million) or second language (<www.ethnologue.com>).

In the 1st millennium C.E., Arabic came to West Africa as the language of trade and Islam. It was the first written language in the region and an important source of loans for Mande languages. In the course of their ongoing contacts with Arab traders, speakers of Mande languages started converting to Islam from the 9th century onward. In the 11th century, al-Bakrī describes how Arab Muslims lived in the Mande region and Mande rulers converted to Islam. Islam and Islamic education became well established in the region in the 14th–16th centuries. Yet, the Bambara did not convert in large numbers until the 18th century.

Bambara probably borrowed from Arabic in the first place through the intermediary of other Mande languages, as well as → Songhay and → Fulfulde, while contact with native speakers of Arabic was limited. The influence of Arabic on Bambara remained restricted to the borrowing of a relatively small number of loans (Baldé 1980: 208; Calvet 1974:205; Toure 1983:34). Since the 1950s, a growing number of Bambara-speakers have achieved fluency in Modern Standard Arabic through modern Arabic schools, a development that might affect the influence of Arabic in the future. The use of Arabic is moreover highly valued within the learned religious elite.

2. Phonological and morphological adaptations

The number of Arabic loans in Bambara is too small to establish rules for phonological and morphological changes (Lagarde 1988:25). Authors also note that their list of loans is incomplete. Bailleul (1996) marks 220 loans, Lagarde (1988:5–18) supplies a list of 219, Baldi (1992:17–24) 140, Baldé (1980:200–205) 78, and Calvet (1974: 213) 50.

Drawing on terms introduced by P. F. Lacroix, Baldé (1980:196; also Toure 1983:35–36) differentiates between 'loans' and 'quotations'. Loans are phonologically and morphologically integrated into Bambara, show semantic shift, and occur in all registers. Quotations remain largely unchanged as 'imitations' of Arabic and are restricted to certain speakers and certain registers. They are replaced by other terms in other registers even by the same speakers. Quotations draw especially on the semantic field of religion and are associated with knowledge of Islam and of sacred language (Baldé 1980: 206–207). In what follows only loans are considered.

Baldi (1992:4), Lagarde (1988:30–39), and Toure (1983:9–20) supply several constants of phonological and morphological change. Most changes are governed by a maximum in phonetic ease, since certain Arabic phonemes are difficult to pronounce for native speakers of Bambara. The consonants *r*, *d*, and *l* are interchangeable in Bambara and, consequently, also in loans. Consonants and vowels in the Arabic word are sometimes deleted in the beginning (*bada* < *ʾabadan* 'always/never'), middle, and especially at the end of loanwords (*naamu* < *nāmūs* 'mosquito(es)'). In initial position only the vowels *a* and *i* are kept. After final consonants vowels are often added; in consonant clusters a vowel is often inserted (*hakili* < *ʿaql* 'mind, reason'). In diphthongs the first vowel is dropped. Long vowels are sometimes preserved and sometimes shortened, whereas short vowels may be lengthened. Metathesis may occur, and even three consonants may change their position (*bàtàki* < *kataba* 'to write'). Prenasalization occurs in initial (*ntamaru* < *tamrun* 'dates'), middle, and final position. Geminate consonants are usually replaced by the single consonant, but sometimes they are preserved; in other cases, the geminate consonant is replaced by a single consonant and a long vowel (*jiidi* < *jiddan* 'very'). Final vowels appear in certain loans (e.g. *kalimu* < *qalamu* 'pen'), probably through consistent final vocalization in Islamic education, while they are omitted in others. Sometimes the nunation changes into the corresponding vowel (*abada* < *ʾabadan* 'always/never') or a nasal sound. The Arabic article *al-* is sometimes omitted, especially when it is assimilated to the first consonant of the noun. When the loanword retains the Arabic article *al-*, the vowel *i* is inserted (*àlìkama* < *al-qamḥ* 'wheat', *àlìjɛnɛ* 'paradise' < *al-janna*). Sometimes, the article is reanalyzed as part of the word, in the form *la-*~*l-* (*lahaji* < *al-ḥājjī* 'pilgrim', *lasiri* < *al-ʾaṣl* 'origin', *làfiya* 'leisure time' < *al-ʿāfiya* 'health, well-being').

The following changes of phonemes are mentioned by Baldi (1992:4-17), Lagarde (1988:26-30), and Toure (1983): *d*, *t*, *r*, *s*, *k*, *l*, *n*, *w* are usually preserved. The glottal stop ʾ becomes *h* (*hadama* < (*ibn*) *ʾādam* ‘human being’); *w* (*balawu* < *balāʾun* ‘misfortune’); *y*, a nasal ending (*kàlan* < *qaraʾa* ‘to read’); a long vowel; or it disappears. *b* is preserved, but sometimes changes into *f* in intervocalic position (*tufa* < *ṭūb* ‘brick(s)’). The interdental *ṯ* becomes *s* or *t*, while *ḏ* is preserved or becomes *j*, *r*, or *l*. The sibilant *z* becomes *j* (*jaka* < *zakāt* ‘alms tax’) or disappears, *š* is preserved or becomes *s* (*sitane* < *šayṭān* ‘Satan’). The emphatic consonants *ṣ*, *ḍ*, *ṭ* become *s*, *d*, *t*, while *ḏ̣* becomes *s*, *j*, or *z*, and *q* becomes *k*, *g*, or a long vowel. The velar *x* becomes *k* (*kibaru* < *xabar* ‘news’), *h*, *g*, or disappears (*àlàmisa* < *al-xamīs* ‘Thursday’); the voiced velar *ġ* becomes *g*, *k*, *y*, or disappears (*baliku* < *bāliġ* ‘adult’). The pharyngeal *ḥ* usually becomes *h* or disappears (*hijàbu* ‘divine protection’ < *ḥijāb* ‘veil, amulet’, *làkìka* ‘true, real’ < *al-ḥaqīqa* ‘truth’), the voiced pharyngeal ʿ usually becomes a long vowel or disappears (*juma* < *jumʿa* ‘Friday’, *àràba* < *ʾarbiʿāʾ* ‘Wednesday’, but *dùba* ‘blessing’ < *duʿāʾ* ‘prayer’, *làánsàra* ‘afternoon prayer’ < *al-ʿaṣr*); *h* is usually preserved, especially at the beginning of the word (*hami* < *hamm* ‘worry, desire’).

3. SEMANTIC DOMAINS OF LOANS

Some loans keep their original meaning, while others add new nuances or meanings (Toure 1983:36). Authors differ as to the proportion of loans relating to specified semantic domains, but general trends emerge. Religion is an important field of borrowing, but altogether non-religious loans outnumber religious loans. Toure (1983:35) mentions 52 percent of loans related to religion, Lagarde (1988:19) 41.2 percent, Baldé (1980:200-205) 32 percent, Baldi (1998: 17-24) 25 percent, and Bailleul (1996) 22.7 percent. Trade and domains related to trade – such as money (e.g., *dɔrɔmɛ* ‘piece of five francs’ < *darāhim* ‘dirhams [pl.]’), measurements, food, and products of material culture – are another important field of borrowing. Other loans relate to semantic domains such as greetings, time and space, economics, politics, legislation, military, and sociocultural, intellectual, and domestic life, including ethical qualifications such as *hasidi* < *ḥāsid* ‘envious’ and *bàkilu* < *baxīl* ‘greedy’. Several verbs and adverbs as well as many personal names are borrowed, too (Baldé 1980: 200-205; Lagarde 1988:19; Toure 1983:20-25), e.g. *hàlàki* < *hallaka* ‘to destroy’, *joora* ‘to visit [a saint’s tomb]’ < *zāra* ‘to visit’, *jiidi* ‘to increase’ < *zāda/yazīdu*. In some cases verbs seem to have been formed from Arabic nominals, e.g. *bawu~bawuli* ‘to urinate’ < *bawl* ‘urine’, *dùrùsi* ‘to memorize’ < *durūs* ‘lessons [pl.]’; adverbs include *hali* ‘even, so’ < (*ʿalā kulli*) *ḥāl* ‘in any case’ (?), *abada* < *ʾabadan* ‘always, never’, and *baasi* ‘bad’ < *baʾs*.

Nowadays, there seems to be a trend to replace Arabic loans with French loans in certain semantic domains. Calvet (1974:211) notes that there are no loans of Bambara origin in Arabic.

BIBLIOGRAPHICAL REFERENCES

Bailleul, Père Charles. 1996. *Dictionnaire bambara–français*. Bamako: Editions Donniya.

Baldé, Abdoulaye. 1980. “Contribution à l’étude de l’impact culturel des emprunts arabes dans les langues manding du Sénégal”. *Itinérances – en pays peul et ailleurs*: *Mélanges réunis par les chercheurs de l’ERA 246 du CNRS à la mémoire de Pierre Francis Lacroix [1981]*. I. *Langues*, 195–209. [Paris]: Société des Africanistes.

Baldi, Sergio. 1992. “Arabic loans in Bambara”. Paper presented at the Cinquième colloque de linguistique nilo-saharienne. Nice.

Calvet, Louis-Jean. 1974. *Linguistique et colonialisme: Petit traité de glottophagie*. Paris: Payot.

Lagarde, Michel. 1988. *Influence de l’arabe sur la langue bambara*. Falajè, Mali: Centre d’Etude de Langue Bambara de Falajè.

Toure, Saliou Almahady. 1983. *Les emprunts à l’arabe communs aux langues nationales: Songhay et bambara*. Mémoire de fin d’études, Ecole Normale Supérieure, Bamako.

DINIE BOUWMAN (Groningen The Netherlands)

Bangladesh

The position of the Arabic language in Bangladesh has undergone a dramatic change over the last three decades, influenced by domestic political changes and the country’s evolving relationship with Arabic-speaking nations.

Bangladesh came into existence in 1971 after a lengthy nationalist movement which emphasized the linguistic identity of the Bengali population of the eastern part of Pakistan. The ethnic homogeneity of the population (about 98 percent of the people are Bengalis and speak Bengali), the low level of literacy, and the linguistic nationalism that inspired the population during the nationalist era have all helped to minimize the use of any other languages in public life. During the Pakistani era (1947–1971) the introduction of Urdu as the sole state language and the government’s insistence on its use as a medium of communication hardened the negative attitude of the literate Bengali community toward both the Urdu and Arabic languages, perhaps because of the misconception that both belong to the same language family. The misconception stems from Urdu's use of Arabic script. However, Arabic continued to enjoy respect as the language of religious texts (i.e. the *Qur'ān*) and was offered as an optional course at secondary level (grades 6–10). Following independence, Arabic lost further ground, and was dropped altogether from the school curriculum. At university level, however, Arabic language and culture remained a subject of study, although relatively few students were attracted to the subject. The Department of Arabic at the University of Dhaka, established in 1921, remained active and continued offering a three-year bachelor's degree and a one-year master's degree.

The situation began to change in the mid-1970s when the Arabic language started to draw the attention of a larger population. This may be attributed to a number of reasons. First, the political changes in 1975 which deposed the nationalist leaders and led to strong ties with Saudi Arabia and other oil-rich Arabic-speaking countries in the Gulf; second, the massive export of manpower to these countries (when this labor migration began in 1976, 217 persons migrated to Saudi Arabia and 1,989 persons to the United Arab Emirates; by 2000, the numbers had reached 144,618 and 34,034 respectively); and third, the proliferation of religious educational institutions (primarily madrassahs). Thus, while the warm relationship between the governments of Bangladesh and the Arabic-speaking countries of the Middle East and the Gulf provided recognition for Arabic as an important language in the global arena, the increased job opportunities for professionals and skilled and unskilled labor have been the impetus for the acquisition of conversational Arabic by many Bangladeshis. It is the latter trend that has fostered the growth of a number of small-scale private Arabic teaching institutes in Bangladesh, mostly set up in the 1980s. The madrassahs, some of which have received financial support from the Gulf nations in general and Saudi Arabia in particular, by contrast have continued teaching religious texts in Arabic without seeking to convey a broader understanding of the language. The post-1975 political environment, which emphasized the 'Islamic identity' of Bangladeshis, also brought changes to mainstream public education. Religious education was reintroduced as a mandatory subject at primary (grades 1–5) and secondary (grades 6–10) levels, and thus Arabic – as an integral part of religious education – returned to the schools. The military regime that came to power in 1975 decreed that Islam should play a pivotal role in the daily lives of Bangladeshis which, it was stressed, was the world's third most populous Muslim nation. In this environment Arabic assumed a new significance and a greater symbolic value.

In 1983, the new education policy of the military regime proposed Arabic as a mandatory language course at primary and secondary levels. The government insisted that learning Arabic would increase job opportunities in Middle Eastern countries, and would also help spread Islamic values. The proposal faced resistance from all political parties and the student community, and was finally withdrawn.

The growth in migration to the Middle East encouraged the teaching of Arabic in specialized colleges. In addition to these privately run small institutions, Arabic language courses are offered by the Modern Language Institute of the University of Dhaka, and the Islamic Foundation of Bangladesh, a government institution established for the purpose of preaching and propagating the values and ideals of Islam. The private institutions and the Modern Language Institute teach Modern Standard Arabic while the focus of the courses of the Islamic Foundation is Classical Arabic. This difference is primarily due to the objectives of the courses and the clientele they serve. The private institutions offer short courses to those who are aiming at or have already been selected for jobs in the Middle East. Their courses are designed to impart spoken Arabic and basic writing skills. The Islamic Foundation offers a two-year, three-tier curriculum. A three-month preparatory course is followed by a nine-month beginner course. Successful completion of these courses allows students to enroll in the one-year advanced course. The focus of the curriculum is the acquisition of written and spoken language skills. The students of the curriculum vary from madrassah students to government officials and professionals. The Modern Language Institute of Dhaka University offers a four-tier curriculum which focuses on language as well as modern Arabic literature. The four tiers comprise Junior Certificate, Senior Certificate, Diploma, and

Higher Diploma. Some of the private universities established in the 1990s also offer courses in Arabic. However, unlike the Department of Arabic at the University of Dhaka, the private universities emphasize the teaching of the Arabic language. A small number of Arabic journals and magazines are published in Bangladesh, most of which started publication in the 1990s. The Department of Arabic of Dhaka University publishes *Al-Majallatu ʾlʿArabaiyah*. The journals include *Ikra* and *Al Islah*, and the monthly magazines include *Al-Kalam* and *Al-Huda*.

→ Bengali, a member of the Indo-European family of languages, has, over the centuries, through cultural exchanges, borrowed a number of words from various languages including Farsi, Turkic, and Arabic during the Mughal period (ca. 12th century to ca. 1563). The Perso-Arabic influence on Bengali vocabulary is also partly due to the spread of Islam throughout eastern Bengal. According to Muhammad Shahidullah (2002: 59), one of the most prominent Bengali linguistics scholars, at least 2,500 Arabic words have been absorbed into Bengali. Words such as *dokan* 'shop' (Arabic *dukkān*), *tarikh* 'date' (*taʾrīx*), *kolom* 'pen' (*qalam*), *bonduk* 'gun' (*bunduq*), *mokaddema* 'litigation' (*muqaddima* 'introduction, premise'), *ain* 'law' (*ʿayn* 'essence, substance'), *kanun* 'convention' (*qānūn*), *adalat* 'court' (*ʿadāla* 'justice'), *hakim* 'magistrate' (*ḥākim*), *vakil* 'lawyer' (*wakīl*), *faisala* 'compromise' (*fayṣal* 'decisive criterion, arbiter'), *afsose* 'to repent' (*ʾāsif* 'repenting') are the most obvious examples. In Bangladesh, as in many other parts of the world, the distinct choice of words is more to do with a political/religious/social inclination than with the level of education. It also reflects attitudes inherent in religious traditions towards texts. Thus, the use of Arabic words in daily discourse reflects the socialization process of the users. This is seen in the greetings of *salaam aleykum* 'Peace be unto you' and the reply *wa aleykum as-salaam* 'Unto you also peace', the choice of names (Mohammed, Tanvir, Khaleda, Fatema), and the names of family members *abba* 'father' and *amma* 'mother'.

BIBLIOGRAPHICAL REFERENCES

Shahidullah, Muhammad. 2002. *Bangla bhasar ittibritta*. Dhaka: Mowla Brothers.

ALI RIAZ (Illinois State University)

Bedouin Arabic

1. INTRODUCTION

Bedouin Arabic is the colloquial Arabic spoken by speakers who are or consider themselves to be of Bedouin origin. Bedouin dialects have been divided into nomadic and semi-nomadic groups, according to their way of living and its effect on their dialects. Bedouin speakers, whether nomadic or not, constitute one of the two major demographically determined dialect groups of colloquial Arabic, the other one being sedentary dialects. In many regions, however, contemporary dialects are mixtures of Bedouin and sedentary dialects, and the distinction between Bedouin and sedentary dialects is not always clear.

Bedouin dialects exist in almost all regions of the Arabic-speaking world. The classification of Arabic dialects into Eastern and Western cuts also across this group. Accordingly, Bedouin dialects share certain common features, but also differ between the Eastern and Western regions. At present Bedouin dialects are considered more conservative than sedentary dialects, since they retain many 'Classical' features, lost elsewhere.

Since the beginning of the 20th century Bedouin tribes' free roaming has decreased immensely. The governments of independent states in the Middle East and North Africa forbade them to cross political borders and encouraged (even forced) their settlement. Bedouin sedentarization was, however, already described as a natural and spontaneous development by Ibn Xaldūn in the 14th century C.E.

Although Bedouin usually keep their own social framework, tribal history, and pride, as well as their language, modern life (employment and intermarriages with sedentary wives) affects their dialects. In such communities, Bedouin dialects reveal many effects of borrowing, code mixing, koineization, and leveling.

1.1 *Sub-groups of Bedouin dialects*

1.1.1 Arabian Peninsula dialects

These are the North Arabian dialects, Hijazi dialects, Southwest Arabian dialects, and Omani dialects.

The North Arabian dialects can be divided into three groups: A. ʿAnazī, B. Šammarī, and C. Syro-Mesopotamian dialects. Cantineau (1936–1937) also described a mixed group (BC). Speakers of dialects A and B are usually camel-herders, whereas the third group includes mainly sheep herders. The tribes of groups A and B roam in the area of the peninsula. Group C tribes are found mainly in the Syrian Desert and in Jordan, although some Šammar type tribes also live in Jordan (e.g. Banī Ṣaxr, or Rwala).

Hijazi dialects have been studied in few papers but certain dialects of the Negev and Sinai are apparently related to them. Southwest Arabian dialects include Yemen, Aden, Hadramawt, and Dhofar. Among Omani dialects there are also Bedouin tribes.

1.1.2 Iraq and Iran

Bedouin *gələt* dialects in Iraq and in southwest Iran (Khuzistan) have been studied and described. These dialects are rather old in the region, dating back to about the 13th century C.E.

1.1.3 Sinai

This area can be considered a link between dialects of the Negev, northeastern Egyptian and Hijazi Bedouin dialects (de Jong 2000).

1.1.4 Egypt

Bedouin dialects in Egypt are found to the east and the west of the Nile. As Egypt is linguistically very complex, recent studies have revealed ex-Bedouin dialect features and mixtures of such dialects.

1.1.5 Western Bedouin dialects in North Africa

These are found in Libya, Tunis, Algeria, and Morocco. They originated in several waves of migration from the Arabian Peninsula starting in the 9th century (Banī Sulaym, Banī ʿUqayl, Banī Hilāl) and some are influenced by local Berber dialects.

1.1.6 Sub-Saharan Bedouin dialects

There are Bedouin dialects found in Sub-Saharan areas, e.g. Sudan, Chad, Nigeria, and Mauritania, whose speakers arrived there apparently within the last 300 years.

2. LINGUISTIC DESCRIPTION

2.1 *Phonology*

2.1.1 Consonants

**q* is a phoneme that marks Bedouin dialects, being articulated as a voiced stop /g/, and an affricate /j/ or a fronted (palatalized) /gʸ/. In some dialects /q/ is used, however, as substitute for /ġ/. **k* is articulated in certain Eastern Bedouin dialects as a fronted affricate /č/ or /ć/. This phoneme, too, is a marker of Bedouin dialects. The affricates /č, j/ occur in the Syrian Desert fringe, in Jordan, northern Israel, and northwestern Arabia; the allophones /g, k/ occur, e.g. in the Hijaz, the Negev, Sinai, and North Africa; and the affricates /dᶻ, ć/ occur in inner Arabian dialects, such as the Šammar dialects.

**j* reveals preference for a voiced affricate articulation. In the Šammar dialects in the Arabian Peninsula, Jordan, and Sinai a fronted (palatalized) /dʸ/ is found. In certain Omani and Bahraini dialects this phoneme is palatalized and fronted up to /y/.

The interdental fricatives /ṯ, ḏ, ḏ̣/, are retained in Bedouin dialects, unlike many urban sedentary dialects.

The glottal stop /ʾ/ usually undergoes processes similar to sedentary Arabic dialects, often becoming a semivowel /w, y/ in the adjacency of the matching vowels (*u*, *i*, respectively). It is also sometimes used in pause forms, unlike sedentary dialects.

The traditional emphatic stops and fricatives remain in Bedouin dialects and are expanded to include new emphatics, such as /g̣, ẉ, ḅ, ṃ, ḷ, ṛ, f̣/; e.g. [g̴ɑːɫ] 'he said'. Emphasis often spreads also to adjacent non-emphatic homophonic phonemes, e.g. [g̴i̴b̴ɑð̴ ~ g̴i̴d̴ɑb̴] 'he seized'.

New foreign phonemes, such as /v, p/, occur in words borrowed from foreign languages such as Persian and Turkish, and nowadays also English in the East, and from Berber, French, or Spanish in the Western dialects.

Table 1 presents the common inventory of consonants and vowels in Bedouin Arabic. (Parentheses mark local allophones or foreign phonemes.)

2.1.2 Vowels

The basic system includes three short and long vowels /i, ī, u, ū, a, ā/. Two new short and long

Table 1. Consonants and vowels of Bedouin Arabic

Feature	Bilabial	Labio-dental	Alveolar	Palatal	Velar	Uvular	Laryngeal	Glottal
Stops	(p) b		t d ṭ	ḍ	k g	(q)		ʾ
Fricatives		ṯ ḏ f (v)	s z ṣ (ḏ̣) ẓ	š ž	x	ġ	ḥ ʿ	h
Affricate			(ŧ) (d^{z}) (d^{y})	č j				
Nasals	m		n					
Trill			r					
Lateral			l					
Semi-vowel			(r)	y				
Vowels	Short i, e, a, o, u	Long ī, ē, ā, ō, ū	Anaptyctic ə (with various timbres)	Diphthongs ay, aw, iw, aw, iy, īye uw				

vowels /e, ē, o, ō/ have been added to the system in most of the Eastern and in some Western dialects. The short and long vowels are not always phonemic, however (e.g. *gōl* 'a saying' vs. *gūl* 'say!', but *kīf* ~ *kēf* 'how?'). In certain Western Bedouin dialects phonemic differences of vowels are neutralized to two or one, although usually Bedouin dialects in this region are more conservative than sedentary ones. In this region, there are more cases of high vowel (/i, u/) neutralization by merging or mixture with /e, o/, respectively, than in the Eastern dialects, although it occurs there as well.

A non-phonemic anaptyctic vowel /ə/ can be added to the list. It resolves consonant clusters within words and between word junctures (sandhi). This vowel may acquire various central allophonic timbres depending on adjacent consonants.

Diphthongs are usually monophthongized (e.g. *sēf* ~ *sīf* 'sword', *yōm* ~ *yūm* 'day', *zēt* ~ *zīt* 'oil'). Sometimes word-final diphthongization is added (e.g. *tuguʿdiy* 'you [sg. fem.] sit', *yuguʿduw* 'they [pl. masc.] sit'), or kept (e.g. *yimšuw* 'they [pl. masc.] walk'). The diphthong /aw/ is retained when the /w/ is a C_1 root consonant as in *mawʿūd* 'promised'. The diphthong /ay/ remains in certain dialects in the dual morpheme (e.g. Negev *šaharayn* 'two months'). The diphthong /iw/ occurs sometimes under specific conditions such as roots with *iw* following the prefix /yi-/, as in *yiwṣal* 'he arrives').

2.1.3 The syllable

Syllable types in Bedouin dialects are similar to those of Eastern sedentary dialects, but differ from patterns in sedentary North African dialects. The basic patterns include Cv, Cv̄, CvC, Cv̄C, CvCC, CCv(C) in words such as these: *ki-tab* 'he wrote', *bin-tiy* 'my daughter', *daḥ-ḥag* 'he saw', *gō-ṭar* 'he went', *gumt* 'I got up', *bint-kam* 'your [pl. masc.] daughter', *gha-wa* 'coffee'. Word-initial clusters are usually avoided by a prosthetic vowel, e.g. *aštaġal* 'he worked'. They are allowed when the clusters are due to morphophonetic rules such as the → *gahawa* syndrome (see 2.1.5).

Vowel shifts may change potentially illegal into legal syllables (e.g. *yiḥ-li-bu(w)* > *yiḥlbu(w)* > *yi-ḥil-bu(w)* 'they milk [a cow]').

2.1.4 Stress

Stress rules in Bedouin dialects are similar in Eastern and Western dialects, unlike Eastern and Western sedentary dialects. Word stress is sometimes phonemic: *fíhim* 'understanding [noun]' vs. *fihím* 'he understood' (Negev). Stress placement depends on syllable structures. The basic rule is: at the end of the word over-long syllables (Cv̄C, CvCC) are stressed, e.g.: *ṣayyād* 'hunter', *ṣayyadīn* 'hunters', *yeʿarfūn* 'they know', *giṭaʿt* 'I cut'. But Bedouin dialects vary as to the possibility to stress word-final long or short syllables (CvC, CCv, Cv or Cv̄) syllables. Thus, for instance, *jimál* and *jímal* 'camel' occur in different Eastern or Western Bedouin dialects.

This fact yields another feature typical of many Bedouin dialects: stress on the definite article and many enclitics and verb prefixes when the enclitic is followed by a short syllable (Cv), e.g. *álʿajiy* ‘the boy’, *áštaġal* ‘he worked’, *yíntegil* ‘he moves away’, *énjeraḥ* ‘he was wounded’.

2.1.5 Phonotactics

The → *gahawa* syndrome, i.e. the insertion of /a/ in a cluster . . . aLC . . . (L = *x*, *ġ*, *ḥ*, *ʿ*, *h*) as in **qahwa* > **gahawa*, is perhaps the most typical Bedouin feature. As in other sequences of *CvCvCV, subsequent elision of the first vowel leads to an initial consonant cluster: **gahawa* > *ghawa* ‘coffee’, **raqaba* > *rguba* ‘neck’, **ġanamak* > *ġnimak* ‘your [sg. masc.] sheep’.

→ *ʾImāla* (/a/-raising). Word-final *ʾimāla* in feminine endings, often found in Eastern sedentary dialects, is usually lacking in Bedouin dialects. Word-internal *ʾimāla* occurs in Bedouin dialects near front consonants, e.g. *kitab* ‘he wrote’, *jimal* ‘camel’; in emphatic or back environments this vowel is /u/ as in *buṣal* ‘onions’.

In, for example, Negev, Sinai, and some Syrian desert fringe dialects /a/ in bound pronoun feminine endings and in certain adverbs is also raised in non-pause utterances such as *bēthiy* ‘her house’, *hniy* ‘here’. *ʾimāla* of long vowels hardly occurs in Bedouin dialects, e.g. *albāriḥ* ‘yesterday’ (cf. sedentary *embēriḥ*).

2.1.5.1 Assimilation and dissimilation

The assimilation of ‘sun letters’ to the preceding definite article *al-* is the general assimilation feature in Bedouin dialects. In certain Eastern Bedouin dialects new sun letters appear: /j, č/.

Other assimilation or elision cases in Bedouin dialects often involve preformatives or affixes and root consonants, as in *gult-lo* > *gut-lo* ‘I told him’, *ṣār-lēna* > *ṣal-lēna* ‘it happened to us’, *yit-jawwaz* > *yijjawwaz* ‘he marries’, *minšān* > *mišān* ‘because’, *nrūḥ* > *irrūḥ* ‘we go’, *tṣīr* > *ṣṣīr* ‘she becomes’, *ridt* > *ritt* ‘I wanted’, *axadt* > *axatt* ‘I took’, *nbūg* > *mbūg* ‘we cheat’, *tzīd* > *dzīd* ‘she adds’, *šēxhum* > *šēxxum* ‘their sheikh,’ *niftaḥhiy* > *niftaḥḥiy* ‘we open it,’ *miʿhiy* > *miḥḥiy* ‘with her’ (de Jong 2000).

Dissimilation, though not very frequent, can be found in cases of geminate, homorganic, or similar consonants, e.g. *fanajīl* ‘cups’, *dilim* ‘dunum’, *mzallaṭ* > *mzalbaṭ* ‘naked’, *daḥḥaj* > *daḥraj* ‘he looked at’ (Rosenhouse 1984).

2.1.5.2 Pausal forms

Many Bedouin dialects are noted for end-of-utterance pause forms. These may take the form of an added glottal stop /ʾ/, e.g. *laʾ* ‘no’ or the voiced laryngeal (*laʿ* ‘no’). Other pausal ‘devices’ delete final consonants such as /t/ of the feminine plural suffix *-āt*, exchange the /t/ with /y/ (*-āy*), drop the sonorants /l, m, n, r/, add a final aspiration *-h* as in *jah* ‘he came’, or add final semivowels as in, for example, *xuṣūṣiy* ‘special,’or *eġnimow* ‘his sheep’.

2.1.5.3 Sandhi

A final vowel may be deleted in sandhi when the following word begins with another vowel. In such cases, /a/ is retained and other vowels are deleted; cf. *ʾani afakkar* > *ʾan-afakkar* ‘I think’, *mā iḥna* > *ma-ḥna* ‘we are not’, *widdo yīži* > *d-īži* ‘he wants to come’. Another sandhi feature is /ʿ/ > /ḥ/, as in *maʿ ḥsēn* > [maħːseːn] ‘with Ḥsēn’. See 2.1.2 and 2.1.6.1 for anaptyctic vowels.

2.1.5.4 Metathesis

Consonantal metathesis occurs sometimes in Bedouin dialects, cf. for example, *daḥḥag* (Classical Arabic *ḥaddaqa*) ‘to look at’, *giḏab* (Classical Arabic *qabaḍa*) ‘to catch, hold’, *yagba* ~ *yigba* (besides *yabqa*) ‘he remains; so then’. etc.

2.1.6 Morphophonology

2.1.6.1 Elision and insertion of vowels

Short vowels are deleted when morphemes are added to the unmarked base-form, e.g.: *saxal* – *sxala* ‘lamb – ewe’, *širib* – *širbit* ‘he – she drank’, *gāʿid* – *gāʿda* ‘sitting [masc.] – [fem.]’.

The difference between Bedouin and sedentary dialects is mainly in the place of the vowel that remains in the end form, after deletion or insertion of an anaptyctic vowel. Mainly high and unstressed vowels are elided. Long /ā/ is usually retained and short /a/ usually remains when other morphophonological rules do not precede this rule and affect the occurrence of /a/. Anaptyctic vowels are added to resolve resulting consonant clusters (reshuffling), e.g. *yuḏ̣rubuw* > *yuḏ̣rbuw* > *yuḏ̣urbuw* ‘they beat’, but this also occurs in certain sedentary dialects.

Typical examples of elision are: *bētk* ‘your [sg. masc.] house’ (Cantineau’s group B) vs. the more prevalent *bētak*; *ghawa* ‘coffee’ (vs. sedentary *qahwa*) of the *gahawa* syndrome, and participle forms such as: *mintišrih* 'widespread [sg. fem.]' < *mintaširih*.

2.1.6.2 Construct state and suffixation

The construct state (annexation) is more prevalent in Bedouin dialects than in sedentary ones. In both groups genitive particles also exist as analytical substitutes to annexation. In Bedouin dialects such particles are, for example, *ḥaggi*, *šēti*, *šuġli* 'mine, my'.

As usual in Arabic, in words ending with the feminine ending *-ah*, *-t* is suffixed in annexation and before suffixed pronouns, e.g. *šjarat attuffāḥ* 'the apple tree', *mratak* 'your wife', and the preceding vowel /a/ is deleted under certain conditions, e.g. *nʿajtak* 'your ewe'. Under other conditions, an anaptyctic is inserted instead of the deleted /a/, e.g. *nʿajitkam* 'your [pl. masc.] ewe'. The structure *abūh la-jāsim* 'Jāsim's father' also exists in Bedouin dialects, as in some sedentary dialects.

When a suffix is added to a feminine singular active participle the vowel preceding the *-t* of the feminine ending often appears as long, e.g. *mšagfī-to* 'having cut [sg. fem.] it into pieces'.

In those Bedouin Galilee dialects and Cantineau's group C dialects where the masculine plural verb suffix is *-am*, this suffix is changed into *-ū* before a bound pronoun, e.g. *libistam* 'you wore' > *libistūha* 'you wore it [sg. fem.]'.

Instead of a double suffixation (a direct object followed by an indirect object), the inflected accusative preposition *iyyā-* is used.

2.2 *Morphology*

2.2.1 Pronouns

2.2.1.1 Personal pronouns

Table 2 presents a representative list of independent and bound pronouns in Bedouin Arabic.

Bedouin Arabic, unlike many urban dialects, retains gender difference in the singular and plural forms of independent and bound pronouns. Certain Bedouin Eastern and Western dialects show *'imāla* in the 1st pers. sg. (*ani* ~ *ane*), the 1st pers. pl. in Western Bedouin dialects (*ḥnē*), and even the bound pronoun of the 3rd pers. sg. fem. (*-hiy*).

In Western Bedouin dialects the influence of local (Berber) languages can be found in forms ending with *-ya* or beginning with *l-* (e.g. *leḥna* 'we'), and in Fezzani Bedouin dialect, for example, one finds *huṃṃ^{w}a*, with rounding after the geminate /ṃ/.

For the 2nd pers. there are variations in the vowel timbre both preceding and following the consonants *-nt-*, cf. forms like: *int*, *inet*, *ənta*, *önt*. The plural form may end in a consonant or be expanded by a following vowel (e.g. *entu*, *entum*, *entam*, *intuw*) or syllable (*entūma*) in the North African area. The feminine form is usually *inti*, *intiy* for the singular, and *intan* for the plural.

The bound pronouns also vary according to dialects, often related to the distinction between 'camel breeders' and 'sheep herders'.

Following a noun ending in a vowel, the 1st pers. sg. bound pronoun may get a /y/ as 'consonantal skeleton' before the suffix vowel which is /i/ or /a/.

Following a verb, the direct object suffix is not only the usual *-ni* but also *-an*.

For the 2nd pers. sg. also the bound pronouns diverge between 'camel breeders' and 'sheep herders', e.g. *betk* ~ *bētak* 'your [sg. masc.] house', *bētkiy* ~ *bētek* ~ *bēteč* 'your [sg. fem.] house', respectively. The suffix consonant is *k*, *č*, or *ć* according to the dialect. When any word ends in a vowel, this vowel is lengthened before the suffix, e.g. *šafūk* – *šafūč* 'they saw you [sg. masc. ~ fem.]'.

For the 3rd pers. sg. masc., usually only a lengthening of the final vowel of the word indicates the suffixed pronoun. The vowel can be high /u, o/ or low /a, e/. Sometimes a weak /h/ can also be heard in this case. In the 3rd pers. sg. fem., the suffix is usually *-ha*, but in some dialects also *-hiy*.

2.2.1.2 Relative pronoun

The uninflected form of the relative pronoun for both genders and numbers in Bedouin dialects is typically *alli* with initial *a-*. In Western Bedouin dialects however, a form such as *eddi* also occurs.

2.2.1.3 Demonstrative pronouns

The demonstrative pronouns for near objects include short forms and long forms; see Table 3. The interdental /ḏ/ in these forms is usually emphatic (velarized) except for the feminine forms, where it is followed by the high /i/. Often also a double demonstrative structure occurs, with the noun between a preceding short pro-

Table 2. Independent and bound pronouns

Independent		Bound		
		Following a noun – consonant	Following a noun – vowel	Following a verb
3rd sg. masc.	*hū*, *huwwa*	*-u*, *-o*, *-uh*, *-ih*, *-a*, *-e*, *-ah*	*-u*, *-o*, *-uh*, *-ih*, *-a*, *-e*, *-ah*	*-u*, *-o*, *-uh*, *-ih*, *-a*, *-e*, *-ah*
3rd sg. fem.	*hī*, *hiyya*, *hīye*	*-(h)a*, *-hiy*	*-ha*	*-ha*, *-he*, *-hiy*
3rd pl .masc.	*hum*, *humme*	*-hum*	*-hum*	*-hum*
3rd pl. fem.	*hin*, *hinna*	*-hen*, *-hin*	*-hen*, *-hin*	*-hen*
2nd sg. masc.	*ʾint(a)*, *ʾinet*, *ənta*, *ənte*, *önt*, *önte*	*-ak*	*-k*	*-ak*, *-k*
2nd sg. fem.	*inti*, *intiy*	*-ek*, *-eč*, *-ič*, *-eć*, *-ić*	*-ki*, *-kiy*	*-eč*, *-ki* *-č*, *-k*
2nd pl. masc.	*entu*, *intuw*, *entum*, *entam*, *entuma*	*-kum*, *-kam*, *-kom*	*-kum*, *kam*, *kom*	*-kum*, *kam*, *kom*
2nd pl. fem.	*intan*, *intin*	*-ken*, *-čin*	*-ken*, *-čin*	*-ken*, *-čin*
1st sg.	*ʾane*, *ʾani*, *ʾanī*	*-i*, *-ni*	*-ya*, *-yi*, *-ay*, *-ayi*	*-ni*, *-an*
1st pl.	*əḥna*, *nəḥna*, *ḥnē*, *leḥna*, *iḥna*	*-na*	*-na*, *-ne*	*-na*, *-ne*

noun and following long pronoun, e.g. *ha-lʿajiy hāḏ̣a* ‘this boy’. Though this structure is also found in some sedentary dialects, it is more characteristic of Bedouin dialects.

Table 3. Demonstrative pronouns

	Near object	Far object
sg. masc.	*ḏ̣a*, *hā*, *hāḏ̣a*	*haḏ̣āk*
sg. fem.	*ḏi*, *hāḏi*	*haḏīk*, *haḏīč*
pl. masc.	*ḏōl*, *haḏōl*	*haḏ̣lāk*
pl. fem.	*hāḏ̣an*	*haḏīkan*, *haḏīčan*

2.2.2 Nouns and adjectives

Noun and adjective patterns do not differ from those in other dialects. Distinctively Bedouin is the frequent use of the diminutive, e.g. *ṭfēl* ‘child’, *ʿnīzih* ‘goat’, *wlīd* ‘boy’, *mwēmti* ‘my small mother’ (a mother’s address to a child), *gṣayyir* ‘short’, *gḷayyil* ‘small’. The *gahawa* syndrome affects nouns and adjectives, as noted, e.g. in *ʾaḥamar* ‘red’, *ʾaʿama* ‘blind’ of the colors and defects patterns. Certain lexical items differ, however, from those in sedentary dialects; e.g. *muġar* ‘caves’, *burgān* ‘coffee pots’, *ʿağawīn* ‘boys’.

2.2.3 Numerals

Special Bedouin numerals include: *ṯnēn*, *ṯintēn*, and *žūž* ~ *zōz* ‘two’ in the Western region; *flāfa* for *ṯlāṯa* ‘three’ in Western Algerian dialects, *samḥa* for *sabʿa* ‘seven’ in the Eastern dialects, *famn-ayyām* ‘eight days’ in the Mzāb (Algeria), *tel*^{e}*t*, *tem*^{e}*n*, *xam*^{e}*s* (three, eight, five) (Marazīg, Tunisia).

The functioning of the morph *-ʿšar* in the numbers 11–19 is as in sedentary dialects. Also sedentary *-ʿšān*, *-ʿšāl*, or *-aš-el-* (with the definite article) of the region occurs in Western Bedouin dialects. *ʾīdaʿš* ‘eleven’ occurs in some Eastern dialects. ‘One hundred’ is *miyye*, and in some places *imya*.

Ordinals take the usual CāCiC pattern, e.g. *ṯāni* [masc.], *ṯānie* [fem.] ‘second’. ‘First’ is *ʾawwal* [masc.], *ʾawwala* [fem.] and in Western dialects *owwal*, *owwul* or *ūl*. ‘Sixth’ is *sādes* [masc.], *sādse* [fem.] in the Eastern dialects, and *sātet*/*sātt* in the Western dialects.

Noteworthy fractions are *nufṣ* and *nfēṣ* for ‘half’ (in Western Bedouin), usually *nuṣṣ* and *noṣf* elsewhere. ‘One seventh’ is *sebḥ* in some dialects.

2.2.4 Verbs

2.2.4.1 Forms and conjugations

Bedouin verb inflections of the perfect and the imperfect comprise three persons, two numbers, and two genders. Inflection of weak or geminate roots basically follows the usual rules of Arabic.

Distinct Bedouin verb forms are due to phonetic and phonotactic factors mentioned above, including *gahawa* syndrome, /a/ raising, and vowel backing, e.g.: *ḥalab/yḥalub* 'he milked, will milk', *ḏ̣urab* 'he hit', *simaʿ* 'he heard', *rḥal* 'he moved', *gāl/ygūl* 'he said, will say', *buga/yibga* 'he remained, will remain', *nseit* 'I forgot', *maṣṣēt* 'I sucked'. Some dialects, mainly in the West, stress the final syllable, which may lead to the loss of the vowel following C_1, e.g. $^{\ddot{o}}$*rčeb* 'he rode', *jhel* 'he did not know' (Cantineau 1936).

In many Bedouin dialects (B, C groups, etc.), the imperfect prefixes are similar to those of Classical Arabic in usually using /a/ instead of the typical sedentary /i/. In some Western dialects this preformative changes according to the vowel of C_2 (vowel harmony), e.g. *yafham* 'he understands', *yomṭor* 'it rains', *yekteb* 'he writes' (Libya).

The suffixes of the 2nd pers. sg. fem., and the 2nd and 3rd pers. pl. masc. of the imperfect form three main groups respectively: (a) those ending in *-īn*, *-ūn*, as in Classical Arabic; (b) those ending in *-i*, *-um/-am*; (c) and those ending in a vowel or diphthong, i.e. *-i(y), -u(w)*. The last group seems to be the most widespread in all the areas. The group with *-īn*, *-ūn* is apparently next in frequency, being found in Western dialects, Eastern Arabia, and in the North of Israel. The last group, with *-i*, *-um/-am*, is the least frequent.

2.2.4.2 Derived Forms

The derived verb Forms use morphemes of several types, as elsewhere in Arabic: gemination of C_2 (Forms II, V); prefixation of *ʾa-* (Form IV) or of *ta-*, *t-* (Forms V, VI); lengthening of the vowel following C_1 (Forms III, VI); prefixing *n-* (Form VII) and *st-* (Form X), often with *a-* as an initial vowel (*an-*, *ast-*); infixing *-t-* (Form VIII), and geminating C_3 (Form IX). Mauritanian Ḥassāniyya dialects innovated the pattern saCCaC. Patterns CCāC, tCCaC also occur in sedentary Western dialects.

Mainly in Arabia but also in other Bedouin dialects internal passives may occur, e.g. *kisar* 'it got broken' (Galilee), *yūšra* 'it will be bought' (Fezzan). The passive of Forms II, III, and saCCvC in the Ḥassāniyya dialects begins with *u-*, e.g. *ubaxxar* 'it was perfumed by incense', *usagbal* 'it was led toward the south'. Elsewhere, the passive is mainly expressed by Form VII *anmasak* 'it was caught', Form V *tīḥakkam* ~ *taḥakkam* 'it was treated medically', or by *t*-passives from Form VIII, e.g. *atwagad* 'it was found'.

After Form I, Form II is the most frequent, as elsewhere. Form IV is used in Bedouin dialects more than in sedentary Arabic, although Form II has generally taken over the causative function.

2.2.4.3 Participles

Active and passive participles are semi-verbal and semi-nominal being inflected to gender and number, and possibly governing an accusative. The active participle functions (a) to express an ongoing (durative) action, when the verb belongs to the verbs of motion, state in space, or senses (e.g. to come, sit, lie down, see, hear) and (b) to express a perfect tense, e.g. *nāyim* 'sleeping, asleep', *mrabbi* 'he has raised, educated (someone)'.

Negev Bedouin and some other dialects use the participle *kāyin* 'there is' (in addition to *kān* 'he was') in narratives, unlike other Bedouin dialects, which hardly ever use the verb in this function (Henkin 1992).

The passive participle refers to a completed action, the result of an action, or a situation, e.g. *maftūḥ* 'open(ed)'. Passive participles of derived Forms (above II) are, however, not frequent.

2.2.4.4 Imperatives

A typical feature of Eastern Bedouin imperatives in roots of verbs III*w/y* is the deletion of the final vowel, yielding, e.g., *imš!* ~ *imiš!* 'go [sg. masc.]!'.

2.2.4.5 Special verbs

The verbs *ʾaxad* 'he took' and *ʾakal* 'he ate' characterize different dialects across the Arabic-speaking world. In these verbs the first syllable (vowel) may be unstressed or entirely lost when the main stress falls on the final syllable. Thus, we find in Bedouin dialects: *ʾakál*, *kála*, *klá* 'he ate'. Rwala dialects (in Arabia) have *ʾ* > *h*, yielding *hökal* 'he ate'. The preformative vowel + C_1 *ʾ* in the imperfect of these verbs is exchanged with a long vowel which may be /ā/ or /ō/ and the vowel following C_2 may be /e/ or /u/, yielding four forms. The verb *jāʾa/ʾaja* 'he came' loses the syllable *ʾa-* in the past in many dialects, thus: *ja*, *je* 'he came', *jam* 'they came', etc.

2.2.4.6 Verbal nouns

Verbal nouns of the derived patterns in Bedouin Arabic are similar to those of sedentary dialects and Old Arabic (→ verbal nouns).

2.3 *Syntax*

An important Bedouin feature is what seems to be a residue of Classical Arabic case endings (*tanwīn*), used under special conditions.

Word order in Bedouin dialects is as elsewhere, either SVO or VSO. In narratives, the subject is often left unmentioned.

A feature that seems especially typical of Bedouin Arabic narrative style though found also in sedentary dialects, is a rhetorical question (with rising intonation), immediately followed by the repeated question (echo question) and its answer, in a statement intonation.

2.3.1 The noun phrase

Syntactically, noun phrases in Bedouin Arabic do not differ from sedentary dialects. The noun may be defined by the article *al-*. The numeral *wāḥad* [masc.], *waḥade* [fem.] 'one, someone' can be considered an indefinite article in certain contexts (→ article, indefinite). An important Bedouin feature is what seems to be a residue of case endings (*tanwīn*). These endings are used under special conditions, mainly in frozen formulae and adverbs, poetry, and noun phrases between a noun and its adjective. The vowel of the *tanwīn* suffix is usually *-en*, *-in*. Some researchers consider this → *tanwīn* a 'specified indefinite.

The construct state is preferred to analytical genitive structures in Bedouin Arabic.

The agreement of adjectives to inanimate plural noun heads fluctuates between the feminine singular and the (masculine or feminine) plural. e.g. *byūt kbār* ~ *kbīra* 'big houses, *ʿajiyyāt farḥānāt* ~ *farḥānīn* 'happy girls'.

2.3.2 Noun phrase negation

The particle *mī* (< *mā hī*) or *mū* (< *ma hū*) 'not', usually negates the noun phrase, though in some dialects *miš* is also used, not only as a koine form. In some dialects the negated pronoun is suffixed to the negative particle, thus e.g. *mint* 'I am not', *mintu* 'you [pl. masc.] are not', *muhū* 'he is not'. Also *mā-hū-b* or *mū-b* are frequent in Saudi Arabian Bedouin dialects. Participles, including such as *lāzim* 'must', *yimčin* 'maybe, possibly', etc., are also considered in this group, since they are negated by *mū* and not by *mā*. The structure *mā mišš* (Johnstone 1967) negates existence.

2.3.3 The numeral phrase

For the singular, the singular noun is used; it may be followed by *wāḥad* as an adjective, e.g. *zalame wāḥad* 'one man'. When the number 'one' precedes the noun it is used as indefinite article, thus *wāḥad zalame* 'someone, a man'.

The dual is often used for 'two' in formulas (*yirudd assāʿ sāʿēn* 'retaliate twice'), as → pseudo dual (e.g. *ḏanēn* 'ears'), and for real dual (*sanatēn* 'two years', *ḥajarēn* 'two stones'). The numeral may also follow the plural noun (*wlād tnēn*, *waladēn tnēn* 'two boys').

The count noun for the numbers 3–10 is in the plural form (*ṯalaṯ imtār* 'three meters'). Nouns in iCCāC pattern may be linked to the numeral by *-t*, *ṯalaṯ-t-iyyām* (cf. Classical Arabic). Above 'ten', the noun is always in the singular, e.g. *arbaʿṭaʿšar yōm* 'fourteen days'.

2.3.4 Relative clauses

Relative clauses usually follow the relative pronoun *alli*. The Classical Arabic rule about deleting the relative pronoun when the antecedent is indefinite, is not always observed.

2.3.5 The verbal phrase

2.3.5.1 Verbal tense, mood and aspect

Bedouin Arabic is basically similar to other Arabic dialects as regards tense, mood, and aspect. Innovations in the Bedouin Arabic expression of moods and aspects are mainly in the lexical domain. The 'narrative imperative', used to enliven the narration when referring to past actions, is described as typical of Bedouin Arabic, although it is also found elsewhere.

2.3.5.2 The indicative mood

The indicative is not usually marked in contrast with the subjunctive in Bedouin dialects. The prefix *b* ~ *bi-* for the indicative is found, for instance, in eastern Egypt, Sinai, and the Negev, but does not usually exist elsewhere (unlike Eastern sedentary dialects). In Arabia the prefix *b* ~ *bi-* indicates the indicative future and originates from a different root (*ʾaba* or *baġa* 'want'). A similar particle is used, for instance, in Fezzan (Western Bedouin).

2.3.5.3 Modal verbs

kān (or *čān*) 'to be', is used to add tense to nominal predicates and aspects (existence, habit,

Table 4. Typical Bedouin lexemes

Class	Gloss	Lexeme
Nouns	hand, handle	*yed(d)*
	nose	*manaxīr*, *xašm*
	mouse, lips	*barāṭim*, *šifāif*
	rain	*maṭar*, *šita*
	riding animals	*zimūl*, *ḏ̣ʿūn*
	camels	*álbil*
	hill	*dabbe*, *rwēsa*
	tent	*bēt šaʿr*
	ewe	*ġnima*
	knife	*xōṣa*
	spoon	*maġrāfe*, *miġrafa*
	good, fine men	*ʾajawīd*
	group, tribe, enemy	*gōm*
	tribe's warriors	*nišāma*
	hunting	*ganaṣ*
	host	*mʿazzeb*
	child	*ʿajiyy*
	children	*ʿiyāl*, *ḏ̣ʿūf*
	food for the journey	*zāda*, *zuwwāda*
Adjective	good	*zēn*
	bad	*šēn*
	small	*zġayyir*
	near	*grayyib*
	a little	*glayyil*
	merciful, kind	*rḥamān*
Verbs	to go	*gōṭar*
	to look	*daḥḥag*, *daḥraj*
	to kill	*kital*
	to go down	*hawwad*, *awṭa*
	to reach, enter	*ṭabb*
	to ask, look for	*nišad (ʿala)*
	to go eastward	*šarrag*
	to hunt	*gannaṣ*
	to shoot at, snipe	*gawwaṣ*
	to wed, marry legally	*ʿagad ʿgādo ʿalēha*
	to want, like, need	*widd*-(+bound pron.), *rād*
Adverbs	now	*(h)alḥīn*, *hassāʿ*
	tomorrow	*bāčir*, *ġodwa*
	here	*hina*, *hniy*,
	there	*hnūh*, *hnāk*, *ġād*
	much, strongly, very	*balḥayl*
	then	*dīma*, *hēḏīkt alḥīn*
	thus	*kiḏiy*
	maybe	*xāf-aḷḷah*
Particles	what?	*šnū*, *šnī*, *wēš*
	who?	*min*, *men*, *min-hu-minh*
	when?	*mita*, *mata*, *wagtēš*
	which?	*ʾayy*, *yāt(u)*, *wein*
	why?	*ʿalāma*, *ʿalāš* ~ *ʿalēš*, *ʿalawēš*
	when (temporal adverb)	*ʿandinma*, *lamman* ~ *lammin*
	and there (it was)	*wlinnih*

Table 4 (*cont.*)

Class	Gloss	Lexeme
	but	*ʾamma* ~ *ʾumma*, *mār*
	so, then	*ʿād*
	so, but	*mār*, *umār*, *mēr*
	and lo (a sudden turn in the narrative)	*wlinn*, *wilin*, *win*, *winno*
	(you are) not	*mint*
	here (he is)	*arʿ*, *harʿ*, *arʿīh*
	because	*ʿaminno*
	if	*iza*, *in*, *kān* ~ *čān*, *inkān* (± bound pronoun)

continuous actions, conditions, wishes, and imminence), *ṣār* 'to begin'. Modal motion verbs, e.g. *jā* 'to come', *rāḥ* 'to go', *naṭṭ* 'to jump', *gaʿad* 'to sit', *gām* 'to get up', vary between the Bedouin dialects, depending on their vocabulary (Ingham 1994).

2.3.5.4 → Pseudo-verbs

This group includes expressions for 'there is', which in Bedouin dialects are *bī*, *bū*, or *fī*. Other pseudo verbs are *ʿando* / *ʾilo* 'he has', *bʿado* '(he) still (is)', *widdo* 'he wants, needs'.

2.3.5.5 Verb negation

Verbs, including pseudo-verbs, are negated in Bedouin Arabic by the particle *mā*. The correlative *-š* suffix, often used in sedentary dialects, is rare in them. *la* + imperfect denotes inhibition or forbidding of an action.

2.3.6 Agreement/concord

The normal rules of Arabic concord prevail in Bedouin dialects. An adjective usually agrees with its noun headword in number and gender. When the subject denotes an animate group of people (whether morphologically singular or plural), there is a strong tendency for the verb or adjectival predicate to take the feminine singular as in, e.g., *ʾijat annās* 'the men came', *nās jelīla* 'few people'.

3. LEXICON

Many lexical items of daily life distinguish Bedouin from sedentary dialects. Differences are found in all the vocabulary domains, including items characteristic of obsolete Bedouin life. In narratives (folk-stories, legends, etc.) and poetry, certain formulaic expressions and Classical Arabic items are also characteristic. Some such typically Bedouin lexemes are listed in Table 4.

BIBLIOGRAPHICAL REFERENCES

Abu al-Fadl, Fahmi. 1961. *Volkstümliche Texte in arabischen Bauerndialekten der ägyptischen Provinz Šarqiyya*. Diss., Münster/Westfalen.

Behnstedt, Peter and Manfred Woidich. 1985. *Die ägyptisch-arabischen Dialekte*. II. *Dialektatlas von Ägypten*. Wiesbaden: L. Reichert.

Blanc, Haim. 1970. "The Arabic dialects of the Negev Bedouins". *Proceedings of the Israel Academy of Sciences and Humanities* 4:7.112–150.

Boris, Gilbert. 1958. *Lexique du parler arabe des Maraziq (Tunis)*. Paris: Klincksieck.

Cantineau, Jean. 1936–1937. "Etudes sur quelques parlers de nomades arabes d'Orient". *Annales de l'Institut d'Etudes Orientales* (Algiers) 2.1–119, 3.119–237.

Caubet, Dominique. 2004. "Les parlers arabes nomades et sédentaires du Fezzan d'après William et Philippe Marçais". *Approaches to Arabic dialects: A collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday*, ed. Martine Haak, Rudolf de Jong, and Kees Versteegh, 67–96. Leiden: E.J. Brill.

Cohen, David. 1963. *Le dialecte arabe ḥassānīya de Mauritanie*. Paris: Klincksieck.

Henkin, Roni. 1992. "'To be or not' in the Sinai Bedouin dialect of the Aḥaywāt". *Mediterranean Language Review* 6.173–197.

Ingham, Bruce. 1994. *Najdi Arabic: Central Arabian*. Amsterdam and Philadelphia: J. Benjamins.

Johnstone, T.M. 1967. *Eastern Arabian dialect studies*. London: Oxford University Press.

Jong, Rudolf de. 2000. *The Bedouin dialects of the Northern Sinai littoral: Bridging the linguistic gap between the eastern and western Arab world*. Leiden: E.J. Brill.

Kampffmeyer, Georg. 1899. "Materialien zum Studium der arabischen Beduinendialekte Innerafrikas". *Mitteilungen des Seminars für Orientalische Studien* 2.143–221.

Marçais, Philippe. 1977. *Esquisse grammaticale de l'arabe maghrébin*. Paris: Maisonneuve.

Maṭar, ʿAbd al-ʿAzīz. 1981. *Lahjat al-badw fī ʾiqlīm sāḥil Maryūṭ*. Cairo: Dār al-Maʿārif.

Musil, Alois. 1928. *The manners and customs of the Rwala Bedouins*. New York: American Geographical Society.

Palva, Heikki. 1980. "Characteristics of the Arabic dialect of the Ṣaxar tribe". *Orientalia Suecana* 29.112–139.

Rabin, Chaim. 1951. *Ancient West-Arabian*. London: Taylor's Foreign Press.

Rosenhouse, Judith. 1984. *The Bedouin Arabic dialects: General problems and a close analysis of North Israel Bedouin dialects.* Wiesbaden: O. Harrassowitz.

Socin, Albert. 1900. *Diwan aus Centralarabien.* Leipzig: Königliche Sächsische Gesellschaft der Wissenschaften.

——— and Hans Stumme. n.d. *Der arabische Dialekt der Houwâra in Wād Sous in Marokko.* Leipzig: Königliche Sächsische Gesellschaft der Wissenschaften.

Stumme, Hans. 1894. *Tripolitanisch-tunesische Beduinenlieder*. Leipzig: J.C. Hinrichs.

Wetzstein, Johann Gottfried. 1868. "Sprachliches aus den Zeltlagern der syrischen Wüste". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 22.69ff.

JUDITH ROSENHOUSE
(Technion – I.I.T. Haifa)

Bedouinization

1. INTRODUCTION

Sociolinguists dealing with Bedouin-sedentary Arabic contact identify several stages of transition from Bedouin to urban Arabic. Cadora (1992:1), for example, cites the following five stages:

Bedouin ⇒ Bedouin-Rural ⇒ Rural ⇒ Rural-Urban ⇒ Urban

This is a gradual process that takes place when nomadic tribes settle in sedentary rural or urban areas and level their speech patterns in the direction of the host communities' dialects. Linguistic leveling may be complete or partial, depending on how far the settlers accommodate to their sedentary environment. What has not been studied in depth, however, is the reverse process from sedentary to Bedouin transition, even though people have been known to move from a sedentary to a nomadic way of life (Ingham 1982:32). Arabic dialect studies, concentrating mainly on stable dialects, either Bedouin or sedentary, "have only seldom provided us with information about the direction of ongoing development processes" (Palva 1994:459).

→ Bedouin dialects have been generally well documented. A number of studies concentrate on the varieties spoken by nomads of the Middle East, providing substantial data on their salient features (Cantineau 1936, 1937; Rosenhouse 1984; Palva 1976, 1980, 1994; Ingham 1997; de Jong 2000). A small number deal with the dialects of North Africa, notably Cantineau's article on Arabic dialects in Oran province (1940). Bedouin Arabic, "the cornerstone of Modern Arabic dialects" (Rosenhouse 1984:3), has influenced sedentary speech varieties through continuous nomadic–sedentary contact in different Arabic-speaking regions (Blanc 1964; Palva 1976). This can be seen in the "many examples of . . . 'mixed urban dialects' incorporating features of both sedentary and Bedouin" speech (Miller 2004:183). Bedouin Arabic is broadly divided into two zones, eastern and western, with the line of demarcation running almost parallel to the Egyptian/Libyan border. In areas where the sedentary population is made up of ex-Bedouin elements, usually in the eastern zone, the dialects "retain more Bedouin features than when the demographical substratum is non-Bedouin or non-Arab, as in the Syro-Lebanese area" (Rosenhouse 1984:260) and the Maghreb.

Rosenhouse (1984:168–169) finds many similarities between north Israeli Bedouin dialects and the spoken varieties of the Lebanese littoral, namely Tyre, Sidon, "and sometimes also Beirut". This similarity can be found in Mount Lebanon where some Arabic dialects spoken by Christians show Bedouin influence, despite the fact that Christians sometimes speak a dialect distinct from their Muslim neighbors (Blanc 1964; Woidich 1997). In North Africa where many nomadic tribes are Berber-speaking, Bedouin influences can still be detected in Arabic dialects, due largely to the Arabization of the region between the 11th and 14th centuries when Bedouin tribes settled there (Singer 1994:273). In the Maghreb, Bedouin borrowings can be found even in sedentary dialects (Marçais 1977), and Eastern Libyan Arabic, described by Owens (1984), shows the effects of Bedouinization in the region (Owens 1984).

This entry considers various aspects of Bedouinization in modern Arabic dialects, ranging from Bedouin features that have infiltrated some sedentary dialects, to a brief survey of the process of Bedounization that has affected the urban Arabic of Bahrain and central Iraq. The entry also focuses on various Bedouin items in the contemporary Nabaṭī poetry of the Arabian Peninsula.

2. Bedouin influence on sedentary Arabic dialects

A characteristic feature of Bedouin speech is the realization of /g/ and /č/ as reflexes of Standard Arabic /q/ and /k/ respectively (Rabin 1951; de Jong, 2004). /g/ < /q/ can be found throughout the Arabic-speaking world. /č/ < /k/ is a characteristic of the dialects of the Arabian Peninsula, most of central and southern Iraq and eastern Syria. In Deir Ezzor, where the actual speech of the town is a *qəltu* variety (Blanc 1964), Jastrow (1978) found that /č/ also occurs there because of the presence of nearby Bedouin dialects.

/g/ is widespread in a number of Arabic dialects in regions where Standard Arabic usually has /q/ or the glottal stop /ʾ/, as in the Arabic of Cairo. /g/ for Cairene /ʾ/ occurs in the spoken varieties of the Nile Delta that "are strongly influenced by the dialects of Bedouins who arrived relatively recently in the region" (Woidich 1997:189). Woidich (1997:186) found, moreover, that although it is expected for a term like Cairene *ʾalb* 'heart' to be realized as *galb* in the Baḥariyya oasis, Standard Arabic *hamza* is also realized as /g/ there, as in the term *hayʾa* 'company', which becomes *hayga*. Egyptian urban Arabic /g/ (< Standard Arabic /j/) has reflex /j/ in Upper Egypt (Woidich 1997:187–188).

/g/ is common in Maghrebi dialects. It is often used interchangeably with /q/. Caubet (1993:12) cites as examples: *qṣəm/gṣəm* 'to share', *fūq/fūg* 'on', *wqəf/wgəf* 'to get up', and Ferré (n.d.) gives *qāl/gāl* 'to say', *qbīla/gbīla* tribe', *zləq/zləg* 'to slide along'.

Owens (1984:6) gives several terms with /g/, including *gabul* 'before', *gassam* 'he divided', *ugrūn* 'horns', *uṭlugat* 'she let go', *azrag* 'blue', showing its widespread occurrence in Libyan Arabic.

In Mount Lebanon /g/ does not occur in everyday speech, but is invariably used in popular songs of Bedouin origin as a reflex of /ʾ/ and Druze /q/, thus: *ʾalbi/qalbi* 'my heart' > *galbi*; *ʾabīli/qabīli* 'tribe' > *gabīli* (data collected in 2003).

Other phonemes that denote Bedouin influence are the interdentals. According to de Jong (2004:155) "most Bedouin dialects have three interdental phonemes /ṯ/, /ḏ/ and /ḏ̣/". In the urban dialects of Egypt and the Levant /ṯ/ and /ḏ/ do not usually occur, except in Modern Standard Arabic loans. However, in some urban areas in Algeria, such as Cherchell, Dellys, and Constantine, both /ṯ/ and /ḏ/ are used (Marçais 1977:9). /ṯ/ and /ḏ/ occur in Libyan dialects also. Owens (1984:7) gives a few examples, including *ṯigīl* 'heavy', *warəṯ* 'he has inherited'; *ḏēl* 'tail', *yikḏəb* 'he lies'. Interdental /ḏ̣/, the reflex of Standard Arabic /ḍ/ and /ḏ̣/ and Egyptian and Levantine /ḍ/ and /ẓ/, is a hallmark of the speech of the Arabian Peninsula, most of Iraq, and parts of Syria.

In some speech varieties the 2nd and 3rd feminine plural pronouns affixed to perfect and imperfect verbs display gender distinction, a characteristic of Bedouin Arabic. In urban Arabic these pronominal affixes are ordinarily of common gender, as can be seen in educated Gulf Arabic where *wugaftaw* 'you [pl.] stood', *tāklūn* 'you [pl.] eat', *gəʿdaw* 'they sat down', *yətərsūn* 'they fill' refer to both genders (Holes 1990:204–205).

Gender distinction occurs in Libyan Arabic, as in the following examples cited by Owens (1984:225): *ətšīlan* 'you [pl. fem.] carry', *ətšīlu* 'you [pl. masc.] carry'; *xəftan* 'you [pl. fem.] fear', *xəftu* 'you [pl. masc.] fear'; *ətzūran* 'they [pl. fem.] visit', *əzūru* 'they [pl. masc.] visit'; *xāfan* 'they [pl. fem.] feared', *xāfo* 'they [pl. masc.] feared'.

Bedouin lexical terms can be found in nearly all sedentary and urban dialects, even in varieties that have not evolved from Bedouin dialects. In Maghrebi Arabic, for example, the term 'to speak' is *hdər*. The derived noun *haḏr* or *hḏāri* 'idle chatter' is common to Bedouin Arabic. Many Maghrebi diminutive forms have equivalents in Bedouin Arabic, such as the Moroccan terms *ḥwīja* 'possession' and *dwīda* 'small worm' (Ferré n.d.). Similarly, *guṣṣa* 'fringe [hair]' (< *guṣṣa*), *məzyān* 'good, beautiful' (< *mazyūn*), and *rqad* 'to sleep' (< *rəgad*) are widespread in Maghrebi dialects.

Bedouin loanwords, moreover, can be found in some Christian sedentary dialects, such as those of Baskinta and Btiġrīn, two towns in the Matn region of Mount Lebanon. The following terms that occur frequently in the Arabic of the two towns are clearly Bedouin borrowings: *zalami* 'man' (< Bedouin *zələm* 'man/men'); *ḥəmi* 'privately owned pasture land' (< Classical Arabic *ḥimā* 'place of pasture prohibited to others'); and the two forms of camel terminology, *ḥəda* 'marching song' (< *ḥudāʾ* or *ḥidāʾ* 'to drive or urge camels by singing to them'), and *barak* 'to be bedridden' (< *baraka* 'to kneel [camel]').

ḥərmi and its variant *ḥərmāyi* 'woman', referring to an elderly village woman, are widely used in the dialects of Mount Lebanon.

3. Bedouinized dialects

Linguistic accommodation resulting in interdialectal forms are a common feature of a number of Arabic speech varieties (→ speech accommodation). Accommodation is often triggered by extra-linguistic phenomena, such as social status, ethnic or religious affiliation, as Holes (1980 1983) has illustrated in his works on Sunni/Shiʿi speech variation in Bahrain. In Baghdad and neighboring central Iraqi towns dialect variation involves a Muslim/non-Muslim dichotomy (Blanc 1964). Within Muslim Arabic, too, there tends to be variation according to urban/Bedouin-rural origin, or Sunni/Shiʿi denomination. Religious affiliation is in fact an important factor in language choice. Woidich (1997:196) found that although many Arabic dialects of Upper Egypt showed distinct Bedouin features, no Bedouinization has occurred in the dialect of the Christian village of ʿIzbit al-Basīlī in the heart of the → Bʿērī region.

3.1 *Bahraini Arabic*

Variation according to denomination exists in the spoken Arabic of Bahrain (Prochazka 1981; Ingham 1982; Holes 1980, 1983). However, contrary to a previously held view that speech variation is usually free, Holes (1980:72) shows that it is "socially constrained". Differentiating between the Shiʿi Baḥārna, the older inhabitants of Bahrain, and the later arrivals, the Sunni Bahrainis, known locally as ʿArab, Holes (1980, 1983) points out that the dialect of the former group is sedentary, while that of the latter is of Bedouin origin. Sunni Arabic is used by some Shiʿis because it is more prestigious than Shiʿi Arabic, since it is the speech of the higher echelons of Bahraini society, including the ruling family, and is, moreover, indicative of Bedouin ethnicity (Holes 1980:81).

The following are some examples of Sunni/Shiʿi variation:

i. /y/ versus /j/. Modern Standard Arabic /j/ which occurs in the speech of the Shiʿi inhabitants of Bahrain has a reflex /y/ in the Sunni dialect. /y/ occurs in any phonological context (Johnstone 1967:20). Thus forms like *yadīd* 'new', *yār* 'neighbor', *daray* 'ladder', *šayar* 'trees' are considered to be more socially prestigious than comparable forms with /j/.

ii. /ṯ/ versus /f/. The Baḥārna dialect spoken by the Shiʿis is characterized by /f/, the allophone of the interdental /ṯ/ (Ingham 1982; Al-Tajir 1982). The Baḥārna realize *ṯūm* 'garlic', *ṯagīl* 'heavy', *ṯarīd* 'meat broth with pieces of bread', and *ṯalāṯīn* 'thirty' as *fūm*, *fagīl*, *farīd*, and *falāfīn*. Al-Tajir (1982:53) points out that in Modern Standard Arabic loans the interdental is usually retained. Thus, it has *ṯyāb* 'clothes' and *ḥadīṯ* 'sayings of the Prophet'. In the Sunni dialect, as in Arabian Bedouin dialects, the three interdentals /ṯ/, /ḏ/ and /ḏ̣/ are preserved.

iii. Final clusters. In CaCC forms final clusters occur in Shiʿi speech, whereas Sunni comparable forms are disyllabic. Thus *ḥabl* 'rope', *raml* 'sand', and *ṭabx* 'cooking' in Shiʿi Bahraini are *ḥabəl*, *raməl*, and *ṭabəx* in the Sunni dialect.

iv. Unstressed syllables. In the Sunni dialect unstressed /a/ and /ə/ are deleted on suffixation, while in the Shiʿi dialect only unstressed short /ə/ is deleted (Holes 1983:17). Thus, Sunni *šrubat* 'she drank', *dfaʿah* 'he pushed him', and *ḥṭuba* 'wood' are *šarabat*, *dafaʿah*, and *ḥaṭaba* in the Shiʿi dialect.

3.2 *Baghdad and central Iraq*

Rosenhouse (1984:169) says that "many words" in the Muslim Arabic of Baghdad are similar to those in north Israeli Bedouin dialects and the Arabic of Jedda, and that Baghdadi Muslim Arabic represents "a mixture of a nomadic-type dialect with sedentary elements . . . due to the settlement of Bedouins in Baghdad, and Iraq in general, during the ages, mainly since the devastation of Baghdad in the 13th century". According to Miller (2004:183–184), Baghdad and other urban centers witnessed the arrival of more Bedouin in the 18th and 19th centuries, and the speech variety of these Bedouin became "first demographically then politically dominant", before gaining status in the 20th century as the "standard urban dialect".

The 20th century represents an ongoing and important process of Bedouinization of the dialects of urban central Iraq, and of Baghdad in particular, that began from about the 1920s when large numbers of Bedouin and rural inhabitants from the south began to settle in Baghdad. During the 1960s new suburbs, among them present-day Ṣadr City in northeast Baghdad, were created to house the influx of southerners. Ṣadr City's two million Shiʿi inhabitants are of Bedouin/rural origin, and have been instrumental in introducing new features of Bedouin/rural speech into urban Baghdadi Arabic. Ṣadr City was called *Madīnat aṯ-ṯawra* 'Revolution City' when it was founded, and later *Madīnat Ṣaddām* 'Saddam City', before it was renamed *Madīnat aṣ-Ṣadr* in April 2003.

The Bedouin elements that have crept into Baghdadi Arabic through contact with the Arabic of Ṣadr City and other urban areas of central Iraq settled by southerners, belong to what Ingham (1982) refers to as the 'Mesopotamian' variety, in contrast to 'Arabian', the variety that comprises the dialects of central Najd and the Gulf region.

The following are some of the 'Mesopotamian' features that now occur in Baghdadi Arabic, predominantly in the speech of the Shiʿi community:

i. One salient feature is the glide vowel /ie/ < long vowel /ē/ (< Standard Arabic diphthong /ay/). /ie/ has been a variant of /ē/ in Baghdadi Muslim speech for a long time. It is typical of the spoken varieties of some inner Baghdad quarters, and is more common in women's speech (Baghdadi informants). Terms like *sief* 'sword', *lammiet* 'I gathered', *hammiena* 'also' can now be heard along with their variants *sēf*, *lammēt*, *hammēna*. The glide vowel does not occur after a guttural or in the environment of an emphatic. In such cases the characteristic vowel is invariably /ē/: *ġēr* 'other', *ḥēl* 'fast', *bēḏ̣* 'eggs', *šēṭān* 'Satan'.

ii. /q/ > /ġ/. is a characteristic of central and southern Iraqi Arabic, and has been well documented, as has /q/ > /k/, as in *kətal* 'to kill, to beat' and *wakət* 'time' (Blanc 1964) and /q/ > /j/, as in *jədər* 'cooking pot', *šarji* 'easterly (wind)' (Johnstone 1967). /q/ is also realized as /ġ/ in Baghdad and central Iraq in forms where it is followed by a long back vowel, as in: *ġūri* 'teapot' (< *qūri*), *ġūṭi* 'box' (< *qūṭəyya*), *ġāṭ* 'floor, storey; suit [clothes]' (< *qāṭ*). The latter term is frequently used in the phrase *pāṣ abu ġāṭēn* 'a double-decker bus'. /q/ > /ġ/ is ordinarily a feature of the speech of elderly people of Bedouin origin.

iii. *makān* 'place', typical of urban Iraqi Arabic, is now in free variation with the southern *mukān* that is also used in the Gulf region (Holes 2001:236).

iv. Stress assignment. In trisyllabic CvCvCCv/CvCCvCv forms it is usually the initial syllable that is stressed in Baghdadi Arabic. Where stress falls on the medial syllable the form is considered to be non-urban. Thus, *wáḥədna* 'by ourselves' and *mádrasa* 'school' occur in the speech of some Shiʿis as *waḥádna* and *madrása*.

v. Ca-/CaCC- for Cə-/CəCC-. Ca-/CaCC-type syllables widespread in southern Iraq, particularly in and around Al-Zubayr and Abu l-Xaṣīb, now occur in free variation with comparable Baghdadi urban forms with initial syllable Cə-/CəC-. Thus: *ʿanab/ʿənab* 'grapes'; *čama/čəma* 'truffles', *rajjāl/rəjjāl* 'man', *sayyāra/səyyāra* 'car'.

vi. CaC- > CCa-. In a number of Arabian and Gulf dialects an initial closed CaC- syllable becomes CCa- if the second consonant is a guttural (Johnstone 1967; Holes 1983; Ingham 1997; → *gahawa*-syndrome). This phenomenon occurs in Baghdadi Arabic in the form *hala* 'welcome' that is in free variation with *ahla* (< *ahlan wa sahlan*). Other similar forms, like *ghawa* 'coffee', *nxala* 'palm tree', can be heard in some parts of Baghdad, although they are still stigmatized as being too rural.

vii. Gender distinction. The 3rd pers. fem. pl. pronominal suffix /-an/ is a feature that is now widespread in Baghdadi Arabic. Thus: *ysayran* 'they go out on visits', *yətsawgan* 'they shop', *jahzan* 'they bought their wedding trousseau', *tlāgan* 'they met each other', occur frequently. Educated speakers who stress their urban origin do not show any such gender distinction (cf. Holes, 1990). However, the 2nd pers. pl. fem. suffix is rarely used.

viii. Lexical items. There are a number of Bedouin lexical items that are now part of Baghdadi Muslim Arabic. For example, *šāl*, originally meaning 'to carry', and now also 'to move house', is in free variation with *tḥawwal* and *ntəqal*. Another Bedouin term, *dašš* 'to enter' is used along with the

Baghdadi *ṭabb*. At one time *ṭabb* was indicative of rural speech occurring with the urban *xašš*. *xašš* is now rarely used, and even Christian Baghdadis are using *ṭabb* in free variation with *daxal*.

4. Nabaṭī poetry

The dialects spoken in the Arabian Peninsula "are clearly" representative of New Arabic (Versteegh (2004:351). This is a view held by many Arabian poets. Apart from Sowayan's (1985) definitive study of Nabaṭī poetry, there is a growing corpus of works by Saudi and Gulf poets who refer to their compositions as 'Nabaṭī' poetry. As-Saʿīd (1987:14–15) describes modern Nabaṭī poetry as "being a type of Bedouin verse whose language is the everyday speech of Bedouin tribes, written as it is spoken". Despite as-Saʿīd's statement, this genre of Bedouin poetry draws also on Modern Standard Arabic where the occasional classical word or expression is inserted into a dialectal sentence:

> *ktabt əj-jawāb muwaḏ̣ḏ̣aḥan ġāməḍ arqāma*
> 'I wrote a reply explaining what is vague in it'
>
> *mā tanālah qadd kutəba lak bəj-jabīn*
> 'What befalls you has (already) been written on your brow'
>
> *yā sīdī hāḏī buyūtan ṭarāʾəf*
> 'My Lord, these are humorous verses'

In the three sentences the Classical Arabic *tanwīn*, particle *qadd* followed by the passive, and medial *hamza* in *ṭarāʾəf* occur in a mainly dialectal context. Medial *hamza* is rarely used, thus: *dāyəm* 'lasting' (< *dāʾim*), *aṣdəqāya* 'my friends' (< *aṣdəqāʾī*), *fwād* 'heart' (< *fuʾād*), *ṣāyəm* 'fasting' (< *ṣāʿəm*), *ġanāyəm* 'booty' (< *ġanāʾəm*). Initial *hamza* is often elided: *l-əbnī* 'to my son' (< *ʾilā ibnī*), *yā hal* 'oh, you people' (< *yā ʾahl*), *mādri* 'I don't know' (< *lā ʾadrī*).

The following hemistichs are all in dialectal Arabic:

> *blād xayr maʿrūfīn bən-naxwa w ʿəzz əj-jār*
> 'It is a bounteous country [whose people] are known to be generous and they honor their neighbor'
>
> *w lā aḥad yunkur ahla. lā. w lā blāda wlā wlāda*
> 'And no one [there] denies his people. No. Nor his country and children'
>
> *marīḏ̣ l-əfwād mdāwəm bəl-ʿənā dawmī*
> 'I am sick of heart and continuing to struggle'

Dialectal plural forms are used throughout, the most common pattern being CaCāCəC: *mayābər* 'needles', *jadāyəd* 'new', *ḥamāyəl* 'burdens', *ḥabāyəl* 'ropes', *rakāyəb* 'mounts, camels', *laḥāyəf* 'blankets, covers'. Less common plural forms are of the patterns CəCāCīC and CəCāCī: *masāyīr* 'guests', *gaṣāṣīb* 'butchers', *lawāhīb* 'hot winds'; *šalālī* 'threads', *namālī* 'ants'. The relative pronoun *əlli* and its abbreviated form *əl*, are used for *allaḏī*. Preverbal negative particle *mā*, as in *mā yənšāf* 'it cannot be seen', *əl-lēl mā anāma* 'I cannot sleep at night' occurs frequently. Nabaṭī poets frequently introduce local dialectal terms, like *dašš* 'to enter', *šāl* 'to go away, to move house/abode', *sawwa* 'to do, to make', *wadda* 'to deliver', *magrūd* 'badly done by', *wəyya* 'with', *yamm* 'at, by', *wāš/wēš* 'what', *yāma* 'how often', *lāzəmlak* 'you must', *mā mīš/mā miš* 'there is not'.

5. Conclusion

Most Arabic-speaking people look favorably on Bedouin culture and speech characteristics. Four Christian Iraqi informants from Basra were proud to claim Bedouin descent, and cited examples of Bedouin Arabic in their everyday speech. In urban Iraqi Arabic *badwi* 'Bedouin' does not have a pejorative meaning like the term *ʿurbi*, which refers to someone poor and uneducated. Iraqis generally speak positively of having been *bəl-ʿarab* 'in the desert; among Bedouin'. Caubet (2004:67–68) referring to a colleague's work among the nomads of the Fezzān region, points out how some nomads and semi-nomads believe they are of a higher social status than neighboring sedentary tribes. In Bahrain, Sunni speech, considered to be of Bedouin origin, is used in the media, in popular plays and poetry "and even newspaper cartoons" (Holes 1980:81), while in both Iraq and Syria plays in Bedouin Arabic on television and radio are popular among a cross-section of listeners (according to informants). After Islam spread to sedentary regions and Bedouin Arabic came into contact with, and acquired features of, sedentary Arabic, Ibn Jinnī and Ibn Xaldūn, among others, thought that it had become "corrupted" (Versteegh 2004:351–352). Such was the high esteem in which Classical scholars held Bedouin Arabic, a status it enjoys to this day.

BIBLIOGRAPHICAL REFERENCES

Al-Tajir, Mahdi A. 1982 *Language and linguistic origins in Bahrain: The Baḥārnah dialect of Arabic*. London and Boston: Kegan Paul International.

Blanc, Haim. 1964. *Communal dialects in Baghdad*. Cambridge, Mass: Harvard University.

Cadora, Frederic J. 1992. *Bedouin, village and urban Arabic: An ecolinguistic study*. Leiden: E.J. Brill.

Cantineau, Jean. 1936. "Etudes sur quelques parlers de nomades arabes d'Orient". *Annales de l'Institut d'Etudes Orientales d'Alger* 2.1–118.

——. 1937. "Etudes sur quelques parlers de nomades arabes d'Orient". *Annales de l'Institut d'Etudes Orientales d'Alger* 3.119–237.

——. 1940. "Les parlers arabes du département d'Oran". *Revue Africaine* 84.220–231.

Caubet, Dominique. 1993. *L'arabe marocain*. 2 vols. Leuven and London: Peeters.

——. 2004. "Les parlers arabes nomades et sédentaires du Fezzān, d'après William et Philippe Marçais". Haak a.o. (2004:67–94).

Dāmiġ, ʾAḥmad ad-. 1990. *aš-Šiʿr an-Nabaṭī fī Wādī l-Fiqī*. 3 vols. Riyadh: Dār ʿĀlam al-Kutub.

Ferré, Daniel. n.d. *Lexique marocain–français*. Morocco: Fédala.

Haak, Martine, Rudolf de Jong, and Kees Versteegh (eds.). 2004. *Approaches to Arabic dialects: A collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday*. Leiden: E.J. Brill.

Holes, Clive. 1980. "Phonological variation in Bahraini Arabic: The [j] and [y] allophone of [j]". *Zeitschrift für Arabische Linguistik* 4.72–89.

——. 1983. "Bahraini dialects: Sectarian differences and the sedentary/nomadic split". *Zeitschrift für Arabische Linguistik* 10.7–38.

——. 1990. *Gulf Arabic*. London and New York: Routledge.

——. 2001. *Dialect, culture and society in Eastern Arabia*, I. Leiden: E.J. Brill.

Ingham, Bruce. 1982. *Northeast Arabian dialects*. London and Boston: Kegan Paul International.

——. 1997. *Arabian diversions: Studies on the dialects of Arabia*. Reading, Berkshire: Ithaca Press.

Jastrow, Otto. 1978. *Die mesopotamisch-arabischen* qəltu-*Dialekte*. I. *Phonologie und Morphologie*. Wiesbaden: F. Steiner.

Johnstone, Thomas M. 1967. *Eastern Arabian dialect studies*. London: Oxford University Press.

Jong, Rudolf de. 2000. *A grammar of the Bedouin dialects of the northern Sinai littoral: Bridging the linguistic gap between the eastern and western Arab world*. Leiden: E.J. Brill.

——. 2004. "Characteristics of Bedouin dialects in Sinai". Haak a.o. (2004:151–175).

Marçais, Philippe. 1977. *Esquisse grammaticale de l'arabe maghrébin*. Paris: Maisonneuve.

Miller, Catherine. 2004. "Variation and change in Arabic urban vernaculars". Haak a.o. (2004:177–206).

Owens, Jonathan. 1984. *A short reference grammar of eastern Libyan Arabic*. Wiesbaden: O. Harrassowitz.

Palva, Heikki. 1976. *Studies in the dialect of the semi-nomadic el-ʿAgārma tribe*. Gothenburg: Acta Universitatis Gothoburgensis.

——. 1980. "Characteristics of the Arabic dialect of the Bani Ṣaxar tribe". *Orientalia Suecana* 29.112–139.

——. 1994. "Bedouin and sedentary elements in the dialect of es-Salṭ". *Actes des premières journées internationales de dialectologie arabe*, ed. Dominique Caubet and Martine Vanhove, 459–469. Paris: INALCO.

Prochazka, Theodore. 1981. "The Shīʿī dialects of Baḥrain and their relationship to the Eastern Arabian dialect of Muḥarraq and the Omani dialect of al-Ristāq". *Zeitschrift für Arabische Linguistik* 6.16–55.

Qaḥṭānī, Muḥammad al-. 1998. *Fayḍ min šāʿir li-ramz al-mašāʿir*. Riyadh: King Fahd National Library.

Rabin, Chaim. 1951. *Ancient West Arabian*. London: Taylor's Foreign Press.

Rosenhouse, Judith. 1984. *The Bedouin Arabic dialects*. Wiesbaden: O. Harrassowitz.

Saʿīd, Ṭalāl ʿUṯmān as-. 1987. *al-Mawsūʿa an-Nabaṭiyya al-kāmila*. Kuwait: Ḏāt as-Salāsil.

Singer, Hans-Rudolf. 1994. "Die Beduinen als Träger der Arabisierung im islamischen Machtbereich". *Akten des III Arabistischen Kolloquiums, Leipzig, 21–22 November 1991*, ed. D. Bellman, 263–274. Stuttgart: Deutsche Morgenländische Gesellschaft.

Sowayan, Saad 1985 *Nabaṭī poetry: The oral poetry of Arabia*. Berkeley, Cal.: University of California Press.

Versteegh, Kees 2004. "Pidginization and creolization revisited: The case of Arabic". Haak a.o (2004: 343–357).

Woidich, Manfred. 1997. "Egyptian Arabic and dialect contact in historical perspective". *Humanism, culture and language in the Near East: Studies in honor of Georg Krotkoff*, ed. Asma Afsaruddin and A.H. Mathias Zahniser, 185–197. Winona Lake, Ind.: Eisenbrauns.

FARIDA ABU-HAIDAR (London, U.K.)

Beirut Arabic

1. GENERAL REMARKS

Beirut has a population of 1,500,000 people. The country's oldest urban communities are the Sunnis and Greek-Orthodox. Successive waves of migration brought Maronites, Shiʿis, Druzes, Greek-Catholic, Syriacs, as well as Syrians, Palestinians, Kurds, Armenians, and others. Colloquial Arabic is the means of communication, *par excellence*. To varying degrees, the population is bilingual: French or English is taught at the primary school level. It is mostly among the upper classes that foreign languages are used, often alternating with Arabic (conversational code-switching).

Within the Syria-Lebanon-Palestine area, the colloquial Arabic of Beirut belongs to those dialects showing strong urban characteristics (Cantineau 1939). Contrary to other Lebanese varieties of Arabic (Feghali 1919; El-Hajjé 1954; Fleisch 1974), it has not been the subject of any exhaustive description, except in the domain of phonology (Mattsson 1911; Naïm-Sanbar 1985). The dialect usually presented in language manuals is the 'Lebanese dialect' (D'Alverny 1970) or 'Lebano-Syrian' (Nakhla 1937), without further specification. No dictionary has been compiled on Beirut usage, but it is possible to find indications in more general dictionaries concerning the Syria-Lebanon-Palestine area or the Lebanon dialect (Barthélémy 1935; Frayha 1973).

In the rest of the country, the Beirut dialect is considered 'affected' (what is urban cannot be authentic), or 'drawling' or 'throaty' (allusions to the linguistic specificity of the Sunni community). No special prestige as capital city language is attributed to it in the media: television programs and soap operas in colloquial Arabic prefer to show particularities typical of certain regions, or 'typically vulgar' as Feghali puts it (1928:x–xii), rather than showing the capital variety, more difficult to characterize, unless through one of its components (the colloquial speech of the old Sunni neighborhood or the French- or English-laden speech of the bourgeoisie).

2. Linguistic description

2.1 *Phonology*

2.1.1 Inventory

2.1.1.1 Consonants

i. The velar /g/ and the uvular /q/

Beirut Arabic differs from Classical Arabic in that the uvular consonant **q* corresponds to the glottal consonant /ʾ/. /g/ appears mostly in older borrowings from Turkish and Persian, and more recently in borrowings from Romance languages and Cairene Arabic: ['gəmruk] 'customs' (Turkish), [ʔar'giːle] 'narguileh' (Persian), [si'gaːra] 'cigarette' (French), [gəlːa'biːye] 'men's clothing' (Cairene); /q/ is found in terms associated with Classical Arabic, *qāṛṛa* 'globe', in administrative vocabulary, *ʿiqaṛāt* 'real estate', or in religious vocabulary, *al-qurʾān* 'the *Qurʾān*', and in military terms (especially since the beginning of the war in Lebanon in 1975), *qannāṣ* 'franc-tireur'. With the exception of the religious domain, which upholds the highest variety of Arabic, this vocabulary is, with use, reinterpreted following the phonology of the dialect with **q* > /ʾ/ or /ʾ̣/ and /g/ > /ž/: *qannāṣ* and *ʾ̣annāṣ* 'franc-tireur', *gəllabiyya* and *žəllabiyya* 'men's clothing' coexisting.

ii. Interdentals

The dental fricatives **ṯ*, **ḏ*, and **ḏ̣* have been replaced by the corresponding alveolar plosives and sibilants (except among the Druze speakers) following a distribution which does not appear to be dictated by the same organizing phonetic or semantic principle:

**ṯ*	/t/ : *talž* 'snow'	–	/s/ : *muʾannas* 'feminine'
**ḏ*	/d/ : *dəhab* 'gold'	–	/z/ : *zabzaba* 'deceitfulness'
**ḏ̣*	/ḍ/ : *ḍahr* 'back'	–	/ẓ/ : *ẓarf* 'envelope'

The pharyngealized sibilant /ẓ/ notably appears in the Turkish words borrowed from Arabic dating from the Ottoman period.

Table 1. Consonant phonemes

	bilabial	labiodental	dental	alveolar	postalveolar	palatal	velar	uvular	pharyngeal	glottal
plosive	b			t d			k g	q		ʾ
pharyngealized	ḅ			ṭ ḍ						ʾ̣
nasal	m			n						
pharyngealized	ṃ									
trill				r						
pharyngealized				ṛ						
fricative		f		s z	š ž		x ġ		ḥ ʿ	h
pharyngealized				ṣ ẓ						
lateral				l						
pharyngealized				ḷ						
approximant	w					y				

iii. Pharyngealization or ‘emphasis’

Pharyngealization undoubtedly constitutes one of the principal characteristics of the Beirut dialect. The number of pharyngealized phonemes has more than doubled, from four in Literary Arabic to nine in the dialect, /ḅ, ṃ, ṭ, ḍ, ṛ, ṣ, ẓ, ḷ, ʾ̣/. /ḅ, ṃ, ʾ̣/ have a low level of distinctiveness. Pharyngealized glottal /ʾ̣/ representing *q is attested in the speech of Sunni men, *ʾ̣alam* [ʡɑɫɑm] ‘pencil’, *ḥaʾ̣ʾ̣* [ħɑʡʡ] ‘law’ (see Naïm-Sanbar 1985).

iv. Sociolinguistic variables

Druze speech is characterized by a relative conservation of /q/ which alternates with /ʾ/ in ordinary vocabulary, and by the absence of the pharyngealized glottal /ʾ̣/, *qaraf* ‘disgust’, *ṭarīq* ‘road’. It also differs through its preservation of the dental fricatives in borrowings from Literary Arabic, *ḏāb* ‘it melted’, *wirāṯa* ‘inheritance.

Among women, the pharyngealized glottal /ʾ̣/ is represented in only a few lexemes; their speech is also characterized by light emphasis (see 2.1.2).

2.1.1.2 Vowels

i. Inventory

There are two vowel series, four short vowels, /i/ /a/, /ə/, /u/ and five long vowels, /ī/, /ā/, /ē/, /ū/, /ō/. These series are not symmetrical: /ə/ has no corresponding long vowel and /ē/ and /ō/ have no corresponding short vowel (see Table 2).

The distribution of short vowels within forms is governed by syllable and stress constraints: /i/ and /u/ only appear in non-stressed closed syllable finals (CvC type), which position is incompatible with the presence of /ə/:

Cv	> /a/, /ə/
Cv́	> /a/, /ə/
Cv́C	> /a/, /ə/
CvC#	> /i/, /u/, /a/
CvCC#	> /a/, /ə/

The current linguistic situation results in the loss of the high vowels /i/ and /u/ of Old and Classical Arabic, in open non-stressed syllables (differential dialects) and in the neutralization of the opposition /i/ ~ /u/ in benefit of a central vowel of medium aperture /ə/. The low vowel /a/ has been better preserved, especially in the context of uvular, pharyngeal, and laryngeal consonants: **lisān* > *lsēn* ‘tongue’; **xudūd* > *xdūd* ‘cheeks’; but **xamīs* > *xamīs* ‘Thursday’.

ii. Short vowels

The vowels /i/ and /u/ are articulated as high central vowels [ɨ] and [ʉ], *žēmiʿ* ['ʒeːmɨʕ] ‘mosque’, *bəlbul* ['bəlbʉl] ‘nightingale’. Besides its unmarked allophone [æ], /a/ has a lower articulation [a], [ɑ] following uvular, pharyngeal, pharyngealized, and laryngeal consonants, *ġarīb* [ɣariːb] ‘foreigner’, *ʿarīḍa* [ʕariːd̴ɑ] ‘wide’ (fem.), *ḍarab* [d̴ɑr̴ɑb̴] ‘he hit’. The vowel /a/ of the feminine suffix *-a* has a higher articulation [e] except following back and pharyngealized consonants, *madrase* [mædræse] ‘school’, *manṭaʾ̣a* [m̴ɑn̴ŧɑʡɑ] ‘region’. These consonants also push back and down /ə/ and /u/, *byədġuṭ* [byədɣɔŧ] ‘he puts pressure on’, *byəʾʿud* [byəʔʕɔd] ‘he sits down’. Due to its position, the central vowel /ə/ shows a wide latitude of high and low articulations, [e], [ɨ], [ʉ], [ɔ], depending on the phonic and syllabic context (Figure 1).

iii. Long vowels

Long vowels have short or semi-long allophonics in pre-stress syllables, *žiṛān* [ʒir̴ɑːn] ‘neighbors’, *ḥalibēt* [ħaliˑ'beːt] ‘milk’; in open word final syllables, they are always short; /ī/ in pausa produces [e]: *ʾahāle#* [ʔæhæːle]#. Before back phonemes, especially pharyngealized phonemes, /ē/ and /ū/ are produced low, open. The vowel /ē/ represents both the reduction of /ay/ and the open vowel /ā/ [æː] of Classical Arabic. Probably through the effects of schooling and modern Arabic used in the media, /ē/ coexists with /ā/ among speakers. The latter is associated

Table 2. Vowels

	i	ī	ē	ə	a	ā	u	ū	ō
backness	-		-	//	-		+		+
height	+		+-	-+	-		+		+-
rounded	-		-	-+	-		+		+
long	-	+	+	-	-	+	-	+	+

Figure 1. Short vowel allophonics

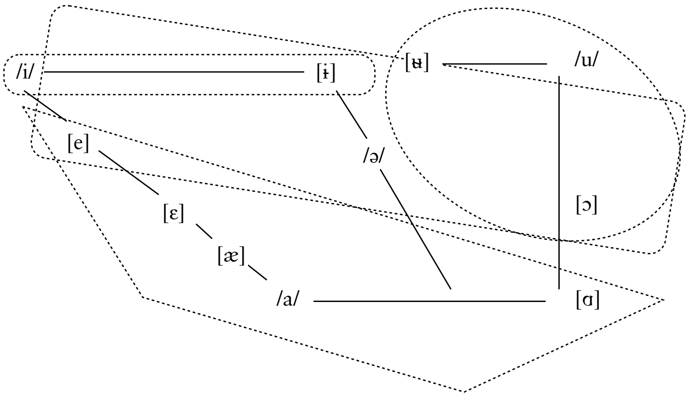
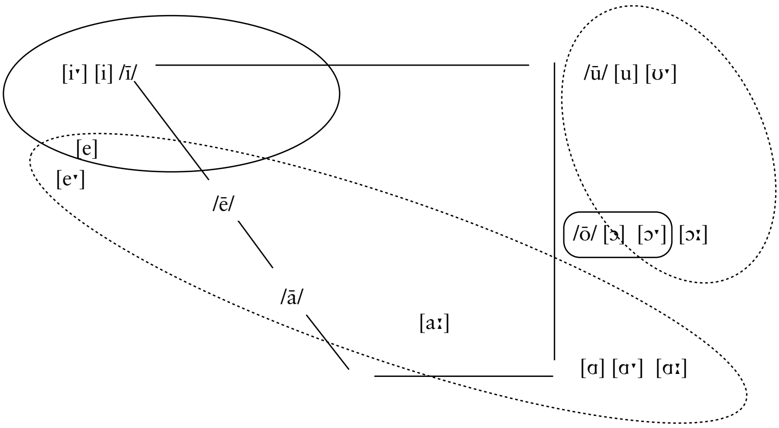


Figure 2. Long vowel allophonics



with formal language, appearing in situations where a questionnaire is being used, whereas /ē/ appears in informal conversation. A sociolinguistic variant has been noted amongst speakers of Palestinian origin in whose speech /ā/ [aː] is the reflex of *ā in all contexts.

2.1.1.3 Diphthongs

The diphthongs /ay/ and /aw/ of Classical Arabic have been reduced to the long vowels /ē/ and /ō/ but an instability is observed in open syllables: *baytēn* 'two houses', *mawže* 'wave'.

2.1.1.4 Syllables

Three types of syllable have been noted, which differ in their structure and in their 'length':

- i. short syllables Cv
- ii. long syllables CvC, Cv̄
- iii. supra-long syllables Cv̄C, CvCC

A syllable may begin with a consonant or a consonant cluster. Supra-long syllables only appear in final position, *mažnūn* 'crazy', *nfalaʾt* 'I am at the end of my tether'.

2.1.1.5 Consonant clusters

Two consonant clusters are well tolerated in initial position. In this position one finds clusters decreasing in closure, *ṭṛāb* 'earth', increasing in closure, *žbēl* 'mountains', and with equal closure, *ḥsēb* 'account'. Through assimilation, homorganic consonants are reduced to geminates necessitating an ultra-short prothetic vowel for their articulation, [ᵊᵵːɑᵲ'ᵲɑjnɑ] 'we had to', [ᵊrːa'biːʕ] 'the spring'. Initial three-consonant clusters are not permitted.

Within a word, three-consonant clusters are tolerated, *bəʿətʾid* and *bəʿtʾid* 'I think', except in the presence of a lateral or trill consonant, [jədᵊrsu] 'they study', [jəħᵊlbu] 'they milk'.

In final position, two-consonant clusters are disjointed before a pause, [laħᵊm]# 'meat' The color of the disjunction vowels depends on the consonantal context, vowel harmony, and rhythm.

2.1.1.6 Stress

Stress is placed automatically, depending on syllable length. It first affects supra-long syllables (Cv̄C and CvCC), which appear only in final position. Stress affects the final syllable if it is supra-long, Cv̄C or CvCC; the penultimate if it is long Cv̄ or CvC and if the final syllable is not supra-long; otherwise antepenultimate. There are some exceptions to this rule, due to morphophonological constraints:

i. forms where a penultimate short syllable is stressed, CvCv́Cv(C), CvCvCv́Cv(C), bearing the 3rd pers. sg. fem. suffix or the 3rd pers. pl. suffix, *ḍaṛáḅa* 'he hit her', *ḍaṛáḅun* 'he hit them', *madrasə́tun* 'their school'. Here, stress indicates the virtual presence of /h/ before the suffixes *-a* and *-un*: *ḍaráb[h]a*, *madrasə́t[h]un*;
ii. forms where a short final syllable is stressed, CvCvCCv́. These are vowel-final forms bearing the 3rd pers. sg. masc. suffix *katabtī(h)* [kætæb'ti:] 'you [fem.] wrote it', *katabnē(h)* [kætæb'ne:] ʿwe wrote itʿ. Here, too, stress indicates the virtual presence of [h] of the 3rd pers. sg. masc. suffix.

2.1.2 Phonotactics

2.1.2.1 Pharyngealization

Beirut usage is characterized by its remarkable faculty for spreading pharyngealization (both progressively and regressively) on the syntagmatic level. Excepting /ī/ and /ū/ and the sibilants /š/ and /ž/, which have the ability to block this spreading, all phonemes may bear pharyngealization within words. It may be limited to one syllable or may cover the entire form, such as when it is constituted by a succession of open syllables. It does not affect the prefixed imperfect marker, *mnəḍṛuḅun* [mnəᵭᵲɔᵬɔʀ̴] 'we hit them', or the feminine suffix *ʾaṣba* [ʡɑsbe] 'reed' (unless it is in direct contact with the pharyngealization inductor phoneme *ḅaʾṛa* [ᵬɑʡᵲɑ] 'cow'), or the vowel [eː], reduction of the diphthong /ay/, *ṣēf* [ᵴeːf] 'summer'.

Among the different groups of speakers, three degrees of pharyngealization are attested: (a) strong, characteristic of men's speech, especially within the Sunni community; (b) medium, characteristic of women with a university education who often speak French or English; and (c) weak, characteristic of groups who usually speak a mixed language, Arabic and French or Arabic and English.

2.1.2.2 Pausal forms

One observes lax articulation at pauses, entailing devoicing and depharyngealization of final consonants, or vowel lengthening. This phenomenon especially affects the suffix *-a* which is pronounced open and long, *zġīra* [zɣiːrɛˑ] 'small', *madrsa* [mædᵊrsɛˑ] 'school', which has earned for Beirut colloquial Arabic the epithet 'drawling'.

2.2 *Morphology*

2.2.1 Pronouns

2.2.1.1 Independent personal pronouns

Closed variants (two degrees of → *ʾimāla*) have been noted in Sunni speech, for the 1st pers. sg. and pl. *ʾane ~ ani*, and *nəḥne*. In the 2nd pers. sg. masc., *ʾənta* and *ʾənte* coexist, attesting to an instability in the distinction of gender. The 3rd persons have two allomorphs, the heavy form (*huwwe*, *hiyye*, *hənne*) is used for insistence (Table 3).

Table 3. Independent pronouns

	3rd pers.	2nd pers.	1st pers.
masculine	*hu ~ huwwe*	*ʾənta ~ ʾənte*	*ʾana ~ ʾane ~ ʾāne ~ ʾani*
feminine	*hi ~ hiyye*	*ʾənte*	
plural	*hənn ~ hənne*	*ʾəntu*	*nəḥna ~ nəḥne*

2.2.1.2 Possessive/object suffixes

These are set out in Table 4

Table 4. Possessive/object suffixes

	3rd pers.	2nd pers.	1st pers.
masculine	*-u* / *-Ø*	*-ak*	*-ī* / *-nī*
feminine	*-a*	*-ik*	
plural	*-un*	*-kun*	*-nā*

The 1st pers. distinguishes between nominal and verbal suffixes. The 3rd pers. sg. masc. suffix is *-u* after consonants, but -Ø after vowels.

2.2.1.3 Demonstratives

There is a distinction between the demonstrative adjective and the demonstrative pronouns. The invariable adjective *ha-* is placed before a noun which is always defined, *ha žžbēl* 'these mountains'. It can combine with a demonstrative pronoun *ha lwalad haydā* 'this boy here'. Gender is formally marked in the singular. Variants, short and long, have been noted in all persons:

i. near: *haydā* masc. ~ *haydī* fem., *hādā* masc. ~ *hādī* fem., *hay* fem. 'this one'; *hadūl*, *haydūl*, *haw* 'these ones [masc. and fem.]'
ii. far: *haydāk* masc. ~ *haydīk* fem., *hadāk* masc. ~ *hādīk* fem. 'that one'; *hadulīk*, *hawdīk*, 'those ones [masc. and fem.]'

A reinforcing element *-e*, appears at the pause for both series: *hadōle*, *hawde*, *hadāke*, *hadolīke*, etc. For expressive purposes (denigration) when the referent is a person, the reverse order 'pronoun + noun' is possible: *hadūl žžirān balā ḥayā* 'these neighbors (are) shameless.

2.2.1.4 Presentatives

Presentatives are composed of either demonstrative + independent pronoun, or demonstrative + the particle *yē* (**ʾiyyā*) + the personal suffix: *haydā hū* ~ *yē* 'it's him, here he is'.

A particle, *yaḥ-*, followed by suffix pronouns is at the origin of a paradigm of exclamatory presentatives: *yaḥḥū*, *yaḥḥī*, *yaḥḥiyun* 'here he is!, here she is!, here they are!'.

2.2.1.5 Relative pronouns

The relative pronoun *yəllī*, [ᵊlli], [ᵊlle], [ᵊl] invariable, *lbint ᵊl kēnit* 'the girl who was', can be combined with the demonstrative *ha*: *hallī baddu* 'the one who wants', *hallī bətrīd* 'what you want', or can be used on its own, *yəlli ṛāḥ ṛāḥ* 'what is gone is gone'. The relative pronoun *mā* is only encountered in set expressions, *mbārak mā žēkun* 'blessed be your newborn'.

2.2.1.6 Interrogative pronouns

i. 'who' *mīn*, can be followed by a person suffix or an independent pronoun, *mīnu/-a* or *mīn huwwe* ~ *hiyye* 'who is he (him ~ her)'; inverse order is used for focus, *huwwe mīn* 'who is it?'.
ii. 'what' *šu*.
iii. 'which' *ʾayya*, can be followed by the numeral *wāḥad, ʾayya wāḥad/wəḥde* 'which one [masc./fem.]'; or by a personal suffix, *ʾayy-ē* 'which one [masc.]', *ʾayyē-ha* 'which one [fem.]', *ʾayyē-hun* 'which ones'.

2.2.2 Adverbs

i. place: 'here' *hōn*; 'there' *honīk*.
ii. temporal: 'now' *hallaʾ*; 'tomorrow' *bukṛa*; 'yesterday' *mbāriḥ*; 'early' *bakkīr*; 'today' *lyōm*.
iii. manner: 'slowly' *ʿala mahl, šway šway, bihudūʾ*; it is generally possible to create adverbs of manner using *bi-* followed by the nominal form, *bsərʿa* 'fast, quickly', *bžarʾa* 'courageously', *brawēʾ* 'softly, calmly'.
iv. affirmation: 'yes' *ʾē*; 'it is so' *mbala*; 'assuredly' *ʾakīd*; 'maybe' *yimkin, barke*.
v. negation: 'no' *laʾ*.
vi. interrogative adverbs: 'where' *wēn*; 'towards where' *la wēn*; 'from where' *mnēn*; 'when' *ʾaymtīn, ʾēmta*; 'how', *kīf*; 'how much', *ʾəddēš*; 'why', *lē(š), lay*.

Interrogative adverbs are usually placed at the head of the utterance. In nominal phrases, 'how' and 'where' can be constructed with a suffix pronoun: *wēna* 'where is she?', *kīfak* 'how are you?'.

2.2.3 *Particles*

In speech, the definite article *al-* [ᵊl], [l] is attached to the preceding lexeme, thus geminates resulting from its assimilation to homorganic consonants are separated *ṛās-əš šažra* 'the top of the tree'.

2.2.3.1 Genitive markers

To express adnominal possession (2.3.1.3), the dialect has recourse to a specific marker *tabaʿ* (< *t-b-ʿ* 'follow') followed by the possessor or its representative suffix.

2.2.3.2 Negation

There are two negation forms, *mā* and *məš*, of which the first, *mā*, is restricted to verbs. The morpheme *məš* is used for nominal negation, *məš kbīr* 'not big', and, under certain conditions, for verbal negation (see 2.3.2.5).

2.2.3.3 Question introducing particles

The most frequently used are: *yā tara*, placed at the head or at the end of an utterance, *ʾəžā yā tara* 'has he come?'; preceded by *hal*, it is placed at the head of the utterance; *šī*, always placed at the end of an utterance *bəddak yē šī* 'do you want it?'; *mədrī* always comes at the head, *mədrī šu baddu* '[who knows] what he wants'

2.2.3.4 Existentials

The existential phrase is constructed with *fī* 'there is' of locative origin **fī* + *h* 'in him', grammaticalized and set (see 2.3.2.6).

2.2.3.5 Conjunctions

i. Coordinating conjunctions: 'and', *w*, can take on a temporal meaning 'then, during . . .' or one of spatio-temporal concomitance, *ʾəža w-əyyē* 'he came with him'. Further, 'with' *maʿ*; 'but' *bass*; 'or' *wəlla* ~ *ʾaw*.
ii. Subordinating conjunctions: 'so that, until' *ta*; 'as soon as' *bass*, *lamma*, *šī* ; 'when' *waʾt əlli*; 'just as soon as, hardly' *ma . . . əlla*; 'that' *ʾənnu*; 'if' *ʾən*, *ʾəza*; 'because' *laʾan[nu]* invariable.

2.2.3.6 Exclamations

yū > admiring; *bəš* > admiring and surprised; *yeh* > surprised; *ya ʾaṭīšī* > surprised and indignant; *yiy*, *tfū* > disgusted; *walaw* > surprised 'well now!'; *ya wayl-* (with a suffix representing the speaker) > suffering, hardship.

2.2.4 Nouns

On the morphological level, one notes the high productivity of the pattern CaCCūC for the affective diminutive, *šaṭṭūr* 'little courageous one', *ḥammūr* 'little donkey', and also the high productivity of the external plural *-ēt*. Combined with collectives or other plural stems it has an affective value *ʾahlētī* 'my parents'. The suffix *-ēt* is also attested in borrowed words, *telefonēt* 'telephones', *kombyuterēt* 'computers'.

The dual serves as plural for paired body parts (→ pseudo-dual), *ʿinayn* ~ *ʿinēn* 'eyes', *ʾəžrayn* ~ *ʾəžrēn* 'feet'. It also expresses an indeterminate quantity, 'some, a little bit', *šī ʾəršēn* 'a little money', *maʿu ʾəršayn* 'he has some money'.

2.2.5 Numerals

On the different uses of 'one' see 2.3.1.1. The number 'two' is usually replaced by the dual suffix *-ayn*.

From 3 to 10 the cardinal numbers have three allomorphs (Table 5): a long form for counting, telling time, etc.; a short form before words beginning with a consonant; a form bearing the suffix -*-t* (**-at*) before words beginning with a vowel.

Table 5. Cardinal numbers

(1) Long forms

1–10	*wāḥad / wəḥda*	*tnēn*	*tlēta*	*ʾarbʿa*	*xamsa*	*sətta*	*sabʿa*	*tmēna*	*təsʿa*	*ʿašra*
11–19	*ḥdaʿš*	*tnaʿš*	*tlēttaʿs*	*ʾarbaʿtaʿš*	*xamstaʿš*	*səttaʿš*	*sabʿtaʿš*	*tmēntaʿš*	*təsʿtaʿš*	*ʿəšrīn*
100–		*miyya*		*mitēn ~ mitayn*			*tlēt miyya*		*ʾarbaʿ miyya*	
1,000		*ʾalf*		*ʾalfēn ~ ʾalfayn*			*tlēt ʾalēf*		*ʾarbaʿt alēf*	

(2) From 3 to 10

Before consonant initial words

tlēt	*ʾarbaʿ*	*xams*	*sətt*	*sabʿ*	*tmēn*	*təsʿ*	*ʿašr*

Before vowel initial words

tlēt	*ʾarbaʿt*	*xamst*	*sətt*	*sabʿt*	*tmēnt*	*təsʿt*	*ʿašrt*

(3) From 11 to 19

ḥdaʿšar	*tnaʿšar*	*tlēttaʿšar*	*ʾarbaʿtaʿšar*	*xamstaʿšar*	*səttʿšar*	*sabʿataʿšar*	*tmēntaʿšar*	*təsʿtaʿšar*

From 11 to 19 all cardinal numbers end with the suffix *-ar*. From 3 to 10, the noun is always in the plural; from 11 on, the noun is in the singular. *miyya* '100' has a short variant *mīt*.

Words referring to body parts, such as *ṛās* 'head', and *ʾəṣbiʿ* 'finger' may be used as count nouns. Their distribution is clearly motivated by the shape of what is being counted, whether it is rounded or oblong, *tlēt ṛūs ġanam* 'three sheep', *ṛās banadūra* 'one tomato', *ʾəṣbiʿ xyāṛ* 'one cucumber'.

All cardinal numbers can be used as pronominal anaphorics: *štarayt təffēḥ w ʾakalt wəḥda* 'I bought some apples and ate one'. The numeral 'one' can carry the dual suffix *-ayn* to signify 'two' *ʾakalt wəḥdtayn* 'I ate two of them', and the plural suffix *-ēt* for feminine nouns only, *ʾakalt xams wəḥdēt* 'I ate five of them'. Used with person suffixes, it takes on the meaning of 'all alone'; it is usually preceded by the dative particle *la, štaġalt la waḥdī* 'I worked all alone'.

2.2.6 Verbs

2.2.6.1 Form I

Today, there are still traces of the semantico-syntactic distinction between active ~ deponent ~ qualitative verbs, historically established on the basis of the distribution of the vowels (*-a-*), (*-i-*) and (*-u*) of the basic verb stem. As in the majority of dialects, the vocalization (*-u-*) has given way to the vocalization (*-i-*). One therefore finds two stems CaCaC and CəCiC.

Under CaCaC one finds active and middle voices, transitive and intransitive verbs; under CəCiC, middle and reflexive verbs, labile intransitive and transitive verbs, semantically grouped into movement verbs, emotion and perception verbs, grooming verbs, and cognition verbs. In general, the vocalizations (*-a-*) and (*-i-*) do not contrast within the same form (contrary to other Lebanese dialects, El-Hajje 1954; El-Zein 1981). Only a few minimal pairs have been noted, *ʾəniʿ* 'to be convinced' ~ *ʾanaʿ* 'to convince', *xərib* 'to deteriorate [intr.]' ~ *xarab* 'to deteriorate [trans.]', *xəliʾ* 'to be born' ~ *xalaʾ* 'to create'.

2.2.6.2 Derived verbs

Middle verbs of the CəCiC type show corresponding causative forms (Form II or Form IV), *šərib* 'to drink' ~ *šarrab* 'to make drink', *həlik* 'to be worn out' ~ *ʾahlak* 'to wear someone out' (Naïm 2001).

The reflexive is expressed by the derived forms, Form V (t-CvCCvC), Form VI (t-Cv̄CvC), and Form VIII (CtvCvC), very much alive in the dialect (see 2.3.2.1).

One notes the loss of Form IV often replaced by Form II which, as has been observed for Syrian Arabic, forms doublets with Form I in characterized roots (Lentin 1991), and the weak productivity of Form VII (*n-* prefix), replaced in asubjectal constructions by the 3rd pers. pl. of the base form, and by Form V for a base formed on Form II.

2.2.6.3 Inflexions of the imperfect and perfect tenses (Table 6)

In Form I of the 'strong' verb, the perfect tense shows two vocalizations CaCaC and CəCiC, and the imperfect tense shows three yəCCaC, yəCCuC and yəCCiC, the vocalization (*-i-*) showing very low productivity.

The simple form of the imperfect tense appears in bound uses and in modal uses such as the optative and the injunctive *ʾaḷḷa yxallīk* 'God preserve you, I beg of you', *yəṣtfill* 'that he fend for himself, too bad for him'. Apart from these constructions, the imperfect preceded by the particle *b-* [b], [bə] expresses the indicative mode.

2.2.6.4 Particles and auxiliaries

The verb 'to be' *kān* serves as temporal auxiliary. It is compatible with other particles such as *ʿam* (progressive) and *raḥ* (periphrastic future),

Table 6. Conjugations of 'strong' verbs

	Perfect				Imperfect	
	3rd pers.	2nd pers.	1st pers.	3rd pers.	2nd pers.	1st pers.
masculine	*katab*	*katabt*	*katabt*	*yəktub*	*təktub*	*ʾəktub*
feminine	*katabit*	*katabtī*		*təktub*	*təktbī*	
plural	*katabū*	*katabtū*	*katabnā*	*yəktbū*	*təktbū*	*nəktub*

which are set. Other particles add temporal or aspectual information to the verb form.

With movement and displacement verbs, the progressive is expressed through the participial form. The other aspectual distinctions are expressed by aspect-indicated verbs:

ṣār (byṣīr) + imperfect > inchoative-durative: *ṣār yʾūl* 'he begins saying'

ballaš (byballiš) + imperfect > inchoative, *bukra byballiš yxabbərun* 'tomorrow he will begin telling them'

ʾəʿid (byəʾʿud) + imperfect > durative-inchoative, *ʾəʿid yxabbərna ḥayētu* 'he began telling us his life story'

ḍall (byḍall) + imperfect > durative, *ḍallit təmšī ta* . . . 'she kept walking until . . .'

bəʾī + imperfect > durative; *yəbʾa* (imperfect) + imperfect (with negation) > injunctive, *mā təbʾa tʾawwī ṣawtak* 'don't raise your voice'; *yəbʾa* + *b*-imperfect > volitional, *mnəbʾa mnəržaʾ sawa* 'we will come back together'

ʾām (byʾūm) + imperfect or perfect > inchoative (sudden beginning of the action), *ʾām*

ḍarabnī 'all of a sudden he hit me'; *yʾūm* (imperfect) + imperfect > injunctive (modal use),

ʾūʿa tʾūm tʾəllu 'careful, don't go and tell him!'

rəžiʿ (byəržaʿ) + imperfect or perfect > iterative *rəžiʿ ʾalla* 'he tells her once again'

ʿād (byʿūd) + perfect > iterative 'once again', in the negative phrase 'any more', *mā ʿādit ʾakalit* 'she didn't eat any more'.

2.2.6.5 The imperative

The imperative (Table 7) is characterized by a lengthening of the vowel of the 2nd pers. sg. masc., /–a-/ [æ] and /-ə-/ [ɔ:] linked to prosodic phenomena characteristic of monosyllables (articulatory energy, pause); these vowels are shortened in the other persons (disyllabic).

Table 7. The imperative

nzal [nzæːl]	*ktəb* [ktoːb]
nzalī [nzæle]	*ktəbī* [ktəbe]
nzalū [nzælu]	*ktəbū* [ktəbu]

For the 1st pers. pl. the exhortative value of the imperative is often expressed in the dialect by the particle *yaḷḷa* or the auxiliary *xallīna*, or also the verb 'to come' in the imperative *taʿā*, *taʿī*, *taʿū* placed before the 1st pers. pl. of the imperfect tense of the verb: *yaḷḷa nəmšī* ~ *xallīnā nəmšī* 'let's go, let's leave', *yaḷḷa* ~ *xallīnā* ~ *taʿā nəmšī* 'let's go, let's leave'.

2.2.7 Weak verbs

2.2.7.1 Geminated verbs

Geminated verbs have /-a-/ in the perfect tense and /-ə-/ in the imperfect tense, rarely /-a-/, *žann/byžənn* 'to go crazy', *naṭṭ/bynəṭṭ* 'to jump', *ʿaḍḍ/byʿaḍḍ* 'to bite'. In the imperative, contrary to 'strong' verbs, the 2nd pers. sg. masc. vowel is not long, *nəṭṭ* 'jump!', *dəll* 'show!'.

2.2.7.2 Verbs Iʾ, I*w* or I*y*

These change their consonants in the imperfect tense in *ē*, *ī*, or *ū*, *byēxud* 'to take', *byūṣal* 'to arrive', *byības* 'to dry up'. The imperatives do not differ from 'strong' verbs.

2.2.7.3 Verbs I*w*

These are of the CaCaC or CəCiC type: *waʿad*, 'to promise', *wərim* 'to swell'. The thematic vowel of the imperfective is (-*a*-). The verbs I*y*, rare (three have been noted), show a perfect tense of the CəCiC type: *yəbis*, 'to dry up', *yətim* 'to be orphaned', *yəʾis* 'to despair'.

2.2.7.4 Verbs II*w*/y

These all show in the perfect tense the long vowel -*ē*- (allophonic [-a:-]) and in the imperfect one of the three vocalizations -*ē*-, -*ū*-, or -*ī*- (the same as in the imperative), *nēm*, *nēmī*, *nēmū* 'sleep', *xāf*, *xāfī*, *xāfū* 'fear', *rūḥ*, *rūḥī*, *rūḥū* 'go away', *žīb*, *žībī*, *žībū* 'bring'.

2.2.7.5 Verbs with III*y*

These belong to the CaCaC, byəCCiC or CəCiC, byəCCaC type: *ramā/byərmī* 'to throw', *nəsī/byənsā* 'to forget', *rəxī/byərxā* 'to soften'. Some verbs of the CəCiC type have their imperfect in byəCCiC: *bəkī/byəbkī* 'to cry', *məšī/byəmšī* 'to walk', *ḥəkī/byəḥkī* 'to speak'. In the imperative the final vowel is long: *rxī*, *rxī*, *rxū* 'to release', *nsā*, *nsī*, *nsū* 'to forget'.

2.2.7.6 Doubly weak verbs

The doubly weak verbs, *ʾəžā/byəžī*, 'to come', *wəṭī/byūṭā* 'to become low', *ʾəwī/byəʾwā* 'to become strong' are conjugated irregularly, for example **žāʾa* 'to come, arrive' (Table 8) that do not distinguish gender in the 2nd pers. sg. of the imperfect.

Table 8. *ʾəžā* ‘to come’

	Perfect			Imperfect		
	3rd pers.	2nd pers.	1st pers.	3rd pers.	2nd pers.	1st pers.
masculine	*ʾəžā*	*žīt*	*žīt*	*byəžī*	*btəžī*	*bəžī*
feminine	*ʾəžit*	*žītī*		*btəžī*	*btəžī*	
plural	*ʾəžū*	*žītū*	*žīnā*	*byəžū*	*btəžū*	*mnəžī*

For the imperative, a different verb is used, *taʿā*, *taʿī*, *taʿū* ‘come!’.

2.3 *Syntax*

2.3.1 Noun phrases

2.3.1.1 Expression of indefiniteness and specificity

For the formal expression of indefiniteness, the numeral ‘one’ is used, *wāḥad*, *wəḥda*, but it never represents an indefinite article. As an indefinite pronoun, *wāḥad* is equivalent to ‘someone’ with a set allophone in the indirect case *ḥadan* (negative variant, *mā ḥada* ‘no one’) in negative, interrogative, and injunctive constructions, *ʿazamnā wāḥad byaʿrfak* ‘we invited someone who knows you’, *ḥadan ʾəžā* ‘did anyone come?’. Before generic person terms, ‘man, woman, child’, *wāḥad* emphatically marks indefiniteness; its presence implies the presence of supplementary information concerning the person represented (speficity): *šəft wāḥad rəžžēl wēʾif ʿaṛāsu* ‘I saw a man [someone] standing on his head’; *šəft rəžžēl wēʾif ʿaṛāsu* ‘I saw a man standing on his head’; (?) *šəft wāḥad rəžžēl* (unacceptable); (?) *šəft wāḥad* (unacceptable).

The pronoun *wāḥad* may be preceded by the particle *šī*, which reinforces its indefiniteness. Preceded by the definite article, *lwāḥad*, it signifies the neutral pronoun ‘one’, *šu bəddu yʾūl lwāḥad*? ‘what can one say?’. Preceded by *kəll* ‘all’ it is equivalent to ‘each’, *kəll wāḥad minkun* . . . ‘each of you . . .’. Combined with the ordinal form of ‘two’, it expresses the distributive, *lwāḥad w ttēnī* ‘one . . . the other’.

2.3.1.2 Construct state

One finds two types of genitive construction. The most common type, which is compatible with all semantic contexts, is the one traditionally designated by → ‘construct state’, *ʾəžr ṭṭāwla* ‘the table’s leg’, *bənt ʾəxtī* ‘my sister’s daughter’. The second, analytical, type makes use of the genitive marker *tabaʿ* ~ *tāʿ* ~ *taʿūl* (the last two forms are typical of the Beirut Sunni dialects), placed between the noun and its determiner, both lexemes being preceded by the definite article, *lktēb tabaʿ lmʿalma* ‘the teacher’s book’. The analytical construction is bound by semantic constraints: it is incompatible with a possessed object belonging to the domain of what is inalienable (kinship terms, part–whole relations etc.). One cannot say *lbənt tabaʿ ʾəxtī* ‘my sister’s daughter’ or *lʾəžr tabaʿ ṭṭāwla* ‘the table’s leg’. Analytical constructs are mostly found in contexts of focus or specification.

The particle *tabaʿ* (lit. ‘following’) followed by person suffixes also serves for the expression of the possessive pronouns *tabaʿī* ~ *tabaʿik* . . . ‘mine, yours . . .’, *haydā taʿūlī* ‘this one is mine’; used in this way it alternates with the dative particle *ʾəl-*, *haydā taʿūlī* ~ *tabaʿī* ~ *ʾəlī* ‘this one is mine’.

2.3.1.3 Quantifiers

There are different forms to express the same notion of quantity. For the notion of ‘paucity’ one finds:

i. *šway* ‘a little’ with an intensive allomorph *šwayye* or *šwayyət*, the bound element is necessarily an indefinite plural or collective, *xōd šway* ‘take some’, *ʾakal šwayye* ‘he ate a little [very little]’, *ʿṭīne šwayyət maṣāre* ‘give me some money’;
ii. *nətfa* ‘very little’ expresses a degree of paucity superior to *šway*. This particle has an affective allophone *nattūfe*; it is mostly used before a collective, *nətfət bann* ‘some coffee’, *nattūfət səkkar* ‘very little sugar’.

Paucity is also expressed by the dual suffix *-ayn*: *maʿu ʾəršayn* ‘he has money’ (2.2.4).

For the plural one finds: *kam* ‘some’ + indefinite singular noun, *ʿəndnā kam təlmīz* ‘we have some students'; *baʿḍ* 'some' + definite plural noun, *baʿḍ lʾaṣḥāb* 'some friends'; *kəll* 'each, every, all' + indefinite noun in the singular but definite in the plural, *kəll walad* 'each ~ every child', *kəll lwlēd* 'all the children'

2.3.2 Verbal phrase

The order in the verbal phrase can be either VSO or SVO. In general, terms governed by a transitive verb follow the verb but there are also constructions with direct government introduced by the dative preposition *la* 'to'. These constructions, which were more frequent formerly, have been explained through Syriac influence (Feghali 1928). Today, they have an expressive value, a rhetorical function of highlighting and insistence, and are found in constructions where the direct complement is represented by a suffix in the verb form, in an anticipatory manner: *hiyye šēftu la lḥaramī* 'she, she saw him, the thief'.

In doubly transitive verb forms, with direct objects anaphorically represented, the suffix representing the argument most closely linked to the verb is separated from it by the particle *yē-*, *ʿṭaytak yēha* 'I gave her to you'. With double object verbs, direct and indirect, it is the indirect object introduced by *l-* that is next to the verb, the direct object is in second position, separated from the verb form by the particle *yē-*, *šūfī-lī yēha maʿa ḥaṛāṛa šī* 'Check if she has a fever' (lit. 'look for me her, has she fever?').

2.3.2.1 Reflexive and reciprocal analytical constructions

Another characteristic of the dialect concerns the analytic expression of reflexive and reciprocal phrases (Naïm 2004). The reflexive is generally expressed by the verb. Form V (tCvCCvC) or Form VIII (CtvCvC), depending on the basic stem. As decausatives, these forms modify the verb's valence and lend the verb an internal orientation, *kaḥḥal* (Form II) 'to put kohl on' – *tkaḥḥal* 'to put kohl on oneself', *farkaš* 'to trip someone' – *tfarkaš* 'to trip'. The reciprocal is usually expressed by Form VI (tCv̄CvC), which gives the verb a bilateral orientation, *ʿabaṭ* 'to hug someone' – *tʿābaṭū* 'to hug each other', *ṣālaḥ* 'to forgive someone' – *tṣālaḥū* 'to become reconciled [reciprocally]'.

Alongside derivatives, the dialect makes use of emphatic or intensive markers following the basic verb form to express reflexives and reciprocals, the intensifier *ḥāl-* (lit. 'state, situation') followed by a suffix representing the subject in the case of reflexives, and the intensifier *baʿḍ-* (2.3.1.3), followed or not by the plural suffix, for the reciprocal*:* *tʿawwad* 'he became accustomed' and *ʿawwad ḥālu* 'he became accustomed'; *tḍāṛabū* 'they hit each other' (Form VI) and *ḍaṛabū baʿḍ(-un)* (Form I) 'they hit each other'.

These constructions show a connection between different strategies of reflexivization and the meaning of the predicate: *ḥāl-* is generally attested with verbs conventionally oriented at others (König and Siemund 1999) such as 'to hit', 'to speak', *ḍarab ḥālu* 'he hit himself', *byəḥkī ḥālu* 'he is talking to himself'; in these constructions, the intensifier occupies the position of object and signals the co-reference of the verb's two arguments. When the verb's orientation is internal (not other-directed), such as with 'to rest', the presence of the intensifier highlights the participation of the initiator and co-reference between the agent and the beneficiary, *baddī rayyiḥ ḥālī* 'I want to rest' instead of *baddī ʾərtēḥ* (Form VIII) 'I want to rest'; the intensifier modifies the argument structure and institutes the presence of two participants whereas only one is required.

Concerning reciprocity, constructions with *baʿḍ* emphasize the sequentiality of the action (Kemmer 1993) instead of considering its simultaneity: *byətʿāḅaṭū* (Form VI) 'they are embracing each other', *byəʿbṭū baʿḍ* 'they are embracing one another'.

2.3.2.3 Performatives

Speech acts are constructed with a verb in the 1st pers. sg. perfect tense, followed by an anaphoric pronoun representing the addressee; in certain cases the presence of the addressee is not formally marked but is simply implied in the enunciation situation. The perfect tense concomitant with the speech act here takes on the value of a 'present of address' (Naïm-Sanbar 1993), *ḥalaft* 'I swear', *ʾəltəllik skətī* 'I tell you: shut up'.

2.3.2.4 Intent, wishes

To express optativity, the imperfect tense is rarely used except in set expressions, *yxallīk* '[God] preserve you', *yṭawwil ʿəmrak* '[God] grant you long life'. In general, particles are used: *(ya) rayt* + noun or pronominal suffix + verb in the perfect tense (for regrets) or verb in the imperfect tense (for wishes); *nšaḷḷa* + perfect or imperfect preceded by *b-* (except in negative sentences) for wishes; *nēwī* (participial form, *nēwya*, *nēwyīn*) + imperfect for intent; *badd-* + pronominal suffix + imperfect for wishes, desires. Used on its own with person suffixes, this particle is equivalent to the verb ‘to want’.

2.3.2.5 Negation

Verbal negation is mostly accomplished by the particles *mā* and *məš*. The particle *lā*, attested in set expressions, is being superseded by *mā*. It is active in double negation phrases. As negation adverb, it takes on three forms: *lā*, *laʾ*, *laʾa* ‘no’.

The particle *mā* is placed before the verb; it is compatible with the perfect and the imperfect tenses, regardless of the temporal value of the latter. It can be reinforced by the particles *baʾa* and *šī*, or by the adverb *ʾabadan*, and can constitute discontinuous negation, ‘not . . . anything’, ‘not . . . any more’ and ‘not . . . at all’.

The particle *məš* is attested with a verb in the present or future tense. For nominal negation, the only particle used is *məš*.

2.3.2.6 Existential, possessive, and locative sentences

In the Beirut dialect, and more generally in the Syria-Lebanon-Palestine linguistic area, existential and possessive constructions have comparable structures: a predicatoid (*fī* for existentials, *ʿənd-* and *maʿ-* for possessives), a theme and a location argument. The theme argument associated with localization is indefinite in both cases: *fī kətub ʿaṭṭāwla* ‘there are books on the table’, *ʿəndī ~ maʿī kətub* ‘I have books ~ with me’. Both constructions also share the word order (Predicatoid + Subject) of the verbal phrase.

The locative construction differs by its definite theme, which occupies the subject function triggering agreement with the temporal exponent *lməḥḥāya kēnit ʿaṭṭāwla* ‘the eraser was on the table’, which is not the case with possessives or existentials. Furthermore, its word order is the same as that of nominal phrases (Subject + Predicatoid). The principal points of convergence and divergence between these three construction types are shown in Table 9.

The distribution of the relaters *ʿənd* and *maʿ* depends on the encoded notions of possession. The relater *ʿənd* encodes ‘permanent’ possession and the relater *maʿ* ‘temporary’ possession; both can serve to encode ‘abstract possession’ (Naïm 2003).

BIBLIOGRAPHICAL REFERENCES

Alverny, André d’. 1970. *Petite introduction au parler libanais*. Beirut: Dar El-Machreq.

Barthélemy, Adrien. 1935. *Dictionnaire arabe–français: Dialectes de Syrie: Alep, Damas, Liban, Jérusalem*. Paris: Geuthner.

Cantineau, Jean. 1939. “Remarques sur les parlers de sédentaires syro-libano-palestiniens”, *Bulletin de la Société de Linguistique de Paris* 11.1:80–88.

El-Hajjé, Hassan. 1954. *Le parler arabe de Tripoli*. Paris: Klincksiek

El-Zein, Abdoul Fattah. 1981. *Le parler arabe des Druzes de Chanay (Liban)*. Thèse de IIIème cycle, University of Paris 3 – Sorbonne nouvelle.

Feghali, Michel. 1919. *Le parler de Kfarʿabîda (Liban-Syrie)*. Paris: Imprimerie Nationale, Éditions E. Leroux.

——. 1928. *Syntaxe des parlers arabes actuels du Liban*. Paris: Geuthner.

Fleisch, Henri. 1974. *Etudes d’arabe dialectal*. Beirut: Dar El-Machreq.

Frayha, Anis. 1973. *A dictionary of non-classical vocables in the spoken Arabic of Lebanon*. Beirut: Librairie du Liban.

Kemmer, Suzanne. 1993. *The middle voice*. Amsterdam and Philadelphia: J. Benjamins.

König, Ekkehard and Peter Siemund. 1999. “Intensifiers and reflexives: A typological perspective”. *Reflexives: Forms and functions*, ed. Zygmunt Frajzyngier and Traci S. Curl, 41–74. Amsterdam and Philadelphia: J. Benjamins.

Lentin, Jérôme. 1991. “A propos de la valeur intensive de la IIème forme verbale en arabe syrien: Modalité et expressivité: Vers un renouvellement du système verbal?” *In honor of Wolf Leslau*, ed. Alan S. Kaye, 891–916b. Wiesbaden: O. Harrassowitz.

Mattsson, Emanuel. 1911. *Etudes phonologiques sur le dialecte arabe vulgaire de Beyrouth*. 2nd ed. Upsala: K.W. Appelberg.

Naïm-(Sanbar), Samia. 1985. *Le parler arabe de Rās Beyrouth*. Paris: Geuthner.

——. 1993. “Performatifs explicites et présent d’allocution”. *Langage et Société*, 41–61.

Table 9. Existential, possessive, locative constructions.

	Predicatoid	Order	Theme
Exist.	*fī*	Predicatoid + Subject (indefinite)	Pt
Poss.	*ʿənd-/maʿ-*	Predicatoid + Subject (indefinite)	Possessed
Loc.	Ø	Subject (definite) + Predicatoid	S

——. 2001. “L’expression du moyen et du réfléchi en arabe yéménite”. *Actes du colloque international Reflexive and Middle: Typological approaches.* Tunis.

——. 2003. “La prédication possessive et l’émergence de formes d’avoir en arabe oriental”. *Bulletin de la Société de Linguistique de Paris* 98.1:359–383.

——. 2004. “L’énoncé réfléchi et les stratégies de réflexivation dans les variétés dialectales de l’arabe”. *Actes du colloque L’énoncé réfléchi.* University of Lille 3 – Charles-de-Gaulle.

Nakhla, P. Raphaël. 1937. *Grammaire du dialecte libano–syrien.* Beirut: Imprimerie Catholique.

Samia Naïm, (CNRS-LACITO)

Bengali

People often make particular linguistic variants straightforward indexes of identity. This lacks analytic validity but reveals the linguistic ideologies upon which the politics of nationalism often turn (Bauman and Briggs 2003). Following Stewart (2001), we should be cautious of modern notions that linguistic form (e.g., Bengali discourse full of Sanskrit- or Perso-Arabic-derived words) directly reflects an author’s politico-religious stance or a Hindu or Muslim identity conceived as a pure essence.

Ask Bangladeshis what divides Muslim from Hindu speech and they will mention *pani* (vs. *jɔl*) ‘water’. This favorite index actually derives from Sanskrit. Yet, the ‘Muslim’ valeur of *pani* is a social fact. Such facts warrant attention to ideological representations of ‘Perso-Arabic’ lexemes in Bengali – and suggest that lists of loanwords require reanalysis in terms of ideologies.

1. Semantic domains

The semantic categories of Arabic loanwords in Bengali reveal the history of Bengali Islam. “The ordinary Bengali words for ‘paper’ *kagɔj* (Arabic *kaġaḏ*) and ‘pen’ *kɔlɔm* (Arabic *qalam*) [are] both . . . corrupted loanwords” (Eaton 1993: 293). Muslims spread literacy in Bengal, and associated terms reflect that fact.

Bengali Muslim kin terms are also mostly Arabic. Muslims usually call fathers *abba*; Hindus use *baba*. Some loanwords like *mullah* or *imām* designate Muslim social categories or reflect institutions of Mughal governance, e.g. the (now honorific) title *qāḍī* (*kāzī*). Then there are labels designating high birth – *sayyid*, *šex*, *ašraf* – which played a significant role in Bengal’s social history (Ahmed 1981). Bengali Muslims use different honorifics from Hindus, e.g. *šaheb* (like ‘Mister’). Muslim names are also typically Arabic. The 19th-century Islamization of Bengal involved rural Muslims rejecting their ‘Hindu’ (Bengali) names (Ahmed 1981:106).

Other salient loanwords denote ritual acts – e.g. *ḥajj*. In late 20th century Dhaka, Bengali newspapers were peppered with such terms; their use peaks during Ramadan. Musa (1995:93) lists 28, including *axeri munajat* ‘final prayer’, *id mobārak* ‘happy Id’, *zakāt* ‘alms’, *janāza* ‘funeral prayer’, and *mīlād mahfil* ‘gathering to celebrate [the Prophet’s] birth’.

2. Phonology and grammatical categories of loanwords

Phonological nativization of loanwords has been the rule in the past. Arabic /a/ in unstressed syllables has followed Bengali rules of vowel harmony to become /o/ in syllables preceding a high vowel (/u/ or /i/). Arabic consonants were generally replaced with their closest Bengali counterparts. The spelling of Arabic-derived terms has recently undergone ‘reform’. The Islamic Preaching Mission, once the Toblig Jamat, is now the Tablig Jamayat; *mowlanas* are now *mawlanas*, at least in writing (Musa 1995:93). Most Arabic loanwords are nouns, typically appearing in otherwise purely Bengali contexts and receiving Bengali affixation (*masjid-e* ‘in the mosque’) rather than Arabic morphology such as the definite article. Phrases like *bissɔ-ijtemāʿ* ‘world gathering’ or *ṣiyām-sadhɔna* ‘fasting-asceticism’ that join Arabic loanwords with Sanskrit derivatives are common. The 17th-century rise in non-nominal Arabic elements borrowed into Bengali was reversed in the 18th century – probably reflecting the declining fortunes of Persian under British hegemony (Mannan 1966:73). Among the non-nominal borrowings is the Arabic *ẓāhir*, used by the early 18th-century poet Vidyapati (Mannan 1966:67) in a verb phrase *karilo ẓāhir* ‘make manifest’. This illustrates the way Arabic loanwords can appear in Bengali verb phrases by virtue of the latter’s capacity to form compound verbs using nouns or adjectives and the Bengali pro-verb *kar* ‘do.’

3. COUNTS AND FREQUENCY OF ARABIC AND ISLAMICATE ELEMENTS IN BENGALI

There are no large corpus-based linguistic studies of Bengali, let alone of the frequency of Perso-Arabic terms in actual instances of contemporary Bengali discourse. Writing in pre-Partition Calcutta, S.K. Chatterji counted 2,500 Perso-Arabic terms in Bengali (Chatterji 1934:210; Ahmed 1981:121). Writing 30 years later in Islamic East Pakistan, Hilali (1967) listed 9,000 such loanwords. But the relation of such 'counts' to actual usage is unknown.

We find a range of loanword frequencies in a small corpus of carefully transcribed, naturally occurring Bengali speech of various registers. In 'Latifa's' 1992 lament (Wilce 1998) only 6 percent of total word tokens were Perso-Arabic loans. By contrast, in the Bengali 'translation' of an Arabic prayer offered at a 1991 wedding (Wilce 2002), about 33 percent of the total words are Arabic loans.

Arabic-laden prayers and other speech registers – and metadiscourses on the frequency of loanwords – reflect linguistic ideologies inseparable from postcoloniality and competing nationalisms (Irvine and Gal 2000). Such ideologies played a clear role in the history of Bengali.

4. HISTORY AND HISTORIOGRAPHY

Apparently, it was the Hindu poet Bharat Chandra in his poem *Mansingha Kāvya* (1752) who coined the term *dobhaṣi* Bangla 'dual language' (Haq 1957:174) for a register using many Perso-Arabic loanwords. Some *dobhaṣi* literature was written in the → *nastaʿliq* script, or in Bengali written from right to left.

Haq argues that *dobhaṣi* reflects the 19th-century Wahhabi movement in southern Bengal. Abdul Mannan, who wrote the definitive treatment of *dobhaṣi* literature in 1966, sees its origins in earlier Mughal patronage of Bengali. The first work on record "which has preserved evidence of the influence of the language of Muslim rulers [on Bengali] is the *Mɔnɔsavijɔyɔ* of Bipradās Piplāi", a Brahmin (ca. 1495 C.E., Mannan 1966:59).

Bharat Chandra wrote the following (from *Onnɔdamɔngɔl*):

na rɔbe prɔsad guṇ	[Persian, Arabic, Hindustani]
na hɔbe rɔsal	lack grace and poetic quality.
ɔtɔeb o kohi bhaṣa	I have chosen, therefore, the
yaboni misal	the mixed language of the Muslims.
ye hok se hok bhaṣa kavyo rɔs lɔye	The ancient sages have declared: "Any language may be used. The important thing is poetic quality" (Mannan 1966: 69–70; emphasis added)

This precolonial aesthetic of mixture gave way to a drive for purification.

In the 19th century, *dobhaṣi* Bengali borrowed even more Perso-Arabic lexemes, perhaps (ironically) reflecting forces unleashed by Halhed's (1969/1778) *Grammar of the Bengal Language*. Halhed considered foreign elements pollutants in the "pure Bengalese". He acknowledged "the modern [mixed] jargon of the kingdom" but declared the loanwords unintelligible outside large cosmopolitan towns (1969:xiv). Following Halhed's lead, British Orientalists and Hindu pundits working in Calcutta (Ft. William College) produced a Sanskritized register successfully promulgated as 'standard Bengali'. The intensification of Perso-Arabic borrowings in 19th-century *dobhaṣi* was thus a reaction to Orientalism and the Sanskritization of Bengali. As emerging Hindu and Muslim leaders competed for populist appeal, they declared the others' favored register (Sanskritized vs. *dobhaṣi*) "unintelligible to the masses".

Some of Halhed's successors – e.g. William Carey – at least for a time rejected linguistic purism. "A multitude of words, originally Persian or Arabic, are constantly employed in common conversation, which perhaps ought to be considered as *enriching* rather than corrupting the language" (Carey 1801:iii; emphasis in original). But Qayyum (1981) notes that later editions of Carey's *Grammar* omitted these words. Around 1850, British missionary James Long dubbed the Islamized form of Bengali "Musalman Bengali" (later called Musalmani Bangla – a form relevant to producing targeted translations of the Bible).

Around 1900, members of the Hindu Bengali intelligentsia, such as Dinesh Chandra Sen and Rabindranath Tagore, made "Bengali literature" central to their "romantic nationalism" (Chakrabarty 2004). They believed that "the national [Bengali] literature" could engender a mystical union of the divergent groups of Bengali speakers, transcending the Hindu-Muslim divide. While they somewhat naively advocated this vision, Muslims in the united British Indian state of Bengal formed a Muslim Literary Association (1911), sensing that the Bengal Literary Academy (formed in 1893) was in some subtle way simply a "Hindu Bengali Literature Society". But it was subtle. Hindu romantic nationalists did not advocate anything like the expurgation of Perso-Arabic words from Bengali. That was not what alienated Muslim literary figures. What the Hindu romanticists did so successfully was to promulgate a lexically Sanskritized Bengali that somehow appeared to be both the unmarked form of the language and the prestige variety.

5. Muslim attitudes to official support of Bengali

Colonial control required understanding and *ranking* various forms of Bengali. Two visions competed, ascribing to Bengali an enduring Hindu 'essence' or a growing Islamic influence. The first branded Musalmani 'unintelligible'. The second prompted colonial officers and some Muslim leaders to propose a 'separate language' for Bengali Muslims (Ahmed 1981:122). But colonial intelligentsia made Sanskritized Bengali represent not only a primordial essence but a prestige standard. Muslim opposition even to a Musalmani variety was a reaction to the putative Hindu essence of Bengali and to Musalmani's reputation as an "unsophisticated *patois*" (Ahmed 1981:126; cf. Qayyum 1981).

That some (not all, Anisuzzaman 1996) Muslims of the mid-20th century rejected Bengali language education indicates Bengali had become a bone of contention. Today, Bengali historians debate whether Partition was the fruit of the Raj's divide and conquer policy or the resolution of 'essential' differences. Metadiscourses about Bengali are part of that tortured history.

6. The status of Bengali in the East Pakistan and → Bangladesh eras

After Partition, the provincial East Pakistan government appointed an East Bengal Language Committee whose policy goals, summarized under the banner *sɔhɔj bangla* 'Simple Bengali', were: "i) that . . . Sanskritization . . . be avoided as far as possible by the use of simple phraseology . . .; ii) that . . . expressions and sentiments of Muslim writers should strictly conform to . . . Islamic ideology; and iii) that the words, idioms and phrases in common use in East Bengal, especially those in the Puthi . . . literatures be introduced in the language more freely" (Chowdhury 1960, as translated by Dil 1986:454).

The reference to the *dobhaṣi* Puthi literature makes clear that the "idioms . . . in common use" were Perso-Arabic. Pakistan had strong motivations for replacing Sanskritic with Islamicate derivatives. Appeals to linguistic 'simplicity' may sound democratic but, in Pakistan and elsewhere, often serve other agendas (Bauman and Briggs 2003).

In the late 1980s, Arabic expressions began displacing Persian ones among Muslim Bangladeshis; Muslims began using *Allāh ḥāfiẓ* rather than the Persian *Xoda ḥāfiẓ* 'go[o]db[ewith]ye'. In 1995, Bangla Academy Director Monsur Musa wrote: "Nowadays, in certain Bengali newspapers, an eagerness to substitute Arabic words for prevailing Persian terms can be seen. These newspapers use *ṣalāt* instead of *namaz*, *ṣiyām* instead of *roja* – and *Allāh* is considered better than *Xoda*" (1995:92; translation mine). Musa noted that the Arabic words in announcements of religious events made them quite hard for the average Bengali to understand – an echo of older claims?

7. Conclusion

While for some, proliferating loanwords represent an impure accretion on the language of the land of Bengal, for others they can signal the true identity of the Bangladeshi nation-state – an Islamic identity (Farukkhi 1990). And there are many positions in between, for example those who celebrate Bengali authors' *playful* use of Perso-Arabic loanwords (Anisuzzaman 1996).

The contemporary Bengali scene is a broad span over rapidly moving *pani*.

Bibliographical references

Ahmed, Rafiuddin. 1981. *The Bengal Muslims, 1871–1906: A quest for identity*. Delhi: Oxford University Press.

Anisuzzaman. 1996. "The Bengali language as a vehicle of creativity: After 1952". *Contemporary Bengali writing: Literature in Bangladesh, Bangladesh period*, ed. K.S. Murshid, 243–250. Dhaka: University Press Limited.

Bauman, Richard and Charles Briggs. 2003. *Voices of modernity: Language ideologies and the politics of inequality*. Cambridge: Cambridge University Press.

Carey, William. 1801. *A grammar of the Bengalee language*. Serampore: Mission Press.

Chakrabarty, Dipesh. 2004. "Romantic archives: Literature and the politics of identity in Bengal". *Critical Inquiry* 30:3.654–683.

Chatterji, Suniti Kumar. 1934. *The origin and development of the Bengali language*, I. London: Allen and Unwin.

Chowdhury, Munier. 1960. "The language problem in East Pakistan". *Linguistic diversity in South Asia: Studies in regional, social, and functional variation*, ed. Charles A. Ferguson and John J. Gumperz, 64–80. Bloomington: Indiana University Press.

Dil, Afia. 1986. "Diglossia in Bangla: A study of shifts in the verbal repertoire of the educated classes in Dhaka, Bangladesh". *The Fergusonian impact*. II. *Sociolinguistics and the sociology of language*, ed. Joshua Fishman, 451–465. Berlin and New York: Mouton de Gruyter.

Farrukhi, Asif Aslam. 1990. "Images in a broken mirror: The Urdu scene in Bangladesh". *Annual of Urdu Studies* 7.83–87.

Halhed, Nathaniel B. 1969. *A grammar of the Bengal language, 1778*. Menston, Yorkshire: Scolar Press.

Haq, Muhammad Enamul. 1957. *Muslim Bengali literature*. Karachi: Pakistan Publications.

Hilali, Shaikh Ghulam Maqsud and Muhammad Enamul Haq. 1967. *Perso-Arabic elements in Bengali*. Dhaka: Central Board for Development of Bengali.

Irvine, Judith and Susan Gal. 2000. "Language ideology and linguistic differentiation". *Regimes of language: Ideologies, polities, and identities*, ed. P. Kroskrity, 35–83. Santa Fe: School of American Research.

Mannan, Qazi Abdul. 1966. *The emergence and development of Dobhāsī literature in Bengal (up to 1855 A.D.)*. Dacca: Department of Bengali and Sanskrit, University of Dacca.

Musa, Monsur. 1995. *Bāṇlādešer rāstrabhaṣā* [The state language of Bangladesh]. Dhaka: Bangla Academy.

Qayyum, Muhammad Abdul. 1982. *A critical study of the early Bengali grammars: Halhed to Haughton*. Dhaka: Asiatic Society of Bangladesh.

Stewart, Tony K. 2001. "In search of equivalence: Conceiving Muslim-Hindu encounter through translation theory" *History of Religions* 40:3.261–288.

Wilce, James M. 1998. *Eloquence in trouble: The poetics and politics of complaint in rural Bangladesh*. New York: Oxford University Press.

——. 2002. "Tunes rising from the soul and other narcissistic prayers: Contested realms in Bangladesh". *Everyday life in South Asia*, ed. D. Mines and S. Lamb, 289–302. Bloomington: Indiana University Press.

James M. Wilce (Northern Arizona University)

Berber

1. Arabic and Berber

At the end of the 7th century, when the first Arabs came to Morocco during the Islamic conquests, Berber was the dominant language in Northern Africa with the exception of some Latinized towns on the Mediterranean coast, whose inhabitants were Romance speakers. After the Islamic conquests, the Arabization of the population seems to have been superficial and limited to the major towns. It was only after the arrival of the Arab tribes of the Banū Hilāl and Banū Sulaym in the 12th century that Arabization progressed in the Maghreb (Colin 1986:1193–1194). More than anything else, the language policies followed by the Maghreb countries in the last decades of the 20th century contributed to a strong Arab advance at the expense of Berber.

In view of the historical evolution and the fact that Berber is in constant contact with Arabic (in both its variants, Classical Arabic and dialects), it is hardly surprising that, after more than 14 centuries, all Berber dialects present a high percentage of Arabic loanwords nowadays. According to Chaker (1995:118) these percentages are 38 percent for Kabyli and 25 percent for Tashelhit. Touareg, being isolated and far from traditional cultural Arabic centers, has no more than 5 percent Arabic loanwords. Kossmann (1997:7) estimates that there are 20 percent loanwords from Arabic in the Berber dialect of Figuig.

Arabic loanwords occur in almost every field, but they are particularly important in religion and trade, as well as in economic and intellectual activities (Chaker 1995:118). In the last decades of the 20th century, as a consequence of Arabization policies, which tended to diminish the influence of French, many loanwords were introduced which were associated with administration and politics.

All examples quoted in this entry are taken from the following works: Laoust (1920), Aspinion (1953), Abdel-Massih (1971), Dallet (1985), Leguil (1985), Amard (1997), Durand (1998), and Kossmann (1997, 2000). Arabic lexical items are always quoted in their Maghrebi dialectal forms except for those specifically noted as a loan from Classical Arabic. Arabic verbs are always quoted in the 3rd person masc. sg. of the perfect and Berber verbs in the 2nd person masc. sg. of the imperative.

2. PHONOLOGY

Arabic phonemes such as /ḥ/, /ʿ/, /h/ have been incorporated in the Berber inventory, as well as the pharyngealized consonants /ṣ/, /ṭ/, /ḍ/ (Chaker 1995:1189, 1989:838; such phonemes may also appear in genuine Berber words: e.g. Riffian *aʿəddis* ‘stomach’); depending on the dialect, they may undergo further change:

i. /ḥ/: Arabic *fəllāḥ* > *afəllaḥ* ‘peasant’ (Tamazight, Tashelhit), Arabic *saḥḥār* > *asəḥḥar* ‘magician’, Arabic *l-ḥāžž* > *rḥažž* ‘pilgrim’ (Riffian), Arabic *ṣ-ṣaḥḥa* > *ṣṣaḥḥt* ‘health’, Arabic *l-bḥar* > *lbḥar* ‘sea’ (Tamazight).

ii. /h/: this phoneme rarely occurs and is frequently elided: Arabic *l-fqīh* > *lfqih* ‘learned, erudite’ (Tashelhit), Arabic *məžhūl* > *aməžhul* ‘irreligious’ (Kabyli), Arabic *l-bhīma* > *ləbhimt* ‘beast of burden’ (Tamazight), Arabic *l-bhāyəm* > *rbāym* ‘herd of goats’ (Riffian), Classical Arabic *šahāda* > *šada* ‘profession of Muslim faith’ (Kabyli).

iii. /ʿ/: Arabic *ʿass* > *ʿass* ‘to watch’, Arabic *ʿāšūra* > *taʿšurt* ‘feast of Ashoura’ (Tashelhit), Arabic *s-sāʿa* > *tassaʿt* ‘then’ (Tashelhit).

iv. /ṣ/: occurs in loanwords like Arabic *mṣalla* > *mṣalla* ‘place for prayer’ (Tashelhit), Arabic *ṣ-ṣadaqa* > *ṣṣadaqat* ‘alms’ (Tashelhit), Classical Arabic *al-qiṣṣa* > *lqiṣt* ‘story’ (Tashelhit). But /ṣ/ is often voiced to /ẓ/: Arabic *ṣām* > *ẓum* (Riffian, Tamazight, Tashelhit) and *uẓum* (Kabyli) ‘to fast’, Arabic *ṣalla* > *ẓẓall* (Tamazight, Tashelhit) and *ẓajj* (Riffian) ‘to pray’, Arabic *l-ḥəmməš* > *lḥiməẓ* (Riffian) and *ḥumməẓ* (Tamazight) ‘chick peas’.

v. /ṭ/: a single /ṭ/ is voiced to /ḍ/: Arabic *ṭbīb* > *aḍbib, iḍbibən* ‘doctor, doctors’, Arabic *xayyāṭ* > *axəyyaḍ* ‘tailor’, Arabic *ṭāləb* > *ḍḍalb* ‘religious teacher’, Arabic *ṭāžīn* > *ḍḍažin* ‘clay pot in which stew is cooked’, Arabic *ṭʿām* > *ḍḍʿam* ‘couscous’, Arabic *fṭar* > *fḍər* ‘to take breakfast, to lunch’, Arabic *l-qəfṭān* > *lqəfḍan* ‘kaftan’, Arabic *qṭaʿ* > *qḍəʿ* ‘to cross’, Arabic *šrəṭ* > *šṛəḍ* ‘to impose a condition’, Arabic *l-ḥīṭ* > *lḥiḍ* ‘wall’, Arabic *ṭabbāx* > *aḍəbbax* ‘cook’ (Tashelhit). But geminated /ṭ/ is never voiced: Arabic *ʿaṭṭār* > *aʿaṭṭar* ‘wandering salesman’, Arabic *ḥaṭṭāb* > *aḥəṭṭab* ‘woodcutter’ (Tashelhit).

vi. /ḍ/: Arabic *ḍədd* > *ḍəṭṭ* ‘against’ (Figuig), Arabic *r-rawḍāt* > *rrawḍat* ‘cemeteries’, Arabic *l-xuḍra* > *lxʷḍərt* ‘legume’ (Tashelhit) and *lxʷəḍṛa* (Kabyli), Arabic *ḍāq* > *ḍaq* ‘to be depressed’ (Figuig), Arabic *mḍiyyəq* > *mḍəyyəq* ‘narrow’ (Riffian), Arabic *ḍawwa* > *ḍawwa* ‘to illuminate’ (Figuig).

vii. /q/: reflexes of Classical Arabic *_q_ are /q/ and /g/. Examples of /q/ are very common: Arabic *l-wərq* >*lwərq* ‘leaves’ (Tashelhit, Kabyli, Shawiya, Nefusi, Siwa), Arabic *qšūr* > *aqšur* ‘bark’ (Tashelhit, Kabyli, Shawiya), Arabic *l-ʿaṛq* > *lʿaṛq* ‘root’ (Tashelhit, Nefusa, Siwa), Arabic *qdūḥ* > *aqduḥ* ‘jug for water’ (Riffian), Arabic *s-sūq* > *ssuq* ‘market’ (Riffian, Tashelhit). Cases in which *_q_ > /g/ are always loans from Hilalian (Bedouin) dialects (this reflex being one of the most characteristic features of these dialects; Heath 2002: 141–142): Arabic *l-gāfla* > *lgafəlt* ‘caravan’ (Figuig), Arabic *l-guddām* > *l-gʷddam* ‘the front, ahead of’ (Tashelhit), Classical Arabic *qaʿūd* > *agʿud* ‘young camel’ (Figuig), Arabic *l-gāyla* > *lgaylət* ‘hottest part of a summer day’ (Figuig).

viii. /j/: /ž/ and, in some cases in Morocco, /g/ normally correspond to the standard pronunciation of Classical Arabic /j/. Examples of /ž/: Classical Arabic *jiḍr* > *žədra* ‘trunk’ (Tashelhit, Kabyli, Shawiya), Classical Arabic *jaḏaba* > *ždəb* ‘to abandon oneself to ecstatic excitation’ (Tashelhit), Classical Arabic *al-jawz* > *lžuž* ‘walnuts’ (Tashelhit, Wargla, Kabyli, Shawiya), Classical Arabic *jurf* > *ažarif* ‘cliff’ (Tashelhit), Classical Arabic *al-jār* > *aržal* ‘neighbor’ (Tashelhit). In some cases, the phoneme /j/ >

/g/, because in Maghrebi Arabic dialects /g/ appears as a dissimilatory deaffrication of /j/ in stems with sibilants (Heath 2002:136–137): Classical Arabic *jazzār* > Moroccan Arabic *gəzzār* > *agəzzar* 'butcher' (Tashelhit, Riffian), Classical Arabic *masjid* > *ṯaməsgiḏa* 'mosque' (Riffian), Classical Arabic *jazīra* > Moroccan Arabic *gzīra* > *ṯagzirt* 'island' (Riffian).

ix. All Classical Arabic interdentals (/ṯ/, /ḏ/, /ḏ̣/) are replaced by the corresponding occlusives: *jiḏr* > *žədra* 'trunk' (Tashelhit, Kabyli, Shawiya), *(aḏ-) ḏahab* > *ddhəb* 'gold' (Kabyli), *ḏāb/yḏūb* > *dub* 'to melt', *ṯalāṯīn* > *tlatin* 'thirty' (Tamazight). In some Berber dialects, however, a restitution of the old Arabic interdentals took place, as a result of secondary affrication of occlusives. This is the case in Riffian: Classical Arabic *ṯalāṯa* 'three', *ṯamāniya* 'eight'> Maghrebi Arabic *tlāta*, *tmənya* > Riffian *ṯlāṯa*, *ṯmənya* (Kossman 2000:160; but *tmənya* in Tamazight, Durand 1998:112).

x. Due to the phonological changes many Arabic loans, especially in Rif Berber, are hardly recognizable: Arabic *qla* > *qra* 'to fry', Arabic *qləb* > *qrəb* 'to overthrow', Arabic *līla* > *jirṯ* 'night', Arabic *fəllāḥ* > *afjaḥ* 'peasant', Arabic *l-lūz* > *jəwz* 'almonds' (not related to Arabic *jawz* 'nuts'), Arabic *ṣalla* > *ẓaj* 'to pray', Arabic *səlləm* > *səjəm* (*səddəm*) 'to greet', Arabic *l-xəll* > *rxəj* 'vinegar', Arabic *mməllaḥ* > *aməjaḥ* 'salted', Arabic *l-bṣəl* > *rəbṣər* 'onions'. In Tamazight, too, loanwords are sometimes difficult to recognize: Arabic *salāma* > *slant* 'peace', Arabic *ḥizām* > *taḥzant* 'belt' (but *tiḥzamin* in plural).

3. MORPHOLOGY OF ARABIC NOMINAL LOANS

Arabic loans have been adapted to the respective Berber nominal patterns:

i. masculine: Arabic *fəllāḥ* > *afəllaḥ*, pl. *ifəllaḥən* 'peasant, farmer', Arabic *bərrāḥ* > *abərraḥ*, pl. *ibərraḥən* 'town crier', Arabic *ḥbīb* > *aḥbib*, pl. *iḥbibən* 'dear friend' (Tamazight), Arabic *ḥūli* > *aḥuli*, pl. *iḥuliyn* 'sheep, ram', Arabic *ḥəžžām* > *aḥəžžam*, pl. *iḥəžžamən* 'barber', Arabic *maḥḍār* > *aməḥḍar*, pl. *imḥəḍar* 'student, pupil', Arabic *musləm* > *amusləm*, pl. *imusəlmən* 'Muslim [masc.]', Arabic *nəššād* > *anəžžad*, pl. *inəžžadən* 'poet, singer', Arabic *səḥḥār* > *asḥḥar*, pl. *isəḥḥarən* 'magician', Arabic *xəbbāz* > *axəbbaz*, pl. *ixəbbazən* 'baker' (Tashelhit), Arabic *məndīl* > *aməndil*, pl. *iməndal* 'scarf', Arabic *ḥfīr* > *aḥfir*, pl. *iḥəfrawən* 'pit' (Riffian), Arabic *xəddam* > *axəddam*, pl. *ixəddamən* 'worker, laborer', Arabic *ḥəddād* > *aḥəddad*, pl. *iḥəddadən* 'blacksmith', Arabic *nəžžār* > *anəžžar*, pl. *inəžžarən* 'carpenter', Arabic *bənnāy* > *abənnay*, pl. *ibənnayən* 'bricklayer, mason', Arabic *xərrāz* > *axərraz*, pl. *ixərrazən* 'shoemaker' (Tashelhit).

ii. feminine: Arabic *qbīla* > *ṯaqbilt*, pl. *ṯiqəbbal* 'tribe', Arabic *mdīna* > *ṯamdint*, pl. *ṯiməddam* 'town' (Riffian), Arabic *ḥmāma* > *taḥəmmamt*, pl. *tiḥəmmamin* 'dove', Arabic *zlāfa* > *tazlaft*, pl. *tazlafin* 'large wooden or clay plate', Arabic *xābya* > *txabit*, pl. *tixabyin* 'big jug', Arabic *bḥīra* > *tabḥirt*, pl. *tibḥarin* 'garden', Arabic *msəlma* > *tansəlmt*, pl. *tinsəlmin* 'Muslim [fem.]' (Tamazight), Arabic *xayma* > *taxyamt*, pl. *tixyamin* 'tent', Arabic *qdīma* > *taddimt*, pl. *tiqdimin* 'old woman', Arabic *šāqūr* > *tašaqurt*, pl. *tišuqar* 'hatchet', Arabic *ḥrīra* > *taḥrirt* 'soup' (Tashelhit).

In some words final *-t* was interpreted as the Berber feminine morpheme (Aspinion 1953:11): Arabic *ḥānūt* > *taḥanut*, pl. *tiḥuna* 'store' (Tashelhit), Arabic *yāqūt* > *talyaqut*, pl. *tilyaqutin* 'sapphire', Arabic *z-zīt* > *zzit* 'oil' (Tamazight, Tashelhit), Arabic *l-mūt* > *lmut* 'death' (Tashelhit), Arabic *l-bīt* > *lbit* 'room' (Tashelhit). In other cases, there is no obvious reason for the change of gender: Arabic *xātəm* > *ṯxaṯəmt*, *talxtamt* 'ring' (Riffian, Tamazight; *xātəm* is also feminine in some Moroccan dialects), Arabic *məsžīd* > *tamzgida*, *timzgidawin* 'mosque' (Tamazight).

iii. loanwords with Arabic article (with respective assimilations), partly with additional Berber feminine morpheme: Arabic *l-xənša* > *talxənšt* 'sack, bag' (Tashelhit), Arabic *l-bərqūq* > *lbərquq* 'plum', Arabic *l-kammūn* > *lkkamun* 'cumin' (Tamazight), Arabic *l-xurṣa* > *talxurṣt* 'ring' (Tashelhit), Arabic *l-ḥāžž* > *rḥažž* 'pilgrim', Arabic *l-qanṭra(t)* > *rqəndarṯ* 'bridge' (Riffian), Arabic *l-kās* > *lkas* 'glass, cup', Arabic *ṣ-ṣīnīya* > *ṣṣinit*

> ‘tray’, Arabic *d-dālya* > *ddilit* ‘vine’, Arabic *d-dunya* > *ddunit* ‘world’, Arabic *l-qāləb d-s-sukkār* > *lqaleb n sekkwar* ‘sugar loaf’, Classical Arabic *al-ʾiḏāʿa* > *lidaʿa* ‘radio station’, Arabic *l-bhīma* > *ləbhimt* ‘beast of burden’ (Tashelhit), Arabic *ṣ-ṣaḥḥa* > *ṣṣaḥt* ‘health’, Arabic *l-ġāba* > *lġabt* ‘forest’ (Tashelhit). Modern loanwords normally take the Arabic article: Classical Arabic *al-muraššaḥ* > *lmuṛəššəḥ*, *lmuṛəššəḥin* ‘deputy’, Classical Arabic *al-jamāʿa al-qarawiyya* > *lžamaʿa lqaṛawyya* ‘the village council’.

Some loanwords keep their original Arabic plurals (Aspinion 1953:59–60; Durand 1998: 97; Kossmann 2000:48): Arabic *z-zənqa*, pl. *z-znāqi* > *zzənqəṯ*, pl. *zznaqi* (Riffian), Arabic *ṭ-ṭāžīn*, pl. *ṭ-ṭwāžən* > *ḍḍažin*, pl. *ḍḍwažən* ‘clay pot in which stew is cooked’, Arabic *l-wāldīn* > *lwaldin* ‘parents’, Arabic *l-bāb*, pl. *l-bībān* > *lbab*, pl. *lbiban* ‘door’, Arabic *l-bəṛma*, pl. *l-bəṛmāt* > *lbəṛma*, pl. *lbəṛmat* ‘pot’, Arabic *l-fənni*, pl. *l-fənniyīn* > *lfənni*, pl. *lfənniyin* ‘technician’, Arabic *l-bhīma*, pl. *l-bhāym* > *lbhimt*, pl. *lbaym* ‘beast of burden’, Arabic *l-lūn*, pl. *lə-lwān* > *llun*, *lalwan* ‘color’, Arabic *l-wuqt*, pl. *l-awqāt* > *luqt*, *lawqat* ‘time’, Arabic *s-sūq*, pl. *lə-swāq* > *ssuq*, pl. *laswaq* ‘market’, Arabic *lə-ḥwāyəž* > *lḥwayž* ‘clothes, things’, Arabic *l-xədma*, pl. *l-xədmāt* > *lxdəmt*, pl. *l-xədmāt*, Arabic *l-ʿīn*, pl. *lə-ʿyūn* > *lʿin*, pl. *lʿyun* ‘spring’, Arabic *lə-bḥər*, pl. *lə-bḥūr* > *lbḥər*, pl. *lbḥur* ‘sea’, Arabic *šāhəd*, pl. *šhūd* > *šahd*, pl. *šhud* ‘witness’ (Tashelhit, Tamazight).

Not only isolated words were borrowed, but also whole syntagms such as genitive constructions: *qəṭṭaʿ əṭ-ṭriq*, pl. *qəṭṭaʿin əṭ-ṭriq* ‘highwayman’ (Figuig). The fact that many loanwords were borrowed with the article, however, sometimes led to deviant Arabic constructions like *lʿid lmulud* (i.e., with two articles in a genitive construction, a construction impossible in Arabic) instead of *ʿīd l-mulūd* ‘the Prophet’s birthday’.

Arabic *bu* ‘father of’ is quite productive and is combined with Arabic, French, and Berber nouns (Aspinion 1953:47; Durand 1998:110–111). Examples from Tashelhit: *bu lbuṣṭa* ‘postman’ (< Arabic *l-buṣṭa* < French *poste*), *bu lḥəmmam* ‘public bath attendant’ (< Arabic *l-ḥəmmām* ‘bath’), *bu lqəhwa* ‘coffee shop owner’ (< Arabic *l-qəhwa* ‘coffee shop’), *bu tuġmas* ‘dentist’ (< Tashelhit *tuġmas* ‘tooth’), *bu tḥanut* ‘shopkeeper’ (< Arabic *ḥānūt* ‘shop, store’), *bu tiyni* ‘tailor’, *bu mḥand*, *bu mḥamməd* ‘hedgehog’ (< Classical Arabic *muḥammad* ‘Muhammad’), *bu təgra* ‘turtle’.

Nouns referring to relatives were borrowed with the Arabic pronominal suffixes (Kossmann 2000:47): Arabic *xāli* > *xali* ‘my maternal uncle’, Arabic *xālti* > *xalti* ‘my maternal aunt’, Arabic *ʿammi* > *ʿammi* ‘my paternal uncle’ (Tamazight, Tashelhit), Arabic *žəddi* > *žəddi* ‘my grandfather’ (Tashelhit). In some dialects, these loanwords retain their original Arabic plurals (Kossmann 2000:48): *ʿmumi* ‘my paternal uncles’ (< Arabic *ʿmūm*), *xwali* ‘my maternal uncles’ (< Arabic *xwāl*).

4. Morphology of Arabic verbal loans

Verbal morphology is doubtless the part of the language least affected by Arabic influence. Verbs are always borrowed as lexical items, which do not affect morphology. Since the Berber verbal system contains verbal forms similar in appearance to those in Arabic (although not necessarily in meaning), their adaptation does not present any difficulty. The verbal pattern R1R2əR3 is represented both by genuinely Berber terms such as *mġər* ‘to grow’ (Figuig), *ffəġ* ‘to go out’ (Tashelhit), and by loans from Arabic such as *fṛəḥ* > *frəḥ* ‘to be happy’, *ddən* > *ddən* ‘to make the call to prayer’. The same phenomenon occurs with other patterns like R1aR2R3, R1āR3, or R1əR2R2əR3: *sawl* ‘to speak’ and Arabic *šāwər* > *šawr* ‘to consult with’, Arabic *sāfər* > *safr* ‘to travel’, *šərrəs* ‘to knot’ and Arabic *kəmməl* > *kəmməl* ‘to finish’, Arabic *ṣḥa* > *ṣḥu* ‘to be healthy’, Arabic *ḏāb/yḏūb* > *dub* ‘to melt’, Arabic *səmma* > *səmma* ‘to name’, Arabic *ḏāq/yḏūq* > *duq* ‘to taste’, Arabic *dāṛ/ydūṛ* > *ḍuṛ* ‘to surround’ (Tashelhit).

5. Other loans

Numerals from four upward are in almost all Berber dialects loans from Arabic (Durand 1998:112–113): *rəbʿa* ‘four’, *xəmsa* ‘five’, *sətta* ‘six’, *səbʿa* ‘seven’, *tmənya* ‘eight’, *təsʿa* ‘nine’, *ʿəšra* ‘ten’

Loans common to almost all Berber dialects are *labas* ‘fine’, *mslxiṛ* ‘good evening’, *fiməṛṛa* ‘at once’, *ḷḷayhənnik* ‘goodbye’, *ḷḷaysəllm* ‘[response to hello]’, *nʿam* ‘yes [in response to someone calling your name]’, *ṣaḥḥa* ‘thank you’, *ṣbaḥəlxir* ‘good morning’, *šḥal* ‘how many, how much?’, *walu* ‘nothing’, *yaḷḷah* ‘let us’, *bsif* ‘by force’, *bzayd* ‘more than’, *iʕni* ‘that is to say’, *linnahu* ‘because’, *təqriban* ‘almost’, *bla* ‘without’, *šway šway* (Riffian), *šḥalmənwaḥəd* ‘how many times?’ (Figuig).

BIBLIOGRAPHICAL REFERENCES

Abdel-Massih, Ernest T. *A reference grammar of Tamazight (Middle Atlas Berber)*. Ann Arbor: University of Michigan, Center for Near Eastern and North African Studies.

Amard, Pierre. 1997. *Textes berbères des Aït Ouaouzguite, Ouarzazate, Maroc*. Aix-en-Provence: Edisud.

Aspinion, Robert. 1953. *Apprenons le berbère: Initiation aux dialectes chleuhs*. Rabat: Félix Moncho.

Chaker, Salem. 1989. “Arabisation”. *Encyclopédie berbère*, VI. Aix-en-Provence: Édisud.

——. 1995. *Linguistique berbère: Études de syntaxe et de diachronie*. Paris and Louvain: Peeters.

Colin, Georges Séraphin. 1986. “al-Maghrib. VII: Aperçu linguistique”. *Encyclopédie de l’Islam*, V. 2nd. ed. Leiden: E.J. Brill.

Dallet, Jean-Marie. 1985. *Dictionnaire français–kabyle*. Paris: Selaf.

Durand, Olivier. 1998. *Lineamenti di lingua berbera: Varietà tamazight del Marocco centrale*. Rome: Università degli Studi ‘La Sapienza’.

Encyclopédie berbère. 1984–2004. 26 vols. Aix-en-Provence: Édisud.

Heath, Jeffrey. 2002. *Jewish and Muslim dialects of Moroccan Arabic*. London: Routledge Curzon.

Kossmann, Maarten. 1997. *Grammaire du parler berbère de Figuig (Maroc oriental)*. Paris and Louvain: Peeters.

——. 2000. *Esquisse grammaticale du rifain oriental*. Paris and Louvain: Peeters.

Laoust, Emile. 1920. *Mots et choses berbères: Notes de linguistique et d’ethnographie: Dialectes du Maroc*. Paris: Challamel. (Repr. Rabat: Société Marocaine d’Edition, 1983.)

Leguil, Alphonse. 1985. *Contes berbères du Grand Atlas*. Paris: Edicef.

JORDI AGUADÉ (University of Cadiz),
PETER BEHNSTEDT (Chipiona, Spain)

Berber Loanwords

This entry deals with contact between → Berber and Arabic, the two major languages used in Morocco, and in the rest of North Africa, with a wide array of language variation. First, an overview is given of the history and sociolinguistic situation of the languages involved. Then a closer look is taken at the contact results between the two languages in the areas of phonology, lexicon, and morphosyntax.

1. A BRIEF HISTORY OF BERBER/ARABIC CONTACTS

The oldest known populations of North Africa are the Libyans or *libici*. Morocco, as part of this area, had its share in the colonization waves under the Phoenicians, the Romans, and finally the Arabs or Muslims in the 7th century. The last wave of Arabic-speaking people arrived in Morocco in the 15th century, as a result of the fall of Muslim Spain. Each of these waves had a large impact on social and cultural life in Morocco. It is clear that Arabic – in its different forms – was the main language the last newcomers brought with them. Thanks to the prestige of Arabic as the language of the *Qurʾān*, and its status as the language of the conquering people, it was naturally especially present in the influential domains of politics and religion. The indigenous population maintained its own language to a certain degree, particularly in the countryside and outside the arena of religion and politics – or administration – where Arabic was and still remains a clear asset (see, e.g., Chaker 1989: 834–842).

Since its arrival on North African territory, Arabic has undergone a number of changes. As is the case with natural languages, some of these changes are caused by natural processes known as internal changes, while others can be traced back to external influences, mainly through contact with Berber.

In present-day Morocco, Berber is still widely spoken, and since September 2003 has also been taught in public primary schools. Berber is commonly subdivided into three main dialects, namely the Tarifit dialect spoken in the northeastern part of Morocco, Tamazight spoken mainly in the Middle Atlas and the neighboring valleys, and Tashelhit spoken mainly in the High Atlas, the Anti Atlas, and the Sous Valley. The number of Berber speakers in Morocco can only be estimated, since a general linguistic survey has never been published. Current estimates range from 40 to 50 percent (Boukous 1997) of the total population of Morocco, which was close to 29 million in 1997. Despite the large number of speakers, and the continuous pleas of cultural associations for an official recognition of Berber, the constitution does not recognize it as an (additional) official language, reserving this status exclusively to Arabic.

It is difficult to provide a well-founded account of where Berber is used. As an oral medium, it remains widely used in informal domains. Its limited presence in the national mass media acts against its spread and the possibility of enlarging the mutual intelligibility among speakers of the different varieties. The presence of Berber in written media is also very scarce, although a large number of literary works, in Latin and Arabic scripts, have proliferated especially in the last decades, accompanying a cultural revival of Berber. The recent introduction of Berber to schools will certainly contribute to the spread of the written text. There are no daily newspapers in any of the Berber dialects.

This brief account concludes with a note about the large number of Berber inscriptions which have been found so far in North Africa. These were written in the Tifinagh script, a Libyan script, dating back to about the 6th century B.C.E., and are mostly inscriptions on gravestones and in caves. More elaborate are the manuscripts in Berber dating back to the 11th, 17th, and 18th centuries C.E.

2. Linguistic borrowing from Berber into Arabic

Borrowing is defined as referring to "linguistic forms being taken over by one language or dialect from another" (Crystal 1980), or the "incorporation of foreign features into a group's native language: the native language is maintained but is changed by the addition of the incorporated features" (Thomason and Kaufman 1988:37). The process of borrowing is assumed to take place at all levels of language: phonology, morphology, syntax, and lexicon, including the levels of semantics and pragmatics. For an elaborate discussion of the issue of borrowing and related phenomena, such as interference, borrowing scales, etc., see Thomason (2001) and Field (2002). Here, the term 'borrowing' will be used to refer to words or utterances that witness some stability in their usage in the borrowing language, viz. Moroccan Arabic. The borrowed item will in general show compatibility with the grammar rules of the borrowing language.

The mutual interference between Berber and Arabic is a good illustration of shift-induced interference and borrowing. The structure of Moroccan Arabic speaks of heavy phonological and morphosyntactic interference and relatively weak lexical borrowing, while Berber shows a large amount of loanwords or lexical borrowing from Arabic. In one of the few works on Berber loanwords in Arabic, Chafik (1999) enumerates close to 1,400 lexical items that he traces back to Berber, although in a few cases he admits that the source–target relationship is not unequivocal. The amount of lexical borrowing from Arabic into Berber is rather large (→ Berber), and amounts to about a third of the lexicon (Taifi 1997). The data presented below are drawn mainly from Chafik (1999).

In the case of well-documented languages, it is relatively less complicated to establish which form has been borrowed by which language, depending mostly on what the researcher is looking for and on the degree of detail found in the transcriptions of the documents. Arabic presents in this respect a particularly interesting case. While there is a large body of documents in Arabic, there is relatively little written in or on Arabic dialects, such as Moroccan Arabic. It is therefore very difficult to make reliable statements about the history of such dialects. The existence of some material in Berber (especially in the variety of Tashelhit in the southern part of Morocco) dating to the 12th century is certainly very helpful in that it opens a window on the state of the language at that time. This material consists mostly of bilingual Arabic–Berber wordlists, and more than 200 texts dating as far back as the 16th century (see van den Boogert 1997 for an overview of Berber manuscripts). Outside these sources there is not much writing which can be used for the purpose of an analysis of borrowing or for the study of the history of the dialects in question.

Before proceeding to a classification of material borrowed from Berber into Moroccan Arabic, a few additional points must be made about the paradox created by the genetic relationship of these two languages: on the one hand it presumably facilitates borrowing because of the congruent structure of the two languages (see Muysken 2000), but on the other hand it makes it difficult to decide unequivocally about borrowed material (see also Taifi 1997 for a similar remark). A concrete example is that of the pre-verb form *lā* in the Chefchaouen variety of

Moroccan Arabic, given by Aguadé (2004). A number of authors have assumed that this preverb form is borrowed from Berber, which possesses a similar form *illa* 'to be, to exist', while others maintain that it might be derived from the Arabic participle *lahi* (Cohen 1924:274, 1963: 113, cited in Aguadé 2004:17). Aguadé maintains that the form in question could simply be derived from the form *raḥ* > *laḥ* > *lā*.

A lexical example of the same problem, cited in Chafik (1999:22), is the word *adfu* referring to a certain type of refreshing food prepared for women who have just given birth. The word is used both in Moroccan Arabic and Berber, and has a corresponding form in Classical Arabic *dafiʾa/yadfaʾu* 'to get warm', and a related one in Berber meaning *idfa/idfu* 'to be strong and energetic'. The source and target language in this case cannot be defined with certainty.

Next to the issue of genetic relationship and the complications it raises, one should also be aware of the fact that Moroccan Arabic as an independent language does have its own grammar system and as a natural language is capable of innovation. This means that forms found in Moroccan Arabic which cannot be traced back to Classical Arabic should not be immediately ascribed to influence from Berber. An example in this connection is the phenomenon of serialization, whereby the verbal predicate is made up of two verbs, of which one might lose its inflected character (Versteegh 1984:100). This phenomenon, although discussed as an illustration of pidginization in Arabic dialects by Versteegh, is also found in Berber and in Arabic dialects in Egypt, Iraq, Lebanon, and Chad, among others (see Versteegh 1984:99–106 for an elaborate discussion; → serial verbs). By looking only at Moroccan Arabic and Berber, and knowing that Classical Arabic has no asyndetic constructions of the type discussed above, one might be tempted to conclude that this construction has made its way into Moroccan Arabic via Berber, which remains a possibility although it cannot be backed with empirical data. Here is an example from Ḥijāzī Arabic and its equivalent in Berber (Tarifit variety) to illustrate how far-fetched such an interpretation would be (Arabic examples from Sieny 1978 cited in Versteegh 1984:102):

Hijāzī Arabic

raḥ	*katab*	*la-na*
go.3s.perf	write.3s.perf.	to-us

'he went (and) wrote to us'

Tarifit Berber

i-raḥ	*i-ura*	*anəġ*	*dd*
3s-go.perf.	3s-write.perf.	to.us	directionality particle

In addition to problems at the level of the structure of the languages involved in a contact relationship, the researcher has to pay attention to the issue of sociolinguistic variation, especially in studying borrowing in the case of a vernacular like Moroccan Arabic: which variety is being examined? Can results in one variety apply to another variety? The most relevant issue here is the degree of formality of the language at hand. The highly formal 'Middle Arabic' would certainly contain fewer Berber borrowings than the everyday variety. Middle Arabic draws more from Classical Arabic or Standard Arabic, especially at the level of the lexicon (see Youssi 1992 for a discussion of the different levels of Moroccan Arabic).

3. PHONOLOGY

At the level of the sound system, the influence of Berber on Moroccan Arabic dialects was noted by several authors. Brunot (1921:351, cited in Chtatou 1997:105) reports that the Branès dialect has conserved some Berber traits, like the spirantization of *b*, *t*, and *d* and the affrication of *t* and *j*. The disappearance of the glottal stop from Moroccan Arabic dialects is considered by Chtatou (1997) to be a direct result of influence from Berber. Here are a few examples, reproduced from Chtatou (1997:107), with the examples in (1) illustrating deletion and those in (2) illustrating deletion and insertion (of a compensating segment):

(1)	*ʾaxaḏa*	*xda*	'to take (something)'
	faʾs	*fas*	'a pick'
	al-māʾ	*lma*	'water'
(2)	*ruʾūs*	*ryus*	'heads'
	masāʾil	*msayl*	'things'
	maṣāʾib	*msayib*	'plights'

At the level of syllable structure, the influence of Berber can also be detected. Although other dialects of Arabic allow certain phonotactic combinations which are unacceptable in Classical Arabic, like initial consonant clusters in the variety spoken in the north of Beirut (e.g.

ḍrabt), we can say with certainty that the Maghreb Arabic dialects are the only ones known to have complete forms – phonological phrases – without any full vowel. The following example from Moroccan Arabic is illustrative: *xdəm-t f l-məktəb* work-1st in the-office 'I have worked in the office'. This high degree of consonantism is a characteristic of Berber, and its 'spread' to Moroccan Arabic is very likely to be due to Berber influence (for more on Berber syllable structure see Dell and El Medlaoui 2002).

4. Berber loan words in Moroccan Arabic

As indicated above, the main source of the data in the following sections is Chafik (1999). A few examples will be reproduced here to illustrate the different categories under which these borrowings fall (in the original work these are listed alphabetically). Chafik (1999) also provides a discussion of areas of grammar in Moroccan Arabic susceptible of having undergone influence from Berber. These will also be dealt with in the appropriate section below.

4.1 *Arabic nouns with Berber morphology*

This category provides examples of forms that have a Berber morphological shape, and an Arabic root morpheme. The forms listed here all have the typical Berber feminine discontinuous marker *t-*... ...*-t*.

4.1.1 Names of professions

taymamt (also *talimamt*) 'profession of imam'; *tabašawt* 'pasha' (< Turkish *paşa*); *tabeyyaˁt* 'spying (< Arabic *bāˁa/yabīˁu* 'to sell'); *tanejjart* 'profession of carpenter'.

4.1.2 Nouns denoting properties

taʾadamit 'humanity'; *taderrit* 'childishness' (< Arabic *durriya* 'offspring').

4.2 *Berber lexemes with Berber morphology*

4.2.1 Nouns denoting properties

tabuhaliyt 'stupidity' (< *abuhali* 'stupid, dumb'); *taduġrit* 'straightforwardness' (< *duġri* 'straight').

4.2.2 Traditional dress and household items

The following terms refer to culturally specific items, which may have been the main reason why they made their way into Moroccan Arabic: *aḥlas* 'dress'; *aˁban* 'dress made of wool'; *agdwar* 'woolen cover'; *asfeṭ* 'treasure box meant for the bride to store her valuables'; *asettour* 'porch for sheep serving as shelter in bad weather conditions'; *afrur* 'pottery'; *afrag* 'fence'.

4.2.3 Food, flora and fauna

This category is similar to the previous one, in that it deals with items that are specific to the culture and environment of the Berbers: *tizbibit* 'black olives ready for consumption'; *azeffan* 'lobster'; *taġzalt* 'kind of fish'; *adfu* 'refreshing food prepared for the woman who has just given birth'.

4.2.4 Place names

tiṭṭawin Tetouan, a city in the North Western part of Morocco, lit. 'eyes'; *anfa* district in the city of Casablanca, lit. 'height'; *asafi* Safi, a city on the Atlantic Coast of Morocco, lit. 'river delta'; *agadir* Agadir, city on the coast of Morocco, lit. 'granary'.

4.3 Berber lexemes with Arabic morphology

zenṭiṭ 'tail'. In this form the disappearance of the Berber prefix *a* (bound determiner morpheme) could be interpreted as a sign of the Arabization of the form in question. The definite form in Moroccan Arabic is *z-zentit*, which clearly bears Arabic morphology. The derived adjectival form *zenṭiṭ-i* 'womanizer' is also an example of a Berber lexeme with Arabic morphology.

A different category of nouns is particularly interesting in that it poses the additional problem of which of the two languages borrowed a given form from a third language. For example, the two Latin words below are, according to Chafik, borrowed by Berber from Latin, and subsequently by Moroccan Arabic from Berber. We simply do not have enough data to verify this claim. Heath (2002, cited in Rosenhouse and Goral 2004), for instance, contends that Arabs conquered ex-Roman garrisons, such as Volubilis in Morocco, and took Roman women as their wives or concubines. Their children apparently used a simplified Arabic dialect mixed with Late Latin substratum, which resulted in the first version of Western (Maghrebi) Arabic dialects. It is very difficult to maintain such an analysis knowing as little as we do about the proportion of these marriages, and who spoke Latin or Berber at the time. Here are two examples of the Latin nouns: *afital* 'bedroom' < Latin *hospitale* 'guest room'; *tayda* 'type of pine' probably < Latin *pinus taeda* 'loblolly pine'.

4.4 *Verbs*

Among the widely used verbs of Berber origin in Moroccan Arabic are the following: *ḥaf/iḥuf* 'to descend'; *saṭ/isuṭ* 'to blow air'; *šaš/išuš* 'to search'; *sas/isus* 'to leave out'; *ḍəhšər* 'to confuse, to overwhelm, to cause dizziness'; *bərnš/ibərnəš* 'to diversify crops'; *šəḥḥer/išəḥḥer* 'to leave a pot of tea on the fire to brew'.

These and other verbs are fully integrated in the grammar of Moroccan Arabic. For example, in the imperfect, *y-* is prefixed to the verb in the 3rd person, as in *ka-y-suṭ* 'he blows/is blowing', whereas in Berber it is *(la)i-tt-suḍ*. Some verbs are productive, like *bərnš*. The noun *l-bərniša* 'land with diverse crops' is derived from this verb. The equivalent in Berber is *taberništ* (Chafik 1999:66). For the verb *šəḥḥer* the corresponding verbal noun is *ttəšḥar*.

4.5 *Calquing*

In Berber the reflexive equivalent for 'itself/himself' is *ixf nnes* (lit. 'his/its head'). This expression may be claimed to have been modeled on the example of Berber. The equivalent of the Classical Arabic *fī ʿayni l-makān* 'in the same place' in Moroccan Arabic is *f dik leblasa b-ras-ha* which is a construction patterned on that found in Berber *deg wemkan nni s ixf nnes*, although in 'higher' forms of Moroccan Arabic one also finds *b-ddat* instead of *b-ras-ha*.

5. MORPHOSYNTAX

The diminutive in Berber is always feminine, whereas in Classical Arabic it keeps the gender of the source word. In Moroccan Arabic, diminutives can have either gender, with a high tendency to use feminine diminutives even when the source word is masculine. This tendency is seen by Chafik (1999) as evidence of borrowing from Berber. The Arabic word *bāb* 'door', for instance, has the masculine diminutive *buwayb* in Classical Arabic, whereas in Moroccan Arabic both *bwiba* [fem.] and *bwiyyeb* [masc.] exist.

The vocative in Moroccan Arabic is usually definite as in *a rražel* 'hey man!', *a lmra* 'hey woman!', whereas in Classical Arabic it is indefinite as in *yā rajul*, *yā mraʾa* 'hey man', 'hey woman!'

It should be noted that Moroccan Arabic has kept the distinction definite/indefinite on the noun as far as form is concerned, but in use the Berber pattern is followed, as is shown in the following example: (Moroccan Arabic) *waḥed r-rajel dxel*; (Berber, Tashelhit) *yan n urgaz ikčem*; (Classical Arabic) *daxala rajul* 'a man entered'.

The prefix *tt-* is used for passive derivation in Moroccan Arabic and Berber, whereas Classical Arabic makes use of specific patterns, like *fuʿila*:

Moroccan Arabic	Classical Arabic	Berber	gloss
kal/tt-kal	*ʾakala/ʾukila*	*eč/i-tt-č*	'to eat'
ḍreb/tt-ḍreb	*ḍaraba/ḍuriba*	*wt/i-tt-wt*	'to hit'

The imperfect form of the verb in Moroccan Arabic is preceded by a prefix *ka-* or *ta-*, and in Berber by *tt-*, as in (1), or a doubling of one of the root consonants as in (2). In Classical Arabic the imperfect has the prefix *y-* in the 3rd pers. masc.

	Moroccan Arabic	Berber	Classical Arabic	gloss
(1)	*ka-yšuf (šuf)*	*i-tt-wala (wala)*	*ya-nḏ̣uru (naḏ̣ara)*	'to look'
	ta-yxaṣem (xaṣem)	*i-tt-menġa (mnġ)*	*yu-xāṣimu (xaṣama)*	'to quarrel'
(2)	*ka-yaxud* (*xud*)	*i-kssi (ksi)*	*ya-ʾxuḏu (ʾaxaḏa)*	'to take'
	ta-yfhem (fhem)	*i-fehhem (fhem)*	*ya-fhamu (fahima)*	'to understand'

The assumption that this is borrowed from Berber is not warranted. Other Arabic dialects use aspectual prefixes, e.g. in Damascus Arabic *b-yəktob* 'he is writing/he writes'. As pointed out earlier, Moroccan Arabic could have developed this system independently from Berber, which is reason enough to speak more of probabilities than categories in borrowing.

The dative construction Subject-Verb-Object 1-to-Object2 can also be dative/accusative in Classical Arabic and Berber. The use of the preposition *l-* in dative constructions in Moroccan Arabic of the type *belleġ l-u lexbar* brought to-him the news 'he gave him the news' is according to Chafik (1999) an instance of Berber influence. The equivalent expression in Classical and Standard Arabic is *ballaġa-hu l-xabara*. The object clitic pronoun *-hu* in Classical Arabic corresponds to *-h* in Moroccan Arabic in accusative constructions, but in dative constructions Moroccan Arabic, unlike Classical Arabic, which uses the same pronoun, makes use of the preposition *l-* 'to', which aligns well with Berber.

6. Borrowings in written texts

Although not very pervasive, this type of borrowing is very interesting. Written texts are usually more conservative than oral conversations, as far as borrowing is concerned. Yet, in his dissertation on Tafilalt, Mezzine (1987) quotes a few examples illustrating the use of Berber within Arabic texts, as in *ya-dfaruna* 'they follow' < Berber *dfer* 'to follow' and *tata* < Berber *taḍa* 'alliance', etc., leading him to conclude that any investigator of these manuscripts needs to know Berber next to Arabic, the language of the manuscripts. This practice of merging the two languages in texts is certainly due to the fact that the writers were of Berber origin and inadvertently employed words from their native language.

7. Summary

The long-term contact between Moroccan Arabic and Berber has left major marks on the structure of both languages. The concern of this entry was to address the issue of borrowing by Moroccan Arabic from Berber. A brief review was given of the difficulties involved in undertaking such an enterprise, such as the lack of written documentation, language variation in Moroccan Arabic, the genetic relationship between the two languages, and finally the issue of independent language change, or internal language change and external language change (how to distinguish between change triggered by Berber – external – and 'normal' change, to which all living languages are subjected).

The most common cases of borrowing from Berber, as listed above, are nouns and verbs. Cases of morphological borrowing are illustrated by examples from diminutives, passive formation, imperfect prefixation, and the use of the preposition *l* in dative constructions.

Bibliographical references

Aguadé, Jordi. 1994. "La formation du réflechi-passif dans le dialecte arabe de Skūra (Maroc)". *Actes des premières Journées internationales de la dialectologie arabe de Paris*, ed. Dominique Caubet and Martine Vanhoeve, 71–78. Paris: Publications Langues'O.

——. 2004. "Zum Verbmodifikator *lā-* in nordmarokkanischen Dialekten". *Approaches to Arabic dialects: A collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday*, ed. Martine Haak, Rudolf de Jong, and Kees Versteegh, 13–19. Leiden: E.J. Brill.

Boogert, Nico van den. 1997. *The Berber literary tradition of the Sous with an edition of* "The ocean of tears" *by Muḥammad Awzal (d. 1749)*. Leiden: Nederlands Instituut voor het Nabije Oosten.

——. 1998. *La révélation des énigmes: Lexiques arabo–berbères des XVII^e et XVIII^e siècles*. Aix-en-Provence, CNRS-Universités d'Aix-Marseille I–III: Institut de Recherches et d'Etudes sur le Monde Arabe et Musulman.

Boukous, Ahmed. 1997. "Situation sociolinguistique de l'Amazighe". *International Journal of the Sociology of Language* 123.41–60.

Brunot, Louis. 1921. "Notes critiques sur l'article de Georges S. Colin (1920)". *Héspéris* 1:13. 350–352.

Chafik, Mohamed. 1999. *ad-Dārija al-maġribiyya majāl tawārud bayna l-ʾamaziġiyya wa-l-ʿarabiyya*. Rabat: ʾAkadīmiyyat al-Mamlaka al-Maġribiyya.

Chaker, Salem. 1989. *Encyclopédie berbère*, VI. Aix-en-Provence: EDISUD.

Chtatou, Mohamed. 1997. "The influence of Berber on Moroccan Arabic". *International Journal of the Sociology of Language* 123.101–118.

Cohen, Marcel. 1924. *Le système verbal sémitique et l'expression du temps*. Paris: Imprimerie Nationale.

Cohen, David. 1963. *Le dialecte arabe ḥassānīya de Mauritanie (parler de la Gəbla)*. Paris: Klincksieck.

Crystal, David. 1980. *A first dictionary of linguistics and phonetics*. London: Deutsch.

Dell, François and Mohamed El Medlaoui. 2002. *Syllables in Tashlhiyt Berber and in Moroccan Arabic*. Dordrecht: Kluwer Academic Publishers.

Field, Frederic, W. 2002. *Linguistic borrowing in bilingual contexts*. Amsterdam and Philadelphia: J. Benjamins.

Heath, Jeffrey. 2002. *Jewish and Muslim dialects of Moroccan Arabic*. London: Routledge Curzon.
Mezzine, Larbi. 1987. *Le Tafilalt: Contribution à l'histoire du Maroc au 17e et 18e siècles*. Rabat: Faculté des lettres, Mohamed V University.
Muysken, Pieter. 2000. *Bilingual speech: A typology of code-mixing*. Cambridge: Cambridge University Press.
Rosenhouse, Judith and Mira Goral. 2004. "Bilingualism in the Middle East and North Africa: A focus on the Arabic-speaking world". *The handbook of bilingualism*, ed. Tej K. Bhatia and William C. Ritchie, 835–868. Oxford: Blackwell.
Sieny, Mahmoud Esma'il. 1978. *The syntax of urban Hijazi Arabic (Saudi Arabia)*. London: Longman.
Taifi, Miloud. 1997. "Le lexique berbère: Entre l'emprunt massif et la néologie sauvage". *International Journal of the Sociology of Language* 123.61–80.
Thomason, Sarah G. 2001. *Language contact: An introduction*. Edinburgh: Edinburgh University Press.
—— and Terrence Kaufman. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
Versteegh, Kees. 1984. *Pidginization and creolization: The case of Arabic*. Amsterdam: J. Benjamins.
Youssi, Abderrahim. 1992. *Grammaire et lexique de l'arabe marocain moderne*. Casablanca: Wallada.

ABDERRAHMAN EL AISSATI (Tilburg University)

Bʿēri Arabic

1. GENERAL

Bʿēri Arabic is a variety of Upper Egyptian Arabic spoken on the West Bank at Luxor between Gurnat Marʿi in the north and Armanṭ in the south in a region named ilBiʿeṛāt (see Ramzī 1963:164). The speakers make a living in agriculture and their number is unknown. They claim Bedouin descent from the historically well known Fazāṛa tribe (see Murray 1935:293f.) and call themselves *afzaṛ* or *nās fuzuṛ*. Their ancestors are said to have arrived in Upper Egypt from the west following the *baʿaṛ* 'droppings' of their camels, hence the region's name ilBiʿeṛāt.

In Behnstedt and Woidich (1985), Bʿēri Arabic is classified as a separate group UE 3, because of the strong admixture of Bedouin elements in its phonology and morphology. Apart from this, it is close to the main Upper Egyptian dialect UE 1. A short account of Bʿēri is Woidich (1974). Texts are available in Fischer and Jastrow (1980) and Behnstedt and Woidich (1988). Audio material can be downloaded from the Semitic Language Archive, <http://www.semarch.uni-hd.de/index>.

2. LINGUISTIC DESCRIPTION

2.1 *Phonology*

2.1.1 Inventory

2.1.1.1 Consonants

Like all other Upper Egyptian dialects Bʿēri Arabic contains 28 phonemes. (Table 1).

As to phonetics, there are three remarkable facts: /ṭ/ is a pharyngealized and glottalized [ᵵˀ], i.e. the glottis is closed and released simultaneously with the articulation of the [ᵵ]. This holds for Upper Egypt to the south of Asyūṭ and for the Awlād ʿAli at the Mediterranean littoral. /j/ is a prepalatal affricate [ɟ] ~ [dʲ] as in the rest of Upper Egypt. Pharyngeals /ḥ/ and /ʿ/ have a relatively strong articulation.

Table 1. Inventory of consonants

	bilabial	labiodental	alveolar	postalveolar	palatal	velar	pharyngeal	laryngeal
plosive								
voiceless, voiced			t, d		j	k, g		ʾ
emphatic	b	ṭ, ḍ						
nasal	m		n					
fricatives								
voiceless, voiced		f	s, z	š		x, ġ	ḥ, ʿ	h
emphatic			ṣ, ẓ					
trill			r					
emphatic			ṛ					
lateral			l					
emphatic			ḷ					
glides	w				y			

Despite Bedouin admixture, interdentals have been replaced by the corresponding plosives. The glottal stop /ʾ/ has been preserved in several items like *iʾāti* ‘is coming closer to’, *iymaʾmaʾ* ‘is bloating’. **q* corresponds to /g/ as in *galb* < **qalb* ‘heart’, **g* to /j/, as in *jamal* < **gamal* ‘camel’. There are several cases of a development **q* > /k/, as in *kadd* ‘as much as’ (**qadr*), and conditioned through devoicing via the imperfect *katal/yiktil* ‘to give a beating’, *kaṭaʿ/yikṭaʿ* ‘to cut’. The root *ġdr* reflects **qdr*, as in *miš ġādra tākil laḥama* ‘she is not able to eat meat’.

2.1.1.2 Vowels

Table 2. Inventory of vowels

short: i u

a

long: ī ū

ē ō

ā

Oppositions /u/ vs. /i/:

tumm ‘two-piaster-coin’ – *timm* ‘make complete!’
šugga ‘scarf’ – *šigga* ‘direction’

/ē/ and /ō/ are the results of monophthongization of /ay/ and /aw/ respectively, cf. *gēḏ* < **qayḏ* ‘summer’, *ḥōš* < **ḥawš* ‘courtyard’.

Phonetic realization: /ā/ is much less affected by the automatic *ʾimāla* typically heard in northern parts of Egypt: *bāb* [baːb], not [bæ̝ːb] as in Cairo.

2.1.1.3 Diphthongs

The two diphthongs *aw* and *ay* are combinations of two phonemes, i.e. /a/ + a glide: /aw/ and /ay/. Intraverbally these occur only in front of a homorganic glide as in *ḍayy* ‘light, *baww* ‘dummy for a calf’, or when preserved by morphological patterns as in *mawlūd* ‘born’, *xayyaṭ* ‘to sow’. Word-finally, *-ay* and *-aw* appear as inflectional suffixes as in *ʿirfaw* ‘they knew’, *tijjuwwuzay* ‘you [fem.] marry’ which are monophthongized when suffixed: *tijjuwwuzēni* ‘would you marry me?’, *ma-yigdaṛōš* ‘they are unable’. For phonetic diphthongs in pause see 2.1.2.6.

2.1.1.4 Syllable

Possible syllables are Cv, Cv̄, CvC, Cv̄C, Cv̄CC, CCv, CCv̄, CCvC, and CCvCC. Final -CvC and -Cv̄C are treated alike and, in contrast to Dakhla, Sudanese, and Meccan Arabic, can both receive consonant-initial suffixes as in *kaṛabna* ‘our yoke’ and *bābna* ‘our door’ (cf. *bābinā* in West Dakhla), with the long vowel remaining long before the cluster. CvC_1C_2 is common word-internally, if C_1C_2 is a geminate or C_1 a liquid or a nasal (see 2.1.1.5): *ʿaddli* ‘tidy up [fem.]!’, *kalbha* ‘her dog’, *yurbṭu* ‘they bind’, *yinzlu* ‘they go down’, *binthum* ‘their daughter’.

2.1.1.5 Vowel elision and consonant clusters

Unstressed /i/ and /u/ in non-final open syllables undergo elision: i > ø / V(C)C_CV, cf. *ʿifiš+a* > *ʿifša* ‘bad [fem.]’, *baxxri bētuk* ‘burn incense in your house!’, *yixbiz+an* > *yixibzan* ‘they bake bread [pl. fem.]’. As the latter example shows, a resulting cluster -(C)CCC- is resolved by inserting /i/ (or /u/ in labial or pharyngealized environments) preceding the second consonant counted from the right: ø > i / . . . $V(C)C\_C_1C_2V$, cf. *nájiʿkum* ‘your village’, *nidifnū(h)* ‘we bury him’, *yúḍurba* ‘he hits him’. The cluster may remain unresolved when C_1 is a liquid or a nasal (see 2.1.1.6): *yinzlu* ‘they go down’, *galbha* ‘her heart’, *silsle* ‘backbone’, *yimsku* ‘they seize’.

/i/ and /u/ resulting from umlaut of **a* remain stable: *buguṛa* ‘cow’, *misikat* ‘she seized’, *da kitīr* ‘this is much’ (cf. Cairo *da ktīr*).

/a/ is not elided after Cv̄, see *gābalaw* ‘they met’, *mēdana* ‘minaret’ (cf. *midna* in Northern Middle Egypt).

2.1.1.6 Word-initial clusters C1C2-

Word-initial clusters are allowed when C_1 is a plosive and C_2 is a liquid, a nasal, or a glide: *bnitta* ‘girls’, *klibba* ‘dogs’, *brīg* ‘teapot’, *dyāba* ‘wolves’.

2.1.1.7 Word-final clusters -C1C2

Word-final clusters $-C_1C_2$ are resolved by vowel insertion, provided there is no morpheme juncture between C_1 and C_2. A vowel is inserted when C_2 is a liquid or a nasal: *šaṭir* ‘teat’, *ʿijil* ‘young bull’, *widin* ‘ear’, *ʿaḍim* ‘bone’. If, at the same time, C_1 is a liquid, a cluster $-C_1C_2$ remains unaffected: *garn* ‘horn’. In contrast, no vowel is inserted, if C_1 is a liquid or a nasal, except when C_2 is also a nasal or liquid: *birj* ‘pigeon tower’, *šilg* ‘thread made of palm fiber’, *ʾinf* ‘nose’, *milḥ* ‘salt’, *wirs* 'heritage', *gumʿ* 'funnel', but *tumun* 'eighth', *gamil* 'lice'. If neither liquids nor nasals are involved, insertion occurs between two voiced stops as in *kidib* 'lie', *ʿagid* 'necklace', and between pharyngeals followed by stops or sibilants as in *tiḥit* 'below', *kiʿib* 'heel', *jiḥiš* 'donkey'. Such insertion does not take place between fricatives or sibilants and stops: *mušṭ* 'comb', *lift* 'tulips'. On the other hand, the cluster is preserved if C_1 is a stop and C_2 is a pharyngeal or a sibilant: *rubʿ* 'quarter' (but *najiʿ* 'village'), *lugḥ* 'pollen', *xabz* 'baking', *baṭš* 'male buffalo calf', as well as between a voiced stop and a voiceless stop: *xabṭ* 'stroke'. The inserted vowel is /u/ with CuCC as in *zuʿur* 'tailless [pl.]', *suxun* 'hot'. It does not lead to morphological restructuring and remains unstressed: *ísimhin* 'their [fem.] name'.

2.1.1.8 Stress: general rules

Word stress falls on the vowel in the sequence -vCC- or -v̄C- closest to the end of the word: *xadámt* 'I worked', *mídrisa* 'school', *imgā́bala* 'encounter, *bā́ṭinha* (< *_baṭnha_ 'her belly'). Stress will not precede the infix *-t-* of Form VIII: *yištíġil* 'he works', except with the verbs III*y*: *yíštiki* 'he complains', nor will it fall on the article. If there is no such sequence, stress advances to the antepenultima as in *káṛab* 'yoke', *búṣula* 'onion'.

As stress assignment precedes vowel elisions and insertions, its position remains stable: *bṓnasa* > *bṓnasti* 'my tomato-bed', *búguṛti* 'my cow, *ríjilha* 'her foot'. See 2.1.3 for suffixed forms *kítibta* 'he wrote it'.

2.1.2 Phonotactics

2.1.2.1 Assimilations: /l/ and /n/ assimilate to the following consonant, in particular to /t/ and to liquids: *gutt* < *_gult_ 'I said', *kutt* < *kunt* 'I was', *bitt* < *bint* 'girl', *irrūḥ* < *inrūḥ* 'I go', *waḳḳanna* < *waḳḳalna* 'he fed us', *tiʿmilinna* < *tiʿmililna* 'you [masc.] do for us'.

2.1.2.2 Sun letters: apart from assimilating to the usual set of dentals and sibilants, /l/ assimilates to /j/ and optionally to /k/: *ijjurn* 'the threshing ground', *ikkīlu* 'the kilo', *ilkurkum* 'the turmeric'.

2.1.2.3 *gahawa* syndrome: Bʿēri displays an explicit → *gahawa* syndrome through which /a/ is inserted after laryngeal, pharyngeal, and post-velar fricatives L = /h, ḥ, ʿ, x, ġ/ when these are preceded by /a/ and followed by any consonant, provided that there is no morpheme juncture in between: ø > a / . . . aL_C. . . . Examples are: *sahal* 'easy', *laḥam* 'meat', *maʿarūf* 'well-known', *axaḍar* 'green', *baġala* 'mule'. This leads to a morphological restructuring of these nouns and subsequent stress adjustment: *ḍáhaṛ* *+ha* > *ḍahā́ṛha* 'her back'.

2.1.2.4 Umlautung: the most conspicuous phonological rule of Bʿēri affecting both verbal and nominal inflection (see the paradigms) is the Umlautung, i.e. /a/ in sequences CaCaC and CaCCaC is replaced by /i/ (or /u/ in emphatic or labial environment) when vowel-initial suffixes are added.

a > i / C_C(C)_Cv
masak+at > *misikat* 'she seized'
našrab+aw > *nušrubaw* 'they drink'
ʿuṛugi 'brandy made of dates'

sanat+ēn > *sinitēn* 'two years
maṛat+i > *muṛuti* 'my wife'
simika 'a fish'

/a/ may be preserved with a following back consonant as in *bilaḥa* 'date'. Exceptions are the numerals *ʾaṛbaʿa* 'four', *ʿašaṛa* 'ten'.

Vowel elision precedes Umlautung and the vowels produced by this rule are not elided (see 2.1.1.5). Umlautung, on the other hand, precedes the *gahawa* rule and stands in a counter-feeding relation to it: *yuxluṣaw* 'they end', but *yaxalaṣ* 'it ends'. This is why there are surface exceptions to Umlautung created by the later application of the *gahawa* rule: *šaḥamāya* 'a piece of fat' (< *_šaḥm+āya_), *naxala* 'date-palm tree' (< *_naxl+a_), but cf. *dihibāya* 'a piece of gold' (< *_dahab+āya_). For more details see Woidich (1973–1974).

2.1.2.5 Pausal forms

Bʿēri, like most rural Egyptian dialects, shows a variety of vowel changes in pause.

-ā́ > -íh : *ištā > ištíh* 'winter',
iġríh 'glue'.
-a > -e ~ -eh ~ -ih : *yākleh* 'he eats it',
milike 'queen',
deh ~ dih 'this'

Final *-i* and *-u* in pause develop phonetic diphthongs as in *yimšu* ['jimʃu^{w}] ‘they go’, *timši* ['timʃi^{y}] ‘you go’; and in closed long syllables: *tilifōn* [tili'foˑ^{a}n] ‘telephone’; *ṇufuṛēn* [n̴u̴f̴u̴reˑ^{a}n] ‘two men’.

2.1.2.6 Pharyngealization tends to spread throughout a word and beyond it: *faṛkạ̄hạ* [f̴ɑr̴k̴ɑːhɑ] ‘twirling it’, *ilṣāḥba* [ɪls̴ɑːħb̴ɑ] ‘to his friend’, *ʿaṣāyạ* [ʕɑs̴ɑːyɑ] ‘stick’.

Depharyngealization takes place in the imperfect of Form II with the /i/ in the ultima: *xaḍḍar*, *iyxaddir* ‘to grow [plants]’, *ṣaffar*, *iysaffir* ‘to whistle’.

2.1.2.7 Labialization of /a/ after *-ū-* is very common: *ʿa-yjībūha* [ʕayɟibuːhɒ] ‘they add it’.

2.1.3 Morphophonology

The construct state of the feminine noun is *-at*: *ṛugubatha* ‘her neck’, *išwayyat ṃayya* ‘a little water’. The active participle lengthens its fem. *-a*: *ṛamyāha* ‘having thrown [fem.] it’.

Long vowels are shortened when stress is removed, but remain long before -CC-: *dābiḥ* ‘having slaughtered’, fem. *dābḥa*, pl. *dabḥīn*. Unstressed /ē/ in open syllables is replaced by /a/: *zagēna+ha* > *zaganāha* ‘we irrigated it’.

The feminine suffix *-a* becomes *-at* with following genitive. With vowel-initial suffixes the *-a-* is elided: *buḥḥat+ēn* > *buḥḥtēn* ‘two ducks’, *ʿammat+i* > *ʿammti* ‘my aunt’; but not with /t/ or /d/: *jaddati* ‘my granny’, *jittati* ‘my body’. Final *-īya* > *-īt-*: *ribʿīya* – *ribʿītha* ‘her kid goat’, final *-ya* > *īt*: *ṭūrya* – *ṭurīti* ‘my hoe’; *sāgya* – *sagīt istanyōs* ‘the water wheel of Istanyōs’.

Suffixation of verbal forms:

3 sg. fem. perf.+ suffix
nuḍurat+a > *núḍurta* ‘she saw him’
libsat+a > *líbista* ‘she put on’
biddilat+a > *bíddilta* ‘she replaced it’
gābalat+ak > *gā́baltak* ‘she met you’
xadat+a > *xídita* ‘she took it’
lammat+a > *límmita* ‘she took it up’
šāfat+a > *šā́fta* ‘she saw it’
ramat+a > *rúmuta* ‘she threw it’

3 pl. f. perf.
nuḍuran+a > *nuḍuránna* ‘they [fem.] saw him’
gābalan+ak > *gabalánnak* ‘they [fem.] met you’

3 pl. masc. perf.
nuḍuraw+ha > *nuḍurōha* ‘they [masc.] saw her’

2 sg. fem.
isʾilay+ha > *isʾilēha* ‘ask [fem.] her!’

2.2 *Morphology*

Bʿēri Arabic makes a gender distinction in the 2nd and 3rd plural forms (see 2.2.1). The active participle receives *-āt* in this case: *saknāt wēn il ʿawajīz dēl* ‘where do these old women live?’.

2.2.1 Pronouns

2.2.1.1 Personal independent pronouns

Table 3. Personal independent pronouns

hū ~ *hūwa*	*inta*	*ana* ~ *āna*
hī ~ *hīya*	*inti*	
huṃṃa	*intu*	*iḥna*
hinna	*intan*	

2.2.1.2 Possessive/object suffixes (Table 4)

-h, *-y*, *-ki* appear after vowels: *abūh*, *abūy*, *abūki*, *yištirūh* ‘they buy it’, with negation *mayištirūhiš* ‘they do not buy it’. The object suffix for the 1st pers. sg. is *-ni*.

Table 4. Possessive/object suffixes

-a ~ *-h*	*-ak*	*-i* ~ *y* ~ *(-ni)*
-ha	*-uk* ~ *-ki*	
-hum	*-kum*	*-na*
-hin	*-kan*	
bēta	*bētak*	*bēti*
bētha	*bētuk*	
bēthum	*bētkum*	*bētna*
bēthin	*bētkan*	

2.2.1.3 Indirect object suffixes

Table 5. Indirect object suffixes

-la	*-lak*	*-li*
-ilha	*-luk*	
-ilhum	*-ilkum*	*-ilna [inna]*
-ilhin	*-ilkan*	
jābla	*jāblak*	*jābli*
jābílha	*jābluk*	
jābílhum	*jābílkum*	*jābílna*
jābílhin	*jābílkan*	

Without *-i-* after vowels: *galōlna* [ga'loːnːa] ‘they told us’, etc. The indirect object suffixes may be added to direct object suffix: *injibālak* ‘I bring it to you’, *tifiršuhāla* ‘you [pl.] spread it for him’, *inbiʿhinlak* [imbœ'ħːelːak] ‘I sell them to you’, but this is not obligatory: *ʿayyṭūh liyya* ‘call him for me!’, *itjība līk* “she brings it to you’.

2.2.1.4 Demonstratives

Table 6. Demonstratives

‘these’	‘those’
da ~ dih	*dukkạ̄ti*
di ~ diy	*dikkīti*
dōl ~ dōla	*dukkuṃṃa*
dēl ~ dēla	*dikkinna*

There is a third deixis *awwēnhūti* (sg. masc.), *awwēnhīti* (sg. fem.) ‘over there’ in reference to something at some distance.

2.2.1.5 Presentatives

Table 7. Presentatives

áha	*ahūwa*	*ahīk*	*ahīni*
	ahīya	*ahīki*	
	ahuṃṃa	*ahīkum*	*ahīna*
	ahinna	*ahīkan*	

Example: *ahuṃṃa nnās ištaṛaw minnīh ilbāb* ‘look, the people bought the door from him’. *áha* is invariant and refers to a general fact: *áha ğğamūs miš kulla wāḥid* ‘as a matter of course, the buffaloes are not all the same’.

2.2.1.6 Relative pronoun

illi *innās illi ʿindīha šabb* ‘the people who have a bull’.

2.2.1.7 Interrogative pronouns

‘who?’	*mīn*
‘what?’	*ē(h)*
‘which?’	*innhi* (invariable and preposed): *innhi waṛad* ‘which boy?’, *innhi nās* ‘which people?’. But with concord when following the noun: *innhū* (sg. masc.), *innhī* sg. fem.), *innhuṃṃa* (pl. masc.), *innhinna* (pl. fem.): *ilwaṛad innhū* ‘which boy?’, *fi lḥitta nnhī* ‘in which place?’.

2.2.2 Adverbs

temporal	*dilgē ~ dilgēti* ‘now’, *innhāṛda* ‘today’, *bukṛa* ‘tomorrow’, *imbāriḥ* ‘yesterday’, *lissaʿ* ‘still; not yet’, *inhạ̄rītha* ‘then’
local	*hēna ~ ihnā* [ɪh'neh] ~ [ɪh'nɪʔ] ‘here’, *ihnạ̄k ~ ihnạkkạ̄ ~ ihnukkạ̄ti* ‘there’, *minna* ‘this way’, *minnạ̄k* ‘that way’
manner	*ikdá* [ɪg'deh] ~ *ikdēti* ‘so’; *gūwi* ‘very’, *wāṣil* ‘totally’
interrogatives	*wēn* ‘where?’, *mēta* ‘when?’, *lē(h)* ‘why?’, *kē(f)* ‘how?’, *kaddē* ‘how much?’, *kām* ‘how many?’

2.2.3 Particles

ihnīn sg. masc., *ihnīt* pl. fem., *ihniyyīn* pl. masc., *ihniyyāt* pl. fem. serve as genitive markers as in *ijjamāʿa hniyyāta* ‘his womenfolk’, but the pan-Egyptian *ibtāʿ* is very much in use, sometimes showing the plural *ibtaʿīn*: *innaṣṣāra lli btaʿīnna* ‘our Christians’.

The common negational particles *ma- . . . š*, *ma-. . .*, and *miš* are used. *ma- . . . š* (*ma- . . . iš* after vCC and v̄C) negates not only verbal forms, but nominals and participles as well: *xašimha ma-ḥaṣilš iššṭūr* ‘her mouth does not reach the teats’, *ma-xabrīniš* ‘they do not know’, *ma-zēniš* ‘it is not good’. In emphatic contexts, it is replaced by *ma-*, which may receive stress in this case: *ʾabadan ma ʿamalt ikdih* ‘I have never done such things!’, *ʿumur ma tḥaddit ḥaddīta zēne* ‘you never tell a nice story!’, *tubt tāni mā́-nrūḥ maṣir* ‘never again shall I go to Cairo!’. *lissaʿ* ‘not yet’ still has its /ʿ/.

Questions may be reinforced by *walla* ‘or’: *kal jiḥšak walla* ‘did your donkey eat or (not)?’.

Wishes may be introduced by *rēt-*: *rētak ma jīt* ‘I wish you hadn’t come!’, *rēta twaffa* ‘I wish he had died!’.

Prepositions not commonly used in Egypt include *ṭūl* ‘besides’: *giʿmizat ṭūl issawwāg* ‘she sat down beside the driver’; *kē* ‘like’: *kitīra gūwi kē lġanam* ‘as many as the sheep’ (but *kēfak* ‘like you’); *xašim* ‘in front of’: *xašm iddikkān* ‘in front of the shop’; *bakān* ‘instead of’: *injīb bakānha ḥajāt jidīda* ‘we bring something new in its place’.

Conjunctions: *lōla* ‘otherwise’: *lāzim yiḍirbūh lōla ma-ygūm* ‘they have to beat it, otherwise it will not stand up’; *ʾadām* ‘since, as [causal]’; *yann* ‘because’: *yanni gasyān mi lmišwāṛ* ‘because I am exhausted by the walk’; *lakan ~*

yakan ‘but’; *ḥaggāš* ‘as long as not, if not’: *ḥaggāš iššamš taṭlaʿ lōla ma-nasraḥ* ‘as long as the sun does not rise, I shall not go to work in the fields’; *la-* ‘lest’: *xušš jawwa la-tizzakk* ‘go inside, lest you catch a cold’; *la . . . la . . .* ‘either . . . or . . .’: *illi ʿāšaṛ ilgōm irbʿīn yōm la ṣār la ṭār* ‘whoever stayed with people for 40 days, will either be gone or will stay for ever’; *in* ‘that’: *nādir in šuftak* ‘I rarely saw you’; *ibnēn ma* ‘everywhere’. *illi* is used as a complementizer ‘that’: *zēn illi . . .* ‘it is good that . . .’.

The vocative particle *ya* may be followed by a definite noun: *ya lʿāṛ* ‘what shame!’, *ya jjráss* ‘what an embarrassment!’. Common exclamations are: *yá-wṛa* ‘oh boy!’, *ya-bá* ‘oh girl!’, *ya-ḅāy* ‘oh father!, *ya-xayy* ‘oh brother!’, *ya-xayyti* ‘oh sister!’, *ya-ḥalāha ~ ya-ḥalayyha* ‘how lovely she is!’, *ya-rawagatha* ‘how nice she is!’.

2.2.4 Noun

Gender: feminine nouns without the marker *-a* are the usual ones such as the body parts *rijil* ‘foot’, *ʾīd ~ yadd* ‘hand’, *ʿēn* ‘eye’, *baṭin* ‘belly’, *ṭīẓ* ‘buttocks’, etc. and *ʾarḍ* ‘soil’, *nāṛ* ‘fire’, *maṛkab* ‘boat’, *balad* ‘village’, *šamš* ‘sun’, and the less common ones *šnāb* ‘moustache’, *sūg* ‘market’, *sikkīn* ‘knife’, *bīr* ‘well’, *gammāri* ‘moon’.

Article: *il-* no assimilation to /g/, optionally to /k/: *ilkalb* ‘the dog’, *ilgarn* ‘the horn’; assimilation with /j/: *ijjurj* ‘the drawer’

Specificity: ‘a certain’ may be expressed by *wāḥid*: *ʿind wāḥid šēx* ‘with a (certain) sheikh’; *wāḥid ṣāḥbak* ‘a friend of yours’.

Besides the usual plural patterns, there are some unusual ones:

- CiCCiyy (sg. CaCāya)
  - *irḥiyy* ‘hand mills’, *iʿṣiyy* ‘sticks’
- CCaCC (sg. CvCCa)
  - *ibsass* ‘cats’, *irkaḅḅ* ‘knees’, *šnaṭṭ* ‘bags’
- CiCCēC for domestic animals:
  - *biṭṭēš* ‘buffaloes’, *šibbēb* ‘bulls’, *jiḥḥēš* ‘donkeys’
- CCāCa (sg. CaCaC)
  - *iġlāga* ‘locks’, *igfāṣa* ‘crates’
- CCūCa for animals:
  - *imhūra* ‘colts’, *sbūʿa* ‘lions’, *idkūṛa* ‘males’
- CCiCCa
  - *bnitta* ‘girls’, *klibba* ‘dogs’, *brigga* ‘jugs’, *iʿyilla* ‘children’
- CiCān ~ CuCān (sg. Cv̄C)
  - *bibān* ‘doors’, *fisān* ‘axes’, *ḥušān* ‘courtyards’, *ʿudān* ‘shadoofs’

Besides the usual → pseudo-dual forms *ʾidēn* ‘hands’, *rijlēn ~ rajalēn* ‘feet’, *ʿenēn* ‘eyes’, there are other body parts following this pattern: *daraʿēn* ‘arms’, *ṣabaʿēn* ‘fingers’, *janaḥēn* ‘wings’, cf. *xamas ṣabaʿēn* ‘five fingers’, *ṛajalēn iḥmāṛ* ‘legs of a donkey’. As usual, final *-n* is dropped with suffixes: *ṣabaʿáyy* ‘my fingers’, *rijlēha* ‘her legs’.

The diminutive is quite productive, mostly with CCēC and CCēCa for 3-radicals: *bṭēš* ‘buffalo calf’, *brēš* ‘mat’, *jḥēš* ‘donkey’, *jdayy* ‘kid’, *dnēša* ‘piece’, *ksēwa* ‘garment’. It may be formed from adjectives: *ḥlēw* ‘nice’, *jdayyid* ‘new’ and with *-a* in the case of feminine nouns: *sūg* from *swēga* ‘teashop in the market’. With words containing a long vowel in the ultima CCayyiC is used: *brayyig* ‘jug’, *rġayyif* ‘loaf of bread’. Particular forms are: *axx* > *xayy* ‘brother’, *uxt* > *xayya* ‘sister’, *bitt* > *bnayya* ‘girl’, *maṛa* > *mṛayya* ‘women’. Some adjectives take an infix *-ēṭa-* or *-aṭṭū-*: *gṣar ~ gṣaṭṭūr* ‘short’, *zġēṭar ~ zġaṭṭūr* ‘small’, *glēṭal ~ glaṭṭūl* ‘little’.

For colors and deficiencies there is the usual pattern aCCaC: *abyaḍ* ‘white’, *aswad* ‘black’, *axaḍar* ‘green’, *aʿaraj* ‘lame’, *azʿar* ‘tailless’; fem. and pl. as elsewhere: *akḥal*, *kaḥala* fem., *kuḥul* pl. ‘deep black’.

The elative pattern is aCCaC: *azyan* ‘better’, *aġalaḍ* ‘thicker’, with *ajdad ~ ajadd* ‘newer’; *aglal ~ agall* ‘less’ for II gem.

2.2.5 Numerals

‘one’ *wāḥid*, fem. *wiḥda ~ waḥda*; ‘two’ *itnēn*, but also *jōz* may be used as in *jōz ifṛāx* ‘two chickens’, *jōz ilfṛāx* ‘the two chickens’.

Numerals 3–10 are like Standard Egyptian, with long and short forms: *talat igrūš* ‘three piasters’, *xamas fidin* ‘five feddans’, *taman isnīn* ‘eight years’, and with reanalysis of feminine suffix *-at* in connection with *aCCāC and aCCuC plurals: *xamas t-iyyām* ‘five days’, *xamas t-ušhur* ‘five months’.

For 11–19 only long forms exist ending in *-ir*: 11 *iḥdāšir* without pharyngealization, *iṭnāšir*, *xamaṣṭāšir*, etc., in some villages with the /ʿ/ preserved as in *tamanṭaʿšir*. For 20–90 as in Cairo, with the exception of *irbiʿīn* ‘forty’.

For 100 *mīya*, construct state *mīt maṛṛa* 'a hundred times', *mitēn*, *tultmīya*, *urbuʿmīya*, *xumsmīya*, etc.; 1,000 *ʾalf*.

2.2.6 Verb

2.2.6.1 Verbal forms

2.2.6.1.1 Form I

Form I displays in the perfect an *a*-type *masak* 'he seized', *rama* 'he threw', and an *i*-type *íšrib* 'he drank', *ílgi* 'he found'.

2.2.6.1.2 The derived forms

The system of derived forms follows the Eastern type, i.e. in contrast to Cairo it displays a morphological distribution of /a/, /i/ in the ultima of Forms II and III, with /a/ in the perfect and /i/ in the imperfect:

II *ṣabbax*, *iyṣabbix* 'to fertilize' III *ʿārak*, *iyʿārik* 'to fight'

Form II replaces IX: *wišša ḥammaṛ* 'his face became red'. Besides its usual semantics, Form II stresses the plurality of subjects and objects, cf. *ilʿanza wildat* 'the goat gave birth to a kid' against *ilʿanzāt wullidan* 'the goats gave birth to kids'. The reflexive passive forms are exclusively formed with an *it*-prefix:

t-I Form	*itrakan*, *yitrikin*	'to lie down'
t-II Form (V)	*itwakkal*, *yitwakkal*	'to set out'
t-III Form (VI)	*ittāwab*, *yittāwab*	'to yawn'

The *ista*-Form (X) is quite productive in the sense of 'to find/consider something . . .': *istaḥsan*, *yistaḥsan* 'to find good', *istabṛad*, *yistabṛad* 'to find sth. cold', *istawʿaṛ*, *yistawʿaṛ* 'to find difficult'. Combinations of *ista*- with Forms II or III occur: *istaṛayyaḥ*, *yistaṛayyaḥ* 'to take a rest', *istabāṛak*, *yistabāṛak* 'to receive a blessing'.

Occasionally, an *in*-Form (VII) and a *-t*-Form (VIII) occur, but these are fixed to certain lexical items and are not productive: *indabaḥ*, *yindibiḥ* 'to be slaughtered'. *-t-* (VIII) *ištaġal*, *yištiġil* 'to work'.

2.2.6.2 Inflection of imperfect and perfect

The imperfect paradigm, which follows the Maghrebinian inflection, shows the prefixes *ya*-, *yi*-, *yu*- distributed according to vowel harmony, i.e. the vowels of the prefix and the imperfect base are alike. *yu*-, however, may replace *yi*- before /w/: *yuwzin* 'to weigh out'. The suffixes, too, follow this harmony with *-i* (2nd sg. fem.), *-u* (3rd pl. masc.) added to bases with high vowels and *-ay*, *-aw* to bases containing /a/ or /ā/ in the last syllable. This spread of *-ay*, *-aw* from III*y* verbs to other verbal classes, again, can be considered a Bedouin feature.

The perfect paradigm is less harmonic, as it shows *-at* (3rd pers. sg. fem.), *-aw* (3rd. pers. pl. masc.), *-an* (3rd pers. pl. fem.) suffixes, all containing /a/, throughout the paradigm irrespective of the quality of the vowel of the base: *libsat* 'she put on' (< *ilbis+at*).

2.2.6.2.1 Imperfect: paradigm

Table 8. Imperfect

yálbas 'to put on'		
yálbas	*tálbas*	*nálbas*
tálbas	*tílbisay*	
yílbisaw	*tílbisaw*	*nílbisaw*
yílbisan	*tílbisan*	
yáġalaḍ 'to become thick'		
yáġalaḍ	*táġalaḍ*	*náġalaḍ*
táġalaḍ	*túġluḍay*	
yúġluḍaw	*túġluḍaw*	*núġluḍaw*
yúġluḍan	*túġluḍan*	
yúrgud 'to lie down'		
yúrgud	*túrgud*	*núrgud*
túrgud	*túrugdi*	
yúrugdu	*túrugdu*	*núrugdu*
yúrugdan	*túrugdan*	
yímsik 'to take'		
yímsik	*tímsik*	*nímsik*
tímsik	*tímiski*	
yímisku	*tímisku*	*nímisku*
yímiskan	*tímiskan*	
yádbaḥ 'to slaughter'		
yádbaḥ	*tádbaḥ*	*nádbaḥ*
tádbaḥ	*tídbaḥay*	
yídbaḥaw	*tídbaḥaw*	*nídbaḥaw*
yídbaḥan	*tídbaḥan*	

Aspectual, temporal and modal prefixes

ʿa- ~ *ʿama*- for present tense: *ʿangullak* 'I tell you', and habitual: *ʿatxaddir min nafsha* 'it grows by itself'.

ḥa- and *ṛāḥ* ~ *ṛaḥa* for the future *miš ḥa-yxallan fīha wala dinša* ‘they will not leave a single piece in it’, *ṛaḥa-yjību* ‘they are going to bring’.

tam- as an intensifier: *tam-tunfux tam-tunfux tam-tunfux, tunfux lamma ʾē, yitnifix maẓbūṭ* ‘you go on blowing, and blowing, and blowing, until it what? it is properly inflated’.

xal- for finality: *hatinna jozēn ganadīl xannišwūhin* ‘bring us some corn cobs so that we can grill them!’.

2.2.6.2.2 Perfect: paradigm (Table 9)

For suffixation see 2.1.3.

Table 9. Perfect

xabaz ‘to bake bread’		
xabaz	*xabazt*	*xabazt*
xibizat	*xabazti*	
xibizaw	*xabaztu*	*xabazna*
xibizan	*xabaztan*	
íšrib ‘to drink’		
íšrib	*išribt*	*išribt*
širbat	*išribti*	
širbaw	*išribtu*	*išribna*
širban	*išribtan*	

2.2.6.3 Participles

The same types exist as in Cairo Arabic for Form I, i.e. CāCiC for the active and maCCūC for the passive participle. In contrast to Cairo Arabic, active participles are formed with /i/ in Form II, passive participles with /a/: *imraggiʿ* ‘having mended’ – *imraggaʿ* ‘mended’ as in *muṛuti mraggʿa jallabīti* ‘my wife has mended my jallabiyya’ vs. *jallabīti mruggaʿa* ‘my jallabiyya has been mended’, or *imgallam* ‘clipped’, *imtallat* ‘tripled’. The passive participle of verbs III*y* in Form II is mCaCCāy as in *imrawwāy* ‘irrigated’, *imġaṭṭāy* ‘covered’. *jā*, *ijī* ‘to come’ forms the active participle *jāy*, fem. *jāya*. Feminine *-a* is lengthened with suffixes: *mistannyāni* ‘she is waiting for me’.

2.2.6.4 Verbal noun patterns

Verbal nouns follow the usual patterns except for the II and t-II Forms. Like elsewhere in Upper Egypt between Sohāg and Edfu (see Behnstedt and Woidich 1985, maps 317–321), the unusual pattern CiCCīC is used: *ṣirrīx* ‘shouting’, *ʿiggīd* ‘tying up’; III*y* verbs: *biggi* ‘condoling’. For verbs with four radicals, correspondingly, *ġirbīl* ‘sieving’.

2.2.7 Weak verbs

2.2.7.1 The perfect of the II gem. verbs shows the normal suffixes originating from the verbs III*y* such as *-ēt* and so on: *lammēt* ‘I took’, *lammat* ‘she took’, *lammaw* ‘they took’. The active participle follows the strong verb: *ḥāṭiṭ* ‘having put down’.

2.2.7.2 Iʾ verbs

Table 10. Inflection of the verb *kal*

kal ‘to eat’					
kal	*kalt*	*kalt*	*yākil*	*tākil*	*nākil*
kalat	*kalti*		*tākil*	*tākli*	
kalaw	*kaltu*	*kalna*	*yāklu*	*tāklu*	*nāklu*
kalan	*kaltan*		*yāklan*	*tāklan*	

The imperative with /u/ deviates from the imperfect with /i/: *kul*, *iklíyy*, *iklúww*, *iklánn*. Active Participle is *wākil*.

2.2.7.3 Verbs I*w* and I*y*

Table 11. Inflection of I*w*/*y* verbs

a-type		
wazan ‘to weigh’		
wazan	*wazant*	*wazant*
wuzinat	*wazanti*	
wuzinaw	*wazantu*	*wazanna*
wuzinan	*wazantan*	
yuwzin	*tuwzin*	*nuwzin*
tuwzin	*túwizni*	
yúwiznu	*túwiznu*	*núwiznu*
yúwiznan	*túwiznan*	
i-type		
íwṣil ‘to arrive’		
íwṣil	*iwṣilt*	*iwṣilt*
wiṣlat	*iwṣilti*	
wiṣlaw	*iwṣiltu*	*iwṣilna*
wiṣlan	*iwṣiltan*	
yōṣal	*tōṣal*	*nōṣal*
tōṣal	*tōṣalay*	
yōṣalaw	*tōṣalaw*	*nōṣalaw*
yōṣalan	*tōṣalan*	
I*y*		
íybis ‘to dry’		
íybis	*iybist*	*iybist*
yibsat	*iybisti*	
yibsaw	*iybistu*	*iybisna*
yibsan	*iybistan*	

Table 11 *(cont.)*

yēbas	*tēbas*	*nēbas*
tēbas	*tēbasay*	
yēbasaw	*tēbasaw*	*nēbasaw*
yēbasan	*tēbasan*	

2.2.7.4 Verbs II*w*/*y*

These are as expected (Table 12).

Table 12. Inflection of II*w*/*y* verbs

gām 'to stand up'		
gām	*gumt*	*gumt*
gāmat	*gumti*	
gāmaw	*gumtu*	
gāman	*gumtan*	*gumna*
iygūm	*itgūm*	*ingūm*
itgūm	*itgūmi*	
iygūmu	*itgūmu*	*ingūmu*
iygūman	*itgūman*	

Imperfects with /ā/ get *-ay*/*-aw* suffixes: *itxāfay*, *itxāfaw* 'you are afraid'. As elsewhere in Upper Egypt, the shortened vowel in derived Forms is /i/, not /a/ as in *irtiḥt* 'I took a rest' *(irtāḥ)* or *inḍimt* 'I got tired' *(inḍām)* (cf. Cairo *irtaḥt*).

2.2.7.5 Verbs III*y*

There is an *a*-type and an *i*-type (Table 13).

Table 13. Inflection of verbs III*y*

maša 'to go'		
maša	*mašēt*	*mašēt*
mašat	*mašēti*	
mašaw	*mašētu*	*mašēna*
mašan	*mašētan*	
yimši	*timši*	*nimši*
timši	*timši*	
yimšu	*timšu*	*nimšu*
yimšan	*timšan*	
ílgi 'to find'		
ílgi	*ilgīt*	*ilgīt*
ligyat	*ilgīti*	
ligyaw	*ilgītu*	*ilgīna*
yalga	*talga*	*nalga*
talga	*talgay*	
yalgaw	*talgaw*	*nalgaw*

2.2.7.6 Irregular verbs (Table 14)

Table 14. Inflection of irregular verbs

jā 'to come'		
jā	*jīt*	*jīt*
jāt	*jīti*	
jāw	*jītu*	*jīna*
jann	*jītan*	
iyjī	*itjī*	*injī*
itjī	*itjī*	
iyjū	*itjū*	*injū*
ijann	*itjann*	
idda 'to give'		
idda	*iddēt*	*iddēt*
iddat	*iddēti*	
iddaw	*iddētu*	*iddēna*
iddan	*iddētan*	
yiddi	*tiddi*	*niddi*
tiddi	*tiddi*	
yiddu	*tiddu*	*niddu*
yiddan	*tiddan*	

2.2.8 4-radicals CaCCaC and itCaCCaC Types

reduplication (diminutive)	*ʿasʿas* 'to grope about' *balbal* 'to wet' *dagdag* 'to smash to pieces' *faʿfaʿ* 'to sniff around' *mahmah* 'to mutter'
onomatopoeic verbs	*kaṛkaṛ* 'to gurgle' *dabdab* 'to knock' *raṭraṭ* 'to chatter'
inserted consonants	*marmaš* 'to bite' *naʿniš* 'to refresh' *daʿbas* 'to grope around' *farjaḥ* 'to spread apart',
derived from nouns	*sabras* 'to hasten' (< English *to express*) *garṭas* 'to wrap' (< **qurṭās*), *magraṣ* 'to strew with bran' (< *magraṣ* 'tray made of clay for baking bread')
reflexive passive	*itkanfil* 'to stumble', *ikkartaḥ* 'to roll down', *itġarbal* 'to be sieved'

The inflection follows Forms II and V, as does the verbal noun: *dibdīb* ‘knocking’, *ġirbīl* ‘sieving’.

2.3 *Remarks on syntax*

Plural nouns of animals and objects can agree with the 3rd pers. pl. fem.: *ilʿanzāt wullidan* ‘the goats gave birth’, *faddanēn ʾarḍ, yaʿni law zaraḥthin ḥašīš* ‘two feddans of land, if I were to sow them with grass’; *ilxibbēza wu lḥummēḍa wu zzabānix dēla kullhin* . . . ‘mallow, sorrel, and spinach, they are all . . .’.

The ‘ethical dative’ is common with verbs such as *ʿāwiz* ‘want’, *ilgi* ‘to find’, *jāb* ‘to bring’, *xad* ‘to take’: *ʿawizla giršēn* ‘he wants some money’, *nalgāli mīt jinēh* ‘I find a hundred pounds’.

To express intensification *dawwaṛ* + verbal noun is used: *duwwaṛat fīh katil* ‘she gave him a good thrashing’. *gaʿad* + imperfect/participle, as in *gaʿad iṛṛājil yaṛgaʿ fī* ‘the man kept on beating him’, expresses prolongation.

In narrative style the ‘narrative verbal noun’ may be used, as in *xašš fi lxēma wu dukkā waṛāh ʿiffīṣ fī ʿiffīṣ fī* ‘he entered the tent, and the other one behind him, keeping on kicking at him’, as well as a ‘narrative imperative’ as in *ǧaybīnlaha ʿirg ifhimt izzāy, dibb fīha, dibb fīha* ‘we took a stick, you understand, and then we keep hitting on it and on it!’. The use of the periphrastic narratives *ṛāḥ*, *jā* is very common too: *lamma ṛāḥat ilgabḍa tḥallat minnīh* ‘when the fastening came loose from him’, *māt ilʿayyil nuṣluxō, wu nǧū kabsīna tibin* ‘the young died, [now] we skin it and then we stuff it with straw’.

Conditional sentences are introduced by *in* or *law: law šuftak tāni ṛāḥ nagṭaʿ zētak* ‘if I see you again, I’ll kill you’, *in zuġt minni ʿārfak* ‘if you slip away from me, I’ll know you’.

The *wi* of syndetic *ḥāl*-sentences is inserted after the subject: *šufta hūwa w ʿayiḥrit* ‘I saw him ploughing’, *tākil minha nnās, hīya w xaḍra* ‘people eat from it when it is green’.

3. Lexicon

For ‘to see’ **raʾā* is common in the negative only: *ma-retāš* ‘I did not see him’, the most common words being *naḍar*, *yunḍur* and the pan-dialectal *šāf*, *yišūf*.

Bedouin words such as *zēn* ‘good’, *šēn* ‘bad’, *ʿadāṛa* ‘women’, *xābir* ‘knowing’, *gēḍ* ‘summer’ are much in use, as well as the common Upper Egyptian items, e.g. *rāyig* ‘fine, good’, *ʿifiš* ‘bad’, *wāṣil* ‘very’, *ʿād* ‘then, yet’, *dinša* ‘piece’, *ʿāza* ‘need’, *ʿayyaṭ ʿala w* ‘to call’, *iddalla* ‘to go down’, *fanas*, *yifnis* ‘to look out’. The latter three, combined in a phrase, are frequently quoted in order to show the incomprehensibility of the Ṣaʿīdi dialect to northeners: *ʿayyaṭt ʿalēh fanas ma-ddallāš* ‘I called at him, he looked out of the window, but did not come down’. Needless to say, the common pan-Egyptian lexical items are – as everywhere – gradually superseding the old expressions.

Bibliographical references

Behnstedt, Peter and Manfred Woidich. 1985. *Die ägyptisch-arabischen Dialekte*. II. *Dialektatlas von Ägypten*. Wiesbaden: L. Reichert.

——. 1988. *Die ägyptisch-arabischen Dialekte*. III. *Texte*. Part 2, *Niltal und Oasen*, 244–275. Wiesbaden: L. Reichert.

Murray, G.W. 1935. *Sons of Ishmael: A study of the Egyptian Bedouin*. London: Routledge.

Ramzī, Muḥammad. 1963. *al-Qāmūs al-juġrāfī li-l-bilād al-miṣriyya*. II/4. Cairo: Maṭbaʿa Dār al-Kutub al-Miṣriyya.

Woidich, Manfred. 1973–1974. “Die 3.sg.f. Perfekt im Dialekt von il-Biʿrāt”. *Mélanges de l’Université Saint Joseph* 48.355–372.

——. 1974. “Ein arabischer Bauerndialekt aus dem südlichen Oberägypten”. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 124.42–58.

——. 1980. “XIV. Text aus il-Biʿrāt”. *Handbuch der arabischen Dialekte*, ed. Wolfdietrich Fischer and Otto Jastrow, 235–242. Wiesbaden: O. Harrassowitz.

——. 1997. “Upper Egyptian Arabic and dialect mixing in historical perspective”. *Humanism, culture and language in the Near East: Studies in honor of Georg Krotkoff*, ed. Asma Afsaruddin and A.H. Mathias Zahniser, 185–197. Winona Lake, Ind.: Eisenbrauns.

Manfred Woidich (University of Amsterdam)

Bilingualism → Multilingualism; Child Bilingualism

Binā’

The term *bināʾ* is used in Arabic grammar to describe words which have fixed end-vowels. It may be generally translated as ‘indeclinability’, and is thus the antithesis of the term for ‘declinability’, → *ʾiʿrāb*. This meaning most probably developed from the sense of ‘phonetic structure’, which *bināʾ* (pl. *ʾabniya*) indicates in expressions like *bināʾ li-l-majhūl* 'passive form', *ʾabniyat al-jamʿ* 'plural patterns', etc.

The centrality of *bināʾ* to grammatical theory is best demonstrated by its discussion, along with *ʾiʿrāb*, by Sībawayhi (d. 180/796) in the very early parts of his *Kitāb* (I, 2ff.). He distinguishes between *ʾiʿrāb* and *bināʾ*, based on whether the *ʿāmil* 'operator' can cause a temporary change (*yazūlu ʿan-hu*) in the final vowel of a word, hence *ʾiʿrāb*, or cannot do so since some words have permanent final vowels (*lā yazūlu ʿan-hu*), which no *ʿāmil* can change, hence *bināʾ* (*Kitāb* I, 3). Sībawayhi enforces this distinction with his terminology, which differentiates between eight 'forms of word endings' in Arabic. These are *naṣb*, *jarr*, *rafʿ*, and *jazm* on the one hand, and *fatḥ*, *kasr*, *ḍamm*, and *waqf* on the other: *wa-ʾinnamā ḏakartu la-ka ṯamāniyata majārin li-ʾafruqa bayna mā yadxuluhu ḍarb min hāḏihi l-ʾarbaʿa li-mā yuḥdiṯu fī-hi l-ʿāmil . . . wa-bayna mā yubnā ʿalayhi l-ḥarf bināʾan* 'I have cited eight forms so as to be able to distinguish between those four forms that are affected by an operator and those [other forms] which are considered as part of the word's pattern' (*Kitāb* I, 3). This rigorous distinction, which was maintained in its entirety throughout the Arabic grammatical tradition, rightly represents, according to Versteegh (1993:128–129), a considerable innovation when compared with previous terms used by commentators, such as Muḥammad ibn as-Sāʾib al-Kalbī (d. 146/763), who did not observe a clear distinction between declensional endings and internal vowels. For the *ʾiʿrābī* and non-*ʾiʿrābī* vowel terminology, see Talmon (2003:238–244).

Words that are indeclinable are referred to as *mabnī*, and include nouns, verbs, and particles. These words, just like those words which are declinable or *muʿrab*, are closely linked with the three parts of speech, i.e. *ism* 'noun', *fiʿl* 'verb', and *ḥarf* 'particle'. As far as nouns are concerned, the grammarians agree that their → *ʾaṣl* stipulates that they be *muʿrab* (declinable) since they imply different meanings – such as those expressed by the subject, the object, and the construct – and hence different case endings had to express those different meanings (Ibn al-ʾAnbārī, *ʾAsrār* 24–25). Accordingly, the grammarians had to deal with those words which they classified as nouns but are *mabnī* nonetheless. These nouns, which are often classified as particles in Western grammars and which are usually described as *ġayr mutamakkin* as opposed to nouns which are *muʿrab* or *mutamakkin*, mainly include conditionals such as *man* 'who', interrogatives such as *ʾayna* 'where?', pronouns such as *huwa* 'he', demonstratives and relatives such as *hāʾulāʾi* 'those' and *al-laḏī* 'who', and adverbs such as *ḥayṯu* 'where' and *munḏu* 'since'. They also include, however, words which are incontestably nouns such as *ʾamsi* 'yesterday' and *sībawayhi* [proper noun].

The grammarians were typically concerned with finding justification for the indeclinability of those nouns which do not agree with the norm for their part of speech. The most widely cited justification is that most of these nouns resemble particles (*ḥurūf*) and are accordingly *mabnī*, as is the norm for all particles (Ibn al-ʾAnbārī, *ʾAsrār* 30–32; Ibn ʿAqīl, *Šarḥ* 32–34). A medial position between *mutamakkin* and *ġayr mutamakkin* was assigned to diptotes, or *mamnūʿ min aṣ-ṣarf*. These are usually referred to as *mutamakkin ġayr ʾamkan* in order to reflect the fact that they have neither a fixed end-vowel nor receive the whole range of end-vowels since their indefinite forms cannot have *kasra* or → *tanwīn*.

A special kind of *bināʾ* in nouns is described by the grammarians as *ʿāriḍ* 'transient'. Among the more common types are the following, all of which are *mabnī* in certain contexts but are otherwise *muʿrab*: (a) single-word nouns in the vocative (e.g. *yā rajulu/muslimūna* 'o man/ Muslims!'); (b) single-word nouns after generic *lā* (e.g. *lā rajula/muslimīna* 'there is no man/are no Muslims'); (c) compound words, including compound numerals (e.g. *xamsata ʿašara* 'fifteen') and compound adverbs (e.g. *layla nahāra* 'day and night'); (d) adverbs like *qablu* 'before' and *baʿdu* 'after' which are not followed by the genitive (i.e. as opposed to *min qabli hāḏā* 'before this' etc.); and (e) *ʾayy* 'whatever/ whoever' in constructions like *iḍrib ʾayyuhum ʾafḍalu* 'hit whomever is best'. For more details, see ad-Dāyil (1990:247–341) and ʿAllūš (1997:188–208).

Contrary to nouns, *bināʾ* is generally said to be the *ʾaṣl* for verbs and particles (Ibn al-ʾAnbārī, *ʾAsrār* 24; Suyūṭī, *Hamʿ* I, 15), although the Kufans reportedly believe that *ʾiʿrāb* is the *ʾaṣl* for verbs as well as for nouns (ʾUšmūnī, *Šarḥ* I, 24). Among the verbs, the perfect or → *māḍī* is said to be uniformly *mabnī* with a final *fatḥa*, *ḍamma*, or *sukūn* (e.g. *kataba*, *katabū*, *katabta* 'he, they, you wrote'), whereas the imperfect or *muḍāriʿ* is interpreted as *mabnī* only when it is followed by the *-na* suffix of the 2nd and 3rd person feminine plural, as in *taktubna* 'you write [2nd person fem. pl.]', *yaktubna* 'you write [3rd person fem. pl.]', or by the energetic *nūn* , as in *la-ʾaktubanna* 'I shall indeed write', but not in forms like *la-taktubunna* 'you shall indeed write [2nd person masc. pl.]', where the *nūn* is said to be not in direct contact with the verb. The grammarians seem here to have disregarded the syllabic and vowel changes triggered by the introduction of the suffix *-na* and energetic *nūn* and explained the verbal form within the general framework of *ʾiʿrāb* and *bināʾ*. As for the imperative (*ʾamr*), the various arguments attributed to the Basrans and Kufans in their theoretical difference as to whether it is *mabnī*, as the Basrans believe, or *muʿrab*, as the Kufans assert, are discussed extensively by Ibn al-ʾAnbārī in *masʾala* 72 of his book on the differences between the Basrans and the Kufans (*ʾInṣāf* I, 524–529; cf. *ʾAsrār* 317–321). All grammarians, however, agree that particles are *mabnī* and that the *sukūn* is the *ʾaṣl* in all *mabnī* words, be they nouns, verbs, or particles (Ibn Yaʿīš, *Šarḥ* III, 82–83; Suyūṭī, *Hamʿ* I, 20–21). For the relationship between *bināʾ* and *ʾiʿrāb*, on the one hand, and the grammarians' distinction between prepositions (*ḥurūf jarr*) and adverbs (*ḏurūf*), on the other, see Levin (1987:354–355).

The picture presented by the grammarians about *ʾiʿrāb* and *bināʾ*, it has been suggested (Baalbaki 1990:17–33), is a simplification of a much more complex dialectal situation which the grammarians recorded but attempted to marginalize for the sake of coherency. In the case of *bināʾ*, the sources report that *al-ʾasmāʾ as-sitta* 'the six nouns [of the type *ʾabūka*, *ḥamūka*, and *ḏū*]', which are normally triptotes, are treated as *mabnī* in some dialects, and that the dual and, to a lesser degree, the sound masculine plural have affinity with *bināʾ* in others. Conversely, some nouns which are *mabnī*, such as *ʾamsi* 'yesterday', *ḏū* '[relative pronoun of Ṭayyiʾ]', *ḥayṯu* 'where', *ladun* 'at, by', as well as proper nouns ending in *-wayhi*, receive various degrees of *ʾiʿrāb* in certain dialects (for details, see Baalbaki 1990:20–21).

BIBLIOGRAPHICAL REFERENCES

Primary sources

Ibn al-ʾAnbārī, *ʾAsrār* = ʾAbū l-Barakāt ʿAbd ar-Raḥmān ibn Muḥammad al-ʾAnbārī, *ʾAsrār al-ʿArabiyya*. Ed. Muḥammad Bahjat al-Bīṭār. Damascus: Maṭbaʿat at-Taraqqī, 1957.

——, *ʾInṣāf* = ʾAbū l-Barakāt ʿAbd ar-Raḥmān ibn Muḥammad al-ʾAnbārī, *al-ʾInṣāf fī masāʾil al-xilāf bayna n-naḥwiyyīn al-Baṣriyyīn wa-l-Kūfiyyīn*. Ed. Muḥammad Muḥyī d-Dīn ʿAbd al-Ḥamīd. 2 vols. Cairo: al-Maktaba at-Tijāriyya, 1955.

Ibn ʿAqīl, *Šarḥ* = Bahāʾ ad-Dīn ʿAbdallāh Ibn ʿAqīl, *Šarḥ Ibn ʿAqīl ʿalā ʾAlfiyyat Ibn Mālik*. Ed. Ramzī Munīr Baʿalbakī. Beirut: Dār al-ʿIlm li-l-Malāyīn, 1992.

Ibn Yaʿīš, *Šarḥ* = Muwaffaq ad-Dīn Yaʿīš ibn ʿAlī Ibn Yaʿīš, *Šarḥ al-Mufaṣṣal*. 10 vols. Cairo: al-Maṭbaʿa al-Munīriyya, n.d.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. 2 vols. Būlāq, 1316–17 A.H.

Suyūṭī, *Hamʿ* = Jalāl ad-Dīn ʾAbū l-Faḍl ʿAbd ar-Raḥmān ibn ʾAbī Bakr as-Suyūṭī, *Hamʿ al-hawāmiʿ šarḥ Jamʿ al-jawāmiʿ fī ʿilm al-ʿArabiyya*. 2 vols. Cairo: Maṭbaʿat as-Saʿāda, 1327 A.H.

ʾUšmūnī, *Šarḥ* = ʾAbū l-Ḥasan ʿAlī ibn Muḥammad al-ʾUšmūnī, *Šarḥ al-ʾUšmūnī ʿalā ʾAlfiyyat Ibn Mālik al-musammā Manhaj as-sālik ʾilā ʾAlfiyyat Ibn Mālik*. Ed. Muḥammad Muḥyī d-Dīn ʿAbd al-Ḥamīd. 3 vols. Cairo: Dār al-Kitāb al-ʿArabī, 1955.

Secondary sources

ʿAllūš, Jamīl. 1997. *al-ʾIʿrāb wa-l-bināʾ: Dirāsa fī naḏariyyat an-naḥw al-ʿarabī*. Beirut: al-Muʾassasa al-Jāmiʿiyya li-d-Dirāsāt wa-n-Našr wa-t-Tawzīʿ.

Baalbaki, Ramzi. 1990. "*ʾiʿrāb* and *bināʾ* from linguistic reality to grammatical theory". *Studies in the history of Arabic grammar*. II. *Proceedings of the 2nd symposium on the history of Arabic grammar, Nijmegen, 27 April–1 May 1987*, ed. Michael G. Carter and Kees Versteegh, 17–33. Amsterdam: J. Benjamins.

Dāyil, ʿAbdallāh ibn Ḥamad ibn ʿAbdallāh ad-. 1990. *al-Bināʾ fī l-luġa qasīm al-ʾiʿrāb*. Riyad: Maktabat ar-Rušd.

Levin, Aryeh. 1987. "The view of the Arab grammarians on the classification and syntactic function of prepositions". *Jerusalem Studies in Arabic and Islam* 10.342–367.

Talmon, Rafael. *Eighth century Iraqi grammar: A critical exploration of pre-Ḫalīlian Arabic linguistics*. Winona Lake, Ind.: Eisenbrauns.

Versteegh, Kees. 1993. *Arabic grammar and Qurʾānic exegesis in early Islam*. Leiden: E.J. Brill.

RAMZI BAALBAKI
(American University of Beirut)

Binding

Nominal expressions may be classified into three categories: anaphors such as reflexives (1a) and reciprocals (1b), pronominals (2), and referential expressions or (R-expressions) (3). The literature on Binding Theory is extensive; see Chomsky (1981), Chomsky (1995), and

Hornstein (2001) for more exhaustive discussions and references. All the examples are given in Lebanese Arabic:

(1) Anaphors
a. *šēf* *ḥāl-o*
saw.3ms state-him
'He saw himself'

b. *ḥiko* *ma* *baˤḍ-un*
spoke.3p with reciprocal.p
'They spoke with each other'

(2) Pronominals
sāmi *šēf-o*
Sami saw 3ms-him
'Sami saw him'

(3) R-expressions
šēf *l-walad*
saw.3ms the-boy
'He saw the boy'

These nominal expressions are distinguished with respect to the contexts in which they must or may have an antecedent. Thus, anaphors must have an antecedent and the antecedent cannot be 'too far' as illustrated in (4a–b) respectively:

(4) a. * *ḥāl-o* *šēf*
state-him saw.3ms
'Himself saw'
b. * *sāmi* *ʾāl* *ʾinn-a* *šēfit* *ḥāl-o*
Sami said that-her saw.3fs state-him
'Sami said that she saw himself'

A pronoun, on the other hand, cannot have an antecedent which is 'too close' as illustrated in (5 a–b):

(5) a. *sāmi* *šēf-o*
Sami saw.3ms – him
'Sami saw him'
b. *sāmi* *ʾāl* *ʾinn-a* *šēfit-o*
Sami said.3ms that-her saw.3fs-him
'Sami said that she saw him'

The pronoun in the direct object position can have *sāmi* as an antecedent in (5b) but not in (5a): it can only be co-referential with *sāmi* in (5b).

Finally, a R-expression cannot be co-referential with a preceding nominal expression when this expression is in a 'high enough' position, as illustrated in (6a–b):

(6) a. *ʾāl* *ʾinn-a* *šēfit* *sāmi*
said.3ms that-her saw.3fs Sami
'He said that she saw Sami'

b. *l-mˤallme* *yalli* *htammit* *fī*
the-teacher.f Rel. took care of.3fs him
'The teacher that took care of him'

ʾālit *ʾinn-a* *šēfit* *sāmi*
said.3fs that-her saw.3fs Sami
'She said that she saw Sami'

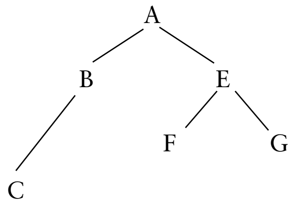
In (6a), the R-expression *sāmi* cannot be co-referential with the non-overt subject pronoun in the matrix clause. It can, however, be co-referential with the pronoun in the relativized clause in (6b).

The Binding Theory deals with the contexts in which a nominal expression must (anaphors), may or may not (pronouns and R-expressions) have an antecedent. It states the following (this formulation is simplified; see Chomsky 1981, 1995, for an extensive discussion of the Binding Theory).

(7) a. An anaphor must have a c-commanding antecedent in the smallest clause or nominal phrase containing this anaphor.
b. A pronoun cannot have a c-commanding an antecedent in the smallest clause or nominal phrase containing this pronoun.
c. An R-expression cannot have a c-commanding antecedent.

C-command (or constituent-command) can be defined as follows. In a tree representation like (8):

(8)



All the nodes are branching except B. A category A c-commands a category B if and only if the first branching category dominating A also dominates B and A and B do not dominate each other. In (8), for instance, C and B c-command E, F, and G and F c-commands B and C.

The terms 'bind' and 'free' can be defined as follows. A nominal expression is bound when it has a c-commanding antecedent; otherwise it is free. The (binding) principles in (7) can be reformulated as (9):

(9) Binding Principles:
a. An anaphor must be bound in the smallest clause or nominal phrase containing it.
b. A pronoun must be free in the smallest clause or nominal phrase containing it.
c. An R-expression must be free.

There is extensive literature on binding in Generative Grammar and space limitation prevents a discussion here of the various issues involving binding. Some of these are dealt with in the references mentioned. Recently, however, the status of the binding principles as independent principles has been questioned. It is argued that at least the binding principles regulating anaphors and pronouns are not primitive principles: they can be derived from other grammatical principles such as the theory of movement (see Chomsky 1995; Hornstein 2001).

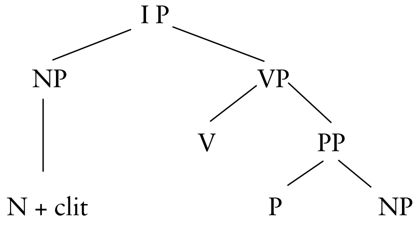
Arabic adds interesting dimensions with respect to binding. To mention some, Mohammad (2000) points out that there is a contrast between (6b) and (10) in Levantine Arabic:

(10) *ʾimm-o htammit bi -sāmi*
mother-his took care of.3fs of -Sami
'His mother took care of Sami'

In (10), *sāmi* cannot be co-referential with the adnominal pronoun. This is surprising since adnominal complements are contained within a nominal phrase. As such, the pronoun in (10) does not c-command *sāmi* and co-reference should be possible.

Mohammad indicates that the adnominal pronoun in (10) is a clitic incorporated into the head noun. This being the case, the representation of the subject nominal phrase is non-branching as in (11 F-2) (non-relevant details omitted):

(11)



In (11) the clitic c-commands the prepositional object and co-reference is prohibited.

Another interesting dimension brought forth by Arabic is the behavior of resumptive pronouns. Resumptive pronouns are used productively in various Arabic dialects. The following examples are in Lebanese Arabic (see Ouhalla 2004; Choueiri 2002).

However, there is dialectal variation that is of interest. As originally pointed out by Eid and Shlonsky (1999), resumptive pronouns are not allowed in Cairene Arabic or in Palestinian Arabic in context (12 a) but are allowed in context (12 b–c). This prohibition, which as Choueiri (2002) indicates does not exist in Lebanese Arabic, is referred to as the Highest Subject Constraint (see McCloskey 1990):

(12) a. *l-walad yalli huwwe rāḥ*
the boy Rel. he left.3ms
'the boy that left'
b. *l-walad yalli šēfit-o*
the-boy Rel. saw.3fs-him
'the boy that she saw'
c. *l-walad yalli fakkarto huwwe rāḥ*
the-boy Rel. thought.2p he left.3ms
'the boy that you thought he left'

The Highest Subject Constraint has a binding flavor: in a relativized nominal phrase, the highest pronominal subject cannot be bound by, must be free from, the relativized nominal element. However, Shlonsky (1992) offers an interesting account for this Highest Subject Constraint in terms of movement. The account is based on the assumption that resumptive pronouns are last resort and occur when movement is not available. The Highest-Subject position in Cairene Arabic or Palestinian Arabic is available for short movement (or A-movement); hence, resumption does not occur. In Shlonsky's account, the Highest Subject Constraint is not an independent principle in the grammar (for other obviation constraints affecting strong pronouns and strong pronouns with epithets → resumption).

In brief, co-referential relations between nominal elements are not free. They are regulated by the binding principles which encode the contexts in which a nominal element must, may, or may not have an antecedent. There are cross-linguistic variations affecting the binding principles. Some that arise in certain varieties of Arabic are discussed here. These variations lend support to the proposal according to which binding principles are not primitive principles, but rather descriptive generalizations. The challenge becomes to derive these generalizations from other means. The debate in the linguistic literature is centered around which theories are most appropriate in deriving the binding generalizations.

Bibliographical references

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

——. 1995. *The minimalist program*. Cambridge, Mass.: MIT Press.

Choueiri, Lina. 2002. *Re-visiting relatives: Issues in the syntax of resumptive restrictive relatives*. Ph.D. diss., University of Southern California.

Hornstein, Norbert. 2001. *Move! A minimalist theory of construal*. Cambridge, Mass.: Blackwell.

McCloskey, James. 1990. "Resumptive pronouns, a-binding, and levels of representation in Irish". *Syntax and semantics of the modern Celtic languages*, ed. Randall Hendrick, 199–248. San Diego Cal.: Academic Press.

Mohammad, Mohammad. 2000. *Word order, agreement and pronominalization in standard and Palestinian Arabic*. Amsterdam and Philadelphia: J. Benjamins.

Shlonsky, Ur. 1992. "Resumptive pronouns as a last resort". *Linguistic Inquiry* 23.443–468.

Sells, Peter. 1984. *Syntax and semantics of resumptive pronouns*. Ph.D. diss., University of Massachusetts, Amherst.

Joseph Aoun
(University of Southern California)

Biradicalism

A minority of Arabic roots are biconsonantal, while the great majority, as in other Semitic languages, consist of three consonants called 'radical consonants' or simply 'radicals' (here symbolized as 'R') and one or two root vowels. Since vowels are subject to morpho-phonemic changes, it is difficult to say precisely which ones are root vowels. Traditionally, Arabic roots, like roots of other Semitic languages, have been analyzed as consisting of consonants only (→ root; → derivation), viz., two, three, or four consonants. Most of the roots consisting of four consonants (relatively numerous but rarely used since only 15 of them occur in the *Qurʾān* against 1,160 triconsonantal roots) are due to partial reduplication and/or dissimilation of some consonants of originally triconsonantal roots or are cases of lexicalization of an affix (usually a prefix) added to triconsonantal roots. Yet, there are also such roots formed by reduplication of original biconsonantal roots, e.g. *ġalġala/wa-ġala* 'to penetrate', *xaḍxaḍa/xaḍḍa/ma-xaḍa* (also *xaṭara* < **xaḍara* < **xaḍr* < *xaḍḍ*) 'to shake' (see Atallah and Ayache 1981).

There is a limited number of nominal roots (37, according to Fleisch 1990:I, 248, 252–254, based on Nöldeke 1910:109–178) consisting of only two consonants. They belong to the basic vocabulary going back to Proto-Semitic and even Proto-Hamito-Semitic/Afro-Asiatic, e.g. *yad* 'hand', *fa-m/fa-mm/fu* 'mouth', *ḥir* 'vulva', *dam* 'blood', *māʾ* 'water', *ibn* 'son', *ism* 'name' (the last two nouns with prothetic *i-*), *lis-ān* 'tongue' (with lexicalized suffix), etc. As far as verbal roots are concerned, medieval Arab lexicographers and grammarians noted that many of these roots have the same or very similar meaning, although they differ only in one consonant. Lists of such roots were collected (e.g. by Ibn as-Sikkīt, *ʾIbdāl* and ʾAbū ṭ-Ṭayyib al-Luġawī, *ʾIbdāl*; → *ʾibdāl*). This discovery was taken over by Hebrew grammarians and via grammars of Biblical Hebrew it became known among European Semitists since the same phenomenon can be observed in other Semitic languages. There is no doubt that much root variation goes back to Proto-Semitic (see Zaborski 1971) and even Proto-Hamito-Semitic (see several studies on Egyptian roots by Belova), but some developments go back only to Proto-Arabic, or can be attributed only to relatively recent dialect differentiation involving, for instance, alternation of /ḏ/ and /d/, /ṯ/ and /t/, etc., as in *ḏāqa/dāqa* 'to taste'. Some of the roots having variants may be considered as a basis for the reconstruction of original biradicals. First of all, there are geminated roots, viz. roots with identical second and third radical (R1VR2R2), having variants with one 'weak' consonant, viz. /w/, /y/, /ʾ/ (→ glide). A few roots have both 'weak' R1 and R3; in this case only one consonant is really 'weak'; for instance, *w-* is 'strong' in *ḥawila/yaḥwalu* 'to be cross-eyed', *ʿawiza/yaʿwazu* 'to be needy, to be poor'. There are also variants in which either the first consonant (sometimes this could shift to the second position due to a metathesis with the second radical) or the third one may be interpreted as an original but later lexicalized ('petrified') affix (or an infix via metathesis). These roots should reasonably be reconstructed as original biradicals, e.g. *kāna* 'to be', *sa-kana* 'to dwell, to live', *kankana* 'to nest, to take a good seat', *kanna* 'to hide, to shelter'; *sa-ruʿa/ha-riʿa* 'to hasten'; *na-ġara/wa-ġara* 'to be enraged'; *ʿatuda* 'to be ready', *ʿadda* 'to make ready'; *zabā/zāba-na* 'to push'. The historically attested affixes are : *t-* (also infixed *-t-*

due to metathesis and then sometimes also assimilated so that we have *hawiša/hāša/hašša* and *hutiša* ‘to be excited’) , *n-*, *m-*, *š-/s-/h-/ʾ-*, *w-*, *y-*, *-t*, *-n*, *-w*, *-y*, *-ʾ* (see Kuryłowicz 1972:6–31; Zaborski 1971). There are several reasons why usually no specific meaning can be attributed to these lexicalized affixes. One of them is that they were used to derive deverbal nouns from original biradical roots, so that triconsonantal roots were further derived from these deverbal nouns (which often disappeared) rather than directly from biconsonantal verbs. Concerning roots with /w/, /y/, /ʾ/ (roots with /ʾ/ sometimes going back to /y/, although the change *-ʾ-* > *-y-* is also attested), without other, viz. geminated or affixed, variants, and geminated roots without /w/, /y/ /ʾ/ variants, it is more reasonable to assume that already in Proto-Semitic if not Proto-Hamito-Semitic, there were some original roots of this kind (e.g. *xadā/xāda/waxada* ‘to go quickly’?), which served as a model for the extension of original biradicals to triradicals (but cf. Chekayri 1995, 1998; Chekayri and Scheer 2003; Voigt 1988). Since in closed syllables (before consonantal and zero endings) long vowels were shortened (e.g. imperfect *ya-qūm-u*, but jussive *ya-qum* from *qāma* ‘to get up’), such roots with *-w-* or *-y-* (in many verbs either *-w-* or *-y-*, see Attalah and Ayache 1972) have been interpreted as biconsonantal, and therefore other biconsonantal roots could be reinterpreted as having variants with *-ū-* < **-w-* or with *-ī-* < **-y-*. As demonstrated by Kuryłowicz (1972: 9–10), the mechanism of the enlargement of R1VR2 > R1VR2R2 was fundamentally the same as the lengthening of the root vowel of other biconsonantal roots. This may be seen in variants R1VR2R2 and R1VR2/*w*/*y*, e.g. *ḥamma/ḥamiya/ḥamā* ‘to be hot’. However, there is no reason to assume that every weak and geminated verbal root was originally biradical, since R1VR2R2 could also go back to R1VR2R3 through dissimilation or through assimilation of the final -R2 to the following *-t*, *-k*, or *-n*. Besides, in forms R1R2VR2 the first -R2- could partially assimilate to R1 or dissimilate with it, resulting in a third consonant in medial position. But the majority of geminated roots were indeed biconsonantal. The only safe hypothesis is that the number of triconsonantal roots in the prehistoric period was smaller and the number of biradicals was larger than in the historic period, but it does not make sense to pretend that we could reconstruct all or even the majority of triconsonantal verbal roots as originally biradical.

There are, however, many cases in which originally triconsonantal roots were differentiated (see Zaborski 1991 for detailed discussion and for an up-to-date bibliography on the problem of biradicalism) in an ‘internal’ phonological process or, rather rarely, through contamination of semantically and/or phonologically similar original triconsonantal roots, e.g. *šābaha* and *šākala* resulting in *šakaha* ‘to be similar’. In such cases, biconsonantal roots cannot be reconstructed on the basis of triconsonantal variants. Assimilation and dissimilation usually worked in direct contact position (although both could also work at a distance), viz. in forms like R1R2VR3, e.g. *ya-sqab-u* > *ya-ṣqab-u*, where both coexisting verbs *saqaba* and *ṣaqiba/ṣāqaba* mean ‘to be near’, *ya-bḥaṯ-u* > *ya-fḥaṣ-u*, where both verbs mean ‘to examine’ (*faḥaṣa* also ‘to examine’), or in forms like R1VR2R3, e.g. *ḥars/ḥarz* ‘watch, guard, control’, *ḥaraz-tu/-ku* > *ḥaras-tu/-ku* ‘I guarded’; before *-kV* of the ending of the Yemenite *k-* perfect, e.g. *battat-ku* > *battak-ku* > *bataka*, but elsewhere *batat-tu* ‘I cut off’; *ġamma/ġamaḍa/ġamada/ġamā* ‘to hide, to cover’, where the last variant is a secondary biradical due to a reinterpretation of *ġamma*. There must also have been cases of spirantization, e.g. *bataʾa/baṯaʾa* ‘to stop’ (see Corriente 1969), as well as metathesis, e.g. *ḥašama/šaḥama* ‘to be fat’. Alternation produced, for instance, *ḥabala/ʿabala* ‘to bind with cords’, which can be also interpreted as due to voicing in *ya-ḥbul-u* > *ya-ʿbul-u*. As a matter of fact, some forms can have two or perhaps even three explanations, and such ambiguity can only contribute to their reinterpretation, e.g. *baʿaṯa/baʿata* can be explained as due either to *baʿaṯ-tu* > *baʿat-tu*, or to spirantization of *baʿata* > *baʿaṯa*, cf. *ʾa-bʿada/ʿa-bʿaṭa* ‘to send away’, which is connected with a different root, viz. *baʿuda* ‘to be far away’. In some triconsonantal nominal roots, like *širš/šilš* ‘root’, *baḍr/baḏr* ‘clitoris’, variants are also due to phonological processes.

Doublets, triplets, etc. are important for the prehistory of Arabic and Semitic phonology: *jabba/qabba* ‘to cut off’, *jaraḥa/qaraḥa* ‘to wound’, *šajja/šaqqa* ‘to split’, *šajā/šaqā* ‘to distress s.o’ (also *šajaba* ‘to grieve s.o.’, where *-b-* may be either originally a preposition or a variant of *-w-*), *lazija/laziqa* ‘to stick’ (but also *lazza/lazaba/lazama* ‘to adhere, to cling’), *zalaja/zalaga/zalaqa/zalaʾa* ‘to slide, to glide, to slip’ go back to different dialectal variants illustrating the history of the pronunciation of /j/ and /q/ as [g] and of /q/ merging with /ʾ/. *Laṭaṣa/laṭaša/laṭaṯa/laṯaṭa/latada/lataza/laṭaḥa/laṭaʿa/lataḥa/lataġa/lataxa/laṭama/latama/laṯama* ‘to hit’ plus *latta* ‘to pound, to hit’ are remnants of old phonological processes. Examples like *šaxxa/ ḍaxxa* ‘to urinate’, *hašama/haḍama* ‘to break’, *waššaḥa/waḍḍaḥa* ‘to comment, to make clear’, *jahaša/jahaḍa* ‘to flee’ are important for the prehistory of sibilants and emphatics. For the history of the latter see, e.g., *ṭafara/ḍafara* ‘to leap’, *ta-qaṭṭaʿa/ta-qaḍḍaʿa* ‘to be cut’ (also *qaṭaba/qaḍaba* ‘to cut’), *qabaṭa/ qabaḍa* ‘to seize with hand, to grab’, *šaṣṣa/šaḏḏa* ‘to be burdensome, to be hard, painful’, *waqaša /waqaḏa* ‘to strike hard’.

Zemánek (1996) has published an important collection and analysis of doublets with the alternation of voiceless emphatic and its non-emphatic counterparts (both voiced and non-voiced), which shows that some triconsonantal doublets may be due to the shift from glottalized to pharyngealized pronunciation of emphatic consonants. Some of his examples can, however, be interpreted as due rather to ‘deemphatization’ and/or different kinds of assimilation or dissimilation (see Kuryłowicz 1972:28–31).

There are unclear cases, like *ta-baʿʿaṣa*, *ta-baʿṣaṣa* ‘to move [intr.]’ and *raʿaṣa* ‘to move [trans.]’ (cf. Zemánek 1996:78), *ṯadaqa/ġadiqa* ‘to be copious, to pour down [rain]’ (Zemánek 1996:79), *habaza/hariza* ‘to die’ (Zemánek 1966:79), which should not be rejected *a priori* and which require further investigation.

When two triconsonantal roots differ only in one consonant which can be neither identified with an attested affix nor explained as being due to specified phonological and/or morphophonological processes (e.g. *ṣaluba/ṣaliba* and *ṣalada* ‘to be hard, firm, solid’ (although neither assimilation in *ṣalib-ta* > *ṣalid-ta*, nor dissimilation in *ṣalad-ta* > *ṣalib-ta* is impossible), some linguists regard such an enigmatic consonant as a ‘root determinative’ (see Hurwitz 1913) or as a ‘root augment’. It is not clear whether the vague idea of ‘root determinatives’ (a notion taken from Indo-European linguistics) or ‘augments’ (by definition with undetermined lexical and/or grammatical function) makes sense at all. Some linguists (especially Ehret 1989; cf. Belova 2004) posit the existence of a very large number, up to 37, of affixes, called by them ‘root determinatives’, to which they assign hypothetical functions and names. In such a model, practically every reconstructed consonant is interpreted as an original affix ‘enlarging’ alleged biconsonantal roots (but cf. Kuryłowicz 1972:26, who states that “most consonants could be used as infixes and determinatives”). Rejecting (morpho)phonological explanations (e.g. Bohas 2000, cf. Zaborski 2002; → lexicon: matrix and etymon model) and using hypothetical etymologies, this approach is very different from the method and the idea of reconstruction of biconsonantal roots, not only in Arabic, but also in other Hamito-Semitic languages. It has to be kept in mind that some triconsonantal variants are also due to ‘Reimwortbildung’, and some can be even neologisms, introduced by poets, or graphemic mistakes, taken over by lexicographers.

Bibliographical references

Primary sources

ʾAbū ṭ-Ṭayyib, *ʾIbdāl* = ʾAbū ṭ-Ṭayyib ʿAbd ar-Raḥmān ibn ʿAlī al-Luġawī, *Kitāb al-ʾibdāl*. Ed. ʿIzz ad-Dīn at-Tanūxī. Damascus, 1960.

Ibn as-Sikkīt, *ʾIbdāl* = ʾAbū Yūsuf Y*a*ʿqūb ibn ʾIsḥāq Ibn as-Sikkīt, *Kitāb al-ʾibdāl*. Ed. Ḥasan M.M. Šaraf. Cairo, 1978.

Secondary sources

Atallah, Wahib and Youssef Ayache. 1972. *L’alternance vocalique dans les racines concaves en arabe classique: Contribution à l’étude de la bilitarité*. (= *Cahiers du Centre de Recherches et d’Applications Linguistiques de l’Université de Nancy II*, 18.) Nancy: University of Nancy II.

——. 1977. *La bilitarité en arabe classique*. (= *Cahiers du Centre de Recherches et d’Applications Linguistiques de l’Université de Nancy II*, 31.) Nancy: University of Nancy II.

——. 1981. *La bilitarité en arabe classique*. II. *Les quadrilitères schématiques*. (= *Cahiers du Centre de Recherches et d’Applications Linguistiques de l’Université de Nancy II*, 37.) Nancy: University of Nancy II.

Belova, Anna G. 2004. *Sravnitel’no-istoričeskij aspekt mnogoznačnosti v korneslove arabskogo jazyka*. [Comparative-historical aspect of homonymy in Arabic roots]. Moscow: Institut Vostokovedenija RAN.

Bohas, Georges. 2000. *Matrices et étymons: Développements de la théorie*. Lausanne: Editions du Zèbre.

Chekayri, Abdellah. 1995. “Le radical et la variation”. *Recherches en linguistique arabe*, ed. A. Chouta and A. Jahfa, 37–55. Casablanca: Kulliyyat al-ʾĀdāb wa-l-ʿUlūm al-ʾInsāniyya, University Bin Imsik.

——. 1999. *L’articulation du lexique et de la morphologie arabes: Des verbes défectueux et particu-*

lièrement les verbes assimilés. Doctorat d'Etat, Mohamed V University, Rabat.
—— and Tobias Scheer. 2003. "The appearance of glides in Classical Arabic defective verbs". *Folia Orientalia* 3.5–34.
Colin, Georges S. 1931–1934. "Recherches sur les bases bilitères en arabe". *Comptes rendus du Groupe Linguistique d'Etudes Chamito-sémitiques* 1.9–10.
Corriente, Federico. 1969. "A survey of spirantization in Semitic and Arabic phonetics". *Jewish Quarterly Review* 60.147–171.
Ehret, Christopher. 1989. "The origin of third consonants in Semitic roots: An internal reconstruction applied to Arabic". *Journal of Afroasiatic Languages* 2.109–202.
El-Berkawy, Abdel Fatah. 1981. *Die arabischen* Ibdāl-*Monographien, insbesondere das* Kitāb al-Ibdāl *des Abū 't-Tayyib al-Lugawi*. Ph.D. diss., University of Erlangen.
Fleisch, Henri. 1990. *Traité de philologie arabe*. 2nd ed. 2 vols. Beirut: Dar el-Machreq.
Hurwitz, Solomon T.H. 1913. *Root-determinatives in Semitic speech: A contribution to Semitic philology*. New York.
Kuryłowicz, Jerzy. 1972. *Studies in Semitic grammar and metrics*. Wroclaw: Wydawnictwo Polskiej Akademii Nauk.
Nöldeke, Theodor. 1910. *Neue Beiträge zur semitischen Sprachwissenschaft*. Strassburg: K. Trübner.
Voigt, Rainer M. 1988. *Die infirmen Verbaltypen des Arabischen und das Biradikalismus-Problem*. Stuttgart: F. Steiner.
Zaborski, Andrzej. 1971. "Biconsonantal roots in Semitic". *Zeszyty Naukowe Uniwersytetu Jagiellońskiego – Prace Językoznawcze* 35.51–98.
——. 1991. "Biconsonantal roots and triconsonantal root variation in Semitic: Solutions and prospects". *Semitic studies in honor of Wolf Leslau*, ed. Alan S. Kaye, II, 1675–1703. Wiesbaden: O. Harrassowitz.
——. 2002. "The problem of biconsonantal roots in Afroasiatic languages". *Sprawozdania z Posiedzeń Komisji Naukowych* 46:1.71–73. Cracow: Polska Akademia Nauk.

Andrzej Zaborski (University of Cracow)

Bleaching → Semantic Bleaching

Bornu Arabic → Subsaharan Arabic

Braille

1. Introduction of Braille in the Arab world

Physically, Braille is a "universally accepted system of writing used by and for blind persons and consisting of a code of 63 characters, each made up of one to six raised dots arranged in a six-position matrix or cell" (*Encyclopedia Britannica* II, 465). Content-wise, Arabic is a six-dot tactile copy of its *schwarzschrift* (normal ink print). The system is divided into the alphabet and its subsystems, the non-alphabetical code systems of contractions, and the mathematical signs and musical notation. One interesting fact is that Braille is a functionally limited system of writing. From its introduction to the Arab world in Egypt in the second half of the 19th century, the system was, and still is, functionally limited to the field of education. Very little non-educational material is printed in Braille in any given year.

The history of the introduction of Braille to the Arab world is vague, perhaps because it was a non-governmental initiative with little documentation (al-Sharkawi 1997:31–32). It was first introduced in the educational system of the visually impaired in Egypt by Muḥammad ʾAnas, an Arabic teacher and private school owner in Cairo. ʾAnas traveled to France where he learned Braille in the same institute where Louis Braille studied and worked (Makhluf 1995). After returning to Egypt, ʾAnas established a school for the blind in his native popular quarter of Šayxūn in Cairo, where Braille was used for the first time as a medium for education. ʾAnas adapted the French Braille system to the Arabic language. Named after its creator, the script he devised came to be known as *al-xaṭṭ al-ʾanasī*. For printing Braille, ʾAnas used the same tools as in Europe, the slate and the stylus. No traces of that adaptation survived because the project came to an end when the school was closed after the death of its owner (al-Sharkawi 1997:34).

Subsequent projects to introduce Braille in Egypt until the first half of the 20th century were sporadic. At the end of the 19th century, a British school run by a Dr. Scott was established and Braille was reinstated as the medium of education. Little is known about the nature of Braille at that time: owing to the rising nationalist spirit of the period, the school was closed at the beginning of the 20th century and Braille faded away (al-Sharkawi 1997:35–36). In 1935, Braille was restored to schools once more, but remained confined to the elementary schools until it was extended to preparatory schools in 1957 and to secondary schools in 1960 (al-Sharkawi 1997:36). The importance of this expansion of Braille through secondary education is that it necessitated devising arithmetical and musical codes, thus enlarging the system. From 1935 onwards, the type of Braille used was the same as the type described in (2) below.

Braille cells are upright rectangular shapes made of two vertical columns. Each is made of three dot positions, which are numerically identified as dots from 1 to 6. Dots 1–3 form the right column, and the dots 4–6 form the left column from the pressed side of the page. Dots are separated from one another by thin vertical and horizontal empty stripes made possible by the metal wall separating the dot cells. Letters and symbols are formed by embossing dots from side A (the upper side) to side B (the lower pressed side) by means of a stylus (a sharp-ended hand tool), which presses against six dot positions on side B. A normal Braille line is made of 30 dot cells.

Through the combination of dot positions and their distribution on the two vertical columns, the symbol takes a distinctive tactile shape. Empty dot positions help the reader identify the embossed positions forming the letters. Between dot cells there is a barrier. The direction of embossing symbols is right to left, and reading goes from left to right, even in Arabic and in top-to-bottom scripts. Groups of symbols that belong to one another are in adjacent dot cells. Between groups of symbols there is a separating empty dot cell. The up and down horizontal contours of cells form the physical borders of lines (al-Sharkawi 1997:10–17).

2. The Arabic Braille alphabet system

The alphabet system in Arabic Braille, albeit for no physical necessity, is divided into three subsystems: the alphabet letters, the short vowels (in addition to case endings, feminine marker, and *ʾalif maqṣūra*), and the *hamzas*. Although all these subsystems can theoretically be represented along the same horizontal line, as in Arabic *schwarzschrift*, the two latter sub-systems are not perceived as letters of the alphabet.

The Arabic Braille alphabet is made of 29 letter symbols, although the letters of the *schwarzschrift* alphabet are only 28. In Braille there is the additional symbol for *lām-ʾalif*. Table 1 gives the dot representations of the alphabet.

Table 1. The Arabic Braille alphabet

Letter Name	Letter Number	Dot Representation
ʾalif	1	1
bāʾ	2	1–2
tāʾ	3	2–3–4–5
ṯāʾ	4	1–2–3–4
jīm	5	2–4–5
ḥāʾ	6	1–5–6
xāʾ	7	1–3–4–6
dāl	8	1–4–5
ḏāl	9	2–3–4–6
rāʾ	10	1–2–3–5
zāy	11	1–3–5–6
sīn	12	2–3–4
šīn	13	1–4–6
ṣād	14	1–2–3–4–6
ḍād	15	1–2–4–6
ṭāʾ	16	2–3–4–5–6
ḏ̣āʾ	17	1–2–3–4–5–6
ʿayn	18	1–2–3–5–6
ġayn	19	1–2–6
fāʾ	20	1–2–4
qāf	21	1–2–3–4–5
kāf	22	1–3
lām	23	1–2–3
mīm	24	1–3–4
nūn	25	1–3–4–5
hāʾ	26	1–2–4
wāw	27	2–4–5–6
lām ʾalif	28	1–2–3–6
yāʾ	29	2–4

The right column is the dominant one from the embossing side, which is the left tactile side. The table also shows that only one letter is represented by one dot position, *ʾalif*; and one letter is represented by the full six dot positions, *ḏ̣āʾ*. Only two letters, *bāʾ* and *yāʾ*, are represented by two dot positions, while the majority of the letters use three, four, or five dot positions. Eleven letters are represented by three dot positions, ten by four, and four by five.

The Arabic Braille letters that stand for the same, similar, or even broadly similar sounds in other languages have the same dot representations. Number 2 in Table 1 above, for instance, stands for the letter *bāʾ* which represents the voiced plosive bilabial /b/. The letter *b* in the languages that use the Latin script, which represents similar sound qualities, has the same dot distribution in Braille.

As in the *schwarzschrift* of Arabic, short vowels are not part of the Braille alphabet. They are the same dot representations given to the case endings, and are therefore categorized with them as elements of *taškīl*. In normal individual writings and printing of books, words are written without short vowels, although there is no physical hindrance to align short vowels along the same horizontal line with consonants. In some cases, however, short vowels are represented inside the word extending its horizontal length (al-Sharkawi 1997:206–210). Short vowels are represented in Table 2.

Table 2. Short vowels

Symbol Name	Dot Representation
fatḥa	2
kasra	1–5
ḍamma	1–3–6

As in the case of the consonants, short vowel dot representations are right-column dominant. The same dot distributions are used to stand for case endings at the end of words. Categorized in the same subcategory are three other scriptural devices: *šadda* 'doubling'; *ʾalif maqṣūra*; and *tāʾ marbūṭa* (the feminine ending) (al-Sharkawi 1997:94–95). Dot representations for these are given in Table 3.

Table 3. Non-Short vowel symbols

Symbol Name	Dot Representation
ʾalif maqṣūra	1–3–5
šadda	6
tāʾ marbūṭa	1–6

Hamza (the glottal stop) is represented by five symbols in Arabic Braille. Four of the five values. represented by these symbols are complex sound values (*hamza* plus a short or long vowel). Although the *hamza* and each vowel have separate dot representations, a sound combination cannot be expressed using two symbols. A *hamza* followed by a short /a/ vowel, for example, is a sound combination expressed by a symbol that is different from both the symbols allocated to the short vowel and the one allocated to the *hamza*. Table 4 gives the Braille dot representations of the *hamzas*:

Table 4. *Hamzas*

Symbol Name	Dot Representation
hamza	3
hamza ʿalā ʾalif	3–4
hamza ʿalā madd	1–2–6
hamza ʿalā yāʾ	1–3–4–5–6
hamza ʿalā wāw	1–2–5–6

Punctuation marks in Arabic Braille are seven in number and are put immediately after the last letter of the word before the blank space which separates words. Physically, punctuation marks in the Braille system are different from the alphabet in that they do not use the dots 1 and 4, leaving the upper part of the dot cell empty. Another difference is that some punctuation marks are represented by two dot cells, while the alphabet letters are represented only by one dot cell. Table 5 presents the punctuation marks.

Table 5. Punctuation marks

Symbol Name	Dot Representation
Comma	5
Full stop	2–5–6
Colon	5–2
Semi-colon	2–3–6
Dash	2–5–2–5
Brackets	2–3–5–6 2–3–5–6
Parentheses	2–3–6 3–5–6

The two cases of the short vowels and the *hamzas* point to the fact that the Braille alphabet system was devised with the purpose of providing a tactile equivalent symbol for each *schwarzschrift* one. Although Braille does not face the physical problems encountered by *schwarzschrift* because it does not need to mount short vowels on *hamzas*, there was no intention to solve in Braille the problems of vowels and symbol complexity in the *schwarzschrift*. Braille has also inherited the *schwarzschrift* problem of the long vowel representation: symbols 27 and 29 represent not only the long vowels /ū/ and /ī/ respectively, but the diphthongs /w/ and /y/ as well.

Braille also has its own physical problems, mirror opposition and short vowel blocking being the two most salient. Mirror opposition is when a certain dot representation is exactly the opposite shape of another dot representation. Eight pairs of letters have this problem: 5-26, 10-27, 6-12, 8-20, 9-11, 13-24, 15-25, 16-18 in

Table 1 are mirror oppositions. Two other letter representations are mirror oppositions of non-alphabet symbols: 29 is a mirror opposite of the *kasra* symbol, and 14 of Table 1 is also a mirror opposite to the *hamza ʿalā yāʾ*. It is a confusing phenomenon, because fast reading depends on readily realizing shapes (al-Sharkawi 1997: 142–147). Vowel blocking happens when contractions are used. Although uncommon, theoretically one can use short vowels in Arabic Braille. If contractions are used for clusters of letters, no short vowel representation is allowed inside the word, nor is a case ending. If short vowels or case endings must be represented, no contraction can take place, and the size of the fully represented words becomes much longer (al-Sharkawi 2002:205–212).

3. CONTRACTIONS

In Braille, a word can take up a large horizontal space on the line. Therefore a system of contraction symbols was devised in order to reduce the number of dot cells needed for a word. Contractions are one or two dot cells used to stand for full words, morphemes in words, or even consonant clusters (al-Sharkawi 1997: 124). They are divided into two categories: the first contains simple contractions, which are one dot cell units. The second contains complex contractions, which are two dot cells for one word. Letters forming one word can be a part of another word. In such a case, however, contraction takes place with certain limitations. If the word or cluster of letters has three or four letters, and if it is attached to a function word, a separation mark (dots 3–6) has to be added before the contracted cluster when the contraction symbol is an alphabet letter. Yet, when the contraction symbol is a non-alphabet letter, there is no limitation. If the contraction symbol is a symbol of punctuation marks or case ending, it cannot be used to contract a letter cluster at the end of a word. If a cluster of letters happens to be composed of the same letters as a functional morpheme, it cannot be contracted in the middle of the word. Therefore, functional morpheme contractions are limited to the end of the word. Finally, if the contraction symbol is in mirror opposition to the preceding letter in the word, contraction is blocked. Contraction conditions are devised to avoid confusion between contraction symbols and single value symbols.

The number of simple contraction symbols is 55. The majority are alphabet dot representations that contract full function words (prepositions, conjunctions, pronouns, demonstratives) and grammatical morphemes in words (definite article, plural and dual morphemes). In most of the simple contractions, the first letter of the word is used as a contraction symbol. When grammatical morphemes are contracted, non-letter symbols are used, and the contracted element remains in its position in the word. Table 6 contains some examples of simple contractions.

Table 6. Examples of simple contractions

Contraction	Contracted word	Meaning
1–2 (*bāʾ*)	*ball*	'but' [conjunction]
2–3–4–5 (*tāʾ*)	*tilka*	'that' [fem. demonstrative]
1–2–3–5–6 (*ʿayn*)	*ʿindamā*	'when' [conjunction]
1–5–6 (*ḥāʾ*)	*ḥattā*	'until' [particle]

Complex contractions are full words contracted in two dot cells: the first part is a non-alphabet symbol, while the second part is a letter in the contracted word. The first part only uses the left vertical column, dots 4–6.The total number of complex contractions is 124.

4. CODE SYSTEMS

Arabic Braille has mathematical and musical codes. Code systems differ from the alphabet structurally in that there are areas in the dot cell they do not use, while the alphabet uses the two vertical columns and the three horizontal lines of the cell. Numerals, not arithmetic signs, use the upper two lines of the dot cell, leaving the bottom dots 3–6 empty, while the musical code uses the bottom two lines, leaving the upper line 1–4 dots empty. In addition, numerals are distinguished by a number marker put before the number to distinguish it from alphabet letters.

Like the alphabet, numerals are written from left to right, and read from right to left. But they use the upper and middle horizontal lines, and not the bottom one. Numerals are clustered beside one another without a space in between. Before the number cluster, there is a number marker. After the cluster ends, there is an empty dot cell. Arithmetic symbols, unlike numerals, use the bottom two lines in the dot cell. Table 7 gives the dot representations of the numerals:

Table 7. Numerals

Number	Dot Representation
0	20405
1	1
2	1–2
3	1–4
4	1–4–5
5	1–5
6	1–2–4
7	1–2–4–5
8	1–2–5
9	2–4
Number symbol	3–4–5–6

The numbers 0, 1, 2, 4, 5, 6, 8, and 9 are dot representations of alphabet letters and short vowels. The numbers 3 and 7, however, are dot representations for contraction symbols. Like the alphabet, numeral dot representations must contain dots in the right vertical column. Table 8 gives the arithmetic symbols in Braille.

Arithmetic symbols are added between numbers without a separating space. After a symbol a number symbol is not necessary.

The musical code of the Arabic Braille system uses the same dot distributions as the numerals, but one line down horizontally. If the number 1 is represented by dot 1, the first note is represented by dot 2. By the same token, if the number four is represented by dots 1-2-4, note d is represented by the dots 2-5-6. Bars are represented by dot representations clustered beside one another, and an empty space stands between bars.

Table 8: Arithmetic Symbols

Arithmetic Symbol	Dot Representation
+	2–6
-	3–5
*	2–5–6
÷	2–3–5
=	2–5 2–5

BIBLIOGRAPHICAL REFERENCES

Sharkawi, Muhammad al-. 1997. *The Arabic Braille: Evaluation and suggestions for modification*. M.A. thesis, American University in Cairo.

——. 2002. "ʾAṣwāṭ al-līn fī ṭarīqat brāyl al-ʿArabiyya". *al-ʿArabiyya: ʾAbḥāṯ luġawiyya wa-jtimāʿiyya wa-tarbawiyya*, ed. Ālaa Elgibali and El-Said Badawi, 207–212. Cairo: Arabic Language Institute, American University in Cairo.

Maxlūf, ʿAbd al-Ḥakam. 1995. *Tarbiyat al-muʿawwaqīn baṣariyyan*. Cairo: al-Nur Institute for the Blind.

MUHAMMAD AL-SHARKAWI
(American University in Cairo)

Bukaṛa-syndrome

The consonant *r* (or velarized *ṛ*) is realized in many dialects with a degree of delay. When *r* directly follows the consonant in a sequence Crv, such delay may result in the realization of an intrusive vowel preceding *r* or *ṛ*. This phenomenon was termed the '*bukaṛa*-syndrome' by Woidich (1978). In allegro speech, however, the syndrome usually remains absent.

Such *bukaṛa*-vowels are often heard in northern and southern Middle Egyptian dialects, including the Fayyūm oasis (see Behnstedt and Woidich 1985:maps 47–49) and in most parts of the oases of the Western Desert of Egypt (see Woidich 1978; Behnstedt and Woidich 1982:50, 1985, map 47). The phenomenon was also observed in several of the Bedouin dialects of Sinai (see, e.g., de Jong 2000:112–118, 266–267, 352, 431–432).

In what is termed the 'simple *bukaṛa*-syndrome' the phonetic quality of the inserted vowel is guided by the vowel following *r* or *ṛ*. The rule for the simple *bukaṛa*-syndrome may be summarized as follows:

Ø > v / - C__r v
[a] [a]

C = any consonant
r = *r* or *ṛ*
[a] = a fixed set of phonetic features

The process entails the following: when a vowel – be it a base vowel or an anaptyctic (see below) – is to be realized following *r*, voicing of this *r* is already being produced before the tongue has been fully brought into position for the actual realization of *r*. Since the realization of the vowel following *r* is already being anticipated, the phonetic quality of the voicing will be guided by this vowel following *r*.

Some examples are (*bukaṛa*-vowels underlined): (from northern Middle Egyptian) (the syndrome's namesake) *bukṛa* > *buka̱ṛa* 'tomorrow', *ḥamra* > *ḥama̱ṛa* 'red [fem. sg.]', (from

Sinai) *yigrib* > *yigi̱rib* ‘he comes near’, *úḏukruw aḷḷah* > *úḏuku̱ruw aḷḷah* ‘pronounce God’s name’, *bizrih* > *bizi̱rih* ‘a seed’, *kiṯrit állaḥam* > *kiṯirit állaḥam* ‘the great quantity of meat’, *bakraj* > *baka̱raj* ‘coffee pot’.

In generative linguistic terms, the (synchronic) simple *bukaṛa*-rule is to be regarded as a late phonetic surface rule; it is applied in the last instance, after rules for → vowel elision and → anaptyxis have been applied (and these are not reapplied, i.e. they are not cyclic).

To illustrate: when we analyze the form *úḏukuruh* ‘pronounce His name’, we see that it is composed of the imperative form *uḏkur* + object suffix *-uh*. In dialects where the ordering of rules is

base form
uḏkur + uh
↓
suffixation
uḏkuruh
↓
elision
uḏkruh
↓
stress
úḏkruh
↓
anaptyxis
úḏukruh

we notice that the outcome is *úḏukruh*, rather than *úḏukuruh*. The rule of the *bukaṛa*-vowel insertion should then follow the anaptyxis rule to arrive at the surface form *úḏukuruh*. We would not arrive at the same surface form if we were to order the simple *bukaṛa*-rule at an earlier stage; if the *bukaṛa*-rule were to precede the anaptyxis rule, there would be no cluster to be resolved and we would therefore not be able to account for the anaptyctic vowel u preceding k in the surface form.

A comparable example form from northern Middle Egyptian (where the anaptyxis rule is Ø > ə / CC_C) is *kasiri̊ dīk* ‘the mating [lit. breaking] of the rooster’ (Behnstedt and Woidich 1988:12, text 45.12; the anaptyctic vowel is here transcribed as *i̊*), where the phonetic quality of the *bukaṛa*-vowel is guided by the anaptyctic vowel. This implies that the latter is inserted before the *bukaṛa*-rule is activated.

An example from the Fayyūm oasis in Egypt clearly illustrates the coloring of *bukaṛa*-vowels with the various vowels following *r*: *midara* ‘winnowing fork’, (with raised T in pause) *midere#* and (a suffixed form) *midiriti* ‘my winnowing fork’ (M. Woidich, p.c.).

The terms ‘expanded’ or ‘greater *bukaṛa*-syndrome’ describe the appearance of a short vowel preceding word-final *r*. In such cases the vowel produced tends to color with its phonetic environment; toward between [ə] and [u] in velarized or labial environments, and between [ə] and [ɪ] in neutral environments.

These terms are used to describe the (diachronic) process producing such intrusive vowels in dialects that eliminate clusters of three consonants by inserting an anaptyctic vowel preceding the first consonant from the right (i.e. Ø > ə / CC_C) (as in Middle Egyptian dialects), or do not eliminate such sequences at all. In the case of northern Middle Egyptian dialects, the syndrome has led to morphological restructuring of base forms, which may be concluded from such forms as *gabírha* ‘her grave’ (see Behnstedt 1979:65).

Many dialects have a rule inserting an anaptyctic vowel preceding the second consonant from the right, i.e. Ø > ə / (C)C_CC. An example is (anaptyctic underlined) *kasi̱r ligzāz*, composed of the base forms (*kasr* + *ligzāz*) ‘the breaking of the glass’. If, however, a vowel appears in, e.g., *kasir igzāz* (composed of the morphological base forms *kasr* + *igzāz* or *kasr* + *gzāz*) this vowel is best ascribed to the influence of the *bukaṛa*-syndrome, since the rule for anaptyxis cannot account for the appearance of a vowel in such a position.

Often, however, such vowels are not only produced preceding *r*, but also preceding other highly sonorous continuants, such as *l*, *n*, and *m*. In the case of Egyptian dialects, northern Middle Egyptian 2 shows ‘expanded’ (or ‘greater’) *bukaṛa*-vowels preceding word-final *r*.

Notice that (as in the case of Egyptian dialects) there is a direct relationship between the dissonorization of voiced consonants in pause and the expanded *bukaṛa*-syndrome: in dialects where dissonorization of final voiced consonants in pause is total, voicing of final *r* will be absent, preventing the expanded *bukaṛa*-syndrome from becoming active. The rules are thus mutually exclusive.

Summarizing the ‘greater *bukaṛa*-syndrome’ we may omit the vowel following *r* from the previous rule, thus:

Ø > v / -C__r

Another aspect of the ‘greater *bukaṛa*-syndrome’ is the realization of vowels preceding word-initial *r* after speech pause, as in (underlined) (*#rama* >) *a̱rama* ‘he threw’, (*# ṛāḥit* >) *a̱ṛāḥit* ‘she went away’, and (*#risil* >) *i̱risil* ‘he sent’ (although such vowel insertion preceding *ri* or *ru* appears to be much less regular than preceding a sequence *ra* or *rā*). Also across word boundaries (though again predominantly when *a* or *ā* follows *r*), an ‘*a*-Vorschlag’ is realized as in, e.g., *gaṭaʿ a̱ragabit ilwizīr* ‘he cut the throat of the minister’ (see Behnstedt 1979:65–66). This *a*-Vorschlag occurs in the dialects of the western oases and in the Nile Valley, roughly between Bani Swayf in the north and al-Minya in the south (around Bani Mazār; northern Middle Egyptian 2). The rule here is:

Ø > v / C/__rV
[a] [a]

C = any consonant or a speech pause
v = short vowel phonetically near or identical to *V*
V = short or long vowel, predominantly *a* or *ā*
[a] = a fixed set of phonetic features (not including length of vowel)

Note that Behnstedt (1979:66) makes an exception for the consonant *n*, which is assimilated to the following *r*, e.g. (*kān rājil* >) *kāṛ ṛājil* ‘he was a man’. The same exception will hold for the consonant *r*: *dāṛ ṛājil* ‘a man’s house’.

A last aspect to be mentioned is the vowel-preserving influence of the *bukaṛa*-syndrome. In some dialects morphophonemic or sandhi elision of short vowels does not take place (especially not in lento speech) when these vowels are followed by *r*. Examples are (preserved vowels underlined) *mitnaṭṭʾi̱rak* ‘waiting for you’, *fināxi̱rak* ‘your nostrils’, and a sandhi example *agōṭi̱r ajīb* ‘I go and bring (back)’ (cf. de Jong 2000:115).

Bibliographical references

Behnstedt, Peter. 1979. “Die nordmittelägyptischen *bukara*-Dialekte”. *Zeitschrift für Arabische Linguistik* 3.62–95.
—— and Manfred Woidich. 1985. *Die ägyptisch-arabischen Dialekte*. II. *Dialektatlas von Ägypten*. Wiesbaden: L. Reichert.
——. 1985. *Die ägyptisch-arabischen Dialekte*. III. *Texte. 2/3 Niltaldialekte Oasendialekte*. Wiesbaden: L. Reichert.
Jong, Rudolf de. 1996. “More material on Fayyūmi Arabic, I”. *Zeitschrift für Arabische Linguistik* 31.57–92.
——. 2000. *A grammar of the Bedouin dialects of the northern Sinai littoral: Bridging the linguistic gap between the eastern and western Arab world*. Leiden: E.J. Brill.
Woidich, Manfred. 1978. “Bemerkungen zu den arabischen Dialekten Mittelägyptens”. *Zeitschrift für Arabische Linguistik* 1.54–63.

Rudolf de Jong
(University of Amsterdam)

C

Cairo Arabic

1. General

Cairo Arabic, often simply called Egyptian Arabic, is the main linguistic vehicle used all over Egypt. Originally the dialect of the capital Cairo, it is being adopted by millions of speakers all over Egypt, in particular in urban areas, as a language of daily life. The number of native speakers is also several millions. The variety described here is spoken by the middle class as it is also commonly heard in the media (radio, television, movies). Owing to the cultural export of Egypt to other Arab countries, Cairo Arabic is widely understood in the whole Arab world.

Cairo Arabic is an urban dialect of the Southeastern Mediterranean and stands not too far from the varieties of Arabic of Jerusalem, Damascus and Beirut, with which it shares an important feature. i.e. the replacement of *_q_ by the glottal stop /_ʾ_/ and the replacement of the historical interdentals by alveolar stops. In contrast to these Levantine cities, however, *_g_ does not show any sign of palatalization in Cairo and is pronounced [g]. Another difference is that stress falls on the penultima in sequences CvCCvCv (*madrása* 'school'), not on the antepenultima as in the Levant (*mádrasa* 'school'). These three features class Cairo Arabic with the Egyptian dialects spoken in the Central Delta and on both shores of the eastern branch of the river Nile up to Damietta. Cairo Arabic differs from these central dialects in its lack of pausal forms for *-a* and the extension of the suffix *-it* of the 3rd pers. sg. fem. perfect to all verbal classes, including verbs III*y* (Woidich 1994): Cairo *baʾit* – *lammit* – *ḍarabit* 'she became' – 'she took up' – 'she hit', vs. *baʾat* – *lammat* – *ḍarabit* in the Delta.

Today, Cairo Arabic is widely used in the Egyptian media, often with an admixture of Standard Arabic (→ diglossia). It has a lively and growing tradition as a written language, which has been developing in particular from the end of the 19th century onwards. Not only folk poetry, dialogues in novels, theater plays etc., but also entire novels and memoirs find their means of expression in the colloquial. Several important texts have come down to us from the Middle Ages (→ dialect literature).

In Cairo itself, linguistic differences between Muslims and Copts are limited to lexical items refering to religion proper. The former Jewish community spoke a *niktib-niktibu* dialect (Blanc 1974) apparently of North African provenance (→ Judaeo-Arabic). A small community of Christians of Syrian and Lebanese background, who arrived some time in the 19th century, still display some pecularities such as *iga* 'he came' (*ga* ~ *gih*) and *ʾultillu (ʾultilu)* 'I told him' (see Tadié 1994). On the impact of Standard Arabic on phonology see Haeri (1996), on argot see Rowson (1981).

Cairo Arabic is one of the most extensively researched and described Arabic dialects to date. The first comprehensive scientific grammar of an Arabic dialect, Spitta (1880), is still a valuable source for 19th-century Cairo dialect. Mitchell (1962) is a comprehensive short reference grammar, Wise (1978) applies generative syntax to Cairo Arabic, and a more detailed reference grammar with the focus on syntax is Woidich (2006). Much detailed research has been done on phonology, syntax, etc. (see, e.g., Harrell 1957, Eisele 1999, Brustad 2000). As for dictionaries, Spiro (1923) is still good for older texts; for more recent ones Hinds and Badawi (1984) is an indispensible resource. Collections of texts can be found in Elder (1927), Hassan (1971), and Prasse a.o. (2000); audio material can be downloaded from the Semitic Language Archive, <http://www. semarch. uni-hd.de/index>.

2. LINGUISTIC DESCRIPTION

This short description is limited to the most essential features of phonology and morphology of Educated Cairo Arabic, followed by some remarks on syntax and lexicon.

2.1 *Phonology*

2.1.1 Consonants (Table 1)

The marginal consonants /v, ž/ are to be found in loans such as *villa* 'villa', *žambōn* 'ham', /q/ [q] in loans from Standard Arabic such as *qaṛya* 'village', *ilqāhíra* 'Cairo'.

/ṛ, ḷ, ṃ, ḅ/ are phonemes by secondary emphasis as can be shown by minimal pairs: *rāʾid* 'lying' – *ṛāʾid* 'major', *walla* 'or' – *waḷḷa* 'by God', *ʾabla* 'before' – *ʾaḅla* 'school-mistress', *yaṃṃa* 'oh mummy' – *yamma* 'side'.

All interdentals are reflected by plosives in true dialect words, in loans from Standard Arabic by sibilants such as *sābit* 'stable', *kizb* 'lie'. The most conspicuous feature is the reflex of /g/ for **g* which in most other Arabic dialects is palatalized. Blanc (1981) explains this as a reintroduction of /g/ for earlier /j/, which he assumes to have taken place from the end of the 18th century onwards; see also Hary (1996) for a comprehensive study. In contrast, Woidich (1996) based on evidence from dialect geography concludes that /g/ must have been prevalent in the Middle Ages: the distribution of /g/ today follows the medieval trade route from Cairo along the eastern branch of the Nile to Damietta, the medieval harbor of Cairo. This is why /g/ must be much older in Egypt than two or three hundred years. It is more likely that /g/ is the original reflex of **g* in the Delta including Cairo, which was pushed back later by /j/-speaking Bedouin moving in from the east and west, which eventually led to the geographical spread seen today.

2.1.2 Vowels

Table 2. Vowels in Cairo Arabic

Short vowels			Long vowels		
i		u	ī		ū
(e)		(o)	ē		ō
	a			ā	

The short high vowels are pronounced lax with allophones ranking between [ɪ] – [e], and [ʊ] – [o] respectively. Unconditioned *ʾimāla* of /a/ [æ] and /ā/ [æː] is usual, often moving further up in the direction of [ɛː] ~ [eː] in the speech of women. It should be noted that this [æ] is far more raised and centralized as in [æ̝̈] than the corre-

Table 1. Inventory of consonants in Cairo Arabic

	bilabial	labiodental	alveolar	postalveolar	palatal	velar	pharyngeal	laryngeal
plosive								
voiceless, voiced	b		t, d			k, g		ʾ
emphatic	ḅ		ṭ, ḍ					
nasal	m		n					
emphatic	ṃ							
fricatives								
voiceless, voiced		f	s, z	š		x, ġ	ḥ, ʿ	h
emphatic			ṣ, ẓ					
trill			r					
emphatic			ṛ					
lateral			l					
emphatic			ḷ					
glides	w				y			

sponding British English sound represented by [æ], see Mitchell (1962:22). All vowel phonemes have pharyngealized (emphatic) allophones. The opposition /i/ – /u/ has full functional load, e.g. *fulla* 'a jasmine flower' – *filla* 'a cork', but less so in unstressed open syllables as in *gimāl* 'camels' – *gumāl* 'beautiful [pl.]'. The [ɪ] inserted into a sequence of -CCC- is written *i̊* here when it happens to separate two words: *ikkalbi̊ da* 'this dog', see 2.1.3.

/e/ and /o/ are marginal and appear only in careful speech as lento-forms and replace /ē/ or /ō/ in unstressed position or before a consonant cluster due to morphophonological changes: *bēt+na* > *betna* ~ *bitna* 'our house', *yisōraʾ+u* > *yisōrʾu* > *yisorʾu* ~ *yisurʾu* 'they faint'.

/ē/ and /ō/ result from the contraction of **ay* and **aw*: *dēl* < **ḏayl* 'tail', *yōm* < **yawm* 'day'. /aw/ and /ay/ are preserved before /w/ and /y/ respectively, or if otherwise the morphological word pattern would become opaque: *ʾawwil* 'first', *ʿayyil* 'child', and *mawlūd* 'born' (maCCūC), *awḥaš* 'worse' (aCCaC).

Final *-*āʾ* and *-*ā* as in **šitā́ʾ* 'winter', **ḥamrā́ʾ* 'red [fem.], **qafā* 'neck', develop into -*a* due to loss of the final glottal stop, stress shift to the first syllable, and shortening of final *-*ā*: *šíta*, *ḥámra*, *ʾáfa*.

2.1.3 Syllable

The possible phonological syllable types are: Cv, Cv̄, CvC, Cv̄C, CvCC. The latter two appear only as last syllables of a word. Cv̄C is treated like CvC, i.e., it gets the same series of consonant-initial pronominal suffixes, *kutub+ha* > *kutubha* 'her books' *kitāb+ha* > *kitabha* 'her book', while CvCC combines with a special set of vowel-initial suffixes, as in *kalb+aha* > *kalb+áha* 'her dog', see 2.4.1.2.

2.1.4 Word stress assignment

In general, stress is determined by a phonological principle, i.e. the occurrence of heavy and light sequences in the last three syllables of a word (not including the article *il*-). Heavy sequences have the form -vCC or -v̄C-. If these occur in a word, stress is assigned to the vowel of this sequence if it is followed by no or only one vowel, as in *ġasált* 'I washed', *maṭā́r* 'airport', *kálbu* 'his dog', *bḗti* 'my house'. Otherwise, the vowel after the heavy sequence receives stress, as in *madrása* 'school', *ištáġalu* 'they worked', *ilqāhíra* 'Cairo'. If there are no heavy sequences, stress falls on the vowel of the first syllable, but will never move further back to the left than the antepenultimate: *ʿásal* 'honey', *sámaka* 'a fish'.

Morphological stress assignment includes the 3rd pers. sg. fem. perfect, which receives stress whenever a suffix is attached as in *balaʿítu* 'she swallowed it', see below 2.3. In addition, there are the plural patterns CuCúCa and CiCíCa, e.g. *bunúka* 'banks', *sibíta* 'baskets', and the presentatives *ahú* sg. masc., *ahí* sg. fem., *ahúm* pl. 'there is . . .', both with phonologically irregular stress assignment.

Stress forms a phonemic opposition in the minimal pair *sínima* 'cinema' and *siníma* 'camel humps'.

2.2 *Phonotactics*

Voice assimilation: in consonant clusters the last consonant determines the sonority of the cluster if it is a plosive or a fricative, making the cluster voiced, as in *yifḍal* ['jɨv̶ᵭɑɫ] 'he remains', *tikdib* ['tɪgdɪb] 'she lies', *masgūn* [mæz'guːn] 'jailed', or unvoiced, as in *madxal* ['mætxæl], *yiġsil* ['jɪxsɪl] 'he washes', *azfat* ['æsfæt] 'worse'. Devoicing also takes place in pause as in *balad* ['bælæt] 'town', *nadġ* [nætx] 'chewing tobacco'.

Sun letters: apart from the usual alveolar stops and sibilants, /l/ of the article *il*- assimilates to /g/ and /k/, e.g. *iggibna* 'the cheese', *ikkōṛa* 'the ball'.

Reflexive passive *it-:* the *it*-prefix of the reflexive passive verbal forms assimilates to alveolar stops, sibilants, and to /g/ and /k/: *iddafan* 'to be buried', *iṭṭafa* 'to be extinguished', *issaraʾ* 'to be stolen', *iššāl* 'to be taken away', *iggawwiz* 'to be married', *ikkaʿbil* 'to stumble'.

Spread of emphasis: suprasegmental spread of pharyngealization starting from an emphatic consonant takes place regularly as in *baṭnaha* [ᵬɑᵵᵰɑhɑ] 'her stomach', *taṣliḥāt* [ᵵɑᵴɫɨħɑːᵵ] 'repairs', *abyaḍ* [ʔɑᵬɟɑᵭ] 'white'.

2.3 *Morphophonology*

Cairo Arabic is a 'parler différentiel', which means that /i, u/ are elided in open unstressed syllables after short open syllables in contrast to /a/: i, u > Ø / . . . vC_CV, e.g. *zinix+a* > *zinxa* 'rancid [fem.]', *širib+it* > *širbit* 'she drank', but *balaḥ+a* > *balaḥa* 'a date'. Both /i/ and /a/ undergo elision after a long open syllable, i.e. . . . v̄C_CV, as in *sāfir+it* > *safrit* 'she traveled', *baniʾādam+īn* > *baniʾadmīn* 'human beings', *miṭēwal+a* > *miṭiwla* 'oblong [fem.]'.

A cluster -CCC- is resolved by inserting a vowel /i/ following the second consonant counted from the left: Ø > i / . . . vCC_C(C)V, cf. I > *ʾultílu* ‘I said to him’, *iššahṛ+da* > *iššahṛída* ‘this month’. This holds in sandhi too where *i̊* is inserted: *kunti̊ ṛāyiḥ* ‘I was going’.

If, as a result of suffixation, a long vowel precedes a consonant cluster, or lands in an otherwise unstressed position, it is shortened: *ṣāḥib+i* > *ṣaḥbi* ‘my friend’, *šāf+ni* > *šafni* ‘he saw me’, *šāfu+na* > *šafūna* ‘they saw us’. /ē/ may be replaced in the latter case by /a/ as in *ḥabbḗtu* ‘I loved him’ but *ma-ḥabbatū́š* ‘I did not love him’.

When suffixes are added to base forms ending in vowels other than *-a* of the feminine suffix, this vowel is lengthened, e.g. *kalti+ha* > *kaltīha* ‘you [fem.] ate it’, *ma-tinsa+š* > *ma-tinsāš* ‘do not forget!’.

In the construct state, the allomorph of the nominal feminine *-a* is *-it*, as in *šuṛba* ‘soup’ but *šurbitʿads* ‘lentil soup’. With possessive suffixes: *ʿizba* ‘farm’ but *ʿizbitu* ‘his farm’; the feminine active participle lengthens this *-a*, as in *samʿa+ha* > *samʿāha* ‘hearing her’.

The suffix of the 3rd pers. sg. fem. perfect receives stress when suffixed: *ġasalítu* ‘she washed it’, *nammaṛítu* ‘she numbered it’, *xadítu* ‘she took it’, *basítu* ‘she kissed him’, *ramítu* ‘she threw it’.

The 3rd pers. sg. masc. object suffix *-(h)* after vowels changes into *-hu-* when other suffixes follow: *ma-nisīti(h)+š* > *ma-nsitihūš* ‘you [fem.] did not forget him’, *warra(h)+lu* > *warrahūlu* ‘he showed it to him’. *-u* may also change into *-hu* when a negational *-š* follows: *ʾuddāmu* ‘in front of him’ but *ma-ʾuddamūš* ~ *ma-ʾuddamhūš* ‘not in front of him’. *-ik* of the 2nd pers. sg. fem. is replaced by *-iki-* in this case and the final vowel is lengthened: *ʿandik* ‘with you’ but *ma-ʿandikīš* ‘not with you’. *-ya* of the 1st pers. sg. may be replaced by *-yi-* as in *maʿāya* ‘with me’ but *ma-maʿayīš* ‘not with me’.

2.4 *Morphology*

Cairo Arabic distinguishes between masculine and feminine only in the 2nd and 3rd persons singular. The verb shows strong synthetic tendencies, as it allows for simultaneous suffixing of direct and indirect object suffixes, as well as negational *-š*, as in *ma-biywarrihalhumš* ‘he does not show her to them’. On the other hand, there are analytical tendencies in the noun phrase, where the use of the genitive particle *bitāʿ* is very common.

2.4.1 Pronouns

2.4.1.1 Independent personal pronouns

Table 3. Independent personal pronouns

	3rd	2nd	1st
sg. masc.	*huwwa*	*inta*	*ana*
sg. fem.	*hiyya*	*inti*	
pl.	*humma*	*intu*	*iḥna*

2.4.1.2 Possessive/object suffixes

There are three series which differ according to the final segments of the word.

after -v

axū(h) ‘his brother’

-(h)	*-k*	*-ya*
-ha	*-ki*	
-hum	*-ku*	*-na*

after -C

xālu ‘his uncle’

-u	*-ak*	*-i*
-ha	*-ik*	
-hum	*-ku*	*-na*

after -CC

uxtu ‘his sister’

-u	*-ak*	*-i*
-aha	*-ik*	
-uhum	*-uku*	*-ina*

A socially lower variant of *-ku* is *-kum*. *-ku* developed from *-kum* by analogy to *-u* of the 3rd pers. pl. masc. of the perfect.

2.4.1.3 Indirect object suffixes

after -v

ʾalūlu ‘they said to him’

-lu	*-lak*	*-li*
-lha	*-lik*	
-lhum	*-lku*	*-lna*

after -C

ʾallu ‘he said to him’

-lu	*-lak*	*-li*
-laha	*-lik*	
-luhum	*-luku*	*-lina*

after -CC
ʾultilu ‘I said to him’
-ilu *-ilak* *-ili*
-ilha *-ilik*
-ilhum *-ilku* *-ilna*

2.4.1.4 Demonstratives

A particular Egyptian feature is the lack of initial *hā-* (Table 4).

Table 4. Demonstratives in Cairo Arabic

	near	far
sg. masc.	*da ~ dawwa*	*dukha*
sg. fem.	*di ~ diyya*	*dikha*
pl.	*dōl ~ dōla*	*dukham*

Extensions with *-t* or *-n* are very common: *dawwat*, *diyyat*, *diyyan*, *dōlan*, etc. *dukha* is mainly used for reference to an item earlier mentioned as in *dukha ʿandu flūs wi da ġalbān* ‘the other one has money and this one is poor’.

The attributive demonstrative follows its noun, e.g. *ilwalad da* ‘this boy’, *ilbinti̊ di* ‘this girl’, *ilmalik dukha* ‘the aforementioned king’. *da* and *di* can be added to the noun as clitics such as *ilḥalaʾā-di* ‘this earring’, *ilmaṛṛā-di* ‘this time’, *iššahrí-da* ‘this month’. After expressions of astonishment, joy, disgust, despair, etc., *da* and *di* may introduce the noun phrase as in *kulli̊ yōm ʿala da lḥāl* ‘every day the same situation!’, and *ya salām ʿala di ʾahwa* ‘wonderful, such a coffee!’ with an indefinite noun.

2.4.1.5 Presentatives

ahú, *ahí*, *ahúm* and *adīni*, *adīk ~ ad-inta*, etc. serve as presentatives: *ilmaṭbax ahú ʾuddāmak* ‘here is the kitchen in front of you!’, *adīni gēt* ‘here I am!’.

2.4.1.6 Interrogative pronouns

mīn ‘who?’, *ʾē* ‘what?’, which take the same position in the sentence as the item they ask for: *mīn ʾallak kida* ‘who told you?’, *akkallim maʿa mīn* ‘whom have I to talk with?’, *kānit hatiggawwiz mīn* ‘to whom was she to be married?’.

‘which?’: fronted *anhi ~ anhu* masc., *anhi* fem., *anhum* pl., postponed *anhū*, *anhī*, *anhūm*, e.g. *anhi srīr ~ issirīr anhū* ‘which bed?’.

2.4.2 Adverbs

Temporal: *imta* ‘when?’, *innahaṛda* ‘today’, *imbāriḥ* ‘yesterday’, *bukṛa* ‘tomorrow’, *baʿdi̊ bukṛa* ‘the day after tomorrow’, *awwil imbāriḥ* ‘the day before yesterday’, *dilwaʾti* ‘now’.

Local: *fēn* ‘where?’, *minēn* ‘from where?’, *hina ~ hinahó* ‘here’, *hināk* ‘there’.

Manner: *izzāy* ‘how?’ (< **ēš zāy*), *kida* ‘so, this way’, *ʾawi ~ xāliṣ* ‘very’.

Causal: *lē* ‘why?’, *ʿašān kida* ‘therefore’.

Number and mass: *kām* ‘how many?’, *ʾaddi̊ ʾē* ‘how much?’.

Interrogative adverbs take the same position in the sentence as the item they ask for: *sākit lē* ‘why are you silent?’.

2.4.3 Particles

2.4.3.1 Article

The article is *il-* ~ *ill-* ~ *l-*, the latter two variants being used with vowel-initial nouns such as *liyyām ~ illiyyām* ‘the days’.

2.4.3.2 Genitive

The genitive marker is *bitāʿ* masc., *bitāʿa* fem., *bitūʿ* pl. and may replace the construct state in complex noun phrases, except with inalienable body parts or family members as a head of the noun phrase: *ʿaṛabiyyiti ~ ilʿaṛabiyya btaʿti* ‘my car’, but always *abūya* ‘my father’.

2.4.3.3 Negation

Negational particles: *miš ~ muš* is mainly used with a nominal predicate, the *ḥa*-imperfect, and the active participle. *ma-* . . . *+š* negates other verbal predicates (perfect, *y*-imperfect, *bi*-imperfect), the predicate in prepositional sentences as in *ma-ʿandīš waʾt* ‘I do not have time’, and personal pronouns when these are subjects: *ma-ntāš ʿārif*? ‘do you not know?’. As to the *bi*-imperfect, there is a tendency nowadays to replace *ma-* . . . *+š by miš*: *ma-byiʿmilši̊ ḥāga ~ miš biyiʿmil ḥāga* ‘he does nothing’. *ma* is limited to assertions introduced by particles such as *waḷḷāhi* ‘by God’, *ʿumru* ‘never’, etc., as in *lā ma-tinzilš, wi nnabi ma tinzil* ‘no, do not go down, by the Prophet, you must not go down!’.

2.4.3.4 Questions

Questions may be introduced by *huwwa*, as in *huwwa ḥaḍritik maṣriyya* ‘are you [fem.] Egyptian?’, or reinforced by *ya-taṛa*, as in *mīn da ya-taṛa* ‘who could this possibly be?’.

2.4.3.5 Prepositions

Cairo Arabic shows the usual set of prepositions. Pronominal suffixes may be connected to some of them by means of *-ī-* (originating from suffixed *fi* ‘in’): *taḥtaha ~ taḥtīha* ‘under her’, *foʾha ~ foʾīha* ‘above/on top of her’.

2.4.3.6 Conjunctions

Besides the usual set of conjunctions, some new ones have developed by grammaticalization: *aḥsan* ~ *laḥsan* 'because, otherwise', *la-* 'lest, that', *illi* 'that, because', *madām* 'since [causal]': *rūḥ baʾa li lmaʿazīm aḥsan itʾaxxaṛt* 'go to the guests now because you are late!'; *xāyif la-ykūn fīha ʿafarīt* 'I fear that there are demons in it'; *kuwayyis illi* . . . 'it is a good thing that . . .'; *ana ġalṭān illi* . . . 'I am wrong in having done . . .'; *ṭabʿan madām ʿayyāna tibʾa la tiġsil wala tuṭbux* 'of course, since she is ill, she neither washes nor cooks'.

2.4.4 Noun

2.4.4.1 Gender

Feminine nouns without the marker *-a* are the usual ones, i.e. female persons, body parts, *šams* 'sun', *ʾarḍ* 'earth', etc., but others as well such as *markib* 'boat', *nāṛ* 'fire'. Others are in fact plurals, or look like such: *filūs* 'money', *manaxīr* 'nose', *baṭāṭis* 'potatoes', *ṭamāṭim* 'tomatoes'.

2.4.4.2 Productive patterns

For instruments muCCāC, CaCCāCa and CuCCēCa: *muftāḥ* 'key', *munfāx* 'bellows', *baxxāxa* ~ *buxxēxa* 'atomizer', *šaffāṭa* 'drinking straw', *nuffēxa* 'balloon'; for waste products CuCāCa: *kunāsa* 'sweepings', *nušāṛa* 'sawdust'. Apart from the common pattern CaCCāC, nouns denoting professionals can be formed from plural nouns by a *nisba*: *saʿāti* 'watchmaker', *muġanniyyāti* 'singer', *ganayni* 'gardener', or with the Turkish suffix *-gi* as in *kababgi* '*kabāb*-seller', *kawalingi* 'locksmith'. The suffix *-iyya* may be used for abstract notions: *ruguliyya* 'manlihood', *ṭarḥiyya* 'yield'.

2.4.4.3 Plural

Plural formation follows the usual rules of Arabic. There are some uncommon plural patterns such as CuCúCa (< *aCCuCa) and CiCíCa (< *aCCiCa): *bunúka* 'banks', *kurúta* 'cards', *ġiríba* 'ravens', *sibíta* 'baskets'.

2.4.4.4 Pseudo-dual

Only three nouns form a → pseudo-dual: *idēn* 'hands', *riglēn* 'feet', *ʿenēn* 'eyes'; these lose the *-n* when suffixed: *riglē*(*h*) 'his feet', *idayya* 'my hands'.

2.4.4.5 Diminutives

These are productive for nicknames only: *ḥammūda* 'Aḥmad', *faṭṭūma* 'Faṭma', *zannūba* 'Zēnab'. The suffix *-u* conveys a hypocoristic sense: *mīdu* 'dear Aḥmad', *xālu* ~ *ʿammu* 'dear uncle!'.

2.4.5 Numerals

1–2: *wāḥid* masc., *waḥda* fem. are used attributively like adjectives; *itnēn* is used mainly as a reinforcement of the dual: *saʿtēn itnēn* 'two hours'; *itnēn* + plural replaces the dual of nouns denoting persons as in *itnēn afandiyya* 'two effendis' (excluding family members).

3–10: long forms: *talāta*, *aṛbaʿa*, *xamsa*, *sitta*, *sabʿa*, *tamanya*, *tisʿa*, *ʿašaṛa*.

The corresponding short forms are *talat*, *aṛbaʿ*, *xamas*, *sitt*, *sabaʿ*, *taman*, *tisaʿ*, *ʿašaṛ*, which are used with counted nouns: *talat waraʾāt* 'three pages', *taman ʾasabīʿ* 'eight weeks'. The system of polarity, which distributed the masculine and feminine allomorphs of the numerals according to the gender of the noun, has been replaced by a system which makes use of the short form only, but prefixes a *t-*. Originally the feminine *-t* of the numeral, it precedes vowel-initial plural patterns aCCāC ~ iCCāC and aCCuC (with initial /a/ changing into /i/, or /u/ when /u/ follows, and a stress on the first syllable): *xámas t-idwāṛ* 'five floors', *sabaʿ t-irwāḥ* 'seven lives', *tisaʿ t-uṣṭur* 'nine lines'.

11–19 *ḥiḍāšaṛ*, *iṭnāšaṛ*. From 13 onwards the short form + *ṭāšaṛ*: *talaṭṭāšaṛ*, *aṛbaʿṭāšaṛ*, etc., except for *siṭṭāšaṛ* 'sixteen'.

100 *miyya*, *mitēn*, *tultumiyya*, *rubʿumiyya*, *xumsumiyya*, etc.

Ordinal numbers follow the usual pattern CāCiC from 2 through 10. From 11 onwards the cardinal number is used attributively: *ilqarn ilʿišrīn* 'the 20th century'.

2.4.6 Verb

2.4.6.1 Forms

The list of verbal forms contains one base form and nine derived forms (Table 5).

Form I: In general, CaCaC is the transitive pattern, CiCiC the intransitive, but there are many exceptions. In quite a few cases, the CaCaC pattern functions as a causative of

Table 5. Derived verbal forms in Cairo Arabic

I	II	III
ġasal/yiġsil ‘to wash’	*tabbil/yitabbil* ‘to season’	*ʾābil/yiʾābil* ‘to meet’
širib/yišṛab ‘to drink’	*naḍḍaf/yinaḍḍaf* ‘to clean’	
xuluṣ/yixlaṣ ‘to be finished’		
t-I	t-II (V)	t-III (VI)
itġasal/yitġisil ‘to be washed’	*ittabbil/yittabbil* ‘to be seasoned’	*itʾābil/yitʾābil* ‘to meet each other’
itšarab/yitširib to be drunk’	*itnaḍḍaf/yitnaḍḍaf* ‘to be cleaned’	
n-I (VII)-	*ista*- (X)	IX
inbasaṭ/yinbisiṭ ‘to enjoy oneself’	*istaġfil/yistaġfil* ‘to act the dimwit’	*ismaṛṛ/yismaṛṛ* ‘to get brown’
	istarxaṣ/yistarxaṣ ‘to consider cheap’	
t-I (VIII)		
iftakaṛ/yiftikir ‘to think’		

CiCiC: compare *tiʿib/yitʿab* ‘to become tired’ – *taʿab/yitʿib* ‘to make tired’, *ʿiyi/yiʿya* ‘to get sick’ – *ʿaya/yiʿyi* ‘to make sick’. An *i*-perfect usually has an *a*-imperfect, whereas the *a*-perfect is often found with an *i*- or *u*-imperfect, except when the final syllable ends with a back consonant (emphatic, pharyngeal), e.g. *katab/yiktib* ‘to write’, *šaḥan/yišḥin* ‘to load’, *ṭalab/yuṭlub* ‘to order’, but *kasaṛ/yiksaṛ* ‘to break’, *nadah/yindah* ‘to call’, *balaʿ/yiblaʿ* ‘to swallow’. There are quite a few exceptions to this rule. In some cases, a verb and its causative differ only by an *i*-imperfect (*Form IV) instead of an *u*-imperfect: *xaṛag/yuxṛug* ‘to leave’ – *xaṛag/yixrig* ‘to bring/take out’, *daxal/yidxul* ‘to enter’ – *daxal/yidxil* ‘to bring in’.

Form II is 2-allomorphic with /a/ or /i/ in the final syllable according to a phonological rule: with /a/ if the final syllable is formed with back consonants, otherwise /i/. Semantically, Form II expresses intensity and plurality, as well as factitivity and may be derived from any noun or verb: *ʾaṭaʿ* ‘to cut’ ⇒ *ʾaṭṭaʿ* ‘to cut into slices’, *niḍīf* ‘clean’ ⇒ *naḍḍaf* ‘to clean’. Form II verbs are not necessarily transitive, but may describe processes undergone by the subjects rather than an action: *zinix* ‘rancid’ ⇒ *zannax* ‘to become rancid’, *kirš* ‘belly’ ⇒ *kaṛṛaš* ‘to develop a belly’. Form II is productively used to derive new verbs from foreign loans: *hannig/yihannig* ‘to hang [computer]’.

Form III is 1-allomorphic with /i/ in the final syllable in both perfect and imperfect.

The t-Forms: Prefixing *it*- to the perfect bases II and III is a productive device for deriving the corresponding reflexive-passive t-II (V) and reciprocal t-III (VI) Forms. The vowels remain unchanged. It is also productive for Form I as a passive, but with a vowel change from /a/ (perfect) to /i/ (imperfect): *itġasal/yitġisil* ‘to be washed’. Forms with (prefixed) *in*- (VII) and (infixed) *-t*- (VIII) are limited to particular lexical items such as *inbasaṭ/yinbisiṭ* ‘to enjoy oneself’, *iftakaṛ/yiftikir* ‘to think’.

The *ista*-Form (X) is fairly common and shows the same distribution of /a/ and /i/ in the final syllable as Form II: *istarxaṣ/yistarxaṣ* ‘to consider cheap’, *istaġfil/yistaġfil* ‘to consider dimwitted’.

There are some cases of combinations of *ista*- with Forms II and III as in *istarayyaḥ* ~ *istirayyaḥ* ‘to rest’ and *istibārik* ‘to seek *baṛaka*’.

2.4.6.2 Inflection of imperfect and perfect

2.4.6.2.1 Imperfect

The base vowel can be /a/, /i/, or /u/, the prefix vowel is /i/ except for the 1st pers. sg., but may be /u/ if the base vowel is /u/ or /ū/. With regard to the prefixes for the 1st pers. sg. and pl. there is no paradigmatic leveling, i.e. they are of the type *aktib-niktib*.

Table 6. Inflection of imperfect in Cairo Arabic

	yišṛab ‘he drinks’		
	3rd	2nd	1st
sg. masc.	*yišṛab*	*tišṛab*	*ašṛab*
sg. fem.	*tišṛab*	*tišṛabi*	
pl.	*yišṛabu(m)*	*tišṛabu(m)*	*nišṛab*

Table 6 (*cont.*)

	yimsik ‘he seizes’		
	3rd	2nd	1st
sg. masc.	*yimsik*	*timsik*	*amsik*
sg. fem.	*timsik*	*timsiki*	
pl.	*yimsiku(m)*	*timsiku(m)*	*nimsik*

The aspectual prefix used for present and habitual is *bi*-. For future and intention it is *ḥa*- or *ha*-.

2.4.6.2.2 Perfect

As with the personal pronouns *(intu ~ intum*, *-ku ~ -kum)*, in the 3rd pers. pl. and the 2nd pers. pl. paradigmatic leveling led to the co-occurrence of *-u ~ -um*, and *-tu ~ -tum*, the forms with final *-m* being the socially lower variants. The suffix of the 3rd pers. sg. fem. is *-it* in all verbal classes.

Table 7. Inflection of perfect in Cairo Arabic

	ġasal ‘he washed’		
	3rd	2nd	1st
sg. masc.	*ġasal*	*ġasalt*	*ġasalt*
sg. fem.	*ġasalit*	*ġasalti*	
pl.	*ġasalu(m)*	*ġasaltu(m)*	*ġasalna*
	širib ‘he drank’		
	3rd	2nd	1st
sg. masc.	*širib*	*širibt*	*širibt*
sg. fem.	*širbit*	*širibti*	
pl.	*širbu(m)*	*širibtu(m)*	*širibna*

2.4.6.3 Participles

A distinction between active and passive participles occurs only in Form I: CāCiC vs. maCCūC. The participle of the verbal forms is derived by replacing the *yi*-prefix of the imperfect by *mi*-, which may be interpreted with transitive verbs as active or passive according to context: *inti mḥawwišālu kām* ‘how many did you [fem. sg.] save for him?’, *ilfilūs ilmiḥawwiša* ‘the saved money’. The passive participle of Form I serves as participle of Form t-I: *itbana* ‘to be built’, but *mabni* ‘built’. The participle of Form t-II (V) may be replaced by the participle of Form II: *mitrayyiš ~ mirayyiš* ‘rich’.

2.4.7 Weak verbs

2.4.7.1 Verbs II gem.

Verbs II gem. are inflected in the perfect like the *a*-type of IIIy, e.g. *ḥaṭṭēt* ‘I put’, etc. The active participle is *ḥāṭiṭ*.

2.4.7.2 Verbs Iʾ

The roots **ʾkl* and **ʾxḏ* are reduced in the perfect to *kal* ‘he ate’, and *xad* ‘he took’ respectively, which are conjugated like the strong verb with forms *kalit* ‘she ate’, *kalu*, *kalt*, etc. *ʾ left its trace in the imperfect prefix *yā*-: *yākul* ‘he eats’, *tākul*, *yaklu*, *ākul*, etc. The imperative is *kul*, *xud*, the active participle *wākil*, *wāxid*, the passive participle *mittākil*, *mittāxid*.

2.4.7.3 Verbs I*w*

These behave like strong verbs (e.g. *waṣal/yiwṣal* ‘to arrive’) except for *wiʾif/yuʾaf* ‘to stand’, *wiʾiʿ/yuʾaʿ* ‘to fall’, with a *yu*-prefix, and *ú*- in the imperative: *úʾaf*, fem. *úʾafi*.

2.4.7.4 Verbs II*w/y*

These follow the general pattern with *ʾām/yiʾūm*, *ʾumt* ‘to stand up’ and *bāʿ/yibīʿ*, *biʿt* ‘to sell’, the vowel of the 1st pers. sg. perfect displaying the same quality as the vowel of the imperfect. Form VIII follows this rule too, with *irtāḥ/yirtāḥ*, *irtaḥt*, whereas *ā*-imperfects of Form I choose either /u/ or /i/: *nām/yinām*, *nimt* ‘to sleep’, but *xāf/yixāf*, *xuft* ‘to fear’.

2.4.7.5 Verbs III*w/y*

These resemble the strong verb Form I with their *a*-type and *i*-type in the perfect. Final *-a* is always elided before vowel-initial suffixes, whereas final *-i* changes to *-y*- before *-it* of the 3rd pers. sg. fem. and *-u* of the 3rd pers. pl. The short paradigms of *rama/yirmi* ‘to throw’ and *nisi/yinsa* ‘to forget’ are given in Table 8.

Table 8. Inflection of verbs III *w/y*

(perfect)	3rd sg. masc.	*rama*	*nisi*
	3rd sg. fem.	*ramit*	*nisyit*
	3rd pl.	*ramu*	*nisyu*
	1st sg.	*ramēt*	*nisīt*
(imperfect)	3rd sg. masc.	*yirmi*	*yinsa*
	2nd sg. fem.	*tirmi*	*tinsi*
	3rd pl.	*yirmu*	*yinsu*

2.4.7.6 Irregular verbs

gih/yīgi ‘to come’ shows some irregularities. The 3rd pers. sg. masc. *gih* seems to be a former pausal form (see Blanc 1973–1974) with *ga* being preserved as a variant and an allomorph when suffixed, e.g. *gāni* 'he came to me'. In general, it follows the III*y* (*a*-type) in the perfect, e.g. *gēt* 'I came', and the 3rd pers. pl. *gum* with an allomorph *gu-* when suffixed, e.g. *ma-gūš* 'they did not come'. Remarkable is the 3rd pers. sg. fem. *gat*, cf. *ramit* 'she threw'. The prefix of the imperfect contains /ī/ as in *yīgi*, *tīgi*, etc. but 1st pers. sg. *āgi* 'I come'.

idda/*yiddi* 'to give' has an irregular base and its initial /i/ is elided when inflected for the imperfect: *yi+iddi* > *yiddi* 'he gives'. The perfect follows the *a*-type of III*y*, e.g. *iddit* 'she gave', *iddēt* 'I gave'.

2.4.8 Quadriradicals

Except for onomatopoeic verbs such as *nawnaw*/*yinawnaw* 'to mew', which always have /a/ in their final syllable, quadriradical verbs display the same distribution of /a/ and /i/ in the ultima as Form II, e.g. *daʾdaʾ* 'to crush', *taftif* 'to splutter [person]'. Reduplication as in these two examples often gives a diminutive and repetitive sense. Another means for forming quadriradical verbs is the insertion of consonants, which may occasionally lead to semi-reduplication as in *radam* ⇒ *dardim* 'to fill up with earth', *dalaʾ* ⇒ *daldaʾ* 'to spill'. Derivation from nouns, including affixal consonants, is productive, e.g. *wasṭan*/*yiwasṭan* 'to center', *itmanẓar*/*yitmanẓar* 'to show off'.

2.5 *Syntax*

2.5.1 Noun phrase

With nouns denoting human beings, specificity may be expressed by means of *wāḥid* as in *hiyya miggawwiza wāḥid muhandis* 'she is married to a [certain] engineer'. *kām* and *kaza* are used as quantifiers: *kām ʾirš* 'some money', *baṭāṭis ikkaza mīza* 'the potatoes with quite a few good characteristics'.

2.5.2 Verbal phrase

Perfect aspect and past tense are expressed by the perfect: *miši* 'he is gone'. Present tense is generally indicated by *bi*-imperfect or the active participle, depending on the lexical aspect of the verbal phrase: *biyzākir* 'he is studying ~ he studies [habitual]', *ana šayfu gayy* 'I see him coming'. *bi*-imperfect may express habituality with any verb: *bašūf kuwayyis* 'I see well'. The simple imperfect without prefixes is non-factual and has a modal meaning 'should, would, may': *yixušš* 'let him enter ~ he may enter'. Intention and future tense is expressed with a *ḥa-* ~ *ha-* prefix: *ḥafrumhum* 'I'll make minced meat of them!' (Sādāt in May 1971) which also may express 'being on the point of' as in *ḥassēt inni̊ nafūxi ḥayṭaršaʾ* 'I felt that my brain was on the point of bursting'. The active participle plays an important role within the verbal system, in its meaning depending on the lexical aspect of the verbal phrase (see Eisele 1999). It may be resultative as in *ana māḍi* 'I have signed', or express present tense as in *huwwa nāyim* 'he is sleeping', or future tense *huwwa gayyi̊ bukṛa* 'he will come tomorrow'.

The combination with *kān*, *yikūn* provides a second point of reference on the time axis: *fiʿlan kunti̊ nsītik* 'indeed, I had forgotten you' (pluperfect), *kān biyittāwib* 'he was yawning' (continuation in the past), *kān ḥayiggannin* 'he nearly went mad'. *kān* + imperative is used as a counterfactive and a desiderative in the past: *kunt istanna šwayya*! 'would that you had waited a little bit'.

Intensity, continuation, and durativity may be indicated by means of a preverb *ʿammāl* or by auxiliary verbs such as *ʾaʿad*/*yuʾʿud* or *fiḍil*/*yifḍal* 'to remain': *ʿammāl yiʿīd nafsu* 'he is continuously repeating himself', *ʾaʿad yidḥak* 'he kept on laughing', *fiḍil māši* 'he kept walking'. For repeated action *rigiʿ*, *yirgaʿ* is used: *rigʿit ṣiḥyit* 'she woke up again'.

2.5.3 Narrative style

Periphrastic *ṛāḥ* and *ʾām* are very common, less so *gih*. The verb either follows in the same tense/aspect or as an active participle: *ṣarraxt*, *ʾāmu ftakaṛu nn iḥna bniġraʾ* 'I screamed, so they thought that we were drowning', *ʾawwil ma yiṣfaṛṛ*, *niṛūḥ ḥaṭṭīn laban* 'as soon as it becomes yellowish, we add milk'.

2.5.4 Word order

In general, word order is SVO, if the subject is definite. With an indefinite subject it is VSO; with a more specific subject, e.g. with an adjective, SVO is possible: *nās kitīr gum saʾalūni* 'many people came and asked me'. For the interrogatives see 2.4.1.6. Grade adverbs always follow their head: *kuwayyis ʾawi* 'very good'.

2.5.5 Prepositional sentences

Prepositional sentences are not only formed with *ʿand*, *maʿa*, *li* ('to have'), but with other local prepositions as well: *ikkalām ma-minnūš* *fayda* ‘talking is worthless’, *da ẓẓāhir ʿalēha gginn ilʾaḥmaṛ* ‘apparently she has got the Red Demon on her’. A definite possessum is possible as in *ana ṛāgil liyya ʾuslūbi* ‘I am a man who has his style’.

2.5.6 Conditional sentences

These are introduced by *iza* or *law*: *iza ṭalabik hizzi dmāġik* ‘if he asks you, shake your head!’, *law šāfak hina tibʾa mṣība* ‘if he sees you here, this will be a catastrophy’. *kān* in the apodosis gives a counterfactual meaning: *law kān ʾalli kunti̊ ʾultilik* ‘if he had told me, I would have told you’.

2.5.7 ḥāl-sentences

The syndetic type *wi* + S + P often conveys a temporal meaning and behaves like a temporal adverbial sentence. It may follow sentence-initial temporal adverbs, as in *xamas t-iyyām wi huwwa ʿayyān* ‘it is five days already that he has been ill’, *min sāʿit inta ma xaragt wi hiyya bitsaṛṛax* ‘since you left, she has been screaming constantly’. Asyndetic *ḥāl*-sentences occur mainly after verbs of motion or perception (see Woidich 1991).

3. Lexicon

A number of lexical items characterize Cairene speech, for example *ʿāwiz* ‘wanting’, *ḥāga* ‘something’, *izzāy* ‘how’, *baʾa* ‘so, then, now’, *baḥr* ‘river, canal’, and many more.

Many a language left its traces in the lexicon of Cairo Arabic. First of all, the Coptic substrate must be mentioned here (→ Coptic loanwords). Other loans are Greek (*taṛabēẓa* ‘table’), Persian (*bass* ‘only’), Mamluk Turkish (*xašš* ‘to enter’), Ottoman Turkish (*kušk* ‘kiosk’), and later Italian (*istabēna* ‘agreed!’), French (*barbarēz* ‘windscreen’), and nowadays, of course, English (*lūri* ‘lorry’, *warša* ‘workshop’). Some older loans entered Egyptian Arabic via Ottoman Turkish, e.g. *ʾirš* ‘piaster’, a back formation from the plural *ʾurūš*, from Ottoman *kuruş* < Latin *grossus* ‘thick’.

Many terms adopted in the 19th century or later are gradually being replaced nowadays by Standard Arabic equivalents: *aksiswaṛāt* ⇒ *kamaliyyāt* ‘accessories’, *abukātu* ⇒ *muḥāmi* ‘lawyer’. On the other hand, new foreign terms enter with modern technology such as *yisayyif* ‘to save [on a computer]’, *yihannig* ‘to hang’ [computer]’, and *mubayl* ‘cellular phone’ which seems to be replacing today the Arabic term *maḥmūl*.

Longer foreign words are often transformed into a pattern CaC(C)aCv̄C/vCC such as *santarafīš* ‘centrifuge’, *garabuks* ‘gear box’, or into a existing pattern such as *gadūn* ‘handlebar’ from French *guidon* following the pattern CaCūC for instruments.

Bibliographical references

Blanc, Haim. 1973–1974. “La perte d’une forme pausale dans le parler arabe du Caire”. *Mélanges de l’Université Saint Joseph* 48.375–390.

——. 1974. “The *nekteb-nektebu* imperfect in a variety of Cairene Arabic”. *Israel Oriental Studies* 4.206–226.

——. 1981. “Egyptian Arabic in the seventeenth century: Notes on the Judeo-Arabic passages of *Darxe Noʿam* (Venice, 1697)”. *Studies in Judaism and Islam presented to Shlomo Dov Goitein*, ed. S. Morag, I. Ben-Ami, and N. Stillman, 185–202. Jerusalem: Magnes Press.

Brustad, Kirsten E. 2000. *The syntax of spoken Arabic: A comprehensive study of Moroccan, Egyptian, Syrian and Kuwaiti dialects*. Washington, D.C.: Georgetown University Press.

Eisele, John C. 1999. *Arabic verbs in time: Tense and aspect in Cairene Arabic*. Wiesbaden: O. Harrassowitz.

Elder, E.E. 1927. *Egyptian colloquial reader*. Cairo: American University in Cairo Press.

Haeri, Niloofar. 1996. *The sociolinguistic market of Cairo: Gender, class and education*. London: Kegan Paul International.

Harrell, Richard S. 1957. *The phonology of colloquial Egyptian Arabic*. New York: American Council of Learned Societies.

Hary, Benjamin. 1996. “The *ǧīm/gīm* in colloquial urban Egyptian Arabic”. *Israel Oriental Studies* 16.153–168.

Hassan, Motie Ibrahim. 1971. *In-nās wil-malik*. Copenhagen: Copenhagen University Publications Fund.

Hinds, Martin and El-Said Badawi. 1986. *A dictionary of Egyptian Arabic: Arabic–English*. Beirut: Librairie du Liban.

Mitchell, T.F. 1962. *Colloquial Arabic: The living language of Egypt*. London: Teach Yourself.

Prasse, Karl-G., Katrine Blandford, Elisabeth A. Moestrup, and Iman El-Shoubary. 2000. *5 Egyptian-Arabic one act plays: A first reader*. Copenhagen: Museum Tusculanum Press.

Prokosch, Erich. 1983. *Osmanisches Wortgut im Ägyptisch-Arabischen*. Berlin: W. Schwarz.

Rowson, Everett K. 1981. “Cant and argot in Cairo colloquial Arabic”. *Bulletin of the American Research Center in Egypt* 122.13–24.

Spiro, Socrates. 1923. *Arabic–English dictionary of the modern Arabic of Egypt*. 2nd ed. Cairo: Elias’ Modern Press.

Spitta-Bey, Wilhelm. 1880. *Grammatik des arabischen Vulgärdialectes von Aegypten*. Leipzig: J.C. Hinrich.

Tadié, Arlette. 1994. "Les Syriens en Égypte". *Actes des premières Journées internationales de dialectologie arabe de Paris*, ed. Dominique Caubet et Martine Vanhove, 481–492. Paris: Institut National des Langues et Cultures Orientales.
Wise, Hilary. 1975. *Transformational grammar of spoken Egyptian Arabic*. Oxford: Blackwell.
Woidich, Manfred. 1991. "Die Formtypen des Zustandssatzes im Kairenischen". *Zeitschrift für Arabische Linguistik* 23.66–98.
——. 1994. "Cairo Arabic and the Egyptian dialects". *Actes des premières Journées internationales de dialectologie arabe de Paris*, ed. Dominique Caubet and Martine Vanhove, 493–507. Paris: Institut National des Langues et Cultures Orientales.
——. 1996. "Rural dialects of Egypt". *Egypte/Monde arabe* 27–28.325–354. Cairo: Centre d'Etudes et de Documentation Economique, Juridique et Sociale.
——. 2006. *Kairenisch Arabisch: Eine Grammatik*. Wiesbaden: O. Harrassowitz.

MANFRED WOIDICH (University of Amsterdam)

Calligraphy → Script and Art

Cameroon Arabic

Literature on the state of Arabic in Cameroon is almost unavailable, if not non-existent. Several works that dwell on the description of Cameroonian languages have little or nothing to offer in this domain, not even the most recently published directory of Cameroonian languages by Bitja'a Kody (2003). Yet the presence of Arabic on Cameroonian soil as an indigenous, vehicular, and foreign language makes this language too important to be overlooked. This entry is therefore an attempt to present a linguistic picture of Arabic as it is known and used in Cameroon.

1. LANGUAGE SITUATION IN CAMEROON

Cameroon is a highly dense multilingual country in which 279 indigenous languages (cf. *Ethnologue* 2004) belonging to three different African language families (the Nilo-Saharan, the Afro-Asiatic, and the Niger-Congo) are spoken. Onto this already complex linguistic situation have been superimposed English and French, the two official languages of colonial heritage, and Cameroon Pidgin English, the leading lingua franca in the country. With a population of about 16,184,748 inhabitants (July 2002 estimate), Cameroon is by all standards a linguistic paradise in view of its sociolinguistic complexity.

Following the latest statistics from *Ethnologue* (2004), languages constituting the Afro-Asiatic group are 60 in number. They are spoken in the northern part of the country. The Nilo-Saharan family is made up of 2 languages, while the Niger-Congo family has 215 languages. Among the indigenous languages, 2 other languages (Bung and Luo) are unclassified, given that they do not fall under any of the above language families. There are also a further 9 languages of wider communication: Basaa, Cameroon Pidgin English, Chowa Arabic, Duala, Ewondo, Fulfulde, Hausa, Kanuri, and Wandala (Breton and Fohtung 1991:20). In this entry the main focus is on the state of Arabic in Cameroon. Consequently, emphasis is laid not only on Chowa Arabic (the local variety of Arabic used in North Cameroon; also spelled Shuwa) but also on Standard Arabic usage as observed within educational and Islamic circles.

2. THE ARABIC LANGUAGE IN CAMEROON: HISTORICAL OVERVIEW

Information on the origins of Arabic in Cameroon is scant, given the present state of scientific research. Baba (1998) contends that the presence of Arabic in Cameroon can be traced as far back as 666 C.E. In the 15th century, the Arabs were nomads between Lake Fitri and Bahr el-Ghazal in the Chowa country. At the beginning of the 18th century, they settled along the banks of Lake Chad. Thus, nomads who came from the east and settled in the area brought Chowa Arabic to the region (Breton and Fohtung 1991:101). According to Hagenbucher (1973:7), the presence of Arab populations south of Lake Chad was first observed in 1823. The Chowa Arabs are described in a testimony presented by Denham a.o. (1931) as an extraordinary race having little in common with the Arabs of the north. Their physiognomy is beautiful and open, their nose aquiline, their eyes wide open, and their body lightly bronzed. They are both cunning and courageous, and speak a type of Arabic that is almost purely Egyptian. 'Chowa' comes from the Arabic word *šāwiyya* 'sheep/cattle rearer' or 'nomad'. However, the Chowa Arabs do not accept the use of the word

'Chowa' (which they consider derogatory), and so simply refer to themselves as Arabs and their language as Arabic (Eldridge 1979).

The history of Arabic in Cameroon is closely linked to the history of Islam. This explains why Arabic in Cameroon is mostly spoken in regions where Islam is the dominant religion. From the time Islam was introduced in North Cameroon, → Fulfulde was used along with Arabic as the language of Islam.

3. Linguistic description of Cameroon Arabic

In Cameroon, two main varieties of Arabic can be distinguished: Literary or Standard Arabic and Chowa Arabic. Literary or Standard Arabic is the variety reserved for literature. It is used for the written medium, be it the *Qurʾān*, textbooks, or other forms of literary expression. This variety is used for formal education, and rarely serves day-to-day interaction. The variety known as Chowa Arabic (= Shuwa Arabic) is an indigenous Cameroonian language, which is presently not written, and serves as a language of wider communication in the northern part of the country.

3.1 *Speakers and regional distribution*

Chowa Arabic, which belongs to the Semitic language family, falls under the Afro-Asiatic phylum. Known variously as Shua Arabic, Shuwa Arabic, and Arabe Choa, it is spoken by 63,600 people in Cameroon (SIL 1982), 754,590 in → Chad (1993 census), 5,000 in Niger (1998), and 100,000 in Nigeria (SIL 1973) (→ Sub-Saharan Africa). Recent statistics advanced by Seignobos and Iyébi-Mandjek (2000:50) hold that the Chowa Arabs of Cameroon are 125,313 in number and they are spread throughout various Kotoko sultanates. In the Logone and Chari division, which is made up of 276,170 inhabitants, the Chowa Arabs form 42 percent of the population, the Kotokos 28 percent, the Bornouans 10.3 percent, and the Musgums 8.7 percent. Geographically, the Chowa Arabs are not limited to the Logone and Chari division, but extend to Mora, the foot of the Mandara mountains, as well as Petté and Bogo in the Diamaré division (Eldridge 1979).

Chowa Arabic is spoken in the Far North Province mostly between Lake Chad and Kousseri, with some pockets of speakers south of Kousseri. Among the divisions in which the language is spoken are Diamaré, Mayo-Danay, Mayo-Sava, and especially Logone and Chari. In the Logone and Chari division, it is not only a mother tongue but also a language of wider communication. In fact, in the rural areas north of the Logone and Chari division, it is the dominant mother tongue, while in the urban areas, it assumes vehicular functions (Breton and Fohtung 1991). Among the leading towns and villages in which it is spoken are: Kousseri, Makary, Goulfey, Logone Birni, Hilé Alifa, Maltam, Fotokol, Cigal, Afade, and Blangoua.

As a vehicular language of the heterogeneous populations of the Logone and Chari division, Chowa Arabic is understood and spoken by speakers of different Kotoko languages such as Mpade, Afade, Maslam, Mser, and Lagwan. As a language of wider communication, the influence of Arabic is felt to the north of the Waza park at latitude 11°30', covering the whole area of Kotoko north (Barreteau and Dieu 2000:68). To the east of the park in areas such as Djégéré and Tchédé, Arabic shares this vehicular function with Fulfulde. The Musgums of Pouss and Maga, who have lived in Kousseri and N'djamena or who have been in contact with Chowa Arabs, as well as the Kotokos of the south, speak Arabic. According to a study carried out by Adji (2002:31–32), Arabic as a language of wider communication is the most widely used language in the Kousseri area. Here, it represents 44.2 percent of daily communication in the neighborhoods, 78.4 percent of interaction in the market-place, 82.4 percent of communication in mosques, and 55.2 percent of communication in the rural areas during electoral campaigns.

3.2 *Sound system and writing*

The sound system of Chowa Arabic as expressed through its consonants and vowels is made up of 20 consonants, 5 vowels, and 2 diacritic signs, (ˆ) for long vowels and (-) for the assimilation of articles. The 20 symbols used to represent consonants are: *b, f, w, m, t, d, s, z, n, l, r, j* [ʒ], *c* [ʃ], *y, k, g, x, h, tc* [tʃ] and *ny* [ɲ]. The 5 symbols used to represent vowels are: a, e, i, o, and u. The consonant and vowel charts of Chowa Arabic are given in Tables 1 and 2.

It should be noted that in Chowa Arabic vowel length in terms of shortness or length determines word meaning. Here are some examples:

Table 1. Inventory of consonants in Cameroon Arabic (adapted from Julien de Pommerol 1997)

		Labial	Apical	Palatal	Velar	Glottal
Occlusive	Voiceless	p	t	tc [tʃ]	k	ʔ
	Voiced	b	d	j [ʒ]	g	
Fricatives	Voiceless	f	s	c [ʃ]	x	h
	Voiced		z			
Nasals	m	n		ny [ɲ]		
Vibrant		r				
Liquids	w	l	y			

Table 2. Inventory of vowels in Cameroon Arabic (adapted from Adji 2002:53)

Back		Center		Front	
Short	Long	Short	Long	Short	Long
i	i:, ī			u	u:, ū
e	e:, ē			o	o:, ō
		a	a:, â		

asala 'python' vs. *āsala* 'sweet thirst'
sel 'ladder' vs. *sēl* 'flood'

Given the present language contact situation in the country, Chowa Arabic, like other languages, has been subjected to linguistic borrowing, loan translation, and interference from other languages in place such as French. Conversely, Chowa Arabic equally exerts some linguistic influences on other languages. Below are examples of various influences of French on Chowa Arabic:

(1) Loans from French (cf. Adji 2002:60–63) *daktor farmasi* < French *docteur pharmacie* 'doctor pharmacist' or 'pharmacist'; a pediatrician is referred to in the language as *daktor hanā liyāl* 'doctor for children'; *samanti* < French *ciment* 'ciment'
(2) Calques from French (cf. Adji 2002:63) *fi dizōrdir* < French *faire le désordre* 'disorderly'; *gandul sigrēt* < French *bâton de cigarette* 'cigarette [stick]'; *gazazt-al biyēr* < French *bouteille de bière* 'bottle of beer'
(3) Lexical interference from French (cf. Adji 2002:66)
   a. *gurus māfi fi* pays *hanana*
   money not in country for us
   'There is no money in the country for us'/ 'Our country is poor'
   b. *Yom* dimanche *ma talga nadum fil xidime*
   The day Sunday not find someone at work
   'There is no one at work on Sunday'

In the two Chowa Arabic utterances (3a) and (3b), the speakers unconsciously use the French words *pays* 'country' and *dimanche* 'Sunday' in their speech. The influence of French on Arabic is not surprising given that French is not only an official language in the area, but also serves as a vehicular language between Francophones and Arabophones.

4. Teaching and learning of Standard Arabic

In North Cameroon, Arabic is taught in Qurʾānic schools using Fulfulde (the local lingua franca) as the medium of instruction (cf. Abdkrm 2000). Unfortunately, the teaching methods employed in these schools are so harsh and inhuman that children end up barely knowing how to read and write Arabic without at all understanding what is being read or written. In short, the pupils end up not mastering the Arabic language itself (Santerre 1973; Adama and Amadou 1998). This is true, too, of the traditional teaching set-up which consists of two levels: the elementary school and the post-elementary school.

At the elementary level, the instructor, referred to as *mallum* (< Arabic *muʿallim*) teaches the children how to read and write Arabic. Children eligible for registration into the school should be at least five years of age. In a typical elementary school, children of different levels sit together in the same classroom, but receive different lessons depending on their level. The elementary cycle generally lasts for three to five years, and the children are expected to have mastered some of the following courses: reading, writing, letters of the alphabet, memorization of some verses of the *Qurʾān*, moral instruction, and manual labor. Since emphasis at this level is on reading and writing, the method commonly used by the instructor is the 'Baghdad' method, an eclectic method of teaching believed to be highly efficient from the pedagogic point of view.

At the post-elementary level, admission is reserved for children who are at least ten years of age. But in principle, most of those who attend the classes are adults. The instructor at this level is referred to as *moodibo* (< Arabic *mu'addib*). In terms of course content, the main subjects taught are Muslim law *(fiqh)*, grammar *(naḥw)*, theodicy *(tawḥīd)*, and *Qurʾān* exegesis *(tafsīr)*. Others include translation, vocabulary, rhetoric, stylistics, pedagogy, and the biography of the Prophet. In these schools, didactic material varies enormously from one teacher to another and from one region to another. Emphasis is laid on writing, reading, memorization, and understanding of texts. Generally, the teachers use the translation method by rendering material in the textbooks or manuals from Arabic into Fulfulde. Thus, the use of Fulfulde in the teaching of Arabic is certainly detrimental to the learning of this language.

Apart from the traditional approach, there is also the modern approach to the teaching of Arabic, which has developed considerably since independence with the assistance of Arabic-speaking countries of North Africa and the Middle East. Today, Arabic is taught in Cameroon not only at Franco-Arabic and Anglo-Arabic bilingual schools, but also at teacher training colleges, government secondary schools, and private language training centers.

Franco-Arabic primary schools were created in 1936 during the French colonial period in Cameroon. The colonial administration put in place a type of local school system (popularly referred to in French as *école de village* 'rural school') in some localities in North Cameroon such as Maroua, Ngaoundéré, and Kousseri (Adama 2001:91). In these schools, covering the first four years of primary education, two teachers taught each class. The French teacher taught during the morning period while the Arabic teacher taught in the afternoon. The latter based his teaching exclusively on the *Qurʾān*. Later on, a regional primary school (known in French as *école régionale*) was opened in Garoua, to which were added the last two classes of primary education (*Cours moyen 1* and *Cours moyen 2*). Thus, the Garoua school provided the only avenue for the pupils to complete their primary education and obtain the *Certificat d'études primaires élémentaires*, the French equivalent of the Anglo-Saxon First School Leaving Certificate. Unfortunately, Arabic was not taught in the Garoua regional school, a situation that created much discontent among the Muslim community, who felt that they were being marginalized by the French colonial administration.

In 1963, the Cameroon government created the Islamic Cultural Association of Cameroon, better known as L'Association culturelle islamique du Cameroun (ACIC). This association, which became official under Law No. 67/LF/19 of 12 June 1967, was responsible for the administration and management of Franco-Arabic and Anglo-Arabic schools throughout the national territory. Among the main objectives of the association were the development of Franco-Arabic and Anglo-Arabic education on the one hand, and the dissemination of Muslim doctrine on the other. In order to accomplish this task, the country was divided into three administrative zones, referred to as secretariats of education. The secretariat of education for the North zone had its headquarters in Garoua and was headed by Abdourahman Abdelkarim; the secretariat of education for the West zone had its headquarters in Kumba and was headed by Ahmed Ibrahim Nzube Epie; and the secretariat of education for the South zone had its headquarters in Yaounde and was headed by Mohamadou Garba (cf. Adama 2001:89).

However, a major problem persisted – that of the lack of availability of qualified personnel needed to teach the Arabic language or use Arabic as a medium of instruction. This led the associations of several towns in the northern part of the country to solicit aid from Saudi Arabia in the training of teachers. The Saudi authorities in Yaounde acceded to this request, and so, selected Cameroonians were sent to Saudi Arabia for further studies. As the trained Cameroonians returned home in the late 1960s and early 1970s, the Franco-Arabic and Anglo-Arabic schools could now boast a real breed of qualified personnel. These new young teachers also brought along with them new teaching techniques and methods radically different from those previously employed by the former *Qurʾān* instructors. Eventually, the new teachers assumed management of the schools and their dynamism greatly contributed to the growth of Franco-Arabic and Anglo-Arabic schools throughout the national territory. In 1965, in North Cameroon there were Franco-Arabic schools in localities such as Garoua, Maroua, Ngaoundéré, Rey Bouba, Kousseri, Kalfou, Banyo, and Makary; and the southern part of the country had altogether 18 schools with a total population of 2,252 pupils (Adama 2001:100). This number increased rapidly as the years passed by. For instance, in 1992, the West zone had 34 officially recognized Franco-Arabic primary schools with a population of 5,022 pupils (3,006 boys and 2,016 girls) spread in 154 classrooms and taught by 188 teachers (Adama 2001:96). In the South zone, made up of the Center, Littoral, South and East Provinces, Franco-Arabic schools are functional only in the Center and Littoral provinces; the South and East provinces have yet to host a single Franco-Arabic or Anglo-Arabic school.

Apart from the role of international cooperation, the local elite also plays a leading role in the training of Arabic teachers. This is the case of Aladji kouotto Malam Atam in the West province and Malam Innua Wirba in the Northwest province. Until his death in 1977, Aladji kouotto Malam Atam was highly instrumental in the development of Franco-Arabic education. He personally contributed immensely to the creation of the first Franco-Arabic private secondary school, opened in Foumban. In this school, Arabic is taught as a foreign language, just like German and Spanish.

Today, most Arabic teachers in the country have been trained in Saudi Arabia, Egypt, or Nigeria. Those trained in Nigeria receive their education in cities such as Yola, Sokoto, Kano, Maiduguri, and Zaria. Notwithstanding this, the problem of qualified teaching personnel continues to be a serious obstacle to the growth of Arabic education in the country. In 1990, for example, the twelve authorized Franco-Arabic schools in the northern part of the country had a total of 50 classrooms and 2,794 pupils; yet the number of qualified teachers who taught in these schools was just 11 as against 76 unqualified teachers – those not officially authorized by the Ministry of National Education to teach because they were found wanting in terms of professional qualification (Adama 2001:107). In terms of qualified personnel, the development of the Arabic language has been relatively slow in the Cameroonian context. This probably explains why, between 1987 and 1993, the national inspector of Arabic in the Ministry of National Education was a Moroccan national, Satifa Mohamed, who was in reality in charge of cultural affairs at the Moroccan Embassy in Yaounde.

From the pedagogic point of view, some important changes took place within the system in the 1970s. Much emphasis was now placed on the teaching of the Arabic language, rather than religious education as was the case before. Another serious development that characterized the Cameroonian school system was the institution of the teaching of Arabic as a foreign language in secondary schools in the north of the country. Some of the localities concerned are Maroua, Garoua, Ngaoundéré, Kousseri, and Mora. Thus, students who choose Arabic as their foreign language can sit for secondary school official examinations of the French system of education such as the *Brevet d'études du premier cycle*, the *Probatoire*, and the *Baccalauréat* examinations where they major in the language. In the southern part of the country, there are no government secondary schools where Arabic is taught as a foreign language. However, during the 1991/1992 academic year, a private secondary school known as Institut pédagogique polyvalent Bella Assan was opened in Yaounde. An Arabic section was formed, featuring subjects taught in Arabic such as mathematics, history, the *Qurʾān*, geography, unicity, science, anatomy, tradition, law, reading, and the biography of the Prophet.

In terms of course content, the weekly program of Franco-Arabic schools for teaching in Arabic covers the following subjects: the

Qur'ān; reading of the *Qur'ān (tajwīd)*; grammar *(naḥw)*; writing and dictation *(qirā'a wa-'imlā')*; theodicy *(tawḥīd)*; composition *(kitāba)*; Islamic law *(fiqh)*; mathematics *(ḥisāb)*; biography of the Prophet *(sīra)* (Adama 2001:116). It should be noted that this program is applicable as from the second year of primary education, while the first year is reserved exclusively for the teaching of the Arabic alphabet and basic aspects of Arabic grammar.

Government yearly subvention has also been instrumental in the development of Franco-Arabic private schools in particular, and Islamic education in general. In the northern part of the country alone, annual state subsidies stand at about 50,000,000 FCFA (approximately U.S. $100,000), an amount that is divided among the three provinces in the following manner: Far North Province 43.75 percent; North Province 28.02 percent; and Adamawa Province 28.23 percent (Adama 2001:108).

At the secondary level, Arabic was introduced as a foreign language in government secondary schools in the 1980s. At the tertiary level, the 1990s saw the introduction of Arabic as a minor in some Cameroon State universities such as the University of Ngaoundéré, the University of Yaounde I, and the University of Douala. Recently, the University of Ngaoundéré has instituted Arabic as a Bachelors degree course in the curriculum. In large towns and cities, several private language training centers in which Arabic is taught have been opened. Some of these centers offer evening classes where people learn how to read, write, and speak Arabic. Unlike in the past, those who are involved in the learning of Arabic nowadays are not only Muslims but also people from other religions who badly need the language for commercial purposes, for tourism, for diplomatic reasons, or for personal educational requirements.

In spite of these efforts, the teaching of Arabic and its use as a medium of instruction in Cameroon has always faced opposition. This was especially true during the French colonial administration, as shown in a circular of 30 August 1910 which appeared in *L'Afrique française*. Its author, William Ponty, argues seriously against Islamic education in the following terms: "Everyone knows that the study of French is the best means to use against religious fanaticism and experience has shown that Muslims who speak our language cause us less prejudice than those who understand only Arabic" (cf. Adama and Amadou 1998:7). There is no doubt that the French colonialists encouraged the use of French rather than Arabic as a medium of instruction. Today, the situation is not very different, given that the official language policy in the country encourages exclusively the promotion of English and French.

5. CONCLUSION

Cameroon Arabic has so far received little or no attention from policy-makers, educational experts, linguists, or researchers. While the teaching of Standard Arabic is somehow being encouraged nowadays by various actors, Chowa Arabic is yet to receive the attention it deserves. Arabic would normally occupy a prominent place in the school system in a country with a substantial number of Arabophones who have come into contact with the language through Islam, but at the pedagogic level, government involvement in Arabic education as well as the teaching of Arabic has all along been very timid. Second, given that Cameroon is part of the Arab League where Arabic is the main language of communication, the promotion of Arabic at the national level should be an important component of language policy, with Arabic being taught as a foreign language to a much greater extent in the school system than it is at present. The teaching of Arabic as a discipline in the Cameroon university system should be reinforced, enabling graduates holding degrees in Arabic to serve as potential teachers of Arabic or diplomats.

In this regard, the Cameroon government should take serious steps in the training of qualified teaching personnel both at home and abroad. Such a measure will in the long run restore Arabic to its much-deserved place in Cameroon not only as a lingua franca but also as a foreign language. Finally, it will be necessary to standardize Chowa Arabic, develop its writing system, and encourage linguistic research on the language. Such endeavors will go a long way to guarantee its survival as an indigenous Cameroonian language.

BIBLIOGRAPHICAL REFERENCES

Abdkrm, Abbo Yerima. 2000. *The teaching of Arabic through Fulfulde in North Cameroon from 1960 to 1984*. D.E.A. diss., University of King Faysal, Sudan.
Adama, Hamadou. 2001. "L'école franco-arabe camerounaise: Bilan et perspectives". *Cameroun 2001: Politique, langues, économie et santé*, 89–135. Paris: l'Harmattan.
—— and Aboubakary Moodiba Amadou. 1998. "Itinéraires d'acquisition du savoir arabo-islamique dans le nord-Cameroun". *Annales de la Faculté des Arts, Lettres et Sciences Humaines de l'Université de Ngaoundéré* 3.5–38.
Adji, Haoua. 2002. *Le français en contact avec l'arabe véhiculaire au Cameroun et au Tchad*. D.E.A. diss., University of Ngaoundere, Cameroon.
Baba, Mamoudou. 1989. *The Arabic language in secondary schools: A case study of North Cameroon*. M.A. diss., Faculty of Education and Social Sciences, University of Ifriqiyya al-Aalamiyya, Sudan.
Barreteau, Daniel and Michel Dieu. 2000. "Linguistique". Seignobos and Iyébi-Mandjek (2000:64–70).
Bitja'a Kody, Zachée Denis. 2003. *Annuaire des langues du Cameroun*. Yaounde: Editions du CERDOTOLA.
Breton, Roland and Bikia Fohtung. 1991. *Atlas administratif des langues nationales du Cameroun*. Yaounde and Paris: CERDOTOLA/CREA; ACCT.
Denham, Dixon a.o. 1931. *Narratives of travels and discoveries in Northern and Central Africa in the years 1822, 1823 and 1824*. London: J. Murray.
Eldridge, Mohamadou. 1979. *Notes sur les arabes choa*. Unpublished manuscript.
Ethnologue. 2004. <http://www.ethnologue.com>. [Website consulted on 12 March 2004.]
Hagenbucher, F. 1973. *Les arabes dits "suwa" du Nord-Cameroun*. Yaounde: ORSTOM.
Julien de Pommerol, Patrice. 1997. *L'arabe tchadien: Emergence d'une langue véhiculaire*. Paris: Karthala.
Santerre, Renaud. 1973. *Pédagogie musulmane d'Afrique noire: L'école coranique peule du Cameroun*. Montreal: Presses de l'Université de Montreal.
Seignobos, Christian and Olivier Iyébi-Mandjek (eds.). 2000. *Atlas de la province extrême-nord Cameroun*. Paris: Editions de l'Institut de Recherche pour le Développement.

GEORGE ECHU AND MOHAMADOU AMINOU
(University of Yaounde I)

Caretaker Talk

1. DEFINITION

The term 'caretaker talk' is used in the areas of developmental psycholinguistics and first language acquisition to describe the language adults use to address young children, and sometimes that of an older to a younger child. It is a register distinct from adult-adult speech (Snow 1986) and describes the modifications found in adult-child speech. It is also known as 'child directed speech' and as 'caregiver' or 'caretaker language'. Some researchers have used 'motherese' to refer to the language used by mothers, as if only mothers use a special way of addressing children. Others have used the fairly general 'parentese', since fathers also adopt this speech style when talking to young children.

Adults speak differently to children depending on the age of the child. Al-Shatti (2003) notices that Kuwaiti mothers and other caregivers, including 'housemaids' (al-Shatti's term), adjust their speech to children according to the child's age, reflecting the mother's response to the developing needs of the child. These adjustments are made to match the child's linguistic level. If adults use very complex structures, children are inattentive and fail to follow the adults' requests or even respond to their questions. This tendency is consistent with studies based on other languages as well. Adults simplify their speech and gradually increase its complexity to match the child's linguistic ability (Snow 1986).

'Baby talk' is a subsystem of caretaker language. The term is often preferred to 'motherese', and the phenomenon is widespread. Baby talk is considered a universal phenomenon, but cross-cultural differences are demonstrated as well (Ferguson 2004). Researchers argue that baby talk has a psychological function to show affection and syntactic simplification. It serves as a transitional period before children start using more adult-like language (Steinberg 1993) and is also used by older children at about three years to address younger children (Ferguson 2004; Steinberg 1999).

Most parents are aware that they modify their speech to children, while few feel that they do not. Haggan (2002) has studied the speech of Kuwaiti parents who deny modifying their language to children and found that their self-perception was valid and baby talk words were rarely produced. However, the language that these parents used showed some of the universal properties of caretaker language to be discussed below.

Adults who deny using baby talk find themselves using it in talking to young children or to a pet animal. Hirsh-Paek and Treiman (1980) argue that there are similarities between the way that adults talk to dogs, which is called 'doggerel', and motherese. (Doggerel is high-pitched and repetitious, with both questions and answers supplied by the speaker.) This implies that motherese is not modified because of the learner's linguistic and cognitive level. Hirsh-Paek and Treiman (1980) suggest that motherese may occur for social reasons and may be considered a marker of affection.

2. MODIFICATIONS OF CARETAKER TALK

A child's input is tailored to the needs of the child and is associated with the process of language development. An overwhelming number of studies on caretaker language describe universal properties of adult-child language. Ferguson (2004) categorizes modifications adults make when talking to children into prosodic, syntactic and grammatical, lexical, phonological, and discoursal. He argues that baby talk assists in the course of language development, in the development of interaction patterns, and the transmission of cultural values, and that it also shows the addresser's relationship to the addressee.

2.1 *Prosodic modification*

Caretaker talk tends to be produced with exaggerated intonation, slowly and in high pitch (Snow 1995). The presence of these features has been confirmed for Kuwaiti Arabic (Haggan 2002), Syrian Arabic (Ferguson 1956, 2004), and Egyptian Arabic (Omar 1973). Exaggerations serve to direct the child's focus on important sentence constituents and to hold their attention (Steinberg 1993). Further, prosodic characteristics make word and phrase boundaries clear to the child and consequently facilitate the child's decoding of the lexicon and grammar (Gleason and Ratner 1998).

Ferguson (2004) considers tone of voice, as in exaggerated intonation and high pitch, to be a universal feature of baby talk. He states three reasons for this type of behavior. Firstly, it could be an imitative act by the adults of what the infants produce because of the physical properties of the children's immature vocal cords. Secondly, children can discern pitch differences and prefer high pitch, which is why it is used in speech directed to children to get their attention. Thirdly, this type of speech may highlight certain grammatical properties and linguistic structures, possibly assisting in the development and comprehension of language. In addition, dysfluencies appear to be few and longer pauses between utterances are common in baby talk (Ferguson 2004).

2.2 *Syntactic and grammatical modification*

Caretaker language is characterized by short sentences, fewer subordinate clauses, and fewer grammatical relations than found in adult-adult speech, and by repetition, as well as omission of function words, inflectional endings, and the verb 'to be' (Ferguson 2004; Hirsh-Pasek and Treiman 1980; Snow 1995). The grammatical complexity of caretaker talk was studied in terms of mean length of utterances, the use of subordinate clauses, mean preverb length, and the use of conjunctions. It appears that caretaker language is simpler and more grammatical than speech addressed to adults. Questions, declaratives, and imperatives are frequently used (Snow 1986). These characteristics are also found in Arabic baby talk (Ferguson 1956; Haggan 2002; Omar 1973). There are no inflections (prefix or suffix) in baby talk words, nor do plural or gender agreement markers appear in baby talk (Ferguson 1956; Omar 1973). Nouns may take the definite article such as *lbuʿbuʿ* for an imaginary being mentioned to frighten children, and *ttišš*, for walk or ride, which is the only inflectional affix that is commonly used with baby talk words (Ferguson 1956).

Kuwaiti mothers show a preference for simple sentences compared to compound or complex sentences (Al-Shatti 2003; Haggan 2002). Al-Shatti (2003) notices an increase in compound and complex utterances in Kuwaiti mothers' speech to their young children at around 14 months. Kuwaiti mothers also show a preference for nouns over verbs. There was no significant difference between mothers and housemaids in their use of verbs; however, the mothers produced a considerably greater number of nouns than the housemaids did. Al-Shatti argues that nouns are important at this stage of language development and that mothers focus on increasing their children's vocabulary. On the other hand, the child-directed speech of Italian mothers is more verb-biased. Mothers talking to their children in their second year of life highlight verbs more than nouns, and morphologically mark verb stems more than they do noun stems. Camaioni and Longobardi (2001) argue that such emphasis on verbs leads to a verb-oriented pattern of acquisition.

Pronouns were used more by mother than by the housemaids; and verbal prescriptive structures such as 'tell me what you did' were very low in general (Al-Shatti 2003). Haggan (2002) noticed the appearance of non-Arabic words in the caregiver's language, which she attributed to the presence of non-Arabic speaking nannies.

2.3 *Lexical modification*

The number of words used in baby talk is small, and they belong to semantic areas related to body parts and functions, children's daily experiences, animals, food, and infant games (Ferguson 2004; Haggan 2002; Omar 1973). Most words are mainly used as imperatives or adjectives, but many are used as nouns.

Most of the baby talk words reported in the literature come from Egyptian Arabic and Syrian Arabic (Ferguson 1956; Omar 1973). They often refer to the same things, but differ in form. Omar (1973) argues that none of the baby talk words she found in Egyptian Arabic resemble their corresponding adult forms. Baby talk words could be used alone or incorporated into an adult sentence, where these words substitute the standard word. The following are examples of baby talk words used in Syrian and Egyptian Arabic respectively: *kixx* and *kuxx* 'dirt, forbidden,' *bīs/bīse* and *bisbis/bissa* 'cat', and *buff* 'bread, food'.

Adults simplify their vocabulary and label items by substituting hypocoristic variation, such as calling a train *choo-choo* on the basis of the sounds it makes, and diminutive formations, often formed by adding a suffix to the word, as in *birdie* for 'bird' and *horsie* for 'horse.' These types of formations seem to be universal (Ferguson 1956, 2004), but individual families may also create their own words. On the other hand, Omar (1973) did not find diminutives in caretaker speech.

Ferguson (1956) identifies hypocoristic variations in Syrian Arabic. These include *kūku* for bird, *tūt* for train (representing train whistle), *nawnaw* for cat, *tiktik* for watch and clock, *dindin* for bell (reflecting its sound), and *čōcčo* for horse, mule, and donkey. He also shows ways of creating nicknames and pet names in Syrian Arabic which are often used in baby talk. One is the hypocoristic suffix /-o/ which is added to a name to create a nickname, for example, *ḥamd-o* for *ḥamīd* or *ʾaḥmad* and *xayy-o* for 'brother'. The other is the application of the reduplicative patterns CuuCu and CiiCi, which are identical in form with a frequent pattern in baby talk (see below), for example, *fūfu* for *fuʾād*. There is no example of CiiCi in the study.

Hayes and Ahrens (1988) examine adult-child speech across the age range of newborns to 12-year-old children in order to see whether adults adjust their lexical choices to the child's age. They found that adults' words to children tend to be much more common compared to those used in adult-adult speech. They provide two reasons for this. Firstly, adult conversation with children is considered ordinary and focused on everyday topics such as household, school, and interpersonal matters. Because of such ordinary speech, common words are used rather than uncommon ones. Secondly, research on word retrieval has shown that it is quicker to retrieve frequently used, common words from memory than less frequently used, uncommon words (Just and Carpenter 1984 cited in Hayes and Ahrens 1988).

2.4 *Phonological modification or simplification*

Baby talk is characterized by consonant cluster simplification, extreme use of reduplication, consonant and vowel harmony, and preference for certain sounds. Studies show (Ferguson 2004; Omar 1973) that there is a tendency to maximize the use of certain sounds. Children omit more complex sounds or replace them with less marked ones. For example, they substitute stops for fricatives, semivowels for liquids, and single sounds for consonant clusters (Ferguson 2004).

Omar (1973) found that geminated consonants are extremely frequent, although they are not acquired early by children. Most of the phonemes are stops and nasals. The earliest phonemes acquired between the ages of 1;0 (one year and no months) and 16 years are the bilabial stop and nasal, pharyngeal consonants, semivowels and primary vowels, such as /a/, /i/ and /u/; some more difficult phonemes are also present, such as /ḥ/ and /x/. Dental stops, sibilants and velar consonants are acquired at the age of 2;0, while fricatives, laterals, and the rest of the vowels are acquired at 2;6 years. None of the emphatic phonemes occur, nor do the difficult phonemes /r/ and /š/ at this point. Similar observations are made by Ferguson (1956) in his study of Syrian Arabic baby talk, in which he observes that phonemes like /b/ /d/ /k/ /n/ are very frequent (no age is given in this study). Velarized phonemes /ḍ/ /ṭ/ /ṣ/ /ẓ/ /ḷ/ /ṛ/, the liquids /l/ and /r/, except the word *kurr* for 'noise said in baby's ear', and the voiced velar stop /g/ (in Egyptian Arabic) are rare or missing in baby talk. /ʕ/ was not attested in Egyptian Arabic data, but it was in Ferguson's data.

Three unexpected results were found in Ferguson's study (1956). First is the frequent use of the pharyngeals /ʕ/ and /ḥ/ in baby talk. It seems that not only are these two phonemes used frequently in Arabic, but they are learned early by children. Second, the velarized labial phonemes /ṃ/ and /ḅ/ are frequent in baby talk although they are rare in adult Arabic. These phonemes occur in the most frequently used forms /ḅāḅa/ for father and /ṃāṃa/ for mother. Third, the study shows the complete absence of the phoneme /ẓ/ in baby talk, although this phoneme is not rare in Arabic. Its voiceless counterpart /ṣ/ does occur, but not often, in baby talk.

The tendency was also found toward maximizing certain canonical syllables and forms such as CVCV. Consonant clusters are avoided except in one form *mbu* for water, a form found in both Syrian Arabic and Egyptian Arabic (Ferguson 1956; Omar 1973). Duplication is universally found in baby talk as illustrated by Arabic *wawa* for water and English *choo-choo* for train. The pattern CVVCV is frequent in baby talk, as in *nūnu* for 'small' and *tēte* for 'grandmother' (Ferguson 1956). The repetition of a word, especially monosyllabic words, such as *daḥḥ daḥḥ* 'pretty, nice, good, clean', is also found in Syrian Arabic baby talk.

2.5 *Discoursal modification*

Certain sentence types are frequently used in baby talk. These include questions, tags, imperatives, and here-and-now sentences (Ferguson 2004; Hirsh-Pasek and Treiman 1980). Al-Shatti (2003) claims that direct imperatives are frequent, which is typical of the mothers of young children. As children get older, their mothers start to use fewer directives and speak more indirectly. Pronoun shift is also common, 3rd person being used for 1st and 2nd person (Ferguson 2004; Hayes and Ahrens 1987), and 1st person plural for 2nd person singular (Ferguson 2004). Recasts such as *ḥamūdi wēnah* and *wēnah ḥamūdi* for 'where is Hamoudi?' are found (Haggan 2002).

Haggan (2002) indicates that in Kuwaiti Arabic words for 'mother' and 'father' are *mama* and *baba*, respectively. Mothers use the word *mama* when they are addressing their children, irrespective of the child's gender. Fathers use the word *baba* when they are addressing their children, irrespective of the child's gender (Haggan 2002; Omar 1973). Haggan (2002) argues that this is a way of showing affection to children.

Finally, the role of caretaker talk in → first language acquisition should be mentioned. Chomsky's theory of Universal Grammar (1986, 1988) revolutionized the understanding of what constitutes knowledge of language, which has led to a search for rules of grammar that explain a child's acquisition and development of language by both psycholinguists and linguists. Chomsky's claim (1965, 1980) that the child's input data is incomplete and may include ungrammatical forms has prompted an interest in caretaker talk. According to Chomsky, a child could not infer the abstract underlying properties and principles of language simply from the input (Harris 1992). Children acquire their language successfully at a fairly young age, although the speech directed to children is characterized as being fragmentary and often ungrammatical, hence 'the poverty of stimulus'.

Crain a.o. (1999) argue that the input does not indicate the possible meaning of the sentence, and it does not inform the child of ungrammatical sentences. Studies show that there is no difference in language development between groups of children whose parents use motherese and those whose parents do not. Subsequent studies have shown that neither expansion nor reinforcement have much effect on children's language production (Cazden 1965). Based on these findings, Crain a.o. (1999) argue that simplified speech does not actually make language acquisition any easier. Other points are still left unresolved, such as the relationship between the properties of caretaker speech and the order of language development, and the difficulty of determining the cause and effect relationship between them.

BIBLIOGRAPHICAL REFERENCES

Al-Shatti, Tariq. 2003. *Differentiating between the interactive styles of two types of caregiver with 11–21 months Kuwaiti children.* Ph.D. diss., University of Reading.

Camaioni, Lugia and Emiddia Longobardi. 2001. "Noun versus verb emphasis in Italian mother-to-child speech". *Journal of Child Language* 15:3. 773–786.

Chomsky, Noam. 1965. *Aspects of the theory of syntax.* Cambridge, Mass.: MIT Press.

——. 1980. Rules and representations. *Behavioral and Brain Science* 3.1–15, 42–61.

——. 1986. *Knowledge of language: Its nature, origin, and use.* New York: Praeger.

——. 1988. *Language and problems of knowledge.* Cambridge, Mass.: MIT Press.

Cook, Vivian and Mark Newson. 1996. *Chomsky's universal grammar: An introduction.* Oxford: Blackwell.

Crain, Stephen and Diane Lillo-Martin. 1999. *An introduction to linguistic theory and language acquisition.* Oxford: Blackwell.

Ferguson, Charles. 1956. "Arabic baby talk". *For Roman Jakobson*, ed. Morris Halle, Horace Lunt, and Hugh McLean, 121–128. The Hague: Mouton.

——. 2004. "Talking to children: A search for universals". *First language acquisition: The essential readings*, ed. Barbara Lust and Claire Foley, 176–189. Oxford: Blackwell.

Gleason, Jean Berko and Nan Bernstein Ratner. 1998. "Language acquisition". *Psycholinguistics*, ed. Jean Berko Gleason and Nan Bernstein Ratner, 347–452. New York: Macmillan.

Haggan, Madeline. 2002. "Self-reports and self-delusion regarding the use of motherese: Implications from Kuwaiti adults". *Language Sciences* 24.17–28.

Harris, Margaret. 1992. *Language experience and early language development: From input to uptake.* Hove, Sussex: Erlbaum.

Hayes, Donald and Margaret Ahrens. 1987. "Vocabulary simplification for children: A special case of 'motherese'?". *Journal of Child Language* 15.395–410.

Just, Marcel A. and Patricia A. Carpenter. 1984. *The psychology of reading and language comprehension.* Boston: Allyn and Bacon.

Omar, Margaret. 1973. *The acquisition of Egyptian Arabic as a native language.* The Hague: Mouton.

Snow, Catherine. 1986. "Conversations with children". *Language acquisition: Studies in first language acquisition*, ed. Paul Fletcher and Michael Garman, 69–89. Cambridge: Cambridge University Press.

——. 1995. "Issues in the study of input: Finetuning, universality, individual and developmental differences, and necessary causes". *The handbook of child language*, ed. Paul Fletcher and Brian MacWhinney, 180–193. Oxford: Blackwell.

Steinberg, Danny. 1993. *An introduction to psycholinguistics.* London: Longman

KHAWLA ALJENAIE (Kuwait University)

Case Roles

Case roles represent deep structure relations between noun phrases and the verb in a given proposition. Case roles, therefore, are not concerned with the traditional grammatical case that is related to inflectional forms designating the nominative, accusative, genitive, etc. While inflections (*ʿalāmāt al-ʾiʿrāb*) express surface cases, are related to surface structure, and are grammatical in nature, case roles express semantic relations or functions, are deep structure relations, and are semantic in nature. In other words, lexical relations are sensitive to case relations rather than grammatical relations. Therefore, semantic functions in Arabic cannot be distinguished by differences in nominal inflections. Robins (1971:228–255), however, warns of the inadequacy of an exclusively semantic definition of grammatical categories (parts of speech). Taylor (1995:196), on the other hand, concludes that semantic criteria are relevant to grammatical categorization and play a role in any intentional definition of word classes.

Case roles are part of the case grammar hypothesis, which is based on Fillmore's use of case relations for semantic functions. A proposition consists of a verb and a set of case phrases, each of which includes a case 'flag' (Latin *casus*) and a noun phrase. There is, however, little agreement on the set of case roles and their nature. At times, some case roles are inconsistently interchanged with each other and at other times their features differ from one linguist to another. Fillmore's case grammar evolved as an alternative to the kind of transformational grammar expounded in Chomsky (1965) (see Anderson 1994). The set of case relations offered in Fillmore (1968:24–25) was tentative and not necessarily intended to be exhaustive. As a result, there has been uncertainty and disagreement over the fixed set of case roles. Anderson (1994:460) believes that much of this disagreement can be understood in terms of diverse interpretations of the distributional and substantive constraints to which case relations conform. This view was expressed earlier by Nilsen (1972:47) and Somers (1987:111).

In 1965, the French linguist Tesnière introduced the notion of → 'valency', which investigates the relationship between a verbal governor and its immediate subordinates. Valency, a notion borrowed from chemistry, is the dependency (*dépendance*) between elements of a sentence that enter into certain governor–subordinate (*régissant-subordonné*) relationships. The most important element in this relationship is the verb (Somers 1987:5). Thus, case roles are case relations conditioned by the semantic features of the verb with which they co-occur. What is important about a verb, Fillmore asserts (1987:29), is its semantic valence, which is a description of the semantic role of its arguments. Case grammarians such as Fillmore, Chafe, Anderson, and Cook claim that case roles are most relevant to the subclassification of verb types. Nominal constituents partake of different case roles in different contexts. Taylor (1995:77) has rightly claimed that words are rather unselective with regard to the types of items they may be adjacent to. A verb-centered system has evolved in which case roles are predetermined by the selectional features of the verb. According to Anderson (1971:10), these cases are deep structure dependency relations of nouns to verbs; they are determined by the semantic features of the verb and specify the role of the noun in the state, process, or action expressed by the verb. Case roles are assigned by the verb to the noun. The verb is central and has one and only one case frame. This case frame represents an explicit array of cases intimately related to the meaning of the verb. The nouns in a given proposition are peripheral; they are not cases but case candidates. The same noun may be used in different contexts as → Agent, Experiencer, Benefactive, etc., depending on its verb. The case role itself is read into the noun from the verb. Proponents of the verb-centered approach include linguists such as Fillmore (1971), Chafe (1970), Cook (1972a, 1972b, 1973), and Anderson (1977, 1994). The following is a set of case roles with their notional characterization adapted from Fillmore (1968), Chafe (1970), and Cook (1972a):

i. Agent is the case required by an action verb which specifies the instigator of the action. This case is typically manifested by a [+ Human], [+ Animate] nominals.
ii. Experiencer is the case required by an experiential verb, and it specifies the undergoer of a psychological event of sensation, emotion, or cognition. There can be no experiential verb without someone to experience the psychological event.
iii. Benefactive is the case required by a benefactive verb which specifies the one in a state of possession, or the one who undergoes gain or loss in the transfer of an object (i.e. thing).
iv. Object with experiencer verbs is the case that specifies the content of, or the stimulus for, the experience. With benefactive verbs, 'object' case specifies what is possessed or transferred. With locative verbs, it specifies the object located, or undergoing change of location. Object case is represented by an animate or inanimate noun (phrase) which is affected by the action or state identified by the verb.
v. Locative is the case required by a locative verb which specifies an object's location or a change in its location. → Locatives are prepositional phrases; they subsume both locative and temporal adverbials, and co-occur with both Agents and non-Agents. Locatives include prepositional phrases like the following: *fī l-madrasati* 'in the school', *ʿalā ṭ-ṭāwilati* 'on the table', *ʿalā sāḥili l-baḥri* 'on the beach', *ʾilā baġdāda* 'to Baghdad', *fī l-qāhirati* 'in Cairo', *fī s-sāʿati l-wāḥidati* 'at one o'clock', *ʿabra l-jibāli* 'through the mountains', *fī ṣ-ṣabāḥi* 'in the morning', as in *waṣala zaydun ʾilā manzilihi fī s-sāʿati l-wāḥidati ṣabāḥan* 'Zayd arrived home at one in the morning', where the adverbials *ʾilā manzilihi* 'to his home', *fī s-sāʿati l-wāḥidati* 'at one o'clock', *ṣabāḥan* 'in the morning' all have the case role of Locative.
vi. Instrumental is a non-agentive nominal; it can be an animate or inanimate entity or a body part which an Agent can intentionally use to accomplish an action or instigate a process.
vii. Goal is the destination, the point of termination of an action, or an entity to which a judgment is directed.
viii. Source refers to a starting point or place of origin.

The term 'coreferential' is employed in the present investigation to refer to a given noun (phrase) that has two case roles that correspond to two distinct semantic functions. Therefore, case roles may be coreferential with each other, in the sense that a noun (phrase) may have the semantic function of two case roles simultaneously. Coreferential case roles are covert, or implicit, semantic relations assigned to noun phrases in a construction and are conditioned by the semantic componential features of the verb. At the deep structure analysis of the verb, covert case roles are crucial to the characterization of agentive verbs; in addition to their agentive meaning, they may have another meaning represented by other case roles such as Object, Source, Goal, or Experiencer. In what follows, a brief discussion is presented of coreferential case roles, defined and illustrated.

i. Agent + Object: Verbs like *rakaḍa* 'to run', *iġtasala* 'to have a bath', and *ʾarʿaba* 'to frighten' are agentive verbs whose Subject noun phrase can be assigned the case roles Agent and Object at the same time, as in *rakaḍa jamālu ʾilā d-dukkāni* 'Jamal ran to the shop', *yaġtasilu jamālu kulla jumʿatin* 'Jamal has a bath every Friday', and *ʾarʿaba jamālu l-ʾaṭfāla* 'Jamal frightened the children', where *jamālu* represents the coreferential, i.e. the underlying, case roles of Agent and Object at the same time in these sentences. *jamālu* is the Agent-as-Object who runs, has a bath, and frightens. With verbs like *ʾarʿaba* 'to frighten', *ʾaḍḥaka* 'to make someone laugh', and *ʾahāna* 'to insult', the Agent is regarded as the stimulus for the psychological experience undergone by the Experiencer like *al-ʾaṭfāla* 'the children.'

ii. Agent + Source: Verbs like *ramā* 'to throw', *ʾaʿṭā* 'to give', and *bāʿa* 'to sell' are agentive verbs whose surface Subject nominal has the coreferential roles of Agent and Source simultaneously, as in *ramā l-waladu l-kurata* 'the boy threw the ball', *daḥraja xālidu ṣ-ṣaxrata* 'Khalid rolled the stone', and *ʾaʿṭat al-bintu l-faqīra xubzan* 'the girl gave the poor man some bread', where *al-waladu* 'the boy', *xālidu* 'Khalid' and *al-bintu* 'the girl' are the Agent-as-Source of throwing, rolling, and giving. Other Agent-as-Source verbs are *barrara* 'to justify', *ittahama* 'to accuse', *taqāyaḍa* 'to barter', and *tanāzala* 'to concede'. Verbs like *intaqada* 'to criticize', *ittahama* 'to accuse', and *lāma* 'to blame' also take the Agent-as-Source, as in *intaqadat/lāmat/ittahamat salmā ṣadīqatahā* 'Salma criticized/blamed/accused her friend', where *salmā* is both an Agent and a Source; the actions of criticism, blame, and accusation have emanated from Salma.

iii. Agent + Goal: Agentive verbs like *saraqa* 'to steal', *iġtaṣaba* 'to take by force', *ištarā* 'to buy', *istaʿāra* 'to borrow', *iqtaraḍa* 'to borrow', and *ʾaxaḏa* 'to take' have their Subject nominals act as an Agent and a Goal at the same time. The Agent-as-Goal of stealing or taking by force, as in *saraqa zaydun al-kitāba* 'Zayd stole the book' and *iġtaṣaba salīmun ʾarḍa l-fallāḥi* 'Salim took the farmer's land by force' where *zaydun* and *salīmun* are assigned the case roles of both Agent and Goal.

iv. Agent + Experiencer: Verbs like *šamma* 'to smell something', *taʿallama* 'to learn something' and *istamaʿa* 'to listen to something' are agentive verbs whose Subject noun phrases can be assigned the case roles of Agent and Experiencer simultaneously, as in *taḏawwaqa z-zawju ṭ-ṭaʿāma* 'the husband tasted the food', *yastamiʿu l-mušāhidūna ʾilā kalimati l-iftitāḥi* 'the viewers are listening to the opening speech', and *yaḥukku l-waladu jildahu* 'the boy is scratching his skin'. Here, the Agent-as-Experiencer has tasted the food, is listening to the speech, and is scratching himself.

v. Agent + Benefactive: Verbs like *qabila* 'to accept', *istalama* 'to receive' are agentive verbs whose surface Subject nominals are Agents and Benefactives at the same time, as in *istalamat salmā hadiyyatan* 'Salma received a present' where Salma assumes the role of an Agent-as-Benefactive.

The above case roles are determined by the verb and are illustrated in the examples below based on verb types.

i. verbs of activity, i.e. action (dynamic) verbs like *banā* 'to build', *dammara* 'to destroy', *qatala* 'to kill', *ʿallama/darrasa* 'to teach', *baʿaja* 'to dent', *ġanna* 'to sing', *raqaṣa* 'to dance', as in *kasara r-rajulu l-bāba* 'the man broke the door' where *ar-rajulu* 'the man' is the Agent of the action verb *kasara* 'to break'.

ii. verbs of motion like *sāfara* 'to travel', *rakaḍa* 'to run', *mašā* 'to walk' are agentive, as in *rakaḍa yūsufu ʾilā l-madrasati* 'Joseph ran to school'.

iii. verbs of transfer like *bāʿa* 'to sell', *ištarā* 'to buy', *ramā* 'to throw', *ʾamsaka* 'to catch' are agentive too, as in *ramā zaydun al-kurata* 'Zayd threw the ball'.

iv. verbs of perception like *ʾaḥassa* 'to feel', *šaʿara* 'to feel', *raʾā* 'to see', *samiʿa* 'to hear', *istamaʿa* 'to listen', *taḏawwaqa* 'to taste', *šamma* 'to smell', as in *raʾat salmā ṣadīqatahā* 'Salma saw her friend'.

v. verbs of cognition like *ṣaddaqa* 'to believe', *ʾāmana* 'to believe', *ʿarafa* 'to know', *taʿallama/darasa* 'to learn', *taḏakkara* 'to remember', as in *ʿarafat salmā ṣadīqatahā* 'Salma recognized her friend'.

vi. verbs of emotion like *ʾaḥabba* 'to like, to love', *kariha* 'to dislike, to hate', *inzaʿaja* 'to get annoyed', *ḍaḥika* 'to laugh', *bakā* 'to cry', *ṣaraxa* 'to yell out', as in *tuḥibbu salmā zawjahā* 'Salma loves her husband', and

vii. verbs of desire such as *ʾamala* 'to hope', *ʾarāda* 'to want', as in *taʾmulu salmā ʾan tusāfira* 'Salma hopes to travel' are all experiential verbs where the surface subject nominal (*salmā*) is the Experiencer in all the above psychological verbs since Salma is the psychologically affected nominal entity.

This analysis is also applicable to nominal sentences (sentences without a main verb) with stative adjectives like *ġaḍbān* 'angry', *saʿīd* 'happy', *ʾaṭraš* 'deaf', *wāṯiqun* 'confident', as in *zaydun saʿīdun/ḥazīnun/xāʾifun* 'Zayd is happy/sad/scared' where *zaydun* is an Experiencer. However, verbs like *rabiḥa* 'to win', *wajada* 'to find' are Benefactive verbs, as in *rabiḥat salmā jāʾizatan* 'Salma won a prize' where *salmā* is a Benefactive only.

According to Fillmore (1987:29), verbs have a deep structure valence that can be expressed in terms of case frames. Unlike English, the Arabic verb *fataḥa* 'to open' takes Agent obligatorily and Object optionally, as in *fataḥa salīmun al-bāba* 'Salim opened the door'. Thus, the case frame for this Arabic verb *fataḥa* is + [___ A (O)]. The Arabic verb *infataḥa* 'to have been opened', however, obligatorily takes an Object case role and optionally an Instrumental, as in *infataḥa l-bābu* 'the door opened' or *infataḥa l-bābu bi-l-miftāḥi* 'the door opened with the key'. The verb's case frame here is + [___ O (I)]. Also, while the case frame of the verb *māta* 'to die' is + [___ E], as in *māta r-rajulu* 'the man died', the verb *qatala* 'to kill' can have the case frame + [___ A, E, (I)], as in *qatala salīmun ar-rajula bi-s-sammi* 'Salim poisoned the man' where *salīmun* is the Agent, *ar-rajula* 'the man' is the Experiencer, and *as-sammi* 'the poison' is the Instrumental, or the case frame + [___ I, E], as in *qatala s-sammu r-rajula* 'the poison killed the man' where the surface Subject *as-sammu* 'the poison' is the Instrumental. However, the passive voice *qutila* 'to be killed' has the case frame + [___ E], as in *qutila r-rajulu* 'the man was killed'. As for verbs like *raʾā* 'to see', they have the case frame + [___ E, O], as in *raʾat salmā l-mudarrisa* 'Salma saw the teacher' where *salmā* is the Experiencer while *al-mudarrisa* 'the teacher' is the Object. It is interesting to note that with action verbs like *ḏahaba* 'to go', we expect two case frames; the first is + [___ A, L], and the second is + [___ O, L]. If the surface Subject is the instigator of the action, as in *ḏahaba zaydun ʾilā l-mustašfā* 'Zayd went to the hospital', the first case frame applies where *zaydun* is the Agent. However, if the surface Subject *zaydun* is taken by someone else, such as an ambulance or a friend, the second case frame applies where *zaydun* is the Object. Similarly, experiential verbs like *xawwafa/ʾarʿaba* 'to frighten' take a double case frame: + [___ O, E] if the action is undeliberate, as in *ʾarʿaba salīmun al-ʾaṭfāla* 'Salim frightened the children' where Salim has not intended to frighten them, and + [___ A, E] if the action is deliberate on the part of Salim. However, verbs like *rawā/ʾaxbara* 'to tell, to report', *qāla* 'to say', *taḥaddaṯa* 'to speak', *saʾala* 'to ask', *istajwaba* 'to question, to interrogate', *xawwa-fa/ʾarʿaba* 'to frighten', *ʾarḍā* 'to please' have the case frame + [___ A, E, O], as in *saʾalat salmā l-walada suʾālan* 'Salma asked the boy a question'. With inanimate surface Subject nominals, verbs like *waṣala* 'to arrive', *ittajaha* 'to head for' have the case frame + [___ O, L], as in *waṣalat risālatun min baġdāda* 'a letter has arrived from Baghdad' and *ittajahat ar-rīḥu šamālan* 'the wind headed north' where *risālatun* and *ar-rīḥu* are Objects.

With nominal sentences, stative adjectives can be sub-classified into: (a) experiencer adjectives like *saʿīd/farḥān* 'happy', *ġaḍbān* 'angry', whose case frame is + [___ E, (O)], as in *salmā saʿīdatun* 'Salma is happy', and (b) agentive adjectives like *raʾūf* 'kind', *ḏ̣ālim* 'tyrant', *qāsin* 'harsh', *mutaʿajrif* 'arrogant', *ʿādilun* 'just' whose case frame is + [___ A (O)], as in *ar-raʾīsu ḏ̣ālimun* 'the president is a tyrant'.

BIBLIOGRAPHICAL REFERENCES

Anderson, John M. 1971. *The grammar of case: Towards a localist theory*. Cambridge: Cambridge University Press.

——. 1977. *On case grammar: Prolegomena to a theory of grammatical relations*. London: Humanities Press.

——. 1994. "Case". *Encyclopedia of language and linguistics*, editor-in-chief R.E. Asher, II, 447–453. Oxford: Pergamon Press.

Chafe, Wallace. 1970. *Meaning and the structure of language*. Chicago: University of Chicago Press.

Cook, Walter. 1972a. "A set of postulates for case grammar analysis". *Language and linguistics: Working papers* IV, 35–49. Washington, D.C.: Georgetown University, School of Languages and Linguistics.

——. 1972b. "A case grammar matrix". *Language and linguistics: Working papers* VI, 15–47. Washington, D.C.: Georgetown University, School of Languages and Linguistics.

——. 1973. "Covert case roles". *Language and linguistics: Working papers* VII, 52–81. Washington D.C.: Georgetown University, School of Languages and Linguistics.

Fillmore, Charles. 1968. "The case for case". *Universals in linguistic theory*, ed. Emmon Bach and Robert Harms, 1–88. New York: Holt, Rinehart and Winston.

——. 1971. "Some problems for case grammar". *Report of the twenty-second annual round table meeting on linguistics and language*, ed. Richard O'Brien. 35–56. Washington, D.C.: Georgetown University Press.

——. 1987. "A private history of the concept of 'frame'". *Concepts of case*, ed. René Dirven and Günter Radden. 28–36. Tubingen: Narr.

Nilsen, Don Lee Fred. 1972. *Toward a semantic specification of deep case*. The Hague: Mouton.

Robins, R.H. 1971. *General linguistics: An introductory survey*. London: Longman.

Somers, H.L. 1987. *Valency and case in computational linguistics*. Edinburgh: Edinburgh University Press.

Taylor, John. 1995. *Linguistic categorization: Prototypes in linguistic theory*. Oxford: Oxford University Press.

Tesnière, Lucien. 1965. *Eléments de syntaxe structurale*. Paris: Klincksieck.

HUSSEIN ABDUL-RAOF (University of Leeds)

Case Theory

Case is a concept with deep historical roots in Western and indigenous Arabic grammatical theory. In modern linguistics, case refers not only to the morphological shape of nominals and their modifiers but also to their syntactic and semantic roles within a sentence (Ura 2003:334), for example subject and object or agent and theme (→ case roles; → theta roles). Arabic grammar, medieval and modern, recognizes three cases of nouns: nominative (*rafʿ* 'promotion'), accusative (*naṣb* 'demotion'), and genitive (*jarr* 'abatement'), realized as the suffixes *-u, -a*, and *-i* respectively (Gaballa 1986:25–30). These are illustrated in (1), from Farghal (1986:8):

(1) *kataba l-walad-u* *d-dars-a*
wrote the-boy-nom the-lesson-acc
fī l-bayt-i
in the-house-gen
'The boy wrote the lesson in the house'

The medieval grammarians posited that each case is assigned under → government; nominative and accusative cases in (1) are assigned by the verb *kataba*, and genitive case by the preposition *fī* (→ *ʿamal*). Each case is associated with a set of syntactic and semantic (participant) roles that a noun phrase (NP) may bear. In general, nominative case is assigned to subjects of root clauses (Abdul-Raof 2001:109–110), accusative case to objects (direct and indirect) of verbs, and genitive case to objects of prepositions (Gaballa 1986:54–56). It should be noted that the accusative has a particularly large set of functions in both Classical and Modern Standard Arabic. Detailed discussion of these accusative functions can be found in Wright (1896:II, A.3) and Cantarino (1975, II, VI) respectively.

Like government, case has evolved into a technical notion in generative grammar. In early work in the 1960s, case received little attention, being simply a feature assigned to a noun through a late transformational rule on the basis of its grammatical function, which was in turn computed from its position in a syntactic configuration (Ura 2003:335). Case assumed prominence in syntactic theory as part of the program in the 1970s of restricting the scope of rule systems in natural language grammars and, more specifically, of accounting for the distribution of NPs (Chomsky 1986:187). This was done in government-binding (GB) theory by drawing a distinction between *morphological* case and *abstract* Case. As the name implies, abstract Case is a property assigned to an NP whether or not it is realized inflectionally as morphological case on its N head (Chomsky 1986:74). Abstract Cases correspond to their concrete counterparts: nominative, accusative, and genitive (or oblique) Case. To ensure that every lexical NP in a sentence bears Case, Chomsky (1981:49, 175) proposed the *Case Filter*:

(2)*NP if NP has phonetic content and has no Case.

(2) allows that NPs that lack a phonetic matrix – empty categories such as PRO, *pro*, NP-trace, and WH-trace – need not be assigned Case. Actually, at this stage of the theory, *pro* was understood to have nominative Case assigned to it as the subject of a finite clause (Chomsky 1982:80–81), and WH- trace was distinguished from NP-trace, in part, by being Case-marked (Chomsky 1981:69).

A second distinction drawn in GB Case theory is between *structural* Case and *inherent* Case (Chomsky 1981:171, 1986:193). Structural Case is assigned solely on the basis of the position of an NP in a syntactic configuration; inherent Case is assigned by a non-nominal category that also assigns its NP a thematic role such as agent or patient (→ case roles; → theta roles). Both structural and inherent Case are assigned under head government. An example of structural Case assignment would be the nominative Case assigned to the subject NP governed by the functional head Inflection (in later work, inspired by Pollock 1989, decomposed into Tense and Agreement). In (1), Inflection rather than the verb would assign nominative Case to *al-walad-u*. An example of inherent Case would be the genitive Case assigned by a preposition to its complement, which it governs and assigns a thematic role; for instance, in (1), *fī* assigns the location role to *al-bayt-i*. Genitive case may also be structural in Arabic (in construct states – see below). Similarly, verbs may assign both inherent and structural Case (Ura 2003:336–337).

In the minimalist program, Case assignment under → government by lexical categories is replaced by Case checking by functional categories such as Agreement (subject and object, abbreviated Agr$_{S}$ and Agr$_{O}$, respectively) and Tense. Structural Case-checking takes place (like agreement) exclusively in a specifier-head relation, with the NP raising to the specifier of the functional head which checks its Case feature (Chomsky 1995:173–74). The process is illustrated schematically in (3) for subject and object NPs, assuming the VP-internal hypothesis (Speas 1990, among others), which locates subjects in the specifier of the verb phrase (VP) at the beginning of a derivation:

(3) a.[$_{\text{AgrSP}}$ [$_{\text{Spec}}$ ——] [$_{\text{AgrS'}}$ [$_{\text{AgrS}}$ [$_{\text{TP}}$ [$_{\text{Spec}}$ ___] [$_{\text{T'}}$ [$_{\text{T}}$ [$_{\text{AgrOP}}$ [$_{\text{Spec}}$ ——] [$_{\text{AgrO'}}$ [$_{\text{AgrO}}$ [$_{\text{VP}}$ [$_{\text{Spec}}$ Subj] [$_{\text{V'}}$ $_{\text{V}}$ Obj]]]]]]

b.[$_{\text{AgrSP}}$ [$_{\text{Spec}}$ Subj] [$_{\text{AgrS'}}$ [$_{\text{AgrS-T}}$ [$_{\text{TP}}$ [$_{\text{T'}}$[t_{T} [$_{\text{AgrOP}}$ [$_{\text{Spec}}$ Obj] [$_{\text{AgrO'}}$ [$_{\text{AgrO-V}}$ [$_{\text{VP}}$ [$_{\text{Spec}}$ t_{Subj} [$_{\text{V'}}$ t_{V} t_{Obj}]]]]]]]

The Agr$_{O}$-V combination licenses accusative Case for the object in the specifier of Agr$_{O}$ Phrase, and the Agr$_{S}$-T combination licenses nominative Case for the subject in the specifier of Agr$_{S}$ Phrase. Recent theoretical innovations, such as eliminating Agr projections in favor of multiple specifiers (Chomsky 1995:285, 349–355) and unifying Case-checking and agreement as reflexes of a single operation Agree (Chomsky 2000:101), have as yet had little impact on the study of Arabic, so these will be put aside here.

We now consider Case checking in Arabic, in three domains: (a) simple verbal and nominal sentences, (b) embedded subjects, and (c) possessive NPs.

Homeidi (1993) is representative of studies of Arabic in the GB framework which assume Case-assignment under government. Homeidi considers how Case is assigned in the traditional taxonomy of clause types: nominal sentences (those beginning with a noun [phrase]) and verbal sentences (those beginning with a verb). He further divides verbal sentences into transitive and intransitive structures. Intransitives need assign only nominative Case to their subjects, which, he proposes (following a suggestion in Bouchard 1984), are governed by a verb phrase-inflection complex rather than inflection alone (1993:126). Transitive verbs assign accusative Case to their object(s) internal to VP (1993:127, PP omitted here):

(4) a. [$_{\text{IP}}$ [$_{\text{VP-I}}$ [$_{\text{V}}$ ʾaʿṭā] [kitāban] [ʾaḥmada]]
gave book-acc Aḥmad-acc
[$_{\text{NP}}$ muḥammad-u]]
Muḥammad-nom

b. [$_{\text{IP}}$ [$_{\text{VP-I}}$ [$_{\text{V}}$ ʾaʿṭā] [$_{\text{NP}}$ muḥammad-u]
gave Muḥammad-nom
[$_{\text{NP}}$ ʾaḥmad-a] [NP kitāb-an]]]
book-acc Ahmad-acc
'Muḥammad gave Aḥmad a book'

In (4a), *muḥammad* is governed by VP-I and assigned nominative Case. *ʾaḥmad* and *kitāb* are complements of V and so assigned accusative Case and transposed under dative shift. Homeidi does not discuss the movement of *muḥammadu* [sic] into the VP.

'Nominal' sentences with *kān-* 'was,' the past tense equivalents of true nominal sentences, such as *aš-šamsu mušriqatun* 'the sun is shining', are derived in the same way (Homeidi 1993:129):

(5)a. [$_{\text{IP}}$ [$_{\text{VP-I}}$ [$_{\text{V}}$ kān-at [$_{\text{AP}}$ mušriqat-an]]
was-3fsg shining-acc
[$_{\text{NP}}$ aš-šams-u]]
the-sun-nom

b. [$_{\text{IP}}$ [$_{\text{VP-I}}$ [$_{\text{V}}$ kān-at [$_{\text{NP}}$ aš-šams-u]
was-3fsg the-sun-nom
[$_{\text{AP}}$ mušriqat-an]]]
shining-acc
'The sun was shining'

Kānat assigns accusative Case to its AP complement, and VP-I assigns nominative Case to the subject *aš-šamsu* before movement into the VP.

Nominal sentences in the present tense like *aš-šamsu mušriqatun* 'the sun is shining,' sometimes called *equational* sentences, pose special challenges for Case theory because the subject and predicate are both nominative but have no overt Case assigners. This fact is highlighted by the emphatic variant with *ʾinna*, which assigns (structural) accusative Case to the subject (Homeidi 1993:130, Abdul-Raof 2001:112):

(6) *ʾinna š-šams-a mušriqat-un*
the-sun-acc shining-nom
'The sun is indeed shining'

Sentences like (6) rule out an analysis of non-emphatic equational sentences in which subject and predicate share nominative Case because they are in an agreement relation. If this were so, we should expect the predicate in (6) to be accusative *mušriqatan*, contrary to fact (Plunkett 1993:248). The nominative Case on the predicate also counts against a null copula analysis, in which a covert present tense *be* assigns Case to the predicate. If this were so, we should again expect the Case to be accusative, as it is with the overt past tense form *kānat* in (5) (Benmamoun 2000:43). To overcome such difficulties, Homeidi (1993:129) proposed that the nominative Case of the subject and the predicate in non-emphatic equationals is due to the default status of that Case, a common theme in the literature (cf. Al-Bayaty 1990:94–95; Plunkett 1993:245; Ouhalla 1997:207 and the references cited there). An improved account is offered by Benmamoun (2000:42, 49), who proposes that Arabic verbless clauses are the projection of Tense with only a determiner feature and no verb feature. This accounts for the nominative Case of the subject (which is licensed by Tense and checks its determiner feature) and the nominative Case on the predicate (which, in the absence of a verb to check accusative Case, surfaces as default nominative).

The notion of nominative as a default Case is more plausible in two other constructions: SVO sentences with full subject-verb agreement and topicalization structures (Homeidi 1993:131–132; Mohammad 1990:101; cf. Moore 1988: 287–288):

(7) *aṭ-ṭullāb-u qaraʾ-ū*
the-students.m-nom read-3mpl
l-kitāb-a fī l-madrasat-i
the-book-acc in the-school-gen
'The [male] students read the book at [the] school'

(8) *al-kitāb-u*$_1$ *qaraʾa-hu*$_1$
the-book-nom read-3msg-it
aṭ-ṭullāb-u fī l-madrasat-i
the-students.m-nom in the-school-gen
'The book, the students read it at [the] school'

(9) *al-bint-u*$_1$ *yabdū ʾanna-hā*$_1$
the-girl-nom seem-3msg that-her.f
sāfar-at
traveled-3fsg
'The girl, it seems that she traveled'

Homeidi asserts that *aṭ-ṭullābu* in (7) bears nominative Case because there is no Case-assigning governor available, since it is outside the VP-I complex (cf. [4]). This conclusion no longer holds under minimalist assumptions: *aṭ-ṭullābu* can move to subject position (the specifier of either Agr$_{\text{S}}$ Phrase or of Tense Phrase) to check nominative Case. The default Case analysis is more secure in the topicalization structures (8)–(9), since neither *al-kitābu* nor *al-bintu* is the subject of the (string) adjacent verb, as the resumptive pronouns make clear. Being left peripheral to the clause (Rizzi 1997), the topics evidently do not receive Case from Agr_S; a reasonable alternative would be to posit an agreement relation with a functional head Topic, which checks nominative Case for the NP in its specifier (al-Shorafat 1999:15–16).

The second domain of application is Case-marking of subjects in complement clauses. As is well known, the complementizer *ʾanna* (like its 'sister' *ʾinna*) assigns accusative Case to the initial NP of the clause that it selects, including the subject (Abdul-Ghany 1981:133; but cf. Shlonsky 2000:332–336):

(10) *yabd-ū ʾanna l-bint-a*
seem-3msg that the-girl-acc
sāfar-at
traveled-3fsg
'It seems that the girl traveled'

If the subject is a pronominal, it must be an accusative clitic, like *-hā* 'her' in (9). A subject pronoun, either overt or covert (*pro*), is impossible, in contrast to root clauses (cf. Harbert and Bahloul 2002:49; Mohammad 1990:100):

(11) a. *hiya sāfar-at*
she traveled-3fsg
'She traveled'
b. **yabd-ū ʾanna hiya sāfar-at*
seem-3msg that she traveled-3fsg
'It seems that she traveled'

(12) a. pro *sāfar-at*
traveled-3fsg
'[She] traveled'
b. **yabd-ū ʾanna* pro *sāfar-at*
seem-3msg that traveled-3fsg
'It seems that she [*pro*] traveled'

The fact that pronominals and lexical NPs both take accusative Case in *ʾanna* clauses favors a unified account. An account is also needed of the fact that an accusative subject occurs in a finite complement clause, an instance of grammatical function splitting (Ura 2003:344). Aoun (1985:56–57) proposes that accusative Case in *ʾanna* clauses is assigned under head government by the complementizer. Aoun's proposal has three consequences: (a) a clitic as well as a lexical NP subject must be in (preverbal) subject position underlyingly so as to be in a governed position. In this position, the closest governor and Case-assigner is not Inflection, since the subject is external to its intermediate projection 'I', but the complementizer (LeTourneau 1993:263–66). This analysis accounts for the accusative subject of a finite verb; (b) the clitic must at the surface incorporate into the complementizer that governs it; and (c) the clitic trace is (the head of) a variable, an empty NP bound by an operator in an A-bar position (Chomsky 1981:185). Its status as a variable is confirmed by its Case-marking (variables being Case-marked traces) and by its binding properties in *wh-* (information) questions (LeTourneau 1993:272–79).

Accusative Case can also be assigned to a complement subject by a verb in the (optional or obligatory) absence of a complementizer (Fassi Fehri 1993:33, 65):

(13) a. *ḥasib-tu r-rajul-a marīḍ-a-n*
thought-1sg the-man-acc sick-acc-indef
'I thought the man [to be] sick'
b. *man*$_i$ *ḥasib-ta (ʾanna) r-rajul-a*
who thought-2msg (that) the-man-acc
ḍarab-at$_i$
beat-3msg
'Who did you think (that) the man beat?'

In (13a), the subject of the complement clause checks its Case in the specifier of Agr_O, as an ordinary direct object would (Lasnik 1999:27–29). The Case is structural because the subject bears no thematic relation to the verb *ḥasibtu*, which selects a clausal complement. How the accusative Case of the predicate *marīḍan* is checked remains an unresolved question. Available proposals – Case marking by the verb (Fassi), treating the predication as a verbless clause headed by Tense (following Benmamoun), or taking *ar-rajulu marīḍan* as a small clause – all suffer from conceptual defects. Turning to (13b), the embedded subject *ar-rajula* occurs in a finite clause and so is excluded from 'I' as in (10); it therefore checks accusative Case with the nearest governing head: *ʾanna*, if present, otherwise *ḥasibta*.

The third application is the distribution of Case in possessive constructions, which come in two varieties in Arabic (and Hebrew, research on which has advanced the study of Arabic analogs): the → *construct state* and the *free state* (Ouhalla and Shlonsky 2002:32):

(14) a. *bayt-u* *l-wazīr-i*
house-nom the-minister-gen
(construct: Standard Arabic)
'the house of the minister'

b. *d-dār* *dyal* *l-wazīr*
the-house of the-minister
(free: Moroccan Arabic)
'the house of the minister'

In the construct state in (14a), the head *bayt* is the possessed and the NP *al-wazīr* the possessor. The head invariably assigns genitive case (and structural genitive Case) to the second term; the Case of the head itself varies with its grammatical context (al-Aboudi 1985:15):

(15) a. *bayt-u* *l-wazīr-i*
house-nom the-minister-gen
jamīl-un
beautiful-nom
'The minister's house is beautiful'

b. *ra'ay-tu* *bayt-a* *l-wazīr-i*
saw-1sg house-acc the-minister
l-yawm-a
the-day-acc
'I saw the minister's house today'

c. *fī* *bayt-i* *l-wazīr-i*
in house-gen the-minister-gen
'in the minister's house'

The free state in (14b) manifests no morphological case on either NP, as is general in colloquial Arabic, but *l-wazīr* presumably receives its abstract Case from *dyal* or its analogs *(bi)tāᶜ* (Egyptian) or *tabaᶜ* (Palestinian, Mohammad 1999:34–35) 'belonging to', *māl* 'property' (Jewish Baghdadi Arabic, Melcer 1995:68–75), or Hebrew *šel* (→ analytic genitive). Modifiers of either the first or second term agree in Case with the noun heading the minimal projection containing both (al-Thalji 1988:51).

The construct state is standardly analyzed as a Determiner Phrase (DP) with a determiner (D) head and an NP complement. In (16a), the possessor *al-wazīr* is in the specifier of NP, and the possessed *bayt* is the N head; the latter raises to the empty D head in (16b) to assign the genitive Case of the possessor and realize surface word order (Benmamoun 1999:623–624; Mohammad 1999:33–34 and the references cited there):

(16) a. [$_{\text{DP}}$ [$_{\text{D}}$ __] [$_{\text{NP}}$ [$_{\text{Spec}}$ [$_{\text{NP}}$ al-wazīr]]
[$_{\text{N}}$ bayt]]]

b. [$_{\text{DP}}$ [$_{\text{D}}$ bayt$_1$] [$_{\text{NP}}$ [$_{\text{Spec}}$ [$_{\text{NP}}$ al-wazīr-i]]
[$_{\text{N}}$ t_1]]]

The Case of the entire DP is checked by the relevant functional head and realized suffixally on the (N in the) D head. Analysis of the free state requires postulating a functional category between D and N, usually called Number Phrase (NumP), on the basis of the following Hebrew data:

(17) [$_{\text{DP}}$ [$_{\text{D}}$ ha- [$_{\text{NumP}}$ [$_{\text{Num}}$ axila$_1$] [$_{\text{NP}}$ [$_{\text{spec}}$
the eating
šel dan] [$_{\text{N'}}$ [$_{\text{N}}$ t_1] [$_{\text{NP}}$ et ha-tapuax]]]]]
of Dan acc the apple
'the eating of Dan [of] the apple'

The D head is occupied by *ha-*, so raising of the N head *axila* 'eating' to that position is barred. However, the thematic relation between *axila* and its complement *(et) ha-tapuax* implies that the noun-subject-object order in (17) is derived by movement of *axila* from the head position of the NP to the empty head Num[ber] of NumP, where it receives its number feature (Ritter 1991:42–44).

Among extensions of N-D raising in construct states, two deserve comment. The first is the analysis of quantified construct states. Quantified construct states parallel lexical ones in that the quantifiers are heads that bear various Cases and assign genitive Case to their second term (Benmamoun 1993:33; 1999:622–623; LeTourneau 1995:33–35):

(18) a. *kull-u* *ṭ-ṭullāb-i*
all-nom the-students.m-gen
jā'ū
came-3mpl
'All the students came'

b. *ra'ay-tu* *kull-a* *ṭ-ṭullāb-i*
saw-1sg all-acc the-students-gen
'I saw all the students'

c. *kitāb-u* *kull-i* *ṭ-ṭullāb-i*
book-nom all-gen the-students-gen
'the book of all the students'

The quantified construct states in (18) are derived in the same way as the lexical construct states in (15): by Q–D (quantifier to determiner) raising, parallel to N–D (noun to determiner) raising in (16). For arguments that genitive Case of the second term is checked overtly in a specifier-head relation, see Benmamoun (1998). Quantified construct states alternate with a variant in which the quantifier attaches to a clitic and agrees in Case with the noun it follows:

(19) a. *aṭ-ṭullāb-u* *kull-u-hum*
the-students-nom all-nom-them
jāʾū
came-3mpl
'All the students came'
b. *raʾay-tu* *ṭ-ṭullāb-a*
saw-1sg the-students-acc
kull-a-hum
all-acc-them
'I saw all the students'
c. *kitāb-u* *ṭ-ṭullāb-i*
book-nom the-students-gen
kull-i-him
all-gen-them
'the book of all the students'

For arguments that the quantified construct state is not derived from the structure in (19), contra Shlonsky (1991), see Benmamoun (1999).

The second extension concerns construct states headed by a process nominal:

(20) *sarr-a-nī* *naql-u*
pleased-3msg-me reporting-nom
l-jarīdat-i *l-xabar-a*
the-newspaper-gen the-news-acc
'The reporting of the newspaper of the news pleased me'

Of interest is the accusative Case on *al-xabara*, which can be assigned only by a verb, here *naqala*, from which the process nominal *naqlu* is derived. Accounts of Case-checking in this construction assume an underlying VP (Mohammad 1999:33), V–N amalgamation (Borer 1996:37–38), an event affix (Fassi Fehri 1993:235–240), or *maṣdar* (nominalization) formation in the syntax rather than at argument structure (Grenat 1996:149, 178–180; Hazout 1990:148–151, 1995:365–370).

BIBLIOGRAPHICAL REFERENCES

Abdul-Ghany, Mohammad Kamal al-Deen. 1981. *Government binding in classical Arabic*. Ph.D. diss., University of Texas at Austin.

Abul-Raof, Hussein. 2001. "On the subject in Arabic". *Journal of Semitic Studies* 46.1:97–120.

Al-Aboudi, Fahmi Basil Shukri. 1985. *The syntax of "idafah": Annexation in Arabic*. Ph.D. diss., Indiana University, Bloomington.

Al-Bayaty, Jumah. 1990. *Adjunction in Arabic, case, and chain theory*. Ph.D. diss., Simon Fraser University, Vancouver.

Al-Shorafat, A. 1999. "Topics in Arabic: A minimalist approach". *al-ʿArabiyya* 32.1–21.

Al-Thalji, Abdel-Majid Ibrahim. 1988. "Head-modifier agreement in case in Arabic". *Archív Orientální* 56:1.42–53.

Aoun, Joseph. 1985. *A grammar of anaphora*. Cambridge: MIT Press.

Benmamoun, Elabbas. 1993. "Null pronominals in the context of NPs and QPs". *Proceedings of the eleventh West coast conference on formal linguistics*, ed. J. Mead, 32–43. Stanford: Center for the Study of Language and Information.

——. 1998. "Spec-head agreement and overt case in Arabic". *Specifiers: Minimalist approaches*, ed. David Adger, Susan Pintzuk, and Bernadette Plunkett, 110–125. Oxford: Oxford University Press.

——. 1999. "The syntax of quantifiers and quantifier float". *Linguistic Inquiry* 30.621–642.

——. 2000. *The feature structure of functional categories: A comparative study of Arabic dialects*. New York: Oxford University Press.

Borer, Hagit. 1996. "The construct in review". *Studies in Afroasiatic grammar: Papers from the second conference on Afroasiatic languages*, ed. Jacqueline Lecarme, Jean Lowenstamm, and Ur Shlonsky, 30–61. Sophia: Antipolis.

Bouchard, Denis. 1984. *On the content of empty categories*. Dordrecht: Foris.

Cantarino, Vicente. 1975. *The syntax of modern Arabic prose*, II. Bloomington: Indiana University Press.

Chomsky, Noam. 1981. *Lectures on government and binding: The Pisa lectures*. Dordrecht: Foris.

——. 1982. *Some concepts and consequences of the theory of government and binding*. Cambridge: MIT Press.

——. 1986. *Knowledge of language: Its nature, origin, and use*. New York: Praeger.

——. 1995. *The minimalist program*. Cambridge: MIT Press.

——. 2000. "Minimalist inquiries: The framework". *Step by step: Essays in minimalist syntax in honor of Howard Lasnik*, ed. Roger Martin, David Michaels, and Juan Uriagereka, 89–156. Cambridge: MIT Press.

Eid, Mushira and Clive Holes (eds.). 1993. *Perspectives on Arabic linguistics*, V. Amsterdam and Philadelphia: J. Benjamins.

Farghal, Mohammad Ali. 1986. *The syntax of wh-questions and related matters in Arabic*. Ph.D. diss., Indiana University, Bloomington.

Fassi Fehri, Abdelkader. 1993. *Issues in the structure of Arabic clauses and words*. Dordrecht: Kluwer Academic Publishers.

Gaballa, Hassan. 1986. *Case marking and Classical Arabic: A semantic and historiographic study.* Ph.D. diss., University of Delaware, Newark.

Grenat, Mohamed Hasan. 1996. *Argument structure and the Arabic masdar*. Ph.D. diss., University of Essex.

Harbert, Wayne and Maher Bahloul. 2002. "Postverbal subjects in Arabic and the theory of agreement". Ouhalla and Shlonsky (2002:45–69).

Hazout, Ilan. 1990. *Verbal nouns: Theta-theoretic studies in Hebrew and Arabic.* Ph.D. diss., University of Massachusetts, Amherst.

——. 1995. "Action nominalizations and the lexicalist hypothesis". *Natural Language and Linguistic Theory* 13.355–405.

Homeidi, M. 1993. "Government and binding and case assignment in Modern Standard Arabic". *Papers and Studies in Contrastive Linguistics* 29.123–139.

Lasnik, Howard. 1999. "Lectures on minimalist syntax". *Minimalist Analysis*, 25–73. Malden, Mass.: Blackwell.

LeTourneau, Mark S. 1993. "Case-marking and binding of subject clitics in Arabic complement clauses". Eid and Holes (1993:261–290).

——. 1995. "Internal and external agreement of quantified construct states" *Perspectives on Arabic linguistics*, VII, ed. Mushira Eid, 29–57. Amster dam and Philadelphia: J. Benjamins.

Melcer, Ioram. 1995. "The analytical genitive in the Jewish-Baghdadi Arabic dialect". *Zeitschrift für Arabische Linguistik* 29.59–76.

Mohammad, Mohammad A. 1990. "The problem of subject-verb agreement in Arabic: Towards a solution". *Perspectives on Arabic linguistics*, I, ed. Mushira Eid, 95–125. Amsterdam and Philadelphia: J. Benjamins.

——. 1999. "Checking and licensing inside the DP in Palestinian Arabic". *Perspectives on Arabic linguistics*, XII, ed. Elabbas Benmamoun, 27–44. Amsterdam and Philadelphia: J. Benjamins.

Moore, John. 1988. "Word order and case assignment in modern standard Arabic". *Papers from the 23rd annual regional meeting of the Chicago Linguistics Society*. II. *Parasession on autosegmental and metrical phonology*, 283–294. Chicago: Chicago Linguistics Society.

Ouhalla, Jamal. 1997. "Genitive subjects and the VSO order". *Studies on universal grammar and typological variation*, ed. Artemis Alexiadou and T. Alan Hall, 197–218. Amsterdam and Philadelphia: J. Benjamins.

—— and Ur Shlonsky (eds.). 2002. *Themes in Arabic and Hebrew syntax*. Dordrecht: Kluwer Academic Publishers.

——. 2002. "Introduction". Ouhalla and Shlonsky (2002:1–43).

Plunkett, Bernadette. 1993. "On the position of subjects in Arabic". Eid and Holes (1993:231–260).

Pollock, Jean-Yves. 1989. "Verb movement, universal grammar, and the structure of IP". *Linguistic Inquiry* 20.365–424.

Ritter, Elizabeth. 1991. "Two functional categories in noun phrases: Evidence from modern Hebrew". *Perspectives on phrase structure: Heads and licensing*, ed. Susan J. Rothstein, 37–62. New York: Academic Press.

Rizzi, Luigi. 1997. "The fine structure of the left periphery". *Elements of grammar*, ed. Lilian Haegeman, 281–337. Dordrecht: Kluwer Academic Publishers.

Shlonsky, Ur. 1991. "Quantifiers as functional heads: A study of quantifier float in Hebrew". *Lingua* 84.159–180.

——. 2000. "Remarks on the complementizer layer of standard Arabic". *Research in Afroasiatic grammar: Papers from the third conference on Afroasiatic languages*, ed. Jacqueline Lecarme, Jean Lowenstamm, and Ur Shlonsky, 325–343. Amsterdam and Philadelphia: J. Benjamins.

Speas, Margaret. 1990. *Phrase structure in natural language.* Dordrecht: Kluwer Academic Publishers.

Ura, Hiroyuki. 2003. "Case". *The handbook of contemporary syntactic theory*, ed. Mark Baltin and Chris Collins, 334–373. Malden, Mass.: Blackwell.

Wright, William. 1896. *A grammar of the Arabic language.* (Repr. Beirut: Librairie du Liban, 1974).

MARK S. LETOURNEAU (Weber State University)

Cataphora

The term 'cataphora' designates anticipatory pronouns and other forward-referring proforms. In this entry, however, the focus is on cataphoric pronouns occurring prior to the nominals they refer to, as opposed to anaphoric pronouns referring back to their antecedents. Compared to European languages, the use of cataphora in Arabic, especially in written Arabic, appears to be uncommon if not rare. Yet, in modern written Arabic, highly influenced as it is by such languages as English and French, cataphoric pronouns are not infrequently attested. By and large, the use of cataphoric pronouns in written Arabic appears to be restricted to the following structures (indexing is marked by subscripts, and cataphoric pronouns, whenever morphologically realized, are boldfaced):

(1) a. OVS:

*ṣadīq-a-**hu**$_i$*	*laqiya*	*zayd$_i$-un*
friend-ACC-**his**$_i$	met	Zayd$_i$-NOM

'Zayd met his friend'

b. VOS:

*laqiya ṣadīq-a-**hu**$_i$*	*zayd$_i$-un*
met friend-ACC-**his**$_i$	Zayd$_i$-NOM

'Zayd met his friend'

c. AdvVSO:

*fī ġurfat-i-**hi**$_i$*	*jalasa*	*zayd$_i$-un*
in room-GEN-**his**$_i$	sat	Zayd$_i$-NOM

'In his room Zayd sat'

(2) a. Clausal predicate preceding the subject:
*ʾab-ū-**hu***$_i$ *muʿallim-un*
father-NOM-**his**$_i$ teacher-NOM
zayd$_i$*-un*
Zayd$_i$-NOM
'Zayd's father is a teacher'

b. Object clause VS:
mā ʾarāda ʾaxaḏa zayd-un
what (**he**$_i$) wanted took Zayd$_i$-NOM
'What Zayd wanted he took'

c. Adverbial clause preceding the main clause:
lammā rajaʿa wajada zayd-un
when (**he**$_i$) returned found Zayd$_i$-NOM
al-kitāb-a
the-book-ACC
'When he returned, Zayd found the book'

Whereas sentences (1) are simple, sentences (2) are complex sentences introduced by a clause rather than by a verb or a phrase. What is common to all six sentences above is that they display a non-basic constituent order. Indeed, one may argue based on the above data that cataphoric pronouns in written Arabic are restricted to topical constituents. Both *ṣadīqa-hu* in (1a) and the adverbial constituent in (1c) function as topics, whereas in (1b) *ṣadīqa-hu* is topical relative to *zayd*. Similarly, in (2) the introductory clause in each of the sentences functions as topic with respect to the remainder of the sentence.

Cataphoric pronouns are inadmissible as part of the subject in sentences displaying the basic VSO word-order pattern. This is borne out by the inadmissibility of sentences such as (3):

(3) **laqiya ṣadīq-u-**hu***$_i$ *zayd-an*$_i$
met friend-NOM-**his**$_i$ Zayd$_i$-ACC
'Zayd's friend met him [= Zayd]'

Sentence (3) is derived from (3a) by pronominalization of the first occurrence of *zayd*.

(3a) *laqiya ṣadīq-u zayd-in*
met friend-NOM Zayd-GEN
zayd-an
Zayd-ACC
'Zayd's friend met Zayd'

Pronominalization here is obligatory, since the two identical nouns in (3a) might be wrongly interpreted as having disjoint reference. The reason for pronominalizing the first rather than the second occurrence of *zayd* is that the second occurrence, being an accusatival noun separated from its verbal head by the subject, cannot be easily pronominalized. Pronominalization thus leads to the unacceptable sentence (3), which then obligatorily undergoes word-order inversion, to yield (3b):

(3b) *laqiya zayd*$_i$*-an ṣadīq-u-hu*$_i$
met Zayd$_i$-ACC friend-NOM-his$_i$
'Zayd's friend met him'

For a discussion of a Classical Arabic verse analogous in referential relationship to (3), see Peled (1992:104, n. 9); for a discussion of this type of cataphora in English, see Bosch (1983:160).

In principle, pronouns should be conceived of as pro-constituents (cf. Radford 1981:63-64) rather than as pronouns. Within the genitival construct, the pronoun is usually linked to the whole noun phrase rather than to the genitival complement. Yet, in this regard a distinction must be drawn between anaphora and cataphora. For, while anaphoric pronouns may, in certain cases, be interpreted as referring to a genitival noun, such an interpretation is excluded in the case of cataphora. Compare the unacceptable (4a) and (5a) below with their respective anaphoric b-versions (see Peled 1992:99–102 for further details):

(4) a. **fī ġurfat-i-**hi***$_i$ *yajlisu ṣadīq-u*
in room-GEN-**his**$_i$ sits friend-NOM
zayd$_i$*-in*
Zayd$_i$-GEN
'Zayd's friend sits in his [= Zayd's] room'

b. *ṣadīq-u zayd*$_i$*-in yajlisu fī*
friend-NOM Zayd$_i$-GEN sits in
ġurfat-i-hi$_i$
room-GEN-his$_i$
'Zayd's friend sits in his [= Zayd's] room'

(5) a. **laqiya-**hu***$_i$ *ṣadīq-u zayd*$_i$*-in*
met-**him**$_i$ friend-NOM Zayd$_i$-GEN
'Zayd's friend met him [= Zayd]'

b. *ṣadīq-u zayd*$_i$*-in laqiya-hu*$_i$
friend-NOM Zayd$_i$-GEN met-him$_i$
'Zayd's friend met him [= Zayd]'

It might be argued that, out of context, the anaphoric relationship in (4b) and (5b) is more likely to be interpreted as in (4c) and (5c) respectively:

(4c) *ṣadīq$_i$-u zayd-in yajlisu fī ġurfat-i-hi$_i$*
(5c) *ṣadīq-u zayd$_i$-in laqiya-hu$_j$*

In (4c), the pronoun is coindexed with *ṣadīq* rather than with *zayd*; in (5c), *zayd* and the pronoun *-hu* are disjoint in reference; the pronoun in this case selects its referent from outside the sentence. This, however, does not rule out (4b) and (5b), where the pronoun refers backwards to a genitival complement rather than to the whole noun phrase. For a Qurʾānic verse in which an anaphoric pronoun refers to a genitival noun, see Q. 74/31.

In sentences (1a–c), which are simple sentences displaying the VS word order, the cataphor is a genitival pronoun attached either to an accusatival or to a genitival noun. This, however, is not necessarily the case when the cataphoric pronoun occurs within a clause occupying sentence-initial position. This can be seen in (2a), where the genitival pronoun is attached to a noun in the nominative, and in (2b, c), where the cataphor is a nominatival pronoun implicit in a verb.

Moreover, there is evidence to suggest that in Classical Arabic a cataphoric pronoun may occur within the first conjunct of a compound sentence without any dislocation involved. Typically in such cases, the cataphor is a nominatival pronoun implicit in a verb; the antecedent occurs explicitly in the second VS conjunct:

(6) *ḍaraba-nī* *wa-ḍarabtu* *zayd-an*
hit (**he**$_i$)-me and-I hit Zayd$_i$-ACC
'Zayd hit me and I hit him'

However, a more common version would be (6a), displaying an anaphoric rather than a cataphoric relationship between pronoun and antecedent:

(6a) *ḍaraba-nī* *zayd$_i$-un* *wa-ḍarabtu-hu$_i$*
hit-me Zayd$_i$-NOM and-I hit-him$_i$

For further discussion of sentences such as (6) and related structures, see Peled (1992:99–100).

Bibliographical references
Bosch, Peter. 1983. *Agreement and anaphora: A study in the role of pronouns in syntax and discourse.* London: Academic Press.
Peled, Yishai. 1992. "Cataphora and *taqdīr* in medieval Arabic grammatical theory". *Jerusalem Studies in Arabic and Islam* 15.94–112.
Radford, Andrew. 1981. *Transformational syntax: A student's guide to Chomsky's extended standard theory.* Cambridge: Cambridge University Press.

Yishai Peled (Tel Aviv University)

Causal Clause → Subordination

Causative

1. Defining the causative construction

The causative construction is a linguistic structure rendering an event viewed by the speaker as being causally dependent on some other event, although not overtly specified (Kemmer and Verhagen 1994:117). Accordingly, the following Literary Arabic sentences express causative constructions:

(1) *jaʿala zaydun hindan taktubu*
'Zayd made Hind write'
(2) *ʾajlastu-hu*
'I made him sit down'
(3) *kasarta l-ġuṣna*
'you broke the branch'

All of these sentences involve events (Hind's writing, his sitting down, the branch breaking) brought about by some other events. The nonspecification of the actual causing event is necessary to differentiate between causative constructions and other modalities of expressing causality. It serves to discard as non-causative expressions like the following:

(4) *dafaʿa-nī fa-saqaṭtu*
'he pushed me so that I fell'
(5) *ʾamara l-maliku l-xādima bi-l-xurūji fa-xaraja*
'the king ordered the servant to go out and he went out'

To be accepted as causative constructions, these expressions have to be reformulated as follows:

(6) *jaʿala-nī ʾasquṭu*
'he made me fall'
(7) *ʾaxraja l-maliku l-xādima*
'the king made the servant go out'

However, a causative making no specification about the way the causativization is produced is hard to find, since force dynamic relations (permission, enablement, etc.) are usually incorporated in the causing predicate (Kemmer and Verhagen 1994:118).

2. SYNTACTIC CONSIDERATIONS

Causative constructions are valency-increasing operations. This was recognized by the ancient Arab grammarians who used one and the same term to denote both transitivity and causativity: *at-taʿdiya* (Larcher 2003:54; → *taʿaddin*). The valency-increasing process is illustrated through the following examples:

(8) *ḍaḥika ʿamrun*
"Amr laughed' (Saad 1982:68)
(9) *ʾaḍḥaka zaydun ʿamran*
'Zayd made ʿAmr laugh' (Saad 1982:68)
(10) *samiʿa zaydun ṣawtan*
'Zayd heard a voice' (Saad 1982:69)
(11) *ʾasmaʿat hindun zaydan ṣawtan*
'Hind made Zayd hear a voice' (Saad 1982:69)

In (8), the verbal predicate *ḍaḥika* 'laughed' is monovalent, i.e. intransitive, since it has only one nuclear argument, the subject, "Amr'. Its causative counterpart (9) displays a bivalent (transitive) verbal predicate *ʾaḍḥaka* 'made laugh'. Its nuclear arguments are the subject, 'Zayd', and the direct object, "Amr'. Sentence (10) contains a non-causative transitive verb *samiʿa* 'to hear' having a subject *zaydun* 'Zayd' and a direct object *ṣawtan* 'a voice'. In (11), the causative counterpart of (10), a causative trivalent (double transitive) predicate appears: *ʾasmaʿat*. It has three nuclear arguments: the subject, 'Hind', and two direct objects, 'Zayd' and 'a voice'. Literary Arabic allows the doubling of the syntactic position of the direct object and, as a consequence, both the first direct object – the causee – and the second direct object of the caused predicate appear in the accusative. This is the only strategy Literary Arabic employs for rendering transitive causative constructions. Semantically-determined variation in the case marking of the causee does not exist.

3. THE TYPOLOGY OF CAUSATIVE CONSTRUCTIONS

A criterion used for formally classifying the causative constructions is the degree of grammatical fusion between the predicate of the causing event, and the predicate of the caused event. Accordingly, causative constructions are divided into lexical, morphological, and analytic (periphrastic). Lexical causatives represent the maximum degree of fusion. It is so great that the two predicates are not overtly discernible (Kemmer and Verhagen 1994:118). In morphological causatives the degree of fusion between these two predicates is looser so that the causing predicate is overtly distinguishable, surfacing as a derivational formative attached to the affected predicate. Analytic causatives display the loosest degree of fusion between the two predicates so that they appear as separate units. However, as is always the case with linguistic categories, they are fuzzy rather than discrete. This frequently makes the classification difficult. In order to overcome these difficulties productivity is taken into consideration. Accordingly, causative constructions are classified as either lexical or derivational. Lexical causatives are causatives generated from their non-causative counterparts by nonproductive means and, consequently, they have to be learnt by the speaker and specified as separate entries in the dictionary. The morphological causatives are those causative constructions that can be regularly generated from their non-causative counterparts by a derivational formative. However, even this latter criterion is gradient because there are not only cases of nonproductivity as opposed to absolute productivity but also different degrees of productivity depending on a given formative.

Causative constructions can also be categorized on semantic grounds. The most useful semantic criterion appears to be conceptual distance, namely the physical distance between the causer and the causee and the temporal distance between the cause and its effect. In line with the iconicity principle, the conceptual distance is mirrored in the degree of fusion between the morphemes expressing the cause and its effect.

4. Causatives in Arabic

Classical Arabic possesses lexical, morphological, and analytical causatives. In the majority of the Arabic dialects the morphological Form IV causative has disappeared except in some dialects (Ingham 1982). The productivity of the Form II causative has also been considerably reduced. New morphological causatives developed in the → Ḥassāniyya dialect (with a prefix *sa-*, of Berber origin) and in → Ki-Nubi creole Arabic (with a suffix *-isha*, borrowed from the Bantu languages). Many dialects have developed their own analytical means of expressing causativity.

5. Lexical causatives in Literary Arabic

Due to the fact that lexical causatives belong to the lexicon and not to the grammar, there is little to be said about them. Examples of lexical causatives are: *qatala* 'to kill', *kasara* 'to break', *mazaqa* 'to tear apart', and *hadama* 'to tear down'.

Literary Arabic also has some verbs that fall somewhere in between the lexical and the morphological causatives. On formal grounds, these verbs can be divided into two categories: the labile (or ambitransitive) verbs, showing no trace of derivation, and the Form I causative verbs, apparently derived by internal flection.

Labile verbs can appear in either intransitive or transitive constructions, apparently without a valency of their own:

(12) *naqaṣa d-dirhamu wa-naqaṣtu-hu*
'the dirham diminished and I diminished it'
(Larcher 2003:40)

Their valency is rather dependent on the context: if they appear with only one nuclear argument they are intransitive, and if they occur with two nuclear arguments they are transitive. Other examples of labile verbs are: *ḥaṭṭa* 'to get down [intrans.], to put (down)'; *dāra* 'to turn [intr., trans.]'; *ḍabḍaba* 'to swing [intrans., trans.]'; *zāda* 'to grow, to make grow'; *rafata* 'to be broken, to break into small pieces'; *ʿadala* 'to be equal, to make equal'; *fatana* 'to be charmed, to charm' (Saad 1982:66). Labile verbs are usually considered to be instances of lexical causatives (Shibatani 2003:3). However, yet another analysis is possible, namely that they are morphological causatives derived by means of a zero formative (Saad 1982:66). Saad considers the direction of derivation to be non-causative > causative. However, there is no evidence for such a direction. Apparently, he chooses this analysis because it confers a systemic character on an alleged morphological derivation of a causative Form I verb by internal flection. The fallacy of this view is attested by the very existence of labile verbs which, in the prefix conjugations, do not have *a* as the vowel of the second consonantal root, as implied by Saad (1982), but rather *u*, as attested by the verb *naqaṣa/yanquṣu* or *i*, for instance *rafata/ yarfitu*. Therefore, the use of a zero formative correlated with the lack of a cross-linguistic directionality for the causative vs. non-causative derivation makes it impossible to determine the direction of the derivation. This and the limited productivity of labile verbs impose their classification as lexical rather than morphological causatives.

Form I causatives have a triconsonantal root and two vocalic templates: *faʿala* (suffix-conjugation)/*yafʿalu* (prefix-conjugation), and *faʿula* (suffix-conjugation)/*yafʿilu* (prefix-conjugation). The latter template is very rarely encountered. Form I causatives are considered to represent an instance of morphological derivation (Saad 1982). Accordingly, internal flection would serve to derive from non-causative Form I verbs, having the templates *faʿala, faʿila*, and *faʿula*, causatives with a template *faʿala*, and from *faʿula* non-causative Form I verbs causative verbs *faʿila: ḥazina* 'to be sad' > *ḥazana* 'to make sad', *xabula* 'to be insane' > *xabala* 'to make insane', *fatana* 'to be charmed' > *fatana* 'to charm', *karuha* 'to be hateful' > *kariha* 'to hate' (Saad 1982:66). However, the existence of two vocalic templates with one and the same semantic value as well as the non-predictability of their use and their reduced productivity are all diagnostic of a lexical formation rather than a derivational one. In fact, they are lexicalizations of different nominal forms (the participial and the verbal adjective *faʿa/i/ul-*) of one and the same basic verb and not morphological derivations. Their emergence was triggered at a much earlier stage, in Western Semitic, by the inclusion of a Proto-Semitic nominal construction, the stative, into the verbal paradigm, i.e. the new perfect. Support for this hypothesis comes from the fact that the two templates of the alleged Form I morphological causatives differ from each other in only one respect, just like the verbal nominalizations they have originated from: the vocalization of the second root consonant.

6. Morphological causatives in Literary Arabic

Morphological causativization is restricted to triconsonantal verbs, but there are also situations when a causative verb seems to be derived from an adjective rather than a verb. This is ascertained by the fact that in such cases there is no corresponding basic form. An example of this is the causative Form II *bayyaḍa* 'to whiten', which seems to be derived from the adjective *ʾabyaḍ-* 'white'. In such an instance, the basic verb has probably disappeared, as the result of the competition with a semantically very similar form, the Form IX verb *ibyaḍḍa* 'to be white'. The morphological causative formatives can also be used for creating denominative verbs. In such events they have as their base for derivation a noun rather than a primitive verb, e.g. *tarraka* 'to Turkify, Turkicize', from *turk-* 'Turk'.

Literary Arabic has two morphological causatives: the derived Form II (or D stem) and the derived Form IV (or H stem), although in some cases Form III (e.g. *nāʿama* 'to make happy') and Form X (e.g. *istaxraja* 'to extract') can express causativity as well.

Form II is derived by a reduplicative formative (called *taḍʿīf* by the Arab grammarians). The second consonant of the root serves as base for → reduplication: *qaṣṣara* 'to make short' (< *qaṣura* 'to be short'), *farraḥa* 'to make glad' (< *fariḥa* 'to be glad'), *jammada* 'to freeze' [trans]) (< *jamada* 'to freeze' [intrans.]) (Saad 1982:66). Form II originally expressed verbal plurality, i.e. action repeated over a period of time. From this meaning a secondary one, causative, developed, most probably through a sociative (assistive) construction (for an analysis of the phenomenon in a generative framework, see Fassi Fehri 2001). A comparison of a Classical Arabic text from the first half of the 9th century C.E. with a modern one shows that roughly half of the occurrences of Form II verbs have causative meaning (Măcelaru forthcoming). Sometimes, the pluractional value is still visible: compare Form II *ʿallama* 'to teach' with Form IV *ʾaʿlama* 'to make know, to communicate'.

Form IV is derived from the basic verb by a prefix *ʾa-*: *ʾasmaʿa* 'to make hear' (< *samiʿa* 'to hear') *ʾakrama* 'to honor' (< *karuma* 'to be noble hearted'), *ʾabṭala* 'to nullify' (< *baṭala* 'to be void') (Saad 1982:67). In both formations, the vowel between the first and the second root consonants is always *a*, while the vowel between the second and the last root consonant is *a* in the suffix-conjugation and *i* in the prefix-conjugation. In the forms of the pre-fix-conjugation, the vowel of the prefix is *u*. When directly preceded by the vowel *u* of the subject-prefix, the Form IV prefix *ʾa-* is elided: **yuʾafʿilu* > *yufʿilu*. Historically, the prefix *ʾa-* originated in the Proto-Semitic causative prefix **ša-*, as a result of a morphologically-conditioned phonological change **š* (phonetically [ꞩ]) > h > ʾ /_V. Traces of it are still extant in Classical Arabic: *saqlaba* 'to throw down' (< *qalaba* 'to turn around, to topple') (Fleisch 1979:II, 282), *harāqa* 'to pour out, to spill' (< *rāqa* 'flow out') (Fleisch 1979:II, 283), *yuʾakramu* for *yukramu* 'he is honored' (< *karuma* 'to be noble') (Fleisch 1979:II, 281).

It is widely accepted that Form II causatives can be derived only from intransitives, but the occurrence of notable exceptions – cognitive verbs (*darasa* 'to learn', *ʿalima* 'to know', *fahima* 'to understand'), ingestive verbs (*ʾakala* 'to eat', *šariba* 'to drink'), verbs that do not seem to reduce to a common semantic feature (*ḥamala* 'carry' [Saad 1982:69], *kataba* 'to write' [Fassi Fehri 2001:13]) – invalidate this assumption. Actually, Form II causatives may be freely derived from any verb as long as it is not a lexical causative.

In contrast, the derivationality of Form IV causatives does not show any restriction: *ʾafraḥa* 'to gladden' (derived from a stative verb), *ʾamāta* 'to make die' (derived from an unaccusative intransitive), *ʾasbaḥa* 'to make swim' (derived from an unergative intransitive), *ʾabnā* 'to make build' (derived from a transitive), *ʾaqtala* 'to make kill' (derived from a lexical causative).

The difference the two morphological causatives display in their derivability is due to the different nature of the formatives they use: Form IV employs a concatenative formative, a prefix, while Form II uses a nonconcatenative one, reduplication. The degree of fusion between a base and a nonconcatenative formative is so high that the resulting word is perceived as monomorphemic. Therefore, the redundant assignment of a grammatical or semantic value (in this case that of CAUSE) to a morpheme is ruled out, since it would be superfluous.

Form II expresses a more direct causation than Form IV (Leemhuis 1973, 1977). This is proved by the following facts:

i. The causer of Form II controls the causation, while the causer of Form IV does not:

(13) ***fassada min ḥayṯu yurīdu l-ʾiṣlāḥa*
'he caused mischief where he meant to make amends' (Saad 1982:74)

The incorrectness of (13) is due to the logical contradiction it expresses: Form II causative *fassada* implies intentionality and, therefore, a situation where mischief is caused on purpose. The only logical possibility that a person causes mischief by trying to make amends is that he does it involuntarily. In such an event, Form IV *ʾafsada* must be used because it implies unwilling action.

ii. Form II may imply coercion, while Form IV does not. Instead, the latter may have a permissive or an assistive reading. Thus, compare the following pairs: *kattaba* 'to make (someone) write by using force, or against his will', *ʾaktaba* 'to dictate'; *kassā* 'to clothe', *ʾaksā* 'to give clothing'.

iii. Usually, the causee of Form II is affected, while the causee of Form IV is not: *kaṯṯara* 'to multiply [by changing the internal structure of the causee, as for instance by dividing it]', *ʾakṯara* 'to multiply [by adding]'; *baddala* 'to change (something)', *ʾabdala* 'to change something for something else, to replace'. However, there are cases when the causee of Form II is not at all affected, as shown in the following sentence:

(14) *ʿullima wa-lam yataʿallam*
'he was taught but he did not learn' (Wright 1896:I, 38)

Although these semantic differences between Form II and Form IV causatives are often visible, there are also instances where such a difference cannot be detected.

The two morphological causatives can also denote a special category of causation, in which the causation takes place at the speaker's mental level and does not have to be real. This causation is represented by the estimative-declarative (tropative as suggested by Larcher 2003:60–61) verbs like *ṣaddaqa* 'to believe someone' (< *ṣadaqa* 'to tell the truth'), *ʾaʿḏ̣ama* 'to consider great' (< *ʿaḏ̣uma*) 'to be great'.

7. Analytic causatives in Literary Arabic

To render causation analytically, Literary Arabic uses the verb *jaʿala* 'to put, to make'. This verb is followed by two direct objects, the former expressing the causee and the latter the caused event. The causee is represented by a noun or a personal pronoun in the accusative. The caused event is rendered either through a finite verbal form (15) or through an active participle (16) or an adjective (17).

(15) *jaʿala zaydun al-walada yaxruju*
'Zayd made the boy go out'
(16) *jaʿala zaydun al-walada xārijan*
'Zayd made the boy go out'
(17) *jaʿala zaydun al-ʾamra sahlan*
'Zayd made the issue facile'

Normally, the analytic causative denotes a less direct causation than the morphological causatives. This is shown by confronting (18), which contains an analytic causative, with (19), which has a Form IV, the least direct morphological causative:

(18) *jaʿala zaydun hindan taḍḥaku*
'Zayd made Hind laugh' (Saad 1982:82)
(19) *ʾaḍḥaka zaydun hindan*
'Zayd made Hind laugh' (Saad 1982:82)

(19) implies that Zayd made Hind laugh by doing a funny thing, while (18) describes a situation where the causer has sent somebody else to make Hind laugh (Saad 1982:82).

The analytic causative is obligatory when a certain morphological causative may not be used with an inanimate causer (the reason for such a restriction seems to be purely lexical). Such a case is illustrated in (20):

(20) *jaʿala l-maṭaru ṭ-ṭifla yanzilu ʿani š-šajarati*
'Rain made the child go down from the tree'
(Saad 1982:81)

If, instead of the analytic causative, the morphological causative *ʾanzala* 'make go down' had been used in (20), the sentence would have been incorrect.

Bibliographical references

Fassi Fehri, Abdelkader. 2001. "Causativity, transitivity, and iterativity as pluralities". *Linguistic Research* (Rabat) 6.1–49.

Fleisch, Henri. 1979. *Traité de morphologie arabe*. II. *Pronoms, morphologie verbale, particules*. Beirut: Dar al-Machreq.

Ingham, Bruce. 1982. *North east Arabian dialects*. London and Boston: Kegan Paul International.

Kemmer, Suzanne and Arie Verhagen. 1994. "The grammar of causatives and the conceptual structure of events". *Cognitive Linguistics* 5:2.115–156.

Larcher, Pierre. 2003. *Le système verbal de l'arabe classique*. Aix-en-Provence: Université de Provence.

Leemhuis, Frederik. 1973. "Sībawaih's treatment of the D stem". *Journal of Semitic Studies* 18. 238–256.

——. 1977. *The D and H stems in Koranic Arabic*. Leiden: E.J. Brill.

Măcelaru, Adrian. Forthcoming. *Causative in Literary Arabic*.

Premper, Waldfried. 1987. *Kausativierung im Arabischen*. (= *Arbeiten des Kölner Universalien-Projekts*, 66). Cologne: University of Cologne, Institut für Sprachwisschenschaft.

Saad, George Nehmed. 1982. *Transitivity, causation and passivization: A semantic-syntactic study of the verb in Classical Arabic*. London: Kegan Paul International.

Shibatani, Masayoshi (ed.). 2002. *The grammar of causation and interpersonal manipulation*. Amsterdam and Philadelphia: J. Benjamins.

Shibatani, Masayoshi. 2002. "Introduction: Some basic issues in the grammar of causation". Shibatani (2002:1–22).

—— and Prashant Pardeshi. 2002. "The causative continuum". Shibatani (2002:85–126).

Wright, William 1896–1898. *A grammar of the Arabic language*. 2 vols. Cambridge: Cambridge University Press.

Adrian Măcelaru
(University of Bucharest)

Chad Arabic → Subsaharan Arabic

Chad Arabic

More than 850,000 Arabs live in the Sahelian zone of Chad. Chadian Arabic is a vehicular language spoken by about 60 percent of the population. It has its own characteristics, which come from the rural nomadic society in which it originated and the society of city-dwellers in which it developed.

1. A mosaic of peoples

The population of Chad, estimated at 8 million persons, is a true ethnic mosaic of diverse peoples spread unequally over a territory of 1,284,000 square kilometers. Several groups of peoples can be distinguished in the three climatic regions of the country.

i. The 'Sara' group in the prefectures of Middle Chari, Western Logone, and Eastern Logone, and the populations of the Mayo-Kebbi and Tandjilé make up about 47 percent of the total population. They live in the tropical region of the south that occupies 10 percent of the country's surface.
ii. The Arabs constitute 11 percent of the total population of Chad. They live amidst many other ethnic groups in the Sahel region which comprises approximately 44 percent of the nation's territory consisting of the following prefectures: Salamat, Guéra, Chari-Baguirmi, Lac, Kanem, Batha, Ouaddaï, and Biltine.
iii. Finally, the Goran, Teda, and Zaghawa ethnic groups, who make up 2 percent of the population, live in the northern prefectures (Borkou, Ennedi, Tibesti, called B.E.T.) in the sparsely populated desert region that covers about 46 percent of the territory of Chad.

2. The Arabs in Chad

The first witnesses to the presence of a sedentary Arab population living in the region of Lake Chad are the explorers Denham and Clapperton in 1823. There are no documents that can testify with certainty to the route followed by the first Arabs to Chad. Nevertheless, according to a still vibrant oral tradition transmitted through stories, epics, and genealogies, the Arabs of Lake Chad, Chari-Baguirmi, and Ouaddaï claim that their ancestors came from Yemen. The earliest written documents, attesting to the presence of Arabs in Chad come from the Arab geographers. A few references go back to the 8th century, but more abundant and accurate records date to the 11th and 12th centuries, with the great works of ʾAbū ʿUbayd al-Bakrī (d. 1094) and al-ʾIdrīsī (d. 1166).

The first Arabs who arrived in Chad came from the Sudan. A first Arab Muslim migration to Egypt took place in 639. "Africa had seen the arrival on its continent of an army of 12,000 foot soldiers and 4,000 horsemen commanded by the general ʿAmr b. al-ʿĀs (d. 42/663)" (Zeltner 1970:114), of the tribe of Qurayš, whose forces reached the borders of Nubia. Up to the 14th century, the great mass of Arab nomads of Egypt remained confined to the region of Aswan. The taking of Dongola by the Mamluks and the fall of the Christian kingdom of Nubia in 1316 led to the most important Arab migration to the Sudan. This was followed in 1504 by a second wave of migration after the fall of the kingdom of ʿAlwa (Sōba), the last Christian bastion in the Sudan. The Arabs then progressed quickly in two directions. Some followed the course of the White Nile and the Blue Nile, while others moved westward toward the southeast frontier of present-day Chad. Among the latter were the Juhayna Arabs.

The Juhayna trace their ancestry through ʿAbdullāh al-Juhaynī all the way to ʿAbd al-Muṭṭalib, grandfather and tutor of the Prophet Muḥammad. "These Arabs claim to be originally from the Yemen and to be descendants of 72 tribes who came to be established in Egypt. They had left their country in order not to pay the tax imposed during the reign (717–720) of the Umayyad Caliph ʿUmar II (b. ʿAbd al-ʿAzīz)" (Carbou 1912:II, 46). They arrived in Chad in the 15th–16th centuries and consider themselves descendants of the four sons of Juhaynī: Ḥamāt, ʿAṭiyya, Barqa Salām, and Rašīd. They thus constitute fractions of the most important tribes in central and eastern Chad: the Wulād Hēmāt, Wulād Atīye, the Salāmāt, and the Wulād Rašīd.

Other Arabs may have arrived in Chad from the north. Three ancient caravan routes linked the Chad valley to the Mediterranean. The Arab geographers knew the routes and provide valuable information when they speak of the Fezzan and discuss relations between the Sahara and the *Dār as-Sūdān*. However, they give very little precise information about the presence of Arab populations located in the north of Chad. They note the presence in the 10th century of Arabs in the Kawar region, which lies on the caravan route linking Lake Chad to the ports of the Mediterranean. In the 11th century, al-Bakrī relates that Kanem was populated by 'unbelieving idol-worshippers' and believes that there also used to exist in this region some descendants of the Umayyads. But this does not constitute conclusive evidence of an Arab presence in the Chadian-Libyan desert before the Muslim conquest. Shortly thereafter, the Muslim kings of Kanem are said to be descendants of the Yemeni Sayf ibn D̠ī Yazan, the last of the Himyarite kings who had freed Yemen from Ethiopian occupation in 575 C.E., and who is said to have come to the region by the Fezzan. Ibn Saʿīd al-Muġrabī (1214–1286) attests to the vitality of this oral tradition in the 12th century. Many have been deceived by this prestigious ancestry, but today historians agree in recognizing that the descendants of Sayf, the Banū Sayf, are not Arabs. "The *Banū Sayf* do not bear the name of an Arab tribe. Moreover, their genealogies are not constructed according to the pattern of Arab genealogies" (Zeltner 1980:44).

However, the Tunjur are of Arab origin and are descendants of the Banū Hilāl. It seems that they had arrived in Africa even before the coming of Islam. They had established themselves in Nubia, where they had accepted Christianity and then moved to the west. Then, toward the end of the 10th century, the Caliph al-ʿAzīz deported them to Upper Egypt. By the middle of the 11th century, in 440 A.H., they were brought to Tunisia by the Fatimid Caliph al-Mustanṣir bi-llāh (d. 487/1094). Even though neither the exact date nor the reason for their return to the south is known, the Tunjur left Tunis and went down to Chad to the region west of Darfour. The Tunjur marked the history of Kanem, Ouaddaï, and Darfour in the 16th and 17th centuries until they were turned back by the scholar and holy man, ʿAbd al-Karīm. Oral tradition in Chad today still evokes a nostalgia for the *tūnes al xadra* (*tūnis al-xaḍrāʾ* 'Tunis the green') or the land of Kerawān (Qayrawān = Kairouan).

The Ḥasawūna consider themselves the descendants of Ḥasan aṣ-Ṣaġīr al-Ġarbī. Their ancestors allegedly passed through Egypt and continued westward to the region of Tripoli. From there they headed south to the region of Lake Chad, where they were given the name of Šuwa Arabs (i.e. ‘owners of sheep’). The Ḥasawūna also bore the surname of Xawalme (Ġawalma). Their oral and written traditions are unanimous in tracing the origin of the tribe to ʿAlī ibn ʾAbī Ṭālib, son-in-law of the Prophet Muḥammad. Their presence in the Lake region is attested as early as the 17th–18th centuries. At this period they had to submit to the domination of the Tunjur in Kanem. They raised camels in the north and lived as nomads in the vast triangle formed by Lake Chad, Lake Fitri, and the Bahr al-Ghazal.

The Wulād Slimān are the children of Sulaymān, who was said to have been charged by the Prophet with bringing Islam to Tripoli. They are the last group of Arabs to arrive in the region of Chad. They are said to have come to Africa at the same time as the Banū Hilāl, in the 11th–12th centuries, passing through northern Egypt before establishing themselves in the Fezzan. They arrived in Chad by way of Tibesti, raising camels and following the length of the great caravan and commercial routes from the desert region of the north. In Chad they are called Fezzanis. They are light-skinned and form a group apart.

The Arabs of Chad are thus found throughout the Sahel region, which covers the whole area north of the Chari River. Table 1 indicates the names of the main tribes and fractions of tribes, as well as their present-day geographical location. (The place names follow the spelling of the maps of Chad).

Table 1. The Juhayna tribes and tribal fractions in Chad

Juhayna tribes	Fractions of tribes	Geographical location
1. Wulād Atīye	Irēgāt	Ati
	Alawne	Oum-Hadjer
	Misirīye humur	Oum-Hadjer
	Misirīye zurug	Am-Timan
	Rizēgāt: Šittiye	Eastern Chad
	Mahāmīd	Arada
	Mahriye	Arada
	Nawaybe	Eastern Chad
2. Wulād Hēmat	Ja’āme	Batha
	Wulād Himād	Batha
	Wulād Salmān (Salmāniye)	Batha
3. Wulād Rašīd	Azid	Abou-Deïa
	Hamīde	Abou-Deïa
	Zabada	Lake Fitri
	Zuyūd	Oum-Hadjer
	Dār Sālem	Am-Timan
4. Salāmāt	Dār Bagli	Chari-Baguirmi
	Wulād Eli	Lake Chad
	Iyēsiye	Salamat
	Wulād Isa	Chari-Baguirmi
	Hammādiye	Massenya
	Wulād Humrān	Guéra
	Wulād Alwān (Alawne)	Salamat
	Sifēra	Salamat
	Sa’ādne	Salamat

Table 2. The Wulād Hasan or Ḥasawūna tribes in Chad

Ḥasawūna tribes	Geographical location
1. Bani Wā’il	Bahar-al-Ghazal, Kanem
2. Wulād Maharēb	Dagana, Bornou, Kanem
3. Wulād Sarrār	Bornou, Kanem
4. Dagana (Wulād abu Digin)	Dagana
5. Wulād abu Xidēr	South bank of Lake Chad
6. Wulād Emir	Bornou
7. Wulād Xānem	Chari-Baguirmi
8. Wulād Sālem	Chari-Baguirmi
Connected with the Ḥasawūna tribes:	
Bani Sēd (from the Bani Wā’il)	Chari
Hemmādiye	South bank of Lake Chad
Wulād Abu Īse	South bank of Lake Chad
Wulād Bilāl	North Manga
Wulād Tālib	Massakory
Wulād Alwān	Massakory
Nawala	Massakory
Asala	South bank of Lake Chad
Šidērāt	South bank of Lake Chad

The number of Arabs in Chad is estimated at 880,000. These Arabs do not form a united group. Until today, their history has been a succession of conflicts with the peoples to the north of the Sahel, internal strife, rebellion against the political power, or challenging of religious authority.

However, they recognize a ‘foundational event’: the camel of an Arab of one tribe was stolen by an

Arab of another tribe, which provoked a bloody battle among the tribes. The tribes then moved apart and gradually dispersed throughout the whole territory of Chad. They also identify themselves as distinct from others by their genealogy. An Arab is able to cite the bonds that link him to his fraction of a tribe and then to the tribe itself, and finally can trace his blood line all the way to the Prophet. From this point of view, one could say that only the nomadic Arabs have been able to preserve their 'Arab' character.

The 'great nomads' *abbāla*, like the Wulād Slimān, the 'small nomads' *baggāra*, like the Wulād Rašīd, the partially sedentary nomads like the Arabs of Kanem, Guéra, Salamat, or the region of Abéché, pasture their flocks between Lake Chad and Waddaï-Darfour, and traverse the Sahel region from north to south between the 11th and 15th parallels. The nomadic Arabs are becoming ever less numerous. More and more they are settling near towns. Their very rapid socioeconomic transformation leads one to expect that within a few years their sedentarization will be complete. This great movement has been accompanied by an interethnic mingling and assimilation among the Arab tribes themselves. The sedentary Arabs forget their history and lose their tribal identity, but they have imposed their language wherever they have established themselves.

Table 3. The Wulād Slimān tribes in Chad

The Wulād Slimān tribes	Geographical location
1. Jebayr	Nokou, Kanem
2. Hemāt	Kanem
3. Myasa	Nokou
4. Šeredāt	Nokou

3. Chadian Arabic

The Arabic dialects in Chad belong to the Sudanese dialects. The nomadic Arab tribes who have come to Chad still keep their Arabic dialect very much alive. They are characterized by their dialect and often differentiated by it. There are about 30 Arabic vernaculars, denoted by the name of the tribe or tribal fraction by which they are spoken.

The dialect of nomads is distinguished from that of sedentary peoples. The dialect of an Arab nomad and that of a sedentary Arab belonging to the same tribe show more differences between them than one would find between two dialects of different sedentary tribes in the same territory. Thus, it is the 'regional' geographical factor rather than the 'tribal' sociological factor that today characterizes the various sedentary Arabic dialects. The mutual comprehensibility between the Arabic vernaculars is total. The variations are mainly phonetic, but the grammatical structure of the dialects is similar; it is only the use of vocabulary proper to the region that gives each dialect its distinctiveness. The lexical particularities tend to disappear in favor of a generalized Chadian Arabic due to the breakdown of traditional society, ethnic mixing, urban concentration, rural exodus, and the media.

In Chad, one can distinguish over 100 'national languages', belonging to 12 quite different linguistic groups. This extraordinary linguistic variety has already favored, in the southern part of the country, the use of vehicular languages such as Bāgirmi (*bārma*), → Fulfulde, and Sango.

Chadian Arabic is dominant in the whole Sahel region of the country (44 percent of the national territory). Farther south, Bāgirmi (*bārma*) and Fulfulde, which are spoken in relatively sparsely populated regions, are more and more losing their importance and receding in the face of vehicular Arabic. The prohibition of the use of Sango, enforced by President Tombalbaye in 1973 in the name of 'cultural authenticity', has strengthened the use of Arabic in all the markets of the south. In the same year, the capital city of Fort Lamy was 'Chadized', becoming N'Djaména (from *anjam-mēna* meaning in Arabic 'We have rested.')

Two national languages are likely to retain their importance in the future: Arabic and 'koine Sara'. In the southern part of the country, the languages of the Sara group are spoken by about 20 percent of the population. In an urban setting the various dialects of the group have become progressively harmonized to form a koine Sara with total mutual comprehensibility among the speakers of various languages. This progressive standardization among the languages of the Sara group remains, however, linked to intra-Sara relations, whereas Chadian Arabic is independent of the Arabic-speaking community that introduced the language to Chad.

One can distinguish several levels of usage in the Arabic spoken in Chad.

The lowest level is called 'Bongor Arabic'. This could be characterized as a type of Pidgin Arabic (→ pidginization), such as might be uttered by those who are not native speakers of Arabic in the markets of the Mayo-Kebbi prefecture. There one might hear, for example: *ana oru gal ke, inti maši hinak! Amis ana kutulu kalib al-addu wiled hanai* 'I told you to go there! Yesterday I killed the dog who bit my child.' Even though they would be understood, these phrases would make Arabic-speakers smile, for the syntax is incorrect, the vowel length is not respected, the imperative is replaced by the present tense of the verb, the verbs do not agree with the subject, and the pronominal possessive suffixes are replaced by redundant prepositional complexes. This level of vehicular Arabic corresponds to the 'Tourkou' Arabic described by Muraz (1926). An Arab of Abéché would have said: *nugūl lēk: amši hināk. Anā katalt al-kalib al-adda wilēd amis.*

The second level of vehicular Arabic is that of the sedentary speakers who are not Arabs but are generally Islamicized, who live in an Arab milieu where the demands of work require the use of vehicular Arabic. This is the case of many speakers who live in the prefectures of Chari-Baguirmi, Guéra, and Ouaddai. For them, vehicular Arabic is an indispensable second language of interethnic communication.

For example, in Guéra one might hear *amis, hu kappalani fat. Banati lisa ma mašatom* 'Yesterday, he frightened me and left. My daughters have not yet brushed their hair.' In this example, the syntax and morphology of Chadian Arabic are generally respected, but the pronunciation reveals the influence of a linguistic substratum proper to the speaker's mother tongue. The [x] has often been transformed into [k] , the [b] and the [f] are confused in a single bilabial fricative close to [p], and the [s] is 'lisped'. An Arab of Ati would say: *amis, xaffalāni fāt. Banāti lissā mā maššatōhum.*

The third level of vehicular Arabic is spoken by sedentary peoples whose father and mother are both Arabs and by those who have been Muslims and Arabized for more than two generations. For them, Chadian Arabic is the mother tongue. Such people are able to recognize grammatical errors and the geographical origins of 'variations' in the pronunciation of certain words. They are even able to point out those who speak well and those for whom Arabic "speaks to the heart as well as to the spirit". This is the Arabic of journalists of the Chad National Radio, of certain commentators on Télé-Tchad, and of the translators of official messages.

Like every living language, Chadian Arabic is continually increasing its vocabulary according to the needs of communication and of transmitting messages which are continually more technical and precise. The vocabulary used today by the media employs a great number of words existing in modern Literary Arabic, which become naturally 'Chadized'. For example: *demoxrātiya* 'democracy', *urubba* 'Europe', *mustaxill* 'sovereign', *naxaba* 'trade union', etc. This media Arabic has become progressively the norm of Chadian Arabic. The standardization of its writing could make it an Arabic koine that would take its place along with French and Literary Arabic as languages taught in schools. It is estimated that 60 percent of the population speak Chadian Arabic in order to make themselves understood in the hospital, police station, or market-place. For the majority of Chadians, Literary Arabic is a foreign language they do not understand. It is to this great mass that the National Radio transmits every day, during the prime listening hours, its 'notices and communications' of deaths, marriages, greetings, condolences, invitations, announcements, technical advice, and the like.

The writing of Chadian Arabic with the letters of Classical Arabic varies from one individual to another. Arabizers are tempted to rediscover the etymology of words in the spoken language by writing them with the vowels and consonants of Classical Arabic, without taking into account the results. Such writing thus varies according to the degree of the writer's erudition, and it is generally only legible by its author or by someone possessing the same level of preparation in Literary Arabic.

However, the writing of Chadian Arabic with the standardized symbols of international phonetics poses less of a problem. It is simply a matter of using the letter *c* to write /š/. A critical study of the graph of the 32 phonemes of Chadian Arabic (22 consonants and 10 vowels) allows one to propose an alphabet of 27 signs:

20 symbols for writing consonantal phonemes: *p b f w m t d s z n l r j c y k g x ʾ h* (*tc* and *ny* being digraphs formed from these symbols)

5 symbols for writing vocal phonemes: *a e i o u*
2 supplementary signs: ^ (circumflex accent) to note the long vowels and – (hyphen) to note the assimilation of the article.

This writing has been used since 1988 for the teaching of Chadian Arabic.

In 1978, Arabic became the second official language in Chad along with French, under the pressure exerted by the former Frolinat rebels who came to power with Hissène Habré. The economic and political pressure from the Islamic countries neighboring Chad and from the Gulf countries resulted in developing the teaching of Arabic literature. In 1962, The 'Éducation nationale' counted 2,500 students in schools where teaching was done in Arabic. This number exceeded 40,000 in 1988 and today is more than 200,000 (almost 3 percent of the population). Literary Arabic is an instrument of communication only for a civil, intellectual, religious or commercial minority, but it gives way to Chadian Arabic in everyday conversations.

French was the official language in Chad long before Arabic. Public instruction in French began to be organized in Chad in 1911, and was restructured after the independence of the country in 1960. Despite all the socioeconomic handicaps that Chad experiences, the number of students in 'French schools' continues to increase. Between 1983 and 1990, their number increased from 133,000 to 492,000 in elementary instruction and today the number exceeds 900,000 (13 percent of the population). It is likely that today the Francophone speakers represent 25 percent of the population. French seems to be better learned and spoken in the southern part of the country. In the central region where Arabic is dominant, even though professors in class might be teaching in French, students tend to speak in the vehicular language (i.e. Chadian Arabic) among themselves during recreation, just as they would do elsewhere outside their homes. Training truly bilingual people would avoid having Literary Arabic and French appear as two mutually competitive and exclusive foreign languages. Those who both speak and write both languages are rare.

Chadian Arabic finds itself at the heart of a very complex social, political, and religious situation. Even though it is spoken by more than half the population, the written language is not easily assimilated by the Arab Muslim community that makes up about 11 percent of the population and is ultimately the only group to master this language. However, in this country where, in 1993, illiteracy was estimated at 70 percent of the population, and where the percentage of illiterates reached 90 percent among women, the conclusions of a study by the Éducation nationale showed Chadian Arabic to be an adequate tool for teaching people to read and write a national language. Writing this language with the standardized international characters would facilitate the passage to learning French for those who still do not know it, and would open the way to literary Arabic for those who would like to extend their roots into the Arabic cultural universe.

4. From Classical Arabic to Chadian Arabic

In Chadian Arabic 22 consonants are recognized (Table 4).

A comparison between the consonantal system of Classical Arabic and that of Chadian Arabic shows:

i. the absence of ten consonants of Classical Arabic: the interdentals /t̠/, /d̠/, /ḏ̣/; the emphatics /ṣ/, /ḍ/, /ṭ/; the back consonants: a velar /ġ/ and a uvular /q/; and two pharyngeals /ḥ/, /ʿ/ , which have disappeared in Chadian Arabic.
ii. the presence of four supplementary consonants unknown to classical Arabic: a voiceless labial /p/ corresponding to /b/; a voiceless palatal /tc/ [tʃ] corresponding to /j/ [dʒ], (/tc/ is found as a palatal occlusive, voiceless affricate as in 'church' in English; /j/ is found as a palatal occlusive, strongly voiced affricate as in the word *jinn*); a voiced velar /g/ corresponding to /k/, (/k/ and /g/ are pronounced farther forward in the mouth than is usual in English, closer to the French); a nasal /ny/ [ɲ], which completes the palatal series.

The absence of interdentals is one of the characteristics of sedentary Arabic dialects, while the pronunciation of the /q/ of classical Arabic as /g/ is seen to be a characteristic of the nomadic dialects. Roth (1970–1971:70) has noted the absence of 'emphasis correlation' in the Arabic dialect of Abéché in Chad, that of Malta, and that of a Maronite village of Cyprus, which

Table 4. The consonants of Chadian Arabic

		labials	dentals	palatals	velars	glottals
plosives						
	voiceless	p [p]	t [t]	tc [tʃ]	k [k]	’ [ʔ]
	voiced	b [b]	d [d]	j [dʒ]	g [g]	
fricatives						
	voiceless	f [f]	s [s]	c [ʃ]	x [x]	h [h]
	voiced		z [z]			
nasals		m [m]	n [n]	ny [ɲ]		
vibrants			r [r]			
laterals			l [l]			
semivowels		w [w]		y [j]		

“have in common their being at the geographical, historical and cultural periphery of the Arabic-speaking world.”

The appearance of the /tc/ and /ny/ phonemes and of the nasal velar /n/ [ ], in the common words *tcatca* ‘to deceive’ and *nyang-nyang* ‘to nibble’, show the contact of vehicular Arabic with the other African languages of the Sahel. The voiceless prepalatal affricate /tc/ and the *mouillée* nasal /ny/ can be found in Kanembou and in the languages of Guéra in the center of the country where the Arabs live. In a collection of more than 7,500 different words, 4.6 percent were borrowed, 2.3 percent from neighboring African languages and 2.1 percent from French.

When the Chadian Arabic dictionary was constituted (Jullien 1999b), the various consonant pronunciations of the same root were compared between Chadian Arabic and Classical Arabic. Some constants emerged:

i. In the first group of words, the fricatives of Classical Arabic /ṯ/, /ḏ/ became in Chadian Arabic occlusives /t/, /d/, and the occlusive /q/ of Classical Arabic remained occlusive on becoming /g/ in Chadian Arabic. This first group consisted of concrete words designating man and his environment and characterizing a nomadic society or that of rural pastoral populations engaged in herding, agriculture, and commerce.
ii. In the second group of words, the same fricatives /ṯ/, /ḏ/ of Classical Arabic remained fricatives on becoming /s/, /z/, and the occlusive /q/ of Classical Arabic became a fricative /x/ in Chadian Arabic. This second group is formed of words that have been newly introduced and concern religion, politics, law, economics, civic life, and other concerns of the modern world.
iii. The same root can be found in words belonging to the first or second lexical group depending on the occlusive or fricative articulation of /q/, /ḏ/, or /ṯ/.

For example:

Root	First group	Second group
w-q-f	*istifag* ‘pact’	*istifax* ‘harmony’
ḏ-k-r	*dakar* ‘male’	*zakar* ‘to pray’
ṯ-w-r	*tor* ‘bull’	*sawra* ‘rebellion’
w-f-q	*fagur* ‘poverty’	*faxir* ‘marabout’

5. Particularities of Chadian Arabic

The study of certain Arabic dialects in western Chad have shown the presence of ‘injectives’ and ‘implosives’ (Zeltner 1971:28–29; Hagège 1973: 14–15, 19–20; Zeltner and Tourneux 1986: 16–17) which are unknown in Classical Arabic and also absent from the vehicular dialect. Other particularities such as the agreement with the feminine plural (Hagège 1973:49) are found in other dialects in the eastern part of Chad, but they tend to disappear. Chadian vehicular Arabic exhibits some original characteristics such as the following:

i. The prefix *ba-* (*bi-*, *bu-*) in the conjugation of the imperfect verb is preferred to the prefix *ya-* (*yi-*, *yu-*) “to express the moment in which the action has a good chance of being completed or show the desire of the speaker”

(Jullien 1999a:206): *yagdar yaxadim* 'he can work [but his ability is not known]'; but *bagdar baxadim* 'he can work [and his ability for good work is known]' (Jullien 1999a:206).

ii. The particle *ke* is frequently used to "introduce indirect discourse," to "designate a direction," "to prolong or intensify the action of the verb" (Jullien 1999a:119): *gāl ke anā mardān* 'he said he was sick [lit. 'I am sick.']'; *fat ke* 'he went in that direction'; *yadurbuh ke* 'they strike him at length and with force' (Jullien 1999a:119).

iii. A great number of exclamations, interjections, onomatopoeic words and ideophones are used: *tcā* 'yuck, ugh!; *hey tara* 'watch out, be careful!'; *alburrāga sawwat wilij* 'the lightning flashed in the sky [lit. the lightning did *wilij*]'; *gamma burdulup* 'he got up with a bound' (Jullien 1999a:122–123).

Certain ideophonic adverbs express the absolute superlative of adjectives of color: *abyad karr* 'very white, pure white'; *ahmar tcu* 'very red, scarlet red'; *azrag litt* 'very black, jet black' (Jullien 1999a:79).

iv. The auxiliary *gā'id*, active participle of the verb *ga'ad*, is placed before another imperfect verb to express the progressive form: *alwilēd gā'id yabki* 'the child is crying' (Jullien 1999a:208).

v. The contrast between the different short and long syllables of the spoken language is marked by a tonic accent which is shown by a 'variation of tonal pitch'. This characteristic accentuation, achieved in an 'increasing tonality' of the final vowel in certain words or transitive verbs, even reveals the presence of the personal pronoun suffix *-h* (Jullien 1999a:21): *akalo* [akalo] 'they have eaten'; *akaloh* [akɑ'lo] 'they have eaten it' (Jullien 1999a:21). This supports "the hypothesis according to which the Arabic of Chad, which has very likely known an ancient accentual state, has seen, little by little, the accent of intensity become a pitch, the stressed syllable thus becoming the stressed syllable (the most stressed) of the word" (Décobert 1985:139).

vi. Finally, "the methods of determination which the Arabic dialects do not traditionally know" have been highlighted in the Arabic dialect of Abéché in Eastern Chad (Roth 1979:79). The language has a tendency today to give special place to the determinative complement with *hanā: Madam Amīna Mūsa tiballix lē axawānha hanā Ridīna tugūl lēhum sadaxa hint al arba'īn hanā xālitha tabga yom al-juma'a* 'Madame Amina Musa informs her brothers in the Ridina Quarter that the 40th anniversary sacrifice for the death of her maternal aunt will be held on Friday.' This word *hanā* (*hint* after a feminine noun), no longer signifying 'thing', is only used in vehicular Arabic to form the complement of a noun. It is placed between the terms of annexation and "permits the introduction of nuances in the determinative complement": *bēt mūsa* 'Musa's house, and we know it belongs to him', but *bēt hanā Mūsa* 'Musa's house, in which he lives but we don't know if he owns it' (Jullien 1999a:183–185.)

Chadian Arabic was not born in the big cities and preserves the originality of the ancient dialects of rural life. It shows its distance and autonomy from Literary Arabic, all the while borrowing new vocabulary from the literary language. It presents the characteristics of an Arabic of merchants and nomads coming from the east, into which are blended certain traits proper to nomads and to sedentary peoples.

Bibliographical references

Carbou, Henri. 1912. *La région du Tchad et du Ouadday*. 2 vols. Paris: Leroux.

Décobert, Christian. 1985. *Phonologies arabes du Tchad*. Paris: Geuthner.

Hagège, Claude. 1973. *Profil d'un parler arabe au Tchad*. Paris: Geuthner.

Jullien de Pommerol, Patrice. 1999a. *Grammaire pratique de l'arabe tchadien*. Paris: Karthala.

——. 1999b. *Dictionnaire arabe tchadien–français, suivi d'un index français–arabe et d'un index des racines arabes*. Paris: Karthala.

Muraz, G. 1926. *Vocabulaire du patois arabe tchadien ou 'tourkou' et des dialectes sara-madjingaye et sara-mbaye (sud-ouest du Tchad)*. Paris: Charles Lavauzelle.

Roth-Laly, Arlette. 1971–1972. "Esquisse de la phonologie du parler arabe d'Abéché". *Comptes rendus du Groupe Linguistique d'Etudes Chamito-sémitiques*, 16. Paris.

Roth, Arlette. 1979. *Esquisse grammaticale du parler arabe d'Abbéché (Tchad)*. Paris: Geuthner.

Zeltner, Jean-Claude. 1970. "Histoire des Arabes sur les rives du lac Tchad". *Annales de l'Université d'Abidjan*, F, II, 2.109–237.

——. 1980. *Pages d'histoire du Kanem*. Paris: l'Harmattan.
——, and Maurice Fournier. 1971. *Notice pour suivre un enregistrement en arabe salamat de la région du lac Tchad*. Fort-Lamy: La Source.
——, and Henri Tourneux. 1986. *L'arabe dans le bassin du Tchad: Le parler des Ulād Eli*. Paris: Karthala.

Patrice Jullien de Pommerol (Ankara, Turkey)

Child Bilingualism

1. The many definitions of bilingualism

Childhood bilingualism is the area of language acquisition concerned with the simultaneous or sequential acquisition of two languages. Simultaneous acquisition usually refers to children who receive input in both languages from birth or before their third birthday, while successive or sequential acquisition takes place where input from a second language (L2) is received after the third birthday (Lyon 1996:47). This, however, constitutes only one way of classifying young bilinguals; different researchers have used different terms and different age limits when describing types of bilingualism. De Houwer (1995:223), for instance, suggests the term 'Bilingual First Language' (BFL) acquirers for situations where the child is regularly exposed to two languages within the first month after birth. She argues that situations where regular exposure to an L2 occurs later than one month after birth and before the age of two should be categorized as 'Bilingual Second Language Acquisition'.

Many neurolinguistic and psychological studies have also distinguished various types of bilingualism (→ multilingualism). One of the earliest distinctions was made by Weinreich (1953), who identified three categories of bilingualism: compact, coordinated, or subordinated bilingualism. The expression 'compact bilingual' refers to an individual who has learnt the two languages simultaneously before the sixth year, normally because they were each spoken by one of the parents. A 'coordinated bilingual' has learnt the L2 before puberty, within or outside the family, for example because the child moved to a foreign country. A 'subordinate bilingual' has one language as the mother tongue and uses the L2 as moderator of the first language. In this type of bilingualism, subjects think of what they want to express in their first language and then translate it into their L2.

Other terms includes 'primary bilinguals', which refers to the acquisition of both languages in natural contexts and usually before the age of three, and 'secondary bilinguals', which refers to cases where one of the languages is acquired after the age of three (Hoffmann 1991:19; Lyon 1996:48). Similar comparisons are drawn using the terms 'early bilingualism', which refers to early acquisition of the two languages, and 'late bilingualism', where the L2 is acquired much later than the mother tongue (though there is no agreed age limit between early and late bilingualism). Other definitions attempt to describe the degree of competence in the two languages. For instance, a 'balanced bilingual' is a subject who has mastered two languages to the same extent, whereas a 'dominant bilingual' is a subject who is more fluent in one language than the other (Fabbro 1999:107). The reason for the various definitions is that bilingualism is hard to classify because the individual circumstances surrounding the language acquisition of every child are different.

2. One or two systems?

The growing number of bilingual speakers all over the world (Holmes 1992; Tucker 1998) has been accompanied by a parallel growth of interest in the study of bilingual children's language development and in cross-linguistic studies of language acquisition. The main question that has occupied researchers since the 1970s is whether bilingual children start by mixing the grammatical systems of the two languages and later separate them during their development or separate the two linguistic systems from the beginning of their language development. The first theory is known as the Gradual Differentiation Theory, and its advocates include Leopold (1947), McLaughlin (1984), Redlinger and Park (1980), Ronjat (1913), Swain (1972), and Volterra and Taeschner (1978). The second is known as the Separate Development Theory, and its advocates include De Houwer (1990), Deuchar and Quay (2000), Genesee (1989), Lindholm and Padilla (1978), Lanza (1997), Meisel (1989), and Pearson a.o. (1993).

The question is a very complicated one, as it touches upon unresolved issues in both monolingual and bilingual acquisition. These relate to infant perceptual abilities and their relation to later production; the organization of language(s) in the brain; the nature of the knowledge (or mental representation) that underies language performance; and the influence of the sociolinguistic environment on the development of language(s) in the child (→ child language). A discussion of these crucial issues in this brief review cannot do them any justice, and the reader is advised to consult the landmark references mentioned in the previous paragraph for a detailed discussion of the arguments for and against the one-or-two-system(s) debate.

3. ARE WE ASKING THE RIGHT QUESTION?

Despite results in recent investigations being largely positioned toward the notion of each language developing independently from a very early age, some researchers note that the lack of precise conceptualization on the nature of 'system', among other issues, makes it impossible to determine what type of data would constitute support for separate versus fused systems (e.g. De Houwer 1995; Deuchar and Quay 2000; Johnson and Lancaster 1998). Different researchers have looked at different levels of analysis (e.g. phonological, lexical, and syntactic) and, as the study by Deuchar and Quay (2000) suggests, differentiation at each level might become apparent at different ages. For instance, signs of lexical differentiation might show earlier than morphosyntactic differentiation. The conflicting results that have been reported in the literature are therefore partly due to the discrepancies in the way language differentiation manifests itself at each level of the grammar.

Within a given level of analysis, for instance in phonology, researchers have examined different issues in order to answer the question, e.g. phoneme repertoires, allophonic distribution, phonetic inventories, phonological errors, phonological processes, and prosody. The early emergence of prosodic features may allow investigations of early stages of children's productions; but with regard to segmental features, some researchers have argued that language-specific features generally appear late, and that it might be fruitless to try and find evidence for systems at an early age. This is mainly due to articulatory maturation and the ongoing change in the child's production and perception, but also to the varied input that children receive (e.g. De Houwer 1995; Deuchar and Quay 2000; Johnson and Lancaster 1998; Pearson a.o. 1993). Moreover, while early differentiation might be present at the perceptual level (e.g. Bosch and Sebastián-Gallés 2002; Eilers a.o. 1981), the majority of investigations of language differentiation have concentrated on production. There are often very few production data from the early stages that can be investigated and labeled as belonging to one system or the other.

More recently, researchers have steered away from the one-or-two-system(s) debate in favor of seeing the two languages of the bilingual as belonging to independent but interactive systems (Bialystok 2001:103; de Groot 1993; Holm and Dodd 1999; Johnson and Lancaster 1998). For instance, a child's phonologies can be differentiated but still show influence from the other language. The competence of bilinguals cannot therefore be considered as the sum of two linguistic codes, nor can it be measured in terms of monolingual standards. This is mainly due to the fact that the linguistic experiences that monolinguals and bilinguals encounter are not directly comparable and take place in different environments. As Grosjean (1989) notes, "the bilingual is not two monolinguals in one person". Rather, bilingual individuals have differentiated needs for their two languages or attribute them to different social/emotional functions (what a language is used for, with whom, where, etc.). Thus, they do not necessarily have to develop perfect knowledge, nor the same level of competence and/or performance in both languages.

This does not suggest that monolingual children are at an advantage in terms of their ability to master a linguistic code. There are speech strategies that are unique to bilinguals and which are used as aids to communicative ability. These are known as → code-switching and code-mixing and are often reported in the literature describing bilingual performance, although they have been used with widely different meanings (see De Houwer 1998 and Grosjean 1995 for a full discussion). In simple terms, they refer to alternation between two languages or linguistic varieties during the same conversation (McLaughlin 1984). This can take place at the intra- or inter-sentential level, the latter involving insertion of a single element or a partial or entire phrase from one language into an utterance in another (Hoffmann 1991:105). Elements can be phonological, morphological, syntactic, lexico-semantic, phrasal, or pragmatic. Research indicates that code-switching is not random, but is grammatically constrained and, furthermore, complies with language-specific characteristics of the participating languages (e.g. Meisel 1989; Myers-Scotton 1997). In Arabic, code-switching may include cross-language and diglossic switching (→ diglossia), the latter referring to the speaker alternating between Standard Arabic and the colloquial varieties of their native dialect (e.g. Abdel Rahman 1991; Eid 1988; Heath 1989).

4. Bilingual studies with Arabic as one of the languages

Investigations of bilingual development with Arabic as one of the languages being acquired are scarce in comparison with reports on other language combinations. This is mainly due to the fact that even studies on the monolingual acquisition of Arabic are rare (e.g. Amayreh and Dyson 1998; 2000; Omar 1973; Shahin 2003; → first language acquisition). In what follows a review of the available literature on bilinguals for whom Arabic is the first language (L1), henceforth, Arabic bilinguals, is presented in themes, ranging from language maintenance and bilingual education to normal and disordered aspects of the bilingual grammar.

4.1 *Bilingual education and its relation to language maintenance and loss*

Language use among bilingual children depends on several factors, including whether both languages are used in their community (e.g. English or French in the postcolonial Arabic world) or whether one remains a minority language (e.g. Arabic in many Western countries). In the latter case, the parents' attitude towards the majority language and towards bilingualism is likely to influence the children's language use, although peer pressure often leads to the children's adoption of the majority language for most communicative needs (Atawneh 1992; Mechta 1976; Rouchdy 1971). In most cases, first-borns will be more proficient in the L1 than their siblings (Shorrab 1986) partly due to the home communication between the offspring shifting towards the L2.

The parental attitude towards bilingualism among immigrant communities is often a positive one (Shorrab 1986; Youssef and Simpkins 1985), especially if families think it may help their children's integration in the host society while preserving their home culture. However, one major concern among parents of bilingual children living in an environment where the home language is a minority language is that the children may experience first language attrition (or → language loss). Parents may have the sole responsibility of maintaining the children's L1, especially if there is no support from a wider network of language users. In this case, the parents may choose Arabic as the home language and teach their children literacy skills. This is crucial to many families who wish to teach religion and culture to their children. In some cases, Arabic youth clubs can be found in the host communities (e.g. Abu-Rabia and Siegel 2002; El-Laithy 2002). Their role is to create extra-curricular activities that would meet the children's interests and needs while also teaching them how to read and write in Arabic. In other cases a bilingual education system is put forward by government authorities and Arabic is taught alongside the majority language in mainstream schools. The aim is to maintain the students' cultural and religious heritage and to promote their language and academic skills (e.g. Altena and Appel 1982; Arraf 1996; Benholz and Lipkowski 1999; Helot and Young 2002; Hertz-Lazarowitz 2004; Rado, 1977; Yan 2002; → Europe), but bilingual programs have had varying degrees of success (Garcia and Molina 2001).

In cases where Arabic is the majority language but another language is taught alongside it due to historic and political events (e.g. French and Berber in Morocco, French and Kabyle in Algeria, French in Tunisia, French and English in Lebanon), policy-makers disagree on whether or not a multilingual education should be maintained. In Lebanon, the national curriculum requires that another language is taught by the 7th grade (Bahous 1999), and many Lebanese children end up learning a third language from age 11 onwards. In the Maghreb countries, however, language policy-makers often oppose bilingual education on grounds of preserving religious, cultural, and national identity (El Kirat 1997). This remains an ongoing debate among practitioners, policy-makers, and the communities involved. For a review of retaining multilingual education in light of academic achievement, social structure, and integration, the work of Boukous (1998), Charmion (1977), Chemski (1985), El Kirat (1997), Marley (2004), Wagner a.o. (1989), and Zouaghi-Keime (1991) may be consulted.

4.2 *Bilingual literacy*

There is little research on how Arabic bilinguals become literate in two scripts that do not share the same alphabet, since Arabic orthography does not use the Latin script. This has often led to worries among parents and educationalists concerning the ability of bilinguals to learn two grapheme-phoneme systems at the same time and the possibility of interference problems (Benholz and Lipkowski 1999; Mirzaei 2003; Teberosky a.o. 2002). More research is therefore needed in this area in order to reassure families raising bilinguals, since recent work on bilingual script learning does suggest that bilinguals are capable of simultaneous acquisition of two different scripts and of developing multiple graphic representations for their acquired symbols (Abu-Rabia and Siegel 2002; Eviatar and Rafiq 2000; Kenner and Kress 2003; Kenner 2004; Kenner a.o. 2004). Kenner a.o. (2004) suggest that bilinguals show better understanding of grapheme-phoneme correspondence when dealing with more than one system; the authors conclude that bilinguals can therefore benefit from cognitive gain when they become biliterate. Similarly, Abu-Rabia and Siegel (2002) and Eviatar and Rafiq (2000) suggest that the acquisition of reading and writing in Arabic improves syntactic and phonological awareness in both languages. This view fits in with evidence for the bilingual's ability to accommodate two different systems and to develop strong metalinguistic awareness (e.g. Bialystok 1991).

4.3 *Aspects of the grammar*

Work on aspects of the grammar is heavily under-researched. In phonology, there is evidence for the development of separate systems from an early age (e.g. Shahin 1995; Khattab 2002a). Shahin (1995) examined the developing system of a Palestinian child who was acquiring Arabic at home and English through childcare in Canada between the ages of 1;11 and 2;8.5 and found evidence for distinct phonologies in his production. Khattab (2002a, 2002b, 2002c, forthcoming a, forthcoming b) found similar evidence in Lebanese–English bilinguals aged between 5 and 10 and growing up in the United Kingdom. Results also showed that the context played an important role in the subjects' production, especially with regard to the language(s) of the interlocutor. For instance, the bilinguals produced English-only features when communicating with monolingual English friends, but both English and Arabic features when communicating with other bilinguals and code-switching between the two languages. This was seen as evidence of the bilinguals' sociolinguistic competence and their ability to use their phonetic/phonological repertoire depending on the needs of the situation. In diglossic situations, however, Embarki (2004) suggests that the two varieties inevitably interact. Embarki presents data from the acquisition of Contemporary Standard Arabic by schoolaged children and notes a gradual modification in the perception of the native Moroccan vowels by subjects as a function of Standard Arabic instruction between the ages of 6 and 16.

Bilingual vocabulary acquisition has often been found to lag behind that of monolinguals (e.g. Eviatar and Rafiq 2000), but this may be due to the uneven amount of input from each language (Rouchdy 1971). It should also be borne in mind that the bilinguals are not expected to have translation equivalents for every item in their L1 or L2. Bilinguals often acquire each of their languages in different contexts and use each for different communicative needs.

Similarly, syntactic structures used by bilinguals have often been judged as lacking the level of complexity that is used by monolinguals (Bos 2001; Rouchdy 1971). For instance, Bos (2001) examined the use of tense, aspect, and temporal adverbials in the narratives of Moroccan-Dutch bilingual children aged 5, 7, and 9 and aged-matched monolingual controls. The author found both developmental and language factors affecting the structures examined, with older children in general using more complex adult-like structures than younger ones, and monolingual children showing more consistency, complexity, and variety of structures than bilinguals. However, in cases where bilinguals are being taught literacy in both of their languages, research shows that their syntactic awareness is greatly improved and matches that of monolinguals (Abu Rabia and Siegel 2002).

4.4 *Code-switching*

As with studies on other language combinations, investigations of code-switching among Arabic bilinguals have been interested in testing whether their data fit within existing models of structural constraints on code-switching (e.g. Myers-Scotton 1997; Poplack 1980) and in the social functions of the bilinguals' code-switches. The types of code-switched utterances found in Arabic studies normally resemble those observed in the code-switching literature; for instance, nouns are the most frequently switched items, followed by verbs and other constituents. However, with respect to linguistic constraints on code-switching, Al-Enazi (2002), Al-Khatib (2003), Atawneh (1992), Bader and Minnis (2000), Bentahila and Davis (1995), and Boussofara-Omar (2003) provide ample evidence from bilingual and diglossic children for violation of most morphological and syntactic constraints available in code-switching models. These include adding English suffixes to Arabic words, English words assimilating into Arabic morphological patterns, and mixing Arabic and English/French word order. This supports Gardner-Chloros and Edwards (2004) and Al-Khatib's (2003) suggestion that it is better to look at social bases for code-switching rather than the structure of the grammar.

With respect to the functions of code-switching, children of immigrants seem to use Arabic for cultural and religious terms, while the other language is often used for academic terms (e.g. Al-Enazi 2002). For children growing up in environments where Arabic is the majority language, the L2 may also be used for household and everyday items that have been encountered in the school environment (Bader, 1998). Factors that may affect the degree and patterning of code-switching include length of contact between the two languages, the role and status of each language or variety, and speakers' relative proficiency in each (Bentahila and Davis 1995).

4.5 *Bilingualism and impairment*

There are very few studies combining bilingual language acquisition with language impairment, and fewer still with Arabic bilingual subjects (Hakansson a.o. 2003; Nettelbladt a.o. 2003; Salameh 2003; Salameh a.o. 2004). Yet, from a clinical point of view, it is essential to differentiate between language problems that are due to impairment and those that are due to inadequate exposure to either of the bilingual's languages. This task is made even harder when working with children, as there are individual differences with respect to developmental stages in both monolingual and bilingual acquisition. What may be perceived as impairment may in fact be due to normal delay in any of the developmental stages. It is therefore important to bear these factors in mind when working with Arabic bilinguals, and to consider input and developmental aspects in the children's production before blaming any delay/impairment on their bilingual background. At the same time, therapists cannot wait too long before referring bilinguals with suspected language impairment to therapy, as earlier intervention may lead to faster improvement (Salameh 2003). Both languages of the bilingual need to be tested in order to find out whether therapy is needed for one, or both, developing languages of the child. This is often difficult to implement due to the difficulty in finding Arabic speech and language therapists in the immigrant community or even language experts who may be able to act as assistant therapists.

5. Future directions

Most of the available information in studies with Arabic bilinguals deals with issues of language maintenance and code-switching. Less is known on bilingual literacy, the bilingual grammar (e.g. phonology, syntax, lexicon, etc.), or bilingualism and impairment. Future research should certainly concentrate on these areas, especially when the other language being acquired is typologically different, as this provides rich grounds for testing theories of bilingual language acquisition.

Bibliographical references

Abdel Rahman, Abou Melhim. 1991. "Code-switching and linguistic accommodation in Arabic". *Perspectives on Arabic linguistics*, III, ed. Bernard Comrie and Mushira Eid, 231–250. Amsterdam and Philadelphia: J. Benjamins.

Abu-Rabia, Salim and Linda S. Siegel. 2002. "Reading, syntactic, orthographic, and working memory skills of bilingual Arabic-English speaking Canadian children". *Journal of Psycholinguistic Research* 31: 6.661–678.

Al-Enazi, Mohamed Husein. 2002. *The syntactic form and social functions of Saudi Arabic-English code-switching among bilingual Saudis in the*

United States. Ph.D. diss. Indiana University, Pennsylvania.

Al-Khatib, Hayat. 2003. "Language alternation among Arabic and English youth bilinguals: Reflecting or constructing social realities?" *Bilingual Education and Bilingualism* 6:6.409–422.

Altena, Nelleke and René Appel. 1982. "Mother tongue teaching and the acquisition of Dutch by Turkish and Moroccan immigrant workers' children". *Journal of Multilingual and Multicultural Development* 3:4. 315–322.

Amayreh, Mohamed M. and Alice T. Dyson. 1998. "The acquisition of Arabic consonants". *Journal of Speech and Language Hearing Research* 41.642–653.

—— and Alice T. Dyson. 2000. "Phonetic inventories of young Arabic-speaking children". *Clinical Linguistics and Phonetics* 14:3.193–215.

Arraf, Shereen. 1996. "The bilingual and compensatory education program of the Dearborn schools, Michigan: A model for systemic change and integration of services". *READ Perspectives* 3:1.63–95.

Atawneh, Ahmad. 1992. "Code-mixing in Arabic-English bilinguals". *Perspectives on Arabic linguistics*, IV, ed. Ellen Broselow, Mushira Eid, and John McCarthy, 219–241. Amsterdam and Philadelphia: J. Benjamins.

Bader, Youssef. 1998. "Lexical code-switching in the speech of an Arabic–English bilingual child". *Interface* 13:1.3–17.

—— and Denise D. Minnis. 2000. "Morphological and syntactic code-switching in the speech of an Arabic-English bilingual child". *Multilingua* 19:4.383–403.

Bahous, Jocelyne. 1999. "What language should a Lebanese child learn first?" Paper presented at the Lebanese American University one-day international conference "Multilingualism and Multiculturalism in Lebanon", 3 December, Beirut, Lebanon.

Benholz, Claudia and Eva Lipkowski. 1999. "Early reading and writing in classrooms with students having different native languages". *Deutsch Lernen* 24:3. 253–272.

Bentahila, Abdelali and Eirlys E. Davies, 1995. "Patterns of code-switching and patterns of language contact". *Lingua* 96:2–3.75–93.

Bialystok, Ellen. 1991. *Language processing in bilingual children*. Cambridge: Cambridge University Press.

—— (ed.). 2001. *Bilingualism in development: Language, literacy, and cognition*. Cambridge: Cambridge University Press.

Bos, Petra. 2001. "Temporality issues in Moroccan Arabic and Dutch". *Narrative development in a multilingual context*, ed. Ludo Verhoeven and S. Stromqvist, 233–254. Amsterdam and Philadelphia: J. Benjamins.

Bosch, Laura and Núria Sebastián-Gallés. 2002. "Early language differentiation in bilingual infants". *Trends in bilingual acquisition*, ed. Jasone Cenoz and Fred Genesee, 71–93. Amsterdam and Philadelphia: J. Benjamins.

Boukous, Ahmed. 1998. "Language policy and education". *Plurilinguismes* 16.119–151.

Boussofara-Omar, Naima. 2003. "Revisiting Arabic diglossic switching in light of the MLF model and its sub-models: The 4-M Model and the Abstract Level Model". *Bilingualism: Language and Cognition* 6:1. 33–46.

Charmion, Claude. 1977. "Bilingualism and cooperation". *Cahiers pédagogiques* 33:153.14–17.

Chemski, Zhor Ben. 1985. "The native language up against the outside: The case of North African children". *Peuples méditerranéens/Mediterranean Peoples* 33.127–138.

De Houwer, Annick. 1990. *The acquisition of two languages from birth: A case study*. Cambridge: Cambridge University Press.

——. "Bilingual language acquisition". *Handbook of child language*, ed. Paul Fletcher and Brian MacWhinney, 219–250. Oxford: Blackwell.

——. 1998. "By way of introduction: Methods in studies of bilingual first language acquisition". *International Journal of Bilingualism* 2:3.249–263.

Deuchar, Margaret and Suzanne Quay. 2000. *Bilingual acquisition: Theoretical implications of a case study*. Oxford: Oxford University Press.

Eid, Mushira. 1988. "Principles for code-switching between Standard and Egyptian Arabic". *al-ʿArabiyya* 21.51–79.

Eilers, Rebecca E., William J. Gavin, and Kimbrough D. Oller. 1981. "Cross-linguistic perception in infancy: Early effects of linguistic experience". *Journal of Child Language* 9.289–302.

El Kirat, Yamina. 1997. "Some causes of the Beni Iznassen Berber language loss". *Peuples méditerranéens/ Mediterranean Peoples* 79.35–53.

El-Laithy, Samira Mohamed Hamed. 2002. *Literacy practices and language ideologies of an Egyptian American Muslim youth group: An exploratory ethnographic study*. Ph.D. diss., Hofstra University, New York.

Embarki, Mohamed. 2004. "Implementation of new phonological categories under the schooling effects: Evidence from Moroccan Arabic vowels". Paper presented at the 15th Sociolinguistics Symposium, University of Newcastle, U.K.

Eviatar, Zohar and Rafiq Ibrahim. 2000. "Bilingual is as bilingual does: Metalinguistic abilities of Arabic-speaking children". *Applied Psycholinguistics* 21:4. 451–471.

Fabbro, Franco. 1999. *The neurolinguistics of bilingualism: An introduction*. Hove, East Sussex: Psychology Press.

Garcia, B. Lopez and Mijares L. Molina. 2001. "Moroccan children and Arabic in Spanish schools". *The other languages of Europe: Demographic, sociolinguistic and educational perspectives*, ed. Guus Extra and Durk Gorter, 279–292. Clevedon and Buffalo: Multilingual Matters.

Gardner-Chloros, Penelope and Malcolm Edwards. 2004. "Assumptions behind grammatical approaches to code-switching: When the blueprint is a red herring". *Transactions of the Philological Society* 102:1.103–127.

Genesee, Fred. 1989. "Early bilingual development, one language or two?" *Journal of Child Language* 16.161–179.

Groot, Annette M.B. de 1993. "Word-type effects in bilingual processing tasks: Support for a mixed representational system". *The bilingual lexicon*, ed. Rob Schreuder and Bert Weltens, 27–51. Amsterdam and Philadelphia: J. Benjamins.

Grosjean, François. 1989. "Neurolinguists, beware! The bilingual is not two monolinguals in one person". *Brain and Language* 36:1.3–15.

——. 1995. "A psycholinguistic approach to codeswitching: The recognition of guest words by bilinguals". *One speaker, two languages: Cross-disciplinary perspectives on codeswitching*, ed. Lesley Milroy and Pieter Muysken, 259–275. Cambridge: Cambridge University Press.

Hakansson, Gisela, Eva-Kristina Salameh, and Ulrika Nettelbladt. 2003. "Measuring language development in bilingual children: Swedish-Arabic children with and without language impairment". *Linguistics* 41:2.255–288.

Heath, Jeffrey. 1989. *From code-switching to borrowing: Foreign and diglossic mixing in Moroccan Arabic*. New York: Routledge, Chapman and Hall.

Helot, Christine and Andrea Young. 2002. "Bilingualism and language education in French primary schools: Why and how should migrant languages be valued?" *International Journal of Bilingual Education and Bilingualism* 5:2.96–112.

Hertz-Lazarowitz, Rachel. 2004. "Storybook writing in first grade". *Reading and Writing* 17:3.267–299.

Hoffmann, Charlotte. 1991. *An introduction to bilingualism*. New York: Longman.

Holm, Alison and Barbara Dodd. 1999. "A longitudinal study of the phonological development of two Cantonese-English bilingual children". *Applied Psycholinguistics* 20.349–376.

Holmes, Janet. 1992. *An introduction to sociolinguistics*. London: Longman.

Johnson, Carolyn E. and Paige Lancaster. 1998. "The development of more than one phonology: A case study of a Norwegian–English bilingual child". *International Journal of Bilingualism* 2:3.265–300.

Kenner, Charmian. 2004. "Living in simultaneous worlds: Difference and integration in bilingual script-learning". *International Journal of Bilingual Education and Bilingualism* 7:1.43–61.

—— and Gunther Kress. 2003. "The multisemiotic resources of biliterate children". *Journal of Early Childhood Literacy* 3:2.179–202.

—— Gunther Kress, Hayat Al-Khatib, Rou Kam, and Kuan-Chun Tsai. 2004. "Finding the keys to biliteracy: How young children interpret different writing systems". *Language and Education* 18:2.124–144.

Khattab, Ghada. 2002a. *Sociolinguistic competence and the bilingual's adoption of phonetic variants: Auditory and instrumental data from English-Arabic bilinguals*. Ph.D. diss., University of Leeds, U.K.

——. 2002b. "/l/ production in English-Arabic bilingual speakers". *International Journal of Bilingualism* 6:3.335–353.

——. 2002c. "VOT in English and Arabic bilingual and monolingual children". *Perspectives on Arabic linguistics*, XIII–XIV, ed. Dilworth Parkinson and El-Abbas Benmamoun, 1–39. Amsterdam and Philadelphia: J. Benjamins.

——. Forthcoming a. "/r/ production in English and Arabic bilingual and monolingual speakers". *VaRiation*, ed. Hans van de Velde, Didier Demolin, Roeland van Hout, and Wim Zonneveld. Amsterdam and Philadelphia: J. Benjamins.

——. Forthcoming b. "Phonological acquisition in Arabic–English bilingual children". *Phonological development and disorders: A cross-linguistic perspective*, ed. Zhu Hua and Barbara Dodd. Clevedon: Multilingual Matters.

Lanza, Elizabeth. 1997. *Language mixing in infant bilingualism: A sociolinguistic perspective*. Oxford: Oxford University Press.

Leopold, Werner F. 1947. *Speech development of a bilingual child: A linguist's record*. II. *Sound learning in the first two years*. Evanston, Ill.: Northwestern University Press.

Lindholm, Kathryn and Amado M. Padilla. 1978. "Language mixing in bilingual children". *Child Language* 5.327–335.

Lyon, Jean. 1996. *Becoming bilingual*. Clevedon: Multilingual Matters.

Marley, Dawn. 2004. "Language attitudes in Morocco following recent changes in language policy". *Language Policy* 3.25–46.

McLaughlin, Barry. 1984. *Second language acquisition in childhood*. I. *Preschool children*. 2nd ed. Hillsdale, N.J.: Erlbaum.

Mechta, Marie Claire. 1976. "The education of immigrant children: Native and learned language". *Societé Alfred Binet et Theodore Simon* 76:3.2–8.

Meisel, Jurgen M. 1989. "Early differentiation of languages in bilingual children". *Bilingualism across the lifespan: Aspects of acquisition, maturity and loss*, ed. Kenneth Hyltenstam and Loraine Obler, 13–40. Cambridge: Cambridge University Press.

Mirzaei, Atousa. 2003. *Development of orthographic knowledge in a consonantal script: Children's invented spellings in Farsi*. Ph.D. diss., University of Southern California, Los Angeles.

Myers-Scotton, Carol. 1997. *Duelling languages: Grammatical structure in code-switching*. Oxford: Oxford University Press.

Nettelbladt, Ulrika, Kjell Norlin, and Eva-Kristina Salameh. 2003. "Assessing phonologies in bilingual Swedish–Arabic children with and without language impairment". *Child Language Teaching and Therapy* 19:3.338–364.

Omar, Margaret Kleffner. 1973. *The acquisition of Egyptian Arabic as a native language*. The Hague and Paris: Mouton.

Pearson, Barbara Zurer, Sylvia C. Fernandez, and Kimbrough D. Oller. 1993. "Lexical development in bilingual infants and toddlers: Comparison to monolingual norms". *Language Learning* 43.93–120.

Poplack, Shana. 1980. "Sometimes I'll start a sentence in Spanish Y TERMINO EN ESPAÑOL: Toward a typology of code-switching". *Linguistics* 18:7–8.581–618.

Rado, Marta. 1977. "The multilingual project: A model of bilingual education". *Regional Language Centre Journal* 8:1.42–50.

Redlinger, Wendy E. and Tschang-Zin Park. 1980. "Language mixing in young bilinguals". *Journal of Child Language* 7.337–352.

Ronjat, Jules. 1913. *Le développement du langage observé chez un enfant bilingue*. Paris: Champion.

Rouchdy, Aleya. 1971. "Competence and performance: Is a child's native language affected by the

acquisition of a second language?" *Word* 27: 1.411–420.
Salameh, Eva-Kristina. 2003. *Language impairment in Swedish bilingual children: Epidemiological and linguistic studies*. Ph.D. diss., Lund University, Sweden.
——, Gisela Hakansson, and Ulrika Nettelbladt. 2004. "Developmental perspectives on bilingual Swedish-Arabic children with and without language impairment: A longitudinal study". *International Journal of Language and Communication Disorders* 39: 1.65–91.
Shahin, Kimari. 1995. "Child language evidence on Palestinian Arabic phonology", *Proceedings of the Twenty-sixth Annual Child Language Research Forum*, ed. Eve Clark, 104–116. Stanford, Cal.: Center for the Study of Language and Information.
——. 2003. "Prosody-segmentism in the acquisition of Arabic: Word-final onsets and no stress effect". *Proceedings of the 2003 UBC Child Phonology Conference*. <http://ling75.arts.ubc.ca/child-phonology>.
Shorrab, Ghazi. 1986. "Bilingual patterns of an Arabic–English speech community". *International Journal of the Sociology of Language* 61.79–88.
Swain, Michael. 1972. *Bilingualism as a first language*. Ph.D. diss., University of California, Irvine.
Teberosky, Ana, Martinez Cristina Olive, Chaves Portilla, and Maria-Luisa Rivero. 2002. "Second language literacy (L2) in a multilingual context/ Alfabetización en una segunda lengua (L2) dentro de un contexto multilingüe". *Anuario de Psicología* 33:4.573–592.
Tucker, Richard G. 1998. "A global perspective on multilingualism and multilingual education". *Beyond bilingualism: Multilingualism and multilingual education*, ed. J. Cenoz and Fred Genesee, 3–15. Clevedon: Multilingual Matters.
Volterra, Virginia and Traute Taeschner. 1978. "The acquisition and development of language by bilingual children". *Journal of Child Language* 5.311–326.
Wagner, A. Daniel, Jennie E. Spratt, and Abdelkader Ezzaki. 1989. "Does learning to read in a second language always put the child at a disadvantage? Some counterevidence from Morocco". *Applied Psycholinguistics* 10:1.31–48.
Weinreich, Uriel. 1953. *Languages in contact: Findings and problems*. New York: The Linguistic Circle of New York.
Yan, Ruth Lingxin. 2002. *Parental perceptions on maintaining heritage languages of CLD students*. Ph.D. diss., University of Northern Iowa, Iowa.
Youssef, Anga and Edward Simpkins. 1985. "Parent attitudes on Americanization and bilingual education: The Dearborn, Detroit, and Farmington study". *The Bilingual Review/La Revista Bilingüe* 12:3.190–197.
Zouaghi-Keime, Marie-Anne. 1991. "Bilingualism and the teaching of French". *Le Français dans le Monde* 18.41–44.

GHADA KHATTAB (University of Newcastle)

Child Language

The analysis of Arabic has yielded important insights into the study of language, particularly in the domains of phonology and morphology. Yet, the acquisition of Arabic as a native language by children has garnered remarkably little attention, although it could shed further light on the structure of Arabic and what it entails for the study of the language faculty more generally. This entry presents a summary of what is currently known about the Arabic spoken by such children – how it differs from that of adult Arabic speakers and what the differences might mean for linguistic theory.

Much of the research on child production of Arabic sounds has focused on establishing a normative progression of acquisition, which may then be used clinically to identify speech-disordered children. Studies have been both longitudinal single-subject and cross-sectional of between 30 and 180 children. As a result, the learning path of phoneme acquisition in Arabic is fairly well documented. This path lends partial support to universalist theories of language acquisition, in that children acquire those sounds that Arabic shares with English in roughly the same order and time-course as English-speaking children do. An exception is the consonant /l/, which is acquired an entire year earlier by Arabic-speaking children, possibly because /l/ is more frequent in Arabic (Omar 1973). The segment /r/, on the other hand, poses greater difficulties for them. These differences demonstrate the relevance of language-specific considerations as well as universal considerations, such as articulatory difficulty and markedness, or phonological complexity.

Up to the age of at least 2;6 (2 years and six months), it appears children acquire features rather than phonemes proper. The progression documented in one Palestinian Arabic-speaking child is given in Table 1.

Table 1. Acquisition of features in Arabic

Feature	labial, pharyngeal	continuant, sonorant	coronal	Voice
Age	1;11	2;0	2;4	2;6

The pharyngeal feature is acquired at a very early age (1;11), which is unexpected given its rarity cross-linguistically and its articulatory difficulty. Arabic is distinguished by having a large set of guttural consonants in its phoneme inventory. These consonants have long been the subject of intense interest because of their acoustic characteristics, the physiology of their production, and their role in the abstract phonological system. The pharyngeal segment /ʕ/ is one of the very first to be produced, indicating the importance of ambient language effects. Despite its difficulty and the cross-linguistic rarity of this speech sound, its frequent presence in the child's linguistic environment leads to relatively early acquisition. Interestingly, its acquisition coincides with the appearance of the laryngeal segments /h/ and the glottal stop. This provides support for theories of feature geometry in which laryngeal and pharyngeal segments pattern together as a class under a single node, rather than one in which laryngeals behave independently (Shahin 1995).

Additional work (Morsi 2003) also suggests that by the age of 3 all segments have been acquired other than /r/, the uvular fricative (*ġayn*), and the continuant emphatics. Productions of /r/ are often deviant at that stage. By 4 years, only the uvular fricative (*ġayn*) and the continuant emphatics have not been acquired. At 5 years, the acquisition of /z/ and its emphatic counterpart remains imperfect. In addition, erroneous devoicing of consonants is typical up to the age of 4 years (Morsi 2003). Interestingly, a third study found that children's production of medial consonants is significantly more accurate than that of word-final or, more surprisingly, word-initial ones (Amayreh and Dyson 1998). This raises methodological questions for researchers, particularly as many studies consider only initial consonants. This study also identified a learning path rather different from the previous two. This path was comparatively delayed, with continuants not acquired until the age range of 4–6;4 years, while voiced fricatives and the pharyngeal /ʕ/ are not mastered until after 6;4. These differences are due to methodology. First, segments were considered in a wider range of positions, as stated above. Second, acquisition rates were based on percentage of correct productions, rather than mere occasional productions of a segment sufficing.

Finally, research with Arabic–English bilingual children shows that when producing sounds that appear in both languages but have different phonetic properties in each (such as voiceless stops, which differ in patterns of voice onset time, or VOT), such speakers are able to control such properties from the age of 5 years, depending on the language in use (→ child bilingualism). Thus, children are not only able to identify and accurately produce language-specific fine phonetic detail but vary it systematically, and from a very young age, prior to the acquisition of some other aspects of morphology, syntax, and semantics (Khattab 2003).

It is important to keep in mind throughout the dialect spoken by the child subjects, and the differences between that dialect and other Arabic varieties, both dialectal and Standard. The uvular stop /q/, for example, is not present in the dialects of any of the children whose data is reported above, and therefore it is not considered.

In the realm of morphology, Arabic is well known for its non-concatenative morphology, which is based on discontinuous roots and patterns rather than affixation to stems. While some linguists have questioned the notion of → root and pattern-based morphology, it is widely taken as a given, and recent psycholinguistic work seems to support it (Safi-Stagni 1990; Boudelaa and Marslen-Wilson 2000; Prunet, Beland, and Idrissi 2000). When testing adults, however, the question is complicated by the fact that the root and pattern system is taught explicitly in schools in language classes. Thus, the existence and/or kind of understanding that pre- and early school-age children have of roots and patterns brings important information to bear on their status.

While work has been done on children's knowledge and use of roots and patterns in Arabic, far more information is available on the topic for Hebrew, a related language that shares this property, and the only other Semitic language for which any research on acquisition has been conducted. Thus, Hebrew data is included below where appropriate.

In studies on roots, Arabic-speaking children were able to identify two words containing the same root as 'related' with more than 80 percent accuracy while still in kindergarten. At the same age, they could also produce another word with the same root more than 70 percent of the time.

Hebrew-speaking children showed similar results for both tasks (Ravid 2003). Children's metathesis errors suggest that they distinguish root from affixal consonants, since they permute root consonants without ever switching one with an affixal consonant (Badry 1983).

Arabic- and Hebrew-speaking children both produce novel verbs and nouns using known roots when as young as 3 years old, despite the greater simplicity of affixed forms for nouns (Ravid 2003). Yet, while the Arabic-speaking children do not use such affixes, they seem to be able to recognize them, as shown by the attempt of some 3-year-olds to form verbs by dropping the nominalizing prefix *mu-* (Badry 1983). Also for Hebrew-speaking children, such coinages fall into semantically appropriate patterns by the age of 4. By 3 years they are also able to interpret novel nouns that use roots known to them. Finally, the acquisition of root-based adjectives in Hebrew precedes that of suffixed adjectives (Ravid 2003).

Kindergarteners then are at or near ceiling in both languages in terms of their root awareness. This suggests that the root exists and is acquired very early. Patterns, however, seem to be more difficult to acquire. They necessitate a longer acquisition period extending through at least the sixth year. One study finds that Arabic-speaking children younger than 6 years use only a restricted set of derived pattern forms, which includes verbs and participles but not nouns. Causative is the first verbal form to appear, and with by far the greatest frequency, while the reciprocal form is acquired latest. Surprisingly, the passive appears much later than in other languages (~2;6 for English). Underived forms are also preferred over derived ones until the age of 9. On the other hand, comprehension far outstrips production, with children understanding the meanings of verbal patterns by 3;5 (Badry 1983).

As for nouns, a suffixed form (feminine *-āt*) has been found to be the default for children as well as adults. However, children between 3 and 6 years also occasionally produce pattern-based for suffixed plurals, or substitute one pattern for another (Ravid and Farah 1999).

In the syntactic domain, data on certain widely discussed phenomena in child syntax such as the Optional Infinitive stage are hitherto lacking with respect to Arabic. However, some data are available on syntactic topics. One cross-linguistic generalization that has been confirmed for Arabic is that children take much longer to acquire Binding Principle B than Principle A. They learn the condition on the use of pronouns (that they must be free in their domain) sooner than on anaphors such as reflexives (that they must be bound). In fact, Arabic-speaking children appear to take even longer than children acquiring other languages, with binding still not acquired at 13 years, versus 7 in other languages (Bolotin 1999). Arabic acquisition research (Khanji and Weist 1996) supports the finding that children acquire spatial terms before temporal ones, and forms involving one reference point (in/on, simple past and future tense) before those involving two (between, before/after). Modals and auxiliaries are virtually absent until at least the age of 2;3 (the extent of this stage is not known). At the same age range, agreement for person and number on the verb appears to be random (Mohamed and Ouhalla 1995).

Arabic-speaking children have also been shown to respect island constraints on wh-movement already in the 3;5–5;0 age range, at least with respect to negation (Abdulkarim a.o. 1998). Thus, they correctly reject utterances involving the movement of question-words such as *why* and *where* out of certain restricted environments, as in utterances like **Why don't you think we can help him?*. While this utterance is grammatical for at least one interpretation, it is *not* possible to answer it with 'because we have the skills to help all sorts of people,' and children appear to know this.

In other respects, however, their use of negation differs from that of adults. Palestinian Arabic-speaking children use the form *la* to negate either a preceding or following utterance. In addition to its negative force, the form may also express the modality of wanting or desiring for children. The sentence *laʾ kola hayy* means 'I don't want this coke' in this stage. This phenomenon has been reported for negative morphemes in the child speech of a variety of languages (Mohamed and Ouhalla 1995).

Another possibility for negation is the use of the particles /m-/ and /-š/, either together (*miš/muš*) or flanking the verb (/m-V-š/). In nominal sentences children, like adults, use only the *miš* variant. In the adult grammar, the two components must occur on either side of the verb, as a result of verb movement above Neg(ation) to I/T (Inflection/Tense). Children, on the other hand, continue to produce the two variants – continuous and discontinuous – in roughly equal proportions. It appears then that this verb movement is optional for children, rather than obligatory as it is for adults (Mohamed and Ouhalla 1995). The implications for linguistic theory remain unclear.

In sum, child learners of Arabic are like those of other linguistic communities in their astonishing facility with their language. They are able to recognize and use complex facets of their language with ease even at very young ages, as shown by the bilingual phonetics data and syntactic movement facts. In addition, their developmental paths have much to reveal specifically about the language they are learning (for example, with the triconsonantal root evidence, and acquisition of pharyngeals). The particular characteristics of Arabic – including such interesting phenomena as the construct state and the longstanding puzzle of reverse agreement in numerals, among others – offer much scope for further research. Such work is bound to increase our understanding of Arabic and of the language faculty in general.

Bibliographical references

Amayreh, Mousa and Alice Dyson. 1998. "The acquisition of Arabic consonants". *Journal of Speech, Language, and Hearing Research* 41.642–653.

Badry, Fatima. 1983. *Acquisition of lexical derivational rules in Moroccan Arabic: Implications for the development of Standard Arabic as a second language through literacy*. Ph.D. diss., University of California, Berkeley.

Bolotin, Naomi. 1999. "Acquisition of binding in Arabic". *Proceedings of the Boston University conference on language development (BUCLD) 23*, ed. Annabel Greenhill a.o., 93–102. Somerville, Mass: Cascadilla Press.

Boudelaa, Sami and William D. Marslen-Wilson. 2000. "Non-concatenative morphemes in language processing: Evidence from Modern Standard Arabic". *Proceedings of the workshop on spoken word access processes (SWAP) 1*, ed. Anne Cutler a.o., 23–26. Nijmegen: Max Plank Institute.

Khanji, Rajai and Richard Weist. 1996. "Spatial and temporal locations in child Jordanian Arabic". *Perceptual and Motor Skills* 82.675–682.

Khattab, Ghada. 2003. "Age, input and language mode factors in the acquisition of VOT by English-Arabic bilingual children". Sole (2003:3213–3216).

Mohamed, Ibrahim and Jamal Ouhalla. 1995. "Negation and modality in early child Arabic". *Perspectives on Arabic linguistics*, VII, ed. Mushira Eid, 69–90. Amsterdam and Philadelphia: J. Benjamins.

Morsi, Ranya. 2003. "Phonological acquisition of normal Egyptian children from the age of two and a half to five years". Sole (2003:1951–1954).

Omar, Margaret. 1973. *The acquisition of Egyptian Arabic as a first language*. The Hague: Mouton.

Prunet, Jean-François, Rene Beland, and Ali Idrissi. 2000. "The mental representation of Semitic words". *Linguistic Inquiry* 31.609–648.

Ravid, Dorit. 2003. "A developmental perspective on root perception in Hebrew and Palestinian Arabic". *Language processing and acquisition in languages of Semitic, root-based, morphology*, ed. Joseph Shimron, 293–319. Amsterdam and Philadelphia: J. Benjamins.

—— and Rola Farah. 1999. "Learning about noun plurals in early Palestinian Arabic". *First Language* 19.187–206.

Safi-Stagni, Sabah. 1990. "Slips of the tongue in Arabic". *Perspectives on Arabic linguistics*, I, ed. Mushira Eid, 271–290. Amsterdam and Philadelphia: J. Benjamins.

Shahin, Kimary. 1995. "Child language evidence on Palestinian Arabic phonology." *Proceedings of the 26th Annual Child Language Research Forum*, ed. Eve Clark, 104–116. Stanford, Cal.: Center for the Study of Language and Information.

Sole, Maria J. a.o. (eds.). *Proceedings of the 15th international congress of phonetic sciences*. Barcelona: The 15th International Congress of the Phonetic Sciences Organizing Committee.

Mary-Ann Walter
(Massachusetts Institute of Technology)

China

Arabic in China has a long history, as a language for economic and cultural exchange and as the religious language of a small but deeply-rooted Chinese Muslim population. Exchanges with 'Western regions' have been attested in Chinese sources since 126 B.C.E. and were pursued during the Sassanid Empire. After the fall of the latter in 651 C.E., the Umayyads and Abbasids sent 88 envoys to China, mostly for trade purposes, under the Tribute system favored by the Chinese Empire. During the Tang (618–907) and Sung (960–1279) Dynasties, commercial exchanges were carried out by road, to the western capital of the Tang Dynasty, Chang'an (now Xi'an), called Xamdān in Arabic sources, or by sea, causing thousands of Arab and Persian traders to settle in the southern ports such as Canton (Xānfū), Quanzhou (Zaytūn), and Yangzhou. During the Mongol Yuan Dynasty (1279–1368), the Mongolians relied heavily on Muslims for administrating China and many of them became top officials. Muslims arrived in large numbers, possibly up to two or three millions, settling all over China, especially in the north and the southwest. They developed medicine, astronomy and calendar calculations, military engineering, architecture, and language study. It is likely that coming essentially from Central Asia and Persia, Persian was their lingua franca. Estefi, a commercial Persian script, was one of the scripts used by the Mongols (Leslie 1986). There is no evidence of an extensive use of Arabic by these Muslims or the Mongol rulers.

Archaeological remains of Arabic script have been found in China, the earliest ones in the coastal ports and in the ancient capital Chang'an. Arabic inscriptions are found on tombstones and inside mosques, mostly Qur'ānic quotations. Mosque steles were erected to commemorate the building or the rebuilding of a mosque, and give accounts about it. The earliest inscriptions in Arabic date from the Sung period. The Yangzhou mosque inscription in Arabic dates from 1272, those of Quanzhou's Shengyu and Qingjing mosques are from 1310–1311, and Canton's Huai-sheng mosque, the earliest in China, has an inscription from 1324 to 1327, with a small part in Arabic. The bilingual stele dated 742 that the Great Mosque of Xi'an claims to possess is considered a forgery as it was most probably carved in the early 15th century (Leslie 1986:46–47, 70). The outside doors of a mosque and the courtyard might have Chinese inscriptions, but the prayer hall bears exclusively Arabic script, especially around the *mihrāb*. Implements such as ceramic or metal plates with Arabic decoration are found from the Ming Dynasty. Nowadays, Arabic script is a marker of Muslim identity. Chinese Muslim calligraphy, a unique style mixing Arabic script and Chinese calligraphy forms, may be found in Muslim households and in mosques (Aubin 1988), and Chinese calligraphers have participated since the late 20th century in international competitions of Arabic calligraphy. Religious books, first of all the *Qur'ān*, due to their scarcity, were treasured. The *Qur'ān* was not printed until 1862 in Yunnan, due to the sacred status of the text. Manuscript copies were kept as precious treasures by mosques and families, constituting an invaluable wedding present. Every religious student in the past had to make his own copy, taking his scrolls on his back when seeking knowledge in different places and with various teachers (Jin 1981).

Arabic, though not spoken, was maintained in China as a sacred language by the tiny minority of people who, first being Muslims in China, slowly became during the Ming Dynasty (1368–1644) Chinese-speaking Muslims and subjects of the Chinese Empire.

Arabic has left its imprint on the language spoken by the Muslims and to a lesser extent on the Chinese language in general. Arabic as well as Persian religious words can be introduced in a phrase, or combined with Chinese words to form a new one, for instance *Gulan-jing* '*Qur'ān*-book' to name the *Qur'ān*. Arabic words may be translated, using the existing Chinese vocabulary, thus creating a new word that refers exclusively to the original one in Arabic. For instance, *allāh* can be phonetically transcribed by two or three characters as *an-la(-hu)*, but it can also be translated as *zhenzhu* 'Real Lord', whereas *tianzhu* 'Lord of Heaven' means 'God' for Christians as well as for Muslims. Phonetic transcription in Chinese characters is problematic, as it differs greatly from Arabic pronunciation, since there is no uniform transcription and since the meaning of the Chinese characters used phonetically does not make sense for the reader, for instance *an-la-(-hu)* 'Peace-pull-family' for 'Allah'. Arabic words are more often used in areas like the northwest with a dense Muslim population which is rural and poorly educated, whereas translated words will be used in the eastern areas where Muslims are less numerous, more urban, and more educated.

The most common Islamic terms, like 'Islam', 'Allah', '*Qur'ān*', have entered the Chinese language. Apart from these common terms, religious words or composite words are used and understood only by Chinese Muslims. When speaking, for instance, Muslims use the word *mumin* (< *mu'min* 'believer'), which fortunately can be written with the character *mu*, used for Muḥammad and *min* for 'people'. They use *gaomu* (< *qawm* 'people') to designate the Muslim population depending on one mosque. *Kafeile* (< *kāfir* 'infidel') is used sometimes of non-Muslims, whereas *munaifeige* (< *munāfiq* 'hypocrite') is used to upbraid a Muslim who is not reliable in his faith or in his relations with his coreligionists. A new expression has appeared for the female students of private confessional schools *talibati* (< *ṭālibāt* 'female students'), whereas *manla* (< *mawlā* 'patron, client, companion') or *halifa* (< *xalīfa* 'successor, caliph') designate male students of traditional religious education (Wang 2001). This is so much the case that in the perspective of seeking to establish Chinese Muslims as a distinct ethnic 'nationality' (Thoraval 1990), this mixed idiom must be regarded as a distinct language, one of the constitutive features of a 'nationality'. It was named *jingtang yu* 'scripture hall language' because it was first shaped and used inside the mosques, in the *jingtang*, the 'scripture hall', a place devoted to Islamic classics and teaching (Yin 1996).

Chinese Muslims, allegedly since the time of Hu Dengzhou (1522–1597), installed in every mosque the so-called 'scripture hall education', in order to train imams, or *ahong* (< Persian *āxūnd*), also named in Chinese 'chiefs of religion'. The traditional course consists of the study of the 'Thirteen Classics', eight in Arabic, including the *Qur'ān*, and five in Persian. Arabic language study comprises the study of grammar, with *'Asās al-ʿulūm*, an abstract of four grammatical treatises, and *Ḍaw' al-miṣbāḥ* by al-Muṭarrizī (1143–1213). Language study also comprises a treatise on rhetoric, *Malā'* or *Šarh al-kāfiya* by ʿAbd ar-Raḥmān al-Jāmī (1397–1477) and one on logic, *Bayān* or *Talxīṣ al-miftāḥ* by Saʿd ad-Dīn at-Taftazānī (1321–1389). Religious studies use excerpts of *ḥadīṯ*: *'Arbaʿūna xuṭab*, which was used in China in Persian translation (it is worth noting that traditionally no great collections of *ḥadīṯ* were available). Religious study includes *kalām* or dogmatic theology with *ʿAqā'id al-'Islām* by ʿUmar an-Nasafī (1068–1142); *fiqh* with *Šarḥ al-wiqāya* by ʿUbaydallāh ibn Masʿūd (d. 1346) with its commentaries; and exegesis of the *Qur'ān*, such as the *Tafsīr al-Jalālayn* by Jalāl ad-Dīn Muḥammad ibn 'Aḥmad aš-Šāfiʿī (d. 1459) and Jalāl al-Dīn as-Suyūṭī (1445–1505). The Thirteen Classics are referred to by the Arabic name of *Sābiqa*, as they are the essential books a prospective imam must study thoroughly. For further study, 70 other books may be used, depending on the knowledge of the teacher and on his sectarian tendency (Yang and Yu 1995). Traditional Islam, called *Qadīm*, adheres to the Ḥanafī *maḏhab*, as is the case with all religious trends in Chinese Islam except the Salafiyya, and has developed the above-mentioned 'scripture hall teaching'. It remains until now the most popular movement in Chinese Islam.

Offshoots of Sufi brotherhoods coming from central Asia, the Kubrawiyya, the Qādiriyya, and the Naqšbandiyya, appeared in the northwest in the second half of the 17th century. These orders called for reform, relying on the support of books. The Jahriyya branch of the Naqšbandiyya introduced in the 18th century a sacred book entitled in Chinese *Mingshale*, an enigma for years for Western scholars, as the title, translated from Chinese characters, means 'Shining Sand' (Lipman 1997:64–72). In reality, *Minšār* 'The Saw' was a small guide to recitations chanted at Sufi gatherings. It was shorter than other books and was violently opposed by the earliest orders, as it gained adherents by reducing the time of rituals, and subsequently the emoluments of religious servants. Thirty years later, a more recent offshoot of the Jahriyya Naqšbandiyya branch introduced new books in Arabic and was met with opposition (Aubin 1990). Rivalries between Sufi orders led to bloodshed and worsened Muslim rebellions in the 18th/19th centuries, which claimed millions of lives.

Founded in the northwest in the late 19th century, the 'Ixwān movement, though claiming to follow the Ḥanafī *maḏhab*, relied on 'Ten Great Books', essentially from the Ḥanbalī *maḏhab* and Wahhābī inspired authors, such as Ibn Taymiyya (d. 1328), or one of his successors, Muḥammad 'Amīn al-Birkawī (d. 1573) and Muḥammad ibn ʿAbd al-Wahhāb (1703–1766). Persian literature was simply abandoned. At the end of the 1930s, a scission of the 'Ixwān, directly influenced by the Wahhābī teaching in Mecca, called itself Salafiyya. Relying on newly imported Arabic religious books, it rejected the authority of the four *maḏāhib*. Nowadays, flooded with Saudi literature, the Salafiyya is one of the most active Islamic trends in China. Arabic religious books are still invested with a sacred meaning.

Although extremely rare, a religious literature in Arabic appeared in China. In the southwestern province of Yunnan from the mid-19th century onward, two scholars, Ma Fuchu (1794–1874) and one of his disciples, Ma Lianyuan (1841–1903), were taught in a traditional Chinese Muslim way but knew Chinese as well, and traveled the Muslim world for pilgrimage and further knowledge (Lin 1990). In order to reform religious teaching, they wrote summaries of Arabic books, for instance grammars (*Kāfiya* and *Muttasiq* books of *naḥw* written by Ma Fuchu, *Hawā* for *ṣarf* by Ma Lianyuan), law books (*ʿUmdat al-ʾIslām* by Ma Lianyuan), and books of rituals, which were translated into Chinese at the end of the 19th or the beginning of the 20th century. Ma Fuchu wrote in Arabic the story of his journeys from 1841 to 1849. Ma Lianyuan, who signed by the name of Muḥammad Nūr al-Ḥaqq ibn Luqmān aṣ-Ṣīnī, translated into Arabic the *Tianfang xingli* 'The philosophy of Arabia', an opus by the most famous Chinese Muslim scholar, Liu Zhi (ca. 1670–ca. 1730) under the title of *Šarḥ al-laṭāʾif* (Wheidi 1994), retranslated in modern Chinese in 1982. A few other books are attested in Arabic, for instance a survey of Islamic doctrine for courses in Arabic language with its translation published in 1937, and in 2000 a small Naqšbandiyya Mujaddidiyya Sufi order published a *silsila* and a pamphlet against Wahhābism, *Šarḥ Sayf ad-Dīn*. Other works may exist, stored by imams or Sufi orders and ignored by Chinese scholars, who until recently did not know Arabic, but there is little evidence that a large body of literature could exist.

Apart from training disciples, imams, together with assistants when the mosque can afford them, have the duty to convey basic Islamic knowledge, rituals, and extracts of the *Qurʾān* to every Muslim depending on the mosque. For children, especially in rural areas, this was the only education available. To support this basic teaching, various kinds of small booklets were written. Various *Haiting* (*Xatm al-Qurʾān*) contain a presentation of the Arabic alphabet with the pronunciation of letters, followed by the last part of the *Qurʾān* and the most famous *ʾāyāt*, which are learnt by heart. Others called *Zaxue* 'Miscellaneous studies' treat rituals, prayers, *ġusl*, fasting, prayers for marriage and funerals, *duʿāʾ* for specific occasions, Qurʾānic names to be given to babies, and so on.

Religious training, even for advanced study, did not include learning the written Chinese language, which was an exception, depending on the sole intention and abilities of the scholar. As explanations of religious books were given in Chinese, Muslims developed, probably from the 18th century, the first alphabetic script for Chinese with Arabic letters written in a Persian style (Bausani 1968; Yin 1996). This script was called *xiao'erjin*, referring to marginal notes in religious books, or *xiaojing* 'secondary book'. It is still used by old people and those of younger generations who did not receive a general education. Booklets may be written in three kinds of script, Arabic, Chinese translation in *xiao'erjin*, and Chinese characters. Some texts are still published in this transcription and, significantly, a translation of the *Qurʾān* was published officially for the first time in 1995 in the three scripts.

Nowadays, Islamic publications are mostly private and mostly printed in Linxia in the northwestern province of Gansu. Publications do not often indicate the author's name (sometimes they only mention the 'Qurʾānic' name, the Arabic name given to Muslim children by the imam), and sometimes not even the date or the publisher's name, because they need an official agreement which private publishers do not care or dare to obtain. Since the mid-1980s, it is likely that all Islamic texts available in China, including some which were spared from destruction during the Cultural Revolution and some brought from the rest of the Muslim world, have been printed. Classical Islamic books in Arabic are copied from the original Egyptian, Turkish, or Indian text for mosque teaching purposes. The Islamic literature in Chinese that appeared from the end of the 16th century, reprinted in late 19th or mid-20th century editions, bears the Arabic script along with the phonetic transcription in Chinese characters of the Arabic words. Other publications are translations, sometimes with a bilingual text. Different compositions are made with Arabic and Chinese: the text can be mostly in Arabic with Chinese marginal notes, or in Chinese with quotations of the Arabic sentences the text translates and explains. Recently, Saudi Arabian publications have been copied and sold mostly in ʾIxwān and Salafī bookstores. Many translations from contemporary Arab or Muslim authors also have been made by Chinese Muslims, who have often studied abroad for several years in Arab or Muslim countries. Institutional scholars look down on this literature, called 'popular', which is badly written because these self-styled translators are not highly educated (Allès a.o. 2003).

Since the beginning of the 20th century, Chinese Muslims have founded modern schools inside or outside mosques in order to offer general education in Chinese and religious education in Arabic, excluding the traditional Persian language. As ties were established with the Muslim world at the beginning of the 20th century, two teachers, ʿAlī Riḍā and Ḥusayn Ḥāfiẓ, were sent in 1907 to Beijing by the late Ottoman monarch, and two others by the Egyptian king in the 1930s, Muḥammad ad-Dālī and Muḥammad ʾIbrāhīm Fulayfil. The two monarchs sent books, stored in the Fūʾād Library in Beijing, which was destroyed during the Cultural Revolution. The People's Republic, founded in 1949, banned private confessional teaching from the early 1950s to the 1980s, until a more liberal stance allowed religious mosque education to resume and private Muslim schools to open. Moreover, except in Xinjiang for fear of secessionist feelings, the government allowed and sometimes encouraged the founding of private Muslim schools in order to provide education for people who could not attend increasingly expensive state schools or who left them early, for lack of money or lack of satisfactory achievements. Not daring to call themselves Islamic schools, these institutions, ranging from kindergartens to secondary-level schools, often call themselves 'Sino-Arabic schools' or 'language' schools (Gladney 1999). A more colloquial Arabic is taught, even though the level remains low. For further studies, students from these schools try to go, at their own expense or with grants provided by Islamic countries, to Malaysia, Pakistan, Egypt, Saudi Arabia, and elsewhere. Some make their living as translators, others become teachers in the private Muslim schools of China, but the vast majority leave the field of Islamic studies altogether.

To develop relations and knowledge of the Arabic world, the government of the People's Republic of China relied in the 1950s on twenty or so Muslim students who had been to study at al-Azhar in the 1930s and early 1940s. They became university teachers, translators, Arabic broadcasting advisors – China had broadcast in Arabic from 1957 onward. The most famous graduate from al-Azhar, Ma Jian (1906–1978), in Arabic Muḥammad Mākīn aṣ-Ṣīnī, translated during his years of study Confucius' *Analects* into Arabic and various other books from Arabic. He was the first Arabic teacher at Beijing University in 1946 and was occasionally a translator for Chairman Mao. He collaborated in a Sino–Arabic dictionary but is much famed for his translation of the *Qurʾān* into colloquial Chinese, which became the main reference for scholars and for Muslims. Only two universities taught Arabic in the early 1950s, out of nine or so universities or colleges (Xiang 1995). The first students of Arabic at Beijing University were all Muslims whereas today Chinese Muslims cannot pass the examination, which is set in English to enter university. Arabic has a modest place in both international relations and academic circles in the People's Republic of China, whereas it remains attractive for Chinese Muslims, who have preserved the traditional mode of learning and attempt to develop knowledge of modern Arabic in modern, but poorly funded schools.

Bibliographical references

Allès Elizabeth, Leila Chérif-Chebbi, and Constance-Hélène Halfon. 2003. "Chinese Islam: Unity and fragmentation". *Religion, State and Society* 1.7-35.

Aubin, Françoise. 1998. "L'art de l'écriture chez les musulmans de Chine". *Horizons maghrébins. Le droit à la mémoire* 35–36.29–43.

——. 1990. "En islam chinois, quels Naqshbandis?". *Naqshbandis: Cheminements et situation actuelle d'un ordre mystique musulman*, ed. Marc Gaborieau, Sacha Popovic, and Thierry Zarcone, 491–572. Paris and Istanbul: Isis.

Bausani, Alessandro. 1968. "Un caso estremo di diffusione della scrittura araba: Il 'sino-arabo'". *Oriente Moderno* 11–12.857–876.

Chen, Dasheng. 1984. *Quanzhou yisilanjiao sheke: Islamic inscriptions in Quanzhou*. Fuzhou: Ningxia and Fuzhou Renmin Chubanshe.

Gladney, Dru. 1999. "Making Muslims in China: Education, Islamization, and representation". *China's national minority education: Culture, state schooling and development*, ed. Gerard Postiglione, 55–87. New York: Garland Press.

Jin, Yijiu. 1981. "The *Qurʾân* in China". *Contributions to Asian Studies* 17.95–101.

Leslie, Donald Daniel. 1986. *Islam in traditional China: A short history to 1800*. Canberra: Canberra College of Advanced Education.

Lin, Chang-Kuan. 1990. "Three eminent Chinese *ʿulamāʾ* of Yunnan". *Journal Institute of Muslim Minority Affairs* 11.100–117.

Lipman, Jonathan N. 1997. *Familiar strangers: A history of Muslims in northwest China*. Seattle and London: University of Washington Press.

Pillsbury, Barbara. 1984. "Muslim history in China: A 1300-year chronology". *Journal Institute of Muslim Minority Affairs* 10–29.

Thoraval, Joël. 1990. "Le concept chinois de nation est-il 'obscur'? A propos du débat sur la notion de *minzu* dans les années 1980". *Bulletin de Sinologie* 65.24–41.

Wang, Jianping. 2001. *Glossary of Chinese Islamic terms*. Richmond: Curzon Press.

Wheidi, Souad. 1994. "Al-falsafa al-ʾislāmiyya fī ṣ-Ṣīn: Waḥdat al-wujūd wa-naḏariyyat al-ʾinsān al-kāmil ʿinda Liū Tšī". *Etudes Orientales. Dirāsāt Šarqiyya* 13–14.72–111.

Xiang, Peike. 1995. "Zhongguo alaboyu jiaoxue shiye de fazhan" [Development of the teaching of Arabic in China]. *Alabo Shijie* 4.45–49.

Yang, Huaizhong and Yu, Zhengui. 1995. *Yisilan yu Zhongguo wenhua* [Islam and Chinese culture]. Yinchuan: Ningxia Renmin chubanshe.
Yin, Zhu. 1996. "Jingtangyu yu xiao'erjin" [Scripture hall language and *xiao'erjin*]. *Zhongguo Yisilan wenhua* [Chinese Islamic culture], 150–153. Beijing: Zhonghua shuju.

Leila Chérif-Chebbi (Paris, France)

Christian Middle Arabic

1. Definition

For methodological reasons, two forms of Christian Middle Arabic are distinguished here, according to the date of the texts written in it: Early Christian Middle Arabic and Late Christian Middle Arabic. Early Christian Middle Arabic refers to the form of → Middle Arabic used in texts and/or manuscripts dating back to the 9th or the 10th century. Many of these texts were copied in monasteries of South Palestine, among which the famous monastery of Saint Catherine in the Sinai (Blau 1966:20). Late Christian Middle Arabic refers to Middle Arabic used in texts and/or manuscripts written after the year 1000 (Knutsson 1974; Bengtsson 1995; Grand'Henry 1984, 1988, 1996). This chronological distinction should not imply that two different linguistic types of Middle Arabic are referred to. Current research on Middle Arabic has not gone into deep enough detail to allow a linguistic definition of different layers of Middle Arabic. Fischer's (1991:432–433) sociolinguistic definition seems to fit all linguistic forms of Middle Arabic encountered:

> I have proposed to call Middle Arabic all those texts written by authors who are not able to reach the high standard of educated Classical Arabic or who do not intend to do so [. . .]. As a result of this sociolinguistic definition, we notice that Middle Arabic texts were not only produced in former times, they are written in our times as well.

At least, Middle Arabic may be identified through two characteristics. On the one hand, the history of Middle Arabic is undoubtedly connected with the history of Arabic dialects. On the other hand, "three standard types can be distinguished in Middle Arabic: Classical Arabic with Middle Arabic admixture; semi-classical Middle Arabic; classicized Middle Arabic" (Blau 1966:50–51).

> Owing to their tendency to use 'correct' language, but hampered by their deficient knowledge of Classical Arabic, Middle Arabic authors used forms which were neither Classical nor Middle Arabic forms and may, accordingly, be called pseudo-correct features. Sometimes these pseudo-correct features take the shape of simple malapropisms [. . .]. In other cases, the pseudo-correct forms appear as hyper-corrections [. . .]: the writer, in his desire to use classical forms and avoid those of Middle Arabic, overshoots the mark and utilizes features peculiar to the language of prestige (i.e. Classical Arabic) even in positions which demand forms found equally in the lower language (i.e. Middle Arabic), thus using features 'too correct' [. . .]. There exists, however, another feature, emerging from the influence of a superior language upon its inferior, which we would propose to call hypo-correction [. . .]. While the authors and copyists of Ancient South Palestinian texts tend to exaggerate the use of 'correct forms', employing hyper-correction, they can also occasionally err in the opposite manner, in 'correcting' vulgar forms only to a partial extent, using hybrid forms 'not sufficiently correct' (Blau 1966:50–51).

1. Orthography and phonetics

1.1 *Vowels*

i. Short *a* may shift to *e/i* and *ā* to *ē/ī* (Blau 1966:63–65, 2002:29). The most important information about the vowels in Early Christian Middle Arabic derives from the publication in 1901 of a bilingual Greek-Arabic fragment of Psalm LXXVIII (Violet 1901), dating back to the 8th century (on this text, see Haddad 1992:159–164: the Psalm book is dated between 680 and 710).

ii. Long diphthongs are shortened: *ḥāʾiṭ* >*ḥāyṭ* 'wall' (Blau 1966:67) or monophthongized: *ʾulāʾika* > هوليك 'those' (Blau 1966:67); *al-mawt* > *al-mūt* 'death' (Bengtsson 1995:105).

iii. Elision of short final vowels: frequent absence of the accusative *ʾalif* in nouns, absence of differentiation between imperfect forms terminating in endings with and without *n*, absence of short final vowels in Greek transcriptions and in Arabic Bible translations (Blau 2002:30; Bengtsson 1995:107).

iv. Lengthening of short vowels: *sā* instead of *sa-* preceding an imperfect as marker of the future tense, sometimes *fal* > *fāl*, *ka* > *kā* and *la* > *lā* (Blau 1966:68–80: orthographic and phonetic features not found so far in Late Christian Middle Arabic; but *wa* > *wā* in Knutsson 1974:52; Bengtsson 1995:100).

v. *u* > *ū* (Blau 1966:73).

vi. *hā* marking final *-a* in the pronominal suffix 1st person sg. *-ya*: رجليه 'my feet' (Blau 1966:75–76).

vii. the Classical Arabic syllable structure has changed: final closed syllabes may contain lo ng vowels in non-pausal position لم يغيب 'it was not hidden' (Blau 1966:77, 2002:30)

viii. *ʾalif maqṣūra bi-ṣūrat al-yāʾ* is often represented by *ʾalif*: اعلا 'higher' (Blau 1966:81, 2002:32; Grand'Henry 1996:5, n.9).

1.2 *Consonants*

i. The glottal stop has weakened and nearly completely disappeared, so that it may be omitted in every position (Blau 1966:83–105, 2002:32–33; Knutsson 1974:60–76): *ḏiʾb* > *ḏīb* 'wolf', *radīʾ* > *radi* 'wicked', *ruʾyā* > *ruyyā* 'vision'; *iʾa* > *iya*, *īʾā* > *īyā*, *uʾa* > *uwa*, *aʾa*, *āʾa* > *ā*, *uʾu* > *ū*, *wʾ* > *ww*, *yʾ* > *yy*.

ii. *ṯ* > *t*, *ḏ* > *d*: this phenomenon is marked as "quite probable" by Blau (1966:106) and is found systematically in some manuscripts in which all the Classical interdentals are superseded by the corresponding dentals (Grand'Henry 1988:234–291, see critical apparatus, manuscripts D and H). Nowadays, this phenomenon regularly appears in most Modern Arabic dialects, mainly in towns. Knutsson (1974:88) uses manuscripts in which a few words only have *ṯ* and *ḏ*, while elsewhere there are dentals. He assumes that in these cases the interdentals represent a sibilant pronunciation, like in modern Egyptian Arabic for some words (*ḥayṯu* > *ḥās*, Knutsson 1974:89), but in certain cases, like دوم/اذوم 'Edom', dentals and interdentals "may have been looked upon as merely alternative spellings" (Knutsson 1974:99).

iii. *ṣ* > *s*: sometimes as an effect of dissimilation يسطبغون 'they are baptized', السور 'the horn' (Blau 1966:110; Knutsson 1974:105–106).

iv. *š* > *s* in loanwords from Aramaic: سليح 'apostle' (Blau 1966:111; Grand'Henry 1996:21, n. 51).

v. *s* > *ṣ*: صخن 'hot' (Blau 1966:113).

vi. *ḍ*: it is probable that all spirants disappeared at one stage (Blau 1966:113, n. 175).

vii. *ḍ* < *ḏ̣* and *ḏ̣* < *ḍ* reflect a form of merger of both consonants, which probably occurred right from the start of Middle Arabic: يغلض 'it renders coarse', ظلت 'it was lost', حضاير 'enclosures', حظر 'it has come' (Blau 1966: 113–114, 2002:34; Knutsson 1974:106–107); when *ḍ* is replaced by *ḏ̣*, which was the marked member of the pair, this tends to be the result of hypercorrection.

viii. *ġayn* instead of *ʿayn*: in *ʿamīq* 'deep', ازلغمق 'the depth' (Blau 1966:115; Knutsson 1974:108).

ix. *tāʾʾ marbūṭa*, *tāʾ ṭawīla*, *ʾalif maqṣūra*, *ʾalif mamdūda* alternate: ابنت ' 'daughter', احدة 'one [fem.]', العذ رة , العذرا 'the virgin', عكةَ, عكا 'Acre' (Blau 1966:119; Knutsson 1974:111).

x. *ʾalif fāṣila* is very often spelled after every *wāw*: للعدوا 'to the enemy', ارجوا 'I hope', نحوا 'toward, about, according to' بنوا, 'sons' (Blau 1966:127–128, 2002:35; Knutsson 1974:114; Grand'Henry 1988: 257, n. 63).

2. MORPHOLOGY

2.1 *Pronouns*

i. Several pronouns are similar to those of Modern Arabic dialects: *ʾantī* 'you [2nd pers. fem. sg.]', *-uh* (3rd pers. masc. sg. suffix) 'him', *naḥnā* 'we' (Blau 1966:133–135; Knutsson 1974:116); this phenomenon is also reflected in Christian Middle Arabic texts written in Coptic characters (Blau 1979:218). Pronominal suffixes in *-kon* and *-hon* appear in Late Christian Middle Arabic instead of Classical *-kum* and *-hum*, as in some Middle Eastern dialects (Blau 1966: 133–135; Knutsson 1974:116–117).

ii. The demonstrative pronoun هؤلاء 'these' > هاولا هولا هاولى هولى هول and أولائك > اوليك 'those' (Blau 1966:127–128; Grand' Henry 1996: 54, l. 14).

iii. Interrogative pronouns: *mā* was superseded by *ʾayy šay* and its developments ايشن, ايش 'what', لايشي 'why' (Blau 1966:142–143).

iv. Indefinite pronouns: شي 'something' and واحد or الواحد 'someone' (Blau 1966:142–143).

2.2 *Verbs*

i. The verb in the perfect: elision of the short final vowels: *semiʾ* 'he heard', *fa-amtanaʿ* 'and he refrained' (8th century, Violet 1901) and suffixes *-tī* (*scriptio plena* for the 2nd pers. fem. sg.); *-tū* (2nd pers. masc. pl. before pronominal suffixes and isolated): ر ايتو ا 'you have seen' (Blau 1966:145–146; Knutsson 1974:119), as in Modern Arabic dialects.

ii. The verb in the imperfect: elision of final short vowels: *yeqdir* 'he can' (8th century, Violet 1901), *yajīb* 'he will bring' (Blau 1966:147–149).

iii. Form I occurs instead of Classical Arabic Form IV, especially in geminate and weak verbs: حسست 'you felt'. Verbs II*w* keep *ī* in the imperfect as in Classical Arabic, but initial *ʾalif* disappears: وطيعوه 'and obey him' (since the shift from Form I to Form IV is widespread in Christian Middle Arabic, the shift from Form IV to Form I could be a hypercorrection, Blau 1966:152–154).

iv. Form II may be used with the same meaning as Form I: ضرر 'to do harm' (Blau 1966:155).

v. Form III may be followed by a preposition marking the person affected by the action, e.g. وخالط الكل بالكل 'and the whole mingled with the whole' (Blau 1966:157).

vi. Form IV appears very often instead of the Classical Form I: اعنا 'to mean', اشفا 'to cure', اقيل 'it was said' (the latter may have been reshaped originally according to the pattern *u-i-a* of the sound verbs: *ʾuqīla* like *kutiba, ʾuktiba*, and then produced new derivations, cf. Blau 1966:157–162). It may be concluded that in spite of the replacement of Form IV by Form I in Christian Middle Arabic, cases of the opposite shift occur either by hypercorrection or by orthographic confusion through the adding of a prosthetic *ʾalif*.

vii. The perfect and imperative of Form V and VI begin with *ʾit* – instead of *ta-*: فاتطهر 'he was clean', اتكلم 'speak!', اتراافت 'I have been gracious', اتباعدوا عني 'let me alone!' (Blau 1966:163–165).

viii. In geminate verbs, the two identical consonants may occur not contracted after long *ā*: محاججة 'argumentation' (Blau 1966:167). In Late Christian Middle Arabic, the 1st and 2nd persons in the perfect are formed according to the pattern of verbs III*y*: كبيبتيني instead of كببتيني 'you have overturned me' (Knutsson 1974:127).

ix. Verbs Iʾ: in Forms II, III, IV, V, and VIII these shift to the category of verbs I*w*: تو امر 'to hold council', اتمن 'to entrust'. Form VIII may have the pattern ايتخذ instead of Classical Arabic اتخذ 'to take' (Blau 1966:168–172, 183) or the *maṣdar* ايتحاد 'unity' (Grand'Henry 1996, 7, n. 11).

x. The imperfect of Form II of verbs Iʾ may be spelled without the first radical consonant, e.g. حتى اسسكم 'that I establish you' (Blau 1966:171).

xi. Verbs IIʾ pass into verbs II*w/y*: تروس 'to become chief' (on the various forms of *raʾā* 'to see' in Christian Middle Arabic, see Blau 1966:173–176).

xii. Verbs IIIʾ often pass into verbs III*y*: اخطينا 'we have sinned' (Classical Arabic *jāʾa* has the forms جيت جا جوا جاوا and *jāʾa bi* 'to come with' produced the new verb *jāb/yijīb* 'to bring' as in Modern Arabic dialects, Blau 1966:176–180).

xiii. Verbs I*w* sometimes retain the *w* in the imperfect: *yajibu* > *yūjib* 'it is necessary', and the passive may be built without *wāw*: *yūṣafu* > *yuṣaf* 'it is described' (Blau 2002:40). The imperatives of Form I are formed with prosthetic *ʾalif*: ارثوا 'inherit!' (Blau 1966:180–184; Knutsson 1974:131).

xiv. Verbs II*w/y*: the breakdown of the mood system results in the occurrence of forms with long vowels where Classical Arabic demands a short one: ولم يقوم 'and he has not got up' (Knutsson 1974:132) (in Early Christian Middle Arabic, hypercorrect *yakun* instead of Classical *yakūnu* occurs as well, Blau 1966:185–188).

xv. Verbs III*w/y*: as in Modern Arabic dialects, there is a total merger of verbs III*w* to verbs III*y*: *ġazawtu* > *ġazayt* 'I raided'. Sometimes, *-iya* > *-ā*: *ʾuʿṭiya* > *ʾuʿṭā* 'is given' (Blau 1966:190–191; Knutsson 1974:134). The 3rd pers. masc. pl. of the perfect of those verbs whose second radical is followed by *i* ends in *-iyū*: بقيوا 'they stayed', نسيوا 'they forgot' (Blau 1966:193; Knutsson 1974:134– 135). Forms like *lam tadrī* 'you did not know' or *ʾaʿṭīnī* 'give me!' may have been formed by analogy to the indicative after the moods had disappeared. Finally, nouns terminating in [-*in*] in Classical Arabic in the nominative/genitive and [-*iyan*] in the accusative, end in [-*ī*] in Christian Middle Arabic: *qāḍin* > *qāḍī* 'judge' (Blau 1966:190–201).

2.3 *Nouns*

i. There are some changes in the gender of the nouns as in modern dialects, e.g. *ḥaql* 'field' and *qamar* 'moon' are used as feminines (Blau 1966:203–204); in the pronouns and the verbs (perfect, imperfect, imperatives, sometimes the participle and the adjective) the feminine plural is replaced by masculine plural: يا بنات اورسلم انظروا وخرجوا 'o you daughters of Jerusalem, go forth and behold!' (Blau 1966:201–209, 2002:41; Knutsson 1974:137).
ii. The dual is about to disappear, being replaced by the plural. As in the masculine sound plural, the oblique case of the dual has superseded the nominative. *-īn* has become the only ending of the masculine sound plural (Blau 1966:224–228, 2002:42; Knutsson 1974:138–139).
iii. In the broken plural, there is a morphological attraction to *ʾafʿāl*: ثلثة ارجال 'three men' (Blau 1966:228).
iv. The *nisba -iyyīn* may be shortened to *-īn*: الكنعانين 'the Canaanites' (Blau 1966:232; Knutsson 1974:144).
v. *ʾafʿal at-tafḍīl* is expressed for 'better/best' and 'worse/worst' by *ʾaxyar* and *ʾašarr* (Blau 1966:233; Knutsson 1974:145).

2.4 *Numerals*

In numerals, *ʾaḥad* > *ḥad* 'one', and the composed numerals of dialectal type appear: *ḥdaʿšar* 'eleven', *iṯnāʿšar* 'twelve'. Numerals are used without Classical Arabic concord (Blau 1966: 236–241; Grand'Henry 1996:9, n. 19).

2.5 *Prepositions*

For prepositions (Blau 1966:241–254), some remarkable uses in Early Christian Middle Arabic should be mentioned: *bi-* may replace *fī*, but more often, *fī* replaces *bi-; fī* may replace *ʿalā; ka-miṯl* replaces *ka-* (*ka-miṯl* was already obligatory in Classical Arabic before pronouns, e.g. *ka-miṯli-hi* instead of **ka-hu*); *li-* replaces *ʾilā; bi-manzila* is used with the sense of 'as, like'; *li-makān* is used as a preposition denoting cause in Ancient South Palestinian (Blau 1966:253), which seems to be related with موضع 'because' in Late Christian Middle Arabic (Grand'Henry 1996:27, n. 1; Lentin 1997:I, 275).

3. SYNTAX

i. Mood endings: The disappearance of mood endings is reflected by the omission of final short vowels and the absence of differentiation between imperfect forms terminating in endings with and without *n*: *lam tuqirrūna* 'you have not acknowledged' (Blau 2002:45); ولم تسمعون لقولي 'but you did not listen to my word' (Knutsson 1974:148); اشربوا من الا وعية الدي يشربوا 'and drink from the vessels, from which my servants are drinking!' (Bengtsson 1995:141).
ii. Concord: whereas in Classical Arabic a verb preceding its subject is put in the singular, in Middle Arabic the preceding verb tends to agree in number with the following subject: ويقولون اناس 'some say' (Blau 1967a:277); فقالوا بنو يهودا 'and the children of Judah said' (Knutsson 1974:152); فتزوجا ابناها 'and her two sons married' (Bengtsson 1995:143). Concord of things is sometimes expressed by a plural: وشفراته التي قد سنهم 'and his swords, which he has sharpened' (Blau 1967a:286); بنوا القرى وسكنوهم 'they erected the villages and dwelt in them' (Knutsson 1974:157).
iii. *aysa* has become invariable: فليس هي والعقل شي واحد 'it and reason are not the same thing' (Blau 1967a:307); فانك ليس تموت الان 'for you are not going to die now' (Knutsson 1974: 158).
iv. Loss of case endings: the accusative *ʾalif* is sometimes omitted in indefinite triptotic nouns: وان كان حسن بهيا 'and if he is beautiful and pretty' (Blau 1967a:324); لان الرب اهلك سبط 'since the Lord has destroyed a tribe' (Knutsson 1974:159). On the other hand, Christian Middle Arabic uses *-an* in short words, irrespective of case: ولم يكن احدا ياكل 'and nobody ate' (Blau 1967b:327); لم يكن ملكا في بيت اسراييل 'there was no king in the house of Israel' (Knutsson 1974:165).
v. *ʾiḍāfa*: sometimes, the definite article is added to the *status constructus*: بالاصطبار والعمل الصالح 'by patience and well doing' (Blau 1967a:351); المراكب الحديد 'chariots of iron' (Knutsson 1974:168).
vi. Syntax of numerals: some changes point to a more dialectal syntax of numerals, combined with hypercorrections: سبع سوابع 'seven weeks', اثنا عشر ساعة 'twelve hours', اثني وعشرين قرية 'twenty-two villages', سبعين

ملك ‘seventy kings’ (Blau 1967a:371–372; Knutsson 1974:168–171; see also Grand’ Henry 1996:9, n. 20, 41, n. 103).

vii. Presentatives: *hāḏā* and *huwa-ḏā* are to be mentioned: هذ ا انا ‘behold, here I am’; هوذا السباع ‘behold, the beasts’ (Blau 1967a: 463, 465); فهذ ا قد ارسلتك ‘and behold, I havesent you’; فهذ ا انا اضع ‘and behold, I will put’ (Knutsson 1974:177).

viii. Subordinate asyndetic clauses: these occur in Classical Arabic as well, but less frequently than in Christian Middle Arabic or in Modern Arabic dialects (→ serial verbs): هذ ا يعلم الناس يعبدون ‘this man teaches men to worship God’ (Blau 1967b:498); فما نقدر نزوجهم ‘we, however, cannot marry them’ (Knutsson 1974:177).

ix. Relative clauses: *allaḏī* has become invariable and no longer agrees with the antecedent in number, gender, and case; جميع الا فاق الذي فر قتكم فيها ‘all the countries where I have scattered you’ (Blau 1967b: 550); الامورانيين الذي اهلكهم الله ‘the Amorites, whom God has destroyed’ (Knutsson 1974:178).

x. Temporal clauses: *ḥayṯu*, which originally had a local sense, has come to be used as a temporal conjunction: حيث ‘[. . .] since, as’ (Blau 2002:198); حيث ضاقت بكم الامور ‘when you are in distress’ (Knutsson 1974:183).

xi. Causal clauses: *ḥīna* seems still to be temporal in Early Christian Middle Arabic (*ḥīna* ‘when’, Blau 2002:198), but may shift to a causal meaning in Late Christian Middle Arabic: يقضي عليه البعل حين عقر مذبحه ‘let Baal pronounce judgment against him, because he has cut down his altar’ (Knutsson 1974:184).

Most of the linguistic peculiarities described here also appear in Christian Middle Arabic texts of Coptic origin (Blau 1979:215–262, 2002:155–167; Retsö 1984:317–337; Czapkiewicz 1984:27–46).

BIBLIOGRAPHICAL REFERENCES

Bengtsson, Per Å. 1995. *Two Arabic versions of the Book of Ruth: Text edition and language studies.* Lund: University of Lund.

Blau, Joshua. 1966. *A grammar of Christian Arabic based mainly on South-Palestinian texts from the first millennium*, I. Louvain: Imprimerie Orientaliste. [= §§ 1–169.]

——. 1967a. *A grammar of Christian Arabic based mainly on South-Palestinian texts from the first millennium*, II. Louvain: Imprimerie Orientaliste. [= §§ 170–368.]

——. 1967b. *A grammar of Christian Arabic based mainly on South Palestinian texts from the first millennium*, III. Louvain: Imprimerie Orientaliste. [= §§ 369–535.]

——. 1979. “Some observations on Middle Arabic Egyptian texts in Coptic characters”. *Jerusalem Studies in Arabic and Islam* 1.215–262.

——. 2002. *A handbook of early Middle Arabic.* Jerusalem: Hebrew University of Jerusalem.

Czapkiewicz, Andrzej. 1984. “The language of al-Maqaffaʿ’s *Sîrat Mârî Marqus*: A contribution to the studies on Christian Middle Arabic”. *Rocznik Orientalistyczny* 43.27–46.

Fischer, Wolfdietrich. 1991. “What is Middle Arabic?”. *Semitic studies in honor of Wolf Leslau on the occasion of his eighty-fifth birthday, November 14th, 1991*, ed. Alan S. Kaye, I, 430–436. Wiesbaden: O. Harrassowitz.

Grand’Henry, Jacques. 1984. “Traits linguistiques de la version arabe du discours 24 de Grégoire de Nazianze”. *Studi in onore di Francesco Gabrieli nel suo ottantesimo compleanno*, ed. Renato Traini, 389–410. Rome: University of Rome.

——. 1988. “La version arabe du discours 24 de Grégoire de Nazianze: Edition critique, commentaires et traduction”. *Versiones orientales, repertorium ibericum et studia ad editiones curandas. Corpus Nazianzenum* (= *Corpus Christianorum, Series Graeca*, 20), ed. Bernard Coulie, I, 197–291. Turnhout: Brepols.

—— (ed.). 1996. *Sancti Gregorii Nazianzeni Opera: Versio arabica antiqua.* I. *Oratio XXI (arab. 20).* (= *Corpus Christianorum, Series Graeca* 34, *Corpus Nazianzenum*, 4, ed.) Turnhout: Brepols and Leuven: University Press.

Haddad, Rachid. 1992. “La phonétique de l’arabe chrétien vers 700”. *La Syrie de Byzance à l’islam VII^{e}–VIII^{e} siècles: Actes du Colloque international Lyon – Maison de l’Orient méditerranéen, Paris – Institut du monde arabe 11–15 septembre 1990*, ed. Pierre Canivet and Jean-Paul Rey-Coquais, 159–164. Damascus: Institut Français de Damas.

Knutsson, Bengt. 1974. *Studies in the text and language of three Syriac–Arabic versions of the Book of Judicum with special reference to the Middle Arabic elements: Introduction, linguistic notes, texts.* Lund: University of Lund.

Lentin, Jérôme. 1997. *Recherches sur l’histoire de la langue arabe au Proche-Orient à l’époque moderne.* 2 vols. Thèse pour le Doctorat d’Etat ès-lettres, Université de la Sorbonne Nouvelle, Paris III. Lille: Atelier National de Reproduction des Thèses.

Retsö, Jan. 1984. “Middle Arabic in a Coptic–Arabic manuscript in the ‘Röhsska Konstlöjdmuséet’, Göteborg, Sweden”. *Studia Orientalia* 55.317–337.

Violet, Bruno. 1901. “Ein zweisprachiges Psalmfragment aus Damaskus”. *Orientalische Litteraturzeitung* 4.384–403; 425–441; 475–488.

JACQUES GRAND’HENRY
(Catholic University of Louvain)

Cilician Arabic

1. GENERAL

1.1 *Area and range*

The Arabic dialects of Cilicia (Southern Turkey) are spoken in the three large cities of the Cilician Plain (Çukurova), namely Adana, Tarsus, and Mersin, as well as in about 25 villages situated to the south of these towns (see map). The total number of an estimated 70,000 Arabic speakers comprises three communities who differ in both religion and dialect: 66,500 Nusayri-Alawis, 4,000 Sunnis and 1,000 Christians (the latter two groups found only in Mersin). Cilician Arabic is isolated both from the Arab countries themselves and from other Arabic speaking minorities in → Turkey (Procházka 2002a:2–12). Though Cilician Arabic is, to a great extent, understandable to the Arabs of Antioch (→ Antiochia Arabic), in their communications with other Arabic speakers, particularly farm workers from the Urfa area, Cilician Arabs use Turkish. There are very few external influences, since Literary Arabic is completely unknown and contacts with other Arabs are rare.

1.2 *Speakers*

Most Arabic speakers in Cilicia are farmers. Those living in towns are usually craftsmen and unskilled workers. For religious reasons, both the Alawis and the Christians are strictly endogamic. Because people under 30 years old have completely switched to Turkish, the dialects of Cilicia are a variety of Arabic in danger of becoming extinct in a few decades (see Procházka 1999).

1.3 *Position and linguistic type*

The dialects of all three communities are sedentary dialects belonging to the Syro-Palestinian group. The Alawi dialects exhibit a striking resemblance to those of their co-religionists in the province of Antioch (Arnold 1998) in particular and to the Syrian coastal dialects in general (Behnstedt 1997: map 501). The Sunni dialect of Mersin shares many features with the vernaculars of the larger towns along the Syrian coast (Lattakia, Banyās), and the mixed Christian dialect of this town shows affinities with the Christian dialect of Antioch. Although not very diversified, and mutually fully understandable, the Alawi dialects of Cilicia can be divided into three groups more or less corresponding to the cities of Adana, Tarsus, Mersin, and their surrounding villages (for details see Procházka 2002:4–14). The image of Cilician Arabic in the surrounding society is very poor. It is never used in education or in the media, nor is it ever written.

1.4 *Historical evidence*

The presence of Arabic-speaking people in the region is attested as early as the 17th century. However, apart from general remarks about the language by early Western travelers, Cilician Arabic was not described in any detail until the 1980s.

1.5 *State of research and main sources*

A preliminary sketch of the dialect of Adana was published by Jastrow (1983). A comprehensive grammar, including numerous texts and an analysis of the Turkish impact, is Procházka (2002a). For a description of the overall sociolinguistic situation in the area, see Procházka (1999). The original recordings of many of the texts in Procházka (2002a) are available in the Semitic Language Archive (<http:// www.semarch.uni-hd.de/index.php4>, under Syrien-Palästina, Türkei-Süd).

2. LINGUISTIC DESCRIPTION

The reference dialect for the following description is the Alawi dialect of the city of Adana. Some differences within the Alawi dialects themselves as well as a few striking features of the Sunni dialect of Mersin will also be discussed. The Christian dialect of this city, because of its heterogeneous character will not be treated.

2.1 *Phonology*

2.1.1 Inventory

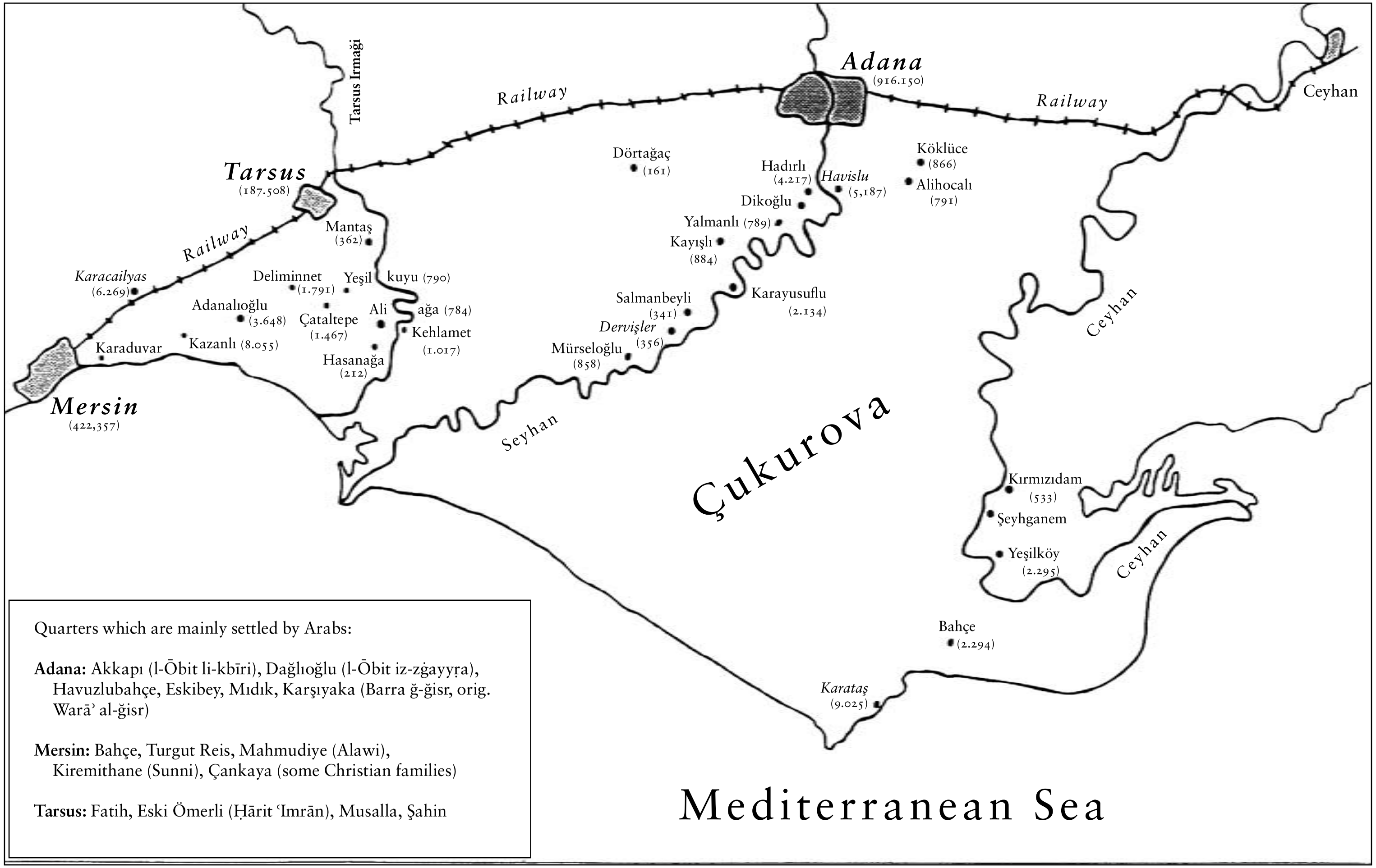
2.1.1.1 Consonants

2.1.1.1.1 Inventory

List of consonants: *b*, *t*, *ṭ*, *d*, *ḍ*, *k*, *q*, *m*, *n*, *r*, *f*, *s*, *z*, *ṣ*, *š*, *j*, *x*, *ġ*, *ḥ*, *ʿ*, *h*, *w*, *y*, *l*.

Marginal consonants: mainly due to assimilatory processes, the following marginal

Map 1. The Cilician Plain


Adana
(916.150)
Railway
Ceyhan
Tarsus Irmaği
Tarsus
(187.508)
Railway
Railway
Köklüce
(866)
Alihocalı
(791)
Dörtağaç
(161)
Hadırlı
(4.217)
Havislu
(5,187)
Dikoğlu
Yalmanlı (789)
Kayışlı
(884)
Karayusuflu
(2.134)
Salmanbeyli
(341)
Dervişler
(356)
Mürseloğlu
(858)
Mantaş
(362)
Karacailyas
(6.269)
Deliminnet
(1.791)
Yeşil
kuyu (790)
Adanalıoğlu
(3.648)
Çataltepe
(1.467)
Ali
ağa (784)
Kehlamet
(1.017)
Kazanlı (8.055)
Karaduvar
Hasanağa
(212)
Mersin
(422,357)
Seyhan
Ceyhan
Çukurova
Kırmızıdam
(533)
Şeyhganem
Yeşilköy
(2.295)
Ceyhan
Bahçe
(2.294)
Karataş
(9.025)
Mediterranean Sea

Quarters which are mainly settled by Arabs:

Adana: Akkapı (l-Ōbit li-kbīri), Dağlıoğlu (l-Ōbit iz-zġayyṛa), Havuzlubahçe, Eskibey, Mıdık, Karşıyaka (Barra ğ-ğisr, orig. Warā᾽ al-ğisr)

Mersin: Bahçe, Turgut Reis, Mahmudiye (Alawi), Kiremithane (Sunni), Çankaya (some Christian families)

Tarsus: Fatih, Eski Ömerli (Ḥārit ʿImrān), Musalla, Şahin

The Cilician Plain (Çukurova) – Turkish Provinces of Adana and İçel
Towns and villages settled by Arabs
(Places with mixed population are in italics; in parentheses are the numbers of inhabitants according to the census of 1990)

consonants are used in Cilician Arabic: ʾ, _v_, _g_, _ẓ_, _č_, _ž_. In addition to these, _p_ is found in borrowings from Turkish.

2.1.1.1.2 Historical remarks

The interdentals have shifted to postdental plosives, both _j_ and _q_ are preserved (only in the Sunni dialect _q_ >ʾ). Typical is velarization and labialization of _b_ (e.g. _bitt ṣaḅ^{w}ī_ 'young girl') and palatalization of _k_ (e.g. _k^{y}ēl_ 'eat!').

2.1.1.1.3 Phonetic realization

Secondary velarization is very common, especially with _r_ (e.g. _baqṛa_ 'cow'), but not phonemic.

2.1.1.1.4 Sociolinguistic variables

By most urban, especially male, speakers _q_ is pronounced as _ḳ_ (Procházka 2002:19). Furthermore, young and urban speakers tend to have a very weak articulation of the pharyngeal and velarized consonants.

2.1.1.2 Vowels

2.1.1.2.1 Inventory

Short: _i_, _a_; long: _ā_, _ē_, _ī_, _ō_, _ū_. Under the impact of Turkish _e_, _ı_, _o_, _ö_, _u_, _ü_ have gained a marginal phonemic status.

2.1.1.2.2 Historical remarks

/i/ < *_i_ and *_u_; regarding the two new phonemes _ē_ and _ō_: /ē/ < *ā by ʾ_imāla_ and, in some cases < *_ay_; /ō/ < *_aw_ and, pre-pausal, < *ū. Final *_ā_ and *_āʾ_ have changed to /i/ if the preceding syllable contains an historical /i/ (not /i/ < *_u_), otherwise they have remained /a/, e.g. _dinyi_ 'world', _šiti_ 'winter', _maʿna_ 'sense', _sama_ 'sky'. Exceptions are the feminine forms of color adjectives where *_āʾ_ > /i/ after front and non-emphatic consonants, e.g. _sawdi_ 'black', _ḥamṛa_ 'red'.

2.1.1.2.3 Phonetic realization

It is worth noting that in the environment of front consonants /ō/ and /ū/ are often pronounced as _ȫ_ and _ǖ_, e.g., _tȫm_ 'twin' and _tkǖn_ 'she is'.

2.1.1.2.4 Distribution

Short /a/ is subject to changes conditioned by syllable types and/or consonantal environment. In pre-tonic closed syllables /a/ becomes /i/ not only when followed by _a_, _ā_, _ē_, _ay_ (e.g. _ʿiṭšān_ 'thirsty', _txibbayt_ 'I hid myself') but also in the patterns *_mafʿūl_ and *_tafʿīl_ (e.g. _miftūḥ_ 'opened').

Restricted to noun patterns, a shift of penultimate /a/ > /i/ occurs in the environment of front consonants, e.g. *_jabal_ > _jabil_ 'mountain'. In Tarsus and several neighboring villages, however, penultimate /a/ is split into _i_, _a_, or _o_, conditioned by the adjacent consonants (details in Procházka 2002a:28–30), e.g. *_laban_ > _labin_ 'yoghurt', *_ʾazraq_ > _ʾazraq_ 'blue', *_ḥaṭab_ > _ḥaṭob_ 'wood'. This dialect also shows the shift of penultimate /ā/ > _ō_ in the environment of back or velarized consonants, e.g. _silṭōn_ 'sultan', _ḥīṭōn_ 'walls'.

2.1.1.3 Diphthongs

There are two diphthongs, /ay/ and /aw/. Cilician Arabic exhibits the rare feature that the two diphthongs have not developed in a parallel way. Roughly speaking, /ay/ is usually retained but /aw/ appears only in open syllables and has been monophthongized in closed syllables, e.g. _sayf_ 'sword', _mōt_ 'death', _dawṛa_ 'excursion'. For the numerous exceptions and the historical development see Procházka (2002a:51–55). For diphthongs in pausal forms, see 2.1.2.2.

2.1.1.4 Consonant clusters

The syllables CCC and -CC remain unchanged unless C2 is _l_, _m_, _n_, _r_, e.g. _kalbkin_ 'your dog', but _baġílkin_ 'your mule'; _milḥ_ 'salt', but _baḥir_ 'sea'. In CC- mostly an anaptyctic vowel appears, e.g. _ᵊbnāt_ 'girls'.

2.1.1.5 Stress

The stress lies on the last syllable if it is long (i.e. v̄C, vCC), otherwise it is on the penultimate syllable.

2.1.2 Phonotactics

2.1.2.1 Conditioned ʾ_imāla_

Cilician Arabic has a conditioned → ʾ_imāla_, i.e. /ā/ usually shifts to /ē/ if there is or was a short or long /i/ (but not /i/ < *_u_, *_a_) in an adjacent syllable, e.g., *_lisān_ > _lsēn_ 'tongue', *_sakākīn_ > _skēkīn_ 'knives', *_fātiḥ_ > _fētiḥ_ 'opening', _nijjār_, but _nijjērīn_ 'carpenters', *_yuʿāwinu_ > _yʿēwin_ 'he helps'. There are many exceptions, some of them phonologically motivated, some of them lexicalized (see Procházka 2002a:40–47), e.g. ʾ_imāla_ never occurs in the sequence _āyi_ and next to _ṛ_. In the Sunni dialects of Mersin the ʾ_imāla_ is more or less restricted to the patterns _fiʿāl_/_fiʿlāl_ (Procházka 2002a:205–206).

2.1.2.2 Pausal forms

Final vowels in pause are lowered and/or diphthongized. Frequently found pausal phenomena, typical for all Alawi dialects, are: *-a* > *aw*, *-i* > *-e/-ey*, *-u* > *-o/ow*, and the same for long vowels. In Tarsus *i* < *i* is not lowered, but only diphthongized: e.g., *bayti* > *baytay* ‘my house’.

Vowels which historically were not in final position are never subject to pausal change. This has led to a secondary phonemization of pausal phenomena in forms such as *nsē* ‘forget [fem.]!’ < **insī*, versus *nsī* ‘forget [fem.] him!’ < **insīh*.

2.1.3 Morphophonology

2.1.3.1 Elision

Cilician Arabic is a *non-différentiel* dialect, i.e., both *a* and *i* are usually elided in open unstressed syllables, e.g. **katábt* > *ktabt* ‘I wrote, **simíʿt* > *smíʿt* ‘I heard’. Exceptions are found among several verbal forms. In the imperfect of Form I the basic vowel is stressed and therefore not elided, e.g. *yismáʿu* ‘they hear’, *yiktíbu* ‘they write’. In the perfect *a* and *i* are elided when an inflectional suffix is attached, but preserved when a pronominal suffix is attached (Procházka 2002a: 32–37, 106–108), e.g. *qatlu* ‘they hit’ versus *qatálu* ‘he hit him’, *simʿu* ‘they heard’ versus *simíʿu* ‘he heard him’. In contrast to other Syrian dialects, this phenomenon occurs also in the derived forms, e.g. *mawwáta* ‘he killed her’, *ʿēwánu* ‘he helped him’.

2.1.3.2 Shortening and lengthening

Both phenomena are restricted to a very few cases, especially the shortening of *ā* before geminated consonants and the lengthening of short vowels in the imperative sg. masc. of Form I (see 2.2.6.2.i).

2.1.3.3 Suffixation

3rd pers. sg. fem. perfect: In Form I, the form remains unchanged when combined with a consonantal suffix, but is lengthened when combined with a vocalic suffix, e.g. *qatlitkin*, *qatlītu* (some speakers *qatiltu*) ‘she hit you/him’. In the derived forms usually both kinds of suffixes cause a metathesis: *ftahmit* > *ftahimtu*, *ftahimtkin* ~ *(ftahmitkin)* ‘she understood him/you’.

3rd pers. pl. perfect: no pecularities, e.g. *qatlu* ‘they hit’, *qatlūwa* ‘they hit her’ (for suffixation to participles see 2.2.6.3).

2.2 *Morphology*

In contrast to the related dialects along the Syrian coast, there is no gender distinction in the 2nd and 3rd persons plural.

2.2.1 Pronouns

2.2.1.1 Personal independent

A typical feature of all Alawi dialects is the initial *h-* in all 2nd persons. This can be explained by analogy with the pronouns of the 3rd persons.

3rd sg .masc. *hūwi~hū*
3rd sg. fem. *hīyi~hī*
3rd pl. *hinni(n)*
2nd sg. masc. *hint*
2nd sg. fem. *hinti*
2nd pl. *hintu*
1st sg. *ana*
1st pl. *niḥna*

In the Sunni dialect the 2nd persons are *inte~itte*, *inti~itti*, *intu~ittu*.

2.2.1.2 Possessive/object suffixes

There exist morphological variants depending on the final sound of the word to which the suffix is attached (for details see Procházka 2002a: 64–67). Worth mentioning is that the historical *h-* in the 3rd pers. sg. masc./fem. and 3rd pers. pl. has disappeared in all forms following a consonant and in most forms following a vowel.

Suffixes after -C:
3rd sg. masc. *-u*
3rd sg. fem. *-a*
3rd pl. *-in*
2nd sg. masc. *-ak*
2nd sg. fem. *-ik*
2nd pl. *-kin*
1st sg. *-i/-ni*
1st pl. *-na*

2.2.1.3 Indirect object suffixes

These are formed by attaching the above-mentioned suffixes to the basis *l-*: e.g. *katabli* ‘he wrote me’. Combined with 3rd pers. sg. fem., suffixes with an initial consonant often cause a metathesis: e.g., *jābiltkin* ~ *jābitilkin* ‘she brought to you’ (Procházka 2002:111–112).

2.2.1.4 Demonstratives

In the region of Adana the forms for the near deixis are *hāda*, *hādi*, *hawdi*, for the far *hāka*, *hāki*, *hawki*. In the regions of Mersin and Tarsus the forms for the near deixis are *hāda*, *hāya*, *hawdin* ~ *hawdīn*, for the far *hāka*, *hāki*, *hawkīn(i)*. The Sunni dialects have *hāda*, *hādi* ~ *haydi*, *hādōl* ~ *hādōn* and *hādāk*, *hādīk*, *hādōk*.

A short form exists for both the near *(ha-)* and the far *(hāk-)* deixis. The latter is not common in other dialects of the Syro-Palestinian group, e.g. *hāk-irrijjāl* 'that man'.

2.2.1.5 Presentatives

The most common form is the invariable *kwā*. When referring to a person, it often gets the corresponding suffix, e.g. *kanītu*, *kanīta*, *kanītin* 'here he, she is/they are!' There exists a variant *kwanīt-*; and in the village of Kazanlı *kahnīt-* is heard (Procházka 2002:150).

2.2.1.6 Relative pronoun

The relative pronoun *il* is identical with the definite article and causes the use of the constructus state (see 2.3.1). Without a head noun *hal* is usually used, e.g. *hal b-yirkiš hēk fī* 'which he digs with'.

2.2.1.7 Interrogative pronouns

mīn 'who?', *šū* 'what?', *ayna* 'which?', e.g. *bi ayna ḥāṛa qēˁid* 'in which neighborhood does he live?' (the Sunnis have masc. *ēnu*, fem. *ēna*), and *šikil* 'what kind of?' (e.g. *b šikil ˁarabāy* 'with what kind of car?').

2.2.2 Adverbs

2.2.2.1 Temporal

Besides the widespread *hallaq* (< *halwaqt*) 'now', Cilician Arabic uses some very special temporal adverbs, among them *killmaš* 'always', *mnistaxx* 'just now' (< **min sāˁit ilāxar*), *baˁda* 'after that', *min ˁatīq* 'formerly', *ilġada* 'at noon', *illēli* 'today', *illēli ˁalambikra* 'this morning' (Procházka 2002:163).

2.2.2.2 Local

hōn(i), in Kayışlı *hōnit* 'here'; *hōnīk(i) hawnīk*, (rarely) *hnīk* 'there'. A third deixis 'just here' exhibits the interesting extended forms of *hōn: hōnstīni*, *hōnstayni*, *hōnaystūni* (Procházka 2002a:133).

2.2.2.3 Manner

hēk(i), in Mürseloğlu *hākihni* 'so'; *ḥabbūši* 'a little'. *kmayn* 'also' is often found with additional endings when emphasized: *kmaynta*, *kmayntanax*.

2.2.2.4 Interrogatives

lašū 'why?'; *škīf* 'how?'; *wayn* 'where?'; *ašqad*, *ašqa*, *ašqay* 'how much?' (Procházka 2002a:135); *īmtan* ~ *ēmtan* 'when?'.

2.2.3 Particles

2.2.3.1 Article

There is only a definite article *il*, shortened after a preceding vowel to *l* and assimilated to the traditional sun-letters and to *j*, *č*.

2.2.3.2 Genitive marker

No genitive marker, but a special analytic form of the genitive is used (see 2.3.1).

2.2.3.3 Negations (see 2.3.5.5)

2.2.3.4 Existentials

fī, *māfi* ~ *ʾāfi* 'there is, there is not'.

2.2.3.5 Prepositions

The following innovations are worth noting: the functions of *la* and *ˁala* have merged to a high degree (Procházka 2002a:136–138), and *qiddām* is used in local and temporal sense, e.g., *qiddām ilġada* 'before noon'. *xalf* 'behind'; *fōq* 'above, upon', e.g. *fōq ilmasāy* 'upon the table'; *jōfāt* 'inside of'; *mḥēdi* 'next to'; *qāršīt* (< Turkish *karşı*) 'opposite to'.

2.2.3.6 Conjunctions

Besides several Turkish loans, e.g. *čünkü* 'because of' and *keške* 'if [for wishes]', Cilician Arabic uses some other peculiar forms: *tak-ma* (Sunnis: *bēn-ma*) 'until', *miššī* 'the moment', *ˁišt* 'because of'.

2.2.4 Noun

2.2.4.1 Gender

Feminina without the fem. marker *-a/-i* are all nouns which denote animate beings of female sex, e.g. *ˁarūs* 'bride', as well as expressions for several paired parts of the body: *dān* 'ear', *ˁayn* 'eye', *ʾīd* 'hand', *ʾijir* 'foot', *ʾiṣbiˁ* 'finger', *ṭīẓ* 'buttocks'. Besides the common unmarked feminine nouns *(nār*, *ʾarḍ*, *šams*, *rūḥ)*, *darb* 'road', *mayy* 'water', *kirsi* 'chair', and – with some exceptions – *sikkīn* 'knife' and *ṭrīq* 'time *[fois]*' are also feminine. Some of these nouns usually get a *t*-suffix in the construct state, e.g. *ʾiṣbiʿti* ‘my finger’.

2.2.4.2 Productive patterns

Cilician Arabic has lost the ability for the creation of new words because lexical gaps are almost always filled by loans from Turkish. Neither diminutives nor the so-called *nisba*-ending are productive any more. Petrified remains, however, are found in a few words, such as the diminutive *xbayzi* ‘bread’ and the *nisba*-form *tirki* ‘Turk’.

2.2.4.3 External and internal plural

Like other peripheral dialects of Arabic, Cilician Arabic shows a tendency toward the use of external instead of internal plurals. Above all, the feminine plural morpheme *-āt* is widely used, not only for loanwords (e.g. *čatalāt* ‘forks’) but also together with or instead of internal plurals (e.g. *jbēlāt* ‘mountains’, *rāsāt* ‘heads’.) Except for *kbār* ‘big’ and *zġār*, ‘small’ internal plurals of adjectives are not found (Procházka 2002:119).

2.2.4.4 Unusual plural patterns

A peculiarity of Cilician Arabic are *faʿal* and *fīʿāl* (Procházka 2002:120). The first corresponds to the singulars *fiʿla* and *faʿla* and thus has replaced the pattern *fiʿal*, found in other Syro-Palestinian dialects, e.g. *ʾibri-ʾabar* ‘needles’, *kirsi-karas* ‘chairs’, *salli-salal* ‘baskets’. *fīʿāl* is only found in *ṣībāʿ* ‘fingers’ and *ḍīfār* ‘fingernails’.

2.2.4.5 Pseudo-dual

The → pseudo-dual appears only with *dīnayn* ‘(two) ears’; all other nouns denoting paired parts of the body have internal plural forms, e.g. *diyyāt* ‘hands’.

2.2.4.6 Colors and deficiencies

Except *ʾaswid* ‘black’, all words of this category have the pattern *ʾafʿal*.

2.2.5 Numerals

The feminine of *wēḥid* ‘one’ has the unique form *wēdi* (discussed in Procházka 2002a:72). Numerals from 3 to 10 and from 11 to 19 have two different sets, depending upon how the number is used: either as a pure number in isolation, or as the modifier of a noun. Thus in 3–10 the isolated forms have the feminine ending (e.g. *tlāti*, *ʾarbʿa* ‘three’, ‘four’) disappearing in the construct state except when combined with *ʾālāf* ‘thousands’ and *ʾīyēm* ‘days’ (e.g. *xamst īyēm* ‘five days’.)

In 11–19, the set used for counting a noun has preserved the final *-r*, e.g., *tminṭaʿš* ‘18’ versus *tminṭaʿšir zalmi* ‘18 men’. The word for ‘100’ is *mī* (<**mīyi*).

2.2.5.1 Ordinal numbers

As in some other peripheral dialects (e.g. → Cypriot Arabic, Borg 1985:129), all ordinal numbers except *awwil* ‘first’ and *tēni* ~ *tāni* ‘second’ have been borrowed from Turkish (see Procházka 2002:149–150,186).

2.2.5.2 Count nouns

Several characteristic count nouns exist to express a unit of a collective, e.g. *qirṣ bandūra* ‘one tomato’, *ʿirq ᵊflayfli* ‘one pepper’.

2.2.6 Verb

2.2.6.1 Forms

2.2.6.1.1 Form I

In both the perfect and the imperfect there are two bases: *faʿal*, *fiʿil* and *yifʿal*, *yifʿil*. Complementary combination of perfect and imperfect forms is preferred, i.e. most verbs are either of the *faʿal/yifʿil* or the *fiʿil/yifʿal* type. Except in combination with final weak roots, the latter type is almost exclusively used for intransitive verbs. In numerous cases Old Arabic intransitive verbs of the *a*-type have been transformed to the *i*-type, e.g. *rijiʿ* ‘to come back’, *rigid* ‘to run’, *šiʿil* ‘to burn [intr.]’ (Old Arabic *šaʿala* ‘to light’).

In only two verbs does vowel alternation in the basis express different concepts: *ġala/yiġli* ‘to boil [trans.]’ versus *ġili/yiġla* ‘to boil [intr.]’.

2.2.6.1.2 Derived Forms

While Form IX *fʿall/yifʿall* is still found in finite verbs (participles are formed in Form II), there is no evidence for Form IV and, except in the Sunni dialect, only one single verb in Form X (see Procházka 2002:96). Form VIII *ftaʿal/yiftiʿil* is quite rare and not productive.

Form II *(faʿʿal/yfaʿʿil)* is very frequent and often causative to Form I, e.g. *fāt* ‘enter’-*fawwat* ‘to let enter’, or denominal. Besides many transitive verbs, a few frequent intransitive verbs are found, e.g. *waqqaf* ‘to stand’. In Forms III and VI some verbs have an *ʾimāla*, e.g. *fāʿal* ~ *fēʿil/yfāʿal* ~ *yfēʿil* and *tfāˁal* ~ *tfēˁil/yitfāˁal* ~ *yitfēˁil*. Form V *tfaˁˁal/yitfaˁˁal* serves mostly for reflexive or passive formations corresponding to Form II, e.g. *xassal* ‘to wash’-*txassal* ‘to wash oneself’. Form VII *nfaˁal/yinfiˁil* is very productive and functions primarily to form passive verbs corresponding to Form I, even if the first radical is *n*, *l*, e.g. *nisi* ‘to forget’-*nnasa* ‘to be forgotten’, *lamm* ‘to collect’ -*nlamm* ‘to be collected’.

2.2.6.2 Inflection of imperfect and perfect

2.2.6.2.1 Imperfect: paradigm

3. sg. masc. *yiftaḥ*
3. sg. fem. *tiftaḥ*
3. pl. *yiftáḥu*
2. sg. masc. *tiftaḥ*
2. sg. fem. *tiftáḥi*
2. pl. *tiftáḥu*
1. sg. *iftaḥ* ‘I open’
1. pl. *niftaḥ*
Imperatives: *ftāḥ*, *ftaḥi*, *ftaḥu*

Thus, Cilician Arabic belongs to the *iktib-niktib* type. In both *a*- and *i*-bases, the form vowel is retained in the whole paradigm by means of stressing it in those forms with an inflectional suffix. Verbal prefixes see 2.3.5.1.

2.2.6.2.2 Perfect: paradigm

3. sg. masc. *fataḥ*
3. sg. fem. *fatḥit*
3. pl. *fatḥu*
2. sg. masc. *ftaḥt*
2. sg. fem. *ftaḥti*
2. pl. *ftaḥtu*
1. sg. *ftaḥt*
1. pl. *ftaḥna*

A very interesting feature of the Sunni dialect of Mersin is that, in most cases, a distinction is made between 1st pers. sg. and 2nd pers. sg. masc., e.g. *ftaḥit* ‘I opened’ versus *ftaḥt* ‘you opened’ (Procházka 2002a:208).

2.2.6.3 Participles

The active participle of Form I is *fēˁil*; for intransitive verbs denoting a condition *fiˁlān* is used, e.g. *mirḍān* ‘being ill’. In the Mersin area the latter occurs in many verbs of the *fiˁil*-type, e.g. *širbān* ‘drinking’. The passive participle has the pattern *mifˁūl*, e.g. *miftūḥ* ‘opened’, exhibiting the unique feature of a *mi*-prefix where other dialects normally have *ma*-. In the derived forms, no difference between active and passive participles exists, e.g. *mˁammar* ‘building, built’, *mˁēwan* ‘helping, being helped’.

Participles of transitive verbs can get suffixes, e.g. *rēmīha* ‘he has left her’. Sg. fem. participles for the 1st and 3rd persons with suffixes are formed regularly, e.g. *qētli* > *qētlitni*, *qētiltu* ‘she has beaten me/him’. For the 2nd person, by analogy with the corresponding finite perfect form, the pattern *fēˁiltī-* is used, e.g. *qētiltīni* ‘you [fem.] have beaten me’.

2.2.6.4 Verbal nouns

Verbal nouns are relatively rare and there is no evidence of them in Forms V–X. In Form I verbal nouns of the following patterns are found: *faˁl*, *fiˁl*, *fˁāl(a)*, *fˁēl(a)*, *fˁūla*, *faˁīla*. For Form II the pattern *tifˁīl* is used, combined with III*w/y* roots *tifˁāy* or *tifˁa/i*, e.g. *tišˁīl* ‘lightning’, *tiswāy* ~ *tiswa* ‘making’, for Form III *mfāˁla/mfēˁla*.

2.2.7 Weak verbs

2.2.7.1 Geminated verbs

All verbs of this category have the same patterns as, e.g. *fazz/yfizz* ‘to jump’, except words meaning ‘to stay’ which have an *a*-imperfect: *yḍamm* (Adana), *yḍall* ~ *ytamm* (Mersin, Tarsus). For the vowel change in the perfect forms, e.g. *ḥaṭṭ*, *ḥiṭṭayt* ‘he, I put’, see 2.1.1.2.4. The active participle is regular, e.g. *ḥēṭiṭ* ‘putting’.

2.2.7.2 Verbs Iʾ

In Form I only two verbs of this type exist: *ʾakal* (variant *kāl*)/*yākil* ‘to eat’ and *ʾaxad/yāxid* ‘to take’. Forms with vocalic inflectional suffixes in the perfect lose their first syllable, e.g. *kalt*, *xadt* ‘I ate, took’. The active participles are *ēkil*, *ēxid* (but Mersin *kāyil*, *xāyid*). The passive participles are formed in Form VIII, e.g. *mittēkal*, *mittēxad*. Imperatives: *kēl*, *kili*, *kilu* ‘eat!’ and *xēd*, *xidi*, *xidu* ‘take!’.

2.2.7.3 Verbs I*w/y*

Except *wijiˁ/yjaˁ* ‘to hurt’, all verbs show the preservation of the initial *w*, although in the imperfect, and above all in the 2nd persons, variants with *ū* are also used, e.g. *wiqiˁ/yiwqaˁ* ~ *yūqaˁ*. The root *wqf* is only combined with Form II, e.g. *waqqaf/ywaqqif* ‘to stop [transitive/intransitive]’. Some Classical Arabic Form VIII verbs appear as Form I verbs with the former *t*-infix as the first root radical, e.g. *taham/yithim* ‘to accuse’ (< *ittahama*).

The only I*y* verb *y* is *yibis/yības* 'to dry'; in the imperfect a restructured variant *yibyas* exists.

2.2.7.4 Verbs II*w/y*

Form I verbs of this category do not show any peculiarities; the imperative has a long vowel in all forms, e.g. *šīl*, *šīli*, *šīlu* 'pull out!' There are, however, a few verbs in which the weak root consonant is retained, e.g. Form I *ṭiwil/yiṭwal* 'to grow up [child]', Form VII *nʿawaj/yinʿiwij* 'to be curved'.

2.2.7.5 Verbs III*w/y*

There exists an *a*- and an *i*-type in both the perfect and the imperfect of Form I, e.g. *baka/yibki* 'to cry', *nisi/yinsa* 'forget'. In the perfect, consonantal inflectional suffixes follow the syllable *-ay-* when combined with *a*-type verbs, and *-ī-* when combined with *i*-type verbs, e.g. *bkayt* 'I cried', *nsīt* 'I forgot'. The root final *y* of the *i*-type is never retained, not even in forms such as *nisu* 'they forgot', where many Syrian dialects (e.g. Damascus) have *nəsyo*. In the imperfect, the forms of the 3rd pers. pl. are the same in both types, e.g. *yinsu* 'they forget' (*a*-type), *yimšu* 'they walk' (*i*-type). The imperatives are *nsā*, *nsī*, *nsū* 'forget!' (*a*-type) and *ʿṭī*, *ʿṭī*, *ʿṭū* 'give!' (*i*-type).

2.2.7.6 Irregular verbs

In Cilician Arabic only the verb 'to come' can be regarded as irregular. The common paradigm for the perfect is sg. *jīt*, *jīt*, *jīti*, *jā*, *jīyit*, pl. *jīna*, *jītu*, *jū*. Worth mentioning is the 3rd pers. sg. fem. *jīyit* 'she came', the possible origin of which is discussed in Procházka (2002a:101). In the dialects of Mersin and Tarsus, however, those forms with a vocalic suffix show a prosthetic vowel, e.g. *ʾija*, *ʾijit*, *ʾiju*. For the imperfect, two bases exist, *yjī* and *yíji*, a fact that can most likely be explained by the mixing of different Syrian Alawi dialects in Cilicia (Procházka 2002a:81). The imperatives are *tʿā*, *tʿī* (in Karaduvar *tʿay*), *tʿū*; the participles are *jāy* (masc./fem.), pl. *jāyīn*.

2.2.8 Quadriradical verbs

These verbs have two forms, *faʿlal/yfaʿlil* and *tfaʿlal/yitfaʿlal*. There is evidence of original quadriradical verbs (e.g. *tbalham* 'to wonder', *tmarjaḥ* 'to roll oneself'), reduplicated verbs (e.g. *tmaṭmaṭ* 'to stretch one's limbs'), onomatopoeic verbs (e.g. *tnaḥnaḥ* 'to clear one's throat'), and loan verbs from Turkish (e.g. *tqaṣqan* 'to be jealous' < *kıskan-mak*). A *fawʿal*-type exists in Turkish loans only, e.g. *tdawšan-yitdawšan* 'to think' (< *düşün-mek*).

A characteristic of the dialects in and around Adana is that their inflection is not parallel to the inflection of Form II, since the original pattern is retained in all forms by the insertion of a stressed anaptyctic vowel in those cases where a vocalic suffix is attached; thus, e.g., *ġarbal*, *ġarbílit* 'he, she sifted' (versus Form II *ʿammar*, *ʿammrit* 'he, she built').

2.3 *Syntax*

2.3.1 Construct state

A typical feature is the use of the construct state together with a determined attribute, e.g. *bayt ilʿatīq* 'the old house', *mġārt lᵊkbīri* 'the big cave'.

A genitive marker does not exist, but if the head noun denotes a human being the noun phrase is often restructured by means of *la* and a proleptic suffix, e.g. *immu la ššayx* 'the mother of the sheikh'.

2.3.2 Elative constructions

Comparative phrases are formed either by an adjective in the → elative followed by *min*, or by a normal adjective followed by *ʿan*, e.g. *halbayt ᵊjdīd ʿan baytna* 'this house is bigger than ours'.

2.3.3 Relative clauses

Contrary to most Arabic dialects the head noun of a determined relative clause appears usually in the construct state followed by the relative pronoun *il* (see also 2.3.1), e.g. *bi mayyt il matiġla* 'with the water which boils', *ʿarabayt il xadta* 'the car which I have bought'.

2.3.4 Verbal phrase

The analytic marking of a direct object by means of the prepositions *la* or *ʿala* is very common in Cilician Arabic, e.g. *šifta la fāṭma* 'I saw Fāṭma', *bi-yḥibbu ʿala xālu* 'he loves his uncle'. Such constructions are extremely frequent when the object denotes a person and, in contrast to other Eastern Arabic dialects (see Levin 1987), this construction is fully grammaticalized and there are no stylistic differences between analytic and synthetic object phrases.

In sentences with two pronominalized objects, the independent pronoun is used for the second object, e.g. *jibtillak hūwi* 'I brought him to you'.

2.3.5 Verbal aspect

2.3.5.1 Indicative prefixes

In the Alawi dialects *ma-* (< *ʿammāl*) is used for the immediate present, e.g. *ma-tiḥmil* 'she is carrying', in the Sunni dialects *ʿab-* or *ʿamb-*. The marker for the habitual present is *b(i)-*, e.g. *b-yiftíhim* 'he understands' (for details see Procházka 2002a:113–116).

2.3.5.2 Future intent prefixes

There exists an invariable prefix *bad-* (in Mersin *baddi-*), e.g. *bad-išrab* 'I will drink', but more frequently used is *badd-* + suffix, e.g. *baddik ᵊtšūfi* 'you [fem.] will see'.

2.3.5.3 Active participle

The active participle is predominantly used in a resultative function (see Wild 1964; Brustad 2000:182–186). However, under the influence of the Turkish evidential (the so-called *miş*-perfect), it is also found instead of a perfect if the speaker wants to indicate that he/she has learned the facts from someone else (see Procházka 2002a:200–201), e.g. in a story about rebirth: *šāyiftu la ššayx*, *ēxdīna lᵊmlāykāt* 'she saw the sheikh, the angels took her'.

2.3.5.4 Auxiliaries

There exist several auxiliaries. Duration is expressed by *ḍamm* or *ḍaqqar* 'to stay', which is followed by an imperfect or a participle, e.g. *sitt ᵊsnīn ḍiqqarna qēʿdīn hōn* 'we have been living there for six years'. The starting point of an action is indicated by a few verbs: *ṣār* 'to become', *qām/yqūm* 'to stand up', *jā/yjī* 'to come', and *rāḥ/yrūḥ* 'to go' (for details see Procházka 2002a:155–156). The invariable *mā baqa* (lit. 'he didn't stay') is used for cessation, e.g. *mā baqa b-tiji* 'she doesn't come any more'.

2.3.5.5 Negation particles

The common negation particle for both nouns and verbs is *mā* or *ʾā*. The latter can be explained by haplology caused by the verbal prefix *ma-* (see 2.3.5.1), i.e. *mā ma-tšūf* > *ʾā ma-tšūf* 'she doesn't see'. *mā/ʾā* is also used for the negation of personal pronouns, e.g. *hint mā kayyis* 'you are not good' (but *māni* 'I am not'). Prohibitions as well as optative and conditional clauses are negated by *lā*.

Together with certain nouns *māla* is used as a general negation, e.g. *māla imkān* 'there is no possibility'.

2.3.6 Word order

Both VSO and SVO word order exists. As in the Eastern Anatolian *qəltu* dialects, SVO has become dominant under the impact of Turkish (see Dahlgren 1998:168), and there are even cases exhibiting the Turkish SOV word order.

2.3.7 Agreement

Plural nouns nearly always agree with plural adjectives and participles and mostly with finite plural verbs, regardless of whether or not they denote human beings, e.g. *byūt ᵊkbār* 'big houses', *lᵊbyūt min waḥil ᵊmsuwwāynīn* 'the houses are made out of mud'; *jū lkaras* 'the chairs came'. Adjectives often agree with feminine plurals, e.g. *bnāt ilkayysāt* 'the beautiful girls', *ḥkāyāt ᵊktīrāt* 'many stories'.

Collective nouns such as *ilʿālam*, *innēs*, *ilmilli* 'the people' agree with either plural or with feminine singular.

2.3.8 'to have'

Possession is expressed by the prepositions *ʿand ~ʿind* and *maʿ* (temporary), *la* indicates inalienable and sometimes also long-term possession, e.g. *ili bitt* 'I have a daughter', *la ššayx bayt ᵊkbīr* 'the sheikh has a big house'.

2.3.9 Adverbial clauses

see 2.2.3.6

2.3.10 Conditional sentences

The conjunctions of the realis are *ʾiz(a)*,*ʾaz(a)*, or *kannu*, followed either by a *b*-imperfect or by a perfect, e.g. *ʾiza bi-trūḥu bi-tšūfū́* 'if you go (there) you will see him'; *ʾaz kān ʾilak dawa* 'if there is a medicine for you'. In the irrealis, invariable *ykūn ~ tkūn* is used, e.g. *ykūn šifta baddak ᵊtqūl* 'if you had seen her, you had said', *lā tkūn jā lḥakīm* 'if the doctor had not come'.

2.3.11 ḥāl-sentences

Syndetic → *ḥāl*-sentences of the structure pronoun-*w*-participle/imperfect are very common, above all in clause initial position, e.g. *niḥna w qēʿdīn hōnīk* 'while we are sitting there'; *santayn hū w mrīḍ* 'he has been ill for two years'; and *hī w ma-tjību* 'while she is giving birth to him'.

Asyndetic constructions are restricted to sentences with two different subjects, e.g. *šāfit wlāda qēˁdīn hōnīk* ‘she saw her children sitting there’.

3. Lexicon

The lexicon contains many historical loans from Aramaic (→ Aramaic loanwords) and a very large and still increasing number of modern loans from Turkish which are not restricted to nouns denoting things of modern life but include numerous verbs, adjectives, and particles (for a detailed analysis see Procházka 2002a:187–199).

Bibliographical references

Arnold, Werner. 1998. *Die arabischen Dialekte Antiochiens*. Wiesbaden: O. Harrassowitz.

Behnstedt, Peter. 1997. *Sprachatlas von Syrien*. Wiesbaden: O. Harrassowitz.

Borg, Alexander. 1985. *Cypriot Arabic*. Stuttgart: F. Steiner.

Brustad, Kristen. 2000. *The syntax of spoken Arabic: A comprehensive study of Moroccan, Egyptian, Syrian, and Kuwaiti dialects*. Washington, D.C.: Georgetown University Press.

Dahlgren, Sven-Olof. 1998. *Word order in Arabic*. Gothenburg: Acta Universitatis Gothoburgensis.

Jastrow, Otto. 1983. “Beobachtungen zum arabischen Dialekt von Adana”. *Zeitschrift für Arabische Linguistik*. 11.72–29.

Levin, Aryeh. 1987. “The particle *la* as an object marker in some Arabic dialects of the Galilee”. *Zeitschrift für Arabische Linguistik* 17.31–40.

Procházka, Stephan. 1999. “From language contact to language death: The example of the Arabic spoken in Cilicia (Southern Turkey)”. *Orientalia Suecana* 48.115–125.

——. 2002a. *Die arabischen Dialekte der Çukurova (Südtürkei)*. Wiesbaden: O. Harrassowitz.

——. 2002b. “Die Alawitendialekte der Çukurova: Vergleichsmaterialien zu Peter Behnstedts *Sprachatlas von Syrien*”. *Wiener Zeitschrift für die Kunde des Morgenlandes* 92.91–109.

Wild, Stefan. 1964. “Die resultative Funktion des aktiven Partizips in den syrisch-palästinensischen Dialekten des Arabischen”. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 114.239–254.

Stephan Procházka (University of Vienna)

Circumstantial Clause → Ḥāl; Subordination

Classical Arabic

1. General definitions

Classical Arabic designates that form of Arabic which was described by the Arab grammarians of the 8th century and called by them → *al-ˁarabiyya*. They regarded this as the only correct Arabic language. Western scholars call it Classical Arabic to differentiate it from the Arabic vernaculars of the neo-Arabic type. It is the language in which the Arabic texts of pre-Islamic and early Islamic times were handed down, first of all the *Qurʾān* and pre-Islamic and early Islamic poetry, but also the historical and legal traditions of that time. In the process of describing Classical Arabic, the Arab grammarians standardized the language, and in this standardized form it became the educated and official language of Islamic civilization and later on, the written standard language of the Arabic-speaking world. From that time to the present, Classical Arabic has remained outwardly almost unchanged. At the latest since the standardization of Classical Arabic, another Arabic tongue has coexisted with this standardized language of school and education, a vernacular Arabic more or less similar to the modern Arabic dialects that existed exclusively in spoken and not in written form. The → diglossia resulting from this situation has been characteristic of the Arabic-speaking world until the present.

Classical Arabic as described by the Arab philologists is characterized by the following archaic features, distinguishing it from other variants of Arabic, especially from most of its pre-Islamic manifestations and the dialects of the neo-Arabic type:

i. The glottal stop is preserved in all positions: *raʾs-un* ‘head’, *saʾala* ‘he asked’, *ʾasʾil-at-un* ‘questions’;
ii. The existence of the phonemes *ḍ* (*ḍād*) in opposition to *ḏ̣* (*ḏ̣āʾ*), which in most of the vernaculars very early on merged into one single phoneme: *ḏ̣*;
iii. The definite article *al-* (with assimilation of *l* to following dentals);
iv. The suffix *-n* (nunation), applied to nouns of the triptotic declension, marking absence of the definite article or an annexed genitive;

v. Distinction of three cases of the noun, marked in the triptotic declension by *-u* (nominative), *-i* (genitive), *-a* (accusative), in the diptotic declension by *-u* (nominative), *-a* (genitive and accusative), in the plural declension by *-ū/u* (nominative) and *-ī/i* (genitive and accusative):

Table 1. Declension of the noun

	triptotic	diptotic	pl. masc.	pl. fem.
sg. nom.	*raʾs-u-n*	*ʿumar-u*	*ban-ūna* 'sons'	*ban-āt-u-n* 'daughters'
gen.	*raʾs-i-n*	*ʿumar-a*	*ban-ī-na*	*ban- āt-i-n*
acc.	*raʾs-a-n*	*ʿumar-a*	*ban-ī-na*	*ban-āt-i-n*

vi. Distinction of three moods in the prefix inflexion of the verb: marked by *-u* (imperfect indicative), by *-a* (subjunctive), by *-Ø* (short imperfect or apocopate):

Table 2. The moods of the verb

	indicative	subjunctive	apocopate
3. sg.	*yaktub-u* 'he writes'	*yaktub-a*	*yaktub*
3. pl.	*yaktub-ūna*	*yaktub-ū*	*yaktub-ū*

vii. The existence of an inner passive voice throughout the verbal system marked by the vowels *u – i* in the perfect and by *a* with *u*-prefixes in the imperfect:

Table 3. Active and passive voice

	active voice	passive voice
perf.	*ḍaraba*	'he struck'
	ḍuriba	'he was struck'
	bādala	'he exchanged'
	būdila	'he was exchanged'
	taʿallama	'he learned'
	tuʿullima	'it was learned'
imperf.	*yaḍribu*	'he strikes'
	yuḍrabu	'he is struck'
	yubādilu	'he exchanges'
	yubādalu	'he is exchanged'
	yataʿallamu	'he learns'
	yutaʿallamu	'it is learned'

viii. The distinction between verbal sentences (word order V-S) with non-agreement or only gender agreement between subject and predicate and nominal sentences (word order S-V) with full agreement between subject and predicate:

Table 4. Verbal sentence and nominal sentence

verbal sentence	nominal sentence
xaraja r-rijāl-u	*ar-rijāl-u xaraj-ū*
went-out the-men-nom	the-men-nom went-out-pl.m.
'the men went out'	'the men went out'
xaraja-(ti) n-nisāʾ-u	*an-nisāʾ-u xaraj-na*
went-out-(f.sg.) the-women-nom	the-women-nom went-out-pl.f.
'the women went out'	'the women went out'

What motivated the Arab grammarians to describe the *ʿarabiyya* was, on the one hand, the wish to protect the high linguistic level existing in the *Qurʾān*, in which it is called 'a clear Arabic language' (Q. 16/103: *hāḏā lisānun ʿarabiyyun mubīn*). Their aim was to prevent the spread of what they called → *laḥn* (Fück 1950:128ff.), which means mistakes in the use of the cases and moods arising from interference with a colloquial form of Arabic more or less corresponding to the modern Arabic dialects. On the other hand, they were faced with the task of teaching Arabic in an educated and standardized form to non-Arabs (Ibn as-Sarrāj, *ʾUṣūl* 35), in particular those who worked as officials (*kuttāb*) in the administration of the empire.

The consequence of this process was that the morphological system and the basic syntactic rules were frozen, so that Classical Arabic was taught according to these rules without any change during the following centuries down to the present time. However, in the course of time there emerged new expressions and syntactic features in areas not evidently regulated by the grammarians. With respect to those developments, one may discern several periods of Classical Arabic (Fischer 1972; → History of Arabic):

i. the period of pre-standardized Classical Arabic, in which it was not entirely regulated and standardized;

ii. the period of standardized Classical Arabic, in which it was used in accordance with the norms laid down by the Arab grammarians;
iii. the period of post-classical language, in which new expressions and syntactic structures appeared;
iv. → Modern Standard Arabic, the literary language of the Arab countries, which is influenced in its expressions and its vocabulary by European languages.

There are no precise chronological boundaries between these periods, especially between standardized Classical Arabic and the post-classical language. On the one hand, archaic forms and unusual syntactic constructions of the pre-classical language may recur in poetry of later generations (Ullmann 1984) and, on the other hand, at all times there were writers who cultivated a strictly classical style alongside others who did not hesitate to use recently developed expressions which had in the meantime become fashionable.

Western scholars use the term Classical Arabic with two slightly different meanings: some have in mind all kinds of Arabic that exhibit the morphological system of Classical Arabic, including Modern Standard Arabic, whereas others restrict the term to the language of the premodern literary language.

2. Pre-standardized Classical Arabic

During the first Islamic century the majority of Arabic poets and Arabic-writing persons spoke Arabic as their mother tongue. Their texts, although mainly preserved in far later manuscripts, contain traces of non-standardized elements in morphology and syntax. In particular the *Qurʾān* contains examples of such non-standardized forms, e.g. the assimilated forms in the V and VI verbal measures like *yaḏḏakkaru* (**yatḏakkaru* = *yataḏakkaru*), *iṯṯāqala* (**itṯāqala* = *taṯāqala*) (Ambros 1993), and the shortened form *ḏ̣altu* (= *ḏ̣aliltu*). The same phenomena are also witnessed in poetry, but with *ḏ̣iltu* instead of Qurʾānic *ḏ̣altu*. In some ancient *Qurʾān* manuscripts the pronouns of the 3rd person are found without assimilation to preceding *i/ī/ay*: *fī dāri-hū, fī dāri-hum* (equivalent to *fī dāri-hī, fī dāri-him* of the standard language), a feature which accords with what the Arab grammarians report about the dialect of the Ḥijāz. A well-known deviation from the standard is the relative pronoun *ḏū* (instead of Classical Arabic *al-laḏī, al-latī* etc,) preserved in poems of the tribe of Ṭayyiʾ. In their poems there are also examples of the elision of *i/u* in unstressed open syllables: *baqiya* > *baqā* 'he remained', *buniya* > *bunā* 'it was built' (Kofler 1941:75). In some poems, verbs with the causative prefix *ha-* (instead of *ʾa-*) have survived: *ha-rāqa* 'he spilt' (Kofler 1941:80). These and other morphological and syntactical phenomena of pre-standardized 'Old-Arabic' (Bloch 1946) have survived despite the process of standardization to which the texts were submitted in the course of being passed down through the ages.

The Arab grammarians recorded many other linguistic features existing in the pre-standardized language, but most of these were eliminated from the texts in the course of standardization. Among such dialectal variants is the so-called → *taltala* of the dialect of Tamīm, i.e. prefixes of the imperfect with *i*: *ni-ʿlamu* 'we know' instead of *na ʿlamu* in the Ḥijāzi dialect (Bloch 1967), a phenomenon which is also attested by some readings of the *Qurʾān* (*qirāʾāt*) and which has survived in one word of the Classical language: *ʾixālu* 'I imagine'. The Arab grammarians accepted such dialectal variants (*luġāt*) as authentic *ʿarabiyya* even if they rejected them during the process of standardization. The standardized form of Classical Arabic is obviously a selection which takes over features from different dialects (Rabin 1955; Corriente 1976:70–71): the vowel of the imperfect prefixes *ya-*, *ta-*, *ʾa-*, *na-* corresponds with the dialect of the Ḥijāz, the assimilation of the vowel of the pronoun of the 3rd person *-hū/-hī* and *-hum/-him* with the dialect of Tamīm. From the dialects of Najd comes the preservation of the glottal stop, whereas the Ḥijāzī dialect changed ʾ after vowels to *w*, *y* or *Ø*: *raʾsun, ruʾūsun, saʾala, suʾālun, suʾila* > Ḥijāzī *rāsun, rūsun, sāla, suwālun, suyila*. In general, standard Classical Arabic shows forms based on regular and unrestricted morphological rules, whereas irregular ones are avoided. For instance, the irregular *ḏ̣iltu, ḏ̣altu* has been re-placed by the regular *ḏ̣aliltu*; the forms of the V and VI verbal stems with assimilation of *-t-* to a dental first radical disappeared and only regular *tafaʿʿala* and *tafāʿala* remained.

How and at what time the process of unification leading to Classical Arabic took place is a matter of much debate. Most Western scholars assume that a relatively uniform intertribal → 'poetic koine' distinct from the spoken vernaculars developed in the late 6th century C.E. based on the Najdī Bedouin dialects, probably in connection with the court of al-Ḥīra (Rabin 1960:565a). They consider it most unlikely that the normal spoken Arabic of the Bedouin tribes at that time retained the archaic case and mood system and assume that these would have continued only in the poetic language. Corriente (1971:39) argues that Old Arabic did not possess the synthetic character often attributed to it and that the case and mood endings were rather redundant. Fück (1950:5) does not share these doubts and sees the unification of the *ʿarabiyya* as taking place in the camps of the Islamic armies, where warriors of different tribes lived together after the Islamic expansion in the second half of the 7th century C.E. In contrast, the Arab grammarians of the 8th and 9th centuries report that their Bedouin informants did not provide them with information only about the poetic language, but also spoke an Arabic dialect akin to Classical Arabic, which could be classed among the correct *ʿarabiyya*. Muslim tradition holds that Classical Arabic is identical to the *ʿarabiyya* spoken by the people of Mecca, the Qurayš. In all events, what must not be forgotten is that the dialectal variants of Classical Arabic are not only reported by the grammarians, but are also attested to by Qurʾānic readings, which clearly date from the 1st and 2nd Islamic centuries. This is a strong argument in favor of the Arabic tradition which puts the process of unification of Classical Arabic into the first centuries of the Islamic period.

The spelling of Classical Arabic is not a direct projection of the language described by the grammarians (Fischer 1992), deviating as it does from the linguistic shape of Classical Arabic in many aspects: on the whole, it reflects the Ḥijāzi change of ʾ to *w, y* with spellings like <swʾl> [suwāl] = *suʾāl-un*, <syl> [suyila] = *suʾila*; *ʾalif* <ʾ> is equivalent to the glottal stop [ʾ] only at the beginning of the word, but within and at the end of the word it stands for *ā*: <ʾn> = *ʾan* or *ʾin*, but <qʾl> = *qāla* 'he said' and <yqrʾ> [yaqrā] = *yaqraʾu* 'he reads'. At the end of the word the glottal stop /ʾ/ is not expressed by any letter after consonants and long vowels: <ʾl-šy> = *aš-šayʾ-u/i/a* 'the thing [nom./ gen./acc.]', <jzʾ> (the *ʾalif* stands for *ā*) = *jazāʾ-un/in* 'requital [nom./gen.]', <bry> = *barīʾ-un/in* 'innocent [nom./gen.]'. The case and mood markers at the end of the word do not usually appear in writing; as seen in the last example, the final *-n* (nunation) is not indicated in the spelling, only in the accusative does *ʾalif* <ʾ> indicate the ending *-an*: e.g. <ktʾb> stands for *kitābun/kitābin* 'a book [nom./gen.]', but <ktʾbʾ> for *kitāb-an* 'a book [acc.]'. Only in one word is the *-n* of the nunation written: <kʾyn> = *ka-ʾayyin* 'like which' (with the meaning 'how many'). The feminine marker *-at-* is in the *Qurʾān*, sometimes written with <t>, but usually with <h>, if it comes in final word position: <snh> represents *san-at-u /-i /-a* and *san-at-un /-in /-an* 'year'. The writing of the initial ʾ of the definite article *al-* and of some anaptyctic vowels is inconsistent with the traditional pronunciation of Classical Arabic: <byt ʾl-ḥsn> = *bayt-u l-ḥasan-i* 'the house of al-Ḥasan', <qʾl ʾbny> = *qāla bnī* 'my son said'. Only *bi-smi* in the formula *bi-smi llāhi* 'in the name of God' is written in accordance with its articulation without <ʾ> as <bsm llh>. In order to adapt these traditional spellings to the standardized form of Classical Arabic, the Arab philologists developed in addition to the vowel signs some supplementary signs: *hamza* to indicate ʾ, *hamzat al-waṣl* to signify that initial *ʾalif* <ʾ> is not to be pronounced, and the putting of two dots over <h> indicating that it stands for the feminine marker *-at-*. Muslim tradition credits ʾAbū l-ʾAswad ad-Duʾalī (d. 69/688) with the first efforts to codify and standardize the language. He is said to be the inventor of the first vowel signs consisting of a dot above the letter for *a*, a dot after the letter for *i* and a dot beneath the letter for *u*, using two dots to mark the nunation. About one hundred years later al-Xalīl ibn ʾAḥmad al-Farāhīdī (d. 175/791) established a new system of vowel signs. It is his system which has remained in use until the present day.

It is obvious that the spelling of ʾ by *wāw* <w> or *yāʾ* <y> is akin to what the grammarians report about the dialect of the Ḥijāz. Nöldeke (1904) explains the other peculiarities by arguing that every word is written down as if it stands alone, i.e. in its pausal form where the final short vowels and the *-n* of the nunation are dropped (Birkeland 1940). Many Western scholars have seen in these spellings reflections of contemporary colloquial Arabic which, in their opinion, even in pre-Islamic times had features akin to the modern Arabic dialects. Vollers (1906) went so far as to assume that the text of the *Qurʾān* was originally written down in the Meccan vernacular, which would explain the lack of case and mood endings. Afterwards the text was – he believed – adjusted to the standardized Classical Arabic. Recent research points out that "the only thing that can be said with any certainty is that the Qurʾānic orthography continues the orthographic conventions of the Aramaic/Nabataean script which were also used in the pre-Islamic Arabic inscriptions" (Versteegh 1997:47). As in the Aramaic orthography *ā* is very often not indicated: <ʾl-ktb> = *al-kitābu*/*al-kitābi*/*al-kitāba* (Diem 1979:242–256). Some Aramaic loanwords retain their original spelling <w>, where Classical Arabic has *ā*, such as <ṣlwh> = *ṣalāt-un* 'prayer' from Aramaic *ṣlōṯ-ā*, <zkwh> = *zakāt-un* 'alms tax' from Aramaic *zkōṯ-ā* (Spitaler 1960). With respect to these spellings the Arab grammarians report the pronunciation with *ō* (*ṣalōt-un*, *zakōt-un*), which might well be a trace of the Aramaic origin of these religious terms. Arabic words occurring in Nabataean inscriptions are sometimes written with a final <w> hinting at the Arabic case ending *-u(n)* (e.g. <qbrw> = *qabr-u(n)* 'grave'). The same is witnessed in Arabic proper names in Palmyrene and other Aramaic inscriptions, where final <y> is also found in compound names indicating the genitive *-i(n)*, (e.g. <ʿbdʾlhy> = *ʿabdullāhi* (Diem 1981: 336–342). These reflexes of the Arabic case system appear in a fossilized form and do not correspond with the specific syntactic requirements. This suggests that those varieties of Arabic which were spoken in the border areas in contact with the Aramaic-speaking population lost their inflectional system very early. It is noteworthy that the orthography of Classical Arabic has kept this fossilized <w> in the spelling of one proper name, <ʿmrw> = *ʿamr-un*, in order to differentiate it from <ʿmr> = *ʿumar-u*. One unsolved problem is the origin and function of what is called in Arabic *ʾalif maqṣūra* 'shortened alif', i.e. the spelling with <y> in cases where it is equivalent to Classical Arabic *-ā* at the end of the word (e.g. <nry> = *narā* 'we see', also <nry-km> = *narāka* 'we see you'), whether it denotes *ē* as an allophone of *ā* (Bergsträsser 1961:36ff.; Hopkins 1984:8), or is merely an orthographic convention.

In fact, one must be very cautious in assuming from these spellings too much information about the linguistic shape of Classical Arabic. However, it is clear that the Qurʾānic orthography was not developed for the standardized form of Classical Arabic; rather, it shows the attempt on the part of writers to utilize a traditional writing system for recording a non-standardized form of Classical Arabic. The case and mood endings are indicated in accordance with their syntactic functions whenever the spelling makes this possible. The case endings are indicated in the masculine sound plural *-ūna*/*-īna* (written <-wn> resp. <-yn>), in the dual *-āni*/*-ayni* (written <-ʾn/-yn>, e.g. <jntʾn> = *jannatāni* 'two gardens [nom.]' vs. <jntyn> = *jannatayni* 'two gardens [gen./acc.]'), in some words with long vowels as case markers (e.g. <ʾbw-km> = *ʾabū-kum* 'your father [nom.]', <ʾby-km> = *ʾabī-kum* 'your father [gen.]', <ʾbʾ-km> = *ʾabā-kum* 'your father [acc.]' and in spellings like <ʾbnʾw-km>, <ʾbnʾy-km>, <ʾbnʾ-km> = *ʾabnāʾu-kum*, *ʾabnāʾi-kum*, *ʾabnāʾa-kum* 'your sons [nom., gen., acc.]'). The imperfect is clearly differentiated from the apocopate in the spelling of the verbs II*wly*, such as <tqwl> = *taqūlu* 'thou sayest' vs. <lʾ tql> = *lā taqul* 'do not say', and subjunctive and apocopate differ from the imperfect in the plural and dual forms.

3. Standardized Classical Arabic

In accordance with the importance of the *Qurʾān*, Classical Arabic became during the first century of the caliphate the official and educated language of the Islamic empire, starting with the order of the Caliph ʿAbd al-Malik (685–705) to shift the administration of the Umayyad Empire from Greek and Middle-Persian (Pahlavi) to Arabic. A number of official texts of that time, such as the inscription of the Dome of the Rock dating from 691 C.E. and the correspondence of the governor of Egypt, Qurra ibn Šarīk, dating from 709–714 C.E. (Abbott 1938), illustrate that the Qurʾānic orthography remained the basis of Arabic spelling during the first Islamic centuries (Hopkins 1984:1–61), and that the administration attempted to preserve the linguistic standard of the language as found in the *Qurʾān* and the poetic tradition. However, the very small number of extant textual documents and the problems of their orthography make it nearly impossible to reconstruct in any definite way the history of Classical Arabic for the first two Islamic centuries.

In the 2nd/8th century, Arabic-Islamic society passed through a transformation which deeply influenced the role of the Arabic language. Until the middle of the 8th century those writing Arabic were usually of Arab origin. The poetic tradition of the Arab tribes which had undergone a revival during the reign of the Umayyad caliphs ended at that time because the tribal elites adopted an urban way of life and lost their connection with the Bedouin traditional language culture. Poets like D̠ū r-Rumma (d. 117/735) and Ibn Mayyāda (d.136/754) are reported to have been the last heirs to the poetic tradition of the Arab tribes.

Yet, the status of Classical Arabic as the language of religious rites, administration, science, and literature remained untouched. Thus, the knowledge of Classical Arabic became a matter of education adopted more and more by the rising class of non-Arab Muslims (*mawālī*), who began to enter into competition in prose and poetry with the Arabs, a trend which increased when the Abbasid dynasty took over power. The appearance of persons like the prose writer and translator from Pahlavi Ibn al-Muqaffaʿ (d. 142/759) and the poet Baššār ibn Burd (d. 167/784), both of Persian origin, are characteristic of this trend.

It seems that the standardization of Classical Arabic had reached completion around the end of the 8th century. The first comprehensive description of the *ʿarabiyya*, called *al-Kitāb* 'the Book', written by the grammarian Sībawayhi (d. 177/793), is a landmark, for it contains the fruits of the linguistic thinking of the Arabs through one hundred years. Sībawayhi's *Kitāb* is based first of all upon a corpus of poetic texts, in addition to the *Qurʾān* and Bedouin informants whom he considers to be *fuṣaḥāʾ* (sg. *faṣīḥ*), that is reliable speakers of the *ʿarabiyya*. As a matter of principle, he accepts all phenomena attested in these sources and presents not only a grammatical description on a high linguistic level, but also considerations about what is to be chosen as the best stylistic usage. In this way he established the norms of Classical Arabic and grammar, which became after him a normative description of the language. Moreover, he tried to give systematic explanations for every linguistic phenomenon. Only rarely did his successors enrich their grammatical treatises with new substantial information. However, they did improve the methods of description and argumentation, so that they reached a very high level of scholarship in developing a linguistic theory (→ Grammatical tradition).

During the 2nd/8th century, the correct knowledge of Classical Arabic had become an essential prerequisite for rising into the higher classes throughout the Islamic world. The high prestige that Classical Arabic had gained as the language of religious rites, administration, and poetry ensured the philologists and grammarians a central position in society. The philological scholars of that time shaped the literary education in an enduring way. Leading philologists like al-Kisāʾī (d. 189/804) or al-Farrāʾ (d. 207/822) were called to the Abbasid court in Baghdad in order to educate the princes. From the second half of the 8th century onwards, philological scholars had begun to gather the legacy of the Arab tribal culture, their poetry, and their tales about famous conflicts between the Bedouin tribes (*ʾayyām al-ʿarab*). This heritage became the basis of the secular literary culture.

At the same time lexicographers began to codify the vocabulary of the tribal linguistic tradition, which had sunk into oblivion as a result of the transition to urban life. Even some Qurʾānic words had become unfamiliar and discussions about their meanings arose among the experts of *Qurʾān* interpretation. At first, scholars like al-ʾAṣmaʿī (d. 213/828) classified the vocabulary according to subject, but after al-Xalīl ibn ʾAḥmad (d. 175/791) had invented a method of analyzing Arabic words on the basis of their root consonants, the way was open for attempts to arrange the vocabulary in an alphabetical order (→ Lexicography, Classical). The lexicographers did not confine themselves to codifying the vocabulary, but just as they wanted to safeguard the language from corruption in its grammatical structure, they tried to stop the uncontrolled expansion of the vocabulary and to bring to a close the infiltration of foreign words. The integration of the cultural achievements of the civilizations now placed under Muslim government required a great number of new terms in Arabic. The simplest way of acquiring these was to borrow the foreign notation with its meaning. In fact, Classical Arabic has at all times received loanwords from other languages and some of them gained permanent acceptance provided they could be adapted to the phonemic and morphological rules of Arabic. In such cases, the foreign origin was forgotten or the words were considered 'Arabized' (*muʿarrab*). The main sources of loanwords in pre-Islamic and early Islamic times were → Aramaic and → Persian. As a result of the systematic translation of works of the Greek sciences (sometimes by way of a Syrian translation), for which the Caliph al-Maʾmūn (813–833 A.D.) founded an academy called *Bayt al-ḥikma* 'House of wisdom', the number of words of Greek origin increased rapidly (→ Greek loanwords). In principle, the translators tried to find Arabic equivalents for the Greek terms, but when they did not succeed at the first attempt, there was a tendency to replace words of foreign origin by Arabic neologisms. Thus the Greek loanword *hayūlā* 'substance' (< Greek ὕλη) was replaced by *mādda, ʾusṭuquss* (< Greek στοιχεῖον) 'element' by *ʿunṣur*, *fanṭāsiyā* 'fantasy' (< Greek φαντασία) by *taxayyul* (Endress 1992:12–23). Several loanwords from Greek like *mūsīqī/ mūsīqā* 'music' (< Greek μουσική), *faylasūf* 'philosopher' (< Greek φιλόσοφος), which formed the basis for the Arabic neologism *falsafa* 'philosophy', *dīmuqrāṭiyya* 'democracy' (< Greek δημοκρατία) as well as many neologisms formed by the translators like *naḏariyya* 'theory', *huwiyya* 'essence, identity', *kiyān* 'substance, being', and others have become a permanent part of the Arabic lexicon. After a period of experimentation Classical Arabic finally became a language able to express all the scientific ideas of the time by its own means.

During the 3rd/9th century, Classical Arabic developed into a universal literary and scientific language used throughout the entire Islamic world. It became the sole language of culture and science employed by Muslim scholars and writers from → al-Andalus to Middle Asia. This development was the joint work of the leading cultural elites: first, the philological scholars and the scholars of the Islamic religious sciences, who established the Islamic law (*šarīʿa*) during the 9th century all over the Islamic world; and second, the government secretaries, who created a high style in writing official documents and were considered the main representatives of the secular literary culture (*ʾadab*) as well as the translators of the Greek scientific works into Arabic. Mastery of Classical Arabic became the highest ideal in education.

As a consequence of the rapid growth of the culture of writing in the Islamic world, the orthography of Classical Arabic was revised in the course of the 9th century C.E. The main change was the constant spelling of *ā* in the middle of the word by *ʾalif* <ʾ>; in some words however, the archaic spelling remained untouched (Fischer 2002:§8). From the 4th/10th century onward this spelling is used also in *Qurʾān* manuscripts.

Classical Arabic became so dominant that even the non-Muslim communities accepted it as the language of education. Around the end of the 8th century, the Melkites began to write in Arabic, the Syrians and the Jews accepted Classical Arabic in the 9th and the Copts in the 10th century. However, they did not participate in the linguistic education of the Muslims; hence, influenced by the Arabic vernaculars, their writings are not always free of morphological phenomena deviating from the classical rules and show a linguistic variety which is called → Middle Arabic. (See also → Christian Middle Arabic, → Judaeo-Arabic)

4. Post-Classical Arabic

In the traditional view of the educated class in the Arab countries, the language of writers from the second half of the 2nd/8th until the 5th/11th or 6th/12th centuries marks the zenith of the Classical Arabic prose style. Authors such as al-Jāḥiẓ (d. 255/868), Badīʿ az-Zamān al-Hamaḏānī (d. 398/1008) or ʾAbū Ḥayyān at-Tawḥīdī (d. 414/1023), are regarded as the exponents of the best Classical Arabic style. What came later, especially after the destruction of Baghdad by the Mongols in 1258 C.E., is considered a period of decline in literary culture.

With the political disintegration of the caliphate, the prestige of Classical Arabic as the sole language of the educated and cultured in the Islamic world began to decrease. Around the end of the 4th/10th century its position was affected by the revival of → Persian as a language of literature under the rule of the Samanids in Transoxania (819–1025 C.E.) and the Ghaznawids in Eastern Iran (977–1187 C.E.), who encouraged the emergence of poetry and historiography in New Persian. After the Mongol invasion of Iran in 654/1256, the Ilkhan rulers made Persian the official language of the government administration and adopted the literary Persian culture. In this context, it is remarkable that the influence of Classical Arabic literature was still so decisive that its prosodic system was taken over by the Persian poets. Later on, in the → Ottoman Empire, Turkish became the language of administration and popular literature. From the 7th/13th century onward, Classical Arabic lost its unique position as the sole language of education in the Islamic world. As a language of literature, it became limited to those countries in which Arabic was also the spoken language, but it maintained its status as the language of the religious sciences all over the Islamic world. It continued to serve as the language of learning, being taught in all colleges of the Islamic religious sciences. However, in all the languages of education which emerged in the Islamic world, from → Hausa in West Africa to Bahasa Indonesia (→ Indonesian/Malay) Classical Arabic has had a major impact, being for these languages what Latin and Greek were for the European languages, providing them with a profusion of scientific terminology and other loanwords.

Although the norms laid down by the grammarians have never been challenged, the linguistic development could not be completely stopped and new expressions and syntactic patterns emerged. Such innovations succeeded wherever the language had not been regulated by mnemonic rules. The writers did not hesitate to use a vocabulary taken from the regional vernaculars (Fück 1950:108–114). Adverbs like *ʾayḍan* 'also', *xāṣṣatan* 'especially', and 'adjectives of relation' (*nisba*) with the ending *-ānī*, like *jismānī* 'bodily', *rūḥānī* 'spiritual' are characteristic innovations of this period. Examples of new syntactic constructions are: *kawn* for subordination of nominal clauses (Diem 1995), subordination by conjunctions like *ʿinda-mā* 'when', *fī-mā* 'while', *bi-mujarradi-mā* 'as soon as', and others. Research in the field of post-Classical Arabic grammar is very limited, so that is impossible to give exact information on its history in this period.

5. Modern Standard Arabic

The modern period has seen Classical Arabic becoming not only the language of the educated classes in Arab countries but also, and above all, the written official language in those countries, a role it shares with Hebrew in → Israel. Modern Standard Arabic is the direct continuation and modern version of Classical Arabic. The literary revival movement of the 19th century (*nahḍa*) aimed at raising linguistic standards by continuing the models of the classical period, but on the whole, Modern Standard Arabic follows the features of the post-Classical language as Blau (1973) has shown (→ Modern Standard Arabic).

Bibliographical references

Primary source

Ibn as-Sarrāj, *ʾUṣūl* = ʾAbū Bakr Muḥammad ibn Sahl Ibn as-Sarrāj, *al-ʾUṣūl fī n-naḥw*. Ed. ʿAbd al-Ḥusayn al-Fatlī. Baghdad, 1987.

Secondary sources

Abbott, Nabia. 1938. *The rise of the North Arabic script and its Kurʾānic development with a full description of the Kurʾān manuscripts in the Oriental Institute*. Chicago: Oriental Institute.

Ambros, Arne A. 1993. "Haplologie und Assimilation im V. und VI. Verbstamm im Koran". *Zeitschrift für Arabische Linguistik* 25.1–16.

Bergsträsser, Gotthelf and Otto Pretzl. 1961. *Geschichte des Qorāns von Theodor Nöldeke*. III. *Die Geschichte des Korantexts*. 2nd ed. Hildesheim: G. Olms.

Birkeland, Harris. 1940. *Altarabische Pausalformen*. Oslo: Norske Videnskaps-Akademi.

Blau, Joshua. 1973. "Remarks on some syntactic trends in Modern Standard Arabic". *Israel Oriental Studies* 3.172–231.

Bloch, Alfred. 1946. *Vers und Sprache im Altarabischen: Metrische und syntaktische Untersuchungen*. Basel: Verlag für Recht und Gesellschaft.

Bloch, Ariel. 1967. "The vowels of the imperfect preformatives in the old dialects of Arabic". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 117.22–19.

Corriente, Federico C. 1971. "On the functional yield of some synthetic devices in Arabic and Semitic morphology". *Jewish Quarterly Review* 62.20–50.

——. 1976. "From Old Arabic to Classical Arabic through the Pre-Islamic koine: Notes on the native grammarians' sources, attitudes and goals". *Journal of Semitic Studies* 21.62–98.

Diem, Werner. 1979. "Untersuchungen zur frühen Geschichte der arabischen Orthographie. I. Die Schreibung der Vokale". *Orientalia* 48.207–257.

——. 1980. "Untersuchungen zur frühen Geschichte der arabischen Orthographie. II. Die Schreibung der Konsonanten". *Orientalia* 49.67–106.

——. 1981. "Untersuchungen zur frühen Geschichte der arabischen Orthographie. III. Endungen und Endschreibungen". *Orientalia* 50.332–383.

——. 1983. "Untersuchungen zur frühen Geschichte der arabischen Orthographie. IV. Die Schreibung der zusammenhängenden Rede. Zusammenfassung". *Orientalia* 52.357–404.

——. 1995. "*kaunahu rasūlan* 'weil er Bote ist' und Verwandtes: Ein Beitrag zur Syntax des nachklassischen Arabisch". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 145.47–108.

Endress, Gerhard. 1992. "Die wissenschaftliche Literatur". *Grundriss der arabischen Philologie*. III.

Supplement, ed. Wolfdietrich Fischer, 1–152. Wiesbaden: L. Reichert.

Fischer, Wolfdietrich. 1972. "Die Perioden des Klassischen Arabisch". *Abr-Nahrain* 12.15–18.

——. 1992. "Orthographie in ihrem Verhältnis zu Phonologie und Morphologie im Klassischen Arazbisch". *Jerusalem Studies in Arabic and Islam* 15.134–144.

——. 2002. *A grammar of Classical Arabic*. New Haven and London: Yale University Press.

Fück, Johann. 1950. *Arabiya: Untersuchungen zur arabischen Sprach- und Stilgeschichte*. Berlin: Akademie-Verlag.

Hopkins, Simon. 1984. *Studies in the grammar of early Arabic based upon papyri datable to before A.H. 300/A.D. 912*. London: Oxford University Press.

Kofler, Hans 1940–1942. "Reste altarabischer Dialekte". *Wiener Zeitschrift für die Kunde des Morgenlandes* 47.61–130, 232–262; 48.52–88, 247–274; 49.15–30, 234–256.

Nöldeke, Theodor. 1904. "Das klassische Arabisch und die arabischen Dialekte". Theodor Nöldeke, *Beiträge zur semitischen Sprachwissenschaft*, 1–14. Strassburg: K.J. Trübner.

Rabin, Chaim. 1955. "The beginnings of Classical Arabic". *Studia Islamica* 4.19–37.

——. 1960. "ʿArabiyya. Arabic language and literatur. I. Pre-Classical Arabic. II (1) Classical Arabic". *Encyclopaedia of Islam*, 2nd ed., I, 564–567. Leiden: E.J. Brill.

Spitaler, Anton. 1960. "Die Schreibung des Typus صلوة im Koran". *Wiener Zeitschrift für die Kunde des Morgenlandes* 56.212–226.

Ullmann, Manfred. 1984. "Vorklassisches Arabisch". *Studi in onore di Francesco Gabrieli*, I, ed. Renato Traini, 807–818. Rome: University of Rome.

Versteegh, Kees. 1997. *The Arabic language*. Edinburgh: Edinburgh University Press. (2nd. ed. 2001.)

Vollers, Karl. 1906. *Volkssprache und Schriftsprache im alten Arabien*. Strassburg: K.J. Trübner.

WOLFDIETRICH FISCHER (Erlangen, Germany)

Classicism

1. INTRODUCTION

The term *classicism* refers here to the incorporation of an item of Classical Arabic into colloquial-based discourse. The term → Classical Arabic is used here as the equivalent of *al-ʿarabiyya al-fuṣḥā* 'the pure classical Arabic language'. Following the sociolinguistically based stylistic classification of Badawī (1973:90–93), it can be divided into *fuṣḥā at-turāṯ* 'the pure Classical Arabic of the legacy', i.e., traditional Classical Arabic, as defined by medieval grammarians, and *fuṣḥā al-ʿaṣr* 'contemporary Classical Arabic', commonly called → Modern Standard Arabic. When the language situation of Arabic is addressed within the theoretical framework of → diglossia, traditional Classical Arabic and Modern Standard Arabic represent the high variety. The low variety consists of different levels of Colloquial Arabic, called by Badawī *ʿāmmiyyat al-ʾummiyyīn* 'illiterate colloquial', *ʿāmmiyyat al-mutanawwirīn* 'enlightened colloquial', and *ʿāmmiyyat al-muṯaqqafīn* 'educated colloquial'. Between these levels no clear-cut divisions are found, but they constitute a linguistic continuum. Every level has, however, its specific combination of linguistic and social characteristics. The most flexible and variegated of the levels is 'educated colloquial', a kind of bridge between the two structurally different linguistic types, and, consequently, also the variety of Colloquial Arabic in which classicisms most frequently appear.

The study of classicisms is restricted to colloquial-based discourse, in which they appear as interferences from Classical Arabic. These may be phonological, grammatical, lexical, or phraseological devices, and they often consist of more than one item. Using Classical elements as a rule reflects a deliberate stylistic shift toward a higher variety. This is not always the case, however. Many Modern Standard Arabic lexical items, for example *mudīr* 'director' and *masʾūl* 'responsible', have been integrated into all stylistic levels of colloquial without assimilating them into the structure of the inherited colloquial, whereas some other items are adapted into it; e.g. in dialects in which the older interdentals have become postdentals, *maṯalan* 'for instance' becomes *matalan* or *masalan*. There are numerous frequently used, frozen phrases which, in spite of their prominent Classical Arabic elements, cannot properly be considered as devices of stylistic modification. Some of them are relatively recent, such as *šukran – lā šukra/šukran ʿalā wājib*, while some are traditional, e.g. *as-salāmu ʿalaykum, ʾaʿūḏu/ʾaʿūzu billāh, aḷḷāhu ʾakbar, waḷḷāhi, al-ḥamdu lillāh, ʾin šāʾ aḷḷāh*, *ʾahlan wa sahlan*. When used in colloquial discourse, all of these have different modifications, yet preserve many strictly Classical Arabic morphemes. For the adaptation of these traditional, mostly religious phrases, no regular pattern seems to exist. Some of them have become shortened, e.g. *waḷḷa*, *ilḥamdillāh*, *inšaḷḷa*, but some have also been completely restructured according to the structure of the colloquial environment. Thus, *ʾaḷḷa yirḥamu* 'may God have mercy upon him', the equivalent of the Classical Arabic *raḥimahu ḷḷāh*, in every respect follows the colloquial structure. Consequently, the use of the Classical Arabic variant of the phrase is in most cases a marked classicism. The same holds true for practically all phrases in which the optative mood is expressed by the perfect in Classical Arabic and by the imperfect in Colloquial Arabic. Exceptions are mainly found in Bedouin dialects, in which the perfect is often used side by side with the imperfect: *gawwak/aḷḷa ygawwīk* 'may God give you strength!', *guwīt* 'may He give strength to you, too' (Banī Ṣaxar, Jordan).

2. Classicization and code-switching

In some cases, especially when academic, political, religious, or other serious topics are treated, or when the situation is intrinsically formal, most speakers tend to classicize their dialect to signal this, using non-dialectal forms or, in the most extreme case, totally replacing the dialect by Modern Standard Arabic (Holes 1995: 294–295). In the latter case, we can speak of → code-switching, "a pattern of textual production in which a speaker alternates between continuous utterance segments in one language L_X and another language L_Y with abrupt and clear-cut switching points, often at phrasal or clausal boundaries" (Heath 1989:23). In Colloquial Arabic discourse, switching to spoken Modern Standard Arabic is often caused by the wish to emphasize a certain idea or proposition, or to foreground information. It may also be indicated by the choice of particular lexical or grammatical items. The two codes are often mixed in one sentence, even at word level. The following examples recorded in Cairo are characteristic of such intrasentential mixing: *ʾayy muḥāfiẓ ʿandína* yurīd ʾan yakūn *ʿindu gamʿa* 'any governor in our country wants to have a university'; *ʾilli nišufha* wa ʾallati naʿīšuha 'which we see and experience' (Mejdell 1999:232–234). Correspondingly, the return to everyday topics may call forth switching back to a basically colloquial language. A typical case for this kind of code-switching is a spontaneous parenthetical comment, injected in a monologue held in Modern Standard Arabic.

As pointed out by Mazraani in her study of language variation in Arabic political speechmaking (1997:213), code-switching is an effective strategy in long speeches. Having finished appealing to the emotions of his audience and established solidarity with them through the dialect, the speaker switches to an approximation of Modern Standard Arabic to re-establish his authority and regain formality. In this respect, the standards are different in different speech communities. In the speeches of the Egyptian, Iraqi, and Libyan leaders analyzed by Mazraani, code-switching appears to be a serviceable rhetorical tool, whereas in other countries, e.g. in Syria and Lebanon, only Modern Standard Arabic is as a rule used in formal monologue (Diem 1974:62–67).

In oral discourse, the choice of the language level does not depend on the register (formal vs. informal) alone, but perhaps even more on the way a topic is dealt with. When an issue is treated in an impersonal way and in abstract terms, stigmatized colloquialisms are suppressed and the basically dialectal structure becomes mixed with certain – mainly phonological and lexical – *fuṣḥā* elements, but as soon as the topic is given a more concrete and personal slant, the style shifts toward the dialect. Holes (1995:287–295) illuminates the complex relationship between the form and the function of a discourse with strategies used by Nasser in his monologues and those used by two Jordanians in a discussion on different social and educational topics, and interestingly finds that the speakers in these substantially different contexts actually used similar means to mark off what they considered 'text' from 'commentary on text' for their audiences. As to the use of classicisms, Holes concludes that it is "more often than not a reflection of the treatment of the topic *per se*" (1995:294).

3. Classicization vs. koineization

Classicisms occur most frequently in the more elevated varieties of Colloquial Arabic in which stigmatized dialectal forms as well as many local features are suppressed, i.e. in koineized Colloquial Arabic. As pointed out for the first time by Blanc (1964:81–85), although classicizing and koineizing tendencies often appear in parallel in interdialectal conversation, they should be kept carefully apart. Koineization is a leveling process which takes place within the sphere of the colloquial language. It leads toward local or regional dialect standards, commonly those of the capital cities, which tend to carry more prestige than other dialects spoken in a country. Through modern media, mainly radio, television, and motion pictures, the linguistic centers spread their influence over national boundaries and evoke another type of dialect leveling. In cross-dialectal conversation between speakers from different Arab countries, many features are suppressed which in one country may be prestige forms but which are felt to be less known elsewhere. Since Cairene is the best-known dialect, its speakers are less inclined to shift away from Egyptianisms (Holes 1995:294). In cross-dialectal conversations the speakers resort to a number of accommodation strategies involving code-switching to another dialect, to Modern Standard Arabic, and to English or French as well (Abu-Melhim 1991:248–249). Classicization, on the other hand, is a feature that tends to narrow the gap between Colloquial Arabic and the oral variety of Modern Standard Arabic, sometimes to such a degree that it may be a very delicate matter to decide whether the linguistic structure of a certain passage should be regarded as Colloquial- or Classical-based. This development of Colloquial Arabic toward Modern Standard Arabic is in a sense counteracted by koineization, because the development of prestigious regional standard dialects undoubtedly strengthens the position of Colloquial Arabic and widens its functional range (Diem 1978:133, 144; Holes 1980:81; Palva 1982:31–32; Miller 2003:179–181).

4. Lexical borrowing and its morphophonological implications

The term ‘borrowing’ refers to the adaptation of a lexical item from one language into another and integrating it into the morphophonological structure of this language (Heath 1989:23). Many borrowed lexical items are integrated into their colloquial environments without substantial changes, e.g. *ʿadad* ‘number [quantity]’, *ʿalam* ‘flag’, *madani* ‘civil’, and *waṭani* ‘national’, whereas many other trigger phonological adaptations. In the urban dialects of Egypt and Greater Syria, in which the old interdentals have become equivalent postdental stops, the interdentals occurring in classicisms are represented by sibilant substitutes. Thus, e.g., in Damascus, *ʾāṯār* ‘relics, antiquities’ becomes *(ʾ)āsār*; *ʾiḏāʿa* ‘broadcasting, radio’ preserves its /i/, which in its dialectal context is phonologically long, and /z/ is substituted for /ḏ/: *(ʾ)izāʿa* /i:za:ʾa/. In the same way, the short /u/ and /i/ in unstressed open syllables of the borrowed words are preserved and become phonologically long: *muraššaḥ* ‘candidate, nominee’ becomes /mu:raššaḥ/, and *bināya* ‘building’ /bi:na:ye/.

Correspondingly, in Morocco the *ž* reflex of **j* assimilates the /l/ of the definite article: *žžml* ‘the camel’, but when it appears as the initial consonant of a borrowed item, the article is not assimilated: *lžumhur* ‘the crowd, audience’, *lžawāʾiz* ‘prizes’. According to Heath (1989:53), this is predictable because the latter items are still felt to be cultivated, classical in nature. Here, too, the /u/ of the /mu-/ morpheme of borrowed participles is preserved, and the reduction of vocalism affects the vowel of the next syllable: *mufattiš* > *mufttiš* ‘inspector’, *mumaṯṯil* > *mumttil* ‘actor’. The prefix *mu-* is now established and any more recent borrowings of this type will adopt it (Heath 1989:63).

In borrowings from Modern Standard Arabic, the inherited glottal stop, which as a rule has disappeared from Colloquial Arabic as an independent phoneme, is restored. Thus, *lājiʾ* ‘refugee’ retains it: *lāžeʾ*, pl. *lāžʾīn*; and in the borrowed stem *s-ʾ-l* ‘to ask’ the glottal stop is in frequent use: *saʾal, yisʾal, suʾāl* /su:ʾa:l/ (Grotzfeld 1965:65). In sedentary dialects spoken in Greater Syria, this verb has almost completely replaced its old dialectal equivalents, e.g. *sāyal*. Consequently, it is no longer felt as a borrowing, as is apparent from its dialectal inflection. Another case in which the glottal stop tends to be introduced in lexical borrowings from Modern Standard Arabic is actually a → hypercorrection or pseudo-classicization, namely, the pronunciation of the infinitives of verbal Forms VII, VIII and X with an initial glottal stop, e.g. *l-ʾiqtiṣād* instead of the Classical *al-iqtiṣād* (Heath 1989:47).

In dialects in which the glottal stop is the reflex of **q*, and /ʾ/ therefore is an independent phoneme, the reintroduction of the glottal stop of Classical Arabic in classicisms may bring about morphological complications. Since the dialect of Damascus, like the majority of the modern sedentary dialects, has lost the verbal pattern *ʾafʿal* as a productive morphological category, many speakers interpret the glottal stop in the borrowed *ʾafʿal* patterns as the first radical of a quadriradical root, and inflect them accordingly: *ʾaṣbaḥ/yʾaṣbeḥ* ‘to become’, *ʾazhar/yʾazher* ‘to flourish’ (Grotzfeld 1964:85).

One of the most important grammatical differences between the *fuṣḥā* and Colloquial Arabic is the lack of internal (apophonic) passives as a productive category in the latter. In the elevated varieties of Colloquial Arabic, some borrowings from this category tend to occur. Perhaps the most frequent of these is the passive imperfect *yuqāl* ‘it is said’, which does not introduce a productive use of this pattern, but can most properly be regarded as a lexical item with passive meaning. Used in colloquial discourse, it is a markedly Classical stylistic device in most dialects of Arabic, not integrated into the morphophonological structure of the dialect. In Cairene, however, it has become common enough to be classified as a lexical borrowing: it may be not only morphologically adapted to the dialect by prefixing the *b(i)*-morpheme (*byuqāl*), but even phonologically by substituting /ʾ/ for /q/ (*byuʾāl*).

The difference between the dialect of Cairo and the majority of dialects in their ability to adapt classicisms is based upon the prestige of the dialects. One way to objectively assess the prestige of a dialect is to examine how high the levels of formality certain dialectal items may pervade are. In Cairo, colloquial morphemes such as the *b(i)*- and *ḥa*- preverbs, as well as the 2nd and 3rd pers. pl. morpheme *-ū* in the imperfect, can relatively freely be combined with Modern Standard Arabic elements. This kind of code-mixing naturally tends to reduce the dialectal markedness of these colloquial elements (Mazraani 1997:76). The occurrence of mixed forms like *wa binaltaqī l-ʾān* ‘and we meet now’ on Egyptian television broadcasts and the absence of similar forms on Syrian television (Kaye 1994:59, with reference to Peter Behnstedt) not only reflects differences in the language policies of the respective television companies, but it also shows that the bridge between the two main varieties in the linguistic continuum is more busily trafficked in Cairo than, e.g., in Damascus, Beirut, or Baghdad.

A category of Classical items frequently occurring in Colloquial Arabic includes a number of denominative adverbs ending in *-an*, e.g. *tamāman* ‘exactly, completely’, *abadan* ‘[with negation] never’, *maṯalan* ‘for instance’, *taqrīban* ‘about, approximately’, *dāʾiman* ‘always’, *ṭabʿan* ‘of course’, *šukran* ‘thank you!’. These are used as lexical borrowings which have been adapted into local colloquials as such or with slight modifications, e.g. *tamāman*, *masalan*, *taʾrīban*, *dayman* (Cairo); *tamamn*, *matalan*, *tqribn*, *daʾiman* (Morocco) (Heath 1989:66). That the final *-an* is felt as an integral part of the lexical item rather than an adverbially used indefinite accusative is apparent, since it has not been adopted as a productive morpheme in colloquial (Palva 1969:33). Because of the well-established, common use of these items on all stylistic levels, in spite of their Classical morphological structure, they have lost their stylistic markedness. On the other hand, the majority of adverbs of the same pattern, e.g. *ḥāliyyan* ‘at present’, *sābiqan* ‘formerly’, *ḥaqqan* ‘really, indeed’, *xuṣūṣan* ‘especially’, *sarīʿan* ‘rapidly’, *ṭawʿan* ‘voluntarily’, *šayʾan mā* ‘to some extent’, *nawʿan mā* ‘in a fashion’, have still preserved their stylistic status as marked classicisms.

Other adverbial items which have found their way into Colloquial Arabic, but which can still be regarded as classicisms deliberately used in order to elevate the stylistic level, are phrases such as *ʾilā ʾāxirih(i)* ‘and so forth’, *bi-dūn šakk* ‘without doubt’, *bi-n-nisba l-* ‘with respect to’, *min nāḥiyat* ‘with respect to’, and *b-sabab* ‘because of’.

5. Stylistic polarization

The classicizing tendency sometimes leads to stylistic polarization. When speakers who wish to elevate the stylistic level of their speech have to make a choice between two variants existing in Classical Arabic, they may prefer the variant which markedly deviates from the variant which is closer to colloquial. An example of this kind of conscious distancing from dialect is the use of the negative construction *lam* plus jussive and the avoidance of *mā* plus perfect, e.g. *lam ʾarahu* vs. *mā raʾaytuhu* ‘I didn’t see him’ (Kaye 1994:57; Holes 1995:263). In the same way, the verb *istaṭāʿa* is in elevated colloquial preferred to *qadira*, because the latter item also belongs to the dialect. In Syrian Arabic, the plural *sanawāt* is regarded as more Classical than the variant *sinūna*, because the *status obliquus* of the latter is identical with the dialectal *s(i)nīn*. This stylistic phenomenon has also been described by Arab authors, among them Tawfīq al-Ḥakīm, Maḥmūd Taymūr, and Muḥammad Mandūr, who give numerous examples of polarized choices of words (Diem 1974:45–46).

6. PHONOLOGICAL CLASSICIZATION

In the dialects which have replaced their inherited interdentals with postdental equivalents, their sibilant substitutes are used as an intermediate phoneme series to fill the gap between the *fuṣḥā* and the *ʿāmmiyya*. The sibilant substitutes are common in both oral Modern Standard Arabic and mixed sequences; stylistically, they are un-marked and imply only a slight downgrading from the interdental realizations. In Cairene Arabic, [q] is a strong Modern Standard Arabic marker which in the monologues analyzed by Mazraani demands co-occurrent use of morpho-phonological *fuṣḥā* devices such as vowel patterning (*yuqāsi*), occasional *ʾiʿrāb* (*yuḥaqqiqa*), *ya*-preformative, and full suffixed personal pronoun (*qudratihi*) (Mazraani 1997:62–63).

In contrast to the interdentals, /q/ has no 'intermediate', stylistically unmarked phoneme between /ʾ/ and /q/, and the choice must be made between two marked variants. Using the ratio of occurrence of the variants as a criterion, Elgibali (1993:83–84) has through statistical methods tested the validity of Badawī's contention that each of the five levels of his stylistic classification of the linguistic continuum of Arabic has its specific combination of linguistic and social characteristics. The five levels were further divided into formal and informal registers. Badawī's description suggested that the ratio of occurrence of a given variant in the informal register would be higher than that of the same variant in the formal register of the immediately lower level. It appeared that such a hierarchy really exists between oral informal Classical Arabic and oral formal Modern Standard Arabic, as well as between oral informal Modern Standard Arabic and formal Educated Colloquial Arabic. On the other hand, no level/register distinction in this respect was observed between informal Educated Colloquial Arabic and formal Enlightened Colloquial Arabic, nor between informal Enlightened Colloquial Arabic and formal Illiterate Colloquial Arabic.

7. HYBRIDIZATION

Most speakers of Arabic make use of different stylistic levels and registers, which also implies use of mixed forms, hybrid combinations of Classical and Colloquial elements in all conversation which moves outside the sphere of everyday concerns. Not all theoretical combinations are, however, acceptable. If the markedly colloquial verb *šāf* is replaced by the markedly Classical *raʾā*, the latter may not only be inflected according to the rules of Classical Arabic, but it may also follow colloquial inflection. For example *raʾēt* 'I saw' is an acceptable hybrid. On the other hand, the colloquial item cannot have Classical Arabic inflectional morphemes, and theoretical hybrid forms like **šāfa* or **lam yašuf* are not acceptable to native speakers. Correspondingly, the dual form **rāḥā* of the dialectal verb *rāḥ* is a stylistically incompatible hybrid, whereas the *b(i)*-imperfect *b-yaḏhab* of the Classical verb *ḏahaba* is acceptable. It thus appears that hybridization is a rule-governed rather than a random process (Schmidt 1974:183–184; Holes 1995:296).

When a markedly Classical lexical item occurs in a colloquial matrix, it may be inflected without constraints according to the structure of the colloquial environment, whereas modifying colloquial discourse by using Classical grammatical items with markedly colloquial lexical items is stylistically unacceptable (Meiseles 1981:1085). The same holds true for phonologically classicized lexical items. The lexical and morphological elements of the Classical forms of *ʿayn*, *ʿaynayn* and the Cairene Arabic *ʿēn*, *ʿinēn* cannot be combined freely, but the Classical form may have either the Classical or the colloquial dual morpheme (*ʿaynayn* or *ʿaynēn*), whereas the colloquial form cannot have the Classical morpheme *-ayn*: **ʿēnayn*, **ʿinayn*. In a similar way, the Classical Arabic *tafakkarat* may be inflected dialectally *tafakkarit*, while the Cairene form *itfakkarit* cannot be modified toward Classical Arabic by using the Classical personal morpheme: **itfakkarat*.

The asymmetric constraint on combinations of stems and suffixes is plausibly explained by the fact that the underlying language of spontaneous speech normally is the language form first acquired by the speaker, in this case Colloquial Arabic. If speakers in a choice between the colloquial and the Classical lexical item select the former, they automatically apply the morphology of the underlying Colloquial as well, but if they select the Classical item, they are free to combine it with Classical or Colloquial grammatical items. Thus there is a hierarchy between the different categories of items: to a native speaker, the lexical choice seems to be stylistically more significant than the grammatical structure. For the pattern of the classicization process, this implies that when Colloquial discourse is modified toward Classical Arabic, lexical items come before grammatical ones in the hierarchy.

The hybrid forms often function as bridges between Colloquial and Classical Arabic. An illuminating example is the short answer given by an interlocutor when asked if he had any comment after a lengthy contribution to a conversation by a senior army officer: *mā fīš ʿandi ḥāga ʾuḍifha ʾila mā qālahu lginirāl* 'I have nothing to add to what the general said'. The answer starts with plain Cairene and switches over to Modern Standard Arabic, not abruptly, but using the hybrid *ʾuḍifha* as a bridge (Mitchell 1986:24–25; Holes 1995:301–302).

When used in colloquial discourse, the passive *yuqāl* can be regarded as a lexical borrowing. Due to the stylistic unmarkedness of the *b(i)*-imperfect, markedly Classical Arabic passive forms frequently occur in the formal register of Educated Colloquial Arabic. Examples of forms recorded in Cairo are *bitunāqaš* 'they are discussed', *bituʿraḍ* 'they are submitted', *fa-bitúṭbaʿu* 'so they are printed' (Diem 1974:76). In constructions of this type, the elements of the two main forms of the language are linked together but both of them retain their identity. Accordingly, Meiseles (1981:1083–1087) defines these forms as 'symbiotic', arguing that a distinction should be made between them and hybrids, which are intermediate forms, not identifiable in any one of the languages involved.

8. In which order do classicisms appear?

The classicisms most likely to appear first in Colloquial Arabic are lexical borrowings. Some of these are technical terms or other lexical items which do not have any counterpart in the dialect (commonly known as *Bedarfsentlehnungen*). Another category consists of Modern Standard Arabic lexical items which for stylistic reasons are substituted for their colloquial counterparts. The adaptation of the borrowed lexemes may affect their environments, since they often contain phonological and morphological elements diverging from the structure of the dialect. The next step in the classicization process is phonological modification toward Modern Standard Arabic, most often affecting the variants of the interdentals as well as of **q*, **k*, and **j*, re-diphthongization of monophthongized diphthongs, and restitution of lost short vowels and glottal stop. In the formal register of Educated Colloquial Arabic, grammatical classicisms sporadically appear, but as soon as more profound grammatical classicisms are used, such as adoption of the Classical tense and negation systems, the language form aimed at is Modern Standard Arabic.

The classicization tendency is certainly not a recent phenomenon. Although we are not able to observe its oral realizations in different varieties of Colloquial Arabic in the past, many parallel features can be traced in literature, in particular in the so-called → Middle Arabic texts, which often reflect the structure of the underlying dialect. The efforts to follow the norms of Classical Arabic often result in hybrid forms, either hyper- or hypocorrections (see, e.g., Blau 1970).

Bibliographical references

Abu-Melhim, Abdel-Rahman. 1991. "Code-switching and linguistic accommodation in Arabic". *Perspectives on Arabic linguistics*, III, ed. Bernard Comrie and Mushira Eid, 231–250. Amsterdam and Philadelphia: J. Benjamins.

Badawī, as-Saʿīd Muḥammad. 1973. *Mustawayāt al-ʿarabiyya al-muʿāṣira fī Miṣr*. Cairo: Dār al-Maʿārif.

Blanc, Haim. 1964. "Style variations in spoken Arabic: A sample of interdialectal educated conversation". *Contributions to Arabic linguistics*, ed. Charles A. Ferguson, 81–156. Cambridge, Mass.: Harvard University.

Blau, Joshua. 1970. *On pseudo-corrections in some Semitic languages*. Jerusalem: Israel Academy of Sciences and Humanities.

Diem, Werner. 1974. *Hochsprache und Dialekt im Arabischen: Untersuchungen zur heutigen arabischen Zweisprachigkeit*. Wiesbaden: F. Steiner.

——. 1978. "Divergenz und Konvergenz im Arabischen." *Arabica* 25:2.128–147.

Eid, Mushira. 1988. "Principles for code-switching between Standard and Egyptian Arabic". *al-ʿArabiyya* 21.51–79.

Elgibali, Alaa. 1993. "Stability and language variation in Arabic: Cairene and Kuwaiti dialects".

Perspectives on Arabic linguistics, V, ed. Mushira Eid, 75–96. Amsterdam and Philadelphia: J. Benjamins.

Grotzfeld, Heinz. 1964. *Laut- und Formenlehre des Damaszenisch-Arabischen*. Wiesbaden: F. Steiner.

——. 1965. *Syrisch-arabische Grammatik (Dialekt von Damaskus)*. Wiesbaden: O. Harrassowitz.

Heath, Jeffrey. 1989. *From code-switching to borrowing: Foreign and diglossic mixing in Moroccan Arabic*. London and New York: Kegan Paul International.

Holes, Clive. 1980. “Phonological variation in Baḥraini Arabic: the [j] and [y] allophones of /j/.” *Zeitschrift für Arabische Linguistik* 4.72–89.

——. 1995. *Modern Arabic: Structures, functions and varieties*. London and New York: Longman.

Kaye, Alan S. 1994. “Formal vs. informal in Arabic: Diglossia, triglossia, tetraglossia, etc.: Polyglossia–multiglossia viewed as a continuum”. *Zeitschrift für Arabische Linguistik* 27.47–66.

Mazraani, Nathalie. 1997. *Aspects of language variation in Arabic political speech-making*. Richmond, Surrey: Curzon.

Meiseles, Gustav. 1980. “Educated spoken Arabic and the Arabic language continuum”. *Archivum Linguisticum* N.S. 11.118–143.

——. 1981. “Hybrid versus symbiotic constructions: A case study of contemporary Arabic”. *Linguistics* 19.1077–1093.

Mejdell, Gunvor. 1999. “Switching, mixing: Code interaction in spoken Arabic”. *Language encounters across time and space*, ed. Bernt Brendemoen, Elizabeth Lanza, and Else Ryen, 225–241. Oslo: Novus Press.

Miller, Catherine. 2003. “Variation and change in Arabic urban vernaculars”. *Approaches to Arabic dialects: A collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday*, ed. Martine Haak, Rudolf de Jong, and Kees Versteegh, 177–206. Leiden: E.J. Brill.

Mitchell, T.F. 1980. “Dimensions of style in a grammar of educated spoken Arabic”. *Archivum Linguisticum* N.S. 11.89–106.

——. 1986. “What is educated spoken Arabic?”. *International Journal of the Sociology of Language* 61.7–32.

Palva, Heikki. 1969. “Notes on classicization in modern colloquial Arabic”. *Studia Orientalia* 40:3.1–41.

——. 1982. “Patterns of koineization in modern colloquial Arabic”. *Acta Orientalia* 43.13–32.

Schmidt, Richard Wilbur. 1974. *Sociostylistic variation in spoken Egyptian Arabic: A re-examination of the concept of diglossia*. Ph.D. diss., Brown University, Providence.

Heikki Palva (University of Helsinki)

Clitic

Clitics are elements that are affixed to another element, usually a head. This very general characterization encodes the fact that a multitude of elements can be affixes, hence clitics (see Kayne 1975). In this entry, the focus is on pronominal clitics in (Lebanese) Arabic that are non-nominative. Nominative affixes (or infixes) are usually viewed as agreement markers and analyzed as inflectional elements that are part of the verbal form. Nominative agreement markers are traditionally the purview of inflectional morphology, whereas pronominal clitics are the purview of derivational morphology. The literature on clitics in Arabic is extensive. The reader is referred to Benmamoun (2000) and the references mentioned there.

Pronominal clitics are unstressed elements that are generally attached to a head, as illustrated in the Lebanese Arabic example (1b). The examples in this entry will all be drawn from Lebanese Arabic.

(1) a. *sāmi* *zār* *karīm*
Sāmi visited.3ms Karīm
‘Sami visited Karim’

b. *sāmi* *zār-o*
Sāmi visited.3ms-him
‘Sami visited him’

The pronominal clitic *-o* ‘him’ in (1b) is attached to the verbal head and cannot bear stress.

Several characteristics of pronominal clitics in Lebanese Arabic are worth noting.

A head can host at most one clitic as illustrated by the non-grammaticality of (2b).

(2) a. *ṭayṭ-o* *ssiyāra*
gave.1s-him the-car.Fem
‘I gave him the car’

b. **ṭayt-o-ha*
gave.1s.-him-it
‘I gave it to him’

c. *ṭayṭ-o* *yēha*
gave.1s.-him dummy-it.Fem
‘I gave it to him’

The dummy form *yē* is used to support the second pronominal clitic as in (2c). Note that in some languages, for instance French, this prohibition does not exist. More than one clitic can be ‘attached’ to a verb, as in the following sentence from French:

Pierre le lui a donné
Pierre it.Masc him aux. gave
'Pierre gave it to him'

Cliticization is local: a clitic is generally attached to the head it is selected by. Thus, a direct (1b) or a relative object (3) are attached to the verb they are complement of, as in (3):

(3) *ḥkēte-llo ḥkeye*
told.1s-him story.Fem
'I told him a story'

Prepositional complements are attached to the preposition selecting them, as in (4):

(4) *sēfart maʕ-a*
traveled.1s with-her
'I traveled with her'

and adnominal complements are attached to the noun selecting them, as in (5):

(5) *šeft siyyārt-o*
saw.1s car-her
'I saw her car'

In many languages, clitics are in complementary distribution with the arguments they 'replace'. In several varieties of Levantine Arabic, this is not the case: clitics can occur with the argument noun-phase, in which case the argument noun phrase is preceded by a preposition *la* 'to'. This phenomenon is referred to as clitic-doubling (see 1).

(6) a. *sāmi zār-o la-karīm*
Sāmi visited.3ms-him to-Karīm'
'Sami visited Karim'

b. *sāmi ḥkēlo la-karīm ḥkēye*
Sāmi told.3ms-him to-Karīm story
'Sami told Karim a story'

c. *sāmi šēf ktēb-o la-karīm*
Sāmi saw.3ms book-his to-Karīm
'Sami saw Karim's book'

d. *sāmi rāḥ maʕ-o la-Karīm*
Sāmi went-3ms with-him to-Karīm
'Sami went with Karim'

Clitic doubling in Lebanese Arabic confirms what is usually referred to as Kayne's generalization (7):

(7) A lexical NP may be doubled by a clitic only if this nominal element is preceded by a (prepositional) case-assigner.

The doubled nominal element is assigned dative case as evidenced by the fact that the doubled pronominal element is morphologically dative:

(8) a. *sāmi zār-o la-ˀilo*
Sāmi visited.3ms-him to-him.Dat
'Sami visited him'

b. *sāmi ḥkēlo la-ˀilo*
Sāmi told.3ms-him.Dat to-him.Dat
ḥkēye
story
'Sami told him a story'

c. *sāmi šēf ktēb-o la-ˀilo*
Sāmi saw.3ms book-his to-him.Dat
'Sami saw his book'

d. *sāmi rāḥ maʕ-o la-ˀilo*
Sāmi went.3ms with-him to-him.Dat
'Sami went with him'

(9) The doubled (pro-)nominal element is assigned dative case.

Nominative subjects cannot be doubled, whether they are in the S(ubject)V(erb) or V(erb)S(ubject) order:

(10) a. (**la-*)*sāmi rāḥ*
to-Sāmi left.3ms
'Sami left'

b. *rāḥ* (**la-*) *sāmi*
left.3ms to-Sāmi
'Sami left'

Non-nominative subjects in the so-called Exceptional Case-Marked Contexts can be doubled. In (11a), the subject receives its case from the matrix verb as evidenced by non-nominative clitic in (11b):

(11) a. *xallēt sāmi yrūḥ*
let.1s. Sāmi leave.3ms
'I let Sami go'

b. *xallayt-o yrūḥ*
let.ls-him leave.3ms
'I let him go'

This non-nominative subject can be doubled:

(12) *xallayt-o yrūḥ la-sāmi*
let.ls-him leave.3ms to-Sāmi
'I let Sami go'

The doubled nominal element is not in the same position as the non-doubled one, contrary to what is usually assumed for doubled elements in Romance languages. The facts establishing this generalization are discussed extensively in Aoun (1999); some are reviewed below. Consider the following sentences:

(13) a. *ṣāḥibt-o šēfit kill walad*
friend.Fem-his saw.3fs every boy
'His friend saw every boy'

b. *ṣāḥibt-o šēfit-o la- kill walad*
friend.Fem-his saw.3fs to every boy
'His friend saw every boy'

(14) a. *ṣāḥibt-o rāḥit maʕ-a kill walad*
friend.f-his went.3fs with every boy
'His friend went with every boy'

b. *ṣāḥibt-o rāḥit maʕ-o la-*
friend.f-his went.3fs with-him to
kill walad
every boy
'His friend went with every boy'

In (13a) and (14a), the bound reading of the pronoun contained within the subject *ṣāḥibto* 'his friend' cannot obtain: the quantifier *kill walad* 'every boy' is not in a high enough position, i.e. it does not c-command the pronoun (→ resumption, → binding for relevant definitions). However, the bound reading of the pronoun within the subject obtains when the quantificational object is doubled as in (13b) and (14b). These sentences indicate that doubled accusatives and doubled prepositional complements are in a position higher than their corresponding non-doubled arguments.

It is possible to show that doubled nominal elements attach to the complete functional complex containing the doubling clitic. A complete functional complex can be described informally as a Determiner Phrase (DP) (6c), a predicate such as VP (6a), or a preposition which takes a subject as in (15):

(15) *sāmi šēf sāmia ḥadd-o*
Sāmi saw.3sm Sāmia next-him
'Sami saw Samia next to him'

In (15), the prepositional phrase takes *sāmia* as its subject. On the other hand, the prepositional phrase in (16) does not take a subject:

(16) *sāmi rāḥ maʕ-o*
Sāmi went.3ms with-him
'Sami went with him'

Accordingly, the doubled element may attach to the prepositioned phrase in sentence (17) but not in sentence (6d), which is repeated here as (18):

(17) *sāmi šēf sāmia ḥadd-o*
Sāmi saw.3sm Sāmia next-him
la-karīm
to-Karīm
'Sami saw Samia next to Karim'

(18) *sāmi rāḥ maʕ-o la-karīm*
Sāmi went.3ms with-him to-Karīm
'Sami went with Karim'

Under standard assumptions according to which extraction processes affect single constituents, the prepositional phase and the doubled element can be treated as a unit and 'fronted' in (17) but not in (18). This is the case as illustrated in (19) and (20):

(19) a. *ḥadd-o la-karīm sāmi šēf*
near-him to-Karīm Sāmi saw.3ms
sāmia
Sāmia
'Near Karim Sami saw Samia'

b. *ḥadd-o la-ʔayya walad sāmi*
near-him to-which boy Sāmi
šēf sāmia
saw.3ms Sāmia
'Near which boy did Sami see Samia?'

(20) a. **maʕ-o la-karīm sāmi rāḥ*
with-him to-Karīm Sāmi went.3ms
'With Karim Sami went'

b. **maʕ-o la-ʔayya walad sāmi*
with-him to-which boy Sāmi
rāḥ
went.3ms
'With which boy did Sami go?'

PP-fronting and wh-movement treat the prepositional phrase and the doubled element in (19a–b), but not in (20a–b), as a unit and raise them to the left edge of the clause.

Determiner Phrases (DPs) form a complete functional complex. As such, a doubled element can be fronted with the DP as in (6c) which is repeated here as (21):

(21) a. *ktēb-o* *la-karīm* *sāmi* *šēf-o*
book-his to-Karīm Sāmi saw.3ms-it
'Karim's book, Sami saw it'
b. *ktēb-o* *la-ʾayya* *walad* *sāmi*
book-his to-which boy Sāmi
šēf
saw.3ms
'Which boy's book did Sami see?'

The construct DP and the doubled element in (20a) and (20b) are treated as a unit and fronted to the left edge of the clause by clitic left dislocation and wh-movement respectively.

Summary
(21) a. Pronominal cliticization is local: clitics are attached to the head that selects them.
b. Clitic doubling is quite productive: non-nominative nominal elements may be doubled.
c. Doubled elements are preceded by a prepositional which assigns the doubled element dative case.
d. Doubled elements are not in the same position as non-doubled ones: they are generated in a position higher than the non-doubled ones.
e. Doubled elements attach to the minimal complete functional complex containing the clitic.
f. As such, doubled elements can form a constituent with a DP or a complex PP but not with a simplex PP.

A more complete analysis of the clitic doubled construction is to be found in Aoun (1999).

BIBLIOGRAPHICAL REFERENCES
Aoun, Joseph. 1999. "Clitic-doubled arguments". *Beyond principles and parameters: Essays in memory of Osvaldo Jaeggi*, ed. Kyle Johnson and Ian Roberts, 13–42 Dordrecht: Kluwer Academic Publishers.
Benmamoun, Elabbas 2000. *The feature structure of functional categories: A comparative study of Arabic dialects*. Oxford: Oxford University Press.
Kayne, Richard. 1975. *French syntax: The transformational cycle*. Cambridge, Mass.: MIT Press.

JOSEPH AOUN
(University of Southern California)

Code-switching

1. INTRODUCTION

'Code-switching' is generally used as a cover term for a linguistic or discourse practice in which elements and items from two or more linguistic systems, or *codes* – be they different languages or varieties of a language – are used in the same language act or interaction. As a field of research it emerged (in the 1970s and 1980s) from the broader field of language contact studies, and scholars insist it must be conducted in that broader context (e.g. Gardner-Cloros 1995).

It is common to distinguish between on the one hand *inter-sentential* code-switching, i.e. switching which occurs at clause boundaries, and on the other *intra-sentential* code-switching, i.e. where switching (or *mixing*) takes place within the domain of a clause. Some analysts prefer to operate with the terms *alternational* code-switching for switching between stretches of speech belonging to one and the other code/language/variety; while *insertional* code-switching denotes single items of one code occurring in stretches of the other code (e.g. Muysken 1995). More recently, certain sociolinguists have proposed to reserve *code-switching* as a term for socially meaningful changes of code in (conversational) interaction; this is in contrast with the switching between elements and items from more than one code where the switches are not in themselves socially meaningful, in which case we have language mixing, a mixed variety, but not code-switching (Auer 1995, 1998).

The sociolinguistic literature on code-switching is extensive. This entry introduces briefly some major theoretical contributions to the field of code-switching, and then some contributions which involve or are related to bilingual code-switching with Arabic as one of the languages; and finally it discusses cases involving what may be called 'diglossic code-switching', i.e. studies of the interaction in speech of standard Arabic and a colloquial variety.

Two major perspectives or approaches to the study of code-switching may be distinguished (although in most studies they are combined to some extent):

i. the discourse/pragmatic perspective, with the main focus on social and communicative functions of, and motivations for, code switching: why and for what purposes do speakers engage in code-switching?

ii. the grammatical/syntactic perspective, with the main focus on linguistic aspects, especially morphosyntactic constraints on intrasentential switching: where in a sentence may or may not a speaker change languages?

2. The discourse/pragmatic perspective

Research reflecting the discourse/pragmatic approach is generally interpretative, i.e. the researcher interprets the meaning of code-switching according to role relationships between speakers, ethnic identities, and communicative functions assigned to the different codes or languages involved by the norms of the bilingual community or different subgroups within it. Typically, a bilingual may use different codes at home and in public, with friends and with a teacher, with the ethnic in-group vs. out-groups.

Instances of code-switching are seen as 'discourse strategies' (Gumperz 1982). Not only may a change in language code reflect a change of domain, for instance topic or situation (situational code-switching), but speakers may switch from one code to another in order to act on, or interfere with, the context, to signal a shift in attitude toward what is going on, influence the tone of the interaction, the formality of the situation and/or the relationship between participants (metaphorical code-switching). A wide range of perspectives is raised in Heller (1988).

Myers-Scotton's work (1993a; 1993b) has been particularly influential in the last decade. She brings together elements from discourse analysis, ethnography of speaking, and accommodation theory. In her model bilingual speakers are assumed to be negotiating the code of their communicative interaction according to what they perceive as the relevant features of the particular situation and the community norm for such an interaction (in the model called the 'Rights and Obligations [RO]' set). Speakers may comply and use the linguistic code which is normally associated with, i.e. considered appropriate for, the actual RO set, i.e. the unmarked choice. They may alternatively make a marked linguistic choice, by choosing or switching to another code, thereby signaling a different attitude, challenging, and changing, the RO set of the interaction. Code-switching often has this function of marking a change in the RO set. Sometimes, however, code-switching itself is the unmarked choice for a particular interaction. This occurs "when the speaker wishes to index two identities or attitudes towards the interaction simultaneously" (1993a:149). Such code-switching modes of speaking can be very frequent in bilingual communities.

3. Grammatical/syntactic approaches

The more linguistically oriented approaches to code-switching are generally concerned with describing and identifying patterns and regularities in the linguistic make-up of (primarily) bilingual speech. The regularities tend to be formulated as restrictions or constraints on the free switching between and mixing of items from the two basic languages or varieties involved. Some are claimed to have universal applicability, i.e. to have predictive force for all language sets. Switching is studied at clause borders (especially involving selection of conjunctions), while the main focus has naturally been on intrasentential code-switching, with constituents and part of constituents from more than one linguistic system combined.

Following Boumans (1998) and Muysken (1995) two main – but sometimes overlapping – trends may be distinguished: on the one hand, the linear or alternational approach, which focuses on patterns and restrictions in terms of the (word) order of categories or constituents that may be switched and on the other hand, the structural, or insertional approach, which explains regularities of code-switching data in terms of structural relations that hold between constituents or parts of constituents. The insertional approach presupposes that one of the two languages or codes involved is structurally dominant and sets the morphosyntactic frame of the sentence, whereas items from the other language/code are inserted into this sentence frame.

The linear approach is reflected in a number of contributions – most famously formulated in Poplack's (1988:219) principles of syntactic constraints on code-switching, based on data from English/Spanish code interaction: the 'free morpheme constraint', "which prohibits mixing morphologies within the confines of the word", and the 'equivalence constraint', "which requires that the surface word order of the two languages be homologous in the vicinity of the switch point". In other words, it

> predicts that code switches will tend to occur at points where the juxtaposition of elements from the two languages does not violate a syntactic rule of either language. [. . .] This means that a language switch ought to take place only at boundaries common to both languages. [. . .] From a cross-linguistic perspective, this means that the more similar two languages are structurally, the more switching sites they should permit (Romaine 2000:58).

Poplack's propositions triggered extensive research applying the principles to other language sets, with the effect of producing cycles of evidence, counter evidence, and further research.

4. The matrix language frame model

Basic to the influential model of Myers-Scotton (1993b and later contributions) is the claim that the languages involved in (intrasentential) code-switching have unequal status and play different roles in the speech interaction, and that only one of the languages – the *matrix* language – provides the basic grammatical structure at a time. More specifically, the model claims that the matrix language (ML) sets the morphosyntactic frame (like word order and inflection) for the sentence. The other language, or code, involved is the *embedded language* (EL), which may provide constituents or single items to be inserted into or embedded in the matrix language base – according to specific requirements elaborated in the model.

Crucial to the model, and to the definition of the matrix code, is the distinction between content morphemes and system morphemes. In very general terms, the distinction corresponds to lexical vs. grammatical or function morphemes. The general insight concerning the asymmetry of the codes involved, and the different workings of lexical items on the one hand, and grammatical/struc-tural items on the other, are generally accepted by researchers in the field. To a considerable extent these notions were also explicitly or implicitly assumed in earlier studies of language contact, e.g. Hasselmo 1970, Petersen 1988 (for a very neat historical survey, cf. Boumans 1998:7–60).

5. Code-switching vs. borrowing

Many studies operate with an analytical distinction between code-switching and borrowing, especially with regard to single lexical items. Among the most common criteria used in the literature to characterize borrowing are: the use of a foreign lexeme, also in the speech of monolingual speakers, often to fill a lexical gap, as when there exists no term for the (new) item or concept in the borrowing language; morphological integration; phonological adaptation; and frequency of occurrence.

While recognizing the usefulness of some kind of differential status for the established loanword and the spontaneous use/transfer of an item from another language in code-switching, the distinction often appears difficult to apply in a systematic way. Poplack introduced the notion of 'nonce borrowings' to account for cases which contradict her free morpheme constraint, e.g. where "*unadapted* English morphemes conjoined with French verbal and participial affixes" – as in *enjoyer* and *drowné*. In matrix language terms, these items would be embedded language content morphemes; in terms of Petersen's (1988) dominant language hypothesis they would be non-dominant language lexical morphemes with dominant language grammatical morphemes.

6. Bilingual code-switching with Arabic

The wide definition of code-switching sketched above might include many – if not most – cases of language contact associated with interference phenomena at various linguistic levels. Rouchdy's (1980) study of Arabic/Nubian (→ Nubian) is one case in point, and several studies of Arabic as a minority language (Owens 2000) and of peripheral Arabic-speaking communities in contact with neighboring languages may fall within such a scope (e.g. Procházka 2002; Miller 2002; Arnold 2002), including the historical cases of Malta (→ Maltese), Sicily (→ Sicilian Arabic), and → al-Andalus. Here, the scope is limited to some studies where code-switching is the explicit field of research and its principles are the focus of attention.

7. Arabic/English

Eid (1992) provides a study of syntactic restrictions on code-switching in the linear approach, based on recorded conversations among bilingual Egyptian Arabic and American English speakers. She examines switching at the syntactic boundaries of coordinate, subordinate (specifically adverbial), relative, and complementary clauses, i.e. whether switching occurs immediately before or after the grammatical markers introducing such clauses. In her data, there appears to be a general restriction precluding switching to Arabic following an English marker. For the categories coordination and subordination, switching may take place immediately before the grammatical marker (the conjunction), whatever the language, and an Arabic conjunction may be followed by either code (*pace* claims for Spanish/English, that "the conjunction always goes with the second switched phrase", cf. Gumperz 1982:88). For relative clauses, however, the relative marker must be followed by a word in the same language. With complementizers, the actual combinations are even more restricted (although the analysis of this issue is complicated by the optionality of complementizers as well as the problem of pronoun doubling in switching before English verbs). The discussion of the latter phenomenon takes Eid outside the limits of the strictly linear approach, and (almost) into acknowledging the unequal status of the languages involved. Pronoun doubling is explained with reference to proposed universal constraints on code-switching between verbs and their pronouns. Myers-Scotton a.o. (1995) is a highly theoretical contribution, presenting and expanding the Matrix Language Frame model, which is applied, e.g., to the same issue of 'pronoun doubling' as well as to embedding of English verbs into an Arabic matrix language frame. The authors claim that bare forms of English verbs are preferred to verbs inflected with Arabic affixes in their Arabic-English data. They do cite one example: *bas* ***cancel****-t-uh* "but I cancelled it" with Arabic 1st pers. suffix *-t* (1995:33) – and one wonders whether the English verb *checked* in the following is not perceived as ending with the same Arabic ending *-t: ruḥt el* ***library*** *imbāriḥ wa* ***checked books*** *ktīri* 'I went to the library yesterday and checked [classified as English past tense suffix] many books' (1995:33).

Rouchdy (1992) reports on English lexical borrowings into Arabic in Arab-American immigrant data (→ English loanwords). She provides some examples of integrated verbs, e.g. *kalnīt* 'I cleaned', *fakkasna* 'we fixed', *la yusammok wa la yudarnik* 'he doesn't smoke and doesn't drink', *nayselluh* 'say something nice to him' [verb formed from the English adjective on the model of *gāmil* 'treat nicely' from *gamīl*?]; derived participles, e.g. *mifarniš* 'furnished', *mihayyat* 'heated'. Rouchdy observes the constraint that verbs never take English pronoun objects, only the Arabic pronoun clitic, and that English adjectives are not inflected (for gender or number), which according to her implies that they are switched rather than borrowed.

Sallo (1994), on code-switching among university students in Mosul, also provides interesting data, e.g. collocations with English head nouns and Arabic modifier adjectives where the adjective shows gender agreement with the Arabic equivalent of the noun: *as-****sensitivity*** *ʿāliya* 'the sensitivity is high [fem.]' (Arabic *ḥassāsiyya* [fem.]), *at-****temperature*** *munxafiḍa* 'the temperature is low [fem.]' (Arabic *daraja* [fem.]) (1994:124). Likewise, in the Iraqi Arabic possesive construction with *māl: ar-****result*** *mālti takūn* ***negative*** 'my result will be negative' [with the fem. form *mālt* + pronoun suffix 1st pers. sg. *-i-*; verb inflected for fem. sg.; Arabic noun *natīja* [fem.]) In the reverse case, the modifying English adjective is not inflected according to Arabic agreement rules, as in the last example above (1994:124). One also notes the use, according to Arabic rules, of the singular form of the counted noun following numerals 11–99: *nāxuḏ xamasṭaʿš* ***rat*** 'we will take fifteen rats' (1994:120).

8. Arabic/French

With French colonial pressure and presence in North Africa, French and the local Arabic vernaculars were brought into close contact. The educational system produced a significant number of bilingual speakers, and code-switching as the unmarked choice appears to be prevalent among professionals and young people in many urban circles. Quoting from a study by Lahlou on linguistic practice among urban educated in Morocco, Caubet (1998:98) writes: "CS [code switching] has become their usual everyday means of interaction with their peers. CS is their 'default mode' of conversation, a mode which is in the middle of their linguistic continuum, with Moroccan Arabic at one end of the continuum and French at the other".

Bentahila and Davies (1983) make a significant contribution to the literature on code-switching, and their Moroccan Arabic/French data is often referred to. While they provide many exceptions to some of the general Equivalence Constraints formulated by Poplack, they provide a refinement of some other constraints, e.g. *subcategorization rules* to account for restrictions on combining certain grammatical structures with French lexicon (*waḥid l-professeur* 'a professor' is acceptable, but not **waḥid professeur*, 1983:321). There is frequent embedding of French verbal lexemes with Arabic inflections, e.g. *tatbqa tatgratter* 'you keep scratching' (1983:315), but subcategorization rules inhibit **taybqa confronter ces idées* 'he keeps opposing these ideas', and also the reverse case, **je dois nṣelli* 'I should pray' – as Arabic requires a finite verb, French an infinite form following the first verb (1983:322). The issue of 'subject pronoun doubling' is treated in terms of Arabic topic pronoun use + clitic French pronoun. The article at several points suggests an inherent asymmetry in the function of the two languages involved (further developed in Bentahila and Davies 1993) – an issue which became crucial in the Matrix Language Frame model.

M'barek and Sankoff (1988) apply the equivalence constraint and the borrowing vs. switching distinction to Moroccan Arabic/French data. They find, however, that the notion of insertion must be introduced to account for the frequent use of NP constituents with French article + French noun in an otherwise Arabic context (Moroccan Arabic is the matrix language of their data).

Heath (1989) is concerned with "the gradual integration of borrowed lexical materials" in Moroccan Arabic, and looks at code-switching – at the phrasal and lexical levels rather than the syntactic level – "as an avenue for more complete integration" (1989:23). Among a wealth of attested borrowings and 'borrowing routines', Heath also observes the retention of the French article with the switched noun – even when the structure otherwise is Arabic, e.g.: *xdəm-t f-waḥəd la société d'assurances* 'I worked in an insurance company', where the French noun with its French (definite) article is inserted into the Moroccan indefinite structure *waḥəd* + article + noun, as in *waḥəd l-bənt* 'a girl'. Integrated borrowings, however, mostly do not retain the French variant of the article – contrary to what he observes in Algeria – but receive the Arabic *l-* (or assimilated variant), which does not show gender accord (1989:34–36).

The superficial phonological similarity between French verbs in their infinitive form and the 'unmarked' form (3rd pers. masc. sg. perfect) of the Arabic verb is seen as a bridge to morphophonological integration: French *changer* 'change' > Arabic *šãža* '[he] changed'. Then other forms of the verb are made with Arabic affixes: *y-šãži* 'he changes'; *n-šãži* 'I change', etc. This productive device is also noted by Caubet (1998), who gives more examples, e.g. French *contacter* > Arabic *kõtakta* [perf.], *ykõtakti* [imperf.]; *šãžēt-u* 'I changed it'; *(bāš) n-developi-hum* '(so that) I can develop them'.This shows a feature of code-switching with Arabic that is specific for this set of languages, while it does not work with English (see above) or Dutch (see below).

9. ARABIC/DUTCH

With new waves of immigrant workers from 'southern lands' into industrialized → Europe from the 1970s and onwards, Western linguists had easy access to interesting language contact phenomena. Of the new bilingual language sets, Dutch and Arabic have been extensively studied by Boumans (1996, 1998) and also by Nortier (1989).

Boumans applies the Matrix Language Frame-work, or rather, critically adapts it to data on Moroccan Arabic in Dutch environments. He adds to the model a distinction between Community Language (the immigrants' language) and Super-imposed Language (of the Dutch language environment): as matrix languages of a stretch of speech they operate differently when it comes to embedding of constituents. He also calls to attention language specific structural differences in the working of certain 'general' principles of the model, comparing his own data with Dutch/Turkish data on the one hand and with French/Moroccan data on the other. He notes, for instance, the omission of the definite prefix before inserted Dutch nouns in Moroccan Arabic matrix (1996 and 1998), contrasting with the regular use of the definite article in insertions of French nouns (as mentioned above). Different strategies come into play for embedding verbs with different language pairs, such as adding Arabic affixes, or restructuring the stem to fit Arabic morphological patterns, or using auxiliary verbs like 'to do', which receive the inflection + a non-finite form of the embedded verb (1996:55–61; Boumans and Caubet 2000).

10. 'DIGLOSSIC' CODE-SWITCHING

It is often assumed by non-Arabist sociolinguists, that 'diglossia' as a language situation type prevents code-switching, in that "there is an almost one-to-one relationship between language choice and social context" so that "only one code is usually employed at any one time" (Romaine 1995:121). Similarly, Myers-Scotton (1993a:128) states: "The expectation is that unmarked CS [code-switching] should not occur at all in narrow diglossic communities (the Arabic-speaking nations of the Middle East, at least)".

Ferguson (1959) already announced the emergence of "unstable intermediate forms of the language", and code-switching as well as code-mixing – intersentential and intrasentential – indeed frequently occurs in many communicative situations. In fact, linguistic data discussed in the literature as stylistic 'levels', 'classicization' (→ classicism), 'colloquialization', and '(sociostylistic) variation' (cf. Holes 1995, ch. 9, for a survey of this literature) borders on, even overlaps with, 'code-switching' in the broad sense adopted here. For instance, the media extracts in Diem (1974), which are classified as 'High' variety with (varying degrees of) 'interference from dialect' or as 'dialect with interference from the High variety', or as 'mixed language', may be treated in a framework of one variety as matrix and the other as embedded, and subjected to constraints of one or the other order. It may often simply be a matter of tradition or preference of conceptual and analytic framework. Of course, switching between varieties of a language, in which there are many points of structural convergence and shared lexicon, offers less clear-cut data for analysis compared to bilingual data. For those same reasons, 'mixing' codes may be extensive, and the identification of a matrix code problematic. On the other hand, in 'diglossic' code-switching cases, it is generally granted, that the 'Low' (vernacular) variety, if not necessarily providing the matrix in all speech events, will at least be the 'dominant' code (Petersen 1988) or 'community' code (in Bouman's 1998 terms), which significantly affects the role of the codes involved for insertional and combinatorial patterns in switching (for a discussion see Mejdell 1999).

The functions and motivations of 'diglossic' code-switching are largely similar to those noted for bilingual switching. Besides the borrowing of items to 'fill a gap in the lexicon', they are of a pragmatic, social, and/or stylistic order. Holes (1993:33) sums up the rhetorical functions of 'level switching' in the late president Nasser's famous public speeches: "*fuṣḥā* is used by Nasir to convey messages which are abstract, idealized or metaphorical [...]. The *ʿāmmiyya*, on the other hand, is used to convey the concrete and the physical, and is strongly associated with the personalization of issues". Wilmsen (1996) points to additional pragmatic functions such as satire and irony, and quotes Badawī as to the creative aspect of code-mixing behavior. Mazraani (1997), Taine-Cheikh (2002), Bassiouney (2003), and Mejdell (2005) are further studies of code alternation and code-mixing as discourse strategies in monologues. Amara (1995) gives evidence of code-switching in classroom situations. While accommodation (Abu Melhim 1991) or speakers' stylistic convergence (Diem 1974) is generally assumed, Mejdell (1999) provides an example of non-accommodating behavior, where two cultured personalities insist on using the standard and the colloquial variety respectively throughout the interaction.

The asymmetrical status of the varieties becomes apparent in the contributions of Eid (1982, 1988) on 'the principles of code-switching between standard and Egyptian Arabic', in the same vein as her (1992) study of bilingual switching (discussed above). The patterns which emerge across the features examined are also similar: the variant of the word preceding the grammatical marker (the 'focal point') is 'free', i.e. not bound to the variety (standard or Egyptian) of the marker, whereas if the marker is standard, the element immediately following it must also be standard (1988:61). If the grammatical marker is Egyptian, "switching was found to be permitted after all focal points *except* after the negative"

(1988:61). This constraint is explained by the structural incongruity of the tense + negative markings between the two grammatical systems. Eid suggests that the general pattern of asymmetric conditioning is "related to the manner of acquisition of each variety: which variety was natively, and which was non-natively learned"; in other words, it is linked up with the 'dominant language principle'.

11. Conclusion

The importance of code-switching for linguistics in general is formulated by Muysken (1995:178) as follows:

> One of the crucial questions in modern linguistics is the division of labor between the lexicon and the grammar of a language. To what extent do we rely on properties of individual words, when we produce and comprehend utterances, and to what extent on general rules of the language we speak? [. . .] When sentences are built up with items drawn from two lexicons, we can see to what extent the sentence patterns derive from the interaction between these two lexicons.

For studies on Arabic, code-switching research challenges notions of 'discreteness' of codes, and provides insights into how the varieties intermingle and mix, in a structured, albeit not predictable way. The social aspects of code-switching behavior will probably only increase in significance in the 'age of globalization'. Finally, there is a cultural and aesthetic role to code-switching, which is captured by Caubet (1998:96): "One must also note that code-switching can be a real pleasure in which true bilinguals like to indulge".

Bibliographical references

Abu-Melhim, Abdel-Rahman. 1991. "Code-switching and linguistic accommodation in Arabic". *Perspectives on Arabic linguistics*, III, ed. Bernard Comrie and Mushira Eid, 232–250. Amsterdam and Philadelphia: J. Benjamins.

Amara, Muhammad Hasan. 1995. "Arabic diglossia in the classroom: Assumptions and reality". *Israel Oriental Studies* 15.131–142.

Arnold, Werner. 2002. "Code switching and code mixing in the Arabic dialects of Antioch". Youssi a.o. (2002:163–68).

Auer, Peter. 1995. "The pragmatics of code-switching". Milroy and Muyskens (1995:115–135).

——. (ed.). 1998. *Code-switching in conversation*. London and New York: Routledge.

Bassiouney, Reem. 2003. "Theories of code switching in the light of empirical data from Egypt". *Perspectives in Arabic linguistics*, XV, ed. Dilworth Parkinson and Samira Farwaneh, 19–39. Amsterdam and Philadelphia: J. Benjamins.

Bentahila, Abdelali. 1992. "Code-switching and language dominance". *Cognitive processing in bilinguals*, ed. R.J. Harris, 443–458. Amsterdam: North-Holland.

—— and Eirlys E. Davies. 1983. "The syntax of Arabic-French code-switching". *Lingua* 59.301–30.

Boucherit, Aziza. 2000. "Réflexions sur le contact de langues à partir du cas d'Alger". Mifsud (2000: 83–88).

Boumans, Louis. 1996. "Embedding verbs and collocations in Moroccan Arabic/Dutch codeswitching". Eid and Parkinson (1996:45–68).

——. 1998. *The syntax of codeswitching: Analysing Moroccan Arabic/Dutch conversations*. Tilburg: Tilburg University Press.

——. 1998b. "Topic pronouns in monolingual Arabic and in codeswitching". Mifsud (1998:89–94).

—— and Dominique Caubet. 2000. "Modelling intrasentential codeswitching: A comparative study of Algerian/French in Algeria and Moroccan/ Dutch in the Netherlands". *Arabic as a minority language*, ed. Jonathan Owens, 113–180. Berlin and New York: Mouton de Gruyter.

Caubet, Dominique. 1998. "Codeswitching in the Maghreb: Why is French Algerianised?". Mifsud (1998: 95–100).

Diem, Werner. 1974. *Hochsprache und Dialekt im Ara-bischen: Untersuchungen zur heutigen arabischen Zweisprachigkeit*. Wiesbaden: F. Steiner.

Eid, Mushira. 1982. "The non-randomness of diglossic variation in Arabic". *Glossa* 16:1.54–84.

——. 1988. "Principles for code-switching between Standard and Egyptian Arabic". *al-ʿArabiyya* 21.51–79.

——. 1992. "Directionality in Arabic–English code-switching". Rouchdy (1992:50–71).

—— and Clive Holes (eds.). 1993. *Perspectives on Arabic linguistics*, V. Amsterdam and Philadelphia: J. Benjamins.

—— and Dilworth Parkinson (eds.). 1996. *Perspectives on Arabic linguistics*, IX. Amsterdam and Philadelphia: J. Benjamins.

Ferguson, Charles A. 1959. "Diglossia". *Word* 15. 325–40.

Gardner-Chloros, Penelope. 1995. "Code-switching in community, regional and national repertoires". Milroy and Muysken (1995:68–89).

Gumperz, John J. 1982. *Discourse strategies*. Cambridge: Cambridge University Press.

Hasselmo, Nils. 1972. "Code-switching as ordered selection". *Studies for Einar Haugen presented by friends and colleagues*, ed. Evelyn Scherabon Firchow a.o., 261–280. The Hague and Paris: Mouton.

Heath, Jeffrey. 1989. *From code-switching to borrowing: Foreign and diglossic mixing in Moroccan Arabic*. London: Kegan Paul International.

Heller, Monica (ed.). 1988. *Codeswitching: Anthropological and sociolinguistic perspectives*. Berlin and New York: Mouton de Gruyter.

Holes, Clive. 1993. "The uses of variation: A study of the political speeches of Gamal Abd al-Nasir". Eid and Holes (1993:13–45).

——. 1995. *Modern Arabic: Structures, functions and* *varieties*. London: Longman. [Especially ch. 9 "Language level".]

Mazraani, Nathalie, 1997. *Aspects of language variation in Arabic political speech-making*. Richmond: Curzon.

M'barek, Mohamed Nait and David Sankoff. 1988. "Le discours mixte arabe/français: Emprunts ou alternances de langue?". *Canadian Journal of Linguistics* 33:2.143–154.

Mejdell, Gunvor. 1999. "Switching, mixing – code interaction in spoken Arabic". *Language encounters across time and space*, ed. Bernt Brendemoen, Elizabeth Lanza, and Else Ryen, 225–241. Oslo: Novus.

——. 2005. *Somewhere between order and chaos: Mixed styles in spoken Arabic in Egypt*. Ph.D. diss., University of Oslo.

Mifsud, Manwel (ed.). 2000. *Proceedings of the Third International Conference of AIDA*. Malta: Salesian Press.

Milroy, Lesley and Pieter Muysken (eds.). 1995. *One speaker, two languages: Cross-disciplinary perspectives on code-switching*. Cambridge: Cambridge University Press.

Muysken, Pieter. 1995. "Code-switching and grammatical theory". Milroy and Muysken (1995: 177–198).

Myers-Scotton, Carol. 1993a. *Social motivations for codeswitching: Evidence from Africa*. Oxford: Clarendon.

——. 1993b. *Duelling languages: Grammatical structure in codeswitching*. Oxford: Clarendon.

——. 1997. "Code-switching". *The handbook of socio-linguistics*, ed. Florian Coulmas, 217–237. Oxford: Blackwell.

——, Janice L. Jake, and Maha Okasha. 1996. "Arabic and constraints on codeswitching". Eid and Parkinson (1996:9–44).

Owens, Jonathan (ed.). 2000. *Arabic as a minority language*. Berlin: Mouton de Gruyter.

Petersen, Jennifer. 1988. "Word-internal code-switching constraints in a bilingual child's grammar". *Linguistics* 26.479–493.

Poplack, Shana. "Contrasting patterns of code-switching in two communities". Heller (1988:215–244).

Procházka, Stephan. 2002. "Contact phenomena, code-copying, and code-switching in the Arabic dialects of Adana and Mersin (Southern Turkey)". Youssi a.o. (2002:133–139).

Romaine, Suzanne. 1995. *Bilingualism*. Oxford: Blackwell.

——. 2000. *Language in society: An introduction to sociolinguistics*. Oxford: Oxford University Press.

Rouchdy, Aleya. 1980. "Languages in contact: Arabic-Nubian". *Anthropological Linguistics* 22: 8.334–344.

——. 1992. "Borrowing in Arab-American speech". Rouchdy (1992:36–49).

——. (ed.). 1992. *The Arabic language in America*. Detroit: Wayne State University Press.

Sallo, Ibrahim-Khidhir. 1994. "Arabic-English codeswitching at the university: A sociolinguistic study". *Language, discourse and translation in the West and Middle East*, ed. Robert de Beaugrande, Abdulla Shunnaq, and Mohamed Helmy Heliel, 115–131. Amsterdam and Philadelphia: J. Benjamins.

Taine-Cheikh, Catherine. 2002. "De la variation linguistique dans le prêche populaire mauretanien". *Language contact and language conflict in Arabic: Variations on a sociolinguistic theme*, ed. Aleya Rouchdy, 179–202. Richmond: Curzon.

Versteegh, Kees. 1997. *The Arabic language*. Edinburgh: Edinburgh University Press. [Especially ch. 12 "Diglossia and bilingualism", and ch. 13 "Arabic as a minority language", with references.]

Wilmsen, David. 1996. "Codeswitching, code-mixing, and borrowing in the spoken Arabic of a theatrical community in Cairo". Eid and Parkinson (1996:69–92).

Youssi, Abderrahim, Fouzia Benjelloun, Mohamed Dahbi, and Zakia Iraqui-Sinaceur (eds.). 2002. *Aspects of the dialects of Arabic today: Proceedings of the 4th Conference of the International Arabic Dialectology Association (AIDA)*. Rabat: Amapatril.

GUNVOR MEJDELL (University of Oslo)

Cognitive Linguistics

The term 'cognition' and its related adjective 'cognitive' are used in different ways by different disciplines. In its narrow sense 'cognition' is essentially synonymous with consciousness or awareness. In a broad sense, it refers to the ability of the brain/mind to perceive, store, process, and use information gathered by the sensory receptors. For a given process to be described as cognitive, it should be comprised of an orderly set of actions implemented in order to achieve a goal. Thus natural languages rank high among the cognitive activities that the human brain/mind undertakes, and are the focus of much research in various disciplines such as linguistics, psychology, neuro-psychology, and neurophysiology, with the view of understanding the connection between language processing and the structure of the brain/mind.

1. COGNITIVE LINGUISTICS: AN OVERVIEW

Cognitive linguistics is a general approach to the study of language emphasizing the internal mental processes involved in language production and language comprehension (Lakoff 1987, 1996; Lakoff and Turner 1989; Lakoff and Johnson 1980, 1999). The guiding principle underlying this approach is that linguistic theory and methodology must be constrained by and consistent with the empirical facts known about the brain/mind. This approach emerged as a reaction against the generative paradigm, which defined language as a system of arbitrary symbols and viewed the task of the linguist as mainly developing a set of formal rules that can provide an explanation of language (Chomsky 1966).

The crucial feature that distinguishes cognitive linguistics from generative grammar relates to the way meaning is viewed in the two theories. In the generative model, the structure of linguistic expression is determined by a formal rule system that is largely independent of meaning. Accordingly, the two sentences (1) *John offered a ring to Mary* and (2) *John offered Mary a ring* are thought to express the same meaning, their syntactic (structural) differences having no impact on semantics since they are hypothesized to derive from the same underlying form.

In contrast, cognitive linguists take linguistic structure to be a mirror of cognition in the sense that a particular linguistic expression is associated with a particular way of conceptualizing a given situation. On this view, the difference between sentences (1) and (2) above is not only one of form (or structure) but also, and critically, one of substance. This is because in some cases only one of the two constructions above is natural. For example, while the sentence (3) *John gave the fence a new coat of paint* is natural, it would be odd to say (4) *John gave a new coat of paint to the fence*. These differences suggest that the two constructions illustrated in (1) and (2) involve different ways of construing the same situation, and that in certain cases only one mode of *construal* is appropriate (Lee 1989, 2001; Ungerer and Schmid 1996). A number of mutually related factors are involved in the choice among alternative construals. Among these are perspective, foregrounding, metaphor, and frame.

'Perspective' refers to the viewpoint according to which a particular situation is construed. For example, sentences (5) *The path falls steeply into the valley* and (6) *The path climbs steeply out of the valley* describe the same scene; however, they do not express identical meaning (Herskovits 1986; Lee 2001). The difference underlying the two sentences is one of perspective: sentence (5) is construed from the point of view of someone looking down into the valley, while sentence (6) is construed from the viewpoint of someone looking up from the valley floor. Importantly, however, the actual position of the speaker in cases such as these is irrelevant. Both sentences in fact may be produced by someone looking at a painting, viewing the scene sideways-on.

'Foregrounding' relates to the relative prominence of the various components of a situation (→ grounding). Compare the sentence (7) *I am standing on the street* with (8) *I am standing in the street*. Clearly the two sentences describe the same situation, but they highlight different aspects of it. In (7) the street is conceptualized as a roadway, and therefore as a supporting surface, while in (8) the buildings on either side are taken into account and the street is conceptualized as a container (Taylor 1995; Lee 2001).

'Metaphor' is essentially a device that involves conceptualizing one domain of experience in terms of another. This implies that for any given metaphor, a 'source domain' and a 'target domain' can be identified. For example, in sentences like (9) *He is a really cold person* and (10) *She gave him a warm welcome*, the source domain is the sense of touch, and the target domain is the more abstract concept of intimacy (Lakoff and Johnson 1980; Langacker 1988; Lee 2001). Metaphor is intimately linked to the notion of construal by virtue of the fact that different ways of thinking about or construing a phenomenon are associated with different metaphors. For example, the concept of intimacy may be construed in terms of heat as in (11) *He is such a cold person* and (12) *He has a very cool manner*, or distance as in (13) *I felt really close to him* and (14) *I found his manner rather distant*.

The final factor to bear on the way a particular situation or event is construed is 'frame' (Fillmore 1985). This refers to the background knowledge a speaker has about a word of his/her language. Consider the aspects of a situation which would be described by the English verb *to buy*, for example. Initially, a person A has some money and another person B owns some goods of which A wants to gain possession. Then, A gives B a certain amount of money and B surrenders the goods. The final state is that A owns the goods, and B owns the money. Thus, one can say that the action category of the verb *to buy* includes a reference to at least four other categories: a buyer, a seller, goods, and money; these make up the frame of the verb *to buy*.

Perspective, foregrounding, metaphor, and frame are not an arbitrary set of concepts. On the contrary, they interact with each other in intricate ways and have significant implications for our understanding of the nature of communication. They suggest that meaning is not a property of the utterance itself, but a product of the interaction between an utterance and a human being's background knowledge. More generally, cognitive linguists believe that mental and linguistic categories are created on the basis of our experience and constrained by the general properties of cognitive system.

Cognitive linguistics is not a homogeneous field of research; rather it involves a number of different views, perhaps the most familiar of which are the Prominence view, the Attentional view, and the Experiental view.

1.1 *The Prominence view*

According to this view an expression gains its meaning by imposing a 'profile' on a 'base'. The base is the underlying set of relevant cognitive domains that is required or evoked in comprehending a given expression. The profile, on the other hand, is the highlighted structure within the base that the expression conceptually designates. For example, the words *parent* and *child* evoke the same base, that of family, but they profile different facets of it. By using such concepts as base and profile, the prominence view explains why, when we look at an object in our environment, we single it out as a perceptually prominent figure standing out from the ground. This approach has been applied to the study of grammatical relations in language (Langacker 1987, 1990, 1991; Brugman 1990; Casad 1993; Linder 1982).

1.2 *The Attentional view*

This view is predicated on the basic properties of the human attentional system, which is relatively limited in resources in the sense that humans cannot attend to all facets of a scene at the same time. The brain may in principle be sensitive to all facets of a given scene but cannot allocate equal attentional resources to all its aspects. Only a sub-part of the scene at hand will be the locus of attention and thus make up the 'figure' of the scene (Fillmore 1985; Talmy 1988, 1991, 1996). The remaining aspects of it will not be focused on; they make up the 'ground' of the scene. Even when the ground recedes to background, it continues to be the point of reference for the figure. A classic example of the figure/ground concept is the Gestalt *Vases/Faces* picture, which is perceived as two vases or two faces depending on what the perceiver relegates to the ground. More generally, this approach assumes that what we actually express reflects which parts of an event automatically attract our attention, or which parts of an event we decide to focus on.

1.3 *The Experiental view*

This view pursues a more practical and empirical description of meaning (Rosch 1977, 1978; Lakoff and Johnson 1980). Experientalists take issue with the objectivist view which depend on truth-values, and take meaning to be a correspondence between an expression and the real, or a possible, world. They also reject the subjectivist view of meaning as unconstrained. Instead, they claim that meanings are relatively commensurable from speaker to speaker in a given speech community because members of the same community will share many experiences of the same type. In other words, meaning is grounded in experience, and is best studied by focusing on the way language users communicate with and understand each other.

2. Arabic cognitive linguistics

Contemporary work on Arabic within mainstream cognitive linguistics is scarce. Among the few exceptions to this general trend is perhaps the work of Maalej (1999a, 1999b, 2000, 2002, 2003, 2004), which focuses on the investigation of metaphor in Tunisian Arabic and Modern Standard Arabic, and compares metaphoric concepts across languages. Following Lakoff and Johnson (1980), Maalej (1999a) explores structural, orientational, and ontological metaphors in Tunisian Arabic. Structural metaphor refers to cases where one concept is structured in terms of another as illustrated in (15) *John demolished Tom's argument.* Orientational metaphor pertains to cases where a whole system of concepts is organized in terms of another as in (16) *Try to pack more thoughts in fewer* *words*. Ontological metaphor encompasses cases where an event, activity, emotion, or idea is conceptualized as an entity or substance in order to make it easier to grasp and understand as illustrated in (17) *We need to combat inflation*. Maalej (1999a) identifies a number of domains used by Tunisian Arabic speakers to build structural and ontological metaphors. According to this author, Tunisian Arabic speakers use a source domain such as that of argument and quarrel, and map it onto a target domain such as that of war or sports. The choice between alternative target domains will depend on the emotional state of the speaker and the gravity of the situation at hand. For example, in a situation where one speaker is simply teasing another by refusing to shake their hand to humiliate them in public, one might say (18) *jbid bīh* 'he side-stepped him'. By contrast, in a more serious situation involving two parties nearly coming to grips with each other, one might say (19) *klālu galbu* lit. 'he ate the heart of him' meaning 'he assailed him with attacks'. According to Maalej (1999), this shows that social discourse may be structured around the conceptual metaphor of ARGUMENT IS SPORT or ARGUMENT IS WAR. Note that in English the ARGUMENT IS WAR schema is the more common one as illustrated by (19) *I demolished his argument*, and (20) *If you use that strategy, he will wipe you out*.

Another domain identified by Maalej (1999a) as giving rise to abundant use of conceptual metaphor is 'time.' Tunisian Arabic speakers very often conceptualize 'time' in terms of 'enmity' as in (21) *lwaqt ma yirḥam* 'time is merciless' or (22) *lwaqt ġaddār* 'time is treacherous'. The process of using 'time' as a source domain and mapping it onto the target domain of 'enmity' seems to have an experiential base in the moral system of the Tunisian Arabic speaker, who is constantly taught not to trust 'time' since in hard times there will be nobody out there to give one help, and his/her enemy might strike (Maalej 1999a).

Focusing on conceptual metaphors in Tunisian Arabic and the problems inherent in attempting to translate into other languages, and in particular into English, Maalej (1999b) remarks that structural metaphors lend themselves more easily to literal translation than ontological metaphors. This is because structural metaphors involve mappings where a source domain is transparently paired with a target domain, whereas ontological metaphors may at times involve culture-specific items which make the mapping of the source domain onto the target domain rather idiosyncratic and difficult to grasp if rendered literally into a different language.

3. The rhetorical Arabic tradition

Metaphor, which is almost the *raison d'être* of cognitive linguistics, lies at the heart of the Arabic rhetorical tradition as developed by al-Jurjānī in the 11th century. In his two major books *ʾAsrār al-balāġa* 'The secrets of eloquence' and *Dalaʾil al-ʾiʿjāz* 'The proofs of inimitability', al-Jurjānī developed his theories of *naḍm* 'construction' and → *istiʿāra* 'metaphor'.

In al-Jurjānī's view, metaphor or *istiʿāra* has three defining features. First, it is based on a fusion of two entities, a *mustaʿār minhu* 'a borrowed from' and a *mustaʿār lahu* 'a borrowed for', and this fusion is predicated on the perceived similarities underlying the attributes or characteristics of the two entities. Second, metaphor relies on the (partial) attribution of a dominant trait (or a set of dominant traits) of the *mustaʿār minhu* to the *mustaʿār lahu*. Third, metaphor ought to result in the creation of a double unit based on the interaction between the *mustaʿār minhu* and the *mustaʿār lahu*. To take an example, a sentence like (22) *I saw a lion*, which may be interpreted literally as referring to an event where a real *lion* has been seen by a speaker, may well in the appropriate context be interpreted metaphorically. Thus, the word *lion* may be taken for a *mustaʿār minhu* referring not to a real *lion*, but to a *man* with some qualities or attributes in common with the *lion*. For this metaphorical reading of the sentence to be possible though, the hearer or reader must go through a number of fundamental processes. Specifically, the *mustaʿār minhu*, that is the *lion*, must be fused in the imagination of the hearer or reader with the *mustaʿār lahu, the man*. Stated differently, the *man* needs to be imagined as part of the *lion*'s species. This fusion between the *man* and the *lion* becomes possible by virtue of the subset of shared attributes between them such as the *daring*, *courage*, *power*, and *attacking force*. Critically, however, the two have a number of other features that they do not share such as the shape of the lion's head, or the size and movement of its tail. For metaphor to be appropriately grasped, the shared attributes between the *man* and the *lion* must be conjured up, while the non-shared attributes must be overlooked. It is this process of activating the shared and relevant while inhibiting the non-shared and irrelevant that underlies the dual nature of metaphor in al-Jurjānī's view.

Istiʿāra as figurative speech that relies on similarity between different entities is fundamentally different from other types of figurative speech that rely on contiguity in space or time between different entities. In particular, similarity and contiguity in al-Jurjānī's view involve different mental activities. Thus if we compare a sentence like (23) *lahu ʿindī yadun* 'he has done me a favor' (lit. 'he has a hand on me'), where the word *yadun* 'hand' is used for 'beneficence,' with sentence (22) above, it is clear that the two sentences do involve different mechanisms. For example, by using the word *yadun* 'hand' in (23), the speaker does not want to attribute 'the quality of being a hand' to the concept of 'beneficence', whereas in (22) the speaker has every intention to confer 'the qualities of being a lion' on the person he describes.

Al-Jurjānī identifies a few other types of figurative speech, such as metonymy and simile, and compares them with metaphor as he conceives of it. Clearly there are some affinities between his treatment of metaphor, and the treatment of very similar phenomena by current cognitive linguists. For example, the *mustaʿār minhu* would certainly be a 'source domain' in cognitive linguistic parlance, and the *mustaʿār lahu* a 'target domain'. Al-Jurjānī's focus on metaphor as deriving from and depending on perceived similarity as opposed to spatio-temporal contiguity between *mustaʿār minhu* and *mustaʿār lahu* has a lot in common with the views of Lakoff and Johnson (1980). There are also differences between the two approaches. Space limitation does not permit a comprehensive discussion of all the interesting aspects of al-Jurjānī's theory, let alone a comparison with contemporary Cognitive Linguistic theory. The reader is encouraged to consult the excellent analysis of al-Jurjānī's *ʾAsrār al-balāġa* and *Dalāʾil al-ʾiʿjāz* by Abu Deeb (1979). More recent work dealing with various aspects of al-Jurjānī's theory from different perspectives includes Leezenberg (2001) and Kamel (forthcoming).

4. ARABIC COGNITIVE PSYCHOLOGY

Cognitive psychology is the psychological science which studies cognition, the mental processes that are hypothesized to underlie behavior. This covers a broad range of research domains, examining questions about the workings of memory, attention, perception, knowledge representation, reasoning, creativity, problem solving, and language. Cognitive psychology research into Arabic is only beginning. The absence of the appropriate statistical tools, such as frequency of occurrence tables and lexical data bases, has made the task of exploring Arabic from a cognitive psychology perspective somewhat daunting. The few existing studies to date focus on the morphophonology of Arabic and the use of 'priming' (Boudelaa and Marslen-Wilson 2000, 2001, 2004a, 20004b; Mimouni, Kehayia, and Jarema 1998), or patient data (Prunet, Beland, and Idrissi 2001; Idrissi and Kehayia 2004). In its most common variant, 'priming' refers to how a single-word context (called the prime) can affect the speed with which a subsequent word (called the target) is processed (Marslen-Wilson a.o. 1994; Forster 1999). The logic underlying priming is that if the mental representations of the prime and target are inter-connected or overlap in some way, activating the representation of the prime should automatically activate that of the target, hence the speeded response and the lower error rate on the target (Forster 1999). Research using this technique has revealed that words sharing a root such as *maktabatun/kitābun* 'library/book', an etymon *mubtallun/wābilun* 'wet/downpour,' a word pattern *kātaba/qābala* 'to correspond with/ to meet' or even a CV-skeleton (the abstract sequence of consonants and vowels), such as CVVCVC for *kūfiʾ/lāzam* 'to be rewarded/to adhere to', prime each other reliably (Boudelaa and Marslen-Wilson 2000, 2001, 2004a, 2004b). This suggests that these various linguistic units function as cognitive units that structure the organization of the Arabic mental lexicon. Pathological data also reveal that in some patients the consonants of the root or etymon may be metathesized while consonants belonging to the word pattern are not. For example, Prunet a.o. (2001) and Idrissi and Keheya (2004) describe an Arabic-speaking aphasic patient who when prompted with a form like *maktabatun* 'library' would produce a form like

**mabkatatun* or **matbakatun*, but would hardly ever output a form like **tabkamatun*, where a consonant of the root is metathesized with a consonant of the word pattern.

Arabic cognitive linguistics and Arabic cognitive psychology still have a long way to go if they are to produce full-fledged models of how Arabic is acquired, represented, and accessed in the brain/mind of its native speakers.

Bibliographical references

Abu Deeb, Kamal. 1979. *Al-Jurjānī's theory of poetic imagery*. Warminster, Wilts.: Aris and Phillips.

Boudelaa, Sami and William D. Marslen-Wilson. 2000. "Non-concatenative morphemes in language processing: Evidence from Modern Standard Arabic". *Proceedings of spoken word access processes*, ed. Anne Cutler, James McQueen, and Rian Zondervan, 23–26. Nijmegen: Max Planck Institute.

——. 2001. "Morphological units in the Arabic mental lexicon". *Cognition* 81.65–92.

——. 2004a. "Allomorphic variation in Arabic: Implications for lexical processing and representation". *Brain and Language* 90.106–116.

——. 2004b. "Abstract morphemes and lexical representation: The CV-Skeleton in Arabic". *Cognition* 92.271–303.

Brugman, Claudia. 1990. "What is the Invariance Hypothesis?". *Cognitive Linguistics* 1.257–266.

Casad, Eugene. 1993. "Locations, paths and the Cora verb". *Conceptualizations and mental processing in language*, ed. Richard Geiger and Brygida Rudzka-Ostyn, 593–645. Berlin and New York: Mouton de Gruyter.

Chomsky, Noam. 1966. *Cartesian linguistics*. New York: Harper and Row.

Fillmore, Charles, C. 1985. "Frames and the semantics of understanding". *Quaderni di Semantica* 6.222–254.

Forster, Kenneth, I. 1999. "The microgenesis of priming effects in lexical access". *Brain and Language* 68.5–15.

Herskovits, Annette. 1986. *Language and spatial cognition: An interdisciplinary study of the prepositions in English*. Cambridge: Cambridge University Press.

Idrissi, Ali and Eva Kehayia. 2004. "Morphological units in the Arabic mental lexicon: Evidence from an individual with deep dyslexia". *Brain and Language* 90.183–197.

Kamel, Salwa. Forthcoming. "The textual component in classical Arabic: Investigating information structure". *Perspectives on Arabic linguistics*, forthcoming.

Lakoff, George. 1987. *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.

——. 1996. *Moral politics*. Chicago: University of Chicago Press.

—— and Mark Johnson. 1980. *Metaphors we live by*. Chicago and London: University of Chicago Press.

——. 1999. *Philosophy in the flesh: The embodied mind and its challenge to Western thought*. New York: Basic Books.

Lakoff, George and Mark Turner. 1989. *More than cool reason: A field guide to poetic metaphor*. Chicago: University of Chicago Press.

Langacker, Ronald. W. 1987. *Foundations of cognitive grammar. I. Theoretical prerequisites*. Stanford, Cal.: Stanford University Press.

——. 1988. "An overview of cognitive grammar". Rudzka-Ostyn (1988:91–125).

——. 1991a. *Concept, image, and symbol: The cognitive basis of grammar*. Berlin: Mouton de Gruyter.

——. 1991b. *Foundations of cognitive grammar. II. Descriptive application*. Stanford, Cal.: Stanford University Press.

Lee, David. A. 1992. *Competing discourses: Perspective and ideology in language*. London: Longman.

——. 2001. *Cognitive linguistics*. Oxford: Oxford University Press.

Leezenberg, Michiel. 2001. *Contexts of metaphor*. Amsterdam: Elsevier Science.

Linder, Susan J. 1982. *A lexico-semantic analysis of English verb particle constructions with* out *and* up. Ph.D. diss., University of California, San Diego.

Maalej, Zouhair. 1999a. "Metaphoric discourse in the age of cognitive linguistics, with special reference to Tunisian Arabic". *Journal of Literary Semantics* 28.189–206.

——. 1999b. "Translating metaphor between unrelated cultures: A cognitive view." Paper presented to The Third International Conference on Researching and Applying Metaphor: Metaphor across Languages – Translation and Intercultural Communication, Including Literature. Tilburg.

——. 2000. "Passives in Modern Standard and Tunisian Arabic". *Proceedings of the third international conference of the Association Internationale de Dialectologie Arabe*, ed. Manwel Mifsud, 21–26. Malta: Salesian Press.

——. 2002. "The image-schematic anchorage of the preposition *fi* (in): The logic of containment in Tunisian Arabic". *Textual secrets: The message of the medium. Proceedings of the 21st PALA conference*, ed. Szilvia Csábi and Judit Zerkowitz, 227–236. Budapest: Eötvös Loránd University.

——. 2003. "Conceptual metaphor as persuasion, with special reference to consumer advertising: A cognitive semantics account". *Pragmatic perspectives on persuasion*, ed. Mounir Triki and Akila Sellami-Baklouti, 121–147. Sfax: Faculty of Letters and Humanities.

——. 2004. "Figurative language in anger expressions in Tunisian Arabic: An extended view of embodiment". *Metaphor and Symbol* 19.51–75.

Marslen-Wilson, William D., Lorraine K. Tyler, Rachelle Waksler, and Lianne Older. 1994. "Morphology and meaning in the English mental lexicon". *Psychological Review* 101.3–33.

Mimouni, Zohra, Eva Kehayia, and Gonia Jarema. 1998. "The mental representation of singular and plural nouns in Algerian Arabic as revealed through auditory priming in agrammatic aphasic patients". *Brain and Language* 61.63–87.

Prunet, Jean-François, Reneé Béland, and Ali Idrissi. 2001. "The mental representation of Semitic words". *Linguistic Inquiry* 31.609–648.

Rosch, Eleanor. 1977. "Human categorisation". *Studies in cross-cultural psychology*, I, ed. Neil Warren, 1–49. London: Academic Press.

——. 1978. "Principles of categorization". *Cognition and categorization*, ed. Eleanor Rosch and Barbara B. Lloyd, 27–48. Hillsdale, N.J.: L. Erlbaum Associates.

Taylor, John R. 1995. *Linguistic categorization: Prototypes in linguistic theory*. 2nd ed. Oxford: Clarendon Press.
Ungerer, Friedrich and Hans-Jörg Schmid. 1996. *An introduction to cognitive linguistics*. London: Longman.

Sami Boudelaa
(MRC-Cognition and Brain Sciences Unit)

Coherence

Coherence refers to the grammatical and semantic interconnectedness between the elements that realize a discourse (discourse is used here to subsume any written or spoken communicative occurrences). Coherence is separate from → cohesion and specifically designates the semantic meanings that hold the discourse together (its flow). Coherence is the sequence of propositions (thematic developments) that realize the web of semantic relations, not the formal meanings. When a discourse seems to lack coherence, receivers resort to inference in order to process the information communicated in order to be able to understand and interpret it. While cohesion is generally indicated by actual markers in the discourse, coherence covers the semantic relationships underlying it, relating discourse propositions to each other and to the overall (macro-) communicative-functional purpose of the discourse.

In the Arabic linguistic tradition, interest in coherence stemmed from the desire to explain the unity (coherence) of the *Qurʾān*. Accordingly, the first generations of Muslim scholars dedicated themselves wholly to this, from which arose the sciences of reading (*ʿilm al-qirāʾāt*), exegesis (*tafsīr*), and jurisprudence (*fiqh*), the basic fields that led to the emergence of a plethora of disciplines, mainly linguistic (grammar, rhetoric, and other allied fields). Consequently, scholars of the early period of Islam were primarily linguists or exegetes who devoted themselves to the study and analysis of the *Qurʾān* as well as the *Ḥadīṯ*. What prompted this dedication was the need to protect the *Qurʾān* against the claim that it lacked coherence. Even among Muslims, serious differences in the interpretation of the *Qurʾān* gave rise to religious sectarianism, with each sect adopting a particular approach to the interpretation of the *Qurʾān* and its overall coherence. One view stipulated that with a sound understanding of the Arabic language one can appreciate coherence in the *Qurʾān*, which is certainly not a haphazard collection of *ʾāya*s and *sūra*s. By taking into consideration the three constituents of order, proportion, and unity, a single and coherent interpretation of the *Qurʾān* was possible (cf. ʿAbd al-Muṭṭalib 1994). What further intrigued and stimulated early Muslim scholars to focus on the *Qurʾān* was that the Book itself forcefully proclaims its inimitability, challenging all (Arabs and non-Arabs alike) to compose even a single *sūra* that would surpass or equal it (→ *ʾiʿjāz al-Qurʾān*).

Like Western linguists, Arab linguists seem to have accepted an apparent confusion as to the actual distinction between cohesion and coherence. In some cases, the notion of cohesion seems to be taken even to subsume that of coherence. But in general, and despite its inclusion as a discourse-centered feature, coherence is taken to relate to the way a discourse is mutually constructed by communicants (producers and receivers). The processing of information in a discourse by the receiver generally involves inferring, and because of its importance in reception, coherence becomes the product of the discourse receiver's evaluation of the information, and ultimately involved with schema theories and other matters related to reception rather than production, as is the case with cohesion (cf. Beaugrande and Dressler 1981).

In Arabic, linguists have found coherence very relevant to their analysis and description of how language elements hang together in discourse. In particular, within the field of rhetoric, medieval linguists considered coherence (*insijām/tamāsuk*) to be the provider of the link between the formalistic and the more contextualized aspects of discourse. Given this dimension, the culture of the user (producer and receiver alike) was seen as vital in processing information in one way or another. In this way, these linguists/rhetoricians understood that discourse-centered features were not enough in dealing with discourse, and that user-centered features were needed as well, such as intentionality of the producer, acceptability of the receiver, and informativity of the discourse itself.

Generally speaking, Arab linguists have treated issues that are outside the internal fabric of discourse (cohesion) under the umbrella of language acts or communicative events, particularly within rhetoric (*ʿilm al-balāġa*, cf. ʿAbd al-Muṭṭalib 1994). To account for coherence, they considered implicatures and illocutionary forces as *ʾaġrāḍ* 'purposes'. In this, they looked at discourse literal meanings (related to cohesion) as meanings retrievable through syntax (arrangements) of the words together, and implied meanings (contextual meanings) that form coherence. Many scholars dealt with the issue of coherence: Ibn Qutayba, al-Jāḥiẓ, Qudāma ibn Jaʿfar, al-ʿAskarī, al-ʾĀmidī, al-Qāḍī al-Jurjānī, al-Qāḍī ʿAbd al-Jabbār, Ibn Rašīq, and Ibn Sinān al-Xafājī, to name but a few. But the most illuminating and important contribution was that of ʿAbd al-Qāhir al-Jurjānī (d. 471/1078). For all authors, the inimitability (*ʾiʿjāz*) of the *Qurʾān* resided not only in its words (*ʾalfāḏ̣*), but also, and most importantly, in its content as embodied in the most eloquent, expressive, and coherent way. Words and meanings were thought to be incapable of achieving great eloquence unless they were linked by a third element called the stringing or joining process (cf. Murād 1983).

The duality of word and meaning (form and content, manner and matter) continued to dominate the thinking of most theologians and writers on rhetoric until al-Jurjānī developed his theory of *naḏ̣m* 'composition, construction', which represented the culmination of an intensive inquiry into the inimitability and coherence of the *Qurʾān* and the beauty and superiority of Arabic literature. The theory of *naḏ̣m* (primarily expounded by al-Jurjānī in his *ʾAsrār al-balāġa* and *Dalāʾil al-ʾiʿjāz*) became the firm basis of the science of meanings (*ʿilm al-maʿānī* 'semantics plus pragmatics'), which has continued to be studied as an independent branch of rhetoric to the present day.

Postulating that language is a system of relations, al-Jurjānī insisted that the only way for words to mean anything is for them to be entered into sets of relations that are mutually constructed according to certain principles of linguistic (grammar, cohesion) and non-linguistic factors (related to discourse users), and then and only then coherence is said to have been achieved. Since the function of discourse is the expression of the human inner state, and the revelation and communication of the hidden contents of human psyche (*nafs*), the realization of this function is achieved only when words are composed in a particular way so that they mean what the producer intended and what the receiver expects and infers from the discourse.

An essential feature of a discourse is that it is conceived of as a whole and as a closely-knit unit. The structural relations involved in expressing its purpose comprise a single interrelated formulation. Each element of this formulation interacts with the other elements, on the basis that its position in the structure is determined by the nature of its relations to the whole. Also, each element determines the position and expressive power of the other parts of this formulation (cf. *ʾAsrār*).

In the analysis of the determinant factors in the realization of the relations Subject-Verb-Object and Subject-Predicate, for example, al-Jurjānī's starting point is the belief that the relations between the units of meaning are determined by the inner state (*nafs*) of the producer. This is implied in his postulate that a statement (*xabar*) and all the meanings of discourse (*maʿānī l-kalām*) are meanings that the producer composes in his psyche, considers in his intellect, contemplates upon, and whispers to his heart (*yunājī bihā qalbahu*), and then reflects upon reactions to them (*yarjiʿu fīhā ʾilayhi*) (cf. ʿAbū Zayd 1996).

This process, however, is determined not by the producer's *nafs* alone, but also by the relationship between the producer and the receiver. The former's awareness of the latter's circumstances and the assumptions made about his possible reactions influence the structure of the psychological and emotional experience of the producer. The syntactic structure *X likes Y*, for example, should not have a fixed structure determined by word-order (who does the liking and who is the liked): more complex elements are involved here, for X and Y are not abstract entities bearing no relations to the discourse users (producer and receiver). They relate to the situational context of the experience and the inter-action between the different elements of the communicative exchange. Thus, the interaction between producer and receiver plays a decisive role in shaping the structure of the former's inner state and, simultaneously, the structure of the expression and its *naḏ̣m* that yields coherence.

Al-Jurjānī distinguishes between meaning (→ *maʿnā*) and purpose (*ġaraḍ*). Meaning is realized through the interaction of contextual elements, and meanings differ even if the purpose is the same. The producer and receiver are given prominent roles in establishing discourse coherence, intentionality and acceptability (grouped as *maʿānī nafsiyya*) are given primacy over *naḏm* in discourse, and word order and even *naḏm* follow meanings. Accordingly, words become mere signs for predetermined meanings and stand for what meanings users of discourse attribute to them. The following points summarize the criteria for coherence:

i. words exist only for their meanings
ii. words serve meanings
iii. words cannot precede meanings in the *taṣaw-wur nafsī*
iv. naming comes after the named is conceptualized and determined by users
v. sign and sign use are linked to *tadāwul* 'pragmatics'
vi. language is realized by discourse in actual contextual settings, involving users
vii. *maqṣadiyya* 'intentionality' is a principle of use (pragmatics)
viii. discourse must be sufficient to meet the requirements of the meanings (quantity and quality)
ix. it should neither exceed its purpose nor fall short of realizing it.

Given its dimensions, coherence is at the heart of *balāġa* 'rhetoric' and is realized in discourse when there is harmony (*talāʾum*) between words and their meanings across the discourse, based primarily on the purpose (goal, intentionality, and acceptability) of the users (producer and receiver). As such, coherence is a branch of *ʿilm al-maʿānī*, whereas eloquence (*faṣāḥa*) is part of → *lafḏ* 'words' and their associated grammar. Language aspects that are usually considered to be components of coherence and seen as important in influencing and/or establishing it include:

i. *faṣl* 'disjunction' and *waṣl* 'conjunction'
ii. *musnad* 'theme' and *musnad ʾilayhi* 'rheme'
iii. deletion of rheme
iv. fronting of rheme
v. deletion
vi. metaphors (metaphors are not understood by the sum total of the denotative meanings of their constituent words, but through the assignment of other meanings, related to intentionality, acceptability, and context of use)
vii. language is linked to thinking
viii. thoughts activate words
ix. meanings come before words
x. words are signs for thoughts (stand for them)

Words merely assist in linking language to thinking in a closely-knit fashion, thus realizing discourse coherence. Here, the concept of *qaṣd* 'purpose' is vital for coherence. It links the meanings of discourse with the producer. Coherence is the product of a triadic network of purpose (*qaṣd*), discourse (*kalām*), and *taʾwīl* 'interpretation/explication'. Another important concept invoked was the *maʿnā nafsī*, the meaning that resides in the producer and is expected by the receiver. Individual words are put together not for their meanings in themselves, but to be strung together in order to communicate the meanings that derive from this stringing. Although they are not mentioned explicitly, one can easily feel the Gricean maxims of quality, quantity, manner, and relevance at play in the treatment of coherence. In its own fashion, al-Jurjānī's theory of *naḏm* deals with all these issues, albeit his motivation was politico-religious.

Modern approaches to coherence in Arabic have either worked on revisiting the classics, particularly the views of al-Jurjānī, or have simply imported and adapted theories of coherence developed within British, Continental European, American, or Russian contexts. The terminology may be different, but the approaches to coherence have primarily focused on discourse-user features: intentionality and acceptability, precisely what al-Jurjānī expounded.

Although assisted by discourse-centered features (cohesion), discourse-user features are not necessarily retrievable from an analysis of the discourse, because they are functions of the communicative interaction between producer and receiver through discourse. As such, these discourse-user features refer to the choices made by the producer and the inferences made by the receiver when they communicate through discourse.

Intentionality subsumes all notions of producer intention (purpose) and is seen as a function of discourse management strategies that guide choices contextualized in terms of the overall discourse plan (communicative purpose) and the situation (context of use). Acceptability is viewed as the attitude on the part of the receiver that a discourse is assumed to be a coherent (and cohesive) communicative event. Acceptability involves the receiver's expectations and decisions about the appropriateness and, above all, the coherence of discourse.

Both intentionality and acceptability subsume knowledge of the real world, the constraints of the language system in use, the arrangement of given/new information in the elements that realize the discourse, and considerations of the discourse types (→ text linguistics).

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Jurjānī, *Dalāʾil* = ʾAbū Bakr ʿAbd al-Qāhir ibn ʿAbd ar-Raḥmān al-Jurjānī, *Dalāʾil al-ʾiʿjāz*. Ed. Yāsīn al-ʾAyyūbī. Beirut: al-Maktaba al-ʿAṣriyya, 2000.
——, *ʾAsrār* = ʾAbū Bakr ʿAbd al-Qāhir ibn ʿAbd ar-Raḥmān al-Jurjānī, *ʾAsrār al-balāġa*. Ed. Helmut Ritter. Istanbul: Wizārat al-Maʿārif, 1953. 2nd ed., Baghdad: Maktabat al-Muṯannā, 1989.

SECONDARY SOURCES

ʿAbd al-Muṭṭalib, Muḥammad. 1994. *al-Balāġa wa l-ʾuslūbiyya*. Cairo: aš-Šarika al-Miṣriyya al-ʿĀlamiyya li-n-Našr. London: Longman.
ʾAbū Zayd, Naṣr Ḥāmid. 1996. *ʾIškāliyyat al-qirāʾa wa-ʾāliyyāt at-taʾwīl*. Casablanca: al-Markaz aṯ-Ṯaqāfī al-ʿArabī.
De Beaugrande, Robert and Wolfgang Dressler. 1981. *Introduction to text linguistics*. London: Longman.
Murād, Walīd Muḥammad. 1983. *Naḏariyyat an-naḏm wa-qīmatuhā l-ʿilmiyya fī d-dirāsāt al-luġawiyya ʿinda ʿAbd al-Qāhir al-Jurjānī*. Damascus: Dār al-Fikr.

SAID FAIQ (American University of Sharjah)

Cohesion

Cohesion is a property of oral and written text or discourse. It refers to relations that exist between (adjacent and) structurally independent clauses or sentences. Cohesion, thus, characterizes a stretch of text that manifests a non-structural relationship. The sentence in (1) does not denote a cohesive relation because of the grammatical dependency that exists between its major component parts, its main and subordinate clauses.

(1) *fawra duxūli-hi l-maqhā badaʾa yudaxxinu sījāra*
'As soon as he entered the café, he started smoking a cigarette'

> The notion of cohesion refers to the explicit signaling in text or discourse of various kinds of inter-clausal and inter-sentential relations. Cohesive relations are expressed by overt devices, markers, or 'ties' that signal surface-structure connectedness with other, mostly preceding, clauses or sentences. As such, cohesion "expresses the continuity that exists between one part of the text and another" (Halliday and Hasan 1976:299).

Cohesion is one of the most important criteria of textuality (de Beaugrande and Dressler 1981), or rather of what forms a text (Gutwinski 1976; Halliday and Hasan 1976; van Dijk 1977). The literature on cohesion discusses several linguistic devices – pertaining to different grammatical categories – that establish ties between sentences and account for textual cohesion in spoken and written text or discourse. It also distinguishes between short-range and long-range cohesive devices (see de Beaugrande and Dressler 1981). Sources of cohesion include reference, substitution, ellipsis, conjunctions, demonstratives, and lexical devices.

The importance of cohesion lies in establishing specific relations among clauses or sentences, and consequently in disambiguating and making explicit the meaning intended by a writer. As such, cohesive devices encode relations that already exist in the underlying semantic structure, i.e. propositions of the text. (A proposition is the semantic equivalent of a clause.) Junctions, for example, show how relations are recovered (de Beaugrande and Dressler 1981:74).

A distinction has been made in the literature between cohesion and → coherence, another important notion that received a great deal of attention in textlinguistic studies. Although the two notions are sometimes conflated, and the presence of cohesive devices is mistaken for textual coherence, basic differences exist between these two text properties as regards the level at which they realize relationships in text or discourse. While cohesion is a property of surface structure connectivity, coherence has to do with the underlying level of semantic relations and does not necessarily depend on the presence in text of grammatical and/or lexical devices that explicitly signal semantic relations. In addition, coherence involves pragmatic factors and processes that lie beyond the text itself and in the context of production and comprehension (as, for example, world knowledge and knowledge of specific situations).

While a text has coherence by virtue of being a text, cohesive devices are not sufficient for creating a text (i.e., for realizing textual well-formedness and acceptability). The sequence in (2) is incoherent, notwithstanding the presence of markers.

(2) *fī ṣ-ṣayfi l-māḍī zurtu madīnat al*
ʾiskandariyya
al-ʾiskandariyya madīna kabīra
al-mudun al-kubrā tuʿānī min at-talawwuṯ
at-talawwuṯ ḍārr bi-ṣ-ṣiḥḥa

'Last summer I visited the city of Alexandria. Alexandria is a big city. All big cities suffer from pollution. Pollution is hazardous to health'

The second, third, and fourth sentences have lexical items that refer to entities or items mentioned earlier. Though cohesive (in the sense that there is a continuity of reference), the sentences in (2) fail to be coherent because they refer to different worlds or different universes of discourse. In the absence of a global topic, the sequence in (2) would not be coherent and would not form a text. The presence of a global topic (usually a proposition expressed in a title or a headline, for example) makes a reader interpret a sequence of sentences as being well-formed and coherent. Surface connectivity by means of lexical repetition does not in itself make an incoherent sequence coherent.

This means that cohesion does not operate in isolation, but rather in conjunction with the underlying semantic or meaning relations between propositions of the text (i.e. coherence). The presence of these relations is more important for textuality than the presence of devices that realize surface connectivity. In fact, a globally-coherent text does not necessarily imply connectivity between juxtaposed sentences (Hendricks 1976:37). In other words, cohesive devices may be absent in a sequence of sentences, but their absence does not necessarily mean that the sequence as a whole is incoherent.

Primary among cohesive devices are anaphoric and lexical devices. Anaphoric devices include pronominal reference, demonstratives, and substitution. A common cohesive device is anaphoric (i.e. backward) reference to an entity in a preceding sentence. In (3) the second sentence contains anaphoric pronouns that refer to the nouns in the first sentence. These pronouns depend on the antecedent for their interpretation.

(3) *muḥammad ištarā hadiyya li-zawjati-hi.*
hiya fariḥat jiddan bihā

'Muhammad bought a present for his wife.
She was very pleased with it'

Demonstratives also express continuity across sentences; in so doing they serve a cohesive function in oral and written text or discourse. This is apparent in (4), where the demonstrative in the second sentence picks up what has preceded (the noun phrase), making it clear that what follows is a continuation of the subject matter introduced earlier.

(4) *ʾaḥraza l-lāʿib natāʾij bāhira.*
wa-hāḏihi n-natāʾij lam takun mutawaqqaʿa.

'The player achieved spectacular results.
These results were not expected'

Substitution is yet another important means to realize cohesion in a sequence of sentences.

(5) *ištarat šanṭa ḥamrāʾ. ṯumma ištarat wāḥida*
zarqāʾ
'She bought a red bag. Then she bought a
blue one'

The word *wāḥida* substitutes the word *šanṭa* in the first sentence and realizes cohesion between the two sentences.

The most common cohesive relation is realized by lexical means. Sequences of sentences usually contain words that express concepts within the same semantic field and hence realize what is referred to as 'collocational cohesion' (Halliday and Hasan 1976:287). Lexical cohesion includes several forms of repetition, such as repetition of the same word or of a synonymous one (e.g. *ʾasad/layṯ* 'lion'), a part-whole relation (*ʿajalat al-qiyāda/as-sayyāra* 'steering wheel/car'), a super-ordinate (*as-sana aṯ-ṯāniya/ad-dirāsa* 'second year/the study'), and antonyms (*nūr/ḏ̣alām* 'light/darkness').

In addition to referential and lexical types of cohesion, various kinds of → connectives, such as conjunctive conjunctions, serve cohesive functions in text or discourse. Al-Batal (1985, 1990), for example, provides a description of the semantic properties of Arabic connectives, including additive (e.g. *wa-* 'and', *kaḏālika* 'also'), adversative (*ʾinnamā* 'however', *lākinna* 'but'), alternative (*ʾaw* 'or'), causal (*fa-* 'for', *li-ʾanna* 'because'), conclusive (*fa-* 'therefore'), explicative (*ʾay* 'namely'), sequential (*ṯumma* 'then'), simultaneity (*wāw al-ḥāl* 'as'), topic introduction/shifting (*ʾammā . . . fa-* 'as for'). Sarig (1995) presents an analysis of some connectives such as *wa-qad* and *fa-qad* as discourse markers in contemporary written Arabic.

Languages vary in the resources they choose to serve cohesive functions. They also vary in their requirement to employ cohesive devices. Within a language, differences also exist at the level of text-types in preference for, and frequency of, use of certain cohesive devices.

In Arabic, cohesion seems to be a text-creating property, essential for text coherence. What is important in Arabic "is not only the presence of the underlying semantic relationships but also the proper presentation, through connectives, of these relationships" (Al-Batal 1990:253). In other words, signaling underlying coherence seems to be a mandatory condition in Arabic texts (Al-Batal 1990:254). Accordingly, signaling the type of relation that holds between various constituents is assumed to be a prerequisite for text coherence as well as text acceptability in many types of Arabic discourse. In this regard, it is recalled that Arabic is a language that depends greatly on connectives, in particular on conjunctive conjunctions. It has been described as a language of connection or junction (*luġat al-waṣl*, ʾAnīs 1975:327). Conjunction in Arabic, unlike English, "is explicit, performing disjunctive and conjunctive functions" (Sa'adeddin 1987:185).

Since many languages have texts that "often contain sections where writers have not clearly specified their plans and leave ambiguous relationships among ideas" (Meyer 1985:66), cohesive devices provide readers with clues as to the type of relation that holds between various constituents and the nature of the underlying coherence. Ḥassān 1973:213 discusses *ar-rabṭ* as a textual indicator in Arabic (*qarīna maqāliyya lafḏ̣iyya*). For a short review of the approach to connectives in Arabic grammar see Al-Batal (1990).

The absence of clear relationships among ideas becomes apparent in surface expression, particularly in sentence-initial position. The absence or omission of markers in this initial position in Arabic leads readers to perceive what is referred to as "a zero signal when they are expecting an overt one" (Callow 1992:359). This may represent a violation of the receivers' expectations about the ways meaning is to be expressed in accordance with the language and conventions that govern text-type. More importantly, it may lead to problems in interpretation and in determination of text-type acceptability.

The notion of cohesion plays a prominent role in second/foreign language learning. Learners of Arabic face a daunting task in creating cohesive texts in translation (see Khalil 1983). This task includes awareness of shifts that may occur in the process of translation as a result of using cohesive devices that affect the level of explicitness of the target text, making it higher or lower than that of the source text (Blum-Kulka 1986). Examples include resorting to lexical repetition instead of making use of pronominalization. Essay-writing is another area where cohesion is "an important property of writing quality" (Witte and Faigley 1981:202). As language users, learners always need to handle stretches of language longer than the single sentence and to refer back and forth in the texts they produce. Hence, they need to know how to achieve cohesion by making use of the appropriate devices that the language offers. In language acquisition, there is a dire need for material that equips learners with the appropriate mechanisms to ameliorate writing quality in connected prose (see, e.g., Al-Warraki and Hassanein 1994).

Devices in Arabic that have the potential to serve a cohesive function are legion. Many of these devices have been investigated in Arabic dialects such as Lebanese (al-Batal 1994) and Iraqi (Aziz 1988). Devices that serve cohesive functions include adverbials and prepositional phrases (*bi-l-ʾidāfa ʾilā ḏālika* 'in addition', *min nāḥiyatin ʾuxrā* 'on the other hand'), as well as the grammatical categories of tense and aspect. Tense maintenance (or shift), i.e., how events are marked temporally, may increase (or decrease) sequential cohesion in text. Similarly, the use of the auxiliary verb *kāna* 'to be' with an imperfective main verb, such as in *kāna yašrab-u* 'he has been drinking', may be crucial for the proper comprehension and interpretation of a certain sequence. An adequate description of the specific cohesive properties of these devices awaits future research.

Bibliographical references

Al-Batal, Mahmoud. 1985. *The cohesive role of connectives in a modern expository Arabic text*. Ph.D. diss., University of Michigan.

——. 1990. "Connectives as cohesive elements in a modern expository Arabic text". *Perspectives on Arabic linguistics*, II, ed. Mushira Eid and John McCarthy, 234–266. Amsterdam and Philadelphia: J. Benjamins.

——. 1994. “Connectives in Arabic diglossia: The case of Lebanese Arabic”. *Perspectives on Arabic linguistics*, VI, ed. Mushira Eid, Vicente Cantarino, and Keith Walters, 91–119. Amsterdam and Philadelphia: J. Benjamins.

Al-Warraki, Nariman Naili and Ahmed Taher Hassanein. 1994. *Connectors in modern standard Arabic*. Cairo: American University in Cairo Press.

ʾAnīs, ʾIbrāhīm. 1975. *Min ʾasrār al-luġa*. Cairo: Maktabat al-ʾAnglū al-Miṣriyya.

Aziz, Yowell Y. 1988. “Cohesion in spoken Arabic texts”. *Pragmatics, discourse and text: Some systemically-inspired approaches*, ed. Erich H. Steiner and Robert Veltman, 148–157. Norwood, N.J.: Ablex Publishing Corporation.

Beaugrande, Robert de and Wolfgang Dressler. 1981. *Introduction to text linguistics*. London: Longman.

Blum-Kulka, Shoshana. 1986. “Shifts of cohesion and coherence in translation”. *Interlingual and intercultural communication*, ed. Juliane House and Shoshana Blum-Kulka, 17–35. Tübingen: Narr.

Callow, Kathleen. 1992. “Non-realised information: A theory for the accurate and natural translation of meaningful zero”. *Translation and meaning*, II, ed. Barbara Lewandowska-Tomaszczyk and Marcel Thelen, 359–368. Maastricht: Rijkshogeschool Maastricht, Faculty of Translation and Interpreting.

Dijk, Teun A. van. 1977. *Text and context*. London: Longman.

Gutwinski, Waldemar 1976. *Cohesion in literary texts*. The Hague: Mouton.

Halliday, M.A.K. and Ruqaiya Hasan. 1976. *Cohesion in English*. London: Longman.

Ḥassān, Tammām. 1973. *al-Luġa al-ʿarabiyya: Maʿnāhā wa-mabnāhā*. Cairo: al-Hayʾa al-Miṣriyya al-ʿĀmma li-l-Kitāb.

Hendricks, William O. 1976. *Grammars of style and styles of grammar*. Amsterdam: North-Holland.

Khalil, Esam N. 1983. “Cohesion in translated texts”. *Arab Journal of Language Studies* 2.171–179.

Meyer, Bonnie J.F. 1985. “Signaling the structure of text”. *The technology of text*, II, ed. David H. Jonassen, 64–89. Englewood Cliffs, N.J.: Educational Technology Publications.

Sa'adeddin, Mohammed Akram A.M. 1987. “Three problem areas in teaching translating to native Arabic literates”. *Anthropological Linguistics* 29.181–193.

Sarig, Lea. 1995. “Discourse markers in contemporary Arabic”. *Zeitschrift für Arabische Linguistik* 30.7–21.

Witte, Stephen P. and Lester Faigley. 1981. “Coherence, cohesion and writing quality”. *College Composition and Communication* 32.189–204.

Esam N. Khalil
(University of Nijmegen)

Collective

The collective, as a linguistic category, conveys the notion of a group of beings, both animate and inanimate (human beings, animals, and objects), that is to say, a plurality, as well as the notion of a species or genre. Examples: *šajar* ‘trees [a group of]’, or the species of the tree, *baqar* ‘cows’, or the notion of cow. It is important to bear in mind that the collective is not a plural. While the plural refers to a group of beings considered as individualized elements, the collective mostly refers to these beings as a single unit, regardless of their specific features. In that sense, the collective, as it appears in Arabic, may be defined as a singular from the point of view of its form and its general semantic implications. However, the collective retains some of the features of the plural, because it conveys the notion of a group of things or beings. This sharing of the two basic numbers (singular and plural) makes the collective an ambiguous category in its linguistic behavior, especially concerning the agreement between collective and verbs or other nouns.

Arab grammarians, from the time of Sībawayhi (*Kitāb* II, 35; IV, 44) onwards, placed the collective in a separate position, rather than a fourth number category. The native terms employed to designate the collective are basically *ism al-jins* and *ism al-jamʿ*, thus reflecting the difference established between two forms of the collective (see Ḍayf 1990:57–58 for more details). The first, which may be translated as ‘the noun of the species’, has a noun of unit (*ism al-waḥda*), designating one individual out of a genus by means of a suffix attached to the collective (e.g. *ḥamām* ‘pigeons’ and *ḥamāma* ‘a pigeon [male or female]’, *naxl* ‘palm trees’ and *naxla* ‘a palm’, with the same suffix). The second native term, that is to say ‘the noun of the plural(ity)’, does not allow for the formation of a noun of unit with a suffix (*qawm* ‘people’, *ʾibil* ‘camels’). A corresponding noun of unit is then obtained from another linguistic root, like *jamal* ‘camel’, or simply does not exist.

For nouns designating things (countable as well as non-countable) or animals with herd instinct (small animals and insects) a variety of forms exists conveying the notion of collective. To all of these forms a suffix *-a(t)* may be attached to obtain the noun of the corresponding unit. Examples: *tuffāḥ* ‘apples’ and *tuffāḥa* ‘one apple’, *naml* ‘ants’ and *namla* ‘one ant’, *waraq* ‘leaves, sheets’ and *waraqa* ‘leaf or sheet’, *ḏahab* ‘gold’, *ḏahaba* ‘a piece of gold’.

For nouns designating ‘bigger’ animals, rational beings, and even some things, a collective form (*ism al-jamʿ*) is used without a corresponding *-at* noun of unit. There are various morphological patterns for these collectives: (a) *faʿl*: *rakb* ‘travelers [on a camel]’, *ṣaḥb* ‘companions’; (b) *faʿal*: *ġanam* ‘sheep’, *ḥaras* ‘guardians’; (c) *fiʿil*: *ʾibil* ‘camels’; (d) *faʿil*: *jamil* ‘a group of camels including owners and shepherds’, *jann* ‘genies, spirits’; (e) *fuʿla*: *ṣuḥba* ‘companions’.

For rational beings two further collective forms exist. The first conveys the notion of a group of people, and has a noun of unit with an *-iyy* suffix. Examples: *al-yāhūd* ‘the Jews’ and *yāhūdiyy* ‘one Jew’, *al-ʿarab* ‘Arabs’ and *ʿarabiyy* ‘one Arab’. The second is formed with the suffix *-at* attached to the noun of the agent (*fāʿil*, the corresponding derivate forms or even the intensive *faʿʿāl*). Examples: *muslima* ‘Muslims’ from *muslim* ‘Muslim’, *mārra* ‘those who pass by’ from *mārr* ‘passer-by’, *najjāra* ‘carpenters’, from *najjār* ‘carpenter’. This collective form is particularly productive in Modern Standard Arabic, and it is often opposed to another plural form, e.g. *muslimūna*, *mārrūna*, and *najjārūna*, referring to the plural, that is to say, the group of beings seen as a combination of individuals which retain their specific features (see Fleisch 1961:301–310).

In some cases the collective may be analyzed as a number category, producing a complex system for a given noun. An example is *naml* ‘ants [collective]’, which is morphologically singular, hence allowing for the formation of a plural *nimāl* ‘groups of ants [plural of abundance]’. The unit noun is *namla* ‘one ant’, which, as a singular, has its own secondary plural, *namalāt* ‘a small group of ants [plural of paucity]’.

One interesting syntactic feature regarding the collective is its agreement behavior when attached to other elements of speech. In general terms, with respect to the varying agreement patterns, a collective may be seen as morphologically singular, but semantically plural. The more the collective conveys the notion of a pure, inorganic mass, the more the language tends to use feminine nouns and verbs combining with it (*an-naml al-ḥamrāʾ* ‘red ants’ *šajar baʿīda* ‘trees far away’. In some instances, however, these collective forms may present another agreeement pattern, with a plural (*šajar biʿād* ‘trees far away’), thus supporting the idea that the collective could be the starting point for the formation of an undetermined plural. On the other hand, human collectives tend to be treated not as a pure mass, but as a combination of distinct individuals. However, some degree of variation is retained, especially in higher registers of the language. For example: *al-ʿarab al-mustaʿriba* ‘the arabicized Arabs’ and *qālat al-ʿarab* ‘the Arabs said’, where the idea of a mass prevails, thus making the feminine the agreement choice. In contrast, one finds *al-ʿarab al-fuṣaḥāʾ* ‘the pure and eloquent Arabs’ and *al-ʿarab qālū* ‘the Arabs said’, where the collective is seen as a combination of individuals, thus making the masculine the preferred agreement choice. The collective here acts more as a plural than as a pure collective. This last form is the preferred choice in the modern stages of the Arabic language.

Bibliographical references

Primary source

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān ibn Qanbar, *al-Kitāb*. Ed. ʿAbd as-Salām Muḥammad Hārūn. 5 vols. Beirut : Dār al-Kutub al-ʿIlmiyya.

Secondary sources

Ḍayf, Šawqī. 1990. *Taysīrāt luġawiyya*. Cairo: Dār al-Maʿārif.

Fleisch, Henri. 1961. *Traité de philologie arabe*. I. *Préliminaires, phonétique, morphologie nominale*. Beirut: Imprimerie Catholique.

Ignacio Ferrando (University of Cádiz)

Collocation

Collocation is the habitual association of two or more words to denote a particular meaning. It is a linguistic phenomenon that exists in Arabic as in other languages and cuts across semantics, lexicography, grammar, translation, and cognitive semantics. It has been studied as part of each of these fields and given different labels accordingly. It is only recently that collocation has been studied on its own as a linguistic phenomenon.

Interest in collocation as a linguistic phenomenon in Arabic dates back to the work of traditional Arab philologists, who noted its existence in Arabic but did not assign it a label. Jāḥiẓ (*Bayān*) noted that certain lexical items in the *Qurʾān* acquire negative or positive connotations when they occur with other lexical items in certain contexts. He cites the two examples of *maṭarat* ‘it rained’ and *ʾamṭarat* ‘it has rained’, where the latter has the *ʾa-* prefixed to the verb. The former occurs within contexts indicating God's granting mercy, while the latter is associated with contexts of God inflicting torture.

Although not studied independently, collocation was still assigned an important focus in Arabic lexicography. This is reflected in the large number of Arabic monolingual thesauri of collocations produced by Arab philologists. Among them are al-Yāzijī's *Najʿat ar-rāʾid*, al-Hamaḏānī's *Kitāb al-ʾalfāḏ̣* and aṯ-Ṯaʿālibī's *Fiqh al-luġa*. Each of these divided the Arabic language into various conceptual topics. Under each topic, several expressions, collocations, vocabulary items, and the synonyms denoting the concept were listed. Early studies either approached collocation contextually or as part of lexicography, but none studied it independently.

Independent study of collocation in Arabic started in the mid-1960s and was influenced by Firth's (1957) 'meaning by collocation'. Unlike early studies, modern studies attempted to coin an Arabic term designating this linguistic phenomenon. According to ʿAbd al-ʿAzīz (1990:60), ʾAbū al-Faraj (1966:111) was the first to introduce the term *al-muṣāḥaba* 'collocation' to Arab readers. Noting that entries in Arabic monolingual dictionaries provide illustrative examples listing the collocants of the lexical item in question, he borrowed Firth's 'meaning by collocation' and argued that the meaning of a lexical item is denoted by collocation. Arguing along the same line was Ezzat (1970, 1971). He used the term *al-muṣāḥaba al-luġawiyya*, lit. 'linguistic co-occurrence' (1971:95) to refer to the phenomenon of a lexical item occurring in the company of another. The two lexical items become so strongly associated in the mind of the user that when one of them is mentioned it calls to the reader's mind the other lexical item. For instance, given the lexical item *jumʿa* 'Friday', a native speaker can list its other collocants, e.g. *ṣalāt* 'prayers' as in *ṣalāt al-jumʿa* 'Friday prayers'. He also noted that collocational patterns differ from one language to another (1970:29). Like ʾAbū al-Faraj (1966), Ezzat believed that grammar does not always account for collocation. He was the first to attempt a stylistic division of collocation into *ʿādiyya* 'normal', that is, collocations that are familiar and well known to the reader, and *ġayr ʿādiyya* 'extraordinary', that is, collocations that are stylistic deviations from familiar ones as they are used in literary texts.

El-Hassan (1982:273) provided a study that attempts to set semantic bases for collocation. Instead of *al-muṣāḥaba*, he coined the term *at-talāzum* 'strict co-occurrence' to denote collocation as a linguistic phenomenon. Like ʾAbū al-Faraj (1966) and Ezzat (1970, 1971), El-Hassan noted that grammar does not account for this linguistic phenomenon. For him, collocation is partly arbitrary and partly semantic. Studying collocational patterns in the *Qurʾān*, he identified three semantic relations connecting collocants. The first is the opposition relation connecting a verb like *yuḥyī* 'he raises to life' and its collocant, the imperfect of its opposite form *yumīt* 'he puts to sleep' or 'he puts to death' as in *yuḥyī wa-yumīt*. The second is synonymy where the meaning of one of the two collocants is synonymous with the meaning of the other, as in *al-mustaqarr wa-l-muqām* lit. 'the settling and residing place'. The third is complementary relation, where the meaning of one of the lexical items complements the meaning of its collocant, as, for example, *as-samāʾ wa-l-ʾarḍ* 'sky (heavens) and earth', where the first lexical item *samāʾ* 'sky, heavens' complements the second collocant, *ʾarḍ* 'earth.' These linguists all conclude that collocation is either arbitrary or determined by semantic relations. They all agree that grammar cannot always account for this linguistic phenomenon. A few studies, however, attempted to show that grammar determines collocation.

Among the earliest studies to show the influence of grammar on collocation was Ibn Fāris' *al-Ittibāʿ wa-l-muzāwaja*. By *ittibāʿ* he meant that a certain lexical item may be followed by one or two lexical items of the same triradical root as *tawkīd* 'corroboration.' He distinguished two types of *ittibāʿ*. The first has a lexical item followed by a meaningful lexical item of the same triradical root but of a different measure. Examples include *laylun lāʾil* and *ṣadīq ṣadūq*. In the first example, a lexical item is followed by the *ism fāʿil* 'active participle' *lāʾil* of the same root *l-y-l*, literally meaning 'nighting night', i.e. 'a very long night'. In the second, a lexical item is followed by the *ism mafʿūl* 'passive participle' *ṣadūq*, derived from the same root *ṣ-d-q*, literally meaning 'friend very friendly', i.e. 'a true friend.' The phenomenon of *ittibāʿ* performs two functions: it emphasizes the meaning of the first lexical item, and it creates a beautiful musical resonance resulting from the repetition of the same triradical root. In the second type of *ittibāᶜ*, the second lexical item may be meaningless, used only to create a musical effect. An example would be *šayṭānun layṭān* 'a devilish Satan', where the second lexical item does not denote anything but merely rhymes with the former as both lexical items end in *-ṭān*. Ibn Fāris lists examples of both types of *ittibāᶜ* and underscores the grammatical factor that might explain the co-occurrence of lexical items in certain collocational patterns.

The grammatical factor was further developed into a set of grammatical rules, or what Ḥassān (1986) refers to as *quyūd intiqāʾiyya* 'selectional restrictions'. Ḥassān (1986:306) uses the terms *tawārud* and *mulāʾama* lit. 'appropriateness', to refer to lexical items that are grammatically and semantically logical and co-occur in collocational patterns in grammatically acceptable and meaningful sentences. *mulāʾama* consists of a set of grammatical and logical rules, *quyūd intiqāʾiyya* 'selectional restriction rules', which function as "constraints on word combinations" (Lehrer 1974:183) determining which lexical items would co-occur to form meaningful sentences. The grammatical rules that he provides are similar to those of Ibn Fāris (*Ittibāʾ* 88). The Arabic grammatical rules of the → *mafᶜūl mutlaq* 'cognate accusative', for instance, require that a verb be followed by a *maṣdar* 'infinitive verbal noun' of the same triradical root, as in *sāra sayran* lit. 'he walked a walking' where the *mafᶜūl muṭlaq* (here *sayran*) is derived from the triradical root *s-y-r* (→ object, absolute). Another grammatical rule is that of *tawkīd lafḏ̣* 'verbal corroboration', in which a lexical item is followed by the same lexical item to emphasize its meaning as in *šayṭān šayṭān* 'devil devil' (1986:309). The semantic (or logical) approach to collocation study accounts for the appropriateness, or inappropriateness, of clause constituency (1986:314–417). A nominal clause that starts with a *mubtadaʾ* 'topic, subject of a nominal sentence' would be logically expected to have a *xabar* 'predicate'. On the other hand, a verbal clause that starts with a verb would be logically expected to have a *fāᶜil* 'agent, subject of a verbal sentence'. Additionally, the action in a grammatically meaningful clause should be assigned to the logically appropriate agent. A verbal clause starting with the verb *sāra* '(he) walked' would be expected to have an animate male human noun following the verb, e.g. *sāra r-rajul* 'the man walked [lit. walked the man]'. If, however, the verb is followed by an inanimate noun like *māʾida* 'table', the result would be *sārat al-māʾida* 'walked the table', a grammatically sound but semantically unacceptable clause. Ḥassān (1986) was the first to introduce and formalize selectional restriction rules for collocation in Arabic.

The notion of selectional restriction in collocation was given a new impetus by El-Gemei (1998). In a contrastive study of discourse-specific collocation in Modern Standard Arabic and American English, El-Gemei (1998:17) pointed out another level of co-occurrence restriction, the conceptual level. Certain semantic fields like *ʾirhāb* 'terrorism' and *kombyūtar vayrūs* 'computer virus' are conceptualized as 'enemy' or 'disease'. This concept explains why the lexical item *ʾirhāb* co-occurs with lexical items that belong to the military semantic field such as *hujūm* 'attack', *jabha* 'front', and *mukāfaḥa* 'anti-, combat' as in the following collocational patterns: *hujūm ʾirhābī* 'terrorist attack', *jabhat al-ʾirhāb* 'the front of terrorism' and *mukāfaḥat al-ʾirhāb* 'combatting terrorism, anti-terrorism'. It also accounts for the occurrence of *kombyūtar vayrūs* 'computer virus' with lexical items that also belong to the military semantic field, e.g. *yaḍribu* 'he hits', as in *al-vayrūs sa-yaḍribu* 'the computer virus will hit, strike', and the lexical item *muḍādd* 'anti-', as in *vayrūs muḍādd* 'anti-virus'. In addition to semantic, grammatical, and arbitrary rules of selectional restriction, El-Gemei's study added the conceptual level as a fourth level of co-occurrence restriction.

Collocation has also been studied as part of the Arabic-English, English-Arabic translation process. Khogali (2004) elaborates on the importance of collocation to translation. Awareness of collocation enables translators to understand the meaning of lexical items. This, in return, enables them to provide a more accurate rendering of the meaning into the target language by selecting appropriate collocants and avoiding literal translations that would otherwise render the target language incoherent or incohesive. Like Baker (1992), Khogali points to an important type of collocation, 'marked collocations', which occur in technical texts and form an integral part of their style and registers.

Economic texts, for example, require the use of certain collocations such as *yaṣrifu šīk* 'he cashes a check' which, if translated literally as 'he issues a check', would not only be meaningless, but would also cause a collocational clash, thereby disrupting the register of the target language.

The realization that collocation impinges heavily on the translation process prompted a surge in studies of collocation in translation. The majority of these studies attempted a semantic or syntactic classification of Arabic collocations with suggestions for overcoming the difficulties encountered in rendering them into the target language. Emery (1988, 1991), and El-Gemei (1998) borrowed Aisenstadt's (1978) and Cowie's (1983) classification of English collocations and mapped them onto Arabic collocations dividing them into three types.

The first is 'open collocations,' in which "each element is used in a common literal sense" (Cowie 1983:xiii). Examples include *waqqaʿa al-muʿāhada* '(he) signed the agreement' where the two collocants can contract collocational relations with numerous other lexical items. The verb *waqqaʿa* 'to sign', for instance, can collocate with nouns like the following: *xiṭāb* 'message, letter', *kitāb* 'book', or *waṯīqa* 'document'. This type is easily translatable into English since such collocations are found in English–Arabic bilingual dictionaries and the two languages allow for them.

The second type is 'restricted collocation', in which one of the two collocants "has a figurative sense not found outside that limited concept" (Cowie 1983:xii). Restricted collocations include examples such as *kabid as-samāʾ* lit. 'the liver of sky', in which the first term *kabid* literally means 'liver' but within this limited figurative sense it denotes the center of the sky. Although this type of collocation is found in Arabic–English dictionaries, it is not easily predictable. Heliel (1990), for example, notes that the word 'heavy' in English would have more than one equivalent in Arabic depending on the collocant. Examples such as 'heavy smoker', 'heavy industries', 'heavy rain' would translate into *mudaxxin mudmin*, *ṣināʿa ṯaqīla* and *maṭar ġazīr*, respectively.

The third type is 'bound collocation' which "exhibits unique contextual determination, in other words, one of the elements is uniquely selective of the other". Derivational richness in Arabic permits a particular root-pattern combination to be earmarked for a specific collocant (Emery 1991:51). A good example here would be verbs that have negative denotations due to their form (or measure), like *waʿada* 'to promise' and *ʾawʿada* 'to threaten'. The prefix *ʾa-* attached to the second verb to derive Form IV of the verb assigns it the negative meaning of making threats. Another example would be *ḥarb ḍarūs* 'horrendous war'. The difficulty in rendering this type of collocation lies in the fact that the target language (here English) lacks exact equivalents that "capture the attitudinal additional meaning" (Emery 1991). Thus the collocation 'horrendous war' is a partial rather than an exact equivalent of the source language collocation, because it does not ring with the connotations of the original Arabic collocation.

Baker (1992) points to another difficulty in translating collocations The difficulty is associated with culture-specific collocations, collocations that reflect certain religious, political, or social traditions unique to the source language community. The collocation 'law and order' in English translates into Arabic *al-ʿādāt wa-t-taqālīd*. The former reflects the English preference for law and order in English-speaking cultures while the latter reflects a preference for customs and traditions in Arabic-speaking cultures (Baker 1992). Another example is the collocation *yišrab šarbāt*, lit. 'he drinks syrup (sherbet)'. This collocation reflects the social custom prevalent among members of the Egyptian (and other Arab) societies: people drink such a beverage on happy occasions including weddings, births, and successes such as passing exams or promotions. According to El-Gemei (1998), the translation technique used in rendering this type of collocation depends on type of text and purpose of translation. If it occurs in a literary text where the purpose of the translation is to provide an exact portrayal of the original text in the source language, the translator is best advised to provide a literal rendition of such a collocation, accompanied by a paraphrase explaining its social connotations. If it is to appear in a non-literary text, the translator may provide a functional equivalent, substituting the collocation with reference to corresponding social habits in the other societies, which, in the case of English-speaking societies, is drinking champagne.

Studies that attempted a syntactic classification of collocations in Arabic include Al-Rawi (2001), Khogali (2004), and Hoogland (2003). Al-Rawi (2001) classifies Arabic collocations syntactically into five patterns. The first represents Verb + Noun collocations which translate into a Verb + Noun as, for example, *yanbaḥu l-kalb* ‘the dog barks’. The second represents cases of Adjective + noun construction as in *dirāsa iqtiṣādiyya* ‘economic study’, where the translator has to make a careful selection of the proper adjective form to convey the appropriate meaning: economic versus economical study. The third pattern consists of a verb (usually transitive) followed by a noun, *ʿaqada ijtimāʿan*, which would simply translate into a verb noun collocation, ‘(he) held a meeting’. The fourth pattern is the Verb + Noun + Adjective collocation, such as *taqaddama taqadduman baṭīʾan*, which would translate into Verb + Adverb ‘(he) progressed slowly, made slow progress’. The last pattern represents the Noun + Noun construction, which includes groups of nouns such as *qaṭīʿ ġanam* ‘a herd of sheep’. Al-Rawi (2001: 26) advises the translator to be careful in making the proper choice in translating this type of collocation.

Khogali (2004:1–2) attempted a more developed syntactic and semantic classification of collocations in Arabic. He divided collocations syntactically into five types based on the categories of the collocants: Noun + Verb *ʾaddā az-zakā* ‘to pay charities’, Noun + Noun *ʾirqāt ad-dimāʾ* ‘blood letting’, Verb + Verb *jaʿala yaqūlu* ‘(he) kept saying/started to say’, Adjective + Noun *ṯāqib ar-raʾy* ‘(being of) an extremely sound opinion’, and Verb + Preposition + Noun *taxarraja fī l-jāmiʿa* ‘(he) graduated from college’. He also divided collocations semantically into three types: *tawārud basīṭ* ‘simple (open) collocation’ where the language user does not strongly associate one lexical item with its collocant since they may collocate with several other lexical items; *tawārud wasīṭ* lit. ‘middle (semi-restricted) collocation’ where one of the collocants is associated with one or more lexical items; and finally *tawārud waṭīd*, lit. ‘strong (restricted) collocation’, in which the two collocants are strongly associated with each other, so that the mention of one recalls the other(s). This division corresponds to Aisenstadt’s (1979) and Cowie’s (1983) division of English collocations.

Hoogland’s (1993) study of collocation is unique in that it provides a more practical strategy for compiling an Arabic–Dutch dictionary of collocations. Ḥāfiẓ (2002) and Heliel (1990) note that bilingual Arabic dictionaries lack any English-Arabic dictionaries of collocation. The surge in studies of collocation in translation has finally led to the production of Arabic–English dictionaries of collocation such as those of Heliel (2000) and Ḥāfiẓ (2003).

The study of collocation in Arabic exists both as an independent field of study and as part of translation studies. It has recently been incorporated in the field of → corpus linguistics. Such studies are likely to contribute to lexicography as well as to the examination of the collocation phenomenon in Arabic.

Bibliographical references

Primary sources

Hamaḏānī, *ʾAlfāḏ* = ʿAbd ar-Raḥmān ibn ʿĪsā al-Hamaḏānī, *Kitāb ʾalfāḏ al-ʾašbāh wa-n-naḏāʾir*. Ed. Zahrān al-Badrāwī. Cairo: Dār al-Maʿārif, 1989.
Ibn Fāris, *Ittibāʿ* = ʾAbū l-Ḥusayn ʾAḥmad Ibn Fāris, *al-Ittibāʿ wa-l-muzāwaja*. Ed. Kamāl Muṣṭafā. Cairo: Maktabat as-Saʿāda, 1947.
Jāhiẓ, *Bayān* = ʾAbū ʿUṯmān ʿAmr ibn Baḥr al-Jāḥiẓ, *al-Bayān wa-t-tabyīn*. Ed. ʿAbd as-Salām Hārūn. Cairo: Maktabat al-Xānjī, 1948.
Ṯaʿālibī, *Fiqh al-luġa* = ʾAbū Manṣūr ʿAbd al-Malik ibn Muḥammad aṯ-Ṯaʿālibī, *Fiqh al-luġa wa-sirr al–ʿarabiyya*. Cairo: ad-Dār al-ʿArabiyya li-l-Kitāb, 1981.
Yāzijī, *Najʿat ar-rāʾid* = ʾIbrāhīm al-Yāzijī, *Najʿat ar-rāʾid wa-sīrat al-wārid fī l-mutarādif wa-l-mutawārid*. Beirut: Maktaba Lubnān, 1970.

Secondary sources

ʿAbd al-ʿAzīz, Muḥammad Ḥasan. 1990. *al-Muṣāḥaba fī t-taʿbīr al-luġawī*. Cairo: Dār al-Fikr al-ʿArabī.
ʾAbū al-Faraj, Muḥammad Aḥmad. 1966. *al-Maʿājim al-luġawiyya fī ḍawʾ dirāsāt ʿilm al-luġa al-ḥadīṯ*. Cairo: Dār an-Nahḍa al-ʿArabiyya.
Aisenstadt, Esther. 1979. “Collocability restrictions in dictionaries”. *Dictionaries and their users*, ed. Reinhard Hartmann, 71–74. Exeter: University of Exeter.
Al-Rawi, Sabah. 2001. “Rendering Arabic collocations into English”. *al-Lisān al-ʿArabī* 52.23–27.
Baker, Mona. 1992. *In other words: Coursebook on translation*. London and New York: Routledge.
Cowie, Anthony P. 1983. “The treatment of collocations and idioms in learner’s dictionaries”. *Applied Linguistics* 2.223–235.
El-Gemei, Dalal. 1998a. “A computer-assisted study of collocation: Two levels of co-occurrence restrictions in discourse”. *Journal of the Faculty of Education (Literary Section)* 4:3.1–72
——. 1998b. *al-Muṣāḥabāt fī l-luġa al-ʿarabiyya wa-muškilāt tarjamatihā ʾilā l-ʾinglīzīyya*. Paper presented at the International Conference on Cultural

Interaction and Translation, al-Azhar University, 22–26 June 1998.

El-Hassan, Shahir A. 1982. "Meaning by collocation". *al-Majalla al-ʿArabiyya li-l-ʿUlūm al-ʾInsāniyya* (Kuwait) 2.273–280.

Emery, Peter. 1988 . "Collocation: A problem in Arabic/English translation". *Quinquerime* 2:2.178–184.

——. 1991. "Collocation in Modern Standard Arabic". *Zeitschrift für Arabische Linguistik* 23.50–60.

Ezzat, Ali. 1970. "an-Naqd al-ʾadabī wa-ʿilm al-luġawiyyāt al-ḥadīṯ". *Majallat al-Majalla* 168.27–31.

——. 1971. "ʿIlm al-ʾuslūbiyyāt wa-mašākil at-taḥlīl al-luġawī". *al-Fikr al-Muʿāṣir* 80.70–95.

Firth, J.R. 1957. *Papers in linguistics 1934–1951*. London: Oxford University Press.

Ghazala, Hassan. Forthcoming. *Dictionary of collocations English–Arabic*.

Ḥāfiẓ, aṭ-Ṭāhir. 2002. "Arabic collocations: The need for an Arabic combinatory dictionary". *International Journal of Arabic–English Studies* 3.94–105.

——. 2003. *al-Ḥāfiẓ dictionary of Arabic collocations*. Beirut: Librairie du Liban.

Ḥassān, Tammām. 1986. "Ḍawābiṭ at-tawārud". *Majalla Muʿjam al-Luġa al-ʿArabiyya* 58.306–331.

Heliel, Mohamed Hilmi. 1990. "Collocations and translation". *FIT Newsletter* 11:3.

——. 2000. *Arabic textual expressions and collocation*. Cairo: Egyptian International Publishing Company – Longman.

Hoogland, Jan. 1993. "Collocation in Arabic (MSA) and the treatment of collocations in Arabic dictionaries". *Proceedings of the colloquium on Arabic lexicology and lexicography*, ed. Kinga Dévényi, Tamás Iványi, and Avihai Shivtiel, 75–93. Budapest: Eötvös Loránd University and Csoma de Kőrös Society. <http://www. let.kun.nl/WBA/content2/1.8.1_collohtm>.

Khogali, Hišām. "al-Mutarjim wa-l-mutawāridāt". <http://www.arabicawata.org/Arabic/ourlibrary/researchesandthesis/2004/january/research1.html>.

Lehrer, Adrienne. 1974. *Semantic fields and lexical structure*. London: North Holland.

Dalal Mahmoud El-Gemei
(al-Azhar University, Women's Branch)

Colloquial

This entry focuses on the variety of Arabic that is used for familiar and informal conversation, which is the primary locus of speech uttered in specific social and situational contexts. Speech in its natural social context is the *sine qua non* for a thorough study of language as a social phenomenon because, as emphasized by Labov (1989:52) "language is not a property of the individual but of the speech community". Language in this sense is redefined by sociolinguists as spoken language, speech discourse, the language used every day (Labov 1972b:258). A 'speech community' is a group of communicating individuals who are potentially able to communicate by their knowledge of language varieties and speaking rules. It is a basic, nonlinguistic, social analytical starting point in any study seeking to relate linguistics to social and political forces (Labov 1972a:43; Milroy 1992:4–5; Gumperz 1996a, 1996b:374; Chambers 1998:269; Ervin-Tripp 1977:192; Fasold 1990:65; Hymes 1986:54).

In the context of sociolinguistics, 'language' refers to every colloquial language variety – familiar conversational speech – which can be traced in a large number of speech communities and which exhibits a structured nature. In the case of Arabic, this structured nature constitutes the comparative basis between the subsystems (Abboud-Haggar 2003). Therefore, Arabic colloquial varieties, or dialects, acquire a special relevance as the only valid linguistic object for sociolinguistics, because they are an authentic reflection of societal situations. The study of this variety is the only way through which modern sociolinguistics can reach its goal of understanding language as a social phenomenon.

Sociolinguistic studies of Arabic colloquials in different speech communities – Cairo (Haeri 1996), Alexandria (Wahba 1996), Amman (Daher 1998), Bahrain (Holes 1987), Tangiers (Herrero 1996), and the region of Jbala in northwest Morocco (Messaoudi 1999), among others – show that, as in the case of other colloquials, there are no single-style speakers, style switching is a reality, variability is inherent in speech, and style stratification has to do with societal factors (Abd el-Jawad 1987:359–360).

Linguistically, Arabic colloquials are part of a language situation characterized by → diglossia, a term defined by Ferguson (1959). A diglossic speech community is one whose speakers use their local dialect (low variety) at home or among family and friends from the same dialect area, but the standard language (high variety) in communication with speakers of other dialects or on public occasions. Each level has its own special uses, depending on the context or the topic treated (Ferguson 1996, 1996a; Myers-Scotton 1986; Mahmoud 1986; Fasold 1990: 34–60). The coexistence of both varieties of the same language is common to all Arabic-speaking societies. The standard variety never functions as the colloquial one, and if some vernacular varieties have certain phonological characteristics in common with the standardized variety – mostly Bedouin and Bedouin type varieties – this does not mean that their speakers speak the 'Classical variety', but only that they approximate normalized Arabic (Abboud-Haggar 2003:92–95). Consequently, most language variation is measured through this reality, where the highest level is the reference point, the variety acquired through education, and the bottom level is the colloquial, the native variety, acquired as a mother tongue (Coulmas 1981).

A five-level scale established by Badawī (1973), whose approach was inspired by the then new trend of sociolinguistics, tried to fix this linguistic pattern and show that attaining the highest level possible depended on education level – school, university – and not on social class or gender. The diglossic situation became embedded in the societal matrix and made it an indispensable element in sociolinguistic analysis in spite of the difficulties inherent in carrying out thorough qualitative studies on collected colloquial data in order to determine its style and establish its level or stratum (Fasold 1984:61–84). In fact, Arabic diglossia does not show two fixed poles from which speakers can choose, since there is a whole continuum of levels of possible variations which depend on many non-linguistic factors such as setting (formal or informal), topic (serious or light), linguistic skill and mastering of Classical Arabic, emotional state of speakers, number of participants in the discussion, function of the discourse, and personal relationship with the audience. Each factor may be counted as an extra-linguistic variable (Badawī 1973; Badawī and Hinds 1986: introduction; Haeri 1996: 69–70, 162–168; Talmoudi 1984; Elgibali 1993; Holes 1993; Hary 1996:76–83; Walters 1996; Wahba 1996:103–104).

Another term introduced by Ferguson from a structural point of view is that of 'bidialectalism', the coexistence of two or more dialectal varieties. In an Arabic speech community, these varieties constitute different systems, where social prestige is determined by the speech community: both varieties can be assigned the same status, or one of them may be given a higher rank or special prestige, for instance, the variety spoken in the capital (Abd el-Jawad 1987: 359–361; Holes 1987).

The study of language in contact, originally linked with structuralist linguistics, but now closely related to sociolinguistics, contributes to the analysis of Arabic colloquial patterns by focusing on the impact of contact with other languages. → 'Code-switching' is one of the results of bilingualism or language contact, widely studied and defined in several ways, although there is no unanimity about its definition, since the term 'code' is used as a cover term for different languages, or dialects of the same language, or styles within a dialect (Myers-Scotton 1997; Franceschini 1998). Given that bilingualism and code-switching characterize various Arabic speech communities (→ multilingualism), especially in cities in Morocco (Forkel 1980), Algiers (Morsly 1986), and Tunisia (Jerad 2002), as well as in those communities that live outside their native countries (Bentahila and Davies 1983), and since bilingual schools in many Arabic countries provide a 'language in contact' type education (Mouatassime 2001), this aspect must be taken into account when establishing a linguistic pattern (Mahmoud 1986).

Sociolinguists working on urban Arabic colloquials – very few compared to the large number of urban Arabic speech communities – followed Labov's interviewing and data collecting methodology, studying variation embedded in the societal and linguistic matrix and selecting variables with numerous and frequent variants (Abd el-Jawad 1981; Haeri 1996; Herrero 1996; Wahba 1996). But in order to apply a strictly sociolinguistic methodology, scholars must try to avoid the strong and persistent influence of diglossia, restricting their research to colloquials – Labov's 'vernacular' – in the dialectological sense proposed by Holes (1987:7): "Variation in dialectal Arabic should not be discussed as 'interference' from the standard, but incorporated into dialectological description since from the speaker's point of view it is every bit as much a part of his speech behaviour as 'the dialect'" (cf. Haeri 1996:16–17).

Variables often chosen are phonological: interdentals /ṯ/, /ḏ/, and /ḏ̣/; uvulars /q/, /x/, and /ġ/, pharyngeals /ḥ/, /ʕ/; velarized /ṭ/, /ḍ/, /ṣ/, and /ḏ̣/, palato-alveolar /j/, and finally emphasis. Morphological variables, such as number, and syntactic variables, such as word order, are also possible, given the supposed frequency of variants (Badawi 1973:120–125; Elgibali 1993:79).

The most frequently selected variable is the realization of the Standard Arabic voiceless, uvular stop /q/, which, on both idiolectal and sociolectal levels, shifts easily from one style to another, depending on who says what to whom and in which context, due to its strong link with Classical Arabic (Badawi 1973; Salam 1980:90; Holes 1987:48–56; Abd el-Jawad 1987:361–364; Haeri 1996:11, 103–158).

The realization of emphasis or velarization is sociolinguistically relevant, too, since it is not linked to diglossic behavior and is especially sensitive to social stratification, as is the case in two Egyptian urban speech communities, Cairene and Alexandrian. Emphasis brings out many societal patterns of the communities studied. Its variants are not related to social class, but to educational level. Generally speaking, educated speakers show a lesser degree of emphasis, tending to avoid similarity with the patterns of Classical Arabic and to select a pronunciation based on the norms of the prestigious colloquial variety, that of the capital. On the other hand, non-educated speakers tend to produce a strong degree of emphasis, tending towards a pronunciation similar to Classical Arabic (Badawī 1973:182–183; Wahba 1996:106–108, 122–123; Haeri 1996:43–100, 1996a). With regard to → gender, as stated by Wahba (1996), females of both educated and non-educated informants tend to produce less emphasis, showing a gender distinction, while heavy emphasis is related to masculinity.

The aim of the sociolinguistic study of Arabic colloquials does not differ from that of the sociolinguistic study of any language and speech community. Its aim is to improve linguistic theory, to acquire a better understanding of the sources of linguistic changes, and establish an empirical linguistic pattern of the spoken language in its social context. Another aim is the application of sociolinguistic data to practical issues, such as education, language acquisition, and institutional language planning, to collaborate in finding solutions for social and ethnographic problems within a speech community, and to help in topics related to the psychology of language that affect families, schools, professions, legal proceedings, etc. (Trudgill 1984; Edwards 1984). In Arabic speech communities, perhaps the most urgent need is for sociolinguistics to help improve education, a field where scholars are debating the best way to enhance the acquisition of Modern Standard Arabic, while at the same time maintaining colloquials as the authentic expression of people's identities (Mahmoud 1986; Benjalloun 2001; Mouatassime 2001).

BIBLIOGRAPHICAL REFERENCES

Abboud-Haggar, Soha. 2003. *Introducción a la dialectología de la lengua árabe*. Granada: El Legado Andalusí.

Abd el-Jawad, Hassan. 1981. *Lexical and phonological variation in spoken Arabic in Amman*. Ph.D. diss., University of Pennsylvania.

——. 1987. "Cross dialectal variation in Arabic competing prestigious forms". *Language in Society* 16.359–368.

Badawī, as-Saʿīd Muḥammad 1973. *Mustawayāt al-ʿarabiyya al-muʿāṣira fī Miṣr*. Cairo: Dār al-Maʿārif.

—— and Martin Hinds. 1986. *A dictionary of Egyptian Arabic: Arabic-English*. Beirut: Librairie du Liban.

Benjalloun, Said. 2001. "L'arabe: De la langue maternelle marocaine à la langue d'enseignement". *Les langues orales dans les pays mediterranéens, La revue des deux rives Europe–Maghreb* 2.9–26.

Bentahila, Abdelali and Eirlys E. Davies. 1983. "The syntax of Arabic-French code-switching". *Lingua* 59.301–330.

Chambers, J.K. 1998. "The variable as a structural unit". *Handbook of dialects and language variation*, ed. Michael D. Linn, 267–283. (2nd ed. London: Academic Press.)

Coulmas, Florian (ed.) 1981. *A festschrift for native speaker*. The Hague: Mouton.

Daher, Jamil. 1998. "Gender in linguistic variation: The variable (q) in Damascus Arabic". *Perspectives on Arabic linguistics*, XI, ed. Elabbas Benmamoun, Mushira Eid, and Niloofar Haeri, 183–296. Amsterdam and Philadelphia: J. Benjamins.

Edwards, John and Howard Giles. 1984. "Applications of the social psychology of language: Sociolinguistics and education". *Applied sociolinguistics*, ed. Peter Trudgill, 119–158. London and New York: Academic Press.

Eid, Mushira and Clive Holes (eds.). 1993. *Perspectives on Arabic linguistics*, V. Amsterdam and Philadelphia: J. Benjamins.

Elgibali, Alaa. 1993. "Stability and language variation in Arabic: Cairene and Kuwaiti dialects". Eid and Holes (1993:75–96).

—— (ed.). 1996. *Understanding Arabic: Essays in contemporary Arabic linguistics in honor of El-Said Badawi*. Cairo: American University in Cairo Press.

Ervin-Tripp, Susan M. 1977. "An analysis of the interaction of language, topic and listener". *Readings in the sociology of language*, ed. Joshua A. Fishman, 192–211. The Hague and Paris: Mouton.

Fasold, Ralph. 1984. *The sociolinguistics of society: Introduction to sociolinguistics*, I. Oxford: Blackwell.

——. 1990. *The sociolinguistics of language: Introduction to sociolinguistics*, II. Oxford: Blackwell.

Ferguson, Charles A. 1959. "Diglossia". *Word* 15.325–340. (Repr. *Sociolinguistic perspectives: Papers on language in society 1959–1994. Charles Ferguson*, ed. Thom Huebner, 25–39. Oxford: Oxford University Press, 1996.)

——. 1996. “Epilogue: Diglossia revisited”. Elgibali (1996:47–67).

Fishman, Joshua A. a.o. (eds.). 1986. *The Fergusonian impact: In honor of Charles A. Ferguson on the occasion of his 65th birthday*. 2 vols. Berlin and New York: Mouton de Gruyter.

Forkel, Fritz, 1980, *Die sprachliche Situation im heutigen Marokko: Eine soziolinguistische Untersuchung*. Ph.D. diss., University of Hamburg.

Franceschini, Rita. 1998. “Code-switching and the notion of code in linguistics”. *Code-switching in conversation: Language, interaction and identity*, ed. Peter Auer, 51–72. London and New York: Routledge.

Gumperz, John J. 1996a. “Introduction to part IV”. Gumperz and Levinson (1996:359–373).

——. 1996b. “The linguistic and cultural relativity of conversational inference”. Gumperz and Levinson (1996:374–406).

—— and Stephen C. Levinson (eds.). 1996. *Rethinking linguistic relativity*. Cambridge: Cambridge University Press.

Haeri, Niloofar. 1996. *The sociolinguistc market of Cairo: Gender, class and education*. London and New York: Kegan Paul International.

Hary, Benjamin. 1996. “The importance of the language continuum in Arabic multiglossia”. Elgibali (1996:69–90).

Herrero Muñoz-Cobo, Beatriz. 1996. *El árabe marroquí: Una aproximación sociolinguística*. Almería: Universidad de Almería.

Holes, Clive. 1987. *Language variation and change in a modernising Arab state: The case of Bahrain*. London and New York: Kegan Paul International.

——. 1993. “The uses of variation: A study of the political speeches of Gamal Abd al-Nasir”. Eid and Holes (1993:13–45).

Hymes, Dell. 1986. “Models of the interaction of language and social life”. *Directions in sociolinguistics: The ethnography of communication*, ed. John J. Gumperz and Dell Hymes, 35–71. New York: Blackwell.

Jerad, Nabiha. 2002. “Langue et identité dans le Magreb post-colonial”. *L’identité, choix ou combat? Actes du colloque international organisé à Tunis en mars 1998*, 175–199. Tunis: University of Tunis, Faculté des Sciences Humaines et Sociales.

Labov, William. 1972a. “The social stratification of (r) in New York City department stores”. *Sociolinguistic patterns*, 43–69. Philadelphia: University of Philadelphia. (Repr. *Handbook of dialects and language variation*, ed. Michael Linn, 221–244. 2nd ed. London, 1998.)

——. 1972b. “The study of language in its social context”. *Sociolinguistic patterns*, by William Labov, 183–259. Philadelphia: University of Philadelphia Press.

——. 1989. “Exact description of the speech community: Short *a* in Philadelphia”. *Language change and variation*, ed. Ralph W. Fasold and Deborah Schiffrin, 1–57. Amsterdam and Philadelphia: J. Benjamins.

Mahmoud, Youssef. 1986. “Arabic after diglossia”. Fishman (1986:I, 239–251).

Messaoudi, Leila. 1999. “Etude de la variation dans le parler des Jbala (nord ouest du Maroc)”. *Estudios de Dialectología Norteafricana y Andalusí* 4.167–176.

Milroy, James, 1992, *Linguistic variation and change: On the historical sociolinguistics of English*. Oxford and Cambridge: Blackwell.

Morsly, Dalila. 1986. “Multilingualism in Algeria”. Fishman (1986:253–263).

Mouatassime, Ahmed. 2001. “Le Magreb entre expression orale et langage écrit”. *Les langues orales dans les pays mediterranéens, La revue des deux rives Europe–Maghreb* 2.53–63.

Myers-Scotton, Carol. 1986. “Diglossia and code switching”. Fishman (1986:403–415).

——. 1997. “Code-switching”, *The handbook of sociolinguistics*, ed. Florian Coulmas, 217–237. Oxford: Blackwell.

Salam, A.M. 1980. “Phonological variation in educated spoken Arabic: A study of the uvular and related plosive types”. *Bulletin of the School of Oriental and African Studies* 43.77–100.

Talmoudi, Fathi. 1984. *The diglossic situation in North Africa: A study of classical Arabic/dialectal Arabic diglossia with sample text in “mixed Arabic”*. Göteborg: Acta Universitatis Gothoburgensis.

Trudgill, Peter. 1984. *Applied sociolinguistics*. London and New York: Academic Press.

Wahba, Kassem M. 1996. “Linguistic variation in Alexandrian Arabic: The feature of emphasis”. Elgibali (1996:103–128).

Walters, Keith. 1996. “Diglossia, linguistic variation and language change in Arabic”. *Perspectives on Arabic linguistics*, VIII, ed. Mushira Eid, 157–197. Amsterdam and Philadelphia: J. Benjamins.

Soha Abboud-Haggar
(University La Complutense)

Communal Dialects

As used in the literature on varieties of Arabic, communal dialects are those associated with specific religious communities across the Arabic-speaking world. As such, they represent a distinct category of social dialects, that is, dialects that tell us something about a speaker’s social background (e.g. sex, age, education, ethnicity, class and/or caste) in the context of the geographic dialect of a specified place. Ferguson and Gumperz (1960), in a paper cited by Blanc (1964), discussed below, explain the complexity of defining rigorously foundational notions like ‘variety’, ‘dialect’, and ‘language’; their observations are as true today as they were in the 1960s.

In his 1953 study of a northern Palestinian Arabic dialect spoken by Druze, Haim Blanc notes that linguistic distinctions across what he termed ‘religio-ethnic communities’ in the Arabic-speaking world were understudied. His 1964 monograph on communal dialects in Baghdad remains the classic work on this topic.

In important ways, it rightly continues to shape how communal dialects are understood.

As Blanc explains, the urban centers of Lower Iraq at the time were characterized by "the unusually profound and sharply delineated dialectal cleavage that divides these populations into three nonregional dialect groups, corresponding to the three major religious communities, namely the Muslims, the Jews, and the Christians" (1964:3). Blanc's choice of label in both his 1953 and 1964 monographs is noteworthy. The earlier label 'religio-ethnic communities' demonstrates not only that distinguishable language varieties of any sort are, by definition, shared by members of a group who are or come to be seen as a community by insiders and/or outsiders but also that in the case of the Arabic-speaking world, such communities are sometimes formulated in terms of religious confession, a social category that interacts with local understandings of ethnicity in complex ways. In other words, religious confession, like ethnicity, remains far more a matter of birth than of choice, as it is generally seen in contemporary America, for example. Blanc's later use of 'communal dialects' reminds us that, historically, at least, these communities, based on religion and ethnicity, lived segregated lives although they interacted in socially prescribed ways. As Blanc's work and that of others make clear, however, the origins of the linguistic differences are not to be found in communal segregation but rather in the combination of settlement history and communal segregation.

This discussion seeks to provide an overview of communal dialects across the Arab world, focusing on the sociolinguistic consequences of such dialects. Thus, it examines Blanc's initial taxonomy of kinds of communal dialects (section 1), sectarian differences among Muslim dialects (section 2), and written → Judaeo-Arabic as it might affect our understanding of → diglossia (section 3). Readers whose primary interest is the linguistic detail of particular cases may consult the works cited on specific varieties.

1. Blanc's (1964) taxonomy of kinds of communal dialects

Linguists expect minor differences in pronunciation and lexis, in particular, across regional and social dialects; while such differences may be salient enough to function as diagnostics for sorting speakers into groups, they often are not. Similarly, when the distinctions among dialects are based on religion, linguists would likely expect to find differences in terms related to the name(s) for the Deity and other phenomena associated with religious practices as well as interactional rituals such as greetings, leave-takings, etc., a point acknowledged by Blanc (1964). Discussions of communal dialects, however, are generally concerned with structural differences in the varieties, that is, salient differences in phonetic or phonological inventories, differences in their morphosyntax, and/or differences in the ways these components of the dialect have changed diachronically, especially in relation to other dialects.

As Blanc notes, if one takes religious affiliation as the relevant axis of social differentiation when distinguishing among dialects, cases from the Arabic-speaking world represent a spectrum of possibilities. Blanc writes of three relative degrees of differentiation: major, intermediate, and minor. As each of these categories is discussed, more recent work or reviews of work on the relevant varieties are mentioned to supplement Blanc's initial characterization.

For Blanc, the most robust category of communal dialects includes those where there is major differentiation between Muslim and non-Muslim varieties. His examples include the Muslim and Jewish dialects of some North African cities, especially Oran and smaller towns near Algiers, as well as those of Muslims and non-Muslims in Lower Iraq. His categorization is based on earlier work on North African varieties of Jewish Arabic (generally termed → 'Judaeo-Arabic') and his own research on Iraqi dialects. Such a categorization relies on two distinct criteria, one structural and the other social: the linguistic differences must be manifest throughout the phonology and morphosyntax of the varieties, and they must correlate completely with membership in the respective community. All these cases involve the presence within a single location of a variety of Arabic that traces its roots historically to the pre-conquest Bedouin dialects of the Arabian Peninsula coexisting with one or more varieties with roots in the sedentary dialects of that period (cf. Versteegh 1997:141–145; Massignon 1924 offers an overview of the spread of these groups of dialects). In North Africa, as in Iraq, the variety spoken by Muslims has Bedouin roots while the one(s) spoken by the non-Muslims represent sedentary types.

Blanc devotes the last chapter of his 1964 study to comparing and contrasting the three varieties of Arabic found in Baghdad and offering a discussion of the settlement history of each group, which helps account for the presence and nature of the communal varieties found there. As Jastrow (2004) explains, inspired by Blanc's work, his own research and that of others on non-Muslim varieties spoken in Iraq have resulted in a situation where far less is known about Muslim varieties of Iraqi Arabic than non-Muslim ones. Building on Blanc's work as well as more recent research, especially that of Abu-Haidar (e.g. 1990, 1991), Holes (1995) summarizes developments in Baghdad since the time Blanc wrote (→ Baghdad Arabic; Baghdad Arabic, Jewish).

The communal dialect situation in Baghdad described by Blanc as a case of major differentiation was far more complex than the mere existence of three distinct varieties, each perfectly correlated with religious affiliation. As he notes, in interactions across community lines, there was a great deal of what we would today term → speech accommodation (Giles, Coupland, and Coupland 1991), which had become completely conventionalized. Blanc explains that the Christian and Jewish varieties were used in domestic contexts and within their respective communities. However, Jews and Christians who spoke the Muslim variety used it in intercommunal and public situations. Thus, many non-Muslims were "nearly perfectly bidialectal" (1964:9; see Blanc 1960 for a detailed example of such bidialectalism). He likewise notes that the Muslim variety was the one most often known by Christians who spoke a language other than Arabic as their first language (e.g. Armenian) and often the only variety of Baghdadi Arabic they knew. Similarly, some non-Muslims, particularly Christians, used the Muslim variety or features of that variety in in-group interactions. In other words, the Muslim variety seemed to serve as a local prestige variety accorded at least covert and perhaps overt prestige by the minority communities, and the spread of features associated with it (or of the variety itself) had consequences for diachronic language change in these varieties via processes of diffusion, as Holes (1995) illustrates. Although Blanc does not explicitly make the point, he implies that Muslims did not speak the Christian or Jewish varieties though we can imagine many might have been familiar enough with them (or at least stereotypical features of them) to employ them in affect-laden situations, whether jocular or patronizing. Such situations of unequal distribution of varieties present in a location – minorities being (expected to be or become) bidialectal, while those of the majority are monodialectal, speaking only the prestige variety – reflect and create social hierarchies of various kinds. From this perspective, it is clear that the notion of communal dialect, as Blanc defines and uses it, results ultimately from the situation of indigenous Christians and Jews as *d̲immī*, or so-called protected minorities. Thus, Blanc's observations on the language of intercommunal interactions offer important insights into the role language played in social differentiation and structuring of Baghdad at the time, while helping us understand how and why change induced by contact and long-term accommodation between the varieties might occur.

Blanc's second category of intermediate differentiation focuses on Jewish and Muslim varieties of urban Arabic in North Africa, noting research from Algiers, Fez, and Tlemcen, to which should be added Cohen (1964, 1975) on Tunis. In discussing intermediate categorization, Blanc explains that although there was clear differentiation by social group, each religious community having a distinguishable variety, the linguistic distinctions were ultimately few in number and, more important, it seems for Blanc, the dialects of both communities were of sedentary origin, that is, both of the same type.

In the cases Blanc considers as illustrating a minor degree of differentiation, he comments that differences that did occur are generally marginal to matters of structure and that the correlation between specific forms and group membership is not robust. Thus, on hearing a tape of someone from the area speaking, a native of the region would not be expected to be able to discern the ethno-religious background of the speaker, assuming, of course, the tape contained no content-related clues to speaker identity or religious practice.

Blanc gives a number of examples of minor differentiation; these include the towns of what he terms Greater Syria, Upper Iraq, the Arabian Peninsula, and Egypt. Citing unpublished work of Piamenta, Blanc explains that "Jerusalem

Arabic, for example, is communally differentiated insofar as some (not all) Christians and Jews deviate from majority usage [i.e., that of Muslims] with respect to intonation patterns, vocal qualifiers, certain allophones, the frequency of certain consonantal assimilations, the proportion of Classical or Neo-Classical vocabulary items, certain idioms, and the proportion and phonetic treatment of European loanwords" (1964:14). The Jewish variety likewise contains vocabulary of Hebrew origin and in some cases shows immigration-related influence from non-local varieties of Arabic (Piamenta 2000 offers a lexical-semantic analysis of aspects of what he terms the Judaeo-Jerusalem vernacular as well as discussion of the community's shift to Hebrew). Particularly significant here is the breadth of Blanc's characterization of the differences among varieties: he is concerned not merely with segmental phenomena that recurred categorically (e.g., the consistent pronunciation of certain sounds) – the focus of most traditional dialectology (see Jastrow 2004 for a recent review of work on Arabic in this paradigm) – but also variable ones (e.g., the frequency with which phonological processes like assimilation occur) – the concern of variationist sociolinguistics. Blanc also notes differences in suprasegmental phenomena like intonation as well as the frequency and treatment of borrowings from learned varieties of Arabic and from European languages. Attention to such detail helps account for the continuing relevance of Blanc's initial account of communal dialects. Other cases of minimal differentiation cited by Blanc include Aleppo and Cairo for Christians and Jews, and Ṣanʿāʾ for Jews, as well as the situation of Christians and Druze in Lebanon and northern Palestine.

An approach that stands outside Blanc's can be found in Heath (2002), a most impressive addition to the work on Moroccan dialects of Jewish and Muslim Arabic that provides detailed information about a score or so of varieties, including those of villages in the south of the country where Jews spoke (Judaeo-)Arabic while local Muslims spoke some variety of Berber. The focus of Heath's work is Jewish and Muslim 'dialect networks', with the goal of providing the necessary linguistic documentation to understand the processes of leveling of dialect features currently occurring in the country's urban centers. As he notes, "Morocco is a special case that resists classification" (2002:2) in discussion of the history of Arabic there, and the same seems true for confessional varieties in that country.

It is important to remember that much of the work cited in this section looks to the past (even more so than language description usually does) because the Jewish communities described, in particular, no longer exist or are very much smaller than they were just over half a century ago. Thus, for example, Heath (2002:14) comments, "The J[ewish] data, mostly from aging émigrés, is 'frozen in time' and represents the state of Judaeo-Arabic around 1950, while the M[uslim] data is from a more diverse set of age grades and reflects the continuing koiné-ization that has been going on in Morocco in recent decades". In similar fashion, Cohen and Piamenta acknowledge the moribund nature of the varieties they describe, and Blanc himself explains that most of his data derived from speakers residing in the United States or Israel. In other words, these researchers generally imagined communities and varieties that in fact no longer existed *in situ*, often with a focus on 'authentic' speakers of 'pristine' varieties (cf. Heath 2002:22), as traditional documentary linguistics and dialectology have generally done. Although certain criticisms can be made of such a homogenizing approach, the current social and demographic realities of these areas remind us of the preciousness of these data and these studies: they constitute our best and likely only representation of the linguistic consequences of prolonged dialect contact and interaction crucial to our understanding of the history of Arabic, especially spoken dialects of Arabic in all these areas.

2. Sectarian differences among Muslims

Blanc (1964:9–10) clearly states that *pace* Massignon (1914), who had claimed there were three Sunni and two Shiʿi dialects of Arabic in Baghdad, he could find no evidence of linguistic differences between the dialects of Sunnis and Shiʿis in Baghdad that correlated with membership of one sect or the other. However, such sectarian differences have been documented in other areas. Citing Blanc, Holes (1983; see also 1987, 1995) labels the differences he found in the Arabic of Sunnis (*ʿArab*, who speak a

Bedouin variety) and Shiʿis (*Baḥarna*, who speak a sedentary variety) in Bahrain a case of "major communal differentiation" because the differences there meet Blanc's two initial criteria: permeation of both the phonology and the morphosyntax of the varieties and full correlation with community membership. Johnstone (1967), Prochazka (1981), and Al-Tajir (1982) had discussed the dialects of this same region, but did so from a more traditional historical perspective. Holes, however, working within the framework of variationist sociolinguistics, provides a very different perspective on communal dialects. Rather than describing the taxonomic differences between the varieties based on elicitation from a very small sample of speakers, as a traditional dialectologist would, Holes offers a quantitative analysis of the behavior of several phonemic and morphophonemic variables in a corpus of data gathered from a much larger sample stratified according to sect, region, literacy, and sex. His careful analysis represents an important contribution to our understanding of how social change can influence the trajectory of communal dialects and social dialects more broadly. As Holes explains, although the Shiʿis, who represent the indigenous population, are more numerous, the Sunnis, who arrived two centuries ago, dominate the government and form the ruling family. Each group remains endogamous, and until about three decades ago, the two communities lived in different villages or quarters of the island's towns. However, changing patterns of employment in the urban areas and the creation of mixed neighborhoods have altered the political economy of language in Bahrain. Like sociolinguists generally, Holes focuses not on the internal homogeneity of the varieties he studies, but on their patterned heterogeneity, correlating it with the social variables he used for informant selection. He documents how and why elderly village women, all illiterate, are most likely to retain and use relic forms while younger speakers of both sects who are literate use a modified form of the Sunni variety, which developed early last century and has become what Holes terms a "neutral 'standard'". The use of this standard results in a pattern of accommodation that recurs in many speech communities around the world, asymmetrical convergence in which the speech of the members of one social group, here, the Shiʿis, moves toward that of the other, here, the Sunnis, while Sunni speakers adjust their speech far less in intergroup interactions. When Sunnis shift, it is toward a form associated with Modern Standard Arabic (MSA), not one associated with the Shiʿis' variety. Thus, Shiʿi speakers, especially those with ties to village life, are bidialectal in the two local varieties in their daily lives. Holes (1995:276) concludes: "Thus, in Bahrain the linguistic effect of urbanization and increased literacy has been to level dialect differences, but do so in a way which reflects local status relations. MSA norms exert only a secondary effect, at least on phonology". Holes is able to track such shifts only because of the quantitative methodology he uses and his focus on language in use rather than the elicited forms preferred by traditional dialectologists. At the same time, although Holes pays great attention to the behavior of specific lexical items and categories of items in a way that sociolinguists working on Western languages generally do not, the results of his research describe and present these varieties in a very different way than would research by a traditional dialectologist. After all, variationist sociolinguists and traditional dialectologists conceptualize the nature of variation and – in many ways – language and linguistic systems in fundamentally different ways, as Walters (1988) details.

3. Written Judaeo-Arabic and our understanding of diglossia

An aspect of communal dialects in the sense of non-Muslim varieties of Arabic and more particularly Judaeo-Arabic that has received little attention among sociolinguists of Arabic is the fact that the varieties of Arabic used by Jews were not only spoken but also written during certain periods and for certain purposes, especially in North Africa and in Tunis in particular. As Sebag (1991:121) explains, Jews were prohibited from using the Arabic script (because of its association with Islam); hence, they used Hebrew characters to write the variety of Arabic they spoke. Chetrit's discussion of Judaeo-Arabic (Bunis, Chetrit, and Sahim 2003) surveys these developments and especially the role of the modern printing press in the creation of novel contexts for the use of the written language (e.g. newspapers, translations of works in European languages, Arabic, and Hebrew, and locally produced literary works), while Sebag (1991) and

Snoussi (2003) focus specifically on Tunis. As Chetrit points out, the spread of modern education in European languages and the advent of printing led to the creation of new varieties of spoken and written Judaeo-Arabic (Hary 1997 offers a sociolinguistic account of the development of written registers of Judaeo-Arabic across the past several centuries). Acknowledging these facts complicates our understanding of the nature of communal dialects (and hence the history of the Arabic language), especially in North Africa, even as it challenges students of diglossia to rethink one of Ferguson's (1959) foundational assumptions in his original formulation of the concept as it has influenced work in Arabic sociolinguistics, namely, that the low variety of Arabic is a spoken variety, one that is written in a very limited set of contexts if at all (Walters 2003). This situation likewise offers sociolinguists an opportunity to study a now-moribund practice of using the script associated with one language for writing what was initially only a spoken variety of another.

4. Conclusion

Even a cursory examination of the existing studies of communal dialects demonstrates how much research remains undone and can, indeed, never be done, a situation all too common with respect to varieties of spoken Arabic. Given the political and economic motivation for Jewish and Christian emigration from countries where Arabic is the dominant language, the size of these communities (if they continue to exist) is drastically smaller today than just over a half a century ago. Hence, researchers can no longer investigate such communal dialects in contact in the way that Holes has done with respect to Sunni and Shiʿi varieties because the multiconfessional communities that gave rise to them no longer exist. Particularly for Jewish varieties, one can minimally seek to reconstruct the near and distant past through the study of the language of emigrants and Judaeo-Arabic texts. It is likewise worth remembering that most research on a communal variety completed before Blanc (1964) and some research completed since then has not been comparative in focus. These facts limit our ultimate understanding of communal dialects, the relationships among them, and, ultimately, the history of Arabic. Finally, it is worth noting the very different projects of descriptive field linguists and traditional dialectologists, on the one hand, and sociolinguists, on the other, when reading and evaluating work on this topic. Although all are concerned with issues of diachronic change, they conceptualize its proper study in very different ways.

Bibliographical references

Abu-Haidar, Farida. 1990. "Maintenance and shift in the Christian Arabic of Baghdad," *Zeitschrift für Arabische Linguistik* 21.47–62.

——. 1991. *Christian Arabic of Baghdad.* Wiesbaden: O. Harrassowitz.

Al-Tajir, Mahdi Abdella. 1982. *Language and linguistic origins in Bahrain: The Baharnah dialect of Arabic.* London: Kegan Paul International.

Blanc, Haim. 1953. *Studies in North Palestinian Arabic: Linguistic inquiries among the Druzes of western Galilee and Mt. Carmel.* Jerusalem: Israel Oriental Society.

——. 1960. "Style variation in spoken Arabic: A sample of interdialectal educated conversation". *Contributions to Arabic linguistics*, ed. Charles Ferguson, 79–161. Cambridge, Mass.: Center for Middle Eastern Studies, Harvard University.

——. 1964. *Communal dialects in Baghdad.* Cambridge, Mass.: Center for Middle Eastern Studies, Harvard University.

Bunis, David M., Joseph Chetrit, and Haideh Sahim. 2003. "Jewish languages enter the modern era". *The Jews of the Middle East and North Africa in modern times*, ed. Reeva Spector Simon, Michael Menachem Laskier, and Sarah Reguer, 113–141. Columbia: Columbia University Press.

Cohen, David. 1964. *Le parler arabe des Juifs de Tunis.* I. *Textes et documents linguistiques et ethnographiques.* Paris: Mouton.

——. 1975. *Le parler arabe des Juifs de Tunis.* II. *Étude linguistique.* The Hague: Mouton.

Ferguson, Charles. 1959. "Diglossia". *Word* 15.325–340.

—— and John J. Gumperz. 1960. "Introduction: Linguistic diversity in South Asia, studies in regional, social and functional variation". *International Journal of American Linguistics* 26:3.1–18.

Giles, Howard, Nikolas Coupland, and Justine Coupland. 1991. "Accommodation theory: Communication, context, and consequence". *Contexts of accommodation: Developments in applied sociolinguistics*, ed. Howard Giles, Justine Coupland, and Nikolas Coupland, 1–68. Cambridge: Cambridge University Press.

Hary, Benjamin. 1997. "On later and modern Egyptian Judeo-Arabic". *Humanism, culture, and language in the Near East: Studies in honor of Georg Krotkoff*, ed. Asma Afsaruddin and A.H. Mathias Zahniser, 199–224. Winona Lake, Ind.: Eisenbrauns.

Heath, Jeffrey. 2002. *Jewish and Muslim dialects of Moroccan Arabic.* London: Routledge Curzon.

Holes, Clive. 1983. "Patterns of communal language variation in Bahrain". *Language in Society* 12.433–457.

——. 1987. *Language variation and change in a modernising Arab state: The case of Bahrain.* London: Kegan Paul International.

——. 1995. “Community, dialect, and urbanization in the Arabic-speaking Middle East”. *Bulletin of the School of Oriental and African Studies* 58.270–287.
Jastrow, Otto. 2002. “Arabic dialectology: The state of the art”. *Israel Oriental Studies* 20.347–363.
Johnstone, Thomas M. 1965. *Eastern Arabic dialect studies*. Oxford: Oxford University Press.
Massignon, Louis. 1914. “Notes sur le dialecte arabe de Baghdad”. *Bulletin de l’Institut Français d’Archéologie Orientale du Caire* 11.1–24.
——. 1924. “Éléments arabes et foyers d’arabisation: Leur rôle dans le monde musulman actuel”. *Revue du Monde Musulman* 57.1–157.
Piamenta, Moshe. 2000. *Jewish life in Arabic language and Jerusalem Arabic in communal perspective: A lexico-semantic study*. Leiden: E.J. Brill.
Prochazka, Theodore. 1981. “The Shīʿī dialects of Bahrain and their relationship to the Eastern Arabian dialect of Muḥarraq and the Omani dialect of al-Ristāq”. *Zeitschrift für Arabische Linguistik* 6.16–55.
Sebag, Paul. 1991. *Histoire des Juifs de Tunisie: Des origines à nos jours*. Paris: L’Harmattan.
Snoussi, Mohamed Larbi. 2003. *La presse judéo-arabe dans la Tunisie coloniale (1884–1896)*. Tunis: MediaCom.
Versteegh, Kees. 1997. *The Arabic language*. Edinburgh: Edinburgh University Press.
Walters, Keith. 1988. “Dialectology”. *Language: The socio-cultural context*, ed. Frederick Newmeyer, 219–239. Cambridge: Cambridge University Press.
——. 2003. “Fergie’s prescience: The changing nature of diglossia in Tunisia”. *International Journal of the Sociology of Language* 163.77–109.

Keith Walters (University of Texas at Austin)

Comoros

1. Arabic in the Comoros

At the entrance of the Mozambique Channel, half-way between Africa and Madagascar lie the four islands that form the archipelago of the Comoros or ‘Islands of the Moon’ (*Juzur al-Qamar*) as medieval Arab geographers and travelers liked to call them. The archipelago has a population of about 700,000 inhabitants, of whom more than 90 percent are Muslim, belonging to the Šāfiʿī *maḏhab*. The common language for all islands is Comorian or Shikomor. Although a different dialect is spoken on each island, understanding between the inhabitants remains strong. Like Swahili, Comorian is an African language (Bantu) with 35 percent of its vocabulary borrowed from Arabic. The three official languages on the Comoros are Comorian, French, and Arabic.

Learning Arabic on the Comoro Islands is inseparably connected with Islamic learning, beginning with the study of the *Qurʾān* at the Qurʾānic school. It is not merely a religious ‘obligation’ but a social necessity, linked to historical political circumstances in the sense that at any given period, the state determined that learning Arabic was part of the general educational program it imposed on the entire country.

Learning of Arabic takes place both within the sphere of the state, i.e. under its control, and beyond it in non-state-controlled educational institutions and structures, in accordance with the fact that the ties between the Comoros and the Arab world are of two kinds. Before the colonization of the archipelago by the French in 1886, the cultural and religious evolution of the islands was linked to Zanzibar under the influence of the Arab Omanis of the Al Busaid. This influence was to last, albeit with less intensity, throughout the colonial period (Martin 1976). The ties between the Comoros and the Arab world during the whole of this period were essentially confined to individual Comorians in search of Islamic learning. Their most frequent destination was Zanzibar, occasionally followed by a stay in the Hadramawt, in Mecca for the *ḥajj* (pilgrimage), or at the Azhar in Cairo (Egypt).

Paradoxically, when the Comoros gained independence in 1975, the political authorities began to encourage and initiate other modes of contact with the Arab world. These were of a diplomatic nature and favored, for example, new cultural and religious ties. Flocks of young students went to Arab countries to receive religious education, while the latter in turn established educational institutions on the Comoros. In September 1993, the Comoros became a full member of the Arab League.

2. Learning Arabic in state institutions: Learning Arabic in the ‘French school’

At the end of the 1950s, or maybe a little later, Arabic was introduced as a modern language at the *lycée* on a par with English and Spanish. The textbook, method of instruction, and Arabic teacher training, hitherto unheard of by Comorians, were innovative for more than one reason. The textbook *Méthode de l'arabe littéral* by the French Arabists Gérard Lecomte and Ameur Ghedira is based on a teaching method conceived by French-speaking scholars for French-speaking students. The addressees of the book pursue their school career by means of the same logic used in the authors' own training. The great majority reason in the same categories of thought (those of Descartes and of the grammar of Port-Royal and Lhomond) imposed on teaching in France and its colonies. The authors basically adopted the terminology employed by the French Orientalist Silvestre de Sacy in his Arabic grammar published in 1820, which inspired many of the subsequent European Arabic grammar books.

The textbook by Lecomte and Ghedira is composed of two volumes, the "first of which can be used in the 4th/3rd grades and the second in 3rd/2nd grades". It was to become a gigantic sales success. After publication of both volumes between 1956 and 1967, it was declared the official textbook for the instruction of Arabic at the French secondary school. Moreover, it also became the prescribed Arabic textbook at the *Ecole Nationale des Langues et Civilisations Orientales* in Paris until the end of the 1980s, which comes as no surprise considering the fact that Lecomte was head of Arabic studies there until the mid-1980s.

The method used does not differ in any way from that of textbooks for other languages at the time. It consists of a number of units with the following structure: a text followed by a vocabulary list drawn from it, comprehension questions (conversation) to be answered by the students, a grammar section, and finally, exercises aimed at reproducing the content of the text previously studied. Based on a passive, non-experimental educational method, the textbook does not leave enough space for active student participation during lessons, nor does it equip students with the ability to work independently at an early stage, e.g., to use an Arabic dictionary.

The cultural context evoked by the textbook is quite different from that of the Comoros, an aspect shared with other contemporary textbooks for all manner of subjects. It is impossible to find a text representing one of the country's great religious manifestations, such as the *mawlid*, the *majlis*, the Friday or ʿĪd prayer, the departure of pilgrims to Mecca and their return to the Comoros, a funeral celebration, life in a Qurʾānic school, or a profane manifestation such as the *twarab* (a genre of sung poetry accompanied by music). Children acquire an Arabic vocabulary that deals with life in the skyscrapers of New York or with the political life of the Abbasid Empire, but will find no mention of things used or gestures made during a local religious ceremony. For obvious reasons this component of the textbook and its educational method were not received favorably by Comorian students. The textbook by Atoui (1978), which replaced that of Lecomte and Ghedira following independence in 1975, was in all likelihood inspired by its predecessor. The resemblance, in both content and form, is striking.

Today, Arabic language instruction, in both the Islamic *madāris* (sg. *madrasa*) and *maʿāhid* (sg. *maʿhad*), and in public and private French secondary schools is mostly in the hands of Arabic-speaking teachers educated at Islamic universities in Arab countries. Nevertheless, this has not solved all the problems. Although they speak Arabic perfectly, which is certainly a good thing, these teachers do not speak a word of French, which is the language of instruction in the country's system of education. Thus, the Arabic *baccalauréat* examination is conducted in French and Arabic.

Consequently, Arabic teaching results are still inadequate today. The poor performance of many students is due less to the qualifications of their teachers (as most speak Arabic well) than to the incoherence and multiplicity of educational methods within one educational system.

Behind the issue of Arabic language teaching methods in schools lies a genuine problem of identity, which is related to the position attributed to each of the three languages – Comorian as the mother tongue, French, and Arabic – in the social and political field. Parallel to this, the question arises as to the future of the Arabic-speaking elites, especially after the admission of the Comoros to the Arab League in 1993.

It should be remarked that language consists of 'words' people use to express 'things' of their past and present, as well as their aspirations. It is not merely a means of speaking, but allows those who command it to express their values and world-views. It is easy to imagine the shock resulting from a confrontation of the different values and ideologies embodied by the three languages in the Comorian context. Developing a Comorian language and introducing it gradually to the educational system with the prospect of one day making it the first national language would, therefore, appear to be a sound idea.

Since Ahmed Abdallah came to power in 1978 and in accordance with the politics of emphasizing the Arabo-Islamic symbolism of Comorian identity, the new regime has encouraged a collective interest in the Arabic language and Islamic learning in Arab countries, above all in Egypt and the Gulf countries, especially Saudi Arabia. Two other factors contributed to this phenomenon. On the one hand, there was the 1974 oil-price explosion that led to the influx of huge amounts of money to the oil-producing countries, enabling them to accept countless students from the Muslim countries of Asia and Africa, among them the Comorians. On the other hand, lack of organization in the educational system and its diminishing quality, both of which were due to the economic crisis, compelled many Comorians to apply for scholarships to Arab countries for their children. The nature of the studies they would pursue mattered little.

Students who had completed their studies abroad returned to the Comoros in the late 1980s. Some began to teach at the *madāris* they had established in their home villages, receiving a salary from the Islamic World League or the Fatwā Academy of Saudi Arabia. Others taught Arabic in the public *lycées* or *collèges*, or at the *maʿāhid* (secondary education institutions) founded in the 1980s and financed by the Islamic World League. Those who did not enter the teaching profession became businessmen. Almost all of them are engaged in *daʿwā* activities, and most are members of the Muslim fundamentalist FNJ (Front national pour la justice) Party (Ahmed 1999).

There are six Islamic World League *maʿāhid*: three located on Ngazidja, two on Anjouan (Ndzuwani), and one on Mohéli (Mwali). All teachers are Comorians who, having been educated in Arab countries, were supported initially by a number of Arabs (Moroccans, Egyptians, and Saudis).

The Comorian elites, who had been educated in the 'French school' and spurred the Comoros into joining the Arab League, were clearly not driven by the idea of culturally integrating their country into the Arab world. Their motivation seems rather to have been the anticipated financial benefits accorded by this organization to its members, which explains why they did not adopt a policy of Arabization toward education or administration, ignoring the advice given to them by the Arab League on this matter.

3. Learning Arabic outside the state institutions: Learning to read the Qurʾān

Unlike other Muslim countries, in particular those in West Africa, the sole aim of learning the *Qurʾān* on the Comoros is to teach children how to read, and to a lesser degree to write, and does not require them to learn it by heart. Instruction takes place at the Qurʾānic school and can be divided into three stages.

The first stage is called *kurasa* from Arabic *kurrāsa* 'notebook, booklet, brochure'. The *kurasa* is a small textbook, printed and edited on the Indian subcontinent. It is composed of two parts; the first consists of made-up words, most of which have no meaning (at least in Comorian or Arabic), and which a child is supposed to repeat until it can decipher complete words. The objective seems to be to make children learn the alphabet in blocks of words. It is not important that the latter have no meaning, provided the children can manage to read and pronounce them correctly.

The *kurasa* is often accompanied by the famous small board used in Qurʾānic schools all over Africa. The child alternately deciphers the words in the *kurasa* and those written on the board by the master. Following the almost universal method of gradually progressing from small to large, the child begins after a year with the second part of the *kurasa*, which extends from the *Fātiḥa* to sura 78 (*ʿĀmma* or *an-Nabaʾ*). The board is then finally abandoned; logic and collective belief hold that the child should now have mastered the reading of the alphabet and its transcription. Nevertheless, the method used in the final part of the *kurasa* differs little from the previous. The master reads out several verses that the pupil is required to learn during the day. The latter repeats them several times and spends all day reading them out loud. The master then repeats the procedure with the next pupil and so on. In the end the pupils all read the 'lesson of the day' out loud individually.

When a child has finished reading the sura *ʿĀmma*, he concentrates on the rest of the

Qurʾān (*msahafu* in Comorian). After the completion of the *msahafu*, only those who are predestined to become *ʿulamāʾ* by virtue of their intellectual capacity and ambition are permitted to stay at the Qurʾānic school. They subsequently pursue a course of studies tailored to this immense goal, which first leads them to the highest level their Qurʾānic master can achieve with his instruction. They then leave to continue their education with the country's great *ʿulamāʾ*, before going abroad (Zanzibar, Hadramawt, Mecca, Medina, or the Azhar in Cairo) to complete their studies. This path taken by the apprentice scholar, leading from the Qurʾānic school to the prestigious Muslim intellectual centers abroad via the local *ʿulamāʾ*, is marked by the study of various religious sciences. The priority clearly lies with Šāfiʿī *fiqh* and Qurʾānic exegesis. In the past, therefore, Arabic was learned outside the state institutions by means of studying Islam.

Today, learning Arabic and even a substantial Arabization is primarily due to the Islamic education in both the *madāris* that are financed and administered by the Islamic World League, and the private *madāris*, founded by many of the returnees from Saudi Arabia in their villages, with the support of the Islamic World League or the Fatwā Academy of Saudi Arabia. During the 1990s, two large Arab *daʿwā* centers were established on the Comoros, the African Muslim Agency and the Muʾassasat al-Ḥaramayn al-Xayriyya. The former is based in Kuwait, where its field of action is sub-Saharan Africa, while the latter is Saudi. Apart from charitable and *daʿwā* activities, both agencies founded educational institutions. The African Muslim Agency opened a bilingual (French-Arabic) school at primary and secondary level with a capacity of 600 male and female students. The Muʾassasat al-Ḥaramayn al-Xayriyya created the Madāris al-ʾĪmān, which offer courses at all levels up to the secondary level diploma.

This 'Arabization from below' on the Comoros could be compared with the case of → Djibouti, a member of the Arab League since 1977. Strictly speaking, Djibouti, with its two large ethnic groups, Afar and Somali, is Arab in neither the cultural nor the linguistic sense. The same is true of → Somalia, which joined the Arab League even earlier than Djibouti. It was the respective religious (the Afar and Somali embraced Islam from the 9th century on) and political considerations that made the two 'sponsors' of Djibouti, Yemen and Saudi Arabia, successfully support its candidature to the Arab League.

As far as the Comoros are concerned, Arabization is taking place to a certain extent, but not to the satisfaction of the protagonists (the French-speaking elites who administer the state, the Arabic-speaking section of the population, and the Arab League). The French-speaking elites favor the *status quo* or at least an Arabic language development that does not cause offence to the French language, essentially to satisfy the demands of the Arab League. Those who speak Arabic and thus automatically contribute to 'Arabization from below' have not yet achieved an obligatory status for Arabic to their satisfaction, insofar as their role remains irrelevant among the political decision-makers.

BIBLIOGRAPHICAL REFERENCES

Ahmed, Chanfi. 1999. *Islam et politique aux Comores*. Paris: L'Harmattan.
——. 2002. *Ngoma et mission islamique (da'wa) aux Comores et en Afrique orientale: Une approche anthropologique*. Paris: L'Harmattan.
Atoui, H. 1978. *L'arabe, langue vivante: Méthode d'enseignement à l'usage des francophones*. 2 vols. Paris: Edicef.
Lecomte, Gérard and Ameur Ghedira. 1980. *Méthode de l'arabe littéral*. 3rd ed. 2 vols. Paris: Klincksieck. (1st ed. 1956; 2nd ed. 1970.)
Martin, Bradford G. 1976. *Muslim brotherhoods in nineteenth-century Africa*. Cambridge: Cambridge University Press.

AHMED CHANFI
(Center for Modern Oriental Studies in Berlin)

Comparative → Elative

Complement Clause → Subordination

Complementizer → Subordination

Compounds

Naḥt is the term Arab grammarians use for the word formational method whereby a new form is coined out of two or more independent words, a process similar to what in English is referred to by such terms as 'blending' and 'fusion'. Morphologically, this term derives from the root *n-ḥ-t* 'to carve, hew, chisel'. Thus, Arabic *manḥūtāt* (pl. of *manḥūt*), '*naḥt*-type constructions', are words, mostly of quadriliteral roots, that are 'chiseled' out of other words, so to speak. These constructions, however, are not all of the same compositional and/or functional nature; they may be said to fall into the following types, as can be gathered from medieval discussions of the subject (Ali 1987:59–85):

i. Acronym-like constructions based on conventional religious expressions or formulae consisting of several elements usually not fully represented in the *naḥt*-construction, e.g. *basmala* < *bi-smi l-lāhi* (*r-raḥmāni r-raḥīm*) 'to say: in the name of Allah (the Beneficent, the Merciful)'; *ḥawqala* < *lā ḥawla wa-lā quwwata* (*ʾillā bi-llāh*) 'to say: there is neither might nor strength (save in Allah)'; *ḥamdala* < *al-ḥamdu li-llāh* 'to say: praise be to Allah'; *damʿaza* <*ʾadāma* (*l-lāhu*) *ʿizzak* 'to say: may Allah preserve your glory!'
ii. Quadriliteral constructions in which a sound is added to a triliteral word, which serves to modify its meaning, usually by introducing an element of intensification, e.g. *raʿšan* 'timorous', from *raʿš* 'shaking'+ /n/ (Ibn Fāris, *Ṣāḥibī* 102).
iii. Relative adjectives (ending in the suffix *-ī*) formed from parts of nouns constituting genitive constructions, e.g. *ʿabdarī* '[a person] belonging to ʿAbd ad-Dār [the name of a family in Mecca]', *ʿabšamī* '[a person] belonging to ʿAbd Šams [name of a clan]', *ʿabqasī* '[a person] belonging to ʿAbd al-Qays [name of a tribe]'
iv. Quadriliteral portmanteau constructions combining parts of two, often semantically overlapping, triliteral words, e.g. *julmūd* 'rock', from *jaluda* 'to be tough' and *jamuda* 'to harden'

A close look at these sub-types of *naḥt*-construction is bound to reveal a number of important facts about the role of this process as a means of lexical expansion in Arabic. *Naḥt*-featured constructions do not constitute a sizeable portion of the Arabic vocabulary; they are far less numerous than those generated by → *ištiqāq qiyāsī* 'analogical derivation'. Those in (i) began to be used with the advent of Islam as abbreviated forms of highly frequent formulaic religious expressions. However, new instances of similarly coined words have failed to appear in the language following this period, which has prompted modern scholars to characterize the process under discussion as artificial, unproductive, and morphologically incompatible with the Arabic system of word formation (Yaʿqūb 1986:209–214; Tarzī 1968:363; Mubārak 1964:148–149).

The process involved in the formation of type (ii) constructions, i.e. the affixation of a single sound (usually /m/ or /n/) to a word in order to change or modify its meaning, has also failed to be productive. This may be accounted for in terms of the fact that a single sound has no meaning in itself.

Relative adjectives such as those given in (iii) above are again extremely rare in present-day Arabic. As the examples show, such formations were confined to the names of some families and tribes in olden times.

Type (iv) constructions, which resemble what in English would be categorized as blendings, like 'brunch' and 'motel', were discussed by several medieval grammarians, among whom the aforementioned philologist Ibn Fāris (d. 1000) stands out as the main contributor with his *Muʿjam maqāyīs al-luġa*. He was one of the leading proponents of the view that most quadriliteral and quinquiliteral words in Arabic are products of the *naḥt* process (Ibn Fāris, *Ṣāḥibī* 263–264). The account he gives, however, is not free from discrepancies and inaccuracies, which has led several modern scholars to question the validity of his claims. A case in point is the word *ṣillidm* 'hard-hoofed', of which two accounts are given, once as a *naḥt*-word from *ṣillad* 'strong' and *ṣadm* 'collision', and once as deriving from the former with the sound *m* suffixed to it (Ibn Fāris, *Ṣāḥibī* 264 and 102 respectively). His enthusiasm for establishing his theory was such that sometimes he mistakenly treated as native *naḥt*-constructions words that Arabic had borrowed from other languages. For example, the Persian loanword *farazdaq* 'leaven; lumps of dough' was thought to be a derivative of Arabic *faraza* 'to divide' and *daqqa* 'to grind [grain, etc.]' (Ibn Fāris *Maqāyīs* IV, 513).

It is obvious from the above discussion that *naḥt* can hardly be said to play a significant role in the configuration of the Arabic lexicon. There has been no direct statement by Arab philologists, not even by Ibn Fāris himself, as to the acceptability or analogicality of this process as a productive model on which new formations could be patterned. There is no clear explanation of the way it works; apart from the statement that it involves the formation of one word out of two others, there are hardly any rules as to the morphological segmentation of the underlying components or the order of the *manḥūt*-constituting elements.

The rarity of *naḥt* words in Arabic may be accounted for in terms of the inherent characteristics of the lexical structure of the language. Unlike the case in many European languages, in which a word may be composed of more than one root morpheme (hence the term 'polymorphemic' unit), the free one-root word is the most frequent lexical unit in Arabic. Arabic words are not structured as combinations of semantically distinct elements, unless they are multiword lexical constructions or set combinations, e.g. *nātiḥāt as-saḥāb* 'skyscrapers', *ġayr qābil li-t-tašabbuʿ* 'unsaturable', etc. (for more examples of such combinations, see al-Xaṭīb 1987; Nafūs 1985).

This being the case, it is no wonder that several modern Arab linguists and language reformers have raised questions about the viability of *naḥt* as a means of lexical creation in Arabic. This method, they argue, is 'alien' to Arabic and distorts the symmetry of its vocabulary (Jawād 1955:86), and it is far less productive than *ištiqāq* 'analogical derivation'. *Naḥt*-featured words, as some would estimate, do not exceed a few dozen (Wāfī 1972:187–189).

However, there have been other voices that are not so strongly opposed to *naḥt*, which, to them, is akin to *ištiqāq*, in that both processes involve the coining of new words from already existing elements (aṣ-Ṣāliḥ 1981:243; ʾAmīn 1958:391; Qaddūr 1993). There are also those who would contend that resorting to *naḥt* constitutes a form of lexical adaptation whereby the language may fill what seems to be a serious terminological gap, particularly in the fields of science and technology, and that traditional methods (i.e. analogical derivation) alone would not solve the problem (Jirjis 1961). A distinguished Arab scholar, al-Ḥuṣrī (1958:130–147), argues that *naḥt* is particularly advantageous in that it produces economical terms, which are preferable to long cumbersome constructions, particularly in the language of science and technology. He prefers, for example, *sarmana*, as an Arabic counterpart of 'somnambulism', from *as-sayru fī l-manām*. Similar, more recent instances of *naḥt*-constructions have indeed been suggested by several language reformers, translators, and compilers of bilingual dictionaries (see, for example, Baʿalbakī 2002).

Naḥt has also been the subject of long discussions in the Arab → language academies (Ali 1987:66–69). The views the academies hold regarding this process are basically the same as those of medieval grammarians. The process, as defined by the Cairo Arab Language Academy (1953:201), is "a kind of abbreviation; it consists in the formation of one word out of two or more others", which is basically identical with the definition given by al-Xalīl ibn ʾAḥmad (Jirjis 1961:63) and Ibn Fāris (*Ṣāḥibī* 263–264) many centuries ago. The Cairo Academy has also specified a number of features as being characteristic of *naḥt*-featured words, namely:

a. the underlying constituents of the *manḥūt* need not necessarily all be represented in it, examples of which we have already seen in type (i) above;
b. the first word in the underlying construction need not be retained in its original shape in the *manḥūt*, as illustrated in type (iii) above;
c. the short vowels (*ḥarakāt*) and zero vowels (*sakanāt*) of the 'letters' (i.e. consonants) constituting the underlying elements need not be observed in the *manḥūt*. For instance, in *maškana* from *mā šāʾa llāhu kān* 'whatever God intends, will happen', the *š* is unvowelled, unlike the case in the underlying word containing it (see al-Ḥuṣrī 1958:140–141).

What may be considered a more significant contribution of the Cairo Academy is that it has authorized the use of *naḥt*-featured words to create new scientific terminology. The authorization comes in a statement issued by a committee comprising a number of academicians especially appointed to assess the role and potential of *naḥt* as a means of lexical expansion. The statement reads: "We [the committee] agree to the permissibility of *naḥt* in scientific and technical disciplines due to the urgent need to express concepts pertaining to these fields in concise Arabic terms". This authorization, however, is subjected to the following restriction: "*Naḥt* is permissible [only] when necessitated by scientific need" (Cairo Arab Language Academy 1953:201–233).

Due to its vagueness, the phrase ‘scientific need’ has given rise to much controversy. Different scholars have interpreted it differently according to their respective attitudes. Those in favor of *naḥt* have taken it to be the long-awaited go-ahead, and set out to apply it unreservedly. On the other hand, those with a more conservative attitude have understood it to be of a rather restrictive force, permitting *naḥt* only in very rare cases.

The Academy has also stipulated, almost equally vaguely, that “*naḥt*-words be kept within the limits of comprehensibility” (Cairo Arab Language Academy 1953:158), meaning that the new forms should not strike the native speaker as entirely unfamiliar or unintelligible. In other words, *naḥt*-produced words should preserve the identity of their underlying constituents as far as possible. Thus, coinages of type (1) below are considered acceptable, whereas those of type (2) are rejected on the grounds of incomprehensibility:

(1) *kahramaġnāṭīsī* < *kahrabāʾī maġnāṭīsī* ‘electromagnetic’
kahraḍawʾī < *kahrabāʾī ḍawʾī* ‘photoelectric’
šibġarawī < *šibh ġarawī* ‘colloidal’
(2) *nazwara* < *nazʿu l-waraq* ‘defoliation’
ḥarsama < *ḥarrara min as-samġ* ‘to degum’
zahraja < *ʾazāla l-hīdrūjīn* ‘to dehydrogenate’

The point should be stressed that the above guidelines are too general to be of any practical use. It is not up to a particular speaker or group of speakers, particularly in the language of science, to mold words in such a way as to make them agreeable to taste, let alone the indefinability of the concept of ‘taste’ itself.

A special type is that of formations like *qab-tārīxī* ‘prehistoric’, *faw-bašarī* ‘superhuman’, *bay-sinnī* ‘interdental’, etc., which constitute what may be called ‘partial *naḥt*-constructions’. These are semi-transparent formations, in which the abbreviated constituents are closed-class items, mostly triliteral prepositions, which are reduced to biliteral units and used as prefixes. It is sometimes argued that such formations are more transparent and hence more likely to be understood than others like *zahraja* ‘to dehydrogenate’, *nazwara* ‘to defoliate’, and are therefore received more favorably. Furthermore, the argument goes, this does not constitute an altogether novel phenomenon in Arabic. The language has a number of prepositions and particles that are always prefixed to a following noun, the only difference being that these are sequences of one, rather than two consonants, plus a short vowel. Examples are: *li-*, as in *li-hāḏā* ‘for this reason’; *bi-*, as in *bi-ḏālika* ‘by that, in that manner’; *ta-*, as in *ta-llāhi* ‘by God’, etc.

A different type of compound is that of words formed with the prefix *lā-* (*lā-nihāʾī* ‘infinite’ etc.): these are what Arab grammarians refer to as instances of *murakkab mazjī* ‘mixed compound’, the formation into one lexical unit of two fully represented words that are otherwise used independently (Ali 1987:80–83). Modern Arabic abounds in new terminological formations in which *lā-* replaces such English negative prefixes as *a-*, *an-*, *anti-*, *non-*, *in-*, *un-*, etc. Examples include: *la-tamāṯul* ‘asymmetry’, *lā-māʾī* ‘anhydrous’, *lā-sāmī* ‘anti-Semitic’, *lā-filizz* ‘non-metal’, etc.

Generally speaking, it may be stated that Arab language academies and a number of Arab linguists and language reformers occupy a conservative position regarding *naḥt* and its viability as a means of lexical expansion, their main argument being that it is uncharacteristic of the native morphological system. There are others, however, who feel that, in a society undergoing an everlasting process of change and development, the need to expand the lexical stock of the language is pressing enough to warrant some sort of adaptation.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Ibn Fāris, *Muʿjam* = ʾAbū l-Ḥusayn ʾAḥmad Ibn Fāris, *Muʿjam maqāyīs al-luġa*. Ed. ʿAbd as-Salām Muḥammad Hārūn. 6 vols. Cairo: Dār al-Jīl.

——, *Ṣāḥibī* = ʾAbū l-Ḥusayn ʾAḥmad Ibn Fāris, *aṣ-Ṣāḥibī fī fiqh al-luġa wa-masāʾilihā wa-sunan al-ʿArab fī kalāmihā*. Ed. O.F. at-Tabbāʿ. Beirut: Maktabat al-Maʿārif.

SECONDARY SOURCES

Ali, Abdul Sahib Mehdi. 1987. *A linguistic study of the development of scientific vocabulary in Standard Arabic*. London and New York: Kegan Paul International.

ʾAmīn, A. 1958. *al-Ištiqāq*. Cairo: Lajnat at-Taʾlīf wa-t-Tarjama wa-n-Našr.

Baʿalbakī, Munīr. 2002. *al-Mawrid: English–Arabic Dictionary*. Beirut: Dār al-ʿIlm li-l-Malāyīn.

Cairo Arab Language Academy. 1953. *Majalla Majmaʿ al-Luġa al-ʿArabiyya*, 6. Cairo.

Ḥuṣrī, Sāṭiʿ al-. 1958. *ʾĀrāʾ wa-ʾaḥādīṯ fī l-luġa wa-l-ʾadab*. Beirut.
Jawād, Muṣṭafā. 1955. *al-Mabāḥiṯ al-luġawiyya fī l-ʿIrāq*. Cairo.
Jirjis, Ramsīs. 1961. "an-Naḥt fī l-ʿarabiyya". *Majalla Majmaʿ al-Luġa al-ʿArabiyya* 13.61–76.
Mubārak, Muḥammad al-. 1964. *Fiqh al-luġa wa-xaṣāʾiṣ al-ʿarabiyya*. 2nd ed. Beirut: Dār al-Fikr al-Ḥadīṯ.
Nafūs, T.A. 1985. *Dictionary of applied scientific terms*. Mosul: Mosul University Press.
Qaddūr, A.M. 1993. *Madxal ʾilā fiqh al-luġa al-ʿarabiyya*. Beirut: Dār al-Fikr al-Muʿāṣir.
Ṣāliḥ, S. aṣ-. 1981. *Dirāsāt fī fiqh al-luġa*. 9th ed. Beirut: Dār al-ʿIlm li-l-Malāyīn.
Tarzī, Fuʾād Ḥannā. 1967. *al-Ištiqāq*. Beirut: Jāmiʿa Bayrūt al-ʾAmrīkiyya, Kulliyyat al-ʿUlūm wa-l-ʾĀdāb.
Wāfī, ʿAlī ʿAbd al-Waḥīd. 1972. *Fiqh al-luġa*. 7th ed. Cairo: Dār Naḥḍa Miṣr li-ṭ-Ṭabʿ wa-n-Našr.
Xaṭīb, ʿAḥmad Šafīq al-. 1987. *A new dictionary of scientific and technical terms*. Beirut: Librairie du Liban.
Yaʿqūb, I.B. 1986. *Fiqh al-luġa al-ʿarabiyya wa-xaṣāʾiṣuhā*. Beirut: Dār al-ʿIlm li-l-Malāyīn.

Abdul Sahib Mehdi Ali (University of Sharjah)

Computational Linguistics

1. Introduction

A general, but operational and even corpus-based, definition of 'linguistics' is: "the study of the way in which language works" (Collins 1997:972–II). Here, the scope of 'language' will be narrowed down to 'Arabic'. Arabic 'linguistics', then, is "the study of the way in which Arabic works". Adding the pre-modifier 'computational', Arabic computational linguistics presupposes two alternatives: (a) the study of the way in which Arabic works, while using the potentiality of computational means; or (b) the study of how computational means are used to process Arabic.

The first alternative mainly concerns end-users of of-the-shelf products, available on high-storage devices or via the Web (→ Internet). Core business is: Arabic (mono- or multilingual) text processing, spelling checkers, spreadsheets, databases, optical character recognition (OCR) software, Arabic Internet browsers, search engines, e-mail facilities, and text or web-page translators. Secondary tools are, for example, Arabic electronically available general or domain-specific dictionaries, concordance and frequency programs, as well as collections of 'raw' (authentic data) or 'annotated' (preprocessed and tagged data) Arabic speech and text corpora (→ corpus linguistics). More information about this type of product can be found in Ali (1988), Ditters (1989–1990, *forthcoming*), Al-Sulaiti and Atwell (2003), Al-Sulaiti (2004), and other publications. The most up-to-date information is available via the Web.

The second alternative comprises developers of the above mentioned products as well as other specialists sharing their interest in the processing of Arabic. At the commercial level, the kitchen door remains, usually, closed. From the market (collaborative projects) as well as from management circles, pressure is put on academic staff to also produce socially relevant short-range product results. The commercial level offers, as yet, a good market for Arabic computational linguistics students and professionals alike. Up-to-date information about this second alternative comes from the Web, and, only interesting for wholesalers, from periodical fairs.

At the academic level, there are psycholinguists, interested in Arabic language understanding, knowledge representation, simulation and stimulation; electrical engineers, interested in Arabic speech recognition, speech generation, system-controlled answering systems, text-to-speech and speech-to-text conversion systems; and computer scientists, interested in the whole field of Arabic language processing (→ automatic language processing). Stimulated by the success of the Web, they are speeding up research on automatic language identification, information retrieval, information extraction, and machine translation. Computational linguists, of course, are interested in Human Language Technology (HLT) and, more specifically, the computational processing of Arabic. The basic information at this level can be obtained from publications, available from specialized periodicals, bookstores and, more and more, via the Web.

The second alternative is the key issue of this entry: what is, linguistically speaking, going on in the computational processing of Arabic? In the following sections, attention is given to an overview of the field, together with a brief history of Arabic computational linguistics; the state of the art as monitored at 2004; some data, tools, resources, and references.

A final remark on the scope of the language under consideration is required. Usually a distinction is made between the written and the spoken variety of Arabic. The written variety (with synonyms, such as Literary Arabic, Standard, or → Modern Standard Arabic, and Contemporary Arabic) is taught and learned at school. It is used for any form of written communication and is the *lingua franca* for educated native speakers of any of the Arabic colloquial dialects. As far as the spoken variety is concerned, the spoken varieties of Literary Arabic should be distinguished from the colloquial dialects, which are almost exclusively learned within the domestic environment of the user community.

In what follows, the emphasis is on linguistic research concerning the processing of written or Modern Standard Arabic (with some sidesteps to theoretical computational linguistics applied to Arabic colloquial varieties). For written Classical Arabic (with or without spoken approximations), including special purpose collections such as the *Qur'ān* and old poetry recitations, see → *Qur'ān*, → poetry). For spoken varieties of Modern Standard Arabic and modern Arabic dialects see → automatic speech processing.

2. Issues in Arabic computational linguistics

A good general introduction to our field of interest remains Winograd (1983); and specifically for Arabic Ennaji and Sadiqi (1994). A good overview of techniques for capturing linguistic knowledge on speech and language processing is given in Jurafsky and Martin (2000). They describe the standard toolkit of computer sciences, mathematics, and computational linguistics including deterministic and non-deterministic procedural and declarative models. Procedural models are finite state automata and transducers, weighed automata, Markov (MM) and hidden Markov (HMM) models.

Declarative models are regular grammars and regular relations, context-free grammars and feature augmented grammars. Knowledge of semantics, pragmatics, and discourse can be captured by logic models, such as first order logic, feature structures, semantic networks, and conceptual dependency. Both procedural and declarative models can and should be expanded with probabilities.

The probability theory and statistics are key issues in the, equally good, overview on natural language processing by Manning and Schütze (2000). Both these authors and Jurafsky and Martin (2000) underline the importance of corpus-based research, not only as an advertisement for the use of authentic data, but also as a test bed for machine learning algorithms for the automated capture and computational representation of linguistic knowledge.

With the processing of (modern written) Arabic as a main topic, one has to account for the production and recognition of words in speech (phonetics and phonology); the production and recognition of words in context (morphology); the production and recognition of words in structured sequences (syntax); the knowledge of the meaning of component words and compound words (lexical and compositional semantics); pragmatics; and the knowledge of discourse conventions. Finally, the main task is to find ways to solve ambiguity at all these individual levels as well as in the overall representation of language.

In Arabic computational linguistics, partially or fully formalized descriptions of Arabic in one or more of the above mentioned fields are crucial. They should follow the principles of a specific linguistic approach (or combine different linguistic theories) and satisfy the conditions of being coherent, consistent, and methodologically adequate, on one hand, and able to be implemented and processed by computational means, on the other. Therefore, a subdivision can be made into descriptions, implementations, and results of implementations.

From the beginning, research on phonetics and phonology have been combined into morphology. This is not so strange in the case of computational linguistics concerning a language as rich in inflection as Arabic. Moreover, in a less finite-state and more declarative and grammar-based approach, phonology and morphology are frequently integrated in formal descriptions of Arabic syntax. In level-organized approaches, even semantics is accounted for as an extension to the syntax level, be it as yet rudimentary. The field of phonology and morphology is of particular importance for text-to-speech and speech-to-text applications as well as in speech generation and speech recognition. In what follows Arabic phonetics, phonology, and morphology are discussed together.

2.1 *Descriptions*

2.1.1 Phonology and morphology

Within computational linguistics, the description of Arabic phonetics, phonology, and morphology is usually worded in terms such as two-level (representing a word on a lexical and a surface level), finite-state (using automata to perform the mapping between the two levels), and non-concatenative (for Arabic or any other root and pattern type language) activity. One finds other terms, synonyms or extensions, such as templatic (= non-concatenative), tiers or multitiered (multi-level), three-level (one on top of the lexical-surface level) (e.g. Narayanan and Hashem 1992, 1993).

Tracing past theoretical developments, a line can be drawn from McCarthy (1979, 1981, 1986, 1990), Kay (1987), via Kiraz (1992, 1996) to a good general overview of research on Arabic computational morphology in the West (Kiraz 1998). In recent developments, Kay (2004) went some distance from statistical and corpus-based approaches of Arabic script-based languages. Kiraz (1997) continued his research of algorithms for the compilation of regular rewrite rules into automata. In the same line of developments must be placed other (successful) attempts (Beesley, Buckwalter, and Newton 1989) to translate the theory into applications.

2.1.2 Syntax

Coinciding with a theoretical revolution in linguistics (Chomsky 1965, 1982, and many others), important technical developments made personal computational tools for linguistic and other research available and affordable. From then on, the factual testing of a formalized linguistic theory made the difference between an elegant linguistic theory and an equally elegant, but verifiable, linguistic hypothesis.

Within the framework of computational linguistics, the structure of a natural language can be described by means of a non-deterministic, declarative, formal rule-system. Whether such a system should be top-down or bottom-up oriented is, linguistically speaking, of less importance than the fact that possible repetitive patterns in language performance can be described and formalized in context-free phrase-structure rules and tested against new data.

In almost all leading currents and subbranches of 'modern' descriptive linguistics, applications for Arabic have been elaborated (Ditters 1992:54–106). Most of them have the following in common: the concept of constituency (a specific coherency of components); the concept of relationships and dependencies between these components; the concept of relationships and dependencies of one constituent with another within a higher unit of linguistic description; the concept of 'unification' (the gradually filling in of forthcoming data within an analyzing process); the notion of slots and fillers within a structured sequence; and a distinction between key-elements (head/modifier) within the analysis process. In what follows, we mention initiatives to describe Arabic according to modern linguistic concepts and initiatives to process these descriptions computationally.

Partially or fully formalized descriptions of written standard Arabic are: Hartmann (1974) and Al-Khuli (1979) within the Transformational-Generative (TG) approach; an early application of TG on spoken Arabic (Wise 1975); and Aoun (1981), closely following Chomsky's journey from Standard and Revised Extended Standard Theory into the Government and Binding (GB) framework (Aoun and Choueri 1999). A similar development can be witnessed in the work of Ayoub (1981, 1997). However, the TG and GB path faced difficulties in satisfying the earlier mentioned 'implementation and processing' condition.

The work of Fassi Fehri (1982, 1985, 1993) started in the framework of Lexical Functional Grammar (LFG) and continued in the GB-direction. Lahmeur (2004) revisited the 'New-Khalilean' school of Haj-Salah (1989) in the framework of Joshi's (1987) LTAG (Lexical Tree Adjoining Grammar), an approach worth following because of interesting 'implementation' compromises between the Arabic grammatical tradition and modern linguistics (see also Debusmann, Duchier, and Niehren 2005).

Other applications to Arabic include Fillmore's Case Grammar (al-Saffar 1979; al-Waer 1989), Hudson's Word Grammar (Chekili 1982, for Tunisian Arabic), Halliday's scale and category grammar (Systemic Grammar) (al-Karouri 1980), Cole and Sadock's Relational Grammar (Salih 1985), and Dik's Functional Grammar (Moutaouakil 1989, 1993).

The same holds for Montague-like, lexicon based, Categorial Grammar and its offspring, Categorial Unification-based Grammar. A more technical approach, but easily implemented in PROLOG, is Warren and Pereira's Definite Clause Grammar (DCG) (Mehdi 1986, 1987). Equally easy to formalize appears to be the approach in Robinson's Dependency Grammar

(DG). Owens (1988) discusses the description of Arabic in terms of constituents and dependency relations. The notion of 'dependency', interpreted according to the Functional Generative Description of the Prague School, successfully combined this theoretical concept with the processing of authentic data resulting in the Prague Arabic Dependency Treebank (Smrž and Pajas 2004; Hajič, Smrž, Zemánek, Šnaidauf, and Beška 2004).

Finally, the authors of the Generalized Phrase Structure Grammar (GPSG), Gazdar, Klein, Pullum, and Sag (1985:15, n. 6), proudly mention Arabic (Barlow 1984; Edwards 1983) among applications of the GPSG formalism to natural languages. One should add updates such as Yusuf (1983), Achit and Azzoune (2004) and others in GPSG or in its kernel-oriented Head-Driven variety (HPSG). These phrase-structure grammars, as well as the earlier mentioned LFG, DG, and DCG formalizations, meet the 'implementation' condition as does the two-level (constituent + dependency) AGFL formalism (Affix Grammar over Finite Lattices) (Ditters 1991, 1992, 2001, 2003).

2.1.3 Beyond syntax

It would seem to be a logical transition from syntax to semantics, pragmatics, and discourse, but this is far from evident. No formal descriptions of (Arabic) syntax structure, neither those accounting for underlying relationships and dependencies, nor those enriched with a 'semantic-feature' level, go beyond the sentence as linguistic unit of description. Formal text grammars, coherent, consistent, and adequate descriptions of semantic features, and outlines for the formal description of general Arabic language pragmatics and discourse particularities (e.g. a 'formalized' Johnstone 1991) are almost non-existent.

Yet, any unification-based approach, with a semantic feature-level, will certainly facilitate the gradual disambiguation of undesired analysis results, while processing raw authentic Arabic language data. Moreover, developments in natural language processing and general formal semantics are preparing for the 'next' step. Progress on the computational side, in hardware development and language engineering (automatic [Arabic] information retrieval, text summarization, machine translation, and data routing), e.g. the dispatch of incoming data to its most likely place of destination within an organization, have made available intelligent counters and database-systems, idiom and data collections, as well as heuristics to prepare the stage for semantic analysis and language generation. All this may well be of use for the description of layers and dimensions beyond the Arabic syntactic sentence level.

While waiting for the breakthrough, mention should be made of some early initiatives in the right direction: the statistical lexical studies on some renowned Arabic lexical compilations (e.g. Mousa 1973); and Al-Muhtaseb's (1988) M.A. thesis on an Arabic knowledge and semantics module with sketches for programmable semantic networks. Other approaches, too, such as the use of semantic frames (Minsky 1975), logic (Haddad and Yaseen 2001, 2003), as well as a differential componential approach (Zouaghi, Zrigui, and Ben Ahmed 2004) have been used for the representation of meaning in Arabic (see below, and → lexicography).

2.2 *Applications*

In what follows an (incomplete) overview is given of what is going on in Arabic speech processing, Arabic text processing, speech ↔ text processing, and resources development. This impression is mainly based on selections from network searching in the field, from contributions to specialized conferences since 2000, from special interest groups, and via 'advanced' Web searches (→ automatic language processing).

2.2.1 Speech processing

In this process audio signals (speech) are first digitized, analyzed, disambiguated, recognized, approximated, translated into (machine) readable sequences, and then transduced into results. The form of these results varies from actions (giving access to what is behind a door) to other processes, such as Arabic Speech-to-Text conversion; Question-Answering Systems, etc. In this domain, research topics are acoustic analysis (Mawhoub 2004), acoustic modeling (Bayeh, Lin, Chollet, and Mokbel 2004; Vergyri and Kirchhoff 2004), as well as diagnostic and remedial applications (Touri, Detsouli, Benkaddour, and El-Kharroussi 2004).

2.2.2 Text processing

i. Analyzers
In any form of text processing, analysis, at different layers, is the core activity. The short-term objective of the analysis usually labels the process as document processing, (cross-lingual) information retrieval, language identification, etc.

ii. Document processing
The final objective of this process may vary between text abstraction, text categorization (El Kourdi, Bensaid, and Rachidi 2004), text classification (Sawaf, Zaplo, and Ney 2001), language identification (El-Shishiny, Troussov, McCloskey, Takeuchi, Nevidomsky, and Volkov 2004), document indexing (Jilani and Haouala 2001), document routing, text summarization (Lehmam and Bouvet 2004; alRahabi, Mourad and Djioua 2004; Douzidia and Lapalme 2004), and 'simple' information retrieval (see below). Two other issues are document or text enrichment (Stede and Wedel 2001), and information retrieval evaluation (Ouard and Gey 2001).

iii. (Cross lingual) Information Retrieval
In this process the general content of an electronic document is estimated by means of different techniques, such as a statistical lexical approach, stemming techniques (Kadri and Nie 2004) (see 2.2.3), or, on the contrary, feature unification by string matching (Schneider, Mandl, and Womser-Hacker 2004), bilingual dictionary look-up (Hasnah and Evens 2001; Zajac, Malki, and Abdelali 2001), fast surface morphological ana-lysis (Al-Sughaiyer and Al-Kharashi 2000), the listing of section headings, summarizing by paragraph-initial sentences, the isolation of head/modifier pairs (Ditters and Koster 2004), etc. (see also Abdelali, Cowie, and Soliman 2004, for Arabic information retrieval perspectives). Depending on the form in which a conclusion is presented, one speaks of document summarizing, text recognition, document routing, etc.

iv. Morphological analyzer
In this process specific, but finite, paradigm changes within Arabic language information interchange are recognized and translated into relevant information. Usually, the underlying layer (Arabic phonetics and phonemics) has already been integrated in the formal morphological account, resulting in a sound and/or text application. The morphological analyzer itself (whether sound or text oriented) is the backbone for possible next steps such as syntactic parsers (Shaalan, Farouk, and Rafea 2000) and part-of-speech taggers (El-Kareh and al-Ansary 2001; Freeman 2001). Also of interest is somewhat related research concerning modular Arabic grammars (Jaccarini 1999) and the parsing of texts without a lexicon (Gaubert 2001) (see 2.2.3).

v. Machine Translation (MT)
In this process, text data is parsed, resulting in a source language parse tree. In interlingua models (Vauqois 1975, 1979, and for Arabic, 1983), from this parse tree a target language parse tree is generated, the last step before the generation of a sequence in target language words. In another approach, a transfer (e.g. a 'minimal' transfer in Hajič 2002) takes place at the level of the source parse tree and the target parse tree. Mono-directional MT systems comprise a single analyzer (for the parsing of the source language) and a single synthesizer (for the generation of the target language). In bi-directional systems the number is doubled. A rough distinction is made between two different approaches: a rule-based and a statistical approach and some 'in-betweens'.

In a rule-based approach (Azzedine 2004) interaction takes place between different modules responsible for the morphological, syntactic, and semantic analysis as well as for the look-up in idiom and general (or specific domain) lexical databases and the transfer into equivalent words and sequences of words in the target language. In more sophisticated applications, a learning algorithm usually allows for the self-enhancement of the system. In a statistical approach (e.g. Marcu, Fraser, Wong, and Knight 2004) training algorithms identify, align, and create a parallel corpus and extract from it probabilistic translation dictionaries, patterns, and rules to be used to translate new texts.

At an early stage, machine translation was taken up in the Arab world (Wahab and Sienny 1986; Aref 1996). Research in this field has been intensified since then (e.g. Chalabi 2001; Yaseen a.o. 2001; Schafer and Yarowsky 2003; Soudi 2004). More recently, attention is being paid to the evaluation of Arabic MT systems (Hamaada 2004; Al-Ansary and El-Kareh 2004).

2.2.3 Resources

Ali (2003), one of the pioneers within (commercial) Arabic natural language processing, anticipated the framework, proposed by Diekema (2004) and the initiatives of ELDA (2005), aiming at a comprehensive listing of existing Arabic language resources and an inventory of those still lacking and the best source of information so far.

i. Analyzers, parsers, taggers and NLP processing environment

One of the first available Arabic morphological analyzers was ALPNET (Beesley, Buckwalter, and Newton 1989), a PC-KIMMO like two-level implementation, now using the Finite-State tools of Xerox (Beesley 2001). Buckwalter (2002) continued with the Arabic Morphological Analyzer version 1.0. Habash (2004) redirected the Buckwalter analyzer into a morphological generator. The morphological analyzer Morpho3 (Attia 2000) is a hybrid mixture of rules and statistics, being further developed by Research and Development International in Egypt. Darwish (2002) developed a 'shallow' analyzer based on automatically derived rules and Freeman (2001) adapted Brill's POS (Parts-of-Speech) tagger for Arabic. A bottom-up chart parser has been developed by Othman, Shaalan, and Rafea (2003). In the framework of the international project DIINAR-MBC (Dichy 1998), aiming at the development of an Arabic toolkit, the morpho-syntactic parser AraParse (Ouersighni 2001) and the syntax parser Larusa (Ditters 2000) have been developed in the AGFL processing environment (AGFL 2005).

ii. Speech databases

Orientel is a recently completed project aiming at the development of speech-driven applications for the Mediterranean and the Middle East (Iskra a.o. 2004). As far as speech-corpora are concerned, the dialectal Arabic telephone speech corpus (Maamouri a.o. 2004), the Broadcast News Speech corpus (Choukri a.o. 2004) and the (commercially developed) Sakhr speech databases (Ghali 2004) should be mentioned.

iii. Machine Translation

An English to Arabic system was designed and implemented by AlNeami (1996). Among a number of (commercially) available MT-systems for Arabic (ELDA 2005) we refer here only to the bi-directional Language Weaver Arabic ↔ English MT (Marcu, Fraser, Wong, and Knight 2004) and Cimos's an-Nakel al-Arabi Arabic ↔ English/ French/German MT-system (Azzedine 2004). Via an advanced search, Web-based Arabic MT-systems (e.g. Systran 2005) can easily be traced.

iv. Lexical resources

Sakhr developed an excellent electronic Arabic lexicon (Chalabi 2004), but are somewhat irresponsive regarding the provision of detailed information about the scientific semantic part of their research. For further developments we therefore rely on publications concerning bi- or multilingual dictionaries with enriched lexical information (El-Katib and Black 2001, 2004; Lelubre 2001), (corpus-driven) lexical databases (Ghazali and Braham 2001), and dictionary definitions with semantic specifiers (Dichy 2000, 2001).

v. Text databases (→ corpus linguistics)

Besides the resources listed in ELDA (2005), reference may be made to research on the exploitation of annotated Arabic text corpora (van Mol 2001), as well as to collections like the Penn Arabic Treebank (Maamouri, Bies, Buckwalter, and Jin 2003–2004; Maamouri, Buckwalter, and Cieri 2004) and the Prague Arabic dependency Treebank 1.0 (Hajič a.o. 2004). Infor-mation about the composition, representativity, verifiability, and related details of a corpus (Zemánek 2001) is always of crucial importance. The importance of monolingual Arabic text-corpora, bi- or multilingual (parallel) corpora (Samy, Sandoval, and Guirao 2004, for Spanish and Arabic) as 'learning' input for any kind of application becomes more and more evident (Diab 2004).

vi. Text-to-Speech conversion

In this process written Arabic is parsed (Ramsay and Mansour, 2004) and synthesized into spoken Arabic (Youssef and Emam 2004). There are different techniques and approaches for the speech synthesis part (Eldin, Abdel Nour, and Rajouani 2001; Saidane, Haddad, Zrigui, and Ben Ahmed 2004; Zemirli and Khabet 2004).

vii. Question answering systems

In the first system (Hammo, Abu-Salem, Lytinen, and Evens 2002), a natural language question is analyzed by information retrieval techniques into words in order to be matched with a small set of ranked documents. Another interesting project is a speech-to-speech translation project (Ganjavi, Georgiou, and Narayanan 2004) aiming at facilitating communication, in this case real-time English–Persian, but modifiable for any other language using the Arabic script. Finally, there is the development of an Arabic chatbot, a special case of question answering system, giving answers from the *Qurʾān* (Abu Shawar and Atwell 2004).

BIBLIOGRAPHICAL REFERENCES

Abdelali, Ahmed, Jim Cowie, and Hamdy Soliman. 2004. "Arabic information retrieval perspectives". Boualem and Chenfour (2004).

Abu Shawar, Bayan and Eric Atwell. 2004. "An Arabic chatbot giving answers from the *Qurʾan*". Boualem and Chenfour (2004).

Achit, Abdelmadjid and Hamid Azzoune. 2004. "Comparative survey between the GPSG and PSG grammars applied to the Arabic language". Abstract available at: <http://www.ulcl.leidenuniv.nl/index.php3?m=3&c=96>.

AGFL. 2005. <http://www.cs.ru.nl/agfl/>.

Ahmed, Mohamed Attia. 2000. *A large-scale computational processor of the Arabic morphology, and applications.* M.A. diss. Faculty of Engineering, Cairo University.

Al-Ansary, Sameh and Seham El-Kareh. 2004. "Arabic-English Machine Translation systems: Discrepancies and implications". Boualem and Chenfour (2004).

Ali, Nabil. 1988. *al-Luġa al-ʿarabiyya wa-l-ḥāsūb (dirāsa baḥṯiyya)*. Cairo: Taʿrīb.

——. 2003. "The second wave of Arabic natural language processing: A content perspective". *Proceedings of the Meeting of the Expert Group on Promotion of Digital Arabic Content, 3–5 June, 2003 in Beirut, Lebanon.* <www.escwa.org.lb/wsis/meetings/3–5june/ docs/18.pdf>.

Alimi, Adel M. and Lamia Hadrich Belguith (eds.). 2000. *Corpora and natural language Processing.* (= *Proceedings of the International Conference on Artificial and Computational Intelligence for Decision, Control and Automation in Engineering and Industrial Applications (ACIDCA' 2000), 22–24 March 2000*, VI.) Monastir, Tunisia.

Al-Karouri, Abdel Muneim. 1980. *The Classical Arabic verb inflection: A WP grammar (with an introductory phonemic investigation).* Ph.D. diss., University of Edinburgh.

Al-Khuli, Muhammad A. 1979. *A contrastive transformational grammar: Arabic and English.* Leiden: E.J. Brill.

Al-Muhtaseb, Husni. 1988. *A natural Arabic language understanding system.* M.Sc. diss., King Fahd University of Petroleum and Minerals.

AlNeami, Ahmed. 1996. *Design and implementation of an English to Arabic machine translation system.* Ph.D. diss., Department of Computer Science, Sheffield University.

Al-Otaibi, Abdulhadi (ed.). 1989. *Proceedings of the 2nd Conference on Arabic Computational Linguistics.* Kuwait.

Alrahabi, Motasem, Ghassan Mourad, and Brahim Djioua. 2004. "Filtrage sémantique de textes en arabe en vue d'un prototype de rèsumè automatique". Boualem et Chenfour (2004).

Al-Saffar, Abdul-Emir. 1979. *A semantico-syntactic study of Classical Arabic in Case Grammar terms.* Ph.D. diss., School of Oriental and African Studies, University of London.

Al-Sughaiyer, Imad A. and Ibrahim A. Al-Kharashi. 2000. "An efficient Arabic morphological analysis technique for information retrieval systems". Alimi and Belguith (2000:61–65).

Al-Sulaiti, Latifa. 2004. "Arabic corpora and corpus ana-lysis tools (Tutorial)". Boualem and Chenfour (2004:547–552).

——, and Eric Atwell. 2003. *The design of a corpus of contemporary Arabic (CCA).* Leeds: School of Computing, Research Report Series, University of Leeds.

Al-Waer, Mazen. 1989. "The syntactic and semantic analysis of the generation of the Arabic sentence". Al-Otaibi (1989:172–220).

Aoun, Joseph. 1981. "ECP, move α, and subjacency". *Linguistic Inquiry* 12:4.637–645.

——, and Lina Choueiri. 1999. "Modes of interrogation". *Perspectives on Arabic linguistics*, XII, ed. Elabbas Benmamoun, 7–26. Amsterdam and Philadelphia: J. Benjamins.

Aref, Mustafa (ed.). 1996. *Machine translation. Proceedings of the 1st KFUPM Workshop on Information and Computer Science, June 9, 1996.* Dhahran: King Fahd University of Petroleum and Minerals.

Ayoub, Georgine. 1981. "Structure de la phrase verbale en arabe standard". Thèse de doctorat du 3$^{\text{ème}}$ cycle, University of Paris 7. *Analyses Théories* 2:1 (*Syntaxe*), i–viii, 1–367.

——. 1997. *La question de la phrase nominale en arabe littéraire: Prédicats, figures, catégories.* 2 vols. Lille: Éditions du Septentrion.

Azzedine, Mohammed. 2004. "Meaning-based translation: CIMOS semantic analyzer". Nikkhou (2004:164–167).

Barlow, Michael. 1984. "Control agreement and syntax". *Proceedings of the Spring Meeting of the Linguistics Association of Great Britain, 1984*, 28–30. Hull: University of Hull.

Bayeh, Rania, S. Lin, Gerard Chollet, and Chafic Mokbel. 2004. "Reconnaissance de l'arabe parlé à partir de modèles acoustiques du français". Boualem and Chenfour (2004).

Beesley, Kenneth R. 2001. "Finite-state morphological analysis and generation of Arabic at Xerox Research: Status and plans in 2001". Krauwer (2001:1–8).

——, Tim Buckwalter, and Stuart Newton. 1989. "Two-level finite-state analysis of Arabic morphology". Ubaydli (1989:63–72).

Boualem, Malek and Noureddine Chenfour (eds.). 2004. *Compte-rendu des Journées d'étude sur la parole (JEP) et Traitement automatique des langues naturelles (TALN) (session arabe), les 19–22 avril, 2004, Fes: University Sidi Mohamed Ben Abdellah.* <http://www.lpl.univ-aix.fr/jep-taln04/proceed/actes/ arabe.htm>.

Buckwalter, Tim. 2002. *Buckwalter Arabic Morphological Analyzer Version 1.0.* Linguistic Data Consortium, catalog number: LDC2002L49 and ISBN 1-58563-257-0.

Cavalli-Sforza, Violetta, Alon Lavie, and Nizar Habash (eds.). 2003. *The MT Summit 9th Workshop on Machine Translation for Semitic Languages: Issues and approaches, Tuesday September 23–28, 2003*. New Orleans.

Chalabi, Achraf. 2001. "Sakhr Web-based Arabic-English MT engine". Krauwer (2001:137–139).

Chekili, Ferid. 1982. *The morphology of the Arabic dialect of Tunis*. Ph.D. diss., University College, London.

Choukri, Khalid, Mahtab Nikkhou, and Niklas Paulsson. 2004. "Net-DC BNSC, an Arabic broadcast news speech corpus". Nikkhou (2004:61–64).

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Mass.: MIT Press.

——. 1982. *Lectures on government and binding: The Pisa lectures*. Dordrecht: Foris.

Collins Cobuild English dictionary. 1997. London: HarperCollins.

Darwish, Kareem. 2002. "Building a shallow Arabic morphological analyzer in one day". Rosner and Wintner (2002).

Debusmann, Ralph, Denys Duchier, and Joachim Niehren. 2004. "The XDG grammar development kit". Van Roy (2005:190–201).

Diab, Mona. 2004. "The feasibility of bootstrapping an Arabic WordNet leveraging parallel corpora and an English WordNet". Nikkhou (2004:71–78).

Dichy, Joseph. 1998. "Some aspects of the DIINAR-MBC research program". Ubaydli (1989:2.8.1–5).

——. 2000. "Morphosyntactic specifiers to be associated to Arabic lexical entries: Methodological and theoretical aspects". Alimi and Belguith (2000: 55–60).

——. 2001. "On lemmatization in Arabic: A formal definition of the Arabic entries of multilingual lexical databases". Krauwer (2001:23–30).

Diekema, Anne R. 2004. "Preliminary lexical framework for English–Arabic semantic resource construction". Farghali and Megerdoomian (2004: 10–14).

Ditters, Everhard. 1998–1990. "Arabic". *The humanities computing yearbook: A comprehensive guide to software and other resources*, ed. Ian Lancashire, 285–292. Oxford: Clarendon.

——. 1991. "A modern standard Arabic sentence grammar". *Bulletin d'Etudes orientales* 43.197–236.

——. 1992. *A formal approach to Arabic syntax: The noun phrase and the verb phrase*. Ph.D. diss., University of Nijmegen.

——. 2000. "The syntactic component of the Larusa parser for voweled or unvoweled written Arabic". CD-Rom, ed. Joseph Dichy, *DIINAR-MBC project*, INCO No 961791.

——. 2001. "A formal grammar for the description of sentence structure in modern standard Arabic". Krauwer (2001:31–37).

——. 2003. "An AGFL for the description of non-coinciding phrasal heads". *Proceedings of the 7th World Multi-conference on Systemics, Cybernetics and Informatics, July 27–30, 2003*, ed. Nagib Callaos, William Lesso, Belkis Sánchez, and Elizabeth Hansen, 107–112. Orlando, Fla.: International Institute of Informatics and Systemics.

——. Forthcoming. "Technologies for Arabic language teaching and learning". *A handbook for Arabic language teaching professionals*, ed. Kassem Wahba, Liz England, and Zeinab Taha. Mahwah, N.J.: Lawrence Erlbaum Associates.

——, and Kees Koster. 2004. "Transducing Arabic text to head/modifier pairs". Nikkhou (2004: 148–154).

Douzidia, Fouad Soufiane, and Guy Lapalme. 2004. "Lakhas, an Arabic summarization system". *Proceedings of the Document Understanding Conference (DUC2004)*. <http://www.iro.umontreal.ca/~lapalme/publications/ConfsSansArbitrage/Lakhas_duc04.pdf.>

Edwards, Malcolm. 1983. "Relative clauses in Egyptian Arabic". Paper presented at the Autumn Meeting of the Linguistics Association of Great Britain, University of Newcastle September 21–23, 1983.

ELDA. 2005. "Survey on Arabic language resources and tools in the Mediterranean countries". *Updated overview of Arabic language resources*, ed. Mahtab Nikkhou and Khalid Choukri. <www.nemlar.org/publications/nemlar-report-survey-final-web.pdf>.

El-Kareh, Seham and Sameh Al-Ansary. 2000. "An Arabic interactive multi-feature POS-tagger". Alimi and Belguith (2000:83–88).

Elkateb, Sabri and Wiliam Black. 2001. "Towards the design of an English-Arabic terminological and lexical knowledge base". Krauwer (2001:113–118).

El Kourdi, Mohamed, Amine Bensaid, and Tajje-eddine Rachidi. 2004. "Automatic Arabic document categorization based on the naïve Bayes algorithm". Farghaly and Megerdoomian (2004:51–58).

El-Shishiny, Hisham, Alexander Troussov, D.J. McCloskey, Mayo Takeuchi, Alex Nevidomsky, and Pavel Volkov. 2004. "Word fragments based Arabic language identification". Nikkhou (2004:90–93).

Ennaji, Moha and Fatima Sadiqi. 1994. *Applications of modern linguistics*. Casablanca: Afrique Orient.

Farghali, Ali and Karine Megerdoomian. 2004. *Proceedings of the Workshop on Computational Approaches to Arabic Script-based Languages, 28 August 2004*, Geneva.

Fassi Fehri, Abdelkader. 1982. *Linguistique arabe: Forme et interprétation*. Rabat: Faculté des Lettres.

——. 1985. *al-Lisāniyyāt wa-l-luġa al-ʿarabiyya: Namāḏij tarkībiyya wa-dalāliyya*. Casablanca: Dār Tūbqāl li-n-Našr.

——. 1993. *Issues in the structure of Arabic clauses and words*. Dordrecht: Kluwer Academic Publishers.

Freeman, Andrew. 2001. "Brill's POS tagger and a morphology parser for Arabic". Krauwer (2001: 148–154).

Ganjavi, Shadi, Panayiotis G. Georgiou, and Shrikanth Narayanan. 2004. "A transcription scheme for languages employing the Arabic script motivated by speech processing application". Farghali and Megerdoomian (2004:59–65).

Gaubert, Christian. 2001. *Stratégies et règles minimales pour un traitement automatique de l'arabe*. Ph.D. diss., University Aix-Marseille 1.

Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum, and Ivan Sag. 1985. *Generalized phrase structure grammar*. Oxford: Blackwell.

Ghali, Michael. 2004. "Sakhr speech databases". Nikkhou (2004:144–147).

Ghazali, Salem and Abdelfattah Braham. 2001. “Dictionary definitions and corpus-based evidence in Modern Standard Arabic”. Krauwer (2001:51–57).

Habash, Nizar. 2004. “Large scale lexeme based Arabic morphological generation”. Boalem and Chenfour (2004).

Haddad, Bassam and Mustafa Yaseen. 2001. “Towards understanding Arabic: A logical approach for semantic representation”. Krauwer (2001:44–50).

——. 2003. “Towards semantic composition of Arabic: A λ-DRT based approach”. Cavalli-Sforza, Lavie, and Habash (2003). <http://www.2.cs.cmu.edu/~alavie/Sem-MT-wshp/Haddad+Yaseen_paper. pdf>.

Haj-Salah, Abdurrahman. 1989. “al-Madrasa al-xalīliyya al-ḥadīṯa wa-mašākil ʿilāj al-ʿarabiyya bi-l-ḥāsūb”. Al-Otaibi (1989:410/1–43).

Hajič, Jan. 2002. “Tectogrammatical representation: Towards a minimal transfer in machine translation”. *Proceedings of the 6th International Workshop on Tree Adjoining Grammar and Related Frameworks*, 101–110. Venice: University of Venice.

Hajič, Jan, Otakar Smrž, Petr Zemánek, Jan Šnaidauf, and Emanual Beška. 2004. “Prague Arabic dependency treebank: Development in data and tools”. Nikkhou (2004:110–117).

Hamaada, Salwa. 2004. “Evaluation of Arabic machine translation systems”. Paper presented at *NEMLAR International conference on Arabic language resources and tools*, 22–23 September 2004. Cairo: Magnet Creative Communications.

Hammo, Bassam, Hani Abu-Salem, Steven Lytinen, and Martha Evens. 2002. “QARAB: A question answering system to support the Arabic language”. Rosner and Wintner (2002).

Hartmann, Regina. 1974. *Untersuchungen zur Syntax der arabischen Schriftsprache: Eine generativ-transformationelle Darstellung*. Wiesbaden: O. Harrassowitz.

Hasnah, Ahmad and Martha Evens. 2001. “Arabic/English cross language information retrieval using a bilingual dictionary”. Krauwer (2001:81–87).

Hudson, Richard. 1984. *Word grammar*. Oxford: Blackwell.

Iskra, Dorota, Khalid Choukri, Ossama Emam, Albino Nogueiras, Oren Gedge, Rainer Siemund, and Robrecht Comeyne. 2004. “OrienTel: More than another SpeechDat for the Mediterranean and Middle East”. Nikkhou (2004:50–54).

Jaccarini, André. 1999. *Grammaires modulaires de l’arabe*. Ph.D. diss., University of Paris 4.

Johnstone, Barbara. 1991. *Repetition in Arabic discourse: Paradigms, syntagms, and the ecology of language*. Amsterdam: J. Benjamins.

Joshi, Aravind. 1987. “Introduction to tree adjoining grammar”. *The mathematics of language*, ed. Alexis Manaster Ramer, 87–114. Amsterdam and Philadelphia: J. Benjamins.

Jurafsky, Daniel and James H. Martin. 2000. *Speech and language processing*. Upper Saddle River, N.J.: Prentice Hall.

Kadri, Youssef and Jian-Yun Nie. 2004. “Traduction des requêtes pour la recherche d’information translinguistique anglais–arabe”. Boualem and Chenfour (2004). <http://www .lpl.univ-aix.fr/jeptalno4/proceed/actes/arabe2004/TFYK31.pdf>.

Kay, Martin. 1987. “Nonconcatenative finite-state morphology”. *Proceedings of the 3rd Conference of the European Chapter of the Association for Computational Linguistics*, ed. Bente Maegaard, 2–10. Copenhagen: University of Copenhagen.

——. 2004. “Arabic script-based languages deserve to be studied linguistically”. Farghali and Megerdoomian (2004:42).

Kiraz, George A. 1992. “Semitic morphology: A linguistic and computational account”. Ubaydli (1992:6.4.1–6.4.10).

——. 1996. *Computational approach to nonlinear morphology*. Ph.D. diss., University of Cambridge.

——. 1997. “Compiling regular formalisms with rule features into finite-state automata”. *Proceedings of the 8th Conference of the European Chapter of the Association for Computational Linguistics*, 329–336. Madrid.

——. 1998. “Arabic computational morphology in the West”. Ubaydli (1998:2.2.1–11). (Also available at <www.bell-labs.com/project/tts/gkiraz.html>.)

——. 2001. *Computational approach to nonlinear morphology*. New York: Cambridge University Press.

Krauwer, Steven (ed.). 2001. *Proceedings of the Workshop on Arabic Language Processing: Status and Prospects: 39th Annual Meeting of the Association for Computational Linguistics (ACL-01) and 10th Conference of the European Chapter, July 6th, 2001*. Toulouse: University of Toulouse, Social Sciences.

Labed Jilani, Lamia and Hichem Haouala. 2001. “A tool for Arabic documents indexing and retrieval from a Web virtual library”. Krauwer (2001: 88–94).

Lahmeur, Karim. 2004. “Parsing Arabic texts with LTAG”. Abstract available at: <http://www.ulcl.leiden-univ.nl/index.php3?m=3&c=96>.

Lehmam, Abderrafih and Philippe Bouvet. 2004. “Watch application, summarization and syndication in Arabic”. Nikkhou (2004:157–162).

Lelubre, Xavier. 2001. “A scientific Arabic terms base: Linguistic approach for a representation of lexical and terminological features”. Krauwer (2001: 66–72).

Maamouri, Mohamed, Anne Bies, Tim Buckwalter, and Hubert Jin. 2003–2004. Arabic Treebank. Linguistic Data Consortium, catalog numbers: LDC2003T06, LDC2004T02, LDC2004T11, LDC2004E71.

——, Tim Buckwalter, and Christopher Cieri. 2004. “Dialectal Arabic telephone speech corpus: Principles, tool design and transcription conventions”. Nikkhou (2004:55–60).

——, Ann Bies, Tim Buckwalter and Wigdan Mekki. 2004. “The Penn Arabic treebank: Building a large-scale annotated Arabic corpus”. Nikkhou (2004: 102–109).

Manning, Christopher D. and Hinrich Schütze. 2000. *Foundations of statistical natural language processing*. Cambridge, Mass.: MIT Press.

Marcu, Daniel, Alex Fraser, William Wong, and Kevin Knight. 2004. “Language weaver Arabic → English MT”. <www.languageweaver.com>.

Mawhoub, Mourad. 2004. “Prosodie et ordre des constituants dans l’énoncé en arabe standard moderne”. Boualem and Chenfour (2004).

McCarthy, John. 1979. *Formal problems in Semitic phonology and morphology*. Cambridge, Mass.: MIT Press.

——. 1981. “A prosodic theory of non-concatenative morphology”. *Linguistic Inquiry* 12.373–418.

——. 1986. “OCP effects: gemination and antigemination”. *Linguistic Inquiry* 17.207–263.

——, and Alan Prince. 1990. “Prosodic Morphology and Templatic Morphology”. *Perspectives on Arabic linguistics*, II, ed. Mushira Eid and John McCarthy, 1–54. Amsterdam and Philadelphia: J. Benjamins. [See also McCarthy and Prince 1996.]

——. 1996. “Prosodic morphology and templatic morphology”. *The Handbook of phonological theory*, ed. John A. Goldsmith, 318–366. Oxford: Blackwell.

Mehdi, Saad A. 1986. “Arabic language parser”. *International Journal of Man-Machine Studies* 25.593–611.

——. 1987. *Computer interpretation of Arabic*. Exeter: University of Exeter Press.

Mousa, Ali H. 1973. *Dirāsa ʾiḥṣāʾiyya li-juzūr muʿjam aṣ-Ṣiḥāḥ (bi-stixdām al-kumbyūtir)*. Kuwait: Maṭbūʿāt Jāmiʿat al-Kuwayt.

Moutaouakil, Ahmed. 1989. *Pragmatic functions in a functional grammar of Arabic*. Dordrecht and Providence, R.I.: Foris.

——. 1993. *al-Waḏīfa wa-l-bunya: Muqārabāt waḏīfiyya li-baʿḍ qaḍāyā t-tarkīb fī l-luġa al-ʿarabiyya*. Rabat: Okad Editions.

Narayanan, Ajit and Lama Hashem. 1992. “A three-level finite-state model for Arabic morphology”. Ubaydli (1992:6.2.1–6.2.23).

——, and Lama Hashem. 1993. “On abstract finite-state morphology”. *Proceedings of the 6th Conference of the European Chapter of the Association for Computational Linguistics 21–23 April, 1993*, ed. Steven Krauwer, Michael Moortgat, and Louis des Tombe, 297–304. Utrecht: OTS, University of Utrecht.

Nikkhou, Mahtab (ed.). 2004. *NEMLAR International conference on Arabic language resources and tools, 22–23 September, 2004*. Cairo: Magnet Creative Communications.

Othman, Emam, Khaled Shaalan, and Ahmed Rafea. 2003. “A chart parser for analyzing modern standard Arabic”. Cavalli-Sforza, Lavie, and Habash (2003). <http://www-2.cs.cmu.edu/~alavie/Sem-MT-wshp/Othman+Shaalan+Rafea_paper.pdf>.

Ouard, Douglas W. and Fredric C. Gey. 2001. “The Trec-2001 Arabic information retrieval evaluation”. Krauwer (2001:95–97).

Ouersighni, Riadh. 2001. “A major offshoot of the DIINAR-MBC project: AraParse, a morphosyntactic analyzer of unvoweled Arabic texts”. Krauwer (2001:9–16).

Owens, Jonathan. 1988. *The foundations of grammar: An introduction to medieval Arabic grammatical theory*. Amsterdam and Philadelphia: J. Benjamins.

——. 1990. *Early Arabic grammatical theory: Heterogeneity and standardization*. Amsterdam and Philadelphia: J. Benjamins.

Ramsay, Allan and Hanady Mansour. 2004. “The parser from an Arabic text-to-speech system”. Boualem and Chenfour (2004).

Rosner, Mark and Shuly Wintner (ed.). 2002. *Proceedings of the Workshop on Computational Approaches to Semitic Languages, ACL-2002, 11 July, 2002*. Philadelphia: University of Pennsylvania.

Safa, Nasser Eldin, Hanna Abdel Nour, and Abdenbi Rajouani. 2001. “Enhancement of a TTS system for Arabic concatenative synthesis by introducing a prosodic model”. Krauwer (2001:97–102).

Saidane, Tahar, Ahmed Haddad, Mounir Zrigui, and Mohamed Ben Ahmed. 2004. “Réalisation d’un système hybride de synthèse de la parole arabe utilisant un dictionnaire de polyphones”. Boualem and Chenfour (2004).

Salih, Mahmud. 1988. “The inadequacy of Case Grammar: Further evidence from Arabic passives” (abstract). *Proceedings of the 2nd International Conference of the Linguistic Society of Morocco, October 11–14*, ed. Abdelkader Fassi Fehri, 99. Rabat: Faculté des Lettres et des Sciences Humaines.

Samy, Doaa, Antonio Moreno Sandoval, and José M. Guirao. 2004. “An alignment experiment of a Spanish–Arabic parallel corpus”. Nikkhou (2004: 85–89).

Sawaf, Hassan, Jorg Zaplo, and Hermann Ney. 2001. “Statistical classification methods for Arabic news articles”. Krauwer (2001:127–132).

Schafer, Charles and David Yarowski. 2001. “A two-level syntax based approach to Arabic–English statistical machine translation”. Cavalli-Sforza, Lavie, and Habash (2003). <http://www-2.cs.cmu.edu/~alavie/Sem-MT-wshp/Schafer+Yarowsky_paper.pdf>.

Schneider, René, Thomas Mandl, and Christa Womser-Hacker. 2004. “Integration of Arabic to a cross-lingual retrieval tool: Challenges and perspectives”. Boualem and Chenfour (2004).

Shaalan, Khaled, Ahmed Farouk, and Ahmed Rafea. 2000. “Towards an Arabic parser for modern scientific text”. Alimi and Belguith (2000:228–235).

Smrž, Otakar and Petr Pajas. 2004. “MorphoTrees of Arabic and their annotation in the TrEd environment”. Nikkhou (2004:38–41).

Soudi, Abdelhadi. 2004. “Challenges in the generation of Arabic from Interlingua”. Boualem and Chenfour (2004).

Stede, Manfred and Gerhard Wedel. 2001. “Turning Arabic texts into active documents”. Krauwer (2001:119–123).

Systran. 2005. <www.systransoftware.com/index.html>.

Touri, Bouzekri, Mohamed Detsouli, Abdel Benkaddour, and Mohamed ElKharroussi. 2004. “Dysphonie: Apport acoustique aux moyens objectifs”. Boualem and Chenfour (2004).

Ubaydli, Ahmad (ed.). 1989. *Proceedings of the Seminar on Bilingual Computing in Arabic and English*. Cambridge: Literary and Linguistic Computing Centre.

——. 1992. *Proceedings of the 3rd International Conference and Exhibition on Multi-lingual Computing (Arabic and Roman Script), 17–18 April 1992*. Cambridge: Centre of Middle Eastern Studies.

——. 1998. *Proceedings of the 6th International Conference and Exhibition on Multi-lingual Computing (Arabic and Roman Script), 17–18 April 1998*. Cambridge: Centre of Middle Eastern Studies.

Van Mol, Mark. 2001. “Exploring annotated Arabic corpora: Preliminary results”. Krauwer (2001: 94–98).

Van Roy, Peter (ed.). 2005. *Proceedings of the 2nd International Mozart/Oz Conference. Brussels, 7–8 October, 2004*. Heidelberg: Springer-Verlag Berlin.

Vauquois, Bernard. 1975. *La traduction automatique à Grenoble*. Paris: Dunod.

——. 1979. *Aspects of mechanical translation in 1979*. Grenoble: ETA.

——. 1983. "Automatic aided translation and the Arabic languages". *Proceedings of the Conference on Applied Linguistics and Signal and Information Processing*, 157–176. Rabat.

Vergyri, Dimitra and Katrin Kichhoff. 2004. "Automatic diacritization for acoustic modeling in speech recognition". Farghaly and Megerdoomian (2004: 66–73).

Wahab, Abdu r-Razzak and Mahmoud Sieny (eds.). 1986. *Studies in machine translation: Proceedings of the International Workshop on Computer Aided Translation, 16–17 March 1985*. Riyadh: King AbdulAziz City for Science and Technology.

Winograd, Terry. 1983. *Language as a cognitive process*. I. *Syntax*. Reading, Mass.: Addison-Wesley.

Wise, Hilary. 1975. *Transformational grammar of Egyptian Arabic*. Oxford: Blackwell.

Wouden, Ton van der (ed.). 2005. *Proceedings of the Conference on Computational Linguistics in the Netherlands (CLIN 2004) December 17, 2004*. Leiden: University of Leiden. Abstracts available at: <http:// www.ulcl.leidenuniv.nl/index.php3?m=3&c=96>.

Yaseen, Mustafa, Bassam Haddad, Harris Papageorgiou, Stelios Piperides, Hattab Hattab, Nick Theophilopoulos, and Steven Krauwer. 2001. "A term base translator over the web". Krauwer (2001: 58–65).

Youssef, Amr and Ossama Emam. 2004. "An Arabic TTS system based on the IBM trainable speech synthesizer". Boualem and Chenfour (2004).

Yusuf, Harun-Al-Rashid. 1983. *Word order variation in Arabic: A generalized phrase structure grammar analysis*. Ph.D. diss., School of Oriental and African Studies, University of London.

Zajac, Remi, Ahmed Malki, and Ahmad Abdelali. 2001. "Arabic–English NLP at CRL". Krauwer (2001:133–136).

Zemanek, Petr. 2001. "CLARA (Corpus Linguae Arabicae): An overview". Krauwer (2001:111–112).

Zemirli, Z. and S. Khabet. 2004. "TAGGAR: Un analyseur morphosyntaxique destiné à la synthèse vocale de textes arabes voyellés". Boualem and Chenfour (2004).

Zouaghi, Anis, Mounir Zrigui, and Mohamed Ben Ahmad. 2004. "Une structure sémantique pour la recherche d'information translinguistique anglais-arabe". Boualem and Chenfour (2004).

Everhard Ditters
(University of Nijmegen)

Computer Linguistics → Computational Linguistics

Concessive Clause

There is no reference to concessive clauses in Arabic grammars, except where conditional clauses are concerned. Their semantics is broadly diversified, featuring two predominant elements: the incausal and the inconditional relationships. In some cases, they may imply an adversative, alternative, consecutive, or comparative relationship. Of these, adversativity is explicitly accentuated through the use of several particles (*fa-*, *fa-qad*, *fa-ʾinna*, *ʾillā ʾanna*, *wa-lākinna*, or *ġayr ʾanna*). In unreal concessive structures, one can also find *la-* with a *mā* negation particle. The scope of the proposition can be divided into at least three basic categories: real, hypothetical, and unreal. Concessive clauses appear as conjunctive clauses, continued sentences, and so-called → *ḥāl*-sentences/clauses. Furthermore, concessivity is indicated by prepositional phrases (→ prepositions).

The most important conjunctions are: *wa-ʾin*, *ḥattā wa-ʾin*, *wa-law*, *ḥattā wa-law*, *ḥattā law*, *wa-ʾiḏā*, *wa-laʾin*, *raġma ʾanna*, and *maʿa ʾanna*, corresponding to '(al)though' and 'even though' in English and 'obwohl', 'wenn auch', or 'selbst wenn' in German. They signal subtle semantic differences, the meaning of which can only be ascertained from the particular context. With the exception of *raġma ʾanna* and *maʿa ʾanna*, they follow the construction rules of conditional clauses. An example of real and adversative concession is: *wa-ʾin lam yafuz fī l-ġaṭs*, *fa-qad fāza fī s-sibāḥa* 'although he didn't win the diving, he won the swimming' (*Marfaʾ* 7). As opposed to this, a hypothetic concession would be: *wa-ḥattā law baqiya l-jinēh al-ʾistarlīnī qawiyyan, sa-yakūnu min aṭ-ṭabīʿī ʾakṯar bi-n-nisba ʾilaynā ʾan nakūna fī sūq al-yūrū* 'even if the pound sterling were to remain strong, it would be more natural for us to be in the Euro-market' (*Waṭan* 17.10.97, 26). Less frequently, this occurs in unreal statements with a past or present meaning, depending on the context: *ḥattā law ʾaradtum ʾan tanʿazilū ʿannā la-mā ʾamkana* 'even if you would like//had liked to be alone, it would not be possible//would not have been possible' (*Gīla* 133/10). Shortened conditional sentences beginning with *wa-law* should also be analyzed as concessive clauses: *wa-li-ḏālika ḥāwala jahdahu ʾan yanāma wa-law li-niṣf sāʿa*. '{and} therefore he tried to sleep, even if (it were to be) only for half an hour' (*Rijāl* 30/10).

wa-ʾiḏā depicts reality with anteriority expressed by compound verb form: *wa-ʾiḏā kānat ar-rūmānsiyya qad ʾaddat dawrahā t-taqaddumī . . . fa-ʾinnahā tūšiku l-ʾān ʾan tafqada waḏ̣īfatahā* 'even though romanticism may have played its progressive role . . . it is close to losing its function now' (*Quṭṭ* 159/9). Similarly *hāḏā ʾiḏā*, where the demonstrative pronoun summarizes and constructs an extreme-case scenario when combined with *iḏā*, rendered via 'wenn überhaupt' in German ('if at all'): *wa-yusayṭiru ʿalayya ʾiḥsās bi-n-našwa ʾaw bi-s-saʿāda allatī iftaqadtuhā munḏu ṭ-ṭufūla, hāḏā ʾiḏā kuntu qad šaʿartu bi-hā ʾaṣlan* 'I am overcome by a feeling of ecstasy or happiness, which I have been missing ever since the time of my childhood if I have ever felt it at all' (*Ḏ̣ilāl* 81/6).

wa-laʾin is less common in Modern Standard Arabic: *wa-laʾin kānat mawāqif al-buldān al-xamsa . . . mutaṭābiqa fī l-bidāya . . . fa-ʾinna tūnis ḏ̣allat waḥduhā ṯābita* 'although the positions of the five countries . . . were identical at the beginning, Tunisia alone maintained its firm position' (*Ḥawādiṯ* 12.02.99, 32).

raġma ʾanna and *maʿa ʾanna* indicate real concessive clauses: *ʾaḥbabtu ʾumayma ʿalā r-raġmi ʾannahā kānat tukabbirunī bi-ʾaʿwām ʿadīda* 'I loved Umayma, although she was several years older than me' (*Ṣahīl* 10/7). Similarly: *hāḏihi l-ḥarb al-jadīda allatī ittaxaḏa qarārahā ḥākim fard bi-raġmi ʾanna l-miʾatay milyūn ʿarabī sa-yataḥāmalūna tabiʿātihā fī yawmihim wa-fī ġadihim* 'this new war, which has been decided upon by one single ruler, although two hundred million Arabs will have to face the consequences in their todays and tomorrows . . .' (*Salmān* 22/2).

mahmā and *ʾayyan* clauses designate an irrelevant circumstance. In terms of structure, they are nominal relative clauses of a generalizing kind: *ʾinna l-murāhana l-waḥīda l-maḍmūnat an-natāʾij – mahma ṭāla z-zaman – hiya l-murāhana ʿalā d-dīmūqrāṭiyya* 'the only thing you can bet on with guaranteed returns – no matter how long it may take – is a wager on democracy' (*Mustaqbal* 3/1992, 89/ 35). Or likewise: *muqāwamat aḏ̣-ḏ̣ulm ʾayyan kāna maṣdaruhu* 'the struggle against injustice whatever its source may be' (*Ṯaqāfa* 3/12, 12).

Double conjunctions describe an alternative possibility that appears irrelevant from the speaker's point of view such as *siwāʾun . . . ʾaw* in*: mimmā taqaddama yattaḍiḥu lanā ʿadam qudrat al-madāris fī filasṭīn ʿalā stīʿāb al-ʾaṭfāl allaḏīna hum fī sinn at-taʿlīm. Siwāʾun kāna ḏālika ʾibān al-ḥukm al-ʿuṯmānī li-l-balad aw xilāl al-intidāb al-briṭānī* 'from these statements it is obvious that the schools in Palestine were incapable of accommodating school-age children both during the Ottoman rule over the country and during the British mandate' (MAfF 1997:102).

Where a *ḥāl*-sentence/clause expresses a simultaneous state which is contrary to that of the basic sentence, it may be interpreted in a concessive sense. Such constructions are stylistically marked and mainly restricted to poetry and fiction: *wa-yastaxdimu l-ʿāmmiyya fī l-masraḥiyyāti l-maḥalliyya al-ʿaṣriyya muḥāfaḏ̣atan ʿalā wāqiʿiyyat al-mawāqif wa-l-ʾašxāṣ . . . wa-laqad katabtu "yā ṭāliʿa š-šajara" bi-l-fuṣḥā wa-hiya maḥalliyya ʿaṣriyya* 'he uses dialect in contemporary national theater plays in order to preserve the authentic character of the conditions and persons . . . I wrote the play "Oh tree climber" in Standard Arabic, although it is a contemporary play' (from an interview with Ṭāhā Ḥusayn). With anteriority: *ʾa-taškū ḥarārat al-jaww wa-qad jāʾat min balad maʿrūfa ḥarāratuhu* 'does she complain about the heat of the climate, although she has come from a country known for the intensity of its heat?' (*ʾAmīr* 123/4).

Continued concessive sentences introduced by *maʿa* or *raġma* and *ḏālika* specify a proposition: *ʾinna n-nisyān huwa ʾaḥsan dawāʾ ixtaraʿahu l-bašar fī riḥlatihim al-marīra, wa-maʿa ḏālika lan ʾansāka* 'to forget is the best remedy people have created on their bitter journey. {And} in spite of that, I won't forget you' (*Rasāʾil*, 93/1).

Bibliographical references

Primary sources

ʾAmīr = Dayzi l-ʾAmīr, *Ṯumma taʿūdu l-mawja*. Beirut, 1969.
Ḏ̣ilāl = Ġassān Kāmil Wannūs, *Ḏ̣ilāl an-našwa al-hāriba* (Qiṣaṣ). Damascus, 1994.
Ġīla = ʿAbdallāh ʿUrwī, *Ġīla* (Riwāya). Casablanca, 1998.
Ḥawādiṯ = *al-Ḥawādiṯ* (ʾUsbūʿiyya siyāsiyya ʾixbāriyya). Beirut (since 1911).
MAfF = *Modernes Arabisch für Fortgeschrittene*. Leipzig, 1997.
Marfaʾ = Ḥannā Mīnah, *al-Marfaʾ al-baʿīd*. Beirut, 1991.

Mustaqbal = *al-Mustaqbal al-ʿarabī*. Paris.
Quṭṭ = ʿAbd al-Qādir al-Quṭṭ, *Fī l-ʾadab al-ʿarabī al-ḥadīṯ*. Cairo, 1978.
Rasāʾil = *Rasāʾil Gassān Kanafānī ʾilā Ġāda s-Sammān*. Beirut, 1992.
Rijāl = Ġassān Kanafānī, *Rijal fī š-šams*. Beirut, 1963.
Ṣahīl = Zakariyyāʾ Tāmir, *Ṣahīl al-jawād al-ʾabyaḍ* (Qiṣaṣ). Damascus, 1978.
Ṯaqāfa = *aṯ-Ṯaqāfa al-waṭaniyya*. Damascus, 1989.
Salmān = Ṭalāl Salmān, *ʿAlā ṭ-ṭarīq. Ḥajar yaṯqub layl al-hazīma*. Beirut, 1992.
Waṭan = *al-Waṭan* (yawmiyya). Kuwait.

Michael Langer (Schlettau/Erzgebirge, Germany)

Conditional Clause → Subordination

Congruence → Agreement

Conjunctions

1. Morphological structure

Conjunctions are a closed class of uninflected words which serve the joining of words, phrases, clauses, or sentences and simultaneously express a specific semantic relationship between the conjoined elements. In most cases Arabic conjunctions lack the otherwise salient root-pattern structure of the language. Accordingly, there is usually no distinction between a consonantal root expressing the basic meaning and a vowel pattern bearing functional value. Even if it is possible to discern in a conjunction a root consisting of three radicals it is difficult or impossible to associate this root with some specific semantic content.

Morphologically, it is possible to distinguish between simple and complex conjunctions. Simple conjunctions consist of only one lexical morpheme. They comprise many different bases such as nouns like *ḥīna* 'when' or *ḥayṯu* 'where, when'; deictic elements like *ʾiḏā* 'when' or *ʾiḏ* 'when, because'; prepositions like *li-* 'in order that' or *ḥattā* 'until, in order/so that'; interrogative particles like *mā* 'what, (the fact) that, as long as' or *matā* 'when(ever)'; and isolated bases such as *wa-* 'and', *fa-* 'then', *ʾaw* 'or', *ʾan(na)* 'that', *law* 'if', *kay* 'in order that', or *lammā* 'when, after'. The latter group contains some of the most important conjunctions of Arabic.

Complex conjunctions are usually combinations of prepositions and simple conjunctions, or of at least two simple conjunctions. Especially common is the compound of preposition and *ʾan(na)* or *mā* where the two simple conjunctions signal the word class change. Examples are *baʿda ʾan*/*baʿda mā* 'after' or *ka-mā* 'just as, same as'. The combination of simple conjunctions show for instance *ḥattā ʾiḏā* 'when, after', *law ʾanna* = *law*, or *li-kay-mā* = *kay*. *ʾIllā* 'if not, except' is formed from *ʾin* 'if' and negative *lā* 'not'. The Arabic script distinguishes between *li-ʾan* = *li-* (spelt <lʾn>) and *la-ʾin* 'if truly' (spelt <lyn>) by means of different *hamza* orthography. Some conjunctions occur much more frequently than synonymous ones, for instance, final syndetic subordinate clauses in Classical Arabic are introduced by *li-* rather than by (*li-*) *kay* or *ḥattā*.

2. Etymology

The etymology of many conjunctions is difficult and sometimes impossible to determine. Some Arabic conjunctions are common Semitic and may be attributed to the protolanguage (Brockelmann 1908–1913:I, 502; Lipiński 1997:470ff., 519ff.), e.g. *wa-*, *ʾaw*, *fa-*, *law*, *kay* (unconvincing Voigt 1999:40), or *ʾin* (Voigt 1995). Others are diachronically nouns in the accusative case and construct state in front of relative clauses, which have synchronically evolved into real conjunctions, e.g. *yawma* '(the day) when' or *ḥīna*. *Ḥayṯu* is a noun in an ancient local case (Brockelmann 1908–1913:II, 533). The same ending can be found in *munḏu* 'since', which may be segmented into the preposition *min* 'from' and a deictic element /ḏ(ā)/ 'from then on' > 'since' (Reckendorf 1895–1898:237; Wright 1975:I, 280). *Muḏ* is the shortened form (< **munḏ*, Fischer 2002:29). Other conjunctions formed on the basis of deictic elements are *ṯumma* 'then' (cf. *ṯamma* 'there') and *ʾiḏ*, *ʾiḏā*, *ʾiḏan* 'then' which are related to Hebrew *ʾåz* and *ʾăzay* 'then' respectively (Brockelmann 1908–1913:II, 594). For the prepositions *li-*, *ladun* 'with, since', and *ḥattā* there do not seem to exist reliable etymologies. *Ladun* may be of nominal origin but in view of its many variant forms it must be considered a rather obscure word (cf. Wright 1975:I, 280). Of the various rather unconvincing proposals see for instance for *li-*

Voigt (1999:41) and for *ḥattā* Fleischer (1885: 403), Rundgren (1955:31), or Fischer (1997: 199). Although it is tempting to relate *ḥattā* to Semitic **ʿaday* ‘until’ (Wright 1975:I, 280), it still remains unexplained. *Lākin* ‘but’ goes back to ‘not so’ (Brockelmann 1908–1913:II, 480; unconvincing Reckendorf 1895–1898:364ff.) while its variant *lākinna* is formed in analogy to the likewise clause initial *ʾinna* ‘indeed’ (→ asseverative). *Bal* ‘but, on the contrary’ is related to Hebrew *ʾăḇål*, which has the same meaning, and originally seems to go back to a negative particle (Brockelmann 1908–1913:II, 200). *ʿAm* ‘or’ in alternative questions is formed from **ʾa-mā* ‘not?’ (Reckendorf 1895–1898:480; Brockelmann 1908–1913:II, 494). The origins of *lammā* and *ʾan(na)* are obscure. Whereas *ʾanna* can be explained as a formal and semantic blending of *ʾinna* and *ʾan* (Nöldeke 1963:40) simple *ʾan* has no apparent etymology (but compare Epigraphic South Arabian *hn* ‘that’ [**han*], Stein, *forthcoming*; (see also Lipiński 1997: 535). Likewise unexplained is *lammā* whose relation to negative *lammā* ‘not yet’ and exceptive *lammā* ‘except’ is not entirely clear (cf. Lipiński 1997:528ff.). Proposals to derive it from **la-mā* (Reckendorf 1895–1898:656) or from **ʾilā* [= *ʾiḏā*] + *mā* (Brockelmann 1908–1913:II, 600) are not convincing.

3. Syntactic and semantic properties

Regarding the syntax and semantics of the various conjunctions, it should be borne in mind that many semantic notions can be conveyed not only by conjunctions but also by means of asyndetic constructions and circumstantial clauses embedded with *wa-*. The asyndetic prefix conjugation, for instance, equals a subordinate final clause with *li-*. Circumstantial clauses with suffix conjugation have past time reference and are introduced by *wa-qad* in which case they have a function similar to temporal conjunctional clauses. Simultaneous actions with a wide range of semantic interpretations (Nebes 1999: 79) are embedded by means of *wa-* and functionally equal several conjunctional clauses.

Wa-, *fa-*, and *ʾaw* coordinate words, phrases, clauses, and sentences, usually syntactically equivalent items. *Bal* and *lākin(na)*, mostly *wa-lākin(na)*, introduce adversative clauses. *Ṯumma* and *ʾiḏan* denote a temporally ordered sequence of events by joining independent sentences. After *ʾanna* and *lākinna* the subject occurs in the manner of *ʾinna* in the accusative case, but not so after the shortened forms *ʾan* and *lākin*. Many conjunctions head complement and adverbial clauses. Complement clauses function as noun phrases and can have subject or object function and follow a head noun or a preposition. They are primarily introduced by *ʾan(na)* and *mā* (Reckendorf 1921:394ff.). Conjunctional clauses often have adverbial function and denote various semantic notions, among others temporal, final, concessive, or conditional. The choice of prefix or suffix conjugation after a conjunction is principally dependent on the intended time reference, except in conditional clauses, which can invariably use suffix conjugation or apocopate in the protasis and apodosis. The former is introduced by *ʾin* and *law* for real and unreal condition respectively. Several conjunctional clauses denoting desired or intended events and therefore having future time reference use the subjunctive, e.g. *ʾan*, *qabla ʾan* ‘before’, *li-*, or *ḥattā* (Reckendorf 1921:454ff.; Fischer 2002:97). The choice between indicative and subjunctive is not always entirely clear, especially after *ʾan* or *ḥattā* (cf. Wright 1975:II, 26; Wild 1980).

The position of the conjunctional clause varies. The overall tendency in accordance with the leftheaded structure of Arabic is for the conjunctionally headed clause to follow its matrix clause, but especially temporal and conditional clauses frequently precede the matrix clause. The distinction between coordinate and subordinate in Arabic often becomes blurred in that several conjunctions have coordinate and subordinate function (→ parataxis). *Wa-* and *fa-* also introduce embedded circumstantial clauses (Fischer 2002:185; Nebes 1999) and *ʾiḏ*, *ʾiḏā*, and *fa-* head the matrix clause of the so-called “Inzidenzschema” (Nebes 2001). Likewise after fronted temporal or conditional clauses *fa-*, *wa-*, and *ʾiḏ* are possible (Reckendorf 1921:482ff.). Even after conjunctions such as *fa-*, *ʾaw*, or *ʾiḏan* the subjunctive may be used (Wright 1975:II, 30ff.). Therefore, it seems at least hazardous to analyze Arabic conjunctions *prima facie* as coordinating or subordinating, although it has to be admitted that indisputable instances of subordinating conjunctions do exist (e.g. complement clauses introduced by *ʾan(na)*, *li-*, or *qabla ʾan* with subjunctive).

4. HISTORICAL DEVELOPMENTS

Generally speaking, Modern Standard Arabic follows the situation in Classical Arabic with only minor changes. It makes use of virtually the same set of conjunctions but also creates new ones by combining nouns, prepositions, or prepositional phrases with *ʾan(na)*, e.g. *waqta ʾan* 'when' (Badawi a.o. 2004:601), *ʿalā ʾanna* 'but, however' (Badawi a.o. 2004:612), or *bi-šarṭi ʾan* 'on the condition that' (Badawi a.o. 2004:603). In addition to Classical (_*ʾin-mā_ >) *ʾimmā . . . (wa-) ʾimmā/ʾaw* 'either . . . or', Modern Standard Arabic also uses *sawāʾan . . . ʾam/ʾaw* 'whether/ either . . . or' (Badawi a.o. 2004:567 ff.). The use of previously rare conjunctions is extended as well, e.g. *ʿinda-mā* 'when' (Reckendorf 1921:471), which has replaced *lammā* to a large extent (Badawi a.o. 2004:637). Likewise, both *ʾiḏā* and *law* have taken over the function of *ʾin* which appears to be less frequently used (Badawi a.o. 2004:636). The temporal conjunction *ʾiḏ* extends its use as explicative 'for, because' (Holes 1995:232). Unlike in Classical Arabic, conjunctional clauses of purpose and reason can precede the matrix clause (Holes 1995:235).

Middle Arabic shows more deviations from Classical usage. *ʾinna* and *ʾan(na)* have merged into invariable <ʾn> (*ʾin(n)*?) in all positions (Blau 1967:510). New conjunctions have been created, e.g. *ʾilā ḥīn* 'until' and *min ḥīn* 'since' (Fück 1950:62) or *li-lā* 'lest' (Blau 1967:539). Simple prepositions without *ʾan* may be used as conjunctions, for instance *dūn* 'without' (Blau 1967:505), whereas others are combined pleonastically with *ʾan*, e.g. *ḥattā ʾan* 'until' (Blau 1967:529). Some conjunctions have acquired new meanings, e.g. *wa-* 'also, or' (Blau 1967: 454–458), *lākin* 'then' after temporal conjunctions (Blau 1967:460) or *ʾin*, *ʾiḏā*, and *law*, for which strict delimitations of usage have been blurred (Blau 1965:96). In other cases, there is a change of syntactic properties. *ʾan* may introduce independent utterances (Blau 1967:516; Diem 2002). Circumstantial clauses with *wa-* are often replaced by conjunctional clauses introduced by *ʿindamā*, *min ḥayṯu*, or *fīmā* (Fück 1950:62) and may precede the matrix clause (Blau 1967:509). Foreign influence is discernible for instance in *sawā wa-* 'as soon as' (< Aramaic *šwē ḏ-*, Blau 1967:457).

Modern Arabic dialects continue the usage of Middle Arabic. Some conjunctions have been lost, e.g. *ʾan(na)*, which is replaced by *ʾilenn* (< *ʾinna*, Kaye a.o. 1997:307), or *ʾiḏ* (Holes 1995: 233). *Fa-* appears to be lost as well (Brockelmann 1908–1913:II, 488) but it is listed in Werbeck (2001:249). Numerous Classical conjunctions have been preserved in altered form, e.g. *w- ~ u- = wa-* (many dialects), *ta = ḥattā* (Blau 1960:233) or *ʾīḏä ~ ʾīlä = ʾiḏā* (Singer 1984:700). Other conjunctions have acquired new meanings, e.g. *kīf* 'when, while' (Singer 1984:689), *ʾəlla ~ ʾalla* 'or' (Wittrich 2001:133, also *willā* Werbeck 2001:250), or *lamma* 'until' (Woidich 1991:175). Many conjunctions have been newly created; examples are *ʿalašān ~ ʿašān* 'in order that, because' (Holes 1995:234), *m-ĕlli* (*mən-ĕlli*) 'since' (Singer 1984:693), or *min sāʿit-ma* 'since' (Woidich 1990:296). The formation of *ma*-compounds has been productive in the dialects in general (Holes 1995:233). Other conjunctions are borrowed from neighboring languages, for instance conditional markers *ʾăgăr* (Sasse 1971:193) and *haka(r)* (Wittrich 2001: 135) from Kurdish. There are also conjunctions of unclear origin like *ʿišt* 'because' in Çukurova Arabic (Procházka 2002:147).

Finally it should be noted that conjunctions in Arabic dialects may show many variant forms (e.g. Marçais 1977:229) and that despite a large number of available conjunctions, in Arabic dialects, as already in Middle Arabic, the use of asyndetic constructions is conspicuous (Brockelmann 1908–1913:II, 474; Blau 1967:487ff.).

BIBLIOGRAPHICAL REFERENCES

Badawi, El-Said, Michael G. Carter, and Adrian Gully. 2004. *Modern written Arabic: A comprehensive grammar*. London: Routledge.

Blau, Joshua. 1960. *Syntax des palästinensischen Bauerndialekts von Bīr-Zēt*. Walldorf-Hessen: Verlag für Orientkunde.

——. 1965. *The emergence and linguistic background of Judaeo-Arabic: A study of the origins of Middle Arabic*. Oxford: Oxford University Press.

——. 1967. *A grammar of Christian Arabic, based mainly on South-Palestinian texts from the first millennium*. Leuven: Secrétariat du Corpus SCO.

Brockelmann, Carl. 1908–1913. *Grundriss der vergleichenden Grammatik der semitischen Sprachen*. 2 vols. Berlin: Reuther und Reichard.

Diem, Werner. 2002. "Nichtsubordinatives modales *ʾan yafʿala*: Ein Beitrag zur Syntax der nachklassischen arabischen Schriftsprache". *"Sprich doch mit deinen Knechten aramäisch, wir verstehen es!": 60 Beiträge zur Semitistik. Festschrift für Otto Jastrow zum 60. Geburtstag*, ed. Werner Arnold and Hartmut Bobzin, 113–145. Wiesbaden: O. Harrassowitz.

Fischer, Wolfdietrich. 1997. "Classical Arabic". Hetzron (1997:187–219).

——. 2002. *Grammatik des klassischen Arabisch*. 3rd ed. Wiesbaden: O. Harrassowitz.

Fleischer, Heinrich L. 1885. *Kleinere Schriften*, I. Leipzig: S. Hirzel.

Fück, Johann. 1950. *Arabiya: Untersuchungen zur arabischen Sprach- und Stilgeschichte*. Berlin: Akademie-Verlag.

Hetzron, Robert (ed.). 1997. *The Semitic languages*. London: Routledge.

Holes, Clive. 1995. *Modern Arabic: Structures, functions and varieties*. London: Longman.

Kaye, Alan S. and Judith Rosenhouse. 1997. "Arabic dialects and Maltese". Hetzron (1997:263–311).

Lipiński, Edward. 1997. *Semitic languages: Outline of a comparative grammar*. Leuven: Peeters.

Marçais, Philippe. 1977. *Esquisse grammaticale de l'arabe maghrébin*. Paris: Librairie d'Amérique et d'Orient.

Nebes, Norbert. 1999. "Das Satzschema *fa-huwa yafʿalu/fāʿilun*/Prädikativ für Vergangenheit in frühklassischer arabischer Erzählliteratur". *Tempus und Aspekt in den semitischen Sprachen: Jenaer Kolloquium zur semitischen Sprachwissenschaft*, ed. Norbert Nebes, 77–100. Wiesbaden: O. Harrassowitz.

——. 2001. "Das Inzidenzschema im klassischen Arabischen: Ein Vorbericht". *Sachverhalt und Zeitbezug: Semitistische und alttestamentliche Studien. Adolf Denz zum 65. Geburtstag*, ed. Rüdiger Bartelmus and Norbert Nebes, 113–128. Wiesbaden: O. Harrassowitz.

Nöldeke, Theodor. 1963. *Zur Grammatik des classischen Arabisch. (Im Anhang: Die handschriftlichen Ergänzungen in dem Handexemplar Theodor Nöldekes bearbeitet und mit Zusätzen versehen von Anton Spitaler)*. Darmstadt: Wissenschaftliche Buchgesellschaft.

Procházka, Stephan. 2002. *Die arabischen Dialekte der Çukurova (Südtürkei)*. Wiesbaden: O. Harrassowitz.

Reckendorf, Hermann. 1895–1898. *Die syntaktischen Verhältnisse des Arabischen*. Leiden: E.J. Brill.

——. 1921. *Arabische Syntax*. Heidelberg: C. Winter.

Rundgren, Frithiof. 1955. *Über Bildungen mit š- und n-t-Demonstrativen im Semitischen: Beiträge zur vergleichenden Grammatik der semitischen Sprachen*. Uppsala: Almqvist and Wiksell.

Sasse, Hans-Jürgen. 1971. *Linguistische Analyse des arabischen Dialekts der Mḥallamīye in der Provinz Mardin (Südosttürkei)*. Ph.D. diss., University of Munich.

Singer, Hans-Rudolf. 1984. *Grammatik der arabischen Mundart der Medina von Tunis*. Berlin: W. de Gruyter.

Stein, Peter. Forthcoming. "Materialien zur sabäischen Dialektologie: Das Problem des amiritischen ('haramischen') Dialektes".

Voigt, Rainer. 1995. "Akkadisch *šumma* 'wenn' und die Konditionalpartikeln des Westsemitischen". *Vom Alten Orient zum Alten Testament: Festschrift für Wolfram Freiherrn von Soden zum 85. Geburtstag am 19. Juni 1993*, ed. Manfried Dietrich and Oswald Loretz, 517–528. Neukirchen-Vluyn: Neukirchener Verlag.

——. 1999. "Die Präpositionen im Semitischen: Über Morphologisierungsprozesse im Semitischen". *Tradition and innovation: Norm and deviation in Arabic and Semitic linguistics*, ed. Lutz Edzard and Mohammed Nekroumi, 22–43. Wiesbaden: O. Harrassowitz.

Werbeck, Wolfgang. 2001. *Laut- und Formenlehre des nordjemenitisch-arabischen Dialekts von Manāḫa*. Münster: Rhema.

Wild, Stefan. 1980. "Die Konjunktion *ḥattā* mit dem Indikativ Imperfekt im klassischen Arabisch". *Studien aus Arabistik und Semitistik: Anton Spitaler zum siebzigsten Geburtstag von seinen Schülern überreicht*, ed. Werner Diem and Stefan Wild, 204–223. Wiesbaden: O. Harrassowitz.

Wittrich, Michaela. 2001. *Der arabische Dialekt von Āzəx*. Wiesbaden: O. Harrassowitz.

Woidich, Manfred. 1990. *Ahlan wa-sahlan: Eine Einführung in die Kairoer Umgangssprache*. Wiesbaden: L. Reichert.

——. 1991. "Zur Entwicklung der Konjunktion *aḥsan ~ laḥsan* im Kairenischen". *Festgabe für Hans-Rudolf Singer: Zum 65. Geburtstag am 6. April 1990*, I, ed. Martin Forstner, 175–193. Frankfurt: P. Lang.

Wright, William. 1975. *A grammar of the Arabic language*. 2 vols. Repr. Cambridge: Cambridge University Press.

Michael Waltisberg
(Philipps University Marburg)

Connectives

1. Introduction

Connectives can be defined as utterance-initial words, phrases, and clauses that share the primary function of linking utterances in discourse. Typical English connectives are: the conjunction *and*, the adverbs *then* and *nevertheless*, the prepositional expression *on the other hand*, and the clause *as I mentioned earlier*. Thus, what matters is function rather than form, in combination with the position in the utterance.

Connective as a linguistic term was borrowed from general philosophy and logic by van Dijk in the 1970s as → text linguistics developed as a discipline. The term was taken to denote the type of ties in the surface structure of natural language which "in traditional grammar are usually called 'conjunctions'" (van Dijk 1977:14). The reason for the change of term was the wish not to confuse the new, discourse-related concept with the traditional word class → 'conjunction'. Van Dijk obviously also wanted to include more material than just conjunctions in the new category, while 'conjunction' could continue to be used in its restricted sense.

The relationships expressed by conjunctions have been seen as logical. However, as pointed out by Van Dijk and others, the so called 'truth-functional' inter-propositional relations of formal logic are essentially different from those of real, natural language, in that the latter also depend on the content of propositions in sequence (Van Dijk 1977:12). Thus, rather than 'standing for' logical relationships themselves, conjunctions seem to function as mere 'markers' or 'deictics' of the relations that hold between propositions in discourse.

Since Schiffrin's (1987) pioneering work on discourse markers in English, this term seems to have been widely adopted – especially by those engaged in pragmatic → Discourse Analysis – for what appears to be the same category as van Dijk's. Blakemore (2002), for instance, has substituted 'discourse markers' for her earlier term 'discourse connectives'. She explains the shift of terms by pointing out that conjunctions and other 'bracketing' words do not have 'descriptive' meaning. Their meaning is rather to indicate how the individual propositions relate to discourse as a whole. Thus 'marker' is the better term. Yet, because they generally also mark connections among units of discourse, the term 'connective' could be used for them, but "since there is no agreement on what counts as a discourse marker, it is difficult to know whether these are two labels for the same phenomena" (Blakemore 2002:1). Schiffrin defines the category as "sequentially dependent elements which bracket units of talk", i.e. utterance-initial items that function in relation to ongoing talk and text. Her analysis includes the expressions *oh*, *well*, *and*, *but*, *or*, *so*, *because*, *now*, *then*, *I mean*, *y'know*, i.e. members of form classes as varied as conjunctions, interjections, adverbs, and lexicalized phrases (Schiffrin 1987:32; 2001:57).

2. The traditional view of connectives

In traditional grammar, connectives are not treated as a unified class. In the Arabic linguistic tradition, they are found within the class of *ḥurūf* 'particles' (→ *ḥarf*), notably within the sub-category of *ʾadawāt ar-rabṭ* 'coordinating particles'. In Western traditional treatments, the majority of them are found under the rubrics of conjunction and adverbs, whereas others would have to be studied individually in different parts of the grammar.

Discussing Ibn Hišām's (d. 761/1359) interest in the particles, Gully (1995:20) marvels at the amount of space devoted by the Arab grammarians to "a class of words whose counterparts in the English language, such as 'at' or 'if', do not appear to ever have attracted anything like the same sort of interest". Ibn Hišām dedicates more than half of his treatise to the particles, mainly because he sees them primarily as *ʿawāmil* 'operators', which govern the different states of the morpho-syntactic category of → *ʾiʿrāb* (both case and mood). He never attempts to treat connectives outside the confines of the sentence. What matters for him is how connectives and other particles affect the *ʾiʿrāb*, and how this reflects the functions of words and consequently, the function of the sentence as a whole. Since declension has no bearing beyond the sentence, the sentence remains the natural domain of the particle.

Al-Batal (1985:22–24) points out that the Arabic medieval literature of the *ʿilm al-balāġa* 'rhetoric' does take an interest in discourse as a whole. The particles *ʾinnamā* and *ʾinna* are treated by the grammarian-*cum*-rhetorician Jurjānī (d. 1078) against the background of his concept of *naḏ̣m* 'logical arrangement' (also 'string of pearls'). This concept combined formal and syntactic features with those of the context. The rhetorician generally used discourse-related concepts, such as *takrār* 'repetition' and *al-waṣl wa-l-faṣl* 'conjunction and disjunction', i.e. connected and disconnected discourse. Although the latter distinction was only applied to the use of *wa-*, "the skill of knowing when sentences should be connected and when they should not was regarded by the rhetoricians as the ultimate degree of eloquence" (Al-Batal 1985:24).

Western Arabic grammars are largely faithful to the treatment of the Arabic conjunction in the Arabic linguistic tradition, although they distinguish clearly between prepositions and conjunctions. The preposition *li-* 'to' and the conjunction *li-* 'in order to', for instance, are no longer considered to be identical, but rather homonyms. The main sources for Wright's grammar (1933, first published 1859–1862) are Ibn Mālik's *ʾAlfiyya* and Zamaxšarī's *Mufaṣṣal* (13th/14th century C.E.). Although there is a separate section for the conjunction, individual expressions are spread throughout the grammar. Conjunctions are mainly discussed in connection with different sentence types. No attempt is made to reach beyond the sentence level of discourse. Wright's examples are borrowed from or modeled on Classical Arabic from the Arab grammarians. Although Reckendorf (1895–1898, 1921) is more systematic in his presentation, his description is still concerned with Classical Arabic, with examples from the *Qur'ān* and the *ḥadīṯ*. Cantarino's work (1974–1975) is the first attempt to describe the syntax of Modern Standard Arabic, with a corpus from the first half of the 20th century, all of it from literary prose. He, too, sorts conjunctions after clause types and never mentions units larger than the sentence.

The most recent grammar of Modern Standard Arabic is that by Badawi, Carter, and Gully (2004). All material is post-1990. The section on the 'Hypersentence and Discourse', includes a paragraph on 'resumptive *fa-*', in which a large number of prepositional phrases followed by *fa-* are listed, e.g. *ʿalā kulli ḥālin fa-* 'in any case . . .', which they assert is an extension of *ʾammā fa-*, influenced by a Western model (Badawi a.o. 2004:723–739). Just as in earlier Western studies, conjunctions are listed according to the clause types they tend to introduce. Clauses and sentences are divided along syntactic lines into coordinated and subordinated entities, without clear criteria for this division. Discourse organization is presented within this same traditional syntactic division, which in Arabic is especially difficult to maintain, since there is no overt marking of it.

For teaching purposes an impressive list of connectives (*ʾadawāt ar-rabṭ*) was established for an exercise book developed at the Arabic Department of the American University at Cairo at the beginning of the 1980s (Warrāqī and Ḥasanayn 1981).

3. A NEW UNDERSTANDING

Since the 1980s a new understanding of connectives has emerged within general linguistics, which is gradually finding its way into language-specific linguistics. A number of studies on Arabic have appeared that attempt to treat conjunctions and conjunctive phrases from a discourse perspective, i.e. as workers of cohesion in text.

Inspired by Halliday and Hasan's (1976) concept of → 'cohesion' and the idea that conjunctions and other items in the surface structure of the language contribute to the cohesion of the text, Al-Batal (1985) sets out to examine the connectives of an excerpt from the writings of the Egyptian prose writer Maḥmūd ʿAqqād. His motivation derives from the observation he has made as a teacher of Arabic that English-speaking students are reluctant to use connectives at the beginning of sentences and paragraphs of their essays, where Arabic would require them. Al-Batal (1985:3–4) feels that the formal properties and syntactic constraints of conjunctions are overemphasized in traditional grammar. The teaching situation calls for an alternative approach where the "semantic, syntactic and discourse properties of connectives" are taken together in a more holistic approach.

Al-Batal (1985:33) aims at describing the various connectives found in the text in terms of their syntactic properties (whether they are coordinating or subordinating conjunctions, adverbs or prepositional phrases); the semantic relationships they signal (e.g. additive, causal, or sequential); their cohesive role in discourse as a whole; and the scope and significance of this cohesive role (at what level it operates, phrase, clause, sentence, paragraph, or discourse). The word 'connective' is used in a rather broad sense to refer to "any element in the text which – regardless of whether or not it belongs to the form class of conjunctions – indicates a linking or transitional relationship between phrases, clauses, sentences and paragraphs exclusive of referential or lexical ties" (Al-Batal 1985:2).

The main results of Al-Batal's investigation can be summarized as follows: connectives fulfill an unexpected number of functions, which are not all covered by Halliday and Hasan's (1976:142–242) four categories (additive, alternative, adversative, and temporal). They appear at all levels in the text. The majority of them operate at sentence level or below (Al-Batal 1985:272), while around twenty occur at text level, which shows that an analysis not taking text level into consideration is incomplete. Sixty percent of the connectives appear at clause level. These connectives seem to have an important function, when it comes to binding together long sentences containing many clauses. In the relation between clauses, coordination dominates over subordination (110 cases of coordination versus 71 cases of subordination), which confirms earlier data on the paratactic nature of Arabic. Furthermore, connectives at sentence level are predominantly coordinating. The connective *wa-* is highly frequent at all levels and therefore crucial for cohesion in Modern Standard Arabic. *wa-* has at all levels an additive function, and at higher levels it signals the uninterrupted flow of the argument. The connective *fa-* has a great variety of functions. There seems to exist a Ø-connective. Sentences without initial connectives signal a special relation to coming sentences. According to Al-Batal, the unmarked sentence signals a change in the progression of the text (the changeover from introduction to the main text or the changeover from question to answer).

The main merit of Al-Batal's work is its profound awareness of the importance of applying a discourse perspective to certain otherwise inexplicable linguistic phenomena, and it is this approach which makes it pioneering in many ways. In a similar study of Lebanese colloquial material audio-recorded in 1989–1991 (Al-Batal 1994), he uses the same linguistic levels as in his earlier study, although he has changed the term 'paragraph' to 'paratone', indicating that it is marked prosodically rather than orthographically. The strength of this work is the useful list of colloquial and diglossic hybrid connectives and the selection of data from different levels of style, with a frequency count. Al-Batal finds that there is a gradual progression along a continuum, where semi-educated Lebanese Arabic represents one end and Modern Standard Arabic the other. As speakers move closer to the written variety, they tend to use more connectives and a greater variety of different connectives. At the same time, sentence structure becomes more and more complex with this movement.

Johnstone (1990) claims that modern Arabic prose is more paratactic than it looks. In Al-Batal's analysis the concept of the syntactic sentence in Modern Written Arabic is not defined, yet he assigns syntactic status to all connectives without clear syntactic and semantic criteria (Al-Batal 1985:115–116). There has been a tendency to decide whether an Arabic conjunction is subordinating or coordinating on the basis of the syntactic status of the translated expression. Johnstone (1990:221–223) points to the fact that there is no overt linguistic marking in Arabic for subordinated clauses. Not even relative clauses are syntactically subordinated, since they are asyndetically joined when indefinite and, when definite, are joined by means of what looks more like a definite article than anything else. In English, asyndetic linkage characterizes main clauses, while syndetic linkage is a sign of subordination. In Arabic some modification clauses (sub-clauses) are conjoined asyndetically, while the vast majority of clauses, whether embedded or not, start with a conjunction or a conjunctive phrase. Thus, there is a 'connective slot', a space reserved for a single conjunction such as *wa-*, or a whole string of connective elements, such as *wa-li-ḏālika fa-* 'and therefore so' or whole phrases, such as *mimmā huwa jadīrun bi-ḏ-ḏikri wa-l-mulāḥaḏāti* 'among that which is worth mentioning and noting' (Johnstone 1990:221).

Such a clause would be labeled 'subject' in a syntactic analysis. However, in discourse it is a connective, i.e., a case of syndetic linkage, which remains outside the propositional core of the sentence. The purpose of this connective seems to be to create a paratactic link between the preceding and the following. Likewise, *ʾanna* 'that' is not only a subordinator, but also a topicalizer which foregrounds the subject (→ grounding). There are quite a few particles behaving in the same way, e.g. *li-ʾanna* 'because' (→ *ʾinna wa-ʾaxawātuhā*), which topicalize the subject, making the word order SV rather than the neutral narrative order V(S) (→ topic/comment). Conversely, to topicalize means to 'pay the price' of subordination, which then becomes a secondary feature. In an instance such as this, the contrast between subordination and coordination is neutralized by the contrast of word order, i.e. neutral vs. marked theme.

The paratactic impression is further enhanced by the appositive conjoining of Arabic relative clauses, where the relative particle does not replace the noun, although it agrees with it in form. Its function is more like that of a definite article. This is proven by the fact that indefinite relative clauses are linked asyndetically. The relationship is signaled by mere juxtaposition: the relative clause defines the noun because it stands next to it. This seems to be a common trait in Arabic modificational syntax. Even the so-called → *ḥāl* clause, i.e. the circumstantial clause, is conjoined in this way when it shares participants with the main proposition. Such clauses are thus semantically subordinated, without being structurally embedded.

Characteristic of Modern Standard Arabic is also, according to Johnstone, its formulaicity and it repetitiousness. Discourse formulas appear at all levels: lexical, morphological, syntactic, and discourse-structural (Johnstone

1990:223). Some of them are 'prior-text formulas', i.e. ready-made chunks of language that recur in a specific discourse type, in texts written by different authors, e.g. the phrase *jadīrun bi-ḏ-ḏikri* 'worth mentioning' in expository prose. The second type is represented by 'emergent formulas', formulas created 'on-line' by frequent use of the same structures. The latter category can be identified with the love of repetition in Arabic, or what Johnstone (1990:224) calls the 'doubling rule' which is especially common with noun phrases containing an adjectival modifier.

Inspired by Schiffrin's (1987) concept of the 'discourse marker', Sarig (1995) discusses *wa-* and *fa-* and a few other Arabic "initial position functionals (IPFs)" with examples from the Egyptian press. Since discourse markers are redundant, the structure and meaning of arguments can be preserved even without them (Sarig 1995:20). They are, thus, neither structural nor semantic, but rather 'deictic' in function. Their role is to "point out the text's rhetorical structure" (Sarig 1995:8). The frequent recurrence of *wa-* and *fa-* at the opening of a chapter, the beginning of a paragraph, or at the head of a new paragraph shows, according to Sarig, that they are deictics. Without defining the term 'connective', she declares that these two conjunctions "are at times inserted where they clearly have no connective function" (Sarig 1995:7). According to this view, the deictic function of *wa-* (and the combination *wa-qad*) is to signal that the discourse following the marker is a continuation of the topic preceding it. Similarly, the deictic role of *fa-* (and *fa-qad*) is to confirm or clarify an earlier proposition (Sarig 1995:8). Yet, her statement of *wa-* and *fa-* 'not having connective function' needs some clarification. If 'connective' is taken to mean a coordinating or subordinating syntactic function, much like that of the traditional class of conjunctions, it is evidently true. However, it is wrong if the broader, discourse-oriented definition of the term is applied, namely that connectives provide linkage among the units of discourse.

Kammensjö (2004) bases her analysis of connectives on a corpus of transcribed audio-recordings from lectures held at the faculties of arts of four Arab universities during 1995–2000. She regards the sentence as the basic significant element in her investigation. On the importance of the sentence, Chafe notes that "the sentence has, in fact, seemed so important to so many that it has been the basic unit of grammatical study from ancient times to the present" (Chafe 1994:140). Halliday and Hasan (1976:8) explain that sentence boundary markers are not cohesive ties as such, but because the sentence is "the highest unit of grammatical structure [. . .] it tends to determine the way in which cohesion is expressed". In later writings Halliday (1994:xxi) develops his idea of the sentence as the largest unit of grammatical structure calling it "a significant border post", which writing systems are sensitive to record. On the linguistic levels below the sentence, the constructional type of organization dominates, i.e. elements are organized as parts of the whole, in a certain variable sequence, where the subject can go either before or after the finite verb depending on the function of the sequence. Above the sentence, the more abstract ties of cohesion take over, e.g. reference, → ellipsis, etc.

Kammensjö's corpus was segmented into sentences primarily on syntactic grounds, but not strictly so, since prosody and semantics sometimes counteract a strict syntactic judgment in oral language. To allow a greater freedom and to avoid the confusion with the graphic sentence, the basic segment of the corpus unit is termed 'utterance', a 'pre-theoretical' concept, to quote Lyons (1977:633–635). Since the utterance can be regarded as the basic unit of discourse, 'utterance starters', i.e., connectives introducing units at sentence level, are assumed to be particularly significant to the coherence of discourse.

By making use of Halliday's thematic structure analysis (Halliday 1994:37–67) and inverting the definition of the thematic slot, the connective slot of the utterances could be successfully delimited. From Halliday's (1994:50) argument it follows that discourse connectives have obligatory thematic status, since they always come first in the sentence, i.e. before the topical theme (the first element belonging to the experiential part of the utterance starting with participant, process, or circumstance; → theme/rheme). The two categories of theme and connective can thus be said to overlap, the theme being what speakers choose to start with. The connective is what speakers choose to use or not. If they use it, it becomes a natural theme by virtue of its position. Consequently, where there are no connectives initially in the utterance, the total thematic force is taken up by the topical theme alone. Hence, the thematic structure and the distribution of connectives must be interdependent.

The first field of the thematic slot is the so-called 'textual theme', composed of continuatives, conjunctions, and conjunctive adjuncts. The members of this category all pass as connectives. The second field is reserved for the interpersonal theme, consisting of vocatives, modal adjuncts, and mood-markers. In this way, Halliday's theme categories have been adapted to form the following five connective classes:

(1) Continuatives: *yaʿnī* 'that is'; *ṭayyib* 'well, OK', etc.
(2) Conjunctions: *wa-* 'and'; *fa-* 'and so, for'; *ṯumma* 'then', etc.
(3) Conjunctive adjuncts: *bi-t-tāli* 'consequently'; *li-ḏalik* 'therefore', etc.
(4) Interpersonal (modal) adjuncts: *ṭabʿan* 'of course'; *ḥaqīqatan* 'truly', etc. (also including vocatives, e.g. *yā jamāʿa* 'you folks').
(5) Connective clauses: various kinds, e.g. *wa-ʾana lastu bi-ḥāja li-ʾan ʾuḏakkirakum ʾinnu* 'I don't need to remind you that'.

This classification demonstrates the function of certain connectives as 'framing' after the topical theme.

Table 1 presents a comparison of the results from the modern corpus with two shorter excerpts representing different Arabic discourse types, the first an extract of about 30 pages from a textbook on human geography for undergraduate level, the second about 44 pages from two different texts by a 9th-century physician. Each consists of 427 utterance segments, based on syntactic criteria rather than punctuation, since full stops mark off paragraphs rather than sentences.

Table 1 compares the results (in percentages) of the three corpora. The most obvious result of the comparison is that they all display a total dominance of *wa-* and *fa-* in class 2. The adverbial phrasal and clausal elements are primarily represented by classes 1, 3, 4, and 5, which taken together represent approximately 46 percent of the total number of the connectives in the corpus. For the textbook the equivalent percentage is 29 percent, and for the Classical text 11 percent. The results may be summarized as follows: both the textbook and the Classical text have higher frequencies for the conjunctions *wa-* and *fa-* than the corpus, the frequency of *fa-* in the Classical text being almost three times larger than in the corpus. These texts also have fewer pre-topical connective clauses than the corpus and no continuatives at all. The Classical text has the lowest number of pre-topical adverbials and clauses.

There is, therefore, a tendency in written Arabic from different periods to have a greater number of *wa-* and *fa-* than in modern oral Arabic. Medieval written Arabic has a substantially larger number of the connective *fa-* by comparison to modern Arabic. Modern Standard Arabic makes use of many more connective clauses than Medieval Arabic, and modern spoken Arabic uses more of them than written Arabic. Written Arabic regardless of period does not seem to make use of continuatives, which suggests that they are to be seen as a pragmatic oral device.

The question is whether Arabic has followed the same trend as English when it comes to allowing more adverbial phrases and clauses in frontal position in more recent periods. Even casual observation of Arabic lecturing discourse does suggest some kind of transfer of patterns common to the global scientific language (Holes 1995:269). English, as a Germanic language, has been the dominant influencer of modern international scientific style. The Germanic languages have witnessed a steady process

Table 1. Percentages by class and by corpus showing *wa-* and *fa-* separately

Class	Corpus %			Textbook %			Classical Text %		
1	8			0			0		
2	54	*wa-*	36	71	*wa-*	54	89	*wa-*	58
		fa-	10		*fa-*	15		*fa-*	28
		other	8		other	2		other	3
3	20			19			7		
4	5			4			0.5		
5	13			6			3.5		

involving the grammaticalization (or pragmaticalization) of adverbs and clauses expressing 'speaker-comments' in the direction of preposed sentential adverbials for several hundred years (Swan 1988:538–539). In English, in the 20th century, there was a virtual boom in the formation of new modal and evaluative utterance-starters. The speech-act adverbials, for instance, were non-existent before the 20th century, e.g. *frankly*, *broadly speaking*, etc. By quantifying sentence adverbials in English corpora from different periods, Swan shows that all classes of sentence adverbials (except subject disjuncts) have increased over the years, as well as the number of members in the different classes. Although most of them occur 'post-subject', the preposed ones have increased as well. Of all speech-act adverbials in English, 62.7 percent are in the initial field (Swan 1988:514–539).

Both Blau (1977) and Kinberg (1985) attest to a more extensive use of fronted adverbial constructions (phrases and clauses) in Modern Hebrew and Modern Standard Arabic. Kinberg adds that there is a greater variety of adverbial clauses in Modern Standard Arabic generally – fronted or not – in comparison with earlier periods. This is due, in his view, to the increasing usage of hypotactic structures in Modern Standard Arabic in contrast with Classical Arabic, which is manifested by the development of many new conjunctions, or the extended function of old conjunctions, as introducers of these dependencies. An example of this trend is that originally temporal conjunctions like *ʾiḏā*, *ʾiḏ* 'then' have also assumed conditional function. When adverbial hypotactic clauses do occur in Classical Arabic, it is generally clause-final. Exceptions to this are the conditional clauses and to a certain extent temporal clauses. Kinberg's observation is further that fronted adverbial clauses (expressing cause, purpose, concession, comparison, etc.) are more common in modern written journalistic style than in *belles lettres* (Kinberg 1985:52).

Blau (1977) believes that the instrument for the change is the construction of topicalization. He explains the conjunctions ending in *-mā*, which are so frequent in Modern Standard Arabic, as a result of a process of → grammaticalization, where a topicalized adverb framed by *mā*, coalesced with *mā*, lost some of its lexical content and became a simple sentence adverb or a connective, e.g. *kaṯīran mā* 'often' or *baynamā* 'while, whereas'. Meanwhile, *mā* lost its productive function as a rheme marker, a function well attested in Classical Arabic, e.g. *fa-bi-ʿizzinā mā tajabbarnā ʿalayhim* 'it was by our power that we evinced haughtiness against them' (Blau 1977:74).

Kinberg (1985) shares the view that topicalization has been instrumental in language change (or at least stylistic change) in Arabic in that it was extended to accommodate more than the usual handful of formulas. Thus, the rather strong constraint on utterance starts in Classical Arabic could be lifted. The typical and most original formula of topicalization is *ʾammā* + NP + *fa-* as rheme marker. The construction could also consist in *ʾinna/ʾanna* + NP + *fa-*, or simply the unframed noun phrase followed by *fa-* (*casus pendens*). In Classical Arabic, adverbial phrases or clauses (mostly conditional) occasionally occupied the framed slot, but with time this usage increased; and in Modern Standard Arabic, *ʾammā* is often replaced by alternative expressions, such as *fīmā yataʿallaqu bi-* or *bi-n-nisbati ʾilā* 'concerning', 'as regards', or nothing at all. After these latter topicalizers, the presence of *fa-* is no longer obligatory. The purpose of Kinberg's study is to show that adverbial clauses are fronted "under conditions similar to those characterizing other types of topics separated from their comments."

The increased use of adverbial connectives is one of the conclusions of the analysis of connectives. If applied to other corpora of different Arabic discourse types (oral and written, old and new) interesting patterns may emerge. Such patterns can be categorized and used as a tool for categorizing discourse in Arabic.

BIBLIOGRAPHICAL REFERENCES

Al-Batal, Mahmoud. 1985. *The cohesive role of connectives in a modern expository Arabic text*. Ph.D. diss., University of Michigan.

——. 1994. "Connectives in Arabic diglossia: The case of Lebanese Arabic". *Perspectives on Arabic linguistics*, VI, ed. Mushira Eid, 91–119. Amsterdam and Philadelphia: J. Benjamins.

Badawi, Elsaid, Michael G. Carter, and Adrian Gully. 2004. *Modern written Arabic: A comprehensive grammar*. London: Routledge.

Blakemore, Diane. 2002. *Relevance and linguistic meaning: The semantics and pragmatics of discourse markers*. Cambridge: Cambridge University Press.

Blau, Joshua. 1977. *An adverbial construction in Hebrew and Arabic: Sentence adverbials in frontal position separated from the rest of the sentence*. Jerusalem: Israel Academy of Sciences and Humanities.

Cantarino, Vicente. 1974–1975. *Syntax of modern Arabic prose*. 3 vols. Bloomington: Indiana University Press. *ūʾ*

Chafe, Wallace. 1994. *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*. Chicago: University of Chicago Press.

Dijk, Teun A. van 1977. "Connectives in text grammar and text logic". *Grammars and descriptions*. I. *Discourse as structure and process*, ed. Teun A. van Dijk and Janos S. Petöfi, 11–63. Berlin: W. de Gruyter.

Gully, Adrian. 1995. *Grammar and semantics in Medieval Arabic: A study of Ibn-Hisham's* 'Mughni l-Labib'. Richmond, Surrey: Curzon Press.

Halliday, M.A.K. 1994. *An introduction to functional grammar*. London: Edward Arnold.

—— and Ruqaiya Hasan. 1976. *Cohesion in English*. London: Longman Group UK.

Holes, Clive. 1995. *Modern Arabic: Structures, functions and varieties*. London and New York: Longman.

Johnstone, Barbara. 1990. "'Orality' and discourse structure in Modern Standard Arabic". *Perspectives on Arabic linguistics*, I, ed. Mushira Eid, 215–33. Amsterdam and Philadelphia: J. Benjamins.

Kammensjö, Heléne. 2004. *Discourse connectives in Arabic lecturing monologue*. Ph.D. diss., Göteborg University.

Kinberg, Naphtali. 1985. "Adverbial clauses as topics in Arabic: Adverbial clauses in frontal position separated from their main clauses". *Jerusalem Studies in Arabic and Islam* 6.353–416. Leiden: E.J. Brill. (Repr. *Studies in the linguistic structure of Classical Arabic*, ed. Leah Kinberg and Kees Versteegh, 43–102. Leiden: E.J. Brill, 2001.)

Lyons, John. 1977. *Semantics*. 2 vols. Cambridge: Cambridge University Press.

Reckendorf, Hermann. 1895–1898. *Die syntaktischen Verhältnisse des Arabischen*. Leiden: E.J. Brill. (Repr. 1967.)

——. 1921. *Arabische Syntax*. Heidelberg: C. Winter.

Sarig, Lea. 1995. "Discourse markers in contemporary Arabic". *Zeitschrift für Arabische Linguistik* 30.7–21.

Schiffrin, Deborah. 1987. *Discourse markers*. Cambridge: Cambridge University Press.

——. 2001. "Discourse markers: Language, meaning, and context". *The handbook of discourse analysis*, ed. Deborah Schriffin a.o., 54–75. Malden, Mass.: Blackwell.

Swan, Toril. 1988. *Sentence adverbials in English: A synchronic and diachronic investigation*. Oslo: Novus Forlag.

Warrāqī, Nārīmān Nāʾlī al- and ʾAḥmad Ṭāhir Ḥasanayn. 1981. *ʾAdawāt ar-rabṭ fī l-ʿarabiyya al-muʿāṣira*. Cairo: Center of Arabic Studies, American University in Cairo.

Wright, William A. 1933. *A grammar of the Arabic language: Translated from the German of Caspari and edited with numerous additions and corrections by William Wright*. 2 vols. 3rd. ed., rev. W. Robertson Smith and M.J. de Goeje. Cambridge: Cambridge University Press.

Heléne Kammensjö (Göteborg, Sweden)

Consonant → Phonetics; Phonology

Consonant Cluster → Epenthesis

Construct State

1. Definition

The construct state (CS) is a syntactic phrase consisting of at least two members, mostly nouns that are in a genitive relation. The sentence in (1) provides a typical example of the construct state, where the first noun carries the main case of the phrase, which can vary according to whether the noun phrase is nominative, accusative, or genitive. The second noun always carries genitive case.

(1) Standard Arabic
kitāb-u/a/i *l-muʿallim-i*
book-Nom/Acc/Gen the-teacher-Gen
'The teacher's book'

The members of the construct state do not have to be nouns. Adjectives can form a construct state complex with a noun, as illustrated in (2).

(2) Standard Arabic
rajul-un *ṭawīl-u* *l-qāmat-i*
man-Nom tall-Nom the-size-Gen
'A tall man'

Moreover, some numerals (3a) and quantifiers (3b) can also enter into construct state formation.

(3) Standard Arabic
a. *xamsat-u* *kutub-in*
five-Nom book-Gen
'Five books'

b. *kull-u* *l-kutub-i*
all-Nom the-books-Gen
'All the books'

Another construct state phrase in Arabic that has figured prominently in syntactic debates involves gerundive nominals (Hazout 1990; Fassi Fehri 1993). In this context, the first member of the construct state behaves like a verbal element in that it can assign accusative case to its object.

(4) Standard Arabic
qaṣf-u *l-ˁaduww-i* *l-madīnat-a*
shelling-Nom the-enemy-Gen the-city-Acc
'The enemy's bombardment of the city'

The properties of the construct state that have attracted attention within modern syntactic theories, particularly within the generative paradigm (Aoun 1978; Borer 1988, 1996; Mohammad 1988, 1999; Ouhalla 1991; Ritter 1991; Fassi Fehri 1993; Siloni 1997; Benmamoun 1997, 2000; Shlonsky 2004) are: (a) the restriction of the (in)definiteness marking to the last member of the construct state; (b) the requirement that the adjectives modifying the members of the construct state follow the whole construct state complex; and (c) the tendency of the construct state complex to display the phonology of words rather than phrases, i.e., prosodically it behaves as a word.

2. The distribution of the (in)definiteness feature

Starting with (in)definiteness, its most intriguing aspect is that the overt marker is restricted to the last rightward member of the construct state. The other members do not carry any marking of (in)definiteness but they are semantically interpreted as definite or indefinite depending on the overt marking on the last member. This is illustrated in (5).

(5) Standard Arabic
a. *kitāb-u* *bn-i* *ˁamm-i*
Book-Nom son-Gen uncle-Gen
ṣadīq-i *ṭ-ṭālib-i*
friend-Gen the-student-Gen
'The book of the son of the uncle of the friend of the student'

b. *kitāb-u bn-i* *ˁamm-i*
book-Nom son-Gen uncle-Gen
ṣadīq-i *ṭālib-in*
friend-Gen student-Gen.Indef
'A book of a son of an uncle of a friend of a student'

That the members other than the last one are semantically interpreted as definite or indefinite depending on the overt marking on the last member is shown by the fact that the adjectives that modify the other members overtly carry the markers of (in)definiteness, depending on their interpretation as illustrated in (6).

(6) Standard Arabic
a. *kitāb-u* *ṭ-ṭālib-i*
book-Nom the-student-Gen
l-jadīd-u
the-new-Nom
'The new student's book'

b. *kitāb-u* *ṭālib-in*
book-Nom student-Gen.Indef
jadīd-un
new-Nom.Indef
'A new student's book'

The distribution of the (in)definiteness marker suggests that the last member acts as a carrier of the feature for the whole construct state. The question then is how to account for the fact that non-last members of the construct state do not carry the feature, even though they are semantically interpreted as if they are carrying the feature.

3. Placement of adjectives within the construct state

Adjectives in Arabic follow the nouns they modify and agree with them in case, number, gender, and (in)definiteness. In the construct state, however, adjectives come after the last nominal member. For example, in (7a) the adjective after the second member of the construct state modifies the first member. Similarly, in (7b) the adjective after the third nominal member of the construct state modifies the second member.

(7) Standard Arabic
a. *kitab-u* *ṭ-ṭālib-i* *l-jadīd-u*
book-Nom the-book-Gen the-new-Nom
'The new book of the student'

b. *ġilāf-u* *kitāb-i* *ṭ-ṭālib-i*
cover-Nom book-Gen the-student-Gen
l-jadīd-i
the-new-Gen
'The cover of the new student's book'

The main issue that arises in the context of data such as (7) concerns the placement of the adjective and how that placement can help capture the concord relation between the adjective and the noun it modifies, given the fact that on the surface the two elements may not necessarily be adjacent.

4. THE PROSODY/PHONOLOGY OF THE CONSTRUCT STATE

The third important property of the construct state is its word-like behavior. This has been argued for on the basis of stress placement and vowel reduction in Hebrew (Borer 1988). With respect to Arabic, it can be illustrated by the distribution of the final consonant of the feminine marker *-at* (Benmamoun 2000). In Moroccan Arabic, as in Standard Arabic and other dialects, the final consonant of the feminine marker on the noun can be deleted (left unpronounced). Thus, *mədrasat* ʿschool' surfaces as *mədrasa*. However, if the noun is followed by a → clitic, the consonant cannot be deleted:

(8) Moroccan Arabic
a. *mədrast-i*
school-my
'my school'

b. **mədras-i*
school-my
'My school'

The distribution of the final *t* can be accounted for by a rule or constraint that bans it from being word-final. Thus, when the noun carries a clitic pronoun, the *t* is no longer word-final and therefore cannot be deleted.

Interestingly, in the construct state, the final *t* is not deleted, as illustrated in (9):

(9) Moroccan Arabic
a. *mədras-t* *nadya*
school-Fem Nadia
'Nadia's school'

b. **mədras-a* *nadya*
school-Fem Nadia
'Nadia's school'

Within the phrasal domain in Arabic this distribution of the final consonant of the feminine marker is restricted to the construct state. In the so-called free state, illustrated in (10), the *t* is deleted. Similarly, if the noun is followed by an adjective or relative clause (11), the *t* is deleted.

(10) Moroccan Arabic
a. *l-mədras-a* *dyal* *nadya*
the-school-Fem of Nadia
'Nadia's school'
b. **l-mədras-at* *dyal* *nadya*
the-school-Fem of Nadia

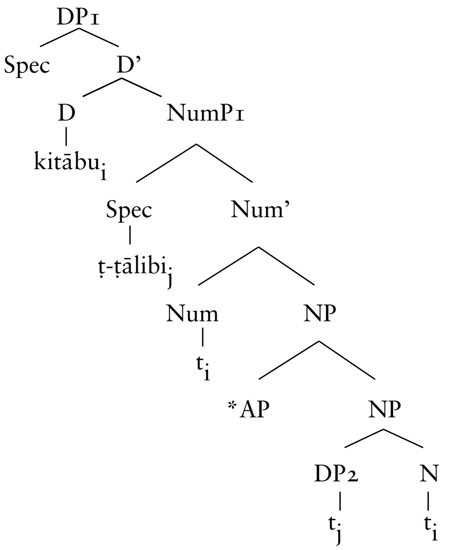
(11) a. *mədras-a* *jdid-a*
school-Fem new-Fem
'new school'
b. *l-mədras-a* *lli* *ħna*
the-school-Fem that here
'The school that is here'

5. SYNTACTIC ANALYSES OF THE CONSTRUCT STATE

The construct state in (4) headed by a gerundive nominal displays a Nominal Subject Object pattern (NSO), which parallels the Verb Subject Object (VSO) pattern in sentences. The VSO pattern in sentences has been argued to be derived from a basic SVO order by verb movement to a functional position higher than the projection containing the subject (Fassi Fehri 1993; Mohammad 1999). A similar analysis has been proposed for the NSO pattern. The underlying order is Subject Nominal Object (SNO), which yields the NSO order by movement of the N to a higher functional projection. The question then is whether this analysis can be extended to other construct state phrases that do not contain gerundive nominal heads.

Most recent analyses claim that the answer is positive. They assume that the construct state contains a lexical projection where the genitive noun phrase is in the Spec(ifier) of a lexical ominal projection headed by the head of the construct state. This lexical projection is in turn dominated by one or more functional projections, one of which, D(eterminer) or Agr(eement), plays a role in genitive case assignment. Most analyses assume the representation in (12) proposed by Ritter (1991) for Hebrew or a variant of the same representation with different labels of the functional categories above the lexical core headed by the noun head of the construct state.

(12)



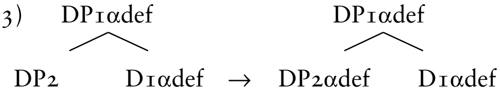
In the representation in (12), the head noun N is generated in the lower N(oun) P(hrase) projection with D(eterminer) P(hrase)2 (the genitive noun) as its specifier. The surface order is derived by movement of the head noun to the DP projection via the number projection.

The adjective modifying the noun is left adjoined to the NP in a position higher than the base position of the genitive noun (DP2). The latter must then move to the Spec of NumP to derive the order whereby the adjective follows both members of the construct state. This analysis does derive the locality of the concord relation between the adjective and the noun it modifies. The adjective is adjoined to the NP and therefore must modify the NP, regardless of where the latter ends up in the structure on the surface.

6. ANALYSES OF THE DISTRIBUTION OF THE (IN)DEFINITENESS FEATURE

How can one then account for the fact that a noun within the construct state can be interpreted as definite or indefinite, despite the fact that it does not carry an overt marker for the feature in question? Some approaches have attempted to provide structural analyses. For example, Fassi Fehri (1993:225–232) argues that because the genitive NP (DP2 in 12) does not overtly raise to the Spec of DP1, which would have allowed it to acquire the definiteness feature from D (the host of this feature), DP2 might raise to the Spec of DP1 at a subsequent abstract level, L(ogical) F(orm). A simplified illustration of the derivation prior and subsequent to the movement of DP2 to Spec DP1 is given in (13).

(13)



The movement of DP2 to the Spec of DP1 only guarantees the proper semantic interpretation but not the overt phonological realization of the feature, given that LF does not feed the morphology, (so-called phonological form [PF] component). Variants of this analysis have been provided for Hebrew as well (Siloni 1997). Analyses vary with respect to details, such as whether the head of construct state enters the derivation already specified for (in)definiteness, which is then checked by the genitive NP, or whether it comes with the feature unspecified and acquires it when it enters into a Spec-head relation with a head that carries the feature. The analyses, though they vary in the details, share the fundamental assumption that the distribution of the (in)definiteness features on the genitive NP are sensitive to the syntactic representation and derivation. In other words, the issue is considered to be purely syntactic.

7. INTERACTION BETWEEN THE PROSODY OF THE CONSTRUCT STATE AND THE DISTRIBUTION OF (IN)DEFINITENESS

Purely syntactic analyses, while somewhat successful in accounting for the distribution and (in)definiteness feature, do not clearly relate this property to the fact that the members of the construct state form a prosodic unit and behave phonologically as a word. Moreover, one challenge that remains for the purely syntactic analysis comes from the (in)definiteness feature on adjectives. The latter carry it overtly, which implies that they do not 'wait' for the NP they modify to acquire it at a subsequent level of analysis after the displacement of the genitive NP.

Benmamoun (1997, 2000) argues that the distribution of (in)definiteness in the context of the construct state is intimately related to its prosodic nature. He specifically argues that the members of the construct state, which are all lexically specified for the (in)definiteness feature, merge to form a single prosodic unit. This in turn obviates the need to realize the (in)definiteness feature on all members. The last member carrying the relevant feature will act as an exponent of that feature on the other members. In other words, the absence of the (in)definiteness marker (definite article or indefiniteness suffix) on a member of the construct state does not imply that it is absent. It just means that it is spelled out differently. Since adjectives may not merge with the NPs they modify in the construct state, the (in)definiteness feature they carry must be spelled out by a morpheme.

Benmamoun further argues that construct state formation does not take place in the lexicon, prior to the syntactic derivation. This is shown by the fact that the construct state unit does not constitute an anaphoric island. For example, the first member can be modified by a relative clause (14a), while the second member can be a negative polarity item licensed by the sentential negative (14b).

(14) Moroccan Arabic
a. *ktab t-tarix lli t-tənšəṛ*
book the-history that Pass-published
lbarəḥ
yesterday
'The history book which was published yesterday'

b. *ma-qrit ktab ḥətta waḥəd*
Neg (I) read book any one
'I didn't read anybody's book'

Moreover, there is no upper limit on the number of construct state members, which is not typical of lexical word formation processes. It also seems that the construct state formation does not take place in the syntax. One main argument against a syntactic analysis is that the construct state involves the combination of a head and a phrase, which is not typical of syntactic derivations. For example, the second member can be a full sentence (Mohammad 1999) as illustrated in (15).

(15) Palestinian Arabic
sāˁet [mā šuft-ak]
hour when saw.1sg-you
'the hour when I saw you'

The most likely scenario is that the merger of the members of the construct state takes place post-syntactically, in the so-called PF interface. The transparency of the construct state follows because all members are independent throughout the syntactic derivation. PF is also the component where it is argued that morphosyntactic features should be spelled out, which then makes it a plausible place for construct state formation, since one of its key features is the peculiar distribution of the (in)definiteness feature.

8. Other remaining issues

Other interesting issues that still need further investigation concern the ordering of adjectives within noun phrases in general and the construct state in particular (→ adjectival clause). It is not clear whether Arabic adjectives adhere to a basic ordering and serialization schema depending on their semantic content, as has been claimed for other languages (Fassi Fehri 1998; Shlonsky 2004). The evidence is difficult to pin down as the judgments of the native speakers are not firm and stable. Another issue concerns the syntax of adjectives, whether they should be viewed as adjuncts or heads of their own syntactic projection (Fassi Fehri 1998; Shlonsky 2004). The latter view has gained some currency within syntactic analyses and has been coupled with the idea that syntactic movement within the construct state does not affect heads but only maximal projections. But it comes at the cost of positing highly abstract syntactic representation with syntactic projections and movement operations that are difficult to motivate independently. Clearly, there are many more questions that are still open despite the fact that the construct state has received a large, if not the largest, share of attention within Semitic syntax.

Bibliographical references

Aoun, Joseph. 1978. "Structure interne du syntagme nominal en Arabe: L'ʔidafa". *Analyses et Theorie* 1.1–40.

Benmamoun, Elabbas. 1997. "Agreement in Arabic and the PF interface". *Proceedings of the West Coast Conference on Formal Linguistics* 15.33–47. Stanford, Cal.: Stanford University, Center for the Study of Language and Information.

——. 2000. *The feature structure of functional categories: A comparative study of Arabic dialects*. Oxford: Oxford University Press.

——, Mushira Eid, and Niloofar Haeri (eds.). 1999. *Perspectives on Arabic linguistics*, XII. Amsterdam and Philadelphia: J. Benjamins.

Borer, Hagit. 1988. "On the morphological parallelism between compounds and constructs". *Year-*

book of morphology, ed. Geert Booij and Jaap van Marle, 45–65. Dordrecht: Foris.
——. 1996. "The construct in review". *Studies in Afroasiatic grammar*, ed. Jacqueline Lecarme, Jean Lowenstamm, and Ur Shlonsky, 30–61. The Hague: Holland Academic Graphics.
Fassi Fehri, Abdelkader. 1993. *Issues in the structure of Arabic clauses and words*. Dordrecht: Kluwer Academic Publishers.
——. 1999. "Layers in the distribution of Arabic adverbs and adjectives and their licensing". Benmamoun a.o. (1999:9–46).
Hazout, Ilan. 1990. *Verbal nouns: Thetatheoretic studies in Hebrew and Arabic*. Ph.D. diss., University of Massachusetts, Amherst.
Mohammad, Mohammad. 1998. "On the parallelism between IP and DP". *Proceedings of the West Coast Conference on Formal Linguistics* 7.241–254.
——. 1999. "Checking and licensing inside DP in Palestinian Arabic". Benmamoun a.o. (1999:27–44).
Ouhalla, Jamal. 1991. *Functional categories and parametric variation*. London: Routledge.
Ritter, Elizabeth. 1991. "Two functional categories in Noun Phrases: Evidence from Modern Hebrew". *Syntax and semantics*, XXV, ed. Susan D. Rothstein, 37–62. New York: Academic Press.
Shlonsky, Ur. 1997. *Clause structure and word order in Hebrew and Arabic*. Oxford: Oxford University Press.
——. 2004. "The form of Semitic noun phrases". *Lingua* 114.1465–1526.
Siloni, Tal. 1997 *Noun phrases and nominalizations: The syntax of DPs*. Dordrecht: Kluwer Academic Publishers.

Elabbas Benmamoun
(University of Illinois at Urbana-Champaign)

Contact → Language Contact

Contrastive Grammar

1. Definitions/preliminaries

The term 'contrastive grammar' connotes an approach to language teaching dominant in the 1960s and 1970s, which has since been discredited because of its extremely narrow definition: structural differences between languages do not exhaustively explain learning difficulties; they are neither necessary nor sufficient conditions for explanation. As a consequence, the pendulum swung the other way in the 1980s, when structural analysis had almost fallen into disuse. In recent contributions, a more differentiated modeling, covering the empirical variation in learner varieties has been attempted, operating with a more complex theoretical framework. Rather than simply comparing the linguistic forms that learners produce with the normative target, the aim is now to reconstruct the various learner varieties ('interlanguages'), taking into account the resources to which learners have recourse, be it their knowledge of their first language (L_1) or the acquired forms of the second language (L_2), and thus making use of the potential space in the learning progress for productive teaching interventions (see the articles in Ritchie and Bhatia 1996 for a survey).

'Contrast' evidently implies a 'comparison', thus in a certain sense all comparative linguistic work is contrastive. With a narrower scope, contrastive analysis focuses on a restricted set of languages, usually a pair of languages. The choice is not arbitrary but generally has practical motivations: while a comparison, say, of Arabic and Inuktitut (Eskimo) might be quite enlightening in a typological perspective, it does not correspond to a relevant practical question. Contrastive analysis is usually implemented in two kinds of institutionalized settings:

i. → Second Language Teaching, where the target language corresponds to political factors (e.g. Arabic as the language of Islam, of an economically important geographical area, etc.), without necessarily implying other cultural and/or social relations (or any special kind of contact between the languages and/or the linguistic communities);
ii. the → language policy of bilingual societal relations, be it in autochthonous constellations (e.g. with Berber, Kurdish, or other non-official languages spoken alongside the official language Arabic), or in allochthonous constellations (e.g., as a consequence of the recent immigration of speakers of Arabic to the West European industrial states).

Evidently, it makes a difference if Arabic is the target or the source language (as in the case of Arabic-speaking immigrants). These constellations set the parameters of contrastive analyses, which run the risk of being incompatible with one another because they focus on the salient features occurring in each situation in an effort to explain difficulties or 'errors' of particular learners. In other words, the analytic framework will be different in the case of a Berber-speaking pupil learning Arabic in a Moroccan school from that of an Arabic-speaking pupil learning German in a German school, or an English-speaking high-school student learning Arabic in the United States. In addition to the dominant social, cultural, and motivational factors that undoubtedly combine to make all these constellations different, structural differences play an important role: Berber and Arabic are to a certain extent structurally homologous languages, which explains the astonishing speed with which monolingual Berber-speaking people become bilingual after migrating to urban centers in Morocco, whereas European residents living there usually remain at the level of a very restricted proficiency in Arabic, despite considerable efforts to learn the language over many years. A more analytic approach demands a typological framework.

What makes the situation so complicated here is the considerable variation within what is referred to as Arabic. The vast literature on → diglossia has shown that this term rather simplifies the situation. There is an enormous distance between the high variety (in the following the Arabic term *fuṣḥā* will be used for this variety) and the colloquial forms (in the following referred to by the Moroccan Arabic term *dārija*; the Egyptian equivalent is *ʿāmmiyya*). The distance can be tested by presenting informants without school training in *fuṣḥā* with spoken forms of Standard Arabic, e.g. television news, and ask them to repeat sentence-by-sentence. The result (at least when this test was applied to Moroccan informants) generally produces rather strange replies. Still, knowledge of the *dārija* provides different resources for learning the *fuṣḥā* from those that speakers of, for instance, English have at their disposal, and any contrastive analysis must take these typological differences into account (see Dichy [1994] for a differentiated grid of distinctions which takes into account the dimension of learnability in different linguistic situations).

An impediment to analysis is the often insufficient differentiation between → register variation in language use, and differences in the linguistic structure that articulate these registers. For the sake of argument it will be sufficient to stick to the traditional ('rhetorical') model of register variation (see Biber [1995] for detailed models and recent research; registers, defined by properties of linguistic structures, are to be distinguished from domains, defined by language external criteria). The model distinguishes between:

i. the *formal register*, which monitors the formal articulation, used in ceremonial contexts as well as for correcting errors, misunderstandings, and the like. Writing, if not practiced in a different language, is based on this formal register;
ii. the *informal register*, used in public places (the market etc.). This register is learnt by taking into account the language of others and therefore, it shows traces of koineization;
iii. the *familiar register*, used in the intimacy of the family and with close friends and peers.

The comparison of these registers is the domain of contrastive analysis as well. The idea that these registers should be articulated by the same langage (representing varieties of this language, usually identified with the national language, and taught as a written language in the school), is a modern concept, bound to the construction of the nation state. Probably, in most parts of the world, these different registers correspond to different languages, people generally being bi- or multilingual, depending on the social domains they participate in. → Diglossia in the narrow sense of the word represents a case between these extremes, where all registers are articulated by the 'same' language, but where most of the population does not participate in domains in which the formal register (the *fuṣḥā*) is used.

There are at least two dimensions that must be taken into account: language external factors which govern the distribution of linguistic varieties (defining the minor and the major languages of a particular society); and the language internal factor of structural distance between varieties, which allows abductive (Andersen 1973) generalizations by which utterances in one variety can be transformed into other varieties, as is the case with modern Arabic varieties. But the empirical situation is more complicated: here, factors of external valorization come into play, such as investing the *fuṣḥā* with emblematic power and devaluating the *dārija* as a kind of corrupted language. Consequently, in Arabic-speaking countries, linguistic knowledge of the *dārija* usually remains strictly procedural, while only the *fuṣḥā* admits conscious cognitive activity, leading to the kind of declarative knowledge that can be used in teaching. Traditional learning of the *fuṣḥā* with examples from the *Qur'ān* can represent a supplementary barrier. There is a tendency in Arabic universities to illustrate discussions of grammatical phenomena with Qur'ānic quotations. This blocks any analytic operations (substituting or permutating forms), since the text cannot be changed. In these cases it is at least very helpful to be prepared for the internal variation in the Qur'ānic text. Descriptive work such as Reuschel (1996) about the verbal system are extremely useful in this respect.

Contrastive analysis must clarify on what register level the comparison is defined. Much of the sampling in modern typological research, even if it is carried out in a sophisticated manner, falls short of this criterion. Recent work in teaching Arabic as a foreign language tries to integrate this register variation (e.g. Younes 1990). For a comprehensive investigation of the varieties of spoken Arabic, see Brustad (2000).

2. Structural vs. piecemeal comparison

Traditional approaches have in common with many recent typologically oriented discussions that they choose particular salient linguistic features for comparison, taking the general linguistic structure for granted. Renewed support for this approach is to be found in the recent focus on universals, which defines empirical questions as a matter of parameter setting within universal grammar. This is in contrast to the structuralist traditions, where linguistic systems are in focus (and thus linguistic variability), characterized by internal structural constraints to be discovered by descriptive work in the wake of Boas and his followers. This way of thinking has recently been matched by systematic psycholinguistic research, where structure building is seen as enforced by cognitive constraints to harmonize the different options in a given linguistic structure, as for instance in the work of Hawkins (1987). The background to this approach is the Greenbergian research program of implicational relations between structural features in language.

These cognitive constraints are at work in any kind of bilingual situation, restricting the potentials for the management of structural diversity. Very instructive is the analysis of → code-switching, where speakers not only cope with the structural demands of different languages in contact but also make use of these to convey a kind of secondary meaning (the 'contextualization' of utterances in the sense of Gumperz [1992]). Code switching in this special sense is to be distinguished from those types of code mixing which result from a lack of resources in the target language (cf. Fakhri 1984). Switching between different varieties or languages presupposes the ability to match the potential structures which articulate the translation equivalents in the varieties in question. A quite elaborate framework for this kind of analysis is provided by Myers-Scotton's Matrix Frame Theory (1993), which has guided much research in contact situations of Arabic, especially in the context of migration (e.g. Nortier 1989; Boumans 1998) and has turned out to be very fruitful in second language acquisition research (see Ritchie and Bhatia 1996), since every kind of bilingual situation (including second language learning) requires the harmonization of the structural resources; see, e.g., Eid (1996) for a study exploring this difference.

3. Structural patterns and language learning

The focus on structural patterns, which led to the pattern drill programs in language laboratories in the 1960s and 1970s, masked the complexity of the phenomenon of language learning. The contextualization of grammatical structures in language learning has only seldom been the subject of research (see Kniffka [1995] for some more anecdotal observations on teaching English and German in Arab universities). Arabic lends itself to a form-biased perspective, given its salient complex morphological architecture, which led Sapir to put it at the top of his complexity ranking scale of linguistic structures (Sapir 1921:142–143). The isolation of patterns has found a certain continuation in computer linguistic work, for instance in the construction of algorithms to compute the non-concatenative morphology of Arabic (see, e.g., Kiraz 1999).

These systemic patterns can play a very different role in language learning: children, learning in a social moratorium and striving for acceptance as native speakers experiment with form differentiations, while adults, learning a (second) language in the context of the requirements for making a living, are content with a pragmatic use of linguistic resources, generally restricting these to the minimum sufficient for conveying their intended message (i.e., what is often termed the 'basic variety' of the target language). This economic alignment to efficient communication with restricted means can seal the fossilization of 'interlanguage' varieties. Children (and learners with a professional motivation), however, strive for refinement of formal patterns. This is clearly shown by the result of various long-term research projects on language learning in an immigrant situation (e.g. Perdue 1996). Investigating these practices shifts the focus from error analysis, i.e. matching the learner's deficient practices to the patterns of the target language (and registering interferences by L_1), to the resources exploited by the learner in an effort to cope with the task at hand. L_1-structures are then not seen as a factor of disturbance but rather as a resource, defining the space of possible progress in learning (Mansouri 2000). In this perspective, for instance Fakhri (1984) analyzes the 'pleonastic' use of full pronouns to secure reference tracking in a case of otherwise insufficient control of Arabic verbal inflection.

At the opposite end are questions of cognitive orientation bound to grammatical structures. Whereas this issue is rather dominant in lay discourse about language, it has for quite a long time been tabooed in professional linguistic circles (cf. Gumperz and Levinson 1999). What has blocked research here is that this issue has traditionally been conceptualized too strictly. As Boas at the beginning of his systematic inquiry already showed a century ago, grammatical patterns (the grammaticization of cognitive structures) serve as default structures in coping with cognitive tasks, not as insuperable barriers. These orientation functions of grammaticized patterns have recently been highlighted in the work of Talmy (2000), whose ideas on different framing of propositional content in typologically different languages has proven to be a very productive guideline for empirical research. Berman and Slobin (1994) show, for instance, that different narrative patternings are preformed by grammatical structures, where language is "thinking for speaking" (Berman and Slobin 1994:594). The difference posited by Talmy between verb-framed vs. satellite-framed structures can be usefully exploited, as Arabic is predominantly verb-framed, coding the core information in the verbal lexeme and conveying information on the mode by secondary constituents. By contrast, in English or German, the core information is conveyed by a particle that is usually articulated in the focus position. Arabic *xaraja* can be translated in English, for instance, by *he went/ran/flew/drove/swam OUT*; *daxala* can be translated in English by *he went/ran/flew/drove/swam IN*, etc. Another example is the different use of verbal categories: the use of aspect in Arabic for staging background vs. foreground in the narrative contrasts with the use of tense marking in, for instance, English or German for staging the chronological order of the event related, anchored in the time of the utterance.

4. The identification of structural constraints

Empirical research must disentangle linguistic structure and non language-specific cognitive factors. The prevailing paradigm of linguistic thinking assumes a structural core (universal grammar) as the initial state of language learning. Such universal structures should restrict second language learning as well, without making L_2-acquistion a simple recapitulation of L_1-acquisition. Second language acquisition has to exploit the resources of L_1 and to harmonize them with the structures of L_2. In the case of conflict, general cognitive constraints (and knowledge of the world) control the interpretation of utterances and in borderline cases grammatical form is dispensed with. This is what is termed → 'pidginization', a process which corresponds to rudimentary steps in second-language learning (see Pienemann [1998] for a detailed modelization and Mansouri [2000] for an application to Arabic as L_2).

Evidently, in this perspective the critical notion is structural correspondence. This is a highly theory-bound concept. The assumption of universal concepts makes comparison easier, but at the same time runs the risk of circular argumentation. There is a widespread postulate of sentence structure as *subject* + *predicate* in accordance with the tradition of European school grammar, which conflates pragmatic conceptualization (→ topic – comment/→ theme – rheme structure), semantic roles (the subject as prototypical actor), and formal constraints (agreement). This conceptualization creates problems with Arabic sentence structure:

i. the *subject + predicate* scheme fits nominal sentences quite well since, evidently, these lack semantic roles, which are bound to the verbal predicate as head of the clause;
ii. this scheme fits verbal sentences as well, when they have a topicalized complement (the Arabic grammatical tradition treats these as nominal sentences); in fact, at least for most New Arabic varieties, most descriptions register a tendency towards SVO-word order (for Moroccan Arabic, this is only true as far as definite 'subjects' are concerned, i.e., in the case of topics, not subjects in the grammatical sense; this is also the gist of most generative analyses made by native speakers, e.g. Lalami 1996);
iii. for all other types of verbal sentences, there are problems, however: for instance, no overt subject constituent is necessary, and verbal agreement is not found when the verb is initial.

There are different options for resolving this conflict found in the literature:

i. a → pro-drop analysis can save the analytic scheme in 'deep structure';
ii. alternatively, the subject can be located in an incorporated pronoun (e.g. *-a* in *xaraja* as a bound allomorph of *huwa* 'he'), which means that a corresponding nominal form must be interpreted as adjunct, possible a topicalized one (*ʿalī xaraja* 'Ali went away');
iii. mismatches in the agreement can be explained away by stripping the verb of critical inflectional categories.

Whatever the merits of these analyses (no detailed references are needed; Fassi Fehri [1988] being an explicit example), a less theoretically biased approach would try to profile the peculiarities of Arabic. 'Subject' is a grammatical notion, which has to be distinguished from non-linguistic concepts, necessary as they may be for explaining linguistic behavior, for instance in learning situations. As Arabic (Classical as well as New Arabic varieties) possesses a category of diathesis, it makes sense to operate with the category of subject in verbal sentences. The subject can then be matched by the complement of the predicate in the nominal sentences. This differs from languages in which syntactic structure is immediately controlled by semantic roles (as, e.g., in Acehnese). From a descriptive point of view, the subject is optional in verbal sentences. If it is expressed, its position is controlled by the information structure of the sentence. Of course, there is more to pro-drop than the 'natural' order of the sentence cherished by medieval speculative grammar: it defines a cluster of syntactic features in a specific theoretical framework. The incorporation analysis depends upon assumptions that must be made explicit, if it is to be more than an *ad hoc* device. It presupposes that the predicate is defined as the non-referring part of the sentence. As a consequence, a referential expression can only be incorporated into the predicate (as in *xaraja* as a complete sentence). The price of this analysis is to treat the agreement marking in *ʿalī xaraja* as merely homophonous with the incorporated pronoun as, for instance, Fassi Fehri (1988) does. A similar analysis has to be made for nominal sentences with referential predicates, where the 'incorporated' pronoun obligatorily has the form of the 3rd pers. *ʾana huwa l-mujrimu* 'I am the culprit' (for the opposite traditional analysis see, e.g., Cantarino 1974–1975: II, 432–436).

The categories of verbal inflection present another much discussed example of the conflation of typologically diverse structures, especially tense and aspect. The assumption of a universal system of temporal differentiation masks the grammatical functions verbal forms play (as mentioned above for the narrative functions). Of course, the temporal anchorage of the interpretation of a sentence is always possible in any language, but the relevant question is whether or to what extent it is grammaticized. Apart from the semantic problems involved, a strict mapping of the tense system of European languages onto Arabic masks the fact that in New Arabic the verbal system has been restructured with the help of analytic forms, already present in Classical Arabic in constructions with → *kāna wa-ʾaxawātuhā*. As a consequence of this bias, it can be rather difficult to treat these questions in teaching at Arab universities, because of the orientation toward the synthetic verb paradigms of the European languages (cf. Maas a.o. 2000).

5. ORTHOGRAPHY

Orthographic questions must be distinguished from writing systems. For the layman, Arabic is bound to the Arabic script, and additionally charged with religious connotations via the transmission of the *Qurʾān*. But the *figurae* of the writing system (the script) have to be distinguished from their use, which is based on language-specific knowledge. Thus, Arabic languages can be written in different scripts (e.g. Maltese in Latin [Roman] script), and other languages can be written in the Arabic script (e.g. Persian, Urdu, Hausa, Berber) (→ Arabic Alphabet for Other Languages). Whereas graphic figures are but a minor burden on memory, habitualized patterns of orthography present much more substantial impediments to second language learning. Arabic orthography has a phonographic basis, similar to Roman (Latin) based orthographies, but it minimizes the graphic expression by maximizing the effect of syllabic and morphological filters: in conventional orthography only syllable margins are represented. In the traditional syllable-based analysis, this is called 'consonantal' writing, with only consonants to be represented, understood as the complements of the syllabic nucleus (the sonant).

Even more important is the morphological filter. Where phonological neutralizations are operative, they can be overwritten to represent lexico-grammatical invariants as, e.g., *ʾalif wa-lām* as an invariant representation of definiteness marking. This orthographic convention is restricted to a special case, word-internal morpheme-boundary. Lexical filters can block the representation of phonological variation, e.g. in the coda position, the place of articulation of nasals is controlled by the following onset, which permits the invariant representation of radicals in, e.g., *ʾanbiʾū* [ʾambiʾu] 'teach! [imper. pl.]' (root *n-b-ʾ*).

Traditionally, the writing system is not used to represent the colloquial languages, but where this is done in marginal practices, such as informal letter writing, quoted oral passages in literary works, filled bulbs in cartoons and comics, etc., it can be fruitfully exploited with but minor accommodations. Yet, the reorganization to a more concatenative structure in most New Arabic varieties leads toward a less grammatically controlled orthography, with a strictly linear phonographic representation, where all 'terminal leaves' of a syllable structure/tree are represented. Consequently, reading a text no longer presupposes its interpretation.

Where language-external considerations are not an impediment, radical changes can occur, as has happened in → Maltese. The situation is more complicated for Muslims with an Arabic linguistic background, for whom this writing system has emblematic functions. Yet, even for them a corresponding switch should not present practical problems, as this kind of notation is familiar from the *taškīl* in Qurʾānic texts as well as in primers. This step has already been taken anyway in cases where the *figurae* of writing were changed, as in the case of editions of the *Qurʾān* printed in Romance script, in Turkey, Malaysia, etc.

A more complex problem is the key concept of an orthographic word, bound to the grammatical structure of the language that is orthographically represented. The primarily suffixing inflection in the Indo-European languages defines the word by its head (the lexeme), such that it comprises a family of forms with variable tails. This is the foundation of alphabetic order for all kinds of dictionary. The predominantly internal variation of grammatical forms together with prefixing inflections requires a different concept of the word in Arabic, not aligned to the left edge. The traditional grouping by radicals in the dictionaries is the answer to this. The conflict between these different logographic orientations becomes evident in the case of immigrants who have to cope with both systems. Thus, children of Moroccan immigrants in Germany are guided in their writing by the concept of a German orthographic word, aligned to the invariant left edge. Interestingly, they stick to this orientation even when writing Arabic (as learned in Qurʾānic schools). If asked to write down sentences in their first language (Moroccan *dārija*) in Arabic letters, they tend to isolate the personal marking of the prefix inflection on the model of German subject pronouns (writing *yi mši* instead of *yimši* 'he goes'), something unheard of in a Moroccan context (see Maas and Mehlem 2003). More research is necessary in this field.

6. THE LINGUISTIC METALANGUAGE

Contrastive analysis must be represented in a metalanguage. While the descriptive linguistic tradition has tried to define its terminology operationally, in an effort not to bias the description, recent linguistic mainstream revives the traditional universalist approach, but in doing so replaces the canonical model of school grammar (Greek or Latin grammar) by theoretically postulated structures (Universal Grammar). This reintroduces the risk of biased descriptions, especially in much graduate work, for instance in dissertations on contrastive topics, in which the analysis is guided by the model language (English, French, German) whose structures are projected onto Arabic via translation equivalents. This is especially true of work in the Generative School that favors the assumption of universals extrapolated from structures of European languages (mostly English).

An example is the categorization of dependent clause constituents as infinitive constructions (and the assumption of infinitives) in Arabic, whereas Arabic actually has (asyndetic) finite clauses. As an example Ennaji (1985) may be quoted, still frequently hailed as one of the rare examples of contrastive syntactic analysis. Ennaji classifies in Moroccan Arabic the constituent *ibqa aḥmad f-ḍ-ḍaṛ* in *xtert ibqa aḥmed f-ḍ-ḍaṛ* 'I chose for Ahmed to stay at home' as an 'infinitival complement' and consequently, he categorizes the finite verb form *ibqa* as an 'infinitive'. This projection of familiar features of European languages onto Arabic misses characteristic differences. The morphological structure in clause junction is matched by the pattern of complex predicate formation in all New Arabic varieties, the prototype of which is already found in Classical Arabic with *kāna wa ʾaxawātu-hā*. The complementary classification of these modifiers as auxiliaries again obfuscates the particularities of Arabic. It is almost an exception to find analyses such as that of Fassi Fehri (1993:156–194), who speaks of "bi-inflectional structures", distinguishing complex sentence formation as in the example just quoted from complex predicates, as e.g. in *kāna ʾakala* 'he had eaten'.

In this regard, it makes sense to revaluate the national grammatical traditions (Kniffka 2001). Some examples may illustrate this (Maas a.o. 2000:15–18). The Arabic definiteness marker is often referred to as 'article', where the Arabic tradition uses the quite flexible term 'instrument of definiteness' (*ʾadāt at-taʿrīf*, from *ʾ-d-w* 'to accomplish'), which denotes a morphological device, not necessarily a word (as the term 'article' does in the Greek tradition of word classes).

Of course, this tradition can be misleading, as in the case of the relative marker. The usual Arabic term for forms like *allaḏī* is *ism mawṣūl*, *mawṣūl* representing the linking function of the form. Yet, *ism* is misleading as it is the term that otherwise designates the noun (substantive). Instead of *ism*, *ḍamīr* is sometimes used, which is not much better as it is the usual term for a pronoun (besides, in the colloquial language it means 'heart, interior', which is somewhat mystifying in this context). Nonetheless, the advantage of this native terminology is to prepare the learner for structural differences: the *ʾadāt at-taʿrīf* is a bound morpheme that is distributed over the elements of the nominal phrase (as agreement marking), rather than a word functioning as the grammatical head of the nominal phrase as, for instance, the article in German. The *ism mawṣūl* is a grammatical instrument (in fact, *ʾadāt* would have been a more suitable term in this context), serving as agreement marker of the definiteness of the noun phrase when the attribute is a clause.

The issue of the 'subject' also can be made less confusing if traditional Arabic terminology is used. Here, the focus is clearly on information structure, since the sentence construction is defined by the partitioning of its field into the initial (thematic) part (→ *ibtidāʾ*), and the final (rhematic) part (→ *xabar*), which also applies to a verbal sentence with a topicalized complement (as in the example quoted above, *ʿalī xaraja*). Yet, the terminology for the constituents of the verbal sentence has to be reanalyzed: instead of reading the representation by *f*, *ʿ*, *l* as the radicals of the verb *faʿala* 'to do', and thus confusing the term for the subject *fāʿil* with its semantic reading 'actor', these forms can be read as algebraic symbols that represent syntactic functions by the corresponding schemes.

Thus, taking recourse to the tradition of national Arabic grammarians can be of considerable help in clarifiying the often confusing issues of grammatical descriptions.

BIBLIOGRAPHICAL REFERENCES

Andersen, Henning. 1973. "Abductive and deductive change". *Language* 49.765–793.

Berman, Ruth A. and Dan Isaac Slobin. 1994. *Relating events in narrative: A crosslinguistic developmental study*. Hillsdale, N.J.: Erlbaum.

Boumans, Louis. 1998. *The syntax of codeswitching: Analysing Moroccan Arabic/Dutch conversation.* Ph.D. diss., University of Nijmegen.

Brustad, Kristen E. 2000. *Spoken Arabic. A comprehensive study of Moroccan, Egyptian, Syrian, and*

Kuwaiti dialects. Washington, D.C.: Georgetown University Press.

Cantarino, Vicente. 1974–1975. *Syntax of modern Arabic prose*. 3 vols. Bloomington, Ind.: Indiana University Press.

Dichy, Joseph. 1994. "La pluriglossie de l'arabe". *Bulletin d'Etudes Orientales* 46.19–42.

Eid, Mushira. 1996. "Things are not what they seem: Pronoun doubling in bilingual and monolingual grammars". *Perspectives on Arabic linguistics*, VIII, ed. Mushira Eid, 7–29. Amsterdam and Philadelphia: J. Benjamins.

Ennaji, Moha. 1985. *Contrastive syntax, English, Moroccan Arabic and Berber: Complex sentences*. Würzburg: Königshausen and Neumann.

Fakhri, Ahmed. 1984. "The use of communicative strategies in narrative discourse: A case study of a learner of Moroccan Arabic as a second language". *Language Learning* 34.15–37.

Fassi Fehri, Abdelkader. 1988. "Agreement in Arabic, binding and coherence". *Agreement in natural language*, ed. Michael Barlow and Charles A. Ferguson, 107–158. Stanford, Cal.: Center for the Study of Language and Information.

——. 1993. *Issues in the structure of Arabic clauses and words*. Dordrecht: Kluwer Academic Publishers.

Gumperz, John. 1992. *Discourse strategies*. Cambridge: Cambridge University Press.

—— and Steven Levinson (eds.). 1999. *Rethinking linguistic relativity*. Amsterdam and Philadelphia: J. Benjamins.

Hawkins, John A. 1987. "Implicational universals as predictors of language acquisition". *Linguistics* 25.453–473.

Kiraz, George Anton. 1999. "Computational tool for developing morphophonological models for Arabic". *Perspectives on Arabic linguistics*, XII, ed. Elabbas Benmamoun, 101–110. Amsterdam and Philadelphia: J. Benjamins.

Kniffka, Hannes. 1995. *Elements of culture-contrastive linguistics*. Frankfurt: P. Lang.

——. (ed.). 2001. *Indigenous grammar across cultures*. Frankfurt: P. Lang.

Lalami, Laïla. 1996. "Clitic left dislocation in Moroccan Arabic". *Perspectives on Arabic linguistics*, IX, ed. Mushira Eid and Dilworth Parkinson, 115–129. Amsterdam and Philadelphia: J. Benjamins.

Maas, Utz, El-Sayed Selmy, and Mostafa Ahmed. 2000. *Perspektiven eines typologisch orientierten Sprachvergleichs Deutsch–Arabisch/Arabisch–Deutsch*. Cairo: Echnaton.

—— and Ulrich Mehlem. 2003. "Schriftkulturelle Ressourcen und Barrieren bei marokkanischen Kindern in Deutschland". Osnabrück: Institut für Migrationsforschung (IMIS).

Mansouri, Fethi. 2000. *Grammatical markedness and information processing in the acquisition of Arabic as a second language*. Munich: Lincom.

Myers-Scotton, Carol. 1993. *Duelling languages: Grammatical structure in codeswitching*. Oxford: Clarendon.

Nortier, Jacomine. 1989. *Dutch and Moroccan Arabic in contact: Code switching among Moroccans in the Netherlands*. Ph.D. diss., University of Amsterdam.

Perdue, Clive (ed.). 1993. *Adult language acquisition: Cross-linguistic perspectives*. 2 vols. Cambridge: Cambridge University Press.

Pienemann, Manfred. 1998. *Language processing and second language development*. Amsterdam and Philadelphia: J. Benjamins.

Reuschel, Wolfgang. 1996. *Aspekt und Tempus in der Sprache des Korans*. Frankfurt: P. Lang.

Ritchie, William C. and Tej K. Bhatia. 1996. *Handbook of second language acquisition*. San Diego, Cal.: Academic Press.

Sapir, Edward. 1921. *Language*. New York. (Repr. 1949. New York: Harcourt, Brace and World.)

Talmy, Leonard. 2000. *Toward a cognitive semantics*. Cambridge, Mass.: MIT Press.

Younes, Munther A. 1990. "An integrated approach to teaching Arabic as a foreign language". *al-ʿArabiyya* 23.105–122.

Utz Maas (University of Osnabrück)

Convergence

1. Definition

The term 'convergence' is not an established term, in either theoretical linguistics in general or Semitic and Arabic linguistics in particular. Neither is the opposite term 'divergence'. However, the term *does* occur in an informal sense in studies concerning koineization phenomena (Ferguson 1959; Palva 1982), as well as pidginization and creolization phenomena of language contact (Gumperz and Wilson 1971; Kossmann 1994; Mous 1994). Lately, with Versteegh's book *Pidginization and creolization: The case of Arabic* (1984), Arabists have become aware of the relevance of this field in the realm of Arabic dialectology. → Juba Arabic (see Kaye and Tosco 2001:85–88 for an overview of the tense and aspect system) and → Maltese (Kontzi 1998b) constitute two well-known cases in point. For scholars working with both Akkadian and Ethio-Semitic, taking into consideration the effects of linguistic substrata constitutes their 'daily bread' (see Haayer 1986 for Akkadian; Leslau 1945 for Ethio-Semitic). The term 'convergence' also occurs in historical Arabic linguistics, for instance in connection with the language of the *Qurʾān* (Nöldeke 1938:137; see below). In the context of Arabic dialectology the term 'convergence' is one of the keywords in the article "Divergenz und Konvergenz im Arabischen" by Diem (1978), who explores the reasons behind common features in contiguous as well as non-contiguous Arabic dialects.

Parallel linguistic developments amounting to linguistic convergence (mainly on the level of syntax) have been investigated across Semitic. Kapeliuk (1989), for instance, investigates parallel features in varieties of Neo-Aramaic and modern Ethio-Semitic, two non-contiguous areas, which nevertheless have been subject to comparable waves of linguistic adstratum. Kropfitsch (1972) equally stresses the importance of parallel linguistic developments. Edzard (1998) constitutes an attempt to put similar observations in a broader context with a focus on the concept of linguistic convergence. Here 'convergence' is used with a view to identifying lines of common development in the following contexts: (a) among historical varieties of Arabic dialects before and during the emergence of Classical Arabic; (b) among modern Arabic dialects and other Semitic languages; and (c) among Modern Standard Arabic, modern Hebrew, and modern European languages.

2. Convergence of linguistic features among historical varieties of Arabic dialects before and during the emergence of Classical Arabic

The problematic concept of a 'proto-language' has been applied by some Semitists and Arabists to sub-groups of Semitic and even to individual Semitic languages, for instance by Owens (1998) to 'Proto-Arabic'. Being aware of the problem of defining a period when the case system of (Classical) Arabic was established (and also when it was lost), Owens (1998) opts for a bipartite system of 'Proto-Arabic', one with case and one without case. The Arabic dialects are then supposed to be descendants of the latter variety.

The beginning of what constitutes → 'Arabs' and 'Arabic' is, of course, a matter of definition, and a certain arbitrariness cannot be avoided in setting up a *terminus post quem* for (Classical) Arabic per se. Suffice it to say here that → Classical Arabic probably must be viewed as an abstracted system on the phonological and morphological levels and as a sort of *Kanzleisprache* on the syntactic level. In no way can it be viewed as a more original, let alone 'prototypical', variety of Arabic. On the contrary, the → *ʿarabiyya* may well be described as a blend of elements of the language of pre-Islamic poetry and grammatical rules inferred from features in the *Qurʾān*. All of this is just to stress that there are no reasons for assuming an earlier linguistic unity in the Arabian Peninsula. Again, it must be emphasized that as far as we can look into the past of Arabic, there is, relative to the smaller geographical territory involved, no less linguistic diversity than can be observed in more recent times.

Rabin (1951) established the major isoglosses distinguishing the Arabic of the two major dialectal groups of the Arabian Peninsula, Ḥijāz and Tamīm, which are associated with West- and East-Arabian, respectively. This dichotomy, which was upheld by the Arab grammarians, is awkward insofar as the first term designates a geographical area and the second one a (group of) tribe(s). Retsö (1989:205), in his comparative treatment of passive and causative constructions in Arabic dialects (and other Semitic languages), also vehemently rejects the idea of a Proto-Arabic language, *koine*, or otherwise, just as he rejects the whole idea of the family tree model (including wave-theoretical emendations).

A later Arabic dialectal koine is accepted, though, by many scholars. The following list constitutes a rearrangement of the features listed in Ferguson 1959 (cf. also Cohen 1970; Versteegh 1984:20–21):

i. Phonological features: *taltala* (/a/ > [i]) in verbal prefixes; pharyngealization of *t* in numbers 13–19; simplification of the *nisba*-ending (*-iyy(un)* > *ī*); phonological merger of /ḍ/ and /ḏ̣/; despirantization of interdental fricatives; loss of *hamza* ('glottal stop'), except for the case where /q/ > [ʔ]; reduction of vowels in short open syllables; phonological merger of short /u/ and /i/; merger of the feminine endings *-a(tun)*, *-ā* and *-āʾ(u)* into *-a*.
ii. Morphological features: loss of the dual in pronouns and verbs; merger of verbs III*w* and III*y*; treatment of geminated verbs (in several derived forms) in analogy to Form II of verbs III*w*/*y*; disappearance of the feminine elative pattern $C_1uC_2C_3\bar{a}$; loss of the *h* in the object suffix *-hu* (3rd pers. masc. sg.) after a consonant; simplification and morphological merger of the $C_1aC_2\bar{a}C_3iC_4(u)$ and $C_1aC_2\bar{a}C_3\bar{\imath}C_4(u)$ plural patterns to

$C_1C_2\bar{a}C_3iC_4$; change of the diminutive pattern $C_1uC_2ayC_3$(un) to $C_1uC_2ayyaC_3$; morphological merger of the {a-i-a} and {a-u-a} verbal patterns; reanalysis of biradical roots as triradical; loss of Form IV (ʾaC_1C_2aC_3a); loss of gender distinction in the plural of pronouns and verbs.

iii. Syntactic features: government of direct objects by the preposition *li-*; morpho-syntactic coalescence of the preposition *bi-* with the verb *jāʾa* 'to come', yielding the new verb *jāb* 'to bring'; loss of gender polarity in number syntax (numbers 13–19); replacement of the series of relative pronouns by the single indeclinable relative pronoun *illī*; emergence of a number of modal verb prefixes (e.g. 'future', 'iterative') such as *b(ayn)-* and *ḥa-*; emergence of a number of → 'analytical' genitive exponents in possessive genitive constructions, replacing the construct case; number agreement between subject and verbal predicate (even with non-human subjects in the plural, which in Classical Arabic take feminine singular agreement); the tendency to shift from VSO to SVO word order; emergence of a number of → serial auxiliary verbs (e.g. 'ingressive') such as *qaʿada*, *qāma*; use of asyndetic modal constructions (e.g. 'necessity') such as *lāzim*, *bədd*.

iv. Lexical features: use of *šāf* instead of *raʾā* 'to see'; nominal periphrasis of interrogative adverbs.

As to these features of the Arabic → dialect koine, Miller (1986) has convincingly shown that they need not, indeed should not, be attributed to a common origin. Miller (1986:56) observes further: "The failure of comparative reconstruction to clarify the origin of the modern Arabic dialects is greatly due to the mobile history of the Arabs". Here, one can go one step further in arguing that such a reconstruction is not only unfeasible and unnecessary (cf. Bloch 1992), but is in fact also undesirable. Rather, the linguistic array of data is far more compatible with the theoretical possibility of a simultaneous emergence of linguistic features in Arabic dialects.

The distribution of 'indefinite markers' *m/n* (→ 'mimation'/→ 'nunation') and 'definite markers' furnishes a good example in this context. With respect to the different forms of the definite articles in pre-Islamic Arabia (*h(n)-*, *(ʾ)l-*, *-n*, and *am-*), Beeston (1981:185–186) provides an analysis in terms of convergence:

> My tentative suggestion is that we should distinguish (a) ancient north-west Arabian, with article *h(n)-*; (b) ancient north-east (?) Arabian, with article *(ʾ)l*; (c) ancient south-west Arabian, split into two branches, the Sayhadic type with article *-n* and the Himyaritic type with article *am-*; (d) ancient west-central Arabian of an indeterminate character constituting a mosaic of north-west, south-west (Himyaritic), and perhaps also some north-eastern, speech forms. In the course of time (though the chronology is impossible to determine), the Sayhadic form has disappeared completely as regards its individual morphological features, though its lexicon has continued to exercise a strong influence; and the remaining speech forms have converged so as to produce the amalgam of dialects which can properly be called Arabic. One of the effects of this convergence has been total elimination of the *h(n)* article in favour of the *(ʾ)l* form, and the present-day restriction of the *am*-article to a few isolated pockets in Yemen.

A further example of linguistic convergence in the history of Arabic is the distribution of the prefix-vowels (/a/ vs. /i/) in the Arabic prefix-conjugation. In Classical Arabic, the preformative vowel is /a/ (*ya-ktubu* 'he writes'); in the dialects it is /i/ (*yi-ktib*, etc.). This puzzling phenomenon, described by the Arab grammarians under the name of → *taltala*, was adduced by Ferguson (1959) in support of his 'Arabic koine' theory. *Contra* Ferguson, Bloch (1967) has shown that the /ī/-preformatives in the modern dialects need not be traced back to an original koine. Rather, an explanation in terms of Barth's Law can be envisaged. Barth's Law (Barth 1894) is the change /a/ > [i] in verbal prefixes when another *a* follows as thematic vowel, e.g. /yaqtal/ > [yiqtal]. If we posit that Barth's Law was in effect in pre-Classical Arabic, then early Arabic would have shown a variety of realizations of the imperfect preformative vowel: /a/ in some environments, /i/ in others. Classical Arabic would then have standardized one option, and the dialects the other. This unpredictable selection of a single option out of an earlier multiplicity of options (and the later distribution of this feature in Arabic) is precisely the scenario of linguistic convergence. We thus do not need to invoke the 'spreading' of an /ī/-vowel, as in the koine-scenario. Rather, the variability was always there as such, as far back as we can look into early Semitic.

The term 'convergence' has also been applied in the context of the language of the *Qurʾān*, or rather the language policies regarding the text of the *Qurʾān*. Nöldeke (1938:137) summarizes the tendencies and, indeed, policies of convergence that were instrumental in creating a *textus receptus* of the *Qurʾān*:

> Die Eliminierung der vom othmanischen *muṣḥaf* abweichenden Varianten und der ohne Rücksicht auf die Tradition frei konstruierten Lesarten ist nur ein Teil des großen Prozesses der Vereinheitlichung von Korantext und Koranlesung, der Schaffung eines *textus receptus*; wirkender Faktor dieses Prozesses ist das Majoritätsprinzip oder allgemeiner die katholische Tendenz, die Konvergenz in der islamischen Entwicklung.
>
> [The elimination of variants that deviate from the ʿUṯmānic *muṣḥaf*, and readings that were construed arbitrarily with no reference to the tradition, is only part of the large-scale process of standardization of the Qurʾānic text and the Qurʾānic recitation, the creation of a *textus receptus*. The driving force behind this process is the 'majority rules' principle, or more generally the 'catholic' tendency, the convergence in the development of Islam.]

Nöldeke's subsequent point is well taken: that even the full list of non-canonical *qirāʾāt* is far more restricted than the breadth of actual language reality described in the contemporary grammatical literature, a circumstance that further underscores the strong forces of linguistic convergence operating on the Qurʾānic text. But even the *textus receptus* contains many features associated with an early *Volkssprache*, as Vollers (1906) and others have cogently shown. An example is the form *yahiddī* (corresponding to standard *yahtadī*) 'he is guided' (Q. 10/35). Similar forms are treated in the chapters 565–571 on → *ʾidġām* 'assimilation' toward the end of Sībawayhi's *Kitāb* (IV, 443–444).

3. CONVERGENCE OF LINGUISTIC FEATURES AMONG MODERN ARABIC DIALECTS AND OTHER SEMITIC LANGUAGES

Diem (1978:129) describes the dynamics of convergence of linguistic features among Arabic dialects in a 'set-theoretical' way. Dialects are defined as carrying a specific selection from a bundle of linguistic features that spread all over various Arabic dialects:

> Dem Dialektologen drängt sich beim Studium arabischer Dialekte immer wieder der Eindruck auf, als ob sich die große Menge arabischer Dialekte voneinander nicht so sehr durch exklusive, ihnen jeweils allein eigene Merkmale abhöben – wenn diese auch durchaus vorhanden sind – als durch die verschiedene, und wie es manchmal fast scheinen mag, vom Zufall bestimmte Auswahl einer begrenzten Zahl von Möglichkeiten, die in verschiedenen Kombinationen auftreten.
>
> [In studying Arabic dialects, the dialectologist is again and again struck by the following impression: The vast array of Arabic dialects are differentiated not so much through exclusive features specific only to the given dialect – although such features do exist – as through different, sometimes apparently even arbitrarily determined, choices from a limited number of options, which appear in different combinations.]

As examples, Diem (1978:129) cites features typically associated with Maghreb dialects which are, however, also attested in other regions, such as the opposition *nəktəb* vs. *nəktəbu* 'I/we write', which is also found in Upper Egypt or the *itfaʿal* passive to the basic form (*Grundstamm*) in the Cairene dialect.

It is clear that linguistic convergence is favored in areas characterized by nomadic lifestyle as opposed to isolated pockets in mountainous areas, for instance, in Lebanon and South Arabia (cf. Diem 1978:132). However, linguistic convergence has also been observed between local dialects of different religious affiliations in one and the same region, e.g. between the Jewish and Christian varieties of Baghdadi Arabic on the one hand and the Muslim variety on the other (cf. Blanc 1964: 164–165; → communal varieties). In modern times, the formation of Arab states may have engendered dialectal convergence in that the dialect of a capital affects other dialects in the same country by virtue of its prestige function, thus justifying terms like 'Algerian Arabic' or 'Syrian Arabic' (cf. Versteegh 2001:140). The same point may hold internationally for dialects with prestige function in the cinema industry.

Parallel innovations in Arabic dialects and other Semitic languages, ancient or recent, have been explained both in the light of older Semitic languages (Kropfitsch 1972; Morag 1989; Rendsburg 1991) and in the light of modern Semitic languages, mainly varieties of Neo-Aramaic (Blau 1966–1967; Diem 1978). Kropfitsch (1972:18– 28) lists the following areas of parallel development: (a) dissolution (expansion) of the three-vowel inventory (/a-i-u/), monophthongization, reduction of the glottal stop, reduction of interdental fricatives, and the emergence of /i/ as vowel of the prefix conjugation (*taltala* in the native Arabic terminology) on the phonological level; (b) loss of case endings, far-reaching loss of the dual in adjectives, pronouns, and verbs, dominance of the oblique case in the dual and sound masculine plural, replacement of the internal (vocalic) passive by derived forms with pre- and infixes, far-reaching loss of the *nūn paragogicum*, and merger of verbs III*w/y/ʾ* (*yāʾ/yōḏ*, *wāw*, and *ʾalif/ʾalef*) on the morphological level; dissolution of the strict SVO word order, emergence of genitive exponents replacing the traditional annexation, and loss of concord within noun-adjective phrases in the dual on the syntactic level.

An example of parallel syntactic innovation is the 'resumptive' use of anaphoric and cataphoric pronouns (→ cataphora). At issue here are constructions of the type seen in Hebrew and Aramaic (*bēṯō šel Dāwīḏ* 'house-his of David' = 'David's house'), which have parallels in many languages, both Semitic and non-Semitic. Syriac, for instance, also exhibits – in addition to the type just mentioned – a cataphoric pronoun in constructions such as *qṭalt-eh l-malkā* 'I:killed-him to-king' = 'I killed the king'. Rendsburg (1991:1270–1271) adduces comparative evidence from Mishnaic Hebrew, modern Arabic dialects, and modern South Arabian languages:

Syriac Aramaic	
qaṭl-eh l-malkā	'he killed-[him] the king'
bayt-eh d-malkā	'the house-[his] of the king'
Gəʿəz	
qatal-o la-nəgus	'he killed-[him] the king'
bet-u la-nəgus	'the house-[his] of the king'
Mishnaic	
ʾāmĕrū ʿala-w ʿal rabbī ḥănīnāʾ	'they said [of him] of Rabbi Hanina'
Iraqi Arabic	
fallšu-ha li-l-madrasa	'they demolished [it] the school'
Jibbali (Modern South Arabian)	
he-s le-ʾem-í	'for [her] my mother'

Another example of parallel syntactic innovation is the emergence of definite noun-adjective phrases of the type Arabic *dār al-bayḍāʾ* 'Casablanca' (lit. 'house the white') or Hebrew *knesseṯ hag-gḏōlā* (lit. 'synagogue the great'), in which the adjective is marked as definite but the noun has 'lost' its definite article. Rendsburg (1991:1268–1269) provides, among others, the following examples:

Syrian (Damascene) Arabic	
bāb əš-šarʾi	'the Eastern gate'
Eastern Neo-Aramaic	
qalpa aw xwāra	'the white hull'
Mehri	
askíin hayd ənóot	'the new knives'
Amharic	
təlləqu bet	'the big house'

4. CONVERGENCE OF LINGUISTIC FEATURES AMONG MODERN STANDARD ARABIC, MODERN HEBREW, AND MODERN EUROPEAN LANGUAGES

The term 'Standard Average European' (SAE), possibly coined by Benjamin Whorf, implies the convergence of linguistic features, mainly on the levels of syntax and the lexicon. Both Blanc (1957:401–402) and Blau (1981:60–141) use this term in referring to common morpho-syntactic and lexical innovations in Modern Standard Arabic and Modern Israeli Hebrew. With respect to the proliferation of Latinate and Greek prefixes as well as prefixes of inner-Semitic origin occurring in modern Hebrew (e.g. *pro-ʾarvi* 'pro-Arabic', *qonter-mahpekhani* 'counter-revolutionary', *anti-mitsri* 'anti-Egyptian', *al-xuti* 'wireless', *xad-tsedadi* 'unilateral') Blanc (1957) formulates:

> The necessity of translating terms from Standard Average European (SAE), ha[s] resulted in the introduction of prefixes, a type of morpheme virtually unknown to Semitic languages and for which there is but the barest precedent in earlier Hebrew; these have been adapted from, or invented on the base of, existing Hebrew and Aramaic particles or words, or lifted bodily from SAE, and today form an extremely important and productive part of the language. Most prefixes are so productive that they can be added, as the need arises, to almost any noun or adjective.

In Modern Standard Arabic such adjectives with prefixes, even though etymologically of Arabic and not of European origin (e.g. *janūb-ʾifrīqī* 'South African', *šarq-ʾawsaṭī* 'Middle Eastern', *ʾafrū-ʾāsiyawī* 'Afro-Asiatic', *šibh-rasmī* 'semi-official', *qab-tārīxī* 'pre-historic', *bar-māʾī* 'amphibian'), are equally in use (see, for instance, Badawi, Carter, and Gully 2004: 751–762 for these and other → compound formations). A typical example of lexical convergence is the following (Blau 1981:62–63): both the Arabic and Hebrew words for 'electricity', *kahrabāʾ* and *xašmal* derive from an ancient word with the meaning 'amber'. The Greek term *ēlektron* 'amber' had also adopted the sense 'electricity'. Accordingly, the Arabic term *kahrabāʾ* (ultimately of Persian origin) was first used in the sense of 'electricity' by Muḥammad ʿAlī's chief translator Rifāʿa Rāfiʿ aṭ-Ṭahṭāwī. The Hebrew term was introduced by the poet Yehuda Leib Gordon with reference to the translation of the Biblical Hebrew term *xašmal* as *ēlektron* and *electrum* in the Septuagint and the Vulgate respectively. Further examples, mainly due to English and French influence, are countless. In the sense of such developments one could well discuss the validity of the concept of a 'Standard Average Semitic' in analogy to Blanc's and Blau's 'Standard Average European'.

5. Conclusions

In conjunction with its opposite 'divergence', 'convergence' has proved to be an important concept in linguistic evolution, both within one and the same language family (here Semitic) and across language families (here Semitic and branches of Indo-European). The scenarios drawn by Diem (1978), Versteegh (1984), and others confirm the current *opinio communis* that the array of historical and modern Arabic dialects cannot be directly derived from Classical Arabic, let alone from some kind of 'Proto-Arabic'.

Bibliographical references

Primary source

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. Ed. Muḥammad ʿAbd as-Salām Hārūn. 5 vols. Cairo, 1968–1977; 2nd ed. Cairo, 1977.

Secondary sources

Badawi, El-Said, Michael Carter, and Adrian Gully. 2004. *Modern written Arabic: A comprehensive grammar*. London: Routledge.

Bakker, Peter and Maarten Mous (eds.). *Mixed languages: 15 case studies in language intertwining*, Amsterdam and Dordrecht: IFOTT/Foris.

Barth, Jacob. 1894. "Zu den Vokalen der Imperfect-Präfixe". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 48.4–6.

Beeston, A.F.L. 1981. "Languages in pre-Islamic Arabia". *Arabica* 28.178–186.

Blanc, Haim. 1957. "Hebrew in Israel: Trends and problems". *Middle East Journal* 11.401–402.

——. 1964. *Communal dialects in Baghdad*. Cambridge, Mass.: Center for Middle Eastern Studies Harvard University.

Blau, Joshua. 1966–1967. *A grammar of Christian Arabic based mainly on South-Palestinian texts from the first millennium*. Louvain: Secrétariat du Corpus SCO.

——. 1981. *The renaissance of modern Hebrew and modern standard Arabic: Parallels and differences in the revival of two Semitic languages*. Berkeley and Los Angeles: University of California University Press.

Bloch, Ariel. 1967. "The vowels of the imperfect preformatives in the old dialects of Arabic". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 117.22–29.

——. 1992. "Different ways, identical results: On parallel retention and parallel innovation in modern Arabic dialects". *Jerusalem Studies in Arabic and Islam* 15.255–269.

Cohen, David. 1970. "Koiné, langues communes et dialectes arabes". *Etudes de linguistique sémitique et arabe*, ed. David Cohen, 105–125. The Hague: Mouton. (First published *Arabica* 9.119–144.)

Diem, Werner. 1978. "Divergenz und Konvergenz im Arabischen". *Arabica* 25.128–147.

——. 1979. "Studien zur Frage des Substrats im Arabischen". *Der Islam* 56.12–80.

Edzard, Lutz. 1998. *Polygenesis, convergence, and entropy: An alternative model of linguistic evolution applied to Semitic linguistics*. Wiesbaden: O. Harrassowitz.

Ferguson, Charles. 1959. "The Arabic koine". *Language* 35.616–630.

Gumperz, John and Robert Wilson. 1971. "Convergence and creolization: A case from the Indo-Aryan/Dravidian border in India". *Pidginization and creolization of languages*, ed. Dell Hymes, 151–167. Cambridge: Cambridge University Press.

Haayer, Geerd. 1986. "Languages in contact: The case of Akkadian and Sumerian". *Scripta signa vocis*, ed. Herman Vanstiphout a.o., 77–84. Groningen: E. Forsten.

Holtus, Günter, Michael Metzeltin, and Christian Schmitt (eds.). *Lexikon der romanistischen Linguistik*, VII. Tübingen: Niemeyer.

Kapeliuk, Olga. 1989. "Some common traits in the evolution of Neo-Syriac and of Neo-Ethiopian". *Jerusalem Studies in Arabic and Islam* 12.294–320.

Kaye, Alan L. and Mauro Tosco. 2001. *Pidgin and creole languages: A basic introduction*. Munich: Lincom Europa.

Kontzi, Reinhold. 1998a. "Arabisch und Romanisch". Holtus a.o. (1998:328–347).

——. 1998b. "Maltesisch". Holtus a.o. (1998:347–365).

Kossmann, Maarten. 1994. "Amarna-Akkadian as a mixed language". Bakker and Mous (1994:169–173).

Kropfitsch, Lorenz. 1972. *Einige ausgewählte parallele Entwicklungen in den altsemitischen Sprachen* *und in den neuarabischen Dialekten*. Ph.D. diss., University of Graz.

Leslau, Wolf. 1945. “The influence of Cushitic on the Semitic languages of Ethiopia: A problem of substratum”. *Word* 1.59–82.

Miller, Ann. 1986. “The origin of the modern Arabic sedentary dialects: An evaluation of several theories”. *al-ʿArabiyya* 19.47–74.

Morag, Shlomo. 1989. “Biblical Hebrew and modern Arabic dialects: Some parallel lines of development”. *Jerusalem Studies in Arabic and Islam* 12.94–117.

Mous, Maarten. 1994. “Ma’a or Mbugu”. Bakker and Mous (1994:175–200).

Nöldeke, Theodor. 1938. *Geschichte des Qorans*, III. 2nd ed., expanded and elaborated by Gotthelf Bergsträsser and Otto Pretzl. Leipzig: Dieterich. (Repr. Hildesheim: G. Olms, 1961)

Owens, Jonathan. 1998. “Case and proto-Arabic”. *Bulletin of the School of Oriental and African Studies* 61.51–73.

Palva, Heikki. 1982. “Patterns of koineization in modern colloquial Arabic”. *Acta Orientalia* 43.13–32.

Rabin, Chaim. 1951. *Ancient West-Arabian*. London: Taylor’s Foreign Press.

Rendsburg, Gary. 1991. “Parallel developments in Mishnaic Hebrew, colloquial Arabic, and other varieties of spoken Semitic”. *Semitic studies in honor of Wolf Leslau on the occasion of his eighty-fifth birthday*, ed. Alan Kaye, II, 1265–1277. Wiesbaden: O. Harrassowitz.

Retsö, Jan. 1989. *Diathesis in the Semitic languages*. Leiden: E.J. Brill.

Versteegh, Kees. 1984. *Pidginization and creolization: The case of Arabic*. Amsterdam and Philadelphia: J. Benjamins.

——. 2001. *The Arabic language*. 2nd ed. Edinburgh: Edinburgh University Press.

Vollers, Karl. 1906. *Volkssprache und Schriftsprache im alten Arabien*. Strassburg: K.J. Trübner.

LUTZ EDZARD
(University of Oslo)

Coordination → Parataxis

Coptic

Table 1. Transliteration of Coptic alphabetic signs

ⲁ *a*	ⲏ *ê*	ⲛ *n*	ⲧ *t*	ϣ *š*
ⲃ *b*	ⲑ *t*^{h}	ⲝ *ks*	(Ⲟ)ⲩ *(o)u, y*	ϥ *f*
ⲅ *g*	(ⲉ)ⲓ *(e)i*	ⲟ *o*	ⲫ *p*^{h}	ϩ *h*
ⲇ *d*	ⲕ *k*	ⲡ *p*	ⲭ *k*^{h}	ϧ *ḫ*
ⲉ *e*	ⲗ *l*	ⲣ *r*	ⲯ *ps*	ϫ *č*
ⲍ *z*	ⲙ *m*	ⲥ *s*	ⲱ *ô*	ϭ *c*

1. HISTORICAL SETTING AND SOURCES

The encounter of Egyptian natives with the Arabs and Arabic in 641 C.E. was the starting point of cultural as well as linguistic developments resulting in the loss of the Egyptian (Coptic) language and the shift of Christian Egyptians to the use of Arabic before ca. 1300 C.E. The socio-linguistic and chronological details of this process are largely unknown and still debated (see Björnesjö 1996; Décobert 1992; Helderman 1997; MacCoull 1985, 1989; Richter 2001; Rubenson 1996). The phenomenon under discussion is based on evidence of a dead written language (on the relevant methodological issues, see Adams a.o. 2002). Language contact phenomena, however plausible in the spoken language, are scarcely visible in the written texts, if at all. There are only two kinds of linguistic interference phenomena that can be observed through the mirror of written evidence: ‘hybrid’ combinations of linguistic and graphical codes not matching each other (‘Karshuni’) and lexical code-switching (borrowing).

Parts of a large 13th-century manuscript written in the Arabic language, but with Coptic signs, have been published (Blau 1988; Burmester 1965–1966; Casanova 1901; Satzinger 1972, 1991; Sobhy 1926; Worrell 1934:134–143). Its Coptic spellings of Arabic words, although based on the phonological system of Bohairic (Lower Egyptian) Coptic, are of some value for comparison with the mostly Fayyumic (Middle Egyptian) and Sahidic (Upper Egyptian) spellings of loanwords (see section 2 below). The same phenomenon of Arabic texts spelled in Coptic letters is further attested in a bilingual curse (Crum 1902a; Blau 1988:189). The reverse phenomenon – the use of Arabic characters to transcribe a Coptic (Bohairic) text – is attested in a manuscript of hymns in honor of the holy virgin (Blau 1988; Galtier 1906; Satzinger 1972, 1991; Worrell 1934:134–143).

→ Coptic loanwords (and some cases of phonological and structural borrowing) in Egyptian Arabic have been dealt with several times over the last century (Praetorius 1901, 1902; Galtier 1902; Sobhy 1950; Bishai 1960, 1961, 1962, 1964; Ishaq 1975, 1991; Behnstedt 1981, 1997; Vittmann 1991; Peust 1999:321–323; Schenkel 2002), although the estimated quantity of these borrowings varies greatly. In contrast, the occurrence of Arabic loanwords in Coptic texts thus far remains almost entirely unresearched (cf. Stern 1885:117–119; von Lemm 1903:xvii.34–36; Chassinat 1921:21–47; Worrell 1934:122–133; Vycichl 1991). In particular, there exists no modern reference tool listing either Arabic loanwords in Coptic or Coptic texts containing them. Both Worrell (1934:122–133) in his phonological study of Coptic and Vycichl (1991) in his brief entry on Copto-Arabic vocabulary dealt with only two Coptic texts. At present, we know more than 90 published Coptic texts and a few unpublished ones which contain Arabic loanwords. It is a very striking fact that almost all of these texts belong to two genres:

i. scientific texts, 18 manuscripts from the 9th to 11th century, among them four large alchemistic treatises (Stern 1885 and the hitherto unpublished Bodleian manuscripts (P)a 1, 2, and 3), a manual providing arithmetical and metrological problems (Drescher 1948–1949), a page dealing with astronomy (Bouriant 1904; von Lemm 1903:35–36), and a few collections of magical (in particular Chassinat 1955) and medical (in particular Chassinat 1921) receipts.
ii. about 80 documentary texts from the 8th to the 12th centuries, including legal documents (Richter 2001, 2003), letters, lists, and accounts, in particular a large 11th-century account book recording income and expenditure of a middle Egyptian monastery (British Library Or. 13885, unpublished).

These texts comprise altogether about 400 transcriptions of Arabic words, which will be called ‘loanwords’ here, regardless of whether they might have been well-established parts of Coptic speech or rather, as is more likely in many cases (see section 5 below), ‘one-time’ or ‘nonce’ borrowings (Field 2002:9–10; van Hout and Muysken 1994:40). The great bulk are nouns; only 19 verbs (about 5 percent) have been identified thus far, and one single function morpheme may be attested. It should be borne in mind, however, that this evidence is not certain, but depends on further philological study of Coptic texts. Linguistic conclusions are necessarily only provisional. Almost all Coptic transcriptions of Arabic words occur in Coptic texts written in a relatively informal orthography and with a relatively low degree of linguistic perscriptivism, closer to the vernacular than any literary composition. The great bulk represents the Sahidic (Upper Egyptian) dialect; a few are Fayyumic (Middle Egyptian) or Bohairic (Lower Egyptian). The linguistic decorum appropriate to the vast number of Coptic (semi-)literary texts demanded the suppression of phonetic, grammatical and, in particular, lexical innovations, so that borrowings from Arabic do not occur in them at all. There are only a few exceptions to this rule, e.g. a magical spell invoking the *roh n-alla* (< *rūḥ allāh*) ‘spirit of god’ (Beltz 1983: 63), a 13th-century hagiographical text (Amélineau 1887), and a 13th-century scribal colophon (Crum 1905, nº 726).

The Arabic underlying these Coptic transcriptions was roughly identified by Worrell (1934: 123) as spoken (or heard) colloquial, rather than Classical Arabic. There are strong affinities between the phonemic correspondences prevailing in the loanwords and those attested in the homogeneous transcription system of a Copto-Arabic Karshuni text (see above). The language of the latter was studied by Satzinger (1972) and Blau (1988), who described it as follows: “Like Middle Arabic texts in general”, this one too “is characterized by freely alternating features of Classical Arabic, Neo-Arabic and pseudo-corrections” (Blau 1988:145).

2. COPTIC SPELLINGS OF ARABIC PHONEMES AND RELATED ISSUES

Some Arabic phonemes have close equivalents in Coptic (e.g. the sonorants *l*, *m*, *n*, *r*); hence the same graphemic correspondences always occur. More commonly, different ways to transcribe a single consonant phoneme are attested even in the same position (note, e.g., the three variants of *f* in *ʿalaf* ‘fodder’: *alêf*, *alêb*, *alêou*). But usually one of these varieties proves to be the most common, regular one.

i. Consonants: ‘ = mostly ø, seldom *a* (*alaasaat* < *al-’asad*), perhaps consonantal gemination (*ammour* < *a’mur*?), perhaps *h* ($k^{h}it^{h}irh$ < *kaṯīrā’*?); *b* = *p* (cf. Hintze 1947b); *t* = *d* or *t*, seldom t^{h} ($alk^{h}iprit^{h}$ < *al-kibrīt*); *ṯ* = t^{h}, seldom *t* (*almatkal* < *al-miṯqāl*); *j* = Sahidic *c* (a palatalized velar), Bohairic *č* (palatal, regularly corresponding with Sahidic *c*); *ḥ* = *h*, seldom ø (*kol* < *kuḥl*); *x* = mostly k^h, in Bohairic *ẖ* (*almairêẖ* < *al-mirrīx*), seldom *š* (*assarnêš* < *az-zirnīx*), perhaps *h* (*arrôham* < *ar-ruxām*?), perhaps ø (*aulen* < *xawlān*?), in Bohairic perhaps once *k* (*allinek* < *al-līnax*?); *d* = *d* or *t*; *ḏ* not attested; *r* = *r*, seldom *l*; *z* = *s*, seldom *z* (*gazouan* < *ġazawān*); *s* = *s*; *š* = *š*; *ṣ* = *s*; *ḍ* = *t* (*apiat* < *ʾabyaḍ*); *ṭ* = *t*, seldom *d* (*hôdôt* < *ḥuṭuṭ*); *ḏ̣* = *t* (*attaheri* < *aḏ̣-ḏ̣āhirī*), perhaps *s* (*naser* < *nāḏ̣ir*?); *ʿ* = mostly ø, sometimes *a* (*alaakrap* < *al-ʿaqrab*, *assalae* < *az-zalʿa*, *arrôpa* < *ar-rubʿ*), seldom *e* (*alceme* < *al-jamʿ*), *(e)i* (*assiri* < *az-zarʿ*) or *ô* (*arrapô* < *ar-rubʿ*) and even *k* (*almaksoul* < *al-maʿsūl*), perhaps consonantic gemination (*alcelle* < *al-jaʿāla*?); *ġ* = *g* (*almoulgam* < *al-malġam*), perhaps *c* (*alcabiri* < *al-ġafīr*?); *f* = *b* (cf. Hintze 1947b), sometimes *f*, seldom *ou* (cf. *alêf*, *alêb*, *alêou* < *ʿalaf*), once (Bohairic) *p* (*espêiteč* < *ʾisfīdāj*); *q* = *k*; *k* = k^h, sometimes *k* (*alkous* < *al-kūz*); *l* = *l*, seldom *r*; *m* = *m*; *n* = *n*, but in contact with labials, Coptic assimilation (*n* becomes *m*) occurs (*assampak* < *azzanbaq*, *assoumpoule* < *as-sunbula*); *w* = *ou*, once (Bohairic) *b* (*iban* <*ʾīwān*); *h* = *h* (*assoouhre* < *az-zuhara*), as feminine ending ø, once *h* (*šetineh* < *šādina*); *y* = *(e)i*. Arabic consonantal gemination is sometimes written (*almousabbi* < *al-muṣaffī*, *almoucarrap* < *al-mujarrab*, *alk*^{h}*ammoun* < *al-kammūn*, *alhôcce* < *al-ḥujja*, *asouk*^{h}*k*^{h}*ar* < *as-sukkar*, *atassa* < *aṭ-ṭāssa*, *aššoukke* < *aš-šuqqa*, *ette* < *ʿidda*), sometimes not (*alcoume* < *al-jummāʿ*, *morape* < *murabbaʿ*, *oušak* < *wuššaq*, *rôman* < *rummān*, *almairêẖ* < *al-mirrīx*), sometimes either way (*alcoup(p)e* < *al-jubba*). Gemination is never spelled in final position (*alhal* < *al-xall*, *alk*^{h}*as* < *al-xazz*, *alhat* < *al-ḥadd*, *almalaf* < *al-milaff*, *armôr* < *al-murr*, *arôs* < *ar-ruzz*). In a few cases it seems to be transcribed as a vocalic ablaut, cf. *aššipe* < *aš-šabb*, *kere* < *qarr*, and *lepe* < *labb*; cf. the proper name *Apoulase* < *ʾAbū Lazz*.

ii. Vowels: *ā* regularly occurs as *a* or *e*, sometimes as *ê*, seldom as *ee* (*alpeep* < *al-bāb*), *i* (*alkili* < *al-qily*), or *ôe* (*almôes* < *al-mās*). There is thus strong evidence of *ʾimāla* (as in Copto-Arabic Karshuni, cf. Blau 1988:152). *ū* occurs as *o*, *oo*, *ou*, *ô*, *ôô*; *ī* occurs as *i* or *ê*. The feminine ending (*tāʾ marbūṭa*) is almost always spelled *e* (again clear evidence of *ʾimāla*), sometimes *a* (*a* and *e* also in Copto-Arabic Karshuni, see Blau 1988:176), seldom *ai* (*almešmelai* < *al-mišmala*), *i* (*almanari* < *al-manāra*), ø (*alpourat* < *al-burāda*), once *eh* (*šetineh* < *šādina*). Other short vowels must be left out of consideration, since both their quality in colloquial Arabic and their Coptic transcriptions show a great deal of variation, so that correspondences remain unclear.

3. MORPHOLOGY OF ARABIC WORDS IN COPTIC

3.1 *Nouns*

As a rule, Arabic nouns taken into Coptic are borrowed in a form beginning with *al-*, less often spelled *ar-*, *el-*, or *er-*. Before the *ḥurūf šamsiyya*, assimilation usually occurs: *an-n . . .*, *ar-r . . .*, *as-s . . .*, *aš-š . . .*, *at-t . . .*, although often spelled haplographically with no gemination: *an . . .*, *ar . . .*, *as . . .*, *aš . . .*, *at. . . .* As in Spanish, this Arabic article does not function as a determiner. Every borrowed Arabic noun, whether prefixed with *al-* or not, was subject to the elaborate Coptic determination system (cf. Layton 2000:35–53), distinguishing, e.g.,

i. definite articles: masc. sg. *p-*, e.g. *taau ehoun e-p-alhal* 'add them to the (*p-*) vinegar (*al-xall*)'; fem. sg. *t-*, e.g. *ci . . . n-t-alpourate* 'take the (*t-*) filings (*al-burāda*)'; and pl. *n-*, e.g. *etbe n-alhecos euhiptôou* 'because of the (*n-*) barriers (*al-ḥajz*) which are on the mountains' (Crum 1902b:no. 290);
ii. indefinite articles: sg. *ou-*, e.g. *eišouei mmof hn-ou-alkous* 'heat it up in a (*ou-*) jug (*al-kūz*)' (Bodl. ms (P)a2, 26) and pl. *hen-*, not attested;
iii. zero-article: ø – as in *ou-alkapele n-at-ø-almisahe* 'a tenancy (*al-qabāla*) without survey (*al-misāḥa*)' (Richter 2003).

These determiners were applied to Arabic nouns according to both the semantic and the syntactic demands of the Coptic language. The use of Arabic nouns without the article *al-* is far less frequently attested. However, this is the standard in a large medical manuscript (Chassinat 1921), and it often occurs in rather early cases of borrowing, as can be shown by the word *(al)para* (< *barāʾa*) 'receipt' which is spelled without *al-* in a number of 8th-century documents (e.g. Kahle 1954:no. 291,5.29, *t-* or *p-para*), while in later (9th- and 10th-century) documents it is always written *t-alpara* (e.g. Crum 1902b:no. 377,9; Crum 1939:no. 49,11.13).

Unlike Arabic nouns, which are subdivided into unmarked masculine vs. marked feminine forms, Coptic nouns have an associated (inherent) gender, which is expressed not by special forms but by masculine vs. feminine determination morphemes (Layton 2000:85–86). In most instances, the choice of a Coptic article matches the respective grammatical gender of the Arabic noun. There are only a few cases of discrepancy between the gender of the noun in Arabic and of the article used in Coptic. The gender of borrowed nouns is often influenced by target language nouns of similar meaning. This may be the case with *p-at*^{h}*askieie* (< *at-tazqiya*) 'purification' (Bodl. ms (P)a1, g11; masculine Coptic equivalent *tbbo*), *p-para* (< *barāʾa*) 'receipt' (Kahle 1954:no. 291,29; masculine Graeco-Coptic equivalent *entagion*), *t-almiret* (< *al-mīrāṯ*) 'heritage' (Richter 2001:80; feminine Graeco-Coptic equivalent *klêronomia*), or *t-almisan* (< *al-mīzān*) 'scales' (Bouriant 1904; feminine Coptic equivalent *maše*). Some words are treated as either masculine or feminine, e.g. *p-* or *t- (al)para* (< *barāʾa*) 'receipt', *p-* or *t- alpourate* (< *al-burāda*) 'filings'.

3.2 *Verbs*

Almost all verbs borrowed from Arabic come from alchemistic treatises, where not only concrete objects like ingredients, utensils, etc., but also certain procedures are designated by technical terms. However, unlike nouns with their common *al-*'prefix' clearly pointing to Arabic etymology, it is not always so easy to make a decision on whether a Coptic-written verbal lexeme comes from Arabic or not, the more so as the morphological richness of the Arabic verb, with its breakdown into stems, conjugations, and verbal nouns, can complicate the identification. Coptic verbal syntax requires only two verbal forms, both operating without inflexion. Verbs borrowed from Greek into Coptic are even restricted to a single basic form: they occur in a non-Classical (Greek) infinitive form and operate within Coptic syntax as (Coptic) infinitives. Similarly, Arabic verbal forms seem to be used in Coptic sometimes in their infinitive forms (see examples [1]–[5] below), although the difficulty of determining vowel qualities (see section 2 above) leaves some uncertainty:

strong verb, Form IV
(1) *akêt* (< *ʿaqada* IV 'to boil down, to thicken') – infinitive: *ʾiʿqād* (cf. imperative *ʾaʿqid*)
(2) *elhêf* (*laḥafa* IV 'to cover') – infinitive: *ʾilḥāf* (cf. imperative *ʾalḥif*)

geminated verbs
(3) *kera*, *kere* (< *qarra* 'to be cold, to be cool') – infinitive: *qarr*
(4) *lepe* (< *labba* 'to stay') – infinitive: *labb*

III*w*
(5): *gazouan* (< *ġazā* 'to conquer, to capture') – infinitive *ġazawān*

In other cases, however (see [6]–[12]), forms similar to the imperative, or even the apocopate imperfect (but without subject prefixes), seem to underlie Coptic transcriptions:

strong verb, Form II
(6) *saeid* (< *ṣaʿida* II 'to sublimate') – imperative: *ṣaʿʿid* (imperfect *yuṣaʿʿid(u)*, but infinitive *taṣʿīd*)
(7) *taperi* (< *dabara* II 'to prepare') – imperative *dabbir* (imperfect *yudabbir(u)*, but infinitive *tadbīr*)

geminated verbs, Form VII
(8) *nhal* (<*ḥalla* VII 'to dissolve') – imperfect: *yanḥall(u)* (imperative *inḥalil*, colloquial also *inḥall*?, but infinitive: *inḥilāl*)

III*w*, Form II
(9) *safbi*, *sabbi* (< *ṣafā* II 'to clean') – imperative: *ṣaffi* (imperfect *yuṣaffi*, but infinitive *taṣfiya*)

III*y*, Form IV
(10) *eišouei* (*šawā* IV 'to roast, to fry') – imperative: *ʾašwi* (or infinitive *ʾišwāʾ*?)
(11) *eicri* (< *jarā* IV 'to carry out') – imperative: *ʾajri* (or infinitive *ʾijrāʾ* ?)

Iʾ
(12) *am(m)our* (<*ʾamara* 'to command') – imperfect: *yaʾmur(u)* (but infinitive *ʾamr*)

3.3 *Function morphemes*

Only one function morpheme probably borrowed from Arabic has been identified thus far. In two Coptic alchemistic treatises, a morpheme *ô-* linking entity terms to each other occurs (Stern 1885 passim; Bodl. ms (P)a2, 70), which is considered to be identical to the Arabic conjunction *wa-*, e.g. *cop p-ašêlas ô p-almêstik*^{h}*e ô p-assampak* 'take the whey (*aš-šīrāz*) and (*wa-*?) the mastic gum (*al-maṣṭakā*) and (*wa-*?) the lily (*az-zanbaq*)' (Stern 1885:VII, 18–19). In an amazing example of written code-switching, the same conjunction, although now written in Arabic script, is used elsewhere to link Coptic-written Arabic nouns: *sincipil wa-houlincan wa-kalanfour wa-soumpoul* 'ginger (*zanjabīl*) and *alpinia officinarum* (*xūlanjān*) and (*wa-*) clove (*qaranful*) and (*wa-*) nard (*sunbul*)' (Chassinat 1921:155).

4. INSERTION OF ARABIC WORDS INTO COPTIC SYNTACTIC STRUCTURES

4.1 *Nouns*

Due to the strong analytic type of Coptic syntax (cf. Hintze 1947a; Loprieno 1995:7), the embedding of Arabic words into Coptic syntactic structures works rather easily. All grammatical categories having to do with entity terms, like gender, number, and determination, are marked exclusively by morphemes belonging to determiner paradigms (see above), while the grammatical function of nouns is indicated by distinctive sentence patterns and function morphemes.

(13) P.Lond.Copt. I 487 (Richter 2003): *ai-ti nak ou-alkapele n-at-almisahe*

a=i- [perfective conjugation base + 1 sg. pronoun] *ti* [predicative infinitive] *na=k* [dative preposition + 2 sg. pronoun] ø- [object position] *ou-alkapele* [indefinite sg. article + noun] *n-* [attributive modifier] *at-* [privative nominal base] ø *-almisahe* [zero-article + noun] 'I gave you a tenancy (*al-qabāla*) without survey (*al-misāḥa*)'.

4.2 *Verbs*

Native Coptic verbal lexemes can be realized in two forms, the infinitive (including the *status absolutus* and two distinct forms indicating close connection with a nominal or a pronominal direct object) and the stative (Layton 2000:124–157). However, Coptic verbs borrowed from Greek are restricted to the basic form, the infinitive *status absolutus*, a form which can function as a verbal predicate of any conjugation pattern, as a verbal noun, and as an imperative, depending on the grammatical context. In the few cases of verbal lexemes borrowed from Arabic, the same technique occurs, as is demonstrated by examples 14–16, each showing an Arabic verbal lexeme used as a (Coptic) infinitive in two functions, the imperative and the verbal predicate of a conjugation pattern:

(14) *akêt* (*ʿaqada* IV) 'to boil down, to thicken'

akêt mmo=ou hičô-ou-kôht e=f-kere . . . e=k-šan-bol=f ebol n-3 n-sop k-akêt nmo=f ša=f-rôše 'boil [imperative] them down on a fire which is cold (*qarra*) . . . if you dissolve it 3 times (and) you boil [conjunctive conjugation] it down, it will be enough' (Bodl. ms (P) a3, 28–30)

(15) *eišouei* (< *šawā* IV) 'to roast, to fry'

eišouei mmo=ou tso=ou kata 3 n-hoou šante=k-eišouei mmo=ou 'fry [imperative] them (and) water them during 3 days, until you have fried [limitative conjugation] them' (Bodl. ms (P)a 1, f 12)

(16) *saeid* (*ṣaʿida* II) 'to evaporate, to sublimate'

saeid mmo=ou 'evaporate [imperative] them' (Bodl. ms (P) a1, a11)

nta=f-saeid n-p-assipak n-7 n-sop '(I saw the master), who evaporated [relative converter + perfect conjugation] the quicksilver (*az-zībaq*) 7 times' (Bodl. ms (P) a1, g1)

5. SEMANTIC ISSUES

In those genres of Coptic texts providing Arabic words at all, a great many of the borrowed terms are in some way technical. In Coptic scientific treatises, we encounter names of planets (e.g. *as-soouhre* < *az-zuhara* 'Venus'), constellations (e.g. *assarataan* < *as-saraṭān* 'Cancer'), plants (e.g. *alk*^{h}*a-bôôr* < *al-kāfūr* 'camphor'), spices (e.g. *alboulboul* < *al-fulful* 'pepper'), minerals (e.g. *assipak* < *az-zībaq* 'quicksilver'), chemicals (e.g. *alk*^{h}*iprit* < *al-kibrīt* 'sulphur'), diseases (e.g. *annikrês* < *an-niqris* 'gout'), and mathematical terms (e.g. *alk*^{h}*ousôr* < *al-kusūr* 'fraction'). Although it is difficult to estimate their linguistic significance, there is good reason to doubt the conclusion drawn by Vycichl (1991:215): “The spoken language was full of Arabic words, as one can see from a medical papyrus or a treatise on alchemy”, for these terms are not part of the vernacular vocabulary. Rather, they belong to specialized taxonomic vocabularies, which in general are subject to special rules of borrowability.

In documentary texts, we meet titles (e.g. *amira* < *ʾamīr* ‘commander’), weights (e.g. *almatkal* < *al-miṯqāl* ‘weight of one dinar’), measures (e.g. *arrôpa* < *ar-rubʿ* ‘quarter’), coins (e.g. *derham* < *dirham* ‘dirham’), book-keeping terms (e.g. *nabaka* < *nafaqa* ‘expenses’), and legal words (e.g. *dyn* < *dayn* ‘debt of money’). Further, there are designations for diverse objects, especially vessels (e.g. *alkaroore* < *al-qārūra* ‘flask’), clothes (e.g. *almicar* < *al-miʿjar* ‘cap’), and textiles (e.g. *alk^{h}as* < *al-xazz* ‘silk fabric’), probably referring to specific qualities of the respective categories flask, cap, etc. in a *genus-pro-specie* way. At any rate, Arabic nouns tend to be used in a specialized, narrower sense when taken into Coptic, e.g. *alkapele* (< *al-qabāla*) ‘obligation, contract, etc.’ in the meaning ‘tenancy’, or *alhat* (< *al-ḥadd*) ‘border’ in the sense of ‘bordering estate’, in keeping with the technical use of these words in corresponding Arabic texts of the same genres.

Although the total amount of Arabic loanwords in Coptic is rather low, there is a conspicuous accumulation of Arabic words in two semantic fields: sciences and economy. The first might point towards a high esteem for Arab natural sciences, established in educated circles of Egyptian Christian society, as an eastern counterpart to the well-known reception of Arabic sciences in medieval Spain (cf. Gallego 2003; Burnett 1997). The latter may indicate widespread commercial transactions between Arabic and Coptic speakers. But this sociolinguistic conclusion remains to be proven by broader evidence.

Bibliographical references

Adams, James, Mark Janse, and Simon Swain. 2002. *Bilingualism in ancient society: Language contact and the written word.* Oxford: Oxford University Press.

Amélineau, Émile. 1887. “Un document copte du XIII^e siècle: Martyre de Jean de Phanidjôit”. *Journal Asiatique* 8th series, 9.113–190.

Atiya, Aziz S. (ed. in chief). *The Coptic encyclopedia.* 8 vols. New York and Toronto: Macmillan.

Behnstedt, Peter. 1981. “Weitere koptische Lehnwörter im Ägyptisch-Arabischen”. *Welt des Orients* 12.81–98.

——. 1997. “Koptisch oder Arabisch?”. *Wiener Zeitschrift für die Kunde des Morgenlandes* 87.31–39.

Beltz, Walter. 1983. “Die koptischen Zauberpapyri der Papyrus-Sammlung der Staatlichen Museen zu Berlin”. *Archiv für Papyrusforschung* 29.59–86.

Bishai, Wilson B. 1960. “Notes on the Coptic substratum in Egyptian Arabic”. *Journal of the American Oriental Society* 80.225–229.

——. 1961. “Nature and extent of Coptic phonological influence on Egyptian Arabic”. *Journal of Semitic Studies* 6.175–182.

——. 1962. “Coptic grammatical influence on Egyptian Arabic”. *Journal of the American Oriental Society* 82.285–289.

——. 1964. “Coptic influences on Egyptian Arabic”. *Journal of Near Eastern Studies* 23.39–47.

Björnesjö, Sofia. 1996. “L’arabisation de l’Égypte: Le témoignage papyrologique”. *Egypte/Monde arabe* 27–28.93–106.

Blau, Joshua. 1988. “Some observations on a Middle Arabic Egyptian text in Coptic characters”. *Studies in Middle Arabic and its Judaeo-Arabic variety*, ed. Joshua Blau, 145–194. Jerusalem: Magnes Press.

Bouriant, Pierre. 1904. “Fragment d’un manuscrit copte de basse époque ayant contenu les principes astronomiques des Arabes”. *Journal Asiatique* 10th series, tome 4.117–123.

Burmester, O.H.E. 1965–1966. “Further leaves from the Arabic MS. in Coptic script of the Apophthegmata patrum”. *Bulletin de la Société d’Archéologie Copte* 18.51–64.

Burnett, Ch. 1997. “Translating from Arabic into Latin: Theory, practice and criticism”. *Editer, traduire, interpréter: Essais de méthodologie philosophique*, eds. S.G. Lofts and P.W. Rosemann, 55–78. Louvain-la-Neuve.

Casanova, P. 1901. “Un texte arabe transcrit en caractères coptes”. *Bulletin de l’Institut Français d’Archéologie Orientale* 1.1–20.

Chassinat, Émile. 1921. *Un papyrus médical copte.* Cairo: Institut Français d’Archéologie Orientale.

——. 1955. *Le manuscrit magique copte N^o 42573 du Musée égyptien du Caire.* Cairo: Institut Français d’Archéologie Orientale.

Crum, Walter Ewing. 1902a. “A bilingual charm”. *Proceedings of the Society of Biblical Archaeology* 24.329–331.

——. 1902b. *Catalogue of the Coptic manuscripts in the collection of the John Rylands Library, Manchester.* Manchester: University Press.

——. 1905. *Catalogue of the Coptic manuscripts in the British Museum, London.* British Museum.

——. 1939. *Varia Coptica.* Aberdeen: University Press.

Décobert, Christian. 1992. “Sur l’arabisation et l’islamisation de l’Égypte médiévale”. *Itinéraires d’Égypte: Mélanges offerts au père Maurice Martin s.j.*, ed. Christian Décobert, 273–300. Cairo: Institut Français d’Archéologie Orientale.

Drescher, James. 1948–1949. “A Coptic calculation manual”. *Bulletin de la Société d’Archéologie Copte* 13.137–160.

Field, Frederic W. 2002. *Linguistic borrowing in bilingual contexts.* Amsterdam and Philadelphia: J. Benjamins.

Gallego, María Angeles. 2003. "The languages of medieval Iberia and their religious dimension". *Medieval Encounters* 9.107–139.

Galtier, Émile. 1902. "De l'influence du copte sur l'arabe d'Égypte". *Bulletin de l'Institut Français d'Archéologie Orientale* 2.212–216.

——. 1906. "Coptica–arabica, III: Un manuscrit copte en caractères arabes". *Bulletin de l'Institut Français d'Archéologie Orientale* 5.91–111.

Helderman, Jan. 1997. "Die Sprache der Ruhe". *Bulletin de la Société d'Archéologie Copte* 36. 105–119.

Hintze, Fritz. 1947a. "Die Haupttendenzen der ägyptischen Sprachentwicklung". *Zeitschrift für Phonetik und Allgemeine Sprachwissenschaft* 1.85–108.

——. 1947b. "Bemerkungen zur Aspiration der Verschlußlaute im Koptischen". *Zeitschrift für Phonetik und Allgemeine Sprachwissenschaft* 1.199–213.

——. 1980. "Zur koptischen Phonologie". *Enchoria* 10.23–91.

Hout, Roeland van and Pieter Muysken. 1994. "Modeling lexical borrowability". *Language Variation and Change* 6.39–62.

Ishaq, Emile Maher. 1975. *The phonetics and phonology of the Bohairic dialect of Coptic, and the survival of Coptic words in the colloquial and classical Arabic of Egypt, and of Coptic grammatical constructions in colloquial Egyptian Arabic.* Ph.D. diss., Oxford University.

——. 1991. "Egyptian Arabic vocabulary, Coptic influence on". Atiya (1991:VIII, 112–118).

Kahle, Paul E. 1954. *Bala'izah: Coptic texts from Deir el-Bala'izah in Upper Egypt.* Oxford: Oxford University Press.

Layton, Bentley. 2000. *A Coptic grammar.* Wiesbaden: O. Harrassowitz.

Lemm, Oskar von. 1903. *Der Alexanderroman bei den Kopten.* St. Petersburg: Académie Impériale des Sciences.

Loprieno, Antonio. 1995. *Ancient Egyptian: A linguistic introduction.* Cambridge: Cambridge University Press.

MacCoull, Leslie S.B. 1985. "Three cultures under Arab rule: The fate of Coptic". *Bulletin de la Société d'Archéologie Copte* 27. 61–70.

——. 1989. "The strange death of Coptic culture". *Coptic Church Review* 10:2.35–45.

Peust, Carsten. 1999. *Egyptian phonology: An introduction to the phonology of a dead language.* Göttingen: Peust und Gutschmidt.

Praetorius, F. 1901. "Koptische Spuren in der aegyptisch-arabischen Grammatik". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 55.145–147.

——. 1902. "Koptischer Einfluss im Aegyptisch-Arabisch". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 56.681–684.

Richter, Tonio Sebastian. 2001. "Arabische Lehnworte und Formeln in koptischen Rechtsurkunden". *Journal of Juristic Papyrology* 31.75–89.

——. 2003. "Spätkoptische Rechtsurkunden neu bearbeitet (III): P.Lond.Copt. I 487: Arabische Pacht in koptischem Gewand". *Journal of Juristic Papyrology* 33.213–230.

Rubenson, Samuel. 1996. "Translating the tradition: Some remarks on the Arabization of the patristic heritage in Egypt". *Medieval Encounters* 2.4–14.

Satzinger, Helmut. 1972. "Zur Phonetik des Bohairischen und des Ägyptisch-Arabischen im Mittelalter". *Wiener Zeitschrift für die Kunde des Morgenlandes* 63–64.40–65.

——. 1991. "Bohairic, pronunciation of late". Atiya (1991:VIII, 60–65).

Schenkel, Wolfgang. 2002. "Glottalisierte Verschlußlaute, glottaler Verschlußlaut und ein pharyngaler Reibelaut im Koptischen: Rückschlüsse aus den ägyptisch-koptischen Lehnwörtern und Ortsnamen im Ägyptisch-Arabischen". *Lingua Aegyptia* 10.1–57.

Sobhy, G.P. 1926. "Fragments of an Arabic ms. in Coptic script". *New Coptic texts from the monastery of Saint Macarius*, ed. H.G. Evelyn White, Appendix I. New York: Metropolitan Museum of Art.

——. 1950. *Common words in the spoken Arabic of Egypt of Greek or Coptic origin.* Cairo: n.p.

Stern, Ludwig. 1885. "Fragment eines koptischen Tractates über Alchimie". *Zeitschrift für Ägyptische Sprache und Altertumskunde* 23.102–119.

Vittmann, Günter. 1991. "Zum koptischen Sprachgut im Ägyptisch-Arabischen". *Wiener Zeitschrift für die Kunde des Morgenlandes* 81.197–227.

Vycichl, Werner. 1991. "Vocabulary, Copto-Arabic". Atiya (1991:VIII, 215).

Worrell, William H. 1934. *Coptic sounds.* Ann Arbor: University of Michigan Press.

Tonio Sebastian Richter (University of Leipzig)

Coptic Loanwords

1. Coptic substratum

If substratal influence is understood as "imperfect group learning during a process of language shift" (Thomason and Kaufman 1988:38), it has to be dealt with separately from 'loanwords' which are simply 'borrowings' (Thomason and Kaufman 1988:20ff.). On the phonological level, there is no evidence for Coptic substratal influence. It is true that in all Egyptian Arabic dialects the interdentals have merged with plosives, a phenomenon to be observed on a large geographical scale in other areas with a non-Arabic substratum population (mainly the Levant, Algeria, Morocco), but since it is found elsewhere in isolated points with no such substratum (Aramaic, Berber), e.g. Mecca, Aden, Bahrain, this might simply be explained by drift (cf. the loss of interdentals in most of the Germanic languages).

Of the five syntactic features allegedly due to Coptic substratal influence, namely *ma tuʾʿud* 'come on, sit down!', *ahu inti gēti, īdik faḍya* ‘and here you are, empty-handed!’ (Hinds and Badawi 1986:43), *d-ana l-malik* ‘but I am the king!’, *kibīr ʿannu* ‘older than him’, and postposition of interrogatives, Diem (1979:50ff.) and Palva (1969–1970) have shown that three cannot be attributed to a Coptic substratum. Diem (1979:52) concludes that only the construction of the type *kibīr ʿannu* has to be seriously taken into consideration as a case of substratal influence. In the case of the postposition of interrogatives, whose Coptic provenance had already been questioned by Munzel (1950), Diem is prepared to accept the effect of a Coptic substratum, but only to the extent that one of two possible constructions was preferred. Or, as Versteegh (1997:106) puts it: “the interference that resulted from language contact may have consisted not in the emergence of new phenomena but in the tipping of the balance towards one of the existing alternatives. In such cases, the learners of Arabic may have been influenced by their first language in the selection of the alternative”.

Postposition of interrogatives is, indeed, rare in Classical Arabic, although it is obligatory in a construct sentence: *mawlā man?* ‘whose client?’ (Reckendorf 1977:74) and correspondingly in Egyptian Arabic: *bēt mīn?* ‘whose house?’. The interrogative in a partial interrogative sentence is therefore found in the same position (slot) where the corresponding answer (filler) appears: *mawlā man?* – *mawlā Muḥammad*; *bēt mīn?* – *bēt Muḥammad*; *ḥatīgi imta?* ‘when will you come?’ – *ḥāgi bukṛa* ‘I shall come tomorrow’. Hence, the postposition of interrogatives in Egyptian Arabic is simply the extension of a rule. But interrogatives cannot be postponed arbitrarily: *ḍarabu mīn?* can only mean ‘whom have they beaten?’ Answer: *ḍarabu Muḥammad*. This cannot mean ‘who has beaten him?’ The latter would be *mīn (illi) ḍarabu?* Answer: *Muḥammad (illi) ḍarabu*. Besides, postposition of interrogatives in Coptic is not the rule. Till (1970:102–103) gives more examples with the interrogative in initial position of the sentence than in final position (*ntok nim?* ‘who are you?’ but *nim ne nai?* ‘who are they?’). As for the sole ‘fairly sure case’ (“einigermaßen sicherer Fall”, Diem 1979:52), namely *kibīr ʿannu* ‘older than him’, the Coptic model would be a construction with the preposition *e~erō* ‘in comparison with’, e.g. ‘A is old in comparison with B’ = ‘A is older than B’ (Till 1970:77–78).

But the Egyptian Arabic construction has also to be compared with similar constructions: in Algerian and Moroccan Arabic: *kbīr ʿlī* ‘older than him’; Nigerian Arabic *kabīr minnu*; in Southern Arabia *akbar ʿannu* (more examples in Procházka 1993:75; also Watson 2004:94 for Yemen); in Syria and Lebanon (*kbāṛ ʿan*, *xfīf ʿan*, Procházka 1993:76); and in the dialects (of Syrian origin) of the Çukurova (Procházka 2002:154 *ha-l-bayt əjdīd ʿan baytna* ‘this house is newer than ours’). Consequently, there is no reason to assume Coptic substratal influence when applying one of Diem’s rules, which says that a substratal phenomenon has to be exclusive. Since *kibīr ʿannu* and similar constructions are also common in Algeria, Morocco, Syria, and elsewhere, *kibīr ʿannu* cannot be attributed to a Coptic substratum in Egyptian Arabic to begin with.

A comparative construction with *ʿan* is found in Classical Arabic, e.g. *taḍīqu jufūnu l-ʿayni ʿan ʿabarātihā* ‘the eyelids are too narrow for her tears’ (Reckendorf 1977:214); *ʾinnī ʾaḥbabtu ḥubba l-xayri ʿan ḏikri rabbī* ‘I loved the agreeable things more than the thinking of my Lord’ (Reckendorf 1977:235); *lī fī ṭilābi l-ʿilmi ġinan ʿan ġināʾi l-ġāniyāti* ‘I find more satisfaction in studying than in the singing of women’ (other translations in Wright 1974:III, 141). Corresponding sentences in Egyptian Arabic are: *yixdim ʿan il-babūṛ* ‘it ploughs better than the tractor’ (Behnstedt and Woidich 1994:109); *il-bihīm l-axṛas di, byaʿrif ʿan il-insān* ‘this dumb animal knows it better than man’ (Behnstedt and Woidich 1987:127, text 23 sentence 5). The comparative of the type *kibīr ʿannu* in Egyptian Arabic is “possible, but less customary” (Mitchell 1956:90) as it is in other Arabic dialects. Were it the normal or the more frequent construction in Egyptian Arabic, then perhaps a Coptic model might have contributed to the tipping of the balance between two possible constructions of Arabic dialects. But this is not the case.

2. Coptic loanwords

One problem when trying to establish the etymology of Egyptian Arabic dialect words of supposed Coptic origin is that a good deal of the lexicon of the Coptic dictionaries is based on literary texts, most of them religious, and that consequently they do not deal extensively with rural or botanical terminology or give a complete list of Nile fishes. Many Coptic etymologies given in various studies have to be rejected, either because they are simply Arabic, onomatopoetic, or are not attested in Coptic (see Behnstedt 1981, 1997; Vittmann 1991). There are, however, many rural, botanical, and miscellaneous lexical items, which are definitely not Arabic and which sound Coptic: *bittāw* 'bread loaf', *bilinf* 'vegetable earth, mould', *bišlūf* 'a kind of dates', *baṛūf* 'small peg on the yoke', *dignāš* 'little sparrow' (Behnstedt and Woidich 1994); or plant names like *amšūt*, *awāy*, *balatāy*, *bašift*, *buṛuwaks*, *daradiks*, *diktāy* (Täckholm 1974). But the Coptic dictionaries cannot help here. The initial *b-* and *d-* might be interpreted as the Coptic masculine and feminine articles *pi-* (*pe-*) and *ti-*.

Intensive research on rural Egyptian Arabic vocabulary carried out by Behnstedt and Woidich 1994, Henein (1988), Laferrière and Ménassa (1974), Riad (1960), Winkler (1936), and others has provided much new vocabulary for which a Coptic origin could be detected, and a recent meticulous examination by Behnstedt and Woidich (2005) gives us some 180 'valid' lexical Copticisms. Assuming that many names of plants and animals, which sound Coptic but are not attested in the Coptic lexica, are of Coptic origin, and taking into account that the rural lexicon has not yet been investigated exhaustively all over Egypt, the total rate might be estimated up to between approximately 250 and 300 loans.

There are some dialect words for which a Coptic origin has been claimed, but which cannot be localized in Egypt (not in Hinds and Badawi 1986; Spiro 1977; Behnstedt and Woidich 1994) like *šalla* 'scorpion' (Ishaq 1991: III [4]; Schenkel 2002:6) < *čʰlē*; *jiffa* 'frost' (Ishaq 1991:IV C [2]; Schenkel 2002:28) < *čaf*, *čef*; *janafōr* 'roof' (Ishaq 1991:S.I V.2; Schenkel 2002:22) < *čenepōr*; *darafs* 'awl, spike' (Ishaq 1991:XX 2; Schenkel 2002:21) < *thraps*, *traps* 'awl, needle'.

The phonological representation of Coptic sounds in Egyptian Arabic is dealt with by Schenkel (2002). With respect to morphology Coptic nouns have been integrated into Egyptian Arabic either with the definite articles *p*(*i-*, *e-*) (masc.) and *-ti* (fem.) or without: *antūb* (rare), *antūt* ~ *bantūt* 'peg on the beam of the plough to which by means of a chain or a rope the yoke is fastened' < *an-tʰōb*, *ham-n-tōb* 'needle' (Behnstedt 1981:83; cf. *misalla* 'big needle' used in Upper Egypt with the same meaning); *dišīda* ~ *pidya* 'sling of the draught animals on the yoke [often made of palm fibres]' < *t-šēte*, *ti-šēti* 'palm fibre'. There is a clear geographical distribution in the Nile Delta for forms with and without article (see Behnstedt and Woidich 1985: maps 492, 506: *antūt* ~ *bantūt*, *šidya* ~ *dišīda*). In some cases, Coptic words have been interpreted as plurals and a new Arabic singular has been formed by reanalysis, e.g. *ginw*, pl. *ignāw*, or sg. *ignāwa* (*ignāwt il-balaḥ*) 'date stalk, or stem of the date stalk' (Behnstedt and Woidich 1994) < *kʰnau*, *knaau* 'yarrow'; or *gaṭāwi*, sg. *gaṭwiyya* 'big basket for transport on camels' < **katooue* 'baskets' (sg. *kat*), or verbs were borrowed as nouns.

The following list of Coptic items in Egyptian Arabic, which is based on Behnstedt and Woidich (2005), only deals with the lexical side and is not exhaustive; it covers the following domains:

i. Christian terminology, e.g. *gabanyōt*, *jabanyōt* 'Lord's prayer' < *če-* + *pe=n(e)iōt* 'our father'; *ṭabḥa* 'prayer' < *tobh* 'to beg'; *amnūt* 'sexton' < *mnout* 'porter, door keeper'; *hōs* 'hymn' < *hōs* 'to sing'; *lubš* 'a kind of prayer' < *lōbš* 'crown' used to refer to a final stanza in certain hymns; *šūrya* 'a vessel for frankincense' < *šourē* 'idem'.

ii. Rural terminology:

   a. measures of capacity: *ardabb* '198 litres' < *artab*; or *rafṭāw* and similar '1/4 of a *kēla*' < *re*, *ra* 'part' + *ftoow* 'four' = 'fourth part'.

   b. names of the months only used in agriculture (cf. Wassef 1971) and farmers' weather maxims: *tūt hāt il-antūt!* 'in the month Tūt bring the small peg [of/to the plough-beam]!' = 'in the month Tūt start ploughing!', *baramhāt rūḥ il-ġēṭ wi hāt!* 'in the month Baramhāt go to the field and bring!' = 'harvest!'.

   c. agricultural tools, or names of plough parts and implements: *ṭūrya* 'hoe' < *tōri* 'axe, spade, hand'; *hōgal*, *hōjal* 'flail' (Hinds and Badawi 1986), 'rake' (Behnstedt and Woidich 1994), 'anchor' (Bishai 1960:42) < *haukʲal* 'anchor, hook'; *nāf* 'yoke' < *nahb*, *nahbef* 'back'; *basxa*, *bisxa* 'plough sole' < *pe-sho*, *pi-sxo* 'handle of the plough' (another possible etymology in Schenkel 2002:43, 51); *bajrūm* 'frog of the plough-beam' < *p-čarōb*, *p-kʲarōm* 'stick'; *dihiks*, *duhuks*, and the like 'the lower and enlarged part of the iron whip-stick which serves to clean the ploughshare from mud' < *ti-hoks* 'scratcher'; *bayš*, *bēš* 'crossbeam on the runners of the threshing-sledge' < *p-oeiš* 'rung of a ladder'.

d. designations relative to irrigation: *dimīra* 'season of the Nile inundation' < *temēre* 'idem'; *šalāw* 'rope of the water-wheel where the jars (*ʿawadīs*, *gawadīs*) are fixed' < *šalau*, *šaloou* 'water-wheel or similar' (Westendorf 1965– 1977:311); *hūdya*, *hudya*, and the like 'rear driving pole on a water-wheel' < *hōte*, *hōti* 'pole'; *šadūf* 'counterpoised implement for raising irrigation water' is not attested in Coptic. Schenkel (2002:24) suggests an Egyptian form **šaṭ=w-u̯'f* 'the one with the bucket [lit. water-hose]?'. The original meaning of *šadūf* in Egyptian Arabic is, however, 'basket' (cf. Behnstedt and Woidich 1994, s.v.) and a basket may be used for drawing the water with the shadoof.

e. terms designating types of fields, parts of them, soil: *barš* 'field free from any traces of last year's crops and which consequently can be tilled and sown' < *porš* 'even surface' (Brunner in Halm 1979:79); *ṭamy* 'mud [of the Nile]' < *t-o(o)me*, *t-aame* 'mud, clay'; *ṭāš* 'boundary balk' < *taš*, *tōš*, etc. 'border'; *bitm*, *bitn* 'earth wall, mould' < *p-hiten*, *p-eitn*, *p-itn* 'soil, ground'.

f. diverse: *hayy*, *hāya* 'heap of cereals' < *hoi* 'idem'; *širš*, pl. *širaš*, *šṛašš* 'yarrow' < *šraš* 'yarrow, bundle'. As for *maxwal* 'feeding trough, hen-house, rabbit-hutch, partition made of clay within the house for storing cereals' < *mahoual* 'nest, dovecot' (Behnstedt 1981:90), Classical Arabic *xawal* 'property' and *maxwal* 'food storage room' used in ʿAsīr (Dostal 1983:31) or Yemeni *maxwalah* 'small bowl, tray' (Behnstedt 1992:351) rather suggest an Arabic origin; a contamination of the two forms is, however, conceivable, especially for the meaning 'hen-house'.

g. rural implements like vessels, baskets: *ḥinn* 'earthenware bowl for milking' < *hin* 'vessel, bowl' (perhaps influenced by *ḥann/yiḥinn* 'to knead the teats of the cow's udder before milking'); *bukla* 'water jar' < *pe-kle* and similar; *bihnāw* 'besom made of the palm-stalk after the dates had been taken off' < *pi-hnau* '(blossoming) palm twig' (thus Westendorf 1965–1977:380); *šinda* 'mat [hung up] in which fresh cheese is wrapped up in order to lose more whey' < *šnte*, *šenti* 'basket-work'.

iii. Names of animals: *baʾrūr* 'toad', *baqrūr* 'frog' (similar forms mainly in the oases like *bagaṛōṛa*, *bagarūr*, *bagaṛūṛa*, *buglul*, elsewhere not attested) < *pe-krour*; *balašōn* 'heron' < *pelkʲōb*; names of Nile fishes like *šāl* 'Synodontis schall' < *čal*; *šilba* 'Silurus mystus' < *kʲ(e)lboou*, *čelfau*; *libīs* 'Cyprinus niloticus' < *labis*; *būri* 'grey mullet' < *bōre*; and others. Cf. also *habya* 'drag net' < *abou*, *abooue* 'idem' (probably interpreted as a plural form).

iv. Names of plants: *burdi* 'papyrus' (Vollers 1896:653), *barsīm* 'clover' < *bersim*; *burnuf*, *barnūf* 'Conyza dioscoridis?' (more details in Behnstedt and Woidich 1994, s.v.) < *pernoufe* 'a plant' (Crum 1939:269); *bašmīn* 'lotus' (Vittmann 1991:221); *sanṭ(a)* 'acacia nilotica' < *t-šante*; *hallūs* 'type of pondweed', 'cobwebs' (Hinds and Badawi 1986, s.v.) < *halous* 'spider web'.

v. Various: *birba* 'temple ruin' < *pe-rpe*; *jukṣ* 'crepitus ventris' < *čoksi*; *bōš* 'size, industrial starch' (Hinds and Badawi 1986:112); *būš* 'porridge' (Bishai 1964:41) < *p-oouš* 'pap'. There are almost no designations of body parts having a Coptic origin, only *bōfa* 'lung of a slaughtered animal' (Behnstedt and Woidich 1994) < *p-ouof*, *bōf* 'lung' is attested.

All told, Coptic loanwords are rare in the basic vocabulary of Egyptian Arabic and mainly cover a field which can be designated by 'rural, agricultural' or 'local natural and local cultural items', which confirms the view of Thomason and Kaufman (1988:117): "If the language of a shifting population did not contribute lexicon to the target language, other than a few words for local natural and cultural items, then we can conclude that the shifting population did not enjoy much social or political prestige". Some Coptic words, however, have been integrated even into Standard Arabic, e.g. *wāḥa* 'oasis', *timsāḥ* 'crocodile' < **ti-msah* (with a problematic feminine article, the Coptic word *msah* in the dictionaries being masculine, so one would expect *bimsāḥ*!) and *haram* 'pyramid'; and one word has acquired some internationality, namely *ṭūb* 'clay', which through Spanish *adobe* has even entered into English.

BIBLIOGRAPHICAL REFERENCES

Atiya, Aziz S. (ed. in chief). 1991. *The Coptic encyclopedia.* 8 vols. New York and Toronto: Macmillan.

Behnstedt, Peter. 1981. "Weitere koptische Lehnwörter im Ägyptisch-Arabischen". *Welt des Orients* 12. 81–98.

——. 1992. *Die nordjemenitischen Dialekte.* II. *Glossar alif – dāl.* Wiesbaden: L. Reichert.

——. 1997. "Koptisch oder Arabisch?". *Wiener Zeitschrift zur Kunde des Morgenlandes* 87.31–39.

—— and Manfred Woidich. 1985. *Die ägyptisch-arabischen Dialekte.* II. *Dialektatlas von Ägypten.* Wiesbaden: L. Reichert.

——. 1987. *Die ägyptisch-arabischen Dialekte.* III. *Texte. I. Delta-Dialekte.* Wiesbaden: L. Reichert.

——. 1994. *Die ägyptisch-arabischen Dialekte.* IV. *Glossar Arabisch-Deutsch.* Wiesbaden: L. Reichert.

——. 2005. *Arabische Dialektgeographie: Eine Einführung.* Leiden: E.J. Brill.

Bishai, Wilson B. 1959. *The Coptic influence on Egyptian Arabic.* Baltimore.

——. 1964. "Coptic lexical influence on Egyptian Arabic". *Journal of the Near Eastern Society* 23. 39–47.

Crum, Walter E. 1939. *A Coptic dictionary.* Oxford: Clarendon Press.

Diem, Werner. 1979. "Studien zur Frage des Substrats im Arabischen". *Der Islam* 56.12–80.

Dostal, Walter. 1983. *Ethnographic Atlas of ʿAsīr: Preliminary report.* Vienna: Verlag der Österreichischen Akademie der Wissenschaften.

Halm, Heinz. 1979. *Ägypten nach den mamlukischen Lehensregistern.* Wiesbaden: L. Reichert.

Henein, Nessim Henry. 1988. *Mārī Girgis: Village de Haute-Egypte* (= *Bibliothèque d'Etude*, 94). Cairo: Institut Français d'Archéologie Orientale.

Hinds, Martin and El-Said Badawi. 1986. *A dictionary of Egyptian Arabic: Arabic-English.* Beirut: Librairie du Liban.

Ishaq, E.M. 1991. "Coptic influence on Egyptian Arabic vocabulary". Atiya (1991:VIII, Appendix: Linguistics).

Laferrière, Pierre and Laïla Ménassa. 1974. *La* sāqia: *Technique et vocabulaire de la roue à eau égyptienne* (= *Bibliothèque d'Etude*, 47). Cairo: Institut Français d'Archéologie Orientale.

Lane, Edward W. 1980. *An Arabic-English lexicon in eight parts.* (Repr. Beirut.)

Littmann, Enno. 1901. "Koptischer Einfluß im Ägyptisch-Arabischen". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 56.681–684.

Maccouli, L.S.B. 1985. "Three cultures under Arabic rule: The fate of Coptic". *Bulletin de la Société d'Archéologie Copte* 27.61–70.

Mitchell, Terence F. 1956. *An introduction to colloquial Egyptian Arabic.* London: Oxford University Press.

Munzel, Kurt. 1950. "Zur Wortstellung von Ergänzungsfragen im Arabischen". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 100.566–576.

Palva, Heikki. 1969–1970. "Notes on the alleged Coptic morphological influence on Egyptian Arabic". *Orientalia Suecana* 18.128–136.

Procházka, Stephan. 1993. *Die Präpositionen in den neuarabischen Dialekten.* Vienna: VWGÖ.

——. 2002. *Die arabischen Dialekte der Çukurova (Südtürkei).* Wiesbaden: O. Harrassowitz.

Reckendorf, Herman. 1977. *Arabische syntax.* 2nd ed. Heidelberg: C. Winter.

Riad, M. 1960. "Native plough in Egypt". *Bulletin de la Société de géographie d'Egypte* 33.241–277, plates I–VIII.

Rizqallah, Fawzeya and Kamel Rizqallah. 1978. *La préparation du pain dans un village du delta égyptien (province de Charqia).* Cairo: Institut Français d'Archéologie Orientale.

Schenkel, W. 2002. "Glottalisierte Verschlußlaute, glottaler Verschluß und ein pharyngaler Reibelaut im Koptischen: Rückschlüsse aus den ägyptisch-koptischen Lehnwörtern und Ortsnamen im Ägyptisch-Arabischen". *Lingua Aegyptia. Journal of Egyptian Language Studies* 10.1–57.

Sobhy, George P.G. 1950. *Common words in the spoken Arabic of Egypt.* Cairo: Publications de la Société d'Archéologie Copte.

Spiro, Socrates. 1977. *An Arabic-English dictionary of the colloquial Arabic of Egypt.* (Repr. Beirut: Librairie du Liban.)

Täckholm, Vivi. 1974. *Student's flora of Egypt.* 2nd ed. Beirut.

Till, Walter C. 1970. *Koptische Grammatik.* 4th ed. Leipzig: VEB Verlag.

Thomason, Sarah Grey and Terrence Kaufman. 1988. *Language contact, creolization and genetic linguistics.* Berkeley: University of California Press.

Versteegh, Kees. 1997. *The Arabic language.* Edinburgh: Edinburgh University Press.

Vittmann, G. 1991. "Zum koptischen Sprachgut im Ägyptisch-Arabischen". *Wiener Zeitschrift zur Kunde des Morgenlandes* 81.197–227.

Vollers, Karl. 1886. "Beiträge zur Kenntnis der l ebenden arabischen Sprache in Aegypten. II. Über Lehnwörter, Fremdes und Eigenes". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 50. 607–657.

——. 1897. "Beiträge zur Kenntnis der lebenden arabischen Sprache in Aegypten. II. Über Lehnwörter, Fremdes und Eigenes". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 51.291–364.

Wassef, C.W. 1971. "Le calendrier copte de l'antiquité à nos jours". *Journal of the Near Eastern Society* 30.1–48.

Watson, Janet. 2004. Review of: Procházka (2002). *Zeitschrift für Arabische Linguistik* 43.93–95.

Westendorf, Wolfhart. 1975–1977. *Koptisches Handwörterbuch.* Heidelberg: C. Winter.

Winkler, Hans Alexander. 1936. *Ägyptische Volkskunde.* Stuttgart: Kohlhammer.

Wright, William. 1974. *A grammar of the Arabic language.* (Repr. Beirut.)

Youssef, Ahmed Abdel-Hamid. 2003. *From Pharaoh's lips: Ancient Egyptian language in the Arabic of today.* Cairo: American University in Cairo Press.

PETER BEHNSTEDT (Chipiona, Spain)

Copula

Arabic root clauses are divided into two distinct categories: verbal and nominal. While the former category exhibits a verbal form, the latter does not. This is shown in (1) and (2) (Assahārīn 2003:46):

(1) *yu-rḍiʿu* *al-ḥūt-u*
Imp.3.m-breastfeed.s the-mammals-Nom
ṣiġāra-hu
babies-his
'Mammals breastfeed their babies.'

(2) Ø *al-ḥūt-u* *ʾaṭwalu* *min*
the-mammals-Nom taller than
at-timsāḥ
the-crocodile
'Mammals **are** taller than crocodiles.'

While the sentence in (1) contains a verb *yu-rḍiʿu* 'breastfeed' inflected for the imperfect tense-aspect category and nominal agreement features such as person, gender, and number, the sentence in (2) appears without any lexically realized verbal form, as indicated by the null symbol Ø. Despite the apparent discrepancy, each sentence is considered complete, for it fulfills the speaker's intended meaning. Such root clauses are common in Arabic and are systematically introduced in grammar and pedagogical textbooks as such. This entry focuses on nominal sentences similar to those in (2). First, the syntagmatic properties of copular sentences are examined. This is followed by a discussion of recent approaches to the analysis of copular sentences. Finally, an analysis of copular sentences in Standard Arabic is presented taking into account a wide range of issues that are otherwise inexplicable and unaccounted for.

1. COPULAR SENTENCES

The copula in Arabic has been a topic of interest and discussion since the early works of the 8th century Arab and Muslim grammarians (e.g. Sībawayhi). The examples in (3) all include a verbal form except (3d):

(3) a. *nām-a* *aṭ-ṭālib-u*
sleep.Pst the-student-Nomin
fī l-faṣl
the-classroom
'The student slept in the classroom'

b. *kāna* *aṭ-ṭālib-u*
be.Pst the-student-Nom
fī l-faṣl
in the-classroom
'The student was in the classroom'

c. *sa-yakūnu* *aṭ-ṭālib-u*
Fut-be the-student-Nom
fī l-faṣl
in the-classroom
'The student will be in the classroom'

d. *aṭ-ṭālib-u* *fī l-faṣl*
the-student-Nom in the-classroom
'The student is in the classroom'

If a verbal form, i.e. a copular verb like that in (3b) and (3c), is used, the sentence becomes ungrammatical, as illustrated by (4).

(4) **yakūnu* *aṭ-ṭālib-u*
be the-student-Nom
fī l-faṣl
in the-classroom
'The student is in the classroom'

The examples in (3a–b) and (3c) appear in the past and future tenses respectively. They all contain verbal forms, *nāma* and *kāna* in (3 a–b), and the copular form *sa-yakūnu* in (3c). Despite their apparent similarities, (3a) contains a regular intransitive verb and (3b) and (3c) make use of the past and future forms of the copular verb *kāna* 'to be'. The example in (3d), however, does not exhibit any verbal form, and as (4) shows, the sentence is ungrammatical once the corresponding copular verb *yakūnu*, the present form of *kāna*, is used.

The examples in (5) illustrate other types of constructions where the present form of the copular verb *yakūnu* is absent.

(5) a. *aṭ-ṭālib-u* *marīḍun*
the-student-Nom sick
'The student **is** sick'

b. *aṭ-ṭālib-u* *jār-ī*
the-student-Nom neighbor-my
'The student **is** my neighbor'

c. *aṭ-ṭālib-u* *hunā*
the-student-Nom here
'The student **is** here'

These, together with (3d), show that sentences sensitive to the presence of the copula cover a wide range of structures in their predicates. The initial Noun Phrase (NP) is followed by a Prepositional Phrase (PP) in (3d), an Adjective Phrase (AdjP) in (5a), a NP in (5b), and an Adverb Phrase (AdvP) in (5c). The diagram in (6) represents this variation.

(6) S
/\
NP XP (PP, AdjP, NP, AdvP, *VP)

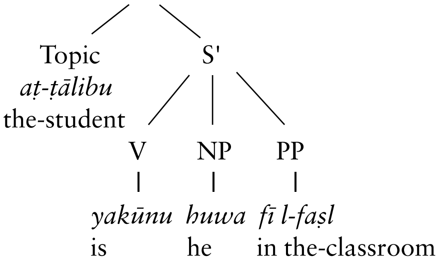
In sum, apart from a Verb Phrase (VP), the initial NP of the copular sentence may be followed by any other lexical category. The intriguing question is how to account for the presence of the copula in past and future tense contexts (3b–c), and its obligatory absence in present tense contexts, as shown in (3d), (4), and (5a–c).

2. Analyses of copular sentences

Within the generative Arabic linguistics tradition, a number of analyses have been suggested to account for the contrast between (3d) and (3b–c), four of which are discussed below.

The Null Copula analysis assumes that at some underlying or representational level the Arabic nominal clause contains a verb (Bakir 1980; Abdul-Ghany 1981; Farghal 1986; Al-Waer 1987). Bakir (1980), for example, argues that copular sentences are S" (S double-bar) projections, with the first NP as topic and its predicate contained within a lower sentence. He argues that the copula and the lower subject undergo deletion, yielding the topic NP and the Predicate at surface structure. According to this analysis (3a) is represented as in (7).

(7)



Likewise, Fassi Fehri (1982) argues for the presence of a phonetically unrealized head, or a copula inserted as null (1993) along with the presence of a Tense Phrase (TP) projection.

Despite the apparent success that such analyses have enjoyed, case assignment and the obligatory presence of the copula in certain contexts constitute a major challenge to the Null Copula analysis. Bahloul (1994:201), for example, shows that the ungrammaticality of sentences such as (8b), where "the presence of the copula in a present/timeless context is not even optional" militates against the Null Copula analysis.

(8) a. *qad takūnu l-ʾarḍ -u*
may be the-earth-Nom
mustadīrat-an
round-Acc
'The earth may be round'

b. **qad ʾarḍ -u mustadīrat-un*
may the-earth-Nom round-Nom

Benmamoun (2000:42–43) questions the validity of the Null Copula analysis on the basis of case assignment. The copula, like transitive verbs and other functional words in Arabic, assigns accusative case to the predicate. See (9), where the ungrammaticality of (9b) is attributable to the accusative case. This establishes the inadequacy of postulating deleted or null copulas in such constructions; a copula should assign the same case, be it overt, deleted, or null.

(9) a. *kāna ṭ-ṭālib-u marīḍ* ***-an***
be past the-student-Nom sick-**Acc**
'The student was sick'

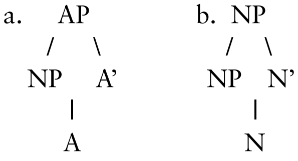
b. *aṭ-ṭālib-u marīḍ-*****an/un***
the-student-Nom sick-***Acc/Nom***

The Small Clause analysis draws on apparent similarities between the absence of a copula in Arabic copular sentences and their counterparts in certain English constructions, specifically 'small clauses', which are bracketed in (10):

(10) a. I find [the course challenging]
b. They consider [Mary an excellent teacher]

The small clause in (10a) consists of a noun phrase and an adjective phrase, while that in (10b) contains a subject noun phrase and a predicate noun phrase. These small clauses look similar to the Arabic copular sentences in (5a) and (5b), respectively. By analogy to such English structures, Mouchaweh (1986) argues that Arabic copular sentences are best analyzed as small clauses. Rapappport (1987) argues the same for Hebrew, whose copular constructions are similar to those in Arabic. Like small clauses, copular sentences lack any verbal or tense phrases. As a result, English small clauses (10a–b) and Arabic copular sentences (5a–b) are all generated within the structures in 11:

(11)



Attractive as this analysis may appear, it falls very short of accounting for certain peculiarities of the Arabic copular construction. Benmamoun (2000:39–42) provides ample evidence against the small clause analysis. He shows, for example, that unlike small clauses embedded under a tensed matrix, copular sentences have fixed temporal reference. He also provides evidence from case assignment, WH-movement, and the distribution of negative polarity items, showing that copular sentences are not small clauses. He proposes instead that they are clauses with tense projections, and establishes a correlation between the inherent features of the functional category 'Tense' as being specified or unspecified for a [+V] feature and the surfacing of the copula to check such features.

Temporal interpretation, however, is not by itself a sufficient condition for the presence or absence of the copula. Aspect, mood, and modal elements play crucial roles as well. Any adequate account of copular sentences must also address the systematic absence of the copula in examples like (12a) and its obligatory use in others like (12b) and (13)–(21). These illustrate contexts that condition the use of copula in Standard Arabic: temporal (12), atemporal (13), aspectual (14–17), modal (18), and mood contexts (19–21).

The examples in (12) show the sensitivity of the copula to the [+/– past] feature. The copula is systematically absent when the described event is not located prior to the moment of speech, as in (12a). On the other hand, the copula is systematically present when the event is located prior to the moment of speech, as in (12b).

i. Temporal context (present, past):

(12) a. *_y-akūn-u r-rajul-u fī_
is the-man-Nom in
d-dār-i
the-house-Gen
'The man is in the house'

b. *_(kāna) r-rajul-u fī_
was the-man-Nom in
d-dār-i
the-house-Gen
'The man was in the house'

ii. Atemporal context (true in the past, present, and future):

The examples in (13) show that events that are true in the past, present, and future do not trigger the use of the copula. In other words, statements with no temporal anchor do not require the presence of a copular verb.

(13) a. *_takūnu l-ʾarḍu mustadīratun_
is the-earth-Nom round
'*The earth is round'

b. *_yakūnu ramaḍānu šahru_
is Ramadan-Nom month
ṣ-ṣiyām-i
the fasting-Gen
Ramadan is the month of fasting'

iii. Aspectual context (habitual):

Referring to recurrent events as in the case of habit description calls for the obligatory presence of the copula, as illustrated in examples (14)–(17). In these examples, the recurrence effect or habitual meaning is mediated through the presence of the adverbs _ʿādatan-mā_ 'usually' as in (14), _ʿindamā_ 'when, whenever' as in (15), _ḥīnamā_ 'when, whenever' as in (16), and _lammā_ 'when, whenever' as in (17). Omitting such temporal adverbs results in ungrammaticality. Their presence is therefore a necessary condition in licensing the copula.

(14) a. **_ʿādatan-ma_** *_(y-akūn-u)_
usually is
r-rajul-u fī d-dār-i
the-man-Nom in the-house-Gen
'The man is usually in the house'

b. *_y-akūn-u r-rajul-u fī_
is the-man-Nom in
d-dār-i
the-house-Gen
'*The man is in the house'

(15) a. ***ʿindamā*** **(y-akūn-u) r-rajul-u*
when is the-man-Nom
marīḍ-an faʾinna-hu lā y-ubālī
sick-Acc then-that-he not 3rd-cares
'When/ever the man is sick, he does not care any more'

b. **y-akūn-u r-rajul-u marīḍ-an . . .*
'*The man is sick'

(16) a. ***ḥīna*** (ʾ-akūnu) fī d-dār-i*
when be.1st.sg in the-house-Gen
ʾ-akūn-u murtāḥ-an
be.1st.s relaxed-Acc
'When/ever I am at home, I feel relaxed'

b. **ʾ-akūnu fī d-dār-i ʾ-akūn-u murtāḥ-an*
'*I am in the house, I am relaxed'

(17) a. ***lammā*** **(y-akūn-u) ṭ-ṭaqs-u*
when is the-weather-Nom
jamīl-an ʾ-akūn-u murtāḥ-an
beautiful-Acc I-am relaxed-Acc
'When/ever the weather is beautiful, I feel relaxed'

b. ** y-akūn-u ṭ-ṭaqs-u jamīl-an . . .*
'*The weather is beautiful . . .'

iv. Modal context (will, may, must, and can):

The examples in (18) show another context where the presence of the copula is obligatory. The use of modals such as *sawfa* 'will' as in (18a), *qad* 'may' as in (18b), *yajibu ʾan* 'must' as in (18c), and *y-astaṭīʿ-u ʾan* 'can' as in (18d) in nominal sentences requires the copula.

(18) a. ***sawfa*** **(y-akūn-u) r-rajul-u*
will is the-man-Nom
wāqif-an
standing-Acc
'The man will be standing up'

b. ***qad*** **(y-akūn-u) r-rajul-u*
may is the-man-Nom
wāqif-an
standing-Acc
'The man may be standing up'

c. ***yajibu ʾan*** **(y-akūn-a) r-rajul-u*
must is the-man-Nom
wāqif-an
standing-Acc
'The man must be standing up'

d. ***y-astatīʿ-u ʿan*** **(y-akūn-a)*
3rdm-can-s is
r-rajul-u mudīr-an
the-man-Nom director-Acc
'The man can be a director'

v. Mood context (interrogatives, conditionals, and imperatives):

Finally, the presence of such mood markers as wh-constituents *matā* 'when' as in (19a), conditionals *ʾin*, *law* 'if' in (20a–b), and imperatives in (21) necessitates the obligatory spelling out of the copula.

(19) a. ***matā*** **(y-akūn-u) ʾabū-ka*
when is father-your
fī d-dār-i
in the-house-Gen
'When is your father at home?'

(20) a. ***ʾin*** **(kāna) l-ʾustāḏ-u*
if be.m.s the-professor-Nom
nājiḥ-an . . .
successful-Acc
'If the professor is successful . . .'

b. ***law*** **(kāna) l-ʾustāḏ-u*
if be.m.s the-professor-Nom
marīḍ-an . . .
sick-Acc
'If the professor were sick . . .'

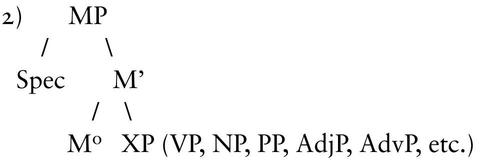
(21) a. **(kun) ʾustāḏ-an nājiḥ-an*
be.m.s professor-Acc successful-Acc
'Be a successful professor!'

b. *lā* (t-akun) ġabiyy-an*
not 2-be.m.s silly
'Do not be silly!'

It should be clear, then, that the presence or the absence of the copula is not limited to cases such as those in (12a–b) where the sentential temporal structure is based on present and past tense interpretations. It must be extended to include correlations with other functional categories, specifically aspectual, modal, and mood markers.

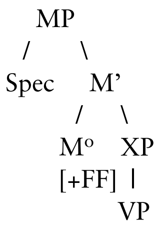
As a result, any adequate analysis of Arabic copular sentences must provide an adequate account of functional categories, since their presence not only affects the syntagmatic properties of sentences (i.e. the presence of a copular verb for grammaticality) but, more importantly, the semantic interpretation and propositional truth value. On that basis, the structure in (22) is hereby adopted as a schematic representation for copular constructions. It stipulates that sentences are projections of Modal Phrases (MP), similar to Inflectional Phrases (IP) but with much more interpretive power, and that modality is a major sentential component which includes functional categories (see Bahloul 1994:7–30 for more details).

(22)

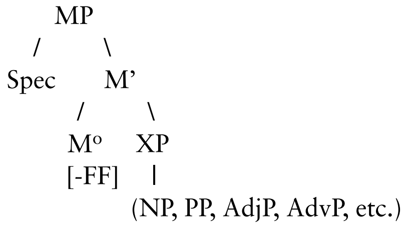


This alternative analysis, based on Bahloul (1994), agrees in principle with the feature-sensitivity spirit advocated in Fassi-Fehri (1993) and Benmamoun (2000). It differs substantially from both in that it relies on the selectional properties of functional categories, and in particular the Modal Phrase (MP). The head of the Modal Phrase (M°), selects a VP in the presence of such modal features as Tense, Aspect, Modality, and Mood. Functional features need a lexical host, hence a verbal form is selected. In the absence of any functional feature in M°, it selects other categories such as NPs, PPs, AdjPs, and AdvPs. A Functional Feature (FF) may therefore be postulated as a triggering mechanism, whereby a positive value specification ([+FF]) yields a VP selection while a negative value specification ([-FF]) results in the selection of other nonverbal categories. This generalization is represented in (23a) and (23b).

(23) a.

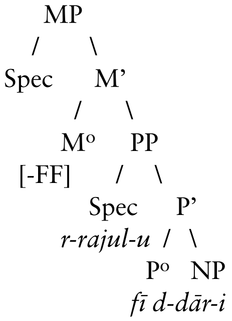


(23) b.

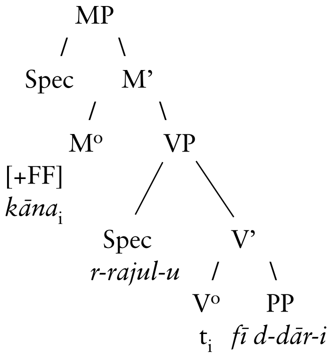


Under this analysis, the original sentences in (12a) and (12b) would have the representations in (24a) and (24b):

(24) a.



(24) b.



3. Conclusion

This entry focuses on copular sentences in Arabic and the contextual conditions which trigger the presence and/or the absence of a copular verb. Syntagmatic properties of copular sentences are first highlighted by showing various contexts under which the copula is not used. A number of approaches to the analysis of copular sentences are then discussed and an analysis of Standard Arabic copular is presented that takes into account a wide range of contextual triggers such as tense, aspect, modal, and mood markers. The analysis advocates a model which places modality at the center of the sentence structure and derives Arabic root clauses, copular and non-copular, on the basis of types of features in M°.

Bibliographical references

Assahārīn, Muḥammad A. 2003. *at-Taṭbīqāt al-luġawiyya*. Abu-Dhabi: Ben Mansour Publishing.
Abdul-Ghany, Mohammed. 1981. *Government binding in Classical Arabic*. Ph.D. diss., Massachusetts Institute of Technology.
Al-Waer, Mazen. 1987. *Toward a modern and realistic sentential theory of basic structures in standard Arabic*. Damascus: Tlass Publishing House.

Bahloul, Maher. 1993. "The copula in Modern Standard Arabic". *Perspectives on Arabic linguistics*, V, ed. Mushira Eid and Clive Holes, 209–229. Amsterdam and Philadelphia: J. Benjamins.

——. 1994. *The syntax and semantics of taxis, aspect, tense and modality in Standard Arabic*. Ph.D. diss., Department of Modern Languages and Linguistics, Cornell University.

Bakir, Murtadha. 1980. *Aspects of clause structure in Arabic*. Ph.D. diss., Indiana University.

Benmamoun, Elabbas. 2000. *The feature structure of functional categories: A comparative study of Arabic dialects*. New York and Oxford: Oxford University Press.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, Mass.: MIT Press.

Eid, Mushira. 1983. "The copula function of pronouns". *Lingua* 59.197–207.

——. 1991. "Verbless sentences in Arabic and Hebrew". *Perspectives on Arabic linguistics*, III, ed. Bernard Comrie and Mushira Eid, 31–61. Amsterdam and Philadelphia: J. Benjamins.

Fassi Fehri, Abdelkader. 1982. *Linguistique arabe: Forme et interprétation*. Rabat: Faculté des Lettres et Sciences Humaines.

——. 1993. *Issues in the structure of Arabic clauses and words*. Dordrecht: Kluwer Academic Publishers.

Mouchaweh, Lubana. 1986. *De la syntaxe des petites propositions*. Ph.D. diss., University of Paris VIII.

Rapapport, Tova. 1987. *Copula, nominal, and small clauses*. Ph.D. diss., Massachusetts Institute of Technology.

MAHER BAHLOUL
(American University of Sharjah)

Corpus Linguistics

1. INTRODUCTION

Corpus linguistics is a branch of → computational linguistics, but its object of research is restricted to large collections (corpora) of authentic language data. Just like computational linguistics, corpus linguistics shares with computer sciences and electronic engineering an interest in natural language processing. Corpus linguistics tries to profit, as much as possible, from computational means and techniques. On the other hand, computer sciences and electronic engineering are more and more aware of the challenges natural language processing offers and, in their turn, they profit from what corpus or computational linguistics bring forth.

When asked about his views on modern corpus linguistics, Chomsky is said to have replied: "It doesn't exist", while characterizing corpus linguistics as butterfly collecting (Aarts 2000: 6–7). More recently, Kay (2004) fulminated against the dominance of statistical and corpus-based approaches to language processing, while Farghaly (2004) observed that, though necessary, statistical and corpus-based approaches "are not sufficient to address all issues involved in building viable applications in Natural Language Processing". On the contrary, two basic introductory works (Manning and Schüze 1999; Jurafsky and Martin 2000) objectively discussed the possibilities and impossibilities of corpus-based research. One gets the impression that corpus linguistics is an issue and therefore does indeed exist.

Arabic corpus linguistics (Ditters 1990) is the study of linguistic phenomena through large collections of Arabic data. In order to handle large collections of data in a methodologically adequate, comprehensive, and consistent way, the application of computational means and the availability of machine readable corpora are prerequisites for corpus linguistics. There is a strong argument in favor of following a corpus-based approach concerning written Arabic, almost a foreign language to all of its users as it is taught and learned at school: without any native speaker, this variety is nonetheless the unique and universal language for the entire Arab world, used in almost any form of written communication.

As far as 'spoken' Arabic is concerned, the term is ambivalent. First, it may refer to a verbal approximation of the literary variety, used, with a varying degree of proficiency, as a formal language in mosques, other religious places, on radio and television, as *lingua franca* between educated Arabs, and on formal occasions. Second, it may refer to a language variety, within the Arab world, almost exclusively learned in a domestic environment and diversified into numerous Arabic dialects, all with different subcategories, each with their own native speakers. As a consequence, 'Arabic speech' data may belong to the first, to the second or, by language interference, to an 'in-between' category. We reserve the term 'spoken literary' Arabic for the former, and 'spoken' Arabic for the latter category of data. 'Blended' spoken Arabic for the 'in-between' category of data will not be taken into consideration here.

Corpus-based research concerning spoken Arabic is still in its early stages. A coherent and univocal formal script for spoken Arabic data has still to be developed and agreed upon. Once this has been done, such a description should be able to account for numerous different dialects of spoken Arabic, each with its own variants. Corpus-based research concerning spoken literary Arabic is usually reduced to a transcript in its textual equivalent. One could qualify this 'approach' as a fatal mistake or conclude that it is better than nothing.

In what follows, McEnery and Wilson (2001) will be followed as a guideline for the discussion of such issues as the collection of data (in itself not a corpus linguistics activity, but rather a prerequisite); data collections; data (pre)processing; data (pre)processing tools; problems in data collecting; and research objectives.

2. The collection of data

The availability or the collection of linguistic data is a prerequisite for corpus linguistics. It is never its final objective. In most cases, data are gathered in the framework of academic or commercial research projects. In subsidized projects, e.g. DIINAR-MBC (2005), data and results are usually available at no charge or for a nominal fee.

An Arabic text or a speech corpus should be machine-readable. There is a general consensus about this feature of corpus linguistics. Matters such as sampling, representativeness (as far as geographic spread, text varieties, semantic domains, age, gender, and target groups are concerned), the form, the script or the size of an Arabic corpus heavily depend on the objectives of the research (see below, 7).

Proposals for the gathering of a representative, machine readable corpus of Arabic data have been made on numerous occasions (e.g. Ditters and Moussa 1995; Zughoul 1997).

Ways to collect data vary. There are the error-prone ways of keying-in or making an analog recording. On the other hand, the scanning in of text data is nowadays of high quality. Moreover, digital recording and the digitization of analog recording are better than before, but here all depends on the quality of the (pre)processing of encoded pertinent data.

The road to obtain data via publishers and broadcast providers is accessible (e.g. Al-Hayat 2005), but this road is still undermined by copyright problems. The Web is a prime source for collection (e.g. Arab Media 2005), apart from the by now almost solved problem of copyrights on web-provided data, given the fact that Web sources deliver materials free or with restricted log-in (see below, 4).

The Arab world itself is eager to make available, via Web-based sources, their sociocultural, scientific, historical, and religious fundamental texts. The textual data of *Qurʾān*, *Ḥadīṯ* and *Tafsīr*, Classical and contemporary Arabic poetry (Arabic Poetry 2005), Arabic literature (e.g. AlAdeeb 2005), as well as many other data are available in character mode via the → Internet.

Arabic text and speech corpora are available from the Linguistic Data Consortium (LDC 2005), the European Language Resources Association (ELRA 2005a), as well as via other sources (see also below, 8). Means used in, and results from, subsidized research projects are usually freely available. For example, in a project (DIINAR-MBC 2005) sponsored by the European Union, a toolkit for the processing of Arabic text data has become available. For research purposes, the IRSIT and the KUN-corpora of electronic textual data (content details can be found in DIINAR-MBC 2005) have also become available, while the impressive data collection of another partner in the project (IERA) had to be kept out of the deliverables, due to copyright issues (see below, 4).

3. Data collections

One of the first electronic collections of Arabic text data is stored in the Oxford Text Archive (Ota 2005). This database of electronic text data comprises transcribed files of Badīʿ az-Zamān al-Hamaḏānī's *Maqāmāt*, early Arabic epistles, and a corpus of modern Arabic prose samples. For more details about this early collection, as well as for information about other, at the time electronically available, Arabic text corpora see Ditters and Moussa (1995).

Many researchers, both within and outside the Arab world, at the academic level or in the commercial sector, use privately collected and electronically stored authentic Arabic data collections for distinct research purposes. The University of Nijmegen houses a collection of digitized text, speech (mostly Moroccan dialects), and digitized video (geographically spread over the Arab world in the form of news and background actualities), gathered for research and teaching purposes. At other centers, similar but as yet unregistered collections are being used.

Some institutions and individuals are active in centralizing information about electronic Arabic corpora and act as providers. Nikkhou and Choukri (ELRA 2005b:36–37) present periodically updated information about available Arabic text and speech corpora together with a short description and mention of the provider. The Linguistic Data Consortium (LDC 2005) houses Arabic raw data collections as well as annotated Arabic corpora (see below, 7). Al-Sulaiti (2005a) includes a link to a listing of existing Arabic linguistic data collections. Moreover, in al-Sulaiti (2005b), everyone is invited to share information about available data with others, by means of a questionnaire.

As far as bilingual and multilingual corpora are concerned, research on a Spanish–Arabic parallel corpus has been commented upon in Samy, Sandoval, and Guirao (2004). International institutions, mainly the United Nations at different levels, produce textual documents in Arabic as well as in a variety of other languages, which enables the compilation of bi- and multilingual parallel corpora (including Arabic), differentiated in a large number of semantic domains. In the literary sector, mainly privately owned, electronically stored Arabic source data, with a translated version in parallel form, are being used for research purposes.

As far as speech databases are concerned, Orientel is a recently completed project aiming at the development of speech-driven applications for the Mediterranean and the Middle East (Iskra a.o. 2004). As for speech-corpora, the dialectal Arabic telephone speech corpus (Maamouri a.o. 2004), the Broadcast News Speech corpus (Choukri a.o. 2004), and the (commercially developed) Sakhr speech databases (Ghali 2004) should be mentioned.

4. Problems in data dissemination

Arabic text and speech data, also in digitized form, are liable to copyright and intellectual ownership restrictions. Commercially available corpora are sold with the approval of the author for use under specific conditions. Organizations such as ELRA (2005a) and LDC (2005) provide 'approved' data collections saving the user much laborious effort.

The 'value' of Arabic corpora depends heavily on the correct digitized reproduction of the source material, as well as on a transparent description of the annotation used. In tagged corpora the reliability of the tagging is as important as the verifiability of the tagging with the source material.

With privately gathered collections a still unsolved problematic area of issues is encountered, such as original copyright; the ownership of an enriched data collection; the right to disseminate an enriched corpus; and the area wherein authentic data or enriched material can be used (publications, the classroom, computer laboratories, or on a local or Web-based server, etc.).

On the other hand, there is a general consensus that Arabic data published via public broadcasting systems, radio and television, satellite, cable, the Web, or otherwise, should be regarded as freely available.

5. Data (pre)processing

Electronically available Arabic text or speech data can be stored as is: it is 'raw' data. In contrast, text or speech data may be 'enriched' or 'annotated' in one way or another. However, according to Leech (1993), it should be possible to remove the annotation from the annotated corpus in order to revert to the raw corpus. The annotation scheme should be well documented or in accordance with standards like the Text Encoding Initiative (Ide and Véronis 1995; TEI 2005; El Hachani 1998; Ramzi 2000) in order to allow for identification of the text or speech sample concerned.

A naive form of enrichment consists in the manual, error-prone, introduction of tags, only to get them electronically reproduced in a subsequent search for these labels later. Another matter is the automatic removal of any non-relevant information from the data (page numbers and any other non-linguistic insertions in linguistic text research, stopgaps and background noise in linguistic speech research) or the automatic insertion of labels relevant to specific research objectives. Practices to enrich Arabic data collections include:

5.1 *Parts-of-speech annotation*

Parts-of-speech annotation refers to the assignment of a code to each lexical unit for its place in a given sequence (Freeman 2001). For some languages (in the case of English text data or with the closed corpus of Latin data), the combination of a parts-of-speech tagger, conjugation, and declension paradigms, as well as lexical listings, a useful enrichment of the authentic data can be obtained. With inflection-rich languages such as Arabic (Khoja 2001), or tonal-based languages the development of a parts-of-speech tagger becomes almost as complicated as a formal grammar describing the morphosyntactic structure of 'raw' linguistic data.

5.2 *Lemmatization and stemming*

The lemmatization and stemming of corpus data refer to the reduction of tokens into types, of words into their respective lexemes or into their basic stem (Dichy 2001; Kadri and Nie 2004; Khoja 2005). This is not a trivial issue, since syntactic parsing relies on a reliable description of the phonomorphological and lexical properties of the language involved, in this case Arabic.

5.3 *Parsing*

Parsing refers to the marking of relationships and dependencies of basic morphosyntactic categories at a higher level of linguistic description. The results of the parsing are usually represented in the form of derivation trees, labeled bracketing, or indented sequences, closely resembling the structure of a tree diagram or derivation tree.

Most of the parsing is done by means of context-free phrase structure grammars (ElNaggar 1990). One may distinguish full parsing (i.e., aiming to provide an analysis of the sentence structure that is as detailed as possible) and skeleton parsing (i.e., aiming to provide only the broad lines of the sentence structure, mainly in terms of constituents).

Instead of a type of context-free phrase structure grammar, a combination of grammar formalisms can be used or a different grammar approach may be applied (Ditters 2005), for instance the combination of immediate constituency with dependency (AGFL 2005; Ditters 2001), constraint grammar (Karlsson a.o. 1995), Definite Clause Grammar (Mehdi 1987), the Tree Adjoining grammar (Lahmeur 2004) and the Prague dependency approach (Hajič a.o. 2004).

Other formal models that comply with the implementation condition (i.e., that a compiler should be available) are (head-driven) generalized phrase structure grammar (Achit and Azzoune 2004), as well as a number of finite-state algorithms (Beesley 2001; Beesley a.o. 1989; Kiraz 1998, 2001).

5.4 *Relationships and dependencies*

The marking of relationships and dependencies between the elements of a constituent or between constituents is one form of semantic annotation (Haddad and Yaseen 2001, 2003). Another form is the marking of semantic features of words in a text. In information retrieval the extraction of head-modifier pairs is an adequate approach. Using this approach on 'raw' data may be the 'best' example of a 'good practice' (Ditters and Koster 2004). At any rate, corpus linguists should for all purposes formulate their questions to data in such a way that 'raw' data can be accessed.

5.5 *Discourse and text linguistic annotation*

Discourse tags frequently occur in the annotation of speech corpora. For Arabic, we are not yet at the level of coherent and consistent discourse marking attained in English (Stenström 1984). For annotated Arabic text corpora a form of marking of colloquial interferences could be meaningful.

Anaphoric resolution is one of the most problematic issues in the automatic parsing of corpora (Chalabi 2004). Therefore, in tagged corpora anaphoric annotation (the marking of pronoun reference) is frequently used. In the parsing of 'raw' data this tricky problem has to be solved at a descriptive level going beyond sentence boundaries.

5.6 *Phonological and phonetic annotation*

During the early stages of the exploitation of computational means, Arabic textual data have been represented in different varieties of scarcely annotated and often inconsistent transliteration and transcription adaptations (OTA 2005). A 'one-to-one' representation of graphemes is, nowadays, a universally accepted standard, except for some, even electronically available, encyclopedic works such as the *Encyclopaedia of Islam*, serial works (*Perspectives on Arabic Linguistics*), periodicals (*Arabica*), and other peer-reviewed publications (Nikkhou 2003;

Krauwer 2004) on Arabic or Arab world related matters. By now, for almost all platforms, a universally accepted and adequate encoding format (Unicode UTF-8, 2005), both for the reproduction of textual data in Arabic characters or in a coherent transcription format, and for the phonetic representation of spoken Arabic data, even with stress marking, is available.

5.7 *Prosody*

Prosody goes beyond the level of segmental differentiation into simple phonemes and concerns the marking of sounds and pertinent features at higher levels of linguistic description, such as stress, intonation, rhythm, meter, and rhyme. In the Arabic domain, there are studies on prosodics (McCarthy 1981; McCarthy and Prince 1990, 1996; Safa, Abdel Nour, and Rajouani 2001). Other studies are related to acoustic analysis and modeling (Mawhoub 2004; Bayeh a.o. 2004). However, researchers in this field could still learn substantially from existing descriptive frameworks for prosody such as O'Connor and Arnold (1961).

5.8 *Problem-oriented tagging*

Problem-oriented tagging (Haan 1984) is the insertion of markers in an annotated or 'raw' data collection for facilitating the individual objectives of a researcher. It is evident that this form of annotation might lack consistency, comprehensiveness, as well as usefulness for other users of the material. However, this form of annotation certainly makes sense. A global search for quotation marks in a machine-readable Arabic text corpus and their subsequent marking could, for example, pinpoint colloquial interferences.

5.9 *Multi-level annotation*

Multi-level annotation is a generic term for the marking of elements belonging to the different levels of linguistic description, for example morphosyntactic description (Khoja a.o., *forthcoming*) and the multi-level annotation as applied by the Prague school (Smrž a.o. 2002; Smrž and Pajas 2004).

6. DATA (PRE)PROCESSING TOOLS

Concordancers (e.g. MonoConc Pro 2005; Collocate 2005; Paraconc 2005), global textual analyzers and frequency software (e.g. WordSmith and MicroConcord), and many other tools are found at Textanalysis (2005); lemmatizers (Stevens 2005), machine translation systems, parsers (analyzers as well as generators for different language levels) (Beesley 2001; Buckwalter 2002; Habash 2004), speech-to-text converters, stemmers, (parts-of-speech) taggers, text-to-speech converters, tokenizers, and other tools for the processing of machine-readable Arabic text and speech data are available or being developed. Al-Sulaiti (2004) and Khoja (2005) provide more information about Arabic corpora and analyzing tools (→ computational linguistics). An advanced Web search will provide up-to-date information (e.g. Buckwalter 2005)

7. RESEARCH OBJECTIVES

For diachronic and/or synchronic lexical purposes a saturation point, concerning the size of a corpus, will never be reached (Hoogland 2003). For syntactic research, a sample of 20,000 words is an adequate starting point, since the descriptive power of a formal grammar should be tested on another sample in a cyclic process of testing and correction. Research on individual authors can be upgraded according to age level and completed with the analysis and final appreciation of their collected works.

Studies on text or speech varieties should be based on a broad range of (identifiable vs. anonymous) authors and speakers, including as many genres and semantic domains of the topic concerned as possible, with mention of the geographical origin of the actors, age, sex of the source and the target group, as well as information about the geographic spread of the target group of receivers. Applications using learning algorithms will start with a few text lines or recorded utterances, being automatically increased during the learning process.

In summary, for the analysis of linguistic data in order to obtain verified results, for the testing of linguistic hypotheses on authentic data in order to arrive at a verified theory, for the construction of particular data collections (electronically available), and for synchronic or diachronic, general or specific domain wordlists and dictionaries, the use of corpora is essential. Moreover, for the development of learning algorithms, tools for information retrieval, and text-to-speech and speech-to-text converters, linguistic knowledge and machine-readable corpora are required.

8. Resources

LDC (2005), ELRA (2005a), some academic institutions (among many others the universities of Lyon-2 and Nijmegen: DIINAR-MBC 2005), as well as individual researchers (Khoja 2005; Al-Sulaiti 2005), are intermediaries in the exchange of 'raw' or (partially and/or fully) annotated collections of authentic, machine readable, Arabic text or speech data.

Usually, the exploitation of available corpus linguistics resources depends on in-house skills, research and/or teaching objectives, available means, as well as long-term objectives. The LDC-approach (Maamouri a.o. 2003–2004; Maamouri a.o. 2004; Maamouri a.o. 2004) remains close to a stable, linguistically 'safety first' strategy. In the meantime, a 'faction' (the Prague School), eager to exploit available data for the testing out of new theoretical concepts in combination with more advanced technological means, has become active (Zemánek 2001; Smrž a.o. 2002; Smrž and Pajas 2004; Hajič a.o. 2004; Smrž 2005). Finally, it is worth mentioning the development of speech-driven applications for the Mediterranean and the Middle East (Iskra a.o. 2004).

Bibliographical references

Aarts, Bas. 2000. "Corpus linguistics, Chomsky and fuzzy tree fragments". Mair and Hundt (2000: 5–14).

Achit, Abdelmadjid and Hamid Azzoune. 2004. "Comparative survey between the GPSG and HPSG grammars applied to the Arabic language". Abstract available at <http://www.ulcl.leidenuniv.nl/index.php3?m=3& c=96>.

AGFL. 2005. <http://www.cs.ru.nl/agfl/>.

AlAdeeb, 2005. <http://www.aladeeb.com/>.

Al-Hayat. 2005. <http://www.alhayat.com>.

Al-Sulaiti, Latifa. 2004. "Arabic corpora and corpus analysis tools (tutorial)". Boualem and Chenfour (2004:547–552).

——. 2005a. <http://www.comp.leeds.ac.uk/latifa>.

——. 2005b. <http://www.personal.leeds.ac.uk/~scsla>.

——, and Eric Atwell. 2003. *The design of a corpus of contemporary Arabic (CCA)*. Leeds: School of Computing, Research Report Series, University of Leeds.

Arabic Poetry. 2005. <http://www.arabicpoems.com>.

Bayeh, Rania, S. Lin, Gerard Chollet, and Chafic Mokbel. 2004. "Reconnaissance de l'arabe parlé à partir de modèles acoustiques du français". Boualem and Chenfour (2004).

Beesley, Kenneth R. 2001. "Finite-state morphological analysis and generation of Arabic at Xerox Research: Status and plans in 2001". Krauwer (2001:1–8).

——. Tim Buckwalter, and Stuart Newton. 1989. "Two-level finite-state analysis of Arabic morphology". Ubaydli (1989:63–72).

Boualem, Malek and Noureddine Chenfour (eds.). 2004. *Compte rendu des Journées d'étude sur la parole (JEP) et Traitement automatique des langues naturelles (TALN) (session arabe), les 19–22 avril, 2004*. Fes: Sidi Mohamed Ben Abdellah University. <http://www.lpl. univ-aix.fr/jep-taln04/proceed/actes/arabe.htm>.

Buckwalter, Tim. 2002. *Buckwalter Arabic Morphological Analyzer Version 1.0*. Linguistic Data Consortium, catalog number: LDC2002L49 and ISBN 1-58563-257-0.

——. 2005. <http://www.qamus.org/>.

Cavalli-Sforza, Violetta, Alon Lavie, and Nizar Habash (eds.). 2003. *The MT Summit 9th Workshop on Machine Translation for Semitic Languages: Issues and Approaches, Tuesday September 23, 2003*. New Orleans.

Chalabi, Achraf. 2004. "Elliptic personal pronoun and MT in Arabic". Boualem and Chenfour (2004).

Choukri, Khalid, Mahtab Nikkhou, and Niklas Paulsson. 2004. "Net-DC BNSC, an Arabic broadcast news speech corpus". Nikkhou (2004:61–64).

COLING. 2004. <http://www.iisco.univge.ch/coling2004>.

COOLOCATE. 2005. <http://www.athel.com>.

Dichy, Joseph. 2001. "On lemmatization in Arabic: A formal definition of the Arabic entries of multilingual lexical databases". Krauwer (2001:23–30).

DIINAR-MBC. 2005. *DIctionnaire INformatisé de l'ARabe, Multilingue et Basé sur Corpus*. CD-rom (INCO-DC no 961791), ed. Joseph Dichy. Lyon: University of Lyon-2. <http://site.univ-lyon2.fr/langues_ promodiinar/accueil.htm>.

Ditters, Everhard. 1990. "Arabic corpus linguistics in past and present". Versteegh and Carter (1990: 129–141).

——. 2001. "A formal grammar for the description of sentence structure in modern standard Arabic". Krauwer (2001:31–37).

——, and Ali Helmy Moussa. 1995. "The compilation of a corpus of modern standard Arabic: The CoCoMSA-project". Ditters (1995:115–132).

—— (ed.). 1995. *Processing Arabic*. Report 8. Nijmegen: Institute for the Languages and Cultures of the Middle East.

——, and Kees Koster. 2004. "Transducing Arabic text to head/modifier pairs". Nikkhou (2004:148–154).

El Hachani, Mabrouka. 1998. "Text encoding initiative: Note bibliographique". DIINAR-MBC (2005).

ElNaggar, Ayman. 1990. "A phrase structure grammar of the Arabic language". Karlgren (1990: 342–344). <http://acl.ldc.upenn.edu//C/C90/C90-3067.pdf>.

ELRA. 2005a. European Language Resources Association. <http://www.elra.info/>.

——. 2005b. "Survey on Arabic language resources and tools in the Mediterranean countries". *Updated overview of Arabic language resources*, ed. Mahtab Nikkhou and Khalid Choukri. <http://www.nem-

lar.org/Publications/NEMLAR-REPORT-SURVEY-FINAL_web.pdf>.

Farghaly, Ali. 2004. “Computer processing of Arabic script-based languages: Current state and future directions”. Farghaly and Megerdoomian (2004:1).

——, and Karine Megerdoomian (eds.). 2004. *Proceedings of the Workshop on Computational Approaches to Arabic Script-based Languages, August 28, 2004*. Geneva: University of Geneva.

Freeman, Andrew. 2001. “Brill’s POS tagger and a morphology parser for Arabic”. Krauwer (2001: 148–154).

Haan, Pieter de. 1984. “Problem-oriented tagging of English corpus data”. *Corpus linguistics*, ed. Jan Aarts and Willem Meijs, 123–140. Amsterdam: Rodopi.

Habash, Nizar. 2004. “Large scale lexeme based Arabic morphological generation”. Boalem and Chenfour (2004).

Haddad, Bassam and Mustafa Yaseen. 2001. “Towards understanding Arabic: A logical approach for semantic representation”. Krauwer (2001: 44–50).

——, and Mustafa Yaseen. 2003. “Towards semantic composition of Arabic: A l-DRT based approach”. Cavalli-Sforza, Lavie, and Habash (2003). <http://www-2.cs.cmu.edu/~alavie/Sem-MT-wshp/Haddad+Yaseen_paper.pdf>

Hajič, Jan. 2002. “Tectogrammatical representation: Towards a minimal transfer in machine translation”. *Proceedings of the 6th International Workshop on Tree Adjoining Grammar and Related Frameworks*, 101–110. Venice: University of Venice.

——, Otakar Smrž, Petr Zemánek, Jan Šnaidauf, and Emanual Beška. 2004. “Prague Arabic Dependency Treebank: Development in data and tools”. Nikkhou (2004:110–117).

Hoogland, Jan, Kees Versteegh, and Manfred Woidich (eds.). 2003. *Woordenboek Arabisch-Nederlands/Nederlands–Arabisch*. Amsterdam: Bulaaq.

Ide, Nancy and Jean Véronis. 1995. *The text encoding initiative: Background and context*. Dordrecht: Kluwer Academic Publishers.

Iskra, Dorota, Khalid Choukri, Ossama Emam, Albino Nogueiras, Oren Gedge, Rainer Siemund, and Robrecht Comeyne. 2004. “OrienTel: More than another SpeechDat for the Mediterranean and Middle East”. Nikkhou (2004:50–54).

Jurafsky, Daniel and James H. Martin. 2000. *Speech and language processing*. Upper Saddle River, N.J.: Prentice Hall.

Karlgren, Hans (ed.). 1990. *Proceedings of the 13th International Conference on Computational Linguistics, August 20–25, 1990*. Helsinki: University of Helsinki. <http://www.essex.ac.uk./linguistics/research/clgroup/ aclbib/coling90.html>.

Karlsson, F., A. Voutilainen, J. Heikkilä, and A. Anttila (eds). 1995. *Constraint grammar: A language-independent system for parsing unrestricted text*. Berlin: Mouton de Gruyter.

Kay, Martin. 2004. “Arabic script-based languages deserve to be studied linguistically”. Farghaly and Megerdoomian (2004:42).

Khoja, Shereen. 2001. “APT: Arabic parts-of-speech tagger”. *Proceedings of the Student Workshop at the 2nd Meeting of the North American Chapter of the Association for Computational Linguistics (NAACL2001), June 2001*. Pittsburgh, Penn.: Carnegie Mellon University. (Available at Khoja 2005.)

——. 2005. <http://zeus.cs.pacific.edu/shereen/ research. html>. [Also an older url: <http:// www .comp. lancs.ac.uk/computing/users/khoja/research.htm>.]

——, and Roger Garside. 2001. “Automatic tagging of an Arabic corpus using APT”. Paper presented at the Arabic Linguistic Symposium (ALS), March 2001. Salt Lake City, Utah: University of Utah.

——, Roger Garside, and Gerry Knowles. Forthcoming. “A tagset for the morpho-syntactic tagging of Arabic”. Paper presented at Corpus Linguistics 2001, Lancaster University, UK, March 2001, to appear in: *A rainbow of corpora: Corpus linguistics and the languages of the world*, ed. Andrew Wilson, Paul Rayson, and Tony McEnery. Munich: Lincom-Europa. (Available at Khoja 2005.)

Kiraz, George. 1998. “Arabic computational morphology in the West”. Ubaydli (1998:2.2.1–11). (Also available at <www.bell-labs.com/project/tts/gkiraz.html>.)

——. 2001. *Computational nonlinear morphology with emphasis on Semitic languages*. Cambridge: Cambridge University Press.

Krauwer, Steven (ed.). 2001. *Proceedings of the Workshop on Arabic Language Processing: Status and Prospects: 39th Annual Meeting of the Association for Computational Linguistics (ACL-01) and 10th Conference of the European Chapter, July 6, 2001*. Toulouse: University of Social Sciences.

Lahmeur, Karim. 2004. “Parsing Arabic texts with LTAG”. Abstract available at <http://www.ulcl.leiden-univ.nl/index.php3?m=3& c=96>.

LDC. 2005. *Linguistic Data Consortium*. <http://www. ldc.upenn.edu/>.

Leech, Geoffrey. 1993. “Corpus annotation schemes”. *Literary and Linguistic Computing* 8:4.275–281.

Maamouri, Mohamed, Anne Bies, Tim Buckwalter, and Hubert Jin. 2003–2004. *Arabic Treebank*. Linguistic Data Consortium, catalog numbers: LDC2003T06, LDC2004T02, LDC2004T11, LDC2004E71.

——, Tim Buckwalter, and Christopher Cieri. 2004. “Dialectal Arabic telephone speech corpus: Principles, tool design and transcription conventions”. Nikkhou (2004:55–60).

——, Ann Bies, Tim Buckwalter, and Wigdan Mekki. 2004. “The Penn Arabic treebank: Building a large-scale annotated Arabic corpus”. Nikkhou (2004: 102–109).

Mair, Christian and Marianne Hundt (eds.). 2000. *Corpus linguistics and linguistics theory: Papers from the 20th International Conference on English Language Research on Computerized Corpora (ICAME 20), Freiburg im Breisgau 1999*. Amsterdam and Atlanta: Rodopi.

Manning, Christopher D. and Hinrich Schütze. 2000. *Foundations of statistical natural language processing*. Cambridge, Mass.: MIT Press.

Marcu, Daniel, Alex Fraser, William Wong, and Kevin Knight. 2004. “Language Weaver Arabic → English MT”. COLING (2004).

Mawhoub, Mourad. 2004. “Prosodie et ordre des constituants dans l’énoncé en arabe standard moderne”. Boualem and Chenfour (2004).

McCarthy, John. 1981. "A prosodic theory of nonconcatenative morphology". *Linguistic Inquiry* 12.373–418.

——, and Alan Prince. 1990. "Prosodic morphology and templatic morphology". *Perspectives on Arabic linguistics*, II, ed. Mushira Eid and John McCarthy, 1–54. Amsterdam and Philadelphia: J. Benjamins. [See also McCarthy and Prince 1996.]

——. 1996. "Prosodic Morphology and Templatic Morphology". *The handbook of phonological theory*, ed. J.A. Goldsmith, 318–366. Oxford: Blackwell.

McEnery, Tony and Andrew Wilson. 2001. *Corpus linguistics*. 2nd ed. Edinburgh: Edinburgh University Press.

Mehdi, Saad. 1987. *Computer interpretation of Arabic*. Ph.D. diss., University of Exeter.

MonoConc Pro. 2005. <http://www.athel.com>.

MicroConcord. 2005. <http://www.lexically.net/wordsmith/>.

Nikkhou, Mahtab (ed.). 2004. *NEMLAR International Conference on Arabic Language Resources and Tools, 22nd–23rd September, 2004*. Cairo: Magnet Creative Communications.

O'Connor, Joseph and Gordon Arnold. 1961. *Intonation of colloquial English*. London: Longman. 2nd ed. 1973.

Ota. 2005. *Oxford Text Archive*. <http://ota.ahds.ac.uk/>.

Paraconc. 2005. <http://www.athel.com>.

Ramzi, Abbes. 2000. "Encodage des corpus de textes arabes en conformité à la TEI: Outils et démarche technique". DIINAR-MBC (2005).

Safa, Nasser Eldin, Hanna Abdel Nour, and Abdenbi Rajouani. 2001. "Enhancement of a TTS system for Arabic concatenative synthesis by introducing a prosodic model". Krauwer (2001:97–102).

Samy, Doaa, Antonio Moreno Sandoval, and José M. Guirao. 2004. "An alignment experiment of a Spanish-Arabic parallel corpus". Nikkhou (2004: 85–89).

Smrž, Otakar. 2005. *Functional Arabic morphology: Formal system and implementation*. Ph.D. diss., Charles University, Prague.

——, and Petr Pajas. 2004. "MorphoTrees of Arabic and their annotation in the TrEd environment". Nikkhou (2004:38–41).

——, Jan Žnaidauf, and Petr Zemánek. 2002. "Prague dependency treebank for Arabic: Multilevel annotation of Arabic corpus". *Proceedings of the International Symposium on the Processing of Arabic, April 18–20, 2002*, 147–155. Manouba: Faculté des Lettres, University of Manouba, Tunisia.

Stenström, Anna-Brita. 1984. "Discourse tags". *Corpus linguistics*, ed. Jan Aarts and Willem Meijs, 65–81. Amsterdam: Rodopi.

Stevens, Vance. 2005. <http://www.geocities.com/vance_ stevens//textanal.htm>.

TEI. 2005. *Text Encoding Initiative*. <www.tei-c.org.uk/>.

Textanalysis. 2005. <http://textanalysis.info/inforet.htm.>

Ubaydli, Ahmad (ed.). 1989. *Proceedings of the Seminar on Bilingual Computing in Arabic and English*. Cambridge: The Literary and Linguistic Computing Centre.

Unicode UTF-8. 2005. <http://www.unicode.org/>.

Versteegh, Kees and Michael Carter (eds.). 1990. *Studies in the history of Arabic grammar*, II. Amsterdam and Philadelphia: J. Benjamins.

WordSmith. 2005. <http://www.lexically.net/wordsmith/>.

Wouden, Ton van der (ed.). 2005. *Proceedings of the Conference on Computational Linguistics in the Netherlands (CLIN 2004), December 17, 2004*. Leiden: University of Leiden. Abstracts available at <http:// www.ulcl.leidenuniv.nl/index.php3?m=3&c=96>.

Zemánek, Petr. 2001. "CLARA (Corpus Linguae Arabicae): An overview". Krauwer (2001:111–112).

Zouaghi, Anis, Mounir Zrigui, and Mohamed Ben Ahmad. 2004. "Une structure sémantique pour la recherche d'information translinguistique anglais–arabe". Boualem and Chenfour (2004).

Zughoul, Muhammad Raji. 1997. "Developing computer based corpora of Arabic: A preliminary proposal". Paper presented at the Conference on Situated Languages, Technology and Communication, held at the Institute of Studies and Research on Arabicization (IERA), Rabat, Morocco, April 1997. <http://members.fortunecity.com/mrzughoul/Developing computer base.htm>.

Everhard Ditters and Jan Hoogland
(University of Nijmegen)

Courtesy Expressions → Politeness; Greetings

Creole Arabic

1. Arabic-based creoles

Arabic-based pidgins and creoles have two profiles. The better-attested one consists of a range of varieties, more or less closely related historically, spoken, or once spoken, in the central and east African countries of the Sudan, Uganda, Kenya, and Chad. In this area the varieties are no older than about 150 years. In true pidgin/creole fashion, the varieties emerged within a short period of time and have developed into a language not mutually intelligible with any other variety of Arabic and having a radically different grammatical structure. The second group consists of fragmented attestations of pidgin varieties. These include an enticing text from the 11th-century geographer al-Bakrī and a short report on a pidgin variety in the contemporary Persian Gulf (→ pidginization).

2. HISTORY

The emergence of the pidgins/creoles in the Sudan is not so much a consequence of events in the 19th century as an embodiment of its history. Its medial origins begin in 1820 when Muhammad Ali, the Khedive of Egypt, started his conquest of the Sudan. Egyptian control of the Sudan spread steadily, and in 1849 a Major Salim led the first expedition, a trading expedition, down the Nile River into the southern Sudan. Thereafter, a floodgate of trading expeditions made their way through the great Sud swamp during its high water period, and in 1854 the first permanent trading settlement was established near present-day Wau. Ivory was the lucrative product of the southern Sudan, and within a decade the region was crisscrossed with trading camps, the German explorer and scientist Schweinfurth reporting that by 1869 camps were located at a distance of every 18–21 miles, connected by a good transportation system.

The establishment of the camps changed the social relationships in the south decisively, and it is here that the origins of the Arabic pidgin/creole varieties are to be found. Certainly, the advent of Muhammad Ali's army in the northern Sudan and its heavy reliance on non-Arabic speakers, especially Nubian and slave recruits, would have led to the spread of second language interference (Wellens 2003:13). However, the classic conditions for the development, spread, and stabilization of a pidgin/creole variety became established only in the southern Sudan.

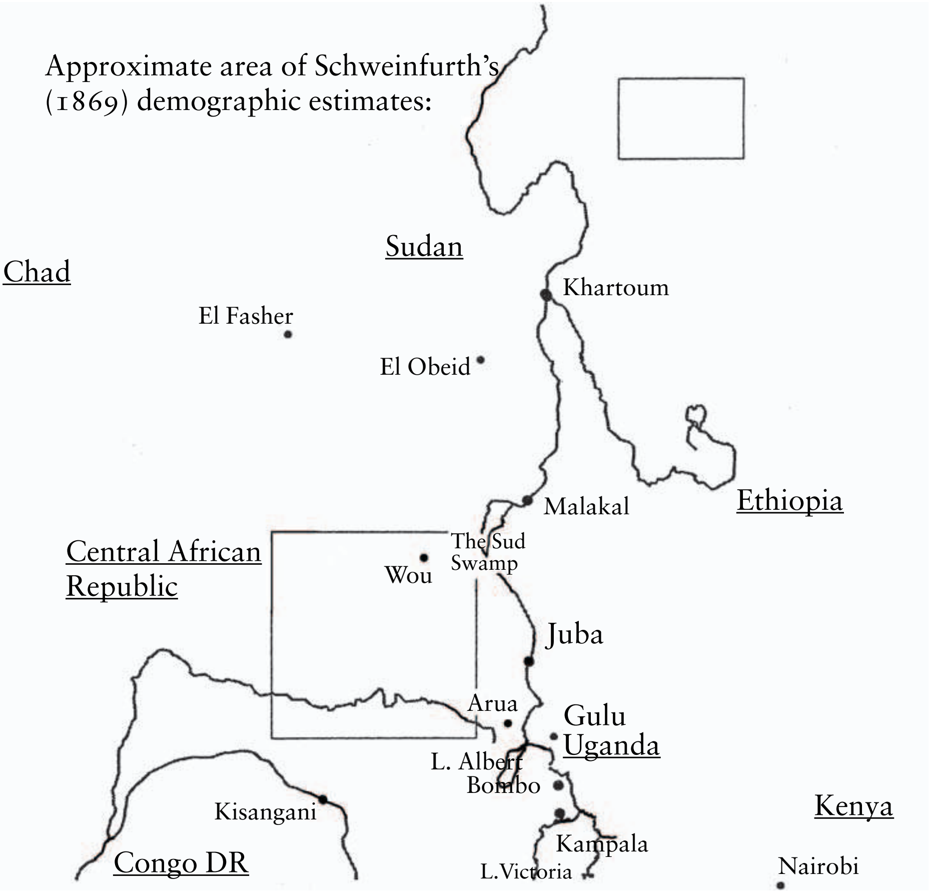
The southern Sudan and the adjoining areas in northern Uganda, eastern Central African Republic and the Congo Republic is, and was in the second half of the 19th century, a highly multilingual region with groups speaking various languages mainly from the central (e.g. Mamvu, Mangbetu, Moru-Madi, Bongo Lendu) and eastern Sudanic (Bari/Kakwa, Lotuko, Dinka, Nuer, Shilluk) language families.

The abrupt opening of the southern Sudan presented a communicative challenge. The trading camps were established throughout the southern Sudan and adjoining regions (see Map 1) and they were linked by a dense communications network. Schweinfurth's statistics (Mahmud 1983; Owens 1990:222) estimated that as early as 1869 nearly a fifth of the entire population, some 60,000 individuals, was resident in the camps. Of these, at most only a quarter were native Arabic speakers from the northern Sudan, Egypt, and elsewhere (officials, itinerant traders from Kordofan/Darfur, soldiers), the rest southerners of various ethnic and linguistic origin. A lingua franca was needed. The situation replicates the unbalanced demographics which have been reconstructed for the development of other creole languages: when the dominant lingua franca, in this case an Arabic dialect, spreads among a population where second language (L2) speakers vastly outnumber first language (L1) speakers, pidginization is assured, and creolization, the nativization of the pidgin is possible. Whether creolization occurs depends on the sociopolitical environment.

Briefly, what probably happened in the southern Sudan between 1854 and 1888 is the following. Three broad social classes emerged in the south: the native population not resident in the camps; southerners, and increasingly their offspring, who served in the camps; and the commanding elite of the camps, composed of Arabs, a few Europeans, and Nile Nubians from the northern Sudan and southern Egypt. The camps themselves changed their status by 1878, when the Egyptian government, under pressure to stop slave trading, transformed the armies of the camp traders into government soldiers. From the relatively plentiful eyewitness reports from the era, many of high scholarly and scientific quality, by authors such as Petherick, Baker, Junker, Schweinfurth, Gessi, Casati, Emin Pasha, Stanley, and Jephson the following can be distilled. Tensions existed between all three groups, and indeed continually increased. The trading camps, originally outposts for gathering ivory, became bases for slave raiding, thus alienating the non-camp population. Within the camps there was an explicit divide and rule policy, whereby different southern ethnic groups were mixed as much as possible to prevent one from becoming dominant. When the Mahdist troops appeared in the southern Sudan in 1885, tensions between the two camp groups came to a head, and a number of mutinies by the southern soldiers occurred. Many of the Arabs and Nubians, moreover, sympathized with, or supported the Mahdi outright, generally against the opposition of the southern soldiers.

The crystallization of a stable Arabic pidgin/creole may be sought in the class of southern

Map 1. The Sudan, Chad, East Africa


Approximate area of Schweinfurth's
(1869) demographic estimates:
Sudan
Chad
Khartoum
El Fasher
El Obeid
Malakal
Ethiopia
Central African
Republic
The Sud
Swamp
Wou
Juba
Arua
Gulu
Uganda
L. Albert
Bombo
Kenya
Kisangani
Kampala
Congo DR
L.Victoria
Nairobi

government soldiers and their families. Linguistically, they were highly heterogeneous, and their numerical predominance, under the best of circumstances, would have hindered a rapid acquisition of something close to a normal Arabic dialect. In any case, conditions for such acquisition were deeply unfavorable, given the social and political cleavage within the camps. It may be assumed that a pidgin/creole Arabic developed among this group. The final act, full creolization as a native language, was played out not in the Sudan, but in East Africa. Emin Pasha, governor of Equatoria, had withdrawn with a remnant troop of soldiers into what it is now Uganda. He was rescued there by Stanley, who took Emin and many of the Egyptian officers and their families to the Indian Ocean coast. Many of Emin's followers, however, stayed on in Uganda, were co-opted into the King's African Rifles, and formed the core of the present-day East African Nubi, speaking the language they brought with them from the Sudan.

3. The varieties emerging from the Sudan

Three distinct varieties of pidgin/creole Arabic, tracing their origins to the southern Sudan of the 19th century, can be distinguished.

i. East African Nubi, also known as → Ki-Nubi, is spoken by between 30,000 and 50,000 people (Khamis 1994:51). Khamis estimates 10,000 of these live in Kenya, where Kibera in Nairobi is their greatest area of concentration, the rest in Uganda, Bombo, and Kampala seeing the largest concentration.
ii. → Juba Arabic. This is a variety mutually intelligible with Nubi, spoken in the southern Sudan. It probably originated at the same time as Nubi, preserved among Emin's soldiers who stayed in the southern Sudan and among other elements of the population who had been part of the camp population. Influence via remigration from Uganda cannot

be ruled out, however. Mahmud (1983) reports that it is nativized in the southern Sudan, i.e. a creole, though at the same time it serves as a lingua franca not only in the southern Sudan, but among the southern Sudanese diaspora in general. Its total population of speakers is therefore hard to discern.

iii. Turku. This is a variety described in a single, rather incomplete work (Muraz 1929). It is a pidgin variety spoken in Chad, which Tosco and Owens (1993) suggest was introduced originally into the area by soldiers of Rabeh, who fled the southern Sudan in 1879 after being defeated by Gessi. In present-day → Chad, lingua franca Arabic is spoken as a non-native variety by nearly half of the population, more than 2,400,000 people (Jullien de Pommerol 1997:96), which is three or four times more than the native Arabic speakers in the country. This Arabic encompasses many varieties, some of which qualify as pidgins. As no descriptions exist, it is impossible to gauge whether Turku can be considered an ancestor of present-day L2 Chadian Arabic varieties.

4. EAST AFRICAN NUBI

The most complete and detailed descriptions are those for East African Nubi (Owens 1977, Heine 1982), particularly Wellens (2003) on Ugandan Nubi, and so this variety will serve as basis for the grammatical sketch, only minimally supplemented by observations from fieldwork.

4.1 *Phonology*

4.1.1 Consonants and vowels

p	t	k	
b	d	g	
m	n	ɲ	
f	s	š	h
v	z		
	c (= tʃ)		
	j (= dʒ)		
	l, r		
w		y	

Wellens (2003:38) notes that Standard Arabic phonemes may, correctly or not, be imported via learned pronunciation, *t̲*, *d̲*, *ḥ*, *x*, *q* (e.g. *bahar*~ *baxar* ~ *baḥar* 'sea, lake'). The tap /r/ may also be realized as retroflexed, and in some dialects varies with /l/. Geminates are very rare, *tenna* 'our' (variant of *ˈtena*).

Nubi has a balanced 5 vowel system.

i u
e o
a

4.1.2 Syllables and stress

(Stress is indicated here by an apostrophe before the stressed syllable.)

The most common syllable types are CV, VC (initially only), V and CVC (see 4.3 below).

ˈam-su-ku 'grab, take', *ju-ˈa* 'houses', *ˈa-ki-li* 'food', *li-ˈ fil* 'elephant'.

Additionally -CC occurs in specific sets of clusters (e.g. NC), and CC- initial is similarly attested, but restricted to specific clusters (e.g. Cw).

ˈbwangiri 'cheeks', *ˈskul* 'school', *ˈsems* 'sun'.

Stress is contrastive, both lexically and morphologically

ˈsaba 'seven', *saˈba* 'morning'
ˈbagara 'cow', *bagaˈra* 'cattle'
ˈkasulu 'wash', *kaˈsulu* 'washing', *kasuˈlu* 'washed', *ˈkásúlu* 'to wash'.

Heine's (1982:27) report of contrastive lexical tone for Kenyan Nubi has never been confirmed (by Owens, Khamis, or Wellens or equally Miller or Watson for Juba Arabic); rather, lexical tone contours are always predictable once the stressed syllable has been identified. Wellens (2003:43), did, however, show that a tonal contrast is the basis of the distinction between the gerund vs. infinitive vs. finite base, *kaˈsulu*/*ˈkásúlu*/*ˈkasulu* as in (11) below.

4.1.3 Phonological processes

The most striking phonological process is the tendency to eliminate vowels in fast speech, particularly high vowels. In fast speech, therefore, CVC syllables are quite common.

ˈgezima → *ˈgezma* 'shoe'
ˈasurubu → *ˈasrubu* ~ *ˈasrub* 'drink'
ˈkasulu → *ˈkasul* 'wash'

The tendency is strongest in unstressed, word-final syllables, though Wellens notes that even in the passive form a final stressed *-u* may be deleted after *-m*, *n*, *l*, *f*, *b*, *gi-rakaˈb* 'being cooked'. Owens reports that final high vowels may be realized voiceless, rather than being fully deleted. There is a tendency to avoid monosyllables, so a word like *geru* 'change' will not necessarily undergo *-u* deletion.

Deletion of the final vowel will lead to a syllable realignment which cuts across word boundaries, as in

(*ˈana*) *ˈam-su-ku ˈi-ta* → *ˈamsuk ˈita* → *ˈam-su-ˈki-ta* 'I caught you'

Wellens notes the entire phonological complex associated with vowel deletions is extremely variable. It may not happen, though if it does it is more likely to occur inter alia in the contexts noted above than elsewhere.

Stem-internal vowels show a strong tendency towards a front-back vowel harmony, with the low mid vowel /a/ occurring in both sets: /i, e, a/ or /u, o, a/, *be̱re̱du* 'wash' vs. *asūrūbu* 'drink'. A similar vowel harmony tendency is attested across morpheme, sometimes word boundaries as well. The verb prefix *gi-* or the preposition *fi*, for example, will assimilate to a following /u/, *gu-loˈgo* 'finding', *fuˈdul* 'in the shade'. In Owens (1977) it is represented as a regular phonological rule, though Khamis (p.c.) reports that in Bombo it is more characteristic of older speakers.

4.2 *Nominals*

4.2.1 Inflectional categories

The only inflectional category is number, SG/PL, though the majority of nouns have no morphological plural. Wellens gives 5 broad classes.

i. Stress shift to final syllable: *yoˈwele/yoweˈle* 'boy(s)'
ii. Ablaut: *keˈbir/kuˈbar* 'big [sg./pl]'
iii. Suffixation of various elements(always accompanied by stress shift), *ˈtajir/tajiˈr-in* 'rich person(s)', *ˈseder/sedeˈr-a* 'tree(s)'
iv. Suppletion: *ˈmarya/nusˈwan* 'woman/women'. This may be supplemented by (iii), *nuswaˈn-a*
v. Prefixation: (Bantu loans only), *muˈze/wa-ˈze* 'old man'

4.2.2 Modifying categories

The word *nas* (cf. *aˈnas* 'people') + sg. or pl. noun may act as a collective, mass, or type marker: *ˈnas sedeˈra* 'trees', *ˈnas ˈakil* 'foodstuffs', *ˈnas kaˈmis* 'those with name of Khamis'.

The demonstratives are as follows.

	near	far
SG	*(u)weˈde*, *ˈde*	*ˈnaˈde* ~ *ˈnade*
PL	*dolˈde*, *ˈdole*	

(1) *kaˈlam weˈde*
'this problem'

Wellens observes that when used predicatively, (*u*) *ˈwede* has initial stress.

Wellens suggests that an opposition *wai/waid* vs. *ˈde* marks an indefinite or non-specific vs. definite opposition. This opposition is not syntacticized, however: indefinite nouns may occur unmarked, and definite nouns (i.e., identified either textually or pragmatically) may be unmarked, or marked by *de*.

Adjectives follow the noun, and may agree with it in number. *al*, *ˈali*, *ab*, *ˈabu* before the adjective mark a relationship as habitual.

(2) *aˈjol ab saˈkar . . . ˈma ˈalisi aˈjol ab keˈbir*
'a young person shouldn't insult an old person' (Wellens 2003:78)

The cardinal numbers are as follows: 1–10: *ˈwai*, *tiˈnin*, *taˈlata*, *ˈarba*, *ˈkamsa*, *ˈsita*, *ˈsaba*, *taˈmaniya*, *ˈtisa*, *ˈašara*; 11–19: *iˈdašar*, *itˈnašar*, *talaˈtašar*, etc.; 20–90: *išiˈrin*, *teleˈtin*, etc., 100: *ˈmia*. Digits precede tens, *ˈkamsa u siˈtin* '65'.

A possessor noun follows the possessed and is marked by *ta*.

(3) *kuˈbaya ta ˈplastik*
'a plastic cup' (Wellens 2003:83)

In what have variously been called compounds (Wellens, Heine) and inalienable possessed nouns (Owens), the *ta* will be omitted.

(4) *suˈnun liˈfil*
'teeth elephant' = 'trunk'

ˈkasma ˈbab
'mouth door' = 'doorstep'

Except for *ˈkila* 'each' and *jina* 'diminutive marker', all modifiers follow the noun.

4.2.3 Pronouns

There are two sets of personal pronoun, which may be termed general and possessive.

general

SG	PL
'*ana*	'*ina*
'*ita*	'*itokum/'itakum*
'*uo/'uwo*	'*umon/'umwon*

possessive

SG	PL
ta'yi	'*tena*
'*taki*	'*tokum/'takum*
'*to*	'*toumon/'tomwon*

The general pronouns are used as subjects, direct objects, and objects of prepositions. The reciprocal pronoun is *ba'dum/'badu*. The reflexive repeats the subject pronoun.

(5) '*ana* '*du'g ana* 'I hit myself'.

4.3 *The verb*

The verb is inflected for tense/aspect, imperative plural, passive, and two degrees of non-finiteness.

The tense/aspect markers are the prefixes *bi-* 'future', *gi-* 'present progressive'. A non-inflected verb stem is basically a past tense. In the active voice a nominal or pronominal (see above) subject is obligatory.

(6) '*uwo 'ja*, '*uwo gi-'ja*, '*uwo bi-'ja*
'he came, he is coming, he will come'

'*ina 'kelemu no-umon*, '*ina gi-'kelemu no-'umon*, '*ina bi-'kelemu no-'umon*
'we talked to them, we are talking to them, we will talk to them'.

The sequence *bi-gi* is possible in principle, '*uwo bi-gi-'arija* 'he will be arriving'. The unmarked tense meaning may be overridden by contextual factors.

A passive-like verb is formed by shifting stress to the final syllable, objects remaining in their post-verb position.

(7) *bi-ku'bu 'moyo* 'water will be spilled'

cf.

(8) *ɲere'ku 'de bu-'kubu 'moyo*
'The child will spill the water'

The broader meaning of this construction is that of an unspecified agent, so even intransitive verbs occur in it.

(9) *gi-ari'ja*
PROG-arrive.PAS
'someone is arriving/they are arriving'

The imperative has the same form as the basic stem. In the plural the suffix *-kum* or *-'tokum* is attached.

(10) '*kelemu no-'umon* *gum-kum*
'speak to them!', 'wake-ye up!'

There are two nominal forms. The gerund is formed by placing stress on the penultimate syllable, while an infinitive, first identified by Wellens (2003), has initial stress, and a raised tone on the second syllable.

(11) *a'rija* 'arriving', '*áríja* 'to arrive'

The semantic difference between the two is slight, though the gerund generally expresses the result of an action, the infinitive the process itself. However, there are contexts where only a gerund can occur.

(12) *a'bidu ta ka'lam*
'beginning [GER] of a problem, the beginning of a problem'

(13) *umon 'rua fi 'gaba fi 'gátá*
they go to forest to cutting.IN
la'kata
Fwood
'They went to the forest to cut wood'.

A small set of verbs, including all loans from Swahili, have a fixed stress on a non-initial syllable, e.g. '*ita ni'situ* 'you forgot', *ni'situ* 'someone forgot/forgetting'.

Reduplication conveys the idea of plurality or diffuseness (see also Miller 2003).

(14) *tu'raal 'kub-ku'bu 'uo ma*
soil which pour-pour.PAS it with
'*namna ta dus'man*
means of war
'The soil which was strewn because of war . . .' (Wellens 2003:141)

Affixes are generally marked only once on a reduplicated verb, as in (14) where only the second verb bears the passive stress shift. Wellens also gives examples, however, where both members shift stress.

Nubi has verbs which besides functioning as main verbs serve as auxiliaries, preceding the main verb and conveying various aspectual and modal meanings, e.g.: *'kan* 'anterior marker', *gurwa* 'immediate future', *'rua* 'intention', *'ja* 'inchoative', *'arija* 'to return, again', *'gum* 'to stand up, inchoative', *'gen/'gai* 'to sit, remain, durative', *'agideri* 'to be able', *'abidu* 'to begin', etc. Segmental inflectional markers usually occur only before the first verb, though with specific verbs may occur after the auxiliary (16). Usually, supra-segmental inflections are marked on each verb in a chain (15).

(15) *'kan gu-ru'wa raka'bu*
if PROG-go.PAS cook.PAS
'akili 'ma
food not
'If food will not have been prepared' (Wellens 2003:117)

(16) *'ter 'gen gi-'guna je'de*
bird sit PROG.sing like that
'The bird keeps singing'

kan + future *bi-* forms a counterfactual.

(17) *kan 'ita bi-'kelem n-ena*
was you FUT-say to-us
'You should have told us' (Wellens 2003: 115)

Non-verbal sentences have a number of special tense and mode markers, including *'kun* 'temporary state, presumptive, inchoative', *'kan* 'past', *'fi* 'existential' *'mafi* 'negative existential'

(18) *'ragi de 'kan fi 'be*
man DEM was at home
'The man was at home'

(19) *'asede 'mafi ka'lam*
now exist not problem
'Now there is no problem'.

4.4 *Other word classes*

Common prepositions include: *fi* 'at', *min* 'from', *gi'dam* 'in front', *'wara* 'behind', *'kabla* 'before', *ma* 'with, by means of', *ze ~ je* 'like' and *na* 'to, for'. Adverbs include *'ini* 'here', *(i)'na* 'there', *bo'yi* 'far', *ge'ri* 'near', *m'bari* 'yesterday', *waltum'bari* 'day before yesterday', *'bukra/sa'ba* 'tomorrow', *'asa/'asede* 'now', *na're/'nare* 'today'.

Question words are *mu'nu* 'who', *su'nu* 'what', *mi'ten* 'when', *ke'fin* 'how', *ma'lu* 'why', *ya'tu* 'which', *kam* 'how many', *musu* 'tag marker'. They remain in the position of the noun they replace.

(20) *ita 'endis 'sente ' kam*
you have money how much
'How much money do you have'?

Conjunctions include *ma/wa/u* 'and', *la'kin* 'but', *gal* 'dependent clause', *ke/ke'de* 'subjunctive marker, *'kan* 'if', *ba'kan* 'when', *ladi* 'until', *ja/je* 'as if, as'. The relative marker is *abu* or *al* as in (14). *'ya* is a focus or topic marker.

(21) *'umon ba'kan sa'kari, ba'ba 'de*
they when young father DET
'ja a'yan
became sick
'When they were young, their father became ill'.

(22) *ɲere'ku' 'ya a'dan 'to ma*
child FOC ears his not
'The child's ears are not there'.

Note that the topicalized NP leaves a pronominal trace.

The negative marker is *'ma or 'mafi*, which occurs sentence final.

(23) *'ina gi-'dusman 'mafi*
we PROG-fight not
'We are not fighting'.

5. East African dialects

Although mutually intelligible, there are dialect differences in Nubi. Kenya Nubi, for instance, maintains etymological **š* to a greater degree than does Ugandan, e.g. *šu'nu* 'what', *šemsi* 'sun', (C)VC syllables are tolerated to a greater degree (*ašrubu* 'drink'), and the negative, usually *ma*, occurs before the predicate. Wellens (2003) remarks in a number of places on differences between Bombo Nubi and that of northern Uganda (e.g. Gulu, Arua). Systematic studies are lacking, however.

6. Juba Arabic: Decreolization

Juba Arabic of the Sudan has essentially the same structure as Nubi (see Watson and Ola 1984) and is mutually intelligible with it. However, it differs in two important ways, both of which greatly increase its overall variability relative to its East African sibling. First, because it is a non-codified, non-implemented (Owens 2004) lingua franca with L2 speakers far outnumbering L1, it is susceptible to various influences, which in their totality tend to increase variability in the language. First, variation may derive from the speakers' L1s (Vincent 1986). Secondly, what might be termed general developmental strategies may be applied differently for different populations of speakers. An example here is Miller (2001), who shows that the complementizer *gali* < *gaal* 'to say' has expanded beyond its original use of 'to say' to mark propositional complements of perception and cognition.

(24) *uwo gi-ayinu gali mumkin*
he PROG-see say perhaps
ita kan bineya bata
you if girl bad
'He realizes that perhaps you are a bad girl'.

However, there is considerable variation, Miller suggesting that urban varieties are more systematic in the contexts where *gale* may be used than are rural ones.

Thirdly, variability derives from the decreolizing pressure of the dominant Sudanese Arabic. Mahmud (1979; also Versteegh 1993) describes how the Creole Arabic verbal prefixes (see above, 4.3) tend to get replaced by the imperfect prefixes of the Sudanese Arabic verb. What Mahmud terms a basilectal variant (i.e., close to the variety described above) will see the prefixes replaced by *ba*, *ta-*, *ya-* among speakers closer to Sudanese Arabic.

(25) *zaman nna gi-atanu mrukaka*
past we PROG-grind grinder
'In the past we used to grind on the grinder'

vs.

(26) *rajab jaman aksan madrasa. awlad*
Rajaf PAST best school boys
ya-alimu kweyis
YA-learn well
'Rajaf was the best school. Boys used to learn well'
(Mahmud 1979:210)

Still, it is not clear that a 'classic' post-creole continuum is in place. Looking at Mahmud's data structurally, the regular ablaut distinctions which are the basis of the perfect/imperfect verb forms in Arabic are not attested, for instance.

A study by Abu Manga and Miller (1992) conducted among ethnic non-Arab southern and western Sudanese immigrants to Khartoum would appear to confirm that the post-creole continuum is not the best model for representing the Juba Arabic/colloquial Arabic contact. In a corpus-based study with eleven participants they note (1992:168ff.) that the perfect/imperfect conjugations, including inflectional elements, are for the most part 'correctly' used. There is no scale-like movement from Juba Arabic to colloquial, such as the post-continuum model requires. There is variation in the dominant Arabic dialect, Western Sudanese or Khartoum, but this is a separate variational parameter. With the exception of one speaker, the Juba Arabic verb does not appear to be an alternative for this group, deviations from a colloquial norm potentially being interpretable as L2 acquisition problems.

At the same time, they note that there are a number of features (e.g. lack of gemination, use of definite article) where the southern segment of the study appears to be influenced by Juba Arabic norms. Furthermore, Miller (2002) observes that Juba Arabic continues to be a symbolic identity marker among southerners in the diaspora. Juba Arabic thus continues to exist in Khartoum but perhaps as an alternative language to colloquial Arabic. More research is needed.

7. Historical, comparative, and creole linguistics

The comparative linguistics of creole languages is more complicated than that of 'regular' languages, as the classic comparative method gives only partially satisfying explanations as to why a creole is as it is. Probably 80 percent of the Nubi vocabulary derives from colloquial Arabic, and sound correspondences can be established according to the classic comparative method (Owens 1985; Pasch and Thelwall 1987). Briefly, correspondences, which are all but perfectly regular, include loss of emphasis, pharyngealization, gemination, and vowel length, and *x/ġ* > *k*: *rudu* 'to return' < *rudd-u*, *kasulu* 'to wash' < *ġasalu* or *xasalu*, *teri* 'bird' < *ṭēr*.

Common, but often idiosyncratic changes include loss of final non-nasal obstruents and a marked tendency towards CV syllable structure: *'be* 'house' < *bēt*, *tu'ra* 'earth' < *turāb*, but also *'danab* 'tail' < *danab*, *'aseti* 'lion' < *asad*.

Defining the precise source of the pidgin/creole structures requires considerable care. Note, for instance, that although Standard Arabic *ṯ* always corresponds to Nubi *t*, *ti'nin* 'two', the correspondence *ṯ* > *t* was not listed as a perfectly regular one, since it should be assumed that Nubi inherited only *t*, the change from *ṯ* > *t* having become established throughout the larger Sudanese dialect area centuries before Nubi came into existence. Considering all dialects in the region, correspondences which look like indicating changes may in fact represent simple inheritance. For instance, most dialects of the western Sudanic region (→ Subsaharan Arabic), including some found in the Sudan, lack pharyngeals altogether, and these could have been a source of ancestral Nubi. It was noted that Nubi displays vowel harmony, but vowel harmony is characteristic not only of Western Sudanic Arabic, but also of most of the substrate languages of the southern Sudan.

Yet, comparisons with varieties of Arabic only partly solve the question of where Sudanic pidgin/creole originates. Certainly, it can be said to have massively simplified its morphology vis-à-vis Arabic. Such simplification is known to be typical of pidgins/creoles generally: indeed, McWhorter (1998) argues that it defines a creole type language. However, creoles are also assumed to eschew supra-segmental lexical and grammatical marking, even the Atlantic creoles with their rich substratum of tonal West African languages. Yet, Nubi uses stress shift, in one case coupled with tonal contrast, to distinguish no less than four categories in the verb (basic stem, passive, gerund, infinitive) and the noun (some singulars vs. plural). This is explicable in terms of neither the classic comparative method, since the pidgins/creoles innovate categories not present in colloquial Arabic, nor in terms of common creolization processes (Owens 2001).

Bibliographical references

Heine, Bernd. 1982. *The Nubi language of Kibera: An Arabic creole*. Berlin: Reimer.

Jullien de Pommerol, Patrice. 1997. *L'arabe tchadien: Emergence d'une langue véhiculaire*. Paris: Karthala.

Kaye, Alan S. 1994. "Peripheral Arabic dialectology and Arabic pidgins and creoles". *Actas del Congreso Internacional sobre Interferencias Lingüísticas Árabo-Romances y Paralelos Extra-Iberos*, ed. Jordi Aguadé, Federico Corriente, and Marina Marguán, 125–140. Zaragoza: Navarro e Navarro.

—— and Mauro Tosco. 1993. "Early East African pidgin Arabic". *Sprache und Geschichte in Afrika* 14.269–306.

Khamis, Cornelia. 1994. *Mehrsprachigkeit bei den Nubi*. Hamburg: Lit.

Mahmud, Ushari. 1979. *Variation in the aspectual system of Juba Arabic*. Ph.D. diss., Georgetown University.

——. 1983. *Arabic in the southern Sudan: History and spread of a pidgincreole*. Khartoum: Khartoum University Press.

McWhorter, John. 1998. "Identifying the Creole prototype: Vindicating a typological class". *Language* 74.788–818.

Miller, Catherine. 1983. "Aperçu du système verbal en Juba Arabic". *Comptes Rendus du Groupe Linguistique d'Études Chamito-Sémitiques* 24–28. 292–315. Paris: Geuthner.

——. 1986. "Un exemple d'évolution linguistique: Le cas de la particule 'ge' en 'juba-Arabic'". *Cahiers du MAS-GELLAS* 3.155–166.

——. 1989. "*Kelem kalam bitak*: Langues et tribunaux urbains en Equatoria". *Matériaux Arabes et Sudarabiques* 2.23–58.

——. 1993. "Restructuration morpho-syntaxique en Juba-Arabic et Ki-Nubi: À propos du débat universaux/substrat et superstrat dans les études créoles". *Matériaux Arabes et Sudarabiques* 5.137–174.

——. 2001 "Grammaticalisation du verbe *gale* 'dire' en Juba-Arabic". *Leçons d'Afrique: Filiations, ruptures et reconstitutions des langues: Un hommage à Gabriel Manessy*, ed. Robert Nicolaï, 455–482. Louvain: Peeters.

——. 2002. "Juba Arabic as a way of expressing a southern Sudanese identity in Khartoum". *Contemporary Arabic dialects: Proceedings of the 4th Aida meeting*, ed. A. Youssi a.o., 114–122. Rabat: AMAPATRIL.

——. 2003. "Reduplication in Arabic-based contact langues". *Twice as meaningful: Reduplication in pidgins, creoles and other contact languages*, ed. Silvia Kouwenberg, 289–299. London: Battle-bridge.

—— and Al-Amin Abu Manga. 1992. *Language change and national integration: Rural migrants in Khartoum*. Khartoum: Khartoum University Press.

Muraz, Gaston. 1929. *Vocabulaire du patois arabe tchadien ou 'tourkou' et les dialectes Sara Madjingaye et Sara Mbaye*. Nancy: Charles Lavauzelle.

Nhial, Abdon. 1975. "Ki-Nubi and Juba Arabic: A comparative study". *Directions in Sudanese linguistics and folklore*, ed. Sayyid Ḥurreiz and Herman Bell, 81–93. Khartoum: Khartoum University Press.

Owens, Jonathan. 1977. *Aspects of Nubi syntax*. Ph.D. diss., London University.

——. 1980. "Monogenesis: The universal and the particular in creole studies". *Anthropological Linguistics* 22.97–117.

——. 1985. "The origins of East African Nubi". *Anthropological Linguistics* 27.229–271.

——. 1989. "Zur Pidginisierung und Kreolisierung im Arabischen". *Afrika und Übersee* 72.91–107.

——. 1990. "East African Nubi: Bioprogram vs. inheritance". *Diachronica* 7.217–250.
——. 1991a. "Nubi, genetic linguistics and language classification". *Anthropological Linguistics* 33.1–30.
——. 1991b. "Local and universal sources of a creole verbal construction". *Semitic studies in honor of Wolf Leslau*, ed. Alan Kaye, 1169–1179. Wiesbaden: O. Harrassowitz.
——. 1997. "Arabic-based pidgins and creoles". *Contact languages: A wider perspective*, ed. Sarah Thomason, 125–172. Amsterdam: J. Benjamins.
——. 2001. "Arabic creoles: The orphan of all orphans". *Anthropological Linguistics* 43.348–378.
Pasch, Thelma and Robin Thelwall. 1987. "Losses and innovations in Nubi". *Bochum-Essener Beiträge zur Sprachwandelforschung*, III, ed. Norbert Boretzky, W. Enniger, and Thomas Stolz, 91–165. Bochum: Brockmeyer.
Schweinfurth, Georg. 1918. *Im Herzen von Afrika.* Leipzig: Brockhaus.
Smart, J. 1990. "Pidginization in Gulf Arabic: A first report". *Anthropological Linguistics* 32.83–119.
Smith, Ian and Morris Ama. 1985. *A dictionary of Juba Arabic and English*. Juba: The Committee of the Juba Cheshire Home.
Thomason, Sarah and Alaa Elgibali. 1986. "Before the lingua franca: Pidginized Arabic in the eleventh century". *Lingua* 68.317–349.
Tosco, Mauro. 1995. "A pidgin verbal system: The case of Juba Arabic". *Anthropological Linguistics* 37.423–459.
—— and Jonathan Owens. 1993. "Turku: A descriptive and comparative study". *Sprache und Geschichte in Afrika* 14.177–267.
Versteegh, Kees. 1984. *Pidginization and creolization: The case of Arabic*. Amsterdam: J. Benjamins.
——. 1993. "Leveling in the Sudan: From Arabic creole to Arabic dialect". *International Journal of the Sociology of Language*. 99.65–79.
——. 2004. "Pidginization and creolization revisited: The case of Arabic". *Approaches to Arabic dialects*, ed. Martine Haak, Rudolf de Jong, and Kees Versteegh, 343–358. Leiden: E.J. Brill.
Vincent, George. 1986. "Juba Arabic from a Bari perspective". *Current approaches to African linguistics*, ed. Gerrit Dimendaal, 71–78. Dordrecht: Foris.
Watson, Richard and L. Ola. 1984. *Juba Arabic for beginners*. Juba: Summer Institute of Linguistics.
Wellens, Ineke. 2003. *An Arabic creole in Africa: The Nubi language of Uganda*. Ph.D. diss., University of Nijmegen.

JONATHAN OWENS (University of Bayreuth & College Park Maryland)

Culture and Language

1. INTRODUCTION

How the Arabic language shapes and is in turn shaped by its various cultural contexts is a complex question. This is so for a number of reasons. Although some aspects of culture become encoded in language, we have not yet reached satisfactory understanding regarding the many ways in which language and culture do (and do not) reflect and shape each other. Culture is not only that which can be described positively on the basis of a set of explicit or implicit patterns of behavior, beliefs, and values, but also that which conceals. All cultures socialize their members into treating their own values and behaviors as 'natural' while alternatives are 'hidden' in that very same process. Hence, what is concealed also needs to be brought out. Comparative perspectives are particularly useful in this regard.

Moreover, there is not just one kind of Arabic language and one Arab culture, while there is also a multiplicity of other languages and cultures that interact with the larger 'Arab' cultures. Kurds, Berbers, Nubians, Copts, Jews, Greeks and Armenians (among others) have been living side by side with Muslim Arabs for centuries. Some of these communities predate the arrival of Islam. Finally, many aspects of the social and cultural lives of Arabic remain to be empirically investigated so that in certain areas our knowledge is insufficient (but see Chejne 1969; Doss and Miller 1996).

Given such limitations and the vast scope of the topic addressed here, it is perhaps wise as a first step to acknowledge the constraints and not aim for an exhaustive but superficial treatment (for a review of sociolinguistic and anthropological works in this area, see Haeri 2000). In this brief entry on language and culture, little is said on what has historically been considered the 'High Culture' of Islamic civilization. There already exist ample resources for consultation (Hodgson 1974; Hourani 1991). The aim here is to explore the contemporary Arab world and to examine the dynamics of language use in several domains of everyday life: art, religion, education, and politics. The majority of sources cited here are limited to scholarly works in the English language.

A brief note needs to be made of the terms used in this entry to refer to the Arabic language. The full name of the written language in the Arab world is *al-luġa al-ʿarabiyya al-* → *fuṣḥā* 'the eloquent Arabic language'. → 'Classical Arabic' will be used for this term. Non-Classical varieties of Arabic are called *ʿāmmiyya* 'common'. For the latter, vernacular Arabic instead of the more usual 'colloquial Arabic' will be employed. The modifier → ‘colloquial’ can describe a style within a language, for example ‘colloquial English’. Otherwise it cannot be applied to entire languages. The term → ‘Modern Standard Arabic’ will be discussed in the section on modernization.

2. The sound of words: Religion, poetry, and aesthetics

Among literary genres, poetry occupies a special place in diverse Arab cultures. Aesthetic sensibilities are created and honed by exposure to poetry recited orally on many occasions. Thousand-year-old epic poems continue to be performed (Reynolds 1995). There is a longstanding tradition of transforming some of this poetry into highly popular songs renovated by new generations who still listen to old but unforgotten singers such as Umm Kulthum in Egypt (Danielson 1997) and Nazim il-Ghazali in Iraq. A profound affinity for poetry feeds into and is in turn inspired by oral recitations of the *Qurʾān*, whose language (Classical Arabic) is viewed as beautiful, powerful, inimitable by humans, and as a reservoir of crucial knowledge for Muslims. Non-Arab Muslims hold similar views.

A number of factors have contributed to shared senses of aesthetics among Muslims with regard to the language. There are highly developed and melodic styles of reciting the *Qurʾān* (Nelson 1985; → *tajwīd*) that are heard in the mosque, on the radio, on television, in recitation competitions, and on special occasions in various venues. The evocative call to prayer (*ʾaḏān*), that marks the time of each daily prayer and more generally organizes the passage of time from early morning to night, is broadcast from mosques, radio, and television. Believing Muslims memorize the five daily prayers as well as Qur*r*ʾʾānic *sūras* ‘chapters’. These are learned at home or at (religious) school. Associations are formed from childhood between calendrical rituals, for example, the fasting month of Ramadan, and special *sūra*s. All these practices of listening and reciting create a consciousness of the language and its beauties. This ‘soundscape’ is integral to a sense of community that, depending on what aspects we examine (poetry and songs with or without the religious uses), contribute to an Arab and/or Muslim identity.

One of the most interesting features of the social life of Classical Arabic is that, on the one hand, it has been the prime vehicle of a large body of highly regarded texts central to Islamic civilization in theology, the sciences, linguistics, and literature. On the other hand, it has had a more inclusive social life as an oral/aural language – in prayers, recitations, poetry, and songs. The memorization and performance of these do not require formal schooling. Hence, across the centuries and at present a far larger number of people participate in the oral life of the language than in reading the historical texts or in writing.

The prominence of poetry in the diverse cultures of the Arab world is demonstrated by the fact that for many it is a part of daily life. Among the Awlad ʿAli Bedouin of Egypt, Abu-Lughod (1986) found that *ǧinnāwa*s ‘little songs’ are used by women to voice sentiments and feelings that they otherwise rarely articulate. According to Abu-Lughod, they are “lyric poems, like Japanese Haiku in form”, whose main themes are love, longing, and honor (1986:27). For the tribes of North Yemen, poetry is a “cultural practice” that is “both the creation of art and the production of social and political reality in the same act of composition” (Caton 1990:21). One of the most important uses of poetry in this region is for the settlement of disputes. Caton demonstrates how feuding parties in North Yemen sometimes use poetry to persuade their enemies rather than coercive force: “What Western people might find strange, Khawlāni tribesmen take for granted, namely, that politics and poetics are inseparable” (1990:155; see also Miller 2002).

While poetry, whether in Classical or vernacular Arabic, is an old and integral part of what Arab cultures share, there are more recent genres of prose that have become increasingly important (Cachia 1990). Among these are plays, short stories (Hafez 1993), novels (Allen 1982; al-Musawi 2003), radio and television serials, films (Shafik 1998; Armbrust 1996), and comic strips (Douglas and Malti-Douglas 1994). The emergence of such genres marks artistic milestones in the cultural histories of the Arab world. As Hafez argues, their forms and content may be analyzed without always attributing their origin either to older Classical literary forms or to the influence of the West.

There are a great many female poets and novelists (Zeidan 1995), a majority of whom write either in Classical Arabic or in a foreign language. In fact, there are far fewer female poets who write in vernacular Arabic. Some famous Arab female novelists write in English or French. Among them are Ahdaf Sueif, Hanan al-Sheikh, and Assia Djebar. While in writing, women seem to choose to write in Classical Arabic as often or more often than men, in speech they use Classical Arabic consistently less than men even when they have the same level of education (Walters 1991; Haeri 2000). Just as the use of language is gendered, language used to discuss men and women also shows gendered aspects (Eid 2001).

3. Arabic in the Muslim world

Muslim countries can be divided into two categories, those where the national and official languages are not genealogically related to Arabic, and those in which various Arabic vernaculars, related to but distinct from Classical Arabic, are spoken. In the first category, we have countries such as → Iran, → Turkey, → Pakistan, → Indonesia, → Senegal, and → Nigeria (among others). In the second, we have the entire Arab world. In both kinds of countries, regardless of mother tongue, a believing Muslim must know some Classical Arabic in order to read the *Qur'ān*, perform the daily prayers and carry out other religious rituals and obligations. In this sense, there is no Islam without Classical Arabic. There would also be no Classical Arabic without Islam; the language would not have survived in the way it has merely because it was the language of high poetry in pre-Islamic Arabia. While the Arabic vernaculars spoken as mother tongues are related to Classical Arabic, knowledge of the former does not translate into knowledge of the latter. Classical Arabic is chiefly a written language whose proficiency requires formal schooling; it is no one's mother tongue.

There are also significant differences among Muslim countries with respect to language. Non-Arab Muslim countries use Classical Arabic only for religious purposes. Otherwise, their national standard languages (at times in conjunction with a former colonial language) are employed for all other domains. But in Arab countries, Classical Arabic is also the official language of the state, of public education and almost all that appears in print. In the countries of the Maghreb, French continues to play a central role in education and public and private life more generally. In most other parts of the Arab world, foreign languages – in particular English – are employed for a variety of purposes, including training in the sciences for higher education and private schools.

One other major difference is that while in non-Arab countries the *Qur'ān* has been translated into the local vernaculars, the holy text has never been translated into (and published in) any of the Arabic vernaculars. Bilingual *Qur'ān*s with line by line or page by page translations are prevalent in countries such as Iran, Turkey, and Pakistan. In Arab countries, the question of the translation of the *Qur'ān* remains an extremely controversial issue.

It seems to be received wisdom to point to 'Islam' as the source for the obligation to perform the daily prayers and read the *Qur'ān* exclusively in Classical Arabic. In the same vein, a disembodied 'Islam' is said to be the cause of the absence of translations of the *Qur'ān* in the vernaculars of Muslims across the world. As mentioned, in Muslim countries whose official languages are not related to Arabic, the *Qur'ān* has been translated. In Turkey, the *Qur'ān* was translated into Turkish in the 12th century (Holt a.o. 1970:684). Translations of parts of the *Qur'ān* into Persian are more than a thousand years old and at present there are numerous complete translations of this text in Iran (Khorramshahi 1997). Bilingual editions are ubiquitous and may be found in most homes. New translations continue to be published.

Although the question of translation has been a point of debate and disagreement since the early centuries of Islam, 'Abū Ḥanīfa, who was Persian-speaking and the founder of one of Islam's legal schools, issued a *fatwā* 'opinion' in the 8th century when asked whether it was permissible to perform the prayers in Persian or Turkish (Khorramshahi 1997:619). At the time, the full form of the prayers had not yet been established. He argued that the recitation in translation of the *Sūrat al-Fātiḥa* (the opening chapter of the *Qur'ān*) for those who do not know Arabic but wish to pray is allowed (1997:691). A number of Qur'ānic verses and *ḥadīṯ*s (the sayings and actions of the Prophet) are cited in support of the permissibility of translation. Among them are: "And we have not sent an Apostle except [to teach] in the language of his own people in order to make [things] clear to them" (Q. 14/41; translation into English from Ali 1992; the Persian translation is somewhat different; in English it would read as "And we have not sent a Messenger [with a message] except in the language of his people in order to make [rulings and realities] clear"; Khorramshahi 1997:255). The other verse that is cited is "God does not demand of anyone except according to his abilities" (Q. 2/286).

The *Qurʾān* refers to its language as a 'miracle', challenges readers to create writing that would rival it, and mentions a number of times that its language is 'clear'. Yet, nowhere does it forbid its own translation. The evidence of the achievements of early Muslim scholars writing in Arabic in many fields suggests that for them the language of the *Qurʾān* was pliable, not untouchable. As is well known, when confronted with words or constructions in the *Qurʾān* that they were unsure of, such scholars sought the aid of Bedouin who spoke the same dialect as the Prophet in order to settle doubts: "The early beginnings of grammar and lexicography began at a time when Bedouin informants were still around and could be consulted. There can be no doubt that the grammarians and lexicographers regarded the Bedouin as the true speakers (*fuṣaḥāʾ*) of Arabic" (Versteegh 1997:63). This method of verification suggests that scholars' conception of the language of the *Qurʾān* was not one of absolute fixity and rigidity and hence it did not preclude asking humans for clarification. It remains for historians to tell us when, and for what reasons, a profound transformation with regard to the translatability and hence negotiability of the language of the *Qurʾān* and of the prayers took place. Versteegh (1997:63) explains that after the 10th century, the tradition of using such informants ceased: "Since there were no longer living informants to provide fresh information, the corpus of the language was closed, and 'fieldwork' could no longer produce reliable results". Perhaps this closing of the corpus and the lack of human input is in part responsible for language ideologies emerging in the following centuries that underline fixity of the language of the *Qurʾān* and reject the possibility of translation.

It seems more tenable to argue that exclusivity with regard to the language of Islam is more a result of the fact that this was the language of an empire whose ruling classes had an interest in not sharing power with the new converts (see also Barakat 1993). If the converts could pray and read the *Qurʾān* in their own languages, less privileged status would accrue to Arabic. Instead, by keeping Classical Arabic as the only permissible language of Islam and its rituals, a hierarchy developed, which is maintained to the present: the center of 'true Islam' is Arab, and others are at various points on its periphery. This discussion brings us to an abiding incongruence with regard to Classical Arabic that continues to affect the language situation. On the one hand, Classical Arabic was the language of a highly successful empire that made vast conquests, one whose scholars produced knowledge in many fields. It furthermore founded many schools of art in calligraphy, architecture, textiles, and so on. On the other hand, many parts of the empire fell under colonial rule and the empire was defeated. Its language came to belong to subjects of European colonial powers. The two pillars of that empire that survived were religion and language. But there had to be far more struggle to keep the language alive than the religion – due to policies under colonial rule, lack of adequate educational institutions to teach the language, the distance between the language and the mother tongues of Arabs, and unfavorable economic and trade relations with Europe. From this point on, the language seems to have become far more politicized than it had ever been as it also came to symbolize resistance to foreign domination. While secular and religious Arabs (Muslims, Christians, and Jews) saw in the language a great potential for unification, only religious Muslim Arabs could use Islam in the same way. Hence, Classical Arabic became the primary symbol, more so than religion.

4. Language and nationalism

More has been written on the centrality of Classical Arabic to pan-Arab → nationalism than on any other aspect of the language situation in the Arab world. Historians, political scientists, and linguists in particular have contributed greatly to the prevalence of this topic. Suleiman (2003:2) argues there is a "glaring weakness" due to general "reluctance to take the study of nationalism in the Arab Middle East into the wider cultural arena of literary production, the arts, film, music, sports, tourism, festivals . . . and other media of symbolic expression" (but see Badran 1995; Khoury 2003; Ahmed 1999; Shafik 1998; Starrett 1998; Shryock 1997). Due to lack of space this issue is discussed here in only broad outlines. Suleiman (2003:224), perhaps the most detailed study to date of language and nationalism in the Arab world, finds that "formulations of Arab nationalism, whether embryonic or fully fledged in character, are invariably built around the potential and capacity of Arabic in its standard form to act as the linchpin of the identity of all those who share it as their common language".

The peak of pan-Arab nationalism was reached in the 1950s and 1960s following the coming into power of Gamal Abdel Nasser and his group of Free Officers. Pan-Arab nationalism was a movement in which many non-Muslim Arabs participated. It was expressly aimed at including Arabs of every faith. The language that all Arabs share is Classical Arabic, not the local vernaculars that are different from each other and therefore 'divisive'. Pan-Arabism's explicit efforts in defining an Arab as anyone who speaks 'Arabic' as a native language was meant to remove religion and race as bases of an Arab identity. Non-Muslim Arabs welcomed this shift and seem to have contributed to its emergence as well. However, as a sociopolitical force that might potentially demand the erasure of local differences, it was resisted by some leftist groups who nevertheless were in agreement with its other principles.

In its main thrust, pan-Arabism was secular and promised a search for alternative and authentic forms of modernity. At the same time, the movement made Classical Arabic its central vehicle and symbol for the staging of this promise. There seems to have been a belief that the historical association between the language and Islam would in time fade away. That belief continues to the present. Suleiman (2003:229) mentions 'Islamic nationalists' as being opposed both to the secular tendency of pan-Arabism and to attempts to appropriate and secularize the language thereby severing its relations to Islam. However, this is not merely the position of Islamists. In Egypt, most 'ordinary' people – those who are outside the reading-writing elites – show great surprise at the suggestion that modernized versions of Classical Arabic are unrelated to Islam. First, they do not make any fundamental distinction between old and new versions of the language. Second, they respond by asking rhetorically: "so where does the language come from then?" (Haeri 2003). There are many layers of ambiguity in the relation between language and nationalism in the Arab world. Further discussion of this point is found below in the section on "Experiences of education".

5. The modernization and simplification of Classical Arabic

In several Arab countries, a movement to modernize Classical Arabic began in the 19th century. The center of these efforts was Syria/Lebanon and Egypt. New institutions of learning were being founded that were independent of mosque-universities. These institutions were for educating soldiers, training nurses, engineers, and other similar professionals. Along with the appearance of newspapers, recent colonial bureaucracies, and new economic activities and trade relations, the influence of foreign languages increased and new demands began to be put on Classical Arabic for use in these novel domains. It is not unreasonable to suggest that the question of the choice of a language must have come up. Apart from those areas where a colonial language was used, in most others vernacular Arabic could have been chosen over Classical Arabic considering, in particular, the very low rates of proficiency in that language. We lack the necessary historical studies on possible debates on this question at the time, particularly at the level of the state.

What we do know is that the idea of modernizing (*taḥdīṯ*) and simplifying (*tabsīṭ*) the language emerged perhaps as part of the general prevalent desire for reform and progress in these countries. Classical Arabic stood tall against colonial claims of a superior civilization. At the same time, it was recognized that the emerging new domains and functions needed a more supple language with vocabulary for the modern sciences and technology and a less complex grammar than that of old texts. State institutions, publishers, journalists, translators and other writers participated in debating the meanings of *taḥdīṯ* and the steps that could be taken. The literary and political *Nahḍa* movement (late 1800s to early decades of the 20th century) contributed to the revival of the language. The other movement that greatly expanded the scope of modernization processes is pan-Arabism.

Many scholars believe that the resulting linguistic changes are primarily due to the influence of foreign European languages – through the translation of texts from these languages into Arabic and through the education of Arabs in foreign schools abroad (see, e.g., Stetkevych 1974). Yet, both the available statistics and the consideration of other factors seem to undermine the exaggerated importance of translation. The influence of the mother tongues of writers was and continues to be stronger than that of foreign languages in particular with respect to grammar (Haeri 2003). In addition, sociological and technological factors having to do with printing and publishing have had an impact on orthography and on the language more generally (Mahdi 1995). In this regard, there was also the change from a scribal tradition of copying and correcting texts to modern correctors, *muṣaḥḥiḥūn*, who are similar to copy-editors (Mahdi 1995). The change has implications for the language of the texts because, whereas well into the 19th century, scribes were exclusively men of religion, modern correctors include many who have been trained at state universities, a majority of whom are not religious scholars (Haeri 2003). More studies on these aspects of print culture need to be made in order for us better to assess their impact on the language.

At present, the contemporary versions of Classical Arabic, found in newspapers, magazines and books, are on linguistic grounds different from the language of a century ago, though some authors continue to write in older styles (for language change in Gulf countries, see Holes 1988). In some respects the language is closer to the grammar and vocabulary of the vernaculars. Still, research on the details of language use today needs to be taken into account. For example, Parkinson (1991:36) found that:

> There are language columns in Egyptian newspapers and magazines that continually blur the line between classical fusha and modern fusha, on the assumption that any word, form, or structure sanctioned long ago is also fine today . . . Even the Arabic Language Academy, whose mission is to modernize fusha, has published a dictionary recently with an absolutely confusing mixture of archaic, classical, and modern meanings under almost every entry, with no marking whatsoever on which are likely to be understood by modern readers, and which are entirely out of date, *as if the unity of classical and modern fusha was a political imperative to be enforced* [emphasis added]

Therefore, on the one hand, change has surely taken place, but on the other hand, disagreement and controversy over what constitutes legitimate change continue to the present.

With regard to modernization efforts, one question that needs to be posed is: what constitutes the modernity of a language? To make a language modern, that is, to render it a congruent and unaffected medium of the contemporary lives of not just highly educated people but also of the whole community, is not simply a matter of linguistic change. A language is not merely a linguistic system. It is braided and intertwined with all aspects of our selves and our societies from religion to culture to politics. Hence, a number of other more important transformations need to take place as well. For example, a language cannot be modern and still continue to have, as its ultimate authority, a sacred text as is the case with all modernized versions of Classical Arabic. No other text or grammatical study that is not based on the *Qurʾān* has appeared to date that is recognized within Arab societies as a rival to the authority of the holy text.

Although few deny that Classical Arabic predates the appearance of the *Qurʾān*, many see the relevant origin of the language in that event, namely, when God chose to speak Arabic to His Prophet. If the language of the *Qurʾān* is God's word, then believers can only be the custodians of that language, not its owners. In contrast, mother tongues are languages that grow up and change with their speakers who have rights over them (Sabri 1967; Haeri 2003; see also Khatibi 2002). It is their speakers who have ultimate authority albeit with various hierarchies of power. Classical Arabic can continue to change linguistically, but without simultaneous changes in its social and cultural contexts, it cannot become modern (see Haeri 2003 for an extensive discussion).

The Classical Arabic of today is linguistically different from a century ago, but the name of the language has not changed. This is not surprising but relevant since it adds to the ambiguity of the language situation. Whether in textbooks or newspapers or in ordinary conversation, this language is still referred to as *al-luġa al-ʿarabiyya al-fuṣḥā* or shortened versions of it. The contemporary versions are routinely referred to, in English-language publications, as → 'Modern Standard Arabic'. The use of this term has been avoided here because in Arabic no equivalent is in current use and because the term takes the modernity of the language as wholly unproblematic. Modern Standard Arabic is a term coined at Harvard University in the 1960s (Charles Ferguson, p.c.). There are terms used by some Arab intellectuals, such as *fuṣḥā al-ʿaṣr* 'contemporary Classical Arabic', *al-fuṣḥā al-ḥadīṯa* 'Modern Classical Arabic' and so on, but these are not in prevalent use among the majority of people. Many educated Arabs, particularly non-religious intellectuals, insist that not only has there emerged a language that is linguistically different but that this new language is independent of religion and has no special relationship to it. Yet, a majority of people disagree with these assertions and find them rather extraordinary (Haeri 2003). There are also other intellectuals who in the past as at present continue to write on the problems and dilemmas of the language situation and do not view the modernization process as an unproblematic success.

In the highly heated debates on these issues, one figure is constantly referenced and spoken of, and that is the figure of the 'educated Arab'. One hears of the views, needs, and practices of the educated Arab in articles, books, and conferences. Even Edward Said, a deservedly respected writer who was acutely aware of power and its unequal workings, spoke exclusively of the educated Arab in a posthumously published article on the language situation (Said 2004). He also acknowledged that after years of trying to gain rhetorical proficiency in the language, he continued to be on its 'margins'. The recently published Arab Human Development Report (UNDP 2002), written by a group of Arab scholars, states that: "About 65 million adult Arabs are illiterate, two thirds of them women. Illiteracy rates are much higher than in much poorer countries" (2002:3). The question remains, therefore, what about the uneducated Arab? On the rare occasions when uneducated Arabs are discussed, the claim is made that they 'understand' Classical Arabic even if they cannot speak it or write it. More recently, the claim is supported by pointing to the popularity of satellite channels such as al-Jazeera. This is taken as proof that even without education everyone understands the language. Yet, no empirical study has been undertaken on this claim. More generally, we must acknowledge that understanding and knowledge of a language are difficult and multifaceted issues to investigate. It is not unlikely that those with very little education develop strategies to decipher some newspaper articles or television programs. Nevertheless, this should cease to be a matter of opinion and ideology and is in sore need of empirical investigation.

6. The experience of education: Replacing 'diglossia' with other dualities and multiplicities

The co-existence of Classical and vernacular Arabic referred to by Arab scholars as *al-izdiwājiyya* 'duality' was analyzed in a landmark article by Charles Ferguson (1959) that was entitled → "Diglossia." While this brief article made a great contribution to Arabic sociolinguistics and helped launch hundreds of other studies (Hudson 1992; Fernández 1993), few of these investigated the actual workings of the contemporary sociolinguistic settings within the Arab world.

One productive alternative is to examine the duality of experience with regard to the official language as people go through the state educational system. Again, ethnographic studies of the experience of public education are lacking (but see Wagner 1993 for Morocco). However, in Haeri's (2003) ethnography of the language situation in Egypt, it was found that adolescents and adults with whom the author spoke described almost antithetical experiences with Arabic in the domain of religion versus all other domains. Most Muslim children first encounter Classical Arabic when they hear prayers and recitations. If their parents are believing Muslims, they are taught to memorize the prayers and short Qurʾānic *sūra*s. Some children are encouraged to recite them for family and guests. Whether or not they attend religious schools (*kuttāb*), so long as they have not entered the school system, they do not have to master (understand and become proficient in) the grammatical details of Classical Arabic. But when they begin school, they are gradually confronted with Classical Arabic as a language that must be learned (grammar, vocabulary, morphology), and produced correctly in reading and writing. Their knowledge of their mother tongues (any one of the Arabic vernaculars) does not prepare them for the significant linguistic differences in Classical Arabic.

While many people across social classes and educational levels speak of the use of Classical Arabic in prayers and recitations as practices that give them calm and peace of mind, they simultaneously state that their grammar classes are too difficult and they dislike reading and writing for fear of making mistakes (Haeri 2003). These antithetical experiences with Classical Arabic in the domain of religion, and with Classical Arabic and its modernized varieties in all other domains are part of the complexity of the language situation in the Arab world. One gleans in them the irony that precisely the language that is supposed to be 'simpler' than the language of the *Qur'ān* is the one people find 'difficult' and 'convoluted'.

This bifurcation of experience is one of the major sources of ambiguity in the language situation. As was mentioned earlier, modernized versions of Classical Arabic are not normally referred to any differently in order to distinguish them from older forms. Moreover, it is not unusual to find that individual educated Arabs at times refer to Classical Arabic as the 'language of my country' or as their native language. Writing on Arabization in Algeria, Berger (2002:2–3) states: "The question was then: Which Arabic [Classical or vernacular] should the Algerians be taught in order for them to become proper Arabophone speakers and hence legitimate citizens of the Algerian nation? This very ambiguity seems to have impinged on the nationalist movement from the start". There is also ambiguity with regard to the Classical Arabic of religion and that of other domains. The last four Egyptian constitutions mention Classical Arabic (*al-luġa al-ʿarabiyya*) as the official language of the nation. But the question that hovers over this definition of citizenship is: which Arabic, the religious one or the modernized one or both?

7. BERBERS IN THE MAGHREB

In addition to vernacular and Classical Arabic, French is an integral part of daily life in the Maghreb. Many newspapers and parts of the education system and of government bureaucracies use French. Long predating Arabic and French is Tamazight, the language of the → Berber populations. In Morocco Berbers comprise about 60 percent of the population and in Algeria about 20 percent, though throughout the centuries there has been intermarriage so that percentages must be understood with caution. In the Maghreb countries, nationalists sought the Arabization of education, bureaucracy, and mass media, thereby marginalizing Berber language and culture (Grandguillaume 1983, 1991). In Algeria, a central and historical demand of Berbers has been the official recognition of their language Tamazight (and its varieties). This demand has been voiced in periodic large-scale demonstrations by Berbers. The law of Arabization that was voted by the Popular National Assembly of Algeria in 1990 – the strictest such law in the Arab world – required not only the Arabization of administration and of tertiary education (by 1997), but also of all imported technology, media, billboards, and road signs (Djité 1992:15). Although the law was supposed to be fully implemented by 1998, it has run into predictable problems. As a result of many struggles, Berber was finally recognized in 2002 as a national but not an official language and steps have been taken to include it in the educational curricula. Saadi-Mokrane (2002) speaks of the violence of the language question in Algeria as *linguicide*, where the death of Arabic, French, and Berber was predicted by the colonizers or Arab nationalists.

In Morocco, there exists a nationalist, Islamist, and governmental discourse on Arabization (Mouhssine 1995). King Hassan II argued against an exclusive reliance on (Classical) Arabic for Moroccans: "We are for Arabization. But if that is a task, bilingualism [Arabic/French] is a necessity" (Mouhssine 1995:49, translated from the French original). In Morocco as well, Berber has begun to be taught at some schools and textbooks in that language went on sale for the first time in 2003.

BIBLIOGRAPHICAL REFERENCES

Abu-Lughod, Lila. 1986. *Veiled sentiments: Honor and poetry in a Bedouin society*. Berkeley, Los Angeles, and Oxford: University of California Press.

Ahmed, Leila. 1999. *A border passage: From Cairo to America – a woman's journey*. New York: Farrar, Straus and Giroux.

Ali, Abdullah Yusuf. 1992. *The Holy Qur'an: English translation of the meanings of the Qur'an with notes*. Indianapolis: H and C International.

Allen, Roger. 1982. *The Arabic novel: An historical and critical introduction*. Syracuse, N.Y.: Syracuse University Press.

Al-Musawi, Muhsin Jassim. 2003. *The postcolonial novel: Debating ambivalence*. Leiden: E.J. Brill.

Armbrust, Walter. 1996. *Mass culture and modernism in Egypt*. Cambridge: Cambridge University Press.

Atiyeh, George (ed.). 1995. *The book in the Islamic world: The written word and communication in the*

Middle East. Albany: State University of New York Press.

Badran, Margot. 1995. *Feminists, Islam, and nation: Gender and the making of modern Egypt*. Princeton, N.J.: Princeton University Press.

Barakat, Halim. *The Arab world: Society, culture and state*. Berkeley: University of California Press.

Belnap, Kirk and Niloofar Haeri (eds.). 1997. *Structuralist studies in Arabic linguistics: Charles Ferguson's papers, 1954–1994*. Leiden: E.J. Brill.

Berger, Anne-Emmanuelle. 2002. *Algeria in others' languages*. Ithaca, N.Y. and London: Cornell University Press.

Cachia, Pierre. 1990. *An overview of modern Arabic literature*. Edinburgh: Edinburgh University Press.

Caton, Steven. 1990. *"Peaks of Yemeni summon": Poetry as cultural practice in a North Yemeni tribe*. Berkeley, Los Angeles, and Oxford: University of California Press.

Danielson, Virginia. 1997. *The voice of Egypt: Umm Kulthum, Arabic song, and Egyptian society in the twentieth century*. Chicago and London: University of Chicago Press.

Djité, Paulin. 1992. "The arabization of Algeria: Linguistic and sociopolitical motivations". *International Journal of the Sociology of Language* 98.15–28.

Doss, Madiha and Catherine Miller (eds.). 1996. *Les langues en Egypte*. Special issue of *Égypte/Monde arabe* 27–28:3–4.

Douglas, Allen and Fedwa Malti-Douglas. 1994. *Arab comic strips: Politics of an emerging mass culture*. Bloomington and Indianapolis: Indiana University Press.

Eid, Mushira. 2001. *The world of obituaries: Gender across cultures and over time*. Detroit, Mich.: Wayne State University Press.

Ferguson, Charles. 1959a. "Diglossia." *Word* 15.325–340. (Repr. *Sociolinguistic perspectives: Papers on language and society, 1959–1994/ Charles Ferguson*, ed. Thom Huebner, 25–39. New York and Oxford: Oxford University Press, 1996.)

———. 1959b. "The Arabic koiné". *Language* 35.616–630. (Repr. Belnap and Haeri 1997: 50–68.)

Fernández, Mauro. 1993. *Diglossia: A comprehensive bibliography 1960–1990*. Amsterdam and Philadelphia: J. Benjamins.

Grandguillaume, Gilbert. 1983. *Arabisation et politique linguistique au Maghreb*. Paris: Maisonneuve et Larose.

———. 1991. "Arabisation et langage maternelle dans le contexte national au Maghreb". *International Journal of the Sociology of Language* 87.45–54.

Haeri, Niloofar. 2000. "Form and ideology: Arabic sociolinguistics and beyond". *Annual Review of Anthropology* 29.61–87.

———. 2003. *Sacred language, ordinary people: Dilemmas of culture and politics in Egypt*. New York: Palgrave Macmillan.

Hafez, Sabry. 1993. *The genesis of Arabic narrative discourse*. London: Saqi Books.

Hodgson, Marshall. 1974. *The venture of Islam: Conscience and history in a world civilization*. 3 vols. Chicago: University of Chicago Press.

Holes, Clive. 1988. *Language variation and change in a modernising Arab state: The case of Bahrain*. London: Kegan Paul International.

Holt, P.M., Ann K.S. Lambton, and Bernard Lewis. 1970. *The Cambridge history of Islam*, IIB. Cambridge: Cambridge University Press.

Hourani, Albert. 1991. *A history of the Arab peoples*. Cambridge, Mass.: Belknap Press and Harvard University Press.

Hudson Alan. 1992. "Diglossia: A bibliographic review". *Language in Society* 21.611–674.

Khatibi, Abdelkebir. 2002. "Diglossia". Berger (2002: 157–160.

Khorramshahi, Bahaeddin. 1997. *Qur'an-e karim: Tarjomeh, tozihat, va vajeh-nameh*. Tehran: Niloofar Publishers.

Khoury, Philip. 2003. *Urban notables and Arab nationalism: The politics of Damascus 1860–1920*. Cambridge: Cambridge University Press.

Mahdi, Muhsin. 1995. "From the manuscript age to the age of printed books". *The book in the Islamic world: The written word and communication in the Middle East*, ed. George Atiyeh, 1–14. Albany: State University of New York Press.

Miller, Flagg. 2005. "Of songs and signs: Audiocassette poetry, moral character and the culture of circulation in Yemen". *American Ethnologist* 32: 1.82–99.

Parkinson, Dilworth. 1991. "Searching for modern *fuṣḥā*: Real life formal Arabic". *al-ʿArabiyya* 24.31–64.

Reynolds, Dwight. 1995. *Heroic poets, poetic heroes: The ethnography of performance in an Arabic oral epic tradition (myth and poetics)*. Ithaca, N.Y. and London: Cornell University Press.

Saadi-Mokrane, Djamila. 2002. "The Algerian linguicide." Berger (2002:44–59).

Ṣabrī, ʿUṯmān. 1967. "Il-luġa il-ʿarabiyya il-ḥadīsa aw il-luġa il-maṣriyya (il-ʿāmmiyya)". *Riḥla fi n-Nīl*, by ʿUṯmān Ṣabrī. Alexandria: Maktabat al-ʿArab al-Miṣriyya. (Repr. *al-Garad* 1994, 79–111, under the title "Fī l luġa al-maṣriyya al-muʿāṣira".)

Shafik, Viola. 1998. *Arab cinema: History and cultural identity*. Cairo: American University in Cairo Press.

Shryock, Andrew. 1997. *Nationalism and the genealogical imagination: Oral history and textual authority in tribal Jordan*. Berkeley: University of California Press.

Starrett, Gregory. 1998. *Putting Islam to work: Education, politics, and religious transformation in Egypt*. Berkeley, Cal.: University of California Press.

Suleiman, Yasir. 2003. *The Arabic language and national identity*. Edinburgh: Edinburgh University Press.

UNDP (United Nations Development Programme). 2002. *Arab human development report: Creating opportunities for future generations*. New York: UNDP.

Versteegh, Kees. 1997. *The Arabic language*. Edinburgh: Edinburgh University Press.

Wagner, Daniel. 1993. *Literacy, culture, and development: Becoming literate in Morocco*. Cambridge: Cambridge University Press.

Walters K. 1991. "Women, men and linguistic variation in the Arab world". *Perspectives on Arabic lin-*

guistics, III, ed. Bernard Comrie and Mushira Eid, 199–229. Amsterdam and Philadelphia: J. Benjamins.

Zeidan, Joseph. 1995. *Arab women novelists: The formative years and beyond.* Albany: State University of New York Press.

Niloofar Haeri (Johns Hopkins University)

Cypriot Maronite Arabic

1. Introduction

Cypriot Arabic is a non-literate vernacular that has been spoken natively in Cyprus for probably well over a millenium and continues to be used by a community of about 1,300 Cypriot Maronites, i.e. the former inhabitants of Kormakiti, or Korucam in Turkish, resettled in the Greek sector of Nicosia, in Larnaca and Limassol, in the aftermath of the Turkish invasion of Cyprus in 1974.

Reference to a Cypriot variety of Arabic occurs in a 13th-century Arabic work entitled *Kitāb as-simāt fī ʾasmāʾ an-nabāt* (Treatise on the characteristics of plant names) by ʿIzz ad-Dīn ʾAbū ʾIsḥāq ibn Muḥammad ibn Tarxān as-Suwaydī (1204– 1292) (see Ullmann 1970:291).

The travel account by the Augustinian monk Frater Jacobus de Verona (1335) states "omnes de Cypro loquuntur grecum, bene tamen sciunt saracenicum et linguam francigenam, sed plus utuntur greca" (Röhricht 1895:178). Whereas the Maronites of Cyprus are frequently mentioned in historical sources, their Arabic speech is mostly overlooked; the earliest modern allusion to it occurs in a brief survey of the language situation in Cyprus by Beaudouin (1884:11). Some 50 years later, Storrs (1930:41) referred to "a bastard Arabic mixed with Greek" spoken by the Cypriot Maronites but, regrettably, linguistic scholars of the time did not follow this lead; Cypriot Arabic has become, in the meantime, a terminal language.

The earliest reliable data on Cypriot Arabic were published in a short but informative sketch by the Greek dialectologist Brian Newton (1964). Tsiapera (1969), purporting to be a structuralist description of this vernacular, is marred by serious factual and methodological shortcomings rendering it virtually unusable as a reference work (see Jastrow 1977). The profoundly Arabic character of Cypriot Arabic first emerged in Arlette Roth's (1973–1975) valuable investigation of its verbal system.

Borg (1985) represents the most extensive and detailed linguistic description of Cypriot Arabic to date and addresses principally phonology and morphology but also provides specimen folk texts in transcription with English translations. Borg (2004) studies this vernacular's lexical inventory from a comparative and historical perspective. Apart from elucidating the highly evolved structural profile of Cypriot Arabic, the upshot of Borg's research, relating principally to the areal links of Cypriot Arabic with other varieties of Eastern Arabic, highlights several unexpected parallels with the so-called *qəltu* dialect group (Blanc 1964; Jastrow 1978) – a factor that raises intriguing questions relating to this vernacular's external history and the evolution of the contemporary Arabic vernaculars spoken along the Anatolian Syrian/Mesopotamian continuum.

The rest of this entry sets out to delineate the significance of Cypriot Arabic for Arabic language history and addresses mainly the most characteristic and areally significant features in the diachronic phonology and morphophonemics of Cypriot Arabic, and the question of Aramaic substrata. On general synchronic aspects of Cypriot Arabic, Roth (1973–1975) and Borg (1985) can be consulted.

2. External History

Striking formal traits of Cypriot Arabic unattested in other Eastern Arabic dialects (such as the systematic fusion of Old Arabic ʿ and ġ > Cypriot Arabic ʿ and occasional reflexes of the *begadkefat* rule indicative of an Aramaic substrate) suggest that this vernacular's medieval antecedent displayed an evolutionary stage chronologically close to the language shift from Aramaic to Arabic. Significantly, Mas-Latrie (1861:106) ascribes to the Maronites an early settlement on the island in the 7th century.

A later chronology for the presence of Maronites in Cyprus is proposed by the Lebanese Maronite historian Pierre Dib who links the emigration of Maronites to the island with "la persécution générale qui eut lieu sous al-Mamoun (813–833) en Syrie et en Palestine" (Dib 1962:71). For the later period, Philip Hitti postulates waves of Lebanese Maronites to

Cyprus in the 12th and 13th centuries (Hitti 1957a:623; 1957b:325). Thus Cypriot Arabic is, in essence, the offshoot of a Medieval Arabic vernacular with a Christian communal imprint concomitant with its speakers' Aramaic ethnic origins and *ḏimmī* status, later evolving in a Greek-speaking cultural milieu.

The distinctive sociohistorical profile of Cypriot Arabic invites comparison with those of other 'peripheral' varieties of Arabic: → Maltese, certain dialects of Anatolian Arabic, and the Arabic vernaculars of → Uzbekistan, → Afghanistan, and Central Africa which are, on historical, typological, cultural, and sociolinguistic grounds, today more realistically regarded as autonomous languages rather than simply as 'Arabic dialects'. Thus, unlike most native Arabs, their speakers tend to be typically bilingual or multilingual which means, *inter alia*, that the notion of → 'diglossia', as commonly discussed in relation to contemporary spoken Arabic, is inapplicable to these special varieties of Arabic.

The sociohistorical parallel between Cypriot Arabic and Maltese is particularly close (cf. Borg 1994) since, in both cases, we are dealing with an Arabic vernacular surviving in complete isolation from the Arabic-speaking world, exposed to interaction with a variety of Indo-European (Italian and English in the case of Maltese, Greek in that of Cypriot Arabic), and spoken by Catholic Christians in a Mediterranean and insular sociocultural habitat.

Cypriot Arabic is a linguistic relic area and consequently an ideal venue for a study contrasting the center with the periphery. Observe, for instance, lexical rarities like Cypriot Arabic *kanirízz*, *pirízz* 'to inherit' < Classical Arabic *razaʾa*; *šaxve* 'hair of head' < Classical Arabic *šaʿafa*; *ʿarra*, *piʿarri* 'to exhaust, consume' < Classical Arabic, Ugaritic *ʿrw*, etc. Furthermore, given its speakers' confessional background, Cypriot Arabic is of special interest to the Arabist since it attests to a form of linguistic usage free of Literary Arabic influences and koineizing trends liable to obscure its pristine colloquial character.

3. PHONOLOGY AND MORPHOPHONOLOGY

The phonological analysis of Cypriot Arabic in Borg (1985:11–74) postulated a segmental paradigm of 18 consonants and 7 vocalic nuclei: /p, t, k, f, ṯ, s, š, x, v, ḏ, r, ž, ʿ, m, n, l, r, y, i, e, a, o, u, ay, and aw/. This is a somewhat reduced inventory when compared with the 36 phonemes of Classical Arabic and with the evolved sound systems of many contemporary Arabic vernaculars, e.g. → Damascus Arabic with its 40-odd sound segments (Cowell 1964:2–9). Significantly, the sound system of Cypriot Greek, a source of adstratal influence on Cypriot Arabic, comprises 25 phonemes. The present Cypriot Arabic sound system is the outcome of paradigmatic shifts examined in 3.2 below.

In essence, it is the evolutionary outcome of two main chronological stages corresponding to a pre-Cypriot period during which its parent dialect was spoken on the mainland, and the Cypriot stage after it became separated from mainstream Arabic and came in contact with Greek.

3.1 *The pre-Cypriot stage*

The earlier linguistic stratum of Cypriot Arabic shows clear diagnostic areal traits typifying mainstream Arabic colloquials spoken along the Syro-Anatolian-Mesopotamian dialectal continuum, i.e. the so-called *qəltu* dialect type. Highly systematic in Cypriot Arabic is the fronting of Classical Arabic *ā* in the historical adjacency of a long or short high front vowel ([i:] or [i]), i.e. the umlaut type of → *ʾimāla* shift systematically exemplified in Jewish and Christian Baghdadi, in Mosul, Aleppo (cf. Blanc 1964:48), and southern Turkey (Jastrow 1969; Procházka 2002:40f):

Cypriot Arabic		Classical Arabic
klep	'dogs'	< *kilāb*
rex	'going [masc.]' *(<* *rēyeḥ*)	< *rāyiḥ*
pkyeter	'threshing-floors'	< *bayādir*
znepíl	'baskets'	< *zanābīl*
pan	'he seemed'	< *bān*
klam	'words'	< *kalām*
xlave	'sweetness'	< *ḥalāwa*

Cypriot Arabic also shows special cases of the *ʾimāla* typifying this dialect group as a whole: Cypriot Arabic *nes* 'people', *žež* 'hens', presumably continuing Classical Arabic **inās* and *dijāj* (~ *dajāj*), respectively. The vocalically conditioned *ʾimāla* also occurs in certain North Syrian dialects showing principally this, but not necessarily other, features of the *qəltu* group. Blanc (p.c.) referred to this subgroup by the term *qētel* vernaculars; Cypriot Arabic is such a dialect.

Certain Lebanese and Syrian Arabic dialects also display an *ʾimāla* shift of *ā* > *ē*/*ī* affecting certain varieties of Lebanese (and Syrian) Arabic: Palmyra *lsēn* 'tongue', *bzēz* 'breasts' (Cantineau 1934:75–76) and the dialects of Beirut (Mattsson 1911:65f), Qarîtēn, Palmyra, and Soukhné (Cantineau 1956:121, 128–129), where it occurs specifically in contact with non-emphatic and non-back consonantal environments.

Cypriot Arabic contrasts stressed /i/, /u/, and /a/ but curtails the ancient phonological distribution of these contrasts via extensive shift of tonic Classical Arabic /a/ > i/: *kilp* 'dog', *sift* 'Saturday', *timm* 'blood' < *damm*, *simm* 'poison' < *samm*. This rare areal trait in the Syro-Anatolian continuum occurs in the Qalamun region: *kilb* 'Hund', *šims* 'Sonne', *kibš* 'Hammel', etc. (Arnold and Behnstedt 1993:19). Reflexes of the stem vowel in reflexes of CuCC are mostly retained distinct in Cypriot Arabic; the outcome is usually /o/ in the adjacency of historical *x*, *ḥ*, ʿ, *ġ*, and before *r*, but /u/ elsewhere: *oxt* 'sister', *xops* 'bread', *koxl* 'antimony', *xork* 'anger' (< *xulq*), *šoʿol* 'work', *ʿorpe* 'foreign parts', *žorne* 'tub', *korne* 'corner', *umm* 'mother', *pukra* 'tomorrow', *župn* 'cheese', *žúmaʿa* 'week', *truntž* 'citrons', etc. These vowels have undergone systemic fusion with historical or secondary long counterparts: *tut* 'mulberries', *žoʿo* 'hunger' (< *jūʿ*), etc.

The following phonological or morphophonemic shifts are particularly significant from an areal standpoint:

i. occasional reflexes of the Aramaic *begadkefat* rule yielding spirants from historical stops (see below), (cf. Muslim Baghdadi *tūṯ* ~ Anatolian Arabic *tūf* < Classical Arabic *tūt* 'mulberries'; Woodhead and Beene 1967:60; Vocke and Waldner 1982:75).

ii. Systematic loss of the word-initial laryngeal spirant in pronominal and deictic forms. As in the Diyarbakır and Siirt dialect groups of Southeast Anatolia, historical treatment of Classical Arabic **h* in Cypriot Arabic discriminates between root-morphemic realizations of this segment and its incidence in indexical forms, the latter being irreducible to the constituent stucture of 'root + stem formative'. The laryngeal is retained in the former but deleted in the latter: *šipex* 'he resembled' < **šibih*; *uo* 'he' < *huwwa*, *ie* 'she' < *hiyya*, etc.

iii. Like Anatolian Arabic, Cypriot Arabic shows permanent agglutination of the feminine ending *(tāʾ marbūṭa)* to certain plural nouns even when no numeral precedes: *tišxúr* 'months', *ṯkyem* 'days' < **tiyyām*, *tižmáʿa* 'weeks', *táʿirfe* 'loaves', etc.; cf. Anatolian Arabic *alf*, pl. *talāf* 'tausend', *šahr*, pl. *təšhor/təšhūr* 'Monat' (Vocke and Waldner 1982:13, 232).

iv. the morphophonemic shift of the verbal proclitics |ta-| and |tta-| ⇒ |te-| and |tte-| (signaling the future) before the underlying 3rd pers. masc. sg. marker |y-| also typifying Anatolian Arabic vernaculars (Jastrow 1978:303; 1999:46).

v. ellipsis of the semivowel /w/ in the imperfect inflection of the Cypriot Arabic verb *sava*, *pisáy* 'to make' < **sawwā*, **bisawwī*: *psay*, *pitsáy*, *pisáy* 'I make, you make, he makes, etc'.

Other areally diagnostic features in Cypriot Arabic enjoy a wider distribution and continue more general isoglosses in the region:

i. the *b*-imperfect, whose geographical distribution extends across most of the Syrian, Lebanese, Palestinian, and Egyptian dialect areas, sedentary and nomadic (cf. Cowell 1964:174; Cantineau 1946:221, 1956:125; Mitchell 1962:81).

ii. vowel alternations in the inflection of the Cypriot Arabic imperative (*xtop*, *xtupi*, *xtupu* 'write!') replicating synchronic vowel changes contingent on closed vs. open syllabicity typifying several Syrian and Lebanese vernaculars: Aleppo *skōb*, *skəbi*, *skəbu* 'pour!' (Sabuni 1980:119). The sensitivity of historically long vowel quality to syllabic structure in Eastern Arabic dialects is commonly attributed to Aramaic substratal influence (Nöldeke 1888:34; Lewin 1969:23).

3.2 *Cypriot stage*

The highly distinctive surface phonology of Cypriot Arabic terms is immediately striking to the observer and attests to extensive paradigmatic restructuring in its sound system induced by language contact with Modern Greek. As already noted, three paradigmatic shifts in Cypriot Arabic phonology merit notice: (a) reduction in points of articulation along the back of the vocal tract arising from unconditioned fusion of three Classical Arabic consonant pairs: ʿ and *ġ*, *ḥ* and *x*, and *k* and *q* into Cypriot Arabic ʿ, *x*, and *k* respectively; (b) fusion of historical emphatics with their plain counterparts; and (c) absolute neutralization of the historical voicing contrast in stops (via contact with Greek) yielding in Cypriot Arabic a set of three voice-indifferent stop segments: /p, t, and k/: /payt/ 'house', /tayn/ 'debt', and /kilp/ 'dog' < Classical Arabic *bayt*, *dayn*, and *kalb* respectively. The following display presents a bird's-eye view of the principal correspondences between the consonant systems of Cypriot Arabic and Classical Arabic:

Classical Arabic		Cypriot Arabic
t, *ṭ*, *d*	>	*t* ~ [θ]
s, *ṣ*	>	*s*
k, *q*	>	*k* ~ [x]
ʾ	>	Ø
b	>	*p* ~ [f]
d	>	*t*
ḏ, *ḏ̣*	>	*ḏ*
j	>	*ž*

The cumulative effect of this systemic streamlining process has occasioned extensive re-etymologization and proliferation of homophones: *prk* 'to bless, to flash [lightning]' < Classical Arabic *brk*, *brq*; *ktp* 'to write, to frown' < Classical Arabic *ktb*, *qṭb*, *xlk* 'to be born', to shave', 'to arrive' < Classical Arabic *xlq*, *ḥlq*, *lḥq*, etc. Historical fusion of root morphemes within the lexicon on a comparable scale, also caused by loss of certain Classical Arabic consonantal oppositions, has occurred in Maltese.

Through contact with Greek, Cypriot Arabic acquired phonotactic constraints untypical of Arabic determining surface realization of consonant clusters; specifically, rules of (a) manner dissimilation transforming 'stop + stop' sequences into 'fricative + stop': Cypriot Arabic *xtuft* 'I wrote' ⇐ |ktupt| < Arabic *katabt*, Cypriot Arabic *fkum* 'I get up' ⇐ |pkum| < Arabic *baqūm*, etc.; and (b) *yod*-occlusivization in sequences of 'obstruent + *y*': *pkyut* 'houses' < Arabic *byūt*, Cypriot Arabic *ṯkyep* 'clothes' < Arabic *ṯyāb*, Cypriot Arabic *pkyara* 'cisterns' < Arabic *byāra*, etc. Cypriot Arabic also tends to palatalize the velar consonants *x* and *k* and to reduce underlying |ly| and |ny| to /yy/ before *i* and *e*. Cypriot Arabic has an unsystematic rule of postnasal epenthesis: *šimps* 'sun', *intsán* 'man', etc.; in many cases, the epenthetic consonant has probably been lexicalized (cf. French *nombre*).

Cypriot Arabic has lost the length opposition in vowels via fusion of Classical Arabic short vowels with their long counterparts: Classical Arabic *i*, *ī* > /i/; Classical Arabic *u*, *ū* > /u/, and Classical Arabic *a*, *ā* > /a/. The historical diphthongs *ay* and *aw* have been extensively retained but, as already noted, occasionally yield /e/ and /o/: *tex* 'his hands' < **īdayh*, *xok* 'on, above' < Classical Arabic *fawq*.

4. ARAMAIC SUBSTRATA

Highly noteworthy in Cypriot Arabic phonology is the aforementioned fusion of Classical Arabic ʿ and *ġ* > Cypriot Arabic ʿ, and of Classical Arabic *ḥ* and *x* > **ḥ* > Cypriot Arabic *x*. The velar reflex here represents a later shift induced by contact with Cypriot Greek /x/ (= Greek χ):

Cypriot Arabic		Classical Arabic
ʿakl	'intelligence'	*ʿaql*
ʿapn	'displeasure'	*ġabn*
saʿet	'he helped'	*sāʿad*
páʿale	'mule'	*baġla*
xilef	'he swore'	*ḥalafa*
xirep	'he destroyed'	*xaraba*
taxak	'he laughed'	*ḍaḥika*
naxal	'he sifted'	*naxala*

Though consonant gemination is common in this vernacular, the voiced pharyngeal fricative ʿ never appears long – a restriction plausibly ascribable to Aramaic substratal influence.

Fusion of Classical Arabic ʿ and *ġ* in Cypriot Arabic is unlikely to have transpired in contact with the Greek sound system since the presence in the latter of the voiced velar fricative [*ġ*] < Ancient Greek *gamma* should logically have promoted retention of Classical Arabic *ġ* rather than that of the voiced pharyngeal fricative ʿ. Of interest here is the shift of Common Semitic **g* >

ġ > ʿ as in the Cypriot Arabic form *ʿaddef* ‘to curse’ < **ġaddef* < **gaddef* (Classical Arabic *jaddafa*), where the voiced velar fricative *ġ* is the substratal outcome of the stop lenition constraint called *begadkefat*. In Biblical Hebrew and literary Aramaic, this replaced postvocalic stops by the corresponding fricatives. Traces of the fusion of Classical Arabic ʿ and ġ occur, on a limited scale, in Lebanese Arabic yielding, as in Cypriot Arabic, retention of Aramaic ʿ where Arabic shows *ġ*: Kafr ʿAbīda: *aʿmaṣ* ‘qui a les yeux chassieux’ ~ Syriac *ʿəmīṣūṯā* ‘lippitudo oculorum’ ~ Old Arabic *aġmaṣ*; *bāʿūt* ‘prière de demande, supplication’ < Syriac *bāʿūṯā* ‘petitio, supplicatio’ ~ Classical Arabic *biġya*; *ʿebe* ‘il rendit épais, dense, serré (blé, bois)’ ~ Syriac *ʿəbā* ‘densus, spissus, crassus fuit’ ~ Classical Arabic *aġbā* ‘épais, touffu, couvert d’épais feuillages (arbre, branche)’ (see Feghali 1918:43).

The systematic character of the correspondence, Classical Arabic *ġ* > ʿ exemplified in Cypriot Arabic is unique in the Arabic *Sprachraum* and plausibly signals the retention in this peripheral dialect of an earlier chronological historical stage in the phonological evolution of Eastern Arabic than is attested in any contemporary mainstream dialect of Arabic described so far, namely, the period that witnessed the initial interaction between Arabic and Aramaic. Retention of the voiced pharyngeal fricative ʿ in Neo-Aramaic is the norm in Western and Central Neo-Aramaic (Arnold 1990: *passim*; Jastrow 1985:2).

Grammaticalization of stress in Cypriot Arabic via accentuation of the heads of certain syntactic constructions (e.g. the negative particles *lá* and *má*) may also be ascribable to Aramaic. Phonological words with a verbal nucleus often retain initial syllable stress even when heavy syllables intervene closer to the final word boundary: *ʿížipna* ‘we liked it’, *táxakilla* (< *ḍ-ḥ-k*) ‘he cheated her’, *šáttimna* ‘he abused us’, *tláxitna* ‘the three of us’, *síttiṯkon* ‘the six of them’, *xálluon* ‘leave them!’ (as opposed to *xallúon* ‘they left them’), etc. Note the closely comparable stress assignment rule typifying certain varieties of Neo-Aramaic where “collocations of two, rarely three words which are closely bound syntactically can form stress groups . . . In stress groups the second word loses its stress, and the main stress of the collocation comes to be on the last syllable of the first word” (Jastrow 1997:353): Ṭuroyo *ló-kəḏʿina* ‘wir wissen nicht’ (Jastrow 1985:30). The behavior of Cypriot Arabic *kull* in *kúyyom* ‘every day’ < Arabic *kull yawm*, *kúlsaʿa* ‘every hour’ < Arabic **kúll-sāʿa*, etc., where the determiner receives stress, is paralleled in at least three Neo-Aramaic languages: Ṭuroyo *kə́l-yawmo* ‘jeder/jeden Tag’ (Jastrow 1985:30), Hertevin *kóyyom* ‘jeden Tag’ (Jastrow 1988:191), Urmi *kudjom* ‘every day’ (Garbell 1965:316). As in Aramaic, Cypriot Arabic *kull* requires a mandatory pronominal suffix copying the number and gender features of a following definite noun (cf. Nöldeke 1904:172). Cypriot Arabic extends this trait to *nuss* ‘half’ < Classical Arabic *niṣf*: *kulla s-sine* ‘the whole year’, *kullon in-nes* ‘all the people’, *nussu lil-ʿaríf* ‘half the loaf’, etc.

A noteworthy Aramaism in Eastern Arabic, including Cypriot Arabic, is the reflex of Aramaic *ʾīṯ b-* (lit. ‘there is in . . .’. = ‘to be able’; cf. Sokoloff 1990:55; Dalman 1967:16) yielding in Cypriot Arabic the paradigm of ‘*fi* + enclitic pronoun’ meaning ‘to be able’: Cypriot Arabic *fini* ‘I can’, *fik* ‘you [masc. sg.] can’, etc. Cognates also occur in several Eastern Arabic dialects and in Neo-Aramaic. Lebanese: *fiyyi*, *fīk* ‘ich kann, du kannst, usw’ (Bišmizzîn, Jiha 1964:149); Syrian Arabic: *fīk ᵊtsāʿədni?* ‘can you help me?’ (Damascus, Cowell 1964:547); Çukurova *fîtni b-sawwi hāda* ‘ich kann das machen’ (Procházka 2002:138); Ṭuroyo *kibi*, *kibux*, etc. ‘können, imstande sein’ (Jastrow 1990:107).

In the syntactic sphere, Cypriot Arabic often implements direct object marking by means of the proclitic particle *l-*, for instance, before an emphatic pronoun: *kíšʿeni ll-ana* ‘it was me he saw’, *rákʿalla ll-ie* ‘he hit her’; it can also appear without the concomitant proleptic pronoun characterizing Eastern Arabic vernaculars that have this trait: Lebanese (Koutsoudas 1967), Palestinian (Levin 1987), Iraqi (Blanc 1964:130), etc.

Cypriot Arabic also utilizes a special genitive construction specifically with inalienable nouns, principally kinship terms, entailing the possessive marker *l-* preceded by an proleptic pronoun: Cypriot Arabic *yapatu l-yorko* ‘George’s father’. On the Aramaic origin of this trait, see Blanc (1964:130); observe the close parallel in Central Neo-Aramaic where it is generally restricted to a small class of nominals comprising inalienables and temporal expressions: *əšmé-d-ḥoṯux* ‘your sister’s name’, *barṯé-d-ʿamti* ‘my paternal cousin’, etc.

5. LEXICON

The Aramaic component in the colloquial Arabic lexicon still awaits systematic treatment (cf. Diem 1979:41f., → Aramaic/Syriac loanwords). The Cypriot Arabic lexicon also shares several Aramaic features with other vernaculars of Eastern Arabic and often with Neo-Aramaic. Above all, its isolation from mainstream colloquial Arabic renders it a potential repository of Aramaisms rare in spoken Arabic.

Cypriot Arabic: *taylep* 'he prepared' ~ Syriac *ṭayyeb* 'paravit' (Brockelmann 1928:270); *ʿapi* 'dense [vegetation]' ~ Palestinian Arabic *ʿabi* 'dicht [von der Saat]' (Bauer 1957:75) ~ Syriac *ʿabī* 'thick, dense' (Payne Smith 1903:395); *mnyeššar* 'abandoned' Jewish Aramaic *nāšrā* 'das Abfallende' (Jastrow 1886:454); *firex* 'he fainted' ~ Aramaic *parəḥā nišmatān* 'their soul fled' (Jastrow 1886:1223; Sokoloff 1990:445) ~ Hebrew *pārəḥāh nišmatō* 'he passed away, he was scared to death, he fainted' (Even-Shoshan 1993:1095) ~ Mlaḥsō *poreḥ*, *priḥle* 'fliegen' (Jastrow 1994:160); Cypriot Arabic *kisx* 'small onion' ~ Syriac *kesḥā* (Brockelmann 1928:338) ~ Galilean Aramaic *qiṣḥā* 'kind of onion' (Sokoloff 1990:501); Cypriot Arabic *patrúr* (unity noun *-e*) 'mushrooms' ~ Syriac *pəṭūrṯā* 'fungi' (Brockelmann 1928:565) ~ Classical Arabic *fiṭr*, etc.

Feghali (1918:94) estimated that 75 percent percent of the 200-odd certain Aramaisms in Lebanese Arabic encode traditional household concepts. In Cypriot Arabic, the number noted so far does not exceed 70 at an outside limit; these predictably pervade the religious domain without being restricted to a particular semantic field. The Aramaean population traditionally consisted of rural communities (Poizat 1979: 355; Anschütz 1984:160) – a trait that also typified the Cypriot Maronites until their recent urbanization. Cypriot Arabic lexical Aramaisms relate to farming, household industries, ethnobotany, and general concepts:

xakle 'field', *sammex* 'to sprout'; *sunt* 'plough'; *moraš* 'threshing sled'; *paytar* 'threshing floor'; *kiten* 'to yoke [oxen]'; *stapl* 'barn, cowshed'; *saʿúr* 'kid'; s*aykún* 'brushwood'; *šummár* 'fennel'; *zaʿrúr* 'hawthorn', *terráš* 'bushes'; *šall* 'to sew'; *šammút* 'spindle'; *kriše* 'preparation for making *plata* cheese'; *tilef* 'it leaked'; *xarkeš* 'he moved'; *šataf* 'he rinsed'; *tallel* 'he thinned out'; *šaxxet* 'he begged for alms'; *šaxxat* 'he threw out, expelled'; *šalax* 'he undressed'; *ʿakkes* 'it stung, bit [insect, snake]'; *ʿarreb* 'he separated'; *tlatteš* 'it got stained'; *zakrúr* 'throat'; *xassek* 'he collected'; *afkax* 'lame'; *šara* 'midday meal', etc.

Covert Aramaisms also occur; note Cypriot Arabic *sak*, *pisúk* 'to plough' < Classical Arabic *sāq (u)* 'to drive' paralleled by Bəḥzāni *ysōq fadān* 'Pflügen' (Jastrow 1981:392); the monolexemic Cypriot Arabic reflex calques the elliptical Syriac expression *dəbar* 'to plough' (for *dəbar paddānā* 'das (Pflug)gespann führen' (Seidel 1988:159).

Aramaic substratal influences also show up in the grammar of function words relating to spatial and temporal categorization. Note, for instance, Cypriot Arabic *kintám* < *quddām*, which in Classical Arabic and many, probably most, modern vernaculars denotes exclusively spatial precedence, but which in Cypriot Arabic and certain varieties of vernacular Arabic (in Southeast Turkey; Vocke and Waldner 1982:143), encode, as in Aramaic, both temporal and spatial precedence.

6. CONCLUSION

The sketch of Cypriot Arabic presented here justifies the assumption of an early formation for this vernacular. Its areally hybrid Arabic character and far-flung dialectal affiliations within Eastern Arabic may hold the key to questions relating to the early history of an extensive *Sprachlandschaft* stretching from the sources of the Tigris and the Euphrates in Southeast Anatolia, across northern Syria and Mesopotamia, on the one hand, and embracing the Arabic dialects of Greater Syria, on the other, such as: (a) the origins and chronology of the dialect boundary dividing the Arabic dialects of Greater Syria from the so-called *qəltu* vernaculars; (b) the putative existence of medieval regional koines of Eastern Arabic cross-cutting later geolinguistic boundaries; and (c) the formal continuities obtaining between vernacular Aramaic and the sedentary Arabic dialects of the region.

BIBLIOGRAPHICAL REFERENCES

Andrews, Peter Alford. 1989. *Ethnic groups in the Republic of Turkey*. Wiesbaden: L. Reichert.

Anschütz, Helga. 1989. "Christliche Gruppen in der Türkei". Andrews (1989:454–471).

Arnold, Werner. 1990. *Das Neuwestaramäische*. V. *Grammatik*. Wiesbaden: O. Harrassowitz.

—— and Peter Behnstedt. 1993. *Arabisch-aramäische Sprachbeziehungen im Qalamūn (Syrien): Eine* *dialektgeographische Untersuchung*. Wiesbaden: O. Harrassowitz.

Bauer, Leonhard. 1957. *Deutsch-arabisches Wörterbuch der Umgangssprache in Palästina und im Libanon*. 2nd ed. Wiesbaden: O. Harrassowitz.

Beaudouin, Mondry. 1884. *Étude du dialecte chypriote moderne et médiéval*. Paris: Ernest Thorin.

Blanc, Haim. 1964. *Communal dialects in Baghdad*. Cambridge, Mass.: Harvard University.

Borg, Alexander. 1985. *Cypriot Arabic*. Stuttgart: F. Steiner.

——. 1994. "Some evolutionary parallels and divergences in Cypriot Arabic and Maltese". *Mediterranean Language Review* 8.41–67.

——. 2004. *A comparative glossary of Cypriot Maronite Arabic*. Leiden: E.J. Brill.

Brockelmann, Carl. 1928. *Lexicon Syriacum, editio secunda aucta et emendata*. (Repr. Hildesheim: Georg Olms, 1966.)

Cantineau, Jean. 1934. *Le dialecte de Palmyre*. II. *Vocabulaire et textes*. Beyrouth: Institut Français de Damas.

——. 1946. *Les parlers arabes du Ḥōran*. Paris: C. Klincksieck.

——. 1956. "Notes sur les parlers arabes des oasis syriennes: Qârîtēn, Palmyre, Soukhné". *Studi orientalistici in onore di Giorgio Della Vida*, I, 120–31. Rome: Istituto per l'Oriente.

Cowell, Mark W. 1964. *A reference grammar of Syrian Arabic*. Washington, D.C.: Georgetown University Press.

Dalman, Gustav. 1938. *Aramäisch-Neuhebräisches Handwörterbuch zu Targum, Talmud und Midrasch*. (Repr. Göttingen: Vandenhoek und Ruprecht, 1967.)

Dib, Pierre. 1962. *Histoire de l'eglise maronite*, I. Beirut: Editions 'La Sagesse'.

Diem, Werner. 1979. "Studien zur Frage des Substrats im Arabischen". *Der Islam* 56.12–80.

Even-Shoshan, Avraham. 1992. *Ha-millon he-ḥadash* [The new dictionary]. Jerusalem: Qiryat Sefer.

Feghali, Michel. 1918. *Etude sur les emprunts syriaques dans les parlers arabes du Liban*. Paris: Champion.

Fraenkel, Siegmund. 1886. *Die aramäischen Fremdwörter im Arabischen*. (Repr. Hildesheim: G. Olms, 1962.)

Garbell, Irene. 1965. *The Jewish neo-Aramaic dialect of Persian Azerbaijan*. The Hague: Mouton.

Hitti, Philip K. 1957a. *History of Syria*. London: Macmillan.

——. 1957b. *Lebanon in history*. London: Macmillan.

Jastrow, Marcus. 1886. *Dictionary of the Targumim, the Talmud Babli and Yerushalmi, and the Midrashic Literature*. London: Luzac.

Jastrow, Otto. 1969. "Die arabischen Dialekte des Vilayets Mardin (Südosttürkei)". *Supplementa I: XVII Deutscher Orientalistentag vom 21. bis 27. Juli 1968 in Würzburg 1969*, II, ed. W. Voigt, 683–88. Wiesbaden: F. Steiner.

——. 1977. "Gedanken zum zypriotischen Arabisch". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 127.258–86.

——. 1978. *Die mesopotamisch-arabischen* Qəltu-*Dialekte*. I. *Phonologie und Morphologie*. Wiesbaden: F. Steiner.

——. 1981. *Die mesopotamisch-arabischen* Qəltu-*Dialekte*. II. *Volkskundliche Texte in elf Dialekten*. Wiesbaden: F. Steiner.

——. 1985. *Laut- und Formenlehre des neuaramäischen Dialekts von Mīdin in Ṭūr ʿAbdīn*. Wiesbaden: O. Harrassowitz. (1st ed. 1967).

——. 1988. *Der neuaramäische Dialekt von Hertevin (Provinz Siirt)*. Wiesbaden: O. Harrassowitz.

——. 1990. *Lehrbuch der Ṭuroyo-Sprache*. Wiesbaden: O. Harrassowitz.

——. 1994. *Der neuaramäische Dialekt von Mlaḥsô*. Wiesbaden: O. Harrassowitz.

——. 1997. "The Neo-Aramaic languages". *The Semitic languages*, ed. Robert Hetzron, 334–377. New York: Routledge/Kegan Paul.

——. 1999. "Verbformen und ihre Funktionen im arabischen Dialekt von Kəndērîb (Südostanatolien)". *Tempus und Aspekt in den semitischen Sprachen: Jenaer Kolloquium zur semitischen Sprachwissenschaft*, ed. Norbert Nebes, 43–53. Wiesbaden: O. Harrassowitz.

Jiha, Michel. 1964. *Der arabische Dialekt von Bišmizzîn*. Beirut: F. Steiner.

Koutsoudas, Andreas. 1967. "Object particles in Lebanese". *Journal of the American Oriental Society* 87.512–517.

Levin, Aryeh. 1987. "The particle *la* as an object marker in some Arabic dialects of the Galilee". *Zeitschrift für Arabische Linguistik* 17.31–40.

Lewin, Bernhard. 1969. *Notes on Cabali: The Arabic dialect spoken by the Alawis of Jebel Ansariye*. Stockholm: Orientalia Gothoburgensia.

Mas-Latrie, M.L. de. 1861. *Histoire de l'île de Chypre sous le règne des princes de la maison de Lusignan*. Paris.

Mattson, Emanuel. 1911. *Études phonologiques sur le dialecte arabe vulgaire de Beyrouth*. Upsala: K.W. Appelberg.

Mitchell, T.F. 1962. *Colloquial Arabic: The living language of Egypt*. London: English Universities Press.

Newton, Brian. 1964. "An Arabic-Greek dialect". *Word* 20:3 (Supplement). 43–52.

Nöldeke, Theodor. 1888. *Kurzgefasste syrische Grammatik*. Leipzig: Tauchnitz.

——. 1904. *Beiträge zur semitischen Sprachwissenschaft*. Strassburg: Trübner.

Payne Smith, Robert. 1903. *A compendious Syriac dictionary*. Oxford: Clarendon Press.

Poizat, B. 1979. "Une bibliographie commentée pour le Néo-Araméen". *Comptes rendus du Groupe Linguistique d'Etudes Chamito-sémitiques* 17–23. 347–414.

Procházka, Stephan. 2002. *Die arabischen Dialekte der Çukurova (Südtürkei)*. Wiesbaden: O. Harrassowitz.

Röhricht, R. 1895. "Le pèlerinage du moine augustin Jacques de Vérone (1335)". *Revue de l'Orient Latin* 3.178.

Roth, Arlette. 1973–1975. "Le verbe dans le parler arabe de Kormakiti (Chypre)". *Epetēris* 8.21–117. Nicosia: Cyprus Research Centre.

——. 1987a. "Quelques observations sur les emplois du morphème *ta* comme relatif dans le parler arabe de Kormakiti". Πρακτικα Β′ διεθνους Κυπριολογικου Συνεδριου, III, 323–331.

——. 1987b. "Questions à propos de quelques usages 'particuliers' de l'inaccompli à préfixe *p-* dans le parler arabe de Kormakiti (Chypre)". *Matériaux Arabes et Sudarabiques* NS 1.21–34.

——. 1995–1996. "Quelques aspects du fonctionnement de la négation dans le parler arabe de Kormakiti (Chypre): Matériaux et hypothèses". *Matériaux Arabes et Sudarabiques*, NS 7.63–97.

Sabuni, Abdulghafur. 1980. *Laut- und Formenlehre des arabischen Dialekts von Aleppo*. Frankfurt: P. Lang.

Seidel, Ulrich. 1988. "Studien zum Vokabular der Landwirtschaft im Syrischen". *Altorientalische Forschungen* 15.133–173.

Sokoloff, Michael. 1990. *A dictionary of Jewish Palestinian Aramaic of the Byzantine period*. Bar Ilan: Bar Ilan University Press.

Storrs, Ronald. 1930. *The handbook of Cyprus*, 9th ed. London: Christophers.

Tsiapera, Maria. 1969. *A descriptive analysis of Cypriot Maronite Arabic*. The Hague: Mouton.

Ullmann, Manfred. 1970. *Die Medizin im Islam*. Leiden: E.J. Brill.

Vocke, Sybille and Wolfram Waldner. 1982. *Der Wortschatz des Anatolischen Arabisch*. Erlangen: W. Waldner.

Woodhead, D.R. and Wayne Beene. 1967. *A dictionary of Iraqi Arabic*. Washington, D.C.: Georgetown University.

ALEXANDER BORG
(Ben-Gurion University of the Negev)

D

Ḍād

Ḍād is the name of the 15th letter of the Arabic alphabet, denoting nowadays a voiced velarized (emphatic) dental stop /ḍ/ IPA [d̴]. The sound denoted by this letter must have had a special status in Classical Arabic since the language is sometimes called *luġat aḍ-ḍād*, which probably indicates that the grammarians believed this particular sound was unique to Arabic (Ibn Jinnī, *Sirr* I, 214.14: *wa-ʕlam ʾanna ḍ-ḍād li-l-ʕArab xāṣṣa*). Sībawayhi (*Kitāb* II, 405.8–9) describes its place of articulation as being "between the first part of the side of the tongue and the adjacent molars" (*min bayna ʾawwal ḥāfat al-lisān wa-mā yalīhi min al-ʾaḍrās*). The exact interpretation of this passage remains controversial. According to Roman (1983:I, 162–206), it describes a voiced emphatic alveo-palatal fricative, somewhat similar to Polish ź (IPA [ʑˤ] or [ʑ̴]). But Cantineau (1960:55) is probably right in interpreting it as a lateral or lateralized velarized voiced interdental fricative /ḍˡ/ IPA [ð̴ˡ], which is also the position taken by Fleisch (*EI*^{2}, s.v. *ḍād*). This would make it, indeed, a unique sound among the world's languages (cf. Ladefoged and Maddieson 1966:154–156). The interpretation as a lateral is also supported by Sībawayhi's statement (*Kitāb* II, 406.23–24) that if the feature of → *ʾiṭbāq* 'velarization' were taken from *ṭāʾ*, it would become *d*, likewise, *ṣād* would become *s*; but if the *ḍād* were to lose its emphatic feature, it would disappear from speech, since there is no corresponding non-emphatic phoneme (i.e., there is no /dˡ/).

The interpretation of the *ḍād* as a lateral(ized) sound is in line with the reconstruction of the phonemic inventory of the Semitic languages. According to Lipiński (1997:129–132) Arabic /ḍ/ goes back to Proto-Semitic /ṣ́/, which he reconstructs as [ɬ], the voiced counterpart of Proto-Semitic /ś/. The latter soon merged with Proto-Semitic /š/ and is not differentiated in most Semitic languages. Reflexes of Proto-Semitic /ṣ́/ are still pronounced as a voiced (non-emphatic) lateral /ź/ in Modern South Arabian.

It is not surprising that this special sound disappeared in the New Arabic dialects, which all merged the reflexes of Classical Arabic *ḍ* and *ḏ̣*, the resulting phoneme being /ḍ/ in all sedentary dialects, and /ḏ̣/ in all Bedouin dialects. As a matter of fact, Sībawayhi describes a variant of the *ḍād*, called *ḍād ḍaʕīfa* 'weak *ḍād*' (*Kitāb* II, 404.23–405.4), which according to Al-Nassir (1993:45) may represent the emphatic interdental *ḏ̣*, as a first sign of the merger of *ḍād* and *ḏ̣āʾ*.

The only alleged exception to the general merger is the dialect of Daṯīna, which according to Landberg (1901, 1905–1913) has /ḷ/ as reflex of Classical Arabic /ḍ/ and /ḏ̣/ as reflex of /ḏ̣/. Outside the Arabic-speaking world, an unexpected place for a differentiated treatment of the two phonemes is the secret language *kalamo* of Madagascar – a variety of → Malegasy with a large number of Arabic loanwords – in which, for instance, Arabic *ʾabyaḍ* 'white' is represented as *alibiavy*, whereas *(aḏ̣-) ḏ̣uhr* 'afternoon' is *azohora* (Versteegh 2001).

The merger of /ḍ/ and /ḏ̣/ must have taken place rather early, since confusion between the two letters *ḍād* and *ḏ̣āʾ* is one of the oldest documented mistakes in Middle Arabic texts. The correction of this error gave rise to an entire genre in the → *laḥn al-ʕāmma* literature, the treatises about the distinction between *ḍād* and *ḏ̣āʾ*,

e.g. Zanjānī's [d. 471/1078] *Farq bayna ḍ-ḍād wa-ḏ̣-ḏ̣āʾ*. In modern Arab education, schoolchildren are still being taught to learn by heart which words are spelled with a 'tail' and which without.

Grammarians, too, remained interested in the pronunciation of the *ḍād*, as evidenced by the large number of treatises on this topic, for instance the *Risāla fī kayfiyyat an-nuṭq bi-ḍ-ḍād* 'Treatise on the nature of the pronunciation of the *ḍād*', by ʿAlī ibn Sulaymān al-Manṣūrī (d. 1150/1737–1738), edited by Kinberg (2001). This treatise opposes the faulty pronunciation of the *ḍād*, and from the description it becomes clear that this involved both pronouncing it as *ḍ* and as *ḏ̣*. It is not completely clear what the author himself regarded as the correct pronunciation: basically, he simply repeats the description by Sībawayhi. Therefore, this genre of linguistic treatises cannot serve as an argument for or against the lateral character of the *ḍ*.

A stronger argument are the Arabic loanwords in some languages (cf. Steiner 1977; Versteegh 1999). Such loanwords are found in the Ibero-Romance languages, e.g. Spanish *alcalde* < *al-qāḍī* 'judge'; *aldea* 'hamlet' < *aḍ-ḍayʿa* 'estate, hamlet' (cf. Corriente 1977:46, n. 1; but see for arguments against this interpretation Roman 1983:I, 194–199). Because of their limited number, the loanwords in Ibero-Romance are not conclusive evidence in themselves. In some African languages, however, /l/ is the regular reflex of /ḍ/, not of /ḏ̣/, e.g. → Fulfulde *waaju* < *waʿḏ̣* 'sermon', but *farilla* < *farḍ* 'moral obligation', and → Hausa *láfazí* < *lafḏ̣* 'speech, pronunciation', but *háyla* < *ḥayḍ(a)* 'menstruation' (likewise, → Yoruba). Arabic loanwords in East African languages such as → Swahili do not exhibit this differential treatment. The lateral reflexes of Arabic /ḍ/ are also found in some Southeast Asian languages, e.g. in → Indonesian/Malay *lahip*, *laip*, *laif* (but also *daif*) < *ḍaʿīf* 'weak', *kali* (but also *kadi*) < *qāḍī* 'judge', *perlu* 'it is necessary' < *farḍ* 'moral obligation' (cf. the more recent loanword *fardu* with the same sense as the Arabic word), as against *zahir* < *ḏ̣āhir* 'manifest', *zohor* < *ḏ̣uhr* 'noon' (but it should be noted that some loanwords with *z* < *ḏ* have variants with *l*, *lahir*, *lohor*).

It is not immediately clear what al-Andalus, West Africa, and Southeast Asia could have in common to explain the occurrence of this lateral reflex of /ḍ/. One possibility might be that in all of them South Arabians were particularly active, for instance the tribe of Kalb in al-Andalus and North Africa, and merchants from Hadramawt in the Indian Ocean trade. Therefore, the loanwords in these regions might be evidence that the lateral pronunciation survived longer in the south.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Ibn Jinnī, *Sirr* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *Sirr ṣināʿat al-ʾiʿrāb*. Ed. Ḥasan Hindāwī. 2 vols. Damascus: Dār al-Qalam, 1985.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. 2 vols. Cairo: Bulaq, 1317/1899–1900. (Repr. Baghdad: Al-Muthanna, n.d.)

Manṣūrī, *Risāla* = ʿAlī ibn Sulaymān al-Manṣūrī, *Risāla fī kayfiyyat an-nuṭq bi-ḍ-ḍād*. Naphtali Kinberg, *Studies in the linguistic structure of Classical Arabic*, ed. Leah Kinberg and Kees Versteegh, 197–267. Leiden: E.J. Brill.

Zanjānī, *Farq* = ʾAbū l-Qāsim Saʿd ibn ʿAlī az-Zanjānī, *al-Farq bayna ḍ-ḍād wa-ḏ̣-ḏ̣āʾ*. Ed. Mūsā Banāȳ ʿAlwān al-ʿAlīlī. Baghdad: Ṁaṭbaʿat al-ʾAwqāf wa-š-Šuʾūn ad-Dīniyya, 1983.

SECONDARY SOURCES

Al-Nassir, Abdulmunim Abdalamir. 1993. *Sibawayh the phonologist: A critical study of the phonetic and phonological theory of Sibawayh as presented in his treatise* Al-Kitāb. London and New York: Kegan Paul International.

Cantineau, Jean. 1960. *Etudes de linguistique arabe (Mémorial Jean Cantineau)*. Paris: Klincksieck.

Fischer, Wolfdietrich. 1968. "Die Position von *ḍ* im Phonemsystem des Gemeinsemitischen". *Studia orientalia in memoriam Caroli Brockelmann*, 55–63. Halle: Martin-Luther University Halle-Wittenberg.

Ladefoged, Peter and Ian Maddieson. 1996. *The sounds of the world's languages*. Oxford: Blackwell.

Landberg, Carlo de. 1901, 1905–1913. *Etudes sur les dialectes de l'Arabie méridionale*. I. *Hadramoût*. II. *Daṯīna*. Leiden: E.J. Brill.

Lipiński, Edward. 1997. *Semitic languages: Outline of a comparative grammar*. Leuven: Peeters.

Roman, André. 1983. *Etude de la phonologie et de la morphologie de la koine arabe*. 2 vols. Aix-en-Provence: Université de Provence.

Steiner, Richard C. 1977. *The case for fricative laterals in Proto-Semitic*. New Haven, Conn.: American Oriental Society.

Versteegh, Kees. 1999. "Loanwords from Arabic and the merger of *ḍ/ḏ̣*". *Israel Oriental Studies* 19.273–286.

——. 2001. "Arabic in Madagascar". *Bulletin of the School of Oriental and African Studies* 64.177–187.

KEES VERSTEEGH (University of Nijmegen)

Damascus Arabic

1. GENERAL

Damascus Arabic (*llahže ššāmiyye*) is spoken in Damascus, capital of the Syrian Arab Republic (1.6 million people in 2004; 3.5 million including Greater Damascus, with an important proportion of non-native speakers: immigrants from various parts of the country, Palestinians). In the old villages of the surrounding Ġūṭa different dialects are spoken, which have not yet been studied. Damascus Arabic is well understood in the whole country, and in Lebanon, Jordan, and Palestine.

Contiguous to the Damascus area are the Qalamūn dialects in the north, the Ḥōrān dialects in the south, and 'steppe-range' and *Šāwi* (Bedouin) dialects in the east. Being the dialect of the capital, and as such largely used in radio and television programs, Damascus Arabic is prestigious, 'neutral' in many respects, and can be qualified as the Syrian koinic dialect. Used between the two World Wars in satirical journals, it is nowadays rarely present in printed literature, except for theatrical plays; cinema journals publish screenplays. Plays of (now extinct) shadow theater have been preserved and published (Kayyāl 1987). Some works are very popular, such as theater and radio plays by Ḥikmat Muḥsin (1910–1968) and plays and films by Durayd Laḥḥām (born 1934), available on tapes, videotapes, and DVDs. Poetry, mainly from the 19th century, is found in manuscripts. Collections of poems are sometimes published.

Damascus Arabic is typically a 'Syro-Lebanese' sedentary dialect. Minor differences, mainly lexical, can be observed between the Muslim, Christian, and (now almost totally expatriate) Jewish communities.

Historical information on the dialect can be gathered from → Middle Arabic texts, sometimes highly colloquializing, written by Damascene authors. Shadow theater plays, and many manuscripts of popular epic literature (Bohas and Zakhariya 2000) date back to at least the 19th century. It is nevertheless difficult to draw a documented history of the dialect, which seems to be, for the two last centuries anyway, remarkably stable.

Damascus Arabic has been studied for over a century, and is one of the best described Arabic dialects. Grammars are Grotzfeld (1965), Ambros (1977), and especially Cowell (1964). Still useful handbooks are Cantineau and Helbaoui (1953), Kuhnt (1958), Kassab (1970), and Ferguson (1961). An important syntactic study is Bloch (1965). Among collections of texts Bergsträsser (1924; with an excellent phonetic study), and Bloch and Grotzfeld (1964) deserve special mention; see also Grotzfeld (1965: 118–149) and, in Arabic script, Turžmān (1978), Kayyāl (1987), and al-ʾAswad (1998, 2002). Stowasser and Ani (1964) is a very good dictionary (English-Damascus Arabic), sometimes reflecting a classicizing layer (the announced Arabic-English version has never appeared); see further Salamé and Lentin (forthcoming).

2. LINGUISTIC DESCRIPTION

2.1 *Phonology*

2.1.1 Consonants

b, *(ḅ)*, *(p)*, *m*, *(ṃ)*, *w*, *f*, *(v)*, *t*, *ṭ*, *d*, *ḍ*, *n*, *l*, *(ḷ)*, *r*, *(ṛ)*, *s*, *ṣ*, *z*, *ẓ*, *š*, *ž*, *(ç)*, *k*, *(g)*, *y*, *x*, *ġ*, *q*, *ḥ*, *ʿ*, *h*, *ʾ*.

/ḅ, ṃ, ḷ, ṛ/ are generally phonetically conditioned variants of /b/, /m/, /l/, /r/ but also (marginal) phonemes.

/p, v, č, g, q/ appear mainly in borrowings: *kɔmpyūtər* 'computer', *pāyp* 'pipe' (/p/ is generally not replaceable by /b/); *vīza* 'visa', *təlvəzyōn* 'television', *vītēs* 'gear change' (/v/ is generally replaceable by /f/ or in older borrowings by /b/; but only e.g. *narvazni* 'he made me nervous'); Turkish /č/ being represented by /š/, /č/ is rather rare; /g/ appears in borrowings: *sīgāra* 'cigarette', *grēfōn* 'grapefruit' (and rarely alternates with /k/) or in 'expressive' words: *garr* 'he jabbered'; /q/ is frequent in borrowings from classical Arabic: *musaqqafīn* 'intellectuals'; it often alternates freely with */ʾ/*: *mūsīqa* ~ *mūsīʾa* 'music'; the distribution is sometimes unpredictable: *ṣadīʾ* 'friend', plural *ʾaṣdiqāʾ*.

The historical interdentals **ṯ*, **ḏ*, and **ḏ̣* correspond with /t, d/, and /ḍ/, the latter also representing **ḍ*. In borrowings from Classical Arabic **ṯ*, **ḏ*, and **ḏ̣* are most often represented by /z, s/ and /ẓ/. Hence, a historical root can split into two distinct ones: *tāni* 'second' – *sānawi* 'secondary', *dāʾ* 'he tasted' – *zōʾ* '(good) taste'. The same situation obtains with /ʾ/ (< **q*) vs /q/: *daʾīʾa* 'a second [of time]' – *daqīqa* 'precise [fem.sg.]', *ʾarᵊn* 'horn' – *qarᵊn* 'century', which establishes the phonemic status of /q/. *ʾ is generally represented by /y/, e.g. in verbs IIIʾ (*xabba* 'he hid') or in the active participle of verbs II*w/y*: *nāyem* ‘sleeping’; it has been preserved in verbs Iʾ (*ʾakal* ‘he ate’) and exceptionally in *saʾal* ‘he asked’. /ʾ/ thus represents both preserved *ʾ and **q* (*ʾaššar* ‘he indicated’ or ‘he pealed’.

The new velarized consonants /ṛ/, /ḷ/, /ḅ/, /ṃ/ have been phonemicized. *ṛ* is much used (also the most frequent phonetic realisation of /r/): *barra* ‘he proved (s.o. to be) innocent’, *baṛṛa* ‘outside’, *nhār* ‘he broke down’, *nhāṛ* ‘daytime’. Examples for the other phonemes: *ʾalla* ‘he told her’, *ʾaḷḷa* ‘God’; *ḅāḅa* ‘daddy’, *bāba* ‘her door’; *mayy* [female first name], *ṃayy* ‘water’. In some roots or words, ‘emphasis’ has disappeared: *rakad* ‘he ran’, *dərs* ‘tooth’, or voicing has occured: *ẓġīr* ‘small’.

‘Emphatic’, i.e. velarized articulation, tends to be more pronounced in traditional quarters like Mīdān, as well as among Christians.

2.1.2 Vowels

Short vowels: *i*, *e*, *a*, *o*, *u*, *ə*; long vowels: *ī*, *ē*, *ā*, *ō*, *ū*.

In French borrowings, nasal vowels /ã/, /õ/, /ɛ̃/ and /ü/ occur: *ʾasãsēr* ‘lift’, *selülēr* ‘mobile phone’. The two main short vowels are /a/ and /ə/; /e/ and /o/ merge into /ə/ in non-final syllables and alternate freely in some verbal forms. Due to the increasing number of borrowings from Standard Arabic, /i/ and /u/ occur in non-final syllables: *bināye* ‘building’, *muxrež* ‘(film) director’ and can switch to /e/ and /o/ or zero; doublets also obtain: *mʿallem* ‘boss’ versus *muʿallem* ‘schoolteacher’.

a* is mainly represented by /a/, **i* by /e/ and /ə/,u* by /o/ and /ə/. /ē/ and /ō/ represent *-*ay* and *-*aw*. Final **āʾ* changed into -*e* in nouns: *kəre* ‘rent’, *šəte* ‘winter’ [construct state -*ā*-], and into -*a* in the feminine of color and deficiency adjectives: *ḥamra* ‘red’.

Although the main phonological opposition is /a/–/ə/, /a/ is often very close, phonetically, to /ə/, especially in closed syllables. Since only /a/ and /ə/ occur in non-final syllables, /ə/ has a wide range of realizations, close to [e] or [i] (after palato-alveolar fricatives), [a] (contiguous to pharyngeals), or [o] or [u] (after labials or ‘emphatics’). The ‘feminine’ suffix *-*a(t)* has split into two allomorphs: -*a* after back and ‘emphatic’ consonants, -*a* and/or -*e* after /r/ according to phonetic environment, -*e* otherwise (but construct state always -*et*). The feminine ending of the 3rd person perfect is also affected by → *ʾimāla* (*-*at* > -*et*).

/i/ and /u/ occur only in final open syllables (except in borrowings); /i/ is also found before -*y(y)*- where it can be considered a realization of /ə/). /e/ and /o/ switch to /ə/ in non-final syllables. Consequently, among short vowels only /a/ and /ə/ can be stressed. Final /i/ and /u/, usually described as shortened **ī* and **ū*, can be given phonemic status on the basis of pairs like *katabti* ‘you [fem. sg.] wrote’ ~ *katabtī* ‘you [fem. sg.] wrote it’, *katabu* ‘they wrote’ versus *katabū* ‘they wrote it’.

‘Popular’ pronunciation is often characterized by a more back articulation of vowels and a kind of vowel harmony. Women often display a more contrasted and less back articulated range of vocalic timbres, giving the false impression that their velarized consonants are less ‘emphatic’.

2.1.3 Diphthongs

aw* and **ay* are monophthongized into /ō/ and /ē/: *tōr* ‘bull’, *zēt* ‘oil’. Morphological constraints can prevent monophthongization: *stawʿab* ‘he grasped the meaning’, *ġayrān* ‘jealous’, *mawžūʿ* ‘being in pain’. Diphthongs also occur in borrowings from Standard Arabic (*žawle* ‘round’) and stylistic doublets obtain: *mōsem* ~ *mawsem* ‘season’, *ḥēwān* ~ *ḥay(a)wān* ‘animal’, *ʾēwa* ~ *ʾaywa* ‘yes’. Irregular are negation *mu* (<maw* <**māhu*), *kīf* (< **kayf*) ‘how?’ and *ši* (< **šayʾ*) ‘thing’.

Synchronically, *ey* and *əy* turn into *ī*: *ḥāki* (< *ḥākī* < **ḥākey*) ‘having spoken’, *bīḍ (<*bəyḍ)* ‘white [pl.]’, and *əw* to *ū*: *tūṣāye* (< **təwṣāye*) ‘order’; occasionally, *əw* is maintained: *lᵊwlād* ~ *lūlād* ‘the children’. The groups -*aww*- and -*ayy*- are preserved (even when closing syllables, then generally pronounced -*aw*-, -*ay*-).

2.1.4 Syllable

Syllable types are Cv, CvC, CvCC (generally pronounced CvC in non final syllable if CC is a geminate); Cv̄, Cv̄C; Cv̄CC (only CāCC, with final geminate: *mawādd* ‘substances’); CCvC, CCvCC; CCv̄, CCv̄C; CCCvC, CCCvCC, CCCv̄, CCCv̄C occur mainly in Form X derived verbs and foreign borrowings.

When the clitic preposition *l*- (‘to’) + personal pronoun is suffixed to a verbal form with final Cv̄C or CCv̄C syllable, v̄ is shortened *(ā > a*, *ī ~ ū > ə)*: *ʾāl* ‘he said’, *ʾallo* ‘he told him’.

A cluster CC (and sometimes CCC) can occur in initial position, and in final position, especially when one of the two consonants is a resonant; more commonly an epenthetic [ə] appears between the two consonants. Epenthetic [ə] appears, in the same conditions, between the two first consonants of a -CCC- cluster (except if they are identical). Clusters of more than three consonants are separated by word boundaries.

Stress falls on the first v̄, or v in vCC, from the end of the word, or, if there is no such vowel, on the first one (without considering proclitics). 'Exceptions' to this rule are imperfect and participles of Forms VII and VIII of sound and roots III*y*: *byəḥtə́rem* 'he respects', *məftə́ker* 'thinking that' (borrowings from Standard Arabic can have two accentuations: *múḥtaram*/*muḥtáram* 'of good quality'); 3rd person feminine singular perfect of certain verb types: *šāfə́to* 'she saw him'. Some speakers also have, for the sound verb: *katbə́to* (~ *kátəbto*) 'she wrote it' and for feminine nouns CaCaCe/a + pronominal suffix with initial V: *baṣlə́ti* (~ *báṣəlti*) 'my onion'. This prevents a morphological mark from totally disappearing.

Voicing and devoicing of consonants occur by assimilation, e.g. in Forms V and VIII: *məddayyeʾ* (< **mətdayyeʾ*) 'feeling uncomfortable'; *štamaʿu* (< **žtamaʿu*) 'they gathered'. /n/ is labialized before /b/: *žamb* 'next to', *məm baʿḍ* 'of each other', and assimilates to /r/: *mərrūḥ* 'we go'. /ʿ/ + /h/ is usually pronounced [ḥḥ]: *maḥḥa* 'with her'. All these processes also occur in sandhi, except those involving velars.

The article *l*- assimilates to the 'sun-letters', which include /ž/ besides the usual consonants. Consonantal assimilation occurs mainly between sibilants and palato-alveolar fricatives: *zazar* 'carrots', *zōzi* 'my husband'; many of these forms are popular and old-fashioned. So are most words with metathesis (*səddāže* ~ *səžžāde* 'mat'), but some are usual: *žōzi* 'my husband' (*zōži* is a bit classicizing) or lexicalized: *starža* (< **stažra*) 'he dared'.

Velarization spreads very easily, often over the whole word; /i, e, ə/ tend to restrain it.

No pausal forms are found (or preserved). Typically Damascene is the lengthening of the last vowel, even epenthetic, of interrogative and exclamative utterances: *šəfə̄t* 'did you see that?!'

Damascus Arabic is a 'différentiel' dialect: /a/ is much more stable than /e/ and /o/. But nonstressed **a* is elided, e.g., in most CaCīC pattern nouns: *ʾžīr* 'apprentice' and in 3rd pers. sg. fem. perfect of Form I CaCaC and Forms VII and VIII verbs: *katbet* 'she wrote', *ḥtarʾet* 'she burned'. Synchronically, /ə/, archiphoneme of /e/ and /o/, is elided in Cv syllables when unstressed: *šāyfak* 'having seen you', *sayyārto* 'his car'.

Final vowels are lengthened by suffixation of a personal pronoun, alone or preceded by *l* (*w*, *y*, or *h* being inserted if the suffix has initial vowel). In the case of feminine singular active participles referring to a 2nd person, the verbal ending *-tī-* is inserted: *šāyəftīha* 'did you [fem.] see her?'.

2.2 *Morphology*

There is no gender distinction in the 2nd and 3rd persons plural of verbs or pronouns.

2.2.1 Pronouns

2.2.1.1 Personal pronouns

Table 1. Independent personal pronouns

		Singular	Plural
1st person		*ʾana*	*nəḥna*
2nd person	masc.	*ʾənte*	*ʾəntu*
	fem.	*ʾənti*	
3rd person	masc.	*huwwe*	*hənne(n)*
	fem.	*hiyye*	

nəḥna has a variant *ləḥna*.

Reflexive pronoun: *ḥāl-* + personal pronoun; *laḥālo* 'on his own; by himself'.

Table 2. Bound personal pronouns

		Singular		Plural
		. . . V(C)C+	. . . V	
1st person	(noun)	*-i*	*-yi*	*-na*
	(verb)	*-ni*		
2nd person	masc.	*-ak*	*-k*	*-kon*
	fem.	*-ek*	*-ki*	
3rd person	masc.	*-o*	*-Ø*	*-(h)on*
	fem.	*-(h)a*	*-(h)a*	

-ni is also used after the particle *yā-*, the preposition *b-* in the idioms *šu bani*ʔ/*ma bani ši* 'what/nothing is wrong with me', *ləssā(ʿ)t-* and *baʿd-* ('still'), *fī-* 'to be able', the presentative *šaʿ-*.

2.2.1.2 Demonstratives

Adjectival demonstratives are followed by the article (incorporated in *hal-*).

2.2.1.2.1 'This'

Adjectives: *hal-*, assimilated to sun-letters; *hāda*, feminine *hādi* or *hayy*, plural *hadōl*. They are always in front, except when they are associated: *harrəžžāl had(a)* 'this man'. Pronouns: *hāda* (or *hād* before #), feminine *hādi* or *hayy(e)*, plural *hadōl*.

2.2.1.2.2 'That'

Adjective (in front) and pronoun: *hadāk*, feminine *hadīk(e)*, plural *hadolīk* or *hadənk(e)*. Variants with *-(e)* are used only for pronouns.

2.2.1.2.3 Presentatives

hayy žəbtəllak yā: 'here you are, I brought it to you'; *lēk-* or *šaʿ-* (< **qšaʿ* 'look at!') + personal pronoun: 'here he is', etc.

2.2.1.3 Relative pronoun

Invariable *ʾəlli/(ə)lli/yəlli/halli*; also used as relative adjective.

2.2.1.4 Interrogative pronouns

Interrogative pronouns: *mīn* 'who?'; *šu* 'what?' (also *ʾēš*, mainly after preposition), normally in front in verbal sentences, free position in nominal sentences. Adjectives *ʾanu* (optional feminine *ʾani*), *ʾay(y)a* (or classicizing *ʾayy*), 'which?'.

2.2.1.5 Reciprocal

baʿḍ (not repeated) + optional (in most cases) bound personal pronoun.

2.2.1.6 Concessive

ʾēmən 'whoever'; *šu ma*, *ʾēšmən* 'whatever'.

2.2.2 Adverbs

hallaʾ (< * *hal-waqt*) 'now'; *baʿdēn* 'afterwards'; *bakkīr* 'early'; *bukra* 'tomorrow'; *mbāreḥ* 'yesterday'; *ʿala bukra* 'early in the morning'; *waʾta* 'at that time'; *ṣṣəbᵊḥ* 'in the morning' or 'this morning'; *ʿala ṭūl* 'always'; *ləssa*: assertive or interrogative 'still' or 'not yet' (according to context; *ləssə((ʿ)t)-* before personal pronouns); *baʿd-* 'still' or 'not yet'; *kəll mā-l-* + personal pronoun: 'he [etc.] never stops doing . . .'; *ʾēmta* 'when?'.

hōn(e): 'here'; *hnīk(e)* (originally non-Muslim variant: *honīk[e]*); 'there'; *wēn* 'where?'; *mnēn* 'where from?'.

ʾawām 'quickly'; *ḥāže* 'enough!'; *bas* 'only'; *kamān(e)* 'also'; *hēk(e)* 'like this'; *ktīr* 'very' (free position); *dəġri* 'straight on'; *šlōn* 'how?' (*šlōn ʿrəft* 'how did you know?', *ʿrəft ᵊšlōn* 'do you see what I mean?'); *lēš* 'why?'; *ʾaddēš* 'how much?'; *kam* 'how many . . .?' must be specified by a singular noun (or by *wāḥed* 'one = unit').

2.2.3 Particles

2.2.3.1 Definite article

The definite article is *l*; it can be separated by epenthetic /ə/ from the first consonant of its noun: *lətlāte* 'the three'.

2.2.3.2 Genitive marker

'Genitive markers': *tabaʿ* (invariable, optionally *tabaʿīt* after a singular [masc. or fem.] and *tabaʿāt* ~ *tabaʿūn/l/t* after a plural); *šīt* (see 2.3.1)

2.2.3.3 Negations

There is no discontinuous negation (with *-š*). *ma* is the verbal negation, *la* is used only with prohibitive, or in *la . . . u la* 'neither . . . nor'; *mu* (less frequently *ma*, *maw*, *mi/may*) is used with nominals; *mu* also negates a whole sentence. In verbless sentences, subject personal pronouns are negated with connective *-l-* or *-n-*: *mālo hōn* 'he's not here'. 'No' is *la (laʾ, laʾa)*. *mnōb* and *ʾabadan* are intensifiers: '(not) at all'. 'No more' is *ma . . . ʿād* or *ma . . . baʾa* (negation + 'he did again' or 'he stayed', conjugated or fossilized). The existential particle is *fi* 'there is'.

2.2.3.4 Prepositions

2.2.3.4.1 Directional

la 'toward, to' (including goal), *ʿa(la)* 'in the direction of', *laʿand* 'to somebody's place'.

2.2.3.4.2 Locative

b(i) 'in, at' (*fī-* before personal pronouns), *ʿa(la)* 'on', *fōʾ* 'above', *žuwwāt* 'inside', *barrāt* 'outside', *ḥawāli* (construct state *ḥawālē-*) and *dāyer ma dār* 'around', *žamb* 'next to', *ʾəddām* 'in front of, next to', *ʿand* 'by'.

2.2.3.4.3 Attributive

la- 'to' (*l-* before article); before personal bound pronoun: *l-* or *laʾəl-* and *-əll-* after -CC when clitic, *ʾəl-* when separated from verb, or predicative in verbless sentences; *m(ən)šān* ‘for the sake of’.

2.2.3.4.4 Others

ʿala ‘in the state of’, *m(ə)n* ‘from’, *ʿan* ‘removed from’, *maʿ* and *wiyyā-* + personal pronoun ‘(together) with’, *b(i)-* ‘with’ (instrumental), *la* and *m(ən)šān* ‘because of’, *mətᵊl* ‘as’, *ʾadd* ‘the size/weight/ . . . of’.

2.2.3.5 Conjunctions

w ~ *u* ‘and’, *ʾaw* ‘or’, *wəlla* and *ləmma* ‘or?’, *ya . . . ya* ‘either . . . or’, *bas(s)* ‘but’, *laʾənno* or *ḥākem* ‘because’. *lamma*, *waʾᵊt*, *waʾtəlli*, *bas(s)* ‘when’, *ṭāla ma* ‘as long as’, *kəll ma* ‘every time that’, *(ʿa)la bēn ma* ‘until’, *ʾaḥsan ma* ‘rather than’, *la*, *ḥatta*, *la ḥatta*, and *mənšān* ‘in order to’. *la* introduces many types of clauses: *ṭalabak la tṣallḥo* ‘he asked you to fix it’, *ʾənšaḷḷa la tžəbli yā* ‘may you bring it to me!’, *lēš la* ‘why . . .?’. Complement clauses are introduced by Ø, *ʾenno* (or *ʾenn-* + personal pronoun). *la-* is used after verbs of fear, *əlli* after verbs, etc. meaning ‘(to be) (un)happy that’: *frəḥt* ~ *mnīḥ əlli* ‘I am happy ~ it is fine that’, and *ma* after ‘to forget’.

The vocative particle is *ya*. Many address forms involving kinship terms are ‘bipolar’: a mother would use *ṃāṃā* ‘mummy’ when talking to any of her children. A few have *-o*: *ʿammo* ‘uncle’(address form to peers or elders).

2.2.4 Noun

Most feminine nouns have the marker *-e/a*, a few do not: *ṃayy*, ‘water’, *ʿēn* ‘eye’, *daʾᵊn* ’beard’, *ʾəmm* ‘mother’, etc. Some nouns have both genders: *məstašfa* ‘hospital’. Some usual patterns are productive: CaCCāC (names of professionals), maCCaC (place names), CaCCāC(e/a) (names of instruments). Suffixes are *-xāne* (place names), *-i* and *-ži* (professional names or pejorative: *səkarži* ‘drunkard’), *-a/e* (*-āye* after vowel) for unit nouns. *-īn*, *-āt*, *-e*, and *-iyye* are ‘external’ plural suffixes. *-āt* is also the mark of ‘deictic’ plural: *xalaṣu lə nbīdāt* ‘is there no wine left?’; *šu haššōbāt* ‘what a heat!’. Mixed plurals (CCūCāt) are frequent. → Pseudo-dual is used for names of body parts: *ʾīdēn* ‘hands’, *ʾīdēhon* ‘their hands’ (*ʾīd(t)ēno* ‘his two hands’). Diminutives are not frequent: *ṭayyūb* ‘very kind’, *baṭṭūn*, *bṭēn*, ‘(sweet) little belly’, except for proper names: *nabbūl(e)* (< *Nabīl*). Particular vocatives: *yāmo* ‘Mum’, *yābi* ‘Dad’.

2.2.4.1 Adjectives

Frequent patterns are CəCᵊC, C(a)CīC, məCCeC. Suffixes are *-i*, *-āni* (*nəṣṣāni* ‘in the middle’), and *-ži*. Color adjectives: ʾaCCaC, feminine CaCCa, plural CəCᵊC (deficiency adjectives have generally CəCCān plurals). Elatives (invariable ʾaCCaC) are based on various or even non attested simple forms: *ʾaždab* (< *ʾaždab*) ‘more stupid’, *ʾaḥsan* ‘better’ (**ḥasan*), *ʾaḥwan* (< *ḥēwān*) ‘more idiotic’, *ʾamyaz* (< *mumtāz*) ‘more perfect’.

2.2.5 Numerals

wāḥed, fem. *waḥde* ‘one’ is the absolute form (otherwise the noun alone is used). Postponed to the noun: *mara waḥde* ‘(only) one women’.

tnēn, fem. *təntēn* (masc. pl. *tnēnāt-*, fem. *təntēnāt-* with personal pronouns) are absolute forms (otherwise the dual is used for nouns) which can appear for insistence, when the noun is not present (*tnēn ᵊkbār* ‘two big ones’), and in particular cases: *təntēn nəswān* ‘two women’ (**martēn*). The dual also means ‘some’: *hal kam bētēn šəʿᵊr* ‘these few verses’, and is an intensifier: *ʾalʿanēn* ‘still worse!’.

From 3 to 10, three series are used:

i. *tlāte*, *ʾarbʿa*, *xamse*, *sətte*, *sabʿa*, *tmāne (tmānye)*, *təsʿa*, *ʿašra* are absolute and attributive forms, also used (like *wāḥed* and *tnēn*) with specificative: *tlāte ʾahwe* ‘three coffees’. *tlāte* and *tmāne* become *tlāta* and *tmāna* before the tens (*tmāna w xamsīn* ‘58’).
ii. *tlət(t)*, *ʾarbaʿ*, *xams*, *sətt*, *sabʿ*, *tmən(n)*, *təsʿ*, *ʿašr* with counted noun in plural (except with *miyye* for hundreds, before which *tlət(t)* and *tmən(n)* become *tlāt* and *tmān)*.
iii. *tlətt*, *ʾarbaʿt*, *xamst*, *sətt*, *sabʿt*, *tmənt*, *təsᵊʿt*, *ʿašᵊrt* with a few counted nouns (*ʾəyyām* ~ *iyyām*, ‘days’, *ʾəšhor* ‘months’, *ʾəṣṭor* ‘lines’, etc.).

From 11 to 19 (forms with *-aṛ* = construct state): *ḥdaʿš(aṛ)* or *ʾida(ʿ)š(aṛ)*, *ṭnaʿš(aṛ)*, *tləṭṭaʿš(aṛ)*, *ʾarbaʿṭa (ʿ)š(aṛ)*, *xamsṭaʿš(aṛ)*, *səṭṭaʿš(aṛ)*, *saba(ʿ)ṭa(ʿ)š(aṛ)*, *tmənṭaʿš(aṛ)*, *təs(ʿa)ṭaʿš(aṛ)*; counted noun in singular.

Tens: *ʿəšrīn*, *tlātīn*, etc. (counted noun in singular), as with *miyye* ‘100’ (*mīt* in construct state) and *ʾalf*, plural *(ʾ)ālāf* ‘1000’.

Ordinals: *ʾaw(wa)lāni*, *tāni*; from 3th to 10th pattern CāCeC; others same as cardinals.

Unit nouns are often formed with count nouns: *ḥabbet baṭāṭa* ‘a potato’, *rġīf xəbᵊz* ‘a [loaf of] bread’, *ʿūd kəbrīt* ‘a match’.

2.2.6 Verb

Form I verbs have two perfect types: CaCaC, generally active verbs and CəCeC, generally ‘middle’ verbs. The imperfect thematic vowel is /a/, /e/, or /o/. The most common types are CaCaC, yəCCeC ~ yəCCoC (yəCCaC before or after pharyngeal or laryngeal) and CəCeC, yəCCaC. The two types can occur for a same root: *xəreb*, *yəxrab* ‘it was destroyed’, *xarab*, *yəxreb* ‘he ruined’.

Form II is the most productive; it can be causative, frequentative, close to Form I but more concrete in the description of the process (*wəṣel* ‘he arrived’, *waṣṣal* ‘he reached his destination’), or denominative. Form IV verbs are borrowed from Standard Arabic and have byəCCeC and/or biʾaCCeC imperfect (*biʾaslem* ‘he becomes a Muslim’); a few are new creations: *biʾaflem* ‘it is mere acting of him’), some are probably phonetically conditioned: *ʾaḥka* = *ḥaka* ‘he talked’. Form V is sometimes passive of Form II, often ‘middle’: *txabba* ‘he hid’. Form IX is derived from a few color adjectives (most deficiency adjectives have VIIth forms). The prefix of Forms V and VI is *t-*. Form VII is passive and is generally freely shaped; it is used in impersonal constructions: *ma byənšəbeʿ mənnon* ‘you never get tired of them’, and often expresses potentiality: *ma byənšəreb* ‘it is undrinkable’. Form VIII is not productive, whereas Form X is, mainly for (often transitive) ‘middle’ verbs: *stafkar* ‘he recalled the memory of’. With same semantism are a few verbs of combined Forms X+II: *stlaʾʾa* ‘he caught in midair’ and X+III: *stnāwal* ‘he grasped’.

2.2.6.1 Paradigms

Imperative (2nd pers.): sg.: masc. *šrāb*, *ktōb*, fem. *šrabi*, *ktəbi*; pl. *šrabu*, *ktəbu*.

Table 3. Imperfect conjugation

			Type 1 *a*	
			b-imperfect	Ø-imperfect
3rd	sg.	masc.	*byəšrab*	*yəšrab*
		fem.	*btəšrab*	*təšrab*
	pl.		*byəšrabu*	*yəšrabu*
2nd	sg.	masc.	*btəšrab*	*təšrab*
		fem.	*btəšrabi*	*təšrabi*
	pl.		*btəšrabu*	*təšrabu*
1st	sg.		*bəšrab*	*ʾəšrab*
	pl.		*mnəšrab*	*nəšrab*
			Type 2 *e/o*	
			b-imperfect	Ø-imperfect
3rd	sg.	masc.	*byəktob*	*yəktob*
		fem.	*btəktob*	*təktob*
	pl.		*byəkᵊtbu*	*yəkᵊtbu*
2nd	sg.	masc.	*btəktob*	*təktob*
		fem.	*btəkᵊtbi*	*təkᵊtbi*
	pl.		*btəkᵊtbu*	*təkᵊtbu*
1st	sg.		*bəktob*	*ʾəktob*
	pl.		*mnəktob*	*nəktob*

The vowel of the personal prefixes is /ə/, except for a few verbs IIʿ (*byaʿref*, *byaʿmel*, *byaʿṭi*).

The standard imperfect is the *b*-form; the Ø-imperfect is the dependent (‘subjunctive’) form.

Table 4. Perfect conjugation

			Type 1 CaCaC	Type 2 CəCeC
3rd	sg.	masc.	*katab*	*šəreb*
		fem.	*katbet*	*šərbet*
	pl.		*katabu*	*šərbu*
2nd	sg.	masc.	*katabt*	*šrəbt*
		fem.	*katabti*	*šrəbti*
	pl.		*katabtu*	*šrəbtu*
1st	sg.		*katabt*	*šrəbt*
	pl.		*katabna*	*šrəbna*

2.2.6.2 Participles

Form I: active: CāCel, and CaCCān for ‘middle’ CəCeC verbs; passive: maCCūC.

Among derived forms are worth mentioning VII mənCəCel and VIII məCtəCeC (passive məCtaCaC).

2.2.6.3 Verbal nouns

Verbal nouns are not very frequently used; original patterns are CCīCe/a (*ktībe* ‘writing’) and, for roots IIIy, CaCwe (*bakwe* ‘crying’). With Form VII verbs, Form I verbal nouns are normally used, and Form II and III (or I) verbal nouns with Forms V and VI verbs.

2.2.6.4 Weak verbs

Initial *ʾ, **w* and **y* are preserved; final *ʾ and **w* have switched to *y*.

Verbs II gem.: the 1st and 2nd persons of the perfect are assimilated to verbs III*y* (type 1): *ḥaṭṭēt* ‘I put’.

The imperfect of verbs Iʾ has *ā*: *byākol* ‘he eats’; imperative: *kōl*.

The imperfect of verbs I*w* is *byūṣal* or *byəṣal* ‘he arrives’.

The imperative of verbs II*w*/*y* is Cū/īC. In the imperfect of Forms VII and VIII, *ā* is shortened into *ə*: *rtəḥᵊt* ‘I relaxed’.

Verbs III*y* have two perfect types: CaCa (generally yəCCi in the imperfect: *ḥaka*, *ḥaket*, *ḥaku*, *ḥakēt*, etc., *yəḥki*); CəCi (generally yəCCa in the imperfect: *nəsi*, *nəsyet*, *nəsyu*, *nsīt*, etc., *yənsa*).

Imperative: ʾəCCv or CCv̄ (exception: *ʿaṭi* ‘give!’); passive participle məCCi.

Imperfect of Forms VII and VIII has two variants: *byənḥaka* ~ *byənḥəki* ‘it is said’.

Verbal noun of Form II: təCCāye.

‘Irregular’ verb: *ʾəža* ‘he came’: *ʾəža*, *ʾəžet*, *ʾəžu*, *(ʾ)žīt*, etc.; *byəži*, *btəži*, *byəžu*, *bəži*, *mnəži*, etc. Active participle: *žāye*, plural *žāyīn*.

Quadriradicals verbs are from genuine roots, reduplicative (especially of roots II gem.), augmented triradicals (CarCaC, CawCaC, CaCwaC, CayCaC, CaCCan . . .), denominative or former Form IV verbs. A few CaCCa are frequently used (*farža* ‘to show’), verbal noun tCəCCi. t-CaCCaC is either passive or synonymous with CaCCaC. Both are conjugated like Form II verbs.

2.3 *Syntax*

2.3.1 Noun

Definiteness: *hal-* is often used as a determiner in vivid narration: *rkəbt bi hal bāṣ* ‘I got on the bus’.

An indefinite phrase (construct state) noun + noun is generally made definite by defining the second one: *šarāb əttūt* ‘the mulberry juice’, *ʾamīṣ nōma* ‘her nightdress’, *maḥḍar haḍḍabᵊṭ* ‘this report’, sometimes both: *kāstak əlḥalīb* ‘your glass of milk’; some noun + article + noun phrases do not vary: *mēlt əlḥəsᵊn* ‘a ~ the cast in one eye’. *Badal*-constructions: *ṭṭāwle lxašab* ‘the wooden table’. The article of a definite noun qualified by an adjective (or relative clause) can be dropped in certain cases (quasi-proper names, designation of an object among a known limited series, etc.); the *t* of feminine *-a(t)* appears then as in construct state: *sənt əžžāye* (= *ssəne žžāye*) ‘next year’, *baṭṭīx əlʾaḥmar* ‘the watermelon’, *yōm əlflāni* ‘such-and-such a day’. It is frequent in toponyms.

Indefiniteness: *wāḥed*: *kān fi wāḥed* ‘there was a guy’; *wāḥed* + singular appositive indefinite noun ‘a, one’; *kam wāḥed* ‘a few’; *ši*: *ši ʿəšrīn* ‘about twenty’, *ši šəġl mrattab* ‘some kind of nice job’; *ši . . . ši* ‘some . . . others’; *ḥayalla wāḥed*: ‘whoever/whatever he/it may be’.

Some adjectives can be variable: *ʾahwe ʿarabi* ‘Turkish coffee’, *sayyāra xuṣūṣi* 'a private car’.

Along with prevailing construct state, two ‘analytic’ constructions are used: (a) noun$_1$/preposition + cataphoric personal pronoun + preposition *la* + noun$_2$ (instead of construct: noun$_1$/preposition + noun$_2$): *sayyārto la ʾaxū* ‘his brother’s car’, *ʾana ḍədda la hal fəkra* ‘I am against this idea’ and (b) the genitive marker *tabaʿ* (or less frequent *šīt*) is used when focusing on the first term (*žžnān tabaʿek* ‘that madness you’re suffering from’); when pronominal pronouns can not be suffixed to the noun (*llandrōver tabaʿi* ‘my Land Rover’); or when two or more construct states would lead to nonsense: *qism əttaržame tabaʿ əžžāmʿa* ‘the translation department at the university’. N.B. Invariable *tabaʿ* also means ‘used for, peculiar to, belonging to’ and ‘presenting such-and-such a feature’: *fanažīn tabaʿ ʾahwe* ‘coffee cups’; *tabaʿ šu* ‘to which (of these) does this belong?’; *tabaʿ əlkahraba* ‘the guy who is looking after electricity’; *lbēt tabaʿ əlbāb əlʾazraʾ* ‘the house with a blue door’.

2.3.2 Quantifiers

kəll + singular indefinite noun ‘every’, + singular definite noun ‘the whole of’, + plural definite noun ‘all the’; *kam* + singular indefinite noun ‘some, a little amount of’; *ktīr* + plural indefinite noun ‘many’; *ši ktīr* ‘in large quantity’; *šwayyet* + singular noun ‘a few, a little’ (also *nətfet*); *hēk* ‘such a’: *hēk ʿālam* ‘people like that’; *kaza* ‘several, such-and-such number of’; *halʾadd* ‘this much’.

2.3.3 Relative clauses

These are introduced by *yəlli* when referring to a definite noun, asyndetic otherwise; resumptive pronoun is always present. The pronominal subject of a verbless relative clause is normally dropped: *lbēt əlli sāken fī* ‘the house I’m living in’.

2.3.4 Verb

When the second pronominal object is separated from the verb, it is introduced by *yā-*: *faržīni yā*: ‘let me see it!’.

The nominal object of a direct or indirect transitive verb often appears first as a cataphoric personal pronoun preceded by the preposition *la*: *šāfo la ʾaxū* ‘he saw his brother’, *ʾallo la ʾaxū* ‘he told his brother’ (*ʾəl-* with pronominal object: *bətšūfīhon ʾəlon* ‘you see them’).

Introduction of direct object by *b(i)-* (or *la*) is rare, and generally associated with aspectual distinctions; the preposition emphasizes the relationship to the object: *ʿam bəʾra bəlᵊktāb*/*ləktāb* ‘I’m reading just now/these days the book’.

Reflexive dative: *stannālo šwayye* ‘he waited for a while’, *ḥkīlak kəlme* ‘say a word!’.

Being enclitics, object personal pronouns, alone or suffixed to *la*, precede the second complement (introduced by *yā-* if pronominal).

2.3.5 Verbal system

2.3.5.1 The verbal system is basically aspectual. For the majority of verbs, the ‘perfect’ is a preterite, the active participle a resultative perfect, the *b*-imperfect a general ‘non-perfect’, the *ʿam*-imperfect a relative (concomitant) non-perfect; the Ø-imperfect is a modal form. Time reference is given by context or words (adverbs). *kān* is often used for reference to actual or fictional past.

The main uses of the *b-* form are: general present; narrative present; ‘planned’ or modal future; potential.

The *ʿam*-imperfect (*ʿammāl*, variable or not, *ʿamma*, more frequently *ʿam*, preceding the *b-* or the Ø-imperfect) is a ‘progressive’ form; the process takes place within an interval of time: *ma ʿam yənzel maṭar* ‘it’s not raining these days/this year...’; *ma ʿam baʿref ʾēmta* ‘I just can’t manage to know when’; *šu ʿam tsāwi* ‘what are you doing?’; *ṣarlo xams ᵊsnīn ʿam yədros ṭəbb* ‘he has been studying medicine for five years’.

ḥa- is used for a future given for certain: *ḥayrūḥ* ‘he will go (for sure)’; *ma ḥaʾəllak* ‘I won’t tell you’.

raḥ(a) (laḥ(a)) is used for intentional or expected future: *raḥ žīb mayy səxne* ‘I’m going to bring hot water’.

bədd- and *b-* are also used for various kinds of future, depending on the person of the verb (often more intentional with the 1st person): *bəktəblak* ‘I will write to you’, *bəddo yūʾaʿ* ‘it is going to fall down any time’.

2.3.5.2 Various → auxiliary verbs, → pseudo-verbs or impersonals express wish, ability, obligation, inchoation, durativity, etc.: *bədd-* + personal pronoun ‘to want’; *ʾəder*, *ḥəsen* ‘he was able’; *lāzem* ‘it is necessary to’; *ṣār* ‘he started to, he got used to doing’; *ballaš* ‘he began to’; *fəḍel*, *bəʾi* ‘he ended up . . .’; *ḍall*, *tamm* ‘he kept doing . . .’; *rəžeʿ* ‘he did/started doing again’. All are followed by an Ø-imperfect, some also by an active participle.

xalla + (pro)noun + Ø-imperfect is factitive ‘he let . . . do’. Its frozen imperative *xallī-* is used as suppletive 2nd and 3rd person imperative: *xallīhon ifūtu* ‘let them in!’.

Among the various uses of the verb *kān* ‘to be’ are: *kān* + Ø-imperfect ‘he used to . . .’ or ‘he was . . . ing’ (imperfect tense); with negation: *ma kənt ᵊtməll mənna* ‘haven’t you got fed up with it?’. *kān* is used as counterfactual: *ižība ʿa lbēt kān* ‘he should have brought it home’.

bəʾi ~ *baʾa*, imperfect *byəbʾa*, is used as verb of existence, and as continuative or inchoative auxiliary; it also appears in the apodosis of conditionals or along with imperatives, meaning ‘well then’.

Conjugated or fossilized *ʾām* is very frequent in narration: *ʾām rann əttelefōn*, ‘at that moment the telephone rang’, especially before the verb *ʾāl* ‘to say’, when reporting a conversation.

2.3.6 Participles

An active participle constitutes by itself a complete sentence. It often has resultative meaning: *ʾāxed bənton* ‘he is married to their daughter’; *wāžəʿni rāsi* ‘I’ve got a headache’; *wēn mxabbīhon* ‘where have you been hiding them?’. With verbs of movement it can refer to past, present, or future: *ʾana ṭāleʿ* ‘I’m going out’; *msāfer* ‘he’s away’; *msāfer bukra* ‘I’m traveling tomorrow’.

Passive participles often have potential meaning: *mahḍūm* ‘very pleasant (person)’; they can have active meaning: *mahyūb* ‘frightening’.

Imperative used as counterfactive: *kənti sʾali ḥālek has suʾāl* ‘you [fem.] should have asked yourself’.

2.3.7 Word order

In verbless sentences, the normal order subject-predicate can be inverted for topicalization. More generally, sentences of all types can begin with a topicalized element: *ləkk ḥžāb ma byaʿref yəktob* ‘Imagine, he’s not even able to write a charm for an amulet!'. A (quasi-)verbal predicate usually precedes its subject, especially indefinite, or when verbs of movement are concerned; the object is between both. Highly definite subjects precede the verb, as do sometimes even indefinite subjects (*walad ᵊb baṭᵊn ʾəmmo byəḥki* 'a baby in his mother's womb should talk?!'). In interrogative sentences and in relative clauses, the verb comes first. Auxiliary *kān* is very often put at the end or after its verb: *byəʾder kān ma yəži* '[had he wanted] he could have not come'; *iḥəṭṭu lʾafaṣ kānu* 'they would put the cage'.

2.3.8 Agreement

Adjectives agree in gender, number, and definiteness with nouns (even referring to non-humans), except for nouns in dual which have adjectives in plural; only adjectives referring to abstract (or concrete nouns considered as 'abstract') have feminine singular agreement. Many adjectives and most participles referring to women have *-īn* (not *-āt*) external plurals; when they refer to things, they generally have *-īn* and *-e*. Verbs agree in genus, number, and person with their subjects, except for abstract nouns (same agreement as adjectives); nouns in dual have a verb in plural; certain nouns referring to human groups (*lʿarab* 'the Bedouin', *nnās* 'the people') also have a 3rd person singular feminine agreement; both patterns can be found side by side: *ṣṣəfražiyye wāʾfe lābsīn ʾabyaḍ* 'the waiters were standing, dressed in white'.

Pseudo-impersonal constructions occur with verbs in masculine singular: *lafat naẓari šaġle* 'something caught my attention'.

Reference to 'neutral' 'that' is in feminine singular: *mu ḥəlwe* 'that's not fine!'; *ṣāret* 'here we go again!'

2.3.9 Existential sentences

Existential sentences are introduced by *(ma) fi* 'there is (not)' (preceded by *kān*: 'there was [not]'). They can have a definite subject in certain conditions (proper names, generic nouns).

'To have' is expressed by existential sentences '(to be) by/with/to' with prepositions + personal pronoun:

ʿand-: *ʿandi bəntēn* 'I have two daughters'; *ʿandi bēt bəššām* 'I have a house in Damascus'; *(fi) ʿandak ʾimkāniyye tānye* 'you've got another possibility'.

maʿ-: *ma maʿi maṣāri* 'I have no money on me' or 'I have no money at all'; *maʿo šhāde* 'he holds a diploma'.

ʾəl-: *ʾəli ʾaxx bi ʾamērka* 'I have a brother in America'; *ʾəlo ʾīdēn* 'it has two handles'; *ʾəlo ṭaʿme ġarībe* 'it has a strange taste'.

2.3.10 Conditional sentences

The conditional particles are *ʾiza* and *law*; *ʾən* is used only in proverbs or idioms. The system of conditional sentences is complex; many possibilities are provided, and the same construction can have different meanings, distinguishable only by the context. *ʾiza* refers mainly to 'realis', and *law* to 'irrealis', but they share many a meaning. Perfect never occurs in the apodosis of *ʾiza* sentences. *b*-imperfect is frequent both in protasis and apodosis, due to its ability to represent situations, even unrealized or unrealizable. *kān* is used to emphasize the hypothetical situation; it combines with *ʾiza* in the new particle *ʾizakān*.

2.3.11 Circumstantial clauses

Circumstantial clauses have the pattern *w* + Subject + Predicate. They are placed before the main clause; the order can be Subject + *w* + Predicate if the subject is an independent personal pronoun: *w huwwe ṭāleʿ* ~ *huwwe w ṭāleʿ* 'while he was leaving'.

3. LEXICON

Like the other dialects of the area, the Damascus Arabic lexicon is full of borrowings, both ancient (from Aramaic, Turkish, and Persian), medieval (from French, Italian, *lingua franca*), and modern (from Ottoman Turkish, French, Italian, English, other Arabic dialects: Bedouin, Egyptian, Lebanese); sometimes other words are derived from them. They are numerous in certain parts of the lexicon (e.g. French words in mechanics and cosmetics).

Many words, especially verbal forms, including borrowings from Standard Arabic, are fossilized and grammaticalized; examples, besides the already mentioned *ʾām*, are: *ʾāl*, used before reporting someone's words or a proverb, *niyālak* 'good for you!', *yaʿni* 'that is, so-so', *ʾabṣar* 'I wonder', *ʾaṣbaḥ* 'in the end', *baʾa* 'so, then', etc.

BIBLIOGRAPHICAL REFERENCES

Ambros, Arne. 1977. *Damascus Arabic*. Malibu: Undena Publications.

ʾAswad, Nizār al-. 1998. *al-Ḥikāyāt aš-šaʿbiyya aš-šāmiyya*, IV. Damascus: published by the author.

——. 2002. *al-Ḥikāyāt aš-šaʿbiyya aš-šāmiyya*, V. Damascus: published by the author.

Bergsträsser, Gotthelf, 1924. *Zum arabischen Dialekt von Damaskus. I. Phonetik – Prosatexte*. Hannover: Orient-Buchhandlung Heinz Lafaire.

Bloch, Ariel. 1965. *Die Hypotaxe im Damaszenisch-Arabischen, mit Vergleichen zur Hypotaxe im Klassisch-Arabischen*. Wiesbaden: F. Steiner.

——, and Heinz Grotzfeld 1964. *Damaszenisch-arabische Texte, mit Übersetzung, Anmerkungen und Glossar*. Wiesbaden: F. Steiner.

Bohas, Georges and Katia Zakhariya (eds.). 2000, 2001, 2002, 2003, 2004. *Sīrat al-Malik az-ẓāhir Baybarṣ ḥasab ar-riwāya aš-šāmiyya*. 5 vols. Damascus: Institut Français d'Etudes Arabes.

Cantineau, Jean and Youssef Helbaoui. 1953. *Manuel élémentaire d'arabe oriental (Parler de Damas)*. Paris: Klincksieck.

Cowell, Mark W. 1964. *A reference grammar of Syrian Arabic (based on the dialect of Damascus)*. Washington, D.C.: Georgetown University Press.

Ferguson, Charles with Moukhtar Ani a.o. 1961. *Damascus Arabic*. Washington, D.C. (Repr. by the Center for Applied Linguistics of the Modern Language Association of America.)

Grotzfeld, Heinz. 1965. *Syrisch-arabische Grammatik (Dialekt von Damaskus)*. Wiesbaden: O. Harrassowitz.

Kassab, Jean. 1970. *Manuel du parler arabe moderne au Moyen-Orient*. Paris: Geuthner.

Kayyāl, Munīr. 1987. *Ḥikāyāt dimašqiyya*. Damascus: published by the author.

Kuhnt, Eberhard. 1958. *Syrisch-arabischer Sprachführer*. Wiesbaden: O. Harrassowitz.

Lentin, Jérôme. 1982 [1984]. "Un ancien système à formes 'lento' et 'allegro' dans le dialecte de Damas?". *Bulletin d'Etudes Orientales* 34.111–139.

Salamé, Claude and Jérôme Lentin (forthcoming). *Dictionnaire d'arabe dialectal syrien (parler de Damas)* [Damascus Arabic–French]. Damascus: Institut Français du Proche Orient.

Stowasser, Karl and Moukhtar Ani. 1964. *A dictionary of Syrian Arabic. English–Arabic*. Washington, D.C.: Georgetown University Press.

Turžmān, Sihām. 1978. *Yā māl əš-Šām*. 2nd enlarged ed. Damascus.

JÉRÔME LENTIN (Institut National des Langues et Civilisations Orientales, Paris)

Ḍamīr

In Medieval Arabic grammatical literature, the concept of *ḍamīr* 'pronoun' (pl. *ḍamāʾir*) constitutes a subcategory of *kināya* 'proform'. The term *kināya* refers to a nominal substitute realized either as a vague noun such as *fulān* 'so-and-so', *kayta wa-kayta* 'such and such' and *kaḏā wa-kaḏā* 'so-and-so many', or otherwise as a pronoun. *Kināya* is thus a device used for the sake of vagueness or economy. The *ḍamīr*, as a special case of → *kināya*, is used as a short version of the noun and often serves for disambiguation. Ibn Yaʿīš (*Šarḥ* III, 84) explains that, whereas in *zaydun faʿala zaydun* 'Zayd, Zayd did' the use of the name *zaydun* twice might be interpreted as referring to two different persons, in *zaydun faʿala*, where *zaydun* occurs first as a full nominal and then as a pronoun implicit in the verb, there is unambiguously only one person involved. The personal pronoun is characterized by the grammarians as a definite (many regard it as the most definite) nominal. Some however indicate that a personal pronoun may refer also to an indefinite noun, in which case it should be regarded as indefinite, as in *laqītu rajulan fa-ḍarabtu-hu* 'I met a man and I hit him' (see for discussion Ibn Yaʿīš, *Šarḥ* V, 87; Ibn ʿUṣfūr, *Šarḥ* I, 404).

The pronominal system is normally presented as composed of three persons: *mutakallim* (1st person), *muxāṭab* (2nd person), and *ġāʾib* (3rd person). The pronouns change form according to case. A genitival (*jarr*) pronoun can only be realized as a suffix attached (*muttaṣil*) to a noun, as in *ʾaxū-ka* 'your brother'. In the nominative and accusative, by contrast, it can be realized either as a separate independent word (*munfaṣil*), or as *muttaṣil*. In the latter case, the pronoun may either have a morphological realization, as is the case with the suffix *-ka* in *ḍaraba-ka* 'he hit you', or otherwise be implicit in the verb, as it is in *qāma* in the sentence *zaydun qāma* 'Zayd stood up'. Thus, for instance, the suffix *-tu* in *ḍarabtu* 'I hit' and the word *ʾanā* are regarded as two versions (*muttaṣil* and *munfaṣil* respectively) of the nominatival pronoun in the first person singular.

The separate pronoun also serves as an emphasizer (*taʾkīd*) for a preceding accusatival or genitival attached pronoun, as in *qumta ʾanta* '**you** stood up', *ḍarabtu-ka ʾanta* 'I hit **you**' and *marartu bi-ka ʾanta* 'I passed by **you**' (see, e.g., Sībawayhi, *Kitāb* I, 344–346). Furthermore, an emphasizing separate pronoun is obligatory in cases where a nominatival noun is to be coordinated with a preceding attached (or implicit) pronoun. Thus, **qumtu wa-zaydun* 'I and Zayd stood up' and **qāma wa-zaydun* 'he and Zayd stood up' are disallowed: the verb in each of these cases must be followed by an emphasizing separate pronoun: *qumtu ʾana wa-zaydun*, *qāma huwa wa-zaydun*. However, an attached accusatival pronoun may be coordinated with an accusatival noun, as in *ḍarabtu-ka wa-zaydan* 'I hit you and Zayd' (for details, see Sībawayhi, *Kitāb* I, 342–344; Ibn as-Sarrāj, *ʾUṣūl* II, 119, and see further below; for the principle of balancing, see Bloch 1986:1–14).

Regarding the 3rd person pronoun, the grammarians differ as to whether the verbal endings *-ā* in [*az-zaydāni*] *qāmā* 'the two Zayds stood up' and *-ū* in [*az-zaydūna*] *qāmū* 'the Zayds stood up' serve as pronouns or just as dual and plural markers (*ʿalāmāt*) respectively. Ibn Yaʿīš (*Šarḥ* III, 87–88) states that these two endings serve both as *ḍamīr al-fāʿil* (*fāʿil* denotes the subject in a VS sentence, sometimes translated as 'agent', a semantic term that does not always correspond with the medieval syntactic usage of *fāʿil*) and number marker. This indeed represents the common view among the grammarians in this regard. He indicates, however, that for other grammarians the endings in question implement only the function of number markers, whereas the *fāʿil* pronoun is implicit in the verb. By and large, the grammarians agree that in cases such as *zaydun qāma* 'Zayd stood up' the verb displays no morphological marker of number, and that the *fāʿil* pronoun is implicit (*fī n-niyya*). A significant distinction is drawn between cases such as *az-zaydāni qāmā*, where the ending *-ā* is interpreted as a *ḍamīr* for *az-zaydāni*, and cases such as *qāmā az-zaydāni*, where the ending is analyzed as a dual marker. The same applies to the plural (see, e.g., Sībawayhi, *Kitāb* I, 201–203). However, the construction in which the *fāʿil* is preceded by a verb in the dual/plural form is considered uncanonical by the grammarians; this kind of usage is termed *luġa ʾakalūnī l-barāġīṯ* 'the fleas ate me up' (for an extensive discussion, see Levin [1989] and his references). A verb preceding its *fāʿil* must, in the standard language, occur in the singular form (masculine or feminine), irrespective of the number inflection of the *fāʿil*. In such sentences as *qāma zaydun/az-zaydāni/az-zaydūna*, the verbal form is said to be devoid of a personal pronoun: the function of the *fāʿil* is only implemented by the nominal following the verb (see, e.g., Sībawayhi, *Kitāb* I, 201). The feminine form *qāmat* displays, in the grammarians' theory, the feminine marker *-t*, which cannot, however, be construed as a pronoun (see, e.g., Sībawayhi, *Kitāb* I, 201–202). The argument is that *qāmat* may be followed by a *fāʿil*, whether in a verbal sentence such as *qāmat hindun* 'Hind stood up', or in a clausal predicate such as [*hindun*] *qāmat*. In the first of the two cases, interpreting the *-t* as a *fāʿil* pronoun would mean that the verb assigns the *rafʿ* case to two different constituents, and that the function of *fāʿil* is performed first by a pronoun and then by a full noun, which is theoretically unacceptable. Moreover, assigning the ending *-t* the function of a pronominal *fāʿil* in cases such as *qāmat hindun* entails an unacceptable preposing of the pronoun to its antecedent (see, e.g., Ibn Yaʿīš, *Šarḥ* III, 88–89; and → cataphora). In the second case, there is a *fāʿil* pronoun implicit in *qāmat*, but here too the ending *-t* is not analyzable as a *fāʿil* pronoun since *qāmat* may be followed by a full-noun *fāʿil*: *hindun qāmat jāriyatu-hā* 'Hind, her maid stood up'.

Pronouns in the *naṣb* and *jarr* cases normally occur as suffixes (*muttaṣil*) and, except for the 1st person singular, they are identical. Thus, the 2nd person masculine singular suffix *-ka* may be attached either to *ḍaraba* as an accusatival pronoun (*ḍaraba-ka* 'he hit you'), or as a genitival pronoun to *bayt* (*baytu-ka* 'your house'). The suffix for the 1st person singular is *-ī* for the genitive (*bayt-ī* 'my house') and *-nī* for the accusative (*ḍaraba-nī* 'he hit me'). The grammarians explain that the actual pronoun in both cases is the same, *-ī*; the *n* in the accusatival suffix is designed to avoid the ending of a verbal form with the vowel *ī*. This *n* is termed *nūn al-wiqāya*, as its function is to retain the original vocalization of the third radical (e.g. the *u* in *yaḍribu-nī*) by preventing an unacceptable vocalic termination of the verbal form (for further details see, e.g., Ibn Yaʿīš, *Šarḥ* III, 89–91, 122–125; cf. Sībawayhi, *Kitāb* I, 338–340).

A pronominal object coreferring with the *fāʿil* is disallowed, unless the verb is a bi-transitive (see below) cognitive verb. Thus, **ḍarabtu-nī* must be replaced by *ḍarabtu nafs-ī* 'I hit myself', but *ḥasibtu-nī munṭaliqan* 'I regarded myself as leaving' is acceptable (for discussion see, e.g., Sībawayhi, *Kitāb* I, 337–338).

An accusatival pronoun may be realized as an independent separate pronoun in cases where the syntactic structure bars the possibility of attaching it to a verb or any other verb-like operator (*ʿāmil*). The grammarians outline three groups of cases:

i. When a pronominal object precedes its verb, as in *ʾiyyā-ka naʿbudu wa-ʾiyyā-ka nastaʿīnu* 'You alone we worship and to you alone we pray for help' (Q. 1/5).

ii. When the pronominal object functions as the second object of a bi-transitive verb, as the second complement of a verbal noun (*maṣdar*) or, otherwise, as the predicate in an auxiliary-verb construction (*xabar kāna* etc.). In the first two cases, the use of *ʾiyyā-* is the norm when each of the objects is a 3rd person pronoun, as in: *ʿallamtu-hu ʾiyyā-hu* 'I taught him it'. When the pronouns involved are of two different persons, the grammarians do not rule out the use of an attached pronoun: *ʾaʿṭaytu-ka-hu* 'I gave you it' and *ḍarb-ī-ka* 'my hitting you' are acceptable constructions, beside the more common *ʾaʿṭaytu-ka ʾiyyā-hu* and *ḍarb-ī ʾiyyā-ka*. It is normally stipulated that the 2nd person pronoun should precede the 3rd, and that the 1st person takes precedence over the other two (Ibn Yaʿīš, *Šarḥ* III, 104–105). This restriction does not apply, however, when *ʾiyyā* is used (for further discussion of pronominal objects in bi-transitive constructions, see Sībawayhi, *Kitāb* I, 335–338). Similarly, a pronominal *xabar kāna* may be realized as either an attached pronoun (*kāna-nī* 'it was me'), or independently with *ʾiyyā*, as in *kāna zaydun ʾiyyā-hu* ('Zayd was him'), *kāna ʾiyyā-ya* ('he was me'). Ibn Yaʿīš (*Šarḥ* III, 107) indicates that the *ʾiyyā* version is the preferred one, primarily because *xabar kāna* is analogous to the → *xabar* in a *mubtadaʾ-xabar* construction, which always occurs as an independent constituent, unattached to the *mubtadaʾ* (for further details, see Ibn as-Sarrāj, *ʾUṣūl* II, 120–121).

iii. When the conjunctive *wa-* or the exceptive particle *ʾillā* bar the use of an attached pronoun: *ḍarabtu zaydan wa-ʾiyyā-hu* 'I hit Zayd and him', *mā ḍarabtu ʾillā ʾiyyā-ka* 'I didn't hit anyone except you'.

In all the above cases, the independent version of the accusatival pronoun is realized by a suffix attached to the bound morpheme *ʾiyyā*. The grammarians differed as to the status of *ʾiyyā*. Some regarded it as a prop word (*ʿimād*) for the pronoun attached to it. The majority of grammarians, however, regarded the suffix attached to *ʾiyyā* as a kind of particle (*ḥarf*) rather than as a pronoun, while *ʾiyyā* itself was described as a pronominal element (*ism muḍmar*) (For further details see, e.g., Ibn Yaʿīš, *Šarḥ* III, 98–104). One case where the accusatival pronoun may be realized as an independent pronoun without *ʾiyyā* is that of exceptive sentences with intransitive verbs: *mā qāma ʾillā ʾanta* 'no one stood up but you' (the pronoun in this case may be interpreted as occupying either a nominatival or an accusatival position). For further discussion of *ʾiyyā* see Sībawayhi, *Kitāb* I, 332–337.

A genitival pronoun, by contrast, must obligatorily take a suffix form (*muttaṣil*), since a constituent in the *jarr* case may not precede or be separated from its operator, whether a preposition or the first member of a genitival construct (*muḍāf*). However, a pronominal suffix may not be attached to the particles *ka-*, *ḥattā*, and *muḏ*: in the first case, **ka-hu* 'like him' is obligatorily substituted by *miṯla-hu*, whereas the last two particles would be followed by *ḏāka* as a substitute for the attached pronoun (*ḥattā ḏāka* 'until this', etc.). For further details see, e.g., Ibn Yaʿīš, *Šarḥ* III, 93; cf. Sībawayhi, *Kitāb* I, 335, 344). Another point usually made by the grammarians regarding the genitival pronoun is that in certain cases it retrieves the original vowel of a segment, as is the case in *la-hu*, etc., where the original vowel *a* of the preposition is retrieved by the suffix: where the preposition is followed by a noun it takes the vowel *i* (*li-zaydin* 'to Zayd') so as to avoid ambiguity with *lām al-ibtidāʾ* (see, e.g., Sībawayhi, *Kitāb* I, 341–342; Ibn as-Sarrāj, *ʾUṣūl* II, 124).

Two special cases of pronominal usage discussed by the grammarians concern the *ḍamīr* following *lawlā* and *ʿasā*. Ibn Yaʿīš (*Šarh* III, 118–119) indicates that since the noun following *lawlā* is assigned the *rafʿ* case by → *ibtidāʾ*, a pronoun in that position must correspondingly be realized as a separate pronoun (*munfaṣil*): *lawlā ʾanta*, *lawlā ʾantum* 'but for you [sg. and pl. masc. respectively]', etc. He notes, however, the use of *lawlā-ka* and *lawlā-ya* by Arabic speakers, and describes the controversy existing in medieval grammatical literature regarding the case of the pronoun following *lawlā*. In Sībawayhi's (*Kitāb* I, 340–341) view, a noun following *lawlā* is assigned the *rafʿ* case, whereas a pronoun in that same position takes the *jarr*. ʾAxfaš is quoted by Ibn Yaʿīš (*Šarḥ* III, 121–122) as advocating the *rafʿ* case for the pronominal suffixes in *lawlā-ka* and *lawlā-ya*, in correspondence with the case of the noun in this position. He argued that there is no one-to-one correspondence between case and pronominal form. In support of this claim he drew an analogy between *lawlā-ka* and *mā ʾanā ka-ʾanta* 'I am not like you'. His argument was that the *rafʿ* pronoun in *lawlā-ka* may take the suffix form, much as the *jarr* pronoun following *ka-* in *ka-ʾanta* takes the independent form of the pronoun normally assigned to the *rafʿ* case.

The controversy over *ʿasā* is related likewise to the fact that Arabic admits two forms of conjugation: *ʿasaytu*, *ʿasayta* 'I may, you may' versus *ʿasānī*, *ʿasāka*. Ibn Yaʿīš' discussion of this issue is interwoven in his discussion of the form of the pronoun following *lawlā* (*Šarḥ* III, 118–122).

The medieval grammarians also deal with non-anaphoric pronouns. Two cases are normally discussed. The first case is that of *ḍamīr aš-šaʾn*. In Medieval Arabic grammatical theory, *ḍamīr aš-šaʾn* is conceived of as a pronoun occupying a *mubtadaʾ* position (*mubtadaʾ* denotes the subject in a non-VS sentence; it is sometimes translated as 'topic', a pragmatic term that does not cover all instances of *mubtadaʾ*). As such, it may occur either as a separate pronoun in sentence-initial position (*huwa zaydun munṭaliqun* 'Zayd is leaving') or, otherwise, as a *ḍamīr muttaṣil*. In the latter case it may be attached to a particle of the *ʾinna* group (*ʾinna-hu man yaʾti-nā naʾti-hi* 'whoever comes to us we will come to him', *ʾinna-hu ʾamatu llāhi ḏāhibatun* 'the female slave of God is going'), or take the position of the first object of a cognitive verb of the *ẓanna* group (*ẓanantu-hu zaydun qāʾimun* 'I thought Zayd was standing', *ḥasibtu-hu qāma ʾaxū-ka* 'I thought your brother was standing'). Another option, still, is for *ḍamīr aš-šaʾn* to be implicit in an auxiliary verb of the *kāna* group (*kāna ʾanta xayrun min-hu* 'you were better than he', *laysa xalaqa llāhu miṯla-hu* 'God has not created anything like him'). The clause following *ḍamīr aš-šaʾn* in all these cases is presented as its *xabar*. The pronominal *mubtadaʾ* refers in a general way to the 'matter' (variously *šaʾn*, *ḥadīṯ*, *qiṣṣa*) conveyed by the following *xabar* clause, and accordingly occurs in the singular. Hence the respective terms *ḍamīr aš-šaʾn*, *ḍamīr aš-šaʾn wa-l-ḥadīṯ*, and *ḍamīr al-qiṣṣa*. In the grammarians' view, the clausal *xabar* functions as an exponent (*tafsīr*) to the 'vague' pronoun filling the *mubtadaʾ* position. Occasionally *ḍamīr aš-šaʾn* takes the feminine form, so as to agree with a feminine post-verbal subject occurring later in the sentence. This is normally exemplified by the Qurʾānic sentence *ʾinna-hā lā taʿmā l-ʾabṣāru* 'it is not the eyes that are blind' (Q. 22/46).

Semantically, the grammarians restrict *ḍamīr aš-šaʾn* to markedly emphatic cases of *tafxīm wa-taʿẓīm* 'emphasis and exaltation'; thus, *huwa zaydun munṭaliqun*, *ẓanantu-hu zaydun qāʾimun* and *ʾinna-hu ʾamatu llāhi ḏāhibatun* are conceived of as the respective emphatic versions of *zaydun munṭaliqun*, *ẓanantu zaydan qāʾiman*, and *ʾamatu llāhi ḏāhibatun*.

As analogous to *ḍamīr aš-šaʾn* the grammarians present the pronoun *-hu* in *rubba-hu rajulan* 'many a man'. Zamaxšarī (Ibn Yaʿiš, *Šarḥ* III, 118) describes this pronoun as non-referential (*yurmā bi-hi min ġayri qaṣdin ʾilā muḍmarin la-hu*), indefinite (*nakira*), and vague (*mubham*). Much like the similarly vague pronoun *ḍamīr aš-šaʾn*, it requires *tafsīr*. Ibn Yaʿīš points out, however, that the two cases differ in the kind of *tafsīr* involved: while *ḍamīr aš-šaʾn* is expounded by a clause (*jumla*), the pronoun attached to *rubba* is expounded by a single word (*mufrad*).

Another type of non-anaphoric *ḍamīr* dealt with by the grammarians is *ḍamīr al-faṣl*. This pronoun is presented as occurring between two definite predicative constituents, in sentences introduced by *ẓanna*, *kāna*, *ʾinna* and their 'sisters', as well as in 'pure' *mubtadaʾ* + *xabar* sentences. In particular, it is stipulated that the *xabar* must be a definite or quasi-definite phrase. In effect, any type of nominal phrase would qualify as a post-*ḍamīr-al-faṣl xabar*, apart from a single-word indefinite nominal (whether a substantive or an adjective) unaccompanied by any modifying complement.

Ḍamīr al-faṣl is normally presented by the grammarians as a disambiguating device: by separating the (definite) *xabar* from its (definite) *mubtadaʾ* it marks the relationship between the two constituents as predicative rather than attributive. The status of *ḍamīr al-faṣl* is often described as that of a particle (*ḥarf*) whose → *ʿamal* is annulled (*ʾilġāʾ*), and is thus devoid of case (*lā yakūnu la-hu mawḍiʿ min al-ʾiʿrāb*, Ibn Yaʿīš, *Šarḥ* III, 113).

Significantly, however, the grammarians view the occurrence of this pronoun as optional rather than as obligatory. They often point out that *ḍamīr al-faṣl* also serves for emphasis, and that as such it corefers with the *mubtadaʾ* (e.g.

Ibn Yaʿīš, *Šarḥ* III, 110). On the other hand, there are cases where the pronoun must be interpreted as *ḍamīr at-taʾkīd* ‘pronoun of emphasis’ rather than as *ḍamīr al-faṣl*. Indeed, Ibn Yaʿīš (*Šarḥ* III, 113–114) is well aware of the structural ambiguity that may arise in cases where a separate pronoun occurs between two nominal elements. He therefore outlines the main differences between *ḍamīr al-faṣl*, a pronominal apposition (*badal*), and a pronominal emphasizer (*taʾkīd*):

i. A pronominal emphasizer may follow only a pronoun, whereas *ḍamīr al-faṣl* may follow either a noun or a pronoun.
ii. A pronominal emphasizer is by definition a nominal, and as such it must agree in case with the emphasized element. *Ḍamīr al-faṣl*, by contrast, is caseless.
iii. A pronominal apposition agrees in case with its head much like a pronominal emphasizer. However, if the apposition modifies an accusatival nominal, it must be attached to *ʾiyyā* (*ḏ̣anantu-ka ʾiyyāka xayran min zaydin* ‘I considered you to be better than Zayd’). By contrast, *ḍamīr al-faṣl* and the pronominal emphasizer invariably take the form of a separate pronoun (without *ʾiyyā*).
iv. *Lām at-taʾkīd* may be attached to *ḍamīr al-faṣl* (*ʾin kāna zaydun la-huwa l-ʿāqila* ‘indeed, Zayd was the intelligent one’), but not to a pronominal apposition or an emphasizer, because separation between head and modifier of a noun phrase is unacceptable.

For illuminating discussions of *ḍamīr al-faṣl* by earlier grammarians, see Sībawayhi, *Kitāb* I, 346–350; Ibn as-Sarrāj, *ʾUṣūl* II, 125–126.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES
Ibn as-Sarrāj, *ʾUṣūl* = ʿAbū Bakr Muḥammad ibn Sahl Ibn as-Sarrāj, *al-ʾUṣūl fī n-naḥw*. Ed. ʿAbd al-Ḥusayn al-Fatlī. Beirut, 1987.
Ibn ʿUṣfūr, *Šarḥ* = ʿAlī ibn Muʾmin ibn Muḥammad ibn ʿAlī Ibn ʿUṣfūr, *Šarḥ Jumal az-Zajjājī*. Ed. Ṣāḥib ʾAbū Janāḥ. Mosul, 1980–1982.
Ibn Yaʿīš, *Šarḥ* = Muwaffaq ad-Dīn Yaʿīš ibn ʿAlī Ibn Yaʿīš, *Šarḥ al-Mufaṣṣal*. Cairo: Maktabat al-Mutanabbī, n.d.
Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. Ed. Hartwig Derenbourg. Hildesheim and New York: G. Olms, 1970.

SECONDARY SOURCES
Bloch, Ariel A. 1986. *Studies in Arabic syntax and semantics*. Wiesbaden: O. Harrassowitz.
Levin, Aryeh. 1989. “What is meant by *ʾakalūnī l-barāġīṯu*?”. *Jerusalem Studies in Arabic and Islam* 12.40–65.

YISHAI PELED (Tel Aviv University)

Ḏ̣arf → Mafʿūl fīhi

Declension

1. DEFINITION

In general, *declension* affects the nouns, adjectives, articles, numerals, and pronouns of a language (as opposed to *conjugation*, which affects finite verb forms). In Classical Arabic, the nouns, adjectives, numerals (except those from 11 to 19), as well as the dual forms of the demonstrative pronoun (near deixis) and the relative pronoun are subject to declension. According to native Arabic grammatical theory, nouns are either *muʿrab* ‘declinable’ or (in rare cases) *mabnī* ‘indeclinable’. Declinable nouns are then *munṣarif* ‘declined with nunation (*tanwīn*)’ or *ġayr munṣarif* ‘declined without nunation’. The characteristic vowels *-u*, *-i*, and *-a* of the three cases nominative, genitive, and accusative are associated with the Arabic terms *rafʿ*, *jarr* (*xafḍ* in the terminology of the Kufan grammarians), and *naṣb* respectively. Depending on whether a given noun appears in two forms (*u*-ending for nominative and *a*-ending for both genitive and accusative) or in three forms (*u*-ending for nominative, *i*-ending for genitive, and *a*-ending for accusative), it is commonly called ‘diptote’ (‘having two cases’) or ‘triptote’ (‘having three cases’) in the Western tradition (the terms deriving from Greek *ptōsis* ‘case’, → diptosis). The scenario ‘diptote’ is sometimes captured by the opposition *casus rectus* for the nominative and *casus obliquus* for both the genitive and the accusative. There are a few indeclinable Arabic nouns of the type CaCāCi which only appear in one surface form; however, these forms are not described as ‘monoptote’.

The term → *ʾiʿrāb* (lit. ‘making Arabic/making clear, manifest’) should – at least from a historical perspective – not automatically be equated with *declension* (and, for that matter, *conjugation*). Fleisch (1998:1249) points out that for the traditional Arab grammarians *ʾiʿrāb* was a purely functional term for the analysis of the syntactic functions that the three vowel markers *ḍamma*, *kasra*, and *fatḥa* indicated in both nouns and verbs.

2. DECLENSION AND CASE IN CLASSICAL ARABIC

Within Semitic, Classical Arabic shares a tripartite desinential declension of the noun together with Akkadian (for a comprehensive overview, see Brockelmann 1908:459–466; Wright 1967:I, 234–252; Huehnergard 1997:606–609). Unlike in Akkadian, the independent personal pronouns and the pronominal suffixes in Arabic do not undergo declension, with the exception of the pronominal suffix of the 1st pers. sg. com., which is *-ī* (allomorph *-ya* after long vowel or diphthong) for the possessive (genitive) on nouns and prepositions and *-nī* for the accusative on verbs, e.g. *kitāb-ī* 'my book' vs. *ḍaraba-nī* 'he hit me'. Also, the dual of the demonstrative pronoun (near deixis) has the oblique forms *hāḏayni* (masc.) and *hātayni* (fem.) which contrast with *hāḏāni* and *hātāni* in the nominative. For the relative pronoun, the corresponding oblique forms are *allaḏayni* (masc.) and *allatayni* (fem.) which contrast with *allaḏāni* and *allatānī* in the nominative. There are also traces of conflicting forms such as *allaḏūna* vs. *allaḏīna* for the relative pronoun masc. pl. While the former form represents the attempt of early Arab philologists to 'reconstruct' the case marking across the board, it was the latter form, pointing to the ascendance of the 'oblique' case at early stages of Arabic, that nevertheless became the accepted form (cf. also Vollers 1906:170).

The singular paradigm features separate nominative (*rafʿ*), genitive (*jarr* or *xafḍ*), and accusative (*naṣb*) forms for Akkadian and Arabic nouns (both masculine and feminine). However, the genitive and accusative forms overlap in both the dual and plural paradigm, as long as the plural is 'sound' (external). (This circumstance is not referred to as 'diptosis'.) Table 1 presents an overview of the Akkadian (Old Babylonian) nouns *malāḫum* 'seafarer [masc.]' and *malāḫtum* 'seafarer [fem.]' in the absolute form with → mimation, as well as the Arabic counterparts *mallāḥun* and *mallāḥatun* in the indefinite form with nunation (while the Akkadian term goes back to a Sumerian compound formation [(LÚ.)MÁ.LAḪ$_{4/5}$], the Arabic term is synchronically associated with the root *m-l-ḥ* as in *milḥ* 'salt').

Arabic nouns with external plural exhibit the same case endings when indefinite or definite, irrespective of number, except for the loss of nunation in the singular and the loss of the *na-* or *ni*-ending in the dual and plural. Thus, as shown in Table 2, taking the paradigm of 'the seafarer' and 'the seafarer of the country', one arrives at the following definite forms, with the definite article or in an annexation (*ʾiḍāfa*), in which the second term could also be replaced by a pronominal suffix.

Nouns with a → 'broken' (internal) plural always appear with *u*-ending for nominative, *i*-ending for genitive, and *a*-ending for accusative in their definite form, i.e. they then are triptote. The indefinite forms of broken plurals may be triptote or diptote (the dual plays no role in this context as it is always externally marked). While

Table 1. Singular, dual, and sound plural in Akkadian and Arabic

		masc.		fem.	
		Akkadian	Arabic	Akkadian	Arabic
sing.	nom.	*malāḫum*	*mallāḥun*	*malāḫtum*	*mallāḥatun*
	gen.	*malāḫim*	*mallāḥin*	*malāḫtim*	*mallāḥatin*
	acc.	*malāḫam*	*mallāḥan*	*malāḫtam*	*mallāḥatin*
dual	nom.	*malāḫān*	*mallāḥāni*	*malāḫtān*	*mallāḥatāni*
	g./a.	*malāḫīn*	*mallāḥayni*	*malāḫtīn*	*mallāḥatayni*
pl.	nom.	*malāḫū*	*mallāḥūna*	*malāḫātum*	*mallāḥātun*
	g./a.	*malāḫī*	*mallaḥīna*	*malāḫātim*	*mallāḥātin*

Table 2. Definite forms in Arabic

		masc.		fem.	
		def. article	annexation	def. article	annexation
sg.	nom.	*al-mallāḥu*	*mallāḥu l-baladi*	*al-mallāḥatu*	*mallāḥatu l-baladi*
	gen.	*al-mallāḥi*	*mallāḥi l-baladi*	*al-mallāḥati*	*mallāḥati l-baladi*
	acc.	*al-mallāḥa*	*mallāḥa l-baladi*	*al-mallāḥata*	*mallāḥata l-baladi*
dual	nom.	*al-mallā ḥāni*	*mallāḥā l-baladi*	*al-mallāḥatāni*	*mallāḥatā l-baladi*
	g./a.	*al-mallāḥayni*	*mallāḥayi l-baladi*	*al-mallāḥatayni*	*mallāḥatayi l-baladi*
pl.	nom.	*al-mallāḥūna*	*mallāḥū l-baladi*	*al-mallāḥātu*	*mallāḥātu l-baladi*
	g./a.	*al-mallāḥīna*	*mallaḥī l-baladi*	*al-mallāḥāti*	*mallāḥāti l-baladi*

nouns in the singular are triptote, certain types of adjectives and certain types of proper nouns are diptote. Whether or not not a given broken plural is triptote or diptote depends on phonotactic criteria (as reflected in the nominal pattern); in rare cases (e.g. *šayʾun*/*ʾašyāʾu* 'things'), diptosis is a lexicalized feature. Diptote nouns and adjectives always become triptote in the definite form. Thus, as shown in Table 3, one arrives at the following forms:

Table 3. Declension of diptote nouns

sg.	nom.	*šayʾun*	*aš-šayʾu*	*šayʾu l-baladi*
	gen.	*šayʾin*	*aš-šayʾi*	*šayʾi l-baladi*
	acc.	*šayʾan*	*aš-šayʾa*	*šayʾa l-baladi*
pl.	nom.	*ʾašyāʾu*	*al-ʾašyāʾu*	*ʾašyāʾu l-baladi*
	gen.	*ʾašyāʾa*	*al-ʾašyāʾi*	*ʾašyāʾi l-baladi*
	acc.	*ʾašyāʾa*	*al-ʾašyāʾa*	*ʾašyāʾa l-baladi*

Generally speaking, the following items are diptote (for further details cf. Wright 1967:I, 239–247):

i. broken plurals of the type CaCāCiCu, CaCāCīCu, CuCaCāʾu, ʾaCCiCāʾu, CaCāCin, CaCCā, CaCāCā; a number of plural forms such as *ʾuwalu* (pl. of *ʾawwalu* (masc.) and *ʾūlā* (fem.) 'first'); *ʾašyāʾu* (pl. of *šayʾun* 'thing');
ii. adjectives of the type CaCCāʾu, Ca/iCCā, ʾaCCaCu (fem. CuCCā or CaCCāʾu), CaCCānu;
iii. proper names of Canaanite origin such as *ʾibrāhīmu*, proper names ending in *ʾalif maqṣūra* or *ʾalif mamdūda* such as *yaḥyā*, proper names ending in *-ānu* such as *ʿuṯmānu*, proper nouns that formally resemble a verb form such as *yazīdu*, proper names ending in *tāʾ marbūṭa* such as *makkatu* 'Mecca', most feminine proper names such as *miṣru* 'Egypt', and many masculine proper names of the type CuCaC such as *ʿumaru*.

As already stated, all items of type (i) and (ii) become triptote when definite. Even the proper nouns of type (iii) may be triptote when used in an indefinite way, e.g. *rubba ʾibrāhīmin laqītuhu* 'many an Abraham have I met' (cf. Wright 1967:I, 245).

Nouns with a weak third radical slightly complicate the situation, as they may have only a virtual (*taqdīrī*) declension and thus may only surface in one or two forms, while still being 'triptote'. Taking, for example, the noun *ʾafʿan* 'viper' one arrives at the paradigm in Table 4.

Table 4. Declension of nouns with weak third radical

sg.	n./g./a.	*ʾafʿan*	*al-ʾafʿā*	*ʾafʿā l-baladi*
du.	nom.	*ʾafʿayāni*	*al-ʾafʿayāni*	*ʾafʿayā l-baladi*
	g./a.	*ʾafʿayayni*	*al-ʾafʿayayni*	*ʾafʿayayi l-baladi*
pl.	n./g.	*ʾafāʿin*	*al-ʾafāʿī*	*ʾafāʿī l-baladi*
	acc.	*ʾafāʿiya*	*al-ʾafāʿiya*	*ʾafāʿiya l-baladi*

Noteworthy in this context are the triptote words *ʾab* 'father', *ʾax* 'brother', *ḥam* 'father-in-law', and *han* 'thing', all of which have a long final vowel in the position of the first term of an *ʾiḍāfa*, e.g., *ʾabū bakrin*, *ʾabī bakrin*, *ʾabā bakrin*, as well as the words *ḏū* 'owner/possessor of' and *fam* or *fū* 'mouth' (cf. Wright 1967:I, 249).

3. TRACES OF DECLENSION AND CASE IN THE ARABIC DIALECTS

In the Arabic dialects the tripartite case system has been lost, and the 'oblique' forms won out in the dual and plural (cf., for instance, Fischer and Jastrow 1980:41). As is well known, the same process can also be observed in the transition from Vulgar Latin to (most of) the Romance languages (e.g. Italian *notte* 'night' deriving from an oblique form of Latin *nox*). Taking again the example of *mallāḥun* 'seafarer' one arrives at the forms in Table 5.

Table 5. Case leveling in the Arabic dialects

		Classical Arabic	Arabic dialects
sing.	nom.	*mallāḥun*	*mallāḥ*
	gen.	*mallāḥin*	*mallāḥ*
	acc.	*mallāḥan*	*mallāḥ*
dual	nom.	*mallāḥāni*	*mallāḥēn*
	g./a.	*mallāḥayni*	*mallāḥēn*
pl.	nom.	*mallāḥūna*	*mallāḥīn*
	g./a.	*mallāḥīna*	*mallāḥīn*

The question as to the historical period during which this happened has been the subject of much debate (see, for instance, Rabin 1951:56–57; Talmon 2002:359–360; studies ranging from Fück 1952 vs. Wehr 1953 to Diem 1973 vs. Blau 1977 up to Owens 1998; cf. also Versteegh 2001:102–113 and *passim*). Fück (1950:2) adduces the following examples from the *Qurʾān*, in which the object precedes the subject; he argues that a reading without *ʾiʿrāb* would have made such examples incomprehensible, and that hence *ʾiʿrāb* could not have been lost in pre-Islamic times:

Q 35/28: *ʾinnamā yaxšā llāha min ʿibādi-hi l-ʿulamāʾu*
'out of his worshipers, only the scholars love God'

Q 2/124: *wa-ʾiḏi btalā ʾibrāhīma rabbu-hu bi-kalimatin*
'when his Lord put Abraham to the test'

Q 4/8: *wa-ʾiḏā ḥaḍara l-qismata ʾūlū l-qurbā* 'and when the relatives are present at the distribution'

Against these examples, however, Wehr (1953: 181) adduces *ʾiʿrāb*-less examples from Egyptian colloquial Arabic that are supposed to weaken Fück's argument, in that the syntax is perfectly comprehensible without *ʾiʿrāb*. Yet, both Fück and Wehr disregard the fact that most of their examples involve a human agent and a non-human non-agent. Ambros (1972) also raises the question of the functionality and redundancy of the Arabic declension.

At any rate, the loss of the case endings has gone hand in hand with the emergence of a range of genitive exponents and object markers in the Arabic dialects, resulting in the opposition of the traditional 'synthetic' constructions vs. the innovative 'analytical' constructions. Nowadays, it is just in adverbs that the *an*-ending of the accusative masc. sg. is retained (or 'borrowed' from the standard language), e.g. *dāyman* (< *dāʾiman* 'always'). Residues of → nunation (*tanwīn*) in a small number of Arabic dialects have usually adopted functions other than case marking (see, for instance, Fischer and Jastrow 1980:96, 120).

4. FUNCTIONS OF DECLENSION AND CASE IN CLASSICAL ARABIC

With respect to the functions of the Arabic terms for the three case endings, it must be kept in mind that the Arabic grammatical term *rafʿ* covers all functions of the *u*-ending, and the Arabic grammatical term *naṣb* covers all functions of the *a*-ending. The common terminology is meant not only to capture the similar vocalic pattern but also, more importantly, to reflect the observation that nominative and indicative, and accusative and subjunctive, have a syntactically comparable function, namely that of independent position vs. complement position respectively (cf., for instance, Bohas a.o. 1990:53–55; Versteegh 2001:79). The question is whether the grammatical analysis of the Arab grammarians was triggered by commonality in the vocalic pattern, or whether it arose independently (in the way that modern Transformational Grammar might approach such a problem). Here is a sketch of the analysis by the Arab grammarians:

u-ending (*rafʿ*)
noun *al-kitāb-u jamīl(un)*
DEF-book-NOM beautiful
'the book is beautiful'
verb *yaktub-u*
he:writes-IND
'he writes'

a-ending (*naṣb*)
noun *ʾurīdu l-kitāb-a*
I:want DEF-book-ACC
'I want the book'
verb *ʾurīdu ʾan yaktub-a*
I:want COMP he:write(s)-SUB
'I want that he write'

While the functional characterization of the *u*-ending (nominative: subject) and the *i*-ending (genitive: possession) is rather straightforward, the characterization of all the functions of the *a*-ending (namely case for the direct object, absolute negation, vocative in the *ʾiḍāfa*, predicative participle (*ḥāl*), predicate of '*kāna* and its sisters' (→ *kāna wa-ʾaxawātu-hā*) and focalized subject preceded by *ʾinna*) as 'accusative' is problematic. This may be illustrated by the following examples (cf. Sasse 1984:119–120):

i. accusative: *ḍaraba zaydan* 'he hit Zayd'
ii. absolute negation: *lā ʾilāha ʾillā llāh* 'there is no god except Allāh'
iii. vocative in the *ʾiḍāfa: yā ʾabā bakr* 'oh ʾAbū Bakr!'
iv. predicative participle: *jāʾa qāʾilan* 'he came saying'
v. predicate of *kāna wa-ʾaxawātu-hā: kāna malikan* 'he was a king'
vi. focalized subject: *ʾinna zaydan kabīr* '(indeed,) Zayd is big'

Among others, Sasse (1981:142) and Lipiński (2001:259–267) have argued that Afro-Asiatic originally had an ergative subject ('agent') case associated with an *u*-ending and an predicative/absolutive ('non-agent') case, which also served for marking the object, associated with an *a*-ending. Along this line of argumentation, the cases (ii)–(vi) can be said to reflect the latter predicate case, of which the 'accusative' is just a functional subset of the array of functions associated with the *a*-ending.

5. CONCLUSION

Even those who believe in the existence of a 'Proto-Semitic' entity have doubts as to whether the tripartite case system, as found in Akkadian and Arabic nouns and adjectives, can be reconstructed for this entity. Other early representatives of Afro-Asiatic such as Ancient Egyptian did not have a case system. Classical Arabic in its traditional description featured a tripartite case system, as outlined above. The case system of Classical Arabic broke down quite early in the history of Arabic and has ceased to exist in the modern Arabic dialects.

BIBLIOGRAPHICAL REFERENCES

Ambros, Arne. 1972. "Funktionalität und Redundanz in der arabischen Kasusdeklination". *Wiener Zeitschrift für die Kunde des Morgenlandes* 63–64. 105–127.
Blau, Joshua. "The beginnings of Arabic diglossia: A study regarding the origins of Neoarabic". *Afroasiatic Linguistics* 4:4.1–28.
Bohas, Georges, Jean-Patrick Guillaume, and Djamel Eddine Kouloughli. 1990. *The Arabic linguistic tradition.* London and New York: Routledge.
Brockelmann, Carl. 1908. *Grundriss der vergleichenden Grammatik der semitischen Sprachen*, I. Berlin: Reuther und Reichard.
Diem, Werner. 1973. "Die nabatäischen Inschriften und die Frage der Kasusflexion im Altarabischen". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 123.227–237.
——. 1975. "Gedanken zur Frage der Mimation und Nunation in den semitischen Sprachen". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 125. 239–258.
Fischer, Wolfdietrich and Otto Jastrow (eds.). 1980. *Handbuch der arabischen Dialekte.* Wiesbaden: O. Harrassowitz.
Fleisch, Henri. 1998. "Iʿrāb". *Encyclopaedia of Islam.* 2nd ed., III, 1248–1250. Leiden: E.J. Brill.
Fück, Johann. 1950. *Arabiya: Untersuchungen zur arabischen Sprach- und Stilgeschichte.* Berlin: Akademie-Verlag.
Huehnergard, John. 1997. *A grammar of Akkadian.* Winona Lake, Ind.: Eisenbrauns.
Lipiński, Edward. 2001. *Semitic languages: Outline of a comparative grammar.* 2nd ed. Leuven: Peeters.
Owens, Jonathan. 1998. "Case and proto-Arabic". *Bulletin of the School of Oriental and African Studies* 61:1.51–73.
Rabin, Chaim. 1951. *Ancient West-Arabian: A study of the dialects of the western highlands of Arabia in the sixth and seventh centuries A.D.* London: Taylor's Foreign Press.
Sasse, Hans-Jürgen. 1981. "Afroasiatisch". *Die Sprachen Afrikas*, ed. Bernd Heine, Thilo Schadeberg and Ekkehard Wolff, 129–148. Hamburg: Buske.
——. 1984. "Case in Cushitic, Semitic, and Berber". *Current progress in Afro-Asiatic linguistics*, ed.

John Bynon, 111–126. Amsterdam: J. Benjamins.

Talmon, Rafael. 2002. "Grammar and the *Qurʾān*". *Encyclopaedia of the Qurʾān*, II, 345–369. Leiden: E.J. Brill.

Versteegh, Kees. 2001. *The Arabic language.* 2nd ed. Edinburgh: Edinburgh University Press.

Vollers, Karl. 1906. *Volkssprache und Schriftsprache im alten Arabien.* Strassburg: K.J. Trübner.

Wehr, Hans. 1952. Review of: Fück (1950). *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 102. 179–186.

Wright, William. 1967. *A grammar of the Arabic language.* Trans. from the German of Caspari. 3rd ed., rev. W. Robertson Smith and M.J. de Goeje. 2 vols. Cambridge: Cambridge University Press.

Lutz Edzard (University of Oslo)

Defective Verbs

Amongst those verbs that have either a semivowel or a *hamza* as one element of their root only a few irregularities tend to occur. Moreover, these irregularities are restricted to certain frequently used verbs, which in some forms lose the weak element, e.g. the loss of *hamza* in some forms of the common verb *raʾā/yar(ʾ)ā* 'to see' (→ weak verbs). Apart from this, there are some verbs and verbal expressions that are defective and/or which can only partly be subsumed under the usual verbal paradigms.

1. Defective verbs

Defective verbs show normal conjugation patterns but certain grammatical forms are never used. Thus, in Classical Arabic there is an imperfect *yadaʿu* with the imperative *daʿ* 'to let alone' and *yaḏaru* with the imperative *ḏar* 'idem'; neither verb has perfect forms **wadaʿa* and **waḏara*. As to its cognate roots in other Semitic languages, presumably, the Hebrew hapax form *mûdå̄ʿîm* (modified from *mûʿådîm*) '[they were] set' (Jer 24:1) is derived from the root *y-d-ʿ* (Sem. **w-d-ʿ*).

Another defective verb is *ʿasā* 'perhaps', of which only a few perfect forms are documented (e.g. 1st pers. sg. *ʿasaytu*). It may be assumed that originally *ʿasā* was a particle, which was later interpreted as a verb and received verbal treatment. From amongst the many usages that Ullmann (1984) collected only the two most frequent will be mentioned here:

i. *ʿasā* ([subject]) *ʾan* [jussive], e.g. *ʿasā llāhu ʾan yatūba ʿalayhim* 'perhaps God will relent toward them' (Q. 9/101);
ii. *ʿasā* ([subject]) *ʾan* [indicative], e.g. *ʿasā qīlun yadūmu lanā* 'perhaps the [extolling] words will outlast us'.

2. Laysa and Lāta

The pseudo-verb *laysa* 'he is not' is not only defective because it only has perfect forms, but it is also irregular. The two morpheme variants are *lays-* (as in *laysat* 'she is not') and *las-* (as in *lastu* 'I am not'). Classical Arabic grammar would like to see in this a shortening from **layisa*. But comparative studies in other Semitic languages suggest that we are dealing with an old compound of the negative particle *lā* and an existential expression (like Hebrew *yẹš* and Aramaic *ʾīṯ*). The oldest cognate expression occurs in Akkadian *lā isû(m)* 'not to have', *laššu* 'not having, a have-not; is/are not' (cf. Barth 1913; Blau 1972; Fleisch 1979; Ullmann 1970, s.v.).

In the *Qurʾān laysa* has an elucidating function, e.g. Q. 88/22: *lasta ʿalayhim bi-muṣayṭirin* 'thou hast no power over them'.

The expression *lāta* 'it is/was all over [usually with expression of time]', now only used as a particle, is etymologically related, cf. Q. 38/2: *wa-lāta ḥīna manāṣin* 'there was no time for deliverance' (see Barth 1913; Ullmann 1970, s.v.).

3. niʿma and biʾsa

The two pseudo-verbs *niʿma* 'how good is . . .!' and *biʾsa* 'how bad is . . .!' are called in Arabic 'verbs of praise and blame'. They represent a shortening of **naʿima* and **baʾisa* respectively. These expressions are normally fossilized, but supposedly other forms in the 3rd pers. do exist (e.g. sg. fem. *niʿmat* and *biʾsat*). Along with the frequent type *niʿma ṣ-ṣāḥibu zaydun* 'how excellent is Zayd as companion!' we also find constructions with *mā* like: *la-biʾsa mā kānū yafʿalūna* 'wretched be what they wrought' (Q. 5/79), cf. also *niʿim-mā* (= *niʿma mā*) in Q. 4/58: *ʾinna llāha niʿim-mā yaʿiḏ̣ukum bihi* 'good is what God admonishes you with!'.

4. ʾafʿāl at-taʿajjub

The name 'verbs of surprise or wonder' covers two different syntactic constructions, which express, however, the same meaning.

i. *mā ʾafʿala zaydan* ‘what an excellent man Zayd is!’
In this construction predominantly active verbs describing a state are used. This is folk-etymologically motivated from its assumed original meaning (‘what has made Zayd excellent?’).

Yet, in reality it is a fossilized → elative (*ʾafʿal*) in the accusative (Brockelmann 1913:12). This explains the special condition that this construction is only possible with triradical verbs, and also its strong inflection with roots II*w/y*. The construction *mā ʾajwada-hu* ‘how generous he is!’ shows clearly that we are not dealing with a perfect form of Form IV, but with an original elative *ʾajwad* ‘excellent’.

ii. *ʾafʿil bi-zaydin*
The alternative construction with the same meaning looks like an imperative of Form IV, e.g. *ʾasmiʿ bihim ʾabṣir* ‘how well they hear, how well they see’ (Q. 19/39). An essential feature is the strong treatment of the second radical, e.g. *ʾajwid bihi*. This construction appears to be a secondary formation, based on the interpretation of *mā ʾafʿala-hu* as a formation of Form IV.

Bibliographical references

Barth, Jakob. 1913. “Arab. lāta ‘es ist nicht’”. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 67.494–496.
——. 1914. “Die Etymologie von arab. *ʾin* ‘nicht’, *laisa* ‘nicht sein’”. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 68.360–364.
Blau, Joshua. 1972. “Marginalia Semitica, II”. *Israel Oriental Studies* 2.57–82 (58–61).
Brockelmann, Carl. 1913. *Grundriß der vergleichenden Grammatik der semitischen Sprachen*, II. Berlin: Reuther und Reichard.
——. 1962. *Arabische Grammatik*. 15th ed. Leipzig: O. Harrassowitz.
Caspari, Carl Paul: see Wright (1964).
Fleisch, Henri. 1961, 1979. *Traité de philologie arabe*. 2 vols. Beirut: Imprimerie Catholique.
Ullmann, Manfred. 1983–. *Wörterbuch der klassischen arabischen Sprache*, II (*lām*), 1–4. Wiesbaden: O. Harrassowitz.
——. 1984. *Arabisch* ʿasā “*vielleicht*”: *Syntax und Wortart*. Munich: Verlag der Bayerischen Akademie der Wissenschaften.
Wright, William. 1964. *A grammar of the Arabic language, translated from the German of Caspari . . .*, I. 3rd ed. Cambridge: Cambridge University Press.

Rainer Voigt (Free University of Berlin)

Deixis

1. Definition

The term ‘deixis’ (also *deictic expressions* or *shifters*) indicates a group of linguistic elements attested in all languages, whose meaning necessarily implies a return to the uttering act in order to find a particular referent. ‘Deixis’ is a borrowing from Ancient Greek, which originally signified ‘the action of showing’. In fact, deixis draws the attention of the interlocutor(s) to a referent that is present in the situation of communication through the aid of specific words, such as demonstratives in *Look at this beautiful painting!*, which might be said while visiting a museum. The referent is shown to the addressees, if need be by pointing a finger. The demonstrative *this*, as the name suggests, is a form of deixis.

The closest Arabic term to ‘deixis’ is *ʾišāra*. The term is most often associated with *ism* ‘noun/nominal’ used as a demonstrative pronoun, which is an element of a deictic nature. The term is also employed by some early Arab grammarians in the basic lexical meaning ‘indication’. Since antiquity, the grammatical tradition has recognized the specificity of these terms, which only have a determined value through the momentary existence conferred on them by the utterance in which they appear and which refers back to different localized referents in the act of communication: the speaker (‘I’), the addressee (‘you’) and, in a non-anaphoric interpretation, an object that is present (‘this painting’). To interpret these deictic expressions, we must refer to the utterance. Their reference varies with each particular situation: *I’ll be back in ten minutes* can be said or written at different moments and in different places by a secretary, a teacher, etc.

Far from being isolated units, these terms form a veritable system, that of “indexical expressions” (Bar-Hillel 1954), or “the formal apparatus of the utterance” (Benveniste 1970), which has an essential role in the mechanism of the function of language (Kuryłowicz 1972) and which allows the speakers to appropriate language in order to convert it into discourse and state their position as speakers. The coded meaning of these terms refers back to their use in utterances. Thus, a particular occurrence of 'I' designates the person who states this occurrence. They are then more or less linked to different constitutive elements in the situation in which the utterance was made.

Besides person, temporality (represented by the present tense, the origin of our bearings in temporality, but also by the circumstants as for example *now* and *tomorrow*) and location (*here* 'where I am', *here*/*in this place* 'where I can be found') likewise permit one to 'anchor' the utterance with regard to the specific situation of use, i.e., with regard to the 'I-here-now' of the speaker.

In a synthetic presentation of the demonstrative in Arabic, Fleisch (1979:28–73) compiled a series of deictic elements, of 'demonstrative bases', in other words morphemes (mono-consonantal for the most part, which can be found in affective language) with a deictic signification: *ḏ* (which appears in demonstratives like *hāḏa* 'this, this one', *ḏū* '[of a person] possessing', relative pronouns such as *allaḏī* 'who, which', interrogative pronouns *māḏā* 'what', *man ḏā* 'who is it that', adverbial markers like *ʾiḏ* 'then'); *t* (in the feminine demonstrative *tā*, *tī*; demonstrative adverbs like *ṯammata* 'here'); *k* (in the particle *kayfa* 'how'); *l* (in demonstratives like *ḏālika* 'that'); *ʾul* (in the demonstrative *ʾulāʾika* 'these, those'); *n* (in demonstrative adverbs such as *hunā* 'here'); *ḥ* (in *ḥayṯu* 'there, where'); *ṯ* (in *ṯamma* 'there'); *m* (in interrogative pronouns such as *man* 'who'); *f* (in the conjunction *fa*- 'next, then'); and *h* (in demonstrative pronouns and adverbs such as *ḥāḏā* 'this, this one', *hunā* 'here'), among others. These bases represent the constitutive elements of deictic words in Arabic, whose long established usage places them outside the general system of language: they are agglutinated lexical forms, without internal inflexion.

Several deictic categories should be considered in Arabic: personal, spatial, and textual.

2. Personal deixis

Although traditional grammar associates the pronouns 'I' and 'you' with 'he', limiting the class of shifters dissociates the 'I–you' couple (the veritable 'persons' in a dialogue) from the pronoun 'he'. The category of personal deixis includes the pronouns at the level of speaking, which necessarily refer to human entities, including what are traditionally called *al-mutakallim* 'the speaker' (i.e. the 1st person) and *al-muxāṭab* 'the spoken to, the interlocutor' (i.e. the 2nd person). *ʾana* 'I' and *ʾanta*/*ʾanti* 'you [masc./fem. sg.]' refer to the roles of the speaker and the addressee, which cannot be dissociated or reversed. What is understood by 'I' and 'you' refers in fact to a class larger than the two corresponding terms, i.e. *aḍ-ḍamīr al-munfaṣil* 'independent pronoun', *aḍ-ḍamīr al-muttaṣil* 'clitic pronoun' like *-ī* 'I', *-ka* 'you [masc. sg.]', *-ki* 'you [fem. sg.]', or the suffixes added to the perfect verbal form *al-māḍī* (*-tu* 'I', *-ta* 'you [masc. sg.]', *-ti* 'you [fem. sg.])' and the personal prefixes attached to the imperfect verbal form *al-muḍāriʿ* (*ʾa-* 'I', *ta-* 'you [sg.])'. They also figure in the plural forms (*al-jamʿ*) and the dual (*al-muṯannā*); independent and clitic forms like *naḥnu*, *-nā*, *na-* 'we'; *ʾantum*, *ʾantunna*, *-tum*/*-tunna*, *ta-* 'you [masc./fem. pl.]'; *ʾantumā*, *-tumā*, *ta-* 'you two [masc./fem]'.

All these personal forms have a deictic value in the sense that their meaning integrates a relationship with the speaker or the addressee(s) different from the 3rd person (*al-ġāʾib* 'the absent') like *huwa*/*hiya* 'he/she/it' and the different morphological variants that Benveniste (1970) prefers to place at the level of what he calls the 'non-person', that of objects of the world other than speakers, given that this also applies to things and abstractions.

3. Spatial deixis

Spatial deixis terms (like *here* and *there*) can be interpreted by taking into account the position of the body of the speakers and their gestures. From a morphosyntactic point of view, spatial deixis appears quite heterogeneous, and terms can be classified basically into two groups: → demonstratives (*ʾasmāʾ al-ʾišāra*), which, according to Brockelmann (1908:296), originate in interjections (like *hāḏā* 'this, this one', *ḏālika* 'that', *ʾulāʾika* 'these, those there'), and adverbials (such as *hunā* 'here', *ṯamma* 'there', *ṯammata* 'over there').

In general, the demonstratives combine lexical meaning and deictic value, and they are used to constitute noun phrases referring to an object present in the situation of communication, especially if they are accompanied by a gesture of showing something: *xuḏ hāḏā l-kitāb/xuḏ hāḏā* 'Take this book/this one!'.

Spatial deixis (adverbs or prepositions used adverbially, adjectives, prepositional phrases), often with the status of circumstantial complements, can be found in diverse micro-systems of opposition: *hunā/hunāka* 'here, there/over there', *ʾamāma/ warāʾa* 'in front/behind', *ʾilā l-yasār/ʾilā l-yamīn* 'right/left', *qabla/baʿda* 'before/after', *hāḏihi z-zuhūr/tilka z-zuhūr* 'these flowers/those flowers', *hāḏāni ṣ-ṣadīqāni/ḏānika ṣ-ṣadīqāni* 'these two friends/those two friends'. They acquire meaning with regard to the gesture, position or orientation of the speaker's body. Any change in one of these parameters correlatively modifies the objects susceptible of being localized: if the speaker should turn around, that which was 'in front' is now 'behind', that which was 'left' is now 'right', etc. Among these semantic axes of opposition, language indisputably favors the opposition between near and far that we find in *hunā/hunāka* 'here/over there', *hāʾulāʾi/ʾulāʾika* 'these here/those there'.

Within this system of opposition, many Arab grammarians such as az-Zamaxšarī (Fleisch 1970:44–45) recognize for demonstrative pronouns a tripartite system, three *marātib* 'positions': *al-qurbā* 'the nearest [to the speaker]' (in the case of simple morphemes such as *ḏā* 'this [masc.]', *ḏāni* 'these [masc. dual]', *tāni* 'these [fem. dual]', *ʾūlā* 'these', etc.); *al-wusṭā* 'the mid-distanced' (for demonstratives which include the deictic base *k* as in *ḏāka* 'that [masc.]', *ḏānika* 'those [masc. dual]', *ʾulāka* 'those'); and *al-buʿdā* 'the farthest' (for demonstratives constructed with the base *l* as in *ḏālika* 'that there [masc. sg.]', *ʾulāʾika* 'those there'). The same is true for the demonstrative adverbs *hunā* 'here, *hunāka* 'there', *hunālika* 'over there'. Fleisch (1970: 45–46), citing as-Suyūṭī, rejects this triple division and pleads in favor of another interpretation, which sees in this remarkable symmetry the result of a dialectal sharing. He considers the demonstratives of Arabic to be organized on two levels: near, centered on the speaker (simple demonstratives), and far, centered on *al-ġāʾib* 'the absent, the non-person' (forms including the bases *k* and *l*). Yet, it should be pointed out that the deictic base *k* is also present in the configuration of the 2nd person of the personal enclitic pronoun, *aḍ-ḍamīr al-muttaṣil al-muxāṭab* (*-ka* 'you [masc. sg.]', *-ki* 'you [fem. sg.]', *-kumā* 'you two [masc./fem.]', *-kum* 'you [masc. pl.]', *-kunna* 'you [fem. pl.]'), which refers back to the addressee situated at a relatively close distance.

Besides the easily recognizable spatial deixis presented so far, there exist deictic phenomena no less frequent, yet more delicate. This is the case in particular for the opposition between verbs such as *ḏahaba* 'to go' and *ʾatā* 'to come'. From an objective point of view, nothing distinguishes *yaḏhabu zayd ʾilā maktabi-hi* 'Zayd goes to his office' from *yajīʾu zayd ʾilā maktabi-hi* 'Zayd comes to his office', but the verb *ʾatā* 'to come' is used only if the agent in the process is moving toward the speaker at the moment this process happens/will happen.

Adverbial demonstratives are also used in the construction of → locative clauses (*al-jumal aḏ-ḏarfiyya*): *hunāka muškila* 'there is a problem [lit. 'there a problem']'. In this example, the localizer does not indicate a precise place; in such a case, *hunāka* 'there' does not designate a place that one could point at. The sense of the sentence is abstract: it expresses the existence of a problem. The demonstrative localizers *hunāka* 'there', *hunālika* 'over there', *ṯamma* 'there' and *ṯammata* 'over there', which express distancing and are frequently employed in Modern Arabic, become the indicators for localization (Anghelescu 1995:66–67), where space is the most spread out to include existence.

4. Temporal deixis

Temporal deixis originates at the moment when the speaker is speaking. The temporal complements are therefore identified in relation to the moment the utterance is made, and can show a coinciding (*now*, *at this moment*, *today*), a previous (*yesterday*, *last month*), or a posterior (*tomorrow*, *in a few days*, *next week*) interval. Besides adverbs, adjectives in phrases like *next/last week* can also have a deictic value.

It is more difficult to limit the class of temporal shifters to elements whose function is that of 'circumstants' and whose statute is that of an adverb (*ʾamsi* 'yesterday') or a prepositional group (*fī yawmayni* 'in two days'). Apart from these, there exists a triplet of shifters as 'essential' as they are 'subtle': the tense markers contained in the morphology of verbs, whether they concern the present, past, or future. A sentence like *raʾaytuhu ʾams* 'I saw him yesterday' includes not one but two temporal deictic elements: the adverb *ʾams* 'yesterday' and the 'past' associated with the paradigm of *al-māḍī*.

Verbal tenses are first identified in relation to the circumstances of utterance but do not situate the process in time in the same manner that noun phrases refer to their referent. Temporal localization is made by the entire utterance, and not by the verb alone, even if the verb has a determining role.

Arabic, which has only two fundamental paradigms, *al-māḍi* ('past' or 'perfect') and *al-muḍāriʾ* ('non-past' or 'imperfect'), attributes to each a vast spectrum of temporal, aspectual, and modal values which intersect and complement each other. In this respect, the terminology of Arab grammarians is heterogeneous, whereas that of linguists working with Arabic is usually homogeneous and focused on the aspectual values. The form *al-māḍī* includes the past or present perfect: *katabtu* can be translated as 'I wrote, I have (just) written'. The expression of the present, indefinite or continuous, but including all the nuances of the future tense, belongs to the *muḍāriʿ*: *ʾaktubu* can be translated as 'I am writing, I write, I shall write.' In fact, the context and value of verbs already used in a text play an important role in the interpretation one must give to a verbal form. Elsewhere, there exists a good number of 'word-tools' clarifying the exact value the speaker intends to give to the verbal form as, for example, the proclitic particles *sa-* or *sawfa-*, which unambiguously mark the future tense of the imperfect form (*sa-ʾaktubu* can only mean 'I shall/I am going to write'), or, in colloquial Arabic, the particles *ġadi* in Moroccan Arabic (*ġadi nəmšiw* 'we are going to leave'), *rāḥ* in Lebanese Arabic (*raḥ yəḍrobna* 'he is going to hit us').

The imperative is also deictic. The verb phrase is directed at one or more addressees at the particular time of utterance, namely the present tense as in *udxul* 'come in!' (in a poster saying *xuḏ* 'take!', for example, the time can be prolonged, but the communicative deictic function is still the same).

5. DEIXIS AND ANAPHORA OR TEXTUAL DEIXIS

Whenever a linguistic expression localizes its referent in the extra-linguistic situation of utterance, the reference is said to be 'exophoric' or 'deictic' (*I am not here*). On the other hand, if the referent in the expression can only be accessed through other segments of the text, the reference is called 'endophoric': 'anaphora' or 'anaphoric' (*Paul left. He forgot his book*), or → 'cataphora' (*Then I saw it. A furry little thing beside the bed*). In general, anaphora seems to be more common than cataphora.

The → coherence of a text depends in part on iteration. Diverse linguistic elements contribute to this. Noun phrases, in particular, realize by their usage and relationships throughout the discourse the recalling of information. The notion of 'anaphora' allows one to describe this aspect of textual organization, defined traditionally as the recalling of a previous element in a text.

Certain terms used to designate an element in a specific situation of utterance also allow anaphoric reference to an element in context. In discourse, for example, context plays a role analogous to that of utterance. If the deixis involves referring back to a referent that must be looked for in the speech situation, because of its extra-linguistic orientation, the anaphora refers back to a word (phrase, sentence, or part of discourse) in the linguistic context in adjacent sentences; such sentences are also part of the situation of the utterance. This explains why anaphora and deixis are generally represented by the same elements – sometimes deictic, sometimes anaphoric. At times this leads linguists to consider anaphora as being "textual deixis" (Levinson 1983:62–63). Yet, there is an obvious difference, noted by Kleiber (1991). Between the deixis and the referent the relationship is direct, but anaphora only makes a reference through the intermediary of a word it calls up. The relationship is therefore indirect. There are several types of 'anaphoric' or 'textual anaphoric deixis'. Among these are the following:

i. *Pronominal anaphora*. The use of a third person pronoun (*aḍ-ḍamīr al-ġāʾib*), a demonstrative (*ism al-ʾišāra*), and a relative pronoun (*al-ism al-mawṣūl al-xāṣṣ*) to avoid the repetition of a noun phrase, a name, a sentence, or a group of sentences: *zayd fī r-ribāṭ. huwa muʿallim* 'Zayd is in Rabat. He is a teacher', *ar-rajulu l-laḏī waṣala ṣadīqī* 'The man who arrived is my friend', *ar-rajulu l-laḏī ḏahabtu maʿa-hu ṣadīqī* 'The man with whom I left is my friend (lit.: ‘The man that I left with-him my friend’)’. The anaphoric usage of personal pronouns is present in the qualifying expansions of the noun phrase with *tawkīd* ‘emphasis’ as in *al-mudīr nafsu-hu* ‘the director himself (lit.: ‘the director soul-[of]-him’)’, *an-nās kullu-hum* ‘all the people (lit.: ‘the people all-them’)’, and adjective clauses (*bint ʾaʿrifu ʾaxā-hā* ‘a girl whose brother I know (lit.: ‘a girl I know brother-[of]-her’)’. Used as pronouns, demonstratives, especially *hāḏā* ‘this’and *ḏālika* ‘that’, call up again one or several sentences already evoked: *ʾanhaytu ʿamalī. hāḏā jayyid* ‘I finished my work. That’s good.’

ii. *Nominal anaphora* refers to noun phrases including definite → determiners such as the definite article, possessive determiners, or demonstratives. These noun phrases can take several forms and maintain several types of relations with their antecedent: the ‘accurate’ anaphora (or the recalling of a name by simply changing the determiner) as in *kataba ḥasan kitāban. al-kitāb/hāḏā l-kitāb/kitābu-hu mumtāz* ‘Hassan wrote a book. The/this/his book is excellent’; the ‘inaccurate’ anaphora (which calls for a lexical change) as in *nagīb maḥfūẓ miṣrī. muniḥa hāḏā l-muʾallif jāʾiza nōbil* ‘Naguib Mahfouz is Egyptian. This writer has been awarded the Nobel Prize’; the ‘conceptual’ anaphora (which condenses and sums up the contents of a sentence or an entire fragment of a previous text) as in *ʾaxū-hu ṣadamat-hu s-sayyāra. hāḏā l-ḥādiṯ taraka la-hu muxallafāt badaniyya* ‘His brother was hit by a car. This accident left him with scars’.

iii. *Adverbial anaphora* refers to instances where anaphora is extended beyond nominal expressions. A locative adverb, for example, can refer to a previously mentioned localization as in *yaʿīšu fī firansā. laʿalla ṭ-ṭaqs jamīl hunālika* ‘He lives in France. Perhaps the weather is nice there’.

iv. *Verbal anaphora* is expressed by using the verb *faʿala* ‘to do’, which represents a verb denoting a process capable of representing a previous verb phrase: *hal katabta r-risāla? naʿam, faʿaltu hāḏā ʾams*. ‘Did you write the letter? Yes, I did it yesterday’.

The semantic relations and anaphoric functions that the deixis of Arabic presents are common among natural languages. Since many linguistic elements in an utterance can refer directly to things in the outside world, deixis may be considered a specific kind of reference.

BIBLIOGRAPHICAL REFERENCES

Anghelescu, Nadia. 1995. “Le système des déictiques en arabe”. *Proceedings of the Colloquium on Arabic Linguistics Bucharest, August 29–September 2, 1994*, ed. Nadia Anghelescu and Andrei A. Avram, 55–72. Bucharest: University of Bucharest.

Bar-Hillel, James. 1954. “Indexical expressions”. *Mind* 63.354–379.

Benveniste, Emile. 1970. “L’appareil formel de l’énonciation”. *Langages* 17.12–18.

Brockelmann, Carl. 1908. *Grundriss der vergleichenden Grammatik der semitischen Sprachen*. I. *Laut- und Formenlehre*. Berlin: Reuther und Reichard.

Fleisch, Henri. 1970. *Traité de philologie arabe*. II. *Pronoms, morphologie verbale, particules*. Beirut: Dar el-Machreq.

Kleiber, Georges.1991. “Anaphore-deixis: Où en sommes nous?”. *L’Information Grammaticale* 51. 3–18.

Kuryłowicz, Jerzy. 1972. “The role of deictic elements in linguistic evolution”. *Semiotica* 5.174–183.

Levinson, Stephen. C. 1983. *Pragmatics*. Cambridge: Cambridge University Press.

MIHAI DAT (University of Bourgogne)

Demonstratives

Demonstratives in Arabic dialects have different forms, depending on the region. In order to describe the main paradigms, this entry generalizes as much as possible and describes their general features rather than individual dialectal details.

In all Arabic dialects, there are two series of demonstratives, one conveying the idea of near deixis, with respect to the speaker, ‘this, these’, the other the idea of far deixis, which is associated with the interlocutor, ‘that, those’. This difference is not limited to a space opposition near/far, it can also refer to a modal opposition of an affective/emotional nature (Caubet 1993:I, 168).

The various existing paradigms can be grouped into three categories, according to different types of dialects: on the one hand, Bedouin dialects, without geographical distinction, on the other hand, sedentary dialects, both eastern and western. It should be emphasized, however, that dialect frontiers are never impermeable.

Demonstratives may have a long form and a short one, and the use of these forms varies according to the region, although generally the long form is the most widely used. However, in some regions the abbreviated forms are the more common, as in Egyptian dialects.

The usual position of demonstratives is before the noun, but in some cases the attributive demonstrative is placed after the noun, as in Egyptian and Sudanese Arabic. In these dialects, the normal pattern is Noun-Demonstrative, although in the near deixis the word order Demonstrative-Noun is also possible, with specific structural and semantic connotations (Woidich 1992). According to Doss (1979:351), this word order is currently a residue of a previous stage in the evolution of Egyptian dialects toward the Cairene koine, which contrasted with the usual Noun-Demonstrative word order. In other dialects, the demonstrative can either precede or follow the noun, as is the case with the Central Arabian dialects, where the southern dialects use both the short form, either preposed or following the noun, and the longer forms, whereas in the northern dialects this is impossible (Ingham 1994:55).

The demonstratives in some Bedouin dialects are as in Table 1.

Table 1. Demonstrative pronouns in Bedouin dialects (Rosenhouse 1984:20; de Jong 2000:43)

	near deixis long form	short form
m. sg.	*hāḏ̣a*	*ḏ̣a*
f. sg.	*hāḏi*, *hēḏiy*	*ḏi*
m. pl.	*hāḏ̣ōl*, *hōḏ̣aḷ*	*ḏ̣ōl*
f. pl.	*hāḏ̣an*, *haḏ̣ōlin*	*ḏ̣ōl*, *ḏ̣ōlin*
	far deixis long form	short form
m. sg.	*haḏ̣āk*	*ḏ̣āk*, *hā*, *ha-*
f. sg.	*haḏīk*, *haḏīč*	*ḏīk*, *ḏīč*, *hāy-*, *hay-*
m. pl.	*hāḏ̣lāk*, *haḏ̣ōlāk*	*ḏūk* *ḏ̣ōlāk*
f. pl.	*haḏīkan*, *haḏīčan*, *haḏ̣ōlič*	*ḏōlīč*

The gender distinction in the singular is widespread in the Arabic dialects, but in the plural it only exists in eastern Bedouin dialects, whereas in the western ones there is a common form under the influence of sedentary neighbors. The same applies to the form of expressing the near and the far deixis, of which the forms of the Ḥassāniyya dialect in Mauretania are a good example: near deixis: *hāḏä* (*ḏä*), *hāḏi* (*ḏi*), *hāḏu* (*ḏu*) and far deixis: *ḏāk*, *ḏīk*, *ḏūk* (Ould Mohamed Baba 1999:155, 2000–2001:261). As for the eastern Bedouin dialects, not all of them have plural feminine forms, namely the dialects of the northern Sinai littoral, some of which have the same plural form for both genders and others have feminine plural forms, e.g. *hādaḷḷāka*, *hadinna*, or *hāḏoḷḷayn* 'these' (de Jong 2000:369–370, 451).

With regard to number, a cross-dialectally uniform characteristic is the absence of a dual demonstrative pronoun, continuously receding in Neo-Arabic. This feature is even found in the Bedouin dialects, where there is still some residue of the dual form in the nominal area. The singular is more stable, while in the plural we find several different forms. Some plural forms are typical of some areas only, as in the case of *ḏanni* and *hāḏanni*, typical of the dialects of the southern area of Mesopotamia (Ingham 1982:83).

Table 2 gives an example of demonstratives in an eastern sedentary dialect.

Table 2. Demonstrative pronouns of the Cairo dialect (Zaki 1972:126)

	near deixis	far deixis
m. sg.	*da*, *dah*	*dokha*
f. sg.	*di*, *diyya* , *diyyat*	*dikha*, *dikhat*
c. pl.	*dōl*	*dokham*, *dokhom*, *dikhom*

Here, we note the distinction of gender in the singular and the use of a common form in the plural, this being a typical feature of the process of reduction that is more typical of sedentary than Bedouin dialects. In the Cairo dialect, the forms for far deixis are less frequent, being replaced usually by those for near deixis, which have a wider semantic value. Besides, the masculine singular form *da* is used to express the neutral gender.

An example of demonstratives in a western sedentary dialect is given in Table 3.

Table 3. Demonstrative pronouns of the Fes dialect (Caubet 1993:168–169)

	near deixis long form	short form	far deixis long form	short form
m. sg.	*hāda*	*hād*	*hādāk*	*dāk*
f. sg.	*hādi*	*hād*	*hādīk*	*dīk*
c. pl.	*hādu*	*hād*	*hādūk*	*dūk*

Gender distinction exists only in the singular, as in the eastern sedentary dialects; concerning number, the only remarkable feature is the absence of a dual form. In this case, the neutral gender is expressed adding the particle *ši* to the masc. sg. form: near deixis *hād ši*; far deixis *dāk ši*.

As in most Moroccan dialects, we note that the use of the long and the abbreviated form differs according to their syntactic function, since the former acts as a pronoun, the latter as an attributive demonstrative. Another special feature is the presence of an invariable form in gender and number in the short form for the near deixis, whereas in the far deixis the form varies both in gender and number. Such a feature is not general, since in some Moroccan dialects, especially the ones in the northern regions, the far deixis has only one invariable form, *dīk* (cf. Heath 2002:272).

All these forms used as demonstratives in different Arabic vernaculars have variant forms, depending on the dialects. For instance, a common feature of almost all Bedouin dialects is the presence of the interdental phonemes, both /ḏ/ and /ḏ̣/, which vary depending on the case. Thus, for example, in some dialects the emphatic does not appear in contact with /i/ (Rosenhouse 1984:20; de Jong 2000:43). In the sedentary dialects, however, this phoneme is normally occlusive, /d/, although there are some exceptions. There are also differences between dialects due to the presence or absence of → *ʾimāla*, Thus, in the dialect of Sūsa some demonstratives are pronounced: *hǟḏä* 'this [masc.]', *hǟḏi* 'this [fem.]' (Talmoudi 1980:148). This is also the case in some Syrian dialects, where we find the forms: *hādi* 'this [masc.]', and *hāde* 'this [fem.]' (Behnstedt 1993:77). Another example, regarding far deixis demonstratives, consists of the different forms of the phoneme /k/, which in some dialects is pronounced as an affricate /č/, a phenomenon known as → *kaškaša*, which occurs in some Bedouin type dialects. On the other hand, in some western dialects, mainly northern sedentary rural ones, this phoneme is fricative /k̠/, because of Berber Rif dialect influence.

Apart from these forms, there is also the demonstrative *ha-* which is found in some eastern and in some western dialects; according to Fischer (1959:45), it occurs in several dialects of Libya, Tunisia, Palestine, Syria, Lebanon, and Mesopotamia. However, its recurrence and form vary, depending on the type of dialect. In Bedouin dialects of northern Israel, there are forms *hā~ha-* for the masculine and *hāi-~hai-* for the feminine (Rosenhouse 1984:20), and in the Bedouin dialects of the northern Sinai littoral the usual form is *hal* (= *ha* + article *al*). De Jong (2000:172) believes this form has less deictic value than the aforementioned demonstrative forms: "a deicticized article, specifying some object(s), person(s), or abstraction(s) not physically present or demonstrable at the moment of the utterance, but which/who is/are present in the mind of the speaker, not in the mind of the hearer".

In some eastern sedentary dialects, this demonstrative also exists, for example in the Syrian ones, where according to Behnstedt (1993:77) it is a more common form in the western Syrian dialects than in those spoken in the eastern regions of the country. In western type dialects where the demonstrative *hā* exists (mainly of Tunisia and Libya) it is invariable in gender and number (Marçais 1977:197).

Apart from the near and far deixis already mentioned, there are other forms that show an intermediate type of deixis (middle deixis), but this can only be found in some Arabic vernaculars. This is the case of the Jiblah dialect in Yemen, where the following forms occur: *hāḏkāh*, *hāḏkíh*, *hāḏkúh* (Fischer and Jastrow 1980:116), Bəḥzānī Arabic in Iraq, with the forms: *hākā*, *haykā*, *hawkā* (Fischer and Jastrow 1980:151), and Daragözü Arabic *ukkā*, *ukkī*, *ukkō* (Jastrow 1973:41).

Furthermore, there are some forms made up of a demonstrative element and some added morphemes that vary depending on the type of dialect. They are part of an independent semantic development and function also as middle deixis. Thus, in Egyptian, the following forms can be found: *ʾahú*, *ʾahí*, and *ʾahúm*, e.g. *ʾahú ga* 'here is, has come', *ʾis-sitt maʿāya ʾahí* 'the woman is with me right here', *ʾahúm il-kutub* 'here are the books', which may refer to a noun or to the subject of a verb and do not have a fixed position within the sentence. In western type dialects, we have the particles *hā-* and *ṛā-* to which the independent or suffixed personal pronouns are added, the latter only in the case of the particle *ṛā* or with the 1st person singular in the case of *hā*, e.g. *ṛāhu* 'there he is', *ṛāk* 'there you are', *ṛāni* 'there I am', *hākum* 'there you are', *hāhūwa* 'there he is', *hā-hūma* 'there they are' (see Caubet 1992). All these forms are used to denote the existence of someone or something, and they also have a demonstrative value, which is sometimes regarded as a middle deixis.

Demonstratives may have two syntactic roles, attributive and pronominal. In the former role they act as attributive adjectives and, depending on the dialect, they may have a long or an abbreviated form and precede or follow the noun, which is always determined, e.g.:

Near deixis:
Bedouin dialect (Ḥassāniyya)*: hāḏä eṛ-ṛāžəl* 'this man'
eastern sedentary dialect (Egyptian): *ir-rāgil da* 'this man'
western sedentary dialect (Moroccan): *hād ṣ-ṣīf* 'this summer'

Far deixis:
Bedouin dialect (Ḥassāniyya): *ḏāk l-gahwa* 'that coffee'
eastern sedentary dialect (Egyptian): *ir-rāgil dokha* 'that man'
western sedentary dialect (Moroccan): *dāk ṣ-ṣīf* 'this summer'

In this case, the long form is the most common in Bedouin dialects, whereas in sedentary dialects the abbreviated form is the more common.

Demonstratives with pronominal value act as a noun, which they replace. The form of the demonstrative also varies, depending on the type of dialect. An important feature is that in their role as a pronoun they may act as the subject of both a nominal or a verbal utterance, e.g.:

Bedouin dialect (Ḥassāniyya): *hāḏu yətkallmu klām yəngāl-lu klām et-təgṛāḥ* 'these speak a language considered pedantic'
eastern sedentary dialect (Egyptian): *da l-mudīr wi-di mrātu* 'that is the director and that is his wife'
western sedentary dialect (Moroccan): *hādi l-bənt dyāli* 'this is my daughter'

Demonstratives in Arabic dialects have other roles too, for example emphasizing a deictic expression, as is the case in Moroccan when a long form demonstrative that agrees in gender and number with the noun is postponed, e.g. *hād əl-bənt hādi* 'this girl over here'. They are also used when forming compound constructions which are adverbial phrases of time, in which the space expressed may be both deictic or figurative, e.g. *hād əl-yūm* 'today' (Moroccan Arabic), *hēḏīkt assāʿah* 'at that moment' (Eastern Bedouin Arabic). Finally, in some dialects they may also form exclamatory expressions, as is the case in the Egyptian dialect by means of *da*. . . . *da*, *(di* . . . *di*, fem.) and (*dūl*. . . . *dūl*, pl.), e.g. *di ʿarabiyya di* 'you call that a car!'.

Finally, it should be pointed out that the demonstratives play a role in sociolinguistic variation, since the use of some of the forms indicates the existence of diastratic variations between the speakers of a certain dialect. This is the case, for example, of the demonstratives formed with the suffix *-ti*, used by the women of Balyāna in South Egypt, (*de/deitei* 'this [masc./fem.]', for women, whereas men use the koineized forms *da/di*, Miller 2003:489). In this case, the use of one demonstrative or another means a diastratic variation related to gender. We can therefore state that the demonstrative system in Arab dialects continues to evolve within a typical general process of koineization (→ dialect koine).

With regard to demonstratives in different Arabic dialects, Fischer (1959) remains the main reference. His general vision must, however, be complemented with more recent data, which can be obtained from the synchronic studies carried out in the various Arabic dialects, and with the information supplied by works dealing with specific aspects of the role and formation of demonstratives (e.g. Woidich 1992; Doss 1979; Behnstedt 1993). Sociolinguistic variation is another source of information on demonstratives in Arabic dialects (e.g. Killean 1980; El-Hassan 1979).

Bibliographical references

Behnstedt, Peter. 1993. "Die demonstrativen Bildungen der syrisch-arabischen Dialekte". *Zeitschrift für Arabische Linguistik* 25.76–94.

Caubet, Dominique. 1992. "Deixis, aspect et modalité: les particules *hā-* et *ṛā-* en arabe marocain". *La deixis*, ed. Mary-Annick Morel and Laurent Danon-Boileau, 139–149. Paris: Presses Universitaires de France.

——. 1993. *L'arabe marocain*, I. *Phonologie et morphosyntaxe*, II. *Syntaxe et catégories grammaticales, textes*, Paris and Louvain: Peeters.

Doss, Madiha. 1979. "The position of the demonstrative *DA, DI* in Egyptian Arabic: A diachronic inquiry". *Annales Islamologiques* 15.349–357.
El-Hassan, Shahir A. 1978. "Variation in the demonstrative system in educated spoken Arabic". *Archivum Linguisticum* 9.32–57.
Fischer, Wolfdietrich. 1959. *Die demonstrativen Bildungen der neuarabischen Dialekte.* The Hague: Mouton.
—— and Otto Jastrow (eds.). 1980. *Handbuch der arabischen Dialekte.* Wiesbaden: O. Harrassowitz.
Heath, Jeffrey. 2002. *Jewish and Muslim dialects of Moroccan Arabic.* London: Routledge Curzon.
Ingham, Bruce. 1982. *North east Arabian dialects.* London and Boston : Kegan Paul International.
——. 1994. *Najdi Arabic: Central Arabian.* Amsterdam and Philadelphia: J. Benjamins.
Jastrow, Otto. 1973. *Daragözü: Eine arabische Mundart der Kozluk-Sason-Gruppe (Südostanatolien)*, Nuremberg: H. Carl.
Jong, Rudolf E. de. 2000. *A grammar of the Bedouin dialects of the northern Sinai littoral.* Leiden: E.J. Brill.
Killean, Carolyn G. 1980. "Demonstrative variation in oral media Arabic in Egypt". *Studies in the Linguistic Sciences* 10:2.165–178.
Marçais, Philippe. 1977. *Esquisse grammaticale de l'arabe maghrébin.* Paris: Librairie d'Amérique et d'Orient.
Miller, Catherine. 2003. "La télévision dans le placard: Parlers féminins à Balyâna (Haute-Égypte)". *Mélanges David Cohen*, ed. Jérôme Lentin and Antoine Lonnet, 481–495. Paris: Maisonneuve and Larose.
Ould Mohamed Baba, Ahmad-Salem. 1999. "Tres cuentos de ʿAbd el-Kärīm: La astucia del pobre". *Estudios de Dialectología Norteafricana y Andalusí* 4.149–156.
——. 2000–2001. "Cuentos de los Wlād Däymān: Sutil socarronería". *Estudios de Dialectología Norteafricana y Andalusí* 5.253–264.
Rosenhouse, Judith. 1984. *The Bedouin Arabic dialects: General problems and a close analysis of North Israel Bedouin dialects.* Wiesbaden: O. Harrassowitz.
Talmoudi, Fathi. 1980. *The Arabic dialect of Sūsa (Tunisia).* Göteborg: Acta Universitatis Gothoburgensia.
Woidich, Manfred. 1992. "Vorangestellte Demonstrativa im Kairenischen". *Jerusalem Studies in Arabic and Islam* 15.195–219.
Zaki, Abdel-Malek. 1972. *The closed-list classes of colloquial Egyptian Arabic.* The Hague and Paris: Mouton.

Ángeles Vicente (University of Zaragoza)

Derivation

1. Introduction

In Indo-European languages, 'derivation' is usually defined as "the formation of a new word or inflectable stem from another word or stem. It typically occurs by the addition of an affix" (<http://www.sil.org/linguistics/Glossary>). In Semitic languages and particularly in (Classical) Arabic, this type of derivation does exist but must be considered marginal. More central is a type of derivation in which a word is not derived from another word, but from a root 'crossed' with a pattern (or 'scheme', from the French *schème*, or 'template').

Where French-speaking scholars, following Cantineau (1950a), use the positive term 'crossing' (*interdigitation*), many English-speaking linguists use the negative term 'non-concatenation', in contrast to the concatenative morphology of the Indo-European languages.

This 'interdigitation' would explain not only the formation of any new word but also its interpretation, because root and pattern are both considered to be what are called 'signs' in the Saussurean tradition, and 'morphemes' in the American linguistic tradition (for recent overviews, see Goldenberg 1994; Shimron 2003).

The radical opposition between the two types of derivation (e.g. Fleisch 1961:247–251) has as a counterpart the possible analogy between them, in the sense that the root is to the stem exactly what the pattern is to the affixes, i.e. a lexical morpheme (or lexeme) to a grammatical morpheme (or, for the sake of brevity, a morpheme). Lexeme and morpheme correspond to what the German linguistic tradition calls *Bedeutungslaute* and *Beziehungslaute*.

Arab grammarians do not derive words belonging to the same lexical family from a root but from a base. This base is generally the *maṣdar* (*nomen actionis*, lit. 'source') of the groundform of the triliteral verb or, even more generally, an *ism jāmid* 'underived noun', since only this *maṣdar*, i.e. the *maṣdar* of the groundform of the triliteral verb, is considered an *ism jāmid*. Ġalāyinī (*Jāmiʿ* I, 214ff., II, 3ff.) points out that the discipline called *ʿilm at-taṣrīf* (→ *ṣarf*) consists of two parts, the second part being dedicated to phonology, the first part to morphology, more exactly to derivational morphology. It is the latter which is called *ištiqāq ṣaġīr* or simply, → *ištiqāq* (see below).

Until the 19th century, Arabists, too, used to derive words not from the root but from a base, unfortunately called → 'root', as clearly appears from the following quotation from the Arabic grammar of Antoine-Isaac Silvestre de Sacy (1758–1838): "Cette racine, en arabe *aṣl*, est toujours la troisième personne du singulier masculin du prétérit de la voix active" (1831:I, 123). The root here is evidently not *k-t-b*, but *kataba* 'he wrote'. This quotation from Silvestre de Sacy easily explains the misunderstanding between Arab grammarians and Western specialists of Arabic grammar as a purely terminological one. In Silvestre de Sacy's usage, the term *racine* is a French translation from the Latin *radix*, translating the Arabic *ʾaṣl*, possibly *via* the Hebrew *šoreš* (see Troupeau 1984).

To avoid this ambiguity, we may simply call *k-t-b* a consonantal root and *kataba* a vocalized root. The terminological misunderstanding was merely the result of the paradigm shift which occurred at the time. The newly-born comparative linguistics of the Indo-European languages introduced the concept of root, but the Indo-European root was described as monosyllabic while the Semitic root was described as purely consonantal (see Rousseau 1980, 1984, 1987). Western Semiticists and Arabists could then re-establish the concept of the root as consonantal, since it had always played a crucial role in Arabic traditional → lexicography as an instrument of classification, even though it was not involved in the derivation of words.

From the early 8th century, Arabic lexicons have been organized on the basis of the consonantal root of the words. This principle of classification implies the extraction of the root and thus the analysis of the word into two elements, a root (*mādda* 'matter' or *jawhar* 'substance') and a pattern (*ṣīġa* 'form'). This does not imply that the word is formed through the synthetic crossing of root and pattern, and consequently, the entry of the lexicon should not be regarded as the basis of the derivational process.

In short, it is not immediately clear whether 'derivation' in Arabic means to derive words from a consonantal root, as most Arabists believe, or to derive them from a concrete base, i.e., from another word or stem, as Arab grammarians and Arabists (until the 19th century) believed, as well as many linguists after them (e.g. Brockelmann 1908).

2. DERIVATION AND *IŠTIQĀQ*

The *maṣdar* being the base of derivation, all words belonging to the same lexical family can be called its derivatives (*muštaqqāt*). Nevertheless, grammarians do not derive all of these directly from the *maṣdar*, but indirectly, by successively deriving them one from the other. Thus, when writing about the derivation of the verb, Ġalāyīnī tells us that the imperative *uktub* 'write!' is derived from the imperfect *yaktub-* 'he writes, he will write', the imperfect *yaktub-* from the perfect *kataba* and the perfect *kataba* from the *maṣdar kitāba* 'writing' (*Jāmiʿ* I, 215). As for verb-related nouns, most of them are derived from the imperfect, especially active and passive participles, *nomen locis et temporis*, *nomen instrumenti*, *maṣdar mīmī*. Thus, for Arab grammarians the *maṣdar* is a semantic base rather than a morphological one, chosen as the source of the derivation because it was the unmarked semantic, not morphological, form (e.g. Ibn Masʿūd; cf. Åkesson 2001). However, it would be possible to unify semantic and morphological bases by choosing the *maṣdar katb* as the source of the derivational chain. The *katab-* ⇒ *yaktub-* derivation could appear somewhat surprising since the ground-form of the triliteral verb has at least two stems, namely the perfect stem and the imperfect stem. Arab grammarians take into consideration not only the ground-form, but also the augmented forms in which the stem is the same, abstracting away from the → apophony (or ablaut): *qattal/yu-qattil*, *qātal/yu-qātil*, *ta-qattal/ya-ta-qattal*, *ta-qātal/ya-ta-qātal*, etc. Hence, the perfect could be regarded as the basic form of the verb.

The central role of the imperfect in derivational morphology is fully respected, however. The notion of *ištiqāq* transcends the traditional distinction between derivation and → inflection. The imperfect prefixes (*ḥurūf al-muḍāraʿa*) are the 'person' features and consequently belong to the conjugation, and hence to the inflection. This shows that derivation and inflection share a morphological operation, the addition of one or more affixes to a stem. It suffices to distinguish between two classes of affixes: the derivational affixes and the inflectional ones. From this point of view, there is no difference between Arabic *ya-ta-ʿallam-u* 'he learns' and English *re-en-larg-ed*, since both juxtapose on the same line inflectional (*y-* and *-u* in Arabic and *-ed* in English) and derivational affixes (*t-* and gemination in Arabic and *re-* and *en-* in English). However, Arabic adds another type of morpheme, well-known in English and some other Indo-European languages, namely apophony (or ablaut). Note that apophony is sometimes included with → gemination in the category of 'internal inflection' by some authors, who refuse to consider gemination as a derivational affix.

In school grammar, *ištiqāq* is used as a synonym of *ištiqāq ṣaġīr*, but scholars regard *ištiqāq ṣaġīr* 'small etymology' as only one of the three levels covered by the term. The two other levels are called *kabīr* 'great' and *ʾakbar* 'greater'. These three levels of derivation were directly or indirectly introduced by Ibn Jinnī (d. 392/1002) in his *Xaṣāʾiṣ* (II, 133–139, 145–152). In the *kabīr* type, the order (*tartīb*) of the radical consonants is no longer required. In the *ʾakbar* type, the phonetic proximity of the point of articulation of one of the three radical consonants suffices (see Versteegh 1986). This theory, although marginal in the Arabic linguistic tradition, is of historical importance because it is the ancestor and the precursor of all theories of the Arabic and Semitic root that have appeared in the last 150 years. These theories have to be considered for what they are, i.e. root-generative models or root-computing models. Nevertheless, because of unsolved methodological problems and 'rough' semantics, these theories bear no direct relation to what it is usually called 'derivational theory', i.e. theories of word-formation and interpretation at the same stage of the same variety of one language.

The fact that *ištiqāq* is translated not only as 'derivation', but also as 'etymology', reminds us that *ištiqāq*, including *ištiqāq ṣaġīr*, is related to what the Western tradition calls → 'etymology'. These terms are not equivalent in the technical sense of etymology, which is diachronic, but in its historical one, i.e. the search for 'the true sense' (Greek *to étumon*) of a word, e.g. *al-jinn* 'jinnis' is 'derived' from *al-ijtinān* 'the fact of being hidden'. It is the consonantal root which functions as a vector for such fallacious etymologies and similar puns, which some researchers invoke as an argument for the 'consciousness' of the root and, in the case of folk etymology, of its semantic unity. Obviously, such unity can always be found, even if it has to be invented.

Arabic *ištiqāq* raises the same questions as 'derivation' in Arabist terminology, i.e. whether it pertains to morphology or semantics, to synchrony or diachrony. D. Cohen has repeatedly and very explicitly defined derivation as involving the crossing of a root with a pattern. He argues that the 'functional reality' and 'linguistic existence' of the consonantal root as the morphological base of derivation is demonstrated by the case of loanwords, *nikl* 'nickel', for instance, being the origin of the verb *nakkala* 'to nickel'. He adds (1964:74) that "etymologically, *nikl* is the mother-form of *nakkala* . . .; but the process implies passing through the root *nikl* > NKL ⇒ *nakkala*". Cohen gets rid of the source-word turning it into the *étumon*, according to Goldenberg's correct expression, who followed Cohen in this (Goldenberg 1994).

The approach of these linguists is reminiscent of the 19th-century practice, when historical linguists, essentially German (e.g. Hermann Paul), stated: "We will not talk about derivation if we cannot prove that a word comes from another one, that '*maisonnette*' comes from '*maison*', and proving this requires that the source-word ('*maison*') pre-exists the derived word ('*maisonnette*')" (Ducrot and Schaeffer 1995:26). But historical linguistics was followed by Saussurean linguistics, for which: "if there is a '*travailler-travailleur*' ('to work-worker') derivation, it fits into a '*manger-mangeur*' ('to eat-eater'), '*lutter-lutteur*' ('to fight-fighter') etc. series, where the verb, in each pair, is an action-verb" (Ducrot and Schaeffer 1995:283).

If *nickel* diachronically precedes *to nickel* (and in Arabic *nikl* precedes *nakkala*), they also co-exist synchronically. Synchronically, it is the verb that is understood as including the semantics of the noun, not the reverse, since the pair *nickel/to nickel* or *nikl/nakkala* fits into a series of denominative action-verbs (English 'to N', French *N-er*, Arabic *faˁˁala*), whose general paraphrase is 'to do what we do with the object referred to by N'. If we refuse to articulate the relations *nikl/nakkala* and NKL/*nakkala* along the diachrony/synchrony distinction, we can then only articulate them along the semantics/morphology distinction.

If *nakkala* is formed morphologically from the root *n-k-l*, the latter is semantically the trace of the nominal base *nikl* within the derived (and denominative) verb. Word-formation may be the crossing of the root with the pattern, but there is no doubt that word-interpretation must relate that root to the source-word. Loanwords (and more generally denominatives), always presented as an *a fortiori* morphological argument that roots are combined with patterns, could thus be viewed as a semantic counter-argument to the same thesis.

This example shows the contradiction existing between the concept of derivation itself and the thesis that Semitic morphology involves the crossing of a root with a pattern, both of them being bound morphemes, and neither of them occurring independently from the other. Derivation presupposes a base. The root/pattern mode of word-formation turns each word into a small system, unrelated to the other words of the same lexical family, except via their common root.

The root/pattern principle is typically transcendental and substantialist. It can neither explain the horizontal relationship among words (e.g., that *makātib* is the plural of *maktab*) nor can it describe the semantic difference between *maktab* and *maktaba-t*. The meanings of these two words, having the same root and the same pattern, should differ only in what results from the addition of the affix *-at*. But everyone knows that *maktaba-t* ‘a library’ is linked to the noun *kitāb* ‘a book’/ *kutub* ‘books’, which implies that *maktab* ‘scriptorium’ is linked to the verb *katab-/yaktub-* ‘to write’ (Larcher 1995). The question then becomes whether or not in the derivational process the derived word is root-based in a morphological sense, since it is never based on the ‘bare’ root semantically.

3. Non root-based word-formation

Affixation to a base (word or stem) plays a relatively small part in the derivation of nouns. Cohen (1964) recognizes only suffixes as morphemes: *-iyy* (e.g. *ṭabīʿa* ‘nature’ ⇒ *ṭabīʿiyy* ‘natural’; the addition of this suffix may cause the deformation of the stem, for instance *madīna* ‘city’ ⇒ *madaniyy* ‘civil’); *-at-*, *-ān*, augmentative (e.g. *sulṭa* ‘power’ ⇒ *sulṭān* ‘sultan’), or adjectival suffix (e.g. *ʿaṭš* ‘thirst’ ⇒ *ʿaṭšān* ‘thirsty [masc.])’, to which *-ā* may be added (*ʿaṭšā* ‘thirsty [fem.]’). Cohen denies the status of morpheme to prefixes such as *m-*, except for active and passive participles, which he relates to verb derivation (for a further discussion of *m-* as a morpheme, see below).

Affixation plays, in fact, a big part in the derivation of verbs, since the so-called derived verb classes are primarily derived from the ground-form, by prefixation and/or infixation. But all these forms can also be secondarily denominative. In this case, the nominal base is generally represented by the root alone and, as a consequence, the grammatical meaning is disconnected from the affix and attributable to the pattern as a whole, so that the analysis/synthesis of the word into two elements (root and pattern) is pushed back to word-formation from a verb itself. This illustrates the paradoxical character of the principle of root/pattern. It is morphologically better suited to word-formation from nouns than from verbs. But, since this principle is clearly semantically inadequate for word-formation from nouns, word-formation from verbs is considered more representative of this principle at both the morphological and semantic level. What Arabists call the ‘meaning of the root’ is generally that of the ground-form of the triliteral verb, for instance *k-t-b* = ‘writing’. As a result of this theory, Western grammars of Classical Arabic usually present a list of forms on the one hand and a list of meanings on the other: no links are proposed between the forms and the meanings.

Some phenomena can be treated under the generic category of ‘formal’ derivation. This terminology is taken from Joüon (1935), but he himself used it for only one of these phenomena (see below). The most productive of these is surely that in which one form is a function of another. This adequately describes the formation of → diminutives and of plural forms, i.e. the processes of *taṣġīr* and *jamʿ* of the Arab grammarians (e.g. *kalb/kulayb*, *šāʿir/šuwayʿir*, *kitāb/kutayyib*, *ʿaqrab/ʿuqayrib*, *ʿuṣfūr/ʿuṣayfir* ‘a (small) dog, poet, book, scorpion, sparrow’; *maktab/makātib* ‘scriptorium/ scriptoria’, *daftar/dafātir* ‘a register/registers’, *ʾakbar/ʾakābir* ‘greater/greater ones’, *muškila/mašākil* ‘a difficulty/difficulties’. This derivational process led some Arabists to conclude that the root/pattern model was inadequate and to search for alternative models, see, e.g., Ratcliffe’s (1997, 1998) work on → ‘broken’ plurals. Joüon (1935, cf. Nöldeke 1897) specifically called such derivations as *bāl* ‘mind’ ⇒ *bāla* ‘to mind’ ‘formal’, where the form of the base conditions that of the derivative, hence the pattern of the derivative has no semantic value but only a semiotic one. This phenomenon concerns not only Form III, but also other forms, especially Form IX, which is always denominative and always linked to a noun *ʾafʿal* (*ʾafʿal al-lawn* and *ʾafʿal al-ʿilla* for colors and defects or deformities), for instance *ʾaswad* ‘black’ ⇒ *iswadda* ‘to be or become black’ and *ʾaʿwar* ‘blind in one eye’ ⇒ *iʿwarra* ‘to be or become blind in one eye’.

Many processes may be called ‘formal derivation’, from simple deformation of a word formally attracted to another by coupling, to true word-formation. The first case is what Barth (1906) called *Formangleichung* ‘formal adjustment’, for instance *bidāya* ‘beginning’, linked to the verb *bada*ʾ*a* ‘to begin’ *vs.* *nihāya* ‘end’, or *zamān* ‘time’ besides *zaman* because of the existence of the pair *az-zamān wa-l-makān* ‘time and space’ (Joüon 1913).

The second case involves such words as *qasāma* ‘group of people who swear, a collective oath’. Arab lexicographers explained its formation by the formal attraction either to *jamāʿa* ‘group’ or to *ḥamāla*/*ġarāma* ‘thing that must be paid’: *qasāma* implies bloodwit (see Ibn Manẓūr [d. 711/ 1311], *Lisān al-ʿarab*, s.v. *q-s-m*). One could similarly explain the formation of *ḏ̣āhara* ‘to say [to one’s wife] “thou art to me like the back of my mother”’ by the formal attraction of *fāraqa* ‘he separated himself from his wife’ and the formation of *istarjaʿa* ‘to say: *ʾinnā li-llāh wa-ʾinnā ʾilayhi rājiʿūn*’ ‘verily, to God we belong and verily, unto Him we return’, which has the value of a prayer, by the formal attraction of other verbs involving prayers such as *istaġfara* ‘to ask for forgiveness’ and *istasqā* ‘to pray for rain’.

Finally, under the same category of formal derivation may also be grouped the formation of words starting from a morphologically equivocal base. For instance, *ʾišāra*, pl. *ʾišārāt* (*maṣdar* of the hollow verb of Form IV *ʾašāra* ‘to indicate’), reinterpreted as a form *fiʿāla*, is the plausible origin of the strong verb of Form II *ʾaššara* of which two deverbal nouns are used, namely *taʾšīr(a)* ‘visa’ and *muʾašširāt* ‘parameters’. This kind of word-formation shows that native speakers of Arabic do not necessarily analyze a word the way Arabists would.

4. ROOT-BASED WORD FORMATION: AN AMENDED THEORY

Arabists sometimes call ‘secondary’ (Mez 1906; Colin 1961–1963) some kinds of word-formation discussed above, through which new roots can eventually come into existence, for instance *ʾišāra* being a bridge between the root *š-w-r* and the root *ʾ-š-r*. These word-formation processes reinstate the root in its proper place, which is not primary, but secondary. A root exists only if its consonants appear in at least two words (e.g. *nikl*/*nakkala*). An already existing root can acquire another meaning when a new association appears, for instance *k-t-b*, after having associated the original *maṣdar kitāb* with the plural *kutub*, changes its meaning from active ‘writing’ to resultative ‘a writing/writings’ before getting the restrictive meaning of ‘book’, to which *maktaba* ‘place for books’ is related.

There is no contradiction in the fact that a root, even if it may be the morphological base, is always the product of derivation. Just as a root comes into existence through the association of at least two words (a noun and a verb derived from this noun; a noun and its plural, etc.), in the same way patterns do not pre-exist as autonomous entities. Instead they progressively gain their autonomy.

The so-called *nomen loci* ‘noun of place’, for instance, occurs in four variants: *mafʿa(i)l(a)*. Both Arab grammarians and Arabists focus on the correlation existing between this nominal form and the imperfect of the verb. What we actually have is *mafʿil(a*), if the imperfect is *yafʿil* (e.g. *yajlis* ‘he sits down’/*majlis* ‘sitting place’; *yanzil* ‘he goes down’/ *manzil* ‘place’ and *manzila* ‘position’). Since the same form is not exclusively a *nomen loci*, but could also be a *maṣdar mīmī* (e.g. *maʿrifa* ‘knowledge’), it could be analyzed as *m* + the stem of the imperfect. This is less true for *mafʿal(a)* which goes with *yafʿal* and *yafʿul*. There is no longer a link with the vowel of the imperfect, but there is still a link with the syllabic structure of its stem. It is no surprise that *mafʿal(a)* was chosen by Arab grammarians as the denominative *nomen loci* (*ism al-kaṯra*, *nomen abundantiae* ‘noun of abundance’).

The nominal base of *mafʿal(a)* is only represented by the root, hence the grammatical meaning is conveyed by the entire pattern but, even in this case, the word is not only the result of the crossing of the root with the pattern at the morphological level. The form itself denotes its denominative origin. For example, *maṭār* ‘airport’ is not a *nomen loci*, but a *nomen abundantiae*, i.e. it is denominative, instead of deverbal. If it had been derived from the verb *ṭāra*/*yaṭīru* ‘to fly’ it would have the form **maṭīr*, just as *bāta*/*yabītu* ‘to pass, or to enter upon, the night’ gives *mabīt* ‘a place in which one passes, or enters upon, the night’. It is actually linked to the noun *ṭāʾira*/*ṭayyāra* ‘an airplane’ pl. *-āt* ‘airplanes’ and it means ‘a place abounding in airplanes’ (see *Lisān al-ʿarab*, s.v. *ṭ-y-r*: *ʾarḍ maṭāra kaṯīrat aṭ-ṭayr* ‘ground abounding in birds’).

The process of pattern derivation is in itself a dynamic phenomenon. It can be interpreted in a historical perspective (e.g. Barth 1894) but may also be observed in the synchrony of Classical Arabic and Modern Standard Arabic. For example, the so-called 'noun of profession' *faʿʿāl* is primarily an intensive form of the active participle (*ʾākil* 'eating'/*ʾakkāl* 'glutton'). If the action verb denotes a professional activity, *faʿʿāl* becomes a noun of profession and such a change can occur at any time: when *ṭāra* 'to fly' becomes 'to fly by plane', *ṭayyār* becomes 'a pilot'. But the form becomes definitively a noun of profession only when it is derived from a nominal base, e.g. *ṯalj* 'snow' ⇒ *ṯallāj* 'snow-trader', today 'ice-cream seller'.

5. CONCLUSION

To assert, following Cantineau (1950b), that the two great crossed orders of roots and patterns account for the entire Semitic lexicon, is a holistic view, which, like all such views, fails to reach simple descriptive adequacy. At the morphological level, word-formation may be based on a root, as well as on a word or a stem. At the semantic level, even when the word is morphologically based on the root, this root never has a meaning of its own but always acquires a meaning as a trace of either a verb or a noun, not necessarily the simplest form of the verb or the noun. In the case of the verb, it may well not be the active but the passive of the ground-form (e.g. *ʾawjada* 'to bring into existence', linked to *wujida* 'to exist', not to *wajada* 'to find') or an augmented form (e.g. *ġallaqa* 'to bolt [door after door]' and *inġalaqa* 'to be bolted', linked to *ʾaġlaqa* 'to bolt') (Larcher 1999).

If at the semantic level, there is no derivation except from word to word and at the morphological level, the derived word may be based either on the source-word (or its stem) or its root, there is no need to raise the question of word derivation and formation in Classical Arabic in terms of a conflict between the two systems ('word to word' or 'root/pattern'), but in terms of a concurrence of the two systems. Hence, the linguist has to understand how one of the two systems can be 'derived' from the other, not only at the synchronic level (since they occur simultaneously), but also in the *continuum* (since there are some mixed forms belonging to both systems).

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Ġalāyinī, *Jāmiʿ* = Muṣṭafā al-Ġalāyinī, *Jāmiʿ ad-durūs al-ʿarabiyya*. 3 vols. 10th ed. Beirut and Sidon: al-Maktaba al-ʿAṣriyya li-ṭ-Ṭibāʿa wa-n-Našr, 1385/1966.

Ibn Jinnī, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. Ed. Muḥammad ʿAlī an-Najjār. 3 vols. Beirut: Dār al-Hudā li-ṭ-Ṭibāʿa wa-n-Našr, n.d.

Ibn Manẓūr, *Lisān* = Jamāl ad-Dīn ʾAbū l-Faḍl Muḥam-mad ibn Mukarram ibn ʿAlī al-ʾAnṣārī al-ʾIfrīqī al-Miṣrī Ibn Manẓūr, *Lisān al-ʿArab al-muḥīṭ*. Ed. Yūsuf Xayyāṭ. 4 vols. Beirut: Dār Lisān al-ʿArab, n.d.

SECONDARY SOURCES

Åkesson, Joyce (2001). *Arabic morphology and phonology based on the* Marāḥ al-ʾarwāḥ *by Aḥmad b. ʿAlī b. Masʿūd presented with an introduction, Arabic edition, English translation and commentary*. Leiden: E.J. Brill.

Barth, Jakob. 1894. *Die Nominalbildung in den semitischen Sprachen*. 2nd ed., Leipzig. (Repr. Hildesheim: G. Olms, 1967.)

——. 1906. "Formangleichung bei begrifflichen Korrespondenzen". Bezold (1906:II, 787–796).

Bezold, Carl (ed). 1906. *Orientalische Studien Theodor Nöldeke zum siebzigsten Geburtstag (2. März 1906) gewidmet von Freunden und Schülern*. 2 vols. Giessen: Tölpelmann.

Brockelmann, Carl. 1908–1913 [1999]. *Grundriss der vergleichenden Grammatik der semitischen Sprachen*. I. *Laut- und Formenlehre*; II. *Syntax*. Berlin: Reuther und Reichard. (Repr. Hildesheim: G. Olms, 1999.)

Cantineau, Jean. 1950a. "Racines et schèmes". *Mélanges William Marçais*, 119–124. Paris: G.P. Maisonneuve.

——. 1950b. "La notion de 'schème' et son altération dans diverses langues sémitiques". *Semitica* 3.73–83.

Cohen, David. 1964. "Remarques sur la dérivation nominale par affixes dans quelques langues sémitiques". *Semitica* 14.73–93. (Repr. David Cohen, *Etudes de linguistique sémitique et arabe*, 31–48. The Hague and Paris: Mouton, 1970.)

Colin, Georges. 1961–1963. "Singuliers secondaires analogiques tirés de faux pluriels en arabe". *Comptes rendus du Groupe Linguistique d'Etudes Chamito-sémitiques* 9.11–15.

Ducrot, Oswald and Jean-Marie Schaeffer. 1995. *Nouveau dictionnaire encyclopédique des sciences du langage*. Paris: Le Seuil.

Fleisch, Henri. 1961, 1979. *Traité de philologie arabe*. I. *Préliminaires, phonétique, morphologie nominale*; II. *Pronoms, morphologie verbale, particules*. Beirut: Imprimerie Catholique.

Goldenberg, Gideon. 1994. "Principles of Semitic word-structure". *Semitic and Cushitic studies*, ed. Gideon Goldenberg and Shlomo Raz, 29–64. Wiesbaden: O. Harrassowitz.

Joüon, Paul. 1913. "Arabica". *Mélanges de l'Université Saint-Joseph* 6.147–159.

——. 1935. "Remarques sur les 3ème et 7ème formes verbales *fāʿala* et *infaʿala* de l'arabe". *Mélanges de l'Université Saint-Joseph* 19.99–116.

Larcher, Pierre. 1995. “Où il est montré qu’en arabe classique la racine n’a pas de sens et qu’il n’y a pas de sens à dériver d’elle”. *Arabica* 41.291–314.

——. 1999. “Vues ‘nouvelles’ sur la dérivation lexicale en arabe classique”. *Tradition and innovation: Norm and deviation in Arabic and Semitic linguistics*, ed. Lutz Edzard and Mohammad Nekroumi, 103–123. Wiesbaden: O. Harrassowitz.

Mez, A. 1906. “Über einige sekundäre Verba im Arabischen”. Bezold (1906:I, 249–254).

Nöldeke, Theodor. 1897. *Zur Grammatik des klassischen Arabisch*. Wien: Carl Gerold’s Sohn.

Ratcliffe, Robert. 1997. “Prosodic templates in a word-based morphological analysis of Arabic”. *Perspectives on Arabic linguistics*, X, ed. Mushira Eid and Robert R. Ratcliffe, 93–122. Amsterdam and Philadelphia: J. Benjamins.

——. 1998. *The “broken” plural problem in Arabic and comparative Semitic: Allomorphy and analogy in non-concatenative morphology*. Amsterdam and Philadelphia: J. Benjamins.

Rousseau, Jean. 1980. “Franz Bopp et la pratique de la grammaire arabe”. *Histoire, Epistémologie, Langage* 2.52–66.

——. 1984. “La racine arabe et son traitement par les grammairiens européens (1505–1831)”. *Bulletin de la Société de Linguistique de Paris* 89.285–321.

——. 1987. “La découverte de la racine trilitère en sémitique par l’idéologue Volney”. *Historiographia Linguistica* 14.341–365.

Shimron, Joseph. 2003. “Semitic languages: Are they really root-based?”. *Language processing and acquisition in languages of Semitic, root-based morphology*, ed. Joseph Shimron, 1–28. Amsterdam and Philadelphia: J. Benjamins.

Silvestre de Sacy, Antoine-Isaac. 1831. *Grammaire arabe à l’usage des élèves de l’Ecole spéciale des langues orientales vivantes*. 2nd ed. 2 vols. Paris: Imprimerie Royale. (Repr. Paris: Institut du Monde Arabe, 1986.)

Troupeau, Gérard. 1984. “La notion de ‘racine’ chez les grammairiens arabes anciens”. *Matériaux pour une histoire des théories linguistiques*, ed. Sylvain Auroux, Michel Glatigny, André Joly, Anne Nicolas, and Irène Rosier, 299–245. Lille: Presses Universitaires.

Versteegh, Kees. 1986. “La ‘grande étymologie’ de Ibn Jinni”. *La linguistique fantastique*, ed. Sylvain Auroux, Jean-Claude Chevalier, Nicole Jacques-Chaquin, and Christiane Marchello-Nizia, 44–50. Paris: Clims and Denoël.

Pierre Larcher
(University of the Provence)

Determiners

Determiners, which are a class of noun modifiers used to express or identify the reference of a noun, include in Arabic articles, possessive pronouns, and demonstratives. Quantifiers, the class of noun modifiers used to specify the quantity of a noun, are sometimes considered as part of the class of determiners, inasmuch as they restrict the reference of a noun to a specific or indefinite quantity. The distribution and syntax of quantifiers differs from that of the other determiners and can therefore be discussed separately. The identifiability of a noun’s reference corresponds to the semantic notion of (in-)definiteness, which is expressed by the use of articles. However, the presence of an article in a noun phrase is not always an indication of semantic (in-)definiteness. This is for instance the case with generic nouns, which take the definite article in Arabic, as in (1).

(1) *ʾarā* *ʾanna* *l-ʾasada* *ḥayawān-u-n*
think.1s that the-lion.acc animal.nom
jamīl-u-n
beautiful.nom
‘I think that the lion is a beautiful animal’

Thus, it could not be said that in (1), the noun phrase *al-ʾasada* ‘the lion’ has an identifiable referent, since it refers to the species rather than to a token animal.

Specificity, a notion related to semantic definiteness, but not necessarily syntactic definiteness, is not grammaticalized in Arabic. That is, the interpretation of a noun’s reference as known by the speaker is not associated with a specific expression in Arabic. Thus, definite noun phrases with an identifiable referent are naturally specific, as illustrated in (2),

(2) *jāʾa* *l-muʿallim-u*
came.3sm the-teacher.ms.nom
‘The teacher came’

where the definite noun phrase *al-muʿallim-u* ‘the teacher’ is also specific. Indefinite noun phrases can also be specific, and that is an interpretation that is available for the indefinite noun phrase *rajul-un* ‘a man’ in (3).

(3) *jāʾa* *rajul-u-n* *li-yabīʿa-nā*
came.3sm man.nom to-sell.3sm-us
kutub-a *t-tārīxi*
books.acc the-history
‘A man came to sell us the history books’

The morphological absence of an article preceding the noun is generally taken to be a marker of semantic indefiniteness in Arabic. Indefinite nouns are also affixed with the → nunation morpheme *-n*, a suffix that has a wide distribution in

Modern Standard Arabic, as well as several functions, and that occurs at the end of nouns and adjectives. The status of the nunation morpheme (or → *tanwīn*, as Arab grammarians refer to it) is still a controversial issue in the context of the discussion on the syntax of noun phrases (Fassi Fehri 1993). Two main positions can be distinguished regarding this element: one which characterizes nunation as an indefinite enclitic article, the counterpart of *al-* 'the', and another which argues that it is an inflectional suffix marking case. The first position is based mainly on the observation that (a) nunation generally appears on indefinite nouns and (b) in those cases it is in complementary distribution with the definite article *al-* 'the' (Kouloughli 2001). Table 1 illustrates the complementarity between the definite article and nunation for the three case forms that are generally available for nouns. In addition, the translations indicate that the nouns occurring with the nunation morpheme are interpreted with indefinite reference.

Table 1. Distribution of nunation

	definite noun 'the man'	indefinite noun 'a man'	
nom.	*ar-rajul-u*	*rajul-u-n*	**ar-rajul-u-n*
acc.	*ar-rajul-a*	*rajul-a-n*	**ar-rajul-a-n*
gen.	*ar-rajul-i*	*rajul-i-n*	**ar-rajul-i-n*

Despite these observations, there is no one-to-one correspondence between the presence of nunation and the (semantic) indefiniteness of nouns. Specifically, some indefinite nouns cannot co-occur with the nunation enclitic morpheme, as the case of diptotes shows (→ diptosis). The case system in Arabic consists of three forms, namely nominative, generally marked by the short vowel /-u/; accusative, generally marked by the short vowel /-a/; and genitive, generally marked by the short vowel /-i/. Most Arabic nouns belong to this three-way inflectional system and are called triptotes. Diptotes are nouns that have only one form for the genitive and the accusative and therefore the three-way distinction in their case system is reduced to a two-way distinction. For common nouns that fall into the category of diptotes, the genitive form merges with the accusative form only when the nouns in question are indefinite. When those nouns occur with the definite article, they become triptotes. The category of diptotes is exemplified by singular nouns like *ṣaḥrāʾ* 'desert' and quadriliteral plural nouns like *mafātīḥ* 'keys'. An important observation about the distribution of nunation is that the morpheme cannot occur on diptotes, as illustrated in the contrast in (4).

(4) a. *min ṣaḥrāʾ-a šāsiᶜat-i-n*
from desert.fs.gen wide.fs.gen
'from a wide desert'
b. **min ṣaḥrāʾ-a-n/-i-n šāsiᶜat-i-n*
from desert.fs.gen wide.fs.gen

This observation leads to the conclusion that nunation need not be present for the indefinite reading of a noun phrase to obtain, and thus nunation is only compatible with indefinite readings of noun phrases.

In fact, there are cases where nunation can be shown to be clearly compatible with the definite reading of a noun phrase. This fact relates to the possibility of nunation occurring on proper names, like *zayd-u-n* or *muḥammad-u-n*. Proper names being semantically definite, it would be contradictory to say that nunation there is a marker of indefiniteness. Therefore, the fact that some indefinite nouns need not occur with the nunation morpheme and that nunation may be compatible with a definite meaning led some grammarians to take the position that nunation is an inflectional suffix that does not mark indefiniteness, but case morphology (Fleisch 1961). It marks the noun to which it attaches as a member of the class of triptotes. Interestingly, most modern Arabic vernaculars have lost the distinction between triptotes and diptotes along with morphological case marking. In parallel, those vernaculars have generally lost the nunation morpheme on indefinite nouns, except in a few (mainly frozen) expressions, e.g. *yawm-a-n ma* 'some day' in Lebanese Arabic. However, nunation still marks adverbial uses of nouns and adjectives, like *ʾabad-an* 'never', *deym-an* 'always', and *ʾahl-an wa-sahl-an* 'welcome', also in Lebanese Arabic. Thus, in the modern Arabic dialects, indefiniteness is associated with the absence of any morphological marking.

Definiteness in Arabic is morphologically marked by the presence of the definite → article, a possessive pronoun, or a → demonstrative.

The definite article *al-* 'the' in Arabic is a proclitic morpheme that immediately attaches to the noun. It is in complementary distribution with possessive pronouns, but must co-occur with demonstratives, as illustrated in (5).

(5) *hāḏā* *l-walad-u* *marīḍ-u-n*
this.ms the-child.ms-nom sick.ms-nom
'This child is sick'

Table 2 shows the distribution of the definite article and the possessive pronoun with respect to the noun and case markers.

Possessive pronouns are morphologically identical to the accusative pronouns, which are also suffixes. Table 3 lists all the possessive/accusative pronouns.

Table 2. Distribution of definite article and possessive pronouns

	definite noun 'the book'	possessive noun 'her book'	
nom.	*al-kitāb-u*	*kitāb-u-hā*	**al-kitāb-u-hā*
acc.	*al-kitāb-a*	*kitāb-a-hā*	**al-kitāb-a-hā*
gen.	*al-kitāb-i*	*kitāb-i-hā*	**al-kitāb-i-hā*

Table 3. Possessive and accusative pronouns

	singular		dual	plural
1st person	*-ī* [poss.]/ *-nī* [acc.]		*-nā*	*-nā*
2nd person	masc.	fem.		
	-ka	*-ki*	*-kum*	*-kum*
3rd person	*-hu*	*-hā*	*-humā*	*-hum*

The only notable difference is in the 1st pers. sg. forms, where the possessive pronoun *-ī* has an accusative counterpart *-nī*. Some grammarians analyze the *-n* morpheme here as a semantically vacuous suffix that breaks down a hiatus, as illustrated in (6) (Fleisch 1961:22–23).

(6) *ḍaraba-n-ī*
hit.3sm-n-I
'He hit me'

Possessive pronouns are in complementary distribution with lexical noun phrases. Thus, (7a) alternates with (7b) and (7c). However it cannot alternate with (8a) or (8b).

(7) a. *kitābu-hā*
book.ms.nom-her
'her book'
b. *kitābu* *zayna*
book.ms.nom Zeina
'Zeina's book'
c. *kitābu* *t-tilmīḏat-i*
book.ms.nom the-student.fs-gen
'The student's book'

(8) a. **kitābu-hā zayna*
b. **kitābu-hā t-tilmīḏat-i*

Those observations indicate (a) that the possessive pronouns are indeed to be classified as pronouns occupying the position of the possessor in a noun phrase, and (b) that the syntax of possessive noun phrases is similar to that of construct state nominals.

→ Demonstratives in Classical and Modern Standard Arabic are independent words that generally precede definite noun phrases, as illustrated in (9).

(9) *hāḏā* *l-kitāb-u* *jayyid-u-n*
this.ms the-book.ms-nom good.ms-nom
'This book is good'

Classical Arabic has two sets of demonstratives that serve to identify proximate objects and distant objects. Table 4 lists those elements.

Table 4. Demonstratives in Classical Arabic

	singular		dual		plural
	masc.	fem.	masc.	fem.	
proximate	*hāḏ-ā*	*hāḏ-ihi*	*hāḏ-āni* [nom.] *hāḏ-ayni* [acc./gen.]	*hātāni* [nom.] *hātayni* [acc./gen.]	*hāʾulāʾi*
non-proximate	*ḏā-ka*	*tīka*	*ḏ-ānika* [nom.] *ḏ-aynika* [acc./gen.]	*tānika* [nom.] *taynika* [acc./gen.]	*ʾulāʾika*
	ḏ-ālika	*tilka*			

It is worth noting that Classical Arabic demonstratives have a dual form, which inflects for case. As such, it can be shown clearly that demonstratives, like adjectives, agree with their corresponding nouns in gender, number, and case (10).

(10) a. *qaraʾtu* *hāḏayni*
read.1s this.dual.acc
l-kitābayni
the-book.dual.acc
'I read these two books'

b. *hāḏāni* *l-kitābāni*
this.dual.nom the-book.dual.nom
jayyidāni
good.dual.nom
'These two books are good'

An older form of the language had a proclitic demonstrative *al-*, which can still be seen in some adverbial uses of nouns like *al-ʾān* 'now', *al-yawm* 'today', and *al-bāriḥa* 'yesterday'. This demonstrative was generalized in Classical Arabic and then in Modern Standard Arabic as the definite article.

Demonstratives function both as determiners (9) and as pronouns (11), and in these cases, they need not be followed by a lexical noun phrase.

(11) *hāḏā* *jayyid-u-n*
this.ms good.ms-nom
'This [one] is good'

Demonstratives can also occur following their corresponding noun (12).

(12) *al-kitāb-u* *hāḏā* *jayyid-u-n*
the-book.ms-nom this.ms good.ms-nom
'This book is good'

This alternation between two positions, one prenominal and the other postnominal, does not obtain freely. Whereas demonstratives can follow any definite noun phrase, they cannot precede a proper name (13b) or a construct state nominal (14b), including possessive noun phrases (15b).

(13) a. *šahida* *zaydun*
witnessed.3sm Zayd
hāḏā *l-ḥādiṯa*
this the-accident
'This Zayd witnessed the accident'

b. **šahida* *hāḏā* *zaydun*
witnessed.3sm this Zayd
l-ḥādiṯa
the-accident
'This Zayd witnessed the accident'

(14) a. *kitābu* *t-tilmīḏati* *hāḏā*
book.ms.nom the-student.fs.gen this
jayyidun
good.ms.nom
'This student's book is good'

b. **hāḏā* *kitābu*
this book.ms.nom
t-tilmīḏati *jayyidun*
the-student.fs.gen good.ms.nom
'This student's book is good'

(15) a. *kitābu-ki* *hāḏā*
book.ms.nom-you.fs this
jayyidun
good.ms.nom
'This book of yours is good'

b. **hāḏā* *kitābu-ki*
this book.ms.nom-you.fs
jayyidun
good.ms.nom
'This book of yours is good'

These observations indicate that it is not sufficient that the noun phrase following the demonstrative be definite; it has to be introduced by the definite article. In other words, the pronominal demonstrative selects the definite article.

Some modern Arabic dialects have developed a proclitic demonstrative, a reduced form of the full demonstrative, which can only occur prenominally. This demonstrative also differs from the full demonstrative in that it does not inflect for number or gender with the corresponding noun and is not able to function as a pronoun (Benmamoun 2000).

An important phenomenon that characterizes the syntax of (in-)definiteness in Arabic is (in-)definiteness agreement or (in-)definiteness spread. That is, adjectives that modify a given noun must agree with that noun, not only in gender, number, and case, but also in (in-) definiteness (16).

(16) a. *qaraʾt-u* *kitāban*
read.1s book.ms.acc.[indef.]
jamīlan
nice.ms.acc.[indef.]
'I read a nice book'

b. *ḍāᶜa* *l-kitāb-u*
lost.3sm the-book.ms.nom
l-jamīl-u
the-nice.ms.nom
'The nice book is lost'

This phenomenon can be used to diagnose the syntactic definiteness of noun phrases, especially those that do not co-occur with the definite article, such as construct state noun phrases. Thus, one can observe that construct state nominals in which the second member is definite must be modified by a definite adjective (17).

(17) *ḍāʿa* *kitāb-u* *t-tilmīḏat-i*
lost.3sm book.ms.nom the-student.fs-gen.
**(l-)jamīl-u*
*(the-)nice.ms.nom
'The student's nice book is lost'

This parallelism between the behavior of (in-) definiteness and that of other phi-features like gender and number has led some researchers to suggest that, in Semitic languages in general, (in-) definiteness is not only a feature of the article, but it is also a lexical feature of the noun and it can function as a trigger of an agreement relation (see Borer 1999 among others).

BIBLIOGRAPHICAL REFERENCES

Benmamoun, Elabbas. 2000. *The feature structure of functional categories: A comparative study of Arabic dialects*. New York: Oxford University Press.

Borer, Hagit. 1999. "Deconstructing the construct". *Beyond principles and parameters*, ed. Kyle Johnson and Ian Roberts, 43–89. Dordrecht: Kluwer Academic Publishers.

Fassi Fehri, Abdelkader. 1993. *Issues in the structure of Arabic clauses and words*. Dordrecht: Kluwer.

Fleisch, Henri. 1961. *Traité de philologie arabe*, I. *Préliminaires, phonétique, morphologie nominale*. Beirut: Imprimerie Catholique.

Kouloughli, Djamel Eddine. 2001. "Sur le statut linguistique du *tanwīn*". *Arabica* 48.20–50.

LINA CHOUEIRI (American University of Beirut)

Dialect → Colloquial

Dialect Geography

1. INTRODUCTION

The Arabic-speaking area is dialectally fragmented, as only few other speech areas are, and many a region remains a terra incognita. This shortcoming is a result of the geographical aspect on one hand: not all localities and regions have been investigated to the same extent. Thus, Khurasan Arabic was discovered only recently (Seeger 2002; → Afghanistan Arabic). On the other hand, the data are heterogeneous as for size, quality, and moment of investigation. This can be explained by the fact that Arabic dialectology is rather a secondary discipline of Arabic or Semitic studies. It is therefore not surprising that Arabic dialectology has not contributed to the general theory of dialect geography. Arabic data are quoted in virtually no introductions to dialectology or dialect geography.

This might be due partly to the political conditions. Because of wars and civil strife, research in many places is almost impossible. Often the local authorities do not provide the necessary research permit, because research on Arabic dialects is considered to be useless or imperialistic and therefore directed against the unity of the Arabic language. Thus, in some cases, research had to be carried out illegally. This may be one of the reasons why existing Arabic dialect atlases do not have such dense nets of research points as European ones. Of the current projects, only the *Atlas linguistique de Tunisie* and a general survey of the Arabic dialects spoken in Israel (see Talmon 2002) have a logistic background comparable to European atlas projects.

2. THE ARABIC DIALECT ATLASES

i. Bergsträsser (1915) is the first Arabic dialect atlas. Bergsträsser attempts a classification of the dialects of Syria and Palestine. For such a vast area the net of research points (67 localities) is too thin, so that no conclusions can be drawn from the isoglosses. The 43 maps deal almost exclusively with phonology and morphology. Actually, the atlas may also be regarded as 'historical' because, due to political changes, in some of Bergsträsser's points Arabic is not spoken anymore.

ii. Cantineau (1940–1946): the atlas is intended as a regional atlas of the Ḥōrān and as a supplement to Bergsträsser (1915) and the monograph *Les parlers arabes du Ḥôrân*. It comprises 60 maps. Almost every locality of the region has been recorded (over 200). Cantineau uses only isoglosses but does not interpret any of the configurations. The maps have been integrated partly into the dialect atlas of Syria (cf. Behnstedt 1997). Cantineau furnishes a classification of the dialects.

iii. Abul Fadl (1960): the 40 dialect maps with 108 research points are a supplement to the text. The isoglosses are not interpreted, but Abul Fadl gives a classification of the dialects of the Šarqiyya. The layout leaves something to be desired. His maps have been integrated into Behnstedt and Woidich (1985).

iv. Behnstedt and Woidich's atlas of the Egyptian dialects (1985) has the densest net of research points (over 800) to date. It deals with 438 themes on 560 maps. The oases are only partly taken into consideration. Isoglosses and synthetic symbols are used. The configurations are interpreted. Cartographically it could be improved and due to too small a format many themes are presented on several maps each, which makes a general view difficult. Part of this atlas is a survey map (*Tübinger Atlas zum Vorderen Orient*. A VIII–12) on which the Egyptian dialects (except the oases) are classified according to a reference system with 50 features.

v. Behnstedt (1985): the atlas deals only with the former North Yemen. The net of points of investigation is, with 165 localities, rather thin. The lexical part is very modest compared to the Egyptian atlas. As a supplement Behnstedt (1987) presents 27 maps with variable data.

vi. Arnold and Behnstedt (1993) is intended as a dialect-geographical investigation of the Qalamūn region (Syria). All 50 localities of the region are registered. It covers 61 themes, especially with respect to language contact. The maps are interpreted, but the layout is wanting.

vii. Behnstedt (1997): with 518 themes the atlas of the Syrian dialects represents the most elaborate Arabic dialect atlas to date. The net of points of investigation is, however, thinner (some 500 localities) than its Egyptian counterpart. Data from Lebanon are partly integrated. Non-Arabic languages are also dealt with (mainly lexicon). There is a classification of the dialects. Some of the figurations are commented on in the *Beiheft*. For one of the disadvantages see Zaborski (1999).

viii. Arnold (1998): the monograph on the Arabic dialects of the Turkish province of Hatay (the research for which was carried out entirely in Germany) contains an appendix with 59 maps for 70 research points. Arnold furnishes a classification of the dialects. The maps are not interpreted. Arnold uses maps with symbols, but the symbols are not constructive since they do not allow recognition of related forms; it is only possible to tell in which religious community the respective forms are used.

ix. The *Atlas linguistique de Tunisie* is the first 'national' project for an Arabic dialect atlas (see Mejri 2000). A total of 250 localities were registered, one for each *sous-préfecture*. Four persons per locality were interviewed (two female, one younger, one older; two male, one younger, one older). There are some 3,000 hours of recordings. For the moment nothing seems to have been published.

x. There is an ongoing project for an atlas of the Arabic dialects spoken in Israel. For a first report, see Talmon (2000).

xi. Seeger has investigated some 68 localities in the surroundings of Ramallah (Palestine) with some quite surprising results (Seeger 2004) and intends to publish a dialect atlas of the area.

xii. A project in Morocco by Behnstedt and Benabbou has failed due to lack of funds. Approximately 150 localities (Arabic and Berber dialects) have been investigated. Part of the material (38 maps) has been published in *Zeitschrift für Arabische Linguistik* 2004.

xiii. Maps for single isoglosses can be found in different monographs and articles, e.g. Cantineau (1940–1946) for Algerian dialects; Fleisch (1974:144) a map of the borders between 'differential' and 'non-differential' dialects in Lebanon; Ingham (1982:36) a map of the reflexes of **j* = /ž/ and /y/ in Southern Iraq and Khuzistan; Ingham (1982:64) a map of Northern Najdi features; Ingham (1982:78) two maps for main isoglosses in the Arabian Peninsula; Jastrow (1990) a map of the *qəltu*-dialects in Mesopotamia and Anatolia; Johnstone (1967:3) a dialect map of Northern Arabia; Johnstone (1967:5) a map of the reflexes of **j* und **q* in the north of the peninsula; and

Johnstone (1967:10) a map of /y/ as reflex of *ǰ. Survey maps are found in Kontzi (1982) with 14 primarily lexical themes, Mörth (1997) with 26 maps of the cardinal numbers, Procházka (1993) with 9 maps of different prepositions, and Taine-Cheikh (1988–1999) with maps of the distribution of the reflexes of *q and the interdentals.

3. Dialect borders

Determining dialect borders is one of the great problems of dialect geography. Some even reject the concept of discrete dialects and accept only transitional areas. In neither German(ic) nor Romance dialect geography do the specialists agree on the part of the grammar which should mainly be dealt with: phonology, morphology, lexicon, or syntax. The problem has to do partly with isoglosses, as "isoglosses usually mark transition zones rather than discrete breaks" (Trudgill 1983:47). Therefore some authors suggest a distinction between 'formal' and 'structural' isoglosses, the latter being more important. Another distinction is the one between 'distinctive' and 'non-distinctive' isoglosses. 'Non-distinctive isogloss' means that one can 'translate' sound by sound from one dialect into another and vice versa, e.g. 'interdentals in dialect A' = 'plosives in dialect B'. As for /q/ - /g/ - /ʾ/ as reflexes of *q the *qāf*-isogloss is non-distinctive in Egypt or Syria, but it is distinctive in many Maghrebi dialects where both /q/ and /g/ are found while their distribution cannot be predicted. The same is true for Standard Iraqi Arabic, some mixed dialects like the dialect of Soukhne (/k/ and /g/), or the dialect of Balāṭ in the Daxla oasis. For Goossens (1969:57), distinctive isoglosses in phonology represent a break and therefore a classification of dialects should be based mainly on these. Since the isoglosses of /q/, /ʾ/, etc. – /g/, interdentals–plosives are non-distinctive in many dialect areas and considered to be 'macro-discrimants' (see Taine-Cheikh 1988– 1989), this approach is certain to be rejected by dialectologists of Arabic.

Another distinction is one between 'primary' and 'secondary' features introduced by Schirmunski (1930). The differences between them are determined by four factors: (a) linguistic aspect = difference from a norm: primary = important (salient), secondary = insignificant; (b) socio-psychological aspect: primary = the speakers are aware of the feature, secondary = the speakers are unaware of it; (c) communicative aspect: primary = impossible, secondary = possible; and (d) areal linguistic aspect: primary = limited distribution of the feature, secondary = broad distribution. Primary features are given up first, for instance, in a situation of → code-switching or dialect contact. The concept is not undisputed in German dialectology, since it implies subjective judgments of the dialectologist. It is applicable to Arabic dialect geography, but it is also intricate. The /ʾ/–/g/ and /g/–/j/ isoglosses (*qāf–jīm*) in the northern Nile valley are non-distinctive. They are primary because /g/–/j/ deviates from the norm (i.e. Standard Egyptian), the speakers are aware of their respective pronunciations, but communication is possible because people know how others speak, and finally the distribution of each features covers a large area. As for Moroccan affricated /t/, which is a primary feature (deviating from an Arabic norm), speakers often are not aware of it, and its regional distribution is large.

For dialect speakers themselves, whose judgment should also be taken into account when fixing dialect borders (see Löffler 1974:136), very often phenomena which are considered by the dialectologist to be minor are seen as typical of another dialect. The speakers' perception of other dialects has been investigated in several dialect areas on 'mental maps', e.g. for the United States by Preston (1988, 1989), for German and Dutch by Weijnen (1968). For Japanese, see <http://nihongo.human.metro-u-ac.jp/long/maps/perceptmaps.htm>. For Romansh, Goebl (2002:192), using dialectometrical methods, comes to the conclusion that the results could be interpreted as if they were based on research with objective data. Applying perceptive dialectology, similarities between dialects and not their differences should be indicated by the informants.

As for isoglosses, Löffler (1974:138) comes to the conclusion that no certain criterion has yet been found for a true dialect border, that the quantitative weighing of the number of lines would be the most objective one, but that there is still no agreement on the grammatical elements which furnish the sharper and linguistically more important borders. The problem may be illustrated by the shibboleth imperfect forms of the Maghrebi dialects *nəktəb-nəktbu*. Are the Egyptian dialects which have similar forms therefore Maghrebi dialects? As for the undoubtedly mixed Maghrebi-Egyptian dialects of the northern oases, Woidich (1993) and Behnstedt (1999) each stress the elements which seemed important to them, Woidich more Egyptian, Behnstedt more Maghrebi.

The fact that on maps with isoglosses very often one cannot see the wood for the trees led Seguy (1973) to renounce them completely and instead to measure the distance between dialects statistically. He introduced the term 'dialectometry', which has become a well-established discipline in dialect geography. There are many different methods of calculating distance or similarity, e.g. with the 'identity test', the 'coherency test', 'multidimensional scaling', or 'dendrograms'. Dutch dialectologists have applied the 'feature frequency method' (Hoppenbrouwers and Hoppenbrouwers 2001) or the 'Levenshtein distance' (Nerbonne and Siedle, forthcoming). The 'Salzburg School', established by Goebl amongst others, works with the RIV ('relative identity value') and the MINMWMAX algorithm. In such measurements all features are weighed equally, a fact which may raise criticism. For the technical side of the 'Salzburg dialectometry', see Goebl (1981, 1982, 1984) and the homepage <http://ald.sbg.ac.at/DM/germ/default. htm>. Of course, calculations and mapping are computerized (cf. also Viereck 1988).

In the so-called 'identity test' a given point of an atlas, which could be a point with the standard dialect or a northernmost point, is chosen and compared with all the rest of the atlas points.

The relative identity is calculated as a percentage. When computerized, any point can be clicked at in order to see the degree of similarity with the rest, as illustrated in Map 1 for the Egyptian oases. The reference point is white. Comparing Cairo with the oases, the nearest dialects are those of Xarga-North (XN) and Xarga-Center (XM) and the least similar those of the Daxla oases (DW, DZW, DZ). Comparing Baḥariyya-Center (BZ) to the others, Baḥariyya-East (BO) is the next, then Baḥariyya-West (BW). Farafra is the most isolated dialect. In this test, all the features have been weighed equally. This merely quantitative test reflects perfectly a traditional classification of the dialects.

Dialectometry is without any doubt an important step forward, and the graphical representation of dialectometrical maps is certainly superior to traditional isogloss maps. See also the perceptive maps in Goebl (2002) with the perceptive map of Weijnen (1968) using the 'little arrow method'.

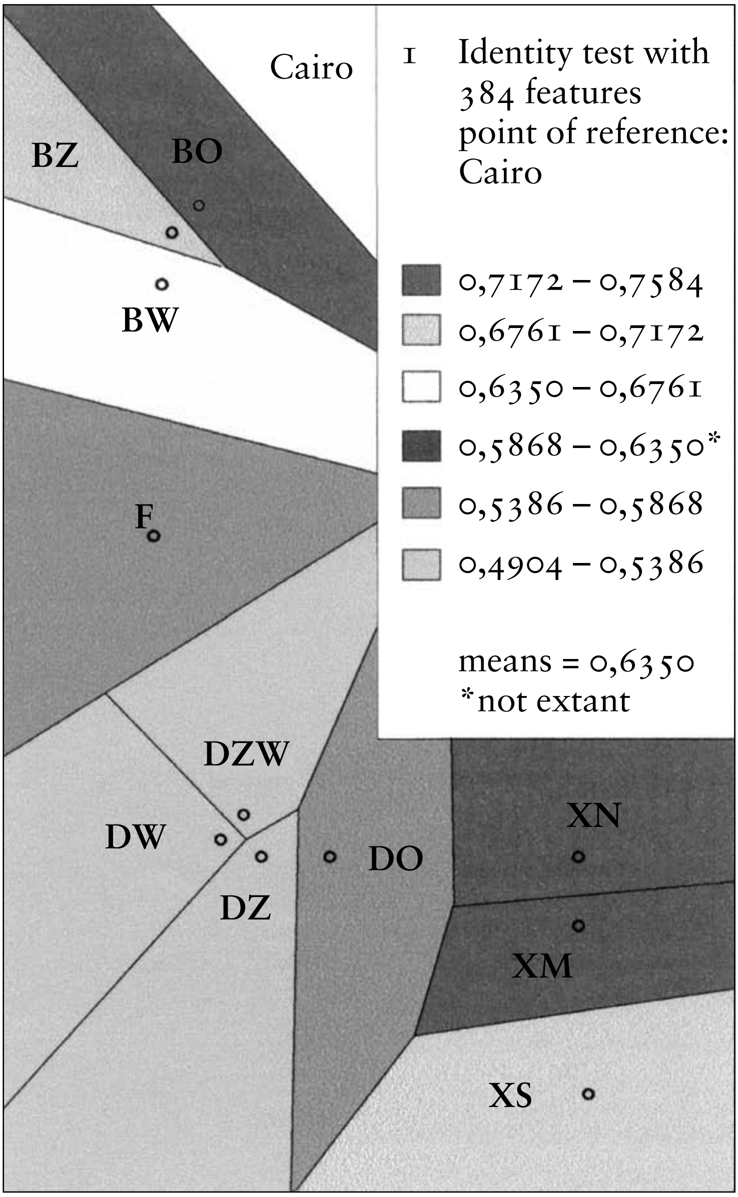
Dialectometrical maps are quantitative. A comparison of Cairene with the dialects of the Nile valley by means of the identity test shows a clear descent in similarity the further south one goes. The important dialect border in the region of Asyūṭ can be seen neatly on the relevant map, but the dialect border between Northern Middle Egyptian and Southern Middle Egyptian does not appear. The reason is evident: Northern Middle Egyptian shows features like the → *bukaṛa* syndrome which neither Cairene nor South Middle Egyptian share. On the other hand, Northern Middle Egyptian has forms of the type *kitīr* 'much', Southern Middle Egyptian *katīr*. The differences with respect to Cairene neutralize each other. Therefore, dialectometry cannot replace traditional qualitative dialect geography (cf. Viereck 1988:547).

4. INTERPRETING DIALECT MAPS

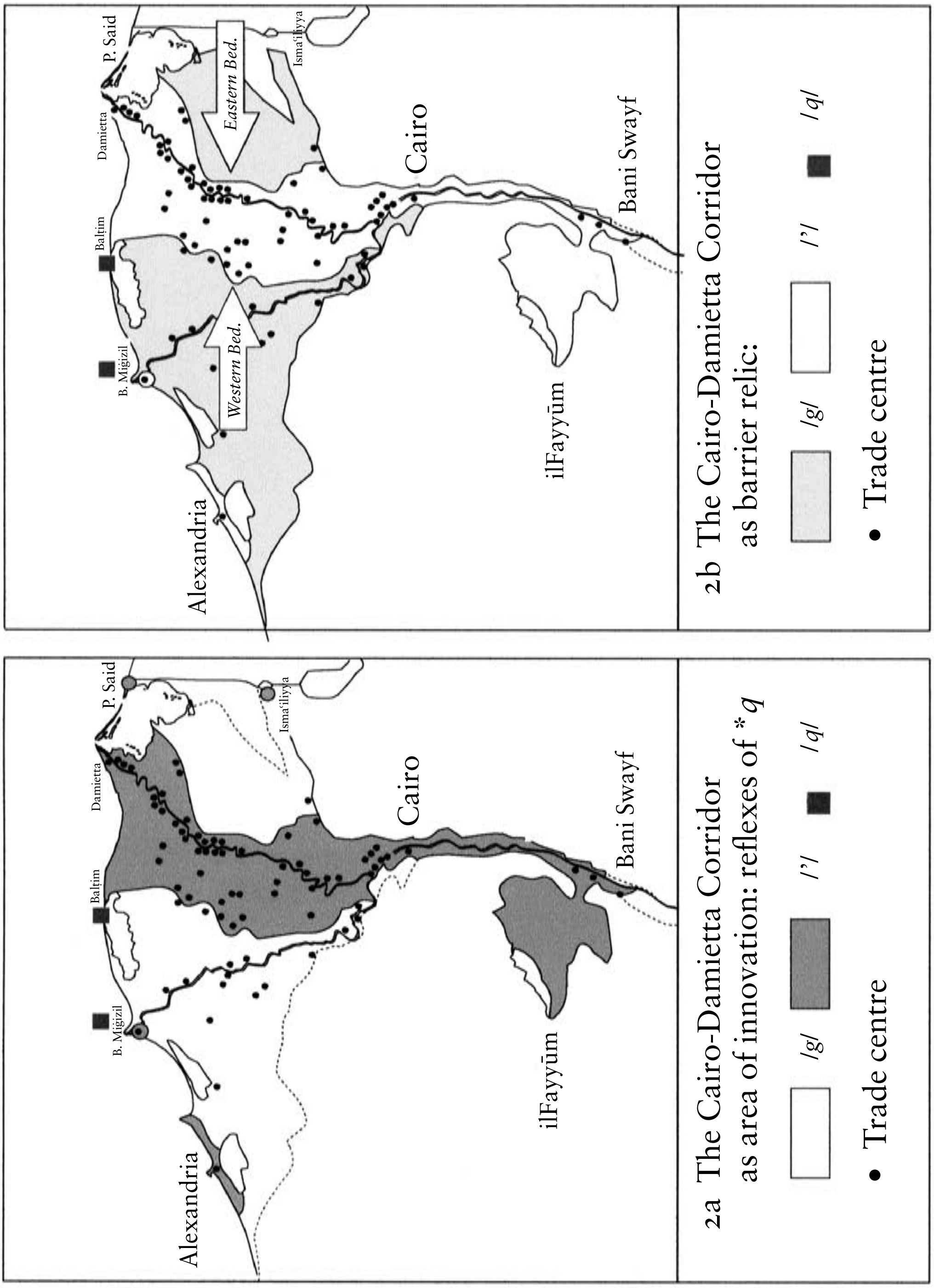
The first attempt to cope with the findings of European dialect geography, which has a long tradition and has developed a large theoretical framework, is to be found in Behnstedt and Woidich (1985). In Behnstedt and Woidich (2005), the authors try to elaborate the subject. Arabic dialect geography is different from European dialect geography in that the latter is concerned with linguistic diffusion, centers of radiation, urban hierarchies, barriers, and the like, but not migration of speakers. This aspect is, however, of extreme importance in Arabic dialect geography and neglecting it would lead to the wrong interpretation of many areal linguistic phenomena. Map 2 shows the complexity of the argument.

The distribution of /g/ and /ʾ/ as reflexes of **g* (*j*) and **q* (*q*) in the Central Delta has the configuration of a corridor (*Schlauch* in German). The first interpretation in Behnstedt and Woidich (1985) followed the diffusion model (see Weinhold 1985), by interpreting Cairene pronunciations on this branch of the Nile as being diffused from Cairo to Egypt's most important harbor in the Middle Ages, namely Damietta. Along this branch, com-

Map 1. Identity test


Cairo
BZ
BO
BW
F
DZW
DW
DZ
DO
XN
XM
XS
1 Identity test with 384 features point of reference: Cairo
0,7172 – 0,7584
0,6761 – 0,7172
0,6350 – 0,6761
0,5868 – 0,6350*
0,5386 – 0,5868
0,4904 – 0,5386
means = 0,6350
*not extant

Map 2. The Cairo–Damietta corridor



2a The Cairo-Damietta Corridor as area of innovation: reflexes of *q

2b The Cairo-Damietta Corridor as barrier relic:

Alexandria
B. Miğizil
Balṭim
Damietta
P. Said
Isma'iliyya
Cairo
ilFayyūm
Bani Swayf
Western Bed.
Eastern Bed.
/g/
/ʾ/
/q/
● Trade centre

mercial centers were concentrated. Cairene pronunciations were supposedly used first in the commercial centers, then spread to the surroundings, and finally all the dots with Cairene pronunciations grew together to form a homogeneous area. This perfectly fits the sociogeographical diffusion models and is proved by Alexandria's dialect. When Napoleon occupied Egypt, Alexandria counted only some 7,000 inhabitants. The reasons for its decay need not be related here. After the repair of the Maḥmūdiyya canal Alexandria recovered and became again the most important harbor of Egypt. With functionaries, traders, and others from Cairo the Cairene dialect was introduced, and completely supplanted the original dialect of Alexandria, which is very close to the rural Buḥayra dialects (see Behnstedt 1980).

This configuration might be interpreted in another way, namely as caused by speakers' migration. Assuming for the Delta the Maghrebi model of Arabization, i.e., a first layer with sedentary type dialects, and a second Bedouin one, one might say that the Delta first had as reflexes of *q and *j /q/ (later /ʾ/)–/g/, and that an originally homogeneous area was squeezed in from two sides. From the beginning of the 12th century, a massive influx of Bedouin from Sinai and Palestine to the Šarqiyya province is attested. This is shown by medieval fiefs (cf. map 552 in Behnstedt and Woidich 1985) and by Syro-Palestinian and Bedouin features in the Šarqiyya dialects. The same applies to the Western Delta, which through the 12th–18th centuries witnessed a constant reflux of Western Bedouin (mainly from Libya) whose dialects clearly show some Maghrebi features. This means that the Cairo–Damietta corridor does not reflect the diffusion of Cairene pronunciation along an important trade route in the Middle Ages, but on the contrary, it is a 'barrier relic' of former pronunciations.

'Dialects', 'dialectal layers', and 'mixed dialects' are often not the result of mixing via vicinity or linguistic diffusion but frequently of migration and mixing of population. The *niktib-niktibu* imperfect in the dialect of the Cairene Jews is not a linguistic innovation or a former stage of Cairene, neither is it a relic enclave (linguistic island) as might be assumed, since the area of the *niktib-niktibu* dialects begins north of Gizeh. Actually, in the 11th century a massive immigration of Maghrebi Jews to Egypt took place when the Maghreb under the Almohades had become 'uninhabitable for non-Muslims' (cf. Goitein 1973:204). The → 'communal dialects', a phenomenon not found in European dialectology, have little to do with religion but rather with distinctive linguistic layers and migration.

The example presented here is an extra-linguistic interpretation of a map. The extra-linguistic method interprets the configuration of maps by political and economic history (trade routes), natural environment, and anthropological, cultural, and topographical facts. It considers language or dialect as one of many areal factors. For example, in northern Yemen the many stereotypical configurations showing a clear division between Tihāma dialects and other dialects correspond to the nature of the country, namely coast vs. mountains. Political history can explain the fact that in Eastern Syria, in an old cultural area, almost no sedentary dialects are spoken, as a result of the invasion of the Mongols (also in Iraq). Towns were destroyed or left by their inhabitants. This empty space was later settled by Bedouin. The plague in the 14th and 15th century had disastrous consequences for Upper Egypt. In the 14th century, between Cairo and Asyūṭ 438 localities are registered, but only 77 in Upper Egypt. The Ottomans compensated for the high human losses by sending immigrants "from neighbouring countries, especially from North Africa; Maghrebi villages were established in Upper Egypt at that time" (Dols 1977:167). In the Šarqiyya province, there is a neat concentration of blood group O in the core area of the dialects of the Eastern Delta (the group designated as OD 1 in Behnstedt and Woidich 1985), which is considered by ʿAmmar (1944) to be the Bedouin blood group (for a similar case in England, see Viereck 1998). Having measured skulls in the northern oases, Mitwalli (1943:119) states for Farafra: "The inhabitants say their ancestors came from Tunisia and the western parts of Tripoli . . . Comparing the curve representing percentage frequencies of the cephalic index for Farafra men with that for the Carthaginian skulls measured by Bertholon and Chantre we find a striking resemblance between the two". Toponymy, especially in Egypt, shows a clear preponderance of Bedouin type toponyms like *Banī, Awlād, al-Faʿālil, al-Faʿālila, Nazla, and Najʿ* where a strong Bedouin element in the local dialects can be detected.

According to Goossens (1969), the extra-linguistic method can be resumed in six points: dialect and trade route; mixed areas; enclaves; the configuration of the areas; graduation of phonological changes; and historical lingustic material.

Bartoli's areal norms should also be seen in this context. They are exhaustively discussed in Weinhold (1985). Actually, only two of his areal norms are still accepted, namely norm 1 'norm of the isolated area', i.e. an area off trade routes, which is normally more conservative than other areas, and norm 4 which is hardly applicable to Arabic dialect geography. In isolated areas, indeed, highly archaic forms can often be detected, as in the ʿAsīr province in Saudi Arabia and northernmost Yemen, but on the other hand, due to their isolation from mainstream dialects, such areas can be very innovative, as for instance the Egyptian oases or those areas which have lost contact with the coherent dialect area of Arabic and are affected by superstrate languages like Turkish, Uzbek, Tajik, Persian, etc. Sometimes Bartoli's norms (also the one concerning the lateral areas) work, sometimes not. The same is true for enclaves. They might be relic enclaves, they might be innovation enclaves, or they might be due to migration (cf. the ʿAjmān in the Gulf area; see Ingham 1982:103). For other configurations and aspects of extra-linguistic interpretations see Behnstedt and Woidich (2005).

The internal linguistic interpretation explains a map by factors within the linguistic system. The most famous example is Gilliéron's map for 'cat' and 'cock', *cattus* and *gattus* in Latin, which in Gascon merged into *gat*, which led to the use of new forms for 'cock'. One should expect many homonymic clashes in Arabic by merging of **l* and **n* in some dialects of the Egyptian oases and some of Morocco, merging of **d*, **ḏ*, and **j* in Upper Egypt, and merging of **s*, **š*, **z*, and **ž* in Maghrebi dialects. As a matter of fact, no such clash has been described until now, only a morphological one in Yemen, namely analogical formation of the pronouns of the 1st sg. *anā-* (masc.) *-anī* (fem.), reanalysis of *-nā* as pronominal suffix of *anā*, *-nī* of *anī*, and filling the gap for the suffix of the 1st person pl. by the pronoun *-iḥna*, *-niḥna*.

Some maps show an evolution or a derivation. According to one's viewpoint one can talk of maps showing how, mainly in the field of phonology and morphology, a former system has developed, or what has become of a former system. Map 258 in Behnstedt (1997) shows a graduation of the merging of masculine and feminine with verbs, pronouns, pronominal suffixes, and demonstrative pronouns. An extra-linguistic interpretation would claim that this is due to dialect contact. An internal interpretation, which is more probable, shows that there is a gradual merging, starting in many places with the 3rd person pl. com. *hinne* 'they'. At the borders of the respective areas in many localities the distinction is maintained in non-noun categories only with adjectives: *kwayysīn* (masc.), *kwayy-sāt* (fem.), no doubt supported by the noun endings *-īn* and *-āt*.

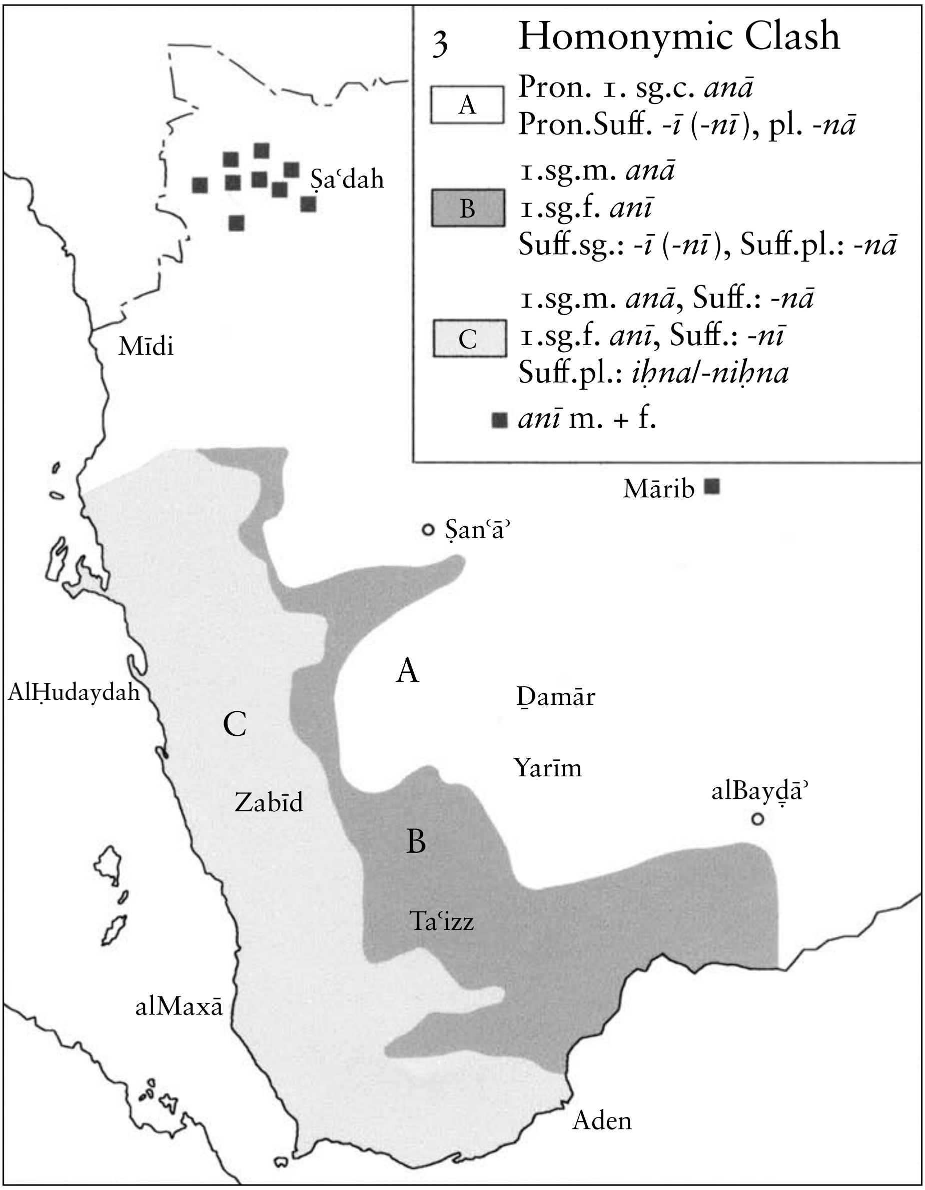
Other maps show differences of systems. In Northern Morocco, there are three areas that have maintained interdentals, namely the northeast, the Jbāla dialects, and the Rīf dialects (Berber). A form *ṯūm* 'garlic' in a Jbāla dialect has to be interpreted within a system other than that for a corresponding form in the northeast. The Jbāla dialects have the same distribution of plosives and interdentals as the Rīf dialects, consequently also *zīṯ* 'oil' vs. northeast *zīt*.

As for the generative interpretation of dialect maps, see mainly Behnstedt and Woidich (1985). It is intricate and very often several interpretations are possible. According to Lang (1982:169ff). there are no advantages of generative dialect geography and traditional dialect geography can interpret maps just as well.

5. Lexical geography

The lexicon is normally neglected in Arabic dialect monographies, and for reasons of space cannot be treated here at length. More recent atlases, however, have accorded it at least one third of the maps' inventory. Lexical maps are normally onomasiological ones, semasiological ones being rare and secondary products. Survey maps of the latter kind, however, might easily be produced with the existing dialect material, e.g. meanings of *bisbās*: 'fennel' in the Maghreb, 'chilis' in Yemen, *xallaṣ* 'to pay' in the Maghreb, 'to finish' in the Mashreq, etc. Onomasiological maps with different designations which, at first glance, have nothing to do with each other may show a common semantic background, e.g. *qarʿa* '(glass) bottle' used in Algeria and

Map 3. Homonymic clash


3 Homonymic Clash
A Pron. 1. sg.c. *anā* Pron.Suff. *-ī* (*-nī*), pl. *-nā*
B 1.sg.m. *anā* 1.sg.f. *anī* Suff.sg.: *-ī* (*-nī*), Suff.pl.: *-nā*
C 1.sg.m. *anā*, Suff.: *-nā* 1.sg.f. *anī*, Suff.: *-nī* Suff.pl.: *iḥna/-niḥna*
■ *anī* m. + f.
Ṣaʿdah
Mīdi
Mārib
Ṣanʿāʾ
A
Ḏamār
AlḤudaydah
C
Yarīm
alBayḍāʾ
Zabīd
B
Taʿizz
alMaxā
Aden

Morocco and *dabbūza* in Tunisia, both originally signifying ‘bottle-gourd’, whereas in other Arabic dialects words for ‘glass’, ‘glass bottle’, or even English ‘bottle’ (*buṭil*) are used.

As for semantic fields, synonyms or rather pseudo-synonyms of the Classical lexicon are often heteronyms in modern dialects, i.e. regional variants. The semantic field ‘nose’ in Classical Arabic consists of *ʾanf*, *maʿṭis*, *manxar*, *nuxra*, *xašm*, *xaṭm*, *xurṭūm*. Out of the seven forms, reflexes of four of them are found in Syria in regional distribution, namely: *xaṭm* and derivatives (mainly Bedouin); *ʾanf* (mainly urban), *manxar* (mainly coast); and *xaṭm* in a small coastal area. There is a similar distribution in Morocco with predominately *nīf*, forms related to *manxar*, forms related to ‘nasality’ (Classical Arabic *xanna* ‘to speak through the nose’) like *xnāna*, *xanfūṛa*.

Bibliographical references

Abul Fadl, Fahmi. 1961. *Volkstümliche Texte in arabischen Bauerndialekten der ägyptischen Provinz Šarqiyya mit dialektgeographischen Untersuchungen zur Lautlehre*. Ph.D diss., University of Münster.

ʿAmmar, ʿAbbas. 1944. *The People of Sharqiyya: Their racial history, serology, physical characters, demography and conditions of life*. 2 vols. Cairo: Société Royale de Géographie de l'Egypte.

Arnold, Werner. 1998. *Die arabischen Dialekte Antiochiens*. Wiesbaden: O. Harrassowitz.

—— and Peter Behnstedt. 1993. *Arabisch–Aramäische Sprachbeziehungen im Qalamūn (Syrien)*. Wiesbaden: O. Harrassowitz.

Behnstedt, Peter. 1980. “Zum ursprünglichen Dialekt von Alexandria”. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 130.35–50.

——. 1985. *Die nordjemenitischen Dialekte*. I. *Atlas*. Wiesbaden: L. Reichert.

——. 1987. *Die Dialekte der Gegend von Ṣaʿdah* (Nord-Jemen). Wiesbaden: O. Harrassowitz.

——. 1997. *Sprachatlas von Syrien*. Wiesbaden: O. Harrassowitz.

——. 1999. “La frontière orientale des parlers maghrébins en Egypte”. *Peuplement et arabisation au Maghreb occidental*, ed. Jorge Aguadé, Patrice Cressier, Ángeles Vicente, 85–96. Madrid and Saragossa: Casa de Velázquez.

—— and Manfred Woidich. 1985. *Die ägyptisch-arabischen Dialekte*. I. *Einleitung und Anmerkungen zu den Karten*. II. *Dialektatlas von Ägypten*. Wiesbaden: L. Reichert.

——. 2005. *Arabische Dialektgeographie: Eine Einführung*. Leiden: E.J. Brill.

Bergsträsser, Gotthelf. 1915. “Sprachatlas von Syrien und Palästina”. *Zeitschrift des Deutschen Palästina-Vereins* 38.169–222.

Cantineau, Jean. 1940. “Les parlers arabes du département d'Oran”. *Revue Africaine* 84.220–231.

——. 1940–1946. *Les parlers arabes du Ḥôrân*. Paris: Klincksieck.

Dols, Michael W. 1977. *The Black Death in the Middle East*. Princeton, N.J.: Princeton University Press.

Fleisch, Henri. 1974. *Etudes d'arabe dialectal*. Beirut: Dar el-Machreq.

Goebl, Hans. 1981. “Eléments d'analyse dialectométrique (avec application à l'AIS)”. *Revue de Linguistique Romane* 45.349–420.

——. 1982. *Dialektometrie: Prinzipien und Methoden des Einsatzes der Numerischen Taxonomie im Bereich der Dialektgeographie*. Vienna: Österreichische Akademie der Wissenschaften.

——. 1984. *Dialektometrische Studien anhand italoromanischer, rätoromanischer und galloromanischer Sprachmaterialien aus AIS und ALF*. 3 vols. Tübingen: Niemeyer.

——. 2002. “Il trattamento dialettometrico della coscienza (meta)linguistica nella Ladinia dolomitica”. *Che cosa ne pensa oggi Chiaffredo Roux? Percorsi della dialettologia precezionale all'alba del nuovo millennio: Atti del convegno internazionale, Bardonecchia 25, 26, 27 maggio 2000*, ed. Monica Cini and Riccardo Regis, 189–201, 203–204. Alessandria, Italy: Edizioni dell'Orso.

Goitein, S.D. 1973. *Letters of medieval Jewish traders*. Princeton, N.J.: Princeton University Press.

Goossens, Jan. 1969. *Strukturelle Dialektgeographie*. Heidelberg: C. Winter.

Hoppenbrouwers, Cor and Geer Hoppenbrouwers. 2001. *De indeling van de Nederlandse streektalen: Dialecten van 156 steden en dorpen geklasseerd volgens de FFM*. Assen: Van Gorcum.

Ingham, Bruce. 1982. *North east Arabian dialect studies*. London: Kegan Paul International.

Jastrow, Otto. 1990. “Die arabischen Dialekte der irakischen Juden”. *Deutscher Orientalistentag vom 26. bis 30. September 1988 in Köln, Ausgewählte Vorträge*, ed. Werner Diem and Abdoljavad Falaturi, 199–205. Stuttgart: F. Steiner.

Johnstone, T.M. 1967. *Eastern Arabian dialect studies*. London: Oxford University Press.

Kontzi, Reinhold 1982. “Maltesisch: Sprachgeschichtliche und areallinguistische Aspekte”. *Die Leistung der Strataforschung und der Kreolistik: Typologische Aspekte der Sprachkontakte: Akten des 5. Symposions über Sprachkontakt in Europa, Mannheim 1982*, ed. P. Sture Ureland, 63–87. Tübingen: Niemeyer.

Lang, Jürgen. 1982. *Sprache im Raum*. Tübingen: Niemeyer.

Löffler, Heinrich. 1974. *Probleme der Dialektologie*. Darmstadt: Wissenschaftliche Buchgesellschaft.

Mejri, Salah. 2000. “L'Atlas linguistique de Tunisie: Le questionnaire morphosyntaxique”. Youssi a.o. (2002:55–59).

Mitwalli, M. 1943. “The population of the Egyptian oases”. *Publications de la Société Royale d'Eygpte* 21.109–312.

Mörth, Karl. 1992. *Die Kardinalzahlwörter von Eins bis Zehn in den neuarabischen Dialekten*. Vienna: WUV-Universitätsverlag.

Nerbonne, John and Christine Siedle. Forthcoming. “Dialektgeographie auf der Grundlage aggregierter Ausspracheunterschiede”. *Zeitschrift für Dialektologie und Linguistik* 2005. [Available at: http://www.let.rug.nl/~nerbonne/papers/nerbonne-siedle-deutsch-aggregiert2004.pdf.]

Preston, Dennis R. 1988. "Methods in the study of dialect perceptions. Methods in dialectology". Thomas (1988:373–395).
——. 1989. *Perceptual dialectology: Nonlinguists' views of areal linguistics*. Dordrecht and Providence, R.I.: Foris Publications.
Procházka, Stefan. 1993. *Die Präpositionen in den neuarabischen Dialekten*. Vienna: VWGÖ.
Schirmunski, Viktor 1930. "Sprachgeschichte und Siedlungsmundarten". *Germanisch-Romanische Monatsschrift* 18.113–122, 171–188.
Seeger, Ullrich. 2002. "Zwei Texte im Dialekt der Araber von Chorasan". "*Sprich doch mit deinen Knechten aramäisch, wir verstehen es!". 60 Beiträge zur Semitistik: Festschrift für Otto Jastrow zum 60. Geburtstag*, ed. Werner Arnold and Hartmut Bobzin, 629–646. Wiesbaden: O. Harrassowitz.
——. 2004. <http://semitistik.uni-hd.de/seeger/english/ ramalla_e.htm.>
Taine-Cheikh, Catherine. 1998–1999. "Deux macro-discriminants de la dialectologie arabe (la réalisation du *qâf* et des interdentales)". *Matériaux Arabes et Sudarabiques, N.S.* 9.11–50.
Talmon, Rafi. 2000. "Preparation of the North Israeli Arabic *Sprachatlas*". Youssi a.o. (2002:68–77).
Thomas, Alan R. (ed.). 1988. *Methods in dialectology: Proceedings of the sixth international conference held at the University College of North Wales, 3–7 August 1987*, Clevedon, Avon and Philadelphia: Multilingual Matters.
Trudgill, Peter. 1983. *On dialect*. Oxford: Blackwell.
Viereck, Wolfgang. 1988. "The computerisation and quantification of linguistic data: Diametrical methods". Thomas (1988:524–550).
——. 1998. "Geolinguistics and haematology: The case of Britain". *Links and Letters* 5.167–179.
Weijnen, Anton. 1968. "Zum Wert subjektiver Dialektgrenzen". *Lingua* 21.594–596.
Weinhold, Norbert. 1985. *Sprachgeographische Distribution und chronologische Schichtung: Untersuchungen zu M. Bartoli und neueren geographischen Theorien*. Hamburg: Buske.
Woidich, Manfred. 1993. "Die Dialekte der ägyptischen Oasen: Westliches oder östliches Arabisch?". *Zeitschrift für Arabische Linguistik* 25.340–359.
Youssi, Abderrahim a.o. (eds.). 2002. *Aspects of the dialects of Arabic today: Proceedings of the 4th Conference of the International Arabic Dialectology Association Marrakech, 1–4 April 2000*. Rabat: AMAPATRIL.
Zaborski, Andrzej. 1999. Review of: Peter Behnstedt, *Sprachatlas von Syrien*. Kartenband, Wiesbaden: O. Harrassowitz, 1997. *Zeitschrift für Arabische Linguistik* 36.98–100.

PETER BEHNSTEDT (Chipiona, Spain)

Dialect Koine

1. INTRODUCTION

The term 'koine' comes from the Greek word *koinē* 'common', referring to the variety of Greek that became the lingua franca, or common language, of the eastern Mediterranean area during the Hellenistic period. It has since been applied to many other languages that share certain features with the original Greek koine. Inspired by the Hellenistic tradition, a number of Arabists used the term to refer to two historical types of Arabic varieties: the pre-Islamic poetic koine and the military or urban dialect koine of the early periods of the Arab conquest. The → poetic koine, which refers to a literary use, will not be dealt with here, although many authors have suggested a dialectal base to this literary koine. The concept of a military or urban koine has been used to explain the emergence and development of what are called the modern Arabic dialects or → Middle Arabic. Not all Arabists agree with this hypothesis and other models have been proposed. The debate around the concepts of koine turns around the role of contact phenomena in the emergence of modern Arabic dialects.

The term 'koineization' refers to a process of interdialect contact leading to an amount of linguistic restructuring. A dialect koine is the stabilized mixed variety that results from this process (Siegel 1993). Koineization usually, but not always, implies that the most peculiar features of each contact dialect are dropped and that the regular/most common features are selected instead. This implies a certain degree of leveling, but without radical restructuring, unlike pidginization. To speak in terms of koineization rather than language mixing or → pidginization means that the varieties in contact are considered to be sub-varieties of the same linguistic system.

The concept of koineization has also been used to describe the changes that are occurring in many contemporary dialects following movements of population and urbanization. This implies that these changes are due to interdialectal contact as much as to Classical–dialectal contact (→ diglossia). The historical situation was dealt with in a number of theoretical articles, which became classic references in the field (Blau 1965; Cohen 1962; Ferguson 1959, 1989; Fück 1950; Larcher 2001; Miller 1986; Rabin 1955; Versteegh 1984, 1997, etc.). The contemporary situation has been approached in numerous works describing specific local situations (urban, rural, or Bedouin), but few papers offer a wider perspective of the contemporary dynamics (Ferguson 1987; Miller 2004; Palva 1982).

2. Dialect koine in the first centuries of the Arab conquest

Most authors agree that the grammatical restructuring that characterized the modern Arabic dialects, compared to Classical Arabic, must have taken place during the first centuries of the Arab conquest, when Arabic spread out from its traditional homeland to expand to vast territories of previously non Arabic-speaking areas. But they disagree on the relationship between Classical Arabic and the modern dialects and on the nature of the restructuring process. The first issue concerns the nature of pre-Islamic Arabic and whether there were pre-Muslim Arabic dialectal varieties or not. The second issue is whether the restructuring processes were (a) induced by external influence (the influence of the local non Arabic languages in a process of second language learning) or (b) the result of an internal drift leading to leveling or (c) a result of interdialectal contact. A number of theories have been put forward but three main streams can be isolated: Ferguson's military/urban koine; Cohen's urban koines; and Versteegh's pidginization/creolization processes.

A number of earlier Arabists assumed that, generally speaking, modern Arabic dialects derive lineally from Classical Arabic or a variety very similar to it (Larcher 2001; A. Miller 1986). Ferguson (1959) refines this hypothesis and postulates that the modern Arabic dialects descend from the earlier language through an Arabic koine. The koine was not identical with any of the earlier dialects and differed in many significant respects from Classical Arabic but was used side by side with the Classical language during the early centuries of the Muslim era. The koine came into existence through a complex process of mutual borrowing and leveling among various dialects and not as a result of a diffusion from a single source. Ferguson postulates one koine and assumes that present dialect differences are innovations that took place following the spread of the koine. He believes that the koineizing process must have begun before the great expansion of Arabic with the spread of Islam, but that the full development of the koine coincided with this expansion, which brought about a mingling of the original dialects and caused large numbers of speakers of other languages to adopt Arabic. The koine developed chiefly in the cities and in the armies and its spread coincided with the spread of an urban Arabo-Islamic culture. He distinguishes this urban koine from Bedouin dialects and assumes that most sedentary dialects came from this urban military koine. Ferguson based his argumentation on a selected list of 14 linguistic features (1 phonological feature, 10 morphological features, 3 lexical features) which, according to him, cannot be analyzed as natural development or drift continuing early trends (which is why the loss of the glottal stop or the reduction of inflectional categories are not included in these 14 features). Among the 14 features are the loss of the dual, → *taltala*, the loss of verbs III*w* verbs, the loss of the feminine comparative, the relative *ʾilli*, etc.

It may be noted that Ferguson was not the first to use the concept of koine, which can be found also in the writings of Fleischer (1847, 1854; see Larcher 2001:595) to designate the common language or Middle Arabic of the period following the conquest. Moreover, the notion, if not the term, of a dialect koine appears also in the Arab grammarians' definition of the Qurayš dialect, the supposed base of the Qurʾānic language and the poetic koine (Larcher 2004).

Ferguson's hypothesis was discussed by Cohen (1962), who questions the existence of a unique dialect koine. He mainly bases his argumentation on the nature of the features selected by Ferguson. In order to sustain the argumentation, those features must be common to most sedentary dialects, should not be found in either Bedouin or Qurʾānic Arabic, and should not be attributable to a general natural drift. Cohen concludes that a number of Ferguson's selected features do not fit these criteria. Many innovations in the sedentary dialects appear to be the result of parallel evolution rather than inheritance from a single koine. Cohen proposes a wave-like diffusion model. A number of different urban koines emerged independently and progressively spread out in various directions.

The koine(s) hypothesis leaves little room for the linguistic influence of local vernacular languages in the early period of the Muslim conquest and stresses the continuity between Old Arabic and the modern dialects. On the contrary, the pidginization/creolization hypothesis defended by Versteegh (1984) considers that non-Arabic speakers played a crucial role in the early restructuring of urban dialects (see also Fück 1950). Versteegh draws a parallel between a number of pidgin/creole contexts and the language situation in the newly conquered cities of the new Arab Empire. He emphasizes that acquisition of Arabic by the majority of the non-Arab urban population was a process of untutored second language acquisition. He concludes that the modern Arabic dialects are the result of an initial process of pidginization/creolization followed by a decreolization trend and a realignment toward the rules of Classical Arabic. Versteegh's hypothesis was much discussed, both at the linguistic and extralinguistic levels. The contemporary dialectal varieties show no definite evidence of an earlier pidginization process. Most features can be analyzed as the result of internal drift (e.g. the case of the dual) or interdialectal contact or universal language trends (Ferguson 1989). Today many authors agree that a less radical process of second language learning can explain the development of a number of modern dialectal features. Historically, there is little evidence that a creolization-like context prevailed in most urban centers of the Arab Empire.

It may be noted that the main arguments raised in favor or against the koineization theory are of a linguistic nature: the presence or absence of such-and-such a feature which could be analyzed as a produce of koineization. Yet, linguistic arguments have to be supported by historical and social data and, for the time being, we still know very little about the social conditions that led to the spread of Arabic in many areas of the Arab Empire (Donner 1981).

3. CONTEMPORARY DIALECT KOINES

Whatever their origin, the modern Arabic vernaculars present a high range of diversity. Dialectologists tend to classify the various vernaculars according to geographical factors (eastern versus western dialects), social factors (Bedouin dialects versus sedentary dialects, and within the latter urban versus rural dialects), or ethnic and religious factors (Muslim dialects versus Christian or Jewish dialects). But in many instances, dialect contact and dialect mixing led to the emergence of mixed dialects. Sedentary dialects have often been leveled by Bedouin dialects following waves of Bedouin settlement in the 10th–13th centuries (Diem 1978; Holes 1995a, 1996). This process was recorded in North Africa following the settlement of the Hilālī tribes through the 12th–13th centuries (Aguadé a.o. 1988), but also in Mesopotamia between the 14th and 18th centuries. A well-described case is Muslim → Baghdadi Arabic, which emerged in the 18th–19th centuries (Blanc 1964). Another case of a 17th–18th-century emergence of a mixed dialect is that of the city of Salṭ in Jordan (Palva 1994). Koineization occurred not only in urban areas, but also in the rural areas lying between different dialect groupings. Examples of such mixed rural dialects are provided by Behnstedt and Woidich in their atlas of Egyptian rural Arabic (Woidich 1997). Many other examples are provided by Johnson (1967) for the Gulf Arabic vernaculars.

In the late 19th century and in the 20th century, rural/urban migration and urbanization led to many cases of dialect mixing and koineization. This koineization trend goes in two main directions:

i. Due to the urbanization of large segments of previously rural speakers, many contemporary urban standards emerge through various degrees of leveling.
ii. The urban dialects of the main cities emerge as national or regional standards and often spread to other cities and to rural areas through the influence of the media; in this respect they are competing with Modern Standard Arabic as prestigious norms (Ferguson 1987; Ibrahim 1986; Palva 1982).

Degrees of koineization and leveling depend on each city history and on the rate of rural/urban migration. Therefore, there is neither a single model nor a common linear development. A number of urban vernaculars, such as Cairo Arabic, went through a process of dialect contact and leveling during the second part of the 19th century, following a significant population renewal (Woidich 1994). Since that time, Cairo Arabic has become more or less established, and migration does not initiate new processes of dialect leveling and koineization. Instead, rural migrants are subjected to a long-term accommodation process to Cairo Arabic. In Morocco, the most typical example of a contemporary koine is the dialect of Casablanca, which emerged in the early 20th century, based on a koineized Hilālī dialect (Aguadé 2003). It is spreading as the national standard and leading to the progressive attrition of the prestigious old urban dialects of the cities of Fes, Rabat, etc. (Messaoudi 2003). This phenomenon is not restricted to Morocco since in most North African old urban centers (Algiers, Constantine, Tetouan, Tunis, etc.), the old urban dialects tend to become restricted to women (Boucherit and Lentin 1989) while the urban koines of the capital cities are expanding, functioning as national dialect koine. In cities, where → communal dialects had coexisted for centuries, a koine tends to become the shared language among the various communities. In Bahrain, for instance, where two communal dialects (the Shiʿi and Sunni dialects) have existed for more than 200 years, the economical changes of the 20th century have led to the emergence of an intercommunal standard urban dialect, spoken in public context and mainly based on the Sunni dialect, due to the political weight of the ruling Sunni families (Holes 1995b). In the relatively recent city of Amman, where different dialects coexist (rural/urban Palestinian and rural/Bedouin Jordanian dialects, cf. Abdel Jawad 1986; Sawaie 1994), it seems that a new urban koine is emerging among the youth who have developed a mixed vowel system (Al Wer 2000). In an expanding city such as Ṣanʿāʾ in Yemen, Ṣanʿānī speakers tend to keep their old vernacular for family interaction and to shift to pan-dialectal items in public settings, although one cannot speak of a Ṣanʿānī koine (Watson 2003).

The development of a koine used in public urban context does not necessarily lead to the loss or total attrition of the different communal dialects. Each urban context needs to be investigated in detail. Likewise, not all urban dialects of the capital cities are imposing themselves as a national standard and there are recorded cases of regional competition within the same country (e.g. Algiers versus Oran). Finally, koineization processes expand sometimes beyond national boundaries. Many non-Egyptian dialects have taken some Cairo Arabic features and there is evidence that some features like the genitive particle *btāʿ* or the verbal prefix *b-* are becoming pan-dialectal features (Palva 1982).

4. CONCLUSION

The past and present states of Arabic dialects/vernaculars indicate that dialect mixing and processes of koineization have been extremely important trends in the development of modern Arabic vernaculars. Many contemporary cases could not be recorded here due to lack of space. Koineization implies a certain degree of leveling and simplification, yet we lack a comprehensive survey of the Arabic linguistic features subjected to leveling. The various examples of dialect koine indicate that all levels of the language can be affected and that the selected koine features are not always the less salient, or the demographically dominant features. Some general trends have been recorded, however, such as the loss of gender distinction for 3rd and 2nd pers. pl. verbal imperfective markers.

BIBLIOGRAPHICAL REFERENCES

Abdel Jawad, Hassan R.S. 1986. "The emergence of an urban dialect in the Jordanian urban centres". *International Journal of the Sociology of Language* 61. 53–63.

Aguadé, Jordi. 2003. "Notes on the Arabic dialect of Casablanca (Morocco)". Ferrando and Sandoval (2003:301–308).

——, Patrice Cressier and Ángeles Vicente (eds). 1998. *Peuplement et arabisation au Maghreb occidental.* Madrid and Saragossa: Casa de Velazquez and Universidad de Zaragoza.

Al Wer, Enam. 2000. "Raising of /a/ and related vocalic movements in the emerging dialect of Amman". *Proceedings of the Third International AIDA Conference 29 March–2 April 1998*, ed. Manwel Mifsud, 77–82. Malta: Salesian Press.

Blanc, Haim. 1964. *Communal dialects in Baghdad.* Cambridge, Mass.: Harvard University.

Blau, Joshua. 1965. *The emergence and linguistic background of Judaeo-Arabic: A study of the origins of Neo-Arabic and Middle Arabic.* Oxford: Oxford University Press.

Boucherit, Aziza and Jérôme Lentin. 1989. "Les dialectes féminins dans le monde arabe: Des dialectes minoritaires et leur évolution". *Genre et langage: Actes du Colloque tenu à Paris X-Nanterre 14–16 Déc. 1988*, ed. Eliane Koskas and Danielle Leeman, 17–37. (= *Linx*, 21.) Paris: Université Paris-10.

Caubet, Dominique and Martine Vanhove (eds.). *Actes des premières Journées internationales de dialectologie arabe.* Paris: INALCO.

Cohen, David. 1962. "Koinè, langues communes et dialectes arabes". *Arabica* 9.119–144.

Diem, Werner. 1978. "Divergenz und Konvergenz im Arabischen". *Arabica* 25.128–147.

Donner, Fred McGraw. 1981. *The early Islamic conquests.* Princeton, N.J.: Princeton University Press.

Ferguson, Charles. 1959. "The Arabic koine". *Language* 35.616–630.

——. 1987. “Standardization as a form of language spread”. *Georgetown University Round Table on Language and Linguistics*, 119–132. (Repr. *Structural studies in Arabic linguistics*, ed. Kirk Belnap and Niloofar Haeri, 69–80. Leiden: E.J. Brill, 1997.)

——. 1989. “Grammatical agreement in Classical Arabic and the modern dialects: A response to Versteegh’s pidginization hypothesis”. *al-ʿArabiyya* 22.5–17. (Repr. *Structural studies in Arabic linguistics*, ed. Kirk Belnap and Niloofar Haeri, 81–98. Leiden: E.J. Brill. 1997.)

Ferrando, Ignacio and Juan José Sánchez Sandoval (eds.). *Aida 5th Conference Proceedings, Cádiz September 2002*. Cadiz: Universidad de Cádiz.

Fück, Johann. 1950. *Arabiya: Untersuchungen zur arabischen Sprach und Stilgeschichte*. Berlin: Akademie Verlag. (Trans. Claude Denizeau, *ʿArabīya: Recherches sur l’histoire de la langue et du style arabe*. Paris: Didier, 1955.)

Holes, Clive. 1995a. *Modern Arabic: Structures, functions and varieties*. London: Longman.

——. “Community, dialect and urbanization in the Arabic-speaking Middle East”. *Bulletin of the School of Oriental and African Studies* 58.270–287.

—— (ed.). 1996. “The Arabic dialects of south eastern Arabia in a socio-historical perspective”. *Zeitschrift für Arabische Linguistik* 31.36–50.

Ibrahim, Muhammad. 1986. “Standard and prestige language: A problem in Arabic sociolinguistics”. *Anthropological Linguistics* 28.115–126.

Johnstone, Thomas M. 1967. *Eastern Arabian dialect studies*. London: Oxford University Press.

Larcher, Pierre. 2001. “Arabe moyen et moyen arabe”. *Arabica* 48.578–609.

—— (ed.). 2004. “Théologie et philologie dans l’islam médiéval: Relecture d’un texte célèbre de Ibn Fâris (X^{e} siècle)”. *Cahiers de l’Institut de Linguistique et des Sciences du Langage de l’Université de Lausanne* 17.101–114.

Messaoudi, Leila. 2003. “Parler citadin, parler urbain: Quelles différences?”. *Sociolinguistique urbaine: Frontières et territoires*, ed. Thierry Bulot and Leila Messaoudi, 105–135. Cortil-Wodon: Editions Modulaires Européennes.

Miller, Ann. 1986. “The origin of the modern Arabic sedentary dialects: An evaluation of several theories”. *al-ʿArabiyya* 19.47–74.

Miller, Catherine. 2004. “Variation and changes in Arabic urban vernaculars”. *Approaches to Arabic dialects: Collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday*, ed. Martine Haak, Rudolf de Jong, and Kees Versteegh, 177– 206. Leiden: E.J. Brill.

Palva, Heikki. 1982. “Patterns of koineization in modern colloquial Arabic”. *Acta Orientalia* 43.13–32.

——. 1994. “Bedouin and sedentary elements in the dialect of es Salṭ: Diachronic notes on the sociolinguistic development”. Caubet and Vanhove (1994:459–469).

Rabin, Chaim. 1955. “The beginning of Classical Arabic”. *Studia Islamica* 4.19–37. (Repr. *Readings in Arabic linguistics*, ed. Salman Al-Ani, 70–88. Bloomington, Ind.: Indiana University Linguistics Club, 1978.)

Sawaie, Mohammed. 1994. *Linguistic variation and speakers’ attitudes*. Damascus: Al Jaffar and Al Jabi Publisher.

Siegel, Jeff. 1993. “Dialect contact and koineization: Review article”. *International Journal of the Sociology of Language* 99.105–122.

Versteegh, Kees. 1984. *Pidginization and creolization: The case of Arabic*. Amsterdam: J. Benjamins.

——. 1997. *The Arabic language*. Edinburgh: Edinburgh University Press.

Watson, Janet C.E. 2003. “Insiders, outsiders and language change in San’a today”. Ferrando and Sandoval (2003:561–570).

Woidich, Manfred. 1994. “Cairo Arabic and the Egyptian dialects”. Caubet and Vanhove (1994: 493–510).

——. 1997. “Egyptian Arabic and dialect contact in historical perspectives”. *Humanism, culture and language in the Near East: Studies in honor of Georg Krotkoff*, ed. Asma Afsaruddin and Mathias Zahniser, 185–197. Winona Lake, Ind.: Eisenbrauns.

CATHERINE MILLER (Aix-en-Provence, France)

Dialect Literature

1. INTRODUCTION

‘Dialect literature’ is defined for the purposes of this entry as materials written in colloquial Arabic to be read rather than heard. Thus, the materials described here are the product of a choice by their authors to disregard the norm, prevalent throughout the history of Arabic literature, that mandates the use of the literary idiom in written communication. Materials composed primarily for oral performance and then transcribed, such as printed traditional poetry and folk stories, have been excluded. Also excluded are materials that mix colloquial and literary features, whether these be → Middle Arabic texts, more recent experiments with a compromise ‘third language’, or novels that use colloquial for dialogue only.

The distribution of such works has been uneven over space and time. It is in Egypt that written colloquial appears to have been produced in the greatest quantity and the polemic around it most elaborated; Lebanon follows. Certainty on this point is hampered, however, by a lack of research, especially outside Egypt. This entry therefore focuses on Egypt and Lebanon. Tunisia and Morocco are dealt with cursorily. For Egypt, the entry reviews both the historical profile of writing in colloquial and the debate over its acceptability.

2. EGYPT

The history of writing in colloquial in Egypt falls spontaneously into two periods. From the 15th till the end of the 18th century fewer than 10 items are known – a small number, even when allowance is made for the vagaries of manuscript survival and the relative lack of interest in colloquial materials among scholars. Following a hiatus of some three-quarters of a century, known production resumes around 1870 and continues to increase gradually, aside from a downswing in the middle of the 20th century, gaining additional momentum in the 1990s; for the period post-1870, some 50 items may be identified with ease and without attempting exhaustive coverage. Since modern prose, drama, and poetry have manifested different dynamics in terms of their use of colloquial, each will be treated separately.

2.1 *Egypt: The early period*

The earliest materials in Egyptian colloquial appear to be the long prose piece *Kitāb Funayn* 'Funayn's letter' in the *Dīwān* of Ibn Sūdūn (1407–1464) and a few verses from the same source (Vrolijk 1998). The former may be a survivor of an established genre of comic colloquial letters – "missives of this type are innumerable" says the 17th-century Yūsuf aš-Širbīnī – and its humor is clownish. The poetry is also mostly light-hearted if sentimental and is placed by the author in the section *Muwaššaḥāt hubāliyya* 'Silly *muwaššaḥs*'. Similar in tone is the *Qaṣīda* on the death of an elephant by an anonymous poet quoted by Ibn ʾIyās in his account of the year 804/1458. A different genre (though the tone is still comic or farcical) is represented by *Misṭarat xayāl munādamat ʾUmm Mujbir* 'The ʾUmm Mujbir party show' by ʿAbd al-Bāqī al-ʾIsḥāqī, the only surviving premodern live-theater text, which dates from between 1644 and 1654 (Moreh 1992). *Hazz al-quḥūf bi-šarḥ qaṣīd ʾAbī Šādūf* 'Brains confounded by the ode of ʾAbū Šādūf expounded' by Yūsuf aš-Širbīnī (written after 1097/1686) stands out in this earlier period for the length of its colloquial passages, which, at almost 7,000 words, provide probably the largest corpus of premodern Egyptian colloquial in a naturalistic style; the tone is satirical (aš-Širbīnī, *forthcoming*). Also from the 17th century (exact date uncertain) is another example of entertainment literature, albeit more serious and formal, in the form of a *Rangstreit* fragment, *Qiṣṣat al-miṣrī wa-r-rīfī* 'The story of the Cairene and the countryman', written in Hebrew characters, from the Cairo Geniza (Goitein 1972). Kahle has published the surviving fragments of different 17th-century versions, by Dāʾūd al-Manāwī and ʿAlī an-Najjār, of a shadow play largely written in strophic verse, entitled *Liʿb at-timsāḥ* 'The crocodile play' (Kahle 1915), as well as of a further shadow play, the *Liʿb al-manār* 'The lighthouse play' by Dāwūd al-ʿAṭṭār and others, also in strophic verse and dating from the early 18th century, which survives in full (Kahle 1930). Also entertainment literature is the final piece from this early period, the anonymous *Ḥikāya Bāsim al-Ḥaddād wa-mā jarā lahu maʿa Hārūn ar-Rašīd* 'The story of Bāsim the Smith and what happened to him with Hārūn ar-Rašīd', a story in the *Arabian Nights* tradition, though, unlike the latter, consistently colloquial in idiom (Landberg 1888). This is the longest colloquial text from the early period. Apparently dating from before 1795, the manuscript is curious in that it contains a second version of the story in Syrian colloquial.

2.2 *Egypt: Modern colloquial prose writing*

Egyptian colloquial prose writing resurfaces in the late 19th century through the conduit of the magazine and the newspaper, new vehicles for colloquial literature that were to play an essential role in its development through the first decades of the following century. The first magazine produced by an Egyptian was *ʾAbū naḏḏāra zarqāʾ* 'The man with the blue glasses' (1877 to at least 1882), and was edited and mostly written by the maverick political satirist Yaʿqūb Ṣannūʿ, who railed against the Khedive and the British largely in colloquial (and French); Ṣannūʿ had earlier introduced drama in (colloquial) Arabic to Egypt (see below). The foundation in 1892 by the reformist journalist ʿAbdallāh an-Nādim of the "humorous, educational (*tahḏībī*), scientific" magazine *al-ʾUstāḏ* 'The Professor', gave a further boost to colloquial writing: up to half of the earlier issues of *al-ʾUstāḏ* consisted of dialogues around important issues of the day designed to reach an audience unfamiliar with literary Arabic (*fuṣḥā*). Though an-Nādim, a staunch defender of the *fuṣḥā*, eventually phased colloquial out of the magazine, humorous but didactic topical dialogues in colloquial became a standard feature of a series of magazines such as *Ḥimārat munyatī* 'The she-ass of my desire' (ca. 1898) and *al-ʾUrġūl* 'The reed pipe' (1894), and this tradition continued, at least intermittently, up to the end of the third decade of the 20th century. The dialogues of Maḥmūd Bayram at-Tūnusī's *is-Sayyid wi-mrātu f-Bārīs* 'The gentleman and his wife in Paris' and *is-Sayyid wi-mrātu f-Maṣr* 'The gentleman and his wife in Egypt', both originally published in his Tunis-based magazine *aš-Šabāb* and subsequently in book form (1923 and 1925) are late and artistically developed examples of the same genre.

A fashion for fictitious memoirs of socially marginal characters 'as told to' editors of magazines, peaking in the late 1920s, produced at least two works in colloquial. *Muzakkarāt fitiwwa* 'Memoirs of a neighborhood bravo' by "Yūsuf ʾAbū l-Ḥajjāj" and *Muzakkarāt naššāl* 'Memoirs of a pickpocket' by "ʿAbd al-ʿAzīz an-Nuṣṣ" were both probably written by Ḥusnī Yūsuf, of *Lisān aš-šaʿb* 'The voice of the people' (Häusler 1989) and were later published in book form (1927/1929, two parts, and 1930). In both cases the tone is implicitly or explicitly critical of the political and social order. Similar are the monologues placed in the mouths of members of the 'traditional' lower classes written by Ḥusayn Šafīq al-Miṣrī that appeared originally in the latter's magazine *al-Fukāha* 'Humor' under the titles *Ḥawādis wi-ʾarāʾ il-Ḥagg Darwīš w-Umm ʾIsmaʿīn* 'Conversations and opinions of Ḥājj Darwīš and ʾUmm ʾIsmāʿīn' and the later *Ḥadīs xalti ʾUmm ʾIbrāhīm wi-xalti ʾIsmāʿīn* 'Conversation between Aunty ʾUmm ʾIbrāhīm and Aunty ʾUmm ʾIsmāʿīn', the former also appearing in 1929 as a book.

Though the period from the 1930s to the 1980s witnessed a decline in the number of works written in colloquial in Egypt, what was produced broke new ground. Muṣṭafā Mušarrafa's *Qanṭara allazi kafar* 'Qanṭara who denied God', written in the 1940s, is the first novel written entirely in colloquial. It is of such quality that, when finally published in the 1960s, its author was hailed by Yūsuf ʾIdrīs as "a giant among writers". Others also experimented with colloquial for longer literary works, which, even if not as successful artistically, extended the range of theorizing around the issue. Thus, in his introduction to his book-length memoir in colloquial *Muzakkarāt Ṭālib Biʿsa* (written in 1942, published in 1965), Luwīs ʿAwaḍ describes himself as having been "kept awake by the problem of linguistic diglossia in Egypt", explaining that, inspired by at-Tūnusī's pioneering use of the colloquial as a language of "narrative, description, and criticism but within the limits of humor", he had attempted to extend the latter's use into the realms of "serious thought, sublime emotions, and even tragedy". Also self-consciously trail-blazing was Ṣabrī ʿUṯmān's *Riḥla fi n-Nīl* 'Journey on the Nile' (1965), which the author sub-titled "the first comic novel in the Egyptian language", and which he prefaced with a lengthy justification of his use of "the modern, or Egyptian, Arabic language . . . as a way of realizing socialism in the field of culture".

The richest and most sustained phase of prose writing in Egyptian colloquial started in the 1990s and continues until the present. The novel is represented by, for instance, Yūsuf al-Qaʿīd's *Laban il-ʿaṣfūr* 'Sparrow's milk' (1994), Bahāʾ ʿAwwād's *Šams il-aṣīl* 'Late afternoon sun' (1998), Ṣafāʾ ʿAbd al-Munʿim's *Ḥalawt ir-rōḥ* 'Zest for Life' (1998), and Sāmiḥ Faraj's *Bahnihōf ištirāsa* 'Bahnhof Strasse' (1999), the latter unusual in that it provides glossaries of specialized, largely automotive, vocabulary. ʾAḥmad Fuʾād Nijm's *il-Fugūmi: Tarīx ḥayāt muwāṭin šāyil fi qalbu . . . waṭan* 'Il-Fugūmi: The life story of a citizen bearing in his heart . . . a homeland' (1992), Muḥammad Nāṣir's *Ūla ʾawwil* 'Hopscotch' (2000), and Fatḥiyya al-ʿAssāl's three-volume *Ḥuḍn il-ʿumr* 'A lifelong embrace' (2002–2003) are autobiographies. Short and experimental fiction in colloquial is to be found in occasional literary magazines such as *Kitāb il-garāḍ* 'The locust book' (January 1996) and among the collection of traditional tales rewritten from a feminist perspective entitled *Malik walla ktāba* 'Heads or Tales?' (2003, e.g. pp. 35–38, Nisma ʾIdrīs's *Zāt ir-ridāʾ il-xafī* 'She of the invisible cloak'). ʿAbd ar-Raḥmān al-ʾAbnūdī's *Jawabāt Ḥarāji il-Guṭṭ* 'Letters of Haraji the Cat' (2001) is unique both for its epistolary form and for its use of the dialect of Upper Egypt. Most innovative, however, is the use of colloquial for expository prose, in a variety of contexts. Masʿūd Šūmān and Majdī al-Jābirī have written literary criticism – *il-Sabʿ taklu l-ḥumāra: Muxtārāt min šiʿr Ibn ʿArūs* 'The lion is eaten by the donkey: Selections from the poetry of Ibn ʿArūs' (1996), while Bayyūmī Qandīl, *Bāb il-Magmaʿ il-Luġawi mxallaʾ* 'The Language Academy's door hangs askew' (2000), and Muṣṭafā Ṣafwān, *il-Kitāba wis-sulṭa* 'Writing and authority', (2001) have written on language and society; the latter applies his theories in *ʿUṭēl*, a translation of Shakespeare's *Othello* (1998). ʿAyda Sayf ad-Dawla has written educational material on gender rights in *Hiyya w-huwwa: zayy baʾḍ* 'She and He: Just the same' (2000). *Barti* 'Party' (from at least 2002), a monthly society news magazine containing articles in colloquial (often mixed with phrases in English and French), and *Iḍḥak li-d-Dunyā* 'Smile for the world', a weekly (from 2005) with regular articles in colloquial, show a renewed interest on the part of the press, though with no pretense to the earlier didacticism. Advertisements increasingly use slogans in colloquial. Finally, the phenomenon, associated with the young, of using colloquial for emails and other electronic communication, often in an *ad hoc* Latin-character transcription, should be noted.

2.3 *Egypt: Modern colloquial drama*

Arabic theater in Egypt was born speaking colloquial – not surprisingly given the fundamental role of dialogue, and hence naturalistic speech, in drama. The first plays to be performed in Egypt in Arabic were those of Yaʿqūb Ṣannūʿ, who between 1871 and 1872 authored and performed in colloquial a large number of sardonic social comedies, of which eight have been published. The more earnest and explicitly didactic Muḥammad ʿUṯmān Jalāl subsequently published translations of Molière – *al-ʾArbaʿ riwāyāt min nuxab at-tiyātrāt* 'The four plays chosen from the best theatrical works' (1873–4) – and Racine – *ar-Riwāyāt al-mufīda fi ʿilm at-tarājīda* 'The useful plays in the science of tragedy' (1883–1884) – in colloquial verse; his works were not performed, however, until the early 20th century. Introducing his Molière, the author asserts that "theatrical pieces are composed to educate . . . and the Europeans did not only newly adopt them but took them in ancient times from the Romans . . . because of their capacity to educate the young and train youth", and in his introduction to Racine he explains that he had "made [the plays'] verse in such as a way as to be understandable to ordinary people, for the everyday language (*al-luġa ad-dārija*) is more appropriate for this situation and has greater impact on the elite and the common people".

From the late 1870s, Egyptian theater was dominated by writers and managers from Lebanon and Syria, who favored Literary Arabic, but by the second decade of the 20th century a new generation of Egyptian writers and actors started to emerge (Badawi 1988:43–67). ʾIbrāhīm Ramzī (1884– 1949) wrote four of his plays (mostly comedies of manners) in colloquial, between 1915 and 1931; his *Duxūl il-ḥammām miš zayy xrugu* 'Getting into the bathhouse isn't like getting out of it' has been described as "the first fully-fledged, truly Egyptian social comedy" (Badawi 1988:76). His contemporary, Muḥammad Taymūr (1892–1921), also wrote four colloquial plays, including comedies such as *ʿAbd is-Sattār ʾAfandi* 'Mr. ʿAbd is-Sattār' (1918) and *il-ʿAšara ṭ-ṭayyiba* 'The ten of diamonds' (before 1921, in collaboration with Badīʿ Xayrī) as well as serious drama on social problems (*il-ʿAṣfūr fi l-qafaṣ* 'The sparrow in the cage', 1918 and *il-Hāwiya* 'The Abyss', 1921) (Badawi 1988:113). The final figure in this generation, ʾAnṭūn Yazbak, also produced serious drama in colloquial with his play *iz-Zabāʾiḥ* 'The sacrifices' (1925). Then as now, 'light' or 'boulevard' colloquial theater coexisted with (and presumably far exceeded in quantity) the 'legitimate' theater, but no textual trace of it remains, the scripts almost never being published. An exception that proves the rule is *Kiškiš Bak ʿUḍw fi l-Barlamān* 'Kishkish Bey, Member of Parliament' (1929): while the authors Najīb ar-Rīḥānī and Badīʿ Xayrī were leading lights of boulevard theater, this particular play is fundamentally didactic in its attempt to explain how parliamentary democracy works.

The establishment in 1935 of a National Theater with a mandate to encourage the use of literary Arabic appears to have stifled the development of colloquial drama over the next twenty years (Badawi 1988:6). Tawfīq al-Ḥakīm, the leading dramatist of the period, wrote only 3 (out of a total of over 80) plays in colloquial: *Ḥayāh taḥaṭṭamat* 'A life destroyed' (1930), *Ruṣāṣa fi* *l-qalb* 'A bullet in the heart' (1931), *az-Zammār* 'The piper' (1932). And al-Ḥakīm's main disciple, Maḥmūd Taymūr, after writing five one-act plays in colloquial during 1941 at the start of his career, switched to Classical Arabic for his 18 other plays, with the exception of *Kidb f- kidb* 'All lies' (1951).

Nuʿmān ʿAšūr's *in-Nās illi taḥt* 'The people downstairs' (1956) marked the start of a return to colloquial as the dominant idiom for both serious and lighter plays sparked by the populism unleashed by the 1952 revolution and which accompanied the revival of the Egyptian theater in the 1950s and 1960s. ʿAšūr's social-realist plays (eight in all) are characterized by "a bold use of the colloquial" (Badawi 1987:143). ʿAšūr was followed in this unapologetic attitude by Saʿd ad-Dīn Wahba (eight plays in colloquial between 1961 and 1970), ʾAlfirīd Faraj (three plays in colloquial between 1956 and 1966; he also wrote in *fuṣḥā*), and Rašād Rušdī, whose use of colloquial for tragic drama in *il-Farāša* 'The butterfly' (1959) marks a return to the practice of such early writers as Yazbak. The most influential writer of the period, Yūsuf ʾIdrīs, wrote all his plays, and most notably the seminal *il-Farafīr* 'The flipflaps' (1964 – the word is invented), in colloquial. The dominant role of colloquial in the legitimate theater was maintained in the second half of the 20th century in the writings of ʿAlī Sālim, Maḥmūd Diyāb, Najīb Surūr, and others, and colloquial is the most common idiom used in the theater today.

2.4 *Egypt: Modern colloquial poetry*

Egypt has a long tradition of written poetry manifesting colloquial features; in common modern usage all such poetry is somewhat imprecisely lumped together under the term *zajal*. Most older *zajal*, however, falls outside our definition of colloquial literature, since it also contains literary features (e.g. desinential inflection [*ʾiʿrāb*], used to regularize the meter and not as a grammatical feature). It is better to think of such verse as → Middle Arabic. Verse that is truly colloquial (with the odd literary poeticism), appears with the first magazines and newspapers (although this implies the existence of undiscovered predecessors) and these magazines promoted *zajal* even more vigorously than they did colloquial prose, since many of them, e.g., the newspaper *as-Sayf* 'The Sword' (from 1910), did not open their pages to colloquial prose but did print examples of *zajal* in almost every issue. Though often anonymous, such verse is sometimes signed by the editors, such as ʿAbdallāh an-Nadīm of *al-ʾUstāḏ*, Muḥammad an-Najjār of *al-ʾUrġūl* or Muḥammad Tawfīq of *Ḥumārat munyatī*. By the 1920s and 1930s newspapers and magazines had their resident *zajal* poets, such as Maḥmūd Ramzī Naẓīm and Muḥammad Yūnus al-Qāḍī in *al-Laṭāʾif al-muṣawwara* 'Illustrated amusing stories', Muḥammad ʿAbd al-Munʿim (ʾAbū Buṭayna) in *al-Fukāha* 'Humor' (from 1926), Faraj as-Sayyid Faraj (ʾAbū Farrāj), who wrote long narrative poems (*qiṣaṣ zajaliyya*), Badīʿ Xayrī (in *as-Sayf*), and others (see Zakariyyā 1980:321–347). This tradition reached its apogee with Maḥmūd Bayram at-Tūnusī (1893–1961), whose poetry of political and social protest "brought Egyptian [*zajal*] to a new level of artistry" (Booth 1998).

Prefigured in the works published in *al-ʿĀmil al-Miṣrī* magazine 'The Egyptian Worker' (from 1930), the Nasserist period produced a populist movement in *zajal* as in the theater that found an outlet in, e.g., *al-Masāʾ* magazine (from November 1957) (Booth 1992). To this point, colloquial poetry "played the part of a public voice; almost exclusively, it addressed issues of public concern" (Booth 1992:423). Today, Egyptian *zajal* continues to flourish, though its practitioners, such as Fuʾād al-Ḥaddād (1927–1975), Ṣalāḥ Šāhīn (1930–1986), and many living poets such as Sayyid Ḥijāb have largely switched their focus to the poet's inner life. ʾAḥmad Fuʾād Nijm and ʿAbd ar-Raḥmān al-ʾAbnūdī (who writes his verse, as he does his prose in Upper Egyptian dialect) are exceptions in their continuing concern with political and social issues.

2.5 *Egypt: The debate*

Whereas the use of colloquial in writing in the premodern period appears to have attracted no special condemnation, or even attention, from the literary elite, who perhaps viewed it as innocuous because restricted to humor and entertainment, the increase in its prominence and diversity in the early modern period did not go unchallenged. Ṭāhā Ḥusayn inveighed against it and Najīb Maḥfūẓ condemned it as a "social disease" (Dawwāra 1996). The most comprehensive attack, however, came in 1964 in the form of an academic study entitled *Tārīx ad-daʿwā ʾilā l-ʿāmmiyya* 'History of the campaign for colloquial' by Naffūsa Zakariyyā. Zakariyyā posits a campaign, mounted by Western Orientalists and British government servants and adopted by Egyptian nationalists, to replace literary Arabic with colloquial, thus precipitating "the most violent crisis that [the Arabic language] has known in its long history" (Zakariyyā 1980:i). According to Zakariyyā, the foreign campaign started with the publication of Wilhelm Spitta's *Qawāʿid al-ʿArabiyya al-ʿāmmiyya fī Miṣr* 'Grammar of Egyptian Colloquial Arabic' in 1880 and culminated in the activities of William Wilcox, initiated by his 1893 lecture *Limā lam tūjad quwwat al-ixtirāʿ ladā l-Miṣriyyīn?* 'Why is there no power of invention among the Egyptians?'. She identifies ʾAḥmad Luṭfī as-Sayyid, Muḥammad Taymūr, and Salāma Mūsā as the campaign's leading Egyptian theorists and lists writers whom she believes were influenced by the campaign. Zakariyyā fails, however, to demonstrate either that the theorists were influenced by, or even aware of, the writings of the foreigners (with the exception of Mūsā, who had read Wilcox), or that most writers in colloquial aimed to oust the literary language from its pre-eminent position.

In retrospect, writing in Egyptian colloquial appears to be a persistent if also sporadic phenomenon; more research is needed to identify the features of the social, intellectual, and political environment that have served to encourage or discourage it. It has been suggested that Egypt's reduction to the status of a province of the Ottoman Empire from the 16th century may have led to a upsurge of interest in colloquial since "when the state is decentralized . . . and the structures at the top are weaker, the cultural forms and patterns from below are more likely to emerge" (Hanna 1998:87). In the late 19th century, the reformist project of the *nahḍa* or Arab Revival included a desire to reach out to those beyond the pale of elite culture; hence the didactic intent and social critical content of much early writing in colloquial from ʿAbdallāh an-Nadīm and Yaʿqub Ṣannūʿ through to the populist theater of the mid-20th century. The adoption of Arab nationalism, however, as an official ideology in the middle of the 20th century seems to have acted as a temporary brake on output. From the middle of the 20th century, the spread of literacy without concomitant initiation into elite culture may be providing an ever-growing pool of writers who use the colloquial as a mark of their identification with the broader culture and likewise of readers for their work. Similarly, the Internet, with its opportunities for informal personal communication, and the market's need for the shortest communicative route to the consumer, may be giving its use further momentum. Booth (1992:419) has noted that the use of colloquial in the modern period may suggest to its readers a number of possibilities: erasure of class borders, the broadening of literary discourse to include all speakers of the language, the assertion of a specific identity, and the reproduction of power relationships (in that educated users may employ the dialect to speak 'to the masses'). Though her remarks were made with regard to poetry, a reading of modern colloquial writing of all sorts raises these issues in constantly shifting combinations throughout the modern period. Now, at the start of the 21st century, it may be asked whether writing in colloquial has achieved sufficient critical mass that it is possible to discern in the nexus of writer and reader a 'culture of the colloquial'.

As for content, it may be noted that prose writing in colloquial – the bellwether for change in this field – which was limited almost entirely to the sphere of entertainment and humor in the early period, has become steadily more diversified in the modern period. Starting in the 1870s with the didactic dialogue, political journalism, and verse concerned with issues of public concern, and with a strong tendency to humor and satire, the field expanded to include fictitious biography and comic monologue in the early years of the 20th century, then embraced the novel, and had moved into the critical area of impersonal expository prose by the end of the century, a change accompanied by its emancipation from an earlier predominance of the first person narrative (Woidich 1994:14–15). Almost all colloquial writing has been in the prestige dialect of Cairo.

3. LEBANON

Lebanon comes closest to Egypt in terms of the age and vigor of its dialect literature tradition, although Lebanon's is dominated by verse. From the late 15th until the late 17th century, Maronite clergy of Mount Lebanon trained in Rome and their associates wrote long poems in Syriac script (*karšūnī*). The earliest and best known of these priest-poets, Jibrāʾīl ibn al-Qilāʿī (ca. 1450–ca. 1516), appears to have written in colloquial; however, the orthographic uncertainties require that the linguistic status of each work be determined separately, a task that is in hand (Kallas 1997, 2000). With the spread of printing in the 19th century, a previously oral *zajal* tradition of village 'minstrels' (Lecerf 1932:220) started to take on a more urban and perhaps literary form, and by the early 20th century *zajal* was being promoted by newspapers and magazines such as *al-Mašriq* 'The Levant' (from 1898) and *ad-Dabbūr* 'The Hornet' (from 1925); the name of ʿUmar az-Ziʿinnī (1895–1961) is particularly associated with this period. The cheap quality of the booklets in which *zajal* was also published up to the 1930s (Lecerf 1932:212) implies, however, a poor and largely illiterate audience. This changed definitively in the 1940s, when a modern school of *zajal*, of whom the best-known proponent is the outstanding lyric poet Mīšāl (Michel) Ṭrād (1912–1998), gained huge popularity (Abdel-Nour 1966:7); Ṭrād sounds a rare note with his explicit rejection of formal Arabic as a literary tool, describing the *fuṣḥā* as "un vieux bostonien sclérose, au dentier doré" (Kallas 2003:450). Beirut at this time boasted five reviews dedicated exclusively to colloquial verse and an association of 200 poets (the *Emirat de zajal*) that convened the first (and only) congress of Arab *zajal*-poets there in 1945 (Abdel-Nour 1966: 10); in 1957, Fuʾād al-Bustānī could write that this school "fit si bien que le *zajal* est devenu méconnaissable: les traits de cette poésie se sont transformés si profondément qu'on est porté a changer son nom spécifique pour celui de poésie tout court" (Abdel-Nour 1966:x). Lebanon's colloquial verse tradition continues to flourish. Saʿīd ʿAql (b. 1912) and his circle are notable for asserting that the Lebanese dialect is of Phoenician rather than Arabic origin and for using a Roman-character transcription (for a list of such works, see ʿAql 1997:34–35). Mūrīs (Maurice) ʿAwwād (b. 1934), after ʿAql perhaps Lebanon's best-known living poet, has collected in his *l-ʾAntulujya l-Libnaniyyi* 'The Lebanese anthology' (1983) the works of 171 colloquial poets from the period 1800–1982.

The Lebanese colloquial prose tradition goes back to at least 1892 when Ṭannūs al-Ḥurr published *Riwāyat aš-šābb as-sikkīr ʾay Qiṣṣat Naṣṣūr as-Sikrī* 'The tale of the drunken youth, or The story of Naṣṣūr the Drunkard'. Of a similar tone are al-Xūrī Šukrī's *at-Tuḥfa l-ʿāmmiyya fi qiṣṣat Finyānūs* 'The popular gem on the story of Finyānūs' (1902) and Ḥannā al-Xūrī al-Fiġālī's series of moralistic tales (Abdel-Nour 1966:5, n. 4), of which the best known is *Rasāʾil Šmūni* 'Letters of Šmūni', an epistolary novel describing the adventures of a village girl who goes to work as a servant in Beirut (Lecerf 1932:102–111). Saʿīd ʿAql and Mūrīs ʿAwwād (see above) have both also written prose works in colloquial, the former publishing a weekly newspaper *Lebnaan* (in Roman transcription) (1983–1988), the latter a short story, *t-Tiṣwīni* 'The enclosure' (1985), a translation of de Saint Exupéry's *Le petit prince* (*l-Amīr iz-zġīr*, 1986), and a translation of the *New Testament* (*l-Injīl*, 2002); all ʿAwwād's works are fully vernacular, from the titles to the typographical information.

4. TUNISIA AND MOROCCO

Maḥmūd Bayram at-Tūnusī (see above), who, though raised in Egypt, was of Tunisian origin, is said to have written verse in Tunisian as well as Egyptian colloquial (Booth 1998:784). The plays *ʾIsmāʿīl Bāšā* 'Ismāʿīl Pasha' (1997) by Muḥammad ʾIdrīs and *Famīlya* 'Family' (1997) by al-Fāḍil al-Jʿaybi are examples of the colloquial theater in Tunisia; unlike their Egyptian and other counterparts, these employ colloquial for stage directions as well as dialogue. Al-Hādī al-Bāliġ has translated de Saint Exupéry's *Le petit prince* into Tunisian colloquial prose as *al-ʾAmīr aṣ-ṣġīr* (1997). In Morocco, written colloquial appears to be associated primarily with poetry (*zajal*) (e.g. ʾAḥmad Limsīḥ, aṭ-Ṭayyib al-ʿAlj, al-Maskīnī al-ʿĀfir, az-Zubayr bin Bu Štā, Murād al-Qādirī, ʾIdrīs al-Misnāwī, Muḥammad Miskīn), while colloquial drama is associated with the names of Yūsuf Fāḍil and Drīs ar-Rūx.

BIBLIOGRAPHICAL REFERENCES

(early works and secondary sources)

Abdel-Nour, Jabbour. 1966. *Étude sur la poésie dialectale au Liban*. 2nd ed. Beirut: Publications de l'Université Libanaise.

Badawi, Mustafa M. 1987. *Modern Arabic drama in Egypt*. Cambridge: Cambridge University Press.

——. 1988. *Early Arabic drama*. Cambridge: Cambridge University Press.
Booth, Marilyn. 1992. "Colloquial Arabic poetry, politics, and the press in Modern Egypt". *International Journal of Middle East Studies* 24. 419–440.
——. 1998. "al-Tūnisī, Maḥmūd Bayram". *Encyclopedia of Arabic literature*, ed. Julie Scott Meisami and Paul Starkey, II, 784–785. London and New York: Routledge.
Dawwāra, Fuʾād. 1996. *ʿAšara ʾudabāʾ yatahaddaṭūn*. 3rd rev. ed. Cairo: al-Hayʾa al-Miṣriyya li-l-Kitāb.
Goitein, S.D. 1972. "Townsman and fellah: A Geniza text from the seventeenth century". *Asian and African Studies* 8.257–263.
Häusler, Martina. 1989. *Fiktive ägyptische Autobiographien der zwanziger und dreissiger Jahre*. Frankfurt: P. Lang.
Kahle, Paul. 1915. "Das Krokodilspiel (*Liʿb et-Timsâḥ*): Ein ägyptisches Schattenspiel nach alten Handschriften und modernen Aufzeichnungen". *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen*, 288–359.
—— with Georg Jacob. 1930. *Der Leuchtturm von Alexandria: Ein arabisches Schattenspiel aus dem mittelalterlichen Ägypten*, 22–26 (Arabic) Stuttgart: Kohlhammer.
Kallas, Elie. 1997. "Manoscritti neoarabo-cristiani del Monte Libano: Idioma, poeti e opera (XV–XVII Sec.)". *Quaderni di Studi Arabi* 15.131–150.
——. 2000. "Ibn al-Qilāʿī (15ème–16ème siècles) pionnier de la literature néoarabe chrétienne du Mont-Liban". *Quaderni di Studi Arabi* 18.221–230.
——. 2003. "Michel Ṭrād, in memoriam". *Arabica* 50. 447–463.
—— (intro. and trans.). 1997. *Saʿīd ʿAql, Yaara. Inno alla donna*. Venice: Cafoscarina.
Landberg, Carlo de (ed.). 1888. *Basim le forgeron et Harun er-Rachid: Texte arabe en dialecte d'Égypte et de Syrie d'après les manuscripts de Leide, de Gotha et du Caire et accompagné d'une traduction et d'un glossaire par le Comte Carlo de Landberg*. Leiden: E.J. Brill.
Lecerf, Jean. 1932. "Littérature dialectale et renaissance arabe moderne, I". *Bulletin d'Etudes Orientales* 2.179–258.
Moreh, Shmuel. 1992. *Live theatre and dramatic literature in the medieval Arab world*. New York: New York University Press.
Širbīnī, Yūsūf aš-. *Forthcoming. Hazz al-quḥūf bi-šarḥ qaṣīd ʾAbī Šādūf*, ed. Humphrey Davies. Leuven: Peeters.
Vrolijk, Arnoud. 1998. *Bringing a laugh to a scowling face: A study and critical edition of* Nuzhat an-nufūs wa-muḍḥik al-ʿabūs *by ʿAlī Ibn Sūdūn al-Bašbūġāwī*. Leiden: Research School CNWS.
Woidich, Manfred. 1994. "Egyptian Arabic as a written language". Unpublished paper delivered in Cracow.
Zakariyyā (Saʿīd), Naffūsa. 1980. *Tārīx ad-daʿwā ʾilā l-ʿāmmiyya wa-ʾāṯāruha fī Miṣr*. 2nd ed. Cairo: Maṭābiʿ Dūr an-Nāšir al-Jāmiʿī.

HUMPHREY DAVIES
(American University in Cairo)

Dialects: Classification

1. GENERAL REMARKS

Generalizing labels such as 'Egyptian Arabic', 'Syrian Arabic', or 'Moroccan Arabic' are commonly used to refer to dialect types spoken in the respective countries. In textbook titles and names of courses in Spoken Arabic, they are used for the sake of convenience, although in fact they often refer to the dialects of the capital cities. This is not merely a simplification but, in a sense, it is also justified because of the ongoing trend toward regional standard dialects with the dialects of the urban centers as the models. A striking example of this development is the influence of Cairene not only on the dialects of the neighborhood, but on the dialects of urban centers such as Alexandria and Port Said as well, in which the structure of Cairene already predominates (Behnstedt and Woidich 1983).

When systematic classifications based on well-defined criteria are aimed at, there is more than one choice for the approach. If the interest is purely synchronic, the classifications can be made on the basis of an adequate selection of synchronically well-documented linguistic variables for each dialect or group of dialects, contrasted with their counterparts in other dialects, without consideration of diachronic and extra-linguistic criteria. If the interest is focused on cultural and historical points of view, diachronic and comparative data play a crucial role. Linguistic features mirroring the movements and interrelationships of various groups, as well as their ecological environments, stand out as relevant criteria for classification.

Dialects are identified and their boundaries defined by means of isoglosses. Drawing a number of isoglosses on a map normally exhibits border areas in which a number of isoglosses lie close enough together to constitute bundles of isoglosses marking boundaries between different dialect areas. The bundles normally reveal the focal area of a dialect, and between the focal areas there are transitional areas in which the isoglosses do not tally with the bundles and in which contrasting items may be used interchangeably. In a parallel way, groups of dialects can be identified. A prerequisite for exact synchronic classifications is thorough dialect-geographical study in which a sufficient number of relevant isoglosses have been drawn on a map. Early dialect maps of Arabic are Bergsträsser's map of Syria and Palestine (1915) and Cantineau's atlas of the Ḥōrān region in southern Syria (1940); more recent works are Behnstedt's dialect atlas of North Yemen (1985), Behnstedt's and Woidich's atlas of Egypt (1985), Behnstedt's atlas of Syria (1997), and the dialect-geographical appendix of de Jong's grammar of the Bedouin dialects of the northern Sinai littoral (2000). An exceptionally illuminating single example of the results of a systematic survey is the colored overview map of the sedentary Arabic dialects of Egypt (Behnstedt and Woidich 1983), which, in addition to twelve distinct dialect areas, specifies two areas in which different dialects coexist, six transitional areas, and a number of locally limited dialects, all defined by means of 50 selected distinctive features. This map, based on purely synchronic material, contains details – among them 'town dialect with Cairene elements' and 'dialect in regression' – which suggest that it can also be read diachronically.

2. Eastern vs. Western dialects

Dialect-geographically, the Arabic-speaking area can be divided into an Eastern *(mašriqī)* and a Western *(maġribī)* dialect group. Until the 1970s, the boundary between the two groups was commonly drawn from the western border of Egypt on the Mediterranean coast in the north to Lake Chad in the south. As the most distinctive individual isogloss, the conservative inflection of the 1st persons sg. and pl. in the imperfect *(aktib*, *niktib)* in the Eastern group, and the paradigmatically leveled innovative inflection *(niktib*, *niktibu/níkitbu)* in the Western dialects was used. However, more detailed investigations have shown that this isogloss runs through the western Delta and follows the Nile Valley between Asyūṭ and Luxor (Behnstedt and Woidich 1985, maps 210–213). Because the dialects of these areas share a significant number of distinctive features with the Egyptian dialects, this isogloss cannot be used as an absolute classificatory criterion, but other typologically prominent differences must be considered as well. Important distinctive features of the Western dialect group include the following: (a) loss of inherited short vowels in medial position; (b) non-phonemic vowel quantity; (c) aspiration of *t* [t^{s}] < **t* and **ṯ*; (d) the syllable patterns CVCC > CCVC (**rijl* > *ržəl*) and CVCV- > CCV- (**katab* > *ktəb*); (e) the use of an innovative indefinite article **wāḥid al-*; (f) the use of *-āš* (**ʾayy šī* < *ʾayy šay*ʾ) to form adverbs and conjunctions (*kīfāš* 'how?', *bāš* 'in order to', etc.); and (g) high ratio of analytical genitives (Marçais 1977:iv–vii). Among these, (a), (b), (c), and (d) apply especially to the Western branch of the *maġribī* dialects. In addition, there are noticeable differences in the vocabulary, both in lexical items and their semantic sense.

3. Bedouin vs. sedentary dialects

Arabic dialects cannot be properly classified without attention to the stratification of society. One relevant point of departure is a sociologically-based grouping of them into sedentary *(ḥaḍarī)* and Bedouin *(badawī)* dialects. The sedentary dialects can further be divided into urban *(madanī)* and rural (*qarawī* 'village' or *fallāḥī* 'peasant') dialects. These divisions reflect the history of settlement and are applicable to the classification of the dialects in virtually the entire Arabic-speaking world, but with a wide range of variation as to the degree of mutual divergences. It has to be emphasized that these designations in this context refer to different dialect types, irrespective of the present-day division between urban, rural, and Bedouin populations. As a result of radical changes in the course of history, in a number of cities the majority of the population speak dialects of a Bedouin, or sedentarized Bedouin type, many villagers speak Arabic of an urban type, and in several old urban centers the inhabitants speak Bedouinized dialects.

In a classification exclusively based on linguistic contrasts, scarcely any single criterion besides the reflexes of **q* distinguishing between the Bedouin-type and sedentary-type dialects can be found. However, there are a number of prominent typological features, some of which are shared by all sedentary dialects, yet without constituting a contrast with all dialects of Bedouin type, and vice versa, and there are features constituting significant partial contrasts between the two groups. In the following list, the kind of contrast is indicated by adding (A) to the criteria which are shared by virtually all the dialects belonging to the group, and (P) to the criteria which are shared by a substantial part of the respective group (Table 1).

Table 1. Features of Bedouin and sedentary dialects

Bedouin dialects	Sedentary dialects
retained interdental fricatives (A)	interdental fricatives > postdental stops (P)
partially retained and generalized indefinite marker *-in (tanwīn)* (P)	no indefinite marker *-in*, except in formulaic expressions (A)
retained gender distinction in plural (P)	no gender distinction in finite verbs and personal pronouns (P)
no verb modifiers in the imperfect (P)	different verb modifiers in the imperfect (A)
internal passives productively used (P)	absence of internal passives (A)
retained productivity of Form IV (P)	absence of Form IV (P)
very low frequency of analytical genitive	common use of analytical genitive structures (A)

Among Bedouin dialects, a division can be made between those which use phonetically-conditioned affrication of *g* and *k* (peninsular Bedouin dialects) and those which do not have affricated allophones (northwest Arabian dialects, Egyptian and North African Bedouin dialects). In the classification of a more restricted group of Bedouin dialects, the contrasting isoglosses are the different variants of the affricated allophones: *g/j* and *k/č* (Syro-Mesopotamian = ‘pre-‘Anazī’) vs. *g/ǵ* and *k/ć* (northern Najdī = ‘Šammarī’ or central Najdī = “Anazī’). Another criterion is the division between the inflectional suffixes in the imperfect: *-īn*, *-ūn* (Najdī, Ḥijāzī Bedouin, the Gulf dialects, sedentary and Bedouin Mesopotamian) vs. *-ī*, *-ū* (urban Ḥijāzī, northwest Arabian, Bedouin of Egypt and North Africa; common sedentary type).

As is obvious from these differences, the Bedouin dialects have retained more morphophonemic categories than the sedentary dialects. However, they also exhibit innovations. Thus, the partial retention of the → *tanwīn* in the Najdī dialects is a striking conservative feature, but this *tanwīn* is different from the Old Arabic indefinite marker. It occurs in new morphological categories (e.g. suffixed to sound pl. masc.: *muslimīnin*) and has partly new functions. Another highly conservative feature is the retention of the internal (apophonic) passive. This category also exhibits innovations: in the northern Najdī dialects the passive vocalism in Form I (act. *kitab*, pass. *ktib* < **kitib*) has been applied to derived forms as well: *libbis* ‘he was dressed’, *ʿīlij* ‘he was treated’, *irsil* ‘he was sent’, *tgīsim* ‘it was shared’, *iʿtibir* ‘it was considered’, *stigbil* ‘he was welcomed’. Another innovation in verbal morphology is the *n*-passive of the *t*-reflexives: *yintalabbas* ‘he can/should be dressed’, *tintigāsimōn* ‘you [pl. masc.] are apt to be shared with’ (Abboud 1979:474, 476–477; cf. the *nt-/tn-* forms in Algeria, Marçais 1958:195–196).

It is a general tendency in Arabic dialects to develop toward more analytical structures. The sedentary dialects have as a rule proceeded farther than the Bedouin dialects in this direction. The domains in morphosyntax most clearly displaying the sedentary vs. Bedouin dichotomy are the aspect/tense system in the imperfect and the genitive structures. In Bedouin dialects of the Najdī type, the old aspect-centered system is preserved, with only incipient development in the macro-structure toward a new tense-based system (Ingham 1994:87), whereas most sedentary dialects have a well-developed relative tense system implying the use of different verb modifiers. The very low frequency of analytical genitive structures is another synthetic trait in Bedouin dialects.

4. The sedentary dialects of the Eastern group

4.1 *Yemen*

Contrary to most sedentary dialects of Arabic, the dialects spoken in the southern parts of the Arabian Peninsula do not result from developments that have taken place as a consequence of the spread of Arabic to the Fertile Crescent and to the African continent. Therefore, they exhibit many archaic features not found in the more leveled dialects of the other parts of the Arabic-speaking world. Among them, the following are attested in different parts of Yemen: the relative pronoun *allaḏī*, *mā* ‘what?’, *ayna* ‘where?’, *ata* ‘to come’, *ams* ‘yesterday’, *maʾ*, *māʾ* ‘water’, and reflexes of **raʾā* ‘to see’. An archaic feature interesting from the comparative Semitic point of view is *-k* (instead of *-t*) in the personal morphemes of the perfect (e.g. *katabku* 'I wrote'). Idiosyncratic innovations include paradigmatically complemented gender distinction in the 1st pers. sg. personal pronoun: *ana* m., *ani* fem., suffixed pronoun 1st pers. sg. masc. *-na*, fem. *-ni*, 1st pers. pl. com. *-ḥna*. Developments typical of sedentary dialects can also be noticed, among them the verb modifiers *bi-*, 1st pers. sg. *bayn-* (present tense), and *ša-*, *š-*, *bā-* (future and volitive/intention), as well as the split-morpheme negation: *mā fīš*, *mā būš* 'there is not'. (Behnstedt 1985)

4.2 *Mesopotamia*

Outside the Arabian Peninsula, most sedentary dialects are descendants of the dialects spoken in these areas in the first Islamic centuries. Since that time they have developed relatively independently of each other in four greater dialect areas: Mesopotamia, Syria, Egypt, and North Africa.

In Mesopotamia the *qəltu* dialects represent the old sedentary dialect type. Their salient features include the retention of the Old Arabic 1st pers. sg. morpheme *-tu* (*qəltu* 'I said') in the perfect; the *q* reflex of **q*; final-stressed feminine forms of color adjectives; and invariable suffixed pronoun *-kī* in the 2nd pers. sg. fem. (Jastrow 1978:1–32).

Characteristic features shared by all Mesopotamian dialects include the 2nd pers. sg. fem. *-īn* and 2nd and 3rd pers. pl. *-ūn* suffixes in the imperfect, which they have in common with the Najdī and Syro-Mesopotamian Bedouin dialects, and the use of analytical genitives, with *māl* as the genitive marker.

One of the most striking sedentary traits of the *qəltu* dialects is the well-developed system of verb modifiers. In this respect, the Anatolian group stands out (Jastrow 1978:299–311). Another Anatolian trait is the substitution of *n* for *m*, not only in personal pronouns as in most sedentary dialects in Syria and Lebanon (Behnstedt 1997, map 257; Jastrow 1978: 223–225), but in inflectional morphemes of the perfect as well: *hənne* 'they', *baytən* 'their house', *əntən* 'you [pl. com.]', *baytkon* 'your [pl. com.] house', *jītən* 'you [pl. com.] came' (Mardin). An additional conspicuous innovation, obviously due to language contact, is the use of a copula in nominal sentences, based on independent personal pronouns: *bayti gbīr-we* 'my house is big' (Mardin town). The Tigris group has a very strong → *ʾimāla*, e.g. *basītīn* 'gardens' (Bəḥzāni, near Mosul) (Jastrow 1978: 26–28), a feature shared by northern and coastal Syrian dialects, and particularly by the dialect of the Suxne oasis (*lsīn* 'tongue', *lībis* 'clothed', *ṯīni* 'second', Behnstedt 1994:30; 1997, maps 43, 45, 48) as well as by the now virtually extinct Jewish dialect of Baghdad (*klīb* 'dogs', *jīmeʿ* 'mosque', *mizīn* 'scale', Blanc 1964:42). A comparison with Cypriot Arabic suggests that this is an old trait of a sedentary Syrian-Mesopotamian dialectal continuum (Borg 1985:156–157).

The *qəltu* dialects differ from the majority of sedentary dialects in that they have retained the interdentals *ṯ*, *ḏ*, and *ḏ̣*. Only in the Christian dialect of Baghdad and in the Anatolian dialect of Diyarbakır have they become postdental stops. In some dialects – mainly in the peripheral Kozluk– Sason subgroup – of the Anatolian group they have developed further to corresponding sibilants (*s*, *z*, *ẓ*: *sawr* 'ox', *zīb* 'wolf', *ḏ̣əhər* 'noon'), in some others – mainly in Siirt – they have become labio-dental fricatives (*f*, *v*, *ṿ*: *fāfe* 'three', *vahab* 'gold', *ṿəhor* 'noon') (Jastrow 1978:34–39). A peculiarity of the Tigris group is the *ġ* reflex of **r*: *ġās* 'head', *ḥāġ* 'hot' (Blanc 1964:20).

4.3 *Greater Syria*

In the Greater Syrian dialect area, the urban dialects distinguish themselves as a group of their own, whereas some rural dialects, for example those spoken in northern Lebanon and in the Damascus Plain, do not essentially differ from them. Some others, for example those spoken in Ḥōrān and central and southern Palestine, are in sharp contrast with the urban dialects. Shared features in the whole area include *b-* as indicative and non-contingency marker in the imperfect, contrasting in function, but probably not in etymology, with the Egyptian present tense marker *b(i)-*. The southern half of the Greater Syrian dialect area, up to Beirut, shares the use of split-morpheme negations with Egyptian and North African dialects, whereas in the northern half, like in the *qəltu* dialects, they are not used (Behnstedt 1997, map 225).

As in many parts of the Arabic-speaking world, in most of Greater Syria the contrast between the urban and the rural dialects has traditionally been noticeable. Using the reflexes of **q* (→ *qāf*) as the most important criterion, Cantineau (1938) divided the sedentary dialects of Greater

Syria into four groups. The first division line goes between the S_1 and the S_2 dialects. Group S_2 comprises the rural dialects spoken in central Palestine as well as the oasis of Suxne in the Syrian Desert. This group uses fronted variants *(k)* of **q*, which has brought about an unconditioned palatalization of **k* to *č*, or the other way round (Behnstedt 1994:8). On the other hand, the S_1-speakers use back reflexes *(q, ʾ)* of **q*: the rural speakers in Syria, southern Lebanon, Galilee, and Jabal Ḥōrān use the [q] reflex, whereas the dialects spoken in Aleppo, Latakia, Hama, Homs, Beirut, Damascus, Saida, Gaza, Jerusalem, and Hebron are 'urban S_1(ʾ) dialects' in which the reflex of **q* is the glottal stop; the Palestinian dialects of Haifa, Safed, Tiberias, and Jaffa belonged to this group. The rural dialects to the north of Damascus and in the northern half of Lebanon are 'rural S_1(ʾ) dialects'. Since the 1930s, the use of the glottal stop reflex of **q* has spread to comprise the whole of Lebanon, the entire area between Damascus and Homs, the southern half of the Syrian coast, and large areas west of Aleppo (Behnstedt 1997, map 9). The same urbanizing development is going on in Palestinian dialects.

Although Cantineau's classification is mainly based on the reflexes of **q*, it actually coincides with a significant number of other prominent isoglosses as well, among them the reflexes of the interdentals (retained in the oasis dialects as well as in Ḥōrān, Jabal Ḥōrān, and rural Palestinian); retention or absence of the *h* in 3rd pers. sg. fem. and 3rd pers. pl. com. suffixed pronouns (retained in the oasis dialects, part of central Syrian rural dialects, Qalamūn dialects, southern Lebanon, southern Syrian, Palestinian); as well as gender distinction in plural forms of finite verbs and personal pronouns (retained in Ḥōrān and rural central and southern Palestinian). The de-affricated *ž* reflex of **j* is an additional S_1(ʾ) feature, used in urban dialects, except in Hama and the Muslim dialect of Aleppo. It occurs in Lebanon and coastal Syrian dialects as well; in rural areas it seems to be a progressive feature (cf. Behnstedt 1997, map 3, and Bergsträsser 1915, map 2).

4.4 *Egypt*

One of the best-known Egyptian Arabic traits is the *g* reflex of **j*, as a matter of fact belonging only to the dialects of Cairo, and the central and northeastern Delta, as well as the dialect of the Fayyūm and Bani Swēf areas, whereas the reflexes in the western and eastern Delta and the whole Nile Valley to the south of Bani Swēf vary between *j* and *d* (Behnstedt and Woidich 1985, maps 10–15). The distribution of the reflexes of **q* is virtually the same: if the reflex of **j* is *g*, the reflex of **q* is the glottal stop, whereas in most of the other dialects it is the Bedouin-type *g*. A trait typical of most Lower Egyptian dialects is the place of the anaptyxis. In them, in contrast to virtually all other sedentary dialects outside the Arabian Peninsula, three-consonant clusters (CCC, #CC, CC#) are broken up so as to form open syllables (maps 51–58). A further salient trait of Lower Egyptian is the Cairene word accent *(bā́ʾaṛa*, *madrása*, *yixbízu)*, contrasting with Upper Egyptian in which the initial syllable is stressed (maps 59–60).

In contrast to Mesopotamian and Syrian dialects, in Cairene and most of the Delta dialects long vowels are shortened in closed syllables: *kātib*, fem. *katba*, pl. *katbīn*, and the monophthongized diphthongs *ē* and *ō* are reduced in closed syllables as well as in pretonic position to *i* and *u*: *bitna*, *yumēn*. In the sg. fem. of the active participle, the vowel of the feminine morpheme is lengthened before suffixed pronouns: *māsik* + *a* > *maska* + *-ha* > *maskāha* (Cairo, Woidich 1980:214).

The short demonstrative pronouns sg. masc. *da*, fem. *di*, pl. com. *dōl* are well-known hallmarks of Egyptian (and Andalusian) Arabic; in Egyptian dialects they are placed after the noun *irrāgil da*, *innās dōl*. The distal demonstrative pronouns in Cairo and central Delta are *dukha*, *dikha*, *dukham*; 'how?' is *izzāy* in all of Lower Egypt; *t* + Form I is used instead of Form VIII in almost the whole country, a feature shared with *maġribī* dialects; the present marker is *bi-* in Lower Egypt, while in Upper Egypt different reflexes of **ʿammāl* are used; split-morpheme negations are used in the whole country; the most common genitive marker is *bitāʿ*, in Upper Egypt *šuġul* and *ihnīn*; 'to give' is *idda* in Lower Egypt except western Delta, both *ʿaṭa* and *idda* occur in Upper Egypt; 'to wish, want' is expressed by the participle *ʿāyiz* in Cairo as well as in the central and western Delta; *kuwayyis* 'good' is an item shared with Greater Syrian dialects. A lexical hallmark of Egyptian Arabic is *issanādi* 'this year'; in *dilwaʾti* 'now' probably an older placement of the demonstrative pronoun is preserved.

A typologically prominent feature of Egyptian Arabic is the word order of interrogative sentences: _tiʿmil ēh_ 'what are you doing?' (Behnstedt and Woidich 1985).

5. The Western dialects

The Western dialects can be divided into two major groups: the so-called pre-Hilālī sedentary dialects and the Bedouin dialects. The former hark back to the first phase of Arab immigration (7th–10th centuries C.E.). The rural dialects of the Jbāla in northern Morocco as well as those spoken around Nedroma in the northwestern corner of Algeria and in the neighborhood of Djidjelli and Collo in northeastern Algeria also belong to this phase. These dialects display considerable substrate influence from Berber languages.

In the 11th century the originally Najdī tribes of Banū Sulaym and Banū Hilāl and the southern Arabian tribe of the Maʿqil moved westward and occupied the North African plains and steppes. At present, Sulaymī Bedouin dialects are spoken in Libya, southern Tunisia, and northeastern Algeria; eastern Hilālī dialects in central Tunisia and eastern Algeria; central Hilālī in central and southern Algeria; northern Hilālī in the northern part of central Algeria; and Maʿqilī dialects in northwestern Algeria and Morocco. The differences between the Bedouin dialects in the whole Western dialect area are relatively slight. In the Maʿqil and northern Hilālī dialects *_j_ > _ž_, _ġ_ is retained, and the 3rd pers. sg. masc. suffix pronoun is _-ah_, whereas the counterparts in the central Hilālī dialects are *_j_ > _ž_, *_ġ_ > _q_, and _-u_ (Grand'Henry 1976:4–5; Fischer and Jastrow 1980:31–38).

As a result of the Bedouin migrations, clear-cut distinctions developed between urban, rural, and Bedouin dialects. The long belt of urban pre-Hilālī dialects begins with the old Tunisian cities of al-Qayrawān, Mahdiya, Sousse, and Tunis. In Algeria it continues with the littoral cities of Skikda, Djidjelli, Dellys, Cherchell, and Ténès, and the interior cities of Constantine, Médéa, Blida, and Miliana. In the westernmost part of Algeria the pre-Hilālī dialects include the dialect of Tlemcen, the old urban center of Orania, surrounded by a wide area of Bedouin dialects, and to the northwest of it, the dialect of Nedroma. In Morocco, old urban dialects are spoken in Old Fes, Rabat, Salé, Taza, Tangier, and Tétouan; these constitute the northern group of urban Moroccan, with the present tense marker _kā-_ as a salient feature, distinguishing the group from the southern urban Moroccan spoken in Marrakesh and New Fes, which have _tā-_. The new cities of Casablanca and Mogador represent Bedouin-type dialects (Fischer and Jastrow 1980:33–35).

The pre-Hilālī dialects of the Maġrib can be divided into an Eastern and a Western branch. The Eastern branch, comprising Libya, Tunisia, and easternmost Algeria, has a more conservative structure, as is apparent from the following phonological traits. The interdental fricatives are retained in all Tunisian dialects except Mahdiya and the Jewish dialects. Inherited short vowels – e.g., in Tunis _a_, _i_, _u_ – are better retained, whereas in the Western group they have been reduced into _a_ vs. _ə_ or only one phoneme, as in Djidjelli. The reflexes of the diphthongs *_aw_ and *_ay_ are _ū_ and _ī_, as in the Western branch, but in Mahdiya they are _ū_ and _ī_, in Sfax _aw_ and _ay_. Also, the women in Tunis and Sousse, as well as the Jews in these towns, have retained the diphthongs (Cohen 1975:65–67; Singer 1980:249–251)

One prominent feature shared by the oldest urban dialects of the area is the glottal stop reflex of *_q_, typical of the urban dialect of Fes, but occurring in Rabat, Tétouan, Tlemcen, and, significantly, also in Maltese. When classified according to the Eastern vs. Western division, Maltese undoubtedly represents the latter. In a contrastive analysis on the basis of 37 isoglosses, Maltese shared 25 with the urban pre-Hilālī _maġribī_ dialects (Vanhove 1998). Its phonological innovations, interesting from the comparative Semitic point of view, include the merger of _x_ and _ḥ_ > _ḥ_, realized as [h], [ħ], or [χ], and of _ġ_ and ʿ > ʿ, and further > Ø, still discernible as pharyngealization of the adjacent vowels (Schabert 1976:45–50).

In Algeria, the old urban dialects of the interior, except the prestigious dialect of Tlemcen, have been influenced by neighboring Bedouin dialects; in Morocco, this is the case with Marrakesh and Meknès. The rural dialects spoken in wide areas adjacent to Djidjelli and Nedroma have exerted a considerable influence on the dialects of these towns; in Morocco, the same development has taken place in Tangier (Fischer and Jastrow 1980:34; Iraqui-Sinaceur 1998:138–139). The population of Algiers, one of the pre-Hilālī urban centers, has during the last few generations grown too heterogeneous to render it meaningful to speak about its dialect any more (Boucherit 2002:24–25). In Libya, the most closely sedentary-type dialect is that spoken in Tripoli, which can be characterized as a Bedouinized former urban dialect. The few extant features of the pre-Hilālī urban type include the postdental stop reflexes of the interdentals and the reflexes of the verb **raʾā* ‘to see’ (Fischer and Jastrow 1980:36).

6. CLASSIFICATION ACCORDING TO RELIGIOUS AFFILIATION

In many Arab cities, religion correlates more or less closely with dialect. One of the most noticeable cases is the situation in the Mesopotamian dialects. The dialect of the Christians of Baghdad differs from that of the Muslims in several points, among them the following: the interdental fricatives have become postdental stops vs. retained interdentals; **r* > *ġ* vs. retained *r*; retained *q* vs. *g*; retained *k* vs. *k* + phonetically-conditioned *č*; use of the present markers *qad-*, *qa-* vs. *gāʿed*, *da-*; use of the optative marker *da-* + 1st pers., e.g., *daqūl* ‘let me say’, *danqūl* ‘let us say’ vs. *xaldangūl* (Blanc 1964:20, 25–26, 115–118; Abu-Haidar 1991:7–9, 88–89). The Jewish dialect, which until the beginning of the 1950s was spoken by a significant number of the population of Baghdad, was to a high degree identical with that spoken by the Christians. Salient Jewish Arabic features were, e.g., retained interdentals, a strong *ʾimāla*, and the future marker *(has)sa-* (Blanc 1964:42–43,117–118; Abu-Haidar 1991:29).

In Mosul, where the whole population irrespective of religious affiliation speaks *qəltu* dialects, the differences are minimal when compared with Baghdad (Jastrow 2004:141–142). In Aleppo differences are also found between the dialects of the Muslims and the Christians. At the beginning of the 20th century there still were different Christian dialects in different quarters, but since then these divisions have blurred (Behnstedt 1989:43–44). Salient Aleppine Christian traits include the retention of the diphthongs *aw* and *ay*; the use of the glottal stop reflex of **q* without the back allophone of *a* and *ā* in juxtaposition to it; and *ʾimāla* in many cases in which the dialect of the Muslims does not have it (Behnstedt 1989:45–63).

The division between two different dialect types in Bahrain is parallel with the earlier development in Lower Iraq. Although representing the Shiʿi–Sunni split, it is in fact a result of two phases of settlement: the Shiʿi population speak the old rural Baḥārna dialect, which displays typical sedentary devices, whereas the Sunni newcomers speak a dialect of the ʿAnazī Bedouin type (Holes 1995:272–273).

In North Africa the Jewish Arabic dialects are of urban type and represent the first phase of Arab settlement. Their phonology is markedly urban: the reflex of **q* is the glottal stop (Algiers); the old interdental fricatives have become postdental stops; *h* has disappeared; *š* and *s*, as well as *ž* and *z* have merged; short vowels have been highly reduced (Marçais 1977: 9–11). In Oran and some towns in the region of Algiers, the Jewish dialects represent the sedentary, and the Muslim dialects the Bedouin type. As pointed out by Blanc (1964: 16), the parallel with the distribution of the *qəltu* vs. *gilit* dialects in Lower Iraq is striking. Here, as in all other cases of dialect differences along the lines of religious affiliation, the differences – besides religious-cultural technical terms – can mainly be attributed to settlement history.

7. CLASSIFICATION OF DIALECTS ON THE EASTERN–WESTERN BOUNDARY

7.1 *The oases of al-Baḥariyya and al-Farāfira*

The dialects of the Egyptian oases of al-Baḥariyya (B) and al-Farāfira (F) are illustrative examples of classification problems. They display several isoglosses of the *maġribī* type, among them the aspiration of *t* [t^s] < **t* and **ṯ* (F), the neutralization of phonological contrasts between *s*/*š* and *z*/*ž* (F), and the paradigmatic leveling *niktib*/*niktibu* (B and F). However, they exhibit many important features of the Egyptian type as well, among them the syllable structure (F); the → ‘*bukaṛa* syndrome’; absence of *hā-* in demonstrative pronouns and placing them after the noun; and the inflection of the verbs *kal* ‘to eat’ and *xad* ‘to take’ (Woidich 1993:343–347). It is therefore obvious that the *aktib*/*niktib* vs. *niktib*/*niktibu* isogloss alone cannot be used as a categorical criterion for grouping these dialects together with the *maġribī* dialects. In a strictly synchronic classification two alternative solutions may be applied: these dialects might be defined as part of a transitional area between the

Egyptian and the *maġribī* dialects, or the question of their belonging to either of them might be solved with reference to the classificatory weight of different isoglosses. However, no satisfactory theory has as yet been created which would give adequate tools for measurement. But as soon as the question is asked, whether these oasis dialects basically belong to the sphere of Egyptian dialects displaying adstrate features of *maġribī* type, or vice versa, diachronic and extralinguistic criteria will be involved. Since there is a gap of one thousand years in our knowledge of the history of the oases and of the dialects spoken in them, different conclusions can be drawn. Woidich regards the dialects of the two oases as isolated and peripheral dialects belonging to the greater Egyptian dialect area, with greatest resemblance to the dialects spoken in Fayyūm and the province of Banī Swēf, while the Western traits are best explained as results of dialect contact (1993:355–356).

Behnstedt (1998:88–92), however, points out that the short demonstrative pronouns and the forms of the verbs *kal* and *xad* are well attested Western forms from al-Andalus, and also the syllable structure in al-Farāfira can be interpreted as retention of a very conservative *maġribī* feature, known from the dialect of al-Andalus. One may also ask why the contrast *žawž* vs. *itnēn* 'two' is not considered as having the same classificatory weight as a morphological contrast, which Woidich considers as having greater classificatory weight. According to Behnstedt, the first Arab immigrants to the oases may very well have been *maġribī* tribes, perhaps speaking a dialect resembling the Andalusian type. Moreover, the *maġribī* traits cannot be solely attributed to the influence of neighboring Sulaymī Libyan tribes, at least not the aspiration of *t and *ṯ.

7.2 *The Chad region*

Besides westward, the spread of Arabic also continued from Egypt southward along the Nile Valley to the Sudan – which was Arabicized also directly from the Ḥijāz – and from there westward to the Lake Chad region. Arabic arrived there from southern Egypt in the 14th century at the latest. In the question of the division of Arabic into Western and Eastern groups, this region is of interest because immigrants from east and west may have met here. This may be reflected by the occurrence of both the Eastern *(b)aktub-naktub* and the Western *baktub-naktubu* imperfect patterns. However, since one and the same speaker will vary across the different paradigms, they cannot be regarded as two isoglosses but rather as variants of a single variable (Owens 1995:323, 330).

8. DE-BEDOUINIZATION, SEDENTARIZATION, AND BEDOUINIZATION DEVELOPMENTS

The dialects spoken in the Arabian Peninsula, except its southwestern parts, are Bedouin or former Bedouin dialects. In sedentary environments the Bedouin dialects tend to adopt reductional and innovative traits, plausibly as results of increased dialect contact. The Meccan dialect, for instance, while displaying several Bedouin-type features such as the *g* reflex of **q* and the productive use of Form IV, at the same time exhibits many traits typical of sedentary dialects, among them the following: absence of interdental fricatives; absence of gender distinction in plural forms of finite verbs and personal pronouns; optional use of the present continuous marker *bi-*; and frequent use of the genitive exponent *ḥagg* (Schreiber 1970:6 and passim; Ingham 1971:273 and passim).

In a corresponding way, the Arabic dialects of the Gulf area, which are a relatively recent offshoot of the ʿAnazī dialect group, during only two centuries of sedentarization have adopted considerable reductional changes, such as elimination of the internal passive, the indefinite marker *-in*, and, mostly, gender distinction in plural forms of finite verbs and personal pronouns. An innovation typical of sedentary dialects is the future and volitive marker *bi-*. These dialects have thus drawn away from their original central Arabian ʿAnazī dialects (Ingham 1982:33; Holes 2001:xviii). In maritime environments, the vocabulary naturally differs noticeably from the one mirroring the traditional nomadic culture. Even the rhythm and intonation patterns are at the present time quite different from the dialects of the ʿAnazī type (Johnstone 1967:18).

The Mesopotamian *gilit* dialects exhibit similar developments in the reductional and innovatory direction. The rural *gilit* dialects have still preserved several prominent features of Bedouin type, contrasting with the urban *gilit* dialects. These include the phonetically-conditioned affrication of the reflex of **q*; the → *gahawa* syndrome; retention of gender distinction in plural forms of finite verbs and personal pronouns; and infrequent use of the verb modifier *gāʿid/jāʿid*. In the Muslim dialect of Baghdad the sedentary-type development has advanced much more, obviously under influence from the old sedentary *qəltu* dialects. Typical sedentary traits, such as the use of the verb modifiers *gāʿed*, *da-*, and *rāḥ/raḥ* with the imperfect, *d(e)-* with the imperative, have been adopted. An important → Aramaic substrate feature has also found its way to the Muslim dialect of Baghdad through the *qəltu* dialects, viz. marking the definite direct object by affixing to the verb a suffixed pronoun referring to the object and introducing the object epexegetically by the preposition *l-*, e.g., *bāʿa lilbēt* 'he sold the house' (Blanc 1964:128–130).

During the Ottoman period, in particular, new Bedouin tribes settled down in the neighborhood of towns and villages lying near the fringes of the Syrian Desert. One of the results was a progressive Bedouinization of the old sedentary dialects in these areas. Examples of this development are the *qəltu* dialects of the Euphrates group and the few sedentary dialects spoken to the east of the Jordan (Blanc 1964:26–27; Jastrow 1978:25–26; Palva 1994:468–469).

9. Isolated dialects

→ Cypriot Arabic, the dialect spoken by a few thousand Maronite Christians in Cyprus, synchronically to be classified as an isolated dialect displaying considerable superstrate influence from Cypriot Greek, can historically be classified as representing the old sedentary dialects of the Fertile Crescent. According to Borg's definition (1985:157), Cypriot Arabic "represents a now superstratally modified variety of a dialectal prototype antedating the present areal configuration obtaining among Arabic-speaking sedentaries in this region". A prominent typological feature which Cypriot Arabic shares with northern Syrian dialects is the vocalically conditioned *ʾimāla* (Borg 1985:156–157; Behnstedt 1997, maps 43–62). Cypriot Arabic also shares a number of salient traits with the southeastern branch of the Anatolian *qəltu* dialects, among them, *-n* in the suffixed personal pronoun of the 2nd and 3rd pers. pl. com.; use of copulas derived from independent personal pronouns; use of reflexes of **ḥattā* as a verb modifier marking the future tense; use of genitive exponents derived from **ḏayl*; and dropping of *h* in personal and demonstrative pronouns. Important affinities to the present-day sedentary dialects of Greater Syria include the *b(i)-* non-contingency marker in the inflection of the imperfect; retention of reflexes of *tāʾ marbūṭa* in numerical constructs; reflexes of **hunnā*, **-kun*, **-hun*; and a genitive marker harking back to *šayʾat-* (Borg 1985:154–155). An interesting trait is the partial retention of **raʾā*, which attests its use in Syrian Arabic during the first Islamic centuries.

Bibliographical references

Abboud, Peter F. 1979. "The verb in northern Najdi Arabic". *Bulletin of the School of Oriental and African Studies* 42.467–499.

Abu-Haidar, Farida. 1991. *Christian Arabic of Baghdad*. Wiesbaden: O. Harrassowitz.

Aguadé, Jordi, Patrice Cressier, and Ángeles Vicente (eds.). 1998. *Peuplement et arabisation au Maghreb occidental: Dialectologie et histoire*. Madrid and Saragossa: Casa de Velázquez.

Behnstedt, Peter. 1985. *Die nordjemenitischen Dialekte*. I. *Atlas*. Wiesbaden: L. Reichert.

——. 1989. "Christlich-aleppinische Texte". *Zeitschrift für Arabische Linguistik* 20.43–96.

——. 1994. *Der arabische Dialekt von Soukhne (Syrien)*. II. *Phonologie, Morphologie, Syntax*. III. *Glossar*. Wiesbaden: O. Harrassowitz.

——. 1997. *Sprachatlas von Syrien: Kartenband: Beiheft*. Wiesbaden: O. Harrassowitz.

——. 1998. "La frontière orientale des parlers maghrébins en Egypte". Aguadé a.o. (1998:85–93).

—— and Manfred Woidich. 1983. "Ägypten: Arabische Dialekte". *Tübinger Atlas des Vorderen Orients*, AVIII-12. Wiesbaden: L. Reichert.

——. 1985. *Die ägyptisch-arabischen Dialekte*. II. *Dialektatlas von Ägypten*. Wiesbaden: L. Reichert.

Blanc, Haim. 1964. *Communal dialects in Baghdad*. Cambridge, Mass.: Harvard University.

Borg, Alexander. 1985. *Cypriot Arabic*. Wiesbaden: F. Steiner.

Boucherit, Aziza. 2002. *L'arabe parlé à Alger: Aspects sociolinguistiques et énonciatifs*. Paris and Louvain: Peeters.

Cantineau, Jean. 1938. "Remarques sur les parlers des sédentaires syro-libano-palestiniens". *Bulletin de la Société de Linguistique de Paris* 40.80–88.

——. 1940–1946. *Les parlers arabes du Ḥōrān*. Atlas. Paris: Klincksieck.

Cohen, David. 1970. *Études de linguistique sémitique et arabe*. Paris: Mouton.

——. 1975. *Le parler arabe des juifs de Tunis*. II. *Étude linguistique*. The Hague and Paris: Mouton.

Fischer, Wolfdietrich and Otto Jastrow (eds.). 1980. *Handbuch der arabischen Dialekte*. Wiesbaden: O. Harrassowitz.

Grand'Henry, Jacques. 1976. *Les parlers arabes de la région du Mzāb* (*Sahara algérien*). Leiden: E.J. Brill.

Holes, Clive. 1995. "Community, dialect and urbanization in the Arabic-speaking Middle East".

Bulletin of the School of Oriental and African Studies 58. 270–287.

——. 2001. *Dialect, culture, and society in Eastern Arabia*. I. *Glossary*. Leiden: E.J. Brill.

Ingham, Bruce. 1971. "Some characteristics of Meccan speech". *Bulletin of the School of Oriental and African Studies* 34. 273–297.

——. 1982. *North east Arabian dialects*. London and Boston: Kegan Paul International.

——. 1994. *Najdi Arabic: Central Arabian*. Amsterdam and Philadelphia: J. Benjamins.

Iraqui-Sinaceur, Zakia. 1998. "Le dialecte de Tanger". Aguadé a.o. (1998:131–140).

Jastrow, Otto. 1978. *Die mesopotamisch-arabischen* qəltu-*Dialekte*. I. *Phonologie und Morphologie*. Wiesbaden: F. Steiner.

——. 2004. "Jüdisches, christliches und muslimisches Arabisch in Mossul". *Approaches to Arabic dialects: A collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday*, ed. Martine Haak, Rudolf de Jong, and Kees Versteegh, 135–150. Leiden: E.J. Brill.

Johnstone, T.M. 1967. *Eastern Arabian dialect studies*. London: Oxford University Press.

Jong, Rudolf de. 2000. *A grammar of the Bedouin dialects of the northern Sinai littoral, bridging the linguistic gap between the eastern and western Arab world*. Leiden: E.J. Brill.

Marçais, Philippe. 1958. *Le parler arabe de Djidjelli (Nord constantinois, Algérie)*. Paris: Adrian Maisonneuve.

——. 1977. *Esquisse grammaticale de l'arabe maghrébin*. Paris: Adrian Maisonneuve.

Owens, Jonathan. 1995. "Minority languages and urban norms: A case study." *Linguistics* 33.305–358.

Palva, Heikki. 1994. "Bedouin and sedentary elements in the dialect of es-Salṭ: Diachronic notes on the sociolinguistic development". *Actes des premières Journées internationales de dialectologie arabe de Paris*, ed. Dominique Caubet and Martine Vanhove, 459–469. Paris: INALCO.

Schabert, Peter. 1976. *Laut- und Formenlehre des Maltesischen anhand zweier Mundarten*. Erlangen: Palm und Enke.

Schreiber, Giselher. 1970. *Der arabische Dialekt von Mekka: Abriß der Grammatik mit Texten und Glossar*. Ph.D. diss., University of Münster.

Singer, Hans-Rudolf. 1980. "Das Westarabische oder Maghribinische". Fischer and Jastrow (1980: 249–276).

——. 1984. *Grammatik der arabischen Mundart der Medina von Tunis*. Berlin and New York: W. de Gruyter.

Vanhove, Martine. 1998. "De quelques traits préhilaliens en maltais". Aguadé a.o. (1998:97–108).

Woidich, Manfred. 1980. "Das Ägyptisch-Arabische". Fischer and Jastrow (1980:207–248).

——. 1993. "Die Dialekte der ägyptischen Oasen: westliches oder östliches Arabisch?" *Zeitschrift für Arabische Linguistik* 25:340–359.

——. 1996. "Rural dialects of Egyptian Arabic: An overview". *Centre d'études et de documentation économique, juridique et sociale: Égypte/Monde arabe* 27–28.325–354.

Heikki Palva (University of Helsinki)

Dialects: Genesis

1. Introduction

Scholarly discussions about the genesis of the Arabic dialects have always been connected by another crucial issue: the linguistic situation in Arabia before Islam and its relation to the linguistic variety of the *Qurʾān* and poems attributed to pre-Islamic poets. Yet, there are very few studies devoted exclusively to dialects in the pre-Islamic period (Rabin 1951; Owens, forthcoming 2).

Arabic dialects appeared after the expansion of the Arabs, which began after the death of the Prophet Muḥammad in 632 C.E. All colloquial varieties or dialects of Arabic are generically called 'New Arabic' or 'Neo-Arabic', comprising the spoken varieties that evolved outside Arabia and those which progressed and changed within that region. These include modern dialects as well as those Arabic dialects which disappeared for historical reasons and can only be studied through documents, like the Andalusian and Sicilian varieties, or those which served as basis for other dialects and disappeared as autonomous languages, as happened in Malta and Iran.

Two issues need to be addressed here. The first is the nature of the spoken language exported by Arabs to conquered lands, because this determined the type of Arabic which developed afterwards and its effect on the genesis of dialects. The second is the new reality after the expansion, in which two linguistic elements were to play a fundamental role: on the one hand, the conquerors' language and its evolution, and on the other hand, the conquered populations and their linguistic reactions.

The term 'linguistic variety' is used here as a neutral linguistic term to be applied to any particular kind of language considered as a single entity. It designates a set of linguistic items with similar social distribution and permits a much wider use than the lay term 'language' (Chambers and Trudgill 1980:5; Hudson 1996:22–24). This is why the term 'pre-Islamic poetic linguistic variety' will be applied to the language used in this poetry, and 'normalized variety' to the standard language.

→ 'Middle Arabic' is a confusing term (Larcher 2001), traditionally used for "the mixed language of medieval texts" (Blau 1988b;

1999:223; see also Fück 1955:87; Fischer 1991), and in need of further specification (see Abboud-Haggar 2003:73–75). It is used here to designate the early medieval (between the 7th and the 10th century) written variety, whose main feature is the interference or mixing between normalized Arabic and non-normalized Arabic. Normalized Arabic refers to what is known as → Classical Arabic, while non-normalized refers to Arabic with grammatical errors or mixed to a certain degree with dialectal features, depending on each individual speaker.

The term 'New Arabic' or 'Neo-Arabic' is used in the same way that Neo-Latin is used to refer to varieties related to Classical Latin (Meillet 1975:310–322; Coseriu 1990:25). It is also used in opposition to → Old Arabic, the pre-Islamic Arabic – a use which should be revised (Versteegh 1997:98). Neo-Arabic is applied to all post-Islamic colloquial varieties – medieval or modern, disappeared or living, inside or outside the Arabian Peninsula – which emerged after the expansion of the Arabs (for their classification, see Abboud-Haggar 2003:84–89, and → Dialects: Classification).

The term → 'Pre-Islamic Arabic' or → 'Old Arabic' is used to refer to all varieties known and used before Islam, a historical epoch fixed approximately during the century and a half before the revelation of the *Qurʾān*. The term comprises all linguistic varieties used then: colloquial, poetic, and commercial, varieties used in religious services, etc.

2. Nature of the exported language

Most scholars concerned with the issue of pre-Islamic Arabic agree that Arab tribes, whether Bedouin nomads dwelling in the Arabian desert or settled in sedentarized nuclei (Retsö 2003: 113–116), spoke colloquial linguistic varieties, which to some extent differed from the variety used in poetry and the *Qurʾān*, and also from one another. The best evidence is that Arab grammarians admit that the literary idiom as used in the *Qurʾān* contained elements of different dialects (Fleisch 1974:13–14; ʾAnīs 1965:43; Rabin, quoting Ibn Jinnī, 1955:19–23, 1951:17; al-Fatlī 1988; Owens 1988:19–21, 1990:5–6).

Arabic sources leave no doubt, indeed, about the existence of many colloquial Bedouin varieties in the period before Islam, although in an unsystematic and often contradictory way (Rabin 1955:1). Direct accounts of phonetic, phonological, morphological, and syntactic particularities provide a rich comparison with the 8th-century normalized classical variety (Owens 1988:4). Thanks to these sources we know, for instance, about → *kaškaša* (Banū Rabīʿa pronouncing /k/ [k] at the end of the word in *waqf* position with *ʾiskān* /š/ [ʃ]); *ʿajʿaja* (Quḍāʿa tribes, pronouncing final *yāʾ* as [ž] instead of [j]) and vice-versa, Tamīm tribes, pronouncing as [j] instead of [ž]; permanence or loss of *hamza*; → *ʾimāla* (pronouncing [a] > [e] > [i]); and → *taltala* (pronouncing the normalized *fatḥa* of the imperfect as [e]), among numerous other phenomena, many of which are still in use in current colloquial varieties. There are also accounts of internal vocalization of words and lexical usage as homonyms, synonyms, and antonyms. This information permitted the division of the old colloquial varieties into Eastern and Western varieties (Vollers 1906:4–23; Sarauw 1908; Rabin 1955:1–13; Corriente 1976:75; aṭ-Ṭaʿʿān 1978:146–152, 171–202, 153–161; ʿAbd at-Tawwāb 1987:108–115, 229–356; Ḍayf 1988: 121–130 quoting Suyūṭī, *Muzhir*; al-Fatlī 1988; Ghul 1993:11–15; Versteegh 1997:41–46; Fischer 1995:75; Abboud-Haggar 2003:55–57).

There are no explicit or direct data about declension or *ʾiʿrāb* – a fundamental trait of the synthetic character of Arabic – as characteristic of any colloquial Bedouin variety, whether Eastern or Western, nomad or sedentarized, in the pre-Islamic period. Besides, no precise information can be extracted from the sources about pause or *waqf* (pronouncing the last consonant without vowel or *ʾiskān*, i.e. without declension or mood-ending; ʾAnīs 1965: 82–84, 145–152; Fleisch 1974:23). This lack of information about declension opened the door to all kinds of speculation, mostly because declension was – and still is – considered the first sign of the highest level of literacy, directly related to the Qurʾānic → poetic koine, the basis of the normalized literary Arabic language, 'Classical Arabic' (Versteegh 1983: 170–171; Hopkins 1984:155, n. 1). Declension, mood endings, and the analytical character of the spoken varieties are fundamental in determining the nature of pre-Islamic colloquials.

In the Muslim tradition, and among most Arab scholars, the colloquial variety of Qurayš, the Prophet's tribe, with all its particularities, including declension, was the basis of the poetic-

Qurʾānic koine (Ḍayf 1988:131–137; ʾAnīs 1965:39–40). Consequently, the literary language was identified with the Qurašī colloquial variety, the linguistic variety in which the *Qurʾān* was revealed (Ḥusayn 1927:103–106; Rabin 1951:24, 1955:21; Zwettler 1978:117). This led to the assumption that all Bedouin colloquial varieties formed a homogeneous group among which one variety excelled – the Qurašī dialect – and that this group never lost the declension until the post-Islamic era, when the *muwalladūn* (converted non-Arabs who learnt Arabic for social, economic, or simply religious reasons) 'corrupted' the Arabic language, *fasād al-luġa* (al-Fatlī 1988:95).

The position of Arab and Muslim scholars was followed by many European linguists. Fück (1955) – preceded by Nöldeke (1904), followed mainly by Ferguson (1959), Blau (1981), and Versteegh (1984) – based his reasoning on the assumption that the *ʿarabiyya* – as he called Classical Arabic – of poetry and the *Qurʾān* was essentially the Meccan language variety and identical to Bedouin vernaculars, thus with declension "[which was] in full flower in Bedouin speech" (Fück 1955:3; Blau 1981:2–3; Fischer and Jastrow 1980:15–19; Versteegh 1984:4–5; see also Zwettler 1978:133–134, 151–152). But Fück and those who followed him did admit that there were some differences between Meccan speech and Bedouin dialects and between the latter themselves (Fück 1955:3; Versteegh 1984:22–23; Fischer 1995:76).

On the other extreme stands Vollers (1906), followed by Kahle (in Zwettler 1978:129–130), ʾAnīs (1966), Diem (1973), Kaye (1972), Zwettler (1978), Bakalla (1984:85), and Retsö (1992). He more or less rejected the existence of declension as an element of spoken vernacular: if it ever existed, it was only among eastern Bedouin tribes, whereas in the more sedentarized west, declension was retained simply as an element of more elevated discourse, borrowed from Bedouin for poetic and other highly formalized purposes, an element which according to Zwettler (1978:170–172) was fossilized for prosodic purposes. According to these scholars, in their everyday speech Bedouin only spoke a colloquial variety lacking declension and nunation (*tanwīn*), as pausal or apocopated forms with *ʾiskān* (*taskīn*) became generalized. They believe that both phenomena are nothing but archaisms, as early as the time of Muḥammad himself and his companions (Vollers 1906; Zwettler 1978:118–130).

The general view is that Vollers went too far in his thesis when he asserted that Muḥammad transmitted the revelations in his vernacular variety, and that they were later transferred into the *ʿarabiyya* (Vollers 1906:80–83; Rabin 1951:23). For this reason his basic thesis, as expressed in his *Volkssprache und Schriftsprache im alten Arabien*, lost support, although it was to constitute an unavoidable reference in all later studies about declension in pre-Islamic colloquials (Zwettler 1978:145, 170–171).

Between Fück's and Vollers's position stand other scholars who can be divided into two groups: on the one hand, the Orientalist scholars of the early 1950s, who worked mainly within the German Neogrammarian framework. They were aware of the importance of spoken dialects for the study of the evolution of languages – in fact this is what they studied at the level of the Semitic languages, including Arabic – but did not reach a conclusive opinion about the role of *ʾiʿrāb* in the pre-Islamic linguistic setting. None of them explicitly broached the question of whether *ʾiʿrāb* was unique to the poetic koine, or whether it was shared by some or all spoken dialects (Zwettler 1978:131; see also Corriente's critical remarks, 1975:45–46), their interests being mainly oriented toward Classical Arabic and the Qur'ānic linguistic variety. The lack of decisive documents made the issue an eternal, but inconclusive one. Among these scholars, mention must be made of Fleisch (1974:17), who believed that the poetic koine was based on one colloquial variety, which was, in an unspecified period, considered to be the "langue artistique, langue commune de la poésie". Geyer accepted the poetic *ʿarabiyya* more or less as the naturally spoken language of Bedouin, whose vernacular was practically identical with this koine, with declension as its main feature (Zwettler 1978:118). Rabin did not give a definitive opinion about *ʾiʿrāb*, as he found no clear evidence about it, although he was more inclined to accept the existence of *ʾiʿrāb*, inasmuch as Muḥammad himself was conscious of using the Bedouin *ʿarabiyya*, with its case endings (Rabin 1955:26–27; 1951:81).

On the other hand, in the 1970s Corriente's investigations showed that, before the standardization of language by grammarians, the *ʾiʿrāb*, especially in prose texts, was largely irrelevant for the functional yield of the language. In his view, even the pre-Islamic poetic koine was less synthetic than commonly thought (1973–1974). Following Jespersen's (1925) theories on language – according to which the natural and universal tendency toward simplification is to fix a word order, thus paving the way for a complete loss of case endings, already weakened by purely phonetic laws and analogy (Jespersen 1964:268–271, 361–364) – Corriente studied al-ʾIṣfahānī's *Kitāb al-ʾaġānī* and showed that the disappearance of *ʾiʿrāb* declension vowels (*ʾiskān*) was actually the natural tendency of a drift that is observed in all Semitic languages, including Arabic. He also showed that *ʾiʿrāb* was in fact in use as a vernacular colloquial register, and posited the use of case endings for some areas at least, but at a rather low functional yield. In his view, two varieties coexisted, one with declension, but with less strict rules than those of the normalized variety, and the other without it. Later on, the variety with declension, Classical Arabic, was normalized by Muslim grammarians on the basis of an already normalized poetic koine (Corriente 1971–1972:29, 41–50, 1973–1974:157, 1976:71–72, 89).

This view was supported by Zwettler, who went even further by saying that any spoken variety with a sort of *ʾiʿrāb* was the special speech-form routinely mastered by poets and *rāwī*s, which could be occasionally employed for poetic purposes. Corriente's thesis supplemented Zwettler's: *ʾiʿrāb* in poetry is the most prominent archaic feature of the language and has nothing to do with spoken language (Corriente 1971–1972:40; Zwettler 1978:148).

Regarding stress and pausal form, Corriente showed that pre-Islamic Arabic, even if considered as a whole, had a combinatory phonetics – assimilation, dissimilation, backformation, haplology, and aphaeresis – and that phonological Arabic terms like *taskīn* (elimination of post-tonic vowel *i/u)*, *ijtizāʾ* (shortening and eventual elimination of unstressed originally long vowels), and *ʾišbāʾ* (lengthening of a short vowel) refer to the degree of intensity of stress. For Corriente, it was almost certain that the pre-Islamic dialects were stressed and had concomitant segmental phenomena, with shifts in pausal forms as in other Semitic languages. These allomorphic solutions – which go against Birkeland's thesis that pre-Islamic Arabic had no stress at all, and that modern dialects do not derive from pre-Islamic Arabic (Birkeland reviewed by Ferguson 1997a, 1997b; Kaye and Rosenhouse 1997:280) – establish a definite link between old dialects and Neo-Arabic post-Islamic medieval and modern dialects (Corriente 1976:78–87).

Pre-Islamic colloquial varieties of Arabic, as part of the → Semitic languages – its 'linguistic stock' (Sapir 1949:153) – pertain to the West Semitic group, most probably the South Semitic subgroup as Diem (1980) and Blau (1978: 29–35) advocate, among others (Faber 1997:12; ʿAbd at-Tawwāb 1987:25–36). They were exposed to the group's natural drift – when language moves with time in a current of its own making (Sapir 1949:150–151) – and variations were assimilated by speakers, becoming, with time, part of its structure.

The results of this drift are reflected in some features of the exported colloquial varieties: evolution of defective verbs, a tendency which abutted the contraction of the diphthong *ay* > *ē* > *ā*; maintenance of initial *hamza* in *ʾafʿal* forms in spite of the natural tendency of the *hamza* toward disappearance – the loss of *hamza*, among the Qurayš and Tamīm tribes, is a development shared with Canaanite (Rabin 1955:65, 83; aṭ-Ṭaʿʿān 1978:194–195); and the syntactic phenomenon of natural agreement between subject and verb, known as *ʾakalūnī al-barāġīṯ* (ʿAbd at-Tawwāb 1980:2–14). Other features were also due to Semitic linguistic drift: disappearance of *ʾiʿrāb* declensional vowels because of *ʾiskān* or apocopation (Corriente 1976:95–96; aṭ-Ṭaʿʿān 1978:40–45, 189–190; Fück 1955: 91); unconditioned change of /p/ to /f/; appearance of broken plurals formed by prefixation and/or internal change rather than by suffixation; the existence of verb stems – Forms III and VI *fāʿala* and *tafāʿala* – with a long first vowel (Diem 1980:68–71; Faber 1997:12–13); and the treatment of dual and → pseudo-dual, more developed in Arabic than in any other Semitic language (Retsö 1995:190).

Apart from the general Semitic drift, the *Qurʾān* itself and grammatical treatises show a pre-Islamic language in evolution with many variants. For instance, Forms V *tafaʿʿala* and VI *tafāʿala* alternated with *itfaʿʿal* and *itfāʿal* forms, and Form VII *infaʿala* was used as an alternative for the internal passive, as in modern dialects (ʿAbd at-Tawwāb 1980:14–17, 1987:47–49; Retsö 1983; Abboud-Haggar 2003:149–150, 235). Sībawayhi's *Kitāb* examines the 'accidents' (*ʾaʿrāḍ*), i.e. the modifications in spoken language of the theoretically expected form, mentioning a series of phonological, morphological, and syntactic variants (Blanc 1979:158–161, 172–173).

Pre-Islamic colloquial varieties also show elements of accommodation between dialects as a result of contacts between speakers of different dialects of the same language, as well as adaptation of loanwords to the Arabic language, mainly from Persian. These elements, like the use of *hamza* and of the verb *raʾā* 'to see', are present in many sources, including the *Qurʾān* (ʿAbd at-Tawwāb 1980:17–20; Blanc 1979:159–160, 162–163; García Yebra 1982:333–341).

Although research on this topic has not reached any definitive conclusion, scholars have pointed out the effects of contact between Arabs of the south, the Qaḥṭān descendants or Yemenis (who were in contact with another Semitic language, South-Arabian), and Arabs of the north, the ʿAdnān descendants, the Arabs *par excellence*. This contact was possible in the pre-Islamic period because of migration from north to south, when the south was prosperous, and vice versa, when the south declined and commerce attracted southern Arabs to the north (Robin 1992:71–88; Ḍayf 1988:55–66; al-Mayyāḥ 2003:67–79; Retsö 2003:34–40). Linguistic effects of the contact between these two groups are to be taken into consideration when speaking about the nature of the exported colloquial varieties because they are reflected in Neo-Arabic features (aṭ-Ṭaʿʿān 1978:203–232; Ghul 1993; Corriente 1992:26–27; Diem 1979:18–41; Ingham 1982:26–32).

On the eve of the Arab expansion, a supratribal linguistic variety used in poetry – and then in the revealed text of the *Qurʾān* – was part of the linguistic reality in the Arabian Peninsula in pre-Islamic times (at-Tikrītī 2003:104). There is no general agreement about the origin of that variety, called → poetic koine or Qurʾāno-poetic koine (Versteegh 1993:66). If we were to follow Ṭāhā Ḥusayn's thesis (1927), we would have to say that most pre-Islamic *Jāhiliyya* poetry is mainly an artificial post-Islamic fabrication of grammarians and is not at all representative of pre-Islamic Arabic. The only authentic source for knowing the linguistic situation of that period is the Qurʾānic text, revealed and conserved in the Qurayš linguistic variety (Ḥusayn 1927:92–111; Hopkins 1984: xxxviii–xi).

It should be mentioned that the term 'koine' – traditionally used to designate the linguistic poetic variety elevated to the highest rank, as the linguistic variety of the *Qurʾān*, normalized by grammarians and considered the literary language of Arabia, the *ʿarabiyya*, Classical Arabic – is in itself a problematic linguistic issue. Many Orientalists have questioned the accuracy of this notion in the case of Arabic (Rabin 1955:24; Cohen 1962:119–120).

Whether it was originally derived from an existing tribal dialect, that of Qurayš, or whether it was from the start a formal or poetic variety separate from all the dialects, most scholars seem to concur that the poetic-Qurʾānic koine never formed the spoken vernacular of any Arabic-speaking group, either before or after Muḥammad. The koine, "a language of poets, made for the poets and comprehended by themselves above all" (Brockelmann, cited in Zwettler 1978:101), had never been the colloquial mother tongue of anybody in Arabia and had to be acquired like a foreign idiom (Rabin 1951:17, 1955:19–23; ʾAnīs 1965:84).

3. Genesis and formation of Neo-Arabic

The formation of Neo-Arabic has to be considered in two areas, the Arabian Peninsula, cradle of the Arabic language, and the area outside it. Corriente's thesis from the 1970s advocates a common commercial urban spoken Arabic variety – labeled 'commercial koine' – which began before Islam and continued its natural evolution after the expansion. Probably a product of linguistic contact between pre-Islamic Arab tribes settled in the vicinity of Aramaean population, the *Nabaṭ* or *ʾAnbāṭ*, this koine developed in the *nabaṭī* small trade settlements in north Arabia, the Syrian desert, down to the *sawād al-ʿIrāq* and the area between Kūfa and Baṣra in Mesopotamia (Retsö 2003:375–382). Its speakers lived in cities – which is why it is also described as an urban or *ḥaḍarī* variety – and were systematically dubbed *ʿulūj* 'persons unable to speak pure Arabic'. This commercial spoken koine, attested in pre-Islamic documents, arose with almost complete loss of declension. Diem, through his studies on the Nabatean area, situated the loss of case ending as early as the late 2nd century C.E. (1973:234). This koine is considered to have been the forerunner of current spoken varieties. According to Corriente, this spoken variety, due to its commercial importance, and its business contact with the west Ḥijāzī area, began expanding and 'contaminating' the spoken varieties of those areas. Meanwhile, the east Arabian population resisted this expansion for some time, but finally adopted it (Corriente 1971–1972, notes 4 and 8, 1973–1974:63, 1976:75, 87–89). A similar conclusion was reached by Janssens (1972) who argued on the basis of Ferguson's (1959) koine material, that a spoken koine originated in the Mediterranean area before Islam – the *nabaṭī* area – and expanded slowly among the sedentarized Bedouin in western Arabia and then among the eastern nomad tribes. Janssens (1972:9–16) asserts that the spoken koine bore Eastern and Western Arabian features. Both Corriente and Janssens hold that the modern colloquials developed from this spoken koine.

A second theory is Ferguson's (1959) concept of a military koine, originating after the Arab expansion in the military camps. The existence of this koine had already been assumed by Fück (1955:7) who thought, contrary to Ferguson's thesis, that it was the basis of Classical Arabic. Ferguson believed that diachronic dialectology should study the growth of partial differences at the expense of similarities and the increase of partial similarities at the expense of differences (Weinreich 1977:309, 314). He listed 14 common features between Arabic dialects (1997c: 55–68). The language exported by Arabs, which spread all over most of the Islamic world in the first centuries of the Islamic era, was a relatively homogeneous koine, totally different from the traditional poetic koine (Ferguson 1959: n. 4). It was a new language variety, not based on the dialect of a single locality, which developed as a conversational form of Arabic just after the conquest, although, as he admits, it seems highly probable that the origin of this koine lies before the great expansion of Arabic. For Ferguson, this spoken koine existed side by side with the standardized variety, even at the time of Muḥammad, and diverged in many parts of Arabia from the standard *ʿarabiyya*, which implies the kind of diglossia that exists nowadays in Arabic-speaking countries (Ferguson 1959, 1996), although he admits that evidence of such continuation on any large scale is hard to confirm. Followed mainly by Janssens and Fischer, Ferguson's structural approach – which was greatly criticized (Blau 1981:14–17) – reached a new conclusion: modern dialects were generated from a spoken koine which existed in pre-Islamic Arabia and was stressed as post-Islamic dialects (Janssens 1972:16–18). Fischer regards this post-Islamic koine as a 'proto Neo-Arabic' variety, which could explain – if confirmed by more historical-comparative studies – similarities between Neo-Arabic dialects, especially those spoken in the central area of the Arabic-speaking world (Fischer 1995:85–86).

Agreement exists that post-Islamic Arabian peninsular spoken varieties, urban or Bedouin, slowly changed to such an extent that, by the 10th century, Arabic grammarians were alarmed by operated changes among their people and complained that nobody spoke Arabic correctly; grammatical treatises are full of anecdotes about Arabs who made grammatical errors → *laḥn* (Owens 1988:21–28; al-Farrāʾ, *Muḏakkar* 24). The same language variation was present among the Arabs who conquered new territories.

Outside the Arabian Peninsula two poles should be considered: the Arab contingents and vernacular populations. Scholars who admit the existence of spoken pre-Islamic varieties different from Classical Arabic, whether emerging from a leveled koine (like Corriente, Janssens, Ferguson, and Fischer) or from a non-specified situation (like Rabin) reached the evident conclusion that Arabs established in the newly conquered land spoke the same speech varieties they used before taking part in these 'military-like' campaigns (Shaban 1971:34–35). Each variety bore all or part of the features described above: loss of declension and mood endings, certain evolutions proper to all Semitic languages, and special tribal isoglosses, probably leveled before the expansion in urban *ḥaḍarī* varieties or leveled after it in military camps – Baṣra, Kūfa, al-Fusṭāṭ, or al-Qādisiyya – transformed, a few years later, in active urban nuclei (Pellat 1953:125). The evolution of these spoken varieties continued and established the basis of spoken varieties which were to constitute, over time, the colloquial varieties of the area with its differentiating features. This explains the similarities between contemporary dialects even when they are geographically very distant from one another (Owens *forthcoming*: "the statistical similarities between Uzbekistan Arabic and western Sudanic Arabic are due to a common pre-diaspora ancestor on the Arabian peninsula"), with no possibility of applying the dialectological theory of dialect continuum and wave diffusion (Hudson 1996:35–36, 39–41). Actually, as suggested by Diem (1979:63), the beginning of Neo-Arabic was much more influenced by the proper evolution of Arabic than by substrata of the conquered populations.

Although Arab tribes, mostly southern Arabs, settled in the valleys of the Tigris and Euphrates in Mesopotamia and the Nile in Egypt, as well as in East Africa during Roman times and had commercial activities until the 2nd century C.E. (al-Jamīlī 2003:94), no linguistic influence was left. So when Arabs expanded outside the limits of the Arabian desert, comprising the Syrian desert and Mesopotamia – geographical precision is important because in the north of the peninsula, as is historically known, Arab tribes lived in contact with Persian and Aramaean populations and might have participated in the expansion – they arrived in totally non-Arabic speaking societies.

For Fück and others supporting his views, who believe that Arabs, as a whole, reached new territories speaking the Classical variety or Classical-like varieties, these non-Arab populations, who came into contact with Arab conquerors, initiated and caused the deterioration of language purity and, most importantly, they were the origin of the formation of dialects. Following this view, conquered populations are to be socially divided into a high class and a low class: the former came into direct contact with the new masters and began to distort Arabic on phonetic and grammatical levels, which led grammarians to write their treatises to protect the language of the *Qur'ān* from corruption. The latter class, a low social stratum mainly constituted by prisoners of war and manumitted slaves, adopted a low variety of spoken Arabic without declension, which was the beginning of the dialects that arose in urban Muslim cities, each with its own properties (Fück 1955:8–11). This is how → 'Middle Arabic' literature emerged as the conquered population ignored the conquerors' Classical Arabic, mostly due to their incapacity to cope with the synthetic (non-analytic) type of Arabic language (Blau 1981: 1–18). In this view, Neo-Arabic emerged from Classical Arabic and non-Arabic populations were the origin of the dialects.

This position led to another thesis based on the same supposition: the genesis of dialects built on a 'pidgin-like' *lingua franca*, proposed by Fück (1955:8–11) and reinforced by Versteegh (1984), a sort of pidgin-Arabic, similar to pidgin-English, which arose as a consequence of the difficulties inherent in learning a totally different language. A pidgin is a language variety created when speakers of one language, for commercial reasons, come into contact with speakers of another and neither knows the other's language (→ pidginization). It is the imperfect version of the language of the stronger, usually the incoming population. Creolization of a pidgin takes place when it becomes the native language of a new generation, the shared home language and mother tongue of children, who later tend to normalize their language in a process called decreolization (Jespersen 1964: 216–236; Fasold 1990:180–220). This theory, based on the homogeneity of the exported language and its atomization in pidgins, reached this conclusion: Arabic was pidginized, creolized, and then decreolized (Versteegh 1984), a theory that has met much criticism because it goes against the observed uniformity between neo-Arabic dialects (Fischer 1995:76–78).

For the other scholars, who took into consideration pre-Islamic spoken varieties introduced by Arabs, vernacular non-Arab populations began a long and slow process of coexistence and contact between their vernacular language and Arabic with linguistic consequences that varied from one locality to another, depending on the local conditions (Jespersen 1964:201; Weinreich 1974:1–70).

Conquered populations exerted a subtle and gradual influence on the spoken variety, i.e. dialects established in the area, as vernacular populations abandoned their original language and adopted that of the conquerors. During this process, linguistically known as substratal influence, people kept many of their speech-habits, especially with regard to articulation and accent, even while using mainly the vocabulary of the new language, which in turn was tinged to a large extent by the old language (Jespersen 1964:191–192, 200–201). In this way the indigenous populations affected not the genesis but the beginning and evolution of Arabic dialects.

The future of the process as a whole, generally called 'Arabicization' – which is not yet finished, as stated by Fischer (1995:77) – depended on many extra-linguistic conditions, but each case needs to be examined individually. Nevertheless, on the whole, it can be said that the conquered populations – Aramaeans (in Iraq and Syria-Lebanon-Palestine), Berbers (Maghreb), Copts (Egypt), Iberians (in the Iberian Peninsula) or

Persians (Persia) – followed the same path, although reaching different results. To some degree, they all affected the already shaped colloquial variety – or varieties – introduced by the Arabs (Diem 1979). An example is that of the Mozárabes in → al-Andalus, who from the very beginning made an effort to learn Arabic and shaped a bilingual society (Millet-Gérard 1984: 49–53; Zwartjes 1997:5–22).

In addition, these conquered populations, which in the second generation became *muwalladūn* as they converted to Islam, took part in a process which was already on its way among the Arabic-speaking population: the formation of Middle Arabic literature. Like Arabs, they began to aspire to a correct use of Classical Arabic – which has never lost its supremacy; on the contrary, it gained in importance – and began using the same grammatical errors, but certainly to a much higher degree. For both Arabs and *muwalladūn*, Classical Arabic was an artificial, acquired language (Carter 1972:92).

4. CONCLUSION

The process, not only of Arabicization of the conquered population but also of language variation among the Arabs, progressed slowly during the first centuries of Islam (Garbell [1958: 303–306] proposed a possible evolution of the entire phonological/phonetic system although the dates are to be revised). This led to the formation of different types of Arabic colloquial varieties, each with its proper features. By the 10th century, dialectal areas were already shaped as al-Muqaddasī described them in 985 C.E. in his geographical treatise (Fück 1955:143).

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Farrāʾ, *Muḏakkar* = ʾAbū Zakariyyāʾ Yaḥyā ibn Ziyād al-Farrāʾ, *al-Muḏakkar wa-l-muʾannaṯ*. Ed. Ramaḍān ʿAbd at-Tawwāb. Cairo: Dār at-Turāṯ, 1989.

Ibn Jinnī, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. Ed. Muḥammad A. an-Najjār. 3 vols. Cairo: al-Hayʾa al-Miṣriyya li-l-Kitāb, 1986.

Suyūṭī, *Muzhir* = Jalāl ad-Dīn ʾAbū l-Faḍl ʿAbd ar-Raḥmān ibn ʾAbī Bakr as-Suyūṭī, *al-Muzhir fī ʿulūm al-luġa*. Ed. Muḥammad ʾAḥmad Jād al-Mawlā a.o. 2 vols. Beirut: Dār al-Fikr, n.d.

SECONDARY SOURCES

Abboud-Haggar, Soha. 2003. *Introducción a la dialectología de la lengua árabe*. Granada: El Legado Andalusí.

ʿAbd at-Tawwāb, Ramaḍān. 1980. "ʾAsbāb aš-šuḏūḏ fī l-luġa". *Studien aus Arabistik und Semitistik Anton Spitaler zum siebzigsten Geburtstag von seinen Schülern überreicht*, ed. Werner Diem and Stefan Wild, 1–22 (Arabic section). Wiesbaden: O. Harrassowitz.

——. 1987. *Fuṣūl fī fiqh al-ʿarabiyya*. 2nd ed. Cairo: Maktaba Miṣr. (1st ed. 1973.)

ʾAnīs, ʾIbrāhīm. 1965. *Fī l-lahajāt al-ʿarabiyya*. Cairo: al-Maktaba al-ʾAnglū al-Miṣriyya.

——. 1966. *Min ʾasrār al-luġa*. Cairo: al-Maktaba al-ʾAnglū al-Miṣriyya.

Bakalla, Muhammad H. 1984. *Arabic culture through its language and literature*. London: Kegan Paul International.

Belnap, Kirk and Niloofar Haeri (eds.). 1997. *Structuralist studies in Arabic linguistics: Charles A. Ferguson's papers, 1954–1994*. Leiden: E.J. Brill.

Blanc, Haim. 1979. "Diachronic and synchronic ordering in Medieval Arab grammatical theory". *Studia orientalia memoriae D.H. Baneth dedicata*, 155–180. Jerusalem: Magnes Press.

Blau, Joshua. 1978. "Hebrew and Northwest Semitic: Reflections on the classification of the Semitic languages". *Hebrew Annual Review* 2.21–44.

——. 1981. *The emergence and linguistic background of Judaeo-Arabic: A study of the origins of Middle Arabic*. 2nd ed. Jerusalem: Magnes Press. (1st ed. Leiden: E.J. Brill, 1965.)

——. 1988a. *Studies in Middle Arabic and its Judaeo-Arabic variety*. Jerusalem: Magnes Press.

——. 1988b. "The state of research in the field of the linguistic study of Middle Arabic". Blau (1988: 118–134.

——. 1988c. "The role of the Bedouins as arbiters in linguistic questions and the *masʾala az-zunbūriyya*". Blau (1988:135–144).

——. 1999. "The status and linguistic structure of Middle Arabic". *Jerusalem Studies in Arabic and Islam* 23.221–227.

Carter, Michael G. 1972. "Les origines de la grammaire arabe". *Revue des Etudes Islamiques* 40.69–97.

Chambers, John K. and Peter Trudgill. 1980. *Dialectology*. Cambridge: Cambridge University Press.

Cohen, David. 1962. "Koiné, langues communes et dialectes arabes". *Arabica* 9.119–144.

Corriente, Federico. 1971–1972. "On the functional yield of some synthetic devices in Arabic and Semitic morphology". *Jewish Quarterly Review* 52.20–50.

——. 1973–1974. "Again on the functional yield of some synthetic devices in Arabic and Semitic morphology". *Jewish Quarterly Review* 64.154–163.

——. 1975. "Marginalia on Arabic diglossia and evidence thereof in the *Kitāb al-Aġānī*". *Journal of Semitic Studies* 22.38–61.

——. 1976. "From Old Arabic to Classical Arabic through the pre-Islamic koine: Some notes on the native grammarians' sources, attitudes and goals". *Journal of Semitic Studies* 21.62–98.

——. 1992. *Árabe andalusí y lenguas romances*. Madrid: Colecciones Maphre.

Coseriu, Eugenio. 1988. *Sincronía, diacronía e historia: El problema del cambio lingüístico*. 3rd ed. Madrid: Gredos.

——. 1990. *Introducción a la lingüística*. Mexico: Universidad Nacional Autónoma de México.

Ḍayf, Šawqī. 1988. *Tārīx al-ʾadab al-ʿarabī: al-ʿAṣr al-jāhilī*. 7th ed. Cairo: Dar al-Maʿārif.

Diem, Werner. 1973. "Die nabatäischen Inschriften und die Frage der Kasusflexion im Altarabischen".

Zeitschrift der Deutschen Morgenländischen Gesellschaft 123.227–37.

——. 1979. “Studien zur Frage des Substrats im Arabischen”. *Der Islam* 56.12–80.

——. 1980. “Die genealogische Stellung des Arabischen in den semitischen Sprachen: Ein ungelöstes Problem der Semitistik”. *Studien aus Arabistik und Semitistik. Anton Spitaler zum siebzigsten Geburtstag von seinen Schülern überreicht*, ed. Werner Diem and Stefan Wild, 65–85. Wiesbaden: O. Harrassowitz.

Faber, Alice, 1997. “Genetic subgrouping of the Semitic languages”. Hetzron (1997:3–15).

Fasold, Ralph. 1990. *The sociolinguistics of language*. Oxford and Cambridge, Mass.: Blackwell.

Fatlī, ʿAbd al-Ḥusayn M. al-. 1988. “Jawānib min al-fawāriq al-lahjiyya fī n-naḥw wa-l-qirāʾāt”. *al-Mawrid* 17:1.95–121.

Ferguson, Charles A. 1959. “Diglossia”. *Word* 15.325– 340. (Repr. *Sociolinguistic perspectives: Papers on language in society, 1959–1994. Charles A. Ferguson*, ed. Thom Huebner, 25–39. New York and Oxford.)

——. 1996. “Epilogue: Diglossia revisited”. *Understanding Arabic: Essays in contemporary Arabic linguistics in honour of El-Said Badawi*, ed. Alaa Elgibali, 49–67. Cairo: American University in Cairo Press.

——. 1997a. Review of Harris Birkeland, *Growth and structure of the Egyptian Arabic dialect*. Oslo: Dybwad, 1954. Belnap and Haeri (1997:37–45).

——. 1997b. Review of Harris Birkeland, *Stress patterns in Arabic*. Oslo: Dybwad, 1954. Belnap and Haeri (1997:46–49).

——. 1997c. “The Arabic koine”. Belnap and Haeri (1997:50–68). (First published 1959.)

Fischer, Wolfdietrich. 1991. “What is Middle Arabic?”. *Semitic Studies in honor of Wolfgang Leslau*, ed. Alan S. Kaye, 430–436. Wiesbaden: O. Harrassowitz.

——. 1995. “Zum Verhältnis der neuarabischen Dialekte zum klassisch Arabischen”. Palva (1995: 75–86).

—— and Otto Jastrow. 1980. *Handbuch der arabischen Dialekte*. Wiesbaden: O. Harrassowitz.

Fück, Johann. 1955. *ʿArabiyya: Recherches sur l'histoire de la langue et du style arabes*. Paris: Didier.

Fleisch, Henri. 1974. “Arabe classique et arabe dialectal”. *Études d'arabe dialectal*, by Henri Fleisch, 3–43. Beirut: Dar el-Machreq.

Garbell, Irene. 1958. “Remarks on the historical phonology of an east Mediterranean Arabic dialect”. *Word* 14.303–337.

García Yebra, Valentín. 1982. *Teoría y práctica de la traducción*. Madrid: Editorial Gredos.

Ghul, Maḥmūd ʿAlī. 1993. *Early Southern Arabian languages and Classical Arabic sources*. Yarmouk: Yarmouk University Publications.

Hopkins, Simon. 1984. *Studies in the grammar of early Arabic, based upon papyri datable to before A.H. 300/A.D. 912*. London: Oxford University Press.

Hudson, R.A. 1996. *Sociolinguistics*. 2nd ed. London and New York: Cambridge University Press.

Ḥusayn, Ṭāhā. 1927. *Fī l-ʾadab al-jāhilī*. Cairo: Dār al-Maʿārif. (2nd ed. 1989.)

Hetzron, Robert (ed.). 1997. *The Semitic languages*. London and New York: Routledge.

Ingham, Bruce. 1982. *North east Arabian dialects*. London and Boston: Kegan Paul International.

Jamīlī, Xuḍayr ʿA. al-, 2003, “Intišār al-ʿArab fī Miṣr wa-šarq ʾIfrīqiyā qabla l-ʾIslām”. al-Majmaʿ al-ʿIlmī al-ʿIrāqī (2003:81–101).

Janssens, Gerard. 1972. *Stress in Arabic and word structure in the modern Arabic dialects*. Leuven: Peeters.

Jespersen, Otto. 1964. *Language, its nature, development and origin*. 12th ed. London: Allen and Unwin. (German original: *Die Sprache: Ihre Natur, Entwicklung und Entstehung*, Heidelberg, 1925; 2nd ed., 2003.)

Kaye, Alan S. 1972. “Arabic /jiim/: A synchronic and diachronic study”. *Linguistics* 79.31–72.

—— and Judith Rosenhouse. 1997. “Arabic dialects and Maltese”. Hetzron (1997:263–311).

Larcher, Pierre. 2001. “Moyen arabe et arabe moyen”. *Arabica* 48.578–609.

Majmaʿ al-ʿIlmī al-ʿIrāqī, al-. 2003. *al-Waṭan al-ʿarabī: an-Nuwāt wa-l-imtidādāt ʿabra t-tārīx*. Beirut: Markaz Dirāsāt al-Wiḥda al-ʿArabiyya.

Mayyāḥ, ʿAlī M. al-. 2003. “al-Qabāʾil al-ʿarabiyya fī l-ʿIrāq qabla l-ʾIslām”. al-Majmaʿ al-ʿIlmī al-ʿIrāqī (2003:57–79).

Meillet, Antoine. 1975. *Linguistique historique et linguistique générale*. Paris: Champion.

Millet-Gérard, Dominique. 1984. *Chrétiens mozarabes et culture islamique dans l'Espagne des VIII^{e} au IX^{e} siècles*. Paris: Etudes Augustiniennes.

Milroy, James. 1992. *Linguistic variation and change*. Oxford and Cambridge, Mass.: Blackwell.

Nöldeke, Theodor. 1904. *Beiträge zur semitischen Sprachwissenschaft*. Strassburg: K. Trübner.

Owens, Jonathan. 1988. *The foundations of grammar: An introduction to medieval Arabic grammatical theory*. Amsterdam and Philadelphia: J. Benjamins.

——. 1990. *Early Arabic grammatical theory: Heterogeneity and standardization*. Amsterdam and Philadelphia: J. Benjamins.

——. Forthcoming. “Pre-diaspora Arabic: Statistics, dialects and historical reconstruction”.

Palva, Heikki. 1995. *Dialectologia Arabica: A collection of articles in honour of the sixtieth birthday of professor Heikki Palva*. Helsinki: Studia Orientalia.

Pellat, Charles. 1953. *Le milieu baṣrien et la formation de Ǧāḥiẓ*. Paris: Librairie d'Amérique et d'Orient, Adrien-Maisonneuve.

Rabin, Chaim. 1951. *Ancient West-Arabian*. London: Taylor's Foreign Press.

——. 1955. “The beginnings of Classical Arabic”. *Studia Islamica* 4.19–37.

Retsö, Jan, 1983, *The finite passive voice in modern Arabic dialects*, Götenburg: Acta Universitatis Gothoburgensis.

——. 1992. “*Iʿrāb* in the forebears of modern Arabic dialects”, *Actes des premières journées internationales de dialectologie arabe de Paris*, ed. Dominique Caubet and Martine Vanhove, I, 333–342. Paris: INALCO.

——. 1995. “Pronominal state in colloquial Arabic: A diachronic attempt”. Palva (1995:183–192).

——. 2003. *The Arabs in antiquity: Their history from the Assyrians to the Umayyads*. London and New York: RoutledgeCurzon.

Robin, Christian. 1991–1993. “Les langues de la peninsule arabique: L'Arabie antique de Karibʾīl à Mahomet. Nouvelles données sur l'histoire des

Arabes grâce aux inscriptions". *Revue du Monde Musulman et de la Méditerranée* 61.89–111.

Sapir, Edward. 1949. *Language: An introduction to the study of speech.* 2nd ed. San Diego, Cal.: Harcourt Brace Jovanovich. (1st ed. 1921.)

Sarauw, Christian. 1908. "Die altarabische Dialektspaltung". *Zeitschrift für Assyriologie und Vorderasiatische Archäologie* 21.31–49.

Shaban, M.A. 1971. *Islamic history A.D.: A new interpretation*. I. *600–750 (A.H. 132)*. Cambridge: Cambridge University Press.

Ṭaʿʿān, Hāšim aṭ-. 1978. *al-ʾAdab al-jāhilī: Bayna lahajāt al-qabāʾil wa-l-luġa al-muwaḥḥada*. Baghdad: Wizārat aṯ-Ṯaqāfa.

Tikrītī, Bahjat K. at- a.o. 2003. "Intišār qabīla ʿAbd al-Qays qabla l-ʾIslām ʾilā nihāyat al-qarn al-ʾawwal al-hijrī". al-Majmaʿ al-ʿIlmī al-ʿIrāqī (2003:103–127). Beirut: Markaz Dirāsāt al-Wiḥda al-ʿArabiyya.

Versteegh, Kees. 1983. "A dissenting grammarian: Quṭrub on declension". *The history of linguistics in the Near East*, ed. Kees Versteegh, Konrad Koerner, and Hans-Josef Niederehe, 167–193. Amsterdam: J. Benjamins.

——. 1984. *Pidginization and creolization: The case of Arabic*. Amsterdam: J. Benjamins.

——. 1993. "Leveling in the Sudan: From Arabic creole to Arabic dialect". *International Journal of the Sociology of Language* 99.65–79.

——. 1997. *The Arabic language*. Edinburgh: Edinburgh University Press.

Vollers, Karl. 1906. *Volkssprache und Schriftsprache im alten Arabien*. Strassburg: K. Trübner.

Weinreich, Uriel. 1974. *Languages in contact: Findings and problems*. 8th ed. The Hague and Paris: Mouton.

——. 1977. "Is a structural dialectology possible?". *Readings in the sociology of language*, ed. Joshua A. Fishman, 305–319. The Hague and Paris: Mouton.

Zwartjes, Otto. 1997. *Love songs from al-Andalus: History, structure and meaning of the kharja*. Leiden: E.J. Brill.

Zwettler, Michael. 1978. *The oral tradition of classical Arabic poetry: Its character and implications*. Columbus, Ohio: Ohio State University Press.

Soha Abboud-Haggar
(La Complutense University)

Diathesis

1. Definition

Diathesis can be defined as the syntactic relationship between the verbal core of a sentence and its nominal constituents, i.e. the verbal predicate and the parts of speech directly relating to it, mainly subject and object. The relationship encompasses both semantic and morphosyntactic categories. A verbal sememe may thus automatically imply a series of nominal complements which differ in number and semantic content depending on the semantic class of the verb. Such nominal complements are named by some linguists 'actants'.

Thus, for instance, a verbal sememe meaning 'to rain' (= the falling of rain) does not normally imply any nominal sememe or actant (Miklosich 1883:1–7). A sememe meaning 'to fall' generates one nominal actant: 'a thing falls'. A sememe meaning 'to hit' generates two nominal actants: 'the man hit the thief'. The sememe 'to give' may generate three: 'the woman gave the book to the man'. It is immediately seen that the nominal actants may have different semantic properties. The nominal actant in the second example, 'the thing', can be said to have the same semantic properties or rather, fulfil the same semantic role as 'the book' in the fourth example, i.e. an object being affected by surrounding events designated by the verbal sememe. The first actant in the three latter examples has a similar semantic role: agent, or source of a verbally designated process or event. In the three latter cases it is obvious that there is a hierarchy between the actants as shown by the fact that usually the verb generates the first actant, the 'subject', but not necessarily the other(s), the 'first and second objects'. The one-actant verbs, i.e. those which only generate one actant, a 'subject', are inherently intransitive verbs, whereas those with two or more actants are transitive ones.

A special case of diathesis is when two of the nominal actants are coreferential, as in 'the mother washed the child' and 'the mother washed herself'. The second actant in both examples has the same position in the hierarchy but refers to different things, in the second case back to the first actant. The term 'reflexive construction' should be reserved for this case. This designation is also often used for some inherently intransitive as well as intransitivized transitives due to the fact that many languages use a similar morphological or syntactic marking of them. Yet, it is wise to keep all three apart.

The different semantic roles of the nominal actants and the relationship between them and the verbal core is marked in various ways on the morphosyntactic level of the language, and languages show great variation in how they handle this. The means used are mainly of three kinds: verbal agreement, word-order, and case-marking. Many languages, including Arabic, have a combination of all three.

2. DIATHETIC CHANGES

It is possible to change the 'normal' diathetic structure of a verbal sentence according to referential and contextual factors. It is, for instance, possible to leave out the first actant of most verbs if the context allows this. A sentence like 'there will be dancing tonight' implies the process designated by the verb, but a specification of the first actant is left out. In such a case, many languages choose not to generate any first actant or 'subject' at all. Many languages generate a 'dummy subject' like English 'it' or 'there'. In a similar way one can state 'there was fighting in the street last night', using a two-actant verb implying the process but without mentioning the source or the agent. These 'subjectless sentences' resemble sentences containing verbs like 'to rain', the difference being that in the latter the absence of the first actant is inherent, whereas in the former it is generated by the context.

Another case is when the second actant is not generated with verbs which may have it, like 'he hit' (= 'he did not caress'), 'she gave' (= 'she did not sell'). This process may be called intransitivization, distinguishing it from the inherent intransitive verbs = one-actant verbs. Many languages tend to employ in such cases the same devices as with the pure reflexive verbs, although not consistently. These cases where the full possible actant structure is not realized often receive special kinds of syntactic and/or morphological marking. A two- or three-actant verb may generate the second actant but not the first. In many languages, including Arabic, an expression like 'the murder of a man happened yesterday' may be construed with a finite verb rather than a noun, 'murder'. Absence of the first actant is marked in different ways, one of which includes the traditional passive construction (Retsö 1982– 1983). This passive construction has the following properties: (a) the first actant of a two- or three-actant verb is absent; (b) the second or third actant is moved upward in the hierarchy assuming the morphological (case-marking) and syntactic (determining verbal agreement) properties of the first actant; and (c) the role of the new first actant in the diathesis is marked morphologically or syntactically by a change in the verb. The absence of the first actant may be conditioned by many factors: the first actant is unknown, the speaker does not want to specify it, it is a general agent (people, usually, etc.), or it is recoverable from the context.

From this construction should be distinguished the one where the first actant reappears in the sentence as an adverbial complement, the so-called agent construction. This is also a diathetic change whereby the first actant is not absent but moved downward in the hierarchy, its former place now being occupied by the second actant. Unlike the preceding construction, this one is a device for → topicalization or thematization (→ theme/rheme) of a constituent in the sentence. The term passive construction is usually employed for both, but it should be kept in mind that they are two different diathetical constructions.

Another diathetic construction is the → causative. This implies the adding of a new first actant to a diathetical structure: 'the tree falls' > 'the man felled the tree'; 'the woman wrote a letter' > 'the man had the woman write a letter'. The addition of a new first actant moves the original first actant one step downward in the hierarchy transforming it into a second actant, an 'object'. A causative of the three-actant verb thus transforms it into a four-actant one. Closely related to the causative is the factitive which is the transformation of a nominal sememe to a verbal one by adding an agent, as when the adjective 'thick' is verbalized into 'thicken' = 'make thick' by the addition of a new first actant.

3. DIATHESIS IN ARABIC

The most important diathetical categories in Arabic are:

(1) the → passive construction, i.e. the case where the absent first actant is replaced by the second, 'disguised' morphologically and syntactically as a first actant;
(2) the passive construction with agent extension, i.e. a topicalization device;
(3) 'the subjectless sentence', i.e. the case when the first actant is absent, not being replaced by any other constituent;
(4) the causative/factitive construction, i.e. when a new first actant is added.

All forms of Arabic show similar patterns in handling diathesis and diathetic changes, which it shares with the other ancient Semitic languages as part of the common heritage. There is, however, considerable morphological variation in marking different kinds of diathetical structures. To the oldest devices belong the use of the *t*-element, prefixed or infixed to the verbal root, which probably originally marked intransitivization. Thus, for instance, the frequent opposition between Forms II and V often exhibits the distinction two-actant verb-intransitivized two-actant verb, like *rakkaz-* ‘to concentrate (something)’, *tarakkaz-* ‘to concentrate (oneself)’, *kabbar-* ‘to increase’, ‘to magnify (something)’, *takabbar-* ‘to be proud, haughty’. This pattern is old (even if the examples may not be).

Another ancient diathetical marking is the *n*-prefix of Form VII. This seems to be an ancient morpheme marking verbs of the type ‘fall’ with a non-agentive first actant. This explains its occurrence in verbs like *-nfajar-* ‘to explode’, *-nṭalaq-*, ‘to emanate, sweep along’, *-nqaṭaʿ*– ‘to break [intr.]’, where we may have traces of its original function, which may be seen operating in, e.g., Akkadian and Ethio-Semitic. The *t-* and *n-* morphemes are found with these functions in all documented forms of Arabic. In the above-mentioned functions they are hardly productive any longer. Both morphemes are, however, widely used in the spoken forms of Arabic marking the verb in a passive construction (1 and 2) as well as in a subjectless sentence (3). Their function as markers of verbs in the passive construction is explicable from the intransitive character of their original function. Both verbs with the *n*-morpheme and with the *t*-morpheme were originally intransitives and mostly still are.

The use of *n-* and *t*-forms in the passive construction and subjectless sentences is found in the *ʿarabiyya* as well, but to a much lesser degree than in the dialects, since the *ʿarabiyya* has another device for this kind of verbal marking, viz. the so-called ‘inner passive’. This is an Ablaut-like device marking the verb in a passive or subjectless construction by a vowel sequence different from the one in an active construction (→ apophony). In the *ʿarabiyya*, this is the regular marking of the verb both in subjectless sentences and passive constructions. In the Ablaut-system the finite verb in an active construction has several different vowel patterns depending on form and tense, whereas the verb in a passive construction always has the same vowel sequence. In the perfect there is *-u-* in all syllables except the last: *qaṭaʿ-/quṭiʿ-*, *šarib-/šurib-*, *kabbar-/kubbir-*, *(i)staktab-/(u)stuktib-*. The imperfect has *-a-* in all syllables except the first (= the marker of the person): *yaqṭaʿ/yuqṭaʿ-*, *yašrab-/yušrab-*, *yukabbir-/yukabbar-*, *yastaktib-/yustaktab-*. The difference in marking of the verb in a passive construction is one of the most salient differences between the *ʿarabiyya* and the modern spoken dialects of Arabic, even if some dialects (Central Arabia, the Gulf region, Yemen, Mauritania) have developed similar systems that operate parallel to the *t-* and *n-* forms. Historically, the Ablaut-marking of the verb in the passive construction in the *ʿarabiyya* is probably connected with the morphology of semantically stative verbs in general, since their common syntactic characteristic is intransitivity. A small group of statives in fact shows the same morphology: *huḏil-/yuhḏal-* ‘to be/become meagre’. Sometimes the difference between *t-* and *n-* forms in their old function and the ablaut of the passive construction can be exploited: *(i)nhazama l-jayšu* ‘the army fled’; *huzima l-jayšu* ‘the army was put to flight’. In Modern Standard Arabic, especially in newspaper style, an analytic passive construction consisting of the (intransitive) verb *tamm-* + a verbal noun (*maṣdar*) as subject is frequently found: *tamma ʾiʿdāmu ṯalāṯati ʾašxāṣ ʾamsi* ‘three persons were executed yesterday’ (Holes 1995:257–260).

4. THE PASSIVE CONSTRUCTION (= 1 AND 2)

Both the *ʿarabiyya* and the dialects have a passive construction with and without agent extension. The former is, however, of lower frequency in Arabic than in European languages. According to Classical Arabic grammar, agent extension does not (or rather should not) exist at all, which lies behind the designation for the passive construction: *majhūl* ‘unknown’, i.e., a verb of which the agent (→ *fāʿil* ‘subject, first actant’) is unknown. It is also often claimed that the *t-* and *n-* form cannot be used in passive constructions in the *ʿarabiyya*, which, however, is not quite correct. Passive constructions with agent extension do occur in all periods of Arabic, from the *Qurʾān* and early poetry to Modern Standard Arabic, e.g. *ʾurḍiʿna bihā* ‘we were nursed by her’, *subiqtu bihi* ‘I was left behind by him’, *(a)llaḏīna yusʾalūna lahu* ‘those who were asked by him’ (cf. Q. 2/178; Retsö 1983:29, n.2). In Modern Standard Arabic agent phrases are quite common, introduced by *min*, *li-*, *min qibal*, *min jānib*, or *ʿalā yad*. Their increased frequency compared to earlier stages is mainly due to influence from English and French. Analysis of the context of passive constructions in Arabic shows clearly that the first actant is often absent not only because it is unkown. In the dialects, too, agent extension occurs: *hāḏā nbahar* ***min*** *jamālha* ‘he was blinded by her beauty’ (Palestine, Bir Zeit); *ssəlṭān taṯṯar* ***biya*** ‘the sultan was impressed by this’ (Baghad Jewish); *nəḥna nətṣarraf* ***fikon*** ‘we are honored by you’ (Damascus). The morphology of the verb in the passive construction in the dialects differs. It is remarkably uniform as far as the derived Forms II and III are concerned. These generally add the *t*-morpheme in the passive construction, i.e., Forms V and VI. The verbs of Form I show three main variants with different geographical distribution. Almost all dialects of the Arabian Peninsula as well as Syria-Mesopotamia use the *n*-form, i.e. Form VII. Only in some regions in Yemen a form with a *t*-infix, i.e. Form VIII, is used. The *n*-form also has a wide distribution in North Africa where, however, some regions, viz. northern Tunisia, northwestern Morocco, parts of the Nile valley including Cairo, use a *t*-prefixed form of Form I. In many areas there are isolated cases of the Ablaut variety limited to certain lexemes like *qitil* ‘he was killed’. In the poetic language of the Bedouin in Arabia these forms have a wider distribution and may be used as free variants to the *t*- and *n*- formations in oral poetry (→ Najdī Arabic). It is doubtful whether these are regularly used in the normal spoken language. The dialect of → Ḥassāniyya in Mauritania has an apophonic system in passive contructions, too. It does, however, also employ the *t*- and *n*- forms in the same function. Analytic passive constructions consisting of an auxiliary verb + a passive participle are documented in at least some dialects, but their frequency and function have not yet been studied: *rāḥ marmi fi l-arḍ* ‘he was thrown to the ground’ (Cairo).

5. SUBJECTLESS SENTENCES (=3)

The verb in a subjectless sentence shows the same morphology as in the passive construction and is probably influenced by the latter since the passive construction is a far more frequent phenomenon than the subjectless one. This is also why the subjectless sentence is sometimes called ‘impersonal passive’. The *ʿarabiyya* may, however, use an ‘active’ form for inherently subjectless constructions with ‘meteorological verbs’ (cf. Miklosich 1883:43–46): *wa-lākinna l-yawma l-laḏī fīhi xaraja lūṭu min sudūma* ***ʾamṭara*** *nāran wa-kibrītan min as-samāʾi* (Luke 17:29) ‘on the day when Lot went out from Sodom it rained fire and sulphur from heaven’. With a normal transitive verb the use of a subjectless construction with the verb in the same form as with a passive construction is occasionally found in the *ʿarabiyya*: *wa-****yuxraju*** *lahu yawma l-qiyāmati kitāban* ‘on the day of resurrection a book will be brought to him’. This is a variant reading of Q. 17/14, where the Vulgate has an active construction: ***nuxriju*** ‘we will bring out’. The second actant, the accusative object, remains in the *a*-case, the normal appearance of the second actant. The third actant, the ‘dative object’, is introduced by the preposition *li-*.

The subjectless construction is quite frequent with intransitive verbs, i.e. one-actant verbs, where it is usually the equivalent of constructions with ‘dummy’ subjects like ‘they’, ‘one’, ‘man’ in Western European languages (cf. Miklosich 1883:58–64). There seems to be a rule that this construction should have an adverbial/prepositional complement of some kind: *furiḥa bihi* ‘there was rejoicing over it’ (with first actant: *fariḥa bihi* ‘he rejoiced over it’); *sīra ʾilā l-ʿirāq* ‘there was a journey to Iraq’, ‘someone went to Iraq’ (with first actant: *sāra ʾilā l-ʿirāq* ‘he went to Iraq’).

The Arabic dialects have the same construction but with a different kind of marking of the verb. Subjectless contructions in the dialects tend to occur in certain standing expressions, very often with a negation: *hal-bēt ma-****yinsikin*** *bi* ‘this house is inhabitable’ (Baghdad Muslim), ***bənṭəleʿ*** *aṣṣṭuḥ* ‘is it possible to ascend to the terrace?’ (Aleppo). In both cases, the absent actant of an intransitive verb is marked in the verb by the *n*-morpheme. The active forms of the verbs would be *yiskin* and *yəṭlaʿ*, respectively. The subjectless construction is also found with transitives but is much rarer and possibly belongs to literary style: *wa-yza bəlbab* ***byənda***” ‘immediately they knocked on the door’ (Damascus, literary). A case standing between construction (1) and (2) is the one where a verb with marking of subjectlessness or passive construction has a clause as its complement, e.g.: *ma **yindiri** rāḥ yistiqil lo yibqa*, 'it is not known if he is going to resign or stay (Baghdad M)'; ***ttesḥab-li** bāb kānet mesdūda* 'I thought [lit. 'it seemed to me'] that the door was closed' (Morocco). It is not possible to determine whether the clause complement should be seen as the subject (first actant) to the verb or as an object complement (second actant). Another construction difficult to analyze is with the verb 'to say' in the meaning 'to call by name': ***yəngalhom** ʿeylt ʿabdalla* 'they were called the family of ʿAbdallah' (Jordan, Ajarma), where it is uncertain whether the name should be seen as the subject to a verb in a passive construction, or as a complement to a subjectless verb.

6. THE CAUSATIVE/FACTITIVE (= 4)

The → causative is a productive category in all known forms of Arabic. In the *ʿarabiyya* it is usually marked by Form IV of the verb, which in the perfect is characterized by a prefix *ʾa-* and in the imperfect by the vowel-sequence *u-i*. A difference between the *ʿarabiyya* and the dialects is that in the latter this form either is of low frequency or absent altogether. It seems to be used only in some dialects of the *badawī*-type. Instead, most dialects employ Form II as causative to Form I. This usage is not unknown in the *ʿarabiyya*, but seems limited to some lexemes. As in many other languages, the factitive in Arabic has the property of often being the verbalization not only of an adjective but also of a stative verb. An adjective like *kabīr* 'big', which may stand as a predicate in a non-verbal sentence has a verbal equivalent, which may replace a non-verbal sentence with a finite verb, which is syntactically intransitive (no second actant): *kabur-*, 'to be/become big'. In opposition to this stands a factitive verb *kabbar-* 'to make big', which implies the addition of a new first actant and the transformation of the original one into a second actant, i.e. an object (Retsö 1989:48–52). Both Forms II and IV are used for both causative and factitive in the *ʿarabiyya*, whereas most dialects have Form II for both.

BIBLIOGRAPHICAL REFERENCES

Holes, Clive. 1995. *Modern Arabic: Structures, functions and varieties*. London: Longman.
Miklosich, Franz von. 1883. *Subjectlose Sätze*. 2nd ed. Vienna: W. Braumüller.
Retsö, Jan. 1982–1983. "Subjectless sentences in Arabic dialects". *Orientalia Suecana* 31–32.71–91.
——. 1983. *The finite passive voice in modern Arabic dialects*. Göteborg: University of Göteborg.
——. 1989. *Diathesis in the Semitic languages: A comparative morphological study*. Leiden: E.J. Brill.

JAN RETSÖ (Göteborg University)

Dictionaries → Lexicography

Ḍidd

In Arabic lexicography this term indicates '(a word) having two mutually exclusive meanings' (pl. *ʾaḍdād*). The compilations concerning the *ʾaḍdād* pertain to the activity of collecting every aspect of the Arabic language undertaken by philologists from the end of 2nd/8th century onwards. The case of the *ʾaḍdād*, according to the Arab philologists, is a particular instance of the lexical category of the → *muštarak* 'the common one', i.e. homonymous polysemic words, such as *ʿayn* 'eye', 'source', 'coin', and so on (cf. as-Suyūṭī [d. 911/1505], *Muzhir* I, 369–386), which in this particular case, have contradictory meanings, e.g. *jawn* which means 'white' and 'black', or *jalal* 'momentous matter' and 'trifle'. Words of this kind are few, according to Ibn al-ʾAnbārī ([d. 328/940] *ʾAḍdād* 6); this author, too, explicitly places the *ʾaḍdād* within the frame of the homonymous polysemic words, which can be understood only by means of their linguistic (or, in poetry, literary) context (Ibn al-ʾAnbārī, *ʾAḍdād* 4–5). The first scholar to whom a *Kitāb al-ʾaḍdād* is ascribed is Quṭrub (d. 206/821). During the 3rd/9th century, several books on the same subject are recorded, and of these at least six have been edited. Works on *ʾaḍdād* have continued to be written till our time (cf. ʾAḥmad 1989:53–57; ʾĀl Yāsīn [1985] reports a list of 33 authors). The great interest this issue raises is explained by some scholars, such as Kamal (1967:62), as a reply to the attacks coming in particular from the Šuʿūbites, who criticized the Arabic language for exhibiting a phenomenon leading to obscurity and misunderstanding (cf. for instance Ibn al-ʾAnbārī, *ʾAḍdād* 1–2; Kofler 1931–1932:389).

Lists of the early edited works on *ʾaḍdād* are found in Cohen (1961) and Weil (1960). Since then, other works have been edited, namely those of ʾAbū ʿUbayd (d. 224/838), at-Tawwazī, (d. 233/847), ʾAbū ṭ-Ṭayyib al-Luġawī (d. 351/962), Ibn ad-Dahhān (d. 569/1174), and al-Munšī (d. 1001/1593); chapters on this topic are found in general works on linguistic or lexicographical questions, such as as-Suyūṭī's *Muzhir* (I, 387–402), which reports examples of *ʾaḍdād* taken not only from specific works but also from very different literary sources.

Some of the works on *ʾaḍdād*, especially the later ones, such as those of aṣ-Ṣaġānī, Ibn ad-Dahhān, or al-Munšī, are mere lists of words collected from previous sources, arranged in alphabetical order, without *šawāhid* 'textual evidence'. Early works are also lists, but they present extensive discussion of the meanings of the words in their different poetic or Qurʾānic contexts. As-Sijistānī (d. ca. 255/869) shows (*ʾAḍdād* 72) clearly the turn of mind which governs these early compilations: to clarify to non-Arab Muslims the exact meaning of each Qurʾānic passage. Since a Qurʾānic passage must have a dogmatically unambiguous meaning, a verb such as *ḏanna* cannot mean 'to presume' in those passages where an article of faith is concerned, as in Q. 72/12, Q. 18/53, or Q. 69/20; therefore, *ḏanna* is given the *ḍidd* meaning of 'to doubt' and 'to be sure' (cf. Ibn al-ʾAnbārī, *ʾAḍdād* 14–15; Kofler 1931–1932:391–392).

As ʾĀl Yāsīn (1979:162) points out, later on a trend to fill out the collections led authors to include among the *ʾaḍdād*, all at the same level, words which only for metaphorical or stylistic reasons could be used with two opposite meanings, or words coming from different readings of the *Qurʾān* (as-Sijistānī, *ʾAḍdād* 203), or from two different forms of a verb. Ibn al-ʾAnbārī disagrees with Quṭrub's opinions, when the latter lists among the *ʾaḍdād* words like *rabʿ* 'dwelling' and *rabʿa* 'quick pace' (Ibn al-ʾAnbārī, *ʾAḍdād* 366) or *xaḏimat an-naʿl* '[the lace of] the sandal became cut' and *ʾaxḏamtu-hā* 'I repaired it' (*ʾAḍdād* 371), because each of the two items has its own single meaning; or words such as *ḥirfa* 'wealth' and 'poverty', because the second meaning is used only by the common people (*ʾAḍdād* 366).

According to Ibn Fāris (d. 395/1004, *Ṣāḥibī* 117) and Ibn Sīda (d. 458/1066, *Muxaṣṣaṣ* XIII, 259), there are 'people' (*nās*) who deny the existence of the *ʾaḍdād*, and according to al-Jawālīqī (d. 539/1144, *Šarḥ* 251), who agrees with them, they are *al-muḥaqqiqūna* 'the experts' in Arabic. Ibn Fāris, who affirms that he has written a book refuting this opinion, does not mention them, but we know from as-Suyūṭī (*Muzhir* I, 396) that Ibn Durustawayh (d. 346/957) composed a work entitled *ʾIbṭāl al-ʾaḍdād* 'Invalidation of the *ʾaḍdād*'. According to Weil (1960), al-Mubarrad (d. 286/900) also held this opinion, but this author himself (*Mā ttafaqa lafḏu-hu waxtalafa maʿnā-hu* 3–11) seems to accept the common view, because he quotes, without mentioning the phenomenon of the *ʾaḍdād*, the usual examples *jalal*, *jawn*, and *ḏanna*.

Many medieval Arab scholars, without denying a phenomenon which in their opinion does not lead to ambiguity, owing to the particular structure of Arabic language (Ibn al-ʾAnbārī, *ʾAḍdād* 1–3), and which on the contrary shows its extensiveness and offers multifarious means of achieving a literary expression, emphasize the origin or the formation of the *ʾaḍdād* and state that the original meaning of a *ḍidd* was often one and the same, because the basic rule of the language is that each nominatum has its particular noun (Ibn al-ʾAnbārī, *ʾAḍdād* 8). Ṯaʿlab (d. 291/904), for instance, says (Ibn al-ʾAnbārī, *ʾAḍdād* 16) concerning *ḏanna* that it indicates 'an inner speech', which may prove to be true, so that the verb may receive the meaning 'to be sure of', or 'to be uncertain', hence the meaning 'to doubt'. The book gives many examples of this kind, e.g., (*ʾAḍdād* 8) about *ṣarīm* 'day' and 'night' and (*ʾAḍdād* 27) about *qurʾ* 'menstruation' and 'state of purity [from a menstruation]'. Al-Jawālīqī (*Šarḥ* 251) quotes Ṯaʿlab's remarks of the same genre, concerning *talʿa* 'elevated ground' and 'depressed ground' and *jawn* 'white' and 'black', and throughout his chapter on this question (*Šarḥ* 251–257) al-Jawālīqī tries to demonstrate that words considered to have contradictory meanings do not in fact have them.

Another argument lending support to those who are inclined to view one meaning only at the origin of a *ḍidd*, is pointed out by those who remark that sometimes the two opposite meanings were attested in the dialect of two different Arab tribes and that they were combined afterwards, e.g. *mušāyiḥ*, which means 'who strives [in fighting]' in the dialect of Huḏayl and 'cautious, fearing' in Najd (Ibn as-Sikkīt [d. ca. 244/858], *ʾAḍdād*, 193; see as-Sijistānī, *ʾAḍdād* 125; al-ʾAṣmaʿī [d. 213/828], *ʾAḍdād* 39), or *sājid*, which means ‘erect, straight’ in the dialect of Ṭayyiʾ and ‘inclined, bowed’ elsewhere (Ibn as-Sikkīt, *ʾAḍdād* 196; al-ʾAṣmaʿī, *ʾAḍdād* 43; Ibn al-ʾAnbārī, *ʾAḍdād* 294). Another reason mentioned is antiphrasis for rhetorical or apotropaic (*ʿalā jihat at-tafāʾul*, Ibn al-ʾAnbārī, *ʾAḍdād* 105, 267) reasons, e.g. *mafāza*, which means ‘place of perdition’ but which is also used to designate a ‘place of safety’ (Quṭrub, *ʾAḍdād* 248; al-ʾAṣmaʿī, *ʾAḍdād* 38; Ibn as-Sikkīt, *ʾAḍdād* 192; Ibn al-ʾAnbārī, *ʾAḍdād* 104–105), *salīm* ‘safe’, which means also ‘bitten, stung’ (as-Sijistānī, *ʾAḍdād* 99, 114; Ibn al-ʾAnbārī, *ʾAḍdād* 105; *mafāza* and *salīm* often occur together). A survey of the various explications which can be given for the presence of a *ḍidd* in the work of Arab medieval authors is found in ʾĀl Yāsīn (1985).

Contemporary Western essays on this phenomenon have tried, in various ways, to reduce the number of words viewed as having two opposite meanings and/or to find a general explanation for all or most of them, sometimes on the basis of more sociopsychological than linguistic considerations. Critical reviews are found in Weil (1960) and Cohen (1961); see also Reig (1971). An introduction to the subject of the *ʾaḍdād* is also found in the articles collected by Charnay (1967). Bausani (1971) represents a commentary on this book; he distinguishes between the conscious use of ambivalent words, for rhetorical or philological purposes, and the possible existence of remnants of an ancient bipolarity in Arabic as well as in other languages. In this case he points out the importance of the gestures that accompany the ambiguous words in some languages. Cohen (1967) emphasizes the sociocultural importance of the reflections of Arab scholars on the question, which he considers more relevant than the linguistic reality of the *ʾaḍdād*. Gardet (1986) links the notions of *ḍidd*, *taḍmīn*, *taxrīj*, and *muqābal* in the frame of a ‘bipolarity’ of Arab thinking.

Bibliographical references

Primary sources

ʾAbū ʿUbayd, *ʾAḍdād* = ʾAbū ʿUbayd al-Qāsim ibn Sallām al-Ḥarawī, *Kitāb al-ʾaḍdād*. Ed. Muḥammad Ḥusayn ʾĀl Yāsīn. *Majallat al-Majmaʿ al-ʿIlmī al-ʿIrāqī* 38 (1987) 257–296.

ʾAbū ṭ-Ṭayyib al-Luġawī, *ʾAḍdād* = ʾAbū ṭ-Ṭayyib ʿAbd ar-Rahmān ibn ʿAlī al-Luġawī, *Kitāb al-ʾaḍdād fī kalām al-ʿarab*. Ed. ʿIzza Ḥasan. 2 vols. Damascus: Majmaʿ al-Luġa al-ʿArabiyya, 1963.

ʾAṣmaʿī, *ʾAḍdād* = ʾAbū Saʿīd ʿAbd al-Malik ibn Qurayb al-ʾAṣmaʿī, *Kitāb al-ʾaḍdād*. Ed. Haffner (1913:51–61).

Haffner (1913) = *Ṯalāṯa kutub fī l-ʾaḍdād*. Ed. August Haffner and ʾAnṭūn Ṣālḥānī. Beirut: Imprimerie catholique, 1913.

Ibn al-ʾAnbārī *ʾAḍdād*, = ʾAbū Bakr ibn al-Qāsim Ibn al-ʾAnbārī, *Kitāb al-ʾaḍdād*. Ed. Muḥammad ʾAbū l-Faḍl ʾIbrāhīm. Kuwait, 1960.

Ibn ad-Dahhān, *ʾAḍdād* = ʾAbū Muḥammad Saʿīd ibn al-Mubārak Ibn ad-Dahhān, *al-ʾAḍdād fī l-luġa: Nafāʾis al-maxṭūṭāt, al-majmūʿa al-ʾūlā*. Ed. Muḥammad Ḥusayn ʾĀl Yāsīn. an-Najaf: al-Maṭbaʿa al-Ḥaydariyya, 1953.

Ibn Fāris, *Ṣāḥibī* = ʾAbū l-Ḥusayn ʾAḥmad Ibn Fāris, *aṣ-Ṣāḥibī*. Ed. ʾAḥmad Ṣaqr. Cairo: Maṭbaʿa ʿĪsā al-Bābī al-Ḥalabī, 1977.

Ibn Sīda, *Muxaṣṣaṣ* = ʾAbū l-Ḥasan ʿAlī ibn ʾIsmāʿīl Ibn Sīda, *Kitāb al-muxaṣṣaṣ*. 17 vols. Būlāq: al-Maṭbaʿa al-ʾAmīriyya, 1316–1321 A.H.

Ibn as-Sikkīt, *ʾAḍdād* = ʾAbū Yūsuf Yaʿqūb ibn ʾIsḥāq Ibn as-Sikkīt, *al-ʾAḍdād*. Ed. Haffner (1913: 163–209)

Jawālīqī, *Šarḥ* = ʾAbū Manṣūr Mawhūb ibn ʾAḥmad al-Jawālīqī, *Šarḥ ʾAdab al-kātib*. Cairo: Maktabat al-Quds, 1350 A.H.

Mubarrad, *Mā ttafaqa* = ʾAbū l-ʿAbbās Muḥammad ibn Yazīd al-Mubarrad, *Mā ttafaqa lafḏu-hu wa-xtalafa maʿnā-hu min al-Qurʾān al-majīd*. Ed. ʿAbd al-ʿAzīz al-Maymanī. Cairo: al-Maṭbaʿa as-Salafiyya, 1350 A.H.

Munšī, *ʾAḍdād* = Muḥammad Jamāl ad-Dīn ibn Badr ad-Dīn al-Munšī, *Risālat al-ʾaḍdād*. Ed. Muḥammad Ḥusayn ʾĀl Yāsīn. Baghdād: Maktabat al-Fikr al-ʿArabī, 1985.

Quṭrub, *ʾAḍdād* = ʾAbū ʿAlī Muḥammad ibn al-Mustanīr Quṭrub, *Kitāb al-ʾaḍdād*. Ed. Hans Kofler. *Islamica* 3 (1931–1932) 241–284, 385–461, 493–544.

Ṣaġānī, *ʾAḍdād* = ʾAbū l-Faḍāʾil al-Ḥasan ibn Muḥammad aṣ-Ṣaġānī, *Kitāb al-ʾaḍdād*. Ed. Muḥammad ʿAbd al-Qādir ʾAḥmad. Cairo, 1989.

Sijistānī, *ʾAḍdād* = ʾAbū Ḥātim Sahl ibn Muḥammad al-Jušamī as-Sijistānī, *al-ʾAḍdād*. Ed. Haffner (1913:71–157).

Suyūṭī, *Muzhir* = Jalāl ad-Dīn ʾAbū l-Faḍl ʿAbd ar-Raḥmān ibn ʾAbī Bakr as-Suyūṭī, *al-Muzhir fī ʿulūm al-luġa wa-ʾanwāʿi-hā*. Ed. Muḥammad ʾAḥmad Jād al-Mawlā, ʿAlī Muḥammad al-Bijāwī, and Muḥammad ʾAbū l-Faḍl ʾIbrāhīm. Cairo: ʿĪsā al-Bābī al-Ḥalabī, n.d.

Tawwazī, *ʾAḍdād* = ʾAbū Muḥammad ʿAbdallāh ibn Muḥammad at-Tawwazī, *al-ʾAḍdād*. Ed. Muḥammad Ḥusayn ʾĀl Yāsīn. *al-Mawrid* 8 (1979) 161–190.

Secondary sources

ʾAḥmad, Muḥammad ʿAbd al-Qādir. 1989. Introduction to Ṣaġānī, *ʾAḍdād*.

ʾĀl Yāsīn, Muḥammad Ḥasan. 1974. *al-ʾAḍdād fī l-luġa*. Baghdad: Maṭbaʿat al-Maʿārif.

——. 1979. *Introduction* to Tawwazī, *ʾAḍdād*.

——. 1985. *Introduction* to Munšī, *ʾAḍdād*.

Bausani, Alessandro. 1971. “Osservazioni sugli *aḍdād* arabi”. *Actes du Ve Congrès international d’arabisants et islamisants, Bruxelles 31 août–6 septembre 1970*, 97–106. Bruxelles: Publications

du Centre pour l'Etude des Problèmes du Monde Musulman Contemporain.
Charnay, Jean-Paul (ed.). 1967. *L'ambivalence dans la culture arabe*. Paris: Anthropos.
Cohen, David. 1961. "*Aḍdād* et ambiguité linguistique en arabe". *Arabica* 8.1–29.
——. 1967. "Ambivalence, indifférence et neutralisation de sèmes: Post-scriptum sur le problème des *Aḍdād*. Charnay (1967:291–295).
Gardet, Louis. 1986. "*ḍidd, taḍmīn* et *taḫrīj, muqābal*". *Mélanges de l'Institut Dominicain d'Etudes Orientales du Caire* 17.67–80.
Kamal, Ribhi. 1967. "Le *tadʾâdd*". Charnay (1967: 51–64).
Khamash, Salim S. al-. 1991. Aḍdād, *a study of homopolysemous opposites in Arabic*. Ph.D. diss., University of Indiana.
Kofler, Hans. 1931–1932. "Anmerkungen". In Quṭrub, *ʾAḍdād* 385–461; 499–544.
Reig, Daniel. 1971. "Antonymie des semblables et corrélation des opposés en arabe". *Bulletin d'Etudes Orientales* 24.136–155.
Weil, Gustav. 1960. "Aḍdād". *Encyclopaedia of Islam*, I, 184. 2nd ed. Leiden: E.J. Brill.

Lidia Bettini (University of Florence)

Diglossia

In his 2001 inaugural speech at the 67th conference of the Arabic Language Academy (19 March–2 April), Šawqī Ḍayf, the president of the academy, openly accused the media of being carelessly oblivious, noting that → *fuṣḥā* is "the language of all the peoples of the *ʾumma* [*luġat šuʿūb al-ʾumma jamīʿan*]" whereas the *ʿāmmiyya* is the "daily language of a single people . . . the local language understood only by its people". He argued that the media has allowed the dialects to gradually but intrusively creep into domains that are traditionally reserved for *fuṣḥā* and eventually claim victory over it. Such intrusion needs to be stopped because it will eventually "dismantle the ties that bond the peoples of the *ʾumma*". Echoing these remarks, the Egyptian minister of higher education, Mufīd Šihāb, noted in his address to the conference participants that *fuṣḥā* is "presently suffering from some ailments", but then reassured them that the Arab Academy was and still is the armor of the Arabic language, "its impenetrable fortress, and its great minaret".

It is quite startling to see how pervasive and still prevalent the exaltation and professing of *fuṣḥā* as the sole unifying force of an otherwise politically and economically divided Arab world is, and how allegiance to 'perfect' *fuṣḥā* (*fuṣḥā salīma*) continues to be constructed as allegiance to the unity of the Arab world, its glorious Golden Age and magnificent heritage, when allegiance to any alliance or unity in the rest of the world is based on economic interests and political ties. It is equally startling to see how pervasive the view of diglossia as a problem and as a phenomenon linked to conflict and 'communicative tensions' between linguistic codes is in the scholarly literature on Arabic in both the Arab world and the United States of America. When discussing Arabic diglossia, a significant number of Arab intellectuals and researchers continue to describe the diglossic situation in terms of a crisis (*ʾazma*), a cause (*qaḍiyya*), or a clash (*ṣirāʿ*). Other scholars, mostly in the United States, construct diglossia as a social problem (Ferguson 1959; Meiseles 1980; Walters 1989; → colloquial). Discussing Arab and non-Arab researchers' response to Ferguson's characterization of diglossia and specifically summarizing European researchers' treatment of diglossia, Walters (1989:54), for example, states that European researchers "remind [us] that when languages, dialects, or varieties are in contact, they are almost always in a very real sense in conflict, both at the level of linguistic system and at the far less tangible level of their place within the society". These observations about diglossia are reminiscent of studies on "bilingualism and multilingualism [that] traditionally have been cast not only in popular belief but also in social and linguistic theoretical perspectives as anomalous, marginal, and in need of explanation" (Woolard 1999:3), obviously oblivious of "the idea that monolingualism [as] the human norm is a myth" (Thomason 2001:31).

It is not the purpose of this entry to conduct a detailed review of the profusely abundant literature on diglossia or to discuss the history of Arabic diglossia as a concept. The purpose is first to discuss some major directions that research in the Anglophone literature on Arabic diglossia has taken, focusing primarily on major developments that have taken place since the publication of Ferguson's (1959) inspirational but controversial article "Diglossia", and second to argue for a reconceptualization of the notion of diglossia. Such a reconceptualization foregrounds different questions that are in dire need of posing and certainly researching.

1. FERGUSON'S DIGLOSSIA AND INSIGHTFUL PREDICTIONS

In a recent issue of the *International Journal of the Sociology of Language* (2003), dedicated to the memory of Ferguson and devoted solely to 'Western' Arabist contributors, two major things are remarkable. First, the contributors credited Ferguson's account of diglossia for its 'prescience' (Belnap and Bishop 2003; Walters 2003), and 'grandeur' (Parkinson 2003), profusely but rightfully. Second, one cannot fail to acknowledge the increasing but cautious efforts to tread on the slippery section of the 'Arabic continuum' (e.g. Parkinson, Belnap, Eisele, and Walters) by focusing on the 'messy' area of diglossia. The 'slippery' and 'messy' area is what Ferguson broadly sketches as "relatively uncodified, unstable, intermediate forms of the language" in his classic paper on diglossia and what caught early research in a sterile debate.

Ferguson's "Diglossia" has become a rite of passage for scholars working on the current linguistic and sociolinguistic situation in the Arab world. His definition of diglossia has become the standard introduction to studies dealing with *izdiwājiyyat al-luġa* 'diglossia' in the Arab world. Ferguson's article has inspired and challenged scholars for four decades and seeded significant research on the linguistic and sociolinguistic changes in Arabic-speaking communities. He characterizes (1959:336) diglossia as:

> a relatively stable language situation in which, in addition to the primary dialects of the language (which may include a standard or regional standards), there is a very divergent, highly codified (often grammatically more complex) superposed variety, the vehicle of a large and respected body of written literature, either of an earlier period or in another speech community, which is learned largely by formal education and is used for most written and formal spoken purposes but is not used by any sector of the community for ordinary conversation.

The article, however, generated an equally strong reaction to his characterization of the linguistic situation in the Arabic-speaking communities (Badawī 1973; El-Hassan 1977; Blanc 1960; Meiseles 1980; Hary 1996). The strongest objections leveled are best summarized in Mahmoud (1986:239). He states:

> Many Arab and non-Arab scholars who have empirically studied the language behaviour of Arabic speakers (El-Hassan 1978, Blanc 1960, Mahmoud 1984, Mitchell 1978) have contended that Ferguson's description of the societal alternation between the two forms of Arabic tended to be too categorical and impressionistic and had overlooked the range of sociolinguistic variation en-countered in the speech of Educated Arabic speakers. They have also contended that the language situations Ferguson has cited as the exclusive domains of each variety are not as hermetically separated as he had thought (El-Hassan 1978:113–6). The emergence of a new, intermediate form of Arabic called Educated Spoken Arabic is commonly cited as evidence that the diglossic situation is undergoing a dramatic change (Abdel-Masih 1975; Bishai 1966; Mahmoud 1984; Mitchell 1962).

It is true that the linguistic situation in the Arab world is no longer (if it ever was) characterized by Classical Arabic/Modern Standard Arabic, on the one hand, and the various regional dialects, on the other. Ferguson's impressionistic and perhaps idealized characterization of the two varieties as being in complementary distribution functionally is removed from the reality of Arabic-speaking communities. It is idealized in the sense that it does not reflect the constant flux and ever-increasing leakage between the two varieties. Nor does it reflect the dramatic social changes that have taken place in the Arab world. "One may ask whether the seal between the two varieties had ever been hermetic", as Walters (1996a) rightly puts it. The linguistic situation in the Arab world has always been permeated by a state of linguistic flux due to the prolonged contact between the two varieties, on the one hand, and between Arabic and a foreign language (typically a former colonial language, e.g. French or English), on the other. The advocacy of universal education and the accessibility and knowledge of *fuṣḥā* have increased over the years. *Fuṣḥā* has increasingly ceased to be used restrictively by a privileged literate elite or to be known passively by a handful of illiterate people. Besides, social changes and growing literacy rates have called for new domains of use of both varieties, resulting in an increasing overlap between the two varieties of Arabic (and hence a significant leakage), and have induced changes in attitude toward the use of either variety in both the written (Daher 1999; Belnap and Bishop 2003) and the spoken mode (Parkinson 1996, 2003; Walters 1996, 2003; Boussofara-Omar 1999, *forthcoming*).

While some scholars may be right in their contentions, they were not perspicacious enough to recognize Ferguson's insightful predictions with regard to at least three major questions. The first prediction relates to the kinds of sociolinguistic changes that would take place (and have indeed taken place) across the Arab world. The second concerns the effects that changes in speakers' attitudes toward each variety may have on variety use as, for example, the emergence of the so-called third language, to which some scholars refer as → Educated (Spoken) Arabic. The third involves the consequences of those changes for the distribution of and differential access to varieties of Arabic.

In his characterization of diglossia, Ferguson (1959:10) argues that diglossia is a source of "communicative tensions [that] may be resolved by the use of relatively uncodified, unstable, intermediate forms of the language". In the case of Arabic he suggests that this is:

> a kind of spoken Arabic much used in certain semi-formal or cross-dialectal situations [which] has a highly classical vocabulary with few or no inflectional endings, with certain features of classical syntax, but with a fundamentally colloquial base in morphology and syntax, and a generous admixture of colloquial vocabulary.

Although impressionistic in nature, Ferguson's characterization of the intermediate forms of the Arabic language offers a significant working definition. It answers questions as to the modality of Arabic used (speaking vs. writing), when and where it is used (semi-formal and interdialectal settings), and what its linguistic features are (colloquial morphology and syntax but Classical Arabic and colloquial vocabulary). Although Ferguson did not relate his description of the intermediate forms to a theoretical linguistic model or offer a principled way of analyzing their nature, he first acknowledged their existence and then paved the way for future research. The majority of early studies inspired by Ferguson's characterization of diglossia have mainly focused on dividing the Arabic continuum into what they claimed would be 'clearly marked middle varieties'. The result was the emergence of a constellation of labels to categorize a tentative taxonomy of 'ill-defined' middle varieties of Arabic, and hence, a failure to articulate their description in a coherent manner or to relate these sets of practices to a theoretical linguistic model that can account for them.

2. A MIDDLE LANGUAGE OR INTERMEDIATE VARIETIES?

Over the last 15 years or so there has been a renewed interest in diglossia (Ferguson 1991; Haeri 2000, 2003; Hudson 1994, 2002; Kaye 2001; Walters 1989, 1996a, 1996b). At least three important linguistic journals have devoted complete issues to this topic (*Language* 1981; *Southwest Journal of Linguistics* 1991; *International Journal of the Sociology of Language* 2002 and 2003). Increasingly, considerable efforts are being made to redefine Ferguson's notion of diglossia, while taking into consideration the sociodemographic and socioeconomic changes that Arab countries have known over the last four decades or so. The new situation has given rise to differential access to and new uses of the varieties of Arabic, resulting in the emergence of 'intermediate forms' of Arabic – as Ferguson (1959) first characterized them – whose boundaries and contours were fuzzy. Because of their fuzziness and fluidity, they constituted a major challenge to all efforts devoted to delineating their boundaries and labeling them accordingly.

The middle language or intermediate varieties have been referred to, for the most part, as Educated (Spoken) Arabic. However, as Parkinson (2003:29) argues: "Everyone claims to believe that Educated Spoken Arabic is rule-governed, but none seems to be able to come up with the rules. Part of the reason for this may be, of course, that Educated Spoken Arabic may not actually *be* anything". Other terms that have been coined include → Middle Arabic (Ferguson 1959; Mahmoud 1978), urban cultivated Arabic (Abdulaziz 1986), interregional standard (Ibrahim 1986), elevated colloquial (Blanc 1960), and *luġat al-muṯaqqafīn*, the language of the educated (Badawī 1973). Some researchers identify a number of intermediate levels ranging from Modern Standard Arabic or a more traditional Classical Arabic to a plain colloquial, vernacular or colloquial of the illiterate (Blanc 1960; Badawī 1973; Meiseles 1980). Still others (El-Hassan 1978; Mitchell 1980; Mahmoud 1984) posit a single intermediate variety, Educated Spoken Arabic, which is, in Mitchell's (1980:13) words, "created and maintained by the constant interplay of written and vernacular Arabic."

Blanc's (1960) paper entitled "Style variation in Spoken Arabic: A sample of interdialectal educated conversation" is considered by El-Hassan (1978) and Mitchell (1986) a pioneering contribution to Educated Spoken Arabic despite its limitations. Blanc proposes five levels, Standard Classical, Modified Classical, Semi-literary or Elevated Colloquial, Koineized Colloquial, and Plain Arabic. He attributes stylistic modification to two major devices: leveling and classicizing. "But since the boundaries between the so-called 'leveling' and 'classicizing' devices are not clearly drawn in the first place . . . it is not clear how these processes are recognized", El-Hassan (1977:119) argues.

From a sociolinguistic perspective, Badawī (1973), unlike Blanc, identifies five levels within the continuum of contemporary Egyptian Arabic. In his analysis, Badawī identifies phonological, morphological, syntactic, and lexical features for each level. Switching is an upward or downward movement from one level to another. Badawī asserts that the speaker moves only from one level to the next (i.e., one level at a time), a claim that Walters (1989:99) characterizes as unclear. He states:

> I have difficulty with the number and characterization of Badawī's levels – I can never decide if the schema in fact applies to style of language use within a speaker (cf. Joos 1967), levels of class-related style (cf. Labov 1966b), or a possible taxonomy for spoken utterances and written language.

Badawī's contention that "these registers do not have clear, permanent boundaries between one another" leaves us with unanswered questions and unclear understandings of how he delineated the boundaries between the varieties that he argues constitute the Arabic continuum.

In his attempt to study variation in contemporary Arabic, Meiseles (1980) proposes four varieties: Literary Arabic or Standard Arabic, Oral Literary Arabic, Educated Spoken Arabic, and Basic or Plain Arabic. He states:

> Intermediate between the two varieties or sets of varieties, relatively 'pure' Classical and Colloquial, there are many shadings of 'middle language.' These intermediate forms, some highly fluctuating and transitional, others more stable, represent these two tendencies: classicization, in which a dialect is modified in the direction of classical, and koineization, in which dialects are homogenized by the modification or elimination of features which are felt to be especially distinctive of a particular dialect.

Within the Tunisian context, the third language is also characterized by many labels, including *la troisième langue* 'third language' (Garmadi 1965), *le troisième registre* 'third register' (Attia 1966), simplified Arabic, elevated or polite dialect (Garmadi 1968), Educated Arabic (Maamouri 1973), Middle Arabic (Chaib 1976), and Educated Tunisian Arabic (Walters 1989). In his characterization of this variety of Arabic, Maamouri (1973:17) writes:

> Its morphology is simplified because it does away with all the inflectional case endings. Its syntax is that of T[unisian] A[rabic] since it adopts, in most cases, the SVO structure and deletes the dual, the feminine plural, and similar unnecessary distinctions . . . Its lexicon is almost equally divided between regular common MSA words, which replace TA synonyms, and TA loan words which do not exist in MSA.

In his description of the third register, Attia (1966) adds some 'features' to the ones Maamouri reported, for example the use of a syllable configuration that does not exist in Classical Arabic, the use of a single Tunisian Arabic relative marker rather than the declined Classical Arabic/Modern Standard Arabic marker, the use of a Tunisian Arabic negation marker, and the use of Classical Arabic lexis. In addition to these features, Chaib (1976) notes others including the use of case endings in 'ossified' expressions, the loss of glottal stop in all cases, and the use of the accusative/genitive cases regardless of function.

Previous research on the alternating use of the two varieties of Arabic does not offer a framework to handle what Meiseles (1980:120) calls the "uncharted sea of intermediate shades, whose overall picture is one of a state of flux". The efforts are predominantly impressionistic in nature and do not provide a coherent theoretical model to predict and explain the structural constraints on the mix of the two varieties of Arabic in a principled manner. "The only possible conclusion", Meiseles (1980:120–21) points out, "is that outlining borders to the different Arabic varieties is not only a very difficult task, but one that seems, *prima facie*, unnecessary and superfluous".

3. NOTION OF INTERMEDIATENESS

The nature and degree of the mix of the two varieties give rise to judgments about the forms produced: whether they belong to *fuṣḥā* or dialect and, more importantly, where they fall on the continuum. The speaker/hearer is said to be able to judge whether intermediate forms are more dialect-like forms than *fuṣḥā* forms and vice versa. Western scholars have underlined native speakers' intuitive perception, and acute awareness of the diglossic nature of their language. Schmidt states (1974:10):

> Although native speakers of Arabic tend to perceive their speech and the speech of others as discrete CA [Classical Arabic] or EC [Egyptian Colloquial], they are able to make judgments, in some cases finely detailed, about intermediate forms and they can arrange these forms into hierarchies.

This statement is echoed by Parkinson (1993:91) who argues that "speakers [of Arabic] themselves are aware of the source of their linguistic material, and can tell you if a particular lexical item, grammatical pattern, or even vowel marker, is dialectal or *fuṣḥā*". In a somewhat similar vein, Walters (1996:404) notes that "speakers of Arabic have consistent intuitions about which forms are H[igh] and which are L[oud], and these intuitions involve more than merely partitioning the lexicon into categories". Parkinson (2003:29) reiterates the same claim stating: "Each specific element of the item is recognized by native speakers as being *fuṣḥā* or colloquial or both". This evidence by no means indicates that languages/varieties are discrete, or "homogeneous, static systems, with a minimum of variation or none at all" (Meiseles 1980:121). It is "unrealistic to assume that all members of a speech community share the same language rules, notably at the level of performance", as El Hassan (1977:117) rightly points out. But one wonders whether researchers are not, in a way, "positively forced to ask [themselves]" (Parkinson 1993:70), not just what Classical Arabic, Modern Standard Arabic, and the dialect are, but what they are for users. In their perception, judgment, and rating of what constitutes Classical Arabic, Modern Standard Arabic, the dialect, and switching patterns, do the phonological, morphological, syntactic, or lexical 'flags' play the same or differential roles? What the statements seem to indicate is that there is some kind of consensus about the native speakers' consistent ability to linguistically differentiate between mixed forms, *fuṣḥā* forms, and dialectal forms, despite their apparent fluidity and elusiveness. Three factors may explain the consistency of native speakers in their judgment of the wide constellation of those mixed forms. The first is the non-randomness in the mix of the two varieties. The alternating use of *fuṣḥā* and dialect falls within the scope of structural constraints on → code-switching. The second factor relates to the nature of combinations of the linguistic levels (whether it is phonological, morphological, syntactic, and/or lexical). The third concerns the degree of their combination (i.e., one or more of those linguistic levels are used at a time). The rating and the hierarchical arrangement of the intermediate forms are dependent upon the nature, degree, and scope of combination of levels.

The early studies on middle language or intermediate varieties led to several conclusions. First, they suggest that these intermediate forms (some highly fluctuating and transitional, others more stable) represent two main tendencies. One tendency consists in elevating the dialect in the direction of the prestigious Classical Arabic/Modern Standard Arabic. The other consists in simplifying Classical Arabic/Modern Standard Arabic in the direction of the dialect. Second, previous studies focused on delimiting boundaries for some variety that is neither *fuṣḥā* nor dialect. Third, the description of the middle varieties did not offer a coherent framework to understand the nature of the mix between the two varieties in the sense that the findings were not articulated in a principled manner or in terms of a theoretical model.

In their efforts to explore the nature of diglossia, earlier studies have tended either to produce orderly, neat, and 'clean' taxonomic descriptions of Arabic diglossia (e.g. Ferguson's taxonomic chart of domains of use of each variety), some of which implicitly or explicitly convey the impression that 'domains of use' of either variety of Arabic "never change, or merge" (Haeri 2003:66), while others tended to describe it in terms of rivalry, tension, and conflict (Ferguson 1959; Stetkevytch 1970; Meiseles 1980). Even in their attempt to understand the changing nature of diglossia, researchers have sought to divide the Arabic spectrum into clearly marked middle varieties, a task that proved to be "superfluous and unnecessary" (Meiseles 1980:121), because it focused solely on the discrete, the homogeneous, and the coherent in language and neglected "the contingent, the hybrid, the contested, and the performed" (Walters 2003:79).

Upon revisiting his 1959 article "Diglossia" and evaluating the overabundant studies that arose after its publication, Ferguson (1991:215) was disheartened to see that "most descriptions of register variations including [his] own, are static descriptions that fail to examine the phenomena of register switching and negotiation of meanings by register variation within a social interaction" (1991:229). He does not fail to reiterate (1991:215) his initial goals:

> I hoped other people would write articles on other clear cases to develop a fairly elaborate taxonomy of language situations. Ultimately, that taxonomy would be replaced by some set of principles or frame of reference in terms of which this kind of thinking about language and this kind of research should be done. My goals, in ascending order, were clear: clear case, taxonomy, principles, theory.

Research on Arabic diglossia has attained the first two goals but seems to be grappling with the 'principles' and 'theory' prongs. Some other studies (Eid 1980, 1988; Walters 1996, 2003; Boussofara-Omar 1999, 2003, *forthcoming*) were conducted in order to go beyond the efforts to divide the Arabic spectrum into 'clearly marked' middle varieties or to produce mere taxonomies. These efforts were an attempt to provide a coherent theoretical framework to understand the nature of the structural mix between the two varieties. This involves using and applying the concept of switching which, in Parkinson's (2003:29) view, "has added a much needed correction to the lens through which we have viewed the diglossic situation", despite his uncertainty that it "has been any more successful than the multiglossic track in characterizing the mixed varieties in a precise, complete and insightful way".

4. THE CODE-SWITCHING APPROACH

Scarce were the efforts to analyze the 'third language' and 'middle varieties' in terms of patterns of code-switching, i.e., switching between *fuṣḥā* and dialect, in the literature on the 'mix' of Classical Arabic/Modern Standard Arabic and dialects. Eid (1982, 1988) was among the first voices to call for a code-switching approach to Arabic diglossia. In the conclusion of her study on Standard Arabic and Egyptian Arabic, Eid (1980:84) states:

> The results obtained from this study then show that switching between varieties of the same language does not proceed randomly; it is governed by a principle sensitive to three types of information: the position of the switch, the type of focal point involved at the switch position, and the variety from which the focal point is drawn.

What is more important to note is her call for further research along the code-switching lines:

> The results also indicate that there is a significant relationship between the kind of switching that takes place between varieties of the same language and that which occurs between different languages in the speech of bilinguals – a relationship that warrants further study in future research.

Walters (1996a, 1996b) was the first to use the label 'Arabic diglossic switching' to characterize the alternating use of Classical Arabic/Modern Standard Arabic and the dialect. Following Walters and in light of Myers-Scotton's original (1993) and refined (Myers-Scotton and Jake 2000, 2001) versions of the Matrix Language Frame Model, Boussofara-Omar (1999, 2003, forthcoming) demonstrates that what researchers have termed 'third language' or 'intermediate varieties' in fact constitute diglossic switching. She also argues that there is no conventionalized variety known as 'third language' or Educated Spoken Arabic. What is being conventionalized, however, are patterns of switching between the two varieties of Arabic where the dialect serves as the matrix variety in which constituents from *fuṣḥā* are embedded. Unlike other researchers (Versteegh 1997:194) who believe that "since language choice [in Arabic] takes place on a continuum, [the] changes do not take the form of code-switching from one variety to another, but manifest themselves in a larger percentage of features from the opposite variety", Boussofara-Omar demonstrates that Arabic diglossic switching is more than a 'Low' variety sprinkled with lexical items, frozen expressions or proverbs from the 'High' variety and involves far more motivations than simply elevating or lowering the level of one's text or speech. It is shaped by morphophonological processes and constrained by grammatical rules. The analysis of the morphophonological processes and the morphosyntactic constraints seems to suggest that in diglossic switching there is a significant interplay between the two varieties, which is quite different from code-switching. Research focused on the nature of the interplay between the morphophonological processes and the morphosyntactic constraints in a diglossic situation is expected to provide a better understanding of diglossic switching.

5. PARKINSON'S STATISTICAL AND VARIATIONIST APPROACH

In his efforts to characterize the middle section of the Arabic continuum, Parkinson (2003) argues for the use of a combination of the statistical and variationist approaches. In his estimation, attempts to define the continuum functionally have been only partially successful. The issues, observations, and major findings remain the same as those reported in the early continuum studies or those based on the code-switching approach. What differs is their enunciation and articulation. Parkinson, however, makes an interesting point in the conclusion to his study:

> I put it out as an unsupported opinion that although there are many performances where a clear matrix and embedded language can be surmised, it may be, in the case of a speaker with a good command of both MSA and colloquial, and with a lifetime habit of mixing the two, that there is simply no matrix for him . . . Thus we get the verb forms with one or more features from both ends as if there were no problem, leading us from intrasentential to intralexemic codeswitching, with no apparent effort to follow the "rule" that system morphemes all come from the matrix language.

Parkinson's concluding remarks on the use of system morphemes from both varieties (and hence the absence of a matrix language/variety) call for comment. Attempts to apply Myers-Scotton's Matrix Language Frame model and its sub-models to Arabic diglossic switching (Boussofara-Omar 1999) gave rise to two sets of problematic patterns which the model neither predicts nor provides a satisfactory explanation for. One set involves utterances in which one finds grammatical morphemes from both varieties (cases similar to Parkinson's) within a single Complement Phrase. The second involves utterances in which the word order and subcategorization rules are those of one variety (e.g. Tunisian Arabic) but system morphemes come from the other variety (i.e. *fuṣḥā*). Boussofara-Omar (2003) revisited these problematic sets in light of Myers-Scotton's latest refinements of her model in order to provide an explanation for their occurrence and to understand better the "messiness of the middle" (Parkinson 2003:40) and limitations of the Matrix Language Frame model. The findings suggest that the complex interactions between *fuṣḥā* and dialect cannot be merely framed within either the Fergusonian idealized paradigm or the vague continuum notion. Nor can they be simplistically characterized as elevating or lowering one's speech. The 'mix' is socially motivated and structurally constrained like 'classic' code-switching (to borrow Myers-Scotton's term) between any language pairs. It is also slightly nuanced because it is a language contact situation. The findings provide possible venues for investigating the nuances between 'classic' code-switching and switching in a language contact situation, for rethinking our conceptualization of the diglossic situation, and for investigating it in light of theoretical models that allow us to articulate our findings in a more principled and coherent manner.

6. RECONCEPTUALIZING DIGLOSSIA

The conceptualization of the coexistence of languages/varieties within a speech community in terms of rivalry, clash, tension, conflict, and constraints alone ignores their fluidity, downplays the dynamically 'positive' nature of the mutual impact on each other, and disallows any effort to explore the conditions under which the languages come together naturally, either through speech or context, and the complex patterns and configurations of use that arise out of their coexistence.

This entry calls for a shift from studying diglossia as a "relatively stable situation" (Ferguson 1959:336) to diglossia as sets of practice. The ways in which members of a community use language as well as their beliefs about language varieties and their ways of speaking shift and change. In the present transnational, globalist world discrete languages/varieties or homogeneous ideologies are far less likely to be wholly discrete. Furthermore, discussions of Arabic varieties, or language varieties in general, need not be based solely on formal linguistic grounds. Socially-based factors/motivations play as significant a role as structural constraints. And the interplay between linguistic constraints and social motivations has been neglected so far.

There has been little, if any, discussion of how Arabic varieties may have influenced one another in some kind of dynamic way. The general tendency has been to perceive and describe the influence as unidirectional, i.e. from *fuṣḥā* to the dialect, in other words from 'High variety' to 'Low variety'. In her analysis of the (socio)linguistic situation in Egypt, Haeri (2003:82) aptly captures the argument when she states: "There has been an ideological interest in not professing the profound actual and potential influences of Egyptian Arabic in the development of a contemporary Classical Arabic". New domains for the use of colloquials are generally perceived as intrusion (Belnap and Bishop 2003:20) into the domains of Modern Standard Arabic although the constant leakage and the great overlap between the two varieties of Arabic appear to have facilitated the flow both ways. The practice of switching is increasingly gaining ground and may eventually give rise to a conventionalized spoken standard Arabic that is dialectal in its underlying structure but *fuṣḥā* in its surface realization. Boussofara-Omar (1999, *forthcoming*) has demonstrated one way in which the grammar of the dialect is influencing *fuṣḥā* in Tunisian speech. Tunisians, as well as speakers of other varieties of Arabic (Egyptian, Syrian, Lebanese), accept as part of *fuṣḥā* syntactic patterns that are unambiguously dialectal but which occur with *fuṣḥā* lexical items and *fuṣḥā* system morphemes.

In the case of Arabic diglossia, the relentless efforts to reinforce the sacred and divine origin of *fuṣḥā* together with the majestic aura in which it is – and must continue to be – shrouded do not allow room for studies of this type. The exaggerated focus on the high reverence that Arabs have for *fuṣḥā*, its perfection and purity of speech or eloquence (*faṣāḥa*), remain as widely prevalent and advocated as they were in the pre-Islamic era. But users have taken an active part in ushering Classical Arabic/Modern Standard Arabic into new "domains of use" (Fishman 1972) and new "spheres of activity" (Bakhtin 1981). Perhaps because Classical Arabic is deeply embedded in the divine, the holy, and the sacred, research on the significance and implications of the 'modernization' of a sacred language or its appropriation by the state is scarce, if not non-existent. That Modern Standard Arabic is the modernized version of Classical Arabic seems to be a *fait accompli* since the frequent use of the term has allowed scholars to take the intricacy/complexity of the interconnectedness between the politics of modernization of Classical Arabic, secularization of the state, and its appropriation of a sacred language for granted.

In reconceptualizing diglossia, a dialectical approach allows an understanding of the processes by which 'boundaries' between Classical Arabic/ Modern Standard Arabic and the various dialects are negotiated, redefined, redrawn, and reproduced by social factors. This approach is "not an antagonistic alternative to the study of systems or structures but a necessary complement to it" (Ortner 1984:147, cited in Wedeen 2002:720).

Diglossia need not necessarily be seen as a problem to be solved, denied, or contained; it is a richness that is often dramatically undervalued. Diglossia can be viewed as "zones of contact" or "zones of relatedness" (Pratt 1987) in which both varieties are continually and dynamically shaping and reshaping each other, while creating new arenas for subsequent interaction, zones of convergence and divergence. Such a perspective foregrounds the workings of dynamic interaction and exchange among varieties of Arabic (instead of terminologies of intrusion and code superiority) and allows for "the contingent, the hybrid, the contested and the performed" (Walters 2003:79) to be explored.

Bibliographical references

Attia, A. 1966. "Quelques registres de l'emploi de l'arabe en Tunisie". *Revue Tunisienne des Sciences Sociales* 8.115–149.

Badawī, as-Saʿīd Muḥammad 1973. *Mustawayāt al-ʿarabiyya al-muʿāṣira fī Miṣr*. Cairo: Dār al-Maʿārif.

Bakhtin, Mikhail M. 1981. *The dialogic imagination: Four essays by M.M. Bakhtin*, ed. Michael Holquist. Austin: University of Texas Press.

Belnap, Kirk R. and Brian Bishop. 2003. "Arabic personal correspondence: A window on change in progress". *International Journal of the Sociology of Language* 163.9–25.

Blanc, Haim. 1960. "Stylistic variation in spoken Arabic: A sample of interdialectal educated conversation". *Contributions to Arabic linguistic*s, ed. Charles A. Ferguson, 81–156. Cambridge, Mass.: Harvard University.

Boussofara-Omar, Naima. 1999. *Arabic diglossic switching: An application of Myers-Scotton's MLF Model*. Ph.D. diss., University of Texas, Austin.

——. 2003. "Revisiting Arabic diglossic switching in light of the matrix language frame model and its submodels: The 4-M and the Abstract Level models". *Bilingualism: Language and Cognition* 6.1.33–46.

——. Forthcoming. "Neither third language nor middle varieties but Arabic diglossic switching". *Zeitschrift für Arabische Linguistik*.

Chaib, M. 1976. "al-ʿArabiyya al-wusṭā wa-mā našaʾa fīhā min tadāxul bayna l-fuṣḥā wa-l-ʿāmmiyya". *Revue Tunisienne des Sciences Sociales* 46.47–66.

Daher, Jamil. 1999. "Written Arabic of personal letters". Paper presented at the Thirteenth Annual Symposium on Arabic Linguistics. Stanford, Calif., March 6.

Eid, Mushira. 1982. "The non-randomness of diglossic variation in Arabic". *Glossa* 16:1.54–84.

——. 1988. "Principles for code-switching between Standard and Egyptian Arabic". *al-ʿArabiyya: Journal of the American Association of Teachers of Arabic* 21.51–80.

El-Hassan, Shahir A. 1977. "Educated Spoken Arabic in Egypt and the Levant: A critical review of diglossia and related concepts". *Archivum Linguisticum* (n.s.) 8:2.112–132.

Ferguson, Charles A. 1959. "Diglossia". *Word* 15. 325–340.

——. 1991. "Diglossia revisited". *Southwest Journal of Linguistics* 10:1.214–234.

Fishman, Joshua A. 1972. "The relationship between micro- and macro-sociolinguistics in the study of who speaks what language to whom and when". *Sociolinguistics*, ed. John B. Bride and Janet Holmes, 15–32. Harmondsworth: Penguin.

Haeri, Niloofar. 2000. "Form and ideology: Arabic socio-linguistics and beyond". *Annual Review of Anthropology* 29.61–87.

——. 2003. *Sacred language, ordinary people*. New York: Palgrave Macmillan.

Hary, Benjamin. 1996. "The importance of the language continuum in Arabic multiglossia". *Understanding Arabic: Essays in contemporary Arabic linguistics in honor of El-Said Badawi*, ed. Alaa Elgibali, 69–90. Cairo: American University of Cairo Press.

Hudson, Alan. 1994. "Diglossia as a special case of register variation". *Sociolinguistic perspectives on register*, ed. Douglas Biber and Edward Finegan, 294–314. New York: Oxford University Press.

——. 2002. "Outline of a theory of diglossia". *International Journal of the Sociology of Language* 157.1–48.

Kaye, Alan S. 2001. "Diglossia: The state of the art". *International Journal of the Sociology of Language* 152.117–129.

Maamouri, Mohammed. 1973. "The linguistic situation in independent Tunisia". *American Journal of Arabic Studies* 1.50–56. (Repr. *Language in Tunisia*, ed. Richard Payne, 12–22. Tunis: Bourguiba Institute of Modern Languages, 1983.)

Mahmoud, Youssef. 1986. "Arabic after diglossia". *The Fergusonian impact: In honor of Charles A. Ferguson on the occasion of his 65th birthday*. I. *From phonology to society*, ed. Joshua Fishman a.o. 239–251. Berlin: Mouton de Gruyter.

Meiseles, Gustav. 1980. "Educated spoken Arabic and the Arabic language continuum". *Archivum Linguisticum* (n.s.) 11.118–148.

Mitchell, T.F. 1986. "What is educated spoken Arabic?". *International Journal of the Sociology of Language* 61.7–32.

Myers-Scotton, Carol. 1993. *Duelling languages: Grammatical structure in codeswitching*. Oxford: Oxford University Press. (2nd ed. 1997.)

Ortner, Sherry. 1984. "Theory in anthropology since the sixties". *Comparative Studies in Social History* 26.126–66.

Parkinson, Dilworth B. 1994. "Speaking *fuṣḥā* in Cairo: The role of ending vowels". *Arabic sociolinguistics: Issues and perspectives*, ed. Yasir Suleiman, 179–211. Richmond, Surrey: Curzon.

——. 2003. "Verbal features in oral *fuṣḥā* in Cairo". *International Journal of the Sociology of Language* 163.27–41.

Pratt, Marie-Louise. 1987. "Linguistic utopias". *The linguistics of writing: Arguments between writing and literature*, ed. Nigel Fabb a.o., 48–66. Manchester: Manchester University Press.

Thomason, Sarah. 2001. *Language contact*. Washington, D.C.: Georgetown University Press.

Versteegh, Kees. 1997. *The Arabic language*. Edinburgh: Edinburgh University Press.

Walters, Keith. 1989. *Social change and linguistic variation in Korba, a small Tunisian town*. Ph.D. diss., University of Texas, Austin.

——. 1996a. "Diglossia, linguistic variation, and language change". *Perspectives on Arabic linguistics*, VIII, ed. Mushira Eid, 157–197. Amsterdam and Philadelphia: J. Benjamins.

——. 1996b. "Intrasentential codeswitching in diglossic settings and its implications for linguistic variation and change". *Sociolinguistic variation: Data, theory, and analysis: Selected papers from NWAV23 at Stanford*, ed. Jennifer Arnold a.o., 401–416. Stanford, Calif.: CSLI.

——. 2003. "Fergie's prescience: The changing nature of diglossia in Tunisia". *International Journal of the Sociology of Language* 163.77–109.

Wedeen, Lisa. 2002. "Conceptualizing culture: Possibilities for political science. *American Political Science Review* 96:4.713–728.

Woolard, Kathryn A. 1999. "Simultaneity and bivalency as strategies in bilingualism". *Journal of Linguistic Anthropology* 8:1.3–29.

Naima Boussofara-Omar
(University of Kansas)

Diminutive

The diminutive is a morphological pattern which expresses diminution, reduction, or lessening. The common Arabic terms for diminution and diminutive are *at-taṣġīr* or *al-ism al-muṣaġġar*, and *at-taḥqīr* and *al-ism al-muḥaqqar* for a pejorative/deteriorative/contemptuous meaning which is sometimes implicit in the form. The diminutive may also be used to express endearment or charity (e.g. Wright 1896:166; Fleisch 1961:380–381, 392; Fischer 2002:51).

Semitic languages, including Arabic, present series of noun patterns, i.e. substantives and adjectives, which express diminution. Brockelmann (1928) has the most comprehensive collection, and he shows that the majority of noun patterns in it hold not only a diminutive meaning but the opposite, i.e. an augmentative meaning as well (named elsewhere 'enhancement'; e.g. Wright 1896:166; Fischer 2002:51; see also Fleisch 1961:390–391). The patterns collected by Brockelmann may be classified into four main pattern groups with respect to their morphological nature.

i. Patterns involving inner vowel change *qutāl* and *qutayl*, and variants in various Semitic languages. These exhibit one or more cases of vowel shortening or change, consonantal geminating, and addition of a semivowel between the first and second consonants instead of the second and third, e.g., *qutal, quttal, quttāl, qaytal, qittawl, qattūl* (Brockelmann 1928:109–117, 1908:351, on diminutives in *qutāl*, 1908:352–353; Fischer 2002:51–53, on diminutives in *qutayl*; Fleisch 1961: 378–380, on *qutāl* and *qutayl*, and 380–381, for references to discussions of Arab medieval grammarians on the meaning of *qutayl*). More on *qutāl* for diminutives in Semitic languages is given in Fox (2003:229–235). Diem (1970:61–65) argues that *qutāl* does not really have a diminutive meaning in Arabic. On the form *qutayl* for diminutives in Semitic languages, see also Moscati (1964:78) and Lipiński (2001:219). Wright (1896:167) states that the *qutayl* pattern might occasionally be pronounced in Arabic with *kasra*, i.e. *qitayl*, when the second consonant of the root is *-y-*. Barth (1894: 312–315) indicates that the form *qutayl* has a diminutive meaning in Arabic, Biblical Hebrew, and Aramaic, but not in Ethiopic and Amharic. Arabic is generally considered the richest Semitic language with respect to diminutive words of this type. Nonetheless, von Soden (1991) shows that the number of Akkadian instances that are probably derived from the pattern *qutayl* is significantly greater than previously thought.

   The main Arabic diminutive pattern is *qutayl*, and it is the only one treated in many Arabic grammars, e.g. *kulayb* from *kalb* 'dog', *ʿuyayd* from *ʿīd* 'holiday', *buyayt/ buwayt* from *bayt* 'house', *jubayl* from *jabal* 'mountain'; *qulayʿa* from *qalʿa* 'fortress', preserving the feminine suffix; *hunayda* from *hind* 'Hind [proper name]' and *šumaysa* from *šams* 'sun', both with the addition of a feminine suffix according to the feminine gender of the original noun; and without such an addition in *ḥurayb* from *ḥarb* 'war'; *ʾumayma* from *ʾumm* 'mother', extracting two consonants from a geminate one and adding a feminine suffix; *ḥumayrāʾ* from *ḥamrāʾ* 'red', preserving the original feminine suffix *āʾ* according to Fischer (2002:52), but replacing it with the regular feminine ending according to Wright (1896:174); *sukayrān* from *sakrān* 'drunk', preserving the suffix *-ān*; *buṣayriyy* from *baṣriyy* 'someone from Basra', preserving the *nisba* suffix*-iyy*. Moreover, in Arabic, in order to fit the above mentioned triconsonantal patterns, biconsonantal roots are extended by the addition of a semivowel, e.g. *dumayy* from *dam* 'blood', *luġayya* from *luġa* 'language'; by creating a semivowel for a two-consonant form originally with a prosthetic *ʾalif*, e.g. *bunayy* from *ibn* 'son'; or by restoring a semivowel that might appear in certain forms of the word, e.g. *ʾubayy* from *ʾab* 'father', *ʾuxayya* from *ʾuxt* 'sister'. Likewise, forms with more than three consonants or additional suffixes, such as feminine, dual, etc., may be reduced by the omission of vowels and consonants in a process called *tarxīm*, e.g. *ʿuṭayf* from *miʿṭaf* 'cloak', *ḥumayd* from *ḥāmid* 'Hamid [proper name]', *suwayd* from *ʾaswad* 'black'. This tendency to adjust roots that are too short or too long to a triconsonantal pattern is a perfect example of the morphologically inherent need in any Semitic language, as explained by Goldenberg (1994), to make a Semitic root enter a three-place pattern. However, Arabic forms with more than four consonants or long vowels frequently employ the secondary patterns *fuʿaylil* or *fuʿaylīl* if the basic form has a long vowel in the last syllable, e.g. *ʿunaykib* from *ʿankabūt* 'spider', involving a consonant omission as well, *šuwayʿir* from *šāʿir* 'poet' (unlike other *fāʿil* forms, which are shortened to *fuʿayl* like *ḥāmid* above), or *ṣunaydīq* from *ṣundūq* 'trunk'. For these and similar instances, and for phonetic and morphological variants, see Wright (1896:166–175), Fleisch (1961:380–389), and Fischer (2002:51–53).

ii. Patterns constructed by duplication of the last consonant or the second and last consonants together (Brockelmann 1928:117–120, 1908:366–367; Moscati 1964:79; Lipiński 2001:221 indicates a diminutive meaning of such forms only in Hebrew for names of colors). Arabic instances are, e.g., *baʿrūr* from *baʿr* 'camel dung', *ḥubrūr, ḥibrīr, ḥabarbar*, and *ḥuburbūr* from *ḥubārā* 'bustard' (these and more in Brockelmann 1908:366–367, 1928:117).

iii. Patterns produced linearly by the addition of special suffixes, some of which are limited to a specific Semitic language or group of languages, while others are common to several Semitic languages, e.g. *-āy*, *-īt* (mostly Tigrē, Brockelmann 1928:120–121, also suggested by Lipiński 2001:230). This *-āy* appears as *-ē* in Amharic according to Brockelmann (1928:121–122), *-ō* (mostly Ethiopic, Brockelmann 1928:122–124). *n* and *l* suffixes, the *n* combined with vowels, i.e. *-an, -ān, -ōn, -ūn*, and *-īn*, appear in several Semitic languages, including Arabic (according to Brockelmann 1928:124–128, 1908:395–396, 402–403; Moscati 1964:82; according to Lipiński 2001:227–229, *-ān* and *-ōn* are attested in Arabic), Arabic instances of which are *ʿuqbūl* 'reminder of illness/pimples on the lips after fever', *ʿaqrabān* 'small scorpion' (however, many such Arabic instances presented in Brockelmann 1908:394, 1928:124–126, are actually *qutayl* patterns affixed by the adjectival *-n* endings, which is part of the original pattern, as suggested by Fleisch 1961:386). Suffixes *-ōs, -ūs* occur in Aramaic, and were also borrowed by Arabic (Brockelmann 1908:395, n. 2, 1928:128), e.g. *qudmūs* 'old', *qarqūs* 'bald plain'.

iv. Patterns produced linearly by adding the feminine *-t/at* and in Arabic *tāʾ marbūṭa* suffix (several Semitic languages including Arabic, according to Brockelmann 1908:420, 1928:129–131). Arabic instances are *ṭāḥūna* 'small mill' from *ṭāḥūn* 'mill', and *rajul rabʿa* 'a man of average build', involving a pejorative meaning.

In Semitic languages, including Arabic, expression of the diminutive was not restricted to nouns but expanded into other forms. Leslau (1945) presents rare instances of a verbal diminutive pattern *qaytala*, occasionally with phonological modifications, in a few modern Arabic dialects, in South Arabian, and in a small number of instances in Ethiopian languages. He suggests that this verb form with *-y-* was developed by analogy to the nominal forms containing *-y-*, *qutayl* and *quttayl*. Johnstone (1973) describes diminutive patterns in the Modern South Arabian languages, among them diminutives produced from adverbs (Johnstone 1973:99). In Classical Arabic, diminutives can be formed from demonstrative and relative pronouns, certain prepositions, and a few verbs of surprise or wonder. These forms usually take the form *qutayl* or employ only its main element *ay*, e.g. relative pronouns: *hāḏayyā* from *hāḏā* 'this' or *al-llaḏayyā* from *al-laḏī* 'which'; prepositions/adverbs: *qubayl* from *qabla* 'before', *buʿayd* from *baʿda* 'after'; verbs of surprise or wonder: *mā ʾuḥaysina-hu* from *mā ʾaḥsana-hu* 'how good he is' (Wright 1896:167, 100, regarding all; Fleisch 1961:368, for demonstrative and relative pronouns; Reckendorf 1921:214, 221 and Fischer 2002:153, for prepositions, which Fischer 2002:171 regards as adverbs). In addition, Fleisch shows that diminutives are produced not just from singular nouns but from certain internal plural forms as well, e.g., *ʾukaylib* from *ʾaklub* 'dogs' (Fleisch 1961:386). Fischer (2002:65) indicates that the plural forms of the *qutayl* diminutives are inflected full plurals only. He also mentions that a small number can be expressed by the plural diminutive, e.g. *sunayyāt* from *sana* 'year', expressing 'several years' (Fischer 2002:65).

The use of diminutives has been productive in Arabic all along. New words in old and new diminutive patterns appear in various Modern Arabic dialects exhibiting phonological changes accordingly. One example of the vitality of diminutives in a modern Arabic dialect is found in Masliyah (1997), who presents a large variety of diminutives of several patterns in spoken Iraqi Arabic. The *qtēl* and *qtēlīl* patterns for three- and four-consonant roots respectively are the most productive in the Iraqi dialect. They are actually similar to the Classical Arabic *qutayl* and *qutaylīl* in involving similar inner vowel change, albeit adjusted to phonological changes of vowel shortening in the first syllable and the contraction *ay* > *ē* in the second syllable. In addition, Iraqi Arabic occasionally employs for diminutives linear patterns with suffixes that might denote diminutive value, i.e. the feminine *tāʾ marbūṭa* and the suffixes *-ūn, -ān, -āya/-ya, -iyya*. Another marginal option represented in Iraqi Arabic is the repetition of a root letter of the basic word. Since it is found only with biconsonantal roots, it should probably be interpreted as a means of expanding the word to fit a triconsonantal pattern rather than a real diminutive duplication. Sometimes strategies are combined, both suffixes and inner vowel changes being used to create diminutives. Finally, Iraqi Arabic marginally employs patterns with inner vowel change other than *qtēl* and *qtēlīl*, like *qutal, qattūl*, *qtāl* (*qutāl* in Classical Arabic). Also, it uses more extensively, though only for endearment of proper names, an additional pattern *qattūli*, constructed by inner vowel change and the *nisba* suffix attached, with no equivalent in Classical Arabic.

Bibliographical references

Barth, Jacob. 1894. *Die Nominalbildung in den semitischen Sprachen*. Leipzig: J.C. Hinrichs. (Repr. Hildesheim: G. Olms, 1967.)

Brockelmann, Carl. 1908. *Grundriss der vergleichenden Grammatik der semitischen Sprachen*, I. Berlin: Reuther und Reichard.

——. 1928. "Deminutiv und Augmentativ im Semitischen". *Zeitschrift für Semitistik und Verwandte Gebiete* 6.109–134.

Diem, Werner. 1970. "Die Nominalform *fuʿāl* im klassischen Arabisch". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 120.43–68.

Fischer, Wolfdietrich. 2002. *A grammar of Classical Arabic*. 3rd. rev. ed., trans. Jonathan Rodgers. New Haven and London: Yale University Press.

Fleisch, Henri. 1961, *Traité de philologie arabe*, I. Beirut: Imprimerie Catholique.

Fox, Joshua. 2003. *Semitic noun patterns*. Winona Lake, Ind.: Eisenbrauns.

Goldenberg, Gideon. 1994. "Principles of Semitic word-structure". *Semitic and Cushitic studies*, ed. Gideon Goldenberg and Shlomo Raz, 29–64. Wiesbaden: O. Harrassowitz. (Repr. Gideon Goldenberg, *Studies in Semitic linguistics*, 10–45. Jerusalem: Magnes Press, 1998.)

Johnstone, Thomas M. 1973. "Diminutive patterns in the modern South Arabian languages". *Journal of Semitic Studies* 18.98–107.

Leslau, Wolf. 1945. "Sur le diminutif verbal en sémitique". *Word* 1.277–280.

Lipiński, Edward. 2001. *Semitic languages: Outline of a comparative grammar*. 2nd ed. Leuven: Peeters.

Masliyah, Sadok. 1997. "The diminutive in spoken Iraqi Arabic". *Zeitschrift für Arabische Linguistik* 33.68–88.

Moscati, Sabatino. 1964. *An introduction to the comparative grammar of the Semitic languages*. Wiesbaden: O. Harrassowitz.

Reckendorf, Hermann. 1921. *Arabische Syntax*. Heidelberg: C. Winter.

Soden, Wolfram von. 1991. "Deminutiva nach der Form *qutail*, *qutīl* und vergleichbare vierkonsonantige Bildungen im Akkadischen". *Semitic studies in honor of Wolf Leslau on the occasion of his eighty-fifth birthday, November 14th, 1991*. ed. Alan S. Kaye, II 1488–1492. Wiesbaden: O. Harrassowitz.

Wright, William. 1896. *A grammar of the Arabic language*, I. 3rd ed. Cambridge: Cambridge University Press.

TAMAR ZEWI (University of Haifa)

Diphthongs

1. Diphthongs in Classical Arabic and its dialects

In the Semitic linguistic domain a vowel + glide (*w* or *y*) compound is called a diphthong. Its Arabic name has the same meaning: *ṣawt murakkab* 'compound sound'. The hypothetical Proto-Semitic diphthongs, **aw*/**ay*, according to the generally accepted view, are conserved in Old Arabic (Cantineau 1960:102), but this conservation is not at all surprising, since Proto-Semitic phonology has been reconstructed mainly from Classical Arabic, as interpreted by comparatively late sources. Early papyri testify that *aw/ay* are preserved and not contracted (Hopkins 1984:17). Long diphthongs resulting from the elision of intervocalic *hamza* are often shortened: *ʿāʾiša* > *ʿāyša* > *ʿayša* [fem. proper name]. The only possible contraction seems to be *ay* > *ā*: *duwābba* < *duwaybba* (Fleisch 1961:69). The same *-ay* > *ā* change may be observed in the prepositions *ʾilā*/*ʿalā* when used separately, but it remains *ay* in connected forms: *ʿalayka*. Brockelmann (1908:I, 90) supposes an **ʿalaya* as the starting point of this shift: *ʿalaya* > *ʿalay* > *ʿalā*. The → *ʾimāla* may have played a decisive role in the monophthongization of *ay/aw* in the hollow verbs, **bayaʿa* > *bayʿa* > *bēʿa* > *bāʿa*. In some modern dialects we can see the same process: *šayʾ* > *šēʾ* > *šāʾ*; *lēh* > *lāh* (Fayyūm, Egypt) (ʾAnīs 1995:66–67).

Words ending in *-ā* (written with *ʾalif maqṣūra*) became diphthongized in some Classical Arabic dialects. An example from Sībawayhi is *ʾalif* > *ay (ey) (ʾafʿā* > *ʾafʿey)* by *ʾimāla* (only in pause, not in context; quoted in Al-Nassir 1993:94). In the Ḥijāz, in some Qays dialects *ā* > *ay* (Rabin 1951:116; Sībawayhi, *Kitāb* II, 349): *ʾafʿay*, *ḥublay*, and some place names. This led Birkeland (1940:76) to consider the *yāʾ* a pausal spelling and to maintain that the context form could only have been *-ē* (becoming later *-ā*).

On the other hand, there are Classical Arabic dialects which had monophthongization in prepositions where the norm is a diphthong: *ay* > *ā*. The Banū l-Ḥāriṯ ibn Kaʿb (North Yemen) had *ʿalāhā/ʾilāhā/ʾilāka/ladāka* (Rabin 1951:65, referring to Ibn Fāris, *Ṣāḥibī*). Ḥijāzī *yawjal* had a *yājal* reflex in the dialect of the Qays, and a *yījil* in that of the Tamīm *(yiw > yiy > yī)* (Jundī 1983:II, 576–577, referring to al-Farrāʾ's *Maʿānī l-Qurʾān*). The word occurs in the *Qurʾān* (15/53). ʾAbū Ḥayyān (*Baḥr* V, 458) mentions this special reading: *qālū lā tawjal, wa-quriʾa lā tājal bi-ʾibdāl al-wāw ʾalifan kamā qālū tāba fī tawba* "They said: 'Fear not!' This is also read as *lā tājal* replacing the *wāw* with the *ʾalif*, just as they said *tāba* instead of *tawba*" (cf. Jundī 1983:II, 578; other cases are mentioned in Åkesson 1996:30).

The monophthongization of the prepositions *ʾilā/ʿalā* is explained by Bravmann (1977:103–104) *(ʿalā < ʿalayhi)* by the process *ʿalay > ʿalē > ʿalā*. According to Bravmann's theory, there had been *ay > ē/aw > ō* changes word-finally, while the diphthongs remained word-medial, e.g. in the construct state of the dual *-ay (yadayhi)*. Then, *ē/ō* secondarily changed to *ā*. Monoyllabic *ay/aw* remained non-contracted: *law*, *ay*, *kay*. Bravmann (1977:105) also explains the change in the forms of III*w/y* ('weak') verbs by positing the context form **ramaya* (like *qatala*) and the pausal form **ramay* (like *qatal*), which then became *ramay > ramē > ramā*. The same with *w*: *-awa/aw > ō > ā*. Probably, word-final *-aw* was contracted to *-ā* at an earlier period than *-ay*, as testified by the Arabic script (Bravmann 1977:106).

Since *w* and *y* are radical consonants, it is doubtful whether there are real diphthongs at all in Arabic. According to Fleisch (1961:67) the words *ṯawb* and *ġayb* demonstrate that they contain real diphthongs by their behavior in contracted forms, since real consonants are impossible at this place in the syllable structure: *ṯawbu bakrin > ṯawbbakrin*; *ġaybu bakrin > ġaybbakrin* (where the diphthongs stand in place of a long vowel in the syllable structure).

The main source of Arabic diphthongs in the verbal system is the declension of the so-called → weak verbs, verbal roots with final *w/y*, where the semivowel is deleted between short and/or long vowels, and the remaining vowels form diphthongs according to the general phonological rule of Arabic that two vowels cannot follow each other (since no syllable may begin with a vowel): *ġazaw* (<*ġazaū* < *ġazawū*); *tansayna* (< *tansaīna* < *tansayīna*). During this morphological process those of the potential diphthongs which have no equivalent in Classical Arabic are monophthongized: *uw/iy* > *ū/ī*; *uy/ūy* > *ī/īy* (= *iyy*) (Cantineau 1960:85–88; Fleisch 1961: 125ff.). Arab medieval grammarians explained this *-aw* in *daʿaw* and other verbal forms on the basis of the following series of shifts: *kataba* + *w* > *katabaw* > *katabuw* > *katabū* (Bohas and Guillaume 1984:30, 291; see also the review of this by Versteegh 1989). They may have done so because in the contemporaneous vernaculars similar forms had already developed. The same phenomenon takes place in some modern Iraqi dialects (see below). On the other hand, the *uw* > *ū* change may be conceived of as a writing convention as well (the two forms are written in exactly the same way).

Secondary diphthongs may have come into being by double consonants merging with the final weak radicals in some verbal stems: *taḏannantu* > *taḏannaytu* (Roman 1983:361, referring to Sībawayhi, who states that these phenomena, though acceptable, were *laysa bi-muṭṭarid* 'not in general use'). In the modern dialects, this phenomenon has become general with a further monophthongization: *raddayt* > *raddēt*.

Notwithstanding the conservatism of the Proto-Semitic diphthongs in Arabic, the frequency of the diphthongs in Classical Arabic texts is very low compared to that of the vowels. In a Qurʾānic text (Q. 2/5–11) chosen at random, the diphthongs *ay* and *aw* occur only twice, compared to 202 short and long vowels (Fleisch 1968:16).

2. Modern Arabic dialects

Modern Arabic dialects are characterized by an overall shift to monophthongization: *aw* > *ō* > *ū*; *ay* > *ē* > *ī*. There are two main systems: one with five long vowels (mainly in the east) and the other with three without diphthongs (mainly in the west).

The monophthongization of the Classical Arabic diphthongs in modern dialects is so widespread that their conservation counts as an exception, for example in some Lebanese dialects. It is not known when this monophthongization happened. Blau (1966: ch. 7.2) thinks that it had already occurred in early Middle Arabic, but for Diem (1985:76) this is not supported by the evidence. He believes that in the 1st century A.H. there were as yet no signs of this monophthongization, on the basis of the evidence of Greek transliterations of Arabic names, and by the treatment of Arabic loanwords.

Monosyllabic words with final diphthongs form a group in themselves in most dialects: *law*, *saw*, *šey*, *ḍaw*, *ḥay* (Bravmann 1977:102). In addition to the normal usage there is also, however, *šē* in the Libyan dialects as a variation. Feghali (1919:83) confirms the existence of diphthongs in Lebanese, e.g. *-aw* in *mawtna* and *-ay* in *layle*. The former proves that there are (phonetically) real diphthongs in the modern dialects, since three consonants could not possibly occur together. Diphthongs after pharyngeal *ḥ*/ʿ are conserved in almost all modern dialects: *ʿayn*/*ḥayṯ* (Cantineau 1960:104). Otherwise, *aw*/*ay* have remained only in some dialects: Sfax, Mauritania, Lebanon, and in women's dialects in Tunis and Sūsa (Fischer and Jastrow 1980: 54–56).

In the Eastern Arabic dialects most of the hypothetical Old Arabic diphthongs have been monophthongized: *ay* > *ē*, *aw* > *ō*. In the Arabian Peninsula, however, diphthongs are not rare, although there are dialects totally lacking them, e.g. Inner Oman (Fischer and Jastrow 1980: 103). In most parts of Mesopotamia there are five long vowels and no diphthongs at all (Fischer and Jastrow 1980:140–141). In the Muslim dialect of Baghdad, however, there are diphthongized verbal forms of the type *ḍurbaw* (Fischer and Jastrow 1980:152). The Syro-Palestinian area is characterized by monophthongization, but in the Nuṣayri region, *aw*/*ay* – *ā* alternation is found as well; *ay* > *ā* (in open syllable) > *ay* (in closed syllable): *bayt* > *bāt* > *bayti* (Fischer and Jastrow 1980:176). Lower Egyptian dialects primarily monophthongize the Old Arabic diphthongs (Fischer and Jastrow 1980: 213). The Old Arabic diphthongs may be further altered by the so-called 'fracture', e.g. in Iraq: **zayn* > *zēn* > *z(i)yēn* (Cantineau 1960:105). There is a special type of partial conservation of the diphthongs with a slight articulation of the glide: *ay* > *ē*^{y}; *aw* > *ō*^{w} in the speech of some North Arabian nomads (Cantineau 1960:105).

In the Western Arabic dialects Ph. Marçais (1977) distinguishes three types of diphthongs: (a) short vowel + glide, mainly after the pharyngeals ʿ, *ḥ*: *ay*/*aw*/*ey*/*ew*/*ow*; (b) long vowel + glide, of non-Classical origin as a mutation of the phonetic character: *āw*/*āy*, *ēw*/*ēy (āwnu*/*lāyži)*; and (c) secondary diphthongization by 'fracture', when a long vowel breaks into two elements: *žūʿ*/*žōʿ* + *žuaʿ*; *mšēt*/*mšīt* + *mšēyt* (Ph. Marçais 1977:15–19). In Algeria and Morocco, mainly **aw* > *ū* and **ay* > *ī* are found; in Libya and south Tunisia **aw* > *ō*, **ay* > *ē*; in south Algeria, in the speech of the camel breeding nomads *aw*/*ay* and *ō*/*ē* occur (Fischer and Jastrow 1980:37). Not infrequently, the conservation of the diphthongs becomes partial: *ay* > *ē*^{y}; *aw* > *ō*^{w} (some North Arabian nomads); in the speech of the Saharan nomads *e*^{i}/*o*^{u}: *be*^{i}*ḍa*/*mo*^{u}*t*; in the northern parts of Algeria *mšē*^{y}*t*/*fō*^{w}*g* (Cantineau 1960:105; W. Marçais 1908:32, 40, 41). There is also a secondary diphthongization in North Africa, when ancient *ū*/*ī* is diphthongized: *ʿūd* > *ʿawd*, as well as a secondary diphthongization of the type *āy*/*āw*: *bqāu*, *mēida* (Cantineau 1960:105).

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

ʾAbū Ḥayyān, *Baḥr* = ʾAbū Ḥayyān Muḥammad ibn Yūsuf al-ʾAndalusī al-Ġarnāṭī, *at-Tafsīr al-kabīr al-musammā al-Baḥr al-muḥīṭ*. 5 vols. Beirut: Dār ʾIḥyāʾ at-Turāṯ al-ʿArabī, 1990.

Ibn Fāris, *Ṣāḥibī* = ʾAbū l-Ḥusayn ʾAḥmad ibn Fāris, *aṣ-Ṣāḥibī fī fiqh al-luġa wa-sunan al-ʿArab fī kalāmihā*. Ed. Muṣṭafā aš-Šūʾaymī. Beirut: Badran, 1964.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān, *Kitāb Sībawayhi*. Ed. Hartwig Derenbourg. Paris: Imprimerie nationale, 1881–1889.

SECONDARY SOURCES

Åkesson, Joyce. 1996. "Conversion of the *yāʾ* into an *alif* in Classical Arabic". *Zeitschrift für Arabische Linguistik* 31.27–33.

Al-Nassir, Abdulmunim Abdulamir. 1993. *Sibawayh the phonologist: A critical study of the phonetic and phonological theory of Sibawayh as presented in his treatise* al-Kitab. London and New York: Kegan Paul International.

ʾAnīs, ʾIbrāhīm. 1995. *Fī l-lahajāt al-ʿarabiyya*. Cairo: al-Maktaba al-ʾAnglū-Miṣriyya.

Birkeland, Harris. 1940. *Altarabische Pausalformen*. Oslo: J. Dybwad.
Blau, Joshua. 1966. *A grammar of Christian Arabic*. Louvain: Secrétariat du Corpus SCO.
Bohas, Georges and Jean-Patrick Guillaume. 1984. *Étude des théories des grammairiens arabes*. I. *Morphologie et phonologie*. Damascus: Institut Français de Damas.
Bravmann, Meier Max. 1977. *Studies in Semitic philology*. Leiden: E.J. Brill. [3. "Some aspects of the development of Semitic diphthongs", 98–123, originally appeared in *Orientalia*, N.S. 8 (1939) 244–260; 8 (1940) 45–60.]
Brockelmann, Carl. 1908. *Grundriss der vergleichenden Grammatik der semitischen Sprachen*. I. *Laut- und Formenlehre*. Berlin: Reuther und Reichard.
Cantineau, Jean. 1960. *Cours de phonétique arabe*. Paris: Klincksieck.
Diem, Werner. 1985. "Die Monophthongisierung der Diphthonge *ay* und *aw* im frühen Neuarabisch". *Zeitschrift für Arabische Linguistik* 14.76–78.
Feghali, Michel T. 1919. *Le parler de Kfarʿabīda (Liban-Syrie)*. Paris: E. Leroux.
Fischer, Wolfdietrich and Otto Jastrow (eds.). 1980. *Handbuch der arabischen Dialekte*. Wiesbaden: O. Harrassowitz.
Fleisch, Henri. 1961. *Traité de philologie arabe*, I. Beirut: Imprimerie Catholique.
——. 1968. *L'arabe classique: Esquisse d'une structure linguistique*. Beirut: Dar el-Machreq.
Hopkins, Simon. 1984. *Studies in the grammar of early Arabic*. Oxford: Oxford University Press.
Jundī, ʾAḥmad ʿAlam ad-Dīn al-. 1983. *al-Lahajāt al-ʿarabiyya fī t-turāṯ*. 2 vols. Tripoli and Tunis: ad-Dār al-ʿArabiyya li-l-Kitāb.
Marçais, Philippe. 1977. *Esquisse grammaticale de l'arabe maghrébin*. Paris: Librairie d'Amérique et d'Orient, Adrien-Maisonneuve.
Marçais, William. 1908. *Le dialecte arabe des Ūlād Brāhīm de Saïda (Département d'Oran)*. Paris: H. Champion.
Rabin, Chaim. 1951. *Ancient West-Arabian*. London: Taylor's Foreign Press.
Roman, André. 1983. *Étude de la phonologie et de la morphologie de la koiné arabe*, I. Aix-en-Provence: Université de Provence.
Versteegh, Kees. 1989. "'Early' and 'late' grammarians in the Arab tradition: The morphology of the hollow verbs". *Zeitschrift für Arabische Linguistik* 20.9–22.

Tamás Iványi (Budapest, Hungary)

Diptosis

1. Introduction

Arabic is a language of case inflections (→ declension). The majority of nouns have three cases: /u/ for the nominative, /i/ for the genitive, and /a/ for complements. This three-case system, or triptosis, contrasts with a two-case system, in which nouns have /u/ for the nominative, and /a/ for the other functions. The most important categories of diptotic nouns in Classical Arabic are (Fleisch 1961:271–276):

i. proper names: names of foreign origin (e.g. *ʾibrāhīmu* 'Abraham'); names of cities and regions (e.g., *baġdādu* 'Baghdad'); names ending in *tāʾ marbūṭa*, both male and female (e.g. *ṭalḥatu* 'Talha', *fāṭimatu* 'Fatima'); all female proper names except those of the pattern *fa/i/uʿl* (e.g. *daʿdu* 'Daʿd', but *hindun* 'Hind'); all proper names with the same pattern as the verbal perfect or imperfect (e.g. *yazīdu* 'Yazid'); all proper names with the ending *-ānu* (e.g. *ʿuṯmānu* 'Uthman'); all compounds (e.g. *ḥaḍramawtu* 'Hadramawt')
ii. common nouns: all broken plurals in patterns containing four consonants (e.g. *madārisu* pl. of *madrasa* 'school', *mafātīḥu* pl. of *miftāḥ* 'key', *salāṭīnu* pl. of *sulṭān* 'sultan'); nouns ending in /-ā'u/ (*ʿaḏrāʾu* 'virgin', *ḥamrāʾu* 'red [fem.]'), including plurals with the same ending (e.g. *wuzarāʾu* 'viziers'); adjectives ending in /-ānu/ (e.g. *ġaḍbānu* 'angry'); the → elative *ʾafʿalu* (e.g. *ʾakbaru* 'bigger'); → numerals when used independently (e.g. *niṣfu sittata* 'half of six').

According to the Arab grammarians, diptosis was the result of the loss of one case-ending in certain nouns deviating from the default case of ordinary nouns. They attributed this loss to the accumulation of properties that differed from the default case, such as feminine gender, proper names, foreign origin, or verbal pattern. Whenever two or more of these properties co-occurred in a noun (e.g. a feminine noun of foreign origin, or a proper name with a verbal pattern; for a list of these *mawāniʿ aṣ-ṣarf* 'factors preventing declension' see Carter 1982:74–77), they assumed the noun became 'less declinable' *(ġayr munṣarif)* and was no longer completely free in its movement *(ġayr mutamakkin)*. As a result, its genitive ending became identical with the accusative ending (→ *ṣarf*).

Diptosis is not completely unknown in other Semitic languages (cf. Moscati 1954, 1958; on possibly diptotic endings in Ugaritic see Gordon 1955:45) and may even be present in some → Afro-Asiatic languages. According to some theories, Proto-Semitic originally had both a triptotic and a diptotic declension. Brockelmann (1908:

461; cf. Fleisch 1961:278; Kienast 2001:142; Baerman 2005:817, n. 10), for instance, believed that the diptotic endings are etymologically different from the normal case-endings. In his view, the diptotic endings started with personal names formed with verbal patterns ending in /-u/ or /-a/, of the type *yazīd-u* 'Yazid' and *šammar-a* 'Šammar', the former modeled on the imperfect, the latter on the perfect verb. These endings were later reinterpreted as part of a declensional system, and other categories of nouns were added to this group by analogical extension, e.g. nouns of the pattern *ʾafʿalu*, possibly because of their already having the ending /-a/ in predicative function, as in *kāna ʾaswad-a* 'he was black'.

Others maintain that Proto-Semitic originally had a two-case system. Lipiński (1977:254–259) compares this two-case system with ergative systems in other languages (cf. Kienast 2001:179–180). In his view, Proto-Semitic shared with Libyco-Berber two endings, /-u/ for the 'active' case, and /-a/ for the non-active or predicative case, the genitive ending /-i/ being a later development, related to the gentilitial suffix /-iyy/ (cf. Petráček 1981). A recent proposal by Baerman (2005) analyzes the diptotic declension as a case of syncretism and connects the diptosis of the diptotic nouns with the diptosis of the sound masculine plural {/-ūna/ /-īna/} by a symmetric rule.

At the level of the historical languages, the diptotic endings in Arabic seem to be closely associated with proper names. In the present entry the hypothesis will be presented that in early Arabic the diptotic declension was still the only set of endings, whereas the triptotic declension represents a later extension of the system (cf. Roman 1996, 2001).

2. Redundancy of the case ending /-i/

Historically, the case ending /-i/ has two functions in Arabic: it is applied to complements and adjuncts of the verb that are introduced by a preposition, and it is applied to attributive elements (complements of the noun, or the amplifying nouns in construct state nominals). It is always redundant. In the former use, the preposition heading the complement phrase is sufficient semantically. Nevertheless, the case ending cannot drop off, since case inflections result from the general organization of the language, which makes them basic functional category markers.

In the latter use, the attributive element (the amplifying noun) to which the case ending is attached replaces the morpheme of non-specific place /-n/, called → *tanwīn*, and it is thereby clearly marked. When an indefinite constituent is without /-n/, then the following constituent is the substitute for /-n/ and is the attribute of the first constituent. For example:

(1) *kalb - u - n* 'dog from a non-specific place' = 'a dog'
(2) *kalb - u daʿd - a* 'Daʿd's dog'
(3) *kalb - u šayx - i - n* '[the] dog of a sheikh'

The case inflection applied to the attributive element was initially the vowel /a/ as in (2). The vowel /i/ as in (3) seems to have been brought about by the prepositions themselves, which made preposition-headed complements, {X [. . .] -/a/}, seem different from prepositionless complements, {Ø [. . .] -/a/}, so that the former became {X [. . .] -/i/}. The vowel /i/, having the same characteristics as /u/ and /a/, took the place of /a/ as a marker of prepositional complements and also as a marker of attributive elements that seemed special due to their place in the *tanwīn*, /-n/, paradigm.

In the framework chosen here, Arabic syntax was originally a two-case system. This early diptosis left a number of disparate traces in the Classical language.

3. Common nouns and proper names

In early Arabic, both concrete and abstract common nouns with a triconsonantal root were built on the pattern R1V1R2R3, represented by /fvʿl/, with no vowel between R2 and R3. In this pattern, V1, the first vowel, between the first root consonant R1 and the second root consonant R2, denoted animacy. All common nouns, being capable of forming a non-locational phrase with *tanwīn* /n/, were indeterminate, for example *kalb-u-n* 'a dog'.

Proper nouns are genealogical. Their primary function is to assert a human being's membership of a social group. In addition to personal names, Arabic treated and, in some cases, still treats as proper nouns certain names of places, animals, numbers, vocatives, for example *daʿd-u*

'Daˁd', *ṯalāṯat-u niṣf-u sittat-a* 'three is half of six', *yā ʾamīr-u* 'oh emir!'.

The nominal system had no way of creating a form for the proper noun that differed from that of the common noun. Both had to be built on a three-consonant root, and both had to code animacy. Thus, common nouns for living beings include *bakr* 'young she-camel', *kalb* 'dog', *šayx* 'old man'; and proper nouns include *bakr*, *kalb*, *taym*, *zayd*.

The difference between the proper noun and the common noun therefore had to be marked syntactically. The syntax marked this crucial difference as follows: since the proper noun could not be coupled to *tanwīn* when it referred to a member of a community, it was the very absence of *tanwīn* that, by evoking the community, turned the common noun into a proper noun.

However, this way of marking the difference precluded proper nouns from the attributive relation. This was due to the loss of the relation with *tanwīn*, which is the very head of the paradigm of attributive elements, and the only indefinite element of this paradigm. Indeed, to use a proper noun as the basis for a construct state nominal would render it indistinguishable from a common noun. Other dependent constructions therefore had to stand in for the attributive construction, for example: *zayd-u bn-u daˁd-a* 'Zayd son of Daˁd', where the proper noun *zayd*, identified as a proper noun by the absence of *tanwīn*, is the base for the apposition *ibn*, 'son', thus avoiding the → construct state *zayd-u daˁd-a*, 'Daˁd's Zayd'.

Nonetheless, a construct state nominal can be based on a proper noun where the context makes it clear that it must be a proper noun, for example *yā taym-u taym-a ˁadiyy-i-n* 'o Taym, Taym [descendants] of ˁAdi!', where *taym* is clearly a proper noun, given the form of direct address.

Diptosis, then, is characterized by two archaic features: the two-case inflection system and the absence of *tanwīn*. These two features have stuck together, so that if a diptote noun bears the definite article, /al-/, or if it bears an attributive element, thus forming a construct state nominal, it becomes triptote. For example:

nominative accusative genitive
{*salāṭīn-u salāṭīn-a salāṭīn-a*} 'sultans'
{*as-salāṭīn-u as-salāṭīn-a as-salāṭīn-i*} 'the sultans'
{*salāṭīn-u d-dawlat-i salāṭīn-a d-dawlat-i salāṭīn-i d-dawlat-i*} 'the sultans of the dynasty'

4. The origin of diptosis

The diptotic paradigm of forms without *tanwīn* stems from proper nouns. Proper nouns retained their special status as long as the language kept the same noun formation rules and as long as the speech community took its proper names solely from the set of animate nouns. Thus, *kalb* became a proper noun without disturbing the system: *kalb-u daˁd-a* could only be 'Daˁd's dog'; *kalb-u bn-u daˁd-a* could only be 'Kalb, Daˁd's son'.

The break came with the loss of the regular noun formation, giving rise to *rajul* 'man', *baˁīr* 'camel', *ḥimār* 'donkey', *ḥūt* 'fish', *ˁaqrab* 'scorpion', *ḥubārā* 'bustard', etc. Released from the noun formation constraints, Arabic-speakers were free to take their proper nouns from their environment, for example *ḥasan-u-n* 'beautiful', *diˁbil-u-n* 'seasoned camel', *ṭalḥat-u* 'acacia'.

The loss of the noun formation rules allowed nicknames to become attributes, for example *zayd-u l xayr-i* 'Zayd of the best'. These attributes, being non-genealogical, contributed to the loss of markedness of proper nouns.

Despite the evolution of the language, however, some elements continued to adhere to the diptotic system because they were associated with a set of formal features. In the historical language, diptotes are proper nouns, common nouns, and adjectives. Diptosis in these elements seems to be due either to the form being alien to its paradigm – it is redeployed outside its paradigm – or to the form being alien to the Arabic language.

i. proper names: examples of native Arabic formations are /fuˁal/, /ʾafˁal/, /faˁˁal/, /yafˁvl/; redeployed outside their paradigms are: *ˁumar-u*, *ʾaḥmad-u*, *šammar-u*, *yazīd-u*. This last name, *yazīd-u* 'Yazid', and several similar names, have traditionally been interpreted as redeployments of the 3rd person masculine singular of the imperfect verb /yafˁvl/. In fact, /y/ preceding the three root consonants is probably an old em-phasis marker. Examples of proper nouns of foreign origin include *ʾibrāhīm-u* 'Abraham', *dāʾūd-u* 'David', *baˁlabakk-u*, 'Baalbek'. When the Arabic noun formation system still adhered to its characteristic pattern, the different structure of foreign names indicated the foreign origin of the men who bore them.

ii. proper names suffixed by /-at/, /-ā/, /-āʾ/, /-ān/ and /-ūn/. Proper names with the suffixes /-at/, /-ā/, and /-āʾ/ occur in the historical language as feminine suffixes. The suffix /-ān/, by contrast, occurs, though not regularly, as a masculine suffix. The suffix /-ūn/ occurs in later usage as amasculine proper noun suffix. Examples of male names bearing a masculine suffix are *marwān-u*, *saʿdūn-u*. Examples of names, male and female, bearing a feminine suffix are *ṭalḥ-a-t-u*, *salw-ā*, *zahr-āʾ-u*.

iii. common nominal plurals, traditionally known as 'quadrisyllabic'; these are also quite distinctive. The singular forms are made up of four consonants: root consonants, R, or non-root consonants, C, which can be represented as: {R1R2R3-C} – {C-R1R2R3} – {CCCC}, for example *sulṭān-u-n*, pl. *salāṭīn-u* (R1aR2āR3īC) 'sultan' (from the root *s-l-ṭ*); *maṯlaṯ-u*, pl. *maṯāliṯ-u* (CaR1a̱R2iR3) 'group of three' (from the root *ṯ-l-ṯ*); *yaʿsūb-u-n*, pl. *yaʿāsīb-u* (CaCāCīC) 'male bee, drone; chief' (from the root *y-ʿ-s-b*). Adding the suffix /-t/ to the form /CaCāCiC-u/ produces a triptotic pattern: /CaCāCiCa-t-u-n/, for example *tilmīḏ-u-n*, pl. *talāmīḏ-u*, *talāmiḏa-t-u-n* 'pupil', *tarjumān-u-n*, pl. *tarājim-u*, *tarājima-t-u-n* 'interpreter'. Note that plurals in which the fourth consonant is one of the unstable consonants /w/ or /y/ are sometimes treated as three-consonant elements, and sometimes as four-consonant elements: in the first case they are triptotic, in the second case diptotic, for example *fatw-ā* 'fatwa, judicial consultation', whose plural is either *fatāw-ā* for each case, or *fatāw-i-n* (< **fatāwiy-u-n*) for the nominative, *fatāw-i-n* (< **fatāwiy-i-n*) for the genitive, *fatāwiy-a* for the accusative; in *fatāwiy-a*, the consonant /y/, although unstable, has been maintained because it provides a linking segment between the syntagmatic vowel /i/ and the case ending /a/; it is thus stabilized. This means that the accusative form is treated as a four-consonant form and therefore a diptotic form, while the other two case forms are treated as three-consonant forms and therefore as triptotic. Another 'quadrisyllabic' plural is that of singular adjectives which follow the /ʾafʿal/ pattern and are treated as nouns, for example *ʾakābir-u* '[the] great ones [of the world]', pl. of *ʾakbar-u*, from the root *k-b-r*. Other diptotic plurals are those with feminine suffixes which are plurals of singulars treated as adjectives. They follow the patterns /faʿl-ā/, /fuʿal-āʾ/ and /ʾafʿil-āʾ/, for example *qatīl-u-n*, pl. *qatl-ā* 'killed (ones)', *ʿālim-u-n*, pl. *ʿulam-āʾ-u* 'scholar', *ṣadīq-u-n*, pl. *ʾa-ṣdiq-āʾ-u* 'friend'. All these plurals are nouns.

iv. Finally, some adjectives have become diptotic by analogy with proper nouns, having the same pattern or the same affix, for example *ʾaswad-u* 'black', *sawd-āʾ-u* 'black [fem.]'; *našw-ān-u* 'drunk [masc.]', *našw-ā* 'drunk [fem.]'.

Plural and dual number are denoted by affixes which bear, by conditioning, a superficial resemblance to original diptosis. The masculine plural is signaled by just two affixes: /-ū(na)/ for the nominative and /-ī(na)/ for the genitive/accusative. Likewise, the dual is signaled by two affixes: /-ā(ni)/ for the nominative and /ay(ni)/ for the genitive/ accusative. These two numbers use the three long vowels of Arabic and, lacking a fourth long vowel, use the sequence of vowel /a/ plus consonant /y/ as a diphthong. This is a departure from the vowel system of the language, which has no → diphthongs.

5. Conclusion

The origin of diptosis must be sought in proper nouns. Proper nouns – genealogical names – name human beings as belonging to a community. Linguistically, original Arabic proper nouns indicated genealogical provenance by detaching from *tanwīn* /n/, the marker of non-specific place. Once its origin was forgotten, diptosis spread to common nouns and adjectives bearing a formal resemblance to the nicknames that had become proper nouns.

Bibliographical references

Baerman, Matthew. 2005. "Directionality and (un)natural classes in syncretism". *Language* 80.807–827. (Also available at <http://muse.jhu.edu/journals/language/ v080/ 80.4baerman.pdf>.)

Brockelmann, Carl. 1908. *Grundriss der vergleichenden Grammatik der semitischen Sprachen.* I. *Laut- und Formenlehre*. Berlin: Reuther und Reichard. (Repr. Hildesheim: G. Olms, 1961.)

Carter, Michael G. 1982. *Arab linguistics: An introductory classical text with translation and notes*. Amsterdam: J. Benjamins.

Fleisch, Henri. 1961. *Traité de philologie arabe.* I. *Préliminaires, phonétique, morphologie nominale.* Beirut: Imprimerie Catholique.
Gordon, C.H. 1955. *Ugaritic manual.* Rome: Pontificium Institutum Biblicum.
Kienast, Burkhart. 2001. *Historische semitische Sprachwissenschaft.* Wiesbaden: O. Harrassowitz
Kuryłowicz, Jerzy. 1951. "Le diptotisme et la construction des noms de nombre en arabe". *Word* 7.222–226.
Lipiński, Edward. 1997. *Semitic languages: Outlines of a comparative grammar.* Leuven: Peeters. (2nd ed., 2001.)
Moscati, Sabatino. 1954. "Sulla flessione nominale dell'arabo classico". *Rivista di Studi Orientali* 38.131–160.
——. 1958. "On Semitic case-endings". *Journal of Near Eastern Studies* 17.142–144.
Petráček, Karel. 1981. "Le système de l'arabe dans une perspective diachronique". *Arabica* 28.162–177.
Roman, André. 1996. "Sur l'origine de la diptosie en langue arabe". *Studies in Near Eastern languages and literatures: Memorial volume of Karel Petráček*, ed. Petr Zemánek, 515–534. Prague: Academy of Sciences of the Czech Republic, Oriental Institute.
——. 2001. *La systématique de la langue arabe.* 2 vols. Kaslik: Université Saint-Esprit de Kaslik, Faculté des lettres, Département d'Interprétation et de Traduction.

André Roman
(Université Lumière – Lyon 2 – CRTT)

Discourse Analysis

1. Introduction

The field of discourse analysis is very heterogeneous and covers a wide range of language studies not only by linguists, but also by scholars from other disciplines such as sociology, rhetoric, and anthropology. For this reason the term 'discourse analysis' has often been assigned a variety of meanings. However, it can reasonably be defined as an area of inquiry dealing with language use rather than language as an abstract system of rules and it focuses on units larger than the sentence such as a piece of conversation, a story, or a series of paragraphs in an essay. This definition is broad enough to include the bulk of what discourse analysts actually do. On the one hand, some of their research is concerned with how sentences are put together to form larger chunks of discourse and with the identification, description, and explanation of systematic patterns of discourse organization. On the other hand, there are studies which investigate how language is used in social interaction and attempt to relate aspects of the structure of discourse to contextual factors such as the purpose of the interaction or interlocutors' traits, shared knowledge, and role relationships. This twofold approach to discourse analysis will serve as the basis for the following discussion of Arabic discourse. This discussion also takes into consideration two important features of the literature on Arabic discourse. First, discourse analysis in general examines both speech and writing, with understandable variation regarding the kind of issues addressed and the methodology deemed appropriate for each medium. In the case of Arabic, given the diglossic situation involved, the selection of speech or writing as a research focus often entails the consideration of a different language variety altogether, since writing is typically associated with Modern Standard Arabic, and speech with regional vernaculars. It is not surprising, then, that research on everyday social interaction typically involves vernaculars, whereas texts written in Modern Standard Arabic serve primarily to investigate discourse patterns such as the linkage among sentences or the discourse functions of particular linguistic elements. Second, there exist many studies which deal with Arabic discourse but only secondarily, their primary focus being issues in applied linguistics, particularly → second language acquisition and pedagogy. Relevant content culled from such studies is also included in this discussion in order to achieve a high degree of comprehensiveness. Thus, the remainder of this entry is organized as follows. Section 2 deals with discourse patterns in texts written in Modern Standard Arabic and section 3 with Arabic discourse in social interaction. The discussion of discourse analysis in pedagogical contexts is provided in section 4 and is followed by concluding remarks in section 5.

2. The discourse of written texts

Arabic prose exhibits a number of salient features which give it a very distinctive quality, especially when compared to commonly-studied Western languages, such as English. Lexical repetition, structural parallelism, and the prevalence of coordination are some of the most easily noticeable and widely investigated features. These have been of interest not only to Arabic linguists, but also to scholars in applied linguistics, contrastive rhetoric, and second language writing (Sa'adeddin 1989; Kaplan 1966; Connor 1996; see also discussion in section 4). In addition to investigations of these major characteristics of written Arabic discourse, there are also other studies that deal with the discourse functions of specific linguistic structures and still others that focus on the organization of particular discourse genres such as academic or business discourse.

Leading the first group is probably Johnstone's seminal work on the ubiquity of coordination and repetition in Arabic discourse (Johnstone 1990, 1991). Using numerous examples from Arabic prose, Johnstone describes in detail how such discourse is highly paratactic, relying heavily on the use of → connectives and coordination (→ parataxis) rather than subordination to link sequences of clauses and sentences into cohesive texts. Arabic texts are also shown to employ a great deal of repetition and formulaic patterns at the morphological, lexical, syntactic, and discourse levels. The following excerpt where an Arab describes the reaction of Great Britain to nationalism illustrates extensive use of structural parallelism involving a number of parallel clauses linked with the conjunction *wa-* 'and'.

ʿāraḍūhā	*fī baʿḍi*	*l-ʾaḥwāli*
opposed-them	in some	the cases

wa-ʾayyadūhā	*fī*	*ʾaḥwālin*
and-endorsed-them	in	cases

wa-qayyadūhā	*bi- baʿḍi*	*l-quyūdi*
and restricted them	with some	the restrictions

l-ʾuxrā	*fī baʿḍi*	*l-ʾaḥwāli*
other	in some	the cases

wa-ltazam	*ḥiyālahā*	*siyāsata*
and advocated	regarding them	policy
l-ḥiyādi	*fī muʿḍami*	*l-ʾaḥwāli*
the neutrality	in most the	cases

'They [the leaders of England] opposed them [nationalistic movements] in some cases, and endorsed them in other cases, and placed some restrictions on them in some cases, and advocated a policy of neutrality with regard to them in most cases' (Johnstone 1991:98–99)

The prevalence of these features is partially attributed to the orality of Arabic discourse which, the author argues, can be traced back to oral styles valued in earlier developments of the Arabic language. It is also attributed to the syntactic mechanisms available in Arabic and to the socio-culturally motivated desire of writers to draw attention to the form of the message itself.

Al-Jubouri (1983) concentrates in particular on repetition and the distinct quality it gives to Arabic discourse. He points out that, at the lexical level, Arabic discourse exhibits frequent strings of two and sometimes three words which are semantically related, such as *al-ḥujjatu wa-d-dalīlu* 'evidence and proof'. Repetition of syntactic patterns which results in partial or complete parallelism is also documented, in addition to repetition of meaning through writers' restatement of the same point. It is claimed that these types of repetition serve as a means of developing arguments, strengthening the writer's point of view, and in general enhancing the persuasive quality of discourse.

Other studies have attempted to provide discourse accounts of particular linguistic devices and grammatical structures as they occur in Arabic texts. Al-Batal (1990) examines the discourse functions of connectives such as *wa-* 'and', *lākinna* 'but', and *fa-* 'therefore', in an essay by the Arab writer and literary critic ʿAbbās Maḥmūd al-ʿAqqād and finds that such connectives encode hierarchical relationships among different sections of a text and indicate thematic continuity or shifts in discourse. According to Al-Batal, the high frequency of these devices in Arabic texts is consistent with the oral tradition in Arabic literary practices and is motivated by the search for rhetorical effectiveness in argumentative writing and by the need to create maximum textual → cohesion. Similarly, Khalil (2000) investigates sentence-initial markers such as the particle *qad* and canned phrases like *min al-maʿrūfi ʾanna* 'it is known that' in news discourse. The function of these markers is to signal the → grounding value of information in discourse; in other words, they distinguish important pieces of information from subsidiary ones. Thus, the expression *min al-maʿrūfi ʾanna* introduces backgrounded secondary information, whereas the particle *qad* indicates a higher degree of grounding or 'midground', to use the author's terminology. Furthermore, such grounding functions are claimed to be associated with and to account for the distribution of Verb-Subject and Subject-Verb word orders in Arabic discourse.

Using Givón's model of Topic Continuity (Givón 1983), which captures the correlation between the linguistic encoding of referents and their degree of persistence in discourse, Fakhri (1995a) provides discourse explanations for the use of a number of morphosyntactic properties of Arabic. He shows that the maintenance or shift of referents in discourse accounts for the variable distribution of Verb-Subject and Subject-Verb word orders in narrative and expository texts, for the behavior of the particle *ʾinna* as a topic marker, for the pattern of subject-verb agreement reduction, and for the constraints regarding the deletion of *wāw al-ḥāl* (the equivalent of 'while') in circumstantial clauses. A discourse perspective is also adopted in Fakhri (1998) to analyze the linguistic devices for reporting the speech of others in newspaper articles. The analysis indicates that the choice of direct or indirect reported speech as well as that of reporting verbs like *qāla* 'to say' and *iʿtarafa* 'to admit' are not arbitrary, but rather motivated by the ideological stance of the writer.

In addition to the research outlined above, a few studies have adopted a genre analysis approach to Arabic discourse. Genre analysis describes and explains discourse patterns and rhetorical strategies utilized in particular types of professional or academic modes of communication (e.g. journal articles or business letters), which are often conventionalized and exhibit a well-defined internal structure (Swales 1990). Using Swales's model for the analysis of introductions in research articles (Swales 1981, 1990), Najjar (1990), and Fakhri (2004) investigate the structure of Arabic introductions. Najjar's analysis of introductions of research articles from the field of agricultural sciences shows that about half of these introductions fit the model developed by Swales for English, where authors apply a number of moves to indicate the importance of their area of inquiry, justify their current contribution, and announce the purpose and the structure of the article. The study also shows that challenges of previous research, which often occur in English academic discourse, are absent in the Arabic data. Najjar suggests that this is due to the applied nature of research in agricultural studies where theoretical argumentation is considered unimportant. In the same vein, Fakhri (2004) studies a sample of introductions from humanities and social sciences articles, with the expectation that such data are bound to exhibit more language- and culture-specific properties than data from the hard sciences, as argued in Duszak (1997:11). The study indicates important differences between these data and the Swales model: only a few Arabic introductions employ the moves predicted by the model, previous scholarship is rarely challenged, and the purpose of articles and their structure are not always explicitly stated. These findings are accounted for in terms of the writers' educational background and the modest expectations of a nascent discourse community.

Haichour (1999) compares English and Arabic genres of business discourse in terms of the following linguistic and discourse properties: agency, causation, metadiscourse, and spatio-temporal indexing. The results indicate striking similarities between English and Arabic and lead the author to conclude that English and Arabic parallel business genres are more similar both structurally and functionally than different genres within the same language. Najjar (1996) examines the discourse organization adopted in an altogether different genre, lists of martyrs, which portray Palestinians killed during the uprising against the Israeli occupation. The lists come from issues of the Palestinian magazine *al-Kātib*. The analysis indicates that they exhibit traditional narrative structure consisting of abstract, orientation, problem, solution, and coda. They also rely heavily on metaphors and imagery to dramatize the plight and struggle of the Palestinian people, which, according to the author, makes them very persuasive. From a functional perspective, the author suggests that the regular publication of these lists serves to explain the martyrs' sacrifice so as to make their deaths tolerable, to routinize the rebellion, and to provide opportunities for commentary on the peace process.

3. ARABIC DISCOURSE IN SOCIAL INTERACTION

People regularly engage in talk to exchange information, express opinions, argue, or simply tell stories and jokes. Such social interactions are conducted so skillfully and effortlessly that they appear very simple. However, this apparent simplicity is quite deceptive. Relevant research consistently points out the complexity of such human endeavors, which is due to a wide range of contextual features that determine the shape of spoken discourse and include, among others, the degree of formality of the speech event, the type of language or dialect selected, the topic of conversation, as well as the gender, status, and role relationships of participants and the background knowledge they share. For example, participants' gender and status may influence the choice of conversation content and how turn-taking is managed, whereas their shared background knowledge is likely to determine the amount of detail necessary to facilitate comprehension. These contextual features are bound to reflect the cultural characteristics and social norms of the speech community. Given the linguistic and sociocultural diversity in the Arab world, a great deal of variation is to be expected regarding the conduct of everyday social interaction. This state of affairs should, in principle, make the study of this aspect of Arabic discourse appealing and stimulating because of the obvious intellectual challenges it poses and the potential sociocultural insights to be gained from it. Unfortunately, research in this area is rather scant, with a few exceptions.

Arent (1998) and Kharraki (2001) investigate negotiations that take place between merchants and customers in the process of bargaining, an all too frequent speech act easily observable in Arab markets. The Arent study is based on data collected from bargaining sessions involving Arab and non-Arab customers in a market in Amman, Jordan. The author focuses particularly on language choice and pragmatic failure or communication breakdown. As expected, Arabic vernaculars are utilized most frequently, with instances of use of French, English, or a mixture of the two. A few cases of pragmatic failure are documented, which vary depending on the nationality of the customers. In the case of Arab customers, pragmatic failure often results from their criticism of product quality, which the merchants naturally deem inappropriate and uncalled for. For non-Arab customers, on the other hand, such failure occurs when vendors, to the dismay of these customers, provide the price of merchandize only after inquiring and obtaining information about their nationality. Kharraki's (2001) study of the speech act of bargaining investigates the similarities and differences between men and women regarding the strategies they resort to in order to obtain merchandize price reduction. The data come from interactions between merchants and customers in markets in eastern Morocco, where, according to Kharraki, women are claimed to be particularly keen on, and skilled at bargaining, a socially stigmatized trait for some. The analysis of the data indicates that, as a prelude to the act of bargaining proper, men attempt to establish familiarity with vendors and create a friendly atmosphere through elaborate greetings. Women, on the other hand, are not supposed to and in fact do not engage in such talk, the purpose being to maintain a certain social distance between themselves and their interlocutors and, thus, ensure a measure of respectability. Regarding bargaining strategies, men often denigrate the quality of the products offered in an attempt to secure price reduction. By contrast, women tend to view bargaining as an opportunity for self-assertion and employ an insisting strategy which consists in continually restating their desire for price reduction, reinforced by multiple threats to do their shopping elsewhere.

Of a more solemn nature are the studies conducted by Nazzal (2001) and Saeed (1997), which explore links between religion and discourse. Nazzal examines the pragmatic functions of religious expressions such as *ʾinšāʾallāh* in everyday social interaction. The main finding of the study is that speakers resort to such expressions as face-saving devices that serve to mitigate the potentially negative impact of their talk on listeners. Saeed (1997) analyzes audio- and video-taped segments of religious discourse in order to determine the pragmatic functions of switching from Modern Standard Arabic to regional vernaculars. The data come from Arab theologians' religious lectures and interactions with their audiences. The results show that → code-switching is very frequent in this type of religious discourse, especially in the question-answer portions of the lectures. They also show that content perceived by the speakers to be positive is expressed in Modern Standard Arabic, whereas content with negative connotations is delivered in the vernacular.

4. Discourse analysis in pedagogical contexts

Several studies present analyses of Arabic discourse within the context of → second language acquisition and pedagogy. These studies typically compare and contrast aspects of text organization in Arabic and English (and French in one case) in order to elucidate patterns of language transfer and make suggestions for improving language instruction. Although their accounts of transfer are interesting and their pedagogical suggestions quite valuable, these issues fall beyond the scope of the present discussion and thus will be mentioned only when necessary for the clarity of presentation. Emphasis is instead shifted to portions of this research dealing with Arabic discourse proper and how it is similar to or different from that of other languages. Furthermore, in order to understand the variation in some of the results reached in this research, it is worth noting that some of the analyses presented are quantitative and based on computing features of texts and comparing their frequencies across languages, while others are qualitative, relying on detailed descriptions of excerpts from texts.

Ouaouicha (1986) and Fakhri (1995b) provide quantitative analyses of English and Arabic argumentative and expository texts. Using insights from the Toulmin model for argumentation (Toulmin 1958), the Ouaouicha study investigates the structure of argumentative essays by Americans and Moroccans in both Arabic and English and finds no significant correlation between language and argument structure. However, the American subjects' texts exhibit more awareness of audience and emotional appeals for persuasion, which seems to be the result of training rather than linguistic determinism. Fakhri's (1995b) study compares the topical structure of expository texts in Arabic and English. Topical structure probes the patterns of topic maintenance or shifts across sentence sequences in a text. The comparison reveals no significant differences between the two languages: both rely overwhelmingly on sequential progressions which involve a series of adjacent sentences with different topics. This result is attributed to the type of writing considered, namely expository writing, where often main topics are subsequently developed and supported by introducing other subtopics and details, resulting in topic shifts and a high frequency of sequential progressions.

The qualitative analyses presented in Holes (1983) and Sheikholeslami and Makhlouf (2000) are also concerned with the transfer of discourse patterns from Arabic into the English writing of Arabic speakers. Holes (1983) examines a sample writing in English by a native speaker of Arabic and shows how it is influenced by Arabic discourse patterns. These influences include the use of long sentences with many instances of coordination, frequent repetition and structural parallelism, and the tendency to adopt a direct and personal style even in academic discourse. Sheikholeslami and Makhlouf (2000) on the other hand analyze two Arabic texts and show that they follow linear organization similar to the one used in English prose. This leads the authors to conclude that the rhetorical problems in the English writing of Arabic speakers do not stem from Arabic interference. They suggest that a more plausible source of these problems is the model school essay frequently used in teaching Arabic writing. In support of this claim, they analyze a model essay from a popular Arabic composition textbook, which exhibits ornate style, frequent repetition, and quotations of Arabic poetry, with little attention given to the overall organization and coherence of the text.

Two other studies, Williams (1983) and Zizi (1987), conduct language comparisons involving Arabic for the purpose of improving the teaching of English writing to native speakers of Arabic. The Williams study compares cohesion patterns and thematic development in English and Arabic texts and reveals a tendency in Arabic to repeat the same theme in successive clauses and to resist ellipsis. Within the context of teaching English for journalistic purposes in Morocco, Zizi (1987) compares Arabic, French, and English with respect to the structure of news reports and analysis. The results indicate that the global structure of newspaper articles is similar in all three languages; however, French headlines tend to be the most ambiguous and, not surprisingly, Arabic texts are found to comprise the highest amount of repetition.

5. Conclusion

While by no means exhaustive, this discussion provides an adequate picture of the state of the research on Arabic discourse, its main concerns, approaches, and findings. An important conclusion that emerges from this discussion is the privileged status assigned to Modern Standard Arabic and written language in both the studies of Arabic discourse proper and those conducted for pedagogical purposes. This is not totally surprising because, in a sense, it is simply the continuation of attitudes adopted in the study of the more traditional linguistic levels of phonology, morphology, and syntax, where Modern Standard Arabic has occupied a prominent position in the research agendas of Arab and non-Arab linguists alike. While this approach may not be viewed as a negative, the relative neglect of spoken discourse using regional vernaculars deprives the field of perhaps the most insightful and advantageous dimension of discourse analysis, its sociocultural dimension. By contrast to the other linguistic levels, the analysis of the discourse of spontaneous everyday interaction not only enhances our understanding of how language works, but also provides an important means of gaining insights into the sociocultural makeup of the speech community, its concerns, norms, values, and aspirations. More research in this area is thus necessary, especially in light of the sociopolitical and cultural animation which currently characterizes the Arab world.

Another observation worth making is that most of the topics, issues, constructs, and approaches selected for investigating Arabic discourse reflect to a large degree those used in the analysis of other languages such as English. This type of replication has turned out to be quite fruitful and is undoubtedly very commendable. However, there seems to be a need for highlighting research aspects peculiar to the conduct of discourse in Arab communities where different varieties of Arabic coexist with other languages such as Berber and French in North Africa or English in some areas of the Arab Middle East. Such uncommon linguistic richness with all its social, political, and cultural implications, needs to be reflected more prominently in the study of Arabic discourse, which will certainly constitute a more original and distinctive contribution to the field of discourse analysis.

Bibliographical references

Al-Batal, Mahmoud. 1990. "Connectives as cohesive elements in a modern expository Arabic text". *Perspectives on Arabic linguistics*, II, ed. Mushira Eid and John McCarthy, 234–268. Amsterdam and Philadelphia: J. Benjamins.

Al-Jubouri, Adnan. 1983. "The role of repetition in Arabic argumentative discourse". Swales and Mustafa (1983:99–117).

Arent, Russell. 1998. *The pragmatics of cross-cultural bargaining in an Amman Suq: An exploration of language choice, discourse structure and pragmatic failure in discourse involving Arab and non-Arab participants (Jordan, Amman)*. Ph.D diss., University of Minnesota.

Connor, Ulla. 1996. *Contrastive rhetoric: Cross-cultural aspects of second language writing*. Cambridge: Cambridge University Press.

Duszak, Anna. 1997. "Cross-cultural academic communication: A discourse community view". *Culture and styles in academic discourse*, ed. Anna Duszak, 11–39. Berlin: de Gruyter

Fakhri, Ahmed. 1995a. "Topic continuity in Arabic narrative discourse". *Perspectives on Arabic linguistics*, VII, ed. Mushira Eid, 141–155. Amsterdam and Philadelphia: J. Benjamins.

——. 1995b. "Topical structure in Arabic-English interlanguage". *Pragmatics and language learning*, ed. Lawrence Bouton, 155–168. Urbana-Champaign: University of Illinois.

——. 1998. "Reported speech in Arabic journalistic discourse". *Perspectives on Arabic linguistics*, XI, ed. Elabbas Benmamoun, Mushira Eid, and Niloofar Haeri, 165–182. Amsterdam and Philadelphia: J. Benjamins.

——. 2004. "Rhetorical properties of Arabic research article introductions". *Journal of Pragmatics* 36: 6.1119–1138.

Givón, Talmy. 1983. *Topic continuity in discourse: A quantitative cross-cultural study*. Amsterdam and Philadelphia: J. Benjamins.

Haichour, El Houcine. 1999. *A corpus linguistic analysis of English and Arabic parallel business discourse domains*. Ph.D diss., Georgetown University.

Holes, Clive. 1983. "Textual approximation in the teaching of academic writing to Arab students: A contrastive approach". Swales and Mustafa (1983: 228–242).

Johnstone, Barbara. 1990. "Orality and discourse structure in Modern Standard Arabic". *Perspectives on Arabic linguistics*, I, ed. Mushira Eid, 215–233. Amsterdam and Philadelphia: J. Benjamins.

——. 1991. *Repetition in Arabic discourse*. Amsterdam and Philadelphia: J. Benjamins.

Kaplan, Robert. 1966. "Cultural thought patterns in intercultural education". *Language Learning* 16.1–20.

Khalil, Esam. 2000. *Grounding in English and Arabic news discourse*. Amsterdam and Philadelphia: J. Benjamins.

Kharraki, Abdennour. 2001. "Moroccan sex-based linguistic difference in bargaining". *Discourse and Society* 12:5.615–627.

Najjar, Hazim, 1990. *Arabic as a research language: The case of agricultural sciences*. Ph.D diss., University of Michigan.

Najjar, Orayb. 1996. "The editorial family of al-Kateb bows in respect: The construction of martyrdom text genre in one Palestinian political and literary magazine". *Discourse and Society* 7:4.499–530.

Nazzal, Ayman. 2001. *The pragmatic functions of Qur'anic verses: The case of* insha'allah *in Arabic* *discourse as a species of indirectness*. Ph.D diss., State University of New York, Albany.

Ouaouicha, Driss. 1986. *Contrastive rhetoric and the structure of learner-produced argumentative texts in Arabic and English*. Ph.D diss., University of Texas, Austin.

Saeed, Aziz. 1997. *The pragmatics of codeswitching from* fusha *Arabic to* ammiyyah *Arabic in religious-oriented discourse*. Ph.D diss., Ball State University, Indiana.

Sa'adeddin, Mohamed. 1989. "Text development and Arabic-English negative interference". *Applied Linguistics* 10.36–51.

Swales, John. 1981. *Aspects of article introductions*. Aston ESP Research Report No. 1, Language Studies Unit, University of Aston, Birmingham, U.K.

——. 1990. *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.

—— and Hassan Mustafa (eds.). 1983. *English for specific purposes in the Arab world*. Birmingham: Language Studies Unit, University of Aston, Oxford University Press.

Sheikholeslami, Cynthia and Nabila Makhlouf. 2000. "The impact of Arabic on ESL expository writing". *Diversity in language: Contrastive studies in Arabic and English theoretical and applied linguistics*, ed. Zeinab Ibrahim, Sabiha Aydelott, and Nagwa Kassabgy, 127–146. Cairo: American University in Cairo Press.

Toulmin, Stephen. 1958. *The uses of argument*. Cambridge: Cambridge University Press.

Williams, M.P. 1983. "A problem of cohesion". Swales and Mustafa (1983:118–128).

Zizi, Khadija. 1987. *Contrastive discourse analysis of argumentative and informative prose in Arabic, French, and English: Suggestions for teaching/learning English as a foreign language for journalistic purposes (EJP) in Morocco*. Ph.D diss., University of Illinois at Urbana-Champaign.

AHMED FAKHRI (West Virginia University)

Discourse Markers → Connectives

Dissimilation

Dissimilation can be viewed as a process by which one segment avoids taking on a feature (or a set of features) of another segment, usually an adjacent one. As such, it may be regarded as a counterprocess to assimilation. This process is not very frequent in either Classical/Modern Standard Arabic or in the Arabic dialects, at least not in its linear manifestation. Generally speaking, dissimilation may be progressive or regressive and may involve consonants and vowels alike. In Classical and Modern Standard Arabic, the main domain of dissimilation is that of words derived from weak roots (containing the semivowels *w* and *y*), especially when the semivowel occurs between two vowels.

Some dissimilative processes are attested from the dialects of pre-Islamic Arabic tribes, such as dissimilation of *m* before *b* (*mā smuka* > *bā smuka* 'what is your name?'), dissimilation of *m* after *w* (*wamad* > *wabad* 'strong heat'), elision of *m* before *n* (*minxafat* > *ʾinxafat* 'air pump') (cf. Fleisch 1961:80).

The behavior of semivowels in weak verbal roots (especially II*w/y*) may also be viewed as affected by a process of dissimilation. In some positions, the underlying *w/y* changes into other consonants, mostly ʾ (glottal stop), as in **muwaqqatun* > *muʾaqqatun* 'temporary', **qāwilun* > *qāʾilun* 'saying [active participle]' or **ʾaʿḍāwun* > **ʾaʿḍāʾun* 'members'. The change *w* > 'can be explained as avoidance of the vocalic characteristics of the semivowels. In the Arabic dialects, the Classical Arabic form **qāwimun* changes to *qāyim* 'standing'.

Several forms involving vowels may be interpreted as dissimilated from an underlying/historical form or in opposition to the singular form. These include both grammatical morphemes and plural formation, especially that of the internal plurals ending with suffixed *-ān*. The following forms belong here:

plural feminine accusative ending: **-ātan* > *-ātin*;
dual ending (both nominal and verbal): **-āna* *-āni*;
change of vowel in plurals: e.g., **jārān* > *jīrān* (sg. *jārun*) 'neighbors'

Usually, dissimilation is based on the concept of adjacency, which is treated in a linear way, but in Arabic, processes of dissimilation can also be observed within an autosegmental tier like the root. The → Obligatory Contour Principle (OCP; cf. McCarthy 1986) prevents the neighboring members (consonants) of the verbal root from being identical or similar (formed at a close point of articulation, e.g. a series of coronal sonorants are avoided). Yet, this constraint is only valid within a given domain (verbal root), and is inactive across the tier boundaries (e.g. in affixes to the stem formed by combination of the root/consonantal and vocalic melody). Such a concept of dissimilation means a projection of an underlying dissimilation rule onto the organization of the lexicon.

Another such type of underlying dissimilation rule might be the tendency observable in the formation of Forms VII (*in-KaTaBa*) and VIII (*iK-ta-TaBa*), which are both used for forming a passive transformation of the basic meaning. Here, one finds that the lexicon avoids combinations of adjacent features in the neighborhood of the prefix/infix; the verbal roots with coronal obstruents in the initial position avoid Form VIII (-t- infix), and verbal roots with coronal sonorants avoid Form VII (*n-* prefix).

Bibliographical references

Fleisch, Henri. 1961. *Traité de philologie arabe*. I. *Préliminaires, phonétique, morphologie nominale.* Beirut: Imprimerie Catholique.

McCarthy, John J. 1986. "OCP effects: Gemination and antigemination". *Linguistic Inquiry* 172.207–263.

Petr Zemánek (Charles University)

Djibouti/Eritrea

Arabic, though not native to the Horn of Africa, is widely spoken in two countries where it is not the majority language: the Republic of Djibouti and Eritrea. This region on the west coast of the Red Sea opposite the Tihama region of Yemen has always been one of contact and exchange. Links between the African and Arabian coasts are attested since antiquity and are doubtless much older.

Arabic was spoken there before the arrival of Islam. The Christianization of Abyssinia was carried out by Arabic-speaking priests from Syria who began to evangelize the hinterland, central Eritrea, and the Tigray region (Abraha Wende 2000; Killion 1998:331) early in the 4th century. Islamization began in the 7th century with the part of the coast between Massawa and Zeila which includes the former Adulis (now Zûla). The spread of Arabic attended the expansion of Islam which reached Djibouti in the 9th century via the merchants from southern Arabia who settled in Tadjoura (Rouaud 1997).

This presence in ancient times is well documented and has lasted until today. In the 19th century, the Rashayda came from Saudi Arabia to settle along the coast north of Massawa; their descendants are today citizens of Eritrea, and one of the national ethnic groups. Around the same time, another Arabic-speaking community settled further south in what was to become the Côte française des Somalis in 1896 and the Territoire français des Afars et des Issas in 1967. From 1892 onward, large numbers of Yemenis, mostly from the Tihama and Hogariya regions and from Aden, were enticed by the French to build the new city of Djibouti. They were at that time the largest community in town (Rouaud 1997:324). When the Territoire became independent in 1977, the Yemenis, most of whom still lived in Djibouti Town, although some had settled elsewhere, became citizens of the new republic.

Arabic has official status in both countries. It is an official language of Djibouti together with French. In the young nation of Eritrea, which became independent in 1993, Arabic is one of the three 'working languages' together with Tigrinya and English. It is also the religious language of almost all Djiboutians: article 1 of the country's constitution makes Islam the state religion, and a little over 40 percent of Eritreans are also Muslims.

Official status means that Arabic (more precisely, Modern Standard Arabic) is used nationally for communication in both oral and written forms as a sort of koine by politicians, intellectuals, and business people. It is taught at school in both countries and used in the media. Dialectal forms of oral Arabic are also the vernacular language of a few communities, and/or may be used as a trade language or → lingua franca for communication between speakers whose mother tongue is not Arabic, but another local language.

In Djibouti, all teaching is done in French, but Modern Standard Arabic is the only one of the three national languages to be taught from the fifth year of primary schooling on. An experiment is currently underway to begin teaching Arabic from the first year of primary education. The entry examination to secondary school includes an obligatory Arabic test, as do the examinations leading to the secondary school diplomas: *brevet d'études* in secondary school, and *baccalauréat* in high school (Ali Awad 2004).

Modern Standard Arabic is used in the written press in the bi-monthly *al-Qarn*. The radio devotes six hours a day (from 8 to 10 a.m. and 4 to 8 p.m.) to broadcasting in Modern Standard Arabic. On television, there are news, entertainment, motion pictures, and some advertising (Kassim Mohamed 2004). An article in the tri-weekly official French-language publication *La Nation* (26 April 2004, 12) envisages a plan to open three FM radio stations broadcasting twelve hours a day in each of the three main local languages: Arabic, Somali, and Afar.

Vernacular forms of Arabic are spoken in Djibouti by 8–10 percent of the population of some 644,000 (Couba 1993:18). "Demographic weakness . . . is made up for by the economic strength" of this community, all of whom originated from Yemen (Couba 1993:18). The mother tongue is used in the family and in relations with other members of the same linguistic community. Each can immediately recognize another's dialectal origins, as characteristic features are not totally blurred when different Arabic speech forms come into contact. This community has also been instrumental in the propagation of a form of trade Arabic used in relations with speakers of other mother tongues (Afar and Somali). Arabic speakers can be found in all population centers, but the great majority are concentrated in the capital. Among the 400,000 inhabitants of Djibouti Town, some Arabic speakers, mostly businessmen and traders, occupy the downtown area while others are found in quarters 1 to 4 and in Ingela where they cohabit with citizens who are speakers of other languages. The members of this community, who originate from Taez in the mountainous region of Hogariya, are called Gibaliya. This designation may also cover the group known as the Bayḍāni, a term used in the 1970s to refer to Arabs from the mountainous area *(manṭaqa al-wusṭā)* on the border between the two Yemens (Rouaud 1997:328).

The suburban town of Ambouli, just four kilometers outside Djibouti, is said to have 2,000 Arabic-speaking inhabitants (Ali Awad 2004). They represent three different communities. The Hakmi (*ḥakmi*), who came from Tihama in Yemen (villages of Dhuhab, Wahiga, and Bab al-Mandeb), live in Ambouli 1 and Ambouli 2 (or Jebel South). They were originally marsh dwellers and fishermen. Jebel quarter is home to the Dureyhimi (*durhīmi*, pl. *darāhīma*), who were formerly farmers and shepherds. As their name suggests, they are from Durayhimi, a town near Zabid in Tihama. Lastly, the Mashlahi are from Mozaʿ and live mostly in Ambouli 1. Rouaud (1997:328) also speaks of the Banu Zīd from the Makha region.

Small numbers of Arabic speakers can be found in the other population centers of the republic: Dikhil, Ali-Sabieh, Tadjoura, and Obock, which has an entire quarter of Darwish surrounded by Afar speakers. Their speech is the only dialect thus far to have been briefly investigated (Simeone-Senelle 2002). It belongs to the Tihami group. Research on the other vernacular forms and on the trade Arabic spoken in Djibouti has barely begun (→ lingua franca: Horn of Africa).

In Eritrea, Modern Standard Arabic is one of the three working languages. During the armed struggle, the Eritrean People's Liberation Front (EPLF) "organized literacy classes . . . three times each week in Tigrinya and Arabic to give people from different communities of Eritrea a common language" (Connell 1997:37). Nowadays, Modern Standard Arabic is used as a medium of instruction "in more than 115 elementary schools, grades 1 to 5". Most of these "are located in the lowland regions, especially in Gash-Barka (where there is a large number of returnees from the Sudan), Ansaba, and north and south Red Sea regions" (Osman Ali 2004). In some regions, Arabic dialects are or are becoming the vernacular language. The island of Dahlak Kebir is an exception. Arabic is used there for primary instruction, although the mother tongue of most of the inhabitants is Dahaalik, an Afro-Semitic language which was discovered by the French research team only in 1996, and is therefore not included among the national languages.

The curriculum board of the Eritrean ministry of education is in charge of the preparation of primary school curricula and textbooks in the nine national languages including Arabic. National television broadcasts in Arabic, in the other working languages, and in Tigre. There is a weekly Arabic-language newspaper, *al-ʾIrītriyā al-ḥadīṯa* (Tigrinya and English journals appear with the same frequency).

There are very few people in Eritrea who have Arabic as their mother tongue. Estimates of their number vary: Killion (1998:354) suggests 1,000, Abraha Wende (2000) 30,000, or nearly 1 percent of a total population of 3,500,000.

The largest group is the Rashayda, who make up 5 percent of the native Arabic speakers (Osman Ali 2004). They are Sunni Muslims and live on the coast north of Massawa. They are camel raisers and trade mostly with the opposite side of the Red Sea. Their dialect is related to those of Saudi Arabia. A different dialect is now the mother tongue of certain Halenqa, who are culturally of Beja origin and live in the western part of Eritrea. Modern Standard Arabic is used as the medium of instruction at the elementary level in their area (Osman Ali 2004). There is also a community of Yemeni origin on the coast, but "their numbers declined during the period of the Independence War (1961–1991)" (Osman Ali 2004). To these may be added the Eritrean citizens who have lived in exile as refugees in the Arab countries (the Sudan, Yemen, Saudi Arabia) over the last 40 years. The number of those who lived in the Sudan and were officially registered as refugees was more than 750,000; about half of them have now come back to their homes. The younger generation of these diaspora groups who were born in the Sudan or other Arab countries learnt Arabic as a first language (Osman Ali 2004).

These historical, political, economic, and religious factors account for the importance and prestige of Arabic in the region. Such is this prestige that some parents choose to have their children taught in Arabic, even when it is not their mother tongue. It is equally further incitement to the use of Arabic as a trade language.

BIBLIOGRAPHICAL REFERENCES

Abraha Wende, Tekle. 2000. "A preliminary survey of Arabic in Eritrea". Paper presented at the International Conference Against all Odds: African Languages into the 21st Century, Asmara, 11–17 January 2000.

Ali Awad, Fahade. 2004. Personal communication, Pôle Universitaire de Djibouti, Djibouti.

Connell, Dan. 1997. *Against all odds: A chronicle of the Eritrean revolution with a new afterword on the postwar transition.* Laurenceville, Asmara: Red Sea Press.

Kassim Mohamed, Souad. 2004. Personal communication, Pôle Universitaire de Djibouti, Djibouti.

Killion, Tom. 1998. *Historical dictionary of Eritrea.* Lanham Md. and London: Scarecrow.

Osman Ali, Mohammed. 2004. Personal communication, March. Curriculum Branch, Asmara.

Rouaud, Alain. 1997. "Pour une histoire des Arabes de Djibouti, 1896–1977". *Cahiers des Etudes Africaines* 146.xxxvii–2, 319–348.

Simeone-Senelle, Marie-Claude. 2000a. "Situation linguistique dans le sud de l'Erythrée". *Proceedings from the 2nd World Congress of African Linguistics, Leipzig July 1997*, ed. H. Ekkehard Wolff and Orin Gensler, 261–276. Cologne: R. Köppe.

——. 2000b. "L'arabe véhiculaire parlé en Erythrée sur la côte de la mer Rouge, de Massawa à Rahayta". *Oriente Moderno*, n.s. 18 (79), 153–180.

——. 2000c. "Les langues en Erythrée". *Chroniques Yéménites* 2000. 167–173.

——. 2002. "L'arabe, langue maternelle de citoyens djiboutiens du nord de la République de Djibouti". *Proceedings of the 4th Conference of the International Arab Dialectology Association (AIDA), Marrakesh, Apr. 1–4. 2000*, ed. Abderrahim Youssi a.o., 140–150. Rabat: AMAPATRIL.

MARIE-CLAUDE SIMEONE-SENELLE
(LLACAN-CNRS, INALCO, University of Paris-7)

Dravidian Languages → Malayalam; Tamil; Telugu

Dual → Number

Dysphemism

Dysphemism (the opposite of → euphemism) is a lexical resource in natural language, whereby lexemes are created for the purpose of combining denotation and negative attitude via a complex process of lexical compression. Just like euphemism, but with an opposite directionality, the process of dysphemizing produces cognitive synonyms that converge on denotation but diverge on attitude. For example, the Jordanian Arabic lexemes *bitsammam* 'he is poisoning himself' and *ingalʕat* 'she has been extracted' are dysphemistic counterparts of the neutral lexemes *bōkil* 'he is eating' and *rawwaḥat* 'she left', respectively. That is, the speaker may utter *bitsammam* instead of *bōkil* to inform the interlocutor that the male denotatum is eating but he (the speaker) does not like the fact that the denotatum is doing so. Hence, the speaker intentionally selects the dysphemistic lexeme, which literally says that 'the denotatum is poisoning himself'. As a result, the speaker effectively conveys the message that the referent is eating, alongside his negative attitude about the state of affairs in question. Similarly, the speaker who employs *ingalʕat* 'she has been extracted' instead of *rawwaḥat* 'she has left', effectively conveys the same denotation plus a negative attitude, namely that the referent's presence had been a burden to the speaker.

Allan and Burridge (1991:26) define dysphemism as "an expression with connotations that are offensive either about the denotatum or the audience or both, and it is substituted for a neutral or euphemistic expression for just that reason". In this way, dysphemism is used to talk about one's adversaries or things one wishes to downgrade or express disapproval of. Although Allan and Burridge's definition roughly captures the purpose of dysphemism, their discussion of dysphemistic data (1991:26–29) fails to distinguish inherently negative lexical items from dysphemistic ones. According to their view, the Arabic lexemes *ʾirhābī* 'terrorist', *kaḏḏāb* 'liar', and *mutahawwir* 'reckless' are dysphemisms, despite the fact that their negativity is inherent rather than lexicalized.

Farghal (1995:52) criticizes Allan and Burridge's account for broadening the concept of dysphemism to include lexical items which are inherently marked for negative connotations, because it weakens the strength of a purely pragmatic analysis of dysphemism. For instance, in most Arabic dialects the lexeme *ġabī* 'stupid' cannot be regarded as a dysphemism because it carries inherent rather than lexicalized negative connotations as an immediate consequence of dichotomizing lexis in terms of positivity and negativity. However, *ḥmār* and *ḥumār* 'stupid', in Jordanian and Egyptian Arabic, respectively, are dysphemistic because they create meaning via dysphemizing. Thus, *ḥmār* or *ḥumār* (lit. 'donkey') effectively conveys the message that the denotatum is extremely stupid by a process of lexicalization. Consequently, the input of dysphemism not only covers neutral and positive lexis, but also includes inherently negative lexis. When negative lexis is the target, dysphemizing is not a matter of adding a negative attitude (which is already there), but rather a matter of either heightening or lessening the degree of negativity. An example of reducing the degree of negativity would include the common avoidance of the taboo Arabic verb *nāk* 'to fuck' in most Arabic vernaculars in favor of a dysphemistic counterpart such as *ṭag* 'to knock' (Jordanian Arabic), *xayyaṭ* 'to sew' (Egyptian Arabic), *ṭaraʾ* 'to knock' (Syrian Arabic), *sāq* 'to drive, ride' (Kuwaiti Arabic), and *zaġab* '?' (Saudi Arabic). (Note that in some dysphemisms it is difficult to spell out the literal meaning of the word; a question mark is used to indicate this.)

Unlike euphemisms, which seek to save face (Brown and Levinson 1987), dysphemisms usually create face affronts in communication. When language users opt for dysphemizing, they are taken to be responding naturally to the psychological and social pressures they are experiencing. The issuance of a dysphemism in interaction is a reflex that represents conscious lexical choice on the part of the speaker. Yet, the pragmatic import, i.e. the illocutionary and perlocutionary forces (Austin 1962) of dysphemisms, correlate with the psychological and social rapport between speaker and denotatum in any given conversation, with their distance or closeness to each other in general, as well as in the particular context. The Egyptian dysphemism *ġūr* 'vanish!', i.e. 'get lost!', for example, creates a face affront or is just a flippancy, depending on the existing psychological and social rapport between speaker and interlocutor. Dysphemism, though psychologically real at the level of lexis in Arabic, is interactionally an utterance level phenomenon, just like conversational implicature and politeness in language (Thomas 1995).

Since they represent an utterance-level phenomenon, it is often difficult to assign dictionary meaning to dysphemistic expressions, despite their pragmatic import. One of the celebrated dysphemisms that came into frequent use during the Third Gulf War (which led to the toppling of Saddam and the occupation of Iraq) is the freshly-revived word *ʿulūj*. Few native speakers of Arabic would have missed that as-Sahhaf (the then Iraqi minister of information) was employing the unfamiliar lexeme *ʿulūj* as a dysphemism to refer to personnel of the allied forces invading Iraq. In fact, the revived word came to be on everyone's tongue in the Arab world. Ironically, and amidst the killing of tens of thousands of Iraqis, however, a row emerged over the original meaning of this word, with Arab writers citing many diverse meanings ranging between 'non-Arab', 'infidel stout men' and 'zebras' (*Lisān al-ʿarab*, 4th ed., Cairo, 3065–3066).When asked about it during and after the war, as-Sahhaf himself could not provide an exact original meaning for the word. He made it clear, however, that the word originally carried negative connotations and that he fell back on the Arab linguistic heritage to create this memorable dysphemism.

As a lexical resource expressing negative attitudes by adding vulgar meanings to existing vocabulary, Arabic dysphemisms are largely a product of colloquial rather than standard Arabic. This does not mean that standard Arabic cannot express negative attitudes; it only means that standard Arabic largely constructs such attitudes through grammaticalization, or what Sinclair (1991) calls the ‘Open Principle’. In other words, standard Arabic expresses negative attitudes by falling back on grammatical rules and existing lexemes to construct meaning (see last example in this paragraph). On the other hand, colloquial Arabic expresses negative attitudes via lexicalization and/or the ‘Idiom Principle’, which mainly create new lexemes or expressions by metaphoring or idiomatizing. By way of illustration, *inʿaṯar* (Jordanian Arabic), *ġār* (Egyptian Arabic), *ḏalaf* (Kuwaiti Arabic), and *walla* (Saudi Arabic) all lexically dysphemize the message that ‘the referent has left’. Similarly, Jordanian and Egyptian Arabic employ the Idiom Principle to dysphemize the same message in the idiomatic expression *warrāna ʿarḍ iktāfuh* ‘lit. he showed us the breadth of his shoulders’. By contrast, standard Arabic largely lacks dysphemism as a lexical resource; instead, it resorts to grammaticalization (the Open Principle) to add a negative attitude to a message like the one in the idiomatic expression above, e.g. *ġādara wa-ʾaḥmadu llāha ʾannahu faʿala ḏālik* ‘he left and I thank God that he did so’. This situation led Farghal (2003) to call for integrating colloquial dysphemisms into standard Arabic because they constitute an important lexical resource that efficiently responds to the psychological and social needs of Arabic speakers.

Arabic dysphemisms target a wide spectrum of words ranging from positive through neutral to negative vocabularies. Neutral lexis, which constitutes a clear example of adding negative attitudes to otherwise attitude-free lexical items, consists essentially of verbal dysphemisms, as nouns and adjectives are usually dichotomized in terms of positivity and negativity in natural language. These dysphemisms may replace a host of neutral verbs in different Arabic vernaculars when the need for a negative attitude arises. Below is a list of neutral verbs used in the 3rd person masculine singular along with their dysphemistic counterparts in Jordanian Arabic (JA), Egyptian Arabic (EA), Kuwaiti Arabic (KA), and Saudi Arabic (SA):

(1) a. *bōkil/bākul* ‘he is eating’: *biddafla* ‘lit. he is eating bamboo’ (JA), *bitsammim* ‘he is poisoning himself’ (EA), *yaʿlif* ‘lit. he is eating animal food’ (KA), *yafrum* ‘lit. he is grinding’ (SA)
b. *nāyim* ‘he is sleeping’: *minġamid* ‘lit. he is put back in his case’ (JA), *maxmūd* ‘lit. he has been extinguished’ (EA), *xāmid* ‘lit. he is motionless’ (KA and SA)
c. *rāḥ/maša/miši* ‘he left’: *ingalaʿ* ‘lit. he was extracted’ (JA), *ġār* ‘lit. he vanished’ (EA), *ḏalaf* ‘?’ (KA), *walla* ‘lit. he left, giving us his back’ (SA)

For their part, positive lexical items are sometimes a target for dysphemizing in vernacular Arabic. An interesting and frequent example in most Arabic dialects is the masculine oriented dysphemism of the adjective *ḥilwa* ‘beautiful [to describe a girl/woman]’ as follows: *šigfih* ‘a piece’ (JA), *muzza* ‘?’ (EA), *giṭʿa* ‘a piece’ (KA and SA). Other things being equal, these dysphemisms can be felicitously utilized only in male talk to indicate the out-of-the-ordinary beauty and/or sexual attractiveness of the referent. It is inappropriate to use it in addressing a girl/woman, because the freshly acquired attitude is dysphemistic in nature, even though a socially positive attribute is being intensified. The following are more examples from Jordanian Arabic: *šabbīḥ* ‘?’ instead of *šāṭir* ‘smart’, *habbāj* ‘?’ for *karīm* ‘generous’, *malʿūn wāldēn* ‘a man with damned parents’ instead of *šujāʿ* ‘brave man’.

Lexical items representing the negative or unfavorable part of human experience are a common input for dysphemism. Some socially or inherently negative attributes can be made more negative by dysphemizing them. Below are two negative attributes along with their dysphemistic counterparts in four Arabic dialects:

(2) a. *nāṣḥa/samīna* ‘fat [fem.]’: *dubba/bagara* ‘she-bear/cow’ (JA), *ʿiglah* ‘she-calf’ (EA), *baṭṭa/dubba* ‘duck/she-bear’ (KA), *fīl* ‘elephant’ (SA)
b. *ġabī/ʾaḥmaq* ‘stupid’: *ḥmār/bhīm* ‘donkey’ (JA), *ḥumār/bahīm* ‘donkey’ (EA), *ʾaṯwal* ‘?’ (KA), *dilx* ‘?’ (SA)

Negative or unfavorable verbs are also a target for dysphemism. Below are two examples:

(3) a. *rasab* ‘he failed in a test’: *kawwaʿ* ‘lit. he rested his head on his arm’ (JA and EA), *gawwaʿ* ‘lit. he fell to the bottom’ (KA and SA)
b. *māt* ‘he died’: *inḥarag* ‘lit. he got burned (JA), *fiṭis* ‘he died [animal]’ (EA), *fangaš* ‘?’ (KA), *faṭas* ‘he died [animal]’ (SA)

Death is a taboo subject and a common target for euphemisms. However, the above dysphemistic death terms intensify the negativity of this part of human experience by adding lexicalized negative rather than positive attitudes.

The negativity of some inherently negative lexical items may be diminished by dysphemizing. The output of this lexical process is what Allan and Burridge (1991) call 'euphemistic dysphemisms'. For example, the Egyptian word *xayyaṭ* 'he sewed' in the utterance *xayyaṭ-ha* 'he sewed her' instead of the taboo word *nāk* 'he fucked' in the utterance *nāk-ha* 'he fucked her' is meant to euphemistically dysphemize the inherently negative lexical item *nāk*. The result is lexicalization that avoids the explicit mention of a taboo lexeme in favor of a dysphemism that succeeds in only lessening the degree of negativity. Similarly, the word *naffas* 'lit. he leaked' in Jordanian Arabic manages to reduce the negativity of the taboo word *fasā* 'he broke wind noiselessly' by a euphemistic dysphemism.

As can be observed from the inter-dialectal data given in this entry, there are striking similarities between dysphemisms among Arabic vernaculars. For example, *faṭas/fiṭis* 'to die [animal]', probably among other variants, is dysphemistic in most Arabic dialects when referring to human deaths. However, there are some words which mean different things in different Arabic dialects and are often used to generate humor. Two interesting examples come from Kuwaiti and Jordanian Arabic. In Kuwaiti Arabic, the verb *ṭāg* means 'to beat/hit', corresponding to the verb *katal* in Jordanian Arabic. Interestingly, however, *ṭāg* is a dysphemistic counterpart for *nāk* 'to fuck' in Jordanian Arabic. Thus, the utterance *ṭāg-ha* means 'he hit her' in Kuwaiti Arabic, but 'he screwed her' in Jordanian Arabic. Another example is the dysphemism *baṭṭa* 'duck' in reference to a girl/woman. In Jordanian Arabic, it indicates the sexual attractiveness of a female, whereas in Kuwaiti Arabic it refers to the plumpness of a female.

Bibliographical references

Allan, Keith and Kate Burridge. 1991. *Euphemism and dysphemism: Language used as shield and weapon*. New York: Oxford University Press.

Austin, J.L. 1962. *How to do things with words*. Oxford: Clarendon Press.

Brown, Penelope and Steven Levinson. 1987. *Politeness: Some universals in language*. Cambridge: Cambridge University Press.

Farghal, Mohammed. 1995a "Euphemism in Arabic: A Gricean interpretation". *Anthropological Linguistics* 37:3.366–378.

——. 1995b. "Dysphemism in Jordanian Arabic". *Zeitschrift fur Arabische Linguistik* 30.50–61.

——. 2003. "Arabic conservatism vs. English openness". *International Journal of Arabic-English Studies* 4.141–151.

Thomas, Jenny. 1995. *Meaning in interaction: An introduction to pragmatics*. London: Longman.

Sinclair, John. 1991. *Corpus, concordance, collocation*. Oxford: Oxford University Press.

Mohammed Farghal (Kuwait University)

E

East Africa

1. INTRODUCTION

Much has been written on the history of Arabic, Arabs, and Islam in East Africa, their influence on the peoples, languages, and cultures of the region, and the status of Arabic and Islam there (Lodhi 1994a; Lodhi and Westerlund 1994 and 1999). Particular attention has been paid to the impact of Arabic on → Swahili. A few recent publications deal with the question of the status of Arabic in East Africa, Arabic lexical borrowings, and structural intrusion in Swahili (Lodhi 1986a, 1986b, 1992, 1994b).

2. THE STATUS OF ARABIC IN EAST AFRICA

Arabic in East Africa has a minimal formal and academic recognition in spite of its historical predominance on the East African littoral and the rim of the Indian Ocean in general. Arabic has had an enormous impact on the languages spoken by Muslim communities of the Indian Ocean lands, and particularly on Swahili, the most widely used literary indigenous language in Black Africa (Polomé 1967; Whiteley 1969; Nurse and Hinnebusch 1993).

In East Africa, Arabic has never been more important than it is today when Arabic items in Swahili are increasing and automatically being loaned into most other languages of the region (Polomé 1980a and 1980b; Lodhi 1986a, 2000). Arabic items in Swahili are not properly documented or satisfactorily analyzed in spite of several existing studies in this field (Krumm 1940; Allen 1945; Růžička 1953; Imberg 1973, 1975, 1977; Baldi 1976, 1988; Zawawi 1979; Cassels 1984; Bosha 1993; Lodhi 2000).

Up to the conclusion of the 'scramble for East Africa' with the signing of the Heligoland Treaty in 1890 whereby the Sultanate of Zanzibar became a British protectorate, Arabic had been the sole language of administration, commerce, diplomacy, education, and liturgy in Muslim East Africa. Swahili gradually replaced Arabic in many fields during the 30 years of German occupation of Tanganyika, but after the First World War and the British takeover of Tanganyika, English was formally encouraged and spread there at the expense of both Arabic and Swahili.

Arabic was replaced in all formal contexts except for the following: (a) constitutionally, Arabic was the first official language of the Zanzibar Protectorate/Sultanate (including the Kenya Coastal Strip Protectorate), followed by English and Swahili in descending order of importance, up to the republican revolution in January 1964 when the linguistic recognition was reversed, i.e. Swahili, English, and lastly Arabic; (b) the national anthem of the Sultanate of Zanzibar was in Arabic; (c) Arabic was a compulsory subject at Swahili-medium primary schools and the Muslim Academy which trained teachers of Arabic, Swahili, and *diana* (Islamic theology). Arabic was the medium of instruction up to Class 4 only at the Arabic-speaking primary school at Vikokotoni in Zanzibar Town; and (d) Arabic was offered as an option at the secondary, high school, and teacher training levels. Some *Qur'ān* schools (*chuo/vyuo/kutab/madrassa*), which were all private, also offered in the afternoon classes a minimum of instruction in Arabic to senior pupils who attended primary schools in the morning and who did not have Arabic as a subject.

During the colonial educational expansion up to the middle of the 1920s in Zanzibar, the educational status of Arabic was drastically reduced. In 1910, some government schools in rural areas closed down because of lack of pupils. Most parents boycotted the schools where English was the medium of instruction from the first year of the primary school, but where Arabic was a compulsory subject. Swahili was offered later, in the Roman script, after the standardization of 1924. However, there was no demand from parents to use Swahili as the medium of instruction in the schools of Zanzibar – they wanted Arabic, "the language of their Prophet" (Bennett 1978:229). According to a Zanzibar Government Report in 1939, 47.5 percent of Zanzibaris were literate in Swahili in the Arabic script. However, the colonial government continued publishing its Swahili documents in the Roman script in spite of the fact that only 2 percent of Zanzibaris were literate in it (Bennett 1978).

After the revolution, Arabic gradually disappeared from the school syllabus in Zanzibar during the eight-year-long anti-intellectual reign of Sheikh Abeid Aman Karume, the first president of the People's Republic of Zanzibar. The Muslim Academy was closed down and its functions partly incorporated with the Nkrumah Teachers College at Beit-el-Ras a few kilometers north of Zanzibar Town. Following the assassination of Sheikh Karume in 1972, during the reign of his successor Alhajj Aboud Jumbe, a revival of Arabic was realized. Since 1980, it has again been a compulsory subject from Class 3 at the primary level, though a number of secondary schools do not offer it for lack of teachers or teaching materials. It is not yet compulsory in the current teacher training programs; instead, Arabic teachers are recruited through the Institute of Kiswahili and Foreign Languages (IKFL) at the State University of Zanzibar (SUZA), and the private Zanzibar University in which the former Islamic College (the re-established Muslim Academy) has been incorporated. Both these institutions provide training in educational theory and practice. It is a compulsory subject at the Zanzibar University, whereas at the IKFL it is optional.

However, during the 2002 calendar year, 46 out of the 300 diploma students at the IKFL had Arabic as their major subject. With this raised status, Arabic is the third official language in autonomous Zanzibar and has a *de facto* importance of its own as a language of religious transmission in a predominantly Islamic country (97 percent Muslim). This has encouraged local and private initiatives to start institutes of Islamic studies in other parts of East Africa. An Islamic university similar to the one in Zanzibar, and having several secular disciplines, has been founded in Mbale, Uganda, to cater for the whole region.

In Mainland Tanzania with a population that is approximatey 50 percent Muslim (Kettani 1982), Arabic has its social status among the Muslims and is taught in the *Qur'ān* schools as also in Kenya, but the teaching is mostly limited to 'reading' the *Qur'ān*. A large proportion of the population of East Africa (i.e. Kenya, Uganda, Tanzania, Comoros, Mozambique, Malawi, Rwanda, and Burundi) is Muslim with Arabic as the spiritual language (Kettani 1982). Arabic is also the 'Latin' of Swahili, the main language in this vast region. Although about 42 percent of the Swahili vocabulary is of Arabic origin (Bosha 1993), Arabic is not offered at any level in the secular educational system of Mainland Tanzania (or in Kenya and Uganda), a system mainly in the hands of numerous Christian missions. It is included in neither the programs of the Institute of Kiswahili Research (IKR), nor the Department of Kiswahili and African Languages at the University of Daressalaam. Only an extramural course is occasionally offered at the Institute of Adult Education in Daressalaam, but at the university, no graduate course in Arabic is offered.

During the 1960s, for a couple of terms only, a short introductory extramural course in Arabic was offered to Swahili language students. These courses were removed from the university program after the socialist/nationalist Arusha Declaration in 1967, in an effort to 'decolonize' Tanzania, in spite of increasing contacts with the socialist Arabic-speaking countries.

At the IKR, no specialist is working with Swahili manuscripts in the Arabic script or Arabic language manuscripts. In the archives of the IKR there are more than a thousand such manuscripts listed in the catalogue of Allen (1970). However, Arabic loans are both important and popular. Some of the many hundreds of recent terms approved by the National Swahili Council (Baraza la Kiswahili la Taifa/BAKITA) are: *mhifadhina* 'reactionary', *msamiati* 'vocabulary', *wakala* 'bill of lading', *mwakilishi* 'elected representative', *dhidi ya* 'against', *thaura* 'political revolution' and *harakati ya tabaka* 'class struggle' (Bosha 1993). These contributions are by Swahili experts with a Muslim background and knowledge of Arabic. Swahili lexicography necessitates at least a working knowledge of Arabic, especially for people without a Muslim background.

The language typology in Tanzania Mainland shows that Arabic appears in bilingual, trilingual, quadrilingual, and plurilingual situations, but not in monolingual ones. Until about three decades ago Arabic appeared in some multilingual shop notices (together with English, Gujarati, and Swahili). Arabic has completely disappeared from auctions and the marketplace, and Arabisms such as *arbata-ashara* 'fourteen', *khamso-ishirin* 'twenty five' and *sitaa-alf* 'six thousand' have been replaced by original Swahili or Swahilized Arabic terms *kumi na nne*, *ishirini na tano*, and *elfu sita* respectively. In the mid-1960s, Arabic terms such as these were frequently used at auctions, just as Hindi/Urdu terms such as *do chai* 'two teas' and *tiin kafi* 'three coffees' were used in many restaurants. "The proportion of those who can read Arabic appears to drop slightly with increasing education from 4% at primary level to 2.8% at secondary Form VI level. Many primary school children have also received instruction at Koranic schools but such schools have not traditionally fostered academic study in the way that Christian Mission schools have" (Hill 1980:223).

The status of Arabic in Kenya, with a population that is approximately 25 percent Muslim (Kettani 1982), is similar to that in Tanzania Mainland, though on the Kenya coast Arabic has more prestige, due to its historical geographical proximity and ethnic affinity to southern Arabia, particularly Hadramawt. In predominantly traditional Swahili societies of the Lamu archipelago and Malindi, Arabic is taught in some primary schools and many parents send their children to the Middle East for further studies. It is not uncommon to find Arabic as a second or third language in the Muslim families of the Kenya coast and among Muslims upcountry. Instruction in Arabic is given at numerous *Qurʾān* schools run by about 120 Islamic societies. The Kenya Muslim Association has plans to establish Muslim secondary and high schools in Mombasa and Nairobi with emphasis on Arabic and Islamic studies.

In Uganda, where approximately 25 percent of the population is Muslim (Kettani 1982), the situation of Arabic is rather similar to that in Kenya and Tanzania Mainland in that it has no official recognition. It is taught only at *Qurʾān* schools and Islamic institutes. Their activities are coordinated by the Uganda Supreme Islamic Council, which has established the Uganda Muslim University at Mbale, a joint project by the Uganda Government and the Organization of Islamic Conference (OIC). A major reason given by Muslim leaders (including two former presidents of Zanzibar, Alhajj Aboud Jumbe and Dr. Salmin Amour) for the establishment of the Muslim universities in Uganda and Zanzibar was that theological colleges and faculties at the established universities in East Africa dealt only with Christian theology, missiology and history of Christianity, and peripherally traditional African religions, but not at all with Islam. However, at the turn of the 19th century, Arabic was important in the military camps in Uganda; Furley (1959:321) suggests poor knowledge of Arabic on the part of the British officers as one major reason for the mutiny in Uganda in 1897.

Malawi and Mozambique also have large Muslim minorities, approximately 20 and 30 percent respectively (Kettani 1982), but the Muslims are loosely organized and instruction in Arabic is almost non-existent. However, knowledge of the Arabic script is widespread. The Muslim clergy from these countries, as well as from Burundi, Rwanda, and Zaire (with less than 5 percent Muslim population in all), usually receive their higher education in Tanzania, Kenya, or Uganda. Instruction in Arabic is offered at most *madrassa*s run by the mosques in these countries.

The Islamic Federal Republic of the → Comoros (the Comoro Islands) is wholly Muslim – the fourth island in the Comorian archipelago, Mayotte/Maore, is almost wholly Muslim and continues to be a part of France. Komorian (Shingazija/Shikomoro) and Arabic are the national languages with French as the second official language of the Islamic Federal Republic. Swahili is generally understood and both Shikomoro and Swahili are written in the Roman as well as Arabic scripts; however, recent linguistic development is pointing toward the spread of the Latin orthography (Ben Ali 1983; Ottenheimer and Ottenheimer 1976), although the republic is the only country in Bantu-speaking Africa to be a member of the Arab League since 1994.

In Madagascar, despite several dozen Arabic loans in the northern dialects of → Malagasy, which was earlier written in the Arabic script, Arabic as a language is almost non-existent with a dwindling Muslim population (Munthe 1987). Arabic script was, however, used by the early Christian missionaries there just as in Tanganyika and Kenya.

The Swahili-Arabic script is still in use among Muslims in private correspondence, poetry-writing, and religious instruction in the whole region, but there has been no newspaper or bulletin issued in this script since 1963, although during 1969/70 the Bible Society in East Africa published in Nairobi the Swahili versions of the Acts, St. Matthew, St. Mark, St. Luke, and St. John in the Swahili-Arabic script, printed in Stockholm for free distribution among old Christian and Muslim East Africans who were literate only in the Arabic script (→ Arabic Alphabet: for other languages).

3. ARABIC LEXICAL LOANS IN SWAHILI

"The most visible sign of outside encroachment in Swahili is in lexis, and the largest identifiable set of borrowed lexis almost certainly stems from Omani Arabic in the last three centuries or so" (Nurse and Hinnebusch 1993:321). Reinhardt (1894) has given an excellent description of this → Omani Arabic dialect as spoken in East Africa during the peak years of Omani dominance there.

The most exhaustive study of Arabic loans in → Swahili is by Bosha (1993), one of the few researchers on the subject who are native speakers of Arabic. The studies undertaken so far on the influence of Arabic and/or Arabic loans in Swahili have concentrated on the loans as such, and on suggesting etymologies of mostly nouns (Krapf 1882; Krumm 1940; Zawawi 1979). Relatively much has been written on the Swahili-Arabic script (e.g. Velten 1901; Allen 1945, 1970; Polomé 1967; Imberg 1975). Those few who have attempted to give the etymologies of Arabic verbs have in some cases given the wrong Arabic verb form as the immediate source (e.g. Johnson 1939; Sacleux 1939). Swahili has borrowed from several of these verb forms. Imberg (1975) has pointed out these shortcomings in his essay, and McCall (1969) in his long article has analyzed Krumm's classic work on Oriental loans from a sociological and historical perspective on borrowing.

The efforts to encourage Swahilization in Tanzania and Kenya and limit borrowing from English, except for stabilized Anglicizations like *kesi* 'case', *kuripoti* 'to report', have increased drastically the number and frequency of both direct and indirect Arabic loans in East Africa, and in some cases even established English loans have been replaced with Arabic, Arabic-Bantu, or purely Bantu elements in the fields of administration, law, mechanics, and even Christian theology (KAMUSI 1981), e.g. (with Arabic elements in heavy type) ***taarifa*** 'report', ***mahakama*** 'court', ***hakimu*** 'judge', *nguvu **farasi*** 'horse power', and ***waraka*** 'epistle' instead of *ripoti*, *korti*, *jaji*, *hosipawa*, and *epistola* respectively. In the various sciences, the specialized terminologies have been greatly expanded with the help of Arabic loans such as *ku**sharabu*** 'to absorb', *ki**sharab**io* 'absorbent', ***usharabu*** 'absorption', ***ukabila*** 'tribalism', ***utaifa*** 'nationalism', *ku**taifi**sha* 'to nationalize', *ku**thibiti**sha* 'to probate', *ma**jarib**io* 'probation, experiments', ***msamaha*** 'amnesty', and ***hisabati*** 'mathematics'.

Numerous new compounds with Arabic elements and affixation with Arabic roots or reductions have given rise to specialized terminologies similar to the Greco-Latin compounds in the European languages (Lodhi 1986a:260): (a) *mwana-**sheria*** 'lawyer', ***kibadili**-mwendo* 'cam', ***nusu**-kipenyo* 'radius'; (b) ***da**kuvu* 'fungicide', ***da**buibui* 'arachnicide', ***da**dudu* 'insecticide', ***da**nyungu* 'nematicide', ***da**kono* 'molluscicide', ***da**gugu* 'herbicide' (the prefix **da**- here is derived from the Arabic loan *dawā'* 'drug, medicine, chemical'); and (c) ***elimunafsi*** 'psychology', ***elimu**viumbe* 'zoology', ***elimujamii*** 'sociology', ***elimu**bantu* 'Bantuistics', ***elimumadini*** 'mineralogy', and ***elimu**mimea* 'botany' (from Arabic *ʿilm* 'science').

A large number of Swahili nouns are derived from Arabic roots. A very common way of producing Swahili lexis is to borrow the various forms already existing in Arabic and Swahilize them with an anaptyctic or epenthetic vowel (Polomé 1967:166–176; Cassels 1984), e.g. ***hisabu/hesabu*** 'counting, accounts' from which *mahisabu* 'figures', and ***hisabati*** 'mathematics' are derived; ***haraka*** 'to hurry' , ***harakati*** 'struggle'; ***safiri*** 'to travel', ***msafiri*** 'traveler', ***safari*** 'a journey', ***msafara*** 'caravan'; ***fikiri*** 'to think', ***fikira/fikra*** 'thought', ***fikara*** 'worries', ***tafkira*** 'reflections'.

Arabic broken plurals are also found as loans in Swahili, though few in number, as synonyms of Swahilized plurals, e.g. for the Swahilized singulars, ***binti*** 'daughter', ***sahaba*** 'companion', ***sahibu*** 'friend', ***walad*** 'child, boy', Swahilized plurals *mabinti, masahaba, masahibu*, *mawaladi* are used along with Arabic broken plural loans ***banati, as-haba*** 'companion of the Prophet Muḥammad', ***as-habu, awlad, uladi*** 'descendants'.

A small number of Arabic plurals also appear as singular nouns in Swahili and are pluralized as Bantu roots or stems, e.g. ***muhajirina/mhajirina*** 'refugee', ***mshirikina*** 'one believing in superstition, magic, etc.', and ***mhifadhina*** 'conservative'. Alternatively, some singular Arabic nouns are treated as plural and singular forms are derived from them, e.g. Arabic ***mismār*** 'nail, pin' > Swahili *misumari* (pl.) > *msumari* (sg.).

Swahili adjectives of Arabic origin outnumber those of Bantu origin. Many of them are borrowed in the adjectival form, whereas others are constructed or derived. Loans such as ***dhaifu*** 'weak', ***ghali*** 'expensive', ***huru*** 'free', ***laini*** 'soft', ***maskini*** 'poor', ***nadhifu*** 'pure', ***safi*** 'clean', and about 50 more items belong to the first category. These are not inflected as is the case with the Bantu adjectives. Constructed or derived adjectives such as ***aminifu*** 'reliable, honest' from ***kuamini*** 'to believe in, to trust' and ***badhirifu*** 'extravagant, prodigal', from ***kubadhiri*** 'to squander, to waste' belong to the second, lesser category. Others are constructed from nouns, e.g. ***fakiri*** 'a pauper' > ***fukara*** 'poor'. However, numerous other adjectival concepts are expressed by phrases constructed with Arabic roots, e.g. *mtu wa* ***haki*** 'a just man', *maneno ya* ***kutibu*** 'soothing words', *mlango wa* ***saba*** 'the seventh door', *nyumba za* ***zamani*** 'old buildings', *gari iliyo****haribika*** 'a broken-down car', *mwenye* ***mali*** 'a wealthy person', *mti mwenye* ***maradhi*** 'a sick tree', *watu* ***tajiri*** 'rich people' and m**tawala dhalimu** 'oppressive ruler'. A large number of these loans (together with other Oriental and European loans) have been borrowed from Swahili into many other languages of East Africa as indirect loans.

4. ARABIC STRUCTURAL LOANS IN SWAHILI

Arabic grammatical or structural loans include Arabic or Bantu-Arabic adverbs, conjunctions, and prepositions and extension of Arabic loan verbs and their extensions in Swahili. A detailed description and analysis of Arabic structural loans and hypotactic structures in Swahili is given by Lodhi (1994b, 2000a:99–120, 2000b). In all three cases of structures with adverbs, conjunctions, and prepositions, there is a marked tendency to first use Bantu constructions, followed by Arabic-Bantu phrases to be finally reduced to Arabic independent function words. This advanced use of Arabic loans simplifies Bantu syntax by reducing subordinate clauses to infinitive phrases.

Arabic loans appear in all word classes in Swahili. In many cases they appear as synonyms to Bantu lexical items; in some cases Bantu-Arabic phrases appear as function words; and in some cases the loans have replaced Bantu items. Moreover, items from all these word classes are spreading further as indirect Arabic loans from Swahili to almost all languages of East Africa.

In many Swahili dictionaries, one or more of the following items belonging to different word classes are omitted, though they are all commonly used by native Swahili speakers and other speakers of Swahili: *abadan* 'never', *aidhan* 'also, moreover', *daiman* 'always', *hususan* 'especially', *mathalan/mathalani* 'as, for example', *shukran* 'thank you', *tabaan* 'of course', *takriban* 'nearly, approximately', *wa baada* 'and then', *wa baadahu* 'and after that', *wa katabahu* 'Yours sincerely [lit. And he who has written is . . .]'. This may be because the foregoing are viewed by many non-native speakers of Swahili as Arabisms because of their Arabic *-an* and *-hu* endings and the forward shift in accent. So far, most Swahili lexicographers and/or their assistants or informants have been non-Muslims or non-native speakers of Swahili, without much exposure to Arabic and Islam.

There is an abundance of Arabic grammatical or structural loans in Swahili, which the other languages of East Africa borrow freely from

Swahili. Yet, syntactic Arabic intrusion in Swahili is not a widespread phenomenon since it is limited to only a few structures in the languages of the region, where no Arabic-based pidgin or creole variants have developed. Swahili borrowed three new vowel sequences (*ai*, *au*, *ei*) and several consonants (*th*, *dh*, *kh*, *gh*) and consonantal combinations (*st*, *sht*) from Arabic, "But Swahili had if anything a substrate relationship to Arabic, and all these features are simply borrowed" (Nurse 1996:280, 291).

Even though among non-native speakers of Swahili, especially Christians and/or people away from the Swahili coast, there has always been a "tendency to use Bantu words which usually are cognate with the inland vernacular in which the Swahili speaker received his first education or with the Bantu languages which he currently uses in his narrow tribal circle" (Polomé 1967:166), there is no conscious negative attitude developing toward Arabic elements of Swahili. On the contrary, Arabic continues to make important contributions to the development of the modern Swahili lexicon, and indirectly the lexicon of other East African languages.

Arabic is unique in this respect since the other major contributor languages in East Africa which came from the Indian Ocean (Persian, Cutchi/ Sindhi, Gujarati, English, and Portuguese) have contributed mostly nominals. There are only a few Standard Swahili verbs of English, Persian, and Indic origin. Swahili has borrowed very few adverbs, adjectives, conjunctions, or prepositions from languages other than Arabic, and this phenomenon has not yet been thoroughly studied.

In theory, it is legitimate to use in Swahili any Arabic word of any word-class, since Arabic is the 'Latin' of Swahili and Persian was the 'Greek' in the past. However, it is English which is the largest language contributor to East Africa today, but its contribution is limited to nominals belonging primarily to the fields of modern technology and science (Lodhi 1986a:256–260).

Many Muslim leaders and scholars in East Africa claim that a bright future for Arabic is associated with the growth and strengthening of Swahili and better educational facilities in the predominantly Muslim areas of East Africa. They also believe that reforming and reorganizing the *Qur'ān* schools would improve the quality of religious instruction as well as strengthen the position of Arabic. The *Qur'ān* school system is the origin of the Islamic colleges and institutes, whether established with or without the medium of Arabic. The stagnation of the *Qur'ān* schools and the afternoon or evening *darsa/madarsa* (sessions connected with the mosque) from the middle of the 1960s to the 1980s dealt a severe blow to further growth of traditional Swahili poetry, intellectual exchange, and scholarly production in the fields of *diana*, philosophy, interpretation of the Islamic *šarīʿa*, and Afro-Oriental herbal medicine. The future development of Swahili and Swahili institutions is closely associated with further progress of Islamic institutions and recognition of Arabic in East Africa. To the East Africans, Arabic is not only a foreign colonial language like English, it is also, unlike English, an integral part of the Swahili language, literature, and culture in general. Most Arabic loans are not considered foreign because of their high frequency and commonness. In light of this, a positive change in the attitude to Arabic has been observed in recent years.

BIBLIOGRAPHICAL REFERENCES

Allen, John. W.T. 1945. *Arabic script for students of Swahili: Manuscripts in the Library of the University of Daressalaam*. Daressalaam: Institute of Kiswahili Research.

——. 1970. *The Swahili and Arabic manuscripts in the Library of the University of Daressalaam: A catalogue*. Leiden: E.J. Brill.

Baldi, Sergio. 1976. *A contribution to the Swahili maritime terminology*. Rome: Istituto Italo-Africana.

——. 1988. *A first ethnographic comparison of Arabic loanwords common to Hausa and Swahili*. (= Suppl. 57 to *Annali del Istituto Orientale di Napoli* 48:4.) Naples.

Ben Ali, Damir. 1983. "Hali na hadhi ya Kiswahili katika Visiwa vya Komoro" [The status of Swahili in the Comoro Islands]. Paper presented at the International Seminar on the Standardisation of Swahili Terminologies, 26/9–1/10 1983, Institute of Kiswahili Research, Daressalaam.

Bennett, Norman R. 1978. *A history of the Arab state of Zanzibar*. London: Methuen

Bosha, Ibrahim. 1993. *Taathira za Kiarabu katika Kiswahili pamoja na kamusi thulathiya* [*The influence of Arabic language on Kiswahili with a trilingual dictionary*], *Kiswahili-Kiarabu-Kiingereza/Swahil-Arabic-English*. Daressalaam: Institute of Kiswahili Research.

Cassels, David A. 1984. "Vowel intrusion and change in Swahili words of Arabic origin". *Swahili Language and Society* 1.32–40.

Furley, Oliver. 1959. "The Sudanese troops in Uganda". *African Affairs* 58.311–28.

Hill, Carol P. 1980. "Library users and their reading preferences". Polomé and Hill (1980:206–228).

Imberg, Gösta. 1973. "Till frågan om swahilispråkets arabiska element" [On the question of Arabic elements in Swahili]. Unpublished seminal paper (Semitic languages D1), Department of Middle Eastern Studies, Lund University.

——. 1975. *Handledning i den swahili-arabiska skriften* [A handbook of the Swahili-Arabic script]. Uppsala University, Department of Asian and African Languages. Compendium.

——. 1977. "Islam i Östafrika: En studie i Östafrikas islam med särskild inriktning på swahili-islamisk litteratur" [Islam in East Africa: A study in East African Islam with special emphasis on Swahili-Islamic literature]. Unpublished seminal paper in History of Religion, Uppsala University.

Johnson, Frederick. 1939. *A standard Swahili-English dictionary*. London: Oxford University Press. (Repr. 1965.)

Kettani, M. Ali. 1982. "Muslim East Africa: An overview". *Journal of the Institute of Muslim Minority Affairs* 1:2.104–119.

Krapf, J. Ludwig. 1882. *A dictionary of the Suaheli Language*. London and New Jersey: Gregg Press. (Repr., 1964.)

Krumm, Bernard. 1940. *Words of Oriental origin in Swahili*. London.

Lodhi, Abdulaziz Y. 1984–1986. "The status of Arabic in East Africa". *On the dignity of man: Oriental and classical studies in honour of Frithiof Rundgren*, ed. Tryggve Kronholm and Eva Riad, 257–262. Stockholm: Almqvist and Wiksell.

——. 1986. "The Arabs in Zanzibar: From the Sultanate to the People's Republic". *Journal of the Institute of Muslim Minority Affairs* 2.404–418.

——. 1992. "The role of Kiswahili in the context of Zanzibar". *Pre-papers of the International Conference on the History and Culture of Zanzibar*, 10–13 December 1992, Part 2:1–9. Zanzibar: Museum and Archives.

——. 1994a. "Muslims in Eastern Africa: Their past and present". *Nordic Journal of African Studies* 3:1.88–99.

——. 1994b. "Arabic grammatical loans in the languages of Eastern Africa". *Working Papers in Linguistics* 22.60–74. Trondheim: Institute of Language and Information Studies.

——. 2000a. *Oriental influences in Swahili: A study in language and culture contacts* (= *Orientalia et Africana Gothoburgensia*, 15.) Gothenburg: Acta Universitatis Gothoburgensis.

——. 2000b. "Arabic loans in Swahili". *Orientalia Suecana* 49.71–82.

—— and David Westerlund. 1994. "Afrikansk Islam i Tanzania". *Majoritetens Islam*, ed. Ingvar Svanberg and David Westerlund, 138–151. Stockholm: Arena.

——. 1999. "Tanzania". *Islam outside the Arab world*, ed. David Westerlund and Ingvar Svanberg, 97–110. Richmond, Surrey: Curzon.

McCall, Daniel F. 1969. "Swahili loanwords: Whence and when". *Eastern African history*, ed. Daniel F. McCall, Norman R. Bennett, and Jeffrey Butler, 28–73. New York, Washington, D.C., and London: Praeger.

Munthe, Ludvig. 1987. "The Arab influence on Madagascar". *Religion, development and African identity*, ed. K. Holst Petersen, 103–111. Uppsala: Nordic Africa Institute.

Nurse, Derek. 1997. "Prior pidginization and creolization in Swahili?". *Contact languages: A wider perspective*, ed. Sarah G. Thomason, 271–294. Amsterdam and Philadelphia: J. Benjamins.

—— and Thomas Hinnebusch. 1993. *Swahili and Sabaki: A linguistic history*. Berkeley: University of California Press.

Ottenheimer, Harriet and Martin Ottenheimer. 1976. "The classification of the languages of the Comoro Islands". *Anthropological Linguistics* 19.408–415.

Polomé, Edgar C. 1967. *Swahili language handbook*. Washington, D.C.: Center for Applied Linguistics.

——. 1980a. "The languages in Tanzania". Polomé and Hill (1980:3–25).

——. 1980b. "Swahili in Tanzania". Polomé and Hill (1980:79–102).

—— and Carol P. Hill (eds.) 1980. *Language in Tanzania*. London: Oxford University Press.

Reinhardt, Carl. 1894. *Ein arabischer Dialekt gesprochen in Oman und Zanzibar*. Stuttgart and Berlin: Spemann.

Růžička, Karel F. 1953. "Lehnwörter im Swahili". *Archív Orientální* 21.582–603.

Sacleux, Charles. 1939. *Dictionnaire swahili-français*. Paris: Institut d'Ethnologie.

Velten, Carl. 1901. *Praktische Anleitung zur Erlernung der Schrift der Suaheli*. Göttingen: Vandenhoeck and Ruprecht.

Whiteley, Wilfred H. 1969. *Swahili: The rise of a national language*. London: Methuen.

Zawawi, Sharifa. 1979. *Loan words and their effect on the classification of Swahili nominals*. Leiden: E.J. Brill.

Abdul Aziz Y. Lodhi (Uppsala, Sweden)

Educated Arabic

The terms 'Educated Arabic' (EA) or 'Educated Spoken Arabic' (ESA) are broad designations that refer to spoken Arabic showing the following features:

i. A higher, more formal register than the colloquial of primary discourses of familiarity among family and acquaintances;
ii. A mix of literary and colloquial lexical items;
iii. Absence of *ʾiʿrāb*, the markers of desinential inflection (case and mood).

Some scholars define Educated Arabic as the higher-register spoken variety of a particular country or region. "Of all spoken varieties it is the most dynamic, versatile, and the one readily understood outside its particular geographical region" (Badawi 1985:15). However, in addition to the educated form of speech used in a particular country, Educated Arabic also refers to a type of spontaneous inter-regional, 'mixed', or ‘koineized’ discourse used among Educated Arabic speakers when they need to accommodate differing regional dialects.

Additionally, in recent usage, Educated Arabic refers more and more frequently to unscripted spoken Arabic used in the Arabic broadcast media (→ ‘Media Arabic’) in interviews and in spontaneous commentary situations.

1. Educated Arabic, diglossia, and koine

It is important to note that the effort to distinguish this particular level of Arabic was influenced by work initiated by Ferguson (1959a, 1959b) on → diglossia and on what Ferguson termed ‘the Arabic koine’, an inter-regional Arabic lingua franca. The term ‘diglossia’ has traditionally been used to refer to the binary split between spoken and written Arabic; however, as several researchers have shown (Badawi, Hary, Mitchell) the differentiation between High (H) (literary) and Low (L) (colloquial) forms is considerably more shaded and nuanced than a straight bifurcation.

A key feature of Educated Arabic is that it is variable and realized differently by different speakers; there is no “institutionalized norm” (Mitchell 1985:53). Some Arabic scholars debate Educated Arabic’s existence as an identifiable register or variant. However, from the 1970s to the 2000s, a small but significant body of work has been growing that deals with the nature and principles of Educated Arabic speech.

2. Badawi’s premises

Perhaps the first to pin down Educated Arabic as an identifiable linguistic phenomenon was El-Said Badawi in his classic 1973 Arabic work, *Mustawayāt al-ʿarabiyya al-muʿāṣira* ‘Levels of contemporary Arabic’. In an attempt to systematically describe the wide range of variation that characterizes the Arabic language, Badawi proposed the concept of a continuum of five sociolinguistic ‘levels’, ranging from the most literary to the most colloquial. His later articles in English (1985, 1995) elaborated on these distinctions and in particular on the importance of level three, or what he termed ‘Educated Spoken Arabic’, for teaching Arabic as a foreign language. Badawi proposes the following levels and labels (1985:17):

Level one:	*fuṣḥā at-turāṯ*	Classical Arabic
Level two:	*fuṣḥā al-ʿaṣr*	Modern Standard Arabic
Level three:	*ʿāmmiyyat al-muṯaqqafīn*	Educated Spoken Arabic
Level four:	*ʿāmmiyyat al-mutanawwirīn*	Semi-literate Spoken Arabic
Level five:	*ʿāmmiyyat al-ʾummiyyīn*	Illiterate spoken Arabic

It is important to note that, as Badawi points out, the levels “are not segregated entities,” (1985:17) but shade into each other gradually, with a “graded continuum of features” (1995:35). He classifies Level two (Modern Standard Arabic) as “mostly written” rather than spoken, and Levels two and three as essentially “in complementary distribution” with each other (1985:19), that is, they function in separate spheres, with some overlap. These linguistic levels are characterized by two different sociolinguistic dimensions. First, they reflect the situations in which speakers find themselves – whether those situations are, for example, religious, formal, academic, casual, or intimate. Second, they are influenced by the educational and regional backgrounds of the individual speakers. A complex interplay of situational and interpersonal factors is therefore involved in the choice of language level in any Arabic speech situation.

In his 1995 article Badawi re-draws his distinction between SMSA (Spoken Modern Standard Arabic), which he considers severely restricted in spoken use, and Educated Arabic, a much more natural form of spoken Arabic for native speakers. Similarly, the Jordanian linguist Muhammad H. Ibrahim points out the difference between ‘standard’ (MSA) and ‘prestige’ (Educated Arabic) language within a diglossic community, affirming that, for speakers of Arabic, “a locally recognized standard of prestige exists apart from the standard H variety” (Ibrahim 1986:118) and that “there is a definite prestige norm in Arabic which is different from H” (1986:119).

3. Educated Arabic research

The research program at the University of Leeds, the Leeds Project on Educated Spoken Arabic, originated by T.F. Mitchell and sponsored by the Social Science Research Council of Great Britain was "based on a corpus of spoken Arabic collected in different parts of the Arab world in the late 1970's" (Mitchell 1994:xiii). This project produced a series of articles by Mitchell (1980, 1985, 1986, 1990) and others (for example, Agiùs 1990; El-Hassan 1978; Ibrahim 1986; Meiseles 1980; Sallam 1979) and one book (Mitchell and al-Hassan 1994) analyzing materials drawn from the corpus. Aside from this important project, there has been little or no systematic effort on the part of academic institutions (Arab or Western) to engage in long-term empirical investigation of spontaneous formalized discourse of educated native speakers, although several dissertations have appeared on the topic (Mehall 1999; Schmidt 1974; Schultz 1981; Wilmsen 1995). Current efforts at the University of Maryland's Center for Advanced Study of Language (CASL) to investigate and analyze spoken media Arabic will certainly add to the knowledge and understanding of Educated Arabic.

4. TERMS FOR EDUCATED ARABIC

The terminology used to refer to Educated Arabic, or Badawi's Level three has not yet been codified. 'Educated Spoken Arabic' (ESA) is the most widely-used term (see Agiùs 1990; El-Hassan 1978; Mahmoud 1982; Meiseles 1980; Mitchell 1980, 1985, 1986, 1990; Mitchell and El-Hassan 1994; Sallam 1979). However, researchers also use a range of other terms: 'Formal Spoken Arabic' (FSA) (Kayyal 1985; Ryding 1990, 1991, 1994, 1995; Ryding and Zaiback 1993; Schultz 1981); 'urban cultivated' Arabic (Abdulaziz 1986), 'middle Arabic' (Hary 1989; Mahmoud 1978a, 1978b; Mansoor 1960), 'pan-Arabic' (Abdelmassih 1975; Mitchell 1985), Standard Spoken Arabic (Haddad and Haddad 1984), the 'inter-regional standard' (Ibrahim 1986), 'supra-dialectal L' (SDL) (Ibrahim 1986), 'inter-Arabic' (Bishai 1966; Sawaie 1980), the 'inter-Arabic koine' (Mitchell 1986), 'the koineized colloquial' and 'the elevated colloquial' (Blanc 1960), the 'international koine' (Jernudd and Ibrahim 1986), and 'prestigious oral Arabic' (Drozdík 1996). The most frequent Arabic terms applied to this variety are *al-luġa al-wusṭā* 'the middle language', and *luġat al-muṯaqqafīn* 'the language of the cultivated'.

5. FEATURES OF EDUCATED ARABIC

Despite Badawi's and others' research and discussion of the various levels of formality, there is little consensus among Arabic linguists about the features of Educated Arabic, or the inter-mediate levels in general. As Elgibali states (1993:76), "we do not . . . have intuition or scholarly consensus concerning the number, discreteness and/or stability of the middle level(s)". Thus, what is realized, if anything, about Educated Arabic is that it is a "fluid norm" (Mitchell 1986:7) that is "tendency-governed" (Mitchell 1986:19) rather than consistent and rule-governed. Some scholars, however, have proposed certain guidelines. For example, Mitchell states definitively that Educated Arabic does not include the used of *ʾiʾrāb*, the markers of desinential inflection (case and mood). "If they use the *iʿraab*, then by definition they are not speaking Educated Arabic", but rather a form of Modern Standard Arabic (Mitchell 1986:19). Hary (1996:83) also notes that selection of certain morphological, phonological, and lexical features shows "the systematic nature" of Educated Arabic.

Ryding (1991) lists characteristics of Formal Spoken Arabic (Levantine) observed in use among educated Arabs. These are divided into categories of inflectional morphology, syntax (verb strings), modals, agreement features, and lexicon. In many ways these forms differ from Spoken Modern Standard Arabic. Some central features include the following:

i. In verb morphology, the dual forms are generally not used, nor are the feminine plurals. Final *nūn*s in present tense verb forms ending in *-ūna* or *-īna* are also omitted. Short vowel mood markers for indicative and subjunctive do not occur.
ii. In Formal Spoken Arabic noun morphology, cases are eliminated; where case is indicated by a long vowel suffix, as in the endings for the dual and sound masculine plural, cases are reduced to one form, *-īn* for the sound masculine plural and *-ēn* for the dual. Another feature is conversion of final nunation on indefinite defective nouns to a long vowel, e.g. *karāsī*, *layālī*.
iii. Reduction of the relative pronouns to one form: *illī*. Metathesis of short vowel endings for the 2nd and 3rd pers. sg. suffix pronouns to *-ak*, *-ik*, and *-uh*.

As Ryding (1991) states, “one of the most distinctive aspects of FSA syntax is the use of verb strings without the use of subordinating particles (which would be necessary in SMSA): e.g.: *aHibb aruuH ashuuf-uh* (‘I’d like to go and see him’)”. Another key feature of Formal Spoken Arabic is the use of verb strings with function words such as those that indicate possibility and necessity, *mumkin* and *lāzim*, without the use of periphrastic phrases or subordinating particles: *mumkin taˁṭī-nī l-flūs?* ‘Could you give me the money?’, or *lāzim ʾarūḥ* ‘I have to go’ (see Ryding 1991:214–216 for a more detailed description of these features).

6. KOINE AND CODE-MIXING

‘Koineization’ is another term used to refer to elevating and leveling the spoken Arabic idiom, especially in interdialectal situations. The term ‘koine’ refers to a type of language used as a *lingua franca*, for wide-ranging communicative situations among and between different speech communities. Ferguson’s 1959 article, “The Arabic koine”, proposed that urban cultivated Arabic throughout the Arab world shares features which do not directly and identifiably descend from Classical Arabic, and therefore that these may have come from a shared interdialectal standard that emerged during the time of the Islamic conquests and has been maintained since.

Other scholars who are reluctant to identify a particular level or register for educated speech prefer to focus on principled ‘code-mixing’, where native speakers of Arabic may shift automatically to a ‘mixed variety’ (mixing Modern Standard Arabic and features of dialectal Arabic) of speech determined by the situation, especially in broadcast media (see Eid 1988). This ‘mixed variety’ is not considered an entity with rules of its own, but rather a spontaneously generated construct. Alosh (1997:345–347) provides a three-page list of “colloquial lexical items in the speech of native speakers in a formal situation”, which includes elements such as function words, verbs, nouns, adverbs, and numbers. He points out that “variation occurs among speakers as well as within the language output of the individual speaker” (1997:110), highlighting the characteristic instability of this variety.

Alosh maintains that “so-called Educated Spoken Arabic” (1997:109) is not so much an entity in itself, as a situational strategy used by Arabic speakers “influenced by contextual, experiential, and personal factors” (1997:110). A number of the lexical items in his list (1997:345–47), however, require inflections specific to Educated Arabic or colloquial Arabic, thus representing considerably more than lexical replacives. For example, Alosh lists common colloquial verbs such as *xallā* ‘to let, permit’, *xallaṣ* ‘to finish’, *inbaṣaṭ* ‘to be pleased’, *ḥaṭṭ* ‘to put’, *ijā* ‘to come’, *rāḥ* ‘to go’, and *jāb* ‘to bring’, and function words, such as the relative pronoun *illī*, the negation words *miš* and *mā*, and the predicator of existence *fī(h)*. The selection of these colloquial lexical items, or ‘C-tokens’, as Alosh terms them (1997:109), entails the use of inflectional and syntactic features that characterize the structure of Educated Arabic and contrast with Modern Standard Arabic. For example, Educated Arabic verbal inflections typically do not include the number inflection for dual or the number and gender inflection for feminine plural; it would sound strange to inflect a verb such as *jāb* in the dual or feminine plural. Also, Educated Arabic geminate verbs (such as *ḥaṭṭ*) are inflected as weak verbs in the past tense (e.g. *ḥaṭṭēt* ‘I have put’). The negators *miš* and *mā* are rule-governed in their functions and distribution.

Thus, the selection of certain key lexical items in Educated Arabic entails inflectional morphology and syntactic structures consistent with those items. As more studies of spoken, unscripted media Arabic and Educated Arabic become available, the principles and features that native Arabic speakers use to raise their speech to more elevated levels should become “an interesting topic for a socio-linguistic study” (Eid 1988:53).

7. EDUCATED ARABIC FOR TEACHING ARABIC AS A FOREIGN LANGUAGE

For most of the history of Arabic teaching in the West, either Classical Arabic (CA) (Badawi’s Level one) or, more recently, Modern Standard Arabic (MSA) (Badawi’s Level two), have been the options of choice for almost all academic programs. As programs and curricula shifted into a more communicative gear in the 1990s, and as oral proficiency became a learning goal, Modern Standard Arabic materials have been adapted to be more communicative in their orientation. However, the fact remains that Modern Standard Arabic (and certainly Classical Arabic) are not spontaneously spoken languages, and not authentic vehicles for spoken proficiency. A few programs in the United States (e.g. the University of Michigan, Brigham Young University, Georgetown University) teach spoken vernacular Arabic in addition to Modern Standard Arabic.

The choice of whether to teach Educated Arabic or Formal Spoken Arabic as a viable spoken medium for non-native speakers has been discussed by Agiùs (1990), Badawi (1985), and Ryding (1991, 1994, 1995), with Ryding advocating its value for learners who need a spoken medium of expression that can be used in a wide variety of places and situations. It is the spoken medium of instruction at the United States State Department's Foreign Service Institute. In terms of Educated Arabic/Formal Spoken Arabic teaching materials, Kayyal (1985) has prepared a two-volume basic course in Formal Spoken Arabic, Ryding has published a course in Formal Spoken Arabic (Ryding 1990), Ryding and Zaiback published a short ('FAST' = Familiarization and Short-Term) course in Formal Spoken Arabic (1993), and Haddad and Haddad have also published a course in 'Standard Spoken Arabic' (1984).

BIBLIOGRAPHICAL REFERENCES

Abdel-Massih, Ernest T. 1975. *A sample lexicon of pan-Arabic*. Ann Arbor: Center for Near Eastern and North African Studies, University of Michigan.

Abdulaziz, Mohamed H. 1986. "Factors in the development of modern Arabic usage". *International Journal of the Sociology of Language* 62.11–24.

Agiùs, Dionisius A. 1990. "Which strategy for teaching Arabic?". *Diglossic tension: Teaching Arabic for communication*, ed. Dionysius A. Agiùs, 3–6. Leeds: Folia Scholastica.

Alosh, Mahdi. 1997. *Learner, text, and context in foreign language acquisition: An Arabic perspective*. Columbus, Ohio: Ohio State University Foreign Language Publications Office.

Badawi, El-Said M. 1973. *Mustawayāt al-ʿarabiyya al-muʿāṣira fī Miṣr*. Cairo: Dār al-Maʿārif.

——. 1985. "Educated Spoken Arabic: A problem in teaching Arabic as a foreign language". *Scientific and humanistic dimensions of language*, ed. Kurt R. Jankowsky, 15–22. Washington, D.C.: Georgetown University Press.

——. 1995. "The use of Arabic in Egyptian TV commercials: A language simulator for the training of teachers of Arabic as a foreign language". *Georgetown University Round Table on Languages and Linguistics 1995*, 33–39. Washington, D.C.: Georgetown University Press.

Bishai, Wilson B. 1966. "Modern inter-Arabic". *Journal of the American Oriental Society* 86.319–23.

Blanc, Haim. 1960. "Style variations in spoken Arabic: A sample of interdialectal educated conversation". *Contributions to Arabic linguistics*, ed. Charles A. Ferguson, 81–161. Cambridge: Harvard University.

Byrnes, Heidi. 2002. "Toward academic-level foreign language abilities: Reconsidering foundational assumptions, expanding pedagogical options". *Developing professional-level language proficiency*, ed. Betty Lou Leaver and Boris Shekhtman, 34–58. Cambridge: Cambridge University Press.

El-Hassan, S.A. 1978. "Educated spoken Arabic in Egypt and the Levant: A critical review of diglossia and related concepts". *Archivum Linguisticum* 8.112–132.

Elgibali, Alaa. 1993. "Stability and language variation in Arabic: Cairene and Kuwaiti dialects." *Perspectives on Arabic linguistics*, V, ed. Mushira Eid and Clive Holes, 75–96. Amsterdam and Philadelphia: J. Benjamins.

Ferguson, Charles A. 1959a. "Diglossia". *Word* 15.325– 340.

——. 1959b. "The Arabic koine". *Language* 35.616–630.

Haddad, Thuraya and Ayed Haddad. 1984. *Standard spoken Arabic*. Washington, D.C.: School of Advanced International Studies.

Hary, Benjamin. 1989. "Middle Arabic: Proposals for new terminology". *al-ʿArabiyya* 22.19–36.

——. 1996. "The importance of the language continuum in Arabic multiglossia". *Understanding Arabic: Essays in contemporary Arabic linguistics in honor of El-Said Badawi*, ed. Alaa Elgibali, 69–90. Cairo: American University in Cairo Press.

Ibrahim, Muhammad. 1986. "Standard and prestige language: A problem in Arabic sociolinguistics". *Anthropological Linguistics* 28:1.115–126.

Jernudd, Björn and Muhammad H. Ibrahim. 1986. "Introduction". *International Journal of the Sociology of Language* 61.4–6.

Kayyal, Ziad N. 1985. *Formal spoken Arabic basic course*. 2 vols. Tunis: Foreign Service Institute.

Mahmoud, Youssef. 1978a. "The emergence of Middle Arabic". Paper presented at N-WAVE Conference, Georgetown University, Washington, D.C.

——. 1978b. "The glottal stop in Middle Arabic: A study in language variation". Unpublished manuscript, Georgetown University.

——. 1982. "Towards a functional Arabic". *al-ʿArabiyya* 15.82–89.

——. 1986. "Arabic after diglossia". *The Fergusonian impact*, ed. Joshua A. Fishman, I, 237–251. Berlin: Mouton de Gruyter.

Mansoor, Menahem. 1960. "Arabic: What and when to teach?". *Report of the Tenth Annual Round Table Meeting on Linguistics and Language Studies*, ed. Richard S. Harrell, 83–96. Washington, D.C.: George-town University Press.

Mehall, David John. 1999. *The verb morphology of unscripted media Arabic*. Ph.D. diss., Georgetown University.

Meiseles, Gustav. 1980. “Educated spoken Arabic and the Arabic language continuum”. *Archivum Linguisticum* 11.118–143.

Mitchell, T. F. 1980. “Dimensions of style in a grammar of educated spoken Arabic”. *Archivum Linguisticum* 11.89–106.

——. 1985. “Sociolinguistic and stylistic dimensions of the educated spoken Arabic of Egypt and the Levant”. *Language standards and their codification*, ed. Douglas J. Woods, 42–57. Exeter: Exeter University Press.

——. 1986. “What is educated spoken Arabic?”. *International Journal of the Sociology of Language* 61.7–32.

——. 1990. “The mixture not as before: In search of an Arabic spoken norm”. *Diglossic tension: Teaching Arabic for communication*, ed. Dionysius A. Agiùs, 18–26. Leeds: Folia Scholastica.

—— and Shahir El-Hassan. 1994. *Modality, mood and aspect in spoken Arabic*. London: Kegan Paul International.

Ryding, Karin C. 1990. *Formal spoken Arabic: Basic course*. Washington, D.C.: Georgetown University Press.

——. 1991. “Proficiency despite diglossia: A new approach for Arabic”. *Modern Language Journal* 75:2.212–218.

——. 1994. “Fostering a learning community for Arabic”. *Theory into Practice* 33:1.23–28.

——. 1995. “Discourse competence in TAFL: Skill levels and choice of language variety in the Arabic classroom”. *Teaching of Arabic as a foreign language: Issues and directions*, ed. Mahmoud Al-Batal, 223–231. Provo, Utah: American Association of Teachers of Arabic.

——. Forthcoming. “Teaching Arabic in the United States”. *A handbook for Arabic language teaching professionals in the 21st century*, ed. Elizabeth England, Kassem Wahba, and Zeinab Taha. New York: Lawrence Erlbaum.

—— and Abdelnour Zaiback. 1993. *Formal spoken Arabic: FAST course*. Washington, D.C.: Georgetown University Press.

Sallam, A.M. 1979. “Concordial relations within the noun phrase in educated spoken Arabic”. *Archivum Linguisticum* 10.20–56.

Sawaie, Muhammad. 1980. *Discourse reference and pronominalization in Arabic*. Ph.D. diss., University of Michigan.

Schmidt, Richard. 1974. *Sociolinguistic variation in spoken Egyptian Arabic: A re-examination of the concept of diglossia*. Ph.D. diss., Brown University.

Schultz, David. 1981. *Diglossia and variation in formal spoken Arabic in Egypt*. Ph.D. diss., University of Wisconsin, Madison.

Widdowson, H.G. 1988. “Language, context and culture in the classroom”. *ERIC/CLL News Bulletin* 12.6–7.

Wilmsen. David. 1995. *The word play’s the thing: Educated spoken Arabic in a theatrical community in Cairo*. Ph.D. diss., University of Michigan.

Karin C. Ryding
(Georgetown University)

Encyclopedia of Arabic Language and Linguistics

Volume II

Eg-Lan

General Editor
Kees Versteegh
(University of Nijmegen)

Associate Editors
Mushira Eid
(University of Utah)
Alaa Elgibali
(University of Maryland)
Manfred Woidich
(University of Amsterdam)
Andrzej Zaborski
(University of Cracow)

Advisory Board
Ramzi Baalbaki *(American University of Beirut)*
Elsaid Badawi *(American University of Cairo)*
Dominique Caubet *(INALCO, France)*
Clive Holes *(University of Oxford)*
Manfred Kropp *(Orient-Institut Der Deutschen Morgenländischen Gesellschaft)*
Jérôme Lentin *(Institut National des Langues et Civilisations Orientales)*
John McCarthy *(University of Massachussetts)*
Jamal Ouhalla *(University College Dublin)*
Jan Retsö *(Göteborg University)*
Sabah Safi *(King Abdulaziz University)*

Copy Editor
Carolyn Russ

ENCYCLOPEDIA OF ARABIC LANGUAGE AND LINGUISTICS

VOLUME II
Eg-Lan

General Editor
Kees Versteegh

Associate Editors
Mushira Eid
Alaa Elgibali
Manfred Woidich
Andrzej Zaborski

BRILL
Leiden – Boston
2007

Cover Illustration:
Details from the inscriptions framing the entrance portal tot the Imām Mosque (formerly Shāh Mosque) in Iṣfahān. Photo Courtesey: Sheila Blair and Jonathan Bloom.

Despite our efforts we have not been able to trace all rights holders to some copyrighted material. The publisher welcomes communications from copyrights holders, so that the appropriate acknowledgements can be made in future editions, and to settle other permission matters.

© Copyright 2007 by Koninklijke Brill NV, Leiden, The Netherlands
Koninklijke Brill NV incorporates the imprints Brill, Hotei Publishing, IDC Publishers, Martinus Nijhoff Publishers and VSP

All rights reserved. No part of this publication may be re-produced, translated, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without prior written permission from the publisher.

Authorization to photocopy items for internal or personal use is granted by Brill provided that the appropriate fees are paid directly to The Copyright Clearance Center, 222 Rosewood Drive, Suite 910, Danvers MA 01923, USA. Fees are subject to change.

PRINTED IN THE NETHERLANDS

ISBN-10: 90 04 14973 2 (Set)
ISBN-10: 90 04 14474 9 (Volume 2)

ISBN-13: 978 90 04 14973 1 (Set)
ISBN-13: 978 90 04 14474 3 (Volume 2)

This book is printed on acid-free paper.

Cover design: BEELDVORM, Pijnacker

List of Contributors

Al-Amin Abu-Manga, University of Khartoum, Sudan
Munira Al-Azraqi, King Faisal University, Hofuf, Saudi Arabia
Mohammad Alhawary, University of Oklahoma, U.S.A.
Enam Al-Wer, University of Essex, United Kingdom
Muhammed Hasan Amara, Bar-Ilan University, Ramat-Gan, Israel
Ahmad Atawneh, Hebron University, Palestine
Georgine Ayoub, Institut des Langues et Civilisations Orientales, Paris, France
Fatima Badry Zalami, American University of Sharjah, United Arab Emirates
Muhammad Hasan Bakalla, Emeritus Professor, King Saud University, Riyadh, Saudi Arabia
Sergio Baldi, Oriental Insitute, University of Naples, Italy
Wafaa ʿAbd al-Fahim Batran Wahba, Women's College, Department of English, Ain Shams University, Egypt
Abdelâli Bentahila, Abdelmalek Essaadi University, Morocco
Elizabeth M. Bergman, Georgetown University, Washington D.C., U.S.A.
Monique Bernards, University of Groningen, The Netherlands
Lidia Bettini, University of Firenze, Italy
Sheila Blair, Fine Arts Department, Boston College, U.S.A.
Václav Blažek, Masaryk University of Brno, Czech Republic
Stuart Campbell, University of Western Sydney, Australia
Ernie Capello, St. Johns University, Jamaica, N.Y., U.S.A.
Michael C. Carter, Sydney University, Australia
Dominique Caubet, INALCO, Paris, France
Dana Chahal, University of Balamand, Lebanon
Abdellah Chekayri, Al Akhawayn University, Ifrane, Morocco
Michael L. Chyet, Library of Congress, Washington D.C., U.S.A.
Guido Cifoletti, University of Udine, Italy
Andries W. Coetzee, University of Michigan, U.S.A.
Federico Corriente, University of Zaragoza, Spain
Sven-Olof Dahlgren, Göteborg University, Sweden
Janusz Danecki, University of Warsaw, Poland
Eirlys E. Davies, King Fahd School of Translation, Tangier, Morocco
Kinga Dévényi, Corvinus University, Budapest, Hungary
Joseph Dichy, University of Lumière-Lyon-2, France
James Dickins, University of Salford, United Kingdom
Madiha Doss, Cairo University, Egypt
Herbert Eisenstein, University of Vienna, Austria
Abderrahman El Aissati, Tilburg University, The Netherlands
Yamina El Kirat El Allame, University of Mohammed V, Rabat, Morocco
Nagwa Elzeiny, Faculty of Arts, Helwan University, Cairo, Egypt
Moha Ennaji, University of Fes, Morocco
Rob Ermers, 's-Hertogenbosch, The Netherlands
Mohssen Esseesy, George Washington University, U.S.A.
Mohammed Farghal, Kuwait University, Kuwait
Samira Farwaneh, University of Arizona,Tucson, U.S.A.
Ignacio Ferrando, University of Cádiz, Spain
Daniela Rodica Firanescu, Dalhousie University, Canada

Aharon Geva-Kleinberger, University of Haifa, Israel
Maik Gibson, Reading, United Kingdom
Alain Girod, University of Aix-en-Provence, France
Jacques Grand'Henry, Catholic University of Louvain, Belgium
Jean-Patrick Guillaume, University of Paris-3, CNRS UMR 7597, Paris, France
Dimitri Gutas, Yale University, U.S.A.
Atiqa Hachimi, University of Florida, U.S.A.
Peter Hallman, McGill University, Canada
Jaakko Hämeen-Anttila, University of Helsinki, Finland
Hassan Salam Bazzi Hamzé, University of Lyon-2, France
Benjamin Hary, Emory University, U.S.A.
Clive Holes, University of Oxford, United Kingdom
Bruce Ingham, The School of Oriental and African Studies, London, United Kingdom
Choukri Iraqi-Houssaini, Faculty of Letters, University of Fes, Fes Saiss, Morocco
Tamás Iványi, Eötvös Loránd University, Budapest, Hungary
Otto O. Jastrow, University of Erlangen, Germany
Éva Jeremiás, Eötvös Loránd University, Budapest, Hungary
Rudolf de Jong, Amsterdan Center for Language and Communication, University of Amsterdam, The Netherlands
Samvel Karabekyan, Yerevan State University, Armenia
Alan S. Kaye, California State University, U.S.A. and United Arab Emirates University, Al Ain, United Arab Emirates
Esam N. Khalil, University of Nijmegen, The Netherlands
Geoffrey Khan, University of Cambridge, United Kingdom
Reinhard Kiesler, Würzburg, Germany
Djamel Eddine Kouloughli, CNRS, France
Darlene LaCharité, Laval University, Canada
Pierre Larcher, University of the Provence, Aix-en-Provence, France
Evelyne Larguèche, CNRS, Laboratory of Social Anthropology, Paris, France
Mark LeTourneau, Weber State University, U.S.A.
Aryeh Levin, The Hebrew University of Jerusalem, Israel
Rochaya Machali, The University of New South Wales, Sydney, Australia
Yaron Matras, University of Manchester, United Kingdom
Catherine Miller, IREMAM, University of the Provence, Aix-en-Provence, France
Ahmed Moutaouakil, Faculty of Letters, Mohammed V University, Rabat, Morocco
Mustafa Mughazy, Western Michigan University, U.S.A.
Carole Paradis, Laval University, Canada
Tahera Qutbuddin, University of Chicago, U.S.A.
Christian Robin, CNRS, Paris, Laboratory for Ancient Semitic Languages, France
Judith Rosenhouse, Swantech Ltd., Haifa, Israel
Samuel Rosenthall, Rochester, MI, U.S.A.
Jan Jaap de Ruiter, Tilburg University, The Netherlands
Karin Ryding, Georgetown University, U.S.A.
Fatima Sadiqi, University of Fes, Morocco
Sabah M.Z. Safi, King Abdulaziz University, Saudi Arabia
Waheed Samy, University of Michigan, Ann Arbor, U.S.A.
Mohammed Sawaie, University of Virginia, Charlottesville, U.S.A.
Kassim Shaaban, American University of Beirut, Lebanon
Irfan Shahîd, Georgetown University, U.S.A.
Muhammad al-Sharkawi, American University in Cairo, Egypt
Marie-Claude Simeone-Senelle, LLACAN-CNRS, INALCO, University of Paris-7, Paris, France
Udo Simon, University of Heidelberg, Germany

Karel Steenbrink, Utrecht University, The Netherlands
Catherine Taine-Cheikh, LACITO-UMR 7017, CNRS, Paris, France
David Testen, Reston, VI, U.S.A.
Rolf Theil, University of Oslo, Norway
Sally G. Thomason, University of Michigan, U.S.A.
Sophia Vasalou, Gonville and Caius College, Cambridge, United Kingdom
Kees Versteegh, University of Nijmegen, The Netherlands
Kassem Wahba, Georgetown University, U.S.A.
Keith Walters, Portland State University, U.S.A.
Ineke Wellens, Ghent, Belgium
Stefan Weninger, Marburg, Germany
Andreas Wetter, Johannes Gutenberg-University of Mainz, Germany
David Wilmsen, The American University in Cairo, Egypt
Manfred Woidich, University of Amsterdam, The Netherlands
Andrzej Zaborski, University of Cracow, Poland
Marat Yavrumyan, Yerevan State University, Armenia
Tamar Zewi, University of Haifa, Israel

E

Egypt

1. GENERAL LINGUISTIC SITUATION

1.1 *Languages spoken in Egypt*

Apart from Arabic, some Afro-Asiatic languages are spoken in Egypt, viz. → Berber in the oasis of Siwa, Bedja (Bišāri) in the Eastern Desert to the south of the Aswān – Berenike line, as well as in Darāw and in the *išŠēx Harūn* quarter of Aswān. Reliable numbers of speakers of these languages are difficult to obtain. Bišāri speakers are estimated at about 15,000 (http://www.ethnologue.com/show_country.asp?name=SD) in Egypt and Sudan; and Sīwi (*tasiwīt*) at between 6,000 (Bliss 1998:37), 10,000 (Miller 1996:420), and 22,000 (Malem 2001).

As for non-Afro-Asiatic languages, → Nubian (Eastern Sudanic) exists in two main dialects, viz. Kanzi ~ Kunūzi (*matoki*) and Fadicca. With the erection of the High Dam at Aswān in 1964 and the inundation of their villages, most Nubians were transferred to New Nubia, close to Kom Ombo, but some of them have returned in the meantime to their old homelands on the shore of what is now Lake Nasser. The northernmost Kunūzi-speaking villages used to be in the First Cataract, i.e. the two villages on Elephantine Island (*aswan arti* in Nubian), on the island of Sehel, and in the village of Ġarb Aswān on the West Bank of Aswān. Owing to migration, Nubian can be heard in any of the larger cities of Egypt, although the younger generation are losing their command of the language (Miller 1996:416; Rouchdy 1991:19ff.). The number of speakers of Nubian in Egypt is estimated at about 50,000 (Rouchdy 1991:1) or at 170,000 (Miller 1996:415). → Coptic has been extinct as a living language since the Middle Ages (12/13th centuries). The details are still subject to discussion (see Rubenson 1996 and Björnesjö 1996), but it is still used as a liturgical language in the Coptic church. Turkish is still spoken at home by a small number of families in Cairo. As for Indo-European languages, small Italian-, Greek-, and Armenian-speaking communities exist in Cairo and Alexandria.

1.2 *Arabic dialects of Egypt*

The varieties of Arabic spoken in Egypt can be divided into urban, rural, and Bedouin dialects. The urban varieties, except those spoken in Cairo and Alexandria, have not been studied very much so far, and little can be said about them. In the northern part of the country they seem to be close to → Cairo Arabic, which serves as Standard Egyptian Colloquial. In Alexandria a Biḥēra dialect (WD 1) was originally spoken and can still be heard today from elderly fishermen in Anfūši (Behnstedt 1980). The modern variety of Alexandrinian is close to Standard Egyptian but with a strong admixture of the Biḥēra features mocked by speakers of Cairo Arabic (*nimši/nimšu, šarab, yōm larbaʿ, turṃāy* fem. etc.).

Bedouin – or *ʿArab*, as they are generally called – live in many places in Egypt. The fact that many villagers in Upper Egypt, the oases, and the Delta claim Bedouin descent does not mean that they speak a → Bedouin dialect today;

indeed, most of them are assimilated and do not differ significantly from their non-Bedouin neighbors. As for the different tribes, their present locations, and their history, Murray (1935) is still authoritative. There can be no doubt that sedentarizing Bedouin contributed substantially to the formation of the local dialects in several regions (western and eastern parts of the Delta, and Upper Egypt to the south of Asyūṭ), in particular the ED 1, UE 1, and UE 3 dialects (→ Bʿeri on the West Bank of Luxor). There are villages where true Bedouin dialects are spoken, i.e. with interdentals, syllable reshuffling, → *gahawa*-syndrome, plural feminine forms, etc., on the fringes of the Nile Valley (Maṭāʿna, Winkler 1936; ʿAzāyza at Gufṭ, Winkler 1936; Ǧihēna on the West Bank of Luxor, (Woidich, field data) and the Nile Delta, in the Fayyum, in the Dakhla Oasis (e.g. Rašāyda in Duhūs close to Mūṭ, and in Bir Biʿēri close to ilBašandi in the Dakhla Oasis), in the Western Desert and on the Mediterranean littoral (Awlād ʿAli) to the west of Alexandria, and in Wādi Naṭrūn. So far only the latter have to some extent been studied (Maṭar 1967; Behnstedt and Woidich 1987:111/1). As to the Eastern Desert, the Maʿāza and Xušmān tribes roam the desert roughly between the Cairo-Suez and Gufṭ-Guṣēr roads (Hobbs 1989), but their dialects have not been investigated to date. Farther to the south and on the littoral of the Red Sea, the ʿAbābda live in the Eastern Desert up to a line between Aswān and al-Ḥalāyib. Being former speakers of a Beja language, their present Arabic dialect is close to Sudanese Arabic (see de Jong 2002; Winkler 1936; Murray 1935).

In → Sinai, formerly a terra incognita, thorough research has been done in the north (de Jong 2000) and is in progress in the south (Nishio 1992; de Jong 2004).

The rural dialects of the Fallahin in the Nile Delta and the Ṣaʿayda in Middle Egypt and Upper Egypt, as well as the dialects of the oases in the Western Desert, were investigated in the 1970s and 1980s (see Khalafallah 1969; Doss 1981; Behnstedt and Woidich 1982, 1985–1999). A simplified list identifies seven major dialect groups, apart from the oases with subgroupings separated by bundles of isoglosses (see Maps 1 and 2). For more details see Maps 554–559 in Behnstedt and Woidich (1985–1999) and Woidich (1996).

Lower Egypt
CD = Central Delta (including Cairo)
NED = Northeast Delta: NED 1 dialects (Daqahliyya center) – NED 2 dialects (Daqahliyya East, ilManzala region)
WD = West Delta: WD 1 (Kafr išŠēx, Ġarbiyya West, Biḥēra) – WD 2 (irRašīd [Rosetta] Balṭīm, Burullus) – WD 3 (Minufiyya Southwest) – WD 4 (Biḥēra, Minufiyya Northwest, Gīza North)
ED = East Delta: ED 1 (east and center of Šarqiyya) – ED 2 (Daqahliyya East, Šarqiyya North) – ED 3 (Šarqiyya Southwest, Daqahliyya South)

Middle Egypt
NME = Northern Middle Egypt: NME 1 (Gīza South, Fayyum, Bani Swēf) – NME 2 (ilFašn to ilMinya)
SME = Southern Middle Egypt (ilMinya to Asyūṭ)

Upper Egypt
UE = UE 1 (Suhāg to Luxor) – UE 2 (Najʿ Ḥammādi to Gina) – UE 3 (West Bank Luxor to Isna) – UE 4 (Isna to Aswān)

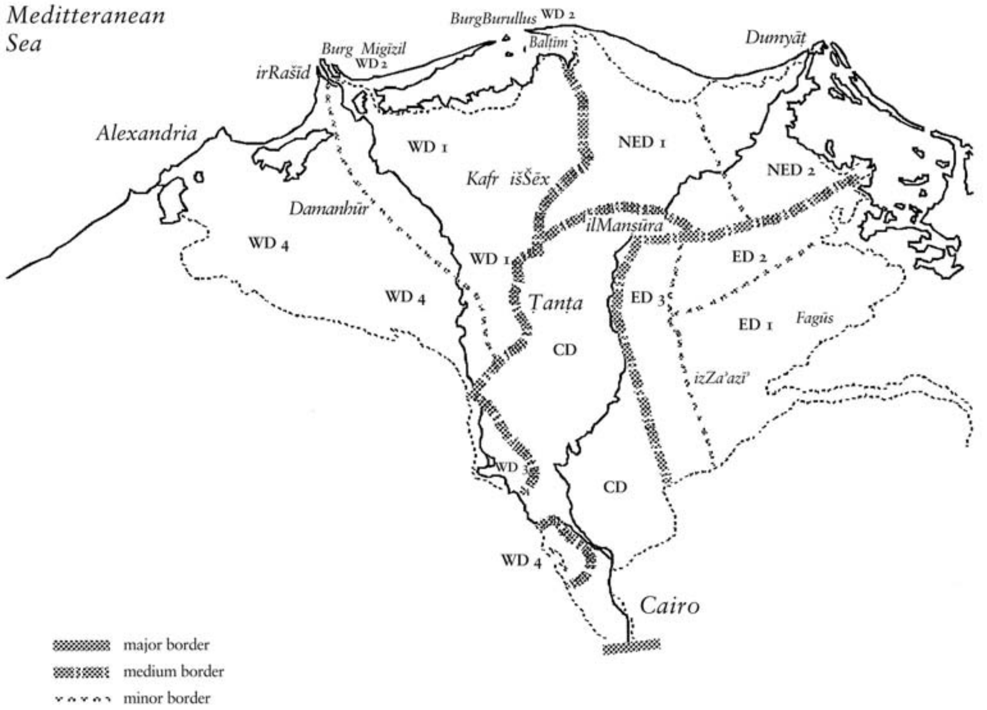
Oases
OAS = BAḤ ilBaḥariyya: BAḤ-E East (Mandīša) – BAḤ-C Center (ilBawīṭi) – BAḤ-W West (ilGaṣir)
FAR alFarāfira
DAX adDāxila: DAX-W West (alMušiyya) – DAX NW Northwest (anʾĀṣər) – DAX-C Center (Mūṭ, Ismint) – DAX-E East (Balāṭ)
XAR ilXārja: XAR-N North (City) – XAR-M Middle (Bulāg) – XAR-S South (Barīs, Dūš)

1.2.1 Lower Egypt – Nile Delta

In the Nile Delta a Central (CD) group is surrounded by Eastern (ED), Western (WD), and Northeastern (NED) groups, each of which can be divided into subgroups (see Map 1).

Cairo Arabic, the Standard Egyptian dialect, belongs to this CD group and differs but little from the rural varieties. Particularly Cairene features are the lack of pausal forms, and the suffix *-it* for the 3rd person singular feminine perfect of III*y* verbs, which is *-at* elsewhere:

Map 1. Dialects of the Nile Delta



ramit ‘she threw’, *ṣallit* ‘she prayed’ vs. *ramat, ṣallat* in rural areas (see Woidich 1997). On the fringes of the Delta, both in the West (WD) and in the East (ED), we encounter a palatalized pronunciation of **g* varying between [dʒ], or /ǵ/ [ɟ], to /ž/ [ʒ], and a voiced palatal stop [g] for **q*, whereas CD has a glottal stop [ʔ] here. Another feature, common in nearly all rural dialects, is the pausal *ʾimāla* for *-a#* as *-e*(*h*)*#* ~ *-i*(*h*)*#*, which Cairo lost at the end of the 19th century (Blanc 1973–74). The dialect groups can be described in more detail as follows (compare Maps 1 and 2).

WD dialects in general display certain phonological features such as pausal glottalization in final long syllables, as in *kitiʾr* [kɪ'tiːʔṛ]; strong pausal *a-ʾimāla* in a non-emphatic context: *kalbe* ~ *kalbi#* ‘bitch’; and strong secondary emphasis under Bedouin influence, as in *muḥṛāt* ‘plough’, *faṛxa* ‘chicken’, *dukkān* ‘shop’. **g* corresponds here to /j/ [dʒ], or /ǵ/ [ɟ], to /ž/ [ʒ] in the westernmost part of the region, apparently imported by North African Bedouin. **q* is /g/, except in Alexandria and irRašīd (Rosetta), which follow the CD dialects with /ʾ/. Short /i/ in open unstressed syllables is not elided: *misikit* ‘she took’, *wisixa* ‘dirty [fem.]’. As to morphology, WD dialects (except WD 3) prefer the *a*-perfect to the *i*-perfect in both strong and III*y* verbs, e.g. *rakab* ‘he mounted’, *šarab* ‘he drank’, *faham* ‘he understood’, *nasa* ‘he forgot’, *maša* ‘he went’, etc. (cf. CD *rikib, širib, fihim, nisi, miši*); the *in*-prefix for the reflexive passive of Form I is more frequent than *it-*, e.g. *inḍarab* ‘he was hit’; the imperfect keeps the vowels of the perfect: *yinḍarab* (CD *yinḍirib*). The most striking feature of WD 4 is certainly the Maghrebi form of the 1st persons of the imperfect *niktib – niktibu*. In the adjacent WD 1 and WD 2 areas (irRašīd only, not so the more eastern towns of Balṭīm and Burullus), the paradigm is *aktib – niktibu*, which could be seen as the first step of the paradigmatic leveling which finally led to the Maghrebi paradigm. Seen from this angle, WD 1 would be a transitional area eventually turning into the CD dialects. A more plausible explanation considers this paradigm as the result of dialect contact. Bedouin from North Africa (*niktib – niktibu*) sedentarized in the western part of the Delta and mixed with the original population (*aktib – niktib*). In this mixed situation *niktib* could be used for both the 1st person singular and the 1st person plural and became ambiguous, thus causing a homonymic clash. Hence *niktib* was eliminated, whereas the two unambiguous forms *aktib* and *niktibu* remained in use and formed the new paradigm (see Behnstedt and Woidich 2005:162). The same developments occur in Upper Egypt in similar situations (see below). In WD 1 the base of Form II exhibits three allomorphs based on a phonologically conditioned distribution of /a/ and /i/, not just two as in CD (Cairo): cf. *yiʿallim* 'he teaches', *yikassaṛ* 'he smashes', but with a third type *yikillim* 'he speaks to' with an /i/ in the penultima, provided there are no back consonants (CD *yikallim*). III*y* verbs still have an *a*-type and an *i*-type as base forms, e.g. *rama* 'he threw' vs. *ġila* 'it became expensive', but the latter is conjugated in the same way as the *a*-type: *ġila, ġilat, ġilu.* Very common in WD dialects is *ilbāriḥ* 'yesterday' instead of CD *imbāriḥ.*

At the northern periphery of the Delta the isolated WD 2 dialects of irRašīd-Burg Miġīzil and Balṭīm-Burullus differ markedly from the adjacent WD 1 dialects not only in preserving /q/ [q] and the diphthongs /ay/ and /aw/ in pausal closed syllables (reduced to /i/ and /u/ respectively when suffixed: *bīti* 'my house', *bitna* 'our house', *bitayn* 'two houses'), but also in their unusual stress pattern, which gives prominence to the penultima in -CvCvCv: *baqára* 'cow', *ḍarábu* 'they hit'.

NED dialects occupy roughly the area between Balṭīm-ilManṣūra, ilMaṭariyya, and Dumyāṭ (Daqahliyya province) and can be divided into two subgroups, viz. NED 1 and NED 2, the latter covering approximately the ilManzala region and the northernmost part of the shores of the eastern branch of the Nile between Širbīn and Dumyāṭ. The two groups differ in their stress patterns: NED 1 stresses *madrása* 'school', *yixbízu* 'they bake bread' on the penultima, following the WD and CD dialects, whereas NED 2 stresses the antepenultima, e.g. *mádrasa, yíxbizu,* a pattern common in the Nile Valley, neighboring Sinai and Palestine, and partly in ED 1. As to phonology, the NED dialects differ from WD and go along with CD: **q* is /ʾ/ and **g* is /g/ [g]; together with CD the two groups form the Cairo-Dumyāṭ corridor described and interpreted in → dialect geography (Maps 2a, 2b). Other prominent features of NED include the following: often /i/ where other dialects have /u/, as in *kint* 'I was', *kill* 'all', *hiwwa* 'he'; the 3rd person plural *-um* perfect, as in *waṣalum* 'they arrived'; all III*y* verbs follow the KiKa pattern, i.e. an *i*-base but an *a*-conjugation: *rima, rimat, rimu, rimēt* 'to throw' and *nisa, nisat, nisu, nisēt* 'to forget'; contraction of *biyimsik* to *bimsik* 'he takes', *biyudrus* to *budrus* 'he studies' (common in ED as well); mostly *iga/yigi* ~ *yāgi* 'to come' instead of *gih; ilwaʾti* 'now' instead of *dilwaʾti.*

The ED dialects occupy a territory covering roughly the province of aš-Šarqiyya and the adjacent parts of ad-Daqahliyya province and can be divided into three subgroups, with ED 1 as the core area and ED 2 and ED 3 dialects at the periphery, i.e. ED 2 close to NED in the northern part and ED 3 bordering on CD in the western part of the area. ED dialects, like WD, have /g/ for **q* and an affricated /j/ [dʒ] ~ [ɟ] due to Bedouin influx, in contrast to the bordering CD and NED dialects. ED is further distinguished by a number of outstanding features not present, or present to a lesser degree, in other Delta dialects. These are most prominent in ED 1: strong secondary emphasis, again under Bedouin influence, in many lexical items (cf. *miḥṛāt* 'plough', *ṛama* 'he threw', *faṛxa* 'chicken', *duḳḳān* 'shop', *kaṃān* 'too', *igṃāš* 'cloth', *xāḷ* 'uncle'); /a/ replaced by /i/ in pretonic closed syllables, as in *fillāḥ* 'farmer', resulting in similar variations in verbal paradigms as in e.g. Lebanese dialects (cf. *lamm* 'he took', but *limmēt* 'I took'); /a/ replaced by /i/ in *aCCaC adjectives of color and deficiency (cf. *iḥmaṛ* 'red', *ismaṛ* 'brown', *iʿma* 'blind'); on the other hand, /a/ is preserved in pretonic open syllables (cf. *katīr* 'much', *jadīd* 'new'). The syllable structure in ED is close to that of Levantine dialects and Upper Egyptian (UE) dialects, with elision of /i/ after -CC- and subsequent insertion of /i/ after the second consonant from the right, i.e. (C)CCC > (C)CiCC, as in *yíxibzu* 'they bake'. Initial and final clusters are resolved in the same way (cf. for the former *ihnạ̄k* 'there', *islāḥ* 'weapon', *itjīb* 'she brings'), and stress can be placed on the inserted vowel, as in *íšta* 'winter', *ídra* 'millet', *írkab* 'knees' (see Woidich 1982). For final clusters, cf. *ṛumuḥ* 'tail-pole', *galib* 'heart', *ʾariḍ* 'ground', where – in contrast to UE dialects – /i/ is inserted even before liquids and nasals. Long vowels are preserved before -CC-, as in *kātba* 'having written [fem.]' or *bētna* 'our house', just as in UE dialects. Pausal *ʾimāla* of *-#a* to *-e#* is as common as in most rural dialects, e.g. *kalbe* 'bitch', *ḥāje* 'something'. Stress assignment shows a split between nouns and verbs: whereas nouns follow the general Delta pattern *madrása,* verbs get the stress on the antepenultima, as in *yíxibzu* 'they bake [bread]'. A morphological peculiarity of ED 1 and ED 3 is the distribution of the allomorphs *-at* and *-it* of the 3rd person singular feminine perfect suffix: *-at* is used with all base forms containing /a/ or /ā/, otherwise *-it* (cf. *ḍarabat* 'she hit', *gālat* 'she said', but *širbit* 'she drank', *nisyit* 'she forgot'). Form VII of the verb has the prefix *in-* as in *inmasak;* it is more common in ED 2 and ED 3, but these still use *in-* much more than e.g. Cairo Arabic; in the imperfect of Forms VII and VIII only the vowel of the ultima is /i/, whereas the one of the penultima remains /a/ (cf. *yinmasik* in ED vs. *yinmisik* ~ *yitmisik* in CD and NED). As in Levantine Arabic the root **g-y-ʾ* 'to come' is extended in the core ED dialect by a prothetic glottal stop, producing *ʾīja, ʾījat, ʾīju* (3rd persons only), but here with a long /ī/ in contrast to NED 2 *ʾiga* with a short one.

ED 3 displays fronting of **k* to /č/ [tʃ], as in *čal* 'he ate', a remarkable feature in this Egyptian context; but it remains /k/ in a *u*-coloring environment (cf. *kul!* 'eat!', *yākul* 'he eats'); just as in NME 1 and ilXārja-South, Form II of the verb contains /a/ in both the perfect and imperfect, as in *čallam, yičallam* 'to speak to', thus showing neither phonological nor morphological variation.

1.2.2 Upper Egypt – Nile Valley

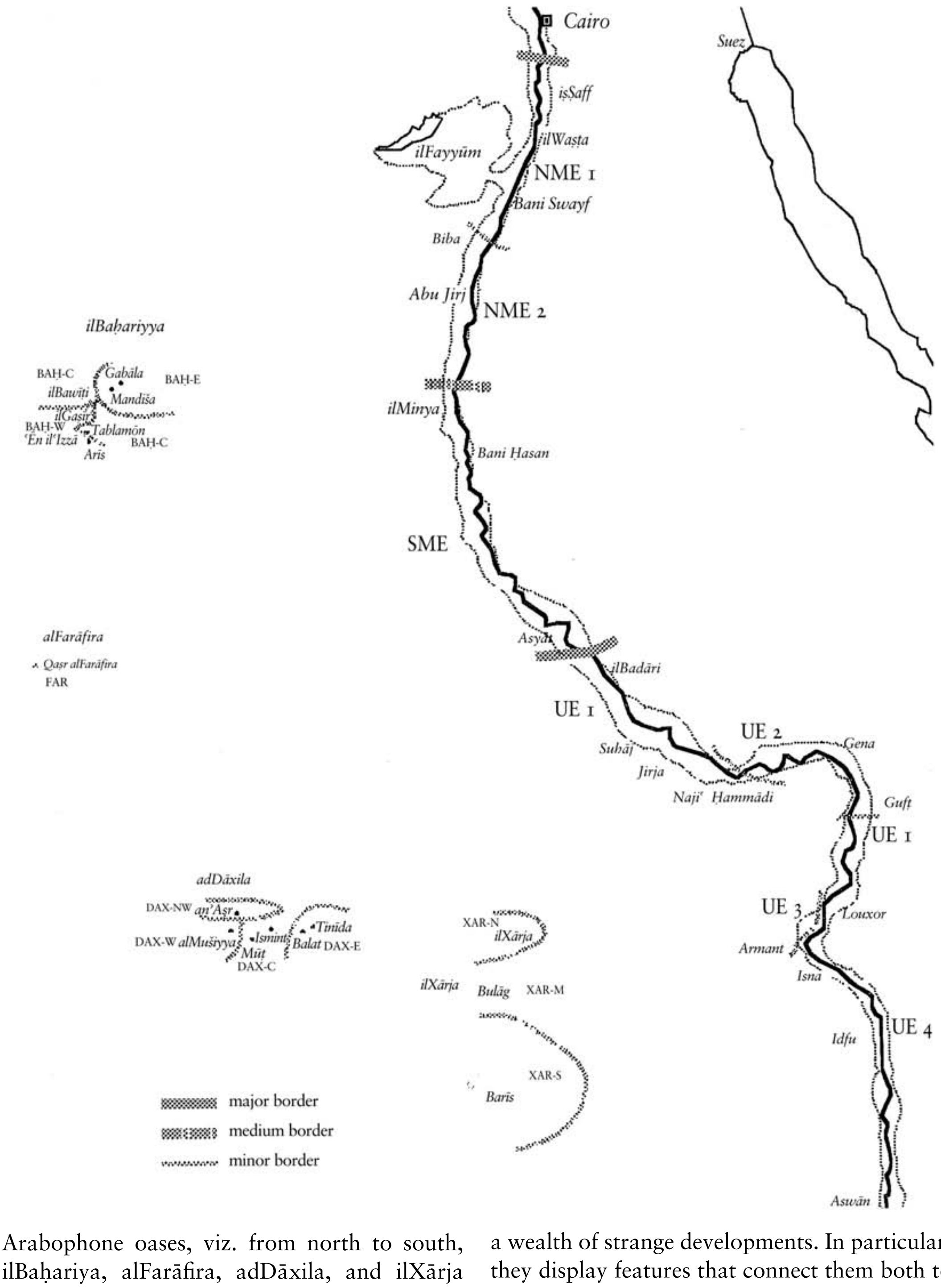
In the Nile Valley, Northern Middle Egyptian dialects start immediately to the south of Cairo, in the southern Gizeh and Bani Swēf provinces, with a transitional area whose dialects are still close to WD 1 phonologically (pausal forms, pausal glottalization, lack of elision of /i/ in open unstressed syllables, as in *misikit*). On the other hand they are already distinguished by the → *bukaṛa*-syndrome, stress on the first syllable in *mádrasa,* preserved diphthongs, and /a/ in the ultima of the base form of Form II throughout, which might be seen as a link to the ED 3 dialects. From here to Aswān *gīt* ~ *jīt* 'I came' is used instead of *gēt,* and long vowels before -CC remain long (*ʿārfa* 'she knows'). NME 1 proper starts in the ilWasṭa area, with its characteristic strong devoicing of the final syllable and the equally strong *bukaṛa*-syndrome, which is a remarkable feature of Middle Egypt and the oases. The Fayyūm deviates from NME 1 of the Nile Valley by such forms as *yāgi* instead of *yīgi.*

Approximately 20 km to the south of the city of Bani Swēf, NME 1 changes into NME 2, which continues southward to the city of ilMinya. In NME 2 we find **q* as /g/ [g] and **g* as /j/ [dʒ]; *i*-elision occurs in *miskit* but not in adjectives such as *wiḥiša.* In the northern part of the region between Biba and Abu Jirj, Form II has three allomorphs, just as in WD, whereas the southern part again follows the one-allomorph system (*yikallam*) in the same way as NME 1, ED 3, and Xarja South. From ilMinya southward to Asyūṭ there follows another group, SME, still characterized by a northern-type syllable structure with preservation of short unstressed /i/ in open syllables after -CC-, as in *yídrisu* 'they thresh'. It is here that the distribution of /a/ and /i/ in the perfect and imperfect of verb Forms II and III, based on morphological conditioning, starts (*baḥḥaṛ/yibaḥḥir* 'to go to the north', *sāfar/yisāfir* 'to travel'), in contrast to the phonological conditioning in the north. South of Asyūṭ, approximately at the rural towns of Abu Tīj and ilBadāri, we enter UE proper, with a glottalized pronunciation of /ṭ/ [ŧˀ] and the elision of short unstressed /i/ in open syllables after -CC- in parallel with the insertion of the intrusive vowel after the second consonant from the right: *yídirsu* 'they thresh'. This UE 1 type prevails in the Nile Valley until far beyond Luxor, where UE 4, a more Sudanese type of dialect, starts, and only with two major areas interspersed, viz. UE 2 between Najʿ Ḥammādi and Gūṣ mainly on the east bank, and UE 3 (→ Bʿeri) on the west bank between Gurna and Esna. UE 2 is close to SME in its syllable structure (*yídrisu*), but otherwise it shares most features with UE 1. UE 3 has a strong Bedouin admixture, with *gahawa*-syndrome and plural feminine forms. UE 4 deviates with its initial *a-* where all other dialects of the Nile Valley have *i-*, as in the article *al-*, *alli, abn, amm anḍarab,* etc. Dialects of the types *aktib-niktib, aktib-nikitbu,* and *niktib-nikitbu* are to be found here side by side. For more details see Woidich (1995).

1.2.3 Oases

In the Western Desert, in addition to the primarily Berberophone Siwa, there are four

Map 2. Dialects of the Nile Valley


Cairo
Suez
iṣṢaff
ilWasṭa
ilFayyūm
NME 1
Bani Swayf
Biba
Abu Jirj
NME 2
ilBaḥariyya
BAḤ-C
Gabāla
BAḤ-E
ilBawīti
Mandīša
ilGasir
BAḤ-W
Tablamōn
ʿĒn ilʿIzzā
BAḤ-C
Aris
ilMinya
Bani Ḥasan
SME
alFarāfira
Qaṣr alFarāfira
FAR
Asyūṭ
ilBadāri
UE 1
UE 2
Suhāj
Jirja
Najiʿ Ḥammādi
Gena
Guft
UE 1
UE 3
Louxor
Armant
Isnā
UE 4
Idfu
Aswān
adDāxila
DAX-NW
anʾAṣr
DAX-W
alMušiyya
Ismint
Tinīda
Balat
DAX-E
Mūt
DAX-C
XAR-N
ilXārja
ilXārja
Bulāg
XAR-M
XAR-S
Barīs
major border
medium border
minor border

Arabophone oases, viz. from north to south, ilBaḥariya, alFarāfira, adDāxila, and ilXārja (see Map 2), whose dialects have been investigated in the last 30 years (see Behnstedt and Woidich 1982; Woidich 1998). Long isolated, they only became accessible in the 1970s. Owing to this isolation and the continuous influx of small groups from outside, they offer a wealth of strange developments. In particular, they display features that connect them both to Middle Egypt and to Western (Libyan) Arabic as possible substrata or adstrata. For diverging views on this subject see Woidich (1993), Behnstedt (1998): the former relates them more to Middle Egypt on structural and phonological evidence, the latter to a North African substrate – at least the two northern ones – on morphological and lexical grounds. No single major discriminant shared by all four dialects, which would separate them as a single entity from the dialects of the Nile Delta or the Nile Valley, could be detected to date. They differ markedly from one another, and even within an oasis the individual villages display distinctive features and can be grouped together into subgroups (see above). Let us now look at this in more detail.

A voiceless reflex of **q*, be it [q] or [ʔ], is attested in all the oases except BAḤ, which has only the voiced [g]; to the south in FAR we encounter [q], corresponding to [ʔ] in DAX-W and DAX-C; in DAX-E [q] occurs again, but only in about a third of the possible roots; the others contain /g/ [g]. This produces a considerable number of minimal pairs between /q/ and /g/, such as *gabba* 'neck [of a *gallabiyya*]' and *qabba* 'block of wood'. XAR has even fewer roots with /q/ (phonetically glottalized [q^ʔ]), the majority displaying [g]. In all the oases, lexical items, such as *gāṛa* 'hill, mountain' and *zagal* (BAḤ, XAR), *žigál* (FAR), *zigā́l* (DAX) 'he threw away', can be found that always have [g], never [q]. This indicates the highly mixed character of the dialects of DAX-E and XAR.

In the Nile Delta and Nile Valley a voiced reflex of **q*, i.e. [g] as a rule implies a palatalized pronunciation of **g* as [dʒ] etc. (see above). This is not true of the oases, where we find a voiced [ʒ] (= **g*) in addition to a voiceless [q] (= **q*) in FAR, as well as [g], [q] (= **q*), and [dʒ] (= **g*) in DAX-E, and [g], [q^ʔ] (= **q*), and [ɟ] (= **g*) in XAR. Only BAḤ conforms to this rule, with [g], [ʒ] in BAḤ-E and BAḤ-W, and [g], [ʑ] in BAḤ-C respectively. In DAX-W, -NW, and -C **q* is represented by /ʾ/ and **g* by /g/. This again suggests a high degree of mixing and the presence of different dialectal layers in these dialects. The presence of /q/ and the stress on the penultima, just as in the WD 2 dialects at the periphery of the Nile Delta, suggest a link to pre-Hilalian Arabic (see Behnstedt 1998).

In the two northern oases of BAḤ and FAR, the sibilants merged to either a postalveolar [ʃ], [ʒ], as in FAR, resulting in e.g. [ʃaːb] 'he left', [ʒajt^s] 'oil', or to an alveolo-palatal [ɕ], [ʑ] in BAḤ-C, giving e.g. [ɕaːb] 'he left', [ʑeːt] 'oil'. This merger is certainly a Maghrebi feature not found elsewhere in Egypt, like the strong aspiration of /t/ [t^h] ~ [t^s] in FAR. Stress assignment, too, follows Maghrebi rules in the BAḤ-C and -W, FAR, and DAX dialects, since they stress the last syllable of a word unless it contains the feminine suffixes *-a* or *-it* (perfect) or the pronominal suffixes of the 3rd person singular masculine *-u* (*-a, -ih*), or the 2nd person singular masculine suffixes *-ak* or *-ik* (cf. BAḤ-W *dikáṛ* 'male' but *wúkkil* 'he fed'), FAR *qamáṛ* 'moon', *ibyáḍ* 'white', DAX-E *ḥalā́q* 'earrings', *gabbā́l* 'he went southward').

Like NME 1, FAR and DAX-W, -NW, and C preserve the diphthongs /aw/ and /ay/. Whereas they are stable in FAR (*bāyt*^{s}, *bayt*^{s}*íy, bayt*^{s}*ihíy* 'house, my house, her house'), they change to /i/ or /e/ in DAX-W depending on the stress and syllable structure (*bāyt, bitḗ, bētihī́*). Another feature common to both NME and the oases' dialects is the → *bukaṛa*-syndrome (cf. BAḤ *itaṛama* 'it was thrown', FAR *nuqaṛa* 'hole', DAX-E *ʾibaṛa* 'needle', XAR-S *yaḥarat* 'he plows'). The syllable structure of all the oases' dialects is also like that of NME, since there is no elision of /i/ in unstressed open syllables after -vC-, let alone after -vCC- (BAḤ and XAR *šíribit,* FAR *širíbit,* DAX *širébit* ~ *širḗbit*). FAR and DAX-West, -NW, and -C do not elide vowels at all, contrary to the common practice in modern dialects, and /i/ is preserved even after -v̄C (cf. DAX *ʿārifa* 'she knows', FAR *qāfila* 'caravan'). The question remains open whether this represents conservation of the older situation or innovation (reintroduction of the base form by paradigmatic leveling). Only BAḤ and FAR exhibit the leveled Maghrebi-type paradigm of the imperfect with *niktib-niktibu;* DAX and XAR follow the general eastern pattern *aktib-niktib*. In contrast to common practice in Egypt, the feminine active participle changes *-a* to *-it* in BAḤ, FAR, DAX when receiving a suffix, e.g. BAḤ *māskitu* 'she has taken it' ~ FAR *ʿāwižit*^{s}*ih* 'she wants him' ~ DAX-W, DAX-C *ʿārifitih* 'she knows him', whereas XAR lengthens this *-a* as is usual in the Nile Valley, e.g. *maskāh* 'she has taken it'. The active participle of III*y* verbs lengthens its final vowel when the feminine suffix is added, resulting in forms such as *mašiyya* 'she is going' in BAḤ, FAR, and DAX; XAR, like the other Egyptian dialects, has *māšya*.

Dialectometrical analyses as presented in Behnstedt and Woidich (2005:108–118) show that BAḤ-E and XAR stand closer to the dialects of the Nile Valley, and FAR and DAX form a separate group characterized by additional features such as a peculiar penchant for nasality (segmental spread in FAR, as in *muwayya* [mũwãyyã] 'water', *šākin* [ʃɛ̃ːkĩ] 'living', *l > /n/ in DAX-NW, pausal nasalization [ĩː] for *-ā* in DAX-C], an article *al-*, and a *u*-vowel in the imperfect of the strong verb and the geminate verbs (cf. DAX *yiktōb* 'he writes' and *yilōmm* 'he collects'). As may have become clear from the evidence adduced above, FAR stands out as the most deviant dialect of the four.

There can be no doubt that Western and Eastern Arabic meet in the oases and that their dialects display features from both sides, including interactions between them that produce rather strange developments (see Woidich 1995–1997, 1997a). In order to understand the situation better we need to know the dialects of the Libyan oases on the other side of the Great Sand Sea in more detail than is currently the case.

2. Sociolinguistics

The discipline of sociolinguistics usually investigates language → variation and change in relation to various socially recognized categories in a speech community such as class, age, gender, and confession. For its part, Arabic sociolinguistics does not usually limit itself, or even apply itself to any large degree, to those particular variables. Some notable exceptions include Blanc (1964), who studied dialect differences among the three main confessional groupings in Baghdad; or, in an Egyptian context, Royal (1985), Wahba (1993), and Haeri (1991), who examined phonological features of men's and women's speech in Cairo; Parkinson (1991), whose study of terms of address in Cairo acknowledged social class as an important variable; Peterson (2000), who observed the jargon of youth in Cairo; or Wilmsen (1999), who studied the variation in and interaction between a syntactic feature of rural and urban dialects. Usually, most attention is paid to the place and function of the spoken vernaculars on one hand and their relation to and interaction with formal written and declaimed Arabic on the other. Studies addressing this issue often assume certain class distinctions among the grades or levels between the ornate, recherché styles of declamation (and of course writing) in the upper registers of Arabic and the purely conversational vernaculars. The assumption is that only the educated (assumed to be the upper classes) will have any great proficiency with the written variety, and only they will be able to declaim extemporaneously in it, while the lower classes will remain limited in their verbal expression to the baser registers of the vernacular.

In reality, the assumption of greater proficiency among the educated upper classes is not entirely accurate. In Egypt it is usually the educated among the lower classes whose facility in written and oral expression in this idealized eloquent variety is more polished. Among the members of the upper classes, the claim of low productive proficiency in written or declaimed Arabic is itself a badge of refinement and breeding. The reasons for this are that the lower classes obtain their education from the state-sponsored schools and universities, where Arabic writing and declamation are integral parts of the curriculum, while the upper classes send their children to private ('language') schools, where European languages predominate. Very often, these same students (young men more often than young women) will gain some of their education – especially its postsecondary stages – abroad.

In an earlier era, the emphasis in these language schools was on French language and education, as it was the short-lived Napoleonic incursion into Egypt in 1798 that first opened Egyptian eyes to the modern West. Despite the brevity of that encounter, for several generations afterward, Egyptian elites would gain their education at French missionary schools at home and their higher education at universities in France.

By the middle of the 20th century, however, and increasingly ever since, English has become the preferred foreign language for everyone, rich and poor alike. French may still be heard on the lips of the remnants of the aristocracy, now declining into their twilight years, and their attendant upper classes, especially in the social venues of upscale neighborhoods of Cairo. A few French-language schools continue to attract students, but most of them, and many others besides, now feature English either as the main language of instruction or as the first foreign language.

Familiarity with foreign languages lends to speakers a certain cachet, not necessarily a class distinction, but surely an air or at least an affectation of sophistication. Almost everyone who goes to school manages to gain some facility with English, if only rudimentary, and many adults continue to pursue language study well beyond their school days. Despite its declining number of speakers, French remains the second most popular foreign language, still retaining some of its older aristocratic associations. For its part, Spanish carries something of an elite air, both for the relatively few people who study it and for its ancient associations with a golden era of Arab civilization; it is apparently gaining in popularity as a third foreign language for adult learners as well. Other European languages, like Italian, German, and Russian, are pursued by those wishing to engage in the hospitality trade, hoping to find work serving the tourists speaking one of those languages who flock in large numbers to the winter resorts on the Red Sea, coming by direct flight to Egyptian beach destinations from their respective countries. For that matter, tourist touts can muster communication in a great many foreign tongues. A smattering of Japanese and increasingly Chinese are sometimes picked up by those engaged in the mule work of importing goods from the East.

It is generally the degree of fluency in a foreign language as much as the actual language spoken that carries with it class connotations. People who are very good with English will give the impression, often a true one, that they have spent large amounts of time abroad and are therefore able to afford such travel. Those who are good with French will more likely have grown up in Egypt in a partially Francophone environment and are, therefore, either from the remnants of the aristocracy – now perhaps of restricted means – or from the Christian upper class, either way only occasionally having spent years abroad. Those fluent in other foreign languages are as likely to be members of the lower middle classes, from which service employees are drawn, as they are to be members of the elite.

Current circumstances aside, Egyptian Arabic has been in contact with foreign languages for centuries, even a millennium or longer, which have left their mark on the language in the form of loanwords. Earlier influences would have come first from Coptic, later Turkish and Persian, and much later French, then Italian and Greek, and finally English, which continues to exert an influence as new concepts, some of them quite trivial, such as jargon from the mass media, enter the language.

New technologies have brought with them entire glossaries, not all of the words of which are technical terms. Non-Arabic names for automobile parts, for instance, are usually French (e.g. *dibriyāž* ‘clutch pedal’, *diriksyōn* ‘steering wheel’, *tablō* ‘dashboard’, *bužehāt* ‘spark plugs’). The more familiar concepts, like windshield (*ʾizāz* ‘glass’) and wheels (*ʿagalāt*) are native Arabic. Those for the computer, a later technological introduction, are English. In this case, it is the operations that are more often English borrowings than the components (although the instrument itself is referred to in speech and often in writing with the English loan *kumbiyūtar*). Otherwise, for example, *sayyif* means ‘to save’, *fayyil* ‘to file’, and *hannig* ‘to hang’, while the more familiar concept of a computer screen is simply labeled *šāša* and the keyboard is the calque *lōḥit ilmafatīḥ*. With an even later technological innovation, the mobile telephone, has come a new set of borrowings. The hand unit itself is referred to in speech as a *mubayl*, or sometimes in the lower registers *mubayyin*. When referred to in writing, the calque *maḥmūl* is more often employed. The process of talking to someone on the telephone is described with the native Arabic *kallim* ‘to speak’, but the operation of sending a text message borrows the English concept to yield *massij*. Similarly, a procedure for avoiding the cost of a call while at the same time alerting friends to one’s availability is to send a missed call: *yibʿatlu mist* ‘he sends him a missed [call]’, or sometimes *yimissi̊ ʿalē* ‘he misses at him’.

The means by which these terms entered the language are instructive. The earliest mobile telephones available in Egypt were incapable of displaying Arabic writing on their screens; as such, the terminology was entirely English. What is more, the high cost of the early units meant that they were accessible only to the affluent, who were generally proficient in reading the English that appeared on the screens. By the time the telephones became affordable to the common people and the technology advanced sufficiently to permit Arabic displays, the terminology was largely fixed in the language. This same process no doubt occurred with other technical terminologies, from automobiles to computers.

This use of direct borrowings tends to appear more in speech than in writing and declamation, wherein loan translations – or calques, whereby a new concept is explained periphrastically in native lexemes (for example, *ḥāsib ʾāli* 'computing machine' for computer) – are preferred. Often this preference is more observed in the breach, with spoken borrowings finding their way into writing, regardless of the prescriptions or preferences of language purists. What appears, then, is a set of parallel lexemes, one used more in speaking and one used more often in writing. This duality of reference in speech on one hand and writing or declamation on the other with reference to some foreign concepts is a further reflection of the acknowledged dichotomy between spoken Arabic vernacular forms and formal written or declaimed forms.

It is this linguistic duality that informs most sociolinguistic investigations into Arabic. Called → 'diglossia' (after Ferguson 1959), it is a characteristic of the language in all parts of the arabophone world. It is marked by a functional distribution of the two varieties of the language, often referred to as high (or simply H), for the written or declaimed variety, and low (or L), for the spoken vernacular. What this implies is that there are domains – or functions – in which one variety or the other is considered appropriate or even obligatory. The H variety is expected in formal situations involving public speaking. In venues such as the Parliament, courtrooms, churches and mosques, official announcements, newscasts, college lectures, etc., it is the H variety that is considered appropriate.

In actual usage in Egypt, however, there is a great deal of overlap and interplay between the two codes at all levels, and accordingly H and L are generally not really mutually exclusive categories. The division of labor between the two varieties is more of an ideal than a reality, reflecting speakers' attitudes toward their language and not their actual behavior with it. True, in some of the venues mentioned above (Parliament, newscasts), the vernacular is never – or hardly ever – heard. In all other situations, however, the vernacular, or L, predominates, even impinging upon the language used in formal situations that might otherwise be considered to be the exclusive domain of H. For instance, although news broadcasts are always delivered in H, interviews might be conducted in a mixture of both, with the announcers hewing more closely to the strictures of declamation in H and their interlocutors holding forth in an amalgam of H and L or remaining completely in L. In man-on-the-street interviews and chats with celebrities, even the announcers will speak in the L variety.

What is more, in teaching at all levels, from the traditional scriptural schools for children (*kuttāb*) to university classrooms, texts are, of course, written and read out in H, but almost all explication and discussion of them is conducted in L. In sermons too the language used is not strictly and exclusively the High variety. Instead preachers shift between H and L for stylistic purposes. For some highly decorous occasions, sermons are written out beforehand and read or recited from memory in the pulpit, in which case the High variety *is,* in fact, employed. The more usual addresses, such as the Friday or Sunday sermons, are not written but delivered extemporaneously, granting preachers freedom to style shift in their oratory. In some types of oral religious discourse, such as scriptural exegesis or hagiographies delivered before live audiences, the Low variety dominates, with speakers only resorting to the H variety when quoting from a text or when driving home a point.

The practice of shifting into a higher register for emphasizing a point is, in fact, one of the motivations for declaiming in H in any discursive context religious or secular. Otherwise, all speakers, regardless of how well educated and how much they employ H in their professional lives, spend most of their time in speech situations in which L predominates. That notwithstanding, most speakers are under the impression that H is an important element of their daily experience, even if most of them do not themselves actively employ it to any large degree. Even so, many would endorse the notion that it should be used in most or all situations. There are always tales, probably apocryphal, of one or two particularly literate individuals who will speak only H, even at home (leading the less reverent to pity their long-suffering spouses).

Despite the predominance of L, that H is paramount is acknowledged by most people, and some of its registers are held in the highest reverence, for instance as the vehicle of Christian and Muslim scripture. What is more, the notion of H as a unifying element of supranational Arab identity is paid a great deal of lip service. Nevertheless, its place as an oral medium, or indeed a vehicle of secular writing, is less widely appreciated and more often contested (Haeri 2003). For one thing, the written Arabic of the daily and weekly press is not generally recognized as meeting the high rhetorical standards of the venerated classical varieties of H (Parkinson 1991). Again, in that respect, the H variety is more ideal than real. It is, nevertheless, an ideal to be encountered daily, in radio and television broadcasts of scriptural recitations and exegesis. On the other hand, as a vehicle of daily speech, it is rarely employed, and attempts to do so are met with a certain amount of derision. This is exploited to great effect in film, wherein characters using or attempting to use H in speech, especially in daily life, are often portrayed as pompous, ridiculous, or sometimes sinister. In a recent comedy, for example, in a scene portraying a meeting at the ministerial level in which an intractable problem is being discussed, an eager up-and-comer announces his elegant solution in flowery H, whereupon the deputy minister, who is chairing the meeting, comments, “Well, I didn’t understand a word of that, but if you all agree, we can give it a try”. This is indeed a paradox: the Arabic of writing and declamation is at once revered and disparaged.

Something similar may be said of the spoken vernacular. There are multitudinous vernaculars spoken in Egypt, displaying wide geographical variation, the principal divisions recognized readily by most speakers being that of Upper Egypt (called *ṣaʿīdi*) and that of the capital city, with Alexandria and the Delta sometimes acknowledged as possessing defining attributes, (see above, Sec. 1). In reality there are many more distinctions to be drawn along the lines of geography, socioeconomic status, age, and gender. Speakers of Egyptian Arabic recognize these to a limited degree, and in order to do so, they must also possess a perception of some standard by which those distinctions are to be contrasted. That standard is the spoken vernacular of the professional classes of the capital city, Cairo (Haeri 1996). Egyptians, especially those who speak this variety with native facility, tend to regard it with a certain pride of ownership, it being a distinguishing mark of identity, both within Egypt and indeed throughout the Arabophone world (El-Hassan 1977). It is this variety that is labeled ‘Egyptian Arabic’ (*maṣri*) by Arabic speakers, even though there are many other varieties that might also lay equal claim to the appellation. What is more, being accessible to almost all Egyptians in any part of the country – anyone possessing a radio or a television – it genuinely is a de facto standard Arabic and is viewed as an appropriate variety for all occasions, up to and including formal situations like meetings and public addresses (where, of course, H may also be employed). Despite this, people will make disparaging remarks about the vernacular varieties, assuming that they are somehow deficient in important respects. Typical positions will be to assert that they lack grammaticality; that they are coarse; or that they are inappropriate for discussions of a scholarly, technical, or high-culture nature. Both the H and the L varieties are, then, valued and demeaned for different reasons.

These attitudes indicate that there do indeed exist some domains in which exclusive use of one or the other of the two varieties are deemed appropriate but that those in which H is actually used exclusively are quite delimited, being only newscasts, official announcements, and public addresses. In all others, the Cairene vernacular variety of L is the standard, if not always acknowledged as such.

Bibliographical References

Behnstedt, Peter. 1980. “Zum ursprünglichen Dialekt von Alexandria”. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 130.35–50.

——. 1998. “La frontière orientale des parlers maghrébins en Égypte”. *Peuplement et arabisation au Maghreb occidental: Dialectologie et histoire,* ed. Jordi Aguadé, Patrice Cressier, and Ángeles Vicente, 85–96. Madrid and Saragossa: Casa de Velázquez.

Behnstedt, Peter and Manfred Woidich. 1982. “Die ägyptischen dasen: Ein dialektologischer Vorbericht”. *Zeitschrift für Arabische Linguistik* 8.39–71.

——. 1983. *Ägypten: Arabische Dialekte.* (= *Tübinger Atlas zum Vorderen Orient*, map A VIII.) Wiesbaden: L. Reichert.

——. 1985–1999. *Die ägyptisch-arabischen Dialekte.* 5 vols. Wiesbaden: O. Harrassowitz.

——. 2005. *Arabische Dialektgeographie: Eine Einführung.* Leiden: E.J. Brill.

Björnesjö, Sophia. 1996. “L’arabisation de l’Égypte: Le témoignage papyrologique”. *Egypte Monde Arabe* 27–28.93–106.

Blanc, Haim. 1964. *Communal dialects in Baghdad*. Cambridge, Mass.: Harvard University.

——. 1973–1974. “La perte d’une forme pausale dans le parler arabe du Caire”. *Mélanges de l’Université Saint Joseph* 48.375–390.

Bliss, Frank. 1998. *Siwa, die Oase des Sonnengottes: Leben in einer ägyptischen Oase vom Mittelalter bis in die Gegenwart*. Bonn: PAS.

Doss, Madiha. 1981. *Le dialecte ṣaʿīdi de la région de Menya*. Thèse de IIIème cycle, Université de la Sorbonne Nouvelle (Paris–3).

El-Hassan, S.A. 1977. “Educated spoken Arabic in Egypt and the Levant: A critical review of diglossia and related concepts”. *Archivum Linguisticum* 8:2.112–131.

Ferguson, Charles A. 1959. “Diglossia”. *Word* 15.325–340.

Haeri, Niloofar. 1991. *Sociolinguistic variation in Cairene Arabic: Palatalization and the* qaf *in the speech of men and women*. Ph.D. diss., University of Pennsylvania.

——. 1996. *The sociolinguistic market of Cairo: Gender, class, and education*. London: Kegan Paul International.

——. 2003. *Sacred language, ordinary people: Dilemmas of culture and politics in Egypt*. New York: Palgrave Macmillan.

Hobbs, Joseph J. 1989. *Bedouin life in the Egyptian wilderness*. Austin: University of Texas Press.

Jong, Rudolf E. de 2000. *A grammar of the Bedouin dialects of the northern Sinai littoral: Bridging the linguistic gap between the eastern and the western Arab worlds*. Leiden: E.J. Brill.

——. 2002. “Notes on the dialect of the ʿAbābda”. *Sprich doch mit deinen Knechten aramäisch: Festschrift für Otto Jastrow zum 60. Geburtstag*, ed. Werner Arnold and Hartmut Bobzin, 337–360. Wiesbaden: O. Harrassowitz.

——. 2004. “Characteristics of Bedouin dialects in southern Sinai”. *Approaches to Arabic dialects: A collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday*, ed. Martine Haak, Rudolf de Jong, and Kees Versteegh, 151–175. Leiden: E.J. Brill.

Khalafallah, Abdelghany A. 1969. *A descriptive grammar of Saεi:di Egyptian Colloquial Arabic*. The Hague and Paris: Mouton.

Malem, Fathi. 2001. *Oasis Siwa from the inside: Traditions, customs and magic*. Cairo: Al Katan.

Maṭar, ʿAbd al-ʿAzīz. 1967. *Lahjat al-badw fī ʾiqlīm sāḥil Maryūṭ*. Cairo: Dār al-Kitāb al-ʿArabī.

Miller, Catherine. 1996. “Nubien, berbère et beja: Notes sur trois langues vernaculaires non-arabes de l’Égypte contemporaine”. *Egypte Monde Arabe* 27–28.411–431.

Murray, G.W. 1935. *Sons of Ishmael: A study of the Egyptian Bedouin*. London: George Routledge.

Nishio, Tetsuo. 1992. *A basic vocabulary of the Bedouin Arabic dialect of Jbāli tribe*. Tokyo: Institute for the Study of Languages and Culture of Asia and Africa.

Parkinson, Dilworth B. 1985. *Constructing the social context of communication: Terms of address in Egyptian Arabic*. Berlin and New York: Mouton de Gruyter.

——. 1991. “Searching for modern Fusha: Real life formal Arabic”. *al-ʿArabiyya* 24.31–64.

Peterson, Jennifer Leigh. 2000. “Contemporary Cairene youth language: Linguistic corruption or social art?” *Fourth International Conference of the Association Internationale de la Dialectologie Arabe* (*AIDA*), 1–4 April 2000, Marrakech, Morocco, ed. Abderrahim Youssi a.o., 422–429. Rabat: AMAPATRIL.

Rouchdy, Aleya. 1991. *Nubians and the Nubian language in contemporary Egypt: A case of cultural and linguistic contact*. Leiden: E.J. Brill.

Royal, Ann M. 1985. *Male/female pharyngealization patterns in Cairo Arabic: A sociolinguistic study of two neighborhoods*. Ph.D. diss., University of Texas at Austin.

Rubenson, Samuel. 1996. “The transition from Coptic to Arabic”. *Egypte Monde Arabe* 27–28.77–91.

Wahba, Kassem. 1993. *A sociolinguistic treatment of the feature of emphasis in Egyptian Arabic*. Ph.D. diss., Alexandria University.

Wilmsen, David. 1999. “*Haaga taaniʔ*: An examination of degendered adjectival agreement in Cairene Arabic”. *al-ʿArabiyya*, 32.217–234.

Winkler, Hans Alexander. 1936. *Ägyptische Volkskunde*. Stuttgart: Kohlhammer.

Woidich, Manfred. 1993. “Die Dialekte der ägyptischen Oasen: Westliches oder östliches Arabisch?” *Zeitschrift für Arabische Linguistik* 25.340–359.

——. 1994. “Cairo Arabic and the Egyptian dialects”. *Actes des premières journées internationales de dialectologie arabe de Paris*, ed. Dominique Caubet and Martine Vanhove, 493–507. Paris: INALCO.

——. 1995. “Al-Saʿid: Dialects”. *Encyclopaedia of Islam*, VIII, 866b–867b. 2nd ed. Leiden: E.J. Brill.

——. 1995–97. “Negation in Dakhla Oasis: The morphological reinterpretation of a phonological rule”. *Mediterranean Language Review* 9.13–28.

——. 1996. “Rural dialects of Egypt”. *Egypte Monde Arabe* 27–28.325–354.

——. 1997. “Upper Egyptian Arabic and dialect mixing in historical perspective”. *Humanism, culture and language in the Near East: Studies in honor of Georg Krotkoff*, ed. Asma Afsaruddin and A.H. Mathias Zahniser, 185–197. Winona Lake, Ind.: Eisenbrauns.

——. 1998. “Aus den Erinnerungen eines Hundertjährigen: Ein Text im Dialekt von Balāṭ in Ost-Dakhla/Ägypten”. *Estudios de Dialectología Norteafricana y Andalusí* 3.7–33.

David Wilmsen (American University in Cairo)
Manfred Woidich (University of Amsterdam)

Elative

1. INTRODUCTION

The word ‘elative’, from the Latin *elatio*, noun of action of the verb *efferre* ‘to elevate’, refers to a morphosemantic entity and expresses in one word what traditional Arabic grammar expresses in two words, *ʾafʿalu at-tafḍīl*. *ʾAfʿalu* indicates, through the conventional *f-ʿ-l* paradigm of Arabic grammar, the word pattern (*aṣ-ṣīġa* ‘pattern, scheme’) and *at-tafḍīl* ‘superiority’ indicates the intended meaning among all the possible different meanings of this pattern (e.g. the masculine singular pattern of the adjective of color, *ʾaswadu* ‘black’, or the 1st person singular imperfect of Form I, *ʾaḏhabu* ‘I go’, etc.). Occasionally, as Wright (1974:I, 140) points out, the intended meaning “has the signification of our comparative and superlative, and is therefore called *ism at-tafḍīl*, noun of pre-eminence, or *ʾafʿalu at-tafḍīl*, the pattern *ʾafʿalu* denoting pre-eminence”. According to an Arab grammarian (Šartūnī 1949:IV, 70) “*ʾafʿal at-tafḍīl* is a pattern [*ṣīġa*] employed to describe something that possesses a ‘plus’ in comparison to something else: *yūsuf ʾakbar min būlus* ‘Joseph is bigger than Paul’”.

2. Origin and evolution

The origin of the elative in Arabic was discussed by Speiser (1952:81), who argues that “Semitic in general had once an elative or emphatic form indicated by a special prefix”. As far as Arabic is concerned, Wehr (1952:34) points out that “the stem of the form, i.e. the part following the prefix *a-*, had originally been a *nomen substantivum* (an abstraction)”. In Bravmann’s view (1968:33), “the form *ʾafʿalu* represents the result of a transformation of a certain basic adjectival pattern (‘positive’) within the context of a comparison of inequality (superiority), i.e. when used with the function of comparative-superlative”.

The semantic evolution of the form *ʾafʿalu* is discussed by Wehr (1952:3), who explains that *ʾafʿalu* originally indicated a positive with a strongly emotional connotation, and by Bravmann (1968:22), who “persists in the opinion that the primary function of *ʾafʿalu* is to indicate a high degree of a quality in comparison with other objects (*ʾafʿal at-tafḍīl*)” and that “the original use of *ʾafʿalu* as a comparative-superlative may secondarily occur in the sense of a positive. Thus, one cannot attribute to *al-akbaru* the primary meaning of ‘the great’ (or with affective connotation ‘the very great’) and the secondary meaning of ‘the greater’ or ‘the greatest’; on the contrary, it is the meaning ‘the great’ (which implies no comparison), which should be considered as secondary”. Parallel to this debate is the controversial *allāhu ʾakbar* (cf. Ibn Manẓūr, *Lisān* III, 211).

3. Formation of the elative

The elative is formed through modification of the triliteral stem of the adjective according to the patterns in Table 1:

Table 1. Patterns of the elative

		masculine	feminine
singular		*ʾafʿalu*	*fuʿlā**
dual	independent	*ʾafʿalāni*	*fuʿlayāni*
	dependent/ oblique	*ʾafʿalayni*	*fuʿlayayni*
plural		*ʾafʿalūna* or *ʾafāʿil*	*fuʿlayāt* or *fuʿal*

* the *ā* is an *ʾalif maqṣūra*

For example, in the case of the adjective *kabīr* ‘big’, the three radical consonants *k-b-r* replace the three consonants of the *faʿala* paradigm (Table 2):

Table 2. Patterns of *ʾakbaru*

		masculine	Feminine
singular		*ʾakbaru*	*kubrā*
dual	independent	*ʾakbarāni*	*kubrayāni*
	dependent/ oblique	*ʾakbarayni*	*kubrayayni*
plural		*ʾakbarūna* or *ʾakābir*	*kubrayā* or *kubar*

We shall further expand on the syntactic reasons which justify the extremely rare occurrence of most of these forms, almost entirely superseded by the masculine singular form *ʾafʿal*.

When the last two consonants of the root are identical, as in *qalīl* ‘few’, the pattern obtained is ʾaC1aC2C3: *ʾaqallu* ‘less/least’; in practice, this form of elative does not occur in the feminine, probably on account of the difficulties in reading such forms without vocalization.

When the last consonant of the root is *w* or *y*, for example *ḥaluw* or *qawiyy*, the pattern obtained is ʾaC1C2 (ā = *ʾalif maqṣūra*) for the masculine, *ʾāḥlā* or *ʾaqwā*; the feminine pattern, C1uC2y (ā = *ʾalif*) rarely occurs, as in *ʾaʿlā*, *ʿulyā* 'higher' or *ʾadnā, dunyā* 'lower'.

Two nouns which do not follow the *ʾafʿal* pattern behave as elatives in comparative constructions: *xayr* 'good' and *šarr* 'bad', *ʾanta xayr min-hu* 'you are better than him', *hiya šarr min ʾuxti-hā* 'she is worse than her sister'. However, according to aš-Šartūnī (1949:IV, 70) "the origin of these two words is *ʾaxyar* and *ʾašarr* (in the *ʾafʿal* elative pattern) but their [initial] *hamza* has been elided due to the frequent use [of these words]". Bravmann, (1968:36) on the other hand, claims that these two nouns do not admit the prefixed prosthetic vowel "on account of the monosyllabic stem of these adjectives" and "because of their extremely frequent use".

Aš-Šartūnī (1949:IV, 70) spells out a number of rules underlying the formation of the elative: the verb "must have a triliteral root (*ṯulāṯī*), a complete conjugation (*mutaṣarrif*), be in the active form (*maʿlūm*), in plain sense (*tāmm*), accept [the degree of] superiority (*qābil li-l-mufāḍala*) and express neither a color nor a defect or ornament (*ḥilya*), for example *ʾanta ʾaʿlam min ʾaxī-ka* 'you are wiser than your brother'".

"It is not possible to form the elative in the following cases: *zāḥama* 'to pile up', because the verb is quadriliteral; *niʿma* 'bravo!', because the verb cannot be conjugated; *ḥumida* 'to be rented', because it is a passive form; *kāna* 'to be', because it does not denote a full meaning; *faniya* 'to disappear' and *māta* 'to die', because they do not admit the superiority [degree]; *xaḍira* 'to be green', because it denotes a color; and *ʿamiya* 'to be blind', because it denotes an illness" (Šartūnī 1949:IV, 70).

All authors record a number of exceptions to the above rule:

i. Examples "formed from the derived forms of the verb, especially from IV: *ʾaṭharu* 'more cleansing' or 'purifying' (*ʾakṯar taṭhīran*), from *ṭahhara* 'to cleanse' or 'purify', II. of *ṭahura* 'to be clean or pure'; [...] *ʾaṯbatu li-* 'making more firm' or 'sure', from *ʾaṯbata*, IV. of *(ṯabata* 'to be firm'; [...] *ʾaḥwalu min* 'more crafty than' from *iḥtāla*, 'to be crafty' VIII. of *ḥāla*." (Wright 1974:I, 141).

   Yet, Blachère (1975:98) argues: "Les grammairiens citent des exemples d'élatifs qui seraient issus de participes ou d'adjectifs verbaux, provenant de verbes à la forme 'nue' ou à une forme dérivée, particulièrement à la $4^{\text{ème}}$, et ayant le sens actif ou passif. Mais il leur est, en général, impossible de donner un exemple du participe ou de l'adjectif au degré simple.

ʾanṣafa être juste	*ʾanṣafu* juste
ʾaqfara être désert	*ʾaqfaru* désert

   On peut se demander [...] si ces verbes d'état, dits de $4^{\text{ème}}$ forme, ne sont pas, au contraire, formés des élatifs, de même que les verbes dits de $9^{\text{ème}}$ forme proviennent des adjectifs de couleur et de difformité".

ii. In both Classical and contemporary Arabic and in the dialects, the adjective of color or deformity *ʾafʿalu* may be employed as an elative; in these cases the elative has a comparative or superlative sense, depending on the context, as such adjectives already bear the pattern *ʾafʿal* (Blachère 1975:98):

   "*ʾabyaḍu-hum* le plus blanc d'entre eux
   abyaḍ min uḫt...plus blanc que la sœur de...(*Cor*. XVII, 74)
   wa-man kāna fi hāḏihi ʾaʿmā fa-huwa fi l-ʾāḫira ʾaʿmā wa-ʾaḍall sabīlan qui sera aveugle dans ce monde, sera, dans l'autre, plus aveugle et plus fourvoyé (*Cor*. XVII, 74)".

   But Blachère (1975:98) clarifies: "Je traduis 'plus aveugle et plus fourvoyé'; ce serait plutôt 'spécialement, complètement aveugle et égaré'", which is confirmed by aš-Šartūnī (1949:IV, 71): "The elative can be deprived of its meaning of superiority and acquire the meaning of the adjective with a nuance of exaggeration."

The loss of the sense of superiority of the elative is frequent in some current expressions, often related to historical events or geographical areas, e.g. *al-ḥarb al-ʿuḏ̣mā* 'the Great War', *al-ʿuṣūr al-wusṭā* 'the Middle Ages', *Barīṭāniyā l-ʿuḏ̣mā* 'Great Britain', *aš-Šarq al-ʾawsaṭ* 'the Middle East', *Miṣr al-ʿulyā* 'Upper Egypt'. These examples show the relational character of the elative, in which the relation of comparison, though still present, does not entail a comparison of degrees.

To sum up, the main rule states that the elative may be formed exclusively from a triliteral adjective or participle (cf. in English, the *-er* and *-est* suffixes); in addition, there are forms such as *ʾakṯaru* 'more numerous' for the comparative and 'the most numerous' for the superlative, or *ʾašaddu* 'stronger' or 'the strongest', followed by the abstract or verbal noun semantically corresponding to the adjective or participle. This verbal noun is analyzed as a specifying complement (*tamyīz*), i.e. dependent case, e.g. for *muzdaḥim* 'congested, cluttered up' *madīnat al-Qāhira ʾakṯar izdiḥāman min ʾAswān* 'the city of Cairo is more congested than Aswan'; for *mujtahid* 'studious' *huwa ʾašadd aṭ-ṭullāb ijtihādan* 'he is the most studious of the students'.

As for the comparatives and superlatives indicating inferiority, they are formed on the same pattern by using the elative *ʾaqallu* 'less' always followed by a verbal noun in dependent case, e.g. *huwa ʾaqall min-hā ijtihādan* 'he is less studious than she', *huwa ʾaqallu-humā ijtihādan* 'he is the less studious of the two'.

4. SYNTAX OF THE ELATIVE

4.1 *The comparative*

The elative "must be deprived of [the article] *al-* and must not be in the annexion state: it is followed by *min*, expressed or omitted, which introduces what the first element is superior to; the elative occurs in the masculine singular, e.g. *al-ʾasad ʾaqwā min al-rajul* 'the lion is stronger than the man', *ar-rijāl ʾafḍal mina l-ʾusd* 'men are superior to lions'" (Šartūnī 1949:IV, 315). The syntagm introduced by *min* may be omitted when answering a question, for example: *hal ar-rajul ʾaqwā min al-ʾasad? lā, al-ʾasad ʾaqwā* 'is man stronger than the lion? No, the lion is stronger'.

4.2 *The superlative*

The elative acquires definiteness through the article: it normally agrees in gender and number, e.g., *al-mudun al-kubrā* 'the biggest cities'; *al-marʾatāni l-fuḍlayāni* 'the two most virtuous women'; *hum al-ʾakābir* 'they are the biggest'; *hunna l-fuḍlayāt* 'they [fem.] are the most virtuous'. In modern written Arabic this construction occurs less frequently than the following ones, with the exception of some current expressions, such as the above-mentioned *aš-Šarq al-ʾawsaṭ*.

The elative is the first element of an annexion in which the second element is indefinite: the elative remains in the masculine singular and "in this case, the second element of the annexion must necessarily be of the same gender as the subject, and it must agree with it in number, as in *az-zaydāni ʾafḍal rajulayni* 'the two Zayds are the most virtuous men', *al-maryamāt ʾafḍal nisāʾ* 'the Maryams are the most virtuous women'" (Šartūnī 1949:IV, 316).

The elative is the first element in an annexion in which the second element is definite: the elative may occur in the masculine singular but it "may also agree in gender and number with the object or the objects spoken of as *hiya fuḍlā n-nisāʾ* 'she is the best of the women'; *humā ʾafḍalā l-qawm* 'these two are the two best of the tribe'; *hum ʾafḍalū l-qawm* or *hum ʾafāḍil al-qawm* 'they are the best of the tribe'; *hunna fuḍlayāt an-nisāʾ* or *hunna fuḍal an-nisāʾ* 'they are the best of the women'..." (Wright 1974: II, 228).

Unfortunately, the above examples apply mainly to nouns denoting human beings (*ʿāqil*). The issue of the agreement with non-human being nouns (*ġayr ʿāqil*) in modern written Arabic appears more complex; examples such as *mā min šakk fī ʾanna min kubrā ihtimāmāt ad-dawla hiya*...'no doubt that the major preoccupations of the State...' (aṣ-Ṣabāḥ, 14 July 2002) lead to the supposition that it is not the object's gender which determines the agreement in the feminine singular (*ihtimām* is in fact a masculine word), but rather the fact that *ihtimāmāt* is a non-human being noun plural (Girod 2000:78).

One can debate the semantic difference between the last two constructions, for example between *ʾakbar madīna* 'the biggest city' and *ʾakbar al-mudun* or *kubrā l-mudun* 'the biggest of the cities'. According to Wright (1974:II, 226), the genitive which follows the elative "is at times indefinite and explicative, at times definite and partitive". Blachère (1975: 366) is less dogmatic: "Parfois le second

terme de l'annexion est un singulier ou un duel indéterminé; on a alors une annexion de qualification notant un superlatif vague que le français rendra, selon le contexte, par un superlatif absolu ou relatif.

ʾašadd ʿaḏāb un extrême tourment/le plus dur tourment
ʾantum xayr ʾumma vous êtes un peuple excellent/le meilleur peuple".

5. CONCLUSION

Despite the relevant stability in the use of the elative throughout the classical and modern period, it is not unrealistic to predict further developments in the near future, e.g. the curious case of the hyperbolic use of the elative in the feminine plural, doubtlessly unacceptable according to the grammatical norm, but which might well become 'jurisprudence' (Girod 2000:78): *al-ittifāq maʿa kubrayāt aš-šarikāt al-ʾamrīkiyya* 'the agreement with the major American companies' (*al-ʾAhrām* 1 April 2000).

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCE

Ibn Manẓūr, *Lisān* = ʾAbū l-Faḍl Jamāl ad-Dīn Mukarram ibn Mukarram Ibn Manẓūr, *Lisān al-ʿArab al-muḥīṭ*. Ed. Yūsuf Xayyāṭ. 4 vols. Beirut: Lisān al-ʿArab.

SECONDARY SOURCES

Blachère, Régis. 1975. *Grammaire de l'arabe classique*. Paris: Maisonneuve et Larose.
Bravmann, Max. 1968. *The Arabic elative: A new approach*. (= *Studies in Semitic Languages and Linguistics*, 2.) Leiden: E.J. Brill.
Ġalāyinī, Muṣṭafā al-. 1973. *Jāmiʿ ad-durūs al-ʿarabiyya*. Beirut: al-Maktaba al-ʿAṣriyya.
Girod, Alain. 2000. *Faits d'évolution en arabe moderne à travers un corpus de presse égyptien*. Ph.D. diss., University of Provence.
Krahl, Günther, Wolfgang Reuschel, and Eckehard Schultz. 1995. *Lehrbuch des modernen Arabisch*. Leipzig: Langenscheidt and Verlag Enzyklopäedie.
Šartūnī, Rašīd aš-. 1949. *Mabādiʾ al-ʿarabiyya fī ṣ-ṣarf wa-n-naḥw li-talāmīḏ as-sana ar-rābiʿa*. Beirut.
Speiser, E.A. 1952. "The 'elative' in West-Semitic and Akkadian". *Journal of Cuneiform Studies* 6.81–92.
Wehr, Hans. 1952. *Der arabische Elativ*. Wiesbaden: Akademie der Wissenschaften und der Literatur in Mainz.
Wright, William. 1974. *A grammar of the Arabic language*. Repr. Beirut: Librairie du Liban.

ALAIN GIROD (University of Aix-en-Provence)

Elision

The two most common terms in the Arabic linguistic tradition for the concept of 'elision' are → *ḥaḏf* lit. 'cutting off, curtailing' and → *ʾiḍmār* lit. 'keeping in mind', but there is a wide range of other expressions for the omission or deletion of linguistic elements in Arabic, and it would be impossible (and indeed undesirable) to equate them strictly with any modern Western terms. Before elision proper can be dealt with, four groups of words will be disposed of, those which (1) are hardly technical in nature, (2) mainly concern the inflectional system, (3) indicate the phonological motives for elision, or (4) refer to the stylistic or rhetorical purposes of elision.

In the first group are *suqūṭ* 'falling away [of a sound]' and *ḏahāb* 'departure [of a sound]', which simply denote the effect of elision. Here may also be included, since their status as technical terms is not clear, some synonyms of *ḥaḏf*, namely *kaff* 'refraining [from saying something]' and *xazala*, *ixtazala* 'to cut off, withhold'. These occur in the context of *ḥaḏf* as if to paraphrase or gloss it (unless they represent an alternative vocabulary from a different grammatical tradition). Thus *kaff* is mentioned by Ibn Fāris as "one of the speech habits (*sunan*) of the Arabs" (*Ṣāḥibī* 197, 240, 256), and in Sībawayhi the elided verb is said to be 'withheld' (*xuzila*) in such exclamations as *subḥāna llāhi* 'praise to God!' (*Kitāb* I, 135, ed. Derenbourg / I, 162, ed. Būlāq). *Xazl* also has a restricted use in metrics to denote the elision of a medial vowel.

The second group comprises *taskīn* and *ʾiskān*, lit. 'silencing', i.e. removing a vowel or voweling with zero, *jazm* 'lopping off [an inflectional vowel]', and *waqf* 'stopping'. While vowellessness occurs in medial or final position in any word, *jazm* is specific to verbs as the marker of the apocopated (*majzūm*) mood, contrast *yaktubu* 'he writes' with *ʾin yaktub* 'if he write'. Note that if the apocopation requires the removal of a final consonant, it is termed *ḥaḏf*, contrast *yaktubūna* 'they write' with *yaktubū* 'they might write'. The phenomenon of *waqf*, generally translated as 'pause', involves a reduction in the word-ending to signal that the speaker is about to take a breath, e.g. (with the removed elements in angle brackets) *hāḏā sayf<un>* 'this is a sword', *fī l-madīna<ti>* 'in the town' (→ pausal forms). It will be seen that the truncation is more extensive than simply dropping the final short vowel.

A third set of terms specifies the articulatory conditions or motives for an elision, viz. *taxfīf* 'lightening [a syllable]' and *istixfāf* 'treating [a syllable] as light', usually in → syncope or apocope, i.e. dropping a vowel or consonant due to frequency of occurrence or lack of accent, as in the common poetic variant *lam yaku* for *lam yakun* 'he was not', called by Ibn Fāris (*Ṣāḥibī* 45) 'elision for the sake of lightening' (*al-ḥaḏf li-t-taxfīf*).

The fourth group consists of essentially rhetorical terms, clarifying the motive or means of an elision, viz. *ʾījāz* 'being succinct', *ittisāʿ* 'exercising latitude', *iqtiṣār* 'keeping short', *ixtiṣār* 'being brief', which occur as often as not in collocation with *ḥaḏf*; thus, *ixtiṣār* is associated with *ḥaḏf* as 'one of the speech habits of the Arabs' by Ibn Fāris (*Ṣāḥibī* 205), quoting among others the famous Qurʾānic example Q. 12/82, *wa-sʾal-i l-qaryata* 'ask <the people of> the village'. Al-Jurjānī (*ʾIʿjāz* 95–114) has an entire section extolling the succinctness of the *Qurʾān*, and the allusive economy of Qurʾānic rhetoric is likewise a major theme for Ibn Hišām (*Muġnī* II, 160; cf. Gully 1994:212f.). In one passage of the *Kitāb* (I, 88, ed. Derenbourg/I, 108, ed. Būlāq), Sībawayhi uses no fewer than three terms in discussing the elliptical *al-qaryata* for '<the people of> the village', which, he says, exploits the latitude (*ittisāʿ*) of the language for the sake of brevity (*ixtiṣār*) and succinctness (*ʾījāz*).

The two most important terms are no doubt *ḥaḏf* 'eliding an element' and → *ʾiḍmār* 'suppressing an element'. It is evident that for the Arabs *ḥaḏf* covered not only 'elision' proper, the omission of parts of words (i.e. syncope and apocope), but also what is differentiated as 'ellipsis', the omission of parts of a syntactic structure. These elisions (the term will be used for both here) are not all morphological but may also be stylistic or hypercoristic, most notably in the curtailing *ḥaḏf* of proper names in the process of *tarxīm* lit. 'softening', i.e. 'shortening a word', as in *yā māli* 'O Māli!', addressing someone called Mālik. *Ḥaḏf* is best understood through a selection of illustrations:

Phonological: *lam yaku* 'he was not', from *yaku<n>*; here may be included the many cases where Classical Arabic prefers an elided form, e.g. the 'Eastern' *jumʿa* 'Friday' over the unelided *jumuʿa* of the 'Western' pre-Islamic dialects. The metrical sense of *ḥaḏf* for eliding the final syllable of a foot can also be mentioned here.

Morphophonological: *yaṣilu* 'he arrives' from *ya<w>ṣilu*, root *w-ṣ-l*; *yaqum* 'he might stand', from *yaqu<w>m*, root *q-w-m*; *lam yaqḍi* 'he did not finish', from *yaqḍi<y>*, root *q-ḍ-y*.

Morphological: *kitābu r-rajuli* 'the book of the man', from *kitābu<n>*, losing the *n* which here is an indefinite marker incompatible with annexation; *yakūnū* 'they might be', marked privatively by elision of final *-na*, contrast independent *yakūnūna* 'they are'.

Syntactic: *lā <baʾsa> ʿalayka* 'there is no <harm intended> to you'; *ʾanta ḏālimun ʾin faʿalta ḏālika <fa-ʾanta ḏālimun>* 'you would be wrong if you did that < you would be wrong>' (an apodosis cannot precede its protasis, and one must be assumed to have been elided here); *hal qāma zaydun? naʿam <qāma zaydun>* 'did Zayd stand up? Yes <Zayd did stand up>'.

Bibliographical references

Primary sources

Ibn Fāris, *Ṣāḥibī* = ʾAbū l-Ḥusayn ʾAḥmad Ibn Fāris, *aṣ-Ṣāḥibī fī fiqh al-luġa*. Ed. Moustafa Chouémi. Beirut: A. Badran, 1964.

Ibn Hišām, *Muġnī* = Jamāl ad-Dīn ʾAbū Muḥammad ʿAbdallāh ibn Yūsuf Ibn Hišām, *Muġnī l-labīb ʿan kutub al-ʾaʿārīb*. Cairo: ʿĪsā al-Bābī al-Ḥalabī, n.d.

Jurjānī, *ʾIʿjāz* = ʾAbū Bakr ʿAbd al-Qāhir ibn ʿAbd ar-Raḥmān al-Jurjānī, *Dalāʾil ʾiʿjāz al-Qurʾān*. Ed. M.R. Riḍā. Cairo. (Repr., Beirut: Dār al-Maʿrifa, 1981.)

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. Ed. Hartwig Derenbourg. Paris: Imprimerie Nationale. (Repr., Hildesheim: G. Olms, 1970.) Ed. Būlāq, 1898–1900. (Repr., Baghdad: al-Muthanna, 1965.) Ed. ʿAbd as-Salām Hārūn, Cairo: Dār al-Qalam, 1968–1977 [and later eds.; see *ḥaḏf* in grammatical index, Vol. V.]

Secondary sources

Carter, Michael G. 1991. 'Elision'. *Proceedings of the Colloquium on Arabic Grammar, Budapest 1–7 September 1991*, ed. Kinga Dévényi and Tamas Iványi, 121–133. Budapest: Eötvös Loránd University and Csoma de Kőrös Society. [The bibliography lists additional sources to those mentioned here.]

Gully, Adrian J. 1995. *Grammar and semantics in Medieval Arabic*. Richmond, Surrey: Curzon.

Jahn, Gustav. 1895–1900. *Sîbawaihi's* Buch über die Grammatik, übersetzt und erklärt. Berlin: Reuther und Reichard. (Repr., Hildesheim: G. Olms, 1969.)

Kinberg, Naphtali. 1996. *A lexicon of al-Farrāʾ's terminology in his* Qurʾān *commentary, with full definitions, English summaries, and extensive citations*. Leiden: E.J. Brill. [s.vv. *ḥaḏafa*, *ḥaḏf*.]

Troupeau, Gérard. 1976. *Lexique-index du* Kitāb *de Sībawayhi*. Paris: Klincksieck. [Lists and counts all relevant terms but does not give page references for the 553 occurrences of *ḥaḏafa* or the 214 of *ḥaḏf*; for these, see ʿUḍayma 1975 and the *Kitāb* edition by Hārūn.]

ʿUḍayma, ʿAbd al-Xāliq. 1975. *Fahāris Kitāb Sībawayhi wa-dirāsa lahu*. Cairo: Maṭbaʿat as-Saʿāda.

Weir, T.H. 1913–1938. "*Ḥaḏh̲f*". *Encyclopaedia of Islam* III, 188. Leiden: E.J. Brill. [No significant mention of this topic in the 2nd ed.]

MICHAEL G. CARTER (Sydney University)

Ellipsis

1. DEFINITION

Ellipsis is "a discourse phenomenon, in the sense that the interpretation of the missing constituent sometimes depends on something said in an earlier sentence – possibly even by another speaker" (Sag and Wasaw 1999:313). For example, in (1) the sentential subject of the embedded clause, *inni aftaḥ il-bāb* 'that I open the door', is elided and only the negated predicate remains. The first clause includes an antecedent, which is morphologically, syntactically, and semantically identical to the missing constituent, hence facilitating its interpretation.

(1) *ḥāwilt* *in-ni aftaḥ* *il-bāb*
tried-1s. that-I open.1s. the-door
bass *ma-ʾamkan-š* (Egyptian Arabic)
but NEG-was.possible-NEG
'I tried to open the door, but I could not'

The antecedent of an ellipsis can be included in preceding discourse rather than a preceding clause. The fragment comprising speaker Y's answer in (2), for example, includes only the perfect form of the auxiliary verb *kān* 'be', which is marked for tense, number, person, and negation, whereas the adverbial predicate *liwaḥdi sāʿit il-ḥadsa* 'alone at the time of the incident' is left out. The missing predicate can be reconstructed in relation to the antecedent predicate, which is provided in speaker X's question, even though the antecedent is not morphologically identical to the elided constituent.

(2) X. *inta* *kunt-ə* *l-waḥdak*
you.2ms were alone
sāʿit il-ḥadsa (Egyptian Arabic)
time the-incident
'Were you alone at the time of the incident?'
Y. *laʾ! ma-kunt-iš.* *ʿali kān maʿā-ya*
no! NEG-was-NEG Ali was with-me
'No! I wasn't. Ali was with me'

Although the missing constituents in an ellipsis and their antecedents usually have the same morphological and syntactic structure, they do not necessarily have the same meaning, i.e., ellipsis sometimes displays 'sloppy identity' where the missing noun phrases are referentially ambiguous (Hardt 1999; Baltin 2003). For example, the deleted constituent in the second conjunct of (3a) is interpreted as including the indefinite noun phrase *jāʾiza* 'a prize', which is ambiguous with regard to its referent. This noun phrase can be interpreted as referring to the same prize that Jamal won or to another one.

(3) a. *fāza* *jamāl bi-jāʾiza,* *lākinna*
won.3ms Jamal with-prize but
ʿumar *lam* *yastaṭiʿ*
Omar NEG.past be.able.to
(Standard Arabic)
'Jamal won a prize, but Omar couldn't'

The ellipsis in the second conjunct of (3b) displays two cases of semantic mismatch between the gap and its antecedent. First, the noun phrase *kitābha* 'her book', which is interpreted as part of the elided structure, presents a case of 'sloppy identity', as the second conjunct could mean that Mona did not read Hoda's book or that Mona did not read her own book. Second, the deleted predicate is negated, as indicated by the grammatical use of the polarity item *lissa* 'yet', whereas its antecedent is not.

(3) b. *huda* *ʾarit* *kitāb-ha,* *bass*
Hoda read.3fs book-her but
muna *lissa* (Egyptian Arabic)
Mona not yet
'Hoda read her book, but Mona hasn't yet'

The interpretation of an ellipsis does not always depend on its structural or semantic identity with a preceding antecedent because the missing constituent can precede the antecedent, as is the case with anaphoric expressions (Lobeck 1995). In other words, the relation between the missing constituent and its antecedent is not necessarily that of syntactic governing. For example, the Standard Arabic sentence in (4) below includes two cases of ellipsis: (a) the clausal complement of the verb phrase *ḥāwaltu* 'I tried' and (b) the predicate of the second conjunct after the auxiliary verb *ʾastaṭiʿ* 'could'. There is no preceding structure in the same sentence or previous discourse that can function as a syntactic antecedent for either ellipsis. Rather, the only possible antecedent is in the lowest clause in the sentence, namely *ʾan yaktuba r-risāla* 'that he write the letter', with the difference in person marking.

(4) a. *ḥāwaltu* *fa-lam* *ʾastaṭiʿ*
tried.1ms but-NEG. past be able to.1ms
'I tried but I couldn't'
b. *fa-saʾaltu* *ṣadīqī*
so-asked.1ms friend-my
ʾan ya-ktub-a *r-risāla* *badalan*
that 3ms-write-subj. the-letter instead
min-nī
from-me
'so I asked my friend to write the letter instead of me'

Barton (1990) and El-Shiyab (1998) demonstrate that having an antecedent in preceding discourse is not a necessary condition for the acceptability of ellipsis, as it is quite often used without any antecedent at all. For example, the sentence uttered by speaker X in (5) below does not include any constituents that can be used to reconstruct the fragment making up speaker Y's response into a grammatical sentence. However, the response is understood as an explanation of speaker X's observation. The acceptability of examples such as (5) suggests that the interpretation of ellipsis involves logical and pragmatic inferences rather than constituent copying under identity.

(5) X. *muna zaʿlān-a* *ʾawi in-nahāṛda*
Mona upset-f. very today
(Egyptian Arabic)
'Mona is very upset today'
Y. *ḥamat-ha* *zayy-ə*
mother-in-law-her as
ma nta *ʿārif*
pro. you.m.sg know$_{\text{(AP)}}$
'Her mother-in-law, as you know'

The fact that there could be semantic and morphological differences between an ellipsis and its antecedent, if there is one, demonstrates that there is no copying process involved in reconstructing the missing constituent(s). Rather, unpacking ellipsis is a pragmatic process of conversational reasoning, where a speaker's intent is interpreted using linguistic and contextual clues. Green (1996) and Levinson (2000) account for ellipsis in terms of pragmatic axioms such as those proposed in Grice's Cooperative Principle (Grice 1975). For example, redundant information that is contextually prominent and can be retrieved from preceding discourse is left out in accordance with Grice's maxim of quantity: "Make your contribution as informative as is required for the current purposes of the exchange" (Grice 1975:45). Moreover, reconstructing ellipsis is based on the assumption that a speaker's fragment is relevant to the discourse content following Grice's maxim of relevance: "Be relevant" (Grice 1975:46). For example, in (5) above speaker Y is understood to be explaining why Mona is upset rather than introducing a new unrelated topic or contradicting speaker A by asserting that it is Mona's mother-in-law who is upset.

2. TYPES OF ELLIPSIS

Several linguistic phenomena are usually described under the category 'ellipsis', including sluicing, ellipsis within a noun phrase, and verb phrase ellipsis. Sluicing differs from other types of ellipsis in that it is constrained to be immediately preceded by a *wh*-element, but not a lexical complementizer (Lobeck 1995), as illustrated by the Standard Arabic examples in (6a) and (6b). In (6a) the question word *limāḏā* 'why' is followed by a gap that corresponds to the preceding clause, whereas in (6b) the sluicing in the first conjunct is introduced by *matā* 'when', and corresponds to the clause following it.

(6) a. *ṭalabat min-nī mūnā ʾan*
asked.3fs from-me Mona that
ʾarḥala, lākin lā ʾaʿrif limāḏā
leave.1s but NEG know.1s why
'Mona asked me to leave, but I don't know why'
b. *lā ʾa-taḏakkar matā*
NEG 1s-remember when
bi-t-taḥdīd, lākin-nī
with-the-precision but-I
ʾaʿṭaytu-hu l-kitāb
gave.1s-him the-book
'I don't know when exactly, but I gave him the book'

Another distinctive criterion of sluicing is that it allows 'pied-piping', where the missing constituent is immediately preceded by a *wh*-element as well as a preposition, as in (7a). However, sluicing in Arabic does not allow stranded prepositions as indicated by the ungrammaticality of (7b), where the gap is introduced by a preposition rather than a *wh*-complementizer.

(7) a. *raḥalat mūnā, lākin lā ʾa-ʿrif-u*
left.3fs Mona but NEG 1s-know-ind
ʾilā ʾayna (Standard Arabic)
to where
'Mona left, but I don't know to where'
b. * *rahalat mūnā, lākin lāʾ aʿrifu*
left.3fs Mona but NEG 1s-know-ind.
ʾayna ʾilā (Standard Arabic)
where to
'Mona left, but I don't know where to'

Ellipsis within a noun phrase is similar to sluicing in that there is a particular class of linguistic forms that signal the syntactic nature of the elided structure. For example, sluicing is marked by a *wh*- word immediately preceding the missing constituent. In the case of ellipsis within a noun phrase, specifiers (e.g., quantifiers and demonstratives) immediately precede the missing constituent. A distinctive property of ellipsis within a noun phrase is that the missing constituent is not a complete phrase, but only the head noun and its modifiers. The examples in (8a) and (8b) demonstrate that quantifiers in Standard Arabic introduce ellipsis within noun phrases provided that the quantifier is marked for → *tanwīn*, which includes the case marking corresponding to the elided head noun as well as the suffix *-n*. This constraint does not apply to cases where the specifier is a demonstrative as in (8c).

(8) a. *zurtu ʾaṣdiqāʾ-ī wa-ʾaʿṭaytu*
visited.1s friends-my and-gave.1s
kull-an hadiyy-a
every-acc. gift-f.
'I visited my friends, and gave each a gift'
b. *qaraʾtu ṣuḥuf aṣ-ṣabāḥ*
read.1s newspapers the-morning
ġayra baʿḍ-in
except some-gen.
'I read the morning newspapers except for some'
c. *qābaltu ṣ-ṣaḥafiyyīn ʿadā*
met.1s the-journalists except
hāʾulāʾi llaḏīna kānū fī l-ijtimāʿ
those who.pl. were in the-meeting
'I met the journalists except for those who were at the meeting'

Missing constituents following universal quantifiers, as in (8a) and (8b), are always interpreted as singular nouns even when the antecedent, if there is one, is plural. Egyptian Arabic, on the other hand, does not allow ellipsis within a noun phrase if the specifier is a universal quantifier, as indicated by the ungrammaticality of (9a) and (9b) compared to the grammatical sentence in (9c), where the quantifier is an existential one. Moreover, not all specifiers allow ellipsis, as it is ungrammatical with *muʿẓam* 'most' in Egyptian and Standard Arabic.

(9) a. * *ʾābilt iḍ-ḍiyūf wi-ʾaʿadt-ə*
met.1s the-guests and-sat.1s
maʿa kull-ə šwayya
with every bit
'I met the guests and sat with each for a bit'
b. * *kull-ə mašġūl fi šuġl-u*
every busy in work-his
'Everyone is busy with his work'
c. *raggaʿt ik-kutub li-l-maktaba*
returned.1s the-books to-the-library
bass-ə xallēt šuwayya f-il-bēt
but kept.1s some in-the-house
'I returned the books to the library, but I kept some at home'

Verb phrase ellipsis is similar to ellipsis within a noun phrase in that both operate on intermediate projections rather than complete phrases. Therefore, the elided structure obligatorily includes the head verb as well as its internal object argument, whereas adverbials are optionally deleted. Moreover, verb phrase ellipsis is allowed only after the auxiliary verb *kān* 'to be' (Kortobi 2002). For example, the ellipsis in the second conjunct in the Moroccan Arabic sentence in (10) involves the deletion of the head verb as well as its complement.

(10) *yasin kān ka-yəlˁəb*
Yasin was imperf.-play.3ms
l-kōra w-yosre kān ḥətta huwa
the-ball and-Yosre was even him
'Yasin was playing football, and Yosre was, too'

Although the different types of ellipsis described above vary with regard to the syntactic structure of the elided strings, they have certain common features. They all operate on intermediate projections, and there are particular classes of linguistic forms that signal the nature of missing structure. Finally, despite the fact that the three types of ellipsis are syntactically constrained, the interpretation of the missing structure is a pragmatic process, since there is not always a one-to-one correspondence between the elided structure and its antecedent, if one is present.

Bibliographical references

Baltin, Mark. 2003. "The interaction of ellipsis and binding: Implications for sequencing of Principle A". *Natural Language and Linguistic Theory* 21.215–246.

Barton, Ellen. 1990. *Nonsentential constituents: A theory of grammatical structure and pragmatic interpretation*. Amsterdam and Philadelphia: J. Benjamins.

El-Shiyab, Said. 1998. "Ellipsis in Arabic and its impact on translation". *al-ˁArabiyya* 31.39–54.

Green, Georgia M. 1996. *Pragmatics and natural language understanding*. Mahwah, N.J.: Erlbaum.

Grice, Paul. 1975. "Logic and conversation". *Syntax and semantics*, III, ed. Peter Cole and Jerry Morgan, 41–58. New York: Academic Press.

Hardt, Daniel. 1999. "Dynamic interpretation of verb phrase ellipsis". *Linguistics and Philosophy* 22.185–219.

Kortobi, Ibtissam. 2002. "Gapping and VP-deletion in Moroccan Arabic". *Themes in Arabic and Hebrew syntax*, ed. Jamal Ouhalla and Ur Shlonsky, 217–240. Dordrecht: Kluwer Academic Publishers.

Lebock, Anne. 1995. *Ellipsis*. Oxford: Oxford University Press.

Levinson, Stephen. 2000. *Presumptive meaning: The theory of generalized conversational implicature*. Cambridge: MIT Press.

Sag, Ivan and Thomas Wasow. 1999. *Syntactic theory: A formal introduction*. Stanford: Center for the Study of Language and Information.

Mustafa Mughazy (Western Michigan University)

Emphasis → Velarization; ʾIṭbāq; Tafxīm

Enclisis

Clitic elements (→ clitics) are those which lack an inherent stress and are therefore found attached to an adjacent word. The value of the term 'enclitic' varies, some authors employing it to refer specifically to an element which follows the element with which it is accentually linked (and hence as a synonym of what is otherwise known as a 'postclitic'), while others use 'enclitic' more broadly to refer to any accentually dependent element, regardless of the linear relation to its accentual host – in the second sense, 'enclitic' may refer to either proclitics or postclitics.

Since Arabic orthography provides no indication of the suprasegmental features of the early language, and since the classical grammarians have left no systematic discussion of the accentuation of the language which they were analyzing, we have no direct contemporary data on the early Arabic accent. Nevertheless, the graphic conventions of written Arabic suggest that clitics existed in the early language, and a sequence of two or more words written without an intervening word boundary has routinely been taken to contain one or more clitics.

The literary Arabic clitics form two discrete sets, a proclitic set composed of prefixes and a postclitic set composed of suffixes. The proclitic set contains a subset of the prepositions, the majority of the conjunctions, and what may be called adverbials, while the postclitic set consists of pronominals. The great majority of the clitics are monomorphemic, and a considerable number are monosyllabic. The clitics are located in strictly defined sites, being attached either to the first element of the sentence or to the syntactic head of the phrase in which they are located.

i. Proclitics

Prepositions: *li-* 'to, for', *bi-* 'at, in, by'; *ka-* 'like' differs from these in that a following pronoun is in the independent shape rather than the suffixed shape (*l-ī* 'to-me', *b-ī* 'in-me', but *ka-ʾana* 'like-me').

Conjunctions: *wa-* 'and' (linking words or clauses), *fa-* 'and (then)' (linking clauses), *li-* 'in order that' (preceding a subordinate clause containing a subjunctive verb).

Adverbials: *ʾa-* (a marker of an interrogative clause), *la-* (the → asseverative particle), *sa-* (a preverbal future-tense marker, the clitic counterpart to *sawfa*), *li-* (the preverbal optative marker).

ii. Postclitic pronominals

The two pronominal-suffix paradigms (accusative vs. genitive) are distinguished only in the 1st person singular (accusative *-nī* vs. genitive *-ī/-ya*), the remaining forms showing no case distinctions: 2nd pers. masc. sg. *-ka*, 2nd pers. fem. sg. *-ki*, 2nd pers. dual *-kumā*, 2nd pers. masc. pl. *-kum*, 2nd pers. fem. pl. *-kunna*, 3rd pers. masc. sg. *-hu*, 3rd pers. fem. sg. *-hā*, 3rd pers. dual *-humā*, 3rd pers. masc. pl. *-hum*, 3rd pers. fem. pl. *-hunna*. The genitive forms are affixed to nouns, prepositions, or the 'dummy' stem *ʾiyyā-* (e.g. *ʾiyyā-ka naʿbudu* 'Thee do we worship', Q. 1/5), while the accusative forms are attached to transitive verbs, or one of various sentence- or clause-initial particles (*ʾinna* 'verily', *ʾanna* 'that', *lākinna-* 'but', *layta-* 'would that…!' etc.); the accusative suffixes (like accusative-shape substantives in general) function as the equivalent of the nominative in certain syntactic situations (*lākinna-nī raʾaytu zaydan* 'but-I [lit. 'me'] saw [1st pers. sg.] Zayd', in contrast to *ʾana raʾaytu zaydan* 'I saw Zayd').

Clitic chains of modest length may be constructed by linking to a tonic word a series of proclitics (*wa-li-zaydin* 'and-to-Zayd', *ʾa-fa-lā* 'so isn't it the case that…?') or postclitics (*ʾaʿṭi-nā-hu* 'give-us-it'), but a sequence composed of a member of each of the clitic sets is rendered as an independent graphic unit (*la-ka* 'to-you'). Certain clitics acquire contextually conditioned alternate shapes as the result of specific morphophonemic processes: the *-u-* of the 3rd-person pronominal suffixes is fronted to *-i-* after a syllable containing *-i-* or *-ī-* (*bi-hi*, *bi-himā*, *bi-him*), the preposition *li-* assumes the shape *la-* before a suffixed pronoun (*li-zaydin* 'to-Zayd', but *la-hu* 'to-him'), and the optative *li-* routinely loses its vowel when preceded by a conjunction (*li-yaqul* 'may he say', but *wa-l-yaqul*). Certain of the postclitic pronouns have preserved a historically underlying long *-ū-* when they are followed by a second suffixed pronoun (*ʾarā-hū-hum* '(he) showed-him-them', *ʾarā-kumū-hum* '(he) showed-you [masc. pl.] -them').

Among the modern Arabic dialects clitics continue to play an important role. New preverbal particles marking aspect, tense, and modality have arisen across the dialects (e.g. Cairo *b-yiktib* 'he writes [indicative]', *ḥa-yiktib* 'he will write'), and several dialects have developed a new set of indirect-object clitics affixed to the end of the verbal complex, e.g. Damascus Arabic *žəbt-əlli* '(you) brought for me', Cairene *ʾult-ulha* '(I) said to her', *gab-hā-lak* '(he) brought her for you'.

David Testen (Reston, VI, U.S.A.)

Energicus

Energicus/energic/energetic and *an-nūn al-muʾakkida/nūn at-tawkīd* (*al-xafīfa wa-ṯ-ṯaqīla*) in Arabic are parallel names for an optional ending of either single or geminate *-n-* which is occasionally suffixed to certain Semitic verb conjugations, particles, and prepositions. In Arabic, the energicus appears mostly in Classical Arabic and is found in many Qurʾānic passages (Wright 1896:61, 1898:24; Brockelmann 1908: 554–555, 1913:159; Reckendorf 1921:16; Fleisch 1979:128–132, 140–141; Fischer 2002: 110, 118, 120, 137, 230; Ambros 1989; Zewi 1999:13–63). The Arabic *-n(n)-* ending is suffixed to the prefix conjugation and to the imperative with a connecting vowel *-a-*. The type of Arabic connecting vowel might suggest that the energicus is affixed in Arabic to the subjunctive, which possibly evolved from an ancient Semitic volitive mood ending in an *-a* vowel. On the origins of the Arabic subjunctive, relating it to an ancient Semitic volitive, see, e.g., Fleisch (1968), but see observations in Blau (1971:144–146), and note the doubts raised by Rainey (1991–1993) regarding the existence of a volitive mood in El-Amarna Akkadian; likewise Testen (1994:158). The connecting vowel of the energicus might also support its comparison with the Hebrew cohortative, especially in its pausal form, which includes a long vowel instead of an *-n-*. On this pausal form see below. On the comparison of the energicus with the Hebrew cohortative see Wright (1890:194). Arguments against it are, e.g., in Blau (1971:135). Also note Testen (1994), who stretches this possibility further by suggesting that the Arabic subjunctive might have evolved from the energicus after elimination of its *-n*-ending, and Zaborski (1996), who regards all modal prefix conjugations forms as Proto-Semitic innovations (→ Afro-Asiatic languages). Wright (1896:61) asserts that the Arabic energicus is added to the jussive, but treats it as an independent mood in his syntactic treatment (Wright 1898:41–43). Nevertheless, semantically, the Arabic energicus might be more related to the indicative sphere, since it mostly expresses modal nuances related to the future, and its modal marking is usually stronger than the one expressed by jussive modal forms. Yet, it might appear in parallel to both jussive and indicative (Wright 1898:24; Zewi 1999:187–192).

Possible Semitic cognates or remnants of cognates to the Arabic energicus appear in several Semitic languages and dialects, i.e. Biblical Hebrew, Phoenician, Old and Imperial Aramaic, the Akkadian of El-Amarna, Taanach, and Kāmid El-Lōz, Ugaritic, and Ancient South Arabian. Some or all of these languages are indicated by, e.g., Wright (1890: 193–194), Lambert (1903), Brockelmann (1908: 554–559), Moscati (1964:135–136), Hetzron (1969), Williams (1972), Muraoka (1975), Rainey (1975, 1986:10–12, 1996:234–244), Blau (1978), Bennett (1984:37–51, 97–102, 143–144, 198), Huehnergard (1988), Testen (1993), Krebernik (1993), Fassberg (1994:63–70), Sivan (1997:98–99, 102–103, 105–106), Zewi (1999), and Lipiński (2001:317, 362–363, 460–461). The connective vowel between the verb forms and the *-n(n)-* endings in these languages varies. Moreover, the exact function of the *-n(n)-* endings attested in these languages is not always clear. In certain Semitic languages and dialects it expresses modality while in others it functions as a mere stylistic variant. Furthermore, Barth (1907:1–10, 1913: 34), who presents a broad variety of *-n(n)-* endings affixed to several verb conjugations other than the prefix conjugation, to infinitives, and to certain particles and prepositions, considers these endings, on account of their variety, dissimilar to the Arabic energicus.

In any case, the function of the energicus in those Semitic languages in which it indeed exhibits a special nuance, including Arabic, is generally regarded as strengthening or emphatic. More precisely, while the energicus is suffixed to prefix conjugation verbs, it is deemed to express modality, i.e. it adds to the verb a nuance of subjective emotional involvement of the speaker, expressing intent, oath, self-encouragement, promise, wish, command, prohibition, threat, warning, affirmation, etc. Energicus forms also appear in several Semitic languages in questions. Arabic instances in all typical contexts are collected in, e.g., Zewi (1999:16–61, 59–61 for prefix verbs accompanied by energicus forms in questions). Classical Arabic instances of prefix conjugation verbs accompanied by an energicus in typical contexts are, e.g., prophecy expressing God's declaration of intent: *la-ʾamlaʾanna jahannama min al-jinnati wa-n-nāsi ʾajmaʿīna* 'I will fill Hell with jinns and men all together' (Q. 11/119, translated by Yusuf Ali 1987), prohibition: *wa-lā taqūlanna li-šayʾin ʾinnī fāʿilun ḏālika ġadan* 'Nor say of anything "I shall be sure to do so and so tomorrow"' (Q. 18/23, translated by Yusuf Ali 1987), an oath containing a conditional: *wa-ʾaqsamū bi-llāhi jahda ʾaymāni-him laʾin ʾamarta-hum la-yaxrujunna* 'and they have sworn by God the most earnest oaths, if you command them they will go forth' (Q. 24/53, translated by Arberry 1955), and a question: *hal yuḏhibanna kayduhu mā yaġīḏ̣u* 'whether his plan will remove that which enrages [him]' (Q. 22/15, translated by Yusuf Ali 1987).

In later stages of Arabic the energicus is rarely used. Hopkins's grammar of Early Arabic (1984), which treats papyri earlier than 912 C.E., mentions only a few exceptions to the general lack of energicus forms in his data, namely the official correspondence of the Aphrodito archive, where the energicus appears regularly (Hopkins 1984:70–71), and some instances of the energicus in the apodosis of certain conditionals (Hopkins 1984:253). A few energicus forms are attested in Saadya Gaon's post-Classical Arabic translation of the Pentateuch (Zewi 2001). Blau (1967, 1980) does not mention the energicus in his grammars of Christian Arabic and Medieval Judaeo-Arabic at all. The energicus does not exist in modern Arabic dialects. The evidence of *-n(n)-* endings affixed to participles in several Arabic dialects probably does not represent an energicus but pronominal elements (Retsö 1988).

The origin of the energicus form is usually considered to be demonstrative, and it is viewed as similar to the *-n-* found in various demonstratives and pronouns, e.g. Hebrew *hen*, *hinnē*, and Arabic *ʾin*, *ʾinna*, and *ʾanna*, e.g. Wright (1890:193) and Barth (1907:7, 1913:34). Other scholars have tried to related its origins to the Hebrew particle *nā* (e.g. Wright 1890:193–194; Fassberg 1994:63, 73).

Paradigms of prefix conjugation verbs and imperatives with the Arabic energicus are found in Tables 1 and 2. The long vowels in the prefix conjugation forms of 2nd pers. fem. sg. (*taqtulīna*), of 2nd and 3rd pers. masc. pl. (*taqtulūna/yaqtulūna*), of imperative 2nd pers. fem. sg. (*uqtulī*), and of 2nd pers. masc. pl. (*uqtulū*) are shortened because the syllable is closed by the energicus consonant *-n-*. The connective *-a-* vowel, which usually appears between the verb and the energicus *-n(n)-* ending, disappears in these forms. The combination of the energicus with III *w/y* verbs is complicated and requires modifications of semi-vowels (see Table 3). Also note that the energicus is occasionally written with a *tanwīn* instead of an *-n-*, and it can also take a pausal form in which the *-n-* becomes silent and the preceding vowel is lengthened (*an* > *ā*, Wright 1890:194, 1896:61; Fischer 2002:8, note 2).

Table 1. Prefix conjugation + long and short energicus

singular	plural	dual
1st pers.		
ʾaqtulanna/	*naqtulanna/*	
ʾaqtulan	*naqtulan*	
2nd pers. masc.		
taqtulanna/	*taqtulunna/*	*taqtulānni*
taqtulan	*taqtulun*	
2nd pers. fem.		
taqtulinna/taqtulin	*taqtulnānni*	
3rd pers. masc.		
yaqtulanna/	*yaqtulunna/*	*yaqtulānni*
yaqtulan	*yaqtulun*	
3rd pers. fem.		
taqtulanna/	*yaqtulnānni*	
taqtulan		

Table 2. Imperative + long and short energicus

singular	plural	dual
masc.		
uqtulanna/	*uqtulunna/*	*uqtulānni*
uqtulan	*uqtulun*	
fem.		
uqtulinna/	*uqtulnānni*	
uqtulin		

Table 3. III *w/y* verbs Form I: Prefix conjugation + energicus

singular	plural	dual
1st pers.		
ʾarmiyanna/	*narmiyanna/*	
ʾadʿuwanna/	*nadʿuwanna/*	
ʾalqayanna	*nalqayanna*	
2nd pers. masc.		
tarmiyanna/	*tarmunna/*	*tarmiyānni*
tadʿuwanna/	*tadʿunna/*	
talqayanna	*talqawunna*	
2nd pers. fem.		
tarminna/tadʿinna/	*tarmīnānni/*	
talqayinna	*tadʿūnānni/*	
	talqaynānni	
3rd pers. masc.		
yarmiyanna/	*yarmunna/*	*yarmiyānni*
yadʿuwanna/	*yadʿunna/*	
yalqayanna	*yalqawunna*	
3rd pers. fem.		
tarmiyanna/	*yarmīnānni/*	
tadʿuwanna/	*yadʿūnānni/*	
talqayanna	*yalqaynānni*	

BIBLIOGRAPHICAL REFERENCES

Ambros, von Arne. A. 1989. "Syntaktische und stilistische Funktionen des Energikus im Koran". *Wiener Zeitschrift für die Kunde des Morgenlandes* 79.35–56.

Arberry, Arthur J. 1955. *The Koran interpreted*. London: Allen and Unwin, New York: Macmillan.

Barth, Jacob. 1907. *Sprachwissenschaftliche Untersuchungen zum Semitischen*, I. Leipzig: J.C. Hinrichs.

——. 1913. Die *Pronominalbildung in den semitischen Sprachen*. Leipzig: J.C. Hinrichs. (Repr. Hildesheim: G. Olms, 1967.)

Bennett, Stephan, F. 1984. *Objective pronominal suffixes in Aramaic*. Ph.D. diss., Yale University. (Repr. Ann Arbor, Mich.: University Microfilms International, 1985.)

Blau, Joshua. 1967. *A grammar of Christian Arabic*. Louvain: Peeters.

——. 1971. “Studies in Hebrew verb formation”. *Hebrew Union College Annual* 42.133–146.

——. 1978. “Pronominal third person singular suffixes with and without N in Biblical Hebrew”. *Eretz Israel* 14.125–131. [In Hebrew.]

——. 1980. *A grammar of mediaeval Judaeo-Arabic*. Jerusalem: Magnes Press. 2nd ed. [In Hebrew.]

Brockelmann, Carl. 1908. *Grundriss der vergleichenden Grammatik der semitischen Sprachen*, I. Berlin: Reuther und Reichard.

Fassberg, Steven. E. 1994. *Studies in Biblical syntax*. Jerusalem: Magnes Press. [In Hebrew.]

Fischer, Wolfdietrich. 2002. *A Grammar of classical Arabic*. New Haven, Conn. and London: Yale University Press. 3rd rev. ed. trans. Jonathan Rodgers.

Fleisch, Henri. 1968. “*Yaqtula* canaéen et subjonctif arabe”. *Wissenschaftliche Zeitschrift der Martin-Luther-Universität Halle-Wittenberg* 17.65–76.

——. 1979, *Traité de philologie arabe*, II. Beirut: Dar El-Machreq.

Hetzron, R. 1969. “Third person singular pronoun suffixes in Proto-Semitic”. *Orientalia Suecana* 18.101–127.

Hopkins, Simon. 1984. *Studies in the grammar of early Arabic*. Oxford: Oxford University Press.

Huehnergard, John. 1988. “The early Hebrew prefix-conjugations”. *Hebrew Studies* 29.19–23.

Krebernik, Manfred. 1993. “Verbalformen mit suffigierten n-Morphemen im Ugaritischen: Überlegungen zur Morphologie des Energikus im Ugaritischen und in anderen semitischen Sprachen”. *Syntax und Text: Beiträge zur 22. Internationalen Ökumenischen Hebräisch-Dozenten-Konferenz 1993 in Bamberg*, ed. Hubert Irsigler, 123–150. St. Ottilien: EOS.

Lambert, Mayer. 1903. “De l’emploi des suffixes pronominaux avec *noun* et sans *noun* au futur et à l’impératif”. *Revue des Études Juives* 46.178–183.

Lipiński, Edward. 2001. *Semitic languages: Outline of a comparative grammar*. 2nd ed. Leuven: Peeters.

Moscati, Sabatino. 1964. *An introduction to the comparative grammar of the Semitic languages*. Wiesbaden: O. Harrassowitz.

Muraoka, Takamitsu. 1975. “The *nun energicum* and the prefix conjugation in Biblical Hebrew”. *Annual of the Japanese Biblical Institute* 1.63–71.

Rainey, Anson. F. 1975. Review of J.W. Wevers and D.B. Redford (eds.), *Studies on the ancient Palestinian world presented to Professor F.V. Winnett on the occasion of his retirement, 1 July 1971*, Toronto: University of Toronto Press, 1972. *Israel Exploration Journal* 25.184–187.

——. 1986. “The ancient Hebrew prefix conjugation in the light of Amarnah Canaanite”. *Hebrew Studies* 27.4–19.

——. 1991–1993. “Is there really a *yaqtula* conjugation pattern in the Canaanite Amarna tablets?” *Journal of Cuneiform Studies* 43–45.107–117.

——. 1996. *Canaanite in the Amarna tablets*, II. Leiden: E.J. Brill.

Reckendorf, Hermann. 1921. *Arabische Syntax*. Heidelberg: C. Winter.

Retsö, Jan. 1988. “Pronominal suffixes with *-n(n)*- in Arabic dialects and other Semitic languages”. *Zeitschrift für Arabische Linguistik* 18.78–94.

Sivan, Daniel. 1997. *A grammar of the Ugaritic language*. Leiden: E.J. Brill.

Testen, David. 1993. “On the development of the energic suffixes”. *Perspectives on Arabic linguistics*, V, ed. Mushira Eid and Clive Holes, 293–311. Amsterdam: J. Benjamins.

——. 1994. “On the development of the Arabic subjunctive”. *Perspectives on Arabic linguistics*, VI, ed. Mushira Eid, Vicente Cantarino, and Keith Walters, 151–166. Amsterdam: J. Benjamins.

Williams, R.J. 1972. “Energic verbal forms in Hebrew”. *Studies on the ancient Palestinian world presented to Professor F.V. Winnett on the occasion of his retirement, 1 July 1971*, ed. John. W. Wevers and Donald. B. Redford, 75–85. Toronto: University of Toronto Press.

Wright, William. 1890. *Lectures on the comparative grammar of the Semitic languages*. Cambridge: Cambridge University Press.

——. 1896, 1898. *A grammar of the Arabic language*. 2 vols. 3rd ed. Cambridge: Cambridge University Press.

Yusuf Ali, Abdullah. 1987. *The Holy Qur’an: Text, translation and commentary*. U.S. ed. Elmhurst, N.Y.: Tahrike Tarsile Qur’an.

Zaborski, Andrzej. 1996. “On the origin of subjunctive and energicus in Semitic”. *Incontri Linguistici* 19.69–76.

Zewi, Tamar. 1999. *A syntactical study of verbal forms affixed by -n(n) endings in Classical Arabic, Biblical Hebrew, El-Amarna Akkadian and Ugaritic*. Münster: Ugarit-Verlag.

——. 2001. “Energicus in Saadya Gaon’s translation of the Pentateuch”. *New data and new methods in Afroasiatic linguistics: Robert Hetzron in memoriam*, ed. Andrzej Zaborski, 214–221. Wiesbaden: O. Harrassowitz.

TAMAR ZEWI (University of Haifa)

English

English may be considered a typical case of a European language indebted to Arabic (often through the intermediary of another language). Many loanwords, direct or indirect, are in the semantic spheres of astronomy (Pei 1967:225 states that 125 out of 183 star names are from Arabic, with 9 more coming from Arabic via Latin), chemistry, agriculture, clothing, commerce, mathematics, military science, the realm of Islam, and so forth.

Arabic is well known as an international language, and Islam’s holy book, the *Qur’ān*, has spread from the western part of the Arabian Peninsula all over the world. Along with the development of Islam, Arabic came in contact with the many local languages of a conquered area. With cultural contact there is, of course, linguistic contact, one of the most important results of which is the phenomenon of linguistic borrowing. In addition to Berber, Arabic has become a major supplier of vocabulary to Swahili in East Africa, Spanish and Portuguese in the Iberian Peninsula, Persian, Turkish, Urdu, Uzbek, Uyghur in the People's Republic of China, Bahasa Melayu and Bahasa Indonesia throughout much of Asia, and others.

Of course, other European languages served as transmitting devices for numerous Arabic loanwords: for example Spanish, Portuguese, and Italian, either directly or indirectly, especially via the Indian subcontinent. In the cases of Spanish and Portuguese, it is easy to understand the impact of Arabic on each because the Moors ruled the Iberian Peninsula for about 800 years. The Arabs also controlled Sicily for over 200 years. Thus, one can appreciate the impact of Arabic on the various Romance languages spoken in that area of hegemony. Since France ruled much of the Middle East and North Africa, it was only natural for numerous Arabic loanwords to penetrate French. During the Crusades, French was the official language in England, and many Arabic loanwords thus surfaced in English from French. Two examples of Arabic loanwords coming into English through the intermediary of French are: *magazine* 'periodical' < French *magasin* < Old Provencal < Arabic *maxāzin*, the plural of *maxzan* 'storehouse, warehouse'; and *admiral* 'the highest ranking officer in the navy, equivalent to a general in the army' < Medieval Latin *admiralis* and Middle French *amiral* < Arabic *ʾamīr al-baḥr* 'commander of the sea', where the last word *baḥr* has been deleted due to the process of clipping (cf. English *deli* < *delicatessen*).

The most common Arabic loanwords in English, based on Cannon (1994), will now be presented. From the realm of anatomy, two loanwords are particularly common. They are: *tripe* 'the wall of a ruminant's stomach, prepared as food; something or someone worthless' < Italian *trippa* probably < Arabic *ṯarb* 'thin layer of fat lining the intestines'; *carcass* 'a corpse of a human or animal' < Middle French *carcasse* < Middle Latin *tarcasius* < Arabic *tarkāš* 'arrow bearing'.

Most of the over 100 astronomical items are little used outside technical terminology, for instance *azimuth* 'a measured arc of the horizon' < Old French *azimat* < Arabic *as-sumūt* 'the azimuths or directions' and *zenith* 'point of heavens directly above observer' < Old French *cenit* or Middle Latin *cenit* < Arabic *samt (ar-raʾs)* 'way above the head'. Over 80 star names came directly from Arabic into English, including: *Dub(b)he* < Arabic *ad-dubb al-ʾakbar* 'the Greater Bear'; the final *-e* is indicative of a feminine in a dialect with vowel raising (*ʾimāla*); *Duhr* < Arabic *ḏ̣ahr al-ʾasad* 'the lion's back'; *Alula Borealis* and *Alula Australis* < Arabic *al-firqa al-ʾūlā* 'the first joint' and Latin *australis* 'southern' < Arabic *al-qafza al-ʾūlā* 'the first leap' and Latin *borealis* 'northern'.

The great majority of the 268 botanical items from Arabic have been part of English for a long time. Among fruits and vegetables (many of which are international), there are *apricot* 'an orange fruit resembling the plum and peach in flavor' < Middle French *abricot* < Arabic *al-barqūq* 'the plum' (itself from Greek *praikokkion* < Latin *praecox*); *artichoke* 'a tall herb resembling a thistle < Italian (dial.) *articiocco* < Arabic *al-xaršūf* 'the artichoke'; *aubergine* 'the fruit of the aubergine, eggplant' < French diminutive of *auberge* < Catalan *alberginia* < Arabic *al-bāḏinjān* 'the aubergine'; *endive* 'a widely cultivated salad plant' < Middle French < Late Latin *endivia* < Latin *intubus* possibly < Arabic *hindab* 'endive'; *lemon* 'an acid fruit containing fragrant lemon oil and often candied or preserved' < Middle French *limon* < Middle Latin < Arabic *laymūn*; *lime* 'the fruit of the lime tree' < French *lime* 'fruit' < Spanish *lima* < Arabic *līma* (sg.), *līm* (coll.) 'citrus fruit'; *orange* 'any of various citrus fruits' < Old French *orenge* < Old Provencal *auranja* < Arabic *nāranj*; *spinach* 'an annual potherb, widely cultivated for its edible leaves' < Middle French *espinache* < Old Spanish *espinaca* and Middle Latin *spinachia* < Arabic *ʾisfānax*; *tangerine* 'one of various cultivated citrus fruits, as a Tangerine orange' < French *Tanger* < Arabic *ṭanja* 'the name of a Moroccan port'.

The following plant-related words have spread internationally: *alfalfa* 'an important forage plant, used as hay; also called *lucerne*' < Spanish < dial. Spanish Arabic *al-faṣfaṣa* 'the alfalfa'; *attar* 'a fragrant oil obtained from rose petals' < Persian < Arabic *ʿiṭr* 'perfume, essence'; *balm* 'an aromatic resinous substance prized for its fragrance and healing powers' < Old French *basme* < Latin *balsamum* probably < Arabic and Hebrew *bāšām* 'spice'; *benzoin* 'a balsamic resin obtained from Southeast Asian trees and used in perfumes, incense, and skin treatment' < Middle French *benjoin* < Old Catalan *benjui* < Arabic *lubān jāwī* 'frankincense of Java'; *calabash* 'gourd, especially the common bottle gourd' < French *calabasse* < Spanish *calabaza* probably < Arabic *qarʿa yābisa* 'dry gourd'; *cane* 'a hollow or jointed stem, used as a walking stick or for flogging' < Middle French < Old Provençal *cana* < Latin *canna* < Greek *kanna* < Semitic, as Arabic *qanā* 'hollow stick or reed'; *caraway* 'a biennial, usually white-flowered herb' probably < Middle Latin *carvi* < Arabic *karawiyā* 'a white-flowered herb, caraway seed'; *cork* 'the bark of the cork oak, as used for stoppers and insulation' probably < Arabic *qurq*; *henna* 'a dye, liquid, powder or paste made from *henna*, used in ceremonies' < Arabic *ḥinnāʾ*; *jasmine* 'any of numerous climbing shrubs with extremely fragrant flowers' < French *jasmin* < Arabic *yāsa(a)mīn*; *lilac* 'a plant of the Syringa genus cultivated for its fragrant flowers' < Arabic *laylak* ~ *līlak*; *safflower* 'the dried petals of *Carthamus tinctorious* or the red dye obtained from it' < Middle French *saf(f)leur* < Old Italian *saffiore*, *zaffrole* < Arabic *ʾaṣfar* yellow [plant]'; *sandal(wood)* 'sandalwood' < Middle French < Middle Latin *sandalum* < Late Greek *sandanon* probably < Arabic *ṣandal* 'sandalwood'; *sarsaparilla* 'a plant of the Smilax genus indigenous to tropical America' < Spanish *zarzaparilla* < *zarza* < Arabic *šaraṣ* 'bush'; *sumac* 'material from a shrub or tree of the Rhus genus used in tanning and dyeing' < Middle French < Arabic *summāq* 'material from a shrub or tree of the Rhus genus'; *tamarind* 'the fruit of the tamarind tree' < Middle Latin *tamarindus* < Arabic *tamar hindī* 'Indian date'; *simsim* 'sesame' < Swahili *simsim* < Arabic *simsim* 'sesame'; *melongena* 'a West Indian name for aubergine, eggplant' < Italian (dial.) *melongiana* < Arabic *bāḏinjān* 'aubergine'.

Turning to the world of birds, 136 words have been recorded, 2 of which are: *saker(et)* 'a large falcon' < Middle French *sacre* < Arabic *ṣaqr* 'falcon'; *albatross* 'a large seabird' < Spanish and Portuguese *alcatraz* 'pelican' probably < Arabic *al-ġaṭṭās* 'the white-tailed sea eagle', lit. 'the diver'.

It is surprising that there are only 29 names of fish which have been borrowed. Only 4 have been directly borrowed from Arabic; the others penetrated via a Romance language. Among the most important are the following: *bolti* 'a cichlid food fish' < Arabic *bulṭi* 'a Nile fish'; *albacore* 'a large pelagic fish of the family Thunidae' < Portuguese *albacor* < Arabic *al-bakūra* 'albacore'; *bonito* 'any of several medium-sized scromboid fishes' < Spanish *bonito* 'beautiful' but possibly < Arabic *baynīṯ* ~ *binnī* 'a Nile fish'; *tuna* 'any of numerous large fish for sport and food' < Spanish *atún* < Arabic *at-tūn*.

There are 82 zoological items. The following are the most common: *giraffe* 'a fleet African ruminant mammal' < Italian *giraffa* < Arabic (dial.) *zirāfa* 'giraffe'; *gazelle* 'a small antelope in Asia and Africa' < Arabic *ġazāl* 'wild goat'; *jerboa* 'any of various small nocturnal rodents inhabiting desert areas of the Old World' < Arabic *yarbūʿ* ~ *jarbūʿ* 'jerboa'; *monkey* 'a member of a primate order excepting humans, and various transferred and figurative uses' probably < Spanish and/or French *mona* 'ape' possibly < Arabic *maymūn* 'ape, monkey'; *popinjay* 'a shade of green or a green parrot's color; a green woodpecker [British dial.]' < Middle French *papejai* < Arabic *babaġā* 'parrot'; *Saluki* 'an old breed of hunting dog formerly called *Persian greyhound*' < Arabic *salūqī* 'of Saluq, an ancient city in Yemen'; *tabby* 'a type of cat, so named because of its striped coat as in the original *tabby taffeta*' < French *tabis* < Middle Latin *attabi* < Arabic *al-ʿattābiyya* 'the name of the Baghdad quarter where this material was originally made'.

Arab Spain was probably the center of the scientific world during its heyday. Thus, 120 chemical terms came into English. The following are the most common: *acetal* 'a colorless, alcohol-smelling liquid used as a solvent' < *acet-* + *alcohol* (see below); *alchemy* 'the medieval science of trying to transfer base metals into gold and of seeking cures for diseases' < Middle French or Middle Latin *alquemie* < Arabic *al-kīmiyāʾ* 'the philosopher's stone' (itself < Greek *khumos*); *alcohol* 'a colorless, volatile, flammable liquid' < Middle Latin < Old Spanish < Arabic *al-kuḥūl* 'the powdered antimony'; *alkali* 'a soluble salt obtained from plant ashes' < Middle Latin *alcali* ~ *alkali* < Arabic *al-qilī* 'the ashes of the saltwort plant'; *amalgam* 'an alloy of mercury with another metal such as gold or copper' < Middle Latin *amalgama* < Arabic *al-malġam* 'an alloy of mercury with another metal such as gold or copper'; *antimony* 'a metalloid element used especially in alloys' < Middle Latin *antimonium*, < possibly Arabic *al-ʾiṯmid* 'the name of the native trisulfide (gray antimony) or stibnite'; *benzine* 'one of various flammable petroleum distillates used especially in solvents or fuels' < Arabic *benzīn* 'benzine'; *borax* 'the best known sodium borate as used in various commercial products' < Middle Latin < Arabic *būraq* 'sodium borate'; *sugar* 'a sweet, primary sucrose substance important in human food' < Middle French *sucre* < Arabic *sukkar* 'sugar'; *tartar* 'a substance that is essentially cream tartar' < Middle French *tartar*, possibly < Arabic *durdī* 'sediment, dregs'; *zirconium* ~ *zircon* ~ *jargo(o)n* 'a colorless, pale yellow, or smoky zircon' < French *jargon* < Italian *giargone* < Arabic *zarqūn* 'bright red'.

There are 48 geological items. Among the most common are: *azure* 'the color of the clear sky' < Old French *azur* probably < Old Spanish *azur* ~ *azul* < Arabic *lāzulaward* 'lapis lazuli'; *coral* 'a skeletal deposit in reefs' < Latin *corall(i)um* < Greek *korallion*, probably < Semitic, as Hebrew *gōrāl* 'pebble' and/or Arabic *garal* 'small stone'; *lapis lazuli* 'a semiprecious stone that is a lazurite with a bright blue color' < Latin *lapis* 'stone' + Middle Latin *lazuli* < Arabic *lāzulaward* 'lapis lazuli'; *talc* 'a cosmetic' < Arabic *talq* 'mica'.

The Arabs were great pioneers in the field of mathematics. The following items are known by many: *algebra* 'a branch of mathematics' < Arabic *al-jabr* 'algebra (lit. 'breaking, solving of an equation)'; *algorithm* 'the system of Arabic numerals, arithmetic' < Arabic *al-xuwārizmī*, named for the 9th-century Persian mathematician al-Xuwārizmī; *cipher* 'zero; naught' < Middle French *cifre* < Middle Latin *cifra* < Arabic *ṣifr* 'empty, zero, cipher'; *sine* 'the mathematical *y* coordinate of a point with certain exceptions' < Middle Latin *sinus* 'the hanging fold of the upper part of a toga' < Latin *curve*, used as a translation of Arabic *jayb* 'sine, bosom of a garment'; *tariff* 'a schedule of rates, as for services, hotel room, train fare, etc.' < Italian *tariffa* < Arabic *taʿrīf(a)* 'information, definition'; *zero* 'the cipher symbol, denoting nought' < French and Italian < Middle Latin *zephirum* < Arabic *ṣifr* 'nothing, cipher'. One should note that *zero* and *cipher* are doublets, the former coming through Italian and the latter through Spanish.

Items relating to health have given English only one common word: *massage* 'the therapeutic manipulation of tissues by various means' < French *masser* 'to massage' < Arabic *massa* 'to stroke, strike'.

The political realm contributed numerous loanwords to English. The most generally widespread are the following: *alcalde* 'an administrative officer of a governmental agency in Spain' < Spanish *alcalde* 'mayor' < Arabic *al-qāḍī* 'judge'; *caliph* 'the title once used in Muslim countries for the chief civil and religious leader, as successor to Muḥammad to lead the Islamic community' < Middle French *calife* < Arabic *xalīfa* 'successor to the Prophet Muḥammad'.

Arabic music has given English the following common terms: *guitar* 'a flat-bodied string instrument of usually six strings' < French *guitar* < Old Spanish *guitarra* < Arabic *qītār* 'guitar'; *lute* 'a stringed musical instrument of Oriental origin' < Middle French *lut* < Old Provencal *laut* < Arabic *al-ʿūd* 'the oud'; *tambour* 'a drum, especially the bass drum' < French *tambour* 'drum' < Arabic *ṭambūr* 'tamboura'; *tambourine* 'a small drum' < Middle French *tambourin*, diminutive of *tambour* 'drum'.

The Arabs' interest in meteorology and climatology is well known. The following are some of the words borrowed in these domains: *monsoon* 'a wind that blows from one direction for part of the year, alternating with one that blows from the opposite direction, rainy season' < Dutch *monsoen* < Portuguese *monção* < Arabic *mawsim* 'season'; *sirocco* 'a hot, oppressive wind from the Libyan desert blowing into Italy, Sicily, etc.' < Italian *sirocco* < Arabic *šarq* 'east'.

In the realm of clothing and cloth, the following are fairly common terms: *chiffon* 'an ornamental addition to a woman's dress, as a knot of ribbons' < French *chiffe* 'old rag' possibly < Arabic *šiff* 'light garment'; *cotton* 'the soft, fibrous substance from the Gossypium plant' < Middle French *coton* < colloquial Arabic *quṭun* < Arabic *quṭn* 'cotton'; *gauze* 'a thin, often transparent woven fabric, used in surgical dressing' < Middle French *gaze*, probably < Arabic *qazz* 'raw silk'; *jacket* 'a male or female's coatlike garment for the upper body' < Middle French *jaquette* < Old French possibly < Old Spanish *jaco* < Old Catalan *jaco* < Arabic *šakk* 'mailcoat'; *mohair* 'any of various yarns or fabric using the hair of the Angora goat' < Italian *mocaiarro* (obsolete) < Arabic *muxayyar* 'choice; select'; *muslin* 'a varied kind of cotton fabric or garment of it' < French *mousseline* < Italian *mussolina* < Arabic *mawṣilī* 'of Mosul, Iraq, where it was formerly made'; *sash* 'a fine Oriental turban, or one who wears it' < Arabic *šāš* 'muslin'; *satin* 'a lustrous, sleek fabric, especially used in lingerie, dresses, and upholstery' < Middle French probably < Arabic (*ʾaṭlas*) *zaytūnī* '(silk) of Zaitun, a Chinese seaport praised by Marco Polo and usually identified as Tsinkiang'.

Arab cuisine is internationally acclaimed. Among the more common food and drink terms are the following: *rice* 'a cereal' < Spanish *arroz* 'rice' < Arabic *ar-ruzz* 'the rice'; *bulgur* 'a cereal food prepared from parched cracked wheat and eaten as a staple in Turkey and elsewhere' < Turkish < Arabic *burġul* 'cracked grain'; *felafel* 'sandwich eaten in some Arabic-speaking countries' < Arabic *falāfil*, pl. of *filfil* 'pepper'; *sherbet* (and *sorbet*) 'a cooling, sweet drink of diluted fruit juice' < Turkish and Persian *êerbet* and *šarbat* < Arabic *šarba* 'drink'; *syrup* 'a preservative or sweetener, especially in confections and drinks' < Middle French *sirop* < Arabic *šarāb* 'syrup, drink'; *coffee* (and *café*) 'a drink made from the seeds of the Coffea plant' < Italian *caffè* < Turkish *kahve* < Arabic *qahwa* 'coffee, wine, the dark brew'.

The Arabs introduced the game of chess to Europe in the 10th century. With it came the following: *checkmate* 'exclamation at chess when an adversary's king is inextricably checked and so has caused the loss of the game' (also just *mate*) < Middle French *escheck mat* < Persian *šāh* 'king' + Arabic *māt* 'he died'.

The following household items are notable: *jar* 'a vessel for holding liquids, without a spout' < Middle French *jarre* < Old Provençal *jarra* < Arabic *jarra* 'earthen water vessel'; *mattress* 'a resilient pad used as a resting place' < Old French *materas* < Arabic *maṭraḥ* 'place where something is thrown'; *sofa* 'an upholstered couch' < Turkish and French < Arabic *ṣuffa* 'long bench, divan'.

In modern times many Islamic terms have entered English as part of the everyday vocabulary, such as *ayatollah*, *hajj*, and *jihad*. Older loanwords deserving an etymological analysis include: *assassin* 'a member of a secret Muslim order who murdered Christians during the Crusades while supposedly under the influence of hashish' < Arabic *ḥaššāšīn* 'those who use hashish (masc. pl., oblique)'; *masjid* 'mosque' < Arabic *masjid* 'mosque'; *mesquita* 'mosque' < colloquial Arabic *masgid* 'mosque'.

Bibliographical references

Cannon Garland (with Alan S. Kaye). 1994. *The Arabic contributions to the English language*. Wiesbaden: O. Harrassowitz.

Pei, Mario A. 1967. *The story of the English language*. Philadelphia: Lippincott.

Alan S. Kaye (California State University and United Arab Emirates University)

English Loanwords

1. English loanwords in Arabic

Borrowing is a natural product of language contact between two communities using different languages. It happens in situations of colonization or when one language fills lexical gaps with words available in another language. It may also take place for reasons of prestige and may appear in the form of calques. The degree of borrowing depends on the intensity and length of time of contact (Thomason and Kaufman 1988:65–109, 215–228). In the case of borrowing from English into Arabic it should be remembered that the British were once colonizers in Arab countries, including Palestine, and that English is the modern language of science and technology, besides being a prestige language. Satellite broadcasting has made the world a small village allowing everybody to see what others have in life.

Little has been written about Arabic borrowing from English. As a result of the diglossic situation in Arabic, borrowing occurs in the vernacular dialects or the low language (Ferguson 1959), which is not used in writing, but only in speech (Nalborczyk 2002). Any new addition to standard Arabic has to go through the Arabic language academy (*majmaʿ al-luġa al-ʿarabiyya*). In everyday life, however, speakers cannot wait for the academy to tell them what words to use, and even if they wait for newly coined words as equivalents to the English ones, such neologisms may be acceptable but rarely used. Cowan (1976:VIII) observes that the impact of Western civilization has confronted the Arab world with the serious linguistic problem of expressing a vast and ever-increasing number of new concepts for which no words in Arabic exist. The Arabic language academies in Damascus and in Cairo coined many new terms in the field of science and technology (e.g. *miḏyāʿ* 'radio', *ḥāsūb* 'computer', and *mirnāh* 'television') but few have gained acceptance.

Without English technical terminology, people would not be able to use computers, which are becoming universal and essential tools in everyday life. A long time before today's scale of computer use, Kachru (1982) observed that the spread of English was as significant in its way as the increased use of computers. When the need for global communication came to exceed the limits set by language barriers, the spread of English accelerated.

The most recent study of Arabic borrowing from English is Daher (2003). He collected his data by interviewing men and women in Damascus to study the linguistic variation between the two genders (Daher 1998). In his interviews, he neither discouraged nor encouraged the use of foreign words. He found that the spoken Arabic of Syria included a vast number of older loanwords from Turkish (Mutawallī 1991) and Persian. More recent loanwords came from English, French, Italian, and Spanish (Ambros 1977). Daher found that many recent borrowings reflect new technology, such as fax, satellite dish, pager, e-mail, cell phone, connection. The export of technology is accompanied by the export of names for that technology. Ngom (2002) found through interviews a connection between borrowing and sociolinguistic variables such as age, culture, and politics.

Some Arabic magazines covering stories about singers and cinema stars include English loanwords in Arabic because interviews with such people are reported as they are spoken in the local dialect. The prescriptive Arab linguists see such magazines as deviant in their use of local dialects. Local Arabic dialects have adopted many English words without replacing them with Arabic approximation. Such borrowing has been responsible for introducing sounds such as /g/ and /v/, as in *hamburger*, *visa*, thus affecting the phonological inventory of the dialect.

2. Motives for borrowing

Borrowing occurs to fill a lexical gap (Atawneh 1992) or for reasons of prestige. In Cameroon, English and French borrow from indigenous languages for local color (Echu 2003). The most obvious motivation for English loanwords is business advertising. American products in foods, clothing, shoes, and other kinds of industry are popular in the Arab market. Goods that carry English loanwords are more in demand and find more customers than local goods; English loanwords add a certain cachet to the advertising process. Advertising uses colloquial language, very subtle yet precise.

A recent survey of 6,250 shops in the city of Hebron in Palestine showed only 100 shops used English names. Hebron is known as a conservative town and a somewhat closed society but with great skill in trade and commerce, which allows for contact with exporting countries for various products. It has a population of about 450,000. Use of English in naming shops started in the 1990s and reflects a rising tendency in borrowing from English. Compared to shops using Arabic names the shops that used English names were also characterized by higher quality of merchandize and better kind of customer. The ones that preferred English names turned out to be those in the shoe industry, sports and dress, like *Rami Sports*, *Tennis Shoes*, *Reem Sports*, *Delux Shoes*, *King Star Shoes*, *Gold Shoes*, *Pretty Woman*, *Backfire Shoes*.

Interviews with shop owners as to why they had chosen to use English names for their shops revealed the appeal and popularity of such names to the public who look for quality and kudos. Such people are middle-class and care about appearances. Most of the shop owners are educated and have some knowledge of English. All of them know the meaning of the names (or titles) they use. Half of them like to use English in their work. About 65 percent of the shop owners use only English in the inscriptions on their shops, which reflects their high status.

Reasons given by owners of shops with English-only names as to why they used English were: English is an international language; goods imported from the West are of better quality; the style and status of English is higher; and English is favored when dealing with women's articles. Reputable names in the West, like *Armand, Teresa, Cinderella, 4-cats, Grandee*, and *Castro* were among the popular names.

3. PHONOLOGICAL ADAPTATION

According to Asher (1994), loanwords are of interest to phonologists for at least two reasons. First, the way in which the loanword is pronounced in the borrowing language is often quite different from its pronunciation in the original language. Second, in many languages loanwords have particular phonological characteristics that make them distinct from the native vocabulary. Haugen's (1950) traditional classification categorizes borrowing into either substitution or importation. Other terms used include such terminological pairs as importation (adoption) vs. substitution, nativization (adaptation).

The degree of adoption of English words in Arabic depends on the following factors:

i. Structural differences between the two languages: the structural difference is not great between the two languages. The disagreements are minimal and mostly relate to vowels.
ii. Quantity of loans from the same source in the borrowing language: the quantity of loans is on the rise due to the fast increase in the Western products that are imported and used by people.
iii. Degree of bilingualism: the more familiar speakers are with foreign words, the more likely they are to adopt such words. Although English is required in schools (7 years), few school graduates are fluent users. Therefore, English is used mostly for reading and writing, not for speaking.
iv. Prestige associated with the lending language: English is the most prestigious language in the world; this is the main reason for the rise in borrowing English (Kachru 1982).
v. Social attitudes toward bilingualism, linguistic nationalism: social attitudes are not against bilingualism, and loanwords in the social domain are the second highest in both males and females, showing that users have a positive attitude toward English.

While the phonological system of Standard Arabic does not include /p/, /v/ or /g/, all three sounds are commonly used by speakers of local Arabic in pronouncing foreign words. The sound systems of English and Arabic do not match; each language contains some sounds and some points of contrast in the consonant patterns (e.g. voicing, uvularization, pharyngealization, velarization) which are not found in the other. This mismatch causes speakers to adapt, rather than simply adopt, borrowed terms. Because the number of Arabic sounds not found in English is greater than the number of English sounds not found in Arabic, Arabic loans in English are more adapted than are English loans in Arabic. In most cases, sounds in borrowed terms that have no matching sounds in the borrowing language are either deleted or replaced by the closest native equivalents.

The phonological inventory of English includes the voiceless and voiced interdental fricatives, /ṯ/ and /ḏ/. While Standard Arabic also includes these sounds, they are generally replaced in the local dialects by the corresponding dental fricatives, /s/ and /z/ in Lebanon and Syria, or by the alveolar /t/ and /d/ respectively in the urban areas of Palestine and Jordan. Villagers usually keep the same standard sounds. Consonantal change in loanwords mainly concerns the /p/ and /b/ difference. English differs from Arabic in showing opposition between voice and lack of voice at the labial place of articulation, i.e., English has both /b/ and /p/ while Arabic has only /b/. As a result, Arabic speakers often replace /p/ in English borrowings with /b/ as shown in Table 1:

Table 1. Change of /p/ to /b/ in English loanwords

English	Arabic	English	Arabic
petrol	/bətrol/	*police*	/bulīs/
pizza	/bītza/	*tape*	/teib/
pendulum	/bəndol/	*lamp*	/ləmbə/
power	/bəwər	*speaker*	/sbīkər/
supply	səblai/		
blouse	/bluzi/	*passport*	/bəsbor/
paradise	/bərədais/	*special*	/sbišəl
reception	/risibšən/	*pardon*	/bərdon/
permit	/birmit/	*packet*	/bəkēt/

At the velar place of articulation, English distinguishes between voiced and voiceless stops, having both voiceless /k/ and voiced /g/. Although Standard Arabic has only voiceless /k/, speakers of local Arabic routinely use /g/ in words of foreign origin. This adoption, rather than adaptation, of /g/ most likely occurs because speakers are already familiar with the sound: /g/ routinely replaces /q/ in Bedouin and village dialects in Palestine and /j/ in the Egyptian dialect. Even though there are few Bedouin around in Palestine, Egyptian movies and television have long been at the forefront of the Arabic-language entertainment industry.

At the labio-dental place of articulation, English distinguishes between the voiceless fricative /f/ and the voiced fricative /v/, while Arabic has only the voiceless /f/. The /v/ in English borrowings is sometimes, but not always, replaced with /f/, e.g., *vīza~fīza* 'visa' and *tilvizyōn~tilfizyōn* 'television'. However, in local dialects, the /v/ sound is adopted besides /g/ in female names like *mervat* and *nivin*. Examples of accepting /v/ and /g/ in the pronunciation of the borrowed forms include the following: *microwave* [maekrowe:v], *vase* [va:zə], *receiver* [risiivər], *hamburger* [haembergər], *goal* [go:l], *garage* [gəra:dʒ], *gallon* [gələn].

The English affricate /j/ is often, but not always, replaced in Arabic with the palatal fricative /j/, as in *djīnz~jīnz* 'jeans'. Such a change is consistent with the practice of speakers in urban areas like Jerusalem, Nablus, and Gaza.

Since the number of vowels in English is double that of vowels in Arabic (12/6), Arab users of English adapt those English vowels that do not exist in Arabic. In Arabic, there is no distinction between /i/ and /e/, which are produced as identical pairs as in /sit/ and /set/ (Kharma and Hajjaj 1989), as in *special* [sbiʃal], *telex* [tiliks]. [i] may also replace English [ɜ] as in [birmit] < *permit*.

The sound [eə] as in [feə] *fair*, [keə] *care* is often replaced by the nearest vowel sound followed by a clear Arabic [r], so that [eə] becomes [eː], e.g. *software* [softweːr]. The sound [əʊ] as in [rəʊz] *rose* is often replaced by the colloquial Arabic vowel /ō/ as in *goal* [goːl], *mobile* [moːbajl]. The sound [ei] as in [leidi] *lady* is replaced by the long colloquial Arabic vowel [eː] as in *cable* [keːbil]. The English shwa [ə] replaces the low front unrounded vowel [a] as in *balcony* [bəlkoːn].

According to Atawneh (2003), the syllable structure of *fuṣḥā* Arabic is either CV, CVC, or CVCC; however, the structure of local dialects is CCV, CVCC, or CCVCC. English syllable structure is CCV, CCVC, or CVCC. That means the onset of a syllable in Standard Arabic is always there, but can only include a single consonant; the coda of a syllable may include one consonant or two. However, in local Arabic, the onset or the coda may have one or two consonants resembling the English syllable structure rather than the *fuṣḥā* structure. This is why loanwords agreeing with the syllable structure of vernacular Arabic will not change or become adapted. Borrowed English names into Arabic are either adopted with no change in their phonological structure or adapted with some change in their syllable structure to suite the Arabic system. Therefore, borrowed words may be classified into two types, adopted and adapted. Adopted words agree totally with the phonology and morphology of Arabic. Table 2 gives some examples of adapted English loanwords:

Table 2. Phonologically adapted English loanwords

mouse [mæws]	*CD* [siːdiː]	*modem* [moːdim]
video [viːdju]	*microwave* [maikroweːv]	*fax* [fæks]
microphone [məkrəfoːn]	*set up* [setəp]	*bomb* [bomb]
software [softweːr]	*C.P.U.* [siːpjuː]	*hard disk* [hard disk]
cassette [kəsit]	*telephone* [tələfoːn]	*headphone* [hedfoːn]
telex [tiliks]	*cable* [keːbil]	*receiver* [risiːvar]

4. SPELLING AND TRANSCRIPTION OF ENGLISH LOANWORDS

In the Hebron survey about 20 percent of the shop owners used English along with Arabic; these shops wanted the customers to be able to read the names in Arabic if they did not have the ability to read English, as seen in the following examples: Garden City for Curtains and Décor جاردن سيتي للبرادي والديكور; Happy Family Exhibition معرض العائلة السعيدة.

A group of about 15 percent transcribed the English names into Arabic. These people were simple, poor, and most of them were not educated. They regarded English as a difficult language besides not having enough time or money to learn English. The following are some examples of writing English loans in Arabic alphabet:

إبراهيم الرجعي للإنترنت (Internet)، سوبر ماركت (Super-
market) الأراضي المقدسة، محلات الدويك للسبورات
(Sports) الجاهزة، بوتيك لايف (Life Boutique)، مون
لايت (Moonlight)، معرض فير ليدي (Fair Lady)،
حلويات براديس (Paradise)، كلاسيك شوب (Classic Shop)

When asked why they wrote English words in Arabic, these shop owners said their knowledge of English was very poor and they were unable to use it correctly. The hierarchy of status between English-using, English-mixing, and English-transcribing owners (65 percent, 20 percent, and 15 percent respectively) shows that English loans are used in degrees relevant to the class of people and their income.

Most borrowed names are common nouns, and compound nouns. To mention but a few: Roaster AmericanC, Red Shoes, Pretty Woman, Cinderella, Beauty Saloon, Yahoo, Castro, Happy Bunny, Five Cats, Internet Café, Coffee Shop, Computer Software, Mobiles, Ceramics, Boutique, Telephone, Toilet, Deodorant, Hamburger, MacDonald, Coca Cola, Macaroni.

The names and borrowings abound with spelling mistakes, e.g. prothers (brothers), jop (job); parquin (parking); chiken (chicken); dinning room (dining); sanwich (sandwich); alyes (eyes); holly (holy); cosmatiks (cosmetics); maneger (manager); magestic (majestic); mak (make); markiting (marketing); angent (agent), taityanic (titanic); bowtik (boutique); midical (medical); roes (rose).

5. SEMANTIC DOMAINS

Wilson (2001) reported that Arabic words were borrowed indirectly into Middle English across all semantic fields. She also found that the majority of borrowed words were nouns and adjectives. Echu (2003) studied indigenous loans into French and English in Cameroon and made four lists of recurrent examples of culture-based areas like gastronomy, traditional titles, dance music, as well as sociocultural institutions, practices, objects, and concepts. He claims that such borrowing adds local color to the official languages. This seems to be true in the case of using English loans in the Arabic dialects, too. In the Hebron survey, semantic domains were tested by means of a questionnaire which requested subjects to give the words of English origin they used in daily life: 1. computer industry, 2. car industry, 3. health, body, foods, 4. kitchen and house, 5. electronic machines, 6. clothes, 7. human relations, and 8. other.

Interestingly, there seems to be a gender difference in the borrowing pattern. In a sample of 30 female and 30 male undergraduates aged between 20 and 23 at Hebron University it turned out that there was a significant difference in the number of words used by females compared to that used by males (Table 3). Females use nearly double the number of English loanwords (370) compared to males (213). This finding is expected, based on the assumption that females are more sensitive to prestige than males. Labov (1966) and Trudgill (1972) found more extreme style shifting toward 'the prestige norm' in women's speech. One of Labov's major hypotheses is that the linguistic insecurity of lower-middle-class women leads to sound change.

Among the seven semantic domains of loanwords 'health, body, foods' turns out to be the domain most used by females (20.30 percent), while it is 'car industry' that is most used by males (19.7 percent). This result seems to reflect the concerns of females versus males, which go with the social patterns of Arab society. It is only recently that women have begun to drive cars and therefore one does not expect their use of car-related words from English to match that of males. The second domain most used by females is the computer industry (15.4 percent), while for males it is electronic machines (12.8 percent). The third for females is social relations (13.2 percent), while it is the computer industry

Table 3. Number of words per semantic category for both genders

Field	Females	%	Rank	Males	%	Rank
1. Health, body, foods	75	20.30	(1)	33	10.50	(4)
2. Computer industry	57	15.40	(2)	26	12.20	(3)
3. Human relations	49	13.20	(3)	17	7.10	(7)
4. Car industry	47	12.80	(4)	42	19.70	(1)
5. Clothes	40	10.80	(5)	23	10.40	(5)
6. Electronic machines	39	10.50	(6)	28	12.80	(2)
7. Kitchen, house	29	7.80	(7)	15	7.40	(6)
8. Other	34	9.20		30	14.40	
Total	370	100.00		213	100.00	

Table 4. Shared versus non-shared words in females and males

Domain	Shared	%	Females	%	Rank	Males	%	Rank
1. car industry	30	18.7	17	08.2	(5)	11	23.9	(1)
2. Health, body, foods	27	16.9	48	23.2	(1)	6	13.1	(3)
3. Electronic machines	26	16.2	13	06.3	(7)	1	02.2	(6)
4. computer industry	23	14.4	33	15.9	(3)	4	08.7	(4)
5. Clothes	20	12.5	20	09.6	(4)	3	06.5	(5)
6. Kitchen, house	15	09.4	14	06.8	(6)	0	00.0	(7)
7. Human relations	10	06.3	38	18.4	(2)	7	15.2	(2)
8. Others	9	05.6	24	11.6		14	30.4	
Total	160	100.0	207	100.0		46	100.0	

(12.2 percent) in males. The fourth rank in both is the category of clothing. Actually, looking at the shared words used by both genders gives a better idea of the differences between the two genders and the differences among the various areas in the sample (see Table 4).

The car industry is the highest shared domain for both genders. However, this same category is highest among males exclusively, while it is the fifth area for females, which shows the lack of interest in this area by females when compared to males. The following are the top 16 frequent English loans in this area: *gear*, *clutch*, *brake*, *motor*, *body*, *handbrake*, *taxi*, *service*, *reverse*, *test*, *bus*, *dynamo*, *switch*, *condition*, *jack*, *garage*, *carburetor*. Some of the loanwords derive from French, like *tableau* 'dash[board]' and [bəskəleːt] < French *bicyclette*.

The category of 'health, body, foods' is rated second highest in shared words, while it is the highest for females (48 words) and third highest in males (6 words). This result agrees with the nature of women's concern with such matters, which is usually greater than it is for males. The following are the top 15 used words: *pizza*, *hamburger*, *ketchup*, *sandwich*, *cake*, *schnitzel*, *mayonnaise*, *chocolate*, *chips*, *Nescafé*, *Coca Cola*, *biscuits*, *doctor*, *steak*, *shampoo*, *diet*.

The kitchen and house is the third domain that differs for females and males. Females have it as sixth in rank. For males it has no single word other than shared words. This result shows also the roles of females compared to males. The following are the top 15 used words: *corner*, *corridor*, *balcony*, *saloon*, *buffet*, *veranda*, *toilet*, *thermos*, *roof*, *jacuzzi*, *shower* (also *dosh* < French *douche*), *décor*, *mug*, *gas*, *villa*.

The similarities between the two genders are in computer industry, clothes, and electronic machines. Here are some examples from each group.

Computer industry: *mouse*, *computer*, *keyboard*, *internet*, *disk*, *cd*, *floppy*, *scanner*, *file*, *printer*, *e-mail*, *hard disk*, *microphone*, *headphones*, *software*, *cable*.

Clothes: *blouse*, *body*, *charleston*, *chiffon*, *coat*, *cowboy*, *cut*, *fashion*, *jacket*, *jeans*, *model*, *overall*, *off-white*, *pajamas*, *shorts*, *sport*.

Electronic machines: *microwave*, *television*, *radio*, *satellite*, *telephone*, *freezer*, *video*, *mole nix* (Moulinex), *microphone*, *receiver*, *mobile*, *lamp*, *digital*, *stereo*, *fax*, *cassette*, *camera*, *remote*, *drier* (also [siʃwaːr] < French *séchoir*).

Among words under the category of 'other' (mainly sports), the following are the top 10 words in frequency: *football*, *goal*, *basketball*, *film*, *flash*, *racket*, *supermarket*, *tennis*, *Visa card*.

6. CONCLUSION

Apparently, borrowing is becoming widespread in Arabic dialects. Loanwords are used in daily life, mostly in the form of nouns and adjectives in relation to kinds of foreign products and social relations. The majority of English loanwords are adopted when they agree with the phonological system of dialect; otherwise they are adapted. A few changes occur in vowels to agree with the system of Arabic. The main motivation for borrowing English according to informants is prestige, which appeals to the younger generation. Users believe that English is an international language and, therefore, they are encouraged to use it in borrowing. Females seem to be more sensitive to prestige than males and thus they use more English loans than males.

BIBLIOGRAPHICAL REFERENCES

Ambros, Arne. 1977. *Damascus Arabic*. Malibu, Calif.: Undena.

Asher, Ron and J. Simpson (eds.). 1994. *The encyclopedia of language and linguistics*. I. Pergamon Press.

Atawneh, Ahmad. 1992. "Code-mixing in Arabic-English bilinguals". *Perspectives on Arabic linguistics*, IV, ed. Ellen Broselow, Mushira Eid, and John McCarthy, 219–241 Amsterdam and Philadelphia: J. Benjamins.

———. 2003. "Epenthesis in local Arabic". *Bethlehem University Journal* 22.108–198.

Cowan, J. Milton (ed.). 1976. *A dictionary of modern written Arabic: (Arabic–English)/Hans Wehr*. Ithaca, N.Y.: Spoken Languages Services.

Daher, Jamil. 1998. *Linguistic variation in Damascus Arabic: A quantitative analysis of men's and women's speech*. Ph.D. diss., New York University.

Echu, George. 2003. "Multilingualism as a resource: The lexical appropriation of Cameroon indigenous languages by English and French". <http://www.inst.at/trans/13Nr/ echu13.htm.>

Ferguson, Charles A. 1959. "Diglossia". *Word* 15.325–340.

Haugen, Einar. 1950. "The analysis of linguistic borrowing". *Language* 26.210–231.

Kachru, B. Braj (ed.). 1982. *The other tongue: English across cultures*. Exeter, U.K.: A. Wheaten.

Kharma, Nayef and Ali Hajjaj. 1989. *Errors in English among Arabic speakers: Analysis and remedy*. London: Longman.

Labov, William. 1966. *The social stratification of English in New York City*. Washington, D.C.: Center for Applied Linguistics.

Mason, Patricia. 1979. "Social implications of borrowing: The Visigothic element in Hispano-Romance". *Word* 30.257–272.

Mutawallī, ʾAḥmad Fuʾād. 1991. *al-ʾAlfaḏ at-turkiyya fī l-lahjāt al-ʿarabiyya wa-fī luġat al-kitāba*. Cairo: Dār az-Zahra.

Nalborczyk, Agatha Skowron. 2002. "The Arabophones in German-speaking communities between diglossia and bilingualism: Lexical borrowings between languages and varieties." *Internet-Zeitschrift für Kulturwissenschaften*. <http://www.inst.at/trans/ 13Nr/ skowron13.htm>.

Ngom, Fallou. 2002. *Lexical borrowings as sociolinguistic variables in Saint-Louis, Senegal*. Ph.D. diss., University of Illinois, Urbana-Champaign.

Rieschild, Verna. 2003. "Words between languages and cultures". <http://www.shlrc.mq.edu. au/style/june2003.htm>.

Robertson, Stuart. 1954. *The development of modern English*. Englewood Cliffs: Prentice Hall.

Thomason, Sarah G. and Terrence Kaufman. 1988. *Language contact, creolization and genetic linguistics*. Berkeley: University of California Press.

Trudgill, Peter. 1972. "Sex, covert prestige, and linguistic change in the urban British English of Norwich". *Language in Society* 1:2.179–195. (Repr. *Language and sex: Difference and dominance*, ed. B. Thorne and N. Henley, 88–104. Rowley, Mass.: Newbury House.)

Wilson, Jessica. 2001. "Arabic in Middle English". <http://www.chass.utoronto.ca/~cpercy/courses/6361Wilson.htm>.

AHMAD ATAWNEH (Hebron University)

Epenthesis

The clustering of consonants at either or both edges of a syllable, which renders it complex (CCVC or CVCC), is disfavored, if not categorically banned, in many languages. Epenthesis (also called → anaptyxis) is one of the repair mechanisms a language may employ to rectify syllabification violations ensuing from undesirable clusters. The surface effect of epenthesis is insertion of a 'helping' vowel, to facilitate proper syllabification of all output consonants, thereby rendering surface phonological representations in tandem with the syllable structure constraints of the language (→ prosody). Cluster simplification via consonant deletion is another repair strategy, but the morphemic status of the consonantal root in Arabic as the main holder of core semantic features guards root consonants against deletion; hence epenthesis remains the optimal option to resolve unwanted complex syllables across the dialect spectrum, with the inclusion of Standard Arabic.

Classical Arabic and its modern offshoot, Modern Standard Arabic, is highly conservative in its tolerance of complex syllables, which are permitted only in pre-pausal position. Epenthesis is not pervasive in Classical Arabic, however, except when a cluster-initial word occurs post-pausally, or at the phrasal level where word concatenation may warrant vowel insertion, as in /l-bayt/ > [ʾalbajt] 'the house', or /hum ntaqal-u/ > [humuntaqalu] 'they moved-3Pl'. Vocalic endings which mark grammatical case in nouns and mood in verbs serve a dual function. In addition to their syntactic function as case and mood markers, they serve a phonological function as a barrier against cluster formation. Deletion of case and mood endings in the spoken dialects (apocope) created potential environments for cluster formation, e.g. Classical [katab-tu] 'wrote-1sg.' vs. dialectal [katabt].

In Modern Arabic dialects, sources of consonant clusters include morpheme concatenation, e.g. /gil-t-l-ha/ 'I said to her'; syncope, which deletes an unstressed high vowel in an open syllable, e.g. /yi-ktib-u/ > /yik.ti.bu/ > /yiktbu/ 'they write'; or, as in the case of some lexical nouns and deverbal nouns (*masḍar*) of Form I, the cluster is provided templatically, as in [kalb] 'dog' or [bint] 'girl' (of the template CVCC).

The dialects diverge in their tolerance of consonant clusters. Some may allow unrestricted clustering only at the right edge of the syllable (complex coda), e.g. Egyptian, while others, guided by the sonority sequencing restrictions (Haddad 1984), break up the coda cluster with an epenthetic vowel. Others permit left edge or onset clusters, e.g. Levantine and Gulf (Qafisheh 1977) dialects. Few allow clustering unrestrictedly at both syllable edges, e.g. North African dialects.

This divergence provided fertile grounds for typological classification (Broselow 1983, 1992; Farwaneh 1995; Itô 1986, 1989; Kiparsky 2003; Selkirk 1981), with epenthesis site serving as the primary axis for the proposed dialect typology, thereby classifying the dialects as 'onset' or 'CV', vs. 'coda' or 'VC' dialects. If the epenthetic vowel lands after the unsyllabified consonant, thereby forming an open syllable, the dialect is classified as an 'onset' or 'CV' dialect. On the other hand, inserting the epenthetic vowel before the unsyllabified consonant to form a closed syllable, identifies the dialect as a 'coda' or 'VC' dialect. This variability in the positioning of epenthesis was not only evident in L1 phonology, but, as observed in Broselow (1983), was also transferred into L2 phonology, thereby producing divergent surface forms such as [filoːr] and [istiriːt] vs. [ifloːr] and [sitriːt] among Egyptian and Iraqi learners of English, respectively.

As correlations were drawn between epenthesis site in medial and initial clusters on the one hand, and between the position of epenthesis and other phonological processes such as → syncope and shortening on the other, it was observed that onset-CV dialects shorten long vowels in closed syllables, and block high vowel syncope in open syllables. Conversely, coda-VC dialects maintain vowel length in closed syllables while deleting high vowels in open syllables, thereby rendering its output subject to subsequent epenthesis. Table 1 exemplifies the typological generalizations.

Although the aforementioned works converge in their typological goals, they diverge in their theoretical treatment of the typological observations. The degenerate syllable camp, spearheaded by Selkirk's work, accounts for the positional variability of epenthesis by stipulating the type of 'degenerate' (nucleusless) syllable an unsyllabified segment is assigned to. This view was later implemented in Abu-Mansour (1990, 1991) and expanded in Broselow (1992). The second camp (Itô 1986, 1989; Farwaneh 1995) resorts to the directionality parameter and its variable settings to account for the same problem. Rightward (onset dialects) or leftward (coda dialects) syllabification places the epenthetic vowel when needed in its proper location.

In a non-derivational theory such as → Optimality Theory, directional syllabification effects are successfully accounted for in terms of an Alignment family of constraints (Mester and Padgett 1994) which stipulate how constituent edges should overlap, e.g. Align left edge of syllable with left edge of word, or Align right edge of root (final root consonant) with right

Table 1. Typological generalizations in Arabic dialect syllable structure

a. Onset-CV Dialects			
Initial epenthesis:	CC > CVC:	/wla:d/ > [wilaːd]	'boys'
Medial epenthesis:	CCC > CCVC:	/dars-na/ > [darsina]	'our lesson'
Syncope:	Ci > *C:	/nizil-t/ > *[nzilt] but [nizilt]	'I descended' [syncope blocked]
		/yi-ktib-u/ > *[yiktbu] but [yiktibu]	'they write' [syncope blocked]
Shortening:	CVVC > CVC:	/xa:l-na/ > [xalna]	'our uncle'
b. Coda-VC Dialects			
Initial epenthesis:	CC > VCC:	/kta:b/ > [iktaːb]	'book'
Medial epenthesis:	CCC > CVCC:	/dars-na/ > [darisna]	'our lesson'
Syncope:	Ci > C:	/nizil-t/ > [nzilt]	'I descended'
		/yi-ktib-u/ > /yiktbu/ > [yikitbu]	'they write' [syncope and epenthesis]
Shortening:	CVVC > *CVC:	/xaal-na/ > *[xalna] but [xaalna]	'our uncle' [shortening blocked]

edge of stem, etc. Attempts are currently being undertaken to eliminate the constraint family Align and replace it with categorical constraints (McCarthy 2003), but the success of this proposal in accounting for all aspects of Arabic dialect epenthesis is yet to be determined. Other optimality-theoretic accounts of epenthesis propose generating the typology within a constraint-based model of lexical phonology and morphology, termed 'stratal' Optimality Theory (Kiparsky 2003:152), which "requires distinct constraint systems for word phonology and sentence phonology, which moreover must interact in serial fashion" (Kiparsky 2003:151).

Despite the divergent ways in which dialects syllabify stray consonants, they converge on the fact that epenthesis is blocked if it splits a geminate (→ gemination). Guerssel (1979) and Abu-Salim (1980) were the first to observe that geminate integrity preserves true geminates from epenthesis split; hence [binit] < /bint/ 'girl' but not *[sitit] < /sitt/ 'woman', derived from the biliteral root /st/. However, fake geminates which arise as a result of morpheme concatenation, e.g. /sakat-t/ > [sakatit], behave on a par with non-identical clusters in inducing epenthesis. Regressive → assimilation, on the other hand, gives rise to true gemination immune to epenthesis (Guerssel 1979). For instance, assimilation of the definite article /l/ in place of articulation to a stem-initial coronal results in an identical sequence which does not yield to epenthesis; thus, [liktaːb] < /l-kta:b/ 'the book', but not *[zizlaːm] < /l-zla:m/ 'the men'; instead, the surface form is [izzlaːm] with pre-geminate epenthesis in some dialects which do not tolerate initial triconsonantal clusters, or [zzlaːm] without epenthesis in others.

Another puzzling issue which captured linguists' interest is the transparency vs. opacity of the interaction between → stress and epenthesis. The interaction of stress and epenthesis is transparent in onset dialects, as the epenthetic vowel is treated on a par with underlying ones. In Egyptian Arabic, for example, penultimate stress targets both underlying and epenthetic vowels, e.g. [mád.ra.sa] 'school' and /bint-ha/ > [bín.ta.ha] 'her daughter'. In Saudi Arabic, where stress falls on the antepenult if both the penult and the ultima are light, both underlying and epenthetic vowels escape stress, e.g. [mád.ra.sa] and /bint-ha/ > [bín.ta.ha]. In Levantine Arabic, by contrast, epenthetic vowels are invisible to stress, even if they fall in what would otherwise constitute a closed, expected to be heavy, syllable, e.g. [ka.táb.na] 'we write' but /dars-na/ > [dá.ris.na], not *[da.rís.na].

Brame (1973), the first to account for this behavioral nonuniformity within the then dominant generative approach of Chomsky and

Halle (1968), captured this opacity through rule ordering and cyclic rule application, thus accounting for such triplets as [fhímna] ‘we understood’ from /fihim-na/, [fihímna] ‘he understood us’ from /fihim-Ø-na/, and [fíhimna] ‘our understanding’ from /fihm-na/. Within Moraic Phonology, both Broselow (1992) and Farwaneh (1995) capture the opaque stress/epenthesis interaction in Levantine Arabic through a typology of moras and syllables, although the two approaches arrive at the same target through heterogeneous parsing processes. An optimality-theoretic account of stress and epenthesis interaction premised on variable constraint ranking and minimal constraint violation is presented in Farwaneh (1996) and Kiparsky (2003).

Like its positional variants, the allophonic variants of the epenthetic vowel received attention in the literature on Arabic dialects, descriptive and analytical alike. The segmental features of the epenthetic vowel in some dialects, e.g. Syrian, Iraqi, Gulf, and Egyptian, are susceptible to alternations under the influence of the place features of neighboring consonants, or the harmonizing features of the preceding or following vowel. Such alternations include lowering in guttural (pharyngeal and laryngeal) contexts (Herzallah 1990; McCarthy 1991, 1994) and rounding in emphatic environments. The examples from Gulf Arabic (Ingham 1982) in Table 2 demonstrate the two types of alternation.

Table 2. Epenthetic vowels in Gulf Arabic

Input	Output	Gloss	Process
ḥaml	*ḥamil*	‘burden’	no alternation
šahr	*šahar*	‘month’	lowering in guttural environment
šaʿr	*šaʿar*	‘hair’	–
barg	*barug*	‘lightning’	rounding in emphatic environment
xaṣm	*xaṣum*	‘enemy’	rounding in emphatic environment

Harmonizing epenthesis wherein the vowel assimilates to a neighboring vowel across consonants occurs in Egyptian and Levantine. In Levantine (Abu-Salim 1987), the epenthetic vowel harmonizes with the preceding vowel. This process, however, is intramorphemic and does not apply intermorphemically (across a stem and a suffix), as shown in Table 3.

Table 3. Epenthetic vowels in Levantine Arabic

Input	Output	Gloss	Process
bint	*binit*	‘girl’	no alternation
baḥr	*baḥar*	‘sea’	lowering in guttural context
ʿumr	*ʿumur*	‘age’	harmony
šuf-t	*šufit* **šufut*	‘I saw’	harmony blocked intermorphemically

In Egyptian, the trigger of harmony is the following vowel, and the process is not blocked by morpheme boundaries, as shown in Table 4.

Table 4. Epenthetic vowels in Egyptian Arabic

Input	Output	Gloss	Process
sib-t-na	*sibtina*	‘you left us’	no alternation
sib-t-ha	*sibtaha*	‘you left her’	lowering
sib-t-hum	*sibtuhum*	‘you left them’	harmony

Dialectal variation in the segmental quality of the epenthetic vowel itself received less attention in the literature than its allophonic variants or landing site. The pervasive assumption is that the quality of the epenthetic vowel is the by-product of a set of language-specific redundancy rules which fill in empty nuclei with the segmental features of the default vowel in the language (Archangeli 1984). The default vowel is a segment present in the vowel inventory of the language. Thus, cross-linguistic differences in the quality of epenthetic vowels follow from arbitrary selection of different default vowels.

Many dialects select a high front /i/ or mid central /ə/; a few, for instance Saudi (Abu-Mansour 1987, 1990), Sudanese (Hamid 1984;

Trimingham 1946), and Ṣanʿānī Yemeni (Watson 2002) opt for a low epenthetic vowel /a/. An attempt to correlate epenthetic quality and epenthetic site is proposed in Farwaneh (1995) who gives a directionality-oriented account motivated by the exclusive occurrence of low vowel epenthesis in onset-CV dialects. The relative weakness of open syllables triggers, as a strengthening device, a highly sonorant vowel (low vowels are more sonorant than high ones), hence the implication that epenthetic low vowels occur in open syllables only. This proposal awaits acoustic verification and cross-linguistic evidence.

BIBLIOGRAPHICAL REFERENCES

Abu-Mansour, Mahasen. 1987. *A nonlinear analysis of Arabic syllabic phonology with special reference to Makkan*. Ph.D. diss., University of Florida, Gainesville.

——. 1990. "Epenthesis, gemination and syllable structure". *Perspectives on Arabic linguistics*, II, ed. Mushira Eid and John McCarthy, 167–191. Amsterdam and Philadelphia: J. Benjamins.

——. 1991. "Epenthesis in Makkan Arabic: Unsyllabified consonants v. degenerate syllables". Eid and Comrie (1991:137–154).

Abu-Salim, Isam. 1980. "Epenthesis and geminate consonants in Palestinian Arabic". *Studies in the Linguistic Sciences* 10:2.1–12.

——. 1987. "Vowel harmony in Palestinian Arabic: A metrical perspective". *Journal of Linguistics* 23:1.1–24.

Archangeli, Diana. 1984. *Underspecification in Yawelmani phonology and morphology*. Ph.D. diss., Massachusetts Institute of Technology.

Brame, Michael. 1973. "On stress assignment in two Arabic dialects". *A festschrift for Morris Halle*, ed. Stephen R. Anderson and Paul Kiparsky, 14–25. New York: Holt.

Broselow, Ellen. 1982. "On predicting the interaction of stress and epenthesis". *Glossa* 16:2.115–132.

——. 1983. "Nonobvious transfer: On predicting epenthesis errors". *Language transfer in language learning*, ed. Susan M. Gass and Larry Selinker, 269–280. Rowley, Mass.: Newbury House.

——. 1992. "Parametric variation in Arabic dialect phonology". *Perspectives on Arabic linguistics*, IV, ed. Ellen Broselow, Mushira Eid, and John McCarthy, 7–46. Amsterdam and Philadelphia: J. Benjamins.

Chomsky, Noam and Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.

Eid, Mushira and Bernard Comrie (eds.). 1991. *Perspectives on Arabic linguistics*, III. Amsterdam and Philadelphia: J. Benjamins.

Farwaneh, Samira. 1995. *Directional effects in Arabic dialect syllable structure*. Ph.D. diss., University of Utah.

——. 1996. "The interaction of stress and epenthesis in Arabic: An optimality-theoretic approach". *Proceedings of the Eastern States Conference on Linguistics*, ed. Lenore Grenoble, Lindsay Whaley, and Marek Przezdziecki, 48–57. Ithaca: Cornell University Press.

Guerssel, M. 1979. *Toward a lexical approach to phonological description*. Ph.D. diss., University of Washington.

Haddad, Ghassan F. 1984. "Epenthesis and sonority in Lebanese Arabic". *Studies in the Linguistic Sciences* 14:1.57–88.

Hamid, Abdel Halim. 1984. *A descriptive analysis of Sudanese colloquial Arabic phonology*. Ph.D. diss., University of Illinois, Urbana-Champaign.

Herzallah, Ruqayyah. 1990. *Aspects of Palestinian Arabic phonology: A non-linear approach*. Ph.D. diss., Cornell University.

Ingham, Bruce. 1982. *North east Arabian dialects*. London and Boston: Kegan Paul International.

Itô, Junko. 1986. *Syllable theory in prosodic phonology*. Ph.D. diss., University of Massachusetts, Amherst.

——. 1989. "A prosodic theory of epenthesis". *Natural Language and Linguistic Theory* 7.217–259.

Kiparsky, Paul. 2003. "Syllables and moras in Arabic". *The syllable in optimality theory*, ed. Caroline Féry and Ruben van de Vijver, 147–182. Cambridge: Cambridge University Press.

McCarthy, John. 1991. "Guttural phonology". Eid and Comrie (1991:63–91).

——. "The phonetics and phonology of Semitic pharyngeals". *Phonological structure and phonetic form: Papers in laboratory phonology*, III, ed. Patricia Keating, 191–233. Cambridge: Cambridge University Press.

——. 2003. "OT constraints are categorial". *Phonology* 20:1.75–138. Cambridge: Cambridge University Press.

Mester, Armin and Jaye Padgett. 1994. "Directional syllabification in generalized alignment". *Phonology at Santa Cruz*, III, ed. Rachel Walker, Motoko Katayama, and Daniel Karvonen, 57–62. Santa Cruz: Linguistics Research Center, University of California.

Qafisheh, Hamdi. 1977. *A short reference grammar of Gulf Arabic*. Tucson: University of Arizona Press.

Selkirk, Elizabeth. 1981. "Epenthesis and degenerate syllables in Cairene Arabic". *Theoretical issues in the grammar of Semitic languages*, ed. Hagid Borer and Joseph Aoun, 3.209–232. Cambridge, Mass.: Massachusetts Institute of Technology, Department of Linguistics.

Trimingham, J. Spencer. 1946. *Sudan colloquial Arabic*. London: Oxford University Press.

Watson, Janet C.E. 2002. *The phonology and morphology of Arabic*. Oxford and New York: Oxford University Press.

SAMIRA FARWANEH (University of Arizona)

Epigraphy (Islamic)

Epigraphy (from the Greek *epigráphein* 'to write on') is the science or study of inscriptions. It is distinguished from calligraphy (lit. 'beautiful writing' → Script and art) by the nature of the physical support on which the writing is inscribed. Epigraphy is executed on durable materials. In the Islamic lands these include buildings made of stone or brick, where the epigraphy is often carved in relief form, and portable objects made of wood, metal, ceramics, or glass and the like. Calligraphy, in contrast, is executed on supple supports. In the Islamic lands it was typically done with a reed pen (*qalam*) on papyrus, parchment, or paper.

Virtually all epigraphy from the Islamic period is executed in Arabic script, most of it in the Arabic language, although vernacular languages were introduced beginning in medieval times. New Persian, for example, is first documented on the tomb of the Qarakhanid ruler Muḥammad ibn Naṣr, erected in the mid-11th century by his son at Safid Buland in the Farghana Valley of Central Asia. Turkish, known since the 13th century, became standard under the Ottomans.

Inscriptions in vernacular languages written in Arabic script, often in verse, decorate secular buildings and a wide variety of objects, from palaces to wine goblets. Arabic, however, has remained the principal language of epigraphy in buildings connected with the practice of the faith, such as mosques and madrasas, and their furnishings, such as *minbar*s or pulpits. Its importance is clear from its wide currency: Arabic inscriptions are found on buildings and objects created across the Islamic lands and in all historical periods.

Scholars often divide the inscriptions into historical or religious categories. These groupings are somewhat artificial, for both types of content can be found within the same inscription. Nevertheless, the divisions are instructive in outlining and understanding the different forms used for each category. Islamic epigraphy is a vast field; the easiest introduction to it is the survey by Blair (1998), from which many of the examples in this entry are drawn.

1. The importance of Arabic epigraphy

Along with vegetal and geometric designs, the extensive use of writing is one of the hallmarks of decoration found on buildings and objects created across the Islamic lands in all historic periods. Epigraphy was already well established in the early Islamic period. The Dome of the Rock in Jerusalem, ordered by the Umayyad caliph ʿAbd al-Malik in 72 A.H./692 C.E., is the first surviving example of Islamic monumental architecture. Although the exterior decoration was completely replaced in the 16th century under the Ottoman sultan Süleiman, the interior survives virtually intact. In both form (a domed ciborium) and materials (limestone decorated with quartered marble and glass mosaic), the building belongs squarely to the Late Antique tradition. The one feature that immediately distinguishes the Dome of the Rock as a building erected for Muslims is the long (240-meter) inscription band that runs around both sides of the ambulatory [Fig. 1]. Materials and technique underscore the prominence of the text: the letters are written in cubes made of gold foil pasted over glass, the most expensive *tesserae* in an already expensive technique, and they are laboriously set at a 30° angle to the surface, in order to better reflect light.

Inscriptions continue to play a prominent role on buildings today, especially religious ones, in which Arabic writing underscores the sanctity of the structure and its importance to the Muslim community. The mosque erected in 1984 at the King Khaled International Airport north of Riyadh, for example, has an enormous epigraphic band encircling the base of the geodesic dome. The inscription, which measures four meters in height, is said to be the largest of its kind ever produced, outdoing the already large ones that had been used in medieval buildings such as the complex built for Sultan Ḥasan in Cairo in the middle of the 14th century [Fig. 2].

Inscriptions were traditionally so important to Muslims that they had them added to objects even in cases where the technique of manufacture rendered their production extremely difficult or costly. Such is the case with textiles, the driving industry of medieval times [Fig. 3]. It is relatively easy to embroider an inscription on an already finished piece but much more difficult to incorporate an inscription during weaving.

Yet, by the end of the Umayyad period, weavers in the Islamic lands had already overcome the limitations imposed by loom technology and learned to incorporate epigraphs into their weaving. These inscription bands (and the textiles in which they are inscribed) became known as *ṭirāz* (< Persian *tarazidan* 'to embroider'). Woven in state factories, they became a hallmark of caliphal privilege, bestowed in large numbers as gifts. The textiles were later cut up and their inscriptions used to drape bodies, perhaps out of a belief in the prophylactic power of the word in warding off evil in the grave and in the afterlife (many examples of these have been preserved in graves in the dry soil of Egypt).

The expression *ṭirāz* became so commonplace that the 15th-century Egyptian historian al-Maqrīẓī used it to indicate the inscription band encircling the funerary complex erected between 682 A.H./1283 C.E. and 684 A.H./1285 C.E. by the Mamluk sultan Qalāʾūn, along the main street (Bayn al-Qaṣrayn) of Cairo [Fig. 4]. The band, which extends more than 67 meters, links the disparate parts of the complex – minaret, tomb, portal, and madrasa – into a unified façade. Its impact was heightened by color, for originally the background was painted red and the letters gilded.

Muslims used writing on buildings where people of other traditions, notably Christians, used images. This is clear when comparing the entrance façades of two medieval religious structures, the small mosque of al-ʾAqmar, erected in 519 A.H./1125–1126 C.E. on the main street of Cairo [Fig. 5], and the Benedictine abbey church of Sainte Foy at Conquès in southern France, whose west portal was completed at approximately the same time.

Both are limestone structures with large doorways into which are set carved tympana. The tympanum over the western door at Conquès depicts the Last Judgment: Christ is seated in majesty, flanked by 124 figures set in riveting scenes depicting heaven and hell. The images are inspired by the Gospel of Matthew, the text of which is inscribed on small banderoles unfurled by angels flanking Christ's head.

On the façade of the al-ʾAqmar Mosque, by contrast, words supplant images. The west façade, set at a 21° angle to align with the main street, is decorated almost exclusively with epigraphy, supplemented by small amounts of floral and geometric decoration. Noticeably absent are any representational figures. The largest band, carved in relief and running some 20 meters along the top of the building, contains the foundation inscription by the Fatimid vizier Maʾmūn al-Baṭāʾiḥī. This text is repeated in a slightly smaller band running across the façade at mid-level.

Perhaps the most stunning element of the decoration on the al-ʾAqmar façade is the pierced medallion set in the center of a scalloped tympanum over the central doorway [Fig. 6]. It contains a circular band inscribed with Q. 33/33: "And God only wishes to remove all abomination from you, ye members of the Family, and to make you pure and spotless, O people of the House". This verse was of particular relevance for Shiʿites, who interpreted it as granting them legitimacy as lineal descendants of the Prophet, and its use here can be construed to support the claims of the Fatimids, who were Sevener Shiʿites. The polemic message is reinforced by the inscription in the center of the medallion naming the prophet Muḥammad and his nephew and son-in-law, ʿAlī – names that are repeated on the chamfered corners of the mosque.

The façades of both Sainte Foy and the al-ʾAqmar Mosque exemplify the superb carving of 12th-century artisans. The tympanum at Conquès contains one of the richest ensembles of Romanesque sculpture to be found. The inscriptions on the al-ʾAqmar façade are equally elaborate. All are carved in sober, angular letters that sprout curved leaves and tendrils. The combination of angular script and curved decoration, as well as the equilibrium between letter and ornament, makes this one of the finest examples of stone epigraphy in the Islamic lands.

In both cases, the carved decoration is intended as proselytism for spreading the faith: Christians used images to explain the Word; Muslims used the Word alone.

2. HISTORIOGRAPHY AND SOURCES TO STUDY HISTORICAL INSCRIPTIONS

Scholars have studied Arabic epigraphy since the 18th century. Inscriptions on coins were the first to be studied, perhaps because Islamic coins, unlike those issued elsewhere – notably in the Classical world, Sasanian Iran, Byzantium, and medieval Europe – are exclusively epigraphic. The legends on Islamic coins typically include the mint and the date, as well as the name of the leader; the striking of coins, along with being named in the Friday bidding prayer (*xuṭba*), was widely recognized as one of the rights of a ruler. In medieval times, Islamic coins were traded extensively in Scandinavia, northern Germany, and Russia, and the large collections in these countries stimulated scholarly interest at an early date. Georg Jacob Kehr's monograph, entitled *Monarchiae asiatico-saracenicae status qualis VIII et IX post Chr. nat. seculo fuit, ex nummis argenteis script. Kufica in littore Maris Baltici prope Gedanum effossis illustratus*, published in Leipzig in 1724, provided readings of the legends on Islamic coins and is often reckoned to be the first scholarly book not only on Islamic numismatics, but on Islamic archaeology in its widest sense.

By the middle of the 18th century, travelers to West Asia began to record the Arabic inscriptions on buildings they saw in the region. The first was the German Carsten Niebuhr, the sole survivor of a scientific mission sent by Frederick V of Denmark in 1761 to explore Egypt, Arabia, and Syria. Niebuhr recognized the importance of Arabic epigraphy, for he brought home with him part of the inscription from the Nilometer in Egypt, and his multi-volume publication of the expedition, entitled *Reisebeschreibung nach Arabien und andern umliegenden Ländern* (Copenhagen, 1774–1778), included drawings of many inscriptions. While this publication was rather piecemeal, Napoleon's expedition to Egypt between 1798 and 1801 carried out a more systematic survey of Egyptian inscriptions. The *Description de l'Égypte*, published between 1809 and 1828 in Paris, was one of the greatest achievements of the encyclopedic tradition of the French Enlightenment, and its ten elephant folios of plates contain reproductions of inscribed objects and monumental inscriptions, including the Nilometer and the long wooden frieze from the mosque of Ibn Ṭūlūn, as well as other texts, now destroyed. The French remained pioneers in the field of Islamic epigraphy; the Orientalist Joseph Touissant Reinaud's 1828 publication of the works of art in the collection of the Duc de Blacas, the first catalog of a collection of Islamic decorative arts, contained substantial work on epigraphy, including the first study of Islamic seals.

Only at the turn of the 19th to the 20th century, however, did the Swiss scholar Max van Berchem, founder and unsurpassed master of the field of Islamic epigraphy, establish a scientific basis for studying Arabic inscriptions. Van Berchem recognized that while any single inscription could be analyzed to furnish specific historical information, only a corpus of inscriptions could help discern the unusual from the standard. He thus initiated the compilation of a regional corpus of Arabic inscriptions. Entitled *Matériaux pour un Corpus Inscriptionum Arabicarum*, the series published the inscriptions from the central region of the Islamic lands, subdivided into Egypt, Syria, Asia Minor (Anatolia), and Arabia (van Berchem a.o. 1894–1985). These weighty tomes went far beyond van Berchem's initial schema for a mere compilation of inscriptions; they contained substantial commentaries on the epigraphy on buildings (and in some cases, objects) arranged chronologically. The lengthy commentaries, however, required time and effort, and only a handful of volumes were ever published. A decade after van Berchem's death, his colleagues Étienne Combe, Jean Sauvaget, and Gaston Wiet took up van Berchem's original idea of a chronological listing of datable historical inscriptions in Arabic. Entitled *Répertoire chronologique d'épigraphie arabe* (typically abbreviated RCEA), this work appeared sporadically throughout the next decades in 18 volumes and with a geographical index (Combe a. o. 1931–1991).

Since 1993, the concept of a chronological survey of Arabic epigraphy has been substantially updated with an electronic version on CD-ROM compiled under the direction of Ludvik Kalus and carried out by Frédérique Soudan. This database, *Thesaurus d'épigraphie islamique*, is designed to bring together inscriptions from the Muslim lands written in Arabic, Persian, and Turkish (as well as in other languages) up to the year 1000 A.H./1591 C.E. (Kalus and Soudan 1993). To date, five installments containing some 15,000 inscriptions have appeared. The first installment covers inscriptions from Morocco, Algeria, Tunisia, and Libya; the second, inscriptions from the Arabian Peninsula (Saudi Arabia, Yemen, Oman, Bahrain); the third, inscriptions in Arabic, Persian, and Turkish from Central Asia (Uzbekistan, Kazakhstan, Kyrgyzstan, Tajikistan, and Turkmenistan); the fourth and fifth, inscriptions from Egypt (funerary and monumental, as well as inscriptions on portable objects), along with updates for Central Asia, the Arabian Peninsula, and the Maghreb. The sixth and seventh installments covering the Indian world (Pakistan, India, Bangladesh, Sri Lanka, and the Maldives) are due out in 2005. The advantages of the electronic format are manifold and manifest, for it is possible to search by various criteria ranging from date, site, and type of inscription to support, material, and current location. By searching the Arabic text, it takes only seconds to locate every inscription containing a particular Arabic word.

3. CONTENT AND FORMAT OF HISTORICAL INSCRIPTIONS

Most Arabic inscriptions begin with the *basmala*, usually written out in its full form, *bismi llāhi r-raḥmāni r-raḥīm* 'In the name of God the Merciful and Compassionate', although sometimes shortened to *bismi llāh* 'in the name of God' where space was short. Locating the *basmala* is easy – it has a distinctive and visually recognizable form – but also important, particularly on circular inscription bands, for this phrase helps the reader to locate the rest of the information, notably the date, which typically occurs at the end.

The *basmala* is sometimes preceded by the rhyming phrase *ʾaʿūḏu billāh min aš-šayṭān ar-rajīm* 'I seek refuge with God from Satan the accursed'. This prefix is common in North Africa, probably because of the conservative affiliations of the patrons there who often belonged to the Maliki school of law. It is found, for example, on the superb *minbar* made in 532 A.H./1137 C.E. for the mosque built by the Almoravid ruler ʿAlī ibn Yūsuf in Marrakesh but later transferred to the Kutubiyya Mosque in the same city. This prefix is also found on madrasas throughout the Islamic lands, presumably as a reflection of the role these colleges played in disseminating the faith. It is used, for example, on the enormous band encircling the *qibla ʾīwān* in the funerary complex erected by Sultan Ḥasan in Cairo in the middle of the 14th century, one of the rare buildings in the Islamic lands that includes four madrasas, for the four major schools of law [Fig. 2].

The *basmala* is typically followed by a statement of commissioning, usually introduced by *mimmā ʾamara* or just *ʾamara* 'ordered', as on the al-ʾAqmar Mosque in Cairo. This opening verb is typically paired with *bi-bināʾ* 'the construction', *bi-ʿamal* 'the making' (used on the al-ʾAqmar Mosque), or *bi-inšāʾ* 'the establishment', a form increasingly common from medieval times onward. The verb is usually followed by the object, sometimes simply a pronoun *huwa* 'it', as on the al-ʾAqmar Mosque, but often a noun specifying the type of building, whether *masjid* 'mosque', *burj* 'tower', *sabīl* 'fountain', or the like, or the type of object, such as *minbar* 'pulpit', etc. At first, nouns were used alone, but by the 11th and 12th centuries, they were often accompanied by adjectives such as *mubārak* 'blessed' or *šarīf* 'noble'.

In a typical foundation inscription, more attention was given to the patron than to the building or object he commissioned, and over the centuries the names and titles of these people became increasingly lengthy. Typically the patrons, whether rulers or their courtiers, were glorified by a series of epithets, often composed in rhyming pairs. On the al-ʾAqmar Mosque, for example (RCEA 3011–3012), the foundation inscription begins with the titles and name of the Fatimid vizier who ordered the building. He is introduced as the servant (*fatā*) of the Fatimid caliph al-ʿĀmir, who is identified as "our lord and master, the imam al-ʿĀmir bi-ʾaḥkām Allāh, son of the imam al-Mustaʿlī billāh, Commander of the Faithful (*ʾamīr al-muʾminin*)". The names of al-ʿĀmir and al-Mustaʿlī are followed by a benediction invoking God's blessings on the two caliphs, their pure ancestors, and their honorable descendants (*ʾābāʾihim aṭ-ṭāhirīn wa-ʾabnāʾihim al-ʾakramīn*), a reference to the Sevener Shiʿite leanings of the Fatimid line, who claimed legitimacy as descendants of the Prophet through his daughter Fāṭima. The vizier himself is hailed as the commander of the army, the sword of Islam, the protector of the imam, the guarantor of the judges of Muslims, and the guide to the missionaries of believers (*ʾamīr al-juyūš, sayf al-ʾislām, nāṣir al-ʾimām, kafīl quḍāt al-muslimīn wa-hudā duʿāt al-muʾminīn*). He is then identified by his full name, including his patronymic (*kunya*), ʾAbū ʿAbdallāh; his proper name (*ism*), Muḥammad; and his epithet of affiliation (*nisba*) al-ʿĀmirī, indicating his relationship to the caliph al-ʿĀmir. Following the vizier's name is a benediction asking God to strengthen him in his faith, make the Commander of the Faithful benefit from prolonged life, extend his power, and raise his word. Such long titles and benedictions served to glorify the patron (and his overlord), who usually undertook the foundation as a pious act that would not only benefit the community but also count as a righteous deed on the Day of Judgment.

Objects, especially large and expensive ones, were often inscribed with similar inscriptions, although the verbs might vary. The most common term is *ʿamila* 'to make', but another option is *ṣanaʿa* 'to fashion', used, for example, on the *minbar* for the Kutubiyya Mosque in Marrakesh [Fig. 7]. Commissioning inscriptions on objects made in multiples, such as glass lamps, often open with the phrase *bi-rasm* 'made/intended for'.

On objects, this part of the inscription may be followed by the place of manufacture. The inscription on the Kutubiyya *minbar*, for example, specifies that it was made in Córdoba, followed by a benediction asking God to protect the city. The inclusion of such optional information signals its importance. In the case of the *minbar*, for example, its manufacture in Córdoba was significant, for the huge inlaid wood construction (it measures almost 4 meters high) had to be shipped some 800 kilometers down the Guadalquivir River, across the Straits of Gibraltar, and then carried by camels or mules over the Atlas Mountains to southern Morocco. Similarly, the benediction following the name of the city might simply be hyperbole, but it also alludes to the insecure political climate in Andalusia in the mid-12th century, when Córdoba changed hands repeatedly.

This epigraphic information can help us to localize craft industries. The inscriptions on a group of ivory boxes made in medieval Andalusia, for example, mention two different locales: Madīnat az-Zahrāʾ, the palace-city founded by ʿAbd ar-Raḥmān, is named on two boxes made in 355 A.H./966 C.E., and Cuenca, a city 300 kilometers northeast of Córdoba near Toledo in Castile, is named on three pieces made in the 11th century. Curiously, Córdoba, the city most often associated with these ivories, is not specifically mentioned in any of the inscriptions.

Historical inscriptions regularly end with the date. Most often, it is introduced by the phrase *fī sana* 'in the year', occasionally preceded by *bi-tārīx* 'dated to' or *fī šuhūr* 'in the months of'. Sometimes a specific month is given, occasionally qualified by an adjective (e.g. 'Ramadan the blessed', *al-mubārak*). Surviving inscriptions suggest that the more specific the information, the more unusual the commission. The inscription on the Kutubiyya *minbar*, for example, tells us that work was begun on 1 Muḥarram 532 A.H./19 September 1137 C.E. It must have been commissioned to celebrate the new year, which begins on 1 Muḥarram.

A building or an expensive object might also bear the name of its maker, who is usually identified by the introductory phrase *ʿamal* 'work of'. Artisans are the least important people named in historical inscriptions, for their names are either tacked on at the end of historical inscriptions or, more frequently, hidden elsewhere in the decoration. Furthermore, on objects they are often incised, a cheaper technique than the relief carving typically used for the main inscription. On the Kutubiyya *minbar*, for example, the craftsman's name 'al-ʿAzīz' is incised on the lower left frame, once hidden behind a carved capital. On the ivory boxes, the names of the artisans are sometimes tucked between the spokes of the hinge on the back.

Sometimes, the artisan's name is hidden in a specific place as a sort of visual pun. Thus, on a large gilt-silver box made for the neo-Umayyad heir-apparent ʾAbū Walīd Hišām in 366 A.H./976 C.E. (RCEA 1869) – a box whose shape shows that it is a copy of an ivory box – the two artisans' names are incised under the clasp. When the clasp was lifted to open the box, the craftsmen, who are identified in the inscription as the caliph's two servants, would have been under the thumb of the ruler, as it were.

Inscriptions on works of art are key in tracing the history of artisans and the organization of crafts, for such people are rarely, if ever, mentioned in texts. Thus, on a box in Pamplona made for the chamberlain ʿAbd al-Malik ibn Manṣūr, the largest and most splendid of all the medieval Andalusian ivories (RCEA 2098), an inscription incised on the inside of the lid records that it was the work of Faraj and his pupils or apprentices (*ʿamal Faraj maʿa talāmiḏihi*). Signatures of five individual craftsmen, each preceded by the word *ʿamal* 'work of', are incised in inconspicuous places in the various figural scenes on the different plaques composing the lid and base of the box. These signatures suggest that each craftsman made a separate plaque and the pieces then fitted together to form the large box.

The artisan's *nisba* 'epithet of affiliation' is sometimes a clue to the place of production, but the *nisba* can be a trademark or even the sign of a workshop. Thus, the *nisba* al-Mawṣilī 'from Mosul' came to indicate a standard of fine metalworking. It is found as part of the signature on at least 30 vessels of inlaid brass dating from the 13th to the early 14th century. At least one piece – the Blacas ewer in the British Museum – was made in Mosul (RCEA 4046), but the inscriptions tell us that others were made in Damascus or Cairo. Similarly, the *nisba* aš-Širāzī (from the city of Shiraz in southwestern Iran) was used by at least three architects working in Timurid Khurasan.

4. Religious inscriptions

Following van Berchem's initiative, scholars have concentrated most of their attention on studying historical inscriptions from the Islamic lands, but many inscriptions also contain other material that is typically, though somewhat artificially, lumped together in the category of religious texts. These inscriptions have sometimes been dismissed as unimportant, even banal, although they are far more numerous than historical texts and often longer and more prominent. In the last decades, however, scholars have begun to recognize the significance of these texts in throwing light on the context in which buildings and objects were made and to distinguish different types of text within this broad category.

Many of these inscriptions contain citations from the *Qur'ān*. Dodd and Khairallah's *Image of the Word* (1981) gathered 4,000 Qur'ānic inscriptions from van Berchem's corpus and other well-known sources into a monograph with an accompanying essay on the significance of Qur'ānic inscriptions on Islamic architecture. They indexed these inscriptions in three ways. The first index lists the citations numerically by *sūra* and verse. The second lists the citations geographically, by country, city, and building. The third lists the citations by building type (madrasa, mausoleum, mosque, and other). Although not without its problems and certainly in need of the electronic format now available for historical inscriptions, this work is the first place to begin any analysis of Qur'ānic inscriptions.

Scholars have just begun to exploit the potential of this rich material. Hoyland (2002), for example, surveyed the field of Qur'ānic epigraphy, with particular emphasis on the ways in which the Qur'ānic text was inscribed on various media (buildings, tombstones and rocks, objects and furnishings, coins, seals, and amulets). Blair and Bloom (2006) enumerated six principles that might govern the choice of a particular chapter or verse in a given situation. Three were general principles, ranging from pragmatic considerations, such as the space available for the inscription and function of the particular building or object, to the ideological goal of glorifying the faith. Three other principles were narrower in scope, adapted to specific historical situations: sectarian ideology, political and current events, and puns or plays on words. These principles are not exclusive but overlapping, and a particular text might be chosen for several reasons.

Coins, seals, and amulets, for example, are often inscribed with *ḥasbiya llāh* 'God is sufficient for me', found in Q. 9/129 and 39/38, a short text that summed up the believer's faith. Q. 9/18, stating that the person to maintain God's mosques is he who believes in God, prays, and gives alms, is the most common verse inscribed on mosques, used four times more frequently than any other Qur'ānic inscription. One of three Qur'ānic texts that refer specifically to God's mosques (*masājid Allāh*), it is the only verse that refers specifically to the duties of Muslims worshipping inside them. This verse was particularly popular in Fatimid times, probably because it includes the word *muhtadin* 'the guided', and was therefore doubly suitable for the Fatimids, descendants of the *mahdī* 'the right guide'.

Parts of buildings were also distinguished by particular Qur'ānic texts. *Miḥrāb*s, for example, are typically adorned with a verse that includes the word for ritual prayer (*aṣ-ṣalāt*), not surprisingly because the term occurs 67 times in the Qur'ānic text. The verse most commonly used is Q. 17/78, in which the believer is enjoined to perform prayer (*aṣ-ṣalāt*) from the setting of the sun to the darkness of the night, as well as the dawn recitation of the *Qur'ān*, for that action is particularly attested. This verse is used, for example, to frame the stunning *miḥrāb* made by Ḥasan ibn ʿArabšāh in 623 A.H./1226 C.E. for the Maydan Mosque in Kashan, the largest and most elaborate example of luster tiling from medieval Iran (RCEA 3961). The reasons for the choice of verse are clear. It is one of only two indisputable instances in the *Qurʾān* – the other (Q. 75/16–18) is not suitable to inscribe on a *miḥrāb* as it refers to moving the tongue – in which the word *qurʾān* functions as a gerund denoting an activity ('reciting') rather than a noun denoting an action ('recitation'). This particular verse was chosen because it conveys the verbal force of Qurʾānic reciting.

In addition to Qurʾānic texts, some inscriptions contain *ḥadīṯ*. Only a very few of these prophetic traditions are found in Wensinck's (1936–1988) extensive concordance compiled from canonical sources. Rather, the *ḥadīṯ*s inscribed on buildings and objects reflect more popular traditions. Some may even have been coined for the occasion and thus provide a rare source for popular religion. The earliest example of an epigraphic *ḥadīṯ*, for example, is the one found on the *minbar* donated in 484 A.H./1081 C.E. by the Fatimid general Badr al-Jamālī to the sanctuary built to hold the miraculously discovered head of Ḥusayn in Ascalon and later moved to the Haram in Hebron (RCEA 2791). The tradition inscribed there, in which the Prophet declares his two legacies to be the *Qurʾān* and his family, vindicates Shiʿite claims to legitimacy and justifies the shrine. Similarly, the doorway to the Shah-i Zinda, the cemetery that grew up outside Samarqand around the tomb of Quṯām ibn ʿAbbās, is inscribed with a *ḥadīṯ* that the person most like the Prophet in character and appearance is Quṯām. In frontier regions such as Central Asia, such inscriptions were drawn from a wide, even eclectic, range of sources. One of the tombs in the Shah-i Zinda, for example, is inscribed with a text from Socrates.

Many portable objects are inscribed with superogatory prayers (*duʿāʾ*) that ask for God's blessings on the owner. These inscriptions typically begin with the phrase *baraka min Allāh li-*... 'blessing from God on...' but can also include other nouns such as *ġibṭa* 'happiness', *surūr* 'joy', *saʿāda* 'felicity', *niʿma* 'favor', and the like. The blessing can be invoked on an anonymous owner (*li-ṣāḥibihi*), but the inscriptions, particularly in the case of luxury objects like the ivory boxes produced in medieval Andalusia, can also carry the name of the recipient. Seven of the ivory boxes were made as gifts for women, who are identified not by their names but by their connections to the neo-Umayyad patrons who commissioned these expensive works of art. Three were given to an unidentified daughter of ʿAbd ar-Raḥmān II and four to al-Ḥakam II's consort, known from written sources as the infamous Ṣubḥ. Inscriptions thus furnish a supplementary, and often unexploited, source for studying history, as these women did not always appear in texts, which were traditionally written by men.

Personal statements drawn from a religious vocabulary were often inscribed on tombstones and rocks. Many texts seek God's forgiveness, approval, mercy, and the like. They typically begin by invoking God's name (*Allāhumma*) or invoking a blessing (*ṣalli ʿalā*... lit. 'pray for...'). Some inscriptions entreat God to admit the inscriber to paradise or to reward him for his good deeds. Other such inscriptions on tombstones and rocks were intended to convey the fundamentals of personal faith. The inscriber often desired to be reunited with the Prophet or to be instructed in God's proof, references to well-known Qurʾānic phrases. Such inscriptions often invoke God's name using Qurʾānic epithets such as 'the Clement', 'the Praiseworthy', 'the Glorious', or 'the Knowing'.

Another category of inscription drawn from the Qurʾānic repertory comprises the 99 Beautiful Names of God (*al-ʾasmāʾ al-ḥusnā*). These names are not found in a single specific place in the *Qurʾān*, but lists of them were often inscribed in tiny script on amulets, particularly those made in later times of semiprecious stones like carnelian or nephrite. The names were typically invoked with the vocative *yā* 'O!' and were sometimes paired with the parallel construction of the 99 Noble Names of the Prophet (*al-ʾasmāʾ aš-širāf*). The latter, however, are not invoked with 'O!' but rather followed by the letter *ṣād*, an abbreviation for the *taṣliya*, the phrase *ṣallā llāhu ʿalayhi wa-sallama* 'may God bless him and give him peace!' that should follow every mention of the Prophet in written or oral discourse.

Sets of these sacred names were also made into calligraphic pictures. The Muḥammadan Rose, for example, is a floral design containing the 99 Beautiful Names of God, the 99 Noble Names of the Prophet, and the names of the Ten to whom Paradise was promised (*al-ʿašara al-mubaššara*), a phrase that turns up on architecture as well. Other phrases favored by Sunnis invoke the names of the Four Rightly Guided Caliphs (*al-xulafāʾ ar-rāšidūn*): ʾAbū Bakr, ʿUmar, ʿUṯmān, and ʿAlī. Shiʿites inscribed the names of the Fourteen Immaculate Ones (Persian *čahardum maʿṣūm*): Muḥammad, his daughter Fāṭima, and the twelve imams who succeeded him. These names attest to confessional allegiances, and when the religious affiliations of a region changed, the inscriptions were often altered. In Iran, now a predominantly Twelver Shiʿite country, for example, the names of the first three of the Rightly Guided Caliphs have been scratched out, leaving only the name of ʿAlī, whom Shiʿites revere as the rightful successor to Muḥammad. These religious inscriptions, then, like the historical ones, can be important documents in charting the religious, social, and intellectual history of Islam.

BIBLIOGRAPHICAL REFERENCES

Berchem, Max van a.o. 1894–1985. *Matériaux pour un Corpus Inscriptionum Arabicarum*. Cairo: Institut Français d'Archéologie Arabe au Caire.

Blair, Sheila S. 1998. *Islamic inscriptions*. Edinburgh: Edinburgh University Press.

——. 2005. "What the inscriptions tell us: Text and message on the ivories from al-Andalus". *Journal of the David Collection* 2.75–100.

——. and Jonathan M. Bloom. 2006. "Inscriptions on art and architecture". *Cambridge companion to the* Qurʾān, ed. Jane McAuliffe, 000. Cambridge: Cambridge University Press.

Combe, Étienne, Jean Sauvaget, and Gaston Wiet. 1931–1991. *Répertoire chronologique d'épigraphie arabe*. 18 vols. Cairo: Institut Français d'Archéologie Arabe au Caire.

Dodd, Erica Cruikshank and Shereen Khairallah. 1981. *The image of the Word: A study of Qurʾānic verses in Islamic architecture*. 2 vols. Beirut: American University of Beirut.

Hawary, Hassan a.o. 1932-1942. *Les stèles funéraires*. 10 vols. Cairo: Institut Français d'Archéologie Arabe au Caire.

Hoyland, Robert. 2002. "Epigraphy". *Encyclopaedia of the* Qurʾān, ed. Jane McAuliffe a.o., II, 25–43. Leiden: E.J. Brill.

Kalus, Ludvik and Frederique Soudan. 1993–. *Thesaurus d'épigraphie islamique*. Geneva: Fondation Max van Berchem. (CD-ROM.)

Sharon, Moshe. 1997–. *Corpus inscriptionum arabicarum Palaestinae*. Leiden: E.J. Brill. [3 vols. published to date.]

Wensinck, Arent Jan. 1936–1988. *Concordance et indices de la tradition musulmane*. Leiden: E.J. Brill. (Repr., 1992.)

SHEILA S. BLAIR (Boston College)

Figure 1. Mosaic inscription inside the Dome of the Rock, Jerusalem (Photo Credit: Sheila Blair and Jonathan Bloom).

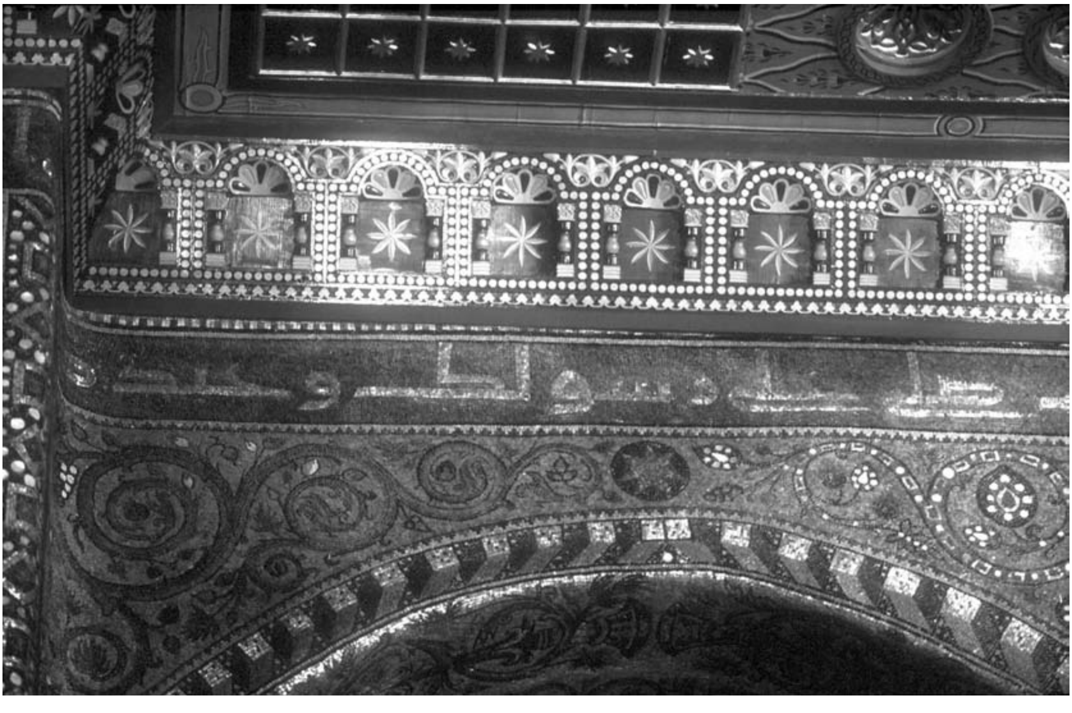


Figure 2. Beginning of the Qurʾānic text on the *qibla ʾīwān* in the Mosque of Sultan Ḥasan, Cairo (Photo Credit: Sheila Blair and Jonathan Bloom).

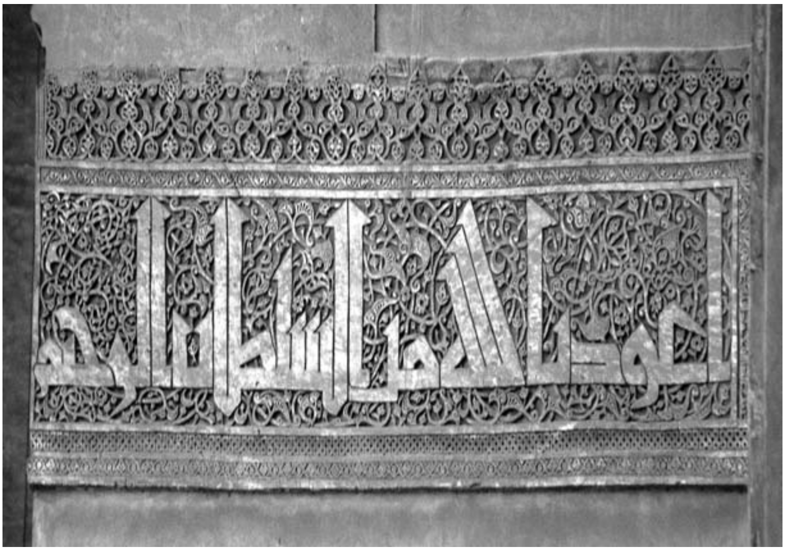


Figure 3. *Ṭirāz* textile made for the Caliph al-Mustanṣir (Photo Credit: V&A Images/Victoria & Albert Museum. Museum reference number: 1381–1888).

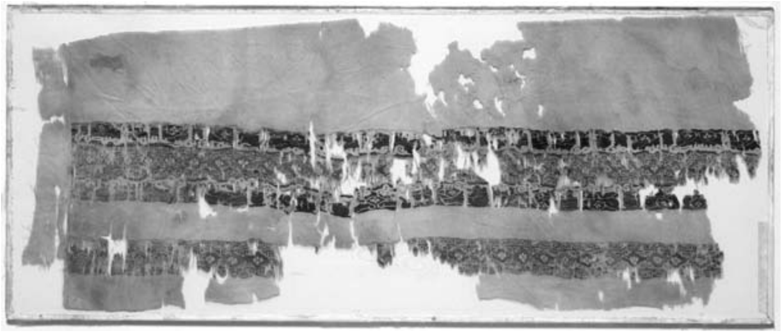


Figure 4. *Ṭirāz* band on the façade of the complex of Qalāʾūn, Cairo (Photo Credit: Sheila Blair and Jonathan Bloom).

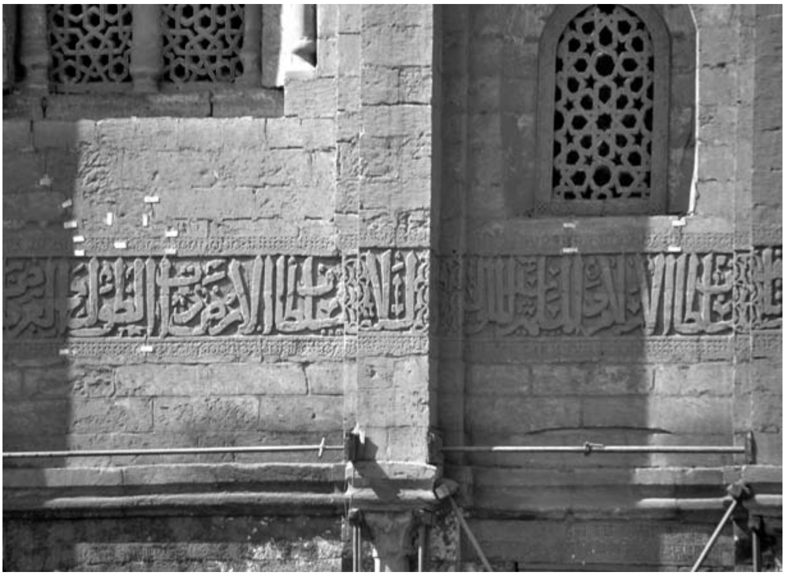


Figure 5. Façade of the al-ʾAqmar mosque, Cairo (Photo Credit: Sheila Blair and Jonathan Bloom).

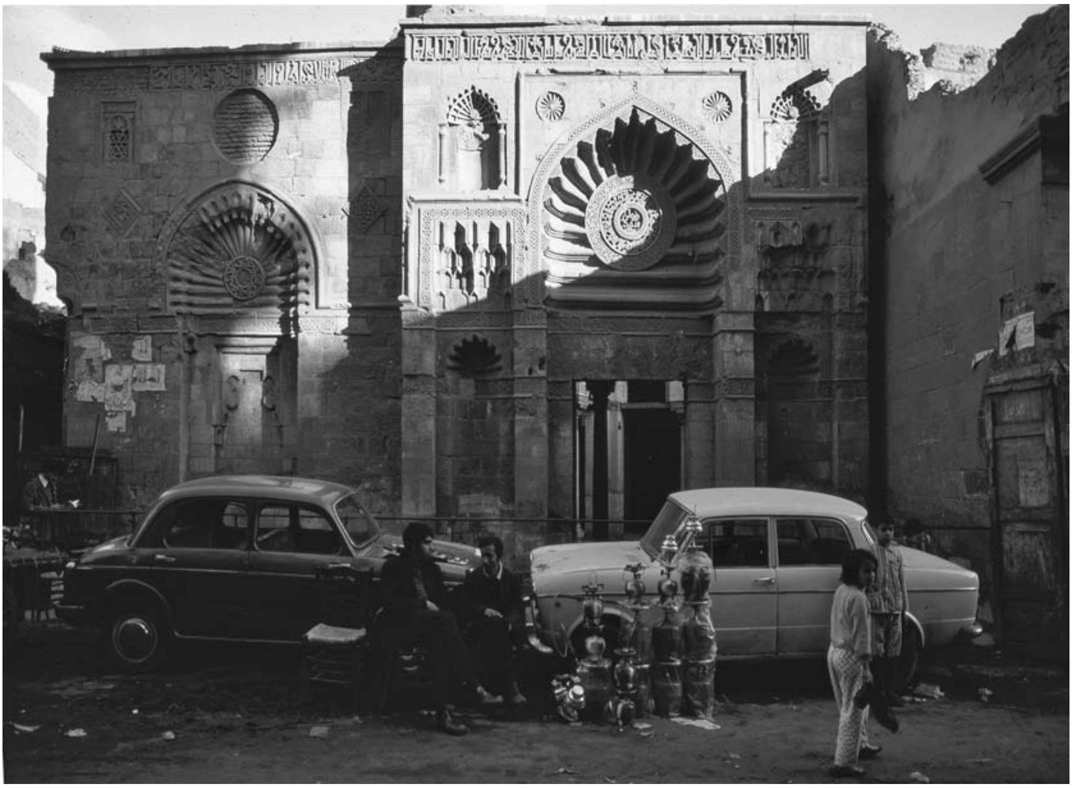


Figure 6. Roundel on the façade of the al-ʾAqmar mosque, Cairo (Photo Credit: Sheila Blair and Jonathan Bloom).

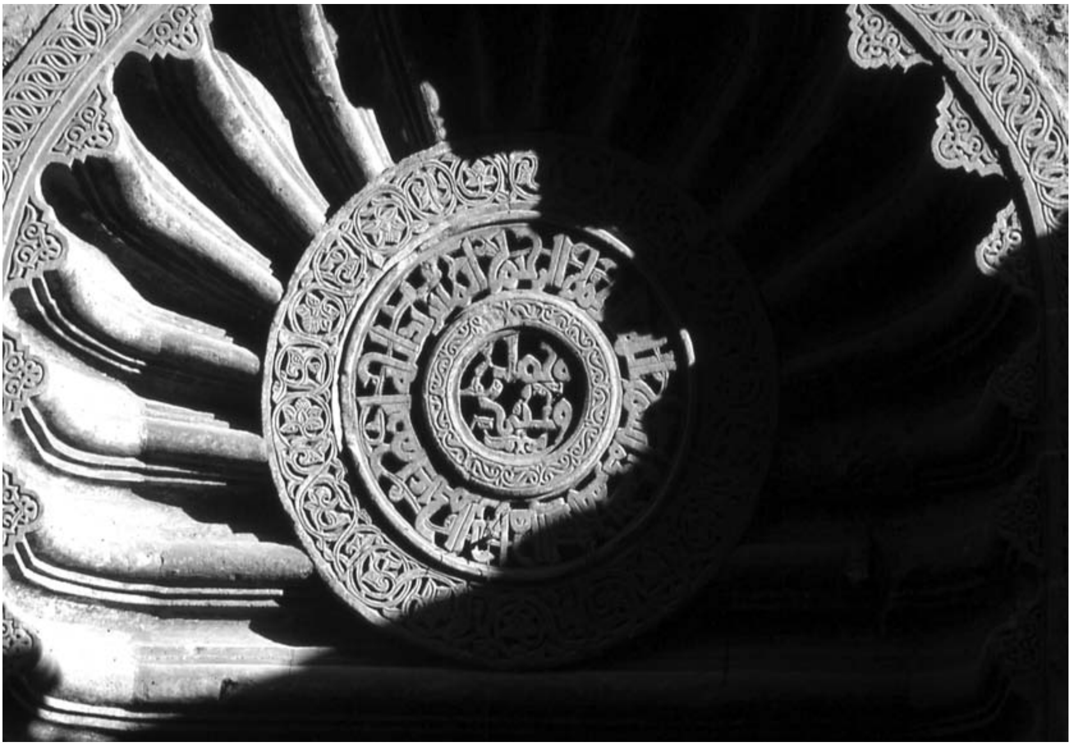
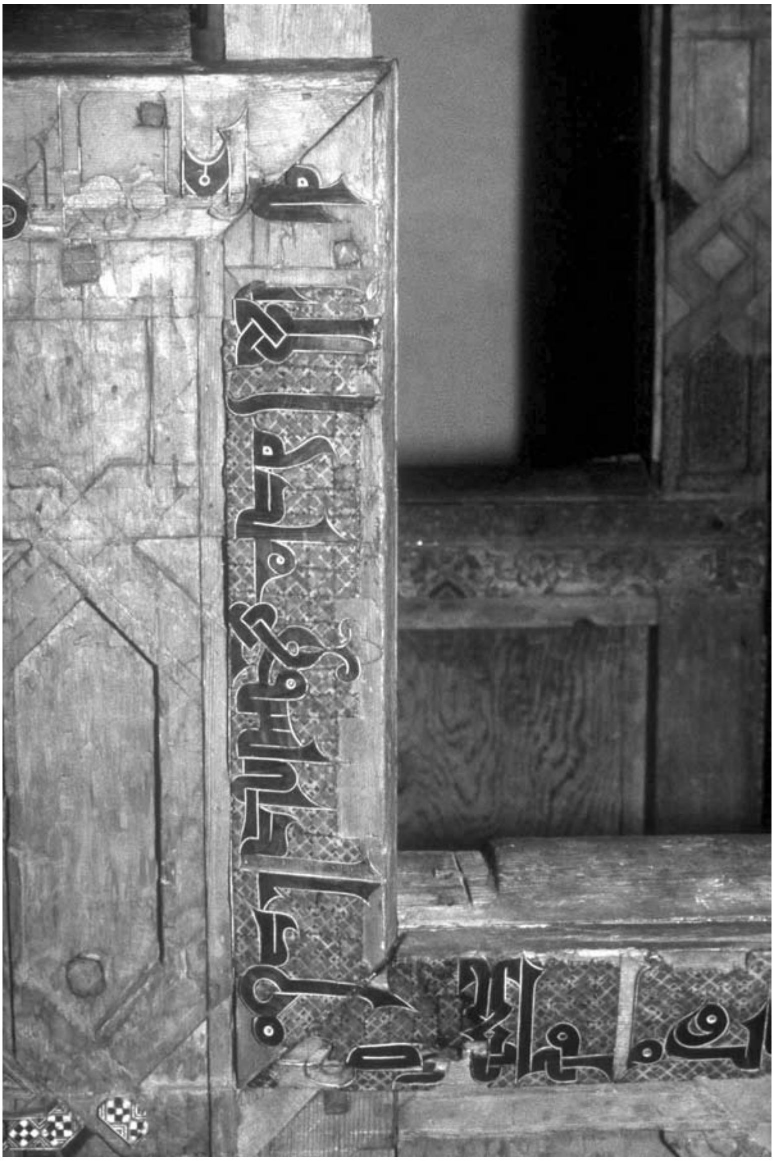


Figure 7. Qurʾānic band on the minbar for the Kutubiyya Mosque, Marrakesh (Photo Credit: Sheila Blair and Jonathan Bloom).



Equative Clause → Copula

Eritrea → Djibonti/Eritrea

Ethiopia

1. ARABIC IN ETHIOPIA

Since ancient times Arabic has played an important role in the social and literary life of Ethiopia. The geographic proximity between the Horn of Africa and the Arabian Peninsula has encouraged cultural and linguistic contacts between the two sides of the Red Sea. As a result, the Arabic language has had a considerable influence on the languages and literature of Ethiopia. Until now, though, only a limited amount of the available data has been studied.

Although Ethiopia is generally known as a Christian country, it is nevertheless host to a rich Islamic culture with a long history. Consequently, Arabic influence is reflected both on the literary level and in everyday life, as in any other Muslim society. The predominantly Muslim areas of Ethiopia are eastern Ethiopia, with the old city of Harar as its cultural center, most parts of the Oromo region in the East and Southeast as well as in the Southwest around the city of Jimma, the eastern part of Wällo in the Northeast, the Somali and Afar regions, parts of the Gurage-Səlṭe region, and the Beni Shangul region in the Far West, at the Sudanese border.

It has been claimed (Ferguson 1970) that Arabic functions as a trade language in Ethiopia, but sociolinguistic investigations (Cooper and Carpenter 1976) do not support this claim. In fact, there exist different spheres of spoken and written Arabic in Ethiopia. First, there are Arabic-speaking immigrant communities, mostly from Yemen, and traders from the Gulf States, Saudi Arabia, and Yemen who regularly come for commercial purposes. Another group are foreign, Arabic-speaking members of *daʿwā* groups who travel around the country, teaching the local Muslim population about the tenets of their religion. Ethiopians who have worked in Arab countries, such as Saudi Arabia, Lebanon, or Yemen, often have some knowledge of spoken Arabic, depending on the length of their stay. Many inhabitants of the Beni Shangul region are bi- or multilingual in Sudanese Arabic and Ethiopian languages.

Furthermore, as in the entire Islamic world, Arabic is the religious language of the Ethiopian Muslims. As children they learn to read and pronounce the signs of the Arabic writing system in *Qurʾān* schools. Although most do not really master the language, some continue their study of Arabic in modern or traditional centers of advanced Islamic learning.

2. CHRISTIAN LITERATURE

The Axumite Kingdom in northern Ethiopia became a Christian country in the 4th century C.E. Its church belonged to Orthodox Christianity and had close connections to Coptic Egypt. Although the Ethiopian Church was theologically independent, its patriarch used to be an Egyptian, sent by the Coptic patriarch in Alexandria.

The Classical Ethiopian language Gəʿəz continued to survive as the literary and liturgical language of the Ethiopian Church. After the rise of the Solomonic dynasty in the 13th century, cultural activities increased, contacts with Egypt were intensified, and many Arab monks, craftsmen, and merchants came to Ethiopia. The Arabic language became the medium of communication of the Ethiopian Christian rulers with their Arab neighbors.

The lasting Arabic linguistic influence on Ethiopian languages of that time can be observed in the literature. The first attested contact between Arabic and Ethiopian languages falls in that period. Many theological works of the Ethiopian Orthodox Church were translated from Arabic into Gəʿəz. In addition to works with religious content, many historical and hagiographical treatises were translated as well. In the process of translation, this literary variety of Gəʿəz was heavily influenced by Arabic. Because Classical Arabic and Gəʿəz have many structural similarities, word-by-word translations seemed to be the easiest strategy for the translators, many of whom were Arabs.

On the lexical level, this resulted in an increase of Arabic loanwords; on the syntactic level, it led to new morphosyntactic structures which were not typical for Gəʿəz and original Arabic (cf. Kropp 1986). It is not easy to give examples of Arabic loanwords which were incorporated into Gəʿəz through direct contact, because the incorporation must have taken place at a time when the language was still spoken, i.e. sometime before the 10th century C.E. In fact, this has to remain speculative, since most Gəʿəz literature was produced after the language had died out as a spoken language and is based mainly on Arabic literature (Kropp 1986:315). Leslau (1990:59) supposes that the language "incorporated into its vocabulary words of the spoken Arabic dialects", but he deals only with the phonetic correspondences between the Arabic loanwords in Gəʿəz and their Arabic etyma. The extensive list of Arabic loanwords at the end of Leslau's (1990) article is structured according to semantic domains but does not mention the origin of individual words. In his article he even contradicts himself by saying that "the Arabic loanwords were taken over in the literary language of Geez and were not adopted by the spoken language" (Leslau 1990:69). Kropp (1986:328) gives some examples of loanwords from Christian Arabic literature: *baʿ* 'ba [length measure]', *ḥəṣn* 'fortress', *sahəl* 'coast', *(əl)-məʿtəzəla* 'secessionist, rebel'.

3. Islamic literature and Classical Arabic

The first contact between Ethiopia and Islam dates back to the time of the Prophet Muḥammad, when a number of his followers found refuge at the Axumite court in northern Ethiopia. However, relatively little is known about the early propagation of Islam among Ethiopians. It is assumed that Islam was spread to the local population by Arabian travelers who had crossed the Red Sea, mainly merchants but also learned men.

The presence of Islamic statehood in Ethiopia beginning in the 9th century C.E. indicates an early use of Arabic in Ethiopia. There are, however, few written sources from that period. In Eastern Shoa, the region neighboring the Christian kingdom to the southeast, Arabic inscriptions on ruins of a mosque give the year 171 A.H. as the date of its construction (*Hawwatoota Godina Oromiyaa* 2000:16).

The first Muslim state in Ethiopia was the Sultanate of Shoa, founded in the year 896 C.E. by the Maxzūmī dynasty. An Arabic document gives some historical accounts of that sultanate (Cerulli 1941). In the following centuries, more Muslim states emerged in the east and south of Shoa (Braukämper 2004). In 1285 C.E. Yīfāt (or Awfāt), which was ruled by the Wālašmaʿ (or Wālasmaʿ) dynasty (cf. Wagner 1976), attacked Shoa and ousted its Maxzūmī ruler. Subsequently, Yīfāt became the most powerful Islamic state in Ethiopia by conquering neighboring Islamic principalities. After years of fierce wars with the Christian kingdom, Yīfāt was defeated in 1415 C.E., and the Wālašmaʿ rulers moved to the east and made the city of Harar the new capital of their kingdom, now called Adal. With the exception of Harar, most of the Islamic principalities in southern and eastern Ethiopia disappeared during a devastating war in the 16th century between the Christian kingdom and Adal and the subsequent Oromo migration. As a result, all the Islamic entities of southern and eastern Ethiopia were extinguished or dramatically reduced. With our current state of knowledge, it is impossible to draw any conclusions about the knowledge of Arabic among the population of these territories prior to these events.

Harar, however, remained a major center of Muslim scholarship in Ethiopia. Being the language of Islam, Arabic was used in religious teaching, liturgy, and administration. But it also functioned as a literary language in a wider sense. The bookbinding tradition, which was remarkably sophisticated, illustrates the appreciation of literature by Harar's population.

The Arabic literature found in Ethiopia can be classified into two types: works originally composed by Arab authors and works written by indigenous scholars. The level of competence in Classical Arabic can be measured by the various genres of Arabic literature imported to Harar. This imported Arabic literature comprises "a fair number of standard works by Arab authors as well as some of the classics of Arabic religious poetry and pious literature, such as Ibn Mālik's *ʾAlfiyya*, Ibn Ḥajar's *Tuḥfa*, Būṣīrī's *Burda* and *Hamziyya* and Jazūlī's *Dalāʾil al-xayrāt*" (Drewes 1976:174).

Most of the known and documented indigenous literary Arabic works were written in Harar. Two authors from the 18th century were ʿAbd al-ʿAzīz ibn ʾAmīr Hāšim and Ḥāmid ibn al-Faqīh Ṣiddīq al-Hararī al-Ḥimyarī (cf. Brunschvig 1974). A large number of manuscripts from Harar were collected and documented by European scholars (cf. Drewes 1983; Wagner 1997).

The life of the most important saint of Harar, Šayx ʾAbādir ʿUmar ar-Riḍā, who came to Harar from the Ḥijāz in the early 13th century, is described in a 19th-century manuscript *Fatḥ madīna Harar al-ʿulyā fī sīrat al-mujtahidīn al-ʾawliyāʾ* (Wagner 1978). Furthermore, the indigenous Arabic literature comprises mainly genealogies of the Harari rulers from the 13th century until the late 19th century (Wagner 1973, 1974a, 1974b, 1991) and songs composed in praise of the Prophet as well as local saints (Wagner 1975).

In the 18th century, centers of Islamic scholarship were also established in the eastern part of Wällo. This region was in direct contact with neighboring Yīfāt and was probably Islamized during the same period. In the 18th century, the Qādiriyya *ṣūfī* order, having been introduced in Wällo from Harar, played a crucial role in the establishment of Islamic centers (cf. Hussein 1988, 2001). These centers of learning were founded by pious individuals and supported by local dynasties. Most of them were located in the narrow lowland area between the highland in the west and the desert in the east. Important locations were Anna in Rayya (northern Wällo), Dana in Yäjju, and Gäddo and Shonke in Däwwe (in southern Wällo). Some of the most influential scholars from Wällo were Muftī Dāwūd (late 18th century) from Gäddo, Šayx Muḥammad al-ʾAnnī from Anna and Šayx ʾAḥmad b. ʾĀdam from Dana (both late 19th century), and Šayx Jawhar ibn Ḥaydar (early 20th century) from Shonke. They taught various fields of Islamic learning and played a crucial role in the revival of Islam in that region of Ethiopia. Furthermore, the scholarly reputation of the *ʿulamāʾ* from Wällo was such that students from other Muslim regions, like Gurage in central Ethiopia and Jimma in the southwest, came to Wällo for higher education.

Many of the Wällo scholars had studied in the Ḥijāz and Yemen. When they returned to Ethiopia, they brought Arabic literature of various genres to Ethiopia. As in Harar, they copied these books for educational purposes. Muftī Dāwūd from Däwwe, for example, produced many copies of well-known books, which today can only be found in private collections. But in addition to this, Wällo scholars composed a large number of works in Classical Arabic. This literature still remains to be studied. It comprises mainly songs and poems in praise of the Prophet, sung during *mawlid* celebrations.

4. ISLAMIC LITERATURE: *AJÄM*

Though nominally Muslim, the majority of the population did not know very much about the basic tenets of their faith. The need to teach and inform the population in its own language was an important concern of the local *ʿulamāʾ*. This was their motivation to compose and write works in Ethiopian languages. Except for the case of Harari (cf. Wagner 1983a), this Islamic literature in Ethiopian languages remains virtually unnoticed. There are works at least in the languages Amharic (cf. Pankhurst 1994), Argobba, Oromo, and Səlṭe (Wagner 1983b). Literature in other Ethiopian languages may also exist (in Afar, Tigrinya, and others). This literature is called *ajäm* in Amharic and is written in the Arabic writing system. In Ethiopia, knowledge of this literature and particularly the application of Arabic script to Ethiopian languages is almost nonexistent. Amharic, the official language of Ethiopia, is usually written with the Ethiopic writing system. However, this script was always regarded as a Christian script and therefore unacceptable to Muslim scholars, whose primary literary language was Arabic. Arabic was the familiar writing system for the authors of the *ajäm* literary works.

Some phonological similarities between Ethio-Semitic (Amharic, Harari, Səlṭe) and Cushitic languages (Oromo) on the one hand and Arabic on the other hand facilitated the application of the Arabic script to Ethiopian languages. Ejective stops, characteristic for the Ethiopian linguistic area, are represented in Arabic script by etymologically corresponding signs: *q* (IPA: kʾ) by ق and *ṭ* (IPA: tʾ) by ط. Only a few consonants do not exist in Arabic and had to be represented by modified letters (Wetter, forthcoming). For Amharic these are the consonants *č*, *č̣*, *ž*, *g*, and *ñ*, and for Oromo the implosive *ḍ*.

Table 1 shows the modified Arabic signs used in Amharic, Argobba, Oromo, and Səlṭe *ajäm*. The consonants *ñ* and *g* have two regional variations. The fourth column contains a description of every modified sign.

Table 1. Modified letters in Arabic script for Ethiopian languages

Ethiopian consonant	IPA representation	*Ajäm* representation	Description
č	tʃ	ش	4 points above *sīn*
č̣	tʃʾ	ڟ	3 points above or under *ṭāʾ*
ž	ʒ	ژ	3 points above *rāʾ*
g	g	چ or ڭ	3 points above *kāf* or under *ḥāʾ*
ñ	ɲ	ني or ۑ	3 points under *yāʾ* or *nūn* + *yāʾ*
implosive d	ɗ	ڎ	3 points above *dāl*

Most of the *ajäm* works in Amharic are composed in various forms of rhyme. The content of these poems is usually religious (for the content of *ajäm* poetry see Pankhurst 1994). This is also the reason why the texts contain more Arabic loanwords than does the spoken language. One of the most important authors of Amharic *ajäm* was the scholar Ṭalḥa Jaʿfar from Argobba in eastern Wällo, whose works still await scientific documentation (cf. Hussein 1989).

While this literature is more traditional, there evolved a modern Islamic literature consisting mostly of theological works translated from Arabic into Amharic and Oromo. An Arabic newspaper (*al-ʿAlam*) has been published by the government since 1942 (Hussein 1994). After 1991, when the Socialist government fell, Islamic newspapers and magazines started to appear.

5. LEXICAL INFLUENCES

The most significant influence of Arabic on Ethiopian languages can be observed on the lexical level. Many Ethiopian languages have a considerable percentage of Arabic loanwords. The incorporation of Arabic loanwords took place in different periods and from different sources. There are differences in number and type of Arabic loanwords according to religious orientation of the respective speaker community. This is the case of Amharic in particular, with Christian and Muslim speaker groups.

An additional distinction is the way loanwords found their way into the various Ethiopian languages. Some Arabic loanwords may have been incorporated into Gəʿəz through language contact with spoken varieties of Arabic before Gəʿəz died out as a spoken language around the 10th century. However, many loanwords were incorporated during the translation of Christian Arabic literary works, when Gəʿəz was merely a literary language. In many cases, one Arabic loanword can be found in many Ethiopian languages, making it difficult to draw conclusions about the way these specific loanwords were incorporated into the individual languages. Some of the loanwords belonging to the more literary language doubtlessly found their way subsequently from Gəʿəz into other languages of Christian Ethiopia, e.g. Amharic and Tigrinya.

A higher number of Arabic loanwords can be observed in languages spoken by Muslim populations. These languages are Harari, the languages of Harar, Afar, and Somali, and a number of Gurage languages like Səlṭe, Argobba, and Oromo, but also regional varieties of Amharic, which are spoken in the predominantly Muslim areas of Wällo. As expected, many of the Arabic loanwords in these languages are semantically connected to a religious context. But there are also many words derived from Arabic that belong to daily life.

As examples of Arabic loanwords in the living languages, the following Argobba words may be cited:

[ʕɑruz]	'groom'	< Arabic *ʿarūz*
[dʒɨsm]	'body'	< Arabic *jism*
[dʒɑhil]	'illiterate'	< Arabic *jāhil*
[sobbɨr]	'patience'	< Arabic *ṣabr*
[sobiy]	'small child'	< Arabic *ṣabiyy*

[tɛʕziyɑ]	‘mourning’	< Arabic *taʿziya*
[t’ɑhɨr]	‘clean’	< Arabic *ṭāhir*
[xɑddɛm]	‘to serve’	< Arabic *xadama*
[xɛttɛw]	‘to write’	< Arabic *xaṭṭ* ‘script’
[zeyyɛr]	‘to pay respect’	< Arabic *ziyāra* ‘visit’

These loanwords occur mostly in such domains as religion, scholarship, social relations, and law, among Muslim populations like the Argobba, Muslim Amharic speakers, Harari, and Səlṭə, but also in Cushitic languages like Afar, Oromo, and Somali. Leslau (1990:9) assumes that Arabic loanwords in spoken Ethiopian languages “were taken from a spoken language and not from classical Arabic” as in Gəʿəz, but some of the loanwords in the Muslim areas must have had their origin in the written Arabic that Islamic scholars used to read.

The phonological integration of Arabic loanwords in Ethiopian languages depends on the structure of the language concerned. Some languages or varieties, for example, have pharyngeal consonants, while some do not, e.g. Argobba *duʿa* and Amharic *du’a* ‘special prayer’. Arabic consonants that do not exist in Ethiopian languages are usually represented by consonants with similar place of pronunciation and articulation (see also Leslau 1990:9), as in the following:

Arabic	Amharic
/ʾ/	zero or /ʾ/
/ʿ/	zero or /ʾ/
/ḏ/	/z/
/ḍ/	/d/, /t’/
/ġ/	/k’/, /k’^{w}/
/ḥ/	/h/
/x/	/k/, /k^{w}/, but also /x/ among Muslim Amharic speakers
/q/	/k’/, /k’^{w}/
/ṣ/	/s/, /s^{w}/
/ṭ/	/t’/

Arabic loanwords are incorporated into the morphological system of the Ethiopian language concerned. In Amharic, many verbs of Arabic origin seem to be derived from Arabic nouns, e.g. *zäyyärä* ‘to pay respect’ < Arabic *ziyāra* ‘visit’, rather than from *zāra* ‘to visit’. In the case of other verbs the origin is not so clear, e.g., the verb *kättäbä* ‘to write’ could have either a verbal or a nominal origin. The verbs take the Amharic morphological structure C1äC2C2äC3- (perfective), -C1äC2C3- (imperfective), and their derived forms are like those of any other Amharic verb, e.g. *täkättäbä* ‘to be written’. Sometimes Arabic plural forms are interpreted as singular and take an Amharic plural suffix *-očč*, e.g. *mäla’ik-očču* ‘his angels’ (< Arabic *malāʾika* ‘angels [pl.]’).

The case of Amharic in particular shows that speakers of Amharic living in Muslim areas tend to use Arabic loanwords even when there is an adequate Amharic term available. Sometimes entire Arabic phrases are mixed with Amharic syntax. This has given rise to a new idiom (Drewes 1976:194), a variety based on religious orientation that has some characteristics of a mixed language. The following lines by She Bashir Umar from Dällämäle in Wärrä Babbo are an example of Amharic *ajäm* poetry:

yämmiša yä’iman yäsləmənnan nägär
läk’ärraw täkätb^{w}*all bäññaw annägaggär*
‘the one who wants [to know about] the matter of faith and Islam
for the one who reads it it has been written in our language’

Arabic loanwords in this fragment include *läk’ärraw* (*k’ärra*) ‘he read, recited’ < Arabic *qaraʾa*; *täkätb*^{w}*all* (*täkättäbä*) ‘it was written’ < Arabic *kitāb* or *kataba*.

BIBLIOGRAPHICAL REFERENCES

Braukämper, Ulrich. 2004. *Islamic history and culture in southern Ethiopia*. Münster: Lit-Verlag.

Brunschvig, Robert. 1974. “L’Islam enseigné par Hamid b. Siddiq de Harar (XVIII^{e} siècle)”. *IV Congreso Internazionale di Studi Etiopici*, I, 445–454. Rome: Accademia Nazionale dei Lincei.

Cerulli, Enrico. 1941. “Il Sultanato dello Scioa del secolo XIII”. *Rassegna di Studi Etiopici* 1.5–42.

Cooper, Robert L. and Susan Carpenter. 1976. “Language in the market”. *Language in Ethiopia*, ed. M. Lionel Bender, J. Donald Bowen, R.L. Cooper, and Charles A. Ferguson, 244–255. Oxford: Oxford University Press.

Drewes, Abraham Johannes. 1976. *Classical Arabic in central Ethiopia*. Leiden: E.J. Brill.

——. 1983. “The library of Muḥammad b. ʿAli b. ʿAbd al-Shakūr, sulṭān of Harar, 1272–92/1856–75”. *Arabian and Islamic studies: Articles presented to R.B. Serjeant*, ed. Robin Leonard Bidwell and Gerald Rex Smith, 68–79. London and New York: Longman.

Ferguson, Charles A. 1970. “The role of Arabic in Ethiopia: A sociolinguistic perspective”. *Languages and Linguistics Monograph Series* 23.355–368.

Hawwatoota Godina Oromoiyaa. 2000 (?). Kemise: Bulchiinsa Godina Oromiyaa. [A publication of the Office for Tourism and Culture of the Oromiya Zone; in Kamisee.]

Hussein Ahmed. 1988. "Traditional Muslim education in Wallo". *Proceedings of the Ninth International Conference of Ethiopian Studies*, ed. Anatoly A. Gromyko, II, 94–106. Moscow.

——. 1989. "The life and career of Shaykh Talha B. Jaʿfar (c. 1854–1936)". *Journal of Ethiopian Studies* 22.13–30.

——. 1994. "*al-ʿAlam*: The history of an Ethiopian Arabic weekly". *Proceedings of the Eleventh International Conference of Ethiopian Studies*, ed. Bahru Zewde, R. Pankhurst, and Taddese Beyene, 155–165. Addis Ababa: University of Addis Ababa.

——. 2001. *Islam in nineteenth-century Wallo, Ethiopia: Revival, reform and reaction.* Leiden: E.J. Brill.

Kane, Thomas L. 1974. "Arabic translations into Amharic". *Bulletin of the School of Oriental and African Studies* 37.608–627.

Kropp, Manfred. 1986. "Arabisch-äthiopische Übersetzungstechnik am Beispiel der Zena Ayhud (Yosippon) und der Tarika Wäldä". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 138.315–346.

Leslau, Wolf. 1990. "Arabic loanwords in Geʿez". Wolf Leslau, *Arabic loanwords in Ethiopian Semitic*, 58–80. Wiesbaden: O. Harrassowitz.

Pankhurst, Alula. 1994. "Indigenising Islam in Wällo: *Ajäm*, Amharic verse written in Arabic script". *Proceedings of the Eleventh International Conference of Ethiopian Studies*, ed. Bahru Zewde, R. Pankhurst, and Taddese Beyene, 257–273. Addis Ababa: University of Addis Ababa.

Wagner, Ewald. 1973. "Eine Liste der Heiligen von Harar". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 123.269–292.

——. 1974a. "Genealogien aus Harar". *Der Islam* 51.97–117.

——. 1974b. "Three Arabic documents on the history of Harar". *Journal of Ethiopian Studies* 12.213–224.

——. 1975. "Arabische Heiligenlieder aus Harar". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 125.28–65.

——. 1976. "Die Chronologie der frühen muslimischen Herscher in Äthiopien nach den Harariner Emirslisten". *Wort und Wirklichkeit: Studien zur Afrikanistik und Orientalistik Eugen Ludwig Rapp zum 70. Geburtstag*, ed. Otto Böcher and Günter Mayer, 186–204. Meisenheim am Glan: Hain.

——. 1978. *Legende und Geschichte: Der Fatḥ Madīnat Harar von Yaḥyā Naṣrallāh.* Wiesbaden: L. Steiner.

——. 1979. "Neues Material zur 'Ausa Chronik'". *Die islamische Welt zwischen Mittelalter und Neuzeit: Festschrift für Hans Robert Roemer zum 65. Geburtstag,* ed. Ulrich Haarmann and Peter Bachmann, 657-673. Beirut: Orient-Institut.

——. 1983a. *Harari-Texte in arabischer Schrift.* Wiesbaden: L. Steiner.

——. 1983b. "Selt'i-Verse in arabischer Schrift aus dem Schlobies-Nachlass". *Ethiopian Studies*, ed. Stanislav Segert and András J.E. Bodrogligeti. Wiesbaden: O. Harrassowitz.

——. 1991. "The genealogy of the later Walashmaʿ sultan of Adal and Harar". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 141.376–386.

——. 1997. *Afrikanische Handschriften: Verzeichnis der orientalischen Handschriften in Deutschland*, XXIV/2. Wiesbaden: L. Steiner.

Wetter, Andreas. Forthcoming. "A comparative description of written Muslim and Christian Amharic". *Papers of the Symposium Cross and Crescent: Christian-Muslim Relations in Ethiopia, Goethe Institut, Addis Ababa, 23–28 September 2002.*

ANDREAS WETTER
(Johannes Gutenberg-University of Mainz)

Ethiopic Loanwords

Ethiopic loanwords in Arabic may derive from Classical Ethiopic, a South Semitic language, attested since the 3rd century C.E., or from related dialects. The question of Ethiopic loanwords in Arabic is complex. Contrary to, for instance, → Greek or → Persian loanwords, Ethiopic loans are not usually revealed by their phonological and morphological shape, so numerous criteria are necessary to determine whether a given word is an Ethiopic loan in Arabic, or vice versa, or of common Semitic stock. These criteria include the lexical isolation and attestation of a root, and its semantic development. In most cases, extra-linguistic factors have to be taken into account, such as the chronology of attestation or the historical background (Weninger 2004).

Probably a number of words common to Arabic and Ethiopic, which were earlier classified as Ethiopic loans in Arabic, are in fact originally from Sabaic, and were absorbed by both Arabic and Ethiopic. After all, Saba and its South Arabian rivals Qataban, Ḥaḍramawt, and Ḥimyar had formed the dominant culture of the region many centuries earlier than the rise of the Aksumite Empire in Ethiopia in the first centuries C.E. and the rise of Islam in the 7th century C.E. Nöldeke (1910:51), for example, classified *xawxa* 'niche, window' as a derivation of Ethiopic *ḏoḏt* 'door, opening in a wall'. But the Sabaic word *ḏḏ* (probably **ḏōḏ*), with a comparable meaning, is attested much earlier (Ja 552/3, 4th or 3rd century B.C.E.). It thus probably derives from a Sabaic word which was taken over by both languages.

Through trade relations and Ethiopian slaves, Ethiopic words denoting cultural items reached Central and North Arabia, for instance the following: *jilbāb* 'garment, gown; shirt', probably a loan from Ethiopic *gəlbāb* 'covering, veil, wrapper' (cf. Ethiopic *galbaba* 'to cover', Nöldeke 1910:53); *waqf* 'bracelet' < Ethiopic *waqf* 'bracelet' (Nöldeke 1910:53); *kabar* 'drum, kettle-drum' < Ethiopic *kabaro* 'drum, timbrel' (Nöldeke 1910:56). A possible candidate for this class of words is also *qārūra* 'bottle [of glass]', a secondary singular, derived from reanalysis of Ethiopic *qʷarir* 'cold, frozen' as a broken Arabic plural *qawārīr*, and then augmented with *tāʾ marbūṭa* (Spitaler 1998:167–168).

The most important group of Ethiopic loans came to Arabia during the first Hijra, when in 615 a group of early Muslims, not protected by tribal law from the hatred of their fellow Meccans, made their way to the court of the Christian king of Aksum in Northern Ethiopia to seek asylum, returning as late as 7/628 to Medina and bringing with them various lexical terms. Most remarkable are words from the sphere of religion like *muṣḥaf* (var. *maṣḥaf*) 'book [especially copy of the *Qurʾān*] < Ethiopic *maṣḥaf* '[any kind of] book'; *minbar* 'pulpit' < Ethiopic *manbar* 'seat, chair, high place, throne' (cf. the verb *nabara* 'to sit', not attested in Arabic); *ḥawārīyūn* 'apostles' < Ethiopic *ḥawārəyā* 'traveler, messenger, apostle (cf. the Ethiopic verb *ḥra* 'to walk', not attested in Arabic); and *fāṭir* 'creator', developed under Ethiopic influence and based on Ethiopic *faṭāri* 'creator'. Derived from *fāṭir* are *faṭara* 'to create' and *fiṭra* 'creation' (Jeffery 1938:221). Among this group should be mentioned some *hapax legomena*, such as *al-jibt* (Q. 4/51) 'idol, worship of idols' < Ethiopic *ʾamāləkta gəbt* 'new gods' (Deut. 32, 17; cf. Hebbo 1984:74–75), a much-disputed word in Arabic philology. Another famous example is *māʾida* 'table [brought down by Jesus for His disciples]'. Although its ultimate origin is obscure (but cf. Kropp 2003), its direct source is Ethiopic *māʾədd* '[the Lord's] table' (Nöldeke 1910:54). Needless to say, the status of a loan is in some cases disputable. *rajīm*, for example, with the meaning 'cursed' as an epithet of Satan, could be influenced by Ethiopic *rəgum* 'cursed' (Nöldeke 1910:25, 47). *rajama* is supposed to mean originally 'to stone'. However, the old *ḥajj*-custom of 'stoning Satan' at al-Minā is an argument for the original Arabic meaning 'stoned' (Hebbo 1984:135). The verb *nāfaqa* 'to be a hypocrite, to behave hypocritically' and the *munāfiqūn* 'hypocrites', an important concept in Qurʾānic and early Islamic reasoning, were for a long time seen as an Ethiopic loanword from *manāfəq* 'hypocrite, heretic' (Nöldeke 1910:48–49; Jeffery 1938:272; Hebbo 1984:356–360). But a close analysis of the Ethiopic and Arabic material by Brockett (1993:562) has seriously challenged this view.

While some Ethiopic loanwords took root in the Arabic language and became integral parts of the ordinary language (e.g. *minbar*), others became through their usage in the *Qurʾān* topics of constant philological and exegetical debate (e.g. *jibt*).

After the 7th century the importance of Ethiopia diminished in relation to the Islamic lands, so that it is unlikely that further Ethiopic words were borrowed by varieties of Arabic. The process was reversed: this was the time when many Arabic words were borrowed by → Ethiopic languages (Leslau 1990).

Bibliographical references

Brockett, A. "al-Munāfiḳūn". *Encyclopaedia of Islam*. 2nd ed. VII, 561–562. Leiden: E.J. Brill.

Hebbo, Ahmed. 1984. *Die Fremdwörter in der arabischen Prophetenbiographie des Ibn Hischām (gest. 218/834)*. Frankfurt: Lang.

Jeffery, Arthur. 1939. *The foreign vocabulary of the* Qurʾān. Baroda: Oriental Institute.

Kropp, Manfred. 2003. "Viele fremde Tische, und noch einer im Koran: Zur Etymologie von äthiopisch *maʾ/.əd(də)* und arabisch *māʾ/.ida/ mayda*". *Oriens Christianus* 87.140–143.

Leslau, Wolf. 1990. *Arabic loanwords in Ethiopian Semitic*. Wiesbaden: O. Harrassowitz.

Nöldeke, Theodor. 1910. "Lehnwörter in und aus dem Äthiopischen". *Neue Beiträge zur semitischen Sprachwissenschaft*, by Theodor Nöldeke, 31–66. Strassburg: Trübner.

Spitaler, Anton. 1998. *Philologica: Beiträge zur Arabistik und Semitistik*, ed. Hartmut Bobzin, with indices by Stefan Weninger. Wiesbaden: O. Harrassowitz.

Weninger, Stefan. 2004. "Anmerkungen zu den arabischen Fremdwörtern im Äthiopischen". *Studia aethiopica in honor of Siegbert Uhlig on the occasion of his 65th birthday*, ed. Verena Böll a.o., 361–369. Wiesbaden: O. Harrassowitz.

Stefan Weninger (Marburg, Germany)

Ethnicity and Language

Among the various identifying characteristics of ethnic groups, such as ancestry, religion, and territory, language is often considered the most prominent. This entry focuses on the extent to which Arabic itself serves as a unifying identity symbol, and then looks at the significance of language variation for some of the minority ethnic groups within the Arab world. Such groups may relate to language in a variety of ways, and distinctions based on language do not always correlate with those based on other criteria.

Before the movement of the Arabs out of Arabia and across the Levant, Mesopotamia, and North Africa, the area now called the Arab world had hosted many other cultures, including the Sumerians, Babylonians, Assyrians, Phoenicians, Ancient Egyptians, Persians, Greeks, and Romans. Before the arrival of Islam, it had seen the birth of Judaism and Christianity. The legacies of these pre-Islamic peoples and cultures did not all simply disappear with the advent of the Muslim Arabs. If the Arab invasions are viewed as a flood arising in Arabia and engulfing the regions from Spain to the Indus, then in parts of these regions the floodwaters bearing Arabic and Islam seem to have entirely submerged what was there before, while other parts were not covered at all, forming islands. In still other places there was a mingling of the floodwaters with lakes or rivers already present, so that the boundaries between the two became fluid. Finally, in some places the floodwaters eventually receded so that lands once under water re-emerged, possibly showing residual effects of the flood. Some peoples of the region resisted the forces of Arabicization, Islamicization, or both; even among those who underwent both these processes, this was not always accompanied by a total abandonment of their earlier culture. Thus, there are still pockets across the Arab world using languages other than Arabic and practicing religions other than Islam, and there are still groups convinced that their ancestors belonged to a people different from those of their neighbors. The extent to which the various groups assert their distinctiveness may vary over time and in relation to circumstances, and individuals may also feel allegiance to more than one group, so that it may be necessary to recognize overlapping identities rather than ones that are wholly incompatible. Our concern here is with the relevance of language to these identities.

Theorists have spent much energy on developing definitions of ethnicity, which is an analytic concept used to describe the bonds which lead certain people to identify themselves as a group. It is generally agreed that ethnic identity is based on some kind of ancestral link – what Fishman (1977:17) has designated as paternity as opposed to patrimony: "Ethnicity is, in part, but at its core, experienced as an inherited constellation acquired from one's parents as they acquired it from theirs, and so on back further and further, *ad infinitum*". Yet according to Fishman, this does not exclude the possibility that certain individuals in certain circumstances may be considered to acquire or lose an ethnicity. It seems important to recognize that the characteristics defining the group need not all be objectively verifiable ones; self-ascription may be important. As Edwards (1985:10) notes, the group's boundary "can be sustained by shared objective characteristics (language, religion, etc.) or by more subjective contributions to a sense of 'groupness', or by some combination of both". Moreover, as Barth (1969) argues, the actual content of a group, such as its use of a particular language, customs, and traditions, may change over time, yet its separate identity may be preserved provided the boundaries remain clear. As for the distinction between ethnicity and nationalism, the two are often considered as points on a continuum, while suggestions about what distinguishes the latter from the former have invoked notions such as degree of self-awareness, organization, mobilization, or ideologization (see, for instance, Connor 1978; Edwards 1985; Fasold 1984; Paulston 1994).

Certain components are consistently evoked as important defining or identifying characteristics of ethnic groups. These include ancestry, language, religion, territory, shared values, and other traditions, such as specific occupations. However, many scholars agree that the content of a particular ethnicity is something malleable: an ethnic group may survive intact despite losing its ancestral territory, abandoning its original language, and changing its religion. Differing views have been adopted on the importance of language to ethnic identity. Fishman (1977:25) describes language as "the quintessential symbol", which functions as "the recorder of paternity, the expresser of patrimony and the carrier of phenomenology". Lieberson (1981:4–5) notes that a difference of language typically signals a difference of ethnic group, insisting that "the overlap between ethnic and linguistic boundaries... is often only partial but never random", and describes language as "an important shield against assimilation". Some imply that language is more essential for nationhood than for ethnicity. Thus, Fasold (1984:4) suggests that the extent to which a group maintains its language rather than abandoning it may be taken as "an indicator of nationality versus simple ethnicity". Kedourie (1961:68) feels that "a group speaking the same language is a nation", whereas Weber (1948:172–173) takes quite the opposite view, claiming that a common language is neither a sufficient nor a necessary condition for nationhood. Ross (1979), while taking the position that language is the most powerful single symbol of ethnicity, emphasizes that the relationship of a group to its language is not static, but may change over time.

Across the Arab world, many groups are commonly identified by reference to their distinctive languages. These include groups such as the Kurds, Berbers, Assyrians, Armenians, Circassians, Nobiin, and Turkmens, at least some of whose members continue to use the traditional language in everyday life. However, ability to speak the language may not be a necessary condition for being recognized as part of the group. Individuals whose families have undergone a process of → language shift, possibly quite some time ago, may still identify with a language they cannot not speak on a symbolic level. In other cases, such as that of Hebrew among some Jewish communities or Coptic among the Egyptian Copts, a language may retain significance among those who cannot speak it because of its role as a liturgical language.

Religion is a crucial distinguishing feature for Christian communities such as the Maronites, Copts, Assyrians, and Armenians, which lay claim to ethnic distinctiveness, and also for sects such as the Druze and Alawis. However, the divisions drawn on linguistic lines do not entirely correlate with religious ones. There are, for instance, Aramaic-speaking Muslims (and there were until recently Aramaic-speaking Jews) in Syria, and Berber-speaking Jews in Morocco, not to mention the large numbers of Arabic-speaking Christians. Maila (1998) points out that many Middle Eastern Christians identify themselves solidly as Arabs, considering religion to be a personal issue rather than a marker of group identity; Christian Palestinians would seem to be a good example of this position. Religious divisions may also split what is felt to be a single ethnic group, as in the case of the Assyrians of Iraq, who are divided between the Nestorian and Chaldean churches. Nisan (1996) suggests that a distinctive religion may actually encourage language shift, noting that communities not distinguished by a separate religion, such as the Kurds and the Berbers, have preserved their language, whereas those with a separate religious identity were more ready to adopt Arabic.

As for ancestry, some groups lay considerable stress on claiming descent from a tribe or civilization different from that of their neighbors. In some cases, the distinction seems uncontroversial, as for the Armenians and Circassians, known to have migrated to the region from a homeland elsewhere, and also along the fringes of the Arab world in states like Sudan and Mauritania, where the tribes of the south are clearly distinguishable from those of the north. In others, it is much less clear. Some Copts, for instance, insist on their descent from the Ancient Egyptians rather than the Arabs, and some Maronites claim to be descendants of the Phoenicians rather than Arabs. Yet Fargues (1998) argues that the idea that such groups descend from ancestors clearly different from those of their Muslim neighbors is a myth, since the immigrants who moved from Arabia to the region were quite simply not numerous enough to constitute the ancestors of all the current Muslim populations. Instead, he accounts for the present demographic situation as the result of conversion and intermarriage. Moreover, even where different ancestry is historically documented, the lack of salient physical differences between the original groups may make it quite impossible to separate them out. Thus, a Moroccan who today thinks of himself as an Arab may have ancestors of Berber or Iberian stock. On the other hand, there are also cases where a group appearing to be racially distinctive plays down this aspect of its identity. Thus, certain black tribes in Sudan choose to identify themselves as Arabs, as do

Moroccans of apparently Sub-Saharan ancestry whose ancestors arrived in the region via the slave trade. Such cases illustrate the extent to which self-ascription may be more important in upholding the boundaries between ethnic groups than scientifically verifiable characteristics.

Finally, territory is of varying importance to the different ethnic groups. Some possess a well-defined heartland, whether this is clearly localized, as with the Maronites in Mount Lebanon, or spread across several national territories, as in the case of Kurdistan, which spans Turkey, Iraq, Iran, and Syria. Others have been displaced from their ancestral lands, and may have reassembled in a new region, as in the case of the Armenians in Lebanon, or remain dispersed, as with the Circassians, spread through Jordan, Syria, and Iraq. The Jews were of course dispersed for centuries, and small groups still remain in the Arab world outside Israel. It is interesting to note that, among those now gathered together in Israel, further layers of identity relating to their previous homelands have become apparent; despite official policies encouraging assimilation, some groups, such as the Jews of Moroccan, Yemeni, and Iranian origin, have nurtured separate ethnic identities relating to these places of origin (Eickelman 1998). A final case is that of the Berber peoples of North Africa, who, though linked through language and ancestry, exist in a number of quite separate pockets from Morocco to Egypt, with no clearly overarching ethnic consciousness linking the various groups together.

It is also worth noting that among those who have emigrated from the Arab world, people who might have felt divided by language, religion, ancestry, or territory at home may come to feel bound together by a very broad shared regional origin. For instance, while the earliest wave of immigrants from the Middle East to the Unites States tended to identify themselves as Syrians or Lebanese rather than Arabs (M. Suleiman 1999), there is apparently today a growing trend for people originating from the region to lay claim to an Arab identity. The 2000 census showed a 62 percent increase in the numbers identifying themselves as Arab rather than by national origin (Clemetson 2004). Yet, those who choose to describe themselves as Arab Americans may be Muslim or Christian, and may or may not speak Arabic (→ Latin America, → North America).

Even from this brief survey it can be seen that, while language is certainly one component contributing to ethnic divisions in the Arab world, it may interact with other components of identity in various ways. Although the scope of this entry does not allow even a cursory look at each of the many ethnic groups already mentioned, it is still possible to consider in a little more detail a few cases which, while not necessarily representative, present contrasting profiles.

It may first of all be worth looking at the label 'Arab' itself. Some scholars tend to use the term 'ethnicity' only when discussing minority groups. Bates and Rassam (2000) suggest that it is not useful to refer to Arabs in the Middle East as an ethnic group, though a subgroup of Arabs living among Persian speakers could be so labeled. However, it is unnecessary to narrow the term's application in this way. In fact, Arab identity might be taken almost as a prototype example of an identity defined mainly through language. From the beginnings of Islam, the → Arabs identified themselves on the basis of their language and contrasted themselves with the *ʿajam*, who spoke other languages (Hourani 1983). Indeed, Mālik cites a *ḥadīṯ* to the effect that anyone who speaks Arabic is an Arab, while the jurist aš-Šāfiʿī, in maintaining that a Muslim could become an Arab by learning Arabic, implied that language took priority over descent in this respect (Y. Suleiman 2003:64). Since then, many prominent Arab thinkers, notably Sāṭiʿ al-Ḥuṣrī (who himself spoke Turkish before learning Arabic), have insisted on the crucial role of the Arabic language in defining the Arabs as a people. To outsiders, the language might not seem such a clearly unifying factor; in fact, the dialects of Arabic spoken by, say, an illiterate Moroccan and his Yemeni counterpart are so different that they may find considerable difficulty in understanding each other. The real link is rather made through Classical Arabic, which is a strong symbol of identity even for people unable to speak or understand it (→ *ʿarabiyya*).

Nowadays, the symbolic value of Arabic tends to be very closely associated with Islam. The prototype Arab has long been assumed to be a Muslim. Yet, the link between Arabic and Islam has not always been felt to be so exclusive. In al-Andalus, Arabic was the native language of Jews as well as Muslims, and Classical Arabic was the written medium for Muslim, Jewish, and Christian scholars. More remarkably, the Mozarab Christians maintained written Arabic as a symbol of their distinctive identity for two centuries after the Reconquista, while they were under Christian rule. Ferrando (2000:69) reports on this rather remarkable "role of Arabic in sustaining and strengthening a *non-Islamic* identity".

Contemporary ethnic allegiances of non-Muslim speakers of Arabic are varied. Some wholeheartedly assert their own Arabness, as in the case of some Moroccan Jews or Lebanese Christians, while others make considerable efforts to distance themselves from the Arabs by claiming separate descent, as in the case of some Maronites. Still others are able to view themselves as possessing a dual identity. Amin Maalouf (1998:24), of Christian Lebanese origin, remarks that as an Arabic speaker he can identify with all other Arabic speakers, and as a Christian with all other Christians, so that the combination of these two elements of his identity allows him to identify with half the human race.

The Arabic linguistic criterion does not necessarily correlate with common descent, as illustrated by the varied make-up of the Arabic-speaking peoples of North Africa. There are also groups who identify themselves as Arabs by descent, yet are unable to speak Arabic at all. Many Arab Americans of course come into this category, while Kieffer (2000) reports on the Persian-speaking Arabs of Afghanistan, who have maintained neither the Arabic language nor distinctive traditions, clothes or way of life, yet, insist on their distinctive ethnic status as descendants of the Prophet.

In fact, then, while much has been made of the role of Arabic as the cement holding Arabs together, it should be recognized that the use of Arabic cannot always be considered as either a necessary or a sufficient criterion for Arab identity. There are Arabic speakers who reject categorization as Arabs and self-ascribed Arabs whose links with the language are tenuous.

For minority groups within the Arab world, the case of the Assyrians may be considered as an example. The traditional heartland of this group lies in the mountains of Northern Iraq, and there are still about one million Assyrians in this region, while others live in Syria, Turkey, Lebanon, and Iran. They claim descent from the ancient Assyrians, and have constituted a minority without a state since the fall of Assyria in the seventh century B.C.E. The Assyrians have long suffered persecution, with the massacre of two-thirds of their number in Iraq and Turkey in 1915, and a further massacre of 3,000 in Iraq in 1933. Under the Baʿath regime in Iraq, Assyrians had to discard their identity and language in order to accede to official posts, and in the 1977 census, they were obliged to identify themselves as either Arabs or Kurds. Unlike the Kurds, they have often failed to obtain recognition as an ethnic group by outside bodies, such as the Arab League, and recently some have expressed fears of finding themselves oppressed by the Kurds if the latter obtain a greater degree of autonomy in postwar Iraq. Since 1991, half of the Assyrian population of Iraq has left, and there is now a diaspora of four million in the West.

Despite all these pressures, the Assyrians have preserved their distinctiveness. They have maintained their language, variously referred to as Modern Assyrian, Neo-Syriac, or Neo-Aramaic, which has a long written history and a highly charged symbolic value as the language spoken by Jesus. Their separateness from their neighbors has also been upheld through their Christian faith, which dates from the 1st century C.E. According to Lewis (2003), the internal divisions between those Assyrians who are Nestorians and those who are Chaldean Catholics have led to some fragmentation of Assyrian identity. However, Deniz (2000) shows how, under the influence of modernization and mass emigration, Assyrian identity has undergone a series of transformations. Intellectual Assyrians in the West have attempted to downplay divisions between the various churches, and some now prefer instead to emphasize their identification with the civilization of ancient Mesopotamia as a unifying element. They fiercely defend their claim to distinct ancestry; most recently, in October 2001, the Coalition of American Assyrians and Maronites lodged an official protest at the Arab-American Institute's tendency to refer to them as Arabs, and totally rejected this label (AINA 2001).

The Assyrians thus offer an example of a group which, through centuries of shifting political and social conditions, loss of land, depleted population and exile, has nevertheless preserved its distinctive character through both linguistic and religious demarcation from other groups. We may compare their status with that of the Berbers of North Africa, who were settled in the region before the Arab invasions. Peoples classified as Berber by virtue of their use of a language belonging to this family are to be found in Morocco, Algeria, Tunisia, Libya, Egypt, Mauritania, Mali, and Niger, but they have been cut off from one another for so long that, until recently at least, no sense of overarching identity remained. As Hart (1973:26), among others, has noted, "the geographical fractionalization of Berber speech areas across the map of North Africa has always been a barrier not only to linguistic comprehensibility but even more so to any concept of 'Berber nationhood'". The discussion here is confined to the Berbers of Morocco, who nowadays prefer to be referred to by the name Amazigh (the term 'Berber' is used here, for clarity's sake, as the term used in earlier research). As the Arabs established their rule over North Africa, the Berber tribes converted to Islam, and many of them also became Arabicized, abandoning their traditional language. Many members of the first Arab forces to arrive in the Iberian Peninsula were of Berber origin, as were two later dynasties that ruled al-Andalus, the Almoravids and the Almohads (both of whom, interestingly, sought to claim Arab lineage for themselves). There are no clear racial characteristics distinguishing the two groups, and after centuries of shared religion and intermarriage, estimates of what proportion of Moroccans today are of Berber descent vary greatly. While some are content to suggest that a majority of Moroccans have Berber ancestry, others point out that, given the relatively small numbers of the invading Arab armies, it is plausible to assume that almost all Moroccans have Berber blood.

Today, then, with no distinct religion and no clear signs of distinct ancestry, language would seem to be the essential defining characteristic of the Berbers, though, as Gellner (1969:13) points out, "for all practical purposes, a 'Berber' is a native Berber speaker who is both Muslim and white. Jewish and Negro Berber-speaking minorities are sociologically distinct". Apart from these exceptions, it is usually taken for granted that families where Berber is still spoken are of Berber descent. However, it cannot be assumed that those whose first language is Arabic are of Arab descent, since they may equally well simply descend from Berber tribes that have long been Arabicized. Nowadays, Berber is still the first language used in a number of communities, mainly to be found in the remote mountainous regions of the Rif and Middle Atlas, the Anti-Atlas, and the plain of Souss. The discontinuous nature of these speech communities, added to the fact that for centuries there was very little writing in Berber, has produced three major dialects, with less than total intelligibility between them. The absence of census questions about knowledge of the language means that estimates of the number of Moroccans who can speak Berber also vary widely, from less than one-sixth to as much as one-half of the population.

Despite French efforts to bring about divisions between Arabs and Berbers during the period of the French protectorate (1912–1956), through the provision of separate education programs and recognition of a distinct legal system for Berber areas, this deliberate manipulation does not seem to have had lasting effects. Several researchers who studied Moroccan society in the postcolonial period concluded that divisions along ethnic lines were not particularly prominent for the Moroccans themselves. Gellner, writing of the linguistic categories of Arab and Berber, remarks that "neither has ever acted or felt as one unit" (1973:12), and insists on the fact that "the Berber sees himself as a member of this or that tribe, within an Islamically-conceived and permeated world – and *not* as a member of a linguistically defined ethnic group, in a world in which Islam is but one thing among others". (1973:13). Likewise, Rosen (1973:173) concludes that even the categories Arab and Berber are not felt to be mutually exclusive, as the French had assumed, and that they are "contingent and partial rather than complete and pervasive features of each man's social identity". This view of the fluidity of the distinction receives some support from a survey of language and identity reported on in Davies and Bentahila (1989) and Bentahila and Davies (1992) in which 15 percent of the Berber-speaking informants actually chose to describe themselves as Arabs. This survey also suggested that the correlation between group identity and language was less than clear. For instance, 30 percent of the Berber-speaking informants said they considered Arabic rather than Berber to be their own language, and in a number of cases Berber speakers referred to Arabic as the language of their ancestors even though their own parents were monolingual speakers of Berber. Such responses may illustrate the powerful symbolic link between Arabic and Islam noted earlier. As Muslims, Berbers may feel they have a share in the Arab heritage, just as the Almoravids and Almohads did nine centuries ago. In addition, 10 percent of those who could not speak Berber chose to identify themselves as Berbers; 83 percent of the respondents claimed that it was not necessary to speak Berber to be a Berber, and interestingly, a full 20 percent of Berber speakers said they did not wish their children to speak this language. Such positions suggest a pragmatic approach to the language, retaining it only where it is useful, rather than clinging to it as an essential symbol of identity.

As late as 1993, Geertz claimed that Morocco is organized "not culturally, or linguistically, or racially, or religiously" (1993:12), but more in terms of personal relationships than in group terms. However, as Crawford (2002) emphasizes, there have been changes since the studies by Gellner (1973) and Rosen (1973), which downplayed the Arab-Berber distinction. For ordinary unmobilized citizens, the boundary between Berber and Arab may still seem relatively unimportant compared to divisions based on tribal, kinship, or occupational links, and bonds of shared nationality and religion, but over the past three decades activist groups seeking to protect and promote Berber language and culture have become more prominent. This more conscious, militant assertion of ethnicity did not spring directly from the rural populations who still use the language in everyday life, but has been largely the work of urbanized intellectuals (including some whose families had already abandoned the use of Berber in the home). Groups based outside Morocco were particularly prominent in the 1980s and 1990s, and links with Berber-speaking communities in other parts of North Africa have been developed through the internet (Almasude 1999). In 2000, a group of intellectuals signed the Berber Manifesto, which made a number of requests mainly concerned with raising the status of the Amazigh language. Significantly, this document explicitly states that people "are Amazighe thanks to their language not to their race… Whoever among them exposes his language to loss is doing the same to his Amazighe existence".

As this ethnic movement has gained impetus, the government position has gradually shifted from studiously ignoring the Berber heritage to increasingly recognizing it. In 1994, King Hassan II acknowledged the need for the Berber language to be used in schools, and television news broadcasts in Berber were introduced. In 2001, his successor Mohamed VI set up the Royal Institute for Amazigh Culture (IRCAM) and announced a program for the introduction of Berber into the school curriculum. Interestingly, the institute has opted to use the ancient alphabet, Tifinagh, rather than the Arabic or Latin alphabet, a move which can be seen as symbolically emphasizing the distinctiveness of the language, at the expense of practical considerations. Present provisions aim at providing teaching of Berber in primary schools across the country by 2008. This step might look like a decisive reinforcement of Berber identity in Morocco. However, it is interesting to note that, rather than providing teaching in Berber only in the areas where it is spoken, which might indeed have emphasized the ethnic division, the government has opted for teaching the language throughout the country and emphasizing the importance of Berber culture as part of the heritage of all Moroccans. The consequences of these new policies remain to be seen, but in fact they may not necessarily strengthen the boundaries between Arab and Berber.

The Moroccan Berbers would thus seem to represent a case where an ethnic group is distinguished by language but in fact not by much else, since they are united with the majority through religion and even ancestry, if we accept that most Moroccans are anyway of Berber descent. Even the linguistic distinction seems not to have been particularly prominent in traditional communities, but it remains to be seen how the recent mobilization among intellectual activists will develop.

Finally, it may be helpful to consider the case of the Jews, who formerly constituted sizeable minorities in a number of Arab countries, including Egypt, Lebanon, Syria, Iraq, and Yemen. The Moroccan Jewish community will serve as an example. The Jewish population in Morocco numbered some 250,000 in the 1950s

(Zafrani 1983), but mass emigration to Israel, Europe, and the Americas has now reduced numbers to around 3,500. It is composed of two distinct components, the Toshabim or native Jews, whose ancestors have lived in Morocco since pre-Islamic times, and the Sephardic Jews who arrived in Morocco from Iberia following the Reconquista. What is striking about this group is the traditional absence of a distinguishing language. While Hebrew was used solely as a liturgical language, the Moroccan Jews have traditionally used either Arabic or Berber as their home language, depending on locality. Close social bonds between Muslims and Jews were not unusual in certain communities and periods (Rosen 1973; Shokeid 1982; Eickelman 1998). However, once the Alliance Israélite Universelle began operating in Morocco, with the opening of schools in Tetouan and Tangier in the 1860s (Stillman 1979), there began a process whereby the Moroccan Jews turned more and more towards the use of French. Under the French protectorate this trend was actively encouraged. El Maleh (1977) describes poignantly the painful process whereby the Jews were impregnated with French and European values and customs, in an attempt to reinforce divisions between Jew and Muslim, in the same way as the French had attempted to divide Berber and Arab. In many Jewish families, there has been a shift over two or three generations from Arabic to French as the home language (Bentahila and Davies 1992).

This seems a rather unusual case, as it involves a minority group, which for centuries had preserved its distinctiveness and hard boundaries without the support of a distinctive language, moving toward what is essentially a foreign language in Morocco and one which historically has perhaps fewer associations with Jewish identity than has Arabic, which was after all the language of Jewish scholarship for centuries in al-Andalus. It may be tempting to see this shift as a kind of symbolic emigration on the part of those who did not leave Morocco. However, in our survey, the Jews questioned insisted on their Moroccan identity, with a majority selecting Arabic as the language Moroccan Jews ought to speak, yet justifying the use of French for its usefulness as a language of wider communication; in no age group did a majority agree on which language they most identified with, and there appeared to be a strong conviction that "language is something quite separate from identity" (Bentahila and Davies 1992:209).

This brief look at the ways in which language relates to ethnic identity in a number of communities within the Arab world may lend support to Ross's (1979:11) claim that "language has no single mode of relationship to collective identity". Ross points out that a group may abandon use of its original language and shift its allegiance to a different one, that the language taken as a symbol of identity need not be spoken by the members of the group, and indeed need not even have been spoken by the group's ancestors. All these possibilities are illustrated by the examples discussed above. The languages of the various minority groups have been left in different circumstances by the metaphorical flood invoked earlier. For the Assyrians, language, assisted by religious distinctions, has remained a landmark which was not swamped by the floodwaters. For the Berbers, to a greater extent, the language forms a layer which was submerged and eroded by the spread of Arabic and Islam yet which remains perceptible and may yet re-emerge. For the Moroccan Jews, on the other hand, a new language, French, has come to rest as a relatively superficial layer on top of those used in the past and now abandoned. The impact of Arabic over the whole region has been variable, sometimes swamping, sometimes merely trickling over other layers. And 13 centuries after the flood of Arabic, the situation is still evolving as elements settle into place, are eroded, submerged, or re-emerge in changing circumstances.

Bibliographical references

AINA (Assyrian international news agency). 2001. "Coalition of American Assyrians and Maronites rebukes Arab American Institute". <http://www.aina.org/releases/2001/ caamletter.htm>.

Almasude, Amar. 1999. "The new mass media and the shaping of Amazigh identity". *Revitalizing indigenous languages*, ed. Jon Reyhner, Gina Cantoni, Robert N. St Clair, and Evangeline Parsons Yazzle, 117–128. Flagstaff, Ariz.: Northern Arizona University.

Barakat, Halim. 1993. *The Arab world: Society, culture and state*. Berkeley: University of California Press.

Barth, Fredrik (ed.) 1969. *Ethnic groups and boundaries: The social organization of cultural difference*. Boston: Little, Brown.

Bates, Daniel G. and Amal Rassam. 2001. *Peoples and cultures of the Middle East*. 2nd ed. Upper Saddle River, N.J.: Prentice Hall.
Bentahila, Abdelâli and Eirlys E. Davies. 1992. "Convergence and divergence: Two cases of language shift in Morocco". *Maintenance and loss of minority languages*, ed. Willem Fase, Koen Jaspaert, and Sjaak Kroon, 197–210. Amsterdam and Philadelphia: J. Benjamins.
Clemetson, Lynette. 2004. "Some younger US Arabs reassert ethnicity". *New York Times*, 11 January.
Connor, W. 1978. "A nation is a nation, is a state, is an ethnic group, is...". *Ethnic and Racial Studies* 1.377–400.
Crawford, David. 2002. "Morocco's invisible Imazighen". *Journal of North African Studies* 7:1.53–70.
Davies, Eirlys E. and Abdelâli Bentahila. 1989. "On mother and other tongues: The notion of possession of a language". *Lingua* 78.267–293.
Deniz, Fuat. 2000. "Maintenance and transformation of ethnic identity: The Assyrian case". <http://aanf.org/America/assyrians/maintenance_transformation.htm>.
Edwards, John. 1985. *Language, society and identity*. Oxford: Blackwell.
Eickelman, Dale F. 1998. *The Middle East and Central Asia: An anthropological approach*. 3rd ed. Upper Saddle River, N.J.: Prentice Hall.
El Maleh, Edmond Amran. 1977. "Juifs marocains et marocains juifs". *Les Temps Modernes* 375 bis.495–520.
Fargues, Philippe. 1998. "The Arab Christians of the Middle East: A demographic perspective". Pacini (1998:48–66).
Fasold, Ralph. 1984. *The sociolinguistics of society*. Oxford: Blackwell.
Ferrando, Ignacio. 2000. "The Arabic language among the Mozarabs of Toledo during the 12th and 13th centuries". Owens (2000:45–63).
Fishman, Joshua A. 1977. "Language and ethnicity". *Language, ethnicity and intergroup relations*, ed. Howard Giles, 15–57. London: Academic Press.
Geertz, Clifford. 1993. "Primordial loyalties and standing entities: Anthropological reflections on the politics of identity". Public lecture No. 7, Collegium Budapest/Institute for Advanced Study. <http://www.colbud.hu>.
Gellner, Ernest, 1969. *Saints of the Atlas*. London: Weidenfeld and Nicolson.
——. 1973. "Introduction". Gellner and Micaud (1973:11–21).
—— and Charles Micaud (eds.). 1973. *Arabs and Berbers: From tribe to nation in North Africa*. London: Duckworth.
Hart, David M. 1973. "The tribe in modern Morocco: Two case studies". Gellner and Micaud (1973:25–58).
Hourani, Albert. 1983. *Arabic thought in the liberal age: 1789–1939*. Reissued with new pref. Cambridge: Cambridge University Press.
Kedourie, E. 1961. *Nationalism*. London: Hutchinson.
Kieffer, Charles. 2000. "The Arabic speech of Bactria (Afghanistan)". Owens (2000:181–198).
Lewis, Jonathan E. 2003. "Iraqi Assyrians: Barometer of pluralism". *Middle East Quarterly* 10:3. <http://www.meforum.org/article/313>.
Lieberson, Stanley. 1981. *Language diversity and language contact*. Stanford: Stanford University Press.
Maalouf, Amin. 1998. *Les identités meurtrières*. Paris: Grasset.
Maïla, Joseph. 1998. "The Arab Christians: From the Eastern Question to the recent political situation of the minorities". Pacini (1998). <http://arabworld.nitle.org/texts. php?module_id=6&reading_id=58&sequence=1>.
Nisan, Mordechai. 1996. "The minority plight". *Middle East Quarterly* 3:3.25–34.
Owens, Jonathan (ed.). 2000. *Arabic as a minority language*. Berlin: Mouton de Gruyter.
Pacini, Andrea (ed.). 1998. *Christian communities in the Middle East*. Oxford: Clarendon Press.
Paulston, Christina Bratt. 1994. *Linguistic minorities in multilingual settings*. Amsterdam and Philadelphia: J. Benjamins.
Rosen, Lawrence. 1973. "The social and conceptual framework of Arab–Berber relations in central Morocco". Gellner and Micaud (1973:155–173).
Ross, Jeffrey A. 1979. "Language and the mobilization of ethnic identity". *Language and ethnic relations*, ed. Howard Giles and Bernard Saint-Jacques, 1–13. Oxford: Pergamon Press.
Shokeid, Moshe. 1982. "Jewish existence in a Berber environment". *Jewish societies in the Middle East*, ed. Shlomo Deshen and P. Zenner, 105–121. Washington, D.C.: University Press of America.
Stillman, Norman. 1979. *The Jews of Arab lands in modern times*. Pennsylvania: Jewish Publication Society.
Suleiman, Michael W. 1999. "Introduction: The Arab American immigrant experience". *Arabs in America: Building a new future*, ed. Michael W. Suleiman, 1–21. Philadelphia: Temple University Press.
Suleiman, Yasir. 2003. *The Arabic language and national identity*. Washington, D.C.: Georgetown University Press.
Weber, Max. 1948. *From Max Weber: Essays in sociology*. London: Routledge and Kegan Paul.
Wexler, Paul. 2000. "Arabic as a tool for expressing Jewish and Romani ethnic identity: A prolegomenon to a typology of Arabic in non-Arabic speaking communities". Owens (2000:65–87).
Zafrani, Haim. 1983. *Mille ans de vie juive au Maroc*. Paris: Maisonneuve and Larose.

Eirlys E. Davies
(King Fahd School of Translation)
Abdelâli Bentahila
(Abdelmalek Essaadi University)

Etymology

Etymology is a linguistic discipline dating from Ancient Greece. Plato, for instance, devoted his dialogue *Kratylos* to the explanation of various Greek words. The word *etumología* 'the analysis of a word so as to find its origin' was first used at the end of the 1st century B.C.E. by Strabo (784) and Dionysius Halicarnassensis (*De compositione verborum* 16); the verb *étumologeō* 'I analyze a word and find its origin' is used only by Athenaeus (35C), living in the 2nd/3rd century C.E. The compound consists of the base *log-* known from the names of various scientific disciplines (*lógos* 'word, promise, discourse, tale', *logízomai* 'I take into account, consider, calculate'). The first component is formed from the adjective *étumos*, *-on* 'true, sure, real' (cf. the substantive *étumon* 'the true literal sense of a word according to its origin', Diodorus I, 11; Athenaeus 571D; and the adverb *étumōs* 'etymologically', Aristotle, *de Mundo* VI, 19).

The main purpose of etymology is to explain the origin of words. There are two possible strategies: (a) internal etymology, limited to one language with its lexicon and grammar; and (b) external etymology, based on a comparison of several related languages, which allows a formulation of phonetic and morphological correspondences. This procedure consists of two steps, called by Otto Dempwolff: (a) the inductive phase, in which similar words with similar meanings are collected; the most frequent sound correspondences may reflect the phonetic rules. For their verification, the second step must be realized: (b) the deductive phase, in which the most probable sound correspondences serve to separate the accidental similarities from the real lexical correspondences inherited from the protolanguage, a common ancestor of the set of studied related languages.

The Arabic grammatical tradition has always preferred the approach based on internal etymology. This is probably the reason why up till now there still is no comparative or etymological dictionary of Arabic. This is rather surprising in view of the general situation in Semitic lexicology: more or less complete comparative dictionaries have already been published for Akkadian, Ugaritic, Phoenician, Hebrew, Aramaic and Syriac, Sabaic, Soqotri, Geʿez, Tigre, Harari, and Gurage. In the early Middle Ages the conditions for a development of comparative Semitic lexicology were very promising. Saʿadyah Ga'on (Saʿadiyya ibn Yūsuf, 892–942), the head of the Jewish community in Babylonia, compiled the *Kitāb ʾuṣūl aš-šiʿr al-ʿibrānī* 'Book of the roots of Hebrew poetry', usually referred to by its Hebrew title *'Agron* 'Compendium'. It was the first Hebrew dictionary, with glosses in Arabic. Yehudah ibn Qurayš, living in Tahort, present-day Algeria, in the 10th century wrote the book *Risāla* 'Treatise, Epistle', in which he compared Biblical Hebrew, Mishnaic Hebrew, Aramaic, Arabic, and other languages including, for example, Berber. Ibn Barūn, living around 1100 C.E. in Saragossa, was the author of the *Kitāb al-muwāzana bayna l-luġa al-ʿibrāniyya wa-l-ʿarabiyya* 'Book of comparison between the Hebrew and the Arabic language', containing sections devoted to the comparative grammar and lexicology of Arabic and Hebrew (Schippers 1998:60, 63). The contribution of Ibn Qurayš in the field of comparative grammar and lexicon of Semitic languages is comparable to that of Gottfried W. Leibniz (1646–1716) in establishing Indo-European comparative linguistics. From this point of view, the Jewish tradition in the Arabic environment had a head start of at least 700 years compared to the study of the Indo-European languages.

The difference between internal and external etymology can be demonstrated by the Arabic root *ṯ-n-y: iṯnāni* masc., *ṯintāni* fem., in compounds *iṯnā-* 'two'. In the Arabic lexicon there is a rich set of derivatives: *ṯiny*, pl. *'aṯnā* 'second child or foal', *ṯiny, ṯun-an, ṯin-an*, pl. *ṯinyat* 'governor' = 'the second person in the kingdom'. The root *ṯ-n-y* does not stand isolated within Semitic. There are cognates in all Semitic languages: Akkadian masc./fem. *šinān/šittān*, Eblaic masc. *šina* (Dombrowski 1994), Ugaritic *ṯnm/ṯtm* [*ṯināmi/ṯittēmi*], Phoenician *šnm* [*šənēm*], late Punic (*l*)*isnim*, Hebrew *šĕnáyim/šĕttáyim*, Arabic *iṯnāni/ṯintāni*, Sabaic *ṯny/ṯnty*, Qatabanian masc. *ṯnw*, all meaning 'two'. In the Ethio-Semitic languages the same root expresses 'the next day' (Geʿez *sānəy*) or 'Monday' (Tigre *säno*, Tigray *sänuy*) (cf. Russian *vtórnik* 'Tuesday' vs. *vtorój* 'second'). On the other hand, the seemingly different second radical in the numeral 'two' in Aramaic (**təron/tarton*) and Modern South Arabian (Mehri *troh/trīt*, Hobyot *θro*, *θroh/θərīt*, Harsusi *θərō/θərə́t*, Jibbali *θroh/θrət*, Soqotri *trøh/trih*) probably has its origin in a rule changing the initial cluster **Cn-* in **Cr-* in Aramaic and Modern South Arabian and in *'iCn-* in Arabic (Testen 1985). The present data may be projected to the Proto-Semitic level, where *_-ā-_ and *_-ay-_ are the dual markers for nominative and oblique respectively, *_-at-_ is the feminine marker and *_-ni_ is the determiner, not appearing before nouns (Table 1).

Table 1. Proto-Semitic 'two'

	nom.	acc./gen.
masc.	*_ṯín-ā-ni_	*_ṯín-ay-ni_
fem.	*_ṯín-at-ā-ni_	*_ṯín-at-ay-ni_

The Proto-Semitic reconstruction is not the final limit. The Egyptian and Berber cognates shift the age of this root to the Afro-Asiatic (= Hamito-Semitic) level: Egyptian masc./fem. _snwj/sntj_, Coptic Sahidic masc./fem. _snau/sənte_ with the root vowel *_-i-_ confirmed by the cuneiform record _ši-na_ from the Amarna tablets || Berber masc./fem. *_sīn_ ~ *_Hissīn_/*_sīnat_, attested in Siwa, Ghadames _sən/-ət_, Kabyle _sin/snat_, Wargla _sin/sent_, Ahaggar _əssin/sänât_, Zenaga _šinan/šenanet_; Guanche of Tenerife _sijn_ 'two', Gran Canaria _smetti_ [= _sinetti_] 'two'. In projecting to the Proto-Afro-Asiatic level, the starting point should be reconstructed as *_čin-(ay-)_. But even the Afro-Asiatic proto-language is not the ultimate limit. If there are promising parallels in language families that are possibly related to Afro-Asiatic (e.g. Indo-European, Kartvelian, Dravidian, Uralic, Altaic, all members of the so-called Nostratic macro-family), it is legitimate to admit a still deeper history of the studied root. In the case of Afro-Asiatic *_čin-_ 'two', one could mention Kartvelian *_č(w)en-_ > Swan _išgen_ '(an)other'.

Another example of internal etymology, but in the Semitic context, is the word for 'tear'. Arabic pl. _damˁ_, nomen unitatis _damˁat_ (Fischer 1972:§84a), differs from other Semitic forms which have the vowel *_i_ in the first syllable: Akkadian pl. _dīmāt-u_, du. _dimā(-šu)_, Eblaic _ì-ti-ma-a-tum_ [_ʻidmaˁātum_], Ugaritic _dmˁt_, pl. _ʻudmˁt_ [_ʻudmaˁāt_], Hebrew coll. _dimˁā_, pl. _dəmāˁōṯ_, Jewish Aramaic (Targum, Babylonian) _dimaḡˁṯ-ā_, West Syriac _demˁəṯō_, Mandaic _dima_, _dimihta_, Neo-Aramaic of Heretvin _dem'a_, Mehri _dämāt_, Harsusi _demāt_, East Jibbali _dämˁat_, Soqotri _ʻedmíˁa_ (Dolgopolsky 1999:20: Proto-Semitic *_dámaˁ-at-_, pl. *_damaˁ-ā́t-_; Militarev and Kogan 2000:49: *_dimˁ-(at-)_, explaining _a_ in the Arabic word as a result of the influence of -ˁ-). This word probably represents a compound consisting of two components, (a) 'blood' and (b) 'eye': (a) *_dám-_ > Akkadian _damu_ ~ _dāmu(m)_, ?Eblaic divine name _Da-mu_, Ugaritic _dm_, Punic _edom_, Hebrew _dām_, pl. _dām-īm_, st. constr. _dam_, pl. _dəmē_, Old Aramaic _dm_, Jewish Aramaic _dam_, st. emph. _dəm-ā_, West Syriac _dem_, st. emph. _dəm-ō_, Mandaic _dma_, Maˁlula _eḏma_, Arabic _dam_, pl. _dimā'_, dialect (North Yemenite) _damm_, Sabaic _dm_, Geˁez _dam_, Tigre, Tigray, Amhara, Harari, Gurage _däm_ 'id.', and Mehri _dəm_, Jibbali _dihm_, Soqotri _dīm_ 'pus' (Dolgopolsky 1999:90; Militarev and Kogan 2000:47–48); (b) *_ˁayn-_ > Old Akkadian, Old Assyrian _ēnu(m)_, Akkadian _īnu(m)_, Eblaic _a-na-a_ = gen.-acc. du. [_ˁayn-ay(n)_], Ugaritic du. st. constr. _ˁn_ [_ˁēnē_], du. st. abs. _ˁn-m_ [_ˁēnēma_], Phoenician du. st. constr. _ˁn_, Hebrew _ˁayin_, pl. (< du.) _ˁēnayim_, Old Aramaic du. _ˁyny_ 'my eyes', Biblical Aramaic _ˁayn-īn_ 'eyes', Jewish Aramaic _ˁayn-ɔ_ ~ _ˁen-ɔ_, Syriac _ˁayn-ō_, Mandaic st. abs. _aina_, st. constr. _ˁin_, Maˁlula _ˁayna_, Arabic _ˁayn_, Sabaic, Minean _ˁyn_, Jibbali, Soqotri _ˁayn_, Mehri, Harsusi _ˁāyn_, Geˁez _ˁayn_, Tigre _ˁən_, pl. _ˁəntat_, Tigray _ˁayni_, Amhara _ayn_, Argobba _en_, Gafat _inä_, Harari, Selti, Zway _īn_, Wolane, Soddo _in_, Chaha, Eža, Muher _en_, Gyeto _ayn_ 'id.' (Dolgopolsky 1999:24, 51, 74, 87; Militarev and Kogan 2000:28–29). A key to the solution consists in the Arabic sg./pl. opposition in the word _dam_, pl. _dimā'_ 'blood' (cf. pl. _nisā'_ 'women'), comparable with _šafat_ 'lip', pl. _šifāh_ or _mā'_ 'water', pl. _miyāh_ (Fischer 1972:§72a, d).

This means that the starting point of the word 'tear' could be formed in two ways: (a) *_dam-_ + *_ˁayn-_, leading to the result attested in Arabic; (b) *_dimā(')_ + *_ˁayn-_, resulting in the most widespread form *_dimaˁ-_. The loss of the final _-ayn-_ could also be explained in two ways: (c) apocope (cf. e.g. Eilers 1984–1986); (d) reanalysis of the compound *_damˁayn-_/ *_dimāˁayn-_ in *_damˁ-_/*_dimaˁ-_, plus the dual oblique marker and determiner *_-ay-nV_ (see above). The metaphor 'tear' = 'blood of the eye' has an analogy e.g. in Hittite _eshahru-_ 'tear' < *$H_1esH_2\mathring{n}$/-H_2 _k̂ ru-_ 'blood-acrid' (cf. Hittite _eshar_, gen. _eshanas_, so first Sapir 1939:181). The common Semitic etymon 'tear' represents a Semitic innovation from the point of view of the Afro-Asiatic etymology, but its components belong to the most archaic part of the Afro-Asiatic lexicon:

i. Semitic **dam-* ‘blood’:

Berber **ā-dīm*, pl. **ī-dāmman* > East: Siwa, Sokna, Fogjaha *idammən*, Awgila *dimmen*, Ghadames *damměn* | North: Kabyle *adim*, pl. *idammen*, Wargla, Mzabi, Ntifa *idammen* | West: Zenaga *dəmm-ən*^{h} (Naït-Zerrad 1999:338; reconstruction after Prasse 1974:196: Ahaggar *aɣil*, pl. *iɣallän* ‘shoulder’ < **ā-ɣîl* < **ā-ɣill* < **ā-ɣaHīl*, pl. **ī-ɣāllan* < **ī-ɣaHālan*).

Chadic: West: Sura *tɔ́ɣɔm*, Ankwe *tyem* ‘blood’; Dera *dôm*, Gerumai *nduma*, Kirfi *ndame*, Bole *dɔm* ‘id.’; Karya *tûm*; Ngizim *dɪdum*, Bade *tədm* ‘id.’ | Central: Tera *tòm* ‘id.’; Gude *idɪna* ‘id.’ (Jungraithmayr and Ibriszimow 1994:30–31; Stolbova 1987: 171 reconstructs West Chadic **dama/ *daHam*);

Cushitic **dim-/*dum-* ‘red’ > Central: Awngi *də́mmí* (Hetzron), Kunfäl *demé* (Cowley) ‘id.’ | East **dim-/*dum-* > Oromo *diim-aa*, Konso *tiim-*, Sidamo *dum-a*, Burji *duww-aa* ‘id.’ (Sasse 1982:59) | South: Qwadza *dimayi-* ‘id.’ (Ehret 1980:325);

Egyptian (from the Old Kingdom) *ʾidmj* ~ *dmj* ‘red linen’ (Erman and Grapow 1971: I, 153; Faulkner 1981:35, 313).

ii. Semitic **ʿayn-* ‘eye’:

Egyptian *ʿjn* *‘eye’ (Erman and Grapow 1971:I, 189), reconstructed on the basis of the sign ‘eye’ determining, for example, the word *ʿjn* ‘limestone’ (Erman and Grapow 1971:I, 191), cf. Arabic *ʿayn aš-šams* ‘limestone’ = ‘eye of the sun’ (Vycichl 1958:381);

Berber *n-H-y* ‘to see’ > North: Middle Atlas *annay*, Ait Warain *inni* | South: Ahaggar *eni*, intens. *hânney*, Ghat *eni*, Adagh *ënhëy*, Awlemidden *ënëy*, Taneslemt *enhy* ‘id.’ (Kossmann 1999:65, 78: **enhey*);

Chadic: West: **ʿayin-* ‘to see’ > Kofyar *naa*, Sura *náá*; Fyer *yaána*, Bole *inne*; Pa’a *hani*; Tule *yaani*, Geji *yenî* | Central: Ga’anda *ànni*, Nzangi *naan*, Bachama *ná*, Bata *nan*, *nī*; Glavda *nagh-* ‘id.’ (Jungraithmayr and Ibriszimow 1994:284–285; Stolbova 1987:228, 1996:78).

The present extra-Semitic parallels for the Semitic words ‘blood’ and ‘eye’ may illustrate the external etymology. Another task of etymology is to differentiate the inherited words from borrowings. On the basis of the regular phonetic correspondences it is possible to determine, for instance, that Arabic *ʿankabūt* ‘spider’ is borrowed from a source of the Aramaic type, which has Semitic **ṯ* >*t*, besides the regular reflexes *ṯ* and *š* in Arabic and Hebrew respectively. Aramaic of Targum *ʿakku/åḇīṯå* and Hebrew *ʿakkåḇiš* indicate Semitic **ṯ* (Fox 1998:28).

In the case of Arabic *timsāḥ*, pl. *tamāsīḥ* ‘crocodile’, it is possible to map a history of this zoonym. It is generally accepted that it is borrowed from the Late Egyptian or early Coptic designation of ‘crocodile’ prefixed by the feminine article *ti-*, cf. Coptic *msaḥ*, Demotic and Middle Egyptian *msḥ*, XVIII Dynasty fem. *msḥ-t*, early Egyptian (Old Kingdom) *mzḥ-t*, fem. *mzḥt*. The Egyptian word was twice transcribed in cuneiform *nam-su-ṭu* and *nim-ša-ḫu* = *n3 msḥ(w)*, where initial *na-/ni-* is the Neo-Egyptian definite article in its plural form, in the annals of the Assyrian King Tiglathpilesar I (Vycichl 1983:123).

Reanalysis of the article is also at stake in Arabic *ʾusquf*, pl. *ʾasāqif* ~ *ʾasāqifa* ‘bishop’, a loan from Greek *epískopos*, which was realized via Coptic mediation, as shown by the apparent identification of *epi-* with the Coptic masc. article *pi-*. Arabic *mīnā*, pl. *miyan* ~ *mawānī* ‘harbor’ is borrowed from Greek *limḗn* thanks to the identification of the first syllable of the Greek word with the Arabic article.

Folk etymology is a false etymology based on naive semantic interpretation of the word, usually neglecting historical context and/or sound laws; for instance the term *al-muʿallaqāt* has been interpreted as ‘suspended [poems]’, with a false reference to *ʿallaqa* ‘to suspend’, rather than *ʿallaqa* ‘to adorn’ < *ʿilq* ‘precious thing, object of value’. Etymological fallacy is a mistake committed when etymological meaning is taken anachronistically for the modern meaning of a word, for instance when the Sufi term *ṭarīqa* is interpreted only as ‘a way’, rather than ‘a religious brotherhood’.

BIBLIOGRAPHICAL REFERENCES

Dolgopolsky, Aron. 1999. *From Proto-Semitic to Hebrew phonology*. Milan: Centro Studi Camito-Semitici.

Dombrovski, Bruno W.W. 1994. "Das System der eblaitischen Zahlen im Vergleich zu anderen, vornehmlich in den semitischen und hamitischen Sprachbereichen". *Folia Orientalia* 30.39–76.

Ehret, Christopher. 1980. *The historical reconstruction of Southern Cushitic phonology and vocabulary*. Berlin: D. Reimer.

Eilers, Wilhelm. 1984–1986. "Apokopierte Vollreduplikation". *Orientalia Suecana* 32–35.86–95.

Erman, Adolf and Hermann Grapow. 1971. *Wörterbuch der aegyptischen Sprache*. Berlin: Akademie-Verlag.

Faulkner, Raymond O. 1981. *A concise dictionary of Middle Egyptian*. 4th ed. Oxford: Griffith Institute.

Fischer, Wolfdietrich. 1972. *Grammatik des klassischen Arabisch*. Wiesbaden: O. Harrassowitz.

Fox, Joshua. 1998. "Isolated nouns in the Semitic languages". *Zeitschrift für Althebraistik* 11.1–31.

Jungraithmayr, Herrmann and Dymitr Ibriszimow. 1994. *Chadic lexical roots*, II. Berlin: D. Reimer.

Kossmann, Maarten. 1999. *Essai sur la phonologie du proto-berbère*. Cologne: Köppe.

Militarev, Alexander and Leonid Kogan. 2000. *Semitic etymological dictionary*. I. *Anatomy of man and animals*. Münster: Ugarit-Verlag.

Naït-Zerrad, Kamal. 1999. *Dictionnaire des racines berbères*, II. Paris and Louvain: Peeters.

Prasse, Karl-Gottfried. 1974. *Manuel de grammaire touaregue (tăhăggart)*. IV–V. *Nom*. Copenhagen: Akademisk Forlag.

Sapir, Edward. 1939. "The Indo-European 'tears'". *Language* 15.180–187.

Sasse, Hans-Jürgen. 1982. *An etymological dictionary of Burji*. Hamburg: Buske.

Schippers, Arie. 1998. "The Hebrew grammatical tradition". *The Semitic languages*, ed. Robert Hetzron, 59–65. London and New York: Routledge.

Stolbova, Ol'ga V. 1987. "Sravnitel'no-istoričeskaja fonetika i slovar' zapadnočadskix jazykov" [Comparative historical phonology and the lexicon of the South Chadic languages]. *Afrikanskoe istoričeskoe jazykoznanie*, ed. V. Ja. Porxomovskij, 230–268. Moscow: Nauka.

——. 1996. *Studies in Chadic comparative phonology*. Moscow: Diaphragma.

Testen, David. 1985. "The significance of Aramaic". *Journal of Near Eastern Studies* 44.143–146.

Vycichl, Werner. 1958. "Grundlagen der ägyptisch-semitischen Wortvergleichung". *Mitteilungen des Deutschen Archäologischen Instituts, Abteilung Kairo* 16.367–405.

——. 1983. *Dictionnaire étymologique de la langue copte*. Leuven and Paris: Peeters.

VÁCLAV BLAŽEK (Masaryk University of Brno)

Euphemism

1. DEFINITION

Euphemism is a lexical resource in language, whereby an offensive or hurtful word/phrase is replaced with one that represents a less direct expression or carries a positive attitude. It is an important vehicle for creating cognitive synonyms in language: the original expression and its euphemistic counterpart come to share denotative meaning but differ in their attitudinal parameter. The two terms *zabbāl* 'garbage man' and *ʿāmil naḏ̣āfa* 'a cleanliness worker', for example, denote the same occupation in Arabic but the second reflects a positive social attitude toward this kind of job which is lacking in the first term. The second alternative is said to euphemize the first. Similarly, the military phrase *ʾiʿādat intišār* 'redeployment' is more acceptable to listeners/viewers than *insiḥāb* 'withdrawal', because it is less direct than the latter, despite the fact that both terms denote the same concept in military affairs.

The term 'euphemism' comes from Greek *euphēmismós*, which means the use of words of good omen. The Random House College Dictionary (1980:455) defines euphemism as "the substitution of a mild, indirect, or vague expression for one thought to be offensive, harsh, or blunt". More recently, Allan and Burridge (1991:14) offer this definition: "Euphemisms are alternatives to dispreferred expressions, and are used in order to avoid possible loss of face". Clearly, both definitions imply the intentional utilization of lexical resources by interactants to achieve the expression of politeness and demureness in human communication. A speaker's use of the Arabic common euphemism *al-marḥūm* 'the person given mercy, i.e. who died', for example, instead of the neutral *al-mayyit* 'the deceased' may be informed by the addressee's relation to the deceased. The speaker/writer will opt for the euphemism in an attempt to prevent loss of face if he/she believes that the addressee cares for the denotatum. In some cases, however, the speaker's use of a euphemism may be instigated by general social mores rather than the addressee's face wants. For example, the speaker may opt to employ the euphemism *ḏawū l-iḥtiyājāt al-xāṣṣa* 'those with special needs' instead of the direct *al-muʿāqūn* 'the handicapped' to express solidarity with the denotatum rather than maintain his/her face wants. Thus, euphemism may express both negative politeness, as illustrated in the former case, and positive politeness, as exemplified by the latter case. (For more information on politeness, see Brown and Levinson 1987).

2. EUPHEMISM IN ARABIC LINGUISTICS

The linguistics of euphemism in Arabic is extremely sparse. There are only a few brief mentions of *at-talaṭṭuf* or *at-talṭīf* (ʿAskarī, *Ṣināʿatayn*; cf. Maṭlūb 1996; al-Jatlāwī 1998). Historically, al-ʿAskarī's term *at-talaṭṭuf*, which fits the term 'euphemism' very well, hardly relates to this phenomenon as we understand it in contemporary linguistics. He defines it as *at-talaṭṭuf li-l-maʿnā al-ḥasan ḥattā tuhajjinuhu wa-l-maʿnā al-hajīn ḥattā tuḥassinuhu* (*Ṣināʿatayn* 482) 'to kindly manage the pleasant meaning to make it objectionable and kindly manage the objectionable meaning to make it pleasant'. His examples show clearly that what he means is the employment of a non-preferred expression in a context where it acquires pleasant connotations, or vice versa. This differs from what we know as euphemism, a resource that necessarily involves the utilization of an alternative expression to replace the original non-preferred one in an attempt to kindly manage meaning via euphemizing.

The lack of a clear treatment of euphemism in medieval rhetoric comes as a great surprise, especially for those who are aware of the striking breadth and depth of this discipline in Medieval Arabic linguistics. However, this absence cannot be attributed to a shortage of euphemisms in Classical Arabic. The *Qurʾān* alone constitutes a rich source for euphemistic expressions intended to avoid blunt or taboo expressions in areas such as sex and bodily effluvia, among others, for instance in the two verses *ʾiḏā jāʾa ʾaḥadukum min al-ġāʾiṭi ʾaw lāmastum an-nisāʾa*… (Q. 5/43) 'If one of you has come back from defecation or you have touched women…'; *fa-lammā qaḍā minhā waṭaran zawwajnākahā* (Q. 33/37) 'After he had got his need from her, we married you to her.' In these verses, sexual intercourse is euphemistically referred to as 'touching' in the first verse and 'getting his need from her' in the second. Similarly, the first verse euphemizes the act of 'shitting' by the employment of the technical term *al-ġāʾiṭ* 'defecation' in order to hide the social taboo regarding this bodily function.

More recently, Farghal (1995) interprets the process of euphemizing in Arabic in terms of conversational implicature (Grice 1975). In particular, he emphasizes the interaction between the politeness principle (Leech 1983) and Grice's maxims of conversation in euphemistic expressions. Euphemisms are viewed as pragmatic mechanisms that reflect the organic interlock between the politeness principle and conversational maxims. By way of illustration, the Arabic euphemism *waḍaʿa ḥaddan li-ḥayātihi* 'he put an end to his life' as a replacement for *intaḥara* 'he committed suicide' flouts both the maxim of quality (by being metaphorical) and the maxim of manner (specifically, the sub-maxim 'Be brief') in order to conversationally imply that the denotatum's life had been full of suffering; hence, from the speaker's point of view, it was good that he killed himself. This conversational implicature is missing in the neutral (but inherently negative) counterpart *intaḥara*. Similarly, the vernacular euphemism *ʾaʿṭāk ʿumru* 'he gave you his age' instead of the neutral *māta* 'he died' flouts the maxim of quality, and as a result, conversationally implies the speaker's wish that the addressee live long.

3. TYPES OF EUPHEMISM

Figurative expressions are the most common device for euphemizing in Arabic in areas such as death, bodily effluvia, sex, and so forth. To observe the richness of metaphor in euphemizing, consider the standard euphemism that views death in terms of a transference to another life and/or joining the supreme Agent, viz. *intaqalat ʾilā raḥmat Allāh/ad-dār al-ʾāxira/dār al-baqāʾ/ ar-rafīq al-ʾaʿlā/jiwārī rabbihā* 'she transferred to the mercy of God/the afterlife/the home of eternity/the supreme comrade/the neighborhood of her Lord'. These standard euphemisms effectively find their way into vernacular Arabic with regional and social phonological variation and may be supplemented with other vernacular death terms, viz. *ʾAlla-xtāru* 'God chose him', *xubzātu xilṣin* 'his bread ran out', *ʿumru ntahā* 'his age [life] ended', and *ʾamr Allāh nāfiḏ* 'God's order is inevitable', for instance, are often heard in many Levantine dialects. The common divider in these death euphemisms is their inherently fatalistic viewpoint, which may be regarded as a hallmark of Arab culture in general (Farghal 1993).

Of particular interest is the use of antonyms in Arabic euphemisms. Examples include *muʿāfā* 'healthy' for *marīḍ* 'sick', *baṣīr* 'sighted' for *ʾaʿmā* 'blind', *ʿāʾidūn* 'returnees' for *lājiʾūn* 'refugees', *majbūr* 'with a healing limb' for *maksūr* 'with a broken limb'. These positive expressions reflect the desired rather than the existing state of affairs and are reminiscent of another deeply-rooted tradition in Arab culture. Ugly personal names such as *Jaḥš* 'Donkey' and *kulayb* 'Doggie' were given upon birth to keep envy away; *Zaynab bint Jaḥš* 'Zaynab, daughter of Donkey' was one of the Prophet Muḥammad's wives. Such proper names are still used in parts of the Arab world. In Egypt, for example, family names such as *al-ḥayawān* 'animal' and *al-ḥimār* 'donkey', still designate big families. Apparently, the use of antonyms in euphemizing has taken an opposite direction from using negative terms, which are meant to drive envy or evil away.

Circumlocutions, another type of euphemism, paraphrase taboos or socially objectionable elements. Examples of cicumlocutions include *lam yuḥālifhu l-ḥaḏ̣ḏ̣* 'luck did not ally with him' instead of *fašila* 'he failed', *bāʾiʿa hawan* 'a seller of love' for *šarmūṭa* 'a prostitute', *ṭāraḥahā l-ġarām* 'he made love to her' instead of *nākahā* 'he fucked her', *ṣāḥibat aṣ-ṣawn wa-l-ʿafāf* 'owner of maintenance and chastity' for *al-ʿarūs* 'bride', and so forth.

Remodeling is a fourth type of euphemism. It essentially belongs to vernacular Arabic and involves the twisting of the phonological structure of existing taboo expressions for a euphemistic purpose. Popular examples in the Levant include *yilʿan dīkak/dīxak* 'damn your rooster/?' for *yilʿan dīnak* 'damn your religion!', *yilʿan ḥarīšak* 'damn your…?' instead of *yilʿan ḥarīmak* 'damn your kinswomen!', *mgayyir* '?' for *mʾayyir* 'horny', and *ganānī* '?' instead of *ʾanānī* 'selfish.' A related euphemizing process is ellipsis. Here the speaker falls short of uttering the complete taboo phrase. Examples of elliptical expressions such as *ʾaxū l-*…'brother of…', *yabn il-*… 'son of…' and *bint il-*…'daughter of…' function as incomplete imprecatives. In some cases and for the purpose of euphemizing, the imprecative formula is completed with a general word instead of an obscene one. Examples include the popular Egyptian euphemistic imprecative formula *yabnil ēh* 'son of what!' and the Levantine flippant imprecative *yilʿan šuġlak* 'damn your work!'

Euphemistic expressions may take the form of understatements. The Arabic word *naksa* 'setback' came into frequent official use after the Arab–Israeli 1967 Six-Day War as a euphemism for *hazīma* 'defeat.' This euphemism was not just a word. It provided the Arab world with a psychological frame of reference through which the late President Nasser of Egypt, King Hussein of Jordan, and al-Atasi of Syria were to emerge as heroes from that humiliating war. Conversely, some euphemisms may be realized as overstatements or hyperboles. Recent examples include *ʾumm al-maʿārik* 'mother of all battles', used by the ex-Iraqi regime instead of the neutral *ḥarb al-xalīj aṯ-ṯāniya* 'Second Gulf War [of 1991]' and *ġazwa wāšinṭūn wa-niyūrk* 'Campaign of Washington and New York' for *hujūm al-ḥādiya ʿašara min sibtambar* 'September 11th attack' in the words of Bin Laden's followers. The latter example delves deep into history in search of a phraseology that would revive Islamic religious sentiment and include fresh positive attitudes. In terms of normative Islamic practice, the use of *ġazwa* is associated only with the campaigns led by the Prophet Muḥammad. The infringement of this tacit agreement among Muslims stems from Bin Laden's awareness of the positive associations of the said term; he used it to euphemize an otherwise objectionable act of terror.

Euphemisms in Arabic may also arise as a result of borrowing of foreign words. One of the most common euphemisms of this type is the use of the loanword *madām* 'madame' for *zawja* or *mara* 'wife' in many urban areas of the Arab world because it carries a more positive attitude. Other examples include *twālēt* 'toilet' for *mirḥāḍ* 'toilet', *kwāfēr* 'coiffeur' instead of *ḥallāq* 'barber' and *sūbarmākit* 'supermarket' for *dukkān* 'shop'. Sociolinguistically, the use of such foreign loans instead of the native counterparts is usually taken to be indicative of the speaker's high level of education and social class (→ English loanwords, → French loanwords).

Finally, Arabic vernaculars as remote from each other as Jordanian Arabic and Moroccan

Arabic utilize what may be called 'euphemizers.' Euphemizers are intended to soften the impact of mentioning a taboo or a socially non-preferred expression. Moroccan Arabic and Jordanian Arabic employ the euphemizer *ḥāšāk/ḥīšāk* 'may this not apply to you!' right after the mention of what is deemed to be socially objectionable, for example, references to shoes, animals such as donkeys and pigs, and negative attributes such as recklessness and stupidity. Other euphemizers from Jordanian Arabic include *balā gāfyih* 'without double meaning, i.e., take what I said at face value', *balā zuġra* when asking someone about his tribal affiliation, and *baʿīd ʿannak* 'may this be far from you!' (for more details, see Farghal 2002).

Bibliographical references

Primary source

ʿAskarī, *Ṣināʿatayni* = ʾAbū Hilāl al-ʿAskarī, *Kitāb aṣ-ṣināʿatayni*. Ed. Qamiha Mufeed. Beirut: Dār al-Kutub al-ʿIlmiyya, 1989.

Secondary sources

Allan, Keith and Kate Burridge. 1991. *Language used as a shield and weapon*. New York and Oxford: Oxford University Press.

Brown, Penelope and Steven Levinson. 1987. *Politeness: Some universals in language*. Cambridge: Cambridge University Press.

Farghal, Mohammed. 2002. "Situational and discoursal social honorifics in Jordan". *International Journal of the Sociology of Language* 158.163–181.

Grice, H. Paul. 1975. "Logic and conversation". *Speech acts*, ed. Peter Cole and Jerry L. Morgan, 41–58. New York: Academic Press.

——. 1995. "Euphemism in Arabic: A Gricean interpretation". *Anthropological Linguistics* 37:3.366–378.

——. 1993. "Arab fatalism and translation from Arabic into English". *Target* 5.43–53.

Jatlāwī, al-Hādī al-. 1998. *Qaḍāyā al-luġa fī kutub at-tafsīr*. Sousse: Kullīyat al-ʾĀdāb.

Leech, G.N. 1983. *Principles of pragmatics*. London: Longman.

Maṭlūb, ʾAḥmad. 1996. *Muʿjam al-muṣṭalaḥāt al-balāġiyya wa-taṭawwuruhā*. Beirut: Lebanon Library Publishers.

Mohammed Farghal (Kuwait University)

Europe

The influx to European countries of (mainly) labor migrants from countries where Arabic is the language of daily communication started in the early 1950s. The present description of the status and development of the varieties of Arabic in Europe is based on studies carried out in the various European countries where Arabic-speaking immigrant groups live. Most relevant studies come from France, Germany, the Netherlands, the United Kingdom, Belgium, and Spain. These studies focus on Arabic as it is spoken by people from the most important Maghreb countries, i.e. Morocco, Algeria, and Tunisia, and to a lesser extent on the Standard variety of Arabic.

1. Introduction: Migration of Arabic-speaking minorities to Europe

A speaker of Arabic can be defined as a national from an Arabic-speaking country, but through processes of nationality erosion, when nationals of Arabic countries become citizens of European countries, the speakers of Arabic disappear from national statistical surveys. The criterion of birthplace has its disadvantages as well, since many second and third generation children born in Europe are not registered in national statistics as 'foreigners'. The combined birth criterion, in which speakers of Arabic are defined as such if they themselves or one or both of their parents are born in an Arabic-speaking country, seems the most plausible way of identifying Arabic-speakers. Based on nationality, EuroStat (1997) registers the following numbers of people originating from Morocco, Algeria, and Tunisia in the West European countries: Belgium 161,588, Germany 133,945, France 1,393,195, Italy 3,656, The Netherlands 167,887, and Great Britain 7,000 (see also Basfao and Taarji 1994; López García 1996; Vermes 1988; Extra and Gorter 2001). According to the Spanish National Institute of Statistics, 199,782 Moroccans live in Spain (Nouaouri Izrelli 2001). In Sweden there is an Iraqi community of 55,696 persons and another Lebanese community of 20,288 persons (statistics based on birth country; Nygren-Junkin and Extra 2003). If the combined birth country-nationality criterion is taken into consideration in the case of Moroccans in the Netherlands their numbers rise by more than a half, from 164,567 to 252,000. This kind of consideration does not hold for Germany though, where it is relatively hard to obtain German citizenship, and therefore the nationality criterion will cover the actual number of speakers of Arabic there reliably. In France, many speakers of Arabic hold French nationality. It is reasonable then to suppose for France as well an addition to the existing numbers of around 30 percent, which implies a total of nearly two million people in France from the Maghreb. The Arabic-speakers from the Near East, living mostly in France, and from the Arabian Peninsula, living mostly in the United Kingdom, are outside the scope of this entry. It is important to note that people from North Africa do not automatically have Arabic as a mother tongue. In Europe, many Moroccans (more than 50 percent) and a minority of Algerians have Tamazight (Berber) as a mother tongue. Most of them speak dialectal Arabic as well, though.

2. Arabic language vitality

Large-scale language surveys executed in several European cities among children of primary school age show that Arabic is a relatively vital language (Extra and Yağmur 2004). One of the goals of these surveys was to establish the language vitality index (LVI) of the languages used and mentioned by the children. For the following cities the LVI for Arabic was calculated (numbers in parentheses are of children who indicated Arabic to be their mother tongue): Hamburg (464) 57; The Hague (1,391) 56; Brussels (1,608) 52; Lyon (2,789) 52; Madrid (662) 69. For all cities combined the LVI of Arabic is 58. The Romani/Sinte language had the highest LVI with 70 and German the lowest with 33. With these scores the Arabic language groups in the surveys hold a middle position among the top 20 languages, indicating that Arabic is a vital language among its speakers. More specific data on language proficiency show that most pupils have a high understanding of spoken Arabic and are quite capable of speaking the language itself but that they have much lower skills in reading and writing. In most cases Arabic is spoken with fathers and mothers but much less with siblings and even less with friends. In general Arabic loses ground with older children. Of the total of 7,787 Arabic-speaking pupils in these surveys 43 percent said they had received lessons in Arabic.

3. Arabic language teaching and education

From the first appearance of Arabic-speaking migrants in European countries there has been a debate on the question of whether or not to teach the children of the migrants Arabic language in primary education in the context of the so-called 'Home Language Instruction' and if so, in what juridical and linguistic contexts (Obdeijn and de Ruiter 1998; Tilmatine 1997). Sweden was the first country to organize this kind of language teaching. As early as 1976 it implemented Home Language Instruction, including the teaching of Arabic to primary school children. Regrettably, at the beginning of the 1990s government spending cuts led to a severe deterioration of the system (Nygren-Junkin and Extra 2003). The effects of this teaching were from the start disputed and criticized. The level of participation differed strongly from country to country, with percentages varying from 70 percent in the Netherlands to 15 percent in France. In Spain, where migration from the Maghreb started later, in the 1980s and 1990s, the teaching of Arabic is partly in the hands of the educational authorities and partly in the hands of non-governmental organizations, such as ATIME (Asociación de Trabajadores Inmigrantes Marroquíes en España; see Broeder and Mijares 2003; López García and Mijares 2001; Franzé and Mijares 1999; López García and Berriane 2005). In Belgium, because of the strict language laws in that country, the teaching of Arabic has never gone beyond the experimental stage (Verlot a.o. 2003). In France, at the beginning of the 3rd millennium, a debate was held over the incorporation of Home Language Instruction into the existing system of 'enseignement de langues vivantes' in primary education (Akinci, de Ruiter, and Sanagustin 2004). In Germany, the policy concerning Home Language Instruction differs from state to state: in Nordrhein-Westphalen it is under the shared responsibility of local authorities and those of the countries of origin, which is also the case in Hamburg, Berlin, and Baden-Württemberg, while in Bavaria it is the responsibility of the local authorities only (cf. Fürstenau a.o. 2003). The Netherlands have decided to abolish Home Language Instruction as of the school year 2004–2005.

Regarding teaching materials, in most countries materials developed in the countries of origin were and are used. Newly developed materials are scarce. The discussion about what variety of Arabic to teach, Standard or dialectal, played a prominent role in the debate in France (Caubet 2001; Caubet, Chaker, and Sibille 2002) and the Netherlands (Boumans and de Ruiter 2002). A European project, Comenius, led to the development of a course in dialectal Moroccan Arabic for elementary and secondary education, suitable for all Western European countries mentioned (Aarts and de Ruiter 1998; Abu Haidar and Bos 1998, 2000a, 2000b; Benjelloun, Bos, and de Ruiter 2001). The dialect for the course was written in Arabic script. In France the Institut du monde arabe (<www.ima.org>) published, not only many materials in Standard Arabic like the periodical *Al-Mukhtaaraat*, but also an interactive CD-ROM for children in Moroccan Arabic (Dumas and Laamiri 1997).

In secondary education, Arabic is taught in France and the Netherlands. The level of participation in both countries has never been high. Absolute numbers indicate that in the school year 2002–2003 some 3,000 pupils in the Netherlands followed this type of education and around 10,000 in France. Only Standard Arabic is taught. In the Netherlands the government decided in 2003 to discontinue financial support for the teaching of Arabic but did not prevent schools from offering it. Yet, the Dutch ministry of education has been very supportive in the development of teaching materials of Arabic and the development of Dutch/Arabic and Arabic/Dutch dictionaries. France offers pupils in secondary education the possibility of sitting for a final examination in more than 60 languages, among which there are 5 dialects of Arabic: Moroccan, Algerian, Tunisian, Egyptian, and Levantine. The measure of participation in these examinations is invariably high, with 5,000 students participating in the 2004 Arabic examinations. Benjelloun (1998a, 1998b, 2000a, 2000b) developed a course in Moroccan Arabic for secondary education.

The debate on writing the informal languages of Moroccans in particular led in Germany to a project aimed at the writing of their mother tongues, Moroccan Arabic and Amazigh (Maas and Mehlem 2003; Maas, Mehlem, and Schröder 2004). This had two goals: a scientific one to establish what processes take place if speakers of a non-codified language start writing their language; and a more applied goal, the codification of both mother tongues of the Moroccans. Research was carried out in both Germany and Morocco. The corpus of the project was formed by 73 spontaneously written texts in Moroccan Arabic and Amazigh. The choice of the writing system was free, but the majority of the children in Germany chose Latin script (62), while only a small group wrote in Arabic characters (11). In Morocco, almost exclusively Arabic script was used, for both Moroccan Arabic and Amazigh. One of the results of the project is that a majority of the Moroccan children who grew up in Germany not only succeeded in acquiring basic orthographical notions of German but also transferred this knowledge when writing spontaneously their non-written vernacular language

The renewed existence of Arabic in Europe led to the establishment of new academic programs in Arabic, its dialectal varieties in particular. In France, INALCO (Institut national des langues et civilisations orientales, <www.inalco.fr>) offers an M.A. program in Maghrebi Arabic. The University of Cadiz (<www.uca.es>) in Spain offers courses in dialectal Moroccan Arabic, Amazigh, and North African culture in its master's program of Arabic and Islamic studies.

4. ARABIC LANGUAGE PROFICIENCY AND LANGUAGE BEHAVIOR

Studies of language proficiency in Arabic concern mostly members of second generation Arabic-speaking children, and to a much lesser extent first generation people. Furthermore, most studies opt for a bilingual or multilingual format. De Ruiter (1989) studied the multilingual development of young Moroccans in Dutch, Moroccan Arabic, and where applicable Berber. Applying a semi-longitudinal model, he measured language proficiency of children and youngsters in four groups aged 7, 11, 14, and 21, each consisting of 20 Moroccans of whom 10 were Arabophone and 10 Amazighophone. The results point to a relatively weak proficiency in Moroccan Arabic compared to Dutch in the three younger groups and a relatively stronger proficiency in the oldest group. The patterns of language use, dominant use of Arabic with parents and much less use of Arabic with siblings and friends are similar to the European patterns found in the Extra and Yağmur data (2004; see above, section 2).

In Germany, Mehlem (1998) performed an elaborate linguistic inventory, comparable to de Ruiter's research of 1989, among 28 children of Moroccan descent who had frequented German schools from the start. The average age of these children was 11 years 8 months. Mehlem tested the proficiency of all these children in their mother tongues and compared it to their proficiency in German. In general, the children attained higher levels of performance in German than in their mother tongues.

In Spain, Nouaouri Izrelli (2001; forthcoming) did a study similar to those of de Ruiter (1989) and Mehlem (1998). He found that young Moroccans, aged 4–17, living in Andalusia had a better proficiency in Spanish than in their mother tongues. Furthermore they used their mother tongue predominantly in contacts with their parents and Spanish only with siblings and friends.

Aarts, de Ruiter, and Verhoeven (1993) report on a study on the proficiency in Standard Arabic of Moroccan children at the end of primary education in the Netherlands in the context of Arabic language teaching (see also section 3). Four language tasks were performed by 222 pupils. The scores of the pupils on word decoding are high at 79 percent. Their scores on reading comprehension are reasonable (50 percent), but the scores on spelling (30 percent) and written vocabulary (34 percent) are extremely low. The authors conclude that the limited amount of Arabic language instruction that the pupils in the Netherlands receive at primary level does not suffice to attain a high level of Arabic language proficiency. Saidi (2001) argues that studies such as those by Aarts, de Ruiter, and Verhoeven (1993) took pupils at random without taking into consideration the often chaotic organization of Arabic lessons. Eager to establish a more reliable picture of the results of Arabic language teaching to Moroccan pupils, he decided to test the proficiency of those Moroccan pupils who had followed seven to eight years instruction in Arabic uninterruptedly within schools. His results point to a higher proficiency in Arabic than in the study of Aarts a.o. (1993). Nevertheless, the successful schools Saidi selected for his study cannot be considered representative for all schools where Home Language Instruction was offered.

Pupils who study Arabic at secondary level in the Netherlands and France, few as they are, are able to attain a reasonably high level. In the Netherlands, Citogroep (<www.citogroep.nl>) is responsible for the development of examinations in Standard Arabic for secondary education, and from the internal guidelines of this institution it is clear that the level of Standard Arabic the pupils attain can be compared to that achieved in French and German in these same schools. Also, in France the levels of the examinations in Standard Arabic point to high proficiency in reading and writing Standard Arabic and the examinations in dialectal Arabic can only be accomplished successfully if candidates have a relatively fluent communicative proficiency in the dialect they opted for. In the Netherlands, Diephuis a.o. (1993) developed a handbook with guidelines for the Arabic examinations for all layers of secondary education. In France, the ministry of education regularly publishes similar guides.

First generation Arabic-speaking migrants have in general a low command of Standard Arabic. Only a few people from this group followed the full educational programs in their native countries or additional education in Europe. Illiteracy is more the rule than the exception, especially for women of this generation (De Ruiter 2000). El Aissati (1997, see also below) shows that the migration setting of young Moroccan adults affected the richness of their language, i.e. Moroccan Arabic, compared to similar young adults in Morocco. Broeder (1992) shows that first generation Moroccans are very poor in Dutch but they can express themselves very well in Moroccan Arabic. With the ongoing reunion of families in all European countries it is observed that higher educated young brides and grooms come to Europe. In general, they have a relatively good command of Standard Arabic. It goes without saying that proficiency in the mother tongues, the diverse Arabic vernaculars, is good in all members of the first generation. Through the 1960s to the 1990s local authorities in the diverse European countries, the Netherlands and Germany in particular, developed information materials in Arabic dialects, both in written and in audiovisual form. At the beginning of the

21st century this kind of information service declined as countries adopted stricter language policies, implying that migrants should know and use the languages of their new countries of residence.

5. Bilingual development and language change

Bos (1997) studied the bilingual development of Arabophone Moroccan children living in the Netherlands. In a pseudo-longitudinal design, she followed the acquisition of grammatical and pragmatic skills in both Moroccan Arabic and Dutch by children aged 4 to 11. Her experimental group consisted of 45 bilingual 4-year-olds and 45 8-year-olds living in different Dutch cities; there was a control group in Morocco. She administered two experimental tasks and a story-telling task, in both Moroccan Arabic and Dutch. The first experimental task concerned sentence-internal anaphoric reference. The youngest informants had the highest scores on the non-reflexive items, while at the age of 6 and 7 they did better on the reflexive ones. All Bos's informants eventually reached a higher than 90 percent score for both sentence types in both languages. However, the bilinguals showed a slower rate, reaching this level at the age of 10, two years after the Dutch and the Moroccan monolinguals. The second experimental task tested the children's understanding of relative clauses. Bos found that the monolingual children in Morocco performed better than their peers in the Netherlands on sentences with OVS order. Remarkably, the bilingual children performed best on SVO sentences, while the monolingual children found OVS sentences easier to process. Finally, Bos administered a story-telling task in order to investigate the children's narrative skills, in particular reference to topics and to temporality. One of the things she investigated was the development of the means of referring to topic characters in the story. She found similar patterns of development over time for both languages and for both the bilingual experimental group and the monolingual control groups. The bilingual informants did not suffer any delay in the acquisition of reference to protagonists (for more details on tense and aspect see Bos 1997). Nouaouri Izrelli (forthcoming) adopted a similar format in his study of the acquisition of temporality in the discourse of bilingual children, speaking Moroccan Arabic and Spanish.

El Aissati (1997) studied Moroccan Arabic in the Netherlands from the perspective of → language loss, as a consequence of the second generation's diminished exposure to the language. His 25 Moroccan informants (aged between 13 and 17) lived in the Netherlands. They filled out a questionnaire reporting on their oral skills in Moroccan Arabic and writing and reading skills in Standard Arabic, as well as on their language choice in various situations. The data on self-assessed proficiency in Moroccan Arabic were supplemented by a panel of two native speaker linguists who rated samples of semi-spontaneous narratives produced by the informants, a procedure which was also followed by Nortier (1990; see also below). Proficiency in Moroccan Arabic turned out to correlate with the age of immigration to the Netherlands and the reported amount of use of the language in daily interactions. The informants' linguistic performance was compared to that of a control group consisting of 30 Moroccans living in the Moroccan cities of Casablanca, Tangier, and Oujda. The experimental and the control group took part in four experimental settings and also produced semi-spontaneous material. First, plural formation was studied on the basis of experimental data. In the experimental setting the informants were asked to provide the plural form of nouns presented to them in the singular and out of context. The outcomes were then compared to those obtained from the control group. The participants of low language proficiency used fewer plural formation strategies. El Aissati concludes that one cannot speak of the emergence of a new, immigrant variety, but rather of individual paradigmatic leveling. Later El Aissati studied language use among Moroccan adolescents in the city of Utrecht in The Netherlands. There was no evidence of a new variety of dialectal Arabic developing in a migration context (El Aissati 2002), which was confirmed by Boumans (2001; see also below).

6. Code-switching

Nortier's study on Dutch/Moroccan Arabic → code-switching is based on the spontaneous conversations of 15 Arabophone Moroccans living in the Netherlands. Boumans (1998) discusses

15 informants, but these come from different parts in the Netherlands. The patterns of code-switching are heterogeneous but similar in the two data corpora. The most conspicuous traits of Moroccan Arabic/Dutch code-switching in comparison with code-switching varieties in general are the omission of the Moroccan Arabic definite article before Dutch nouns, and the use of an auxiliary verb in combination with Dutch verbs. One typically finds forms like Moroccan Arabic/Dutch *waḥed Ø-gesprek* 'a conversation', *dak Ø-examen* 'that examination' instead of Moroccan Arabic *waḥed l-ḥiwaṛ* and *dak l-imtiḥan*. The use of Moroccan Arabic auxiliary verbs is illustrated in (1):

(1) škun ġadi y-dir-hom *controler-en*?
who FUT 3-do-3PL supervise-INF
'Who is going to supervise them?'
(Moroccan Arabic/Dutch, Boumans 1998:231)

When Dutch verbs are used in otherwise Moroccan Arabic clauses, some speakers use the Moroccan Arabic verb *dar/ydir* 'to do' (less commonly *ʿmel*) which is inflected for all verbal categories like subject agreement and tense and aspect. The Dutch verb is inserted in its infinitive form. Some speakers also mark the pronominal object of the inserted verb as a suffix on *dar*, as in the above example. This way of treating Dutch verbs is far more frequent and widespread in Boumans's text corpus than in the earlier data described by Nortier (1990). At first sight, this might reflect a regional distribution of the construction or else its spread and growing conventionalization in the short time between the first and the second data collection in 1986 and 1991–1992. However, it is much more likely that the random and rather small sample of informants caused this difference between the two data sets. In Boumans's text corpus, almost all types and tokens of the *dar* plus infinitive construction are found in the speech of four informants who are siblings. From other observations and recordings by Boumans (2001; see also Boumans and Caubet 2001) it was found that the construction as such is common in Utrecht and elsewhere, although not in use by all code-switching bilinguals. The use of object suffixes like *-hom* in (1) has thus far been attested only for a couple of speakers, however.

Nortier (1990) relates individual speakers' code-switching patterns to their competence in Moroccan Arabic and Dutch. A combination of a number of criteria (self-report, actual language choice during the recordings, and evaluation of text samples by panels of native speakers of Moroccan Arabic and Dutch) assessed the bilingual competence, i.e. competence in both languages, of the individual informants. Relating competence to code-switching patterns, Nortier concludes that speakers with a high degree of bilingual competence produce relatively many switches within sentences, whereas Dutch-dominant and Moroccan Arabic-dominant speakers produce more switches between sentences and between sentences and discourse markers. Both studies on Moroccan Arabic/Dutch are for a large part concerned with the discussion of code-switching in general and of how it should be analyzed. Nortier sets her data against the various constraints on code-switching that had been proposed at the time. Boumans advocates viewing code-switching in terms of hierarchically ordered insertion rather than alternating language systems.

Wernitz (1993) reports on research on code-switching in French/Moroccan Arabic speech among young Moroccans living in France. She is concerned with the motivation of speakers to speak one language or the other, but she also pays attention to some linguistic characteristics of her data. She notes, for example, that in Moroccan Arabic/French speech French nouns may be modified by Arabic adjectives, while the reverse, an Arabic noun with a French adjective, does not occur (cf. Boumans 2002). Canut and Caubet (2002) deal with dialectal Arabic/French code-switching as recorded in France. Their conclusions go in the same direction as Boumans's (1998) analysis of code-switching in terms of matrix and embedded language.

7. Arabic language in the cultural scene

Many artists from North African origin are active in European theatre, music making, and literature. Although their output in the form of theatre plays, songs, and novels and poetry is mostly in the language of the European country in question, some productions are composed in Arabic, be it dialectal or Standard. The productions of these artists have started to attract the cultural interest of researchers in West European countries. The output of this new kind of artists adds to the existing culture in Europe and leads to new forms of intercultural arts. Caubet (2004) presents some of the leading artists of Maghrebi origin in France, among them Fellag, Baâziz, and Cheb Sahraoui. The book describes the development of these artists and how they view their contribution to French art and culture. Caubet (2005) studies artists with a Moroccan background in the Netherlands. She interviewed, among others, writers such as Abdelkader Benali and Hafid Bouazza, both of whom won important literary prizes in the Netherlands for their novels in Dutch.

Bibliographical references

Aarts, Rian and Jan Jaap de Ruiter. 1998. *Guide to the series "My own language"*. Tilburg: Syntax Datura.

——, Jan Jaap de Ruiter, and Ludo Th. Verhoeven. 1993. *Tweetaligheid en schoolsucces.* (= *Studies in Meertaligheid*, 4.) Tilburg: Tilburg University Press.

Abu-Haidar, Farida and Petra Bos. 1998, 2000a, 2000b. *L-lugha dyâl-i (My own language. Moroccan Arabic for primary education)*, Texts 1, 2; Exercises 2 (with tapes). Tilburg and Oisterwijk: Syntax Datura and Dutch University Press.

Akinci, Mehmet-Ali, Jan Jaap de Ruiter, and Floréal Sanagustin. 2004. *Le plurilinguisme à Lyon: Le statut des langues à la maison et à l'école.* Paris: L'Harmattan.

Basfao, Kacem and Hind Taarji (eds.). 1994. *L'annuaire de l'émigration.* Rabat: Ministère de l'émigration.

Benjelloun, Said 1998a, 1998b, 2000a, 2000b. *L-lugha dyâl-i (My own language. Moroccan Arabic for secondary education)* Guide, 1, 2; Texts 1, 2 (with tapes). Tilburg and Oisterwijk: Syntax Datura and Dutch University Press.

——, Petra Bos, and Jan Jaap de Ruiter. 2001. *L-lugha dyâl-i (My own language. Moroccan Arabic for primary education)*, 3. Tilburg and Oisterwijk: Babylon and Dutch University Press.

Bos, Petra. 1997. *Development of bilingualism: A study on Moroccan children living in the Netherlands.* (= *Studies in Multilingualism*, 8.) Tilburg: Tilburg University Press.

Boumans, Louis. 1998. *The syntax of codeswitching: Analysing Moroccan Arabic/Dutch conversations.* (= *Studies in Multilingualism*, 12.) Tilburg: Tilburg University Press.

——. 2001. "Moroccan Arabic and Dutch: Languages of Moroccan youth in the Netherlands". *Languages and Linguistics* 8.99–122.

——. 2002. "Possessive constructions in Morocco and in the Netherlands". *In honor of David Cohen: Aspects of the dialects of Arabic today*, ed. Abderrahim Youssi a.o., 265–275. Rabat: AMAPATRIL.

—— and Dominique Caubet. 2000. "Modelling intrasentential codeswitching: A comparative study of Algerian Arabic/French in Algeria and Moroccan Arabic/Dutch in the Netherlands". *Arabic as minority language*, ed. Jonathan Owens, 113–180. Berlin: Mouton de Gruyter.

—— and Jan Jaap de Ruiter. 2002. "Moroccan Arabic in the European diaspora". *Language contact and language conflict phenomena in Arabic: A variation on a sociolinguistic theme*, ed. Aleya Rouchdy, 259–285. Londen: Curzon.

Broeder, Peter 1992. *Talking about people: A multiple case study on language acquisition.* Amsterdam: Swets and Zeitlinger.

—— and Laura Mijares. 2003. *Plurilingüismo en Madrid: Las lenguas de los alumnos de origen inmigrante en primeria.* Madrid: Centro de Investigación y Documentación Educativa.

Canut, Cécile and Dominique Caubet (eds.). 2002. *Comment les langues se mélangent: Codeswitching en francophonie.* Paris: L'Harmattan.

Caubet, Dominique. 2001. "Maghrebine Arabic in France". Extra and Gorter (2001:261–277).

——. 2004. *Les mots du bled: Création contemporaine en langues maternelles, les artistes ont la parole.* Paris: L'Harmattan.

——. 2005. *Shouf shouf Hollanda: L'importance des artistes d'origine marocaine sur la scène culturelle néerlandaise.* Casablanca: Tarik Editions.

———, Salem Chaker, and Jean Sibille. 2002. *Codification des langues de France.* Paris: Harmattan.

Diephuis, Rob, Maaike Hajer, Theun Meestringa, and Hetty Mulder. 1993. *Kerndoelen Arabisch in het basisonderwijs en de eerste fase van het voortgezet onderwijs: Advies van het Instituut voor Leerplanontwikkeling SLO aan de ABOP.* Enschede: Instituut voor Leerplanontwikkeling (SLO).

Dumas, François and Abdeljalil Laamiri. 1997. *Jamal et l'ogresse.* CD-ROM. Paris: Institut du monde arabe.

El Aissati, Abder 1997. *Language loss among native speakers of Moroccan Arabic in the Netherlands.* (= *Studies in Multilingualism*, 6.) Tilburg: Tilburg University Press.

——. 2002. "Moroccan languages and identity in a multicultural neighborhood". *Meetings at the crossroads: Studies of multilingualism and multiculturalism in Oslo and Utrecht*, ed. Anne Hvenekilde and Jacomine Nortier, 249–259. Oslo: Novus Forlag.

Extra, Guus and Durk Gorter (eds.). 2001. *The other languages in Europe.* Clevedon: Multilingual Matters.

—— and Kutlay Yağmur (eds.). 2004. *Urban multilingualism in Europe: Immigrant minority languages at home and school.* Clevedon: Multilingual Matters.

EuroStat. 1997. *Migration statistics 1996: Statistical document 3A.* Luxemburg: EuroStat.

Franzé, Adéla and Laura Mijares. 1999. *Lengua y cultura de origen: Niños marroquíes en la escuela española.* Madrid: Ediciones del Oriente y del Mediterráneo.

Fürstenau, Sara, Ingrid Gogolin, and Kutlay Yağmur. 2003. *Mehrsprachigkeit in Hamburg: Ergebnisse einer Spracherhebung an den Grundschulen in Hamburg.* Münster: Waxmann.

López García, Bernabé (ed.). 1996. *Atlas de la inmigración magrebí en España.* Madrid: Ministerio de Asuntos Sociales and Universidad Autónoma de Madrid.
—— and Laura Mijares. 2001. "Moroccan children and Arabic in Spanish schools". Extra and Gorter (2001:279–292).
—— and Mohamed Berriane (eds.). 2005. *Atlas de la inmigración marroquí en España.* 2nd ed. Madrid: Ministerio de Asuntos Sociales and Universidad Autónoma de Madrid.
Mehlem, Ulrich. 1998. *Zweisprachigkeit marokkanischer Kinder in Deutschland.* Frankfurt: P. Lang.
Maas, Utz and Ulrich Mehlem. 2003. *Schriftkulturelle Ressourcen und Barrieren bei marokkanischen Kindern in Deutschland.* Universität Osnabruck: Institut für Migrationsforschung und Interkulturelle Studien.
——, Ulrich Mehlem, and Christoph Schröder. 2004. "Mehrsprachigkeit und Mehrschriftigkeit bei Einwanderern in Deutschland". *Migrationsreport 2004: Fakten – Analysen – Perspektiven. Für den Rat für Migration*, ed. K.J. Bade, M. Bommes, and R. Münz, 117–149. Frankfurt and New York: Campus.
Nortier, Jacomine 1990. *Dutch-Moroccan Arabic code switching among Moroccans in the Netherlands.* Dordrecht: Foris.
Nouaouri Izrelli, Nadi Hamdi 2001. "Escolares marroquíes en la provincia de Cádiz: Algunos aspectos de su perfil sociolingüístico". *Al Andalus Maghreb* 8–9.145–176.
——. (forthcoming). *The acquisition of temporality in bilingual children's discourse.* Ph.D. diss.
Nygren-Junkin, Lilian and Guus Extra. 2003. *Multilingualism in Göteborg: The status of immigrant minority languages at home and at school.* Amsterdam: European Cultural Foundation.
Obdeijn, Herman and Jan Jaap de Ruiter (eds.). 1998. *Le Maroc au coeur de l'Europe: L'enseignement de la langue et culture d'origine (ELCO) aux élèves marocains dans cinq pays européens.* Tilburg: Tilburg University Press and Syntax Datura.
Ruiter, Jan Jaap. 1989. *Young Moroccans in the Netherlands: An integral approach to their language situation and acquisition of Dutch.* Ph.D. diss., Rijksuniversiteit Utrecht.
——. 2000. "The position of Muslim migrants in the European Union: Emancipation or marginalization?". *Cambridge Review of International Affairs* 12:2.254–266.
——. 2006. *Les langues des jeunes marocains.* Paris: L'Harmattan.
Saidi, Redouan 2001. *Modern Standard Arabic instruction for Moroccan pupils in elementary schools in the Netherlands: A study on proficiency, status and input.* Ph.D. diss., Tilburg University.
Tilmatine, Mohand (ed.). 1997. *Enseignement des langues d'origine et immigration nord-africaine en Europe: Langue maternelle ou langue d'Etat?* Paris: INALCO/CEDREA-CRB.
Verlot, Marc, Kaat Delrue, Guus Extra, and Kutlay Yağmur. 2003. *Meertaligheid in Brussel: De status van allochtone talen thuis en op school.* Amsterdam: European Cultural Foundation.
Vermes, Geneviève (ed.). 1988. *Vingt-cinq communautés linguistiques de la France: Les langues immigrées*, II. Paris: L'Harmattan.
Wernitz, Corinna 1993. *Bedingungen und Voraussetzungen für Sprachwechsel: Eine Untersuchung zum Sprachwechsel bei bilingualen Marokkanern in Frankreich.* Frankfurt: P. Lang.

JAN JAAP DE RUITER (Tilburg University)

Exclamation

The traditional opinions of Arabic scholars about exclamation (*taʿajjub*) in Classical Arabic are found in various discourses, not only in that of the grammarians, in fragments inserted in works belonging to grammar (*naḥw*), rhetoric (*balāġa*), foundations of grammar (*ʾuṣūl an-naḥw*), scholastic theology (*kalām*), Qurʾānic exegesis (*tafsīr*), etc. We find in fact two ways of speaking about exclamation in the Arabic sciences of the language:

i. the first in a limited acceptance of this concept, referring to an evaluative act, i.e. not only an act of admiration, but in general, that of preferring (*tafḍīl*) something above others in the same class, which expresses wondering and perplexity. In this case, the Arabic authors tend to use the term *taʿajjub*
ii. the second in an extended sense, linked to what modern studies on exclamation in various languages call 'expressiveness', 'affective speech', or 'affection'. In this case, Arabic authors do not systematically use the term *taʿajjub* to refer to the exclamatory formulas, structures, or turns (cf. Firanescu 2003 for a synthetic view on exclamation in Literary Arabic).

All approaches to exclamations in the Arabic grammatical tradition deal with both the 'form' (→ *lafḏ*) and the 'meaning' (→ *maʿnā*) in order to explain the exclamatory character of a structure or utterance, but the approaches differ. Some focus on the form, the expression, in order to arrive at the sense. This morphosyntactic perspective (approximately through the 8–10th centuries) is represented by scholars such as al-Farrāʾ, al-Kisāʾī, al-Māzinī, Ṯaʿlab, etc. Other approaches start from the sense, being interested only marginally in the form. This semantic-pragmatic perspective (11–14th centuries, with an intermediate stage) is represented by scholars such as al-Jurjānī, az-Zamaxšarī, Ibn Yaʿīš, as-Sakkākī, al-Qazwīnī, etc. The two perspectives may be combined in one and the same work, as is the case with Sībawayhi or al-Jurjānī (see Baalbaki 1983:12; Versteegh 1992:119), whereas in others they are easier to distinguish from each other.

Several structures are constantly discussed by Arabic grammarians as a 'conventional' part of the linguistic system, when dealing with exclamation as an evaluative act. These may be regarded as the prototypical expression of exclamation or 'the hard core' of the concept of exclamation (cf. Firanescu 2003:127–128):

i. the two structures with the same meaning that are called *fiʿlā t-taʿajjub* 'the two verbs of exclamation' or *ṣīġatā t-taʿajjub* 'the two exclamatory formulas': *mā ʾafʿala* + direct object (noun; if it is a pronoun, the formula becomes *mā ʾafʿala-hu*), which is the most frequent one in Modern Literary Arabic, and *ʾafʿil bi-hi*: *mā ʾakrama zaydan!* or *zaydun, mā ʾakrama-hu* 'how generous is Zayd!'; *ʾakrim bi-zaydin!* or *zaydun, ʾakrim bi-hi* 'how generous is Zayd!'. Some authors add a third formula, the pattern *faʿula*, which is rare in the texts of the grammarians and can be applied to verbs of Form I only.
ii. the 'blame and praise verbs' *niʿma* (for admiration, praise) and *biʾsa* (for rejection, blame): *niʿma r-rajulāni!* 'how good are these two men!'; *biʾsat al-jāriyatu* 'how bad is this [female] slave!'
iii. the compound particle *ḥabbaḏā: ḥabbaḏā l-ḥālu* 'how nice is the situation!'

Linked to these structures, as secondary exclamatory formulas, Arabic scholars mention certain other marked expressions:

i. vocatives (*nidāʾ*), marked by the particle *yā* (*yā la-l-ʿajabi* 'Oh, how wonderful!', *yā la-ka šāʿiran* 'What a wonderful poet you are!', *yā la-hu min rajulin* 'What a man!')
ii. expressions introduced by the relative *ʾayy* expressing admiration (*marartu bi-rajulin ʾayya rajulin* 'I have visited a man, what a man!')
iii. oath formulas, marked by the particle *li* (*li-llāhi, lā yuʾajjalu l-ʾajalu* 'By God, the appointed time will be not adjourned!')
iv. formulas of compassion and deploring (*nudba*) marked by *yā* or *wa-* and the 'affective' suffix *-āh* (*wa-Muḥammadāh* 'Oh, poor Muḥammad!')
v. 'nouns of verbs' (*ʾasmāʾ al-ʾafʿāl*, sg. → *ism al-fiʿl*) in a fixed form (*mabniyya*), expressing a request (*sukūtan* 'Silence!', *ruwaydan!* 'Easy!'), or an intensive, expressive assertion (*hayhāt al-ʾamal* 'He is so far away!', *šattāna mā bayna zayd wa-xālid* 'Zayd and Khalid are so different!', *surʿāna/buṭʾāna* 'How fast!/slow!')
vi. various oath formulas with exclamatory meaning containing the name *ʾAllāh*: *ʾayman ʾallāhi* 'I swear on God's blessing!'), etc.

Exhaustive lists including the exclamatory structures and expressions may be found in several modern works in Arabic (Ḥassān 1973; Hārūn 1979; Sāmarrāʾī 1990; and others) or other languages (Fleisch 1961; Cantarino 1974–1975). These works systematize from a purely descriptive morphologic-syntactic perspective the enormous quantity of information and linguistic facts linked to the exclamation within Arabic grammar. They speak (e.g. Sāmarrāʾī 1990:651–709) about two categories of exclamatory expressions: those which are conventional and treated in special chapters by the Arabic grammarians (*at-taʿajjub al-mubawwab la-hu*), and those that are not conventional, but can have an exclamatory meaning in context (*lā tadullu ʿalā t-taʿajjub waḍʿan bal bi-l-qarīna*).

Yet, Arabic grammarians observed that there were a great number of unmarked expressions which could acquire an exclamatory sense within an appropriate context, if uttered with a 'certain intentional meaning' or 'speaker's purpose' (*murād* or *qaṣd*). Therefore, a semantic-pragmatic perspective on exclamation was developed by the Arabic grammarians, who speak about the 'act' of exclamation in terms surprisingly reminiscent of those used in modern pragmatic theories, such as 'performative verbs', 'indirect speech acts', 'illocutionary acts', 'conversational implicatures', etc. (Austin 1962; Searle 1970; Searle and Vanderveken 1985).

Sībawayhi (*Kitāb* I, 303–330) presents several vocative formulas (*nidāʾ*) as possibly exclamatory, conveying by the speaker's illocutionary intent such meanings as threat, menace, pride, affliction, and complaint, which are transmitted through expressive speech acts. Ibn Fāris (*Ṣāḥibī* 183–194) developed a complex discourse about indirect conventional and non-conventional speech acts, among them various expressive acts realized by uttering assertive, interrogative, and imperative statements which acquire an exclamatory contextual value. As-Sakkākī (*Miftāḥ* 305–306), speaking about the ‘semantic generation’, points to the expressive component of certain illocutionary acts, realized by uttering exclamatory sentences (cf. Buburuzan [Firanescu] 1993, 1995, 2003). From the second half of the 13th century onward, the discourse of the rhetoricians (ʾAstarābāḏī, *Šarḥ al-Kāfiya*, al-Qazwīnī, *al-ʾĪḍāḥ*), on the concept of ‘performative’ (*ʾinšāʾ*) inaugurated an extensive discussion on the ‘subjective mood of speech’, including exclamation as a principal feature (cf. Larcher 1991:257–263).

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

ʾAstarābāḏī, *Šarḥ al-Kāfiya* = Raḍī d-Dīn al-ʾAstarābāḏī, *Šarḥ Kāfiyat Ibn al-Ḥājib*. 2 vols. Istanbul, 1310 A.H. (Repr. Beirut: Dār al-Kutub al-ʿIlmiyya, n. d.)

Ibn Fāris, *Ṣāḥibī* = ʾAbu l-Ḥusayn ʾAḥmad Ibn Fāris, *aṣ-Ṣāḥibī fī fiqh al-luġa al-ʿarabiyya wa-sunan al-ʿarab fī kalāmi-hā*. Ed. ʿUmar Fārūq aṭ-Ṭabbāʿ. Beirut: Maktabat al-Maʿārif, n.d.

Qazwīnī, *ʾĪḍāḥ* = Jalāl ad-Dīn Muḥammad ibn ʿAbd ar-Raḥmān al-Qazwīni al-Xaṭīb, *al-ʾĪḍāḥ li-muxtaṣar Talxīṣ al-miftāḥ*, in *Šurūḥ at-Talxīṣ*. 4 vols. Cairo: Maṭbaʿa ʿĪsā al-Bābī al-Ḥalabī, 1937.

Sakkākī, *Miftāḥ* = ʾAbū Yaʿqūb Yūsuf ibn ʾAbī Bakr Muḥammad ibn ʿAlī as-Sakkākī, *Miftāḥ al-ʿulūm*. Ed. Naʿīm Zarzūr. Beirut: Dār al-Kutub al-ʿIlmiyya, n.d.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. 5 vols. Ed. ʿAbd as-Salām Hārūn. Cairo: Būlāq, 1966–1977.

SECONDARY REFERENCES

Austin, John Langshaw. 1962. *How to do things with words*. London: Oxford University Press.

Baalbaki, Ramzi. 1983. “The relation between *naḥw* and *balāġa*: A comparative study of the methods of Sībawayhi and Ǧurǧānī”. *Zeitschrift für Arabische Linguistik* 11.7–23.

Buburuzan [Firanescu], D. Rodica 1993. “Exclamation et actes de langage chez Sībawayhi”. *Revue Roumaine de Linguistique* 38:5.421–437.

——. 1995. “Significations des énoncés et actes de langage chez Ibn Fâris”. *Proceedings of the Colloquium on Arabic Linguistics*, Bucharest, 29 August to 2 September 1994, ed. Nadia Anghelescu and Andrei Avram, I, 103–114. Bucharest: University of Bucharest, Center of Arab Studies.

Cantarino, Vincente. 1974–1975. *Syntax of modern Arabic prose*. 3 vols. Bloomington: Indiana University Press.

Firanescu, Daniela Rodica [see also Buburuzan] 2003. *Exclamation in modern literary Arabic: A pragmatic perspective*. Bucharest: Editura Universităţii din Bucureşti.

Fleisch, Henri. 1979. *Traité de philologie arabe*. II. *Pronoms, morphologie verbale, particules*. Beirut: Imprimerie Catholique.

Hārūn, ʿAbd as-Salām. 1979. *al-ʾAsālīb al-ʾinšāʾiyya fī l-luġa al-ʿarabiyya*. Cairo: Maktabat al-Xānjī. 2nd ed. 1989.

Ḥassān, Tammām. 1973. *al-Luġa al-ʿarabiyya: Maʿnā-hā wa-mabnā-hā*. Cairo.

Larcher, Pierre. 1991. “Quand, en arabe, on parlait de l’arabe… II. Essai sur la catégorie de *ʾinšāʾ* (vs. *ḫabar*)”. *Arabica* 38.246–273.

Sāmarrāʾī, Fāḍil Ṣāliḥ. 1990. *Maʿānī n-naḥw*, IV. Baghdad: Jāmiʿa Baġdād.

Searle, John R. 1970. *Speech acts: An essay in the philosophy of language* Cambridge: Cambridge University Press.

—— and Daniel Vanderveken. 1985. *Foundations of illocutionary logic*. Cambridge: Cambridge University Press.

Versteegh, Kees. 1992. “Grammar and rhetoric: Ǧurǧānī on the verbs of admiration”. *Jerusalem Studies in Arabic and Islam* 15.113–133.

DANIELA RODICA FIRANESCU
(Dalhousie University)

Existential → Locative

Extension → Semantic Expression

F

Fāʿil

1. DEFINITION

The *fāʿil* (lit. 'he who does') corresponds, in the analysis of the Arab grammarians, to the protagonist of the verb. It is the primary element to which the verb relates (Sībawayhi, *Kitāb* I, 33–34), meaning that the verb "does not go without it" (Sīrāfī, *Šarḥ* II, 267). In fact, the verb (→ *fiʿl*) and its *fāʿil* constitute a pair "each of whose two elements cannot go without the other, and which the speaker cannot do without" (*lā yaġnā wāḥidun min-humā ʿan al-ʾāxar wa-lā yajidu l-mutakallim min-hu buddan*; Sībawayhi, *Kitāb* I, 23). For the Arab grammarians, the *fāʿil* is a syntactic function expressed in logical, semantic terms. "[Each noun] for which a verb is constructed and which is governed by that verb in the nominative case is called *fāʿil* from the point of view of syntax, not of the reality of the action" (Sīrāfī, *Šarḥ* II, 266–267). This is the way ʾAbū ʿAlī al-Fārisī (Jurjānī, *Muqtaṣid* I, 327) defines the *fāʿil*, hence the early grammarians' definition of *fāʿil* as "each noun postpositive to a verb and to which this verb is predicated and related" (Ibn Jinnī, *Lumaʿ* 13). In this sense, the term *fāʿil* applies to an active as well as a passive verb (Ibn Xālawayhi, *ʾIʿrāb* 70). Later grammarians called the passive verb *nāʾib al-fāʿil* 'substitute for the *fāʿil*'. This term, coined by Ibn Mālik in the 13th century (ʾAhdal, *Šarḥ al-Kawākib* 82–83), came to compete with longer syntagms in use at the time, such as *mā lam yusamma fāʿilu-hu* 'that whose *fāʿil* is not indicated' (Ibn Hišām, *Šarḥ Šuḏûr aḏ-ḏahab* 159), and which definitively replaced them in the works of the very late grammarians.

2. THE IMPLIED *FĀʿIL*

Unlike the first two personal pronouns present in the act of speech, Arab grammarians call *ġāʾib* 'absent' the morpheme of the 3rd person pronoun, considered by Benveniste (1966:228) as a non-person (see refutation in Joly 1973:59–97). In fact, Arabic places *kataba-ø* 'he wrote' in opposition to *katab-tu* 'I wrote' and *katab-ta* 'you wrote'. In the case of *kataba-ø*, Sībawayhi (*Kitāb* II, 6, 352) refers to a pronoun with no sign, or whose sign is not indicated (*al-ʾiḍmār allaḏī laysat la-hu ʿalāma ḏ̣āhira*, or *allaḏī lā ʿalāmata la-hu*), it being understood that the absence of a sign is considered by Arab grammarians to be a sign (Ibn al-ʾAnbārī, *ʾInṣāf* I, 46), which is not very far from the idea of the zero-significant morpheme in modern linguistics.

The *fāʿil* and its substitute are not deleted (*maḥḏūf*) but always implied (*mustatir*) and never used explicitly. Thus in a statement such as *kataba-ø* + *huwa* = 'he wrote, him', the apparent pronoun *huwa* 'him' would be quite rightly analyzed as an epithet (*ṣifa*) or corroboration (*tawkīd*) of the *fāʿil* /ø/ = *huwa* = 'he', still being implied (Sībawayhi, *Kitāb* II, 351, 378), because this implied pronoun /ø/, which carries no sign, is considered to have the status (*bi-manzila*) of a pronoun with a sign (*Kitāb* II, 351). In fact, the free pronoun *huwa* cannot be substituted for the attached, implied pronoun /ø/ which represents the same person, exactly as the first two pronouns, *ʾana* and

ʾanta, as in *katab-tu* + *ʾana* ‘I wrote, me’ and *katab-ta* + *ʾanta* ‘you wrote, you’, cannot be substituted for the equivalent attached pronouns: *-tu* ‘me’ and *-ta* ‘you’, which are always *fāʿil*: **kataba* + *ʾana* and **kataba* + *ʾanta*.

3. The *fāʿil* and the personal morpheme

Unlike the first two persons, which do not replace nouns and which are in a way nouns, the implied pronoun of the 3rd person is anaphoric in Arabic grammatical thinking. It therefore needs to refer to an antecedent. If you were to say *kataba* ‘he wrote’ “without referring to someone in particular and without the person you are speaking to knowing that you are indicating someone, it would not be a [complete] statement (*lam yakun kalāman*)” (Ibn as-Sarrāj, *ʾUṣūl* I, 41). Consequently, in a statement such as *kataba* + *zayd-u-n* ‘he wrote, Zayd [nom.]’, the proper noun *zayd*, which becomes an indispensable element “for lack of conditions permitting the use of a personal morpheme as an anaphoric or a deictic” (Touratier 1989:351), is analyzed as the *fāʿil* of the verb.

Such an analysis of verbs with a suffix could be extended to verbs with a prefix. Since a verb can have only one *fāʿil*, the prefix *y-*, as in *y-aktubu* + *zayd-u-n* ‘he writes, Zayd [nom.]’, cannot be identified as a personal morpheme. The same would apply to the prefixes *-ʾ*, *t-*, and *n-* as in *ʾ-a-ktub-u* ‘I write’, *t-a-ktub-u* ‘you write’, and *n-a-ktub-u* ‘we write’, which are identified as *ḥurūf* ‘particles’, but indicating, as al-ʾAstarābāḏī (*Šarḥ al-Kāfiya* I, 10) puts it, the meaning of a concrete noun, that of an implied pronoun, which is the *fāʿil* of the verb. In modern linguistic analysis, these prefixes, just as the suffixes, are analyzed as personal morphemes. However, *y-* of *y-aktub-u* is an epenthetic element in order to avoid a syllable that would otherwise have been anomalous (Roman 1983:873).

4. Position of the *fāʿil*

The position of the noun in relation to the verb is a determining factor in the concept of the *fāʿil* in the Arabic grammatical tradition (Hamzé 1999:127–149). In fact, the two statements *kataba* + *zayd-u-n* and *zayd-u-n* + *kataba* are not equivalent syntactically or even semantically (Jurjānī, *Dalāʾil al-ʾiʿjāz* 85–87; Ayoub and Bohas 1993:31–48). Quite rightly, two different analyses apply. In the first statement *zayd* is the *fāʿil*, but not in the second. Indeed, two types of substitution can show that the noun placed in front of the verb cannot be the *fāʿil* of that verb, since the *fāʿil* is always placed after the verb: *zayd-un kataba-ø* ‘Zayd, he wrote’ vs. *zayd-un kataba ʾabū-hu* ‘Zayd, his father wrote’ and *zayd-un kataba-ø* vs. *ʾana katab-tu* ‘me, I wrote’. If the element *zayd-un* placed in front of the verb were the *fāʿil*, it would be acceptable to say **ʾana kataba* *‘me, wrote’. This is a valid argument, the Ẓāhirite Ibn Maḍāʾ admits. But as the Ẓāhirites believe in the importance of the external meaning (*ḏāhir*) of the text of the *Qurʾān* for ideological reasons, Ibn Maḍāʾ is opposed to the assumption of implied elements (Versteegh 1997:146–148) and argues that proceeding by analogy from the 2nd and 1st persons to the 3rd person is not irrefutable proof. We may find the noun placed before the verb sufficient in the 3rd person, but not in the others (Ibn Maḍāʾ, *Radd* 92).

Bibliographical references

Primary sources

ʾAhdal, *Šarḥ al-Kawākib* = Muḥammad ibn ʾAḥmad ibn ʿAbd al-Bārī al-ʾAhdal, *Šarḥ al-Kawākib ad-durriya ʿalā Mutammimat al-ʾĀjurrūmiyya*. Cairo: al-Bābī al-Ḥalabī, 1937.

ʾAstarābāḏī, *Šarḥ al-Kāfiya* = Raḍī d-Dīn Muḥammad ibn al-Ḥasan al-ʾAstarābāḏī, *Šarḥ Kitāb al-Kāfiya fī n-naḥw*. Beirut: Dār al-Kutub al-ʿIlmiyya, n.d.

Ibn al-ʾAnbārī, *ʾInṣāf* = Kamāl ad-Dīn ʾAbū l-Barakāt ʿAbd ar-Raḥmān ibn Muḥammad ibn ʾAbī Saʿīd al-ʾAnbārī, *al-ʾInṣāf fī masāʾil al-xilāf bayna n-naḥwiyyīna l-baṣriyyīna wa-l-kūfiyyīn*. Ed. Muḥammad Muḥyī d-Dīn ʿAbd al-Ḥamīd. Cairo: Dār al-Fikr, n.d.

Ibn Hišām, *Šarḥ Šuḏūr aḏ-ḏahab* = Jamāl ad-Dīn ʾAbū Muḥammad ʿAbdallāh ibn Yūsuf Ibn Hišām, *Šarḥ Šuḏūr aḏ-ḏahab fī maʿrifat kalām al-ʿArab*. Ed. Muḥammad Muḥyī d-Dīn ʿAbd al-Ḥamīd. Cairo: al-Maktaba at-Tijāriyya al-Kubrā, 1965.

Ibn Jinnī, *Lumaʿ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Lumaʿ fī l-ʿarabiyya*. Ed. Hadi Kechrida. Uppsala: Almqvist and Wiksell, 1976.

Ibn Maḍāʾ, *Radd* = ʾAbū Jaʿfar ʾAḥmad ibn ʿAbd ar-Raḥmān Ibn Maḏāʾ al-Qurṭubī, *Kitāb ar-radd ʿalā n-nuḥāt*. Ed. Šawqī Ḍayf. Cairo: Dār al-Maʿārif, 1982.

Ibn as-Sarrāj, *ʾUṣūl* = ʾAbū Bakr Muḥammad ibn Sahl Ibn as-Sarrāj, *al-ʾUṣūl fī n-naḥw*. Ed. ʿAbd al-Ḥusayn al-Fatlī. Beirut: Dār ar-Risāla, 1988.

Ibn Xālawayhi, *ʾIʿrāb* = ʾAbū ʿAbdallāh al-Ḥusayn ibn ʾAḥmad Ibn Xālawayhi, *Kitāb ʾiʿrāb ṯalāṯīna* *sūra min al-Qurʾān al-karīm*. Beirut: al-Maktaba aṯ-Ṯaqāfiyya, 1991.

Jurjānī, *Dalāʾil al-ʾiʿjāz* = ʿAbd al-Qāhir al-Jurjānī, *Dalāʾil al-ʾiʿjāz*. Ed. ʾAḥmad Muṣṭafā al-Marāġī. Cairo: al-Maktaba al-Maḥmūdiyya at-Tijāriyya, n.d.

——, *Muqtaṣid* = ʿAbd al-Qāhir al-Jurjānī, *Kitāb al-muqtaṣid fī šarḥ al-ʾĪḍāḥ*. Ed. Kāzim Baḥr al-Marjān. Baghdad: Dār ar-Rašīd, 1982.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān ibn Qanbar Sībawayhi, *al-Kitāb, Kitāb Sībawayhi*. Ed. ʿAbd as-Salām Hārūn. Cairo, 1977–1979.

Sīrāfī, *Šarḥ* = ʾAbū Saʿīd as-Sīrāfī, *Šarḥ Kitāb Sībawayhi*, II, ed. Ramaḏān ʿAbd at-Tawwāb. Cairo: al-Hayʾa al-Miṣriyya al-ʿĀmma li-l-Kitāb, 1990.

SECONDARY SOURCES

Ayoub, Georgine and Georges Bohas. 1993. "La phrase nominale et le bon sens". *The history of linguistics in the Near East*, ed. Kees Versteegh, Konrad Koerner, and Hans-Josef Niederehe, 31–48. Amsterdam: J. Benjamins.

Benveniste, Emile. 1966. *Problèmes de linguistique générale*. Paris: Gallimard.

Hamzé, Hassan. 1999. "La position du sujet du verbe dans la pensée des grammairiens arabes". *Langage et linéarité*, ed. Pierre Cotte, 127–149. Lille: Presses Universitaires de Septentrion.

Joly, André. 1973. "Sur le système de la personne". *Revue des Langues Romanes* 80.59–97.

Roman, André. 1983. *Etude de la phonologie et de la morphologie de la koinè arabe*. Marseille: J. Lafitte.

Touratier, Christian. 1989. "Structure de la phrase simple en arabe". *Bulletin de la Société de Linguistique de Paris* 84.345–359.

Versteegh, Kees. 1997. "Ibn Maḍāʾ and the refutation of the grammarians". *The Arabic linguistic tradition*, 140–152. London and New York: Routledge.

HASSAN SALAM BAZZI HAMZÉ
(University of Lyon-2)

Farsi → Persian

Faṣīḥ

Ibn Jinnī (d. 392/1002) defines grammar (*naḥw*) as follows: "It is to follow the way the Arabs speak… so that the non-Arabs might have access to the Arabs' *faṣāḥa*" (*Xaṣāʾiṣ* I, 34). More than a thousand years later, written Arabic is still called *al-luġa al-fuṣḥā*. This shows how the notion of *faṣāḥa* is an essential component of Arab language thinking.

The root *f-ṣ-ḥ* is very ancient and is found in other Semitic languages. From *f-ṣ-ḥ* is derived *fiṣḥ* 'Jewish Passover', also 'Christian Easter' (Ibn Manẓūr, *Lisān*, s.v.). In some Semitic languages, *f-ṣ-ḥ* is explicitly associated with something clear, or bright: in Assyrian, *piṣū* signifies 'pure; bright'; in Aramaic, *paṣṣiḥ* signifies 'pure; radiant'. In 7th-century Arabic the notion refers to something pure, faultless, unaltered (*faṣḥ*). The verb *ʾafṣaḥa* means 'to become limpid [urine]; to be skimmed of its froth [milk]'; it refers to clearness, to the dazzling morning light (*ʾafṣaḥa ṣ-ṣubḥu*), and to a horse or donkey whose whinnying or braying is clear (*ʾafṣaḥa l-farasu wa-l-baʿīru*). Linguistically, *faṣuḥa wa-ʾafṣaḥa r-rajulu* refers to an enunciation both pure and clear. This seems to be the best match for classical texts, with the notion of correctness added. It is also the meaning retained by Blachère (1952:I, 119) when he translates the expression *fuṣaḥāʾ al-ʿArab* as 'the Arabs with pure and correct speaking'. According to as-Suyūṭī (d. 911/1505), the linguistic usage is a metaphor derived from the concrete meaning of the word. In Classical Arabic, it implies at the same time correctness of language and its aesthetic quality.

1. PRE-CLASSICAL LINGUISTIC USAGE OF THE NOTION *FAṢĀḤA*

The linguistic notion of *faṣāḥa* has a long history. In pre-Islamic usage and that of the 1st century A.H., the main sense is that of 'clearness' or 'intelligibility', rather than 'purity' (Ayoub 2003b). In fact, *faṣīḥ*'s antonym is *ʾaʿjam*, defined by Ibn as-Sikkīt (d. 244/858) and then by Ibn Sīda (d. 458/1066; *Muxaṣṣaṣ* I, 113) and Ibn Manẓūr (d. 711/1311; *Lisān* I, 2825) as "the one whose speaking is not clear, whether he is of Arab or foreign origin" (*allaḏī lā yubayyinu l-kalāma min al-ʿArab wa-l-ʿajam, allaḏī lā yufṣiḥu*). Ibn as-Sikkīt defines *al-faṣīḥ* as 'the one whose speaking is clear' (*al-bayyin*; Ibn Sīda, *Muxaṣṣaṣ* I, 112). In the entry *f-ṣ-ḥ*, the *Lisān* quotes a line by the poet ʾAbū n-Najm (d. 130/747) in which the poet describes a donkey as *ʾaʿjam* for human beings but *faṣīḥ* to the ears of its 'lover' (*aʿjama fī ʾāḏāni-hi faṣīḥa* 'unintelligible, but to her ears, of a dazzling clearness'; I, 2825). Another line quoted by the *Lisān* confirms this antonym in pre-classical usage: "It is a vital source for all creatures, the ideal of every unintelligible and intelligible creature" (*muntahā kull ʾaʿjam wa-faṣīḥ* lit. 'those expressing themselves clearly'; I, 2825).

After Ibn Sīda (*Muxaṣṣaṣ* I, 112–113), Ibn Manẓūr mentions another meaning for *faṣīḥ*, which is equivalent to being endowed with language, or being human: “It has been said that living beings are of two kinds: *ʾaʿjam* and *faṣīḥ*. The *faṣīḥ* is the living being gifted in language, whereas the *ʾaʿjam* is every living being not endowed with language” (*Lisān* I, 3419).

Obscurity of speech is to be understood in relation to non-Arabic speech (*kalām ʾaʿjam: yuḏhabu bi-hi ʾilā kalām al-ʿajam*; Ibn Sīda, *Muxaṣṣaṣ* I, 121). Additionally, the designation of *ʿajamī* refers to the foreigner’s language. This view was expressed at the end of the 8th century by ʾAbū ʿAmr aš-Šaybānī (d. 206/821), one of the first lexicographers, born in Kūfa and a contemporary of al-Xalīl, whose words are reported in the *Lisān* (Ibn Manẓūr, I, 2826): *ʾAʿjama*, he says, is *ʾabhama* ‘to make something indeterminable, obscure’, and “the *ʿajamī* has an obscure discourse; it is not clearly expressed” (*wa-l-ʿajamī mubham al-kalām, lā yubayyinu kalāma-hu*). Furthermore, *ʾaʿjam* signifies ‘foreigner’ (*man laysa bi-ʿarabī*), as Ibn Manẓūr points out under the lexical definition of *ʿajam*. Pre-Islamic poetry seems to confirm this. *Muxaṣṣaṣ* and *Lisān* quote lines by the poet ʾAbū l-ʾAxzar: “Oh hail! Sallūm, would you be among the non-Arabs [*al-ʾaʿjam* ‘those who speak unintelligibly’], with the Byzantines, the Persians, or the Daylamites?” (*Lisān* I, 2825). Likewise, *ʿujma* refers to both the quality of a speech that lacks clearness (*fī lisāni-hi ʿujma*) and the foreign origin of the word (*ʿujmat al-ism*) (Suyūṭī, *Muzhir* I, 270).

The sense of ‘clearness’ explains why *faṣuḥa* and *ʾafṣaḥa* may be used for the speech of both Arabs and non-Arabs. Ibn Sīda (*Muxaṣṣaṣ* I, 112–113) notes, following *Kitāb al-ʿayn,* that *faṣuḥa l-ʾaʿjam*, used to describe a non-Arab speaker, means ‘to speak Arabic’, whereas *ʾafṣaḥa* implies a better quality of enunciation (*izdāda faṣāḥatan*) in an Arab speaker (the reverse in Suyūṭī, *Muzhir* I, 184). In the former case, the enunciation becomes intelligible or clear, while in the latter, it becomes clearer or more eloquent.

This usage also seems to correspond to the Qurʾānic usage of *ʾafṣaḥ*, the masculine elative of *faṣīḥ*. Actually, the only attestation of a word derived from *f-ṣ-ḥ* in the *Qurʾān* is Q. 28/34: *wa-ʾaxī Harūnu huwa ʾafṣaḥu minnī lisānan fa-ʾarsil-hu maʿī* (uttered by Moses). In this context, *ʾafṣaḥ* must be understood in opposition with the way Moses describes himself as being tongue-tied (*lā yanṭaliqu lisānī... fa-ʾarsil ʾilā Hārūna,* Q. 26/13; *wa-ḥlul ʿuqdatan min lisānī yafqahū qawlī*, Q. 20/27). Aaron has a better quality of enunciation than Moses: his words are ‘more understandable’ (the root *f-q-h*) than those of Moses.

Pre-classical usage of *faṣāḥa* thus refers to clear, intelligible enunciation rather than to purity in the sense of absence of linguistic crossbreeding. Rabin (1960:579) believes that this was indeed the meaning of *faṣāḥa*, whatever the period considered. The pre-Islamic Arabs seem to have paid careful attention to the clearness of the enunciation and, consequently, to the language.

2. The classical notion of *faṣāḥa*

In Classical Arabic the most striking feature of *faṣāḥa* is that linguistic correctness, the quality of the enunciation, and its truthfulness are inextricably linked (Ayoub 2001). This is indeed a component of all Classical Arabic thinking on language. Thus, according to Ibn Jinnī, the verb *ʾaʿraba ʿan aš-šayʾ* ‘to express something clearly’ is a denominative verb derived from the term *ʿArab* ‘Arabs’ “because of all the pure elocution (*faṣāḥa*), limpid expression (*ʾiʿrāb*), and clear enunciation (*bayān*) ascribed to them” (*Xaṣāʾiṣ* I, 36). Therefore, *kalām al-ʿArab* is the equivalent of a ‘clear, eloquent language’. But the same verb, *ʾaʿraba*, also refers to correctness: *ʾaʿraba* is to speak the way the Arabs speak, using syntactic endings (*ʾiʿrāb*). In the same way, the notion of *faṣāḥa* denotes the correct and pure usage, as codified by the *ʿarabiyya*, and as such it is one of the basic notions of non-Greek rhetoric. For al-Ḥarīrī (d. 516/1122) in his *Durra*, correct speaking (*aṣ-ṣawāb*) is inextricably connected with eloquent speaking (*al-bayān, al-faṣāḥa, al-kalām al-faṣīḥ*). It is also pure, uncontaminated with any regional feature: although all dialectal variants collected by the grammarians are theoretically legitimate (*al-luġāt kullu-hā ḥujja*; Ibn Jinnī, *Xaṣāʾiṣ* I, 257), many of these variants are decried (Ayoub 2001:112–117). Furthermore, in the 8th century, the collection of linguistic data was directed against any crossbreeding, the philologists purposefully avoiding those tribes that might have borrowed from other languages. This is what the philosopher ʾAbū Naṣr al-Farābī (d. 339/950) asserts in a well-known text (Suyūṭī, *Iqtirāḥ* 17, *Muzhir* I, 211; cf. Renan 1863:451ff.; Rabin 1951:193; Blachère 1952:71). The great debate about foreign words in the *Qurʾān* as reported by as-Suyūṭī (*Muzhir* I, 266–294) confirms this anti-crossbreeding tendency (Kopf 1956; Haywood 1965; Versteegh 1990; Ayoub, forthcoming).

The notion of *faṣāḥa* is commented upon by grammarians, rhetoricians, poets, lawyers, and theologians alike. There is no *ʿarabiyya* without *faṣāḥa*, and there is no religious or legal science, no *ʾadab* without *ʿarabiyya*, a language stamped with the grammatical norm. Definitions of *faṣāḥa* are numerous, qualifying both the utterance and the speaker. Ṯaʿlab (d. 291/904) seems to link *faṣīḥ* with common use: "This book presents the *faṣīḥ* exclusively, what is common (*mā yajrī*) in people's speech and in their written work" (*Faṣīḥ* 2). In the following lines, he makes a distinction between various degrees of *faṣīḥ* (*faṣīḥ* and *ʾafṣaḥ*) according to this criterion. Going back to this usage, as-Suyūṭī (*Muzhir* I, 185) specifies that *faṣāḥa* qualifies those words used most frequently by Arabs whose language is reliable (*kaṯrat istiʿmāl al-ʿArab la-hā*). In the classical era, the *faṣīḥ* is linked to *kalām al-ʿArab*, the corpus of references whose pillars are the *Qurʾān* and pre-Islamic poems (Ḥarīrī, *Durra*, 116). This corpus, which is the basis of the → *ʿarabiyya* (Ibn al-ʾAnbārī, *Lumaʿ*), lays down the rules of any speech.

In lexicography, Ṯaʿlab's choice of 'the most frequent' (*al-ʾakṯar*) is a well-founded epistemological choice. It is to be understood in the light of the linguistic situation of ancient Arabia, when many dialects could be found. The description of dialectal features (→ pre-Islamic Arabic) is recorded in numerous classical works as early as in the first grammatical treatise, the *Kitāb* by Sībawayhi (d. 177/793), and is the object of many studies (Rabin 1951; Blachère 1952; Fück 1955). The methodological choice of the most common turns of phrase must be seen in the light of this dialectal variation. The same choice is made by several medieval scholars, who seem to distinguish clearly between the *ʿarabiyya* and the *kalām al-ʿArab*, the *ʿarabiyya* being the language codified by the grammarians (Ayoub 2001:95, 2003a:42). This is suggested by the way ʾAbū l-ʾAswad ad-Duʾalī (d. 62/681) is described as the first scholar to have established the *ʿarabiyya*: *ʾawwal man waḍaʿa l-ʿarabiyya* (Sīrāfī, *ʾAxbār* 13.2, 3, 5; Zubaydī, *Ṭabaqāt* 21.9; Suyūṭī, *Muzhir* II, 345), or put differently: *ʾawwal man waḍaʿa (rasama) n-naḥw* (Sīrāfī, *ʾAxbār* 10.3; Zubaydī, *Ṭabaqāt* 21.12; ʾAbū ṭ-Ṭayyib, *Marātib*, 27.1). There is no trace whatsoever in these texts of the expression *ʾawwal man waḍaʿa kalām al-ʿArab*. The distinction between *ʿarabiyya* and *kalām al-ʿArab* is implicitly made by az-Zubaydī (*Ṭabaqāt* 21.4–5): ʾAbū l-ʾAswad is the first to have established the *ʿarabiyya*, when the *kalām al-ʿArab* had been altered (*ʾawwal man ʾassasa l-ʿarabiyya… wa-ḏālika ḥīna iḍṭaraba kalām al-ʿArab*), a point explicitly made by az-Zubaydī (*Ṭabaqāt* 22, 39) and as-Suyūṭī (*Muzhir* I, 134). The debate deals precisely with the criteria that led ʾAbū l-ʾAswad – or ʿĪsā ibn ʿUmar – to establish the language he called *ʿarabiyya*: what makes it different from *kalām al-ʿArab*? The grammarian replies that the *ʿarabiyya* is the most common usage of *kalām al-ʿArab (al-ʾakṯar)*, whereas other usages are called *luġāt*. Hence, *ʿarabiyya* cannot be understood as 'pure Arabic' (Blachère 1952:71), at least not in the first centuries. The text emphasizes twice that this is a necessary epistemological choice: only by opting for the common usage was the grammarian able to write a grammar.

As-Suyūṭī (*Muzhir* I, 185–187) adds that iden-tifying the common usage in such a remote past is not easy, which is why scholars set up criteria for lexical items. He revives the criteria applied by al-Qazwīnī, a rhetorician of the 14th century (739/1338), in his *ʾĪḍāḥ*: sounds must be harmonious, the term must not be a rare word requiring long research work in dictionaries, it must comply with the general rule (*qiyās*), and it must not represent a marginal form of the language. The demand for immediate intelligibility of the meaning resembles the pre-classical notion of *faṣāḥa*, but the criterion of compliance with the *qiyās* bears the stamp of the norm, whereas the harmony of sounds refers to an explicit aesthetic concern.

3. THE ROOT *F-Ṣ-Ḥ* IN THE *KITĀB* BY SĪBAWAYHI

The elative *ʾafṣaḥ*, a concept not used in pre-Islamic poetry, appears in the *Kitāb* as a criterion for correct speech. But the *Kitāb* seems to distinguish *kaṯīr, ʾafṣaḥ*, and *qiyās*, as in the following passage: "It is the way of speaking of most of the Arabs, the Arabs with the *ʾafṣaḥ* way of speaking; it is also what complies with the general rule" (*wa-huwa kalāmu ʾakṯari l-ʿArab wa-ʾafṣaḥi-him wa-huwa l-qiyās*; Sībawayhi, *Kitāb* I, 184).

Occurrences of *faṣāḥa* in the *Kitāb* are fairly rare. They often qualify the speakers (*al-ʿArab al-fuṣaḥāʾ; fuṣaḥāʾ al-ʿArab*). Unlike later texts, the Arabs with the *ʾafṣaḥ* manner of speaking are called upon neither for the most common turns of phrase nor for forms that comply with the rule. 'The Arabs' (*ʿāmmat al-ʿArab; Kitāb* I, 252.10, 426.15) or 'the Arabs whose *ʿarabiyya* is reliable' (*man yūṯaqu bi-ʿarabiyyati-hi min al-ʿArab, Kitāb* I, 318.21; *al-ʿArab al-mawṯūq bi-ʿarabiyyati-him, Kitāb* I, 128.9; etc.) are the ones called upon for these turns of phrase. The *fuṣaḥāʾ al-ʿArab* are quoted for turns of phrases which are marginal but nevertheless used by them, and which must therefore be taken into account (Ayoub 2003:49–51). Thus, in the *Kitāb* (Sībawayhi II), ordinary usage suppresses the → *tanwīn* from nouns followed by the *kunya* (e.g. *hāḏā ʾabū ʿamri* [not * *ʿamrin*] *bni l-ʿalāʾ*). Yet, Sībawayhi quotes a line recited to him by *fuṣaḥāʾ al-ʿArab* where the *tanwīn* is kept for metrical reasons. In the *Kitāb* (II, 299.18), the *ʿArab fuṣaḥāʾ* say *min-a bni-ka*, instead of *min-i bnika*. Sībawayhi calls this usage a remarkable exception, after the fashion of *tazdīr* used by the *fuṣaḥāʾ* for *tasdīr* (*Kitāb* II, 477.2). A passage in the *Kitāb* (I, 426) explicitly expresses the discrepancy between the *kaṯīr* and *qawl fuṣaḥāʾ al-ʿArab: samiʿnā fuṣaḥāʾ al-ʿArab yaqūlūna wa-laysat fī kalām kull al-ʿArab*. But Sībawayhi resorts to them in order to strengthen a morphological hypothesis: "This line was recited to me in this way by one of the *ʾafṣaḥ*-speaking Arabs (*ʿarabī min ʾafṣaḥ an-nās*) asserting that it was his father's poetry" (*Kitāb* II, 48). To set up a *qiyās*, he asserts: "We heard it used by *fuṣaḥāʾ al-ʿArab,* and they don't accept anything else" (*samiʿnā ḏālika min fuṣaḥāʾ al-ʿArab lā yaʿrifūna ġayra-hu; Kitāb* II, 20.20). In short, Sībawayhi calls upon the authority of the *fuṣaḥāʾ al-ʿArab* as the final arbiters of the correctness of the language he studies (see also *Kitāb* I, 91.18; II, 40.20). The way of speaking of the *fuṣaḥāʾ al-ʿArab* defines what belongs to the language, even if it is not what is most common.

In fact, as explained by Ibn an-Nadīm (d. 385/995; *Fihrist*, 66–72; cf. Blachère 1950; Blau 1963), the *fuṣaḥāʾ al-ʿArab* were professional Bedouin informants of the grammarians from the 2nd through the 4th centuries A.H. Their knowledge of the grammatical metalanguage was often scant (Ibn Qutayba [d. 276/889], *ʿUyūn* II, 173; Ibn Fāris [d. 395/1004], *Ṣāḥibī* 35–36; Suyūṭī, *Muzhir* II, 343). Therefore, they were tested before their judgment about linguistic matters was asked (Ṯaʿlab in *Muzhir* II, 337). But they serve as the absolute reference as far as language is concerned, which means that they are the ones designated as *al-ʿArab al-mawṯūq bi-ʿarabiyyati-him*. Sībawayhi highlights their quality as *fuṣaḥāʾ* whenever he wishes to take into account some exceptional terms, which proves that the quality of *faṣāḥā* is a decisive criterion in selecting utterances.

4. *Faṣāḥa* and the epistemological foundations of Arab thought

In Sībawayhi's *Kitāb*, the *faṣīḥ* is far from being always the most regular and the most common spoken language. Yet, it defines what can be said in the language, i.e. what is correct, even though it goes beyond it. It refers to a sense of value exceeding what is correct. In fact, the choice of the corpus of references already presupposes the notion of *faṣāḥa*: only *al-kalām al-faṣīḥ*, rather than any *kalām* uttered by Bedouin, is taken as evidence of *kalām al-ʿArab*. As-Suyūṭī defines *samāʿ* (i.e. the corpus collected by the grammarians) as "what has been established (*ṯabata*) as being the *kalām* of those whose *faṣāḥa* is unquestionable" (*Iqtirāḥ*). The frequency of the usage and the value of the enunciation happen to be reconciled in a grammarian's definition (Suyūṭī, *Muzhir* I, 187): you can tell a word is *faṣīḥ* when it is more common among the *fuṣaḥāʾ* whose Arabic is reliable and it is used more frequently than other words (*ʾakṯar ʿalā ʾalsinat al-fuṣaḥāʾ*).

Actually, *faṣīḥ* refers to a sense of value that goes beyond the grammatical text, for the latter, in fact, is founded on it. Hence the *Kitāb* never defines the *faṣīḥ* but admits it as a presupposition shared with the person addressed. Presumably, at the origin of every grammar lies a set of privileged sentences with which a community identifies itself. These 'sentences' establish the language. Furthermore, they also set up the community with the highly symbolic part they play. In Arabic, these words are not only literary texts but also a sacred text. Hence, the dimension of the correctness of the language is not only aesthetic but also ethical and ontological.

In classical usage, *faṣāḥa* shares in the sacred nature of the language. With this notion, rhetoric and linguistics meet theology. The answer to the question asked over and over again in philological works, *man ʾafṣaḥ al-ʿArab?* 'who among the Arabs speaks Arabic most correctly?' is always the same: the Prophet (Suyūṭī, *Muzhir* I, 209). Qurʾānic usage, even if it is not the most common usage, is at the base of the *afṣaḥ*. Ḥarīrī expresses this explicitly (*Durra*, 129). Ibn Xālawayh (d. 370/980) asserts it is an *ʾijmāʿ*: a word that appears in the *Qurʾān* is necessarily more *faṣīḥ* (*ʾafṣaḥ*) than its synonym that does not appear in it (Suyūṭī, *Muzhir* I, 213), even if it does not comply with the rule (*Muzhir* I, 188). In the wake of the Prophet, the tribe of Qurayš is *ʾafṣaḥ al-ʿArab*. This distinction must be understood in relation to the dialects of ancient Arabia. The classical view is expressed by Ṯaʿlab in his *ʾAmālī*: Qurayš has been able to rise above disparaged dialectal features (Suyūṭī, *Muzhir* I, 211; Ibn Jinnī, *Xaṣāʾiṣ* II, 11; cf. Rabin 1951; Blachère 1952:70–84; Versteegh 1984:1–14). According to the grammatical tradition, the tribes considered for the collection of data are those of central and eastern Arabia, essentially Qays, Tamīm, and ʾAsad (→ pre-Islamic Arabic). Does this mean that eastern and western tribes spoke two different types of Arabic? Versteegh (1984:5) points out that since the Ḥijāzī dialect was profoundly influenced by the central and eastern dialects (Rabin 1951) and as there is no doubt that Ḥijāzī is the best-known dialect of all pre-Islamic dialects, this could mean that there was no discrepancy at all between eastern and western dialects.

Related to *faṣāḥa*, the question of Arabic's incommensurability with respect to other languages arises quickly. Arabic cannot be compared with other languages: it is infinitely superior, a perfect language (Ibn Fāris, *Ṣāḥibī* 40–41). Az-Zubaydī praises God for having made Arabic "the most eloquent and melodious language in its sounds, the most balanced in its order, the clearest in its expressions, the most varied in its modes of discourse" (*Ṭabaqāt* 11). This perfection, added to its untranslatability and immutability, is linked to the sacred status of the written language during the classical period.

In these texts, the antonym of *faṣīḥ* is *qabīḥ* or *rakīk*, and not *ʾaʿjam* as in pre-classical usage. The notion of intelligibility and clearness remains essential in the meaning of the word. In discussions about the need to reject the use of rare words, *faṣīḥ* is the antonym of *waḥšī*, *nādir* (Suyūṭī, *Muzhir* I, 233). The purity of the language – its lack of disharmonious sounds, rare words, and disparaged dialectal features – is the condition of its clearness.

As the grammatical epistemology draws its inspiration from the epistemology of legal science, *faṣīḥ* is modeled on *ṣaḥīḥ* in the *Ḥadīṯ*. There are different degrees in *faṣīḥ* (*faṣīḥ* and *ʾafṣaḥ*), which leads to a distinction between *faṣāḥa* and linguistic correctness, an expression being either correct or incorrect. Yet Ibn Fāris, in accordance with the definition of *faṣīḥ* as being what is correct, associates *ʾaṣaḥḥ wa-ʾafṣaḥ*, as if the *ʾafṣaḥ* were the most correct expression (*Ṣāḥibī* 73; Suyūṭī, *Muzhir* I, 261). Likewise, *qalīl* is associated with *radīʾ*, i.e., usage of rare words entails poor quality of an expression (*ʾaqallu-hā wa-ʾardaʾu-hā; Muzhir* I, 226).

The concern for the aesthetic prevailed from the 9th century onward and gave rise to a thorough questioning of the notion of *faṣāḥa* throughout the following four centuries. Several problems dominated the rhetorical treatises, especially in the debate about form (→ *lafḏ̣*) and meaning (→ *maʿnā*). Does *faṣāḥa* fall within the domain of *lafḏ̣* or *maʿnā*? One says *lafḏ̣ faṣīḥ* but not *maʿnā faṣīḥ*, and yet, there cannot be any *faṣāḥa* without *maʿnā*. Many other questions are related to this question: is *faṣāḥa* different from *bayān*, the clear enunciation that immediately discloses the meaning? If these two notions were synonymous, every clear enunciation, even if unsightly, would fall under *faṣīḥ*. Can *faṣāḥa* be tantamount to harmony of sounds and words? If this were the case, a foreigner who does not know Arabic should be able to identify the *faṣīḥ*. Finally, what is the difference between *faṣāḥa* and *balāġa* 'rhetoric'?

Linked to a concern for linguistic correctness and aesthetics, a search for the pure appears, a purism that becomes an essential connotation of the *faṣīḥ*, characterizing the love of the language in classical Arab culture. It becomes the agony of every speaker of written Arabic that the feeling for the language has been lost, when the system of pre-Islamic Arabic based on syntactic endings collapsed as early as the conquests. Ever since the 2nd century A.H. and the establishment of the *ʿarabiyya*, it was feared that scholars might admit fictitious expressions. As-Suyūṭī reports that in response to such accusations, al-ʾAṣmaʿī (d. 213/828) admitted nothing but the *luġa fuṣḥā*, the purest, the clearest dialectal variant.

The more we progress in time, the more this nostalgia for the pure deepens. The history of the meaning of the expression *luġa fuṣḥā* shows this ever-increasing nostalgia (Ayoub 2003:51). In the 8th century, through a necessary epistemological process, the grammarians brought out both the common usage on which the *ʿarabiyya* is based and its dialectal variants (*luġāt*). In this perspective, the expression *luġa fuṣḥā* is a description; it consists in indicating for a given expression the most appreciated variant (the clearest, the purest one) among all the dialectal variants available, all of which are perceived as correct. Quite soon, this most appreciated variant becomes the most common one (Ṯaʿlab), then the only one. From now on, → *luġa* also means language rather than a dialectal variant, since only one variant is accepted. *Luġa fuṣḥā* becomes a designation of written Arabic: a harmonious and pure idiom. Interestingly, in the popular literature, such as the *Arabian Nights*, *faṣāḥa* lost this connotation of correctness and purity. What remains is eloquence, the beauty of the expression associated with the beauty of faces.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

ʾAbū ṭ-Ṭayyib, *Marātib* = ʿAbd al-Wāḥid ibn ʿAlī ʾAbū ṭ-Ṭayyib al-Luġawī, *Marātib an-naḥwiyyīn*. Ed. Muḥammad ʾAbū l-Faḍl ʾIbrāhīm. Cairo: Dār Nahḍa Miṣr li-ṭ-Ṭibāʿa wa-n-Našr, 1974.

Ḥarīrī, *Durra* = ʾAbū Muḥammad al-Qāsim ibn ʿAlī al-Ḥarīrī, *Durrat al-ġawwāṣ fī ʾawhām al-xawāṣṣ*. Ed. H. Thorbecke. Leipzig, 1871. (Repr., Baghdad.)

Ibn al-ʾAnbārī, *Lumaʿ* = Kamāl ad-Dīn ʾAbū l-Barakāt ʿAbd ar-Raḥmān ibn Muḥammad al-ʾAnbārī, *Lumaʿ al-ʾadilla fī ʾuṣūl an-naḥw*. Ed. Attia Amer. Stockholm: University of Stockholm, 1963.

Ibn Fāris, *Sāḥibī* = ʾAbū l-Ḥusayn ʾAḥmad Ibn Fāris, *aṣ-Sāḥibī fī fiqh al-luġa wa-sunan al-ʿArab fī kalāmi-hā*. Ed. Mustafā aš-Šuʾaymī. Beirut: Muʾassasat Badrān li-ṭ-Ṭibāʿa wa-n-Našr.

Ibn Jinnī, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. Ed. Muḥammad ʿAlī an-Najjār. Cairo, 1952–1956. (Repr., Beirut: Dār al-Hudā, n.d.)

Ibn Manẓūr, *Lisān* = Jamāl ad-Dīn ʾAbū l-Faḍl Muḥammad ibn Mukarram Ibn Manẓūr, *Lisān al-ʿArab*. Cairo: Dār al-Maʿārif, n.d.

Ibn an-Nadīm, *Fihrist* = ʾAbū l-Faraj Muḥammad ibn ʾIsḥāq al-Warrāq Ibn an-Nadīm, *al-Fihrist*. Ed. ʾIbrāhīm Ramaḍān. Beirut: Dār al-Maʿrifa, 1997.

Ibn Qutayba, *ʿUyūn* = ʿAbdallāh ibn Muslim Ibn Qutayba al-Marwazī, *ʿUyūn al-ʾaxbār*. 4 vols. Ed. Cairo, 1934–1949.

Ibn Sīda, *Muxaṣṣaṣ* = ʾAbū l-Ḥasan ʿAlī ibn Ismāʿīl Ibn Sīda al-ʾAndalusī, *al-Muxaṣṣaṣ*. Ed. Muḥammad ʾAmīn a.o. Cairo, 1903. (Repr., Beirut: Dār al-Fikr, 1978.)

Qazwīnī, *ʾĪḏāḥ* = al-Xaṭīb al-Qazwīnī, *al-ʾĪḏāḥ fī ʿulūm al-balāġa*. Ed. Muḥammad Xafājī. Beirut: Dār l-Kitāb al-Lubnānī, 1975.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān ibn Qanbar Sībawayhi, *al-Kitāb*. Ed. Hartwig Dérenbourg. Paris: Imprimerie Nationale, 1881. (Repr., Hildesheim and New York: G. Olms, 1970.)

Sīrāfī, *ʾAxbār* = ʾAbū Saʿīd al-Ḥasan ibn ʿAbdallāh as-Sīrāfī, *ʾAxbār an-naḥwiyyīna l-baṣriyyīn*. Ed. Ṭāhā M. az-Zaynī and Muḥammad ʿA.-M. Xafājī. Cairo: Muṣṭafā al-Bābī al-Ḥalabī, 1955.

Suyūṭī, *Iqtirāḥ* = Jalāl ad-Dīn ʿAbd ar-Raḥmān ibn ʾAbī Bakr as-Suyūṭī, *Kitāb al-iqtirāḥ fī ʿilm ʾuṣūl an-naḥw*. Aleppo: Dār al-Maʿārif, n.d.

——, *Muzhir* = Jalāl ad-Dīn ʿAbd ar-Raḥmān ibn ʾAbī Bakr as-Suyūṭī, *al-Muzhir fī ʿulūm al-luġa wa-ʾanwāʿi-hā*. Ed. ʾAbū l-Faḍl ʾIbrāhīm a.o. Cairo: Dār ʾIḥyāʾ al-Kutub al-ʿArabiyya, n.d.

Ṯaʿlab, *Faṣīḥ* = ʾAbū l-ʿAbbās ʾAḥmad ibn Yaḥyā Ṯaʿlab, *Kitāb al-faṣīḥ*. Ed. J. Barth. Leipzig: Hinrich'sche Buchhandlung, 1876.

Zubaydī, *Ṭabaqāt* = ʾAbū Bakr Muḥammad ibn al-Ḥasan az-Zubaydī, *Ṭabaqāt an-naḥwiyyīn wa-l-luġawiyyīn*. Ed. Muḥammad ʾAbū l-Faḍl ʾIbrāhīm. Cairo: Dār al-Maʿārif, 1973.

SECONDARY SOURCES

Ayoub, Georgine. 2001. "Le tout de la langue ou le malheur de l'infini (une étude de *Durrat al-ġawwāṣ* de Ḥarīrī)". *Paroles, signes, mythes: Mélanges J.-E. Bencheikh*, ed. Floréal Sanaugustin, 67–141. Damascus: Institut Français d'Etudes Arabes de Damas.

——. 2003. "Un idiome harmonieux et pur: Le trésor enseveli". *Cent titres à l'usage des bibliothécaires, libraires et amateurs, poésie de langue arabe*, ed. Jean-Charles Depaule, 29–58. Marseille: Centre International de Poésie.

——. 2004. "*Wajh malīḥ wa-lisān faṣīḥ*: La langue des Nuits". *Les mille et une nuits en partage, Proceedings of the Carma's Colloquium 25–29 May 2004*, ed. Aboubakr Chrabi, 491–515. Paris: Actes Sud.

——. Forthcoming. "*Lisān ʿarabī, lisān ʾaʿjamī*: La question des emprunts dans la tradition médiévale". *Proceedings of the Colloquium "Les interférences linguistiques: Deuxièmes rencontres internationales» 2–3 June 2003*, ed. Michel Bozdémir. Istanbul.

Blachère, Régis. 1950. “Les savants iraqiens et leurs informateurs bédouins aux II^{e}–IV^{e} siècles de l’Hégire”. *Mélanges William Marçais,* 37–48. Paris: G.-P. Maisonneuve. (Repr., *Analecta.* Damascus: Institut Français de Damas, 1975.)

——. 1952. *Histoire de la littérature arabe des origines à la fin du XV^{e} s. de J.-C.* I. Paris: Adrien Maisonneuve. (Repr., 1980.)

Blau, Joshua. 1963. “The role of the Bedouins as arbiters in linguistic questions and the *masʾala az-zunbūriyya*”. *Journal of Semitic Studies* 8.42–51. (Repr., *Studies in Middle Arabic and its Judaeo-Arabic variety.* Jerusalem: Magnes Press, 1988.)

Fück, Johann. 1955. *ˁArabīya: Recherches sur l’histoire de la langue et du style arabe.* French transl. Claude Denizeau. Paris: M. Didier.

Haywood, John. 1965. *Arabic lexicography: Its history and its place in the general history of lexicography.* 2nd ed. Leiden: E.J. Brill.

Kopf, Lothar. 1956. “Religious influence on Medieval Arabic philology”. *Studia Islamica* 5.33–59.

Rabin, Chaim. 1951. *Ancient West Arabian.* London: Taylor’s Foreign Press.

——. 1960. “*ˁArabiyya*”. *Encyclopaedia of Islam* I, 579–585. 2nd ed. Leiden: E.J. Brill.

Renan, Ernst. 1863. *Histoire générale et système comparé des langues sémitiques.* Paris: Michel Lévy. (Repr., *Oeuvres complètes de Ernest Renan.* Paris: Calmann-Lévy, 1958.)

Versteegh, Kees. 1984. *Pidginization and creolization: The case of Arabic.* Amsterdam and Philadelphia: J. Benjamins.

——. 1990. “Grammar and exegesis: The origins of Kufan grammar and the *Tafsīr Muqātil*”. *Der Islam* 67.206–242.

GEORGINE AYOUB (Institut des Langues et Civilisations Orientales)

Feminine → Gender

Fiˁl

1. DEFINITION: MORPHOLOGY AND TENSE

Fiˁl (etymologically ‘fact, operation’) is generally translated by ‘verb’. It is not, as one might have assumed, the noun depicting the process (*maṣdar*) of the verb *faˁala* ‘to do’, since the verbs constructed in the form CaCaCa derive their *maṣdar* in the form CaCC (al-Xalīl, *al-ˁAyn*, s.v. *f-ˁ-l*), apparently with the exception of three verbs whose *maṣdar* takes the form CiCC (Ibn Manẓūr, *Lisān*, root *f-ˁ-l*).

Sībawayhi’s *Kitāb*, dating from the 8th century, was the first book in the Arabic grammatical tradition. It begins by dividing speech into three parts: “noun (*ism*), verb (*fiˁl*), and particle (*ḥarf*), which is used to contribute to a meaning and which is neither a noun nor a verb” (*Kitāb* I, 12). In this first chapter, the verb is defined as a sum of “paradigms (*ʾamṯila*) issued from nouns depicting the process and formed to indicate what has been, what will be but has not [yet] happened, and what is but has not been completed” (*Kitāb* I, 12).

Sībawayhi’s definition is interesting in many respects. It states at the outset that the peculiarity of the verb is to be a sum of paradigms, that the verb is derived from the *maṣdar*, which is a subclass of the noun, and that it is constructed in different forms to express time which has or has not elapsed (*Kitāb* I, 35) or, in as-Suhaylī’s interpretation, events which have occurred and those which have not occurred, i.e. the change in the state of an event (*Natāʾij al-fikr* 388–389).

This definition, which is essentially based on morphological criteria (Hamzé 1994:93–115), changed under later grammarians, even those who claimed to follow Sībawayhi’s teaching (Hamzé 2002:577–579). The fundamental morphological aspect of the *Kitāb* disappeared, and the semantic values indicated by the verb, the process, and time came to the fore: “[C]onventionally, according to the grammarians, the verb is what indicates a process and past or future time. . . . This is what Sībawayhi meant by ‘as for the verb, it is a sum of paradigms issued from nouns depicting process and formed to indicate what has been, what will be but has not [yet] happened, and what is but has not been completed’” (Zajjājī, *ʾĪḍāḥ* 52).

The standard definition in the grammatical tradition refers to these semantic values. It has its roots in Ibn as-Sarrāj’s *Kitāb al-ʾuṣūl*, dating from the beginning of the 10th century: “The verb is what indicates meaning and time, the past, present or future tense” (*ʾUṣūl* I, 38). One cannot help noticing how this definition resembles the one given by al-Fārābī, Ibn as-Sarrāj’s contemporary, his teacher in logic and his disciple in grammar (Ibn ʾAbī ʾUṣaybiˁa, *Ṭabaqāt al-ʾaṭibbāʾ* 560), in his commentary on Aristotle’s *Hermeneutics* (Fārābī, *ˁIbāra* I, 133). It has the advantage of justifying the division of speech into three parts and of giving coherent definitions of these parts: if the word does not have a meaning in itself, then it is a particle; and if it does have a meaning, then this meaning is either related to time, in which case it is a verb, or not related to time, in which case it is a noun (ʾAstarābāḏī, *Šarḥ al-Kāfiya* I, 2–7).

2. CONJUGATION AND TENSE

The conjugation of Arabic verbs shows a remarkable economy. It places a form with suffix, *faʿal-ta* 'you did', in opposition to a form with prefix, *ta-fʿal-u* 'you do'. This opposition, which from the 19th century onward has been regarded in Arabic and Semitic linguistics as an aspectual division between accomplished and unaccomplished verbs, or as a mixed aspectual/temporal opposition (→ aspect; Versteegh 1997: 84), is analyzed by the Arabic grammatical tradition as a temporal division.

After the 10th century, Arab grammarians, probably under the ever-greater influence of Aristotelian logic, tried, not without difficulty, to change the binary division of Arabic verbs into a ternary division of time (Fleisch 1965; Versteegh 1997:84). Ibn Yaʿīš takes this correspondence a long way: one, and only one, verbal form must correspond to each part of extralinguistic time: "Since time is divided into three parts, past, present and future – for time represents the movements of the stars: a movement which has taken place, a movement which is yet to take place, and a movement separating these two – verbs are divided likewise into past, future, and present" (*Šarḥ al-Mufaṣṣal* VII, 4). In his commentary on Ibn Jinnī's *Lumaʿ*, which divides verbs into three categories according to the division of time, Ibn al-Xabbāz goes even further in this logic by saying that the division of verbs into three categories is necessary because the tense of the action must either coincide with the moment of enunciation or not. If it coincides, then it is the present; if it does not, then it is either posterior or anterior. If it is posterior, then it is the future; if it is anterior, then it is the past. This limitation to three categories is necessary because, according to these criteria, the division can only be made based on the answer 'yes' or 'no' (*Tawjīh al-Lumaʿ* 100).

The temporal criterion present in Sībawayhi's definition, whereby the paradigms serve to express time, does not aim at a correspondence between the forms *faʿala-ø* and *y-a-fʿal-u*, which distinguish the conjugation of Arabic verbs, and a ternary division of time (past, present, and future). In his division of the verb, Sībawayhi does not use a clearly temporal terminology for the present, "what is, but has not been accomplished", nor for the future, "what will be, but has not yet happened" (Guillaume 1988:29). To express both present and future, he uses the form *y-a-fʿal-u*, without resorting to the use of the specific modalities of the future, *sa* and *sawfa*. The third verbal form Sībawayhi mentions is that of the imperative *(i)fʿal*, which is associated with the form *y-a-fʿal-u* to indicate "what will be, but has not yet happened".

Another important fact presented in the *Kitāb* often goes unnoticed in studies on the Arabic tradition. Sībawayhi says that the verb "is constituted to indicate which time is past or not past". This wording, which is repeated in the *Kitāb* (I, 34, 35, 36) and which suggests a division between past vs. not past, corresponding to two verbal forms *faʿala-ø* vs. *y-a-fʿal-u*, does not correspond exactly to a division between past, present, and future. It is worth noting that the Arabic tradition after Sībawayhi has never sought to find a simple term for the expression *mā lam yamḍi* 'what has not passed', unlike the expression *mā maḍā* 'what passed', which led to the simple term *al-māḍī*. The usual division *al-māḍī* vs. *al-muḍāriʿ* used to designate the opposition between the two forms of Arabic verbs is clearly heterogeneous (→ *māḍī/muḍāriʿ*. Again according to Sībawayhi, the second term, *al-muḍāriʿ*, is situated not on the temporal level but on the level of the governance theory. The aim of the term *muḍāriʿ* "similar [to the noun of agent]" (*Kitāb* I, 13) is to justify the change in the final vowels of the form of the verb with prefix: *y-a-fʿal-u*. From a temporal point of view, Arab grammarians use in opposition to the past several terms indicating either the present (*ḥāl, ḥāḍir, ʾān*, etc.) or the future (*mustaqbil, ʾātī*, etc.), but never both, which is a clear indication of the direction post-Sībawayhi grammarians took.

Dissymmetry between the two forms of Arabic verbs and the three tenses, past, present, and future, has been the subject of much debate in the grammatical tradition. Some grammarians denied the existence of a present tense for Arabic verbs, while others denied even the existence of an extralinguistic present time (Ibn ʿUṣfūr, *Šarḥ al-Jumal* I, 127–128). The philosophers considered the present to be merely a point separating the past from the future. Regarding the separation between two tenses as a third tense would be tantamount to regarding every binary division as a ternary one (ʾAstarābāḏī, *Šarḥ al-Kāfiya* II, 226).

The linguistic present tense of the Arab grammarians cannot therefore correspond to that point separating real time, which is the moment of enunciation. For az-Zajjājī, the present does not have the depth necessary to be expressed in its own right (Suyūṭī, *Hamᶜ* I, 17-18). When Arab grammarians speak of the present tense of a verb, they are referring to that which is situated on either side of the moment of enunciation, which separates the past from the future (*mā ᶜalā janbatay al-ʾān*). This is what permits us to say that a verb like *yuṣallī* 'he prays' is in the present, whereas a part of the prayer is in the past and the other part is in the future (Zajjājī, *ʾĪḍāḥ* 87; ʾAstarābāḏī, *Šarḥ al-Kāfiya* II, 226).

3. MOOD AND GOVERNANCE

Unlike the verb in the past tense, the verb in non-past tense is *muᶜrab* 'declined', i.e., its final vowel changes under the influence of governors, whether apparent or supposed. It has three forms, *y-a-fᶜal-u* vs. *y-a-fᶜal-a* vs. *y-a-fᶜal-ø*, which, in Western school-grammar, are often called *indicative*, *subjunctive,* and *jussive* or *apocopated moods*. Arab grammarians did not deal with the change in the last consonant of the verb in terms of mood: they established a connection with the last consonant of the noun. The same terminology is used for case and modal vowels: as with the noun, the verb is *marfūᶜ* 'with a vowel /u/', and *manṣūb* 'with a vowel /a/'. According to Blachère and Gaudefroy-Demombynes (1975:37), this connection was due not only to similarity in form between the final consonants of these two parts of speech but also to a functional resemblance. The *marfūᶜ* is used in a 'main or isolated clause', like the noun in the nominative case, and the *manṣūb* is employed in a 'subordinate clause'. In fact, case or modal changes are analyzed within the same theory, that of governance (→ *ᶜamal*), according to which the change in the final vowel, when it is not accidental (i.e. when it is not due to phonetic or morphological conditions) is due to a governor. The use of the same terminology for nouns and verbs cannot be attributed to a mere formal likeness of their final consonants. In fact, going by the governance theory, Arab grammarians use two different terms for the same vowel inside the same part of speech: *rafᶜ* vs. *ḍamm* for the vowel /u/, *naṣb* vs. *fatḥ* for the vowel /a/, etc. (Sībawayhi, *Kitāb* I, 13), depending on whether this vowel is declensional. The connection established between the noun and the verb is based on the notions of similarity (*muḍāraᶜa*) and symmetry, or on noun-verb opposition (*munāḏara*), which is the basis for the tripartite division of speech (Sībawayhi, *Kitāb* I, 13–23). A comparison is made between the final vowel /u/ of the verb and the noun, as in *yaktub-u zayd-un* 'Zayd writes Zayd' and *al-kātib-u zayd-un* 'the writer [is] Zayd', and the vowel /a/ as in *yurīdu ʾan yaktub-a* 'he wants that he write = he wants to write' and *yurīdu l-kitābat-a* 'he wants the writing'.

The theory of governance perceives formal resemblance between the verb and the noun as a consequence of the functional resemblance. The total absence of resemblance to the noun – as in the case of the imperative – results in the form farthest removed from the noun, which is characterized by the total absence of a final vowel /ø/ (*sukūn*), an impossibility for nouns. The same applies to the apocopate verb (*majzūm*). Once again, the absence of any functional resemblance explains the absence, in verbs, of a final vowel /i/ specific to nouns: the genitive in the case of the noun and the apocopate in the case of the verb are in opposition (*naḏīr*; Sībawayhi, *Kitāb* I, 14; Zajjājī, *ʾĪḍāḥ* 102–120). A reduced functional resemblance gives a reduced resemblance of form: the verb in the past tense, *faᶜala*, has a final vowel like the noun, but this vowel is invariable.

In the theory of governance, which is the backbone of syntactic analysis for the Arab grammarians, the verb is considered to be the most powerful governor because it necessarily governs a noun, its *fāᶜil*, and can govern one or more complements: object, adverbial phrase (place, time, manner), etc. Given this power, the verb governs, whether preposed or postposed, and whether it is next to or separated from that which it governs. The verb's capacity to govern depends on its variation, that is, its capacity to be conjugated and to belong to a network of derivation, i.e. to have a noun of process (*maṣdar*), an active participle, a passive participle, etc.

4. Simple forms and augmented forms

Arabic verbs can take two forms: a simple, basic form called *mujarrad* 'naked', usually of three consonants, CvCvCv, but sometimes of four, CvCCvCv, and an augmented form called *mazīd*, formed by adding one or more elements to the root consonants of the simple form. The final vowel of the three-consonant verb CvCvCv is a syntagmatic vowel. It is dropped when the verb is conjugated in the 1st or 2nd person, *faʿal-tu* 'I did' or *faʿal-ta* 'you did', to avoid a succession of four short syllables, CvCvCVCV, since the verb is regarded as a single word with its personal pronoun → *fāʿil* (Zajjājī, *ʾĪḍāḥ* 75). The first vowel allows an opposition to be established between the active form (*faʿala*, *faʿila*, *faʿula*) and the passive form (*fuʿila*). Early grammarians called the active form *mabnī li-l-fāʿil* 'constructed for the *fāʿil*' – the *fāʿil* being the protagonist of the verb – and later grammarians called it *maʿlūm* '[verb whose *fāʿil* is] known'. The early grammarians called the passive form *mabnī li-l-mafʿūl* 'constructed for the *mafʿūl*', since the object becomes the protagonist of the verb, or *mā lam yusamma fāʿilu-hu* 'that whose *fāʿil* is not designated', and later grammarians called it *majhūl* '[verb whose *fāʿil* is] unknown'. Arab grammarians emphasized the use of the form *faʿula* for qualities and *faʿila* for illnesses, for suffering, and for colors. The form *faʿala* is used for various meanings since it is the lightest form (Ibn Yaʿīš, *Šarḥ al-Mufaṣṣal* VII, 156–157), the vowel /a/ being lighter than the other two, /u/ and /i/. ʾAstarābāḏī says that "when the form is light, it becomes more frequent and likely to vary" (*Šarḥ aš-Šāfiya* I, 70).

In addition to the simple form of the verb, Arabic has developed a considerable number of augmented forms: twenty-five forms constructed on three-consonant verbs and two forms on four-consonant verbs (Zamaxšarī, *Mufaṣṣal* 369, 375). The simple form has a lexical value, and its domain is the dictionary. However, the value added in the augmented form is often predictable or belongs to a set of predictable values; it is used to add nuance to the lexical meaning provided by the simple form. The traditional grammatical literature (e.g. Sībawayhi, *Kitāb* IV, 55–80; ʾAstarābāḏī, *Šarḥ aš-Šāfiya* I, 70–113) lists the main values of the most frequent augmented forms: *fāʿala* for reciprocity, e.g. *ḍāraba* 'he hit [someone] and he was hit [by that someone]'; *ifʿalla* for taking color, e.g. *ibyaḍḍa* 'he turned white'; *istafʿala* for asking for something, e.g. *istaxbara* 'he asked to be informed'; and so on. As with the median vowel in simple three-consonant forms, augmented forms permit remarkable economy within the system (Fleisch 1965).

5. Types of verbs

From a morphophonological point of view, Arabic verbs were divided into *ṣaḥīḥ* 'healthy' and *muʿtall* 'weak', based on whether they contained elements likely to be transformed. Other subdivisions were made within each of these two categories in order to explain conjugation and the principles of morphophonological variation of the verb (Ibn al-Muʾaddib, *Daqāʾiq at-taṣrīf* 147–360; Liblī, *Buġyat al-ʾāmāl*).

However, a syntactic-semantic criterion is often followed in classifying the chapters on verbs. The domination by syntax is obvious in the twelve chapters on the verb in az-Zamaxšarī's *Mufaṣṣal* (219–275), probably the first gram-mar to be organized according to the three parts of speech. Ever since Sībawayhi's *Kitāb*, the classification had been according to the theory of governance: verbs which do not govern a *mafʿūl*, i.e. intransitive verbs, and verbs governing one, two, or three objects, with subdivisions relating to the possibility of elision of the *mafʿūl* and to the types of relations between them (*Kitāb* I, 33–54).

The grammatical tradition devotes a separate chapter to verbs called *ʾafʿāl al-qulūb* (lit. 'verbs of the heart') because they are related to intimate thoughts and their meaning lies in the heart, such as *ḍanna* 'to believe', *raʾā* 'to see'. Already in the *Kitāb*, this category, which requires two *mafʿūl*s, is quite separate from other verbs that take two *mafʿūl*s, such as *ʾaʿṭā* 'to give', *kasā* 'to dress', because it requires two accusative nouns, and one object does not suffice (*wa-laysa la-ka ʾan taqtaṣira ʿalā ʾaḥad al-mafʿūlayni dūna l-ʾāxar; Kitāb* I, 37).

The chapter on the 'verbs of the heart' is also distinct from that on verbs known as *nāqiṣ* 'incomplete', such as *kāna* 'to be', *ṣāra* 'to become', not only because of the difference in case vowels – the verb governs two accusative cases in the first category, a nominative and an accusative case in the second – but also because of the different functions in the two types of sentence.

Other chapters consider verbs from other aspects, such as the imminence of the process (*ʾafˁāl al-muqāraba*) or its beginning (*ʾafˁāl aš-šurūˁ*). In either case, not only semantic considerations lead to the adoption of such categories but also their effect on the sentence structure.

The same is true for chapters devoted to verbs known as *jāmid* 'fixed', such as verbs of → exclamation (*ʾafˁāl at-taˁajjub*) or verbs of praise and blame (*ʾafˁāl al-madḥ wa-ḏ-ḏamm*), which are typically found in the section on syntax in grammar books because of the effect their fixed character has on their behavior, which was the main preoccupation of Arab grammarians.

In addition to semantic and morphological differences, each of these categories of the verb has specific syntactic properties which justify a separate chapter being devoted to them (Zamaxšarī, *Mufaṣṣal* 345–368). These properties are generally linked to the governance theory or to the position of different elements within the sentence. This position is fixed in verbs of exclamation such as *mā ʾakrama zayd-an* or *ʾakrim bi-zayd-in* 'how generous he is, Zayd!'. In sentences of praise and blame, such as *niˁma r-rajul-u zayd-un* 'what an excellent man he is, Zayd!', the discussion concentrates on the fixed form of verb and, fundamentally, the specific structure of the sentence in order to justify it (Bazzi-Hamzé 2004:272–292).

6. TRANSITIVITY AND THE PASSIVE

The Arabic technical term → *taˁaddī* (lit. 'exceeding [a limit]') is not exactly identical to the term 'transitivity', which comes from Latin *transire* 'to pass'. The verb, which necessarily governs its *fāˁil*, goes beyond that in order to govern other elements of speech. Levin translates *al-fāˁil allaḏī yataˁaddā-hu fiˁlu-hu ʾilā mafˁūl* (*Kitāb* I, 34) by "the subject, the grammatical effect of whose verb passes over to a direct object" (Levin 1998:194). Moreover, *taˁaddī* does not apply to direct objects exclusively but also to any other complement, such as object, state, adverbial phrase, etc. In this regard, *taˁaddī* is equally valid for transitive and intransitive verbs. However, the technical term *fiˁl mutaˁaddī* applies most particularly to transitive verbs, which can have one, two, or three *mafˁūl*s, i.e. complements likely to become pronouns in the accusative case.

The term 'passive' is objected to by Fleisch (1957:151–170), who, in the absence of a better alternative, adopts the term used by the later Arab tradition, *majhūl* '[verb whose *fāˁil* is] unknown'. The passive is generally considered to be a secondary form constructed from an active, transitive verb "by placing a /u/ vowel on its first consonant and an /i/ vowel on its second, and by removing its *fāˁil* and putting the direct object (*mafˁūl*) in its place" (Zajjājī, *Jumal* 77). In this regard, it would be 'absurd' to form a passive verb from an intransitive one for want of a direct object likely to become the *fāˁil* of the verb (Ibn as-Sarrāj, *ʾUṣūl* I, 77). However, the morphological rules referred to permit the construction of the passive, thanks to internal inflection, whether the verb is transitive or intransitive. This construction is clearly shown in action or reaction verbs, which Fleisch (1965:918) calls "verbs with agent". It is accepted by grammarians such as Sībawayhi, who forms passive verbs from intransitive verbs like *qaˁada* 'to sit' and *ḍaḥika* 'to laugh'. Arab grammarians would later justify this construction, admitted by Sībawayhi, by resorting to an elliptic element, the noun of process (*maṣdar*), implied by the verb and referring to something known and usual (Zajjājī, *Jumal* 77; ʾAstarābāḏī, *Šarḥ al-Kâfiya* I, 85).

Yet, indirect transitive verbs and intransitive verbs, when in the passive, always have an expansion, which may be either the *maṣdar* of a passive verb, as in Q. 69/13: *fa-ʾiḏā nufixa fī ṣ-ṣūr-i nafxat-un wāḥidat-un* 'and when the trumpet is blown once', or an adverbial phrase, as in the example *sīra yawm-u l-jumˁat-i* 'it was walked the day of Friday', or a preposition followed by a noun as in this verse by al-Farazdaq: *yuġḍī ḥayāʾan wa-yuġḍā min mahābat-i-hi* 'he looks away out of diffidence and one looks away out of respect for him'.

Because the verb cannot be without a → *fāˁil*, Arab grammarians regard the expansion as a replacement for the passive verb's *fāˁil* (Ibn Jinnī, *Lumaˁ* 14). For Ibn Hišām (*ʾAwḍaḥ al-masālik* I, 371–377), the *fāˁil* of this type of verb is the implied pronoun referring to the noun of process (*maṣdar*) implied by the verb. Others consider that the preposition followed by a noun functions as a *fāʿil* because this syntagm is considered to have the status of direct object of a transitive verb. Indeed, Arab grammarians insist on three methods to transform an intransitive verb into a transitive one: (a) doubling the second root consonant: *fariḥa* 'to be cheerful' > *farraḥa* 'to make cheerful'; (b) adding an initial hamza: *ḏahaba* 'to go away' > *ʾaḏhaba* 'to make go away'; and (c) adding a preposition: *xaraja* 'to go out' > *xaraja* + *bi-* 'to take out' (Zamaxšarī, *Mufaṣṣal* 341). According to this interpretation, *xurija bi-zayd-in* becomes equivalent to *ʾuxrija zayd-un* 'Zayd was taken out' (Jurjānī, *Muqtaṣid* I, 347).

It is conceivable that on the logical-semantic level the preposition followed by a noun is mandatory. In fact, the absence of a reference element would produce a sentence with a semantic content too vague to satisfy the interlocutor. Since it is impossible to identify the morpheme of the agent (*fāʿil*), the verb being impersonal, information and identification are given in an expansion, which then appears to be an indispensable element. It is around this known element that information is articulated and the message is conveyed to the interlocutor (Hamzé 1993:53).

Bibliographical references

Primary sources

ʾAstarābāḏī, *Šarḥ al-Kāfiya* = Raḍī d-Dīn Muḥammad ibn al-Ḥasan al-ʾAstarābāḏī, *Šarḥ Kitāb al-Kāfiya fī n-naḥw*. Beirut: Dār al-Kutub al-ʿIlmiyya, n.d.
——, *Šarḥ aš-Šāfiya* = Raḍī d-Dīn Muḥammad ibn al-Ḥasan al-ʾAstarābāḏī, *Šarḥ Šāfiyat Ibn al-Ḥājib*. Ed. Muḥammad Nūr al-Ḥasan, Muḥammad az-Zafzāf, and Muḥammad Muḥyī d-Dīn ʿAbd al-Ḥamīd. Beirut: Dār al-Kutub al-ʿIlmiyya, 1975.
Fārābī, *ʿIbāra* = ʾAbū Naṣr Muḥammad ibn Muḥammad al-Fārābī, *al-Manṭiq ʿinda l-Fārābī*. Ed. Rafīq al-ʿAjam. Beirut: Dār al-Mašriq, 1985.
Ibn ʾAbī ʾUṣaybiʿa, *Ṭabaqāt al-ʾaṭibbāʾ* = Muwaffaq ad-Dīn ʾAbū l-ʿAbbās ʾAḥmad ibn al-Qāsim as-Saʿdī al-Xazrajī Ibn ʾAbī ʾUṣaybiʿa, *ʿUyūn al-ʾanbāʾ fī ṭabaqāt al-ʾaṭibbāʾ*. Ed. Muḥammad Bāsil ʿUyūn as-Sūd. Beirut: Dār al-Kutub al-ʿIlmiyya, 1998.
Ibn Hišām, *ʾAwḍaḥ al-masālik* = Jamāl ad-Dīn ʾAbū Muḥammad ʿAbdallāh ibn Yūsuf Ibn Hišām, *ʾAwḍaḥ al-masālik ʾilā ʾAlfiyyat Ibn Mālik*. Ed. Muḥammad Muḥyī d-Dīn ʿAbd al-Ḥamīd. Beirut: Dār ʾIḥyāʾ at-Turāṯ, 1966.
Ibn Jinnī, *Lumaʿ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Lumaʿ fī l-ʿarabiyya*. Ed. Hadi Kechrida. Uppsala: Almqvist and Wiksell, 1976.
Ibn Manẓūr, *Lisān* = Jamāl ad-Dīn ʾAbū l-Faḍl Muḥammad ibn Mukarram Ibn Manẓūr al-ʾIfrīqī al-Miṣrī, *Lisān al-ʿArab*. Beirut: Dār Ṣādir, n.d.
Ibn al-Muʾaddib, *Daqāʾiq at-taṣrīf* = al-Qāsim ibn Muḥammad ibn Saʿīd al-Muʾaddib, *Daqāʾiq at-taṣrīf*. Ed. ʾAḥmad Nājī al-Qaysī, Ḥātim Ṣāliḥ aḍ-Ḍāmin, and Ḥusayn Tūrāl. Baghdad: Arab Academy of Iraq, 1987.
Ibn as-Sarrāj, *ʾUṣūl* = ʾAbū Bakr Muḥammad ibn Sahl Ibn as-Sarrāj, *al-ʾUṣūl fī n-naḥw*. Ed. ʿAbd al-Ḥusayn al-Fatlī. Beirut: Dār ar-Risāla, 1988.
Ibn ʿUṣfūr, *Šarḥ al-Jumal* = ʾAbū l-Ḥasan ʿAlī ibn Muʾmin Ibn ʿUṣfūr al-Ḥaḍramī al-ʾIšbīlī, *Šarḥ Jumal az-Zajjājī, aš-Šarḥ al-kabīr*. Ed. Ṣāḥib ʾAbū Janāḥ. Baghdad: Wizārat al-ʾAwqāf wa-š-Šuʾūn ad-Dīniyya, 1980.
Ibn al-Xabbāz, *Tawjīh al-Lumaʿ* = aš-Šayx Šams ad-Dīn ʾAḥmad ibn al-Ḥusayn Ibn al-Xabbāz, *Tawjīh al-Lumaʿ*. Ed. Fāyiz Zakī Muḥammad Diyāb. Cairo: Dār as-Salām, 2002.
Ibn Yaʿīš, *Šarḥ al-Mufaṣṣal* = Muwaffaq ad-Dīn Yaʿīš ibn ʿAlī Ibn Yaʿīš, *Šarḥ al-Mufaṣṣal*. Beirut: ʿĀlam al-Kutub, n.d.
Jurjānī, *Muqtaṣid* = ʿAbd al-Qāhir al-Jurjānī, *Kitāb al-Muqtaṣid fī šarḥ al-ʾĪḍāḥ*. Ed. Kāẓim Baḥr al-Marjān. Baghdad: Dār ar-Rašīd, 1982.
Liblī, *Buġyat al-ʾāmāl* = ʾAḥmad ibn Yūsuf al-Liblī al-ʾAndalusī, *Buġyat al-ʾāmāl fī maʿrifat mustaqbal al-ʾafʿāl*. Ed. Jaʿfar Mājid. Tunis: ad-Dār at-Tūnisiyya, 1972.
Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān ibn Qanbar, *al-Kitāb*. Ed. ʿAbd as-Salām Hārūn. Cairo: al-Hayʾa al-Miṣriyya al-ʿĀmma li-l-Kitāb, 1977–1979.
Suhaylī, *Natāʾij al-fikr* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān ibn ʿAbdallāh as-Suhaylī, *Natāʾij al-fikr fī n-naḥw*. Riyadh: Dār ar-Riyāḍ, 1984.
Suyūṭī, *Hamʿ* = Jalāl ad-Dīn as-Suyūṭī, *Hamʿ al-hawāmiʿ fī šarḥ Jamʿ al-jawāmiʿ*. Ed. ʿAbd al-ʿĀl Sālim Makram. Beirut: Muʾassasat ar-Risāla, 1992.
Xalīl, al-*ʿAyn* = ʾAbū ʿAbd ar-Raḥmān al-Xalīl ibn ʾAḥmad al-Farāhīdī, *Kitāb al-ʿAyn*. Ed. Mahdī al-Maxzūmī and ʾIbrāhīm as-Sāmarrāʾī. Beirut: Muʾassasat al-ʾAʿlamī, 1988.
Zajjājī, *ʾĪḍāḥ* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān ibn ʾIsḥāq az-Zajjājī, *al-ʾĪḍāḥ fī ʿilal an-naḥw*. Ed. Māzin al-Mubārak. 3rd ed. Beirut: Dār an-Nafāʾis, 1979.
——, *Jumal* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān ibn ʾIsḥāq az-Zajjājī, *al-Jumal fī n-naḥw*. Ed. ʿAlī Tawfīq al-Ḥamad. Beirut and Irbid: Muʾassasat ar-Risāla and Dār al-ʾAmal, 1984.
Zamaxšarī, *Mufaṣṣal* = Jār Allāh ʾAbū l-Qāsim Maḥmūd ibn ʿUmar az-Zamaxšarī, *al-Mufaṣṣal fī ṣanʿat al-ʾiʿrāb*. Ed. ʿAlī Bū Mulḥim. Beirut: Dār wa-Maktabat al-Hilāl, 1993.

Secondary sources

Bazzi-Hamzé, Salam. 2004. "Louange et blâme en arabe: Structure sémantique et structure syntaxique des énoncés avec *niʿma* et *biʾsa*". *Le voyage et la langue*, ed. Joseph Dichy and Hassan Hamzé, 272–292. Damascus: Institut Français d'Etudes Arabes de Damas.
Blachère, Régis and Maurice Gaudefroy-Demombynes. 1975. *Grammaire de l'arabe classique*. Paris: Maisonneuve et Larose.

Fleisch, Henri. 1957. "Etudes sur le verbe arabe". *Mélanges Louis Massignon*, 151–170. Damascus: Institut Français d'Etudes Arabes de Damas.

——. 1965. "*Fiʿl*". *Encyclopedia of Islam* II, 916–919. 2nd ed. Leiden: E.J. Brill.

Guillaume, Jean-Patrick. 1988. "Le discours tout entier est nom, verbe et particule". *Les parties du discours*, ed. Bernard Colombat. (= *Langages* 92.25–36.) Paris: Larousse.

Hamzé, Hassan. 1993. "La traduction de la voix objective entre le français et l'arabe". *Aspects du vocabulaire*, ed. Pierre Arnaud and Philippe Thoiron, 49–61. Lyon: PUL, Université de Lyon, Centre de Recherche en Terminologie et Traduction.

——. 1994. "Les parties du discours dans la tradition grammaticale arabe". *Les classes des mots: Traditions et perspectives*, ed. Louis Basset and Marcel Perennec, 93–115. Lyon: PUL, Université de Lyon.

——. 2002. "Fī l-qirāʾa al-manṭiqiyya li-nuṣūṣ an-naḥw". *Annales de l'Université de Tunis* 46.561–581.

Levin, Aryeh. 1998. "The meaning of *taʿaddā al-fiʿl ʾilā* in Sībawayhi's *al-Kitāb*". *Arabic linguistic thought and dialectology*, 193–210. Jerusalem: Hebrew University.

Versteegh, Kees. 1997. *The Arabic Language.* Edinburgh: Edinburgh University Press.

Hassan Salam Bazzi Hamzé
(University of Lyon-2)

First Language Acquisition

Sībawayhi's *al-Kitāb* and Ibn Jinnī's *Sirr ṣināʿat al-ʾiʿrāb* (Bakalla 1994) are prominent examples of distinguished scholarship that demonstrate a long investigative tradition related to the study of the Arabic language. In contrast, the acquisition of Arabic as a native language by children received no particular attention in the past – apart from some passing references, such as in al-Jāḥiẓ's encyclopedic book *al-Bayān wa-t-tabyīn* – and continues to be a relatively neglected area of study. This marginalization may be explained, in part, by the reluctance of traditional Classical Arabic scholars to consider the spoken vernaculars as worthy of true scholarship. And since studying acquisition means studying the dialects, as all Arabic-speaking children are first exposed to a dialectal variety of Arabic acquired as their mother tongue, the field of language development has not represented a central preoccupation in the study of Arabic. Although such an attitude is slowly changing, the continuing lack of attention given to this field is compounded by the unavailability of accepted standards against which → child language can be measured. Dialects generally have no codified grammar and are subject to many dialectal variations. As a result it is difficult to evaluate child speech and decide what is correct and what is not. This hesitation to consider the dialects as serious linguistic varieties worthy of scholarship by Arabs may be one of the reasons why grammars of the Arabic dialects have mainly been written by non-Arab Arabists, e.g. Cowell (1964) for Syrian, Erwin (1963) for Iraqi, Holes (1990) for Gulf, Mitchell (1956) for Egyptian, and Harrell (1962) for Moroccan Arabic.

The first study of the acquisition of Arabic was that of Egyptian Arabic, published by Margaret Omar in 1973. To this day it remains the only book that provides an investigation of the acquisition of all components of Arabic (phonology, morphology, and syntax). However, this state of affairs is progressively changing as many Arab graduates devote their doctoral research to the acquisition of their native Arabic dialects. Such new interest stems from recognition in the field of psycholinguistics that the validity of an acquisition theory necessitates that it be based on evidence from languages typologically different from English and other European languages. To test various hypotheses about the universality of processes and principles underlying language acquisition, psycholinguists have stressed the need to obtain data from structurally different languages to allow them to generalize their conclusions. This is a move away from the paradigm that dominated language development studies in the 1960s and 1970s based on the early Chomskyan theory of the language faculty as an innate ability that is universal and that is not profoundly affected by the type of input received by the child.

As the crosslinguistic approach to language acquisition took hold in the late 1970s, interest in languages other than English, and particularly languages from different typological families, grew. While still maintaining that there are universal principles (Universal Grammar, UG) which guide the acquisition of widely different languages, researchers progressively acknowledged that typological characteristics of the language being acquired must also play a role in how and what children acquire first and what they acquire later. This realization that neither universality (of grammar) nor developmental stages can be based on English and other Indo-European languages alone encouraged investigation of other languages. The crosslinguistic project started by Slobin (1985) led to research in languages from different language families. The first study of Arabic in this tradition was undertaken to show that although there are universal conceptual prerequisites that underlie the acquisition of the Arabic lexicon, in that the same concepts are expressed at similar developmental stages regardless of the language being acquired, the formal properties of the Arabic lexicon influence the strategies adopted by children in using the specific formal properties in word formation (Badry 1983) and sensitize them more to the more productive linguistic processes in their language.

In the last two decades there has been an increase in doctoral studies devoted to the acquisition of different dialects, including Kuwaiti, Moroccan, Egyptian, and Saudi dialects. These investigations, however, have focused on specific aspects in the dialect being acquired rather than providing a comprehensive study of all of its characteristics. Moreover, they remain unpublished and thus difficult to access for those who want to build on them. Information on Arabic acquisition is also available in some published articles about different areas of linguistic development. All the available resources are therefore rather recent and can be framed within modern theoretical frameworks dominant in linguistics and child language development studies. There is also a growing interest in studying bilingual development of children of migrant Arab communities in Europe and Palestinians in Israel (→ child bilingualism). Several studies have been carried out on Arabic-speaking children in the Netherlands (Altena and Appel 1982) and Sweden (Håkansson a.o. 2003), for example (→ Europe). The focus of these studies, however, tends to be limited either to children with → language impairment (Salameh a.o. 2004) or to literacy development among bilinguals (Abu-Rabia 1995). The following sections provide an overview of information available on acquisition of Arabic as a native language by children without language impairment.

1. ACQUISITION OF PHONOLOGY

Research in phonological acquisition addresses several fundamental issues related to speech perception and production. Investigators are interested in finding answers to questions such as the following: What is the phonetic inventory of children at different stages of their development? What are the stages of phonological development? What phonological processes are applied by children in their acquisition of the phonological system of their language? What is universal and what is particular in phonological acquisition? What is the relationship between the prelinguistic (babbling) and linguistic stages? And what factors (physiological, perceptual, and environmental) affect the order of acquisition?

Considering the paucity of research in Arabic acquisition, phonological development is probably one of the areas that have received adequate attention from researchers. Amayreh and others have collected and analyzed data from children acquiring Jordanian Arabic between the ages of 14 months and eight years. Their reports address many of the above questions in different published articles (Amayreh and Dyson 1998, 2000; Dyson and Amayreh 2000; Amayreh 2003). Their findings reveal that children acquiring Jordanian Arabic follow stages in the development of their phonetic inventory that are similar to those of children acquiring a variety of other languages including English, the most studied language in this area. At the same time, the sounds specific to Arabic, such as emphatics, which are physiologically more complex because they involve a secondary articulation, are acquired much later, and their acquisition is not completed before the age of eight. The authors also found that some sounds considered to be late in acquisition are found earlier in phonetic inventories of Jordanian Arabic children. Before discussing the sounds acquired at each stage, it is important to keep in mind that the complexities brought about by the multiglossic Arabic situation and the resulting important free variation observed in adult speech make it very difficult and even sometimes arbitrary to define what is meant by acquisition of a particular sound.

Data available from the prelinguistic stage come from two infants between the ages of 6

Table 1. Developmental stages in the acquisition of Arabic consonants (based on Omar 1973 and Amayreh and Dyson 1998)

	babbling	14–24ms	2–3;10yrs	4–6;4yrs	6;5–8yrs
stops	b, ʔ	b, d, t, ʔ	k, q, g		t̴, d̴
fricatives/affricates	h	ʃ, ʕ, ħ, h	f	s, χ, ð, ɣ, θ, ʤ, s̴	ð̴, z
sonorants/liquids	m	m, n, l		r	
glides	w, y	w, y			
totals		13	+3	+8	+4 = 28

and 10 months, which was reported in Omar's study of the acquisition of Egyptian Arabic. The phonetic inventory of these Arabic-speaking children shows the production of mid and back high and low vowels [ə, ɛ, e, ʊ, u, æ, a] with a conspicuous absence of the high front vowel [i]. Their consonantal inventory included stops [p, b, d, ʔ], fricatives [ɣ, h, v, z], nasals [n, m], and glides [w, y]. Findings from Jordanian Arabic-speaking children between the ages of 14 and 24 months show that labial and dental-alveolar stops [b, p, d, t], nasals [m, n], glides [w, y], fricatives [ħ, ʃ], liquids [l], and the glottal stop [ʔ] are produced in both initial and final syllable position (Amayreh and Dyson 2000). The production of [l], [ħ], and [ʔ] at this early stage distinguishes Jordanian Arabic-speaking children from children of similar ages acquiring English, where these sounds usually appear at a later age. The earlier appearance of these consonants is explained by their relatively high frequency in the input and their high functional load in the language, as they are part of function words commonly used in adult speech. The ease of articulation of [ʃ] is also called upon as a possible explanation for its early appearance in the Jordanian Arabic data. On the other hand, the conspicuous absence of [f], [k], [g], and [s] from data reported by Amayreh and Dyson (2000) is odd because these sounds (except for [s]) are reported to appear early in most languages. Moreover, casual observation of Arabic-speaking children from other dialects indicates their presence in their phonetic productions by age two. Children also produce sounds that are not part of Arabic inventories, such as [p], [ts], [ŋ], [ɫ], [pf], [β], [θ], which lends support to the universality hypothesis proposed by Jakobson (1968) that all children start off with the same phonetic inventories but later hone in to those sounds present in the language they are exposed to and drop off those that are not used in their native language. Findings from Arabic are also in line with generative phonologists' proposals (Vihman a.o. 1986; Mowrer and Burger 1991) that phonological acquisition proceeds through five levels, where, in each successive stage, the additional consonants acquired are characterized by more complex consonant feature contrasts. Table 1 shows the addition of consonants by Jordanian Arabic-speaking children to their inventory at each stage.

Stops are the first consonants to be acquired, while the mastery of fricatives spans several years to be completed beyond age six and a half. The last consonants to be acquired are the emphatic ones. Their late acquisition has generally been explained by their articulatory complexity. Around age two, children produce the six phonemic vowels of Arabic, [a, i, u] and their long counterparts, along with their six allophonic variations. The presence of two emphatic allophonic vowels, [ɑː] and [ɑ], is particularly interesting in light of the absence at this stage of the emphatic consonants from children's inventories. Omar (1973) reports the presence of five diphthongs, /æj/, /aj/, /ij/, /uj/, and /aw/; two initial clusters, /st/ and /ht/; and medial clusters /bl/, /xt/, /bt/, /ft/, /xn/, /sf/, and /dl/, where /l/ is a substitute for /r/ in input language. Comparison of the inventories reported for Jordanian and Egyptian dialects reveals some important discrepancies between the ages of acquisition of the phonemes listed in Table 2.

Table 2. Differences between ages of acquisition as reported by Omar (1973) and Amayreh and Dyson (2000)

Phonemes	Omar	Amayreh and Dyson
/s/, /z/	2;8	4–6;4
/š/	4;6	14–24ms
/ġ/	2;6	4;0–6;4
/ṣ/	3;6	4;0–6;4
/ẓ/, /ṭ/, /ḍ/, /ḷ/	3;6	6;5–8;0
/ž/ (j)	4;0	4;0–6;4
/ʿ/	4;6	2;0
/ṯ/, /ḏ/	not reported in inventory	4;0–6;4
/ḏ̣/	not reported in inventory	6;5–8;0
/q/	6;6	2;0–3;10

These differences could be explained in terms of elicitation methods. They may also be due to the functional load of the particular sounds in the dialect being studied as well as the socioeconomic background of the children investigated. For example, the phoneme /q/ is hardly used in dialectal Egyptian, and children are not likely to be exposed to it before they reach school age. On the other hand, although /q/ is not frequent in the Jordanian dialect, the children tested by Amayreh and Dyson are middle-class urban children who are probably exposed to story reading from their parents and thus are likely to hear this sound, unlike rural children from an isolated village in Upper Egypt.

2. COMPREHENSION

Egyptian children's acquisition of vocabulary items seems to be in line with children acquiring other languages. At the two-word stage (around age two), most words produced are concrete nouns referring to familiar people and objects, but some imperative forms and pronouns do occur in speech samples. In comprehension, children respond correctly to naming and directives. Omar (1973) tested comprehension of locatives, color terms, adjectival modifiers, gender and number agreements, word order, tense contrasts, voice (passive/active), and negation in children between the ages of 2;8 and seven. She identified the following order of comprehension. The youngest children tested were able to understand the affirmative/negative contrasts. The passive/active contrasts are also acquired before age three. Gender marking of nouns is also acquired early. Prepositions representing 'in', 'on', 'beside', and 'under' are also acquired before age three.

Such findings are in line with the developmental sequences reported from other languages. Omar also found that adjective/noun agreements are differentiated at about age 4;6, while gender marking of the verb is understood at around age five and color terms are mastered at around age six. Numbers beyond the number 'two' were difficult and do not seem to be acquired before age six, when children enter school. Omar stressed the influence of environmental factors as playing an important role in the rate of acquisition. Most of the children she studied may have had health and/or affective problems. In addition, their exposure to rich linguistic input was limited, as their parents did not give them much attention.

Similarly, in a later study Al-Akeel (1998) tested the comprehension of possessives, prepositions, and complex commands by children acquiring the Saudi dialect of Arabic and concluded that possessives are understood before age three while the comprehension of prepositions spans from three to six, which is comparable to orders reported for English except for the prepositions 'under' and 'between', comprehended earlier in English (3 and 3;6 respectively). Al-Akeel reported the following order for the Saudi Arabic-speaking children tested: *fi* 'in' and *ʿala* 'on' before age three; *taḫt* 'under', *žanb* 'beside', *wara* 'behind', and *gudām* 'in front of', between the ages of four and five; and finally 'between' after age six. He also reported that reversible passives where both agent and patient are animate were comprehended beginning at age 4;6, but agentless passives were understood earlier, at age three.

3. ACQUISITION OF SYNTAX

Around age one, children begin to use single words to communicate their intentions. As their memory and processing capacity develop, their sentences become longer and more complex. There are several stages in syntactic development beginning with the one-word stage, followed by the two-word stage around the 18th month and the multi-word stage beginning around age three. In early syntactic development, one-word utterances stand for whole intentions and function as full sentences. Psycholinguists have argued that child speech at each stage is rule-governed and that the progression to adult grammar is systematic (Slobin 1973, 1985). Observed children's reductions are due to processing constraints, such as limited memory span. Children conceptualize the whole proposition to be communicated, but because of processing constraints they generally choose the word or words that express the most important and new information in the context and leave out background information shared by the participants.

In her study of syntactic development of Egyptian Arabic-speaking children, Omar (1973) describes similar developmental stages. When children in the early stages of syntactic development are asked to imitate adult utterances, their imitations are reductions of adult utterances and tend to preserve adult word order. The words selected are content words that usually carry a heavy semantic load in the communication. Thus, these early utterances do not contain any function words such as articles and prepositions. However, given the more synthetic structure of Arabic, in which pronouns and other function morphemes are realized as affixes, children's one-word utterances may correspond to two-word utterances in more analytical languages, such as English. Omar also points out that because Arabic equational sentences do not have a copula verb (which is omitted in English-speaking children's two-word utterances), the utterances produced in child Arabic at the two-word stage correspond to the adult grammar. These and other typological differences make it difficult to carry out crosslinguistic comparisons.

As children enter the multi-word stage, their utterances become longer. In addition to content words produced in previous stages, prepositions, articles, cardinal numerals, demonstratives, other modifiers, and negative particles are used.

Arabic sentences can be negated with the anaphoric negator *laʾ* placed as a free morpheme before the sentence. In addition, dialects use the negative double particle *m-…-š* combined with the allophonic variants [i, u, a] as intervocalic elements to negate sentences. The two particles are attached to each other [m-i/u/a-ši] when used before a predicate phrase in nominal sentences, or used as a discontinuous morpheme, prefixed and suffixed to a verbal head *ma*-V-*š* (Mohamed and Ouhalla 1995). Most dialects share these characteristics with some variations.

Mohamed and Ouhalla (1995) discuss the acquisition of negation by Palestinian children aged 1;10 to 2;7. They found that "while nominal sentences are invariably 'correctly' negated with the pattern *m-š* XP, verbal sentences are freely negated either with the 'incorrect pattern *m-š* XP or the correct pattern *m*-V-*š*'" (1995:88). On the basis of this analysis, they argue that this is evidence that children at this stage have not yet developed the obligatory verb movement responsible for a productive derivation of the form *m*-V-*š* in adult Palestinian Arabic. They found that the data collected from the children investigated supports their hypothesis that there are two patterns of negation which are determined by specific syntactic contexts, one that carries a meaning load and one that is functional. During the early stages of negative acquisition, children express only negation that carries meaning, while acquisition of negation falling under functional categories is still missing.

In her investigation of the acquisition of negation by Egyptian Arabic-speaking children, Omar identifies three stages. She notes that *laʾ* 'no' is the first particle to be understood and produced appropriately but that it is also inappropriately overgeneralized to express all types of negation. At this stage, *laʾ* is attached to the sentence without modifying it. The second stage is when children add the negative particle *miš* and use it by tagging it to their utterance. In the final stage in the acquisition of negation, children use the discontinuous negative particle *ma-…-š* appropriately affixed to the verb.

The same developmental stages are observed with interrogatives. First, children ask quesions by rising intonation, then they start tagging question words to either the beginning or ending of utterances, and finally, they use interrogative words with prepositions and adult stylistic placement of question words in various positions in the sentence.

This sequence of development is similar to that observed in the acquisition of negation and interrogation in English and other languages and supports the operating principle proposed by Slobin (1973), which states that one of the strategies used by children in acquisition is initially to avoid interruption of utterances.

4. Acquisition of morphology

Children begin to use inflections productively in their speech around age 2;6½ when their utterances contain more than two words. Even before this age, however, there are no words in their speech that are bare stems. Children tend to omit inflections that are prefixed more than those that are suffixed, supporting the operating principle that ends of words are more perceptually salient than beginnings and therefore tend to be perceived earlier. Children produce nouns in the singular form and inflect them for gender and possession.

4.1 *Noun and adjective inflections*

The acquisition of the Arabic plural is another area that has gained attention in studies of acquisition given the complexity of its morphology. Evidence from Arabic plurals acquisition is used to support or refute theories in morphology relating to the debate between proponents and opponents of the dual route or the connectionist theories in morphology. Arabic-speaking children, like children acquiring other languages, first express their newly developed concept of plurality by using modifiers such as *kulluhum* 'all' before the singular form, or a numeral followed by a singular noun *tlata ktab* 'three book', or by repeating the singular noun by using a coordinate noun phrase, e.g. *di gutta w di gutta* 'this is a cat and this is a cat' (Omar 1973).

As early as 1;8, children start producing plural nouns with all plural markings. In both longitudinal and experimental studies, Ravid and Farah (1999, 2001) found that, in speech samples from children acquiring Palestinian Arabic as their first language, at around age two the broken plural category is the most frequently produced, followed by sound feminine plurals. Sound masculine plurals are the last to appear in speech samples, occurring at around age 2;6, and are rare (2% of all plural forms). Duals and collective nouns are also late and rare acquisitions. By age three, the sound plural marker seems to reach its peak productivity.

In the process of acquiring their pluralization rules, children make two types of overgeneralization errors. They first overgeneralize the sound feminine plural marker *-āt* to broken plural forms, and later, around 2;5, they overgeneralize broken plural patterns to adult sound feminine plural nouns. Their productions reveal that the broken plural patterns and the sound feminine suffix are the most productive pluralization processes. Other pluralization processes, the masculine sound plural, the dual, and the collective forms, are used later and without errors, suggesting that they are not yet productive in the child's system. However, errors with broken plurals continue well beyond age seven.

The overgeneralized use of the sound feminine form is explained by the fact that it is the least constrained semantically and formally. The overgeneralization of the feminine marker *-āt* seems to persist as a default pluralizer even in older children. It is also preferred for pluralizing nonce words. The late and limited use of the sound masculine form among young children, despite its regularity and transparency, is attributed to the fact that this inflection is restricted to nouns referring to human agents or patients (active and passive participles). On the other hand, despite its complexity, the early use of broken plural forms is the result of rote-learned forms which only later become analyzed and categorized as subpatterns in a network (Ravid and Farah 2001). Omar (1973) hypothesizes that these irregular plurals are probably acquired item by item, given their idiosyncrasy and the large number of irregular patterns. Errors of using wrong irregular patterns for irregular nouns, however, suggest that children organize irregular plurals in subclasses of schemas (Ravid and Farah 2001), similar to those proposed to explain the acquisition of irregular English past tense verbs (Bybee and Slobin 1982). A complete mastery of irregular plurals goes well beyond school years.

4.2 *Noun phrase agreement marking*

The plural system of Arabic is characterized by another complexity in number agreement between cardinal numbers and the nouns they modify. Nouns used with numbers are in the plural after the numbers 3–10 but in the singular after the numbers 11 and above. Children as old as 15 years still make errors by overgeneralizing the use of plural nouns after all cardinal numbers above 2. The same developmental trends in pluralization of nouns were also observed in noun-adjective agreements. Children overgeneralize the affix *-īn* to pluralize their adjectives. Unlike number marking, gender is correctly marked early on. Omar found that children were correctly using gender marking by age 2;8, the age at which she tested them. She also reports no errors in noun-adjective agreements except for color adjectives in any of the age groups studied.

4.3 *Verb inflections*

In her study of four children (aged 2 to 2;6) acquiring verb inflections in Kuwaiti Arabic as their first language, Aljenaie (2001) found that children produced verbs with perfect markings earlier than the imperfect markings. The 1st person singular inflections were overgeneralized to other contexts. They were followed by the 3rd person singular marking, which appeared before the 2nd person and plural affixes in the children studied. Masculine marking of the verb appeared before the feminine affixes. Aljenaie notes that while some verb stems appeared to be unmarked, unmarking was limited to those verb stems that were homophonous with imperative forms in the adult language. None of the children used a bare stem that does not correspond to an acceptable verb form used elsewhere in the adult paradigm. Children's performance was interpreted as support for Universal Grammar (UG) in that children do not construct 'wild grammars' and that children know the constraint that prevents non-adult bare stem forms like *šrab* 'drink' from surfacing.

Kuwaiti children also produced suffixed markers (gender and number) earlier than prefixed (tense/aspect) ones. This order is in line with acquisitional data from other languages and has been explained by invoking perceptual saliency of ends of words as well as processing-span constraints (Slobin 1973). Aljenaie (2001) argues that children "retain the suffix which marks the gender of the 2nd person and the number of the 3rd person to avoid confusion with other grammatical markers". Such an argument is in line with the role of functional load of linguistic units as a factor in their earlier acquisition.

4.4 *Derivational morphology*

Badry's (1982, 1983, 2004) investigation of the acquisition of lexical derivational processes was based on data from children 3;5 to 9;9 years old who were acquiring Moroccan Arabic as their first language. It shows that children develop both horizontal and vertical derivational strategies that allow them to form words from their roots and from other surface forms. Using both spontaneous and elicited speech samples, as well as nonce words, she found that children go through four main stages in acquisition of verbal and nominal pattern derivations. In investigating the acquisition of causative, reciprocal, and middle voice patterns, Badry found that the causative pattern was the first to be used productively by children, followed by the reciprocal and then the middle voice patterns. She also notes that children at all ages studied preferred to use the basic pattern *faʿala* (Form I or PI) "to express several semantic and syntactic relations in spite of the availability, in their repertoire, of more specialized verbal patterns" (Badry 2004:140). The progression observed is explained by conceptual development and formal structure as well as pragmatic factors. The early expression of causation is attributed to its earlier conceptual development, while the later production of reciprocals is due to the complex semantic relations of simultaneity which they express (Berman 1985; Bowerman 1982; Slobin 1985). However, the later appearance of the middle voice pattern, appearing earlier in data from other languages, was attributed to the experimental testing environment, where children were presented with pictures in which agents were present and the children therefore had no need to disclaim personal responsibility. In the production of all patterns, children's errors were with verbs derived from weak (one of the three consonants is a glide) or irregular roots (one of the three consonants is a geminate). Children tended to supply a third consonant or metathesize consonantal roots. Such errors were interpreted as evidence for the psychological reality of the → root in the process of word formation (→ derivation). Children use both horizontal and vertical derivational processes in forming new words. However, when surface forms are not transparent, the knowledge that three consonantal roots underlie most content words in Arabic is called upon to reconstruct the word skeleton. The case of erroneous surface forms is due to the lack of specific knowledge about defective or weak roots.

5. Conclusion

The investigation of the acquisition of Arabic is a promising field of study on both practical and theoretical grounds. It promises to shed light on the various stages of linguistic development and uncover processes developed by children that could provide useful insights into the learning of Standard Arabic. On the theoretical level, the acquisition of Arabic structure can enrich the debate in several subfields of linguistics.

Bibliographical references

Abu-Rabia, Salim. 1995. "Learning to read in Arabic: Reading, syntactic, orthographic and working memory skills in normally achieving and poor Arabic readers". *Reading Psychology* 16.351–394.

Al-Akeel, Abdulrahman I. 1998. *The acquisition of Arabic language comprehension by Saudi children.* Ph.D. diss., University of Newcastle upon Tyne.

Aljenaie, Khawla. 2001. *The emergence of tense and agreement in Kuwaiti children speaking Arabic.* Ph.D. diss., University of Reading.

Altena, Nelleke and René Appel. 1982. "Mother tongue teaching and the acquisition of Dutch by Turkish and Moroccan immigrant workers' children". *Journal of Multilingual and Multicultural Development* 3:4.315–322.

Amayreh, Mousa. 2003. "Completion of the consonant inventory of Arabic". *Journal of Speech, Language, and Hearing Research* 46.517–529.

—— and Alice Dyson. 1998. "The acquisition of Arabic consonants". *Journal of Speech, Language, and Hearing Research* 41.642–653.

——. 2000. "Phonetic inventories of young Arabic-speaking children". *Clinical Linguistics and Phonetics* 14:3.193–215.

Badry, Fatima. 1982. "The centrality of the root in Semitic lexical derivation". *Papers and reports on child language development*, 9–15. Stanford, Calif.: Stanford University.

——. 1983. *Acquisition of lexical derivational rules in Moroccan Arabic: Implications for the development of Standard Arabic as a second language through literacy.* Ph.D. diss., University of California, Berkeley.

——. 2004. *Acquiring the Arabic lexicon: Evidence of productive strategies and pedagogical implications.* Bethesda, Md.: Academica Press.

Bakalla, Muhammad H. 1994. "Arab and Persian phonetics". *Encyclopedia of language and linguistics*, ed. Ron Asher, I, 187–191. Oxford: Pergamon.

Berman, Ruth. 1985. "The acquisition of Hebrew". *The crosslinguistic study of language acquisition*, ed. Dan I. Slobin, 255–372. Mahwah, N.J.: L. Erlbaum.

Bowerman, Melissa. 1982 "Evaluating competing models of language acquisition data: Implications of developmental errors with causative verbs". *Quaderni di Semantica* 3.5–66.

Bybee, Joan L. and Dan I. Slobin. 1982. "Rules and schemas in the development and use of the English past tense". *Language* 58.265–289.

Cowell, Mark W. 1964. *A reference grammar of Syrian Arabic (based on the dialect of Damascus).* Washington, D.C.: Georgetown University Press.

Dyson, Alice and Mousa Amayreh, 2000. "Phonological errors and sound changes in Arabic-speaking children". *Clinical Linguistics and Phonetics* 14:2.79–109.

Erwin, Wallace M. 1963. *A short reference grammar of Iraqi Arabic.* Washington, D.C.: Georgetown University Press.

Håkansson, Gisela, Eva-Kristina Salameh, and Ulrika Nettelbladt. 2003. "Measuring language proficiency in bilingual children: Swedish-Arabic bilingual children with and without language impairment". *Linguistics* 41:2.255–288.

Harrell, Richard. 1962. *A short reference grammar of Moroccan Arabic.* Washington, D.C.: Georgetown University Press.

Holes, Clive. 1990. *Gulf Arabic.* London and New York: Routledge.

——. 1994. "Arabic". *Encyclopedia of language and linguistics*, ed. Ron Asher, I, 191–194. Oxford: Pergamon.

Jakobson, Roman. 1968. *Child language, aphasias and phonological universals.* Transl. A. Keiler. The Hague: Mouton.

Mitchell, T.F. 1956. *An introduction to Egyptian Colloquial Arabic.* London: Oxford University Press.

Mohamed, Ibrahim and Jamal Ouhalla. 1995. "Negation and modality in early child Arabic". *Perspectives on Arabic linguistics*, VII, ed. Mushira Eid, 69–90. Amsterdam and Philadelphia: J. Benjamins.

Mowrer, D.E. and Burger, S. 1991. "A comparative analysis of phonological acquisition of consonants in the speech of 2½–6 year old Xhosa- and English-speaking children". *Clinical Linguistics and Phonetics* 5.139–164.

Omar, Margaret 1973. *The acquisition of Egyptian Arabic as a native language.* The Hague: Mouton.

Ravid, Dorit and Lubna Hayek. 2003. "Learning about different ways of expressing number in the development of Palestinian Arabic". *First Language* 23.41–63.

—— and Rola Farah. 1999. "Learning about noun plurals in early Palestinian Arabic". *First Language* 19.187–206.

——. 2001. "The early plural lexicon of Palestinian Arabic: A longitudinal case study". *Early Lexicon Acquisition 2001 Proceedings*. Lyon: Institut des Sciences de l' Homme, Université Lumière.

Salameh, Eva-Kristina, Ulrika Nettelbladt, and Kjell Norlin. 2003. "Assessing phonologies in bilingual Swedish-Arabic children with and without language impairment". *Child Language Teaching and Therapy* 19:3.338–365.

——, Gisela Håkansson, and Ulrika Nettelbladt. 2004. "Developmental perspectives on bilingual Swedish-Arabic children with and without language impairment: A longitudinal study". *International Journal of Language and Communication Disorders* 39:1.65–90.

Slobin, Dan I. 1973. "Cognitive prerequisites for the development of language". *Studies of child language development*, ed. Charles A. Ferguson and Dan I. Slobin, 175–208. New York: Holt.

—— (ed.). 1985. *The crosslinguistic study of language acquisition*. 5 vols. Mahwah, N.J.: Erlbaum.

Vihman, Marilyn M., Charles A. Ferguson, and M. Elbert. 1986. "Phonological development from babbling to speech: Common tendencies and individual differences". *Applied Psycholinguistics* 7.3–40.

Fatima Badry Zalami (American University of Sharjah)

First Language Teaching

This entry gives a brief overview of the present situation of Arabic language teaching and learning at primary, preparatory, and secondary schools in the Arab world. Although the examples concern only a few Arab countries, many issues hold true for the rest of the Arab countries. The following issues are discussed: (1) the teaching and learning of Arabic in the Arab world; (2) the structure of the school system; (3) the place of the Arabic language in the school curriculum; (4) the goals and content of the Arabic language curriculum at each educational level; (5) methodology and professional training of teachers of Arabic; and (6) conclusion and critical observations. Issues of Arabic language assessment and design of Arabic language materials are left aside.

1. The teaching and learning of Arabic in the Arab world

In the 19th and early 20th centuries, Arab scholars attempted to modernize the Arabic language in an effort to cope with modern civilization (for more discussion see Stetkevych 1970; M.M. Badawi 1993). Since then, there have not been many attempts to modernize the teaching and learning process, i.e. its pedagogy. A few studies discuss education in the Arab world in general (e.g. Matthews and Akrawi 1949; Massialas and Jarrar 1983, 1991), and the place of the Arabic language in the curriculum in particular (e.g. Altoma 1957; 1970; Mujāwir 1974; Amara and Mar'i 2002), but there are no comprehensive studies about teaching and learning Arabic as a first language. Many issues need to be reconsidered: needs, goals, what is to be expected from the students in terms of proficiency, the contents of what should be taught, methods of teaching, material design, ways of assessment, and professional training of teachers of Arabic. While steps in this direction have been taken for the teaching of Arabic as a foreign language, teaching Arabic as a first language still requires much attention.

Since the rise of Islam, the aim of those learning Arabic has often been to gain access to the knowledge that was introduced by the Islamic religion. The language of the *Qur'ān* and Classical Arabic literature represents a linguistic standard for which learners of Arabic continue to strive. This linguistic standard has been documented by vast amounts of linguistic descriptions by the classical Arab grammarians (E. Badawi 2002: 157–158). The rise of the non-religious schools and the establishment of media in the Arab world (written and spoken), when combined with attempts made by Arab countries to eradicate illiteracy and to modernize Arabic in order to cope with modern civilization, have generated a new language standard for Arabic. The first standard currently represents an 'ideological' standard, and the new standard represents an 'organic' one. While the ideological standard embodies the aim of Arab native speakers, the organic standard represents the actual practice of Arabic in its oral and written expression.

The present language situation in the Arab world has resulted in three varieties of Arabic. Two of them represent the literary language; the third variety represents the native regional Arabic dialect. For further discussion, see Ferguson (1959) and E. Badawi (2002). There are not many descriptive and educational studies that document the organic language standard in its current form in the same manner as is the case for the ideological one. Because of the influence of religious beliefs and nationalistic motives, proficiency in the ideological standard is still reflected as an educational goal for native Arab speakers in primary, preparatory, and secondary schools in the Arab world.

2. The structure of the school system

Exploring the three basic educational stages reveals that most of the Arab countries follow a similar educational pattern, which is 6–3–3 in terms of the total number of years that the student spends across the three educational stages: the primary, preparatory, and the secondary. Due to the influence of either French or the British occupation, some Arab countries such as Tunisia, Palestine, Lebanon, Sudan, Kuwait, Somalia, and Djibouti follow different patterns (Massialas and Jarrar 1983:41–42). Egypt briefly experimented with a 5–3–3 pattern in 2001 but returned to the common 6–3–3 pattern in 2005. For further discussion of the school ladder in the Arab world, see UNESCO Institute for Statistics (2002); ALECSO (1986).

3. The place of the Arabic language in the school curriculum

An examination of the school curriculum across several Arab countries reveals that Arabic language instruction takes a considerable amount of time across the three school levels per week. The importance attached to the teaching of Arabic as a first language is shown by the total number of weekly periods allocated to Arabic in the school schedule. In Table 1 the total number of periods per week is compared across the three school levels with the total number of periods per week for five subjects. The time of the class period ranges from 30 to 45 minutes at the primary stage, and from 40 to 55 at both the preparatory and secondary stages (ALECSO 1981:36). As can be noted, the time allocated for teaching Arabic ranges from one-third of the time across the curriculum, as in countries such as Kuwait, Iraq, Yemen, and Saudi Arabia, to less than one-third of the time in countries such as Libya and Bahrain.

As seen in Table 1, most countries devote a greater number of periods for Arabic at the primary level than at the other two levels. In addition, Table 1 illustrates that more time is generally spent on Arabic at both the entry and the exit levels than at the middle level. By exception, Egypt and Jordan (see Table 1) foster a gradual system of education, where the numbers of hours decrease from the primary stage to the preparatory to the secondary stage (Gezi 1979:9–10).

4. Arabic language curriculum goals and content at each educational level

In the Arab world, ministries of education are responsible for determining the goals of the teaching and learning of Arabic, the content of the instruction, and how and by whom it is taught in elementary, primary, and secondary schools. The system is structured linearly in the sense that each stage leads to the next and at the conclusion university students are produced. The curriculum is structured around the various subjects taught in the Arabic language lessons.

The goals for the teaching and learning of Arabic vary from one country to another in terms of clarity and specifications. In some cases, they focus more on national and cultural particularities than on the language skills themselves. Few studies currently exist that specifically mention the aims of the teaching and learning of Arabic (see, e.g., Khater 1963; Altoma 1957, 1970).

National and cultural goals have been expressed as, for example, assisting in the development of the learner's ability intellectually, physically, socially, emotionally, and rationally, or instilling in the youth the values of their Arab society and cultural heritage in order for them to become good citizens (Suleiman 1999:106). Other noteworthy aims of teaching Arabic include an effort to reinforce belief in God and to impart religious values.

As for language goals, according to Altoma (1957:84), the ultimate aim of teaching Arabic in Iraq is "the enabling of the pupils to read, write and converse in a simple correct language". Amara and Mar'i (2002:67, quoting Al-Haj 1996) mention that the goal for teaching Arabic to Palestinians in Israel during the 1950s and 1960s was to enable students to have a "correct reading and comprehension for the

Table 1. Arabic language instruction in nine Arab countries: Number of periods per week

Country	School levels			Total periods per week for Arabic	Total periods per week for five subjects*
	Primary	Preparatory	Secondary		
Saudi Arabia	57	18	38	113	290
Iraq	54	36	21	111	196
Kuwait	49	32	28	109	264
Yemen AR	54	18	32	104	223
Jordan	48	21	15	84	201
Qatar	53	7	19	79	223
Tunisia	30	–	42	72	148
Bahrain	11	7	35	52	153
Libya	10	7	8	25	151

* Islamic studies, geography, science, math, and foreign language
Source: ALECSO (1981)

written and spoken language; clear, exact and logical formulation of ideas and feelings, orally and in writing, the ability to understand and evaluate good literature; cultural and literary consciousness of the past and the present".

A comparison of the goals of teaching Arabic in Iraq and in Israel reveals that while both of them emphasize correct reading, writing, and speaking, they vary in their focus. In the case of Iraq, the first goal is to aim at correctness, but this gives no indication of any real-world objectives or functions in terms either of what the learners will be able to do as a result of the language instruction, or any cultural content. Moreover, it focuses more on the form and accuracy of the message being taught than on its communicative meaning. As for the Palestinians in Israel, while the learning goals are more elaborate in terms of what is expressed and comprehended, they fail to provide guidance as to the degree of proficiency a learner will attain at the end of each stage. Moreover, there is no reference to the use of Arabic functionally. Interestingly, neither goal refers to any practical learning needs. Students who drop out at the end of any of the three stages and attempt to use their Arabic functionally in their career find themselves unqualified to do so.

4.1 *Teaching Arabic in primary school*

According to Khater (1963:2), the goal of teaching Arabic, in Egypt as well as in most of the Arab countries, is to develop in learners the abilities and basic skills that enable them to participate effectively in reading and writing activities as carried out by most of the literate population in Arab societies. What is expected from the students in terms of reading is to be able to comprehend a few lines in an Arabic newspaper and to be able to read aloud in order to facilitate correction. In writing, students are expected to be able to write a short letter in which they express their ideas clearly. In speaking, they are expected to articulate their thoughts with "a reasonable degree of clarity and correctness" (Khater 1963:2). In other countries, such as Iraq, the goal of teaching Arabic is correct reading, writing, and conversation (Altoma 1957:85).

Looking at the curricula of Arabic in several Arab countries at the primary school level, we see that language education is divided into subjects, each of which contains exercises for the students to practice. The distribution of these subjects varies from one country to another. For example, in 1981 the Egyptian Arabic curriculum had the following subject divisions: reading and writing, composition and stories, memorization and prose, dictation, handwriting, and grammar (see Table 2). Iraq's 1971 curriculum had almost exactly the same division, but it also emphasized conversation (see Table 3).

Table 2. Primary school curriculum for Arabic (Egypt 1971): Number of periods per week

Branches/Grades	1	2	3	4	5	6
Reading and writing	6	6	3	3	3	3
Composition and storytelling	2	2	2	2	1	1
Memorization and chants	2	2	2	2	1	1
Dictation	–	–	2	1	1	1
Handwriting	–	–	1	1	1	1
Grammar	–	–	–	1	2	2

Source: Khater a.o. (1981:58)

Note that in Egypt composition is introduced in the first grade, while in Iraq it is introduced in the third grade. In addition, the curriculum in Egypt lacks conversation in *fuṣḥā*, while this is readily available in Iraq's curriculum (see Table 3).

Table 3. Primary school curriculum for Arabic (Iraq 1970): Number of periods per week

Branches/Grades	1	2	3	4	5	6
Reading and (writing + dictation)	8	8	4	4	1	1
Memorization	3	2	2	1	1	1
Composition and dictation	–	–	2	4	2	2
Grammar	–	–	–	1	2	2
Conversation	–	2	2	–	–	–

Source: Altoma (1970:701)

Reading and writing are introduced at the beginning of the primary stage in both Iraq and Egypt, which is the case in most Arab countries. Other similarities include grammar, which is introduced as a separate subject in the fourth grade, and dictation, which is introduced in the third grade in both Egypt and Iraq. However, there is no clear justification as to why composition is introduced in one country in the first grade while it is postponed to a later point in the curriculum in another.

Literature at the primary level is introduced in the form of poetry and prose texts for students to memorize. Arab countries vary in their introduction of poetry and prose throughout the six grades that compose the primary stage. Some countries prefer to introduce poetry from the first grade and postpone prose to later years, for instance to the fifth grade in the case of Sudan, while other countries like Saudi Arabia introduce neither poetry nor prose in the first two years (see Table 4).

The literary texts expose students to a variety of topics dealing with the values of religion, nationalism, and manners (Al-Tahir Mikky a.o. 1986:69). A review of the topics offered in the Arabic curricula of four Arab countries (Saudi Arabia, Sudan, Egypt, and Bahrain) at this stage reveals the focus and importance of each topic (see Table 5).

Note that in Table 4 religion ranks fourth after nationalism. Description comes in fifth, where the student is exposed to topics dealing with concrete objects rather than abstract concepts, such as describing what exists in the student's environment. Also, topics that indicate the value of work and science occur less often than topics like nature or nationalism (for more discussion, see Al-Tahir Mikky a.o. 1986:69–82).

4.2 *Teaching Arabic in preparatory school*

According to Khater (1963:7), while the goal of teaching Arabic at the primary level is to build basic language skills, the goal at the preparatory stage is to build upon and master those language skills for the purpose of writing, reading, and oral communication. Thus, at this level, the learners are expected to be prepared to function in different social situations and to begin their exploration of Arabic literature.

At the end of this stage, the students are expected to be able to cope with situations that require the use of Arabic for written and oral expression, for instance for delivering speeches. In writing, the learners are expected to write compositions and personal letters. In reading, the students must be able to read a variety of material and to interpret and summarize what is read. At this stage, a student is expected to cover most of Arabic grammar (Khater 1963:7–9).

An analysis of the curricula of Egypt and Iraq reveals that at the preparatory level attention is paid to the development of oral communication (see Tables 6 and 7). In addition, more

Table 4. Number of poetry (Py) and prose (Pr) texts in the curriculum of four Arab countries at the primary level

Primary grades	1		2		3		4		5		6	
Country/ literary texts	Py	Pr	Py	Pr	Py	Pr	Py	Pr	Py	Pr	Py	Pr
Saudi Arabia	–	–	–	–	388	8	19	379	19	515	159	12
Sudan	92	–	95	–	136	–	–	309	92	41	190	8
Egypt	24	–	–	–	88	–	–	131	–	72	–	82
Bahrain	–	–	–	–	100	1	6	105	120	4	85	4

Source: Al-Tahir Mikky a.o. (1986:68)

Table 5. Topics in the primary-level Arabic curriculum in four Arab countries

Country/ topics	Manners	Nature	Nationalism	Religion	Description	Work*	Family	Social Science	Science	Ideals	*Totals*
Saudi Arabia	448	151	148	133	153	109	69	82	46	102	1,441
Sudan	295	235	193	81	56	41	16	44	53	–	1,014
Egypt	163	59	32	49	10	27	41	8	15	6	410
Bahrain	60	128	82	46	1	19	24	16	12	–	388
Totals	966	573	455	309	220	196	150	150	126	108	3,253

* Work: Topics that motivate students to work

Source: Al-Tahir Mikky a.o. (1986:69)

emphasis is placed on reading than on other language skills, while grammar still occupies a place in the curriculum. As for composition, emphasis is placed more on writing than on oral expression.

Table 6. Preparatory school Arabic curriculum in Egypt (1981): Number of periods per week

Grades/Branches	1	2	3
Reading	2	2	1.5
Composition	1	1	1
Texts	1	1	1.5
Grammar and application	1	2	2
Handwriting and calligraphy	2	1	–

Source: Khater a.o. (1981:71)

Table 7. Preparatory school Arabic curriculum in Iraq (1970): Number of periods per week

Grades/Branches	1	2	3
Reading	1	1	1
Literature	–	–	2
Memorization	1	1	–
Composition and dictation	2	2	1
Grammar	2	2	2

Source: Altoma (1970:701)

The study of literary texts, including modern and classical poems and prose, for comprehension and appreciation is essential at this level. It paves the way for the study of literature at the secondary stage. In the preparatory stage, three topics are added to those from the primary stage: nationalism, love, and elegiac poetry. Religious texts from the *Qur'ān* and the *Ḥadīṯ* are incorporated during the third and final year of the preparatory stage (see Al-Tahir Mikky a.o. 1986:82–88 for more information).

There is a tendency at this stage to train students for description, thereby moving from describing concrete objects to more abstract notions such as the beauty of nature. As pointed out by Al-Tahir Mikky a.o. (1986:85), not many texts reflect the development of personality and maturity at this stage: for example, few poetic texts are offered that provide a thematic representation of love, while the *Qurʾān* and *Ḥadīṯ* texts are presented throughout the three years.

4.3 *Teaching Arabic in secondary school*

According to Khater (1963:7–9), the primary goal at the secondary stage is to prepare students to master the reading and writing skills they need in order to succeed in the future, either in the workforce or at the university. By the end of this stage, students are expected to have developed their communicative skills, their appreciation for literary works, and their understanding of Islamic history.

Again comparing Egypt and Iraq, Tables 8 and 9 reveal that the curriculum in both countries is divided into two sections in the secondary stage: literary studies and science. However, where students study literature under both sections in Egypt, in Iraq literature is only studied in the literary section. As for literature, students begin the first year of the secondary stage in Egypt by studying the pre-Islamic, Islamic, and Umayyad periods. Students then go on to learn more about the Abbasid period during the second year and round out their education by studying contemporary literature in the third year. Students are trained to analyze texts across different literary periods, focusing on meaning, vocabulary, structure, images, and metaphors. Texts include novels, short stories, essays, and dramas. The aim of reading such literature is to teach students to appreciate a variety of literary works.

Table 8. Secondary school Arabic curriculum in Egypt (1981): Number of periods per week

Branches/grades		Literary section		Science section	
	1	2	3	2	3
Reading	2	2	2	2	2
Composition	1	1	1	1	1
Literature and texts	3	4	4	2	2
Criticism, rhetoric, grammar	1	–	–	–	–

Source: Khater a.o. (1981:68)

Table 9. Secondary school Arabic curriculum in Iraq (1970): Number of periods per week

Branches/grades	Literary section		Science section	
	1	2	1	2
Reading	1	1	1	2
Composition and dictation	1	1	1	1
Literature	2	2	–	–
Rhetoric	1	1	1	–
Memorization	1	1	1	1
Grammar	1	1	1	1

Source: Altoma (1970:701)

5. METHODOLOGY OF ARABIC INSTRUCTION AND PROFESSIONAL TRAINING OF TEACHERS

It is evident from past complaints about the quality of standards of instruction and the low level of achievement (Heyworth-Dunne 1939) as well as present complaints (M. Ibrahim 1993) that radical reforms are needed in Arabic language teaching and learning methodology.

Since the rise of Islam, mosques have been the schools where students study Arabic within an Islamic context. The medieval practice of paying schoolmasters according to their results remains a traditional practice in the learning of Arabic. Listening, recitation, and memorization of Arabic have been common methods for learning Arabic. To some extent those methods are still practiced at the present time, for instance in the religious education provided in Egypt.

Moreover, this approach has been working well because it assures educational results (Spolsky 1999). The results are manifested in the students' performance in memorizing parts of the *Qurʾān* or even the entire Qurʾānic text.

While changes in the methodology of teaching Arabic as a foreign language have reflected the move toward language proficiency in general and oral proficiency in particular, the teaching of Arabic as a first language has reflected common goals of formal reading and writing communication. Generally, the methodology of teaching Arabic may be characterized as traditional in the sense that the teacher plays a central role in the classroom. The teacher's role is that of a knowledge giver. This concept is realized in the form of a lecture in which the focus is the language lesson itself, rather than the development of a student's ability to use Arabic in class. The student's role is to listen, recite, and memorize the lesson rather than to express, discuss, and be creative with the language. Since the primary focus is reading and writing, listening and speaking skills are not emphasized in the methodology of teaching Arabic. For example, speaking *fuṣḥā* is not a priority in the classroom. Arabic grammar is taught and learned through heavy reliance on rote memorization of grammatical rules. Reading aloud is still a common practice in the teaching of Arabic, while reading for comprehension receives far less attention in the curriculum. Arabic is not taught as a tool for communication. Compared with other languages, there is not much classroom-based research available concerning Arabic language teaching (see A. Ibrahim 1966 for more discussion of the methods of teaching Arabic). Most teachers derive their knowledge of the teaching and learning of Arabic from their own experience after graduating from the university or higher institutes.

The field of Arabic language teaching has few professionally trained teachers who are able to manage effective instruction in class. There are not many training programs for Arabic language teachers. Many of the programs available for teacher training, for instance at Cairo University, still follow traditional methods of training Arabic teachers. The ability to teach Arabic requires not only knowledge of the Arabic language but also teaching skills in terms of classroom management. According to England (2006), practice means that teachers of Arabic should have pedagogical skills, i.e., they need to know how to apply their knowledge about Arabic to learners of Arabic and how to manage the classroom effectively. The role of the teacher in the classroom has changed from an authority figure and provider of knowledge to the one who facilitates the learning of Arabic. Traditional techniques such as text recitation, memorization, and drills should be expanded to include classroom interaction where the students use Arabic in order to create and produce ideas, sentences, and questions. In addition, Arabic language teachers need to have workshops to learn new methods and share techniques in the classroom. According to Massialas and Jarrar (1992:39), the training of teachers in the Arab world faces a number of problems. Primary school teachers are not professionally prepared to teach Arabic, and there is a common assumption in the Arab world that primary school teachers do not need professional training, unlike secondary school teachers. Secondary school teachers, except those trained in faculties of education or departments of Arabic, are not properly trained in pedagogy to teach Arabic. Moreover, Arabic teachers do not have a role in formulating the curriculum and selecting textbooks. Everything in the teaching and learning process is prescribed by the country's ministry of education.

6. Conclusion and critical observations

Pedagogy in the teaching and learning of Arabic as a first language faces many challenges:

i. The educational policy that governs the Arabic curriculum at the different school levels needs to be modified to fit the needs of Arab society. Current policy lacks a comprehensive vision of a functional curriculum and objectives. The lack of accountability in terms of practical skills in Arabic language curricula has made the present curricula inadequate for attaining the desired goals and objectives of the learners. The Arabic curriculum is based on the assumption that students should master the various subjects that make up the whole of language. However, this approach is directed primarily toward the attainment of reading and writing skills, while neglecting the advancement of speaking, problem solving, or critical thinking skills. Therefore, the curricula written for Arabic have proved to be ineffective. Except for a few attempts by Arab countries, e.g. the Gulf Arab States (1984–1985) and Egypt (1995), to reform the Arab curricula, not many initiatives have been taken to establish general parameters related to a student's use of *fuṣḥā*, either in the area of general education or in communication connected to academic content areas, to meet sociolinguistic or pragmatic objectives in terms of both oral and written skills. It should be noted that the attempts made by the Arab countries to renew the curricula of Arabic fall short of specifying standards of linguistic performance that are expected to be attained by each student at the conclusion of each year and at the end of each stage.

ii. The goals of teaching Arabic as a first language have to be defined clearly. A comparison of the problems faced by native Arab students in learning Arabic 50 years ago with those met by contemporary students reveals that few fundamental modifications have been made, since the same issues continue to surface. It is doubtful that all the goals mentioned in the curricula are adequately met at each educational stage. While religious education has clear and specific goals with a defined standard, represented in the *Qur'ān*, the goals of education systems that are not religiously affiliated lack this clarity. The goals of instruction in Arabic as a first language in elementary, preparatory, and secondary schools are not accompanied by either proficiency scales or guidelines. Accordingly, curricula are not designed to meet such scales. Assumptions about the goals of learning and teaching Arabic and the needs of the students in using Arabic are made at a distance, by ministries of educations. No description is provided of what educated native learners of Arabic as a first language should be able to accomplish at the end of each educational level. A description might follow the guidelines proposed by ACTFL (American Council on the Teaching of Foreign Languages), or the one proposed by the Council of Europe (2001) for learners of foreign languages. Such a description would constitute an important step in defining the goals of the teaching and learning of Arabic as a first language and the stages to be followed to meeting them. (For further discussion of scales and guidelines, see Spolsky 1999:391.)

iii. The issue of → diglossia poses its own problems in the Arabic language classroom, due in part to the differences between Classical Arabic and the colloquial dialects. Problems begin to surface as soon as Arab children start learning Classical Arabic in schools (for more discussion see Altoma 1969; Salegh-Haddad 2003). One example of this problem can be found in reading, where Arab children face an uneasy transition when attempting to relate what they have read to what they have heard and used in daily situations. These situations clearly affect word recognition and language comprehension. As Maamouri (1998:45) states, "The need for language comprehension as a prerequisite to the acquisition of the decoding skills is an aberration" in the continuous development of Arabic reading skills by children in schools. As a result, a feeling of linguistic insecurity is manifest among young learners of Arabic.

iv. Arabic language and textbook materials are not optimal learning tools. The way language materials are presented in textbooks fails to motivate students to learn Arabic because the books fail to deal with authentic contemporary topics. In addition, the quality of the paper and print makes the textbooks unattractive for the study of Arabic. Many of the pictures presented in textbooks are neither expressive nor related to the text, as becomes clear from a comparison of the Egyptian and Jordanian textbooks for the three educational stages for the school year 1995–1996. For more discussion of Arabic language materials, see Barhum (1997:337–348); Al-Kurdi (1997: 350–400). In addition, many of the Arabic-language materials focus on reading and writing skills to the detriment of listening and oral skills in Arabic.

BIBLIOGRAPHICAL REFERENCES

ACTFL. 1989. “ACTFL Arabic Proficiency Guidelines”. *Foreign Language Annals* 22.373–392.

ALECSO. 1981. *Muḥāwalāt tawḥīd wa-taṭwīr al-manāhij ad-dirāsiyya fī l-bilād al-ʿarabiyya*. Tunis: ALECSO.

——. 1986. *Tawḥīd as-sullam at-taʿlīmī fī l-bilād al-ʿarabiyya*. Tunis: ALECSO.

Al-Haj, Majid. 1996. *Education among the Arabs in Israel: Control and social change*. Jerusalem: Magnes Press.

Al-Kurdi, Wasim. 1997. “A critical review of the Arabic language textbooks from grade 7 to grade 12 in the West Bank and Gaza Strip”. Palestinian Curriculum Development Centre (1997:351–400).

Al-Tahir Mikky, Ahmad [aṭ-Ṭāhir ʾAḥmad Makkī], ʾAḥmad Ḥasan Ḥannūra, Muḥammad aṭ-Ṭayyib ʿAbdallāh, and Wadāʿa Muḥammad al-Ḥasan ʿAkūd. 1986. *Manāhij taʿlīm al-ʾadab wa-n-nuṣūṣ fī marāḥil at-taʿlīm al-ʿāmm fī l-waṭan al-ʿarabī*. Tunis: ALECSO.

Altoma, Salih J. [Ṣāliḥ aṭ-Ṭumā]. 1957. *The teaching of Classical Arabic to speakers of the colloquial in Iraq: A study of the problem of linguistic duality and its impacts on language education*. Ph.D. diss., Harvard University.

——. 1969. *The problem of diglossia in Arabic: A comparative study of Classical and Iraqi Arabic*. Cambridge, Mass.: Harvard University Press.

——. 1970. “Language education in Arab countries and the role of the academies”. *Current trends in linguistics*, VI, ed. Thomas A. Sebeok, 690–720. The Hague: Mouton.

Amara, Muhammad and Abd Al-Rahman Mar’i. 2002. *Language education policy: The Arab minority in Israel*. Dordrecht: Kluwer Academic Publishers.

Badawi, Elsaid. 2002. “In the quest for level 4+ in Arabic: Training level 2–3 learners in independent reading”. *Developing professional-level language proficiency*, eds. Betty Lou Leaver and Boris Shekhtman, 156–176. Cambridge: Cambridge University Press.

Badawi, Muhammad Mustafa. 1993. *A short history of modern Arabic literature*. Oxford and New York: Oxford University Press.

Boktor, Amir. 1963. *The development and expansion of education in the United Arab Republic*. Cairo: American University in Cairo Press.

Barhum, Musa. 1997. “Evaluation of the Arabic language curriculum for grade 1–6 in the West Bank and Gaza Strip”. Palestinian Curriculum Development Centre (1997:337–348).

Council of Europe. 2001. *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.

Egypt. Ministry of Education. 2003–2004. *Curri-culum of Arabic language*. Cairo: Ministry of Education.

England, Liz. 2006. “Methodology in Arabic language education”. *Handbook for Arabic language teaching professionals in the 21st century*, eds. Liz England, Zeinab Taha, and Kassem Wahba, 419–436. Mahwah, N.J.: Erlbaum.

Ferguson, Charles A. 1959. “Diglossia”. *Word* 15.325–340.

Gezi, Kal. 1979. *The educational system of the Arab Republic of Egypt*. Washington, D.C.: U.S. Department of Health, Education and Welfare.

Gulf Arab States Educational Research Center. 1984–1985. *United formula for goals of subjects in general education stages in the Arab Gulf States*. 4 vols. Riyadh: Arab Bureau of Education for the Gulf States.

Heyworth-Dunne, James. 1939. *An introduction to the history of education in modern Egypt*. London: Luzac.

ʾIbrāhīm, ʿAbd al-ʿAlīm. 1966. *Al-Muwajjih al-fannī li-mudarrisī l-luġa al-ʿarabiyya*. Cairo: Dār al-Maʿārif.

Ibrahim, Muhammad H. 1993. Foreword to *Toward a literate world: Television in literacy education – lessons from the Arab region*, ed. Edward L. Palmer, ix–x. Oxford and Boulder, Colo.: Westview Press.

Khater, Mahmoud Rushdi [Maḥmūd Rušdī Xāṭir]. 1963. *The teaching of Arabic in the Arab world*. Washington, D.C.: Center for Applied Linguistics.

——, Yūsuf al-Ḥammādī, Muḥammad ʿIzzat ʿẠbd al-Mawjūd, Rušdī ʾAḥmad Ṭiʿīma and Ḥasan Šiḥāṭa. 1981. *Ṭuruq tadrīs al-luġa al-ʿarabiyya wa-t-tarbiya ad-dīniyya fī ḍawʾ al-ittijāhāt at-tarbawiyya al-ḥadīṯa*. 2nd ed. Cairo: Dār al-Maʿrifa.

Maamouri, Mohammed. 1998. “Arabic diglossia and its impact on the quality of education in the Arab region”. Discussion paper prepared for the World Bank Mediterranean Development Forum.

Marlow-Ferguson, Rebecca and Chris Lopez. 2002. *World education encyclopedia: A survey of educational systems*. Detroit, Mich.: Gale Group.

Massialas, Byron G. and Samir A. Jarrar. 1983. *Education in the Arab world*. New York: Praeger.

——. 1991. *Arab education in transition: A source book*. New York: Garland.

Matthews, Roderic D. and Matta Akrawi. 1949. *Education in Arab countries of the Near East*. Washington, D.C.: American Council on Education.

Medley, Frank W. 1985. “Designing the proficiency-based curriculum”. *Proficiency, curriculum, articulation: The ties that bind*, ed. Alice C. Omaggio, 13–40. Middlebury, Vt.: Northeast Conference on the Teaching of Foreign Languages.

Mujāwir, Muḥammad Ṣalāḥ. 1974. *Dirāsa tajrībiyya li-taḥdīd al-mahārāt al-luġawiyya fī furūʿ al-luġa al-ʿarabiyya*. Kuwait: Dār al-Qalam.

NCEART. 1978. *A report of the National Council of Education, Academic Research and Technology*. Cairo: Specialist National Council.

Nucho, Leslie S. (ed.). 1993. *Education in the Arab world*, I. Washington, D.C.: Amideast.

Palestinian Curriculum Development Centre. 1997. *A comprehensive plan for the development of the First Palestinian Curriculum for Central Education*. Ramallah: PCDC.

Richards, Jack C. 2001. *Curriculum development in language teaching*. Cambridge: Cambridge University Press.

Salegh-Haddad, Elinor. 2003. “Linguistic distance and initial reading acquisition: The case of Arabic diglossia”. *Applied Psycholinguistics* 24.431–451.

Spolsky, Bernard (ed.). 1999. *Concise encyclopedia of educational linguistics*. Oxford: Elsevier Science.

Stetkevych, Jaroslav. 1970. *The Modern Arabic literary language: Lexical and stylistic developments*. Chicago: University of Chicago Press.
Suleiman, Yasir. 1999. "Language education policy: Arabic-speaking countries". Spolsky (1999:106–116).
UNESCO Institute for Statistics. 2002. *Arab States Regional Reports*.

KASSEM M. WAHBA (Georgetown University)

Focus

1. DEFINITION

'Focus' belongs to the field of text-linguistic or discourse-pragmatic phenomena, or more precisely, that of information structure. Although focus is determined on a different level than syntactic relations, such as subject, object, and adverbials, it nevertheless belongs to the grammar and syntax of a language.

Focus is often defined as the new element in a sentence. According to Bolinger, "It marks the 'point' of the sentence where there is the greatest concentration of information, which the hearer would be least likely to infer without being told" (1954:152). According to Halliday, "What is focal is 'new' information; not in the sense that it cannot have been previously mentioned, although it is often the case that it has not been, but in the sense that the speaker presents it as not being recoverable from the preceding discourse" (1967:204ff.). Jackendoff (1972:230) speaks about the "presupposition of a sentence" as "the information in the sentence that is assumed by the speaker to be shared by him and the hearer", whereas "the focus of a sentence" is the information in the sentence that is assumed by the speaker not to be shared by him [or her] and the hearer". Lambrecht (1994) elaborates on these concepts and defines focus as "the semantic component of a pragmatically structured proposition whereby the assertion differs from the proposition". He defines the (pragmatic) assertion as "the proposition expressed by a sentence which the hearer is expected to know as a result of hearing the sentence uttered". His definitions exclude the possibility of segmenting a proposition into elements of new and old information. Focus is a pragmatic relation that combines the presupposition and the assertion into a new proposition, as in (1).

(1) Q: Where did you go last night?
A: I went to the movies.

The information conveyed by the answer is the abstract proposition 'The place I went to last night was the movies', not 'movies' or 'to the movies'.

2. FOCUS DOMAIN

Focus is marked through a prominent accent on a word or a minimal constituent. In a sentence, the syntactic domain that expresses the focus component is the 'focus domain'. The focus domain is always a phrasal category (verb or adjective phrase, noun phrase, prepositional phrase, or sentence), not a lexical category, which would not be able to express the relations in question. This is evident from (2), where the capitalized word has focus accent.

(2) Q: Which shirt did you buy?
A: I bought the GREEN one
(or: The GREEN one, but not: *GREEN)
(Lambrecht 1994:216)

The focus domain is *the GREEN one* or the VP *bought the GREEN one*, but not *GREEN*.

As is obvious from the example, the focus domain may contain nonfocal elements; the NP *the GREEN one* is in focus, but the constituents *the* and *one* are topical, which is also why the accent falls on the penultimate element.

There are three principles that determine the placement of the accent (Lambrecht 1994:238–257). First, according to the Iconicity Principle, the prosodic peak falls on the most important communicative element. In this case, the relationship between prosodic prominence and communicative importance is at least partly iconic. Second, the General Phrasal Accent Principle locates the phrasal accent at the right boundary focus domain, marking the end of it. This can be checked by the Unaccentable Element Principle, according to which there are elements, such as topical expressions, that are not accented. Then, by default, the accent is moved to the next element to the left, as happens in (2).

3. FOCUS TYPES

Lambrecht (1994:221–238) discerns three focus-structure categories: predicate focus, which occurs in topic/comment sentences; argument focus, which occurs in identificational sentences; and sentence focus, which appears in event-reporting sentences (→ topicalization for more information on these three sentence types). The examples in (3)–(5) illustrate these categories.

(3) PREDICATE-FOCUS STRUCTURE
(Topic-comment sentence)
(What happened to your car?) My car/it broke DOWN.

(4) ARGUMENT-FOCUS STRUCTURE
(Identificational sentence)
(I heard your motorcycle broke down?) My CAR broke down.

(5) SENTENCE-FOCUS STRUCTURE
(Event-reporting sentence)
(What happened?) My CAR broke down.

In the identificational sentence (4), everything is presupposed except the argument in focus. In the event-reporting sentence (5), on the other hand, nothing is presupposed; no topic is present and, consequently, the whole sentence constitutes the focus domain. The event-reporting type belongs to the category of 'thetic' sentences, which, according to Lambrecht (1994:144), lack topical information in that they are all-new in character. Presentational sentences and sentences with weather verbs are also thetic. The former present a new 'entity' or referent, whereas the latter presents a new 'event'.

The placement of the accent follows the Iconicity Principle in both (4) and (5). The accent in these sentences does not disambiguate between sentences like (4) and (5). This becomes clear from studying the pragmatic structure, through identifying what is presupposed and what is the assertion (Lambrecht 1994:307–311). For example, the sentence in (6) –

(6) *BAQARATUN takallamat*
an ox has spoken
'An OX has spoken' (Wright 1975:II, 263)

– may be either an identificational sentence ('It was not a human being, *an OX has spoken*') or an event-reporting sentence ('What happened? *An OX has spoken*'). Lambrecht (1994:264ff.) observes that nominal referents receive an accent to a much larger degree than verbs, which explains why the accent is on the subject in an event-reporting sentence.

From the examples it may be observed that English has exclusively prosodic mechanisms for marking focus. In French, on the other hand, we find different constructions whereby the argument-focus structure would be *C'est ma VOITURE qui est en panne* 'My CAR broke down', and the sentence-focus structure would be *J'ai ma VOITURE qui est en PANNE* 'My CAR broke down'. Italian uses word-order variation: *Si è rotta la mia MACCHINA* 'My CAR broke down' (argument focus) and *Mi si è rotta (ROTTA) la MACCHINA* 'My CAR broke down' (sentence focus).

The predicate-focus structure is considered to be the unmarked focus structure, argument-focus and sentence-focus being the marked ones. Thus a predicate-focus structure has more than one interpretation. To preclude its 'default' topic/comment reading, the predicate must be prosodically marked by the absence of prominence, which in most cases means an accent on the subject. Such readings are contextually determined. Sentences with unaccented predicates will then be either of argument-focus or sentence-focus structure. However, in topic/comment clauses (with predicate-focus structure), both subject and predicate can be accented, as is the case with contrastive topics.

(7) I saw Mary and John yesterday, SHE says HELLO, but HE's still ANGRY at you. (Lambrecht 1994:291)

4. ARABIC

According to Moutaouakil (1989), Modern Standard Arabic may use word order variation to express focus, as in cases where a preposed constituent may constitute an argument-focus structure. This is illustrated by examples (8) and (9).

(8) *RIWĀYATAN ʾallafat zaynab-u*
novel-Acc wrote3fs Zaynab-Nom
(lā QAṢĪDAT-AN)
(not poem-Acc)
'It was a NOVEL that Zaynab wrote (not a POEM)'

(9) *LAYLĀ ʿašiqa qays-un*
Layla loved3ms Qays-Nom
(lā ZAYNAB-A)
(not Zaynab-Acc)
'It was LAYLA that Qays loved (not Zaynab)'

However, it appears to be very rare to prepose an object without using a suffixed pronoun that refers to this element in the following clause, the object being left-dislocated (Dahlgren 1998:176; also → word order). For focalized objects, Verb-Subject-Object is a possible, but marked order, as illustrated by example (10) from the *Qurʾān*.

(10) *fa-ʾawjasa fī nafs-i-hi*
and-he conceived in soul-Gen-his
xīfat-an mūsā
fear-Acc Moses
'and Moses conceived A FEAR within him' (Q. 20/67)

Postposed independent personal pronouns may also mark argument focus:

(11) *ʾammā huwa fa-yaddaʿī ʾanna-hu*
but he then -he claims that-he
lam yabṣuq ʿalā mirʾāt-i-hā
not-past he spat on mirror-Gen/of it/3fs
'but he claimed that he did not spit at her mirror'

bal baṣaqa fī wajh-i-hi huwa
but he spat in face-Gen/of-his he
ḥīna ṭālaʿa-hu fī l-mirʾāt
when he inspected-it/3ms in the-mirror
'but he spat at HIS face when he inspected it in the mirror' (Bloch 1974:57)

Another type of argument focus is cleft focus, which is construed with *allaḏī* in Arabic, as in (12).

(12) *fa-huwa l-ʾān-a yartādu ʾaġlab-a*
so he now-Acc frequents most-Acc
ʾamkinat-i l-lahw-i
places-Gen the-amusement-Gen
'So he now frequents most places of entertainment'

wa-yaṭlubu mā yurīdu dūna
and demands what he-wants without
ʾan yajruʾa ʾaḥad-un
that he dares anybody-Nom
'and demands whatever he wants, without anybody daring'

ʿalā l-iʿtirāḍ-i ʾaw al-muṭālaba
to the-objection-Gen or the-demand
'to object or demand anything'

bal huwa llaḏī yataqāḍā min
but he the one he claims from
ʾaṣḥāb-i-hā l-ʾitāwāt-i
owners-Gen-their3fs the-tax-Gen
'for HE is the one who claims from their owners taxes'

wa-l-murattabāt-i li-ḍamān-i
and-the-salaries-Gen to-guaranteeing-Gen
l-hudūʾ-i fī hāḏihi l-maḥāll
the-calm-Gen in these3fs the-places
'and salaries to guarantee calm in these places' (Bloch 1974:87)

In the proposition in question it is only *huwa* which is asserted; the rest is presupposed information from the preceding context.

4.1 *Sentence-focus and focus markers*

According to Ouhalla (1997:20-25), particles like *ʾinna(mā)* and *(la)qad*, known as *ḥurūf at-tawkīd* 'particles of corroboration/confirmation', are used to 'reinforce/confirm' the propositional content of a given sentence. These are to be seen as marked forms of topic/comment sentences, as illustrated in (13) and (14), where FM represents focus marker:

(13) *ʾinna zayd-an muhājir-un*
FM Zayd-Acc emigrating-Nom
'Zayd IS emigrating'

(14) *(la)qad ʾarsalat zaynab-u risālat-an*
FM sent3fs Zaynab-Nom letter-Acc
'Zaynab DID send a letter'

However, *la-* may be used independently as a 'constituent focus marker', i.e. an argument focus, in Lambrecht's terminology, as in (15) and (16).

(15) *ʾinna zayd-an la-MUHĀJIR-un*
FM Zayd-Acc FM-emigrating-Nom
'Zayd is EMIGRATING'

(16) *la-MUHĀJIR-un zayd-un*
FM-emigrating-Nom Zayd-Nom
'Zayd is EMIGRATING'

4.2 *Interrogative focus*

Classical and Standard Arabic have two yes/no-question particles: *hal* and *ʾa* (→ interrogative sentence). According to Ouhalla (1997:26–31), the scope of the former covers the whole sentence, whereas the scope of the latter is the constituent. He presents the examples (17)–(20):

(17) *hal ʾallafat zaynab-u l-qaṣīdat-a*
Q wrote3fs Zaynab-Nom the-poem-Acc
'Did Zaynab write the poem?'

(18) *ʾa ʾallafat zaynab-u l-qaṣīdat-a*
Q write3fs Zaynab-Nom the-poem-Acc
(ʾam ʾalq-at-hā)?
(or read-3fs-it)
'Did Zaynab WRITE the poem (or READ it)?'

(19) *ʾa zaynab-u ʾallafat al-qaṣīdat-a*
Q Zaynab-Nom write3fs the-poem-Acc
(ʾam laylā)?
(or Layla)
'Did ZAYNAB write the poem (or LAYLA)?

(20) *ʾa l-qaṣīdat-an ʾallafat zaynab-u*
Q poem-Acc write3fs Zaynab-Nom
(ʾam riwāyat-an)?
(or a novel-Acc)
'Was it a POEM Zaynab wrote (or a NOVEL)?'

4.3 *Negation and focus*

In the use of negation, only the asserted portion of the corresponding affirmative is denied, while the presupposition stays outside the negative scope. There is a close connection between the focus domain and the scope of the negation. In nominal clauses in Standard Arabic the scope of the negation is introduced by the particle *bi-*, as in (21):

(21) *mā zaydun bi-ʿālim-in*
not Zayd-Nom PRT-scientist-Gen
(bal šāʿir-un)
(but poet-Nom)
'Zayd is not a SCIENTIST (but a poet)'

4.4 *Intonation in Moroccan*

There is a general lack of phonetic studies on intonation in the modern Arabic dialects. For Moroccan, though, Benkirane (1998) finds that 'sentence accent' – corresponding to Lambrecht's General Phrasal Accent and in accordance with it – falls on the penultimate of the word of a phrase or utterance farthest to the right. Benkirane also finds the Iconicity Principle to be in operation when the speaker deviates from the unmarked pattern to modify and focus a particular constituent in an utterance. One of his examples is given in (22), where capitalization indicates pitched constituent.

(22) *ʾamina ma-zäla mreḍa*
Amina not-finished sick
'Amina is ill AGAIN'

Bibliographical references

Anshen, Frank and Peter Schreiber. 1968. "A focus transformation of Modern Standard Arabic". *Language* 44.792–797.

Benkirane, Thami. 1998. "Intonation in Western Arabic (Morocco)". *Intonation systems: A survey of twenty languages*, ed. Daniel Hirst and Albert Di Cristo, 345–359. Cambridge: Cambridge University Press.

Bloch, Ariel. 1974. *A chrestomathy of Modern Literary Arabic.* Wiesbaden: O. Harrassowitz.

Bolinger, Dwight. 1954. "English prosodic stress and Spanish sentence order". *Hispania* 37.152–156.

Cantarino, Vicente. 1974–1975. *The syntax of Modern Arabic prose.* 3 vols. Bloomington, Ind.: Indiana University Press.

Comrie, Bernard. 1989. *Language universals and linguistic typology.* 2nd ed. Oxford: Blackwell.

Dahlgren, Sven-Olof. 1998. *Word order in Arabic.* Göteborg: Acta Universitatis Gothoburgensis.

Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar.* Cambridge, Mass.: MIT Press.

Lambrecht, Knud. 1994. *Information structure and sentence form: Topic, focus and the mental representations of discourse referents.* Cambridge: Cambridge University Press.

Moutaouakil, Ahmed. 1989. *Pragmatic functions in a functional grammar of Arabic.* Dordrecht: Foris.

Ouhalla, Jamal. 1997. "Remarks on focus in Standard Arabic". *Perspectives on Arabic linguistics*, X, ed. Mushira Eid and Robert A. Ratcliffe, 9–46. Amsterdam and Philadelphia: J. Benjamins.

Van Valin, Robert D. and Randy J. Lapolla. 1997. *Syntax: Structure, meaning and function.* Cambridge: Cambridge University Press.

Wright, William. 1975. *A grammar of the Arabic language.* 3rd ed. Repr. Cambridge: Cambridge University Press.

Sven-Olof Dahlgren (Göteborg University)

Foregrounding → Grounding

Foreigner Talk

Foreigner Talk is a continuum of formal and discourse modifications used by native speakers in communicating with nonnative speakers/ learners (Ellis 1994:247). It is an automatic process triggered by the native speaker's realization that the nonnative speaker's proficiency level is low (Gass and Varonis 1985:149–162). The degree of modification is determined by the level of proficiency of the particular nonnative interlocutor in a certain interactive context (Gass 1997:66). Native speakers' modifications can affect both linguistic and discourse aspects of language (Long 1983a:177–193). Empirical studies indicate that adjustments on both levels are common among all languages and that this type of language modification is governed by universal mental rules (Tweissi 1990:297).

Research on the formal aspects of Foreigner Talk concentrates on grammaticality issues (Ferguson 1971, 1975; Ferguson and DeBose 1977; Long 1980; Larsen-Freeman and Long 1991) and the nature of linguistic modifications (Ellis 1994:254–257). Research on discourse aspects concentrates on the nature of discourse modifications (Long 1983b); the structure of Foreigner Talk modifications (Arthur a.o. 1980; Derwing 1989); and the functions and triggers of Foreigner Talk (Gass and Varonis 1984; Varonis and Gass 1985).

Different methods have been used for the collection and analysis of Foreigner Talk data, the differences being partly connected with the purpose of the research. Direct audio and/or video recording is one method that concentrates on the spontaneous aspects of a real interaction between interlocutors. It allows the documentation of language data that can be analyzed on all linguistic levels, including not only the phonological, morphological, and syntactic levels but also the discourse level. Some researchers record controlled conversations (Tweissi 1990; Arthur a.o. 1980) and others record free conversation (Sharkawi 2005), while yet others practice a degree of control on both ends (Henzl 1979; Håkansson 1986). The data produced by semi- or uncontrolled conversations provide the most natural modifications native speakers produce since no pressure or direction is imposed on the situation.

Self-reports (Sharkawi 2005; Ferguson 1971) and scripted data (Mühlhäusler 1984, 1986) are used to measure interlocutor awareness of the modification and response to it. This type of data can also show how far a native speaker may go in modifying the language, and the extent to which these modifications may be conscious or merely spontaneous adjustments.

Research in registers modified by native speakers represents a new field of inquiry in Arabic linguistics. Two data collections (Tweissi 1990 and Sharkawi 2005) are available to provide an initial overview of Foreigner Talk features in modern Arabic dialects. Both focus on describing the formal modifications of native Arabic speakers in talking to nonnative speakers. The available data are limited to the two urban dialects of Amman and Cairo, respectively, and may therefore not be representative of other urban dialects.

Two basic strategies have been observed in the collected data of Foreigner Talk in Arabic: explanation of lexical items and a tendency toward structural saliency. These two strategies are reflected at all levels of linguistic analysis. Native speakers may achieve their goal of modification without restructuring their native language, but here, the focus will be on structural modifications. On the phonetic level, Arabic Foreigner Talk is characterized by a slower speech rate than native-speaker to native-speaker talk (Tweissi 1990:305). In Foreigner Talk an average of 3.06 syllables per second was found, as opposed to 5.27 syllables in the case of native-speaker speech. Along the same lines, more primary stress on words was observed in the case of Foreigner Talk than in native-speaker talk. In Foreigner Talk, 2.31 stressed words were found, as opposed to 1.35 words in native-speaker talk per T-unit (a single main clause and the subordinate clauses or non-clauses attached to it). As expected, more pauses and less phonetic and phonological processing were found in Foreigner Talk than in native-speaker talk (Tweissi 1990:305).

In words containing more than one morpheme, short vowels in the word are not deleted, thus allowing a vowel barrier to separate the component morphemes of the word (Sharkawi 2005). This phenomenon is especially clear in the case of verbs. In example (1) the native speaker of Egyptian Arabic asks the nonnative speaker a question (Sharkawi 2005:109).

(1) *bi-ti-ˤraf* *ti-ṭbux*
HAB-2ms 2ms-cook-know
'Can you cook?'

The word *bi-ti-ˤraf* contains three morphemes. The first, *bi-*, is a habitual/continuous mood prefix in Egyptian Arabic. The second, *t-*, is the 2nd person masculine singular imperfect. The third morpheme is the verb stem. The added *-i-* functions as a morpheme boundary to clarify where a meaningful morpheme ends and where another starts.

The same native speaker directs a comment to another native speaker in the same situation in which he expresses his surprise that the nonnative speaker learns Arabic and cooks at the same time (2). In this exchange with another native speaker, the short vowel is deleted (Sharkawi 2005:109).

(2)	*b-yi-tʿallim ʿarabi*	*wi-b-yu-ṭbux̣*
	PROGR-3ms	and-PROGR-3ms
	learn	cook
	'He learns Arabic and cooks'	

All strategies to make the pronunciation more salient are related to one another because a slower speech rate leads to more primary stress and to more pauses. It is to be noted that modification at the level of phonology is produced by means of processes integral to the processes of Arabic phonology.

Interestingly, such measures do not appear in the reports of foreigners who volunteered their experiences with the phenomenon in Arabic. Neither did these modifications appear in scripted data collected from the Egyptian movies (Sharkawi 2005). It is also interesting that the collected data do not attest to phonemic changes that are found in the movies, such as the regular shift of /ḥ/ to /x/, e.g., *xarāmī* instead of *ḥarāmī* 'thief'. This change is not found in the spontaneous data in Sharkawi (2005), nor in Tweissi (1990). The collected data also fail to confirm another phenomenon in Foreigner Talk in movies, which has to do with the place and manner of articulation. Foreigner Talk in the movies replaces /ʿ/ with a glottal stop /ʾ/; the proper noun *ʿimād*, for instance, is regularly converted into *ʾimād*.

The phonetic modifications undertaken by the native speakers of Arabic in the data reported above help nonnative speakers identify major constituents and word boundaries, and give them more processing time. Phonetic modifications in other languages serve the same purpose (Hatch 1983a:66, 1983b:158). The attested modifications in the Arabic Foreigner Talk data are mere suprasegmental changes rather than phonemic or allophonic changes in the features of the phoneme system of Arabic. Modifications representing phonemic changes are not attested in natural data, although they regularly occur in the scripted Foreigner Talk data of the movies.

At the morphosyntactic level, Arabic Foreigner Talk in the collected data is characterized by structural saliency. Among the modification features are redundancy of elements; avoidance of certain elements and morphological forms; and a general tendency toward the use of analytical structures. Generalization of elements is another tendency added by self-reports of nonnative speakers. Native speakers in Arabic Foreigner Talk use slightly longer multi- and single-clause T-units than the ones they use with other native-speaker interlocutors. The average of words per multi-clause T-units is 8.66 to 8.47 words. In single-clause T-units the average of words was 4.72 to 4.59 words per unit (Tweissi 1990:311). At the level of word order, no significant difference is observed between inter-native-speaker talk and Foreigner Talk.

There is, however, a significantly smaller number of main clauses per T-units in Foreigner Talk than in inter-native-speaker talk (Tweissi 1990:314). This means that Foreigner Talk utterances are structurally simpler and more linear than inter-native-speaker talk. This last phenomenon is witnessed in all the sources of data available. Utterances are simple, short sentences. The beginning of each sentence is a redundant nominal or pronominal head. Very few relative sentences and/or embedding are attested in Sharkawi (2005). Whenever relativization occurs, it remains confined to the subject position.

There is also less structural complexity at the level of the individual sentence constituents. In the data collected from Egyptian Arabic, there is a consistent use of redundant independent pronouns after nouns and prepositions that are already modified by a suffix pronoun (Sharkawi 2005:110). The same phenomenon is also attested by Tweissi (1990:313). An example is given in (3).

(3)	*ʾana*	*ha-ʿallimak*	*ʾinta*	*ʿarabi*
	I	FUT-teach1s-2ms	you	Arabic
	'I will teach you Arabic'			

The same type of redundancy is attested at the level of the verb, where, in native speech, the conjugated verb does not need a preceding independent pronoun. In the Arabic Foreigner Talk data, all conjugated verbs are preceded by an independent pronoun, as in (4) (Sharkawi 2005:110).

(4) *ʾiḥna* *bi-n-ḥibb* *id-dawri* *il-ʾīṭālī*
we PROGR-1p-like the-league the-Italian
'We like the Italian league'

This tendency is alluded to by Tweissi (1990:313) and frequently attested in recorded data and the scripted data of Foreigner Talk. Interestingly, nonnative speakers who volunteer their experiences do not mention this tendency.

Another way of using fewer morphemes in one word is the preference of the genitive exponent over the synthetic noun construction. In example (5), the native speaker is talking about the Dutch league of football, when he asks the nonnative speaker about the name of a certain player (Sharkawi 2005:111).

(5) *mīn* *il-lāʿib* *bitāʿ* *ik-kōra*?
who the-player POSS the-football
'Who is the football player?'

Another strategy on the part of the native speaker is avoidance. It occurs at the word level as well as at the sentence level. Native speakers avoid the use of derived verb stems. In example (6), the native speaker asks the nonnative speaker why he did not catch a tan although he spent so much time under the hot Cairo summer sun. He initially asks the question using the derived verb *tismarr*. Quickly, however, he rephrases his question using a clause instead of the verb (Sharkawi 2005:111).

(6) NS *ʾummāl* *ma-smarrit-š*
then NEG-get.a.tan2ms-Neg
yaʿni
it.means
'You have not gotten a tan'
NNS *ha*
[hun?]
'What?'
NS *lēh* *ma-baʾit-š*
why NEG-become2ms.-Neg
ʾasmar?
brown
'Why haven't you got a tan?'

Another avoided element from the recorded native-speaker Foreigner Talk data is the dual ending *-ēn*. In the recorded data the number 'two' is regularly expressed by *ʾitnēn* followed by the noun in plural. In example (7), the native speaker talks about two football players from Egypt who are playing in the German league (Sharkawi 2005:111).

(7) *fīh* *ʾitnēn* *laʿība* *min* *maṣr*
there.are two players from Egypt
'There are two Egyptian players'

Again, it is not clear whether a deliberate avoidance strategy is behind this preference to use the numeral instead of the noun followed by the dual ending, especially since Tweissi's data do not refer to the issue at all.

An interesting phenomenon reported in volunteer reports but not witnessed in the audio-recorded data is the reduction of the category of verb conjugation in the imperfect. Volunteered reports by nonnative speakers claim that native speakers of Arabic delete the imperfect 2nd and 3rd person prefixes on the stem of the verb in speech to the nonnative speakers. The 2nd person *ti-* and 3rd person *yi/ti-* prefixes have been deleted. Native speakers allegedly use forms such as *ʾinta ʾišrab* 'you [masc.] drink', *ʾinti ʾišrabi* 'you [fem.] drink', and *huwwa ʾišrab* 'he drinks'. Although this phenomenon has not been attested in any other data source, it gains credibility because it is mentioned six times by six different self-reporters. It is also reported from the language used by Philippine housemaids in Beirut (Ramzi Baalbaki p.c.), where utterances such as *ana rūḥi* 'I go' with the feminine imperative used as finite verb are used to stereotype the pidginized register of these speakers but reportedly also occur in the Foreigner Talk of native speakers when addressing the housemaids.

In self-reports, but not in the collected data, there is constant reference to the use of one single form for the noun after numerals. All reports list examples (a total of seven) of a numeral being followed by a noun in the singular (Sharkawi 2005). A comparison of these examples shows that in the case of the numerals above 10, like in the Arabic dialects, the numeral is followed by a singular noun. Unlike the dialects, however, in the case of the numerals from 3 to 10, the modifying noun is also in the singular (Sharkawi 2005).

Two phenomena occur in the scripted Foreigner Talk but do not appear anywhere else in the available data, namely, the use of independent pronouns after nouns, prepositions, and particles rather than a pronominal suffix. This occurs in examples (8)–(10) from Sharkawi (2005:112).

(8)	*mrāt*	*ʾinta*	
	wife	you	
	'your wife'		

(9)	*il-gōz*	*bitāʿ*	*ʾinta*
	the husband	POSS	you
	'your husband'		

(10)	*huwwa*	*šāf*	*hiyya*
	he	saw	she
	'He saw her'		

The other phenomenon that appears only in the scripted data is the drastically reduced verb conjugation. The 2nd person feminine singular is used for both masculine and feminine, with the three persons. None of the reports mentions the use of a similarly reduced variety of Arabic in addressing nonnative speakers of Arabic.

The extreme restructuring in the Foreigner Talk of the scripted data is interesting, as it reflects the native speaker's conscious views of the modifications that must be applied when talking to nonnative speakers. Even more interesting, the collected data and the reports do not reflect these extreme modifications. Self-reports and recordings agree on certain modifications as opposed to scripted data. In addition, and most importantly, these modifications never go to the point of heavy restructuring. Native speakers of both Jordanian Arabic and Cairene Egyptian Arabic agree on the points in which they feel they have to modify their language, especially at the level of phonetics. These modifications at the phonetic and morphosyntactic level are real simplifications of the language which make sounds more distinct and structures more transparent.

In addition to the above-mentioned structural modifications, lexical modifications are reported. Foreigner Talk data include the use of foreign lexical items (Tweissi 1990:308). This phenomenon need not detain us here, since such lexical use is a universal phenomenon and does not seem to cause any structural modification of the language. Semantic modifications of Arabic words, however, are relevant because they use aspects within the system.

Tweissi (1990:310) notes that native speakers use what he calls a 'lower type-token' ratio, which means they repeat words the nonnative speakers have heard before in the conversation. In addition, they do not use synonyms and antonyms in explaining words that the nonnative speakers apparently do not understand. Instead, foreign words are used to solve the problem. One is tempted to assume, based on the lower 'type-token' ratio in talking to nonnative speakers, that this phenomenon affects the level of elaboration of the utterance. If native speakers find it difficult for the nonnative interlocutors to comprehend aspects of lexical elaboration (such as synonyms, antonyms, and the use of dependent pronouns and relative clauses and adjectives), and use foreign words to solve a problem, utterances must be short and lexical items repetitive.

The general features listed above reflect a desire on the part of the native speakers to render their output comprehensible to the nonnative interlocutor. Since Foreigner Talk data in Arabic are scarce, it is useful to group the above features into tendencies that may guide further data collection and analysis and may help in understanding the discrepancies between native-speaker interaction with a native speaker and the use of a special register with nonnative interlocutors.

On the phonetic level, Foreigner Talk in the collected data tends to make the sounds more distinct by applying primary stress to them. Sound combinations (words) are also made clearer by adding the full vowel combination without deletion to separate between morphemes. Word boundaries are also marked by pauses. In the natural data, no articulatory modifications were recorded. Native speakers also did not resort to any alteration of the phonological features of sounds. Such modifications and alterations were only represented in the scripted data, as in the case of /ḥ/ changing into /x/.

One of the main tendencies at the morphosyntatic level is the movement toward analytic typology in order to make syntactic relations salient by reducing the number of functions a single word assumes and by expressing syntactic functions by separate words. Hence, the longer multi- and single-clause T-units mentioned in Tweissi (1990:311). This tendency is manifested in the use of analytical structures such as the periphrastic dual, genitive exponents, the use of an auxiliary verb plus an adjective instead of the geminated derived verb, and the use of redundant independent pronouns with nouns modified by a suffix pronoun and conjugated verbs.

The existence of the same tendency of moving toward more analytical structures in the scripted Foreigner Talk in movies leads one to assume that this tendency is a conscious move on the part of the native speakers to simplify their speech. The use of independent pronouns instead of object suffix pronouns on nouns, verbs, and particles is an extreme case of this phenomenon. Although the collected data and the scripted data share this tendency, native speakers seem to avoid completely ungrammatical analytical structures.

In the collected data, speakers tend to produce syntactically and morphologically correct utterances, contrary to popular expectation and unlike the movie Foreigner Talk data. There is no evidence in the collected data of the use of incorrect structures or drastically reduced morphological categories or syntactic structures. Arabic Foreigner Talk in this respect is similar to other languages, where grammatical use of the language is the norm, not the exception (Arthur a.o. 1980:111–112).

The last tendency to be observed in the data is the avoidance of structures that are presumably difficult in favor of other presumably more straightforward structures, usually analytical ones. This can be seen in the absence of relative clauses and passive voice from the collected data. In addition, if we assume that self-reports are representative of Arabic Foreigner Talk in general, then this would represent another tendency toward generalization of certain aspects of verb conjugations and agreement patterns. This is evident in the generalization of the 3rd person singular to the 2nd and 3rd person masculine and feminine.

The study of Foreigner Talk in Arabic is interesting from both the synchronic and historical perspectives. In the first place, it shows that speakers of Arabic exhibit the same universal modifications and simplification in Foreigner Talk as speakers of other languages. It also shows that each language has its own specific treatment of modification. The study of Foreigner Talk in Arabic may also contribute to some of the undecided issues in the field, such as the grammaticality issue (see Ellis 1994:252–257), and it may enrich the discussion with new ideas about the typological implications of Foreigner Talk modifications. In the second place, studying Foreigner Talk strategies in Arabic can shed light on the historical development of Arabic in the period immediately after the Arab conquests of the Middle East, when Arabs and non-Arabs had to communicate extensively in Arabic by providing a scenario of the manner in which such communication may have taken place on the part of the native speaker.

Bibliographical references

Arthur, Bradford, Richard Weiner, Michael Culver, Young Ja Lee, and Dorina Thomas. 1980. “The register of impersonal discourse to foreigners: Verbal adjustments to foreign accent”. *Discourse analysis in second language research*, ed. Diana Larsen-Freeman, 111–124. Rowley, Mass.: Newbury House.

Derwing, Tracey M. 1989. “Information type and its relation to nonnative speaker comprehension”. *Language Learning* 39.157–172.

Ellis, Rod. 1994. *The study of second language acquisition*. Oxford: Oxford University Press.

Ferguson, Charles A. 1971. “Absence of copula and the notion of simplicity”. *Pidginization and creolization of language*, ed. Dell Hymes, 141–150. Cambridge: Cambridge University Press.

——. 1975. “Towards a characterization of English Foreigner Talk”. *Anthropological Linguistics* 17.1–14.

——. and Charles E. DeBose. 1977. “Simplified registers, broken language, and pidginization”. *Pidgin and creole linguistics*, ed. Albert Valdman, 99–125. Bloomington, Ind.: Indiana University Press.

Gass, Susan. 1997. *Input, interaction, and the second language learner*. Mahwah, N.J.: Erlbaum.

—— and Marlos Varonis. 1984. “The effect of familiarity on the comprehensibility of nonnative speech”. *Language Learning* 34:1.65–89.

——. 1985. “Variation in native speaker speech modification to non-native speakers”. *Studies in Second Language Acquisition* 7:1.35–57.

Håkansson, Gisela. 1986. “Quantitative studies of teacher talk”. *Learning, teaching and communication in the foreign language classroom*, ed. Gabriele Kasper. Aarhus: Aarhus University Press.

Hatch, Evelyn. 1983a. “Simplified input and second language acquisition”. *Pidginization and creolization as language acquisition*, ed. Roger Andersen, 64–86. Rowley, Mass.: Newbury House.

——. 1983b. *Psycholinguistics: A second language perspective*. Rowley, Mass.: Newbury House.

Henzl, Vera. 1979. “Foreigner Talk in the classroom”. *International Review of Applied Linguistics* 17:2.159–167.

Larsen-Freeman, Diane and Michael Long. 1991. *An introduction to second language acquisition research*. London: Longman.

Long, Michael. 1980. *Input, interaction, and second language acquisition*. Ph.D. diss., University of California, Los Angeles.

——. 1983a. "Linguistic and conversational adjustments to nonnative speakers". *Studies in Second Language Acquisition* 5:2.177–193.
——. 1983b. "Native speaker/nonnative speaker conversation and the negotiation of comprehensible input". *Applied Linguistics* 4:2.126–141.
Mühlhäusler, Peter. 1984. "Tracing the roots of Pidgin German". *Language and Communication* 4.27–57.
——. 1986. *Pidgin and creole linguistics*. Oxford: Blackwell.
Sharkawi, Muhammad al-. 2005. *Arabicization: A case of second language acquisition.* Ph.D. diss., University of Nijmegen.
Tweissi, Adel. 1990. "'Foreigner Talk' in Arabic: Evidence for the universality of language simplification". *Perspectives on Arabic linguistics*, II, ed. Mushira Eid and John McCarthy, 296–326. Amsterdam and Philadelphia: J. Benjamins.
Varonis, Marlos and Susan Gass 1985. "Miscommunication in native speaker/nonnative speaker conversation". *Language in Society* 14:3.327–343.

Muhammad al-Sharkawi
(American University in Cairo)

Formulaic Speech → Frozen Expressions

French Loanwords

1. Introduction

The literature in Roman alphabet on French loanwords in Arabic is not very voluminous. Noteworthy exceptions are Heath (1989) and Benzakour a.o. (2000), who devote a large part of their work to French loanwords in Moroccan Arabic. Both include a lexicon, the former with phonetic transcriptions. This entry discusses French loanwords only in Moroccan Arabic because it is the Arabic dialect that has been the most documented from the point of view of borrowings, and it is the dialect for which we have a large original corpus. This corpus, which is supervised by Carole Paradis at Laval University, belongs to Project CoPho. The observations in this entry focus exclusively on phonological adaptation of lexical forms.

2. History of French borrowings in Moroccan Arabic

The contact between French and Moroccan Arabic and, consequently, borrowings from the French language began in earnest when Morocco was made a protectorate of France in 1912. Following 1912, French became the official language of Morocco, a role that was played out in those political and administrative bodies put in place by France. Traditional administrative structures existed alongside colonial ones, so there was no need for the majority of Moroccans to know French to conduct their daily business, including legal and financial business. French, however, served as the principal interface language between Moroccans and Europeans and was the main foreign language taught in schools after 1912.

The number of French-medium schools rapidly multiplied after 1912, although, as noted by Benzakour a.o. (2000), very few Moroccans attended schools where French was either the medium of instruction or where it was taught as a foreign language. Indeed, the increase in such schools was initially linked to increases in the number of foreigners – mainly French – in Morocco. French-medium schools were intended for their children. French was also taught in institutions that were designed to train translators and civil servants to meet the practical, political, and social needs of the colonial administration. For many years, though, only a few Moroccans, drawn from those members of the Moroccan elite who were judged to be sympathetic to the colonial regime, had access to an education where French instruction was provided. According to Benzakour a.o. (2000), on the eve of independence in 1956, only 15 percent of school-age children were attending what we might call French-style schools.

It is clear, however, that French enjoys an important status in Morocco (see, e.g., Marley 2002, 2005). After independence, French assumed the status of preferred foreign language, providing access to the Western world, including access to scientific and technical information. It is perceived as the language of modernity and continues to be a mark of social prestige, as evidenced by the fact that it continues to engender a notable body of literature and to be the medium of instruction in institutions of higher learning and the language of diplomacy. Somewhat paradoxically, knowledge of French has become much more widespread since Morocco gained its independence, due to the fact that an education featuring French language instruction subsequently became available to the masses (Marley 2005). Although it is sometimes claimed that only a minority of Moroccans is completely fluent in all aspects of French, French is widely spoken and understood in Morocco, and knowledge of French is considered to be important.

3. Phonological description of Moroccan Arabic

Moroccan Arabic has 28 distinctive consonants, including pharyngealized consonants (which are different from the phonetic emphatic variants). The consonant inventory of Moroccan Arabic is provided in Table 1.

Moroccan Arabic also has three full distinctive vowels, /a, u, i/, which are realized [e, o, ạ] before or after a pharyngealized consonant and often, too, before and after a guttural one. Elsewhere /a, u, i/ are pronounced [æ, u, i] (Heath 1989:19). There are also two short vowels, /ə/ and /ʷ/, whose structural status is not always clear. /ə/ can be heard as [ă], [ĕ], or [ĭ], depending on consonantal environment, but these are always quite short, as mentioned by Heath (1989). When it is syllabic or the onset of a syllabic sonorant, /ʷ/ occurs as [ʊ]; when syncope applies, it is heard as a labialized release of a consonant or a labialized transition between two consonants. The /ʷ/ phoneme usually occurs next to a velar or uvular consonant (Heath 1989:19). According to Heath, there is a third short vowel, which is marginal and can be identified as a clear consonantal release between two segments. Since this vowel is not found in the Project CoPho corpus, it is not discussed further in this entry. Moroccan Arabic also has four diphthongs: /aj, aw, ej, ew/.

4. Project CoPho database of French loanwords in Moroccan Arabic

Project CoPho's corpus of French loanwords in Moroccan Arabic comprises 1,127 borrowings which yield 2,682 borrowing forms, collected between 1994 and 1995. The loanwords were culled from a variety of written and oral sources and their pronunciations were then verified with native speakers of the borrowing language. Forms were elicited via picture naming, fill-in-the-blanks, definitions, etc. For more detail on the methodology for Project CoPho's loanword database see Paradis and Prunet (2000) and Paradis and LaCharité (2002). A 'borrowing form' is a borrowing as it was pronounced by one of three native Moroccan Arabic-speaking consultants (one from Rabat, another from Casablanca, and the third from Tangier). When a consultant provided more than one form, each was calculated as an independent form. The borrowing forms include 4,250 cases of malformations (i.e. foreign phonemes or structures), either segmental (3,676 cases) or syllabic (574 cases). For the most part, borrowers – by whom we mean those who introduce and adapt loanwords – adapt them on the basis of phonology, indicating that loanwords are 'Arabicized' by those with a good knowledge of both French and Moroccan Arabic. We refer to all such adaptations as

Table 1. Moroccan Arabic consonant inventory

	labials	coronals		velars	uvulars	pharyngeals	glottals
		+ant	-ant				
stops	b	t [ts] d		k g	q		ʔ
fricatives	f	s z	ʃ ʒ		χ ʁ	ḥ ʕ	h
pharyngealized		t̴ d̴					
		s̴ z̴					
nasals	m	n					
liquids		l r					
pharyngealized		r̴					
glides	w	j					

phonological cases. Of 3,676 malformations, 3,441 (93.6%) are treated phonologically. There are some nonphonological cases, such as missed targets, or cases influenced by nonphonological factors such as analogy. However, cases whose adaptation in Moroccan Arabic is influenced by factors other than phonology are rare (6.9%). Among the segmental malformations handled by phonology, there are 738 cases of ill-formed consonants (the French consonants (/ɲ/, /p/, /ɥ/ and /v/), 2,245 cases of oral vowels (the French vowels /e/, /ɛ/, /y/, /o/, /ɔ/, /ø/, /œ/, and /ə/), and 693 cases of nasal vowels (the French vowels /ɑ̃/, /ɛ̃/, and /ɔ̃/). As shown in Table 2, ill-formed segments are usually adapted (84.3% of the cases) or imported (14.9% of the cases). Deletion is rare in this corpus (0.8% of the cases), as in the Project CoPho database of loanwords more generally (2.9%, i.e. 1,398/ 47,624 cases).

5. ADAPTATIONS OF FRENCH PHONEMES

Detailed statistics, as well as the nature and examples of the adaptations for each ill-formed segment, are provided in Table 3. Some of the adaptations described below are also reported in Driss (1997).

Table 2. Statistics regarding loanword adaptation

Ill-formed segments	Cases	Phonological cases	Adaptations	Importations	Deletions	Nonphonological cases	
						Total	Orthography
Consonants	738	735/738 99.6%	289/735 39.3%	431/735 58.6%	15/735 2.1%	3/738 0.4%	0/3 0%
Oral vowels	2,245	2,015/2,245 89.8%	1,965/2,015 97.5%	36/2,015 1.8%	14/2,015 0.7%	230/2,245 10.2%	0/230 0% 0%/2,245
Nasal vowels	693	691/693 99.7%	648/691 93.8%	43/691 6.2%	0/691 0%	2/693 0.3%	0/2 0% 0%/693
Total	3,676	3,441/3,676 93.6%	2,902/3,441 84.3%	510/3,441 14.9%	29/3,441 0.8%	235/3,676 6.4%	0/235 0% 0%/3,676

Table 3. Adaptations, statistics, and examples

Consonant /ɲ/		ɲ	→	nj	25	52.1%
Loanwords	14	ɲ	→	n	21	43.8%
Forms	48	ɲ	→	lj	2	4.0%
Adaptation cases	48					

Examples of the two main adaptations

Fr. *beignet*	[bɛɲɛ]	'doughnut'	→	MA	[binji]
Fr. *poignée*	[pwaɲe]	'knob'	→	MA	[pwanji]
Fr. *champagne*	[ʃɑ̃paɲ]	'champagne'	→	MA	[ʃɑmpan]
Fr. *peignoir*	[pɛɲwaʀ]	'dressing gown'	→	MA	[pɪnwaʀ]

Consonant /p/		p	→	b (108), bb (18), pb (1)	127	93.4%
Loanwords	154	p	→	t	3	2.2%
Forms	431	p	→	f	2	1.5%
Adaptation cases	136	p	→	k/g	3	2.2%
		p	→	l	1	0.7%

Examples of the main adaptation

Fr. *place*	[plas]	'square'	→	MA	[*blas̴* -a]
Fr. *papa*	[papa]	'dad'	→	MA	[baba]

Table 3 (*cont.*)

Consonant /ɥ/		ɥ	→	w	19	100%
Loanwords	10					
Forms	34					
Adaptation cases	19					

Examples

Fr. *biscuit*	[biskɥi]	'cookie'	→	MA	[biskwi]
Fr. *juillet*	[ʒɥijɛ]	'July'	→	MA	[ʒwiji]

Consonant /v/		v	→	b (61), p (1)	62	72.1%
Loanwords	70	v	→	f	24	27.9%
Forms	202					
Adaptation cases	86					

Examples of the main adaptation

Fr. *bravo*	[bʀavo]	'bravo'	→	MA	[brabʊ]
Fr. *service*	[sɛʀvis]	'service'	→	MA	[sərbis]

Vowel /e/		e	→	i (253), ɪ (134), j (9)	396	93.1%
Loanwords	185	e	→	a	10	2.4%
Forms	508	e	→	ɛ (9), ə (5)	14	3.3%
Adaptation cases	425	e	→	ɛn	2	0.5%
		e	→	u/ʊ (1)	3	0.7%

Examples of the main adaptations

Fr. *béret*	[beʀɛ]	'beret'	→	MA	[biri]
Fr. *casquette*	[kaskɛt]	'cap'	→	MA	[kaskɪŧ]

Vowel /ɛ/		ɛ	→	ɪ (251), i (197)	448	90.3%
Loanwords	187	ɛ	→	ə	33	6.7%
Forms	525	ɛ	→	a	14	2.8%
Adaptation cases	496	ɛ	→	ɛ̃	1	0.2%

Examples of the main adaptations

Fr. *bordel*	[bɔʀdɛl]	'brothel'	→	MA	[bʊrđɪl]
Fr. *briquet*	[brikɛ]	'lighter'	→	MA	[briki]

Vowel /y/		y	→	ʊ (40), u (26), w (3)	69	47.6%
Loanwords	59	y	→	i (54), ɪ (12)	66	45.5%
Forms	157	y	→	ə	7	4.8%
Adaptation cases	145	y	→	a	3	2.1%

Examples of the two main adaptations

Fr. *luxe*	[lyks]	'luxury'	→	MA	[luks]
Fr. *terminus*	[tɛʀminys]	'terminus'	→	MA	[ŧɪrminʊs]
Fr. *buffet*	[byfɛ]	'buffet'	→	MA	[bifi]
Fr. *culotte*	[kylɔt]	'trousers'	→	MA	[kɪlʊŧ]

Vowel /o/		o	→	ʊ (197), u (29), w (2)	228	97%
Loanwords	79	o	→	a	3	1.3%
Forms	227	o	→	ɪ	2	0.85%
Adaptation cases	235	o	→	wa	2	0.85%

Examples of the main adaptation

Fr. *bateau*	[bato]	'boat'	→	MA	[baŧʊ]
Fr. *chômeur*	[ʃomœʀ]	'unemployed'	→	MA	[ʃumur]

Table 3 (*cont.*)

Vowel /ɔ/		ɔ	→	ʊ (391), u (71), w(3)	465	97%
Loanwords	160	ɔ	→	ə	6	1.3%
Forms	450	ɔ	→	ɪ (5), i (3)	8	1.7%
Adaptation cases	479					

Examples of the main adaptation

Fr. *bottes*	[bɔt]	'boots'	→	MA	[bʊŧ]
Fr. *police*	[pɔlis]	'police'	→	MA	[bulis]

Vowel /ϕ/		ϕ	→	ʊ (15), u (7)	22	78.6%
Loanwords	10	ϕ	→	i	5	17.8%
Forms	30	ϕ	→	ə	1	3.6%
Adaptation cases	28					

Examples of the main adaptation

Fr. *deux cheveaux*	[dϕʃ(ə)vo]	'CV (car)'	→	MA	[duʃuvu]
Fr. *pneu*	[pnϕ]	'tire'	→	MA	[bnʊ]

Vowel /œ/		œ	→	ʊ (72), u (35)	107	96.4%
Loanwords	38	œ	→	ɛj	3	2.7%
Forms	113	œ	→	a	1	0.9%
Adaptation cases	111					

Examples of the main adaptation

Fr. *docteur*	[dɔktœʀ]	'doctor'	→	MA	[đʊktʊʀ]
Fr. *meubler*	[mœble]	'to furnish'	→	MA	[mubl-a]

Vowel /ə/		ə	→	ʊ (33), u (7)	40	87%
Loanwords	34	ə	→	i	6	13%
Forms	110					
Adaptation cases	46					

Examples of the main adaptation

Fr. *remise*	[ʀəmiz]	'presentation'	→	MA	[rʊmiz]
Fr. *recette*	[ʀəsɛt]	'recipe'	→	MA	[rʊsɪŧ]

Nasal vowel /ɑ̃/		ɑ̃	→	VN	190	66.7%
Loanwords	105	ɑ̃	→	a	91	31.9%
Forms	294	ɑ̃	→	ar	3	1.1%
Adaptation cases	287	ɑ̃	→	al	1	0.3%

Examples of the main adaptation

Fr. *bandit*	[bɑ̃di]	'gangster'	→	MA	[bɑnđɪ]
Fr. *manteau*	[mɑ̃to]	'coat'	→	MA	[mɑnŧʊ]

Nasal vowel /ɛ̃/		ɛ̃	→	ɛN (97), iN (2)	99	85.3%
Loanwords	43	ɛ̃	→	V	17	14.7%
Forms	122					
Adaptation cases	116					

Examples of the main adaptation

Fr. *blindé*	[blɛ̃de]	'armored'	→	MA	[blɛndi]
Fr. *coussin*	[kusɛ̃]	'cushion'	→	MA	[kusɛn]

Table 3 (*cont.*)

Nasal vowel /ɔ̃/		ɔ̃	→	uN, ʊN	182	74.6%
Loanwords	90	ɔ̃	→	V	62	25.4%
Forms	258					
Adaptation cases	244					

Examples of the main adaptation

Fr. *bidon*	[bidɔ̃]	'can'	→	MA [bidun]
Fr. *trombone*	[tʀɔ̃bɔn]	'trombone'	→	MA [ʈrʊmbʊn]

Examples of importations

/p/	Fr. *pantalon*	[pɑ̃talɔ̃]	'pants'	→	MA [panʈalʊn]
	Fr. *papillon*	[papijɔ̃]	'butterfly'	→	MA [papijʊ]
/v/	Fr. *devise*	[dəviz]	'currency'	→	MA [dʊviz]
	Fr. *rendez-vous*	[rɑ̃devu]	'appointment'	→	MA [randɪvu]
/ɑ̃/	Fr. *flan*	[flɑ̃]	'custard tart'	→	MA [flɑ̃]
	Fr. *transmission*	[trɑ̃smisjɔ̃]	'transmission'	→	MA [ʈrɑ̃smijɔ̃]
/ɔ̃/	Fr. *crevaison*	[kʀəvɛzɔ̃]	'flat'	→	MA [kʀʊvizɔ̃]
	Fr. *gazon*	[gazɔ̃]	'lawn'	→	MA [gazɔ̃]

6. Importations of French phonemes

French phonemes are not always adapted; in nearly 15 percent of the cases in the database, they are imported. The phonemes that are most often imported are the two obstruents /p/ (320/456 cases, 70.2%) and /v/ (111/197 cases, 56.3%) and the nasal vowels /ɑ̃/ (26/313 cases, 8.3%) and /ɔ̃/ (16/261 cases, 6.1%). Examples of borrowing forms containing these often-imported French sounds are given in Table 3.

7. The influence of orthography

Orthography influence refers to an adaptation that is based on a graphemic representation rather than on the spoken form. The influence of orthography is generally scarce in Project CoPho's loanword database, but it is nonexistent in the adaptations of the malformations of the corpus of French loanwords in Moroccan Arabic. It occurs outside malformations, but it is rare, affecting only 20 out of 2,682 forms (0.7%; e.g. French *casino* [kazino] > Moroccan Arabic [kasino], French *cornet* [kɔrnɛ] > Moroccan Arabic [kʊrnɪta]; jeune [ʒœn] > Moroccan Arabic [jœn]). This type of orthography influence is obviously based on the graphophonemic correspondence rules of French.

8. The syllabic structure of Arabic and the syllabic adaptations

Like French, Moroccan Arabic allows the following syllables: CV, CVC, CCVC, CVCC. The syllables *V, *CV:CC, *CVCCC, *CCVCC are disallowed. In other words, a complex (branching) coda is permitted only if there are no other consonant clusters within the syllable.

Moroccan Arabic also imposes restrictions on the content of branching codas. Thus, a French binary coda can be ill-formed in Moroccan Arabic even though the two consonants included are each permitted separately. This is the case of the following codas of French: /bl, bʀ, dn, dʀ, fl, gl, gm, gʀ, kl, km, kʀ, ks, kt, lg, lm, ls, mn, ʀg, ʀk, ʀl, ʀm, sm, st, tm, tʀ/. Despite being ill-formed in Moroccan Arabic, these codas are nonetheless often imported (52.7%), as shown in Table 4. Otherwise, they undergo vowel insertion (i.e. adaptation via epenthesis, 20.7%) or deletion of one of the consonants (26.6%).

Examples of insertions, deletions, and importations are provided in Table 4. As shown, the epenthetic vowel is the short vowel /ə/.

In fact, most deletion cases might be importations, since the final consonant is often deleted in some codas in casual speech in French. This is indicated by the parentheses around these final consonants in the examples above. Thus,

Table 4. Statistics on ill-formed binary codas

Cases	Phonological cases	Adaptations (insertions)	Nonadaptations (importations)	Deletions	Nonphonological cases		
					Total	Missed targets	Morphological influence
209	150/209 (71.8%)	31/150 (20.7%)	79/150 (52.7%)	40/150 (26.6%)	59/209 (28.2%)	6/59 (10.2%)	53/59 (89.8%)

Examples of insertions in French branching codas

/dʀ/	Fr. *cadre*	[kɑd(ʀ)]	'frame'	→	MA [kađər]
/kʀ/	Fr. *chancre*	[ʃɑ̃k(ʀ)]	'canker'	→	MA [ʃɑnkər]
/bl/	Fr. *immeuble*	[imœb(l)]	'building'	→	MA [mubəl]
/tʀ/	Fr. *mètre*	[mɛt(ʀ)]	'meter'	→	MA [mitər]
/bʀ/	Fr. *timbre*	[tɛ̃b(ʀ)]	'stamp'	→	MA [tsənbər]

Examples of deletions in French branching codas

/tʀ/	Fr. *arbitre*	[aʀbit(ʀ)]	'referee'	→	MA [larbɪŧ]
/kt/	Fr. *contact*	[kɔ̃tak(t)]	'contact'	→	MA [kʊnŧak]
/st/	Fr. *cycliste*	[siklis(t)]	'cyclist'	→	MA [siklis]
/fl/	Fr. *rafle*	[ʀaf(l)]	'raid'	→	MA [laraf]
/ʀk/	Fr. *remorque*	[ʀəmɔʀk]	'trailer'	→	MA [rmʊk]
/zm/	Fr. *rhumatisme*	[ʀymatizm]	'rheumatism'	→	MA [rʊmatiz]
/st/	Fr. *touriste*	[tuʀis(t)]	'tourist'	→	MA [tsuris]

Examples of importations of French branching codas

/tʀ/	Fr. *mètre*	[mɛt(ʀ)]	'meter'	→	MA [mitr]
/kt/	Fr. *acte (de mariage)*	[ak(t)]	'act (of marriage)'	→	MA [lakt]
/ʀm/	Fr. *alarme*	[alaʀm]	'alarm'	→	MA [lalarm]
/ks/	Fr. *boxe*	[bɔks]	'boxing'	→	MA [buks]
/bl/	Fr. *câble*	[kɑb(l)]	'cable'	→	MA [kɑbl]
/dʀ/	Fr. *cadre*	[kɑd(ʀ)]	'frame'	→	MA [kađr]
/ʀk/	Fr. *cirque*	[siʀk]	'circus'	→	MA [sirk]

Moroccan borrowers who, like all borrowers, are bilingual, are often likely to import the casual pronunciation of French. If this is the case, the true rate of segment deletions in branching codas of French loanwords is, in fact, much lower than 26.6%.

Regarding structural restrictions on syllables, there are three types of possible syllabic malformations in French borrowings: ternary codas, hiatus, and onsetless syllables at the beginning of words. There are only twelve cases of ternary codas in the database. Eleven are deletion cases from two borrowings that could actually be importations of French casual pronunciations: French *orchestre* [ɔʀkɛs(t(ʀ))] > Moroccan Arabic [lʊrkis(t)] and *perdre* [pɛʀd(ʀ)] > Moroccan Arabic [bərd-a]. As for hiatus, statistics are provided in Table 5.

As shown, hiatus usually submits to adaptation (70.5%), and less often to deletion (29.5%). Adaptation can consist of epenthesis (of a glide) or substitution. Examples of adaptation through epenthesis are provided in Table 5.

Insertion here is, in fact, the propagation of the articulator of one of the vowels to the empty onset in the hiatus, which results in a glide, either /w/ or /j/. As for substitution, it consists in realizing one of the two vowels of the hiatus as a glide (devocalization), as can also be seen in Table 5.

Examples of vowel deletion in a situation of French hiatus are shown right after. As indicated, vowel deletion occurs more often between words than word internally.

By disallowing onsetless syllables (*V(C)), Moroccan Arabic also prohibits words that begin with a vowel. Statistics on the number of adaptations, nonadaptations, and deletions of vowels at the beginning of words are provided in Table 6.

Adaptation consists in inserting a consonant in the empty onset, usually the French definite article, or the glottal stop, when the borrowing is followed by the Moroccan Arabic indefinite article *(wahda/wεhid)*, as shown in the examples after Table 6.

Examples of vowel deletion at the beginning of French borrowings are provided right after.

Deletion of initial vowels is proportional to the number of syllables within the word. In other words, as shown in Table 7, the longer the word, the greater the likelihood of initial vowel deletion.

This indicates that there is a metrical constraint at work, with a clear preference for words under three syllables. This hypothesis is supported by the fact that deletion of a syllable in other positions than word-initially is also correlated with the number of syllables included in the borrowing. Again, the longer the word, the more syllable deletions there are (Table 8).

Examples of such deletions are provided after Table 8.

Table 5. Statistics on the adaptation of French hiatus (*VV)

Cases	Phonological cases	Adaptations	Nonadaptations (importations)	Deletions
44	44/44 (100%)	31/44 (70.5%)	0/44 (0%)	13/44 (29.5%)

Examples of adaptation through insertion in French hiatus (VV)

/au/	Fr. *caoutchouc*	[kautʃu]	'rubber'	→	MA [kεwεtʃu]
/ea/	Fr. *clé (à molette)*	[kleamɔlεt]	'wrench'	→	MA [klijamʊnɪŧ]
/ai/	Fr. *mosaïque*	[mɔzaik]	'mosaic'	→	MA [mʊzajik]
/ɔε/	Fr. *Noël*	[nɔεl]	'Christmas'	→	MA [nʊwɪl]
/eo/	Fr. *video*	[video]	'video'	→	MA [vidijʊ]

Examples of adaptation through substitution in French hiatus (VV)

/ea/	Fr. *baccalauréat*	[bakalɔʀea]	'high school diploma'	→	MA [bakalʊrja]
/ɔε/	Fr. *Citroën*	[sitʀɔεn]	'Citroën'	→	MA [ꞩɪŧərwɪn]
/eɑ̃/	Fr. *fénéant*	[feneɑ̃]	'lazy person'	→	MA [fənjεn]
/eε/	Fr. *CTM*	[seteεm]	'CTM'	→	MA [sεtjεm]
/eo/	Fr. *video*	[video]	'video'	→	MA [vidjʊ]

Examples of vowel deletions in French hiatus (VV)

/aε /	Fr. *chambre à air*	[ʃɑ̃braεʀ]	'inner tube'	→	MA [ʃɑmbrɪr]
/ɑ̃a/	Fr. *ciment armé*	[simɑ̃arme]	'reinforced concrete'	→	MA [simɑrmi]
/ea/	Fr. *clé (à molette)*	[kleamɔlεt]	'wrench'	→	MA [klamʊnɪŧ]
/eε/	Fr. *dmc*	[deεmse]	'dmc'	→	MA [dimsi]

Table 6. Statistics on the adaptation of French word-initial vowels (*#V)

Cases	Phonological cases	Adaptations	Nonadaptations (importations)	Deletions
309	309/309 (100%)	159/309 (51.4%)	8/309 (2.6%)	142/309(46%)

Examples of adaptation through insertion in French before word-initial vowels

/#a/	Fr. *adresse*	[adʀɛs]	'address'	→	MA [lɛdris-a], [ʔadris]
/#ɛ/	Fr. *essence*	[ɛsɑ̃s]	'gas'	→	MA [lɪs̵ans], [ʔɪsans]
/#ɛ̃/	Fr. *internat*	[ɛ̃tɛʀna]	'boarding school'	→	MA [lɛnŧɪrna], [ʔɛntɪr̵na]
/#e/	Fr. *étage*	[etaʒ]	'floor'	→	MA [lɪŧaʒ], [ʔɪŧaʒ]
/#o/	Fr. *hôtel*	[otɛl]	'hotel'	→	MA [lʊŧɪl], [ʔʊŧɪƚ]
/#ɔ/	Fr. *omelette*	[ɔmlɛt]	'omelette'	→	MA [lʊmlɪŧ], [ʔʊmlɪŧ]
/#y/	Fr. *urgence*	[yʀʒɑ̃s]	'emergency'	→	MA [lʊrʒans̵], [ʔɪrʒans]

Examples of vowel deletion at the beginning of words

/#a/	Fr. *accélérateur*	[akseleratœʀ]	'accelerator'	→	MA [ksiratsur]
/#ɑ̃/	Fr. *ampoule*	[ɑ̃pul]	'bulb'	→	MA [bʊla]
/#e/	Fr. *écurie*	[ekyri]	'stable'	→	MA [kuri]
/#ɛ̃/	Fr. *infirmier*	[ɛ̃firmje]	'nurse'	→	MA [fərmli]
/#y/	Fr. *humidité*	[ymidite]	'humidity'	→	MA [miditsi]
/#i/	Fr. *immeuble*	[imœbl]	'building'	→	MA [mubəl]

Table 7. Word-initial syllable deletions correlated with the number of syllables

1 syllable	2 syllables	3 syllables	4 syllables	5 syllables
0/2	26/103	83/160	29/38	7/9
0%	25.3%	51.9%	76.3%	77.8%

Table 8. Syllable deletions in other word positions than word-initially, correlated with the number of syllables included in the borrowing

Number of syllables	Number of borrowings with	Number of forms with	Number of syllable deletions in
2 syllables	485	1,394	26/1,394 (1.9%)
3 syllables	219	610	54/610 (8.9%)
4 syllables	37	102	24/102 (23.5%)
5 syllables	5	14	7/14 (50%)
6 syllables	1	1	1/1 (100%)
Total	747	2,121	112/2,121 (5.3%)

Examples of syllable deletions in other word positions than word-initially correlated with the number of syllables included in the borrowing

3 syllables	Fr. *millionnaire*	[miljɔnɛʀ]	'millionaire'	→	MA [mlɛjni]
	Fr. *numéro*	[nymero]	'number'	→	MA [nəmra]
4 syllables	Fr. *électricien*	[elɛktrisjɛ̃]	'electrician'	→	MA [trisjɛn]
5 syllables	Fr. *accélérateur*	[akseleratœʀ]	'accelerator'	→	MA [ksiratsur]
	Fr. *électricité*	[elɛktrisite]	'electricity'	→	MA [trisintsi]
6 syllables	Fr. *municipalité*	[mynisipalite]	'town'	→	MA [manisippʊ]

9. CONCLUSION

This entry focuses on the phonological adaptation of French loanwords in Moroccan Arabic. The picture that emerges from this study is much the same as that for the several other large corpora that have been studied by Project CoPho. In sum, it is generally true that loanwords are borrowed and adapted by bilinguals. The adaptations are mostly phonological in nature; they are rarely due to an inability to 'hear the word properly' or to knowing the word only in written form. The general findings presented here are true not only of Moroccan Arabic; they are true of loanword adaptation generally. Thus, the Moroccan Arabic treatment of French loanwords can be taken as being representative of the treatment of French borrowing in Arabic.

BIBLIOGRAPHICAL REFERENCES

Bentahila, Abdelâli. 1983. *Language attitudes among Arabic-French bilinguals in Morocco.* Clevedon: Multilingual Matters.

Benzakour, Fouzia, Driss Gaadi, and Ambroise Queffélec. 2000. *Le français au Maroc: Lexique et contacts de langues.* Brussels: De Boeck and Larcier.

Driss, Rhaid. 1997. "L'emprunt linguistique dans le discours d'enseignants marocains de français au Maroc." *Cahiers Linguistique d'Ottawa* 1–30.

Heath, Jeffrey. 1989. *From code-switching to borrowing: Foreign and diglossic mixing in Moroccan Arabic.* London: Kegan Paul International.

Marley, Dawn. 2002. "Diversity and uniformity: Linguistic fact and fiction in Morocco." *French in and out of France* ed. Kamal Salhi, 335–376. Oxford: P. Lang.

——. 2005. "Official and unofficial attitudes towards 'own' and 'other' languages in Morocco." *Standard variations and language ideologies in different language cultures around the world*, ed. Rudolf Muhr, 183–202. Frankfurt: P. Lang.

Paradis, Carole and Darlene LaCharité. 2001. "Guttural deletion in loanwords." *Phonology* 18: 2.255–300.

Paradis, Carole and Jean-François Prunet. 2000. "Nasal vowels as two segments: Evidence from borrowings." *Language* 76: 324–357.

CAROLE PARADIS AND DARLENE LACHARITÉ
(Laval University)

Fronting → Vowel Fronting

Fronting

Fronting (*taqdīm*) is "an informal term to denote a movement operation by which a word or phrase is moved to the front of some phrase or clause" (Radford 1997:261). The fronting process has been given several terms in the literature, such as → 'topicalization' and → 'focus'. This entry investigates the syntactic notion 'fronting' in Arabic syntax within two frameworks: the Arabic grammatical tradition, represented partly by al-Jurjānī (d. 471/1078), and Chomsky's Minimalist Program. Section 1 deals with fronting in declarative sentences; Section 2 explores the strength of functional categories used in negation and yes/no questions in relation to fronting.

1. FRONTING IN DECLARATIVE SENTENCES

In the linguistic literature, fronting is sometimes analyzed as a syntactic process by which the speaker attempts to draw the attention of the addressee to the significance of the fronted element:

(1) a. *ḍarab-tu* ***zayd-an***
hit-I Zayd-Acc
'I hit Zayd'
b. ***zayd-an*** *ḍarabt-u*
Zayd-Acc hit-I
'Zayd, I hit'

(2) a. *xaraja* *l-walad-u* ***maʿa ʾabī-hi***
left the-boy-Nom with father-his
'The boy left with his father'
b. ***maʿa ʾabī-hi*** *xaraja* *l-walad-u*
with father-his left the-boy-Nom
'With his father the boy left'

(3) a. *jāʾa-t* *al-bint-u* ***ḍāḥikat-an***
came the-girl-Nom laughing-Acc
'The girl came laughing'
b. ***ḍāḥikat-an*** *jāʾat* *al-bint-u*
laughing-Acc came the-girl-Nom
'Laughing came the girl'

The (a) sentences in (1)–(3) represent the unmarked word order as assumed by the Arabic grammatical tradition. The (b) sentences begin with the focused element (in bold) that undergoes the fronting (movement) process. The fronted constituent functions as a direct object in (1b), as a prepositional phrase in (2b), and as a → *ḥāl* (secondary predicate) in (3b). The *ḥāl* is a secondary predicate in that it has no tense (a type of small-clause predicate); its subject (external → theta-role) must be in the theta grid of the primary predicate, as illustrated in the following paradigm:

(i) *jā'at* *al-bint-u* *ḍāḥika-t-an*
came the-girl-Nom laughing-Acc
'The girl, laughing, came'

(ii) *jā'at* *al-bintu* *wa-hiya*
came the-girl-Nom and-she
ḍāḥikat-un
laughing-Nom

The *ḥāl ḍāḥikatan* in (i) is assigned the accusative case by the preceding verb, and its subject *al-bint* (called *ṣāḥib al-ḥāl*) is the subject of the main verb as well; in (ii), however, the *ḥāl* functions as an embedded predicate in a nominal sentence preceded by the complementizer *wa-* (called *wāw al-ḥāl*).

In all three cases the focused element retains its grammatical function as represented by the case marker it carries. In generativist terms, the (b) sentences are derived from the (a) sentences via a movement rule, which simply moves the focused element to sentence-initial position. A different analysis, however, is proposed by the Arabic grammatical tradition when the fronted element changes its case, as illustrated by the following example (Jurjānī, *Dalā'il* 107):

(4) *zayd-un*$_{i}$ *ḍarab-tu-hu*$_{i}$
Zayd$_{i}$-Nom hit-I-him$_{i}$
'Zayd, I hit'

In (4) the fronted object is co-indexed with a co-referential resumptive pronoun (*ḍamīr ʿā'id*; → resumption). *zayd*, the thematic object of the verb 'I hit', is not a fronted object in (4), as it does not carry the accusative case. It rather functions as a *mubtada'* 'topic' since it carries the nominative case assigned to it by 'initiation' (→ *ibtidā'*). The rest of the sentence (i.e. the verbal sentence) forms its predicate (→ *xabar*), as clarified by the following diagram:

(5) $[_{\text{S}}$ $[_{\text{Topic}}$ zayd-un$_{\text{i}}$] $[_{\text{Comment}}$ $[_{\text{Verbal S}}$ ḍarab-tu-hu$_{\text{i}}$]]]

In Chomsky's framework, (5) involves no movement as it is base generated.

1.1 *Fronting in non-initial positions*

Assuming that VSO is the unmarked word order in Arabic, al-Jurjānī (Lāšīn 1980:141) considers fronting the object to a position immediately following the verb to be a fronting process. Consider the following sentences:

(6) a. *qatala* *zayd-un* *al-xārijiyy-a*
killed Zayd-Nom the-Kharijite-Acc
'Zayd killed the Kharijite'
b. *qatala* *al-xarijiy-a* *zayd-un*
killed the-Kharijite-Acc Zayd-Nom
'The Kharijite, Zayd killed'

According to al-Jurjānī's analysis, (6b) is derived from (6a) by moving the object to a position immediately following the verb.

1.2 *Fronting and indefiniteness*

The syntax of Arabic has a general constraint according to which an indefinite subject cannot occur in sentence-initial position.

(7) a. *jā'a-nī* *rajul-un*
came-me man-Nom
b. **rajul-un* *jā'a-nī*
man-Nom came-me
'A man came to me'

According to al-Jurjānī, fronting of the subject in (7b) is blocked due to the indefiniteness constraint. This constraint is also observed in nominal (verbless) sentences as illustrated by the following contrast:

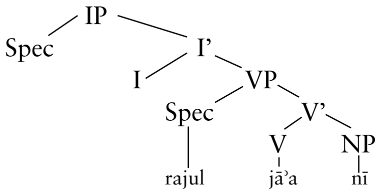
(8) a. **rajul-un* *fī* *l-bayt-i*
man-Nom in the-house-Gen
b. *fī* *l-bayt-i* *rajul-un*
in the-house-Gen man-Nom
'A man is in the house'

The indefinite subject in (8a) is obligatorily postposed to a position after the predicate (*mubtada' mu'axxar*) in (8b). Here, the Arabic grammatical tradition provides two distinct analyses to account for the indefiniteness constraint: a fronting analysis for verbal sentences as in (7) and a postposing analysis for nominal sentences as in (8). The following section provides a unifying analysis that captures the syntactic behavior of the indefiniteness constraint along the lines of Chomsky's theory.

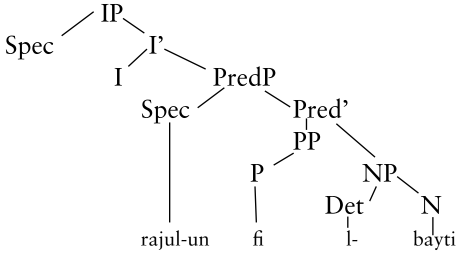
1.3 *Predicate raising*

A unified analysis can be provided to account for the indefiniteness constraint observed in both nominal and verbal sentences in Arabic. The generalization is that an indefinite subject cannot occur initially regardless of its predicate, be it verbal or nominal. To capture this generalization, a verb-raising analysis (Mohammad 1989; Ouhalla 1999) can be extended to embrace nominal (nonverbal) predicates as well. According to this analysis, it is assumed that Arabic is an SVO language in the underlying structure and that the verb is raised to I(nfl) to get the VSO word order. To capture the indefiniteness constraint, it is assumed that the predicate phrase is obligatorily raised to a place outside the sentence (IP) via Chomsky adjunction, if the subject is indefinite. Thus, (7a) and (8b) will have the corresponding underlying trees in (9) and (10) respectively:

(9)

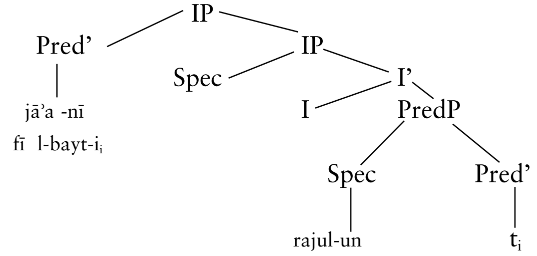


(10)



Predicate Raising will result in the following unified surface structure:

(11)



jāʾa-nī rajul-un
'A man came to me'

fī l-bayt-i rajul-un
'A man is in the house'

The tree in (11) captures the generalization that the predicate phrase, verbal or nonverbal, is obligatorily raised to a sentence-initial position if the subject is indefinite.

2. FRONTING IN NONDECLARATIVE SENTENCES

Al-Jurjānī was the first grammarian in the Arabic linguistic tradition to claim that fronting is not a mere stylistic operation limited to declarative sentences. He investigated two syntactic constructions that also involve fronting: declaratives with negative particle *mā* and yes/no question constructions with the particle *ʾa-*. He claims that fronting triggered by these functional particles does affect the semantics of the sentence. The particles *mā* and *ʾa-* have two basic properties in common. First, both act as complementizers since they cannot occur inside the sentence (IP). Second, both trigger fronting, i.e., they have the property of hosting other arguments from inside the sentence they head.

2.1 *Fronting and negation*

Unlike other negators (*lam* 'did not', *lan* 'will not', and *lā* 'do not', which only precede the verb), the negative operator *mā* behaves as a complementizer since it cannot occur inside the sentence (IP) it heads, i.e., it always occurs in sentence-initial position (→ negation):

(12) a. *ḍarab-tu* *zayd-an*
hit-I Zayd-Acc
'I hit Zayd'
b. *mā ḍarab-tu* *zayd-an*
no hit-I Zayd-Acc
'I did not hit Zayd'

According to al-Jurjāni, (12b) involves no fronting because the negator *mā* is followed by the unmarked word order VSO. Accordingly, the entire sentence in (9b) is negated. However, if the subject is fronted to a position immediately after the negator *mā*, a totally different reading results:

(13) *mā* *ʾana* *ḍarab-tu* *zayd-an*
not I hit Zayd-Acc
'It was not I who hit Zayd'

The negator *mā* in (13) negates the subject only while the rest of the sentence (the verb + the object) is affirmed. The sentence in (13) actually means 'someone hit Zayd, but it was not I who did it'. Al-Jurjānī observes that, unlike stylistic fronting in declarative sentences, fronting in negated sentences does alter the meaning of the sentence. Fronting the object in (12b) above, for example, yields different results:

(14) *mā* *zayd-an* *ḍarab-tu*
not Zayd-Acc hit-I
'It was not Zayd that I hit'

If the object *zayd* is fronted to a position immediately after the negator *mā*, as in (14), only the object is negated; the rest of the sentence (the subject + the verb) remains affirmed. The general meaning in (14) is 'I hit someone, but it was not Zayd'.

Al-Jurjānī assumes that VSO is the unmarked word order of Arabic and, as a result, the fronting of the subject or the object to a position immediately after the negator *mā* changes the meaning of the sentence. The question particle *ʾa-* exhibits a similar behavior, which is discussed in Section 2.2.

2.2 *Fronting and yes/no questions in Arabic*

One strategy for forming yes/no questions in Arabic is by prefixing the particle *ʾa-* to the initial constituent in the sentence, as illustrated in (15):

(15) *ʾa-ḍaraba* *zayd-un* *hind-a*
QP-hit Zayd-Nom Hind-Acc
'Did Zayd hit Hind?'

The Question Particle (QP) is followed by the unmarked VSO word order. The answer is 'yes' or 'no' followed by VSO:

(16) a. *naʿam* *ḍaraba* *zayd-un* *hind-an*
yes hit Zayd-Nom Hind-Acc
'Yes, Zayd hit Hind'

b. *lā* *lam* *yaḍrib* *zayd-un* *hind-an*
no did-not hit Zayd-Nom Hind-Acc
'No, Zayd did not hit Hind'

The subject and the object are fronted in (17) and (18), respectively:

(17) *ʾa-zayd-un* *ḍaraba* *hind-an*
QP-Zayd-Nom hit Hind-Acc
'Is it Zayd who hit Hind?'

(18) *ʾa-hind-an* *ḍaraba* *zayd-un*
QP-Hind-Acc hit Zayd-Nom
'Is it Hind whom Zayd hit?'

The scope of the QP is not the entire sentence but rather the fronted element only, as was the case in negated constructions in (13) and (14) above. The speaker in (17) does not question the act of hitting but rather asks a question about its agent. The difference between (15) and (17) lies in the target of the QP. In (15) the act of hitting can be denied or affirmed, as illustrated in (16). In (17) the act of hitting is affirmed, while the target of the QP concerns the fronted subject, i.e. the doer of the action. The answer either affirms or negates the identity of the subject, which could be Zayd, or anyone else for that matter. In (18) the object is the target of the QP as both the subject and the verb are affirmed. The speaker in this case would be asking about the recipient of the action.

Thus, al-Jurjānī makes a distinction between fronting in declarative sentences and fronting in nondeclarative sentences. Fronting in the former expresses stylistic variation in relation to the constituent being emphasized in a sentence. In nondeclarative sentences, however, fronting changes the meaning of the sentence due to the presence of such functional categories as *mā* and *ʾa-*, discussed in the following section in an attempt to provide a unified analysis based on the notions 'scope' and 'c-command' (May 1985).

2.3 *Functional categories as operators*

Al-Jurjānī's insight into the syntactic as well as the semantic behavior of fronting in the presence of the negative particle *mā* and the yes/no question particle *ʾa-* can be captured by the notions of 'scope' and 'c-command'. These notions are structurally based and defined as follows:

(19)
Scope:
The scope of **a** is the set of nodes that **a** c-commands.

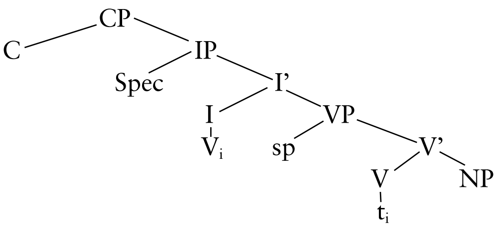
The notion 'c-command' is defined as follows:

(20)
a c-commands **b** iff
(i) The first branching node dominating **a** also dominates **b**
(ii) **a** does not dominate **b**

The particles *mā* and *ʾa-* function as operators whose meaning is determined by their scope, i.e. the domain they c-command. These operators originate in Comp and exercise *wide* scope over the entire sentence (IP) when no fronting is involved, as illustrated by the examples in (21) and their corresponding tree structure in (22):

(21) a. *mā* *ḍaraba* *zayd-un* *hind-an*
not hit Zayd-Nom Hind-Acc
'Zayd did not hit Hind'
b. *ʾa-ḍaraba* *zayd-un* *hind-an*
QP-hit Zayd-Nom Hind-Acc
'Did Zayd hit Hind?'

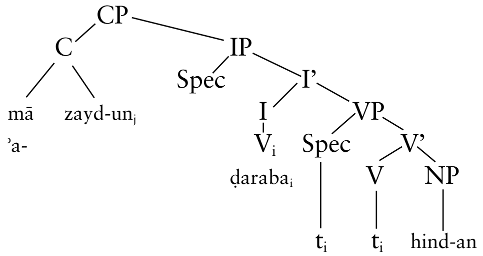
(22)



mā/ʾa *ḍaraba* *zayd-un* *hind-an*
not/QP hit Zayd-Nom Hind-Acc
'Zayd did not hit Hind'
'Did Zayd hit Hind?'

The operator (the negator *mā* or the question particle *ʾa-*) exercizes *wide* scope over the IP since there is no fronting. When fronting takes place, the fronted element moves to Comp where it lies within the *narrow* scope of the operator. The tree in (23) illustrates the structure where the subject is fronted:

(23)



mā zayd-un ḍaraba hind-an
'It was not Zayd who hit Hind'

ʾa-zayd-un ḍaraba hind-an
'Was it Zayd who hit Hind?'

The fronted subject *zayd* in (22) is c-commanded by the operator and is said to lie within its *narrow* scope. The IP lies outside the scope of the operator since it is not c-commanded by it as per the definition in (20) above. The same analysis is obtained when the object is fronted to give the readings in (24):

(24) a. *mā* *hind-an* *ḍaraba* *zayd-un*
not Hind-Acc hit Zayd-Nom
'It was not Hind that Zayd hit'
b. *ʾa-hind-an* *ḍaraba* *zayd-un*
QP-Hind-Acc hit Zayd-Nom
'Was it Hind that Zayd hit?'

Thus, the structural notions of scope and c-command uniformly capture the syntactic and semantic behavior of the operators *ma* and *ʾa-*.

3. CONCLUSION

Fronting in Arabic syntax has been discussed within two distinct grammatical perspectives: the Arab grammatical tradition, as represented by al-Jurjānī, and the Chomskyan linguistic tradition. Al-Jurjānī made a distinction between fronting in declarative sentences and fronting in nondeclarative sentences. The former is stylistic in nature and has no bearing on the basic meaning of the sentence. The latter alters the meaning of the sentence. In minimalist terms, fronting in declaratives, including the indefiniteness constraint, has been given a unified analysis along the lines of predicate raising. Fronting in nondeclarative sentences has been unified through the notions of 'scope' and 'c-command'.

Bibliographical references

Primary sources

Jurjānī, *Dalā'il* = ʾAbū Bakr ʿAbd al-Qāhir ibn ʿAbd ar-Raḥmān al-Jurjānī, *Dalā'il al-ʾiʿjāz*. Ed. Maḥmūd Muḥammad Šākir. Cairo: al-Hay'a al-Miṣriyya al-ʿĀmma li-l-Kitāb, 2000.

Secondary sources

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

——. 1986. *Barriers*. Cambridge, Mass.: MIT Press.

——. 1995. *The minimalist program*. Cambridge, Mass.: MIT Press.

Lāšīn, ʿAbd al-Fattāḥ. 1980. *at-Tarākīb an-naḥwiyya min al-wijha al-balāġiyya ʿinda ʿAbd al-Qāhir*. Riyadh: Dār al-Marrīx.

May, Robert. 1985. *Logical form*. Cambridge, Mass.: MIT Press.

Mohammad, Mohammad A. 1989. *The sentence structure of Arabic*. Ph.D. diss., University of Southern California, Los Angeles.

Ouhalla, Jamal. 1999. *Transformational grammar: From principles and parameters to minimalism*. 2nd ed. London: Arnold.

Radford, Andrew. 1997. *Syntax: A minimalist program*. Cambridge: Cambridge University Press.

Wafaa ʿAbd al-Fahim Batran Wahba
(Ain Shams University)

Frozen Expression

Frozen expressions are also known as 'set expressions' or 'frozen structures'. They are "[a] group of words standing in a fixed association" (Crystal 2001:304–305). Examples of frozen expressions include a number of structures and genres. They may have general applicability, as do phrasal verbs (such as *daʿā li-* 'to pray for' and *daʿā ʿalā* 'to curse') and other → collocations, or be restricted to particular events and situations, as are certain courtesy expressions (such as the Levantine *yaʿṭīk ilʿāfiya* 'may God give you strength', said to a person who is working; → greetings). They can be prayers (such as the familiar *Fātiḥa*), or curses (such as *yilʿan abūk* 'may God curse your father'; → insults). They are as short as a proverb (such as the Egyptian *ga ykaḥḥalha ʿamāha* 'he came to apply kohl to her eyes and he blinded her'), or as long as a tale or epic (such as *Sīrat ʿAntar*).

The term 'frozen expression' has wide applicability but is not widely accepted as a technical term. It does not occur frequently in linguistic research on Arabic or other languages. Infrequent use means that 'frozen expression' does not appear in most English language dictionaries and encyclopedias of linguistics, nor does it occur in bilingual (English-Arabic) dictionaries of linguistics.

'Frozen expression' is, it appears, less a technical term than a loose description of the feature shared by the genres listed above and others. These genres do not share the paradigmatic nature of ordinary language or free discourse; they do not allow for substitution of elements. An example might be an Egyptian version of a well-known proverb *ištiri ijjār ʾabl iddār* 'buy the neighbor before the house'. However one might feel about the importance of the building's *bawwāb* 'doorkeeper', it is not possible to replace the neighbor with the doorkeeper in this proverb by saying *ištiri ilbawwāb ʾabl iddār* 'buy the doorkeeper before the house'. Rather, it is possible, but this new utterance is not a proverb; it is ordinary speech. The elements of other genres, in much the same way, cannot be replaced. One might say *yaʿṭīk iṣṣiḥḥa* 'may God give you good health' rather than *yaʿṭīk ilʿāfiya* 'may God give you strength'. This is, however, inappropriate in certain circumstances, and at worst might be considered an insult. To change the name of ʾAbū Zayd al-Hilālī, the epic hero, means that one is not reciting the *Sīrat Banī Hilāl* but another epic altogether. This feature, the feature of 'frozenness', has yet to be investigated in depth and as it applies to multiple genres of Arabic frozen expressions.

Published research to date has focused on particular types or genres of frozen expression, such as the → proverb, the curse, and the epic. The frozen expression as a class, however, has not been the subject of much published research. One interesting exception is Youssi's (1994) article on the 'frozen structure', as he calls the frozen expression. He outlines a categorization of frozen expressions (1994:138–139). His framework for analyzing them, however, is not based on the fact that these forms are frozen. Instead, he considers them from the point of view of semiotics (1994:139). This approach may hold promise for shorter genres, such as the proverb. Longer forms, especially narratives, may be resistant to semiotic analysis.

BIBLIOGRAPHICAL REFERENCES

Crystal, David. 2001. *A dictionary of language*. 2nd ed. Chicago: University of Chicago Press.

Youssi, Abderrahim. 1994. “Linguistic and semio-cultural properties of frozen structures.” *Actes des Premières Journées Internationales de Dialectologie Arabe de Paris*, ed. Dominique Caubet and Martine Vanhove, 135–149. Paris: Publications INALCO.

ELIZABETH M. BERGMAN (Georgetown University)

Fulfulde

Fulfulde, a language belonging to the Atlantic branch of the Niger-Congo family, is spoken by approximately 20 million people in West Africa, chiefly in the Sahel region. From Maasina (Mali) and eastward, the name of the language is Fulfulde; west of Maasina it is called Pulaar, except in Fuuta Jaloo (Guinea), where it is called Pular. Compare English *Fulani* (< Hausa *Filā̀nī*, pl. of *Bàfilācḕ* ‘Pullo’) and *Fula* (from a Mande language), French *peul* (< Wolof *pël* ‘Pullo’), and German *Ful* (the root of *Fulfulde, Pullo,* etc.). Speakers of Fulfulde call themselves *Fulɓe* (pl. of *Pullo*); the most common English name is Fulani (sg./pl.).

The Fulɓe, traditionally cattle nomads who started to spread out from Senegal and western Mali early in the 2nd millennium C.E., are predominantly Muslims. Some individuals may have adopted Islam already in their contact with the → Mali empire of the 11th and 12th centuries. The contact with Arabic has primarily been through Qurʾānic schools and Islamic studies. Direct contact with speakers of Arabic is roughly limited to Mauritania, Chad, and the Republic of the Sudan.

Fulfulde has the widest geographical distribution of all African languages south of the Sahara, and dialects often have different loanwords or the same loanwords in different forms. Arabic loanwords in Fulfulde have often come via other languages. There is considerable dialect variation. This entry concentrates on the well-described Adamawa dialect of eastern Nigeria and northern Cameroon, where the Arabic loanwords have come via → Hausa and → Kanuri.

Fulfulde is written in the Arabic and Latin alphabets. The Arabic alphabet has been used for several centuries, and the Latin alphabet was introduced in the late 19th century. In Adamawa, the Arabic alphabet is better known than the Latin alphabet.

1. PHONOLOGY

Fulfulde phonemes are presented in Table 1, orthographically. The symbols are self-explanatory, except for *c* /tʃ/, *j* /dʒ/, *ɓ* /b̰/ and *ɗ* /d̰/, *ƴ* (creaky voiced palatal semivowel), and *ʾ* /ʔ/.

All phonemes except those indicated by *w*, *y*, *h*, *f*, *mb*, *nd*, *nj*, *ng* have distinctive quantity oppositions (*s* in loanwords only), expressed orthographically by doubling.

Arabic consonants foreign to Fulfulde are replaced by native ones (see Table 2, where the names of the Arabic letters representing these consonants are also included, in Arabic and Fulfulde forms). Examples of words with these sounds are presented in Table 3. The nominative ending *-u* is added in parentheses to an Arabic noun when the ending is borrowed into Fulfulde. The consonant *z* occurs only in learned pronunciations of some loanwords, and is indicated in parentheses.

When an Arabic consonant has several Fulfulde representations, the loanwords in which it occurs may have been borrowed via different languages. Arabic *ḍ* becomes *d, l*, and *b* in Fulfulde; *b* is found, among other places, in *baadi*, the Fulfulde name of the Arabic letter

Table 1. The phonemes of Fulfulde

p	t	c		k	ʾ	i		u
b	d	j		g		e		o
ɓ	ɗ		ƴ				a	
mb	nd	nj		ng				
f	s				h			
w	l		y					
	r							

Table 2. Arabic consonants not occurring in Fulfulde and their replacements

Arabic		Fulfulde	
Letter names	Sounds	Sounds	Letter names
ṯāʾ	*ṯ*	*s, t*	*samablu, samamlu*
ḥāʾ	*ḥ*	*h*	*haa baalol* 'haa without dot'
xāʾ	*x*	*h*	*haa toɓɓungol* 'dotted haa'
ḏāl	*ḏ*	*j (z)*	*zaali*
zāy	*z*	*j (z)*	*zaayra*
šīn	*š*	*s*	*siini*
ṣād	*ṣ*	*s*	*saadi*
ḍād	*ḍ*	*d, l, b*	*baadi*
ṭāʾ	*ṭ*	*t, ɗ*	*ɗaadi*
ḍāʾ	*ḍ*	*j (z)*	*zadi*
ʿayn	*ʿ*	*ʾ, h*	*ayni*
ġayn	*ġ*	*ng*	*ŋiini, angiini*
qāf	*q*	*k, g*	*gaafu*

Table 3. Fulfulde nouns borrowed from Arabic, illustrating consonant replacements

	Arabic		Fulfulde	
ṯ > s	*ṯaman*	'price, cost, value'	*saman*	id.
ṯ > t	*al-iṯnayn*	'Monday'	*altine*	id.
ḥ > h *ṯ > s*	*ḥadīṯ*	'hadith'	*hadiisewol*	id.
x > h	*xabar(u)*	'news; information; predicate [gram.]'	*habaru*	id.
x > h *z > j*	*xinzīr(u)*	'pig'	*hinjiiru*	id.
ḏ > j	*ḏimmī*	'free non-Muslim living in a Muslim country'	*jimmadunkeejo*	id.
z > j	*zakāt*	'alms tax'	*jakka (zakka)*	id.
š > s *q > k*	*šaqīq*	'full brother'	*sakiikeejo*	'full sibling'
ṣ > s	*ṣawt(u)*	'sound; voice; noise'	*sawtu*	id.
ḍ > d	*ḍamān*	'guaranty'	*dammaana*	id.
ḍ > l	*ḍamīr*	'conscience; personal pronoun [gram.]'	*lamiiri*	id.
ḍ > b	*ramaḍān*	'Ramadan'	*Ramabaana*	id.
ṭ > ɗ *ʿ > ʾ*	*ṭabīʿa*	'nature'	*ɗabiʾa*	id.
ṭ > t *q > k*	*biṭāqa*	'slip of paper, tag; card'	*bataakewol*	'letter; note'
ḍ > j	*ḍuhr*	'midday prayer'	*juura (zuura)*	id.
ʿ > ʾ	*sāʿa*	'time; hour; watch'	*saaʾa*	id.
ʿ > h *q > k*	*ʿaql*	'sense, reason, intelligence'	*hakkiilo*	'attention; intelligence; prudence'
ġ > ng	*maġrib*	'sunset'	*mangariba*	id.
q > g	*bunduqiyya*	'rifle, gun'	*bunndugaaru*	id.

ḍād. According to Klingenheben, “Wörter, in denen Arabic *ḍ…l* gesprochen wird, sind über das H[ausa] ins Ful gekommen” (1963:1). But a word like *alkaali(ijo)* ‘judge’ < Arabic *al-qāḍī* occurs in most Fulfulde dialects, including those spoken outside the area of Hausa influence. Arabic *ḍ* is apparently found as *d* as well as *l* in most or all dialects of Fulfulde.

Phonological adaptations also occur in consonant clusters foreign to Fulfulde (see Table 4). A vowel splits the cluster, the first consonant is changed, or the first consonant (a laryngeal or pharyngeal consonant) disappears and lengthens the preceding vowel. Often, a vowel (whose quality varies according to principles that are not well understood) is also added word-finally.

Vowels usually do not change. However, in some words Fulfulde *e* corresponds to Arabic *a*; cf. Fulfulde *deftere* ‘book’ < Arabic *daftar* ‘notebook’. This word, found in all dialects in the same form (except Maasina *dewtere*, due to a recent sound change), may be among the oldest Arabic loanwords in Fulfulde; the vocalism may indicate oral borrowing.

The initial consonant of Arabic may change in Fulfulde consonant alternations; e.g. *faama* [sg. subject] ~ *paama* [pl. subject] ‘to understand’ < Arabic *fahima/yafhamu*.

In Adamawa, Arabic personal names are borrowed via Hausa and retain Hausa tones, e.g. Fulfulde *Iisaa* (HH) < Hausa *Īsā* < Arabic *ʿĪsā*; Fulfulde *Yuusufu* (HHL) < Hausa *Yūsufù* < Arabic *Yūsuf(u)*; Fulfulde *Umaru* (LHL) < Hausa *Ùmarù* < Arabic *ʿUmar(u)*.

2. Morphology

Morphological adaptation of Arabic loanwords involves loss of Arabic morphology and adoption of Fulfulde morphology. Typically, Arabic verbs are borrowed in the imperfect without

Table 4. Adaptation of Arabic consonant clusters

	Arabic		Fulfulde	
Ø > *i* /C_C	*fajr*	‘dawn; morning prayer’	*fajiri*	‘early morning’
	qurṣ	‘plate, disk, tablet’	*gurus*	‘dollar [esp. the Maria Theresa dollar]’
	waqt	‘time’	*wakkati*	‘time [esp. of appointed time]’
ḍ > *y* / _ C	*(ʿīd) al-ʾaḍḥā*	‘the Feast of Sacrifice’	*layha*	id.
*V*ʿ > *VV* /_C	*fiʿl*	‘verb’	*fiiliwol*	id.

Table 5. Arabic verbs in Fulfulde: Regular

Arabic			Fulfulde	
perfect	imperfect			
zāra	*yazūru*	‘to visit’	*juuroo*	‘to visit returned pilgrim or saint’
tāba	*yatūbu*	‘to repent’	*tuuba*	id.
dāma	*yadūmu*	‘to last’	*duuma*	id.
ṣāma	*yaṣūmu*	‘to fast’	*suumoo*	id.
fassara	*yufassiru*	‘to explain’	*fassira*	‘to explain a text; to translate’
ḥanna	*yaḥinnu*	‘to pity, have mercy’	*hinna*	id.
jarraba	*yujarribu*	‘to test; to try; to put to the test, tempt’	*jarriboo*	id.
māla	*yamīlu*	‘to bend; to bow down’	*miiloo*	‘to bend toward, decline’
darasa	*yadrusu*	‘to learn, study’	*dursa*	‘to know by heart; to recite’
fahima	*yafhamu*	‘to understand’	*faama*	id.
sajada	*yasjudu*	‘to bow down, bow in worship’	*sujida*	id.
nafaʿa	*yanfaʿu*	‘to be of use’	*nafa*	id.

Table 6. Arabic verbs in Fulfulde: Irregular

Arabic			Fulfulde	
perfect	imperfect			
ʿafā	*yaʿfū*	‘to eliminate; to forgive	*yaafoo*	‘to forgive’
ʾajāba	*yujību*	‘to answer’	*jaaboo*	id.
waʿaḏa	*yaʿiḏu*	‘to preach; to admonish’	*waaja*	id.
ṣaḥḥa	*yaṣiḥḥu*	‘to be healthy; to be true, authentic’	*saaha*	‘to be clear, correct, genuine’
šahida	*yašhadu*	‘to witness, be witness’	*seedoo*	id.
talifa	*yatlafu*	‘to perish’	*tilfa*	id.
xalafa	*yaxlufu*	‘to be the successor, follow’	*halfoo*	‘to bring up; to guard’

affixes (cf. the regular cases in Table 5 and the exceptional cases in Table 6). Fulfulde verbs are cited in the active (*-a*) or middle (*-oo*) form of the singular subjunctive.

In imperfect stems with two initial consonants (see the last four examples in Table 5), there is a metathesis from CCVC to CVCC: *drus* > *durs*, *fham* > *fahm*, *sjud* > *sujd*, *nfaʿ* > *nafʿ*; *fahm* is further changed into *faam* (loss of *h* and compensatory lengthening), *sujd* into *sujid* (epenthesis), and *nafʿ* > *naf* (loss of final ʿ).

Arabic verbs in Fulfulde are probably imperfect forms stripped of affixes. This hypothesis is challenged by the fact that in several derived verbs, an Arabic imperfect without affixes is identical to an imperative. Fulfulde *jarriboo* ‘to test, try, etc.’ may come from Arabic imperfect *yujarribu* or imperative *jarrib* ‘id.’. However, affix stripping is required in nonderived verbs even if the imperative is the source; compare Fulfulde *dursa* ‘to know by heart; to recite’ to Arabic imperfect *yadrusu* and imperative *udrus* ‘learn!; study!’. The Arabic imperative could not always be the masculine singular, whose vocalism in hollow verbs differs from that of the imperfect; compare Fulfulde *tuuba* ‘to repent’ to Arabic imperfect *yatūbu* masculine singular imperative *tub* and feminine singular *tūbī*. From a semantic point of view, an imperative is only a likely source in oral borrowing, but in Fulfulde many Arabic loanwords seem to have been borrowed from the written language; in such a situation, a feminine, dual, or plural imperative is a less probable source than a singular masculine.

Morphological adaptation of nouns is primarily the acquisition of one of 20 noun classes (see Table 7). Nouns usually have a class suffix, except for recent loanwords, which lack a suffix and belong to the human O class. Many loanwords get a class suffix on a semantic basis, or the final consonant and/or vowel is reinterpreted as a class suffix.

Some Fulfulde nouns have a final vowel *-u* whose source is probably the nominative suffix of written Arabic, cf. *habaru* ‘news; predicate [gram.]’ (< Arabic *xabar(u)*), where neither the phonology nor the morphology of Fulfulde can explain the presence of *-u*; Fulfulde words may end in *r* (cf. the alternative forms *habar* and *kubar* and imperatives like *war* ‘come!’), and *habaru* belongs to the O class. The analysis is supported by women’s names ending in *-atu*; cf. *Faaɗimatu* < Arabic *Fāṭima(tu)* and *Eysatu* < Arabic *ʿĀʾiša(tu)*. In Fulfulde *hinjiiru* ‘pig’ < Arabic *xinzīr(u)*, the *-u* may also be due to morphological reanalysis creating the *-ru* suffix variant of the NDU class, which is also the class of *gaduuru* ‘warthog, wild pig’ < Hausa *gàdū*. In Fulfulde *sawtu* (O class) ‘sound; voice; noise’ < Arabic *ṣawt(u)*, the *-u* may also be explained as a vowel added because Fulfulde does not accept codas with two consonants.

Many Arabic loanwords have a petrified definite article *al-* ~ *aC-* ~ *l-* (see Table 8). In all dialects of Fulfulde some words are borrowed with the definite article. Further research is required to discover possible diachronic or oral/written differences.

3. SEMANTICS

Arabic loanwords in Fulfulde cover a broad semantic spectrum including *ammaa* ‘but; or’ < Arabic *ʾammā* ‘as for; but’; *jaaboo* ‘to answer’ < Arabic *ʾajāba/yujību; bikriijo* ‘virgin’ < Arabic *bikr;* as well as *sawtu* ‘sound; voice; noise’ < Arabic *ṣawt(u)*. However, in some semantic domains Arabic loanwords are particularly dominant: theology and religion (see Table 9); traditional schools, reading, and writing, including grammar

Table 7. Morphological adaptation of nouns

Arabic		Fulfulde		
ḥāla	‘condition, state; situation; case’	*haala* or *haal-a*	‘talk; discussion, palaver; case’	O class (no suffix) or KA class (*-a*), for phonological reasons
tabūt(u)	‘coffin’	*tabuutu-wal*	id.	NGAL class (*-wal*) of things made of wood or having the shape of a tree trunk
bunduqiyya	‘rifle, gun’	*bunndugaa-ru*	id.	NDU class (*-ru*) of cylindrical objects
tamra	‘date [fruit]’	*tamaroo-re*	id.	NDE class (*-re*) of fruits
ʿinab	‘grape’	*inaboore*	id.	and small spherical things
daftar	‘booklet, notebook’	*deft-e-re* pl. *deft-e*	‘book’	NDE class (*-re*), for phonological reasons
ʾaṣl	‘root; origin’	*asli* pl. *aslii-ji* or *as-ngol* pl. *as-li*	‘root; origin; noun class [gram.]’	O class (no suffix) or NGOL class (*-wol*) of long, thin objects

Table 8. Loanwords with a petrified definite article

Arabic		Fulfulde	
al-ʿāda	‘custom, habit’	*alʾaada*	id.
al-qāḍī	‘judge’	*alkaali(ijo*	id.
al-lawḥa	‘board, slate, tablet’	*alluha*	‘wooden board used for writing, slate’
an-nūr	‘light’	*annoora*	id.
ar-rāʾ	‘the letter *rāʾ*’	*arre/arrewol*	id.
as-samāʾ	‘sky, heaven’	*asama*	id.
at-tājir	‘merchant’	*attaajiriijo*	‘rich and influential merchant’
al-ʾadab(u)	‘culture, refinement’	*ladabu*	‘respect, politeness’
al-ʾimām	‘imam’	*liman(jo)*	id.
al-ʾinjīl	‘gospel’	*linnjiila*	‘the book of the gospels’

Table 9. Theology and religion

Arabic		Fulfulde	
(ʿīd) al-ʾaḍḥā	‘the Feast of Sacrifice’	*layha*	id.
al-ʾimām	‘imam’	*liman(jo)*	id.
baraka	‘blessing, benediction’	*barka*	‘blessing; happiness; affluence’
duʿāʾ	‘prayer’	*doʾa*	id.
ḥadīṯ	‘hadith’	*hadiisewol*	id.
ḥajja/yaḥujju	‘to make the pilgrimage to Mecca’	*hijja*	id.
malʾak, pl. *malāʾika*	‘angel’	*malaaʾikaajo*	id.
rūḥ(u)	‘breath of life, soul; spirit’	*ruuhu*	‘soul’
dīn	‘religion’	*diina*	id.
rizq(u)	‘property, wealth; boon, blessing [of God]’	*risku*	‘prosperity, riches’
at-tawrāt	‘the Pentateuch’	*tawreeta*	id.
al-ʾinjīl	‘gospel’	*linjiila*	‘the book of the gospels’

Table 10. Traditional schools, reading, and writing

Arabic		Fulfulde	
biṭāqa	‘slip of paper, card; ticket; label’	*bataakewol*	‘letter; note’
daftar	‘booklet; notebook’	*deftere*	‘book’
ḥarf	‘letter [of the alphabet]’	*harfeere*	id.
ʾalif	‘alif’	*aliifi*	id.
ṣifa	‘adjective’	*sifa*	id.
mufrad(u)	‘singular’	*mufradu*	id.
darasa/yadrusu	‘to learn, study’	*dursa*	‘to know by heart; to recite’
fassara/yufassiru	‘to explain’	*fassira*	‘to explain a text; to translate’
muʾaddib	‘educator; teacher in Qurʾānic school’	*moodibbo(ojo)*	‘learned man’
muʿallim	‘teacher’	*mallumjo*	‘marabout’
faqīr, pl. *fuqarāʾ*	‘poor man; Sufi mendicant’	*pukaraajo, pl. fukaraaɓe*	‘student, pupil’
taʾrīx	‘history’	*taariiha(awol)*	id.

Table 11. The days of the week and other terms from the temporal domain

Arabic		Fulfulde	
al-ʾaḥad	‘Sunday’	*alad*	id.
al-iṯnayn	‘Monday’	*altine*	id.
aṯ-ṯulaṯāʾ	‘Tuesday’	*salaasa*	id.
al-ʾarbiʿāʾ	‘Wednesday’	*alarba*	id.
al-xamīs	‘Thursday’	*alhamiisa*	id.
al-jumʿa	‘Friday’	*jumʾaare~jumɓaare*	id.
as-sabt	‘Saturday’	*asawe*	id.
waqt	‘time’	*wakkati*	‘time [esp. of appointed time]’
sāʿa	‘time’	*saaʾa*	‘hour; clock, watch’
qarn(u)	‘century’	*karnuwol*	id.

(see Table 10); and time, including the days of the week, but not day, night, and the seasons (see Table 11).

Many Arabic loanwords in Fulfulde belong to a learned style rather than to the colloquial language. *Wakkati* ‘time’, *faama* ‘to understand’, and the names of the days are colloquial, and *karnuwal* ‘century’, *bikriijo* ‘virgin’, and grammatical terminology are learned, while some words, such as *bataakewol* ‘letter’, and many religious terms occupy an intermediary position.

BIBLIOGRAPHICAL REFERENCES

Baldi, Sergio. 1996. “On Arabic loans in Ful”. *Annali del Istituto Orientale di Napoli* 56:3.388-406.

Klingenheben, August. 1927. *Die Laute des Ful.* (= *Beihefte zur Zeitschrift für Eingeborenen-Sprachen*, 9.) Berlin and Hamburg: D. Reimer and C. Boysen.

——. 1963. *Die Sprache der Ful (Dialekt von Adamaua): Grammatik, Texte und Wörterverzeichnis.* Hamburg: J.J. Augustin.

Labatut, Roger. 1983. “Les emprunts du peul à l’arabe”. *Langue arabe et langues africaines*, 41-67. Paris: Conseil International de la Langue Française.

Noye, Dominique. 1989. *Dictionnaire foulfouldé-français: Dialecte Peul du Diamaré, Nord-Cameroun.* Garoua and Paris: Procure des Missions and Librairie Orientaliste Paul Geuthner.

Osborne, Donald W., David J. Dwyer, and Joseph I. Donohoe Jr. 1993. *A Fulfulde (Maasina)-English-French lexicon.* East Lansing: Michigan State University Press.

Taylor, F.W. 1932. *A Fulani-English dictionary.* Oxford: Clarendon Press. (Repr., New York: Hippocrene Books, 1995.)

Zubko, G.B. 1980. *Dictionnaire peul (fula)-russe-français.* Moscow: Langue Russe.

ROLF THEIL (University of Oslo)

Functional Grammar

1. Overview

Functional Grammar as an instantiation of the Functional Paradigm is a pragmatically oriented linguistic theory meant to describe and explain the grammatical organization of natural languages primarily conceptualized as instruments of social interaction. It is commonly opposed to the Formal Paradigm as represented, for example, by Generative Grammar. An excellent account of the theory of Functional Grammar is found in Mackenzie (1995). Functional Grammar was initially proposed by Simon Dik in 1978. It has been further developed by Dik and other collaborators from such countries as the Netherlands, Belgium, England, Spain, and Morocco. Functional Grammar has been applied to the analysis of typologically different languages, including varieties of Arabic (see among others Cuvalay-Haak 1996; Moutaouakil 1984, 1988, 1989, 1993, 1996, 1998, 2000, 2004, 2005), and has been subject to several modifications, extensions, and improvements over the past two decades. Since its introduction in Morocco in the early 1980s, it has been used as a theoretical framework for the description of the pragmatic, semantic, and morphosyntactic aspects of Standard Arabic and some colloquial varieties of Arabic, as well as features of their typology and historical change.

Functional Grammar has evolved through three main stages, based on conceptualization of its general organization: Pre-standard version (Dik 1978), Standard version (Dik 1997a, 1997b), and Post-standard version. This entry takes as a general framework Dik's entire work, the Functional Discourse Grammar model recently proposed by Hengeveld (2004a, 2004b), and the Arabic Functional Grammar literature mentioned above. It focuses on those aspects of Functional Grammar that have been extensively studied in functional studies of Arabic. The main target of Functional Grammar is the description and the explanation of discourse phenomena. But since this has not yet been extensively applied to whole texts, especially Arabic texts, the discussion is restricted to the approach it provides for clause and sentence structures.

2. Clause structure

In Functional Grammar, a clear distinction is made between Clause and Sentence, the latter being understood as a clause with which external constituents are associated. The clause is a hierarchically organized structure consisting of two underlying levels (Interpersonal and Representational), representing the pragmatic and semantic features of linguistic expressions, and a surface (Structural) level where their morphosyntactic and phonological correlates are specified. The two underlying levels contain substructures called 'layers'. Each layer consists of three main components: (i) a Nucleus, which is a verbal, nominal, or adjectival predicate with its arguments; (ii) an Operator triggering morphosyntactic and/or phonological processes; and (iii) some optional Satellites (typically adverbial). It is assumed in Rijkhoff (1992), Dik (1997a), and Moutaouakil (2003) that a certain parallelism holds between the different parts of discourse (Phrase, Predication, Clause, and Text): they all tend to display the same layers, although at different degrees.

The Interpersonal level is meant to represent the relationships between Speaker and Addressee, on one hand, and between Speaker and clause content, on the other. To fulfill this task, the Interpersonal level is provided with an illocutionary layer and a modal layer and with pragmatic functions as well. Three basic clause types are distinguished: declarative, interrogative, and imperative. The following discussion focuses on properties of interrogative clauses.

As an illocution, Interrogation can take in its scope either the whole clause or one of its terms. In Arabic, the interrogative term operator is typically expressed by one of the *man* 'who'-paradigm interrogative pronouns. It can also be expressed by the particle *ʾa*. The difference is due to the type of Focus assigned: requestive in the former case and contrastive in the latter. The clausal interrogative operator surfaces as question word *hal*. Examples (1a)–(1c) illustrate the formal expression of the interrogative term and clause operators.

(1a)	*man*	*ʾanbaʾa*	*man*
	who	informed	whom

'Who has informed whom?'

(1b) *ʾa* ***fahdan*** *tahwā zaynabu (ʾam bakran)*
Q Fahd-Acc loves Zaynab-Nom (or Bakr-Acc)?
'Is it Fahd that Zaynab loves (or Bakr)?'

(1c) *hal qābalta laylā*
Q met-you Laylā
'Did you meet Layla?'

Interrogative constructions can also carry various 'derived' (implicated) illocutions. The derived illocution can be understood only from the context. By uttering (2), for example, one can perform an Offer instead of a real Question. The Offer illocution is rendered possible only by an appropriate situational context:

(2) *hal tašrabu š-šāya*
Q drink-you the-tea-Acc
'Do you drink tea?'

It can also have formal correlates, such as the initial particle *ʾa-wa* and the morpheme *min*, which may only occur in interrogative constructions carrying Disapproval and Denial illocutions, respectively, as illustrated in examples (3) and (4).

(3) *ʾa-wa taštumu ʾabāka*
Q insult-you father-Acc-you
'Do you insult your father?!'

(4) *hal zāra-nī min zamīlin*
Q visited-me of colleague-Gen
'Did any colleague pay a visit to me?!'

The nonliteral illocution not only can determine the occurrence of given particles and morphemes, it can also explain the grammaticality of constructions where two clauses with different illocutions are coordinated, as shown in (5).

(5) *ʾa lam ʾūʿṭika l-māla wa*
Q Neg gave-I-you the-money-Acc and
wahab-tuka d-dāra
bequeathed-I-you the-house-Acc
'Didn't I give you my money and bequeath you my house?!'

What legitimates the coordination in (5) is that the interrogative first clause implies an Assertion, i.e. an illocution compatible with the literal illocution of the second clause.

The derived illocution can lose its marked character over time and become grammaticalized. In such a case, the literal illocution disappears, leaving room for the derived one to become the only illocutionary meaning of the construction. A well-known example is negated interrogative constructions, as in (6), whose actual illocutionary force is a reinforced Assertion.

(6) *ʾa lam ʾuʿir-ka kulla kutub-ī*
Q Neg lent-I-you all-Acc books-me
'Haven't I lent you all my books?!'

The modality layer is meant to account for Speaker's different subjective attitudes toward the content of the clause. Modal features, such as epistemic, volitional, and emotional, are underlyingly handled by operators and realized as particles and/or satellite adverbial expressions. For example, Reinforcement modality is expressed by the particle *ʾinna* 'it is certain' and/or adverbial satellites like *fiʿlan* 'really'. The expression of volitional modality is accomplished through particles like *layta* 'I hope' and *laʿalla* 'I wish'.

Exclamation is, as argued in Moutaouakil (1999), not an illocution but rather an (emotional) modality, which can take in its scope the entire clause, its predicate, or one of its terms, as shown in (7a), (7b), and (7c), respectively.

(7a) *tilka l-madīnatu rāʾiʿatun*
that the-city-Nom marvelous-Nom
'That city is marvelous!'

(7b) *mā ʾajmala hindan*
what beautiful Hind-Acc
'How beautiful Hind is!'

(7c) *ʾayya šāyin šaribtu*
what-Acc tea-Gen drank-I
'What a tea I drank!'

Arabic has many (positionally undifferentiated) sentence/clause-initial particles. The layering approach advocated in Functional Grammar accounts for the differences by analyzing these particles as coming from different underlying layers. Thus, *ʾa* and *hal* are analyzed as illocutionary operators, while *ʾinna*, *laʿalla*, and *layta* are derived in the modal layer.

Dik (1997a) defines Topic and Focus as pragmatic functions whose main task is to ensure discourse → coherence. Topicality and Focality are conceived as continuums of Topic and Focus types within which different types of languages select their grammatically relevant topical and focal distinctions (→ focus; → topicalization; → topic/comment). Typologically, Arabic belongs to a language type that strongly exploits the two continuums. It has a relatively large number of Focus constructions, including Fronting, Pseudo-cleft, and Negative-restrictive, which actualize various types of Focus (New Focus and Contrastive Focus, with subtypes such as Selecting Focus, Replacing Focus, and Restricting Focus).

(8a) *man qābalta*
whom met-you
'Whom did you meet?'

(8b) *qābaltu* ***hindan***
met-I Hind-Acc
'I met Hind'

(9) ***hindan*** *qābaltu (lā zaynaba)*
Hind-Acc met-I (not Zaynab-Acc)
'It was Hind that I met (not Zaynab)'

(10) *allatī qābaltu-hā* ***hindun***
who met-I-her Hind-Nom
'(The person) whom I met was Hind'

Examples (8), (9), and (10) show that in Arabic the Focus function can manifest itself by a mere prosodic prominence, a special constituent order, or special Focus constructions, mostly Pseudo-cleft constructions like (10).

The scope of Focus can be a term, as in (8b), (9), and (10), or a predicate or whole clause, as in (11) and (12b), respectively.

(11) *hudima l-baytu* ***hadman***
was-destroyed the-house-Nom destruction-Acc
'The house was completely destroyed'

(12a) *mā jarā*
what happened
'What happened?'

(12b) **rasabat zaynabu**
failed Zaynab-Nom
'Zaynab has failed'

The assignment of Focus function obeys the following constraint.

(13) **Focus Assignment Constraint**

(i) New Focus may be assigned to more than one constituent in the same clause; and
(ii) One and only one Contrastive Focus may be assigned to/within a clause.

In (14a) and (14b), New Focus is placed on three constituents, which is allowed by constraint (13i). In (15), two constituents carry Contrastive Focus, which is a clear violation of constraint (13ii).

(14a) *man raʾā man matā*
who saw whom when
'Who saw whom and when?'

(14b) *raʾā* ***fahdun hindan al-yawma***
saw Fahd-Nom Hind-Acc the-day-Acc
'Fahd saw Hind today'

(15) ****hindan al-yawma*** *raʾā fahdun*
Hind-Acc the-day-Acc saw Fahd-Nom

The Representational Level is a Predication designating a State of Affairs, which can be an Action, a Process, a Position, or a State. Predication consists of a predicate, which may be verbal, nominal, adjectival, or adverbial, and a given number of terms (argument and satellite) distributed over a Locality, a Quantity, and a Quality layer. Predicates are categorized according to their valency, both quantitative (number of arguments) and qualitative (type of predicate, the semantic functions of Agent, Goal, Recipient, etc., carried by the arguments, and the selection restrictions imposed by the predicate on its arguments). The Aspect-Mood-Tense-(positive/negative) Polarity properties are accounted for in the Quality, Quantity, and Locality layers. The unified underlying structure of (16a), resulting from the mapping of the Interpersonal level onto the Representational level, is roughly represented in (16b):

(16a) *ʾinna fahdan qābala hindan fī*
that Fahd-Acc met-he Hind-Acc in
tilka d-dāri
that the-house-Gen
'Certainly, Fahd met Hind in that house'

(16b) ([DECL Ei:[REINF Pi:[IND PAST PERF
ei:[qbl-fāʿal$_v$ (xi:Fahd) Top Ag
(xj:Hind) Go (yi:dār) New Foc Loc]]]])

The qualitative valency includes, among other features, information on predicate type, which is based on whether a predicate is basic or derived as, for example, Passive, Causative, and Reflexive predicates, and whether it takes as arguments a phrase, a predication, or a whole clause. These issues are extensively discussed in connection with derived and complex clauses in Moutaouakil (1988) and Dik (1997b).

According to the parallelism hypothesis, the underlying term structure conforms to the same general schema as the underlying predication structure. A term contains a nominal nucleus (or Head) and layers of Quality, Quantity, and Locality which represent (i) nominal aspectuality features such as Mass nouns and Count nouns; (ii) Quantifiers and Numerators; and (iii) Demonstrative and Definiteness/Indefiniteness oppositions. The structure in (17), for example, can be taken as a rough underlying representation of the Locative constituent *tilka d-dāri* 'that house' in (16).

(17) (DEM DEF SING yi:dār) New Foc Loc

Moutaouakil (2000) argues that even individual terms can be modalized, i.e. can display a modal layer. His argument is based on data such as (18a) and (18b), where the modal particle (appreciative/depreciative) has in its scope only the subsequent constituent.

(18a) *ʿindaka niʿma l-jawādu*
have-you good the-horse-Nom
'You have a good horse!'

(18b) *fī dārika biʾsa z-zāʾiru*
in house-Gen-you bad the-visitor-Nom
'You have a bad visitor in your house!'

In this approach to term structure, restrictive relative clauses are located in the Quality layer together with (non-appositional) adjectives.

The underlying representation is mapped onto a fully specified formal structure through morphosyntactic and phonological Expression Rules. Some of the most salient features of Arabic morphosyntax are syntactic functions assignment, case marking, and constituent ordering.

Only two syntactic functions are recognized in Functional Grammar: Subject and Object. Both are relevant for the description of Arabic clause structure. In general, their assignment is monitored by the following (probably universal) hierarchy.

(19) **Subject/Object Assignment Hierarchy**
(i) Subject:
Topic Agent > Non-Topic Agent > Non-Agent
+ + +
(ii) Object:
Focus Recipient > Non-Focus Recipient >
+ +
Focus Goal > Non-Focus Goal > Others
+ + +

As (19i) shows, Topic, Agent, and Subject functions tend to coincide, yielding what is commonly called 'Prototypical Subject'.

Case marking in Modern Standard Arabic distinguishes two types of case: Functional cases vs. Structural cases and Underlying cases vs. Surface cases. Functional cases (Nominative and Accusative) are determined by the syntactic (or otherwise the semantic) functions the constituents have in the underlying clause structure. Nominative is assigned to the Subject, and Accusative is assigned to the other (non-Subject) constituents. An example of the assignment of these two cases is given in (20b), which represents a pre-surface structure for (20a).

(20a) *kataba bakrun risālatan*
wrote Bakr-Nom letter-Acc
'Bakr wrote a letter'

(20b) ([DECL E:[ei:[katab$_v$ (SING PROP xi:
Bakr) Top Ag Subj-nom
(INDEF SING xj:risālat) New Foc Go
Obj-acc]]])

Structural case (Genitive and Accusative) is assigned by certain prepositions (e.g. *min* 'from', *fī* 'in'), particles (e.g. *ʾinna* 'that', *layta* 'wish, hope'), and Auxiliary verbs (e.g. *kāna* 'to be', *bāta* 'to become'). It is assigned configurationally within the so-called Annexive phrases. When a constituent bears structural in addition to functional case, the former always 'masks' the latter. This is illustrated in (21b) and (22b), where the Nominative functional case is masked by the Accusative and the Genitive structural cases assigned by the particle *ʾinna* 'that' and the preposition *min* 'from', respectively.

(21a) *hindun nāʾimatun*
Hind-Nom sleeping-Nom
'Hind is sleeping'

(21b) *ʾinna hindan nāʾimatun*
that Hind-Acc sleeping-Nom
'Certainly, Hind is sleeping'

(22a) *hal sāʿadaka rafīqun*
Q helped-you friend-Nom
'Did a friend help you?'

(22b) *hal sāʿadaka min rafīqin*
Q helped-you of friend-Gen
'Did any friend help you?'

Surface cases are morphemes (Arabic endings) by means of which underlying cases are realized. The distinction between Underlying case and Surface case is justified by the possible 'discrepancies' between underlying and surface cases: the former can receive no formal expression and can be expressed by morphemes other than the expected ones. Nouns without → nunation and sound feminine plurals are well-known examples in this respect.

The underlying clause structure is conceived of as an unordered network of (pragmatic and semantic) features and relations. It is mapped onto a linear sequence by a set of position-assigning rules called 'Placement Rules'. These rules obey general principles defining permissible and impermissible sequences of constituents and combinations thereof. They operate on the basis of language-specific templates. Examples of constituent-ordering principles include the Principle of Pragmatic Highlighting (PPH) and the Principle of Increasing Complexity (PIC). According to the Principle of Pragmatic Highlighting, constituents with special pragmatic functions (e.g. Topic and Contrastive Focus) take 'special positions' including clause-initial position(s). In Arabic, Focus or Topic constituents are placed in the second initial position in a clause, regardless of their semantic or syntactic status. The fronted constituent in (23b), standing as a corrective answer to (23a), is placed in this special position according to its pragmatic function (Contrastive Focus).

(23a) A: *ʿašiqa fahdun* ***zaynaba***
loved Fahd-Nom Zaynab-Acc
'Fahd loved Zaynab'

(23b) B: *lā,* ***hindan*** *ʿašiqa fahdun*
no Hind-Acc loved Fahd-Nom
'No, it was Hind that Fahd loved'

The Principle of Increasing Complexity stipulates that constituents tend preferably to be sequenced in an order of increasing complexity. Compare, for example, (24a) with (24b).

(24a) *balaġa hindan ʾanna maryama*
arrived-at Hind-Acc that Maryam-Acc
rasabat
failed
'Hind was informed that Maryam has failed'

(24b) ??*balaġa ʾanna maryama rasabat*
arrived-at that Maryam-Acc failed
hindan
Hind-Acc

The grammaticality of (24b) is doubtful when it is compared with (24a). In competing situations certain principles neutralize the effect of others. For instance, in (24a) the inverted constituents *hindan* and *ʾanna maryama rasabat* (a phrase and an embedded clause, respectively) are placed, under the pressure of the Principle of Increasing Complexity, in positions other than those expected on the basis of their semantic or syntactic functional status. The assignment of appropriate positions takes place according to given syntactic templates. In Arabic, the relevant template for a verbal clause is given in (25):

(25) P1 PO V S (O) (X)

This template is to be read as follows:

Clause-initial particles and subordinators ('complementizers') go to P1;

Q-constituents and Topic or Contrastive Focus constituents go to PO;

P1 and PO can each house only one constituent;

V, S, and O are the positions of the verbal predicate, the Subject and the Object constituents, respectively; and

Constituents without pragmatic or syntactic function are hosted in position X.

3. SENTENCE STRUCTURE

In Functional Grammar, a discourse category standing between the clause and the text is recognized and commonly referred to as 'Sentence'. The general format of sentence structure is given in (26).

(26) (ECCs) CLAUSE (ECCs)

Schema (26) shows that a sentence results from adding Extra Clausal Constituents (ECCs) to a clause. Extra Clausal Constituents do not belong to the clause proper. They can take a pre-clausal or a post-clausal position and may also occur as parenthetical elements. Their role in discourse is limited to four main functions: (i) interaction management, as in greetings, leave-takings, and summonses; (ii) attitude specification as in the expression of Speaker's emotional state; (iii) discourse organization; and (iv) discourse execution, as in responses and tags.

Two Extra Clausal Constituents fulfilling the discourse organization macrofunction, the Theme and Tail constituents, are exemplified in (27a) and (27b), respectively.

(27a) *fahdun,* *raʾaytu-hu* *l-yawma*
Fahd-Nom saw-I-him the-day-Acc
'Fahd, I saw him today'

(27b) *nāmū,* *al-ʾawlādu*
slept-they the-children-Nom
'They are sleeping, the children'

The constituent with Theme function is defined as a constituent designating the 'universe of discourse' with respect to which it is relevant to utter the subsequent clause. The structure involved in Theme constructions such as (27a) can be represented as in (28).

(28) Theme-nom (xi), (…(xi)…)

From representation (28), four main properties of Theme constructions can be deduced. First, the Theme constituent has an autonomous intonational contour marked by a comma. Second, although the clause is independent of the preceding Theme, it must be relevant to be predicated to it. Compare, for example, (29a) with (29b). (Here and elsewhere 'Co-' in the morpheme translation stands for 'coordinator'.)

(29a) *ʾammā* *marrākušu,* *fa-ʾinna*
as-for Marrakesh-Nom Co-that
manāratahā *mašhūratun*
Menara-Acc-it famous-Nom
'As for Marrakesh, its Menara is famous'

(29b) **ʾammā* *fāsun,* *faʾinna*
as-for Fes-Nom Co-that
manārata-hā *mašhūratun*
Menara-Acc-it famous-Nom
*'As for Fes, its Menara is famous'

Third, the Theme is typically resumed by a pronoun within the subsequent clause, as in (27a), for example. However, the resumptive pronoun is not always necessary, as, for example, in (30).

(30) *as-samnu,* *al-kīsu*
the-butter-Nom the-bag-Nom
bi-ʿišrīna *dirhaman*
with-twenty-Gen dirham-Acc
'Butter costs twenty dirhams a bag'

In these cases, the Theme-clause link is ensured only by the relevance pragmatic relationship. In Arabic, unlike languages such as Chinese, this kind of construction is rather rare.

Fourth, the Theme constituent typically takes Nominative case. Given the externity of this constituent, the case it carries is assigned either by default or by the Theme pragmatic function itself. Other markers may characterize the Theme constituent. These include the well-known embracing morpheme *ʾammā… fa* 'as for…', occurring in constructions like (29a) and (29b) (→ theme/rheme).

The Tail function is assigned to the constituent that presents, as an afterthought, information meant to clarify, modify, or correct the content of the clause, or a constituent included in it. The most pervasive type of Tail construction is the so-called Right-dislocation illustrated in (27b), whose rough configuration is given in (31).

(31) (…(xi)…), Tail-nom (xi)

The Tail constituent in these constructions is intonationally set off from the clause proper by a comma, as representation of intonational pause. The Tail constituent is represented within the clause by a cataphoric pronoun. This pronoun is a full argument with semantic, syntactic (Subject/Object), and pragmatic (Topic) functions.

Certain Extra Clause Constituents can be integrated into the clause proper. The Theme and the Tail constituents are more likely to undergo this progressive diachronic process, which can be illustrated by the examples in (32a)–(32c).

(32a) *bakrun* *rajaʿa*
Bakr-Nom came-back
'Bakr came back'

(32b) *fahdan* *raʾaytu-hu*
Fahd-Acc saw-I-him
'I saw Fahd'

(32c) *nāmū* *l-ʾawlādu*
slept the-children-Nom
'The children are sleeping'

Three factors favor the integration: (i) the 'demarking' (or 'loss of markedness') process that the constructions often undergo due to frequency of use; (ii) the pressure that the predicate exercises on the Extra Clause Constituents in order to draw them into the clause and convert them into fully governed arguments; and (iii) the lack of integration-blocking barriers such as Theme markers (mentioned above) and clause-initial particles.

The integration process may have two kinds of effects, 'local' and 'global' (Moutaouakil 1993). Locally, at the clause level, the integrated constituent (Theme or Tail) receives an argument status. When drawn into the clause, the Theme becomes a Topic-Agent-Subject or a Topic-Goal-Object, as in the constructions exemplified in (32a), and the so-called *ištiġāl*-constructions, as in (32b). As a consequence, the resumptive pronoun becomes a mere Subject or Object agreement marker. The integrated Tail in constructions like (32c) is analyzed as a postponed Subject while the cataphoric pronoun acts as a Subject agreement marker (→ cataphora).

Globally, the integration of the two constituents at hand results in crucial changes in both the constituent ordering and the pronominal system of the language in which this process takes place. In Arabic, the absorption of the Theme constituent is leading to a change from VSO to SVO word order, as evidenced by the increasing frequency of verbal clauses with initial Subject such as (32a) in Modern Standard Arabic and the dialects. Number agreement between the verb and its postponed Subject in the Arabic dialects, illustrated in (32c), can be viewed as a direct consequence of the progressive integration of the Tail constituent (Moutaouakil 1993).

Bibliographical references

Cuvalay-Haak, Martine. 1996. *The Arabic verb*. Berlin: Mouton de Gruyter.

Dik, Simon, C. 1978. *Functional grammar*. Amsterdam: North-Holland.

——. 1997a. *The theory of functional grammar*. I. *The structure of the clause*. 2nd rev. ed., ed. Kees Hengeveld. Berlin: Mouton de Gruyter.

——. 1997b. *The theory of functional grammar*. II. *Complex and derived constructions*. 2nd rev. ed., ed. Kees Hengeveld. Berlin: Mouton de Gruyter.

Hannay, Mike and A. Machtelt Bolkestein. 1998. *Functional grammar and verbal interaction*. Amsterdam and Philadelphia: J. Benjamins.

Hengeveld, Kees. 1989. "Layers and operators in Functional Grammar". *Journal of Linguistics* 25:1.127–157.

——. 2004a. "The architecture of functional discourse grammar". Mackenzie and Gómez-González (2004:1–22).

——. 2004b. "Epilogue". Mackenzie and Gómez-González (2004:365–378).

Mackenzie, John Lachlan. 1995. "Functional Grammar". *Handbook of pragmatics*, ed. Jef Verschueren, Jan-Ola Ostman, and Jan Blommaert, 286–293. Amsterdam and Philadelphia: J. Benjamins.

—— and María de los Ángeles Gómez-González. 2004. *A new architecture for functional grammar*. Berlin: Mouton de Gruyter.

Moutaouakil, Ahmed. 1984. "Le Focus en arabe: Approche fonctionnelle". *Lingua* 64.115–175.

——. 1988. *al-Jumla al-murakkaba fī l-luġa al-ʿarabiyya*. Rabat: Okad.

——. 1989. *Pragmatic functions in a functional grammar of Arabic*. Dordrecht: Foris.

——. 1993. *ʾĀfāq jadīda fī naḏ̣ariyyat an-naḥw al-waḏ̣īfī*. Rabat: Publications de la Faculté des Lettres.

——. 1996. "On the layering of the underlying structure in FG". *Complex structures: A functionalist perspective*, ed. Betty Devriendt, Louis Goossens, and Johan van der Auwera, 201–227. Berlin: Mouton de Gruyter.

——. 1998. "Benveniste's *récit* and *discours* as discourse operators in Functional Grammar". Hannay and Bolkestein (1998:25–42).

——. 2000. *Reflections on the layered underlying representation in Functional Grammar*. Casablanca: Afric-Orient.

——. 2004. “Discourse structure, the generalized parallelism hypothesis and the architecture of functional grammar”. Mackenzie and Gómez-González (2004:351–380).

——. 2005. “Exclamation in Functional Grammar: Sentence type, illocution or modality?”. *Morphosyntactic expression in Functional Grammar*, ed. Casper de Groot and Kees Hengeveld, 351–381. Berlin: Mouton de Gruyter.

Rijkhoff, Jan. 1992. *The noun phrase: A typological study of its form and structure*. Ph.D. diss., University of Amsterdam.

Ahmed Moutaouakil (Mohammed V University)

Fuṣḥā → Classical Arabic; Faṣīḥ; ʿArabiyya

G

Gahawa-Syndrome

In a feature typical of Bedouin dialects, a short vowel *a* follows a morpheme-internal back spirant (*X*), whenever this *X* is preceded by *a*. This type of vowel insertion had been described previously, in varying degrees of detail, by, for example, Wetzstein (1868:185–186, 191), Cantineau (1936:66), Mitchell (1960:388), and Johnstone (1964:80), before Blanc dubbed it the '*gahawah*-syndrome' (1970:125–127).

Although known as a general characteristic of 'Bedouin' dialects – Blanc states that it is found "only in *gāl* dialects" (1970:127, n. 29) – the syndrome has been reported for sedentary dialects as well, e.g. dialects of the Egyptian Nile Valley, roughly south of Asyūṭ (in Upper Egyptian 1, 3, and less regularly also in 4; see Behnstedt and Woidich 1988, 1985, maps 45–46), and among sedentary speakers of the Najd (cf. Blanc 1970:127, n. 29). In such cases, the *gahawa*-syndrome is best interpreted as evidence of contact with dialects of the Bedouin type (on Bʿēri or Upper Egyptian 3, for example, see Woidich 1997:195).

The rule may be summarized as follows:

Ø > a / (C)aX_C(V)

X = *h*, *ḥ*, *r*, *x*, or *ġ* (i.e. pharyngeal, laryngeal, uvular/velar fricatives)
C = any consonant
V = any short or long vowel

The examples below include the syndrome's namesake:

**qahwa*	*gahwah*	>	*gahawah*	'coffee'
**naʿja*	*naʿjah*	>	*naʿajah*	'ewe'
**naxl*	*naxḷ*	>	*naxaḷ*	'palm trees'
**baḥr*	*baḥṛ*	>	*baḥaṛ*	'sea'
**baġla*	*baġḷah*	>	*baġaḷah*	'female mule'
**yaxbuṭ*	*yaxbuṭ*	>	*yaxabuṭ*	'he knocks'
**taḥt*	*taḥt*	>	*taḥat*	'under'

The phonetic quality of the *gahawa*-vowel in nonvelarized environments is near IPA front [a], even if the unstressed short *a* preceding X may be nearer to centralized and slightly raised IPA [ɐ] (e.g. in dialects that have a CaCáC(v) stress type, such as some dialects in the Negev and the Sinai, but also when stress shifts due to suffixing, e.g. CaXaC+*ha*). Examples are:

kaʿk#	>	*kaʿak*	['kaʕak]	or [kɐ'ʕak]	'cookies'
taḥt#	>	*taḥat*	['taħat]	or [tɐ'ħat]	'under'
baxt#	>	*baxat*	['baxat]	or [bɐ'xat]	'luck'

In velarized environments the *gahawa*-vowel tends to be nearer to back IPA [ɑ]. The preceding *a* is usually around the same phonetic quality, even when unstressed, e.g.:

baʿḏ̣#	>	*baʿaḏ̣*	['bɑʕɑð̴]	or [bɑ'ʕɑð̴]	'each other'
baḥṛ#	>	*baḥaṛ*	['bɑħ̴ɑr̴]	or [bɑ'ħ̴ɑr̴]	'sea'
naxḻ#	>	*naxaḷ*	['nɑxɑɫ]	or [nɑ'xɑɫ]	'palm trees'
baġḻ#	>	*baġaḷ*	['bɑɣ̴ɑɫ]	or [bɑ'ɣ̴ɑɫ]	'mule'

In the dialect of the Cyrenaican Jebel (see Mitchell 1960:388), raising of the vowel in the first unstressed syllable is more extreme, so that we find forms like *buḥáṛ, nuxáḷ, buġáḷ* (velarization is marked here in the phonetic quality of the *a* of the second syllable, which is transcribed as [ɑ]), *fiḥál* 'stud camel', and *liḥám* 'meat'.

Once the *gahawa*-vowel has become stable, the resulting sequence CaXaC(V) (< CaXC(V)) may become subject to other rules as well, rules by which 'original' CaXaC(V) sequences are also affected.

The Najdi type of resyllabification of CaCaCV sequences is an example of the *gahawa*-syndrome in a 'feeding' role (in generative terms). In Najdi-type dialects, CaCaCV se-quences are resyllabified as CCvCV (a 'phono-tactic constraint' bars the occurrence of CaCaCV sequences; see, e.g., Ingham 1982:37). Sequences such as *katabat* and *zalamah* are thus resyllabified as *ktibat* and *zlimah*, i.e., the vowel *a* of the initial open syllable is dropped and the vowel *a* of the next open syllable is raised to *i*, while in velarized and/or labial environments it tends to be raised to *u*, as in, for example, *ḥṭuḅah* (< *ḥaṭabah*) 'piece of firewood'.

After the *gahawa*-syndrome has produced CaCaCV sequences (i.e. CaXaCV) and they have become stable as morphophonemic bases, these too become subject to this Najdi resyllabification rule. Often, however, the mouth is in an open position, which leads to the creation of the *gahawa*-vowel as *a* in the first place. This fact also prevents the raising of the *gahawa*-vowel in the second open syllable (as is often but not always the case in forms like *ṣḥabat* instead of *ṣḥibat* 'she sat', i.e. with the 'original' *a* of *ṣaḥab* + *at*; see Ingham 1982:49 and also Johnstone 1967:6 for comparable forms), although the vowel *a* of the first open syllable is elided in conformity with the resyllabification rule. Examples are *rḥámah* 'compassion' and *ghawah* 'coffee', and verb forms such as *yʿágid* 'he ties', *yʿárif* 'he knows', *yʿáṛag* 'he sweats' (many such examples in dialects of the Arabian Peninsula may be found in Prochazka 1988:36–37, 143). Mitchell (1960:389) reports forms from the Cyrenaican Jebel like *inḥálih* 'bee' and *umġaṛaf* 'ladle', while Johnstone (1967:14) gives forms from the ʿAnayza like *nxalih* 'date palm', *nʿajih* 'she-goat', and *ghawih* 'coffee'. In many dialects *hamza*-initial forms lose the *hamza* together with the vowel from the first syllable (i.e., the entire syllable is dropped), resulting in X as the initial consonant, e.g. *halu* (< * *ʾahal* + *u*, instead of *ʾhalu*) 'his family' and *ḥamar* (< **ʾaḥamar*, instead of *ʾḥamar*) 'red'.

Another example of the *gahawa*-syndrome in a 'feeding' role is found in the dialect of the Mzēnih of southern Sinai (see de Jong, forthcoming). *Gahawa*-forms are treated in the same manner as 'original' CaCaCV base forms. In this dialect, a rule specifies the resyllabification of sequences of the type CaCaCT+v(C) (like *ragab*T+*uh*) as CaCCitv(C), thus producing the proper Mzēni form *rágbituh* 'his neck'. Since in Mzēni the *gahawa*-syndrome has caused morphological restructuring of base forms of the pattern CaXCT as CaXaCT, a *gahawa*-form like *naxaḷ*T+*uh* will be treated in the same manner, resulting in *náxḷituh* 'his date palm'. Paradoxically, after having been created by the *gahawa*-syndrome, in this dialect the *gahawa*-vowel is dropped in such sequences, since the rule specifies that morphophonemic *a* (here underlined) in sequences CaCa̱CT+v(C) is to be elided when vowel-initial suffixes follow. (On the other hand, proper Mzēni forms with consonant-initial suffixes are, for example, *ṛagabatha* and *naxaḷatha*.)

In many dialects, the incorporation of *gahawa*-vowels into new morphophonemic bases appears to be somewhat problematic when imperfect verb forms are involved. For instance, in several Sinai Bedouin dialects (see de Jong 2000:109) and also in that of the Mzēnih, *gahawa*-vowels – concluded from the fact that they are not stressed – tend to behave more like anaptyctics than morphophonemic base vowels. Forms like (*gahawa*-vowels underlined) *yáʿa̱rfih* 'he knows him' and *yáḥa̱rṯuw* 'they plow' (instead of *yaʿá̱rfih* and *yaḥá̱rṯuw*) are thus comparable to forms such as (anaptyctic vowels underlined) *yíki̱tbih* 'he writes it' and *yúḍu̱rbuw* 'they hit [imperf.]'.

In the dialect of ilBiʿiṛāt (→ Bʿēri) on the west bank of the Nile opposite Luxor, the *gahawa*-syndrome is also active, with forms such as *yaxlaṣ* and *naxḷa* appearing on the surface as *yaxalaṣ* 'it ends' and *naxaḷa* 'date palm'. However, in Bʿēri, the *Umlautung*-syndrome 'counter-feeds' the rule of *Umlautung*. Forms like *masak+at* and *yaḍrab+aw* are normally 'umgelautet' to appear on the surface as *misikat* and *yušṛubaw*, and likewise a form such as *yaxlaṣ+aw* will surface as *yuxluṣaw* 'they end'. A *gahawa*-form like *yaxalaṣ* 'it ends', however, is left unaffected by the *Umlautung*-rule (and does not surface as *yuxulaṣ*). From this it follows that, in terms of rule ordering, the *gahawa*-rule is preceded by the rule of *Umlautung* in Bʿēri (see Woidich 1973–1974, 1974).

Loans from the standard language or Classical Arabic are often unaffected by the *gahawa*-syndrome, e.g. *maḥkamah* 'court', *arraḥmān* 'the Merciful'. The *gahawa*-syndrome also usually remains inactive in derived forms and quadriliteral verbs, e.g. (Form IV) *aʿṭa* 'he gave', (Form *ista*-I) *istahbal* 'he wondered', (quadriliteral) *zaġraṭat* 'she ululated'. The syn-drome does not reach beyond the morpheme boundaries of the verbal stem of the perfect, e.g. *rajaʿt* 'I returned', *balaġna* 'we reached', *najaḥtuw* 'you [pl.] succeeded', nor beyond those of the noun, e.g. *balaḥna* 'our dates'.

Considering the vastness of the geographical area where the *gahawa*-syndrome is known to be present in dialects, it must be of considerable antiquity, and it almost certainly antedates the spread of Arabic.

BIBLIOGRAPHICAL REFERENCES

Behnstedt, Peter and Manfred Woidich. 1985. *Die ägyptisch-arabischen Dialekte*. II. *Dialektatlas von Ägypten*. Wiesbaden: L. Reichert.

——. 1988. *Die ägyptisch-arabischen Dialekte*. III. *Texte*. II. *Niltaldialekte*. III. *Oasendialekte*. Wiesbaden: L. Reichert.

Blanc, Haim. 1970. "The Arabic dialect of the Negev Bedouins". *Proceedings of the Israel Academy of Sciences and Humanities* 4.112–150.

Cantineau, Jean. 1936. "Etudes sur quelques parlers de nomades arabes d'Orient". *Annales de l'Institut d'Études Orientales d'Alger* 1.119–237.

Ingham, Bruce. 1982. *North east Arabian dialects*. London and Boston: Kegan Paul International.

——. 1991. "Camel terminology among the Āl Murra Bedouins". *Zeitschrift für Arabische Linguistik* 22.69–78.

Johnstone, Thomas M. 1964. "Further studies on the Dōsiri dialect of Arabic as spoken in Kuwait". *Bulletin of the School of Oriental and African Studies* 27.77–113.

——. 1967. "Aspects of syllabication in the spoken Arabic of ʿAnaiza". *Bulletin of the School of Oriental and African Studies* 30.1–16.

Jong, Rudolf E. de. 2000. *A grammar of the Bedouin dialects of the northern Sinai littoral: Bridging the linguistic gap between the eastern and western Arab world*. Leiden: E.J. Brill.

——. Forthcoming. "The dialect of the Mzēnih of South Sinai".

Mitchell, T.F. 1960. "Prominence and syllabication in Arabic". *Bulletin of the School of Oriental and African Studies* 23.369–389.

Prochazka, Theodore Jr. 1988. *Saudi Arabian dialects*. London: Kegan Paul International.

Wetzstein, János Gottfried. 1868. "Sprachliches aus den Zeltlagern der syrischen Wüste". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 22.69–194.

Woidich, Manfred. 1973–1974. "Die 3.sg.f. Perfekt im Dialekt von il-Biʿrāt". *Mélanges de l'Université Saint Joseph* 48.355–372.

——. 1974. "Ein arabischer Bauerndialekt aus dem südlichen Oberägypten". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 124.42–58.

——. 1997. "Upper Egyptian Arabic and dialect mixing in historical perspective". *Humanism, culture and language in the Near East: Studies in honor of Georg Krotkoff*, ed. Asma Afsaruddin and A.H. Mathias Zahniser, 185–197. Winona Lake, Ind.: Eisenbrauns.

RUDOLF DE JONG (University of Amsterdam)

Geʿez → Ethiopia; Ethiopic Loanwords

Gemination

Geminate consonants in Classical Arabic are not contrastive, i.e., there are no two monomorphemic words that contrast single and geminate consonants. Gemination of consonants, however, is associated with a number of morphological contexts. Along with cases of morphologically conditioned gemination, there are cases of phonologically conditioned gemination, occurring as a consequence of satisfying the templatic conditions of stems in Arabic (McCarthy and Prince 1990a, 1990b). The targets of morphological and phonological gemination differ: morphologically conditioned gemination involves geminating the medial consonant of a triliteral root, whereas phonologically conditioned gemination involves geminating the final consonant of the root.

Morphologically conditioned gemination occurs in both the verbal and the nominal morphology. A typical case of gemination is seen in Form II verb stems, shown in (1).

(1)		Form I		Form II	
	k-t-b	*katab*	'to write'	*kattab*	'to cause to write'
	j-l-s	*jalas*	'to sit'	*jallas*	'to cause to sit'

Medial gemination in Form II is analyzed prosodically, following McCarthy and Prince (1990a, 1990b). The CVCCVC structure of Form II is understood as a disyllabic template consisting of a bimoraic syllable and a monomoraic syllable (the final consonant is extrasyllabic). The geminate consonant is represented as a consonant autosegmentally linked to a mora in the syllable coda and the onset of the following syllable. This association is determined by the Medial Gemination Rule (see McCarthy and Prince 1990b), shown in Figure 1 and exemplified with *kattab*.

Triliteral roots in the nominal morphology also have medial gemination. This is found in nouns of profession *kallāf* 'stablehand', *xabbāz* 'baker'; habitual action *kawwāy* 'slanderer'; and a very small number of underived nouns, such as *jabbār* 'giant'. The geminate is the result of a triconsonantal root, e.g. *k-l-f*, associating to a template that contains two bimoraic syllables (see McCarthy and Prince 1990b). The root-medial consonant associates to the second mora of the first heavy syllable by the Medial Gemination Rule (see Fig. 1).

Another case of medial gemination, also morphologically conditioned, occurs in the plural of lexicalized active participles, for example *bāhil/buhhal* 'free' and *'ābiq/'abbāq* 'fugitive'. Note that the initial syllable of the singular is bimoraic, which is realized as a long vowel. The plural for this class must have a geminated medial consonant, which is concomitant with shortening the initial vowel in the singular. The prosodic shape of the singular and the plural is similar insofar as both forms contain a bimoraic initial syllable, but they differ with respect to autosegmental association to the template. Following McCarthy and Prince (1990b), the plural is formed by the Medial Gemination Rule, which associates the medial consonant to the second mora of the initial syllable. The singular does not have this rule, so the vowel associates to both moras of the initial syllable.

The Medial Gemination Rule only demands that a consonant link to a mora, but the relation to morphological context is arbitrary. The connection between gemination and morphology is found when the derivation of these stems is considered. McCarthy (1992) proposes that Form II stems are derived by affixing a mora to the Form I stem. This mora, which is the morpheme associated with Form II, is realized as an infix, via prosodic circumscription (see McCarthy and Prince 1990a; McCarthy 1992). The mora is linked to the consonant at the left edge of the circumscribed domain, shown in parentheses in (2).

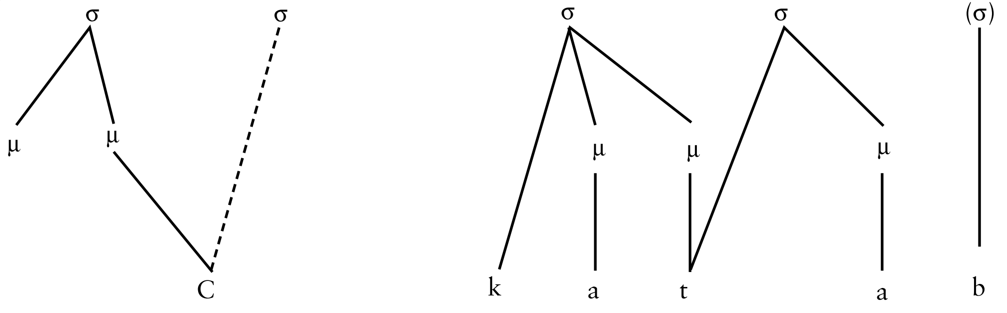
(2) Form II: μ + Form I
k-t-b ⇒ katab ⇒ ka$^{*\ \mu}$(tab) ⇒ kattab

This analysis can be extended to gemination in noun stems. The CVCCVVC template of a noun of professon is derived by affixing a mora to a disyllabic base that contains a monomoraic syllable followed by a bimoraic syllable with an extrasyllabic final consonant. The base is an iambic foot, in accordance with the prosodic morphology hypothesis (McCarthy and Prince 1990a, 1990b). The derivation of *kallāf*, using prosodic circumscription, is shown in (3).

(3) Noun of profession: μ + [F $^{\sigma}$μ $^{\sigma}$μμ]
k-l-f ⇒ kalāf ⇒ ka$^{*\mu}$(lāf) ⇒ kallāf

Medial gemination occurs in certain morphological contexts because the mora that triggers gemination is part of the morphemic representation. The plurals of lexicalized participles can also be derived in a similar way, and the small percentage of underived nouns must have a lexicalized form of this rule.

Figure 1. Medial Gemination Rule



Medial Gemination Rule

kattab

Attributing medial gemination to a mora, which is part of the morphology, reveals the prosodic organization of the stem and is superior to an analysis involving autosegmental association to a CV template (see McCarthy 1979). The autosegmental approach requires one-to-one, left-to-right association and then subsequent delinking and relinking.

Phonological gemination, in contrast to morphological gemination, is triggered by the satisfaction of conditions on the stem template. There are two template conditions that compel gemination: (1) the template must be maximally filled, and (2) the stem must be consonant-final (see McCarthy and Prince 1990a, 1990b). As mentioned above, phonological gemination applies to the final radical of the consonantal root and is triggered by the fact that there are fewer consonants than positions in the stem template. The final consonant will geminate to ensure satisfaction of the conditions on the template.

The CVCVC shape of Form I verbs is derived by a disyllabic template and consonant finality. McCarthy (1979) argues that geminated verb stems, e.g. *jarar* 'to pull' and *samam* 'to poison', are derived from the biliteral roots *j-r* and *s-m*, respectively. Association of the consonants to the appropriate positions in the stem template fails to satisfy consonant-finality of the stem. As a result, the second consonant doubles to satisfy this condition (see Fig. 2).

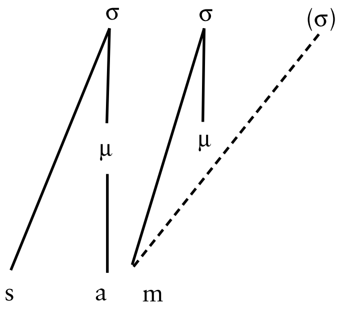


Figure 2. Gemination in biliteral roots

The alternation between geminates and doubled consonants associated with these stems, e.g. [jarra] ~ [jararta], [samma] ~ [samamta], is the result of a subsequent rule (see McCarthy 1979).

Noun stems with a geminate, e.g. *tall* 'hill', *barr* 'reverent', are derived from biliteral roots *t-l* and *b-r* respectively. The CVCC shape of these stems is the result of satisfying a bimoraic minimal template (see McCarthy and Prince 1990a, 1990b), as in Figure 3.

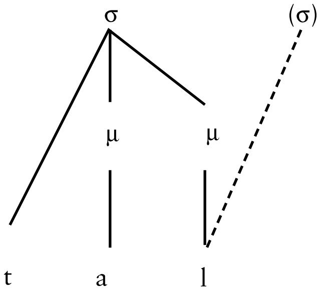


Figure 3. Geminate noun stems

The final consonant is associated to the second mora of the template and the extrasyllabic position.

Gemination also occurs with the prefix /al/. This is a segmentally conditioned process: a word-initial coronal consonant geminates, e.g. [aʃ-ʃams] 'the sun', [ad-daːr] 'the house', [aθ-θawb] 'the garment', but [al-qamar] 'the moon', [al-faras] 'the mare'.

Bibliographical references

McCarthy, John J. 1979. *Formal problems in Semitic phonology and morphology*. Ph.D. diss., Massachusetts Institute of Technology.

——. 1992. "Template form in prosodic morphology". *Papers from the Annual Meeting of the Formal Linguistics Society of Midamerica* 3.187–218.

——. and Alan Prince. 1990a. "Foot and word in prosodic morphology: The Arabic broken plural". *Natural Language and Linguistic Theory* 8.177–208.

——. 1990b. "Prosodic morphology and templatic morphology". *Perspectives on Arabic Linguistics*, II, ed. Mushira Eid and John J. McCarthy, 1–54. Amsterdam and Philadelphia: J. Benjamins.

Samuel Rosenthall (Rochester, MI, U.S.A.)

Gender

1. THE NATURE OF GRAMMATICAL GENDER

Arabic has two genders, conventionally known as feminine and masculine. It has no neutral. In general, masculine nouns are not marked for gender, but feminine nouns may or may not be marked. Every animate and inanimate noun must have a grammatical gender irrespective of whether the noun is marked or not.

Gender distinction coincides with natural sex division in nouns that denote animates; thus, nouns denoting female humans and animals are feminine, while nouns denoting male humans and animals are masculine. However, gender distinction in inanimate nouns is more problematic since it does not refer to natural gender. In fact, it is purely conventional, and it is agreement that disambiguates masculine nouns from unmarked inanimate feminine nouns.

Like other gender languages, questions about the emergence of gender in Arabic have led to two conflicting theories, namely the sex-based theory and the grammar-based theory. The sex-based theory attributes the existence of linguistic gender to speakers who classify everything in nature into masculine and feminine based on natural gender. The name associated with this theory in Arabic and in Semitic languages in general is that of William Wright, who clearly states that "the vivid imagination for the Semite conceived all objects, even those that are apparently lifeless, as endowed with life and personality. Hence for him there are but two genders, as there exist in nature but two sexes" (1896:131). This theory has been criticized for being speculative in nature, and not many contemporary linguists seem to subscribe to it today. The grammar-based theory about Arabic gender, on the other hand, argues à la Brugmann for the independence of gender from social and psychological factors, giving primacy instead to linguistic factors. A number of scholars (e.g. Brockelmann 1908:418–426; Féghali and Cuny 1924; Speiser 1936), influenced by advances in the study of Indo-European gender, have sought explanations of the development of Arabic and Semitic gender in purely linguistic changes. Most of these diachronic studies found evidence for the feminine suffix *-at*. Ibrahim sums up this view as follows: "Grammatical gender is merely a means for classifying nouns according to their suffixes without in the beginning any allusion to sex; the sex reference of gender was always posterior to the emergence of grammatical gender" (1973:50). On the other hand, the ancient Arab grammarians did not provide any theories about the origin of Arabic gender, and their accounts on the topic in general were strictly descriptive (cf. Ibrahim 1973).

2. Description and functions of the feminine gender in Arabic

The most common marker of the feminine is the suffix *-at*, which is usually added to the masculine form to derive the feminine. This is known as *tāʾ at-taʾnīṯ* 'the *t* of femininity' when referring to its grammatical function, or *tāʾ marbūṭā* 'bound *t*' when referring to its orthographic form. Its pronunciation varies between *-t-* when followed by declensional endings, and *-a* when it is in pausal form. For instance, in the sentence *madrasat-u ṭiflat-ī baʿīda* 'my daughter's school is far', the *-t* is pronounced before the nominative case marker *-u* in the first noun of the construct state, and before the possessive suffix *-ī* in the second noun of the construct *ṭiflat-ī*, but it is silent in *baʿīda*.

Besides *-at* there are at least three other less common feminine endings, namely *ʾalif maqṣūra* and *ʾalif mamdūda*, both of which are transcribed as *ā*, as in *taqwā* 'piety' and *ruʾyā* 'vision', and *āʾ* as in *samāʾ* 'sky'. The *āʾ* ending is very common in feminine adjectives, such as those denoting color terms: *bayḍāʾ* 'white [fem.]', *sawdāʾ* 'black [fem.]', *ḥamrāʾ* 'red [fem.]'; it is also found in other descriptive adjectives, such as *ḥasnāʾ* 'pretty', *ʿaḏrāʾ* 'virgin', and in proper names like *lamiyāʾ*. It has been argued that even the noun *ṣaḥrāʾ* 'desert' is an old adjective from *ʾaṣḥar* 'fawn-colored' (Féghali and Cuny 1924:18). This type of lexicalization of descriptive adjectives seems to be a very common phenomenon, not only in Arabic but in Semitic languages in general.

Of all the Semitic languages, Literary Arabic in particular has preserved the largest number of feminine endings, but in the modern Arabic dialects these endings have merged into one ending, *-a*, e.g. Classical Arabic *samāʾ* 'sky', Moroccan Arabic *sma*; Classical Arabic *ḥamrāʾ*, Syrian Arabic *ḥamra*.

It is widely accepted in the Arabic grammatical tradition that the feminine suffix *-at* has several other functions besides the marking of feminine gender (see Suyūṭī, *Muzhir* II, 222; Wright 1896; Féghali and Cuny 1924; Fleisch 1961; Moscati 1964; Ibrahim 1973; Drozdík 1998, to cite just a few). The feminine suffix is also used to build *nomina unitatis*, or singulatives, collectives, abstract nouns, diminutives and intensives.

Ibn Qutayba (d. 276/889) must be credited with discovering the singulative function of *-at* (cf. Ibrahim 1973:48ff.) or the use of the feminine marker *-at* to indicate the *nomen unitatis* of collectives, e.g. *dajāj/dajāj-at-* ‘hen’, *naml/naml-at-* ‘ant’, *bayḍ/bayḍ-at-* ‘egg’, *ward/ward-at-* ‘flower’, *šajar/šajar-at-* ‘tree’, *tuffāḥ/tuffāḥ-at-* ‘apple’. These unit nouns are consistently grammatically feminine, whereas their base generic collective nouns are usually treated as singular masculine. Conversely, the suffix *-at* is also used to derive collectives out of participles, which is conspicuously the precise opposite of singulatives, e.g. *kāfir/kafarat-* ‘unbeliever/s’, *sāḥir/saḥarat-* ‘magician/s’, *xāʾin/xawanat-* ‘traitor/s’. This derivation is quite productive with the plural pattern *ʾafāʿilat-*, which is found both in common nouns, e.g. *ʾasātiḏat-* ‘teachers’, *ʿamāliqat-* ‘giants’, *ʿabāqirat-* ‘geniuses’, and in ethnic and other groups, e.g. *ʾafāriqat-* ‘Africans’; grammatically, these are treated as masculine (Badawi a.o. 2004:92).

Another very important function of *-at* is intensification, which is found primarily in a special category of words usually denoting an excess of a certain feature in a male referent. These nouns usually follow the emphatic pattern *faʿʿālat-*, as in *ʿallāmat-* ‘a man of great learning’, *raḥḥālat-* ‘a widely traveled man, an explorer’, and also *nābiġat-* ‘a genius’. There are other words of this type that are not commonly used today, such as *maddāḥat-* ‘a man who praises a lot’, *nawwāḥat-* ‘a great mourner, a great elegiac poet’, *ʿayyābat-* ‘a great fault-finder’, *ʾimmāʿat-* ‘a characterless person’ (cf. Idriss 1999:42). In addition, the suffix *-at* is also used to form abstract nouns, e.g. *ʿunṣuriyyat-* ‘racism’, from *ʿunṣuriyy-* ‘racist’, and diminutives, as in *ʾuḏun/ʾuḏaynat-* ‘ear’.

It is important to note that this multiplicity of functions of the feminine marker is not limited to *-at*. It has been noted that the other feminine endings, *āʾ* and *ā*, are also found in some broken plurals, e.g. in *šuʿarāʾ* ‘poets’, *fuqarāʾ* ‘poor people’, *šuhadāʾ* ‘martyrs’, *ʿulamāʾ* ‘scholars’, *kuramāʾ* ‘generous people’, and *sukārā* ‘drunks’.

Based on the multifunctional usage of the so-called feminine marker, it is recognized today that Arabic, and Semitic gender in general, might have its origin in non-gender nominal classes. It seems that originally, Semitic languages relied on size and importance to distinguish between various classes of words: “those denoting large, important objects on the one hand, and those denoting small, insignificant objects on the other. The latter category also included such words as diminutives, abstract nouns and collectives; words in this category were marked with suffixes as *-t, -ā, -ay, -āʾu*, which later became the suffixes for the feminine gender” (Versteegh 1997:18). This indicates that the suffix *-at* did not develop its function as a marker of feminine gender until a relatively late period in history.

It is also evident that there is a close connection between the feminine gender and the grammatical category of number. This association of gender and number in Arabic has informed recent discussions about the function of the grammatical category of gender crosslinguistically (see Unterbeck 2000). In an attempt to go beyond the agreement-creating effect that has dominated discussions of gender since Corbett’s (1991) definition of gender, Weber refers to Arabic to demonstrate the true function of gender besides classifying nouns according to their suffixes and creating agreement. He proposes that “gender has the function of qualitatively more precisely defining a quantity. Gender offers the opportunity to refine the crude perspective of number – singular versus plural – into distributive versus collective plural. It is this aspect of quantity that links gender so closely to number” (Weber 2000:506).

Early Arab grammarians recognized that the ‘feminine’ gender is more complex than the masculine. They distinguished between three types of feminines: *muʾannaṯ ḥaqīqī* ‘true feminine’, *muʾannaṯ majāzī* ‘metaphorical feminine’, and *muʾannaṯ lafḏ̣ī* ‘morphological feminine’ (cf. Ibrahim 1973). Under the category of true feminines they include marked and unmarked nouns and proper names denoting biological females, e.g. *ʾumm* ‘mother’, *naḥlat-* ‘bee’, *hajar* ‘Hajar [fem. proper name]’. The metaphorical feminine includes inanimate nouns with or without a feminine ending, such as *jannat-* ‘heaven’, *šams* ‘sun’, and the morphological feminine includes masculine nouns that have a feminine ending, e.g. *xalīfat* ‘caliph’, *ʿumdat-* ‘mare’ and several proper names for men, e.g. *ṭalḥat-*, *ʿubādat-*, *ʿantarat-*, *ʿuqbat-*. These names were very common in the pre- and early Islamic era, and some are still widely used today, e.g. *ḥamzat-*, *riḍā*, *mūsā*, *ʿīsā*. Grammatically, these nouns are always masculine despite their seemingly feminine ending, e.g. *xalīfa ʿādil* ‘just caliph’. The early Arab grammarians’ approach to gender has been criticized as being merely descriptive, its primary objective being to list and classify nouns into gender classes without explaining how lexical and grammatical genders in Arabic emerged, developed, and changed.

3. Lexical-semantic gender in Arabic

In pre-Classical Arabic, natural gender nouns were commonly formed by using different words for male and female humans and animals rather than by adding *-at*. In this respect, neither the real feminine nor the masculine were marked for gender (see Table 1).

Table 1. Natural gender nouns

masculine	feminine	
ʾab	*ʾumm*	‘father/mother’
šayx	*ʿajūz*	‘old man/woman’
ḥiṣān	*faras*	‘stallion/mare’
ḥimār	*ʾatān*	‘[male/female] donkey’
xuzāz	*ʾarnab*	‘[male/female] rabbit’

The fact that this category of words where gender is most natural was not overtly marked, particularly in the feminine, has led scholars to conclude that the feminine marker *-at* did not develop from naturally female beings in Arabic, and in Semitic languages in general (see Ibrahim 1973:40–50 for a review of several theories on the developmental stages of *-at*). It is worth noting, however, that there are pairs in this category where the masculine and the feminine are lexically different but the words denoting the female do have the suffix *-at*, such as *qird* ‘he-monkey’ vs. *qiššat-* ‘she-monkey’ and *ḏiʾb* ‘he-wolf’ vs. *ṯurmulat-* ‘she-wolf’. There exists also a less common category of natural gender nouns where the masculine is marked, e.g. *ʿuqrubān* ‘he-scorpion’, whereas the feminine is unmarked *ʿaqrab* ‘she-scorpion’ (examples from Suyūṭī, *Muzhir* II, 222, cited in Idriss 1999).

Furthermore, there is another class of words denoting natural gender in Arabic in which one word refers to both the masculine and the feminine members of a pair. Some of these unigender words are feminine in form since they are marked with *-at*, e.g. *ḥayyat-* ‘snake’, *naʿāmat-* ‘ostrich’, while others have a zero suffix, e.g. *ʾinsān* ‘human being’, *baʿīr* ‘camel’, *zawj* ‘husband/wife’, but grammatically these tend to remain faithful to their morphological form (examples from Suyūṭī, *Muzhir* II, 222, cited in Ibrahim 1973).

Throughout the history of Arabic a gradual shift from lexical gender to morphological gender demarcation has been taking place. One of the common changes is from unmarked feminine to masculine such as *ʾarnab* and *ʿaqrab*, which have shifted gender from feminine to masculine while their original masculine counterparts are hardly heard today. This change is attributed to a tendency for simplification of the linguistic system (cf. Ibrahim 1973; Procházka 2004). The symmetry of the nominal system is enhanced because unmarked feminine nouns are brought in line with the rest of the nouns, which are unmarked when they are masculine and end in -at- when they are feminine. Procházka (2004) found this shift from feminine to masculine to be the most common gender change in modern Arabic dialects.

The other historical change that has affected natural gender is from unmarked feminine to marked feminine, evidenced by the use of *-at* to generate feminine nouns by attaching it to already feminine nouns. This lexical hypercharacterization seems to be quite common in gender languages, particularly with unmarked feminine or gender neutral nouns denoting female persons and has been explained by Jespersen as the outcome of “a natural tendency to bring about conformity between gender and sex” (1924:230; cited in Ibrahim 1973:53). The most frequently cited naturally feminine words that have undergone hypercharacterization in the spoken dialects are Classical Arabic *ʿarūs* ‘bride’ and *ʿajūz* ‘old woman’, which have become *ʿarūsa* ‘bride’ and *ʿajūza*. This change seems to be common in the Maghreb, probably because it is an old feature of Sicilian and Andalusian Arabic (cf. Drozdík 1973:228; Agiùs 1991:2; Procházka 2004). The word *ʿarūs* is particularly interesting because originally it was a gender-neutral word referring to both ‘bride’ and ‘groom’, but in Literary Arabic today the word *ʿarūs* is reserved for ‘bride’ while *ʿarīs* is used for ‘groom’. It has been observed that this is the only pair where the masculine and the feminine are rendered through vowel alternation rather than suffixation in Literary Arabic, but this is not the case in spoken Arabic, where the feminine suffix *-at* is attached to *ʿarūs* as in *ʿarustak* 'your bride' and *ʿarustēn* 'two brides' (cf. Ibrahim 1973:83). Other words of the gender-indifferent type that have developed a feminine form in later phases of Arabic include *zawj* 'husband/wife', which has developed *zawjat-* for 'wife' and kept the unmarked *zawj* for 'husband', *ʾinsānat-* 'female human being' from *ʾinsān*, and also the passive participle *qatīl* 'killed [fem./masc.]', which now has the feminine *qatīla* 'killed [fem.]'. A diachronic account of this change from unmarked to marked *-at* suggests that in pre-Classical Arabic the change must have started in words that belong to the root system or adjectives, e.g. *xārij/xārijat-* 'going out [masc./fem.]', and later through a process of analogy spread to words belonging to the noun system, e.g. *ḥimārat-* from the masculine *ḥimār* (instead of *ʾatān*) (Hämeen-Anttila 2000).

There is a tendency for words that denote inherently feminine states and experiences to resist lexical hypercharacterization, e.g. *murḍiʿ* 'breastfeeding woman' and *ḥāʾiḍ* 'menstruating', *ḥāmil* 'pregnant', *ṭāmiṯ* 'menstruating', *ʿāqir* 'barren', *ṭāliq* 'divorced woman', *nāhid* 'full-breasted', *kāʿib* 'buxom', *ʿānis* 'spinster', *nāšiz* 'recalcitrant wife' (examples from Idriss 1999). In this category of words, we find a set of adjectives with feminine ending as well, such as *ḥublā* 'pregnant', *ʿaḏrāʾ* 'virgin', *muʿallaqat-* 'stranded woman [between marriage and divorce; cf. Q. 4/129]'. However, even words that denote specifically feminine states are being subjected to *-at* as, for instance, *murḍiʿa* coexisting with *murḍiʿ* in Literary Arabic, and *ḥaml-a* instead of Classical Arabic *ḥāmil* in Moroccan Arabic. It seems that among natural gender nouns the word *ʾumm* 'mother' and *bint* 'girl, daughter' are the most resistant to lexical hypercharacterization because they are very old basic vocabulary items, which are not derived from a root (cf. Hämeen-Anttila 2000). This origin has also been suggested for unmarked inanimate feminines.

Gender assignment in animate nouns is less problematic than in inanimate nouns, because of the former's association with biological gender. Marked inanimate feminines such as *qaryat-* 'village', *maxaddat-* 'pillow' are overtly marked with *-at*, while inanimate masculine nouns have a zero suffix as in *jabal* 'mountain', *lawn* 'color', and as such, neither poses a real problem. Ambiguity arises with the large number of inanimate feminines without an overt feminine marker, such as *šams* 'sun', *ʾarḍ* 'earth', *dār* 'house', *ḥarb* 'war', *nār* 'fire', *sūq* 'market', *ḥāl* 'situation, state', *xamr* 'wine', *balad/bilād* 'country', *ṭarīq* 'road', *rūḥ* 'soul', *nafs* 'self', *biʾr* 'well', *rīḥ* 'wind', *sikkīn* 'knife', *faʾs* 'axe', *kaʾs* 'cup', as well as some body parts that come in pairs, *ʿayn* 'eye', *ʾuḏun* 'ear', *yad* 'hand', *rijl* 'leg' (see Fleisch 1961:311–338 for Classical Arabic; Badawi a.o. 2004:93 for Modern Standard Arabic). Early Arab grammarians, especially in the 8th and 9th centuries, provided strictly descriptive 'explanations' for unmarked inanimate feminines. According to them, these words are feminine, first "because they were heard from the Arabs as feminine" (*li-ʾannahā sumiʿat ʿan al-ʿArab muʾannaṯa*); second, because the diminutives of these nouns are feminine, such as *ʿayn/ʿuyaynat-* 'eye', *sūq/suwayqat-* 'market'; and third, because of their association either with a synonym or by omission – in meaning rather than in ending (cf. Ibrahim 1973:22–23). An instance of association with a synonym is the masculine *lisān* 'tongue', which becomes feminine when it is associated with the feminine *luġat-* 'language'. Association by omission applies, for instance, to names of cities, which acquire feminine gender because the word *madīnat-* 'city' is feminine, and even when this word is omitted, it is still implied.

In fact, unmarked inanimate feminines are the category that has generated by far the most problems for Classical Arab grammarians and lexicographers. They list a total of about 240 words but agree on 100 words only, which all grammarians treat as feminine (cf. Ibrahim 1973:47). They disagree about the gender assignment of approximately 140 words, which some list as feminine, others as masculine or masculine/feminine depending on the dialect of their Bedouin informants.

Among unmarked nouns that are bi-gender, i.e. that can be treated both as a masculine and a feminine, one finds *silm* 'peace', *ṭarīq* 'road', *darb* 'street', *ḥānūt* 'shop', *ʾanf* 'nose'. A comparison of old and new Arabic texts shows that early Arabs had a far greater number of bi-gender unmarked inanimate words, which is still the case in the spoken dialects, unlike

Modern Standard Arabic (Idriss 1999). Idriss notes that early Arabs used to treat collectives such as *ʾunās* ‘people’, *tīn* ‘figs’, *baqar* ‘cows’, *ḥamām* ‘pigeons’, *tamr* ‘dates’, as masculine and feminine interchangeably, referring to *sūrat an-Naḥl* in the *Qurʾān*. He argues that this category of collectives merged toward the masculine as a result of “the gender demarcation process that turned open categories of nouns to masculine only” (Idriss 1999). Idriss refers to extralinguistic factors to explain the change from unmarked feminines to masculines: “This lingo-cultural process reached its climax when the Arabs decided to dethrone their archdeity, the feminine ʿAllaat, and worship Allaah, the supreme ‘He’ instead.... It is more likely that, as far as gender is concerned, Islam came to conclude, rather than initiate or even expedite, a process of masco-centralization that had been going on for many centuries” (Idriss 1999:40). These types of explanations are reminiscent of the speculative theory of gender, which is not accepted by formal linguists who explain such a change from unmarked feminine to masculine as a tendency toward simplifying the linguistic system, zero suffix meaning masculine and *-at* feminine.

In a recent comparative examination of gender changes of unmarked feminines in modern Arabic dialects, Procházka (٢٠٠٤) found that the loss of unmarked feminines is not as common as has often been claimed in Arabic dialectology. He notes, for instance, that the tendency to mark Classical Arabic unmarked feminine nouns in the spoken dialects is shown in only two body parts, *kibda* ‘liver’ and *sinna* ‘tooth’, noted by Heath (٢٠٠٠) for Moroccan Arabic; and in a few other words for tools and vessels such as *qidra* ‘pot’ and *sikkīna* ‘knife’, observed in urban dialects of the eastern Mediterranean. Contrary to earlier observations, only four body parts have feminine gender in the spoken dialects: *ʿayn, ʾuḏun, yad*, and *rijl*. The majority of the dialects, Procházka adds, tend to preserve the gender of basic words, such as those for ‘sun, moon, earth, hand, eye, fire’, which is apparently a universal linguistic tendency. He also notes unusual changes in some basic and frequent vocabulary items, despite the claimed universal linguistic tendency for stability, such as *šamsa* instead of *šams* in eastern and Upper Egypt, as well as in the Anatolian dialect of Daragözü.

The dialects also show gender changes from forms that were masculine in Classical Arabic to feminine forms as the outcome of replacement, semantic and phonological analogy, and influence from a substrate or adstrate language (Procházka 2004). A case of gender replacement may be seen in eastern Arabia, where Classical Arabic feminine *nār* ‘fire’ is replaced by *ḍaww* (< *ḍawʾ*), which is now treated as a feminine despite its masculine gender in Classical Arabic (example from Holes 2001:313, cited in Procházka 2004). Semantic analogy comes about through attraction to the gender of synonyms or semantically related words, as for instance the word *markib* ‘ship’, which has attracted its feminine gender in Egyptian and Sudanese dialects from the feminine of *falūka* ‘boat’ (or Classical Arabic *fulk*, which is feminine) and/or *safīna*. On the other hand, masculine words may become feminine due to phonological analogy, when they resemble feminine nouns in shape. One example comes from Cilician Arabic *kirsi* ‘chair’, which has become feminine because of its feminine ending (cf. Procházka 2002:117; 2004). Finally, substrate or adstrate language influence on gender transformation is difficult to prove (see Diem 1979), but Procházka (2004) notes the obvious Berber influence on the Maghribi Arabic dialects in the feminine gender of *ṣūf* ‘wool’ and in food names like *ʿasal* ‘honey’, and the case of *paxr* (< *baḥr* ‘sea’) in Cypriot Arabic, which is feminine because of the feminine gender of Greek *thálassa*. In addition, Procházka’s comparative study has shown that the Bedouin dialects are probably not more conservative than the sedentary dialects in preserving the Classical Arabic gender of nouns, although he notes that they tend to retain the feminine gender in nouns associated with nature, whereas the sedentary dialects tend to maintain feminine nouns related to trade.

In an attempt to account for unmarked inanimate nouns, Idriss (1999) proposes that there is a tendency for the feminine to denote wider semantic references than the masculine, e.g. *dār* (fem.), *bayt* (masc.), and *manzil* (masc.), which all have the meaning of ‘house’, but only the feminine *dār* is used in the expression *dār al-ʾIslām*. The same applies to *ḥubb* (masc.) ‘love’ and *maḥabbat-* (fem.), the former being used for romantic love, while the latter includes all kinds of love.

4. GENDER AGREEMENT IN ARABIC

The Arabic language shows gender → agreement between subject and verb, between head noun and modifiers, in pronominal anaphora, and in the possessive particle with the possessed noun. The only exception to gender agreement is the definite and indefinite marking. Unlike French and Spanish articles, the Arabic definite article *al-* is gender-indifferent, e.g. *al-qalam* 'the pen [masc.]' and *al-waraqat-* 'the paper [fem.]', *al-ḥarb* 'the war [fem.]', and the same goes for the Arabic indefinite ending *-n*.

4.1 *Pronominal and verbal agreement*

Personal pronouns, which include subject, object, and possessive pronouns, are marked masculine or feminine. Personal pronouns agree with the gender of the head noun. Subject pronouns in all varieties of Arabic fall under two types, independent and dependent pronouns. Independent subject pronouns such as *huwa* 'he' and *hiya* 'she' are primarily used for emphasis or clarity, since the verb form itself usually indicates its subject. Dependent subject pronouns are prefixes or suffixes that attach to verbs to mark the person of the subject, e.g. *-at* (3rd pers. fem. sg. perfect), *ya-* (3rd pers. masc. sg. imperfect). In all varieties of Arabic, there is no gender distinction in the 1st person singular and plural in subject, object and possessive pronouns (e.g. *ʾana* 'I [masc./fem.]' and *naḥnu* 'we [masc./fem.]'), as well as in the 2nd and 3rd person dual (*ʾantumā* 'you both [masc./fem.]' and *humā* 'they both [masc./fem.]'). The only exception is found in some Gulf Arabic dialects that have a fem. 1st pers. *ani* (Zaborski 1995). For the other persons, there is usually a marking of the gender.

Several gender changes have taken place in the pronominal and verbal systems between Classical Arabic and the spoken Arabic dialects. For instance, in Classical Arabic masculine and feminine forms are used for the 2nd and 3rd person plural of pronouns and verbs, but most spoken sedentary varieties have lost this gender distinction, the masculine form being used for both genders. Bedouin dialects, on the other hand, tend to be more conservative than the sedentary dialects with respect to this gender distinction, for instance, in the dialect of the Najd, which distinguishes between *ktibaw* 'they [masc.] wrote' and *ktiban* 'they [fem.] wrote', where related dialects in Iraq have only *ktibaw*, like all sedentary dialects in the Arab world (cf. Versteegh 1997:100).

Furthermore, some urban varieties of North Africa have lost the gender distinction in the 2nd person singular as well. Tunisian sedentary dialects, for instance, have neutralized gender distinction in the 2nd person in both the pronominal and the verbal systems toward the feminine, the feminine form *inti* being used to address a male or a female. In the verbal system, gender distinction has been neutralized toward the masculine. Thus, the suffix *-i* that marks the feminine is absent in the imperfect, perfect, and imperative (Gibson 1996). On the other hand, in sedentary dialects of Moroccan Arabic, gender distinction has been neutralized toward the feminine in the perfect, e.g. *klit-i* 'you ate [fem./masc.]'. Both Tunisian and Moroccan Bedouin dialects maintain the distinction, and in this case they are similar to Classical Arabic and Modern Standard Arabic and to the eastern Bedouin dialects.

Possessives in Arabic attach to nouns and prepositions and are differentiated by the gender of the referent, e.g. *maktabu-hu* 'his office' and *maktabu-hā* 'her office', *min-kum* 'from you [masc. pl.]' and *min-kunna* 'from you [fem. pl.]'. In the spoken dialects, there is variation between synthetic and analytic expression of possession, but the gender marker is attached to the genitive exponent, instead. In Egyptian Arabic, for instance, there is variation between *ilmaktab bitāʿu/bitāʿha*, in Lebanese Arabic *ilmaktab tabaʿu/tabaʿha*, and in Moroccan Arabic *lməktəb dyalu/dyalha*.

Demonstratives in Arabic agree with the gender of their head nouns in the singular. In Classical Arabic and Modern Standard Arabic, there is a gender distinction between *hāḏā* (masc.), *hāḏihi* (fem.) 'this', and *ḏālika* (masc.), *tilka* (fem.) 'that', which has also been preserved in the spoken dialects, albeit in slightly different forms (→ demonstrative pronouns). This is not the case with respect to relative pronouns, where most spoken Arabic dialects have lost Classical and Modern Standard Arabic gender (and number) distinction in relative pronouns, *allaḏī* (masc. sg.), *allatī* (fem. sg.), *allaḏīna* (masc. pl.), *allawātī* (fem. pl.) being replaced by the neutral form *(i)lli*.

4.2 *Adjectival and verbal agreement*

In his discussion of agreement between subject and verb and between head noun and modifier in Arabic, Ferguson (1989) distinguishes between 'strict', 'deflected', and 'equivocal' types of agreement. Strict agreement occurs particularly with head nouns denoting human beings and checks all noun modifiers for gender in the singular and plural, as in (1)–(4).

(1) *hāḏā* *huwa* *l-muhandisu*
this-Masc.Sg. he the-engineer-Masc.Sg.
l-faransiyyu *llaḏī*
the-French-Masc.Sg. who-Masc.Sg.
ṣammama *bināyata-hu*
designed-3Masc.Sg. building- his
'This is the [male] French engineer who designed his building'

(2) *hāḏihi* *hiya* *l-muhandisatu*
this-Fem.Sg. she the-engineer-Fem.Sg.
l-faransiyy-atu *llatī*
the-French-Fem.Sg. who-Fem.Sg.
ṣammama-t *bināyata-hā*
designed-3Fem.Sg. building-her
'This is the [female] French engineer who designed her building'

(3) *hāʾulāʾi* *hum*
this-Masc.Pl. they-Masc.Pl.
al-muhandis-ūna *l-faransiyy-ūna*
the-engineer-Masc.Pl. the-French-Masc.Pl.
llaḏīna ṣammam-ū *bināyata-hum*
who-Masc.Pl. designed-3Masc.Pl.
building-their-Masc.Pl.
'These are the [male] French engineers who designed their building'

(4) *hāʾulāʾi* *hunna* *l-muhandis-ātu*
this-Fem.Pl. they-Fem.Pl. the-engineer-Fem.Pl.
l-faransiyy-ātu *llawātī*
the-French-Fem.Pl. who-Fem.Pl.
ṣammam-na *bināyata-hunna*
designed-3Fem.Pl. building-their-Fem.Pl.
'These are the [female] French engineers who designed their building'

With nonhuman singular referents, gender agreement is often neglected, for instance in menus *qahwa baladī* 'coffee country-style', in fashion terms like *malābis jāhiza rijālī wa-ḥarīmī* 'clothes for men and women', and also in some loanwords *al-fatāt al-mudirn* 'the modern girl'. Similarly, some adjectives do not take a feminine form, e.g. *xāʾina laʿūb* 'treacherous flirt', *al-insāna al-ḥanūn* 'the tender person', *al-ʿaqaba al-kaʾūd* 'insurmountable obstacle' (examples taken from Badawi a.o. 2004:106).

Unlike strict agreement, both 'deflected' and 'equivocal' types of agreement are associated with plural controllers only. In deflected agreement, a plural nonhuman head noun or subject often requires a feminine singular form in adjectives, pronouns, and verbs, as in (5).

(5) *al-mudunu* *l-ʾajnabiyy-atu* *llatī*
the-cities the-foreign-Fem.Sg. that
zurt-u-hā *kānat*
FemSg. visit-1Sg.-Obj3Fem.Sg. was
barīd-atun
cold-Fem.Sg.
'The foreign cities that I visited were cold'

However, there are some variations between strict or plural agreement and deflected or feminine singular agreement with human collective nouns, e.g. *nās muxtalifa* (fem. sg.)/*muxtalifūna* (masc. pl.), *šuʿūb maskīna* (fem. sg.)/*masākīn* (masc. pl.) (Badawi a.o. 2004:104). Other head nouns that show this variation are broken plurals of human reference, such as *rusul* 'messengers', *junūd* 'soldiers', *ʾabṭāl* 'heroes', *ʾaqārib* 'relatives', and *rijāl* 'men' (Belnap 1991).

In equivocal agreement, an initial verb precedes an indefinite nonhuman subject and may have three alternative patterns of agreement. Ferguson (1989) illustrates equivocal agreement from Damascene Arabic, where the initial verb can be masculine singular *ʾajā-na makatīb ktīr*, feminine singular *ʾajat-na makatīb ktīr*, or plural *ʾajū-na makatīb ktīr* 'many letters reached us'.

It is well established that the Arabic dialects are more conservative than Literary Arabic (which uses categorically feminine singular agreement with head nouns denoting nonhumans), and they are much closer to Old Arabic as far as variable agreement patterns are concerned. Belnap and Gee (1994) attribute the transition to the new rule of categorical feminine singular agreement with nonhuman heads to overgeneralization by second-language learners. Nonnative writers, they argue, tended to resort to avoidance strategy of the variable and complex patterns of Old Arabic, because they were writing in a second language and avoided any questionable usage.

Yet, a question that has raised much controversy is whether feminine singular agreement with head nouns denoting nonhumans is an innovation in the Arabic dialects (Versteegh 1984), or whether it is a remnant from Old Arabic (Ferguson 1989). Versteegh (1984:103–105) suggests that deflected agreement was replaced by strict agreement in Classical Arabic and that the presence of deflected agreement in New Arabic is the result of decreolization or a classicism. Against Versteegh, Ferguson (1989) points out that the pure colloquial plural pattern in Damascene Arabic *ʾajūna makatīb ktīr* not only did not disappear as a result of classicization, as one might expect, but is even more popular than the patterns *ʾajāna makatīb ktīr/ʾajatna makatīb ktīr* that are found in Classical Arabic. Versteegh (1997) notes that a quantitative study of agreement patterns is important to support Ferguson's observations and apparently agrees with the argument that "it should not be taken for granted that all movement on the continuum between dialect and standard is upwards". He adds, however, that "in some contexts, it is perfectly possible that there is a movement towards the dialect pattern. In other cases, interference from the standard language leads to a redistribution of grammatical functions. In the case of the agreement in Syrian Arabic, there probably is a semantic difference in that the plural is used for countable entities, whereas the feminine singular is used for non-countable or collective plurals" (Versteegh 1997:111).

6. Conclusion

Arabic gender is a complex category that has attracted much-deserved attention from medieval and contemporary scholars. Several questions regarding its origin, development, function, variation, and change have been thoroughly studied, and we have come a long way toward understanding some of the complexities of Arabic gender. Yet, it seems that the jury is still out on several questions. In particular, one general question about gender in Arabic still puzzles linguists working on the grammatical category of gender, namely "how the quantitative collective/singulative notion is related to the notion of female sex.... This is an open question and it is – also beyond Arabic – one of the crucial questions of the gender discussion. An answer to this question would explain how the general count-mass distinction of nominal classification has been expanded to the double feature of [+count-mass] [+sex]" (Unterbeck 2000b:xxxii).

It remains to be seen whether the answer to this question will be found in Arabic or in another language.

Bibliographical references

Primary source

Suyūṭī, *Muzhir* = Jalāl ad-Dīn ʾAbū l-Faḍl ʿAbd ar-Raḥmān ibn ʾAbī Bakr as-Suyūṭī, *al-Muzhir fī ʿulūm al-luġa*. Ed. Muḥammad ʾAḥmad Jād al-Mawlā Beg, ʿAlī Muḥammad al-Bagāwī, and Muḥammad ʾAbū l-Faḍl ʾIbrāhīm. 2 vols. Sayda and Beirut: al-Maktaba al-ʿAṣriyya, 1406/1986.

Secondary sources

Agiùs, Dionisius. ١٩٩١. "Focus of concern in Ibn Makkī's *Taṯqīf al-lisān*: The case of gender in the medieval Arabic of Sicily". *Proceedings of the Budapest Colloquium of Arabic Grammar*, ed. Kinga Dévényi and Tamás Iványi, 1–7. Budapest: Eötvös Loránd University and Csoma de Kőrös Society.

Badawi, El-Said, Michael G. Carter, and Adrian Gully. 2004. *Modern Written Arabic: A comprehensive grammar*. London and New York: Routledge.

Belnap, R. Kirk. 1991. *Grammatical agreement variation in Cairene Arabic*. Ph.D. diss., University of Pennsylvania.

——. 1999. "A new perspective on the history of Arabic variation in marking agreement with plural heads". *Folia Linguistica* 33:2.169–185.

—— and John Gee. 1994. "Classical Arabic in contact: The transition to near-categorical agreement patterns". *Perspectives on Arabic linguistics*, VI, ed. Mushira Eid, Vicente Cantarino, and Keith Walters, 121–149. Amsterdam and Philadelphia: J. Benjamins.

Brockelmann, Carl. 1908. *Grundriss der vergleichenden Grammatik der semitischen Sprachen*, I. Berlin: Reuther und Reichard.

Corbett, Greville G. 1991. *Gender*. Cambridge: Cambridge University Press.

Diem, Werner. 1979. "Studien zur Frage des Substrats im Arabischen". *Der Islam* 56.12–80.

Drozdík, Ladislav. 1973. "Grammatical gender in Arabic nouns". *Graecolatina et Orientalia* 5.217–248.

——. 1974. "Collective and unit nouns as sex-gender pairs in Arabic". *Asian and African Studies* 10.41–48.

——. 1998. "Functional variations of the so-called feminine gender in Arabic". *Asian and African Studies* 7:1.23–44.

Eid, Mushira, Vicente Cantarino, and Keith Walters (eds.). 1994. *Perspectives on Arabic linguistics*, VI. Amsterdam and Philadelphia: J. Benjamins.

Eid, Mushira and Dilworth Parkinson (eds.). 1996. *Perspectives on Arabic linguistics*, IX. Amsterdam and Philadelphia: J. Benjamins.

Féghali, Michel and Albert Cuny. 1924. *Du genre grammatical en sémitique*. Paris: P. Geuthner.
Ferguson, Charles A. 1989. "Grammatical agreement in Classical Arabic and the modern dialects: A response to Versteegh's pidginization hypothesis". *al-ʿArabiyya* 22.5–17.
Fleisch, Henri. 1961. *Traité de philologie arabe*. I. *Préliminaires, phonétique, morphologie nominale*. Beirut: Imprimerie Catholique.
Gibson, Michael. 1996. "Implicational dialectology: Second person pronouns and suffixes in Tunisian Arabic". *Perspectives on Arabic linguistics*, IX, ed. Mushira Eid and Dilworth Parkinson, 95–114. Amsterdam and Philadelphia: J. Benjamins.
Haak, Martine, Rudolf de Jong, and Kees Versteegh (eds.). 2004. *Approaches to Arabic dialects: A collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday*. Leiden: E.J. Brill.
Hämeen-Anttila, Jaakko. 2000. "Grammatical gender and its development in Classical Arabic". Unterbeck (2000:595–608).
Heath, Jeffrey. 2002. *Jewish and Muslim dialects of Moroccan Arabic*. London: RoutledgeCurzon.
Holes, Clive. 2001. *Dialect, culture and society in eastern Arabia*. I. *Glossary*. Leiden: E.J. Brill.
Ibrahim, H. Mohammed. 1973. *Grammatical gender: Its origin and development*. The Hague and Paris: Mouton.
Idriss, Mohammed. 1999. "The cell and the umbrella: Gender in the Arabic language". *Zeitschrift für Arabische Linguistik* 37.32–53.
Jespersen, Otto. 1924. *The philosophy of grammar*. London: Allen and Unwin.
Moscati, Sabatino (ed.). 1964. *An introduction to the comparative grammar of the Semitic languages: Phonology and morphology*. Wiesbaden: O. Harrassowitz.
Procházka, Stephan. 2002. *Die arabischen Dialekte der Çukurova (Südtürkei)*. Wiesbaden: O. Harrassowitz.
——. 2004. "Unmarked feminine nouns". *Approaches to Arabic dialects: A collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday*, ed. Martine Haak, Rudolf de Jong, and Kees Versteegh, 238–262. Leiden: E.J. Brill.
Speiser, E.A. 1936. "Studies in Semitic formatives". *Journal of the American Oriental Society* 56.22–46.
Unterbeck, Barbara (ed.). 2000a. *Gender in grammar and cognition*. Berlin and New York: Mouton de Gruyter.
——. 2000b. "Gender: New light on an old category". Unterbeck (2000a:xv–xlvi).
Versteegh, Kees. 1984. *Pidginization and creolization: The case of Arabic*. Amsterdam: J. Benjamins.
——. 1997. *The Arabic language*. Edinburgh: Edinburgh University Press.
Weber, Doris. 2000. "On the function of gender". Unterbeck (2000:496–509).
Wright, William. 1896. *A grammar of the Arabic language*. 3rd edition. Cambridge: Cambridge University Press.
Zaborski, Andrzej. 1995. "First person pronouns in Arabic in the light of Arabic and Hamitosemitic dialectology". *Studia Orientalia* (Helsinki) 75:289–294.

Atiqa Hachimi (University of Florida)

Genitive Construction → Annexion; ʾIḍāfa

Gilit Arabic → Iraq

Glide

1. Definition

Glides in Arabic are *w* and *y* (the *ʾalif* [aː] is not considered a glide here; unlike *w* and *y*, it does not appear in any phonological representation of any defective form). Unlike sound consonants, e.g. /b/, /t/, the glides, *w* and *y*, play a double role. They can be consonants, as in, for example, ***wajada*** 'he found, ***yusr*** 'ease', or long vowels, as in, for example, *taqūlu* 'she says', *yasīru* 'he moves on', or both.

Graphically speaking, Arabic transcription does not distinguish between long vowels [uː = و] and [iː = ي] and their respective glides [w = و] and [j = ي], since both are written the same way. When [و] or [ي] is vocalized, e.g. *raḍiya* 'he is satisfied', *saruwa* 'he left', *daʿwa* 'invitation', it is perceived as a glide; when the glide is preceded by [a] and is not followed by a vowel, it forms a diphthong, e.g. *qawl* 'saying, utterance', *sayr* 'trip, tour'.

Roots containing *w* or *y* are called *ʾuṣūl muʿtalla* 'weak roots' (→ *ṣarf*). This class of roots divides into two categories. The first category hosts a glide in the following contexts:

i. C1 'assimilated roots', e.g. *w-j-d* 'to find', *y-s-r* 'to be or become easy'
ii. C2 'hollow roots', e.g. *q-w-l* 'to say', *s-y-r* 'to move on'
iii. C3 'defective roots', e.g. *k-b-w* 'to stumble, slip', *k-f-y* 'to be enough'

The second category is doubly weak, showing glides in the following:

i. C1 and C3 'assimilated and defective roots', e.g. *w-l-y* 'to be near someone or something'
ii. C2 and C3 'hollow and defective roots', e.g. *l-w-y* 'to curve'

Weak forms are subject to *ʾiʿlāl* 'defectiveness' (→ *ʿilla*), the change that takes place in a word in

Table 1. Derivation rules for weak roots (after Bohas 1982:283)

1. Underlying representation	/qawaltu/	/bayaˁtu/
2. Template change	/qawultu/	/bayiˁtu/
3. Vowel truncation	/qwultu/	/byiˁtu/
4. Vowel transfer	/quwltu/	/biyˁtu/
5. Glide elision	/qultu/	/biˁtu/
6. Phonetic representation	[qultu] ‘I said’	[biʕtu] ‘I sold’

which a glide is subject to phonological processes, such as *qalb* ‘mutation’, → *ḥaḏf* ‘truncation’, and/or *ʾiskān* ‘vowel truncation’. This topic is treated in grammatical treatises such as Sībawayhi’s *Kitāb*, al-ʾAstarābāḏī’s *Šarḥ aš-Šāfiya*, and Ibn Jinnī’s *Munṣif* (→ *ṣarf*), and in modern studies such as Brame (1970), Fleisch (1979), Kouloughli (1979), Guillaume (1981, 1982), Bohas (1982, 1985), Angoujard (1984), Mokhlis (1997), and Chekayri and Scheer (2003, forthcoming).

2. THE PHONOLOGICAL REPRESENTATION OF WEAK FORMS

In major Arabic dictionaries, roots are listed in alphabetical order. In texts, though, forms may not give an indication as to the original root form. The identification of the correct root allows the user to find out about the root, its meaning, derivation, context of use, and so on. Mastering the derivation process is crucial to finding similarities between forms and retrieving the appropriate triradical root. Based on the comparison between weak and sound roots, Arab grammarians define circumstances under which a glide can persist, be changed, or be deleted, and they explain Arabic language structure by setting up rules governing its use, through → analogy (→ *ištiqāq*).

For illustration, Arab grammarians consider that the phonological representation of forms such as [yas̴ifu] ‘he describes’, [qaːla] ‘he said’, and [sirtu] ‘I went’ are underlyingly /yawṣifu/, /qawala/, and /sayartu/, respectively. Although the glides do not appear at the phonetic level, these forms are derived from triradical roots: /w-ṣ-f/, /q-w-l/, and /s-y-r/. Arab grammarians used analogy to restore the underlying level and retrieve the triradical root. The proof that these forms are derived from /yawṣifu/, /qawala/, and /sayartu/ resides in their *maṣdar* ‘nominal verb’: /waṣf/, /qawl/, and /sayr/ (cf. Ibn al-ʾAnbārī, *ʾInṣāf*, and Ibn Yaˁīš, *Šarḥ al-Mufaṣṣal*; Bohas 1982). In order to get the phonetic representation of /qawaltu/ and /bayaˁtu/, whose template is *faˁaltu*, these phonological forms are subject to the rules presented in Table 1.

The underlying representations, i.e. abstract forms, are determined on the one hand by the requirement that they must be joined to the forms phonetically attested by natural phonological processes, and on the other hand by the general systemic criteria of consistency and elegance (cf. Foley 1985).

3. PRESENCE VS. ABSENCE OF GLIDES

3.1 *Assimilated forms*

Assimilated forms host *w* or *y* in the first position of the template. The triradical assimilated roots with *w* total some 378 occurrences, while those with *y* total some 30 occurrences (cf. Al-Bawab a.o. 1996; Chekayri 1999, 2001). Chekayri (1999, 2001) and Chekayri and Scheer (1996) state that assimilated verbs with V_2 = [a] do not show the glide in imperfect forms. In fact, this is true for verbs only. The glide is present in almost all verbs with V_2 = /i/ or /u/. Consider the numerical proportions presented in Table 2.

Thus, in 248 out of 262 verbs (94.65%) with V_2 = [a], the glide is absent in imperfect forms. On the other hand, 104 out of 118 verbs (88.1%) with V_2 = /i/ or /u/ do present the glide in imperfect forms, with only 14 glideless verbs. This distribution is not fully complementary, but almost. In any event, it is significant enough to be regarded as nonaccidental.

Verbs with *y*, however, do not lose their first radical in the imperfect. The glides in assimilated verbs with V_2 = /i/ or /u/ that express involuntary actions are regarded as sound consonants. The mediopassive meaning, which is an important factor for the appearance

Table 2. Distribution of glides in assimilated roots

	$V_2 = \emptyset \rightarrow i$	$V_2 = a \rightarrow u$ $a \rightarrow a$	$V_2 = i \rightarrow a$	$V_2 = u \rightarrow u$	Total
Glide present in imperf.	1	11	64	40	116
Example	*waʿak/ya-wʿik* ‘to be very hot’	*wajaz/ya-wjuz* ‘to be brief, concise’ *wazar/ya-wzar* ‘to sin; to err, slip’	*waʾib/ya-wʾab* ‘to be angry’	*wafur/ya-wfur* ‘to be abundant’	
Glide absent in imperf.	205	43	12	2	262
Example	*wazan/ya-zin* ‘to weigh’	*wajad/ya-jud* ‘to find’ *wahab/ya-hab* ‘to give, donate’	*watig/ya-tag* ‘to be guilty’	*waxuš/ya-xuš* ‘to be vile’	

Table 3. Distribution of glides in hollow verbs

No. of verbs with *w* or *y*			Example			
Class	*w*	*y*	perf. 1st pers.	perf. 3rd pers. masc.	imperf. 3rd pers. masc.	Gloss
1	–	203	*sir-tu*	*sār-a*	*ya-sīr-u*	‘to move on’
2a	–	015	*hib-tu*	*hāb-a*	*ya-hāb-u*	‘to fear’
b	016	–	*xif-tu*	*xāf-a*	*ya-xāf-u*	‘to be frightened’
3	262	–	*lum-tu*	*lām-a*	*ya-lūm-u*	‘to blame’
4	001	–	*ṭul-tu*	*ṭāl-a*	*ya-ṭūl-u*	‘to be long’
5	052	–	*ḥawir-tu*	*ḥawira*	*ya-ḥwaru*	‘to have eyes with a marked contrast of white and black’
6	–	022	*ġayid-tu*	*ġayida*	*ya-ġyadu*	‘to be thin’
Total	331	240				
Total		571				

of the glide, makes the glides behave as sound consonants, e.g. *wajila/yawjalu* ‘to be afraid’, *wafura/yawfuru* ‘to be abundant’. Thus, they are not defective at all.

3.2 *Hollow forms*

Hollow forms host a glide in the second position of the template. The class is called ‘hollow’ because the glide never appears in any verb conjugation. The glide in hollow verbs is replaced by the long or the short vowel. Table 3 shows the numerical proportion of each class in Arabic.

In Classes 1–4, the long vowel in the middle is always [aː] in the perfect (3rd pers. sg. masc.). In the imperfect it is an [iː] if the glide in the middle is /y/; it usually has an [uː] if the second root letter is /w/. When it is an [aː] in the imperfect, the glide in the middle is /y/ or /w/.

Three basic generalizations can be made directly from the data in Table 3:

i. Hollow verbs possess one single vowel within the template.
ii. This vowel is short if and only if the suffix begins with a consonant; it is long if and only if the suffix is vowel-initial.

Table 4. Distribution of glides in defective verbs

	Verbs with		Example				
V_2	*w*	*y*	perf. active		imperf. active		Gloss
a	230		*danaw-tu*	1st sg.	*ya-dnu-u*	3rd m. sg.	'to be close'
ø	–	118	*ramay-tu*	1st sg.	*ya-rmi-i*	3rd m. sg.	'to throw'
i	–	132	*raḍiy-a*	3rd m. sg.	*ya-rḍa-a*	3rd m. sg.	'to be satisfied'
u	13	–	*saruw-a*	3rd m. sg.	*ya-sru-u*	3rd m. sg.	'to leave, go'
Total	493						

iii. Classes 5 and 6 (*ḥawira* and *ġayida* classes) stand apart: they disobey both of the above generalizations in that they always host two distinct vowels within the template and are completely insensitive to the kind of suffix present. In sum, they behave exactly like a sound triliteral.

Thus, whatever the quality of the root vowel in any instance of a hollow verb, its length is conditioned by a simple parameter: if the suffix begins with a vowel, it is long; otherwise, it is short.

As stated in the generalizations above, Classes 5 and 6 are in sharp contrast with the rest of hollow verbs. They are included in the discussion on hollow verbs because they possess a glide in C2. But it is this very fact that disqualifies them:

i. Classes 1–4 never show a glide on the surface in any position anywhere in the paradigm, whereas a glide appears in Classes 5 and 6 in all conjugated forms.
ii. The typical alternation in vowel length, controlled by the kind of suffix added, is not observed in Classes 5 and 6 at all.
iii. There is only one vowel within the template in Classes 1–4, against two vowels in Classes 5 and 6, e.g. *ḥawira* and *ġayida*. These two vowels are never long.

3.3 *Defective forms*

Defective forms have a glide in the third position of the template. They are characterized by the presence of a glide, e.g. [ramaj̯tu], [danaw̠tu], or a long vowel, e.g. [rama̠ː], [yadnu̠ː], [yarmi̠ː]. The only difference is the spelling of the long vowel [aː]: *ʾalif mamdūda* 'elongated *ʾalif* [ا aː]' for roots with /w/, *ʾalif maqṣūra* 'shortened *ʾalif* [ى iː]' for roots with /y/. The distribution of glides in defective verbs is as follows (Bayyūmī a.o. 1989:114ff.; see also Chekayri and Scheer 1996).

Since verbs with V_2 = [a] admit both [j] and [w] in numerically significant proportions, it does not seem possible to predict the glide from the second vowel of the stem. However, it may be observed that the distribution of the glide for V_2 = [a] is exactly parallel to that of V_2, not in the forms of the perfect but in those of the imperfect. Indeed, for verbs with perfect V_2 = [a], all and only those that exhibit a [j] show an [i] in imperfect V_2, and all and only the verbs whose glide is [w] present [u] in imperfect V_2. Hence, the glide is predictable for all defective verbs on the basis of the imperfect value of V_2. Distributionally, the glide thus obeys the same regularity as the derivation of the imperfect V_2 from its perfect input.

The deletion approach makes the prediction that a given verb may not exhibit more than one glide throughout the conjugation: if the weak root is recorded in the lexicon under one entry, e.g. *d-n-w*, the /w/ may not be replaced by the other glide, e.g. /y/, in a particular inflectional form. Consider the forms of the verb *d-n-w* 'to be close' (perf. vowel *a*, imperf. vowel *u*) in Table 5.

Table 5. Conjugation of the verb *d-n-w* 'to be close'

		Phonological representation	Phonetic representation
a.	perf. act. sg. 1st	*danaw-tu*	*danaw̠tu*
	perf. act. du. 3rd m.	*danaw-ā*	*danā*
b.	perf. pass. sg. 3rd m.	*duniw-a*	*duniy̠a*
c.	imperf. pass. pl. 3rd. fem.	*tu-dnaw-na*	*tu-dnay̠na*

If the lexical representation of this verb is *d-n-w*, [j] appearing in (5b) and (5c) begs the question. Whatever the nature of the glide in the phonological representation, the sequence [iwa] will appear in a passive form as [ija], i.e. [duniȷ̱a] instead of *[duniw̱a]. In other words, the perfect active form [danaw̱tu] and the perfect passive form [duniy̱a] are derived from the same root *d-n-w*. In case the odd [w] is adjacent to an [i], as in (5b), assimilation rules of the kind [/w/ → [j] / i__] are commonly invoked, for example by Ibn Yaʿīš (*Šarḥ al-Mufaṣṣal*), Brame (1970), Kouloughli (1979), Bohas (1982), Guillaume (1982), and others.

Chekayri and Scheer (1996) have established that the distribution of [w] and [j] among weak verbs is a function of V_2. Table 6 shows underlying and surface forms of V_2 for every verbal class, as well as the glide that is observed on its righthand side. The underlying identity of V_2 for the root *n-h-w* (vowels *a/a*) may not be determined in the usual way because the guttural inhibits → apophony to the effect that no alternation in imperfect forms occurs. The glide [j], however, points to a lexical $V_2 = \phi$ for this verb, hence it would be of the *ḍaraba* class.

Clearly, as seen in Table 6, the distribution of [w] and [j] is a function of V_2: if the latter is subject to variation as in active forms, the glide also alternates. If, on the other hand, V_2 hosts the invariable passive marker [i] (perf.) or [a] (imperf., i.e. the result of an apophonic derivation on perf. [i]), the glide would be [y] at the phonetic level.

Table 6. Underlying and surface forms of V_2 for each verb class and the accompanying glide

	Root	pf V_2	imperf.	Glide (G)
perf. act.	*d-n*-G	a	u	w
	r-m-G	ϕ	i	y
	s-r-G	u	u	w
	n-h-G	a	a	y
	r-ḍ-G	i	a	y
perf. pass.	*d-n*-G	i	a	y
	r-m-G	i	a	y
	s-r-G	i	a	y
	n-h-G	i	a	y
	r-ḍ-G	i	a	y

It could be argued that the invariable [y] in perfect passive forms is simply a copy of the passive [i]. However, in imperfect passive forms whose V_2 is [a], the invariable [y] would remain a mystery. If, on the other hand, it is assumed that in both cases the passive [i] is the derivational basis for apophonic glide appearance (and for the apophonic derivation of the imperf. pass. [a]), the invariable appearance of [y] is regular. Moreover, the following unifying generalizations can be made (Chekayri and Scheer 1996, 2003, forthcoming):

i. All imperfect forms are derived from their corresponding perfects.
ii. Perfect V_2 is the apophonic origin for both imperfect V_2 *and* glides.

4. Conclusion

In the preceding sections, generalizations have been established regarding the appearance of glides in the conjugation of defective verbs in Arabic forms. A very limited anomaly in the conjugational pattern is in fact the trace of a fundamental distinction in the Arabic verbal system that opposes verbs with V_2 = [a] on one side to those with V_2 = [i] and [u] on the other. This contrast has semantic, morphological, and phonological manifestations.

Defective forms, too, demonstrate the split between verbs with V_2 = [a] and those showing [i] and [u] in V_2. Indeed, verbs with V_2 = [a] lose the glide in most conjugations (perfect, imperfect, and imperative). The glides in verbs with V_2 = [i] or [u] that express involuntary actions, e.g. a quality, a defect, or a color, are analyzed as sound consonants, e.g. *ḥawira/yaḥwaru*, *ġayida/yaġyadu*, *wajila/yawjalu*, *wafura/yawfuru*. Thus, they are not defective at all.

It has been shown that the distribution of [y] and [w] is predictable. That is, the glide appearing in some forms of a given verb is the output of a derivation originating in the lexical vowel. The nature of this derivation is apophonic in the sense of Guerssel and Lowenstamm (1996). Considering the predictability of the glide from the lexical vowel, i.e. V_2, one may ask questions about the structure of weak roots: Are they triradicals or biradicals underlyingly (→ biradicalism)? If the appearance and absence of glides is predictable, why are weak forms considered to be triradical and not biradical? (see also → weak verbs)

Bibliographical references

Primary sources

ʾAstarābāḏī, *Šarḥ aš-Šāfiya* = Raḍī d-Dīn al-ʾAstarābāḏī, *Šarḥ Šāfiyat Ibn al-Hājib*. Beirut: Dār al-Kutub al-ʿIlmiyya, 1975.

Ibn al-ʾAnbārī, *ʾInṣāf* = Kamāl ad-Dīn ʾAbū l-Barakāt ʿAbd ar-Raḥmān ibn Muḥammad Ibn al-ʾAnbārī, *al-ʾInṣāf fī masāʾil al-xilāf bayna n-naḥwiyyīna l-baṣriyyīna wa-l-kūfiyyīn*. Cairo: Maṭbaʿat as-Saʿāda, 1961.

Ibn Jinnī, *Mulūkī* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *at-Taṣrīf al-Mulūkī*. Damascus: Dār al-Maʿārif, 1970.

——, *Munṣif* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Munṣif, šarḥ li-Kitāb at-taṣrīf li-ʾAbī ʿUṯmān al-Māzinī*. Cairo: Muṣṭafā al-Bābī al-Ḥalabī, 1954.

——, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. Beirut: Dār al-Hudā, n.d.

Ibn Yaʿīš, *Šarḥ al-Mufaṣṣal* = Muwaffaq ad-Dīn Yaʿīš Ibn Yaʿīš, *Šarḥ al-Mufaṣṣal*. Beirut: ʿĀlam al-Kutub.

——, *Šarḥ al-Mulūkī* = Muwaffaq ad-Dīn Yaʿīš Ibn Yaʿīš, *Šarḥ al-Mulūkī fī t-taṣrīf*. Aleppo: al-Maktaba al-ʿArabiyya, 1973.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿUṭmān ibn Qanbar Sībawayhi, *al-Kitāb*. Ed. ʿAbd as-Salām Hārūn. Cairo: al-Xānjī, 1988.

Secondary sources

Al-Bawab, Marwan, Yahya Mir Alam, Mohamad Mrayati, and Mohamad Hassan Al-Tayyan. 1996. *Arabic verbs in a computerized dictionary*. Beirut: Librairie du Liban.

Angoujard, Jean-Pierre. 1984. *Aspects d'une micro-prosodie, le modèle arabe*. Ph.D. diss., Université Paris VIII.

Bayyūmī, H., Xalīl Kilfat, and A. aš-Šāfiʿī. 1989. *Muʿjam taṣrīf al-ʾafʿāl al-ʿarabiyya*. Cairo: Dār ʾIlyās al-ʿAṣriyya.

Bohas, Georges. 1981. "Quelques aspects de l'argumentation et de l'explication chez les grammairiens arabes". *Arabica* 28.204–221.

——. 1982. *Contribution à l'étude de la méthode des grammairiens arabes en morphologie et en phonologie, d'après des grammairiens arabes 'tardifs'*. Lille: Presse Universitaire.

——. 1985. "L'explication en phonologie arabe". *Zeitschrift für Arabische Linguistik* 15.45–52.

Brame, Michael. 1970. *Arabic phonology: Implications for phonological theory and historical Semitic*. Ph.D. diss., Massachusetts Institute of Technology.

Chekayri, Abdellah. 1994. *La structure des racines en arabe*. Thèse de Doctorat, Université Paris VIII.

——. 1997. "Le médiopassif en arabe". *Place des grammaires traditionnelles dans la linguistique moderne*, ed. M. El-Ouadi and A.-E. Bouchikhi, 21–39. Meknes: Université Moulay Ismail, Faculté des Lettres.

——. 1999. *L'articulation du lexique et de la morphologie arabes: Des verbes défectueux et particulièrement les verbes assimilés*. Doctorat d'Etat diss., Mohamed V University.

——. 2001. "Verbes assimilés en arabe, de l'empirie à la théorie". *Langues et Littératures du Monde Arabe* 2.33–77.

——. Forthcoming. "Weak verbs in Arabic". *Proceedings of the Conference on Arabic Language and Linguistics, May 2002*. Cairo: American University in Cairo.

—— and Tobias Scheer. 1996. "The apophonic origin of glides in the verbal system of Classical Arabic". Lecarme a.o. (1996:62–76).

——. 2003. "The appearance of glides in Classical Arabic defective verbs". *Folia Orientalia* 39.7–33.

——. Forthcoming. "The conjugation of hollow verbs in Classical Arabic".

Fleisch, Henri. 1961. *Traité de philologie arabe*. I. *Préliminaires, phonétique, morphologie nominale*. Beirut: Imprimerie Catholique.

——. 1979. *Traité de philologie arabe*. II. *Pronoms, morphologie verbale, particules*. Beirut: Dar El Machreq.

Foley, J. 1985. "Quatre principes de l'analyse morphologique". *Langages* 78.57–72.

Guerssel, Mohand and Jean Lowenstamm 1996. "Ablaut in Classical Arabic measure I active verbal forms". Lecarme a.o. (1996:123–134).

Guillaume, Jean-Patrick. 1981. "Le statut des représentations sous-jacentes chez Ibn-Ginni". *Arabica* 28.222–241.

——. 1982. *Quelques aspects de la théorie morphophonologiques d'Ibn Jinnī, à propos des verbes à glide médian*. Lille: Presse Universitaire.

Kouloughli, Djamel Eddine. 1979. "Sur le traitement des glides dans la phonologie de l'arabe standard". *Théorie Analyse* 1.54–93.

Lecarme, Jacqueline, Jean Lowenstamm, and Ur Shlonsky (eds.). 1996. *Studies in Afro-Asiatic grammar*. The Hague: Holland Academic Graphics.

Mokhlis, Hassan. 1997. *Théorie du* taṣrīf *et traitement du lexique chez les grammairiens arabes*. (= *Publications Universitaires Européennes*, Série 21, *Linguistique*, Vol. 185.) Frankfurt a. Main: P. Lang.

Abdellah Chekayri (Al Akhawayn University)

Glottal Stope → Hamza

Government

Government is a concept in Arabic grammatical theory, both traditional and modern, with a long history. For instance, Wright (1974:3. I.A.3, B.2) discusses functions of the cases under the rubric of verbal and nominal governance. The medieval Arab grammarians designated it by the terms *ʾiʿmāl* and → *ʿamal*, together with the corollary role assignments *ʿāmil* 'governor [operator]' and *maʿmūl* 'governed [operand]' (Farghal 1986:7; Gaballa

1986:24; Haq 1998:61–63). In both traditions, the word designates the relationship holding between a verb or, secondarily, a preposition, and its argument(s); the verb (or preposition) is the *ʿāmil*, and its argument, typically a noun phrase (NP), is *maʿmūl.* Example (1) illustrates this relationship (Farghal 1986:8).

(1) *kataba* *d-dars-a* *fī*
wrote-3ms the-lesson-Acc in
l-bayt-i *l-walad-u*
the-house-Gen the-boy-Nom
'The boy wrote the lesson in the house'

The verb *kataba* 'he wrote' governs both the subject NP *al-waladu* and the direct object *ad-darsa*, while the preposition *fī* governs its object *al-bayti.* The tradition assumes a VSO word order as basic, with government of subject and object proceeding on the basis of left-to-right linear order. The latter assumption is necessary in the presumed absence of a verb phrase (VP) constituent, as is found in the English gloss *[wrote [the lesson][in the house]].*

These informal characterizations of governance are predecessors to its usage in generative grammar. Government is also a technical notion in government-binding (GB) theory (Chomsky 1981). In GB theory, the grammar of a human language consists of a lexicon and a rule component, the latter consisting of phrase structure and transformational subcomponents. The operation of the rule component is constrained by a modular array of subtheories belonging to Universal Grammar, each of which regulates some aspect of the form or interpretation of sentences:

(i) → Case theory licenses the assignment of abstract Case and morphological case to nominal expressions;
(ii) θ-theory (→ theta roles) determines the assignment of semantic roles such as Agent, Theme, Goal to NPs;
(iii) → Binding theory regulates the assignment of indices to NPs to express referential (in)dependencies;
(iv) Bounding theory limits movement to positions that are no more than two 'bounding nodes' away from the position from which movement originates;
(v) Control theory indicates which overt NP in a sentence may or must antecede the empty category PRO, the subject pronoun in nonfinite clauses *(John$_i$ intends [PRO$_i$ to go]);*
(vi) Government theory specifies the conditions under which one category governs another.

Although itself one of the modules of GB theory, government functions as a unifying principle for the others. Thus, abstract Case and θ-roles are assigned to NPs by their V or P governors; binding relations between NPs are computed within their 'governing category', the minimal phrase or clause containing the relevant NPs and a governor; and control theory includes the PRO theorem, the requirement that PRO be ungoverned (Chomsky 1981:191). In later work, bounding theory was assimilated to the theory of government as a special case of a moved category crossing a phrasal node that disallows government across that node (Chomsky 1986b:28–31). The role of government in Case-assignment or checking and in θ-role assignment is covered in other entries (→ case theory; → theta roles); this entry therefore is devoted to its role in other grammatical constructions and processes. It is first necessary to review some central features of the technical definition of government.

Definitions of government are generally derived from the primitive relation of 'c-command'. A node α *c-commands* a node β if neither α nor β contains the other and the first branching node (in some formulations, the minimal maximal projection) dominating α dominates β (Rinehart 1983; Chomsky 1981:166, 1986a:162, 1986b:8). So defined, c-command is an asymmetric relation: α c-commands β but not conversely. As a more local relation, government is then defined as minimal and symmetrical c-command: α *governs* β only if α c-commands β and conversely, and there is no γ such that (i) α c-commands γ and (ii) γ c-commands β (Farghal 1986:152). Definitions that restrict the government relation to categories contained in the same (minimal) maximal projection have the same effect, even when they omit any reference to c-command in the definition (Aoun and Sportiche 1983:214). These locality conditions on government, while necessary, are not sufficient because they allow phrases to govern heads as well as conversely. Thus, not only would the verbal head V of a verb phrase (VP) govern its NP complement, but the complement would also govern its head. To avoid this, the choice of governors is limited substantively to the categories N (noun), V (verb), A (adjective), P (preposition), and I for inflection as the functional head of IP (Inflection Phrase), the endocentric reanalysis of a root clause (Chomsky 1986b:3).

Government is distinguished from 'proper government', a narrower relation that divides into two types. 'Head-government' holds between the head of a phrase and its complement; for example, in the VP *[[*$_V$ *hit] [*$_{NP}$ *the ball]]*, the V *hit* head-governs its NP complement *the ball*. 'Antecedent-government' is a relation of coindexing that holds between a moved category and its 'trace', a copy of the moved category left behind at the movement site, represented by *t*. For example, in the question *Whom has Emily seen?*, *whom* has moved out of the position following *seen* in the VP and antecedent governs the trace left in that position: *whom*$_i$ *has [*$_{IP}$ *[*$_{NP}$ *Emily] [*$_{VP}$ *[*$_V$ *seen] t*$_i$*]]*. Note that the trace t_i is also head-governed by *seen*. The condition that traces must be properly governed is called the Empty Category Principle (ECP). The trace must not be separated from its antecedent by 'too many' maximal projections, certain of which can act as 'barriers' to government; otherwise, it will not be properly governed. In a similar vein, government is subject to a 'Minimality Condition' (Chomsky 1986b, Section 8): There can be no closer governor γ of the same kind (head or antecedent) as a potential governor α intervening between α and its governee β (Rizzi 1990:7, 2003:90).

We now turn to applications of government to Arabic, under three headings: (i) licensing of traces in subordinate clauses, (ii) licensing of *pro* in finite clauses, and (iii) minimality effects.

Several researchers have investigated the possibility of extraction of interrogative (WH-) arguments out of subordinate clauses in varieties of Arabic. As regulated by the Empty Category Principle, such movement exhibits an asymmetry between subject and object extraction. Consider first the Standard Arabic example in (2) (Mohammad 1999:57).

(2) *man*$_i$ *qāla* *l-walad-u*
who said the boy-Nom
ʾinna *r-rajul-a* *raʾā* *t*$_i$
that the man saw
'Who did the boy say that the man saw?'

Man 'who' has been extracted from the subordinate clause introduced by *ʾinna* 'that', specifically from direct object position. This is allowed because its trace is head-governed by the verb, satisfying the Empty Category Principle. The trace must be head-governed to satisfy the principle because *ʾinna* is a barrier for antecedent-government by *man*.

A more complex pair of examples is (3), adapted from Farghal (1993:106–107), in which a trace has been added to (3b).

(3) a. *māḏā*$_i$ *ḏ̣anna* *xālid-un* *ʾan*
what thought Khalid-Nom that
qatal-a *l-walad-u* *t*$_i$
killed the-boy-Nom
'What did Khalid think that the boy killed?'

b. **māḏā*$_i$ *ḏ̣anna* *xālid-un* *ʾan*
what thought Khalid-Nom that
qatal-a-hu$_i$ *l-walad-u* *t*$_i$
killed-him the-boy-Nom
'What did Khalid think that the boy killed [it]?'

The complementizer *ʾan* here is called 'lightened' (*muxaffafa*) *ʾan* by some scholars (Farghal 1986: 181, n. 2; cf. Abdul-Ghany 1981:8–9). Like the subjunctive *ʾan*, it selects VS word order; like *ʾanna*, it may select perfective aspect (Cantarino 1975:III.107). Example (3a) is a grammatical extraction of the interrogative *māḏā* 'what' from the subordinate clause because, as direct object, the trace is head-governed by *qatala*. Example (3b) is ungrammatical. The reason might seem to be that it adds the resumptive pronoun *-hu* to the subordinate clause, but this is in fact required in contexts such as (4) (Majdi 1990:146–147).

(4) a. *man*$_i$ *ta-ḏ̣unnu* *[salīm-an*
who 2ms-think Salīm-Acc
raʾā-hu$_i$ *t*$_i$*]*
saw-him
'Who(m) do you think Salīm saw [him]?'

b. *man*$_i$ *ta-ḏ̣unnu* *[ʾanna* *salīm-an*
who 2ms-think that Salīm-Acc
raʾā-hu$_i$ *t*$_i$*]*
saw-him
'Who(m) do you think that Salīm saw [him]?'

The trace of *man* 'who' in (4) is again a direct object, which is presumably head-governed by the verbal complex *raʾāhu* in both sentences. This is so whether there is an overt complementizer, as in (4b), or not, as in (4a); the occurrence of the complementizer has no bearing on the head-government relationship between the verb and its complement. The trace is also antecedent-governed by the interrogative operator in (4a), since there is no complementizer to serve as a barrier to government. Perhaps the difference in grammaticality between (3b) and (4b) lies in the fact that in the former, *al-waladu* intervenes between *-hu* and the trace, blocking a government relation between the clitic and the trace that licenses coindexing between them. No such blocking occurs in (4b), allowing the coindexing to proceed successfully there. Extraction of WH-objects is therefore licensed under head-government.

By contrast, extraction of a WH-operator that functions as a subject produces sharply divergent results in the presence or absence of *ʾanna* (Majdi 1990:147):

(5) a. *man*$_i$ *ta-ḍunnu* *[t*$_i$ *raʾā*
who 2ms-think saw
salīm-an]
Salīm-Acc
'Who do you think saw Salīm?'
b. **man*$_i$ *ta-ḍunnu* *[ʾanna t*$_i$
who 2ms-think that
raʾā *salīm-an]*
saw Salīm-Acc
'Who do you think that saw Salīm?'

Example (5a) is grammatical because, in the absence of the complementizer, the trace is properly (head-) governed by the matrix verb *ḍanna* as well as antecedent-governed by *man*, as in (4a). Example (5b) is ungrammatical because the trace is not properly governed. It is not head-governed because the complementizer intervenes between the matrix verb and the subject trace, blocking the government relation between them. It is not antecedent-governed for the same reason: *ʾanna* is a barrier to antecedent-government as well as head-government and is, moreover, not itself a proper governor (Shlonsky 2000:340). There is, however, a repair strategy available, shown in (6), that uses the resumptive pronoun (Majdi 1990:148; cf. Farghal 1993:107).

(6) *man*$_i$ *ta-ḍunnu* *[ʾanna-hu*$_i$
who 2ms-think that-him
raʾā *salīm-an]*
saw Salīm-Acc
'Who do you think that [he] saw Salīm?'

Since *-hu* is an overt pronoun, the Empty Category Principle does not apply, and the violation of it in (5b) is voided in (6). Why the embedded subject is an accusative clitic pronoun rather than the independent form *huwa* will be considered under the rubric of minimality effects.

The subject-object asymmetry with respect to extraction and the Empty Category Principle does not arise with the subjunctive complementizer *ʾan* (Majdi 1990:144–145).

(7) a. *man*$_i$ *tu-rīdu* *ʾan*
who 2ms-want that
ya-ktub-a *t-taqrīr-a* *t*$_i$
3ms-write-Subj the-report-Acc
'Who do you want to write the report?'
b. *māḏā* *tu-rīdu* *ʾan*
what 2ms-want that
ya-ktub-a *t*$_i$ *salīm-un*
3ms-write-Subj Salīm-Nom
'What do you want Salīm to write?'

Majdi argues that the underlying word order for Standard Arabic is VOS, hence the final (subject) trace in (7a) and the medial (object) trace in (7b). Under the VOS analysis, the VO sequence forms a VP constituent that excludes the subject (cf. Mohammad 1999, Chap. 2). Therefore, the direct object in each sentence, *at-taqrīra* in (7a) and the trace of *māḏā* in (7b), is properly governed by the verb. Since the subject is external to the VP, it cannot be governed by the verb. Nor can it be antecedent-governed in (7a) by *man*, since *ʾan* is a barrier to government. In (7b), *salīmun* is governed by I(nflection), as its nominative case indicates (→ case theory). For (7a), Majdi (1990:145) proposes that the subject *man* undergoes NP Preposing to adjoin to the verb in order to be head-governed by it. With the Empty Category Principle thus satisfied, the operator raises to the sentence-initial position it occupies in (7b).

'Lightened' *ʾan* patterns with the subjunctive form in (7) against *ʾanna* with respect to subject extraction (Fassi Fehri 1993:66; cf. Abdul-Ghany 1981:113).

(8) a. *ʾayy-u* *rijāl-in*$_i$ *ḥasib-ta*
which-Nom men-Gen thought-2ms
ʾan t_i *jāʾ-ū*
that came-3mpl
'Which men did you think came?'
b. **ʾayy-u* *rijāl-in*$_i$ *ḥasib-ta*
which-Nom men-Gen thought-2ms
ʾanna t_i *jāʾ-ū*
that came-3mp
'Which men did you think came?'

Fassi Fehri speculates that the complementizer *ʾan* can host an adjacent trace because, unlike *ʾanna*, it is not a Case-assigner. This is plausible in view of the ungrammaticality of **ʾayy-u rijālin ḥasibta ʾan-hum jāʾū*: the sentence is ungrammatical because *-hum* lacks (accusative) Case. But C(omplementizer) is not a member of the class of head-governors above, hence not a proper governor, so there should be an Empty Category Principle violation in (8a); in fact, *ʾan* is a governor only when followed by a subjunctive (Cantarino 1975:III, 105). Should *ʾan* be admitted to the class of head-governors to accommodate this case, new problems arise: (i) the SV order assumed in (8) must be motivated in view of the VS order in (3); and (ii) assuming SV order entails that if *ʾan* assigns subjunctive mood to the verb under government and strict adjacency (Aoun 1985:57), the intervening trace will block government and so mood assignment.

An alternative analysis of (8) is suggested by Kenstowicz's (1989) study of WH-extraction in two Arabic dialects, Levantine and Beni-Hasan. Kenstowicz (1989:263) proposes that Levantine Arabic is a partly null-subject language and Beni Hasan a fully null-subject language, on the basis of criteria devised for null subject Romance languages: (i) null subjects of finite clauses, (ii) extraction from the subject position of a *that* clause (i.e., absence of *that-t* effects), and (iii) free inversion of the subordinate clause subject. Both dialects exhibit the first property while Beni Hasan also exhibits the second in (9a) and the third in (9c) as against (9b) (Kenstowicz 1989:265–266).

(9) a. *wayy* *bint*$_i$ *farīd* *gāl* *innu*
which girl Farīd said that
t_i *ištarat* *allibās*
bought the-dress
'Which girl did Farīd say bought the dress?'
b. *farīd* *gāl* *innu* *albint*
Farīd said that the-girl
ištarat *allibās*
bought the-dress
c. *farīd gāl innu ištarat albint allibās*
'Farīd said that the girl bought the dress'

Neither (ii) nor (iii) holds in Levantine Arabic (Kenstowicz 1989:264–265):

(10) a. **ayy* *bint*$_i$ *farīd* *kāl* *innu*
which girl Farīd said that
t_i *ištarat* *alfustān*
bought the-dress
'Which girl did Farīd say that bought the dress?'
b. *farīd* *kāl* *innu* *albint*
Farīd said that the-girl
ištarat *alfustān*
bought the-dress
c. **farīd kāl innu ištarat albint alfustān*
'Farīd said that the girl bought the dress'

Following Rizzi (1982), Kenstowicz (1989:267) proposes that Beni Hasan permits subject extraction as in (9a) because inversion as in (9c) permits the subject to originate in a position in which it is head-governed by the verb, satisfying the Empty Category Principle. This analysis extends straightforwardly to (8) if we allow that, by virtue of its not being a Case-assigner, *ʾan* requires inversion of the subject, so that its trace is postverbal and properly governed by *jāʾū* (cf. Plunkett 1993:237). Word order with overt subjects confirms this conjecture for both 'lightened' *ʾan* (cf. [3a]) and subjunctive *ʾan* (Alsayed 1998:151[51]):

(11) *ʾarad-tu* *ʾan* *yu-ġādir-a* *xālid-un*
wanted-1sg that 3ms-leave-Subj Khalid-Nom
'I wanted Khalid to leave'

Applications of the Empty Category Principle to Arabic are more extensive than so far indicated. For discussion of subject relativization from complement clauses, see Alsayed (1998); for discussion of how the principle constrains operator movement at Logical Form in Iraqi Arabic, see Wahba (1991).

The second application of government is the licensing of *pro*, the null subject of finite clauses in null-subject languages (Chomsky 1982:77–78). Benmamoun (1995) endorses Rizzi's proposal (1986, 1990:32) that empty categories are subject to two licensing conditions, parallel to the Empty Category Principle: 'identification' by a local head with which the empty category agrees in its Φ-features (person, gender, number), comparable to antecedent government, and 'formal licensing' by Case-assignment under government (Alsayed 1998:28), comparable to head-government. Thematic *pro* is licensed by both identification and Case, while expletive *pro* is licensed only by Case. Benmamoun takes exception to Rizzi's requirement that the same head-governor license and identify *pro*. Standard and Moroccan Arabic provide counterexamples to the requirement:

(12) pro	*lam*	*ya-dxul-ū*	(Standard
	Neg-Past	3-come-mp	Arabic)
	'They did not enter'		

(13) pro	*ġad-a*	*te-mši*	(Moroccan
	will-fsg	3fs-go	Arabic)
	'She will go'		

In (12), the tense feature of the negative head *lam* assigns nominative Case to license *pro*, while agreement features to identify it are carried by the main verb. Similarly, in (13) only *temši* carries the person feature that identifies the pronoun; the auxiliary *ġada*, by hypothesis, occupies T(ense) (Inflection in earlier work) and assigns nominative Case for licensing (actually, under specifier-head agreement [→ case theory] rather than government). The implication of the analysis is that the two forms of proper government sanctioned by the Empty Category Principle are as independent and as necessary as the licensing and identification conditions on *pro*.

The final set of examples illustrate the Minimality Condition on government, the requirement that there be no closer governor between a governor and putative governee. Recall the question of why pronominal subjects in *ʾanna* clauses must be accusative clitics rather than independent pronouns. Licensing the latter would require government by I(nflection, head of Inflection Phrase) to assign nominative Case. LeTourneau (1993) argues that *ʾanna* is a closer governor under Minimality than I(nflection) because although the clitic originates in subject position (the specifier of IP) at D-structure (and later incorporates into C), it is still external to the I' projection headed by I. Since *ʾanna* is the closer governor and is an accusative Case-assigner (cf. [6]), the subject receives accusative Case and surfaces as a clitic (LeTourneau 1993:263–266).

Harbert and Bahloul (2002:51) defend a partial government analysis of cases of "first conjunct agreement" in Standard Arabic VS sentences with conjoined subjects:

(14) a.	[$_{IP}$ [$_{I}$ xaraj-at]	[$_{CjP}$ [al-bint-u]
	left-3fs	the-girl-Nom
	[$_{Cj'}$[$_{C}$ wa]	[al-walad-u]]]]
	and	the-boy-Nom
	'The girl and the boy left'	
b.	[$_{IP}$ [$_{I}$ xaraj]	[$_{CjP}$ [al-walad-u]
	left-3ms	the-boy-Nom
	[$_{Cj'}$ [wa]	[al-bint-u]]]]
	and	the-girl-Nom
	'The boy and the girl left'	

Why do gender and number agreement hold only between the verb and the first conjunct of the compound subject? In agreement with Aoun a.o. (1994) and others, Harbert and Bahloul assume that coordinate structures conform to the X-bar schema, with the first conjunct in the specifier of Conjunction Phrase (CjP) and the second as the complement of the conjunction, which heads the phrase. I(nflection) is a minimality governor for CjP, there being no closer c-commanding head to govern the phrase, and so governs its specifier as well (Chomsky 1986b:11). The latter government relation permits agreement between I and the first conjunct (*al-bintu* or *al-waladu*), yielding gender agreement between the verb and the NP. No agreement holds between the verb and the second conjunct because the latter is in the domain of a more local head, Cj.

BIBLIOGRAPHICAL REFERENCES

Abdul-Ghany, Mohammad Kamal al-Deen. 1981. *Government binding in Classical Arabic.* Ph.D. diss., University of Texas at Austin.

Alsayed, Adnan. 1998. *A government and binding approach to restrictive relatives, with particular reference to restrictive relatives in Standard Arabic.* Ph.D. diss., University of Essex.

Aoun, Joseph. 1985. *A grammar of anaphora.* Cambridge, Mass.: MIT Press.

—— and Dominique Sportiche. 1983. "On the formal theory of government". *The Linguistic Review* 2.211–236.

Aoun, Joseph, Elabbas Benmamoun, and Dominique Sportiche. 1994. "Agreement and conjunction in some varieties of Arabic". *Linguistic Inquiry* 25.195–220.

Benmamoun, Elabbas. 1995. "The conditions on *pro* and the ECP". *Proceedings of the 13th West Coast Conference on Formal Linguistics*, ed. Raul Aronovitch, William Byrne, Susanne Pruess, and Martha Sentoria, 173–188. Stanford, Calif.: Center for the Study of Language and Information.

Cantarino, Vicente. 1975. *Syntax of Modern Arabic prose,* III. Bloomington: Indiana University Press.

Chomsky, Noam. 1981. *Lectures on government and binding: The Pisa lectures.* Dordrecht: Foris.

——. 1982. *Some concepts and consequences of the theory of government and binding.* Cambridge, Mass.: MIT Press.

——. 1986a. *Knowledge of language: Its nature, origin, and use.* New York: Praeger.

——. 1986b. *Barriers.* Cambridge, Mass.: MIT Press.

Eid, Mushira and Clive Holes (eds.). 1993. *Perspectives on Arabic linguistics*, V. Philadelphia and Amsterdam: J. Benjamins

Farghal, Mohammad Ali. 1986. *The syntax of Wh-questions and related matters in Arabic.* Ph.D. diss., Indiana University.

——. 1993. "The theory of PRO and Arabic empty categories". *Papers and Studies in Contrastive Linguistics* 28.99–108.

Fassi Fehri, Abdelkader. 1993. *Issues in the structure of Arabic clauses and words.* Boston: Kluwer Academic Publishers.

Gaballa, Hassan. 1986. *Case marking and Classical Arabic: A semantic and historiographic study.* Ph.D. diss., University of Delaware.

Haq, Anwar ul. 1998. *The theory of governance in Arabic traditional grammar.* Ph.D. diss., Indiana University.

Harbert, Wayne, and Maher Bahloul. 2002. "Postverbal subjects in Arabic and the theory of agreement". *Themes in Arabic and Hebrew syntax*, ed. Jamal Ouhalla and Ur Shlonsky, 45–69. Boston: Kluwer Academic Publishers.

Jaeggli, Osvaldo and Kenneth J. Safir. 1989. *The null subject parameter.* Boston: Kluwer Academic Publishers.

——. 1989. "The null subject parameter and parametric theory". Jaeggli and Safir (1989:1–44).

Kenstowicz, Michael. 1989. "The null subject parameter in Modern Arabic dialects". Jaeggli and Safir (1989:263–275).

LeTourneau, Mark S. 1993. "Case-marking and binding of subject clitics in Arabic complement clauses". Eid and Holes (1993:261–290).

Majdi, Basim. 1990. "Word order and proper government in Classical Arabic". *Perspectives on Arabic linguistics,* I, ed. Mushira Eid, 127–153. Philadelphia and Amsterdam: J. Benjamins.

Mohammad, Mohammad A. 1999. *Word order, agreement, and pronominalization in Standard and Palestinian Arabic.* Amsterdam and Philadelphia: J. Benjamins.

Plunkett, Bernadette. 1993. "The position of subjects in Modern Standard Arabic". Eid and Holes (1993:230–260).

Rinehart, Tanya. 1983. *The syntactic domain of anaphora.* Chicago: University of Chicago Press.

Rizzi, Luigi. 1982. *Issues in Italian syntax.* Dordrecht: Foris.

——. 1986. "Null objects in Italian and the theory of *pro*". *Linguistic Inquiry* 17.501–558.

——. 1990. *Relativized minimality.* Cambridge, Mass.: MIT Press.

——. 2003. "Relativized minmality effects". *The handbook of contemporary syntactic theory,* ed. Mark Baltin and Chris Collins, 89–110. Malden, Mass.: Blackwell.

Shlonsky, Ur. 2000. "Remarks on the complementizer layer of Standard Arabic". *Research in Afroasiatic grammar: Papers from the Third Conference on Afroasiatic Languages*, ed. Jacqueline Lecarme, Jean Lowenstamm, and Ur Shlonsky, 325–343. Amsterdam and Philadelphia: J. Benjamins.

Wahba, Wafaa Abdel-Faheem Batran. 1991. "LF movement in Iraqi Arabic". *Logical structure and linguistic structure: Cross-linguistic perspectives,* ed. C.-T. James Huang and Robert May, 253–276. Dordrecht: Kluwer Academic Publishers.

Wright, William. 1974. *A grammar of the Arabic language.* Beirut: Librairie du Liban. (Repr. of the ed. Cambridge 1896.)

MARK S. LETOURNEAU (Weber State University)

Government and Binding → Binding

Grammatical Tradition: Approach

1. SOURCES AND HISTORICAL OVERVIEW

Throughout its historical development (→ grammatical tradition: history), the Arabic grammatical tradition's approach to language and language description was founded on a remarkably self-consistent set of general principles (of axioms, so to speak) defining its object, its aims, and its methods. These principles, however, were not explicitly and systematically set forth by the first generations of grammarians, who usually took them for granted, or referred to them casually when and insofar as they were relevant to the discussion of a specific question or piece of data. This situation changed rather drastically in the 4th/10th century, when the diffusion of the Greek philosophical and scientific heritage, and mostly logic, presented grammarians with new questions and challenges (Versteegh 2000; Bohas a.o. 1990:8–14). Seen from the grammarians' point of view, the debate revolved around two main issues, the status of grammar as an autonomous science, and the specificity of its object, the 'speech of the Arabs' (*kalām al-ʿArab*). These issues figure prominently in az-Zajjājī's (d. ca. 340/950) *ʾĪḍāḥ* (Versteegh 1995) and Ibn Jinnī's (d. 392/1002) *Xaṣāʾiṣ* (Guillaume 2000), which treat a wide range of epistemological and methodological questions in an attempt to demonstrate that grammar was not a mere utilitarian discipline (as most logicians claimed) but rather an authentic speculative science, offering deep insights into the nature of language. The same preoccupation with clarifying the foundations of grammatical theory and with finding new, more explicit ways to formulate it is also perceptible in Ibn as-Sarrāj's (d. 316/928) *ʾUṣūl*, a descriptive trea-tise following an entirely new and systematic order of exposition where the place of each category is defined by its position within the conceptual organization of the theory. With some readjustments, this order was gradually adopted by later grammarians and became the canonical mode of exposition for grammatical treatises. By the end of the 4th/10th century, this process of 'standardization' (Owens 1990) of grammatical doctrine had practically achieved its ends, and no major evolution occurred in subsequent centuries. In the 6th/12th century, though, the jurist and grammarian Ibn al-ʾAnbārī (d. 577/1181) endeavored to launch a new grammatical discipline, *ʾuṣūl an-naḥw*, on the model of the *ʾuṣūl al-fiqh*. He composed two short treatises on this subject, the *Lumaʿ*, mainly devoted to problems relative to the transmission of linguistic data and to grammatical reasoning (*qiyās*), and the *ʾIġrāb*, about the methodology of grammatical disputation. This attempt, how-ever, did not really succeed; Ibn al-ʾAnbārī had very few followers, the most notable being as-Suyūṭī (d. 911/1505), whose *Iqtirāḥ* follows rather closely the *Lumaʿ* and indeed quotes extensively from it, as well as from a number of other texts, among them az-Zajjājī's *ʾĪḍāḥ* and Ibn Jinnī's *Xaṣāʾiṣ*.

2. THE OBJECT LANGUAGE: *KALĀM AL-ʿARAB*

The most common expression used by Arabic grammarians to refer to their object is *kalām al-ʿArab* 'the speech of the Arabs', by which they meant the linguistic usage of the original inhabitants of the Arabian Peninsula, as opposed to the new, reputedly 'degraded' variety of Arabic that had evolved in the conquered provinces soon after the first great expansion of Islam (Versteegh 1997:93–113). Although the expression *kalām al-ʿArab* seems to refer to the living usage of the Bedouin Arabs, it should be taken in the restrictive sense of the literary variety of Arabic reflected in the *Qurʾān* and the ancient poetry, which is generally considered by contemporary Arabic linguists as clearly different from the language that was used in everyday communication in pre-Islamic Arabia (Zwettler 1978; Versteegh 1997:46–51). This distinction, however, was never acknowledged by the Arabic grammarians (or indeed by Classical Arabic culture as a whole), who, on the contrary, insisted on the fundamental unity of *kalām al-ʿArab*. In their view, the Arabs' mastery of their language was innate, in the sense that it was not acquired through an explicit, conscious process of learning, and came from their 'natural genius' (*ṭabʿ*). Thus, the most primitive and unsophisticated Bedouin were believed not only to be able to express themselves in the purest and most elegant kind of Arabic but also to distinguish with total accuracy between correct and incorrect forms or constructions, so that they had to be considered as infallible authorities on language matters. Actually, while early grammarians and philologists frequently referred to the living usage of Bedouin Arabs (in the case of Sībawayhi, see Levin 1994) and used them as informants or even (reportedly at least) as arbiters in controversies about the acceptability of a given piece of data (see, however, Talmon 1986), it should be kept in mind that these 'native speakers' were probably poets or transmitters of ancient poetry. In any case, this practice soon dwindled away, the reason given being that even the desert-dwelling Arabs had lost their infallible linguistic intuition owing to their frequent contact with the 'corrupt' speech of the sedentary populations.

By the end of the 3rd/9th century, *kalām al-ʿArab* had come to denote in effect a virtually closed set of transmitted textual data, consisting mainly in quotations from the *Qurʾān* (in all its canonical variants, *qirāʾāt*), from ancient poetry (i.e. prior to the 130s/750s and composed by poets born and raised in the desert), and, in a much lesser measure, of old Bedouin proverbs and sayings. These 'primary' data, however, usually play a specific part in grammatical texts, namely to attest that a given form or construction does exist in the *kalām al-ʿArab*, which normally implies that it is somehow deviant from the most common usage, or difficult to analyze in terms of the basic grammatical rules. This function appears quite clearly in the technical term used for poetic quotations referred to in grammatical texts, *šawāhid* (sg. *šāhid*, 'witness'). On the other hand, in order to illustrate and discuss more common facts (or, conversely, purely hypothetical ones, such as could not be directly attested in actual usage; see Baalbaki 2003), the grammarians had recourse to artificially constructed examples (Fournier a.o. 2006). No less than the transmitted data, which knew very few additions after the 3rd/9th century, the 'technical' examples show a remarkable permanence throughout the history of grammar, which tended to lend them practically the same degree of authority.

It should be noted that the variety of Arabic described by the grammatical tradition, with its heavy reliance on ancient poetry and Qurʾānic variants, offered a high degree of variability and heterogeneity, and exhibited a large quantity of archaisms, poetic license, and generally deviant forms and constructions that did not fall easily within the system of rules devised by the grammarians on the basis of the most common usage. These data, however, could not be simply rejected as irrelevant or agrammatical (although some exceptions occur, see Baalbaki 1985 and Guillaume 2006), since they were attested in the *kalām al-ʿArab*. The technical solution to this problem consisted in dividing linguistic facts into 'regular' (*muṭṭarid*, *qiyāsī*) and 'irregular' (*šāḏḏ*) ones, the principle being that irregular facts could not be used as counterexamples to rules built on well-attested regularities, nor could any valid generalization be built on them (Bohas a.o. 1990:17–20). Of course, grammarians did not always agree on which facts were regular and which were not, but they had at least a practical way to neutralize marginal pieces of data. But, at another level, these deviant and irregular facts played a crucial part in the theory. Arabic grammarians had a very strong sense that the *kalām al-ʿArab* formed a totally coherent and harmonious system in which every detail had its place and its reason for being (see below, Sec. 3). In consequence, they felt that even the most irregular and apparently aberrant facts could, and should, be accounted for by the theory; this was usually done by demonstrating that the apparent irregularity could be explained in terms of a deeper structural principle that had prevented the normal application of a rule or, conversely, had caused its application out of its normal context (see examples in Bohas a.o. 1990:17–20; and Guillaume 2006). This technique, known as *taʿlīl* (Suleiman 1999), came to play a larger and larger part in grammatical texts from the 4th/10th century onward, as the gradual standardization of the theory left few other possibilities for grammarians to display both their deep knowledge of the 'speech of the Arabs' and their technical mastery of the theory's potentialities. Their efforts, although usually demonstrating a remarkable degree of ingenuity and subtlety and sometimes bringing to light some interesting properties of Arabic (notably the case with Ibn Jinnī's *Xaṣāʾiṣ*), very often strike one as purely academic displays of a somewhat gratuitous dialectic skill.

3. The aims of grammar

According to a widespread tradition, grammar was first 'invented' by ʾAbū l-ʾAswad ad-Duʾalī (d. 69/688?) in an attempt to correct the 'corruption of speech' that appeared among the descendants of the Arab conquerors after the first expansion of Islam in the second half of the 1st/7th century. Although this account is most certainly legendary (Talmon 1985), it is consistent with a claim grammarians never ceased to make, that the basic purpose of their discipline was to teach the rules of correct linguistic usage and enable learners to avoid solecisms (*laḥn*), notably in the recitation of the *Qurʾān*; indeed, grammar was regarded as a propaedeutical science, whose mastery was a prerequisite to any kind of intellectual career. But, at the same time, it has often been remarked that even the earliest grammatical treatises, such as Sībawayhi's *Kitāb*, al-Farrāʾ's *Maʿānī* (which is actually a grammatical commentary on difficult passages of the *Qurʾān*; see Kinberg 1996) or al-Mubarrad's *Muqtaḍab*, could hardly be used as teaching grammars, and were certainly never meant for that purpose. On the other hand, a number of short treatises addressed to beginners were written roughly during the same period, such as the *Muqaddima*, attributed to Xalaf al-ʾAḥmar (d. 180/796), which was probably composed in the late 2nd/8th or the early 3rd/9th century (Talmon 1990), and some others, written between the 3rd/9th and the 7th/13th century. By and large, however, pedagogical attainability does not appear to have been the primary objective of the overwhelming majority of the Arabic grammarians (Baalbaki 2005). Instead, their main preoccupations seem to have been preserving and recording the 'linguistic heritage' of old Arabia, in all its richness and intricacy (hence the dominant role played, in treatises such as the *Kitāb* or the *Muqtaḍab*, by rare and archaic facts); and devising a coherent and comprehensive theory in which the most minute and the most deviant examples out of a huge mass of data could find their place and be accounted for. That such an objective went beyond a merely utilitarian conception of grammar was explicitly acknowledged by many grammarians, especially in the 4th/10th century. Indeed, they were eager to establish that there was more to their discipline than a mere set of prescriptive rules, and they insisted that, at a higher level, grammar was able not only to describe linguistic facts but also to explain them, and, by so doing, to reveal the deep hidden harmony and 'wisdom' (*ḥikma*) that, in their eyes, uniquely characterized the Arabic language. Perhaps the most perceptive expression of this idea is to be found in a well-known chapter of Ibn Jinnī's *Xaṣāʾiṣ* (I, 48–96), where he claims that the grammarians' explanations (*ʿilal*, sg. → *ʿilla*) are closer to those of the speculative theologians (*mutakallimūn*) than to those of the jurists (*fuqahāʾ*). According to Ibn Jinnī, the *fuqahāʾ* limit themselves to normative statements based on the Revelation and are not supposed to delve into the deep motivations of those statements, while the *mutakallimūn*, in their attempt to vindicate the revealed truths on rational grounds, must necessarily rely on common sense, whether perceptual or rational. In the same way, the grammarians, or more exactly 'the most clever and competent ones', are able not only to formulate adequate rules describing the correct usage but also to justify these rules. Ibn Jinnī exemplifies this claim by showing that all morphophonological processes in Arabic ultimately rely on the immediate sensory perception that some sounds or sequences of sounds are 'heavier' or 'lighter' than others (i.e. that their pronunciation entails a greater or smaller expenditure of energy for the speaker; see Bohas 1981; Bohas a.o. 1990:80–92). Moreover, the way in which morphological rules apply or, in some cases, do not apply, is, according to him, governed by a general constraint optimizing the ratio between the quantity of energy necessary to produce a given form and the amount of meaning it conveys. For instance, the scheme *fuʿil* (CuCiC), which is 'heavy', is never used for nouns but only for the passive of verbs in the perfect tense, since verbs, which express both a process and a time, are 'heavier' than nouns. According to Ibn Jinnī, this correspondence between sound and meaning is only one example among many of the perfect harmony and equilibrium that pervades the *kalām al-ʿArab* and shows its superiority over other idioms. Just as Ibn Xaldūn (d. 757/1356) would do several centuries later, he explains this unique characteristic by the 'innate genius' of the native Arabs, whose natural keenness of perception and sense of harmony have not yet been spoiled by the sophistication of civilized life. An alternative explanation (also sporadically referred to by Ibn Jinnī, e.g. *Xaṣāʾiṣ* I, 40–48), more commonly found within the tradition, involves the myth of the 'institution of language' (→ *waḍʿ al-luġa*): while grammarians and theologians were divided as to the question of the human or divine origin of the language (see Weiss 1974; Loucel 1963–1964), and while most of them concluded on a suspension of judgment between the two hypotheses, they all agreed on the fact that the *kalām al-ʿArab* had been instituted, once and for all, by one or several beings of superior forethought and wisdom, who had planned it even to its minutest details (e.g. Zajjājī, *ʾĪḍāḥ* 55–56; and Versteegh 1995:89, 94). Consequently, the highest aim of grammar should be the reconstruction of the hidden 'intentions of the founder of the language' (*ʾaġrāḍ wāḍiʿ al-luġa*), whose 'wisdom' could be discovered not only in general rules and regular facts but also in the most recondite and apparently deviant pieces of data.

4. GRAMMATICAL EXPLANATION

This conception of the *kalām al-ʿArab* as forming a perfectly coherent and harmonious system is reflected in the explanatory method followed by the grammarians. Although it can be applied in many different ways according to the facts or classes of facts under examination and to the specific point the grammarian wants to make, it relies on a rather simple conceptual pattern. The basic idea is that the properties and behavior of an entity (be it an individual fact, a class of facts, or an abstract category) are defined by its position within the general system of the language and, conversely, that entities occupying identical or similar positions tend to exhibit identical or similar properties (Bohas a.o. 1990:22–26). It follows that each major category is associated with a set of properties or kind of behavior that is considered to be 'normal' or, as we would say nowadays, 'prototypical'. It is, for instance, 'normal' for a noun to have a three-case declension (→ *ʾiʿrāb*), or for a verb to be conjugated. Yet, when an entity departs from the prototypical behavior of the category it belongs to (for instance, in the case of indeclinable nouns and nonconjugated verbs), an explanation is required. Usually, this explanation consists in showing that the element or elements under consideration occupy a marginal position within their category and as such do not enjoy the full privileges associated with more prototypical members (most indeclinable nouns and nonconjugated verbs are actually highly grammaticalized morphemes, which have much in common with particles). One of the most common terms used by Arabic grammarians to refer to the prototypical behavior of a category is → *ʾaṣl* (lit. 'roots and trunk [of a tree]'), as in the following quotation from az-Zajjājī (*ʾĪḍāḥ* 51), which perfectly embodies this kind of reasoning:

> A thing [i.e. a category] can have a generally accepted prototypical behavior (*ʾaṣl mujtamaʿ ʿalay-hi*), and then, some of its elements can depart from it because of a weakness [→ *ʿilla*; it should be noted that this term also means 'cause' and 'explanation'] which affects them.

This approach is explained and illustrated in a more detailed manner in another text of az-Zajjājī (*ʾĪḍāḥ* 64–65; see Versteegh 1995:86–89; Bohas a.o. 1990:25–26), where he distinguishes three different levels of grammatical 'causes' or 'explanations' (*ʿilla*). The first one is the 'didactic explanation' (*ʿilla taʿlīmiyya*), which is all but a statement of a general rule, such as saying that the assertive particle *ʾinna* (→ *ʾinna wa-ʾaxawātuhā*) governs the accusative in the topic of a nominal sentence and the nominative in its comment, for example in *ʾinna zayd-an qāʾim-un* 'verily, Zayd-ACC [is] standing-NOM'. This 'explanation' simply represents what one has to know in order to use correctly this kind of particle.

The second step, the 'analogical explanation' (*ʿilla qiyāsiyya*), consists in explaining that *ʾinna* governs the accusative because it 'resembles' a verb. By the same token, the topic and the comment are 'assimilated', respectively, to the object and subject/agent of a verbal sentence: *ʾinna zaydan qāʾimun*, then, is 'similar' to a sentence like *ḍaraba ʿamran zaydun* 'he hit ʿAmr, Zayd did', where the object precedes the subject. This similarity, of course, is purely formal and only takes into account the nature and distribution of case endings. But it provides a first hypothesis about the position of *ʾinna* in the general system of the language, as a kind of 'verb-like particle'.

The third step, called by az-Zajjājī the 'dialectic explanation' (*ʿilla jadaliyya*), consists in testing this hypothesis, or more exactly in answering the possible objections against it. For instance, the relevance of the similarity between *ʾinna zaydan qāʾimun* and *ḍaraba ʿamran zaydun* can be challenged on the grounds that the latter does not exhibit the basic word order in a verbal sentence, which is Verb-Subject-Object. To put it differently, if *ʾinna* had the same behavior as a verb, it should be expected to govern first the nominative and then the accusative, not the other way around. The canonical answer (referred to by Zajjājī, *ʾĪḍāḥ* 135) is that *ʾinna*, being a particle, does not enjoy the same 'versatility' (*taṣarruf*) as the verb, so that "all that is permitted with a verb is not permitted with it". In other terms, *ʾinna* is only weakly related to the verb and as such cannot exhibit all its prototypical properties, except one, namely the government of the accusative. One could be tempted, at this level, to compare the status of the *ʾinna*-class particles with that of the *kāna*-class of auxiliary verbs (→ *kāna wa-ʾaxawātuhā*). Like the former, the latter can only affect a nominal sentence, governing the nominative in the topic and the accusative in the comment (i.e. the reverse of *ʾinna*), as in *kāna zaydun qāʾiman* 'Zayd was standing'. Consequently, *kāna* is analyzed as being less 'verb-like' than 'normal', prototypical verbs, which express both a process (*ḥadaṯ*) and a time (*zamān*). On the other hand, since they do express a time and have a full conjugation, they are more markedly 'verbal' than *ʾinna*, which explains why they govern the nominative and the accusative exactly as the verb does. Moreover, *kāna* sentences admit the moving of the element governed in the accusative (corresponding to the complement of the verbal sentence) before the element in the nominative (corresponding to the subject/agent), which is impossible with *ʾinna* sen-tences: *kāna qāʾiman zaydun* 'he was standing, Zayd was' is possible just as *ḍaraba ʿamran zaydun*, while **ʾinna qāʾiman zaydun* is completely ungrammatical.

As should by now be clear, explaining a fact in this framework usually consists in relating it to another, supposedly more basic or prototypical one. The closer this relation is, the more properties the fact under discussion will have in common with its prototype. Moreover, the nature of this prototype can vary according to the property under discussion: for example, while the distribution of case markers in *ʾinna* and *kāna* sentences is explained by their formal affinity to the verbal sentence, they can also be analyzed as variants of a basic nominal sentence (e.g. *zaydun qāʾimun* 'Zayd is standing'), since they exhibit the same topic/comment structure. In some domains of morphophonology (*taṣrīf*), however, a variant of this approach can be observed, notably in the treatment of → 'weak' verbs (i.e. those whose triconsonantal root contains a glide). In this case, the explanation consists in reconstructing an underlying form (*ʾaṣl*) and then identifying the phonological processes that have affected it in order to bring forth the surface form. For instance, *qāla* /qa''la/ 'he said' is analyzed as derived from /qawala/ through a general rule predicting that *wāw* and *yāʾ* are changed into an *ʾalif* when they are preceded by an *a* and followed by a vowel (Bohas and Guillaume 1984:375–467; Åkesson 2001:282–296).

In any case, this approach basically relies on the idea that grammatical categories and concepts are organized hierarchically. This conception, already present in an incipient form in Sībawayhi's *Kitāb* (Baalbaki 1979), is reflected in a more systematic way in the organization of the later grammatical treatises, where the order of exposition is intended to represent, in a wholly transparent and predictable way, the conceptual hierarchy of the theory. The basic principles of this method were first developed and implemented in the early 4th/10th century by Ibn as-Sarrāj in his *ʾUṣūl* and, with some minor readjustments and variants, became the generally accepted norm for all later grammatical treatises. They are usually divided into three main parts, devoted respectively to the noun, the verb, and the particle. This is the order of treatment that is always maintained; it is justified by the fact that the noun, which signifies only its meaning, is in a way a more prototypical part of speech than the verb, which signifies both its meaning (i.e. the process) and a time past, present, or future, while the particle, which can only signify when it is associated to something else (a noun, a verb, or a sentence) is less prototypical than the other two. In the first section, the declinable nouns are dealt with before the indeclinable ones, as declension (*ʾiʿrāb*) is a prototypical property of the noun. The nouns governed in the nominative have precedence because they belong to the predicative core of the sentence, while those governed in the accusative and the genitive do not, so that they come later. This method offered many obvious advantages, among them the fact that it could be used in any kind of grammatical writing, from the basic compendium to the most exhaustive and theoretically ambitious treaty. At the same time, it led to an increasing degree of standardization and predictability of the theory, which ultimately caused the decline of this method after the 8th/14th century.

5. CRITICISM OF THE STANDARD APPROACH WITHIN THE GRAMMATICAL TRADITION

Some aspects of the approach theorized by the 4th/10th century grammarians were criticized by a number of later scholars. The most radical was certainly Ibn Maḍāʾ (d. 606/1208), who vehemently rejected some of the basic concepts of the canonical doctrine, notably the theory of government (→ *ʿamal*) and reproached the grammarians for their overindulgence in far-fetched explanations and fruitless speculation (Suleiman 1999:145-164). Although his criticism often agrees with common sense, it is mostly negative, and his *Radd ʿalā n-nuḥāt* 'Refutation of the grammarians' sounds more like the work of a religious polemist than of a reformist grammarian. In a much less polemical vein, as-Suhaylī (d. 581/1192), another Andalusian scholar, offers in his *Natāʾij* an original and interesting attempt at simplifying the canonical theory (Baalbaki 1999), leaving alone many traditional issues of purely academic interest. But perhaps the most far-reaching endeavor to reform grammar, and indeed to found a new approach to language, is represented by al-Jurjānī's *Dalāʾil*. Strongly criticizing the grammarians for their narrowly formalistic outlook and deriding their taste for 'abstruse questions' and their hazardous speculations on the 'intentions of the founder of language' (*Dalāʾil* 30–36), he reproaches their lack of interest in the semantic aspect of the concepts and categories they use (Kouloughli 1985; Bohas a.o. 1990:116–117).

Although some of al-Jurjānī's ideas found their way into the canonical doctrine, if rather superficially and imperfectly, and influenced later grammarians such as az-Zamaxšarī (d. 539/1143) and al-ʾAstarābāḏī (d. 686/1288), these attempts, by and large, remained isolated and had few, if any, consequences for the evolution of grammar. Moreover, they never challenged some of the basic attitudes underlying the tradition's approach to language, notably its claim of the uniqueness and superiority of the *kalām al-ʿArab*, which is perhaps the tradition's major limitation. The effects of the approach of the grammarians were twofold: they were not interested in other languages than Arabic, and they were unable to take into account historical change in the only language they were interested in, Arabic. It is remarkable that, although many Arabic grammarians knew several languages (not a few were indeed non-native Arabic speakers), and although Arabic grammar served as a model to describe several languages (notably Hebrew), references to other languages or to other varieties of Arabic than the *kalām al-ʿArab* are practically nonexistent in the whole corpus of Arabic grammatical texts, extensive as it is. But then, this exclusive interest in the peculiarities of a single language, at the expense of its most universal traits, is perhaps a common characteristic of independent grammatical traditions.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

ʾAnbārī, *Lumaʿ* and *ʾIġrāb* = ʾAbū l-Barakāt Abd ar-Raḥmān ibn Muḥammad al-ʾAnbārī, *Lumaʿ al-ʾadilla fī ʾuṣūl an-naḥw*, followed by *al-ʾIġrāb fī jadal al-ʾiʿrāb*. Ed. Saʿīd al-ʾAfġānī. Damascus: Dār al-Fikr, 1957.

Farrāʾ, *Maʿānī* = ʾAbū Zakariyyāʾ Yaḥyā ibn Ziyād al-Farrāʾ, *Maʿānī l-Qurʾān*. Ed. Muḥammad ʿAlī an-Najjār. 3 vols. Cairo: ad-Dār al-Miṣriyya, 1955–1972. (Repr., Beirut: ʿĀlam al-Kutub, 1983.)

Ibn Jinnī, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. Ed. Muḥammad ʿAlī an-Najjār. 3 vols. Cairo, 1952–1956. (Repr., Beirut: Dār al-Hudā, n.d.)

Ibn Maḍāʾ, *Radd* = ʾAbū l-ʿAbbās ʾAḥmad Ibn Maḍāʾ al-Qurṭubī, *ar-Radd ʿalā n-nuḥāt*. Ed. Šawqī Ḍayf. Cairo: Dār al-Maʿārif, 1982.

Ibn as-Sarrāj, *ʾUṣūl* = ʾAbū Bakr Muḥammad ibn as-Sarī Ibn as-Sarrāj, *al-ʾUṣūl fī n-naḥw*. Ed. ʿAbd al-Ḥusayn al-Fatlī. 3 vols. Beirut: Muʾassasat ar-Risāla, 1985.

Jurjānī, *Dalāʾil* = ʾAbū Bakr ʿAbd al-Qāhir ibn ʿAbd ar-Raḥmān al-Jurjānī, *Dalāʾil al-ʾiʿjāz*. Ed. Muḥammad Riḍwān ad-Dāya and Fāyiz ad-Dāya. Damascus: Dār Qutayba, 1983.

Mubarrad, *Muqtaḍab* = ʾAbū l-ʿAbbās Muḥammad ibn Yazīd al-Mubarrad, *al-Muqtaḍab*. Ed. Muḥammad ʿAbd al-Xāliq ʿUḍayma. 4 vols. Cairo: Dār at-Taḥrīr, 1965–1968. (Repr., Beirut: ʿĀlam al-Kutub, n.d.)

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn Qanbar Sībawayhi, *al-Kitāb*. Ed. Hartwig Derenbourg. Paris: Imprimerie Nationale, 1881–1889. (Repr., Hildesheim: G. Olms, 1970.)

Suhaylī, *Natāʾij* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān ibn ʿAbdallāh as-Suhaylī, *Natāʾij al-fikr fī n-naḥw*. Ed. ʿĀdil ʾAḥmad Ibn al-Mawjūd and ʿAlī Muḥammad Muʿawwaḍ. Beirut: Dār al-Kutub al-ʿIlmiyya, 1992.

Suyūṭī, *Iqtirāḥ* = Jalāl ad-Dīn ʾAbū l-Faḍl ibn ʾAbī Bakr ʾIbrāhīm as-Suyūṭī, *al-Iqtirāḥ fī ʿilm ʾuṣūl an-naḥw*. Ed. ʾAḥmad Muḥammad Qāsim. Cairo: Maṭbaʿat as-Saʿāda, 1976.

Xalaf al-ʾAḥmar, *Muqaddima* = ʾAbū Muḥriz Xalaf ibn Ḥayyān al-ʾAḥmar al-Baṣrī (?), *al-Muqaddima fī n-naḥw*. Ed. ʿIzz ad-Dīn at-Tanūxī. Damascus: Wizārat aṯ-Ṯaqāfa wa-l-ʾIršād al-Qawmī, 1961.

Zajjājī, *ʾĪḍāḥ* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān ibn ʾIsḥāq az-Zajjājī, *al-ʾĪḍāḥ fī ʿilal an-naḥw*. Ed. Māzin al-Mubārak. Beirut: Dār an-Nafāʾis, 1970.

SECONDARY SOURCES

Åkesson, Joyce. 2001. *Arabic morphology and phonology based on the* Marāḥ al-Arwāḥ *by Aḥmad b. ʿAlī b. Masʿūd*. Leiden: E.J. Brill.

Baalbaki, Ramzi. 1979. "Some aspects of harmony and hierarchy in Sībawayhi's grammatical analysis". *Zeitschrift für Arabische Linguistik* 2.7–22.

——. 1985. "The treatment of *qirāʾāt* by the second and third century grammarians". *Zeitschrift für Arabische Linguistik* 15.11–32.

——. 1988. "A contribution to the study of technical terms in early Arabic grammar: The term *ʾaṣl* in Sībawayhi's *Kitāb*". *A miscellany of Middle Eastern articles: In memoriam Thomas Muir Johnstone*, ed. A.K. Irvine, R.B. Serjeant and G. Rex Smith, 163–177. Essex: Longman.

——. 1999. "Expanding the *maʿnawī ʿawāmil*: Suhaylī's innovative approach to the theory of regimen". *al-Abḥāth* 47.23-58.

——. 2003. "Some considerations on word order in *kāna* constructions". *Romano-Arabica* N.S. 3.41–58.

——. 2005. "Theoretical coherency versus pedagogical attainability: The conscious bias of Arab grammarians". *Alltagsleben und materielle Kultur in der arabischen Sprache und Literatur: Festschrift fur Heinz Grotzfeld zum 70. Geburstsdag*, ed. Thomas Bauer and Ulrike Stehli-Werbeck, 39–68. Wiesbaden: O. Harrassowitz.

Bohas, Georges. 1981. "Quelques aspects de l'argumentation et de l'explication chez les grammairiens arabes". *Arabica* 28.204–221.

—— and Jean-Patrick Guillaume. 1984. *Etude des théories des grammairiens arabes: Morphologie et phonologie*. Damascus: Institut Français d'Etudes Arabes.

Bohas, Georges, Jean-Patrick Guillaume, and Djamel Eddine Kouloughli. 1990. *The Arabic linguistic tradition*. London and New York: Routledge. (Repr., Washington, D.C.: Georgetown University Press, 2006.)

Fournier, Jean-Marie a.o. 2006. "L'exemple dans quelques traditions grammaticales (formes, fonctionnement, types)". *Langages* (forthcoming).

Guillaume, Jean-Patrick. 2000. "La nouvelle approche de la grammaire au IV^{e}–X^{e} siècle: Ibn Ğinnī (320/932–392/1002)". *History of the language sciences*, ed. Sylvain Auroux, Konrad Koerner, Hans-Josef Niederehe, and Kees Versteegh, I, 273–280. Berlin and New York: W. de Gruyter.

——. 2006. "Comment se débarrasser des 'mauvais' exemples? Le cas de la tradition grammaticale arabe". *Langages*, forthcoming.

Kinberg, Naphtali. 1996. *A lexicon of Farrāʾ's terminology in his* Qurʾān *commentary*. Leiden: E.J. Brill.

Kouloughli, Djamel Eddine. 1985. "À propos de *lafẓ* et de *maʿnā*". *Bulletin d'Etudes Orientales* 35.43–63.

Levin, Aryeh. 1994. "Sībawayhi's attitude to the spoken language". *Jerusalem Studies in Arabic and Islam* 17.204–243.

Loucel, Henri. 1963-1964. "L'origine du langage d'après les grammairiens arabes". *Arabica* 10.188–208, 253–281; 11.52–72, 151–187.

Owens, Jonathan. 1990. *Early Arabic grammatical theory: Heterogeneity and standardization*. Amsterdam and Philadelphia: J. Benjamins.

Suleiman, Yasir. 1999. *The Arabic grammatical tradition: A study in* taʿlīl. Edinburgh: Edinburgh University Press.

Talmon, Rafael. 1985. "Who was the first Arab grammarian? A new approach to an old problem". *Zeitschrift für Arabische Linguistik* 15.128–45.

——. 1986. "The *masʾala zunbūriyya* and the authenticity of *aḫbār* about the early controversies between the Baṣran and Kūfan schools of grammar". *al-Karmil* 7.131–63.

——. 1990. "Kitāb *Muqaddima fī n-naḥw* al-mansūb ʾilā Xalaf al-ʾAḥmar: Dirāsa wa-fihrist muṣṭalaḥāt". *al-Karmil* 11.129–99.

Versteegh, Kees. 1995. *The explanation of linguistic causes: Az-Zağğāğī's theory of grammar*. Amsterdam and Philadelphia: J. Benjamins.

——. 1997. *The Arabic language*. New York: Columbia University Press.

——. 2000. "Grammar and logic in the Arabic grammatical tradition". *History of the language sciences*, ed. Sylvain Auroux, Konrad Koerner, Hans-Josef Niederehe, and Kees Versteegh, I, 300–306. Berlin and New York: W. de Gruyter.

Weiss, Bernard G. 1974. "The Medieval discussions of the origin of language". *Zeitschift der Deutschen Morgenländischen Gesellschaft* 125.33–41.

Zwettler, Michael. 1978. *The oral tradition of Classical Arabic poetry: Its character and implications*. Columbus: Ohio State University Press.

JEAN-PATRICK GUILLAUME
(University Paris-3 and CNRS UMR 7597)

Grammatical Tradition: History

Arabic is unique among languages as the chosen medium of divine communication in a direct, complete revelation exclusively to a single prophet. That revelation has been preserved to this day in the document known as the *Qurʾān*.

The special character of Arabic did not discourage Muslims from exploring the language as a purely human vehicle, and they were easily able to separate the celestial from the sublunar Arabic to describe and analyze the language spoken in this world: for them Adam was certainly the first created person to speak Arabic in heaven, but on earth he spoke Syriac, and

Arabic was just one of many earthly languages that evolved naturally. These two kinds of Arabic were then recombined in the person of Muḥammad, whose mother tongue was miraculously identical with the language of God and Adam before the Fall.

It is not surprising, therefore, that almost every branch of Muslim scholarship involves some aspect of language, from Qurʾānic textual criticism and exegesis, theology, logic, jurisprudence, and legal semantics to the specifically linguistic sciences of rhetoric, poetics, syntax, phonology, morphology, and lexicography.

Of these areas of study, only syntax will be dealt with here in detail. Articulatory phonetics was highly developed in the Middle Ages (though not enough to constitute an independent discipline as in the modern *ʿilm al-ʾaṣwāt* 'the science of sounds') and served the dual purpose of ensuring correct pronunciation in Qurʾānic recitation and accounting phonologically for numerous morphological processes. Morphology itself became a specialized field of study by the 3rd/9th century as *ʿilm aṣ-ṣarf* lit. 'the science of the way [forms of words] circulate'. Morphology can fairly claim to have identified every known word pattern in the entire Arabic vocabulary (→ *ṣarf*). This vocabulary was in turn the object of lexicography, *ʿilm al-luġa* 'the science of [spoken, specifically idiomatic] language' (→ lexicography: Classical Arabic). Its data being purely empirical, lexicography allied itself methodologically with the science of *Ḥadīṯ*, so vocabulary items were collected and authenticated in the same way as the sayings of the Prophet, by observation and memorization, relying for their evidential value on the probity of the individuals in the chain of transmission.

Syntax, being the study of the arrangement of the elements of verbal communication, normally in the form of statements, offers much more scope for theoretical explorations than phonology or morphology, and it touches the central preoccupations of all the text-based religious sciences. Theologically, the nature of God's speech, *kalām Allāh*, and its relationship to the physical text of the *Qurʾān* were contentious issues, as was the problem of the origins of human language per se. Legal theory could not begin without first agreeing on the way meaning is expressed, both by God and His addressees, and how laws are derived from what God says to us in His *xiṭāb* 'allocution'. And, as we shall see, philosophers and logicians came into conflict with grammarians over the universality of logic against the specificity of the Islamic/Arabic revelation.

The grammatical tradition, therefore, is part of a larger development in which grammar marched in step with the other sciences to construct a doctrinal system in which every discipline had a complementary and mutually supportive role. Unlike modern theoretical linguistics, which aims, rightly or wrongly, to be context-free, Arabic grammar, even at its highest level of abstraction, must always justify itself by its relevance to Islamic beliefs.

Seven phases are used as reference points herein, without implying that the grammarians consciously located themselves at any particular point in the series, although many did: (1) primitive grammar, (2) the first systematic grammar, (3) the beginnings of pedagogy, (4) the evolution of a general theory, (5) the assimilation of grammar and law, (6) the elaboration of scholastic grammar, and (7) grammar since the Middle Ages.

1. Primitive grammar

In the decades after the death of the Prophet in 11/632, the immediate concern was to preserve the record of the revelation, which was threatened with loss and corruption as Muḥammad's surviving contemporaries died off. There was no single authorized text of the *Qurʾān*, and all efforts went into stabilizing the earliest versions, at first memorized and only later written down, during which the archaic and ambiguous orthography was enhanced by the addition of diacritical points and vowel markers. This was completed by about the middle of the 2nd/8th century, the same time in which the first exegetical works were written and, not unrelated, the first collections of pre-Islamic poetry were made in the search for the semantic bedrock of the Arab(ic) revelation. The undertaking presumes some awareness of linguistic entities, and there can be no doubt that a number of important grammatical terms came into use in this period, largely taxonomic, such as → *ḥaraka* 'vowel', → *ṣifa* 'adjective', and *waqf* 'pausing', as well as general notions of number, case, and gender. But research by Versteegh (1993) has shown that there was virtually no scientific analysis or processing during this stage, simply an honest and diligent compilation of data.

2. The first systematic grammar

With the arrival of Sībawayhi in Baṣra in the middle of the 2nd/8th century, the picture changed. Sībawayhi was attracted to Baṣra by the presence there of circles of scholars engaged in relatively advanced discussions of language, law, and creed, among them al-Xalīl ibn ʾAḥmad, the genius of his era. It was he who definitively revised the vowel markers of the *Qurʾān*, formalized the metrical system, and initiated the ordered collection of vocabulary in the first Arabic dictionary, and it was his vast personal corpus of linguistic data that was the main source for Sībawayhi, who became his devoted pupil.

But al-Xalīl did not create a science of grammar: this was Sībawayhi's achievement after his association with al-Xalīl and other like-minded intellectuals, several of whom are named in the *Kitāb* 'Book', as Sībawayhi's untitled treatise came to be known after his death.

The *Kitāb* is a massive exercise of induction based on the data supplied mostly by Sībawayhi's scholarly informants but also by a number of Bedouin native speakers. The result is a near-exhaustive description of Arabic within a coherent theoretical framework that treats language as a form of behavior. Speech (→ *kalām* 'talking') is categorized into about 80 linguistic acts, e.g. → *ibtidāʾ* 'starting [an utterance]', *taṯniya* 'making dual', *nafy* 'negating', *waṣf* 'describing [adjectivally]', *istifhām* 'asking a question'. These are all subjected to the same ethical criteria as legal acts, hence an utterance is called *ḥasan* 'good, beautiful' when it is structurally well formed, *qabīḥ* 'bad, ugly' if not, *mustaqīm* 'right, straight' when it conveys the intended meaning, and *muḥāl* 'wrong, perverted' when it fails to convey any meaning at all. Significantly, a *mustaqīm* utterance that conveys its intended meaning can still be *jāʾiz* 'permissible' even if it is *qabīḥ* 'ill formed', as often happens in poetry (→ poetic license). Structural correctness is determined by the *mawḍiʿ* 'place' in which an element occurs, and its range of functions by its *manzila* 'status'. Formally, there are only three categories of speech elements, → *ism* 'name [in grammar: noun]' (which includes adjectives and most adverbials), → *fiʿl* 'action [in grammar: verb]', and the → *ḥarf* 'bit, particle', an amorphous group that also comprises morphemes and phonemes. Speech is linear, so these elements occur in a string, with inflections resulting from the → *ʿamal* 'operation, effect' of one element upon (usually) the next. The guiding principle of language use, both for speakers and observers, is analogy (→ *qiyās*), often intuitive (see also → grammatical tradition: approach).

Sībawayhi accomplished an unprecedented systematic, comprehensive study of language; after him a science of grammar came into existence that was eventually termed (though not by Sībawayhi) *ʿilm an-naḥw* 'the science of the way [people speak]'.

3. The beginnings of pedagogy

The *Kitāb* of Sībawayhi is far too descriptive to be of use for teaching, but grammarians after him were soon faced with the need to promote a standard Arabic in order to maintain both the Islamic religion and the Muslim state. Here we can agree with Ibn Xaldūn about the corruption of the language arising from the conversion to Islam of more and more non-Arabs; the need for Arabic instruction led to the emergence of a professional class of Arabic teachers, with all the attendant rivalries and power struggles abundantly recorded in biographical literature.

Already within decades of Sībawayhi's death there are signs of pedagogical activity, and the earliest nonanecdotal evidence of Arabic being taught professionally (to children) is in a work of Ibn Saḥnūn, written before 256/870 (Lecomte 1953). The first pedagogical texts were in circulation soon after, such as the *Muxtaṣar fī n-naḥw* (Compendium on grammar) of Luġda (d. late 3rd/9th century) and the *Muwaffaqī* (named after his patron) of Ibn Kaysān (d. between 299/912 and 320/932), probably written for children. A number of more advanced grammars were created in the 4th/10th century and are still useful today, e.g. the *Mūjaz* 'Condensed' of Ibn as-Sarrāj (d. 316/929), the *Jumal fī n-naḥw* 'General statements about grammar' of az-Zajjājī (d. 339/949 or 340/950), and the *Lumaʿ* 'Illuminating flashes' of Ibn Jinnī (d. 392/1002). It is probably about this time that the first versified teaching grammars appeared, to judge from fragments attributed to Qalfāṭ (d. 302/914–915) in late sources (didactic poems credited to 2nd/8th-century grammarians are unconvincing). But these are not the great pedagogical masterpieces, in prose or verse, composed when the systematization of grammar was complete.

4. THE EVOLUTION OF A GENERAL THEORY

By the 3rd/9th century, Sībawayhi's type of grammar was under review, indeed threatened, from two sides. Among the grammarians there was a growing tension between those who regarded Sībawayhi's data as more or less exhaustive and those who believed that more data could always come to light. And from outside the grammatical community came the challenge from the logicians that they were better qualified than the grammarians to control the Arabic language and with it the Islamic ideology.

The issues were connected, as they stemmed from the realization that every science, such as grammar had now become, requires a sound theoretical basis. This had not been a problem for Sībawayhi because he simply transferred the ethico-legal reasoning of his day from the regulation of human behavior to linguistic behavior, but not long after his death the (re)translation of a number of Greek works forced the Arabs to take a position on the nature of the Islamic sciences, especially those dealing with theology, law, and language.

The internal conflict among the professional grammarians was sparked off by the question of authority. They had learned from the logicians that rules depend for their validity on the data from which they are inductively derived and that only a closed corpus could guarantee that these rules could never be overturned by new data. To their credit, everyone was well aware of this: the controversy, which would result in the famous division into 'Baṣran' (closed corpus) and 'Kūfan' (open corpus) grammatical schools, named after the two leading cultural centers before the foundation of Baghdad, was long and acrimonious, but grammarians never lost sight of the fact that grammatical science must draw its authority objectively from its logical structure and not, as had formerly been the case, subjectively from the personal prestige and strength of character of its leading practitioners.

It was inevitable that the Baṣrans would prevail, as their attitude was in harmony with parallel developments among theologians and jurists, who responded to the same problem with the well-known 'closing of the gate of *ijtihād*', deliberately restricting the corpus of religious texts from which they could derive the law by the exercise of their personal reasoning (*ijtihād*). The Baṣrans' way of closing the linguistic corpus was effectively to define it as the contents of Sībawayhi's *Kitāb*, to which hardly anything had been or ever would be added: as a result they could claim, as did the lawyers, that the proper use of analogical reasoning applied to a well-defined and authoritative text could provide answers to all linguistic or juridical questions. This left the Kūfans on the outside as nonconformists, and they never afterward played any significant role, although it is also true that allegiance to one or another school (there was also a 'Baghdad' school and perhaps others) was seldom crucial and often very inconsistent.

At the same time as the grammarians and others were dealing with the need to close the corpus, far more complex issues were being raised both internally among grammarians and externally in court circles, where philosophers and logicians publicly challenged the grammarians' authority. In the end, the grammarians were forced literally to organize their methodology according to the logical principles set out in Aristotle's *Organon* and related works.

There is no compelling evidence that Sībawayhi was aware of the exiguous literature in logic available to Arabic readers in the 2nd/8th century, and what few signs of logical concepts there are in the *Kitāb* can only have been absorbed informally, as there is no trace of any literary influence. But it was not long (and we can admire the Abbasid courts for their open and enthusiastic patronage of intellectual curiosity) before such notions entered grammatical works. Al-Mubarrad (d. 285/898) may be the first to offer the Aristotelian definition of a sentence as that which can be true or false, and other evidence of philosophical contacts is scattered throughout his large grammatical work *al-Muqtaḍab* 'Pruned'; (referring to his virtual abridgement of Sībawayhi's *Kitāb*), such as his formulation that the circumstantial qualifier (→ *ḥāl*) can be used to express only a transient quality, not a *jawhar*, a philosophical term meaning 'substance, *ousía*'.

In the following century, the rivalry between grammarians and logicians created a small literary genre recording their hostile confrontations. The most famous is the battle of words between ʾAbū Saʿīd as-Sīrāfī (d. 368/979) and the Christian ʾAbū Bišr Mattā ibn Yūnus (d. 328/940), which took place in the presence of the wazīr Ibn al-Furāt in 320/932 (ʾAbū Ḥayyān, *ʾImtāʿ* I, 107–129). The symbolism of the debate is at least as important as its content, which must here be reduced to a single issue, namely as-Sīrāfī's refutation of ʾAbū Bišr's claim that Arabic is only a particular instance of a universal logical code. This was an argument he was bound to lose. As a Christian and the leading Aristotelian scholar of the day, ʾAbū Bišr represented a double threat to Islam, as the sources of both his faith and his reasoning were non-Arab, in a period when the identification of Islam with the Arabs was at its peak. Not surprisingly, as-Sīrāfī tried to disqualify him from putting his case at all by declaring that he did not speak Arabic well enough, a not uncommon debating trick in such circles.

On the positive side, there is no doubt that the conflict between grammarians and logicians, like that between Baṣrans and Kūfans, resulted in radical changes in grammar as a science. While the grammarians eventually agreed to differ on the fundamental issue of induction from a closed corpus, the logicians taught them a great deal about categories and methods.

For categories, it will suffice to mention two kinds of innovation that came about during this phase. First, the gaps in Sībawayhi's terminology were filled, partly, perhaps, for pedagogical reasons but also because the imported definitions of the sciences presumed that their vocabulary was exhaustive. So we find → *tamyīz* 'specifying element' for structures such as *ʾašaddu ḥumratan* 'redder', lit. 'more intense as to redness'; *lā li-nafy al-jins* for 'categorical negative *lā*'; *ʾafʿāl al-qulūb* 'verbs of the heart' for mental verbs; and other neologisms for items that Sībawayhi never bothered to name, although they are all dealt with in the *Kitāb*. Several abstract nouns were coined for the same reasons, e.g. *fiʿliyya* 'verbality' for the quality of being a verb, *ḍarfiyya* for the quality of being a *ḍarf* 'adverbial complement', etc.

Second, there was a complete revision of the concept of communication. For Sībawayhi, the purpose of language was essentially ethical and pragmatic, namely, for the speaker to satisfy the listener's expectations by accurately conveying the speaker's intention (*murād* 'what is meant'), and it was linguistically irrelevant whether the utterances were true or false and even less so that they should be structurally complete or free of formal defects. For the grammarians of the 4th/10th century (perhaps even earlier, though less systematically), the unit of discourse was no longer → *kalām* 'talking' but the → *jumla* 'sentence', with a minimum of a subject and a predicate, and which, to qualify as a 'sentence' at all, had to be falsifiable, like a logical proposition. And the pragmatic criterion of satisfying the listener's expectations was replaced by the semantic (*scil.* abstract) prerequisite that the sentence/proposition should deliver *fāʾida* 'information'.

The origin of this new sense of *jumla* is obscure. Although it is common in all periods in the meaning of 'aggregate, general summary, totality', it entered the grammatical vocabulary only hesitantly in the meaning of 'sentence' in the early 3rd/9th century, and *kalām* remained in use alongside it for a long time until it eventually yielded to *jumla*. After this, *kalām* preserved only the overarching meaning of undifferentiated speech, with *jumla* covering all the subtypes of utterance we call sentences and clauses.

Methodologically, there was also a total rethinking because grammar now had to conform with universal scientific principles. Hitherto, it had been taken for granted that language was a rational phenomenon because it is an activity of rational beings, which made it possible to infer linguistic rules directly from the behavior of speakers. But Ibn as-Sarrāj (d. 316/929) introduced the fine distinction between the principles (*ʾuṣūl*) that a speaker applies to produce correct utterances and those the grammarian uses to account for the correctness of an utterance. The former are prescriptive, pedagogical, and deductive, while the latter are inductive and ensure that the science of grammar itself is rational.

These *ʾuṣūl* were the outcome of discussions of grammatical causes (→ *ʿilla*, pl. *ʿilal*), another concept brought in from Greek. In the earliest grammar, linguistic cause and effect was limited to the 'operation' (*ʿamal*) of one word upon another (often misinterpreted as 'government'). No abstract reasons for a phenomenon could be given beyond the practical (e.g. frequency, phonological convenience) or the psychological (e.g. communicative efficiency, pressure of analogy, the speaker's motives): one could say empirically, for example, that agents of verbs always occur in the independent (*rafʿ*) form, but one could not give a logical reason for it. By the 4th/10th century three kinds of causality were distinguished (even more were introduced later). Lowest in the hierarchy was the *ʿilla taʿlīmiyya* 'pedagogical reason', which was all that the language user had to know, e.g. that nouns operated on by *ʾinna* 'verily' take dependent (*naṣb*), i.e. direct object form. Then came the *ʿilla qiyāsiyya* 'analogical reason', which enabled both speakers and linguists to make sense of the rules for their own purposes, e.g., the reason *ʾinna* takes nouns in direct object form is that it resembles a verb. Finally, there was the *ʿilla jadaliyya naḏariyya* 'speculative-dialectical reason', e.g. what kind of verb *ʾinna* resembles and in what way (Suleiman 1999).

Needless to say, the search for the speculative-dialectical reason led to complex discussions among grammarians as well as between them and their rivals, the logicians. In the internal arguments between adherents of the two schools, the Kūfans matched their Baṣran opponents point for point in obscurity and hair splitting. From the *ʾĪḍāḥ* 'Clarification' of az-Zajjājī, it is likewise clear that grammarians and logicians too attained a very high theoretical level as they challenged each other over the fundamentals of grammatical science, although in the end az-Zajjājī, speaking no doubt for all his colleagues, felt obliged to declare that "[the logicians'] aims are not ours and our aims are not theirs" (*ʾĪḍāḥ* 38). Wholly unsympathetic to this line of inquiry were those few grammarians who rejected the notion of linguistic causality altogether on the grounds that every human act, including speech, is predetermined: the short and pungent *ar-Radd ʿalā n-nuḥāt* 'Refutation of the grammarians' by Ibn Maḍāʾ al-Qurṭubī (d. 592/1195–6) is the best known representative of this somewhat marginal view of language.

5. THE ASSIMILATION OF GRAMMAR AND LAW

At the same time that Ibn as-Sarrāj was writing on the principles of grammar (*ʾuṣūl an-naḥw*), his contemporaries in the legal sciences were occupied with a similar task, which came to fruition in works on the *ʾuṣūl al-fiqh* 'principles of jurisprudence'. What these disciplines had in common is that both depended on the interpretation of a textual corpus to derive rules for human behavior. They differed, of course, in the nature of their corpus, the corpus of the law being divine inspired while that of the language was Bedouin speech (the *Qurʾān* could not be the primary source of data for the grammarians). But it is not an exaggeration to say, indeed it was said by the Arabs themselves, that correct grammar (*naḥw*) was a subset of the orthodox practice (*sunna*) of the good Muslim. Nor is it a coincidence that *naḥw* and *sunna* are synonymous, both meaning 'way', none other than the *ṣirāṭ mustaqīm* 'straight path' that Muslims are enjoined to follow in the opening verses of the *Qurʾān*.

This shared character led to a kind of symbiosis between law and grammar, and increasingly there was a professional overlap in the two disciplines, so that a scholar might function as a judge and write grammatical works as well. The relationship was not always amicable. ʾAḥmad ibn Fāris (d. 395/1004) was very critical of the linguistic inadequacies of his legal brethren in his treatise *Kitāb aṣ-Ṣāḥibī fī fiqh al-luġa wa-sunan al-ʿArab fī kalāmihim*, whose title alone reveals the interpenetration of language and law: 'Book [dedicated to the wazīr] aṣ-Ṣāḥib on the jurisprudence of language and the habitual ways of the Arabs in their speech', where *sunan* 'habitual ways' is simply the plural of *sunna* 'orthodox Muslim practice'. He makes a strong plea for greater competence in Arabic among the jurists, from which we may infer that in his time some of them did not live up to that standard.

Full integration of grammar and law, both in goals and methods, is argued explicitly by Ibn al-ʾAnbārī (d. 577/1181) in his *Lumaʿ al-ʾadilla fī ʾuṣūl an-naḥw* 'Illuminating flashes on the evidence[s] for the principles of grammar', which sets out to demonstrate that the value of linguistic and legal evidence and the interpretation of the data are identical in both disciplines.

There is no better indicator of this relationship than the term *šāhid* 'legal [eye]witness', which also stood for 'item of linguistic testimony' centuries before Ibn al-ʾAnbārī. So close, in fact, are the two sciences that it is even possible to discern a correlation between the scholar's legal affiliation (there were important technical differences between the schools of law) and his grammatical preferences, for example over the degree of latitude permitted in making a discretionary choice where two pieces of equally valid evidence conflict.

6. The elaboration of scholastic grammar

Ibn al-ʾAnbārī wrote in a time of two great changes in Islamic civilization, one architectural, the other intellectual. Sometime in the late 4th/10th century the first dedicated educational buildings began to appear. Previously, teaching had been done in the mosque or the scholar's home, but although both continued to be used, the desire for specialized accommodation led to the establishment of the *madrasa* lit. 'place of study', often loosely translated 'college'. Its main purpose was to train jurists in the various schools of law, but the syllabus was quite broad, and there were professorial chairs, student stipends, libraries, and lodgings. Since it was a pious act to endow a madrasa, madrasas were soon found in every major town, often several, although, curiously, they never flourished in al-Andalus, where teaching remained in the mosques.

The intellectual analogue of these new buildings was scholasticism, a system of thought no less rigidly constructed than the bricks and stones of the madrasa. Scholasticism was a response to the pressure for knowledge to be packaged for the curriculum, requiring not only a sound theoretical basis, which had been largely worked out in the 4th/10th century, but also a style of presentation suitable for classroom teaching at different levels.

It is impossible here to do justice to the complexity of the process by which grammatical theory developed to its scholastic maturity. It was a vast communal exercise in which all the Islamic sciences consolidated their place in the educational system, each with its own definition, method, and technical vocabulary. This could not be accomplished until the sciences had become self-conscious enough to assert their own autonomy in the pivotal 4th/10th century. The *Mafātīḥ al-ʿulūm* 'Keys to the sciences' of al-Xwārazmī (written between 366/976 and 387/997) documents the advanced state of organized knowledge in this crucial stage.

A grammarian whose contribution has been undervalued is al-Xwārazmī's contemporary, ʾAbū ʿAlī al-Fārisī (d. 377/987). He was a prolific author, a serious student of the history of his discipline, possessor of several manuscripts of Sībawayhi, tireless in his investigation of grammatical phenomena, and a respected teacher who traveled widely – in short, the very best kind of committed scholar. But he has been overshadowed by his own pupil, the insatiably curious Ibn Jinnī (d. 392/1004), who frequently acknowledged his debt to his master. Another grammarian from the same period whose reputation outshines that of al-Fārisī is the Muʿtazilite ar-Rummānī (d. 384/994), famous for his relentless application of pure logic to the facts of grammar in his search for the ultimate *ḥikma* 'wisdom' of language. These scholars are witness to the almost unlimited intellectual freedom enjoyed before scholasticism imposed itself on the form and content of Muslim thought.

The first pedagogical grammars written expressly for the madrasa date to the 5th/11th century. The *Miʾa ʿāmil* 'One hundred operators' of al-Jurjānī (d. 471/1078) ruthlessly cuts up the whole subject into exactly one hundred very short pieces. The *Muqaddima* 'Introduction' of Ibn Bābašāḏ (d. 469/1077) is remarkable for distributing its contents neatly into ten categories, and the *Mufaṣṣal* of az-Zamaxšarī (d. 538/1144), as its title 'The subdivided' implies, consists of 759 well-planned but still somewhat arbitrary paragraphs. These works, which completely subordinate the natural language to the demands of pedagogical arrangement, are worlds apart from the textbooks of previous centuries. Moreover, the rewards of teaching at the madrasa encouraged scholars to produce more than one version of the same book, short, medium, and long, to suit the curriculum, and even to write commentaries on themselves. Needless to say, this was also the heyday of the supercommentary, gloss, and paraphrase industry.

The apogee of pedagogical grammar was reached in the 7th/13th century, in the works of the three great masters Ibn al-Ḥājib (d. 646/

1249), Ibn Mālik (d. 672/1274), and Ibn Hišām (d. 761/1360).

Two short treatises by Ibn al-Ḥājib, one on syntax, *al-Kāfiya* 'The sufficient', the other on morphology, *aš-Šāfiya* 'The effectual', represent the art of compression at its best, confirmed by the number of commentaries they have inspired, among them the *Kāfiya* by Raḍī d-Dīn al-ʾAstarābāḏī (d. 684/1285 or 686/1288), which still commands respect for its profound and as yet largely unappreciated subtlety. Ibn Mālik is famous for his use of verse as a pedagogical medium, and his *al-Xulāṣa al-ʾalfiyya* 'The thousand-line digest', better known simply as the *ʾAlfiyya* 'The thousand-liner', though not the only work of its kind (and far from the first), is the best-known instructional poem in the genre and has been the object of numerous commentaries, some of which have become as famous as the original, especially those by Ibn ʿAqīl (pupil of Ibn Mālik, d. 769/1367) and al-ʾUšmūnī (d. 872/1467). Ibn Hišām completes the trio with a series of pedagogical works that are such masterly statements of the rules and principles that they earned him the reputation of being 'a better grammarian than Sībawayhi'. Two of his best known works are his *Qaṭr an-nadā* 'Drops of dew' and *Muġnī l-labīb ʿan kutub al-ʾaʿārīb* 'All the intelligent man needs instead of books about inflections'.

These achievements were eclipsed by an obscure Moroccan schoolmaster who boiled down the syntax of Arabic to a dozen pages easy enough to be memorized, if not necessarily understood, by a seven-year-old child. The *Muqaddima al-ʾĀjurrūmiyya* 'The ʾĀjurrūmī introduction', named after its author, Ibn ʾĀjurrūm, of Fes (d. 723/1327), is without doubt the most widely known textbook of its kind and has spawned more than 60 commentaries. It was not the first elementary grammar to appear in this period: there is the *Miṣbāḥ* 'The lamp' of al-Muṭarrizī (d. 610/1213), composed for his young son, and the *Muqaddimat al-Ḥarīrī* 'Ḥarīrī's introduction' of al-Quhandizī (d. 666/1267), also written for juveniles. But the *ʾĀjurrūmiyya* was one of the first two grammars to be published in Europe (Rome 1592; the other was Ibn al-Ḥājib's *Kāfiya*), after which it came to be regarded, quite mistakenly, as typical of Arabic grammar and has had a disproportionate influence on Western attitudes toward the topic ever since (→ Arabic studies in Europe).

7. GRAMMAR SINCE THE MIDDLE AGES

After the 8th/14th century, serious and valuable works, invariably commentaries, continued to be produced. These include works by, among others, ad-Damāmīnī (827/1424), al-ʾAzharī (d. 905/1499), as-Suyūṭī (d. 911/1505), aš-Širbīnī (d. 977/1570), and aṣ-Ṣabbān (d. 1206/1792), all perpetuating the medieval scholastic mode, although the individuality of the author occasionally breaks through. Even when Lebanese scholars began to revive interest in the Arabic literary heritage, they expressed themselves in the medieval style, as in the grammatical works of Germanus (Jarmānūs) Farḥāt (d. 1732), Nāṣīf al-Yāzijī's (d. 1871), and Fāris aš-Šidyāq (d. 1887). By this time we are well into the colonial era, when the Arabic language began to fall under the intellectual dominance of the West. Establishment of the Arab academies in the early 20th century and the increase in vernacular literature are both symptoms of the impact of Western cultural values on the Arab world. In the second half of the 20th century, a new factor emerged: the large number of Arab linguists trained in the West, whose role in the teaching and preservation of Classical Arabic (as it should now be called) is far from clear. But some do claim the right to play a part, and their dealings with the traditional grammarians are likely to be as tense as those encounters between grammarians and logicians in early Islam, and for similar reasons.

To date the most striking postcolonial phenomenon is the movement to simplify Arabic, going back at least as far as ʾIbrāhīm Muṣṭafā, whose *ʾIḥyāʾ an-naḥw* 'Revival of grammar' was first published in 1937 and sparked a series of attempts at language reform that are still being energetically but inconclusively pursued. Ironically, the Andalusian fundamentalist Ibn Maḍāʾ al-Qurṭubī has been involuntarily coopted as the patron saint of reform, even though he would probably not be in favor of anything that threatened to draw people away from the purest classical language, which he himself had no intention of simplifying. It has to be said that with the present increase in fundamentalism, efforts to simplify or modernize Arabic are likely to be resisted if they seem to weaken the bond between contemporary Arabic and the

Islamic revelation, but this simply brings us back to our starting point: Arabic grammar is an integral part of the Islamic faith.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

The most popular works appear in many editions but are more usually read with their commentaries and supercommentaries, especially true for the *ʾĀjurrūmiyya* and the *ʾAlfiyya.*

ʾAbū Ḥayyān, *ʾImtāʿ* = ʾAbū Ḥayyān at-Tawḥīdī, *Kitāb al-ʾimtāʿ wa-l-muʾānasa.* Ed. ʾAḥmad ʾAmīn and ʾAḥmad az-Zayn. Cairo, 1939–1944. (Repr., Saida and Beirut: al-Maktaba al-ʿAṣriyya li-ṭ-Ṭibāʿa wa-n-Našr, 1953.)

ʾAstarābāḏī, *Šarḥ al-Kāfiya* = Raḍī d-Dīn Muḥammad ibn Ḥasan al-ʾAstarābāḏī, *Šarḥ Kāfiyat Ibn Ḥājib.* 2 vols. Beirut: Dār al-Kutub al-ʿIlmiyya, 1969.

——, *Šarḥ aš-Šāfiya* = Raḍī d-Dīn Muḥammad ibn Ḥasan al-ʾAstarābāḏī, *Šarḥ Šāfiyat Ibn al-Ḥājib.* Ed. Muḥammad Nūr al-Ḥasan, Muḥammad az-Zafzāf, and Muḥammad Muḥyī d-Dīn ʿAbd al-Ḥamīd. 4 vols. Repr. from the Cairo ed., Beirut: Dār al-Kutub al-ʿIlmiyya, 1982.

ʾAzharī, *Ḥāšiya* = aš-Šayx Xālid al-ʾAzharī, *Ḥāšiya ʿalā l-ʾĀjurrūmiyya.* Cairo, n.d.

——, *Taṣrīḥ* = aš-Šayx Xālid al-ʾAzharī, *at-Taṣrīḥ bi-maḍmūn at-Tawḍīḥ.* Cairo, n.d.

Ibn ʾĀjurrūm, *al-Muqaddima al-ʾĀjurrūmiyya.* (Numerous editions; see the commentaries by al-ʾAzharī and aš-Širbīnī.)

Ibn al-ʾAnbārī, *ʾInṣāf* = ʾAbū l-Barakāt ʿAbd ar-Raḥmān ibn Muḥammad al-ʾAnbārī, *Kitāb al-ʾinṣāf fī masāʾil al-xilāf bayna n-naḥwiyyīna l-basriyyīn wa-l-kūfiyyīn.* Ed. Gotthold Weil. Leiden: E.J. Brill, 1913. (Ed. Muḥammad Muḥyī d-Dīn ʿAbd al-Ḥamīd. 2 vols. Cairo: al-Maktaba at-Tijāriyya, 1955.)

——, *Lumaʿ* = ʾAbū l-Barakāt ʿAbd ar-Raḥmān ibn Muḥammad al-ʾAnbārī, *Lumaʿ al-ʾadilla fī ʾuṣūl an-naḥw.* Ed. Saʿīd al-ʾAfġānī. 2nd ed. Beirut: Dār al-Fikr, 1971. (Ed. Attia Amer. Stockholm: Almqvist and Wiksell, 1963.)

Ibn ʿAqīl, *Šarḥ* = Bahāʾ ad-Dīn ʿAbdallāh Ibn ʿAqīl, *Šarḥ Ibn ʿAqīl ʿalā ʾAlfiyyat Ibn Mālik.* Ed. Ramzī Munīr Baʿalbakī. Beirut: Dār al-ʿIlm li-l-Malāyīn, 1992. (German trans. Friedrich Dieterici. *Ibn ʿAkils Commentar zur* Alfiyya *des Ibn Mâlik.* Berlin, 1852.)

Ibn Bābašāḏ, *Muqaddima* = Ibn Bābašāḏ, *al-Muqaddima al-muḥsiba.* Ed. H. ʿAbd al-Karīm. Kuwait, 1976–1977.

Ibn Fāris, *Ṣāḥibī* = ʾAbū l-Ḥusayn ʾAḥmad Ibn Fāris, *Kitāb aṣ-Ṣāḥibī fī fiqh al-luġa wa-sunan al-ʿArab fī kalāmihim.* Ed. Moustafa Chouémi [Muṣṭafā aš-Šuʾaymī]. Beirut: A. Badran, 1964.

Ibn Hišām, *Muġnī* = Jamāl ad-Dīn ʾAbū Muḥammad ʿAbdallāh Ibn Hišām, *Muġnī l-labīb ʿan kutub al-ʾaʿārīb.* Cairo, 1885. (Ed. Māzin al-Mubārak and Muḥammad ʿAlī Ḥamdallāh. Damascus: Dār al-Fikr, 1969; numerous reprints of Cairo edition.)

——, *Qaṭr* = Jamāl ad-Dīn ʾAbū Muḥammad ʾAbdallāh Ibn Hišām, *Qaṭr an-nadā wa-ball aṣ-ṣadā.* French trans. Antonin Goguyer. Paris, 1888. (Numerous editions.)

Ibn Jinnī, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ fī n-naḥw.* Ed. Muḥammad ʿAlī an-Najjār. 3 vols. Cairo, 1371/1952–1376/1956.

Ibn Kaysān, *Muwaffaqī* = ʾAbū l-Ḥasan Muḥammad ibn ʾAḥmad Ibn Kaysān, *Kitāb al-Muwaffaqī fī n-naḥw.* Ed. ʿAbd al-Ḥusayn al-Fatlī and H.T. Šilās. *al-Mawrid* 4:2 (1975), 103–124.

Ibn Maḍāʾ, *Radd* = ʾAbū l-ʿAbbās ʾAḥmad Ibn Maḍāʾ al-Qurṭubī, *Kitāb ar-radd ʿalā n-nuḥāt.* Ed. Šawqī Ḍayf. Cairo: Dār al-Fikr al-ʿArabī, 1947.

Ibn Mālik, *ʾAlfiyya* = Jamāl ad-Dīn ʾAbū ʿAbdallāh Muḥammad ibn ʿAbdallāh Ibn Mālik, *al-ʾAlfiyya.* Cairo: Maṭbaʿa ʿĪsā al-Bābī al-Ḥalabī, n.d.

Ibn as-Sarrāj, *Mūjaz* = Muḥammad ibn as-Sarī Ibn as-Sarrāj, *Kitāb al-mūjaz fī n-naḥw.* Ed. Muṣṭafā aš-Šuʾaymī and Bin Sālim Dāmirjī. Beirut: al-Maktaba al-Luġawiyya al-ʿArabiyya, 1965.

——, *ʾUṣūl* = ʾAbū Bakr Muḥammad ibn Sahl Ibn as-Sarī, *al-ʾUṣūl fī n-naḥw.* Ed. ʿAbd al-Ḥusayn al-Fatlī. 3 vols. Beirut: Muʾassasat ar-Risāla, 1985.

Ibn Yaʿīš, *Šarḥ* = Muwaffaq ad-Dīn ʾAbū l-Baqāʾ Yaʿīš Ibn Yaʿīš, *Šarḥ al-Mufaṣṣal.* 2 vols. Beirut, n.d. (Ed. Gustav Jahn. Leipzig, 1882–1886.)

Jurjānī, *Miʾa ʿāmil* = ʾAbū Bakr ʿAbd al-Qāhir ibn ʿAbd ar-Raḥmān al-Jurjānī, *Miʾa ʿāmil.* [Many editions; 1st European ed. Erpenius, Leiden, 1617.]

Luġda, *al-Muxtaṣar* = ʾAbū ʿAlī al-Ḥasan ibn ʿAbdallāh Luġda, *al-Muxtaṣar fī n-naḥw.* Ed. ʿAbd al-Ḥusayn al-Fatlī. *al-Mawrid* 3:3 (1974), 221–246.

Mubarrad, *Muqtaḍab* = ʾAbū l-ʿAbbās Muḥammad ibn Yazīd al-Mubarrad, *al-Muqtaḍab.* Ed. Muḥammad ʿAbd al-Xāliq ʿUḍayma. 4 vols. Cairo: Dār at-Taḥrīr, 1965–1968.

Quhandizī, *Muqaddima* = al-Quhandizī, *Muqaddimat al-Ḥarīrī.* Lucknow, 1262 A.H. (Also many other editions.)

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb.* Ed. Hartwig Derenbourg. *Le livre de Sîbawaihi (Kitāb Sībawaihi), Traité de grammaire arabe.* 2 vols. Paris, 1881–1889. (Repr. 1970.) (Ed. Bulaq, 2 vols., 1316 A.H., repr. Baghdad: al-Muthanna Library, n.d.; ed. ʿAbd as-Salām Hārūn, 5 vols., Cairo, 1966–1977.)

Širbīnī, *Nūr* = Muḥammad aš-Širbīnī al-Xaṭīb, *Nūr as-sajiyya fī ḥall ʾalfāḏ al-ʾĀjurrūmiyya.* Ed. and English trans. Michael G. Carter. *Arab linguistics: An introductory classical text with translation and notes.* Amsterdam: J. Benjamins.

Suyūṭī, *ʾAšbāh* = Jalāl ad-Dīn ʾAbū l-Faḍl ʿAbd ar-Raḥmān ibn ʾAbī Bakr as-Suyūṭī, *al-ʾAšbāh wa-n-naḏāʾir.* 4 vols. Ed. Beirut: Dār al-Kutub al-ʿIlmiyya, 1984. [Ed. Hyderabad, 1940–1942.]

ʾUšmūnī, *Manhaj* = ʾAbū l-Ḥasan ʿAlī ibn Muḥammad al-ʾUšmūnī, *Manhaj as-sālik ʾilā ʾAlfiyyat Ibn Mālik.* Ed. Muḥammad Muḥyī d-Dīn ʿAbd al-Ḥamīd. Cairo: Maktabat an-Nahḍa al-Miṣriyya, 1970. [Also many other editions.]

Xwārazmī, *Mafātīḥ* = ʾAbū ʿAbdallāh Muḥammad ibn ʾAḥmad al-Xwārazmī, *Kitāb mafātīḥ al-ʿulūm.* Ed. Gerlof van Vloten. Leiden: E.J. Brill, 1895.

Zajjājī, *ʾĪḍāḥ* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān ibn ʾIsḥāq az-Zajjājī, *al-ʾĪḍāḥ fī ʾilal an-naḥw.* Ed. Māzin al-Mubārak. Cairo: Dār al-ʿUrūba, 1959. (English trans. Kees Versteegh. *The explanation of* *linguistic causes: az-Zağğāğī's theory of grammar*. Amsterdam and Philadelphia: J. Benjamins, 1995.)

——, *Jumal* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān ibn ʾIsḥāq az-Zajjājī, *Kitāb al-jumal fī n-naḥw*. Ed. Muhammad Ben Cheneb. Paris: Klincksieck, 1957. (Ed. ʿAlī Tawfīq al-Ḥamad. Beirut and Irbid: Muʾassasat ar-Risāla and Dār al-ʾAmal, 1984.)

Zamaxšarī, *Mufaṣṣal* = ʾAbū l-Qāsim Maḥmūd ibn ʿUmar az-Zamaxšarī, *al-Mufaṣṣal*. Ed. Jens Peter Broch. Christianiae: Libraria P.T. Mallingii, 1859. (2nd ed., 1879.)

Secondary sources

The titles below are a few items that survey the whole tradition and the principal reference works where more detailed information on names, concepts, and further reading can be found. Two specialist works cited in the article are also listed.

Auroux, Sylvain, Konrad Koerner, Hans-Joseph Niederehe, and Kees Versteegh (eds.). 2000. *History of the language sciences*, I. Berlin and New York: de Gruyter. (Section 9, "The establishment of Arabic linguistics", 245–336.)

Bohas, Georges, Jean-Patrick Guillaume, and Djamel Eddine Kouloughli. 1990. *The Arabic linguistic tradition*. London and New York: Routledge.

Carter, Michael G. 1990. "Grammar". *The Cambridge history of Arabic literature: Religion, learning and science in the ʿAbbasid period*, ed. M.J.L. Young, J.D. Latham, and R.B. Serjeant, 118–138. Cambridge: Cambridge University Press. [Bibliography, 532–533.]

——. 1994. "Writing the history of Arabic grammar". *Historiographia Linguistica* 21.387–416.

——. 2004. *Sībawayhi*. London and New York: I.B. Tauris.

Fleischer, Heinrich Leberecht. 1885–1888. *Kleinere Schriften*. Leipzig.

Goldziher, Ignaz. 1994. *On the history of grammar among the Arabs, an essay in literary history*. Transl. and ed. Kinga Dévényi and Tamás Iványi. Amsterdam and Philadelphia: J. Benjamins.

Lecomte, Gérard. 1953. "Le livre des règles de conduite des maîtres d'école par Ibn Saḥnūn". *Revue des Etudes Islamiques* 21.77–105.

Meisami, Julia S. and Paul Starky (eds.). 1998. *Encyclopedia of Arabic literature*. New York: Routledge.

Sezgin, Fuat. 1984. *Geschichte des arabischen Schrifttums*. IX. *Grammatik bis ca. 430 H.* Leiden: E.J. Brill.

Stammerjohann, Harro (ed.). 1996. *Lexicon grammaticorum: Who's who in the history of world linguistics*. Tübingen: Niemeyer. (2nd ed., 2005.)

Suleiman, Yasir. 1999. *The Arabic grammatical tradition: A study in* taʿlīl. Edinburgh: Edinburgh University Press.

Versteegh, Kees. 1987. "Arabische Sprachwissenschaft (Grammatik)". *Grundriß der arabischen Philologie*, II, ed. by Helmut Gätje, 148–176. Wiesbaden: L. Reichert.

——. 1993. *Arabic grammar and Qurʾanic exegesis in early Islam*. Leiden: E.J. Brill.

——. 1997. *The Arabic linguistic tradition*. London and New York: Routledge.

Michael G. Carter (Sydney University)

Grammaticalization

Grammaticalization (Arabic *intiḥāʾ*) is commonly defined, in Kuryłowiczian (1965) terms, as a gradual evolutionary process of change whereby contentive lexical units and structures acquire grammatical meanings and functions and less-grammatical forms become more grammatical. The term 'grammaticalization', while selected here for its widespread use and without theoretical predilections, has alternatives: 'grammacization' (Hopper 1991) and 'grammatization' (Matisoff 1991), which stand for divergent theoretical underpinnings.

The term 'grammaticalization' is French in origin, owing its first use to the Indo-Europeanist Meillet in 1912. Nonetheless, the notion of grammaticalization outside of Western scholarship harks back to 10th-century China (see Heine a.o. 1991; Lehmann 1995, for a historical overview), or even farther back to the Greeks (see Harris and Campbell 1995). While there is growing research interest in this subfield of linguistics, systematic studies of grammaticalization in Arabic are conspicuously lacking, despite widespread evidence of the phenomenon in Modern Standard Arabic (MSA) and modern dialects alike.

Grammaticalization is acknowledged generally at synchronic and diachronic levels. Synchronically, it accounts for the existence of multiple usages of a single form at a given point in time and their relative degrees of grammaticalization, hence assuming dynamism in language motivated by semantics and pragmatics. Diachronically, it sheds light on the evolutionary tracks of a given form. In most accounts treating the grammaticalization phenomenon, a panchronic account combining the two perspectives is used (see Hopper and Traugott 1993). In the literature on grammaticalization, crosslinguistic frequency in the development of grammaticalized forms (called 'grams') suggests that the transition from lexical to grammatical status or from grammatical to even more grammatical status follows a predictable conceptual track called pathway, cline, channel, and chain. Crosslinguistic attestations of the evolution of grammatical forms along these pathways preclude the coincidence of grammaticalization. Along these grammaticality pathways, progression from concrete sources toward more abstract end points or 'targets', with numerous intermittent salient functions that cluster at various points with those closer to the lexical source, are more concrete; the closest to the target are more mature abstract forms. The relative order of a given form to the one preceding or succeeding reflects the degree of grammaticality. The precise location of these points is hard to determine, a fact that underlies in the study of grammaticalization the lack of clear boundaries among categories establishing a conceptual continuum in their evolution. The changes (semantic, morphosyntactic, and phonological) that each lexical or less-grammatical item undergoes on its path toward grammaticalization are said to be unidirectional, as shown in Figure 1 (Hopper and Traugott 1993:7).

Figure 1. Hopper and Traugott's cline of grammaticality

content word > grammatical word > clitic > inflectional affix

Basic to the grammaticalization phenomenon is the unidirectionality principle entailing the movement away from the erstwhile lexical form and progressing gradually toward a more grammaticalized morpheme status; for example, once an originally lexical item becomes a clitic, it does not regress to one of its earlier forms. This does not exclude the possibility of the erstwhile lexical item coexisting with the non-lexical grammaticalized form (dubbed 'divergence' in Hopper 1991:24). This evolutionary process figures in the development of the future marker in Modern Standard Arabic, *sa-/sawfa*. This marker has its origin in a noun denoting 'patience, procrastination' that showed full declension and took the definite article, as in *fulānun yaqtātu s-sawfa* 'so-and-so lives on wishes' (Ibn Manẓūr, *Lisān* III, 2152). As a future marker, it lost its ability to be definite and to decline for case and became restricted to co-occurrence with verbs with imperfect stem having indicative inflection. Advancing toward further grammaticalization, the reduced form *sa-* became a particle prefixed to the same class of verbs. Although the noun is no longer used in Modern Standard Arabic, this does not necessarily mean that the original lexeme ceases to exist as a result of advancing toward cliticization or affixation in all cases. In a variety of instances, the two forms – the lexical (or the less-grammaticalized form) and the grammatizalized form – coexist side by side. Such is the case of *sawfa* and *sa-*, which are still used interchangeably in Modern Standard Arabic and negated by means of *lā* 'not' and *lan* 'will never'.

Another important observation about grammaticalization is the fact that it usually begins with a lexical item but need not terminate in an inflectional affix. Lexical items exit their pathway of grammaticalization at points prior to completion of full grammaticalization, as is the case of various nouns and adjectives in Arabic that grammaticalize as adverbs (a 'category conversion' termed 'adverbialization'; see Lehmann 1982:172 cited in Lessau 1994: I, 102), and do not progress further along the grammaticalization pathways (e.g. *ʾaḥyānun* 'times' > *ʾaḥyānan* 'sometimes', *kaṯīrun* 'much' > *kaṯīran* 'a lot, a great deal'). More important, inflectional affixes are not perceived as the end result of grammaticalization (Givón 1979:209; Lehmann 1995:13), which suggests that morpheme loss or reduction to zero is possibly a final destination for some grammaticalized forms, as Figure 2 depicts.

Figure 2. Givón's grammaticalization cycle

discourse > syntax > morphology > morphophonemics > zero

As forms and structures change functions and become more grammatical, they undergo changes in semantic, pragmatic, syntactic, morphological, and phonological features and distribution, which entail losses as well as countervailing gains (Hopper and Traugott 1993:87–88). In the case of the future marker discussed above, among the most notable losses are significant reduction of semantic content (variously labeled → 'semantic bleaching', 'fading', and 'weakening'); membership in open word classes; syntactic independence; and phonological material. Among the gains are change in meaning (also called 'pragmaticization'; see Lessau 1994:II, 675–676); widened range of contexts and greater frequency; increase in semantic generality; and becoming a member of a paradigm (see below, Sec. 2.5).

Different researchers have paid attention to different aspects of the lexical items and constructions recruited for grammaticalization. Heine a.o. (1991:32–33), for instance, emphasize the high level of generality of semantic content, concreteness, and basicness to human experience and hence the cultural independence of these sources. Bybee a.o. (1994:11), on the other hand, ascertain that source concepts "encode major orientation points in human experience", and furthermore, these authors broaden their view by taking the morphology and syntax of source constructions into consideration. Consistent with these observations and crosslinguistic evidence, certain lexical items are recruited while others are excluded from the grammaticalization process: body part terms used as nouns for signaling spatial relations grammaticalize as prepositions (e.g. *fū* 'mouth' > *fī* 'in, inside'; *wasaṭ* 'waist' > *wasṭa* 'middle, in the middle of'; *xalfu* 'rear' > *xalfa* 'in back of'; *qadam*, 'foot' > *quddām* 'in front of'), but not semantically specific ones (e.g. *rimš* 'lashes', *surra* 'navel', or *kāḥil* 'heel'). Likewise, *kāna* 'to exist, be', *ʾaxaḏa* 'to take, begin', and *qaʿada* 'to sit down; to continue' are grammaticalized as auxiliaries, but not *saraqa* 'to steal', *qarfaṣa* 'to squat'. Such selectional restrictions entail the nonarbitrariness of grammaticalization (see Lessau 1994:I, 58–59).

Although the foregoing criteria and examples may suggest a one-to-one correspondence between a source concept and a target, the relationship contrariwise is one to many. Grammaticalization of a single source could result in several targets: the grammaticalized forms resulting from the noun *sawāʾu* 'similar, same' include the clause linker *sawāʾun* 'whether', the exceptive particle *siwā* 'except', and *lā siyyamā* 'particularly, in particular'. These cases exemplify what Craig (1991) terms 'polygrammaticalization'. Syntactic and semantic ambiguity may become the consequences of polygrammaticalization. Such a state of affairs obtains in the usage of *ḥattā* as a preposition in *ʾakaltu s-samakata ḥattā raʾsihā* 'I ate the fish to its head' (Mubarrad, *Muqtaḍab* II, 37), which contrasts with its use as a particle in *ʾakaltu s-samakata ḥattā raʾsahā* (Ibn Hišām, *Muġnī* I, 115), where the meaning becomes 'I ate the fish, even its head'.

After assuming grammatical functions, grams, as a token of their increased grammaticality, may continue to enter into other morphosyntactic relations that were not available to them previously. For example, the verb *ʾaxaḏa* 'to take', when used as a lexical verb, e.g. *ʾaxaḏa l-waladu l-kitāba* 'the boy took the book', may take a rational being as its subject; as such, it may inflect for tense, change word order with its subject, undergo passivization, form an imperative, and be negated. When grammaticalized (e.g. *ʾaxaḏat ar-rīḥu taʿṣifu* 'the wind began to storm'), its syntactic position becomes restricted. It also embeds another finite verb and may not inflect for other tenses. But the grammaticalized *ʾaxaḏa* may take an inanimate subject while embedding a finite verb. In the latter case, *ʾaxaḏa* belongs to a closed class of verbs traditionally labeled 'verbs of beginning' (*ʾafʿāl aš-šurūḥ/al-ʾinšāʾ*).

Grammaticalization very often involves the entire construction in which the gram occurs (see Hopper and Traugott 1993:82). For instance, the active participle *lāzim* 'necessary' is grammaticalized as part of a construction that includes a preposition and a particle *min al-lāzimi ʾan* 'it is necessary that'. In advanced stages of grammaticalization, such paraphrases become severely reduced. This occurs when only *lāzim* substitutes for the entire construction in modern Arabic dialects. Reduction in the structural scope of this kind is labeled 'condensation' (Lehmann 1995:143, 164).

1. Mechanisms for grammaticalization

Several mechanisms motivating semantic change leading to grammaticalization of lexis and constructions have been suggested (→ mechanisms of linguistic change), chief among them metaphor (Heine a.o. 1991), metonymy (Traugott and König 1991; Heine a.o. 1991), → reanalysis, and → analogy (Hopper and Traugott 1993). While most researchers agree that these mechanisms are frequently involved in grammaticalization, consensus on sequence and relative significance of mechanisms has not yet emerged. Since these mechanisms occur outside of grammaticalization, it cannot be concluded that they are a necessary or sufficient condition for grammaticalization. For example, the use of the body part *raʾsun* 'head' in metaphorical constructions such as *raʾsumālin* 'financial capital' or *raʾsu s-sana*

'New Year's Day' did not result in the body part *ra's* 'head' assuming a grammatical function.

That said, interaction between one or more mechanisms can result in grammaticalization, as illustrated in the example below of the development of the negative suffix *-š* in some spoken Arabic varieties. It is plausible to reconstruct an initial stage where preverbal negation particle *mā* was used alone in negation of the perfect or verbal nouns as in *mā bi-wuddī* 'I do not want/like'. Negation structure of this type could have been strengthened optimally by the addition of the noun *šay'un* 'thing' in the object (postverbal position). This is in line with the use of negation emphasizers in such position in Classical Arabic, such as *qaṭṭ* 'time', *al-batta* 'decidedly', as in *mā ra'aytu miṯlahu qaṭṭ* 'I have not seen the like of him' (Ibn Manẓūr, *Lisān* V, 3672). Thus, *mā bi-wuddī šay'un* underwent reanalysis where the preposition *bi-* is reanalyzed as part of the nominal stem *wudd,* resulting in *bidd-ī* or *biddī* 'I want/desire'. It is highly likely that *šay'un* became *šay'an* – having lost its inflectional variability with regard to case, number, and definiteness prior to changing from optional emphasizer, via reanalysis, to an obligatory part of the negation construction. Such a use of *šay'an* is attested in the 3rd century A.H. in Ibn Wahb (*Jāmiʿ* 38.2, cited in Hopkins 1984:167) *wa-lam yuṣibhu šay'an* 'and nothing struck him'. In later stages, *šay'an* underwent phonological reduction or attrition resulting in the enclitic *-š* (for phonological attrition, see Lehmann 1995:126). This step marks a loss of categorial status from a noun to a clitic and simultaneous gain in frequency.

The next step is the emergence of the discontinuous negation pattern, the circumfix *mā* X *-š*. A further development yielded a construction where the original negation particle *mā* becomes optional and the clitic *-š* becomes sufficient for carrying out the negation function alone in certain contexts such as *biddīš* (Holes 1995:202, n. 54).

Via analogical extension, the enclitic *-š* became a fossilized component of the derivative negation particle *miš* in Egyptian and its variant *māši* in Moroccan Arabic, and is further used in predicate negation constructions involving the future in some Arabic dialects, particularly Egyptian. Due to the high level of internal cohesion of *miš*, it is not easily broken down into its componential parts (Brustad 2000:313). The choice of the lexical form *šay'un* 'thing' for grammaticalization is likely because of its general semantic content, which determined its suitability for grammatical functions in varying constructions. When it underwent phonological attrition, it also lost its semantic content, evidencing desemanticization (see Sec. 2.1), by which the original general meaning is lost altogether, resulting in greater abstraction of the enclitic remnant *-š* to the point that it now cooccurs alongside *ḥāga* 'thing' in spoken Egyptian (*mafīš ḥāga* 'there is nothing'), without any hint of redundancy. The presence of *šay'* 'thing' along with the grammaticalized enclitic form *-š* in Arabic may be regarded as an instance of what Hopper and Traugott (1993) label 'splitting' or 'divergence'.

2. MAJOR PRINCIPLES FOR GRAMMATICALIZATION

2.1 *Desemanticization*

'Desemanticization' first appeared in Lehmann (1995); it refers to the partial or total loss of referential meaning or erstwhile semantic sense of a lexical item. Desemanticization marks the first step and continues along the path to grammaticalization. When Ibn Barrī describes auxiliaries such as → *kāna wa-'axawātuhā* 'to be and its sisters' as well-grammaticalized verbs of motion, e.g. *jā'a* (details in Ibn Manẓūr, *Lisān* V, 3962), he might well be speaking of this aspect of the grammaticalization phenomenon when he describes their emptied semantic content and their obligatory occurrences in constructions to mark grammatical relations. An auxiliary such as *kāna*, originally denoting existence, is 'bleached' out of its semantic content. The intimate connection between semantic and pragmatic factors in grammaticalization (hinted at in Hopper and Traugott 1993 and in Bybee a.o. 1994) correctly predicts that the more generalized a gram is, the wider its domain of applicability, which holds true for *kāna*. As an auxiliary, *kāna* embeds a wide range of word classes, including participles and verbs inflected for past and non-past. In some instances, it marks grammatical relations even without contributing to the sentence or clause meaning, e.g. *yajibu 'an yakūna ʿindahu ʿuḏrun* 'he must have an excuse/he should have an excuse', where the embedded auxiliary *kāna* is semantically vacuous but ensures grammaticality of the sentence.

2.2 *Layering*

The existence of forms and constructions with nearly identical function in Arabic may be attributable to successive layers of grammaticalization. An example is the existence of a construction of the type verb + verbal noun (*maṣdar*) *ʾuḥibbu l-qirāʾata* 'I like reading' and its near-equivalent type, the particle *ʾan*, which triggers the use of the subjunctive as in *ʾuḥibbu ʾan ʾaqraʾa*. Other examples are the use of relative pronouns in clauses such as *al-muraššaḥu llaḏī yutawaqqaʿu ntixābuhu* 'the candidate whose election is expected' and its equivalent expressed with a participial form, *al-muraššaḥu l-mutawaqqaʿu ntixābuhu*, and the use of passive verbs with internal vowel pattern alteration such as *uftutiḥa l-maʿriḍ* 'the exhibition was opened', alongside the periphrastic construction *tamma ftitāḥu l-maʿriḍ*, in which the verb is active in form but with acquired passive sense. Another verb in the incipient stage of its grammaticalization in Modern Standard Arabic is *qāma* 'to stand up'. When used in the Arab media, scripted and otherwise, it forms a periphrasis with the following prepositional phrase, for example, *qāmat al-ḥukūmatu bi-ʾiʿdādi taqrīrihā* 'the government prepared its report'. In such a construction, the erstwhile postural verb has lost much of its semantic content and is used to provide the time contour to the event itself. This newer analytic construction is, nevertheless, equivalent to or substitutable by the older synthetic *ʾaʿaddat al-ḥukūmatu taqrīrahā*, in which the full lexical verb carries its own tense marking while having the same function as the periphrastic counterpart. In these examples, layering – as Hopper and Traugott label it – of similar constructions may show a difference along the pragmatic dimension (degree of formality) and with respect to other discourse factors. In such cases, a cline of grammaticality is construed where clauses with overt relative pronouns, periphrastic passive constructions, to exemplify, are rendered newer than the corresponding synthetic counterparts (verbal noun, participial, and internal passive), since the latter exhibit a higher degree of syntheticity and internal unity.

2.3 *Specialization*

The pool of grammaticalized elements serving a particular function may be reduced diachronically to the extent that a single element takes over and becomes the focus grammatical formative. Among the pool of interrogatives, *mā* 'what', *matā* 'when', *man* 'who', which were later grammaticalized as conditional particles expressing unchanging truth value, for example, *man jāla nāla* 'he who roams will reach something' (as cited in Fischer 2001:227), only *mā* survives in the construction of derived conditionals of the type *kullamā*, 'whenever', *mahmā* 'whatever', *ḥayṯumā* 'wherever' *ṭālamā* 'as long as', *ʾaynamā* 'wherever'. In these forms, *mā* assumes a function relatively more abstract than that of conditional particles, since the latter is far more complex cognitively. The same particle specializes alone in the construction of exclamation expressions (e.g. *mā ʾajmalahā* 'how beautiful she is!'), to the exclusion of all other interrogatives. In like manner, only the preposition *min* 'from' specializes in the creation of the comparatives as in *ʾaṭwalu min* 'taller than'. Within the → negation paradigm in Classical and Modern Standard Arabic, several negation particles exist: *lā, lan, lam*, as well as *mā* and the → defective verb *laysa*. Diachronically, however, *mā* and its variants, including the discontinuous *mā…š* or merely *mā* or *miš/māši,* came to enjoy a privileged status as the most widespread negation particle across all modern Arabic dialects. The particle in major dialects is used to negate personal pronouns functioning as copula (Brustad 2000:296), imperatives, and participles, as well as verbal phrases. From the grammaticalization perspective, *mā* already had distinguished itself from other markers of the negation paradigm even in Classical and Modern Standard Arabic because it occurs not only with verbs marked for past and nonpast (e.g. *mā ʾaḥadun yašukku fī qawlika* 'no one is questioning what you said' (Fassi Fehri 1993:165) but also with nominal sentences (e.g. *mā ʾana qādirun ʿalā hāḏā* 'I am not capable of (doing) this' (Fassi Fehri 1993:165). The diachronic reduction in the membership of the paradigm of negation markers corresponds to what Lehmann (1995:139) labels 'obligatorification', which corresponds closely to Hopper and Traugott's 'specialization'.

2.4 *Persistence*

Although lexical sources, when grammaticalized, lose much of their lexical substance, in many cases their erstwhile meaning persists to a varying degree and may continue to have an influence on the grammatical functions that the gram assumes during the course of its grammaticalization (a phenomenon called 'persistence' by Hopper 1991:28). The postural verb *qāma* 'he stood up' is among a class of verbs traditionally called 'verbs of beginning'. As such, they occur in the perfect and usually embed other verbs in the imperfect indicative, and regardless of their original lexical meanings, when grammaticalized, they mark the beginning of an action: *qāmat al-marʾatu tanūḥu* 'the woman began to wail aloud' (Wright 1982:II, 109). The original sense of this postural verb is not entirely lost and continues to impose restrictions on the range of semantic relations for its grammaticalized form, which does not occur with verbs that semantically contradict such a sense: **qāmat al-marʾatu tanāmu* 'the woman began to sleep' is ungrammatical as a result of the contradiction between *qāmat* 'she stood/got up' and *tanāmu* 'she sleeps', when compared with the grammatical construction *qāmat al-marʾatu min an-nawm* 'the woman woke up, got up, from sleep'.

2.5 *Paradigmaticity*

That paradigms containing similarly functioning forms (inflections, personal and other pronouns, prepositions) are formed and recognized even by medieval Arabic grammarians is a testimony to the presence and awareness of the cumulative effect of the grammaticalization process in the field of Arabic linguistics. Admittance of like grammatical forms into paradigmatic relations is facilitated by grammaticalization, and the formation of paradigms is the result of this process. Lehmann (1995) shows that as members of a given paradigm advance toward the end of their grammaticalization, they become radically reduced in number and thus constrict 'intraparadigmatic variability'. This constitutes the highest degree of paradigmaticity, beyond which the whole paradigm becomes extinct. In Arabic, the size of monosyllabic and proclitic primary prepositions (e.g. *li-* 'for', *bi-* 'by', *ka-* 'as', *ta-* 'by') is considerably smaller than other primary, yet autonomous, ones (e.g. *min* 'from', *bayna* 'between'), which are juxtaposed with their dependents. This is exemplified in the case of the relative pronoun paradigm in Modern Standard Arabic, which is radically reduced to merely the single member *ʾillī* and its allomorphs in the modern spoken Arabic varieties. Size alone is an insufficient variable for measuring the degree of 'paradigmaticity' (one of Lehmann's terms); other criteria, such as the degree of homogeneity in functional and formal properties among the members of the paradigm, are shown to be more reliable (Lehmann 1995:134).

3. MOTIVATIONS FOR GRAMMATICALIZATION

That the process of grammaticalization recurs crosslinguistically in a more or less uniform fashion propelled the search for adequate explanations or motivation for its systematicity. Among the proposed explanations are considerations of pragmatic and semantic factors, as well as use of metaphor and metonymy that effect change and its direction. Central to all these factors are the role of context and communicative strategies used by participants, possibly conducive to grammaticalization. The utility of existing basic lexical items in the evolution of more abstract grammatical or functional categories is seen by some researchers (Heine a.o. 1991) as a 'problem-solving strategy', in which creativity in the use of extralinguistic processes such as metaphor and metonymy in a permitting context would result in grammaticalization. Other suggested factors that take semantic and pragmatic functions into account include increased expressivity, at least in the initial stages of grammaticalization, and later 'routinization', that is, grammaticalized forms that are semantically bleached coexist side by side with their erstwhile lexical forms (Hopper in Hopper and Traugott 1993:68). A more reconciliatory approach between the metaphoric and "strengthening of informativeness" as enabling factors that result in grammaticalization is hypothesized (Traugott and König 1991:190), in which the two are seen as complementary, rather than as substitutes for each other.

4. COUNTEREXAMPLES TO GRAMMATICALIZATION

What sets grammaticalization apart from other types of language change lies in the notion of unidirectionality. Few crosslinguistic cases challenge the notion that grammaticalized forms return to lexicalized form (see Campbell 2001). For instance, unidirectionality presupposes that the direction of evolution in the typical case is from analytic to synthetic constructions (Lehmann 1995:21). Cases that run counter to this notion are found in Arabic, particularly in the genitive construction (→ *ʾiḍāfa*) in modern dialects, which have in addition to the synthetic construction (*maktabti l-gamʿah* 'the university library') an alternative construction (→ analytical genitive), in which a group of similarly functioning particles (e.g. *bitāʿ* 'belonging to', *tabaʿ* 'following', *māl* belonging to') intervene between the head noun and the annexed one (e.g. *il-maktaba bitāʿit l-gamʿah* 'the library belonging to the university' = 'the university library'). Likewise, → passive verbs in Classical and Modern Standard Arabic are derived from active ones via alternation of internal vowels (e.g. *kataba* 'he wrote', *kutiba* 'it was written'). These currently have an alternative analytic structure (e.g. *tamma kitābato* 'it was written'), where the erstwhile lexical verb *tamma* 'he completed', which is in the active voice, is grammaticalized to serve this passivizing function. Prepositions such as *li-* 'for, to', *fī* 'in, at', *ʿinda* 'at', as well as personal pronouns, exhibit verb-like syntactic behavior as in *mafīš muškila* 'there is no problem' (Brustad 2000:152; → pseudo-verb). That said, when comparing the number of such cases of what Lehmann (1995:16) labels 'degrammaticalization' with the grammaticalized forms, it becomes clear that these cases constitute a very small and insignificant argument against grammaticalization (see Haspelmath 2004). This does not detract from the fact that such linguistic manifestations pose a challenge to the canons of grammaticalization and need to be accounted for in further research. Examination of the boundaries of grammaticalization may yield more vexing problems and perhaps would bring about modification to some of its well-established principles.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Ibn Hišām, *Muġnī* = Jamāl ad-Dīn ʾAbū Muḥammad ʿAbdallāh ibn Yūsuf Ibn Hišām, *Muġnī l-labīb*. Ed. Muḥammad al-ʾAmīr. 2 vols. Cairo: Dār al-Kitāb al-Miṣrī, n.d.

Ibn Manẓūr, *Lisān* = ʾAbū l-Faḍl Jamāl ad-Dīn Mukarram ibn Mukarram Ibn Manẓūr, *Lisān al-ʿArab*. Ed. ʿAbdallāh ʿAlī al-Kabīr a.o. 3rd ed. Cairo: Dār al-Maʿārif, n.d.

Mubarrad, *Muqtaḍab* = ʾAbū l-ʿAbbās Muḥammad ibn Yazīd al-Mubarrad, *Kitāb al-muqtaḍab*. Ed. Muḥammad ʿAbd al-Xāliq ʿUḍayma. 4 vols. Cairo: Dār al-Kitāb al-Miṣrī, 1978.

SECONDARY SOURCES

Brockelmann, Carl. 1982. *Grundriß der vergleichenden Grammatik der semitischen Sprachen*, II. Repr. Hildesheim: G. Olms.

Brustad, Kirsten. 2000. *The syntax of spoken Arabic*. Washington, D.C.: Georgetown University Press.

Bybee, Joan, Revere Perkins, and William Pagliuca. 1994. *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago and London: University of Chicago Press.

Campbell, Lyle. 2001. "What's wrong with grammaticalization?". *Grammaticalization: A critical assessment*, ed. Lyle Campbell, 113–162. *Language Sciences* 23 (special issue).

Craig, Colette. 1991. "Ways to go in Rama: A case study in polygrammaticalization". Traugott and Heine (1991b:II, 455–492).

Fassi Fehri, Abdelkader. 1993. *Issues in the structure of Arabic clauses and words*. Dordrecht: Kluwer Academic Publishers.

Fischer, Wolfdietrich. 2001. *A grammar of Classical Arabic*. Trans. Jonathan Rodgers. New Haven and London: Yale University Press.

Givón, Talmy. 1971. "Historical syntax and synchronic morphology: An archaeologist's field trip". *Chicago Linguistic Society* 7.394–415.

——. 1979. *On understanding grammar*. New York and London: Academic Press.

Harris, Alice C. and Lyle Campbell. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.

Haspelmath, Martin. 2004. "On directionality in language change with particular reference to grammaticalization". *Up and down the cline: The nature of grammaticalization*, ed. Olga Fisher, Muriel Nord and Harry Perridon, 17–44. Amsterdam and Philadelphia: J. Benjamins.

Heine, Bernd. 1993. *Auxiliaries: Cognitive forces and grammaticalization*. New York and Oxford: Oxford University Press.

Heine, Bernd, Ulrike Claudi, and Friderike Hünnemeyer. 1991. *Grammaticalization: A conceptual framework*. Chicago and London: University of Chicago Press.

Holes, Clive. 1995. *Modern Arabic: Structures, functions, and varieties*. London and New York: Longman.

Hopkins, Simon. 1984. *Studies in the grammar of early Arabic based upon papyri datable to before 300 A.H./912 A.D.* London and New York: Oxford University Press.
Hopper, Paul. 1991. "On some principles of grammaticalization". Traugott and Heine (1991b:I, 17–35).
—— and Elizabeth Closs Traugott. 1993. *Grammaticalization.* Cambridge: Cambridge University Press. (Repr., 2000.)
Kuryłowicz, Jerzy. 1965. "The evolution of grammatical categories." *Diogenes* 51:55–71.
Kuteva, Tania. 2001. *Auxiliation: An enquiry into the nature of grammaticalization.* New York and Oxford: Oxford University Press.
Lehmann, Christian. 1995. *Thoughts on grammaticalization.* Munich and Newcastle: Lincom Europa.
Lessau, Donald. 1994. *Dictionary of grammaticalization.* 2 vols. Bochum: Brockmeyer.
Matisoff, James A. 1991. "Areal and universal dimensions of grammatization in Lahu". Traugott and Heine (1991b:II, 383–454).
Meillet, Antoine. 1912. "L'évolution des formes grammaticales". *Scientia (Rivista di Scienza)* 12.384–400. (Repr., Antoine Meillet, *Linguistique historique et linguistique générale*, I, 130–148. Paris: Champion, 1948.)
Ramat, Anna Giacalone and Paul Hopper (eds.). 1998. *The limits of grammaticalization.* Amsterdam and Philadelphia: J. Benjamins.
Svorou, Soteria. 1994. *The grammar of space.* Amsterdam and Philadelphia: J. Benjamins.
Traugott, Elizabeth Closs and Bernd Heine. 1991a. Introduction to Traugott and Heine (1991b:I, 1–14).
—— (eds.). 1991b. *Approaches to grammaticalization.* 2 vols. Amsterdam and Philadelphia: J. Benjamins.
Traugott, Elizabeth Closs and Ekkehard König. 1991. "The semantic-pragmatics of grammaticalization revisited". Traugott and Heine (1991b:I, 189–218).
Wright, William. 1982. *A grammar of the Arabic language.* 3rd ed. 2 vols. Repr. Beirut: Librairie du Liban.

Mohssen Esseesy
(George Washington University)

Greek Loanwords

Despite the very extensive contacts between speakers of Greek and Arabic for over a thousand years between Alexander the Great and the advent of Islam, the paradoxical fact is that the forms of the two languages extant in recoverable documents show very few traces of such contacts. In an area where mutual influence was presumably most profound, that of spoken forms of the languages, no documentation has survived for either the Greek or the Arabic spoken in places where such interpenetration could be expected, as, for example, in Petra and Palmyra in the first three centuries of the Christian era or among the soldiers on the eastern frontier of the Byzantine Empire in the 4th through the 6th centuries. Thus, colloquial syntax and everyday vocabulary and expressions in civilian or military life, where one would expect the greatest contact, have left no traces. Written Arabic is not documented, either; for the duration of the entire millennium from Alexander to Muḥammad, there is no extant Arabic literature. The Nabataeans used Aramaic in their administration, and although the occurrence of Greek words in Nabataean inscriptions would tend to indicate their presence in Nabataean Arabic as well (Jeffery 1938:17 n. 2), this remains a conjecture. The *Qur'ān* is, essentially, the first Arabic book, and it is written in the language of central and southern Ḥijāz in the Arabian Peninsula (Mecca and its environs), which, unlike the northern part, never came under Roman domination. Pre-Islamic poetry, transmitted orally and eventually recorded only after the 1st Islamic century, is of restricted usefulness for our purposes even if the authenticity of its linguistic forms is accepted: its stylized conventions and high language exclude the more colloquial and provincial elements which might betray foreign influence in syntax and style, i.e. in anything other than vocabulary. Apart from these documents for → pre-Islamic Arabic, the only other sources at our disposal are inscriptions and graffiti in various forms of what has been called 'early North Arabic' (Frühnordarabisch) by W. Caskell: Thamūdic, Liḥyānite, Ṣafāitic, and Ḥasāitic. These inscriptions, the product of nomadic tribes in north Arabia, are mostly commemorative, funereal, or apotropaic; they are important for allowing the reconstruction of the grammar of these varieties of early Classical Arabic and provide a wealth of information about Arabic proper names, but beyond that, their usefulness is very limited (Müller 1982a). Briefly put, we do not have much evidence for pre-Islamic Arabic, and the little that we do have does not come from those Arabs who had the most extensive contacts with Greeks (Gutas 2001:629–631).

The documents at our disposal are accordingly totally inadequate for a study of the history of contacts between the two languages until the rise of Islam. A comparison of the vocabularies of their extant forms, because of the disparities of the levels of the two languages that are being compared, yields few results and is not representative of the extensive contacts between Greek and Arabic speakers throughout the millennium. A major characteristic of Greek loanwords in pre-Islamic Arabic in the *Qurʾān*, in the earliest poetry, and in whatever prose from the first two Islamic centuries can be confidently assumed to reflect pre-Islamic usage is that they were not borrowed directly from Greek but rather through the intermediacy of Aramaic or Persian (references in Endress 1992:14, n. 88). In other words, they are not the direct result of the contact between Greek and Arabic speakers but rather the result of the Hellenization of the Near East after Alexander and the eventual permeation of such culturally significant terms into the Arabic represented by our earliest sources. For example, *dirham* 'monetary unit' in Arabic is a singular back-formation from *darāhim* in Pehlevi (Middle Persian), taken over into Arabic as plural from the Greek δραχμή (i.e. *dirham*, sg. Arabic < *darāhim*, pl. Arabic < *darāhim*, sg. Pehlevi < δραχμή). Similarly, Arabic *ʾistār*, in the meaning of 'four', comes from στατήρ over Syriac *estērā*; and even Arabic *ʾinjīl* ('gospel') for εὐαγγέλιον is transmitted through Ethiopic *wangēl*. As these examples indicate, in the areas of trade and religion, Greek loanwords in the Arabic of those Arabs who had not been in direct contact with the Greeks come through the intermediacy of peoples who had (Gutas 2001:632).

Preliminary studies for Greek loanwords in this earliest recoverable form of Arabic exist only for the *Qurʾān* and Ibn Isḥāq's (d. 150/767) biography of the Prophet, where the yield is relatively meager. Of all the foreign words in the *Qurʾān*, those which with relative certainty can be traced to a Greek origin come to about only 17, a very small number in a text of 340 large pages in Flügel's edition. Since these borrowings are almost without exception early, i.e., they had become part of the Arabic language of the Ḥijāz long before Muḥammad insofar as they are mostly attested also in pre-Islamic poetry, it is worth listing them here. Other than *dirham* and *ʾinjīl* mentioned above, there are *ʾiblīs* 'devil' < διάβολος; *burj* 'tower' < πύργος; *zawj* 'one of a pair' < ζεῦγος; *sīmāʾ* 'sign' < σημεῖον – σῆμα*;* *fulk* 'ship' < ἐφόλκιον; *qirṭās* 'paper' < χάρτης; *quraysˇ qarš* 'shark' < καρχαρίας; *miqlād* 'key' < κλείς; *qalam* 'pen' < κάλαμος; *qamīṣ* 'shirt' < καμίσιον; *kūb* 'cup' < κοῦπα < κύμβη; *yāqūt* 'gem, sapphire' < ὑάκινθος; *yaqīn* 'icon' < εἰκών. Among these, only the nautical term *fulk* and possibly the religious one *ʾiblīs* may be considered as having come directly from the Greek; the rest entered Arabic through the mediation of other languages, primarily Aramaic/Syriac and Pehlevi. Conversely, Greek itself acted as the intermediary for the transmission of a few → Latin loanwords into Arabic: *balad* 'territory; dwelling' < παλάτιον < *palatium*; *ṣirāṭ* 'road' < στράτα < *strata*; *qisṭās* 'scales' < ξέστης < *sextarius*; *qaṣr* 'castle' < καστρον < *castrum*; *qinṭār* 'weight measure' < κεντηνάριον < *centenarium* (Jeffery 1938: s.vv.). Similar, proportionately, is the presence of Greek words in the Prophet's biography: in a text covering 1,000 pages, only 29 words are of Greek origin (Hebbo 1984). It is again obvious from the nature of the Greek loanwords just listed from the *Qurʾān* that this borrowed vocabulary is broadly cultural and not specific, reflecting the dominance of Graeco-Roman institutions and activities in the Near East after Alexander (Gutas 2001:632–633).

After the advent of Islam, and throughout the Umayyad period (661–750), Arabic was in constant contact with Greek through the incorporation of millions of Greek speakers in Syro-Palestine and Egypt within the borders of the Islamic empire. Although the numerous Greek loanwords that must have entered spoken Arabic do not reveal themselves in the documents of the high language that have survived from the earliest period of Islam, an idea of their nature and extent can be gleaned from the administrative and commercial Graeco-Latin terms that entered spoken Egyptian Arabic and are preserved in papyri documents (Schall 1982:149). Most important, however, this contact was official in the chanceries of the Umayyad state in governmental centers, including Damascus, where the administrative personnel, kept over from the Byzantine period, used Greek as the official language until the Arabization reforms of ʿAbd al-Malik (r. 685–705; Gutas 1998:17–18). The reforms apparently worked very well, for high Arabic absorbed few Greek administrative loanwords, notably the name for land tax, *xarāj* < χορηγία, apparently also through Syriac mediation (EI^2 s.v. k̲h̲arād̲j̲).

It was in the subsequent period, during the early Abbasids (750–1000), under whom an extensive Graeco-Arabic translation movement was instituted (Gutas 1998), that numerous Greek words appeared in Arabic scientific and philosophical literature. However, Arabic speakers of the time reacted to this influx of foreign words and concepts in ways that, as in pre-Islamic times, resulted in the absorption of few Greek loanwords in the high language. The loanwords that were accepted in the sciences and philosophy were, again, for the most part mediated primarily through Syriac but also through Persian. This is clearly due to the fact that the vast majority of the Graeco-Arabic translators were native Syriac speakers, while the Persians, who were responsible for the Perso-Arabic translations, may have been transferring into Arabic Greek words that could have passed from the Greek into Middle Persian. Such loanwords became naturalized in Arabic through morphological accommodation to Arabic patterns of nominal, adjectival, and verbal formation. Thus, from the Syriac, we have such words as *jins* 'genus' (pl *ʾajnās*) < Syriac *gensā* < γένος; *hayūlā* 'matter' (adj. *hayūlānī*) < Syriac *hywlā* (for υ > *yw* see Daiber 1980:44–45; Endress 1992:14, n. 89) < ὕλη, eventually replaced by *mādda*; *ʾusṭuquss* 'element' (pl. *-āt*) Syriac *esṭūksā* < στοιχεῖον, replaced by *ʾunṣur*; *faylasūf* 'philosopher' (pl. *falāsifa*, noun *falsafa*, verb *tafalsafa*, adj. *falsafī* < Syriac *pīlōsōpā* < φιλόσοφος, replaced by *ḥakīm* 'philosopher, physician, sage'; *ʾistāḏiyā* 'stadium' < Syriac *estadyā* < στάδιον (cf. modern Arabic *ʾistād*, this time from French *stade*); *ʾaqrābāḏīn* 'dispensatory' < Syriac *grāpā/īdīn* < γραφίδιον; *kīmiyāʾ* 'alchemy' (adj. *-iyāʾī*, *-iyāwī*, *-āwī*) < Syriac *kīmīyā* < χημεία. Through the mediation of Persian we have such words as *qawlanj* 'colic' < κωλική and *ʾiyāraj* 'divine [remedy]' < ἱερά (Endress 1992:14–15).

In the early phases of the translation movement, numerous Greek words without obvious Arabic counterparts were simply transliterated. This practice, however, proved unsatisfactory, primarily because of the inherent unwieldiness of such Greek words and the impossibility of their morphological and syntactic adaptation in an Arabic linguistic environment, and they were eventually almost all replaced by native Arabic words (Heinrichs 1978:260–261, n. 26). For example, 'syllogism' (συλλογισμός), a word which proved compatible with patterns in Latin and other Indo-European languages of Europe, was at first simply transliterated as *sīlūjismūs* (Vagelpohl 2002:256), later replaced by Arabic *qiyās*; *f/banṭasiyā* 'imagination' < Syriac *panṭasiyā* < φαντασία, replaced by *tawahhum*, *taḫayyul* (although in this case it is interesting to see that the Greek word was retained by Ibn Sīnā but given a different technical sense, Heinrichs 1978:260 n. 25); *ʾanṭalāšiyā* 'entelechy' < ἐντελέχεια replaced by *kamāl*, *tamām*; *qatafasīs* 'affirmation' < κατάφασις, replaced by *ʾījāb*; etc. (Endress 1992:16). Of the words transliterated in the learned literature there were eventually tolerated only a few, primarily those that referred to specific disciplines and subjects of study, e.g. *ʾariṯmaṭīqī* 'arithmetic' < ἀριθμητική, *Qāṭāġuriyās*, 'Categories' < Κατηγορίαι of Aristotle, *Sūfisṭīqā* 'Sophistics' (nouns *safsaṭa*, *sufisṭāʾī*) < Σοφιστική (scil. ἐπιστήμη, for Σοφιστικοὶ ἔλεγχοι of Aristotle), etc., although even these were given Arabic counterparts, *ʿilm al-ḥisāb*, *Maqūlāt*, and *tamwīh* 'misrepresentation' and *muġālaṭa* 'inducing error', respectively (Endress 1992:16–17). In the end, it appears that the Greek loanwords of wider currency that were accepted into Arabic vocabulary, both those that entered through the mediation of Syriac and Persian, given above, and those that entered through the Graeco-Arabic transliterations in the learned tradition, show two characteristics: they entered through the mediation of Syriac (and, to a lesser extent, Persian) pronunciation, which apparently made them phonetically more acceptable, and almost all eventually acquired Arabic equivalents (Heinrichs 1978:261 n. 26; Endress 1992:20), so that, in principle rendered superfluous, their use could be regarded as a sign of learned jargonism.

By contrast, loanwords that referred to highly technical terms in most fields, but especially in botany and pharmacology (Dietrich 1988:II, 708–714, 728–746), were accepted as such and have formed part of the Arabic vocabulary. (For example, of the words listed above, *ʾaqrābāḏīn, kīmiyāʾ*, and *qawlanj*, and the word in wide use to this day in the entire Near East, *tiryāq* 'antidote, panacea' < Syriac *teryāqī* < θηριακή [ἀντίδοτος], which, through Persian *teryāq/k* 'antidote, wine, opium' and the adjective *teryāq/kī* 'addict, opium eater' and Turkish *tiryāq/k* and *tiryāq/kī* (with the same meanings as in Persian) has reentered Modern Greek as θεριακλής 'addicted, aficionado [of tobacco and coffee]', with the addition of -λης in an apparently improper back-formation for the Turkish associative suffix *-li*. In at least one case, that of *ʾasṭurlāb* 'astrolabe' < ἀστρολάβιον, the acceptance of the loanword was accompanied by attempts to find an 'Arabic' etymology for it (Endress 1992:17, n. 128, referring to D.A. King). In general, however, and technical terminology apart, written Arabic proved to be not very hospitable to Greek loanwords. Instead, the need to incorporate new words into Arabic to meet the demands of the translated literature and the rapidly expanding scientific and philosophical endeavors during the early Abbasids was met in at least four ways: approximate translations of Greek terms (occasionally accompanying the transliterated word), etymological translations (calques), assignment of new meanings to existing Arabic words, and derivation of new words (Endress 1992:17–23; Schall 1982:150–151).

The phonology of the loanwords and transliterated words oscillated between the two major sources from which these words came into Arabic, through Syriac and directly from Greek. The Syriac phonology is partly due to its impact on the language of the translators, as noted, and partly to their familiarity with the transcription of numerous Greek words in Syriac texts (den Heijer 1991:104). The Greek phonology is directly derived from the spoken Greek of early Islamic times, which appears to be very close to that of Modern Greek: the *spiritus asper* is almost universally suppressed, iotacisms – insofar as vowels can be determined from the Arabic script – would seem to abound, and the consonants clearly have the values of Modern Greek: *gamma* (γ) is transcribed mostly with a *ġayn*, *delta* (δ) with a *ḏāl*, and *chi* (χ) oscillates between the Syriac transcription of *kōf*, which appears in Arabic as *kāf*, and the actual sound of the letter (especially before α, ο, ω, ου), transcribed with a *xāʾ*. The letter *sigma* (σ, ς), normally transcribed with a *sīn*, in one instance gives a clue to the transcription of the letters on the basis of actual pronunciation: in the translation of Aristotle's *Generation of Animals* (763b 1), the name of the island Λέσβος is transliterated as *Lāzbūṣ*, indicating both the actual /z/ phoneme of the letter before *bēta* (presumably, at that time, pronounced /v/) and the emphasized phoneme /s/ characteristic of final *sigma* (Vagelpohl 2002:143–153).

BIBLIOGRAPHICAL REFERENCES

Bowersock, Glenn. 1983. *Roman Arabia*. Cambridge, Mass.: Harvard University Press.

Bubenik, Vit. 1989. *Hellenistic and Roman Greece as a sociolinguistic area*. Amsterdam and Philadelphia: J. Benjamins. [Pp. 264–276: "Hellenistic Koine in contact with Phoenician, Aramaic and Arabic".]

Daiber, Hans. 1980. *Aetius Arabus: Die Vorsokratiker in arabischer Überlieferung*. Wiesbaden: F. Steiner.

Dietrich, Albert. 1988. *Dioscurides Triumphans*. Göttingen: V & R.

Endress, Gerhard. 1992. "Die wissenschaftliche Literatur". Fischer (1992:3–152).

Fischer, Wolfdietrich (ed.). 1982. *Grundriss der arabischen Philologie*, I. Wiesbaden: L. Reichert.

—— (ed.). 1992. *Grundriss der arabischen Philologie*, III. *Supplement*. Wiesbaden: L. Reichert.

Gutas, Dimitri. 1998. *Greek thought, Arabic culture*. London and New York: Routledge.

—— 2001. "Ελληνική και αραβική – πρώϊμες επαφές" [Greek and Arabic: Early contacts], Ιστορία της ελληνικής γλώσας: Από τις αρχές ως της ύστερη αρχαιότητα [*History of the Greek language: From the beginning to late antiquity*], ed. A.-Ph. Christidis, 629–634, 643–644. Salonica: Center for the Greek Language. [English ed. forthcoming, Cambridge: Cambridge University Press.]

Hebbo, A.I. 1984. *Die Fremdwörter in der arabischen Prophetenbiographie des Ibn Hischam (gest. 218/834)*. Frankfurt a.M. and New York: P. Lang.

Heijer, Johannes den. 1991. "Syriacisms in the Arabic version of Aristotle's *Historia Animalium*". *Aram* 3.97–114.

Heinrichs, Wolfhart. 1978. "Die antike Verknüpfung von phantasia und Dichtung bei den Arabern". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 128.252–298.

Jeffery, A. 1938. *The foreign vocabulary of the* Qurʾān. Baroda: Oriental Institute.

Müller, Walter W. 1982a. "Das Frühnordarabische". Fischer (1982:17–29).

——. 1982b. "Das Altarabische der Inschriften aus vorislamischer Zeit". Fischer (1982:30–36).

Schall, Anton. 1982. "Geschichte des arabischen Wortschatzes, Lehn- und Fremdwörter im Klassischen Arabisch". Fischer (1982:142–153).

Schmitt, Rüdiger. 1980. "Die Ostgrenze von Armenien über Mesopotamien, Syrien bis Arabien". *Die Sprachen im Römischen Reich der Kaiserzeit* ed. G. Neumann and J. Untermann, 187–214. Cologne and Bonn: Rheinland Verlag.

——. 1983. "Sprachverhältisse in den östlichen Provinzen". *Aufstieg und Niedergang der Römischen Welt*, ed W. Haase, II, 29/2:554–586. Berlin and New York: W. de Gruyter.

Vagelpohl, Uwe. 2002. *The early Greek-Arabic translation phenomenon: The case of Aristotle's* Rhetoric. Ph.D. diss., Cambridge University.

Werner, J. 1992. "Zur Fremdsprachenproblematik in der griechisch-römischen. Antike". *Zum Umgang mit fremden Sprachen in der griechisch-römischen Antike* ed. C.W. Müller, K. Sier, and J. Werner, 1–20. Stuttgart: F. Steiner.

Dimitri Gutas (Yale University)

Greetings

1. The definition of greetings

The term 'greetings' refers to any verbal behavior that a speaker engages in upon recognizing another, or one that has the function of recognition of an encounter with a person as socially acceptable (Firth 1972:1). Greetings are also defined as the set of linguistic and/or nonlinguistic devices used for the initial management of encounters (Yusuf a.o. 1976), that is of paramount significance in everyday interaction. They are aspects of politeness routines (Ervin-Tripp 1964:195) that are tied to conversational exchange; hence, their manipulation is instrumental to power relationships and solidarity. They are also considered markers of social identity (El-Zeini 1985). While their presence usually guarantees continuation of social interaction, their absence, or the failure to respond to them, is a violation that could be interpreted as rudeness or breakdown in the social relationship between parties or individuals (Chaika 1989:44).

In most cultures, greetings are usually routinely performed formulaic expressions, "culturally valued patterns of speaking preserved for their own sake, and in which few changes can be made, with the exception of features of pitch and loudness" (Caton 1986:296). A formulaic linguistic unit is also "one in which the referential value of the speech signal is either minimized or non-existent" (Yūsuf a.o. 1976:812). Routines do not have a manifested topic, hence they are considered contentless speech that can be replaced by gestures (Ervin-Tripp 1964:195).

The systematic study of greetings started toward the beginning of the 1970s (Kendon and Ferber 1973) and may be attributed to the growing interest in the ethnography of speaking and conversational analysis. They are usually studied within the framework of linguistics and anthropology. In conversational interaction they are termed 'adjacency pairs' (Sacks a.o. 1974), "utterances produced by two successive speakers such that the second utterance is identified as related to the first as an expected follow up". They are also called a 'dialogic couplet' (Goffman 1981), since a greeting is a two-part structure. Yet another term used in reference to greetings is 'exchange' (Mauss 1967:3), in which a formula uttered by the speaker triggers a standard reply uttered by the addressee. In sociolinguistics, they are also referred to by the term 'summons' (Schegloff 1968); they are repeated until they receive a response and are then followed by the formulaic routines as a way of carrying on with social interaction. In the field of discourse analysis, greetings are classified as expressive speech acts (Searle 1979), that is, acts that express emotions and feelings.

In anthropology, greetings are considered aspects of phatic communion (Malinowski 1927) that set the tone for communication but have no cognitive content. They also fit into 'rituals', arbitrary communicative procedures that regulate social situations (Firth 1972). Ritual is also viewed as that aspect of customary behavior that makes statements about the hierarchical relations between people. The form and function of greetings and their interpretation vary across cultures, a variation that has prompted a number of researchers to study them in different languages.

2. Greetings in the Arab world

Greetings in the Arab world are related to various layers of cultural meaning. Reducing greetings to a matter of etiquette or politeness without pointing out their religious significance, as noted in the literature on the subject (e.g. Caton 1986), would yield an impoverished analysis. This religious significance is immediately noticed in the triradical root of the nomenclature/word *as-salām* in Arabic, *s-l-m*. This root denotes 'being sound and void of defects, weaknesses, or any sort of imperfection', e.g. *bi-qalbin salām* 'with a clean heart' (Q. 26/89), i.e. 'with a heart void of hypocrisy and infidelity' (Zabīdī, *Tāj* s.v.).

The religious significance of greetings may also be readily grasped in the Qurʾānic verse in which greetings should follow the principle of 'better or same response': "When you are greeted with a greeting, greet in return with what is better than it, or [at least] return it equally" (Q. 4/86). Hence, performing greetings is in fact a fulfillment of a religious duty in Islam. Moreover, *as-salāmu ʿalaykum* is the greeting that should be used among Muslims (El-Zeini 1985:71–72; Caton 1986:294). Furthermore, the structure of a greeting, i.e. who initiates it and how to respond to it, is rule-governed, as is clearly exemplified by the following Prophetic *ḥadīṯ*: "The rider should first greet the pedestrian, and the pedestrian the one who is seated, and a small group should greet a larger group with *as-salāmu ʿalaykum*" (Muslim, *Ṣaḥīḥ* Book 26, Chap. 1, no. 5374). Another Prophetic *ḥadīṯ* that shows the importance of greetings in Islam says: "Give currency to [the practice of paying salutation to one another by saying] *as-salāmu ʿalaykum*" (*ʾafšū s-salāma baynakum*; Muslim, *Ṣaḥīḥ*, Book 1, no. 0096). Such *ḥadīṯ*s attest to the fact that greetings in Islam are deeply rooted in its religious credo. The meaning of the Muslim greeting itself also carries religious connotations. It is an invocation for peace to spread onto the addressee and it is reversed in the response to invoke peace for the initiator by saying 'and upon you be peace!'. Such connotations are further intensified by adding words like *wa-raḥmatu llāhi wa-barakātuh* 'and God's mercy and His blessings'. In this way, the respondents heap more blessings on the addressee than they receive, which attests to the better or same response stated in the *Qurʾān* as indicated above, which reinforces the religious value of Islamic greetings.

The religious input in Islamic greetings is also reflected in the greeting formulas that Muslims use. For instance, Islamic greetings are often mentioned in association with the Name of God, as in responses to *kēfak*, *ʾizzayyak* 'how are you?', such as *naškur Allāh*/*al-ḥamdu li-llāh* 'thanks, praise be to God', respectively. As such, they differ from other greetings used in the Arab world that have secular values of beauty or flowers, such as those used in Egypt and in Syria, e.g. *ṣabāḥ il-ward*/*ṣabāḥ il-full* '[may you have] a morning of roses/jasmine'.

The greeting *as-salāmu ʿalaykum* is also the one used by pious Muslims in particular (Caton 1986:294), as prescribed in the Qurʾānic verse (Q. 6/54). Moreover, the greeting *ṣabāḥ in-nūr*, as a response to *ṣabāḥ il-xēr*, derives religious significance through the word 'light', which is also one of God's attributes. In addition there is a *sūra* in the *Qurʾān* called *an-Nūr* (Q. 18/35). Thus, to fully understand the greeting behavior in the Arab world, it is important that the study be contextualized within such a religious framework.

The religious background of greetings in the Arab world manifests itself through actual greetings, as in the case of *as-salamu ʿalaykum*. The use of a response better than the initiation, irrespective of the nature (religious or nonreligious) of the greeting itself, reflects an Islamic influence. Thus, it is not only the greeting per se that is religious, but also the way it is structured and manipulated.

Ferguson (1967) points out that Arabic has a number of 'root-echo responses', e.g. *Allāh yibārik fīk* 'may God bless you', which may be modified by "the addition, deletion, permutation or the ringing of paradigmatic changes of some kind" (Ferguson 1976:143). For example, *as-salāmu ʿalaykum* 'peace be on you' receives the response: *wa ʿalaykumu s-salām* 'and upon you be peace'. Another multipurpose response is *Allāh yiḥfaḍak* 'may God protect you', which fits in case no specific response is expected.

The use of greetings in Egypt varies according to education, sex, religion, and degree of familiarity between participants. Greetings also allow for a degree of creativity reflected in the use of metaphorical and flowery expressions. Members of the lower socioeconomic strata commonly use ornate greetings, such as *ṣabāḥ il-full* 'morning of jasmine', *ṣabāḥ il-ʾišṭa* 'morning of cream', *ṣabāḥ il-ward* 'morning of roses'. These same greetings are used among familiars, such as friends and colleagues. They would be inappropriate in formal situations as, for example, from an employee to a boss; such greetings as *ṣabāḥ il-xēr* and *as-salāmu ʿalaykum* are seen as being more suitable for such formal relationships.

In Egyptian society, a speaker's sex influences the use of greetings. A highly educated male speaker is not expected to greet his female colleague with *ṣabāḥ il-full* or *ṣabāḥ il-ʾišṭa*, unless the two have a very close relationship. Otherwise, the woman may feel embarrassed. Christians in Egypt often use the Muslim greeting of *salām*. One possible explanation attributes this usage to certain behaviors of minority groups: minority groups are liable to be influenced by the majority. According to Firth (1972:33–34), "As a rule, adoption of the new forms of greeting has been credited with the prestige of association with a larger universe, or alternatively, has been stimulated by the wish to develop patterns of weight, such as Moslem parallels to Christian forms". Another possible explanation attributes this usage to the word *salām* itself. The *salām* greeting, according to this interpretation, is similar in its reference to 'peace' to the Biblical greeting "And into whatsoever house ye shall enter, first say, peace be to this house" (John 14:27).

Muslims, in general, vary greeting responses in accordance with the Qurʾānic verse stated above, that is, they should use a better greeting than the initiation, which implies that variation is recommended and preferred to using the same response. They use *ṣabāḥ il-full/il-ward/il-ʾišṭa/in-nūr* 'morning of jasmine/roses/cream/light', respectively, all associated with the color white, which is a sign of purity, good omens, and well-wishing. Christians also vary their patterns of greeting responses as well, in spite of the fact that they are not religiously ordered to do so. But, as Ferguson (1976) notes, such a phenomenon is natural because gradually these greetings are integrated within the culture, and people become unaware of the religious origin of the variation in greeting responses.

Related to variation in greeting responses is the creative use of greetings typical of Egyptians. For instance, the greeting *ṣabāḥ is-sukkar* 'morning of sugar' has come into use on a television program in Egypt. Another interesting creative greeting is *ṣabāḥ in-neskafē* 'morning of Nescafé', also used on a TV promotion. Creativity here refers to flexibility of insertion, addition, or replacement in the usual fixed greetings, depending on one's individual choice and context, e.g. *ṣabāḥ ʿasal* 'morning of honey'. This is not usually the case in Western greetings, for instance, where greetings consist of fixed adjacency pairs. In Arab countries, however, they are not fixed but more varied and likely to change according to variables of context, personal mode, and other factors.

The greetings *marḥaba/marḥabtēn/mīt marḥaba* 'hello/two hellos/a hundred hellos' and the plural formula *mīt ṣabāḥa* 'hundred mornings', are typically Levantine. Greetings for the morning in the Levant are used between waking up and midday. Other more flowery responses include *ṣabāḥ il-ward* 'morning of roses' and *ṣabāḥ il-full* 'morning of jasmine', which are common family greetings.

However, variation of greetings in Syria, according to Ferguson (1967), is not predictable in any universal sense; rather, what is universal is the correlation between the structure of the greeting formula and the sociospatial dimension. The one who initiates the greeting cannot anticipate the response, because of the creative nature of greeting usage in Syria (as well as in some other Arab countries). Nonetheless, one can predict certain greeting formulas to be associated with certain social classes and certain contexts. For instance, an informal greeting like *ṣabāḥ il-full* is likely to occur in a café between friends of the same sex, but not in a lecture from a professor to students. Furthermore, such variation is not restricted to the Syrian society; it applies to most of the other Arab countries, such as Egypt.

Caton's (1986:290–303) study of greetings in Yemen focuses on a Hijrah (Caton 1986:292) village whose inhabitants are reputed to be descendants of the Prophet Muḥammad. The greetings used in this village are deeply connected with the Islamic credo (Caton 1986:294). Caton also mentions *ʾahlan wa-sahlan* lit. 'you have come to your people and it is easy to welcome you' or simply 'feel at home'. Although this greeting is not the typical *salām* greeting, it is a pan-Arab greeting.

The intensity of greetings in Yemen is indexed in various ways. One such strategy is the metaphoric use of number categories, such as the dual and plural forms of *marḥabtēn* (Caton 1986:298). A further intensification strategy of greetings in Yemen is to add more words by performing two acts in the second turn: reply to the first exchange and initiate another one, e.g. initiation *ʾahlan wa-sahlan;* response *ʾahlan bīk, wa-ḥayya llāh man ḥayyāk* 'hello, hello to you; may God greet him who greets you'. Other intensification strategies in Yemen mentioned by Caton (1986:299) include the use of an intensifier phrase, such as the phrase *bi-l-xēr wa-l-ʿāfiya*, and extending the blessing beyond the immediate party of the greeting to encompass the whole Muslim community. Furthermore, some greetings can be used only by an individual in saluting a group, especially the *salām taḥiyya* (Caton 1986:303).

In Saudi Arabia, according to Hassanayn (1994), the address term *ʾaxx* 'brother' or *ʿamm* 'uncle' is used with the greeting when addressing a male older than the greeter, and *ʾuxt* 'sister' and *ʿamma* 'aunt' for the female counterpart (→ terms of address). Moreover, some gestures have special significance for Saudi Arabian greetings. For example, handshakes show intimacy and are used after the verbal greeting in male/male interaction, whereas in male/female interaction, the handshake takes place only among relations. In addition, kissing the forehead, nose, and right hand of the person who is being greeted is a sign of respect and loyalty. Gestures also replace the verbal greeting when there are constraints on the utterance of greetings, e.g. in extreme silence as in funerals.

Moroccans greet friends and acquaintances with *ça va?* or *lā bas?*, stop in order to shake hands, then ask about the other person's family, children, and health. The exchange extends to include a flow of questions without actually waiting for a response, which reflects their phatic nature, and it ends when one of the parties says *bārak Allāh fīk* (Lowless 2004).

3. GREETINGS IN GENERAL

On the whole, Arabic displays more elaborate and varied greeting structures than other languages, especially in the 'how are you?' phase, which includes an exchange of phatic questions about one's family members. Despite the specific aspects of greeting behavior in the Arab world, there are other general features that cannot be ignored. These include some typical aspects of greetings, strategies of intensification, social functions, nonverbal behavior, and channels of exchanging greetings. Despite the variation of greetings across cultures, there is, to a great extent, some common ground for greetings everywhere, for instance, the factors affecting their intensity, strategies of greeting intensification, and other shared features.

Factors that influence the intensity of greetings include such salient features as length of time elapsed since previous encounters, distance between communicators, number of individuals in the relevant groups, relative social status of the communicators, and social power and solidarity between them.

There are several strategies for intensifying greetings. One of these is the metaphoric use of number categories in the noun (singular/dual/plural), e.g. *ʾahlēn*, the dual form of *ʾahl* in many Arab countries. Another strategy involves the use of an intensifier phrase in addition to the blessing mentioned in the initiation. A third is the use of a more metaphoric lexical item than the one used in the initiation. Still another common strategy to intensify greetings is extending the greeting beyond the immediate party being greeted to include the entire Muslim community. For example, *ʿīd mbārak* 'blessed holiday' receives the response *ʿalā l-jamīʿ* 'to all'. Repetition of the greeting formula itself, e.g. *ʾahlan ʾahlan* 'welcome welcome', is also an effective strategy for intensifying the greeting.

Greetings perform several social functions. They may be a prelude for social interaction, and for opening a sequence of communicative acts. They also set the frame for interaction (Firth 1972). However, one of the significant functions of greetings is that of indexing the hierarchical relationship of the users and reflecting their status through the variation of the verbal behavior of greetings (El-Zeini 1985). Greetings also perform a pious act, since their structure is deeply implicated in Islamic piety (Caton 1986), and they reflect social values and politeness norms. Other functions of greetings include expressing solidarity through the use of address terms, titles versus first names, and honorific *tu/vous* distinctions common in French and other European languages, and through showing friendliness, respect, deference, or decency toward the greeted party (Chaika 1989).

Nonverbal behavior may accompany greetings, replace them, or precede them. However, this depends on spatiotemporal aspects or context of use as well as on cultural background. Nonverbal greetings may include both facial expressions and body language. A wave, a smile, a nod, a gesture, a bow, and even a clap may replace the verbal greeting. The signs and gestures may communicate in much the same way a word does. However, both verbal and nonverbal greetings may co-occur or overlap. For instance, on recognizing someone at a distance, a smile is the first nonverbal greeting to acknowledge the other. Yet, coming closer, both parties exchange eye contact and extend a hand for shaking. At this point, the verbal greeting is uttered and the nonverbal behavior may last a little bit longer or end, depending on the degree of familiarity between the two parties involved. The closer the participants, the more nonverbal behavior of greeting is observed, namely through hugging, kissing, and touching, although this varies across cultures. Nonetheless, if in a hurry, the two parties may exchange the nonverbal greeting or the distant salutation, as it is sometimes called. It is worth noting here that all nonverbal behavior in greetings is culture-bound. For instance, the number of kisses varies across cultures, as does the part of the face being kissed. Egyptians, for example, usually exchange two kisses, one on each cheek, whereas Saudis may kiss the nose, the forehead, or the cheek depending on the level of familiarity between participants. What is more, the ritual of greeting should be observed strictly, since it is rule-governed. Any change in the nonverbal behavior of greeting would result in embarrassment, and would also label the greeter as a stranger who does not know the routine of that specific culture.

Greetings may be expressed through different channels: oral in the presence of another party, as in spoken, everyday face-to-face interaction, or in one another's absence, as in radio programs and phone calls. Despite their short duration, for example, phone calls may not take place without the greeting exchange, otherwise, the caller is considered extremely rude. Yet another way is the written channel, as in exchanging greetings via greeting cards or e-greetings, which emphasizes the social role that greetings play in various cultures. Greeting cards help to maintain social solidarity, even though the two parties involved are usually absent at the time of the greeting exchange.

4. Conclusions

A number of important conclusions can be drawn regarding the use and manipulation of greetings in the Arab world. Knowledge of the speakers' religious background is crucial to understanding some aspects of the greeting behavior in the Arab world. The greeting *as-salāmu ʿalaykum* is regarded as the pan-Arab/Islamic greeting and is recommended for use by conservative Muslims. The definition of greetings itself is fuzzy in the available literature; it is sometimes used to refer to other fixed routine expressions such as compliments, as in the work of Hassanayn (1994) and Caton (1986). The available literature puts more emphasis on the verbal aspects of greetings. Verbal greetings are accompanied by nonverbal behaviors such as hugging, kissing, and handshaking which vary significantly across cultures. These, however, have not received adequate attention in available literature.

Finally, intensification strategies in the Arab world include quantification, better responses to greetings, and use of flowery language. Overall, further empirical research is needed in the field of greetings, particularly in the Gulf area and in North Africa.

Bibliographical references

Primary sources

Khān, M.M. and M.T. Hilāli. 1996. *Interpretation of the meanings of the Noble Qur'ān in the English language*. Riyadh: Maktaba Dār as-Salām.

Muslim, *Saḥīḥ* = Siddīqī, ʿAbdul Ḥamīd. 1993. *Saḥīḥ Muslim*. Available at http://www.muhaddith.org.

Zabīdī, *Tāj* = as-Sayyid ʾAbū l-Fayḍ Muḥammad Murtaḍā az-Zabīdi, *Tāj al-ʿarūs*. Cairo: al-Maṭbaʿa al-Xayriyya, 1306/1888.

Secondary sources

Caton, Steven. 1986. "*Salām taḥiyya*: Greetings from the highlands of Yemen". *American Ethnologist* 13.290–308.

Chaika, Elaine. 1989. *Language: The social mirror*. New York: Newbury House.

El-Zeini, Nagwa Ṭāhā. 1985. *Greetings and farewells in Cairene Egyptian society*. M.A. thesis, American University in Cairo. [Published by Cairo: Rare Collection Publications Center, American Uni-versity in Cairo Press.]

Ervin-Tripp, Susan M. 1964. "An analysis of the interaction of language, topic and listener". *American Anthropologist* 66.86–102. (Repr., *Readings in the sociology of language*, ed. Joshua A. Fishman, 192–211. Paris and The Hague: Mouton, 1968.)

Ferguson, Charles A. 1967. "Root-echo responses in Syrian Arabic politeness formulas". Linguistic studies in memory of Richard Slade Harrell, ed. Don Graham Stuart, 35–45. Washington, D.C.: Georgetown University Press.

——. 1976. "The structure and use of politeness formulas". *Language in Society* 5.137–151.

Firth, Raymond. 1972. "Verbal and bodily rituals of greetings and parting". *The interpretation of ritual: Essays in honor of Ian A. Richards*, ed. Jean La Fontaine, 1–38. London: Tavistock.

Goffman, Ervin. 1981. *Forms of talk*. Philadelphia: University of Pennsylvania Press.

Hassanayn, Khalid S.A. 1994. "Saudi mode of greeting rituals: Their implications for teaching and learning". *International Review of Applied Linguistics* 32:1.68–77. Mecca: Makka Teachers Training College.

Kendon, Adam and Andrew Ferber. 1973. "A description of some human greetings". *Comparative ecology and behavior of primates*, ed. Richard P. Michael and John Hurrell Crook, 591–668. London: Academic Press.

La Fontaine, Jean (ed.). 1972. *The interpretation of ritual: Essays in honor of Ian A. Richards*. London: Tavistock.

Lowless, K. Laura. 2004. *Moroccan culture theories: Meetings and greetings*. http://French. about. com/library/travel/pl-ma-greetings. htm? terms= greetings+people.

Malinowski, Bronislaw. 1927. "The problem of meaning in primitive language". *The meaning of meaning*, ed. C.K. Ogden and Ian A. Richards. London: Routledge.

Mauss, Marcel. 1967. *The gift*. Trans. Ian Cunnison. New York: Norton.

Sacks, Harvey, Emanuel Schegloff, and Gail Jefferson. 1974. "A simplest systematics for the organization of turn-taking for conversation". *Language* 50.696–735.

Schegloff, Emanuel. 1968. "Sequencing in conversational openings". *American Anthropologist* 70. 1075–1095.

Searle, John R. 1979. *Expression and meaning: Studies in the theory of speech acts*. Cambridge: Cambridge University Press.

Yusuf, Ibrahim A.G., Allen D. Grimshaw, and Charles S. Bird. 1976. "Greetings in the desert". *American Ethnologist* 3.797–824.

NAGWA ELZEINY (Helwan University)

Grounding

Grounding is a discourse semantic notion. It pertains to the organization of meaning in terms of a foreground/background structure. This structure is not a binary opposition but rather a gradual scale of meaning distribution and distinction among whole propositions in terms of 'grounding values', based on the assignment of degrees of importance to information. A grounding value is a measure of the relative worth of a textual proposition on the foreground/background gradient. Roughly, a proposition is the semantic equivalent of a clause. On the grounding scale, foreground meaning is high and background meaning is low. Between both grounding values are meanings (viz. propositions) that occupy various positions such as midground.

The organization of meaning as a foreground/background structure is a fundamental property of text or discourse, and a language-independent, universal principle of text organization. The terms 'text' and 'discourse' are used interchangeably. Discourse does not refer only to conversation, and a written text may also be a "monologue discourse" (van Dijk 1977:8). The scalar foreground/background structure makes texts or discourses interesting. It is also consistent with the assumption that human communication does not manifest the same grounding value throughout a text or texts and that events and participants referred to in text do not enjoy the same significance. A text that was a "story in which every character was equally important and every event equally significant can hardly be imagined" (Callow 1974:49). Language users, and in particular readers and writers of texts, "lend more importance to some information than to other information" (Wallace 1982:208). They assign variant grounding values to the semantic meaning or content of text.

Propositions, and events referred to in them, vary in their significance and grounding values, as illustrated in (1).

(1) *ʾAdrusu l-ʾāna fī jāmiʿat laydin. Wa-kuntu qad darastu min qabl fī jāmiʿat ʾamstirdām.*
'I study now at the University of Leiden. Earlier, I studied at the University of Amsterdam.'

Writers or speakers consider the meaning or proposition expressed in the first sentence as being more important and more relevant to them (and to readers or listeners) than the meaning or proposition expressed in the second sentence, hence, they assign it a foreground value or interpretation. The meaning expressed in the second sentence, which reports information about earlier activities, serves a background function. This means that the writer or speaker assigns different grounding values to what is written or said. It also means that the grounding structure is part of the communicative strategy used for the realization of intended meaning.

A burgeoning interest in the phenomenon of grounding has emerged during the last three decades. Grounding has been studied from different perspectives: linguistic, literary, and psycholinguistic (Dry 1992). Several linguistic studies have characterized foreground in terms of events and background in terms of nonevents or states (Labov and Waletzky

1967; Grimes 1975), or in terms of a contrast between "sequenced events" and "nonsequenced states and actions" (Hopper 1983; Fox 1983; Reinhart 1984; Thelin 1984; Flashner 1987; Thompson 1987; Couper-Kuhlen 1989). This characterization is typical for narrative discourse, where narrative foreground consists of the plot or of sequentially organized events, and narrative background consists of descriptive material.

Much of the work done on grounding has also conflated the distribution of the foreground/background structure (at the semantic level) with manifestations of grounding in surface structure (the expression of the foreground/background structure in clauses or sentences). Furthermore, research on grounding was focused primarily on narrative and conversational types of discourse. Other types, which may very well have different characteristics, remained largely unexplored. Inevitably, this has led to a restricted view of grounding and to problems when other types of discourse, such as news, are analyzed. See Khalil (2000) for a short critical survey of the work done on grounding.

In news discourse, as in other types of discourse, the main, thus most important, and recent events are usually expressed in the beginning of the text (in the headline or the lead sentence), and the less important, old, or known, events about details as well as information about the spatio-temporal setting of main events are expressed later (toward the end of the text). The typical grounding structure of this type of discourse is based on a gradual departure from foreground to background meaning.

Generally speaking, the writer's perspective on grounding and its structure may determine the order in which sentences appear in text. This means that the writer may manipulate the order of sentences in order to signal the foreground/background structure and to provide readers with clues as to how they (should) interpret the grounding structure in a given situation. Example (2) illustrates this point.

(2) a. *Waqaʿa hujūmun ʿala l-madīna. Wa-li-hāḏā istaslamat al-ḥāmiya.*
'There has been an onslaught on the city. As a result the garrison has surrendered.'
b. *Istaslamat al-ḥāmiya. Fa-qad waqaʿa hujūmun ʿala l-madīna.*
'The garrison has surrendered. There has been an onslaught on the city.'

In (2a), the sequence of sentences expresses a consequence relation, denoted by *wa-li-hāḏā*. The sequence of sentences in (2b), which is the opposite of (2a), expresses an explicative relation (denoted by the particle *fa-* and the preverbal particle *qad*), providing the cause of the event referred to in the preceding sentence. In foreground/background terms, one may assume that the proposition expressed in the first sentence of (2a) and (2b) – about the cause and the consequence of the event – has more importance, hence, it is assigned foreground interpretation and, as a result, is more prominently expressed in the sequence. (Depending on the type of text, other interpretations of the significance of sentence order might be possible.)

In addition to sentence order, syntax can also serve an important discourse function, that is, to signal distinctions in grounding values. Thus, grounding may be locationally and syntactically encoded as shown in (3).

(3) *Al-munaḏḏama, allatī tuʿānī min mašākila māliyya kabīra, ʾaʿlanat ʾannahā ḥaṣalat ʿalā musāʿadatin qayyima.*
'The organization, which suffers from huge financial problems, announced that it has received valuable assistance'

The structure of the sentence signals the distinction in grounding values. The main clause expresses meaning about the main and most important information (viz. receiving assistance), and hence, it has a foreground value or interpretation. The relative clause expresses meaning about less important information (viz. contextual information about current financial problems), and hence, it encodes a background proposition.

Languages vary in the explicitness of marking relative grounding values in text or discourse. They also vary in the devices they employ and in the contribution of these devices to the grounding-signaling function. In many types of English texts, for example, expressions such as lexical repetition, → pronominalization, and renominalization are in general sufficient to make clear the different grounding values of the underlying propositions. Example (4) illustrates this point.

(4) *Sandra said that…*
She explained that…
Her explanation…
Sandra has been…

Introduced by a personal pronoun, the second sentence expresses specifics of the event referred to in the first sentence; hence, it denotes a midground proposition. Later in the sequence, a proper noun is introduced again, denoting a shift in the level of description and a lower grounding value.

Arabic texts employ devices that play a crucial role in signaling the foreground/background distinction across sentence boundaries. Arabic also makes extensive use of extra words and prefatory expressions such as spatiotemporal and circumstantial expressions to perform important grounding-signaling functions. For details, see Khalil (2000). Among the many expressions that serve a grounding-signaling function in Arabic are prepositional phrases and adverbials such as *mimmā yuḏkaru ʾanna* 'among things to be mentioned is that', and *fī l-masāʾ* 'in the evening'. In sentence-initial position, these expressions make explicit the writer's perspective on the (presentation of the) subject matter and his or her communicative intentions regarding how the foreground/background structure is to be interpreted. Put differently, the choice of a certain expression is determined by the grounding value that the proposition is intended to serve. *Mimmā yuḏkaru ʾanna* signals the writer's perspective on the underlying proposition, that is, that it is tangential and marginally related to other meanings expressed previously. Preposed adverbials of time, place, and manner – included among "syntactic foregrounding devices" (Fareh 1995) – may be due to constraints to express background meaning about the context of the event before expressing other higher grounding values.

Aspectual markers can also have a grounding-signaling function. For example, *qad* and *kāna qad* perform grammatical functions in text that are distinct from their sentential (viz. temporal or circumstantial) functions. Customarily, the preverbal particle *qad* signals a higher grounding value than the grounding value signaled by *kāna qad*. This is illustrated in (5).

(5) *Waṣala r-raʾīsu ʾilā l-maṭār.*
Wa-qad istaqbala-hu ʿadadun kabīrun min an-nās.
Wa-kāna r-raʾīsu qad zāra baʿḍa d-duwal al-ʿarabiyya.
'The president arrived at the airport.
A large number of people received him.
The president had visited a number of Arab countries.'

In the second sentence, *qad* introduces a development, a subevent, of the event reported in the first sentence, thus signaling a grounding value lower than foreground, a midground. The third sentence is introduced by *kāna qad*, and it encodes a countersequential event that is also a typical background function. The marker expresses a temporal relation of antecedence between two past events, a case of "non-sequenced states and actions" as per Hopper (1982b:6).

Expressions occurring in sentence-initial position have been related to three levels of narration (Hatim 1997). The first is introduced by the simple past, and it indicates that the narrative is progressing; the second and third are introduced by *qad* and *kāna qad* respectively. *Qad* introduces the sequences as one aspect of the same set and "at the same time pushes the narration back slightly from the basic level" (Hatim 1997:70). *Kāna qad* "pushes the narration even further back" (Hatim 1997:71). For functional differences between these two discourse markers, see Khalil (1985).

The notion of grounding plays a prominent role in second/foreign language learning. It underlies several textual and linguistic problems, particularly in translation. The learner/translator has to cope with grounding constraints that lie beyond the sentence boundaries and with ensuing problems in recasting the way meaning has been expressed in the source language text. One source of grounding problems in translation involves misinterpreting the grounding function of propositions in the source language text. A second source of grounding problems relates to the absence or omission in the translated text of expressions that may be particularly important for grounding-signaling functions. The absence of these expressions may lead to an imbalance between source and target language texts (see Khalil 2001). An account of how Arabic distinguishes and signals grounding values in different types of text would provide a valuable contribution to studies of second/foreign language.

BIBLIOGRAPHICAL REFERENCES

Callow, Kathleen. 1974. *Discourse considerations in translating the word of God.* Grand Rapids, Mich.: Zondervan.

Couper-Kuhlen, Elizabeth. 1989. "Foregrounding and temporal relations in narrative discourse". *Essays on tensing in English*. II. *Time, text and modality*, ed. Alfred Schopf, 7–29. Tübingen: Niemeyer.

Dijk, Teun van. 1977. *Text and context: Explorations in the semantics and pragmatics of discourse.* London: Longman.

Dry, Helen Aristar. 1992. "Foregrounding: An assessment". *Language in context: Essays for Robert E. Longacre*, ed. Shin Ja J. Hwang and William R. Merrifield, 435–450. Dallas, Tex.: Summer Institute of Linguistics.

Fareh, Shehdeh. 1995. "Foregrounding devices and their functions in Arabic and English written discourses". *Papers and studies in contrastive linguistics*, XXX, ed. Jacek Fisiak, 119–130. Poznan: Adam Mickiewicz University.

Flashner, Vanessa. 1987. "The grammatical marking of theme in oral Polish narrative". Tomlin (1987:131–156).

Fox, Barbara A. 1983. "The discourse function of the participle in Ancient Greek". Klein-Andreu (1983:23–41).

Grimes, Joseph E. 1975. *The thread of discourse.* The Hague: Mouton.

Hatim, Basil. 1997. *Communication across cultures: Translation theory and contrastive text linguistics.* Exeter: University of Exeter Press.

Hopper, Paul J. (ed.). 1982a. *Tense-aspect: Between semantics and pragmatics.* Amsterdam and Philadelphia: J. Benjamins.

——. 1982b. "Aspect between discourse and grammar: An introductory essay for the volume". Hopper (1982a:3–18).

——. 1983. "Ergative, passive, and active in Malay narrative". Klein-Andreu (1983:67–88).

Khalil, Esam N. 1985. "News discourse: A strategy of recasting". *Journal of Pragmatics* 9.621–643.

——. 2000. *Grounding in English and Arabic news discourse.* Amsterdam and Philadelphia: J. Benjamins.

——. 2001. "Grounding and the reconstruction of meaning: A text linguistic perspective on translation". *Translation and meaning*, V, ed. Marcel Thelin and Barbara Lewandowska-Tomaszczyk, 71–83. Maastricht: Maastricht School of Translation and Interpreting.

Klein-Andreu, Flora (ed.). 1983. *Discourse perspectives on syntax.* New York: Academic Press.

Labov, William and Joshua Waletzky. 1967. "Narrative analysis: Oral versions of personal experience". *Essays on the verbal and visual arts*, ed. June Helm, 12–44. Seattle: University of Washington Press.

Reinhart, Tanya. 1984. "Principles of Gestalt perception in the temporal organization of narrative texts". *Linguistics* 22.779–809.

Thelin, Nils. 1984. "Coherence, perspective and aspectual specification in Slavonic narrative discourse". *Aspect bound: A voyage into the realm of Germanic, Slavonic and Finno-Ugrian aspectology*, ed. Casper de Groot and Hannu Tommola, 225–238. Dordrecht: Foris.

Thompson, Sandra A. 1987. "'Subordination' and narrative event structure". Tomlin (1987:435–454).

Tomlin, Russell S. (ed.). 1987. *Coherence and grounding in discourse.* Amsterdam and Philadelphia: J. Benjamins.

Wallace, Stephen. 1982. "Figure and ground: The interrelationships of linguistic categories". Hopper (1982:201–223).

ESAM N. KHALIL (University of Nijmegen)

Gulf States

1. ARABIC AND MINORITY LANGUAGES

A number of languages in addition to Arabic are used in all the modern Gulf States:

Persian. Persian is widely understood throughout the Persian Gulf, though perhaps less than was the case fifty years ago. Older immigrants of Iranian origin, born in Iran in the first half of the 20th century, have retained their Persian alongside the Arabic they learned on the Arab side of the Gulf, but the younger generations, brought up and educated through the medium of Arabic, are well assimilated and many are either losing their Persian or use it only as a domestic language with their parents and grandparents. Over the centuries, Persian has contributed many loanwords to the Gulf dialects, especially words for foodstuffs, domestic goods, textiles, and building and architectural terms (Holes 2001), e.g. *rubyān* 'prawns', *rwēd* 'radishes', *mēwa* 'fruit', *šakkar* 'sugar', *dōšag* 'mattress', *nihāli* 'carpet', *zari* 'gilded cord', *lās* 'silk', *šigirdi* 'building laborer', *gōni* 'builder's set-square', *dirīša* 'window', *bādgīr* 'wind-tower'.

Indian languages. Hindi/Urdu and, to a lesser extent, Punjabi and some South Indian languages have always been widely understood by sections of the population in all the Gulf States. Trade links with India have been strong for centuries if not millennia, and for a hundred years, until

1947, the whole area was governed by British India from Bombay. The doctors, nurses, minor civil servants, engineers, and teachers who came to the area from India to build its infrastructure also incidentally made Indian languages more widely understood. Gulf Arabs working in certain trades, such as jewelry making, have long been familiar with these languages because India is a main source of gold and precious gems, and many local businesses employed Indian craftsmen. Before the advent of modern education in the Gulf, it was normal for the comfortably well off to send their children to India to be educated, and until the 1970s, many Gulf residents would routinely go to India for the medical treatment until then unavailable in the Gulf. Indian films, not necessarily with Arabic subtitles, have always been very popular in the Gulf States. India has also historically been a source of cheap marriage partners for Gulf men unable to afford the high dowries demanded by the fathers of local brides. More recently, there has been an influx of less-educated cooks, maids, and laborers from India, Pakistan, and Bangladesh, which has led to the formation of a Gulf pidgin. As a consequence of these various types of long-standing contact, the Gulf Arabic dialects are permeated with Indian borrowings, especially in employment-related vocabulary and terms for domestic equipment, clothing, and cooking (Holes 2001). Some examples: *krāni* 'clerk', *kačča* 'form, protocol', *banka* 'fan', *čūla* 'stove', *čirfāya* 'bedstead', *bijli* 'torch', *jūniyya* 'sack', *sirwāl* '(women's) trousers', *binjiri* 'bangle', *jūti* 'shoes', *ālu* 'potato', *ṣālūna* 'curry, stew'. Some of these words are now beginning to drop out of use with the general rise in literacy and exposure to Modern Standard Arabic forms.

There are some other languages whose use is limited to one or two of the Gulf States:

Baluchi. Until 1958, the Baluchi-speaking Gwadur area of Pakistan was an Omani possession. Baluchi is the language of a portion of the population in Oman, well-represented in the police force and the army, which is concentrated in the Capital Area and on the Bāṭina coast. Its use is limited to domestic contexts and in-group conversation. Baluchi speakers are in virtually all cases fully proficient in Omani Arabic. There are smaller groups of Baluchi speakers in Bahrain and the southern Gulf States.

Swahili. Swahili is widely understood and spoken in Oman, particularly in the Capital Area. After the revolution that brought Sultan Qaboos to power in 1970, a large number of Zanzibari Omanis, many of them well educated, returned to Oman, their ancestral home, from East Africa and other Gulf States. The East African contingent spoke little or no Arabic. A UNESCO-funded literacy program to teach these 'returnees' (*ʿāʾidīn*) Arabic ran throughout the 1970s.

The Modern South Arabian languages. In the south of Oman, several non-Arabic Semitic languages are spoken: Mahrī, Ḥarsūsī, and Jibbālī (Johnstone 1977, 1981, 1987; Stroomer 1999). These languages, the remnants of the languages spoken before the area became fully Arabized (probably before the Islamic conquests), are now spoken only by a few thousand speakers each (only a few hundred in the case of Ḥarsūsī). All have been heavily infiltrated by Arabic vocabulary, and virtually all speakers are bilingual in their local language and Arabic (→ Modern South Arabian).

Šiḥḥī. This term refers not to a single language but to a heterogeneous collection of dialects, many of them not mutually comprehensible, which are spoken by a grouping generically known as the Šiḥūḥ, tribesmen who live in the remote mountainous regions of the Musandam peninsula in northern Oman and the United Arab Emirates. Some of these appear to be Arabic dialects, but at least one, Kumzārī, is structurally a variety of Persian and is apparently of some antiquity. None of these dialects/languages have been properly described (Jayakar 1904; Thomas 1930).

2. HISTORY OF ARABIC IN THE REGION

Little is known for certain about the linguistic situation in the Gulf before Arabic became the dominant language after the Islamic conquest, but it is beyond doubt that, alongside the Old Arabic tribal dialects, Persian was in use as a trading language and Syriac was used as a language of liturgy and ecclesiastical correspondence by the Christian church in the whole area. Some short funerary inscriptions in Aramaic have been found, and it is likely that Aramaic was in use as a vernacular language by the sedentary agriculturalists of eastern Arabia. An inscription in Greek has been found at Failaka, off the coast of Kuwait. In the south of Oman, as in the whole of southern Arabia, the forerunners of the present-day Modern South Arabian languages must have been in use in ancient times. Monumental cuneiform inscriptions in Hadramitic, an ancient South Arabian language, have been found at the ruined seaport of Samārum (modern Khōr Rōrī) east of Ṣalāla in southern Oman, founded by colonists from the Wadi Hadramawt toward the end of the 1st century B.C.E.

For centuries before Islam, the coast of eastern Arabia provided a south-north corridor for tribal movements, with Taġlib, Bakr bin Wāʾil, and ʾAzd ʿUmān all gradually moving north along it. There were also major migrations, probably from around the 4th century C.E., west to east from Yemen into Oman, and northeast from Yemen into ancient Bahrain and then south into what is now the United Arab Emirates (UAE). These ancient population movements help explain the present-day wide dispersal of certain 'southern' dialect features (see below). All of the Gulf Arabic dialects, including that of northern Oman, also contain words, especially in agricultural and seafaring terminology, that seem to have come into them from Mesopotamia via Akkadian and/or Aramaic (Holes 2002), e.g. *ṭubaʿ* 'to sink', *xinn* 'hold [in a ship]', *xaṣīn/ṣaxxīn* 'axe, digging tool', *zabīl* 'basket'. This ancient linguistic influence, if that is what it is, is not surprising, given the length of time that these coastal regions were under Babylonian commercial and political influence. It became an *idée reçue* among the medieval Arab grammarians that the speech of Bedouin tribesmen from the Gulf Coast – the ʿAbd al-Qays (regarded by the [Shiʿa] Baḥārna of modern Bahrain as one element of their ancestry) and ʾAzd ʿUmān are usually mentioned by name – was the least 'pure' of all because of their contact with Persians and Indians.

3. Modern Arabic dialectology

From the point of view of phonology and morphology, there are two major dialect types in the Gulf region, although each type exhibits a degree of geographical variation (→ Kuwaiti, Bahraini, Omani Arabic). These are usually referred to in the literature as the *badawī* or 'Bedouin' (B) type, and the *ḥaḍarī* or 'sedentary' (S) type. Despite the fact that, with the arguable exception of southeastern Oman, all forms of nomadic pastoralism have ceased in the Gulf, these two dialect types still remain distinct, harking back to a time when they marked important differences in the culture and social organization of the population of the area (and to a limited extent still do). The B/S distinction cuts across national boundaries (which are a relatively recent invention) and still survives in the collective memory. Speakers in most areas of the Gulf will, if asked, unhesitatingly classify themselves, their community, and their speech as either belonging to the *badu* (an alternative term is *ʿarab*) or the *ḥaḍar*. In the case of Bahrain, this distinction is largely coterminous with a sectarian one, but it is the lifestyle distinction that is historically primary, as it is in the other major case where it is still linguistically salient, Oman. Systemic linguistic differences of this type, like differences in dress, are badges of communal identity, and drawing attention to them is met with official coolness, since it tends to subvert modern governmental efforts to create a unified national identity to which all can sign up (see Holes 2005b for an example of how this is manifested in modern media productions).

Many features differentiate the two dialect types, but the key phonological distinction is their reflex of Old Arabic *q*. The Bedouin dialects have a voiced reflex, which may be a uvular *G* or velar *g*, and in the latter case has undergone a further development, conditioned affrication to *j*. The sedentary dialects, on the other hand, have a voiceless reflex, which may be a uvular *q* as in much of Oman, or a velar *k* as in the village farming communities of Bahrain and in some mountain villages of northern Oman. In the sedentary dialects which have this latter *k* < Old Arabic *q* development, original Old Arabic *k* has developed, possibly as a result, a more fronted reflex, which in some dialects is palatalized and in others affricated to *č*, but, either way, occurs in *all* environments, not just in front-vowel ones, as is the case of the Bedouin dialect treatment of Old Arabic *k*. Some further typical distinctions between the Gulf B and S dialects are given in Table 1.

Table 1. Distinctions between Bedouin and sedentary dialects

Features	B dialects	S dialects
Old Arabic *j*	*y*	*g* (Oman) or *j* (Bahrain)
gahawa-syndrome	yes	no
CCvCvC-type verb forms, e.g. *ktibat*	yes	no: *kitbat/katabat*-type
2nd pers. sg. fem. suffix	*-ič**	*-iš*

* But see comments on Oman below.

In other areas of phonology and morphology, a Bedouin/sedentary distinction exists in one area of the Gulf but is absent in another. For example, in Oman, all the B dialects have a final *-n* in 2nd person singular feminine and 2nd/3rd person plural imperfect verbs, as in *tikitbīn*, *tikitbūn*, *yikitbūn*, and the S dialects do not have this *-n*, whereas in Bahrain there is no distinction: both B and S types have the forms with *-n*. Conversely, in Bahrain the B dialects retain the Old Arabic interdentals *ṯ*, *ḏ*, *ḏ̣* while the S dialects have *f*, *d*, *ḍ*, whereas in Oman the B and S dialects all retain the Old Arabic interdentals.

The B dialects from Kuwait to the UAE as a group are all structurally very similar, although they differ slightly one from another in vocabulary. They have become the dominant speech type in the whole Gulf area and are usually what is being referred to by the shorthand term 'Gulf Arabic'. This linguistic dominance has arisen as a result of pulses of migration to the coast that have occurred over a long period. The most recent of these, in the 18th century, gave rise to the current ruling families of Kuwait, Bahrain, Qatar, and the United Arab Emirates. Compared with the Central Arabian dialects from which they are descended (Ingham 1994), these Gulf B dialects have lost a number of morphological distinctions (cf. Ingham 1982:33–62). They do not distinguish gender in the 2nd and 3rd persons of the verb; have become less synthetic in structure, as evidenced by the loss of the internal passive and the heavier use of the analytic genitive; and absorbed much foreign vocabulary.

Although historically, the origin of the difference between them is geographical, the two dialect types, B and S, have acquired new social connotations in some areas. In Bahrain, the S dialect is associated with the socially disadvantaged (Shiʿa) Baḥārna and is somewhat stigmatized, while the B dialect is that of the largely Sunni social, commercial, and political elite (although of course there are also many poor Sunnis). In contrast with the situation in the northern Gulf, it is an S dialect, that of the Capital Area of Muscat and Mutrah, which is, sociolinguistically speaking, the dominant one in Oman.

In Oman, the B/S dialect distinction continues (just) to correspond to differences of geography and lifestyle, with tribes such as the ʾĀl Bu Šāmis, the Durūʿ, the ʾĀl Wahība, and the Janaba still being seminomadic. However, as noted in Table 1, the B/S distinctions in Oman are not always the same as those which typically distinguish Bedouin and sedentary descended groups in northern Arabia, and there are some important structural features that all, or virtually all the dialects of Oman share, whether they are B or S, and which should be thought of as geographically 'southern' rather than lifestyle related (Holes 1989, 1996). The *-iš* 2nd person feminine singular pronoun ending is one such (→ *kaškaša/kaskasa*), the sole exception being the B dialect of the ʾĀl Wahība of the southeastern Oman, which has *-ik*; another 'areal' feature is the *-in(n)-* infix in active participle + object pronoun forms, such as *šāyfinnah* 'I have/you [masc.] have/he has seen it'. Both these 'southern' features are found as far north as the (Shiʿa) Baḥārna dialects of Bahrain and eastern Saudi Arabia, in some coastal dialects of the UAE, and in parts of southern Yemen. This interrupted pattern of dispersion suggests that the present-day sedentary dialects of the periphery of the Gulf and southern Arabia may be the residue of a homogeneous dialect continuum that was once more continuous and connected than it is today, after centuries of Bedouin migrations from Central Arabia.

4. Diglossia

The domains of use of Modern Standard Arabic are very much the same as they are in the rest of the eastern Arab world: formal written texts and their oral performance, as in news bulletins and other formalized media contexts, and 'set piece' scripted public speaking of any kind. In educational contexts, the age of the audience, the nationality of the teacher, and the subject matter are key factors: at the primary and secondary school level, where in some Gulf States the teachers are now mainly local, an 'educated' variety of Gulf dialect is the norm in speech. Expatriate Arab teachers, often Egyptians or Palestinians, will similarly use their colloquial dialect, especially in what are perceived to be culturally neutral subjects like science and mathematics. But at the university level, where there are very many expatriate lecturers (and nowadays these can be from as far afield as Morocco or Tunisia), a more 'standardized' speech style is the norm, particularly in the teaching of subjects like Arabic literature and Islamic studies. Nonscripted speech in formal contexts where the speakers are all local, such as parliamentary debates, are normally in an educated form of the local dialect, as is also the case in the nonformulaic parts of court proceedings (witness statements, for example). Mosque sermons and religious discourse in general (for example in discussions on television) are normally in Modern Standard Arabic. Television plays, soap operas, and comedies with a Gulf setting are very popular, and invariably are performed in dialect. The language used is often a kind of regional koine, not identifiable with the speech of any one state, and making use of features common to the whole area (as described in Holes 1984, 1990), since such media products need to have as broad an appeal as possible if they are to sell. Popular music sung in Gulf dialect has a big following, alongside the better-known products of Egypt and the Levant.

Ordinary speech throughout the Gulf is permeated with words borrowed from English, especially in the spheres of technology, industry, and office employment in which that language is internationally dominant (see Smeaton 1973 for an account of borrowing in the oil industry in eastern Saudi Arabia). The following speech extract, from a description of his job given in the 1970s by a Bahraini worker in an oil refinery with a secondary school certificate-level education, is typical. The underlined words are English borrowings in varying degrees of assimilation to Arabic phonology and morphology:

> *al-ḥīn nāxid fi tonki yaʿni krūd, maxlūṭ ʾāyil u dīzal, u hāy kull šay maxlūṭ...yaʿni nāxdəh min il-bambāt u ndaxləh fi ikstinjah, miṯil šay ṭawīl yaʿni fīh māy, fīh tyūbāt... baʿadēn nxallīh, ndaxləh dāxil hītar*
>
> 'Now we take the crude [oil] in the tank, a mixture of oil and diesel, all completely mixed up...we take it from the pumps and pass it into the (heat) exchanger, that's like a long thing in which there's water, and tubes...then we make it –...we pass it into a heater'.

Nor is it just a question of borrowed nouns. Verbs like *fannaš* 'to dismiss, fire someone from a job', *čayyak* 'to check', *layyak* 'to leak', *kansal* 'to cancel, abolish' respectively from English *finish*, *check*, *leak* and *cancel* are still in common use throughout the Gulf. However, as a consequence of education and increased exposure to Modern Standard Arabic, many older borrowings are now being replaced by Modern Standard Arabic neologisms in the speech of educated younger speakers, e.g. *mōtir* (< English *motor*) is giving way to *sayyāra* 'car', *sbētar* (< English *hospital*) to *mustašfa*, *drēwil* (< English *driver*) to *sāʾiq*. But as fast as the older borrowings disappear, new ones are coming in, e.g. *rimūt* 'remote control', *dīš* 'satellite dish', *jinz* 'jeans', even *jaksan* 'Afro hairstyle' (< the pop singer Michael Jackson, who in the 1970s and 1980s had this hairstyle).

As elsewhere, informal written Arabic containing dialectal forms is common in personal letters, notes, and other ephemera such as newspaper cartoons. More importantly, there are whole genres of popular literature, especially poetry, that are composed and published in a written form of Gulf dialect, using methods of rendering the sounds and morphological forms of the dialect that are more or less standardized throughout the Gulf. This kind of literature is sometimes frowned upon by the language purists, but it is very popular among ordinary people nonetheless.

The presence of many immigrant laborers from the Indian subcontinent who work in an Arabic-speaking environment but know little or no Arabic has led to the formation of an Arabic pidgin in some of the Gulf States – Bahrain, the

United Arab Emirates, and Oman, for certain (documented by Smart 1990). This pidgin is typically used in market transactions, work environments, and other limited speech contexts (it can be heard used by Bahraini farmers giving instructions to their Pakistani laborers, for example) and is remarkably uniform. Like all pidgins, it exhibits a total loss of inflectional morphology and a drastic reshaping of syntactic structures through the grammaticalization of lexemes.

A non-Arabic (Hindi/Urdu) set of syntactic templates and word order seems to underlie it, with 'frozen' verbal, nominal, demonstrative, copular, and negative Arabic elements being slotted in. The fact that there are quite a large number of words of Indian origin in the normal Gulf colloquial, such as *sīda* 'straightaway, directly' in one of the examples below, no doubt assists comprehension. Some examples:

> *ana fīh maˁlūm, hāḏa mā fīh maˁlūm*
> 'I know how [to do that], but he doesn't'
>
> *inta rūḥ sīda mā fīh yarjaˁ*
> 'Go right away and don't come back'
>
> *hāḏa nafarāt sēm sēm haḏāk nafarāt*
> 'These people are the same as those people'

Trade jargons also exist. One such, recorded in the 1970s but now more or less dead, is that of traditional wooden boatbuilders (*galālīf*). Many of the same jargon terms were used in Bahrain and Kuwait (Holes 2005a:34–35), e.g. *waḥḥar* 'to work', *rābaṣ* 'to sit with', *ṭārūd* 'man', *mingar* 'woman'. The jargon was employed in the presence of outsiders in order to prevent them understanding the boatbuilders' talk.

5. 'NATIONAL' DIALECTS

Since the early 1980s, there has been a conscious attempt to develop a regional Gulf identity. Moves toward political integration, embodied in the Gulf Co-operation Council (set up in 1981), have been largely superficial, but there have been attempts to harmonize educational standards across the Gulf and establish prestige regional institutions like the Arabian Gulf University (based in Bahrain). Sporting links within the Gulf abound, and since as long ago as the 1970s there has been an annual football tournament for the Arabian Gulf Cup, competed for by the national teams of the area. The physical isolation of the individual Gulf states, one from another, has been completely removed by a network of fast, metalled roads that now allow one to drive from Muscat to Kuwait (something unthinkable even 25 years ago) in a matter of hours, and even from mainland Arabia to Bahrain via the causeway opened by King Fahd in 1986. Soon there is to be another bridge linking Bahrain and Qatar, symbolizing the newfound friendship between states that until 1999 were bitterly disputing sovereignty over the Hawar Islands before the International Court of Justice in The Hague.

These developments are having their linguistic impact. There is now more contact between Arabs from different Gulf States than ever before, and a type of dialect has arisen, which, while regionally 'Gulf' in pronunciation, morphophonology, 'core' syntactic structures, and vocabulary, eschews obvious localisms. Thus, in conversations between people from different areas of the Gulf, region-wide forms like *baġa, yabbi* 'to want', *bannad* 'to close', *baṭṭal* 'to open' are usually maintained, but more localized forms like (*m*)*āku* (Kuwait), (*mā*) *hast, miš* (Bahrain, Qatar), all meaning 'there is/are (not)' may be dropped in favor of the more regional (*mā*) *fīh*. At a more formal level, the influence of Modern Standard Arabic kicks in, for example in the broadcast media in discussion programs on all kinds of topics. The excerpt below is taken from a Bahrain radio broadcast of the 1980s; it illustrates how Modern Standard Arabic phraseology from a 'literate' frame of reference (here, the theater) is slotted into a dialectal syntactic framework even here, where both interlocutors are Bahrainis. The sections in bold are dialectal, the rest is in an only slightly modified Modern Standard Arabic:

> *bi n-nisba li taqabbul il-jumhūr….* ***fī ḏāk il-wagt****… li ˀayyi masraḥiyya, ṭabˁan, yaˁni kān il-mustawa mustawa taqabbul il-jumhūr – hal kān, yaˁni, bi l-mustawa* ***illi iḥna*** *al-ˀān* ***nčūfah****, mustawa il-jumhūr li taqabbul ˀayy masraḥiyya tuˁraḍ li ˀayy masraḥ, in kān fi l-baḥrēn…*
>
> 'As far as the receptivity of the public is concerned, **at that time**, to any play at all, I mean, was the level of public receptivity, was it, I mean, er…at the same level **that we see** now, the public level of receptivity to any play which is put on, in any theater, if it is in Bahrain?'

6. LANGUAGES OF WIDER COMMUNICATION

Today, English is the major language of wider communication in the area. It impinges in every sphere of the globalized economies of the Gulf States and serves as the lingua franca between the Gulf Arabs and the large multinational communities of educated expatriates they play host to, from Filipino nannies to European bankers to American military personnel, and between such groups. Semipermanent expatriate communities have been a feature of Gulf society for well over a century, their composition a reflection of the political and economic conditions of the time. The English language first arrived in the area in the 19th century as the language of the British imperial authorities, the protecting power that eventually entered into treaty relations with each of what would later become the modern Gulf States. As modern education systems began to take shape, English was introduced as the sole foreign language taught in the school system. English was the language that the American and British owners/operators of the first industrial concerns, the oil companies, brought with them when they arrived in the area in the 1930s and 40s, and which took over as the main employers of local labor with the decline of local industries such as pearl fishing and agriculture. The industrial training schemes run by these companies were in English, and the British Council has since the 1960s been actively involved in English-language teaching in all the Gulf States both in adult education and inside the school system. In the sphere of higher education, the establishment of universities in the Gulf States (the first, Kuwait University, opened in 1966) has seen a further expansion in the role of English. Many science-based subjects, such as medicine and engineering, are routinely studied through the medium of English even at the undergraduate level. The vastly increased influence in the region – commercial, political, and military – of the United States since the formal withdrawal of Britain from the area in 1971 has accelerated the penetration of English into every sphere of life.

BIBLIOGRAPHICAL REFERENCES

Holes, Clive. 1984. *The colloquial Arabic of the Gulf and Saudi Arabia.* London: Routledge.

——. 1989. "Towards a dialect geography of Oman". *Bulletin of the School of Oriental and African Studies* 52.446–462.

——. 1990. *Gulf Arabic.* London: Routledge.

——. 1996 "The Arabic dialects of south-eastern Arabia in a socio-historical perspective". *Zeitschrift für Arabische Linguistik* 31.34–56.

——. 2001. *Dialect, culture and society in eastern Arabia.* I. *Glossary.* Leiden: E.J. Brill.

——. 2002. "Non-Arabic Semitic elements in the Arabic dialects of eastern Arabia". *"Sprich doch mit deinen Knechten aramäisch, wir verstehen es!" 60 Beiträge zur Semitistik. Festschrift für Otto Jastrow zum 60. Geburtstag*, ed. Werner Arnold and Hartmut Bobzin, 269–280. Wiesbaden: O. Harrassowitz.

——. 2005a. *Dialect, culture and society in eastern Arabia.* II. *Ethnographic texts.* Leiden: E.J. Brill.

——. 2005b. "Dialect and national identity: The cultural politics of self-representation in Bahraini *musalsalāt*". *Monarchies and nations: Globalization and identity in the Arab states of the Gulf*, ed. Paul Dresch and James Piscatori, 52–72. Reading: I.B Tauris.

Ingham, Bruce. 1982. *Northeast Arabian dialects.* London: Kegan Paul International.

——. 1994. *Najdi Arabic: Central Arabian.* Amsterdam and Philadelphia: J. Benjamins.

Jayakar, Atmaram. 1904. "The Shihee dialect of Arabic". *Journal of the Bombay Branch of the Royal Asiatic Society* 246–277.

Johnstone, Thomas Muir. 1967. *Eastern Arabian dialect studies.* London: Oxford University Press.

——. 1977. *Ḥarsūsī lexicon*, London: Oxford University Press.

——. 1981. *Jibbāli lexicon.* London: Oxford University Press.

——. 1987. *Mehri lexicon.* London: School of Oriental and African Studies.

Smart, Jack. 1990. "Pidginization in Gulf Arabic: A first report". *Anthropological Linguistics* 32.83–119.

Smeaton, B. Hunter. 1973. *Lexical expansion due to technical change as illustrated by the Arabic of al-Hasa, Saudi Arabia.* Bloomington, Ind.: Indiana University.

Stroomer, Harry. *1999. Mehri texts from Oman, based on the field materials of Prof. Johnstone.* Wiesbaden: O. Harrassowitz.

Thomas, Bertram. 1930. "The Kumzari dialect of the Shihuh tribe, Arabia, and a vocabulary". *Journal of the Royal Asiatic Society* 785–854.

CLIVE HOLES (University of Oxford)

Ġunna → Nasalization

Gypsy Arabic

'Gypsies' is an ambiguous term. It is used on the one hand as a universal term to denote ethnically and linguistically diverse populations of commercial nomads (also known as service nomads, itinerants, or peripatetics). In a more restricted sense, it often refers specifically to the Řom or Romanies of Europe, a population of Indian origin whose language is (or, in the case of some communities, was) a dialect of Romani – *řomani čhib* or *řomanes,* as it is usually referred to by its speakers (see Matras 2002). A further, mixed reading of the term 'Gypsies' might include populations of commercial nomads outside of India who, like the Romanies, are of Indian origin, but who speak an Indian language that is not a dialect of Romani. This includes the populations known as Dom (also Duman, Qurbāṭ, see Karači*)* in the Middle East, whose language is known as Domari (see Matras 1999), populations like the Jat of Afghanistan (Rao 1995) or the Ḍum of the Hunza Valley (Lorimer 1939), who speak Central Indian languages, and perhaps also the Lom or Boša of Anatolia and Armenia (Finck 1907; Patkanoff 1908), who speak Armenian but retain a distinct in-group vocabulary of Indo-Aryan origin known as Lomavren. The broader interpretation of the term is followed here for the purpose of this description, associating it with populations of commercial nomads, irrespective of origin or ethnicity, in the Arabic-speaking area.

Linguistically, three separate phenomena potentially merit attention: (1) the use of an in-group special vocabulary of limited size and usually of limited communicative functions, by groups whose everyday family and community language is a form of Arabic; (2) the incorporation of Arabic structures into the speech of peripatetic communities that constitute linguistic minorities in the stricter sense, that is, who speak a language other than Arabic among themselves but use Arabic in interaction with outsiders: (3) the kind of Arabic dialect, sociolect, or ethnolect used by minority peripatetic communities. The discussion in this entry is limited to the first two phenomena, in the absence of any data on the third.

The use of special vocabulary to cover everyday, nontechnical meanings (or 'basic' vocabulary) in group-internal communication is a well-known, universal feature of peripatetic communities and is documented among diverse communities in many regions and on different continents. Examples are English Cant, Hiberno-English Gammon, Spanish Germanía, German Rotwelsch, Czech Hantýrka, Dutch Bargoens, and more. Such speech varieties are often referred to as → 'secret languages' and, to the extent that they draw on vocabulary deriving from a particular second language, as 'mixed languages'. Their status as full-fledged languages, however, is disputable. Essentially we are dealing with a fixed, albeit often flexible and volatile set of lexical items covering a limited range of meanings, and so with something that might rather be defined as a 'disguised vocabulary' – a reservoir of lexical items that are known only to group members. Its primary function is to exclude outsiders from key portions of the discourse, by disguising key meanings in the sentence. Sometimes, special vocabularies are also used to establish group membership, to flag group identity, or to mark out the dichotomy between insiders and outsiders (see Hanna 1993:80–83). Compared to 'languages' in the normal sense, special vocabularies are thus structurally and functionally restricted. Grammatical structures usually remain unaffected by the special vocabulary. The occasional confusion of special vocabularies with pidgins or creoles is therefore incorrect.

Only limited documentation exists on Arabic-based special vocabularies. It is nevertheless clear that different groups use different sets of vocabularies, although there is quite often some overlap. A clear-cut taxonomy relating groups to types of special lexicon is made difficult due to the paucity of material, and the fact that there is only partial overlap between group name and the composition of the various special vocabulary sets. Thus, any two groups known by names such as Ġajar, Ḥalab, Nawar, Qurbāṭ, or Bahlawān may have identical, partly overlapping, or even entirely different special vocabularies.

The special vocabulary items themselves may be divided into different types. The first type are language-internal formations that have their origin in Arabic itself and derive from a deliberate attempt to disguise everyday Arabic words. This procedure is well attested in other special vocabularies (e.g. 'pig Latin'), and is sometimes referred to as 'cryptolalic formation'. Vycichl (1959) had already presented an overview of different cryptolalic techniques in what he calls the 'slang' of the *Ḥalab is-Sūdān*, whom he encountered in the vicinity of Luxor, Egypt. They tend to match cryptolalic formations that appear in wordlists collected among other peripatetics, for example by Newbold

(1856) among the Ḥalab of Egypt, by von Kremer (1860) among the Ġajar of Upper Egypt, by Hanna (1993) among the Ġajar of Cairo, or by Streck (1996) among the Ḥalab of Sudan. Although cryptolalic formations have their origin in lexical camouflage strategies, the fact that we encounter the same items in various locations and among different groups indicates that the formation strategies are not usually on-the-spot productive techniques. Rather, they belong to the diachrony of the word, having been formed at some earlier point and then transmitted from one generation of users to the next.

Morphological distortion of words is a common cryptolalic formation. Vycichl (1959) mentions the pattern CuCCāC – *ṭurrāg* ‘road’ *(ṭarīg), ṭubbāx* ‘cooked vegetables’ *(ṭabīx)*. A widespread pattern is the insertion of the root of the target word into a special derivation pattern involving *m-* and a suffix *-iš: mubwābiš* ‘door’ (*b-w-b*), *muftāḥiš* ‘key’ (*f-t-ḥ*) (Vycichl 1959); *menáhriš* ‘day’ (*n-h-r*), *maḥráriš* ‘hot’ (*ḥ-r-r*), *mebrádiš* ‘cold’ (*b-r-d*) (Newbold 1856); *maxšábeš* ‘wood’ (*x-š-b*), *midhábeš* ‘gold’ (*d-h-b*), *migbáliš* ‘mountain’ (*g-b-l*), *mutwariš* ‘bull’ (*t-w-r*), *minxališ* ‘palm’ (*n-x-l*) (von Kremer 1860); *maṣabīʿaš* ‘finger’ (*ṣ-b-ʿ*), *madahaībš* ‘gold’ (*d-h-b*), *maxtīamš* ‘ring’ (*xtm*) (Hanna 1993). There are corresponding feminine forms: *mubṭānše* ‘belly’ (*b-ṭ-n*), *misnānše* ‘tooth’ (*s-n-n*), *muwdānše* ‘ear’ (*w-d-n*) (Vycichl 1959); *mubṣālše* ‘onion’ (*b-ṣ-l*), *mubgarše* ‘cow’ (*b-g-r*), *mudānše* ‘ear’ (*w-d-n*) (von Kremer 1860). The two camouflage morphemes may also appear independently. Hanna (1993) notes *manūra* ‘light’ (*n-w-r*), as well as *maʾʾaṣbāḥ* ‘morning’ (*ṣ-b-ḥ*), *maʾʾakbīr* ‘big’ (*k-b-r*), *maʾʾaṣġīr* ‘small’ (*ṣ-ġ-r*). Plain addition of a camouflaging suffix *-ayiš* is noted by von Kremer (1860): *ḥuṣānayiš* ‘horse’, *šagaráyiš* ‘tree’, *ḥadīdáyiš* ‘iron’, *dībáyiš* ‘wolf’; cf. also *aswádiš* ‘black’ (Newbold 1856). While *m-* is clearly the Arabic nominal/participial marker, with *maʾʾ* deriving from the exclamative/emotive form (‘what a…’), the suffix *-iš,* which Vycichl(1959:224) speculates might be an Indo-European nominative ending, is strongly reminiscent of the Domari nominalizer *-iš:* cf. Domari *mang-* ‘to beg’, *mangiš* ‘begging’. Littmann (1920), on the other hand, derives it from *šīʾ* ‘thing’. Phonological distortions are widespread with numerals: *tulit* ‘three’, *rúbiʿ* ‘four’, *xúmis* ‘five’, *sutet* ‘six’, *súbiʿ* ‘seven’, *túmin* ‘eight’, *tiwaʿ* ‘nine’, *ʿušir* ‘ten’ (von Kremer 1860; Streck 1996).

Another widespread cryptolalic strategy is the functionalization of figurative and metaphorical constructions. Von Kremer (1860) notes *mumešayāt* ‘feet’ (from *m-š-y* ‘to walk’), and paraphrases such as *maʾáḥli* ‘dates’ (‘sweet stuff’), *el-ma-ʾasfar* ‘gold’ (‘the yellow stuff’), and *magaswade* < *ma-ʾaswad* ‘coffee’ (‘black stuff’). Metaphors combined with camouflage morphology are found in *baḥarayiš* ‘north’ (*b-ḥ-r* ‘sea’ toward the Mediterranean), *kiblayiš* ‘south’ (toward Mecca, the *qibla)*. In Vycichl’s (1959) list, a special morphological derivation – *mukafʿal* – is employed with metaphorical associations: *mukabwaḍ* ‘eggs; milk’ (*b-y-ḍ* ‘white’), *mukaswade* ‘coffee’ (*s-w-d* ‘black’), *mukaḥmar* ‘one pound’ (*ḥ-m-r* ‘red’ – ‘gold coin’), *muganwara* ‘lamp, light, fire’ (*n-w-r* ‘light, fire’), *mukabwar* ‘fish’ (*būri*, a Nile fish). Other metaphorical extensions include forms such as *yamūy* ‘to drink’ *(mūy* ‘water’) or *sabsab* ‘hair’ *(ysibsib* ‘to comb’) (Hanna 1993). Word derivation may combine figurative or paraphrase formations with generic or dummy words, such as *māx* ‘thing’ or *anta* ‘place’: *māx l-mōya* ‘well’ (‘water thingy’), *anta l-kabīr* ‘town’ (‘big place’), *anta ṣ-ṣaġīr* ‘village’ (‘small place’). Some vocabulary items appear to derive from local usages and idiomatic expressions. Winkler (1936:389, cited in Streck 1996:300), for example, derives the Ḥalab word for money, *buṭūqa* or *baṭqa,* from the Cairene name of a Spanish coin – *abū ṭāqa* – which pictures fields that appear as ‘windows’. Sudanese Ḥalab *kūšī* ‘Black African’ (Vycichl 1959) can be derived from Kush.

Internal (Arabic-based) cryptolalic formations are rarely the sole basis of the special lexicon. Most vocabulary sets also show words that appear to be of foreign origin, although in many cases their etymology remains unclear. Widespread non-Arabic items of unknown origin in the speech of the Ġajar and Ḥalab of Egypt and Sudan, for instance, include *raxwa* ‘food’ and *raxxa* ‘to eat’, *watab* ‘to come’ and *wattab* ‘to bring’, *kodde* ‘woman’, *anta* ‘place’, *dāzī* ‘policeman’, *xušni* (pl. *xašāna)* ‘non-Ġajar’, *hidīd* ‘night’, *māx* ‘thing’ (also ‘one’), and more. Recognizable etymologies show a range of contributor languages. Vycichl (1959) names Aramaic *damax* ‘to sleep’ and *muṭallim* 'blind', and Nubian *tōd* 'boy', *buru* 'girl', and *amanga* 'water'. The Nubian influence can be attributed to contacts with other Sudanese peripatetic groups that are or were Nubian speaking. The Aramaic component, on the other hand, is found in special vocabularies of peripatetic groups as far away as Iran and Afghanistan. It is likely to derive originally from the use of Aramaic as a trade language or lingua franca in the region, although the concrete diffusion of individual Aramaic-origin lexemes into various special vocabularies of present-day peripatetic groups may be much more recent and attributable simply to contact among the various groups and to vocabulary borrowings or admixture of the special lexicon sets. A small number of words of Iranian origin are likely to have been adopted in a similar fashion. While items such as *piyaz* 'onion', *gošt* 'meat', or *deh* 'ten' could be of either Persian or Kurdish origin, others, such as Newbold's (1856:295) Nawar numerals *suso* 'three' and *čar* 'four' point somewhat more clearly to a Kurdish origin. The source of at least some of this vocabulary may in fact be a peripatetic group of Kurdish origin. The *krād* 'Kurds' of the Palestinian West Bank are itinerant metalworkers who speak Arabic but have a special vocabulary that is based partly on Domari and partly on Kurdish. Palestinian Domari itself also contains many Kurdish loans, which is indicative of a prolonged stay in Kurdish-speaking territory prior to immigration into the present location. Among the isolated items of Turkish origin we find *gemi* 'ship' and, especially widespread, *kapi* 'door', which also appears in Domari. A number of items in Newbold's lists of Nawar and Ġajar words, notably *namak* 'salt' and *thoraki* 'a little', appear to be of Hindi origin and may have similarly been transmitted into the special vocabularies of Egypt via other special lexicons.

An interesting contribution to the Arabic-based special lexicon is that made by (European) Romani to the vocabulary of the Ġajar of Egypt, as documented first by Newbold (1856) and later confirmed by Streck (1996) for the Ġajar of Sudan. The two vocabularies share many similarities, including the same deviations from the common Romani shape of the word, for example *gaziye* 'wife' (Newbold) *qazihe* 'woman' (Streck), Romani *gaži* 'woman, wife'; *marey* (Newbold), *mariʾ* (Streck), Romani *maro* 'bread'; *reibo* 'king' (Newbold) *raibó* 'policeman; non-Gypsy' (Streck), Romani *raj* 'non-Gypsy official'. Both vocabularies are mixed and contain also non-Romani items, including, in Newbold's list, widespread items like *kuddi* 'mother' (elsewhere 'woman'), as well as items derived from Domari, such as *bakra* 'sheep' (Domari *bakra*, Romani *bakro*), *sir* 'head' (Domari *sir*, Romani *šero*), *kustúr* 'hand' (Domari *xastúr* 'your hand', Romani *vast*). The phonology of some words, however, points very clearly to a European Romani origin: *mar-* 'bread' (Romani *maro,* Domari *mana*); *šawe* 'boy', *čavo* and *čai* 'girl' (Romani *čhavo* 'boy', *čhave* 'boys', *čhaj* 'girl'); *kam* 'sun' (Romani *kham*); *ker/kir* 'house' (Romani *kher*); *kalo* 'black' (Romani *kalo*); *lašo* 'good' (Romani *lačho*); *manuš* 'man' (Romani *manuš*); *rátsi* 'night' (Romani *ratʾi*); *yag* 'fire' (Romani *jag*); *kagniye* 'fowl' (Romani *kaxni* 'chicken'). The word *balamo/balamu* 'Christian' is a specifically Balkan Romani term denoting 'Greeks'. The presence of *enna* 'nine' (Romani *enja,* from Greek) in Newbold's list further confirms the Balkan Romani origin. Sampson (1928) had, on this basis, suggested that the Egyptian Ġajar were the descendants of Moldavian Romanies who had been taken prisoner by the Ottomans and deported to Egypt as slaves. Sampson was skeptical about some of the items on Newbold's list, suggesting contamination with George Borrow's lists of the Romani vocabulary of Spanish Gypsies, to which Newbold had access. Thus, the verb *sobelar* 'to sleep' appears in its Spanish Gypsy (Caló) form, with a Spanish infinitive ending. However, other items on the list suggest replication of Romani inflected verb forms, which are not present in Caló. Thus we find on Newbold's list words spelled as *khaba* 'to eat', *chúrábi* 'to rob', *laba* 'to bring', which remind us of Romani *xava* 'I eat', *čorava* 'I rob', *lava* 'I take'. The presence of inflected items is partly confirmed by Streck's list, where we find *besheba* 'sit down!' (*beš-* 'to sit', possibly *bešava* 'I sit down') and *awela* 'come here!' (*avela* 'she/he comes'). The data thus suggest at least contacts with a community of Romani speakers, and so a Romani immigration from the Balkans to Egypt, although the circumstances of this immigration cannot be inferred from the linguistic data alone.

A further conspicuous contributor language is Domari, the full-fledged language of Gypsy groups that are scattered throughout the Middle East. Streck's (1996:295–297) wordlist of the Sudanese Bahlawān is almost entirely derived from Domari. Many words closely resemble the citation form known to us from Jerusalem Domari (see Matras 1999; Macalister 1914): *tmaliyen* 'police', *kušketi* 'small', *díes* 'two', *taran* 'three', *aštar* 'four', and more. Most nouns in the list, however, appear in the Domari accusative form (masc. *-as*, also *-es*, fem. *-(i)a*): *qaras* 'donkey', *santas* 'dog', *kuturyes* 'European', *šunes* 'man', *šunya* 'woman', *masiya* 'meat', *jimariya* 'chicken', and more. Some inalienable possessives (body parts, kin) appear with a possessive marker: *ikyos* 'eye' (lit. 'his/her eye'). Verbs tend to appear in the Domari 3rd person singular present form: *sutari* 'to sleep', *qotari* 'to steal'. This selective replication of inflected forms, both nouns and verbs, suggests that the ancestors of the present-day Bahlawān had access to an inflected language and so to a form of Domari that was in everyday use as a full-fledged language. It appears that the special vocabulary was retained following a shift in the community language from Domari to Arabic, a process that is well attested in many Romani communities of Europe. Domari has also enriched the special vocabularies of other groups that were not previously Domari speakers, and we find items like *bakra* 'sheep', *gora* 'horse', *sir* 'head', *sanota* 'dog' in various special Arabic-based vocabularies, but also in Iran and the Caucasus.

There are thus at least three pools of lexicon on which Arabic-based special vocabularies draw. The first is the indigenous, cryptolalic component, the roots of which appear to be old or even ancient, but the patterns may still be productive and allow speakers to create new lexical items. Some vocabularies draw on this source as a primary reservoir – notably the speech of the Sudanese and Egyptian Ḥalab described by von Kremer (1860), Vycichl (1959), and Hanna (1993). Others may incorporate a selection of items, apparently as a result of contact with these user groups. The second source of lexical enrichment comes from occasional contacts with other peripatetic groups and possibly also settled populations who speak another language, such as Nubians, Kurds, Romanies, or Dom, as well as with peripatetics who are users of a different special lexicon. It is through the latter type of contacts that individual items of vocabulary may be diffused far beyond the area reached by speakers of the actual contributor language. The diffusion area of some vocabulary items may therefore cover wide regions in the Middle East and Central Asia (see Windfuhr 2002). Finally, a third source of vocabulary items stems from the selective retention of lexicon from a former separate language spoken by earlier generations, e.g. Domari, Kurdish, or Romani. Streck (1996:302) suggests a three-way classification of special languages, based on the type and sources of vocabulary, which he labels according to the word for 'Christian; European', as the *xašāna-group* for the *luġa ḥalabiyya* (in which most items are internal cryptolalic formations), the *kuttur-group* for speakers of the *luġa bahlawāniyya* (containing Domari-derived items), and the *balamo-group* for the *luġa ġajariyya* (containing a significant number of words of Romani origin, and otherwise a mixed vocabulary). Although the classification is useful, the various patterns of contact among the groups and the layered vocabulary borrowings that result from them complicate the real picture considerably.

A final point for consideration is the Arabic influence on Domari, the archaic Indo-Aryan language spoken by populations throughout the Arabic-speaking regions and beyond, which are known by various names. Descriptions of the language that is called here Domari appear in Pott (1844–1845, 1846), based on a list by Seetzen from Nablus, Palestine, in Newbold's (1856) description of the speech of the Kurbat of Aleppo and Antioch and the Duman of Baghdad, in Groome's (1891) list from Beirut, and in Patkanoff's (1908) essay based on materials collected among the Karači of Tabriz (Iranian Azerbaijan), Maraş, and Antep (eastern Anatolia). Other Domari-speaking communities are known to exist in Lebanon, Syria, and Jordan. The most extensive documentation of the language so far is based on the Jerusalem dialect (Macalister 1914; Matras 1999). As an archaic New Indo-Aryan language, Domari retains the Old Indo-Aryan present conjugation of the verb and passive and causative valency morphology, as well as consonantal case endings. At the same time, like Romani, it renews the past-tense conjugation in a way that is reminiscent of northwestern Indian or 'Dardic' languages like Kashmiri and shows, again like Romani, agglutinated case affixes. In vocabulary, Domari shows Turkish and Kurdish influences and an immense Arabic impact, with some 50 percent of the Swadesh list of 'basic' lexicon (in the long, 207-item version) deriving from Arabic. Arabic loans tend to keep their Arabic phonology, which means that Domari speakers, who have been bilingual in Arabic for many generations now, also have the full range of Arabic phonemes at their disposal. It is noteworthy, however, that some Arabic loans in Jerusalem Domari retain a pronunciation reflecting, presumably, an earlier Arabic contact variety. Domari has, for instance, *qahwē* 'coffee', cf. Jerusalem Arabic *ʾahwe*. Arabic phonology partly influences the pre-Arabic or Indo-Aryan component as well. Intonation and prosody are the most strongly convergent with Arabic, and in some words pharyngealization of stops is also found in the inherited component. As in Jerusalem Arabic, the affricates *j* and *č* are undergoing a shift, to *ž* and *š* respectively. Although *b* and *p* remain distinct, there is a strong tendency toward lenization of *p*. Arabic verb roots are adapted to Domari through a strategy reminiscent of most Indo-Iranian as well as Turkic languages, whereby a 'carrier' verb, either transitive (from *kar-* 'to do') or intransitive (from *hr-* 'to become') carries the inflection (→ Persian). The Arabic base that is selected is not, as in many other languages, a nominal form or *maṣdar*, but a reduced form of the imperfect/imperative: *štrī-karami* 'I buy', *fhim-homi* 'I understand'.

Jerusalem Domari has in effect undergone what might be referred to as structural fusion with Arabic in the domain of clause combining. All connectors, conjunctions, interjections, and discourse markers derive from Arabic, and word order in the basic and complex clause is virtually identical to that in Arabic. Inflected Arabic conjunctions and particles, such as *inn-* 'that' or the resumptive pronoun *iyyā-*, retain their Arabic agreement inflection, leading for instance to the introduction in Domari of 3rd-person gender distinctions that are otherwise not present in the inherited (Indic) pronominal system. A further domain of near-complete fusion is the area of modality. Domari retains its own tense and modality inflection, but all modal and aspectual auxiliaries, with the exception of *sak-* 'to be able to', are borrowed from Arabic and retain, wherever relevant, Arabic person and tense inflection: *biddī laham* 'I want to see', *lāzim džam* 'I must go', *ṣārat rowari* 'she began to cry'. Most sentential adverbs, as well as many temporal expressions, are Arabic, and almost the entire inventory of prepositions is borrowed from Arabic, with the exception of several person-inflected forms (such as 'for-', 'about-', and 'with- me, you, etc.'): *maʿ* 'with', *la* 'to', *fī* 'in', *bēn* 'between', *min* 'from', *ʿind* 'at', etc. While Macalister's (1914) description of Jerusalem Domari still shows the full Indic series of numerals, present-day Domari (Matras 1999) has retained only the Indic numerals for one through five, ten, and one hundred, replacing the others by Arabic numerals. A lexical-typological oddity is the wholesale borrowing of the comparative-superlative form and with it the Arabic lexical form of the adjective, so that all Domari adjectives, even the basic adjectives that are Indic, have suppletive, Arabic-derived, non-positive forms: *tilla* 'big' – *ʾakbar* 'bigger', *kištota* 'small' – *ʾaẓġar* 'smaller'.

Finally, Domari is undergoing further convergence with Arabic by generalizing those structures that are closest to the counterpart Arabic configuration. In the domain of adjective attribution, the inherited Indic word order is Adjective-Noun: *tilla zara* 'the big boy', *tillī šonī* 'the big girl'. However, there is a clear preference toward the use of predicative adjectival constructions in place of the normal attributive construction, for the former agree in their word order with the Arabic pattern. Thus, we normally find *zarēk tillēk* 'the big boy' (lit. 'the boy, being big'), *šōnik tillik* 'the big girl' (lit. 'the girl, being big'). In the possessive attribu-tive construction, the formation Possessor-Head (*båyim kuryos* 'my father's house', *barim kuryos* 'my brother's house') is being replaced by the construction Head-Possessor: *kury-os båyim-ki* 'my father's house' (lit. 'his-house of-my-father'), cf. Jerusalem Arabic *bēt-o la- ʾabūy*. Although word order in the verb phrase is generally identical to Arabic, Domari has retained just one trace of the Indic verb-final order, namely the present-tense enclitic copula, which in Arabic is matched by the nominal clause: *ama mišta-hromi* 'I am ill', *pandži mišt-ēk* 'he is ill', cf. Jerusalem Arabic *ʾana marīḍ-Ø, huwwe marīḍ-Ø*. In the other tenses, Arabic auxiliaries are employed to maintain the similarity to the Arabic construction:

ama kunt mišta-hroma 'I was ill', *pandži kān mišt-ēk* 'he was ill', cf. Arabic *ʾana kunt marīḍ, huwwe kān marīḍ*. Those areas of structure that are not prone to convergence with, or substitution through, Arabic include a selection of basic vocabulary, perhaps as few as 500–600 lexical roots, among them most body parts and verbs of movement and physical activity, some but not all kinship terms, and the numerals under 'five'; gender, number, and person agreement rules, and corresponding nominal and verbal (and to a lesser extent adjectival) inflectional morphology, including subject and object concord and possessive inflection; synthetic valency-derivation, aspect, tense, and modality formation in the verb; pronouns and demonstratives, as well as place deixis and some time deixis expressions; and some basic expressions for local and spatial relations, including some inflected prepositions that derive from them. Arguably, this is an extremely limited set of structural features and a limited vocabulary range. The primary function of such a system appears to be to maintain and flag group separateness, which makes it functionally related, albeit only partially, to the special vocabularies discussed above.

Bibliographical references

Burton, Richard F. 1898. *The Jew, the Gypsy, and El Islam*. London: Hutchinson.

Canova, Giovanni. 1981. "Note sulle tradizioni zingare in Egitto attraverso la testimonianza di un capo Nawar". *Lacio Drom* 17:6.4–25.

Finck, Franz N. 1907. "Die Grundzüge des armenisch-zigeunerischen Sprachbaus". *Journal of the Gypsy Lore Society* (New Series) 1.34–60.

Galtier, Émile. 1912. "Mémoires et fragments inédits." *Mémoires de l'Institut Français d'Archéologie Orientale du Caire* 27.1–9.

Groome, Francis H. 1891. "Persian and Syrian Gypsies". *Journal of the Gypsy Lore Society* 2.21–27.

Hanna, Nabil Sobhi. 1993. *Die Ghajar*. Munich: Trickster.

Kruse, Fr. 1854. *Ulrich Jasper Seetzen's Reisen durch Syrien, Palästina, Phönicien* [...]. II. Berlin: Reimer.

Littmann, Enno. 1920. *Zigeuner-Arabisch: Wortschatz und Grammatik der arabischen Bestandteile in den morgenländischen Zigeunersprachen*. Bonn and Leipzig: Kurt Schroeder Verlag.

Lorimer, D.L.R. 1939. *The Ḍumāki language*. Nijmegen: Dekker en van de Vegt.

Macalister, R.A.S. 1914. *The language of the Nawar or Zutt, the nomad smiths of Palestine*. Edinburgh: Edinburg University Press.

Matras, Yaron. 1999. "The state of present-day Domari in Jerusalem". *Mediterranean Language Review* 11.1–58.

—— 2002. *Romani: A linguistic introduction*. Cambridge: Cambridge University Press.

Newbold, F.R.S. 1856. "The Gypsies of Egypt". *Journal of the Royal Asiatic Society of Great Britain and Ireland* 16.285–312.

Patnakoff, K.P. 1908. "Some words on the dialects of the Transcaucasian Gypsies: Boša and Karaci". *Journal of the Gypsy Lore Society* (New Series) 1.229–257, 2.246–266, 2.325–334.

Pott, August F. 1844–1845. *Die Zigeuner in Europa und Asien*. Halle: Heynemann.

—— 1846. "Über die Sprache der Zigeuner in Syrien". *Zeitschrift für die Wissenschaft der Sprache* 1.175–186.

Rao, Aparna. 1995. "Marginality and language use: The example of peripatetics in Afghanistan". *Journal of the Gypsy Lore Society* (Fifth Series) 5.69–95.

Sampson, John. 1928. "The Ghagar of Egypt: A chapter in the history of Gypsy migration". *Journal of the Gypsy Lore Society* (Third Series) 7.78–90.

Streck, Bernhard. 1996. *Die Ḥalab*. Munich: Trickster.

Von Kremer, Alfred. 1860. *Aegypten: Forschungen über Land und Volk während eines zehnjahrigen Aufenthalts*. Leipzig: Brockhaus.

Vycichl, Werner. 1959. "The slang of the Ḥalab is-Sūdān". *Journal of the Sudan Antiquities Service* 7.223–228.

Windfuhr, Gernot. 2002. "Gypsy dialects". *Encyclopædia Iranica*, ed. Ehsan Yarshater, XI, 415b–421b. London and New York: Routledge and Kegan Paul.

Winkler, Hans A. 1936. *Ägyptische Volkskunde*. Stuttgart: Kohlhammer.

Yaron Matras (University of Manchester)

H

Ḥaḏf

There is a tendency in Semitic languages (notably in Geʿez) for short vowels *i* and *u* to become *ə* and, often, to disappear. This tendency is apparent in Arabic in a restricted number of cases only, a process the Arab grammarians call *ḥaḏf* → 'elision'. For *huwa* and *hiya* preceded by *fa-, wa-, la-, ʾa-* , an elision may occur: *wa-huwa* or *wa-hwa, fa-hiya* or *fa-hya, la-huwa* or *la-hwa, ʾa-hiya* or *ʾa-hya*. In the jussive, forms like *fa-li-yaqtul* and *wa-li-yaqtul* may have an elision of *i*: *fa-l-yaqtul, wa-l-yaqtul*. As for verbs and nouns, in verbs of the pattern *faʿula, faʿila (maʿlūm)* and of the pattern *fuʿila (majhūl)*, elision of *u/i* may occur: *karuma* ~ *karma* 'to be generous, magnanimous', *ʿalima* ~ *ʿalma* 'to know', *ʿuṣira* ~ *ʿuṣra* 'to be pressed (out), squeezed (out) [grapes, etc.]', *kabid* ~ *kabd* 'liver', *ʿaḍud* ~ *ʿaḍd* 'upper arm' (Sībawayhi, *Kitāb* II, 277.21, 22, 23; 278.2ff.). Sībawayhi (*Kitāb* II, 277.23) regards these elisions as *luġāt* 'idioms' which are transmitted from the tribes Bakr ibn Wāʾil and many Banū Tamīm (*wa-hiya luġa Bakr ibn Wāʾil wa-ʾunās kaṯīr min Banī Tamīm*). Some other sporadic examples of elision of *i* are mentioned by the ancient Arab grammarians: *ʾarāka muntafxan* 'I see you swollen', *muntaṣban* instead of *muntaṣiban* 'set upright', *inṭalqa* instead of *inṭaliq* 'go away!', and in the proverb *lam yuḥram man quṣda lahu* (instead of *quṣida*) 'the one who lives frugally does not feel frustrated' (Sībawayhi, *Kitāb* II, 277.23, 278.1; incorrectly Fleisch 1961:157d, who reads *fuṣda* and *fuṣida* instead of *quṣda* and *quṣida*). Although disappearance of *u/i* in the forms mentioned here seems to be a general tendency in the Semitic languages, *faʿl* is not necessarily a reduced or elided form of *faʿil* in all Semitic languages. Arabic *malik* and Hebrew *malk* 'king' may be lexically different base forms (Fleisch 1961:158). It should be added that Sībawayhi (*Kitāb* II, 274.5ff.), in the chapter on the six back consonants (*al-ḥurūf as-sitta*: *x, ġ, ḥ, ʿ, ʾ, h*), mentions a usage specific to the Banū Tamīm, namely the possibility for those nouns with the *wazn faʿil* and verbs with the pattern *faʿila* that have one of the *ḥurūf sitta* as a second radical consonant to appear in four different forms. For the verb these are *šahida, šahda, šihida, šihda* 'to witness'; for the noun: *faxiḏ, faxḏ, fixiḏ, fixḏ* 'thigh'. In *šihda/fixḏ* elision of *i* occurs (from *šahida/fixiḏ*), and *šihida/fixiḏ* are formed through vowel harmonization on the basis of *šahida/faxiḏ* (Sībawayhi, *Kitāb* II, 274.6–7; Fleisch 1961:158–159).

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCE

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. Ed. Hartwig Derenbourg. 2 vols. Paris: Imprimerie Nationale, 1889.

SECONDARY SOURCE

Fleisch, Henri. 1961. *Traité de philologie arabe.* I. *Préliminaires, phonétique, morphologie nominale.* Beirut: Imprimerie Catholique.

JACQUES GRAND'HENRY
(Catholic University of Louvain)

Hadrami Arabic → Wadi Hadramawt Arabic

Ḥāl

Ḥāl (mostly fem.) is literally the 'state' or 'situation' of someone or something subject to change. The plural *ʾaḥwāl* is used as a plural of multitude, hence 'circumstances' (Lane II, 675). *Ḥāl* as a grammatical term is used for an adverbial expression or a nominal or verbal phrase denoting the circumstances of either the subject or object (*ḏū l-ḥāl* or *ṣāḥib al-ḥāl*), or both, of the act taking place. It is complementary to the sentence and answers the question 'how?', hence the English denotation of 'circumstantial' accusative or circumstantial clause. *Ḥāl* is rendered variously as "denotative of state" (Howell 1990:I, 238–239); "state or condition" (Wright 1967:II, 113, 115); *jumla ḥāliyya* 'circumstantial clause' (Wright 1967:II, 333); "circumstantial accusative" (Cantarino 1975:II, 186); "circumstantial clauses" (Cantarino 1975:III, 242); "circumstantial qualifier" (Badawi a.o. 2004:156, 456); "circumstantial qualifying clauses or phrases (Badawi a.o. 2004:579); cf. Wehr 1994:252; Cachia 1973:34; Dahdah 1988:88. In German, *ḥāl* is rendered as "Zustandsausdruck, Zustandssatz" (Reckendorf 1977:97, 447; cf. Kluge 1999:43) and in French as "complément circonstantiel, proposition circonstantielle" (Blachère and Gaudefroy-Demombynes 1975:397; Blachère 1985:150).

In Standard Arabic, the circumstantial clause may be an asyndetic construction beginning immediately with an imperfect verb, as in *xaraja yaḥmilu l-kitāb fī yadihi* 'he went out carrying the book in his hand'. The circumstantial clause may be a syndetic construction as well, introduced by *wa-*. This introductory *wāw* indicates the simultaneousness of the main clause and the circumstantial clause, as in the English 'while, when, although', e.g. *jarat hāḏihi l-waqāʾiʿu wa-l-ḥarbu qāʾimatun* 'these events occurred while the war was going on'; *kataba maktūban wa-huwa malikun* 'he wrote a letter, although he was king'; or *raʾaytu ṭāriqan fī l-bayti wa-huwa yanhāru* 'I saw Ṭāriq in the house while it [the house] was falling down'. However, when *wa-qad* followed by a perfect verb is used to introduce the circumstantial clause, the circumstance or situation is the result of an act anterior in time to the main statement: *jarat hāḏihi l-waqāʾiʿu wa-qad zālat al-ḥarb* 'these events occurred when the war had ceased'; *sa-yastaqbilunī wa-qad qaraʾa maktūbī* 'he will receive me now that he has read my letter'; *kāna yaʾinnu wa-qad waḍaʿa yadahu ʿalā qalbihi* 'he was groaning, having put his hand on his heart' (examples taken from Bateson 1967:47; Beeston 1975:81–82; Cantarino 1975:III, 242–244; Badawi a.o. 2004:579, 586).

Usually, the syndetic circumstantial clause cannot precede the main clause, but it may be, and frequently is, inserted in the main clause immediately following the subject or object whose specific circumstance it describes (i.e. the *ḏū l-ḥāl*): *sāfartu wa-ʾana fī hāḏihi l-ḥāli ʾilā l-baḥrayni* 'I departed, still in this condition, for Bahrain'; *sa-tanqilu wa-ʾanta fī najd ʾilā l-qarni l-xāmis* 'while you are in Najd, you will be transported back to the 5th century' (Cantarino 1975:III, 278–279). However, Badawi a.o. (2004:584) mention the possibility of inversion of main and circumstantial clauses for emphasis, as in *wa-hum yatanāwalūna l-qahwata taḥaddaṯa ʿan baʿḍi l-ʾanmāṭi llatī qābalahā fī s-sijni* 'while they were drinking coffee, he talked about some of the types he had encountered in prison' – an illustration of Arabic as a living language. When introduced by the particle *ʾammā*, precedence of the circumstantial clause is allowed, a phenomenon that is even "becoming increasingly common" in *ʾammā…fa-* constructions when using the syndetic *ḥāl* introduced by *wa-qad*: *ʾammā wa-qad futiḥat ʾabwābu l-qaṣri l-jumhūriyyi li-l-ḥiwāri…fa-ʾinna mā ḥadaṯa…* 'as for the doors of the republican palace having been opened for dialogue…then what happened…' (Badawi a.o. 2004:586–587; cf. Cantarino 1975:III, 279).

An indefinite accusative, mainly of an active or passive participle, may replace the circumstantial clause: *intaxabūhu wa-huwa ġāʾibun ʿan al-ʿāṣima* 'they elected him while he was absent from the capital' is equivalent to *intaxabūhu ġāʾiban ʿan al-ʿāṣima* (Beeston 1975:94). The following examples illustrate further usage of the circumstantial accusative: *mašaytu mutamahhilan* 'I walked slowly' (like *ġāʾiban*, an active participle); *wa-tuhādā n-naʿšu maḥmūlan ʿalā l-ʾaʿnāqi* 'the bier slowly moved off, carried on shoulders', with the passive *maḥmūlan* (Badawi a.o. 2004:156). Sometimes, a substantive, an infinitive, or, rarely, an adjective is used: *ʾaxaḏa yatamaššā fī l-qāʿati ḏahāban wa-jīʾatan* ‘he began to walk back and forth in the room’; *min al-muḥzini ʾannaka lam tūlad bintan!* ‘it is a pity that you were not born a girl!’; *taxarraja l-waladu ḍābiṭan fī l-baḥriyyati* ‘the boy graduated as an officer in the navy’ (Cantarino 1975: II, 195; Badawi a.o. 2004:157). Although, as stated above, the circumstantial accusative is normally indefinite, it may be defined by a following substantive noun, basically forming an improper annexation, as in *fa-jalastu ḍayyiqa ṣ-ṣadri* ‘so I sat down with a heavy feeling about me’; *ṣirtu ʾasīru fī l-ḥārati kasīra r-rūḥi* ‘I started going around the quarter with a broken spirit’ (Badawi a.o. 2004:158, 580).

Sequences of circumstantial clauses and/or circumstantial accusatives occur asyndetically, as in *al-waʿdu llaḏī qaḥaʿahu ʿalā nafsihi muxliṣan ṣādiqan* ‘the promise he made to himself sincerely and honestly’ and *wa-ʾaqifu ḥāʾiran ʾasʾalu nafsī*…‘and I would stand confused, asking myself…’; and syndetically, as in *wa-btasama samʿānu muʾānisan wa-mušajjiʿan* ‘and Samʿān smiled in a friendly and encouraging way’. Combinations of asyndetic and syndetic constructions are found frequently: *qālat hāmisatan wa-hiya tabtasimu* ‘she said, whispering, as she smiled’ (Badawi a.o. 2004:157, 582, 584; also Cantarino 1975:II, 190-191, 496–497).

Ḥāl may denote the circumstances of objects as well, for instance in *biʿtu š-šāʾa šātan wa-dirhaman* ‘I have sold the sheep at a dirham apiece’ or *bayyantu lahu ḥisābahu bāban bāban* ‘I explained his account to him item by item’; or there may be more *ḥal*s referring to both the subject and the object: *laqītu hindan muṣʿidan munḥadiratan* ‘I, going up, met Hind [a woman’s name] coming down’. When both *ḥāl*s agree in gender and number, confusion may arise unless the circumstantial accusatives are placed behind the *ḍū l-ḥāl*: *laqītu muṣʿidan zaydan munḥadiran* ‘I, going up, met Zayd coming down’ (Wright 1967:II, 115; Howell 1990:I, 240–241).

All preceding examples are *ḥāl* expressions in the affirmative, but negative circumstantial clauses occur as well, e.g. *wa-ʾajabtu wa-ʾana lā ʾaʿrifu li-māḏā yuʿākisunī ramzī hākaḏā* ‘and I answered without knowing why Ramzī was bothering me in this way’; *qad marra bī ʾakṯaru min ʿāmin wa-lam ʾaraka* ‘it has been more than a year since I saw you last’; *kāna llāhu wa-lā šayʾa maʿahu* ‘God existed with nothing else (existing) with him’ (Cantarino 1975:III, 272; Badawi a.o. 2004:585).

In the Arabic linguistic tradition, the nature of the *ḥāl* as an adverbial expression is often explained alongside → *tamyīz*, the accusative of ‘specification’ (as in *ṭāba l-wardu lawnan* ‘the rose is charming in color’; cf. Wright 1967:II, 122). A neat summary is given by the 8th/14th-century scholar Ibn Hišām al-ʾAnṣārī (d. 756/1355), whose eminence ultimately earned him the title of ‘second Sībawayhi’ (Gully 1995:7–8). Ibn Hišām (*Muġnī* II, 532–535; cf. Kluge 1999:43–50) finds five points of resemblance between *ḥāl* and *tamyīz* and seven points of difference, as follows: *Ḥāl* and *tamyīz* agree in that they are both nouns (*ʾasmāʾ*), indefinite (*nakira*), dispensable (*faḍla*, i.e. redundant for rendering a grammatically correct sentence), and in the accusative (*manṣūb*), and they both serve to clarify (*rāfiʿa*) what is unclear. The seven distinctions between *ḥāl* and *tamyīz* can be summarized as follows: (1) the *ḥāl* may be a sentence, a *ẓarf* (i.e. adverbial accusative), or a prepositional expression, whereas the *tamyīz* can only be a noun; (2) sometimes the meaning of the sentence depends on the *ḥāl*, which is not the case with *tamyīz*; (3) the *ḥāl* clarifies exterior circumstances whereas the *tamyīz* clarifies inner essences; (4) the *ḥāl* may consist of several components (because it clarifies exterior circumstances, which can be many), unlike *tamyīz*; (5) the *ḥāl* can precede its regent and the *tamyīz* cannot; (6) in principle, the *ḥāl* is derived from a verb and *tamyīz* is a noun proper (but sometimes it is the other way round); and (7) while *ḥāl* can be used to strengthen or intensify its regent, *tamyīz* has no strengthening power.

The basic rules of the use and the nature of circumstantial expressions as sketched above (in Classical as well as in Modern Arabic) are laid down by Sībawayhi, with references to *ḥāl* dispersed throughout his *Kitāb*, beginning with Chapter 16 (*Kitāb* I, 15–16), titled, “This is the chapter dealing with what the verb puts in the accusative case because it is the *ḥāl* in which the action takes place” (*hāḏā bāb mā yaʿmalu fīhi l-fiʿlu fa-yantaṣibu wa-huwa ḥālun waqaʿa fīhi l-fiʿl*). Here, Sībawayhi explains that for instance in *ḏahaba zaydun rākiban* ‘Zayd came riding’, the verb *ḏahaba* puts *rākiban* in the accusative not as a direct object – *ḏahaba* is intransitive – but as a *ḥāl*, indicating the way

Zayd came (for the main references to *ḥāl* in the *Kitāb*, see I, 155–156, 158–159, 161–163, 168–169, 169–170, 211–212, 218–219, 221–222, 222–224). This at once goes to the heart of the matter as far as the discussion of *ḥāl* within the Arabic linguistic tradition is concerned. The discussion focuses mainly on the circumstantial accusative and evolves around its regent, i.e. what puts the *ḥāl* in its characteristic accusative case, something that has a direct bearing on the *ḥāl*'s position in the sentence as a whole.

In Chapter 16 Sībawayhi introduces the notion of 'intervention, separation' (*ḥāla bayna*) to explain the occurrence of the circumstantial accusative, even if no verb or *ʿāmil* carrying the sense of a verb is present to act as regent and to put the *ḥāl* in the accusative. The underlying principle of this notion of *ḥāla bayna* is identified by Carter (1972) in his classic analysis "'Twenty dirhams' in the *Kitāb* of Sībawayhi". Carter explains that the phrase *ʿišrūna dirhaman* 'twenty dirhams' is used by Sībawayhi to indicate that regency is not restricted to verbs or nonverbal regents carrying the sense of a verb, but that certain classes of words have the same influence, though less powerful.

In Sībawayhi's example mentioned above (*ḏahaba zaydun rākiban*), the expression *ḏahaba zaydun* renders a grammatically correct sentence whose completeness is marked by *tanwīn*, the *nūn* of *zaydun*. Just like any other adverbial accusative, *rākiban* is redundant, not identical with or included in its antecedent, and as such it is a surplus to an already complete utterance. In the expression *ʿišrūna dirhaman*, *ʿišrūna* represents the completeness of *ḏahaba zaydun*, and *dirham* represents the redundancy of *rākiban*. The *nūn/tanwīn* of *ʿišrūna*, indicating completeness, separates the two elements of the sentence and prevents the genitive case in *dirham*. *Dirham* subsequently obtains the accusative case in what Carter (1972:485) has dubbed a *tanwīn-naṣb* construction (called *tanwīn-ʾalif* construction by Blau 1981:183, 204, 206–207; cf. Hopkins 1984:168–169). The *nūn/tanwīn* of *zaydun* in *ḏahaba zaydun rākiban* stands – just as the *nūn/tanwīn* of *ʿišrūna* does in *ʿišrūna dirhaman* – between the *ʿāmil*, the verb *ḏahaba* (which gives *zayd* the nominative case as its subject) on the one hand, and the redundant *rākiban* on the other. It is this sense of separation and redundancy that gives *ḏahaba* the power to put *rākiban* as *ḥāl* in the accusative case (Carter 1972:488–490; cf. Owens 1990:107ff., who calls it the "separation and non-identity principle, SNIP for short"; see also Talmon 1993, Talmon 2003:245ff.).

In this light, Sībawayhi's initial explanation – that in the expression *ḏahaba zaydun rākiban* the verb *ḏahaba* is the regent that puts *rākiban* in the accusative as a *ḥāl* – clearly implies that the regent of the *ḥāl* may be an expression without a verb, not even carrying the sense of a verb. This is, for instance, the case in *ḏālika ʿabdullāhi ḏāhiban* 'that is ʿAbdullāh coming', or *fīhā ʿabdullāhi qāʾiman* 'in it ʿAbdullāh is standing' (*Kitāb* I, 218–220, 222–224). The lack of a verbal sense in the demonstrative *ḏālika* and in the prepositional clause *fīhā* has direct bearing on the position of the *ḥāl* in the sentence. That is, the *ḥāl* may precede its regent only on the condition that this regent is a fully inflected verb. This principle is based on the fact that *rākiban* in *ḏahaba zaydun rākiban* is comparable to, for instance, *ʿamran*, which is the direct object in *ḍaraba zaydun ʿamran* 'Zayd hit ʿAmr'. In the latter case, precedence of the direct object is allowed – *ʿamran ḍaraba zaydun* is grammatically correct – and, therefore, precedence of the *ḥāl* is allowed in similar constructions. However, when the regent is not a fully inflected verb, the comparison cannot be made, and as a consequence, precedence of the *ḥāl* is not permitted. Hence, **ḏālika ḏāhiban ʿabdullāhi* and **fīhā qāʾiman ʿabdullāhi* are rejected, just as **ḏāhiban ḏālika ʿabdullāhi* and **qāʾiman fīhā ʿabdullāhi* are unacceptable (see Sībawayhi, *Kitāb* I, 237–238; cf. e.g. Mubarrad, *Muqtaḍab* III, 36–37, IV, 168–169; Ibn Wallād, *Intiṣār* 38–39, 79–81, 96–97; Ibn Jinnī, *Lumaʿ* 26–27; ʾAbū Ḥayyān, *Manhaj* 194–195, 228; Zamaxšarī, *Mufaṣṣal* 28, 30; Ibn Yaʿīš, *Šarḥ* 57, 73). For the interpretation of *ḥāl* regency and its position in the sentence in Western grammars, see, e.g., Wright 1967:II, 218–220; Howell 1990:I, 244–251).

The general rules on *ḥāl*, as found in Sībawayhi's *Kitāb* and, more crystallized, in later grammatical works, go back to the earlier grammatical tradition which Talmon (2003) identifies as the Old Iraqi school of grammar (including both Baṣran and Kūfan scholars; for a discussion of differences of opinion on various issues concerning *ḥāl* in the light of the

Baṣran/Kūfan dichotomy, see Ibn al-ʾAnbārī, *ʾInṣāf* 112–118, *masāʾil* nos. 31–33). Underlying the analysis of *ḥāl* is the notion of *xabar manṣūb* lit. 'predicate in the accusative case', or *xabar al-maʿrifa* 'predicate of something known', labeled by Talmon (1993:95, 2003:40) "transformed predicate", i.e. a predicate that is separated from its subject and "transforms consequently to a 'second rate' predicate position marked grammatically by *naṣb*". As appears from Talmon's research, the earliest recorded analysis of this kind of structure in terms of *xabar al-maʿrifa* comes from al-ʾAxfaš al-ʾAwsaṭ (d. between 210/825 and 221/835), who refers to previous generations of linguists including Ibn ʾAbī ʾIsḥāq and ʿĪsā ibn ʿUmar (as well as to the *Qurʾān* reader and Companion of the Prophet, Ibn Masʿūd, regarding his interpretation of Q. 36/57 *wa-lahum mā yaddaʿūna salāman*; Talmon 2003:117, 185–194; see also Owens 1990:119–120). The notion of 'transformed predicate' is treated by al-Farrāʾ (d. 207/822) in terms of either → *xabar* or *fiʿl* (Kinberg 1996:205, 595–596; cf. Talmon 2003:40, 190).

Sībawayhi (*Kitāb* I, 218), probably following his most important teacher, al-Xalīl (see Talmon 2003:40), uses the term *xabar al-maʿrifa/xabar li-l-maʿrūf* to explain the difference between a predicate of someone or something not known to the listener and hence in the nominative, and a predicate of someone or something known to the listener, which has a clarifying or strengthening function and obtains the accusative. Sībawayhi identifies this notion of *xabar* more than once as *ḥāl*, thus reformulating the *ḥāl* category into the standard concept it has remained until this day (Talmon 1993:96, 2003:295–297; cf. Levin 1979:193ff.).

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

ʾAbū Ḥayyān, *Manhaj* = ʾAbū Ḥayyān Muḥammad ibn Yūsuf al-Ġarnāṭī, *Manhaj as-sālik fī l-kalām ʿalā ʾAlfiyyat Ibn Mālik*. Ed. Sidney Glazer. New Haven, Conn.: American Oriental Society, 1949.

Ibn al-ʾAnbārī, *ʾInṣāf* = ʾAbū l-Barakāt ʿAbd ar-Raḥmān ibn Muḥammad ibn ʾAbī Saʿīd al-ʾAnbārī, *Kitāb al-ʾinṣāf fī masāʾil al-xilāf bayna n-naḥwiyyīna l-Baṣriyyīna wa-l-Kūfiyyīn*. Ed. Gotthold Weil. Leiden: E.J. Brill, 1913.

Ibn Hišām, *Muġnī* = Jamāl ad-Dīn ʾAbū Muḥammad ʿAbdallāh ibn Yūsuf Ibn Hišām, *Muġnī l-labīb ʿan kutub al-ʾaʿārīb*. Ed. Muḥammad Muḥyī d-Dīn ʿAbd al-Ḥamīd. 2 vols. Beirut: al-Maktaba al-ʿAṣriyya, 1991.

Ibn Jinnī, *Lumaʿ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *Kitāb al-lumaʿ fī n-naḥw*. Ed. Hadi M. Kechrida. Uppsala: Almqvist and Wiksell, 1976.

Ibn Wallād, *Intiṣār* = ʾAbū l-ʿAbbās ʾAḥmad ibn Muḥammad Ibn Wallād, *Kitāb al-intiṣār*. Ed. Monique Bernards. Leiden: E.J. Brill, 1997.

Ibn Yaʿīš, *Šarḥ* = Muwaffaq ad-Dīn Yaʿīš Ibn Yaʿīš, *Šarḥ al-Mufaṣṣal*. 10 vols. in 5. Cairo: ʾIdārat aṭ-Ṭibāʿa al-Munīriyya, n.d. [*Fahāris Šarḥ al-Mufaṣṣal*. Damascus: Majmaʿ al-Luġa al-ʿArabiyya, 1990.]

Mubarrad, *Muqtaḍab* = ʾAbū l-ʿAbbās Muḥammad ibn Yazīd al-Mubarrad, *Kitāb al-muqtaḍab*. Ed. Muḥammad ʿAbd al-Xāliq ʿUḍayma. 4 vols. Cairo: Maṭābiʿ al-ʾAhrām, 1979.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *Kitāb Sībawayhi*. Ed. Hartwig Derenbourg. 2 vols. Paris: Imprimerie Nationale, 1881. (Repr., Hildesheim and New York: G. Olms, 1970.)

Zamaxšarī, *Mufaṣṣal* = ʾAbū l-Qāsim Maḥmūd ibn ʿUmar az-Zamaxšarī, *al-Mufaṣṣal fī ʿilm al-ʿArabiyya*. Beirut: Dār al-Jīl, n.d.

SECONDARY SOURCES

Badawi, Elsaid, Michael G. Carter, and Adrian Gully. 2004. *Modern Written Arabic: A comprehensive grammar*. London and New York: Routledge.

Bateson, Mary Catherine. 1967. *Arabic language handbook*. Washington, D.C.: Center for Applied Linguistics.

Beeston, A.F.L. 1968. *Written Arabic: An approach to the basic structures*. Cambridge: Cambridge University Press. (Repr., 1975.)

Blachère, Régis. 1985. *Éléments de l'arabe classique*. 4th ed. Paris: Maisonneuve and Larose.

—— and Maurice Gaudefroy-Demombynes. 1975. *Grammaire de l'arabe classique*. 3rd ed. Paris: Maisonneuve and Larose.

Blau, Joshua. 1981. *The emergence and linguistic background of Judaeo-Arabic: A study of the origins of Middle Arabic*. 2nd ed. Jerusalem: Ben Zvi Institute.

Cachia, Pierre. 1973. *The monitor: A dictionary of Arabic grammatical terms*. Beirut: Librairie du Liban and London: Longman.

Cantarino, Vicente. 1975. *Syntax of Modern Arabic prose. II. The expanded sentence. III. The compound sentence*. Bloomington and London: Indiana University Press.

Carter, Michael G. 1972. "'Twenty dirhams' in the *Kitāb* of Sībawayhi". *Bulletin of the School of Oriental and African Studies* 35.485–496.

Dahdah, Antoine el-. 1988. *A dictionary of terms of declension and structure in universal Arabic grammar*. Rev. ed. Martin J. McDermott and Elias Matar. Beirut: Librairie du Liban.

Gully, Adrian. 1995. *Grammar and semantics in Medieval Arabic: A study of Ibn-Hisham's 'Mughni l-Labīb'*. Richmond, Surrey: Curzon Press.

Hopkins, Simon. 1984. *Studies in the grammar of early arabic based upon papyri datable to before 300 A.H./912 A.D.* Oxford: Oxford University Press.

Howell, Mortimer Sloper. 1883. *A grammar of the Classical Arabic language. Translated and compiled from the works of the most approved native or naturalised authors by Mortimer Sloper Howell.*

(2nd ed., 1986, with a new preface by Satkari Mukhopadhaya, 4 vols. in 7 parts. New Delhi: Gian Publishing House; repr. 1990.)

Kinberg, Naphtali. 1996. *A lexicon of al-Farrāʾ's terminology in his* Qurʾān *commentary, with full definitions, English summaries, and extensive citations*. Leiden: E.J. Brill.

Kluge, Eva-Maria. 1999. *Syntaktische Bestimmungen im Arabischen: Das vierte Kapitel aus dem* "Muġnī l-labīb ʿan kutub al-aʿārīb" *des Ibn Hišām al-Anṣārī, Übersetzung und Kommentar*. Leiden: Nederlands Instituut voor het Nabije Oosten.

Lane, Edward William. 1863–1893. *An Arabic-English lexicon derived from the best and most copious Eastern sources*. 8 vols. London and Edinburgh: Williams and Norgate.

Levin, Aryeh. 1979. "Sībawayhi's view of the syntactic structure of *kāna wa-ʾaxawātuhā*". *Jerusalem Studies in Arabic and Islam* 1.185–213.

Owens, Jonathan. 1990. *Early Arabic grammatical theory: Heterogeneity and standardization*. Amsterdam and Philadelphia: J. Benjamins.

Reckendorf, Hermann. 1977. *Arabische Syntax*. 2nd ed. Heidelberg: C. Winter.

Talmon, Rafael. 1993. "The term *qalb* and its significance for the study of the history of early Arabic grammar". *Zeitschrift für Geschichte der arabisch-islamischen Wissenschaften* 8.71–113.

——. 2003. *Eighth-century Iraqi grammar: A critical exploration of pre-H̱alīlian Arabic linguistics*. Winona Lake, Ind.: Eisenbrauns.

Wehr, Hans. 1994. *A dictionary of Modern Written Arabic (Arabic-English)*. Ed. J. Milton Cowan. 4th ed. Ithaca, N.Y.: Spoken Language Services.

Wright, William. 1859–1862. *A grammar of the Arabic language, translated from the German of Caspari and edited with numerous additions and corrections*. 3rd ed., rev. W. Robertson Smith and M.J. de Goeje. 2 vols. in 1 part. Cambridge: Cambridge University Press. (Repr., 1967.)

Monique Bernards (University of Groningen)

Hamza

The primary reference of the term *hamza* is the letter ء. However, it is also used, especially in the Arabic grammatical tradition, to refer to a specific speech sound, the glottal stop. Since the letter *hamza* is not always pronounced as a glottal stop, and since some letters other than the *hamza* are sometimes pronounced as the glottal stop, these two referents of the word *hamza* should be carefully distinguished.

1. Orthography

The *hamza* is usually not written on its own but is supported by one of the consonants ى, و or ا. The supporting consonant is known as the *kursī* 'chair' of the *hamza*. The rules about how to write the *hamza* are complicated, but see Wright (1967, Paragraphs 15–22, 131–139) or Mitchell (1953:20–21, 39–40, 79–81) for a good discussion. The following are the most important rules: The default *kursī* is the ا. However, when preceded or followed by the *u* or *i*, the و or ى sometimes serves as *kursī*. There are also contexts in which the *hamza* is not supported by a *kursī* – when preceded by a long vowel or a closed syllable, especially in word-final position. Lastly, there are instances in which the *hamza*, while pronounced, is not written. When it occurs in phrase-initial position, it is customary to write only the vowel associated with the *hamza* directly over or under the ا that serves as the *kursī*.

2. Place of articulation

Consonants are produced by obstructing airflow somewhere in the speech tract. They are classified into places of articulation according to the locus of the narrowest constriction of airflow. Therefore, one question about the *hamza* that needs to be answered is what its place of articulation is. The early Arab grammarians disagreed about the constriction in the articulation of the *hamza*. One of the earliest statements of Arabic phonetics is the introduction that al-Xalīl (d. 175/793) prefaced to his dictionary *Kitāb al-ʿayn* (Sara 1991, 1993). In this treatise, al-Xalīl divided the speech tract into eight regions, each of which was known as a *ḥayyiz* 'locale'. With the exception of the *wāw, yāʾ, ʾalif*, and *hamza*, al-Xalīl assigned each of the Arabic consonants to one of these locales, based on the locus of the narrowest constriction in the speech tract during the articulation of the consonant. However, *wāw, yāʾ, ʾalif*, and *hamza* were classified by him as *hawāʾ* 'air/cavity' sounds (al-Xalīl, *ʿAyn* I, 58.13; Roman 1983:I, 216). This means that he considered the *hamza*, like the glides *wāw* and *yāʾ*, as a consonant with no clear constriction, i.e. no clearly identifiable place of articulation. Sībawayhi (d. 180/796), al-Xalīl's student and the most famous of the Arab grammarians, espoused a different view of the *hamza* in his grammar, *al-Kitāb* (III, 541–556; cf. Al-Nassir 1993:81–90; Sara 1993, 1996). He assigns the *hamza* to a region of the speech tract that he calls the *ḥalq* 'throat'. He then divides the *ḥalq* into three subregions, and assigns *hamza* and *hāʾ* to the subregion that is the furthest back, most probably corresponding to the glottis. Sībawayhi did consider the *hamza* to be a true consonant, formed with a glottal constriction. He also describes an allophonic variant of the *hamza* (*hamza bayna bayna* 'intermediate *hamza*'), which according to Al-Nassir (1993:81–82) was a spirantized glottal consonant (cf. Roman 1983:I, 322–348). Sībawayhi's influence in the Arabic grammatical tradition was significant enough that his views on the *hamza* were accepted by his successors.

In the modern linguistic tradition, it is generally accepted that the *hamza*, together with the other so-called gutturals (see McCarthy 1991, 1994), is formed with a constriction in the postvelar region of the speech tract. However, there is disagreement about where in this region the constriction is made and about which articulators are responsible for the constriction.

One view is that the constriction is made at the glottis (Kästner 1981:47; McCarthy 1991:78, 1994:193). This view is based on the results of Klatt and Stevens (1969) who found that the *hamza* does not influence the formant structures of adjacent vowels. Any constriction above the glottis will influence formant structures of adjacent vowels, and a constriction in the postvelar region will in particular raise F_1 and depress F_2 (Stevens and House 1955).

More recently Zawaydeh (1999) and Shahin (1997, 2002) found evidence that the laryngeal consonants (*hamza* and *hāʾ*), do influence the formant structures of adjacent vowels. Both of them found that the laryngeals raise F_1, and Shahin also found that laryngeals depress F_2 in non-low vowels. Zawaydeh and Shahin ascribe the findings of Klatt and Stevens (1969) to the fact that they investigated only vowels in stem-final position – a position that is phonologically immune to co-articulation with adjacent consonants.

The fact that the *hamza* does influence the formant structure of adjacent vowels implies that it must be pronounced with a constriction somewhere between the velum and glottis. The question then becomes exactly where and by which articulators this constriction is made. Shahin assumes that a constriction is made in the pharynx by retracting the tongue root, basing this assumption on analogy with the pharyngeals *ḥāʾ* and *ʿayn*. The pharyngeals are pronounced with a retracted tongue root, and they are associated with raised F_1 and depressed F_2 (Al-Ani 1970:60, 63–64; Butcher and Ahmad 1987). Since the *hamza* shares these acoustic properties with the pharyngeals, Shahin assumes that it also shares the articulatory properties with these sounds.

However, Zawaydeh (1999) performed an experiment in which she inserted a fiberoptic endoscope through the nasal passage into the pharynx, enabling her to observe the movement of the pharyngeal articulators during speech. Zawaydeh found evidence for narrowing of the pharynx (through tongue root retraction) during the pronunciation of the pharyngeals, emphatics, and uvulars. However, she found no pharyngeal narrowing for the laryngeals *hāʾ* and *hamza*. She therefore hypothesizes that the constriction for the *hamza* is made closer to the glottis, in the area of the aryepiglottic folds, just above the glottis. There are Caucasian languages that distinguish two kinds of laryngeal consonants, pure glottal consonants and consonants with a constriction at the aryepiglottic folds (Nolan 1995). It is therefore possible that the Arabic laryngeals, including the *hamza*, have a constriction above the glottis at the aryepiglottic folds. However, this issue still needs to be settled by careful physiological investigation.

3. CONTEXTUALLY DETERMINED VARIATION

The *hamza* is subject to contextual variation in the degree of obstruction associated with its pronunciation. This is not unexpected – the glottal stop is an unstable consonant and is subject to weakening (widening of the closure) in many languages (Ladefoged and Maddieson 1996:75). There is likely to be dialectal variation concerning this aspect of the *hamza*'s pronunciation. Unfortunately, few of the grammars of colloquial Arabic dialects report in detail on this aspect of pronunciation (but see Al-Ani 1970:60–62; Kästner 1981:46–48; and Watson 2002:18, for some discussion). The discussion below is based primarily on what we know about Classical Arabic from Sībawayhi's *Kitāb*.

Sībawayhi devotes a separate section of his grammar to the realization of the *hamza* (Al-Nassir 1993:81-90). In general, the *hamza* is pronounced as a stop only in utterance-initial position (*ʾabun* [ʔabun] ‘father’). In other contexts, it is subject to various degrees of weakening. Intervocalically, there are three ways in which it can be realized: as a glottal fricative [h] (*saʾala* [sahala] ‘he asked’) or as one of the glides [w, j] (*suʾālun* [suwaːlun] ‘question’). The fricative is the default value in this context, with the glides being more likely to occur if one of the flanking vowels is *u* or *i*. When the *hamza* is preceded or followed by another consonant, it is usually deleted (*raʾsun* [raːsun] ‘head’). In Classical Arabic, words did not typically end in consonants, so that the *hamza* could not occur in word-final position. However, modern colloquial Arabic has lost many of the word-final vowels of Classical Arabic, so that words often end in consonants in colloquial Arabic. Al-Ani (1970:62) and Kästner (1981:48) claim that *hamza* in word-final position is pronounced as a stop (i.e. with total glottal closure), either with or without release of the glottal closure.

4. OTHER PHONOLOGICAL CHARACTERISTICS

The *hamza* is a member of the natural class of guttural sounds (McCarthy 1991, 1994). McCarthy has detailed a cluster of phonological phenomena associated with this class of sounds in Semitic. The two that are most relevant for Arabic are that with very few exceptions, consonantal roots are allowed to contain only one guttural consonant – i.e. there exists no Arabic root that contains both a *hamza* and another guttural; and that vowels that occur in the context of the gutturals are often lowered. For instance, in Classical Arabic the thematic vowel of the Form I imperfect of verbs with a guttural as second or third radical is always *a* rather than the more typical *i* or *u*, e.g. *yasʾalu* and not **yasʾilu* or **yasʾulu* (*s-ʾ-l* ‘to ask’). Vowel lowering associated with gutturals in general and with the *hamza* in particular is also observed in modern colloquial Arabic – see, for instance, Zawaydeh (1999) on Jordanian Arabic, and Shahin (2002) and Younes (1982) on Palestinian Arabic (→ vowel backing).

Traditional Arabic grammars distinguish two kinds of *hamza, hamzat al-qaṭʾ* ‘the *hamza* of separation’ and *hamzat al-waṣl* ‘the *hamza* of connection’ (Haywood and Nahmad 1965:10–11, 114 etc.; Wright 1967, Paragraphs 18–21). The *hamzat al-waṣl* is found only word-initially. It occurs in a few nouns (*(ʾ)ibnun* ‘son’), in the definite article (*(ʾ)al* ‘the’), and in some verbal forms – Form I imperatives (*(ʾ)uktub* ‘write!’), and all forms from VII onward that do not take an imperfect or participial prefix (*(ʾ)inkasara* ‘it broke’). When any of these words occurs in non-phrase-initial position, the *hamzat al-waṣl*, together with its accompanying vowel, is deleted (compare *(ʾ)inkasara* ‘it broke’ with *wa-nkasara* ‘and it broke’). The *hamzat al-qaṭˁ*, on the other hand, can occur word-initially (*ʾabun* ‘father’) and word-medially (*suʾila* ‘it is asked’), and is never deleted. A word like *ʾabun* with the *hamzat al-qatˁ* is therefore pronounced with the *hamza* when preceded by another word, while *(ʾ)ibnun* with the *hamzat al-waṣl* is pronounced without the *hamza* in this context (*li-ʾabin* ‘for a father’, but *li-bnin* ‘for a son’). Coetzee (1998) and Gadoua (2000) independently argue against this traditional view. They argue that the *hamzat al-waṣl* is not part of the underlying form of any word. Words that are traditionally assumed to start on the *hamzat al-waṣl* should rather be seen as starting on a consonant cluster. Classical Arabic, and many modern colloquial dialects as well, do not allow tautosyllabic consonant clusters. When any of these words occur phrase-initially, a syllable consisting of a *hamza* and a vowel is inserted in order to prevent the word from starting on a consonant cluster (/bnun/ → [ʔib.nun]). However, when such a word is preceded by another word, resyllabification across the word boundary resolves the consonant cluster (/li + bnin/ → [lib.nin]). Under the traditional view it is not possible to explain adequately why *hamza* deletes phrase-initially in some words but not in others. Under the alternative view, this is no longer a problem. The *hamza* is never deleted but is rather inserted phrase-initially just in those cases where it is necessary to resolve a tautosyllabic consonant cluster.

5. DIALECTOLOGY

Modern colloquial dialects have not all preserved the glottal stop in their phonemic inventories (Fischer and Jastrow 1980:52–53; Kaye and Rosenhouse 1997:277). According to Cadora (1992:13), it is mostly the Bedouin dialects that have retained the glottal stop, while urban dialects have lost it. However, this is probably an oversimplification, especially since the distinction between Bedouin and urban dialects is becoming more blurred with increased urbanization. The following dialects have all preserved at least some glottal stops: Cairene (Watson 2002:20), Iraqi (Al-Ani 1970:29), Jordanian (Zawaydeh 1999), Lebanese (Obrecht 1968:19), Maltese (Borg 1997b:249), Palestinian (Shahin 2002), and Ṣanʿānī Arabic (Watson 2002:19). Dialects that have lost the glottal stop include Cypriot (Borg 1997a:222) and Moroccan Arabic (Heath 1997:207). For the pharyngealized glottal [ʡ] in the speech of Sunni men in Beirut → Beirut Arabic.

There are also dialects in which the uvular stop *qāf* of Classical Arabic has been replaced by the glottal stop. Cadora (1992) claims that this has happened mostly in the urban dialects. See also Fischer and Jastrow (1980:52) and Watson (2002:17), who state that this change affected mostly the larger cities around the Mediterranean – Cairo, Jerusalem, Damascus, and Beirut.

Bibliographical references

Primary sources

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. Ed. ʿAbd as-Salām Muḥammad Hārūn. 5 vols. Cairo: al-Hayʾa al-Miṣriyya al-ʿĀmma li-l-Kitāb, 1966–1977.

Xalīl, *ʿAyn* = ʾAbū ʿAbd ar-Raḥmān al-Xalīl ibn ʾAḥmad al-Farāhīdī, *Kitāb al-ʿayn*. Ed. Mahdī al-Maxzūmī and ʾIbrāhīm as-Sāmarrāʾī. 8 vols. Beirut: Muʾassasat al-ʾAʿlamī li-l-Maṭbūʿāt, 1988.

Secondary sources

Al-Ani, Salman H. 1970. *Arabic phonology: An acoustical and physiological investigation*. The Hague: Mouton.

Al-Nassir, Abd al-Munʿim ʿAbd al-Munīr. 1993. *Sibawayh the phonologist: A critical study of the phonetic and phonological theory of Sibawayh as presented in his treatise* Al-Kitab. London and New York: Kegan Paul International.

Borg. Alexander. 1997a. “Cypriot Arabic phonology”. Kaye and Daniels (1997:219–244).

——. 1997b. “Maltese phonology”. Kaye and Daniels (1997:245–285).

Butcher, Andrew and Kusay Ahmad. 1987. “Some acoustic and aerodynamic characteristics of pharyngeal consonants in Iraqi Arabic”. *Phonetica* 44.156–172.

Cadora, Frederic J. 1992. *Bedouin, village and urban Arabic: An ecolinguistic study*. Leiden: E.J. Brill.

Coetzee, Andries W. 1998. “The phonology of the two *hamza*’s of Qurʾānic Arabic”. *Theoretical Linguistics* 24.219–244.

Fischer, Wolfdietrich and Otto Jastrow. 1980. *Handbuch der arabischen Dialekte*. Wiesbaden: O. Harrassowitz.

Gadoua, Abdulhamid H. 2000. “Consonant clusters in Quranic Arabic”. *Cahiers Linguistiques d’Ottawa* 28.59–85.

Haywood, John A. and Hayim M. Nahmad. 1965. *A new Arabic grammar of the written language*. London: Lund Humphries.

Heath, Jeffrey. 1997. “Moroccan Arabic phonology”. Kaye and Daniels (1997:205–217).

Kästner, Hartmut. 1981. *Phonetik und Phonologie des modernen Hocharabisch*. Leipzig: VEB Verlag Enzyklopädie.

Kaye, Alan S. 1997. “Arabic phonology”. Kaye and Daniels (1997:187–204).

Kaye, Alan S. and Peter T. Daniels. 1997. *Phonologies of Asia and Africa (including Caucasus)*, I. Winona Lake, Ind.: Eisenbrauns.

Kaye, Alan S. and Judith Rosenhouse. 1997. “Arabic dialects and Maltese”. *The Semitic languages*, ed. Robert Hetzron, 263–311. New York: Routledge.

Klatt, Dennis H. and Kenneth N. Stevens. 1969. “Pharyngeal consonants”. *MIT Research Laboratory of Electronics Quarterly Progress Report* 93.208–216.

Ladefoged, Peter and Ian Maddieson. 1996. *The sounds of the world’s languages*. Oxford: Blackwell.

McCarthy, John J. 1991. “Semitic gutturals and distinctive feature theory”. *Perspectives on Arabic linguistics*, III, ed. Bernard Comrie and Mushira Eid, 63–91. Amsterdam and Philadelphia: J. Benjamins.

——. 1994. “The phonetics and phonology of Semitic pharyngeals”. *Phonological structure and phonetic form: Papers in laboratory phonology*, III, ed. Patricia A. Keating, 191–233. Cambridge: Cambridge University Press.

Mitchell, Terence F. 1953. *Writing Arabic: A practical introduction to* ruqʿah *script*. London and New York: Oxford University Press.

Nolan, Francis. 1995. “The role of the jaw – active or passive? Comments on Lee”. *Phonology and phonetic evidence: Papers in laboratory evidence*, IV, ed. Bruce Connell and Amalia Arvanti, 361–367. Cambridge: Cambridge University Press.

Obrecht, Dean H. 1968. *Effects of the second formant on the perception of velarization consonants in Arabic*. The Hague: Mouton.

Roman, André. 1983. *Etude de la phonologie et de la morphologie de la koine arabe*. 2 vols. Aix-en-Provence: Université de Provence.

Sara, Solomon I. 1991. “al-Ḫalil: The first Arabic phonologist”. *International Journal of Islamic and Arabic Studies* 8.1–57.

——. 1993. “The beginning of phonological terminology in Arabic”. *Proceedings of the Colloquium on Arabic Lexicology and Lexicography Budapest, 1–7 September, 1993*, ed. Kinga Dévényi, Tamás Iványi, and Avihai Shivtiel, 181–194. Budapest: Eötvös Loránd University and Csoma de Kőrös Society.

——. 1996. “Changes in the phonological perception of Classical Arabic”. *Proceedings of the Colloquium on Logos, Mythos in the Middle East and North Africa. I. Linguistics and Literature, Budapest, 18–22 September, 1995*, ed. Kinga Dévényi and Tamás Iványi, 3–14. Budapest: Eötvös Loránd University and Csoma de Kőrös Society.

Shahin, Kimary N. 1997. "Acoustics of pharyngealization vs. uvularization harmony". *Perspectives on Arabic linguistics*, X, ed. Mushira Eid and Robert Ratcliffe, 215–237. Amsterdam and Philadelphia: J. Benjamins.

——. 2002. *Postvelar harmony*. Amsterdam and Philadelphia: J. Benjamins.

Stevens, K. and A. House. 1955. "Developments of a quantitative description of vowel articulation". *Journal of the Acoustical Society of America* 27.484–494.

Watson, Janet C.E. 2002. *The phonology and morphology of Arabic*. Oxford: Oxford University Press.

Wright, William. 1967. *A grammar of the Arabic language*. 3rd ed. Repr. Cambridge: Cambridge University Press.

Younes, Munther Abdullatif. 1982. *Problems in the segmental phonology of Palestinian Arabic*. Ph.D. diss., University of Texas at Austin.

Zawaydeh, Bushra Adnan. 1999. *The phonetics and phonology of gutturals in Arabic*. Ph.D. diss., Indiana University.

Andries W. Coetzee (University of Michigan)

Ḥaraka

In Arabic linguistic terminology, the term *ḥaraka* lit. 'movement' indicates a vowel or, more precisely, the phonemes that are known in the Western tradition as 'short vowels'. It contrasts with the term → *ḥarf* 'consonant'. Sībawayhi distinguishes three vowels, /a/, /u/, and /i/, called *fatḥa, ḍamma*, and *kasra*, respectively (cf. Al-Nassir 1993:28–35). The vowels are not phonemic entities in themselves; their sole function is to make the pronunciation of the consonants possible, a statement attributed by Sībawayhi (d. 177/793?) to his teacher al-Xalīl (d. 175/791): "Al-Xalīl asserted that the *fatḥa*, the *kasra* and the *ḍamma* are additions; they are attached to the consonants so that these can be pronounced" (*zaʿama l-Xalīl ʾanna l-fatḥa wa-l-kasra wa-ḍ-ḍamma zawāʾid wa-hunna yalḥaqna l-ḥarf li-yūṣala ʾilā t-takallum bihi; Kitāb* II, 315.2–3; Troupeau 1989).

The names of the vowels are explained by the Arabic tradition in articulatory terms. The legendary founder of the linguistic tradition, ʾAbū l-ʾAswad ad-Duʾalī, is said to have instructed a scribe as follows:

> When you see me opening my mouth, write a dot above the letter, and when you see me contracting my mouth, write a dot within the letter, and when you see me folding my mouth, write the dot beneath the letter (*ʾiḏā raʾaytanī qad fataḥtu famī bi-l-ḥarf fa-nquṭ nuqṭa ʿalā ʾaʿlāhu wa-ʾiḏā ḍamamtu famī fa-nquṭ nuqṭa bayna yaday al-ḥarf wa-ʾiḏā kasartu famī fa-jʿal an-nuqṭa taḥta l-ḥarf*). (ʾAbū ṭ-Ṭayyib, *Marātib* 10–11)

The etymology given here for the names of the vowels is probably spurious, but it seems to be connected with the phonetic terminology for the vowels in the Syriac linguistic tradition, in which the short vowels /a/, /i/, and /u/ were called *pĕṯāḥā* 'opening', *ḥĕḇāṣā* 'pressure, pushing', and *ʿĕṣāṣā* 'contraction' (Versteegh 1993:28–32). Apparently, the Syriac names for the vowels are the source for the Arabic terms, which in itself is quite plausible in view of the early contacts between Arabic and Syriac grammarians (Revell 1975).

Originally, the terms for the vowels and those for the vocalic declensional endings were confused. In the earliest Qurʾānic commentaries, terms like *xafḍ, kasr*, and *jarr* are used indiscriminately for all instances of the vowel /i/, both within the word and as an ending, and regardless of whether or not they are declensional (Versteegh 1993:125–130); likewise, *naṣb* and *fatḥ* are used for the vowel /a/, and *ḍamm* and *rafʿ* for the vowel /u/. A word like *muxliṣīna* 'being sincere [acc.]', for instance, is distinguished from the passive form *muxlaṣūna* with the expression *bi-xafḍ al-lām*. This practice was continued by most Kufan grammarians. In al-Farrāʾ, the confusion is no longer conceptual but purely terminological; he does distinguish between declensional and nondeclensional vowels but calls the latter indiscriminately *ḍamma/rafʿ, fatḥa/naṣb*, and *kasra/xafḍ* (Owens 1990:159). One of the innovations in Sībawayhi's grammatical system was probably the introduction of a distinction between the two sets. This distinction must be attributed to Sībawayhi alone because in the *Kitāb al-ʿayn* that is attributed to al-Xalīl (→ lexicography; → Classical Arabic), the names of case endings are still used for non-final vowels, and occasionally the vowel names are used for case endings (Talmon 1997:194–197).

Right from the beginning of the *Kitāb* (3.1–5), Sībawayhi takes great care to distinguish between the vowels as phonological units and the vowels that constitute the declensional endings *-a, -u*, and *-i* (→ *ʾiʿrāb*). He states that there are eight different endings (*majārin*) in Arabic:

> These eight endings are combined in four categories in speech: *naṣb* and *fatḥ* are one category in speech, *jarr* and *kasr* are one category, likewise *rafʿ* and *ḍamm*, and *jazm* and *waqf*. I mentioned eight endings in order to distinguish between [words] receiving one of these endings because they are affected by a governor…and [words] whose final consonant always stays with the same [vowel] (*wa-hāḏihi l-majārī ṯ-ṯamāniya yujmiʿuhunna fī l-lafḏ ʾarbaʿ ʾaḍrub fa-n-naṣb wa-l-fatḥ fī l-lafḏ ḍarb wāḥid wa-l-jarr wa-l-kasr ḍarb wāḥid wa-kaḏālika r-rafʿ wa-ḍ-ḍamm wa-l-jazm wa-l-waqf wa-ʾinnamā ḏakartu laka ṯamāniya majārin li-ʾafruqa bayna mā yadxuluhu ḍarb min hāḏihi l-ʾarbaʿa li-mā yuḥdiṯu fīhi l-ʿāmil…wa-bayna mā yubnā ʿalayhā l-ḥarf bināʾan lā yazūlu ʿanhu*).

In Sībawayhi's phonological analysis, phonetically long vowels are analyzed as combinations of a vowel with one of the three → glides (*ḥurūf al-līn wa-l-madd* or *ḥurūf muʿtalla*; → *ʿilla*), /w/, /y/, and /”/ (*fa-l-fatḥa min al-ʾalif wa-l-kasra min al-yāʾ wa-ḍ-ḍamma min al-wāw; Kitāb* II, 315.4). According to this analysis [uː] is /uw/, [iː] is /iy/, and [aː] is /a"/. The glide that is combined with the vowel /a/ (transcribed here as /"/) is called *ʾalif*, an abstract element that does not surface phonetically and only serves to preserve the transparency of the triradical structure on the underlying phonological level, for instance when comparing *ʾaswadu*, pl. *sūdun* 'black' with *ʾaḥmaru*, pl. *ḥumrun* 'red', whose relationship becomes much clearer when they are compared on the underlying level, /suwdun/ and /ḥumrun/ (cf. Bohas 1985; Bohas and Guillaume 1984:241–267).

The phonetic definition of the vowels is derived from that of the glides of which they are a part (*Kitāb* II, 285.12ff.). The three glides are defined as *ġayr mahmūsāt* 'not voiceless', and they are called *ḥurūf al-līn wa-l-madd* 'consonants of softness and lengthening'. Their place of articulation is the widest of all consonants for the passage of the air (*wa-maxārijuhā muttasiʿa li-hawāʾ aṣ-ṣawt wa-laysa šayʾ min al-ḥurūf ʾawsaʿ maxārija minhā wa-lā ʾamadd li-ṣ-ṣawt*). The interpretation of long vowels as combinations of (short) vowels and glides implies that length as a contrastive feature in vowels does not play any role in Sībawayhi's analysis. Cases that are described in Western grammars of Arabic as a shortening of a long vowel because of the constraint against long vowels in closed syllables, for instance in *fī l-kitābi* [filkitaːbi], are analyzed as the deletion of a glide, i.e. /fiylkita"bi/ ⇒ /filkita"bi/.

The changes in the combinations of vowels and glides were explained by the grammarians within the framework of an opposition in terms of heavy (*ṯaqīl*) and light (*xafīf*). They assigned to each sound a certain position on a scale that went from the lightest element to the heaviest. Certain constraints in phonological structure were then explained as measures to avoid excessive heaviness (*istixfāf*), for instance when a word contains the combination /a-"-a/, which is therefore changed into /a"/, realized as *ā* (→ *ʿilla*).

Vowels are not always realized the same way. Sībawayhi attributes these phonetic differences to the preceding consonant (cf. Schaade 1911:23). The main modifications that affect vowels are → *ʾimāla* and → *tafxīm* (Al-Nassir 1993:91–107). In certain contexts, the sequence /a"/ is said to contain an *ʾalif al-ʾimāla*, which then spreads to the short *a* preceding it. In his view, this is part of a process of → *iddiġām* in which the *ʾalif* is made similar to a *yāʾ*, and as a result the short *a* is made similar to a short *i* (cf. Ibn Jinnī, *Xaṣāʾiṣ* II, 141). In combinations of a consonant with a following short *a*, it is again the consonant that is said to undergo *ʾimāla* (e.g. *Kitāb* II, 270.19: when you pronounce *ʿamrin*, you apply *ʾimāla* to the /ʿ/). Likewise, *tafxīm* is regarded as a feature of the consonant, which spreads to the vowel.

Later grammarians modified this analysis by stating that the difference resided in the phonetic realization of the short vowels. According to Ibn Jinnī (*Xaṣāʾiṣ* III, 120–121; cf. Méhiri 1973:227ff.), for instance, between each pair of vowels there is another one, so that the total number of vowels is actually six. Between /a/ and /i/, there is the *fatḥa* that is followed by an *ʾalif* with *ʾimāla*, whose pronunciation is between [a] and [i], i.e. [ɛ] or [æ], called elsewhere (*Sirr* I, 52.4) *al-fatḥa al-mašūba bi-l-kasra* '*a* mixed with *i*'. Between /a/ and /u/, there is the *fatḥa* followed by an *ʾalif at-tafxīm*, which is probably pronounced [ɔ]. Between /i/ and /u/, there is the *kasra* in words like *qīla*, which is pronounced with *ʾišmām* of the *kasra*, i.e. probably as [y]. In this analysis, the different realizations are seen as allophones of the three phonological vowels.

In later analyses, the views on the relationship between vowels and consonants and that between vowels and glides also changed. In the *Kitāb al-ʿayn* the glides were seen as the product of the (short) vowels (*al-wāw xuliqat min aḍ-ḍamma* 'the *w* is created from the *u*'; VIII, 195; cf. Talmon 1997:134), but there is no explanation of the process by which this 'creation' takes place. Apparently, some grammarians believed that phonetically long vowels are the result of a 'lengthening' (*ʾišbāʿ*) of the short vowels. This term is derived from metrical theory, where it is used to indicate the lengthening of a vowel for metrical reasons. According to Ibn al-ʾAnbārī (*ʾInṣāf* 6.12–12.25), the grammarian al-Māzinī (d. 249/863) applied it to the declension of the so-called 'six nouns' (*al-ʾasmāʾ as-sitta*) *ʾabun* 'father', *ʾaxun* 'brother', *ḏū* 'possessor of', *ḥamun* 'father-in-law', *hanun* 'thing', *fū-* 'mouth', whose case endings have long vowels in the construct state (*ʾabū, ʾabī, ʾabā*, etc.). The same theory is mentioned by Ibn Jinnī (*Xaṣāʾiṣ* III, 136.1) who probably also quotes from al-Māzinī (about al-Māzinī's theory of morphology, see Versteegh 1985).

The *ʾišbāʿ* theory applied only to a limited class of nouns, but it may be connected with a general theory that made the long vowels in the sound dual and masculine plural endings (*-āni; -ūna/-īna*) 'substitutes of the vowels' (*ʾabdāl al-ḥarakāt*). This theory is attributed by az-Zajjājī (*ʾĪḍāḥ* 141.10ff.; cf. Versteegh 1995:254) to the Kufan grammarian Ṯaʿlab (d. 291/903). Az-Zajjājī states that Ṯaʿlab held that "the *ʾalif* in *az-zaydāni* is the equivalent of two *ḍamma*s… and the *wāw* in *az-zaydūna* is the equivalent of three *ḍamma*s" (*al-ʾalif fī az-zaydāni badal min ḍammatayni…wa-l-wāw fī az-zaydūna badal min ṯalāṯ ḍammāt*).

Throughout the Arabic tradition there are traces of such a different concept of length in vowels. This concept may be connected in its turn with the theory about the vowels that is mentioned by al-Xwārizmī (*Mafātīḥ* 46.3–10) in connection with Greek logic. He says:

> According to the Greek logicians the nominative is a defective *w*, and so is the *u* and the related sounds we have mentioned; the *i* and related sounds are for them a defective *y*, and the *a* and related sounds are for them a defective *ʾalif*. You could also call the glide *w* a lengthened *u*, and the glide *y* a lengthened *i*, and the glide *ʾalif* a lengthened *a*" (*ar-rafʿ ʿinda ʾaṣḥāb al-manṭiq min al-Yūnāniyyīn wāw nāqiṣa wa-ka-ḏālika ḍ-ḍamm wa-ʾaxawātuhu l-maḏkūra wa-l-kasr wa-ʾaxawātuhu ʿindahum yāʾ nāqiṣa wa-l-fatḥ wa-ʾaxawātuhu ʿindahum ʾalif nāqiṣa wa-ʾin šiʾta qulta al-wāw al-mamdūda al-layyina ḍamma mušbaʿa wa-l-yāʾ al-mamdūda al-layyina kasra mušbaʿa wa-l-ʾalif al-mamdūda fatḥa mušbaʿa*).

The term *ʾišbāʿ* provides the link between this 'Greek' theory and the ideas of Ṯaʾlab and al-Māzinī mentioned above. In a different context, the same concept of short and long vowels in Arabic is found in Jābir ibn Ḥayyān (2nd/8th century; cf. Kraus 1942:244, n. 1).

In a different manner Ibn Jinnī reaches the same conclusion; he says that when you lengthen the vowel, the homorganic (*min jinsihā*) glide is produced, which is why "the *ḍamma* is called a small *wāw*, the *kasra* a small *yāʾ*, and the *fatḥa* a small *ʾalif*" (*Xaṣāʾiṣ* II, 121–124 *fī maṭl al-ḥarakāt*). The connection with Greek theories is also apparent in Ibn Sīnā's *Risāla* (85.2–6):

> The case of these three [the glides /w/, /y/, /"/] is difficult for me to understand. But I know for a fact that the lengthened voiced *ʾalif* takes up twice as much or more time than the *fatḥa* and that the *fatḥa* just takes up the minimum amount of time that is necessary to go from one consonant to another. The relationship between the voiced *w* and the *ḍamma* and between the voiced *y* and the *kasra* is the same (*ṯumma ʾamr hāḏihi ṯ-ṯalāṯa ʿalayya muškil wa-lākinnī ʾaʿlamu yaqīnan ʾanna l-ʾalif al-mamdūda al-muṣawwita taqaʿu fī ḍiʿf ʾaw ʾaḍʿāf zamān al-fatḥa wa-ʾanna l-fatḥa taqaʿu fī ʾaṣġar al-ʾazmina llatī yaṣiḥḥu fīhā l-intiqāl min ḥarf ʾilā ḥarf wa-ka-ḏālika nisbat al-wāw al-muṣawwita ʾilā ḍ-ḍamma wa-l-yāʾ al-muṣawwita ʾilā l-kasra*).

In a different recension of the *Risāla* (126.2–7; Semaan 1977:48), Ibn Sīnā even uses the terms *ṣuġrā* and *kubrā* to indicate short and long vowels. He avoids the term *ḥarakāt* altogether and combines the three glides and the three short vowels under the term *muṣawwitāt*.

In mainstream Arabic grammar, meanwhile, the focus was on the relationship between consonants and vowels. In a series of chapters in his *Xaṣāʾiṣ*, Ibn Jinnī deals with this issue, in particular the question of when the vowels are pronounced: before, after, or together with the consonant (*Xaṣāʾiṣ* III, 121–127; Mehiri 1973:219-225). According to Sībawayhi, the vowel is pronounced after the consonant. Ibn Jinnī discusses the alternative theory according to which the vowel precedes the consonant and demonstrates that this does not make sense, because in that case phenomena of assimilation could not be explained. In /miwza"n/ the change of the *w* to *y* is caused by the preceding *i*, resulting in /miyza"n/, which is realized as *mīzān*. If the *i* were to precede its consonant, it would not be contiguous to the *w* and could not cause its change.

According to some grammarians, the vowel and the consonant are pronounced simultaneously. Ibn Jinnī cites his teacher ʾAbū ʿAlī al-Fārisī (d. 377/987), who claimed that this is the only way to explain the pronunciation of the /n/: without a vowel it is produced through the nose, with a vowel through the mouth. If the vowel followed the consonant, even the /n/ with a vowel would have to be pronounced through the nose. Ibn Jinnī himself does not accept this argument, because sounds may affect preceding sounds (as, for instance, the assimilation of the /m/ in *ʿanbar*, which is pronounced as /m/ because it is affected by the following /b/). His own view is that vowels are pronounced after the consonant, and he adduces as one of his arguments the relationship between vowels and glides. The glides are lengthened vowels, therefore, the combination C*a*" cannot be analyzed as *a*C", because in that case the glide is no longer in contact with the vowel /a/ and therefore cannot surface as a lengthened vowel. This means that vowels indeed come after the preceding consonant.

A different division of sounds occurs in the tradition that is connected with Greek philosophy and logic, for instance in Fārābī's (d. 339/950) *Kitāb al-mūsīqā al-kabīr*. He divides the *ḥurūf* (here 'phonemes', rather than 'consonants') into *muṣawwit* 'voiced' and *ġayr muṣawwit* 'voiceless'; he states further that the *muṣawwit* can be divided into long and short ones and adds that the short ones are what the grammarians call *ḥarakāt*. Ibn Sīnā avoids the term *ḥarakāt* altogether, as seen above.

The Arabic concept of vowels was taken over by all linguistic traditions that borrowed the Arabic model for the description of their own language, for instance Turkic, Coptic, and Persian. As these languages possess more vocalic phonemes than Arabic does, there was a problem, which was solved in different ways. For Coptic grammarians it was relatively easy to find a solution because they used the Coptic alphabet to represent the sounds of Coptic. In their description of Coptic they called the seven vowels of Coptic *ʾaḥruf ṣawtiyya* or *ʾaḥruf nawāṭiq*, a direct translation of Greek *phōnēenta (stoicheía)* 'sounding (elements)' (Bauer 1972:147–148), thus avoiding the problem of having to assign them to the three *ḥarakāt* of the Arabic system.

The Turkic grammarians, beginning with ʾAbū Ḥayyān al-ʾAndalusī (d. 745/1344), who wrote the first grammar of a Turkic language, were confronted with a vowel system with a twofold opposition between front and back vowels and between rounded and unrounded vowels (cf. Ermers 1999:93–128), giving a total of eight vowels, five more than in the Arabic system. They solved this problem by assigning the vowel quality to the preceding consonant. Thus, for instance, the opposition between the back vowels /u, o, a, ı/ and the front vowels /ö, e, i, ü/ is equated with that between velarized (*mufaxxam*) and palatalized (*muraqqaq*) consonants, which implied that the Turkic phonemic inventory had to be enlarged with a number of velarized consonants compared to the Arabic inventory. The rounded/unrounded vowels were more difficult to accommodate in this system. According to Ermers (1999:121–128), some grammarians applied the label of velarization (*tafxīm*) to a velar consonant to indicate the vowel /o/, and to a palatal consonant to indicate /ö/. Likewise, palatalization of a palatal consonant indicated /ü/. In all these attempts, they followed the basic rule of Arabic grammar, according to which the quality of the vowel depends on the preceding consonant.

Bibliographical references

Primary sources

ʾAbū ṭ-Ṭayyib, *Marātib* = ʾAbū ṭ-Ṭayyib ʿAbd al-Wāḥid ibn ʿAlī, *Marātib an-naḥwiyyīn*. Ed. Muḥammad ʾAbū l-Faḍl ʾIbrāhīm. Cairo: Maktaba Nahḍa, 1955.

Fārābī, *Mūsīqā* = ʾAbū Naṣr Muḥammad ibn Muḥammad al-Fārābī, *Kitāb al-mūsīqā al-kabīr*. Ed. Ġattās ʿAbd al-Malik Xašaba. Cairo, 1967.

Ibn al-ʾAnbārī, *ʾInṣāf* = ʾAbū l-Barakāt ʿAbd ar-Raḥmān ibn Muḥammad al-ʾAnbārī, *Kitāb al-ʾinṣāf fī masāʾil al-xilāf bayna n-naḥwiyyīna l-Baṣriyyīn wa-l-Kūfiyyīn*. Ed. Gotthold Weil. Leiden: E.J. Brill, 1913.

Ibn Jinnī, *Sirr* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *Sirr aṣ-ṣināʿa*. Ed. Muḥammad as-Saqā, ʾIbrāhīm Muṣṭafā, Muḥammad az-Zafzāf, and ʿAbdallāh ʾAmīn. Cairo, 1954.

——, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. Ed. Muḥammad ʿAlī an-Najjār. 3 vols. Cairo, 1952–1956. (Repr., Beirut: Dār al-Hudā, n.d.)

Ibn Sīnā, *Risāla* = ʾAbū ʿAlī al-Ḥusayn ibn ʿAbdallāh Ibn Sīnā, *Risāla ʾasbāb ḥudūṯ al-ḥurūf*. Ed. Muḥammad Ḥassān aṭ-Ṭayyān and Yaḥyā Mīr ʿAlam. Damascus: Majmaʿ al-Luġa al-ʿArabiyya, 1983.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. 2 vols. Ed. Bulaq, 1316 A.H. (Repr., Baghdad: Al-Muthanna Library, n.d.)

Xalīl, *ʿAyn* = ʾAbū ʿAbd ar-Raḥmān al-Xalīl ibn ʾAḥmad al-Farāhīdī, *Kitāb al-ʿayn*. Ed. Mahdī al-Maxzūmī and ʾIbrāhīm as-Samarrāʾī. 8 vols. Beirut: Muʾassasat al-ʾAʿlamī li-l-Maṭbūʿāt, 1988.

Xwārizmī, *Mafātīḥ* = ʾAbū ʿAbdallāh Muḥammad ibn ʾAḥmad al-Xwārizmī, *Kitāb mafātīḥ al-ʿulūm*. Ed. Gerlof van Vloten. Leiden: E.J. Brill, 1895.

Zajjājī, *ʾĪḍāḥ* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān ibn ʾIsḥāq az-Zajjājī, *al-ʾĪḍāḥ fī ʿilal an-naḥw*. Ed. Māzin al-Mubārak. Cairo: Dār al-ʿUrūba, 1959.

Secondary sources

Al-Nassir, A.A. [ʿAbd al-Munʿim ʿAbd al-ʾAmīr an-Nāṣir]. 1993. *Sibawayh the phonologist: A critical study of the phonetic and phonological theory of Sibawayh as presented in his treatise* Al-Kitab. London and New York: Kegan Paul International.

Bauer, Gertrud. 1972. *Athanasius von Qus,* Qilādat at-taḥrīr fī ʿilm at-tafsīr*: Eine koptische Grammatik in arabischer Sprache aus dem 13./14/Jahrhundert.* Freiburg: K. Schwarz.

Bohas, Georges. 1985. "L'explication en phonologie arabe". *Zeitschrift für Arabische Linguistik* 15.45–52.

—— and Jean-Patrick Guillaume. 1984. *Etude des théories des grammairiens arabes.* I. *Morphologie et phonologie*. Damascus: Institut Français de Damas.

Ermers, Robert. 1999. *Arabic grammars of Turkic: The Arabic linguistic model applied to foreign languages and translation of ʾAbū Ḥayyān al-ʾAndalusī's* Kitāb al-ʾidrāk li-lisān al-ʾatrāk. Leiden: E.J. Brill.

Kraus, Paul. 1942. *Jābir ibn Ḥayyān, contribution à l'histoire des idées scientifiques dans l'Islam: Jābir et la science grecque.* Cairo. (Repr., Paris: Les Belles Lettres, 1986.)

Mehiri, Abdelkader. 1973. *Les théories grammaticales d'Ibn Jinnī.* Tunis: Université de Tunis.

Owens, Jonathan. 1990. *Early Arabic grammatical theory: Heterogeneity and standardization.* Amsterdam and Philadelphia: J. Benjamins.

Revell, E.J. 1975. "The diacritical dots and the development of the Arabic alphabet". *Journal of Semitic Studies* 2.178–190.

Schaade, A. 1911. *Sībawaihi's Lautlehre.* Leiden: E.J. Brill.

Semaan, Khalil I. 1977. *Arabic phonetics: Ibn Sīna's* Risālah on the points of articulation of the speech-sounds *translated from Medieval Arabic.* 2nd ed. Lahore: Sh. Muhammad Ashraf. (1st ed., 1963.)

Talmon, Rafael. 1997. Kitāb al-ʿayn *and its attribution to Ḫalīl b. Aḥmad.* Leiden: E.J. Brill.

Troupeau, Gérard. 1989. "Voyelles et semi-voyelles dans le *Kitāb* de Sībawayhi". *Jerusalem Studies in Arabic and Islam* 12.31–39.

Versteegh, Kees. 1985. "The development of argumentation in Arabic grammar: The declension of the dual and the plural". *Studies in the history of Arabic grammar*, II, ed. Michael G. Carter and Kees Versteegh, 152–173. Amsterdam: J. Benjamins.

——. 1993. *Arabic grammar and Qurʾānic exegesis in early Islam.* Leiden: E.J. Brill.

——. 1995. *The explanation of linguistic causes: Az-Zağğāğī's theory of grammar.* Amsterdam and Philadelphia: J. Benjamins.

Kees Versteegh (University of Nijmegen)

Ḥarf

The term *ḥarf* (pl. *ḥurūf, aḥruf*) 'part, particle, edge, end, boundary' is used in Arabic linguistic terminology to indicate (1) the final segment formed as a result of the linear segmentation of the Arabic word; (2) a component of the prosodic, morphological, and lexical pattern of a word; (3) any discrete unit of an Arabic text that has a linguistic function (word, morpheme); (5) a certain class of linguistic units; or (6) one of the parts of speech, against nouns and verbs.

As a polyvalent notion, the term *ḥarf* has no equivalent in the conceptual system of European linguistics (Frolov 1991:57). Its polysemy derives from the systematic use of the same term at different levels, not only within one scientific domain but also across a broad range of subjects, a characteristic typical of the entire spectrum of medieval Arabic science (*al-ʿulūm al-ʿarabiyya*). This is not the consequence of undifferentiated functional contents of the denoted units; rather, it stems from the conceptual perception of language as a unitary process (Carter 2004:53), whereby the model of language structure is linear and one-dimensional and lacks multi-tier organization (Frolov 1991:134). The term *ḥarf* can be applied to elements of any size and length, as long as it denotes a quantum of enunciated and hence recorded information that is small in scale but not strictly limited (Weiss 1910:375–379; Fischer 1989:140), its actual meaning depending entirely on the context. Building upon the perception of the term *ḥarf* as an ultimate unit (Ibn Jinnī, *Sirr* I, 15–19) with its own meaning (Fleisch 1986:204b), it may be treated as a segment of speech with a semiological value, specified both semantically and semiotically. Consequently, two general meanings of the word *ḥarf* can be viewed as most prominent in early Arabic texts: a unit with a syntactic status (semantically specified), on the one hand, and a unit of a phonological (scriptural) character (semiotically specified), on the other (Owens 1990:245).

Both meanings derive from the most comprehensive definition of the term *ḥarf*, by Sībawayhi (*Kitāb* I, 1), as a word that "has a meaning, but is neither a noun nor a verb" (*ḥarf jāʾa li-maʿnan laysa bi-smin wa-lā fiʿl*). The second part of this definition, 'neither a noun nor a verb', implies that the notion *ḥarf* may apply to both nouns and verbs. It is used to indicate the class of words in general, which can be broken down further into three subclasses. The third subclass is represented by those units from the *ḥarf* category that are neither nouns nor verbs (a particular usage of the term is the subclass of particles lacking a common semantic characteristic and morphological regulation), because noun words and verb words have as their basis a precisely defined set of morphological models and a set range of meanings that only vary insignificantly. Particles do not have any structured form or model whatsoever in their basis, and they may have a wide range of meanings (Carter 2004:88).

As far as the first part of the definition is concerned (a word that 'has a meaning'), this feature implicitly puts the *ḥarf* in opposition to the eponymous category of units, which even though lacking semantic meaning (cf. Versteegh 1977:44, 45; Carter 2004:75; but according to Levin [2000:45], the correct interpretation of this phrase in Sībawayhi is 'which occurs in order to denote a meaning and nothing else [except this specific meaning]'), nonetheless have a semiotic function. It is the nature of that unit that generates the most frequent use of the term *ḥarf*. By virtue of the ambiguity of its content, this use has produced a wide array of explanations and interpretations in Western studies of Arabic, because in this connotation it correlates with a whole number of similar (in terms of the system) concepts of theoretical linguistics, such as phoneme, grapheme, prosodeme, and morpheme (see Fischer 1989 for a general review). Attempts at viewing any single one of these aspects independently face considerable obstacles, created by their interconnected nature, because such an approach goes against the functional universality of a unit postulated in the Arabic system of grammar.

The syntactic role of the *ḥarf* (in the narrow sense of 'particle') is defined by its instrumental function, a fact supported by the alternative name of this subclass, *ʾadā* lit. 'tool' (Carter 2004:74). In many cases, *ḥarf* manifests itself as an operational element (*ʿāmil* 'operator'; → *ʿamal*). Since particles are the exclusive instrument for performing linguistic functions, each of them can be defined by its function, for example *ḥarf istifhām* 'interrogative particle', *ḥarf nidāʾ* 'vocative particle', *ḥarf qasam* 'particle of oath, vow', and so on (Carter 2004:88). These functions form specific subclasses of particles: *ḥarf aš-širāk* (Sībawayhi, *Kitāb* I, 211.10, 264.15, 382.17) 'coordinate particle'; *ḥarf jarr* (*Kitāb* I, 244.1, 252.12, 329.14) or the possessive *ḥarf* (*Kitāb* I, 276.16), the 'preposition' that controls the *-i* inflection. *ʾInna* and the particles that govern two nominal complements (→ *ʾinna wa-ʾaxawātuhā*) are also called *ḥarf* (*Kitāb* I, 241.13, 244.14). Other subclasses of particles called *ḥarf* include particles of warning (*tanbīh; Kitāb* I, 277.18), particles of exception (*Kitāb* I, 314.17), and verb operators (*Kitāb* I, 361.13, 363.11, 406.1), realized at the beginning of the sentence. Other particles include those that occur sentence-initially (*Kitāb* I, 244.12, 367.15, 391.10, 429.20), those that must be followed by a verb (*Kitāb* I, 407.16), and conditional particles (*Kitāb* I, 398.8).

In the phonological sense, the term *ḥarf* is viewed as a phoneme. This is validated by data coming from early philological texts, which present the behavior of sounds in phonetic processes through the term *ḥurūf*, as well as by research done by later philologists, such as az-Zamaxšarī (see Ibn Yaʿīš, *Šarḥ* X, 120–155, 124.7-8; Carter 2004:120–121). In descriptions of their articulatory characteristics, reference is made to the place of articulation (*maxraj* lit. 'place from which [the sound] exits') of the *ḥurūf*. It may be concluded from this that Arab philologists view the *ḥarf* as a discrete, phonetically differentiated sound (*ṣawt*). In this sense, *ḥarf* is a subclass of *ṣawt*, a sound with specific and discrete features (Owens 1988:91, 95). The correlation of the notions of *ḥarf* and → *ḥaraka* is explained through the vowel/consonant contrast (Levin 1986:425). In terms of the graphical realization (*ḥarf yatahajjī bihi* or *ḥurūf li-l-hijāʾ*; see Weiss 1910:357, 359–360), this approach leads to the perception of the *ḥarf* as the equivalent of the 'letter of the alphabet' concept (Fleisch 1986:204b; Al-Nassir 1993:9–10).

On the other hand, it is noteworthy that in more recent grammatical treatises (Ibn Sīnā, *Risāla*, Chap. 2), the terms *ḥarf* and → *ḥaraka* are used to indicate consonants and vowels when they are viewed functionally as components of a word, but when the sounds of speech are viewed as independent units of the phonological system, the terms *ṣawt ṣāmit* 'consonant' and *ṣāʾit* or *muṣawwat* 'vowel' (Axvlediani 1981:93; Bravmann 1934:7–18, 112–135) are preferred. The structure of a word in the Arabic linguistic tradition is described and mapped in *ḥurūf*, which suggests that → *ḥaraka* should be regarded as a vowel element, integrated into the *ḥarf*, rather than a vowel sound that is added to *ḥarf*. Consequently, sound segments, such as /ba/, /bu/, /bi/, /bØ/ (i.e. *b* with *sukūn*), should be interpreted as different variants of the same *ḥarf* that emerge as a result of a variation in the vowel component of the *ḥarf*, while the consonant remains invariant. According to al-Xalīl ibn ʾAḥmad, every *ḥarf* has its own sound and variation (Sībawayhi, *Kitāb* II, 342.21–23: *wa-zaʿama l-Xalīl ʾanna l-fatḥa wa-l-kasra wa-ḍ-ḍamma zawāʾid wa-hunna yalḥaqna l-ḥarf li-yūṣala ʾilā t-takallum bihi wa-l-bināʾ huwa s-sākin allaḏī lā ziyādata fīhi* 'al-Xalīl claims that the /a/, the /i/, and the /u/ are additions that attach themselves to the consonant so as to make it possible to pronounce it; the pausal form is the vowelless consonant without addition'). Variation is revealed in the *ḥaraka* (Gabučan 1965:121).

Alternatively, it has been noted that, unlike the classic definition of the phoneme as a unit with differentiating features as well as an integrative (morpheme-forming) function, the consonantal component *ḥarf* typically has only differentiating features, whereas the implementation of the integrative function is delegated to the element *ḥaraka*. This is why the status of the phonological segment capable of forming the exponent of a morpheme requires the combination of both elements (Karabekyan 2004:510). With this approach, *ḥarf*, corresponding not only to a separate consonant (*ḥarf sākin*), but also to a combination of sounds (*ḥarf mutaḥarrik*), can be correlated with the notion of 'grapheme' rather than that of 'letter' (Frolov 1991:56, 57). Conceptually, this is linked with the general methodological premise of the Arab linguistic tradition, which considered the spoken word to be prior to the written (Owens 1988:284). On the other hand, it should be noted that as an element of the script system, *ḥarf* is in the first place a coding unit. Any segment of Arabic speech can be coded as a sequence of *ḥurūf*. Through the application of minor modifications, the structure of the Arabic written language is derived from this code system (Sančes 1968:91). Others have suggested that as a term, *ḥarf* does not indicate any physical substrate at all but rather is a diacritic element, capable of transforming into one physical substrate or another, acoustic or scriptural (Gabučan 1965:120).

The analysis of the basic notions of the Arabic prosodic system (*ʿarūḍ*), in which *ḥarf* occurs as the main operational notion (Sančes 1968:86), also suggests that *ḥarf* is a unit, largely similar to the notion of 'mora' (Trubetzkoy 1977:169–179; Al-Nassir 1993:9–10). As such, it is used both in morphology, where it has a quantitative parameter, and in metrics, where *ḥarf* indicates metric feet and their components (Frolov 1991:54). This conclusion is supported by the analysis of the terminological pair *ḥarf mutaḥarrik* and *ḥarf sākin*. Underlying this contrast is the polarity of the prosodic functions of two types of *ḥarf*, rather than the modality of their phonetic realization. It is no coincidence that despite their reduced consonantal status, the so-called *ḥurūf al-madd wa-l-līn* are thought to belong to the *ḥarf sākin* class on the basis of their prosodic function.

In morphological terms, the notion *ḥarf* can be interpreted as a (morphological) position. Relevant here is the distinction between basic *ḥurūf* (*ḥurūf ʾaṣliyya*: root positions of the base depicted through the symbols *fāʾ*, *ʿayn*, *lām*, which are open to realization by phonologically perceived segments) and augmented *ḥurūf* (*ḥurūf zāʾida*: positions added to the base, closed to realization, and represented by one segment from a given list, abstracted from its exact phonetic meaning, occurring only as a quantitatively structuring morphological model of a diacritic notion). One of the key questions here is the morphological status of *ḥurūf ʾaṣliyya* and *ḥurūf zāʾida*. On the one hand, both the *ḥurūf ʾaṣliyya* aggregate (as root morphemes) and the separate *ḥurūf zāʾida* (as analogues of affixes) are customarily explained as morphemes. At the same time, the notions *ḥurūf ʾaṣliyya* and *ḥurūf zāʾida* are a complex morphological characteristic (*faʿala; tafaʿØʿala*, etc.) of the segment of speech they indicate (*kataba, jalasa, rasama; tarakØkaba, tanafØfasa*, etc.), rather than the segment of speech itself. Therefore, it is not the *ḥurūf ʾaṣliyya* and *ḥurūf zāʾida* that have the status of a morpheme, but this status is rather assigned to segments that are formed by a certain series of concrete realizations of the open positions (*ḥurūf ʾaṣliyya*), as well as at a certain localization of the closed positions (*ḥurūf zāʾida*). Furthermore, the *ḥurūf zāʾida* of this type (*tāʾ, sīn, hamza*, etc.) differ in function from the *ḥurūf zāʾida* that additionally have their own discrete nominal or nominal-relative value (*ḥurūf al-muḍāraʿa, tanwīn, tāʾ at-taṯniya*, etc.). The latter are considered to be positions distributed around the basis that shape the word form. Unlike the *ḥurūf zāʾida* of the former type, they are partially open to realization by a strictly defined set of segments that modify the overall meaning of the given position. For example, the position *tanwīn* may be realized in two ways, the *nØ* (*nūn sākina*) and the null one, whereas the position of, for instance, the *ḥurūf al-muḍāraʿa* allows for four realizations, and so forth.

This leads us to believe that the morphological analysis of the Arabic linguistic tradition is based on a range of interrelated factors whereby the variability of characters in the *ḥurūf ʾaṣliyya* and their invariance in the *ḥurūf zāʾida*, on the one hand, and the preservation of the quantitative integrity of *ḥarf* as a unit of morphological calculation, on the other, form a complex system that constitutes the main mechanism for forming both the expression and the content plan of a word within, as well as outside, the root base. The functioning of this system itself is based on the conceptual nature of *ḥarf* as a universal operational unit of linguistic analysis.

Bibliographical references

Primary sources

Ibn Jinnī, *Sirr* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *Sirr ṣināʿat al-ʾiʿrāb*. 3 vols. Ed. ʾAḥmad Farīd ʾAḥmad. Cairo, n.d.

Ibn Sīnā, *Risāla* = ʾAbū ʿAlī al-Ḥusayn ibn ʿAbdallāh Ibn Sīnā, *Risāla ʾasbāb ḥudūṯ al-ḥurūf*. Ed. Vladimer Grigor'evič Axvlediani, *Fonetičeskij traktat Avicenny*. Tbilisi: Mecniereba, 1966.

Ibn Yaʿīš, *Šarḥ* = Muwaffaq ad-Dīn Yaʿīš Ibn Yaʿīš, *Šarḥ al-Mufaṣṣal*. 10 vols. Beirut and Cairo: ʿĀlam al-Kutub and Maktabat al-Mutanabbī, n.d.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. 2 vols. Ed. Hartwig Derenbourg. Paris: Imprimerie Nationale, 1881. (Repr., Hildesheim and New York: G. Olms, 1970.)

Secondary sources

Al-Nassir, ʿAbd al-Munʿim ʿAbd al-ʾAmīr. 1993. *Sībawayh the phonologist: A critical study of the phonetic and phonological theory of Sībawayh as presented in his treatise* Al-Kitab. London and New York: Kegan Paul International.

Axvlediani, Vladimer Grigor'evič. 1981. "Arabskoe jazykoznanie srednix vekov" [Arabic linguistics of the Middle Ages]. *Istorija lingvističeskix učenii: Srednevekovyj Vostok* [History of linguistics: The medieval East], ed. Agnija V. Desnickaja and Solomon D. Kacnel'son, 53–94. Leningrad: Nauka.

Bravmann, Max. 1934. *Materialien und Untersuchungen zu den phonetischen Lehren der Araber*. Göttingen: W.F. Kaestner.

Carter, Michael G. 2004. *Sībawayhi*. Oxford: Oxford University Press.

Fischer, Wolfdietrich. 1989. "Zur Herkunft des grammatischen Terminus *ḥarf*". *Jerusalem Studies in Arabic and Islam* 12.135–145.

Fleisch, Henri. 1986. "*Ḥarf*". *Encyclopaedia of Islam* III, 204. 2nd ed. Leiden: E.J. Brill.

Frolov, Dmitrij V. 1991. *Klassičeskij arabskij stix* [Classical Arabic verse]. Moscow: Nauka.

Gabučan, Gračija Mikajelovič. 1965. "K voprosu o strukture semitskogo slova (v svjazi s problemoj 'vnutrennej fleksii')" [On the question of the structure of a Semitic word (in relation to the issue of the 'inner inflection')]. *Semitskie Jazyki* 2:1.114–127.

Karabekyan, Samvel B. 2004. "Urovni lingvističeskogo analiza i osnovnyje jedinicy opisanija v tradicionnom arabskom jazykoznanii (k postanovke problemy)" [The levels of linguistic analysis and the basic units of description in the Arabic linguistic tradition]. *Countries and Peoples of the Near and Middle East* (Yerevan) 23.507–518.

Levin, Aryeh. 1986. "The Mediaeval Arabic term *kalima* and the modern linguistic term morpheme: Similarities and differences". *Jerusalem Studies in Arabic and Islam* 7.423–446.

——. 2000. "The meaning of *ḥarf ǧāʾa li-maʿnan* in Sībawayhi's *al-Kitāb*". *Jerusalem Studies in Arabic and Islam* 24.22–48.

Mosel, Ulrike. 1975. *Die syntaktische Terminologie bei Sībawaih*. Ph.D. diss., University of Munich.

Owens, Jonathan. 1988. *The foundations of grammar*. Amsterdam and Philadelphia: J. Benjamins.

——. 1990. *Early Arabic grammatical theory: Heterogeneity and standardization*. Amsterdam and Philadelphia: J. Benjamins.

Sanĉes [Sánchez], Alcaén A. 1968. "K voprosu o suščnosti sistemy arabskoj metriki" [On the question of the essence of the Arabic metrical system]. *Arabskaja filologija: Sbornik statej*, ed. Aleksandr Aleksandrovič Kovalev and Gračija Mikajelovič Gabučan, 86–95. Moscow: University Press.

Trubetzkoy, Nikolaj S. 1977. *Grundzüge der Phonologie*. 2nd ed. Repr. Göttingen: Vandenhoeck und Ruprecht. (1st ed., 1958.)

Versteegh, Cornelis [Kees] H.M. 1977. *Greek elements in Arabic linguistic thinking*. Leiden: E.J. Brill.

Weiss, Josef. 1910. "Die arabische Nationalgrammatik und die Lateiner". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 64.349–390.

Samvel Karabekyan and Marat Yavrumyan
(Yerevan State University)

Ḥassāniyya Arabic

1. GENERAL INFORMATION

Ḥassāniyya (or *klām əl-Bīḏān* 'language of the Whites') is the mother tongue of the Arabic-speaking population of the western Sahara, especially the Moors (Bīḏān) of Mauretania and the former Spanish Sahara (from the Sagya el-Ḥamra and the Rio del Oro). It is difficult to draw the precise geographical limits of this dialect, but its approximate borders are Goulimine in the north, Tindouf in the northeast, Tombouctou in the southeast, and the Senegal River in the south. The percentage of Ḥassāniyya speakers is highest in the central regions. There are about 3 million speakers, around 2 million of them living in Mauretania, out of a total estimated population of 2.9 million. By the middle of the 20th century, most speakers still had a Bedouin lifestyle.

The origin of the Ḥassāniyya is linked (as indicated by its name) to the arrival of the Banū Ḥassān, a branch of the Maʿqil Arabs – who were themselves linked to the movements of territorial expansion of the Banū Hilāl and the Banū Sulaym. Ḥassāniyya is a Bedouin dialect that is part of the western dialect group (Maghrebi) and has developed in a Berber-speaking environment. Despite the influence of the substrate and because of its Bedouin nature, it has more in common with eastern Arabic dialects, most of which (though not all) are Bedouin, than with most of the Maghrebi dialects like Moroccan and Algerian.

Ḥassāniyya is rarely used as a lingua franca, even if certain Black African Mauretanians are more or less proficient in it. It is not used as a means of written communication; Literary Arabic (Classical or Standard) or a foreign language, in particular French, are used for writing. An important oral literature exists, most of it poetry, but the conditions that gave rise to this literature are rapidly changing (→ Mauretania).

Ḥassāniyya has been studied extensively, but no recent manuals exist, except in xeroxed form. A reference grammar was published by Cohen (1963). Although its subtitle is 'Dialect of the Gəbla', it may be regarded as typical of the dialect variety that is spoken in Mauretania as well as in the former Spanish Sahara. An overview of the grammar is given in the introduction to Taine-Cheikh's Ḥassāniyya/French dictionary (1988:I, CIII). For grammatical details see Taine-Cheikh's articles, published in particular in *Matériaux Arabes et Sudarabiques*.

Ḥassāniyya shows an exceptional unity, with a few exceptions. Genuine Ḥassāniyya speakers, as well as unilingual or, more often, bilingual people speaking a variety that differs from the standard dialect, can be found at the Moroccan, Malian, and probably Algerian borders. The dialects of these speakers are so different that they are virtually incomprehensible to the uninformed Ḥassāniyya speaker (Heath 2002, 2004). Nowadays, this variation is largely individual, but formerly it was often linked to the history of certain tribal groups, such as the Tekna of Morocco or the Brābīč and the Kunta of Mali.

In the 1970s, under the influence of Arab nationalist movements and in the context of forced settlement, intellectuals and their students developed a form of Arabic that was a mix of dialect and Modern Standard Arabic, used in political discussions (Taine-Cheikh 1978). This 'middle' Arabic has spread gradually, introducing many new words into the traditionally very rural dialectal lexicon.

2. LINGUISTIC DESCRIPTION

Ḥassāniyya is a relatively conservative language (at least when compared to other Maghrebi dialects), but it has also developed several innovations, in particular certain morphosyntactic patterns (Taine-Cheikh 1991). The Berber substrate does not seem to have had a large influence on the structure of Arabic. Its presence may be perceived in the lexicon, but since most loanwords receive special treatment, their influence is usually limited. Where common forms exist between Ḥassāniyya and Mauretanian Berber (Zenaga), it is often hard to attribute the source to either language. In several cases, a parallel development may have taken place, facilitated by the remote genealogical relation between Arabic and Berber. This development may have been furthered by the progressive disappearance of Zenaga and a process of osmosis between Arabic- and Berber-speaking groups.

Table 1. Inventory of consonants

	velarized	labial	inter-dental	dental	pre-palatal	post-palatal	velar	pharyngeal	laryngeal
plosive	-	b		d	d^y	g			
voiced	+	ḅ		ḍ					
plosive	-			t	t^y	k			(’)
voiceless	+			ṭ			q		
continuant	-	v	ḏ	z	ž		ġ	ʿ	
voiced	+	(ṿ)	ḏ̣	ẓ					
continuant	-		ṯ	s	š		x	ḥ	h
voiceless	+			ṣ					
nasal	-	m		n	n^y				
	+	ṃ		(ṇ)					
lateral	-			l					
	+			ḷ					
vibrant	-			r					
	+			ṛ					
semivowel		w			y				

2.1 *Phonology*

2.1.1 Consonants

2.1.1.1 Inventory (Table 1)

/ġ/ is attested only among certain speakers (mainly in the west, southwest, and northwest) when it is not geminated. /ġ/ and /q/ have merged, being realized as [q] by the other Ḥassāniyya speakers (central, eastern, and northeastern Mauretania, Mali, and Algeria), and by all in case of gemination (/ġġ/ realized [qq] as in Zenaga).

The phonological system is rich. Ḥassāniyya tends to phonemicize the opposition emphatic/non-emphatic (especially for the vibrants, /ṛ/ being more frequent than /r/) and to preserve the pronunciation of some of the borrowed lexemes, hence the presence of phonemes from Classical Arabic (/ḍ/, /q/, or even /’/), Zenaga (/ẓ/, /d^y/, /t^y/, /n^y/), and even from Black African languages (/q/ and the palatalized consonants).

Yet, certain phonemes remain marginal, especially in loanwords and the velarized *ḅ, ṃ, ṇ, ṿ*). It is often difficult to find minimal pairs, except for *l~ḷ*, *g~q*, and above all *r~ṛ*: *dār/idīr* 'to put' vs. *dāṛ/idōṛ* 'to want'; *gās* 'to go toward' vs. *qās* 'measure'; *gām* 'to get up' vs. *qām* 'to prepare the tea'; *ḍall* 'to err [in religion]' vs. *ḏ̣all* 'to spend the day'; *zanga* 'to skirt around while going up' vs. *ẓanga* 'to make someone pay a tribute'; *walla* 'to come back' vs. *waḷḷa* 'or'; *bərga* 'shacks' vs. *ḅəṛga* 'bad-quality tea'; *tamātāya* '(a) gum tree' vs. *taṃātāya* '(a) tomato'.

Even though the status of some phonemes is problematic (especially in the case of *ṇ* and *ṿ*), this does not challenge the existence of emphasis (for an opposite position see Zavadovskij 1981:26–27).

2.1.1.2 Historical remarks on the inventory

Ḥassāniyya is characterized by the realization of *qāf* as [g], the maintaining of the interdentals (/ḏ̣/ being the reflex of most words with *ḍ* in Classical Arabic), and the disappearance of the *hamza* (often compensated, at the end of the syllable, by lengthening the preceding vowel).

2.1.1.3 Phonetic realization

/j/ is realized as a palato-alveolar fricative [ʒ]. The labial spirant is realized preferentially as a voiced consonant [v], except in contact with a voiceless consonant or when it is geminate. This realization is particular to this dialect (with the exception of the Ḥassāniyya of Mali).

2.1.1.4 Distribution

The emphasis of the vibrant varies sometimes according to the context. Emphasis may be lost or absent in the presence of /y/, /ī/, or even /ə/: *ʿašṛa* 'ten', *ʿəšrīn* 'twenty', *ʿāšər* 'to count by tens'. More generally, the assimilation of sonority and emphasis – in particular for /s/~/ṣ/, /z/~/ẓ/ – is frequent among the consonants, whether in contact or not: *ʿrīs* 'bridegroom', but *ʿṛūṣ* 'bride'. Extended to all the forms of the same root, this helps to explain certain differences with Classical Arabic (sometimes common to other dialects), e.g. *t-f-l*: *dvəl* 'to spit' (but in the east and in Mali: *tfəl*), *q-t-l*: *ktəl* 'to kill', *ṣ-ġ-r*: *sġayyər* 'small', *ṣ-ʾ-y*: *ẓwä* 'to twitter'. There are also some conditioned alterations between sibilant and palato-alveolar fricatives belonging to the same root (assimilation with loss of the palato-alveolar fricative), thus *n-s-z*: *näzz* 'to weave', *ž-z-z*: *zäzz* 'to shear'.

2.1.1.5 Sociolinguistic variables

Some tendencies appear among the least educated groups, e.g. the marginalization of /ḍ/; in the southwest, the merger of /q/ and /ġ/; and, only in a limited region, the tendency to emphasize *t* in contact: [ŧɍaːb] 'earth'.

2.1.2 Vowels (Table 2)

Table 2. Vowels

long	short	
	open syllable	closed syllable
/ī/	/i/	/ə/
/ū/	/u/	
/ā/	/a/	/a/

The merger of /i/ and /u/ in closed syllables is characteristic of the nomadic dialects (Cohen 1970). The short phonemes /a/ and /ə/ are realized variably, according to context. In a neutral context, /a/ undergoes *ʾimāla* and is realized more centralized (transcribed *ä*). Long vowels have variable length: long under the accent, average apart from the accent, and short in final position. Final long vowels are lengthened again before a suffix: *šāvu* 'they have seen', *šāvū-h* 'they have seen him/her/it'. When /ā/ is realized as a short vowel, it undergoes *ʾimāla*: *žä* 'he came', *žā-h* 'he came to him'.

2.1.3 Diphthongs

The four former diphthongs are preserved: /ay/, /aw/, /iy/, and /uw/. However, the realization of /ay/ and /aw/ sometimes tends toward [eː] and [oː].

2.1.4 Syllables

Because of the general preference for closed syllables, short vowels in open syllables are rare, apart from loanwords and in final position. They are found, however, in several initial syllables in which the short syllable represents a first radical *w/y* (*uvä* 'he is over') or plays an important morphological role (*aḥmaṛ* 'red', *ikättäb* 'he makes [them] write', *udägdäg* 'to be broken').

The most frequent syllabic type is CVC and CVV, but syllables with double coda CVCC or double-onset CCVC are frequent. Closed syllables with long vowels (CVVC) are attested, especially in the participles: *kātbīn* 'writing [pl.]'. Several open syllables have a secondary origin (3rd radical *w/y*): *žäru* < *žərw* 'young dog' (fem. *žərwä*).

2.1.5 Consonant clusters

The general rule for consonant clusters is to introduce epenthetic vowels after elision of short vowels in an open syllable, *malḥafa* > /malḥfa/ 'veil [of the women]' (realized [mælᵊħfæ]). In monosyllabic nouns, metathesis is regular, except in loans from Classical Arabic: [ŧfəl] 'boy' (but [ŧəvl-u] 'his boy').

2.1.6 Stress

Stress is on the third mora from the end of the word, e.g. on the first syllable in *mäktäb* 'desk', on the second in *mäktūb* 'written'. It is strongly marked only on long vowels (accent of length rather than intensity). Many grammatical morphemes are clitics.

2.2 *Morphology*

2.2.1 Pronouns

The gender opposition is never marked in the 1st person.

2.2.1.1 Personal independent pronouns (Table 3)

Table 3. Personal pronouns

		singular	plural
3rd	masc.	*huwwä*, *hūwä*	*hūmä*
	fem.	*hiyyä*, *hīyä*	*hūmāti*
2nd	masc.	*(ə)ntä*	*(ə)ntūmä*
	fem.	*(ə)nti*, *(ə)ntiyyä*	*(ə)ntūmāti*
1st		*ānä*	*(ə)ḥnä*, *nəḥnä*

2.2.1.2 Possessive/object suffixes (Table 4)

Table 4. Possessive/object suffixes

			singular		plural
			after a consonant	after a vowel	
1st		object suffixes	*-ni*		*-nä*
		possessive suffixes	*-i*	*-yä*	
2nd	masc.		*-ak*	*-k*	*-kum*
	fem.		*-ək*		
3rd	masc.		*-u* (*-ū-*)	*-h*	*-hum*
	fem.		*-hä* (*-hā-*)		

The clitic pronoun of the 1st person singular has two different forms, after a verb (*šāv-ni* 'he saw me') or after a preposition (*vī-yä* 'in me') or noun (*ktāb-i* 'my book'). In certain special contexts (e.g. after *mā-* 'not', *mən-* 'who?'), a short form of the independent pronoun is used for the 3rd person singular: masc. *-hu*, fem. *-hi*. The possessive pronouns consist of a base that varies according to gender in the singular (masc. *līl-*, fem. *līlt-*; pl. *lwāyl-*) and of an affix pronoun: *līl-i* 'mine'.

2.2.1.3 Demonstratives (Table 5)

Table 5. Demonstratives

	masc. sg.	fem. sg.	pl.
proximity or neutral use 'this'	*ḏä*	*ḏi*	*ḏu*
proximity (*hā-*) 'this one, this'	*hāḏä*	*hāḏi*	*hāḏu*
distance (*-k*) 'that one, that'	*ḏāk*	*ḏīk*	*ḏūk*

2.2.1.4 Presentatives

The presentatives consist of an independent personal pronoun (in the sg., short or long form), preceded by a demonstrative or a particle with a verbal origin: *ḏähu(wwä)* 'here', *ḏākhu(wwä)* 'there', *(a)ṛahu(wwä)* 'there he is', *ḏīkhi(yyä) maṛyäm* 'there is Maryem'.

2.2.1.5 Relative pronoun

The relative pronoun is invariable in gender and in number: *lli* (sometimes *əl*) 'who, what'.

2.2.1.6 Interrogative pronouns

Interrogative pronouns include *mən* 'who?' (*mən-hu* 'who is it?'); *š(ə)-* 'what?' and its variants: *-āš* after a preposition; *šən-* in *šən-hu* 'what is it?'; *äyy* (invariable) 'which one?' (*äyy-kum* 'which one of you?').

2.2.2 Adverbs

i. Interrogative adverbs: *mnäyn* and *wäyn* 'where?', *äyntä* 'when?', *kämm* 'how much?', *əʕl-āš* 'why?', *škīv* 'how?'
ii. Adverbs of place: *hūn*, *hūnāti(yyä)* 'here', *vämm*, *vämmāti(yyä)* 'there', *hōʷk*, *hōʷkāti(yyä)* 'over there', *ilāh* 'toward there', *l-gäddām* 'in front of', *ət-taḥt* 'under', *əl-väwg* 'above '
iii. Adverbs of time: *ḏaṛk*, *ḏaṛkāti(yyä)* 'now', *l-yäwm* 'today', *ġdä* 'tomorrow', *əṣṣəbḥ* 'tomorrow morning', *yāməs* 'yesterday', *əlbārəḥ* 'yesterday night', *əlläylä* 'tonight'
iv. Adverbs of quantity: *yāsər* 'a lot', *ḥattä* 'very', *šwäyy* 'little, few'

2.2.3 Particles

The *l* of the definite article *əl-* assimilates to all 'sun letters' and to *ž*. There is no indefinite article and no particle of the genitive. The verbal negative form is *mā* in assertive sentences (without a second element), *lā* with the imperative. In a nominal sentence, the negative form is combined with the suffix pronouns (*mā-ni*, *mā-n-ak*, etc.).

i. Prepositions: *və* (*vī-* + pronoun) 'in', *ʕlä* (*əʕlī-*) 'on', *bə* (*bī-*) 'with', *mən* 'from, of', *ʕand* 'by', *šäwr* 'toward', *gäddām* 'in front of', *uṛa* 'behind, after', *sābəg* 'before', *taḥt* 'under', *väwg* 'above'
ii. Subordinating conjunctions: *ʕan*, *änn* 'that'; *mnäyn* 'when'; *bī(h)əlli* 'because'; *äyyāk*, *bbāš* 'for'; *ilā*, *iḏa* 'if' (condition); *(yä)kān* 'whether' (indirect interrogation)
iii. Coordinating conjunctions: *wə/u* 'and', *waḷḷa* 'or', *(ya)ġäyr/(ya)qäyr* 'but'

2.2.4 Nouns

The singular feminine form of nouns and adjectives ends in *-a(t)* with a few exceptions such as *ʿanz* ‘goat’, *xādəm* ‘woman slave’, *ʿayn* ‘eye’, *dāṛ* ‘house’, *xandūd* ‘good milker’, *ḥāməl* ‘pregnant’ (but *ḥāmlä* ‘[who] wears’).

Apart from the broken plurals there is an external plural: masculine *-īn* and feminine *-āt*; there are some → pseudo-duals such as *uḏnäyn* ‘ears; two ears’.

Numerous Berber loanwords have special affixes: prefixes in *a(a)-/i(i)-* for masculine nouns, *ta(a)-/ti(i)-* for feminine nouns; suffixes *-t* for feminine singular nouns and *-ən* for plural nouns.

The pattern C1aC2C2āC3 (nouns of habit, profession) is very frequent: *käḏḏāb* ‘liar’.

Adjectives of color and defect: masc. sg. aCCaC, *aḥmaṛ* ‘red’, fem. sg. CaCCa, *ḥamṛa*, comm. plural CəCC, *ḥəmṛ*.

aCCaC is also the pattern of the comparative form (invariable): *akbaṛ* ‘taller’, *aḥmaṛ* ‘more red’.

The diminutive formation is very productive and very differentiated for nouns and adjectives: CCayC *kläyb* (< *kälb* ‘dog’), CCayyəC *ktäyyəb* (< *ktāb* ‘book’), CCayCəC *ʿgäyrəb* (< *ʿagrəb* ‘scorpion’), CCayCīC *bzäyzīl* (< *bäzzūl* ‘udder’), aCayCəC *aḥaymər* (< *aḥmaṛ* ‘red’).

2.2.5 Numerals

Cardinals 1 and 2 agree in gender: 1 masc. *wāḥəd*, fem. *waḥdä*; 2 masc. *äṯnäyn*, fem. *ṯäntäyn*. The dual is still productive: *kälb-äyn* ‘two dogs’. Certain cardinals have two forms. From 3 to 10, the long form in *-a* is used in the absolute state.

	absolute state	construct state
3	*(ä)ṯlāṯä*	*äṯlət*
4	*aṛbʿa*	*aṛbaʿ*
5	*xamsä*	*axməs*
6	*səttä*	*sətt*
7	*säbʿa*	*äsbaʿ*
8	*(ä)ṯmānyä*	*äṯmən*
9	*təsʿa*	*ətsaʿ*
10	*ʿašṛa*	*əʾšəṛ*

From 11 to 19, the cardinals are used without final *-ər* in the absolute state.

	absolute state	construct state
11	*aḥdaʿš*	*aḥdaʿšəṛ*
12	*aṯnaʿš*	*aṯnaʿšəṛ*
13	*aṯləṭṭaʿš*	*aṯləṭṭaʿšəṛ*
14	*arbaʿṭaʿš*	*arbaʿṭaʿšəṛ*
15	*axməsṭaʿš*	*axməsṭaʿšəṛ*
16	*səṭṭaʿš*	*səṭṭaʿšəṛ*
17	*äsbaʿṭaʿš*	*äsbaʿṭaʿšəṛ*
18	*äṯmənṭaʿš*	*äṯmənṭaʿšəṛ*
19	*ätsaʿṭaʿš*	*ätsaʿṭaʿšəṛ*

The number 100 is *miyyä* in the absolute state and *mīt* in the construct state. The other cardinals have an invariable form: 20 *ʿəšrīn*, 30 *ṯlāṯīn*, 40 *aṛbʿīn*, 50 *xamsīn*, 60 *səttīn*, 70 *säbʿīn*, 80 *ṯmānyīn*, 90 *təsʿīn*, 200 *mītäyn*, 1,000 *älv*.

With the exception of *äwwäl* ‘first’, the ordinals have the pattern of the participle C1āC2əC3: *ṯāni* ‘second’.

2.2.6 Verbs

2.2.6.1 Patterns/stems

2.2.6.1.1 Pattern I: Triradicals

The stem vowels are /a/ or /ə/; there are three subclasses, the most frequent being that of verbs with a harmony between the two vowels: type a: *ktəb/yəktəb* ‘to write’ (often roots without back consonant); type b: *vtaḥ/ yavtaḥ* ‘to open’; type c (mixed): *ṛgaṣ/yərgəṣ* ‘to dance’.

The diminutive form aC1ayC2aC3/yaC1ayC2aC3 *äkäytäb/yäkäytäb* ‘to write with a bad handwriting’ is rare.

2.2.6.1.2 Pattern I : Quadriradicals

This type is unique, with two stem vowels /a/: *gaṛmaṣ/igaṛmaṣ* ‘to pinch’. Particular cases are (R2=R4) *baxbax* ‘to burst out laughing’, (R3=R4) *baḥšäš* ‘to strangle’.

2.2.6.1.3 Derived patterns (Table 6)

The derived forms are numerous and productive. In many cases there is a correlation between active and reflexive (forms with/without *t-* or *-t-*) and a systematic relation between active and passive (forms with/without *n-* or *u-*). The internal passive has disappeared in Ḥassāniyya (as in most dialects), but a new

Table 6. Derived forms

	active meaning*	reflexive meaning	passive meaning
triradicals	I C1C2ə/aC3	VIII (ə)C1tC2ə/aC3	VII (ə)nC1C2ə/aC3
doubling R2	II C1aC2C2aC3	V tC1aC2C2aC3	uC1aC2C2aC3
lengthening V after R1	III C1āC2əC3	VI tC1āC2əC3	uC1āC2əC3
prefix *s*-	'IV' saC1C2aC3	X staC1C2aC3	usaC1C2aC3
lengthening V after R2	'IX' (ə)C1tC2āC3	—	—
quadriradicals	C1aC2C3aC4	tC1aC2C3aC4	uC1aC2C3aC4

system of formal oppositions has developed to denote the distinction between middle/reflexive/reflexive-passive, on the one hand, and a real passive form, on the other. Only Form 'XI' (expressing a change of state) is isolated in this system.

The stem vowel of the derived forms is always that of the perfect of Form I. Form VIII (rare as middle, *əštġal* 'to work') is used as passive of Form I if the first radical is *l*, *m*, *n*, *r*, *ṛ*, *w*, or an original *hamza*: *(ə)rtdəm* 'to be buried', *(ə)ltʿan* 'to be cursed'. Form VII is the regular passive of Form I: *(ə)nktəb* 'to be written', *ənvtaḥ* 'to be opened'.

Form II is very frequent as causative-factitive and iterative of Form I or denominative: *gaṭṭaʿ* 'to have it cut; to cut in small pieces', *baṛṛag* 'to make it shine'. Form V is frequent as middle or reflexive of Form II: *tgaṭṭaʿ/yətgaṭṭaʿ* 'to cut itself in small pieces'; *u*-II: *ugaṭṭaʿ* 'to be cut in small pieces; to be made to go across'.

Form III is quite frequent as extensive of Form I or causative-factitive of Form VI: *vārəg* 'to separate one from the other'. Form VI is frequent as reciprocal, middle, or reflexive of Form III: *tvārəg* 'to separate ourselves from one another'; *u*-III: *uvārəg* 'to be separated from one another'.

Form 'IV' has some causatives-factitives of Form X: *saʿṛab* 'to arabize', *saḥmaṛ* 'to get it reddish'. Form X is quite frequent as reflexive, middle, and inchoative: *staʿṛab* 'to arabize itself', *staḥmaṛ* 'to become reddish'.

Form 'XI' is rare: *gṣāṛ* 'to become short'.

Examples of derived forms of quadriradical verbs are *baʿṛaṣ* 'to fluster', *tbaʿṛaṣ* 'to fluster oneself', *ubaʿṛaṣ* 'to be flustered'.

2.2.6.2 Inflection of aspects and moods

2.2.6.2.1 Perfect (Table 7)

Table 7. Perfect verb

	Triradicals: type a	Triradicals: types b and c	Quadriradicals
3rd sg. masc.	*ktəb*	*vtaḥ*	*baʿṛaṣ*
3rd sg. fem.	*kətbət*	*väthət*	*baʿᵊṛṣət*
2nd sg. masc.	*ktəbt*	*vtaḥt*	*baʿṛaṣt*
2nd sg. fem.	*ktəbti*	*vtaḥti*	*baʿṛaṣti*
1st sg.	*ktəbt*	*vtaḥt*	*baʿṛaṣt*
3rd pl.	*kətbu*	*väthu*	*baʿᵊṛṣu*
2nd pl.	*ktəbtu*	*vtaḥtu*	*baʿṛaṣtu*
1st pl.	*ktəbnä*	*vtaḥnä*	*baʿṛaṣnä*

2.2.2.6.2 Imperfect (Table 8)

The prefix vowel of the imperfect is always /ə/ for derived verbs beginning with two consonants (V, VI, VII, VIII, X, and 'XI'), *yətbaʿṛaṣ* 'he flusters himself'. The vowel *u*- of the passive form is constant, *yubaʿṛaṣ* 'he was flustered'.

Table 8. Imperfect verb

	Triradicals: types a and c	Triradicals: type b	Quadriradicals
3rd sg. masc.	*yəktəb*	*yävtaḥ*	*ibaˁṛaṣ*
3rd sg. fem.	*təktəb*	*tävtaḥ*	*tbaˁṛaṣ*
2nd sg. masc.	*təktəb*	*tävtaḥ*	*tbaˁṛaṣ*
2nd sg. fem.	*təkᵊtbi*	*tävᵊtḥi*	*tbaˁᵊṛṣi*
1st sg.	*nəktəb*	*nävtaḥ*	*nbaˁṛaṣ*
3rd pl.	*yəkᵊtbu*	*yävᵊtḥu*	*ibaˁᵊṛṣu*
2nd pl.	*təkᵊtbu*	*tävᵊtḥu*	*tbaˁᵊṛṣu*
1st pl.	*nəkᵊtbu*	*nävᵊtḥu*	*nbaˁᵊṛṣu*

2.2.6.2.3 Imperative (Table 9)

Table 9. Imperative

	Triradicals: types a and c	Triradicals: type b	Quadriradicals
2nd sg. masc.	*ktəb*	*avtaḥ*	*baˁṛaṣ*
2nd sg. fem.	*kᵊtbi*	*ävᵊtḥi*	*baˁᵊṛṣi*
2nd pl.	*kᵊtbu*	*ävᵊtḥu*	*baˁᵊṛṣu*

2.2.6.3 Participles and elatives

All derived verbs have a participial form in *m*- except VII and 'XI'. The participles form their plural form with the suffixes *-īn* and *-āt*. The suffix of the feminine (**-at*) is in *-a*, except before a direct object clitic pronoun: *hiyyä šārbä* 'she drank', *hiyyä šārᵊbt-u* 'she drank it'.

Elatives (invariable) exist for all participles, with an identical form for the forms with/without *t* (Taine-Cheikh 1984:284–290): I (*vāhəm*) *äfhäm mən* 'understanding better than'; VIII (*məštmaṛ*) *äšmaṛ mən* 'demonstrating more courage than'; II/V (*mˁalläm, mətˁalläm*) *äˁalläm mən* 'teaching/studying better than'; III/VI (*mvārəg, mətvārəg*) *ävārəg mən* 'separating (themselves) from one another better than'; 'IV'/X (*msaˁṛab, məstaˁṛab*) *äsaˁṛab mən* 'arabizing (oneself) more than'; quadriradicals (I and with *t*-) (*mbaˁṛaṣ, mətbaˁṛaṣ*) *abaˁṛaṣ mən* 'flustering (oneself) more than'.

2.2.6.4 Verbal nouns (Table 10)

With the exception of Form 'XI', verbal nouns usually exist for all verbs with a non-passive meaning. The forms vary for I (long or short vowel): *vähm* 'understanding', *ḥsāb* 'act of counting', *ġrīg* 'act of sinking', *rsūl* 'act of sending'. One form prevails for all other cases, often common to verbs with/without *t*: II/V təC1C2āC3, *təbṛāg* 'act of making something shine'; III/VI tC1āC2īC3, *tvārīg* 'mutual separation; act of separating from one another'; 'IV'/X staC1C2īC3, *staˁrīb* 'arabization; act of arabizing oneself'; quadriradicals (I and with *t*-) tC1aC2C3īC4, *tbaḥšīš* 'strangling; the act of strangling (oneself)'.

The instance noun is generally in *-a*. It may be accompanied by a pattern change for I (*räslä* '(a) sending') and regularly so for II/V: təC1C2īC3a, *təbṛīgä* 'act of making something shine once'.

Table 10. Verbal nouns

	active	reflexive	passive	elative
—	I C1āC2əC3	VIII məC1tC2ə/aC3	mäC1C2ŪC3	I and VIII aC1C2aC3
doubling R2	II mC1aC2C2aC3	V mətC1aC2C2aC3	muC1aC2C2aC3	aC1aC2C2aC3
lengthening V after R1	III mC1āC2əC3	VI mətC1āC2əC3	muC1āC2əC3	aC1āC2əC3
prefix *s*-	'IV' msaC1C2aC3	X məstaC1C2aC3	musaC1C2aC3	asaC1C2aC3
quadriradicals	mC1aC2aC3aC4	mətC1aC2aC3aC4	muC1aC2aC3aC4	aC1aC2aC3aC4

2.2.6.5 Weak verbs (Table 11)

Table 11. Weak verbs

	geminate	I *w/y*	II *w/y*	III *w/y*
type a	2 cases: *ḏall/iḏall* ‘to while away the time’, *tämm/itämm* ‘to go on’	R1=*w* *uṣal/yäwṣal* ‘to arrive’	(rare) R2=*w*: *xāf/ixāf* ‘to be scared’ R2=*y*: *bāt/ibāt* ‘to spend the night’	R3=*w/y* (rare) *nsä/yänsä* ‘to forget’
type b	—	R1=*w*: *uzən/yūzən* ‘to weigh’ R1=*y*: *ibəs/yībəs* ‘to dry’	R2=*w*: *gāl/igūl* ‘to tell’ R2=*y*: *gās/igīs* ‘to head for’	R3= *w/y* (frequent) *šrä/yəšri* ‘to buy’
type c	R2=R3 *baṭṭ/ibəṭṭ* ‘to beat’			
particularities	1st, 2nd pers. perf. in *-äy-*: *baṭṭäyt*	3rd pers. fem., pl. perf. in *w-*: *waṣlət* *wəznət/ūznət*	1st, 2nd pers. perf. in *-ə-*: *gəlt*	3rd pers. pl. imperf. type a: *yänsāw* type b: *yəšru*
active participle	*bāṭṭ*	*wāzən*	R2=*w* > y: *gāyəl*	*nāsi* (fem. *nāsyä*) *šāri* (fem. *šāryä*)
passive participle	*mäbṭūṭ* (pl. *mbaṭṭa*)	*mäwzūn*	R2=*w* > y: *mägyūl*	*mänsi(yy)* *məšri(yy)*

The derived verbs are generally well attested.

i. Geminated verbs: Derived forms include Forms III and VI: *sātt/isātt* ‘to count by sixes’, *tmāss mʿa* ‘to adjoin something’; X: with a joint form *stäxaff* ‘not to take something seriously’, or disjunctive *stäglāl* ‘to regard as rare’.
ii. I *ʾalif*: In integrated borrowings from Classical Arabic, the first radical of these verbs is represented by *ā*: *āḏən əl* ‘to authorize’; X: *stāxaṛ* ‘to move back’.
iii. I *w/y*: Form VIII has a passive meaning: *ūtzən/yūtzən* ‘to be weighed’.
iv. II *w/y*: The alternation *ā~ə* attested in the perfect of Form I likewise occurs in the derived forms VII, VIII, and IX; VIII: *ḥtāl* ‘he was crafty’, *ḥtəlt* ‘I was crafty’; Form X: with *ā*, *stävād* ‘to take advantage of’; with *w* or *y*, *stälyän* ‘to become more supple’.
v. III *w/y*: The vowel of the 3rd person singular is always *ā* in the perfect. In the imperfect *ī* is used in Forms II, III, ‘IV’, and single quadriradicals verbs; *ā* is used in Forms V, VI, VII, VIII, X, quadriradicals verbs with *t-*, and all the *u*-forms.
vi. Irregular verbs: There are two verbs with ‘mixed’ conjugation (perfect of II *w*, imperfect of I *w*): *kāl/yäwkäl* ‘to eat’ (participles *wākəl* and *mäwkūl*) and *xāḏ/yäwxaḏ* ‘to leave’; the verb *žä/iži* ‘to come’ has the participle *žāy*.

2.3 *Syntax*

The syntax of the dialect shows only a few specificities when compared with the other Bedouin dialects of the Maghreb.

2.3.1 Noun phrase

The main characteristic of the noun phrase is the preservation of the synthetic construction.

2.3.1.1 Expression of (in)definiteness

The presence/absence of the definite article *əl* marks definiteness, except in the construct state and with certain masculine nouns borrowed from Berber (generally beginning with *a(a)-* or *i(i)-*): *kälb* ‘(a) dog’ ~ *əlkälb* ‘the dog’, *ävūk* ‘(a ~ the) veal’.

The demonstrative generally precedes the noun, which is always determined: *ḏīk-əddāṛ* 'this house'. Sometimes it follows the noun, especially with proper nouns: *ṭəvl aḥmäd ḏāk* 'this son of Ahmad', 'the son of this Ahmad'. There is no indefinite article. The partitive is expressed with *mən* 'of': *wāḥəd mən ləktūb* 'one of the books', *ktāb mən ləktūb* 'any of the books'.

2.3.1.2 Construct state

Possession is expressed by the construct state: *ktāb əṭ-ṭfəl* 'the book of the boy', *ktāb-u* 'his book'. There is no genitive particle in Ḥassāniyya, except, infrequently, in Morocco: *dyal* (Taine-Cheikh 1999:98–99), *ntaˁ* (Heath 2002:7).

2.3.1.3 Numeral phrase

If the counted noun is indetermined, the numerals (from 3 upward) are always constructed as nouns in annexion: *ˁašrīn žmäl* 'twenty camels'. The form used is the one of the construct state: from 3 to 10 short and/or contracted (*axməs äklāb* 'five dogs'; with a *-t* suffix before some masculine forms with an original *hamza*, *axməs-t äyyām* 'five days'); from 11 to 19 long with *-ər* (*axməsṭaˁšər ktāb* 'fifteen books').

If the counted noun is determined, the numeral is constructed as an adjective (invariable in gender from 2 upward): *läklāb läṯnäyn* 'the two dogs', *ktūb aḥmäd əlˁašrīn* 'the twenty books of Ahmed', *žmāl-u l-miyyä* 'his hundred camels'. The form used is the one of the absolute state: from 3 to 10 long in *-a* (*läklāb lxamsä* 'the five dogs'); from 11 to 19 short without *-ər* (*ləktūb laxməsṭaˁš* 'the fifteen books').

2.3.1.4 Adjective phrases

The order is noun+adjective. Adjectives agree in gender and number with the noun they determine: *ṛāžəl msäggäm* '(a) fair man', *ṛažžālä msäggmīn* '(some) fair men', *mṛa msäggmä* '(a) fair woman', *ˁläy(y)āt msäggmāt* '(some) fair women'. They are preceded by the definite article when the noun is determined, *ənnāgä lbäyḏ̣a* 'the white female camel', *nyāg aḥmäd əlbīḏ̣* 'the white female camels of Ahmad', or when it is highly referential, *maṛyäm əššäybāniyyä* 'Maryem, the old woman'.

2.3.1.5 Elative constructions

Followed by *mən* (introducing the second term of the comparison), the elative expresses the comparative: *äkbaṛ mən xū-h* 'taller than his brother'. Followed by a determined noun (or pronoun), it expresses the relative superla-tive: *äkbaṛ-hum* 'the taller among them', *äkbaṛ əṭ-ṭavilāt* 'the taller of the girls'. When it is definite, it expresses the absolute superlative: *läkbaṛ* 'the tallest', *ləmṛa läkbaṛ* 'the tallest woman'.

2.3.1.6 Relative clauses

The relative pronoun does not appear with an undetermined antecedent: *mšä mˁa ḥadd mā gaṭṭ šəfnā-h* 'he is gone with someone we have never seen', but *mšä mˁa əṛṛāžəl lli vətnä šəfnä* 'he is gone with the man we had already seen'. Note the absence of the referential pronoun in relative clauses with *lli*.

2.3.2 Verbal phrase

The direct object precedes the indirect object, which is introduced by *əl*: *əˁṭa žmäl l-xāl-u* 'he gave a camel to his maternal uncle', *əˁṭā-h l-u* 'he gave it to him'. However, the indirect object is expressed without *əl* when it is the only suffix: *əˁṭā-h žmäl* 'he gave him a camel'.

2.3.3 Verbal aspect: Time and tense

Innovations are very limited. There is no indicative prefix.

2.3.3.1 Future intent prefixes

The predicted future is expressed with the invariable particle *lāhi* (+ imperfect), which corresponds etymologically to the participle of *lhä* 'to keep oneself busy doing something': *lāhi ngīs nwākšōṭ əššhar əddāxəl (in-šā-allāh!)* 'I'll go to Nouakchott next (God willing!)'. Combined with past modality (perfect of *kān* 'to be'), *lāhi* express the future in the past, *kənt lāhi ngīs nwākšōṭ mnäyn ˁədt mäwžūˁ* 'I was about to go when I fell ill'. The imperfect of *dāṛ* (verb of desire) is also used, especially for the intended future: *ndōṛ nəmši* 'I am going (willing) to leave'.

2.3.3.2 Use of active participle

In its predicative use, the active participle has the meaning of a concomitant action. Depending on the verb, it expresses a concomitant action in the unaccomplished [= progressive present or past] (*huwwä ṭāləˁ əlkədyä* 'he is going up the mountain', *hiyyä kānət ṭālˁa əlkədyä* 'she was going up the mountain'), or in the accomplished [= resultative perfect] (*huwwä mətġaddi* 'he had lunch' [= he is not hungry anymore]). The participle is sometimes used for an imminent action, as if the agent was already engaged in the action: *ānä māši (ṣṣəbḥ)* 'I am going (tomorrow)'.

2.3.3.3 Negation

The usual negative form is *mā*, but *lā* is used in several cases: for prohibitions, with the imperfect (*lā təbki!* 'don't cry!'); after the coordinator *wə/u* 'and' (*mā žä u lā ktəb* 'he didn't come and he didn't write'; for a negative wish, with the perfect, in some expressions (*lā qərzu nyāghum!* 'I hope the milk of your camels doesn't dry up!'); sometimes, as an 'expletive' negation after *xāf* 'to fear'.

The negation is not discontinuous, but the affix pronoun appears regularly in the absence of a conjugated verbal form (non-verbal predicate or presence of *lāhi*): *mān-ak gaṛṛāy* 'you are not a teacher', *mā-ni xāyəv* 'I am not afraid', *(aḥmäd) mā-hu lāhi yəbki* '(Ahmad) he won't cry'.

2.3.4 Word order SVO, VSO

In the absence of thematization, the order of the verbal sentence is VSO. This is the only possible order if the subject is totally undetermined (Taine-Cheikh 1998). The interrogatives *mən* 'who?' and *əš* 'what?' come at the beginning of the sentence.

2.3.5 Existential sentences

The dialect has several 'pseudo-verbs' consisting of a preposition and an affix pronoun, in particular *ʿand-* (possession, hence 'to have'), *l-* (attribution and alienable possession), *mʿa-* 'with', *vī-* (location). The order is VO (SVO with a thematized 'subject' noun), and the negative form is *mā*: *(ʿayšä) mā ʿand-hä vaḏḏa* '(Aïsha) she doesn't have money'. The present participle of *xləg* 'to exist' expresses existence: *(mā-zāl) xāləg mbūṛu* 'there is (still) bread'.

3. LEXICON

The Ḥassāniyya lexicon is rich and well structured around a relatively limited number of high-frequency schemes (cf. Taine-Cheikh 1988–1998). It seems always to have had a tendency to enrich itself, either through borrowings (from Berber, closely related African languages, Literary Arabic, French, etc.) or through internal development. Yet, the majority of the lexicon, at least 80 percent of the lexical items and maybe 90 percent of the roots, is still Arabic in origin. A quite important part of this vocabulary is more or less characteristic of the Maghrebi dialects, especially of the Bedouin type, e.g. *ʿatrūs* 'goat', *bġa* 'to desire', *gḏəv* 'to vomit', or *yāməs* 'yesterday'.

The Arabic core is very stable, and it is found, for its essential parts, across the whole Ḥassāniyya-speaking area. Borrowings, calques, neologisms, semantic shifts, and other innovations are often less stable and more localized (for the eastern region, cf. the lexicons of Pierret 1948 and Heath 2004). The history of these forms, hard to reconstruct, is often more autochthonous than it seems. The Ḥassāniyya dialect has certainly borrowed more from Berber (especially Zenaga) than from any other non-Arabic language, but many creations seem to have appeared, if not first in Ḥassāniyya, at least simultaneously in Ḥassāniyya and in Zenaga. Indeed, they are often absent from both the Berber and Arabic lexicons.

BIBLIOGRAPHICAL REFERENCES

Cohen, David. 1963. *Le dialecte arabe ḥassānīya de Mauritanie (parler de la Gəbla)*. Paris: Klincksieck.

——. 1970. "Le système des voyelles brèves dans les dialectes maghrébins". *Etudes de linguistique sémitique et arabe*, 172–178. The Hague and Paris: Mouton.

Heath, Jeffrey. 2002. *Jewish and Muslim dialects of Moroccan Arabic*. London and New York: Routledge Curzon.

——. 2004. *Hassaniya Arabic (Mali)-English-French dictionary*. Wiesbaden: O. Harrassowitz.

Nicolas, Francis. 1953. *La langue berbère de Mauritanie*. Dakar: Institut Français d'Afrique Noire.

Pierret, Roger. 1948. *Etude du dialecte maure des régions sahariennes et sahéliennes de l'Afrique Occidentale française*. Paris: Imprimerie Nationale.

Taine-Cheikh, Catherine. 1978. *L'arabe médian parlé par les arabophones de Mauritanie: Etude morphosyntaxique*. Thèse de 3ème cycle, Université Paris V.

——. 1984. "Un exemple de créativité lexicale: L'élatif en hassaniyya". *Arabica* 31.274–305.

——. 1988–1998. *Dictionnaire ḥassāniyya-français*. Paris: Geuthner. [8 vols., *hamza – gāf*, published to date.]

——. 1990. *Lexique français-ḥassāniyya*. Nouakchott: Centre Culturel Français Saint-Exupéry. (Repr., Paris: Geuthner, 2004.)

——. 1991. "L'arabe des Bīḏān, un dialecte bédouin de Maghreb occidental". *Semitic studies in honor of Wolf Leslau*, ed. Alan S. Kaye, 1528–1548. Wiesbaden: O. Harrassowitz.

——. 1998. "Topicalisation, thématisation et anaphore en arabe". *La thématisation dans les langues:*

Actes du colloque de Caen (1997), ed. Claude Guimier, 247–261. Bern and Paris: P. Lang.

——. 1999. "Les hassanophones du Maroc: Entre affirmation de soi et auto-reniement". *Peuples méditerranéens (Langues et stigmatisations sociales au Maghreb)* 79 (April–June 1997), 85–102.

Zavadovskij, Ju. N. 1981. *Mavritanskij dialekt arabskogo jazyka: Xassaniya* [The Mauretanian dialect of Arabic: Xassaniya]. Moscow: Izdatel'stvo 'Nauka'.

Catherine Taine-Cheikh
(CNRS, LACITO-UMR 7017)

Hausa

1. Introduction

Most scholars who speculate on the origin of the Hausa agree that the Hausa's ethnic composition includes a Hamitic element. If Palmer's (1967:95) view is taken as reasonably true, the Hausa people developed from a mixture of groups migrating from the central Sahara (due to desertification) to the central savanna in the south during the 1st millenium C.E. The new group which emerged out of that contact was relatively more sophisticated and later on absorbed a number of other small ethnic groups, all constituting together one cultural and linguistic entity, with the Hausa language as a unifying factor. Therefore, the term 'Hausa' is in actual fact more a linguistic than an ethnic term, and the Hausa people can be regarded as a nation rather than a tribe.

Spoken by well over 80 million people, the Hausa language is the first lingua franca of West Africa. The Hausa migrations for the purposes of trade, pilgrimage, and preaching of Islam led to the spread of their language beyond the original homeland (Hausaland: present Nigeria and Niger Republic). Hausa also has a significant presence in Ghana, Togo, Chad, Cameroon, Central African Republic, southern Libya, and Sudan, and is known as far as Ethiopia and Saudi Arabia.

In Nigeria, Hausa exists in two major dialects: the eastern (or Kano) and the western (or Sokoto) dialects. Each of them comprises a chain of further subdialects and geographically extends to the neighboring parts of the Niger Republic. Although the eastern dialect has been affected by more phonological and morphological erosion and simplification than the western dialect, it is still the variety on which Standard Hausa is based.

2. Hausa and Arabic

Hausa is one of the few languages that enjoys a strong relationship with Arabic at three distinct levels: common descent (i.e. genetic relationship), indirect contact through literary traditions (in West Africa), and direct contact through human migrations to Arabic-speaking lands (e.g. Sudan, Libya, and Saudi Arabia).

With regard to the first level, it is now unanimously accepted among Africanists that Hausa shares a common descent with Arabic. The two languages are classified as members of the old Hamito-Semitic family reorganized and renamed by Greenberg (1966) as 'Afro-Asiatic': Arabic as Semitic and Hausa as Hamitic (Meinhof 1912) or Chadic (Greenberg 1966). The Ethnologue (www.ethnologue.com 2000) lists it as Chadic, West.

The common linguistic heritage of the two languages is reflected in various elements. Examples of the most salient of these include:

i. (V+)t as a feminine marker: Hausa *ya/ta zo*, Arabic *jāʾa/jāʾat* 'he/she came'
ii. Some object and possessive pronouns: Hausa *ya ba ni/ka/ki/ku*, Arabic *ʾaʿṭā nī/ka/ki/kum* 'he gave me/you [sg. masc.]/[sg. fem.]/[pl.]'; Hausa *gidanka/ki/ku*, Arabic *baytuka/ki/kum* 'your [sg. masc.]/[sg. fem.]/[pl.] house'
iii. Some forms of broken plural: Hausa *doki (<dawki)/dawaki* 'horse/horses', Arabic *zawraq/zawāriq* 'boat/boats'
iv. The morpheme m+V for derivation of noun of agent, place, or instrument: Hausa *rubuta/marubuci* 'to write/writer', Arabic *qātala/muqātil* 'to fight/fighter'
v. A few cognates: Hausa *ƙashi* 'bone', Arabic *qaṣ* 'chest bone'; Hausa *afa* 'to put in the mouth', Arabic *fā* 'mouth'; Hausa *tofa* 'to spit', Arabic *taffa* 'to spit'; Hausa *yau* 'today', Arabic *yawm* 'day'

When studying the Arabic loanwords in Hausa, their common heritage should be taken into consideration, and inherited features should be left out of the discussion. Note that the Nigerian Hausa orthography, which does not mark vowel length, is used for the language examples.

3. Contacts between Arabic and Hausa

The indirect contact between Hausa and Arabic in West Africa (through commercial transactions, literary traditions, and intermediary languages) extends over six centuries, during which the Hausa language borrowed and integrated a considerable corpus of words and concepts from the Arabic. Arabic loanwords in Hausa have been historically identified and registered to varying degrees in all the Hausa dictionaries (Mischlich 1906; Robinson 1913; Bargery 1934; Abraham 1962; Newman 1977; McIntyre and Meyer-Bahlburg 1991), in addition to a few research papers and monographs. The most comprehensive lists of these loans are those of Greenberg (1947) and Baldi (1988), comprising 455 and 1,245 items, respectively. The large discrepancy in size between the two lists may be attributed partly to the difference in sources from which each author drew his data and partly to the time span separating the compilation of their lists (35 years).

Arabic words are borrowed into Hausa from both written and oral sources, directly from the target language and indirectly through other intermediary lingua francas, mainly → Berber and → Kanuri. The phonological features of a substantial number of these loans, coupled with some extralinguistic evidence, indicate that the North African (Maghrebi) dialects of Arabic were the major suppliers of the orally borrowed words. In fact, the North African factor in Hausa Islamic culture is very significant. This is in addition to the early commercial links which existed between Hausaland and North Africa, reflected in the many loanwords denoting North African commodities, such as *susiyya* < *sūsiyya* 'purple cloth or thread' (from Sousse in Tunisia), *za'afaran* < *zaʿfarān* 'saffron', *zaitun* < *zaytūn* 'olive'. Another piece of evidence is that of words borrowed with the undoubtedly North African contracted definite article *al*. This article is usually realized in the North African dialects as *li-* before plosives and *l-* before continuants. Thus, the Standard Arabic form *al-kitāb* 'the book' becomes *liktāb*, and *al-xayma* 'the tent, umbrella' becomes *lxayma*. The form of these two words in Hausa and the different phonological changes they underwent are as follows: Arabic *al-kitāb* > North African dialect *liktāb* > Hausa *liktabi* > *liktafi* > *littafi* (addition of the vowel suffix *-i*, change of *b* into *f*, and regressive assimilation of *k* to *t*); Arabic *al-xayma* > North African *lxayma* > Hausa *laima* (deletion of *x* to avoid a consonant cluster in a word-initial syllable).

Relying on their phonetic behavior vis-à-vis their Arabic etyma and on other cultural evidence, Greenberg (1947) assigned all Arabic loanwords in Hausa to two major groups. Loans in Group I were borrowed earlier than those of Group II, and mainly from colloquial sources, whereas those of Group II were more recent and borrowed mainly from written sources. The characteristics of Group I can be summarized as follows:

i. Loans display irregular treatment of the Arabic sounds: e.g. Arabic *b* rendered by Hausa *f*, as in *aljifu* < *al-jayb* 'pocket'; *ṯ* and *ḏ* represented by *t* and *d*, as in *talata* < *ṯulaṯāʾ* 'Tuesday'; *dara'a* < *ḏirāʿ* 'cubit'.
ii. The definite article is usually borrowed with the noun in a contracted form *l-* or *li-*, as in the above examples of *laima* and *littafi*.
iii. They comprise terms of everyday life, trade, and technology and elementary aspects of Islamic religion.

Loans of Group II, on the other hand, are characterized by the following:

i. They display more regular treatment of the Arabic sounds, e.g. Arabic *b* > Hausa *b*, as in *aibi* < *ʿayb* 'fault, defect'; *ṯ* > *s*, as in *wasiƙa* < *waṯīqa* 'letter'; *ḏ* > *z*, as in *zamba* < *ḏanb* 'fraud, swindling'.
ii. The article, if borrowed with the nouns, is in its complete form, *al*, as in *alkali* < *al-qāḍī* 'judge'.
iii. They include words that refer to the more recondite aspects of Islam and technical terms of pseudosciences (grammar, astrology, etc.).

The characteristics of these groupings exhibit a high degree of regularity, but exceptional cases can also be encountered. A few loans may be detected with characteristics of both groups, e.g. *aljifu* < *al-jayb* 'pocket', with the complete form of the article *al* (Group II) and the representation of *b* by *f* (Group I).

4. Intermediary languages

Many loans, especially those identified by unusual treatment of the Arabic sounds, reached Hausa through intermediary languages. This is why Wexler (1980) emphasizes the role of African lingua francas in the diffusion of Arabic loanwords in Central and West Africa, and the importance of this fact in retracing the different paths followed by these words into the recipient languages. With regard to Hausa, three intermediary languages have been shown to provide loans: Berber, Kanuri, and → Fulfulde. Loans passed through Berber can be recognized by the Berber feminine marker *ta-*, as in Hausa *tad(d)awa* < Berber *taduat* < Arabic *dawāt* 'inkstand', and replacement of *ṣ* by *z* as in Hausa *azurfa* 'silver' < Berber *azref* < Arabic *aṣ-ṣarf* 'changing money [formerly in silver]'. Kanuri, on the other hand, contributed loans such as Hausa *kasuwa* < Kanuri *kasuwu* < Arabic *as-sūq* 'market', and Hausa *sirdi* < Kanuri *sirdi* < Arabic *sarj* 'saddle'. Finally, in words such as Hausa *hubbare* < Arabic *qubba* 'tomb of a religious leader', the class suffix *-re* speaks for Fulfulde as an undoubted intermediary channel.

5. Phonology

The phonological adaptation of the Arabic loanwords in Hausa relates basically to the process of altering the Arabic sounds that do not exist in Hausa and the unusual behavior of those that do exist in Hausa. The variations concern only consonants, since all the Arabic vowels (in terms of quality as well as quantity) exist in Hausa and therefore do not undergo any significant change in the recipient language, with the exception of a few isolated cases.

Hausa and Arabic share 17 consonants: *b t d k ʾ f h s z š j m n r l w y*. There are 11 consonants in Arabic that do not exist in Hausa: *ṭ ḍ q ṯ ḏ ḏ̣ ṣ x ġ ḥ ʿ*, while Hausa includes 14 consonants that do not exist in Arabic: *ɓ ɗ c* t^s k^y k^w *ƙy ƙ ƙw g gy* g^w *ʾy ɽ*. In terms of their treatment in Hausa, Arabic sounds can be divided into three categories:

i. Sounds that do not exist in Hausa and that are usually replaced by the phonetically nearest Hausa sounds. These include *q ṯ ḏ ḏ̣ s x ġ ḥ* and ʿ. Examples:

q > *k* or *ƙ*: *kabila* < *qabīla* 'tribe', *fasiƙi* > *fāsiq* 'profligate'
ṯ > *s* or *t*: *wasiƙa* < *waṯīqa* 'letter', *talata* < *ṯulaṯāʾ* 'Tuesday'
ḏ > *z* or *d*: *kazafi* < *qaḏf* 'false accusation', *idan* < *ʾiḏā* 'if'
ḏ̣ > *z*: *azahar* < *aḏ̣-ḏ̣uhr* 'noon (prayer)'
x > *h* or Ø: < *hatimi* < *xātim* 'seal, stamp', *lahira* < *al-ʾāxira* 'the hereafter'
ġ > *g*: *gaibi* < *ġayb* 'the unknown'
ḥ > *h*: *haji* < *ḥajj* 'pilgrimage'
ṣ > *s* or *z*: *nasiha* < *naṣiḥa* 'advice', *azumi* < *aṣ-ṣawm* 'fasting'
ʿ > ʾ: *jamaʾa* < *jamāʿa* 'the public, crowd, community'

ii. Sounds that do not exist in Hausa and are replaced by phonetically less related Hausa sounds. These include *ṭ* and *ḍ*. Examples:

ṭ > *ɗ* (alveolar voiced implosive) or *ts* or *ʾy* (dialectal variations): *ɗibbu/tsibbu/ʾyibbu* < *ṭibb* 'medicine', *ɗahara* < *ṭahāra* 'ritual purity'
ḍ > *l*: *laʾifi* 'sexually impotent' < *ḍaʿīf* 'weak', *alkali* < *al-qāḍī* 'judge'

According to the description by the Arab grammarians, these two sounds are realized as lateral (→ *ḍād*) and voiced, respectively. As such, their nearest Hausa correspondents are indeed *l* and *ɗ*, respectively (and not *d* and *t* as expected).

iii. Sounds that do exist in Hausa and yet sometimes are also replaced by phonetically less-related Hausa sounds. These include *b m n s r j*. Examples:

b > *f* (intervocalic): *aljifu* < *al-jayb* 'pocket'
m > *b* (one instance): *albashi* < *al-maʿāš* 'salary, wage'
s > *š* (before front vowel): *numfashi* < *nafas* 'breath'
n > *l*: *lakadan* < *naqdan* 'in cash'
r > *n* (one instance): *alharini* < *al-ḥarīr* 'silk'
j > *d* (from a Sudanese dialect through Kanuri) or *z* (through Fulfulde?): *sirdi* < *sarj* 'saddle', *zuwaira* < *juwayriyya* '[female's name]'

The irregular behavior of some Arabic sounds in Hausa primarily reflects the intermediary languages, but in some cases it is the result of phonological constraints in Hausa, e.g. Hausa *t* > *c* (before front vowel), hence *lokaci* < *al-waqt* 'time'; or of internal sound shift, e.g. *b* shifted to *w*, hence *allura* (< *alliwra*) < *al-ʾibra* 'needle'. In other cases it is the result of dialectal variations, e.g., *f* in the eastern dialects is realized as *h* in the western dialects, hence *sahu* < *ṣaff* 'row, a line of people'.

6. MORPHOLOGY

Like most African languages, Hausa words as a rule end in open syllables. Therefore adaptation of Arabic loanwords in Hausa involves opening the final closed syllables through attachment of the vowel suffix *-i* for masculine substantives (e.g. *alkalam-i* < *al-qalam* 'pen') or, in a few cases, the nominative case marker *-u*, as in *sahu* < *ṣaff* 'row, line of people', especially with Arabic proper names (*umaru*, *bashiru*, etc.) Final closed syllables can also become open through the deletion of the final consonant, as in *albasa* < *al-baṣal* 'onion'. Feminine nouns are borrowed with their feminine ending *-a*, which is identical with the Hausa feminine marker (common heritage). Otherwise, nominals systematically attach *-i* when masculine and *-a* when feminine, as in *jaʾir-i*/*jaʾir-a* < *jāʾir* 'a shameless man/woman'. Non-final closed syllables, too, sometimes become open through insertion of an epenthetic vowel, as in *lakadan* < *naqdan* 'in cash'. The Hausa nominalizing morpheme *-ci* is suffixed to loan adjectives to derive nouns, as in *haʾinci* 'dishonesty' < *xāʾin* 'dishonest', or to loan nouns for their further integration, as in *hukunci* < *ḥukm* 'judgment, verdict'. Likewise, the verbalizing morpheme *-ta* is attached to loan nouns to derive verbs, as in *hukunta* 'to pass judgment' < *ḥukm* 'judgment, verdict'. Some Arabic verbs are used in Hausa as nouns, as in *ƙaddara* 'fate' < *qaddar* 'to destine'. Derived verbs behave like proper Hausa verbs; they receive their tonal patterns in accordance with their assimilation to one or the other Hausa verbal classes or grades; e.g. *bayyana* (Grade II) 'to make clear' < *bayyana*. The root then takes various affixes and tonal changes, such as *ya bayyana* 'it appeared', *ya bayyana mishi* 'he explained to him', *bayyana mishi* 'explain to him!'. However, assignment of tone to nominals follows a different pattern in that the stressed syllable of the Arabic word is often (though not always) assigned a high tone and the following (but not necessarily the preceding) syllable always carries a low tone.

7. SEMANTICS

Arabic loanwords in Hausa cover a wide range of semantic fields and touch upon almost all aspects of the life of Hausa Muslims. However, more than half of the loanwords recorded by Greenberg and Baldi derive from the Islamic religion in its broad sense, i.e. as a way of life embodying not only a system of belief but also a series of social institutions. Therefore, this macrosemantic field can be detailed in a number of related fields, such as social, political, and legal domains. Other important semantic fields include trade, literacy, and sciences, as well as numerals and time reckoning.

i. Religion (faith, practice, and concepts): *imani* < *ʾīmān* 'belief', *addini* < *ad-dīn* 'religion', *salla* < *ṣalāt* 'prayer', *adili* < *ʿādil* 'just', *alƙurʾani* < *al-qurʾān* 'the *Qurʾān*'
ii. Social system: *sadaki* < *ṣadāq* 'bride-money', *iyali* < *ʿiyāl* 'family', *balaga* < *balaġ-* 'to reach puberty', *likkafani* < *al-kafan* 'shroud', *taʾaziyya* < *taʿziya* 'condolence'
iii. Political organization: *siyasa* < *siyāsa* 'politics', *waziri* < *wazīr* 'vizier', *jamhuriyya* < *jumhūriyya* 'republic', *jamʾiyya* < *jamʿiyya* 'political party', *mulki* < *mulk* 'rule, reign'
iv. Legal system: *alkali* < *al-qāḍī* 'judge', *tuhuma* < *tuhma* 'suspicion', *shaida* < *šahida* 'to witness', *hukunci* < *ḥukm* 'judgment, verdict'
v. Trade and imported commodities: *attajiri* < *at-tājir* 'trader; wealthy man', *mizani* < *mīzān* 'scales', *riba* < *ribḥ* 'profit', *asara* < *xusāra* 'loss', *sandal* < *ṣandal* 'sandalwood', *zaitun* < *zaytūn* 'olive'
vi. Literacy and sciences: *littafi* < *al-kitāb* 'book', *alkalami* < *al-qalam* 'pen', *karatu* < *qirāʾatu* 'reading; learning', *hisabi* < *ḥisāb* 'arithmetic; astrology', *fikhu* < *fiqh* 'jurisprudence', *ɗibbu* < *ṭibb* 'medicine'
vii. Numerals and time reckoning: *ashirin* < *ʿišrīn* 'twenty', and all decimal numbers up to ninety; *laʾasar* < *al-ʿaṣr* 'late afternoon (prayer)', and all prayer times; *asabar* < *as-sabt* 'Saturday', and the rest of the days of the week

Other minor fields include: household utensils, e.g. *sahani* < *ṣaḥn* 'plate'; tools, e.g. *allura* < *al-ʾibra* 'needle'; warfare items, e.g. *bindiga* <

bunduqiyya ‘gun’; and a large number of miscellaneous items such as abstract concepts, e.g. *annashawa* < *an-našwa* ‘joyful feeling’, *niyya* < *niyya* ‘intention’, *niʾima* < *niʿma* ‘bounty’.

A few cases of reborrowing have also been recorded whereby the same loan exists in two morphological forms indicating different sources and times of borrowing, either with the same or with a slightly different meaning. Examples of these are *cazbi/carbi* and *tasbaha* < *tasbīḥ* ‘rosary’; *lissafi* ‘calculation’ and *hisabi* ‘astrology’ < *(al-)ḥisāb* ‘calculation’.

Most loanwords are used in Hausa with their original Arabic meanings, but some have undergone varying degrees of semantic modifications: (a) by semantic extension, e.g. *attajiri* ‘trader; wealthy man’ < *at-tājir* ‘trader’; *alhaji* ‘pilgrim; socially distinguished personality’ < *al-ḥājj* ‘pilgrim’; (b) by semantic shrinking: *bidiʾa* ‘innovation in religious practices; merrymaking; drumming’ < *bidʿa* ‘innovation in religious practices’; *sunna* ‘prophetic tradition; sexual intercourse with a wife’ < *sunna* ‘prophetic tradition’; *laʾifi* ‘sexually impotent’ < *ḍaʿīf* ‘weak’; (c) by semantic intensification: *alkawari/alkawali* ‘solemn promise’ < *al-qawl* ‘word, promise’; (d) by semantic devaluation: *wasiƙa* ‘letter’ < *waṯīqa* ‘document’; *fitina* ‘sedition, troublesomeness’ < *fitna* ‘sedition’.

Unlike the above modifications, which developed spontaneously, the choice of an item such as *daƙiƙa* < *daqīqa* ‘minute’ to stand in Hausa for ‘second’ is the direct result of a decision by the Hausa Language Board established in the mid-1970s at Bayero University, Kano (for ‘minute’ Hausa uses the English loan *minti*). Arabic *ṯāniya* (expected to be realized in Hausa as *saniya*) ‘second’ seems to have been excluded to avoid confusion with Hausa *saniya* ‘cow’.

Not all documented loanwords in Hausa are commonly used. In fact, these loanwords range from being dormant to being very current. Generally, words are used with varying degrees of frequency according to the relevance of their semantic field to the actual conditions of life of Hausa speakers. Therefore, a number of loanwords recorded in the early Hausa dictionaries have now become obsolete because of the disappearance of the context in which they were used, e.g. *hindi* < *hindī* ‘a type of [Indian] sword’, *addaʾira* < *ad-dāʾira* ‘small dependent kingdom’. The largest number of infrequently used loanwords belong to the specialized vocabulary used among narrow scholarly circles or small social sectors, such as traditional scholars or medicine men. Most of the loans relating to the areas of grammar, astrology, minerals, and precious stones fall in this category, e.g. *liʾirabi* < *al-ʾiʿrāb* ‘declension’, *zaharatu* < *zuhrā* ‘Venus’, *zabarjad* < *zabarjad* ‘topaz’. Some loans have been replaced by other words, either native or borrowed from other languages, mainly English (or French), e.g. *munzari* < *minḍār* ‘eyeglasses’, replaced by *tabaru*; *alhanzir* < *al-xinzīr* ‘pig’, replaced by *alade*; *gahawa* < *qahwa* ‘coffee’, replaced by *kofi*.

With the spread of the Western type of education toward the end of colonial rule in the Hausa-speaking states (Nigeria and Niger Republic), Arabic was relegated to third place, behind English and French, as a source of borrowing for Hausa, but it was not completely displaced. After these states achieved independence in the 1960s, Arabic resumed its role, operating on an almost equal footing with English and French. From that time, borrowing from Arabic into Hausa has been, to a large extent, monitored by academic institutions (e.g. the Center for the Study of Nigerian Languages at Bayero University, Kano) and other relevant bodies, especially the Hausa Language Board and Hausa radio stations. Thus, many words have recently been borrowed from Arabic into Hausa within this framework to meet the needs of some new semantic fields. These include, for example:

i. The modern political system: *jamhuriyya* < *jumhūriyya* ‘republic’, *ƙuriʾa* < *qurʿa* ‘vote’, *milkin mallaka* < *mulk at-tamalluk* ‘colonization’
ii. The modern system of education: *jamiʾa* < *jāmiʿa* ‘university’, *ɗalibi/ɗaliba* < *ṭālib/ṭāliba* ‘male/female student’
iii. Mass media: *jarida* < *jarīda* ‘newspaper’, *mujalla* < *majalla* ‘magazine’, *basasa* ‘civil war’ < *(ḥarb al-)basūs* ‘a famous war between two tribes in Arabia before Islam’, *taʾaddanci* ‘terrorism’ < *taʿaddā* ‘to aggress’, *ƙalu-bala* ‘challenge, confrontation’ < *qālū balā* ‘they said yes in a friendly manner’
iv. Other fields: *lahani* ‘harm, damage’ < *laḥn* ‘tune; distortion in pronunciation’, *naʾura* < *nāʿūra* ‘machine’, *annoba* < *an-nawba* ‘epidemic, plague’

The recent borrowings display minimal phonological and morphological changes in the recipient language. From the semantic point of view, loans such as *basasa* 'civil war' and *ƙalu-bala* 'challenge, confrontation' show clearly that they were introduced by literate and knowledgeable people and endorsed by specialized (academic) institutions.

8. DIRECT CONTACTS BETWEEN ARABIC AND HAUSA

The third type of relation between Hausa and Arabic is that of close or direct contact, where speakers of the two languages live in contiguous proximity under conditions favoring close and active interaction. This is the situation under which Hausa communities in some Arab countries such as Sudan, Libya, and Saudi Arabia are found. Such a situation leads to the emergence of various sociolinguistic phenomena: bilingualism, intensive borrowing, interference, loan translation (calques), code-switching, and language shift (to Arabic). In some cases the recipient language undergoes a kind of → pidginization – though never creolization – before the total shift of its speakers to Arabic.

In Sudanese Hausa, for instance, all the abovementioned phenomena have been observed among its speakers to varying degrees according to a number of variables: place of living (urban/rural area), age, level of education, profession, etc. However, the great majority of the Sudanese Hausa can be said to be bilingual in their mother tongue and Arabic, irrespective of the above variables.

At the phonological level, Sudanese Hausa speakers differ from speakers in West Africa in the way they realize individual Arabic sounds. As a rule, all emphatic Sudanese Arabic consonants are replaced by their non-emphatic correspondents (*ṭ* > *t*, *ḍ* > *d*, *ṣ* > *s*, *ḏ̣* > *z*). But this pattern is not always regular because not all Sudanese Hausa speak Arabic with the same degree of fluency and perfection. So, their realization of these sounds may vary from the above pattern to perfect pronunciation and even to hypercorrection. Thus, for a Sudanese Arabic loanword such as *ʿadas* 'lentil', one may hear *ʾadas, ʿadas*, or even *ʿaḍaṣ*.

Borrowing of Sudanese Arabic words is intensified through social and cultural assimilation whereby entire groups of words pertaining to certain adopted customs or professions are borrowed wholesale. For example, the speech of those Sudanese Hausa communities that have adopted the Arab marriage system includes Arabic loanwords such as *xutuba* < *xuṭūba* 'engagement (gifts), *shela* < *šēla* 'marriage gifts', *darira* < *ḍarīra* 'plaster of local perfume applied on the head of the bridegroom', *ʾazuma* < *ʿazūma* 'feast', *shahar al-ʾasal* < *šahr al-ʿasal* 'honeymoon'. Such intensive use of Sudanese Arabic words in Sudanese Hausa extends over a number of grammatical categories, including:

i. Verbs: *ya xataba* < *xaṭab* 'he got engaged', *ya jaddada ruxsarsa* < *jaddad ruxṣatahu* 'he renewed his (driving) license'
ii. Adjectives: *ʾawira* < *ʿawīra* 'foolish [fem.]', *muʾaddaba* < *muʾaddaba* 'polite [fem.]'
iii. Adverbs: *koyis* < *kōyis* 'well', *tawwali* < *ṭawwāli* 'ahead; immediately', *ʾaslu* < *ʾaṣlu* 'in fact, actually'
iv. Prepositions: *lahaddi* < *laḥaddi* 'until'
v. Conjunctions: *lakin* < *lākin* 'but', *ʾaw* < *ʾaw* 'or'
vi. Interrogative pronouns: *malu* < *mālu* 'why?', *le* < *lēh* 'why'

Some of these words have already been well integrated in Sudanese Hausa, whereas many others appear merely as cases of interference on the way to integration.

Another remarkable phenomenon characterizing Sudanese Hausa is loan translation (calques), whereby Arabic concepts are expressed by Hausa words (literal translation). Even if such an expression does not contain any Arabic loanwords, it may still be difficult or even impossible for the West African Hausa speaker to understand the message. Loan translations involve mostly idiomatic expressions, such as *buɗe mishi*, translated from *aftaḥ lēhu* lit. 'open for him', i.e. 'forget about him, just neglect him'.

A kind of pidginized Hausa has been observed in the speech of educated Hausa speakers, especially females, living in urban centers, exemplified by utterances such as *ku faḍḍala su cikin ṣalun, ku gaddama musu sharbat da ḥalawa* (Arabic words underlined) 'Welcome them in the sitting room and offer them soft drinks and sweets'. In families where Hausa is spoken in this way, one usually finds the children speaking Arabic as their first language, and the succeeding generations may be expected to shift completely to Arabic.

BIBLIOGRAPHICAL REFERENCES

Abraham, Roy Clive. 1962. *Dictionary of the Hausa language*. London: University of London Press.

Abu-Manga, Al-Amin. 1999. *Hausa in the Sudan: Process of adaptation to Arabic*. Cologne: R. Köppe.

Baldi, Sergio. 1988. *A first ethnolinguistic comparison of Arabic loanwords common to Hausa and Swahili*. (= Suppl. to *Annali del Istituto Universitario Orientale di Napoli* 48:4.) Naples: Istituto Universitario Orientale.

——. 1989. "On the semantics of Arabic loan words in Hausa". *Current Progress in Chadic Linguistics*, ed. Zygmunt Frajzyngier, 285–301. Amsterdam and Philadelphia: J. Benjamins.

Bargery, G.P. 1934. *A Hausa-English dictionary and English-Hausa vocabulary*. London: Oxford University Press.

Greenberg, Joseph. 1947. "Arabic loan-words in Hausa". *Word* 3.85–97.

——. 1966. *Languages of Africa*. The Hague: Mouton.

McIntyre, Joseph and Hilke Meyer-Bahlburg. 1991. *Hausa in the media: A lexical guide*. Hamburg: Helmut Buske Verlag.

Meinhof, Carl. 1912. *Die Sprache der Hamiten*. Hamburg: L. Friederichsen.

Mischlich, Adam. 1906. *Wörterbuch der Hausaprache*. I. *Hausa-Deutsch*. Berlin: Reimer.

Newman, Paul. 1977. *Modern Hausa-English dictionary*. Ibadan: Oxford University Press.

Palmer, H.R. 1967. *Sudanese memoirs*, III. London: F. Cass.

Robinson, Charles Henry. 1913. *Dictionary of the Hausa language*. I. *Hausa-English*. 3rd rev. ed. Cambridge University Press.

Wexler, Paul. 1980. "Problems in monitoring the diffusion of Arabic into West and Central African languages". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 130.522–556.

AL-AMIN ABU-MANGA (University of Khartoum)

Head → X-bar Syntax

Hebrew → Ivrit

Ḥimyaritic

Ḥimyaritic is a Semitic language that was spoken in the mountains of Yemen during the first centuries of Islam. The name is derived from the tribe of Ḥimyar, whose origins lie in the region of Ẓafār (125 kms south of Sanaa) in the southern highlands. This tribal group gradually extended its power across the whole of Yemen and eventually exercized authority over the southwestern half of the Arabian Peninsula (1st century B.C.E.–6th century C.E.).

The adjective 'Ḥimyarite' was coined by traditional Arab and Islamic scholars and grammarians who preserved and transmitted a limited number of words and a small corpus of short texts that they had heard (proverbs, sayings, and conversations from daily life) or read (epitaphs), in order to emphasize the strangeness of this language to an Arab ear. The term 'Ḥimyarite' can also be applied to two inscriptions of pre-Islamic Yemen.

1. BEFORE ISLAM

For the pre-Islamic period, the term 'Ḥimyarite' cannot be used without being defined and explained. According to author or context, it possesses a political meaning (texts coming from the kingdom of Ḥimyar) or a linguistic one (written evidence exhibiting certain morphological and lexical features, as well as a specific syntax). Even in the latter sense, the adjective may refer to different varieties. One must keep in mind that in the pre-Islamic period the terminology was not yet fixed. The confusion is the result of a rather complex situation. Although the Ḥimyarites left a great number of inscriptions, sometimes of considerable length (Gajda 1997), these inscriptions were written in Sabaean, the language of the kingdom of Saba, of which the Ḥimyarites considered themselves the rightful heirs. However, the Sabaean language used by the Ḥimyarites was not perfectly regular. Even a superficial examination reveals lexical, morphological, and syntactic particularities that were infrequent before the 4th century C.E. and became increasingly common later on (Robin 1991:96). The *Sabaic dictionary* (Beeston a.o. 1984) uses a special symbol for inscriptions of the 380–560 C.E. period. To add to the confusion, authors apply various names for the Sabaean used by Ḥimyarites: 'Ḥimyaritic', 'Raydānitic' (after D̲ū Raydān, the name of the tribal confederation formed by the princes of Ḥimyar, whose name is derived from the Raydān palace in Ẓafār), 'Sabaeo-Ḥimyaritic', or 'Sabaeo-Raydānitic'.

Two inscriptions discovered in Ḥimyarite territory are not written in Sabaean. They come from peripheral areas (Qāniya and Širjān, located 100 and 150 kms, respectively, from Ẓafār, in the territory of the tribes of Radmān and Maḏḥā). It is therefore difficult to assess to what extent these inscriptions accurately reflect the language spoken by the Ḥimyarites, in the strict sense of inhabitants of Ẓafār. The inscriptions are of very different genres. The first text is a hymn of 27 rhyming verses in honor of the goddess Šams (Robin 2001:516–521, with bibliography), while the second one is a triumphal song written in the first person by a great lord (Robin 2001:523, n. 47). It is not even possible to establish whether they were written in the same language. The first text dates from around 100 C.E. and the second from around 300 C.E.

The most noteworthy features of these two inscriptions are the occurrence of three sibilants (*s*^{1}, *s*^{2}, and *s*^{3}, probably articulated as /š/, /ś/, and /s/) and the form *fʿlk* for the 1st and 2nd person singular perfect (Arabic *faʿaltu, faʿalta, faʿalti*), as in Sabaean. However, the article *hn-* precedes the noun, whereas in Sabaean the article is an *-n* suffix. Most of the lexicon is unknown from Ancient South Arabian texts and contains several words whose roots are not attested in any other Semitic language, not even in Arabic.

The difficulties increase when uncertain readings are taken into consideration. The Qāniya poem was carved in a very shallow manner, so that only a few passages have been deciphered securely. All in all, only two verses (2 and 13) have been interpreted convincingly. The first verse plausibly alludes to a ritual hunt, a sign of abundant rains:

> *(b-)ṣyd Ḫnw*^{n} *ʾt ns*^{3}*ḥk*
> 'During the hunt of Xinwān, you have bled a hundred [beasts]'

The second alludes to wine:

> *w(y)[n]mzr kn k-s*^{2}*qḥk*
> 'The vine became wine after you shone'

Note the rhyme *-ḥk* at the end of the poem's verses.

The reading of the Širjān text is not as problematic. Carved high on a rock face and always in the shade, it is still difficult to decipher from photographs, which explains the many mistakes in the early translations. However, the Qatabān expedition was able to read it directly in the field in 1992, by looking at the original stone copy and at a squeeze, and thus obtained a more thorough and exact translation. The beginning of the text is deciphered without problems and is easily understood:

> *S*^{3}*mdt S*^{1}*ʿd*^{m} *Yhs*^{1}*kr bn Ḥṣbḥ* | *s*^{2}*mk-ʾn b-Ẓnt s*^{1}*wr Bnʾ ms*^{1}*qt l-ḏ* | *w-brʾk-h ʾn ḏ-tns*^{1}*r ṯw-s*^{2}*qr…*
> 'Song of Saʿd^{um} Yuhaskir ibn Haṣbāḥ: I myself have erected in *Ẓnt* the *Bnʾ* wall in order to irrigate it; and I, in command [?, read *tansīr* or *tanassur*?] have built it myself, right up to its summit…'

Compare *s*^{3}*mdt* and *tns*^{1}*r* with Yemeni Arabic *samada* 'to sing' and *mansar* 'command [*qiyāda*]' (Al-Selwi 1987:113, 201–202).

2. After Islam

If the traditional Arab and Islamic scholars are to be believed, the Ḥimyaritic language was incomprehensible to a speaker of Arabic. This is stated as common knowledge by a large number of authors (Rabin 1951:49). The most frequently quoted anecdote states that a desert Arab visited a king of Ḥimyar, who politely invited him to sit down, saying *ṯib* (imperative of Sabaean *wṯb*, Hebrew *yašav*). The man understood 'jump!' (imperative of Arabic *waṯaba*) and obeyed by leaping out the window. The story's function is to illustrate the proverb with which it ends: 'Whoever enters Ẓafār has to "Ḥimyarize" himself' (*man daxala Ḍafār taḥammara*), which means that 'he has to learn Ḥimyaritic' (*fa-l-yataʿallam al-ḥimyariyya*) (according to the Yemeni author Našwān al-Ḥimyarī [d. 573/1178]; see Ahmad 1916:113, s.v. *w-ṯ-b*). Of course, this story does not reflect a true historical incident but is a pun, based on the fact that there are 'false friends' (identical words with very different meaning) between Ḥimyaritic and Arabic.

About 200 years before Našwān, another Yemeni, al-Ḥasan al-Hamdānī (d. after 360/971) wrote a description of the languages spoken in the Arabian Peninsula. Quite understandably, this description is much more detailed as to Yemen, an area that was of more interest to al-Hamdānī and, moreover, exhibited a greater diversity of dialects. For al-Hamdānī, whose native language was Arabic (his family came from al-Marāšī, 130 kms north of Sanaa, from the tribe of ʾArḥab, to which he claimed to belong), the language of reference was Standard (*faṣīḥ*) Arabic. The other languages were defined by him according to their divergence from this model. The most important text is found in the *Ṣifa jazīrat al-ʿArab* (D. Müller 1877:I, 134–136). It was studied and translated by Rabin (1951:43–44), Robin (1991:103–105, 83 with map), and Belova (1996:15–16).

Concerning the tribe of Ḥimyar in the narrow sense, al-Hamdānī observes that "from Ḥaql Kitāb to Ḏamār, pure Ḥimyaritic is spoken, which is hard to understand [*al-ḥimyariyya al-quḥḥa al-mutaʿaqqida*]". The last word, *mutaʿaqqid*, is translated by Rabin (1951:45) as 'halting', i.e. 'with Ḥimyaritic rhythm and intonation', which could mean 'without tonic accent' (Rabin 1951:49). Belova (1996) renders the same word as *maloponjatnyj* 'incomprehensible'.

Al-Hamdānī is the author with the greatest amount of information about the Ḥimyaritic language. Drawing from the *Kitāb al-ʾiklīl*, of which only Books I and II (genealogies of Ḥimyar), VIII (antiquity of Yemen), and X (genealogies of Hamdān) have survived, he cites many Ḥimyaritic texts – all of them quite short – and comments on certain traits of the language. No manuscript of Book IX, *Fī ʾamṯāl Ḥimyar wa-ḥukmi-hā bi-l-lisān al-ḥimyarī wa-ḥurūf al-musnad* (Ḥimyarite wisdom and proverbs in Ḥimyaritic language and South Arabian script), has yet been found.

The most recent observations are by a Yemeni ruler of Turcoman-Syrian origin, the Rasulid ʾAbū l-Fatḥ ʿUmar ibn al-Malik al Muẓaffar Yūsuf ibn ʿAlī ibn Rasūl, often called al-Malik al-ʾAšraf, who reigned from 1295 to 1297 C.E. In a book written around 1271, the author characterizes the Ḥimyaritic language as having a "strong foreign deformation" (*ʿujma*), resulting from the combined influence of Abyssinian and Arabic. He illustrates this with a surprising anecdote: while traveling in Yemen, a person with good knowledge of Arabic heard Ḥimyaritic spoken and thought it was Abyssinian (Saliba 1985). Al-Malik al-ʾAšraf reproduces the sentence that this person heard, <ḥssnʾ-h yxsm mʿ-nʾ sw sʾmnʾ (f-)dw ʾsynʾ-h>, and translates it as 'they inquired about the man so that he would eat with them until satisfaction, but they did not find him' (*iltamasū r-rajul yaʾkul maʿa-hum ʾilā ʾan saʾimū fa-lam yajidū-hu*). Walter Müller (1989), who reconstructed that Ḥimyaritic text, translates it as 'we looked for him so he could eat with us, until he could do so no more; but we have not found him'.

The Austrian David Heinrich Müller was the first scholar interested in the Ḥimyaritic language as presented by Arab authors from the Islamic era. In 1877, he published a brief inventory of small texts taken from al-Hamdānī; these are short quotations, proverbs, or alleged epitaphs, all of them more or less corrupted. This seminal study was the basis for the work of Chaim Rabin (1951:42–53), who suggested that Ḥimyaritic did not differ essentially from Arabic, except for certain features that appeared exotic to speakers of Arabic. According to him, close scrutiny revealed that these specific features were few in number and of little consequence, even though the Arab philologists had made the most of them. Fifty years earlier, Landberg (1898:110–119) had reached the same conclusion.

Since Rabin's study, which remains generally valid, several unknown works of al-Hamdānī have been published (not always, however, in a critical edition), offering new quotations and allowing a better understanding of those already known. An inventory and study of these quotations was published by the Yemeni scholar ʾAsmahān al-Jahrū (1987). To these may be added the sentence reproduced by al-Malik al-ʾAšraf that was mentioned above (Saliba 1985; W. Müller 1989).

Knowledge of the Ḥimyaritic calendar has also improved. The names of the months were already known in consonant graphs from pre-Islamic inscriptions. Their vocalization and order were established with the help of a medieval Yemeni Arabic poem relating to agriculture, and subsequently by a work on astronomy, the *Kitab at-tabṣira fī ʿilm an-nujūm* by al-Malik al-ʾAšraf, who mentions (in Chap. 32) the corresponding months in the Syrian solar calendar (Beeston 1974; al-ʾAkwaʿ 1981; Saliba 1985; Varisco 1994). The above-mentioned sentence in Ḥimyaritic is found in al-Malik al-ʾAšraf's text.

Lastly, knowledge of Yemen's historical geography has much improved. This has enabled scholars to map more accurately the spatial and tribal distribution of linguistic phenomena, which was still very approximate in Rabin's time.

The study of these materials is facilitated nowadays by a better understanding of ancient inscriptions, especially those in Sabaean, as shown by the publication of the *Sabaic dictionary* (Beeston a.o. 1984). Studies on the Yemeni lexicon in Classical Arabic texts (Al-Selwi 1987) or in contemporary dialect (Piamenta 1990; ʿIryānī 1996) also contribute to the understanding of these materials.

Robin's (1991) study, which surveys the languages of Arabia, has a note summarizing what is known about the Ḥimyaritic language from the Arab sources (Robin 1991:107–108).

As already underlined by Rabin, all texts in Ḥimyaritic quoted by Islamic authors seem to exhibit the same differences with Arabic: the negation *daw* (Arabic *lā*); the relative *ḏī* (Arabic *allaḏī*); the definite article *an-* or *am-* (Arabic *al-*); the ending *-k* of the 1st and 2nd person singular perfect of the type *faʿalku, faʿalka, faʿalki* (instead of *-t* as in Arabic); the regular endings *-an/-anna* of the imperfect (as in the → energicus in Arabic); and finally, words belonging to an exotic lexicon, such as *bahala* 'to say', *halla* 'to be', *šaʾama* 'to buy', *ʾawwala* 'to bring', *ʾasiya* 'to find', *ṯaw* 'until', *ḥinj* 'as'.

Yet, the real impetus for research into Ḥimyaritic was given by Anna Belova, a Russian scholar who published several short studies focusing on aspects of this language and has now written the first comprehensive monograph on the language, *The Himyaritic language: Regional studies on the history of the Arabic language* (1996). She has discovered numerous data related to Ḥimyar in Arabic sources, and undertook the ambitious project of presenting a synthetic and comparative view of Ḥimyaritic phonetics and phonology (Chap. 1), morphology (Chap. 2), and syntax (Chap. 3), supplementing these with tables comparing the Ḥimyaritic data with other Semitic languages.

Admittedly, Belova's work is only a first step. It could be argued that a simple description of the language would have been safer at this stage. It is easy to see that quotations from Ḥimyaritic offer very different readings according to the manuscripts. Belova does not give the full recorded body of material with a critical edition for each of them but only provides details for five examples, offering her translation without showing the possible variants (Belova 1996:139–169). In order to have an idea of the different readings, one has to analyze and explore the commentaries of the first three chapters. Not surprisingly, the selected readings often differ from those chosen by other scholars, which obviously weakens the linguistic reconstruction (compare, for instance, the text of Daybajat in Robin 1991:107 with that in Belova 1996:151). The publication's date is deceptive, as the work must have been available in draft form for a long time and must have been stored for a considerable period of time on the shelves of the Academy. Consequently, several of the studies mentioned above, such as W. Müller (1989), Robin (1991), and Varisco (1994) do not appear in the bibliography (for a supplement to the bibliography, see Belova 1998).

There is nowadays a certain consensus about the intelligibility of Ḥimyaritic during the first centuries of Islam. It should not have presented extraordinary difficulties for an Arab with an agile mind, even if certain varieties of the language may have been particularly difficult to understand; al-Hamdānī describes one of these varieties as 'unintelligible' (*ġutm; ʾIklīl* II, 246). This is also the case for some of the contemporary dialects, especially those of the Tihāma and the valleys of western Yemen. This alleged lack of mutual intelligibility between certain varieties of Ancient Arabic is not without parallel: the Arabic spoken by the Jews of the Ḥijāz during the 7th century C.E., called *yahūdiyya*, was also called 'gibberish' (*raṭan*; Gil 1984:205–206, esp. n. 10 for full references).

One last point remains to be explained: the relationship between the Ḥimyaritic of the Islamic period and the Ḥimyaritic of the ancient inscriptions. This relationship must have been a close one, as demonstrated by common features, such as the article *hn-* in the Qāniya inscription and *an-* (but also *am-*) in the works of traditional Islamic scholars (Belova 1996:115–121). Yet, substantial differences are apparent, which may have been the result of linguistic evolution. This is the case with the system of sibilants, of which there are three in the Qāniya and Širjān inscriptions while only two seem to occur in the Ḥimyaritic texts of the Islamic period. This is a field of research that still needs more investigation. Likewise, Ḥimyaritic loanwords in Arabic may be mentioned; they are sometimes referred to (Rabin 1984; Al-Selwi 1987) but have never been treated comprehensively.

3. A ḤIMYARITIC SAYING IN ANCIENT MANUSCRIPTS

An example from Book X of al-Hamdānī's *ʾIklil* may serve to illustrate the difficulties inherent in the study of the Ḥimyaritic dialect of the Islamic period. There are two editions of this work, one by an Egyptian scholar (Muḥibb ad-Dīn al-Xaṭīb) and the other by a Yemeni scholar (Muḥammad al-ʾAkwaʿ). The former could be called a critical edition, even though it is based on only three manuscripts. The latter derives from the first and contains many corrections, often justified by the author's familiarity with Yemeni toponymy and onomastics, but also demonstrating certain biases on the editor's part.

Al-Hamdānī quotes a saying by a certain al-Labaxī (also mentioned in Book VIII of the *ʾIklil*, but otherwise unknown), who appears to have been one of his masters.

> The Ḥimyarite says in the Ḥimyaritic language, quoting the *ʾanwāʾ*: *ʾqsmn ʾnjwm ʾrbʿ ḏw tġyb lw yrwy sd btʿ* [al-Xaṭīb's edition] *ʾqsmn ʾm ʾnjm ʾm ʾrbʿ dw tġyb lw yrwy sd btʿ* [al-ʾAkwaʿ's edition] '[The Bataʿ dam] is located between Ḥāz [a village 27 kms northwest of Sanaa] and Bayt Dafaʿ [a place not yet identified]'; *ḏū/dū* means 'no' and *lw* 'until'. Al-Ḥasan [al-Ḥamdānī] quotes [this saying] in Book IX of *al-ʾIklil*, and it is to be understood as follows: 'The four stars – they constitute [the constellation] of aṣ-Ṣawāb – will not disappear at the moment of the dawn prayer, before the Bataʿ be fed by strong March rains, and this is the regular pattern'.

Belova (1996:166–169) uses the text of al-Xaṭīb with only minor changes. There are only two ways to reconstruct the original text of the saying. The first is to compare the different manuscripts, something that warrants a specific focus for research; the second is to restore the original script by assuming that it derives from words known from pre-Islamic inscriptions and that some Arabic letters were subjected to normal alterations (change in diacritical points and transformation of a letter into one with a similar shape).

The article in this saying is probably *an-* (Robin 1989:5), a form mentioned by both al-Hamdānī and Našwān (see Ahmad 1916, root *ʿ-b-b*: "Among these [Ḥimyarites] there are some who change the *l* of the article into *mīm* and others who change it into a *nūn*"). One must therefore correct *ʾjwm ʾrbʿ* into *ʾ(n)-njwm ʾ(r)-rbʿ*.

The negation *ḏw* is probably written this way because of contamination with the pronoun *ḏū* and should therefore probably be read as *dw*. Compare with this the negation *dw* in late Ḥimyaritic inscriptions and the negations *daʾ*, *daw*, and *duwway* attested in modern southern Yemeni dialects (Behnstedt 1985:170, table 117).

The conjunction *lw* may be compared with the preposition/conjunction *ṯw* of Ḥimyaritic inscriptions (three occurrences published and a fourth one still unpublished), which means 'until, up to' (Beeston a.o. 1984; to the references quoted there add the Širjān text mentioned above). The unpublished inscription commemorates the conclusion of the hydraulic works in Wādī Ḍuraʾ 'from its upper to its lower course' (*mn ʿly-hw ṯw s*^{1}*fl-hw*; Ḍuraʾ 4/3–4). In contemporary Yemeni dialects, *ṯw* survives under the form *taww* 'until, toward' (Deboo 1989:258; Piamenta 1990:54). The passage from *ṯ* to *t* (*ṯw, taww*) to *l* (*lw*) can be explained without much difficulty by the particularities of the Arabic script: there may have been a certain measure of confusion between *ṯ/t*, written without diacritical dots, and a short *l*.

The saying quoted by al-Hamdānī should probably be read as *ʾqsmn ʾ(n)-njwm ʾ(r)-rbʿ dw tġyb tw yrwy sd btʿ*. Accordingly, its meaning probably was 'I swear that the four stars will not set until the Bataʿ dam will be filled'. There is still one obscure point, the precise analysis of *qsmn*, for which al-Hamdānī does not give any explanation. Is the verb *ʾaqsama* a 1st person singular imperfect form with the ending *-n*, which is mentioned by Arab lexicographers as typical of the Ḥimyaritic language (*ʾuqsiman* 'I swear')? Or is it a 1st person plural perfect (*ʾaqsamnā* 'we have sworn')? The former seems more likely, because al-Hamdānī prefers to give examples that illustrate the largest possible number of Ḥimyaritic particularities.

Bibliographical references

Primary sources

Hamdānī, *ʾIklīl* = Lisān al-Yaman ʾAbū Muḥammad al-Ḥasan ibn ʾAḥmad ibn Yaʿqūb al-Hamdānī, *Kitāb al-ʾiklīl min ʾaxbār al-Yaman wa-ʾansāb Ḥimyar*. Vol. II. Ed. Muḥammad ibn ʿAlī al-ʾAkwaʿ al-Ḥiwalī (= *al-Maktaba al-Yamaniyya*, 3.) Cairo: as-Sunna al-Muḥammadiyya, 1386/1967. Vol. X. Ed. Muḥibb ad-Dīn al-Xaṭīb. Cairo: as-Salafiyya, 1368/1948–1949. Vol. X. Ed. Muḥammad ʿAlī al-ʾAkwaʿ. Sanaa: Maktabat al-Jīl al-Jadīd, 1410/1990.

Hamdānī, *Ṣifa* = Lisān al-Yaman ʾAbū Muḥammad al-Ḥasan ibn ʾAḥmad ibn Yaʿqūb al-Hamdānī, *Ṣifa jazīrat al-ʿArab*. Ed. David Heinrich Müller, *al-Hamdânî's Geographie der arabischen Halbinsel*. 2 vols. Leiden: E.J. Brill, 1884–1891. (Repr., 1968.)

Secondary sources

Ahmad, Azimuddin. 1916. *Die auf Südarabien bezüglichen Angaben Našwān's im* Šams al-ʿulūm, *gesammelt, alphabetisch geordnet und herausgegeben*. (= *E.J.W. Gibb Memorial Series*, 24.) Leiden: E.J. Brill; and London: Luzac.

ʾAkwaʿ, Muḥammad ʿAlī…al-Ḥiwalī al-. 1981. “Qaṣīdat al-baḥr an-naʿāmī fī l-ʾašhur al-ḥimyariyya wa-mā yuwāfiqu-hā min ʾaġḏiya”. *al-ʾIklīl* (Sanaa), I/3–4, 9–17.

Al-Selwi, Ibrahim [ʾIbrāhīm aṣ-Ṣilwī]. 1987. *Jemenitische Wörter in den Werken von al-Hamdāni und Našwān und ihre Parallelen in den semitischen Sprachen*. Berlin: D. Reimer.

Beeston, A.F.L. 1974. “New light on the Ḥimyaritic calendar”. *Arabian Studies* 1.1–6.

——, M.A. Ghul, Walter W. Müller, and Jacques Ryckmans. 1984. *Sabaic dictionary (English – French – Arabic) – Dictionnaire sabéen (anglais – français – arabe) – al-Muʿjam as-sabāʾī (bi-l-ʾinjilīziyya wa-l-firansiyya wa-l-ʿarabiyya)*. Louvain-la-Neuve: Éditions Peeters; and Beirut: Librairie du Liban.

Behnstedt, Peter. 1985. *Die nordjemenitischen Dialekte*. I. *Atlas*. Wiesbaden: L. Reichert.

Belova, Anna G. 1996. *Himʾjaritskij jazyk: Arealnye issledovanija k istorii arabskogo jazyka* [The Ḥimyaritic language: Regional studies on the history of the Arabic language]. Moscow: Rossijskaja Akademija Nauk, Institut Vostokovedenija.

——. 1998. “Nazvanija sezonov v južnoj Aravii: Opyt sravnitel’nogo analiza” [The names of the seasons in South Arabia: An attempt at comparative analysis]. *Epigrafika Vostoka* 25.5–15.

Deboo, Jeffrey. 1989. Jemenitisches Wörterbuch: Arabisch – Deutsch – English. Wiesbaden: O. Harrassowitz.

Gajda, Iwona. 1997. *Himyar gagné par le monothéisme (IV^{e}–VI^{e} siècle de l’ère chrétienne): Ambitions et ruine d’un royaume de l’Arabie méridionale antique*. Thèse de doctorat, Université de Provence.

Gil, Moshe. 1984. “The origin of the Jews of Yathrib”. *Jerusalem Studies in Arabic and Islam* 4.203–224.

ʿIryānī, Muṭṭahar ʿAlī al-. 1996. *Fī l-luġa wa-t-turāṯ: Ḥawla mufradāt xāṣṣa min al-lahjāt al-yamaniyya*. (= *al-Muʿjam al-yamanī*, I). Damascus: Dār al-Fikr.

Jahrū, ʾAsmahān al-. 1987. *Les antiquités du Yemen dans l’oeuvre d’al-Hamdānī*. Ph.D. diss., Université de Paris-1.

Landberg, Carlo (comte de). 1898. *Arabica*, V. Leiden: E.J. Brill.

Müller, David Heinrich. 1877. “Südarabische Studien”. *Sitzungsberichte der philosophisch-historischen Classe der kaiserlichen Akademie der Wissenschaften* 86.103–182. [“II. Die ḥimjarischen Schriftdenkmäler in der arabischen Überlieferung”, 115–134.]

Müller, Walter W. 1989. “Ein himjarischer Satz bei al-Malik al-Ašraf”. *On both sides of al-Mandab: Ethiopian, South-Arabic and Islamic Studies presented to Oscar Löfgren on his ninetieth birthday, 13 May 1988, by colleagues and friends*, II, 49–43. Istanbul: Swedish Research Institute in Istanbul.

Piamenta, Moshe. 1990, 1991. *Dictionary of post-Classical Yemeni Arabic*. 2 vols. Leiden: E.J. Brill.

Rabin, Chaim. 1951. *Ancient West Arabian*. London: Taylor’s Foreign Press.

——. 1984. “On the probability of South-Arabian influence on the Arabic vocabulary”. *Jerusalem Studies in Arabic and Islam* 4.125–134.

Robin, Christian J. 1989. Review of: Al-Selwi (1987). *Bulletin Critique des Annales Islamologiques* 6.3–6.

——. (ed.). 1991. *L’Arabie antique de Karib ʾîl à Mahomet: Nouvelles données sur l’histoire des Arabes grâce aux inscriptions. Revue du Monde Musulman et de la Méditerranée* 61:3.

——. 2001. “Les inscriptions de l’Arabie antique et les études arabes”. *Arabica* 48.509–577.

Saliba, George. 1985. “A medieval note on the Ḥimyarite dialect”. *Journal of the American Oriental Society* 105.715–717.

Varisco, Daniel M. 1994. *Medieval agriculture and Islamic science: The almanac of a Yemeni sultan*. Seattle and London: University of Washington Press.

Christian Julien Robin (CNRS,
Laboratory for Ancient Semitic Languages)

Hindi → Urdu/Hindustani

Hindustani → Urdu/Hindustani

History of Arabic

This entry describes the evolution of the Arabic language through all its phases, paying special attention to the causes and implications of the changes which have taken place in the language. In order to place Arabic in its broad context and to offer a diachronic insight into the history of a language which is spoken today by approximately 200 million people worldwide and is the preferred religious language of all Islamic countries, it is necessary to consider its historical setting and present-day situation.

From the outset it is necessary to consider the scope and limits of the term ‘Arabic language’. Should it comprise mainly the Classical variety in both its ancient and modern stages, or should the term be applied to all of the known varieties of the language, including vernacular or colloquial varieties past and present? The answer to this question, quoting Holes, may be that “much of the work done in Arabic, it seems to me, has suffered, and continues to suffer, from a reluctance on the part of native and western linguists alike to recognise the contemporary linguistic realities of the Arab World and take these as the starting point for their descriptive and analytical researches… [and to] view the language as an integrated whole which is, after all, how its native speakers use, experience and think about it” (1995:

viii). It should not be forgotten that Standard Arabic has always been subject to the influence of other varieties of the language. To study Standard or Classical Arabic as independent entities means to disregard the broad range of interaction among the different varieties of the language from the very beginning to the modern stages of Arabic. The interdependent approach is taken, among others, by the general surveys of Versteegh (1997) and Ferrando (2001). A history of Arabic, then, must emphasize that Arabic is a dynamic language, rather than a monolithic, immutable reality. When dealing with the different varieties of the language complex, it seems inaccurate to rely on the Classical or Qurʾānic language as a model from which all other varieties have been derived. One frequently encounters unsubstantiated claims such as "This X form in a given Arabic dialect comes from the Classical Y through a process of change (or even corruption)". It would seem more precise to attempt an integration of facts in a less treelike linguistic model, emphasizing alternative explanations such as language drift, convergence and/or divergence processes, and adstratal, substratal, and hyperstratal influx.

Nonetheless, describing the history of Arabic and its evolution through all its stages is not an easy task. The main reason is that the Qurʾānic variety and the canonized Classical Arabic became a model to imitate and also a pattern not to deviate from. As the language of Islam, Classical Arabic is considered the language of God; hence, a reluctance to change the contents of this linguistic inheritance persists. This does not mean, of course, that Arabic remains today as it was in the 7th century. However, it is undeniable that, compared with other language groups (e.g. Romance and Germanic), language change and variation in Standard Arabic has been reduced to a minimum, especially in aspects of morphology and phonology.

1. Arabic in its linguistic setting: The position of Arabic within the Semitic phylum

Arabic belongs to the group of the so-called → Semitic languages, formed by about 70 different languages spoken mainly in the Middle East. The temporal frame of this language phylum dates from the 3rd millennium B.C.E. and continues up to the present. This group includes some well-known and widespread languages (Hebrew, Aramaic, Arabic), together with many varieties no longer extant or very poorly attested to. Within this broad context, Arabic may be placed, together with Ethiopian, Modern South Arabian, and Epigraphic or Ancient South Arabian (Sayhadic in the terminology used in Hetzron 1997:241), in the Southwest Semitic languages, according to the traditional classification supported in recent times by scholars like Diem (1980) and Corriente (1996:12–13; 2003). Others, like Hetzron (1974), Voigt (1987), and Rodgers (1991), prefer to separate Arabic from the → South Semitic languages and place it in a central position, alongside Hebrew, Aramaic, and other close varieties. Both positions account for the fact that Arabic, in addition to sharing important isoglosses linking it to the basic linguistic structure of the South Semitic languages, shows some particular evolutionary features akin to North Semitic languages (→ Northwest Semitic). Its peripheral location may have given Arabic these ambivalent, particular features, typical of a transition-like status. Nevertheless, while Arabic is a Semitic language and the structure of the language includes the main features of other Semitic languages, it differs from other Semitic varieties in that it shows a wide range of presumably ancient linguistic forms that can be traced back to an early stage of Semitic. Arabic may be labeled, in this respect, a conservative Semitic language. However, the fact that Arabic is the best-known of the Semitic languages does not necessarily mean that it is equivalent to Proto-Semitic.

2. Early Arabic (or Proto-Arabic)

The Arabian Peninsula, especially in its southern region, was inhabited by the South Arabian people, who developed advanced forms of civilization between the 4th century B.C.E. and the 6th century C.E. However, there are records of several other groups in the same region speaking different languages at the beginning of the 2nd millennium B.C.E., and even more to the middle and the north of the peninsula. These groups, who, according to the sources, were identified as → *ʿArab*, began making use of other languages (South Arabian, Greek, and Aramaic) in writing their inscriptions and monuments. The presence of their own Arabic language may be traced in these foreign inscriptions, according to Müller (1982:26–28), Diem (1973), and Robin (1993:122–125). From the 2nd century B.C.E. onward, however, inscriptions attest to the use of these very language varieties by these Arabs. According to Robin (1993), these inscriptions allow for a division of → Proto-Arabic into four varieties, Ṯamūdic, Liḥyānitic, Ṣafāʾitic, and Ḥaṣāʾitic, all of them showing an *h*-prefixed article (→ North Arabian). Later, from the 3rd century C.E. onward, a variety of inscriptions are attested to in various parts of the Arabian Peninsula. The language of these records is more akin to the language patterns of Classical Arabic in its known structure. In these inscriptions the first attestations of an *l*-prefixed article appear, as well as a causative prefix ʾ- and other features closer to Standard Arabic. The most famous of these inscriptions is that of an-Namāra, 120 km southeast of Damascus, dating from the 4th century C.E. Robin (1993:116–117), who carefully studied its language, has demonstrated the existence of a close affinity between an-Namāra Arabic and Classical Arabic, the most relevant common features being an article *ʾl-*, the negational particle *lam* with the imperfect verb, and a case of internal verbal complement. Other features, however, link this inscription to the Aramaic language, which was dominant in the region at that time (→ Old Arabic). It should be noted that South Arabian or Nabataean alphabets were generally used in these inscriptions, with the exception of some late (4th and 5th centuries C.E.) Syrian and Jordanian inscriptions in the new Arabic alphabet, which, according to most scholars, e.g. Gruendler (1993), is probably based on Nabataean script (→ Arabic alphabet: origin.)

3. The Arabian Peninsula just before the rise of Islam: Linguistic panorama

The next stage in the development of the Arabic language is represented by the variegated Arabic varieties employed throughout the Arabian Peninsula just before the rise of Islam, that is to say, during the 5th and 6th centuries C.E. (→ pre-Islamic Arabic). The speakers of these varieties were called the Arabs (→ *ʿArab*), pertaining to numerous tribes, and their languages are generally labeled as *kalām al-ʿArab*, the starting point for the Arabic of the *Qurʾān* and the pre-Islamic poetry, the two major sources for the Classical Arabic language. Since the available data concerning the dialects of pre-Islamic Arabia are fragmentary, it is not easy to set up a dialectal map of the region. The valuable work of Rabin (1951) stands out as a basic tool for an understanding of the particular features of some Western varieties where a basic distinction between Eastern Arabic and Western Arabic is postulated. Eastern Arabic, represented by the dialect of Tamīm, is said to be closer to the language of poetry, whereas the Western dialect, represented by the dialects of the Ḥijāz, including the famous variety of Qurayš, is supposed to be the basis for the language of the *Qurʾān*. Another important group is that of the South or Yemenite dialects. The principal differences between the three groups are found in phonological and lexical aspects, although, as clearly stated by as-Sāmarrāʾī (1994:16–18), the data remain unclear and even contradictory in many cases, which makes a complete description of the broad linguistic panorama of pre-Islamic Arabia particularly difficult. Some of the most commonly mentioned features are the insertion of a vowel in final consonant clusters (West); vowel harmony or assimilation (East); → *ʾimāla* or fronted pronunciation of [a] toward [e] and [i] (East); the presence of the phoneme [e] (West); voiceless realization of [q] in the East but voiced in the West; the loss of the glottal stop (*hamz*) in the West as opposed to its retention in the East; imperfect prefix with *-i-* (the so-called → *taltala*) in the East but with *-a-* employed in the West.

Other features certainly exist, but the absence of a complete list makes an exhaustive survey of the linguistic situation of ancient Arab dialects impossible at present. The question of whether these Arabic varieties were conservative (retaining the full declension system, called → *ʾiʿrāb*) akin to the Classical synthetic type, or innovative (not showing a coherent declension system), akin to the Neo-Arabic language type, has led to considerable debate among scholars. The most widely accepted view is that both types coexisted and interacted during this period. The synthetic varieties were, however, considered to be more 'pure', rapidly becoming more prestigious throughout the Arabian Peninsula.

This accounts for the adoption of these varieties in the two major works of pre-Islamic Arabic, the *Qurʾān* and pre-Islamic poetry.

4. Qurʾānic Arabic and the Arabic of pre-Islamic Poetry

The primary source for the study of the Standard Arabic variety consists of a corpus of poetry composed during the 5th and 6th centuries C.E. and at the beginning of the 7th century by Arabs originating from tribes located in different places in the Arabian Peninsula, with a predominance of poets coming from eastern regions. This corpus was orally transmitted until it was finally recorded and encoded in the 8th century. The most outstanding feature of this poetry is its full linguistic homogeneity: all the recorded material conforms to morphological and syntactic patterns with no apparent deviations. Based on this observation, Ṭāhā Ḥusayn (1926:32–35), among other scholars, did question its authenticity. If, as has been clearly demonstrated, different tribes employed different varieties of Arabic, why was this not reflected in their poetry? Ḥusayn was inclined to believe that this material underwent reelaboration after Islam. However, an alternative answer to the question suggests that the poets resorted to a supratribal register of the language, a pre-Islamic koine that allowed for communication among all Arabs (→ poetic koine). The language of this poetry is essentially the same as the language of the *Qurʾān* in its morphology and syntax. Some phonetic differences presumably exist, e.g. the status of *hamz*, although the absence of contemporary records prevents further in-depth analysis in this domain. Lexical material also differs slightly: pre-Islamic poetry reflects the nomadic life in the Arabian deserts and plains, hence the archaic flavor of a good deal of the lexicon.

Although the *Qurʾān* stands out as the first monument of Arabic prose, several other pieces of prose, some of which can be traced back to pre-Islamic times, show linguistic and stylistic forms similar to those found in the *Qurʾān* (see Mannāʿ 1993). The *Qurʾān* is reputed to be the highest and purest performance/reflection of the Arabic language and thus impossible to imitate or translate into other languages (→ *ʾiʿjāz*). As most Arab speakers see in the language of the *Qurʾān* the perfect model of → *al-ʿarabiyya*, a consensus arose that this language variety had to be preserved and cared for in order to avoid the corruption of the Arabic language. Its rhymed prose (→ *sajʿ*), in a generally concise style, brought to the Arabic language a new stock of lexicon and phraseology that set it apart from the Bedouin ambiance of pre-Islamic poetry and allowed new tools to develop for the creation of new lexical items, such as the Arabicization of loanwords, new morphological derivations, and semantic extensions. The fact that it was originally meant to be recited gives it a new rhythm, leaving a powerful acoustic impact on every audience. Many scholars accept that the language of the *Qurʾān* is similar to that of pre-Islamic poetry, a sort of unified register or *luġa muwaḥḥada*, leaving aside some orthographic devices and the question of the *hamz*, a symbol which was probably added at the time of the official recording of the *Qurʾān* at about 650 C.E., under the auspices of ʿUṯmān, the third caliph of Islam. A basic reference for the study of the textual history and the language of the *Qurʾān* is Nöldeke (1938). As to the status and identification of the language of *Qurʾān*, various theories have been put forward. Some scholars see it as the very dialect of Qurayš (the tribe of the Prophet), while others consider it an extended variety based on the Ḥijāzī dialect. Others point out that some ingredients of the prestigious Eastern Arabic varieties are also present in the Qurʾānic text. Whatever its origins, Qurʾānic Arabic became, in a very short time, the undisputed reference model for good Arabic, and this is still the case today (→ *Qurʾān*).

5. The classical period: Codification of the language corpus and grammar

During the 7th century, the new Islamic empire grew very quickly, reaching the farthest regions of the known world. The development of Arabic as the official language of this empire, a step not taken until the times of ʿAbd al-Malik, around the end of the century, required some kind of language standardization. First, a complete set of orthographic norms was required when recording the sacred text of the *Qurʾān* and the pre-Islamic poetry. Second, Arabic had to go through a transition stage from being a language with no cultural tradition to being a powerful new cultural tool for a growing empire. Arab grammarians played a crucial role in this process by creating an accurate description of Arabic and by setting out linguistic rules so as to avoid corruption or mistakes. Of great importance were the recording and study of the lexical material, giving rise to several impressive and influential works. In addition, lexical borrowings from prestigious neighboring languages, such as Persian, Greek, and Aramaic, were promoted, giving Arabic the flexibility and prestige it needed. The development of a written literary and scientific style relied on the emergence of a literary prose corpus consisting of translations from Persian, Aramaic, and Greek. Arabic became the language of prestige throughout the empire and was employed for all religious, cultural, and administrative purposes. Other languages such as Persian, Greek, or Aramaic were finally discarded by Arab people, who clearly identified themselves as Arab speakers.

Arab grammarians produced impressive works dealing with practically all aspects of the language. Their main task was to record and preserve all data stemming from the speakers of the true or pure Arabic, that is to say, *kalām al-ʿArab*, the language of the Bedouin of Arabia. The language varieties of the urban centers were considered corrupt and were therefore not included in the grammarians' material. As the Bedouin came more and more in contact with speakers of other varieties of Arabic, their language also underwent a process of corruption. Lexicographers and grammarians tended not to accept Bedouin data stemming from the 4th century A.H. and later. This meant that the linguistic inheritance of Classical Arabic consisted of a closed corpus. On the other hand, the dialects of Arabic evolved in a more natural way, as they did not show the same degree of reluctance to change or to incorporate foreign borrowings.

6. Post-Classical Arabic: Neo-Arabic and diglossia

As indicated by Versteegh (1997:93–98), important changes occurred in the Arabic language as a consequence of its spread over an enormous territory and its contact with many different languages (South Arabian, Persian, Greek, and Berber). The acquisition of Arabic by a large number of speakers of other languages had a considerable impact on the language which the conquerors brought with them, especially in the formation of the different colloquial varieties. However, the process of pidginization, creolization and de-creolization which Versteegh (1984) puts forward is unlikely to have occurred on a large scale. Rather, it is probable that Arabic, as the language of power and prestige, was retained by the Arab rulers and learned by the local population. The substratum effect, important as it is, should be kept in proportion, according to Diem's contentions (1979; → substrate).

An important issue is that of the emergence of the so-called Neo-Arabic language type, an analytic language structure attested in most dialects of Arabic. This variety would be one of the two sides of a linguistic situation known as → diglossia, that is to say, the use of a High variety for formal and literary purposes along with a Low variety for everyday communication. The old type of Arabic is a synthetic one, bearing, among other features, nominal declension and internal morphological devices, in contrast to the Neo-Arabic type, which is more analytic, having no declension and showing a strong tendency to make use of external morphemes for the expression of syntactic relations between words. In order to trace the coexistence of these two varieties diachronically, the emergence of the new type of Arabic needs to be considered. Two different theories have been proposed to account for this. The first one holds that diglossia was present well before Islam, with the old type being restricted to literary and formal domains and contexts. The second theory claims that diglossia did not emerge until the spread of Islam, as a consequence of the linguistic contact between Arabic and other languages. According to this second approach, well argued by Blau (1977), the Old Arabic type was the language of all Arabs for almost all purposes before Islam. As we do not have clear evidence supporting either one of the two theories, both theories must be taken into account when considering the history of Arabic.

7. Middle Arabic

→ Middle Arabic is a cover term for the language variety employed in all texts with deviations from Classical Arabic grammar. This term does not, however, refer to a particular period in the history of the language, as is the case for Old, Middle, and Modern English. Texts which may be labeled as Middle Arabic exist from the first years of the Islamic era to the present day, and even some forms of oral production could be regarded as Middle Arabic. It is also important to note that Middle Arabic does not constitute a discrete variety of the language with its own set of rules, nor does the term refer to an intermediate stage between Classical and dialectal Arabic. While it is true that some Middle Arabic texts are close to the High acrolect and others to the Low variety, they are neither truly dialectal nor truly classical. Moreover, the amount of deviation varies from one text to another, and even in the same text from one sentence to another. It is therefore more accurate to see Middle Arabic as a continuum, rather than an established register of the language. For further information on these aspects see Blau, especially (1988) and (1999).

The aforementioned deviations from Classical Arabic grammar are mainly due to the fact that the authors had an imperfect knowledge of Classical Arabic. When trying to write correctly, the authors of Middle Arabic texts often used forms that were neither Standard nor colloquial. These forms are called pseudocorrections (→ hypercorrection). An example of this may be the use of a dual verb *našrabāni* 'we two drink', with a suffix used only with the 2nd and 3rd persons in Classical Arabic. The author, wishing to be seen to be a good Classical Arabic writer, incorrectly extends the use of the suffix to the 1st person, thus producing an aberrant form. Another kind of deviation is the presence in the texts of dialectal elements which originate from the author's mother tongue. These dialectal elements are often used unconsciously, although some are used deliberately to depict, for example, a lively dialogue where Classical Arabic would be inappropriate.

Middle Arabic texts are very important for the study of the history of Arabic, especially for the history of Arabic dialects, because they provide us with early attestations of dialectal features and allow for a diachronic insight into the origins of Arabic dialects and Neo-Arabic lingual forms.

There is a special branch of Middle Arabic which is usually called → Judaeo-Arabic, referring to the variety of Middle Arabic used by Jewish communities. It is characterized by two particular features: the use of Hebrew script and the presence of a large number of Hebrew loanwords and borrowings. These texts are usually less influenced by the Classical Arabic rules than those written by Muslim authors. See Blau (1981) for a full and accurate description of Judaeo-Arabic. Another branch is the so-called → Christian Middle Arabic, a group of texts which are often translations of religious texts from Syriac or Greek sources, often neglecting Classical Arabic norms and offering a unique flavor which sets it apart from general Middle Arabic. See Blau (1966–1967) for a good study of these texts.

8. THE EMERGENCE OF MODERN STANDARD ARABIC

From the 13th century onward, Arabic began to lose its position of prestige and its consideration as the language of culture and power. Consequently, its use in official and high domains gradually declined. With some outstanding exceptions, Arabic became restricted to everyday communication. This is the so-called dark period of the language. However, starting from the end of the 18th century, Arabic began to recover its status as a language of culture. The contacts between Arabs and other cultures, especially in Egypt and the Syro-Lebanese area, favored a renaissance (*nahḍa*) in which the Arabic language was an important symbol. The development of many cultural activities, such as translation, publication of Arabic journals and books, and the foundation of → language academies, led to the emergence of a 'new' Arabic based mainly on the inherited classical language. Such a new cultural tool had to be adapted to the reality of the new era, and to achieve this goal Arabic needed new words to designate modern concepts and realities. This was one of the main objectives of the Arabic → language academies, which tried to introduce the new terms into the language without altering the very nature of Arabic. The preferred method for new coinages was to create a derivate word from a preexisting Arabic root (e.g. *miṣʿād* 'elevator', formed with the root *ṣ-ʿ-d* 'to raise' + the pattern *mifʿāl* for names of instruments). Another method was to adapt the foreign word to Arabic morphological patterns (e.g. *tilfāz* 'television set', with the foreign consonants *t-l-f-z* + the pattern for instruments). The compounding of two or more words (→ compounds) and simple borrowing were two methods not favored by the academies. The emergence of the mass media and a new Arabic literature during the 20th century have brought Arabic to its current position as a powerful cultural language well adapted to the requirements of modernity.

The main differences between Modern Standard Arabic and Classical Arabic are of a lexical nature. A large number of foreign terms and expressions have entered the language, in many cases through a process of adaptation. Concerning syntax, some ancient structures were abandoned and a new style of connection between sentences was favored, heavily influenced by the language of the press, which incidentally suffered the impact of European linguistic influence. Various tendencies toward a more analytic language are registered in the modern language. Phonetics and morphology remain essentially the same, except where ancient patterns have been abandoned in favor of more productive modern forms.

9. The dialects of Arabic

Arabic dialects constitute an integral part of the language complex. They are the favorite everyday communication tool for all Arab speakers. Although they are usually studied as a separate phenomenon, they should perhaps be considered in a more comprehensive way, together with Modern Standard Arabic, a line taken by Holes (1995) in his complete survey.

Concerning the origins of Arabic dialects, three major theories have been proposed. The first one, represented by Ferguson (1959), states that all dialects derive from a koine, a hypertribal language in opposition to Classical Arabic, originating in the first military camps just before the great conquests of Islam. This would explain the structural differences between Classical Arabic and all Arabic dialects. Versteegh (1984) proposes that the process of acquisition of Arabic by a large number of foreign speakers (pidginization) is at the base of a structural reduction attested to in many dialects. A third theory sees in the dialects of Arabic the natural evolution of Old Arabic dialects by a process of drift. According to this view, because the Old Arabic dialects were diverse and tended toward the use of analytic structures, unlike Classical Arabic, it is hardly surprising that the modern Arabic dialects continue this trend.

10. The present-day situation

Arabic is today the native language of more than 200 million people, from the Atlantic coast to the plains of Mesopotamia. As the most outstanding identifier of the Arab people, the Arabic language is considered a symbol of Arab → nationalism. It is the official language of at least 23 countries, and since 1974 one of the official languages of the United Nations. In addition to this, Arabic is a minority or second language in many parts of the world, either as a native language or as a religious or cultural means of expression.

Bibliographical references

Blau, Joshua. 1965. *The emergence and linguistic background of Judaeo-Arabic.* London: Oxford University Press.

——. 1966–67. *A grammar of Christian Arabic, based mainly on South-Palestinian texts from the first millennium.* 3 vols. Louvain: Imprimerie Orientaliste.

——. 1977. "The beginnings of Arabic diglossia: A study of the origins of Neo-Arabic". *Afroasiatic Linguistics* 4.1–28.

——. 1988. *Studies in Middle Arabic and its Judaeo-Arabic variety.* Jerusalem: Magnes Press.

——. 1999. "The status and linguistic structure of Middle Arabic". *Jerusalem Studies in Arabic and Islam* 23.221–227.

Corriente, Federico. 1996. *Introducción a la gramática comparada del semítico meridional.* Madrid: Consejo Superior de Investigaciones Científicas.

——. 2003. "On the degree of kinship between Arabic and Northwest Semitic". *Proceedings of the Fifth Conference of AIDA (Association Internationale de Dialectologie Arabe), Cádiz 2002*, ed. Ignacio Ferrando and J.J. Sánchez Sandoval, 187–194. Cadiz: Servicio de Publicaciones de la Universidad de Cádiz.

Diem, Werner. 1979. "Studien zur Frage des Substrats im Arabischen". *Der Islam* 56.12–80.

——. 1980. "Die genealogische Stellung des Arabischen in den semitischen Sprachen: Ein ungelöstes Problem der Semitistik". *Studien aus Arabistik und Semitistik Anton Spitaler zum siebzigsten Geburstag von seinen Schülern überreicht*, ed. Werner Diem and Stefan Wild, 64–85. Wiesbaden: O. Harrassowitz.

Ferguson, Charles A. 1959. "The Arabic koine". *Language* 35.616–630.

Ferrando, Ignacio. 2001. *Introducción a la historia de la lengua árabe.* Saragossa: Universidad de Zaragoza.

Gruendler, Beatrice. 1993. *The development of the Arabic scripts: From the Nabatean era to the first*

Islamic century according to dated texts. Atlanta, Ga.: Scholars Press.
Hetzron, Robert. 1974. "La division des langues sémitiques". *Actes du premier Congrès International de Linguistique Sémitique et Chamito-Sémitique, Paris 1969*, ed. André Caquot and David Cohen, 181–194. The Hague and Paris: Mouton.
——. 1997. *The Semitic languages*. London and New York: Routledge.
Holes, Clive. 1995. *Modern Arabic: Structures, functions and varieties*. London and New York: Longman.
Ḥusayn, Ṭāhā. 1926. *Fī š-šiʿr al-jāhilī*. Cairo: Maktaba Dār al-Kutub al-Miṣriyya.
Mannāʿ, Hāšim Ṣāliḥ. 1993. *An-Naṯr fī l-ʿaṣr al-jāhilī*. Beirut: Dār al-Fikr al-ʿArabī.
Müller, Walter M. 1992. "Das Altarabische der Inschriften aus vorislamischer Zeit". *Grundriß der arabischen Philologie*, ed. Wolfdietrich Fischer, I, 30–36. Wiesbaden: L. Reichert.
Nöldeke, Theodor. 1938. *Geschichte des Qorans*. III. *Die Geschichte des Korantextes*. Leipzig. (Repr., Hildesheim: G. Olms, 1961.)
Rabin, Chaim. 1951. *Ancient West Arabian*. London: Taylor's Foreign Press.
Robin, Christian. 1993. *L'Arabie antique de Karibʾīl à Mahomet: Nouvelles données sur l'histoire des Arabs grâce aux inscriptions*. Aix-en-Provence: Editions Edisud.
Rodgers, J. 1991. "The subgrouping of the Semitic languages". *Semitic studies in honor of Wolf Leslau*, ed. Alan S. Kaye, II, 1323–1336. Wiesbaden: O. Harrassowitz.
Sāmarrāʾī, ʾIbrāhīm, as-. 1994. *Fī l-lahajāt al-ʿarabiyya al-qadīma*. Beirut: Dār al-Ḥadāṯa li ṭ-Ṭibāʿa wa-n-Našr wa-t-Tawzīʿ.
Versteegh, Kees. 1994. *Pidginization and creolization: The case of Arabic*. Amsterdam and Philadelphia: J. Benjamins.
——. 1997. *The Arabic language*. Edinburgh: Edinburgh University Press.
Voigt, Rainer M. 1987. "The classification of Central Semitic". *Journal of Semitic Studies* 32.1–21.

Ignacio Ferrando (University of Cádiz)

Homonym → Muštarak

Horn of Africa

While Arabic is the native language of only a small minority of the citizens of Eritrea and Djibouti (→ Djibouti/Eritrea), it is nevertheless used as a trade language by almost everyone. The multilingualism in these two countries (three national languages in Djibouti, ten in Eritrea) has favored the development of Arabic as a lingua franca. This is the situation in Djibouti and in the lowland coastal region of Eritrea, including the islands such as the Dahlak archipelago, together with a part of the western highlands. The current situation has its historical, geographical, and economic roots in the strength of the centuries-old links between the African and Arabian coasts of the Red Sea and the subsequent presence of Arabs on the islands and in the ports on the African side. The preference for Arabic is reinforced by its religious dimension, as the coastal populations are all Muslim. In Eritrea, the war of independence has also had a decisive impact. For 30 years, this war was responsible for the emigration of a considerable number of inhabitants, mainly from non-Arabic-speaking groups, toward the surrounding countries, Saudi Arabia, Yemen, and particularly the Sudan. These emigrés are now returning home having learned Arabic abroad. In addition, the liberation army included soldiers of diverse ethnic origin, and there has been internal displacement of populations toward regions where a different language is spoken. The result has been the emergence of an Arabic lingua franca spoken by people whose mother tongues are Afro-Semitic and Cushitic. The lingua franca displays a north-to-south divergence deriving from both linguistic factors such as differing roles of the mother tongue, substrates, and influences of the Arabic dialects with which the speakers have been in contact, and extralinguistic factors such as the speakers' class, personal experience, and the circumstances in which they learned Arabic and now use it.

The following discussion is based on data collected on the island of Dahlak Kebir and on the coast between Massawa and Djibouti from speakers whose native tongue is either an Afro-Semitic language such as Dahalik (at Dahlak Kebir) or a Cushitic language such as Saho (between Massawa and Irafayle in Eritrea) or Afar (on the coast between the Bori peninsula in Eritrea and the city of Djibouti). More is known of the Arabic lingua franca spoken in Eritrea than of any of the other varieties (Simeone-Senelle 2000a, 2000b).

1. Phonological features

The characteristic features of the Arabic lingua franca and the influence of the speakers' mother tongues are most immediately perceptible in the phonetics and the phonological system of the language. The consonant system is composed

Table 1. Primary correspondences Standard Arabic/lingua franca

	emphatic			interdentals		apico-alveolars	palato-alveolars		velars		uvulars
Standard Arabic	ṭ	ṣ	ḍ, ḏ̣	ṯ	ḏ	z	š	ž	ġ	x	q
lingua franca	(ṭ) t	(ṣ), s	(ḍ) d, z	(ṯ) t, s	(ḏ) d, z	z, d	(š) s, ss, s̃	j, y, g	ġ, q, g	k	(q) g, ġ, ʁ, χ, k, ʔ

Table 2. Primary assimilations

regressive assimilation	*n-d* > *dd*: *ʿáddahā*	*t-d* > *dd*: *d*^{ə}*dáwwer*	*d-t* > *tt*: *āgétto*
	'with her'	'you[sg.] look for'	'I found him'
	t-z > *zz*: *yazzáwwej*	*ž-d* > *dd*: *taddil*	*l-n* > *nn*: *en-nūn*
	'he gets married'	'you [sg.] throw'	'the color'
progressive assimilation	*s-t* > *ss*: *yissénni*	*b-l* > *bb*: *kébb-i*	
	'he stays'	'before me'	

of 25 elements, including the semiconsonants, or 24, when the emphatic dental stop and interdental fricative merge, as they did in most Arabic dialects. The native languages of lingua franca speakers include Afro-Semitic Dahalik and Tigre, which have an equivalent 24-consonant system including two palatoalveolars (*š*, IPA ʃ, and *ž*, IPA ʒ), a voiced apico-alveolar (*z*), and a voiced velar (*ġ*). While the Tigre system has four ejectives, *ṭ, ṣ, ḳ, č̣* (Raz 1983:4), Dahalik has only ejective *č̣* (in borrowings?), the other three being velarized. Cushitic Afar and Saho, on the other hand, have smaller systems: 15 in Afar and in Saho 16, although some scholars have noted up to 22 (with ejectives) according to dialect (Morin 1994:257–258, 1995:68–69; Welmers 1952:145). Afar and southern Saho have no emphatic consonants, no alveolars (apico- and palato-alveolars in Afar, while Saho has only /z/), and no dental fricatives. Only Saho has a velar, voiceless fricative, but both languages have a retroflex stop (*ḍ*). With the exception of the retroflex, the features of these languages can be found in the speakers' lingua franca systems.

A common feature is that, whatever the speakers' mother tongue, the velarized emphatic consonants of Arabic are either unstable or lost. While Saho speakers at Irafayle and Dahalik speakers tend to retain emphatics, realizations elsewhere vary over retention and change. Thus a single speaker may use both *manṭaqa* or *mantaqa* 'region', and both *ʿasab* or *ʿaṣab* 'Asab' in the same utterance. Where the emphatic consonant is retained, it always has a velarized realization as in Arabic, rather than a glottalized one as in the Afro-Semitic contact languages (Tigre and Tigrinya). More often the realization is indistinguishable from the corresponding non-emphatic stop or fricative /ṭ/ > [ŧ], [t]; /ṣ/ > [ꞩ], [s] (*mabsūtīn* '[they are] well'; *sádəf* 'shell'). Generally speaking, the archiphoneme Ḍ is realized as a dental stop or more rarely as a voiced apico-alveolar /ḍ/ and /ḏ̣/ > [d] and [z]: *ard* 'earth', *baʿd* 'some of them', *del* ~ *zel* 'shadow'.

None of the speakers' mother tongues have the dental fricatives *ṯ* and *ḏ*. A few speakers use them in the lingua franca, but they usually merge with the apico-alveolars in line with the Afro-Semitic systems: /ṯ/ > [s] and /ḏ > [z] (*ízā* 'if', *masalan* 'for example', *hēze* 'this, that'). Many speakers produce a dental stop instead of a dental fricative, as in the Arabic dialects which have no dental fricatives (*tiyāb* 'garment', *hēde* 'this, that'). Nevertheless, a single speaker may use all three pronunciations (in Dahlak Kebir, *keṯīr, kesīr, ketīr* 'many').

Native Afar speakers often merge the voiced alveolar fricative, not found in Afar, with the voiced dental stop /z/ > [d] (*derīʿa* 'agriculture').

Saho and Dahalik speakers, who have a *z* in their own language, produce one in their variety of the lingua franca.

Standard Afar has no voiceless palato-alveolar fricative *š, ž*; only a few dialects have incorporated *š* in a small number of lexical items. This sound exists in some Saho dialects (Morin 1995:68), although its status is uncertain. Native Saho speakers retain the unvoiced palato-alveolar fricative of Arabic in the lingua franca. Afar speakers either palatalize the unvoiced consonant as *s̃* (IPA ç) or pronounce it as an apico-alveolar fricative [s]: *mus* 'isn't?'. A single speaker may use all three realizations in the same utterance (Dahlak Kebir: *šajara* ~ *s̃ajara* ~ *sajara* 'tree'). Others pronounce the unvoiced consonant as a geminate apico-alveolar (in Galaʿlo: *tamúšu* ~ *tamússu* 'you [pl.] walk').

/ž/, which is not found in either Afar or southern Saho, allows of varying degrees of palatalization: /ž/ > [ʤ], [dy], [j] (*jedīra* 'island', *yed* 'grandfather', *yayi* 'he comes'). This phenomenon has also been observed in some Tihama Yemeni dialects; the realization of */ž/ as [j] is a feature of the Hadramawt dialects. Another pronunciation, infrequent among Afar speakers but more widely used by Saho speakers, is the velar stop [g] (*nargaʿ* 'we come back'). This can be found in many Arabic dialects, particularly in southwest Yemen where it is common. Three pronunciations can be observed in the production of a single speaker (*jibūti* ~ *dyibūti* ~ *gibūti* 'Djibouti').

The voiceless uvular plosive /q/ has a wide variety of realizations. Unlike the situation which prevails in the Tihama region (Behnstedt 1985:41), it is rare to find a speaker of any of the native languages who will pronounce this consonant as a voiceless uvular plosive outside such frequently spoken items as *qāl* 'he said'. It usually appears as the voiced velar (*g*), as in some forms of 'Bedouin' Arabic (*wálad ḥagg-i* <boy/of-me> 'my son', *gábel* 'before', *gárya* 'village'), or as the unvoiced velar (*k*) (*taktaʿ* 'you [sg.] cut'). Strong palatalization of the velar may move it to a palatal fricative (*y*) (/qaʿāda/ > *gaʿāda*, but also *jaʿāda* ~ *dyaʿāda* ~ *yaʿāda* 'traditional bed/bank', in Eritrea). This phenomenon can be found in some Tihama dialects (Simeone-Senelle a.o. 1994:219). The most widespread realizations are fricative, velar, or uvular, either voiced or unvoiced according to the other consonants and/or vowels in the context, although these may alternate with plosion. This is true even of Saho and Afar speakers, whose own system has no velar or uvular fricatives (*qanūn* ~ *ʁanūn* 'law', *naġder* 'I am/we are able', *wáxta-nā* 'our times', *qalīl* ~ *ʁalīl* ~ *ġalīl* 'few'). The same alternating variants (*q, ġ, ʁ*) can be found in some Yemeni dialects (Vanhove 1995a:146). /q/ is rarely pronounced as a glottal stop [ʔ] unless under the influence of an Arabic dialect encountered in emigration (Cairo or Middle East) or some western Yemeni dialects (Behnstedt 1985:44). There are abundant examples showing that all these realizations can be used by a single speaker, often in the same utterance (*et-taʾālid w-al-kawānin ḥagg-al-ʿáfar* 'the customs and laws of Afar'; *lá-ho ʾúmnā ka-qárya* <for-him/we set up/like-village> 'for that reason, we built a kind of village').

Concomitantly, the etymological voiced velar fricative *ġ* is realized as a uvular stop, whether unvoiced as in the southern Yemeni dialects (Vanhove 1995a:147), where /q/ > *ġ* and /ġ/ > *q* (the most frequent pronunciation for Saho speakers: *qāli* 'expensive', *ṣaqīra* 'small [fem. sg.]', *qálat* 'he is wrong'), or as voiced *g* (*ṣagīra*, *áṣgar min-el-ġazāl* '[it is] small [fem. sg.], smaller than the gazelle'). A single speaker may use both pronunciations in the same utterance. The unvoiced velar fricative in turn is most often realized as a velar stop (*lákam* 'shark'; *dākil* 'inside'; *karūf* 'lamb'; *kállast* is the general way of saying 'I have finished').

The lingua franca has another feature common to many Arabic dialects: the spirantization of bilabials (*ßə-báḥar* rather than *bᵊ-báḥar* 'at sea'; *φē* 'there is'). As in the Tihama dialects, particularly among women of the Khokha region, dental and bilabial stops are not released in word-final position and almost all consonants are lost before pause (*báʿdeʾ* 'after', *basīʾ* 'simple', *nḥásaʾ* 'I/we count').

Reduplication and assimilation are widespread in consonant groups. As in many other Arabic dialects, the suffixation of a personal pronoun to a preposition with a final consonant results in gemination of the latter. All speakers say *ménni* 'from me', *mínnahā* 'from her', *mínno* (or *múnno* in Dahlak Kebir) 'from him'. Assimilation is systematic whenever one of two consonants is dental or apicoalveolar (*l* or *n*). The direction is usually regressive except when a dental or the bilabial *b* follows an apico-alveolar (*s* or *l*).

There are six short vowel realizations: *i*, *e*, *a*, *o*, *u*, and unstressed *ə*, except in certain verb paradigms. Some Afar speakers, particularly at Dahlak and Assab, tend to nasalize prepausal vowels. The nasalization is sometimes velarized as in the western Yemeni dialects (*lámmã*^{ŋ} ~ *námmã*^{ŋ} 'when'; Simeone-Senelle a.o. 1994:220; Behnstedt 1985:57). Most speakers contrast length for all stressed vowels.

For most speakers, diphthongs are systematically reduced: /aw/ > [ō]/[ū], and /ay/ > [iː]/[eː], e.g. *mūj* 'waves', *dē* (cf. *zey*) 'like'. This phenomenon is attested in many Yemeni dialects including most Tihama dialects. A single speaker may nevertheless pronounce either *ʿáydān* or *ʿī́dān* 'pieces of wood'.

Syllable structure is generally either CVC or CVV. Most etymological consonant clusters are split, and there are no heavy CVCC syllables in final position (*láḥam* 'meat', *sámis* 'sun', *al-íbin* 'the son').

2. MORPHOLOGICAL FEATURES

2.1 *Pronouns*

Table 3. Personal pronouns

singular	independent	dependent
3	*húwa/hū*	*-o/-oh/-ho*
3 f.	*híya*	*-hā*
2	*ínta*	*-k*
1	*ána*	*-i* (*-ni* after verb)
plural		
3	*húmma*	*-hum/-hun*
2	*éntun/íntu*	*-kum*
1	*náḥna*	*-nā*

As in Afar and Saho, there is no gender distinction in the 2nd person.

The reflexive is either *b-nafs* + pronominal suffix or *b-rūs* + pronominal suffix (*bi-náfsa-nā* '(by) ourselves', *b-rūs-hum* '(by) themselves'). Note that *nafs* and *raʾas* are also used to form the reflexive pronouns in Tigre (Raz 1983:40).

The relative pronoun is *elli*, almost always invariable. The animate associative pronoun is *men* or *elli*, but *mā* is used for an inanimate referent (*ʿalāqa élli térbūt* 'relation which binds', *élli ʿéndihum sámbuk* 'those who have a boat', *men dáraʿo* 'the one who has cultivated it', *mā gúlta* 'what I/you [sg.] have said').

The demonstrative is *hēda/hēza/*(*hēḏa*); infrequently one finds *hēde/hēze/*(*hēḏe*) for the feminine. It is extremely rare to find a distinction between proximate and distant deictics. So far, only one Afar speaker has been found in Eritrea using *dē/dī/da* as a base for the deictic, which is also found in the Khokha region of Tihama (Simeone-Senelle a.o. 1994:223) and occasionally forms a plural (*en-nās dīla (~ zēla)…fī l-amākin da* 'these people…in these places').

2.2 *Nouns*

Some nouns have plural forms of a type uncommon or even unknown in Modern Standard Arabic (*mutaṭállabāt* 'requests', *qaryāt* 'villages', *ayāmāt* 'days', *nuxūl* 'palm trees').

2.3 *Verbs*

The imperfect conjugation is characterized by the use of a single form for both singular and plural 1st person subjects. This feature, already recorded by Cohen (1931–1934:31), is confined to this form of Arabic. It is rare in the north and only becomes general in the southern part of the region (*naġder* 'I am/we are able').

In the perfect, there is a single form for masculine and feminine subjects in the 1st and 2nd persons. It has final *-a* (*samaʿta, šúfta, fahámta* 'I/you heard, I/you saw, I/you understood'), as in some Tihama dialects (Behnstedt 1985:117) and in Khartoum (Persson and Persson n.d. [1979]:29).

The apophonous passive is very rare. The passive is a derived form obtained by prefixing or infixing *t*, generally in a passive or middle sense (*fēn atwáladt?* 'where were you born?', *yetwāgid* 'there is/it is found', *yébtaʿ* 'it is sold', *néntaqəl* 'we move').

3. VERBAL MORPHOSYNTAX

A set of particles, preverbs, and auxiliaries, together with *kān*, is used to express tense and mode in verb phrases. The attested constructions are mostly comparable to those found in Yemeni dialects (Simeone-Senelle 1996).

The prefix *ba-* in the imperfect expresses concomitance (simultaneity of the verb process with the moment of utterance or some other reference point fixed in the utterance). This particle, which is also attested in eastern dialects, is used in Yemeni dialects to express concomitance and impending future (Behnstedt 1985:132; Jastrow 1980:122; Simeone-Senelle a.o. 1994:204): *bá-ʿtì lu-kum* 'I am giving/transmitting [it] to you', *íḏe kúnna ba-nakátter-o* 'if we were currently increasing it'.

Concomitance can also be expressed by the participle (generally variable in gender and number) of the verb *gaʿad* 'to stay, remain', followed or preceded by a verb in the imperfect: *el-ʿáməl élli íntu gā́ʿadīn tšū́fu-hā* 'the labor which you are watching', *tagúl-l-ak gā́ʿad* 'she is telling you'. The same speakers use *gaʿad* in the perfect as an auxiliary expressing past inchoative. The main verb takes the imperfect: *gaʿádna nəbni fi-hā* <we stayed/we build/in-her> 'we started to build there'. Elsewhere (at Galaʿlo), another verb, *bagā* 'to stay', is used to express concomitance in the imperfect: *nábgā nagūl bi-l-ʿáfar* 'I am/we are now saying it in Afar' = 'I am/we are now speaking Afar'.

ḥa- followed by the imperfect expresses the future in southern Eritrea in the Afambo region and at Rahayta: *yagūlu ḥa-yákūn əmṭār katīra* 'they say there will be a lot of rain'. This particle is attested in other Arabic dialects, in Yemen (Vanhove 1995b:266), in Egypt, and in Sudan. It is also found in the dialects of Oman.

The use of *gāma*, etymologically 'to stand up', in the imperative as an interjection or exhortation is another feature shared by the Tihama Yemeni dialects (Simeone-Senelle 1996: 232–233) and the lingua franca of the Horn of Africa: *gū́mu!* 'go ahead!' (inviting the addressees to speak into the microphone without moving from where they are).

The use of *bagā/baga* 'to stay' in the 3rd person masculine singular of the imperfect as a connective in a list seems to be limited to native Dahalik speakers on the island of Dahlak Kebir; it has not been observed in Yemen or in other varieties of the lingua franca, e.g. in a genealogy, *bága héḏi yaʿīdi jāb aḥmed saġīr* 'and then, that Ya'idi gave [birth to] Ahmed Junior', or in a description,... *bága-l-láḥam yiṭálləʿu-ho* '...and then the flesh, they removed it'.

4. SYNTACTIC FEATURES

4.1 *Noun syntax*

The constituent order is Modified + Modifier. The nominal or pronominal modifier may be constructed synthetically or analytically. The direct construction is used for kinship terms, body parts, and origin (*ibn ʿámm-o* 'the son of his uncle', *ism-o(h)* 'his name', *ṣulṭān ʿáṣab* 'the sultan of Asab'). This rule is not applied strictly, however, since the same speakers may say *qalabíyyet-nā* 'the majority of us'. The analytic construction is more common, even for the expression of possession: the personal pronoun referring to the possessor is usually suffixed to the preposition (*ḥagg/ḥaqq*): *wálad ḥagg-i* 'my son', *ʿiyēl ḥágg-ak* 'your [sg.] children', *er-rīša ḥagg-o* 'its [masc.] wing', as in the Tihama dialects, but unlike most other Yemeni dialects (Piamenta 1990–1991). Use of *men* is rare and requires reversal of constituent order (*mén-ni awlād* 'my children'). In Afambo, a speaker who had been in contact with the Cairo dialect used the connective *bitāʿat* (*el-mawāši bitāʿat-nā* 'our cattle').

The definite article is *el-/al-*. There being no article in Afar or Saho, the rules governing its use in the lingua franca are often different from those which hold in the dialects of native Arabic speakers (*ed-dīn islāmi* 'the Muslim religion', *et-tejémmoʿ axwān* 'the gathering of the brothers', *min el-jild el-ġánem* 'out of sheepskin').

4.2 *Agreement*

As in most trade languages, agreement in person, number, and gender is sporadic, even between subject and verb or auxiliary verb. The relative pronoun is usually invariant, and very few speakers use varying forms of the demonstrative (see above). Apart from this, only a few examples can be given, without any claim to systematicity: *el-lúqat* [fem.] *el-ʿáfari* [masc.] 'the Afar language'; *məntága ḥágganā, zerāʿa mā-fì-ho* [instead of *fi-hā*] 'in our region, there is no agriculture'; *ánā kúnta ʿajibt* 'I was surprised', but also: *námšī kān yémen* 'I was going to Yemen'; *atfāham*, *atfággu* 'they understood each other/one another, they agreed'; *yáqdar yaqburu* 'they can bury'. Any type of agreement with a collective noun is possible: as masculine or feminine singular, or as plural (*nās kúll-o* 'all the people, everybody', *en-nās élli mawjūda gábel kída, ḥáddadu el-ʁanūn* 'the people who were there beforehand, they laid down the law'. As in Afar, an ethnonym takes feminine singular agreement (*ʿáfar tugūl* 'the Afar say', *el-ʿáfar ʿinda-hā qanūn sābit* 'the Afar have firmly established laws'). After a numeral, a noun can be either singular or plural (in the native languages, the counted noun is always singular).

4.3 *Sentence syntax*

The predominant constituent order in the sentence is Subject – Object/Adverbial – Verb, which seems attributable to the speakers' native languages. The most striking feature is the placement of the auxiliary, usually after the modified verb and even in proposition-final position (*ḥarb isáwwū wu-igéṭṭəʿū kā́nū en-nās* <war/they do/and-they cut into pieces/they were/ART-people> 'they were waging war, and they were slaughtering people'; *kulli-yōm yáʿmil-l-oh kída kān* 'he behaved thus with him every day'), as in the Cushitic and Afro-Semitic languages of the region.

Nominal predication is used to express existence. For all speakers, the locative preposition *fī* 'in' is the existential copula 'there is'. The 3rd person masculine singular independent personal pronoun (at times supported by *fī*) can have the same role. The verbal copula *kān*, with or without *fī*, allows existence to be situated in the past (*fī ḥarīm, fī kamān ríjel* 'there are women, and also men'; *árbaʿ w-ášrīn kōkab húwa* 'there are twenty-four stars'; *húwa fī ʿalāqa* [fem.] 'there is a relationship'; *kān ḥarb hīna* 'there was war here').

Existential possessive propositions are construed almost exclusively with *ʿind-/ʿend-* 'at (the home of), for' + pronominal suffix referring to the subject/possessor (*ʿinda-hā ʿiyēl* 'she has children', *élli ʿénd-o ḥúkum* 'the person who is in power').

4.4 *Negation*

The negative particle is *mā/má*. It is always stressed as in Afar, where the same particle is used. When the scope of the negation is the entire predicate, the particle precedes the verb or the copula (*mā-yetḥasel* 'it doesn't happen', *má-kān dirāsa* 'there was no teaching'). Negation of existence is placed at the end of the proposition (*yemeniyūn mā-fi* 'there are no Yemenis/Yemenis, there are none'). On rare occasions, *mā-fi-š/mā-fi-ši*^{n} expresses absolute negation ('there is nothing'). In a noun phrase, *mūš/muš/mus* before a noun or an adjective functions as negation (*mūš húmma* 'it's not them', *mūš kwéys* 'it's not well'). One of the characteristic features of both the Afro-Semitic languages and the Arabic dialects of the region is the placement of this particle in proposition-final position as a rhetorical interrogative negation punctuating an assertion. It is so used by all speakers without regard to native language. All languages in the region tend to use it as punctuation after each phrase in the utterance.

5. LEXICAL FEATURES

The lexicon reveals the various Arabic dialects with which the speakers have come into contact. Many terms are common to the lexicon of the Yemeni dialects, particularly those of the Tihama region. Thus, the most widespread verb, *išti* 'he wants', is used in almost all of western Yemen (Behnstedt 1985:202). At Dahlak Kebir, some people also use *bġa*, as in the Hadramawt region (*eš tabġa* along with *eš tišti* 'what do you [sg.] want?'). At the same time, a speaker at Afambo in Eritrea who spent some years in Egypt makes nearly systematic use of the participle *ʿāwez: fī wāḥad ʿāwez yazzáwwej* 'there is one who wants to get married'. As in the Khokha region, in Yemen, *ata* 'to come', is used along with *agā/ajā*. In Eritrea, particularly in the Tiʾo region, *estanna* does not mean 'to wait' but rather 'to live, stay'; when followed by a participle, it denotes a durative aspect 'to do over time': *arum, húwa yisténni fi-l-baḥar* 'the dugong [sea cow] lives in the sea'; *yestánnu máwgudin fi l-mántaqa* 'they still live in the region'.

Speakers vary greatly in their choice of nouns and prepositions; even individuals may vary from one occasion to another. According to the region, 'water' is either *mōya* or *māʾ*; in Dahlak Kebir, *mōya* alternates with *mō*. *katkūt* (pl. *katākīt*) denotes 'chick' as in some quarters of Aden (the term is not recorded by Piamenta 1990–1991). *bēn* is used in Dahlak, as in the Khokha region of Yemen and in Jiblih (Jastrow 1980:127), with the sense of 'in, inside' along with *fī*: *bēn el-fərūm* 'in the oven', *fī l-baḥar* 'in the sea'. Under the influence of an Egyptian dialect, one speaker uses *júwwat el-baḥ*^{a}*r* in the same context (Egyptian *guwwa*; Hinds and Badawi 1986, s.v.). *daḥḥin* alternates with *delḥin*, or *déḥin* in some cases, for 'now'.

6. Discourse-related and sociolinguistic features

Discourse-related features found in the lingua franca are similar to those found in many trade languages. One characteristic feature is the tendency on the part of all speakers to repeat themselves as they proceed. Such redundancy makes the content clearer and does away with ambiguity. Lexical, syntactic, and phonetic procedures are all used. Word order may be rearranged (*élli yisáwwu kānu fi-hā ʿásel áwwel, héḏe áwwel élli kānu yisáwwu-hā* 'what they put in there was first of all honey, that was what they put in there first'); prepositions, whatever their meaning, may be reinforced by *fī* 'in' (*fī dākil* 'inside', *fī men* 'outside'); an adverb may be supported by another one with a similar meaning (*kamān bárdo* 'likewise, also'); phonetic variants of the same term may be repeated within the same sentence (*sámis*, *sámi* 'sun'; *túmun qirāṭ, núṣṣ-u-ʁirāṭ, gerāṭ* 'one-eighth of a carat, half a carat, a carat'); and different words for the same referent may be used within the same utterance (*bābur, safīna, markab*, all denoting a kind of boat).

Hypercorrect forms involve mainly phonetic realizations (*xatīr* for /kaṯīr/).

Instances of semantic interference are few in number; a typical example is provided by a speaker who uses *ʿīdan* [pl.] for 'medicines'. In Afar, his native tongue, *ḥaḍa* (pl. *ḥooḍ*) means both 'tree, wood' and 'medicine'.

Codeswitching is frequent with speakers who begin a sentence in Arabic and then switch to their native tongue (Arabic//Afar: *el-gélid ḥágg-o//goobu-h haan* 'its skin,//they make a shield out of it'; *taġrīban kébbi kam ʿām//yekkee* 'how many years before me [i.e. my birth]//did it happen?'.

Evidently, Arabic as a lingua franca is not a homogeneous variety. While some features may fluctuate less than others, it is an unstable variety, varying from one group of speakers, or even from one speaker, to another. This is the inevitable consequence of the fact that its development is intimately connected with both the sociolinguistic history of the users and the structure of their native tongues (Abu Manga and Miller 1995–1996:185).

In the coastal region of Eritrea and in Djibouti, fishermen and seafarers in general are in regular contact with Yemen, the Tihama region, Aden, and the Hogariya region, as well as with some Hadramawt dialects on the coast of the Indian Ocean. The Arabic lingua franca in Eritrea is also dependent on the dialect learned and used during the independence war (1961–1991), on the front or in the refugee camps, or in exile in a neighboring or more distant Arabic-speaking country. Overall, the Arabic lingua franca strikes one as a speech form in constant change. Nevertheless, it has its own characteristic features, which identify it as originating in this region of the Horn of Africa.

Bibliographical references

Abu Manga, Al-Amin and Catherine Miller. 1995–1996. "Contact between Arabic and Hausa in Sudan: A semantic perspective". *Matériaux Arabes et Sudarabiques* n.s. 7.183–204.

Behnstedt, Peter. 1985. *Die nordjemenitischen Dialekte. I. Atlas*. Wiesbaden: L. Reichert.

Hinds, Martin and El-Said Badawi. 1986. *A dictionary of Egyptian Arabic: Arabic – English*. Beirut: Librairie du Liban.

Cohen, Marcel. 1931–1934. "Sur l'arabe parlé à la côte des Somali et au Harar". *Comptes Rendus du Groupe Linguistique d'Etudes Chamito-Sémitiques* 1.30–31.

Jastrow, Otto. 1980. "Text aus Yarīm, Nordjemen", "Texte aus Giblah (Jiblih), Nordjemen". *Handbuch der arabischen Dialekte*, ed. Wolfdietrich Fischer and Otto Jastrow, 122–129. Wiesbaden: O. Harrassowitz.

Morin, Didier. 1994. "Dialectologie de l'afar-saho". *Semitic and Cushitic studies*, ed. Gideon Goldenberg and Shlomo Raz, 252–266. Wiesbaden: O. Harrassowitz.

——. 1995. *Des paroles douces comme la soie: Introduction aux contes dans l'aire couchitique (bedja, afar, saho, somali)*. Louvain: Peeters.

Persson, Andrew and Janet Persson [with Ahmad Hussein]. N.d. [1979]. *Sudanese Colloquial Arabic for beginners*. Summer Institute of Linguistics.

Piamenta, Moshe. 1990–1991. *Dictionary of post-Classical Yemeni Arabic*. Leiden: E.J. Brill.

Raz, Shlomo. 1983. *Tigre grammar and texts*. Malibu: Undena.

Simeone-Senelle, Marie-Claude. 1996. "Le fonctionnement de certains verbes de mouvement dans des dialectes de la Tihama du Yémen". *Proceedings of the 2nd International Conference of l'Association Internationale pour la Dialectologie Arabe held at Trinity Hall in the University of Cambridge, 10–14 September 1995*, ed. Joe Cremona, Clive Holes, and Geoffrey Khan, 227–235. Cambridge: Cambridge University Publications Center.

——. 2000a. "La situation linguistique dans le sud de l'Erythrée". *Proceedings of the 2nd World Congress of African Linguistics at the University of Leipzig, 27th July–3rd August 1997*, ed. H. Ekkehard Wolff and O.D. Gensler, 261–276. Cologne: R. Köppe.

——. 2000b. "L'arabe véhiculaire parlé en Erythrée sur la côte de la mer Rouge, de Massawa à Rahayta". *Oriente Moderno*, n.s. 19 (80):1.153–181.

——, Martine Vanhove, and Antoine Lonnet. 1994. "Les dialectes arabes de la Tihaama: Diversité et caractéristiques". *Actes des Premières Journées Internationales de Dialectologie Arabe de Paris (27–30 janvier 1993)*, ed. Dominique Caubet and Martine Vanhove, 217–231. Paris: INALCO.

Vanhove, Martine. 1995a. "Notes on the Arabic dialectal Area of Yâfiʿ (Yemen)". *Proceedings of the Seminar for Arabian Studies* 25.141–152.

——. 1995b. "À propos du verbe dans les dialectes arabes de Yāfiʾ (Yémen)". *Dialectologia Arabica: A collection of articles in honour of the sixtieth birthday of Professor Heikki Palva*, 257–269. Helsinki: Finnish Oriental Society.

Welmers, W.E. 1952. "Notes on the structure of Saho". *Word* 8:2–3.145–162, 236–281.

Marie-Claude Simeone-Senelle
(LLACAN-CNRS, INALCO, University of Paris-7)

Hypercorrection

Hypercorrection (also called overcorrectness) is one kind of 'linguistic correction', best termed 'pseudocorrection' (Blau 1970). Pseudocorrections result from speakers' and writers' desire to speak and to write a more prestigious variety and to avoid stigmatized forms. For example, in England, tension between social dialects has persistently caused speakers and writers to employ various hypercorrections. One important determinant of social status has been the pronunciation of the glottal fricative [h] rather than its omission. A phrase such as [amend-hɛgz] 'am and heggs', therefore, is a hypercorrection because the speaker wants to use the prestigious glottal fricative [h] but fails to use it in the appropriate place. Another example, from the dialect of children in Reading, England (Cheshire 1984:551–552), appears where the marked verbal *-s* suffix form is generalized for all forms, thus, the dialect calls for *I knows, he knows*, and *they knows*. In this example of hypercorrection, the children aspired to use the prestigious form of the verbal *-s* suffix in all forms. However, with growing criticism of this phenomenon by their educators, the children omitted the verbal *-s* suffix from all forms, as it was widely criticized; their efforts resulted in another form of hypercorrection, as they began to omit it even when it was actually needed in the standard variety, resulting, for example, in *it taste all rich and creamy*, this time in writing. In other dialects, what is in Reading, England, a hypercorrection is simply a dialectal feature, as in *he go* in deep African American Vernacular English. Another example comes from written British Black English (Fasold 1996:280), where the plural marking *-s* is not present in the dialect. Thus, *golds* results, not from a desire to use the dialectal zero plural suffix, but rather to employ the prestigious plural marking *-s* in all positions, even when it is unnecessary in the standard variety as well.

The phenomenon of hypercorrection is not at all unusual in most languages. Hypercorrection is especially important in Arabic because of its continuuglossic situation (Hary 2003), and because the phenomenon is a mechanism for change in all languages (Labov 1972:178–180).

Whenever a variety of a language with social, religious, economic, or other prestige comes into contact with a variety without such prestige, speakers and writers of the latter will, at times, try to use forms of the former even if the forms are unnecessary in that linguistic environment. They will change or 'correct' forms of the prestigeless variety and utilize the forms of the language with the prestige. However, sometimes these speakers do not have sufficient knowledge of the prestigious variety. Thus, they change or 'correct' forms that do not need to be changed even according to the prestigious variety, and may arrive at forms that are 'too corrected' (hence, hypercorrections) or 'halfway corrected' (or 'not corrected enough'; hence, hypocorrections). Furthermore, speakers and writers may sometimes mix forms. Such pseudocorrections stem from the desire of the speakers to 'decorate' themselves with forms of the prestigious variety and to avoid the dominant usage of the prestigeless variety, often out of over-self-denial, in order to gain prestige in their speech or writings. At times, such corrections can become standardized in the prestigeless variety, thus contributing to language change (see below the example of *lam* followed by the perfect and the example of *hamza*).

Arabic serves as a perfect example of frequent pseudocorrections because historically Classical Arabic has enjoyed tremendous prestige compared to colloquial Arabic. For example, the pronunciation of the uvular stop [q] in Classical Arabic becomes a glottal stop [ʔ] in many urban Arabic dialects. Therefore, one can hear the pronunciation of a woman's name [ʔumm fuʔɑːd] 'mother of heart' in a hypercorrected form as [ʔumm fuqɑːd] 'mother of bereavement' (Garbell, as quoted in Kutcher 1959:23, n. 4). The speakers may have heard the word [fuʔɑːd] 'heart' but assumed that every glottal stop [ʔ] in the dialects stems from a literary uvular stop [q], thus replacing the glottal stop [ʔ] with [q], where it is not needed even by the literary variety, which, in effect, changes the name and its meaning in an awkward way. Because of the desire of the speakers to 'decorate' themselves with the prestigious literary pronunciation of [q] and to avoid using the glottal stop, they created the form [fuqɑːd] 'bereavement'. This phenomenon of hypercorrection can be also analyzed in terms of markedness. The glottal stop [ʔ] is the unmarked equivalent of the marked uvular stop [q]. Thus, speakers in a language contact situation involving issues of prestige try to avoid the unmarked forms and use the marked ones.

Prestige plays a key role in terms of background analysis of the creation of pseudocorrections; in addition, the issue of 'authenticity' must be taken into account. In other words, if speakers want to sound authentic in a language or a dialect that may be foreign to them, they might create a pseudocorrection. The following example from South African English illustrates this point. A visitor from Johannesburg regarded American English as a more prestigious variety and wanted to sound authentic and fluent in it. In his dialect, the word *guess* for *think*, as in the sentence *I guess he's coming tomorrow*, does not exist. Thus, instead of saying, *I don't think he's coming tomorrow*, he substituted *guess* for *think*, only because *think* is dominant in his dialect and he knew that in American English one may use *guess*. However, he used *guess* in the wrong environment (*guess* for *think* cannot be used in a negative sentence in American English) and came up with a pseudocorrected form: *I don't guess he's coming tomorrow*. This process may be best termed as hyperadaptation. In Arabic, we see hyperadaptations and authenticity in pseudocorrections in the contact between dialects. Baghdadi Christians and Jews use in their dialect the velar fricative [ɣ] to replace Classical Arabic /r/. Baghdadi Muslims, on the other hand, use /ġ/ and /r/ in accordance with Classical Arabic. When Baghdadi Christians and Jews wish to avoid their characteristic marked feature [ɣ] and sound more 'authentic' in the Muslim variety, they replace it with the unmarked /r/, even when it is not needed by the dominant Muslim dialect and Classical Arabic. Since many of these speakers are 'literate', these examples are quite rare: Christians use [qɑʃuːra] 'spoon' for Muslim [qɑʃuːɣa] (Blanc 1964a:21), and Jews use in their literary idiom [ləjərsəl] 'that he may wash' (Blanc 1964b:21, n. 8). In fact, the verb *ġasal* is perfectly accurate even in Muslim Baghdadi dialect (as is the case in Classical Arabic). Because it had the marked feature [ɣ], though, it was replaced by the unmarked /r/, even when it was not needed. This case of hyperadaptation can also be called 'false regression' (Blau 1970:16) or 'false restitution' (1970:109).

The above example, typical of hyperadaptation, is triggered not just by prestige but also by the motivation for authenticity. In other words, both prestige and authenticity are important factors in the creation of pseudocorrections.

Hypercorrections are only one kind of pseudocorrection made by speakers. Other kinds are hyperadaptations, hypocorrections and mixed forms. While it is possible and important to distinguish between hyper- and hypocorrections (Blau 1981:27–34; Hary 1992:62–69, 313–314) and between hypercorrection and hyperadaptation, it is often difficult to identify mixed forms or simple mistakes. Differentiating several kinds of pseudocorrections is valuable since it aids in examining the different elements exhibited by the various kinds. The following paragraphs exhibit examples for the various pseudocorrected forms, accompanied by an analysis of the difference between hyper- and hypocorrections.

Hypercorrection examples (1) and (2) come from a 16th-century Egyptian Judaeo-Arabic text (published in Hary 1992:144–203):

(1) *wa-ṭalab minhum mablaġ miya wa-xamsūn ʾalf šarīfī ḏahab* 'and he asked from them the sum of one hundred and fifty thousand coins' (Hary 1992:313). The form *wa-xamsūn* is a hypercorrection: in the writer's Egyptian Judaeo-Arabic dialect, as in most Arabic dialects, speakers use the sound plural form *-īn* for all occurrences, but here the writer uses *-ūn*, knowing that Classical

Arabic has the latter form, considered by him to be more prestigious. However, he fails to use the form *-ūn* in the correct syntactic environment, as even in Classical Arabic the form *-īn* would be employed in this example because of the oblique case. In terms of markedness, the oblique form *-īn* is used everywhere in his dialect, so the author chooses the marked form *-ūn* but uses it in the inappropriate syntactic environment.

(2) *naḥwa ʿan iṯna ʿašar rajul* 'about twelve thousand men' (Hary 1992:313). The numeral *iṯnā ʿašar* 'twelve', which is the Classical Arabic nominative form, is a hypercorrection. The author, wanting to use the prestigious marked Classical Arabic form, not current in the colloquial, fails to use it in the correct syntactic environment. In this case the unmarked colloquial form *iṯnay ʿašar* would be appropriate here even in Classical Arabic because it follows the preposition *ʿan* 'about'.

Example (3) presents a special kind of hypercorrection, 'false regression' or 'false restitution':

(3) Tunisian Judaeo-Arabic frequently exhibits the omission of the glottal fricative [h]. Yet, in their writings, Tunisian Jews often reinstitute the glottal fricative, even when not needed, so we find [nhɑːɍ] for 'fire' where [nɑːɍ], which is found in both Tunisian Judaeo-Arabic and Classical Arabic, would have been perfectly appropriate (Cohen 1964:17.4). Out of over-self-denial and in their desire to sound more classical, Tunisian writers embellish their writings when not needed.

Another kind of hypercorrection is 'hyperforeign' (Blau 1970:17, 104–105), as in example (4):

(4) Speakers of Colloquial Arabic may hyperforeign words borrowed into their dialect from foreign languages when their knowledge of the foreign language is insufficient and when they desire to embellish their speech with foreign elements, considered prestigious by them. Thus, in Baghdad and Damascus one may hear [pas] for English *bus*. As a rule, in Arabic dialects the bilabial stop [p] occurs in the dialect only when there is a direct borrowing from a foreign language. In the example of [pas], speakers hyperforeign, as they mistakenly think that the [b] in *bus* comes from the bilabial stop [p] and was assimilated into the Arabic phonemic inventory (/p/ > /b/).

Examples (5) and (6) are hypocorrections:

(5) In a Christian Arabic text, we find *ʿīda ʾilayhi r-rasūl ṯāniya* 'the messenger was sent to him again' (ms. British Museum Or. 5008, 7b, 9, quoted in Blau 1966:153). In Christian Arabic the verb *ʿāda* means 'he returned [trans.]', as it is common to use Form I for Form IV (cf. *bād* 'he destroyed' for *ʾabād*). The writer evidently wants to use an internal passive form, unusual in his dialect. However, instead of using the internal passive of Form IV, as required by Classical Arabic, he goes only halfway and uses the internal passive of Form I. Consequently, the form *ʿīda* is not regular in Christian Arabic (as the internal passive is scarcely used), or in Classical Arabic (where Form IV would have been used). Because the correction is made only halfway, it is a hypocorrection.

(6) In a Judaeo-Arabic text, the phrase *hum bāqiyūn* 'they remain' appears (Blau 1981:29). In the dialects, the sound plural suffix *-īn* supersedes the Classical Arabic sound plural suffix *-ūn* in all occurrences. Therefore, the author does not want to use the form *bāqiyīn* with the prevalent suffix *-īn* of his dialect, because he knows of the Classical Arabic suffix *-ūn*, not employed in his dialect. Therefore, he substitutes *-ūn* for his commonly used suffix *-īn*. However, this is only a 'half correction', since in Classical Arabic the form is *bāqūn* (the *yāʾ* is elided). Consequently, the form *bāqiyūn* appears neither in a Judaeo-Arabic dialect nor in Classical Arabic. Because it is corrected only halfway, it is a hypocorrection.

Both hyper- and hypocorrections stem from the speakers' desire to use a linguistic form of a prestigious variety. However, there are some fundamental differences:

i. The underlying form (to be changed in order to assimilate it to a more prestigious form) in the hypercorrected case does not differ from the prestigious form, i.e., it is perfectly standard usage and should not have been 'corrected' at all, whereas the underlying form in the hypocorrected case differs from the prestigious form. Thus, in example (1) above, the word *wa-xamsīn* 'and fifty' is the underlying form to be changed and is not different from the use in the prestigious variety, since even in the prestigious variety it is used in this syntactic environment. Similarly, in example (2), the underlying form *iṯnay ʿašar* 'twelve' is not different from the Classical Arabic form. The same is true for the other hypercorrection examples. In the hypocorrection example (5), though, the underlying reflexive verbal form differs from the internal passive form used in the prestigious variety. Similarly, the underlying form in example (6) *bāqiyīn* is different from the prestigious word *bāqīn*.

ii. The resulting (i.e. surface) form in the hypercorrected case does not contain any dialectal element, whereas in the hypocorrected case, it contains at least one vernacular feature, since only one feature was changed, thus not changing the whole word enough to equal the prestigious form. Accordingly, in example (1) above, the resulting form *wa-xamsūn* 'and fifty' does not contain any dialectal feature; the same is true for *iṯnā ʿašar* 'twelve', in example (2), whereas in example (5), the resulting form contains a vernacular feature, the use of Form I rather than Form IV, which is typical of Arabic dialects in general. Similarly, in example (6) the resulting form *bāqiyūn* contains a dialectal element (the preservation of the *yāʾ*).

iii. The resulting form in the hypercorrected case goes too far, whereas the resulting form in the hypocorrected case does not go far enough. Thus, in example (1), the resulting form *wa-xamsūn* 'and fifty' goes too far, since the underlying word *wa-xamsīn* is sufficient, even by the grammatical rules of the prestigious variety. Likewise, in example (2) the resulting form *iṯnā ʿašar* 'twelve' goes too far, as the colloquial form *iṯnay ʿašar* should not have been corrected in this environment. However, in the hypocorrected examples, the resulting forms do not go far enough: in example (5) *ʿīda* 'was sent' does not go far enough in the direction of Form IV, and in example (6) *bāqiyūn* does not go far enough because of the preservation of the *yāʾ* rather than its deletion.

iv. The surface form in hypercorrected cases may exist in the prestigious variety in another environment or it may not exist at all in the prestigious variety, whereas the resulting form in hypocorrected cases does not exist in what the speakers consider the prestigious variety, nor does it exist in the colloquial. Consequently, in example (1) the resulting form *wa-xamsūn* 'and fifty' exists in the prestigious variety but in a different syntactic environment, and the same is true for *iṯna ʿašar* 'twelve' in example (2), whereas the resulting forms in example (5), *ʿīda* 'was sent', and in example (6), *bāqiyūn* 'they remain', do not exist in Classical Arabic (the prestigious variety), nor do they exist in the Judaeo-Arabic dialect.

Sometimes, however, it is not so easy to distinguish between the various kinds of pseudocorrections, e.g. example (7), [ŧɑriːʔ] 'way', in Damascene Arabic. Is it a hypocorrection, as it has both a vernacular feature (/q/ > /ʾ/) and a Classical Arabic feature (the preservation of /a/, as we would have expected [ŧriːʔ] in the dialect)? Or is the word simply a mixed form (Blau 1970:107)? The same question holds true for example (8), *faqaṭ bi-l-luġa al-ʿāmmiyya* 'in the colloquial only' (Blau 1970:107), where the Classical Arabic word *faqaṭ* 'only' is used with colloquial word order. Similarly, example (9), the number [θalatiːn] *ṯalatīn* 'thirty', used in Egyptian radio Arabic, may be interpreted as a hypocorrection where the speaker does not go far enough, or as a mixed form of a literary element (the use of the interdental fricative /ṯ/) and a colloquial element (the alveolar stop /t/).

Example (10) from Lebanon [ʕalakəllen] 'at any rate' (Bloch and Grotzfeld 1964:66, n. 5, quoted in Blau 1970:105) is more clearly a mixed form. It contains a colloquial element, the dialect [ə] corresponding to Standard /u/, and a literary element, the → *tanwīn*, although there are many dialects where /kull/ occurs with frozen *tanwīn*.

When a pseudocorrection becomes prevalent in the variety, at one point it ceases to be a pseudocorrection and becomes an accepted form of the variety. In Later Egyptian Judaeo-Arabic there are countless examples where the negative particle *lam* is followed by the perfect: *lam ʿarafūk* 'they did not know you', *lam nadahu* 'they did not call', *lam ʿaṭānī* 'he did not give me' (examples taken from ms. 3, Cairo Collection, Hary 1995:84, n. 18; see also Hary 1992:294–295, 314; Blau 1981:223). The examples are hypocorrected forms that have probably been later standardized in the variety. In the dialect, the regular particle to negate the past is *mā*. The writer does not choose it because it is the unmarked form, dominant in his dialect, and thus not prestigious. He chooses instead the prestigious marked Classical Arabic negation particle *lam*. However, he 'corrects' only halfway: he only changes *mā* to *lam*, but does not replace the perfect form that follows it with a jussive, as required by standard Classical Arabic. This example follows the criteria for hypocorrections: the underlying form *mā* differs from the form in the prestigious variety *lam*; the resulting form contains a vernacular feature (the use of the perfect form, not the jussive), it does not go far enough (to change to the jussive), and the form *lam* followed by the perfect does not exist in the prestigious variety nor in the dialect. However, since these forms were regularly used in Later Egyptian Judaeo-Arabic, they were standardized at one point in writing (as is evident in many manuscripts from that period), and probably reflected also a living usage of *lam* followed by the perfect in this Judaeo-Arabic dialect. This is a case where a pseudocorrection is created, used quite frequently, and then becomes standardized to be part of the dialect.

Another example comes from orthographical evidence. In the old Arabic dialect of the Ḥijāz, the glottal stop was reduced and deleted, thus *raʾs* 'head' became *rās*. Official Qurʾānic reading, though, as a hypercorrection restored the glottal stop and added the orthographic symbol of the *hamza*, probably under the influence of Eastern dialects, where the glottal stop was still retained. Consequently, the *hamza* with its glottal stop pronunciation was added to words even when not needed by the dialect, and since this was prevalent, it became standardized in Arabic. Thus, *kās* 'cup', *radīy* 'bad', and *tawrīx* 'history, date', which were standard in pre-Ḥijāzī dialect, became *kaʾs*, *radīʾ* and *taʾrīx*, respectively, as this hypercorrection standardized in the language (Blau 1970:16–17, 56).

Pseudocorrections are crucial to the study of Arabic. For example, there are many texts of literary written Middle Arabic exhibiting a mixture of Classical Arabic elements, vernacular features, pseudocorrected forms, and standardization of such forms. We can extract the characteristics of the Middle Arabic dialects only by isolating the Classical Arabic elements, on the one hand, and the pseudocorrected forms, on the other hand, leaving us with the dialectal features of the text. Only a careful examination of Middle Arabic texts can reveal dialectal or pseudocorrected features and distinguish between them.

Bibliographical references

Blanc, Haim. 1964a. *Communal dialects in Baghdad*. Cambridge, Mass.: Harvard University Press.

——. 1964b. "Notes on the literary idiom of the Baghdadi Jews". *For Max Weinreich on his seventieth birthday: Studies in Jewish languages, literature, and society*, ed. Lucy S. Dawidowicz a.o., 18–30. The Hague: Mouton.

Blau, Joshua. 1966. *A grammar of Christian Arabic based mainly on South-Palestinian texts from the first millennium*, I. Louvain: Secrétariat du CorpusSco.

——. 1970. *On pseudo-corrections in some Semitic languages*. Jerusalem: Israel Academy of Sciences and Humanities.

——. 1981. *The emergence and linguistic background of Judaeo-Arabic*. 2nd ed. Jerusalem: Ben-Zvi Institute.

Bloch, Ariel and Heinz Grotzfeld. 1964. *Damaszenisch-arabische Texte mit Übersetzung, Anmerkungen und Glossar*. Wiesbaden: Deutsche Morgenländische Gesellschaft.

Cheshire, Jenny. 1984. "Indigenous nonstandard English varieties and education". *Language in the British Isles*, ed. Peter Trudgill, 546–558. Cambridge: Cambridge University Press.

Cohen, David. 1964. *Le parler arabe des Juifs de Tunis: Textes et documents linguistiques et ethnographiques*. The Hague: Mouton.

Fasold, Ralph. 1996. *The sociolinguistics of language*. Oxford and Cambridge, Mass.: Blackwell.

Hary, Benjamin. 1992. *Multiglossia in Judeo-Arabic, with an edition, translation and grammatical study of the Cairene Purim scroll*. Leiden: E.J. Brill.

——. 1995. "Judeo-Arabic in its sociolinguistic setting". *Israel Oriental Studies* 15.129–155.

——. 2003. "Judeo-Arabic: A diachronic reexamination". *International Journal for the Sociology of Language* 163.61–75.

Kutcher, Eduard Yechezkel. 1959. *The language and linguistic background of the Isaiah scroll*. Jerusalem: Magnes Press. [In Hebrew.]

Labov, William. 1972. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

Benjamin Hary (Emory University)

I

ʾIbdāl

ʾIbdāl or *badal* ‘permutation’ is a term used for two distinct phenomena in Arabic linguistics. *ʾIbdāl naḥwī* ‘grammatical *ʾibdāl*’ refers to certain, mainly morphophonological, changes in verbs, nouns, and particles, whereas *ʾibdāl luġawī* ‘lexical *ʾibdāl*’ refers to phonologically and semantically related doublets, triplets, or longer series in the lexicon. In both cases, only consonants (as defined by Arab grammarians, i.e. including the ‘weak consonants’ *w, y, ʾalif*) are concerned; variation of short vowels is not called *ʾibdāl*, the starting point for the medieval authors being the codified written language rather than the spoken vernacular language. (For a survey of changes affecting individual consonants, see Cantineau 1960:17–88.)

1. Grammatical *ʾIbdāl*

Various lists are given of consonants subject to grammatical *ʾibdāl*. Sībawayhi (*Kitāb* II, 313–315) lists *ʾ, ʾalif, h, y, t, m, j, n, l, w*, and Ibn Manẓūr (*Lisān*, root *b-d-l*) lists *ʾ, ʾalif, y, w, m, n, t, h, ṭ, d, j*, remarking that if one adds to these *s* and *l* while removing *ṭ, d*, and *j*, the list is identical with the *ḥurūf az-ziyāda*. Al-Qālī (*ʾAmālī* II, 186) gives the mnemonic phrase *ṭāla yawmun ʾanjadtuhū* (*ṭ, ʾalif, l, y, w, m, ʾ, n, j, d, t, h*).

The cases of *ʾibdāl* listed by Sībawayhi include both standard morphophonological forms (*ʾ* < *y*, as in *qaḍāʾ* < **qaḍāy*) and rare variants caused by partial assimilation (*fuzdu* < *fuztu*, 1st pers. sg. perf. *f-w-z*). *Dunaynīr* is taken as an *ʾibdāl* because the diminutive would imply **dinnār* instead of *dīnār; māʾ* is an example of *ʾ* < *h*, on the basis of the plural *miyāh*, implying a root *m-w-h*; and *fam* is an *ʾibdāl m* < *w* (*fū-*).

Even partial assimilation not visible in the script (as in [ʿambar] for [ʿanbar], or [ʾaždaq] for [ʾašdaq]; *Kitāb* II, 427) is taken by Sībawayhi as an *ʾibdāl*, but total assimilation is discussed under another term, → *ʾidġām*. In a few cases, such as *ʿalijj-* for *ʿalī(y)*, Sībawayhi accepts as *ʾibdāl*s cases that come closer to lexical *ʾibdāl*. In cases like *ḍūriba-ḍāraba*, modern terminology would speak of a vocalic change (*ā-ū*), but for Sībawayhi, this is a consonantal *ʾibdāl* (*w* < *ʾalif*).

2. Lexical *ʾIbdāl*

ʾIbdāl luġawī received less attention in native theory, but this lack of attention was compensated for by monographs listing hundreds of cases. Ibn as-Sikkīt's (d. 857) soundly critical *Kitāb al-qalb wa-l-ʾibdāl* influenced all later studies, but the largest, though completely uncritical, collection was ʾAbū ṭ-Ṭayyib al-Luġawī's (d. 962) *Kitāb al-ʾibdāl* (see El Berkawy 1981). The mutual relations of these and other *ʾibdāl* works, or chapters in other works, and their later influence have been studied in Hämeen-Anttila (1993).

The terminology of the lexicographers fluctuated, and one finds as synonyms for *ʾibdāl* also *qalb* (used by both Sībawayhi and Ibn as-Sikkīt; later usually in the sense of ‘metathesis’), *muʿāqaba,* and sometimes *naḍāʾir* (az-Zajjājī, *ʾIbdāl* 1).

Ibn as-Sikkīt lists pairs that, if taken as examples of phonological changes, often match modern linguistic criteria. Thus, a case typical of Ibn as-Sikkīt is *baʿṯara* ~ *baḥṯara* (*Qalb* 86), where *baḥṯara* is quite obviously a variant caused by the following voiceless consonant. This example contains a salutary caveat, though: the pair clearly derives from the *Qurʾān*, where the rather rare *buʿṯira(t)* occurs twice (Q. 82/4, 100/9). The pair of Ibn as-Sikkīt does not stem from spoken language but from the recitation of the *Qurʾān* – the relevant article *b-ʿ-ṯ-r* in Ibn Manẓūr's *Lisān* begins symptomatically by mentioning this *qirāʾa*. Whether the variant *baḥṯara* really existed – it may well have – in other varieties of Arabic cannot be proven on the basis of Ibn as-Sikkīt. Many examples have their origin in texts, whether Qurʾānic or poetic (e.g. → *rajaz*, notorious for its extravagances). Thus, even Ibn as-Sikkīt's examples have to be used with caution.

ʾAbū ṭ-Ṭayyib, on the other hand, gives most improbable pairs, aiming at systematically listing every theoretically possible pair and finding examples for it. Thus, we find in his book seven cases of *ʾibdāl* between *q* and *m* (*ʾIbdāl* II, 365–368), which he gets through devices such as interpreting *ʾitbāʿ* as *ʾibdāl* (as in *ʾinnahā tamīsu maysan wa-taqīsu qaysan; ʾIbdāl* II, 367, taken from a *ḥadīṯ*) or searching for various meanings for words until he finds one common to two etymologically unrelated, though phonetically similar, words. ʾAbū ṭ-Ṭayyib (as quoted in as-Suyūṭī, *Muzhir* I, 460, from the initial lacuna of *Kitāb al-ʾibdāl*), though, does not claim that the members of the pairs would be etymologically or historically related. In his opinion, *ʾibdāl* is an accidental phenomenon, caused by the fact that various tribes used similar words to denote one thing. The same idea is seen in *al-Muxaṣṣaṣ* (XIII, 274), where Ibn Sīda introduces his chapter on lexical *ʾibdāl* as *Bāb mā yajīʾu maqūlan bi-ḥarfayn wa-laysa badalan* 'Chapter of what is said in two different ways but is not [grammatical] *badal*'.

Variation introduced after the beginning of the Islamic period is not called *ʾibdāl*, but → *laḥn*. Thus, *ʾibdāl* words are correct Arabic, and *laḥn* words are mispronunciations and fall outside correct language. From a modern point of view, *ʾibdāl* and *laḥn* (as far as the latter concerns phonological matters) differ from each other only in chronology: pre-Islamic phonetical variation is called *ʾibdāl*, Islamic variation *laḥn*.

A special case are the orthographic *ʾibdāl*s, i.e. pairs that are due to misspellings of rare words, one of the pair being a ghost word. Still, *ʾibdāl* monographs contain much material valuable for the study of early dialectal variation and historical linguistics, but they must be used with circumspection.

Bibliographical references

Primary sources

ʾAbū ṭ-Ṭayyib, *ʾIbdāl* = ʾAbū ṭ-Ṭayyib ʿAbd ar-Raḥmān ibn ʿAlī al-Luġawī, *Kitāb al-ʾibdāl*. Ed. ʿIzz ad-Dīn at-Tanūxī. 2 vols. Damascus: al-Majmaʿ al-ʿIlmī al-ʿArabī, 1379–1380/1960–1961.

Ibn Manẓūr, *Lisān* = ʾAbū l-Faḍl Jamāl ad-Dīn Mukarram ibn Mukarram Ibn Manẓūr, *Lisān al-ʿArab*. Ed. ʾAlī Šhīrī. 18 vols. Beirut: Dār ʾIḥyāʾ at-Turāṯ al-ʿArabī, 1408/1988.

Ibn Sīda, *Muxaṣṣaṣ* = ʾAbū l-Ḥasan ʿAlī ibn ʾIsmāʿīl al-Maġribī Ibn Sīda, *al-Muxaṣṣaṣ*. 3rd ed. 17 vols. Cairo: Dār al-Fikr, 1970.

Ibn as-Sikkīt, *Qalb* = ʾAbū Yūsuf Yaʿqūb ibn ʾIsḥāq Ibn as-Sikkīt, *Kitāb (al-qalb wa)-l-ʾibdāl*. Ed. ʿAlī an-Najdī Nāṣif. Cairo: Majmaʾ al-Luġa al-ʿArabiyya, 1398/1978.

Qālī, *ʾAmālī* = ʾAbū ʿAlī ʾIsmāʿīl ibn Qāsim al-Qālī, *Kitāb al-ʾAmālī*. 2 vols. N.p., n.d.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. 2 vols. Cairo: Būlāq, 1316 A.H.

Suyūṭī, *Muzhir* = Jalāl ad-Dīn ʾAbū l-Faḍl ʿAbd ar-Raḥmān ibn ʾAbī Bakr as-Suyūṭī, *al-Muzhir fī ʿulūm al-luġa wa-ʾanwāʿihā*. Ed. Muḥammad ʾAḥmad Jār al-Mawlā Beg, ʿAlī Muḥammad al-Bijāwī, and Muḥammad ʾAbū l-Faḍl ʾIbrāhīm. 2 vols. Beirut: al-Maktaba al-ʿAṣriyya, 1406/1986.

Zajjājī, *ʾIbdāl* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān ibn ʾIsḥāq az-Zajjājī, *Kitāb al-ʾibdāl wa-l-muʿāqaba wa-n-naḍāʾir*. Ed. ʾIzz ad-Dīn at-Tanūxī. Damascus: al-Majmaʿ al-ʿIlmī al-ʿArabī, 1381/1962.

Secondary sources

Cantineau, Jean. 1960. *Etudes de linguistique arabe*. Paris: C. Klincksieck.

El Berkawy, Abdel Fatah. 1981. *Die arabischen* Ibdāl-*Monographien insbesondere das* Kitāb al-Ibdāl *des Abū ṭ-Ṭayyib al-Luġawī*. Ph.D. diss., University of Erlangen.

Hämeen-Anttila, Jaakko. 1993. *Lexical* ibdāl. I. *Introduction. Source studies*. (= *Studia Orientalia*, 71.) Helsinki: Finnish Oriental Society.

Jaakko Hämeen-Anttila
(University of Helsinki)

Ibero-Romance

1. Introduction

The multifarious contacts between Arabic and Ibero-Romance have led to prolific cultural exchange. The linguistic aspect of this exchange materializes as mutual influence between different varieties of Arabic and the Ibero-Romance languages and dialects. Arabic loanwords are defined as 'words introduced directly from Arabic into another language'. The superordinate term is Arabism, "an Arabic word or meaning introduced into another language, or a typical Arabic syntagm imitated in another language" (Kiesler 1994:44), e.g. Spanish *azúcar* < *as-súkkar*, *llenar el ojo* 'to delight' ← *mala'a l-ʿayna*. Arabisms extend semantically from *limón* 'lemon' to *alboroto* 'uproar' and Andalusian *califa* 'rascal'. Formally, most of them are substantives, although there are adjectives like *alazán* 'chestnut', verbs like *halagar* 'to flatter', adverbs like *de marras* 'long ago', the preposition *hasta* 'up to', the demonstrative particle *he* (*aquí*) '(here) is', and interjections like *ojalá* 'I hope so!'. Chronologically, they reach from *barrio* 'district' (10th century) and *aduana* 'customs' (13th century) to contemporary *jalifa* 'Moroccan caliph [in the former Spanish protectorate]' (20th century?). Geographically, they remain restricted to local or regional diffusion (Andalusian *aljofifa* 'floorcloth'), or have spread nationally (*aduana*) or internationally (*taza* 'cup'); many internationalisms, e.g. *alcohol* and *algebra*, have been diffused via Middle Latin. As for stylistic levels, we find learned words like *almagesto* 'almagest', technical terms like *alfiz* 'square [of an arch]', everyday words like *aceite* 'oil' and *aldea* 'village', colloquialisms like *jamar* 'to chew', and vulgarisms like *cipote* 'tool'. There are even loanwords in the terminology of children's games, e.g. *alelevi* 'hide-and-seek' (Corriente 1999:65, emphasizing the role of Moorish nursemaids).

The following types of Arabisms can be distinguished (cf. Kiesler 1994:35–57). First, there are direct (*alfombra* 'carpet' < *al-xúmra*) and indirect Arabisms, Indirect Arabisms have either been introduced via third languages (e.g. Spanish *tarifa* 'tariff' < Catalan *tarifa* < *taʿrīfa*; many words passed through Middle Latin), or they are intralinguistic formations (*alquilar* 'to rent' ← *alquilé* [< *al-kirā'* 'rent'] + *-ar*). In what follows, only direct Arabisms will be considered. Second, there are loanwords and loan names: different sorts of toponyms (*Alhambra*, *Alcalá*, *Almería*, *Guadalquivir*, [Plaza de] *Bibarrambla* in Granada; and names of more than 170 cities founded by the Arabs in the Iberian Peninsula) and learned astronyms (*Aldebarán*, *Algol*, *Altaír*, *Rigel*). Third, we must distinguish between loanwords and calques (see below, Sec. 5). Finally, Arabisms can be fully integrated, e.g. *arroz* 'rice' < *ar-ruzz*, or be 'foreign words', e.g. *Islam* < *'islām* (where *m* = [n]).

2. ARABISMS IN PORTUGUESE, SPANISH, AND CATALAN

Arabic influence in the Iberian Peninsula has differed in intensity according to place and time. It was very important in the Middle Ages, with a majority of first documentations occurring in the 13th century, although special studies are lacking (Kiesler 1994:78–80). Geographically, this influence is stronger in the south than in the north, which can easily be illustrated by toponyms, e.g. by the names of the five big rivers in the peninsula (*Guadalquivir* is entirely Arabic, *Guadiana* is only half Arabic, and *Tajo* is phonetically influenced, while *Duero* and *Ebro* are wholly Romance), and, naturally, by the Arabic loanwords in the southern dialects; they very probably are stronger in the western than in the eastern part of the peninsula. There is some evidence of a deeper level of Arabization in the west (Kiesler 2003): the higher number of loanwords existing only in Portuguese (22 vs. only 17 in Spanish, and only 13 in Catalan); the number of Arabisms in the basic vocabulary (1.13% in Portuguese vs. 0.92% in Spanish and 0.75% in Catalan; Kiesler 1994:70); the high number of loanwords preserved in Portuguese with corresponding forms abandoned in Spanish; a possibly higher ratio of toponyms of Arabic origin in Portugal than in Spain; and a probably higher number of calques in Portuguese. Finally, the linguistic contacts between Arabic and Portuguese continued for a long time even after the Reconquista, in Africa and in the Orient, while those between Arabic and Spanish were much less intensive after 1492.

Among 300 direct Arabisms – 100 in each language – there are 248 that appear in two or three of the main Ibero-Romance languages, while 52 exist in only one of these languages. Among these 248 Arabisms, 166 have similar forms and meanings, while the remaining 82 show semantic, formal, diachronic, or diatopic differences or differences in use. Portuguese *taça*, for instance, means 'glass', while Spanish *taza* and Catalan *tassa* have the meaning 'cup'. Sometimes there are different Arabisms for the same concept, as in Portuguese *alfândega* 'customs' (according to Corriente 1999, s.v. < **al* + *ḥaḏ̣ḏ̣* + -IQA) and Spanish *aduana*, Catalan *duana* < (*ad*-) *dīwān*. Diachronic differences appear when one language substitutes a new word for an Arabic loanword, the most frequent cases seeming to be those where Portuguese conserves Arabisms and Spanish gives up corresponding forms, as occurred with the words for 'flock', 'lettuce', and 'secondhand dealer', where Portuguese retains *alcateia* < *al-qaṭīʿa*, *alface* < *al-xass*, *algibebe* < *al-jabbāb*, whereas Spanish has substituted *manada*, *lechuga*, *ropavejero* for Old Spanish *alcatea*, *alfaça*, *aljabibe*. Diatopic differences can be seen where an Arabism exists only in certain dialects of one language, as in *aletría* 'noodle' in the Spanish of Murcia and in Portuguese *aletria* < *al-ʾiṭriya*, but Spanish *fideo*(*s*) (for which see Corriente 1999, s.v. *fideu*). Differences in use concern those cases where an Arabism exists in one language only as an unusual variant, e.g. Portuguese *xairel* 'caparison' and Spanish *jirel* < *jilāl*, but the usual word in Spanish is *gualdrapa*, probably of Latin origin, or, vice versa, Spanish *alquiler* 'rent(ing)' and Portuguese *alquilé*(*r*) besides usual Portuguese *aluguer* < Latin LOCARIUM.

Interestingly, the number of Arabisms shared by Portuguese and Spanish (83) is much higher than the number of Arabisms shared by Spanish and Catalan (27; Kiesler 1996:478). The typical cases may be illustrated by Tables 1 and 2.

3. SEMANTIC FIELDS OF IBERO-ROMANCE ARABISMS

Like other loanwords, Ibero-Romance Arabisms are usually classified according to semantic fields. The most important of these fields are the following (Kiesler 1994:80–82, 115–135):

i. Names of plants: Portuguese *alcaçuz* 'licorice' < *ʿirq as-sūs*, *alecrim* 'rosemary' < *al-ʾiklīl*, *alface* 'lettuce', *cenoura* 'carrot' < *safunāriya*, *tremoço* 'lupin' < *turmūs*; Spanish *acelga* 'beet' < *as-sílqa*, *adelfa* 'rosebay' < *ad-dífla*, *alubia* 'bean' < *al-lūbiyā*,

Table 1. Arabisms shared by the three Ibero-Romance languages (Kiesler 1996:473)

Arabic	Portuguese	Spanish	Catalan	meaning
(as-)súkkar	*açúcar*	*azúcar*	*sucre*	'sugar'
(al-)quṭún	*algodão*	*algodón*	*cotó*	'cotton'
(al-)maxzan	*armazém*	*almacén*	*magatzem*	'warehouse'
al-fānīd	*alfenim*	*alfeñique*	*alfenic*	'almond paste'
al-qawwād	*alcaiote*	*alcahuete*	*alcavot*	'procurer'
ar-ruzz	*arroz*	*arroz*	*arròs*	'rice'
bárri	*bairro*	*barrio*	*barri*	'district'
laymūn(a)	*limão*	*limón*	*llimona*	'lemon'
makīla	*maquia*	*maquila*	*maquila*	'multure'

Table 2. Arabisms shared by Portuguese and Spanish (Kiesler 1996:474)

Arabic	Portuguese	Spanish	Catalan	meaning
al-xuzāmā	*alfazema*	*alhucema*	*(espígol)*	'lavender'
al-muxádda	*almofada*	*almohada*	*(coixí)*	'pillow'
al-búnduqa	*almôndega*	*albóndiga*	*(pilota)*	'meatball'
ḥáttā	*até*	*hasta*	*(fins)*	'up to'
at-tūn	*atum*	*atún*	*(tonyina)*	'tuna'
az-zayt	*azeite*	*aceite*	*(oli)*	'oil'
fulān	*fulano*	*fulano*	*(en tal)*	'so-and-so'
law ša lláh	*oxalá*	*ojalá*	*(tant de bo!)*	'I hope so!'
zárqa	*zarco*	*zarco*	*(blau clar)*	'light blue'

berenjena ‘aubergine’ < *bāḏinjān*, *retama* ‘broom’ < *ratama*; Catalan *alfàbrega* ‘basil’ < *al-ḥabaqa*, *alfals* ‘alfalfa’ < *al-faṣfaṣ*, *atzerola* ‘haw’ < *az-zaʿūra*, *bacora* ‘early fig’ < *bākūra*, *llessamí* ‘jasmine’ < *al-yāsamīn*.

ii. Names of victuals, clothes, etc.: Portuguese *acepipe* ‘tidbit’ < *az-zabīb*, *aletria* ‘noodle’, *alféloa* ‘icing’ < *al-ḥalwā*, *fatia* ‘slice’ < *fatīla*, *moxama* ‘salted tuna’ < *mušammaʿ*; Spanish *albóndiga* ‘meatball’, *alhaja* ‘jewel’ < *al-ḥāja*, *alpargata* ‘canvas shoe’ < Hispano-Arabic *al-parġāt*, *jofaina* ‘washbasin’ < *jufayna*, *zaragüelles* ‘wide-legged pants’ < *sarāwīl*; Catalan *alamara* ‘braid’ < *al-ʿamāra*, *arracada* ‘pendant earring’ < *qarraṭ*, *arrop* ‘boiled must’ < *ar-rubb*, *barnús* ‘bathrobe’ < *barnūs*, *escabetx* ‘marinade’ < *sikbāj*.

iii. Designations in the spheres of agriculture (where the terminology of irrigation is especially important), handicrafts, trade, and household: Portuguese *alfaiate* ‘tailor’ < *al-xayyāṭ*, *almece* ‘whey’ < *al-mayṣ*, *almofaça* ‘currycomb’ < *al-miḥassa*, *almofate* ‘awl’ < *al-mixyaṭ*, *ceifa* ‘harvest’ < *ṣayfa*; Spanish *aceña* ‘water mill’ < *as-sāniya*, *acequia* ‘irrigation ditch’ < *as-sāqiya*, *albañil* ‘bricklayer’ < *al-bannāʾ*, *almoneda* ‘auction’ < *al-munādā*, *noria* ‘waterwheel’ < *nāʿūra*; Catalan *almàssera* ‘oil mill’ < *al-miʿṣara*, *almodí* ‘corn exchange’ < *al-mudī*, *assut* ‘dam’ < *as-sudd*, *caduf* ‘scoop’ < *qādūs*, *naquera* ‘trough’ < *naqīr*.

iv. Designations in the sphere of social organization: Portuguese *aldeia* ‘village’ < *aḍ-ḍayʿa*, *alfândega* ‘customs’, *algoz* ‘executioner’ < *al-ġuzz*, *bairro* ‘district’, *refém* ‘hostage’ < *rahn*; Spanish *aduana* ‘customs’, *albacea* ‘executor’ < *al-waṣiyya*, *alcalde* ‘mayor’ < *al-qāḍī*, *arrabal* ‘suburb’ < *ar-rabaḍ*, *marchamo* ‘seal’ < *maršam*; Catalan *aljama* ‘Arab quarter’ < *al-jamāʿa*, *almogàver* ‘raider’ < *al-muġāwir*, *atzucac* ‘blind alley’ < *az-zuqāq*, *rambla* ‘avenue’ < *ramla*, *tàvega* ‘dungeon’ < *ṭabaqa*.

v. Names of colors: Spanish *alazán* ‘chestnut’ < *al-ʾazʿar*, *azul* ‘blue’ (according to Corriente 1999, s.v. *atzur,* < Middle Latin), *carmesí* ‘crimson’ < *qirmizī*, *escarlata* ‘scarlet’ < *siqirlāṭ*, *turquí* ‘indigo’ < *turkī*, *zarco* ‘light blue’, and corresponding Portuguese *alazão*, *azul*, *carmesim*, –, *turqui*, *zarco*.

vi. Names for measures and weights, often obsolete today, e.g. Spanish *adarme* ‘dram’ < *ad-dirham*, *almud* ‘almud’ < *al-mudd*, *fanega* ‘fanega’ < *fanīqa*, *quilate* ‘carat’ < *qīrāṭ*, *quintal* ‘quintal’ < *qinṭār*, *resma* ‘ream’ < *rizma*, and corresponding Portuguese and Catalan words.

vii. There are a few Arabisms for feelings, emotions, etc. (Kiesler 1994:82), e.g. *alborozo* ‘joy’ < *al-burūz* and *zalema* ‘cajolery’ < *as-salām ʿalayk*.

4. Linguistic adaptation of Arabic loanwords

The adaptation of Arabic loanwords concerns the phonological, morphological, and lexical levels. The phonological adaptation is treated extensively in Corriente (1999:22–50; for Catalan cf. Kiesler 1995). The vocalism of Arabic loanwords “is not exactly predictable from Arabic with unequivocal generative rules” (Corriente 1999:25). Arabic /a/ is conserved (Hispano-Arabic *maṭráqa* > *matraca* ‘rattle’), or rendered as /e/ or /i/ in palatalized contexts or through → *ʾimāla*: (*as-*)*sāqiya* > *acéquia*, *acequia*, *sèquia*; *al-mihrās* > *almofariz*, *almirez* ‘mortar’. Similarly, Arabic /i/ > /i/ or /e/: (*al-*)*qaṭīfa* > *alcatifa*, Catalan *catifa* ‘carpet’; *al-misk* > *almíscar*, *almizcle*, *almesc* ‘musk’; *taʿlīqa* > Old Portuguese, Spanish *talega*, Catalan *taleca* ‘bag’. And Arabic /u/ > /u/ or, frequently, /o/: (*as-*)*súkkar* > *açúcar*, *azúcar*, *sucre* besides Hispano-Arabic (*al-*)*quṭún* > *algodão* (with assimilation to *-ão*), *algodón*, *cotó*. Diphthongs are frequently conserved in Portuguese (where *ou* > /ọ/) but monophthongized in Spanish and Catalan: *aḍ-ḍáyʿa* > *aldeia* but Spanish and Old Catalan *aldea*, *as-sáwṭ* > *açoute* vs. *azote*, *assot* ‘whip’.

As for consonants, the plosives /b, d, t, ḍ, ṭ, k, q/ are normally substituted with the corresponding Romance phonemes, while /ʾ/ is not reflected in loanwords. Thus *al-birqūq* > *albricoque, albaricoque, albercoc* ‘apricot’, (*ad-*)*dīwān* > *aduana, aduana, duana*, *turkī* > *turqui, turquí, turquí*, *aḍ-ḍáyʿa*, *al-qāḍī* (see above; sometimes /ḍ/ > *-ld-*, see Corriente 1999:31; Kiesler 1994:161, 173), *ṭássa* > *taça* ‘glass’, *taza, tassa* ‘cup’, *kubāba* > *cubeba* ‘cubeb’, Hispano-Arabic *ʾispinā̄x*(*a*) > *espinafre*, *espinaca*, *espinac* ‘spinach’. The affricate

/j/ regularly > (*d*)*ž* (> Spanish /χ/): *jarra* > *jarra*, *jarra*, *gerra* ‘jug’. Among the fricatives, /f, š, ġ/ are normally rendered as /f, š (> Spanish /χ/), g/, while /ṯ, ḏ, ḏ̣/ are rare in etyma of Arabisms, and /ʿ/ and /h/ are generally not reflected (sometimes /h/ > /f/, as in *Faro* < *Hārūn*): *al-fīl* ‘elephant; bishop [in chess]’ > *alfil*, *alfil*, *alfil*, in Romance only ‘bishop [in chess]’, *šabbāk* > *xaveco*, *jabeque*, *xabec* ‘xebec’, *ġadāmasī* > *guadamecil*, *guadamecí*, *guadamassil* ‘embossed leather’; Hispano-Arabic *al-bárdaʿa* > *albarda* ‘packsaddle’, *az-zahr* > *azar*, *azar*, *atzar* ‘chance’. Arabic /s, ṣ, z/ are frequently rendered as affricates in Old Portuguese (> /s/) and Old Spanish (> /θ/), as /s/ in Catalan: (*as-*)*sāniya* > *azenha*, *aceña*, *sínia* ‘watermill’, (*aṣ-*)*ṣíbar* > *azebre*, *acíbar*, *sèver* ‘aloe’, *az-zahr* (see above). For /x/ and /ḥ/ there are different solutions, i.e. /f/, often /h/ in Spanish, more rarely /k/ (Corriente 1999:37–38): Hispano-Arabic *xaláq* > *afagar*, *halagar* (Old Spanish *falagar*), *afalagar* ‘to flatter’ (Corriente 1999, s.v. *afagar*), (*al-*)*xaršūfa* > *alcachofra*, *alcachofa, carxofa* ‘artichoke’, *al-ḥabaqa* > *alfavaca, albahaca*, *alfàbrega* ‘basil’, Hispano-Arabic *al-maṭráḥ* > *almadraque*, Old Spanish *almadraque*, Old Catalan *almatrac* ‘mattress’. Nasals and liquids are normally conserved, see examples above. For prosodics and combinatory phonetics, see Corriente (1999:39–50).

Morphological adaptation: Spanish and Catalan – but not Portuguese – have borrowed the *nisba*-suffix *-í*, still productive today for the formation of Oriental(izing) gentilics like *andalusí* and *bengalí* (Corriente 1999:51–52; Kiesler 1994:42–43). Borrowed verbs, whose number is considerably higher than traditionally assumed (Corriente 1999:53), naturally take Romance endings, thus *_en_ + _balá_ + _-ar_ > _embelecar_ ‘to deceive’, *_en_ + _juf_ + _-ar_ > _enchufar_ ‘to connect’, *halagar* (see above). A much-discussed problem is the ‘agglutination’ of the Arabic article *al-* in many Spanish and Portuguese Arabisms, somewhat less in Catalan, which distinguishes Ibero-Romance Arabisms from Italian ones, cf. *açúcar*, *azúcar* vs. *sucre*, Italian *zucchero*; see the discussion in Noll (1996), Corriente (1999:57–63). The explanation – though some questions remain – is to be seen in the stable form of Hispano-Arabic *al-* (no elision), the special sociolinguistic situation in al-Andalus, where the article was used as a prestige marker, and the role of the only superficially Arabicized Berbers, who frequently agglutinated *al-* in the Arabic they spoke. Some typical examples can be seen in Table 1.

At the lexical level, we may distinguish contamination through mere analogy from remotivating folk etymology (cf. Corriente 1999:19, n. 3). Examples of contamination are *aljama* ‘Arab quarter’ < *al-jamāʿa* ‘community’ × *al-jāmiʿ* ‘mosque’, *algoritmo* ‘algorithm’ < *al-Xuwarízmi* × Greek ἀριθμός ‘number’, *guadamecí* < *ġadāmasī* × *guad-* as in *Guadalquivir*, etc. Folk etymology may be illustrated with *caparrosa* ‘vitriol’ < (*zāj*) *qubrusī*, *mamarracho* ‘ninny’ < *momarracho* (< Hispano-Arabic *_muharráj_ + _momo_ ‘gesture’) + *mamar*, *matalahuva* ‘anise’ < *matahalúa* < Hispano-Arabic *ḥabbat ḥalúwwa*. The importation of Arabic loanwords has endowed the Ibero-Romance languages with new homonyms and synonyms. Curiously, there are a good many cases of homonymy where both words derive from Arabic, e.g. Portuguese *albardão* ‘cheeky’ < *al-bardān*, *albardão* ‘packsaddle’ ← *albarda* < *al-bárdaʿa*, Portuguese *alifafe* ‘bedspread’ < *al-liḥāf*, *alifafe* ‘windgall’ < *an-nafáxa*; Spanish *limón* ‘lemon’ < *laymūn*, *limón* ‘shaft’ < *al-ʿamūd*. Synonyms are either both of Arabic origin or not. Thus Portuguese *adelo* ‘secondhand dealer’ < *ad-dallāl* ~ *algibebe* ‘id.’ < *al-jabbāb*; Spanish *orozuz* ‘licorice’ < *ʿurūq sūs* ~ *alcazuz* ‘id.’ < *ʿirq as-sūs* ~ *rabazuz* ‘id.’ < *rubb as-sūs* besides *regaliz* ‘id.’ < Late Latin LĬQUĬRĬTĬA. For different types of semantic change, see Kiesler (1994, index).

5. A TENTATIVE TYPOLOGY OF IBERO-ROMANCE ARABISMS

Apart from loanwords, the loan affix *-í*, and loan names, there are other types of linguistic loans, which can preliminarily be classified in three groups.

Loan blends are ‘partial translations’: *māʾ zahár* > *agua de azahar* ‘orange-flower water’; *_māʾ an-nafḥ_ > Catalan _aiguanaf_ ‘id.’. Arabic *jild ġadāmasī* > *cuero guadamecí*, Arabic *fī l-bāṭil* > *en balde* ‘in vain’ (Kiesler 1994:158); Hispano-Arabic *wáḥida bi-wáḥida* > *guájete por guájete* (Corriente 1999, s.v.) ‘one thing for another’; Arabic *jawz aš-šarq* > Catalan *nou d’eixarc(h)* (Corriente 1999, s.v. *axarque*) ‘nutmeg’.

Calques can be distinguished in two types. On the word level there are (a) *loan translations* (‘literal translations’), e.g. Portuguese *mãe d’água* ‘fountain’ ← *ʾumm al-māʾ*; (b) *loan meanings* (‘borrowed senses’), e.g. Navarrese *fila* (*de agua*) ‘portion [of irrigation water]’ ← *xayṭ* (Corriente 1999, s.v.) and Old Spanish *casa* ‘house’ + ‘town’ *dār*; and (c) hardly explored *loan creations* (‘formally independent formations’), e.g. Old Spanish *endereçamiento* ‘straightening’ ← *endereçar*, in imitation of Arabic *taqwīm* ← *qawwama*. On the level of syntagms there are influences in the realms of (a) *phraseology*, e.g. *saíu ao pai, salió a su padre* ‘he took after his father’ ← Hispano-Arabic *xaraj li-wildihi*; also many formulas and proverbs (Kiesler 1994:46–47); (b) *word formation*, e.g. Old and Classical Spanish, *fi de nemiga* ‘enemy’, *hijo del naipe* ‘who always wins’, imitating Arabic *ibn al-layl* ‘thief’, *ibn as-sabīl* ‘wanderer’; (c) *syntax*, e.g. the Old Spanish type *el caballo del rey y su espada* ‘the king’s horse and sword’ ← *farasu l-maliki wa-sayfuhū*, and probably Portuguese *falar em* (*um tema*) ‘to talk about (a subject)’ ← *takallama fī*; and (d) the use of the *figura etymologica,* e.g. in Llull (d. 1316).

Finally, there are borrowings of literary themes and motifs, as Galmés de Fuentes (1999:186) has convincingly demonstrated, e.g. in Spanish mysticism and in Ramón Llull, where we find motifs like ‘the blind obedience to the beloved’, ‘the joyful suffering’, or ‘the unrequited love’. These types of borrowings represent a challenge for contact linguistics insofar as their linguistic repercussions cannot be seized by current classifications of loans. It is true, for instance, that we can classify the designation *amic* ‘friend/lover’ in Llull ← *muḥibb* as loan meaning or perhaps better as loan usage, but there is no such generally accepted linguistic category for the borrowing of a motif like ‘the blind obedience to the beloved’ (Kiesler 2002).

Thus, the revision of the typology of loans is one of the tasks of future research. Another task is the description of the exact filiation of Arabic loanwords. Many of the Arabic words that have passed to the Ibero-Romance languages are not autochthonous but are taken from other languages (cf. Corriente 1999:22), especially from Persian (*taza* < *ṭássa* < Pehlevi *tašt*), Sanskrit (*azúcar* < *as-súkkar* < Pali *sakkharā*), and Greek (*albóndiga* < Hispano-Arabic *al-fúndaq* < Arabic *funduq* < Aramaean *pendĕqā* < Greek πανδοκεῖον). On the other hand, Ibero-Romance languages have served as transmitters for Arabic loanwords introduced into → English and other languages, even if → Italian seems to have played a more important role in this sense, e.g. English *apricot* < Spanish or Portuguese, *aubergine* < French < Catalan, *spinach* < French < Old Provençal < Catalan, *typhoon* < Portuguese.

The study of Arabic influence on Ibero-Romance languages and literatures has made considerable progress in the last decades, thanks to scholars like Corriente and Galmés de Fuentes; still, much remains to be done. We know, for instance, rather little about “the marginal and final phenomena of the history of Romance Arabisms”, such as their introduction into thieves’ slang and into the language of children’s games (Corriente 1999:65). Yet, it is clear that the linguistic and extralinguistic influences – like influence on the origin of universities – of Arabo-Islamic civilization on Medieval Europe, and especially on the Ibero-Romance languages, have been greater and more profound than generally thought.

Bibliographical references

Corriente, Federico. 1999. *Diccionario de arabismos y voces afines en iberorromance* [Dictionary of Arabisms and related words in Ibero-Romance]. Madrid: Gredos.

Galmés de Fuentes, Álvaro. 1999. *Ramón Llull y la tradición árabe: Amor divino y amor cortés en el* ‘Llibre d’amic e Amat’ [Ramón Llull and Arabic tradition: Divine and courtly love in the ‘Book of the friend and the Beloved’]. Barcelona: Quaderns Crema.

Kiesler, Reinhard. 1994. *Kleines vergleichendes Wörterbuch der Arabismen im Iberoromanischen und Italienischen*. Tübingen: Francke.

——. 1995. “Zur Phonetik der katalanischen Arabismen.” *Zeitschrift für Romanische Philologie* 111. 171–205.

——. 1996. “La repartición de los arabismos en iberorromance” [The distribution of Arabisms in Ibero-Romance]. Lüdtke (1996:471–479).

——. 2002. Review of: Galmés de Fuentes (1999). *Zeitschrift für Romanische Philologie* 118.301–305.

——. 2003. “¿Hay más arabismos en español o en portugués?” [Are there more Arabisms in Spanish or in Portuguese?]. *Actas del XXIII Congreso Internacional de Lingüística y Filología Románica*, ed. Fernando Sánchez Miret III, 263–266. Tübingen: Niemeyer.

Lüdtke, Jens (ed.). 1996. *Romania Arabica: Festschrift für Reinhold Kontzi zum 70. Geburtstag.* Tübingen: Narr.

Noll, Volker. 1996. “Der arabische Artikel *al* und das Iberoromanische.” Lüdtke (1996:299–313).

Reinhard Kiesler (Würzburg, Germany)

Ibero-Romance Loanwords

1. Language contacts between Arabic and Romance

At the beginning of the 8th century, armies of Muslim Arabs and somewhat Arabicized Berbers invaded the Iberian Peninsula. The lexical interference of Ibero-Romance with Arabic was the unavoidable consequence of contacts between these armies, who imported Arabic dialects, and the local inhabitants, who spoke Proto-Romance dialects derived from Low Latin. The invaders succeeded in creating a new geopolitical entity, soon called → al-Andalus by its mixed population. The guest language obviously occupied a dominating position in the ensuing situation of bilingualism. In addition to being the tongue of the ruling classes, administration, and religious services, Arabic was soon to become the main linguistic link with the rest of the Islamic world, with which trade and intercourse would be increasingly frequent, and even between the different ethnic groups of al-Andalus.

However, as the newcomers had not brought along women, they took native women as wives, thus creating bilingual households. The local Hispanic population, which could not, would not, and did not forsake their ancestral tongue immediately, nevertheless soon had to communicate with the monolingual newcomers. The only possible outcome of this situation was that in a matter of a few years almost everybody became more or less bilingual. It took the guest language, Arabic, much longer to drive Romance out of fashion even in low registers, and four or five centuries finally to reign alone in the country.

In the meantime, of course, there was interference between the spoken dialects of Arabic and Romance, above all lexical borrowing in both directions, to the point that this borrowing became one of the most characteristic features of both → Andalusi Arabic and Andalusi or Southern Romance (formerly misnamed 'Mozarabic'). The earliest attempt at producing a survey of the matter was Simonet (1888), to which Corominas (1951) and Griffin (1961) added much-needed methodological approaches, while Corriente (1992:125–142) constitutes the latest account of this subject.

Romance loanwords in Andalusi Arabic can be studied from diverse linguistic points of view, some diachronic, e.g. their chronology, some synchronic, e.g. their degree of integration into the lexical stock and their semantic distribution, and even panchronic, like their frequency ratio in the texts and over the lexicon, or their functional distribution, from the point of view of the relative frequency of the different grammatical categories.

2. The diachronic study of Romance loanwords in Arabic

From a diachronic viewpoint, it is noteworthy that Romance loanwords may belong to at least three different chronological strata:

i. The earliest borrowings are substratal and consist of signifiers for concepts unknown to the newcomers, such as animal and plant names (e.g. /lúp/ 'wolf' < Latin *lupus*, /qalápaq/ 'tortoise' [cf. Castilian *galápago*], /istípa/ 'rockrose' < Low Latin *stippa*, /mullún/ 'melon' < Low Latin *melon[em]*), or those belonging to semantic fields where the Romance spoken by the mothers won the day because of their predominant role in nursing and rearing children, as is patent in the names of parts of the body (e.g. /imlíq/ 'navel' < *umbilicus*, /múčča/ 'breast' < Latin *mulcta*, /paččáyna/ 'eyebrow' [cf. Castilian *pestaña*]); the same may also be the case of some taboo words (e.g. /píšš/ [cf. Castilian *picha*] or /qaráyl/ 'penis' [cf. Portuguese *caralho*], /búlba/ 'vulva' < Latin *vulva*), where the Romance words might have served at the beginning as euphemisms in a society aesthetically dominated by Arabic-speaking fathers. In all such cases, the morphophonemic integration into Arabic patterns is perfect or nearly perfect, as is obvious, in particular, from the fact that the Castilian suffixes were recognized or metanalyzed (e.g. /qaláp+áq/, /pačč+áyna/, /qar+áyl/, etc.). However, in some instances, the Arabic equivalents often appear to have survived as well (e.g. /súrra/ 'navel', /ṯádd/ 'breast', /zúbb/ 'penis', /ḥírr/ 'vulva', undoubtedly current, in spite of their rudeness, as some of them would wind up in later stages of the Romance languages, e.g. Castilian *zorra*, *zupo*, and *herre que herre* 'stubbornly').

ii. The next oldest group of loanwords, in all likelihood adstratal, is integrated by words proceeding from Southern Romance as well, but only imperfectly integrated into Arabic, e.g. /apúryu/ 'stick [used to goad animals]' < Low Latin **aporrigium*, /rumíškal/ 'rorqual' < Hispanic Romance **lu musklu*, /marrúyu/ 'horehound' < Latin *marrubium*. Most of them appear to belong to the special lexica of certain crafts and rural milieus, where Romance technical terms were likely to have survived better and longer without strong Arabicizing pressures. It must be acknowledged, though, that their distinction from the previous group is not always easy.

iii. The youngest group of Romance loanwords in Andalusi Arabic is superstratal and results from its contact with the languages spoken by the Christian powers, Castilian, Portuguese, or Catalan, as a consequence of the Reconquista, i.e. the occupation of lands formerly held by the Muslims, in a reversed situation of bilingualism, where Arabic became the dominated language. In this case, their origins, semantics, and scarce adaptation to Arabic morphophonemics are generally quite obvious and characteristic, e.g. /nifindír/ 'I defend' < Castilian *defender*, /qalunjíyya/ 'canonry' < Old Castilian *calongia*.

3. THE SYNCHRONIC STUDY OF ROMANCE LOANWORDS IN ARABIC

The degree of morphophonemic integration of these borrowings, although legitimately construed as proof of their age, cannot always be considered of decisive value, without resorting to other gauging devices. It may happen, for instance, that a certified late loanword like /laṣqúna/, from Castilian *azcona* 'dart', on account of its casual fit in the template {1a23ú4(a)}, is assimilated at once and given a broken plural, /laṣáqin/, an assumed hallmark of total integration, which would suggest a much earlier date than the actual ones. A more accurate methodology should take into account as well whether the borrowed item proceeds from Low Latin, Early Romance, Southern or Andalusi Romance, or the modern Romance languages, and seek some philological information about the dates of its earliest appearance in texts of general or specialized contents. Of course, it will often happen that some of these details or indeed most of them are presently not available, which will necessarily lead to uncertain conclusions.

The semantic distribution of Romance loanwords is of paramount importance in assessing the overall contribution of the native Hispanic population to the culture, society, and institutions of al-Andalus, in the same manner that Arabic loanwords in the modern Romance languages of the Iberian Peninsula are witnesses to their debt to Arab and Islamic culture. Practical considerations make it advisable to adopt the following semantic classification.

Onomatopoetic borrowings have been transferred from one language to another, without regard to lexical necessity, simply because of their phatic value or synaesthetic expressiveness, such as /bába/ 'daddy', /páppa/ 'pap', /čawčál/ 'to whisper', /karkál/ 'to trample', / zázza/ 'slap on the neck' (cf. Castilian *papá*, *papa*, *chuchear*, and *zas* and Latin *calcare*). Some of them might be instances of baby talk, easily borrowed in situations of bilingualism.

Conceptual borrowings, designed to fulfil the speakers' lexical needs not satisfied by the target language, are related to physical objects, such as humans themselves, their inanimate setting, and their biological environment, or to social realities.

i. Physical objects. For the reasons previously advanced, loanwords related to humans are relatively abundant in the case of names of parts of the body (e.g. /búff/ 'lungs' [cf. Castilian *bofe*], /pulliqár/ 'thumb' < Latin *pullica[is]*, /furáṭ/ 'anus' < Latin *forat[us]*, /qubṭál/ 'elbow' [cf. Castilian *codillo*]), but rather scarce in the case of diseases and their cures (e.g. /rábyana/ 'scabies' < Latin *robigin[em]*, /russál/ 'sty' < Latin *hordeolus*), apparently because imported Eastern medicine was more developed than the native notions in this realm.

ii. Inanimate settings. Loanwords related to inanimate settings such as dwellings, sites, and places are also quite frequent (e.g. /pílč/ 'latch' < Latin *pestul[um]*, /parṭál/ 'porch' < Low Latin *portale*, /párčala/ 'garret' < Romance **bárčena*, /purṭál/ 'mountain pass' < Low Latin *portel[um]*, /šimṭáyr/ 'path' < Low Latin **semitarius*). It is also noteworthy that some meteorological phenomena have kept their Romance names (e.g. /číqa/ 'fog' < Latin *caeca*, /čírč/ 'north wind' < Latin *cercius*, /labáč/ 'south wind' < Latin *libyce*, /šulúq/ 'wind from the sea' < Greek *sálos* with a Romance suffix).

iii. Biological environment. The biological environment is the largest contributor to the Andalusi Arabic lexicon. This may have been because the Arabic names of animals and plants were often inadequate for the local species (e.g. /párṭal/ 'sparrow' < Latin *pardalus*, /kúrs/ 'roebuck' < Low Latin **curtiu*, /náġra/ 'crow' < Latin *nigra*, /qunílya/ 'rabbit' < Latin *cuniculus*, /paníč/ 'millet' < Latin *panicium*, /bíban/ 'willow' < Latin *vimen*, /aplantáyn/ 'plantain' < Latin *plantagin[em]*), or simply because the direct care of lands and beasts was left in the hands of the local population, so that such words had become a species of craft jargon by the time they were Arabicized.

iv. Social realities. When dealing with loanwords related to the social infrastructure of al-Andalus, it is obvious that legal language is almost entirely free of Romance borrowings (in spite of an exceptional /pársana/ 'accusation' < Low Latin *perdition[em]*), although there are a few among kinship terms (e.g. /šúqr/ 'father-in-law' < Latin *socer*, /núra/ 'daughter-in-law' < Latin *nurus*, /antináṭ/ 'stepson' < Low Latin *antenatus*). There is a much larger number of borrowings connected with folklore, games, and music (e.g. /fálya/ 'bonfire' < Latin *facula*, /maránda/ 'snack between lunch and dinner' < Latin *merenda*, /fáṭa/ 'fairy' < Latin *fata*, /dúrqa/ 'witch' < Low Latin **turica*, /pandáyr/ 'drum' < Latin *pandorium*, /ġayṭa/ 'bagpipe' [cf. Castilian *gaita*], /malandíyya/ 'melody' < Latin *melodia*), as well as with armies and weapons (e.g. /ṭárġa/ 'shield' [cf. Castilian *targa*], /ṭištány/ 'helmet' < Low Latin **testaneu*, /ġírra/ 'war' [cf. Castilian, Catalan, and Portuguese *guerra*], /pannún/ 'banner' < Latin *pinna* with a Romance suffix).

There is a host of borrowings related to household accessories and clothing (e.g. /pinnís/ 'jug' [cf. Castilian *pañés*], /iškála/ 'cup' < Low Latin *scala*, /qúbb/ 'bucket' < Latin *cupus*, /ṭábla/ 'table' < Latin *tabula*, /ṭirpáč/ 'candelabrum' < Low Latin **tripediu*, /ġaškún/ 'shirt' < Latin *vascon[em]*), in addition to the names of the months of the Christian calendar, used for agricultural purposes, and a considerable array of names of instruments of the diverse crafts and trades (e.g. /fúrka/ 'pitchfork' < Latin *furca*, /duntál/ 'plowshare bed; share-beam' < Latin *dentale*, /fawčíl/ 'sickle' < Latin *falc[em]*, /rúkka/ 'distaff' [cf. Portuguese *roca*], /šuqúr/ 'axe' < Latin *securis*, /barrína/ 'gimlet' < Latin *veruina*, /pála/ 'shovel' < Latin *pala*), agricultural and industrial products (e.g. /purčín/ 'certain fruits of low quality' < Latin *porcinus*, /fuqqún/ 'figs' < Low Latin *ficon[em]*, /iškirláṭ/ 'scarlet' < Latin *sigillatus*, /ušṭúp/ 'tow' Latin *stuppa*, /ṭápya/ 'adobe wall' [cf. Castilian *tapia*], /ṭúnna/ 'cask' < Low Latin *tunna*), and food specialties (e.g. /pulyáṭ/ 'porridge' < Latin *puleiatus*, /mirkás/ 'sausages' [cf. Portuguese *morcela*], /pišmáṭ/ 'biscuit' < Greek *paxamádion*).

Romance loanwords related to social, cultural, or religious aspects of the Andalusi people's life are practically nonexistent, as those areas were ruled by Islamic principles, and therefore their concepts were expressed in Arabic. The few exceptions are technical terms applied to typically Christian or Jewish institutions, such as /qúmṭ/ 'count' < Latin *comit[em]*, /šunúġa/ 'synagogue' < Latin *synagoga*, /paṭríq/ 'patriarch' < Latin *patricius*.

4. The panchronic study of Romance loanwords in Arabic

The frequency ratios of Romance loanwords in the texts and over the lexicon are extremely difficult to calculate because the number of Andalusi Arabic texts having reached us is exceedingly small, and even these are most of the time interspersed with Classical Arabic. As this last feature also impairs the quality of most of the extant medieval lexical inventories, we are not always in a position to determine whether a given item really belongs to Andalusi Arabic or is a mere classicism embedded in a low-register utterance. Besides, we are far from possessing a nearly complete lexical list of this dialect bundle. Corriente (1992:142) reports that the ratio of Romance loanwords in the *Vocabulista in arabico* was about 2.7 percent and that the 5.3 percent exhibited by Alcalá's lexicon is undoubtedly swollen by the author's habit of introducing unassimilated Castilian words whenever he does not know or decides to ignore their Arabic equivalents, if they existed at all. Taking into account those reservations, an educated guess is that, regarding the admission of foreign lexical stock, Andalusi Arabic stood in a position intermediate between the dialects of eastern Arabia, almost free from that kind of lexical interference, and the somewhat hybrid Maltese and Central Asian dialects.

As for the functional distribution of Romance loanwords in Andalusi Arabic, it is noteworthy that the loanwords are overwhelmingly substantives, followed at a considerable distance by verbs, most often denominative, and that adverbs (e.g. /makkár/ 'at least' [cf. Old Castilian *maguer*], /yá/ 'already' < Latin *iam*) and interjections (/áya/ 'come on!' < Latin *eia*) are oddities.

BIBLIOGRAPHICAL REFERENCES

Corominas, Joan. 1951. *Diccionario crítico etimológico de la lengua castellana*. Bern: Francke.

Corriente, Federico. 1978. "Dos nuevos romancismos del árabe hispánico, >*aḏālah*< y >*aḏāqal*< = /*yaḏḏ(a)l*/". *al-Andalus* 43.423–426.

——. 1980. "Notas de lexicología hispanoárabe. II. Los romancismos del *Vocabulista in arabico: Addenda et corrigenda*". *Vox Romanica* 39.194–210.

——. 1981. "Notas de lexicología hispanoárabe. III. Los romancismos del *Vocabulista*". *Awrāq* 4.5–27.

——. 1992. *Árabe andalusí y lenguas romances*. Madrid: MAPFRE.

——. 2000. "El romandalusí reflejado por el Glosario Botánico de Abulxayr". *Estudios de Dialectología Norteafricana y Andalusí* 5.93–241.

Ferrando, Ignacio. 1995. "Los romancismos de los documentos mozárabes de Toledo". *Anaquel de Estudios Árabes* 6.71–85.

Griffin, David. 1961. *Los mozarabismos del* Vocabulista *atribuido a Ramón Martí*. Madrid: Maestre.

Simonet, Francisco Javier. 1888. *Glosario de voces ibéricas y latinas usadas entre los mozárabes, precedido de un estudio sobre el dialecto hispano-mozárabe*. Madrid: Fortanet.

FEDERICO CORRIENTE
(University of Zaragoza)

Ibtidāʾ

Ibtidāʾ is the term used in the Arabic grammatical tradition for the commencement of an utterance by putting a noun in the initial position. The notion of *ibtidāʾ* is used in two main domains of the Arabic language sciences, grammar and *Qurʾān* reading. In grammar, it is used in governance (→ *ʿamal*) theory and in sentence grammar (*naḥw al-jumal*). It is an ambiguous notion that was much disputed in the Arabic grammatical tradition. While it is still used in school programs and in university departments of Arabic, the notion of *ibtidāʾ* plays hardly any role in modern linguistic approaches to Modern Standard Arabic.

1. *IBTIDĀʾ* IN GOVERNANCE

The *ibtidāʾ* indicates the governor (*ʿāmil*) of the first noun of the nominal sentence, which is called *mubtadaʾ*, like *zaydun* in *zaydun munṭaliqun* 'Zayd is going'. The nominal case of *zaydun* is assigned by a posited 'abstract governor', "characterized by the fact that it is phonetically void" (Bohas a.o. 1990:60, 69–70). In standard Arabic grammar, the *ibtidāʾ* operators both the *mubtadaʾ* and the attribute (*xabar*), in the same way as the governance of operators like *kāna* 'to be', *ḏanna* 'to believe', and *ʾinna* 'indeed' (emphatic particle) (→ *kāna wa-ʾaxawātuhā;* → *ʾinna wa-ʾaxawātuhā*).

It is possible for a sentence to contain more than one *ibtidāʾ*, as in *zaydun ʾabūhu munṭaliqun* 'Zayd's father is going', where both *zaydun* and *ʾabūhu* 'his father' are *mubtadaʾ*. In this case, we have two *ibtidāʾ*s for the two *mubtadaʾ*s. However, the governance of the *ibtidāʾ* can be canceled if a verbal or particle modifier (*nāsix;* → *nawāsix*) is introduced in the nominal sentence, e.g. *kāna ʿabdullāhi munṭaliqan* 'Abdallah was going', where *ʿabdullāhi* is the agent of *kāna* and no longer a *mubtadaʾ* (on the *nawāsix*, see Versteegh 1995:92–93, n. 3; Owens 1988:239–242; Peled 1992:148–150). The *mubtadaʾ* can be introduced by some particles (*ḥurūf al-ibtidāʾ*), which are, however, devoid of governance (Peled 1992:148–150). Historically, the term *ibtidāʾ* was used to denote the *mubtadaʾ* as well, for instance by Xalaf al-ʾAḥmar (*Muqaddima* 51).

It is generally admitted that the *ibtidāʾ* is an abstract governor (*ʿāmil maʿnawī*). This abstract idea (*maʿnā*) is not the result of elision (→ *ḥaḏf*), nor of suppression (→ *ʾidmār*), and it cannot be restored with a paraphrase (→ *taqdīr*). Its main characteristic is that it has neither equivalent nor substitute in the actual utterance; "it is impossible to pronounce it or to translate it" (*lā ḥaḏḏa li-l-lisāni fīhi* or *lā yutalaffaḏu bi-l-lisān*), as defined by Jurjānī in his commentary on the *ʿawāmil* (*ʿAwāmil* 312). The concept of such a type of governor is posited by the so-called Baṣra school in reply to the question 'What governs the *mubtadaʾ*?'. But the so-called Kūfan school rejected the idea of *ibtidāʾ* and opted for a bidirectional governance: "The topic and the comment govern[ed] each other in the nominative" (Owens 1988:52). This provides the point of departure for the analysis in terms of *ʿumda/faḍla* 'indispensable/optional elements [of the sentence]', which was developed by al-ʾAstarābāḏī (see Guillaume 1997:55). The majority of the Baṣrian grammarians believed that it was the abstract governor that governed the *mubtadaʾ*, but they disagreed about its definition (on these differences, see Ibn al-ʾAnbārī, *ʾInṣāf* I, 44–53; Baṭalyūsī, *Ḥulal* 144–149; ʿUkbarī, *Tabyīn* 224). Three tendencies may be distinguished in the definition of *ibtidāʾ*: the abstract-governance approach, the syntactic approach, and the functional approach, which integrates the 'pragmatic' dimension of the definition of *ibtidāʾ*.

The abstract-governance approach is based on the opposition between the morphological governor (*ʿāmil lafḏī*) and the abstract governor (*ʿāmil maʿnawī*). It is widely agreed that an abstract governor is a 'pure meaning' (*ʾinnamā huwa maʿnan*; Jurjānī, *Muqtaṣid* I, 214). The specificity of this 'meaning' is the lack (*taʿarrī* lit. 'nakedness') of an *ʿāmil lafḏī*, but this lack does not have the same status as an *ʿāmil lafḏī*: "The lack is not a word [and is not comparable to a morphological governor] like *ʾinna*" (*wa-laysa t-taʿarrī bi-lafḏin ka-ʾinna; Muqtaṣid* I, 214). Al-Jurjānī insists on the special status of the abstract governor and refuses to integrate the notion of → *ʾisnād* 'predication' in its definition, because the presence of predication is a condition on sentence construction but not on governance. According to him, the *ʾisnād* operation must be accomplished before the introduction of the abstract governor. The role of this governor in the grammatical system consists in effecting the nominative case on the *mubtadaʾ*. Governance is an abstract structure, in which syntactic and semantic considerations are irrelevant. Thus, the meaning of *ibtidāʾ* is neither semantic nor predicative.

The syntactic approach focuses on the syntactic characteristics of the *mubtadaʾ*. It stipulates that the reason why the *ibtidāʾ* makes the noun first in the sentence is in order to attribute a predicate to it (ʿUkbarī, *Tabyīn* 224). In this approach, the *ibtidāʾ* is not dissociated from the *ʾisnād*. This approach also emphasizes the fact that the achievement of the *ʾisnād* is a condition on the governance of the *ibtidāʾ*. Even if the *ibtidāʾ* is taken as an abstract governor, this abstraction is interpreted here as resulting from the absence of a morphological governor (*ʿāmil lafḏī*), such as the verbal operators *kāna* and *ḏanna*, or the particle modifiers *ʾinna* and its sisters (→ *ʾinna wa-ʾaxawātuhā*). The *ibtidāʾ* is no longer a 'pure meaning' with its own independent structure, as in the abstract-governance approach. The majority of the Arab grammarians adopt the syntactic approach, which uses the term *ibtidāʾ* but without its theoretical meaning. Az-Zajjājī (*Jumal* 36), for example, defines it as the resemblance to the agent, because for him the verbal sentence constitutes the basic structure of the Arabic language (cf. Baṭalyūsī, *Ḥulal* 144). Ibn Yaʿīš summarizes the theoretical criteria of this approach: "The *ibtidāʾ* means that you are focused on the noun and give it a first position [in the sentence] with respect to a second [noun], which is its attribute. Being in first position is an abstract idea that gives it strength, while the other noun [i.e. the predicate] depends on it" (*al-ibtidāʾu ihtimāmuka bi-l-ismi wa-jaʿluka ʾiyyāhu ʾawwalan li-maʿānin kāna xabaran ʿan-hu, wa-l-ʾawwaliyyatu maʿnan qāʾimun bi-hi yuksibuhu quwwatan ʾiḏ kāna ġayruhu mutaʿalliqan bi-hi*; *Šarḥ* I, 85). This means that the *ibtidāʾ* is replaced by a syntactic characteristic, 'being first, in initial position'. This notion is all but equivalent to the *mubtadaʾ*. We have moved from *ibtidāʾ* to *mubtadaʾ*.

As a historical consequence of this approach, some grammarians consider that it is useless to speculate about the identity of *ibtidāʾ*. This is the position that as-Suyuṭī (*Hamʿ* I, 308) attributes to ʾAbū Ḥayyān (*qāla ʾAbū Ḥayyān wa-hāḏa l-xilāfu lā yuġdī fāʾidatan*). Some grammarians practically ceased to use the term *ibtidāʾ* and dealt with the *mubtadaʾ* directly (e.g. Ibn Mālik, *ʿUmda* I, 256).

The functional approach may be exemplified by al-Mubarrid, who defines the *ibtidāʾ* first as awareness (*tanbīh*) and then as lack of governors. He clarifies it through an example: "The *ibtidāʾ* is when you say *zayd*, and when you express it, you do so for the listener, so that he expects the information you are bringing him concerning it [sc. *zayd*]; when you then say 'is going', or something like this, the meaning of the speech is fulfilled" (*al-ibtidāʾ naḥwu qawlika ʻzaydunʼ fa-ʾiḏa ḏakartahu fa-ʾinnamā taḏkuruhu li-s-sāmiʿi li-yatawaqqaʿa mā tuxbiruhu bi-hi ʿanhu, fa-ʾiḏā qulta munṭaliqun ʾaw mā ʾašbahahu ṣaḥḥa maʿnā l-kalām*; Mubarrid, *Muqtaḍab* IV, 126).

Al-Baṭalyūsī considers that "the best interpretation of *ibtidāʾ* is to say that the meaning that assigns to the *mubtadaʾ* the nominative case is the special interest uttered by the speaker" (*fa-ʾaḥsanu mā qīla fī ḥaqīqati r-rāfiʿi li-l-mubtadaʾi ʾanna l-maʿnā r-rāfiʿa la-hu ʿināyatu l-mutakallimi wa-htimāmuhu; Ḥulal* 147). What is stressed here is the pragmatic dimension of the notion of *ibtidāʾ*: the special attention of the speaker is the reason why the nominative case is assigned. In both texts, the speaker and the listener are part of the definition of this notion. In this conception, the grammarians do not see any conflict or heterogeneity between governance analysis and pragmatic dimension. The *ibtidāʾ* is indeed a governor, but it is more than a standard one, because it links the different aspects (levels) of the grammatical analysis. In order to understand this characteristic we have to go back to Sībawayhi, who puts forward such aspects, according to Carter (1968:219, 247–248).

2. *Ibtidāʾ* in Sībawayhi

Sībawayhi believes that in the nominal utterance the *ibtidāʾ* governs the topic (*mubtadaʾ*), which in its turn governs the attribute. In his *Kitāb*, in the chapter on *ibtidāʾ*, he defines the relation between these three notions as follow: "The topic is every noun you begin with in order to construct an utterance on it. The topic, and what is constructed on it, has the nominative case; the *ibtidāʾ* cannot exist without a construct on it" (*fa-l-mubtadaʾu kullu smin ubtudiʾa bi-hi li-yubnā ʿalayhi kalām, wa-l-mubtadaʾu wa-l-mabniyyu ʿalayhi rafʿun, fa-l-ibtidāʾu lā yakūnu ʾillā bi-mabniyyin ʿalayhi*; *Kitāb* II, 126). Thus, the comment depends on the *mubtadaʾ*. This definition suggests that the *mubtadaʾ* is the organizer of the grammatical relations in nominal utterances, just as the verb in verbal sentences organizes the grammatical relations upon which agent and complement depend (cf. Alaoui 1987:25). Sībawayhi's governance system is not based on lexical or morphological properties of the governors (as is the case in standard Arabic grammar), but on its capacity to organize the positions and assign declensional endings. He states that "just like the verb governs necessarily, the *ibtidāʾ* governs necessarily" (*fa-kamā lā tajidu buddan min ʾiʿmāli l-fiʿli kaḏālika lā tajidu buddan min ʾiʿmāli l-ibtidāʾ*; Sībawayhi, *Kitāb* I, 128). There is no 'conflict' between the two types of governance; each has its specific governance domain. The governance is a necessity for the *ibtidāʾ* even in the presence of a verb or interrogation (*Kitāb* I, 128): "You produce the interrogation when you have finished with the governance of *ibtidāʾ*" (*ʾinnamā tajīʾu bi-l-istifhām baʿdamā tafruġu min al-ibtidāʾ*). Moreover, Sībawayhi's use of *ibtidāʾ* has to be understood within a paradigm of notions like governance interruption (*qaṭʿ*), continuity of the governance relation (*ʾišrāk*), separation (*faṣl*), and connection (*waṣl*). The domain of this paradigm is mainly the relation between sentences. The conditionals are the best example for the use of this paradigm. In a case like *ʾin taʾtinī ʾātīka wa-ʾuḥaddiṯuka* 'if you come to me, I will come to you and I will tell you', the third verb *ʾuḥaddiṯu* is not in the apocopate because the governance of the particle *ʾin* 'if' is interrupted, and a new utterance with a new governance begins. In this new utterance, the indicative (*rafʿ*) is used because of the governance of the *ibtidāʾ* (*yurfaʿu ʿalā qaṭʿ wa-btidāʾ*; Sībawayhi, *Kitāb* I, 49). Here, the *ibtidāʾ* governs the entire utterance and not just a word (Iraqi 1992:58–60). For Sībawayhi, it organizes the assignment of the declensional endings and regulates the relation between the utterances. The *ibtidāʾ* governs the entire utterance and is a universal governor, which is necessary for every governance structure.

3. *Ibtidāʾ* in sentence *(jumla)* theory

Naḥw al-jumal is a grammar of text. It is an extension of 'word grammar', whose main focus is on the use of declensional endings. *Naḥw al-jumal* deals with the grammatical status of these endings and their function within the text. For instance, a sentence is called *ibtidāʾiyya* in two cases: at the beginning of a speech or text, and when a sentence is independent and 'cut off from a previous sentence' (*munqaṭiʿa ʿammā qablahā*). Ibn Hišām (*Muġnī* 500) calls it 'recommencement' (*istiʾnāf*; for the use of *istaʾnafa* in early exegesis, see Versteegh 1993:132–136, 196ff.). The *ibtidāʾiyya* sentence belongs to the category of nondeclined sentences, to which no case is assigned (*lā maḥalla lahā min al-ʾiʿrāb*).

4. *Ibtidāʾ* in *Qurʾān* reading

The meaning of *ibtidāʾ* in Qurʾānic reading is 'starting; starting again'. It is related to the notion of pause (*waqf*). The main goal is to show the reader of the *Qurʾān* when it is possible to make a pause and how to start again. For Ibn al-Jazarī (*Našr* I, 230), the *ibtidāʾ* is optional, being conditioned by autonomy of meaning and noncontradiction with text 'intention'. This use of *ibtidāʾ* is not associated with grammatical categories, such as the noun or the sentence. Generally, *ibtidāʾ* rules depend on pausal (*waqf*) rules (*Našr* I, 231).

Bibliographical references

Primary sources

ʾAstarābāḏī, *Šarḥ al-Kāfiya* = Raḍī d-Dīn Muḥammad ibn al-Ḥasan al-ʾAstarābāḏī, *Šarḥ Kāfiyat Ibn al-Ḥājib*. Ed. Yūsuf Ḥasan ʿUmar. 4 vols. Beirut: Dār al-Kutib al-ʿIlmiyya, 1969.

Baṭalyūsī, *Ḥulal* = ʾAbū Muḥammad ʿAbdallāh ibn Muḥammad ibn as-Sayyid al-Baṭalyūsī, *Kitāb al-ḥulal fī ʾiṣlāḥ al-xalal min kitāb al-Jumal*. Ed. Saʿīd ʿAbd al-Karīm Saʿʿūdī. Beirut: Dār aṭ-Ṭalīʿa, n.d.

Ibn al-ʾAnbārī, *ʾInṣāf* = ʾAbū l-Barakāt ʿAbd ar-Raḥmān ibn Muḥammad al-ʾAnbārī, *Kitāb al-ʾinṣāf fī masāʿil al-xilāf bayna n-naḥwiyyīna l-baṣriyyīna wa-l-kūfiyyīn*. Beirut: al-Maktaba al-ʿAṣriyya, 1987.

Ibn al-Jazarī, *Našr* = al-Ḥāfiẓ ʾAbū l-Xayr Muḥammad ibn Muḥammad ad-Dimašqī, *an-Našr fī l-qirāʾāt al-ʿašr*. Ed. ʿAlī Muḥammad aḍ-Ḍabbāʾ. Beirut: Dār al-Fikr, n.d.

Ibn Mālik, *ʿUmda* = Jamāl ad-Dīn Muḥammad Ibn Mālik, *Šarḥ ʿUmdat al-ḥāfiḏ wa-ʿuddat al-lāfiḏ*. 2 vols. Ed. ʿAdnān ʿAbd ar-Raḥmān ad-Dūrī. Baghdad, 1975.

Ibn Yaʿīš, *Šarḥ* = Muwaffaq ad-Dīn Yaʿīš ibn ʿAlī Ibn Yaʿīš, *Šarḥ al-Mufaṣṣal*. 2 vols. Beirut: ʿĀlam al-Kutub, n.d.

Jurjānī, *ʿAwāmil* = ʾAbū Bakr ʿAbd al-Qāhir ibn ʿAbd ar-Raḥmān al-Jurjānī, *al-ʿAwāmil al-miʾa an-naḥwiyya fī ʾuṣūl ʿilm al-ʿarabiyya*, with commentary by aš-Šayx Xālid al-ʾAzharī al-Jarjāwī. Ed. Badrāwī Zahrān. Cairo: Dār al-Maʿārif, 1988.

——, *Muqtaṣid* = ʾAbū Bakr ʿAbd al-Qāhir ibn ʿAbd ar-Raḥmān al-Jurjānī, *Kitāb al-muqtaṣid fī šarḥ al-ʾĪḍāḥ*. Ed. Kāẓim Baḥr al-Marjān. 2 vols. Baghdad: Wizārat aṯ-Ṯaqāfa wa-l-ʾIʿlām, 1982.

Mubarrid, *Muqtaḍab* = ʾAbū l-ʿAbbās Muḥammad ibn Yazīd al-Mubarrid, *al-Muqtaḍab*. Ed. Muḥammad ʿAbd al-Xāliq ʿUḍayma. Beirut: ʿĀlam al-Kutub, 1963.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. 5 vols. Ed. ʿAbd as-Salām Hārūn. Cairo: al-Hayʾa al-Miṣriyya al-ʿĀmma li-l-Kitāb, 1966–1977.

Suyūṭī, *Hamʿ* = Jalāl ad-Dīn ʿAbd ar-Raḥmān ibn ʾAbī Bakr as-Suyūṭī, *Hamʿ al-hawāmiʿ fī šarḥ Jamʿ al-jawāmiʿ*. Ed. ʾAḥmad Šams ad-Dīn. 4 vols. Beirut: Dār al-Kutub al-ʿIlmiyya, 1997.

ʿUkbarī, *Tabyīn* = ʾAbū l-Baqāʾ ʿAbdallāh ibn al-Ḥusayn al-ʿUkbarī, *at-Tabyīn ʿan maḏāhib an-naḥwiyyīna l-baṣriyyīna wa-l-kūfiyyīn*. Ed. ʿAbd ar-Raḥmān ibn Sulaymān al-ʿUṯaymīn. Beirut: Dār al-Ġarb al-ʾIslāmī, 1986.

Xalaf al-ʾAḥmar, *Muqaddima* = ʾAbū Muhriz Xalaf ibn Ḥayyān al-ʾAḥmar al-Baṣrī, *Muqaddima fī n-naḥw*. Ed. ʿIzz ad-Dīn at-Tanūxī. Damascus: Wizārat aṯ-Ṯaqāfa wa-l-ʾIršād al-Qawmī, 1961.

Zajjājī, *Jumal* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān ibn ʾIsḥāq az-Zajjājī, *al-Jumal fī n-naḥw*. Ed. ʿAlī Tawfīq al-Ḥamad. 2nd ed. Beirut: Muʾassasat ar-Risāla, 1985.

Secondary sources

Alaoui, Ahmed, 1987. "ʾĀyat al-fikr wa-kibriyyāʾ an-naḏar". *al-Mawqif* 1.18–44.

Bohas, Georges, Jean-Patrick Guillaume, and Djamal Eddine Kouloughli. 1990. *The Arabic linguistic tradition*. London and New York: Routledge.

Carter, Michael G. 1968. *A study of Sībawayhi's principles of grammatical analysis*. Ph.D. diss., University of Oxford.

——. 1991. "Elision". *Proceedings of the Colloquium on Arabic Grammar, Budapest, 1–7 September 1991*, ed. Kinga Dévényi and Tamás Iványi, 121–133. Budapest: Eötvös Loránd University and Csoma de Kőrös Society.

Guillaume, Jean-Patrick. 1997. "Les discussions des grammairiens arabes à propos du sens des marques d'*iʿrāb*". *Actes du Colloque sur la flexion casuelle: Concepts et méthodologie. 1–3 December 1994*, ed. Choukri Iraqi Houssaini, 43–57. Fes: Faculté des Lettres et des Sciences Humaines.

Iraqi Houssaini, Choukri. 1992. "Mafhūm al-ibtidāʾ ʿinda Sībawayhi". *Recherches sur le patrimoine linguistique arabe*, ed. Djamal Eddine Kouloughli, 53–62. *Linguistica Communicatio* 4:2.

Owens, Jonathan. 1988. *The foundations of grammar: An introduction to Medieval Arabic grammatical theory*. Amsterdam and Philadelphia: J. Benjamins.

Peled, Yishai. 1992. "*ʿAmal* and *ʾibtidāʾ* in Medieval Arabic grammatical tradition". *Abr-Nahrain* 30.147–171.

Versteegh, Kees. 1993. *Arabic grammar and Qurʾānic exegesis in early Islam*. Leiden: E.J. Brill.

——. 1995. *The explanation of linguistic causes: Az-Zaǧǧāǧī's theory of grammar*. Amsterdam and Philadelphia: J. Benjamins.

Choukri Iraqi Houssaini
(University of Fes)

ʾIḍāfa

1. The *ʾiḍāfa* in the Arabic linguistic tradition

Two Arabic nouns may be linked together in a noun phrase in such a way that the second noun in the sequence determines the first by limiting, identifying, possessing, defining, or amplifying it. The two nouns in this phrase function as a closely knit syntactic unit. In Arabic grammatical terminology, this structure is referred to as *ʾiḍāfa* 'annexation; addition'; the first noun in the structure is *muḍāf* 'annexed' to the second noun, which is the *muḍāf ʾilay-hi* lit. 'the added-to (or 'annexing') noun'. The annexing noun is in the genitive case. In traditional English descriptions of Arabic grammar, this unit is normally termed 'construct state', 'genitive construct', 'construct phrase', or 'annexation structure'.

The grammarians also acknowledged the existence of what they called an 'improper annexion' (*ʾiḍāfa ġayr maḥḍa*; Ibn as-Sarrāj, *ʾUṣūl* II, 6–10); this included the construction of a participle with a following genitive (*hāḏā ḍāribu zaydin ġadan* 'this is the one who will hit Zayd tomorrow'); an adjective with a following genitive (*marartu bi-rajulin ḥasani l-wajhi* 'I passed a man with a beautiful face'); an elative with a following genitive (*zaydun ʾafḍalu l-qawmi* 'Zayd is the best of the people'); and the slightly controversial construction of a noun with a following adjective in the genitive (e.g. *ṣalāt al-ʾūlā* 'the first prayer', *masjid al-jāmiʿ* 'the Friday mosque').

The genitive construct is a central component of Arabic syntax, and the term *ʾiḍāfa* occurs frequently in grammatical treatises. In the *Maʿānī l-Qurʾān* by the Kūfan grammarian al-Farrāʾ (d. 207/822), the verb *ʾaḍāfa ʾilā* is used both for the annexion of a noun to another noun (for which al-Farrāʾ also uses the term → *ʾisnād*) and for the construction of a preposition with a noun (Kinberg 1996:436–445).

In the *Kitāb Sībawayhi, ʾiḍāfa* and cognate terms are very frequent (Mosel 1975:205–207): the verb *ʾaḍāfa* occurs 233 times, *ʾiḍāfa* 243 times, and *muḍāf/muḍāf ʾilayhi* 183 times (Troupeau 1976:132). They do not always indicate the same phenomenon, however. Troupeau distinguishes between the translations 'annexer à' and 'relier (un individu) à' for the verb *ʾaḍāfa*, probably in order to differentiate between noun/noun constructions and preposition/noun constructions, but it is not quite clear which criteria he uses in assigning the loci to either sense. Talmon (2003:236–238) points out that *ʾaḍāfa* may be used generally for any preposition linking a verb with a noun (e.g. *Kitāb* I, 177.11 *yuḍāfu bihā ʾilā l-ism mā qablahu ʾaw mā baʿdahu*). Hence, all prepositions may be called *ḥurūf al-ʾiḍāfa* (e.g. *Kitāb* II, 146.11; Owens 1990:14–17). Elsewhere, Sībawayhi restricts this use to the preposition *li-*, which is called *lām al-ʾiḍāfa* (*Kitāb* II, 331.2).

A special case is the use of *ism ʾiḍāfa* for the relative adjective (→ *nisba*), referring to what is currently called the *yāʾ* of *nisba* as *yāʾ al-ʾiḍāfa* (cf. Wright 1967:II, 225). According to Talmon (2003:132), this is a relic of an earlier use of *ʾiḍāfa*, which he connects with Syriac grammar. The same term, *yāʾ al-ʾiḍāfa*, is used in the *Maʿānī l-Qurʾān* by al-ʾAxfaš al-ʾAwsaṭ (d. 215/830?) for a different meaning, namely the suffix *-ī* of the 1st person (*Maʿānī* II, 375.5).

In the case of an annexion between a preposition and a noun, no explanation was needed of the genitive case in the noun, because particles have the right to govern. But in the case of the syntactic relationship between two nouns in an *ʾiḍāfa*, the question of the governor of the genitive in the second noun was more difficult to answer, at least from the perspective of later grammarians. For Sībawayhi (*Kitāb* I, 177) this was apparently not a problem, since he regards the first noun in the *ʾiḍāfa* as the *ʿāmil* of the genitive case in the second noun (cf. Mosel 1975:86–87):

> The genitive occurs in any noun that is annexed; you must know that the annexed noun is put in the genitive by one of three items: a word that is neither a noun nor an adverbial adjunct, a word that is an adverbial adjunct, and a word that is a noun and not an adverbial adjunct (*wa-l-jarr ʾinnamā yakūnu fī kull ism muḍāf ʾilayhi wa-ʿlam ʾanna l-muḍāf ʾilayhi yanjarru bi-ṯalāṯa ʾašyāʾ: bi-šayʾ laysa bi-sm wa-lā ẓarf wa-bi-šayʾ yakūnu ẓarfan wa-bi-sm lā yakūnu ẓarfan*)

The three instances mentioned in this definition may be exemplified by the following expressions: *marartu bi-zayd-in* 'I passed Zayd', *taḥta zayd-in* 'under Zayd', and *ḥimār-u zayd-in* 'the donkey of Zayd', respectively. In the third example, the noun is clearly recognized by Sībawayhi as governor. Note that in all three instances, it is the second noun that is called *muḍāf ʾilayhi* 'the [noun] to which [the first noun] is annexed'.

Al-Mubarrad (d. 285/898), too, regarded the first noun as the governor (*Muqtaḍab* IV, 136; Owens 1988:153–154, 242–243, 1990:15), but later grammarians rejected this government by a noun. They believed that the only elements in language capable of governing are the verbs and the particles, which meant that in the case of the *ʾiḍāfa* another element than the first noun had to be made responsible for the genitive ending in the second noun. Just as Sībawayhi did in the *Kitāb*, al-Mubarrad (*Muqtaḍab* IV, 143) paraphrases the relationship between the two nouns by a construction with the preposition *li-* indicating possession: "Nouns that are annexed to other nouns do so with the meaning of *li-*" (*wa-ʾammā l-ʾasmāʾ al-muḍāfa ʾilā l-ʾasmāʾ bi-ʾanfusihā fa-tadxulu ʿalā maʿnā l-lām*). This reflects the old connection between prepositional phrases and noun/noun phrases. But the semantic equivalence does not imply that a preposition has to be posited in the underlying structure as the implicit governor.

Some grammarians were content to limit their analysis to the observation that in possessive constructions the second noun has a genitive; this is the case, for instance, in the *Kitāb al-jumal fī n-naḥw*, a treatise that has sometimes been attributed to al-Xalīl ibn ʾAḥmad (Ryding 1998:109; about the authorship of the treatise see also Owens 1990:180). For the genitive of annexion, the author simply observes (*Jumal* 173.9–10) that in an expression like *dār-u zayd-in*, the second noun receives a genitive ending because of the annexion of the first noun to it (*qawluhum dāru zaydin wa-ġulāmu ʿamrin xafaḍta zaydan bi-ʾiḍafa dārin ʾilayhi*).

Later grammarians, however, were concerned with determining the government relations in the possessive construction. Ibn as-Sarrāj (*ʾUṣūl* II, 5–6) distinguishes within the proper annexion (*ʾiḍāfa maḥḍa*) between two possibilities, depending on whether the meaning of the construction was that of *li-* or *min*. Clearly, he believed that the semantics of the construction could be expressed by prepositional meanings. But the expression *bi-maʿnā* 'with the meaning of' does not imply that there is an underlying preposition acting as implicit governor. According to Owens (1990:16), it was not until al-Jurjānī (d. 471/1078) that the first explicit statement appeared to the effect that the noun cannot be the governor and that the underlying preposition is the governor of the construction, the noun having "the meaning of the genitive particle" (*maʿnā ḥarf al-jarr; Muqtaṣid* II, 871).

Apart from the government relations, grammarians also became interested in the various types of annexion. Az-Zajjājī (d. 339/949) devotes an entire chapter to the question of why verbs cannot enter in an annexing construction, and in this connection he classifies the *ʾiḍāfa* into three types (*ʾĪḍāḥ* 108; Versteegh 1995:190–215): annexion of a possession to its possessor, e.g. *hāḏā dāru zaydin* 'this is Zayd's house'; annexion of something to someone who is entitled to it or connected with it, e.g. *al-ḥamdu lillāhi* 'praise be to God', *marartu bi-zaydin* 'I passed Zayd'; and annexion of something to its genus, e.g. *bābu sājin* 'a door made of teak'. The inclusion of the second type in az-Zajjājī's analysis demonstrates that the original connection with prepositional phrases as in the *Kitāb* persevered in the Arabic tradition.

Later grammarians elaborated on this by translating the various semantic functions of the genitive construct into different prepositional governors; along with *li-* for possession, prepositions like *min* were also adduced. In his *Irtišāf* (II, 501–502), the Andalusian grammarian ʾAbū Ḥayyān (d. 745/1344) quotes az-Zajjāj as being the first to assign to the noun in the *ʾiḍāfa* the meaning of *li-*. ʾAbū Ḥayyān then lists the various types of *ʾiḍāfa* proposed by other grammarians, not only with an underlying *min* or *li-*, but also with other prepositions, such as *fī*, *ʿalā*, and even *ʿinda*. In his own view, all these shades of meaning can be subsumed under one heading, that of *ixtiṣāṣ* 'specification', which he believes is the core meaning of the *ʾiḍāfa*.

2. The *ʾiḍāfa* in modern grammars of Arabic

An extensive examination of *ʾiḍāfa* constructions is contained in Cantarino (1974–1975:II, 92–119). See also Wright (1967:II, 198–234) on the rules for Classical Arabic. Fleisch (1971) is concise and informative. For discussion in

Arabic of the genitive construct, Ḥasan (1987: III, 1–180) provides an in-depth analysis. Just as the Arabic grammarians did, modern grammarians attempt to set up a semantic framework for the different shades of meaning in the *ʾiḍāfa* construction. However, the term is no longer used for constructions with a preposition as in the Arabic grammatical tradition.

The first noun in the genitive construct, the *muḍāf*, is distinguished by the fact that it carries neither the definite article nor nunation (the word-final inflectional marker of indefiniteness) because it is determined by means of the second noun. “In Arabic it is the amplifying term whose definitional status yields the definitional status of the whole phrase: consequently, an annexed substantive will not itself have the article” (Beeston 1970:46). However, as the head noun of the phrase, the first noun in the genitive construct may be in any case: nominative, genitive, or accusative, depending on the function of the *ʾiḍāfa* unit in a sentence structure. The first term of a construct phrase cannot have a possessive pronoun suffix.

The first term of the construct carries a case marker (overt or implied) determined by the syntactic role of the phrase in the sentence or clause, e.g. *ḥaḍar-a zuʿamāʾ-u l-qabāʾil-i* ‘the leaders of the tribes came’; *nu-qaddim-u mūjaz-a l-ʾaxbār-i* ‘we present the news summary’; *fī šimāl-i l-ʿirāq-i* ‘in the north of Iraq’.

The restriction on nunation of the first term of the *ʾiḍāfa* applies also to the final inflectional *nūn*s of the dual (*-āni/-ayni*) and the sound masculine plural (*-ūna/-īna*) suffixes, e.g. *wazīr-ā l-xārijiyyat-i* ‘the two foreign ministers’; *li-wazīr-ay-i l-xārijiyyat-i* ‘for the two foreign ministers’; *muraššaḥ-ū l-ḥizb-i* ‘the party’s nominees’; *min muraššaḥ-ī l-ḥizb-i* ‘from the party’s nominees’.

Another traditional restriction on the first term of the *ʾiḍāfa* is that it may not be conjoined; if more than one noun is to be included in the first element of the phrase, then the surplus nouns follow the *ʾiḍāfa* and refer back to it by means of a resumptive pronoun suffix, e.g. *bi-n-nisbat-i ʾilā ʾasātiḏat-i l-luġat-i wa-ṭullāb-i-hā* ‘in relation to professors and students of the language’. In the grammatical tradition, the insertion of a surplus noun was called *ʾiqḥām* ‘intrusion, invasion’ (Ryding 1992:272), and it was generally regarded as incorrect. This rule is still active in Modern Standard Arabic, but in current usage it is sometimes broken, in Media Arabic especially. This conjoining of the first term is referred to by Badawi a.o. (2004:138) as ‘binomial/polynomial annexation’, and by Ryding (2005:217) as ‘joint annexation’ (after Beeston 1970:48), e.g. *qaddam-ū šukr-a wa-taqdīr-a š-šaʿb-i* ‘they offered the thanks and appreciation of the people’; *fī šimāl-i ʾaw ġarb-i l-minṭaqat-i* ‘in the east or west of the region’.

A construct phrase may not be interrupted by modifiers for the first term. Adjectives or other modifiers applying to the first term of the *ʾiḍāfa* follow the last term. Modifiers for the first term agree with it in gender, number, case, and definiteness, e.g. *ʾarkān-u l-ʾislām-i l-xamsat-u* ‘the five pillars of Islam’; *jawāzāt-u s-safar-i l-jadīdat-u* ‘the new passports’; *fī kutub-i t-tārīx-i hāḏihi* ‘in these history books’.

The second or determining noun in the *ʾiḍāfa* is in the genitive case and is marked either for definiteness or indefiniteness, thereby determining the definiteness or indefiniteness of the entire phrase. It may carry a suffixed pronoun; it may also be a demonstrative pronoun, e.g. *wādī n-nīl-i* ‘the Valley of the Nile’; *ṭabīb-u ʾasnān-in* ‘a dentist’; *ḥall-u muškilāt-i-nā* ‘the solution of our problems’; *maʿnā hāḏā* ‘the meaning of this’. Some grammarians also consider a noun followed by a possessive pronoun suffix to be in the construct state (see, for instance, Fischer 2002:89).

The second term of the construct may be modified by adjectives directly following it and agreeing with it in definiteness, gender, number, and case, e.g. *fī minṭaqat-i š-šarq-i l-ʾawsaṭ-i* ‘in the region of the Middle East’; *ḥaqq-u l-lujūʾ-i s-siyāsiyy-i* ‘the right to political asylum’. It may be preceded directly by a demonstrative pronoun plus definite article, e.g. *qīmat-u hāḏihi l-maxṭūṭāt-i* ‘the value of these manuscripts’.

More than two nouns may occur in a string of construct relationships; this is called a ‘multi-noun construct’ or ‘extended annexation’ by Badawi a.o. (2004:133), and a ‘complex or multi-noun construct’ by Ryding (2005:215). In this multi-noun construct, the first term (as the head noun) carries the case marker determined by the role of the phrase within a sentence; all subsequent nouns are in the genitive case. For non-final nouns, the restrictions on nunation and the definite article apply. Only the final noun in the string may carry the markers of definiteness: either the definite article or nunation, e.g. *raʾīs-u taḥrīr-i l-jarīdat-i* 'the editor in chief of the newspaper'; *jamīʿ-u qarārāt-i majlis-i l-ʾamn-i* 'all the resolutions of the security council'.

An adjective or participle may appear as the first term of a construct phrase instead of following the noun as a modifier. In these phrases the adjective remains in the masculine gender, but it may be singular or plural. These expressions are often set phrases, e.g. *fī muxtalif-i l-manāṭiq-i* 'in various regions'; *fī qadīm-i z-zamān-i* 'in olden times'.

In the adjective *ʾiḍāfa* (*ʾiḍāfa ġayr ḥaqīqiyya*), the first term is an adjective or participle and may carry the definite article if it modifies a definite noun, hence its label of *ġayr ḥaqīqiyya* 'unreal, false *ʾiḍāfa*', e.g. *fī l-ʿalāqāt-i l-mutaʿaddidat-i l-ʾaṭrāf-i* 'in multilateral relations'; *al-ʾarḍ-u mustadīrat-u š-šakl-i* 'the earth is circular in shape'.

The noun *ġayr*, which denotes a meaning of 'non-', 'dis-', or 'un-' may form the first term of an *ʾiḍāfa* whose second term is an adjective, e.g. *ġayr-u rasmiyyin* 'unofficial'; *ġayr-u munāsib-in* 'unsuitable'; *ġayr-u ʾislāmiyy-in* 'non-Islamic'; *ġayr-u marġūb-in fīhi* 'undesirable'.

The meanings assigned to an *ʾiḍāfa* in modern grammars of Arabic are wide-ranging, and occasionally the categories are hard to delimit; Beeston (1970:46) calls this the "semantic polyvalency of the annexation structure". Some general categories are listed here; for a more detailed list see Ryding (2005:205–211).

i. Identity relation. In this broad category, the second term specifies, defines, limits, or explains the purpose of the first, e.g. *madīnat-u l-xarṭūm-i* 'the city of Khartoum'; *ṭullāb-u t-tārīx-i* 'history students'; *ḥalīb-u l-baqar-i* 'cow's milk'; *ṭāʾirat-u ʾinqāḏ-in* 'a rescue plane'.
ii. Possessive relation. In this kind of annexation structure, the first term can be interpreted as belonging (in the very broadest sense) to the second term. In certain respects, it is very close to the next category, the partitive relationship, and it is sometimes difficult to draw a line between the two, e.g. *maktabat-u l-jāmiʿa* 'the university library'; *bayt-u ʾabī* 'my father's house'.
iii. Partitive relationship. Here the annexed term (the first term) serves as a determiner for a part or quantity of the annexing term. This includes the use of quantifier nouns, certain numbers and fractions, and the superlative construction, e.g. *kull-u l-qawāmīs-i* 'all of the dictionaries'; *muʿḍam-u l-ʾaṣwāt-i* 'most of the votes'; *xams-u mudun-in* 'five cities'; *miʾat-u marrat-in* 'a hundred times'; *nafs-u š-šayʾ-i* 'the same thing'.
iv. Constructs with deverbal nouns. In this type of construct, the first term is a verbal noun (→ *maṣdar*), the name of an action, and the second term is the agent or doer of the action, e.g. *muġādarat-u s-safīr-i* 'the departure of the ambassador'. When the object of the verbal action is mentioned in addition to the doer of the action, the object may follow the *ʾiḍāfa* construction and is in the accusative case (as object of the underlying transitive verb), e.g. *muġādarat-u s-safīr-i l-ʿāṣimat-a* 'the ambassador's leaving the capital'. In an object relationship, the second term of the construct is the object of an action, and the first term is either the name of the action (*maṣdar*), or an active participle (*ism fāʿil*) referring to the doer of the action, e.g. *rukūb-u l-xayl-i* 'riding horses'; *rafʿ-u l-ʿālam-i* 'the raising of the flag'; *musāʿid-u wazīr-in* 'an assistant minister'.
v. Measurement, composition, contents. In these structures, the second noun of the construct expresses measurement, contents, or nature of the first, e.g. *muddat-a sanatayni* '[for] a period of two years'; *finjān-u qahwat-in* 'a cup of coffee'; *bāqāt-u zuhūr-in* 'bouquets of flowers'; *ʿunqūd-u ʿinab-in* 'a bunch of grapes'.
vi. Title or quote. When a title or quote is the second term of an *ʾiḍāfa*, it is considered isolated from the case-marking requirements of the second term and it is inflected independently, not necessarily in the genitive, e.g. *kitāb-u ʾalf-u laylat-in wa-laylat-un* 'the book "The thousand and one nights"'.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

ʾAbū Ḥayyān, *Irtišāf* = ʾAbū Ḥayyān Muḥammad ibn Yūsuf al-Ġarnāṭī, *Irtišāf aḍ-ḍarab min lisān al-ʿArab*. Ed. Muṣṭafā ʾAḥmad an-Nammās. 3 vols. Cairo: Maṭbaʿat al-Madanī, 1984–1989.

ʾAxfaš, *Maʿānī* = ʾAbū l-Ḥasan Saʿīd ibn Masʿada al-ʾAxfaš al-ʾAwsaṭ, *Maʿānī l-Qurʾān*. Ed. Fāʾiz Fāris. 2 vols. Kuwait, 1979. (2nd. ed., Beirut: Muʾassasat ad-Dirāsa, 1974.)

Ibn as-Sarrāj, *ʾUṣūl* = ʾAbū Bakr ibn as-Sarī Ibn as-Sarrāj, *Kitāb al-ʾuṣūl fī n-naḥw*. Ed. ʿAbd al-Ḥusayn al-Fatlī. 3 vols. Beirut: Muʾassasat ar-Risāla, 1985.

Jurjānī, *Muqtaṣid* = ʾAbū Bakr ʿAbd al-Qāhir ibn ʿAbd ar-Raḥmān al-Jurjānī, *Kitāb al-muqtaṣid fī šarḥ al-ʾĪḍāḥ*. Ed. Kāẓim Baḥr al-Marjān. 2 vols. Baghdad: Wizārat aṯ-Ṯaqāfa wa-l-ʾIʿlām, 1982.

Mubarrad, *Muqtaḍab* = ʾAbū l-ʿAbbās Muḥammad ibn Yazīd al-Mubarrad, *al-Muqtaḍab*. Ed. Muḥammad ʿAbd al-Xāliq ʿUḍayma. 4 vols. Cairo: Dār at-Taḥrīr, 1965–1968.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. Ed. Hartwig Derenbourg. 2 vols. Paris: Imprimerie Nationale, 1970.

[Xalīl], *Jumal* = *Kitāb al-jumal fī n-naḥw*. Ed. Faxr ad-Dīn Qabāwa. Beirut: Muʾassasat ar-Risāla, 1987. [Also published as *al-Muḥallā: Wujūh an-naṣb*, ascribed to Ibn Šuqayr, ed. Fāʾiz Fāris, Beirut: Muʾassasat ar-Risāla, 1987.]

Zajjājī, *ʾĪḍāḥ* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān Ibn ʾIsḥāq az-Zajjājī, *al-ʾĪḍāḥ fī ʿilal an-naḥw*. Ed. Māzin al-Mubārak. Cairo: Dār al-ʿUrūba, 1959.

SECONDARY SOURCES

Badawi, El-Said, Michael G. Carter, and Adrian Gully. 2004. *Modern Written Arabic: A comprehensive grammar*. London: Routledge.

Beeston, A.F.L. 1970. *The Arabic language today*. London: Hutchinson University Library.

Cantarino, Vicente. 1974, 1975, 1976. *The syntax of Modern Arabic prose*. 3 vols. Bloomington: Indiana University Press.

Fischer, Wolfdietrich. 2002. *A grammar of Classical Arabic*. 3rd rev. ed. Transl. by Jonathan Rogers. New Haven and London: Yale University Press.

Fleisch, Henri. 1971. "*Iḍāfa*". *Encyclopaedia of Islam*, III, 1008. 2nd ed. Leiden: E.J. Brill.

Ḥasan, ʿAbbās 1987. *an-Naḥw al-wāfī*. 4 vols. Cairo: Dār al-Maʿārif.

Kinberg, Naphtali. 1996. *A lexicon of al-Farrāʾ's terminology in his* Qurʾān *commentary, with full definitions, English summaries, and extensive citations*. Leiden: E.J. Brill.

Mosel, Ulrike. 1975. *Die syntaktische Terminologie bei Sībawaih*. Ph.D. diss., University of Munich.

Owens, Jonathan. 1988. *The foundations of grammar: An introduction to medieval Arabic grammatical theory*. Amsterdam and Philadelphia: J. Benjamins.

——. 1990. *Early Arabic grammatical theory: Heterogeneity and standardization*. Amsterdam and Philadelphia: J. Benjamins.

Ryding, Karin C. 1992. "Morphosyntactic analysis in *al-Jumal fii l-naḥw*: Discourse structure and metalanguage". *Perspectives on Arabic linguistics*, IV, ed. Ellen Broselow, Mushira Eid, and John McCarthy, 263–277. Amsterdam and Philadelphia: J. Benjamins.

——. 1998. "Aspects of the genitive: Taxonomy in *al-Jumal fī al-naḥw*". *Early Medieval Arabic: Studies on al-Khalīl ibn Aḥmad*, ed. Karin C. Ryding, 92–142. Washington, D.C.: Georgetown University Press.

——. 2005. *A reference grammar of Modern Standard Arabic*. Cambridge: Cambridge University Press.

Talmon, Rafael. 1993. "Two early 'non-Sībawaihian' views of *ʿamal* in kernel-sentences". *Zeitschrift für Arabische Linguistik* 25.278–288.

——. 2003. *Eighth-century Iraqi grammar: A critical exploration of pre-Ḫalīlian Arabic linguistics*. Winona Lake, Ind.: Eisenbrauns.

Troupeau, Gérard. 1976. *Lexique-index du* Kitāb *de Sībawayhi*. Paris: Klincksieck.

Versteegh, Kees. 1995. *The explanation of linguistic causes: Az-Zaǧǧāǧī's theory of grammar*. Amsterdam and Philadelphia: J. Benjamins.

Wright, William. 1967. *A grammar of the Arabic language, translated from the German of Caspari and edited with numerous additions and corrections*. 3rd ed. rev. W. Robertson Smith and M.J. de Goeje. 2 vols. in 1 part. Cambridge: Cambridge University Press. (1st ed., 1859–1962.)

KARIN C. RYDING (Georgetown University)
KEES VERSTEEGH (University of Nijmegen)

ʾIdġām

The term *ʾidġām* (or *iddiġām*) lit. 'insertion' denotes in Arabic phonetics different types of assimilatory processes (→ assimilation), which lead either to → gemination (prolongation) of consonants or their change. The term *ʾidġām* was used by the Kufan school of grammar, whereas *iddiġām* is said to have been used by Sībawayhi and the Basrans (Suyūṭī, *Hamʿ* II, 225 quoting Ibn Ḥayyān). In fact, Troupeau (1976) has a total of 217 occurrences for the verb *ʾadġama* in the *Kitāb* and 141 for the masdar *ʾidġām; iddaġama* occurs once, in the passive (Sībawayhi, *Kitāb* I, 386.12 *ʾanna n-nūn qad tuddaġamu maʿa l-lām*); in al-Farrāʾ's *Maʿānī l-Qurʾān* (cf. Kinberg 1996:250–252) *ʾidġām* is very frequent as well; to express the notion 'to be assimilated', al-Farrāʾ uses *indaġama* once (*Maʿānī* I, 279.6). The analysis of *ʾidġām* takes up the last six chapters of Sībawayhi's *Kitāb* (*Kitāb* II, 404–431; cf. Al-Nassir 1993:56–80); in this *bāb al-ʾidġām*, Sībawayhi also includes his description of the entire phonetic system of the Arabic language. *ʾIdġām* is regarded as a subject belonging to *taṣrīf* 'inflection' (→ *ṣarf*), or rather, in modern terms, to morphophonology. It is said to occur more often in verbs, since they have more irregular forms than do nominal parts of speech (Suyūṭī, *Hamʿ* II, 225).

It should be noted that although some cases of *ʾidġām* would be described in modern phonetics as assimilation, this does not mean that the two terms are identical; in a case like *ʿanbar* > *ʿambar* ‘amber’, Sībawayhi calls the assimilation of the /n/ to the place of articulation of the /b/ → *ʾibdāl*, because the two consonants do not share the same place of articulation. In principle, *ʾidġām* is reserved for those cases where identical consonants or consonants with adjacent places of articulation are affected. Between *ʾibdāl* and *ʾidġām* there is yet another category of phonotactics, *ʾixfāʾ* lit. ‘concealment’ (Al-Nassir 1993:58); this term is used for combinations of CvCv, in which the duration of the intervening short vowel is reduced without disappearing completely. *ʾIxfāʾ* takes place in cases where *ʾidġām* is not allowed, e.g. in *ismu mūsā* ‘Moses’ name’, where *ʾidġām* would lead to an illicit cluster of three consonants.

Early grammarians like al-Xalīl had already pointed out the connection between gemination and *ʾidġām* (*ʿAyn* I, 50.4 *fa-t-tašdīd ʿalāmat al-ʾidġām*), but in later grammatical theory *ʾidġām* was identified with the result of complete assimilation, i.e., it was used to mean gemination or geminated consonants. Ibn Jinnī (*Xaṣāʾiṣ* II, 139–145) distinguishes between *iddiġām ʾaṣġar* ‘partial assimilation’ and *iddiġām kabīr*, which is only used for complete assimilation (cf. Méhiri 1973:183–186). Al-ʾAstarābāḏī (*Šarḥ* III, 233–235; cf. Bohas and Guillaume 1984:295–307) stresses the fact that the consonants have to be identical; were they not, no *ʾidġām* would result. This is different from the early tradition in the *Kitāb* where *ʾidġām* covers both gemination of identical consonants and assimilation (change) of nonidentical consonants.

When two consonants (→ *ḥarf*) have the same or an adjacent place of articulation, they are assimilated and form a geminated *ḥarf*. The first *ḥarf* loses (or is without) the following vowel (i.e., it becomes *sākin*), while the other has to be followed by a vowel (*mutaḥarrik*), i.e., it forms an open syllable.

ʾIdġām is classified by Sībawayhi as acceptable (*yajūzu*), not approved (*lā yaḥsunu*), or unacceptable (*lā yajūzu*). In later theory, other terms were used: obligatory (*wājib*), unacceptable (*mumtaniʿ*), and acceptable (*jāʾiz*). In Sībawayhi’s *Kitāb* it denotes:

i. Sandhi phenomena
   a. With accompanying haplology: two identical consonants form one geminated (long) consonant, as in *jaʿala laka* → *jaʿallaka* or *faʿala labīdun* → *faʿallabīdun*. According to Sībawayhi, *ʾidġām* in these cases is the result of a constraint in Arabic against four consecutive open short syllables (*ḥurūf mutaḥarrika*; *Kitāb* II, 407). If there are four or more, haplology sets in, as in the above examples.
   b. With accompanying assimilation:
      n + *l* → *nn̩*: *buyyina lahum* > *buyyinnahum* (بُيِّنَ لَهُمْ ← بُيِّنَّهُمْ)
      x + *ġ* → *ġġ*: *islaḥ* × *ġanamaka* > *islaġġanamaka* (اسلخ غنمك ← اسلغّنمك)
      t + *s* → *ss*: *ḏahabat salmā* > *ḏahabassalmā* (ذهبَتْ سلم ← ذهبسّلمى)
      d + *s* → *ss*: *qad samiʿta* > *qassamiʿta* (قد سمعت ← قسّمعت)

ii. Haplology
   Obligatory: *humā yuḏ̣limāninī* > *humā yuḏ̣limānnī*
   Acceptable: *tatakallamūna* > *takallamūna; tatarassūna* > *tarassūna*; *tatanazzalūna* > *tanazzalūna*
   Obligatory *ʾidġām* is explained by reference to underlying structures. For instance, forms of verbs such as *radda, madda, wadda* are analyzed as *radada* > *radda*, on the basis of existing forms such as *radadtu (> radattu).* In forms such as *yamuddu*, the underlying process is described as transfer (*naql*) of the vowel *-u-* from the *du* syllable to *mu: yamdudu* > *yamuddu.*

iii. Progressive assimilation
   Between two consonants having the same place of articulation: *mudtakir* > *muddakir* > *muḏḏakir* (but also regressive: > *muddakir*); *iṣtabara* > *iṣṭabara*; *muḍtajiʿ* > *muḍṭajiʿ* > *muḍḍajiʿ/muṭṭajiʿ*; *iṭtaʿanū* > *iṭṭaʿanū*; *idtānū* > *iddānū*. Although progressive assimilation of verbal endings (as in *faḥaṣtu* > *faḥaṣṭu*, فَحَصْتُ > فَحَصْطُ; *ḥafiḏ̣tu* > *ḥafiṭṭu*, etc.) sometimes occurs, it is deemed rare. When it occurs, it is accepted but not approved, since the *-tu* is a personal verbal ending having the status of a pronoun (*ʿalāmat al-ʾiḍmār*); therefore, it should retain its initial form. However, regressive assimilation of verbal endings is acceptable since the initial form of the ending is preserved: *madadtu* > *madattu.*

iv. Regressive assimilation
tasdīr > *tazdīr* (تزدير < تسدير)
tasdulu > *tazdulu* (تزدل < تَسْدُل)

As a matter of fact, almost all instances of *ʾidġām* in the *Kitāb* concern regressive assimilation. Al-Nassir (1993:80) calculates that out of 123 cases of *ʾidġām* only 31 are cases of progressive assimilation. of which only 6 concern total progressive assimilation; the rest are cases of regressive assimilation. This is in agreement with Sībawayhi's repeated rule that *ʾidġām* is the assimilation of the first element with the second one (e.g. *Kitāb* II, 254.5–7 *wa-l-ʾidġām ʾinnamā yadxulu fīhi l-ʾawwal fī l-ʾāxir wa-l-ʾāxir ʿalā ḥālihi wa-yuqlabu l-ʾawwal wa-yadxulu fī l-ʾāxir ḥattā yaṣīra huwa wa-l-ʾāxir min mawḍiʿ wāḥid*). This first element is, in the overwhelming majority of cases, a coronal consonant (Al-Nassir 1993:80).

A special case is that of the assimilation of the consonant /n/ to medial consonants (Al-Nassir 1993:68–69), e.g. *man kāna* 'who was', *man jāʾa* 'who came', etc. The resulting value of /n/, [ŋ] or [ɲ], respectively, is regarded by Sībawayhi as one of the 'acceptable consonants' of Arabic. In his view, the original place of articulation of /n/ is 'concealed' (*ʾixfāʾ*), but its release through the nasal cavity remains, and this is enough to preserve its status as a separate consonant.

In accordance with the use of *ʾidġām* in later Arabic grammatical theory, modern Arabic linguistic terminology usually employs the term to mean gemination resulting from assimilation. Related terms are *mudġam* 'the assimilated [consonant]' and *mudġam ʾilayhi* '[the consonant] to which something is assimilated'.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

ʾAstarābāḏī, *Šarḥ* = Raḍī d-Dīn Muḥammad ibn al-Ḥasan al-ʾAstarābāḏī, *Šarḥ Kāfiyat Ibn al-Ḥājib*. Ed. Muḥammad Nūr al-Ḥasan, Muḥammad az-Zafzāf, and Muḥammad Muḥyī d-Dīn ʿAbd al-Ḥamīd. 2 vols. Beirut: Dār al-Kutub al-ʿIlmiyya, 1975.

Farrāʾ, *Maʿānī* = ʾAbū Zakariyyāʾ Yaḥyā ibn Ziyād al-Farrāʾ, *Maʿānī l-Qurʾān*. Ed. Muḥammad ʿAlī an-Najjār. 3 vols. Cairo: ad-Dār al-Miṣriyya, 1966–1972.

Ibn Jinnī, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. Ed. Muḥammad ʿAlī an-Najjār. 3 vols. Cairo, 1952–1956. (Repr., Beirut: Dār al-Hudā, n.d.)

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. 2 vols. Būlāq, 1318 A.H.

Suyūṭī, *Hamʿ* = Jalāl ad-Dīn ʿAbd ar-Raḥmān ibn ʾAbī Bakr as-Suyūṭī, *Hamʿ al-hawāmiʿ fī šarḥ Jamʿ al-jawāmiʿ*. Repr. Beirut n.d.

Xalīl, *ʿAyn* = ʾAbū ʿAbd ar-Raḥmān al-Xalīl ibn ʾAḥmad al-Farāhīdī, *Kitāb al-ʿayn*. Ed. Mahdī al-Maxzūmī and ʾIbrāhīm as-Sāmarrāʾī. 8 vols. Beirut: Muʾassasat al-ʾAʿlamī li-l-Maṭbūʿāt, 1988.

SECONDARY SOURCES

Al-Nassir [an-Nāṣir], ʿAbd al-Munʿim ʿAbd al-ʾAmīr. 1993. *Sibawayh the phonologist: A critical study of the phonetic and phonological theory of Sibawayh as presented in his treatise* Al-Kitab. London and New York: Kegan Paul International.

Bohas, Georges and Jean-Patrick Guillaume. 1984. *Étude des théories des grammairiens arabes*. I. *Morphologie et phonologie*. Damascus: Institut Français de Damas.

Kinberg, Naphtali. 1996. *A lexicon of al-Farrāʾ's terminology in his* Qurʾān *commentary with full definitions, English summaries, and extensive citations*. Leiden: E.J. Brill.

Méhiri, Abdelkader. 1973. *Les théories grammaticales d'Ibn Jinnī*. Tunis: Université de Tunis.

Troupeau, Gérard. 1976. *Lexique-index du* Kitāb *de Sībawayhi*. Paris: Klincksieck.

JANUSZ DANECKI
(University of Warsaw)

ʾIḍmār

Although the term *ʾiḍmār* lit. 'keeping in mind' sometimes occurs interchangeably with → *ḥaḏf* '→ elision', it represents an entirely different phenomenon. Whereas *ḥaḏf* denotes an omission at the surface level, as a purely phonological event that leaves the utterance formally incomplete, *ʾiḍmār* refers to the mental act of suppressing an element at what might now be called the deep-structure level, independent of any phonological realization, and not necessarily producing a formally incomplete utterance (→ ellipsis). Moreover, it is a feature of *ʾiḍmār* that, unlike *ḥaḏf*, only complete morphemes (mostly whole words) can be 'suppressed'. Significantly, the concept of 'suppression' also supplies the cognate term for 'pronoun', → *ḍamīr* lit. 'thing kept in the mind' (in Modern Standard Arabic also 'conscience'), so that *muḍmar* can mean both 'suppressed' and 'pronominalized'. One class that is regularly 'suppressed' is the agent pronoun of certain verb forms: thus, *ḍaraba-nī* is interpreted by default as 'he struck me' in the absence of an overt agent (*muḏhar* 'expressed openly', antonym of *muḍmar*), but in the sentence *ḍaraba-nī wa-ḍarabtu zaydan* lit. 'struck me and I struck Zayd', i.e. '[Zayd] struck me and I struck Zayd', the agent of *ḍaraba-nī* 'struck me' must be interpreted as elided (*maḥḏūf*) and not pronominalized (*muḍmar*), because a pronoun normally cannot refer cataphorically to a following noun (→ cataphora), so the full sentence is *ḍaraba-nī [zaydun] wa-ḍarabtu zaydan*. *ʾIḍmār* also denotes the 'keeping in mind' of other elements, particularly the subordinating conjunction *ʾan* 'that', thus accounting, *inter alia*, for the ability of the preposition *li-* 'for, to' to subordinate verbs; hence, *li-yaktuba* 'so that he may write', is analyzed as *li-[ʾan] yaktuba* lit. 'for that he may write'. Finally, *ʾiḍmār* in metrics refers to the shortening of a prosodic pattern, e.g. the foot *mutafāʿilun* to *mustafʿilun* (= *mut[a]fāʿilun*). It will become apparent that the distinction between *ḥaḏf* 'ellipsis' and *ʾiḍmār* 'suppression' is sometimes blurred. However, although as a rule of thumb all elided elements can be thought of as suppressed, not all suppressed items are elided; thus, the suppressed agent pronoun in *ḍaraba* 'he struck' cannot be restored in the surface structure (*ḍaraba huwa* lit. 'struck he' is an emphatic expression meaning 'he and not someone else struck'). The agent pronoun in *ḍaraba* is called *mustatar* 'concealed' (→ *ḍamīr*).

The term *ʾiḍmār* belongs to the oldest layer of Arabic grammatical terminology: it occurs already in the early exegetical literature, for instance in Muqātil's (d. 150/767) *Tafsīr* (Versteegh 1993:146–151). In the majority of passages in which Muqātil uses this term, it indicates the suppression of an attributive or prepositional phrase, e.g. Q. 33/50 *fa-sajada l-malāʾikatu* 'and the angels bowed down', to which Muqātil (*Tafsīr* III, 653.11) adds *allaḏīna kānū fī l-ʾarḍ: ʾiḍmār* 'the ones that were on the earth: suppression'. In such cases, *ʾiḍmār* indicates the suppression of an element that may be reconstructed from the context but which is not necessary for the syntactic construction. In later exegetical literature, the meaning of *ʾiḍmār* shifts to those instances of suppression where the suppressed element is necessary for the explanation of the surface structure. ʾAbū ʿUbayda (d. 210/825), for instance, uses the term *ḍamīr fiʿl* to explain the dependent (accusative) form in Q. 2/135 *bal millata ʾIbrāhīma*: "It is in the dependent form because there is a *ḍamīr fiʿl* in it, as if its actual meaning is 'follow Ibrahim's religion'" (*intaṣaba li-ʾanna fīhi ḍamīr fiʿl ka-ʾanna majāzahu bal ittabaʿū millata ʾIbrāhīma*; *Majāz* I, 57.10–11). Of course, *ḍamīr* could also refer here to some kind of concealed agent pronoun.

Muḍmar is a common term among the Kūfan grammarians (cf. Kinberg 1996:430–433). According to Dévényi (1990:104), al-Farrāʾ (d. 207/822) uses *muḍmar* to refer to words that are 'hidden' in one reading of the *Qurʾān* but appear in another reading (e.g. *Maʿānī* I, 141). In her view, this differs from the use of the term in the *Kitāb Sībawayhi*, where *muḍmar* does not refer to any actual reading in which a word appears but rather to a word that is suppressed in the surface structure, and which may be restored in the underlying structure.

In the *Kitāb*, *muḍmar* and its cognates are used frequently, both in the sense of 'suppression' and of 'pronominalization' (Troupeau 1976:132; see the index by ʿUḍayma 1975, s.v.). Ayoub (1990) interprets this double meaning by focusing on the role of the speaker. In her view, the common factor is that in both cases the speakers leave out something in their speech because they know that the interlocutor will understand what is being referred to. This is why *ʾiḍmār* is often used in conjunction with the expression *fī n-niyya* 'in the intention' (Ayoub 1990:3–4; e.g. Sībawayhi, *Kitāb* I, 106 ed. Derenbourg/I, 127 ed. Bulaq). In all instances, the choice between using a verb or suppressing it is connected with a pragmatic motivation on the part of the speaker who has to decide whether or not to use the verb. The intended sense of the utterance may be reconstructed by the grammarian through *tamṯīl* (called in later grammatical treatises → *taqdīr*). The grammarians may then use this reconstructed element to explain the syntactic structure, for instance to explain the *-a* ending in *li-yaktub-a* by reconstructing a suppressed *ʾan*.

Elision or suppression is not unconditioned. Sībawayhi laid down the principle that there are three categories: those where it is incorrect to suppress the verb (*lā yaḥsunu ʾiḍmāruhu*); those where the verb is usually elided; and those where it is always elided, the last also including elliptical expressions that have become proverbial (*Kitāb* I, 125 ed. Derenbourg/I, 149 ed. Bulaq):

> Know that…the verb has three different ways with respect to the noun: a verb that is used explicitly and whose suppression is incorrect; a verb that is suppressed but may be used explicitly; and a suppressed verb that is never used explicitly (*fa-ʿraf…ʾanna l-fiʿl yajrī fi l-ʾasmāʾ ṯalāṯata majārin: fiʿl muḏ̣har lā yaḥsunu ʾiḍmāruhu wa-fiʿl muḍmar mustaʿmal ʾiḏ̣hāruhu wa-fiʿl muḍmar matrūk ʾiḏ̣hāruhu.*)

Sībawayhi then gives examples of these three categories. The first case is when you say to someone *zaydan* without any mention of hitting; in such a case the use of a verb is compulsory because otherwise the interlocutor would not know that he was meant to hit Zayd. The second category is when you say *zaydan* to someone when hitting has been mentioned, so that it is obvious that you intend him to hit Zayd; here, the verb may or may not be used. The third category is that of phrases like *marḥaban* 'welcome', where it would be incorrect to add an explicit verb.

The listeners' ability to recover the missing elements based on their own knowledge (*ʿilm al-muxāṭab*) was schematized by later grammarians into the various kinds of 'contextual indicators' (*dalīl*), both linguistic and situational, which accompany elision and suppression and through which the countless elliptical passages in the *Qurʾān* were reconstructed by the process of → *taqdīr* 'suppletive insertion'. The treatise on the inflection of the *Qurʾān* attributed to az-Zajjāj, for instance, is a rich compendium of grammatically classified material in which numerous instances of *ʾiḍmār* (and *ḥaḏf*) are dealt with (*ʾIʿrāb*, Chaps. 1, 2, 4,10, 14, 15, 16, 20, 31, 32, 33, 47, 52, 61, 63, 66, 68, 71; in Chap. 84 → cataphora is treated; cf. the classified index, pp. 973–1054).

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

ʾAbū ʿUbayda, *Majāz* = ʾAbū ʿUbayda Maʿmar ibn al-Muṯannā at-Taymī, *Majāz al-Qurʾān*. Ed. Muḥammad Fuʾād Sazkīn [Fuat Sezgin]. 2 vols. Cairo: Muḥammad Sāmī ʾAmīn al-Xānjī, 1954.

Farrāʾ, *Maʿānī* = ʾAbū Zakariyyāʾ Yaḥyā ibn Ziyād al-Farrāʾ, *Maʿānī l-Qurʾān*. Ed. Muḥammad ʿAlī an-Najjār. 3 vols. Cairo: ad-Dār al-Miṣriyya, 1955–1972.

Muqātil, *Tafsīr* = ʾAbū l-Ḥasan Muqātil ibn Sulaymān al-Balxī, *at-Tafsīr*. Ed. ʿAbdallāh Maḥmūd Šiḥāta. 4 vols. Cairo: al-Hayʾa al-Miṣriyya al-ʿĀmma li-l-Kitāb, 1980–1987.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. Ed. Hartwig Derenbourg. Paris: Imprimerie Nationale, 1881–1889. (Repr., Hildesheim and New York: G. Olms, 1970.) Ed. Būlāq, 1898–1900. (Repr., Baghdad: Muthanna [1965].)

(Zajjāj), *ʾIʿrāb* = (attrib.). *ʾIʿrāb al-Qurʾān, al-mansūb ʾilā z-Zajjāj*. Ed. Ibrāhīm al-ʾAbyārī. Cairo: al-Hayʾa al-ʿĀmma li-Šuʾūn al-Maṭābiʿ al-ʾAmīriyya, 1963–1964.

SECONDARY SOURCES

Ayoub, Georgine. 1990. "De ce qui 'ne se dit pas' dans le *Livre* de Sībawayhi: La notion de *tamṯīl*". *Studies in the history of Arabic grammar*, II, ed. Michael G. Carter and Kees Versteegh, 1–15. Amsterdam and Philadelphia: J. Benjamins.

Carter, Michael G. 1991. "Elision". *Proceedings of the Colloquium on Arabic Grammar, Budapest 1–7 September 1991*, ed. Kinga Dévényi and Tamás Iványi, 121–133. Budapest: Eötvös Loránd University and Csoma de Kőrös Society.

Dévényi, Kinga. 1990. "On Farrāʾ's linguistic methods in his work *Maʿānī l-Qurʾān*". *Studies in the history of Arabic grammar,* II, ed. Michael G. Carter and Kees Versteegh, 101–110. Amsterdam and Philadelphia: J. Benjamins.

Fleisch, Henri. 1971. "*Iḍmār*". *Encyclopaedia of Islam,* III, 1027–1028. 2nd ed. Leiden: E.J. Brill.

Gully, Adrian J. 1994. *Grammar and semantics in Medieval Arabic*, 207–217. Richmond: Curzon.

Kinberg, Naphtali. 1996. *A lexicon of al-Farrāʾ's terminology in his* Qurʾān *commentary, with full definitions, English summaries, and extensive citations*. Leiden: E.J. Brill.

Troupeau, Gérard. 1976. *Lexique-index du* Kitāb *de Sībawayhi*. Paris: Klincksieck.

ʿUḍayma, ʿAbd al-Xāliq. 1975. *Fahāris Kitāb Sībawayhi wa-dirāsa lahu*. Cairo: Maṭbaʿat as-Saʿāda.

Versteegh, Kees. 1993. *Arabic grammar and Qurʾānic exegesis in early Islam*. Leiden: E.J. Brill.

MICHAEL C. CARTER
(Sydney University)
KEES VERSTEEGH
(University of Nijmegen)

ʾIʿjāz

The term *ʾiʿjāz* is a technical term referring to the status of the *Qurʾān* as a miracle (*muʿjiz* or *muʿjiza*) resisting imitation and confirming the Prophet's mission. The doctrine of the inimitability of the *Qurʾān* was founded on a number of Qurʾānic verses in which the authenticity of the Prophet's mission was linked to a challenge (*taḥaddī*) addressed to unbelievers to produce a likeness (*miṯl*) of a specified portion of it: "Or do they say, 'He fabricated the [Message]?' Nay, they have no faith! Let them produce a recital unto it – if [it be] they speak the Truth!" (Q. 52/33–34; cf. Q. 2/23–24, 10/38, 11/13, 17/88). This notion of matching or emulating (*muʿāraḍa*) reflected the competitive practice of pre-Islamic poets, and there are several reports of attempts to meet the Qurʾānic challenge by the Prophet's contemporaries, such as Musaylima ibn Ḥabīb, whose verses earned him the taunts of later writers. But it was in the 3rd/9th century that the doctrine took shape and found its place at the confluence of several Islamic sciences, such as exegesis, theology, and the nascent practice of literary criticism, whose resources it both used and helped extend.

The development of the doctrine must be seen as a response to several social, political, and intellectual contexts, such as the heightened interest in the notion of prophethood and in miracles as the grounds of prophetic claims which was stimulated by interreligious (especially Muslim and Christian) polemics in the 3rd/9th century, as well as by the challenges arising from within the Islamic community itself, where attacks on the notion of prophethood were often combined with criticisms of the *Qurʾān*. Ibn ar-Rāwandī (d. 245/860 or ca. 298/912) is the most famous example, but similar challenges were ascribed to representatives of the *šuʿūbī* movement in the same century, whether poets or secretaries, several of whom attracted charges of freethinking or dualism (*zandaqa*; Abdul Aleem 1933; Martin 1980, 2002). Thus, the doctrine of *ʾiʿjāz* might be seen in its connection with the battles over language which were bound up with larger questions about cultural and ethnic allegiance involved in *šuʿūbiyya*, and hence as a force cementing a sense of community – a community ranged around a text whose fixedness the doctrine could be seen as reflecting and celebrating (cf. Audebert 1982:10–11). In addition, the doctrine had close links with the theological controversies concerning the status of the *Qurʾān* (whether created or eternal) and the divine attribute of speech that dominated the 3rd/9th century and constituted a formative influence on the perspective on language developed across a variety of Islamic sciences. Theological commitments on these questions were an important factor affecting positions adopted on the question of *ʾiʿjāz* (Larkin 1995).

Several unpreserved works appeared in the 3rd/9th century under the title of *Naḏm al-Qurʾān*, among them one by the renowned Muʿtazilite littérateur al-Jāḥiẓ (d. 255/868–869), and have been interpreted as precursors of the later treatments of the topic, initiating an exploration of the *Qurʾān*'s literary qualities (see Audebert 1982:57–71). Special treatises under the title of *ʾiʿjāz al-Qurʾān* begin to be written in the 4th/10th century, with the Muʿtazilite ar-Rummānī's (d. 384/994) *an-Nukat fī ʾiʿjāz al-Qurʾān* and the traditionalist al-Xaṭṭābī's (d. 386/996 or 388/998) *Bayān ʾiʿjāz al-Qurʾān,* while the works often deemed to consummate the treatment of the doctrine appear soon after, in the contributions of the Ashʿarite authors ʾAbū Bakr al-Bāqillānī (d. 403/1013; *ʾIʿjāz al-Qurʾān*) and ʿAbd al-Qāhir al-Jurjānī (d. 471/1078; *Dalāʾil ʾiʿjāz al-Qurʾān*), on which the greatest focus will be placed here.

Investigations into the Qurʾānic miracle had to engage with several distinct questions, of which the most important concerned the nature of the miracle and the means by which knowledge of it was attained. The first question was directly invited by the terms of the challenge itself, which challenged unbelievers to produce the 'like' of the *Qurʾān*. But 'like' in what respect? Two approaches came to be distinguished early in the history of the debates, one of which located the miracle in certain features of the *Qurʾān* itself, while the other located it in God's act of averting or preventing (*ṣarafa*) human beings from attempts to rival the *Qurʾān* by depriving them of motivation and interest. Known as the doctrine of *ṣarfa,* the latter was notably propounded by the Baġdādī Muʿtazilite ʾAbū ʾIsḥāq an-Naẓẓām (d. between 220/835 and 230/845), and its corollary was the claim that "had God left [the Arabs] to their own devices, they would have been capable of producing a *sūra* which was like it in eloquence" (Šahrastānī, *Milal* I, 56–57). Like others who followed his lead (notably ar-Rummānī), he was happy to combine this view with an approach of the first kind, and here he claimed that the feature internal to the *Qurʾān* which constituted its miraculous nature was the knowledge of *ġuyūb* – information inaccessible to human beings, such as prophetic statements – which it contained.

An-Naẓẓām's view remained popular with Baġdādī Muʿtazilites as well as Imamite Shiʿites, but most writers on the topic – in particular the Baṣran Muʿtazilites, Zaydite Shiʿites, and Ashʿarites – rejected it on both counts. On the one hand, they took issue with its implied ascription of miraculous quality to God's act of averting as against the *Qurʾān* itself, and some went so far as to claim consensus for the latter position (Suyūṭī, *ʾItqān* II, 231; cf. al-Bāqillānī's discussion of the *ṣarfa* view in *ʾIʿjāz* 41–43, and ʿAbd al-Jabbār's in *Muġnī* XVI, 217–220, 323–328). Equally important, the dominant view within the second approach was to identify the miraculous quality, not primarily within the content of the message, but within its form and, in particular, within its singular and insuperable eloquence.

The long probing of Muslim scholarship into the stylistic wonders of the *Qurʾān* which thus began would play host to several key themes that linked it to other domains of scholarship in various ways. One prominent theme, which brought investigations into *ʾiʿjāz* into the intellectual orbit of literary criticism, appeared in answer to the second question mentioned above, concerning the means by which knowledge of the miracle could be gained. Insofar as the miracle was to serve as a proof (*ḥujja*) and a sign of prophethood (one of the Prophet's *dalāʾil an-nubuwwa*), knowledge of it needed to be readily accessible – its target audience was not the believer but the unbeliever. Indeed, "The underlying assumption of the challenge was that the merit and beauty of the *Qurʾān* could be appreciated even by those outside the hold of faith" (Mir 1988:51). The fact of its accessibility was urged by reports of spontaneous responses to the beauty of the Qurʾānic language by the Prophet's contemporaries, which could produce conversions and secure grudging admissions even from the Prophet's opponents (see az-Zarkašī's paradigmatic description of al-Walīd ibn al-Muġīra's awestruck response in *Burhān* II, 110–111). At the same time, the fact that knowledge of the *Qurʾān*'s aesthetic excellence required cultivation was made evident in certain writers' classification of the miracle as one known through a process of proof (*istidlāl*). The more nuanced position on the topic was that an immediate knowledge was possible to those of trained literary judgment and ability, whereas others would need to rely on proof and avail themselves of the indirect or circumstantial evidence provided by the fact that the Arabs, acknowledged masters of their eloquence, had not produced a successful match for the *Qurʾān* (Bāqillānī, *ʾIʿjāz* 393; cf. the discussions in Ġazālī, *Iqtiṣād* 206–208; Juwaynī, *ʾIršād* 288–295; Weiss 1992:74–79). This understanding of aesthetic judgments was in line with developments in the field of literary criticism, where the notion of the critic as an authority whose literary judgment overruled the subjective response of ordinary people, and of criticism as a craft (*ṣināʿa*) requiring professional training and producing specialized knowledge, had emerged clearly, beginning with Ibn Sallām (d. ca. 232/847) in the 3rd/9th century and carried forward by his successors in the 4th/10th (Abu Deeb 1990, esp. 348–349). While never losing sight of the limits of explanation, these developments led away from aesthetic subjectivism and toward a view of literary judgments as justifiable by reasons.

A second theme carried discussions of the Qurʾānic *ʾiʿjāz* into the range of *kalām*, where the debates concerning the nature of divine speech – chiefly polarized between Muʿtazilite and Ashʿarite theologians – had issued in theologically mindful accounts of the nature of language and the relation between word (→ *lafḏ*) and meaning (→ *maʿnā*). These accounts provided the backdrop for a debate to which early critics such as al-Jāḥiẓ and Ibn Qutayba (d. 276/889) had made formative contributions which sought to identify the locus of eloquence, and in particular whether it lay in the words of a text or its meaning – a contrast between form and content different from, but not entirely unrelated to, the debate between an-Naẓẓām and his detractors.

These themes provide some important prisms for presenting and contrasting the works of al-Bāqillānī and al-Jurjānī in this field. An accomplished theologian, al-Bāqillānī's *ʾIʿjāz al-Qurʾān* was a layman's foray into the field of literary studies (cf. the comments of van Gelder 1982:100; Grunebaum 1950:xx), and he draws heavily on the work of predecessors such as ar-Rummānī, Qudāma ibn Jaʿfar (d. 337/948), and al-ʾAskarī (probably ʾAbū ʾAḥmad, d. 382/993). He identifies several elements which might constitute the *Qurʾān*'s miraculous character (*ʾIʿjāz* 48–71), but his main focus is on its eloquence (*faṣāḥa*, *balāġa*), which he claims surpasses that of any other speech. His account of this position is multifaceted, but his urge throughout is to present the *Qurʾān* as a literary phenomenon transcending the categories of human literary creation. Thus, one of the grounds on which al-Bāqillānī argues the case of the *Qurʾān*'s unsurpassable eloquence is in its sui generis construction (*naḏm*), which defies classification and departs from all accustomed styles, modes, or genres, such as poetry or rhymed prose (→ *sajʿ; ʾIʿjāz* 51–52, 76ff.). The same desire to steer clear of category is revealed in al-Bāqillānī's discussion of the types of *badīʿ* – which he lists in some detail – where again he insists that, while the *Qurʾān* partakes of such figures, its eloquence cannot be reduced to them. Employing many of the concepts familiar in the literary thinking of his contemporaries, he likens such figures to a craft attainable through training and argues that the grounds of the *Qurʾān*'s miraculous character must lie beyond the reach of human artifice (Bāqillānī, *ʾIʿjāz* 162).

The challenge, then, is to place the Qurʾānic eloquence beyond human artifice, yet still within human capacities of aesthetic appreciation, beyond human literary categories, yet still within the reach of judgments of value which stand in need of such categories. However, this is not ultimately the task that al-Bāqillānī has set himself. Whether it was in awareness – or, if not, then in token – of his status as a layman in the field of literary studies (a status expressed in the pervasive mode of third-person reference to the judgment of the specialists or *ʾahl aṣ-ṣanʿa*), or whether it was a result of his stress on the transcendence of the Qurʾānic eloquence, al-Bāqillānī's positive strategy for revealing this quality does not involve an analytical approach to the text (cf. the remarks in ʿAbbās 1971:353–354; his negative strategy, consisting of a critique of the *Muʿallaqa* of Imruʾ al-Qays, among the 'Ancients', and a poem by al-Buḥturī, among the moderns, is examined by van Gelder 1982:100–107, Mir 1990, and Grunebaum 1941). Diffuse and rhetorically effusive, its task is not to educate the reader's judgment through literary analysis and lead him through a perception of the grounds of eloquence to an aesthetic response. Within this framework, the capacity for responding to the beauty of the *Qurʾān* is demanded as a prerequisite – for otherwise al-Bāqillānī's invitations to contemplate examples of it (*ʾIʿjāz* 279ff.) would meet with no response – as, indeed, is the willingness to engage in the attempt, insofar as this cannot be urged on by reason-giving. Both demands are evident in al-Bāqillānī's final admonition to his reader, should he fail to perceive the beauty toward which al-Bāqillānī points, to "take one's seat among the *muqallidīna*" (*ʾIʿjāz* 370), contenting himself with *taqlīd* – the mode of knowledge which consists of accepting the judgment of authority. This authority must be understood at the same time as literary and religious authority (that the latter is also at issue is revealed further in al-Bāqillānī's references to the authority, internal to faith, of consensus [*ʾijmāʿ*]: *ʾIʿjāz* 389). At his hands, the inquiry into the *Qurʾān*'s unsurpassable eloquence, formally conceived as a probative sign addressed to the unbeliever, becomes a form of *fides quaerens intellectum* – a devotional practice in which the miraculous character of the sacred text is presupposed in the attempt to perceive it.

Al-Jurjānī – fellow Ashʿarite, philologist, literary theorist, and thus member of the specialist *ʾahl aṣ-ṣanʿa* which is the subject of al-Bāqillānī's allusions – approaches the topic in a way which shares many of the elements of al-Bāqillānī's outlook while displaying fundamental differences in others. He, too, locates the miracle in the *Qurʾān*'s inimitable eloquence (*faṣāḥa, balāġa, bayān*), which he examines in the context of a multilayered project addressing the nature and locus of eloquence, the (closely linked) questions of the relation between meaning and words, and the means by which eloquence may be known. Rejecting narrow views of eloquence which locate it *merely* at the level of isolated words, of the aural qualities of speech, or of its semantic content, he argues for a more holistic view of both eloquence and language. Eloquence is a matter of the *naḏm* – the ordering or construction – of speech, where *naḏm* is defined in terms of the features of grammar (*maʿānī n-naḥw*) and is seen as the carrier of an indissoluble connection between thought and language, insofar as the order of words is determined by the order of meanings in the mind. Al-Jurjānī's position on the semantic dimension of eloquence is supported by his discussion of figurative language (→ *majāz*), arguably a showpiece of eloquent speech, claiming that devices such as → *istiʿāra* 'metaphor', *tamṯīl* 'analogous comparison', → *kināya* 'metonymy' or 'allusiveness' are only intelligible in terms of second-order signification (*Dalāʾil,* passim; Larkin 1988, 1995; Abu Deeb 1979).

Al-Jurjānī's stress on the unity of thought and language and on the semantic dimension of eloquence aims in part to heal the rift, fraught with theological significance, pervasive in Muʿtazilite views of the nature of language and, as a corollary, of the nature of eloquence. Part of his project constitutes a polemical engagement with the views of the Baṣran Muʿtazilite ʿAbd al-Jabbār (d. 415/1025; Larkin 1988, 1995; for his views see briefly Rahman 1996). Al-Jurjānī's stress on the semantic component may be seen simultaneously as supporting an Ashʿarite understanding of language in terms of a connection between uncreated internal speech (*kalām an-nafs* or *nafsī*) and its external expression as phonic speech (*kalām lafḏī* or *lisānī*) which buttressed Ashʿarite claims about the eternity of the *Qurʾān* as God's speech (Larkin 1995; cf. the presentation of the Ashʿarite position in Weiss 1992:65–69). Simultaneously, in overcoming the duality of thought and language through a conception of grammar in which the structure of the linguistic utterance is based on the structure of a mental act, al-Jurjānī was helping to overcome a view of language – made compelling by a prevalent conception of its institution – as a naming mechanism intelligible in terms of a correlation between words and objects (see especially his remarks in *Dalāʾil* 374ff.).

Al-Jurjānī's opposition to placing the *Qurʾān*'s eloquence in the realm of the ineffable subjective reactions and his commitment to the justifiability of aesthetic response in terms of reasons and grounds are expressed clearly in his work, in the insistence that "for every instance of discourse which you approve (*tastaḥsinuhu*) . . . your approval of it must have a known ground and intelligible cause (*ʿilla*)" (*Dalāʾil* 85). The language of reason, cause, and proof (*dalīl*, *ʿilla*, and *sabab* all appear in al-Jurjānī's idiom), carried over to an investigation of the *Qurʾān*'s eloquence, pulls in the opposite direction from al-Bāqillānī's transcendentalism. Here, al-Jurjānī's interest in producing a unified, general account of the qualities determining literary value could be deemed to frustrate an appreciation of the special otherness of the *Qurʾān*, insofar as it is handled in terms of its instantiation of grounds of value which it shares with other literary works and is thus assimilated to human literary creations (Vasalou 2002). Taken as a remark about the absence, in al-Jurjānī's work, of an explicit analytical statement of the features constituting the *Qurʾān*'s miraculous eloquence, this appraisal may rest on a conception of the aim as that of *saying* or of *stating,* whereas al-Jurjānī's work might be cast as an effort to rather *show*, which, by training one in the recognition of literary beauty through the recognition of general grounds exemplified by means of a variety of literary works, cultivates the taste that equips one to perceive the *Qurʾān*'s inimitable beauty. Later writers using al-Jurjānī's tools, such as az-Zamaxšarī (d. 538/1144) in his Qurʾānic commentary *al-Kaššāf* (Boullata 1988:146–7), may be said to exemplify this training.

On the other hand, careful attention to al-Jurjānī's method will reveal that his strategy in examining literary passages – whether from poetry or from the *Qurʾān* – is often to begin from the aesthetic response (of approval, of delight, of wonderment) and, working backward, to provide a literary analysis which comes as an interpretation of this response by uncovering its latent reasons (for an example, see *Dalāʾil* 282). The question which such a strategy raises is faced boldly by al-Jurjānī when he confronts the possibility (*Dalāʾil* 284) of a failure to respond differently to qualitatively different types of utterance – to have the response that becomes material for interpretation. Here, acquiescing to the logic of his specialized craft, he accepts that the capacity for such response may be confined to the people of taste and knowledge (*ʾahl aḏ-ḏawq wa-l-maʿrifa*) and, regarding it as a given which one either has or fails to have, countenances the existence of those who altogether lack the instrument (*ʾadā*) and the sense (*ḥāssa*) by which such knowledge is attained, whom he would advise against seeking to enlighten ("how little words avail with such a one!"). Whether or not al-Jurjānī's professional exclusiveness here prevented him from doing sufficient justice to his own work and its capacity to cultivate aesthetic taste and judgment, these features of his project render it a contribution, less to the articulation of an apologetic theological doctrine than to the development of the sciences of rhetoric (*ʿilm al-maʿānī* and *ʿilm al-bayān*), where he can rightly be credited with having bequeathed a towering presence to the generations that followed.

Bibliographical references

Primary sources

ʿAbd al-Jabbār, *Muġnī* = ʾAbū l-Ḥasan ʿAbd al-Jabbār ibn ʾAḥmad al-ʾAsadābādī, *al-Muġnī fī ʾabwāb at-tawḥīd wa-l-ʿadl*. 16 vols. Various editors. Cairo: Various publishers, 1960–1969.

Bāqillānī, *ʾIʿjāz* = ʾAbū Bakr Muḥammad ibn aṭ-Ṭayyib al-Bāqillānī, *ʾIʿjāz al-Qurʾān*. Ed. as-Sayyid ʾAḥmad Ṣaqr. Cairo: Dār al-Maʿārif, 1954. [Translated selections in Gustav E. von Grunebaum, *A tenth-century document of Arabic literary theory and criticism*. Chicago: University of Chicago Press, 1950.]

Ġazālī, *Iqtiṣād* = ʾAbū Ḥāmid Muḥammad ibn Muḥammad al-Ġazālī, *al-Iqtiṣād fī l-iʿtiqād*. Eds. Ibrahim Agâh Çubukçu and Hüseyin Atay. Ankara: Nur Matbaası, 1962.

Jurjānī, *Dalāʾil* = ʾAbū Bakr ʿAbd al-Qāhir ibn ʿAbd ar-Raḥmān al-Jurjānī, *Dalāʾil al-ʾiʿjāz*. Ed. Muḥammad ʿAbd al-Munʿim Xafājī. Cairo: Maktabat al-Qāhira, 1969.

Juwaynī, *ʾIršād* = ʾAbū l-Maʿālī ʿAbd al-Malik ibn ʿAbdallāh al-Juwaynī, *Kitab al-ʾiršād ʾilā qawāṭiʿ al-ʾadilla fī ʾuṣūl al-iʿtiqād*. Ed. ʾAsʿad Tamīm. Beirut: Muʾassasat al-Kutub aṯ-Ṯaqāfiyya, 1996.

Šahrastānī, *Milal* = Tāj ad-Dīn ʾAbū l-Fatḥ Muḥammad ibn ʿAbd al-Karīm aš-Šahrastānī, *al-Milal wa-n-niḥal*. 2 vols. in 1. Ed. Muḥammad Sayyid Kīlānī. Cairo: Maktabat Muṣṭafā al-Bābī al-Ḥalabī, 1967.

Suyūṭī, *ʾItqān* = ʾAbū l-Faḍl ʿAbd ar-Raḥmān ibn ʾAbī Bakr as-Suyūṭī, *al-ʾItqān fī ʿulūm al-Qurʾān*. 2 vols. Ed. Muḥammad Sālim Hāšim. Beirut: Dār al-Kutub al-ʿIlmiyya, 2000.

Zarkaši, *Burhān* = Badr ad-Dīn ʾAbū ʿAbdallāh Muḥammad ibn ʿAbdallāh, *al-Burhān fī ʿulūm al-Qurʾān,* 4 vols. Ed. Muḥammad ʾAbū l-Faḍl ʾIbrāhīm. Cairo: Dār ʾIḥyāʾ al-Kutub al-ʿArabiyya, 1957–1958.

Secondary sources

ʿAbbās, ʾIḥsān. 1971. *Taʾrīx an-naqd al-ʾadabī ʿinda l-ʿArab: Naqd aš-šiʿr min al-qarn aṯ-ṯānī ḥattā l-qarn aṯ-ṯāmin al-hijrī*. Beirut: Dār aṯ-Ṯaqāfa.

Abdul Aleem. 1933. “Iʾjazuʾl-Quran”. *Islamic Culture* 7.64–82, 215–233.

Abu Deeb, Kamal. 1979. *al-Jurjānī's theory of poetic imagery*. Warminster: Aris and Phillips.

——. 1990. “Literary criticism”. *ʿAbbasid belles-lettres,* ed. Julia Ashtiany a.o., 339–387. Cambridge: Cambridge University Press.

Audebert, Claude-France. 1982. *al-Ḫaṭṭābī et l'inimitabilité du Coran: Traduction et introduction au* Bayān ʾiʿǧāz al-Qurʾān. Damascus: Institut Français de Damas.

Boullata, Issa J. 1988. “The rhetorical interpretation of the *Qurʾān*: *Iʿjāz* and related topics”. *Approaches to the history of the interpretation of the* Qurʾān, ed. Andrew Rippin, 139–157. Oxford: Clarendon Press.

Gelder, Geert Jan H. van. 1982. *Beyond the line: Classical Arabic literary critics on the coherence and unity of the poem*. Leiden: E.J. Brill.

Grunebaum, Gustave E. von. 1941. “Arabic literary criticism in the 10th century A.D.”. *Journal of the American Oriental Society* 61.51–57.

——. 1950. *A tenth-century document of Arab literary theory and criticism*. Chicago: University of Chicago Press.

Larkin, Margaret. 1988. “The inimitability of the *Quʾrān*: Two perspectives”. *Religion and Literature* 20.31–47.

——. 1995. *The theology of meaning: ʿAbd al-Qāhir al-Jurjānī's theory of discourse*. New Haven, Conn.: American Oriental Society.

Martin, Richard C. 1980. “The role of the Basrah Muʿtazilah in formulating the doctrine of the apologetic miracle”. *Journal of Near Eastern Studies* 39.175–189.

——. 2002. “Inimitability”. *Encyclopaedia of the Qurʾān,* ed. Jane Dammen McAuliffe a.o., II, 526–536. Leiden: E.J. Brill.

Mir, Mustansir. 1988. “The *Qurʾān* as literature”. *Religion and Literature* 20.49–64.

——. 1990. “Bāqillānī's critique of Imruʾ al-Qays”. *Studies in Near Eastern culture and history: In memory of E.T. Abdel-Massih,* ed. James A. Bellamy, 118–130. Ann Arbor: Center for Near Eastern and North African Studies, University of Michigan.

Rahman, Yusuf. 1996. “The miraculous nature of Muslim scripture: A study of ʿAbd al-Jabbār's *Iʿjāz al-Qurʾān*”. *Islamic Studies* 35.409–424.

Vasalou, Sophia. 2002. “The miraculous eloquence of the *Qurʾān*: General trajectories and individual approaches”. *Journal of Qurʾanic Studies* 4.23–53.

Weiss, Bernard G. 1992. *The search for God's law: Islamic jurisprudence in the writings of Sayf al-Din al-Amidi*. Salt Lake City: University of Utah Press.

Sophia Vasalou
(Gonville and Caius College)

ʾIlġāʾ

The term *ʾilġāʾ* lit. ‘nullification, annulment’ is used in Arabic grammar to denote the opposite of → *ʿamal* ‘government’. It is applied to sentences in which the expected government relations have been canceled. The term is related to *laġw* (synonyms *ḥašw* and *zāʾid*), which is used for redundant elements in the sentence (on the meaning of *laġw*, see Talmon 2003:222–223).

ʾIlġāʾ was analyzed by Carter (1973:156), who translates it as ‘neutralization’ and defines it as a process “by which elements are deprived of their operative effect”. He refers to Sībawayhi (*Kitāb* I, 243.6), who explains the fact that the comment in the sentence *fīhā ʿabdullāhi qāʾimun ġadan* “Abdallāh is in it standing tomorrow’ does not have the expected accusative by saying, “because the adverbials are annulled, so that it is as if the speaker did not mention them in this position” (*li-ʾanna ḏ̣-ḏ̣urūf tulġā ḥattā yakūna l-mutakallim* ka-ʾannahu lam yaḏkurhā fī hāḏā l-mawḍiʿ*). The term appears rather frequently in the *Kitāb* (according to Troupeau 1976:190, *ʾilġāʾ* occurs 9 times, the verb *ʾalġā* 25 times, and the passive participle *mulġā* ‘neutralized element’ 4 times).

The prototypical use of *ʾilġaʾ* is in cases of disrupted government in sentences like *zaydun ḏ̣anantu munṭaliqun* ‘Zayd, I believe, is leaving’ (Owens 1988:50–51), as against *zaydan ḏ̣anantu munṭaliqan*, in which the verb *ḏ̣anna* – one of the → *nawāsix* – governs the underlying topic/comment construction. In such a sentence, the topic and the comment are in principle governable, but the potential governor *ḏ̣anna* may be inserted as a canceling element that has no structural relations with the rest of the sentence. Another example is given by aš-Širbīnī (Carter 1981:114–115), *ʾiḏan ʾukrima-ka* ‘therefore, I shall honor you’, as against *ʾiḏan ʾana ʾukrimu-ka*; in the latter case, it is the intervening personal pronoun that prevents the governor *ʾiḏan* from governing.

Ibn as-Sarrāj (*ʾUṣūl* II, 257–261) has a separate chapter about the *ʾilġāʾ* in which he defines it as follows: “*ʾilġāʾ* is when you use a word that has no place in the declension, if it is something that [normally] causes declension; when it is deleted from speech, the sentence does not become faulty” (*al-ʾilġāʾ ʾinnamā huwa ʾan taʾtiya bi-kalima* [read so, instead of *l-kalima*] *lā mawḍiʿa lahā min al-ʾiʿrāb ʾin kānat mimmā tuʿribu wa-ʾinnahā matā ʾusqiṭat min al-kalām lam yaxtalli l-kalām* [read so, instead of *li-kalām*]; *ʾUṣūl* II, 286.2–3). This definition also explains the connection with the meaning of *laġw* denoting a redundant element in the sentence.

By extension, *ʾilġāʾ* is also used in cases of deviating constructions, for instance when the absolute negation *lā* is followed not by an accusative but by a nominative, e.g. *lā rajulun fī d-dāri wa-lā mraʾatun* ‘there is no man in the house, nor a woman’ (Carter 1981:414–415); or in exceptive sentences when *ʾillā* is not followed by an accusative, as in *mā qāma ʾillā zaydun* ‘no one stood up except Zayd’ (Carter 1981:400–401). In the latter example, *zaydun* functions as agent of *qāma* and the government of *ʾillā* is neutralized (the author refers to Q. 54/50 *wa-mā ʾamrunā ʾillā wāḥidatun* ‘and Our command is not but one’ for another example of this).

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Ibn as-Sarrāj, *ʾUṣūl* = ʾAbū Bakr ibn as-Sarī Ibn as-Sarrāj, *Kitāb al-ʾuṣūl fī n-naḥw*. Ed. ʿAbd al-Ḥusayn al-Fatlī. 3 vols. Beirut: Muʾassasat ar-Risāla, 1985.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. 2 vols. Bulaq, 1316 A.H. (Repr., Baghdad: Al-Muthanna Library, n.d.)

SECONDARY SOURCES

Carter, Michael G. 1973. “An Arab grammarian of the eighth century A.D.: A contribution to the history of linguistics”. *Journal of the American Oriental Society* 93.146–157.

——. 1981. *Arab linguistics: An introductory classical text with translation and notes.* Amsterdam: J. Benjamins.

Owens, Jonathan. 1988. *The foundations of grammar: An introduction to Medieval Arabic grammatical theory.* Amsterdam and Philadelphia: J. Benjamins.

Talmon, Rafael. 1997. *Arabic grammar in its formative age:* Kitāb al-ʿayn *and its attribution to Ḫalīl b. Aḥmad.* Leiden: E.J. Brill.

——. 2003. *Eighth-century Iraqi grammar: A critical exploration of pre-Halīlian Arabic linguistics.* Winona Lake, Ind.: Eisenbrauns.

Troupeau, Gérard. 1976. *Lexique-index du* Kitāb *de Sībawayhi.* Paris: Klincksieck.

KEES VERSTEEGH (University of Nijmegen)

ʿIlla

The term *ʿilla* lit. ‘cause; illness’ is used for two central notions in technical linguistic terminology. On the one hand, it means ‘(linguistic) cause’, and in this sense it is connected with such terms as *iʿtilāl* ‘argumentation’, *maʿlūl* ‘caused’. On the other hand, it is used to indicate the effect of the ‘weak consonants’; in this sense it is connected with such terms as *muʿtall* ‘weak’, *iʿtilāl* ‘weakness’.

The connection between the two central senses in which *ʿilla* is used is not immediately clear. According to Lane (1863–1893:V, 2124), *ʿilla* is “an accident that befalls an object and causes its state, or condition, to become altered”, hence “a disease that diverts [from the ordinary occupations]”, hence it may be “an accident, or event, that diverts the person to whom it occurs from his course” and even “a cause [and particularly an efficient cause]”. The verb *ʿalla/yaʿullu* means ‘to give someone to drink the second time’, hence ‘to divert someone from his want'; the verb *ʿalla/yaʿillu* means 'to become sick, to fall ill'. What is probably behind the double sense in which *ʿilla* is used is that in speech, accidents may 'befall' the sounds, so that the shape of the word is altered. These accidents may be regarded as the causes of linguistic change, and even of linguistic structure.

The phonetic sense of *ʿilla* is probably the older one. It occurs in al-Xalīl's *Kitāb al-ʿayn* (e.g. I, 59–60) to indicate the effect of the three → glides, /w/, /y/, and /ʾ/, which are related to the → *ḥarakāt* 'vowels'. They are often called *ḥurūf al-līn wa-l-madd* 'consonants of softness and lengthening', because they represent the lengthening of vowels. But at the same time, they are responsible for most of the phonetic changes words undergo, disguising the transparency of Arabic, which is why they are known as *ḥurūf al-ʿilal*. Words containing one of the glides are called *muʿtall* 'sick, ill', and words without any glide are called *ṣaḥīḥ* 'healthy, sound'. If the intended metaphor in characterizing the phonetic behavior of morphological forms is indeed one of 'illness/health', it may have been transmitted to Arabic grammar from Greek theories of linguistic pathology, through the intermediary of Syriac grammarians. In Greek grammar, the term *hugiēs* 'healthy' indicates a word that is not affected by any change, and *páthos* 'illness; accident; affection; suffering' is used for phonetic changes affecting the shape of the word (Versteegh 1977:26). This parallel is, however, rejected by Guillaume (n.d.:29–35), who points out that the *páthē* of Greek linguistics are used for all kinds of linguistic change, whereas in Arabic grammar *ʿilal* are closely connected with the presence of glides. Guillaume (n.d.:9–47) does not accept the sense of 'illness' for *ʿilla* and, in fact, rejects the traditional interpretation of its meaning, which in his view does not involve any polysemy. He proposes for *ʿilla* and the adjective *muʿtall* the interpretation of "une préoccupation ou une distraction secondaire qui vient détourner quelqu'un de son occupation principale" (Guillaume n.d.:36); in linguistics, this translates as a transfer from an underlying, base form (*ʾaṣl*) to a secondary form. The *ḥurūf al-ʿilla* are those elements that effect this transfer, and they have nothing to do with the notion of 'illness'. The use of *ṣaḥīḥ* 'sound' to indicate the opposite of *muʿtall* is in this view secondary, inspired by the other meanings of *ʿilla* in everyday speech.

Guillaume is certainly right when he says that for the grammarians there was no ambiguity in the use of the term *ʿilla*: for them, 'cause' and 'accident befalling a word' were identical notions. It is also true that *ʿilla* has a connotation of 'deviation' from the rules that has to be explained. As a matter of fact, Ibn Jinnī (*Munṣif* II, 262; Méhiri 1973:187) states that *ʾiʿlāl* should be avoided as much as possible, and Arabic grammarians always take great care to posit as few *ʿilal* as possible. In some cases, however, they are the lesser evil, when no other explanation can be found.

In Sībawayhi's *Kitāb*, according to Troupeau (1986:144–145), *ʿilla* is used as a methodological term, translated by him with 'raison', 23 times, and as a morphological term, translated with 'débilité', 11 times; the related term *iʿtilāl* always means 'débilité' (50 times), and other related terms, like *ʿalla* (2 times), *iʿtalla* (97 times), and *muʿtall* (95 times) are all connected with the morphological sense. In the sense of 'cause', *ʿilla* is used in the *Kitāb* in a nontechnical way to indicate general causes in linguistic argumentation, for instance when a structure like *yā zaydu* is said to have a 'reason' that will be dealt with in another chapter (*Kitāb* I, 147.9). In discussing the declension of the imperfect, Sībawayhi (*Kitāb* I, 409.12) combines *ʿilla* with *sabab* 'reason' when he says that the verb has the nominative because it is used in a syntactic slot (*mawḍiʿ*) of the noun (*wa-hiya sabab duxūl ar-rafʿ fīhā wa-ʿillatuhā*); the same combination is used for the cause of a physical phenomenon (*Kitāb* I, 430.16). Elsewhere, *ʿilla* indicates the cause of a phonetic change, for instance when Sībawayhi compares the relational adjective of *sana* 'year', *sanawiyy*, with that of *ʾab* 'father', *ʾabawiyy*, and states that the 'cause' (*ʿilla*) of this change is the same in both words (*Kitāb* II, 80.19).

In 4th/10th-century grammar, *ʿilla* acquired a new, crucial meaning when it began to be used in discussions about the epistemological status of linguistic arguments. Ibn as-Sarrāj (d. 316/928), in a famous passage (*ʾUṣūl* I, 35), states that there are two kinds of linguistic argumentation (*iʿtilālāt*). The first kind consists of grammatical rules, such as 'all agents are in the nominative'. The second kind is called by him the *ʿillat al-ʿilla* 'the cause of the cause'. This type of reasoning provides an explanation of the linguistic rules and explains, for instance, why an agent has to have a nominative (Guillaume n.d.:94–106).

In his discussion of the dichotomy, Ibn Jinnī (d. 392/1002) shows that ultimately it leads to an argument ad infinitum (*Xaṣāʾiṣ* I, 173–174; cf. Méhiri 1973:139–149). He himself contributed to the discussion by comparing grammatical *ʿilal* to those used in theology and legal theory (*Xaṣāʾiṣ* I, 48–95). True to his Muʿtazilite leanings, he concludes that like the theologians, linguists look for the underlying rational 'causes', whereas legal scholars work with rules that have once been instituted and cannot be explained rationally (Suleiman 1999: 64–108).

The most important contribution to the debate is found in az-Zajjājī's (d. 339/949) *ʾĪḍāḥ fī ʿilal an-naḥw* (The elucidation of the causes of grammar), which devotes an entire chapter to the nature of linguistic causes. Az-Zajjājī attempted to get out of the vicious circle by introducing three levels of linguistic reasoning. The first two levels, the *ʿilal taʿlīmiyya* 'causes connected with acquisition', and the *ʿilal qiyāsiyya* 'causes connected with analogy', correspond to Ibn as-Sarrāj's *ʿilal* and *ʿilal al-ʿilal*. But with the third level, that of the *ʿilal naḏ̣ariyya wa-jadaliyya* 'speculative causes' (Versteegh 1995:90–92; Guillaume n.d.:144–158; Suleiman 1999:43–63), he attempted to legitimize linguistic argumentation by appealing to causes outside the system of grammar: as a true rationalist (like most 4th–10th-century grammarians, he was a Muʿtazilite), he believed that without such outside arguments the entire reasoning would collapse. Therefore, the third level of the *ʿilal naḏ̣ariyya wa-jadaliyya* seeks arguments in objective facts, such as natural phenomena, the wisdom of the Arabs, the nature of sound, the hierarchy scale of sonority, or the laws of logic.

In the later systematization of linguistic argumentation along the lines of legal theorizing, *ʿilla* came to mean the compelling cause that constitutes the basis for a sound analogy (→ *qiyās*). This is what Ibn al-ʾAnbārī (d. 577/1181) holds in his *Lumaʿ al-ʾadilla* when he says (*Lumaʿ* 42.10–11; cf. Suleiman 1999:109–144) that each valid analogy (*qiyās*) consists of a basic rule (*ʾaṣl*), a derived rule (*farʿ*), a compelling cause (*ʿilla*), and a conclusion (*ḥukm*). This became the most obvious sense in which *ʿilla* was used in later grammar, always along with the standard terminology of the *ḥurūf al-ʿilla*. The notion of a 'compelling cause' is precisely what led to the rejection of this concept by some theorists (Ibn al-ʾAnbārī, *Lumaʿ* 44–48, 'refutation of those who reject reasoning by analogy'; cf. Arnaldez 1956:165–193), and in particular, to the rejection of the entire structure of linguistic reasoning by scholars like the Ẓāhirī grammarian Ibn Maḍāʾ (d. 592/1196) in his *Kitāb ar-radd ʿalā n-nuḥāt* (cf. Arnaldez 1956:89–97; Wolfe 1984; Suleiman 1999:145–177).

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Ibn al-ʾAnbārī, *Lumaʿ* = ʾAbū l-Barakāt ʿAbd ar-Raḥmān ibn Muḥammad al-ʾAnbārī, *Lumaʿ al-ʾadilla fī ʾuṣūl an-naḥw*. Ed. Hadi M. Kechrida. Stockholm: Almqvist and Wiksell, 1963.

Ibn Jinnī, *Munṣif* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Munṣif šarḥ li-Kitāb at-taṣrīf li-l-Māzinī*. Ed. ʾIbrāhīm Muṣṭafā and ʿAbdallāh ʾAmīn. 3 vols. Cairo: Muṣṭafā al-Bābī al-Ḥalabī, 1954.

——. *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. Ed. Muḥammad ʿAlī an-Najjār. 3 vols. Cairo, 1952–1956. (Repr., Beirut: Dār al-Hudā, n.d.)

Ibn Maḍāʾ, *Radd* = ʾAbū l-ʿAbbās ʾAḥmad Ibn Maḍāʾ al-Qurṭubī, *Kitāb ar-radd ʿalā n-nuḥāt*. Ed. Šawqī Ḍayf. Cairo: Dār al-Fikr al-ʿArabī, 1947.

Ibn as-Sarrāj, *ʾUṣūl* = ʾAbū Bakr ibn as-Sarī Ibn as-Sarrāj, *Kitāb al-ʾuṣūl fī n-naḥw*. Ed. ʿAbd al-Ḥusayn al-Fatlī. 3 vols. Beirut: Muʾassasat ar-Risāla, 1985.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. 2 vols. Bulaq, 1316 AH. (Repr., Baghdad: Al-Muthanna Library, n.d.)

Xalīl, *ʿAyn* = ʾAbū ʿAbd ar-Raḥmān al-Xalīl ibn ʾAḥmad al-Farāhīdī, *Kitāb al-ʿayn*. Ed. Mahdī al-Maxzūmī and ʾIbrāhīm as-Sāmarrāʾī. 8 vols. Beirut: Muʾassasat al-ʾAʿlamī li-l-Maṭbūʿāt, 1988.

Zajjājī, *ʾĪḍāḥ* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān Ibn ʾIsḥāq az-Zajjājī, *al-ʾĪḍāḥ fī ʿilal an-naḥw*. Ed. Māzin al-Mubārak. Cairo: Dār al-ʿUrūba, 1959.

SECONDARY SOURCES

Arnaldez, Roger. 1956. *Grammaire et théologie chez Ibn Hazm de Cordoue: Essai sur la structure et les conditions de la pensée musulmane*. Paris: J. Vrin.

Guillaume, Jean-Patrick. n.d. *La 'cause' des grammairiens: Etude sur la notion de* ʿilla *dans la tradition grammaticale arabe (fin IIIè/IXè – milieu du IV/X s.)*. Thèse de 3ème cycle, Université de Paris-3.

Lane, Edward William. 1863–1893. *An Arabic-English lexicon derived from the best and most copious Eastern sources*. 8 vols. London and Edinburgh: Williams and Norgate.

Méhiri, Abdelkader. 1973. *Les théories grammaticales d'Ibn Jinnī*. Tunis: Université de Tunis.

Suleiman, Yasir. 1999. *The Arabic grammatical tradition: A study in* taʿlīl. Edinburgh: Edinburgh University Press.

Troupeau, Gérard. 1976. *Lexique-index du* Kitāb *de Sībawayhi*. Paris: Klincksieck.

Versteegh, C.H.M. (Kees). 1977. *Greek elements in Arabic linguistic thinking*. Leiden: E.J. Brill.

——. 1995. *The explanation of linguistic causes: Az-Zağğāğī's theory of grammar*. Amsterdam and Philadelphia: J. Benjamins.
Wolfe, Ronald G. 1984. *Ibn Maḍāʾ al-Qurṭubī and the* Book in refutation of the grammarians. Ph.D. diss., Indiana University.

Kees Versteegh
(University of Nijmegen)

ʾImāla

The term *ʾimāla* is used by the medieval Arab grammarians to denote the fronting and raising of Old Arabic *ā* toward *ī*, and the old short *a* toward *i* (Levin 1992:74, esp. n. 1). Although the term *ʾimāla* denotes the fronting and raising of both *ā* and *a*, the ancient Arabic sources almost completely ignore the *ʾimāla* of short *a*, and the term in these texts usually denotes the *ʾimāla* of long *ā* (Levin 1992:74). The few grammarians who mention the *ʾimāla* of short *a* deal with it only when it occurs in the proximity of *r* (Sībawayhi, *Kitāb* II, 293.6–294.5; Zamaxšarī, *Mufaṣṣal* 160.10–11; Ibn Yaʿīš, *Šarḥ* II, 1265.19–1266.7; Ibn ʿAqīl, *Šarḥ* II, 527.6–528.4). Literally, *ʾimāla* means 'inclination'. It was known by this name because the person who pronounces *ā* and *a* with *ʾimāla* 'inclines' the *ʾalif* in the direction of *yāʾ* and the *fatḥa* in the direction of *kasra* (Levin 1992:74–75, esp. n. 3).

The Arab grammarians saw the *ʾimāla* as a phenomenon close to that known today as 'vowel harmony'. In their view, *ʾimāla* and → *ʾidġām* 'assimilation of consonants' show much similarity (Sībawayhi, *Kitāb* II, 279.11–16; Ibn Jinnī, *Xaṣāʾiṣ* II, 141.16; Zamaxšarī, *Mufaṣṣal* 158.11–13; Ibn Yaʿīš, *Šarḥ* II, 1252.22–1253.14).

In Sībawayhi's (d. 180/796) view, the occurrence of *ʾimāla* includes both *ā* and the preceding consonant (Sībawayhi, *Kitāb* II, 284.8–11, 292.13–14, 293.6–294.2). This suggests that he noticed that the phenomenon of the fronting of *ā* is combined with the fronting of the point of articulation of the preceding consonant (→ vowel raising).

The grammarians regarded *ʾimāla* as a legitimate phenomenon from the normative point of view when it occurred in certain conditionings, but cases deviating from these conditionings were regarded as errors (Levin 1992:75, n. 5). In the *Kitāb, al-ʾalifu llatī tumālu ʾimālatan šadīdatan* 'the long *ā* which is pronounced with strong *ʾimāla*' is counted among the legitimate sounds of Arabic, which may be used in reading the *Qurʾān* and poetry (Sībawayhi, *Kitāb* II, 452.8–10).

Sībawayhi, al-Mubarrad, and Ibn as-Sarrāj do not say anything about the quality of the vowel of the *ʾimāla*. It is therefore impossible to get any idea of this quality in the 8th and 9th centuries and to judge whether the vowel of the *ʾimāla* was closer to *ē* or *ī*. According to as-Sīrāfī and Ibn Jinnī (10th century), the vowel of the *ʾimāla* was pronounced somewhere between *ā* and *ī* (Sīrāfī, *Šarḥ* I, 306.3–4; Ibn Jinnī, *Sirr* I, 50.1–2; Ibn Jinnī, *Xaṣāʾiṣ* III, 120.10–13; Ibn Jinnī, *Munṣif* I, 42.11–17). Similar information is given by Ibn Yaʿīš (*Šarḥ* II, 1252.12–1253.1). This suggests that the vowel of the *ʾimāla*, at least in most dialects, was *ē*.

ʾImāla occurred in Old Arabic (Levin 1971: I, 9–61), but it was not a general phenomenon, occurring only in some of the old dialects (Sībawayhi, *Kitāb* II, 284.1–6). These dialects sometimes also differed from each other with respect to *ʾimāla*, since the *ā* shift was not homogeneous in all of them (Sībawayhi, *Kitāb* II, 284.1–6, 284.11–13). It also occurred in some Middle Arabic dialects (Levin 1971:I, 62–73; Levin 1975:261–265) and in many modern Arabic dialects (Levin 1971:I, 79–412).

ʾImāla and the factors conditioning its occurrence were described for the first time by Sībawayhi. His description of *ʾimāla* is based on the linguistic situation prevailing in his time and environment, mainly al-Baṣra and its surroundings in southern Iraq. This is confirmed by evidence in the *Kitāb* (Sībawayhi, II, 294.4–5; Levin 1992, n. 14). The description of *ʾimāla* by all later grammarians is based on that of Sībawayhi.

ʾImāla occurred both in medial and final position. The factors conditioning medial *ʾimāla* differed from those of final *ʾimāla*.

According to Sībawayhi, there were three types of medial *ʾimāla* in the Arabic spoken in Iraq in the 8th century:

i. Medial *ʾimāla*, conditioned by the vocalic environment of medial *ā*. This *ʾimāla* occurs when the vowel of the syllable adjacent to *ā* is *i* or *ī*, e.g. *ʿimạ̄d* 'pillar' (Sībawayhi, *Kitāb* II, 279.17), *kilạ̄b* 'dogs' (Sībawayhi, *Kitāb* II, 279.21), *šimlạ̄l* 'brisk camel' (Sībawayhi, *Kitāb* II, 279.21), *ʿā̩bid* ‘worshipper’ (Sībawayhi, *Kitāb* II, 279.11), *ʿā̩lim* ‘one who knows’ (Sībawayhi, *Kitāb* II, 279.11), *mafā̩tīḥ* ‘keys’ (Sībawayhi, *Kitāb* II, 279.11). Note that *ʾimāla* occurs irrespective of whether *i* or *ī* follows or precedes *ā* (Sībawayhi, *Kitāb* II, 279.10–21).

With some speakers, *ʾimāla* also occurs when *ā* is preceded by *y* (Sībawayhi, *Kitāb* II, 281.19–282.3), e.g. *bayyā̩ʿ* ‘seller’ (Sībawayhi, *Kitāb* II, 281.20), *kayyā̩l* ‘grain measurer’ (Sībawayhi, *Kitāb* II, 281.20), *šaybā̩n* [proper name] (Sībawayhi, *Kitāb* II, 282.2).

This type of *ʾimāla*, conditioned by the occurrence of *i* or *ī* in the syllable adjacent to *ā*, does not occur if one of the *ḥurūf al-mustaʿliya*, i.e. one of the emphatic consonants *ṣ, ḍ, ṭ, ḏ̣* or one of the back consonants *q, ġ, x* is adjacent to *ā* (Sībawayhi, *Kitāb* II, 285.17–286.6), e.g. *qāʿid* ‘sitting’ (Sībawayhi, *Kitāb* II, 285.19), *ġāʾib* ‘absent’ (Sībawayhi, *Kitāb* II, 285.19), *ṭāʾif* ‘walking around something’ (Sībawayhi, *Kitāb* II, 285.19), *ʿāṭis* ‘sneezing’ (Sībawayhi, *Kitāb* II, 286.5). However, Sībawayhi mentions that *ʾimāla* does occur in the immediate proximity of these consonants in the dialect (or dialects) of certain people whose speech cannot be considered as an example of good Arabic (Sībawayhi, *Kitāb* II, 285.17–286.6).

Sībawayhi also mentions the exceptions *an-nā̩s* ‘the people’ (Sībawayhi, *Kitāb* II, 285.7) and the proper name *al-ḥajjā̩j* (Sībawayhi, *Kitāb* II, 285.8), where *ʾimāla* occurs although no *i* or *ī* is found in the proximity of *ā* (Sībawayhi, *Kitāb* II, 285.5–8).

ii. Medial *ʾimāla* occurring in the 3rd person masculine singular perfect of II*w/y* verbs, as in *ṭā̩ba* ‘he was good’ (Sībawayhi, *Kitāb* II, 281.13), *xā̩fa* ‘he was afraid’ (Sībawayhi, *Kitāb* II, 281.13), *hā̩ba* ‘he feared’ (Sībawayhi, *Kitāb* II, 281.13). This *ʾimāla* occurs only in verbs where the first vowel in the 1st and 2nd person is *i*, as in *ṭibtu* and *xiftu*. It is only found in the speech of some of the people of al-Ḥijāz (Sībawayhi, *Kitāb* II, 281.10–14) and is not conditioned by the vocalic or consonantal environment.

Sarauw (1908:35) was the first to notice that Sībawayhi distinguishes between these two types of *ʾimāla*. In his view, *ʾimāla* conditioned by the vocalic environment occurred in the dialects of the eastern tribes Tamīm, Qays, and ʾAsad, although Sībawayhi does not explicitly say so (Sarauw 1908:33). However, it can be inferred from his remarks (Sībawayhi, *Kitāb* II, 36.20–37.21) that this *ʾimāla* occurred in the dialect of Tamīm (Levin 1971:II, 422–424, n. 10).

iii. The third type of medial *ʾimāla*, occurring in nouns of II*w/y* roots, is conditioned by the consonantal environment of *ā*. It occurs only in the proximity of front consonants, not in that of the *ḥurūf al-mustaʿliya* (Sībawayhi, *Kitāb* II, 285.10–16, 289.15–18), e.g., *bā̩b* ‘door’ (Sībawayhi, *Kitāb* II, 285.10), *mā̩l* ‘flocks’ (Sībawayhi, *Kitāb* II, 285.10), and *ʿā̩b* ‘disgrace’ (Sībawayhi, *Kitāb* II, 285.10) are pronounced with *ʾimāla*, as against *sāq* ‘leg’ (Sībawayhi, *Kitāb* II, 289.16), *qār* ‘large herd of camels’ (Sībawayhi, *Kitāb* II, 289.16), *ġāb* ‘forest’ (Sībawayhi, *Kitāb* II, 289.16). Sībawayhi does not say in which dialect this particular *ʾimāla* occurred.

Final *ʾimāla* occurred only in some of the Iraqi dialects in the 8th century (Sībawayhi, *Kitāb* II, 281.4–5, 281.9–10). It was not usually conditioned by the presence of *i* or *ī* in the syllable preceding final *ā*, and it was not precluded in the immediate proximity of the *ḥurūf al-mustaʿliya* (Sībawayhi, *Kitāb* II, 287.15–20), e.g. *ḥublā̩* ‘pregnant’ (Sībawayhi, *Kitāb* II, 281.6), *muʿṭā̩* ‘given’ (Sībawayhi, *Kitāb* II 287.18), *saqā̩* ‘he gave to drink’ (Sībawayhi, *Kitāb* II, 287.19), *daʿā̩* ‘he called’ (Sībawayhi, *Kitāb* II, 280.18). The pausal form of the feminine ending *-at-* was also affected by it, as in *ḍarbā̩h* ‘one act of striking’ (Sībawayhi, *Kitāb* II, 292.14) and *ʾaxḏā̩h* ‘one act of taking’ (Sībawayhi, *Kitāb* II, 292.14). As-Sīrāfī, Sībawayhi’s commentator (10th century), states that the *ʾimāla* of this ending was widespread in al-Baṣra, al-Kūfa, and Mosul and their surroundings (Sīrāfī, *Šarḥ* I, 349.3).

In most cases there was no difference between the behavior of the endings ا and ى. With some speakers, *ʾimāla* of final *ā* in certain endings occurred in pause. The endings Sībawayhi mentions in this respect are the accusative and genitive pronoun suffix of the 3rd person singular feminine *-hā*, and the genitive pronoun suffix of the 1st person plural *-nā* (Sībawayhi, *Kitāb* II, 284.16–21).

ʾImāla occurred also in some Middle Arabic dialects, and the phenomenon is also known today in many modern Arabic dialects, both in medial and final position. The medial *ʾimāla* vowel in most modern dialects is *ē*, but in some of them it is *ī* (Levin 1971:I, 387–388; also for the occurrence of diphthongization after the *ʾimāla* vowel in Cyrenaica and usually in Malta). The final *ʾimāla* vowel is in most cases a short, unstressed vowel. In some of the dialects the vowel is *i*, in others *e*, and in some of them the vowel is between *i* and *e* (Levin 1971:388–390; also for other variants).

The data gathered from some Middle Arabic and modern Arabic dialects shows that Sībawayhi's description of the *ʾimāla* is authentic and accurate (Levin 1992:80–88). This inference is mainly based on a discovery by Blanc (1964:42–49) that in the modern *qəltu* dialects of Iraq and Anatolia and in the modern dialect of Aleppo, the factors conditioning the medial *ʾimāla* correspond to those described by Sībawayhi in the 8th century. The same type of medial *ʾimāla* also occurs today in Turkey, in the dialects of the districts of Alexandretta and → Antiochia (Arnold 1996:4), and in the dialects of the districts of Adana and Mersīn (Procházka 1996:192). Medial *ʾimāla* in these modern dialects is conditioned by the historical vocalic environment: *ʾimāla* usually occurs when the historical vowel of the syllable adjacent to *ā* was *i* or *ī*, e.g. *klīb* 'dogs' in Jewish Baghdadi, *klēb* in Christian Baghdadi, Mosul, Anatolia, and Aleppo; *uḥēm* (<**wiḥām*) 'carving during pregnancy' in Aleppo (Barthélemy 1935–1954:887); *jīməʿ* 'mosque' in Jewish Baghdadi, *jēməʿ* in Christian Baghdadi and in Mosul and Anatolia; *mafētīḥ* 'keys' in the Jewish dialect of Mosul, *mfētīḥ* in Aleppo (Levin 1994:219); *sakīkīn* 'knives' in the Muslim dialect of Mosul, *sakēkīn* in the Jewish dialect of Mosul (Jastrow 1989:285). These examples show that in these modern dialects, medial *ʾimāla* occurs in the proximity of old *i*, even if it has dropped out or changed. The historical character of the factors conditioning medial *ʾimāla* in these dialects is also shown by the fact that it does not occur in the proximity of *ə* < *a* or *ə* < *u*, as in the examples *xəbbāz* 'baker' and *səkkān* 'inhabitants' in Jewish Baghdadi.

Evidence supporting the authenticity of Sībawayhi's description of *ʾimāla* is also furnished by texts in Middle Arabic dialects. Recently, Blau and Hopkins noticed in unpublished manuscripts of early Judaeo-Arabic texts from the 9th and 10th centuries the occurrence of medial *ʾimāla* under conditions very similar to those described by Sībawayhi (Hopkins 2005:195, 211). The occurrence of the type of medial *ʾimāla* described by Sībawayhi is also attested by examples found in the text of the vernacular poetry of Ṣafiyy ad-Dīn al-Ḥillī in the 14th century (Levin 1975:261–264).

The inference that Sībawayhi's description of *ʾimāla* is authentic and accurate is also supported by his information that some people pronounce the form *an-nās* 'the people' with *ʾimāla*, as an exception to the usual conditioning factors. The same exception is found today in some of the *qəltu* dialects: the form *nēs* 'people' occurs in Christian Baghdadi and in the dialects of northern Iraq and Anatolia (Levin 1994:220).

The above inference is also confirmed by Sībawayhi's description of the final *ʾimāla*: in comparing final *ʾimāla* as described by him with that occurring today in some modern dialects, it can be inferred that the situation in Sībawayhi's period was, in general, similar to that prevailing in the modern *qəltu* dialects and in the dialect of Aleppo (Levin 1992:86–88): in the *qəltu* dialects of Iraq and Anatolia and in Aleppo, just as in Sībawayhi's period, final *ʾimāla* is not conditioned by the existence of historical *i* or *ī* in the syllable preceding final *ā*, e.g. *ḥəblē* 'pregnant' in Jewish and Christian Baghdadi and in the Jewish dialect of Mosul (Blanc 1964:150), and *ḥəble* in Aleppo (Barthélemy 1935–1954:144). This example corresponds to *ḥublā* in Sībawayhi's period (see above). Other examples are *aʿmi* 'blind' in Jewish Baghdadi (Blanc 1964:84), *skāġi* 'drunkards' in Jewish Baghdadi, and *sakāġi* in Christian Baghdadi and in the Muslim dialect of Mosul (Blanc 1964:82). One of the most striking points of resemblance between final *ʾimāla* in some old and some modern dialects is the behavior of the final ى: in some dialects in Sībawayhi's time, the *ʾimāla* of this *ā* occurred only in nouns and adjectives, and not in verbs (Sībawayhi, *Kitāb* II, 284.13–16). In the modern *qəltu* dialects and in Aleppo the situation is exactly the same, as illustrated by the examples *ḥəblē* and *aʿmi* vs. *banā* 'he built' (Levin 1992:87–88).

There is, however, another type of medial *ʾimāla*, not mentioned by Sībawayhi, which occurs today in many modern dialects outside Iraq. This medial *ʾimāla* is conditioned by the consonantal environment of *ā*, and it occurs in many Lebanese dialects, in the Druze dialects of the Ḥōrān and the Golan, in the dialects of the Syrian desert oases Qarītēn, Palmyra, and Suxne, in the Bedouin dialects of Sāḥil Maryūṭ in Egypt, and in the Jabali dialect of Cyrenaica (Levin 1971:I, 186–234). In the dialects of Malta, medial *ʾimāla* occurs in any consonantal or vocalic environment, but sometimes it is conditioned by the historical consonantal environment of old *ā* (Levin 1971:I, 295–302).

The factors conditioning the occurrence of the medial *ʾimāla* in the above-mentioned dialects, except for those of Malta, can be summarized as follows:

i. *ʾImāla* usually occurs when the old *ā* is surrounded by two front non-emphatic consonants or by *h* and a front non-emphatic consonant (the furthest back consonant to be considered front according to this definition being *k*), e.g. *kēn* 'he was' in Beirut; *bēb* 'door' in Sāḥil Maryūṭ; *ljēm* 'bit' in Beirut; *čuhhēl* 'youngsters' in Palmyra; *člīb* 'dogs' in Suxne (Levin 1971:I, 209).

ii. *ʾImāla* is precluded when one of the historical emphatic consonants *ṣ, ḍ, ṭ, ḏ, ẓ* occurs in the immediate proximity of the old *ā*, e.g. *ḥīṭān* 'walls' and *ḥṣān* 'horse' in Palmyra; *naḏāfi* 'cleanliness' in Beirut; *aʿḏām* 'bones' in Palmyra (Levin 1971:I, 209).

iii. *ʾImāla* is frequently precluded in words containing an emphatic consonant, even when it does not occur in the immediate proximity of old *ā*. Usually this phenomenon occurs when one of the labials *b, m, w, f* or the lateral consonant *l* is adjacent to the old *ā*. For example: *ṭwāl* 'high [pl.]' in Beirut; *ṣəbāya* 'women' in Cyrenaica (Levin 1971: I, 210).

iv. When one of the back consonants *q, g < q, ʾ < q, x, ḥ, ʿ, ʾ* or one of the consonants *w, r, k* occurs in the immediate proximity of old *ā*, *ʾimāla* occurs in some cases but not in others. The factors conditioning the occurrence of *ʾimāla* in this position are very complicated, and there are many differences between the various dialects in this respect (Levin 1971:210, 212–233).

In the modern dialects where medial *ʾimāla* is conditioned by the consonantal environment, final *ʾimāla* is also known. This final *ʾimāla* is conditioned by the consonantal environment of old *ā* (Levin 1971:I, 235–251). This *ʾimāla* also occurs in dialects where medial *ʾimāla* does not occur at all (Levin 1971:I, 304–326). *ʾImāla* of the feminine ending *-at-* occurs in some dialects where final *ʾimāla* of other endings usually does not occur (Levin 1971:I, 326–327, 341–351). In the dialects of Malta, final *ʾimāla* usually does not occur (Levin 1971:I, 302). Final *ʾimāla* occurs only in pause in the dialects of Upper Egypt (Levin 1971:I, 371–376) and in the Lebanese dialect of Šḥīm (Levin 1971:I, 262–263, 273–278).

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Ibn ʿAqīl, *Šarḥ* = Bahāʾ ad-Dīn ʿAbdallāh Ibn ʿAqīl, *Šarḥ ʾAlfiyyat Ibn Mālik*. Ed. Muḥyī d-Dīn ʿAbd al-Ḥamīd. 14th ed. Cairo: Maṭbaʿat as-Saʿāda, 1384 A.H.

Ibn Jinnī, *Munṣif* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Munṣif, šarḥ li-Kitāb at-taṣrīf li-ʾAbī ʿUṯmān al-Māzinī*. Ed. ʾIbrāhīm Muṣṭafā and ʿAbdallāh ʾAmīn. 3 vols. Cairo: ʾIdārat ʾIḥyāʾ at-Turāṯ al-Qadīm, 1954–1960.

——, *Sirr* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *Sirr ṣināʿat al-ʾiʿrāb*. Ed. Ḥasan Hindāwī. 2 vols. Damascus: Dār al-Qalam, 1985.

——, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. Ed. Muḥammad ʿAlī an-Najjār. Beirut: Dār al-Hudā, n.d.

Ibn Yaʿīš, *Šarḥ* = Muwaffaq ad-Dīn ʾAbū al-Baqāʾ Yaʿīš Ibn Yaʿīš, *Šarḥ al-Mufaṣṣal*. Ed. Gustav Jahn. 2 vols. Leipzig, 1882–1886.

Ṣafiyy ad-Dīn al-Ḥillī = Hoenerbach, Wilhelm, *Die vulgärarabische Poetik Al-Kitāb Al-ʿĀṭil Al-Ḥālī Wal-Muraḫḫaṣ Al-Ġālī des Ṣafīyaddīn Ḥillī*. Wiesbaden: F. Steiner, 1956.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. Ed. Hartwig Derenbourg. 2 vols. Paris: Imprimerie Nationale, 1881–1889.

Sīrāfī, *Šarḥ* = ʾAbū Saʿīd al-Ḥasan ibn ʿAbdallāh as-Sīrāfī, *Šarḥ Kitāb Sībawayhi*. Ed. ʿAbd al-Munʿim Fāʾiz Masʿad, *as-Sīrāfī an-Naḥwī fī ḍawʾ šarḥihi li-Kitāb Sībawayhi*. 2 vols. N.p., n.d. [Includes as-Sīrāfī's commentary on Chaps. 432–512 of the *Kitāb*.]

Zamaxšarī, *Šarḥ* = ʾAbū l-Qāsim Maḥmūd ibn ʿUmar az-Zamaxšarī, *al-Mufaṣṣal fī n-naḥw*. Ed. Jens Peter Broch. 2nd ed. Christianiae: Libraria P.T. Mallingii, 1879.

SECONDARY SOURCES

Arnold, Werner. 1996. "Arabian dialects in the Turkish province of Hatay". Cremona a.o. (1996:1–10).

Barthélemy, A. 1935–1954. *Dictionnaire arabe-français, dialectes de Syrie: Alep, Damas, Liban, Jerusalem*. Paris: Paul Geuthner.

Blanc, Haim. 1964. *Communal dialects in Baghdad*. Cambridge, Mass.: Harvard University Press.

Cremona, Joe, Clive Holes, and Geoffrey Khan (eds.). 1996. *Proceedings of the 2nd International*

Conference of l'Association Internationale pour la Dialectologie Arabe held in Trinity Hall in the University of Cambridge, 10–14 September 1995. Cambridge: University of Cambridge.

Hopkins, Simon. 2005. "On *imāla* of medial and final *ā* in Early Judaeo-Arabic". *Sacrum Arabico-Semiticum: Homenaje al profesor Federico Corriente en su 65 aniversario*, ed. Jordi Aguadé, Ángeles Vicente and Leila Abu-Shams, 195–214. Zaragoza: Navarro y Navarro.

Jastrow, Otto. 1978. *Die mesopotamisch-arabischen Qəltu-Dialekte*. I. *Phonologie und Morphologie*. Wiesbaden: F. Steiner.

——. 1989. "Notes on Jewish Maṣlāwī". *Jerusalem Studies in Arabic and Islam* 12.282–293.

Levin, Aryeh. 1971. *The* ʾimāla *in the Arabic dialects*. 2 vols. Ph.D. diss., Hebrew University of Jerusalem. [In Hebrew; English summary, XIII–LIV.]

——. 1975. "The vernacular poetry of Ṣafiyy al-Dīn al-Ḥillī, a source for the reconstruction of Iraqi Arabic in the 14th century". *Israel Oriental Studies* 5.259–276.

——. 1978. "The *ʾimāla* of *ʾalif fāʿil* in Old Arabic". *Israel Oriental Studies* 8.174–203.

——. 1992. "The authenticity of Sībawayhi's description of the *ʾimāla*". *Jerusalem Studies in Arabic and Islam* 15.74–93.

——. 1994. "Sībawayhi's attitude to the spoken language". *Jerusalem Studies in Arabic and Islam* 17.204–243.

Procházka, Stefan. 1996. "The Arabic dialects of Cilicia (southern Turkey)". Cremona a.o. (1996: 189–196).

Sarauw, Christian. 1908. "Die altarabische Dialektspaltung". *Zeitschrift für Assyriologie und Verwandte Gebiete* 21.31–49.

Aryeh Levin
(The Hebrew University of Jerusalem)

Impairment → Language Impairment

Imperative → Mood

Imperfect → Tense

Impersonal Verb

1. Introduction

According to a classic definition, "An impersonal verb is a verb that occurs only in third person singular forms, has no specified agent, and has a dummy subject or no subject" (www.sil.org/linguistics/GlossaryOfLinguisticTerms/WhatIsAnImpersonalVerb.htm). The notion of 'impersonal verb' seems to have been unknown to Arab grammarians, who assumed that all verbs have either a 'real' subject (→ *fāʿil*) or a 'substitute for the subject' (*nāʾib fāʿil*). Blachère (1975:266) highlights this ambiguity as follows: "L'appellation de *verbe unipersonnel* paraît devoir être préférée à celle de *verbe impersonnel*". There exists, therefore, a difference between unipersonal verbs, which occur only in 3rd person singular forms whose subject has a referent, and impersonal (or not-personal) verbs, whose subject has no referent. Badawi a.o. (2004:394) provide an example of such an ambiguity when they refer to "impersonal verbs, particularly those which have a subordinate clause as their agent, 'it is necessary that', etc.". They quote, among other examples: *yajibu ʾan natawaqqafa ʿinda-hum qalīlan* 'we must pause with them a little', where the verb *yajibu* has the subordinate clause as its subject. Although this kind of verb is now perceived as impersonal, it has not always been treated as such, as shown in Ibn Jubayr (*Riḥla* 199): *hāʾulāʾi mimman inqaṭaʿa ʾilā llāh fa-tajibu mušārakatu-hum* 'they are among those who dedicated themselves to God, therefore, it is necessary to help them'. The feminine *tajibu* proves that this verb, undoubtedly unipersonal, is neither impersonal nor 'non-personal' and has a real subject (*mušārakatu-hum*).

2. Unipersonal verbs

Western authors agree that "our impersonal actives indicating natural phenomena, such as *it snows*, *it rains*, etc., are always expressed by the Arabs personally. They say either *ṯalaja ṯ-ṯalju* 'the snow snows', *maṭara l-maṭaru* 'the rain rains', or *ṯalajat as-samāʾu* 'the sky snows', *maṭarat as-samāʾu* 'the sky rains'" (Wright 1974:II, 271).

Wright continues: "In the case of words like *yajūzu* 'it is allowed', *yajibu* 'it is necessary', yanbaġī 'it behooves', *etc.*, followed by *ʾan* with the subjunctive, the subject naturally is the following clause, and therefore the verb does not come under the head of impersonal".

In their recent grammar, Badawi a.o. (2004:395) confirm that "verbs denoting 'must', 'ought', 'may', 'suffice', etc. have *ʾan* clauses as their agents"; they implicitly acknowledge that since these verbs have a non-dummy subject, they are not impersonal verbs.

The above-mentioned verb 'to suffice' (*kafā*/*yakfī*) is quite controversial. According to Wright (1974:II, 161), in the example "*kafā bi-*

llāhi šahīdan ‘God sufficeth as a witness’...the preposition *bi-* is redundant after *kafā*, its function being simply that of emphasizing the subject-predicate relation, in which *allāhu*is the agent”. Wright expresses the opinion that “it is better, however, to take *kafā* as containing its subject in itself, viz. *kifāya* and thus being used impersonally”.

According to Brahim (1996:30), “Dans le même exemple *wa-kafā bi-llāhi šahīdan* ‘Suffit d’*Allah* (comme) témoin’ l’équivalence entre groupe prépositionnel et nom au nominatif (*bi-llāhi* = *allāhu*) utilisée comme argument pour justifier l’analyse de ces groupes comme des *fāʿil* (sujets) ou *nāʾib fāʿil* (substituts de sujets) se fonde sur une correspondance distributionnelle trop partielle pour permettre une quelconque généralisation”.

Ibn Hišām (*Muġnī* 106) stresses that the particle *bi-* is “a superfluous augment (*zāʾida*) which occurs predominantly (*al-ġāliba*) with the subject of *kafā* as in *kafā bi-llāhi šahīdan*”; the nonagreement of the verb in the feminine in the example he quotes subsequently, *kafā bi-hindin* ‘Hind suffices’, does not constitute for him a counterexample because “according to Ibn as-Sarrāj, the subject is the pronoun [which refers to the *maṣdar*] *al-iktifāʾ*”.

Although classical Arabic grammarians often disagree on the analysis of such examples, they nevertheless admit that the presence of a subject or a semantically charged substitute for the subject is obligatory, be it apparent or underlying. In other words, the notion of impersonal verb seems to be incompatible with the classical Arabic grammatical theory. There exists a real dummy pronoun (*-hu*), the *ḍamīr aš-šaʾn* or *ḍamīr al-qiṣṣa* as in *qāla ʾinna-hu yanbaġī ʾan yarjiʿū* ‘he said that it was necessary for them to return’ (→ *ḍamīr*), but the *ḍamīr aš-šaʾn* only plays the role of a semantically empty embedding element aimed at ensuring the grammaticality of the sentence. The particle *ʾinna* cannot be directly followed by a verb, and therefore, the subject of *yanbaġī* remains the *ʾan* clause.

3. IMPERSONAL VERBS

According to the above definition, one gets the impression that in Arabic impersonal verbs are generally passive. Šarṭūnī (1969:188) describes the passivization process as follows:

> If the subject (*fāʿil*) is deleted and the verb has a complement (*mutaʿaddiyan ʾilā mafʿūl*), the complement takes the place of the subject and takes the subjective case; it is called ‘substitute for the subject’ (*nāʾib al-fāʿil*) and the verb is in the passive voice.
>
> If the sentence contains no complement, the subject can be replaced (*nāba ʿan al-fāʿil*) by a declinable element such as the *maṣdar* or a circonstant, but the two must be semantically defined (*muxtaṣṣ*) and be able to function as a support (*ʾisnād ʾilay-hi*), as in *ḍuriba ḍarbun šadīdun* ‘there have been violent blows/some violent blows have been stricken’ or in *suhirat laylatu l-ʾuns* ‘we spent the night in a sociable atmosphere/the night was spent in a sociable atmosphere’, where the agreement of the verb in the feminine and the subjective case placed on *layla* keep the ambiguity on the meaning of the sentence and do not allow for any particular choice.
>
> The subject can be replaced by a prepositional group (*al-majrūr wa-l-ḥarf*) as in *murra bi-l-bustāni* ‘we passed through the garden’ and *nuḏ̣ira fi l-ʾamri* ‘we examined the matter’; *al-bustāni* and *al-ʾamri* are virtually in the subject case as substitutes for the [subject] (*ʿalā n-niyāba*).

According to as-Suyūṭī (*Hamʿ* I,162), the analysis is more complex in the latter case:

> As to the prepositional group, if the preposition is redundant (*zāʾid*), there is no disagreement on the fact that it can take the place [of the subject] and that it is positionally in the subject case (*fi maḥall rafʿ*), like the word *ʾaḥad* in a context such as *mā ḍuriba min ʾaḥad* ‘nobody has been beaten’. However, if the preposition is different, things change: most scholars view the prepositional group as positionally in the subject case, hence, they believe it to be the substitute for the subject, as in *sīra bi-zayd* ‘there was a trip with Zayd’, regarding the preposition as superfluous. Some others, like Hišām, claim that the substitute for the subject is a nonspecified pronoun, dissimulated (*mubham mustatar*) in the verb, its function being that of referring back to the *maṣdar* signified by the verb. Others, with al-Farrāʾ, consider the preposition, positionally in the subject case, as the veritable substitute for the subject. Finally, Ibn Durustawayhi, as-Suhaylī, and ar-Rundī maintain that the substitute for the subject is a pronoun that refers to the *maṣdar* of the verb. What is underlying, then, is *sīra huwa ʾay as-sayr* ‘it has been traveled, i.e. the traveling’, because if the substitute for the subject was the indirect complement (*majrūr*), one could say *sīrat bi-hind* ‘there was a traveling [they traveled] with Hind’ and *julisat fi d-dār* ‘there was a sitting [they sat] in the house’.

This last argument is refuted by as-Suyūṭī, who quotes as a paradigmatic example *kafā bi-hindin*, where *hind*, though the subject of *kafā*, does not impose on the latter the feminine gender. The argument highlighted so far suggests a circular argument, caused by the fact that the grammarians felt that in Arabic there always needs to be a substitute for the subject, be it real or virtual.

Orientalist grammarians are rather cautious when talking about impersonal verbs. Blachère (1975:261) stresses that "les verbes arabes peuvent fournir un passif impersonnel", but that this passive refers to an active verb "figé à la troisième personne masculin pluriel dont le sujet virtuel serait *nās* 'gens': *furiḥa bi-hi* (= *fariḥū bi-hi*) 'on se réjouit de le voir', *sīra* (= *sārū) ʾilā l-ʿIrāq* 'on alla en Iraq'". Wright (1974:II, 268) notes:

> If the *impersonal* form of expression [by which he means '*our* impersonal form in English'] is to be employed, the Arabs use the third person sing. masc. of the passive voice, whether of a transitive or of an intransitive verb…; as *kutiba bi-l-qalam* 'it has been written, it is written with the writing reed'; *sīra ʾilā l-ʿirāq* 'there was a traveling, they traveled towards Iraq'.… Verbs thus used are always of the masculine gender, which the Arabs frequently employ where we should use the neuter.

He cautiously continues:

> If a passive, which is, according to our ideas, impersonal, governs an object by means of a preposition – as *ġušiya ʿalay-hi* 'he fainted' (lit. 'there was a covering thrown over him') – this subject becomes virtually the subject of the passive voice, just as it was virtually the object of the active.

Badawi a.o. (2004:389) mention a rather recent use of the passive with an impersonal appearance that "often functions as a kind of imperative, especially in labels and instructions for use: *yuḥfaḍu fi ṯallājatin* 'to be kept in refrigerator', lit. 'is kept', *turajju qabla l-istiʿmal* 'to be shaken before use'". These passive forms are not really impersonal; their subject is, as Arab grammarians put it, the pronoun dissimulated in the verb, which refers back to the concrete object on which the instruction is written.

Brahim (1996:34) attempts an analysis of impersonalization:

> En s'inspirant de la 'théorie de la valence' de L. Tesnière, on peut affirmer que la diathèse du *mağhūl* est essentiellement une 'diathèse récessive' (Tesnière 1976:272) dont la fonction caractéristique est la réduction de la 'valence verbale' par la suppression du 'prime actant' initial, de sorte qu'un verbe 'divalent' (transitif à complément unique) devienne 'monovalent' et qu'un verbe monovalent (intransitif) devienne 'avalent'. Dans ce premier cas, la fonction syntaxique du prime (en fait le seul) actant non exprimé est prise en charge par un élément postiche (*dummy element*) représenté en surface par une forme neutre (homonyme du pronom de l'absent [*ḍamīr al-ġāʾib*] masculin singulier) intégrée au verbe.… Mais on peut envisager une analyse encore plus conséquente avec la notion de réduction de valence et ce qu'elle implique quant à la valence zéro du verbe intransitif mis à la forme *fuʿila*. Au lieu de chercher coûte que coûte à trouver un sujet, même postiche, à ce verbe, il est possible de partir de l'idée que c'est la forme même du verbe qui 'impersonnalise' la désinence neutre (homonyme de la désinence de troisième personne) et non l'inverse, c'est-à-dire de considérer cette désinence comme une simple marque syntaxique nécessaire, en arabe, pour maintenir la 'verbalité' du verbe.

4. CONCLUSION

The impersonal verb does not have, or at least does not yet have, a real status in Arabic grammar. In this connection the verb *taqarrara*, which occurs more and more frequently in the Arabic press, may be mentioned. The 'mental' calque seems evident, whether from English *it was decided that* or from French *il a été décidé de/que*. Although *taqarrara* has a formal subject, a *maṣdar* or a clause, it is rather perceived as an impersonal verb. In the example *wa-ṣarraḥa s-sayyid [fulān] bi-ʾanna-hu taqarrara taklīf ittiḥādāt*…'Mr [X] declared that it had been decided to assign to the unions…' (*al-ʾAhrām* January 2000), the meaning is not 'the fact of assigning to the unions…has been decided', but rather 'it has been decided to assign to the unions…'. This simple calque would suffice to justify the use of the *ḍamīr aš-šaʾn*, which takes the place – 'mentally', not yet grammatically – of the subject *it* or *il* of the English or French impersonal verb. The *ḍamīr aš-šaʾn* is evidently required as an embedding element because of the – phraseologically, not grammatically – imposed Verb + Subject order.

Nevertheless, in the case of a sentence such as *qāla ʾinna-hu yajibu ʾan yušārika [fulān] fī*…'he said that it was necessary for [x] to take part in…', it is not clear what is perceived as the subject of *yajibu*, the pronoun *-hu* or the *ʾan yušārika* clause. Finally, in the sentence *qāla ʾinna-hu yajibu mušārakatu-hu* (**mušārakata-hu?*)…"he said that it is necessary for him to take part in…', it is not clear what the 'mental subject' is (Girod 2000:223).

Bibliographical references

Primary sources

Ibn Hišām, *Muġnī* = Jamāl ad-Dīn ʾAbū Muḥammad ʿAbdallāh ibn Yūsuf Ibn Hišām al-ʾAnṣārī al-Miṣrī, *Muġnī l-labīb ʿan kutub al-ʾaʿārīb*. Cairo: Maṭbaʿat al-Madanī, n.d.

Ibn Jubayr, *Riḥla* = ʾAbū l-Ḥusayn Muḥammad ibn ʾAḥmad Ibn Jubayr, *Riḥla*. Cairo: Dār at-Taḥrīr li-ṭ-Ṭabʿ wa-n-Našr, 1968.

Suyūṭī, *Hamʿ* = Jalāl ad-Dīn ʿAbd ar-Raḥmān ibn ʾAbī Bakr as-Suyūṭī, *Hamʿ al-hawāmiʿ šarḥ Jamʿ al-jawāmiʿ*. Beirut: Dār al-Maʿrifa li-ṭ-Ṭibāʿa wa-n-Našr.

Secondary sources

Badawi, El-Said, Michael G. Carter, and Adrian Gully. 2004. *Modern Written Arabic: A comprehensive grammar*. London: Routledge.

Blachère, Régis. 1975. *Grammaire de l'arabe classique*. Paris: Maisonneuve et Larose.

Brahim, Ahmed. 1996. *L'occultif*. Le Bardo: Association Tunisienne de Linguistique.

Galāyīnī, Muṣṭafā al-. 1973. *Jāmiʿ ad-durūs al-ʿarabiyya*. Beirut: al-Maktaba al-ʿAṣriyya.

Girod, Alain. 2000. *Faits d'évolution en arabe moderne à travers un corpus de presse égyptien*. Ph.D diss., Université de Provence.

Šarṭūnī, Rašīd aš-. 1969. *Mabādiʾ al-ʿarabiyya fi ṣ-ṣarf wa-n-naḥw*. Beirut: Dar el-Mashreq.

Tesnière, Lucien. 1976. *Eléments de syntaxe structurale*. Paris: Klincksieck.

Wright, William. 1974. *A grammar of the Arabic language*. Repr. Beirut: Librairie du Liban.

Alain Girod
(University of Aix-en-Provence)

Implicational Scale

Implicational scaling is a very useful language analysis tool when it is suspected that two variables are in a dependency relationship with one other. This entry explores the nature and proper (and improper) use of implicational scaling. It also gives some examples from Gibson's work in Tunisian Arabic, where ordering of changes in the dialect demonstrate what Weinreich a.o. (1968) call 'structured heterogeneity'.

1. The implicational relationship

Implication is the relationship between two variables such that we can make a generalization about the absence or presence of one from the absence or presence of another. An example from Gibson (1998) is given in Table 1 to demonstrate this relationship.

Table 1. Co-occurrence of gender marking in the 2nd person in Tunisian Arabic

		PRONOMINAL SYSTEM	
		no gender marked	gender marked
VERBAL SYSTEM	gender not marked in perfect	165	9
	gender marked in perfect	0	68

Table 1 sums up data taken from 242 speakers of different varieties of Tunisian Arabic. It shows that the vast majority of speakers use either a fully urban variety, i.e. a variety that does not mark gender anywhere in the 2nd person (represented by the top left cell) or a fully Bedouin one, i.e. a variety that does (as shown in the bottom right cell). We also have in the top right cell nine speakers who use gender marking in the relevant pronouns, but not consistently with the verbs – a mixed system. The bottom left cell represents a combination with no speakers.

The data in Table 1 lead to two absolute implicational statements: (i) if speakers use gender marking in the verb, then they also distinguish for gender in the pronouns; and (ii) a lack of gender marking in the pronouns implies the same in the verbs. In any such table with one empty cell, there are always two implicational relationships. These two may, at first sight, appear to be independent statements, but they are inevitably each other's corollary. In each case, the inverse of the implication does not hold; we can say nothing about use of the verbal forms if we find marking in the pronouns (both options are available), nor can we predict what marking there is among the pronominal forms if there is no marking in the verbal system. The empty cell is essential for there to be a real or strong implicational relationship. (Weak or statistical implication is discussed below.)

The implicational relationship describes a situation that is clearly different from that of free variation, where all four cells of Table 1 would be filled in some way. It is also different from *equivalence* or *bilateral implication*, where there is complete codependency between two variables, exemplified by Table 2.

Table 2. Equivalence or bilateral implication

		VARIABLE A	
		a_1	a_2
VARIABLE B	b_1	found	not found
	b_2	not found	found

Here, the presence of b_1 implies a_1, whereas in Table 1 we cannot make the inverse implication, that a_1 also implies b_1. The same relationship holds between a_2 and b_2. Whereas the normal implicational relationship is that expressed by *if* in logic, the equivalence relationship is the logical *iff*, i.e. *if and only if*: a bidirectional relationship. Bilateral implications denote a discrete split between language varieties, rather than a continuum. Henceforth, any reference to implication refers to the unilateral implication demonstrated in Table 1.

2. Implicational scales

The use of implicational scales in social science stretches back to the Guttman scale (from its first use in Guttman 1944). The word 'scale' denotes a one-dimensional measurement, but the term 'implicational scale' is often used by sociolinguists for the matrix (e.g. Table 3 below); this might be better named an 'implicational matrix'.

In an implicational matrix, one dimension represents the individuals involved, and the other represents different linguistic environments. The data are then arranged so that the data appear, as much as possible, to pattern in an implicational manner, where implicational relationships may hold between each environment. A neat example is Table 3, Fasold's (1990:191) restructuring of De Camp's (1971a:355) portrayal of an implicational continuum holding between seven speakers of Jamaican Creole.

Table 3. An implicational scale or matrix

	Feature					
Speaker	B	E	F	A	C	D
1	+	+	+	+	+	+
2	+	+	+	+	+	-
3	+	+	+	+	-	-
4	+	+	+	-	-	-
5	+	+	-	-	-	-
6	+	-	-	-	-	-
7	-	-	-	-	-	-

The presence of a plus sign means in this case that the relevant speaker uses the standard form of the variable, and the minus sign the Creole form. In this example, apparent implicationality can be seen between all environments.

3. Implications in linguistic research

The notion of implication in linguistics has been particularly associated with the study of language typology and universals, e.g. Greenberg (1962:83, n. 1), where he acknowledges the work of Roman Jakobson in pointing him toward implicational (as opposed to absolute) universals of language. This use of implicational universals has continued in the typological and generative paradigms, where it has been complemented by the use of parameters. Drawing on these frameworks in an attempt to lay the bases of a theory that could predict the route of language change, Weinreich a.o. (1968) invoke interconnected changes as phenomena that can delimit possible types of linguistic change; these are in fact implications. In a similar vein, Greenberg (1978) proposes that implicational universals not only account for the state of languages but also limit the routes from one state to another, i.e. the route of change, and that therefore there should be universals of linguistic change.

In the late 1960s, many students of language variation became interested in implication as a tool of analysis. For example, Labov (1969) uses implicational analysis to examine the social meaning of certain linguistic features in Spanish. But it was in the field of creolistics that implicational analysis was initially popularized, starting with De Camp (1971a), and where its ramifications were most deeply considered. The phenomenon under study was the 'post-creole continuum'. Taking Jamaica as an example, this continuum of varieties resulted from continued contact between Standard English and (English-based) Jamaican Creole.

De Camp finds that most speech in Jamaica cannot be analyzed as either purely standard or purely creole, and he further finds that the use of one variant often implies the use of another, logically independent, variable. This implies a linear continuum of varieties, the alternative to each rule being 'blind' to the application of other rules, which would give a much greater number of possible combinations of variants in an utterance. De Camp (1971b:34) describes the community grammar as follows: "Implicational analysis attempts not to describe a set of speech acts but to model the idealized competence of the persons involved in those speech acts". He does not, however, attempt to find linguistic explanations for these relationships; there is no a priori reason why the word *child* should be more restricted to near-standard contexts than *eat*.

Outside Creole studies, Kristensen and Thelander (1984) find that the use of dialect variants can also be analyzed in terms of implicational scales. In a work that provides a very useful framework for examining contact-induced dialect change, Auer (1997) proposes a fixed route in the shift toward a standard language, with a strong claim that a phonological dialectalism implies a morphological dialectalism. Dialect shifting toward the standard must, herefore, occur first in the morphology before it affects the phonology. In the field of Arabic linguistics, the earliest example of implicational analysis in a contact situation is Holes (1987: 81–91).

4. The implicational relationship in gender marking in Tunisian Arabic

Now we turn to examine an Arabic implicational relationship in detail, taken from Gibson (1998). Table 1 shows a clear implicational relationship between the marking of gender in the pronoun and the perfect form of the verb in Tunisian Arabic. In fact, the correlation is highly significant, with a security level of 0.004. Similar absolute relationships hold between the pronoun and the verb in the imperfect and the imperative but do not occur between the different aspects and moods of the verb.

Gibson (1998, 2002) shows that many Tunisian Arabic speakers accommodate their Bedouin dialect in the direction of the dialect of Tunis (this variety, unlike Modern Standard Arabic and most modern Arabic dialects, does not mark gender in the 2nd person). On this variable (unlike any of the others examined in the same works), there is a categorical implication among Bedouin-origin speakers:

genderless pronoun $\supset$ *genderless verbal form*

or its converse,

gender-marked verbal form $\supset$ *gender-marked pronoun*

Hence any intermediate system will consist of a gender-marked pronoun and a fully or partially genderless verbal system. This implies that in the shift from a system that marks gender, this distinction will be lost first in the verbal system, and only then in the pronominal system. This is not to say that gender marking is merely more likely to be lost in the verbal system, but that its loss in the pronominal system is dependent upon its disappearance in the verbal system. This is what the implication established shows – a fixed route in accommodation to a Tunis-like variety. What is interesting is why the relationship between gender marking on pronouns and verbs should be so strong.

The discovery of such a strong implicational relationship motivates a further quest for explanation: this is not what we expect when correlating logically independent variables. The relationship here is one that holds over a variety of dialects. In fact, mixed dialects were found among speakers from all over the Tunisian interior. Explanations such as one variable having special social marking are not sufficient. In fact, there is no evidence of particular stigmatization of one feature, and even if this were to be the case, we would still want to look for a linguistic explanation behind it. Some such potential reasons are examined in Gibson (1998:216–222) and are found to be unconvincing.

Labov (1972:300) asserts that in contact situations the simpler system is more likely to win out. A study of linguistic universals and typology can help us define what constitutes a simpler system. In → pro-drop languages such as Arabic, we often find a strong match in person, number, and gender marking between the verbal and pronominal systems. The unmarked form of a sentence is one with no lexicalized subject pronoun, the features of the subject being marked on the verb. Hence, pronouns have a different pragmatic role from that found in languages with an obligatory overt subject. In Tunisian Arabic, for example, pronouns are mainly used either with the copula or to mark focus and contrastive topics. So if there is any difference between the number of forms available to verbal morphology and this number in the pronominal system, then we would expect to find finer distinctions in the pronominal system, where contrast and focus are expressed. This generalization will be referred to as the 'feature implication hypothesis'. It is supported by a typological study of 49 languages made by Perkins (1992:184–187) for inclusive/exclusive and dual marking on verbs and pronouns. The evidence is not entirely conclusive; Perkins does find some counterexamples. In fact, Modern Standard Arabic constitutes a counterexample in the conjugation of dual forms that are differentiated for gender (unlike the pronoun), as does the dialect of Jenin, Palestine, where gender is marked in the verb morphology of the 2nd person, but not on the equivalent pronoun.

The feature implication hypothesis, a strong universal tendency, is followed strictly in the Tunisian case of contact-mediated change. The change in the verbal forms is a necessary prerequisite of the change in the pronominal system: a rare example of something that fulfils the aims of Weinreich a.o. (1968). So we have a fixed-route process for the loss of gender, with all varieties remaining within the limits of the preferred language type delimited by the feature implication hypothesis. This contact-mediated change has taken the route of the simplest system. Changes that are not due to contact are more likely to result in less simple systems. There is no strong implication between the forms of the verb itself; these are not linked by any universal tendencies.

Here we have a clear case in which the use of implicational analysis gives us a richer understanding of the data than that which could be obtained through other statistical methods. Had it not been used, we might not have been driven to suggest an explanation of why this implication should hold so strongly.

5. STATISTICAL IMPLICATION

When using the notion of implication, creolists do not adhere to the notion that implication is necessarily an absolute relationship with an empty cell in the matrix, as in Table 1. The measure of the extent to which a relationship of covariation can be defined by implication is referred to by Bickerton as 'scalability', a figure representing the number of cells in an implicational matrix that fit the proposed pattern of implication; in the case of strong implication (as in Table 1), this figure is 100 percent. However, Bickerton, along with other creolists, permits a certain number of exceptions under his working definition of implication. For an implication still to be valid, he says that "in practice, figures around 90% can be regarded as adequate" (Bickerton 1973:647). The use of the implicational scale has come under much criticism from some sociolinguists (see, e.g., Romaine 1982:177–82; Fasold 1990:199), but none have shown that Bickerton's 90 percent figure for validating an implicational matrix is fallacious. An example disproving Bickerton's contention is now given.

For there to be a genuine relationship of correlation, if we find variant a_1, then variant b_1 must be more likely to occur than if we had a_2 in the same environment. As a hypothetical example, we will take two linguistic variables, <A> and <B>. Innovation a_1 occurs in 70 percent of possible instances, and innovation b_1 in 20 percent. Assuming that the two variables pattern independently of each other, we would get something like the covariation pattern in Table 4:

Table 4. The matrix of probabilities given independent variation of variables <A> and < B>

	a_1	a_2
b_1	14%	6%
b_2	56%	24%

Table 4 satisfies Bickerton's criteria for an implicational relationship to hold, with 94 percent of cells satisfying the implication that the presence of b_1 implies a_1 (and a_2 b_2); it would seem that the innovation in <A> is preceding that in <B>. In order to claim this, however, we need to establish whether the presence of b_1 correlated with that of a_1. Of course there is no such correlation; the figure of 6 percent of counterexamples arises from the covariation pattern of two independent variables. We have thus been able to show that a scalability measure of 90 percent does not correctly delimit real implicational relationships; in fact, a percentage figure is not a suitable measure at any level. An implicational relationship is of necessity a relationship of dependency.

In Table 4, innovation in <A> is found more often than innovation in <B>. This is very different from saying that there is a strong or even weak implicational relationship between them. What we have is purely and simply a relationship of differing strengths of rule application. From this we have no basis to propose any temporal or other ordering of the two variables with respect to each other. This can only be proposed in the case of a statistically justified implicational relationship.

A better basis for proposing a genuine but statistical implication is needed. In the case of strong implication, the correlation coefficient is 1 or -1. In the example given in Table 4, it is 0, as there is no codependency. Furthermore, a correlation coefficient in itself is not enough to propose an implicational relationship, as we need to measure the chance of such a coefficient occurring as a consequence of the random patterning of two independent variables. Random variation will normally give some small level of correlation even for totally independent variables, so statistical tests (in this case an exact-probability test) are used to assess the likelihood of such variation being due to chance. Results that have less than a 5 percent chance of occurring by chance are normally accepted as being sufficient to show some level of correlation between the two variables. Where such statistical testing does show some level of codependency, we may propose statistical implication, so we are not led to reject implicational analysis: it can still offer an insight into ordering of changes.

6. The variable <Q> in Tunisian Arabic: A case of statistical implication

The second variable to be examined is more fully discussed in Gibson (1998); the data used here are taken from interviews. It again concerns the adoption of an urban sedentary variety by Bedouin-origin speakers and concerns a variable familiar to any student of Arabic dialectology, the reflexes of Old Arabic → *qāf*. In both urban and Bedouin varieties of Tunisian Arabic, we find two phonemically contrastive reflexes of this variable: /q/ and /g/. However, in the majority of these words, urban varieties have /q/ while the Bedouin varieties have /g/. Table 5 presents which form was used in particular roots and words.

The matrix is arranged so that the use of /g/ is maximal to the top and left. The bold line is an attempt to separate uses of /g/ from /q/. Starred cells are those that do not fit the proposed pattern, while one speaker does not seem to be ranking more privileged environments for use of /q/ in a manner similar to other speakers.

Submitting potential implications between columns to statistical evaluation, we find that none are secure at the 5 percent level. The closest we get to this is the relationship between *l-q-a* and /qad/, with all seven cases of mismatch being in the predicted direction, but where independent variation would predict a probability of this happening in each case by chance as 0.766, giving us the overall probability of 0.766 to the power of 7, which is 0.156, equivalent to a significance level of 15.6 percent. The comparatively small amount of data makes it difficult to establish individual implications with any level of certainty, as does the increased number of starred cells. But if we further combine the above categories in the manner shown in Table 6, we find more individuals that can be examined for each implication.

The significance values of the proposed implications in Table 6 are given in Table 7.

Hence, even though the data sample is quite small, we can still show one implicational relationship to be valid if we group larger numbers of words together. What we have here is not an absolute implication but rather a statistical one; there are some exceptions. However, we have established that a speaker's use of /g/ in a word such as /qad/ has a statistically

Table 5. An implicational matrix for /g/-use

	q-w-l	*l-q-a*	*q-m-ḥ* *q-ʿ-d* *ṭ-ʿ-l-q* *ḥ-r-q* *q-d-m*	*z-r-q* *ḏ-w-q* *qad* *q-l-q*	*q-t-l* *s-w-q* *qbilīka* *q-l-l* *qrīb*	*fūq* *qbal* *ḥ-q-q* *q-s-m*
BenGardaneM	g	g	g			g
JerbaM2	g		g		g	g
ZarzisM3	g	g	g	g	g	q
AlamM2	g	g		g	g	q
OmAdhmaM	g	g	g/q*		g	q
KefF3	g		g	g	g/q	
AgegchaM		g	g	g/q	g/q	q
TalaM	g	g		g	q	q
GafsaM3		g	g	g	q	q
ChniniGabesF	g/q*	g	g	q/g	q	
TeborsoukF	g	g	g	q	q	q
GhomrassenM1	g		g		q	
SidiHassouneM	g	g	g			q
GafsaM2		g	g			q
TnGafsaM1	g	g/q*	g	q		q
NeftaM	g	g	g	q		
BirLahfeyM2	g	g				q
JendoubaM	g	g/q	g/k		q	q
ZaghouanM	g				q	q
TnGafsaM3	g/q	g/q	q	q	q	
TnJerbaM1	q (g)	q/g	q	q		q
ZarzisM2	g/q	q	q (g)*	q		q
Misfit						
JerbaM3	g (q)		g	q		q

significant relationship on concomitant use of /g/ in a root such as *q-w-l*. Note that we have not validated the entire chart, but just one of the relationships within it.

Such data, among others found in Gibson (1998), show that there are groups of words that do pattern together and vary in similar ways to other members of that group. When taken as a group, they act in a hierarchical way in relation to other groups of words, the use of one feature in one group implying to some extent its use in another group. As to why the words are in a particular group, there are some pointers. Loans from Modern Standard Arabic and more technical words tend to be found with /q/, while many of the most common words, such as *q-w-l* 'to say' and *q-ʿ-d* 'to sit, stay' (and marker of progressive aspect), are found with /g/, along with agricultural words such as /qamḥ/ 'wheat'. This finding is similar to what Holes (1987:49–57) notes in Bahrain, where he says that the dialectal 'core-items', such as /gāl/ 'to say' (also the word most resistant to change in Tunisia), show no variation, despite morphosemantic congruity with Modern Standard Arabic, which he otherwise finds to be a predictor of change. He claims that these words are "too much a part of the core of the dialect for replacement to occur in anything but the most formal styles of speech" (Holes 1987:53). Thus, frequency

Table 6. A combined implicational matrix for /g/-use

	q-w-l *l-q-a*	*q-ʿ-d* *qad*	*qrīb* *qbal*
BenGardaneM	g	g	g
JerbaM2	g	g	g
ZarzisM3	g	g	g/q
AlamM2	g	g	g/q
OmAdhmaM	g	g/q	g/q
KefF3	g	g	g/q
AgegchaM	g	g	g/q
TalaM	g	g	q
GafsaM3	g	g	q
ChniniGabesF	g (q)	g (q)	q
TeborsoukF	g	g/q	q
GhomrassenM1	g	g	q
SidiHassouneM	g	g	q
GafsaM2	g	g	q
TnGafsaM1	g (q)	g/q	q
NeftaM	g	q	
BirLahfeyM2	g		q
JendoubaM	g (q)	g/k	q
ZaghouanM	g		q
TnGafsaM3	g/q	q	q
TnJerbaM1	g/q	q	q
ZarzisM2	q (g)	q (g)	q

Table 7. Significance values for Table 6

	q-ʿ-d qad	*qrīb qbal*
q-w-l l-q-a	*0.040	0.252
q-ʿ-d qad		0.247

of occurrence is a barrier to such change in two very different dialects of Arabic.

7. CONCLUSION

The use of implicational analysis has given significant insights into the patterning of two linguistic variables. In the case of gender marking we have seen that there is strict ordering of the categories in which it was lost, and that this is a rare case where we can predict the route that a linguistic change will take. The reasons for such a strict ordering seem to be found within language typology, and we do not find such strict ordering in the route of acquisition of a new dialect in the variable <Q>; however, we still find evidence for a less-than-perfect continuum of varieties here. We can conclude that implicational analysis is indeed a very useful tool in studying variation, but one to be used more carefully than has often been the case in the past.

BIBLIOGRAPHICAL REFERENCES

Auer, Peter. 1997. “Restrictions of co-occurrence between linguistic variables: A case for social dialectology, phonological theory and variation studies”. *Variation, change and phonological change*, ed. Frans Hinskens, Roeland van Hout, and W. Leo Wetzels, 69–99. Amsterdam and Philadelphia: J. Benjamins.

Bickerton, Derek. 1973. “The nature of a creole continuum”. *Language* 49:3.640–669.

De Camp, David. 1971a. “Towards a generative analysis of a post-creole speech continuum”. *Pidginization and creolization of languages*, ed. Dell Hymes, 349–370. Cambridge: Cambridge University Press.

——. 1971b. “Implicational scales and sociolinguistic linearity”. *Linguistics* 73.30–43.

Fasold, Ralph. 1990. *The sociolinguistics of language.* Oxford: Blackwell.

Gibson, Michael [Maik]. 1998. *Dialect contact in Tunisian Arabic: Sociolinguistic and structural aspects.* Ph.D. diss., University of Reading.

——. 2002. “Dialect levelling in Tunisian Arabic: Towards a new spoken standard”. *Language contact and language conflict in Arabic: Variations on a sociolinguistic theme*, ed. Aleya Rouchdy, 24–40. London: RoutledgeCurzon.

Greenberg, Joseph H. 1962. “Some universals of grammar with particular reference to the order of meaningful elements”. *Universals of language*, ed. Joseph H. Greenberg, 55–90. Cambridge, Mass.: MIT Press.

——. 1978. “Diachrony, synchrony and language change”. *Universals of human language*, ed. Joseph H. Greenberg, I, 62–91. Stanford, Calif.: Stanford University Press.

Guttman, Louis. 1944. “A basis for scaling qualitative data”. *American Sociological Review* 9.139–150.

Holes, Clive. 1987. *Language variation and change in a modernising Arab state.* London: Kegan Paul International.

Kristensen, Kjeld and Mats Thelander. 1984. “On dialect levelling in Denmark and Sweden”. *Folia Linguistica* 18.223–246.

Labov, William. 1969. “Commentaries: How can we describe and measure the behavior of bilingual groups?”. *Description and measurement of bilingualism*, ed. Louis G. Kelly, 250–255. Toronto: University of Toronto Press.

——. 1972. *Sociolinguistic patterns.* Philadelphia: University of Pennsylvania Press.

Perkins, Revere D. 1992. *Deixis, grammar and culture.* Amsterdam and Philadelphia: J. Benjamins.

Romaine, Suzanne. 1982. *Socio-historical linguistics: Its status and methodology.* Cambridge: Cambridge University Press.

Weinreich, Uriel, William Labov, and Martin Herzog. 1968. "Empirical foundations for a theory of language change". *Directions for historical linguistics: A symposium*, ed. Winfred P. Lehmann and Yakov Malkiel, 95–189. Austin: University of Texas Press.

Maik Gibson (Reading, United Kingdom)

India

From the 8th century onward India has had a close association with Arabic. In its classical form, with some regional linguistic features, the language has played a significant liturgical and religio-scholarly role in the subcontinent. It has generally not been utilized for routine communication or other secular functions. Viewed as the sacred language of Islam and the *Qurʾān*, its primary use is by Muslims, who equate Arabic scholarship with Islamic learning and regard both the language and its scholars with veneration. Indian Muslims currently number approximately 133.54 million (Shahabuddin 2003, from *Census of India* 1991, 2001). The vast majority of Indian Muslims have, at the very least, a basic liturgical association with Arabic, and some study the language in *maktab*s or madrasas. Accordingly, India has one of the largest numbers of Arabic users – albeit as a subsidiary language – in the world. Indeed, through the centuries Indian scholars and poets have proffered a weighty, though largely unrecognized, contribution to the Arabic library.

1. Historical development

Indians made their first, rudimentary acquaintance with Arabic sometime in the 3rd century C.E., close to the beginning of the time it appeared as a distinct language in the Arabian Peninsula. This linguistic contact transpired through Arab sailors who, from antiquity, had been docking at ports on the southwestern coast of India to acquire spices. From the 8th century onward, the history of Arabic in India (Yusuf 1967) became closely linked with the development of Islam in the subcontinent (Wink 1990–2004; Schimmel 1980; overview in Burton-Page 1971).

In 711 C.E., the Arab-Muslim Umayyad commander Muḥammad ibn al-Qāsim aṯ-Ṯaqafī conquered and colonized the western Indian province of Sind. Over the next four centuries of Arab rule, Arabic was probably the language of administration, and the cities of Multan, Mansura, and Daibul became key centers of Arabic-Islamic scholarship. According to the medieval geographers Ibn Ḥawqal and al-Muqaddasī, the people of these towns were bilingual, speaking both Sindhi and Arabic (Yusuf 1967:56). Southward, maritime trade continued. Early on, several small trader settlements of Muslim Arabs from coastal Yemeni and Iraqi backgrounds sprang up on the Malabar coast (Wink 1990:I, 67–86; Koya 1988; Bouchon 1986). These settlers came to form distinct ethnic groups, including the Navāyat or Naitias (from the Arabic *nūtī*, pl. *nawāʾiṭ* 'mariner') of Maharashtra and Karnataka (Bahadur 1902; Poonawala 1993), the Mappilas of Kerala (Miller 1991), and the Ilappais or Labbais of Tamil Nadu (Mines 1986). Arabic was initially their mother tongue, but they gradually phased it out, switching to the local languages.

In 1193, the Turkish sultan Muḥammad Ghuri conquered Delhi and established definitive Muslim dominion in India. From the 13th through the mid-19th centuries, Muslim rulers controlled almost the entire subcontinent. Over time, large numbers of the local populace converted to Islam and gained an interest in learning the language of its scriptures, Arabic. Following the initial Umayyad conquest, the majority of the incoming monarchs, including the Great Mughals, were Central Asian Turks, whose mother tongue was not Arabic. As such, they neither cultivated Arabic as a language of daily use in India nor used it as an official court language (the latter role was filled by Persian; for the parallel development of Persian in India, see Alam 2003). Thus, contrary to the case in lands conquered by Arab Muslims, such as Egypt, Arabic did not displace the local Indian languages. However, the Mughals and other Muslim sovereigns fostered Arabic as a language of religion and scholarship by patronizing Arabic-Islamic scholars and establishing a large number of Arabic madrasas and *maktab*s.

By the 10th century, the first ad hoc *maktab*s – associated with local mosques – had already been established in Sind, particularly in the towns of Mansura and Multan. In the last decade of the 12th century, Muḥammad Ghuri instituted formal madrasas in the town of Ajmer, and early in the 13th century, Iltutmish founded the first madrasa in Delhi and one in Badaun. In the following decades, madrasas sprang up all over the north. There were fewer elsewhere, but Ibn Baṭṭūṭa (*Riḥla* IV, 66 = *Travels* 218), who visited the city of Ḥinawr on the Malabar coast in the 14th century, does mention that he saw there something he had not seen in other parts – 23 *maktab*s for boys and 13 for girls; he marvels that women memorized the (entire?) *Qurʾān*. Over the next five centuries of partial or full Muslim rule, madrasas, many of them independent of mosques, proliferated in all parts of India into the hundreds.

In 1612, the East India Company was formed, and the British gradually took over Mughal power until, in 1857, they deposed the last emperor and declared India a colony. They promoted Western-style secular education, particularly English, and the Arabic language (and Persian) diminished in importance. Many madrasas were adversely affected, but Islamic revivalists instituted several new ones deliberately to counter the colonial approach and bolster the traditional education of Indian Muslims. In these institutions, "Arabic, being the language of the original sources of Islam, was to be the major focus of study. It was, so to speak, not only a language, but the major linguistic symbol of Islamic identity and Muslim resistance to modernity" (Rahman 2000:411). Most of the important (Salafī and other) madrasas existing today were established during British rule in the 19th century: Dār al-ʿUlūm in Deobandh (Metcalf 1982; al-Fārūqī 1990); Jāmiʿa Sayfiyya in Surat; and, in Lucknow (Farooqi 1999), Dār al-ʿUlūm Nadwat al-ʿUlamāʾ (Zaman 2001), Madrasat al-Wāʿiḏīn, and Jāmiʿa Nāḏimiyya. (For details of all Indian madrasas and Muslim universities, see Kaur 1990; Desai 1978; Rahman 2000; Isḥāq 1996.)

While deemphasizing religious madrasa education, the British – indirectly, through modernist Muslim reformers – created three institutions of secular learning, largely for Muslims, although not restricted to them: Aligarh Muslim University (founded 1875; see T. Wright 1966), Jamia Millia Islamiyya (founded 1920), and Jamia Osmania University (founded 1917). The Arabic curriculum in these institutions was to some extent detached from Islam; nevertheless, students of Arabic were mostly Muslims who were interested in the language for religious reasons.

In 1947, after independence and partition into India and → Pakistan (and later → Bangladesh), Arabic usage in all three nations developed in somewhat different directions. For example, in the 1950s and 1970s, various political groups in Pakistan voiced the proposal that Arabic be adopted as the national language (H. Nadvi 1975; Rahman 2000:416). In India, the central government has, in the main, not promoted Arabic scholarship (it does patronize the study of Sanskrit, the sacred language of Hinduism, through scholarships, literary prizes, and organizations such as the Rashtriya Sanskrit Sansthan). Being an intrinsic part of Muslim religiosity, however, Arabic usage continues to be important. Institutions of Arabic-Islamic learning flourish – the number of full-time Arabic madrasas in 1996 is listed by a modern scholar as 757 (Qamar ad-Dīn 1996:70). Under the auspices of the madrasas and the various Indian Muslim communities, Arabic scholarship, too, thrives.

2. DISTINCTIVE FEATURES OF INDIAN ARABIC

Because of its association with religious tradition, Indian Arabic has preserved classical features and archaic forms. It has almost totally resisted the penetration of dialectal elements and modifications discernible in Modern Standard Arabic. Adherence to Classical Arabic is manifest in three major areas:

i. Orthography and pronunciation. Rather than being changed to *ʾalif*, the archaic *wāw* is preserved in the orthography of certain Aramaic-origin words (صلوة, not صلاة; زكوة, not زكاة). The *ʾalif* is usually not written in certain words (قيمة, not قيامة; سموت, not سماوت). Long vowels are generally preferred to the *hamza* in the *tafʿīl* form of verbal nouns (*tāwīl* vs. *taʾwīl*, *tārīx* vs. *taʾrīx*); and (usually in pronunciation only), in the *mufʿil* form of the active participle (*mūmin* vs. *muʾmin*).
ii. Reading and speech (uttering of short phrases). Except in rhyming prose and

expressions ending in *Allāh*, the pausal form is rare. The *hamzat al-waṣl* is carefully maintained. And instead of being changed to a long *ā*, the *tanwīn fatḥa* at the end of a sentence is often sustained.

iii. Grammar. Rather than being substituted by the masculine plural, the *nūn an-niswa* and the dual are regularly used. The *lām al-ʾamr* and *nūn at-tawkīd* can be found in most texts.

Contrary to standard Middle Eastern pronunciation, Indians use the Persianate pronunciation of Arabic. They pronounce all emphatic consonants nonemphatically (*ṣ* > *s*, *ḏ̣* > *z*, *ḍ* > *d*, *ṭ* > *t*, *ḷ* > *l* in the word *Aḷḷāh*). They have difficulty with the velars and pharyngeals (*hurūf ḥalqiyya*), some always (ʿ > ʾ , *ḥ* > *h*, ʾ dropped from *ʾalif mamdūda*), and others occasionally (*x* > *kh*, *ġ* > *g*). They substitute some nonemphatic apicals with other, similar-sounding consonants (*ṯ* > *s*, *ḏ* > *z*). From time to time they mispronounce three other consonants (*q* > *k*, *w* > *v*, *z* > *j*). They treat the *tāʾ marbūṭa* as a *tāʾ maftūḥa*, pronouncing it as *t* rather than *h*, even outside an *ʾiḍāfa* construct (*daʿwat*, *jannat*). Their articulation of vowels is also unusual: *u* and *i* at the end of words become *ō* and *ē* (*kuntu* > *kuntō*, *kunti* > *kuntē*); and *a* at the beginning or middle of a word often becomes *e* (ʾAḥmad > Ehmad, Zahrāʾ > Zehra, *ṣaḥrāʾ* > *sehra*).

Textual composition is dominated by the enormously difficult and often artificial ornate Ḥarīrian *badīʿ* style. Authors regularly employ what appear to be pure verbal acrobatics, such as restricting the text to undotted letters of the alphabet (see the *Qurʾān* commentary *Sawāṭiʿ al-ʾilhām* by Akbar's court poet Fayḍī, d. 1595), or avoiding the use of letters like the *ʾalif* or *fāʾ*. There is often, however, a philosophical rationale – such as a *tawḥīd*-based cosmological underpinning – behind this manner of writing (some long opening *taḥmīd* sections of the *Rasāʾil Ramaḍāniyya* of the Ṭayyibī *dāʿī* Ṭāhir Sayf ad-Dīn, d. 1965, are good examples).

3. Curricula and teaching methodology

Arabic is taught in India for the most part in religious schools called *maktab*s and madrasas, occasionally at home, *maktab*-style, by *mulla*s or *moulvi*s, and, to a small extent, in secular universities. *Maktab*s impart primary learning, including *Qurʾān* recitation and memorization, *Šarīʿa* precepts relating to the ritual prayer, and basic reading knowledge of the Arabic script. In madrasas, which are generally for more advanced religious learning, Arabic is an important component of the syllabus, and many of the schools have 'Arabic madrasa' as part of their name (Persian madrasas exist as well).

The curriculum followed currently in most madrasas is the Dars-i Niẓāmī, the revised version proposed in the 18th century by Mulla Niẓām ad-Dīn of Sihali (near Lucknow) of an earlier, more ad hoc program. It includes texts from the Middle Eastern Arabic canon in the religious sciences of *Qurʾān* exegesis, *ḥadīṯ*, jurisprudence, Sufism, theology, and history; the related subjects of grammar, rhetoric, and prosody; the rational sciences of logic and philosophy; medieval science (added in the 15th century); and *belles lettres* (with a view to understanding better the literary features of the *Qurʾān*). Niẓām ad-Dīn incorporated a large number of books authored by early-18th-century Indian savants such as Mulla Jīwan of Amethi, Mīr Muḥammad Zāhid al-Harawī, and Mullā Maḥmūd Jawnpūrī. Later, secular fields such as mathematics and English were included as peripheral subjects. In the 19th and 20th centuries, Salafī madrasas expurgated Sufi material. Shiʿite madrasas follow different curricula with regard to religious texts, but they usually use the same works for the study of Arabic grammar, rhetoric, and literature.

The method of teaching Arabic in these madrasas is text based, where Islam is the subject and Arabic the tool. Madrasa students pay particular attention to the study of philology, as this is vital in learning to decipher the Islamic scriptures and theological texts. Speaking skills are not emphasized, although writing skills (termed *ʾinšāʾ*) are given some attention. Generally, modern proficiency-based techniques are not used, although there is a gradual move toward them. Rote memorization is favored over analysis, and large selections are learned by heart.

4. Scholarship

Through the centuries, a large number of Arabic books on various subjects have been composed in India (listed in several bibliographical, anthological, and descriptive studies:

Table 1. Selected texts from the Dars-i Nizāmī (full list in Qamar ad-Dīn 1996:345–352)

Qurʾān exegesis	*Tafsīr*s of al-Jalālayn, al-Bayḍāwī, az-Zamaxšarī, and Ibn Kaṯīr	*Ḥadīṯ*	Mālik, *Muwaṭṭā* Buxārī and Muslim, *Ṣaḥīḥayn*; al-Xaṭīb at-Tibrīzī, *Miškāt al-Maṣābīḥ*
Jurisprudence	al-Marġīnānī, *Hidāya fī l-furūʿ* Mullā Jīwan, *Nūr al-ʾanwār*	Sufism	ʾAbū Najīb Suhrāwardī, *ʿAwārif al-maʿārif* Ibn al-ʿArabī, *Fuṣūṣ al-ḥikam*
Logic	Quṭb ad-Dīn ar-Rāzī, *Quṭbī* Qazwīnī, *Šarḥ Šamsiyya* Taftāzānī, *Tahḏīb*	Philosophy	Jawnpūrī, *Šams al-bāziġa* Mullā Ṣadra, *Ṣadra*
History	Suyūṭī, *Tārīx al-xulafāʾ* Ibn al-Xayyāṭ, *Tārīx*	*Belles lettres*	*al-Muʿallaqāt as-sabʿ* al-Mutanabbī, *Dīwān* ʾAbū Tammām, *Ḥamāsa* Ḥarīrī, *Maqāmāt*
Rhetoric	Taftāzānī, *Muxtaṣar al-maʿānī* and *Muṭawwal* Qazwīnī, *Talxīṣ al-miftāḥ*	Grammar	Ibn al-Ḥājib, *Kāfiya* and *Šāfiya* Ḥusayn ibn Ṭawqānī, *Hidāyat an-naḥw* and *Šarḥ miʾa ʿāmil*
Mathematics	Nāṣir ad-Dīn aṭ-Ṭūsī, *Taḥrīr ʾUqlīdus* Bahāʾ ad-Dīn al-ʿĀmilī, *Xulāṣat al-ḥisāb*	Theology	Dawwānī, *Šarḥ ʿaqāʾid Jalālī* Ibn aš-Šarīf, *Musāmara*

Brockelmann 1938:309–312, 598–628, 849–864; Ahmad 1946; ʾAḥmad 1977; ʾIdrīs 1998; Schimmel 1973:1–8, 48–52; Poonawala 1977; Haroon 1996; Ishaq 1955; Kokan 1974). The majority of these works are on topics of an Islamic nature: *Qurʾān* and *Ḥadīṯ* studies, jurisprudence, Sufism, theology, and the lives of saints. They include original religious books; commentaries, glosses, and superglosses on classical religious texts; translations into Arabic from Persian Sufi works; and religious praise poetry. Several works have been composed on Islamic philosophy and history, Graeco-Arabic medicine, Arabic grammar and rhetoric, and classical-style *belles lettres*. Also numerous are Indian-language commentaries on, and translations of, Classical Arabic religious texts. Some Arabic works are by Arab immigrants, but the bulk are by scholars of Indian ethnicity, a few of them trained in Mecca or Baghdad. A large number of authors are Sufi, many of whom composed their *Xilāfat Nāmah*s in Arabic.

The best known Indian *Qurʾān* commentary is the two-volume *Tafsīr ar-raḥmān wa taysīr al-mannān* by the Navāyat scholar ʿAlāʾ ad-Dīn Mahāʾimī (d. 1431). Eminent compilations of *ḥadīṯ* include the *Mašāriq al-ʾanwār* by Ḥasan aṣ-Ṣaġānī of Lahore and the *Kanz al-ʿummāl fī sunan al-ʾaqwāl wa-l-ʾafʿāl* by ʿAlī al-Muttaqī of Burhanpur (d. 1568). A sizable *fiqh* work is the multiauthored Ḥanafī law book commissioned by Aurangzeb (r. 1754–1760), titled *al-Fatāwā al-hindiyya* (or *Fatāwā-yi ʿālamgīrī*). In theology, an important work is the *Hujjat allāh al-bāliġa* of Šāh Walī Allāh of Delhi (d. 1760). *Belles lettres* in prose includes ʾAbū Bakr ibn Muḥsin's (d. 1715) *al-Maqāmāt al-hindiyya* (study and translation, Ebeid and Young 1978).

Of the Arabic poetry composed in India, a large proportion is in praise of the prophet Muḥammad and his family. The prolific poet and author Ġulām ʿAlī Āzād Bilgrāmī (d. 1785) of Aurangabad in the South was given the honorific *Ḥassān-i Hind* 'the Ḥassān of India', after the Prophet's chief panegyrist. Several poets of the Twelver Shiʿite Deccan kingdom of Golconda in southern India – such as the Ḥijāzī poet Sayyid ʿAlī ibn Maʿṣūm (d. 1705) – spent a large proportion of their literary energies in praising ʿAlī ibn ʾAbī Ṭālib and the Shiʿa imams (Khan 1963). In western India, several of the religious leaders of the Dāʾūdī Bohra (Ṭayyibī Mustaʿlawī ʾIsmaʿīlī Shiʿite) community, especially the *dāʿī*s ʿAbd ʿAlī Sayf ad-Dīn (d. 1817) and Ṭāhir Sayf ad-Dīn (d. 1965), were notable poets who composed poetry in praise of the Prophet, the imams, and the *dāʿīs*, elegies for Ḥusayn, and poetry in communion with God, called *munājāt*.

Numerous libraries in India house extensive collections of Arabic works by Indian and Middle Eastern scholars, including tens of thousands of manuscripts, some quite valuable (S. Nadvi 1945, 1946; Desai 1978:95–125). Several of the libraries are affiliated with madrasas and universities or with shrines of saints (*dargāh*), and others are independent, either public or private. In northern India, the most important are the Rampur Raza Library in Rampur; Maulana Azad Library, Aligarh Muslim University, Aligarh; and Kutubkhana-i Nāṣiriyya (Twelver Shiʿite), Lucknow. In western India, significant libraries are Rajasthan Oriental Research Institute, Tonk; Jāmiʿa Sayfiyya Library (Dāʾūdī Bohra), Surat; and Hazrat Pir Muhammad Shah Dargah Library, Ahmedabad. Mumbai has the Jami Masjid Library, the library of Bombay University, and the Ṭayyibī Daʿwat (Dāʾūdī Bohra) Library. In eastern India, sizable libraries include the Khuda Bakhsh Oriental Public Library, Patna; the Library of the Asiatic Society of Bengal, Calcutta; and the Oriental Public Library, Bankipore. In southern India, libraries with large collections include three in Hyderabad: the Salar Jang Museum Library (Twelver Shiʿite), the State Central Library, and the Kutubkhana-i-Saidiyya.

A number of publishing houses take a special interest in publishing editions of Arabic texts. The foremost is the Dāʾirat al-Maʿārif al-ʿUṯmāniyya, Hyderabad-Deccan (founded 1888). Other publishers include university-affiliated and government-sponsored houses such as the Institute of Islamic Studies, Muslim University, Aligarh; Osmania University, Hyderabad; Government Oriental Manuscripts Library, Madras; and Government of Bihar Institute of Post-graduate Studies and Research in Arabic and Persian, Patna. The University of Lucknow, the University of Delhi, and Madras University also publish studies on Arabic works. A few publishers are associated with madrasas, such as the Dār al-Muṣannifīn (also called Shibli Academy), Azamgarh (founded 1915).

5. Usages

In addition to scholarship, there are numerous other usages of Arabic in India, mostly connected with religion: liturgy, inscriptions, nomenclature, borrowing of vocabulary and phraseology (→ Urdu/Hindi; → Bengali), and the use of the Arabic script to write Indian languages (→ alphabet, Arabic for other languages).

The foremost usage of Arabic is liturgical, primarily in the form of Qurʾānic recitation. Since most Indians do not understand the language, their recitation is, more often than not, without an understanding of the literal meaning of the verses. They nevertheless consider it a source of divine grace (*baraka*). Arabic litanies (*tasbīḥ*) are common, as is the recitation in Arabic of prayers (*duʿāʾ*), Sufi ritual chants (*ḏikr*), and religious poetry (called *qaṣīda* or *naʿt*).

Arabic is utilized for inscriptions on mosques, mausoleums, graves, madrasas, palaces, forts, and other monuments, as well as on coins (overview by Burton-Page 1986; extensive catalogs, e.g. Desai 1989, 1999; Abdur Rahim 2000; Abdul Karim 1992; N. Wright 1972a, 1972b; Whitehead 1914). The inscriptions are of an essentially religious nature, with Qurʾānic verses taking pride of place. They are used for dedications, ornamentations, and epitaphs. The earliest Arabic inscription found in India is from a 2nd/8th-century mosque in Kovalam, South India (Chaghatai 1978 from *Majalla Ṭilsānīn* I, 51).

Indian Muslims frequently adopt names of Arabic derivation, and these names often have a religious association. Names from the family of the Prophet and of the early Companions are common. Many males use Muḥammad as the first component of a compound name. Mixed Arabic and Persian or Indian language names are also found, such as Ġulām ʿAlī 'servant of ʿAlī'. Pseudo-Arabic names that are semantically difficult to fathom sometimes crop up, such as Samīyullāh 'God's namesake [?]', ʾIslām ad-Dīn 'the Islam of Religion [?]', and Qiyām ad-Dīn 'arrival of Judgment Day [?]'.

Bibliographical references

Primary sources

Abdul Karim (ed.). 1992. *Corpus of the Arabic and Persian inscriptions of Bengal*. Dhaka: Asiatic Society of Bangladesh.

Abdur Rahim, Syed (ed.). 2000. *Arabic, Persian and Urdu inscriptions of Central India: A topographical list*. New Delhi: Sundeep Prakashan.

Desai, Ziyaud-Din (ed.). 1989. *A topographical list of Arabic, Persian and Urdu inscriptions of South India*. New Delhi: Sundeep Prakashan.

—— (ed.). 1999. *Arabic, Persian and Urdu inscriptions of West India: A topographical list*. New Delhi: Sundeep Prakashan.

Ibn Baṭṭūṭa, *Riḥla* = ʾAbū ʿAbdallāh Muḥammad ibn ʿAbdallāh Ibn Baṭṭūṭa, *Riḥla Ibn Baṭṭūṭa al-musammā Tuḥfat an-nuḏḏār fī ġarāʾib al-ʾamṣār wa-ʿajāʾib al-ʾasfār*. Ed. and transl. C. Defremery and B.R. Sanguinetti, *Voyages d'Ibn Batoutah*. Paris: Chez Benjamin Duprat, 1888. (Ed., transl., and abr. Tim Mackintosh-Smith, *The Travels of Ibn Battutah*. London: Macmillan, 2002.)

Whitehead, R.B. (ed.). 1914. *Catalogue of coins in the Panjab Museum*. II. *Coins of the Mughal emperors*. Lahore and Oxford: Clarendon Press.

Wright, Nelson (ed.). 1972a. *Catalogue of the coins in the Indian Museum Calcutta*. II. *The Sultans of Delhi, contemporary dynasties in India*. New Delhi: Indological Book Corporation.

—— (ed.). 1972b. *Catalogue of the coins in the Indian Museum Calcutta*. III. *Mughal emperors of India*. Lahore and Oxford: Clarendon Press.

Secondary sources

ʾAḥmad, Jamīl. 1977. *Ḥarakat at-taʾlīf bi-l-luġa al-ʿarabiyya fī l-ʾiqlīm aš-šimālī al-hindī fī l-qarnayn aṯ-ṯāmina ʿašara wa-t-tāsiʿa ʿašara*. Damascus: Wizārat aṯ-Ṯaqāfa.

Ahmad, M.G. Zubaid. 1946. *The contribution of Indo-Pakistan to Arabic literature*. Lahore: Sh. Muhammad Ashraf. (Repr., 1968.)

Alam, Muzaffar. 2003. "The culture and politics of Persian in precolonial Hindustan". *Literary cultures in history: Reconstructions from South Asia*, ed. Sheldon Pollock, 131–198. Los Angeles and London: University of California Press.

Asani, Ali. 1998. "India". *Encyclopedia of Arabic Literature*, ed. Julie Meisami and Paul Starkey, 395–396. London and New York: Routledge.

Bahadur, Nawab Aziz Jang. 1902. *Tārīx an-nawāʾiṭ*. Hyderabad: Villa Academy. (Repr., 1976.)

Bouchon, G. 1986. "Quelques aspects de l'islamisation des régions maritimes de l'Inde à l'époque médiévale (XII–XVI^e s.)". *Purūṣārtha* 9.29–36.

Brockelmann, Carl. 1938. *Geschichte der arabischen Literatur*. Supplementband II. Leiden: E.J. Brill.

Burton-Page, J. 1971. "al-Hind: Languages". *Encyclopaedia of Islam* III, 412–428. 2nd ed. Leiden: E.J. Brill.

——. 1986. "Kitābāt (i) inscriptions (10) India". *Encyclopaedia of Islam* V, 231–233. 2nd ed. Leiden: E.J. Brill.

Chaghatai, M. Abdullah. 1978. "Khaṭṭ (iv) In Muslim India". *Encyclopaedia of Islam* IV, 1126–1128. 2nd ed. Leiden: E.J. Brill.

Desai, Ziyaud-Din. 1978. *Centres of Islamic learning in India*. New Delhi: Publications Division, Ministry of Information and Broadcasting, Government of India.

Ebeid, R.Y. and M.J.L. Young. 1978. "Arabic literature in India: Two *maqāmāt* of Abū Bakr al-Ḥaḍramī". *Studies in Islam* 15.14–20.

Farooqi, F.U. 1999. *Lucknow: A centre of Arabic and Islamic studies during the nineteenth century*. New Delhi: Falāh-e-Dārayn Trust.

Fārūqī, Zubayr ʾAḥmad al-. 1990. *Musāhama Dār al-ʿUlūm bi-Deoband fī l-ʾadab al-ʿarabī ḥattā ʿām 1400 H./1980 C.E.* [The contribution of Dār al-ʿUlūm of Deoband to Arabic literature until 1400/1980]. New Delhi: Dār al-Fārūqī.

Haroon, Mohammed. 1996. *Islamic literature: Indian contribution*. New Delhi: Indian Bibliographies Bureau.

ʾIdrīs, ʾAḥmad. 1998. *al-ʾAdab al-ʿarabī fī šibh al-qārra al-hindiyya ḥattā ʾawāʾil al-qarn al-ʿišrīn*. Cairo: ʿAyn li-d-Dirāsāt wa-l-Buḥūṯ al-ʾInsāniyya wa-l-Ijtimāʿiyya.

Ishaq, Muhammad. 1955. *India's contribution to the study of hadith literature*. [Dhaka:] University of Dacca.

Isḥāq, Muḥammad Qamar. 1996. *Hindustān ke ahamm madāris* [Important madrasas of India]. New Delhi: Institute of Objective Studies.

Kaur, Kuldip. 1990. *Madrasa education in India: A study of its past and present*. Chandigarh: Centre for Research in Rural and Industrial Development.

Khan, M.A. Muid. 1963. *The Arabian poets of Golconda*. Bombay: University of Bombay.

Kokan, Muhammad Yousuf. 1974. *Arabic and Persian in Carnatic 1710–1960*. Madras: Hafiza House.

Koya, Mohammad. 1988. *Muslims of Malabar with special reference to their distinctive character*. Ph.D. diss., Calicut University.

Metcalf, Barbara Daly. 1982. *Islamic revivalism in British India: Deobandh, 1860–1900*. Princeton, N.J.: Princeton University Press.

Miller, R.E. 1991. "Mappilla". *Encyclopaedia of Islam* VI, 458–466. 2nd ed. Leiden: E.J. Brill.

Mines, M. 1986. "Labbai". *Encyclopaedia of Islam* V, 582–583. 2nd ed. Leiden: E.J. Brill.

Nadvi, Habibul Haq (ed.). 1975. *Arabic in Pakistan*. Karachi: University of Karachi, National Congress for Promotion of Arabic in Pakistan.

Nadvi, S.A. Zafar. 1945, 1946. "Libraries during the Muslim rule in India". *Islamic Culture* 19:4.329–347, 20:1.3–20.

Poonawala, Ismail K. 1977. *Biobibliography of Ismaili literature*. Malibu, Calif.: Undena.

——. 1993. "Naitias". *Encyclopaedia of Islam* VII, 919–920. 2nd ed. Leiden: E.J. Brill.

Qamar ad-Dīn. 1996. *Hindustān kī dīnī darsgāhēn* [Religious schools of India]. New Delhi: Hamdard Education Society.

Rahman, Tariq. 2000. "The teaching of Arabic to the Muslims of South Asia". *Islamic Studies* 39:3.399–443.

Robinson, Francis. 2001. *The ʿulama of Farangi Mahall and Islamic culture in South Asia*. Delhi and Lucknow: Orient Longman.

Schimmel, Annemarie. 1973. *Islamic literatures of India*. (= *A history of Indian literature*, ed. Jan Gonda, VII:5.) Wiesbaden: O. Harrassowitz.

——. 1980. *Islam in the Indian subcontinent*. Leiden: E.J. Brill.

Shahabuddin, Syed. 2003. "Approximate Muslim population in India [2001]". *Milli Gazette* 2:18, 14 January 2003.

Wink, André. 1990–2004. *al-Hind: The making of the Indo-Islamic world*. 2nd ed. 3 vols. Leiden: E.J. Brill.

Wright, Theodore P. Jr. 1966. "Muslim education in India at the crossroads: The case of Aligarh". *Pacific Affairs* 39.50–63.

Yusuf, S.M. 1967. "Arabic language and literature in the Indo-Pakistan subcontinent". *Iqbal* 16:1.54–66.

Zaman, Muhammad Qasim. 2001. "The role of Arabic and the Arab Middle East in the definition of Muslim identity in twentieth century India". *Muslim World* 87:272–298.

Tahera Qutbuddin
(University of Chicago)

Indicative → Mood

Indirect Speech

The function of reporting speech in any language consists of a speaker conveying or reflecting what another speaker said, wrote, or thought, either in the original speaker's terms or in the reporter's words. Direct speech corresponds to the former case, by which the exact terms of the speaker are quoted, whereas indirect speech (also called 'reported speech') corresponds to the latter case. As noted by Coulmas (1986:1), the notion of 'verbatim rendition' of direct-speech utterances varies from one culture to another. Also subject to variation from one language to another are the grammatical modifications involved by the indirect-speech utterance. The main grammatical modifications involved by indirect speech generally concern the pronoun, the verbal tense, and mood, as well as the form of the question and command (Jespersen 1965:290–300). Yet, not all languages require the same rules of modification.

The question of reported speech is generally discussed by the Arab grammarians under the title of *maqāl al-qawl* 'the utterance of saying' or *al-jumla al-marwiyya* 'the reported sentence'. Two types of storytelling are identified: *ar-riwāya bi-l-lafḏ̣* 'the story told in its literal words', which corresponds to direct speech, and *ar-riwāya bi-l-maʿnā* 'the story told according to its meaning', which corresponds to indirect speech (Abdul Aziz 2003:268). The direct-speech utterance is usually introduced by a verb of saying, such as *qāla*, followed by the quotation. The quotation marks, part of a punctuation system introduced in Arabic only in the 20th century, are not thoroughly used in marking direct-speech quotations. The indirect-speech phrase is also introduced by a verb of saying, such as *qāla, ṣarraḥa, ʾakkada, zaʿama*. It can be introduced by a complementizer, such as *ʾanna*, followed by a pronoun, but it can also be asyndetic. No modifications are required in the verbal tense or mood, but only in the form of the pronoun, which shifts from 1st to 3rd person.

In texts from the Classical Arabic tradition, direct-speech utterances seem to be more frequent in comparison to indirect-speech ones. Since research on reported speech in Arabic is very scarce (Doss 2000:32), the following remarks on the ratio of direct to indirect speech in Classical Arabic texts are based on direct perusal of the sources. Concerning the *Qurʾān*, the findings concur with those of Hatim, who observes that the Qurʾānic text "abounds with examples of direct speech" (1997:124). In fact, very few occurrences of 'indirect speech' are to be found in the *Qurʾān*. When found, indirect-speech utterances are generally represented by short forms, such as the ones appearing in the examples below, and are by far outnumbered by direct quotes or direct-speech utterances. An example of indirect speech from the *Qurʾān* appears in sura *ʾIbrāhīm* (Q. 14/31), where the construction is asyndetic: *qul li-ʿibādī llaḏīna ʾāmanū yuqīmū ṣ-ṣalāta wa-yunfiqū mimma razaqnāhum* 'Tell My bondmen who believe to establish worship and spend of that which We have given them' (Pickthall 1956:189). Direct-speech phrases, however, are abundant in the *Qurʾān*, as in the following verse from sura *al-Baqara: wa-ʾiḏā qīla lahum lā tufsidū fī l-ʾarḍ qālū ʾinnamā naḥnu muṣliḥūna* (Q. 2/11) 'And when it is said unto them: Make not mischief in the earth, they say: We are peacemakers only' (Pickthall 1956:34).

This observation equally applies to texts from the Classical Arabic period, such as Ibn al-Muqaffaʿ's (d. 760) *Kalīla wa-Dimna*, where, in a random sampling of two stories, "Bāb al-qirdi wa-l-ġaylami" (The story of the monkey and the tortoise) and "Bāb al-jurḏi wa-s-sinnawri" (The story of the rat and the cat), 63 occurrences of direct speech were found, versus 5 of indirect ones, the latter all being introduced by the verb of saying *zaʿama*, as can be seen in the following example, where both forms appear: *qāla l-faylasūfu: zaʿamū ʾanna nāsikan min an-nussāki kāna bi-ʾarḍi Jurjāna wa-kānat lahu mraʾatun jamīlatun* 'The philosopher said: They claimed that a pious man lived in the land of Jurjan, and that he had a pretty wife' (Ibn al-Muqaffaʿ, *Kalīla* 240).

In Modern Standard Arabic, the two forms of reported speech (direct and indirect) are used in addition to 'free indirect speech', a form which combines the characteristics of both direct and indirect reported utterances. This form, also called 'represented speech' by Jespersen (1965:290), appears particularly in literary writings; it is not introduced by a verb of saying or by a complementizer, nor is it marked by quotation marks, but it nevertheless represents the particularity of direct speech, and it is understood by the context, as appears in the example below from Edouard Kharrat.

Unlike Classical Arabic, indirect speech utterances in Modern Standard Arabic outnumber direct ones. In a survey conducted of newspaper Arabic, Hatim (1997:123, 125) gives the figure of 98 percent for the frequency of indirect-speech utterances vs. direct speech ones. In translating news from texts written in Western languages, and particularly from English into Arabic, the indirect form is preferred (Hatim 1997:123). Indirect speech is considered more appropriate for the "pedestrian occasions such as that of reporting the news", as opposed to direct speech, which is reserved for "loftier forms of expression" (Hatim 1997:137, 138). The absence of punctuation associated with indirect speech is provided as a partial explanation for this preference. According to this explanation, punctuation has become redundant because of the dependence of Arabic on oral/heard rather than on visual effects like punctuation and diacritical marks (Hatim 1997:125).

A study based on an observation of journalistic writings (from the daily newspaper *al-Ḥayāt*) gives the following most frequent verbs of saying: *qāla* 'to say', *ṣarraḥa* 'to declare', *ʾaḍāfa* 'to add', *ʾaʿlana* 'to announce', *ʾakkada* 'to emphasize', by order of frequency out of a list of 29 verbs in 500 instances of reported statements (Al-Kasimi 1966:73). In a short unpublished search conducted by Doss on texts from the newspapers *al-ʾAhrām* and *al-ʾAxbār*, these same verbs were found to appear, as well as *ʾašāra* 'to point out', *ʾablaġa* 'to inform, notify', and *ḏakara* 'to mention'. According to Al-Kasimi (1966:75–76), verbs of saying are followed by two forms of particles, *ʾanna (ʾinna)* and *bi-ʾanna*, the first usually following the verb *qāla*, the second following the verb *ṣarraḥa*.

Free indirect speech is a stylistic device considered by some to be a "European technique" (Essawi 1996:3). It appears in contemporary Arabic literary texts, possibly as the result of the influence of literary translations. The following extract from *az-Zaman al-ʾāxar*, a novel written by the Egyptian novelist Edouard Kharrat, illustrates this stylistic device:

> *wa-humā yatabādalāni as-suʾāla ʿan al-ʾaḥwāli, wa-limāḏā lam taštarikī fī l-muʾtamari? li-ʾannahā ʿādat munḏu ʾayyāmin faqaṭ min biʿṯati taftīšin ʾuxrā fī l-wāḥāt, wa-ʿindahā lajnatun fī l-wizārati, wa-mā ʾaxbāru l-matḥafi fī l-ʾIskandariyya, wa-t-tarmīmāti l-jadīdati li-l-ḥāʾiṭi š-šarqiyy? wa-ʾidmāji ʿanāṣiri l-ʾaʿmidati llatī ktušifat fī Maryūpūlīs, hal najaḥa? wa-ʾaxbāru qirāʾātihā l-ʾaxīrati li-l-bardiy-yāti l-yūnāniyyati? wa-yaḍḥakāni min ʾaxbāri z-zumalāʾ wa-r-ruʾasāʾ al-judud fī l-maṣlaḥa, wa-yansilāni rīšahum qalīlan, bi-stimtāʿ* 'And they engage in questions about everything, and why is it that you're not participating at the Conference? Because she has returned only a few days ago from another exploration mission at the oases, and she is in charge of a committee at the Ministry. And what is the news of the Alexandria Museum, and the Eastern Wall's new restoration? And the element insertion in the columns discovered at Mariopolis, did it succeed? And news about her last reading of the Greek papyri? And they laugh, recalling the news of new colleagues and superiors, and they enjoyably 'pluck their feathers'

A hybrid form of reported speech remains to be mentioned. This involves an indirect-speech utterance abruptly interrupted by a direct one. These forms are frequent in Middle Arabic texts, as well as in contemporary journalistic writing. Both are illustrated in the following two excerpts; the first excerpt is from al-Jabartī (1756–1825; *ʿAjāʾib* III, 5.10–12):

> *fa-kallamūhum wa-staxbarūhum ʿan ġaraḍihim fa-ʾaxbarū ʾannahum ʾInkilīz ḥaḍarū li-t-taftīš ʿala l-Firinsīs li-ʾannahum xarajū bi-ʿimāratin ʿaḏ̣īmatin yurīdūna jihatan min al-jihāti wa-la nadrī ʾayna qaṣdihim fa-rubbamā dahamūkum fa-lā taqdirūna ʿalā dafʿihim wa-lā tatamakkanūna min manʿihim* 'So they [Umar Makram and high dignitaries] talked to them [the English] and asked them about their aims, they said that they were English, coming in search of the French, since these had sailed in a huge vessel toward a certain direction, and we don't know where they are planning to go: they could attack you and you would not be able to push them off or to prevent them'.

The second excerpt is from the daily newspaper *al-ʾAhrām* (February 9, 2004):

> *wa-ṭalabat minhu ʿadam al-ḥuḍūr ʾilayhā marratan ṯāniyatan bi-šiqqatihā ʿinda ḏālika janna junūnahu baʿda ʾan ġadarat bī wa-qatalat ṭumūḥī* 'And she asked him not to visit her in her flat another time; it is then that he lost his mind after she has deceived/betrayed me and killed my ambition'.

This form, which we designate as 'hybrid', is sometimes attributed to the rhetoric figure known as *iltifāt*, in which certain stylistic devices concerning mainly pronominal shifts are used to attract the attention of the reader. In this case, it would be the change of pronominal reference within the same utterance. In the example, we see the story told by the narrator (who in this situation is the journalist), who uses the third person, then shifts to the first. The construction observed can also be attributed to lack of attention and disregard for punctuation in journalistic writing (Doss 2003:191–201).

In the Arabic dialects, *ʾinna* is the only conjunction (or complementizer) used to introduce indirect speech: *biyʾūl innu nāwī ysāfir barra* 'he says he is planning to travel abroad'. However, indirect-speech utterances often appear in the asyndetic construction: *ʾultilha tīgi tāxud ḥāgitha magatš* 'I asked her to come and pick up her things, but she didn't'. In general, the conjunction *ʾinna* is less frequent in oral discourse.

Bibliographical references

Primary sources

Ibn al-Muqaffaʿ, *Kalīla* = ʿAbdallāh ibn Ḏādawayhi Ibn al-Muqaffaʿ, *Kalīla wa-Dimna*. Beirut: Dār at-Tawfīq, 1978.

Jabartī, *ʿAjāʾib* = ʿAbd ar-Raḥmān ibn Ḥasan ibn Burhān ad-Dīn al-Jabartī, *ʿAjāʾib al-ʾāṯār fī t-tarājim wa-l-ʾaxbār*. Bulaq, 1297/1878–1879.

Kharrat, Edouard, *Az-zaman al-ʾāxar*. Cairo: Dār Šuhdī li-n-Našr wa-t-Tawzīʿ, 1985.

Pickthall, Mohammed Marmaduke, *The meanings of the Glorious Koran*. New York: New American Library, 1956.

Secondary sources

Abdul Aziz, Mohammad Hassan. 2003. *ar-Rabṭ bayna l-jumal fī l-luġa al-ʿarabiyya*. Cairo: Dār al-Fikr al-ʿArabī.

Al-Kasimi, Ali. 1966. *Reported statement: A contrastive analysis of reported statement in American and British English with special reference to Arabic*. M.A. thesis, American University of Beirut.

Coulmas, Florian. 1986. "Reported speech: Some general issues". *Direct and indirect speech*, ed. Florian Coulmas, 1–28. Berlin and New York: Mouton de Gruyter.

Doss, Madiha. 2000. "Analyse comparée d'une forme hybride du discours rapporté". *al-Loġa* (Cairo) 29–56.

——. 2003. "Forms of reported speech in Arabic within the perspective of diglossia and oral writing". *Mélanges David Cohen*, 191–201. Paris: Maisonneuve et Larose.

Essawi, Hala. 1996. *Le style indirect libre dans* Madame Bovary *de Flaubert: Etude contrastive du texte français et des traductions arabes*. M.A. thesis, Cairo University.

Ḥasan, ʿAbbās. 1968. *an-Naḥw al-wāfī*. Cairo: Dār al-Maʿārif.

Hatim, Basil. 1997. *Communication across cultures: Translation theory and contrastive text linguistics*. Exeter: University of Exeter Press.

Jespersen, Otto. 1965. *The philosophy of grammar*. New York: Norton Library.

Madiha Doss (Cairo University)

Indonesia

Since the arrival of Islam in the Indonesian archipelago, Arabic has been used as a religious language for the basic rituals and for technical instruction of specialists in the religious sciences. It has, however, remained restricted to the religious realm, the major vehicle of contact being → Tamil rather than Arabic. Many Indonesian commercial terms bear witness to the influence of southeast Indian Muslims in the period before Malay developed into the lingua franca of the archipelago and, after independence in 1945, into modern standard Indonesian. During the 16th–18th centuries, Portuguese was an important medium for trade and cultural contacts in mixed ethnic groups, but from the 17th century onward, Dutch increasingly took over the role of Portuguese, while English has become important as an international language since 1945. Notwithstanding efforts by Muslim modernists in the 20th century to promote Arabic as a second language (after Malay/Indonesian) for Indonesian culture in general, Arabic has not been able to compete with the Western languages, and its use has remained restricted to the religious domain.

1. The arrival of Islam in Indonesia

The first Muslims to arrive in the archipelago were part of an expanding Islamic network of trade in the larger Indian Ocean. They were often not of Arab descent but were Gujeratis, Tamils, Bengalis, or Chinese and Champa (Vietnamese) Muslims. The arrival of Islam caused a stream of Indonesian Muslims to go to the Middle East, mostly Mecca and Medina, for religious studies, bypassing India, where until the 14th century Indonesians had studied Hinduism and Buddhism (Steenbrink 1988). With the early conversions, Arabic became known as the language of Islam, as reflected in the conversion myth of the ruler of the Acehnese kingdom of Pasai, whose tomb bears an inscription from the year 696/1297. The ruler, Merah Silau, saw in a dream the prophet Muḥammad, who spat in his mouth. When the ruler woke up, the taste in his mouth was rich and sweet. The following day, a ship arrived from Mecca that had picked up a certain Sultan Muhammad, who came to preach Islam. At his arrival, Merah Silau was already able to recite the profession of faith and the thirty *juz'* of the *Qur'ān* "although he had never been taught by anyone. At this the people and the chiefs exclaimed: As for this that our ruler is reciting, we do not understand a single word of it" (Jones 1979:134). Whether acquired through a miracle or a long process of study, the Arabic of the *Qur'ān* and other texts would remain a sacred language never truly understood by most Indonesians. Some of the conversion stories add: "But apart from that the ruler did not alter a single one of his heathen habits." (On the conversion of the ruler of Patani, see Jones 1979:143.)

Right after the first conversions, Arabic names became quite common, as many of the conversion stories show, more so in Malay culture than in Javanese, Buginese, and some other societies of the vast and diverse archipelago. Especially in the Malay world, an Arabic name was a sign of Islamic identity, like giving up idol worship and pork.

2. Arabic in literary texts

The oldest preserved literary texts in Malay were mostly, besides original chronicles of the major realms, translations from Persian rather than Arabic. The first major Islamic epic work to be translated into Malay, probably in the 15th or 16th century, is the *Hikaya Muhammad Hanafiyya* (about the half-brother of the grandsons of the Prophet, Ḥasan and Ḥusayn; cf. Brakel 1975). The stories of 'Amīr Ḥamza and Alexander the Great, and the *Book of the Thousand Questions* (about a discussion between a Jew, ʿAbdallāh ibn Salām, and the prophet Muḥammad) were also translated from Persian before 1600. There is only one major work in this tradition that was translated directly from Arabic, the *Ḥikāya 'Iskandar Ḏū l-Qarnayn*. Only in the 17th century do we find more translations directly from Arabic, such as the abstracts from Ṭabarī's *Ta'rīx* in the work of Raniri.

The first author whose mystical works in Malay have been preserved, Hamzah Fansuri (probably active around 1600 C.E.), wrote in a very refined Malay, with many loanwords from Persian and even more from Arabic. Hamzah Fansuri was a well-traveled person. He almost certainly learned Persian in Ayuthia, the ancient capital of Siam. This shows the connections between Indonesian Islam and East Asia, which are further documented in the influence of Chinese and Champa or Vietnamese Muslims on the spread of Islam in Southeast Asia. It is not certain whether Hamzah Fansuri visited Mecca or other places in Arabia, but he testified that he had been in Baghdad. His life story is reflected in his style of 'international' Malay, with many Persian and Arab words but still so elegant that it remains obligatory reading for all students in contemporary Indonesian high schools.

To write Malay, an adapted form of the Arabic script was and sometimes still is used (→ alphabet, Arabic for other languages). Major differences from the original Arabic script are an *ʿayn* with three dots for *nga*; a *fā'* with three dots for *pa*; a *nūn* with three dots for *nya*; a *kāf* with one dot on top for the letter *ga*; a *jīm* with three dots in its 'heart' for *cha*. Some regional languages, like Javanese and Buginese, continued to use a script derived from Sanskrit (Devanagari) in addition to the adapted Arabic script. Malay, written with adapted Arabic script, is called Jawi script (cf. Jones 1983), while Javanese with Arabic script is called *pegon*. In direct translations from Arabic many loanwords were introduced in → Indonesian/Malay, as well as some morphological and syntactic structures taken from Arabic. This is even seen in the older translations of the Bible by missionaries, which are full of Arabicizing tendencies. Yet, the general impression is that Arabic has had little impact on the overall structure of the Indonesian language, although about 10 percent of its vocabulary originates from Arabic, directly or indirectly (van Ronkel 1899; Landelinus 1938; Drewes 1950).

From the 17th century onward, many Islamic scholars wrote in both Arabic and local or regional languages. Šams ad-Dīn as-Samatrānī, who held a prominent position at the Acehnese court in the early 17th century and died in 1630, wrote mystical works in both Arabic and Malay. One of the major scholars of the 17th century was Nūr ad-Dīn ar-Ranirī (d. 1659), born in Gujerat, probably of mixed Arab-Malay descent. In Malay he wrote a large encyclopedic work of some 1,600 pages (*Bustanus Salatin*) and a *fiqh* work, but his doctrinal and mystical works are partly in Arabic and partly in Malay. He held a prominent position at the Acehnese court. His successor as scholar and politician at that court was ʿAbd ar-Rāʾūf as-Singkīlī (ca. 1620–1693), who stayed for 19 years in Medina, where he studied with al-Qušāšī (d. 1660), ʾIbrāhīm al-Kurānī, and others. He translated the *Qurʾān* into Malay with the full *Tafsīr al-Jalālayn* and many additions from other commentaries. This Malay translation and commentary only became popular after it had been printed at the expense of the Ottoman sultan Abdülhamid in Istanbul in 1884 and was reprinted in various places like Bombay, Singapore, Penang, and Jakarta.

Another 17th-century scholar who wrote in Arabic was Muḥammad Yūsuf al-Makassarī (1627–1699). He was born into a noble family in South Celebes that had accepted Islam in 1605. He studied Arabic and religious disciplines in his hometown of Goa in Southwest Celebes with an Arab teacher from Yemen, Bā ʿAlwī ibn ʿAbdallāh al-ʿAllāma aṭ-Ṭāhir, until the age of 15, when he went traveling to study, at the same time pursuing a political career. In 1644 he sailed to Arabia, where he first followed courses in Zabid, Yemen, and later in Mecca and Medina, where he was together with the Acehnese as-Singkīlī. Al-Makassarī even went to study in Damascus and visited Istanbul before he returned to the Indonesian archipelago between 1664 and 1672, having lived in the Arab world for at least 20 years. In the 1670s he settled in Banten, West Java, where he became the dominating political figure in the anti-Dutch party. In 1683 he was arrested by the Dutch and sent into exile to Sri Lanka. In Sri Lanka he was still considered dangerous because of the many people who came to see him. Therefore, in 1694 he was expelled to Cape Town, South Africa, where he died five years later. Yūsuf al-Makassarī wrote many mystical works, some in his native language, Makassarese, and some in Malay, but quite a few also in Arabic. Like other major 17th-century scholars, he could address his readers in the international language, Arabic, but also in some local Indonesian languages (Riddell 2001; also Azra 2004 for other 17th- and 18th-century authors).

Manuscript collections may give an idea of the spread and use of Arabic in the Indonesian archipelago. The collection of colonial Batavia (now Jakarta) is of special interest because it contains documents brought together from the whole archipelago. The catalogs of the Batavia Society of Sciences and Arts show us that with the exception of a few manuscripts with poetry and folktales, most manuscripts were devoted to the religious sciences. Out of more than one thousand Arabic manuscripts, only fifteen contain poetry, and even these are about religious topics: four copies of the *Qaṣīdat al-Burda* by al-Busīrī, three copies of the poem about the miracles performed by Muḥammad, the *Hamziyya*, etc. The collection catalogued by the Batavia Society of Sciences and Arts includes more than one hundred full or incomplete copies of the *Qurʾān*, often with interlinear translation in one of the Indonesian languages. The collection also contains more than seventy manuscripts of *ḥadīṯ*, twelve of them copies of the *Forty [ḥadīṯ] of Nawawī*, the others offering selections from the six great collections, according to specific topics. It is likely that no full text of al-Buxārī's *Ṣaḥīḥ* or of any other collections was available anywhere in the archipelago before interest in the topic rose together with 20th-century reformism. Works on mysticism were very well represented (about 250), even somewhat more frequently than those on *fiqh* or Islamic law (about 180). Besides smaller numbers of manuscripts on history (the life of the Prophet and his companions, saints, and Sufis) and on 'sciences' (mostly the calendar, astrology, astronomy, and medicine in combination with divination and amulets or charms), there is quite a large section of philology.

3. The study of Arabic

The approximately 150 texts of the Batavia Society of Sciences and Arts section on philology mostly concentrate on grammar, and more specifically the quite complicated books in use for Indonesian students who have wanted to master Arabic. There were as yet no Arabic/ Indonesian dictionaries, and only a few Arabic/

Arabic works. The total number of Arabic works is quite striking: the society had a special interest in local languages, but in reality it possessed fewer Malay than Arabic manuscripts. This underlines the fact that in more advanced religious education, Arabic texts were the rule and full translations the exception. Even interlinear translation was an important means but not the rule.

Religious works in translation or in more or less original form represent mainstream Sunnī Islam, but several of the original works, especially Javanese texts, show an aversion to strictly Arabic idiom. One of the most prominent examples is the minor prince of Surakarta, Mangkunagera IV (1811–1881), who wrote angrily in a major poem: "Many are the young people who boast of their theological knowledge / Though not yet qualified / they are in a hurry to show off, / The way they interpret the Arabic texts / is like a Sayyid from Egypt/ every time they belittle the abilities of others". In the same vein, perhaps with even more nationalist sentiment, the notorious text of the *Suluk Gatoloco*, also from the second part of the 19th century, rejects Islam as an Arab religion: "Know, the religion of Muhammad / is the religion of the Arabs. Since you invoke a foreign people, / again you're simply proven to be thieves" (references and more examples in Steenbrink 1999). These are not isolated texts, criticizing Arabic language and influence. Already in the 17th century, a Javanese poem circulated that even more specifically criticized the study of Arabic grammar, the *Suluk Wujil*, transmitted in a manuscript dated 1607, but probably much older. Stanza 88 reads in translation: "Therefore people quarrel, desiring to outperform other people. They cling to the letter, scrupulously follow the rules. Day and night they study *bayān-maʾāni, saraf, nahu*" (Poerbatjaraka 1938:172). A similar anti-Arab mood is perhaps found in the curious debate in the late 1990s about a statement made by Abdurrahman Wahid, the leader of the major organization of the mainstream Muslim leaders, *Nahdlatul Ulama*, who had stated on several occasions that Indonesians should not use the Arab greeting *as-salāmu ʿalaykum* but rather the common Indonesian expression *selamat pagi*, because there was no need to use Arabic in nonreligious, everyday use. Yet Wahid himself, the Indonesian president from October 1999 until July 2001, was not consistent in the application of this bold statement and usually started and interspersed his speeches with many phrases in Arabic.

Over the centuries, Islamic religious knowledge was more or less identical with skill in reading and even writing in Arabic. The lower-level Qurʾānic courses provided the basis for reading and reciting. The boarding school of the *pesantren* started with the teaching of Arabic, usually with the *Miʾa ʿāmil* by ʾAbū Bakr ʿAbd al-Qāhir ibn ʿAbd ar-Raḥmān al-Jurjānī (d. 1078) and commentaries on this work by various authors. Another extremely popular work was the *ʾĀjurrūmiyya* by ʾAbū ʿAbdallāh Muḥammad ibn Dāʾūd as-Sanhājī Ibn al-ʾĀjurrūm (d. 1322). Only in the 1930s did Indonesians start to write their own grammars and Arabic dictionaries. Probably the first series of basic teaching material was the four volumes of Mahmad Junus, *Peladjaran Bahasa Arab*, published in the 1930s. For dictionaries, advanced students in the 20th century could use the well-known *Munjid* by the Lebanese priest Louis Ma'lūf. The first more or less complete dictionary in 1,701 pages was compiled by the traditional scholar Ahmad Warson Munawwir of the religious school (*pesantren*) of Krapyak, Yogyakarta, in 1984 under the title *al Munawwir, Kamus Arab-Indonesia*. An important means to understand Arabic better consisted in presenting texts (first in handwriting and then, since the late 19th century, also in print) in two languages, an Arabic sentence immediately followed by a translation in one of the Indonesian languages. A special method of this learning-by-translation is the so-called *jenggot* (lit. 'beard') method: under each Arabic word, an Indonesian translation is written or printed [Fig. 1]. Many basic books still circulate in *jenggot* copies, both in standard Indonesian and in Javanese. In the late 20th century, this method came to be used also for editions of the *Qurʾān*, in which every word in Arabic was accompanied by a Latin transcription of the Arabic and a translation. Obviously, a *jenggot* text without further explanation is not very helpful.

4. ARABS IN THE ARCHIPELAGO

There were never many Arab migrants in Indonesia. Those who came (mostly in the second half of the 19th century and the early 20th century) married Indonesian women. The migrants were often Ḥaḍramī peasants who became traders or religious leaders and soon adapted to their new country (de Jonge and Kaptein 2002). In 1936, 115,535 people counted as *Vreemde Oosterlingen* 'foreign Easterners' in the Dutch colony, including women and children. Of a total population of more than 60 million in 1936, this amounted to only 0.19 percent. The Chinese, another category in the ethnic composition of the Dutch East Indies, constituted more than 2 percent. Van den Berg (1886) published a monograph on the largest group of Indonesian Arabs, who most often originated from Hadramaut, the northern mountainous section of Yemen. After the opening of the Suez Canal in 1869, their migration became more and more significant until in the early 1930s about 25 percent of the population of Hadramaut lived outside their homeland, the majority of them in Southeast Asia (de Jonge and Kaptein 2002:2–3).

Within the family, this group sometimes continued the use of Arabic, and therefore several members of this community have played an important role in diplomatic relations between Indonesia and the Arab countries. Recently, two Indonesians of Ḥaḍramī descent served as minister of foreign affairs: from 1987 to 1999, Ali Alatas (b. 1932), succeeded by Alwi Abdurrahman Shihab (b. 1949; Ph.D. in religious studies, Temple University, 1995) for the tumultuous period 1999–2000 under President Abdurrahman Wahid. Alwi's older brother, Muhammad Quraish Shihab (b. 1944), used his fluency in Arabic to study in Egypt, where he took a doctorate in Qurʾānic studies at al-Azhar University. He was minister of religion in the last year of the Suharto government, 1998–1999.

After independence in 1945, the Arab community quickly mixed with the Indonesian nation, and most of its members are no longer easily recognizable as Arabs, unlike the much larger group of ethnic Chinese, who remained much more separate from the Indonesian population.

5. The study of Arabic today

In the first decades of the 20th century, reformist teaching started to pay attention to Arabic outside the realm of religion. In Minangkabau, West Sumatra, the modernizing reformist Zainuddin Labai el-Junusi introduced Arabic schoolbooks from Egypt for history and geography. His students complained that in Dutch schools the pupils had to learn all the Dutch towns, while in the 'Arab school' all towns and villages along the Nile had to be memorized! Haji Abdulmalik ibn Karim ibn Amrullah (1908–1982, better known under his acronym Hamka), who studied with Zainuddin Labai, is the only Indonesian author who wrote about modern Arabic literature; he even got in trouble because of accusations that he had plagiarized a popular love story by Muṣṭafā Luṭfī al-Manfalūṭī (Teeuw 1979:I, 69–72). The most outspoken advocate of Arabic in Indonesian public life was Ahmad Hassan (1887–1958), himself of mixed Tamil-Indonesian descent. As leader of the reformist movement *Persatuan Islam*, he wrote a political pamphlet in 1947 in which he argued that the 90 percent of Indonesia's population who accepted Islam, a religion based on an Arab revelation, had a personal duty (*farḍ ʿayn*) to learn Arabic as well as possible: "This language should not be studied for the sake of religion alone; also philosophy, medicine, ethics, and many secular sciences can be studied in this language". Against this the lawyer and Ahmadiyyah sympathizer Hasbullah Bakry argued in 1972 that a *ḥadīṯ* orders that the prayer be performed "as we saw it being performed by the Prophet Muḥammad" and not "as we heard it". Therefore, even the *ṣalāt* may be performed in Indonesian, with the exception of a few essential Arab phrases. As an army chaplain, Hasbullah Bakry had suggested this easier style to soldiers who found it difficult to perform the ritual prayer in Arabic. He stated that Šāfiʿī law did not allow this modification, but Ḥanafī rules do (cf. Steenbrink 1974:179–188).

The debate about the role of Arabic in modern Indonesian society is also visible in official terminology. For the national Parliament, Arabic terminology has become dominant: *Majelis Permusyawaratan Rakyat* or MPR. In institutions related to religion, Arabic phrases are very often used, but the Department of Education usually prefers neologisms rooted in Sanskrit. The state ideology is called *Pancasila* 'Five Pillars', and its introductory course for all government officials is called *Ekaprasetya Pancakarsa* 'Five Noble Works for One Final Goal', a quite mysterious expression for modern Indonesians, taken from Sanskrit. The army, defender of national unity and free from religious bounds, in its official terminology also often borrows new words from Sanskrit. Against this tendency Arabic loanwords are constantly produced in Islamic religious circles, like the terminology for the State Islamic Institute, the major academic training center in Islam: *al-Jāmiʿa al-ʾIslāmiyya al-Ḥukūmiyya*.

Indonesian education, from kindergarten through the secondary level, is divided into two types. One system, under the Ministry of Education, is called *sekolah* after the Dutch word *school*. This secular education continues the system started by the colonial government after 1900. In modern Indonesian schools religious education is obligatory at all levels, but restricted to about two hours per week. The second type, the *madrasah* system, falls under the Ministry of Religion. It provides primary and secondary education for about 15 percent of the children in this age group, and its schools pay much more attention to religious matters; this automatically means attention to Arabic, which is taught intensively at all levels. Religious (and Arabic) classes constitute 25–40 percent of the curriculum of the *madrasah*. In the year 2000–2001, nearly 3 million pupils at the primary school level were enrolled in the *madrasah* system, while 1,888,576 students were enrolled in the junior high schools and 576,221 students in the senior high schools. The *madrasah* system is supported by a third type of schools, the *pesantren*. These latter are boarding schools, often including a secondary *madrasah*. Indonesian students who follow the *madrasah* system at the secondary level in the boarding facilities of the *pesantren* sometimes attend courses restricted to reading and memorizing the *Qurʾān*. At other institutions, several thousand students acquire the ability each year to read Arabic books on religion. There are about 6,000 larger *pesantren* in Indonesia that provide lodging for about 700,000 students, as well as a favorable atmosphere for religious studies, including the study of the Arabic language. The system is often criticized as irrelevant for modern Indonesian society. Therefore, since the 1970s both government and nongovernmental organizations have introduced many agricultural and technical training programs into this type of education. The *pesantren* is also seen as a cheap alternative for dropouts from the secular system. In general, however, the debate about this type of education has not changed much since the 1950s. In 1953, Kiyahi Haji Abdul Wahid Hasjim, minister of religion and father of the later president Abdurrahman Wahid, stated that "matters of Islam and Arabic are two different fields. Of course, I agree that the religion of Islam must be acquired with the help of the language that has brought it, Arabic. But I do not agree that people try to 'Arabicise' our society by the use of Arabic language and customs, that are different from our Indonesian language and customs" (Steenbrink 1974:192). Another scholar, Nurcholis Madjid (b. 1939), studied between 1955 and 1960 at the Darul Salam Pesantren of Gontor, where it is obligatory to use Arabic (besides some English, but no Indonesian) in class and also at leisure time during meals and sports. As a cultural and political leader, Madjid propagated the idea that Islam in Indonesia should be 'secularized' in the sense of 'finding a local expression, in a local context'.

English has become by far the most popular foreign language in Indonesia. In the 1980s and 1990s, many Muslim organizations complained that state television regularly gave English courses and even news in English but paid no attention to the study of Arabic. The complaints did not help. From the very start, the use of Arabic was restricted in Indonesia to the religious domain, and this has not changed. Books in Arabic, in a mixture of Arabic and Indonesian languages, or simply in one of the Indonesian languages with Arabic script, are still published and used. They are called 'yellow books' (*Kitab Kuning*) because they are popular only when printed on yellowish paper (cf. van Bruinessen 1990).

In the 1990s and the early 21st century, an orthodox style of religious school developed, called *pesantren salafiyah*. Its main characteristic is the use of 'classical textbooks', written in Malay or modern standard Indonesian but printed with Arabic characters or, for the higher level of learning, in full Arabic. When *šarīʿa* law was promulgated as the law for the province of Aceh on 15 March 2002, it generated little change in the region. There was much stress on Islamic clothing and limitation of free movement for women. One of the effects was the adoption of Arabic script to indicate the function of the various offices in government buildings. Yet, Arabic not only remained restricted to the religion of Islam, it was increasingly identified with its more traditional and orthodox aspects.

BIBLIOGRAPHICAL REFERENCES

Azra, Azyumardi. 2004. *The origins of Islamic reformism in Southeast Asia: Networks of Malay-Indonesian and Middle Eastern ʿulamāʾ in the seventeenth and eighteenth century*. Leiden: KITLV.

Berg, L.W.C. van den. 1886. *Le Hadhramout et les colonies arabes dans l'archipel indien*. Batavia: Imprimérie du Gouvernement.

Brakel, Lode Frank. 1975. *The* Hikayat Muhammad Hanafiyyah: *A medieval Muslim Malay romance*. The Hague: Nijhoff.

Bruinessen, Martin van. 1990. "Books in Arabic script used in the *pesantren* milieu". *Bijdragen tot de Taal-, Land-, en Volkenkunde* 146.226–269.

Drewes, G.W.J. 1950. "De herkomst van het voegwoord *bahwasanja*: Bijdrage tot de kennis van het Kitab-Maleis". *Bingkisan Budi: Een bundel opstellen door vrienden en leerlingen aangeboden aan Dr. Philippus Samuel van Ronkel*, 104–116. Leiden: Sijthoff.

Jones, Russell. 1979. "Ten conversion myths from Indonesia". *Conversion to Islam*, ed. Nehemia Levtsion, 129–158. New York: Holmes and Meier.

——. 1983. *Hikayat Sultan Ibrahim*. Dordrecht: Foris.

Jonge, Huub de and Nico Kaptein. 2002. *Transcending borders: Arabs, politics, trade and Islam in Southeast Asia*. Leiden: KITLV.

Landelinus, P. 1938. "Arabismen in het katholieke taaleigen". *Koloniaal Missietijdschrift* 21.204–210.

Meuleman, Johan M. 1994. "Arabic in Indonesia". *Indian Journal of Applied Linguistics* 20.13–34.

Poerbatjaraka. 1938. "De geheime leer van Soenan Bonang (Soeloek Woedjil)". *Djawa* 18.321–361.

Riddell, Peter. 2001. *Islam and the Malay-Indonesian world: Transmission and responses*. London: Hurst.

Ronkel, Philippus Samuel van. 1899. "Over de in-vloed der arabische syntaxis op de maleische". *Tijdschrift van het Bataviaas Genootschap* 41.498–528.

Steenbrink, Karel A. 1974. *Pesantren, madrasah, sekolah: Recente ontwikkelingen in Indonesisch Islamonderricht* [*Pesantren, madrasah, sekolah*: Recent developments in Indonesian Islamic education]. Meppel: Krips Repro.

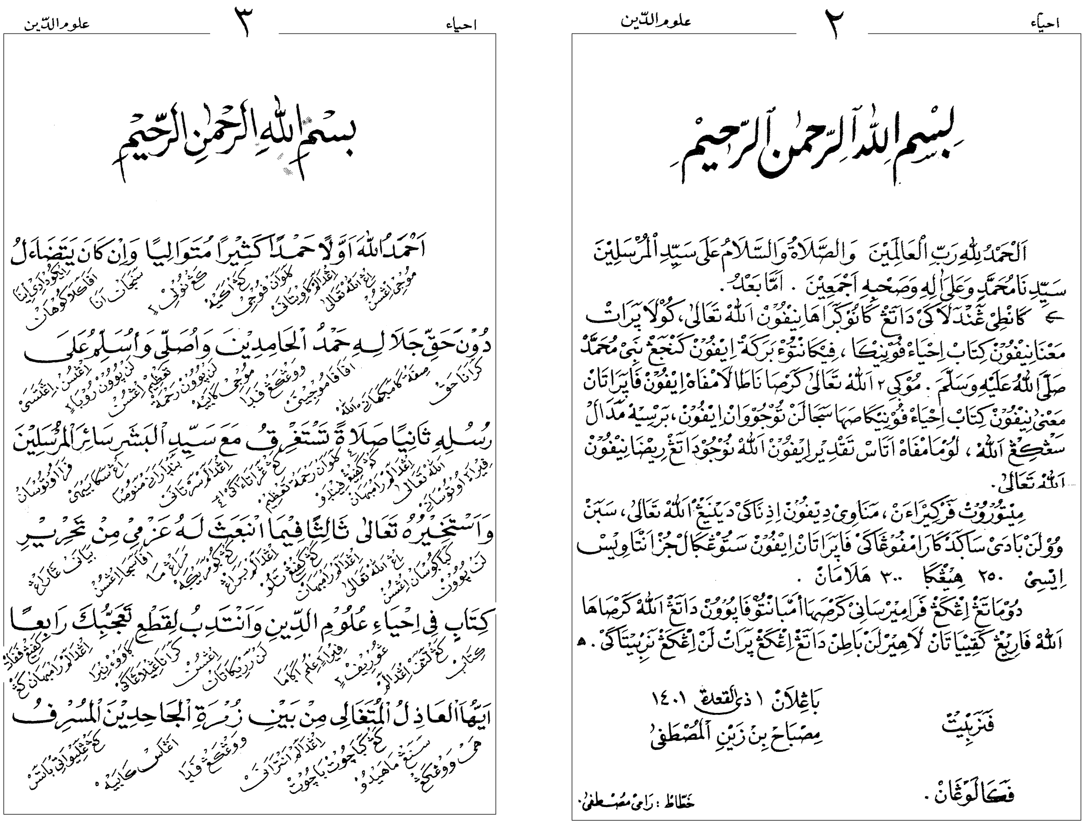
——. 1988. "Indian teachers and their Indonesian pupils: On intellectual relations between India and Indonesia, 1600–1800". *Itinerario* 12.129–142.

——. 1999. "Opposition to Islamic mysticism in nineteenth-century Indonesia". *Islamic mysticism contested*, ed. Frederick de Jong and Berndt Radtke, 687–704. Leiden. E.J. Brill.

Teeuw, A. 1979. *Modern Indonesian literature*. 2 vols. The Hague: Martinus Nijhoff.

Karel Steenbrink (Utrecht University)

Figure 1. *Kitab jengot* (litt. 'beard-book'), first page of Abu Hamid al-Ghazali, *Ihya*, in Arabic, with below each word the Javanese translation by Kiyahi Haji Misbakh ibn Zain Mustafa Bangilani. Published ca. 1988 in Pekalongan: Maktab Raja Murah.



احياء ٢ علوم الدين

بسم الله الرحمن الرحيم

الحمد لله رب العالمين والصلاة والسلام على سيد المرسلين
سيدنا محمد وعلى اله وصحبه اجمعين . اما بعد .
كانطى ڠندلاكى داتڠ كانوگراهانيفون الله تعالى، كولا ڽرات
معنانيفون كتاب احياء فونيكا ، فيكانتوء بركة ايفون كنجڠ نبي محمد
صلى الله عليه وسلم . موگى ٢ الله تعالى كرصا ناطالامفاه ايفون فايراتان
معنى نيفون كتاب احياء فونيكا صهاسجالن توجووان ايفون، برسيه مدال
سڠكڠ الله ، لومامفاه اتاس تقدير ايفون الله نوجود اتڠ ريضانيفون
الله تعالى .
مينوروت فركيراءن ، مناوي ديفون اذناكى ديڠ الله تعالى، سبن
وولن بادى ساكد كارامفوڠاكى فايراتان ايفون سوڠكال جزانتاويس
اينسى ٢٥٠ هيڠڬا ٣٠٠ هلامان .
دوماتڠ اڠكڠ فراميرساني كرصها امباننتوفايوون داتڠ الله كرصاها
الله فاريڠ كقينياتان لاهيرلن باطن داتڠ اڠكڠ ڽرات لن اڠكڠ نربيتاكى .

باڠلان ١ ذي القعدة ١٤٠١
مصباح بن زين المصطفى

فنربيت
فكالوڠان .

خطاط : راعي مصطفى .

احياء ٣ علوم الدين

بسم الله الرحمن الرحيم

الحمد لله اولا حمدا كثيرا متواليا وإن كان يتضاءل
دون حق جلاله حمد الحامدين واصلي واسلم على
رسله ثانيا صلاة تستغرق مع سيد البشر سائر المرسلين
واستخيره تعالى ثالثا فيما انبعث له عزمي من تحرير
كتاب في احياء علوم الدين وانتدب لقطع تعجبك رابعا
ايها العاذل المتغالي من بين زمرة الجاحدين المسرف

Indonesian/Malay

1. DEFINITION

Indonesian and Malay (officially known as Bahasa Indonesia and Bahasa Malaysia) contain hundreds of loanwords of Arabic origin. Many early loans came via intermediary languages influenced by Persian, if not Persian itself, while recent loans come directly from Arabic. Most of the loanwords are connected in some way to Muslim life, belief, and practice. Arabic loanwords mostly function as nouns in Indonesian and Malay, and few Arabic morphological rules have been borrowed. Loans are assimilated to the Indonesian/Malay phonological patterns, for example with the virtual loss of the plain/emphatic distinction in consonants. Malay was routinely written in a Persianized Arabic script until the end of the 19th century, and the script is still taught in Malaysian schools. Very little colloquial Arabic was borrowed; rather, loans are generally learned forms. The Arabic loanwords are symbolic of the radical shift in the Malay cultural sphere occasioned by the adoption of Islam and are used to convey ideas that were absent in the Hinduized and animist cultures of the region before Islam.

2. HISTORICAL BACKGROUND

The Arabic loanwords in Indonesian and Malay are a consequence of the Islamization of the Malay world. From the late 13th century, accounts by travelers as well as gravestone inscriptions point to the gradual spread of Islam first in North Sumatra, next in northeast Malaya and East Java, and, by the 15th century, in Malacca. While the main source of Muslim influence was from the west, there are claims of Muslims reaching the Malay world from the north via southern China and via Champa (Fatimi 1963; Chen Da-Sheng 1992). Islam is said to have been brought to the Malay world by traders and missionaries, and there are indeed accounts of intermarriage between foreign Muslims and local women, local dynasties headed by foreign Muslims, and conversion of local rulers to Islam. Indonesian and Malay chronicles recount stories of conversions and miraculous events surrounding the adoption of Islam, although the relatively recent manuscript versions of these old tales make them less than reliable as primary historical sources. Suggestions that Arabs were the sole vectors of Islam (e.g. Shellabear 1901:75) have been disputed by Drewes (1968), and it is generally agreed that Arab, Persian, and Indian Muslims must have played a role. Al-Attas (1969), while insisting that Hadramaut was the source of the early missionaries, concedes that they may have come via India. Indeed, the etymology of the Arabic loanwords in Indonesian and Malay is a crucial part of the historical evidence for the Islamization of the region.

Arabic loanwords in Indonesian and Malay show evidence of layers of borrowing over time (Bausani 1974; Campbell 1996; Versteegh 2003). Conclusive periodization and identification of the layers is problematic, but there is agreement that an early layer dated from the arrival of Islam and involved borrowing directly from Arabic and from other, evidently Persianized, Muslim languages, and a more recent and continuing layer has borrowed directly from Standard Arabic. Versteegh (2003) has proposed four elements of the older layer, i.e. words deriving from the South Arabian, North Indian, South Indian, and Chinese connections.

There is little reliable historical evidence about early language contact between the Arabic language and Malay speakers that would help to answer questions such as: (a) Were those who imported Arabic native speakers of Arabic or of something else? (b) If those people were native speakers of Arabic, did they speak with Malays in some form of colloquial Arabic? (c) Was there any Arabic-Malay bilingualism among the Malays, for example among the offspring of mixed marriages? (d) How important was translation as a vector of Arabic loanwords? In this case the Arabic loanwords must speak for themselves: (a) Many Arabic loanwords carry the phonological traces of Persianized languages, indicating that at least some of those who introduced Arabic were not native speakers. (b) There are very few unequivocally colloquial Arabic loanwords in Indonesian and Malay, suggesting that colloquial Arabic was not the medium of discourse with Malays. (c) If there was any bilingualism among the foreigners it was short-lived and left no obvious traces in Malay; otherwise, the loan stock would contain such evidence as a wider range of word classes and more examples of colloquial Arabic. (d) The semantic and grammatical characteristics of the Arabic loanwords confirm that many of them functioned as neologisms in Malay texts with Islamic themes or in translations of such texts from other languages into Malay (see Teeuw 1959:150 for a discussion of both types of text). The language contact situation with regard to Arabic seems unlike that of Chinese. Jones (2003), extending to the Malay context Van Coetsem's (1988) theory of two types of loanword transfer, argues that Chinese loanwords probably entered Indonesian and Malay via Chinese speakers who were more proficient in Malay (the recipient language) than Chinese (the source language). Using this framework we might argue that while some early loans may have been introduced by non-Malays who were more proficient in Arabic (the source language) than Malay (the recipient language), on the whole the loanwords were (and continue to be) introduced by Indonesian and Malay native speakers with a range of proficiency in Arabic as a learned language.

Our understanding of more recent borrowing from Arabic is much clearer, especially in the late 19th and early 20th centuries, when Malay began to be seen as an essential element of nation building and provided the foundation for Bahasa Indonesia and Bahasa Malaysia. Abas (1987:59) summarizes the development of these two national languages, including the conversion to Roman orthography (Abas 1987:104) and procedures for coining new terminology (Abas 1987:125). Lexicographers and language planners have proposed guidelines for adopting and spelling foreign loanwords, including Arabic words (e.g. *General guidelines for the formation of terms in Malay* 1992:20), and the etymology of Arabic words has been examined and recorded by native lexicographers and linguists (e.g. Ngajenan 1987). New Arabic loanwords are much more likely to appear in Bahasa Malaysia than in Bahasa Indonesia, and Jawi script is actively taught in Malaysian schools. Large Malay colonies evolved in Arabia itself in the 20th century (Meulen and Von Wissman 1932; Matheson and Hooker 1988:13–14), and they have presumably provided a channel of recent influence directly from Arabic to Malay rather than via an intermediary language.

The size and currency of the Arabic loan stock is a source of contention. Beg (1979:81) mentions eight scholarly works between 1801 and 1975 that have provided estimates ranging from 150 to 1,125 items, and claims that there are about 1,000 such words in current use (Beg 1979:78). Jones (1978) lists 2,750 Arabic and Persian loans, but on testing the list with three young university lecturers in Indonesia, it was found that they could only recognize about 10 percent of the words (Campbell 1996:26). On the whole, authorities may be more or less liberal with what they consider to be Arabic loanwords; Kasimin (1987) includes many new and unassimilated loans in his lists. (Unattributed examples of loanwords in this entry are adapted from Echols and Shadily 1992.)

3. The linguistic characteristics of Arabic loanwords

From the semantic point of view, Arabic loanwords tend to refer to abstract concepts rather than concrete objects, and most loanwords are concerned with aspects of Muslim belief and practice, or with the arts and sciences that accompanied Islam to the Malay world. Semantic classifications have been proposed by Jones (1984), Beg (1979), Tham Seong Chee (1990), and others. Beg, speaking of Malaysia rather than Indonesia, mentions the categories of "knowledge and science, architecture, nomenclature, greetings, feasts and festivals, state affairs, economy, common feelings and sentiments, canon law (Shariʿah), and rites and rituals of religion" (Beg 1979:89). Beg's work also includes an interesting discussion of Arabic loanwords in Malay proper names, historic coins of Malaya, and Malay state anthems – again with the focus on Malaysia (Beg 1979:90–93). Tham Seong Chee (1990:86–93) proposes a cognitive classification of Arabic loanwords. A 'central domain' includes words such as *Allah* 'God', *firdaus* 'paradise' < *firdaws*, *nabi* 'prophet' < *nabī*, and *wahyu* 'revelation' < *waḥy*. There are then subdomains: Islamic jurisprudence, e.g. *dakwa* 'accuse' < *daʿwā*; Muslim worship, e.g. *kalimah* 'creed' < *kalima*; Islamic mysticism, e.g. *makrifat* 'enlightenment' < *maʿrifa*; Islamic rituals, e.g. *khatan* 'circumcision' < *xatn*; family life, marriage, and inheritance, e.g. *nikah* 'wedding' < *nikāḥ*; Muslim education, e.g. *makalah* 'essay' < *maqāla*; and Islamic institutions or institutional ideas, e.g. *mahkamah* 'court' < *maḥkama*.

Jones (1984:13–17) proposes the following semantic fields: Islamic religion, including examples such as *balig* 'of the age of legal maturity' < *bāliġ* and *lahad* 'the hollow made in a grave on the *qibla* side' < *laḥd*; abstract and philosophical terms, e.g. *adil* 'just' < *ʿādil* and *ilmu* 'science' < *ʿilm*; euphemisms, e.g. *hamil* 'pregnant' < *ḥāmil*; political and military, e.g. *daerah* 'district' < *dāʾira* and *saif* 'sword' < *sayf*; botanical and zoological, e.g. *tufah* 'apple' < *tuffāḥ* and *zaitun* 'olives' < *zaytūn*; anatomy and medicine, e.g. *kulup* 'foreskin' < *ġulfa* and *haid* 'menses' < *ḥayḍ*; times, dates, and numerals, e.g. *Muharram* 'first lunar month' < *muḥarram* and *Selasa* 'Tuesday' < *yawm aṯ-ṯalāṯāʾ*; education, books, and writing, e.g. *murid* 'pupil' < *murīd* and *kalam* 'pen' < *qalam*; cultural innovation, e.g. *kursi* 'chair' < *kursī* and *salju* 'snow' < *ṯalj*.

There are some instances of semantic change or specialization. While Arabic *kitāb* is a general word for 'book', *kitab* refers to a religious book, with the English loan *buku* used as the general word; while Arabic *kalima* means 'word', *kalimat* is used for 'sentence', with the indigenous words *kata* or *perkataan* used for 'word'. There are a number of etymological doublets such as *fardu* 'obligation under Islam' and *perlu* 'necessary', both from *farḍ*; similarly, *syajarah* 'tree' and *sejarah* 'history' are derived from *šajara* (Jones 1984:17).

With regard to word classes, Indonesian and Malay have generally borrowed nouns and nominal forms derived from verbs. Among the borrowed Arabic verbs are *yakni* 'that is' < *yaʿnī* 'it means' and *nukil* 'quotation' < *nuqila* 'it was transmitted'. Arabic verbal nouns are a source of loans, such as *iklan* 'advertisement' < *ʾiʿlān*; *tafakur/tafakkur/tafkur* 'contemplation' < *tafakkur*; *mufakat/muafakah/muafakat* '(to hold) discussion to make agreement' < *muwāfaqa*; *tertib* 'order' < *tartīb*; *bina* 'building, construction' < *bināʾ*. The *fāʿil* form of the Arabic verb provides loans such as *arif/arip* 'learned' < *ʿārif* and *hadir* 'to be present' < *ḥāḍir*, while the *mafʿūl* form gives us *maklum* 'to know' < *maʿlūm* and *masyhur* <well known < *mašhūr*. While mostly nominal forms have been borrowed, Indonesian and Malay have where necessary grammatically reclassified words. Note among the foregoing examples the shift from Arabic nominal form to Indonesian/Malay verb in *mufakat*, *hadir*, and *maklum*.

Besides single words, Indonesian and Malay also include numerous loan phrases, apparently of two types. There are formulaic expressions of everyday Muslim discourse such as *astagfirullah* 'God forbid' < *ʾastaġfiru llāh* and *alhamdulillah* 'Praise be to God' < *al-ḥamdu lillāh* (Jones 1984:16); and there are borrowed *ʾiḍāfa* phrases which are lexicalized in Indonesian and Malay, e.g. *usuludin* 'one of the branches of theology pertaining to law and philosophy' < *ʾuṣūl ad-dīn,* and *malaikatulmaut* 'angel of death' < *malāʾikat* [pl.] *al-mawt*.

Two phenomena are relevant to the morphology of Arabic loanwords: the extent to which Arabic morphological processes have been borrowed, and the extent to which loanwords have been assimilated into the morphological system of Indonesian and Malay. The derivational and syntactic morphology of Indonesian and Malay is quite different from that of Arabic; derived forms in Indonesian and Malay arise from prefixing and suffixing stems, and there is virtually no inflectional morphology. Arabic loanwords mostly occur in Indonesian and Malay as unanalyzed, lexicalized forms. Nouns, for example, are usually borrowed in their singular forms, and where broken plurals do occur they are likely to be lexicalized as singular forms, e.g. *ulama* 'scholar of Islam', routinely used as a singular despite its origins in the Arabic broken plural *ʿulamāʾ*. While there is an Indonesian/Malay word *alim* which is derivable from the Arabic singular *ʿālim* (as well as from *ʿalīm*), *alim* does not function as the singular of *ulama*. In all likelihood, Arabic singular nouns and Arabic plural nouns have been mostly borrowed as unrelated items.

While a range of derived Arabic forms have been borrowed, these being mostly verbal nouns and participial forms, the Arabic root + measure derivational system is entirely unproductive in Indonesian and Malay loans. Thus, while words like *takrif* 'definition' < *taʿrīf*, *arif,* and *makrifat* coexist, they are lexical isolates rather than members of a paradigm. The forms of the feminine ending *tāʾ marbūṭa* in Indonesian and Malay loans differ in distribution from Arabic. While Arabic *tāʾ marbūṭa* has a citation form *-a(h)* and a construct state form *-at*, Indonesian/Malay has Arabic loans with either *-ah* or *-at*, or both spellings, e.g. *noktah* 'dot' < *nuqṭa*; *saat* 'moment' < *sāʿa* 'hour'; *ibadah/ibadat* 'act of devotion' < *ʿibāda*.

Verhaar (1984) identifies three Arabic affixes that have some degree of morphemic status. The definite article *al-* occurs in many Arabic loanwords and, if not productive, is at least identifiable as a morpheme; the oblique masculine plural suffix *-īn* appears in *Muslimin* < *muslimīn* 'the Muslims' and *hadirin* < *ḥāḍirīn* 'those present', but it cannot be said to be productive. The feminine plural suffix *-āt* similarly appears in *Muslimat* < *muslimāt* 'Muslim women' and *hadirat* < *ḥāḍirāt* 'those women present'. The *-ah* variant of the feminine ending *tāʾ marbūṭa* is found in a small number of words with natural feminine gender such as *almarhumah* < *al-marḥūma* 'the female deceased' alongside *almarhum* < *al-marḥūm* 'the male deceased'. Other perhaps morphemically identifiable but unproductive suffixes are *-iah* in *ilmiah* 'scientific' < *ʿilmiyya* and *harafiah* 'literal' < *ḥarfiyya,* and *-i* in words such as *abadi* 'everlasting' < *ʾabadiyy* and *falsafi* 'philosophical' < *falsafī* (Verhaar 1984:22). No gender distinction is signaled in Indonesian and Malay by *-iah* and *-i*.

Traces of the Arabic noun inflections remain, notably in construct phrases like *malaikatul-maut*. A handful of loanwords end in *-u* and/or *-i*, e.g. *napsu*/*nafsu* 'natural appetite or desire' < *nafs*, *perlu*, *wahyu*, *salju*/*salji*. While some scholars have suggested that the *-u* ending is evidence of a South Indian intermediary source, Versteegh (2003) has more plausible explanations, e.g. naive attempts to emulate case endings or a reflex of the 3rd person masculine suffix *-hu*.

Many Arabic loanwords in Indonesian/Malay undergo modification through prefixation and/or suffixation. An example is *khusus* 'special' < *xuṣūṣ*, which appears in the derived forms *khususnya* 'in particular', *mengkhususkan* 'set aside', *terkhusus* 'particularly', *kekhususan* 'specific characteristics of', *pengkhususan* 'specialization'. Less transparent because of phonological assimilation are derived forms of *pikir* 'opinion, idea' < *fikra*, e.g. *berpikir* 'to think', *berpikiran* 'to have a certain thought', *mempikir* 'to think about, worry over', *mempikirkan* 'to think about, meditate over', *terpikir* 'to come to one's mind', *terpikirkan* 'can happen to be thought about', and others. The accretion of derived forms is a likely indicator of the age of Arabic loanwords; the more derived forms evidence older loans.

From the perspective of phonology, Arabic loanwords in Indonesian and Malay lose their vowel length distinction and the distinction between plain and emphatic consonants, between alveolar stops and interdental fricatives, and between the glottal stop and the voiced pharyngeal fricative. Gemination of consonants is not usually represented in loanwords. The phonology of Arabic loanwords in Indonesian/Malay reflects a complex history of language development, and only the broad issues can be summarized here. Among the phenomena to influence the way Indonesian and Malay speakers pronounce Arabic loanwords are the following:

i. Layers of borrowing. There are differences between the way that recent and older borrowings represent some Arabic phonemes. Arabic /ḍ/, for example, corresponds to Indonesian/Malay /l/ or /dl/ in some older loans, e.g. *perlu*, *Nahdlatul Islam* 'Association of Muslim Scholars' < *nahḍat al-ʾislām*. Elsewhere /ḍ/ occurs as /d/, e.g. *darurat* 'emergency' < *ḍarūra*. Versteegh (2003) argues that the /l/ and /dl/ reflexes are evidence of a very old, pre-Persianized layer of borrowings.

ii. Degrees of assimilation. Some Arabic loans are more assimilated than others, e.g. *pisah* 'legal separation of a couple' < *fasx*, cf. *khurafat* (pronounced *kurafat*) 'myth' < *xurāfa*.

iii. Change of script. Toward the end of the 19th century, Jawi script was replaced by Roman orthography, with the loss in writing of some Arabic phonemic distinctions and the confusion of others. One such area of confusion is the glottal stop and the voiced pharyngeal fricative, for which the new orthography introduced the letter *k* in some contexts, perhaps with the intention that it be pronounced as an unreleased velar stop to approximate a glottal stop. Although the Arabic glottal stop /ʾ/ is omitted in some loanwords, e.g. *masalah* 'problem' < *masʾala*, it occurs in others as *k*, e.g. *mukmin* 'the believers' < *muʾmin*. However, *k* is also commonly used in some words as the counterpart of the voiced pharyngeal fricative /ʿ/, e.g. *yakni*, although in other contexts /ʿ/ may be omitted, e.g. *saat*.

The Trengganu Stone, which contains the earliest inscriptions in the Malay language written in Arabic script, dates the adoption of the Arabic or Jawi writing system to no later than the 14th century. A tradition of fine Malay manuscripts in Jawi script dates back to the early 17th century (see Gallop 1991). The principles and practice of writing in Jawi are popular pursuits in Malaysia, and works such as *Daftar ejaan Rumi-Jawi* (1994) exist to guide the student. Many Arabic loanwords are written in contemporary Malaysia as they are written in Arabic. However, highly assimilated loanwords such as *perlu* are written as pronounced, not as their Arabic reflexes; some possibly older loans have Malayized spellings with additional vowels, such as *mukmin*, which has an additional *yāʾ* before the *nūn*. The *tāʾ marbūṭa* is spelled somewhat inconsistently. No vowel pointing is used in Jawi, and it is presumed that Malaysian readers of Jawi infer the voweling of Arabic loans primarily from their knowledge of the Roman spelling or from their knowledge of Arabic per se. For the serious scholar, *Pedoman transliterasi huruf Arab ke huruf Rumi* (1992) has instructions on writing Arabic words in a detailed Roman transcription (→ alphabet, Arabic for other scripts).

4. ROLE OF INTERMEDIARY LANGUAGES AND COLLOQUIAL ARABIC

In a historiographical analysis of the evidence for the coming of Islam to the region, Drewes (1968) cites sources claiming Muslim influence from Bengal, Gujerat, and South India. From a linguistic perspective, we need to exercise caution in looking for evidence of North and South Indian languages in the forms and meanings of Arabic loanwords. The most reliable conclusion that can be drawn is that numerous Arabic loanwords have come to the Malay world via languages influenced by Persian or via Persian itself; this conclusion is made on the basis of the distribution of the *-at* and *-ah* variants of *tāʾ marbūṭa*, which corresponds to the distribution in Persian or Persianized Indian languages (Perry 1991:158). Campbell argues that the distribution of *tāʾ marbūṭa* endings proves the existence of an early Persianized loan stock, but he warns of the difficulty of untangling the Indian provenance of such words (Campbell 1996:37).

As to the possibility of colloquial Arabic influence, Versteegh (2003) cites a few examples of words with /g/ for /j/, e.g. *gamal* < Egyptian Arabic *gamal* 'camel' and *gengsi* 'prestige, status' possibly from Egyptian Arabic *gins* 'race', as possible evidence of an Egyptian source. However, there is no trace of the high frequency and characteristically urban colloquial verbs such as *šāf* and *rāḥ*, which would indicate a more extensive and systematic evidence of colloquial Arabic influence.

5. CULTURAL DIMENSIONS OF BORROWING FROM ARABIC

The Arabic element in Indonesian and Malay is a potent symbol of Islam in the region and its connection with the wider Muslim world; indigenous scholars see in the Arabic loanwords the evidence of a fundamental shift in Malay culture resulting from Islamization. Arabic names are especially important: Beg (1979:81) notes that each of the states of Malaysia has an Arabic name as well as its Malay name, e.g., Kedah is also known as *darul-amān* 'the abode of peace'. He talks of the three basic types of Malay proper names: Arabic names, e.g. Hamid bin Dollah (i.e. Abdullah); combined Arabic/Malay names, e.g. Yusof bin Long; and names such as Mastom bin Tumingam which are Malay but nevertheless contain words like *bin* 'son of' and *binti* 'daughter of' (Beg 1979:91). Tham Seong Chee (1990:82) is especially interested in the cognitive restructuring of the Malay world that he claims was triggered by the arrival of Islam, and remarks in relation to the acceptance of Muslim names that "no other area of Malay life exemplifies such an extensive cultural-psychological transformation". A third indigenous commentator should have the last word: Al-Attas says of the "Islamized Malay language" that "it would be there that the revolutionary changes in world view...would be preserved and effected" (1969:22).

BIBLIOGRAPHICAL REFERENCES

Abas, Husen. 1987. *Indonesian as a unifying language of wider communication: A historical and sociolinguistic perspective.* Canberra: Australian National University Research School of Pacific Studies.

Al-Attas, Syed Naguib. 1969. *Preliminary statement on a general theory of the Islamization of the Malay-Indonesian archipelago.* Kuala Lumpur: Dewan Bahasa Dan Pustaka.

Bausani, Alessandro. 1974. "Notes sur les mots persans en malayo-indonésien". *Acta Iranica*, première série, 2.347–379.

Beg, Muhammad Abdul Jabbar. 1979. *Arabic loanwords in Malay*. Kuala Lumpur: University of Malaya Press.

Campbell, Stuart. 1996. "The distribution of *-at* and *-ah* endings in Malay loanwords from Arabic". *Bijdragen tot de Taal-, Land- en Volkenkunde* 152:1.23–44.

Coetsem, Frans van. 1988. *Loan phonology and the two transfer types in language contact*. Dordrecht and Providence, R.I.: Foris Publications.

Daftar ejaan Rumi-Jawi [Roman-Jawi spelling list]. 1994. Kuala Lumpur: Dewan Bahasa dan Pustaka.

Da-Sheng, Chen. 1992. "A Brunei sultan in the early 14th century: Study of an Arabic gravestone". *Journal of South East Asian Studies* 23:1.1–13.

Drewes, G.W.J. 1968. "New light on the coming of Islam to Indonesia?". *Bijdragen tot de Taal-, Land- en Volkenkunde* 124.433–459.

Echols, John M. and Hassan Shadily (eds.). 1992. *Kamus Indonesia-Inggris: An Indonesian-English dictionary*. Jakarta: Gramedia.

Fatimi, S. Qudratullah. 1963. *Islam comes to Malaysia*. Singapore: Malaysian Sociological Research Institute.

Gallop, Annabel Teh. 1991. *Golden letters; Writing traditions of Indonesia: Surat emas; Budaya tulis di Indonesia*. London and Jakarta: British Library and Yayasan Lontar.

General guidelines for the formation of terms in Malay. 1992. Kuala Lumpur: Dewan Bahasa dan Pustaka.

Jones, Russell. 1978. *Arabic loan-words in Indonesian*. London and Paris: Indonesian Etymological Project and Archipel.

——. 1984. "Loan-words in contemporary Indonesia?". *Towards a description of contemporary Indonesian: Preliminary studies,* II, ed. John W.M. Verhaar, 1–38. Jakarta: Universitas Atma Jaya.

——. 2003. *Chinese loan-words in Malay and Indonesian: A background study*. Kuala Lumpur: University of Malaysia Press.

Kasimin, Amran. 1987. *Perbendaharan kata Arab dalam Bahasa Melayu* [Arabic vocabulary in Malay]. Bangi: Universiti Kebangsaan Malaysia.

Matheson, Virginia and M.B. Hooker. 1988. "Jawi literature in Patani: The maintenance of an Islamic tradition". *Journal of the Malaysian Branch of the Royal Asiatic Society* 61:1.1–86.

Meulen, Daniël van der and Herman von Wissman. 1932. *Hadramaut: Some of its mysteries unveiled*. Leiden: E.J. Brill.

Ngajenan, Mohamad. 1987. *Kamus etimologi Bahasa Indonesia* [Etymological dictionary of Bahasa Indonesia]. Semarang: Dahara Prize.

Pedoman transliterasi huruf Arab ke huruf Rumi [Arabic-Roman transliteration manual]. 1992. Kuala Lumpur: Dewan Bahasa dan Pustaka.

Perry, John R. 1991. *Form and meaning in Persian vocabulary: The Arabic feminine ending*. Costa Mesa: Mazda.

Shellabear, W.G. 1901. "The evolution of Malay spelling". *Journal of the Royal Asiatic Society, Straits Branch* 36 (July), 75–135.

Teeuw, A. 1959. "The history of the Malay language: A preliminary survey". *Bijdragen tot de Taal-, Land- en Volkenkunde* 115.138–161.

Tham Seong Chee. 1990. *A study of the evolution of the Malay language: Social change and cognitive development*. Singapore: Singapore University Press.

Verhaar, John W.M. 1984. "Affixation in contemporary Indonesian". *Towards a description of contemporary Indonesian: Preliminary studies*, I, ed. Bambang Kaswanti Purwo, 1–26. Jakarta: Universitas Atma Jaya.

Versteegh, Kees. 2003. "The Arabic component of the Indonesian lexicon". *Rintisan kajian leksikologi dan leksikografi*, ed. Lilie Suratminto and Munawar Holil, 216–229. Jakarta: Fakultas Ilmu Pengetahuan Budaya Universitas Indonesia.

Stuart Campbell
(University of Western Sydney)

Inflection

INTRODUCTION

Inflection is one of the two major processes of word formation, the other being → derivation. It is generally assumed that inflection serves to signal grammatical functions (such as case marking, plural formation, and verb conjugation), while derivation is used to generate new word classes (verbs from nouns, adverbs from adjectives, etc.). This is not the case in Arabic, where inflection plays a considerable part in the formation of the lexicon.

Languages vary widely as to the importance and complexity of their systems of morphological marking. Typologists have therefore proposed to classify them along a continuum ranging from languages in which words undergo almost no morphological change to those in which variations are numerous and complex. The first case, typically exemplified by Chinese, corresponds to uninflected languages. The second, in which Classical Arabic is ranked with Latin, Classical Greek, and modern languages such as German or Russian, corresponds to inflected languages. Languages like French or English stand somewhere between these two extreme positions.

Morphological changes undergone by words in inflected languages are not discrete: one cannot establish a one-to-one relationship between a morphological variant and a precise functional value. Consequently, although an inflected word necessarily contains more than one morpheme (since it is composed of a basic lexical unit plus various grammatical markers), one cannot, as a rule, establish a precise correspondence between the morphemes involved and the linear sequence of segments which make up the word. This trait typically differentiates inflected languages from 'agglutinative' ones. It results both from coalescence, which tends to merge into a single unanalyzable sequence morphemes that were distinct at an earlier stage of the history of the language, and from the fact that some inflected languages make use of morphophonological processes affecting the word globally, for instance vowel alternation (ablaut or → apophony), which modifies the overall vocalic melody of a word to express some of its functional variations. At all documented stages of its long history, Arabic has made massive use of this kind of process, both to develop its lexicon (which is mainly the domain of 'internal' inflection) and to mark several grammatical functions (the preferential domain of 'external' inflection).

A fundamental feature of Arabic is that no word belonging to an open class is formed by a single morpheme. Every such word is minimally made up of a → root morpheme and a pattern morpheme intertwined with it to form the base of the word. This base, nominal or verbal, constitutes the domain of internal inflection. On the other hand, affixes are normally necessary to actualize nouns and verbs in discourse. At this level, external inflectional phenomena take place.

1. INTERNAL INFLECTION

1.1 *General processes of internal inflection*

In Arabic, internal inflection essentially involves three formal processes, used either separately or in combination:

i. Vowel insertion in different positions of the 'consonantal skeleton' formed by the consonants of the root and those of the affixal augments. When it involves several vowels, this insertion obeys constraints imposed both by principles governing vowel alternation (ablaut) and by certain rules of phonological alternation. Vocalic alternation may be purely qualitative (only vowel quality is involved) or quantitative (vowel length is also used). The two parameters may be combined.
ii. Consonant → gemination, essentially concerning the second or third consonant in triconsonantal roots.
iii. Addition of prefixal, infixal, or suffixal augments to give substance to the consonantal skeleton of a base. Such augments may themselves bear a vowel which becomes an integral part of the overall vocalic melody of the base involved.

1.2 *Internal vs. external inflection*

In his study of the development of the Arabic lexicon, Fleisch (1968:49ff.) defines 'internal' inflection as including only those processes which involve vocalic alternations within root consonants; he calls 'external' all processes involving augments (which he simply calls affixes).

It seems, however, that limiting 'internal' inflection in the way proposed by Fleisch constitutes a rather artificial distinction and does not shed any light on the way the system actually works. Many facts suggest that augments are an integral part of the base of the word, even though the root consonants may still be recognized as such, at least as long as the link is transparent between the augmented form and the members of the same morphological family. In fact, as soon as this link is obscured, some augments tend to be reanalyzed as root consonants, leading to the formation of a new root. For example, at first a word like *miskīn* 'poor', but originally 'immobilized', was a qualifying noun formed with the intensive pattern miR$_1$R$_2$īR$_3$ and associated with the triconsonantal root *s-k-n* (general meaning 'to be motionless'). In the subsequent history of the language, it tended to get free from its original root and started its own family, based on the (reinterpreted) quadriconsonantal root *m-s-k-n*, thus generating forms like *maskana* 'poverty' and *tamaskana* 'to play the poor'.

In the same way, most processes affecting the domain of the root can also be shown to play a role outside this domain, suggesting that the speaker does not care about such a distinction. For example, the tendency, observed in many dialects, to change R$_1$iR$_2$āR$_3$ into R$_1$uR$_2$āR$_3$, e.g. *ḥimār*/*ḥumār* 'donkey', which is 'internal' to the root, is quite analogous to the change of *miftāḥ* to *muftāḥ* 'key', which is 'external'.

Likewise, in most cases there seems to be no point in differentiating phonological processes that take place inside the word base (nominal or verbal) from those which take place outside it. For example, a sequence like /awu/ is normally transformed into [aː], whether it occurs within a base, e.g. /ṭawula/ ‘to be long’ which changes into [ŧaːla], or at a ‘borderline’, e.g. /yudˁaw+u/ ‘it is called’, which becomes [judʕaː]. It is true that a few morphophonological processes only apply within the boundaries of the root, such as the change of /buyḍ/ ‘white [fem. pl.]’ to [biːđ], contrary to the general tendency to change /uy/ sequences to [uː]. But precisely such processes are irregular and have to be listed as exceptional and probably belonging to an older layer of rules that are no longer productive and whose output is learned as such by the users of the language.

It has even been suggested (Saguer 2002) that most roots beginning with /m/ or /n/ and analyzed today as triconsonantal are very likely to derive from biconsonantal bases with an old nasal augment, which was integrated to generate a new triconsonantal root.

Last but not least, from a methodological point of view, separating augments from bases, whether nominal or verbal, requires a segmentation procedure producing two strings with a definite linguistic status (albeit an abstract one). This is the case when one analyzes a nominal base, say *kitāb* ‘book’, into a root morpheme *k-t-b* and a pattern morpheme $R_1iR_2āR_3$, the first one conveying the idea of ‘writing’ and the second one being identified as a nominalizing pattern found in other nouns like *ḥiṣān* ‘horse’ or *silāḥ* ‘weapon’. But nothing similar is possible with the *mi++* augment of a word like *miftāḥ* ‘key’, for although one does recognize in this word the root *f-t-ḥ* (general idea: ‘opening’), it would be completely misleading to link the idea of ‘instrument’ to the sole augment *mi++*, for it is the complete pattern $miR_1R_2āR_3$ that conveys this notion.

Consequently, it seems more appropriate to limit the notion of ‘external inflection’ to those inflectional phenomena which involve clearly identified affixes segmentable into discrete morphemes to which a definite form and function may be attributed. In Arabic verbs, this is the case for the affixes of conjugation, and in nouns for the morphemes of case and other affixal morphemes, such as the /+iyy/ marking origin (*nisba*), the /+at/ marking individuation or gender, and the suffixes of ‘external’ number (dual and masc. or fem. external plurals).

2. Inflection in verbs

In the verbal domain, due to the joint influence of many historical and structural factors, such as the deep renewal of paradigms, the pressure of → analogy, and the tendency to paradigmatic uniformity, inflection has undergone an evolution tending to limit its proliferation and enhance its coherence. As a result the verbal system of Arabic appears less diversified and more systematic than the nominal one.

2.1 *Internal inflection in verbs*

The general processes of internal inflection can be seen at two levels in the verbal system: the organization of ‘primary’ (i.e. unaugmented) verb bases in several classes of vocalic alternations on the one hand, and that of ‘secondary’ or augmented verbs on the other hand.

In the primary form, the verbal base of the perfect presents a $R_1vR_1vR_3$ structure, while that of the imperfect has a $R_1R_1vR_3$ structure. Between these two kinds of bases there exists a system of vowel alternations involving both the two vowels of the perfect and the only vowel of the imperfect. Table 1 lists the four main classes of this alternation and, when relevant, the proportion of verbs involved, as computed from Wehr’s *Dictionary of Modern Written Arabic* (note that there is a small class with *a/a* alternation, for phonological reasons, representing 12% of the verbs, and a ‘rest’ category of irregular patterns, representing 8%):

Table 1. Main verb classes

$R_1aR_2aR_3$	→	$Ca+R_1R_2i/uR_3$	(59%, of which 27% intransitive)
$R_1aR_2iR_3$	→	$Ca+R_1R_2aR_3$	(16%, of which 68% intransitive)
$R_1aR_2uR_3$	→	$Ca+R_1R_2uR_3$	(5%, of which 97% intransitive)
$R_1uR_2iR_3$	→	$Cu+R_1R_2aR_3$	(passive verbs)

There is more or less agreement about the nature of the semantic categorization underlying the last two classes: the last one is recognized as a ‘passive’ conjugation, or, to stick to the intuitions of the Arab grammarians, an ‘objective’ one promoting the semantic ‘patient’ to the status of grammatical subject, and the last but one is recognized as a stative conjugation whose members are normally intransitive. On the other hand, the meaning of the semantic categorization of the first two classes, and especially the second one, has been the object of many speculations. Fleisch (1968:115f.) proposes to analyze the first class as that of 'purely and simply agentive processes', while the second would subdivide into a class of 'processes with concerned agents', e.g. *rabiḥa* 'to gain' or *sakira* 'to get drunk', and a class of 'qualitative processes', practically melting into the third class, e.g. *fariḥa* 'to be glad' or *kabira* 'to get old'. This classification is rather unsatisfactory. In the first place, it splits the second class into two, thereby making it impossible to determine what the common factor was between all R$_1$aR$_2$iR$_3$ verbs that could have led the speakers of the language to group them together. Moreover, the proposed analysis of R$_1$aR$_2$aR$_3$ verbs is far from convincing, for even disregarding the classic case of *māta* 'to die', it remains unclear how such unintentional verbs as *waqaʿa* 'to fall' or *jaḥaḍa* 'to bulge' could be regarded as 'purely and simply agentive'. Although the point cannot be argued here in detail, it seems more likely that this classification is based essentially on the orientation of the process: R$_1$aR$_2$aR$_3$ verbs, whether intentional or not, may be regarded as having their 'starting point' in the grammatical subject (agent or not). In contrast, the subject of R$_1$aR$_2$iR$_3$ verbs, whether agentive or not and whether transitive or not, may be analyzed as being the 'arrival point' of the process, as can be seen in verbs like *fahima* 'to understand' or *mariḍa* 'to fall ill'. Finally, R$_1$aR$_2$uR$_3$ verbs are characterized by the fact that their subject is both the starting and the ending point of the process.

There are 14 types of augmented verb bases, but only about 10 are really productive. Their traditional classification rests on a simple enumeration which sheds no light on their linguistic organization. Consequently, Fleisch (1968: 121–122) is quite right to reorganize augmented verbs 'according to their morphological connections', which gives approximately the listing in Table 2.

Table 2. Augmented verb classes

Unaugmented or simply augmented forms	Doubly augmented forms
R$_1$vR$_2$vR$_3$ → Cv+R$_1$R$_2$vR$_3$ (Form I)	R$_1$++ta++R$_2$aR$_3$ → Cv+R$_1$++ta++R$_2$iR$_3$ (Form VIII)
R$_1$aR$_2$R$_2$aR$_3$ → Cu+R$_1$R$_2$R$_2$iR$_3$ (Form II)	ta++R$_1$aR$_2$R$_2$aR$_3$ → Ca+ta++R$_1$aR$_2$R$_2$aR$_3$ (Form V)
R$_1$āR$_2$aR$_3$ → Cu+R$_1$āR$_2$iR$_3$ (Form III)	Ta++R$_1$ā R$_2$aR$_3$ → Ca+ta+R$_1$ā R$_2$aR$_3$ (Form VI)
ʾa++R$_1$R$_2$aR$_3$ → Cu+(ʾa++)R$_1$R$_2$iR$_3$ (Form IV)	s++ta++R$_1$R$_2$aR$_3$ → Ca++s++ta++R$_1$R$_2$iR$_3$ (Form X)
n++R$_1$R$_2$aR$_3$ → Ca+n++R$_1$a R$_2$iR$_3$ (Form VII)	

As a rule, simple augmentations add to the basic verbal notion a special modality: conative/factitive (Form II), participative/afficient (Form III), causative (Form IV), and passive (Form VII). Double augmentation, always marked with the *ta*++ augment, adds to the first modality a reflexive one. Form VII, being in itself a reflexive-passive transform of Form I, has no correspondent in the doubly augmented forms.

Some formal characteristics of this table call for additional comment. In the first place, doubly augmented forms associated with Form I and Form IV exhibit an infixal ++*ta*++ augment, in contrast to all other cases, where it is prefixal. The historical evolution of the language transformed former *ta++R$_1$vR$_2$vR$_3$ and *ta++sa++R$_1$R$_2$aR$_3$ into R$_1$++ta++R$_2$aR$_3$ and s++ta++R$_1$R$_2$aR$_3$. In this last case, it must be added that the correspondence between ʾa++R$_1$R$_2$aR$_3$ and s++ta++R$_1$R$_2$aR$_3$ also rests on diachronic considerations: the prefixal causative augment *ʾa*++ which shows up in Form IV was historically in competition with other augments with the same meaning, *ha*++ and *sa*++. It is the latter which finally won in doubly augmented forms. Such transformations of prefixal augments into infixal ones stress the strong tendency to incorporate into the base components that were originally external to it. This constitutes one of the deep tendencies of Arabic inflectional morphology.

Table 2 displays a vocalic alternation within the prefix of conjugation of the imperfect: it is everywhere C*a*+ except in Forms II, III, and IV, where it shows up as C*u*+ (as in the objective conjugation). A close scrutiny of the bases involved shows that they are the only ones exhibiting a long syllable (Cvv or CvC) after the prefix in question. This correlation is no coincidence and shows that a phonological process involving the vowels of the prefix of conjugation ('external inflection') crucially depends on the structure of the verb base ('internal inflection'), again revealing the narrow relationship between the two domains and stressing still more the strong morphophonological unity of the word.

The study of the so-called rare augmented forms reveals other applications of the mechanisms of internal inflection: gemination of R_3 (as in Form XI *iḥmarr* 'to become red'), diphthongization (as in Form XII *iḥdawdaba* 'to become hunchbacked'), and insertion of an /++n++/ infixal augment and duplication of R3 (as in Form XIV *iḥlankak* 'to be deep black'). These forms are no longer productive in Modern Arabic and are mere survivors in the lexicon.

Such is not the case for verbs formed on quadriconsonantal roots, which represent more than 500 units in Wehr's *Dictionary of Modern Written Arabic*, i.e. about 6 percent of the total verbal lexicon. These roots result from different processes, mainly addition of a consonant (generally a liquid) to a triconsonantal root, or partial or total duplication of a biconsonantal one. In Wehr's dictionary 30 percent of these verbs have an augmented form with a prefixal *ta*++ and a reflexive value.

2.2 *External inflection in verbs*

The two basic conjugations of Arabic are realized through a set of suffixes for the perfect and prefixes and suffixes for the imperfect. When the suffixes of the perfect or imperfect are appended to a 'weak' third-root consonant, that is, a /w/ or /y/ → glide, the ensuing sequence is in general phonologically unstable and coalescence phenomena tend to occur. Thus, a theoretical sequence like /daʿaw+a/ 'he called' is realized as [daʕaː], and an underlying sequence /ya+rmiy+uw+na/ 'they throw' is realized [jarmuːna] (→ weak verb). The rules involved are basically the same as those postulated for phonological processes within the verb base, but also for all analogous phenomena in verbs and nouns. This highlights the fact that the same basic phonological processes are involved in internal and external inflectional phenomena, whether in verbs or in nouns.

3. INFLECTION IN NOUNS

Inflection in nouns is much more productive and even more prolific than in verbs and has retained many more archaic traits. As a result, inflection in nouns is much more difficult to present in a structured and coherent form. The following representative samples will serve to clarify the general logic underlying its development.

3.1 *Internal inflection in the noun*

3.1.1 Nominal bases

Although it is impossible to give, in a short article, a complete survey of the nominal bases generated by the mechanisms of internal inflection in Arabic, the following Tables and the accompanying comments give a general idea of the process and its results. Table 3 presents the nouns with a nominal basis containing only one thematic vowel.

Table 3. Nominal basis with one thematic vowel

Vowel insertion	Vowel alternation	Augments
$R_1vR_2R_3$	$R_1aR_2R_3$ $R_1uR_2R_3$ $R_1iR_2R_3$	
$R_1R_2vR_3$		$Cv{+}{+}R_1R_2aR_3$ $Cv{+}{+}R_1R_2uR_3$ $Cv{+}{+}R_1R_2iR_3$

The 'minimal' triconsonantal nominal bases have CVCC patterns. A statistical survey of Wehr's *Dictionary of Modern Written Arabic* shows that 2,172 nouns (including all subclasses) have such patterns. As for their voweling, their distribution is as follows: 1,478 have /a/ (68%), 364 have /i/ (16%), and 330 have /u/ (15%). The three vowel qualities of the language are not distributed randomly, even in this elementary word pattern. The obtaining proportions actually reflect the old 'lightness scale' established by the Arab grammarians, that is a > i > u, which plays such an important part in their explanatory system of the morphophonological properties of the language.

As can be seen in the second column of Table 3, inserting a vowel after the second root consonant is only possible if the base is preceded by a prefixal augment, since initial consonant clusters are precluded by Arabic's general constraints on → syllable structure.

The next step in the logical process of development of nominal bases is the insertion of two vowels: one after R$_1$, the other after R$_2$. From this step on, the process of vowel alternation (or ablaut) comes into play, giving its characteristic aspect to the lexicon of Arabic.

For nominal bases containing two thematic vowels, there exist, in principle, 36 potential patterns resulting from the insertion after R$_1$ and after R$_2$ of the six vowels (three short and three long) of Arabic. Here too, in order to give an idea of the generative capacity of the mechanism in play, close examination is limited to one of the cases, that of R$_1$vR$_2$vR$_3$ patterns. In theory there are nine of these. Out of these nine potential bases, one, namely R$_1$iR$_2$uR$_3$, is precluded because it violates general constraints on ablaut. In the terminology of the Arab grammarians, it is 'too heavy'. A second one, R$_1$uR$_2$iR$_3$, is reserved to verbal bases, with one or two exceptions, e.g. *duʾil* 'a small weasel-like animal', according to Ibn Xālawayhi (d. 370/980; *Laysa* 65). The seven remaining patterns are realized in 1,088 nouns in Wehr's dictionary (all subtypes included), and their distribution is as follows: R$_1$aR$_2$aR$_3$: 711 (65%), R$_1$aR$_2$iR$_3$: 260 (24%), R$_1$uR$_2$uR$_3$: 45 (~4%), R$_1$iR$_2$aR$_3$: 40 (~4%), R$_1$uR$_2$aR$_3$: 20 (~2%), R$_1$aR$_2$uR$_3$: 8 (~0%), R$_1$iR$_2$iR$_3$: 4 (~0%). Here, too, the vowel patterning is subject to severe constraints, and 'light' vowel patterns have the lion's share.

The data so far concern only the insertion of short vowels. But nothing precludes the use of long vowels, either in only one of the two positions, with the second keeping a short vowel, which gives patterns R$_1$vvR$_2$vR$_3$ and R$_1$vR$_2$vvR$_3$, or in both positions, producing patterns like R$_1$vvR$_2$vvR$_3$. All these possibilities are used, again with the prohibition of some sequences and a definite preference for some of the possible patterns over others.

The study of the other 27 possibilities would reveal the same general tendencies: the exclusion of sequences considered 'too heavy', and assigning quite variable productivity to those retained.

Among the possible reasons to explain the highly variable productivity of attested patterns, Fleisch (1968:63) mentions a possible preference of speakers for 'iambic' rhythmical patterns like R$_1$vR$_2$vvR$_3$ as compared to 'trochaic' ones like R$_1$vvR$_2$vR$_3$. As a matter of fact, a statistical comparison of the two categories reveals that, among trochaic patterns, only R$_1$āR$_2$iR$_3$ represents a significant part of the nominal patterns, probably because it spread as the *nomen agentis* of the unaugmented verb.

Another massively underrepresented category is R$_1$vvR$_2$vvR$_3$. The only base in this class to occur with some frequency is that of nominal plurals of pattern R$_1$āR$_2$āR$_3$, and then most of them are in fact ʾaR$_1$R$_2$āR$_3$ forms whose first root consonant is a glottal stop /ʾ/. This forms in a closing-syllable position a long vowel, following a well-documented process. In order to account for the scarcity of nominal bases of this type, Fleisch (1968:67) mentions a supposed 'preference for consonants', which would have led to R$_1$vR$_2$R$_2$vvR$_3$ patterns with gemination of R$_2$ rather than R$_1$vvR$_2$vvR$_3$, with lengthening of the vowel. It seems that one could invoke, with as much likelihood, the stress rules of Arabic (which Fleisch ignores): in a word with a long vowel after R2, stress has no chance to go beyond there, and the preceding long vowel, always unstressed, will most likely tend to shorten. The productivity of such a patterns is, therefore, bound to be very low.

The speakers have tended to diphthongize formerly long vowels, as an expressive variant tending to become independent. This possibility has of course been taken advantage of to further enrich the lexicon. Thus, in addition to a R$_1$uR$_2$āR$_3$ pattern, there developed a R$_1$uR$_2$ayR$_3$ pattern. Nöldeke (1904), who carefully studied this case, showed that R$_1$uR$_2$āR$_3$ must have been the older diminutive pattern of Arabic but that it was supplanted in this function by R$_1$uR$_2$ayR$_3$, which came to be felt as more expressive by the speakers. R$_1$uR$_2$āR$_3$ was consequently confined to limited uses such as the expression of physical ailments (e.g. *suʿāl* 'cough', *zukām* 'influenza'), or verbal nouns expressing physiological sounds (e.g. *ṣurāx* 'shout', *ḍurāṭ* 'fart'). Only rare vestiges of the former value of the diminutive of this pattern are found, e.g. *qurāba* 'small goatskin'. This case sheds some light on one of the linguistic mechanisms underlying the development of the Arabic lexicon, a mechanism which may be called 'expressive renewal'. It consists in adjoining to an existing form, felt to be worn-out and lacking expressive force, a new one which only differs from the first by a minimal addition of phonetic material (vowel lengthening, diphthongization, gemination) and which is consequently felt to be more expressive. With the passing of time, this new form ends up ousting the first one from its previous function and marginalizes it in secondary functions. By then the conquering form, having become dominant, loses in turn its expressive force and ends up being felt itself as unsufficiently expressive. A new cycle of renewing may then start. As a matter of fact, the diminutive pattern $R_1uR_2ayR_3$, having become 'normal', has tended in turn to lose its expressive force and has been replaced, at least in North African Arabic dialects (Maghreb and Egypt), by $R_1uR_2ayyiR_3$, which has more 'body'.

Table 4 gives an idea of the generative power of the mechanism of root consonant gemination (patterns preceded by an asterisk are not attested, and G stands for 'glide').

Table 4. Nominal bases with root consonant gemination

Vowel insertion	Gemination of R_2	Gemination of R_3
$R_1vR_2R_3$		
$R_1R_2vR_3$		$Cv{+}{+}R_1R_2aR_3R_3$
		$*Cv{+}{+}R_1R_2uR_3R_3$
		$Cv{+}{+}R_1R_2iR_3R_3$
$R_1vR_2vR_3$	$R_1aR_2R_2aR_3$	$R_1aR_2aR_3R_3$
	$*R_1aR_2R_2uR_3$	$R_1aR_2uR_3R_3$
	$*R_1aR_2R_2iR_3$	$R_1aR_2iR_3R_3$
	$R_1uR_2R_2aR_3$	$R_1uR_2aR_3R_3$
	$R_1uR_2R_2uR_3$	$R_1uR_2uR_3R_3$
	$R_1uR_2R_2iR_3$	$R_1uR_2iR_3R_3$
	$R_1iR_2R_2aR_3$	$R_1iR_2aR_3R_3$
	$R_1iR_2R_2uR_3$	$*R_1iR_2uR_3R_3$
	$R_1iR_2R_2iR_3$	$R_1iR_2iR_3R_3$
$R_1vR_2vvR_3$	$R_1aR_2\bar{a}R_3$	
	$R_1aR_2R_3$	
	$R_1aR_2\bar{\imath}R_3$	
	$R_1uR_2\bar{a}R_3$	
	$R_1uR_2\bar{u}R_3$	
	$R_1uR_2\bar{\imath}R_3$	
	$R_1iR_2\bar{a}R_3$	
	$R_1iR_2\bar{u}R_3$	
	$R_1iR_2\bar{\imath}R_3$	
$R_1vR_2R_2vGR_3$	$R_1uR_2R_2ayR_3$	

Here, the general process consists in geminating either the second radical to generate the $R_1vR_2R_2v(v)R_3$ class, or the third one to generate the $R_1v(v)R_2vR_3R_3$ class. In practice, one observes a significant gap between the theoretical possibilities of the system and those actually realized. On the one hand, the possible succession of vowels is severely restricted ($R_1iR_2iR_3R_3$ is possible, $R_1iR_2uR_3R_3$ is not); on the other hand, the preference for some rhythmic patterns is here again sharply marked: patterns like $R_1vR_2R_2vR_3$ are rarer than those with rhythm $R_1vR_2R_2vvR_3$ or even $R_1vR_2vR_3R_3$. It should be noted that here too the possibility of replacing a long vowel by a diphthong has been exploited.

The proliferation of patterns and hence of nominal bases did not take place in a random and anarchic way. Well-documented cases suggest that the development of the lexicon took place under the control of two major principles: the principle, already referred to, of expressive amplification, and the principle of analogical leveling, which is both its dialectic negation and its logical continuator. One aspect of the pressure of analogy is the fact that there exist what may be called 'categorial quasi exclusions', which tend to preclude some patterns from applying to given categories. One example is, according to Ibn Yaʿīš (d. 643/1245; *Šarḥ al-Mufaṣṣal* VI, 112), the fact that only one adjective is realized with pattern $R_1iR_2aR_3$, namely *ʿidan* 'hostile' (from the root *ʿ-d-w*). Ibn ʿUṣfūr (d. 669/1270; *Mumtiʿ* I, 62–63) manages to cite only one other adjective, *ziyam*, glossed as '[a dwelling place] whose people are scattered'.

The best example of the leveling effects of analogy remains that of verbal nouns (*maṣdar*s) of unaugmented verbs. In Classical Arabic, and in sharp contrast to the case of augmented verbs (for which there is practically only one pattern of *maṣdar* for each augmented form), the *maṣdar*s of unaugmented verbs used to exhibit a large variety of forms. Ibn Yaʿīš (*Šarḥ al-Mufaṣṣal* VI, 43), following Sībawayhi, cites as many as 32 different patterns. But although this remains the case in Wehr's dictionary of the modern language, most of these patterns have only limited use: a large number of unaugmented verbs select their *maṣdar* from a very small subset of nominal bases. Thus, the pattern $R_1aR_2R_3$, which represents 40 percent of singular nouns in Wehr's dictionary, stands for 62 percent of all *maṣdar*s. If one restricts the inquiry to *maṣdar*s of R$_1$aR$_2$aR$_3$ verbs, the number jumps to 87 percent. Likewise, the R$_1$aR$_2$aR$_3$ pattern represents 'only' 19 percent of triconsonantal singular nouns, but it constitutes 64 percent of the *maṣdar*s of verbs of the R$_1$aR$_2$iR$_3$ pattern.

3.1.2 Nominal augments

The augment *m*v++ has a nominalizing value, which shows up in the role it plays in the formation of participles. Augments *y*v++ and *t*v++, found in such names as *yaġūṯ* '[name of a pagan god]' and *tanḍub* '[name of a thorny tree]', could well be the prefixes of conjugation of the imperfect incorporated to the base following the nominalization of verbal forms. The latter has witnessed an expansion of its use in the formation of *maṣdar*s. The augment ʾv++ is widely used not only in the formation of the pattern of high degree, ʾaR$_1$R$_2$aR$_3$ (as in *ʾakbar* 'greater'), but also in the formation of many internal plurals, particularly those of 'small number': ʾaR$_1$R$_2$āR$_3$, ʾaR$_1$R$_2$uR$_3$, ʾaR$_1$R$_2$iR$_3$at.

Infixed augments are similar to those already mentioned in the verbal lexicon. They seem to result either from the incorporation of former prefixal augments (as is the case for infix ++*t*++), or from diphthongization processes, or else from the dissimilation of long vowels.

Only one suffixed augment seems to play an important role in the lexical development of the language, to wit ++*n*, variably realized as ++*an*, ++*ān*, or even ++*un*, ++*in*, ++*ūn*, or ++*īn*. It is found in adjectives and *maṣdar*s of augmentative formations.

3.1.3 Internal plurals

One of the characterizing features of Arabic nominal morphology consists in associating to a singular noun a plural formed on the same root but on a totally different base. Such plurals are called 'internal plural' to indicate that they do not use any suffixing mechanism, or → 'broken plural', a term taken from the ancient Arab grammarians and indicating that the nominal base of the singular has, so to say, 'exploded' in this process of plural formation. This process constitutes one of the most typical fields of application of internal inflection. Such plurals are traditionally divided into 'plurals of small number', whose main patterns have been presented in relation with the ʾv++ augment, and 'plurals of large number', whose patterns are extremely diverse. Here too, however, the profusion of possibilities is considerably reduced in practice. Thus, in Wehr's dictionary 49 types of pattern realize the plurals of all triconsonantal nouns, and only three of these patterns are really frequent: ʾaR$_1$R$_2$āR$_3$, ʾaR$_1$R$_2$uR$_3$, and R$_1$iR$_2$āR$_3$. Here again, analogy seems to have played a crucial role in the regulation of this lexical profusion: 27 percent of nouns of the type R$_1$aR$_2$aR$_3$ have a plural of type ʾaR$_1$R$_2$āR$_3$ (e.g. *faraḥ/ʾafrāḥ* 'joy'). Plurals of type R$_1$uR$_2$ūR$_3$ have, in 80 percent of cases, a singular R$_1$aR$_2$R$_3$ (e.g. *šaʿb/šuʿūb* 'people'). Moreover, the so-called quadriconsonantal nouns all form their internal plurals on the maR$_1$āR$_2$iR$_3$ pattern (maR$_1$āR$_2$īR$_3$ if their last vowel is long). Significantly, words formed on actual quadriconsonantal roots, e.g. *ʿaqrab/ʿaqārib* 'scorpion', are treated in the same way as words formed on triconsonantal roots with augments, e.g. *maktab/makātib* 'office', which reveals once again the integration of augments in the word base.

A word in passing concerning what is termed 'plural of plural' in traditional Arabic grammar (*jamʿ al-jamʿ*): in some cases, a former plural having ceased to be felt as such, as in the case of *bilād* ('country', originally the plural of *balad* 'country'), the speakers formed on the older plural a newer one felt to be a 'real plural', in our example *buldān*. In other cases, the speakers, probably driven by 'expressive wit' felt the need to add to the existing plural a new one supposed to stress even more the idea of plurality. This was especially frequent when the older plural was a 'plural of small number': thus, singular *kalb* 'dog' has a 'small-number plural' *ʾaklub* (supposedly referring to no more than ten individuals) and a plural-of-plural *ʾakālīb* (supposedly applying to a larger number of animals). The mechanism involved is strongly reminiscent of the processes of expressive renewing already referred to. Grammarians stress that such plurals may not be freely produced and that only the existing ones, inherited from the old language of the Bedouin, are to be used.

3.2 *External inflection in nouns*

The noun base may be modified by elements in prefixal position, but one should rather qualify these elements as → clitics rather than affixes because they are not indispensable to the morphological autonomy of the base. These clitic elements are the definite article /al#/ (whose initial vowel behaves like an epenthetic one, although it is quite likely that it was originally lexical, as the older form of the article could well have been /*hal/) and a small number of functional words: coordinating conjunctions (e.g. /fa#/ and /wa#/), thematic particles (e.g. /la#/), and prepositions (e.g. /bi#/, /li#/, or /ka#/), whose common characteristic is their CV struc-ture, a sequence too small to constitute an Arabic word, hence their cliticization.

External inflection in the noun essentially shows up in suffixal position and may be divided into three main functional slots: the suffix morpheme /+iyy/, marking origin (*nisba*), the suffix morpheme /+at/, marking individuation or gender (which may alternate with the suffix of 'external' fem. pl. /+āt/), and the morphemes of case. In words like /ʿarab+iyy+at+u#n/ 'Arab [fem., sg., nominative, indefinite]' or /ʿarab+iyy+āt+u#n/ 'Arab [fem., pl., nominative, indefinite]', all three slots are filled as well as a fourth one, occupied actually not by a suffix but by a clitic, the famous → *tanwīn*, marking indefiniteness in the morphologically regular common noun. This succession is regular and shows that Arabic presents to some degree morphological properties found in agglutinative languages.

On the other hand, forms like /muslim+ā+ni/ 'Muslim [dual, nominative, unspecified status marker]' or /muslim+ū+na/ 'Muslim [pl., nominative, unspecified status marker]' show that case markers sometimes coalesce with number, a morphological accident which seems to have caused the reinterpretation of a former *tanwīn* as a mere marker of unspecified status (for more details cf. Kouloughli 2001).

The declensional suffixes deserve some special attention. In the general case, that of fully declinable nouns, they simply consist of the three short vowels of Arabic added in final position (pre-final if the *tanwīn* follows), /+u/ marking the nominative, /+a/ the accusative, and /+i/ the genitive. In the so-called diptotic declension (→ diptosis), the /+a/ suffix covers the last two cases, and after the feminine plural suffix /+āt/ it is the /+i/ case suffix which fills this role. In the case of duals and external masculine plurals, case, together with number as we have just seen, is marked by long or diphthongized vowel suffixes, again in a reduced binary formal opposition: /+ā/ in the dual and /ū/ in the plural for the nominative, /+ay/ in the dual and /+ī/ in the plural for the oblique case. In all cases, if the base of the word has a glide as its third root consonant, the contact with vowel suffixes triggers, just as in verbs, phonetic mutations. Here again, the same basic rules apply to generate stable surface representations manifesting the great morphophonological unity of the language and the absence of essential differences between the processes involving the base and those involving the affixes.

It must be stressed, in this connection, that the suffixal nature of the phonetic material involved in marking case inflection in Arabic, together with its 'lightness' (short vowels in most cases), has had considerable consequences for the morphosyntactic evolution of the language. The main reason for this is that the → stress rules of the language, together with its pausal rules (→ pausal forms), implied that the case suffixes, when consisting of short vowels, were doomed to become imperceptible in pausal position. It is generally assumed that children elaborate their grammar on the basis of the pausal forms. It was consequently inevitable that the grammar of the language tended to shift progressively toward a form devoid of case markers. Epigraphic data suggest that this progressive loss of case marking started very early in peripheral Arabic dialects. It is only due to the deeply conservative character of the Bedouin dialects of the central domain (the high plateaus of the Najd) that the → poetic koine, which was the basis of elaboration of the literary language, preserved the use of case endings up to the times when the language was codified by grammarians. It is all the more remarkable that the grammarians turned these inflectional endings, which by then were disappearing from daily use even in Bedouin dialects, into the nodal point of all their theoretical work on the language, so much so that the word → *ʾiʿrāb* came to mean both case marking and syntax in general.

In conclusion, we may note that the spoken varieties of the language, even if they all discarded case marking, have nevertheless continued to make systematic use of the mechanisms of internal inflection: qualitative and quantitative vowel alternation, consonant gemination, and augments. Just as in earlier stages of the language, these mechanisms remain essential for the development of the living lexicon of today's Arabic, even in the borrowing of foreign words. Thus, on a productive pattern of singular/plural alternation like CāC/CīCāC (as in *bāb/bībān* 'door(s)'), Maghreb Arabic has integrated the French word *car* 'coach' as *kār/kīrān*. In the same way, Cairo Arabic forms on the borrowed word *gurnān* 'journal' the quadriconsonantal plural with vowel-lengthening *garānīn*. All this shows that a language may be massively inflectional and yet make little or no use of case endings in marking its grammatical functions.

Bibliographical references

Primary sources

Ibn ʿUṣfūr, *Mumtiʿ* = ʾAbū l-Ḥasan ʿAlī ibn Muʾmin Ibn ʿUṣfūr al-ʾIšbīlī, *al-Mumtiʿ fī t-taṣrīf*. Beirut: Dār al-ʾĀfāq al-Jadīda, 1978.

Ibn Xālawayhi, *Laysa* = ʾAbū ʿAbdallāh al-Ḥusayn ibn ʾAḥmad Ibn Xālawayhi, *Laysa fī kalām al-ʿArab*. Beirut: Dār al-ʿIlm li-l-Malāyīn, 1979.

Ibn Yaʿīš, *Šarḥ* = Muwaffaq ad-Dīn Yaʿīš Ibn Yaʿīš, *Šarḥ al-Mufaṣṣal*. Beirut: Dār al-Kutub, n.d.

Secondary sources

Bohas, Georges. 1979. "Contribution à l'étude de la méthode des grammairiens arabes en morphologie et en phonologie d'après des grammairiens arabes 'tardifs'". Georges Bohas and Jean-Patrick Guillaume, *Etude des théories des grammairiens arabes. I. Morphologie et phonologie*, 1–333. Damascus: Institut Français d'Etudes Arabes, 1984.

Fleisch, Henri. 1968. *L'arabe classique: Esquisse d'une structure linguistique*. Beirut: Dar El-Machreq.

Kouloughli, Djamel Eddine. 2001. "Sur le statut linguistique du *tanwīn*: Contribution à l'étude du sys-tème déterminatif de l'arabe". *Arabica* 48.20–50.

Nöldeke, Theodor. 1904. *Beiträge zur semitischen Sprachwissenschaft*. Strasbourg: Trübner.

Saguer, A. 2002. *Ḏāhirat al-ʾisbāq fī l-judūr al-ʿarabiyya*. Agadir: Université d'Agadir, Faculté des Lettres et Sciences Humaines.

Wehr, Hans. 1960. *A dictionary of Modern Written Arabic*. 3rd ed. Ed. J. Milton Cowan. Wiesbaden: O. Harrassowitz.

Djamel Eddine Kouloughli (CNRS)

ʾInna wa-ʾaxawātuhā

Among the particles in Arabic grammar there is one group that has a special status because of the way it affects syntactic relations in the sentence. This group includes the particles *ʾinna, ka-ʾanna, lākinna, layta*, and *laʿalla*, collectively called *ʾinna wa-axawātuhā* 'ʾ*inna* and its sisters'. Most later grammarians count *ʾanna* and its derivatives, such as *li-ʾanna*, among the *ʾaxawāt ʾinna*, but according to ʾAbū Ḥayyān (*Manhaj* 72.1), these were excluded by Sībawayhi because *ʾanna* is an accidental form (*ʿaraḍ*) of *ʾinna* (for a detailed analysis of the difference between *ʾinna* and *ʾanna*, see ʾAbū Ḥayyān, *Manhaj* 74–78; Bergter 1988:64–85). All particles in this group act on a topic/comment sentence; the former topic (*mubtadaʾ*) is put in the accusative and the former comment (*xabar*) in the nominative, as in (1).

(1) *ʾinna zayd-an munṭaliq-un*
'Indeed, Zayd is leaving'

Traditionally, *ʾinna* is translated in English with asseverative adverbs like 'verily, indeed, surely, certainly', but as Bloch (1986:102–136) explains, "*Inna* was originally a presentative in a primary, nuclear sentence-structure of the type *inna Zaydan*" (1986:136). He compares *ʾinna* with Hebrew *hinnē* 'here is…!', drawing attention to isolated examples in Arabic in which *ʾinna* is followed by only one noun (Bloch 1986:113–115). Sībawayhi (*Kitāb* I, 283.15–16; cf. Zamaxšarī, *Mufaṣṣal* 15.8–12) quotes the following example: *ʾinna mālan wa-ʾinna waladan wa-ʾinna ʿadadan* 'there is money, there is a boy, there is [large] number'. Bloch regards such examples, in which *ʾinna* functions as a → presentative, as traces of the original usage of *ʾinna*. The grammarians interpreted them as cases of deletion of the comment; underlyingly, *ʾinna mālan* means *ʾinna lanā mālan* 'there is money for us, we have money' (Bloch 1986:123).

The presentative function of *ʾinna* was already mentioned by Reckendorf (1921:127); according to him the main function of *ʾinna* is to present a subject about which new information is given in the sentence. Accordingly, the emphasis is in many cases on the predicate (comment):

> Es legt bei weitem nicht mehr immer einen Nachdruck auf den Akk.; im Gegenteil ist der Akk. oft eine aus der Situation bereits wohlbekannte Vorstellung, über die etwas Neues und Wichtiges erst jetzt hinzugefügt wird, so daß der Schwergewicht gerade auf diesem anderen Teile des Satzes liegt.

This means that it is not quite accurate to say that *ʾinna* "introduces independent sentences with emphasis or focus on the subject", as a recent grammar of Modern Standard Arabic maintains (Badawi a.o. 2004:320).

The other particles in this group have various meanings: 'but' (*lākinna*), 'that [introduction of sentential complement with verbs of information and sensory perception]' (*ʾanna*), 'as if' (*ka-ʾanna*), 'because' (*li-ʾanna*), 'perhaps' (*laʿalla*), 'may [wish]' (*layta*).

The grammarians assert that there is a shorter form of *ʾinna*, *ʾin* (*ʾin al-muxaffafa*), which has the same meaning as *ʾinna* but without its syntactic effect (ʾAbū Ḥayyān 82–84; cf. Bergter 1988:106–111; Badawi a.o. 2004:321; Nebes 1982, 1985, 1987). This form is mainly known from its use in the *Qurʾān* (e.g. Q. 86/4 *ʾin kullu nafsin la-mā ʿalayhā ḥāfiḏun* 'verily, every soul has its own guardian'; cf. Q. 36/32, and Q. 7/102, 26/186, where it is followed by a verb; Nebes 1982:11). Sībawayhi (*Kitāb* I, 283.7ff.) mentions expressions of the type *ʾin zaydun la-ḏāhibun* 'indeed, Zayd is leaving', in which *ʾin* has the same function as *ʾinna*; in this form it is followed obligatorily by *la-* (*lām al-fāriqa*). According to some grammarians, even with *ʾin* the topic noun may have the accusative (Rabin 1951:168–170, quoting Sībawayhi, *Kitāb* I, 283, and Zamaxšarī, *Mufaṣṣal* 137). Ibn Hišām (*Muġnī* I, 22) says that Kūfan grammarians did not accept this usage. It should be added that Nebes (1982, 1985, 1987) interprets these instances of *ʾin* as negative *ʾin*. The particle *lākinna* also has a shortened form, *lākin*, which occurs much more frequently than *ʾin*; it is always followed by a verbal sentence (Reckendorf 1921:130–131). Obviously, the short form of the particle *ʾanna*, *ʾan*, is always followed by a verb, or as Ibn Mālik expresses this, *wa-ʾin tuxaffif ʾanna fa-smuhā stakan / wa-l-xabara jʿal jumlatan min baʿdi ʾan* 'if you shorten *ʾanna*, its noun is concealed; turn its comment into a sentence following *ʾan*' (see ʾAbū Ḥayyān's commentary, *Manhaj* 84.1ff.; Bergter 1988:111–112).

Some variation in the syntactic behavior of *ʾinna* is reported from → pre-Islamic Arabic. Ibn Hišām (*Muġnī* I, 35) adduces a *luġa* from the Ḥijāz in which both topic and comment after *ʾinna* have the accusative (Rabin 1951:173; and cf. ʾAbū Ḥayyān, *Manhaj* 72). Conversely, a famous controversy surrounded the interpretation of the Qurʾānic verse *ʾinna hāḏāni la-sāḥirāni* (Q. 20/63 of 66) 'indeed, these two are sorcerers', in which *ʾinna* is followed by two nominatives (Rabin 1951:156). Ibn Hišām (*Muġnī* I, 37) explains this deviation from the normal rules by a morphological peculiarity of the dialect of Mecca, in which, so he claims, the nominative and the oblique case of the dual both end in *-āni*.

The government of *ʾinna* and its sisters is explained by their formal resemblance to a verb: just like a verb, they exclusively govern nouns, and just like a verb in the perfect tense, they end in *-a*. The exact government relations, in particular the government of the predicate, are a point of controversy (cf. Ibn al ʾAnbārī, *ʾInṣāf* 81–84; ʾAbū Ḥayyān, *Manhaj* 72): according to the Baṣran grammarians the comment in the sentence with *ʾinna* is put in the nominative by the particle *ʾinna*, just as the topic is put in the accusative by this particle. According to the Kūfan grammarians, however, the former governor of the comment, i.e. the topic, continues to exert its influence on it, even when it is affected itself by its new governor, *ʾinna* (→ *ʿamal*). Az-Zajjājī (*Majālis* 132–133) attributes the Baṣran and Kūfan points of view to al-Māzinī (d. 248/862) and al-Kisāʾī (d. 183/799), respectively.

In later Arabic grammar (Carter 1981:222–229), *ʾinna wa-ʾaxawātuhā* are dealt with in the chapter about the → *nawāsix* (cf. Bergter 1988:169–171), i.e. those grammatical elements that change the regular relationship between topic and comment, and they are treated together with the class of verbs like *kāna* (→ *kāna wa-ʾaxawātuhā*), and the class of verbs like *ḏanna* 'to think' that are followed by a topic/comment sentence as sentential complement.

There are two ways of looking at the relationship between *ʾinna* and the nouns dependent upon it. The first analysis calls the first noun – the original topic – the *ism ʾinna*, while the second noun – the original comment – is called *xabar ʾinna*. This is the analysis found in Sībawayhi's *Kitāb*. In his view, the second noun in the construction with *ʾinna* is the *xabar* of the governor *ʾinna*, just like *munṭaliqan* is *xabar* in the construction with *kāna* (Sībawayhi, *Kitāb* I, 280.7; Mosel 1975:283). Sībawayhi explains that in the sentence (2a)

(2a) *ʾinna zaydan aḏ-ḏarīfa munṭaliqun*
'Indeed, Zayd, the charming one, is leaving'

when *munṭaliqun* is omitted, *aḏ-ḏarīf* becomes the comment, as in (2b)

(2b) *ʾinna zaydan aḏ-ḏarīfu*
'Indeed, Zayd is the charming one'

This is exactly like the syntactic relationship in sentences (3a) and (3b).

(3a) *kāna zaydun aḏ-ḏarīfu ḏāhiban*
'Zayd, the charming one, was going away'
(3b) *kāna zaydun aḏ-ḏarīfa*
'Zayd was the charming one'

Sībawayhi states that the accusative *aḏ-ḏarīfa* in (3b) is in the same *manzila* 'category' as the nominative with *ʾinna* in (2b). The implication is that *ʾinna* governs both the topic and the comment that fall under its scope. If another constituent intervenes between *ʾinna* and its comment, as in (4)

(4) *ʾinna fīhā zaydan/zaydan qāʾiman*
'Indeed, in it is Zayd standing'

two analyses are possible: *zayd* may be analyzed as being governed by *ʾinna*, in which case it receives the accusative, or the intervening adverbial *fīhā* may annul the government (→ *ʾilġāʾ*), in which case it receives the nominative as the topic of the sentence.

Likewise, when the adverbial intervenes between the topic and the comment, as in (5)

(5) *ʾinna zaydan fīhā qāʾimun/qāʾiman*
'Indeed, Zayd is standing in it'

there is a choice in the case ending of the comment: it may be interpreted as being governed by *ʾinna*, in which case it receives the nominative, or it may be regarded as a *ḥāl*, comparable to its interpretation in a simple topic/comment sentence like (6):

(6) *zaydun fīhā qāʾimun/qāʾiman*
'Zayd is standing in it'

In grammatical treatises in the 4th/10th century, the formal resemblance between *ʾinna wa-ʾaxawātuhā* and a verb is handled differently. Some grammarians draw from the case relations in the sentence with *ʾinna* the conclusion that the noun governed in the accusative must be its object, and the second noun, which is governed in the nominative, its agent. Thus, a linguistic theorist like az-Zajjājī (d. 339/949) simply states: *fa-l-manṣūb bihā mušabbah bi-l-mafʿūl lafḏan wa-l-marfūʿ bihā mušabbah bi-l-fāʿil lafḏan* 'the word in the accusative [i.e. the topic] is formally likened to the object, and the word in the nominative [i.e. the comment] is formally likened to the agent' (*ʾĪḍāḥ* 64.17). His preoccupation here is with the epistemological structure of linguistic argumentation (cf. Versteegh 1995:93); elsewhere, az-Zajjājī (*Jumal* 64.6) takes care to add that this formal resemblance exists in spite of the meaning because obviously the topic governed by *ʾinna* does not have the meaning of an object (cf. Ibn al-Warrāq, *ʿIlal* 333.7–8).

Owens (1988:223, 240–241) analyzes the relationship between the simple sentence *zaydun munṭaliqun* 'Zayd is leaving' and the sentence with *ʾinna* as governor in connection with his comparative treatment of Arabic grammar and modern linguistics. He states that at first sight the construction with *ʾinna* might be seen as a derivation in the sense of transformational grammar: the topic/comment sentence is transformed in a sentential complement of *ʾinna* or *ʾanna*. But as it turns out, the relationship is analyzed by the Arab grammarians in a different way: they say that these governors 'enter' (*daxala*) the simple sentence, just like any other non-basic element may enter a construction. In this sense, the sentences with *ʾinna* etc. are similar to those with *kāna* (→ *kāna wa-ʾaxawātuhā*). In Owens' view a transformational interpretation of this process would be inappropriate: "There is no technical sense, no specific rule by which the Arabic grammarians derive a sentential complement from a basic topic-comment construction" (1988:242).

In connection with the constructions with *ʾinna wa-ʾaxawātuhā*, several aspects were analyzed by the grammarians. The syntactic difference between constructions with and without *ʾinna* are dealt with, for instance, by al-Baṭalyūsī (d. 521/1127) in his commentary on az-Zajjājī's *Jumal* (*Ḥulal* 158–166; cf. Owens 1988:241). The difference had already been noted by al-Mubarrad (quoted in Zajjājī, *ʾĪḍāḥ* 135.8–9), who attributes it to the fact that *ʾinna* and its sisters have less 'freedom of action' (*taṣarruf*) than the verbs which they formally resemble. It is, for instance, impossible for *ʾinna* to follow its *ism*, whereas a verb can follow its noun; therefore, a sentence like **zaydan ʾinna qāʾimun* is unacceptable. Likewise, it is impossible to say **inna qāʾimun zaydan*, fronting the comment before the topic, even though in a verbal sentence the object normally follows the agent.

By contrast, it is allowed to say *ʾinna fīhā zaydan (istaqarra)* 'indeed, Zayd is (residing) in it', where the adverbial *ḏarf* (→ *mafʿūl fīhi*) intervenes between *ʾinna* and its noun (a word order that is even deemed obligatory by ʾAbū Ḥayyān, *Manhaj* 73.2 to avoid → cataphora in cases like *ʾinna fī d-dāri sākinahā* 'indeed, in the house is its inhabitant'). An original explanation of this construction is given by Ibn al-Warrāq (d. 381/991). He states (*ʿIlal* 333–339) that the adverbial is governed by the underlying verb *istaqarra* and may be fronted to it just like any other object may be fronted to its verb. It is not directly governed by *ʾinna*, and consequently, it may be moved. A verb, as in *ʾinna zaydan kataba* 'indeed, Zayd wrote', is not directly governed by *ʾinna*, either, but it is in the same slot (*mawḍiʿ*) as the comment, and therefore, it cannot be moved. Consequently, a sentence like **ʾinna kataba zaydan* is not allowed. The reason why *fīhā* may intervene between *ʾinna* and its noun is that it is connected with the comment (the underlying verb *istaqarra*), so that it is not really a 'foreign' (*ʾajnabī*; → *sabab*) element, and may be allowed to intervene between *ʾinna* and its noun (cf. Versteegh, forthcoming).

The early grammarians did not focus on any semantic differences between the particles in the group of *ʾinna wa-ʾaxawātuhā*, and they did not use these differences to explain any differences in their syntactic behavior. For later grammarians, however, this became a major issue. Ibn al-Warrāq, for instance, explains the difference in syntactic behavior between *lākinna* 'but' and *ʾinna*, on one hand, and the rest of the *ʾaxawāt kāna*, on the other, by referring to the semantic difference between these two groups (*ʿIlal* 339–344). In (7a)

(7a) *lākinna zaydan qāʾimun wa-ʿamrun munṭaliqun*
'but Zayd is standing, and ʿAmr is leaving'

the two sentences can be coordinated because *lākinna* signals 'topicality' (→ *ibtidāʾ*) in the following noun (i.e. it signals that the noun following it may be interpreted as a topic). With *ka-ʾanna* 'as if', such a coordination is impossible, as in (7b)

(7b) **ka-ʾanna zaydan qāʾimun wa-ʿamrun munṭaliqun*
'as if Zayd is standing and ʿAmr is leaving'

In (7b), the topicality of *ʿamrun* clashes with the meaning indicated by *ka-ʾanna*, and, therefore, the coordination fails. Some grammarians even condemned the use of *lākinna* in coordinative structures because they maintained that the topicality of the second part clashed with the meaning expressed by *lākinna*, viz. *istidrāk* 'rectification' (e.g. Ibn ʿUṣfūr, *Šarḥ* I, 451–457; cf. Larcher 1990). Ibn al-Warrāq's explanation is one of the first attempts to define the semantic aspects of the construction with *ʾinna wa-ʾaxawātuhā*. A thorough semantic analysis was developed in later grammar, starting with grammarians like al-Jurjānī (d. 471/1078) in his *Muqtaṣid* (I, 451ff.), and culminating in Ibn Hišām's discussion of the particles in his *Muġnī* (I, 36–38).

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

ʾAbū Ḥayyān, *Manhaj* = ʾAbū Ḥayyān Muḥammad ibn Yūsuf al-Ġarnāṭī an-Naḥwī, *Manhaj as-sālik fī l-kalām ʿalā ʾAlfiyyat Ibn Mālik*. Ed. Sidney Glazer. New Haven, Conn.: American Oriental Society, 1947.

Baṭalyūsī, *Ḥulal* = ʾAbū Muḥammad ʿAbdallāh ibn Muḥammad Ibn as-Sīd al-Baṭalyūsī, *Kitāb al-ḥulal fī ʾiṣlāḥ al-xalal min Kitāb al-jumal*. Ed. Saʿīd ʿAbd al-Karīm Saʿʿūdī. Baghdad: Dār ar-Rašīd li-n-Našr, 1980.

Ibn al-ʾAnbārī, *ʾInṣāf* = ʾAbū l-Barakāt ʿAbd ar-Raḥmān ibn Muḥammad al-ʾAnbārī, *Kitāb al-ʾinṣāf fī masāʾil al-xilāf bayna n-baḥwiyyīna l-baṣriyyīna wa-l-kūfiyyīn*. Ed. Gotthold Weil. Leiden: E.J. Brill, 1913.

Ibn Hišām, *Muġnī* = Jamāl ad-Dīn ʾAbū Muḥammad ʿAbdallāh ibn Yūsuf Ibn Hišām, *Muġnī l-labīb ʿan kutub al-ʾaʿārīb*. Ed. Māzin al-Mubārak and Muḥammad ʿAlī Ḥamdallāh. Damascus: Dār al-Fikr, 1969.

Ibn as-Sarrāj, *ʾUṣūl* = ʾAbū Bakr ibn as-Sarī Ibn as-Sarrāj, *Kitāb al-ʾuṣūl fī n-naḥw*. Ed. ʿAbd al-Ḥusayn al-Fatlī. 3 vols. Beirut: Muʾassasat ar-Risāla, 1985.

Ibn ʿUṣfūr, *Šarḥ* = ʾAbū l-Ḥasan ʿAlī ibn Muʾmin Ibn ʿUṣfūr al-ʾIšbīlī, *Šarḥ Jumal az-Zajjājī*. Ed. Ṣāḥib ʾAbū Janāḥ. 2 vols. Baghdad: Dār al-Kutub li-ṭ-Ṭibāʿa wa-n-Našr, 1980.

Ibn al-Warrāq, *ʿIlal* = ʾAbū l-Ḥasan Muḥammad ibn ʿAbdallāh al-Warrāq, *ʿIlal an-naḥw*. Ed. Maḥmūd Muḥammad Maḥmūd Naṣṣār. Beirut: Dār al-Kutub al-ʿIlmiyya, 2002.

Jurjānī, *Muqtaṣid* = ʾAbū Bakr ʿAbd al-Qāhir ibn ʿAbd ar-Raḥmān al-Jurjānī, *Kitāb al-muqtaṣid fī šarḥ al-ʾĪḍāḥ*. Ed. Kāẓim Baḥr al-Marjān. 2 vols. Baghdad: Wizārat aṯ-Ṯaqāfa wa-l-ʾIʿlām, 1982.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. 2 vols. Bulaq, 1316 A.H. (Repr., Baghdad: Al-Muthanna Library, n.d.)

Zajjājī, *ʾĪḍāḥ* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān Ibn ʾIsḥāq az-Zajjājī, *al-ʾĪḍāḥ fī ʿilal an-naḥw*. Ed. Māzin al-Mubārak. Cairo: Dār al-ʿUrūba, 1959.

——, *Jumal* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān Ibn ʾIsḥāq az-Zajjājī, *Kitāb al-jumal*. Ed. Muhammad Ben Cheneb. Paris: Klincksieck, 1957.

——, *Majālis* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān Ibn ʾIsḥāq az-Zajjājī, *Majālis al-ʿulamāʾ*. Ed. ʿAbd as-Salām Muḥammad Hārūn. Kuwait: Wizārat al-ʾIršād wa-l-ʾAnbāʾ, 1962.

Zamaxšarī, *Mufaṣṣal* = ʾAbū l-Qāsim Maḥmūd ibn ʿUmar az-Zamaxšarī, *Kitāb al-mufaṣṣal fī n-naḥw*. Ed. Jens Peter Broch. Christiania: Libraria P.T. Mallingii, 1870.

SECONDARY SOURCES

Badawi, Elsaid, Michael G. Carter, and Adrian Gully. 2004. *Modern Written Arabic: A comprehensive grammar*. London and New York: Routledge.

Bergter, Annette. 1988. *Das Kapitel* inna wa-aḫawātuhā *aus dem* “Manhağ as-sālik” *des Grammatikers Abū Ḥayyān al-Ġarnāṭī (1256–1344)*. Hildesheim: G. Olms.

Bloch, Ariel A. 1986. *Studies in Arabic syntax and semantics*. Wiesbaden: O. Harrassowitz.

Carter, Michael G. 1981. *Arab linguistics: An introductory classical text with translation and notes*. Amsterdam: J. Benjamins.

Larcher, Pierre. 1990. “Eléments pragmatiques dans la théorie grammaticale arabe post-classique”. *Studies in the history of Arabic grammar*, II, ed. Michael G. Carter and Kees Versteegh, 195–214. Amsterdam and Philadelphia: J. Benjamins.

Mosel, Ulrike. 1975. *Die syntaktische Terminologie bei Sībawaih*. Ph.D. diss., University of Munich.

Nebes, Norbert. 1982, 1985, 1987. “*In al-muxaffafa* und *al-lām al-fāriqa*, I, II, III”. *Zeitschrift für Arabische Linguistik* 7.7–22, 14.7–44, 16.7–30.

Owens, Jonathan. 1988. *The foundations of grammar: An introduction to medieval Arabic grammatical theory*. Amsterdam and Philadelphia: J. Benjamins.

Rabin, Chaim. 1951. *Ancient West-Arabian*. London: Taylor's Foreign Press.

Reckendorf, Hermann. 1921. *Arabische Syntax*. Heidelberg: C. Winter.

Versteegh, Kees. 1995. *The explanation of linguistic causes: Az-Zağğāğī's theory of grammar*. Amsterdam and Philadelphia: J. Benjamins.

——. Forthcoming. “A new treatise about the *ʿilal an-naḥw*: Ibn al-Warrāq on *ʾinna wa-ʾaxawātuhā*”.

KEES VERSTEEGH
(University of Nijmegen)

ʾInšāʾ

Lexically speaking, *ʾinšāʾ* is the *maṣdar* of the verb *ʾanšaʾa* ‘to create’. In Arabic linguistic terminology, it is the opposite of → *xabar*. *Xabar* designates the assertive (*xabariy(ya)*) utterance (→ *kalām*) or sentence (→ *jumla*); *ʾinšāʾ* designates the non-assertive utterance or sentence (*ġayr al-xabariy(ya)*). According to the encyclopedist al-Kafāwī (d. 1094/1683), it consists of two subdivisions, *ṭalabī* ‘jussive utterance’ and *ʾīqāʿī* ‘performative utterance’ (Kafāwī, *Kulliyyāt* I, 332). *ʾInšāʾ* is sometimes referred to as *balāġī* ‘rhetorical’ for two reasons: first, to distinguish it from the *ʾinšāʾ kitābī* ‘composition’, which belongs to the *kātib*'s art; second, because rhetoric (*ʿilm al-balāġa*) is the linguistic discipline in which the opposition *xabar*/*ʾinšāʾ* stands out most clearly. Since al-Qazwīnī's (d. 739/1338) *Talxīṣ*, this opposition has shaped the first part (*ʿilm al-maʿānī*) of the tripartite science of rhetoric.

The late introduction of this category may explain why it is still to some degree unknown to Western scholars of Arabic. The entry “*Inshāʾ*” in the 2nd edition of the *Encyclopaedia of Islam* does not mention it. The few scholars of Arabic who are aware of it provide a totally or at least partially inadequate description of it. Antoine-Isaac Silvestre de Sacy (1758–1838) refers to it twice in his *Grammaire arabe* (1831:I, 147, II, 513), in the form of *jumla ʾixbāriyya* ‘proposition énonciative’/*jumla ʾinšāʾiyya* ‘proposition tendant à produire une action ou une manière d'être’ (also called ‘productive’ or ‘volitive’). His pupil Heinrich Leberecht Fleischer (1801–1888), who knew the definitions and subdivisions of this category in the Arabic linguistic tradition, was well aware that de Sacy's terminology resulted in a confusion between *ʾinšāʾ* and one of its two sections (*an-nawʿ aṭ-ṭalabī* ‘jussive utterance’), overlooking the other section (*an-nawʿ al-ʾīqāʿī*). Unfortunately, though describing the meaning of the latter, he proposes for *ʾixbār* and *ʾinšāʾ* two terms that are not very helpful, either: *Objectivitätssatz* and *Subjectivitätssatz* (Fleischer 1885–1888:I, 779–780).

Paradoxically, one can find the most faithful description of the *ʾinšāʾ ʾīqāʿī* in the Arabic grammar by Caspari/Wright (1896–1898:II, 1): “The perfect, *al-māḍī*…indicates:…(d) an act which is just completed at the moment, and by the very act of speaking; as *ʾanšadtuka-llāha* I conjure thee by God; *biʿtuka hāḏā* I sell thee this". The paradox is that Caspari/Wright do not refer at this point to the *ʾinšāʾ ʾīqāʿī*, and that the only time they do mention the *ʾinšāʾ* (1896–1898: II, 73), they mean de Sacy's interpretation of *ʾinšāʾ*, i.e. 'a command or wish', as opposed to *ʾixbār* 'statement of fact'. Caspari/Wright's description immediately identifies the *ʾinšāʾ ʾīqāʿī* as what Western linguistics after Austin (1962) has called 'performatives' and, more specifically, 'juridical performatives'. This interpretation has been developed in contemporary Arabic linguistics (Larcher 1980; summed up in a series of articles – see, among others, Larcher 1998), followed by contemporary Arab linguistics (e.g. Mīlād 2001).

The first part of this entry shows how an originally juridical category has become a linguistic one. The second, shorter part suggests that this juridical category represents a strong and original pragmatic link between the different sciences that deal entirely or in part with language in the Islamic world and in later times.

1. From *fiqh* to grammar and rhetoric via the *ʾuṣūl al-fiqh*

The term *ʾīqāʿ* appears in the treatises of *fiqh* 'jurisprudence' as early as aš-Saybānī's (d. 189/805) *al-Jāmiʿ al-kabīr*. It designates what in post-classical times will be defined by az-Zarkašī (d. 794/1392) as "the accomplishment through words of something one aims at, the existence of which is linked to the existence of these words, such as the accomplishment of a sale by *biʿtu* 'I sell', of a marriage by *tazawwajtu* 'I take as a wife', and of a repudiating act by *ṭallaqtu* 'I repudiate'" (*ʾīqāʿ lafḏ̣ li-maʿnan yuqārinuhu fī l-wujūd ka-ʾīqāʿ al-bayʿ bi-biʿtu wa-n-nikāḥ bi-tazawwajtu wa-ṭ-ṭalaq bi-ṭallaqtu; Manṯūr* I, 205). In later treatises of *fiqh, ʾinšāʾ* appears as a synonym of *ʾīqāʿ*, but also in opposition to *ʾixbār*, in order to distinguish between two possible ways of uttering the same sentence, for example *ʾaʿtaqtu-ka*: this can be either a performative of the freeing of a slave with present-time meaning ('I free you'), or a statement with past-time meaning ('I freed you') (Kāsānī [d. 587/1189], *Badāʾiʿ* IV, 46). In the related discipline of the *ʾuṣūl al-fiqh* 'foundations of jurisprudence', al-ʾĀmidī (d. 631/1233) uses *ʾinšāʾ* in this way, but he also distinguishes between imperative and performative of order (*ʾamartu-ka* 'I order you' or *ʾanta maʾmūr* 'you are ordered to…'), qualifying the former as "instituted for the *ʾinšāʾ*" and the latter as "used for the *ʾinšāʾ*" (al-ʾĀmidī, *ʾIḥkām* II, 131–132). This means that he uses *ʾinšāʾ* in a broader sense, since this can be said not only of the *ʾīqāʿ* but also of the *ṭalab* 'request'. The opposition between 'institution' (*waḍʿ*) and 'usage' (*istiʿmāl*) reveals the relationship with rhetoric. Yet, the only division of utterances known in rhetoric at this time is that into *xabar* and *ṭalab* (Sakkākī [d. 626/1229], *Miftāḥ* 71). One can therefore propose the following hypotheses:

i. *ʾInšāʾ* has its roots in *fiqh*. Its juridical origins are corroborated by a short entry in al-Kafāwī's *Kulliyyāt*:

> The act of the tongue is to say and not to do; similarly, the act of all the other organs is to do and not to say. Despite that, the law has made the act of the tongue a legal act, so that it has become similar to the acts of all the other organs (*fiʿl al-lisān huwa li-l-ʾixbār lā li-l-ʾinšāʾ kamā ʾanna fiʿl sāʾir al-jawāriḥ li-l-ʾinšāʾ lā li-l-ʾixbār lākinna š-šarʿ jaʿala fiʿl al-lisān ʾinšāʾ šarʿan fa-ṣāra ka-sāʾir ʾafʿāl al-jawāriḥ*). (V, 314)

ii. *ʾInšāʾ* has experienced a broadening of scope toward the *ṭalab* in the *ʾuṣūl al-fiqh*, for in this discipline, the *ṭalab*, which includes in the first place *ʾamr* 'order' and *nahy* 'prohibition', represents a crossing between language and law.

iii. Finally, it was thanks to such versatile writers as Ibn al-Ḥājib (m. 646/1249), *ʾuṣūlī* and grammarian, that the category of *ʾinšāʾ* has been expanded from the juridical to the lin-guistic sciences. In the latter, *ʾinšāʾ*, though still used to indicate a performative act of uttering, both in a strict sense (*ʾinšāʾ* = *ʾīqāʿ*) and in a broader one (*ʾinšāʾ* vs. *ʾixbār*), is used more frequently by metonymy for the performative utterance (= sentence) in a strict as well as in a broad sense: *ʾinšāʾ* has therefore four extensions. Ibn al-Ḥājib's main commentator in the field of grammar, al-ʾAstarābāḏī (d. 688/1289), uses *ʾinšāʾ* either in opposition to *xabar* and *ṭalab*, or in opposition to *xabar* only, but in this case he divides *ʾinšāʾ* into *ṭalabī* and *ʾīqāʿī* (*Šarḥ al-Kāfiya* I, 8, II, 221). Likewise, Ibn Hišām al-ʾAnṣārī (d. 761/1361) first divides utterances into *xabar*, *ṭalab*, and *ʾinšāʾ*, then he classifies them further as *xabar* or *ʾinšāʾ*, thereby transferring the definition of the first *ʾinšāʾ* to the second: “to bring its expression into existence is to bring what it aims at into existence” (*ʾījād lafḏihi ījād li-maʿnā-hu; Šarḥ* 31–32). Using Austin’s (1962) terminology, this is a typically ‘illocutionary’ definition, whereas the definition of *ṭalab*, in the first classification, was a typically ‘perlocutionary’ one, concerning the realization of the requested thing (*maṭlūb*) and not the act of requesting itself. Although the four extensions are constantly and simultaneously employed, it is nevertheless the fourth and last one that has become the ‘classic’ one: in this sense, *ʾinšāʾ* (vs. *xabar*) is the equivalent, by extension, of Austin’s ‘performative utterance’ (vs. ‘constative utterance’).

Although the interpretation in terms of performativity and illocutionary activity is the only one that allows an adequate comprehension of the entire category of *ʾinšāʾ*, one finds in the primary sources other interpretations of this concept, which cannot be discussed here in detail. It should be signaled, however, that the opposition *ʾixbār*/*ʾinšāʾ* is sometimes reduced to two types of *ʾixbār* ‘assertion’, one concerning the outside world (*ʿāmma fī l-xārij*), the other concerning the speaker’s inner world (*ʿāmma fī l-bāṭin*). This conception may be at the origin of Fleischer’s above-mentioned terminology.

2. Juridical and linguistic sciences in post-classical Islam

In *fiqh*, the category of *ʾinšāʾ* occupies a central place within the section of the *muʿāmalāt* ‘transactions’ and, more specifically, in the debate on the *ṣiyaġ al-ʿuqūd wa-l-fusūx*, i.e. the expressions allowing to bind or unbind juridical ties. In the *ʾuṣūl al-fiqh*, this is the key concept of juridical hermeneutics, i.e. the derivation of norms (*ʾaḥkām šarʿiyya*), in particular the prescriptive ones (*taklīfiyya*), from Qurʾānic or *ḥadīṯ* utterances having a jussive form or meaning. In both disciplines, the interpretive mechanism is actually a ‘rhetorical’ one: the meaning of the utterance is its practical value when it is regarded as an ‘address’ (*xiṭāb*). The *ʿilm al-maʿānī*, in its two sections, *xabar* and *ʾinšāʾ* (the latter subdivided into *ṭalabī* and non-*ṭalabī*, or *ġayr ṭalabī*), is nothing but a pragmatic calculation of the meaning. In the interpretation of Qurʾānic utterances, as well as in the arguments (*qiyās*) provided by the *ʾuṣūliyyūn*, there appear some ‘abstract’ performatives in the form *faʿaltu* (e.g. *ḥarramtu* ‘I forbid’, where *-tu* refers to the ‘Lawgiver’). This form of concrete juridical performatives is the one most frequently used, though not exclusively. The concept of *ʾinšāʾ* and the form *faʿaltu* reveal the parallelism made by Islamic law between contract creation and norm creation. Finally, the same couple appears in grammar, namely in al-ʾAstarābāḏī’s *Šarḥ al-Kāfiya*: *faʿaltu* represents the ‘speaker’s act’ (*fiʿl al-mutakallim*), which is constituted by the uttering of any sentence. Via the exclamative utterances – *ʾinšāʾ juzʾu-hu* (‘partly’) *l-xabar* – al-ʾAstarābāḏī enlarges the category so as to include the most descriptive statement, *zaydun qāʾimun* ‘Zayd is standing’, to which an act of assertion (*ʾixbār*) may correspond. He enlarges it even to the point of including the connectives (e.g. *lākinna* ‘but’/*istadraktu* ‘I rectify beforehand’; *Šarḥ al-Kāfiya* I, 93; II, 311, 346).

3. Conclusion

At first sight, the category of *ʾinšāʾ* (vs. *xabar*) occupies a modest place as a simple classification of utterances that originated at a late date. Actually, it has a long and complex history, and synchronically it symbolizes well the systematization and interaction of the different disciplines, as well as the predominance of the pragmatic dimension. Its importance has recently begun to be recognized by scholars of Arabic, not only by specialists of language sciences, rhetoric (e.g. Simon 1993; Jenssen 1998), or grammar (e.g. Firănescu 2003; Versteegh 2005), but also of Islamic law (e.g. Johansen 1996).

Bibliographical references

Primary sources

ʾĀmidī, *ʾIḥkām* = Sayf ad-Dīn ʾAbū l-Ḥasan ibn ʾAbī ʿAlī ibn Muḥammad al-ʾĀmidī, *al-ʾIḥkām fī ʾuṣūl al-ʾaḥkām*. 4 parts in 2 vols. Cairo: Muʾassasat al-Ḥalabī, 1487/1967.

ʾAstarābāḏī, *Šarḥ al-Kāfiya* = Muḥammad ibn al-Ḥasan Raḍī d-Dīn al-ʾAstarābāḏī, *Šarḥ Kāfiyat Ibn al-Ḥājib*. 2 vols. Istanbul: Maṭbaʿat aš-Šarika aṣ-Ṣiḥāfiyya al-ʿUṯmāniyya, 1275 A.H. (2nd ed., 1310 A.H.; repr. Beirut: Dār al-Kutub al-ʿIlmiyya, n.d.).

Ibn Hišām, *Šarḥ* = Jamāl ad-Dīn ʾAbū Muḥammad ʿAbdallāh ibn Yūsuf ibn ʾAḥmad Ibn Hišām al-ʾAnṣārī, *Šarḥ Šuḏūr aḏ-ḏahab fī maʿrifa kalām al-ʿArab*. Ed. Muḥammad Muḥyī d-Dīn ʿAbd al-Ḥamīd. Cairo, n.d.

Kafāwī, *Kulliyyāt* = ʾAyyūb ibn Mūsā al-Ḥusaynī al-Kafāwī, *al-Kulliyyāt*. Ed. ʿAdnān Darwīš and Muḥammad al-Maṣrī. 5 vols. Damascus: Wizārat aṯ-Ṯaqāfa wa-l-ʾIršād al-Qawmī, 1981.

Kāsānī, *Badāʾiʿ* = ʿAlāʾ ad-Dīn ʾAbū Bakr ibn Masʿūd al-Kāsānī, *Kitāb badāʾiʿ aṣ-ṣanaʾiʿ fī tartīb aš-šarāʾiʿ*. 7 vols. Cairo: Maṭbaʿat al-Jamāliyya, 1327–1328 A.H.

Qazwīnī, *Talxīṣ* = Jalāl ad-Dīn Muḥammad ibn ʿAbd ar-Raḥmān al-Qazwīnī, *Talxīṣ al-Miftāḥ*. Ed. ʿAbd ar-Raḥmān al-Barqūqī. Cairo: al-Maktaba at-Tijāriyya al-Kubrā, n.d.

Sakkākī, *Miftāḥ* = ʾAbū Yaʿqūb Yūsuf ibn ʾAbī Bakr Muḥammad ibn ʿAlī as-Sakkākī, *Miftāḥ al-ʿulūm*. Cairo: Maṭbaʿat at-Taqaddum al-ʿIlmiyya, 1348 A.H. (Repr. Beirut: Dār al-Kutub al-ʿIlmiyya, n.d.)

Šaybānī, *Jāmiʿ* = ʾAbū ʿAbdallāh Muḥammad ibn al-Ḥasan aš-Šaybānī, *al-Jāmiʿ al-kabīr*. Ed. Riḍwān Muḥammad Riḍwān. Cairo: Maktabat al-Istiqāma, 1356 A.H.

Zarkašī, *Manṯūr* = Badr ad-Dīn Muḥammad ibn Bahadūr az-Zarkašī, *al-Manṯūr fī l-qawāʿid*. Ed. Taysīr Fāʾiq ʾAḥmad Maḥmūd. Kuwait: Wizārat al-ʾAwqāf wa-š-Šuʾūn al-ʾIslāmiyya, 1402/1982.

Secondary sources

Austin, John Langshaw. 1962. *How to do things with words*. Oxford: Oxford University Press.

Firănescu, Daniela Rodica. 2003. *Exclamation in Modern Literary Arabic: A pragmatic perspective*. Bucharest: Editura Universităţii din Bucureşti.

Fleischer, Heinrich Leberecht. 1885–1888. *Kleinere Schriften*. 3 vols. Leipzig: S. Hirzel. (Repr., Osnabrück: Biblio Verlag, 1968.)

Jenssen, Herbjøn. 1998. *The subtleties and secrets of the Arabic language: Preliminary investigations into al-Qazwīnī's* Talḫīṣ al-Miftāḥ. Bergen: Center for Middle Eastern and Islamic Studies.

Johansen, Baber. 1996. "Vérité et torture: *Ius commune* et droit musulman entre le X^e^ et le XIII^e^ siècle". *De la violence*, ed. Françoise Héritier, 123–168. Paris: Odile Jacob.

Larcher, Pierre. 1980. *Information et performance en science arabo-islamique du langage*. Ph.D. diss., Université de Paris-3.

——. 1998. "Une pragmatique avant la pragmatique: 'Médiévale', 'arabe' et 'islamique'". *Histoire Epistémologie Langage* 20:1.101–116.

Mīlād, Xālid. 2001. *al-ʾInšāʾ fī l-ʿarabiyya bayna t-tarkīb wa-d-dalāla: Dirāsa naḥwiyya tadāwuliyya*. Tunis: Jāmiʿa Manūba wa-l-Muʾassasa al-ʿArabiyya li-t-Tawzīʿ.

Silvestre de Sacy, Antoine-Isaac. 1831. *Grammaire arabe à l'usage des élèves de l'Ecole spéciale des langues orientales vivantes*. 2nd ed. 2 vols. Paris: Imprimerie Royale. (Repr., with a preface by Gérard Troupeau, Paris: Institut du Monde Arabe, n.d.)

Simon, Udo Gerald. 1993. *Mittelalterliche arabische Sprachbetrachtung zwischen Grammatik und Rhetorik:* ʿIlm al-maʿānī *bei as-Sakkākī*. Heidelberg: Heidelberger Orientverlag.

Versteegh, Kees. 2005. "Meanings of speech: The category of sentential mood in Arabic grammar". *Le voyage et la langue*, ed. Joseph Dichy and Hassan Hamzé, 269–287. Damascus: Institut Français du Proche-Orient.

Wright, William. 1896–1898. *A grammar of the Arabic language*. Transl. from the German of Caspari and ed. with numerous additions and corrections. 3rd ed. 2 vols. Cambridge: Cambridge University Press. (1st ed., 1859–1862.)

Pierre Larcher
(University of the Provence)

Insults

Classical Arabic uses two more or less synonymous words that are equivalent to 'insults': *šatm* and *sabb*. There are only a few definitions of the word *šatm*, and they mention practically only one synonym, that of *sabb*. The term refers to ugliness, or more exactly to the expression of a feeling of repulsion or rejection, whether in the spoken word or physical appearance. *Šatm* indicates bad, ugly language, *qabīḥ al-kalām*. It is sometimes specified (Ibn Manẓūr [d. 705/1311], *Lisān* s.v.) that the reference is to words that are "bad but not accusatory" (*wa-laysa fī-hi qaḏf*). In the case of *sabb*, once the synonymy with *šatm* is made clear, other synonyms seem just as important and their definitions have very diverse meanings. The closest synonym found is *qaṭaʿa-hu* for *sabba-hu*, which carries some idea of a cutting; or else *ṭaʿana-hu bi-* 'to hit with a sword' or 'to strike with a lance', the expression *ṭaʿana fī-hi* (or *ʿalay-hi*) *bi-l-qawl* being translatable as 'to administer tongue lashings'.

Apart from these two words, there are others whose definition shows more or less explicitly a link with *šatm* and *sabb* (Ibn Manẓūr, *Lisān* s.v.; Bustānī, *Muḥīṭ* s.v.). For example, the verb *šāna* emphasizes the idea of making ugly, disfiguring, or spoiling in the physical but also moral sense of debasing; *ʿāba* implies a notion of fault and is used to denigrate, to accuse someone of misbehaving; *qadaḥa* includes the notion of drilling, gnawing, disintegrating by penetrating inside, and, when followed by *fī fulān*, it means 'to denigrate; to destroy'; *ḏamma* is defined by *ḍidd madaḥa-hu* 'the opposite of praising' and refers to the notion of blame and criticism; *qaḏaʿa* is associated with ideas of indecency, coarseness (*faḥš*), or dirt (*qaḏr*), but it also evokes words thrown out one after another. Ibn Sīdah's (d. 458/1066) thematic dictionary *al-Muxaṣṣaṣ* includes in the same entry the words 'insult', 'blame [or: reproach]' and 'prejudice': *aš-šatm wa-l-lawm wa-l-ʾaḏā*ʾ.

Reference to these terms, as well as many others, makes it possible to define the limits of a lexico-semantic field of insults in Arabic. Although this procedure is quite interesting in itself, it is a type of investigation that may lead to circular reasoning, avoiding the insult phenomenon itself, as it is evident in social life.

The laws surrounding insults provide more information. In their analysis, Yahya Ould al-Barra and Abdel Wedoud Ould Cheikh (2004), extensively quoted below, show clearly that insults cannot be easily isolated from a whole set of regulations concerning the use of spoken or written language. Yet, two distinct kinds of offense may be distinguished. Generally speaking, for the *fuqahā*ʾ 'theologian-legal scholars', insults are associated with " 'dangers' (*ʾāfa*) that threaten through words the moral rectitude of someone who would wish to stay or become a good Muslim" (Ould al-Barra and Ould Cheikh 2004:59). "Complete self-control, and even complete silence, are essential in order to avoid the pitfalls and *faux pas* that go hand in hand with the oral expression of language" (2004:60). In the detailed list of the twenty 'dangers' drawn up in particular by al-Ġazālī (d. 505/1111) in the *ʾIḥyāʾ ʿulūm ad-dīn*, insults (*sabb*) rank seventh along with coarseness (*faḥīš*) and verbal aggression (*baḏaʾat al-lisān*), all of which, he believes, "come from the low moral standards and spiritual turpitude of those who are tempted by them" (2004:60).

A quick look at the other 'dangers' (Ould al-Barra and Ould Cheikh 2004:60–64) shows that these words are to be avoided because they are of no concern to anyone, are futile, and deal with dubious subjects. They also have a kind of affectation, or else they are used to tease, argue, joke, and banter aimlessly. Singing and poetry, without being completely prohibited, belong to these dangers. Cursing (*laʿn*) has no effect against the 'infidel', whereas cursing a believer is equivalent to killing him. To make fun of someone, to betray him (*namīma*), to give away a secret, to break a promise, to lie and give false testimony, or to use deliberately ambiguous words and flattery (*madhī*) are clearly to be regarded as 'dangers'. To speak ill of someone and make unkind comments about someone who is absent (*ġība*) is compared to 'eating the flesh of a dead person'. On the other hand, to talk about God's qualities in a casual way or to ask questions about them relates more particularly to the notion of *ridda* 'apostasy'.

When studying Ġazālī's treatise and others like it (cited by Ould al-Barra and Ould Cheikh 2004), one cannot help thinking about the 'sins of language' analyzed and listed by Christian theologians in medieval culture (Casagrande and Vecchio 1991), where insults also figured prominently. On the other hand, the *Kitāb al-kabāʾir* (Book of sins) by aḏ-Ḏahabī (d. 748/1348) shows an interest in language only as far as ten sins are concerned, out of the seventy listed. If it deals with lying (*kiḏb*), abusive invective (*ʾaḏā*), and insults (*šatm, sabb*), it is mainly insofar as they are perpetrated against God or His Envoy, His saints and His worshippers, or the Prophet's companions. Denunciation and malediction are also mentioned, but those practicing them are referred to as *an-nummām* 'those who betray' and *al-liʿān* 'those who curse'.

Apart from its general criminal character, the insult represents for legal scholars a quite specific offense, also linked to the use of words, but clearly of a different type: it does not represent a possible, misguided use of language, implying a lack of self-control, but rather a voluntary act with a precise intention. This offense is the *qaḏf* explained by Ould al-Barra and Ould Cheikh (2004:66–79) as follows: "The verb *qaḏafa* is given as a synonym of *ramā*, which means 'to shoot' (a target with a bow, etc.), 'to aim', 'to throw towards', 'to hit with a projectile', but also throwing hurtful or insulting words at someone, for *qaḏf* is a synonym of *as-sabb* (insult). In the vocabulary of the *fuqahā*ʾ, *qaḏf* defines mainly the accusation of fornication (*zinā*)". The condemnation that follows is not just a moral one; it is a criminal sanction that, the authors claim, originates in the Qurʾānic formulation "defining the sentence [eighty lashes] risked by the accuser of fornication toward 'protected women' (*muḥaṣṣanāt*) [Muslim women of good repute]". In fact, the *qaḏf* leads to a whole system in which specialists in theological law study in detail conditions, cases, degrees of sanction, etc., and strive to distinguish whether an accusation is justified, whether it is merely calumny, whether it is an insult, and whether it is necessary, lawful, or forbidden. "What is above all called into question, what is used as the focal point of verbal aggression is women's 'virtue', and the prejudice incurred by their male relatives because of the insulting suspicions they might attract". This is why *qaḏf* is particularly representative of what is at stake in society, because "insults, taken in the context of the *qaḏf*, turn out to be an expression of the threat that weighs on the social order, to the precise extent in which they affect, through the questioning of female sexuality, the 'honor' of individuals and groups or even more that of 'corporate groups organized along patrilineal lines' (*ʿaṣabāt*) which forms the initial base of social order in the Arabo-Muslim world" (2004:79). From this perspective it is possible, for example, to interpret the acts of violence against women in Algeria in the 1990s as insults against the group, encouraging similar reprisals against the opposing group, and, from a more individual, less collective point of view, as an affirmation (or a defense) of manhood degraded precisely by these offenses against honor (Moussaoui 2004).

The difference in character between the offense of *šatm* and that of *sabb*, made obvious by the legislation, can be found in society itself. In a society like the Lebanese, which is not exclusively Muslim, these same basic ideas are deeply rooted in the culture. On the one hand, there is a kind of repulsion and fear of words and expressions with a sexual content, mostly considered to be vulgar and coarse, and in some ways unpronounceable without risking contamination. On the other hand, there is an awareness of the seriousness of any violation of *ʿirḍ* 'women's sexual honor' (Kanafani-Zahar 2004), which, in a way, represents the group's honor as a whole, whether it be its lineage, the space it occupies, the members who make it up, or the God and prophets it worships.

Studies concerned with the concept of insults in Arabic culture are mostly scattered in general anthropological works, sometimes in ethnographical documents or even in research in semantics. The collective work *L'injure, la société, l'islam* (Larguèche 2004a) focuses on this theme while tackling the subject from a pragmatic point of view through an interest in the effect of insults, rather than confining itself to a collection of terms and expressions that are regarded as insults, independent of any social context or specific situations in which they are used. A number of generalizations emerge from the studies cited in this article that raise further questions and allow for the formulation of specific hypotheses regarding insults.

The moral taboo of insults is quite often coupled with a social taboo. In highly hierarchical societies, members who are at the top are precisely those who know how to use a word properly and are, more fundamentally, the representatives of moral standards. In Tunisia, for example, a kind of social control of insults can be implicitly determined on the basis of "a system of appropriateness allowing people to distinguish between 'those who can be insulted' (and insult each other) and those who cannot", and consequently, a space begins to emerge where it is possible to talk about 'tolerated' insults under specific conditions (Roth 2004). This hierarchical dimension is expressed through social classes; it also structures the relationship between generations and the sexes, thereby establishing a certain systematic usage of insults, with its own conditions, its obvious evasions, its allusive processes, etc. As a result, prohibition depends not only on the type of words or expressions used but also on the relative hierarchical position of the persons involved.

When an insult is examined within its situational context, the moral taboo itself appears to be part of an honor system that is not identical to that of the group but rather is one in which a person's own self-image plays a particularly important role. In Mauritania, for example, insulting someone not only brings discredit to the one doing the insulting (since that person does not respect propriety) but also causes loss of respectability to their own image, their body itself (Taine-Cheikh 2004). An insulting person is said to have a 'foul-smelling' mouth, or one that is 'eaten by termites'. Insulted persons certainly perceive themselves as debased, or at least smeared; but they are also torn between the necessity of retorting to defend themselves and the impossibility of replying in a way that will not soil their own mouth. Thus, the use of insults is not at all recommended, as is borne out by numerous Moorish proverbs and popular sayings. Strategies are therefore developed in order to manage this phenomenon between equals and between persons from different social backgrounds. This is often done through literary forms of expression, poems, and songs, the main consideration being to preserve a good image of oneself. Praise itself, in a particular nuance of this concept of honor, can be interpreted as an insult, precisely for the one who does not receive it.

From the point of view of the modalities of insults, cursing (*laʿn* or, often in dialects, *naʿal*) is a very common, if not the main form for expressing insult. Accordingly, it remains an issue with respect to the taboo formulated by legislation. Divine invocation, if not explicitly mentioned, is always implied. Wishing someone ill can take many different forms; its target may be a particular person, his or her relatives, or anything closely or remotely linked to that person. Cursing compels its recipient to retort in similar or even stronger terms than the original insult (usually just common, unpleasant words) would require. It would appear that in the act of cursing, divine invocation – be it explicitly or implicitly evoked – makes the mouth powerful and malevolent, so that an immediate reply is necessary, not in order to protect one's image but to exorcize evil by sending it back to the other.

The points discussed in this article and the examples provided are only some among many others still too scattered to summarize. They show how much light the phenomenon of insults can shed on a set of problems far more complex than the sole notion of moral and social taboo usually evokes. The most obvious is the double-based system of honor. On the one hand, it considers honor in the context of a person who is an inseparable part of the group, as exemplified by the 'double' insult (Kanafani-Zahar 2004), in which the named target is part of a group (mainly based on lineage and religion). On the other hand, insults refer to the honor of a person (perhaps independent of the group), whose image in society is determined by the person's own language and actions. This image is what is at stake in the competitive behavior among persons with an equal status with respect to the honor of the group, sharing rights and duties toward women, revenge, etc. Through this image, anyone may temporarily acquire honor and prestige, thereby distinguishing themselves from the others (Bonte 1998, 2005).

The effect of an insult (injury) is above all social, in the collective and individual sense of comparing one's own image to that of others and, through a kind of internalization of the social dimension, of comparing one's image to oneself. This may recall other references to honor, represented, for example, by respect for a pledge that has been made.

Bibliographical references

Primary sources

Bustānī, *Muḥīṭ* = Buṭrus al-Bustānī, *Muḥīṭ al-muḥīṭ*. Beirut: Maktaba Luḇnān, 1977.

Ḏahabī, *Kabāʾir* = Šams ad-Dīn ʾAbū ʿAbdallāh Muḥammad ibn ʾAḥmad aḏ-Ḏahabī, *Kitāb al-kabāʾir*. Beirut: Dār al-Kitāb al-ʿArabī, 1980.

Ġazālī, *ʾIḥyāʾ* = ʾAbū Ḥāmid Muḥammad ibn Muḥammad al-Ġazālī, *ʾIḥyāʾ ʿulūm ad-dīn*. Beirut: Dār al-Fikr, 1997.

Ibn Manẓūr, *Lisān* = ʾAbū l-Faḍl Jamāl ad-Dīn Mukarram ibn Mukarram Ibn Manẓūr, *Lisān al-ʿArab*. Beirut: Dār Sādek et Dār Bayrūt, 1955.

Ibn Sīdah, *Muxaṣṣaṣ* = ʾAbū l-Ḥasan ʿAlī ibn ʾAḥmad al-Mursī Ibn Sīdah, *Kitāb al-muxaṣṣaṣ*. Beirut: The Trading Office, n.d.

Secondary sources

Bonte, Pierre. 1998. *L'émirat de l'Adrar: Histoire et anthropologie d'une société tribale saharienne*. Thèse de Doctorat d'État en ethnologie et en anthropologie sociale, Ecole des Hautes Etudes en Sciences Sociales, Paris.

——. 2005. "La notion de 'biens de prestige' au Sahara occidental". *Journal des Africanistes*, 75:2. To appear, 2006.

Casagrande, Carla and Silvana Vecchio. 1991. *Les péchés de la langue: Discipline et éthique de la parole dans la culture médiévale*. Paris: Editions du Cerf. (Transl. from Italian, *I peccati della lingua: Disciplina ed etica della parola nella culturamedievale*, Rome: Istituto dell'Enciclopedia Italiana, 1987.)

Kanafani-Zahar, Aïda. 2004. "L'injure dans la société libanaise: Les mots, le sens". Larguèche (2004a:153–164).

Larguèche, Evelyne (ed.). 2004a. *L'injure, la société, l'islam: Une anthropologie de l'injure. Revue des Mondes Musulmans et de la Méditerranée* 103–104.

——. 2004b. "L'injure comme objet anthropologique". Larguèche (2004a:29–56).

Moussaoui, Abderrahmane. 2004. "La politique de l'injure: Une décennie meurtrière en Algérie". Larguèche (2004a:165–179).

Ould al-Barra, Yahya and Abdel Wedoud Ould Cheikh. 2004. "L'injure comme délit: L'approche des *fuqahâ'*, théologiens-légistes musulmans". Larguèche (2004a:57–80).

Roth, Arlette. 2004. "Entre les contraintes de l'éthique musulmane et le système de l'honneur, peut-il y avoir un usage toléré de l'injure?". Larguèche (2004a:81–101).

Taine-Cheikh, Catherine. 2004. "L'injure en pays maure ou 'qui ne loue pas critique'". Larguèche (2004a:103–126).

Evelyne Larguèche (CNRS, Laboratoire d'Anthropologie sociale)

Interface Linguistics

Grammar is modularized, as evidenced by the fact that distinct types of grammatical processes impose their effects autonomously, unaware of the contributions of other types of processes. Generalizations of the form 'subjects precede verbs' have no counterparts of the form 'terms containing the phoneme /p/ precede verbs'. Nor do generalizations such as 'stops become voiced intervocalically' have counterparts of the form 'stops become voiced in subjects'. This division of labor is nonetheless the apportionment of a common goal, the production of a linguistic expression in which the work of the modules is brought together. In points of intersection, processes are at work that belong neither entirely to one module nor to the other but lie in the interface, the space between modules where their autonomy melts away. The following discussion examines salient interface phenomena in Arabic in the phonetics-phonology interface, the phonology-morphology interface, the morphology-syntax-semantics interface, and the syntax-semantics interface.

1. The phonetics-phonology interface

Much recent work in phonology revolves around the issue of the extent to which phonological processes are guided by articulatory and perceptual (i.e. phonetic) considerations (Ohala 1974, 1983; Archangeli and Langendoen 1994; Steriade 1995; Jun 1995; Kaun 1995; Flemming 1995; Silverman 1995; Kirchner 1998; Hayes 1999). Phonological processes that appear to be motivated by phonetic considerations are said to be 'grounded'. Grounded phonological processes lie at the interface between phonetics and phonology. Davis (1995) characterizes the opacity of certain segments to emphasis spread in Arabic as a grounded phenomenon. In a dialect of Palestinian, the phonemes [i], [j], [ʃ], and [ʤ] block the rightward spread of retraction of the tongue root (the feature RTR), hence, e.g., [ʔ**ɑɫfɑːl**] 'children' and [**s̴ɑbɑːħ**] 'morning' (where the bold characters are those pharyngealized by emphasis spread – RTR assimilation – including the trigger), but [**ɫ**iːnak] 'your mud' and [**s̴ɑ**jjaːd] 'hunter'. The class of opaque phonemes share a high tongue body position, which is antagonistic with the low tongue position associated with RTR (Archangeli and Langendoen 1994). The blocking effect is a physiological antagonism. It is not a categorical impossibility, since it constrains only progressive spreading. Regressive spreading is unconstrained, hence [**xɑjjɑːɫ**] 'tailor' and [**nɑʃɑːɫ**] 'energy' (where the bold characters are those pharyngealized by leftward spreading, including the trigger, which is the last segment in these words). Hence, regressive and progressive emphasis spreading are distinct processes, one grounded and the other not, meaning emphasis spreading is a true interface phenomenon, neither purely phonological nor physiologically epiphenomenal.

McCarthy (1994) proposes that physiological considerations define the class of gutturals ([ʔ], [h], [ħ], [ʕ], [x], and [ɣ]). The gutturals do not appear to share an articulator or a place of articulation, the first pair being articulated in the larynx by the vocal chords, the second in the pharynx by the tongue root and epiglottis, the third at the uvulum by the tongue dorsum, but they nonetheless behave as a phonological class. McCarthy proposes, following Perkell (1980), that distinctive features correspond to patterns of orosensory feedback. The lack of phonological distinctiveness among the gutturals is directly due to poor neural innervation of the posterior region of the vocal tract, which obscures definition in the feedback from that area. The orosensory feedback from the posterior region consists of proprioceptive input from an undifferentiated posterior region and auditory input from the high F_1 formant typical of the gutturals (Klatt and Stevens 1969; Al-Ani 1970; and others). The proprioceptive indistinguishability of the gutturals, an epiphenomenon of the neural structure of the vocal tract, defines the gutturals as a phonological class, which McCarthy terms 'pharyngeal'.

2. THE PHONOLOGY-MORPHOLOGY INTERFACE

Phonological processes are commonly sensitive to the morphological context in which they apply. Level ordering of phonological rules is a paradigm case. Arabic displays level-ordering effects in stress placement and ellipsis as described in Brame (1971). In his review of Abdo (1969), Brame points out that the elision of unstressed high front vowels in Palestinian Arabic, as in (1b) and (1c), is blocked when a word is lengthened by clitic affixation but not by agreement affixation, a phonological distinction related to the morphological structure of the word.

(1) a. [$_{V}$ símiˁ-Ø]
hear-3sg/sub
'He heard'

b. [$_{V}$ smíˁ-t]
hear-1sg/sub
'I heard'

c. [$_{VP}$ [$_{V}$ simíˁ-Ø]-kum]
hear-3sg/sub-2pl/obj
'He heard you [pl.]'

Affixation of *-kum* in (1c) shifts stress to the right after initial stress is established in the base, the constituent Brame labels V, as in (1a). The prior placement of stress on the initial syllable manifests itself at the VP level as a secondary stress on that syllable, which in turn blocks the vowel deletion seen at work in (1b). A categorial morphological distinction (agreement inflection vs. clitic pronoun) affects the morphological constituency of the word and, in turn, phonological processes sensitive to that constituency, such as stress, a morphology-phonology interaction.

Phonological processes in Arabic appear to be sensitive to lexical semantic classes as well. Regular rules of glide deletion are suspended in the class of verbs of color or defect, e.g. *sawida* 'to become black' and *ˁawira* 'to become one-eyed'. These forms evade rules that normally delete glides in the phonological contexts presented there. The same roots are subject to glide deletion when another interpretation is available for them. For example, the root *s-w-d* surfaces as *sāda* when construed to mean 'to reign' (underlyingly /sawada/), and a morphological augmentation of *ˁawira*, *ˀaˁāra* (underlyingly /ˀaˁwara/), means 'to loan' (Brame 1970). These facts demonstrate a phonological process that is sensitive to the meaning of the term it applies to, not merely its phonological and morphological form, information that is outside the domain of the phonology module proper. Brame notices that a metathesis rule is subject to the same exception. Verbs constructed from roots with identical final consonants require them to be adjacent in certain phonological contexts, hence, underlying /madada/ 'to spread' surfaces as *madda*. Verbs of defect evade metathesis, e.g. *ṣakika* 'to be knock-kneed', *ˀalila* 'to smell bad'. Brame concludes that glide deletion involves metathesis. At the relevant level they are one and the same rule.

A great deal of Arabic morphology is templatic, a quasi-productive derivational morphology in which conditions on phonological and phonotactic well-formedness play a substantial role in the morphological expression of syntactic and semantic features, a paradigmatic case of phonology-morphology interaction. Because of the derivational significance of templatic morphology, it is discussed in Section 3 in the context of the morphology-syntax-semantics interface.

3. THE MORPHOLOGY-SYNTAX-SEMANTICS INTERFACE

The morphology-syntax-semantics interface deals with the manner in which a word's morphophonological characteristics influence its syntactic distribution and/or semantic interpretation (the term is not meant to subsume the syntax-semantics interface, discussed in Section 4). As mentioned above, phonology and morphology are unusually intimately intertwined in Arabic. Such an interaction is exemplified by the Arabic verb forms. Verbs are constructed from a three- (but sometimes two- or four-) consonant root filling consonantal positions in a prosodic template. The templates differ from one another in their phonological structure and the manner in which root consonants associate with prosodic positions. The templates are not productive, and not all templates are attested for all roots. But when a root is extant in more than one template, the meanings of the words so formed and the syntactic frames in which they occur differ in systematic ways. That is, alternations in the phonological and prosodic structure of verbs correlate with units of meaning and associated syntactic properties.

For example, doubling of the middle consonant of the root yields either a → causative form (compare *fariḥa* 'to be happy' with *farraḥa* 'to make happy') or an intensive or repetitive form (compare *ḍaraba* 'to beat' with *ḍarraba* 'to beat violently/repeatedly'). Prefixation of *ʾa-* forms a noncoercive causative (compare *jalasa* 'to sit down' with *ʾajlasa* 'to bid one to sit down' or *ʿalima* 'to know' with *ʾaʿlama* 'to inform'). Causativization is valency-increasing, i.e., the derived forms occur in a syntactic frame with one additional argument vis-à-vis the underived form.

(2) a. *fariḥa* *samīr-un*
be:happy Samir-NOM
'Samir was happy'
b. *farraḥa* *ʾaḥmad-u* *samīr-an*
make:happy Ahmed-NOM Samir-ACC
'Ahmed made Samir happy'

(3) a. *ʿalima* *samīr-un* *al-ḥukm-a*
know Samir-NOM the-judgment-ACC
'Samir knew/learned about the judgment'
b. *ʾaʿlama* *al-muḥāmī* *samīr-an* *al-ḥukm-a*
inform the-attorney Samir-ACC the-judgment-ACC
'The attorney informed Samir of the judgment'

Other templates preserve the argument structure of the base but affect its temporal/aspectual character. Doubling of the final consonant invariably forms a verb from an adjective (i.e. a finite form from a non-finite form; compare *ʾaḥmar* 'red' with *iḥmarra* 'to become red' or *ʾaʿwaj* 'curved' with *iʿwajja* 'to become curved'). Lengthening of the first vowel derives an activity from an accomplishment (compare *qatala* 'to kill' with *qātala* 'to fight with' or *sabaqa* 'to outrun' with *sābaqa* 'to run a race with').

Morphologically unaugmented verbs vary in the 'stem vowel', i.e. the second vowel of the stem, which may be *i, a*, or *u*. The stem vowel *u* systematically occurs in verbs that describe permanent states (Wright 1981:2.38), e.g. *ḥasuna* 'to be beautiful', *ṯaqula* 'to be heavy', *kabura* 'to be big'.

There is no universal generalization about the meanings of verbs with stem vowel *i* or *a*, but the following implicational generalization is systematic. Whenever the same root occurs with both *i* and *a*, the *i*-form is unaccusative (→ middle verb) and the *a*-form is its causative (Fassi Fehri 1987).

(4) a.	*ḥazina*	'to be sad'
	ḥazana	'to cause to be sad'
b.	*hadima*	'to fall apart'
	hadama	'to demolish'
c.	*wajira*	'to be frightened'
	wajara	'to frighten'
d.	*xariba*	'to go to ruin'
	xaraba	'to destroy'
e.	*najiza*	'to be implemented'
	najaza	'to implement'

See Doron (2003) for a detailed semantic analysis of the form-meaning correspondences in Semitic template systems, and McCarthy and Prince (1990a, 1990b), Ussishkin (1999, 2001), Idrissi (1997), and Ratcliffe (1998) for discussions of templatic morphophonology in Semitic.

Morphosyntactic regularities are found in templates other than the verbal templates. In some varieties of Arabic, agentivity and eventiveness are morphologically marked in passive participles. In Lebanese Arabic, passive participles of basic (morphologically unaugmented) verbs may be formed on either the template *minfaʿel* or *mafʿūl*, the latter more productive than the former (where *f*, *ʿ*, and *l* stand for the first, second, and third consonants of a triconsonantal root). Examples (5) and (6) below demonstrate that the *mafʿūl* participles license agent-oriented adverbs like *bi-diʾʾe* 'carefully', a diagnostic of agentivity (Manzini 1983), and are grammatical in the progressive, a diagnostic of eventiveness (Vendler 1957). The *minfaʿel* participles fail these diagnostics as in (7) and (8).

(5) a. *š-šrīṭ* *kēn* *maʾtūʿ* *bi-diʾʾe*
the-wire was cut carefully
'The wire was cut carefully'

b. *š-šrīṭ* *ʿam* *byikun* *maʾṭūʿ*
the-wire PROG is cut
'The wire is being cut'

(6) a. *l-bēb* *kēn* *maʾfūl* *bi-diʾʾe*
the-door was opened carefully
'The door was opened carefully'
b. *l-bēb* *ʿam* *byikun* *maʾfūl*
the-door PROG is opened
'The door is being opened'

(7) a. **š-šrīṭ* *kēn minʾeṭiʿ* *bi-diʾʾe*
the-wire was cut carefully

b. **š-šrīṭ* *ʿam* *byikun* *minʾeṭiʿ*
the-wire PROG is cut

(8) a. **l-bēb* *kēn* *minʾefil* *bi-diʾʾe*
the-door was closed carefully

b. **l-bēb* *ʿam* *byikun* *minʾefil*
the-door PROG is closed

4. THE SYNTAX-SEMANTICS INTERFACE

Gottlob Frege's Principle of Compositionality (Frege 1892) states that in languages that obey the principle (clear cases being manufactured logical languages), the meaning of an expression is a function of the meaning of its parts. Natural languages appear to obey this principle to a great extent, although it is unclear whether they can be described as fully compositional (Higginbotham 1986). In a fully compositional language, the only kinds of ambiguity that are expected to arise are lexical ambiguities, e.g. the different meanings of 'bank' in English ('riverbank' vs. 'financial institution'). The meaning of a sentence is otherwise pinned down by its structure. But there are many cases in natural language in which the interpretation of an expression appears to diverge from its structure and therefore from a one-to-one relationship between structure and meaning. These divergences occur at the interface between syntax and semantics and constitute the focus of linguistic inquiry on the relationship between structure and meaning.

There is not uncommonly a divergence, for example, between the apparent hierarchical order of quantifiers in a sentence (based on their linear order) and their relative scope. A → quantifier is in the scope of another quantifier if the interpretation of the first is subordinate to the interpretation of the second. In (9a) below, for example, the indefinite *qaṣīdatan* 'a poem', which asserts the existence of a poem, is (most saliently) interpreted subordinate to *kullu šābbin* 'every boy', so that a potentially different poem is asserted to exist for each boy, the poems varying with the boys. In (9b), on the other hand, *šābbun* 'a boy' asserts the existence of only one boy; it is not subordinate to any other quantifier.

(9) a. *qaraʾa kull-* *šābb-in* *qaṣīdat-an*
recited every-NOM boy-GEN poem-ACC
'Every boy recited a poem'

b. *qaraʾa* *šābb-un* *kull-a*
recited boy-NOM every-ACC
qaṣīdat-in
poem-GEN
'A boy recited every poem'

But the sentences above are ambiguous. In both cases, the opposite scopal interpretation is available. That is, (9a) may mean 'a (particular) poem is such that every boy recited it', and (9b) may mean 'every poem is such that a (potentially different) boy recited it'. In these interpretations, the scopal order is the inverse of the linear order. The fact that the hierarchical arrangements that instantiate the scope of quantifiers in a sentence may diverge from those hierarchical arrangements that instantiate their linear order suggests that sentences are systematically related to 'logical forms', representations that feed the interpretive component of language use but not the articulatory component.

If this is so, the syntax-semantics interface lies in the relationship between surface representations, or 'phonological forms', and these compositional semantic representations, or 'logical forms'. A conventional line of reasoning about this relationship is that it is transformational (May 1985). Logical forms are derived from phonological forms by the same transformations that derive phonological forms from other phonological forms, i.e. syntactic displacement. In (9a) and (9b), the object may (covertly) raise to the sentence-initial position (a transformation known as 'quantifier raising'), where the subject falls in its scope. This view preserves the Principle of Compositionality by casting the ambiguity in (9a) and (9b) as syntactic, not semantic. Each syntactic structure is associated with one and only one meaning, but a sentence may be associated with more than one structure at the syntax-semantics interface.

Properties of Arabic support this view. The availability of the object-wide scope reading of the sentences in (9) is contingent on the position of the subject with respect to the verb. If the subject precedes the verb, the inverse scope reading is unavailable. Note that in Standard Arabic, an indefinite subject must be modified in order to be licit in the preverbal position, another syntax-semantics interaction (Mohammad 1999).

(10) *šābb-un* *ṣaġīr-un* *qaraʾa*
boy-NOM young-NOM recited
kull-a *qaṣīdat-in*
every-ACC poem-GEN
‘A specific young boy recited every poem’
Not: ‘For each poem, a potentially different young boy recited it’

That is, the availability of an inverse scope reading depends on the syntactic form of the sentence, indicating that certain syntactic structures block certain interpretations, which correlates structure and interpretation. It indicates, in particular, that quantifier raising may raise an object to a position superior to a postverbal subject, but not a preverbal subject, and more generally, that quantifier raising is not unbounded but targets particular syntactic positions, as in Beghelli and Stowell (1997), Kitahara (1996), and Hornstein (1999).

Quantifier raising lies at the interface of syntax and semantics, as does its inverse, ‘reconstruction’. When a displaced term is interpreted, for the purposes of scope or binding, as if it occurred in its canonical position, it is said to have been reconstructed. Mohammad (1989) reports that a term that binds a pronoun in its clause must either be structurally superior to the pronoun’s canonical position (its position prior to any surface word order changes), or precede it in the surface order (see also Fassi Fehri 1993). Since the canonical order (subject > object) can be reinstated by reconstruction, a subject may bind a pronoun in an object regardless of surface precedence (11a) and (11b), but an object may only bind a pronoun in a subject if the object precedes the subject (12a), not if it follows (12b). The judgments below reflect the reading in which the pronoun is bound by the NP *ʾaḥmad*.

(11) a. *ḍaraba ʾaḥmad-u ṣadīq-a-hu*
hit Ahmed-NOM friend-ACC-his
‘Ahmed hit his friend’
b. *ḍaraba ṣadīq-a-hu ʾaḥmad-u*
hit friend-ACC-his Ahmed-NOM
‘Ahmed hit his friend’

(12) a. *ḍaraba ʾaḥmad-a ṣadīq-u-hu*
hit Ahmed-ACC friend-NOM-his
‘Ahmed’s friend hit him’

b. **ḍaraba ṣadīq-u-hu ʾaḥmad-a*
hit friend-NOM-his Ahmed-ACC
‘Ahmed’s friend hit him’

That reconstruction is impossible in (13), taken from Ouhalla (1994), suggests that recipient > theme is the canonical order for objects.

(13) **ʾaʿṭay-tu ṣāḥib-a-hu*
gave-1SG owner-ACC-its
kull-a kitāb-in
every-ACC book-GEN
‘I gave its owner every book’

Aoun and Benmamoun (1998) and Aoun a.o. (2001) point out that in Arabic, reconstruction is more restricted than displacement itself. Aoun and Benmamoun show that in Lebanese Arabic, a displaced term may index a syntactic position across what is normally a syntactic island (a barrier for displacement), if the indexed position is identified by a pronominal clitic (as opposed to a gap). That is, islands do not restrict displacement when the displaced term binds a clitic. However, antecedent-clitic chains only display reconstruction when not separated by an island. Hence, the possessive pronoun *-un* ‘their’ in (14a) may be bound by the quantifier *kell l-mʿallmēt* ‘the teachers’, but not in (14b). The difference is that the term containing the pronoun, *tlēmiz-un z-zġār* ‘their young students’, is separated from the clitic pronoun it binds by an island (a relative clause) in (14b) but not (14a).

(14) a. *[tlēmiz-un*$_j$ *z-zġār]*$_i$ *ʾəl-to*
students-their the-young said-2SG
[kəll l-mʿallmēt]$_j$ *bifaḍḍluw-un*$_i$

all the-teachers prefer-them
‘Their young students, you said all the teachers prefer them’

b. *[*tlēmiz-un*$_j$ *z-zġār*]$_i$ *ʾəl-to*
students-their the-young said-2SG
[*kəll l-mʿallmēt*]$_j$ *ʾaṣaṣ-o*
all the-teachers punished-3PL
l-wlēd [$_{\text{Island}}$ *yalli ḍarab-uw-un*$_i$]
the-children that hit-3PL-them
‘Their young students, you said all the teachers punished the children that hit them’

(14b) is grammatical if the pronoun *-un* in *tlēmiz-un z-zġār* is not bound by *kell l-mʿallmēt*. That is, displacement itself is not blocked, but reconstruction is. Aoun and Benmamoun claim that since reconstruction is a property of movement chains (Hornstein 1984; Barrs 1986; Chomsky 1993), the ‘displaced’ term *tlēmiz-un z-zġār* has moved only in (14a), not (14b). In (14b) it is base-generated at the left clause edge and so does not reconstruct, because it has not been moved. It is displaced, then, only in the sense of not occurring in the canonical object position, which is postverbal.

Aoun a.o. (2001) show a similar effect for strong (nonclitic) subject pronouns in Lebanese Arabic. But interestingly, a quantifier may bind a strong pronoun only when the two are separated by an island (again a relative clause in (15b)).

(15) a. **kəll muttaham-e*$_i$ *ʿrəf-to ʾənno*
each suspect-F know-2SG that
hiyye$_i$ *nḥabas-it*
she imprisoned-3FSG
‘Each suspect, you know that she was imprisoned’

b. *kəll muttaham-e*$_i$ *šəf-to*
each suspect-F saw-2SG
l-muḥāme yalli byaʿrif
the-attorney that know
ʾənno hiyye$_i$ *harab-it*
that she ran:away-3SF
‘Each suspect, you saw the attorney that knows that she ran away’

Again, reconstruction is impossible in the context presented in (15b), demonstrated in (16), suggesting that the relation between the displaced term and the pronoun is not a movement relation when the two are separated by an island, as expected, since islands block movement.

(16) *[*təlmīz-a*$_i$ *l-kəslēn*]$_j$ *ma bədd-kun*
student-her the-bad not want-2PL
txabbr-o wala mʿallme$_i$ *ʿan l-bənt*
tell-2PL no teacher about the-girl
yalli huwwe$_j$ *zaʿbar maʿ-a*
that he cheated with-her
b-l-faḥṣ
in-the-exam
‘Her bad student, you don’t want to tell any teacher about the girl with whom he cheated on the exam’

Aoun a.o. (2001) claim that (15a) is ungrammatical for interface reasons as well. In the base structure for the derivation, the pronoun *huwwe* is an appositive modifier of the quantifier *kəll muttahame* before the quantifier is displaced. Following Emonds (1979), Aoun a.o. propose that appositive modifiers are interpreted as separate clauses conjoined with the clause in which they occur in the surface representation, so that binding between the quantifier and the pronoun fails in the logical form, as the pronoun is then not structurally subordinate to the quantifier. In summary, the breakdown in each of the ungrammatical sentences in (12)–(16) arises because the logical forms of these sentences do not instantiate the relevant configurations. These sentences cannot be mapped to logical forms that license the relevant configurations because of constraints on the interface between syntax and semantics, in particular constraints on displacement. These facts lend credence to the hypothesis described above that the constraints on the interface are the same as those that constrain the derivation of surface forms.

Another syntax-semantics interaction typical of Arabic is the manner in which inflectional distinctions in the finite verb reflect the position and/or interpretation of its subject. Again to take Lebanese as an example (see Hoyt 2000 on Palestinian), unaccusative verbs may optionally fail to agree with a postverbal indefinite subject. However, the presence of agreement correlates with the discourse semantic property of specificity. A noun phrase (NP) interpreted specifically must refer back to a previously mentioned discourse referent, while an NP interpreted nonspecifically must introduce a new discourse referent (Enç 1991). The discourse in (17)–(19), in Lebanese Arabic, demonstrates this; (17i) and (17ii) are two possible continuations of the sentence in (17).

(17) *šōför l-baṣ ʾəll-na ʾənno ʿiddet*
driver the-bus told-us that several
baṣēt ʿilʾ-o bi-ʿažʾet sēr
buses stuck-3PL in-jam traffic
'The bus driver told us that several buses were stuck in traffic'
(i) *baʿdēn smiʿ-na ʾənno wəṣil*
later heard-1PL that arrived
tlēt baṣēt mʾaxxarīn
three buses late
'Later we heard that three buses arrived late'
(ii) *baʿdēn smiʿ-na ʾənno wəṣl-o*
later heard-1PL that arrived-3PL
tlēt baṣēt mʾaxxarīn
three buses late
'Later we heard that three buses arrived late'

(18) *ʿrif-t sēʿita ʾənno hūle kēn-o*
knew-1S then that these were-3PL
min bayn l-baṣēt lli
of among the-buses that
ʾəll-na ʿan-un š-šōför
told-us about-them the-driver
'Then I knew those were the buses that the driver was talking about'

Sentence (17) introduces several buses as discourse-referent. In the continuation in (17i), the verb *wəṣil* 'arrived' does not agree with the indefinite subject *tlēt baṣēt* 'three buses' (cf. [17ii]), and the indefinite is interpreted as introducing a new discourse referent – three additional buses. In the continuation in (17ii), the verb agrees with the indefinite, and the indefinite is interpreted as referring back to a previously introduced discourse referent, the previously mentioned buses in this case. That is, (17ii) asserts that the buses that arrived late were among the buses the driver said were stuck in traffic, while (17i) does not make this assertion. As a result, the assertion in (18) is infelicitously redundant following continuation (17ii) (it asserts only what [17ii] already asserted), but informative following continuation (17i) (it asserts more than what [17i] asserted). (17i) and (17ii) differ only in the presence of agreement on the verb, meaning this inflectional morphological characteristic has a significant impact on the interpretation of the sentence.

The phenomenon in (17) and (18) illustrates a case in which a particular morpheme, the agreement affix, affects the interpretation of another constituent in the sentence, the subject. The fact that subjecthood is a syntactic notion suggests that the dependency is mediated structurally, and indeed the relationship is impacted by word order. When an indefinite subject precedes the verb, the verb obligatorily agrees (and consequently, as expected, the subject is interpreted specifically).

(19) *tlēt baṣēt wəṣl-*(o) mʾaxxarīn*
three buses arrived-*(3PL) late
'Three buses arrived late'

5. SUMMARY

The phenomena discussed above illustrate interactions between phonetics and phonology, phonology and morphology, morphology and syntax, and syntax and semantics. These interactions demonstrate limits to modularity and reify the coherence of language. Interface linguistics in Arabic is a rich and varied domain, with unique empirical contributions to the theory of language and with substantial complexities yet to be explained.

BIBLIOGRAPHICAL REFERENCES

Abdo, Daud. 1969. *On stress and Arabic phonology: A generative approach*. Beirut: Khayats.

Al-Ani, Salman. 1970. *Arabic phonology*. The Hague: Mouton.

Aoun, Joseph and Elabbas Benmamoun. 1998. "Minimality, reconstruction and PF movement". *Linguistic Inquiry* 29.569–597.

Aoun, Joseph, Lina Choueiri, and Norbert Hornstein. 2001. "Resumption, movement, and derivational economy". *Linguistic Inquiry* 32.371–403.

Archangeli, Diana and D. Terence Langendoen. 1994. *Grounded phonology*. Cambridge, Mass.: MIT Press.

Barrs, Andrew. 1986. *Chains and anaphoric dependence: On reconstruction and its implications*. Ph.D. diss., Massachusetts Institute of Technology.

Beghelli, Filippo and Timothy Stowell. 1997. "The syntax of distributivity and negation". *Ways of scope taking*, ed. Anna Szabolcsi, 71–108. Dordrecht: Kluwer Academic Publishers.

Brame, Michael. 1970. *Arabic phonology: Implications for phonological theory and general Semitic*. Ph.D. diss., Massachusetts Institute of Technology.

——. 1971. "Stress in Arabic and generative phonology". *Foundations of Language* 7.556–591.

Chomsky, Noam. 1993. "A minimalist program for linguistic theory". *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger*, ed. Kenneth Hale and Samuel Jay Keyser, 1–52. Cambridge, Mass.: MIT Press.

Davis, Stuart. 1995. "Emphasis spread in Arabic and grounded phonology". *Linguistic Inquiry* 26.465–498.

Doron, Edit. 2003. "Agency and voice: The semantics of the Semitic templates". *Natural Language Semantics* 11.1–67.

Emonds, Joseph. 1979. "Appositive relatives have no properties". *Linguistic Inquiry* 10.211–243.

Enç, Mürvet. 1991. "The semantics of specificity". *Linguistic Inquiry* 22.1–25.

Fassi Fehri, Abdelkader. 1987. "Anti-causatives in Arabic, causativity and affectedness." Ms., Massachusetts Institute of Technology.

——. 1993. *Issues in the structure of Arabic clauses and words*. Dordrecht: Kluwer Academic Publishers.

Flemming, Edward. 1995. *Perceptual features in phonology*. Ph.D. diss., University of California Los Angeles.

Frege, Gottlob. 1892. "Über Sinn und Bedeutung". *Zeitschrift für Philosophie und Philosophische Kritik* 100.25–50.

Hayes, Bruce. 1999. "Phonetically driven phonology: The role of optimality theory and inductive grounding". *Functionalism and formalism in linguistics*. I. *General papers*, ed. Michael Darnell, Edith Moravcsik, Frederick Newmeyer, Michael Noonan, and Kathleen Wheatley, 243–285. Amsterdam and Philadelphia: J. Benjamins.

Higginbotham, James. 1986. "Linguistic theory and Davidson's program in semantics". *Truth and interpretation: Perspectives on the philosophy of Donald Davidson*, ed. Ernest LePore, 29–48. Oxford: Blackwell.

Hornstein, Norbert. 1984. *Logic as grammar*. Cambridge, Mass.: MIT Press.

——. 1999. "Minimalism and quantifier raising". *Working minimalism*, ed. Samuel Epstein and Norbert Hornstein, 45–75. Cambridge, Mass.: MIT Press.

Hoyt, Fred. 2000. *Agreement, specificity effects and phrase structure in rural Palestinian Arabic existential constructions*. M.A. thesis, Cornell University.

Idrissi, Ali. 1997. "Plural formation in Arabic". *Perspectives on Arabic linguistics,* X, ed. Mushira Eid and Robert Ratcliffe, 123–145. Amsterdam and Philadelphia: J. Benjamins.

Jun, Jongho. 1995. *Perceptual and articulatory factors in place assimilation: An optimality theoretic approach*. Ph.D. diss., University of California Los Angeles.

Kaun, Abigail. 1995. *An optimality-theoretic typology of rounding harmony*. Ph.D. diss., University of California Los Angeles.

Kirchner, Robert. 1998. *An effort-based approach to consonant lenition*. Ph.D. diss., University of California Los Angeles.

Kitahara, Hisatsugu. 1996. "Raising quantifiers without quantifier raising". *Minimal ideas*, ed. Werner Abraham, Samuel Epstein, Höskuldur Thráinsson, and C. Jan-Wouter Zwart, 189–198. Amsterdam and Philadelphia: J. Benjamins.

Klatt, Dennis and Kenneth Stevens. 1969. "Pharyngeal consonants". *MIT Research Laboratory of Electronics Quarterly Progress Report* 93.208–216.

Manzini, Maria Rita. 1983. "On control and control theory". *Linguistic Inquiry* 14.421–467.

May, Robert. 1985. *Logical form: Its structure and derivation*. Cambridge, Mass.: MIT Press.

McCarthy, John. 1994. "The phonetics and phonology of Semitic pharyngeals". *Phonological structure and phonetic form: Papers in laboratory phonology*, ed. Patricia Keating, III, 191–233. Cambridge: Cambridge University Press.

—— and Alan Prince. 1990a. "Prosodic morphology and templatic morphology". *Perspect-ives on Arabic linguistics*, II, ed. Mushira Eid and John McCarthy, 1–54. Amsterdam and Philadelphia: J. Benjamins.

——. 1990b. "Foot and word in prosodic morphology: The Arabic broken plural". *Natural Language and Linguistic Theory* 8.209–283.

Mohammad, Mohammad. 1989. *The sentential structure of Arabic*. Ph.D. diss., University of Southern California.

——. 1999. *Word order, agreement and pronominalization in Standard and Palestinian Arabic*. Amsterdam and Philadelphia: J. Benjamins.

Ohala, John. 1974. "Phonetic explanation in phonology". *Papers from the Parasession on Natural Phonology*, 251–274. Chicago: Chicago Linguistic Society.

——. 1983. "The origin of sound patterns in vocal tract constraints". *The production of speech*, ed. Peter MacNeilage, 189–216. New York: Springer.

Ouhalla, Jamal. 1994. "Verb movement and word order in Arabic". *Verb movement*, ed. David Lightfoot and Norbert Hornstein, 41–85. Cambridge: Cambridge University Press.

Perkell, Joseph. 1980. "Phonetic features and the physiology of speech production". *Language production*. I. *Speech and talk*, ed. B. Butterworth, 337–372. London: Academic Press.

Ratcliffe, Robert. 1998. *The 'broken' plural problem in Arabic and comparative Semitic: Allomorphy and analogy in non-concatenative morphology*. Amsterdam and Philadelphia: J. Benjamins.

Silverman, Daniel. 1995. *Acoustic transparency and opacity*. Ph.D. diss., University of California Los Angeles.

Steriade, Donca. 1995. "Underspecification and markedness". *The handbook of phonological theory*, ed. John Goldsmith, 114–174. Oxford: Blackwell.

Ussishkin, Adam. 1999. "The inadequacy of the consonantal root: Modern Hebrew denominal verbs and output-output correspondence". *Phonology* 16.401–442.

——. 2001. *The emergence of fixed prosody*. Ph.D. diss., University of California, Los Angeles.

Vendler, Zeno. 1957. "Verbs and times". *Philosophical Review* 66.143–160.

Wright, William. 1981. *A grammar of the Arabic language*. Beirut: Librairie du Liban.

Peter Hallman (McGill University)

Interference

Broadly speaking, there are two sources of interference in Arabic, one extraneous and one indigenous, affecting both the spoken and written forms of the language. Extraneous interferences come from the foreign languages with which speakers of Arabic come into contact. Indigenous interferences come from local spoken varieties of Arabic and the written form of the language itself. Contact interference is not uncommon in other languages. In the case of Arabic, however, the matter is more complex, owing to factors peculiar to the language and region itself.

Arabic has been in contact with foreign languages since its first appearance as a world lan-guage, which coincided with the expansion of Arab/Islamic civilization in the late 7th century C.E. (or the early 1st century A.H.). As Arabic-speaking armies and migrants spread into neighboring regions and beyond, they came into contact with foreign languages (→ language contact).

It can be argued that even before the coming of Islam, foreign languages had an impact on Arabic, although the extent of their interference is difficult to assess. Speakers of Arabic resided outside of the Arabian Peninsula well before the Islamic era. From as early as the 1st millennium C.E., and possibly well before that, Arabic speakers were in contact with peoples of the Fertile Crescent and the Levant, who spoke related languages, such as → Aramaic and Hebrew (Bowersock 1983; Abbot 1939). However, because these languages are so closely related, and early historical records are so sparse, it is difficult to determine the extent to which they affected one another.

A noticeable effect of these early contacts can be found in the first lengthy Arabic writing, the *Qur'ān*, encoded in the late 7th century C.E. It exhibits scores of borrowings from languages such as Greek (→ Greek loanwords), Persian (→ Persian loanwords), and other more closely related languages, such as Hebrew and Aramaic (→ Aramaic/Syriac loanwords). Early scholars produced catalogs of foreign borrowings found in the *Qur'ān* (Jeffery 1938; Rippin 1981).

It is, moreover, likely that the spoken varieties of the Arabian Peninsula differed one from another. It is also likely that all of them differed from the formal, elegiac language of poetry (Rabin 1951; Zwettler 1978), which, by the time the Arabs had migrated in large numbers into neighboring lands, had become the language of religious scripture (→ poetic koine).

With the expansion of the Arab/Islamic civilization outside of the Arabian Peninsula, Arabic became the language spoken over a vast geographical area. With a resulting need for written records, Arabic also became a formal written language, patterned after the Arabic used in scripture, elocution, and poetry.

As Arabs settled into the lands of the Fertile Crescent and North Africa, the spoken varieties of Arabic diverged even further from each other and from the variety that had become the written language. Exposure to Persian also increased under the Arab/Islamic civilization, with some of the eastern varieties adopting large Persian vocabularies. Later, the roughly five hundred years of Ottoman Turkish dominance of the Arab world provided a new avenue of entry for Persian, as well as Turkish, words into Arabic (→ Turkish loanwords).

In some areas, the existing (or → substrate) languages were entirely or almost entirely replaced by Arabic. In other areas, adstrate languages continued to exist alongside Arabic. Those becoming extinct or nearly extinct include Aramaic and → Coptic. Aramaic is still spoken in small holdout communities in the Levant and Iraq, while Coptic survives as a purely liturgical language in the Christian Church in Egypt. The spoken Arabic of Egypt contains thousands of loanwords from Coptic (→ Coptic loanwords), although they are not frequently used by all Egyptian speakers, especially urbanites (Bishai 1959; Versteegh 2001). A great many of the Coptic borrowings are related to agricultural products, processes, and implements; clothing and adornment; and Christian dogma and practice.

The Moroccan vernaculars provide an excellent illustration of the types of language interference to which Arabic is subject. Prior to Arab expansion into the area, → Berber was the spoken language of the region. Today, large Berber communities retain their language, which has in turn exerted a great deal of influence on the Arabic dialects of the region (→ Berber loanwords). French occupation of Morocco in the 19th and 20th centuries exerted a strong influence on colloquial Arabic. The combined interference from Berber and French sources has given the Moroccan dialect a reputation for incomprehensibility to speakers of Arabic from outside areas.

In truth, all of the regional dialects of Arabic are to some extent mutually incomprehensible, and the Moroccan dialect presents a good illustration of the principles involved to one degree or another in most of them. The impression in Arabic-speaking countries outside is that Moroccans cannot complete a sentence without lapsing into French. In reality, a great many French words have been adopted into the language and are used as if they are native elements (→ French loanwords). Moreover, French continues to exert an influence on Moroccan Arabic dialects, and new French words continue to enter the language. The manner in which this occurs involves a complex process of → code-switching and code mixing – the first involving an alternation between the two languages, often at clause and sentence boundaries, and the second involving a blending of the two, often within a single word (Heath 1984). This contributes to the perception among outside Arabic speakers that Moroccan dialects are incomprehensible hybrids.

Interference from European languages is by no means unique to Morocco. Various regional dialects of Arabic exhibit interference from modern foreign languages, depending on the recent history of the countries in which they are spoken. For example, the Algerian and Tunisian dialects, closely related to those of Morocco, also exhibit influence from French. For its part, Libyan Arabic contains many borrowings from Italian, arising out of Italian occupation of the country in World War II (→ Italian loanwords). Egyptian Arabic has borrowed words from French and English (→ English loanwords). This is because of the 19th-century orientation of Egyptian elites toward France in matters having to do with education and culture, engendered by the short-lived Napoleonic invasion of Egypt in 1798, and due to the British occupation, which began in the late 19th century. For similar reasons of occupation and cultural orientation, the Levantine dialects of Lebanon and Syria also demonstrate tremendous influence from French. The Gulf dialects contain many words of Persian origin, because of the Gulf's close proximity to and cultural interaction with Iran. Finally, the ascendance of English as an international language has influenced all Arabic dialects.

The path by which unfamiliar concepts are introduced involves what initially could be regarded as single-word switches into one of these second languages, immediately followed by a switch back to the matrix language. This switching into another code may have several motivations; one of the most important of these is that the concepts under discussion may be more easily expressed in one code than the other. In this case, Arabic speakers with European education may be more accustomed to expressing certain concepts in a foreign language than in Arabic. The same principle applies in contending with concepts that are entirely foreign to Arabic, such as new technologies or ideologies introduced from abroad.

This process can be observed in Arabic writing as well as in Arabic speech, whereby a single foreign word might appear in a sea of Arabic print, either written in its native alphabet or transliterated into Arabic script and often enclosed in quotation marks or parentheses. Of necessity, many of the borrowed concepts are nouns, expressing concepts unfamiliar to traditional Arabic culture. Prime targets are concepts from various new technologies such as those involving automobiles or electronics. A borrowed concept will undergo assimilation to Arabic phonology, and if the borrowing is completely successful, its lexeme may be reanalyzed as other parts of speech, and Arabic paradigms may be applied (Wilmsen 1996).

Phonological modifications may include, among others, /p/ > /b/, /v/ > /f/ or /w/, /ü/ > /u/ or /i/. Different regional varieties will treat foreign sounds differently according to the repertoire available to them. For example, Gulf Arabic tends to modify /v/ as /w/, as in *rewas* for *reverse*, while others favor /f/, as in the Egyptian *filla* for *villa*. Moroccan Arabic, with its extensive exposure to French, exhibits a much wider range of acceptance of foreign phones than other varieties do. Moreover, as most varieties of Arabic cannot tolerate consonant clusters of more than two – and in the word-initial position none at all – epenthetic vowels will be interpolated between two – often the first and second – consonants in a cluster. Word-initial consonant clusters beginning with sibilants may have a vowel placed in the word-initial position, as in *ʾistudiyu* 'studio.'

The amount of exposure to foreign languages attained by individual speakers will also dictate their own articulation of foreign phonological patterns. The word for 'rehearsal' in Egyptian Arabic, for example, is *birova* (< Italian *prova*), with an epenthetic vowel between the /b/ and the /r/, but, unusually, the /v/ is retained. More highly educated speakers can handle the pronunciation *prova* without difficulty. Similarly, the name of the make of the Peugeot automobile is generally pronounced *bījū*, but the more sophisticated speakers can muster something approximating the canonical French pronunciation, perhaps without the /ö/.

Often, but not always, a borrowed term with more than four consonants or three syllables will undergo syllable reduction and/or consonant elision to bring it into conformity with the triliteral or quadriliteral root system of Arabic, as in the treatment of the word 'television' as *tilfaz* in Peninsular varieties (but elsewhere it is *tilifizyōn* or *tilifizyō*).

Some nouns already containing only three or four consonants become so completely assimilated as to be assigned broken plurals, the default plural form for most foreign borrowings being the feminine sound plural *-āt*. Familiar examples of broken plurals applied to foreign nouns are *ʾaflām* (sg. *film*) and *bunūk* (sg. *bank*). Apparently, these particular examples apply in all regional varieties. Others may appear in one region but not in another, as, for example, the Egyptian *falātir* (sg. *filtir*, 'filter').

Well-assimilated nouns may be reanalyzed as verbs, very often of Form II. The process by which this happens seems initially to involve the use of a dummy verb compounded with a borrowed noun, as in *ʿamal sēf* 'to save [a computer file]', which with familiarity with the concept is reanalyzed to *sayyif*. These processes occur in an ad hoc fashion across the Arabophone world, with some regions adopting the foreign term and others supplying a loan translation (calque) or Arabic analogue. Sometimes, the same term can be used to denote different, even opposite, meanings in different regions. For example, in the Gulf, the borrowing *fannaš* means 'to resign', whereas in Egypt, if used at all, it is pronounced *fanniš*, meaning either 'to dismiss someone from employment' or more commonly 'to apply a finish'. The verbal noun *tafnīš* 'finishing' also exists and is in fact more often used than the verb itself to indicate the final coat of varnish (*warnīš*, which also means 'nail polish') on a piece of furniture or the final detailing of a garment.

It is worth noting that the same ad hoc procedure applies in the borrowing of → technical terminology into the written language. Various techniques are employed in the borrowing of unfamiliar concepts from other codes. The easiest of these is transliteration, whereby the foreign term is simply rewritten in Arabic letters, perhaps with some modifications. A further step involves 'Arabization' – exploiting the derivational properties of Arabic to coin new terms and introduce them into the Arabic lexicon. This may involve adopting some or all consonants of the borrowed term and incorporating them into an Arabic format. Alternatively, it may involve constructing a novel derivation of a native triliteral root conveying something of the semantic domain of the borrowed term. The latter alternative is in itself an example of another much-used technique: calquing or loan translation, wherein the meaning of a novel concept is translated into a native Arabic construction, often in periphrasis employing novel collocations.

A well-known example of these techniques may be seen in the concept of 'computer', which is borrowed 'as is' in Egypt, to give the term *kombiyutir*, adopted as an Arabized analogue in North Africa *ḥāsūb* 'something that by its very nature computes well', *rattāba* in the Levant (< French *ordinateur*), and as the calque *ḥāsib ʾālī* 'a computing machine' in the Arabian Peninsula. A similar term, *ʾāla ḥāsiba* 'calculating machine', is used in the Levant to indicate a calculator.

All told, borrowing and assimilation of foreign nouns affects the spoken language more than it does the written, inasmuch as concepts from fields such as consumer electronics and automobiles, food and fashion, or arts and entertainment are likely to be spoken about more often than written about. In addition, the pace of adoption and spread of foreign terms is slower in formal written Arabic because more rigid standards of acceptability are exercized in that medium. Be that as it may, many novel concepts will remain in the domain of formal spoken or written discourse.

The various → language academies meet regularly to discuss new concepts entering the language and to propose and publish lists of Arabized alternatives for the foreign terms associated with them. Aside from some of the laughable constructions they come up with – the classic example being their proposal for 'sandwich', *šaṭr wa-mašṭūr wa-baynahumā ṭāzij*, roughly 'two slices with something fresh between them', the glacial pace at which they contend with the flood of new terminology renders their efforts all but irrelevant.

At the same time, a variety of organizations and entities, ranging from international donor agencies to newspaper editorial boards, have a pressing need for formal Arabic technical terminology. They therefore engage in their own construction of concepts, very often in the form of loan translations or coinages. This duplication of efforts leads to variation in standard terminologies exhibited between regions and organizations.

It is especially in the domain of formal discourse over technical subjects that spoken varieties of Arabic interfere with formal written Arabic, and vice versa. To be more precise, the norms of spoken Arabic may interfere with the canons of written Arabic when that variety is spoken or declaimed. At the same time, canonical forms of the written language may impinge upon speech, especially in formal or technical discourse otherwise conducted in a vernacular (i.e. spoken) variety.

In either direction, the interaction can be complex. The consonants of the written variety in reading aloud, recitation, or declamation may be rendered as their colloquial analogues. This tendency generally applies to the alveolar and dental consonants, so, for example, /ḏ/ may be pronounced /z/, as it would be in colloquial analogues, producing such hybrid forms as *hāza* for the canonical *hāḏā* 'this' and *allazīna* for *allaḏīna* 'those which'.

Additionally, lexemes from one code may be inserted for effect into discourse conducted in the other. This is especially evident in speech, but it may occur in writing as well, if less often. Discourse conducted in the vernacular will incorporate specialized terminology from the written code, as appropriate. This can result in the incorporation of other features of the written code while declaiming aloud. An example might appear like this:

issabab inn ilʾaflām ilmulawwana btitkallif kitīr ʾawi
'The reason is that color film costs a whole lot'
wi lfilm ilmaṣri byuʿraḍ fi lbilād iššarqiyya faqaṭ
'And Egyptian films are shown in Eastern countries only'
wi da ma ykaffī š innafaqāt ilbāhiḏa illi titkallafha ilʾaflām ilmulawwana
'And that is not enough to cover the exorbitant expenses that burden color films'

As the above example demonstrates, the mixing between two codes is complex, and yet, the utterance is not at all unusual. Of the 21 words of which it consists (counting the words *mulawwana, film* along with its plural *ʾaflām*, which might also be considered a borrowed technical term, and the conjunction *wi* only once each, while the clitic pronoun *-hā*, in any case a shared feature of both codes, is left uncounted), eleven are lexemes common to both the written and the spoken forms of the language. Seven of those are marked by vernacular features such as the affixation of the vernacular variant of the definite article *il-* as opposed to the canonical written form *al-*. Six more of them arise unambiguously from the vernacular (those being *tikallif, kitīr, ʾawi, da, ykaffī, illi*). Only four are clearly drawn from writing, albeit three of them with colloquializing elements, such as the definite article on the word for 'expenses' *nafaqāt* (the preferred vernacular reflex being *takālīf*), the indicative prefix *bi-* in the passive voice verb *yuʿraḍ* (for which the vernacular variant *yitʿiriḍ* exists) and the verb *takallaf* 'to burden' with its 3rd person feminine singular clitic *ti-* drawn from the vernacular, the canonical form of the conjugated verb being *tatakallaf*. The sole lexeme drawn unambiguously from the written code is *faqaṭ*. Meanwhile, the single instance of negation, *ma-ykaffīš*, conforms to the vernacular grammar, even though the negative particle *ma-* is a shared feature of both codes. This is basically a vernacular utterance that incorporates a few features from the written code.

Nevertheless, in more formal speech declaimed in a spoken form of the written code, the vernacular language also interferes. This may be seen in the following example:

ʾinta ḥa-tuʿallim-ni ēh huwa lmugtamaʿ ilmadani? laʾ miš maʾʿūl!
‘Are you going to teach me what civil society is? No, unbelievable!’
ʾilmugtamaʿ ilmadani huwa da illi byiʿmal kida; byiʿmal xadamāt biyiʿmal waẓāʾif ʾibdāʾiyya
‘Civil society is precisely that which does such things; it offers services, it creates entry-level jobs’
biyiʿmal muḥāwalāt at-taġyīr, kull šeʾ. ma tkallimš ʿan nuṣṣ wi tṣīb innuṣṣ ittāni
‘It attempts to work change, everything. Don’t talk about one half and leave off the other half’
ʾilmugtamaʿ ilmadani wiḥda lā tatagassad
‘Civil society is an undifferentiated unit’

Using the same measure as that used with the previous utterance, it may be seen that of the 28 tokens in this utterance, 17 are shared by both the vernacular and written codes, generally marked by vernacular features. Seven are drawn from the vernacular; two are verbs from the written code (albeit one with the vernacular clitic *ḥa-* indicating futurity applied to it); and two other tokens, *ilmugtamaʿ ilmadani* ‘civil society’ and *waẓāʾif ʾibdaʾiyya* ‘entry-level jobs’, are calqued technical terms and can thus be considered to have been drawn from the formal written code. The second is not even a standard technical term, appearing to be a coinage of the speaker (the standard would be *waḏ̣āʾif al-mustajaddīn*).

It is worth noting that all of the verbs in both of the above examples exhibit vernacular features of one sort or another. Especially common in this type of formal or semiformal discourse is the *bi-* clitic, indicating the habitual and inchoate in the vernacular. In formal speech, this may be applied in some instances in a manner uncharacteristic of vernacular grammars (Doss 1987). Indeed, one of the features of this type of speech is that it regularly produces forms that violate the prescriptions of either code. This has led some to propose a third intermediate variety between the formal written code and the unadulterated vernacular. Others argue that these are nothing more than unstable, intermediate forms (→ educated Arabic; → *luġa wusṭā*).

Whatever the case, interference from the vernacular appears in declaimed formal Arabic, and interference from formal written Arabic occurs as well in the vernacular. This does not mean that entire discourses in either variety may not be conducted without interference; such occurrences are common and entirely appropriate, depending upon the situation in which the speech acts occur. On the other hand, interference between the codes is also quite common.

Generally, the perception and preference of native speakers of the language is toward a strict separation of the two. The two samples above, however, demonstrate that some leeway is allowed in extemporaneous speech, where the deliberate mixing of codes may be employed for effect. Inserting a vernacular expression into an otherwise staid and punctiliously correct formal declamation often serves to bring a point home or invest it with some emotive force. Similarly, to punctuate vernacular discourse with formalities from the written code may serve to elevate the discussion – or just the opposite: to express irony or ridicule. Deliberate style shifting aside, the sheer effort of declaiming extemporaneously in formal Arabic may prove too burdensome for speakers unaccustomed to the exercise, and vernacular features may begin to interfere as a consequence.

It also happens that vernacular features may find their way into serious writing. The following example comes from a feature article in a newspaper discussing the sale of rancid butter and other spoiled foodstuffs on the open market:

wa-lā baʾs ʾan yaʾkulahu an-nās al-galaba al-maḥrūmūn min al-laḥm fī baladinā
‘And it’s okay that poor people deprived of meat in our country eat it’
wa-miʿdatuhum taʾkulu z-zalat
‘Their stomachs are made of cast iron’ (lit. ‘their stomachs eat pebbles’)
wa-xalli n-nās taʾkul wa-tfalʿas
‘So let them eat and get fat’

Here, the writer appears to have employed colloquialisms deliberately, lapsing almost entirely into the vernacular by the end of the paragraph with the use of the thoroughly vernacular lexeme *falʿas* ‘to get fat’. It is also worth noting the deflected agreement (appearing as the fem. sg.), employed with the plural noun *nās* 'people', in the verbs *tfalʿas* 'get fat' and *taʾkul* 'eat', which, nevertheless, appear in canonical written form (the vernacular being *tākul* without the glottal stop), while displaying deflected agreement with the human plural. This appears to be an instance of morphosyntactic interference of spoken norms in writing.

Acceptable in either code, deflected agreement with human plurals occurs far more frequently in speech than in writing (Cachia 1969). Even so, that form of agreement may also appear in more straightforward writing, without exhibiting style shifting, as in the following sentence from a newspaper columnist largely considered to be a superb stylist:

ʾinna n-nās lā tuṣāb bi-ḥāla junūn mufājiʾ
'People are not afflicted with sudden madness'
ʾaw tuṣarrif ġayr mutawaqqaʿi bi-dūn ʾasbāb
'or act unexpectedly without reason'

Here, both verbs carry deflected agreement. Whether this amounts to interference from the vernacular or whether it is purely a stylistic alternative is unclear. It is seen as perfectly acceptable written style; all the same, the possibility that it represents interference from spoken Arabic cannot be dismissed.

Either way, just as speakers will style shift for effect, so too will writers in certain genres (e.g. humor and irony) deliberately insert colloquialisms into their writing (Rosenbaum 2000). Viewed from afar, this should not appear so terribly remarkable. Like any living language, Arabic will exhibit ranges of style. It will register between the formal and the casual and between spoken and written norms. So, too, will it be affected by events and concepts impinging on it from outside its native regions. What makes Arabic remarkable is its speakers' expressed preference for – and perception of – the strict segregation between the spoken and written codes. Regardless of such normative sentiments, the two codes do interfere and interact with each other.

Bibliographical references

Abbott, Nabia. 1939. *The rise of the North Arabic script and its Ḳurʾānic development, with a full description of the* Ḳurʾān *manuscripts in the Oriental Institute*. Chicago: University of Chicago Press.

Bishai, Wilson B. 1959. *The Coptic influence on Egyptian Arabic*. Ph.D. diss., Johns Hopkins University.

Bowersock, G.W. 1983. *Roman Arabia*. Cambridge, Mass.: Harvard University Press.

Cachia, Pierre J.E. 1969. "The use of colloquial in modern Arabic literature". *Journal of Near Eastern Studies* 87.12–22.

Doss, Madiha. 1987. "Further remarks on the use of b-imperfect in spoken literary Arabic". *Zeitschrift fur Arabische Linguistik* 17.93–95.

Heath, Jeffrey. 1984. *Functional grammar of Nunggubuyu*. Canberra: AIAS.

——. 1989. *From code-switching to borrowing: Foreign and diglossic mixing in Moroccan Arabic*. New York: Kegan Paul International.

Jeffery, Arthur. 1938. *The foreign vocabulary of the* Qurʾan. Baroda: Oriental Institute.

Rabin, Chaim. 1951. *Ancient West-Arabian*. London: Taylor's Foreign Press.

Rippin, Andrew. 1981. "Ibn ʿAbbās's *al-Lughāt fī ʾl-Qurʾān*". *Bulletin of the School of Oriental and African Studies* 44.15–25.

Rosenbaum, Gabriel M. 2000. "'Fusḥāmmiyya': Alternating style in Egyptian prose". *Zeitschrift fur Arabische Linguistik* 38.68–87.

Versteegh, Kees. 2001. "Linguistic contacts between Arabic and other languages". *Arabica* 48.470–509.

Wilmsen, David. 1996. "Code-switching, code-mixing, and borrowing in the spoken Arabic of a theatrical community in Cairo". *Perspectives on Arabic linguistics*, IX, ed. Mushira Eid and Dilworth Parkinson, 69–92. Amsterdam and Philadelphia: J. Benjamins.

Zwettler, Michael. 1978. *The oral tradition of Classical Arabic poetry: Its character and implications*. Columbus: Ohio State University Press.

David Wilmsen
(American University in Cairo)

Interjection

An interjection is a word, typically of indeclinable form, that in most languages has no grammatical or syntactic relationship to any other words or parts of a sentence (Hartmann and Stork 1972:115). An interjection may express emotional reactions and sensations, but also surprise, including curse and malediction, and it may serve to contact someone (Bussmann 1983:216–217). Therefore, it is used to exclaim, to protest, or to command. It is generally short, one syllable or word, and cannot be inquired after.

Classical Arabic possesses a large number of interjections (listed especially in Wright 1962: I, 294–296), e.g. *yā* 'oh' (as a particle also used for the vocative) and, with similar meaning, *wā, hā, ʾa*, and *ʾay* or *way* 'woe'. Most of these interjections have lost their importance in Modern Standard Arabic, as also in the modern Arabic dialects, where only a few have survived.

Some interjections have variant forms. They may occur with or without final vowels, which may be quite different. Others occur with or without the final *n* of the → nunation (→ *tanwīn*): *ʾah, ʾahi, ʾahin* 'oh', *ʾāh, ʾāhi, ʾāhin* (equal meanings). Further variation characterizes those interjections, for 'oh' also may be expressed by *ʾawih, ʾawha* or *ʾawhu, ʾawwah* or *ʾawwih, ʾawwāh, ʾawwāhi* or *ʾawwāhu*, as also by *ʾawwatāh, ʾawwi, ʾāwi, ʾāwin*, etc.

Numerous variations are known to express 'ugh, faugh, fie': *ʾuf, ʾuffu, ʾuffun, ʾuffi, ʾuffin, ʾuffa, ʾuffan, ʾuffatan*; to express 'excellent, bravo': *wāha, wāhi, wāhan, wāhā, wāh*; or 'well done, bravo': *bax, baxi, baxin, baxxin*.

The interjections already mentioned may be regarded as often repeated. Also used frequently are *ṣah* and *ṣahin* 'hush!, silence!', *mah* and *mahin* 'stop!, give up!, let alone!', *ʾīhan* 'be silent!, give up!', and also *ʾīhi* and *ʾīhin* 'go on!, proceed!'.

Interjections with more than one syllable are also in frequent use; see the enlargements of one-syllable words mentioned above, as well as other examples: *hayhāta, hayhāti* or *hayhātu* 'away with it; impossible', and *hayya, hayyi*, or *hayyā* 'make haste!', followed by a personal suffix.

Unlike interjections in other languages, Arabic forms may be followed by personal suffixes or certain grammatical cases. For example, *hāka* 'take!' is built from *hā* plus a suffix and in addition may receive suffixes: *hākahā* 'take her'; *hālumma* is followed by *ʾilā* in the meaning of 'come here!' but by the accusative in the meaning of 'bring here!'; *wayḥa* 'woe!' is followed by a genitive or a personal suffix, and *hayya* 'make haste, come!' is followed by *ʿalā*.

Some of these interjections are frequently used in reduplicated form: *hayyā hayyā* (same meaning as *hayyā* 'make haste!') and *bax bax* (same meaning as *bax* 'well done!'), and some of them imitate sounds, like *ṭīxi ṭīxi* (laughter) and *ʿuʿ ʿuʿ* and *huʿ huʿ* (vomiting).

Classical Arabic possesses a wide range of compound interjectional forms like *ʾallāhumma* 'oh God'; *ʾayyuhā* and *yā ʾayyuhā* 'oh' (also used as a vocative particle, requiring a definite article in the following noun, which must have a nominative case ending); *ʾayā* 'oh'; *ʾalā* and *ʾamā* 'oh no'; *yā la-* (used as a cry for help when followed by the genitive case, and as a cry of surprise when followed by a personal suffix or the accusative case or by the preposition *min*); and *way la-* 'woe' (with personal suffix added), as also *waylummi*, which is contracted from *waylun li-ʾummihi* (shows shock of admiration and is followed by personal suffix or the accusative case).

In certain atypical cases, long words are originally derived from short words, after enlargement: *ḥayyahala, ḥayyahalan, ḥayyahlā*, etc. 'make haste!; keep to!; call!', followed by personal suffix and *ʾilā, ʿalā*, or *bi-*.

Some Arabic forms are used as interjections only according to their function, not their grammatical origin, like *daʿ* 'let it!', or *taʿāla* 'come, go on!'.

A special group of interjections in Classical Arabic is that of calls to animals, especially to domesticated ones (an exhaustive list is to be found in Schulthess 1912). The sounds of the animals themselves are not to be counted as interjections in the real sense of the word. Among the interjections serving as calls to animals are those used in connection with driving different animals: *halā* (horses), *ʿadas* (mules), *hayda, hīda, ḥāʾi*, and *ḥawba* (camels), *hus* (sheep or goats); those in connection with making camels knee: *ʾīnix, hīxi, naxxi*; and those for calling camels to water: *jawta*. A dog is called by *qūsi* and driven away by *haj*.

Classical Arabic grammarians also may deal with interjections to an abounding extent (→ *ism al-fiʿl*). Az-Zamaxšarī (*Mufaṣṣal* 61–67) gives information about the Classical Arabic interjections, dividing them into certain groups, the calls to animals included. In addition to the examples given above, he also refers to adverbs like *ruwayda* and *tayda*, prepositional expressions like *ʿalayka* and *ʿalayya*, and to *šattānā, sarʿānā*, and *waškānā*, and he deals at length with *faʿāli* forms like *nazāli, tarāki, barāki*, etc.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCE

Zamaxšarī, *Mufaṣṣal* = ʾAbū l-Qāsim Maḥmūd ibn ʿUmar az-Zamaxšarī, *Kitāb al-mufaṣṣal fī n-naḥw*. Ed. J.P. Broch. Christiania: P.T. Mallingii, 1879.

SECONDARY SOURCES

Bussmann, Hadumod. 1983. *Lexikon der Sprachwissenschaft*. Stuttgart: A. Kröner.

Hartmann, R.R.K. and F.C. Stork. 1972. *Dictionary of language and linguistics*. London: Applied Science Publishers.

Schulthess, Friedrich. 1912. *Zurufe an Tiere im Arabischen*. Berlin: Preussische Akademie der Wissenschaften.

Wright, William. 1962. *A grammar of the Arabic language*, I. 3rd ed. Cambridge: Cambridge University Press.

Herbert Eisenstein (University of Vienna)

Internet

The discussion of Arabic and the Internet consists of three main parts. The first part, "Configuring the components", is a discussion about configuring a basic set of tools – software applications – and enabling them to support Arabic for browsing the Internet as well as storing and exchanging documents and doing e-mail, thereby enabling communication and interchange between students and experts. Configuration issues are only discussed as necessary for enabling Arabic support. The second part, "The Arabic Internet", outlines a small subset of Arabic site genres: searching, portals, news media, multimedia organizations, and chat forums. The last part deals with capturing data for processing and exchange.

1. CONFIGURING THE COMPONENTS

At the time of this writing, the two most commonly used operating systems for Arabic Internet information storage and retrieval systems are Windows and Mac. Regrettably, these two systems are not immediately compatible for data exchange because their schemes for Arabic text encoding are different.

The repertoire of Arabic alphanumeric characters is encoded as computer character sets using more than one standard encoding scheme. Consequently, applications using different encoding schemes for Arabic – as is the case with the Windows and Mac operating systems – cannot exchange Arabic documents unless the documents are translated. The Windows Arabic encoding scheme is called cp1256; the Mac Arabic encoding scheme is a superset of the ISO 8859–6 encoding scheme. For further information regarding encoding, see the links section below. In the future, as different operating systems move more fully to implementing the Unicode encoding standard, sometimes referred to as UTF, it may become possible to overcome current problems of Arabic document interchange.

In view of the foregoing, a selection has to be made regarding which of the three following operating systems to use: Windows, Mac, or Linux/Unix. This selection must be informed by the target audience. Who will read your files, whose files do you want to read? The selection must also be informed by the type of files you wish to create or use: will these be HTML, will they be word processor documents, or will they be application programs. Other important questions include: what tools and development environments will be used; will they be programming languages, or authoring environments; what kind of support will be provided; how large is the community of users/experts involved; and many other important questions. The selection thus has far-reaching consequences, and once made, it becomes a limiting one, and exchanging documents – which is what computers and the Internet are all about – across different platforms requires the extra, intermediate, and possibly problematic translation step.

In the Arab world, Windows is the choice because it is – by far – the most widely used operating system – if not the only one. This situation does not appear likely to change any time in the foreseeable future. Mac users are locked out of exchanging documents with Windows users. AWstats (advanced Web statistics) is a tool that generates advanced Web statistics about visitors to a site. A site running AWstats can determine a visitor's operating system and browser (see the AWstats site for details). For a glimpse of the situation in Norway, consider the *Al-Mashriq* site's statistical page showing the percentage of visitors to the site, categorized by operating system and by browser: The Windows operating system is in the first four positions – Windows XP at 47.4 percent, Windows 98 at 16.6 percent, Windows 2000 at 15.1 percent, Windows ME at 7.2 percent – comprising 86.3 percent of all operating systems accessing the site, followed by Mac in the 6th position at 2.5 percent. For browsers, the site reports Internet Explorer at 89 percent, with Netscape in third position at 3.4 percent, and Safari in 7th position at 0.5 percent (results obtained November 12, 2003). Although *Al-Mashriq* is neither an Arabic site nor located in the Arab world, the numbers do indicate a predominance of Windows and the Internet Explorer browser. The AWstats demo site enables one to compare the statistics of their visitors with those of *Al-Mashriq*. They are similar, and also show Windows and Internet Explorer to be dominant. No Arab host site running AWsats or a similar tool could be found. At any rate, until the time of this writing in late 2003, browsing the Internet using Windows components has been a perfunctory affair for several years; for the Mac, browsing Arabic sites has been plagued with problems, impeding efficient and productive work. Viewing Arabic sites under the Mac OS X, with the Safari browser, said to be the best available at this time, displays Arabic letters with varying sizes within the sentence.

Based on the number of users and the problems associated with the Mac, Windows is the operating system to use for the Internet in Arabic.

2. The Arabic Internet

2.1 *Searching*

Search engines such as Google, Vivisimo, and Copernic support searches with Arabic strings. To perform an Arabic search in any of these search engines, just type the search string in Arabic. It is also possible to customize search engines. For example, under Preferences or Language Tools, Google can be tailored to specific languages preferences, such as Arabic. Vivisimo has a useful feature, a clustering engine, which groups hits in the lefthand side pane of the browser, allowing one to select specific groups and display the hits in the righthand side pane. Copernic integrates with the Internet Explorer browser, providing its own search box on the browser's toolbar. Copernic comes in three versions: Basic, which is free, and Personal and Professional. The latter is able to track changes in the contents of Web pages one wishes to watch.

2.2 *Portals*

According to Berkman (2000), "A Web portal is a 'supersite' on the Internet that provides a comprehensive entry point for a huge array of resources and services. Portals typically contain news, free e-mail services, search engines, online shopping, chat rooms, discussion boards, and links to other sites". There are several Arabic portals, including *Bawwābat al-ʾAhrām, al-ʾImārāt li-l-ʾintarnit wa-l-waṣāʾil al-mutaʿaddida, Šabakat al-ʾintarnit li-l-ʾiʿlām al-ʿarabi*, and *Šabakat al-maʿlūmāt al-ʿarabiyya.*

2.3 *News media*

Commenting on the media in the Arab world, Lynch (2003) states:

> Whereas the broadcasting of the 1950s had been in the service of powerful states, the new media (both television and press) have self-consciously portrayed themselves as a mouthpiece for an Arab public deeply frustrated with all Arab regimes and beholden to none of them. Based primarily in London, the elite Arab press has been able to escape direct government control while drawing on writers and journalists from all over the world. Regular news roundups broadcast on the new satellite stations, along with the increasing availability of newspapers on the Internet for a small but growing younger following, have allowed this Arab press to reach a large audience. As a result, the staid and politically conservative national television stations have been rapidly losing market share and political significance. Yemeni President Ali Abdallah Salih, for example, once famously admitted that he watched the Qatar-based independent satellite network al Jazeera more regularly than he did official Yemeni TV.

Using a search string such as "Arabic newspapers" or "Arabic press" or "الجرائد العربية" (al-jaraʾid al-carabiyya) or "الصحافة العربية" (*aṣ-ṣaḥāfa al-ʿarabiyya*) will provide several hits, such as *Daily Sahafa* of Sahafa.com and several others containing links to Arabic newspapers indexed by country name. Viewing newspapers by means of an intermediate redirecting site – such as Sahafa.com – is slow, particularly if the redirecting site has bandwidth problems. Therefore, a direct link is always better. Yet, it is a useful strategy to use indirect sites to determine names of newspapers and then to use the search engines to find the URLs of these publications.

There are online versions of newspapers published in the Arab world as well in non-Arab countries. Internet news organizations are a rich source of editorial and op-ed opinions about local or non-local issues – such as the attacks against the United States on the 11th of September, or the invasion of Iraq. Using these Internet resources it is possible to compare ideologies across different publications within one country, such as *al-ʾAhrām, al-ʾAxbār*, and *aš-Šaʿb* in Egypt; or between different Arab countries, such as Egyptian newspapers, the Lebanese *an-Nahār*, and the Saudi *Jarīdat aš-Šarq al-ʾAwsaṭ*; or between publications based in non-Arab countries, such as *al-Ḥayāt* and *al-Quds al-ʿarabī*; or between them all.

These and other media organizations are a rich source of Arabic news discourse and culture: short stories, poetry, religious discussions, history, analyses of current affairs, and so on. Editorials and op-ed type articles can be investigated for ideology, hegemony, discourse structures, and layout (see, e.g., Bell and Garret 1998). Not only are these opinions present in the daily issues, but also in the weekend editions, such as *ʾAxbār al-Yawm*, which is published on Saturdays, and periodicals, such as *Markaz al-ʾAhrām li-d-Dirāsāt al-ʾIstrātijiyya*, which has analyses on issues such as the attacks on New York on September 11, 2001. One of the great features of *al-ʾAxbār* is their search facility, which provides archived editions, free of charge.

2.4 *Multimedia organizations*

Multimedia sites require programs for playing audio and video files. The most common are Windows Media Player, which comes with the Windows operating system, and Real Player and Apple's Quick Time Player, both of which can be downloaded from the Internet. BBCArabic.com uses Real Player, and *al-Jazīra* uses Windows Media Player.

Like newspapers and newsmagazines, multimedia sites such as *al-Jazīra* and BBCArabic.com also provide news. Unlike newspaper sites, these multimedia sites continually update their news content, and are a good place for breaking news.

These organizations and others provide discussion forums. BBCArabic.com hosts a section entitled *Šārik bi-raʾyik* 'Share your opinion', which is a forum for visitors to write their views on various political, religious, and cultural issues. BBCArabic.com also has an excellent audio page, *Istamiʿ ʾilā ʾiḏāʿat Bi Bi Si al-ʿarabiyya* 'listen to the BBC Arabic broadcast', which is accessible from a link on their site. At this time, the site features 19 programs, including such classics as *as-Siyāsa bayna s-sāʾil wa-l-mujīb* 'Politics Q&A', *ʿĀlam aḏ̣-ḏ̣āhira* 'The world at noon', and *Ḥiṣād as-sinīn* 'Harvest of the years'. BBC Arabic radio aficionados will recognize these programs and be happy to know they are accessible on the Internet as well.

Al-Jazīra also includes a discussion forum entitled *Muntadayāt al-Jazīra* '*Al-Jazeera* forums', which, like the BBC's *Šārik bi-raʾyik*, involves Arabs from varying backgrounds discussing political and cultural issues. *Al-Jazīra* is particularly distinguished because in addition to the latest news, its Internet site also hosts well-known, widely watched phone-in discussion panel programs that cover a wide array of topics including – to mention only 5 out of the 40 currently existing programs – news and opinions *ʾAkṯar min raʾy* 'more than one opinion', religion *aš-Šarīʿa wa-l-ḥayāt* 'life and the *Šarīʿa*', books *al-Kitāb xayru jalīs* 'a book is man's best companion', sports *Suʾāl fi r-riyāḍa* 'a sports question', and culture *al-Mašhad aṯ-ṯaqāfi* 'the cultural scene'. Not only does *Al-Jazīra* provide complete audio recordings of these programs, as well as sometimes providing short video clips, it also provides the complete text transcript. These resources are a great source of data for discourse-analytic research. For example, the panel discussions demonstrate mixing, overlap, and fusion of colloquial and Modern Standard Arabic in syntactic, morphological, lexical, and phonological terms. These panel discussions can also be investigated in conversation-analytic terms, as Clayman and Heritage do in *The news interview* (2002), focusing on features peculiar to the interview as an interactional forum, such as its turn-control system, management and progression by questions and answers, the art of questioning and interrogation, how public figures deal with journalists' questions, the conduct and strategies of the interviewees in order to serve their agendas, how they deal with each other at different points during the game, and other distinctive features.

2.5 *Chat forums*

In order to find chat sites a search engine, such as Vivisimo, may be used with a search string such as: منتدى (*muntadā* 'place of gathering or assembly') or حوار (*ḥiwār* 'dialogue') or حلقة (*ḥalqa* 'circle'). There are numerous such chat sites in various Arab countries. In contrast with opinion forums such as BBCArabic.com's *Šārik bi-raʾyik* or *al-Jazīra*'s *Muntadayāt al-Jazīra*, these *muntadayāt* are often multitiered and extensive, containing several top-level topic nodes, such as the 17 top nodes at the site Al-Muntada.com: *Muntadā al-ʾaxbār* (news), *al-Muntadayāt al-ʿāmma* (general), *al-Muntadayāt al-ʾislāmiyya* (Islam), *al-Muntadayāt as-siyāsiyya* (politics), *al-Muntadayāt al-ʾadabiyya* (literature), *Muntadayāt al-marʾa* (women), *al-Muntadayāt at-tarfīhiyya* (entertainment), *Muntadayāt al-kumbyūtar wa-l-ʾintarnit* (computers and Internet), *Muntadayāt ʾajhizat al-ittiṣālāt wa-l-ʾajhiza aš-šaxṣiyya* (personal digital assistants), *Muntadā l-hiwāyāt* (hobbies), *al-Muntadayāt at-taʿlīmiyya* (education), *al-Muntadayāt ar-riyāḍiyya* (sports), *Muntadayāt Sinama* (cinema), *Muntadayāt mawqiʿ al-ʾalʿāb al-ʿarabī* (gaming), *Muntadayāt as-siyāḥa wa-s-safar* (travel), *Muntadayāt as-sayyārāt* (automotive), and *Muntadā l-iqtirāḥāt wa-š-šakāwā* (suggestions and complaints). Each one of these top-tier topic mother nodes can comprise numerous daughter nodes, with each daughter node having several daughters of its own. For example, at the time of this writing, *al-Muntadayāt al-ʾadabiyya* includes two topics: *Muntadā š-šiʿr*, which shows 194 subtopics with 1,045 replies, and *Muntadā r-riwāya wa-l-qiṣṣa al-qaṣīra*, which has 86 topics and 491 replies.

The largest topic appears to be *Muntadayāt mawqiʿ al-ʾalʿāb al-ʿarabī*, with 107,873 topics and 1,081,707 replies. *Muntadayāt al-kumbyūtar wa-l-ʾintarnit* has the greatest number of daughters under the heading: 8 daughters, some of which are mothers themselves, with thousands of topics and responses.

The site *al-Qiṣṣa al-ʿarabiyya* comprises two sister nodes: a literary chat forum (*muntadā*), and a site (*mawqiʿ*) where authors from 22 Arab countries post short stories. Members discuss and criticize these stories, and post inquiries about them to the authors.

In the Holy Koran.net site, clicking any word performs a search and provides a list of the locations – *sūra* (chapter) and *ʾāya* (verse) – of that word in the *Qurʾān*, as well as a link to a choice of three *tafāsīr* ʻinterpretationsʼ. An alternative method of searching is also provided, permitting one to locate words or strings of one's own choice in the *Qurʾān* by typing the required term in *al-Munaqqib al-Qurʾānī* ʻthe Qurʾānic searcherʼ. It is possible to select the entire text of any *sūra*, or portions of it, and to paste it into a word processor or other application.

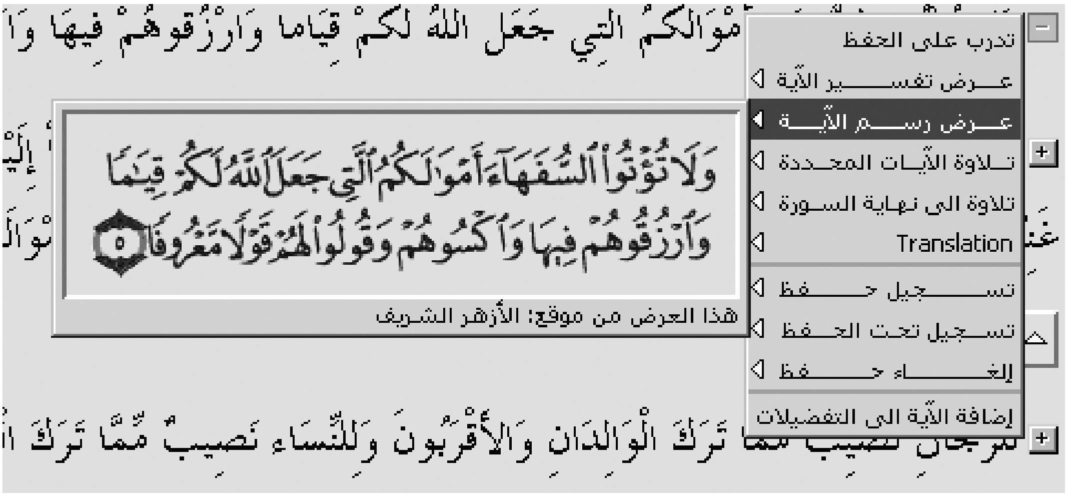
The Koran.muslim-web.com site has regular and calligraphic lettering, audio, a search facility, *tafsīr* ʻinterpretationʼ, translation, and a memorizing tool. For each *sūra*, there is a choice of recordings – from the most famous readers of the *Qurʾān* – to listen to. The audio is not locked, so it is also possible to record it by right-clicking the audio icon and selecting Save Target As. The memorizing tool plays one verse at a time and pauses for the learner to type it, using either an on-screen keyboard or the regular one. The result can be checked and kept in a database file that tracks the learner's progress. There is also a choice of translation to English, French, German, Dutch, or Spanish. The calligraphic-style lettering displays the selected *ʾāya* in the attractive familiar traditional style. This view is displayed in another window (see Fig. 1). This calligraphy is a link to the English site of al-ʾAzhar University. The *tafsīr* offers the following four selections: Ibn Kaṯīr, al-Jalālān, aṭ-Ṭabarī, and al-Qurṭubī. Each one of these four *tafāsīr* is a link to the *tafsīr* page of the al-ʾIslam.com site.

Al-ʾIslam.com is probably the largest site dedicated to Islam, and it is vast. The top-level Arabic page includes links to the following daughter node sites: *ʾArkān al-ʾIslām* ʻthe pillars of Islamʼ, *al-Ḥadīṯ* ʻthe sayings of the Prophetʼ, *al-Fiqh* ʻjurisprudenceʼ, *al-Qāmūs al-ʾIslāmī* ʻthe Islamic dictionaryʼ, *as-Sīra an-Nabawiyya* ʻthe tradition of the Prophetʼ, *at-Tārīx al-ʾIslāmī* ʻIslamic historyʼ, including the *Qurʾān* as well as a few others. Some of these daughter sites are large. For example the *as-Sīra* section includes excerpts from the following classics of biographical works: *Sīrat Ibn Hišām, ar-Rawḍ al-ʾAnaf, Muxtaṣar as-Sīra, Zād al-Maʿād*, and *al-Maġāzī*. The history section also includes several classics, and the dictionary section includes the following dictionaries: Arabic to English, Malay, or Indonesian; and English, Malay, or Indonesian to Arabic.

3. Capturing data for processing and exchange

It may be necessary to capture a document displayed in the browser and save it to the hard drive before it is deleted from its location on the server, or archived. The displayed document can be directly saved from the browser to HTML or text format files. Alternatively, the displayed document or portions of it can be

Figure 1. Calligraphic-style lettering in the site Koran.muslim-web.com



ولا تؤتوا السفهاء أموالكم التي جعل الله لكم قيما وارزقوهم فيها واكسوهم وقولوا لهم قولا معروفا ٥

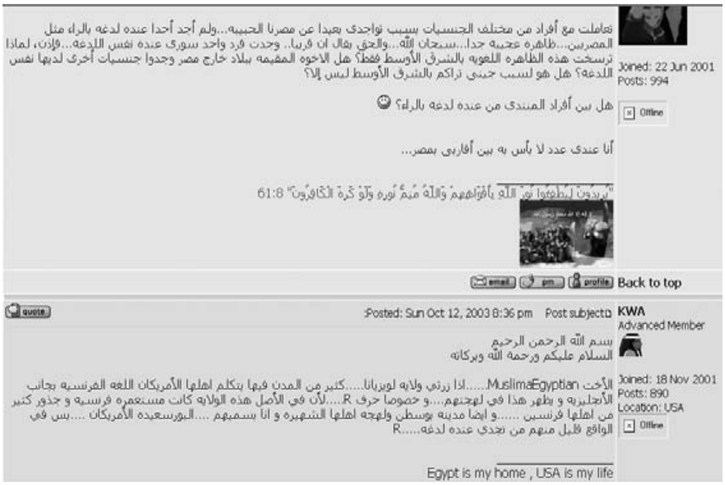
هذا العرض من موقع: الأزهر الشريف

pasted into a word processor. Whether the captured document should look exactly as it appears depends on the kind of analysis required. If, for example, the analysis requires examining layout – see Kress and van Leeuwen (1998:186–216) – the document should be saved directly from the browser as HTML. On the other hand, if the analysis does not call for displaying the document layout, the displayed document should be saved directly from the browser as "text" format. An alternative method is to select the entire displayed document, or parts of it, and paste the selection into Word, then to save the Word document in the required format. Before saving large amounts of data, in one or several files, one could experiment by saving a document displayed in the browser, or parts of it, to different file formats. File sizes, and appearances, may be compared in order to verify that the final format choice meets the required goals. Text files require the smallest storage overhead, and they are suitable for corpus linguistics analysis of different genres, such as fiction vs. academic texts (Biber a.o. 1998). On the other hand, text format documents are neither visually pleasing nor able to be suitably formatted for presentations. Thus, while Word documents impose more storage requirements, they allow more control over document formatting and processing. If document size becomes a problem, it is possible to select only the required portion from the browser display, for instance when doing Critical Discourse Analysis as ideology analysis (van Dijk 1995, 1998), where it is possible to select only the required part of the displayed document and to paste it into Word. If necessary, the document can then be reformatted. For example, editorials are often displayed in columnar or tabular format in HTML documents. Tables and columns may not serve a useful purpose in a word processor, and they can impose impediments to processing that document. Therefore, once pasted into Word, it is possible to use the Convert Table to Text function of Word, under Table, to get rid of unwanted tables. Similarly, in doing conversation analysis, for example to investigate the applicability of structural-functional categories of chat dialogue, including features such as topic management, pre-closing sections (Levinson 1983), or discourse markers (Schiffrin 1988), with large amounts of chat room text, it is possible to select the desired thread from a chat room, paste it into Word, and finally remove buttons, separators, icons, and any other superfluous objects (see Figs. 2, 3, 4). For issues regarding what to capture and what to omit, see for example Preston (1985), and Have (1999).

4. CONCLUSION

Having a set of tools configured for Arabic makes it possible to browse Arabic sites, store and retrieve documents, do e-mail, participate in chat forums, exchange attachments, and buy books. The right set of tools configured for Arabic increases productivity and enables research, networking, and collaboration in a community of experts.

Figure 2. Chat room data, original layout (Source: http://www.montada.com)



تعاملت مع أفراد من مختلف الجنسيات بسبب تواجدى بعيدا عن مصرنا الحبيبه...ولم أجد أحدا عنده لدغه بالراء مثل
المصريين...ظاهره عجيبه جدا...سبحان الله...والحق يقال ان قريبا.. وجدت فرد واحد سورى عنده نفس اللدغه...فإذن، لماذا
ترسخت هذه الظاهره اللغويه بالشرق الأوسط فقط؟ هل الاخوه المقيمه ببلاد خارج مصر وجدوا جنسيات أخرى لديها نفس
اللدغه؟ هل هو لسبب جينى تراكم بالشرق الأوسط ليس إلا؟

هل بين أفراد المنتدى من عنده لدغه بالراء؟

أنا عندى عدد لا بأس به بين أقاربى بمصر...

"يُرِيدُونَ لِيُطْفِئُوا نُورَ اللَّهِ بِأَفْوَاهِهِمْ وَاللَّهُ مُتِمُّ نُورِهِ وَلَوْ كَرِهَ الْكَافِرُونَ" 61:8

Joined: 22 Jun 2001
Posts: 994

Posted: Sun Oct 12, 2003 8:36 pm Post subject:

KWA
Advanced Member

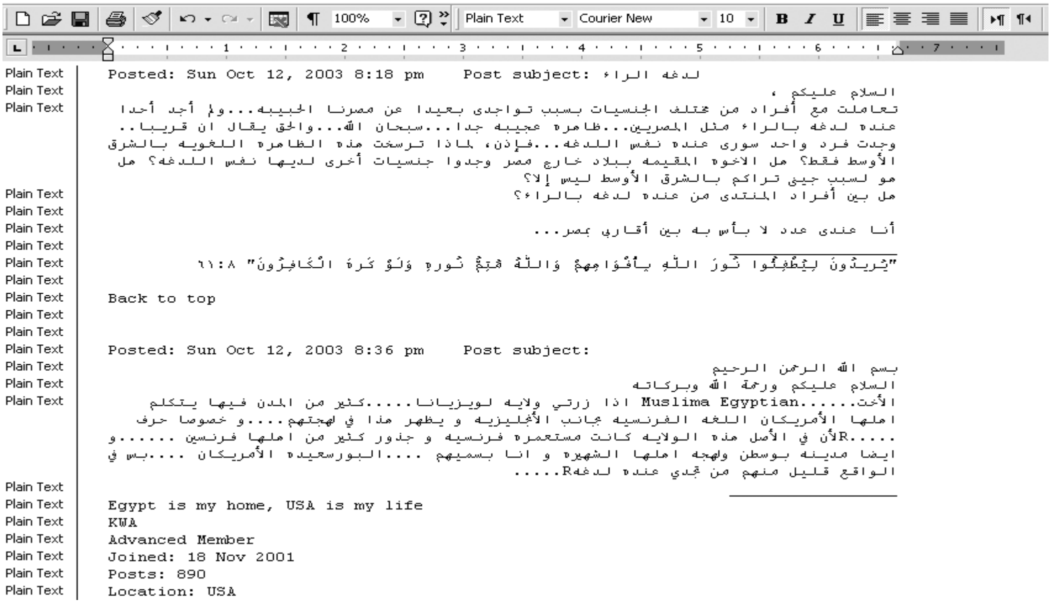
بسم الله الرحمن الرحيم
السلام عليكم ورحمة الله وبركاته

الأخت MuslimaEgyptian......اذا زرتي ولايه لويزيانا.....كثير من المدن فيها يتكلم اهلها الأمريكان اللغه الفرنسيه بجانب
الأنجليزيه و يظهر هذا في لهجتهم....و خصوصا حرف R.....لأن في الأصل هذه الولايه كانت مستعمره فرنسيه و جذور كثير
من اهلها فرنسيين ......و ايضا مدينه بوسطن ولهجه اهلها الشهيره و انا بسميهم ....البورسعيده الأمريكان ....بس في
الواقع قليل منهم من تجدي عنده لدغه.....R

Egypt is my home , USA is my life

Joined: 18 Nov 2001
Posts: 890
Location: USA

Figure 3. Chat room data, text format (Source: http://www.montada.com)



Posted: Sun Oct 12, 2003 8:18 pm Post subject: لدغه الراء

السلام عليكم ،
تعاملت مع أفراد من مختلف الجنسيات بسبب تواجدى بعيدا عن مصرنا الحبيبه...ولم أجد أحدا
عنده لدغه بالراء مثل المصريين..ظاهره عجيبه جدا...سبحان الله...والحق يقال ان قريبا..
وجدت فرد واحد سورى عنده نفس اللدغه..فإذن، لماذا ترسخت هذه الظاهره اللغويه بالشرق
الأوسط فقط؟ هل الاخوه المقيمه ببلاد خارج مصر وجدوا جنسيات أخرى لديها نفس اللدغه؟ هل
هو لسبب جينى تراكم بالشرق الأوسط ليس إلا؟
هل بين أفراد المنتدى من عنده لدغه بالراء؟

أنا عندى عدد لا بأس به بين أقاربي بمصر...

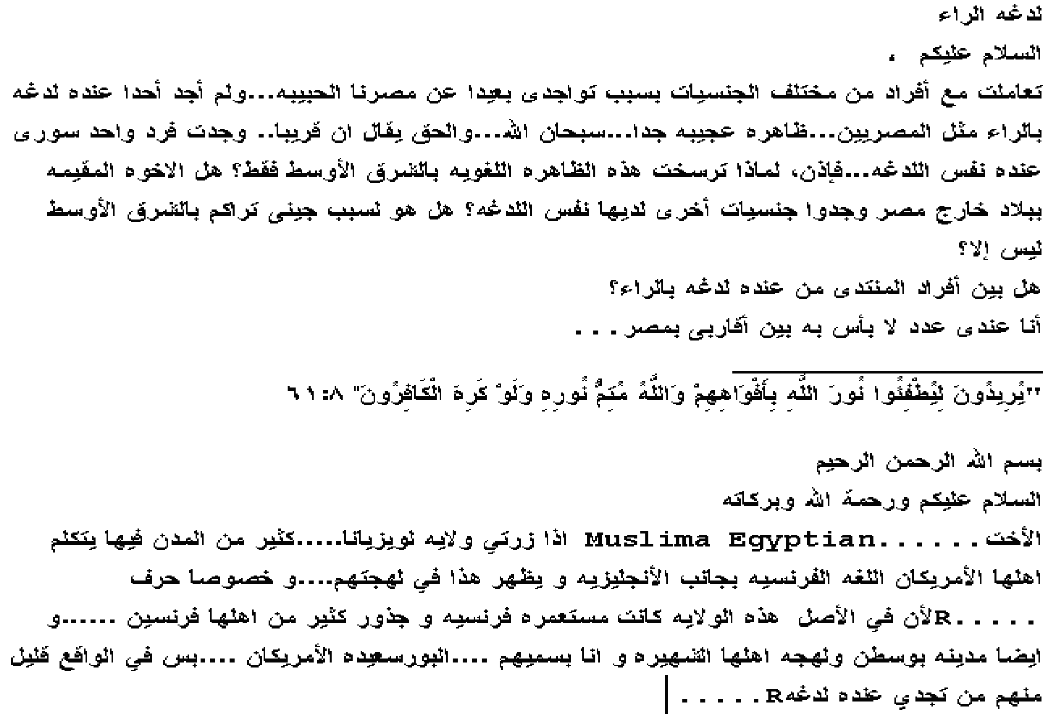
"يُرِيدُونَ لِيُطْفِئُوا نُورَ اللَّهِ بِأَفْوَاهِهِمْ وَاللَّهُ مُتِمُّ نُورِهِ وَلَوْ كَرِهَ الْكَافِرُونَ" ٦١:٨

Back to top

Posted: Sun Oct 12, 2003 8:36 pm Post subject:
بسم الله الرحمن الرحيم
السلام عليكم ورحمة الله وبركاته
الأخت.......Muslima Egyptian اذا زرتي ولايه لويزيانا.....كثير من المدن فيها يتكلم
اهلها الأمريكان اللغه الفرنسيه بجانب الأنجليزيه و يظهر هذا في لهجتهم....و خصوصا حرف
R.....لأن في الأصل هذه الولايه كانت مستعمره فرنسيه و جذور كثير من اهلها فرنسين ......و
ايضا مدينه بوسطن ولهجه اهلها الشهيره و انا بسميهم ....البورسعيده الأمريكان ....بس في
الواقع قليل منهم من تجدي عنده لدغهR.....

Egypt is my home, USA is my life
KWA
Advanced Member
Joined: 18 Nov 2001
Posts: 890
Location: USA

Figure 4. Chat room data, reformatted, more legible text (Source: http:www.montada.com)



لدغه الراء
السلام عليكم ،
تعاملت مع أفراد من مختلف الجنسيات بسبب تواجدى بعيدا عن مصرنا الحبيبه..ولم أجد أحدا عنده لدغه
بالراء مثل المصريين...ظاهره عجيبه جدا...سبحان الله...والحق يقال ان قريبا.. وجدت فرد واحد سورى
عنده نفس اللدغه...فإذن، لماذا ترسخت هذه الظاهره اللغويه بالشرق الأوسط فقط؟ هل الاخوه المقيمه
ببلاد خارج مصر وجدوا جنسيات أخرى لديها نفس اللدغه؟ هل هو لسبب جينى تراكم بالشرق الأوسط
ليس إلا؟
هل بين أفراد المنتدى من عنده لدغه بالراء؟
أنا عندى عدد لا بأس به بين أقاربى بمصر . . .

"يُرِيدُونَ لِيُطْفِئُوا نُورَ اللَّهِ بِأَفْوَاهِهِمْ وَاللَّهُ مُتِمُّ نُورِهِ وَلَوْ كَرِهَ الْكَافِرُونَ" ٦١:٨

بسم الله الرحمن الرحيم
السلام عليكم ورحمة الله وبركاته
الأخت Muslima Egyptian...... اذا زرتي ولايه لويزيانا....كثير من المدن فيها يتكلم
اهلها الأمريكان اللغه الفرنسيه بجانب الأنجليزيه و يظهر هذا في لهجتهم....و خصوصا حرف
R.....لأن في الأصل هذه الولايه كانت مستعمره فرنسيه و جذور كثير من اهلها فرنسين ......و
ايضا مدينه بوسطن ولهجه اهلها الشهيره و انا بسميهم ....البورسعيده الأمريكان ....بس في الواقع قليل
منهم من تجدي عنده لدغهR.....

BIBLIOGRAPHICAL REFERENCES

Bell, Allan and Peter Garrett (eds.). 1998. *Approaches to media discourse.* Oxford: Blackwell.

Berkman, Eric. 2000. "What is a portal?" *Darwin.* http://www.darwinmag.com/learn/curve/column.html?ArticleID=13 (accessed November 5, 2003).

Biber, D., S. Conrad, and R. Reppen, 1998. *Corpus linguistics: Investigating language structure and use.* Cambridge: Cambridge University Press.

Clayman, Steven and John Heritage. 2002. *The news interview: Journalists and public figures on the air.* Cambridge: Cambridge University Press.

Dijk, Teun van. 1995. "Discourse analysis as ideology analysis". *Language and peace*, ed. Christina Schäffner and Anita L. Wenden. Brookfield, Vt.: Dartmouth Publishing Company.

——. 1998. "Opinions and ideologies in the press." Bell and Garrett (1998:21–63).

Freesoft.org. *Connected: An Internet encyclopedia.* http://www.freesoft.org/CIE/index.htm (accessed October 30, 2003).

Have, Paul ten. 1999. "Transcribing talk-in-interaction". *Doing conversation analysis: Practical guide* 47–98. London: Sage.

Kress, Gunther and Theo van Leeuwen. 1998. "Front pages: (The critical) analysis of newspaper layout". Bell and Garrett (1998:186–219).

Levinson, Stephen C. 1983. "Conversational structure". *Pragmatics*, Chap. 6, 284–370. Cambridge: Cambridge University Press.

Lynch, Mark. 2003. "Taking the Arabs seriously". *Foreign Affairs.* http://www.foreignaffairs.org:80/20030901faessay82506/marc-lunch/taking-arabs-seriously.html (accessed September 9, 2003).

Preston, D.R. 1985. "The Li'l Abner syndrome: Written representations of speech". *American Speech* 60:4.328.

Schiffrin, Deborah. 1988. *Discourse markers.* Cambridge: Cambridge University Press.

LINKS

The following links were all live at the time of writing.

AWSTATS

The AWstats official Web site: http://awstats.sourceforge.net/

AL-MASHRIQ

al-Mashriq: http://almashriq.hiof.no
AWstats page: http://almashriq.hiof.no/cgi-bin/awstats.pl?config=almashriq.hiof.no

LINUX

http://www.arabeyes.org/
http://www.langbox.com/arabic/
http://www.linux-egypt.org/
http://www.linuxjournal.com/index.php
http://www.linux4arab.com/

ENCODING

ISO 8859–6 (Arabic): http://www.microsoft.com/globaldev/reference/iso/28596.htm
Windows Arabic: http://www.microsoft.com/globaldev/reference/sbcs/1256.htm
Mac Arabic: http://www.unicode.org/Public/MAPPINGS/VENDORS/APPLE/ARABIC.TXT

Transliteration fonts

Font Creator: http://www.high-logic.com/fcp.html
SIL International: http://www.sil.org/computing/index.html

Search engines

Vivisimo: http://vivisimo.com/
Copernic: http://www.copernic.com/en/index.html
Google: http://www.google.com/

Portals

Bawwābāt al-ʾAhrām al-ʾiliktrūniyya: http://portal.ahram.org.eg
al-ʾImārāt li-l-ʾIntarnit wa-l-Wasāʾiṭ al-Mutaʿaddida: http://www.albahhar.com/
Šabakat al-ʾintarnit li-l-ʾIʿlām al-ʿarabī: http://www.amin.org/
Šabakat al-Maʿlūmāt al-ʿarabiyya: http://www.moheet.net/

News media

Daily Sahafa: http://www.sahafa.com/daily.asp
al-ʾAhrām: http://www.ahram.org.eg/
al-ʾAxbār: http://www.elakhbar.org.eg/index2.html
aš-Šaʿb: http://www.alshaab.com/
an-Nahār: http://www.annaharonline.com/
Jarīdat aš-Šarq al-ʾAwsaṭ: http://www.asharqalawsat.com/
al-Ḥayāt: http://www.daralhayat.com/
al-Quds al-ʿarabī: http://www.alquds.co.uk/
ʾAxbār al-Yawm: http://www.akhbarelyom.org.eg/akhbarelyom/
ʾAxbār al-ʾAdab: http://www.akhbarelyom.org.eg/adab/
Markaz ad-Dirāsāt as-Siyāsiyya wa-l-ʾIstratījiyya: http://www.ahram.org.eg/ACPSS/
Barīd al-ʿArab: http://www.arabmail.de/

Multimedia

Audio and video players:
Real: http://www.real.com/
Quick Time: http://www.apple.com/quicktime/
Sites:
BBCArabic.com: http://news.bbc.co.uk/hi/arabic/news/
BBCArabic.com audio: http://www.bbc.co.uk/arabic/radio/aod/arabic promo.shtml
Al-Jazira: http://www.aljazeera.net/index.htm

Chat

Al-Muntadā: http://www.montada.com/
Muntadā l-Miṣriyyīn: http://www.egyptiantalks.org/bforums/
al-Qiṣṣa al-ʿarabiyya: literary forum: http://www.arabicstory.net/forum/index.php

Literature/culture

al-Qiṣṣa al-ʿarabiyya: short stories: http://www.arabicstory.net/index.php
ʾAxbār al-ʾadab: http://www.akhbarelyom.org.eg/adab
al-Ḥayāt: http://www.daralhayat.com/culture/

Heritage

The *Qurʾān*: http://quran.muslim-web.com/index.htm
The *Qurʾān*: http://www.holyquran.net/quran/index.html
Al-Islam.com top-level Arabic entry point: http://quran.al-islam.com/arb/
Zaydān: http://www.ziedan.com/index_o.asp
al-Warrāq: http://www.alwaraq.cc/

Books online

an-Nīl wa-l-Furāt: http://www.neelwafurat.com/
Kawkab: http://en.kawkab.com/egi-bin/switch?lang=ar

Translation

Al-Misbār (requires a subscription): http://www.almisbar.com

Dictionaries

ʿAjib: http://lexicons.sakhr.com/

Recipes

Dalīl aṭ-Ṭabx al-ʿarabī: http://acookweb.hawaaworld.com/index.php
Zurūna: http://216.203.153.240/pages/arabic/foodand-pastries/recipes/more2.html
Ḥiwārāt al-maṭbax: http://www.hwarat.com/index.php?s=3687424a302a6bb1b714339c46b8c542
Horus: http://www.horus.ics.org.eg/arabic/html/recipes.html

Waheed Samy
(University of Michigan)

Interrogative Pronoun

The interrogative pronouns in Classical Arabic are *man* 'who?' and *mā* 'what?', which are indeclinable (Wright 1964:I, 274–275; II, 311–315; Fleisch 1979:74–78); the interrogative adjective is *ʾayyun* 'which?', which has a feminine form *ʾayyatun* and is declined (Wright 1964: I, 275–276; II, 315–317; Fleisch 1979:78–81). In combination with prepositions, *mā* is often shortened (*bima*, *lima*, etc.); before verbs, it is usually strengthened by *ḏā: mā ḏā ṣanaʿta* 'what have you done?'.

According to Rabin (1951:189), the pronoun *man* was inflected in the pre-Islamic dialect of the Ḥijāz, with forms like *manū*, *manī*, *manā*, *manūna*, which may be related to interrogative pronouns in other Semitic languages, such as Akkadian *mīnu*. The Arabic grammarians give an entire paradigm of this form of the pronoun (cf. Fleisch 1979:II, 78; Wright 1964:I, 275). According to az-Zamaxšarī (*Mufaṣṣal* 59), the ending of inflected *man* reflects the ending of the questioned word: in reaction to *raʾaytu zaydā* 'I have seen Zayd', one might say *manā* 'whom?', and in reaction to *jāʾanī rajulun* 'a man came to me', one might say *manū* 'who?'. In the dual and the plural, these forms are *manān*, *manūn*/*manāt* (the variant *manūna ʾantum* 'who are you?' in a poetic line is rejected by az-Zamaxšarī). If this information is correct, these forms are not connected with the phrasal interrogatives (< *man huwa*, *man hiya* in modern Arabic dialects; see below).

In the Arabic linguistic tradition, *man* and *mā* are regarded as nouns with a special function, that of *istifhām* 'questioning, interrogation'. All grammarians devoted much attention to the syntax of the interrogative pronouns, especially the interrogative adjective *ʾayyun*, but in later grammar there is also considerable interest in the pragmatic aspects of interrogation. Ibn Hišām (d. 761/1359), for instance, investigates in detail the various pragmatic categories of yes/no questions in his *Muġnī l-labīb* (see Gully 1995:138–140, 145–146, 189–191). With respect to the interrogative pronouns, he observes that they indicate the meaning of *istifhām* in something else, a property they share with the particles (Gully 1995:127–128). In the framework of the Arabic linguistic tradition, this explains why they are indeclinable.

In the modern dialects, the pronoun for 'who?' goes back to four different forms (Singer 1958:93ff.): *man* (e.g. Gulf Arabic *man* ~ *min*), *mīn* (e.g. Egyptian Arabic *mīn*), *ʾayyu šayʾin huwa* (e.g. Daṯīna *šūḏī*), and *ʾayyu šayʾin yakūnu* (e.g. Moroccan Arabic *škun*). The first two are regarded by Singer as original doublets (1958:129–130, 137); the other two are phrasal interrogatives. In some dialects, the forms *man* or *mīn* also develop into phrasal interrogatives, e.g. in Gulf Arabic, where *minhu* occurs along with *min*, and in Šukriyya *minū* (< *man huwa*; Reichmuth 1983:116–117). The origin of Chadian Arabic *yātu* is not completely clear; according to Roth-Laly (1979:170–171) it goes back to *ʾayyu* + *t* + *hu*; a feminine form *yāti* is also attested.

For the pronoun 'what?', almost all modern dialects have a form containing the word *šayʾ* 'thing'. Reflexes of Classical Arabic *mā* are found only in Yemeni Arabic (Behnstedt 1985, Map 59), where *mā* is found alone or in combination with a personal pronoun: *mā huwa* > *maw*, *mō*, *mū* and *mā hiya* > *mī*; in some areas *mā huwa* is combined with *-ši*, e.g. Ristāq *mhūšši* (Singer 1958:173).

In the other dialects, the neutral interrogative pronoun has developed from a combination *ʾayyu šayʾin* 'which thing?'. Even in Classical Arabic this form is not unknown, possibly as an emphatic variant of the more usual *mā*, but mostly as the result of interference from the vernacular language in Middle Arabic texts. In the modern dialects, it has become grammaticalized as the only way to express 'what?'. There are two types, distinguished by the presence or absence of the ending *-in*. Forms with this ending are, for instance, Iraqi Arabic *šinu*, Tunisian Arabic *āšnūa*, Moroccan Arabic *āšnhuwa*, Andalusian Arabic *ašan*, *aššan*. According to Singer (1958:209), they represent an older type, while the forms without *-in*, for instance Egyptian Arabic *ʾēh*, Sinai Arabic *ʾēš*, Syrian Arabic *šū*, Moroccan Arabic *ʾāš*, Maltese *xi*, are the result of a later development. In many dialects, phrasal forms occur along with nonphrasal forms, e.g. in Gulf Arabic *šinhu* along with *wēš*; other dialects have only phrasal forms (e.g. Ḥassāniyya *šənhu*, *šənhi*). In some dialects, for instance in Aden Arabic, the interrogative pronoun begins with *w-*, *weš*, *waš*, *wuš*. These forms are explained by Singer (1958:211) as the result of an alternation *w-~ʾ*, rather than a combination with the conjunction *wa-*. The Moroccan Arabic question particle *wāš* (→ interrogative sentences), on the other hand, may have developed out of such a combination.

For the interrogative adjective, the dialects use various forms that can be divided into forms with and without an element *-n-*, often in combination with a suffixed pronoun. In Egyptian Arabic, for instance, Woidich (2006:35) mentions *ʾayy* as a borrowing from Standard Arabic. The original forms were *ʾanhu/ʾanhi* before the questioned noun, or *anhū/anhī/ anhūm* after the questioned noun. An alternative form is *āni*, which is not inflected at all. Forms with *-n-* also occur in Syrian Arabic (*ʾanu/ʾani/ʾanon* along with *ʾayy*; Cowell 1964:573) and Moroccan Arabic (*īna/āna* or *fīna/fāna*; Caubet 1993:I, 172). According to Fischer and Jastrow (1980:86), they go back to Aramaic *aynā*. Forms without *-n-* occur in Gulf Arabic (*ayhu/ayhi/ayhum* or *yahu(m)/ yahi/yahum* as independent forms, and *ay* as adjective; Holes 1990:177–178) and in Iraqi Arabic (*yāhu*; Erwin 1963:295).

Forms like *ʾayyu šayʾin* could be called bimorphemic: they contain a question word and a questioned word (Versteegh 2006). Through the question word they are connected paradigmatically – at least partly in some dialects – with interrogative adverbs, e.g. Iraqi Arabic *šwakit* 'when?' (< *ʾayy šayʾ waqt*), *šlōn* 'how?' (< *ʾayy šayʾ lawn*), *šgad* 'how much?' (< *ʾayy šayʾ qadr*). There are some indications that the bimorphemic forms represent an older type. In Egypt and Syria, for instance, the usual form for ‘when?’ is a reflex of Classical Arabic *matā* or **ʾayy matā*, but some peripheral areas use forms going back to *ʾayyu (šayʾ) waqtin* ‘which time?’, e.g. *waktē, waxtē* in the oasis of Kharga and in the eastern Delta (Behnstedt and Woidich 1985, Map 185), or *šwakit*, *šokt* in the Qamişli area in northeast Syria (Behnstedt 1997, Map 290). Such peripheral forms are probably traces of an earlier language type in which the interrogative words were more transparent (Muysken and Smith 1990; on the development of bimorphemic interrogatives, see Bruyn 1991).

For the syntax of interrogative sentences in modern dialects, see → interrogative sentence.

Bibliographical references

Primary source

Zamaxšarī, *Mufaṣṣal* = ʾAbū l-Qāsim Maḥmūd ibn ʿUmar az-Zamaxšarī, *Kitāb al-mufaṣṣal fī n-naḥw*. Ed. Jens Peter Broch. Christiania: Libraria P.T. Mallingii, 1870.

Secondary sources

Behnstedt, Peter. 1985. *Die nord-jemenitischen Dialekte*. I. *Atlas*. Wiesbaden: L. Reichert.

——. 1997. *Sprachatlas von Syrien Kartenband + Beiheft*. Wiesbaden: O. Harrassowitz.

—— and Manfred Woidich. 1985. *Die ägyptisch-arabischen Dialekte*. I. *Einleitung und Anmerkungen zu den Karten*. II. *Dialektatlas von Ägypten*. Wiesbaden: L. Reichert.

Bruyn, Adrienne. 1991. “Question words in 18th-century and 20th-century Sranan”. *Historical linguistics 1991: Papers from the 10th International Conference on Historical Linguistics, Amsterdam, 12–16 August 1991*, ed. Jaap van Marle, 31–47. Amsterdam and Philadelphia: J. Benjamins.

Caubet, Dominique. 1993. *L'arabe marocain*. 2 vols. Paris and Louvain: Peeters.

Cowell, Mark W. 1964. *A reference grammar of Syrian Arabic*. Washington, D.C.: Georgetown University Press.

Erwin, Wallace M. 1963. *A short reference grammar of Iraqi Arabic*. Washington, D.C.: Georgetown University Press.

Fischer, Wolfdietrich and Otto Jastrow. 1980. *Handbuch der arabischen Dialekte*. Wiesbaden: O. Harrassowitz.

Fleisch, Henri. 1979. *Traité de philologie arabe*. II. *Pronoms, morphologie verbale, particules*. Beirut: Dar El-Machreq.

Gully, Adrian. 1995. *Grammar and semantics in medieval Arabic: A study of Ibn-Hisham's* ‘Mughni l-Labib’. London: Curzon Press.

Holes, Clive. 1990. *Gulf Arabic*. London and New York: Routledge.

Muysken, Pieter and Norval Smith. 1990. “Question words in pidgin and creole languages”. *Linguistics* 28.883–903.

Rabin, Chaim. 1951. *Ancient West-Arabian*. London: Taylor's Press.

Reichmuth, Stefan. 1983. *Der arabische Dialekt der Šukriyya im Ostsudan*. Hildesheim: G. Olms.

Roth-Laly, Arlette. 1979. *Esquisse grammaticale du parler arabe d'Abbéché (Tchad)*. Paris: P. Geuthner.

Singer, Hans-Rudolph. 1958. *Neuarabische Fragewörter: Ein Beitrag zur historischen und vergleichenden Grammatik der arabischen Dialekte*. Ph.D. diss., University of Erlangen.

Versteegh, Kees. 2006. “What's where and how's what? Interrogatives in Arabic dialects”. *Estudios de Dialectología Norteafricana y Andalusí* 8.000.

Woidich, Manfred. 2006. *Grammatik des Kairenisch-Arabischen*. Wiesbaden: O. Harrassowitz.

Wright, William. 1964. *A grammar of the Arabic language*. 3rd ed. rev. W. Robertson Smith and M.J. de Goeje. Repr. 2 vols. Cambridge: Cambridge University Press.

Kees Versteegh (University of Nijmegen)

Interrogative Sentences

Interrogative sentences in Classical Arabic are divided into yes/no questions, and questions with an → interrogative pronoun or adverb (Reckendorf 1921:31–41; Wright 1964:II, 306–317). Yes/no questions are usually introduced by a question particle *ʾa-* or *hal*. In Classical Arabic, *ʾa-* is the general interrogative particle, but it may be used to express astonishment or indignation: *ʾa-jāʾa zayd* ‘Zayd didn't come, did he?’. With a negation, a positive answer is often anticipated, e.g. *ʾa-lā tarā ʾanna*… ‘don't you see that…; surely, you can see that…’; *ʾa-lā* or *ʾallā* may also express a request, e.g. *ʾa-lā tuqātilūna qawman nakaṯū ʾaymānahum* ‘won't you fight a people who have broken their oaths?’, in the sense of ‘fight those people!’ (Q. 9/13, quoted by Wright 1964:II, 311). The particle *hal* is often used when a positive answer is anticipated, e.g. *hal ʾaqsamtum kulla muqsamin* ‘didn't you swear every oath?’ (Zuhayr, *Muʿallaqa* 26, quoted by Reckendorf 1921:36), *wa-hal ʾatāka ḥadīṯu mūsā* ‘haven't you heard the story of Moses?’ (Q. 20/9, quoted by Wright 1964:II, 308). With the negation *lā* it is used in the sense of ‘why not?’, e.g., *fa-hal-lā taškuru lī ʾiḏ qataltu zuhayran* ‘why aren't you grateful to me that I killed Zuhayr?’ (*ʾAġānī* X, 17.15, quoted by Reckendorf 1921:36). The structure of interrogative sentences in Modern Standard Arabic is essentially identical with that of Classical Arabic (Badawi a.o. 2004:685–710), except for the fact that the distinction between the two question particles has become blurred.

Interrogative sentences in modern Arabic dialects are dealt with at length in most reference grammars, e.g. Libyan Arabic (Owens 1984:101–103); Egyptian Arabic (Woidich 2006, Syntax IV); Iraqi Arabic (Erwin 1963:292–295); Syrian Arabic (Cowell 1964:566–577); Gulf Arabic (Holes 1990:2–16); Yemeni Arabic (Watson 1993:127–128, 292–294, 404–405, 408–410). The present entry gives a detailed description of the major types of direct and indirect questions in one dialect, Moroccan Arabic. Like most Arabic dialects, this dialect has → WH-movement in questions containing an → interrogative pronoun or adverb (Egyptian Arabic is an exception, because it has *in situ* position of the interrogative). It differs, however, from most other dialects in that it has a special question particle in yes/no questions, *waš*.

Moroccan Arabic has two basic word orders: an unmarked Verb-Subject-Object order used in any context, and a marked Subject-Verb-Object order used only under specific discourse conditions. Furthermore, Moroccan Arabic is not a case language, i.e., there is, for instance, only one single form for all noun phrases (NPs), as in (1) and (2).

(1) *šaf ʿali l-wəld*
saw Ali the-boy
'The boy saw Ali'

(2) *xrəž l-wəld*
went.out the-boy
'The boy went out'

As the examples above illustrate, verbs in Moroccan Arabic are highly inflected. For instance, all the verb forms are complex and tensed, unlike English; they inflect for number, gender, and person of the subject and carry tensed endings (e.g. in the past, future tense, etc.).

On the other hand, prepositions cannot be stranded in Moroccan Arabic. Thus, whereas (3a) is grammatical, (3b) is ill formed because Moroccan Arabic prepositions cannot be left without an overt complement.

(3a) *ḥaṭṭ ʿali lə-ktab fuq ṭ-ṭabla*
put Ali the-book on the-table
'Ali put the book on the table'

(3b) **ṭ-ṭabla ḥaṭṭ ʿali lə-ktab fuq*

Like English, Moroccan Arabic exhibits 'Pied-Piping', i.e., the movement of the WH-phrase along with the preposition is obligatory in Moroccan Arabic because it cannot be stranded, as in (4a) and (4b).

(4a) *xəlla ʿali l-kas fuq aš?*
left Ali the-glass on what
'Ali left the glass on what?'

(4b) *fuq aš xəlla ʿali l-kas?*
on what left Ali the-glass
'On what did Ali leave the glass?'

Furthermore, there is no analogue in Moroccan Arabic of Subject-Auxiliary-Inversion in questions of any sort. By way of example, consider (5) and (6).

(5a) *ġadi marya tšuf ʾaḥmed*
will Maria see Ahmed
'Maria will see Ahmed'

(5b) *škun ġadi marya tšuf?*
who will Maria see
[Subject-Auxiliary-Inversion]
'Who will Maria see?'

(6a) *xəlla ʿali lə-ktab fuq ṭ-ṭabla*
put Ali the-book on the-table
'Ali put the book on the table'

(6b) *fin xəlla ʿali lə-ktab?*
where put Ali the-book
'Where did Ali leave the book?'

Note that questions in Moroccan Arabic have the same word order(s) as statements, e.g. Verb-Subject-Object in (5) and (6) above.

By contrast, indirect questions are characterized by the fact that the interrogative structure is a dependent clause and is at the same time the complement of verbs such as *sewwel* 'to ask', *xammem* 'to wonder', etc., as illustrated by the bold clause in (7).

(7) *suwlat-ni maryam* ***ʾimta ržəʿ***
asked-me Maria when came.back
yusif
Youssef
'Maria asked me when Youssef had come back'

Note the contrast between (7) and (8).

(8) *suwlat-ni maryam, ʾimta ržəʿ*
asked-me Maria when came back
yusif?
Youssef
'Maria asked me, "When did Youssef come back?'"

The direct WH-question in (8) is distinguished from the indirect one in (7) by intonation.

WH-words in Moroccan Arabic behave in more or less the same way as they do in English, except for the fact that Moroccan Arabic does not distinguish between WH-phrases referring to humans and those referring to nonhumans. In addition, WH-words in Moroccan Arabic are not phonetically homogeneous but behave functionally in the same manner as in English, as illustrated in Table 1.

By way of example, consider (9) – (11).

(9a) *aš ṭra?*
what happened
'What happened?'

(9b) *aš šəfti?*
what saw2s
'What did you see?'

(9c) *b-aš ḥalliti l-bab?*
with-what opened2s the-door
'With what did you open the door?'

(10a) *škun ža?*
who came3ms
'Who came?'

(10b) *škun šəfti?*
whom saw2s
'Whom did you see?'

(10c) **l-škun hḍərti?*
to-who spoke2s
'To whom did you speak?'

(11a) **-mən rḥəl?*
whom moved3ms
*'Whom moved?'

(11b) **-mən šəfti?*
whom saw2s
'Whom did you see?'

(11c) *l-mən hḍərti?*
to-whom spoke2s
'To whom did you speak?'

From these data, it follows that *škun* and *aš* are NP (pronouns) while *-mən* is an NP complement of a preposition, as in (11). The interrogative *-mən* is a bound morpheme derived from Classical Arabic *man* 'who?', which is used as an interrogative and as a relative pronoun. In Moroccan Arabic, *-mən* never occurs alone. It is a bound morpheme to which enclitic prepositions like *l-* 'to', *mʿa* 'with', etc. are usually attached when it is used as an interrogative, as in (11c), as well as a relative pronoun, e.g. *r-ražel l-mən hḍərti.*

Although *-mən* is restricted to only some case functions, this is not a counterexample to the fact that Moroccan Arabic is not a case language, because *-mən* is a clitic bound morpheme.

Table 1. Syntactic functions of WH-words in Moroccan Arabic

Interrogative PRO	Subject	Direct object	Prepositional object	Adverbial form
aš 'what?'	✓	✓	✓	*
škun 'who?'	✓	✓	*	*
-mən 'whom?'	*	*	✓	*
fin 'where?', *ʾimta* 'when?', *ʿlaš* 'why?', *kif* 'how?'	*	*	*	✓

✓ = acceptable; * = unacceptable

The pronoun *aš* is apparently used in five contexts:

(12) *r-ražel* *l-aš* *hḍərt*
the-man to-whom spoke2s
'The man to whom you spoke'

(13) *ʿlaš* *žiti-ni* *f* *had* *ḍ-ḍuruf?*
why came2s in these circumstances
'Why did you come to me under these circumstances?'

(14) *rma* *yusif* *l-fardi* *b-aš*
threw away Youssef the-gun with-which
qtal *marya*
killed Maria
'Youssef threw away the gun with which he killed Maria'

(15) *bġit* *yusif* *baš* *iqra*
want1sg. Youssef for study3ms
məzyan
well
'I would like for Youssef to study hard'

(16) *waš* *nta* *mažnun?*
whether you crazy
'Are you crazy?'

However, apart from the WH-phrases *b-aš* 'with what' (not identical with the complementizer in (15)) and *l-aš* 'to whom', which can be segmented into *b-*, *l-*, and *aš* (meaning synchronically and literally 'with what' and 'to what' respectively), it seems that WH-phrases like *ʿlaš* 'why?', *baš* 'for', and *waš* 'whether' cannot be divided into enclitic prepositions and the WH-element *aš*, and must be assumed to be simple forms rather than compounds. One could, for instance, divide *waš* into *w-aš* (meaning historically and literally 'and what') and *baš* into *b-aš* 'with what', but it is not clear that synchronically there is any relation between these enclitic prepositions (namely *w-*, *b-*) and *aš* since *waš* corresponds simply to 'whether' and *baš* to 'for, in order to'.

As in English, the interrogative pronouns *fin* 'where?', *ʾimta* 'when?', *ʿlaš* 'why?', and *kif* 'how?' are used as adverbial pronominal forms but never as subject, direct object, or prepositional object forms, because they are not WH-NP forms. Note that except for the pronominal quantifier *šḥal* 'how much/many?', which can be used in exclamations as well, and *-aš*, which can also appear in relative clauses, all these interrogative pronouns are used only in interrogative constructions.

In Moroccan Arabic, there are no interrogative determiners. For example, the WH-phrases *aš* and *šḥal* cannot be used as interrogative determiners:

(17) *šḥal* *mən* *ktab* *qriti?*
how.many of book read2s
'How many books did you read?'

(18) *aš* *mən* *ktab* *qriti?*
what of book read2s
'Which book did you read?'

The preposition *-mən* lit. 'of' goes along here with the head noun phrase *ktab*, not with the WH-words *aš* or *šḥal*, unlike the WH-phrase 'whom' in the configuration in Table 1. The use of *-mən*, which is also used as a preposition meaning 'from' as well as 'than' after comparatives, is similar to the French preposition *de*, which precedes NPs, as in (19) (for a detailed discussion of French syntax, see Kayne 1975).

(19) *combien* *de* *livres* *a-t-il* *achetés?*
how.many of books has-he bought
'How many books did he buy?'

In fact, we might have either of the following structures:

(20) i. [*šḥal* [*mən* *ktab*]]
NP PP

ii. [*šḥal*] [*mən* *ktab*]
NP PP

In (20i), *šḥal mən ktab* is one constituent, whereas in (20ii) *šḥal* and *mən ktab* are two separate constituents. The same thing can be stated about (18): either *aš mən ktab* represents a single constituent as in:

i. [*aš* [*mən ktab*]]
NP PP

or it consists of *aš* + *mən ktab:*

ii. [*aš*] [*mən ktab*]
NP PP

These two alternative structures could clearly make different predictions about extraction. But in neither case would *šḥal mən/aš mən* be one constituent; hence, it can never undergo preposing. Consider (21a)–(21d).

(21a) *qriti aš/šḥal mən ktab?*
read2s what/how.many of book
'Which book/how many books did you read?'

(21b) *aš/šḥal mən ktab*
what/how.many of book
qriti -----?
read2s -----
'Which book/how many books did you read -----?'

(21c) **aš/šḥal mən qriti -----*
what/how.many of read2s -----
ktab?
book

(21d) *aš/šḥal qriti mən ktab?*
what/how.many read2s of book
'Which/how many books did you read?'

In sum, all these WH-words seem to behave analogously within the framework of WH-Movement, as exemplified below. Hence, on the basis of the typological differences between Moroccan Arabic and English, it is not unexpected that the rule of WH-Movement behaves in a different way in Moroccan Arabic.

Interrogatives in Moroccan Arabic can be divided into various types. One fundamental typological classification, for instance, is between yes/no questions and WH-questions. The former yield 'yes' or 'no' as appropriate answers. WH-questions are so called because in English they make use of an interrogative word starting with *wh-*, e.g. 'what', 'where', 'who', 'how', etc. They are usually employed to inquire about the identity of some entity in the sentence. A second typological classification of questions is between echo and non-echo questions. Echo questions involve someone echoing the speech of another person:

(22) speaker A: *šrit ktab*
bought1s book
'I bought a book'

speaker B: *šriti ktab?*
bought2Sg. book
'You bought a book?'

In this case, speaker B is reiterating a statement made by A by means of yes/no questions. A WH-question could have been used, as in (23):

(23) speaker A: *šrit ktab*
bought1s book
'I bought a book'

speaker B: *šriti aš?*
bought2s what
'You bought what?'

Non-echo questions do not echo statements made by another person. Unlike echo-questions, which cannot initiate a conversation because they are generally used to echo somebody else's statement, non-echo questions can be used to initiate a discourse.

A further typological division is the traditional one between direct and indirect questions. In direct questions, the interrogative construction constitutes an independent sentence, as in (24).

(24) *ʾimta ržəʿ ʾaḥmed?*
when return Ahmed
'When did Ahmed return?'

There are various interrogative words in Moroccan Arabic. Each form corresponds to the questioned element, which could be a subject, an object, an adverb of time or place, an adjective, etc. The basic question morphemes include the following: *škun* 'who?', *šnu* or *ʾaš* 'what?, *ʿlaš* 'why?', *kif* or *bḥalaš* 'how?', *fin* 'where?', *fuqaš* or *ʾimta* 'when?, *kifaš* 'how?'.

Indirect interrogatives are instances of complex sentences, which consist of more than one clause: a main clause and at least one subordinate clause. They are also referred to as interrogative complement clauses. By interrogative complement clauses we mean indirect questions, which are of two major types: (i) yes/no questions, which are obligatorily introduced by the neutral complementizer *waš*; and (ii) WH-questions, which are introduced by different WH-complementizers.

(25a) *ma-ʿraft-š waš had*
not-know1s-not whether this
lə-xbar ṣḥiḥ
the-news correct
‘I don’t know whether this information is correct’

(25b) *sewwel-ni ʿali waš ža*
asked-me Ali whether came
ṣaḥb-u
friend-his
‘Ali asked me whether his friend came’

The matrix verbs in (25) select the complementizer *waš*, the absence of which leads to ungrammaticality:

(26) **sewwel-ni ʿali ža ṣaḥb-u*

Note that when the complementizer *waš* in (25a) is replaced with *bəlli*, the complement clause becomes noninterrogative:

(27) *ma-ʿraft-š bəlli had lə-xbar ṣḥiḥ*
‘I don’t know that this information is correct’

In this case, the complement clause implies that the information is correct, while in (25a) the complement clause implies that the information may be correct or not. However, in the case of (25b), the interrogative verb *sewwel* or *seqsa* ‘to ask’ does not permit the occurrence of *bəlli*, which appears exclusively with declarative clauses:

(28) *sewwel-ni ʿali waš ža ṣaḥb-u*
‘Ali asked me whether his friend had come’

As to WH-interrogative sentences, they may be exemplified by the following structures:

(29a) *ma-ʿraft-š škun ḥall*
not-know1s-not who open3ms
l-bab
the-door
‘I don’t know who opened the door’

(29b) *kan xammem ʿlaš baġi*
was3ms wonder3ms. why wanting
yə-mši l-xariž
he-goes the-abroad
‘He wondered why Ali wanted to go abroad’

(29c) *kan tsa’el fuqaš/kifaš ʿali ġadi*
yə-mši l-xariž
‘He was asking himself when/how Ali wants to go abroad’

(29d) *kan tsa’el fuqaš/fin/kifaš ʿali ġadi*
yə-mši l-xariž
‘He was asking himself when/where/how Ali will go abroad’

Worthy of notice here is that the embedded subject in the interrogative clause can be preverbal, as in (29b) and (29c), or postverbal, as in (30):

(30) *kan xammem šnu ġa-ydir*
was3ms wonder3ms what FUT-do3ms
ʿali f l-xariž
Ali in the-abroad
‘He wondered what Ali would do abroad’

This possibility of placing the subject before or after the verb follows from the fact that Moroccan Arabic allows both SVO and VSO orders.

Thus far, we have seen interrogative complement clauses requiring a particular complementizer, depending on the selectional properties of the matrix verb. With regard to finite interrogative complement clauses, the complementizers involved are of two major types:

Yes-No questions = *waš*
WH-questions = WH-elements (*škun, šnu, fin, ʿlaš, kifaš*, etc.)

The first type of questions is introduced by *waš*. The latter is optional when the question is simple, e.g. (*waš*) *safer mḥammed?* ‘has Mohamed traveled?’ vs. *ma-ʿraft-š waš mḥammed safer* ‘I don’t know whether Mohamed has traveled’. As to WH-questions, they are obligatorily introduced by a WH-element, irrespective of whether the question is direct or indirect; see Ennaji (1985) for more details.

BIBLIOGRAPHICAL REFERENCES

Badawi, Elsaid, Michael G. Carter, and Adrian Gully. 2004. *Modern Written Arabic: A comprehensive grammar*. London and New York: Routledge.

Cowell, Mark W. 1964. *A reference grammar of Syrian Arabic*. Washington, D.C.: Georgetown University Press.

Kayne, Richard S. 1975. *French syntax: The transformational cycle*. Cambridge, Mass.: MIT Press.

Ennaji, Moha. 1985. *Contrastive syntax: English, Moroccan Arabic and Berber complex sentences*. Würzburg: Königshausen und Neumann.

——, Ahmed Makhoukh, Hassan Es-saiydy, Mohamed Moubtassime, and Souad Slaoui. 2004. *A grammar of Moroccan Arabic*. Fes: Faculty of Letters, University of Fes.

Erwin, Wallace M. 1963. *A short reference grammar of Iraqi Arabic*. Washington, D.C.: Georgetown University Press.

Holes, Clive. 1990. *Gulf Arabic*. London and New York: Routledge.

Owens, Jonathan. 1984. *A short reference grammar of Eastern Libyan Arabic*. Wiesbaden: O. Harrassowitz.

Reckendorf, Hermann. 1921. *Arabische Syntax*. Heidelberg: C. Winter.

Watson, Janet C.E. 1993. *A syntax of Ṣanʿānī Arabic*. Wiesbaden: O. Harrassowitz.

Woidich, Manfred. 2006. *Grammatik des Kairenisch-Arabischen*. Wiesbaden: O. Harrassowitz.

Wright, William. 1964. *A grammar of the Arabic language*. 3rd ed. rev. W. Robertson Smith and M.J. de Goeje. Repr. 2 vols. Cambridge: Cambridge University Press.

MOHA ENNAJI (University of Fes)

Intonation

1. INTRODUCTION

Intonation is defined as the linguistic variation of pitch that applies to an utterance as a whole, the resultant tune of which gives the utterance a particular meaning independent of that ascribed by the text (Ladd 1996). The study of intonation has been approached from different perspectives, some scholars investigating it in terms of attitudinal and sociolinguistic factors (e.g. Bolinger 1989), others examining its interface with other parts of the grammar (e.g. Selkirk 1984). More recently, researchers have come to a consensus that a phonological level of abstraction must be posited, mediating between intonation and the other linguistic spheres (Ladd 1996).

In this phonological perspective, different models have been proposed for the description of intonation in various languages, and the description of Arabic intonation has been largely influenced by these models. The American Structuralist school views tunes as being composed of four level tones. The Dutch and British schools of intonation divide the pitch contour configurationally, in terms of dynamic falling-and-rising pitch movements. Currently, the Autosegmental-Metrical (AM) model is a widely accepted model of intonation (Bruce 1977; Pierrehumbert 1980; Gussenhoven 1984; Pierrehumbert and Beckman 1988; Ladd 1996). This model resolves issues raised by previous theories by analyzing apparent rises and falls in intonational contours as a local and linear sequence of two level tones, High (H) and Low (L) (Ladd 1996).

Studies that investigate Arabic intonation point to the fact that Arabic, like intonational languages such as English and Dutch, uses pitch postlexically. Utterances in Arabic thus show various F0 contours, which give them additional meaning. Furthermore, in Arabic an intimate relationship exists between intonation on the one hand, and prosody in its prominence and phrasing functions on the other. Some of the tones composing an intonational contour attach to prominent syllables in the prosodic hierarchy (pitch accents), while others attach to the edges of phrasal constituents in this hierarchy (edge tones). Thus intonation in Arabic fulfills both a 'prominence-lending' and a 'demarcative' function (Beckman 1996).

2. INTONATION AND PROMINENCE

It is widely observed that in the various dialects of Arabic, certain F0 peaks or valleys in intonational contours occur around the lexically stressed syllable of a word (Cairene Arabic: Norlin 1989; Standard Arabic: Haydar and Mrayati 1985; Jordanian Arabic: Rammuny 1989, De Jong and Zawaydeh 1999; Moroccan Arabic: Benkirane 1998; Lebanese Arabic: Chahal 2001). This is taken as evidence of the association of intonational tones to these lexically stressed syllables. Lexical → stress thus plays an important role in the prominence hierarchy of Arabic, since lexically stressed syllables form the potential landing sites of intonational accent in the language.

The location of the lexically stressed syllable in Arabic is predicted by rule. In Levantine dialects, lexical stress assignment rules require that the word-final superheavy syllable (CVVC or CVCC) be stressed, otherwise the heavy penultimate (CVV or CVC), otherwise the antepenultimate (whether heavy or light CV). CV, CVC, and CVV form the maximal word-internal syllables (McCarthy 1979). The superheavy syllable is restricted in distribution, occurring only in word-final position, where it is always stressed (Kenstowicz and Abdul-Karim 1980; Kiparsky 1979; McCarthy 1980; Selkirk 1981; Halle and Vergnaud 1987; Halle and Kenstowicz 1991).

Lexically stressed syllables that get attached to intonational tones are conventionally called pitch-accented syllables. They are more prominent than lexically stressed syllables that are not attached to such accents. In addition, among the various pitch-accented syllables occurring in an utterance, research on Arabic intonation finds that the positionally final pitch accent – the nuclear accent – carries the most prominence (Rammuny 1989; De Jong and Zawaydeh 1999; Chahal 2001). There are thus three paradigmatic levels of prominence in Arabic: lexical stress, (prenuclear) pitch accent, and nuclear accent, the latter two levels being ascribed by intonation.

While syllables bearing pitch-accent or nuclear-prominence levels are characterized by their association to tonal targets, they may also bear nontonal cues. In Lebanese Arabic, for example, it is found that nuclear-accented, accented, and unaccented syllables are significantly distinguished in terms of a combination of Fo, duration, RMS, F1, and F2 correlates (Chahal 2003). The higher the prominence level of a particular vowel, the higher its pitch and amplitude, the longer its duration, and the more peripheral its spectral realization in the vowel space. These results reiterate findings on other dialects of Arabic (Al-Ani 1992; Belkaid 1984; De Jong and Zawaydeh 1999; Haydar and Mrayati 1985), thereby confirming the typological classification of Arabic as a stress-accent language, similar to English (Beckman 1986).

The idea that the nuclear accent in Arabic constitutes the final pitch accent in an intonational phrase is evidenced by various studies on intonational focus (Cairene Arabic: Norlin 1989; Lebanese Arabic: Chahal 2001; Moroccan Arabic: Benkirane 1998). When a declarative utterance receives broad focus in Arabic, it shows one or more prenuclear accents and a final nuclear accent in a predominantly falling contour. When a particular item of this utterance is narrowly focused, the narrowly focused item bears a nuclear accent and receives an expanded Fo value, while the postfocus material is invariably realized in a compressed and level pitch range. This postfocal compression is suggestive of de-accenting, thereby making nuclear accent assignment in Arabic inherently right-headed.

3. Intonation and phrasing

In addition to prominence-lending tones, Arabic displays intonational edge tones that carry out a phrasing function. These edge tones group utterance segments into phrasal constituents and delimit the right-edge boundaries of these phrasal constitutents. Studies investigating Arabic suggest three intonationally relevant phrasal constituents. The first is the prosodic word, which is a widely implied constituent since it defines the unit within which lexical stress is assigned. The remaining two phrasal constituents are generally referred to as minor and major boundaries (Alharbi 1991; Rammuny 1989). Using Autosegmental-Metrical terminology, these correspond to the intermediate and intonational phrase respectively.

According to Autosegmental-Metrical analyses of Arabic (De Jong and Zawaydeh 1999; Chahal 2001), complex utterance-final pitch patterns in the language (such as falling-rising contours) can be explained in terms of a sequence of two distinct edge tones. The first is a phrase accent that associates itself to the edge of the intermediate phrase, while the second is a boundary tone that associates itself to the right edge of the intonational phrase. Besides being demarcated by a right-edge tonal event, the intermediate phrase defines the domain within which relative prominence patterns are established (each intermediate phrase is composed of at least one pitch accent) and within which pitch is reset. The intonational phrase, on the other hand, is a higher-level constituent composed of one or more intermediate phrases; it is demarcated by a boundary tone and defines the domain within which global declination and local final lowering occurs.

While the intermediate phrase and the intonational phrase are primarily defined by the presence of edge tones, these tonally marked boundaries receive junctural cues additionally delimiting their right edge. It is widely noted that syllables occurring at the end of intonational phrases in Arabic receive phrase-final lengthening (Haydar and Mrayati 1985; Alharbi 1991; Al-Ani 1992; De Jong and Zawaydeh 1999) and pausing phenomena (Rammuny 1989). There is also evidence of segmental sandhi-blocking phenomena of the type discussed in Nespor and Vogel (1986), whereby voiceless alveolar stops are heavily aspirated when they occur at the right edge of major boundaries (El-Imam 1990).

In addition, junctural phenomena in Arabic seem to be affected by boundary strength. Accented syllables located at the end of intonational phrases are significantly longer in duration than corresponding syllables found at intermediate phrase boundaries, and the latter are in turn significantly longer than corresponding syllables found at prosodic word edges. Similarly, Fo accentual peaks occurring at the end of intonational phrases are aligned earlier within the bounds of the syllable they are associated with than those occurring at the edge of intermediate phrases, while the latter are earlier aligned than peaks occurring at the edge of prosodic word boundaries (Chahal 2001).

4. TUNES

The combination of pitch accents and edge tones discussed above constitutes the tune of a particular intonational phrase. Various studies describe the nuclear tunes found in different varieties of the Arabic language, most using analyses in the British school style. In Modern Standard Arabic, Haydar and Mrayati (1985) note that nominal affirmative, negative, and exclamatory sentences show a falling configuration. They also found that interrogatives display a rising configuration in this dialect. Kharrat (1994) reports similar findings for this dialect, additionally noting the usage of level tunes.

A study of Syrian Arabic shows that it displays three phonological nuclear melodies – falling, rising, and level (Corvetto 1982). These melodies exhibit phonetic differences, depending on the configurations of syllables occurring before the nuclear word (the heads) and those occurring after it (the tails). The three nuclear tunes are employed in diverse sentence types: falling melodies are used in assertions, commands, and exclamations; rising melodies are used in WH-questions and yes/no questions; and level contours are used in incomplete phrases, questions with alternative answers, and vocative exclamations.

In Jordanian Arabic, Rammuny (1989) identifies (1) falling contours, which start from high or mid and subsequently fall to low; (2) rising contours, which start from low and then rise to mid or high; (3) level contours, which start and remain either high or mid in the pitch range; (4) rising-falling contours, which start from low, then rise to high or mid level, then fall back to low; and (5) rising-falling-rising contours. Investigating the same dialect, De Jong and Zawaydeh (1999) describe four major tunes occurring in single-word intonational phrases: (1) a rising-falling tune, which is realized as an Fo peak followed by a subsequent fall at the right edge of the utterance, indicating general statements; (2) a rising tune, which is generally produced as a rising pitch configuration beginning with either low or high pitch, and indicating question tunes; (3) a rise-plateau tune, which is composed of an Fo peak followed by a high plateau extending until the end of the utterance, and indicating incompleteness; and (4) a rising-falling-rising tune, which shows a similar Fo peak followed by a fall and a subsequent rise at the right edge of the utterance, and may be used in question forms.

For the Kuwaiti dialect and for utterances containing a single nucleus, Alharbi (1991) proposes a set of falling, rising, and level simple nuclear tunes – tunes occurring across utterances displaying a single nucleus and unidirectional pitch movement, and falling-rising and rising-falling complex tones – tunes occurring across utterances containing a single nucleus and bidirectional pitch movement. Alharbi notes that the final pitch movement following the nucleus constitutes the most important pitch movement in an utterance.

The Autosegmental-Metrical analysis of Lebanese Arabic (Chahal 2001) identifies (1) declarative tunes, which show a falling configuration at the edges of phrases, taking the shape of a 'pointed hat' or a 'flat hat' ('t Hart a.o. 1990), and indicating neutral statements; (2) question tunes, which consist of a high rising edge configuration usually preceded by low pitch occurring on the nuclear-accented syllable and which are used in yes/no question forms; (3) plateau tunes, which illustrate a high plateau extending from the nuclear accent to the edge of the intonational phrase and which indicate an unmarked type of incompleteness; (4) marked continuation tunes, which consist of a falling-rising edge and indicate a marked type of continuation; and (5) downstepped or stylized plateau tunes, consisting of a plateau the elbow of which is realized in the middle of the speaker's range (it is downstepped), and which conveys a sense of polite and mild reproach.

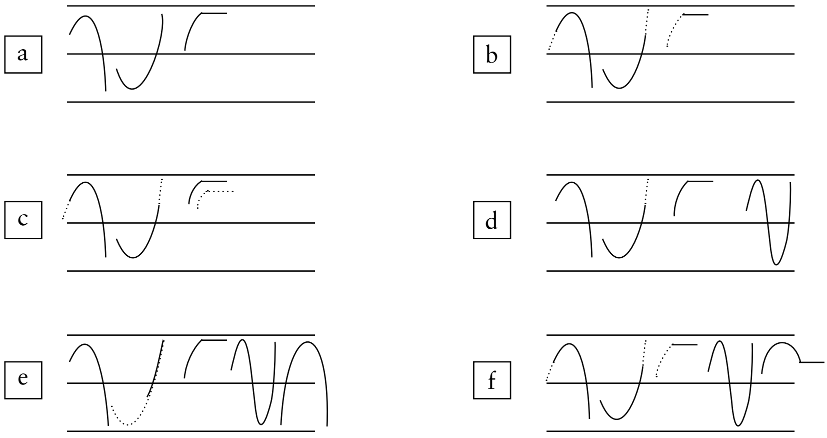
Figure 1 schematizes the various nuclear tunes found for different dialects of Arabic.

In order to account for tunes found in particular dialects of Arabic, certain Autosegmental-Metrical studies posit a tonal inventory for the pitch accents and edge tones composing these tunes. For example, De Jong and Zawaydeh (1999) find that Jordanian Arabic tonal patterns may be explained in terms of a pitch-accent inventory that includes a high and a low monotonal pitch accent (H* and L*) and a bitonal pitch accent, composed of a low tonal target that occurs within the bounds of the associated syllable and is preceded by high-level pitch (H+L*). The tonal inventory also includes a high and a low phrase accent (H- and L-) and a high and a low boundary tone (H% and L%).

The proposed model for Lebanese Arabic (Chahal 2001) posits a tonal inventory composed of six pitch-accent types (H*, L+H*, L*, !H*, L+!H*, and H+!H*), three phrase-accent types (H-, L-, and !H-), and two boundary tones (L% and H%). Like Jordanian Arabic, pitch accents in Lebanese Arabic may be bitonal. However, it is the rightmost tone that associates with the accented syllable in this dialect. High pitch accents and phrase accents (!H* and !H-) may be affected by downstep, a local tonal-implementation rule that considerably lowers the pitch range of these high tones. Boundary tones are affected by upstep, another tonal-realizational rule, which raises the Fo value of boundary tones following H- phrase accents. Upstep accounts for the existence of high-rising boundary configurations and the lack of rise-fall boundaries in this dialect (similar to American English). L% tones are optionally affected by final lowering effects.

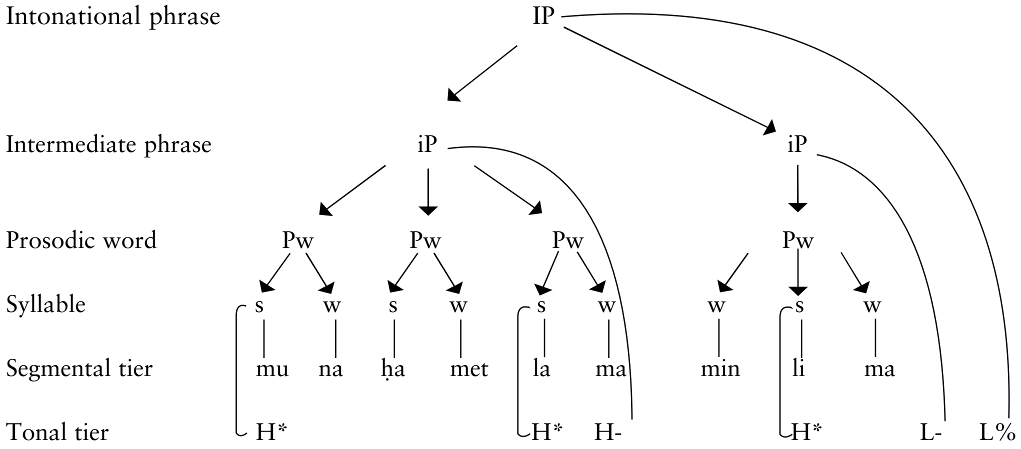
Figure 2 provides a tonal representation of the intonation of an Arabic utterance using the Lebanese Arabic inventory. It also represents the association of tonal targets with their relevant prominence and phrasal constituents, conventionally using prominence grids and metrical trees.

Figure 1. Schematization of the various nuclear tunes found for different dialects of Arabic



(a) Pitch contours for Standard Arabic (Haydar and Mrayati 1985; Kharrat 1994)
(b) Pitch contours for Syrian Arabic (Corvetto 1982)
(c) Pitch contours for Jordanian Arabic (Rammuny 1989)
(d) Pitch contours for Jordanian Arabic (De Jong and Zawaydeh 1999)
(e) Pitch contours for Kuwaiti Arabic (Alharbi 1991)
(f) Pitch contours for Lebanese Arabic (Chahal 2001)
Dotted lines indicate optional pitch configurations.

Figure 2. A representation of the intonation and prosodic hierarchy of Lebanese Arabic, illustrated for the utterance /muna ḥamet lama min lima/ ‘Muna protected a llama from Lima’



Nuclear accent	[				x	]	[	x	]
Pitch accent	X				x			x	
Lexical stress	X		x		x			x	
Syllable	X	x	x	x	x	x	x	x	x
Utterance	mu	na	ḥa	met	la	ma	men	li	ma

The utterance in Figure 2 is produced as two intermediate phrases within a single intonational phrase. In the first intermediate phrases, /muna/ and /lama/ are pitch accented, and in the second, /lima/ is accented. The intonational structure of the utterance is: H* H* H- H* L-L%. In the prominence grid, square brackets indicate intermediate phrase boundaries.

5. Comparative studies

A small body of work on Arabic intonation is comparative in nature, delineating similarities and differences between the intonation of Arabic and that of the English language. El-Hassan (1988), for instance, compares the intonation of questions in Standard Arabic with English counterparts. The author finds that in WH-questions, both languages predominantly use a falling nuclear melody. However, whereas in English a low fall is interpreted as hostile and a high fall as brisk and business-like, in Arabic a low fall represents an ordinary question, with no connotations of hostility, while a high fall suggests a high degree of interest verging on objection and disbelief. In yes/no questions, El-Hassan finds that both Arabic and English usually employ a rising tune. But he argues that English uses a variety of falling melodies (such as falling tunes in rejoinders and tag questions, and rise-fall tunes) that are alien to Arabic yes/no questions.

In a comparison of Lebanese Arabic with American English (Chahal 2001), it was found that while mid-scaled (or downstepped) plateau contours occur in both languages, in English they are characteristic of stylized calling contours (Ladd 1978), whereas in Lebanese Arabic they denote statements conveying a sense of mild reproach. Also, whereas incompleteness is indicated by falling-rising tunes in English, in Lebanese Arabic it is indicated by plateau tunes. Phonotactically, both English and Lebanese Arabic require similar rules of association of pitch accents to lexically stressed syllables, and nuclear accents to right-most positions.

One phonotactic difference between the two languages, however, is that L* and H* are restricted in distribution in rising phrase-final syllables in Lebanese Arabic (i.e., they do not occur in such phrasal contexts), whereas they do not show such distributional constraints in English. As for phonological contrasts, it was found that whereas English contrasts left-headed vs. right-headed bitonals (e.g. H*+L and L*+H and H+L* vs. L+H*), bitonal accents in Arabic do not show such contrasts, as they are all right-headed. Also, English displays a categorical contrast dependent on the phonetic alignment of one tone within the boundaries of the accented syllable (e.g. L+H* vs. L*+H), while no evidence for such contrast is found in Lebanese Arabic.

6. CONCLUSION

Based on the above literature, it is found that Arabic falls within the typology of intonational languages where the linguistic usage of pitch applies at the level of the utterance as a whole. Arabic thus differs from languages such as Mandarin Chinese, where tone is used lexically. Pitch in Arabic also carries out both prominence and grouping functions. Certain tonal events associate with lexically stressed syllables, thereby contributing to the relative prominence hierarchy of the language. Other tones demarcate the right edge of certain intonational boundaries occurring in the constituency hierarchy of the language. Arabic is thus similar to languages such as English and Dutch, which carry out these functions, and is dissimilar to languages like French or Korean, where pitch is seen to fulfill a purely delimitative function. Related to the prominence-enhancing function is the classification of Arabic as a stress-accent language; intonationally assigned prominence patterns in the language demonstrate nontonal stress cues.

The by-product of the combination of prominence and edge tone types is a tune that gives the utterance additional meaning not predicted by that of the text. Three basic contours are shared by all dialects of Arabic: falling, rising, and level or plateau contours. Additional tunes may be found, depending on the particular dialect examined. While formally these tunes occur in other languages, such as English, they may differ in the meanings they convey, in the phonotactic constraints they follow, and in the phonological contrasts they display.

BIBLIOGRAPHICAL REFERENCES

Al-Ani, Salman. 1992. "Stress variation of the construct phrase in Arabic: A spectrographic analysis". *Anthropological Linguistics* 34.256–276.

Alharbi, Lutfi. 1991. *Formal analysis of intonation: The case of the Kuwaiti dialect of Arabic.* Ph.D. diss., Herriot-Watt University.

Beckman, Mary. 1986. *Stress and non-stress accent.* Dordrecht: Foris.

——. 1996. "The parsing of prosody". *Language and Cognitive Processes* 11.17–67.

Belkaid, Yamina. 1984. "Les voyelles de l'arabe littéraire moderne: Analyse spectrographique". *Travaux de l'Institut de Phonétique de Strasbourg* 16.217–240.

Benkirane, Thami. 1998. "Intonation in Western Arabic (Morocco)". *Intonation systems: A survey of twenty languages,* ed. Daniel Hirst and Albert Di Cristo, 345–359. Cambridge: Cambridge University Press.

Bolinger, Dwight L. 1972. "Accent is predictable (if you're a mind reader)". *Language* 48.633–644.

——. 1989. *Intonation and its use: Melody in grammar and discourse.* Stanford, Calif.: Stanford University Press.

Bruce, Gosta. 1977. *Swedish word accents in sentence perspective.* Lund: Gleerup.

Chahal, Dana. 2001. *Modeling the intonation of Lebanese Arabic using the Autosegmental-Metrical framework: A comparison with English.* Ph.D. diss., University of Melbourne.

——. 2003. "Phonetic cues to prominence in Lebanese Arabic". *Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona, 1 August-7 August 2004.* http://shylock.uab.es/icphs.

Corvetto, Ines. 1982. "L'intonazione nell'arabo siriano" [The intonation of Syrian Arabic]. *Lingua e Stile* 17:2.371–394.

De Jong, Kenneth and Bushra Zawaydeh. 1999. "Stress, duration, and intonation in Arabic word-level prosody". *Journal of Phonetics* 27.3–22.

El-Hassan, Shahir. 1988. "The intonation of questions in English and Arabic". *Papers and Studies in Contrastive Linguistics* 22.97–108.

El-Imam, Youssef. 1990. "Speech synthesis using partial syllables". *Computer Speech and Language* 4.203–229.

Gussenhoven, Carlos. 1984. *On the grammar and semantics of sentence accents.* Dordrecht: Foris.

Halle, Morris and Michael Kenstowicz. 1991. "The free element condition and cyclic versus noncyclic stress". *Linguistic Inquiry* 22:2.457–501.

Halle, Morris and Jean-Roger Vergnaud. 1987. *An essay on stress.* Cambridge, Mass.: MIT Press.

Hart, Johan 't, Rene Collier, and Antonie Cohen. 1990. *A perceptual study of intonation: An experimental-phonetic approach to speech melody.* Cambridge: Cambridge University Press.

Haydar, Y. and M. Mrayati. 1985. "Etude de l'intonation, la courbe melodique de phrases de l'arabe standard". *Travaux de l'Institut de Phonétique de Strasbourg* 17.75–113.

Kenstowicz, Michael and Kamal Abdul-Karim. 1980. "Cyclic stress in Levantine Arabic". *Studies in the Linguistic Sciences* 10:2.55–75.

Kharrat, Muhammad. 1994. *A descriptive analysis of Standard Arabic intonation with implications for teaching English as a second language.* Ph.D. diss., University of Kansas.

Kiparsky, Paul. 1979. "Metrical structure assignment is cyclic". *Linguistic Inquiry* 10.421–442.

Ladd, Robert D. 1978. "Stylized intonation". *Language* 54.517–539.
——. 1996. *Intonational phonology*. Cambridge: Cambridge University Press.
McCarthy, John. 1979. "On stress and syllabification". *Linguistic Inquiry* 10.443–465.
——. 1980. "A note on the accentuation of Damascene Arabic". *Studies in the Linguistic Sciences* 10.77–98.
Nespor, Marina and Irene Vogel. 1986. *Prosodic phonology*. Dordrecht: Foris.
Norlin, K. 1989. "A preliminary description of Cairo Arabic intonation of statements and questions". *Speech Transmission Quarterly Progress and Status Report* 1.47–49.
Pierrehumbert, Janet. 1980. *The phonology and phonetics of English*. Ph.D. diss., Massachusetts Institute of Technology.
—— and Mary E. Beckman. 1988. *Japanese tone structure*. Cambridge, Mass.: MIT Press.
Rammuny, Raji. 1989. "Instrumental and auditory analysis of colloquial Jordanian Arabic, an update". *Journal of Arabic Linguistics* 20.23–42.
Selkirk, Elizabeth. 1981. "Epenthesis and degenerate syllables in Cairene Arabic". *Theoretical issues in the grammar of Semitic languages*, ed. Hagit Borer and Joseph Aoun, 209–232. (= *MIT Working Papers in Linguistics*, III.) Cambridge, Mass.: MIT Press.
——. 1984. *Phonology and syntax: The relation between sound and structure*. Cambridge, Mass.: MIT Press.

DANA CHAHAL (University of Balamand)

ʾIʿrāb

The verb *ʾaʿraba*, from which the term *ʾiʿrāb* is derived, means 'to use good Arabic style, to express one's mind clearly, to make known' and, in a technical sense, 'to pronounce the final short vowels of a word', hence the usual translation of *ʾiʿrāb* as 'declension'. The *ʾiʿrāb* is the main distinctive feature of the *ʿarabiyya*, the language used by the Arabs of the desert, and in particular the form of expression of the oldest poetry, transmitted by the *ruwāt* 'reciters' of certain trustworthy Bedouin tribes. It is usually regarded as a synonym of *bayān* 'clear expression' (e.g. Zajjājī, *ʾĪḍāḥ* 91; cf. Carter 1981:34).

The explanations given concerning the original meaning of this word show its central place in Arab grammar. Arabic sources date its first appearance to the beginnings of Arab linguistic activities, specifically to ʾAbū l-ʾAswad ad-Duʾalī, the putative father of Arab grammar (Suyūṭī, *Tuḥfa* 49). Reportedly, he was persuaded to write a grammar of Arabic when he heard someone reciting Q. 9/3 with a faulty *ʾiʿrāb* ending (ʾAbū ṭ-Ṭayyib, *Marātib* 24–29). Different sayings recommending the use of *ʾiʿrāb* are attributed to prominent persons, among them for instance the 8th-century jurist Mālik ibn ʾAnas, who said: "*ʾIʿrāb* is the jewelry of your tongue, so do not deprive your tongue of its jewelry" (Zubaydī, *Ṭabaqāt* 13 *al-ʾiʿrāb ḥaly al-lisān fa-lā tamnaʿū ʾalsinatakum ḥalyahā*). Versteegh (1977:61ff.) following Merx (1889:56–62) argued that, notwithstanding their different technical meanings, *ʾiʿrāb* is a calque of the Greek term *hellènismós*, but this is rather improbable in view of the occurrence of the term *ʾiʿrāb* in one of the earliest Qurʾānic commentaries, Muḥammad al-Kalbī's (d. 146/763) *Tafsīr*, where it has the sense of 'speaking Arabic correctly as a Bedouin' (Versteegh 1993:127–128). This obviates the need for an explanation in terms of foreign origin.

The emphasis placed by the Arabic grammarians on the correct use of the case endings shows that its original function had become obscured by the 8th century, or even by the end of the 7th century, if we give some credit to the traditions about ʾAbū l-ʾAswad. Vollers (1906:169) maintains that the use of declensional endings (*ʾiʿrāb*) had died out by 600 C.E. from spoken Arabic. In combination with the presence of alternating endings in the corpus of → *kalām* studied by the grammarians (including *Qurʾān* and poetry), this explains the central position of *ʾiʿrāb* in Arabic grammar. It also accounts for the fact that grammarians accept that in a given construction different *ʾiʿrāb* endings of a word may be found; if one of the alternatives is favored, for instance because of theological considerations, this ending is then explained painstakingly (cf. Dévényi 1987–1988:201). Such an approach would have been impossible if the endings still functioned as real case markers.

The central place of *ʾiʿrāb* became even more conspicuous in shorter grammatical treatises written by later grammarians. Even in a very late author like the 19th-century Nāṣif al-Yāzijī, we find the following definition: "Grammar is the discipline by which the states of the endings of the words are known regarding *ʾiʿrāb* and *bināʾ*" (*Faṣl* 94 *an-naḥw ʿilm tuʿrafu bihi ʾaḥwāl ʾawāxir al-kalim min jihat al-ʾiʿrāb wa-l-bināʾ*). A series of present-day works announce the central role of *ʾiʿrāb* already in their titles, e.g. *Qāmūs al-ʾiʿrāb* (al-ʾAsmar 1986).

ʾIʿrāb is regarded as an essential characteristic of Arabic, or more precisely of *kalām*. It is dealt with by *naḥw* 'grammar' or 'syntax', and in this sense it is contrasted with → *ṣarf/taṣrīf*, which deals with morphological and phonological changes in the declension of nouns and the conjugation of verbs, whereas *ʾiʿrāb* deals with syntactic changes. Changes in person and number are not part of *ʾiʿrāb*.

The term *ʾiʿrāb* is used first of all for the short vowel endings, which alternate according to the underlying rules of the language. These changes were generally regarded as correlating with syntactic functions. The task of the 8th-century grammarians was to reveal the causes and circumstances of these changes, to find satisfactory explanations for *ʾiʿrāb* on the basis of the vast corpus that was partly collected by the philologists and partly memorized by the transmitters. The raison d'être of Arabic grammarians was to sort out this corpus and state the rules governing the short vowel endings. Their task did not consist in deciding whether in a given line of poetry a word should have an *-a* or an *-u* ending, or both, but in explaining each alternative. Al-Farrāʾ's explanations in his *Maʿānī l-Qurʾān* typically contain such phrases as: if it ends in *-u*, there are three ways (*wajh*) of explanation; if it ends in *-a*, there are also three ways of explanation, and if there is a reading with an *-i* ending, there are also three ways. The meaning of the expression was not necessarily at stake here, but often remained the same in all alternatives.

During the early period of Islamic civilization, possibly during the 7th century, two sets of labels had been established for the three short vowels (→ *ḥarakāt*) that were not expressed in script: for *-u, rafʿ* and *ḍamm* were used; for *-a, naṣb* and *fatḥ*; for *-i, jarr* (or *xafḍ*) and *kasr*; and for the lack of a vowel, the terms *jazm* and *waqf* (later *sukūn*) were used. Both sets were still used indiscriminately for final and non-final vowels by al-Farrāʾ (Owens 1990:159), and even by al-Xalīl, Sībawayhi's teacher (Talmon 1997:194–197), and in a few cases by Sibawayhi himself (Talmon 2003:239–244). The names for the various endings, whether declensional or non-declensional, possibly derive from Syriac grammar (Versteegh 1995:159–162; → *ḥaraka*).

This lack of distinction between declensional and non-declensional endings disappeared with Sībawayhi's innovatory approach to the analysis of language. The second chapter of his *Kitāb* (*Bāb majārī ʾawāxir al-kalim min al-ʿarabiyya* 'Chapter on the ending of words in Arabic') is dedicated to the definition of *ʾiʿrāb* and its counterpart, → *bināʾ*. Sībawayhi divides the three short vowel endings and the zero morpheme (*-u, -a, -i, -ϕ*) into two sets of four endings, one for *ʾiʿrāb* endings (*rafʿ, jarr, naṣb, jazm*) and the other for *bināʾ* endings (*ḍamm, fatḥ, kasr, waqf*). His innovation consisted in assigning to the members of the former set (*rafʿ, naṣb, jarr, jazm*) a syntactic status, while the latter set served as the names for short vowels both in non-final position and in those endings that did not change (*bināʾ*).

As the next step, the grammarians identified other endings – long vowels and consonants, expressed in Arabic script by letters (*ḥurūf*) – as being analogically similar to the short vowels. This applied, for instance, to the declension of words like *ʾabun* 'father' (*ʾabū-ka, ʾabī-ka, ʾabā-ka*); their endings contained a short vowel accompanied by a glide (/uw/, /iy/, /a''/) and could be analyzed as having a partly consonantal ending. Likewise, in the plural endings *-ūna/-īna* and the dual endings *-āni/-ayni*, some grammarians posited a declension by means of consonants (cf. Zajjājī, *ʾĪḍāḥ* 72–75; Ibn al-ʾAnbārī, *ʾInṣāf* 6–18). Thus, the terms used for the short vowel *ʾiʿrāb* endings were expanded to cover these consonants as well, although the short vowels remained the basic and primary (*ʾaṣl, ʾawwal*) *ʾiʿrāb* markers.

Already in the *Kitāb*, a further step was made to expand *ʾiʿrāb* to cases where no material marker at all was present. The method used by the grammarians to establish the declensional status in such cases was the restitution of the underlying structure (→ *taqdīr*); this became one of the most important tools of linguistic analysis. If a whole phrase or even clause is used in a place where originally a single word occurred with a definite *ʾiʿrāb* category, this phrase would be considered as having implicitly the same *ʾiʿrāb* category. In this way, even clauses could be said to be in *rafʿ*, although originally this term was reserved for words having the ending *-u*.

Sībawayhi not only distinguished between declensional and non-declensional endings, he also explained the declension as the result of the action (→ *ʿamal*) of another word. The usual definition of *ʾiʿrāb* in medieval Arabic grammars runs as follows: "*ʾIʿrāb* is the alteration of the endings of a word because of the variation of the regents entering upon it, either verbally or implicitly" (*al-ʾiʿrāb taġyīr ʾawāxir al-kalim li-xtilāf al-ʿawāmil ad-dāxila ʿalayhā lafḏan ʾaw taqdīran*; Ibn ʾĀjurrūm, *Muqaddima*, Chap. 2; see Carter 1981:34). In this definition, *taġyīr* 'alteration' is the most important element, since without alteration there is no *ʾiʿrāb*. What does not change is called *bināʾ* and is less problematic because there is no need for seeking the ways and causes of its variants. But the definition also introduces the cause of the alteration, which was found in the 'regents' or 'operators' (*ʿawāmil*) that operate on the words with *ʾiʿrāb*. Finding and explaining these regents became one of the central preoccupations of the grammarians to such an extent that Arabic grammar almost became identified as *ʿilm al-ʿawāmil* 'the study of the regents', and the regents sometimes seem to be more important than the *ʾiʿrāb* endings themselves. Yet, this is not quite true. Arabic grammar had started as the study of *ʾiʿrāb*, and these endings remained its main focus throughout its history. As al-ʾArdabīlī (d. 617/1220), the commentator of az-Zamaxšarī's (d. 538/1144) *ʾUnmūḏaj* (3) puts it: "The purpose of [grammar] is [to acquire] the knowledge of *ʾiʿrāb*" (*al-ġaraḍ minhu maʿrifat al-ʾiʿrāb*). The study of regents was part of linguistic theory, while the *ʾiʿrāb* endings were regarded as legitimized by the speakers themselves. Grammarians who did not recognize the limits of their theorizing are criticized by Sībawayhi (*Kitāb* I, 157.12) when he says that one or the other form or phrase is only the result of a linguistic rule but is not acknowledged by informants of the *ʿarabiyya*: "This is [only the grammarians'] example, but not used in the speech [of the Arabs]" (*fa-hāḏā tamṯīl wa-lākinnahu lam yustaʿmal fī l-kalām*).

In the long history of Arabic grammar writing, surprisingly few alternatives have been brought forward to the theory of *ʾiʿrāb* and *ʿawāmil*. One grammarian who voiced a different opinion concerning the status and origin of *ʾiʿrāb* endings was Sībawayhi's student Quṭrub (d. 206/821). He maintained that different *ʾiʿrāb* endings are not used to indicate different meanings (*lam yuʿrab al-kalām li-d-dalāla ʿalā l-maʿānī wa-l-farq bayna baʿḍihā wa-baʿḍ*) and do not correlate with the action of a regent, but their function is to ensure the alternation of vowelless consonants and consonants with a vowel (Zajjājī, *ʾĪḍāḥ* 70–71; cf. Versteegh 1981:407–408).

Another alternative to the theory of *ʿawāmil* explains the *ʾiʿrāb* endings by proximity (*ʿalā l-jiwār*). Although this explanation has not gained prominence, it is present throughout the history of Arabic grammar as a less speculative and more descriptive way of explanation (see e.g. Ibn Jinnī, *Xaṣāʾiṣ* III, 218–227, Chap. 147 *Bāb al-jiwār*; cf. Dévényi 1987–1988). An example of this principle is when an attribute to the first member of an *ʾiḍāfa* attracts the genitive ending of the second member, e.g. *nasja al-ʿankabūti l-murmali* 'the fabric [acc.] of the spider [gen.], the woven [gen.]' (Ibn Jinnī, *Xaṣāʾiṣ* III, 221.2).

The term *ʾiʿrāb* is characteristic of Arabic grammar and does not lend itself easily to identification with any European grammatical term or expression. Rabin (1969:190) has drawn attention to the fact that a forced correspondence between Indo-European and Semitic case endings might lead to a false interpretation of these endings. His reevaluation of the Semitic system led him to the reconstruction of an original system of case endings (1969:201). The most fundamental difference between Arabic *ʾiʿrāb* and Indo-European declension is that *ʾiʿrāb* covers both nominal and verbal endings, so that the notion of *rafʿ* '-*u* ending, nominative' is applied to the -*u* ending of both nouns and (imperfect) verbs (i.e. both the 'nominative case' and the 'indicative mood'). Whereas nouns possess *ʾiʿrāb* by nature (→ *ʾaṣl*), verbs acquire it only by their similarity (*muḍāraʿa*) to the nouns. This resemblance is found in the imperfect verb because it can be used in some of the syntactic contexts in which a noun can be used, as in *ʾinna zaydan la-yaktubu*/*ʾinna zaydan la-kātibun* 'Zayd is writing' (→ *māḍī*/*muḍāriʿ*).

The notion of 'resemblance' also plays a role in the nominal declension. Nouns are called *munṣarif* when they receive all *ʾiʿrāb* endings, including the → *tanwīn*; those nouns that lose part of their *ʾiʿrāb* (the so-called diptotic nouns; → diptosis) are called *ġayr munṣarif*. These nouns resemble verbs or particles because of their phonological shape (e.g. a proper name like *yazīd* which looks like an imperfect verb), and for that reason they lose (part) of their *ʾiʿrāb*. Note that the term *munṣarif* is of the same root as → *ṣarf* 'morphology' and *taṣrīf al-ʾafʿāl* 'conjugation of verbs', which have nothing to do with *ʾiʿrāb*, according to the grammarians. On the concepts behind Sībawayhi's theoretical system with special regard to *ʾiʿrāb* and *bināʾ*, see Baalbaki (1990).

The *ʾiʿrāb* endings that express the syntactic functions do not always have the same phonetic shape. In nouns, the 'declension by means of consonants' was mentioned above; in verbs, on the analogy of *yaktubu* (*rafʿ*), *yaktuba* (*naṣb*), *yaktub* (*jazm*), a similar series may be modeled from verbal forms ending in consonants: *taktubīna* (*rafʿ*), *taktubī* (*naṣb, jazm*). Whether or not the endings apart from a syntactic function also stand for an underlying meaning is controversial (cf. Ermers 1999:170ff.). In the case of phrases or clauses having a certain syntactic function, their *ʾiʿrāb* status is implicit and acquired by analogy. In the nominal sentence *zaydun qāʾimun* 'Zayd is standing', for instance, the nominal predicate (*xabar*) *qāʾimun* 'standing' is in *rafʿ*, so when a whole clause occupies its place it will also be designated as *rafʿ*. In *zaydun qāma ʾabūhu* 'Zayd, his father stood up', the clause *qāma ʾabūhu* acquires *rafʿ*, implicitly (*taqdīran*), without having an explicit formal marker. Inside this clause, *ʾabūhu* 'his father' is also in *rafʿ*, because it is the agent of the verb. In these constructions, the assignment of a case is purely formal and does not correlate with any underlying 'meaning'. According to az-Zamaxšarī (*Mufaṣṣal* 10.19ff.), the endings of the noun indicate a → *maʿnā* 'meaning' (*kull wāḥid minhā ʿalam ʿalā maʿnan*); in (imperfect) verbs, on the other hand, the endings are again purely formal and do not stand for any meaning (*Mufaṣṣal* 109.8ff.: *wa-laysat hāḏihi l-wujūh bi-ʾaʿlām ʿalā maʿānin ka-wujūh ʾiʿrāb alʾism*).

ʾIʿrāb and *bināʾ* are an integral part of the structure of language and cannot exist outside of a context. An individual word is deprived of any endings, al-ʾArdabīlī (*ʾUnmūḏaj* 3) says: "*ʾIʿrāb* and *bināʾ* cannot exist without entering into a syntactic structure that only exists in the [intended] speech [i.e. *al-ʿarabiyya*]" (*al-ʾiʿrāb wa-l-bināʾ lā yūjadu ʾillā fīmā yaqaʿu fī t-tarkīb al-ʾisnādī allaḏī lā yūjadu ʾillā fī l-kalām*). This is the reason why grammars always exemplify *ʾiʿrāb* endings in context: *zaydun munṭaliqun, raʾaytu zaydan, marartu bi-zaydin* 'Zayd is departing; I saw Zayd; I passed by Zayd' for the nominative, accusative, and genitive cases, respectively.

The strict relation between → *kalām* and *ʾiʿrāb* is exemplified very well by the discussion on the priority of the one over the other, best summed up by az-Zajjāji (*ʾĪḍāḥ* 67). According to one view, the *kalām* has precedence over the *ʾiʿrāb*, since we see *kalām* without *ʾiʿrāb* endings: *zayd wa-muḥammad* has the same meaning with and without *ʾiʿrāb*, and the two words remain nouns. According to other views, however, it is only *ʾiʿrāb* that makes *kalām* what it is, since *kalām* always has a composite nature and a compound cannot do without *ʾiʿrāb*.

Short vowel endings, both those of *ʾiʿrāb* and *bināʾ*, may be dropped in pause (*waqf*). This shows that a fundamental difference exists between *ʾiʿrāb* with *ḥaraka* and *ʾiʿrāb* with *ḥarf* in spite of their being brought together in common categories. At the same time, there are three different sets of → pausal forms: pausal forms in prose, in poetry in *qāfiya muṭlaqa* 'fettered rhyme', and in *qāfiya muqayyada* 'loose rhyme'. In poetic pause, there may occur more than simply the dropping of the vowels: *ʾiʿrāb* endings may be shifted backward, resulting in *bakur* instead of *bakru* (ascribed by Ibn Jinnī, *Xaṣāʾiṣ* III, 220, to the working of *jiwār*). If the same phenomenon occurs inside the line, it is called *ḍarūra* (→ poetic license). For the first exposition on the *qāfiya* and *waqf*, see Sībawayhi (*Kitāb* II, 507–508, Chap. 325; II, 493ff., Chap. 306ff.). The best overview of pausal forms in European literature is still Birkeland (1940).

The preservation of the case endings in the spoken dialects in the 8th century and even earlier is a controversial issue. In contemporary dialects, instances of the use of → *tanwīn* endings in some of the Bedouin dialects (→ Najdi Arabic) are sometimes advanced as evidence of the preservation of the case system among the Bedouin. Yet, this misrepresents the nature of the *ʾiʿrāb* in Classical Arabic grammar: if a word is supplied with one ending only, it cannot be taken as a sign of *ʾiʿrāb*, since *ʾiʿrāb* means first of all the change of endings (cf. Ingham 1994:47–51). For a detailed study of the usage of *ʾiʿrāb* in the various registers of Modern Standard Arabic, see Aamer (1980), describing the usage in Cairo.

The special connection of *ʾiʿrāb* with Arabic is often underlined by the Arabic grammarians. Ibn Fāris (d. 395/1004), for instance, states that Arabic is the only language in which such a system of declensional endings governed by *ʿawāmil* exists (*Ṣāḥibī* 42.6). A similar view is found in the descriptions of other languages within the framework of Arabic grammar. Grammarians of Turkic, for instance, hesitate to apply the notion of *ʾiʿrāb* to the structure of Turkic and prefer to use instead the notion of *ʿalāma*, which in Arabic grammar usually means 'marker' of categories that are not affected by the action of an *ʿāmil* (Ermers 1999:163). Ermers (1999:286–290) explains this hesitation to use the concept of *ʾiʿrāb* by pointing out the difference between Arabic and Turkic morphology and syntax. The Turkic possessive construction, for instance, does not have the fixed relationship with an *ʿāmil* that exists in Arabic *ʾiʿrāb* (→ *ʾiḍāfa*).

Bibliographical references

Primary sources

ʾAbū ṭ-Ṭayyib, *Marātib* = ʾAbū ṭ-Ṭayyib al-Luġawī, *Marātib an-naḥwiyyīn*. Ed. Muḥammad ʾAbū l-Faḍl ʾIbrāhīm. 2nd ed. Cairo: Dār Nahḍa Miṣr li-ṭ-Ṭibāʿa wa-n-Našr, 1974.

Farrāʾ, Maʿānī = ʾAbū Zakariyyāʾ Yaḥyā ibn Ziyād al-Farrāʾ, Maʿānī l-Qurʾān. Ed. Muḥammad ʿAlī an-Najjār. 3 vols. Cairo: ad-Dār al-Miṣriyya, 1955–1972.

Ibn al-ʾAnbārī, *ʾInṣāf* = ʾAbū l-Barakāt ʿAbd ar-Raḥmān ibn Muḥammad al-ʾAnbārī, *Kitāb al-ʾinṣāf fī masāʾil al-xilāf bayna n-naḥwiyyīna l-baṣriyyīna wa-l-kūfiyyīn*. Ed. Gotthold Weil. Leiden: E.J. Brill, 1913.

Ibn Fāris, *Ṣāḥibī* = ʾAbū l-Ḥusayn ʾAḥmad Ibn Fāris, *aṣ-Ṣāḥibī fī fiqh al-luġa*. Ed. Moustafa Chouémi [Muṣṭafā aš-Šuʾaymī]. Beirut: A. Badran, 1964.

Ibn Jinnī, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. 3 vols. Ed. Muḥammad ʿAlī an-Najjār. Beirut: Dār al-Hudā, 2nd ed. N.d. (1st. ed., Cairo: Dār al-Kutub al-Miṣriyya, 1952–1956).

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. Ed. Hartwig Derenbourg. 2 vols. Paris: Imprimerie Nationale, 1881–1889.

Suyūṭī, *Tuḥfa* = Jalāl ad-Dīn ʾAbū l-Faḍl ʿAbd ar-Raḥmān ibn ʾAbī Bakr as-Suyūṭī, *at-Tuḥfa al-bahiyya wa-ṭ-ṭurfa aš-šahiyya*. Istanbul: Maṭbaʿat al-Jawāʾib, 1302/1882.

Yāzijī, *Faṣl* = Nāṣif al-Yāzijī, *Kitāb faṣl al-xiṭāb fī ʾuṣūl luġat al-ʾiʿrāb*. Beirut, 1834.

Zajjājī, *ʾĪḍāḥ* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān ibn ʾIsḥāq az-Zajjājī, *al-ʾĪḍāḥ fī ʿilal an-naḥw*. Ed. Māzin al-Mubārak. 3rd ed. Cairo: Dār an-Nafāʾis, 1979.

Zamaxšarī, *Mufaṣṣal* = ʾAbū l-Qāsim Maḥmūd ibn ʿUmar az-Zamaxšarī, *Kitāb al-mufaṣṣal fī n-naḥw*. Ed. Jens Peter Broch. Christiania: Libraria P.T. Mallingii, 1870.

——, *ʾUnmūḏaj* = ʾAbū l-Qāsim Maḥmūd ibn ʿUmar az-Zamaxšarī, *al-ʾUnmūḏaj fī n-naḥw*. Ed. Ḥusnī ʿAbd al-Jalīl Yūsuf. Cairo: Maktabat al-ʾĀdāb, 1990. [With the commentary of Jamāl ad-Dīn Muḥammad al-ʾArdabīlī.]

Zubaydī, *Ṭabaqāt* = ʾAbū Bakr Muḥammad ibn al-Ḥasan az-Zubaydī, *Ṭabaqāt an-naḥwiyyīna wa-l-luġawiyyīn*. Ed. Muḥammad ʾAbū l-Faḍl ʾIbrāhīm. Cairo: Dār al-Maʿārif, 1984.

Secondary sources

Aamer, Mahmoud. 1980. *Aussprachepraxis des Hocharabischen in Kairo: Phonetik, Phonologie, Behandlung des Iʿrāb, Einfluss des Dialekts auf die Morphologie*. Ph.D. diss., University of Erlangen. (Publ., Erlangen: Höfer und Limmert.)

ʾAsmar, Jirjis ʿĪsā al-. 1986. *Qāmūs al-ʾiʿrāb*. Beirut: Dār al-ʿIlm li-l-Malāyīn.

Baalbaki, Ramzi. 1990. "*Iʿrāb* and *bināʾ*: From linguistic reality to grammatical theory". *Studies in the history of Arabic grammar*, II, ed. Michael G. Carter and Kees Versteegh, 17–33. Amsterdam and Philadelphia: J. Benjamins.

Birkeland, Harris. 1940. *Altarabische Pausalformen*. Oslo: J. Dybwad.

Carter, Michael G. 1981. *Arab linguistics: An introductory classical text with translation and notes*. Amsterdam and Philadelphia: J. Benjamins.

Dévényi, Kinga. 1987–88. "*Muǧāwara*: A crack in the building of *ʾiʿrāb*". *Quaderni di Studi Arabi* 5–6.196–207.

Ermers, Robert. 1999. *Arabic grammars of Turkic: The Arabic linguistic model applied to foreign languages and translation of ʾAbū Ḥayyān al-ʾAndalusī's* Kitāb al-ʾidrāk li-lisān al-ʾatrāk. Leiden: E.J. Brill.

Ingham, Bruce. 1994. *Najdi Arabic: Central Arabian*. Amsterdam: J. Benjamins.

Merx, Adalbertus. 1889. *Historia artis grammaticae apud Syros*. Leipzig. (Repr., Nendeln: Kraus, 1966.)

Owens, Jonathan. 1990. *Early Arabic grammatical theory: Heterogeneity and standardization*. Amsterdam and Philadelphia: J. Benjamins.

Rabin, Chaim. 1969. "The structure of the Semitic system of case endings". *Proceedings of the International Conference on Semitic Studies held in Jerusalem, 19–23 July 1965*, 190–204. Jerusalem: Israel Academy of Sciences and Humanities.

Talmon, Rafael. 1997. *Arabic grammar in its formative age*: Kitāb al-ʿayn *and its attribution to Ḫalīl b. Aḥmad*. Leiden: E.J. Brill.

——. 2003. *Eighth-century Iraqi grammar: A critical exploration of pre-Ḫalīlian Arabic linguistics*. Winona Lake, Ind.: Eisenbrauns.

Versteegh, C.H.M. [Kees]. 1977. *Greek elements in Arabic linguistic thinking*. Leiden: E.J. Brill.

——. 1981. "A dissenting grammarian: Quṭrub on declension". *The history of linguistics in the Near East*, ed. Konrad Koerner, Hans-Josef Niederehe, and Kees Versteegh, 403–429. Amsterdam: J. Benjamins.

——. 1993. *Arabic grammar and Qurʾānic exegesis in early Islam*. Leiden: E.J. Brill.

——. 1995. *The explanation of linguistic causes: Az-Zağğāğī's theory of grammar*. Amsterdam and Philadelphia: J. Benjamins.

Vollers, Karl. 1906. *Volkssprache und Schriftsprache im alten Arabien.* Strasbourg.

Kinga Dévényi
(Corvinus University)

Iran

1. Introduction

New → Persian, including its most recent phase, Modern Persian, the official language of Iran and the main or second language of huge surrounding territories, has absorbed a large number of foreign words. These loanwords were borrowed from various northwestern and eastern Iranian dialects or from Western languages in the modern period, but the most effective and influential source was Arabic. Arabic loanwords constitute more than 50 percent of the contemporary Persian vocabulary, and in elevated styles it may exceed even 80 percent (Pisowicz 1985). The reason for the presence of such a staggering number of loanwords goes back to the earliest period of Arabicization of the country in the 7th century. After the fall of the Sasanian empire, Arabic was introduced as the main language of administration, just as happened in all conquered territories (about the period of the conquests and their effects in Persia, see Zarrīnkūb 1975). During the first centuries of the Islamic period, Arabic remained the dominant language in administration, religion, theology, science, and culture.

There was, however, a crucial difference between Persia and other Islamic provinces with respect to the fate of the language originally spoken there. In Persia, the indigenous aristocracy remained an important factor; Persian nobles (*dihqān*s) were put in charge of taxation and held on to their influential position, even to the point that they collected taxes from Arab settlers, for instance, in the province of Khurasan. At first, the language of the Sasanian empire, Middle Persian, retained its original position as the administrative language, but this changed in the course of the first two centuries of Islam, when it was replaced by Arabic. All in all, both socially and politically the indigenous culture was more prestigious in Persia than was the case in other provinces.

Persian also remained in use as one of the vernacular languages in the new cities of the eastern Islamic empire, and the sources contain many anecdotes about the Persian spoken in the streets of cities like Basra (e.g., Fück 1955:12–18; Pellat 1953:128). During this period, many → Persian loanwords were introduced into Arabic and became an accepted part of the Classical Arabic lexicon. Yet, despite this survival of the colloquial language, the cultural role of Persian was diminished by the dominant position of Arabic, and no doubt many Persians became bilingual in Arabic and Persian or even switched to Arabic entirely. Some of the most important scholars of Arabo-Islamic culture had Persian as their mother tongue, among them the first grammarian of Arabic, Sībawayhi (d. 177/793?), whose very name was Persian. But in spite of their Persian background, these scholars did not doubt the cultural and religious superiority of the Arabic language. Sībawayhi hardly ever mentions his native language in the *Kitāb*, and a later grammarian, al-Fārisī (d. 377/987), when asked about the Persian language, is said to have stated that Arabic surpassed it by far in elegance (*luṭf*; Ibn Jinnī, *Xaṣāʾiṣ* I, 243).

The spoken language of the Sasanian empire, Dari (*Pārsī-i Darī*) never went out of use completely, but it remained a colloquial language because of the adoption of Arabic as a cultural language, even by Iranian intellectuals (Lazard 1975). This changed with the emergence of Persian poetry in Dari, in particular through the success of Firdawsī's *Šāhnāma*. This New Persian language became the preferred language of the court of the newly independent Iranian dynasties like the Samanids (10th century). After the fall of Baghdad in 1258, Arabic lost its foothold in the eastern provinces, and eventually New Persian (*fārsī*) was adopted as the language of administration in the Seljuq period (11th century) and replaced Arabic entirely. This may have been caused in part by the rift between Shiʿis and Sunnis, and the introduction of Persian as the new language went hand in hand with the introduction of Shiʿite Islam as the 'national' religion of Iran (Lazard 1975). Because of the role Persian missionaries played in the spread of Islam to the East, Persian also became the main vehicle of the Islamic *daʿwā* in South Asia and Southeast Asia (→ Urdu/Hindustani; → Indonesian/Malay).

Obviously, the original shift in status of the Persian language also had an impact on the influx of Arabic loanwords, and this did not change when it was reinstated as the language of culture and religion. The earliest loanwords began to penetrate New Persian in the 9th/10th centuries (20–30 percent), but their proportion increased heavily in the subsequent 11th/12th centuries (ca. 50 percent) and continued until quite recently. The majority of Arabic loanwords were already incorporated into New Persian by the late 12th century and showed a remarkable stability until recently, as the statistics of Arabic loanwords based on sources of various styles, genres, and authors illustrate (see Moïnfar 1970; Skalmowski 1961; Koppe 1959; Utas 1978; Lazard 1965; Telegdi 1974; Ṣādiqī 1975; etc.). This large presence of Arabic loanwords demonstrates that, by reason of rational, aesthetic, or religious motivation, the Iranians' attitude toward the Arabic language may be characterized as hospitable, as early and subsequent sources of Persian literature, lexicography, and science testify. For instance, the author of the *Qābūsnāma* in the 11th century, while treating the art of letter writing, says: "Do not write pure Parsi, for it is unpleasant" (translation Lazard 1993:27), and Shams-i Faxrī, the famous lexicographer, writes in the 14th century: "The word can be of two [kinds]: a strange Arabic word and a pearl-like Darī". Further testimonies can also be found in the lexicographical literature, such as the *Farhang-i Jahāngīrī* in the 17th century, or Ḥabīb Iṣfahānī's *Dastūr-i suxan* in the 19th century (see Jeremiás 2003, Sec. II "The impact of Arabic"). Even though the proportion of Arabic loanwords may vary according to age, genre, social context, or even idiolect, a Modern or Classical Persian style entirely deprived of Arabic loans is almost impossible, despite the purists' reawakening movement over centuries (Koppe 1959:593).

This difference in treatment of loanwords can also be found in the attitude toward Arabic in modern Iran. In spite of the religious importance of Arabic, there is a certain ambivalence in the attitude toward the Arabs' claim about the uniqueness of the language. Even within the Iranian clergy there seem to be differences, some religious leaders preferring to use Persian words while others always use Arabic words for certain fundamental concepts. Surprisingly, the word for 'God' is always Persian *Xodā* rather than Arabic *Allāh*. Whether or not a choice for Persian words reflects a political or ideological attitude is difficult to say because no study has yet been made of the religious vocabulary. During the Iran-Iraq war, the friction between Arabs and Persians was not often alluded to in the propaganda, but the underlying tone often made clear that there was not much sympathy for the Arab monopoly on religious matters (Gieling 1999:85). Religious leaders in Iran resisted the opinion sometimes expressed in Iraqi war propaganda that Persians were not Muslims because the *Qurʾān* had been revealed in Arabic. In reply, the ayatollah Khomeini stressed the universal character of Islam, rather than playing on Iranian nationalist feelings.

Most common believers in Iran, in spite of compulsory education in Arabic, are unable to read Arabic texts, let alone to conduct a conversation in Arabic. Arabic is not taught as a living language but purely as a means to learn passages from the *Qurʾān* by heart (Ingham 1994:104). This applies even to the clergy, although there are exceptions, usually because of contacts with Shiʿite clerics in the holy cities of Karbala and Najaf in Iraq. Especially those among the religious establishment who had to flee the country during the shah's regime and stayed in Iraq or other Arabic-speaking countries are fluent in Arabic. In their Friday speeches, the religious leaders sometimes translate Qurʾānic verses into Persian (Gieling 1999:93).

The situation is even more complicated because of the presence of an Arabic-speaking minority within the boundaries of Iran. In the province of → Khuzistan (Arabic ʿArabistān), most people speak Arabic, but this dialect is not recognized as a minority language; to some extent, it is even suppressed because of a deep distrust on the part of the Iranian government, whether in the period of the shah or after the Islamic Revolution, toward the Arab neighbors. Because of the large-scale immigration of Persian families, the larger towns in Khuzistan have now become Persian-speaking, but in the countryside many people still speak (and dress) as Arabs (Ingham 1994:107–108). The younger generation grows up learning

Persian and speaking Arabic only as a colloquial language, with little prestige. Elsewhere in Iran, Arabic is spoken on the coast of the Persian Gulf in the region of Fars, and possibly also in the area of Gumbad-i Qabus in Khurasan (Ingham 1994:106), but not much information is available on these areas. Speakers of Arabic on the coast have regular trade relations with speakers of Arabic on the other side of the Gulf (Ingham 1994:112), which makes it easier for them to hold on to their language.

2. Arabic linguistic influence

Whatever the attitude toward Arabic, the influence of the Arabic language has always been a crucial factor in Modern Persian. The following section deals with its influence on script, phonology, and morphosyntax, with special attention to the undeniable register differences in each domain.

2.1 *Script*

The linguistic influence of Arabic is most clearly detectable in the vocabulary of Persian, due to the huge number of Arabic lexical items, and somewhat less so in phonology and morphosyntax. But these are not the only fields where the impact of Arabic can be felt. The first step was obviously the script (→ Arabic alphabet for other languages). After the Islamic conquest, New Persian, including its ancient and modern varieties (except → Tajik in the 20th century) in the larger Persian-speaking area (Iran, Afghanistan, and Central Asia), began to use a slightly modified Arabic script. This script has 32 letters, 28 taken over from Arabic and 4 new letters supplied with three dots to denote Persian phonemes that are missing in Arabic: *p* پ, *č* چ, *ž* ژ, and *g* گ = archaic ک (here, the letters *g* and *k* could have the independent form with *hamza* in it). However, there is no one-to-one correspondence between letters and consonantal phonemes in Persian because not all the Arabic letters distinguish phonemes in Persian; some of them were introduced mainly, though not exclusively, through Arabic loanwords, and (probably) gradually assimilated to the pronunciation of the Persian phonemes (see ذ, ض, ظ = /z/, ث, ص = /s/, ط = /t/, غ, ق = /ġ/ or /q/, ح = /h/, ع, ء = /ʾ/ or zero. Some of these letters (ذ, ث, ص, ط, غ, ق) do occur in words of Persian origin as well as reflexes of earlier dialects (e.g. پذیرفتن *paḏīruftan* 'to accept'), or they are due to the unfixed orthography (e.g. *Tihrān* ~ *Ṭihrān*). The new Perso-Arabic script as it is used today gained ground gradually by the 12th century, but some intricate problems of orthography have remained unsettled until recently (for details, see Horn 1898–1901: I/2, 12; Lazard 1963, *passim*; Meier 1981:71; Hashabeiky 2005, *passim*). These difficulties derived mainly from the typological difference between the two languages: the Arabic script, being a Semitic alphabet, consists only of consonantal signs (*ḥurūf*), and the vowels are represented only partially and variously, that is, by certain consonantal letters (*ʾalif*, *yāʾ*, *wāw*), which serve to denote long vowels and vowel sequences called diphthongs, by orthographic devices (*hamza*, *ʾalif*, etc.), and by optional superscript signs (*fatḥa, ḍamma, kasra*, etc.). The Iranians, however, whose language is of an Indo-European type, in which the vowels and consonants are of equal value, introduced some innovations into the Semitic system – in addition to the newly invented letters – in order to adapt the script to the characteristics of the Persian language. For instance, the system of denoting word-final short vowels had gradually developed (partly following an already existing Arabic tradition), and word-final *-a, -u* (Modern Persian *-e, -o*) came to be written by the so-called silent letters *hāʾ* or *wāw* (*bayān-i ḥaraka*), e.g. <nʾmh> = *nāma* (Modern Persian *nāme*) 'book', <tw> = *tu* (Modern Persian *to*) 'you' (for more details, see Jeremiás 2003, V "Scripts").

2.2 *Phonology*

Because of the impact of Arabic loanwords, the phonological inventory of Classical Persian was augmented with new phonemes as compared to Middle Persian. The most characteristic new phoneme is the glottal stop, which originated in two separate Arabic phonemes /ʾ/ and /ʿ/, represented by the sign *hamza* and the letter *ʿayn*. In Modern Persian, the pronunciation of the Arabic-voiced pharyngeal /ʿ/merged entirely with the glottal stop /ʾ/ (their phonemic status cannot be stated with certainty for the earlier stages), and its occurrence as a separate phoneme is restricted by both phonetic and social factors: it is pronounced only in certain medial and final positions before or after consonants in careful speech (e.g. *maʿlūm* > Modern Persian /maʾlum/ ‘known’; *al-ʾān* > Modern Persian /alʾān/ ‘now’) or, rarely, in intervocalic position in place of an etymological *ʿayn* (e.g. *sāʿat* > /sāʾat/ ‘hour’). In Modern Persian, however, there is a non-phonemic indigenous glottal articulation in a vocalic onset after a pause or between two consecutive vowels, represented by the signs *hamza* or *ʾalif* , which may have helped the incorporation of the glottal stop into Persian (e.g. أمروز *imrūz* > Modern Persian /emruz/ > [ʔɛmruz] ‘today’, پائیز (or پاییز) *pāīz* > Modern Persian /pāiz/ > [paːʔiz] ‘autumn’). The appearance or disappearance of this phoneme with a compensatory lengthening of the preceding vowel or consonant (e.g. *maʿlūm* > Modern Persian /maʾlum/ > [maːlum]) or its substitution with hiatus or an intrusive element such as *y*, or rarely *h* and *w* (e.g. Modern Persian /pāiz/ > [paːʔiz, paːjiz] ‘autumn’ greatly varies according to register (see Pisowicz 1985, *passim*; Windfuhr 1979:139; Lazard 1992, *passim*; Jahani 2005). In the spoken register, the glottal stop as a phoneme does not occur at all (Lazard 1992:12). Like the glottal stop, the uvular voiceless plosive /q/ became a separate phoneme in Classical Persian under the influence of Arabic (see its development in Modern Persian in Pisowicz 1985:42–47; Ṣādiqī 1975).

2.3 *Morphosyntax*

Through loanwords, some grammatical elements of Arabic were also transmitted into Persian, especially in nominal morphology, such as regular and broken plurals, *ʾiḍāfa*-structures, feminine gender, and gender agreement, but also in the creation of compounds and derivations. However, it is hard to decide whether these phenomena pass beyond pure lexical borrowings to function as creative linguistic rules that generate Arabicized structures. Occasionally, the latter could be the case, but the majority of constructions containing Arabic linguistic elements appear to involve lexical borrowing that has become integrated into the morphology of Persian to various extents (see Sec. 2.3.1, regarding plural formation of the noun). It is also true that the Arabic component of Persian has shown a great variety both historically and stylistically. Certain hybrid constructions containing both Arabic and Persian elements are used mainly, but not exclusively, in highly elevated style (occasionally they can evoke humorous or satiric effects), while others occur in the colloquial register in most cases. The productivity of the rules imported from Arabic is a matter of debate (Windfuhr 1979:80–83).

2.3.1 Plural formation

The plural formation of nouns exhibits the characteristic linguistic contribution brought about by Arabic loans. New Persian (Classical and Modern) uses two kinds of Persian plural markers: *-hā* (with all kinds of nouns) and *-ān* (with animate beings). In addition, many Arabic loans have preserved their original plural formation and, moreover, extended their use to words of Persian origin (sometimes in Persian garb as *-jāt*). This steadiness in use and the creativity of Arabic plural formation in producing doublets, sometimes with differing meanings, shows the deep and effective impact of Arabic, even though some of the Arabic plural forms were substituted by Persian plural formations or were kept as alternative (formal or colloquial) stylistic variants throughout the different periods of Persian. Among the types of Arabic plurals there are nouns with regular plural endings, such as *-āt, (-jāt), -(ī)yāt, -īn, -yūn* (e.g. *kalima* ‘word’, pl. *kalimāt; intixāb* ‘election’, pl. *intixāb-āt; rubāʿī* ‘tetrastich’, pl. *rubāʿī-yāt; musāfir* ‘traveler’, pl. *musāfir-īn; inqilābī* ‘revolutionary’, pl. *inqilābī-yūn*), or nouns with irregular, broken plural (e.g. *kitāb* ‘book’, pl. *kutub*; *vazīr* ‘minister’, pl. *vuzarā,* but also *vazīr-hā, vazīr-ān*). From among the plural markers, the morpheme *-āt* appears to be mostly accepted by Persian speakers: in the post-Classical formal language it began to appear with non-Arabic words as well (e.g. *farmāyiš* ‘order’, pl. *farmāyiš-āt; dih* ‘village’, pl. *dih-āt* ‘villages’ [see also *dihātī* ‘villager’]; *tiligrāf/tilgirāf* ‘telegraph’, pl. *tiligrāf-āt*), but some of them are already attested in early sources (e.g. *bāġāt* ‘gardens’ in the *Safarnāma* of Nāṣir-i Xusraw in the 11th century). Its variant *-jāt* is used with Persian words (e.g. *mīva-jāt* ‘fruits’), but there are some occurrences with words of Arabic origin (e.g. *qalʿa-jāt* ‘fortresses’ vs. *qalʿa-hā, qilāʿ*) as well. The acceptance of the latter formation is disputed among men of letters: some allow the formation (*-jāt*) only with Persian words with a final vowel, while others accept it only with foreign words. Modern writers have started, however, to use such plurals regularly. There are some Persian words that can be used with a double plural formation (Arabic *-(j)āt* and Persian *-hā* or *-ān*), and the nouns so produced can exhibit a semantic difference of genus (*jins*) vs. species (*nawʿ*), e.g. *rūznāma-jāt* 'the press, newspaper' vs. *rūznāma-hā* 'the various individual copies of newspapers'. Sometimes, these doublets represent only various registers, e.g. *farmāyiš-āt* (vulgar) vs. *farmāyiš-hā* (cf. Haïm 1988, s.v.; but Ṣadrī Afšār a.o. 2002, s.v., give the two plural forms without register specification), or *dihāt-ī* 'villagers' (vulgar; cf. Dihxudā 1999, s.v.). The plural marker of Arabic origin *-īyāt* is attached to denote a special literary genre in most cases, e.g. *pand* 'advice', pl. *pand-īyāt* 'moral poems'. Broken plurals were widely used in Classical Persian and have been retained in modern formal style. Some words occurring with both regular Persian and Arabic (broken) plurals may display a semantic difference, according to some linguistic sources, e.g. *ḥarf* 'letter, particle, word, etc.', pl. *ḥurūf* 'letters', *ḥarf-hā* 'particles', etc. (for different views see Ṣadrī Afšār a.o. 2002, s.v.; Anwarī and Aḥmadī Gīwī 1999:II, 4), while others appear to signal only various registers, e.g. *musāfir-īn*, *musāfir-hā*, *musāfir-ān* 'travelers'; *ʿamala* 'workman', pl. *ʿamala-hā, ʿamala-jāt*. Sometimes, Arabic broken plurals are further expanded by Persian plural markers, combining the two types of plural formation; these occur in very early sources of prose and poetry, e.g. *malik* 'king', pl. *mulūk, mulūk-ān* in *Tarjuma-i tafsīr-i Ṭabarī*; *kawkab* 'star', pl. *kawākib, kawākib-hā* in Farruxī (see Anwarī and Aḥmadī Gīwī 1999:II, 341 *jamʿ-i jamʿ*). In the latter case, the broken plural may be regarded as a singular noun (see Dihxudā 1999, s.v. *aṭrāf*; but modern sources like Haïm 1988 and Ṣadrī Afšār a.o. 2002 treat *aṭrāf* as a plural). In the classical period, certain Persian words were reborrowed from Arabic with their broken plurals modeled on patterns of Arabic morphology (where they remain in use to date), such as *farmān* 'order' and *farāmīn* 'orders'; *bustān* 'garden' and *basātīn* 'gardens', etc. (see more details in Faršīdward 1988:151; Anwarī and Aḥmadī Gīwī 1999:II, 87–93; Lazard 1989:60–62).

2.3.2 Concordance

Gender is not marked in Persian morphologically, but via the transference of an Arabic syntactic rule of concord (*muṭābiqa(t)-i ṣifat wa mawṣūf*), a certain type of *ʾiḍāfa*-structure (adjectival modification) began to appear in early prose from the 12th century onward (e.g. in the *Čahār maqāla* of Niẓāmī ʿArūżī (Arabic ʿArūḍī), where a feminine noun or a broken plural was followed by an adjective supplied with the ending of the Arabic feminine *-at* (> Classical Persian *-a*, Modern Persian *-e*) in Persian (e.g. *mulūk-i māżiya* 'past kings'; see Bahār 1958:II, 306–307). This construction occurred especially in the terminology of science (e.g. *quwwa-i ḥāfiẓa* 'retentive power', *quwwa-i darrāka* 'perceptive power', etc.), but structures of Persian type – without this gender concordance – were most commonly used (e.g. *ḥawās-i bāṭin* 'internal senses'). In Modern Persian, however, these kinds of construction have begun to occur again with great frequency as translations of Western political terms (e.g. *umūr-i xārija* 'foreign affairs'), used in parallel with constructions of Persian type (e.g. *umūr-i ijtimāʿī* 'social affairs'; see Faršīdward 1988:24, 155; Perry 1991:16). These adjectival phrases with the Arabic feminine concord rule seem to have been regarded as directly borrowed lexical units from Arabic rather than generated by a creative syntactic rule. On the other hand, the two constituents of the phrase are connected through a Persian *iḍāfa*-structure which signals their embedding into Persian morphosyntax, even though it has a marginal significance, due to its rare occurrence.

2.4 *Word-formation (derivation and compounding)*

2.4.1 Derivation

Some characteristic features of word-formations through derivation are the following. Nouns (abstract nouns in most cases) can be derived from words of Arabic or Persian origin with certain patterns or suffixes (sometimes with an uncertainty in vocalization) that do not exist in Arabic or, if they do exist, now convey a new meaning (e.g. *ḥifāẓat* 'care, custody', *dixālat/daxālat* 'interference; participation', *xijālat/xajālat* 'shame', etc.). The ending *-īyat* creates abstract nouns from both Arabic (e.g. *jamʿ*

'united [adj.]' ⇒ *jamʿīyat* 'crowd'; see Telegdi 1962:330, n. 15; Perry 1991:25–26) and Persian words (e.g. *mard* 'man' ⇒ *mard-īyat* 'manliness [rare]'). Curiously, some Persian words create derivatives with certain patterns of Arabic morphology in addition to their regular Persian plural forms (e.g. *nāzuk* 'tender' ⇒ *nāzukī* 'tenderness' or *nazākat* 'id.'; see Dihxudā 1999; Haïm 1988; Ṣadrī Afšār a.o. 2002; Faršīdward 1988 ss.vv.; *nard* 'backgammon', *narrād* 'backgammon player'; *muhr* 'seal', *mamhūr* 'sealed'; see more details in Faršīdward 1988:151–152, etc.). Some comparatives are derived by a double formation, and occasionally, the Arabic elative forms are further extended by the Persian suffixes of comparative degree *-tar* (e.g. *awlā-tar* 'prior, better', *aʿlam-tar* 'more/most learned [very formal]', *muhim(m)* 'important' ⇒ *aham(m)* 'more/most important' or *muhim(m)-tar* 'more important', *muhim(m)-tarīn* 'most important'; see Jazayery 1977:133). The abstract words and comparatives of the above formation mostly belong to a very formal style. Besides, one of the oldest ways of deriving Persian verbs was by way of adding the suffix *-īdan, -āndan* to Arabic *ism*-forms, e.g. *fahm* 'comprehension' ⇒ *fahm-īdan* 'to comprehend' or *fahm-āndan* 'to cause to comprehend [causative]'. However, this synthetic method was gradually replaced by the more popular analytical method of creating new verbal forms by compounding (see below, Sec. 2.4.2).

The Arabic accusative forms with *tanwīn,* generally used in Persian as adverbs, belong to the oldest layer of Arabic loans (e.g. *ʿamd-an* 'intentionally', *ḥaqq-an* 'justly'). Their proliferation began from the 12th century onward, and they integrated into Persian vocabulary to the extent that Persian words (and also Western loans) were derived analogically by the suffixation of the morpheme *-an* (e.g. *jān-an* 'wholeheartedly', *nāčār-an* 'helplessly', in addition to the adverbs of Arabic origin that are not used in Arabic (e.g. *akṯar-an* 'mostly'). In Modern Persian they are widely used in every register (e.g. *taqrīb-an* 'approximately', *kāmil-an* 'completely'; see Faršīdward 1988:152–154), even though these forms are not recommended by purists.

2.4.2 Compounding (word-compound, phrasal-compound)

Compounding was one of the most developed procedures of enlarging vocabulary with Arabic loans in New Persian. Its capability for forming compounds made the indigenous grammarians consider Persian to be a 'compounding language' (*zabān-i tarkībī*) and Arabic a 'derivational language' (*zabān-i ištiqāqī*). This process created word-compounds and phrasal-compounds by combining two (or, rarely, more) base forms (also called free morphemes), but the distinction between them cannot always be established clearly. These compounds are of a heterogeneous nature, representing various types according to the syntactic and semantic relations between their constituents, according to the word class of the constituents, and according to the word class produced by composition. Constituents of Arabic origin appear in every kind of old and new compounds from the early beginning of New Persian, substituting Arabic loans for the Persian elements of old compounds or creating new compounds modeled on Persian patterns. There are various types of compounds (nouns or adjectives in most cases), where the first or the second constituent or sometimes both constituents are of Arabic origin. Some compounds are made up of two nominal parts, e.g. *dawlat-xāna* 'wealth [Arabic]' + 'house [Persian]' ⇒ '(your) house [polite]', *kutāh-naẓar* 'short [Persian]' + 'mind [Arabic]' ⇒ 'narrow-minded', *kam-ʿaql* 'few [Persian]' + 'mind [Arabic]' ⇒ 'foolish', *nāqis-ʿaql* 'defective [Arabic]' + 'mind [Arabic]' ⇒ 'foolish'. The latter represents the Persianized form of the well-known Arabic structure called *ʾiḍāfa taqdīriyya* (e.g. *nāqiṣ ul-ʿaql*), which abounded in Persian formal style (e.g. *zāyid ul-waṣf* 'beyond description').

Another common type of verbal compound consists of an Arabic *ism* followed by a Persian verbal stem (past or present) whose relation is equivalent to a verb and an object (or another verbal complement); the compounds so produced can have active or passive (if the second constituent is the past stem) meaning. The Persian verb-stems that can appear in such compounds are *ālūd-* (⇒ *ālūdan* 'to soil'), *šinās-* (⇒ *šināxtan* 'to know'), *ārā(y)-* (⇒ *ārāstan* 'to decorate, adorn'), *gūy-* (⇒ *guftan* 'to say'), *xʷāh-* (⇒ *xʷāstan* 'to wish'), etc., e.g. *xāk-ālūd* or *xāk-ālūda* 'earth-soiled ⇒ earthy, dusty', *ḥaqq-šinās* 'justice (duty)-knowing ⇒ grateful', *ʿālam- ārā(y)* 'world-adorning', *sajʿ-gū(y)* *ʿsaj* '-saying' ⇒ 'who uses rhyme in prose' ⇒ 'rhymer', *maṣlaḥat-xʷāh*, 'benevolence-seeking' ⇒ 'well-wishing'. The old type of compounds that combined an adjective and a noun produced an adejctival compound, such as *xujasta-pāy* or *xujasta-qadam* 'blessed [Persian]' + 'foot(step) [Persian or Arabic]' ⇒ 'of blessed foot(step)' (see Utas 1978:138). This type shows the possibility of selecting either a Persian or an Arabic word for constructing the same compound (see Lazard 1992:261–292).

The phrasal compounds originating as syntactic phrases are not simple loanwords: sometimes it is hard to establish the distinction between regularly used syntactic constructions held together by stress and intonation and lexicalized, 'fossilized' compounds. The compound verbs (also called 'verbal phrases') constitute the largest group of these phrasal compounds, which may appear as single lexemes or occasionally as syntactic groups. Their behavior differs according to historical stages, stylistic levels, dialects, or idiolects. Historically, the possibility of constructing this type of compound facilitated the infiltration of the majority of Arabic loans from the first centuries of the Islamic period. This heterogeneous group, including very peculiar verbal constructions, consists of two or more lexemes combined with the 'base' verb. The most common type is formed with verbs of exclusively Persian origin that serve as a kind of auxiliary, such as *kardan* 'to do', *zadan* 'to cut', *dādan* 'to give', *dāštan* 'to have', etc., following a nominal form of Arabic origin in most cases (noun, participle, adjective, etc.), e.g. *ḥarf zadan* 'to speak', *ṣuḥbat kardan* 'to talk' (or 'to keep company [archaic]'), *najāt kardan* 'to save' or *najāt dādan* 'id.', *xalāṣ šudan* 'to be rescued', *xalāṣ kardan* 'to rescue', *šurūʿ šudan* 'to begin [intrans.]', *šurūʿ kardan* 'to begin [trans.]', *maḥfūẓ dāštan* 'to protect'. As these examples show, the selection of the verbal constituents may express various verbal categories, such as voice (active or passive), or many other nuances of meaning. The instability of these phrasal compounds appears clearly in cases where their constituents can be broken up, and the phrasal compounds are constructed as syntactic phrases. In such cases, the nominal part may be followed by certain (but not all) grammatical morphemes of the nominal inflection, such as the indefinite article *-ī* (e.g. *ḥarf-ī zad* 'he said (s.th.)', adjective modifiers such as noun, adjective, pronoun, etc. (e.g. *xwāja…bisyār dar najāt-i mardum kard* 'X… made (*kard*) a lot (*bisyār*) for the salvation (*najāt*) of (*-i*) people (*mardum*)'), or verbal complements such as direct or indirect objects (e.g. *xalāṣ-am kun* 'rescue me (*-am*)'). The variations of such verbal phrases (either as single lexical items or as syntactical phrases) are almost unlimited in Persian, and they may occur in every possible register, irrespective of historical periods. Yet, some of them seem to belong exclusively to the most elevated speech, while others belong to the colloquial style. Moreover, a tendency is felt in Modern Persian to substitute verbal phrases made of Arabic nouns and Persian auxiliaries for single Persian verbs of ancient origin (e.g. *paẕīruftan* (Persian) ⇒ *qabūl* (Arabic) *kardan* (Persian) 'to accept', especially in the modern standard. Nearly all old Persian verbs have similar synonyms constructed from an Arabic noun and a Persian auxiliary (see Jazayery 1977:129).

Other phrasal compounds of heterogeneous nature, which appear rather as stock phrases, occur abundantly from an early date, such as prepositional phrases (*fawq ul-ʿāda* 'unusual', *fī l-jumla* or *bi-l-jumla* 'in short', *bayn ul-milal* [modern *beynolmelal*] 'international' or *bayn ul-milal-ī* 'id.', which is extended by the Persian adjective suffix); determinative juxtaposed constructions (*qābil-i taḥammul* 'bearable' or *ġayr-i qābil-i taḥammul* 'unbearable', *ʿadam-i muwāfiqat* [archaic *muwāfaqat*] 'disagreement'); rank-expressions (*amīr ul-umarāʾ* 'the prince of princes' ⇒ 'prime minister'); and even whole sentences used as single lexical units, nouns or adverbs in most cases (e.g. *mā jarā* 'what happened' ⇒ 'event, accident', *mā yaḥtāj* 'what is needed' ⇒ 'necessaries', *kamā huwa ḥaqquhu* 'as ought to' ⇒ 'duly', e.g. *man ū-rā kamā huwa ḥaqquhu na-šināxtam* 'I did not recognize him as I should' (Haïm 1988 s.v.). Historically, some phrasal compounds changed both constituents or the type of constructions, supposedly in order to bring the phrase closer to Persian context, e.g., *qarīb* 'near, close' could be used either with an *ʾiḍāfa* construction as *qarīb-i* or with a preposition *qarīb ba-* meaning 'approximately'.

There are peculiar examples of calque-like phrases in Persian in which the borrowed Arabic lexeme (a participle used as an adjective or a prepositional phrase, in most cases) requires an obligatory prepositional extension and the Persian phrase employs the Persian equivalent of the original Arabic preposition, e.g. *ʿalā mūjib* (Arabic) ⇒ *bar* (Persian archaic) *mūjib-i* or *ba* (Persian) *mūjib-i* 'according to, by virtue of'.

The types and examples listed here do not cover all the possible variations of smaller or larger units of compounds constructed wholly or partly out of Arabic constituents. There was practically no limit to creating such constructions, including even such strange phrases as *aʿam(m) az mard-u zan* ‘both men and women’, or mixed constructions of Classical Persian literary sources, which reflected the underlying Arabic structures truly, such as *baʿżī* [*az īn murakkabāt*] *muxtalaf fīhi bāšad* ‘some [of these compounds] are disputed’, in which *muxtalaf fīhi* means ‘[what is] disputed upon’ (13th century).

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCE

Ibn Jinnī, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. Ed. Muḥammad ʿAlī an-Najjār. 3 vols. Cairo, 1952–1856. (Repr., Beirut: Dār al-Hudā, n.d.)

SECONDARY SOURCES

Anwarī, Ḥ. and Aḥmadī Gīwī. 1378/1999. *Dastūr-i zabān-i fārsī*. 2nd rev. ed. 17th print. 2 vols. Teheran.

Bahār, M.T. 1337/1958. *Sabk-šināsī yā tārīx-i taṣaw-wur-i naṣr-i fārsī*. 2nd ed. 3 vols. Teheran.

Dihxudā, ʿA.A. 1999. *Luġatnāma*, eds. M. Muʿīn and J. Šahīdī. Repr. Teheran.

Faršīdward, Kh. 1367/1988. *ʿArabī dar fārsī*. 5th ed. Teheran.

Fück, Johann. 1955. *ʿArabīya: Recherches sur l’histoire de la langue et du style arabe*. Trans. Claude Denizeau. Paris: M. Didier.

Gieling, Saskia. 1999. *Religion and war in revolutionary Iran*. London and New York: I.B. Tauris.

Haïm, Soleiman. 1367/1988. *New Persian-English dictionary*. 7th ed. 2 vols. Teheran.

Hashabeiky, Forogh. 2005. *Persian orthography: Modification or change-over*. Ph.D. diss., University of Uppsala.

Horn, Paul. 1898–1901. “Neupersische Schriftsprache”. *Grundriss der iranischen Philologie*, I/2, ed. Wilhelm Geiger and Ernst Kuhn, 1–200. Strassburg: K. Trübner.

Ingham, Bruce. 1994. “The Arabic language in Iran”. *Indian Journal of Applied Linguistics* 20.103–116.

Jahani, Carina. 2005. “The glottal plosive: A phoneme in spoken Modern Persian or not?”. *Linguistic convergence and areal diffusion: Case studies from Iranian, Semitic, and Turkic*, ed. Éva Ágnes Csató, Bo Isaksson, and Carina Jahani, 79–96. London: RoutledgeCurzon.

Jazayery, Mohammad Ali. 1977. “Some aspects of lexical synonymy in contemporary Persian”. *Studies in historical and descriptive linguistics: Festschrift for Winfred P. Lehmann*, ed. Paul J. Hopper, 125–142. Amsterdam: J. Benjamins.

Jeremiás, Éva M. 2003. “New Persian”. *Encyclopaedia of Islam*, 2nd ed. *Supplementum*, Fascicules 7–8, *Iranian languages* III, 426–448. Leiden: E.J. Brill.

Koppe, R. 1959. “Statistik und Semantik der arabischen Lehnwörter in der Sprache ʿAlawī’s”. *Wissenschaftliche Zeitschrift der Humboldt-Universität zu Berlin* 9.585–619.

Lazard, Gilbert. 1963. *La langue des plus anciens monuments de la prose persane*. Paris: Klincksieck.

——. 1965. “Les emprunts arabes dans la prose persane du X^{e} au XII^{e} siècle: Aperçu statistique”. *Revue de l’Ecole Nationale des Langues Orientales* 2.53–67.

——. 1975. “The rise of the new Persian language”. *Cambridge history of Iran*. IV. *The period from the Arab invasion to the Saljuqs*, ed. R.N. Frye, 566–594. Cambridge: Cambridge University Press.

——. 1989. “Le persan”. *Compendium linguarum iranicarum*, ed. Rüdiger Schmitt, 263–293. Wiesbaden: L. Reichert.

——. 1992. *A grammar of contemporary Persian*. Trans. Shirley A. Lyon. Costa Mesa, Calif.: Mazda.

——. 1993. *The origins of literary Persian*. Bethesda, Md.: Foundation of Iranian Studies.

Meier, Fritz. 1981. “Ausssprachefragen des älteren Neupersisch”. *Oriens* 27–28.70–176.

Moïnfar, Mohammad Djafar. 1970. *Le vocabulaire arabe dans le Livre des Rois de Firdausī*. Wiesbaden: O. Harrassowitz.

Pellat, Charles. 1953. *Le milieu baṣrien et la formation de Ǧāḥiẓ*. Paris: Librairie d’Amérique et d’Orient A. Maisonneuve.

Perry, John R. 1991. *Form and meaning in Persian vocabulary*. Costa Mesa, Calif.: Mazda.

Pisowicz, Andrzej. 1985. *Origins of the New and Middle Persian phonological systems*. Cracow: Nakład Uniwersyteta Jagiellońskiego.

Ṣādiqī, ʿA.A. 1975. “L’influence de l’arabe sur le système phonologique du persan”. *La linguistique* 11:2.145–152.

——. 1986. “Arabic elements in Persian”. *Encyclopaedia Iranica*, ed. Ehsan Yarshater, II, 229–231. Columbia: Columbia University Center for Iranian Studies.

Ṣadrī Afšār, Gh.Ḥ., Nasrī Ḥakamī, and Nastaran Ḥakamī. 1381/2002. *Farhang-i muʿāṣir-i fārsī-i imrūz*. 2nd rev. ed. Teheran.

Skalmowski, W. 1961. “Ein Beitrag zur Statistik der arabischen Lehnwörter im Persischen”. *Folia Orientalia* 3.171–175.

Telegdi, Zsigmund. 1962. “Zur Unterscheidung von Substantiv und Adjektiv im Neupersischen”. *Acta Orientalia Academiae Scientiarum Hungaricae* 15.325–336.

——. 1974. “Remarques sur les emprunts arabes en persan”. *Acta Iranica* 2.337–345.

Utas, Bo. 1978. *A Persian Sufi poem: Vocabulary and terminology, concordance, frequency word-list, statistical survey, Arabic loan-words and Sufi religious terminology in* Tariq ut-Tahqiq *(A.H. 744)*. (= *Scandinavian Institute of Asian Studies Monograph Series*, 36.) London and Malmö: Curzon Press.

Versteegh, Kees. 1997. *The Arabic language*. Edinburgh: Edinburgh University Press. (2nd ed., 2001.) [“Arabic in Iran”, pp. 232–234.]

Windfuhr, Gernot L. 1979. *Persian grammar: History and state of its study*. The Hague and Paris: Mouton.

Zarrīnkūb, ʿAbd al-Ḥusain. 1975. "The Arab conquest of Iran and its aftermath". *Cambridge history of Iran*. IV. *The period from the Arab invasion to the Saljuqs*, ed. R.N. Frye, 1–56. Cambridge: Cambridge University Press.

Éva M. Jeremiás
(Eötvös Loránd University)

Iranian Arabic → Khuzestan Arabic

Iraq

1. General

Arabic is the majority language of Iraq, the most important minority language being Kurdish; lesser minority languages are Neo-Aramaic (in its two main varieties, Chaldean and Ashuri) and Turkoman. Geographically, Arabic is mostly found in the lower-lying regions of Iraq, while the minority languages are predominantly found in the mountainous regions of the north and northeast (i.e. Iraqi Kurdistan). As in all Arab states, a situation of → diglossia obtains. Modern Standard Arabic is the official language and the language of instruction and the media, while in everyday life the various dialects are spoken. The use of dialects for literary purposes is rare (e.g. vernacular poetry of the Bedouin type), and dialect texts are rarely committed to writing. There is no tradition of written dialect literature.

Arabic in Iraq is spoken by Muslims (Sunnis and Shiʿis) and by the religious minorities of the Christians and Jews; in the area of Kurdistan, however, the Christians and, to a lesser degree, the Jews speak Aramaic rather than Arabic. In the vicinity of Mosul, there are some Arabic-speaking Yezidi villages. During the second half of the 20th century the Christian and Jewish populations of Iraq were greatly reduced. The Jewish population has emigrated to Israel since the 1950s, and today there are no Jews reported to be still living in Iraq. Likewise, the Christian population has been greatly diminished by emigration.

The Arabic dialects spoken in Iraq fall into two main categories, called *qəltu* and *gələt*. This division was first described by Haim Blanc in his well-known book *Communal dialects in Baghdad* (1964). Blanc coined the terms *qəltu* and *gələt* (for typographical reasons he used the symbol *e* instead of *ə*). The key words *qəltu* and *gələt* (both harking back to Old Arabic *qultu* 'I said') encapsulate two distinguishing linguistic features: the treatment of the Old Arabic uvular stop *q* and the shape of the inflectional suffix of the 1st person singular perfect. The linguistic division is, following Blanc's formulation, partly regional, partly social. While Jews and Christians speak *qəltu* dialects, and nomadic, sedentarized nomadic, and Bedouinized populations speak *gələt* dialects everywhere, the dialects of the sedentary Muslim population (mostly city dwellers) follow a geographical pattern of distribution: north of the Samarra-Fallujah line *qəltu* dialects prevail, and south of the same line *gələt* dialects prevail. In the city of Baghdad, there was a unique linguistic situation, with Muslims speaking a *gələt* dialect and Christians and Jews speaking two different *qəltu* dialects, dubbed Muslim, Christian, and Jewish Baghdadi (→ Baghdad Arabic; → Baghdad Arabic, Jewish). The linguistic panorama that formed the subject of Blanc's book has since been disturbed by the emigration of religious minorities.

Neither *qəltu* nor *gələt* dialects are confined to the political borders of Iraq. *Qəltu* dialects are also spoken in southeastern Anatolia (→ Anatolian Arabic) and, to a lesser extent, in north-eastern Syria (Der izZor, Khatuniya); they share some significant features with → Uzbekistan Arabic (Jastrow 1997, 1998). *gələt* dialects, on the other hand, extend into Kuwait, the Persian Gulf, the Iranian province of Khuzestan (→ Khuzestan Arabic) and, again, northeastern Syria; more generally, they are akin to the Bedouin dialects of the Fertile Crescent and the Arabian Peninsula.

The history of Arabic in Iraq that led to the intricate present-day dialect distribution has best been summed up by Blanc (1964:168ff.). According to this view, the *qəltu* dialects continue the Arabic vernacular spoken in Abbasid Iraq, whereas the *gələt* dialects penetrated into Iraq only after the Mongol raids, due to a re-Bedouinization of central and southern Iraq, with a subsequent sedentarization of the Bedouin in rural settlements. In the Ottoman period, the influx of rural population into the growing towns led to a Bedouinization of the urban dialects as well. Only Christians and Jews who did not mix socially with the Muslims retained their older *qəltu*-type dialects.

The state of research on Iraqi Arabic is rather unbalanced. Interestingly, the dialects of the Christian and Jewish minorities have received more attention than those of the Sunni and Shiʿi population (cf. Abu-Haidar 1991; Blanc 1964; Jastrow 1989, 1990a, 1990b, 1991, 1993; Khan 1997; Mansour 1991). This is due to the fact that the Jews and part of the Christian population have left Iraq, and thus have been easier to access. A noteworthy exception is the majority dialect of the capital, Baghdad (Muslim Baghdadi), for which a number of excellent books, including grammars and dictionaries, exist (see Clarity a.o. 1964; Erwin 1963; Malaika 1963; Woodhead and Beene 1967). There are three articles on the slightly divergent Jewish, Christian, and Muslim dialects spoken in the northern town of Mosul (Jastrow 1979, 1989, to appear). The dialect of the most important city of the south (and second largest city of Iraq), Baṣra, so far has been described only in an unpublished dissertation (Mahdi 1985).

The following is a short dialect classification of Iraqi Arabic. Note that the classification of *gələt* dialects is still tentative.

Qəltu dialects

i. Tigris group
  Mosul and surroundings (Muslims, Christians, Jews, Yezidis)
  Tikrit and surroundings (Muslims)
  Baghdad and southern Iraq (Jewish and Christian communities only)

ii. Euphrates group
  ʿĀna (Muslims, Jews)
  Hīt (Muslims, Jews)

iii. Kurdistan group (Jewish communities only)
Northern Kurdistan (Sendor, ʿAqra, Arbil)
Southern Kurdistan (Kirkuk, Tuz Khurmatu, Khanaqin)

Gələt dialects
i. Northern and central Iraq
  Rural dialects of northern and central Iraq Sunni area around Baghdad Muslim Baghdadi

ii. Southern Iraq
  Rural dialects of southern Iraq
  Urban Muslim dialects

2. LINGUISTIC DESCRIPTION

2.1 *Phonology*

2.1.1 Consonants

Muslim Baghdadi has the inventory of consonant phonemes shown in Table 1.

Identical or very similar consonant systems are found in all Iraqi dialects. The following divergencies may be noted:

i. The voiceless bilabial stop /p/ is a stable phoneme in most Iraqi dialects. It has been introduced via loanwords from Turkish and Iranian (Persian, Kurdish) but also English, e.g. Muslim Baghdadi *pāča* 'traditional dish made from the head, feet, and stomach of sheep', *parda* 'curtain', *pančar* 'puncture, flat tire'.
ii. The interdental fricatives /ṯ/, /ḏ/, /ḏ̣/ (the latter being the joint reflex of Old Arabic *ḍād* and *ḏ̣āʾ*) have been retained in the vast majority of dialects, and constitute one of the hallmarks of Iraqi Arabic speech. They have shifted to the dental stops /t/, /d/, and /ḍ/ in Christian Baghdad and in the Jewish dialects of Kurdistan, with the exception of the northernmost dialect of Sendor, e.g. Arbil *talj* 'snow', *dahab* 'gold', *ḍəfər* 'fingernail' vs. Muslim Baghdad *ṯalij, ḏahab, iḏ̣fir.*

Table 1. Inventory of consonants in Muslim Baghdadi

	bilabial	labio-dental	apical	palatal	velar	uvular	pharyngeal	glottal
stop	p b		t d ṭ		k g	q		(ʾ)
affricate				č j				
fricative		f	ṯ ḏ ḏ̣ s z ṣ		x ġ		ḥ ʿ	h
nasal	m		n					
lateral			l ḷ					
vibrant			r ṛ					
semivowel	w			y				

iii. Alongside the lateral /l/ and the vibrant /r/ there exist the emphatic counterparts /ḷ/ and /ṛ/, which, in many Iraqi dialects, have a marginal phonemic status, e.g. Muslim Baghdad *čalib* 'dog' but *gaḷub* 'heart', *ḥāl* 'condition, state' but *xāḷ* 'maternal uncle'.

iv. Old Arabic /r/ has shifted to a velar fricative /ġ/ (phonetically identical with original /ġ/) in the *qəltu* dialects of the Tigris group and the southern part of the Kurdistan group, e.g. Muslim Baghdad *čiṯīr* vs. Mosul *kṯīġ* 'much'. In loanwords, including loans from Modern Standard Arabic, an apical /r/ is nevertheless pronounced. The shift /r/ > /ġ/ is very old and has been attested in vernacular Arabic of the Abbasid period; it is one of the hallmarks of most Iraqi *qəltu* dialects.

v. The voiced affricate /j/ has been preserved as such throughout most of Iraq; however, in Samāwa (situated on the Euphrates, halfway between Baghdad and Baṣra) a voiced fricative /ž/ is the common reflex. The south, with Baṣra as its center, has shifted /j/ to a palatal approximant, /y/, e.g. *yimal* 'camel', *ʿayūz* 'old woman'.

vi. The voiceless palatal affricate /č/ is a stable phoneme in most Iraqi dialects. It was introduced via loanwords from Turkish and Iranian (Persian, Kurdish) but also from English, e.g. Muslim Baghdad *čōl* 'desert', *čāy* 'tea'. In *gələt* dialects /č/ may also be a reflex of Old Arabic /k/, e.g. Muslim Baghdad *čam* 'how many', *simča* 'a fish' (see below under vii).

vii. Old Arabic /k/ has been preserved as a velar stop in the *qəltu* dialects but in the *gələt* dialects has shifted to the palatal affricate /č/ in the vicinity of front vowels, thus, e.g., Muslim Baghdad *čān* 'he was' vs. *ykūn* 'he shall be', *čīs* 'sack' vs. *kull* 'all'. The affrication of /k/ is more widespread in the rural *gələt* dialects, whereas in Muslim Baghdad /k/ has been preserved (or restituted) in many cases due to the stronger impact of Modern Standard Arabic.

viii. As implied by Blanc's terminology, Old Arabic /q/ has been preserved as a voiceless uvular stop in the *qəltu* dialects but shifted to a voiced velar stop /g/ in the *gələt* dialects, thus, e.g., Jewish Baghdad *qām* 'he got up', *qāl* 'he said' vs. Muslim Baghdad *gām, gāl.* Contrary to the rural *gələt* dialects, however, /q/ in Muslim Baghdad has been preserved (or restituted) in quite a number of words, e.g. *biqa* ~ *buqa* 'he stayed', *qira* 'he read', *qarya* 'village'. Parallelling the affrication of Old Arabic /k/ to /č/ there has been affrication of /g/ to /j/ in rural *gələt* dialects in the vicinity of front vowels, e.g. *jarya* 'village', *ṯijīl* 'heavy', but *ṯuguḷ* 'weight'. This feature is all but absent from Muslim Baghdad, e.g., *ṯigīl* 'heavy'; a noteworthy exception is the name of the Old City quarter, *Bāb išŠarji* (< *šarqi*) 'the East Gate'. In some rural and Bedouin dialects of northwestern Iraq, Old Arabic /ġ/ has shifted to /q/ (phonetically identical with the original /q/), while Old Arabic /q/ is represented by /g/, thus, e.g., *qanam* 'sheep' (< Old Arabic *ġanam*) but *gām* 'he got up'. In all Iraqi dialects /g/ may appear in foreign loans, e.g. the widespread words *gḷās* '(drinking) glass', *gḷōb* 'light bulb'.

ix. The glottal stop /ʾ/ is a marginal phoneme in city dialects, e.g. Muslim Baghdad *siʾal* 'he asked'. In most rural dialects it has been replaced by a glide, e.g. *sāyal, sayyal* 'he asked'.

2.1.2 Vowels

2.1.2.1 Long vowels and diphthongs

The inventory given in Table 2 is widespread, although the historical correspondences vary widely.

The Old Arabic diphthongs /ay/ and /aw/ have been preserved in two Jewish *qəltu* dialects in the north (Səndōṛ and ʿAqra) and in Tikrit but have been monophthongized to /ē/ and /ō/ in the rest of the country. In *gələt* dialects there

Table 2. Inventory of long vowels and diphthongs in Iraqi dialects

ī				*ū*		
	ē		*ō*		*ay*	*aw*
		ā				

is, however, a widespread occurrence of *-aw* as the ending of the 3rd person plural masculine perfect, e.g. Muslim Baghdad *ḏ̣urbaw* 'they shot'; *-aw* changes to *-ō-* when suffixes are added, e.g. *ḏ̣urbōni* 'they shot at me'.

Other than by monophthongization of /ay/ and /aw/ the mid long vowels /ē/ and /ō/ have entered the inventory by the following processes:

i. Via borrowing from Turkish, Iranian, English, etc., e.g. *čōl* 'desert', *xōš* 'good', *mēz* 'table', *mēwa* 'fruit'.
ii. In the *qəltu* dialects, by lowering of /ū/ and /ī/ in contact with emphatic and back consonants, e.g. Jewish Arbīl *daqēq* 'flour', *xəyōṭ* 'threads' (< Old Arabic *daqīq, xuyūṭ*).
iii. In the *qəltu* dialects, by the so-called → *ʾimāla*, i.e. the conditioned shift of Old Arabic /ā/ to /ē/ or /ī/ if the preceding or following syllable contains an *i* or *ī* vowel. In the Jewish *qəltu* dialects of Kurdistan and in Christian Baghdad, the *ʾimāla* only reaches /ē/, e.g. Jewish ʿAqra *basētīn* 'fruit gardens', *jēməʿ* 'mosque', pl. *jawēməʿ* < Old Arabic *basātīn, jāmiʿ, jawāmiʿ* (Jastrow 1990a:30), while in Jewish Baghdad it reaches /ī/, e.g. *jīməʿ*, pl. *jwīməʿ* (Blanc 1964:42). An interesting situation obtains in Mosul. While the Jewish dialect only has /ē/, the Muslim dialect has /ī/ when the *ʾimāla* has been triggered by Old Arabic /ī/, otherwise /ē/, e.g. Jewish Mosul *jawēməʿ*, *basētīn*, but Muslim Mosul *jawēməʿ, basītīn* (Jastrow 1989:285).

2.1.2.2 Short vowels

Qəltu dialects have a system of two short vowels, while the *gələt* dialects typically display a system of three short vowels (Table 3).

In open unstressed syllables the Old Arabic short high vowels /i/ and /u/ have been elided in the majority of Iraqi dialects, e.g. *flūs* 'money' < Old Arabic *fulūs*; in Jewish Baghdad, Jewish Kirkuk, and Tikrit, short /a/ has been elided as well, under the same conditions, e.g. Jewish Baghdad *ftaḥtu* 'I opened'. In all *qəltu* dialects /i/ and /u/ have merged into a single phoneme /ə/, e.g. Christian Baghdad *əbən* 'son', *əxət* 'sister' < Old Arabic *ibn, uxt*. In the *gələt* dialects the situation is much more complicated. Thus, in Muslim Baghdad /i/ and /u/ have been redistributed according to the phonetic environment. In what Blanc (1964:37) called 'u-coloring environments' (back, emphatic, and labial consonants), both older /i/ and /u/ now appear as /u/, e.g. *xubuz* 'bread' < *xubz, ṣudug* 'truth' < *ṣidq*. In so-called i-coloring environments (neither back nor emphatic and labial consonants), both older /i/ and /u/ now appear as /i/, e.g. *ṯiliṯ* 'third' < *ṯulṯ, isim* 'name' < *ism*. There is a third, less frequent 'color-preserving environment' in which original /i/ and /u/ are preserved as such, e.g. *yākul* 'he eats' < *yaʾkul* vs. *mākil* 'eating, having eaten' from a dialectal **mākil* (cf. Old Arabic *ʾākil*). Not only /i/ and /u/ have been redistributed in Muslim Baghdad; stressed /a/ in open syllable, if immediately followed by another /a/, has shifted to either /i/ or /u/ according to the phonetic environment, e.g. *simač* 'fish' < *samak*, but *buṣal* 'onion' < *baṣal; diras* 'he studied' < *daras*, but *ḏ̣urab* 'he hit' < *ḏ̣arab*. Old Arabic /a/ in closed syllables has generally been preserved, but in Tikrit, Jewish Kirkuk, and Jewish Baghdad, it has been raised to /ə/ when unstressed, e.g. Jewish Baghdad *fatḥət* 'she opened', but *fəṭḥə́tu* 'she opened it'.

Table 3. Inventory of short vowels in Iraqi dialects

qəltu	*gələt*
ə	*i u*
a	*a*

2.1.3 Suprasegmental features

2.1.3.1 Stress

In all Iraqi Arabic dialects stress is usually determined by syllable structure, according to the rule of thumb that stress will fall on vKK or v̄K closest to the end of the word, otherwise on the first v. In the *qəltu* dialects there is an additional morphological rule that requires the stress always to be on the last syllable of a nominal or verbal form if a pronominal suffix is added, e.g. Muslim Mosul *nəxlə́ṭu* 'we mix it', *dašaġə́lak* 'I shall make you [masc.] work', *sayyə́di* 'my grandfather'.

2.1.3.2 Consonant clusters and anaptyxis

Final clusters of two consonants are in general separated by an anaptyctic vowel (*ə* in the *qəltu* dialects, *i, u*, or *a* in the *gələt* dialects), e.g. Muslim Baghdad *miliḥ*, Christian Baghdad *məlaḥ* 'salt' < Old Arabic *milḥ*; Muslim Baghdad *uxut*, Christian Baghdad *əxət* 'sister' < Old Arabic *ʾuxt*; Muslim Baghdad *šahar*, Christian Baghdad *šahəġ* 'month' < Old Arabic *šahr*; Muslim Baghdad *dirásit*, Christian Baghdad *darásət* 'I studied' < Old Arabic *darast*. Jewish Baghdad, on the other hand, inserts an anaptyctic *ə* only when the last consonant is a liquid (*r, ġ, l, m, n*), thus *məlḥ, əxt, drast* but *əbən, šahəġ*. Word-initial clusters of two consonants and word-internal clusters of three consonants have an anaptyctic vowel after the first consonant (word-initial after the word juncture). In the *gələt* dialects, the anaptyctic vowel can be stressed, e.g. Muslim Baghdad *yuḏrub* + *-a* > **yuḏrba* > *yuḏúrba* 'he shoots him'. Only some Jewish *qəltu* dialects of the north (Arbīl, Kirkuk) have in this case the anaptyctic vowel after the second consonant, e.g. Arbīl *təqūm* 'she gets up', *ʿəndəkəm* 'with you'. In the imperfect, all *qəltu* dialects of the north retain the *ə* vowel of the base, e.g. *yəktəbūn* 'they write'.

2.1.3.3 Pausal forms

Final *-a* in Muslim Baghdad, irrespective of its origin, is pronounced as a slightly raised and centralized [ä] vowel. In some rural *gələt* dialects, the raising can reach *-e* (according to the notation of some sources, e.g. Salonen 1980). These pausal allophones tend to be generalized and thus can also occur in context.

2.1.4 Diachronic changes in syllable structure

2.1.4.1 Syllables of the type CaCaCa(C) have changed to *CvCvCa(C) in the *gələt* dialects; v symbolizes either /i/ or /u/ depending on the phonetic environment. In the urban *gələt* dialects, for instance Muslim Baghdad, the second vowel has been elided, thus *samaka* > *simča* 'fish', *qaṣaba* > *guṣba* 'reed'. In the rural *gələt* dialects the first vowel has been elided: *smiča* 'fish', *gṣuba* 'reed'; likewise in the verb: *ḏarabat* 'she hit' > *ḏurbat* (Muslim Baghdad), *ḏrubat* (rural).

2.1.4.2 The elision of short /i/ and /u/ in open syllables in the *gələt* dialects leads to anaptyctic vowels that acquire phoneme status and eventually bear the stress; this produces the effect called *Vokalumsprung*, e.g. **yidris-ūn* > **yidrs-ūn* > *yidirsūn* > *ydirsūn* 'they study'; **yuḏrub-a* > **yuḏrb-a* > *yuḏurba* > *yḏurba* 'he hits him'. In the *qəltu* dialects, /ə/ is usually not elided in the imperfect forms, e.g. Christian Baghdad *yəktəbūn* 'they write'. When a pronominal suffix is added, the stress remains on the last vowel of the preceding form (see 2.1.3.1), e.g. Jewish, Christian Baghdad *yəktə́bu* 'he writes it'.

2.2 *Morphology*

Gender distinction in the 2nd and 3rd person plural in verbs and pronouns has been preserved in rural *gələt* dialects, while it is absent both in urban *gələt* and in all *qəltu* dialects. Generally speaking, the former masculine forms have been generalized as the new communis forms.

2.2.1 Pronouns

2.2.1.1 Independent pronouns

Table 4 shows the independent personal pronouns in two *gələt* dialects, one rural and one urban, and in two *qəltu* dialects.

Table 4. Independent personal pronouns in Iraqi dialects

		Širgāṭ	Muslim Baghdad	Muslim Mosul	Jewish Mosul
3rd	sg. masc.	*huwwe*	*huwwa*	*huwwa (hīnu)*	*hūwe*
	sg. fem.	*hiyye*	*hiyya*	*hīya*	*hīye*
	pl. masc.	*humme*	*humma*	*humma (hīyəm)*	*həmme*
	pl. fem.	*hinne*			
2nd	sg. masc.	*inte*	*inta*	*ənta*	*ənta*
	sg. fem.	*inti*	*inti*	*ənti*	*ənti*
	pl. masc.	*intu*	*intu*	*əntəm*	*əntəm*
	pl. fem.	*intin*			
1st	sg.	*āni*	*āni*	*ana*	*ana*
	pl.	*iḥne*	*iḥna*	*nəḥna*	*nəḥna*

For the final *-e* in the Širgāṭ forms, see 2.1.3.3. The 1st person singular *āni* is a hallmark of *gələt* dialects and may be explained by analogy to the pronominal suffix *-il-ni*. Equally significant is the *gələt* form *iḥna*, as opposed to *qəltu nəḥna* (with initial *n*). In Muslim Mosul, the expected form *hīye* (cf. Jewish Mosul *hīye*) has become *hīya* by analogy to the pronominal suffix 3rd person singular feminine *-a*. The forms *hīnu* and *hīyəm* in turn are back-formations from *hīya*, by attaching to a basis *hīy-* the respective pronominal suffixes *-nu* and *-əm*; they are now becoming obsolete and are replaced by the respective forms of Muslim Baghdad.

2.2.1.2 Copula

In the Jewish dialects of Iraqi Kurdistan and in Jewish Mossul, a copula is used in nominal sentences. It consists of the unstressed and sometimes shortened forms of the independent pronouns that follow the predicate enclitically, e.g. Jewish Mosul *malīḥ-ūwe* 'he is good', *malīḥa-ye* 'she is good', Jewish ʿAqra *rəjēl-antəm* 'you are (real) men', *šabāb kəwayysīn-əme* 'they are good-looking young men'. While Jewish Baghdad does not have a copula, Christian Baghdad uses a copula consisting of the base *yā-* plus pronominal suffixes, e.g. *hāy šlōn bənət yāha?* 'what sort of girl is she?'

2.2.1.3 Pronominal suffixes

The pronominal suffixes (attached to *bēt* 'house' and *abū-* 'father') are set out in Table 5.

In Muslim Mosul verbal imperfects like *yšūfūnu* (< *yšūfūn* + *-u*) 'they see him' have been reanalyzed as *yšūfū-nu*, thus giving rise to a new pronominal suffix, *-nu*, which is attached to forms ending in a long vowel, e.g. *abūnu* 'his father', *nsammīnu* 'we name him', *ʿalēnu* 'on him', *šāfūnu* 'they saw him' etc. The suffix *-nu* is also found in Jewish and Christian Baghdad and seems to be a hallmark of the Tigris group. There are, however, differences in the extent to which *-nu* is applied. Thus, Jewish Mosul has *abūnu* but *nsammīyu, ʿaláyu*. On the other hand, the suffixes of the 3rd person replace the initial *h-* by a glide in Jewish Mosul, e.g. *abūwa* 'her father', *nsammīya* 'we name her', *abūwəm* 'their father'.

2.2.1.4 Relative pronouns

The relative pronoun is generally *illi* (*qəltu* dialects: *əlli*), with a frequent allomorph *il-* (*qəltu* dialects: *əl-*) which is identical with the definite article (2.2.3.1), cf. Muslim Baghdad *tuʿruf kull irriyājīl ič-čānaw bi lijtimāʿ*? 'do you know all the men who were at the meeting?'.

2.2.2 Particles

2.2.2.1 The definite article is *il-* (*gələt* dialects) or *əl-* (*qəltu* dialects). The *l* is usually assimilated to following 'sun letters', e.g. Muslim Baghdad *iššaxiṣ* 'the person', however, in Jewish ʿAqra and Arbīl there is a tendency to retain the *l-* unassimilated, e.g. *əlṣāla* 'the synagogue'.

2.2.2.2 Iraqi Arabic, both in its *qəltu* and *gələt* varieties, has an indefiniteness marker *fad* (< *fard*) 'a, one; some, a certain', e.g. Muslim Baghdad *fad yōm fad rajjāl gālla lxādma* 'one day a (certain) man told his servant' (→ article, indefinite). *Fad* is a prime

Table 5. Pronominal suffixes in Iraqi dialects

		Širgāṭ	Muslim Baghdad	Muslim Mosul
3rd	sg. masc.	*bētu/abū*	*bēta/abū*	*bētu/abūnu*
	sg. fem.	*bētha/abūha*	*bētha/abūha*	*bēta/abūha*
	pl. masc./comm.	*bēthum/abūhum*	*bēthum/abūhum*	*bētəm/abūhəm*
	pl. fem.	*bēthin/abūhin*		
2nd	sg. masc.	*bētak/abūk*	*bētak/abūk*	*bētak/abūki*
	sg. fem.	*bētič/abūč*	*bētič/abūč*	*bētki/abūki*
	pl. masc./comm.	*bētkum/abūkum*	*bētkum/abūkum*	*bētkəm/abūkəm*
	pl. fem.	*bētčin/abūčin*		
1st	sg.	*bēti/abūya*	*bēti/abūya*	*bēti/abūyi*
	pl.	*bētna/abūna*	*bētna/abūna*	*bētna/abūna*

hallmark of Iraqi speech as it is hardly found outside the Iraqi borders. Only → Uzbekistan Arabic (which ultimately hails from Iraq; see Jastrow 1998) has *fat* in the same function, e.g. *fat amīr kōn* 'there was an emir'. The lexeme *faqet* 'one, a certain' used in the Diyarbakır group of Anatolian Arabic seems to be a cognate of *fad* as well.

2.2.2.3 The most frequent genitive marker in Iraqi Arabic is *māl*, with an allomorph *mālat* which is, however, not subject to gender agreement, e.g. Muslim Baghdad *fad gahwa māl ḥaššāša* 'a café of hashish smokers'. In the northern *qəltu* dialects other genitive markers can be found, e.g. Jewish ʿAqra *līt*, pl. *lāt*.

2.2.3 Nouns

2.2.3.1 Feminine forms

In the *gələt* dialects the feminine ending is always *-a* (with phonetic variants as described in 2.1.3.3), e.g. Muslim Baghdad *simča* 'fish', *marga* 'stew', *ḫunṭa* 'wheat', while the *qəltu* dialects have two allomorphs: *-e*/*-i* after front consonants and *-a* after emphatic and back consonants, e.g. Jewish ʿAqra *samake* 'fish', *ḫənṭa* 'wheat', *maraqa* 'soup, stew'. The variant *-e* is predominant in the northern *qəltu* dialects, while in the south we find *-i*. In Mosul the Jewish dialect has the feminine ending *-e*, e.g. *jēje* 'hen', while the Muslim dialect shows *-i* (*jēji*). The distribution of the allomorphs *-i* and *-a* in Jewish Baghdad is particular insofar as it is not determined by the preceding consonant but by the presence of *i*, *ī*, *ē*, or *y* in the preceding syllable, in which case the *-i* allomorph occurs, e.g. *jīji* 'hen' but *samka* 'fish'.

2.2.4 Numerals

The numerals from one to ten are given in Table 6.

The numeral 'four' in Mosul has the peculiar form *ōbʿa* (< **aġbʿa*). In the Muslim dialect it has been replaced by the Baghdad form *arbaʿ-* ~ *arbā-* in the derived numerals *arbāṭáʿəš* 'fourteen', *arbaʿīn* 'forty' and *arbaʿmīyi* '400', while the Jewish dialect still has *ōbaʿṭaʿš, ōbʿīn*, and *ōbʿamīye*.

Wāḥid ~ *wēḥəd* is constructed as an adjective but can either precede or follow a noun. The number 'two' is frequently expressed by the dual, which is not restricted to time units, e.g. *yōmēn* 'two days', *sāʿtēn* 'two hours', but can be used with ordinary nouns as well, e.g. *waladēn* 'two boys', *bintēn* 'two girls'. When combined with a counted noun (always in the plural) the numerals from three to ten have shorter forms without the final vowel, e.g. Muslim Baghdad *tlaṯ/arbaʿ/xamis banāt*, Mosul (Jewish and Muslim) *ṯāṯ, xams banāt*, but Jewish *ōbaʿ banāt*, Muslim *arbaʿ banāt*.

A small list of nouns that originally began with *ʾv-* in the plural have special count plurals with initial *t-*; they are used after the numerals from three to ten, e.g. Muslim Baghdad *xamis talāf* 'five thousand', *xamis tiyyām* 'five days', *xamis tušhur* 'five months'.

The numerals from 11 to 19 only have a single form, e.g. Muslim Baghdad *xmuṣtaʿaš* 'fifteen', *xmuṣtaʿaš sana* 'fifteen years'.

Table 6. Numerals in Iraqi dialects

	Muslim Baghdad	Muslim Mosul	Jewish Mosul
1	*wāḥid*, fem. *wiḥda*	*wēḥed*, fem. *wəḥdi*	*wēḥed*, fem. *wəḥde*
2	*ṯnēn*, fem. *ṯintēn*	*ṯnēn*, fem. *ṯəntēn*	*ṯnēn* (no fem.)
3	*tlāṯa*	*ṯāṯi*	*ṯāṯe*
4	*arbaʿa*	*ōbʿa*	*ōbʿa*
5	*xamsa*	*xamsi*	*xamse*
6	*sitta*	*sətti*	*sətte*
7	*sabʿa*	*sabʿa*	*sabʿa*
8	*ṯmānya*	*ṯmēni*	*ṯmēnye*
9	*tisʿa*	*təsʿa*	*təsʿa*
10	*ʿašra*	*ʿašġa*	*ʿašġa*

2.2.5 Verbs

2.2.5.1 Derivation

2.2.5.1.1 Form I

Most *qəltu* dialects preserve two different vowel patterns in the perfect, reflecting Old Arabic 'transitive' *faʿala* and 'intransitive' *faʿila/faʿula* patterns. In the imperfect, the stem vowel is *ə* (< Old Arabic *i, u*) or *a*. In the *gələt* dialects the vowel qualities have been redistributed (2.1.2.2) and accordingly only a single pattern with predictable vowel variation remains, e.g. Muslim Baghdad *širab* 'he drank', *kitab* 'he wrote', *ḍurab* 'he shot', with Jewish ʿAqra, Arbīl *šərəb, katab, ḍarab*.

2.2.5.1.2 Derived forms

Form IV survives to a significant extent only in rural *gələt* dialects, in urban *gələt* and in *qəltu* dialects it survives only in some fossilized forms. The internal passive has all but disappeared; only in rural *gələt* are there still a few traces.

In Jewish ʿAqra and Arbīl, Forms II, III, V, VI, and X have identical inflectional bases for perfect and imperfect, the last syllable always being vocalized with *ə*, e.g. *ʿalləm/yəʿalləm* 'to teach', *tʿalləm/yətʿalləm* 'to learn', *ʿēwən/yəʿēwən* 'to help'.

2.2.5.2 Inflection

2.2.5.2.1 Table 7 contains the paradigm of the Form I imperfect of the strong verb.

2.2.5.2.2 Treatment of IIIy verbs

In the *gələt* dialects IIIy verbs are no longer distinguished from strong verbs; the inflectional endings of the strong verb have been generalized. In most of the *qəltu* dialects, however, the distinction is maintained, as in Table 8.

In Jewish Baghdad, Jewish Kirkuk, and Tikrit, on the other hand, the endings of the IIIy verbs have been generalized, thus Kirkuk, Jewish Baghdad *tkətbēn, ykətbōn*, Tikrit *tkətbayn, ykətbawn*.

Most *qəltu* dialects distinguish between strong and IIIy verbs also in the perfect, e.g. Jewish and Muslim Mosul *bano* < Old Arabic *banaw* 'they built'. In the *gələt* dialects, strong and IIIy verbs have been thrown together; in this case, however, the ending of the 3rd person plural masculine of the IIIy verb has been generalized, e.g. Muslim Baghdad *binaw* 'they built', *kitbaw* 'they wrote'.

2.2.5.2.4 Perfect forms with the empty morph *-ē-*

An almost universal feature of modern Arabic dialects, the insertion of an empty morph *ay* ~ *ē* in the perfect between an inflectional base ending in -CC and an inflectional suffix with an initial consonant, is also current in Iraqi Arabic, thus Muslim Baghdad *dazz* 'he sent', *dazz-ēt* 'I sent'. By analogy this formation has sometimes been extended to hollow verbs in Forms VII and VIII, e.g. Muslim Baghdad *ndāṛ* 'he turned', *ndāṛēt* 'I turned'. In southern Iraq, including the city of Baṣra, the use of the empty morph *-ē-* can also be found in Form I of the strong verb, alongside the regular inflection: (see Table 10)

Table 7. Imperfect of the strong verb (Form I) in Iraqi dialects

		Rural *gələt*	Muslim Baghdad	Muslim Mosul	Jewish Baghdad
3rd	sg. masc.	*yiktib*	*yiktib*	*yəktəb*	*yəktəb*
	sg. fem.	*tiktib*	*tiktib*	*təktəb*	*təktəb*
	pl. masc.	*yikitbūn*	*yikitbūn*	*yəktəbūn*	*ykətbōn*
	pl. fem.	*yikitbin*			
2nd	sg. masc.	*tiktib*	*tiktib*	*təktəb*	*təktəb*
	sg. fem.	*tikitbīn*	*tikitbīn*	*təktəbīn*	*tkətbēn*
	pl. masc.	*tikitbūn*	*tikitbūn*	*təktəbūn*	*tkətbōn*
	pl. fem.	*tikitbin*			
1st	sg.	*aktib*	*aktib*	*aktəb*	*aktəb*
	pl.	*niktib*	*niktib*	*nəktəb*	*nəktəb*

Table 8. Inflectional endings of strong and III*y* verbs

			Old Arabic	Muslim Baghdad	Muslim Mosul
3rd	pl. masc.	strong verb	*yaktubūna*	*yikitbūn*	*yəktəbūn*
		III*y*	*yansawna*	*yinsūn*	*yənsōn*
2nd	sg. fem.	strong verb	*taktubīna*	*tikitbīn*	*təktəbīn*
		III *y*	*tansayna*	*tinsīn*	*tənsēn*

Table 9. Perfect of the strong verb (Form I) in Iraqi dialects

		Rural *gələt*	Muslim Baghdad	Muslim Mosul	Jewish Baghdad
3rd	sg. masc.	*kitab*	*kitab*	*katab*	*katab*
	sg. fem.	*ktibat*	*kitbat*	*katabət*	*katbət*
	pl. masc./comm.	*ktibaw*	*kitbaw*	*katabu*	*katbu*
	pl. fem.	*ktibin*			
2nd	sg. masc.	*kitábit*	*kitábit*	*katábit*	*ktabt*
	sg. fem.	*kitabti*	*kitabti*	*katabti*	*ktabti*
	pl. masc./comm.	*kitabtu*	*kitabtu*	*katabtəm*	*ktabtem*
	pl. fem.	*kitabtin*			
1st	sg.	*kitábit*	*kitábit*	*katabtu*	*ktabtu*
	pl.	*kitabnē*	*kitabna*	*katabna*	*ktabna*

Table 10. Perfect of the strong verb (Form I) in Baṣra Arabic

3rd	sg. masc.	*kitab*
	sg. fem.	*kitbat*
	pl. masc./comm.	*kitbaw*
	pl. fem.	*kitban*
2nd	sg. masc.	*kitábit* ~ *kitbēt*
	sg. fem.	*kitabti* ~ *kitbēti*
	pl. masc./comm.	*kitabtu* ~ *kitbētu*
	pl. fem.	*kitabtan* ~ *kitbētan*
1st	sg.	*kitábit* ~ *kitbēt*
	pl.	*kitabnā* ~ *kitbēna*

2.3 Syntax

2.3.1 Noun phrase

While the use of the definite article *il- or əl-* (2.2.2.1) does not call for any comment, Iraqi Arabic is remarkable for its use of an indefiniteness marker, *fad* (2.2.2.2), e.g. Muslim Baghdad *šifna fad ibnayya kulliš ḥilwa* 'we saw a very pretty girl' (Erwin 1963:355).

2.3.2 Verbal phrase

2.3.2.1 Object marking

A definite object noun is often announced by an anticipatory object suffix on the preceding verb, e.g. Jewish Baghdad *ʿəqəb ma axḏūwa ləflūs* 'after they took [lit. 'they took it'] the money', Jewish ʿAqra *kūharrbūwa lbənt* 'they have kidnapped the girl'. In Muslim Baghdad the object noun itself is frequently introduced by the particle *l-*, e.g. *axaḏa l-xāla* ‘he took his uncle’. The same construction is also used with indirect objects, e.g., Muslim Baghdad *ḥičāla li-lqāḏ̣i ḥčāyta* ‘he told [lit. ‘he told him’] his story to the qadi’.

2.3.2.2 Expression of tense and aspect

2.3.2.2.1 Present tense

The Jewish dialects of Iraqi Kurdistan have a present tense marker *kū-* like some of the Anatolian dialects, e.g. Jewish ʿAqra *aš kūtrīd?* ‘what do you want?’, *kōqəllək* (< *kū* + *aqəllək*) ‘I am telling you’. The *qəltu* dialects of the Tigris group have a present tense marker *qa-* (< *qāʿid* ‘sitting’) with gemination of the following consonant, e.g. Jewish Baghdad *qayyəḥki* ‘he is talking’. The corresponding particle of Muslim Baghdad is *da-*, which is, however, used much more rarely, e.g. *aku ḥarāmi dayungub ḥāyiṭ bēthum* ‘there is a robber who is making a hole in the wall of their house’.

2.3.2.2.2 Future and intent

In Jewish Arbīl and ʿAqra the future is expressed by a prefix *tə-* ~ *də* - and in Mosul (all communities) by *də-*, e.g. ʿAqra *nəḥna tənəbqa yəhūd* ‘we shall become Jews’, *dīrmawn* (< *də-yərmawn*) ‘they will shoot’. When prefixed to the 1st person singular, the *ə* is elided, e.g. Muslim Mosul *daġōḥ* ‘I shall go’. In the *gələt* dialects invariable *rāḥ* is used, as is *ġāḥ* in the *qəltu* dialects of central and southern Iraq. In Muslim Baghdad *da-* is used not for the future but for intent only, e.g. *faš-šaxuṣ rāḥ da-yiṣbaḥ b ilḥammām* ‘a person went to the hammam to take a bath’.

2.3.2.2.3 Perfective

The perfective (resultative) aspect in the *gələt* dialects is expressed by the active participle, e.g. Muslim Baghdad *sāmiʿ ilʾaxbār?* ‘have you heard the news?’ In Mosul for the same purpose *kən-* is prefixed to the past tense, e.g. Muslim Mosul *kənmāt abūnu* ‘his father has died’. The Jewish dialects of Kurdistan, like some of the Anatolian dialects, use a prefix *kū-* e.g. Jewish ʿAqra *abūyi kūqāl* ‘my father has said.’

3. LEXICON

3.1 *Borrowings*

The most important source for lexical borrowing into Iraqi Arabic is Ottoman Turkish, second come Iranian (Modern Persian) and English. Many borrowings from these sources are part of the core lexicon and are still in current use, e.g. from Ottoman Turkish: *ūtači* ‘presser [of clothes]’, *bōyinbāġ* ‘tie’, *purtaqāl* ‘oranges’, *pōṣṭa* ‘post office’, *pōṣṭači* ‘postman’, *jōrāb* ‘pair of socks’, *junṭa* ‘suitcase’, *čāy* ‘tea’, *šarbat* ‘sherbet’, *qāṭ* ‘suit’, *qappūṭ* ‘overcoat’, *qazāṭma* ‘a dish of meat and rice’, *qišla* ‘barracks’, *qundara* ‘shoe’, *qundarči* ‘shoemaker’, *qūzi* ‘roasted lamb’, *gazīno* ‘open air café’, *gējaluġ* ‘nightgown’, *nēšan* ‘to aim’. From Persian: *parda* ‘curtain’, *xōš* [unvariable, preceding a noun] ‘good’, *mēz* ‘table’, *hamm* ‘also, too’. From English: *bōy* ‘waiter’, *pāṣ* ‘bus’, *talafōn* ‘telephone’, *sīnama* ‘cinema’, *gḷās* ‘glass’, *gḷōb* ‘light bulb’.

3.2 *Characteristic lexical items*

Characteristic lexical items include *fad* ‘a, one’ [indefiniteness marker, 2.2.2.2), *māl* (genitive marker, 2.2.2.3), *aku* ‘there is’, *maku* ‘there is not’, *hwāya* ‘much, a lot’, *wājid* ~ *wāyid* ‘a lot’ (in the south), *kulliš* ‘very’, *zēn* ‘good, nice’, *hamm* ‘also, too’, *bāčir* ‘tomorrow’, *hassa* ‘now’, *hna* ‘here’, *hnāk* ‘there’, *hīč* ‘thus, so’.

BIBLIOGRAPHICAL REFERENCES

Abu-Haidar, Farida. 1991. *Christian Arabic of Baghdad*. Wiesbaden: O. Harrassowitz.

Blanc, Haim. 1964. *Communal dialects in Baghdad*. Cambridge, Mass.: Harvard University.

Clarity, Beverly E., Karl Stowasser, and Ronald Wolfe. 1964. *A Dictionary of Iraqi Arabic: English-Arabic*. Washington, D.C.: Georgetown University Press.

Erwin, Wallace M. 1963. *A short reference grammar of Iraqi Arabic*. Washington, D.C.: Georgetown University Press.

Ingham, Bruce. 1973. “Urban and rural Arabic in Khuzistan”. *Bulletin of the School of Oriental and African Studies* 36.535-553.

Jastrow, Otto. 1979. “Zur arabischen Mundart von Mossul”. *Zeitschrift für Arabische Linguistik* 2.36-75.

——. 1983. “Tikrit Arabic verb morphology in a comparative perspective”. *al-Abḥāth* 31.99-110.

——. 1989. “Notes on Jewish Maṣlāwi”. *Jerusalem Studies in Arabic and Islam* 12.282-293.

——. 1990a. *Der arabische Dialekt der Juden von ʿAqra und Arbil*. Wiesbaden: O. Harrassowitz.

——. 1990b. “Die arabischen Dialekte der irakischen

Juden". *XXIV. Deutscher Orientalistentag vom 26. bis 30. September 1988 in Köln.* Ausgewählte Vorträge, ed. Werner Diem and Abdoljavad Falaturi, 199-206. Stuttgart: F. Steiner.

——. 1991. "Unheimliche Begebnisse in Sendor". *Semitic studies in honor of Wolf Leslau on the occasion of his eighty-fifth birthday, November 14th,* 1991, ed. Alan S. Kaye, I, 773-784. Wiesbaden: O. Harrassowitz.

——. 1993. "Geschichten aus Səndōṛ". *Zeitschrift für Arabische Linguistik* 25.161-177.

——. 1997. "Wie arabisch ist Uzbekistan-Arabisch?" *Built on solid rock: Studies in honour of Professor Ebbe Egede Knudsen on the occasion of his 65th birthday, April 11th, 1997*, ed. Elie Wardini, 141-153. Oslo: Novus.

——. 1998. "Zur Position des Uzbekistan-Arabischen". *Annäherung an das Fremde: XXVI. Deutscher Orientalistentag vom 25. bis 29. 9. 1995 in Leipzig*, ed. Holger Preissler and Heidi Stein, 173-184. Stuttgart: F. Steiner.

——. Forthcoming. "Jüdisches, christliches und muslimisches Arabisch in Mossul".

Khan, Geoffrey. 1997. "The Arabic dialect of the Karaite Jews of Hīt". *Zeitschrift für Arabische Linguistik* 34.53-102.

Mahdi, Qasim R. 1985. *The spoken Arabic of Baṣra, Iraq: A descriptive study of phonology, morphology and syntax.* Ph.D. diss., University of Exeter.

Malaika, Nisar. 1963. *Grundzüge der Grammatik des arabischen Dialekts von Bagdad.* Wiesbaden: O. Harrassowitz.

Mansour, Jacob. 1991. *The Jewish Baghdadi dialect.* Or-Yehuda: Babylonian Jewry Heritage Center.

Salonen, Erkki. 1980. *On the Arabic dialect spoken in Širqāṭ (Assur).* Helsinki: Suomalainen Tiedeakatemia.

Woodhead, Daniel R. and Wayne Beene. 1967. *A dictionary of Iraqi Arabic: Arabic-English.* Washington, D.C.: Georgetown University Press.

OTTO O. JASTROW
(University of Erlangen)

Ism

Ism, pl. *ʾasmāʾ*, is the technical term for 'noun(s)', the first of the three major parts of speech traditionally recognized in Arabic grammar (the other two being → *fiʿl*, pl. *ʾafʿāl* 'verb(s)', and → *ḥarf*, pl. *ḥurūf* 'particle(s)'; → parts of speech).

Morphosyntactically, the category of *ʾasmāʾ* is characterized by the following four features:

i. Nouns have declension, either full (*munṣarif*), partial (*ġayr munṣarif*), or invariable (*mabnī*);
ii. Nouns may be marked for definiteness (by the prefixed *al-*) or indefiniteness (by → *tanwīn* 'nunation');
iii. Nouns have three numbers, singular, dual, and plural;
iv. Nouns have two genders, masculine and feminine.

The last two features are not restricted to nouns, since verbs inflect for gender and number as well. A fifth feature, unique to Arabic nouns in the plural, may be added, namely animacy. Whether or not a noun refers to human beings determines the way it agrees with other elements of the phrase or clause (Badawi a.o. 2004:25; Ryding 2005: 119; also Abboud a.o. 2002:57–8).

Arabic nouns occur in a wide range of types identified by a variety of morphological patterns, such as → verbal nouns (*maṣdar*); active and passive → participles (*ism al-fāʿil/al-mafʿūl*); nouns of place (*ism makān*), instrument (*ism al-ʾāla*), profession, intensity, or repetition (of the *faʿʿāl*-pattern); abstract nouns (marked by the *nisba*-ending); and diminutives (*taṣġīr*; Badawi a.o. 2004:49–50; Ryding 2005:74–92).

The category of *ʾasmāʾ* comprises the subclasses of nouns (common and proper nouns, including cardinal numbers), pronouns (either free or bound), adjectives (including → elatives and ordinal numbers), adverbials, demonstratives, relatives, and nominal prepositions (i.e. nouns that function as prepositions; Badawi a.o. 2004:25; Abboud a.o. 2002:67–8).

In general, Western grammars of Modern Arabic lack a straightforward definition of *ism*; rather, they describe the category of nouns in terms of a set of grammatical properties as outlined above. Arab grammars and grammatical studies, on the other hand, usually start with a definition that is anchored in the Arabic linguistic tradition. According to this tradition – a consensus reached after much debate – a noun is "a word [→ *kalima*, also *lafḏ* 'sound'] indicating a meaning in itself and not containing any reference to time" (*ġayr muqtarin bi-zamān*, also *wa-laysa z-zamān juzʾan minhā* 'time is not part of it'; e.g. Dahdah 1990:29; cf. Ryding 2005:74, n. 1; Howell 1990:I, 1) or, alternatively, "a word indicating in itself something perceptible (*šayʾ maḥsūs*) or something imperceptible (*ġayr maḥsūs*) known by understanding (*bi-l-ʿaql*), both not containing any reference to time" (Ḥasan 2004:26).

This definition is typically followed by five characteristics that describe the category in the manner Western grammars generally do and that are based on grammatical properties considered to be unique to nouns. These features include the following:

i. Nouns may have the genitive case ending (*jarr*) by one of the grammatical reasons (*li-dāʿin min ad-dawāʿī n-naḥwiyya*) restricted to nouns, for instance by a preposition (*ḥarf al-jarr*);
ii. If fully inflected, nouns may have nunation;
iii. Nouns can be used as a vocative (*nidāʾ*);
iv. The *ʾalif lām* of definiteness may be prefixed to nouns; and
v. Nouns may be used as subject of announcement (*mubtadaʾ* → *ibtidāʾ*) or receive a predicate (*musnad ʾilayhi* → *ʾisnād*; Dahdah 1990:28–29; Ḥasan 2004:26–28; cf. Howell 1990:I, 1–3).

An additional set of features includes the fact that nouns may be governing words in a genitive construction (*muḍāf*), that personal pronouns (*ḍamīr/ḍamāʾir*) may refer to them, and that they form diminutives (Ḥasan 2004:29).

The tripartite division of words (*ʾaqsām al-kalām*) into nouns, verbs, and particles, though not always easy to maintain, has never really been challenged in the Arabic linguistic tradition. Only a few examples that indicate otherwise are known; as-Suyūṭī (*ʾAšbāh* III, 2) mentions *xālifa*, a category reportedly added by one Jaʿfar ibn Ṣābir; a category of *xawālif* is also recognized by al-Farābī, who has four more categories (cf. Guillaume 1988:25; Versteegh 1995:28). For a discussion of the number of parts of speech restricted to three, see az-Zajjājī (*ʾĪḍāḥ* 41–45). The introduction of this tripartite division into Arabic grammar is ascribed to either Caliph ʿAlī or his Baṣran judge ʾAbū l-ʾAswad ad-Duʾalī (see, e.g., Ibn al-ʾAnbārī, *Nuzha* 3–8).

That the basic tenet of the three parts of speech was taken from Greek grammar can no longer be sustained – the beginnings of the Arabic tradition lie before the introduction of Greek writings in the Arab world, even if we take Sībawayhi's *Kitāb* as the starting point. Yet, some traces of influence from Greek grammar, which may have entered the Arabic tradition through knowledge of the Hellenistic education system (Versteegh 1977:40, 1995:33), can, however, be recognized in the examples Sībawayhi uses to illustrate the noun category. Sībawayhi gives no real definition of *ism* but simply states: *fa-l-ismu rajulun wa-farasun* 'the noun is man and horse' (*Kitāb* I, 1). A third example, *ḥāʾiṭ* 'wall', has found its way into the printed editions of the *Kitāb*, but as Humbert (1995:57–58) has shown, this was a later addition. The manner in which grammarians of Arabic came, through trial and error, to formulate definitions of the parts of speech shows how, in the course of time, Arabic grammar yielded to the influence of Greek logic. Various theories on the introduction of the tripartite division of parts of speech in Arabic linguistic thinking have been put forward, for instance by Weiss (1976); Versteegh (1977:38–54); Guillaume (1988); Owens (1989); Suleiman (1990); for a general overview, see Versteegh (1995:27–30; → parts of speech).

Although Sībawayhi does not define *ism*, the *Kitāb* in several disconnected passages presents numerous properties of the noun, mainly based on their syntactic functions or on the basis of their morphological forms in relation to these functions. Thus, we learn that nouns have *tanwīn* (suffixed *-n*) as a sign of their being *mutamakkina* 'fully declinable' (lit. 'firmly established'), that they may have *jarr* (genitive case ending), as in the *muḍāf ʾilayhi* ('to which is annexed', i.e. the second part of a genitive construction), and that they can take the *ʾalif lām* of definiteness (prefixed *al-*, i.e. the article), to give but a few examples. In the same inventory-like manner, Sībawayhi identifies, among others, *kayfa* 'how?' and *ʾayna* 'where?' as nouns, which made it very complicated for representatives of the later grammatical tradition to formulate definitions of the noun (*Kitāb* I, 2.3–5, 7, 9–10, 20–21).

Subsequently, the grammarians applied the principle of similarity (*tašbīh, muḍāraʿa*; also → *qiyās* 'analogy'; cf. Carter 2004:82–86) to the relationship between words from the category of nouns and those from one of the other two categories, and in this way they identified subclasses of nouns with deviating characteristics. This allowed them to rank the three categories and the various subclasses in terms of heaviness (*ṯiqal*) or lightness (*xiffa*). The following example illustrates the working of this principle. Nouns constitute the lightest category of words due to their property of having *tanwīn*, which makes them unrestricted in their movements (*taṣrīf, taṣarruf*). Some nouns, however, show the same pattern as the *fiʿl muḍāriʿ*, the imperfect verb ('that resembles the noun'): nouns like *ʾabyaḍu* 'white' or *ʾaḥmaru* 'red', for instance, have the same *ʾafʿalu* pattern as *ʾaḏhabu* 'I go' and *ʾaʿlamu* 'I know'. This similarity makes them share the verbal property of being too 'heavy' to bear *tanwīn*. Consequently, they lose their ability to be fully inflected (*mutamakkin*, also *munṣarif*) and become less free in their movements, and hence they are called *ġayr mutamakkin* or *ġayr munṣarif* (i.e. diptotes instead of triptotes; → diptosis; Sībawayhi, *Kitāb* I, 5.8–14; II,1.11ff.). Az-Zajjāj's *Kitāb mā yanṣarif wa-mā lā yanṣarif* was a work entirely devoted to this subject (cf. Carter 2004:68, 104–106, 108; Versteegh 1995:174, n. 3). Ḥadīṯī (2003:296–297) even credits Sībawayhi with 'inventing' the science of *ṣarf*. For an overview of the subclasses of nouns as defined by the grammatical tradition, see Howell (1990:I, 1–1813 [Parts 1–4]). Specific classificatory problems concern adverbs of time and place, *ḏ̣arf/ḏ̣urūf* (Versteegh 1995:68; Owens 1988:131–138; cf. Howell 1990:I, 217–227; Wright 1967:II, 109–112; → *mafʿūl fīhi*), and → *ṣifa* 'adjective, attribute' (Diem 1974; Versteegh 1977:49–50, 1995:29–30).

Other early grammarians and those of the first generations after Sībawayhi more or less followed his cue in describing rather than defining the class of nouns in Arabic. Some of these descriptions are the following: *al-ism mā wuṣifa* 'the noun is that to which [something] can be attributed' (al-Kisāʾī, d. 183/799); *mā ḥasuna fīhi 'yanfaʿunī wa-yaḍurrunī' fa-huwa ism* 'that about which it is allowed to say "it benefits me" or "it hurts me" is a noun', i.e., as az-Zajjājī (*ʾIḍāḥ* 49) explains, *mā jāza ʾan yuxbara ʿanhu* 'about which it is allowed to predicate something' (al-ʾAxfaš al-ʾAwsaṭ, d. ca. 215/830); *al-ism mā ḥtamala t-tanwīn ʾaw al-ʾiḍāfa ʾaw al-ʾalif wa-l-lām* 'the noun is what bears the nunation or the annexion or the article' (al-Farrāʾ, d. 207/822); *kull mā daxala ʿalayhi ḥarf min ḥurūf al-jarr fa-huwa ism fa-ʾin imtanaʿa ḏālika fa-laysa bi-sm* 'everything to which a preposition can be attached is a noun, and if this is impossible, it is not a noun' (al-Mubarrad, d. 285/989).

A neat overview of these early, as well as some later, attempts to formulate quasi-definitions of *ism* is presented by Ibn Fāris (*Ṣāḥibī* 89–91). He himself is satisfied with none of them because they do not offer any room for *kayfa, ʾayna*, and the like to be included in the category of nouns.

When referring to al-Mubarrad's above-mentioned description of noun, Ibn Fāris apparently leaves out the first part, found at the very beginning of the *Muqtaḍab* (I, 1). Al-Mubarrad first restricts the noun to the *ism mutamakkin* 'fully declinable noun', and then proceeds by saying *ʾammā l-ʾasmāʾ fa-mā kāna wāqiʿan ʿalā maʿnan naḥwa rajul wa-faras* 'nouns are what touches upon ('what denotes', Jarrar 1992:134; 'what is applied to', Versteegh 1995:51) a meaning, like "man" and "horse"'. It is precisely the inclusion of → *maʿnā* 'meaning' in the description that initiates the formulation of the first true definition of *ism* by the late- 3rd/9th-early- 4th/10th-century grammarian az-Zajjāj, which runs as follows: *ṣawt muqaṭṭaʿ mafhūm dāll ʿalā maʿnan ġayr dāll ʿalā zamān wa-lā makān* 'an articulated understandable sound, indicating [or: implying] meaning and not indicating [or: implying] time or place'. Comparing this with the Aristotelian definition of the noun, "a noun is a sound with meaning [by agreement] without time, no part of which has meaning in isolation", Fleisch (1978:181–182) recognizes the Greek logical basis of the definition, and, consequently, he makes a clear distinction between the grammarians before and after az-Zajjāj; 'by agreement' belongs to the discussion about the origin of language, whether by institution or by revelation (for a discussion of variant formulations of Aristotle's definition, see Versteegh 1977:138, 1995:58).

Ibn Fāris (*Ṣāḥibī* 2) is equally dissatisfied with az-Zajjāj's definition of *ism*, this time because it does not exclude particles. His objections clearly show how difficult it is to stick to the rationale of categorization, namely that words belong to one category and one category only. The many subclasses of noun appeared to be very hard to capture in one overall definition, so that even az-Zajjāj's rather sophisticated definition had to be modified and improved on. Criticisms of and attempts to improve this definition basically focus on two aspects: one having recourse to *maʿnā* for the purpose of excluding particles, the other dealing with aspects of time to distinguish nouns from verbs.

Ibn as-Sarrāj (d. 316/928), for instance, modifies *maʿnā* by adding *mufrad* 'a single meaning', *wa-ḏālika l-maʿnā yakūnu šaxṣan wa-ġayr šaxṣin* lit. 'and this meaning is a person or not a person'. From his examples it becomes clear that he refers to concrete (e.g. *rajul, faras*) and abstract (*ḍarb* 'hitting', *ʾakl* 'eating') nouns. In an elaborate discussion of the matter at hand, az-Zajjājī remarks that this modification is not an improvement because the definition still allows meaningful particles, like *ʾan* 'that' or *lam* 'not', to be included in the category of nouns. Reference to the aspect of time in az-Zajjāj's definition is formulated as *maʿnan ġayr muqtarin bi-zamān muḥaṣṣal min maḍiyyin aw ġayrihi* 'a meaning not associated with a specified time, be it past time or another', for instance in as-Sīrāfī (d. 368/958; Ibn as-Sarrāj, *ʾUṣūl* I, 36–37; Zajjājī, *ʾĪḍāḥ* 48; Sīrāfī, *Šarḥ* I, 53–54; on the fundamental distinction between concrete and abstract nouns, see Versteegh 1995:61).

Az-Zajjājī provides us in the *ʾĪḍāḥ* (48–52) with a thorough discussion of the definitions based on logic that were known to him at the time, but he himself explicitly prefers to remain on the solid ground of linguistics and formulates the following definition on a syntactic basis: "A noun in the language of the Arabs is what is an agent or an object or what replaces the agent and the object" (*al-ism fī kalām al-ʿarab mā kāna fāʿilan aw mafʿūlan aw wāqiʿan fī ḥayyiz al-fāʿil wa-l-mafʿūl bihi*). At this point, az-Zajjājī underlines his preference for grammatical rather than logical reasoning by adding: "This definition is in accordance with grammatical standards and conventional rules. It includes absolutely all nouns and excludes everything which is not a noun" (Versteegh 1995:49–53, 56–64; Jarrar 1992:129–135). For the interpretation of *fāʿil* and *mafʿūl* in this context as 'something active or something passive' because the definition deals with nouns as representatives of substances, see Versteegh (1995:57). Otherwise, Versteegh argues, words such as *kayfa* and *ayna*, which denote substances and hence may be regarded as something active or something passive, are not covered by the definition. For further details see, however, Owens (1989:215–219), who compares az-Zajjājī's definition with earlier versions ascribed to Luġda (d. 310/922) and his contemporary Ibn Kaysān.

Ibn al-ʾAnbārī (*ʾAsrār* 9–10) remarks that in due course up to seventy definitions of *ism* had been formulated (which Versteegh 1995:56 roughly divides into the five main categories of morphological, ontological, logical, syntactic, and exemplifying categories). Some grammarians even assert that no clear-cut definition can be established, and they venture to say that this is the reason why Sībawayhi did not even try to give one, sticking to a description and a few illustrations. And, indeed, in the Arabic linguistic tradition a gradual movement to what may be termed a two-pronged approach may be observed: a definition including a reference to 'meaning without time', followed by an enumeration of properties to make certain that everything is absolutely clear. For instance, az-Zamaxšarī (d. 538/1143–1144) rephrases the definition as follows: *al-ism mā dalla ʿalā maʿnan fī nafsihi dalālatan mujarradatan ʿan al-iqtirān [bi-zamān] wa-lahu xaṣāʾiṣ minhā jawāz al-ʾisnād ʾilayhi wa-duxūl ḥarf at-taʿrīf wa-l-jarr wa-t-tanwīn wa-l-ʾiḍāfa*. Here, we find *māʿnan fī nafsihi* 'a meaning in itself', in order to avoid overlap with the category of particles and *mujarradatan ʿan al-iqtirān [bi-zamān]* 'free from association with time', in order to make sure verbs are excluded; this is followed by a selection of the most characteristic properties of the noun (Zamaxšarī, *Mufaṣṣal* 6; Ibn Yaʿīš, *Šarḥ* I, 22–25; Xwārizmī, *Taxmīr* I, 157–158). The same manner of defining the noun, outlined above, is embraced by contemporary Arab grammarians.

Finally, in the Arabic grammatical tradition, the etymology of the word *ism* belongs to the traditional Baṣran-Kūfan debate (recorded as the first *masʾala* in Ibn al-ʾAnbārī's *ʾInṣāf* and in his *ʾAsrār* 5–10; summarized by, for example, Zabīdī, *Tāj* IX, 538, followed by Lane 1863–1893:IV, 1435). The word has found its way into the Modern Arabic dictionary of Wehr (1994) under the radicals *s-m* (cross-referenced with *s-m-y* from *sammā* 'to name, designate'). Yet, consensus follows the Baṣran theory that *ism* (or, according to some, the tribes of Tamīm and Quḍāʿa say *usm*) is derived from *s-m-w* 'to be elevated, raised', and the original *simw* or *sumw* is a thing raised into notice, thus denoted and made known. The last radical, *wāw*, has been dropped and compensated by a prefixed *hamza*; consequently, its plural is *ʾasmāʾ* and its diminutive *sumayy*. According to some Kūfans, however, the word is from *w-s-m, wasm* (meaning *ʿalāma* 'sign, indication’, synonym of *ism*), the first radical *wāw* being substituted by *hamza*. This theory is considered to be weak because *wasm* would form the plural *ʾawsām* and the diminutive *wusaym*. Fleisch (1978:181, 1961:252–261) treats the word as biliteral and considers it as such to belong to ‘ancient linguistic stock’. The etymology of the word *ism* has some bearing on the discussion about whether *ism* and *musammā*, the sign and the designated, are identical, an issue that belongs to the origin of speech (see, e.g., Ibn Jinnī, *Xaṣāʾiṣ* I, 40–47; Ibn Manẓūr, *Lisān al-ʿArab* III, 2109–2110; Versteegh 1995:34, 38).

Bibliographical references

Primary sources

Ibn al-ʾAnbārī, *ʾAsrār* = ʾAbū l-Barakāt ʿAbd ar-Raḥmān Muḥammad ibn ʾAbī Saʿīd al-ʾAnbārī, *ʾAsrār al-ʿArabiyya*. Ed. Muḥammad Bahja al-Bīṭār. Damascus: Maṭbaʿat az-Zaqī, 1957.
——, *ʾInṣāf* = ʾAbū l-Barakāt ʿAbd ar-Raḥmān Muḥammad ibn ʾAbī Saʿīd al-ʾAnbārī, *Kitāb al-ʾinṣāf fī masāʾil al-xilāf bayna n-naḥwiyyīna l-Baṣriyyīna wa-l-Kūfiyyīn*. Ed. Gotthold Weil. Leiden: E.J. Brill, 1913.
——, *Nuzha* = ʾAbū l-Barakāt ʿAbd ar-Raḥmān Muḥammad ibn ʾAbī Saʿīd al-ʾAnbārī, *Nuzhat al-ʾalibbāʾ fī ṭabaqāt al-ʾudabāʾ*. Ed. Attia Amer. Stockholm: Almquist and Wiksell, 1963.
Ibn Fāris, *Ṣāḥibī* = ʾAbū l-Ḥusayn ʾAḥmad Ibn Fāris, *aṣ-Ṣāḥibī [fī fiqh al-luġa]*. Ed. ʾAḥmad Ṣaqr. Cairo: Maktaba ʿĪsā al-Bābī al-Ḥalabī, 1977.
Ibn Jinnī, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. Ed. Muḥammad ʿAlī an-Najjār. 3 vols. Beirut: Dār al-Hudā, n.d.
Ibn Manẓūr, *Lisān* = Jamāl ad-Dīn ʾAbū l-Faḍl Muḥammad ibn Mukarram ibn Manẓūr, *Lisān al-ʿArab*. Ed. ʿAbdallāh ʿAlī al-Kabīr a.o. 6 + 3 vols. Cairo: Dār al-Maʿārif, n.d.
Ibn as-Sarrāj, *ʾUṣūl*, = ʾAbū Bakr Muḥammad ibn Sahl Ibn as-Sarrāj, *al-ʾUṣūl fī n-naḥw*. Ed. ʿAbd al-Ḥusayn al-Fatlī. 3 vols. Beirut: Muʾassasat ar-Risāla, 1987.
Ibn Yaʿīš, *Šarḥ* = Muwaffaq ad-Dīn Yaʿīš Ibn Yaʿīš, *Šarḥ al-Mufaṣṣal*. 10 vols. in 5. Cairo: ʾIdārat aṭ-Ṭibāʿa al-Munīriyya, n.d. [*Fahāris Šarḥ al-Mufaṣṣal*, Damascus: Majmaʿ al-Luġa al-ʿArabiyya, 1990.]
Mubarrad, *Muqtaḍab* = ʾAbu l-ʿAbbās Muḥammad ibn Yazīd al-Mubarrad, *Kitāb al-muqtaḍab*. Ed. Muḥammad ʿAbd al-Xāliq ʿUḍayma. 4 vols. Cairo: Maṭābiʿ al-ʾAhrām, 1979.
Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr b. ʿUṯmān Sībawayhi, *Kitāb Sībawayhi*. Ed. Hartwig Derenbourg. 2 vols. Paris: Imprimerie Nationale, 1881. (Repr., Hildesheim and New York: G. Olms, 1970.)
Sīrāfī, *Šarḥ* = ʾAbū Saʿīd al-Ḥasan ibn ʿAbdallāh as-Sīrāfī, *Šarḥ Kitāb Sībawayhi*. Ed. Ramaḍān ʿAbd at-Tawwāb. 6 vols. (all published). Cairo: al-Hayʾa al-Miṣriyya, 1986–.
Suyūṭī, *ʾAšbāh* = Jalāl ad-Dīn ʿAbd ar-Raḥmān as-Suyūṭī, *al-ʾAšbāh wa-n-naḏāʾir*. Ed. ʾIbrāhīm Muḥammad ʿAbdallāh. 3 vols. Damascus: Dār al-Maʿārif, 1985–1986.
Xwārizmī, *Taxmīr* = ʾAbū Muḥammad al-Qāsim ibn al-Ḥusayn al-Xwārizmī, *Šarḥ al-Mufaṣṣal fī ṣanʿat al-ʾiʿrāb al-mawsūm bi-t-Taxmīr*. Ed. ʿAbd ar-Raḥmān ibn Sulaymān al-ʿUṯaymīn. 4 vols. Beirut: Dār al-ʿArab al-ʾIslāmī, 1990.
Zabīdī, *Tāj* = Murtaḍā az-Zabīdī, *Tāj al-ʿarūs*. Ed. ʿAlī Šīrī. 20 vols. Beirut: Dār al-Fikr, 1994.
Zajjāj, *Mā yanṣarif* = ʾAbū ʾIsḥāq ʾIbrāhīm ibn as-Sarī az-Zajjāj, *Mā yanṣarif wa-mā lā yanṣarif*. Ed. Hudā Maḥmūd Qarāʿa. Cairo: Maṭābiʿ al-ʾAhrām, 1971.
Zajjājī, *ʾĪḍāḥ* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān ibn ʾIsḥāq az-Zajjājī, *al-ʾĪḍāḥ fī ʿilal an-naḥw*. Ed. Māzin al-Mubārak. Beirut: Dār an-Nafāʾis, 1986.
Zamaxšarī, *Mufaṣṣal* = ʾAbū l-Qāsim Maḥmūd ibn ʿUmar az-Zamaxšarī, *al-Mufaṣṣal fī ʿilm al-ʿarabiyya*. Beirut: Dār al-Jīl, n.d.

Secondary sources

Abboud, Peter F., Aman Attieh, Ernest McCarus, and Raji Rammuny (eds.). 2002. *Intermediate Modern Standard Arabic*. Rev. ed. Ann Arbor, Mich.: Center for Middle Eastern and North African Studies.
Badawi, Elsaid, Michael G. Carter, and Adrian Gully. 2004. *Modern Written Arabic: A comprehensive grammar*. London and New York: Routledge.
Carter, Michael G. 2004. *Sibawayhi*. Oxford: Tauris.
Dahdah, Antoine el-. 1990. *A dictionary of universal Arabic grammar*. Rev. ed. George M. Abdul-Massih and Elias Matar. Beirut: Librairie du Liban. (In Arabic.)
Diem, Werner. 1974. “Nomen, Substantiv und Adjektiv bei den arabischen Grammatikern”. *Oriens* 23–24.312–323.
Fleisch, Henri. 1961. *Traité de philologie arabe*. I. *Préliminaires, phonétique, morphologie nominale*. Beirut: Librairie du Liban.
——. 1978. “*Ism*”. *Encyclopaedia of Islam* IV, 181–182. 2nd ed. Leiden: E.J. Brill.
Guillaume, Jean-Patrick. 1988. “Le discours tout entier est nom, verbe et particule: Elaboration et constitution de la théorie des parties du discours dans la tradition grammaticale arabe”. *Langages* 22.25–36.
Ḥadīṯī, Xadīja al-. 2003. *ʾAbniyat aṣ-ṣarf fī* Kitāb *Sībawayhi*. 2nd ed. Beirut: Librairie du Liban. (1st ed., Baghdad, 1965.)
Ḥasan, ʿAbbās. 2004. *an-Naḥw al-wāfī*. 4 vols. Cairo: Dār al-Maʿārif.
Howell, Mortimer Sloper. 1990. *A grammar of the Classical Arabic language, translated and compiled from the works of the most approved native or naturalised authors, with a new preface by Satkari Mukhopadhaya*. 4 vols. in 7 parts. Repr. New Delhi: Gian Publishing House. (1st ed., 1883; 2nd ed., 1986.)
Humbert, Geneviève. 1995. *Les voies de la transmission du* Kitāb *de Sībawayhi*. Leiden: E.J. Brill.
Jarrar, Walid Sadeq. 1992. *The grammatical concept of causality in Arabic: A study and an annotated translation of al-Zajjājī’s book ‘al-Iḍāḥ fī ʿilal al-*

naḥw’ (On explaining grammatical causes). Ph.D. diss., Indiana University.

Lane, Edward William. 1863–1893. *An Arabic-English lexicon derived from the best and most copious Eastern sources*. 8 vols. London and Edinburgh: Williams and Norgate.

Owens, Jonathan. 1988. *The foundations of grammar: An introduction to medieval Arabic grammatical theory*. Amsterdam and Philadelphia: J. Benjamins.

——. 1989. “The syntactic basis of Arabic word classification”. *Arabica* 36.211–233.

Ryding, Karin C. 2005. *A reference grammar of Modern Standard Arabic*. Cambridge: Cambridge University Press.

Suleiman, M. Yasir. 1990. “Sībawayhi’s ‘parts of speech’ according to Zajjājī: A new interpretation”. *Journal of Semitic Studies* 35.245–263.

Versteegh, Kees. 1977. *Greek elements in Arabic linguistic thinking*. Leiden: E.J. Brill.

——. 1995. *The explanation of linguistic causes: Az-Zaǧǧāǧī’s theory of grammar*. Amsterdam and Philadelphia: J. Benjamins.

Wehr, Hans. 1994. *A dictionary of Modern Written Arabic (Arabic-English)*. Ed. J. Milton Cowan. 4th ed. Ithaca, N.Y.: Spoken Language Services.

Weiss, Bernard. 1976. “A theory of the parts of speech in Arabic (noun, verb and particle): A study in *ʿilm al-waḍʿ*”. *Arabica* 23.23–36.

Wright, William. 1967. *A grammar of the Arabic language, translated from the German of Caspari and edited with numerous additions and corrections*. 3rd ed. rev. W. Robertson Smith and M.J. de Goeje. 2 vols. in 1 part. Cambridge: Cambridge University Press. (1st ed., 1859–1962.)

MONIQUE BERNARDS (University of Groningen)

Ism al-fāʿil

It is characteristic of medieval terminology not to be specific to one analytical level, thus, for example, → *ḥarf* ‘particle’ can refer to anything from phonemes to paragraphs. The word *fāʿil* is no exception. It is discussed here on four levels.

1. Lexically, → *fāʿil* means ‘doer, person doing’. An obscene metaphor, doubtless coined by a grammarian, pairs it with another technical term, *mafʿūl bihi* ‘done to’, viz. ‘direct object’, for the active and passive members of a homosexual relationship.

2. At the morphological level, *fāʿil* has three applications:

(a) *Fāʿil* stands for any word in the template $C_1\bar{a}C_2iC_3$, where the root *f-ʿ-l* ‘to do’ acts as a morphological symbol, hence *xāmis* ‘fifth’ is said to have the pattern *fāʿil*.

(b) It denotes the pattern of the agent of a simple verb, rather as one says that the agent in English is formed by adding *-er* to the stem, e.g. *kātib* ‘writer’, *mālik* ‘owner’, *ḏ̣ālim* ‘wrongdoer’. Not every word of this pattern is necessarily an agent, e.g. the toponym *qādis* ‘Cadiz’.

(c) By extension *fāʿil* stands for the whole morphological category of words denoting agents regardless of the pattern, which varies according to the verb stem, e.g. *mudarris* ‘instructor’, *murāsil* ‘correspondent’, *mudīr* ‘director’, *mutaʿallim* ‘learner’, etc. The allomorphs *faʿʿāl*, *fiʿʿāl*, etc., which denote an agent in an exaggerated sense, e.g. *kaḏḏāb* ‘great liar’, *širrāb* ‘heavy drinker’, are also covered by this term.

3. At the syntactic level, the phrase *ism al-fāʿil* denotes the ‘agent noun’ in all aspects: word order, case, agreement, and other features (→ *fāʿil*). Apart from the Qurʾānic quotations, the examples below are artificial, and the translations are literal enough to expose the structures under review (often in bold type).

(a) Word order. By observation the grammarians knew that the agent noun always follows its verb, e.g. *ḍaraba zaydun* ‘Zayd struck’. The agent pronouns, being fundamentally suffixes, morphologically confirmed this order, e.g. *ḍarabtu/ḍarabū* ‘I struck/they struck’, analyzed as *ḍarab+tu* ‘struck-I’/*ḍarab+ū* ‘struck-they’. With passive verbs, e.g. *ḍuriba zaydun/ḍuribtu* ‘Zayd was struck/I was struck’, the grammatical agent is called *nāʾib ʿan al-fāʿil* ‘deputy agent’, among other terms.

(b) Case. The agent noun takes the independent (*rafʿ*) case, in common with subjects and predicates. Several functional equivalents for the agent noun were identified by the grammarians, such as prepositional phrases, e.g. the Qurʾānic *mā jāʾanā min bašīrin* lit. ‘not came to us [anyone] by way of messenger’ (Q. 5/13) and *kafā bi-llāhi šahīdan* lit. ‘there is enough with God as a witness’, interpreted as ‘God suffices as a witness’ (Q. 4/79, 166); nominalized clauses, e.g. *yabqā ʾan ʾaḏkura* lit. ‘that I mention remains’, i.e. ‘it remains for me to mention’; and free pronouns in certain contexts, e.g. *mā faʿala ʾillā ʾana* ‘none did [it] but I’.

(c) Agreement is the most outstanding feature of verbal syntax: the verb precedes its agent and remains singular, agreeing with the agent only in gender, e.g. *qāma r-rajulu/r-rajulāni/*

r-rijālu 'the man/two men/men stood up [masc. sg.]', *qāmat al-marʾatu/al-marʾatāni/an-nisāʾu* 'the woman/two women/women stood up [fem. sg.]'. Historically, this is a regularization of a more fluid situation allowing other forms of agreement, notably the dialectal *ʾakalūnī l-barāġīṯu* 'the fleas consumed [masc. pl.] me', and the Qurʾānic *qāla niswatun fī l-madīnati* 'women in the town said [masc. sg.]' (Q. 12/30); agreement with the sense or by attraction is not rare, e.g. *jāʾathu kitābī* 'my letter [masc.] came [fem.] to him', where the speaker is thinking of feminine *risāla* 'message'. Much discussion was provoked among the grammarians by agreement problems in elliptical coordinated verbal sentences of the type *marartu wa-marra bī zaydun* lit. 'I passed and Zayd passed by me', similar to English *They do not and have not written together* (the example is genuine).

(d) The overt agent noun does not have to follow its verb immediately, but pronoun agents are incorporated in the verb, reinforcing their mutually indispensable formal and semantic relationship: this indispensability also exists between subjects and predicates, and it is one reason why agents, subjects, and predicates all have the same independent (*rafʿ*) inflection. There are, however, two fundamental differences between agents and subjects. The first difference is that agents, being integral with their verbs, can never be elided, while subjects and predicates can be left unsaid. Every verb is in effect a complete sentence, and there is no infinitive in Arabic, only fully inflected verbal nouns (→ *maṣdar*).

(e) Nominal vs. verbal sentence. The second difference between subjects and agents is that their status is determined by word order. Since the order verb+agent is fixed, any sequence in which the verb is preceded by a noun must be analyzed differently: *zaydun qāma* can only be parsed as 'Zayd, he stood up', i.e. an equational sentence whose predicate is a verbal sentence. This is clear from the obligatory agreement in the postposed verb, e.g. *ar-rajulāni qāmā/ar-rijālu qāmū* 'the two men, they [du.] stood up/the men, they [pl.] stood up'. Agreement is obligatory because all predicates in Arabic must contain a pronoun linking them to the subject, and it is always possible at *qām-* to introduce a new agent, e.g. *ar-rijālu qāmat ʾummuhum* 'the men, their mother stood up [fem. sing.]' with the binding pronoun now elsewhere.

The Arab grammarians maintained a strict terminological distinction between the structure verb+agent, which they called a 'verbal sentence' (*jumla fiʿliyya*), and the binary unit of subject+predicate, e.g. *zaydun qāʾimun* 'Zayd [is] standing', which they termed a 'nominal sentence', *jumla ismiyya*. Regrettably, *fāʿil*, which in Arabic can only denote the agent of a verb, is all too often rendered inappropriately as 'subject', thereby obliterating the indigenous grammatical theory altogether.

4. At the semantic and theoretical levels, various issues were raised by the grammarians, some with important logical and theological implications.

(a) Miscellaneous theoretical topics. The assignment of independent case to agents, subjects, and predicates was accounted for by their structural similarities; in addition, for reasons of systematic coherence, the agent noun was said to take the (phono)logically more significant case (independent) over the dependent (*naṣb*) case of objects and complements because there is only one agent but multiple complements. The latter moreover are intrinsically redundant elements (*faḍla*), while the agent is a structural necessity (*ʿumda*).

The agreement system led to the inference that the agent pronoun of the 3rd person singular verb was always latent (*mustatir* 'hidden'). Hence *qāma/yaqūmu* only means 'he stood/stands' by default, when not followed by an overt agent. Western segmentation of these verbs as *qām+a*, fem. *qām+at* gives a wrong impression of the status of these endings: for the Arabs they are merely gender markers, unlike the true agent suffixes in *qum+tu/qum+nā* 'I stood/we stood', etc. The same applies to the 3rd person prefixes *ya-* [masc.] and *ta-* [fem.] in the imperfect.

The logical priority of agents and subjects was disputed. Some argued that the agent is logically prior to the subject (*mubtadaʾ bihi* 'thing started with'; → *ibtidāʾ*) because it has a formal operator (*ʿāmil*), i.e. the verb, while the subject has no overt operator. The opposing view was that the subject has priority because it remains a subject even in inversion (*fī d-dāri rajulun* 'in the house [is] a man'), while an agent becomes a subject when fronted. Expressed so briefly here, the argument may seem trivial, but it had pedagogical consequences affecting the arrangement of material, and it was also part of the larger topic of predication theory.

The standard definition of the *fāʿil* as someone who performs the act denoted by the verb was an irresistible challenge. How, it was asked, can there be an agent when no act occurs, such as dying? Further, how can there be an agent when the act is prohibited or negated? This was easily answered by pointing out that at the grammatical level all verbs must have agents regardless of their meaning. But it was not a complete irrelevance: the lawyers applied similar reasoning to conclude that Qurʾānic prohibitions reduce to positive commands to refrain from the forbidden action. More subtly, the grammarians recognized that there is a category of verbs which cannot have an agent, namely verbs of existence, e.g. *kāna* 'to be' (→ *kāna wa-ʾaxawātuhā*). These do have the same syntax as transitive verbs, hence their predicate has dependent form like a direct object (*kāna zaydun qāʾiman* 'Zayd [indep.] was standing [dep.]'), but they have no *fāʿil*; instead, the element in that position is called the 'noun of *kāna*' (*ism kāna*), and the dependent complement its 'predicate' (*xabar kāna*).

(b) The semantic status of the agent noun in the sense of participle is ambiguous. In the first place, its time reference is controversial: some Kūfan grammarians called this form *fiʿl dāʾim* lit. 'lasting act[ion]', perhaps in response to a debate about the unreality of the present as the dimensionless point where the past meets the future. If so, the Kūfan term implies a rejection of that sophism. Incidentally, there can be no discussion here of the role of the *fāʿil* in the tense system, not well developed in Classical Arabic but a prominent feature of some modern dialects (→ participle; → aspect).

Secondly, it combines both verbal and nominal properties. It does loosely correspond to both the English 'present participle' and 'agent', but, unlike the participle, it always remains a noun. So *ʾana kātibun* can represent 'I am a person writing or going to write' or 'I am a person having written, a writer'. This duality is reflected syntactically in the alternative verbal and nominal constructions seen in *ʾana kātibun al-kitāba* 'I am a person going to write the book' (marked as direct object) or *ʾana kātibu l-kitābi* 'I am the writer of the book' (marked as possessed). The lawyers were aware of the implications for contracts, as the first implies future and the second past time, but they could do little with the Qurʾānic *kullu nafsin ḏāʾiqatu l-mawti* (Q. 3/185, 21/35, 29/57), lit. 'every soul has tasted death' in the 'past' construction, where one might expect *ḏāʾiqatun al-mawta* 'is going to taste death'. In inversions of the type 'are they standing, the two Zayds?', the preposed predicate may display either partial, verbal agreement, viz. *ʾa-qāʾimun az-zaydāni* 'are the two Zayds standing [masc. sg.]?', or nominal, full agreement *ʾa-qāʾimāni z-zaydāni* [masc. du.].

(c) The inflection of the imperfect verb is ascribed to its resemblance to the agent noun, for which reason it is called *al-fiʿl al-muḍāriʿ li-sm al-fāʿil* 'the verb which resembles the agent noun' (→ *māḍī/muḍāriʿ*), referring to the striking distributional and semantic similarities between this verb and agent noun: both occur as predicates and can be prefixed with the emphatic *la-*, e.g. *ʾinna r-rajula la-yaktubu/la-kātibun* 'verily the man, indeed he writes/[is] indeed someone writing', and both occur attributively, as in *hāḏā rajulun yaktubu/kātibun* 'this is a man who writes/a man, a person writing'. As a result the verb inflects like a noun, sharing two of the three nominal case endings.

(d) The agent noun is central to the theory of predication. The theory is based on two axioms: first, as there is no copula, the union of subject and predicate is marked by a referential pronoun in the latter; and second, verbs function exclusively as predicates of their agents. Moreover, all verbs incorporate an agent pronoun, either implicit (*mustatir*, see above), as in *qāma r-rijālu* 'the men stood [Ø-marker]', or externalized, as in *ar-rijālu qāmū* 'the men, they stood [masc. pl.]'. Because agent nouns have verbal meaning they also are marked for the gender and number of their agent, as in *hiya qāʾimatun* 'she is a person standing [fem. sg.]' (remember that the *ism fāʿil* is by definition a noun: even attributively it remains a noun, in apposition, e.g. *marartu bi-rajulin kātibin* 'I passed by a man, a person writing [masc. sg. obl.]').

The grammarians generalized the number and gender agreement in verbs and agent nouns to explain the agreement of a third category, the qualifiers we call adjectives (→ *ṣifa*). These, like the *fāʿil*, also alternate with verbs, e.g. *wajhun ḥasanun/yaḥsunu* 'a beautiful face/face which is beautiful', and between pseudoverbal syntax, as in *ḥasanun wajhan* 'beautiful as to face [dep.]'

(cf. *kātibun al-kitāba* 'writing the book'), and pseudonominal, as in *ḥasanu l-wajhi* 'beautiful of face [obl.]' (cf. *kātibu l-kitābi* 'the writer of the book'). For this they are termed 'the adjective assimilated [by the speaker] to the agent noun' (*aṣ-ṣifa al-mušabbaha bi-sm al-fāʿil*). It is important to stress that this is not a subset of adjectives; it covers all adjectives. Since they are morphologically related to agent nouns in being deverbal (*muštaqq* 'derived'), adjectives too are assumed to contain a pronoun reference to their head which accounts for their agreement. Significantly, agent nouns and adjectives are virtually in complementary distribution, thus *kātib* 'person writing, writer' has no corresponding adjective and *ḥasan* 'beautiful' no corresponding agent noun. The symmetry is completed by the occurrence of agent nouns in attributive adjectival positions and of adjectives in isolation as nouns, e.g. *al-ḥasanu* 'the beautiful [one]'.

The pronominal trace in verbs, agent nouns, and adjectives acts as the copula in equational sentences, except when the predicate is not deverbal, e.g. *zaydun ʾaxūka* 'Zayd is your brother' (subject and predicate are identical), or is locative, e.g. *zaydun fī d-dāri* 'Zayd is in the house' (mostly explained as elliptical for 'Zayd [is situated] in the house', using the agent noun *mustaqirrun* 'person situated'). The theological implications of all this cannot be explored here: it must suffice to allude to the problem of God's attributes. The theologians were aware of the risk of hylomorphic dualism in predicating qualities of God through verbs, since these are a compound of verb (scil. 'accident') and agent pronoun (scil. 'substance'). But the pronoun trace in the agent noun and adjective is never realized except as agreement, and since moreover the agent noun is temporally and aspectually indeterminate, the theologians consciously preferred it over the verb for enumerating and discussing God's attributes.

BIBLIOGRAPHICAL REFERENCES

All the material above comes from the works listed here, which provide easy and direct access to the primary sources.

Badawi, El-Said, Michael G. Carter, and Adrian Gully. 2004. *Modern Written Arabic: A comprehensive grammar*. London and New York: Routledge.

Carter, Michael G. 1981. *Arab linguistics: An introductory classical text with translation and notes*. Amsterdam: J. Benjamins.

——. 1995. "Predication tests, copula, and a possible connection with Ašʿarism". *Budapest Studies in Arabic* 13–14.23–36.

Fleisch, Henri. 1961, 1979. *Traité de philologie arabe*. 2 vols. Beirut: Imprimerie Catholique.

Howell, Mortimer Sloper. 1886–1911. *A grammar of the Classical Arabic language translated and compiled from the works of the most approved native or naturalized authorities*. Allahabad: Northwestern Provinces and Oudh Govt. Press.

Reckendorf, Hermann. 1898. *Die syntaktischen Verhältnisse des Arabischen*. Leiden: E.J. Brill. (Repr., 1967.)

——. 1921. *Arabische Syntax*. Heidelberg: Walter.

Suleiman, Yasir. 1999. *The Arabic grammatical tradition: A study in* taʿlīl. Edinburgh: Edinburgh University Press.

Versteegh, Kees. 1995. *The explanation of linguistic causes: Az-Zaǧǧāǧī's theory of grammar*. Amsterdam and Philadelphia: J. Benjamins.

Wright, William. 1896–1898. *A grammar of the Arabic language*. 3rd ed. Rev. W. Robertson Smith and M.J. de Goeje. Cambridge: Cambridge University Press. (Many reprints.)

MICHAEL G. CARTER (Sydney University)

Ism al-fiʿl

The term *ism al-fiʿl* (pl. *ʾasmāʾ al-fiʿl*) is used for → interjections conveying the sense of a verb. Most of them denote an imperative, e.g. *nazāli* 'go down!', *ʾilayka* 'go away!', *hālumma* 'come here!'. However, some *ʾasmāʾ al-fiʿl* have the sense of a declarative sentence (*xabar*), e.g. *hayhātu* 'far from it!'. The plural form is mentioned for the first time in the 8th century, by Sībawayhi (*Kitāb* I, 102.8). The singular *ism al-fiʿl* or *ism fiʿl* occurs only in later sources. Interjections known by this term are also called *ʾasmāʾ li-l-fiʿl* (sg. *ism li-l-fiʿl*) and *ʾasmāʾ al-ʾafʿāl* (Levin 1991:247).

In the Arab grammarians' classification, *ʾasmāʾ al-fiʿl* as imperatives are divided according to derivational criteria into three groups:

i. *ʾAsmāʾ al-fiʿl* on the *faʿāli* pattern, derived from triliteral roots, e.g. *nazāli* 'go down!', *ḥaḏāri* 'beware, be cautious!', *naḏ̣āri* 'look!' (Sībawayhi, *Kitāb* II, 34.11–35.8).
ii. *ʾAsmāʾ al-fiʿl* formed by a combination of a preposition and a genitive pronoun of the 2nd person, e.g. *ʾilayka* 'go away!'; *ʿalayka zaydan*, *ʿindaka zaydan, dūnaka zaydan* 'seize Zayd!' (Sībawayhi, *Kitāb* I, 105.15–107.10). This category of *ʾasmāʾ al-fiʿl* is called by some grammarians *ʾiġrāʾ* lit. 'incitement' (Zajjājī, *Jumal* 144; Ibn

al-ʾAnbārī, *ʾInṣāf* 99.15–16, *ʾAsrār* 67.2–5). As a technical term *ʾiġrāʾ* means 'words inciting to do a certain act'.

iii. *ʾAsmāʾ al-fiʿl* as isolated forms, e.g. *ḥayya* 'come!' (Zamaxšarī, *Mufaṣṣal* 63.1), *ḥayahalla* 'come quickly!' (Sībawayhi, *Kitāb* I, 102.10), *ʾīhi* 'tell more!' (Zamaxšarī, *Mufaṣṣal* 61.13). These isolated forms are called *ʾaṣwāt* (sg. *ṣawt*) by some grammarians (Levin 1991:249). This term is rarely used by Sībawayhi (*Kitāb* II, 40.22–41.8, 49.3–5).

Some interjections belonging to the category of *ʾasmāʾ al-fiʿl* take the place of a declarative sentence: *hayhātu* 'far from it!' (Zamaxšarī, *Mufaṣṣal* 61.18), *šattāna* 'how different is...!' (Zamaxšarī, *Mufaṣṣal* 61.18), *ʾuffin* (an interjection expressing anger; Zamaxšarī, *Mufaṣṣal* 61.20), *ʾawwah* (an interjection expressing pain; Zamaxšarī, *Mufaṣṣal* 61.20).

According to some grammarians, the meaning of the term *ʾasmāʾ al-fiʿl* is 'the proper names of the verbs', which derives from their notion that certain verbs have a proper name. Sometimes, instead of using a given verb, the speaker uses its proper name as a sign denoting this verb, just as the proper name of a given person is a sign denoting that person. The form *ḥayyahala*, for instance, is the proper name of the verb *ʾīti* 'come!', just as the name *ḥassānu* is the proper name of a given person. Hence, *ḥayyahala*, when used by the speaker, denotes the verb *ʾīti*, just as *ḥassānu* denotes a person known by this name. For this interpretation of the term *ʾasmāʾ al-fiʿl*, see Mubarrad (*Muqtaḍab* III, 202.1–4); Ibn Xaššāb (*Murtajil* 248.1–11); Ibn Yaʿīš (*Šarḥ* I, 494.17–19, ed. Jahn = *Šarḥ* IV, 25.12–14, Cairo ed.; *Šarḥ* I, 496.14–15, ed. Jahn = *Šarḥ* IV, 29.6–7, Cairo ed.); Ibn ʾAbī r-Rabīʿ (*Basīṭ* 163.17–164.1; *Ḍabṭ* 347.2–8, 348.12–349.1). For a translation of some of the excerpts mentioned in the above references, see Levin (1991:249–251).

The interpretation of the term *ʾasmāʾ al-fiʿl* by later grammarians shows that Wright's editor is mistaken when he says that *ʾasmāʾ al-fiʿl* are called thus because they have 'a certain verbal force' (Wright 1951:296). Lane is also mistaken when he calls various forms of *ʾasmāʾ al-fiʿl* 'verbal noun' (1863–1893:V, 2146) or 'imperative verbal noun' (1863–1893:I, 86, 305; II, 534).

Since the Arab grammarians hold that *ʾasmāʾ al-fiʿl* are proper names of verbs, they classify them as nouns (Sībawayhi, *Kitāb* I, 102.17–18; Mubarrad, *Muqtaḍab* III, 202.1–4). As-Suyūṭī (*ʾAšbāh* III, 2.8–9) quotes from ʾAbū Ḥayyān (13th century) the exceptional view of Ibn Ṣābir that the *ʾasmāʾ al-fiʿl* form a special part of speech called *xālifa* 'the word which takes the place of verbs'.

In the grammarians' view, words belonging to the category of *ʾasmāʾ al-fiʿl* are nouns conveying the sense of a verb (Sībawayhi, *Kitāb* I, 102.6–7, 102.23–24; II, 38.12–13), and the place occupied by them in speech is that of a verb (Sībawayhi, *Kitāb* I, 102.6–7). Some grammarians emphasize that *ʾasmāʾ al-fiʿl* are not *maṣādir* 'verbal nouns' (Sībawayhi, *Kitāb* I, 102.23–24; Mubarrad, *Muqtaḍab* III, 202.1–3). Every *ism fiʿl* conveys the sense of the verb named by it; *ḥayyahala* 'come!', for instance, is the *ism* of the imperative *ʾīti* (Sībawayhi, *Kitāb* I, 102.9–11).

ʾAsmāʾ al-fiʿl of verbs that take an accusative as a direct object can also govern a direct object (Sībawayhi, *Kitāb* I, 102.7–16; Ibn as-Sarrāj, *ʾUṣūl* I, 141.2–9), as in the examples *hālumma zaydan* 'bring Zayd!' (Sībawayhi, *Kitāb* I, 102.9) and *ḥayyahala ṯ-ṯarīda* 'come quickly to have soup!' (Sībawayhi, *Kitāb* I, 102.10). Such *ʾasmāʾ al-fiʿl* can take an accusative pronoun as a direct object, as in *tarākihā* 'leave them!' (Sībawayhi, *Kitāb* I, 102.13) and *manāʿihā* 'defend them!' (Sībawayhi, *Kitāb* I, 102.15). An *ism fiʿl* of an intransitive verb cannot govern a direct object (Sībawayhi, *Kitab* I, 102.7–9; Ibn as-Sarrāj, *ʾUṣūl* I, 141.2–4), as illustrated by the examples *ṣah* 'shut up!' and *mah* 'stop!' (Sībawayhi, *Kitāb* I, 102.7).

In Sībawayhi's view, *ʾasmāʾ al-fiʿl* on the pattern *faʿāli*, e.g. *nazāli*, are feminine nouns (Sībawayhi, *Kitāb* II, 37.21–22). Some of the *ʾasmāʾ al-fiʿl* that form a combination of a preposition and a 2nd person genitive pronoun can take a direct object, as in *ʿalayka zaydan, ʿindaka zaydan*, and *dūnaka zaydan* 'seize Zayd!' (Sībawayhi, *Kitāb* I, 105.15–107.10; see Sībawayhi, *Kitāb* I, 105.15–17; Ibn as-Sarrāj, *ʾUṣūl* I, 141.11–16). Some of these *ʾasmāʾ al-fiʿl* can govern a preposition + genitive, as in the examples *ʾilaykum ʿannī* 'go away from me!' (ʾIṣbahānī, *ʾAġānī* II, 36.31), *ʿalaykumā bi-n-nāqati llatī ṣifatuhā kaḏā* 'go, the two of you, to the she-camel whose description is such and such' (ʾIṣbahānī, *ʾAġānī*

II, 182.17), and *ʿalaykum bi-n-nāri* ‘come to the fire!’ (ʾIṣbahānī, *ʾAġānī* XVI, 109.10).

Most of the *ʾasmāʾ al-fiʿl* are indeclinable forms that can denote any number or gender of the 2nd person, e.g. *ḥayyahala t-tarīda* ‘come [pl.] quickly to have soup’ (Sībawayhi, *Kitāb* I, 102.10) and *ḥayya ʿalā ṣ-ṣalāti* ‘come [pl.] to prayer!’ (Zamaxšarī, *Mufaṣṣal* 63.1). However, in certain dialects the form *hālumma* was declined like an imperative (Sībawayhi, *Kitāb* I, 106.24–25). The 2nd person genitive pronoun suffix in those *ʾasmāʾ al-fiʿl* that combine a preposition with a pronominal suffix is declinable for gender and number, e.g. *ʾilayka* ‘go away!’ (Sībawayhi, *Kitāb* I, 105.23) and *ʾilaykum ʿannī* ‘go away from me!’ (ʾIṣbahānī, *ʾAġānī* II, 36.31).

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Ibn ʾAbī ar-Rabīʿ, *Basīṭ* = ʾAbū l-Ḥusayn ʿUbaydallāh ibn ʾAḥmad al-Qurašī al-ʾIšbīlī as-Subtī Ibn ʾAbī r-Rabīʿ, *al-Basīṭ fī šarḥ Jumal az-Zajjājī*. Ed. ʿAyyād ibn ʿĪd aṯ-Ṯābītī. 2 vols. Beirut: Dār al-Ġarb al-ʾIslāmī, 1986.

——, *Ḍabṭ* = ʾAbū l-Ḥusayn ʿUbaydallāh ibn ʾAḥmad al-Qurašī al-ʾIšbīlī as-Subtī Ibn ʾAbī r-Rabīʿ, *al-Mulaxxaṣ fī ḍabṭ qawānīn al-ʿArabiyya*. Ed. ʿAlī b. Sulṭān al-Ḥakamī. N.p., 1985.

Ibn al-ʾAnbārī, *ʾAsrār* = Kamāl ad-Dīn ʾAbū l-Barakāt ʿAbd ar-Raḥmān ibn Muḥammad Ibn al-ʾAnbārī, *ʾAsrār al-ʿarabiyya*. Ed. Christian Friedrich Seybold. Leiden: E.J. Brill, 1886.

——, *ʾInṣāf* = Kamāl ad-Dīn ʾAbū l-Barakāt ʿAbd ar-Raḥmān ibn Muḥammad Ibn al-ʾAnbārī, *Kitāb al-ʾinṣāf fī masāʾil al-xilāf bayna n-naḥwiyyīna l-Baṣriyyīna wa-l-Kūfiyyīna*. Ed. Gotthold Weil. Leiden: E.J. Brill, 1913.

Ibn as-Sarrāj, *ʾUṣūl* = ʾAbū Bakr ibn as-Sarī Ibn as-Sarrāj, *Kitāb al-ʾuṣūl fī n-naḥw*. Ed. ʿAbd al-Ḥusayn al-Fatlī. 4th ed. 3 vols. Beirut: al-Resalah Publishers.

Ibn Xaššāb, *Murtajil* = ʾAbū Muḥammad ʿAbdallāh ibn ʾAḥmad Ibn Xaššāb, *al-Murtajil*. Ed. ʿAlī Ḥaydar. Damascus, 1972.

Ibn Yaʿīš, *Šarḥ* = Muwaffaq ad-Dīn ʾAbū l-Baqāʾ Yaʿīš Ibn Yaʿīš, *Šarḥ al-Mufaṣṣal*. Ed. Gustav Jahn. 2 vols. Leipzig, 1882–1886. (Cairo ed. 10 vols. Cairo: ʿĀlam al-Kutub.)

ʾIṣbahānī, *ʾAġānī* = ʾAbū al-Faraj al-ʾIṣbahānī, *al-ʾAġānī*. 20 vols. in 10. Repr. of Būlāq ed., Beirut: Ṣalāḥ Yūsuf al-Xalīl and Dār al-Fikr li-l-Jamīʿ.

Mubarrad, *Muqtaḍab* = ʾAbū l-ʿAbbās Muḥammad ibn Yazīd al-Mubarrad, *Kitāb al-muqtaḍab*. Ed. Muḥammad ʿAbd al-Xāliq ʿUḍayma. 4 vols. Cairo: Lajna ʾIḥyāʾ at-Turāṯ al-ʾIslāmī, 1385–1388 A.H.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. Ed. Hartwig Derenbourg. 2 vols. Paris: Imprimerie Nationale, 1881–1889.

Suyūṭī, *ʾAšbāh* = Jalāl ad-Dīn ʾAbū l-Faḍl ʿAbd ar-Raḥmān ibn ʾAbī Bakr as-Suyūṭī, *al-ʾAšbāh wa-n-naḍāʾir fī n-naḥw*. 2nd ed. 4 vols. Hyderabad, 1359–1361 A.H.

Zajjājī, *ʾĪḍāḥ* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān ibn ʾIsḥāq az-Zajjājī, *al-ʾĪḍāḥ fī ʿilal an-naḥw*. Ed. Māzin al-Mubārak. Cairo: Dār al-ʿUrūba, 1959.

——, *Jumal* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān ibn ʾIsḥāq az-Zajjājī, *Kitāb al-jumal fī n-naḥw*. Ed. ʿAlī Tawfīq al-Ḥamad. Beirut and Irbid: Muʾassasat ar-Risāla and Dār al-ʾAmal, 1984.

Zamaxšarī, *Mufaṣṣal* = Jār Allāh ʾAbū l-Qāsim Maḥmūd ibn ʿUmar az-Zamaxšarī, *al-Mufaṣṣal fī n-naḥw*. Ed. Jens Peter Broch. 2nd ed. Christianiae: Libraria P.T. Mallingii, 1879.

SECONDARY SOURCES

Lane, Edward W. 1863–1893. *Arabic-English lexicon*. 8 vols. London and Edinburgh: Williams and Norgate.

Levin, Aryeh. 1991. “The category of *ʾasmāʾ al-fiʿl* in Arabic grammar”. *Proceedings of the Colloquium on Arabic Grammar,* ed. Kinga Dévényi and Tamás Iványi, 247–256. Budapest: Eötvös Loránd University and Csoma de Kőrös Society.

Wright, William. 1951. *A grammar of the Arabic language*. 3rd ed. 2 vols. Cambridge: Cambridge University Press.

ARYEH LEVIN (Hebrew University of Jerusalem)

ʾIsnād

The term *ʾisnād* and words derived from it indicate in Arabic grammar the connection between a noun and its predicate, or the act of assigning a predicate to a subject (Levin 1981:157). The central point in the analysis of the sentence in Arabic grammar is the distinction between nominal (*ʾismiyya*) and verbal (*fiʿliyya*) sentences, each with their own constituents. The nominal sentence is built on a topic/comment structure, with the topic (*mubtadaʾ*; → *ibtidāʾ*) and the comment (→ *xabar*) as basic constituents, whereas the verbal sentence consists of a verb (→ *fiʿl*; this term is synonymous with the name of one of the three parts of speech) and an agent (→ *fāʿil*). This means that in traditional Arabic syntax there is no room for a notion of ‘subject’ (cf. Ayoub and Bohas 1983). In the later tradition of linguistic analysis, which was heavily influenced by Greek logic, terms like → *mawḍūʿ* ‘subject’ and → *maḥmūl* ‘predicate’ are sometimes mentioned, but these remained alien to the indigenous tradition of analyzing sentence structure.

Yet, from an early stage onward, there was a definite awareness among Arab grammarians about the basic resemblance between the constituents of the nominal sentence and the constituents of the verbal sentence. This resemblance focused on the notion of 'predication', for which such verbs as *ʾaxbara* and *ḥaddaṯa* were used. These terms were semantic in nature: they expressed the fact that the basic purpose of a sentence is to provide new information about something already known. Alongside this terminology, there was another set of terms, derived from the verb *ʾasnada* 'to make something lean on something', which stressed the syntactic relationship between the main constituents in the sentence. This set was possibly connected with early Arabic logic. In his exposé of Aristotle's *De interpretatione*, Ibn al-Muqaffaʿ (d. 2nd half 8th century C.E.) uses *musnad ʾilayhi* in the sense of 'predicate' (Talmon 1987:215; Goldenberg 1988:45). He says about the verb (in his terminology the *ḥarf*): *lā yakūnu ʾillā maḥmūlan ʿalā ġayrihi musnadan ʾilayhi* 'it is always predicated about something else, being made to lean against it' (Ibn al-Muqaffaʿ, *Manṭiq* 28.12). Here, *musnad ʾilayhi* is the element that is made to lean upon another element, being supported by it. This other element – the subject – could then appropriately be called *musnad* 'support, prop'. The term *musnad* also occurs in the works of a later philosopher, al-Fārābī (d. 339/950), who states (*ʾAlfāḏ̣* 57) that *musnad/musnad ʾilayhi* for subject/predicate is synonymous with *muxbar ʿanhu/muxbar bihi*, *xabar*, and *mawṣūf/ṣifa*, the latter used especially in theological discourse (Versteegh 1995:67).

In grammatical treatises, the terms derived from *ʾasnada* also occurred, but they were superposed on the usual terms for the syntactic relations in nominal and verbal sentences. In a sentence like (1)

(1) *zayd-un ḍaraba*
'Zayd hit'

the topic *zaydun* acts as the support for the comment, which may be a verb as in (1), or a noun as in (2)

(2) *zayd-un rajul-un*
'Zayd [is] a man'

In a verbal sentence like (3)

(3) *ḍaraba zayd-un*
'Zayd hit'

the agent *zaydun* is still credited with the act of hitting. This means that at a deeper, semantic level, the agent and the topic are identical. In general, one could say that the grammarians were only interested in the analysis of the syntactic level. Nonetheless, they did speculate about the nature of this identity and the nature of the tie between the constituents in both types of sentence. This tie was the *ʾisnād* 'support, the act of making lean against', and the two participants in the *ʾisnād* relationship were called *musnad* and *musnad ʾilayhi*.

Sībawayhi dedicates a separate but unfortunately rather short and not very informative chapter to the *musnad* and the *musnad ʾilayhi* (*Kitāb* I, 7). He begins by saying that the *musnad* and the *musnad ʾilayhi* need each other (*lā yastaġnī wāḥid minhumā ʿan al-ʾāxar*), and then exemplifies this mutual need with the example of the *mubtadaʾ* and the *mabniyy ʿalayhi* in *ʿabdallāh ʾaxūka* "Abdallāh is your brother' (for the term *mabniyy ʿalā* cf. Levin 1985). The problem arises when he then gives a second example, *yaḏhabu zaydun* 'Zayd goes away', in which the verb *yaḏhabu* has the same need for the following noun as the first noun for the second noun in the first example. This is the basis for the identification of the *musnad* as topic/verb and the *musnad ʿilayhi* as comment/agent.

Elsewhere, Sībawayhi (*Kitāb* I, 256.17–18) uses *musnad* in the statement *fa-l-mubtadaʾ musnad wa-l-mabniyyu ʿalayhi musnad ʾilayhi* 'the topic [or: the word with which the sentence begins] is being leaned upon and the [constituent] that is built on it [i.e. the predicate] is made to lean upon it' (Goldenberg 1988:43). From this statement it is not clear, either, what his position vis-à-vis these two terms is when applied to a verbal sentence. According to Levin (1981) and others, *musnad* is always literally the first constituent in the sentence, which implies that in a verbal sentence the *musnad* is the verb and the *musnad ʾilayhi* the agent. But Goldenberg (1988:44) argues that it is more likely that both *xabar* and *fiʿl* are subsumed under *musnad ʾilayhi* because they both function as predicates to another word, which is then the *musnad* 'the constituent upon which something else [sc. the predicate] is made to lean'. He supports this with a phrase from Ibn as-Sarrāj (*ʾUṣūl* I, 63.7–8), *al-mubtadaʾ yubtadaʾu fīhi bi-l-ism* *al-muḥaddaṯ ʿanhu qabla l-ḥadīṯ* ‘in the topic, one begins with the noun of which something is reported before the report’, whereas according to Ibn as-Sarrāj, *al-fāʿil mubtadaʾ bi-l-ḥadīṯ qablahu* ‘with the agent, one begins with the report before it’. At the very least, this quotation confirms that there is a semantic resemblance between the topic and the agent, which are both constituents about which something else is reported (*muḥaddaṯ*); in other words, they are both constituents on which a predicate is made to lean (*musnad ʿilayhi*).

In fact, these terms are hardly ever used by Sībawayhi. According to Troupeau (1976:112), *musnad* is used four times and *musnad ʾilayhi* three times, and *ʾisnād* does not occur at all in the *Kitāb*; the verb *ʾasnada ʾilā* is used only once. Talmon (2003:163–166) believes that this terminology belongs to an earlier stage of grammar, called by him ‘the Iraqi school’, which was transcended by Sībawayhi’s distinction between nominal and verbal sentences. This probably implies that the few instances of the *ʾasnada* set in the *Kitāb* are traces from the earlier tradition, which somehow remained in the *Kitāb* but did not play an important role. The same terms occur only twice in al-Xalīl’s *Kitāb al-ʿayn* (VII, 68, 228; VIII, 11; Talmon 1997:209).

In al-Farrāʾ’s *Maʿānī l-Qurʾān*, terms derived from *ʾasnada* are used somewhat more frequently: *ʾasnada* is used nine times and *musnad ʾilā* four times (Kinberg 1996:266–267). According to Owens (1990:103–107, 249–250), al-Farrāʾ uses these terms in two different senses. In nominal constructions, the relationship between an annexed noun and the noun to which it is annexed is expressed by the verb *ʾasnada*, for instance, when he compares the two expressions *huwa rajul dūna-ka* ‘he is a man inferior to you’ and *huwa rajul dūnun* ‘he is an inferior man’ (al-Farrāʾ, *Maʿānī* I, 119.5); the former is said to have been annexed to the pronoun (*ʾusnida ʾilā šayʾ*). For the same relationship al-Farrāʾ also uses, and more frequently, the term → *ʾiḍāfa*. The second sense in which *ʾasnada* is used by al-Farrāʾ is more related to its use in Sībawayhi’s *Kitāb*, viz. to indicate the relationship between a verb and its agent. When comparing the expressions *ḍāqa ḏarʿī bihi* ‘my ability [to deal] with him became narrow [i.e. ‘I became fed up with him]’ and *ḍiqta bihi ḏarʾan* ‘I became annoyed with him’ (*Maʿānī* I, 79.9; cf. II, 270.1), al-Farrāʾ states that in the latter expression the notion of *ḍīq* is made to lean against the 2nd person (*jaʿalta ḍ-ḍīq musnadan ʾilayka*). Owens (1990:249) is probably right in regarding *musnadan ʾilayhi* here as the description of a process, rather than a fixed grammatical term. The analysis of all passages in which this terminology is used by al-Farrāʾ makes it clear that he regards the first word (the verb) as the constituent that acts as support for the second word (the agent), in line with the way Sībawayhi uses these terms.

In later grammar, *ʾisnād* is used in almost the same sense as *ʾixbār*, for instance by al-Jurjāni (*Muqtaṣid* I, 76.19–20; Goldenberg 1988:48), the only difference being that *ʾixbār* is usually restricted to cases where the truth value of a statement can be determined, so that imperatives do not commonly fall under the definition of *ʾixbār*, yet they still represent a case of *ʾisnād*, i.e. a connection between a subject and its predicate. The basic identity between the topic and the agent, on one hand, and the verb and the predicate, on the other, is stressed by many authors, e.g. by Ibn as-Sarrāj: *al-fāʿil muḍāriʿ li-l-mubtadaʾ min ʾajli ʾannahumā jamīʿan muḥaddaṯ ʿanhumā* ‘the agent resembles the topic in that they both are being predicated about’ (*ʾUṣūl* I, 39.10–11; cf. Zajjājī, *Jumal* 48.6; Goldenberg 1988:47).

However, with respect to the names of the two constituents of the *ʾisnād* connection, Goldenberg (1988:45; cf. Levin 1981:151–153) notes that their meaning was switched in the later grammatical tradition. Originally, *musnad* indicated either the subject or, according to another interpretation of its early use, the first element in the sentence, and *musnad ʾilayhi* the predicate that was ‘made to lean’ on the subject, or the second element in the sentence. The two terms may, however, be interpreted in a diametrically opposed way: *musnad* could be the constituent that is being made to lean against something else; *musnad ʾilayhi* would then be the constituent upon which this other constituent is made to lean. In this interpretation, *musnad* corresponds to the predicate and *musnad ʾilayhi* is the subject. Levin (1981:151–157) has shown that this new interpretation of the two terms can first be attested in the 10th century C.E., for instance in as-Sīrāfī’s (d. 368/978) commentary on the

Kitāb (*Šarḥ* II, 59) where he explicitly refers to a difference of opinion about the interpretation of the terms. He states that in his view the best interpretation is the one according to which the *musnad ʾilayhi* is the *muḥaddaṯ ʿanhu*, i.e. the thing about which something is told, and the *musnad* the *ḥadīṯ* or *xabar*, i.e. the thing that is told. He adds that the *musnad ʾilayhi* is thereby identical with both the topic and the agent, and the *musnad* with the comment and the verb. Similar statements are made by many other later grammarians, e.g. by ʾAbū Ḥayyān (*Manhaj* 4.13–15). The inherent ambiguity of the terminology probably made these terms less than useful for syntactic analysis (a similar ambiguity reigned in the case of the two terms *muḍāf* and *muḍāf ʾilayhi*; → *ʾiḍāfa*).

Bibliographical references

Primary sources

ʾAbū Ḥayyān, *Manhaj* = ʾAbū Ḥayyān Muḥammad ibn Yūsuf al-Ġarnāṭī an-Naḥwī, *Manhaj as-sālik fī l-kalām ʿalā ʾAlfiyyat Ibn Mālik*. Ed. Sidney Glazer. New Haven, Conn.: American Oriental Society, 1947.

Fārābī, *ʾAlfāḏ* = ʾAbū Naṣr Muḥammad ibn Muḥammad al-Fārābī, *Kitāb al-ʾalfāḏ al-mustaʿmala fī l-manṭiq*. Ed. Muḥsin Mahdī. Beirut: Dār al-Mašriq, 1968.

Farrāʾ, *Maʿānī* = ʾAbū Zakariyyāʾ Yaḥyā ibn Ziyād al-Farrāʾ, *Maʿānī l-Qurʾān*. Ed. Muḥammad ʿAlī an-Najjār. 3 vols. Cairo: ad-Dār al-Miṣriyya, 1966–1972.

Ibn al-Muqaffaʿ, *Manṭiq* = ʾAbū Muḥammad (ʾAbū ʿAmr) ʿAbdallāh Ibn al-Muqaffaʿ, *al-Manṭiq*. Ed. Muḥammad Taqī Danešpažuh. Teheran: Imperial Iranian Academy of Philosophy, 1398/1978.

Ibn as-Sarrāj, *ʾUṣūl* = ʾAbū Bakr ibn as-Sarī Ibn as-Sarrāj, *Kitāb al-ʾuṣūl fī n-naḥw*. Ed. ʿAbd al-Ḥusayn al-Fatlī. 3 vols. Beirut: Muʾassasat ar-Risāla, 1985.

Jurjānī, *Muqtaṣid* = ʾAbū Bakr ʿAbd al-Qāhir ibn ʿAbd ar-Raḥmān al-Jurjānī, *Kitāb al-muqtaṣid fī šarḥ al-ʾĪḍāḥ*. Ed. Kāẓim Baḥr al-Marjān. 2 vols. Baghdad: Wizārat aṯ-Ṯaqāfa wa-l-ʾIʿlām, 1982.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. 2 vols. Bulaq, 1316 A.H. (Repr., Baghdad: Al-Muthanna Library, n.d.)

Sīrāfī, *Šarḥ* = ʾAbū Saʿīd al-Ḥasan ibn ʿAbdallāh as-Sīrāfī, *Šarḥ Kitāb Sībawayhi*, II. Ed. Ramaḍān ʿAbd at-Tawwāb. Cairo: al-Hayʾa al-Miṣriyya al-ʿĀmma li-l-Kitāb, 1990.

Xalīl, *ʿAyn* = ʾAbū ʿAbd ar-Raḥmān al-Xalīl ibn ʾAḥmad al-Farāhīdī, *Kitāb al-ʿayn*. Ed. Mahdī al-Maxzūmī and ʾIbrāhīm as-Sāmarrāʾī. 8 vols. Beirut: Muʾassasat al-ʾAʿlamī li-l-Maṭbūʿāt, 1988.

Zajjājī, *Jumal* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān Ibn ʾIsḥāq az-Zajjājī, *Kitāb al-jumal*. Ed. Muhammad Ben Cheneb. Paris: Klincksieck, 1957.

Secondary sources

Ayoub, Georgine and Georges Bohas. 1983. “Les grammairiens arabes, la phrase nominale et le bon sens”. *The history of linguistics in the Near East*, ed. Konrad Koerner, Hans-Josef Niederehe, and Kees Versteegh, 31–48. Amsterdam: J. Benjamins.

Goldenberg, Gideon. 1988. “Subject and predicate in Arab grammatical tradition”. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 138.39–73.

Kinberg, Naphtali. 1996. *A lexicon of al-Farrāʾ's terminology in his* Qurʾān *commentary with full definitions, English summaries, and extensive citations*. Leiden: E.J. Brill.

Levin, Aryeh. 1981. “The grammatical terms *al-musnad*, *al-musnad ʾilayhi* and *al-ʾisnād*”. *Journal of the American Oriental Society* 101.145–165.

——. 1985. “The syntactic technical term *al-mabniyy ʿalayhi*”. *Jerusalem Studies in Arabic and Islam* 6.299–352.

Owens, Jonathan. 1990. *Early Arabic grammatical theory: Heterogeneity and standardization*. Amsterdam and Philadelphia: J. Benjamins.

Talmon, Rafael. 1987. “*Musnad*, *musnad ilayhi* and the early history of Arabic grammar: A reconsideration”. *Journal of the Royal Asiatic Society* 207–222.

——. 1997. *Arabic grammar in its formative age:* Kitāb al-ʿayn *and its attribution to Ḫalīl b. Aḥmad*. Leiden: E.J. Brill.

——. 2003. *Eighth-century Iraqi grammar: A critical exploration of pre-Ḫalīlian Arabic linguistics*. Winona Lake, Ind.: Eisenbrauns.

Troupeau, Gérard. 1976. *Lexique-index du* Kitāb *de Sībawayhi*. Paris: Klincksieck.

Versteegh, Kees. 1995. *The explanation of linguistic causes: Az-Zaǧǧāǧī's theory of grammar*. Amsterdam and Philadelphia: J. Benjamins.

Kees Versteegh (University of Nijmegen)

Israel

1. Introduction

Arabic is the mother tongue and the main national language of more than one million Palestinian citizens in the state of Israel. It is also a community language of hundreds of thousands of Sephardic Jews. Arabic serves as the sole official language of Israel's neighboring countries and enjoys a special status in most Muslim countries. The case of Arabic in Israel is unique. It is a minority language and is legally recognized as a second official language. Arabic was a majority language, as well as one of three official languages, in the British mandate to Palestine, until the founding of Israel, when by changed sociopolitical circumstances it became a secondary language. Arabic is learned as the first language in all Arab schools in Israel from the first through the twelfth grades, and in several institutes for teacher training.

The perception of Arabic and its status as an additional language among the Jews of Israel creates a paradox. On the one hand, the status of the language is inferior and there are only minimal requirements to learn it, but on the other hand the study of Arabic is closely connected to the army and the intelligence service. That is to say, learning Arabic as a foreign language in Israel is largely influenced by the military needs of the state (Amara and Mar'i 2002).

Spolsky and Shohamy (1999a:103) estimate that 4,500,000 Israelis have functional competence in Hebrew vs. 2,000,000 in Arabic. According to these figures, Arabic is the second major language of Israel. The overwhelming majority of Palestinians in Israel demonstrate high levels of functional competence in Hebrew because they learn it as a subject in their schools beginning in the third grade and in their daily contact with Israeli Jews in various domains of life (Amara 2002). About one million Israeli Jews have functional competence in Arabic, whether they learned it at school or as a community language, as in the case of Sephardic Jews.

The official status of Arabic in Israel is far from clear. When researchers talk about its status, they mention mainly Article 82 of the Palestine Order in Council 1922 (e.g. Landau 1987; Koplewitz 1992; Spolsky 1994; Spolsky and Shohamy 1999a, 1999b). However, to understand the real status of Arabic in Israel, one needs to consider also the influence of sources such as the Supreme Court and the attorney general, and basic laws in Israel such as the one relating to 'human dignity'.

Examining the official status of Arabic is of particular interest in Israel because Israel is defined and perceived as an ethnic national state, and Jews are the only nationality recognized by law in Israel. Arabs are legally recognized as a religious and cultural minority but not as a national minority. Rights granted to the Arabs are mainly at an individual rather than communal level (Smooha 1999).

2. The status of Arabic as an official language

The official status of Arabic was bestowed in the mandatory legislation, and Arabic was left on its pedestal after the establishment of the state of Israel. Article 82 of the Palestine Order in Council 1922 referred (under the marginal heading 'Official Languages') to English, Arabic, and Hebrew, and determined that:

> All ordinances, official notices and official forms of the government and all official notices of local authorities and municipalities in areas to be prescribed by order of the High Commissioner shall be published in English, Arabic, and Hebrew. The three languages may be used in debates and discussions in the Legislative Council and subject to any regulations to be made from time to time, in the Government offices and the Law Courts.

After Israel became independent, all mandatory legislation was retained except for a few laws. Paragraph 15(b) of the ordinance, pertaining to the use of English, was canceled, but no legislation, then or later, canceled the official status of the Arabic language as fixed in Article 82 (Kretzmer 1990; Rubinstein and Medina 1996).

The three issues dealt with in the Order in Council 1922, Article 82, are (1) obligations regarding the languages in which the central government must carry out central functions; (2) the languages in which official notices must be issued by the local authorities; and (3) the languages in which an individual is able to access public services of the central government and perhaps also the language of the workplace of public servants in government offices and the courts.

The structure of the language arrangements in these areas is not uniform. The operating language of the central government, according to law, is bilingual: there are two official languages. The operating government language of the local authorities, though, is bilingual only in those areas where it is fixed by decree. As for the accessibility of the central government's services for the public, here the normative standard prescribes bilingualism.

The status of Arabic as an official language means that there exists a clear and far-reaching communal right. It is far reaching because the official status is not just the grant of a multiethnic right but is an even more radical right: the communal right for a special allocation. The potential importance of the official status of Arabic extends from symbolic to pragmatic dimensions, economic, social, and political.

As for the symbolic dimensions of Article 82, the possible implications are even more far-reaching. The potential is embedded in law in that Arabic is an 'official' language. That is to say, not just a language, or a recognized language, nor even a language with a special status which fixes comprehensive rights and obligations in its use. An official language touches on the essence of the state, hence official bilingualism is a clear manifestation of a binational state. It appears in all the binational states in which the national communities have separate languages, as in Switzerland, Belgium, and Canada. Yet, multilingualism exists not only in multinational states. It also exists in states with a multicultural citizenry, outstanding examples being the new South Africa and India, and also Italy/southern Tirol, and Spain/Catalonia and the Basque country, although there it is limited to the area of the minority itself (Kaufman 1999; Alcock 1979; Hannum 1996; *World Directory of Minorities* 1997).

It may be assumed that the historical-statutory background greatly facilitated the maintenance of the official status of Arabic. First of all, its status as an official language was already anchored in mandatory legislation. Second, to cancel this status would have elicited a negative reaction on the part of the international community, for the partition decision of the United Nations in November 1947 directed the protection of a series of human and collective rights, including linguistic rights, in the two states – the Arab and the Jewish – that were to be established in Palestine.

However, a more comprehensive and logical explanation lies in the practical meaning of the official status of Arabic. Instead of going through a problematic process of canceling the official status of Arabic, Israel satisfied itself with a de facto emptying of Arabic's official status from almost all practical arenas in Israel. That is to say, Arabic was deleted from almost all practical sociolinguistic uses in the public language landscape. The following examples illustrate the various meanings and dimensions of the official position of Arabic.

i. An Arabic version of state legislation does indeed appear in official publications, but only many months after publication in Hebrew. This delay does not affect the validity of the legislation: the effective date of the legislation is not conditioned on its publication in law books in both languages but on "publication in the law books" (Paragraph 10 of the Rules of Order of the Government and the Courts, 1948). In any case, there is a preference for Hebrew, which results from the force of the interpretive principle as set out in Paragraph 24 of the Law of Interpretation (1981), according to which the two versions do not have equal weight, but rather the legislation is to be interpreted according to the original language in which it was passed (Rubinstein and Medina 1996:99).
ii. In practice, one cannot appeal in Arabic to the central government authorities, which are above the level of local branches in the Arab settlements. There is no basic organization that enables this. Many official forms, perhaps most of them, do not appear in the Arabic version.
iii. Individuals may testify in Arabic at their own trial or that of someone else, and they are entitled to an interpreter if they are a defendant (see *Criminal Procedure Law* 1982), but in practice, one cannot conduct a criminal or civil procedure mainly in Arabic, or submit court documents in Arabic.
iv. Official announcements of the government, especially those of the local authorities where the population is Jewish or mixed, appear almost always only in Hebrew (Barzilai 2003). Until recently, the road signs were almost all in Hebrew (sometimes English was added, but only rarely Arabic). Street signs and by-laws of the 'Jewish' authorities and the mixed ones are still given only in Hebrew, except for the cities of Jerusalem and partially also Haifa and Acre.

Over the course of the years, quite a bit of legislation has been added to Article 82 dealing with the question of language in the framework of its arrangements. This legislation deals, among other things, with the obligation of various bodies to publish public notices regarding procedures and safety instructions and with the language of examinations and conditions for qualification in certain regulated professions and even conditions for naturalization. This legislation does not reveal any uniform approach with reference to Arabic. On the one hand, there is occasionally an explicit obligation to publish in Arabic as found, for instance, in Article 76 of the Knesset and Prime Minister Elections [Consolidated Version], 1969 (proclaiming that the ballots must be in both languages). On the other hand, there is often an obligation to publish in Hebrew (or a requirement of sufficient knowledge of Hebrew) without any parallel obligation as to Arabic. A characteristic example appears in Paragraph 9(b) of the Decree of Local Councils A, 1950: "A notice for a public tender shall be made by its publication in at least two newspapers printed in Israel, of which at least one is in Hebrew". In other words, there is a requirement for publication in Hebrew, and against this there is only the possibility for publication in Arabic (or Russian, or some other language). Another example with much symbolic significance expressing a clear preference for Hebrew is the instruction in Article 5 (a) (4) of the Citizenship Law, 1952, that requires "a certain knowledge of Hebrew" as a condition for naturalization.

The practical significance of the status of Arabic has been mainly confined to the domain of protection of the internal life of the Arab minority, especially the right to education in Arabic. There exists a full-fledged public education system (mainly at the primary and secondary levels), in which more than a half million Arab pupils receive their education in Arabic. The right to operate private Arab schools is being upheld, and since the 1980s the schools have been, to a large extent, state financed.

The language situation in Israel has nevertheless changed somewhat in recent years. Developments have taken place in several contexts. The government entities involved in these developments are especially the Supreme Court, the attorney general, and the government legal advisor – all of whom have been motivated for their involvement by the activities of human rights organizations and reactions from members of the minority group. Thus, for example, the Supreme Court certified the right of citizens to publish in Arabic alone on bulletin boards of local authorities, even of those in which the majority or even the entire body of residents are Jews. Similarly, following petitions to the court from Adalah (the Legal Center for the Defense of Arab Minority Rights in Israel) and the Society for Civil Rights, there occurred a strengthening of the appearance of Arabic on road signs on interurban highways – including the commitment to go over to fully bilingual signposts within a relatively short period of years – and also the wider use of Arabic in the signposts on streets and highways in two mixed cities, Haifa and Acre. The attorney general entered a non-negligible change when in 1995 he put into force a requirement for publication also in Arabic of announcements of tenders published by government offices. The government legal advisor also intervened a number of times in defense of the Arabic language, as when he made clear that government offices were required to publish in Arabic notices inviting citizen groups to submit requests for state support.

The above-mentioned developments are joined to a new understanding that began to crystallize in recent years, according to which the source of requirements in the matter of Arabic is not in Article 82 of the Order in Council alone (or the legislation added to it). In recent judgments, a more general and intangible requirement was pointed out – the obligation derived from 'human dignity' and the linkage between human dignity and cultural affiliation.

3. Conclusion

Although recent legal developments regarding the status of Arabic are impressive, their influence on the sociolinguistic actuality of Israel is limited. The state of Israel, and Israeli society even more so, is very far from being bilingual. More concretely, the status of 'official language', given to Arabic in Israeli law, is still empty of practical meaning at the level of public life, that is to say, in the common arena, the general society of Jews and Arabs in Israel. Hebrew is, in practice, the sole language of the society at large. This is the language in which the public front of Israel speaks in the preponderant portion of the public arena, as well as being the governmental bureaucratic language, the language of higher education, the language of most of the public electronic communication in Israel, and most important, the language of large sections of the business market open to the minority. In other words, in Israel there is an integration of a state and civil society that works almost exclusively in the language of the majority community. Together they push toward a society characterized by asymmetrical bilingualism and biculturalism: unilateral bilingualism and biculturalism of the minority (Smooha 1980; Al-Haj 1996). The main significance of the status of Arabic in Israel appears, then, not in regard to society as a whole but to the extent of the protection it affords to the internal life of the minority, especially in regard to the right to education in the minority tongue.

BIBLIOGRAPHICAL REFERENCES

Alcock, Antony. 1979. "The development of governmental attitudes to cultural minorities in Western states". *The future of cultural minorities*, ed. Antony E. Alcock, Brian K. Taylor, and John M. Welton, 102–119. New York: St. Martin's Press.

Al-Haj, Majed. 1996. *Education among the Arabs in Israel: Domination and social change.* Jerusalem: Magnes. (In Hebrew.)

Amara, Muhammad Hasan. 2002. "Hebrew among the Arabs in Israel: Sociolinguistic aspects". *Teuda* 18.85–105. (In Hebrew.)

—— and ʿAbd al-Rahman Mar'i. 2002. *Language education policy: The Arab minority in Israel.* Dordrecht: Kluwer Academic Publishers.

Barzilai, Gad. 2003. *Communities and law: Politics and cultures of legal identities.* Ann Arbor: University of Michigan Press.

Criminal Procedure Law (Consolidated Version). 1982. *Sefer ha-Xukim* [Law book]. 1982.

Hannum, Hurst. 1996. *Autonomy, sovereignty, and self-determination: The accommodation of conflicting rights.* 2nd ed. Philadelphia: University of Pennsylvania Press.

Kaufman, Elana. 1999. "The option of the Israeli state". *Seven ways: Theoretical options of the Arab status in Israel*, ed. Sara Ozacky-Lazar a.o., 201–242. Giv'at Haviva: Institute for Peace Studies. (In Hebrew.)

Koplewitz, Immanuel. 1992. "Arabic in Israel: The sociolinguistic situation of Israel's Arab minority". *International Journal of the Sociology of Language* 98.29–66.

Kretzmer, David. 1990. *The legal status of the Arabs in Israel.* Boulder, Colo.: Westview.

Landau, Jacob. 1987. "Hebrew and Arabic in the state of Israel: Political aspects of the language issue". *International Journal of the Sociology of Language* 67.117–133.

Minority Rights Group (ed.). 1997. *World directory of minorities.* London: MRG International.

Rubinstein, Amnon and Barak Medina. 1996. *The constitutional law of the State of Israel.* 5th ed. Tel Aviv: Shoken. (In Hebrew.)

Smooha, Sami. 1980. "Existing and alternative policy toward the Arabs in Israel". *Megamot* 26.7–36. (In Hebrew.)

——. 1999. *Autonomy for the Arabs.* Beit Berl: Center for Arab Studies. (In Hebrew.)

Spolsky, Bernard. 1994. "The situation of Arabic in Israel". *Arabic sociolinguistics: Issues and perspectives,* ed. Yasir Suleiman, 227–234. London: Curzon Press.

—— and Elana Shohamy. 1999a. "Language in Israeli society and education". *International Journal of the Sociology of Language* 137.93–114.

——. 1999b. *Languages of Israel: Policy, ideology and practice.* Clevedon: Multilingual Matters.

MUHAMMAD HASAN AMARA
(Bar-Ilan University)

Istiʿāra

Istiʿāra is used as a technical term in the Arabic linguistic and literary tradition for a figure of speech or mode of expression that is usually described as a metaphor in classical and modern rhetorical theory. The literal meaning of *istiʿāra* is 'borrowing', which in the given context means borrowing the name or an attribute of something to stand for something else, as in *raʾaytu ʾasadan* 'I saw a lion', *ʾasad* being borrowed to stand for a brave man, to cite the simplest form. This linguistic phenomenon may be found in Western manuals also under the heading of simile or allegory, when related to the level of a sentence. As in Western rhetoric, the concept of metaphor and its definitions vary with time and theorist. Within the mainstream theory of metaphor in the East, i.e. the school of as-Sakkākī (d. 626/1229) and his commentator al-Qazwīnī (d. 739/1338), whose works are based mainly on the thoughts of the master in the field, ʿAbd al-Qāhir al-Jurjānī (d. 471/1078), *istiʿāra* is essentially an intensified comparison in which one term is concealed. In the canonical system, the study of *istiʿāra* falls under the heading of → *majāz* 'trope', and *istiʿāra* is the most frequently discussed. *Majāz* in its turn is part of *ʿilm al-bayān*, a discipline that studies the presentation of thoughts in various degrees of explicitness and the corresponding modes of indirect expression. Usually, *istiʿāra* is treated subsequent to the section on simile, *tašbīh*.

The *balāġiyyūn*, or theorists of eloquence, use a set of terms in order to describe the mechanism and functioning of *istiʿāra*. The *mustaʿār minhu* is the thing or notion the metaphorical meaning is borrowed from, while the thing or notion the metaphorically used word is given to is called *mustaʿār lahu*. The transferred or borrowed word itself is referred to as *mustaʿār*. One might draw a parallel between *mustaʿār lahū* and 'tenor'. For both *mustaʿār minhu* and *mustaʿār*, however, 'vehicle' is the only equivalent. In a simple word metaphor like *raʾaytu ʾasadan* 'I saw a lion', the *mustaʿār* is simply regarded as the noun that represents the *mustaʿār minhu*; thus, the difference in descriptive approach might not be obvious at first sight. Another standard example, *wa-štaʿala r-raʾsu šayban* 'The head flared up in white hair' (Q. 19/4), makes it clearer. Here, the *mustaʿār minhu* is 'fire', the *mustaʿār lahū* is 'white hair', and the *mustaʿār* borrowed from 'fire' is 'flared up'. The ground or aspect of similarity is called *wajh aš-šibh* or *jāmiʿ*. There must be an indication in the situation, *qarīna ḥāliyya*, or the speech sequence itself, *qarīna maqāliyya*, that blocks the conventional meaning of the word, e.g. *raʾaytu ʾasadan yarmī* 'I saw a lion throwing'. The adjunct *yarmī* indicates a nonliteral usage of *ʾasad*.

An important difference between *istiʿāra* and the common concept of metaphor is that most theorists do not treat an expression combining both tenor and vehicle, e.g. *zaydun ʾasadun* 'Zayd is a lion', as *istiʿāra*. If you say '*zayd*', one knows that you want to assert something about that person. A sentence like *zaydun ʾasadun*, however, makes no sense when taken literally. It is perceived and interpreted immediately as a comparison. Once the compared thing is mentioned, it can no longer be imagined as belonging to the genus of the object of comparison. The identification is weakened, if not made impossible, by mentioning the entity being compared, thus bringing it as something real and existing to the consciousness of the hearer. While in the example above the metaphorical character is blocked, due to the predicative relationship between the thing compared that functions as subject of the sentence and the thing compared with as its predicate, the example *raʾaytu ʾasadan* constitutes a separate case. The comparison is not obvious, and the thing compared appears in a completely new shape. It is as if you have given someone the robes of a king. If you just give him some of the clothes while attributes of a common man remain, you cannot make him appear as a king. Here, the concept of transference, *naql*, turns into a claim of identity (*iddiʿāʾ*). Despite these differences between the canonical concepts of East and West, it should be emphasized that, contrary to a widespread assumption, from Aristotle up to Quintilian, metaphor was not seen mechanically as a shortened simile, nor did the formula 'A is B' exclude the interpretation as a simile without particle (cf. Eggs 2001).

The criterion of originality was taken into consideration when commonly known metaphors were treated not as *istiʿāra* but as similes. Several further issues are discussed: the requirements for the beauty of a metaphor; its strangeness and unfamiliarity (*ġarāba*); the relationship between metaphorical wording and lying; the nature of the *qarīna*; the problem of distinguishing between *istiʿāra* and the use of synonyms or the use of a general term instead of a specific one; the grammatical categories of metaphorically used words; and the like.

The first studies of *istiʿāra* confined themselves to providing examples and did not go into further theoretical considerations. Usage of metaphor is termed *istiʿāra*, or sometimes *maṯal*. On the other hand, *istiʿāra* is used to designate semantic figures in a broader sense, especially in Qurʾānic studies. A more specific meaning of the term emerges only in later studies. According to Heinrichs, who summarizes opinions on metaphor before ʿAbd al-Qāhir al-Jurjānī, *istiʿāra* is used in old textbooks for a concept based on analogy, rather than a simple comparison. Comparison and metaphor seem to have been regarded as two separate concepts. *Istiʿāra* is described as *jaʿlu š-šayʾi li-š-šayʾi* 'making something belong to something else', like borrowing an object from its owner and giving it to someone else, as in the famous line 'the reins of the morning were lying in the hand of the northwind'. Here, 'reins' and 'hand' are borrowed objects. This form of metaphor frequently results in a genitive construction. In scriptural exegesis, however, the term *istiʿāra* meant that a name is borrowed and given to a new owner. This comes closer to a concept of *istiʿāra* designed for a word metaphor like *raʾaytu ʾasadan*. Here, something is declared something else, hence the formula is *jaʿlu š-šayʾi š-šayʾa* 'making something become something else', and the problem of identity and similarity comes into the focus.

ʾAbū ʿAmr ibn al-ʿAlāʾ (d. 154/750) is said to have been the first scholar who used the term *istiʿāra*. Al-Jāḥiẓ (d. 255/868) describes *istiʿāra*, among other figures of speech, in a more generic way as "naming something with the name of something else when it takes its place" (*Bayān* I, 86). Ibn Qutayba (d. 276/889) means by *istiʿāra* figurative use of language including metonymy and other rhetorical devices. He introduced the approach of Qurʾānic studies into the discussion and is regarded as the first representative of a Qurʾānic tradition within Arabic rhetoric. Metonymy is also included in Ibn Durayd's (d. 321/933) concept. This is paralleled by the formative phase of Western rhetoric, in which no clear distinction was made between metonymy and metaphor. *Istiʿāra* apparently signified what was later classified *majāz*, while *majāz* was used in an even broader sense than in later periods.

Ṯaʿlab (d. 291/904) seems to have felt the need for a unifying theory, covering different forms of metaphor in his *Qawāʿid aš-šiʿr*, which marks the first study on the field of Arabic poetics. Ibn al-Muʿtazz, (d. 296/908) defines *istiʿāra* in a work exclusively dedicated to rhetorical figures as "borrowing of a word for something that is not known under this word from something that is" (*Badīʿ* 2). It is listed among eighteen means of figurative expression associated with the so-called new poets of his time, although, as the author points out, they may be found in the *Qurʾān*, old poetry, and other contexts as well. Understanding those early theorists' definitions is difficult, and translations are doubtful because of the lack of distinction between words as formal units (→ *lafḏ*) and their meaning (→ *maʿnā*). The relation between simile and metaphor is discussed by ʿAlī ibn ʿĪsā ar-Rummānī (d. 384/994) within the study of → *ʾiʿjāz al-Qurʾān*, the unsurpassable rhetorical qualities of the Holy Book. Similarity is seen as the basic principle and the only necessary condition for the transference (*naql*), which constitutes a metaphor. For aṯ-Ṯaʿālibī (d. 429/1038), comparisons without the particle of comparison can be called *istiʿāra*. The views of ʿAlī ibn ʿAbd al-ʿAzīz al-Jurjānī (d. 392/1001) on simile and metaphor influenced the distinctions made by ʿAbd al-Qāhir al-Jurjānī and later theorists. In the *Kitāb aṣ-ṣināʿatayn* of al-ʿAskarī (d. after 395/1004), we find *istiʿāra* described as a motivated transference of an expression from its normal use in language to another one. The intention might be to make something more distinct, to stress it, or to intensify its presentation. The *istiʿāra* can touch the soul of the listener in a way the straightforward expression cannot (ʿAskarī, *Sināʿatayn* 295).

ʿAbd al-Qāhir al-Jurjānī, the key figure in the field of Arabic rhetoric, deals with metaphor in both his major works. In particular his *ʾAsrār al-balāġa* contains a theory of *istiʿāra* that provided the material for the next generations. ʿAbd al-Qāhir shares with Aristotle an interest in cognitive processes and discusses the psychological aspects of metaphorical speech. Expressions that evoke images appeal to the senses and thereby move the soul. Strangeness and alienation (*ġarāba*) elevate their effects. Al-Jurjānī first follows the lines of his predecessors by defining *istiʿāra* as a word that is temporarily lent to something it does not designate conventionally in the system of language, and by discussing the concept of transference (*ʾAsrār* 29). Then he points out that the information implied in an *istiʿāra* is a comparison (*ʾAsrār* 31). Since increased intensity and conciseness are further purposes, metaphor meets three goals with one rhetorical operation. Al-Jurjānī sees comparison as a basic mental operation that results in a corresponding semantic pattern, *maʿnan min al-maʿānī* (*ʾAsrār* 222). However, neither the idea of comparison or analogy nor the concept of transference can explain the effect of metaphor. Its appeal is rather based on the claim (*iddiʿāʾ*) that two disparate entities are identical. Al-Jurjānī distinguishes between a metaphor that provides some new information, *istiʿāra mufīda*, and an exchange of words that, without being strictly synonyms, have the same basic meaning and thus do not carry original new information, *istiʿāra ġayr mufīda*. For example, *mišfar* 'lip of a camel' differs from *šafa* 'lip of a human being' by selectional restrictions only. The *istiʿāra mufīda* is of two kinds: the borrowed word stands for something particular and known as real, e.g. 'lion' for a brave man; or there is no concrete equivalent in reality, as in the example where a hand is attributed to the northwind, or when 'light' stands for 'right guidance'. Cases like the latter are also named *tamṯīl*.

It should be noted that the subject of transference is always supposed to be a noun, while verb metaphors are to be explained via the verbal noun. Further distinctions focus on how easily the point of resemblance of an *istiʿāra* may be understood. One type of *istiʿāra* remains within the same genus by naming one of its species by another, e.g. ‘flying’ for ‘running’. The difference between this type and an *istiʿāra ġayr mufīda* lies in the intention to stress the specificity of the borrowed notion. The criterion of a second type is that a certain quality is shared between the *mustaʿār* and the *mustaʿār minhu,* like ‘bravery’ in the lion example. The third type is a fictitious intellectually construed image (*ṣūra ʿaqliyya*). The aspect of comparison does not lie in genus or in shared characteristic features, form, or spatial arrangement. This is what ʿAbd al-Qāhir declares to be the purest form of *istiʿāra*. Here, (1) something perceivable by the senses might be applied to something abstract, like ‘balance’ for ‘justice’, or ‘light’ for ‘evidence’; (2) something perceived by the senses stands for something else that is also perceived by the senses, but the similarity can only be inferred intellectually, as in ‘he is a honey, if you meet him halfway’; or (3) there is a transference within the level of intellectual concepts. The most common form of this latter type plays with the category of quantity by ascribing a degree of existence or nonexistence of a certain quality to something, as in Q. 6/122, where a human being living in ignorance is said to be dead. Al-Jurjānī equally uses the term *tamṯīl* for a metaphor that is based on abstract similarity, for the corresponding explicit simile, and for an analogy-based sentence metaphor. Here, a multiplicity of relations must be taken into account, e.g. when a noble and generous man is described as a full moon high above and yet close. A special type of *tamṯīl* is *maṯal*, a proverbial saying whose words retain their original meaning.

Faxr ad-Dīn ar-Rāzī (d. 606/1209) criticizes ar-Rummānī’s definition of *istiʿāra* as the use of an expression for something it does not denote in normal language. According to ar-Rāzī, this definition is deficient for four reasons: (1) every trope would be an *istiʿāra*; (2) proper names that also have a conventional meaning would be tropes; (3) every erroneous use of a word would be an *istiʿāra*; and (4) the so-called *istiʿāra taxyīliyya* would not be included. Then Ar-Rāzī defines *istiʿāra* as “mentioning something under the name of something else and attributing to it what belongs to something else, in order to intensify a comparison” (*ḏikru š-šayʾi bi-smi ġayrihī wa-ʾiṯbātu mā li-ġayrihī lahū li-ʾajli l-mubālaġati fī t-tašbīhi*). The first part of the definition should exclude patterns where the thing compared is explicitly named, as in *zaydun ʾasadun*. The reference to *ʾiṯbāt* is needed to cover the *istiʿāra taxyīliyya*. Then he refers to older concepts, stating that ‘making something become something else’ would also be a valid definition, e.g. making a brave man a lion as in *laqītu ʾasadan*, as would ‘making something belong to something else’, e.g. ascribing a hand to the northwind in order to compare it to someone who controls things. In both cases an enhanced comparison is intended (Rāzī, *Nihāya* 81–82). Faxr ad-Dīn ar-Rāzī proved to be influential in the field of systematization and technical vocabulary, shaping pairs of antonymic terms out of categories he found mostly in the works of his predecessors, e.g. *istiʿāra ʾaṣliyya/istiʿāra ṭabaʿiyya* or *taršīḥ al-istiʿāra/tajrīd al-istiʿāra.*

Finally, metaphor is described in terms of simile. Once *istiʿāra* is based on comparison, *mušabbah bihī* ‘the thing compared with’ and *mustaʿār minhu*, on the one hand, and *mušabbah* ‘the compared thing’ and *mustaʿār lahū*, on the other, are paralleled. The term *mustaʿār* has no counterpart. Again, this approach threatens to mix the formal and the conceptual levels, a danger not unknown in Western linguistics.

With as-Sakkākī (d. 626/1229), a new stage of systematization is reached. While al-Jurjānī sometimes has different names for one type of *istiʿāra* or uses one and the same technical term for different types, as-Sakkākī establishes strict and rigorous categorizations. The main criterion is now the mention of either the *mušabbah* or the *mušabbah bihi* of the implicit comparison, which is combined with the idea of *iddiʿāʾ*. If both are mentioned, it is a simile without a particle. As-Sakkākī defines *istiʿāra* by saying *hiya ʾan taḏkura ʾaḥada ṭarafay at-tašbīhi wa-turīdu bihī ṭ-ṭarafa l-ʾāxara muddaʿiyan duxūla l-mušabbahi fī jinsi l-mušabbahi bihi dāllan ʿalā ḏālika bi-ʾiṯbātika li-l-mušabbahi mā yaxuṣṣu l-mušabbaha bihī*, i.e., one mentions one side of a comparison while the other one is intended, claiming that the thing compared falls into the same genus as the thing likened to, and indicating this by ascribing a special property of the thing likened to the thing compared (Sakkākī, *Miftāḥ* 369). This may be the application of its noun, in the simplest case, or of something else, e.g. ascribing claws to death in order to equate death with a beast of prey. As for the opposite concepts of linguistic trope corresponding to word metaphor (*majāz luġawī*) and intellectual trope (*majāz ʿaqlī*), developed from al-Jurjānī's ideas, as-Sakkākī tries to establish a unifying theory. Anchoring *majāz* in language itself, he postulates a broader meaning of a word comprising all physical and abstract properties of the *denotatum*, whereas a narrower meaning would cover only abstract features. In a given speech act the metaphorical claim for identity might be restricted to some incorporeal attributes which are indicated by context.

Another step toward definitive formalization of the *istiʿāra* concept is seen in Ibn az-Zamlakānī's (d. 651/1253) *Tibyān*. The classificatory division into two categories, the word metaphor and the analogy-based metaphor, appears to be firmly grounded. Views on *istiʿāra* as shortened similes are known from Ḍiyāʾ ad-Dīn Ibn al-ʾAṯīr (d. 637/1239) and Yaḥyā ibn Ḥamza al-ʿAlawī (d. 749/1348).

In the works of Jalāl ad-Dīn al-Qazwīnī, who comments critically on as-Sakkākī and al-Jurjānī, we find the material systematized in a way that represents the dominant tradition passed on in lectures up to modern times. In essence, *istiʿāra* expresses a relationship of comparison between the conventional meaning of a word and the metaphorical one that is actualized in a particular case. Al-Qazwīnī distinguishes a word metaphor with a verifiable *mušabbah* in the background, whether it is sensory ('I saw a lion') or intellectual ('light' for 'proof'), from a compound trope, *majāz murakkab*, a complete analogy or allegory that is realized in a sentence as a whole ('one step forward and one back'). An intermediate position has the combined *istiʿāra bi-l-kināya wa-l-istiʿāra t-taxyīliyya* (the northwind example). Subdivisions are made considering the nature of both sides of the implied comparison, of the *tertium comparationis*, of all these three components, of grammatical and formal aspects of the word, *lafḏ*, and so on. For example, whenever the *jāmiʿ* is easily understood, al-Qazwīnī speaks of an *istiʿāra ʿāmmiyya*; otherwise, it is called a special one, *istiʿāra xāṣṣiyya*. In case the *jāmiʿ* belongs as inherent quality to both sides, he speaks of an *istiʿāra dāxiliyya*, otherwise of an *istiʿāra ġayr dāxiliyya*. When both sides are compatible ('make someone alive' for 'guiding him the right way'), they constitute an *istiʿāra wifāqiyya*, while a logically contradictory combination ('death' for 'living as an ignorant') is called an *istiʿāra ʿinādiyya*.

Just as in Western rhetoric, formal, syntactic, semantic, and modal subdivisions may be found in handbooks. It should be noted that the parameter 'animate/inanimate' is not used as a distinctive feature in Arabic rhetoric. Instead, the dichotomy 'perceptible by the senses vs. intellectual' plays an important role as classificatory criterion. As a result of the history of the concept of *istiʿāra* outlined above, the following technical terms may be found in later sources.

An *istiʿāra taṣrīḥiyya* or *muṣarraḥ bihā* is one which explicitly names the thing compared with, i.e. the *mušabbah bihī* of the implied comparison. This type is called *taḥqīqiyya* if the suppressed *mušabbah* is something perceivable by the senses or something one is certain about.

In case of an *istiʿāra makniyya* or *istiʿāra bi-l-kināya*, the *mušabbah* appears while the *mušabbah bihī* is concealed, as is the comparison itself, and is only hinted at by something associated with it (→ *kināya*). In the canonical system, this type is always combined with an *istiʿāra taxyīliyya*, which means the aspect of attributing to the *mušabbah* a characteristic it does not have in reality and thus creating an illusion.

Considering grammatical categories, an *istiʿāra* realized through a nonderived noun, *ism jāmid*, is named *istiʿāra ʾaṣliyya*. However, realized through a verb or a verbal derivative or a particle – e.g. *wa-lammā sakata ʿan mūsā l-ġaḍabu* 'when anger grew silent before Moses' (Q. 7/154), *sukūt* 'growing silent' being the *mušabbah bihī* according to one of the possible interpretations, and the ending of anger the *mušabbah* – it is named an *istiʿāra ṭabaʿiyya*.

As for the 'divested' metaphor, *istiʿāra mujarrada*, the *mustaʿār lahū*, i.e. the *mušabbah*, can be inferred from a qualification added to the *mustaʿār*, e.g. 'I saw a lion groaning under his armor'.

In case of the 'prepared' or 'developed' metaphor, *istiʿāra muraššaḥa*, something that fits the *mušabbah bihī*, i.e. the *mustaʿār minhu* from which the metaphorical word is borrowed, is added to speech.

The term *istiʿāra muṭlaqa* is applied to cases that are neither *mujarrada* nor *muraššaḥa* or, especially in late compendiums, to cases that are both at the same time. That means that either there is no specification at all or both sides of the implied comparison are characterized, e.g. ‘a lion, bristling with weapons, whose talons have not been cut’.

The ironic or ‘salty’ metaphor, *istiʿāra taḥakkumiyya* or *istiʿāra tamlīḥiyya*, is a play with opposites in which one substitutes for the other, e.g. calling a coward a lion.

In an *istiʿāra tamṯīliyya* or *tamṯīl ʿalā sabīl al-istiʿāra*, the point of resemblance is not a simple notion but a relation in itself. As mentioned above, the meaning of a whole sentence is likened to another constellation of things, e.g. ‘he whose mouth is bitter from malady will find bitterness even in pure water’.

Given the overall relevance of poetry and language as a source of cultural pride and the construction of Arabic identity, metaphor as a creative mode of expression has always played an important role in the linguistic habit of the Arabs. While in pre-Islamic times a preponderance of similes and genitive metaphors are stated, in the Abbasid period poetry, influenced by urban milieus and the influx of Persian cultural elements, started to change. A higher degree of rhetorization came into fashion, with metaphors becoming keener and sometimes far fetched. In the discussion on literary standards, a line was drawn between modern (*muḥdaṯūn*), and more traditionally orientated poets (*mutaqaddimūn*). Muslim ibn al-Walīd (d. 208/823) is said to have been the first to cultivate the new style, *badīʿ*, of which ʾAbū Tammām (d. 230/846) is another master. In the 10th century an often manieristic style was established, characterized by the anthropomorphization of nature and the construction of highly imaginative worlds where a poetic image provides the background for another (‘the cheek of a rose is glowing with shame’). The role of panegyric, especially in poetic production, was another factor that stimulated hyperbolic metaphorical speech and the construction of fantastic etiologies. Even official writings were expected to meet literary standards. In the milieu of learned secretaries, often closely connected with literary circles, the ideal of *ʾadab* was held up, and artistic prose, including verses, assonances, tropes, and rhetorical figures, was of common usage. A rhetorically highly elaborated genre of Arabic literature was the *maqāma,* represented in classical form in the works of al-Ḥarīrī (d. 516/1122). Due to the unrivaled artificiality and enigmatic character of later *maqāmāt*, even the well-educated reader of that time was not easily able to get along without commentaries on metaphors, allusions, and the like. Here, verbal art tends to become an end in itself. Literary virtuosity in all genres of writing, down to the most prosaic ones, has been an ideal throughout Islamic history, or at the least, a learned person was supposed to have knowledge of grammar and rhetoric.

Recently, the use of metaphorical language in contemporary official, public, and everyday speech in the Arabic world has come into the focus of study and needs further research.

Bibliographical references

Primary sources

ʿAskarī, *Ṣināʿatayn* = ʾAbū Hilāl al-Ḥasan ibn ʿAbdallāh ibn Sahl al-ʿAskarī, *Kitāb aṣ-ṣināʿatayn*. Ed. Mufīd Qumayḥa. Beirut: Dār al-Kutub al-ʿIlmiyya, 1984.

Ibn al-Muʿtazz, *Badīʿ* = ʿAbdallāh Ibn al-Muʿtazz, *Kitāb al-badīʿ*. Ed. Ignatius Kratchkovsky. Damascus: Dār al-Ḥikma, n.d. (1st ed. London, 1935.)

Jāḥiẓ, *Bayān* = ʾAbū ʿUṯmān ʿAmr ibn Baḥr al-Jāḥiẓ, *al-Bayān wa-t-tabyīn*. Beirut: Dār al-Kutub al-ʿIlmiyya, n.d.

Jurjānī, *ʾAsrār* = Majd ad-Dīn ʾAbū Bakr ʿAbd al-Qāhir ibn ʿAbd ar-Raḥmān al-Jurjānī, *ʾAsrār al-balāġa*. Ed. Helmut Ritter. Beirut: Dār al-Masīra, 1983. (1st ed. Istanbul, 1954. German transl. Hellmut Ritter, *Die Geheimnisse der Wortkunst des ʿAbdalqāhir al-Curcānī*. Wiesbaden: F. Steiner, 1959.)

——, *Dalāʾil* = Majd ad-Dīn ʾAbū Bakr ʿAbd al-Qāhir ibn ʿAbd ar-Raḥmān al-Jurjānī, *Dalāʾil al-ʾiʿjāz*. Ed. Muhammad Rašīd Riḍā. Beirut: Dār al-Maʿrifa, 1978.

Qazwīnī, *ʾĪḍāḥ* = Jalāl ad-Dīn ʾAbū ʿAbdallāh Muḥammad ibn ʿAbd ar-Raḥmān al-Qazwīnī, *al-ʾĪḍāḥ fī ʿulūm al-balāġa*. Ed. Muḥammad al-Xafājī. Beirut: Dār al-Kitāb al-Lubnānī, 1985.

Rāzī, *Nihāya* = Faxr ad-Dīn ʾAbū ʿAbdallāh Muḥammad ibn ʿUmar ar-Rāzī, *Nihāyat al-ʾījāz fī dirāyat al-ʾiʿjāz*. Cairo: Maktabat al-ʾĀdāb, 1909.

Sakkākī, *Miftāḥ* = Sirāj ad-Dīn ʾAbū Yaʿqūb Yūsuf ibn ʾAbī Bakr as-Sakkākī, *Miftāḥ al-ʿulūm*. Ed. Naʿīm Zarzūr. Beirut: Dār al-Kutub al-ʿIlmiyya, 1983.

Secondary sources

Bonebakker, Seeger Adrianus. 1973. “*Istiʿāra*”. *Encyclopaedia of Islam* IV, 248–252. 2nd ed. Leiden: E.J. Brill.

Eggs, Ekkehard. 2001. “Metapher”. *Historisches Wörterbuch der Rhetorik*, V, ed. Gert Ueding, 1099–1186. Tübingen: Niemeyer.

Heinrichs, Wolfhart. 1977. *The hand of the northwind: Opinions on metaphor and the early meaning of istiʿāra in Arabic poetics*. Wiesbaden: O. Harrassowitz.

Udo Simon (University of Heidelberg)

Ištiqāq

1. Ištiqāq

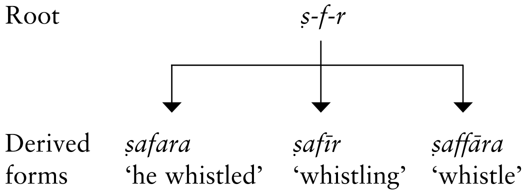
The term *ištiqāq* lit. ‘splitting’ is a grammatical term that translates into English as ‘derivation’; in some respects it could be viewed as the equivalent of the notion of → etymology. *Ištiqāq* means that one word is derived from another, or that the two are derived from a common source, called *ʾaṣl* ‘root’ (cf. ʾAstarābāḏī, *Šarḥ aš-Šāfiya* II, 334). In order to understand the derivation of new words from the root, the root should be defined in relation to another concept called *wazn* ‘template’.

2. Root

Since there is no infinitive in Arabic, the verb in the 3rd person singular masculine perfect, e.g. *kataba*, corresponds to the citation form of the verb ‘to write’, even though grammatically it refers to a conjugated verb ‘he wrote’. *Ṣafara* ‘he whistled’, *ṣafīr* ‘whistling’, *ṣaffāra* ‘whistle’ are words that share the three consonants *ṣ, f, r* and the underlying idea of ‘whistling’. These three consonants in this particular order are known as the → ‘root’. The root is composed of consonants referred to as radicals (*ḥurūf ʾaṣliyya*), which denote a general meaning. Altering the position of any of the radicals would cause a complete change in the meaning (cf. Ibn Jinnī, *Xaṣāʾiṣ* II, 134, about *ištiqāq ʾaṣġar* ‘small derivation’). The root refers to the common denominator shared by a number of words connected synchronically by meaning. Vowels and affixes are introduced to derive actual words.

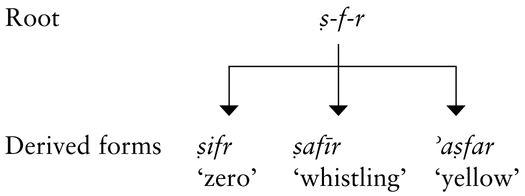
Some Arab grammarians consider the semantic relationship between the root and its derived forms to be crucial. Only when the semantic relationship is satisfied are derived forms regarded as belonging to this root. This approach is called derivational and is found, for instance, in Ibn as-Sarrāj (*ʾUṣūl*) and as-Sakkākī (*Miftāḥ*). Figure 1 illustrates this approach.

Figure 1. Derivational approach



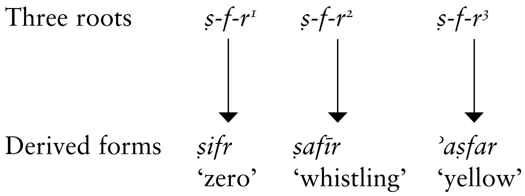
Forms such as *ṣifr* ‘zero’, *ṣafara* ‘he whistled’, *ṣafīr* ‘whistling’, *ṣaffāra* ‘whistle’, and *ʾaṣfar* ‘yellow’ do not necessarily all have a semantic relationship, although these forms have one common root, i.e. *ṣ-f-r*. Some Arab grammarians (e.g. Ibn Jinnī, *Xaṣāʾiṣ*) regard the derived forms as derived from a formal root. This view does not require the derivation to be governed by semantic considerations. Figure 2 illustrates this point of view.

Figure 2. Formal approach



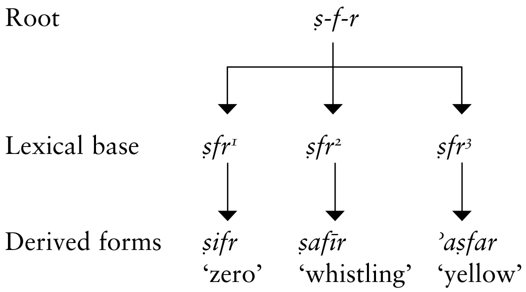
Since the forms *ṣifr*, *ṣafīr*, and *ʾaṣfar* do not have any semantic relationship, each derived form could be analyzed as having its own root, as in Figure 3.

Figure 3. Three different roots



Both approaches, the derivational and the formal, could be combined, as in Figure 4.

Figure 4. Combination of derivational and formal approaches



The root *ṣ-f-r* could be seen as the neutral and abstract source of derivation for all the forms in Figure 4, without giving any specific consideration to the semantic or the formal root (cf. Mokhlis 1997:37-38).

The three consonants *ṣ, f, r* are stable in different templates. This class of roots is called *ṣaḥīḥa* 'sound' because the three consonants appear regularly without showing any change and are not subject to elision. The total number of triradical sound roots is about 3,218, while the number of quadriradical roots is about 1,293 (cf. Al-Bawab a.o. 1996).

Another class of roots is called *muʿtalla* 'weak' (→ *ʿilla*). This class shows semivowels (or → glides), i.e. *w* and/or *y*, among its radicals. This class subdivides into two categories. The first cate-gory shows a glide in R_1, the 'assimilated words', in R_2, the 'hollow words', or in R_3, the 'defective words'. The second category is doubly weak and shows glides either in R_2 and R_3, e.g. R_1*wy* (*lafīf maqrūn*), or in R_1 and R_3, e.g. *w*R_2*y* (*lafīf mafrūq*). These last two classes are subject to *ʾiʿlāl* 'defectiveness': "the phonological change that is carried out in a word in which a weak radical is counted as unsound" (→ *ṣarf*; Åkesson 2001). This weak radical is subject to morphological and phonological processes such as mutation, truncation, and/or vowel truncation (cf. Brame 1970; Bohas 1981, 1982, 1985; Kouloughli 1979; Guillaume 1982; Angoujard 1984; Mokhlis 1997; Chekayri and Scheer, forthcoming). Comparing weak and sound roots, the Arab grammarians define circumstances under which a glide could persist, be changed, or be deleted, and they explain the morphological structure of the language by setting up rules governing its use.

According to the Arab grammarians, both defective and sound verbal phonological representations are subject to → *qiyās* 'measurement'. Translated into English as → 'analogy', *qiyās* is a method used to explain apparent deviations from the rules in certain phenomena by referring to their resemblance to other phenomena (cf. Versteegh 1997:47). This concept is one of the basic instruments for finding similarities between forms. It became an important explanatory principle at the theoretical level and was used for induction as well as for formulating general principles in the language. As an illustration, Arab grammarians consider that the phonological representation of forms such as [jas̴ifu] 'he describes', [qɑːla] 'he said', and [sirtu] 'I went' are /yawṣifu/, /qawala/, and /sayartu/, respectively. Although the glides do not appear at the phonetic level, these forms are derived from triradical roots: *w-ṣ-f, q-w-l*, and *s-y-r*. The Arab grammarians use analogy to restore the underlying level and retrieve the triradical root. In this view, grammar defines in which contexts glides are deleted. According to Arab grammarians, the proof that these surface forms are derived from /yawṣifu/, /qawala/, and /sayartu/ resides in the form of their → *maṣdar* 'verbal noun': *waṣf, qawl*, and *sayr* (cf. Ibn al-ʾAnbārī, *ʾInṣāf*; Ibn Yaʿīš, *Šarḥ al-Mufaṣṣal*; Bohas 1982, regarding the debate about whether the verb or the *maṣdar* is the source of the derivation).

All derived forms are ultimately traceable to a triradical root. The verb *istaqbala* 'he received' breaks down to a triradical root *q-b-l* 'to accept.' In the same way, the imperfects *yaṣifu* 'he describes', *yasmū* 'he ascends', *yamšī* 'he walks' derive from the triradical roots *w-ṣ-f, s-m-w*, and *m-š-y*, respectively. As for the verb *tadaḥraja* 'he rolled himself', it is listed under a quadriradical root, *d-ḥ-r-j* 'to roll.'

3. *WAZN* TEMPLATE

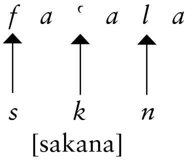
The Arabic and Semitic grammatical tradition use the three consonants *f*, *ʿ*, and *l* to describe forms according to specific templates, called *wazn, ṣīġa*, and *binya*. The template represents an abstract concept for formal description. The template denotes the morphological representation that substitutes for a given lexical form. Every template carries a grammatical sense that is combined with the basic meaning of the root.

To make a distinction between the root radicals and affixes, Arab grammarians list ten segments that are susceptible to being affixes or augments (*zawāʾid*): *ʾalif* 'long *a*', *y, w, m, t, n, s, h, l*, and the glottal stop *ʾ*. The question that arises is how to identify these affixes.

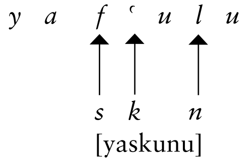
Three main arguments are given by Arab grammarians to distinguish between the root and the augment segments. The identification of augment segments is made by (1) *ištiqāq* 'splitting', (2) *istidlāl bi-l-miṯāl* 'arguing by similar form', and (3) *istidlāl bi-l-kaṯra* 'arguing by frequency'.

The first argument is related to 'splitting'. The root *s-k-n*, which carries the semantic load 'to live', is associated with the template *faʿala*. The first radical, *s*, of the root *s-k-n* must occupy the position of the *f* of the template *faʿala*. The second radical of the root, *k*, fits in the *ʿ* position of the template. Finally, the third radical of the root, *n*, must replace the *l* position of the template. The result of this association is the output form *sakana* 'he lived'. It is necessary to note that the vowels of the template remain unchanged in their features, quantity, and position. If one says *yaskunu* 'he lives', then the template is *yafʿulu*, the radicals corresponding to *f*, *ʿ*, and *l*. The *y* and the vowels are considered to be augments in this case. They are kept identical to the ones in the template (cf. Ibn Yaʿīš, *Šarḥ al-Mulūkī* 118–121):

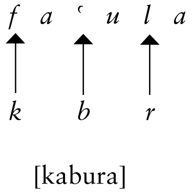
(1)



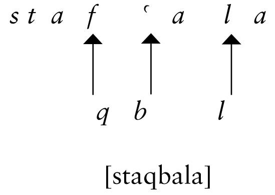
(2)



(3)



(4)



It should be noted that forms may be derived not only from a root but from other derived forms as well. The reflexive form *tafaʿʿal* is derived from *faʿʿal*, and *tafāʿal* is generated from *fāʿal*; the passive is derived from the active, the imperfect is derived from the perfect, the passive participle is derived from the passive imperfect, and the active participle is derived from the active imperfect, etc. (cf. Bohas 1982).

The second argument is referred to as *istidlāl bi-l-miṯāl* 'arguing by similar forms'. It is a method used to identify segments susceptible to being an augment in a given word. These opaque segments can be radicals or augments: if *ištiqāq* cannot help in extracting the root, then *istidlāl bi-l-miṯāl* may be used as an analogical process based on comparing forms to similar ones. Ibn Yaʿīš gives two main examples to illustrate this process (cf. Ibn Yaʿīš, *Šarḥ al-Mulūkī* 120–121; Bohas 1982:180–182; Mokhlis 1995:94).

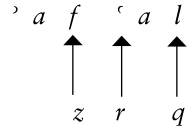
i. The form *ʿantar* [proper noun] has two segments, *n* and *t*, that are susceptible to being augments. Ibn Yaʿīš compares *ʿantar* with a similar form such as the proper noun *jaʿfar*, whose segments are all radicals. He then concludes that *ʿ, n, t*, and *r* are also radicals.

ii. The form *ʿanbas* [attribute of lions] contains a segment *n*, which might be considered to be an augment. Although this form has the same template as *jaʿfar*, i.e. *faʿlal*, the *n* in *ʿanbas* does not appear in the verbal form *ʿabas* 'to scowl' and in the *maṣdar ʿubūs* 'scowling'. Thus, the *n* in *ʿanbas* is an augment. Since *ʿanbas, ʿabas*, and *ʿubūs* have a related meaning, they are considered to be derived forms from the same root *ʿ-b-s*.

The third argument is referred to as *istidlāl bi-l-kaṯra* 'arguing by frequency'. An example is the segment *ʾa*, which is analyzed as a prefix because of its frequent occurrence in forms of the template *ʾafʿal*, which denotes nouns of color or a comparative or superlative form, e.g. *ʾazraq* 'blue', *ʾakbar* 'biggest', etc. Assigning these words to the template *ʾafʿal* makes it possible to establish the difference between the elements that are analyzed as augments and those that are radicals. It then becomes obvious that *ʾa* in *ʾazraq* is an augment.

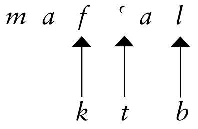
The repetition of *ma* in *masbaḥ* 'swimming pool', *malʿab* 'stadium', *maktab* 'office', etc. identifies it as a prefix. These forms are derived from the template *mafʿal*, which refers to a noun of place. We thus have:

(5)



[ʔazraq]

(6)



[maktab]

To restore the underlying level (phonological form) and retrieve the correct root, it is crucial to know the principles governing Arabic word structure (morphology). In this context, three possible ways of arguing are given: (1) splitting words (*ištiqāq*); (2) comparing forms to similar ones; and (3) determining frequency. Arab grammarians consider *ištiqāq* the most accurate tool to find out about the source of derivation. Other methods may have different results. Raḍī d-Dīn al-ʾAstarābāḏī gives in his *Šarḥ aš-Šāfiya* (II, 335) the example of *ʾalandad*, which has the template *ʾafanʿal*. *ʾAlandad* and *yalandad* have the same meaning as *ʾaladd* 'fierce'. The three forms *ʾalandad, yalandad*, and *ʾaladd* are derived from the root *l-d-d*, which carries the semantic load 'to dispute violently'. According to *ištiqāq*, the glottal stop at the beginning of the word followed by three consonants is analyzed as an augment. The nonvocalized *n* in the third position and the reduplicated radical are perceived as augments as well.

Raḍī d-Dīn al-ʾAstarabadi considers two possibilities: *ʾ* and *n* are augments because *ʾalandad* is derived from the root *l-d-d*; and *ʾ* and one *d* are augments because *ʾalandad* is derived from the root **l-n-d*. He chooses the first possibility, in which *ʾ* and *n* are augments. In his view, *ištiqāq* is still the best way to identify augment segments accurately.

The choice of three segments, *f*, *ʿ*, and *l*, as the template of the lexical forms in Arabic is based on the occurrence and the frequency of triradical forms compared to other structures such as quadriradical or quinqueradical forms (Ibn Yaʿīš, *Šarḥ al-Mulūkī* 116). Table 1 gives an estimate of the number of roots, based on an inventory from Classical Arabic dictionaries (cf. Ḥilmī 1973; Mrayati 1987).

This explains why the Arab grammarians believe that the minimal template in Arabic is triradical and why they deny the existence of biradical roots in the lexical system of the Arabic language. In modern linguistic analyses this is sometimes regarded as one of the weaker points of the Arab grammarians' reflections on the lexicon (→ biradicalism; cf. Chekayri forthcoming; Mokhlis 1997:57).

4. CONCLUSION

Being able to recognize word templates and roots is very important for three principal reasons:

i. In the usually unvocalized text, the recognition of a known template will enable the learner to pronounce a particular word and learn it without having to puzzle out which vowels go where.
ii. In many cases the template of a word will suggest its function within the sentence and help the learner to determine its meaning.
iii. In case of doubt, 'splitting' confirms the existence of a word (cf. Ibn Jinnī, *Xaṣāʾiṣ* I, 369).

Not all derived verbs exist for all roots. Most roots (1,212) have at least one, and a few (5) have nine derived forms. This obviously causes problems for the learner, and the only way to overcome them is to study the templates of Arabic words and thus gain experience in distinguishing roots from augments. Students

Table 1. Numbers of roots in dictionaries

Author(s)	Dictionaries inventoried	Roots			Total
		three radicals	four radicals	five radicals	
Ḥilmī (1973)	*Tāj al-ʿarūs*	7,597	4,081	300	11,978
Mrayati (1987)	*al-Jamhara* *at-Tahḏīb* *al-Muḥkam* *Lisān al-ʿArab* *al-Qāmūs*	7,198	3,739	295	11,232

of the Arab grammarians, confronted with the same problems, were presented with complicated exercises to learn this procedure of ‘splitting’ the word.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

ʾAstarābāḏī, *Šarḥ aš-Šāfiya* = Raḍī d-Dīn Muḥammad ibn Ḥasan al-ʾAstarābāḏī, *Šarḥ Šāfiyat ibn al-Ḥājib*. Ed. Muḥammad Nūr al-Ḥasan, Muḥammad az-Zafzāf, and Muḥammad Muḥyī d-Dīn ʿAbd al-Ḥamīd. 4 vols. Beirut: Dār al-Kutub al-ʿIlmiyya, 1975.

Ibn al-ʾAnbārī, *ʾInṣāf* = Kamāl al-Dīn ʾAbū l-Barakāt ʿAbd ar-Raḥmān ibn Muḥammad ibn ʾAbī Saʿīd al-ʾAnbārī, *Kitāb al-ʾInṣāf fī masāʾil al-xilāf bayna n-naḥwiyyīna l-baṣriyyīna wa-l-kūfiyyīn*. Cairo: Maṭbaʿat as-Saʿāda, 1961.

Ibn Jinnī, *Taṣrīf* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *at-Taṣrīf al-Mulūkī*. Damascus: Dār al-Maʿārīf, 1970.

——, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. Ed. Muḥammad ʿAlī an-Najjār. 3 vols. Cairo, 1952–1956. (Repr., Beirut: Dār al-Hudā, n.d.)

Ibn as-Sarrāj, *Ištiqāq* = ʾAbū Bakr ibn as-Sarī Ibn as-Sarrāj, *Risālat al-ištiqāq*. Damascus: Dār Majallat aṯ-Ṯaqāfa, 1972.

——, *ʾUṣūl* = ʾAbū Bakr ibn as-Sarī Ibn as-Sarrāj, *Kitāb al-ʾuṣūl fī n-naḥw*. Ed. ʿAbd al-Ḥusayn al-Fatlī. 3 vols. Beirut: Muʾassasat ar-Risāla, 1985.

Ibn ʿUṣfūr, *Mumtiʿ* = ʾAbū l-Ḥasan ʿAlī ibn Muʿmin Ibn ʿUṣfūr al-ʾIšbīlī, *al-Mumtiʿ fī t-taṣrīf*. Ed. Faxr ad-Dīn Qabāwa. 4th ed. 2 vols. Beirut: Dār al-ʾĀfāq al-Jadīda, 1979.

Ibn Yaʿīš, *Šarḥ al-Mufaṣṣal* = Muwaffaq ad-Dīn Yaʿīš Ibn Yaʿīš, *Šarḥ al-Mufaṣṣal*. 10 vols. Beirut: ʿĀlam al-Kutub, n.d.

——, *Šarḥ al-Mulūkī* = Muwaffaq ad-Dīn Yaʿīš Ibn Yaʿīš, *Šarḥ al-Mulūkī fī t-taṣrīf*. Aleppo: al-Maktaba al-ʿArabiyya, 1973.

Sakkākī, *Miftāḥ* = ʾAbū Yaʿqūb Yūsuf ibn ʾAbī Bakr as-Sakkākī, *Miftāḥ al-ʿulūm*. Ed. Nuʿaym Zarzūr. Beirut: Dār al-Kutub al-ʿIlmiyya, 1983.

SECONDARY SOURCES

Åkesson, Joyce. 2001. *Arabic morphology and phonology based on the* Marāḥ al-ʾarwāḥ *by ʾAḥmad b. ʿAlī b. Masʿūd*. Leiden: E.J. Brill.

Al-Bawab, Marwan, Yahya Mir Alam, Mohamad Mrayati, and Mohamad Hassan Al-Tayyan. 1996. *Arabic verbs in a computerized dictionary*. Beirut: Librairie du Liban.

Angoujard, Jean-Pierre. 1984. *Aspects d'une micro-prosodie: Le modèle arabe*. Ph.D. diss., Université Paris VIII.

Bohas, Georges. 1981. “Quelques aspects de l'argumentation et de l'explication chez les grammairiens arabes”. *Arabica* 28.204–221.

——. 1982. *Contribution à l'étude de la méthode des grammairiens arabes en morphologie et en phonologie, d'après des grammairiens arabes ‘tardifs’*. Lille: Presse Universitaire.

——. 1985. “L'explication en phonologie arabe”. *Zeitschrift für Arabische Linguistik* 15.45–52.

Brame, Michael. 1970. *Arabic phonology: Implications for phonological theory and historical Semitic*. Ph.D. diss., Massachusetts Institute of Technology.

Chekayri, Abdellah. 1994. *La structure des racines en arabe*. Ph.D. diss., Université Paris VIII.

——. 1999. *L'articulation du lexique et de la morphologie arabes: Des verbes défectueux et particulièrement les verbes assimilés*. Doctorat d'Etat, Mohamed V University, Rabat.

——. Forthcoming. “Weak verbs in Arabic”.

—— and Tobias Scheer. Forthcoming. “The conjugation of hollow verbs in Classical Arabic”.

Fleisch, Henri. 1973. “*Is̲h̲tiḳāḳ*”. *Encyclopaedia of Islam* IV, 122. 2nd ed. Leiden: E.J. Brill.

——. 1979. *Traité de philologie arabe*. II. *Pronoms, morphologie verbale, particules*. Beirut: Dār El Machreq.

Foley, James. 1985. “Quatre principes de l'analyse morphologique”. *Languages* 78.57–72.

Guillaume, Jean-Patrick. 1982. *Quelques aspects de la théorie morpho-phonologique d'Ibn Jinnī, à propos des verbes à glide médian*. Lille: Presse Universitaire.

Ḥilmī, Mūsā. 1973. *Dirāsa ʾiḥṣāʾiyya li-juḏūr muʿjam* Tāj al-ʿarūs *(bi-stixdām al-kumbyūtar)*. Kuwait: Maṭbūʿāt Jāmiʿat al-Kuwayt.

Kouloughli, Djamel Eddine. 1979. “Sur le traitement des glides dans la phonologie de l'arabe standard”. *Analyse Theorie* 1.54–93.

Mokhlis, Hassan. 1997. *Théorie du* taṣrīf *et traitement du lexique chez les grammairiens* arabes. (= *Publications Universitaires Européennes, Série 21, Linguistique*, Vol. 185.) Frankfurt a. Main.: P. Lang.

Mrayati, Mohamad. 1987. “Statistical studies of Arabic language roots”. *Applied Arabic linguistics and signal and information processing*, ed. Raymond Descout, 97–103. Berlin: Springer.

Smart, Jack R. 1986. *Arabic*. (= *Teach Yourself Books*.) London: Hodder and Stoughton.

Versteegh, Kees. 1997. *The Arabic linguistic tradition*. London and New York: Routledge Kegan Paul.

ABDELLAH CHEKAYRI
(Al Akhawayn University)

Ištirāk → Muštarak

Istiṯnāʾ

Istiṯnāʾ, a verbal noun of Form X from the radicals *ṯ-n-y*, literally means ‘setting aside as excluded; exclusion, exception’ (Lane 1863–1893:I, 357; Wehr 130). The term is used in Arabic grammar to denote an exception or an exceptive sentence (*jumlat al-istiṯnāʾ*), which basically consists of two parts: the general term from which the exception is made (*al-mustaṯnā minhu*) and the exception itself (*al-mustaṯnā*). The relationship between these two parts of the sentence is made through the use of an exceptive particle (*ḥarf al-istiṯnāʾ*), mostly *ʾillā* (a compound of *ʾin* and *lā* and traditionally considered the *ʾaṣl* 'source, basic principle' of the exceptive particles). Other expressions are used in the sense of 'excluding, excepting' as well, for example nouns like *siwā* 'another [besides so-and-so]' and *ġayr* 'other than, different from', or the verbal clauses *mā xalā* 'what is free from', *mā ʿadā* 'what goes beyond', and *ḥāšā* 'he excepted' (Cachia 1973:19–20 [A-E], 27 [E-A]; Wright 1967:II, 335–343, and Howell 1990:I, 296–319, both of whom also mention the verbal expressions *laysa* and *lā yakūnu* to denote exceptions in Classical Arabic; Fischer 2002:168–170; Cantarino 1975:III, 192–193, 338–352; Dahdah 1988:180–181; Reckendorf 1967:712–726, "Exzeptivsätze"; Reckendorf 1921:502–512, "Ausnahmesätze"; Blachère and Gaudefroy-Demombynes 1975:405–406, 445–446, "phrase exceptive"; Blachère 1985:156; Badawi a.o. 2004:671–684).

The general principles of exceptive sentences are usually explained based on the way the most widely used exceptive particle, *ʾillā*, occurs in Arabic. *ʾIllā*'s natural position in the sentence is after the general term from which the exception is made; the syntactic function of the exception following *ʾillā* depends on the relationship between the general term and the excepted element. There are three types of exceptive sentences, called 'void exception' (*istiṯnāʾ mufarraġ*), 'joined exception' (*istiṯnāʾ muttaṣil*), and 'severed exception' (*istiṯnāʾ munqaṭiʿ*); they are translated, respectively, as 'exhaustive', 'continuous', and 'discontinuous' exceptions in Badawi a.o. (2004:671).

The void-exceptive sentence – the most frequently used in Modern Standard Arabic – is always negative and has no general term; the exception obtains whatever case the general term would have had were it expressed. Hence, in *mā qāma ʾillā zaydun* 'no one stood up except Zayd', *zayd* assumes the nominative case ending because, if expressed, the general term would have taken the nominative, e.g. *mā qāma l-qawmu* 'the people did not stand up'. The following examples provide further illustration: *mā ʿalayka ʾillā tajdīdu malābisika* 'you only have to change your clothes', *lam yakun ʾillā majnūnan* 'he was nothing but insane', *ʾamrāḍun lā tuʿālaju ʾillā bi-d-dawāʾi l-mustawradi* 'illnesses that can only be treated with imported medicine' (Cantarino 1975:III, 342; Badawi a.o. 2004:673, 674).

The joined exception is characterized by the fact that the general term is explicitly mentioned and the exception belongs to the same genus or generic category as the general term, as is the case with *qawm* 'people' and *zayd* 'Zayd'. The syntactic function of the excepted element depends on whether the antecedent containing the general term is positive or negative. If positive, the exception expresses an exclusion from the general term and obtains the accusative case: *qāma l-qawmu ʾillā zaydan* 'the people stood up except Zayd', i.e., everyone stood up, but Zayd was excluded from them and did not stand up. When the antecedent is negative, however, the exception might be considered a substitute for the general term and takes over its syntactic function. In *mā qāma l-qawmu ʾillā zaydun* 'the people did not stand up, except Zayd', Zayd is the only one standing up from among the people, taking their place as it were. Examples of positive joined exceptions are *wa-qad turjimat kutubuhu fī ḥayātihi ʾillā kitāban wāḥidan* 'all his books were translated during his life except one' and *taraka l-ʿamala ʾillā ʾašġālan xafīfatan* 'he gave up work except small jobs'. Examples of negative joined exceptions are *lam yatruk lahum šayʾan ʾillā naṣībahu fī baytin qadīmin* 'he did not leave anything for them except for his share in an old house' and *laysa lī naṣībun min ad-dunyā ʾillā hāḏihi s-sāḥatu ḍ-ḍayyiqatu l-jāmida* 'I have no other property on earth than this narrow, dry patch of land' (Cantarino 1975:III, 340; Badawi a.o. 2004:675).

In a severed exception, the antecedent containing the general term is negative (as in the last mentioned examples), but the exception belongs to a fundamentally different category than this general term. As a consequence, the exception expresses exclusion and cannot take the syntactic function of the general term, so it is always put in the accusative case. In *mā jāʾanī ʾaḥadun ʾillā ḥimāran* 'no one came to me except a donkey', the donkey cannot substitute for a human being, and *ḥimār* obtains the accusative case ending to emphasize its being different. Although the concept of the severed exception is prominent in the discussions about exceptive sentences in the Arabic grammatical tradition (which may be due to the fact that initially the distinction between severed and void exceptions was not made; cf. Sībawayhi, *Kitāb* I, 319ff.), the grammatical construction itself is rare in Standard Arabic (no examples were found in Badawi a.o. 2004; cf. 672, 675). For other usages of *ʾillā* (e.g. *ʾillā wa-*, *ʾillā* with clauses introduced by *ʾan* or *ʾanna*, *ʾillā* with adverbial clauses), the rules for *ġayr* and *siwā* (which resemble those for *ʾillā*), and the use of verbal clauses with the sense of exception, see Cantarino (1975:III, 338–352) and Badawi a.o. (2004:675–684).

The principal grammatical rules for exceptive sentences, as formulated in Sībawayhi's *Kitāb* based on the use of the exceptive particle *ʾillā* (*Kitāb* I, 314–329; cf. Carter 1975; Bernards 1997:16–18), were not subject to fundamental disagreement between grammarians of Arabic, but the interpretation and application of the rules left ample room for discussion. Especially the governance of the exceptive particle *ʾillā* was elaborated. In Sībawayhi's view, for example, the noun following *ʾillā* in a severed exception is governed by the antecedent in the same way that *dirhaman* is governed by *ʿišrūna* in the expression *ʿišrūna dirhaman*, where the *nūn/tanwīn* of *ʿišrūna* separates the two elements of the sentence and prevents the genitive case in *dirham* (Sībawayhi, *Kitāb* I, 315; cf. Carter 1972 for further analysis of what he has coined the *tanwīn-naṣb* construction; Owens 1990:107ff., who calls it the 'separation and non-identity principle', SNIP; see also Talmon 1993, 2003:245ff.). *Mā jāʾanī ʾaḥadun* 'no one came to me' is a grammatically correct and complete sentence – the *nūn/tanwīn* indicates this completeness. That a donkey did come (*ʾillā ḥimāran*) stands by itself. The donkey cannot take the place of the general term, a human being; the donkey cannot even be considered to be a description of the general term as an adjective (*waṣf/ṣifa*), so in this case there can be no grammatical agreement between *ʾaḥad* and *ḥimār*. Hence, *ʾillā* assumes the meaning of *wa-lākinna*: 'no human being came to me, but a donkey did', and the accusative is preferred for the excepted noun (Sībawayhi, *Kitāb* I, 319).

Among the first generations of grammarians after Sībawayhi, not only the nature of a severed exception was contested but the regency of *ʾillā* as well. In a marginal note to the *Kitāb Sībawayhi*, al-Māzinī (d. 248/862) argues that human beings and donkeys belong to the same category of living creatures; and in the opinion of al-Mubarrad (d. 285/898; *Muqtaḍab* IV, 390), *ʾillā* takes the place of a verb and gives what follows its case ending. Moreover, the regency of *ʾillā*, more specifically in positive exceptive sentences like *qāma l-qawmu ʾillā zaydan* 'the people stood up, except Zayd', is the topic of one of the *masāʾil ixtilāfiyya*, the *Streitfragen* between Baṣran and Kūfan grammarians as it appears from Ibn al-ʾAnbārī's *ʾInṣāf* (118–122, *masʾala* no. 34). Subsequent generations of grammarians pondered the issue, refining and clarifying Sībawayhi's theories without, however, changing the basic underlying concepts (see, e.g., Zajjājī, *Jumal* 235–236; Ibn Jinnī, *Lumaʿ* 28; ʾAbū Ḥayyān, *Manhaj* 162ff.; Ibn Yaʿīš, *Šarḥ* II, 79–81; cf. Bernards 1997:16–22).

Bibliographical references

Primary sources

ʾAbū Ḥayyān, *Manhaj* = ʾAbū Ḥayyān Muḥammad ibn Yūsuf al-Ġarnāṭī, *Manhaj as-sālik fī l-kalām ʿalā ʾAlfiyyat Ibn Mālik*. Ed. Sidney Glazer. New Haven, Conn.: American Oriental Society, 1949.

Ibn al-ʾAnbārī, *ʾInṣāf* = ʾAbū l-Barakāt ʿAbd ar-Raḥmān Muḥammad ibn ʾAbī Saʿīd al-ʾAnbārī, *Kitāb al-ʾinṣāf fī masāʾil al-xilāf bayna n-naḥwiyyīna l-Baṣriyyīna wa-l-Kūfiyyīn*. Ed. Gotthold Weil. Leiden: E.J. Brill, 1913.

Ibn Jinnī, *Lumaʿ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *Kitāb al-lumaʿ fī n-naḥw*. Ed. Hadi M. Kechrida. Uppsala: Almqvist and Wiksell, 1976.

Ibn Yaʿīš, *Šarḥ* = Muwaffaq ad-Dīn Yaʿīš Ibn Yaʿīš, *Šarḥ al-Mufaṣṣal*. 10 vols. in 5. Cairo: ʾIdārat aṭ-Ṭibāʿa al-Munīriyya, n.d. [*Fahāris šarḥ al-Mufaṣṣal*. Damascus: Majmaʿ al-Luġa al-ʿArabiyya, 1990.]

Mubarrad, *Muqtaḍab* = ʾAbū l-ʿAbbās Muḥammad ibn Yazīd al-Mubarrad, *Kitāb al-muqtaḍab*. Ed. Muḥammad ʿAbd al-Xāliq ʿUḍayma. 4 vols. Cairo: Maṭābiʿ al-ʾAhrām, 1979.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *Kitāb Sībawayhi*. Ed. Hartwig Dérenbourg. 2 vols. Paris: Imprimerie Nationale, 1881. (Repr., Hildesheim and New York: Georg Olms Verlag, 1970.)

Zajjājī, *Jumal* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān ibn ʾIsḥāq az-Zajjājī, *Kitāb al-jumal fī n-naḥw*. Ed. ʿAlī Tawfīq al-Ḥamad. Beirut: Muʾassasat ar-Risāla, 1986.

Zamaxšarī, *Mufaṣṣal* = ʾAbū l-Qāsim Maḥmūd ibn ʿUmar az-Zamaxšarī, *al-Mufaṣṣal fī ʿilm al-ʿarabiyya*. Beirut: Dār al-Jīl, n.d.

Secondary sources

Badawi, Elsaid, Michael G. Carter, and Adrian Gully. 2004. *Modern Written Arabic: A comprehensive grammar*. London and New York: Routledge.

Bernards, Monique P.L.M. 1997. "Al-Mubarrad as key-figure in the development of early Arab grammatical science". *Zeitschrift für Arabische Linguistik* 34.7–30.

Blachère, Régis. 1985. *Eléments de l'arabe classique*. 4th ed. Paris: Maisonneuve and Larose.
—— and Maurice Gaudefroy-Demombynes. 1975. *Grammaire de l'arabe classique*. 3rd ed. Paris: Maisonneuve and Larose.
Cachia, Pierre. 1973. *The monitor: A dictionary of Arabic grammatical terms*. Beirut: Librairie du Liban and London: Longman.
Cantarino, Vicente. 1975. *Syntax of modern Arabic prose*. II. *The expanded sentence*. III. *The compound sentence*. Bloomington and London: Indiana University Press.
Carter, Michael G. 1972. "'Twenty dirhams' in the *Kitāb* of Sībawayhi". *Bulletin of the School of Oriental and African Studies* 35.485–496.
——. 1975. "A note on Classical Arabic exceptive sentences". *Journal of Semitic Studies* 20.69–72.
Dahdah, Antoine el-. 1988. *A dictionary of terms of declension and structure in universal Arabic grammar*. Rev. ed. Martin J. McDermott and Elias Matar. Beirut: Librairie du Liban.
Fischer, Wolfdietrich. 2002. *A grammar of Classical Arabic*. New Haven and London: Yale University Press.
Howell, Mortimer Sloper. 1883. *A grammar of the Classical Arabic language. Translated and compiled from the works of the most approved native or naturalised authors by Mortimer Sloper Howell. With a new preface by Satkari Mukhopadhaya*. 4 vols. in 7 parts. New Delhi: Gian Publishing House. (2nd ed. 1986, repr. 1990.)
Lane, Edward William. 1863–1893. *An Arabic-English lexicon derived from the best and most copious Eastern sources*. 8 vols. London and Edinburgh: Williams and Norgate.
Owens, Jonathan. 1990. *Early Arabic grammatical theory: Heterogeneity and standardization*. Amsterdam and Philadelphia: J. Benjamins.
Reckendorf, H. 1895. *Die syntaktischen Verhältnisse des Arabischen*. Leiden: E.J. Brill. (Repr., 1967.)
——. 1921. *Arabische Syntax*. Heidelberg: C. Winter. (Repr., 1977.)
Talmon, Rafael. 1993. "The term *qalb* and its significance for the study of the history of early Arabic grammar". *Zeitschrift für Geschichte der Arabisch-islamischen Wissenschaften* 8.71–113.
——. 2003. *Eighth-century Iraqi grammar: A critical exploration of pre-H̲alīlian Arabic linguistics*. Winona Lake, Ind.: Eisenbrauns.
Wehr, Hans. 1994. *A dictionary of Modern Written Arabic (Arabic-English)*. Ed. J. Milton Cowan. 4th ed. Ithaca, N.Y.: Spoken Language Services.
Wright, W. 1859–1862. *A grammar of the Arabic language. Translated from the German of Caspari and edited with numerous additions and corrections*. 3rd ed. Rev. W. Robertson Smith and M.J. de Goeje. 2 vols. in 1 part. Cambridge: Cambridge University Press. (Repr., 1967.)

Monique Bernards
(University of Groningen)

Italian

The history of the Italian territory bears a certain resemblance to that of the Iberian Peninsula because both were occupied by Arabs at the time of their expansion. The Arab influence was most strongly felt in Sicily, which was conquered in 827 and remained under Arab rule until some years after the fall of Palermo to the Normans in 1071. The effect of Arab civilization in Italy was much smaller than in Spain, and, accordingly, the number of Arabic loanwords is much smaller. A number of Arabic loanwords are attested in Sicilian dialects, especially in medieval dialects, but the majority of them were never accepted in the literary language, which is founded on the medieval dialect of Florence (a town not in direct contact with Mediterranean trade). In the late Middle Ages, linguistic contact was established between Arab populations and some maritime Italian cities, especially Pisa, Genoa, and Venice. During the Ottoman period, from the early 16th century to 1830 (when the French occupation of Algeria began), relations between Italy and the southern Mediterranean coast were mostly hostile, and the Arab world was despised and seen as the worst enemy. In the late 19th century new occasions of contact with the Arabic language arose with the colonization of Eritrea, where a variety of Sudan Arabic was employed as a trade language, and, in 1911, with the occupation of Libya.

In his seminal work *Gli arabismi nelle lingue neolatine con speciale riguardo all'Italia*, Pellegrini (1972:70–71) focuses on four types of Arabic loanwords in Italian: (1) words spread from Sicily; (2) scientific words, known in Medieval Latin and also in European languages; (3) words employed in medieval trade, spread through maritime cities and sometimes restricted to regional varieties of Italian; and (4) some foreign words (*voci peregrine*), reported by travelers and chroniclers. Perhaps one could add another type: (5) words typical of the colonial experience.

The first group is probably the largest, and examples may be quoted from many semantic fields: plants, e.g. *carciofo* 'artichoke' from Arabic *xaršūf*, *melanzana* 'eggplant' from *bad̲injān* (with a paretymological assimilation to *mela* 'apple'), *spinaci* 'spinach' from *ʾisbināx*, *carruba* 'carob' from *xarrūba*, *gelsomino* 'jasmine' from *yasmīn* (with a paretymological assimilation to *gelso* 'mulberry'), *limone* 'lemon' from *līmūn*, *cotone* 'cotton' from *quṭn*; pottery containers, e.g. *giara* (Arabic *jarra*) 'jar', *tazza* (Arabic *ṭāsa*) 'cup', *ziro* 'a container for olive oil' from *zīr*; fishing terms, e.g. *ràisi* 'the leader of a group of tuna fishermen' from Maghrebi Arabic *rāyis* 'captain', *sciàbica* 'a kind of net' from *šabaka*; some objects, like *coffa* 'a big basket of woven palm leaves; and, in ancient ships, 'the top of a mast' from *quffa*, or *risma* 'ream' from *rizma*. In ancient Sicilian documents, many other Arabic loanwords were employed (see Caracausi 1983; Scholz 1996).

Among scientific words of Arabic origin, the best known is *algebra*, coined, according to Pellegrini (1972:79), by Leonardo Fibonacci, a citizen of Pisa who was born in Bugia (today Bijāya, an Algerian town); it appears in his *Liber Abaci*, adapting the Arabic expression *ʿilm al-jabr wa-l-muqābala* 'science of reduction and comparison'. The word *algoritmo* or *algorismo* occurred in the title of a book by Gherardo da Cremona, *Liber alchorismi de iebra et almucabala*; it acquired the meaning of 'algorithm' but was originally the name of the mathematician al-Xwārizmī; the words *cifra* 'numeral' and *zero* were two different adaptations of Arabic *ṣifr*. Many names of stars are derived from Arabic, e.g. *Aldebaran* from *al-dabarān* (without assimilation of the article), *Vega* from *al-wāqiʿ*, and also words connected with astronomy, e.g. *zenit* 'zenith' from *samt* (English *azimuth* is derived from the plural *as-sumūt* Italian *azimut*), *alidada* 'alidade' from *al-ʿidāda*, and *almanacco* 'almanac' from *al-manāx* (about these 'cultivated' loanwords see also Gleßgen 1996).

Many words are typically connected with medieval trade. One could mention *fondaco*, which was a sort of accommodation for traders, with a warehouse and the possibility of selling. It derives from Arabic *funduq*, which has come to mean 'hotel' in Egyptian Arabic but in Tunisian Arabic retains the meaning 'caravanserai'. One could also quote the modern Italian word *magazzino* 'warehouse', derived from *maxzin*, for which Pellegrini (1972:105) mentions an eastern Algerian form, *maġzin*, as the true etymon, but a regressive assimilation like this can be found almost anywhere in substandard Arabic; *dogana* (in Medieval Italian the form *duana* was used), derived from Arabic *dīwān* 'office'; and *tariffa* 'tariff rate', derived from Arabic *taʿrīf* 'notification'. In Italian, one also finds many names of merchandise known through medieval trade: some spices, e.g. *zafferano* 'saffron' from *zaʿfarān*, *tamarindo* 'tamarind' from *tamar hindī*, *curcuma* 'turmeric' from *kurkum;* and also *zucchero* 'sugar' from *sukkar*, *ambra* 'amber' from *ʿanbar*, *sandalo* 'sandal(-wood)' from *ṣandal*; and probably also some minerals, e.g. *catrame* 'tar' from *qaṭrān*, *talco* 'talc' from *ṭalq*; perhaps also *soda*, which in Medieval Italian had the meaning of 'alkali' and may be derived from Arabic *sawwād*, the name of a plant whose ashes contain sodium. Very typical are the variants *darsena* 'dock' and *arsenale* 'shipyard', both derived from *dār aṣ-ṣināʿa*, the former being a word from Pisa, the latter from Venice.

All these 'ancient' words exhibit some phonetic adaptations (not phonetic laws). As noticed by many scholars, most Arabic loanwords in Spanish have an assimilated article (*aceite* from *az-zeit*, *alcázar* from *al-qaṣr*), whereas Arabic loanwords in Italian do not, with the exception of 'scientific' words. Final *-ī* is assimilated to the suffix *-ino*, which in Italian can form diminutives, but also some ethnonyms, such as *triestino* from Trieste; with this suffix are formed *tunisino* from *tūnisī* 'Tunisian', *beduino* 'Bedouin' from *badawī*, *assassino* 'murderer' from *ḥaššāšī* 'member of the sect of Assassins', perhaps also *facchino* 'porter' from *faqīh* 'expert in Islamic law' (final *-h* in Arabic is pronounced very weakly; for the change of meaning see below). Some words show a dialectal origin from the Maghreb: the model of *beduino* probably was not the classical form *badawī* but an occidental form *[bɛdwi]; also *benzoino* 'benzoin' is probably not derived from *lubān jāwī*, literally 'incense of Java', but from a Maghrebi form *[lbɛːn ʒaːwiː], where initial *l-* was mistaken for the article and then eliminated; and the expression *a bizzeffe* 'in large quantity' is clearly from Algerian *bizzāf* 'a lot', derived from Classical Arabic *bi-l-juzāf* 'in a big amount'.

During the long period of cultural and economic decline of the Arab world, some words were borrowed with pejorative meanings. The most typical case is *facchino*, attested in the 15th century: in some Arabic documents of the late Middle Ages the word *faqīh* is attested with a generic honorific meaning, but *facchino* was employed in Italy as a word for qualifying foreign workers, like French *sidi*; now it means 'porter'. *Cabibbo*, from Arabic *ḥabīb*, is a semi-slang word to indicate people from southern Italy. *Crumiro*, from the Tunisian tribe of Khumirs, means 'blackleg, one who works during a strike' (these etymologies are treated in Pellegrini 1972:503–523; see also Cifoletti 1984; Orioles 1984).

Arabic loanwords of the colonial period have received less attention, but some of them are easily recognizable: *ghirba* 'water skin', attested since 1881, is probably from Sudan Arabic *girba* (classical *qirba*); *carcadè* or *karkadè* 'roselle tea, made with flowers of a kind of hibiscus' is a well-known word in Egyptian and Sudanese Arabic, pronounced [karkadeː]; *uadi* 'a valley without river, typical of dry climates', attested since 1895, probably goes back to Egyptian Arabic *wādī* (in Sudan, the current word for this meaning is *xōr*; in the Maghreb this word is pronounced [wæːd] and means 'river'); and *ghibli*, designation of the south wind in Libya, attested since 1912, derives from a local pronunciation for *qiblī*.

In more recent times, some Arabic loanwords are reproduced in the lexicon of the press and international politics, which is why many Italians know what *jihad* means (in Italian, it is treated mostly as a feminine substantive, under the influence of *guerra* 'war'), or *raìs* (a president of an Arabic republic), or *fatwa*. Probably most of these words have been introduced through English or French. There is perhaps one word that has gained a certain popularity because of the popularity of the object: the typical headgear of Palestinian men, called (in Italian) *kefiyyah* or *kefiah*.

Bibliographical references

Caracausi, Girolamo. 1983. *Arabismi medievali di Sicilia.* Palermo: Centro di Studi Filologici e Linguistici Siciliani.

Cifoletti, Guido. 1984. "L'etimologia di 'facchino'". *Incontri Linguistici* 9.155–158.

Gleßgen, Martin-Dietrich. 1996. "Die mittelalterliche Übersetzungsliteratur und ihre sprachlichen Auswirkungen: Forschungslinien in der Italia Arabica". Lüdtke (1996:191–200).

Lüdtke, Jens (ed.). 1996. *Romania Arabica, Festschrift für Reinhold Kontzi.* Tübingen: Narr.

Orioles, Vincenzo. 1984. "Italiano 'crumiro' come effetto di irradiazione sinonimica". *Incontri Linguistici* 9.159–164.

Pellegrini, Giovan Battista. 1972. *Gli arabismi nelle lingue neolatine con speciale riguardo all'Italia.* Brescia: Paideia.

Scholz, Arno. 1996. "Gli arabismi siciliani: prospetto riassuntivo dei principali studi". Lüdtke (1996:169–189).

Guido Cifoletti (University of Udine)

Italian Loanwords

This entry addresses Italian loanwords in some Arabic dialects, especially Egyptian Arabic (the dialect of Cairo, with some mention of Alexandria) and Tunisian Arabic (the dialect of Tunis). Because of its proximity, Tunisia has always been in contact with Italy; likewise, the dialect of Libya has been strongly influenced by Italian, but not much is known about it, and after Rossi (1933) the topic has hardly been studied. For Italian borrowings in some Oriental dialects, see Abou Abdallah (1981), Butros (1973), and Behnstedt (1996).

In the centuries when the Arab civilization flourished, borrowings were mostly from Arabic to the European languages, especially Medieval Latin (which only became well differentiated from Italian in the 13th century). There was a long period in which Italian was the most prestigious of the Mediterranean languages. Although not much is known about this period, it probably began in the early 15th century and ceased about 1850. During this time, many treaties, diplomatic correspondence, and official documents were written in Italian, or at least in what may count as Italian (cf. Bruni 1999, 2003). Especially in the Ottoman Empire, Italian was the language of foreign relations: some of the most important documents, for example the decree that authorized Lord Elgin to take the sculptures of the Parthenon to England, were written in Italian, probably by Greek secretaries. Moreover, from the beginning of the 15th century the maritime language of the Mediterranean developed on the basis of Italian (cf. Kahane a.o. 1958); this lingua franca was observed and partially recorded by some German travelers and pilgrims (Battisti 1962; Milani 2000; Röll 1967).

For the Arab world, this was perhaps the period in which cultural stagnation and economical decline were at their worst and the prestige of the Arabic language was at its lowest level. When the modern age began in Egypt with the government of Muhammad Ali, Italian was still the current language for contacts with the thousands of foreigners who settled in the country, many of whom were Italians. Until 1860, Italian was the high-level language for Egypt, when the country was organizing all public services like mail, railways, hospitals, and courts of law. The subsequent loss of influence of Italian and the increase in the French and English influence are well documented (Sammarco 1937; see Balboni 1906 for an eyewitness report of the decline of Italian in Egypt). Some of the loanwords in use in Egypt at the beginning of the 20th century are now obsolete (see Spiro 1904), but many others are still current.

Borrowing was facilitated by some accidental similarities between Italian and Arabic: both languages have geminate consonants, both have a feminine ending *-a*, and the Italian masculine singular *il* is very similar to the dialectal Arabic article. The problem was that many Italian words were too long for Arabic; in this case, aphaeresis often presented a solution. Italian *insalata* 'salad', for instance, is known in Modern Arabic as *salāṭa* (note that *salata* is used in some Italian dialects, in particular Venetian); Italian *arrosto* is in Egypt *rustu* 'roast', which is nearer to Venetian *rosto*. In some cases the first Italian syllable is simply omitted, without any model in Romance languages. Tunisian Arabic *furnāṭa* 'batch', for instance, is derived from Italian *infornata*, but with aphaeresis; likewise, Egyptian Arabic *sibinsa* 'caboose, guard's van [of a train]' probably derives from Italian *dispensa* 'pantry'. In all these cases, the deleted syllable corresponds to an Italian preposition or to a prefix.

Another difficulty is the presence in Italian of an ending *-o*, which usually marks the masculine. In spoken Arabic, the ending *-u* can be the ending of some plural forms of the verb, or the possessive suffix for the 3rd person masculine. Substantives with such an ending are exceptional. Only in Tunisian Arabic there are some words that probably derive from the Latin substrate, like *furnu* 'oven' < Latin *furnum, knāstru* 'a traditional basket of the bride' < Latin *canistrum*, *ginnāru* 'chicken coop' < Latin *gallinarium*, etc. Perhaps these words opened the way for the introduction of new words with the same ending. This explains why many Italian words lost the ending *-o*, as in Egyptian Arabic *baskūt* 'biscuit(s)' < *biscotto*, or Tunisian Arabic *jīlāt* < *gelato*. Some other words changed the vowel into *-a*, like *ricamo* 'embroidery', which became *rukāma* in Egypt. But the majority maintained final -o, pronounced [u] in Tunisia or [ʊ] in Egypt, in words like (Egyptian and Tunisian) *kámbiyu* 'change [of money]' < *cambio; brutistu ~ burtistu* 'protest' < *protesto; abukātu* 'lawyer' < *avvocato; karru* 'animal-drawn cart' < *carro*. In most cases, a final *-e* of Italian was rendered as *-a*, which in many Arabic dialects has a fronted pronunciation (in a non-emphatic context): so Italian *cartone* 'cardboard' > Arabic *kartōna*, or *balcone* 'balcony' > *balkōna*.

Some Italian consonants have no corresponding sounds in Arabic: /p/, /v/, /tš/, /dj/, /ts/, /dz/, /lj/, /nj/. While in some Arab countries Italian /dj/ corresponds to the normal realization of *jīm*, in others it is a foreign sound. The sound /p/ is normally realized as /b/ and is pronounced voiceless only by cultivated persons. For /v/ the situation is more complex, because in the past (probably during the Ottoman period) it was rendered as /w/, and this pronunciation survives in some words like *lawanda* 'lavender' < *lavanda*. In Egypt, there is also a popular pronunciation as /b/, which is prevalent in words like *bitillu* 'veal' < *vitello* 'calf, veal'. /f/ for /v/ is very common in Egypt and also in Tunis, in words like Egyptian *vanilya ~ fanilya* 'vanilla' < *vaniglia, karafatta* 'necktie' < *cravatta*, or Tunisian *fīsta* 'clothing' < *veste*. The sound /č/ was rendered normally with /š/ as in Modern Arabic *šīkūrya* 'chicory [used as a substitute for coffee]' < *cicoria*; but when geminated, it is rendered by /tš/ (which in Tunisian and Egyptian Arabic is not a phoneme but a cluster), as in Egyptian *bilyatšu* 'clown' < *pagliaccio*, Tunisian *fātša* 'face' < *faccia*, or Tunisian *fatšāṭa* 'façade' < *facciata*; and also in the recently spread (Tunisian and Egyptian) *kabutšīnu* < *cappuccino*. Tunisian Arabic has the sound /j/ as the normal realization of *jīm* and employs it for Italian /dj/. The Cairene dialect (which has /g/ for *jīm*) has a phoneme /j/ which occurs only in foreign words and did not become popular until the 20th century (also for Italian words, such as *jilāti* 'ice cream' < *gelati*, or *jakitta* 'jacket' < *giacchetta*). In older loanwords the pronunciation /g/ is kept, as in *gamadāna* 'demijohn' < *damigiana* (with metathesis), or *garunya* 'geranium' < *geranio*. Complex Italian sounds like /ts/ and /dz/ are normally rendered by sibilants: Italian *pinza* 'pliers' (with /ts/) becomes *binsa* in Egyptian;

dozzina 'dozen' (with geminated /dz/) becomes *dazzīna* in Alexandria (not used in Cairo) and *tuzzīna* in Tunis; *gazzosa* 'soda water' (also with geminated /dz/) is *kazōza* in Egypt and *gazūz* in Tunisia; *terrazzina* 'terrace' is *tarasīna* in Cairo; only the word *mozzarella*, recently spread, is *mutsarilla* in Egypt and Tunisia with the cluster /ts/ for a geminate Italian *z*. Palatal continuous consonants of Italian are rendered with the clusters /ly/ and /ny/: for example, Egyptian *bastilya* 'small sugar candies, pastilles' < *pastiglia* 'pastille'; Tunisian *trilya* 'mullet' < *triglia*; and *bunya* 'punch' (known in Egypt, Tunisia, and also in Algiers and Lebanon) < *pugno*. Sometimes in Italian there is a consonantal cluster (two consonants at the beginning, or three within the word), which is not tolerated in Egyptian Arabic (nor in Classical Arabic). This cluster is dissolved by the insertion of a vowel: so we have in Cairo *kirēma* 'cream' < *crema; garanīṭa* 'grated-ice drink' < *granita*; and *kuntirātu* 'agreement' < *contratto*. Cases like *istubba* < *stoppa* 'oakum', or *istabēna* from the expression *sta bene* 'that's all right', probably go back to an earlier period.

Italian stress is almost always maintained: when it is on the penultimate syllable and followed by only one consonant, in Egyptian Arabic, the stressed vowel is lengthened, as in *kabīna* < *cabina* 'cabin; bathing hut', or in *nōta* < *nota* 'note'.

The great majority of Italian loanwords are nouns. Some of them are without plural, as for example all names of diseases: Egyptian *malaria* < *malaria; azma* 'asthma' < *asma* (this word is homophonous to another *azma* 'crisis', so that some speakers believe the borrowed meaning is a semantic expansion). Some Italian plurals are taken in a collective sense, as Tunisian *bambalūni* < *bomboloni* [pl.] 'a kind of pancake', but some loanwords have developed a plural in *-āt*, e.g. Egyptian *kambiyāla* 'bill of exchange' < *cambiale*, pl. *kambiyalāt*. In some masculine nouns this formation can be interesting, as in Egyptian *banyu* 'bathtub' < *bagno* 'bath', pl. *banyuhāt*, or Tunisian *būnu* 'voucher' < *buono*, pl. *būnuwāt*. Many loanwords have a broken plural, e.g. Egyptian *kabīna* 'cabin', pl. *kabāyin*; or Tunisian *kalsīṭa* 'sock', pl. *klāsiṭ* < *calzetta*. In Egyptian Arabic some words have been integrated in the paradigm nomen unitatis/collective, e.g. *bilya* < *biglia* 'small ball, marble', pl. *bily; kartōna* < *cartone,* which means 'one box of cardboard', while the material is *kartōn*; also *benzīna* < *benzina* 'gasoline', which means 'gas station', while *benzīn* means 'gasoline'.

The great majority of borrowed nominals are substantives, but there are also some adjectives: they are integrated in a category which is common in many Arabic dialects, that of the uninflected adjectives. Both Egyptian and Tunisian Arabic have *fīnu* < *fino* 'fine, fine-textured'; also *falṣu* < *falso*, meaning 'false, counterfeit'. Egyptian has also *birīmu, sikundu, tirsu* < *primo, secondo, terzo* 'first', 'second', 'third', meaning 'first-, second-, third-class'; Tunisian has *dilikātu* < *delicato* 'delicate, weakly'. No Italian verbs have been borrowed.

Semantic changes are frequent. Some are obvious, as in the case of *ballerina* 'dancer', which became in Egyptian Arabic *ballirīna* 'ballet dancer'; others are more interesting, e.g. Egyptian *manafilla* < *manovella* 'crank', which is used for the crank of old cars and also for the foot lever to switch some motorcycles. Egyptian *ballu*, at the beginning of the 20th century, was a dance soirée in European style but now is only a hullabaloo; Tunisian *armatūra* < *armatura* probably originally indicated all equipment of a ship but now may be used also for the goods of a shop. Some words retain the various meanings of Italian: *kabīna* < *cabina* in Egypt and also in Tunisia; it can be a ship's cabin, a telephone kiosk, or a bathing hut. Also *baṭṭāriyya* < *batteria*, known in Egypt and Tunisia, can be the battery of a car, a battery of guns or of drums, or (only in Egypt) an electric torch. Some Italian words have become current in Modern Standard Arabic, e.g. *mōḍa* 'fashion' < *moda, mobilya* 'furniture' < *mobilia* (now obsolete), *fātūra* 'invoice, bill' < *fattura, ubira* 'opera' < *opera, ṣāla* 'hall' < *sala, kumudīnu* 'bedside table' < *comodino*, etc.

The semantic domains in which most Italian borrowings are found are gastronomy and the old maritime terminology, but in the past there were many other domains, especially in Egyptian Arabic. The many commercial terms include, besides the above-mentioned *fātūra, kambiyāla, kuntirātu, kambiyu, istabēna, burṣa* 'stock exchange' < *borsa* (still current both in Egypt and Tunisia), *brutistu* ~ *burtistu* 'protest' < *protesto, marka* 'mark, check' < *marca, busṭa* 'post, post office' < *posta,* and (only Tunisian) *afā̊r* 'bargain' < *affare*, also the expression (Egyptian, but also well known in Tunis) *āla ūna, āla duwwa, āla trē*, which closes an auction sale, < *uno, due, tre* 'one, two, three'. Nouns referring to music and theater, like the above-mentioned *ubira, ballirīna*, and *nōta*, include (Egyptian) *kumidya* 'comedy' < *commedia, urkístira* 'orchestra' < *orchestra, mayístiru* 'conductor' < *maestro, prīma dunna* ~ *birīmadunna* ('prima donna' < *prima donna*), *trumbēta* 'trumpet' < *trombetta*.

Finally, it is surprising to find some words of Italian origin for elementary concepts that could have been expressed by Arabic words. Because of their diffusion in many modern dialects, these words were probably borrowed at an early period, e.g. *familya* 'family' < *famiglia*, surely a non-necessary loanword, and also *bunya* 'punch' < *pugno* (in other Romance languages, corresponding words mean only 'fist') or *mākīnā* (pronounced in Egypt *mákana*) < *macchina* 'engine'.

BIBLIOGRAPHICAL REFERENCES

Abou Abdallah, Albert. 1981. "Prestiti italiani nel dialetto libanese." *Quaderni Patavini di Linguistica* 2.189–222.

Baccouche, Taïeb. 1994. *L'emprunt en arabe moderne*. Tunis: Institut Bourguiba.

Balboni, Luigi. 1906. *Gli Italiani nella Civiltà Egiziana del secolo XIX*. 3 vols. Alexandria: Società Dante Alighieri.

Battisti, Carlo. 1962. "Termini marinareschi italiani in una canzone di Osvaldo di Wolkenstein". *Il Cristallo* 4.23–42.

Behnstedt, Peter. 1996. "Romanisches Lehngut im Syrisch-Arabischen". *Romania Arabica: Festschrift für Reinhold Kontzi*, ed. Jens Lüdtke, 63–72. Tübingen: G. Narr.

Bruni, Francesco. 1999. "Lingua d'Oltremare: Sulle tracce del 'Levant Italian' in età preunitaria". *Lingua Nostra* 60.65–79.

——. 2003. "Italiano all'estero e italiano sommerso: Una lingua senza impero". *Storia della lingua e storia*, ed. Gabriella Alfieri, 179–198. Firenze: F. Cesati.

Butros, Albert. 1973. "Turkish, Italian and French loanwords in the colloquial Arabic of Palestine and Jordan". *Studies in Linguistics* 23.87–104.

Cifoletti, Guido. 1986. *Prestiti italiani nel dialetto del Cairo*. Milano: UNICOPLI.

——. 1998. "Osservazioni sugli italianismi nel dialetto di Tunisi". *Incontri Linguistici* 21.137–153.

Kahane, Henry, Renée Kahane, and Andreas Tietze. 1958. *The lingua franca in the Levant: Turkish nautical terms of Italian and Greek origin*. Urbana: University of Illinois Press.

Milani, Celestina. 2000. "Da Venezia al Vicino Oriente: Veneziano e lingua franca in diari di viaggio in Frühneuhochdeutsch". *Processi di convergenza e differenziazione nelle lingue dell'Europa medievale e moderna*, ed. Fabiana Fusco, Vincenzo Orioles, and Alice Parmeggiani, 353–374. Udine: Forum.

Röll, Walter. 1967. "Zur Lingua franca". *Zeitschrift für romanische Philologie* 83.306–314.

Rossi, Ettore. 1933. "Vocaboli stranieri nel dialetto arabo della città di Tripoli". *Atti del IV Congresso Internazionale dei Linguisti*, 186–193. Rome.

Sammarco, A. 1937. *Gli italiani in Egitto*. Alexandria: Edizioni del Fascio.

Spiro Bey, Socrates. 1904. *Note on the Italian words in the modern spoken Arabic of Egypt*. Cairo.

GUIDO CIFOLETTI (University of Udine)

ʾIṭbāq

The term *ʾiṭbāq*, derived from the Arabic root *ṭ-b-q*, generally means 'covering, e.g. a lid covering a pot'. As a phonetic term, it is defined by Sībawayhi as "the raising of the (back of) the tongue toward the velum" (*Kitāb* IV, 436). While some modern phoneticians call *ʾiṭbāq* 'emphasis' (Vollers 1893:147), others speak of 'velarization' (Gairdner 1925:20). Gairdner defines velarization as an articulation in which "the back of the tongue is raised towards the back of the velum, i.e. the extreme back of the palate. The tongue feels as if it 'fills the mouth'. This velarization was described by the old Arab phoneticians as a 'lidding' – they said that the tongue seems to fill the cavity above like a lid".

There are four velarized consonantal phonemes in Arabic, namely /ṣ/ (*ṣād*) ص, /ṭ/ (*ṭāʾ*) ط, /ḍ/ (*ḍād*) ض, and /ḏ̣/ (*ḏ̣āʾ*) ظ. Sībawayhi describes them as velarized consonants (*ḥurūf muṭbaqa*) in contrast with the rest of the Arabic consonants, which are described as nonvelarized (*ḥurūf munfatiḥa*). In his explanation of the articulatory processes of the velarized consonants, Sībawayhi states: "In these four consonants, if you apply your tongue in their place, it will close on them from their primary positions of articulation up to that part of the tongue opposite the soft palate, toward which you raise the tongue. Applying the tongue this way the sound will be enclosed between the tongue and the soft palate on the one side, and the places of the consonants on the other side" (*Kitāb* IV, 436). Again, Gairdner explains this process as follows: "We have called it velarization because the most prominent feature of the phenomenon is the raising of the back of the tongue towards the soft palate. But in addition to this, the blade of the tongue is tensed and spread, which involves a raising, and the middle part of the tongue, though lower than the back, is doubtless raised also" (1935:250).

Two analytic systems for describing the speech sounds of Arabic are employed in Classical Arabic phonetic scholarship: the description according to place and manner of articulation, on the one hand, and the description according to binary or tertiary feature analysis, on the other.

As for /ṣ/, Sībawayhi describes it, together with the other sibilants, /s/ and /z/, as being articulated "between the tip of the tongue and the upper part of the incisors" (*Kitāb* IV, 433ff.). In addition, /ṣ/ is described as a voiceless, velarized, fricative consonant.

The consonant /ḍ/ is described as being articulated "between the front part of the edge of the tongue and the molars lying next to it" (→ *ḍād*; Sībawayhi, *Kitāb* IV, 433). In terms of manner of articulation, it is a voiced, velarized, lateral, fricative consonant (Sībawayhi, *Kitāb* IV, 435ff.). In modern Arabic dialects, /ḍ/ is pronounced either as a plosive or fricative consonant. There is hardly any trace of laterality in the modern Arabic dialects.

The consonant /ṭ/, just like /d/ and /t/, is described as being articulated "between the tip of the tongue and the tooth-holes/tooth ridge of the incisors" (Sībawayhi, *Kitāb* IV, 433), i.e. alveolar. It is also described as a voiced, velarized, plosive (Sībawayhi, *Kitāb* IV, 434ff.). In the modern pronunciation of Arabic, however, /ṭ/ is a voiceless consonant.

As for /ḏ̣/ ظ, Sībawayhi describes it as being produced "by the tip of the tongue and the edges of [the upper and lower of] the front incisors" (*Kitāb* IV, 433), just like /ḏ/ and /ṯ/. It is also described as a voiced, velarized, fricative consonant (Sībawayhi, *Kitāb* IV, 434ff.).

In all these cases, the Arab and Muslim phoneticians made it very clear that the feature of *ʾiṭbāq* or velarization is like an additional characteristic of the four consonants. Sībawayhi points to the fact that in the pronunciation of each of those four consonants the articulation is made in two places. He remarks: "Each of these four [consonants] has two places [of articulation] in the tongue" (*Kitāb* IV, 436). In modern phonetics these are termed 'primary and secondary articulations'. Ladefoged analyzes velarization as a secondary articulation or coarticulation which "involves raising the back of the tongue. It can be considered as the addition of a [u]-like tongue position, but without the addition of the lip rounding that also occurs in [u]" (1975:208). According to Sībawayhi, the difference between the nonvelarized consonant and the velarized counterpart lies in the absence of *ʾiṭbāq* in the former and its presence in the latter. Thus, he states: "Without *ʾiṭbāq*, /ṭ/ would be a /d/, /ṣ/ would be an /s/, /ḏ̣/ would be a /ḏ/, and /ḍ/ would not exist, because it has no nonvelarized counterpart" (*Kitāb* IV, 436). The statement applies to Classical Arabic only, as both /ṭ/ and /ḍ/ have undergone phonetic changes which make them different from their cognates in Modern Arabic.

Velarization, however, is not as simple as that. It is a more involved process. Apart from the raising of the back of the tongue, it also involves multiple articulations. Other secondary articulations such as pharyngealization (constriction of the pharynx), labialization (some degree of lip rounding), and glottalization (a simultaneous glottal constriction) may contribute to velarization.

A form of binary distinctive feature analysis is also concomitantly employed in the traditional Arabic phonetic description, as is obvious from the statement cited above, according to which velarized consonants are distinguished

Figure 1. Matrix of distinctive features of velarized and nonvelarized consonants

features	ṣ	ḍ	ḏ̣	ṭ	s	d	ḏ	t	b	ṯ	j	ḥ	x	r	z	š	ʾ	ġ	f	q	k	l	m	n	h	w	y	a	a:	u	u:	i	i:	ʿ
consonantal	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	+
vocalic	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	-
velarized	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
voicing	-	+	+	+	-	+	+	-	+	-	+	-	-	+	+	-	o	+	-	-	-	+	+	+	+	-	+	+	+	+	+	+	+	+

key to symbols
+ = presence of the feature
- = absence of the feature
o = neutral

from nonvelarized ones. The matrix in Figure 1 outlines some of the distinctive features distinguishing the four velarized consonants from their nonvelarized counterparts.

Acoustically, velarization manifests itself in a spectrogram as a lowering of the second formant in speech segments. Concluding his experiments on velarized vs. nonvelarized pairs of Arabic segments, Obrecht remarks:

> Though F2 transitions are obviously a powerful cue in the perception of velarization, a separate transitional segment cannot be identified and said to function with constant effectiveness throughout the system. The degree of separability of a visible transition portion of a pattern depends, naturally, on the type of consonant involved and its vocalic environment. (1968:39)

The analysis by the Arab and Muslim phoneticians of the phenomenon of *ʾiṭbāq* or velarization constitutes a significant contribution to the study of Arabic speech sounds, which helps us to reconstruct the sound system of the time. It is due to their meticulous phonetic description that the modern Arabic sound systems can be understood.

Bibliographical references

Primary source

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān ibn Qanbar Sībawayhi, *al-Kitāb*. Ed. ʿAbd as-Salām Muḥammad Hārūn. Cairo: al-Hayʾa al-Miṣriyya al-ʿĀmma li-l-Kitāb, 1975.

Secondary sources

Gairdner, W.H. Temple. 1925. *The phonetics of Arabic*. Humphrey, Milford: Oxford University Press.

——. 1935. "The Arab phoneticians on the consonants and vowels". *Muslim World* 25.242–257.

Ladefoged, Peter. 1975. *A course in phonetics*. New York: Harcourt Brace Jovanovich.

Obrecht, Dean H. 1968. *Effects of the second formant on the perception of velarization consonants in Arabic*. The Hague: Mouton.

Vollers, Karl. 1893. "The system of Arabic sounds, as based upon Sibawaih and Ibn Yaish". *Transactions of the Ninth International Congress of Orientalists*, ed. E. Delmar Morgan, II, 130–154. London.

Muhammad Hasan Bakalla
(Emeritus Professor, King Saud University)

Ivrit

The penetration of Arabic loanwords into Modern Hebrew (Ivrit) is vast. Some of these loanwords derive from *fuṣḥā* (Literary Arabic), some from the dialects, especially the Palestinian dialects, and a rare few from the Jewish Arabic dialects of other regions. All layers of Israeli society, including the media and modern literature and poetry, rich and poor, Jews of Ashkenazi or Sephardi origin, use Arabic loanwords in their speech. There seems to be no geographical connection with the degree of usage of Arabic loanwords, but it is possible that younger people, especially those who have been in the army, use a higher percentage of such words in comparison with other age groups.

The phonological developments in Ivrit – the loss of emphatic and guttural consonants such as *ḥ* and *ṭ*: *q* > *k*, *ṭ* > *t*, *ḥ* > *x*, ʿ > ʾ for most of the speakers; some changes in the vowel system, such as the loss of vowel length; and the shift of the stress – have led to some misconceptions about the etymology of loanwords of Arabic origin. Therefore, the Ivrit words in this entry will be written out according to the modern pronunciation. An example would be the Arabic loanword *ṭafrān* 'penniless' (Ivrit *tafrán*, with the same meaning). The root *t-f-r* (and not *ṭ-f-r*) exists in Ivrit with the meaning 'to sew'; with the loss of emphatics in Ivrit, the Arabic loanword in *faʿlān* created another meaning to this loanword, namely 'a tailor', hence 'poor as a tailor'. The word *tafrán* has become naturalized to such an extent that Israeli speakers do not think of it as an Arabic loanword. Arabic loanwords that have changed their original vowels have entered Ivrit, e.g. *yaʾéni* 'as though', derived from the Arabic *yaʿnī* 'that is'. Some Ivrit speakers attribute a folk etymology to this Arabic loanword and think that it is derived from the Arabic word *ʿayn* 'eye'. Apart from *yaʾéni*, the Arabic filling word *yaʿnī* exists in Ivrit with the same meaning as in Arabic (pronounced *yaʾni*).

When dealing with Arabic loanwords in Ivrit, a distinction must be made between loanwords and loan translations. The Arabic words that have penetrated into Ivrit can be divided into two groups: a natural penetration through the spoken language and Modern Hebrew literature, and a planned insertion of Arabic loanwords, especially by numerous innovators beginning with Eliezer Ben-Yehuda, the Hebrew Language Council (Vaʿád ha-Lašón ha-Ivrít) in the late 1920s, and, since 1952, the Academy of Hebrew Language (ha-Akadémya la-Lašón ha-Ivrít).

Ben-Yehuda often borrowed from Literary Arabic. His innovations, based on Literary Arabic vocabulary, occurred on different levels. One level is that of direct copies of the Arabic word or root, e.g. *adív* from Arabic *ʾadīb* 'polite', *mehagér* < *muhājir* 'immigrant', *retsiní* 'serious' from Literary Arabic *raṣīn*, *bubá* 'doll' from *buʾbuʾ* (lit. 'pupil of the eye'), *ribá* 'jam' < *murabbā*, *mivrák* 'telegram' < *barqiyya*, *kidmá* 'progress' < *taqaddum*. A second level is that of loan translations based on Arabic, e.g. *afór* 'gray' from Hebrew *éfer* 'ash', since the Arabic word for 'gray', *ramādī*, is related to *ramād* 'ash'. The third level of Ben-Yehuda's innovations was based on morphology, e.g., the Ivrit word *miṭriyá* (pronounced *mitriyá*) 'umbrella' has the suffix *-iya*, imitating the Arabic word *šamsiyya* (lit. 'parasol'), with a similar suffix *-iyya*. Another example is *iriyá* 'municipality' from the Hebrew *ʿīr* (pronounced *ʾir*) 'city', based on the Arabic *baladiyya*.

Ben-Yehuda was not the only one to innovate words based on Arabic. Mirkin (1902) proposed the combination *ʾeglát kitór* for the English *steam car* (German *Dampfwagen*; French *wagon à vapeur*). Yalin objected to this innovation and said that this new loanword based on European idioms could be understood in the developing Ivrit as 'vapor'. Therefore, he preferred Arabic *qiṭār*, which could be built according to the Hebrew morphological structure of *faʿʿāl*, hence *qattār* (today *katár*, meaning 'locomotive'). In the 1930s and 1940s some neologisms were based directly on Arabic but did not survive, e.g. *baddorá* from Arabic *bandōra* 'tomato'. The word used today is *agvanyá*, derived from the Hebrew root *a-g-v* 'to lust for', based on the supposition that the tomato is an aphrodisiac. Another example is *ʾegóz hódu* (lit. 'nut of India'), meaning 'coconut', as a loan from the Arabic *jawz hindī*. The prevalent Ivrit word today is *egóz kókus*. Two other examples of Ben-Yehuda's innovations that have not survived are *latíf* < *laṭīf* 'gentle, cute', today *nexmád*; and the invention *mózen ha-ʾavír* from Arabic *mīzān naql al-hawāʾ* 'barometer', a term that was not accepted at all, as Ivrit today uses *barométer*. Piamenta (1961) quotes Ben-Yehuda as saying that Arabic holds out hope for the enrichment of the Modern Hebrew language because its words preserve the ancient etymologies of the Semitic languages.

Already at the beginning of the Hebrew Language Council's work one finds words derived from Arabic, e.g. *ṭappīl* 'parasite' (nowadays *tapíl*) from *ṭufayliyy*. In recent years there has been a tendency in the Academy of Hebrew Language to derive Arabic loanwords, particularly in the field of botany, e.g. *xilbá* 'fenugreek' from Yemenite Arabic *ḥilbe*. In 2003 the Academy decided that the official word for 'pine nut' should be changed from *ʾóren ha-sélaʾ* to *tsnovár* (pronounced today *tsnóbar*), from Arabic *ṣanawbar*, a word which had been naturalized by the Israeli public in any case. The plant *prosopis* (English 'burdock') was named by the Academy after the Arabic *yanbūṭ* (with the same pronunciation, *yanbūṭ*, Spoken Hebrew *yanbút*), although this plant is mentioned in some old Hebrew sources as *kalīs*.

Blanc (1955) divides the Arabic loanwords in Ivrit into groups:

i. Words for food, e.g. *xúmus* < *ḥummuṣ*, *txína* < *ṭaḥīna*. *Faláfel*, the representative and symbol of Israeli food, is a loanword from the Arabic *falāfil* (pl.). It is used in the singular, not as in Arabic, e.g. *faláfel taʾím* 'a tasty falafel'.
ii. Words connected to Israeli society, especially those adopted before the establishment of the state of Israel in 1948, e.g. *xamsín* < *xamsīn* 'hot and dry weather', or the head covering *kafiya* < *kufiyye* 'kefiyyeh'.
iii. Blessings, curses, and interjections, e.g. *ʾáhlan* or *ʾálen* < *ʾahlan* 'welcome, hello!', *ʾináʾal ʾabúk* < *ʾinnaʿal ʾabūk* 'may your father be cursed!' or the well-known *kuss ʾimmak* 'motherfucker!', or *xabíbi* < *ḥabībī* 'my dear'.
iv. A wide range of adjectives or adjuncts, e.g. *zift* 'a very bad thing' < *zift* (etymologically 'pitch') or *jéda* 'a strong and vigorous woman' < *jadaʿ* 'a vigorous and courageous man'.

Ben-Amotz is responsible for a large contribution of Arabic loanwords in Ivrit. His dictionary of spoken Ivrit (1972), written together with Ben-Yehuda, is a landmark in the history of the language for its audacity and daring. Many words in this dictionary are curses belonging to certain sociolinguistic registers, like army language and street language, e.g. *ʾaxú lmanyúki* < *ʾaxú šarmúta* < *ʾaxū š-šarmūṭa* ‘the brother of the prostitute’. A very interesting phenomenon is that many curses derived from Arabic have both negative and positive meanings in Ivrit. For example, one can say *ʾasíti ʾaxú lmanyúki šel ʾavodá*, meaning ‘I did a very good job’. It is worth noting that various entries in Ben-Amotz’s dictionary are not known to many people nowadays and it even seems that some of the entries were already idiolects or obsolete words at the time and are considered archaic today, e.g. *furšáye* ‘bad and superficial work’; in some Palestinian Arabic dialects, this word has the meaning of ‘partial sexual intercourse’, whereas in all Palestinian dialects it means ‘toothbrush’.

Some loanwords were taken from Palestinian Arabic dialects like that of the Triangle Area in modern Israel (*al-muṯallaṯ*), e.g. *číl̀ba* ‘to be angry with’ from *čílba* (Literary Arabic *kalba(tun)* ‘a bitch’) or *čizbát* ‘fib’, used as a singular noun in Ivrit, from *čizbāt* ‘fibs’ (Literary Arabic *kiḏbāt*). We occasionally find Arabic loanwords in Ivrit derived from Jewish Arabic dialects, e.g. *šlóx* ‘a shabby person’ after the name of a Berber tribe in Morocco.

The use of Arabic components is found not only in the vocabulary but also in loan translation (calque) and external borrowing, e.g. *hu ʾaxál otá* < *ʾakalha w širib ʿalēha mayy ṣāfi* ‘he got done’, *ʾal ha-bóker* < *ʿala ṣṣubiḥ* 'first thing to do early in the morning', *hu met aléha*< *bimūt ʿalēha* ‘he is dying for her’. However, Blau (1974) notes that it is often quite difficult to locate the real origin of foreign borrowings into Ivrit. He observes that borrowed phrases like *hitsíl et ha-matsáv* (Literary Arabic *ʾanqaḏa l-mawqif*) exist in several languages, e.g. German *die Situation retten*, French *sauver la situation*, and English *save the situation*.

BIBLIOGRAPHICAL REFERENCES

Academy of Hebrew Language. 1994. *Laméd lešonxá* [Teach your language]. Jerusalem: Academy of Hebrew Language.

Avishur, Yitshak. 2003. “ha-Markív ha-ʾaravi ba-lašón ha-ʾivrít bat zmanénu u-va-sifrut” [The Arabic component in contemporary Hebrew language and literature]. *ha-ʾIvrít va-ʾAḥyotéha* (Haifa) 2–3.9–50.

Avneyon, Eitan. 2000. *Milá be-milá* [Word for word]. Holon: Eitav.

Barthélemy, A. 1935. *Dictionnaire arabe-français: Dialectes de Syrie: Alep, Damas, Liban, Jérusalem*. Paris: Librairie Orientaliste Paul Geuthner.

Ben-Amotz, Dan and Netiva Ben-Yehuda. 1972. *Milón ʾolamí le-Ivrít medubéret* [A universal dictionary of spoken Hebrew]. Jerusalem: Lewin-Epstein.

Ben-Yehuda, Eliezer. 1948. *Milón ha-lašón ha-ivrít ha-yšaná ve-ha-ḥadašá*. Jerusalem: Ben-Yehuda Hozaa La’Or, Le zecher Eliezer Ben-Yehuda.

Berman, Ruth A. and Dorit D. Ravid. 2000. “Acquisition of Israeli Hebrew and Palestinian Arabic: A review of current research”. *Hebrew Studies* 41.83–98.

Blanc, Haim. 1955. “la-Yesód ha-ʾaravi še-ba-dibúr ha-yisre’elí” [The Arabic background of Israeli speech]. *Leshonenu la-ʾAm* 6, commentarius 1.6–14; commentarius 3–4.27–32.

——. 1964. *Communal dialects in Baghdad*. Cambridge, Mass.: Distributed for the Center for Middle Eastern Studies of Harvard University Press.

Blau, Joshua. 1974. “Txiyát ha-ivrít ve-txiyát ha-ʾaravít ha-sifrutít” [The revival of Ivrit and the revival of Literary Arabic]. Jerusalem: Academy of Hebrew Language.

——. 1991. *The emergence and linguistic background of Judeo-Arabic*. 2nd ed. Jerusalem: Ben-Zvi Institute.

Burla, Yehuda. 1944. *Kitvéy Yehudá Búrla* [Yehuda Burla’s writings]. Tel Aviv: Massada.

Even-Shushan, Avraham. 1985. *ha-Milón he-xadáš*. 8 vols. Jerusalem: Kiryat Sefer.

Fischer, Wolfdietrich and Otto Jastrow. 1980. *Handbuch der arabischen Dialekte*. Wiesbaden: O. Harrassowitz.

Fraenkel, Siegmund. 1982. *Die aramäischen Fremdwörter im Arabischen*. Hildesheim: G. Olms (Repr. of 1886 ed.)

Geva-Kleinberger, Aharon. 2004. *Die arabischen Stadtdialekte von Haifa in der ersten Hälfte des zwanzigsten Jahrhunderts*. Wiesbaden: O. Harrassowitz.

Hebrew Language Council. 1946. *Tsimxéy Eretz Israel* [The plants of the land of Israel]. Jerusalem: Hebrew Language Council.

Landau, Ya’akov M. 2002. Review of: Bernard Spolsky and Elana Shohamy, *ha-Safót be-Yisraʾél: Mediniyút, ʾidiológya ve-nóhag* [The languages in Israel: Policy, ideology and practice]. *Bit’ón ha-Morím le-Arvít ve-Islám* 27.194–203.

Mansour, Jacob. 1991. *The Jewish Baghdadi dialect*. Or Yehuda: ha-Makhon le-Ḥeker Yahadut Bavel.

Oren, Shalom. 2002. *Bimló muván ha-milá: ʾAl motsaʾán šel milím beynleʾumiyót* [In the full sense of the word: Words and their origins]. Jerusalem: Keter.

Piamenta, Moshe. 1961. *Hašpaʾat ha-Arvít al xidušéy ben Yehuda* [The influence of Arabic on the innovation of Ben Yehuda]. Jerusalem: Leshonenu La’am.

Shapira, Anita. 1975. “ʾÉven va-síd: Parašát šutafút Yehudít-Arvít” [Stone and whitewash: An affair of Jewish-Arabic partnership]. *Ma’asáf* 7.118–129.

Stahl, Abraham. 1995. *Milón dvir du-lešoní etimológi le-Aravít Meduběret u-le-Ivrít* [Bilingual etymological dictionary of spoken Israeli Arabic and Hebrew]. Tel Aviv: Zmora-Bitan.
——. 1999. *Motsá ha-milím: Mekoroteyhén ve-gilguleyhén šel ha-milím še-be-fínu* [The origin of words: Origins and transformations of the words we use]. Tel Aviv: Dvir.

Aharon Geva-Kleinberger
(University of Haifa)

Ivrit Loanwords

Historically, Hebrew was extensively influenced by Arabic, especially during the Middle Ages, with the result that many lexical items entered Hebrew from Arabic. Since the establishment of Israel, though, the Arabic spoken by Palestinians who became Israeli citizens has absorbed many items from Modern Hebrew (Ivrit).

1. Knowledge and use of Ivrit

Although Palestinian Arabs in Palestine had some contact with the Jews even before the establishment of Israel in 1948, extensive contacts developed in subsequent years. Between 1948 and 1966 the encounter was limited to young Palestinian Arab men coming in contact with Israeli Jews at work. The military government that controlled the Palestinian Arabs limited direct contact with the Israeli Jewish population not just in workplaces but also in other areas. Today, however, in Israel, there is intensive contact with Jews almost daily in all areas of life: at work, in institutions of higher education, in government offices, in health institutions, and even, though minimally, in social relations. Without taking into account age, gender, or education, most Palestinians in Israel know and use Ivrit. Modern Israeli Hebrew has become an important language in the linguistic repertoire of the Palestinian Arabs in Israel.

While the purpose of teaching Ivrit to Jewish immigrants is to replace their original languages with Ivrit, the purpose of teaching Ivrit to Palestinian Arabs is additive, teaching Ivrit as an additional language.

For most Palestinian Arabs, Ivrit is the most important second language, even more so than English, and at times and in some domains it is even more important than Arabic (Shohamy and Donitsa-Schmidt 1998; Amara and Mar'i 2002). Not knowing Ivrit handicaps a Palestinian Arab in Israel, especially in contacts with government offices, in employment, and in higher education.

Ivrit is learned formally and informally. Since Ivrit is one of Israel's two official languages, the Palestinian Arabs in Israel learn it as the language of the state (Winter 1981; Hallel and Spolsky 1993; → Israel). Indeed, Ivrit is studied in Arabic-medium schools from the third grade and, in many schools, even from the second grade, but the influence of informal studies and external contacts is much greater (Reves 1983), thanks to continuous contact with Israeli Jews. All age groups maintain contact with Israeli Jews at different levels and degrees. Therefore, the use of Ivrit words, phrases, and even expressions is common among Palestinian Arabs in Israel. The extent of this use reflects the level of familiarity of the Palestinians in Israel with Israeli Jewish culture (Amara 1986, 1995; Amara and Spolsky 1986).

The degree of proficiency in speaking is the result of the extent and efficacy of the connection with the Jewish society. For example, Palestinian Arabs in the cities (such as Haifa, Jaffa, Ramla, and Lydda) and also Druze Arabs and the Bedouin who serve in the army use more Ivrit lexical items in their Arabic than other Palestinian Arabs (Abu Rabi'a 1996:7).

Ivrit for the Palestinian Arabs in Israel is getting stronger both at the qualitative level – higher levels of knowledge of the various varieties of Ivrit – and the quantitative level – the number of speakers and the frequency of speaking it.

2. Ivrit loanwords

Palestinian Arabs in Israel come in daily contact with the Hebrew language and Israeli Jewish culture not only in speech but also with the written varieties of the language, such as in literature, science, and journalism. Consequently, spoken Arabic in Israel constantly absorbs words from all varieties of Hebrew in different domains of life (Mar'i 1998). Among the Palestinian Arab population there is an increasing trend to borrow words from Ivrit and integrate them into their native language, Arabic. This phenomenon is not specific to a particular level in the Palestinian Arab society but is spread throughout all levels and ages. It is especially salient among academics. Integrating Ivrit words into Arabic is not considered alien but constitutes an integral part of the spoken language.

Based on various studies carried out by Amara (1986, 1991, 1995, 1999a, 1999b), Ben-Rafael (1994), Badeir (1990), and Koplewitz (1990) on Ivrit lexical integration, the following conclusions can be drawn.

i. Ivrit is now the main source of innovation, not just for Ivrit words but also for original English words.

ii. The level of diffusion and integration of Ivrit items differs according to domain. In the traditional domains (such as kinship and animal names), there is little influence from Ivrit. In modern domains (such as transport and electricity), there is extensive borrowing from Ivrit. Mixed domains (such as construction and health services), which existed before contact with Ivrit but in which change has occurred as a result of this contact, show appropriate intermediate levels of integration of Ivrit items.

iii. The levels of diffusion and integration of Ivrit items differ according to the social characteristics of the individual: education and the opportunity for outside contact are the main contexts for borrowing. The brokers of Ivrit diffusion are those who work outside their places of residence and their children, those who go to school, and especially those who pursue higher education in Israeli Jewish institutions.

Although Palestinians in Israel code-switch frequently from Arabic to Ivrit in many topics (e.g. shopping, construction, electricity, health care, transportation, schooling), the items reported here represent cases of borrowing. Borrowing from Ivrit takes place in all speech categories. Most of the studies show (although there are no comprehensive studies showing the frequency of the various types used) that nouns are the most frequent words borrowed. Items are borrowed from Ivrit into Arabic in several ways:

i. Many words are borrowed from Ivrit into Arabic without any change, preserving their Ivrit phonological form and meaning. Examples: *deši* ‘grass’, *širutīm* ‘toilets’, *ʾeisik* ‘business’, *mʾalit* ‘elevator’, *bagrūt* ‘matriculation examination’, *šalat* ‘remote control’, *otomāt* ‘automatic’ (originally borrowed into Ivrit from English – notice that the Ivrit form is borrowed into Arabic and not the English one; in the West Bank and Gaza Strip, the form *otomātīk* is used, revealing a direct borrowing from English).

ii. Many borrowed words have been adapted to Arabic phonology and morphology but preserve the original Ivrit meaning. Examples: *yiṭabbal* < *yitapel* ‘to take care’, *yišaxbel* < *yišaxpel* ‘to duplicate’, *yimraḥ* < *yimrāx* ‘to smear’. Note that the borrowed words are verbs.

iii. There are borrowed words from Ivrit that may be used with both Ivrit and Arabic suffixes. The use of one suffix or another depends on the social characteristic of the speaker and context. Highly educated people, for instance, tend to use the Ivrit suffix in most contexts and words. Examples: *tlūš/ʾitlūš*, pl. *ʾitlūš-āt* or *tlūš-īm* (with Ivrit suffix) ‘coupon’; *maḥsom*, pl. *maḥsom-āt*, *maḥāsīm*, or *maḥsom-īm* (with Ivrit suffix) ‘roadblock’.

iv. Not only single words are borrowed but also phrases. Examples: *taxanāt delek* ‘gasoline station’, *bank diskont* ‘discount bank’, *ʾīr taxtīt* ‘downtown’.

v. Many loan translations are borrowed from Ivrit into both the spoken and the written varieties of Arabic. Examples are found in Table 1.

vi. There are words that did not exist in Arabic and were borrowed from Ivrit as loan translations; these are mainly academic or scientific terms. Ivrit originally borrowed many of these words from English, and Arabic in

Table 1. Loan translations

Ivrit	Arabic loan	Standard Arabic	Translation
ʾavar et ha-bxina	*abar il-imtiḥān*	*najaḥa fī l-imtiḥān*	‘he passed the test’
sagar et ha-ʾiskāh	*sakkar ʾiskāh*	*ʿamila ʾiskāh*	‘he has made a deal’
giyūs ʾiksafīm	*taġnīd ʾamwāl*	*istiqṭāb ʾamwāl*	‘raising money’

Israel borrowed them from Ivrit. Presumably, in other Arab dialects these words were originally borrowed from either English or French. Examples are found in Table 2.

vii. It is important to note that many of the words borrowed from Ivrit and preserving Ivrit forms are also used in the written varieties of Arabic in Israel, mainly in newspapers. Frequently used words: *bagrūt* ‘matriculation exam’, *ramzor* ‘traffic light’, *histadrūt* ‘federal organization of workers’, *kupāt xolīm* ‘clinic’.

viii. Sometimes words existing in Arabic receive an expanded meaning borrowed from Ivrit. Examples: *naqqāl* < *nayad* ‘cell phone’, *imtiḥān al-ʾataba* < *baxīnat sāf* ‘placement test’.

ix. Some items are borrowed that are not familiar in the Arab world. Examples are found in Table 3.

3. Conclusion

Ivrit is now the main source of innovation, not just for Ivrit words but also for words that originally derive from English. The various linguistic aspects borrowed from Ivrit into Arabic in Israel reveal that borrowing is motivated not only by need but also by taboo and by desire for prestige. Several types of loanwords occur, i.e. loanwords preserving original Ivrit form and meaning in Arabic, loan translations, and borrowing of single words and also phrases and expressions.

However, in spite of the increasing use of Ivrit features in Palestinian Arabic in Israel, the existing tension between the Israeli and Palestinian identities among the Palestinian Arabs (because of the Israeli-Palestinian conflict, the definition and perception of Israel as a Jewish state, and the physical separation of Palestinians and Jews in patterns of settlement) has limited the convergence to the Ivrit language and the language of the dominant Jewish culture in Israel. The Palestinian Arabs in Israel have opted for the strategy of linguistic integration rather than linguistic assimilation. On the one hand, they attempt to acquire high sociolinguistic competence in Ivrit in order to be connected to and easily function in the wider social network mainly shaped by the majority culture, but on the other hand they preserve their Palestinian Arab identity by maintaining their Arabic mother tongue (Amara 2002). The nature of loanwords and their types from Ivrit in Palestinian Arabic in Israel give support to the conclusion that most of the borrowed items are lexical items, mainly nouns. The borrowing of Ivrit morphemes is minimal, and so far no influence on Arabic syntax has been documented.

Table 2. Loan translations as neologisms

Ivrit	Arabic loan translation	Gloss
tahlīx	*sayrūra*	‘process’
mašov	*imtiḥān al-mardūd*	‘feedback test’
mipūy	*masḥ*	‘surveying’

Table 3. Special expressions in Arabic in Israel

Ivrit	Arabic in Israel	Arabic in the Arab world	
mifakēx	*mufattiš*	*muwajjih fannī*	‘inspector’
menahēl	*mudīr*	*nāḍir*	‘headmaster’
bēt-sefer jisodi	*madrasa ibtidāʾiyya*	*madrasa ʾasāsiyya*	‘elementary school’
bēt-sefer-xativat benayim	*madrasa ʾiʿdādiyya*	*madrasa mutawassiṭa*	‘high school’
lōx	*lawḥ*	*sabbūra*	‘blackboard’
misrād ha-xinūx vi-ha-tarbūt	*wizārat at-tarbiya wa-t-taʿlīm*	*wizārat at-taʿlīm wa-t-tarbiya*	‘Ministry of Education’

BIBLIOGRAPHICAL REFERENCES

Abu-Rabi'a, Salim. 1996. "Learning Hebrew in Israel: The relationship of attitudes, cultural background, and interest of material to reading comprehension in a second language". *Journal of Multilingual and Multicultural Development* 17:5.1–12.

Amara, Muhammad Hasan. 1986. *The integration of Hebrew and English lexical items into the Arabic spoken in an Arab village in Israel*. M.A. thesis, Bar-Ilan University, Ramat-Gan.

——. 1991. *Sociolinguistic aspects of variation and change: A study of the languages spoken in an Arab village in Israel*. Ph.D. diss., Bar-Ilan University, Ramat-Gan.

——. 1995. "Hebrew and English lexical reflections of socio-political changes in Palestinian Arabic". *Journal of Multilingual and Multicultural Development* 15:6.165–172.

——. 1999a. "Hebrew and English borrowings in Palestinian Arabic in Israel: A sociolinguistic study in lexical integration and diffusion". *Language and society in the Middle East and North Africa*, ed. Yasir Suleiman, 81–103. London: Curzon Press.

——. 1999b. *Politics and sociolinguistic reflexes: Palestinian border villages*. Amsterdam and Philadelphia: J. Benjamins.

——. 2002. "Hebrew among the Arabs in Israel: Sociolinguistic aspects". *Teuda* 18.85–105. (In Hebrew.)

—— and Abd Al-Rahman Mar'i. 2002. *Language education policy: The Arab minority in Israel*. Dordrecht: Kluwer Academic Publishers.

Amara, Muhammad Hasan and Bernard Spolsky. 1986. "The diffusion and integration of Hebrew and English lexical items in spoken Arabic in an Israeli village". *Anthropological Linguistics* 28:1.43–58.

Badeir, Salih. 1990. *Positions and exposure to the target language and the connection between these and success in learning a second language*. M.A. thesis, Tel Aviv University. (In Hebrew.)

Ben-Rafael, Eliezer. 1994. *Language, identity and social division: The case of Israel*. Oxford: Studies in Language Contact.

Hallel, Michael and Bernard Spolsky. 1993. "The teaching of additional languages in Israel". *Annual Review of Applied Linguistics* 13.37–49.

Koplewitz, Immanuel. 1990. "The use and integration of Hebrew lexemes in Israel spoken Arabic". *Multilingual Matters* 71.181–195.

Mar'i, ʿAbd Al-Rahman. 1998. "The status of Arabic and Hebrew in the past and the present". *ar-Risāla* 7.257–276. (In Arabic.)

Reves, Thea. 1983. *What makes a good language learner?: Personal characteristics contributing to successful language acquisition*. Ph.D. diss., Hebrew University of Jerusalem.

Shohamy, Elana and Smadar Donitsa-Schmidt. 1998. *Jews vs. Arabs: Language attitudes and stereotypes*. Tel Aviv: Tami Steinmetz Center for Peace Research, Tel Aviv University.

Winter, Michael. 1981. "Basic problems in the educational system". *The Arabs in Israel: Continuity and change*, ed. Aharon Lish, 168–179. Jerusalem: Magnes Press. (In Hebrew.)

MUHAMMAD HASAN AMARA (Bar-Ilan University)

J

Jargon

1. Definition and name

'Jargon' may be broadly defined as "the modifications that a socioprofessional group brings to the national language (especially in vocabulary and pronunciation)" (Ducrot and Todorov 1979:59). These modifications appear to arise from the particular or specific nature of the topic, the need for group members not to be understood by others, or the wish to identify the group as somehow different (Ducrot and Todorov 1979:59). Jargon thus includes → technical terminology and specialist vocabulary, such as the terms defined in this encyclopedia. It also includes words and phrases that stand in for or replace commonly used words and phrases that already exist.

The literature on Arabic jargon published in Western languages uses a number of terms to name its subject matter. These terms include 'jargon', 'argot', 'cant', and → 'secret language', and their equivalents. 'Argot' and 'cant' usually refer to the jargon of criminals and other groups whose professional activities depend on secrecy. 'Secret language' is a more general term. It includes jargons as well as 'play language'. Play language is one of a number of terms that describe language disguised by regular phonological change, like Pig Latin in English or *javanais* in French. Play language in Arabic provides useful data to test assumptions about Arabic phonology (see Heath 1989:185–199). A particular jargon may contain lexical items that appear to be produced by phonological change similar to that of play language. The lexicon of most jargons, however, derives from other processes as well. Play languages rely only on phonological change for word disguise. Further, a play language tends to be shared by a social group rather than a professional or special interest group (for an example of play language in Saudi Arabia, see Bākallā 2000, 2002).

There is no generally accepted equivalent for 'jargon' in Arabic, although several equivalents occur in bilingual dictionaries. These include *raṭāna* 'jargon; lingo, gibberish' (Baʿalbakī 1996:489). This alludes to the incomprehensibility of jargon to outsiders. Also found is *luġa xāṣṣa* 'special language, specialist language' (Bākallā 1983:45) and *lahja ḥirfiyya* 'professional language' (Alkhuli 1982:142). These are technical terms, the jargon of linguists and linguistics.

Nontechnical names for jargon in Arabic are numerous. One is *luġa sirriyya* 'secret language'. The term *sīm* or *sīn* is attested from the 19th century (Rowson 1983:11). Spoken Moroccan Arabic uses *ġūṣ* or *ġawṣ*. These terms refer to language used within a specific professional or social group, as well as to play speech. A specific jargon or the jargon of a particular profession or group is often named by the term *luġa* 'variety; dialect', sometimes followed by the name of the group, as in *luġat an-naššālīn* 'the language of pickpockets'.

2. Jargon in medieval Arabic

Research on jargon in Arabic has focused less on conventional or mainstream professions than on groups that are marginal within society.

This tradition has a long history, examined by Bosworth (1976) in his study of the Banū Sāsān. 'Banū Sāsān' is a general term for the group that includes beggars, tricksters, swindlers, etc. The jargon of the Banū Sāsān (*luġa Banū Sāsān* or *manāġāt Banū Sāsān*) is preserved in two jargon poems, each titled *al-qaṣīda as-sāsāniyya* 'the Sasanian poem'. One was written by ʾAbū Dulaf al-Xazrajī (fl. 4th/10th century) and the other by Ṣafī d-Dīn al-Ḥillī (d. ca. 750/1349). Other sources, among them the works of al-Jāḥiẓ (160/776–255/868), the *maqāmāt* of Badīʿ az-Zamān al-Hamaḏānī (d. 398/1008) and al-Ḥarīrī (d. 516/1122), the shadow plays of Ibn Dāniyāl (d. 710/1310), and *Kašf al-ʾasrār* by al-Jawbarī (fl. 7th/13th century), use jargon to describe the tricks of the Banū Sāsān.

The jargon of the Banū Sāsān shares certain features with jargon in modern Arabic, particularly processes of word formation. Bosworth describes as phonological change the regular morphological derivation that results in the verb *faššaša* 'to fart in the mosque, thus annoying the worshippers into giving money to make the beggar go away', from *fašša* 'to belch, fart softly and gently' (1976:159). Semantic change or shift narrows the sense of the verb *fakkaka* 'to escape from bonds or chains as a feat of skill' from its more general meaning of 'to separate two things fastened together' (1976:159). Finally, jargon may borrow words from other languages, as the word *kurs* 'hunger' from the Persian *gurs* (1976:160).

Although the Banū Sāsān no longer exist, many of their tricks and some of their jargon live on. Rowson (1983:17–18) compares the jargon terms he collected in Cairo with earlier sources. These include the jargon of the shadow puppet theater (*xayāl aḏ-ḏill*) from Kahle's (1926) list and the shadow plays of Ibn Dāniyāl, as well as the jargon poems of ʾAbū Dulaf al-Xazrajī and Ṣafī d-Dīn al-Ḥillī. In spite of the time and in some cases geographical distance that separate the historical sources from contemporary Cairene jargon, Rowson finds correspondences. He cites, for example, the term *bišbāša*, which means 'moustache' in his data and 'beard' in earlier sources (Rowson 1983:18).

3. JARGON IN MODERN ARABIC

Treatments of jargon in modern Arabic are not numerous. This may be due to the difficulties of gathering data from groups of users with a vested interest in maintaining the secrecy of their jargons. Jargons in modern Arabic can be divided into the following categories: jargons of Egypt, jargons of North Africa, jargons of Arabic-speaking Jews, and jargons of Islamic scholars.

3.1 *Jargon in Egypt*

Research on jargon in Egypt has to date focused on social as well as linguistic variety. Based on fieldwork conducted mainly in Alexandria, ʿĪsā (1988) identifies nine distinct jargons: *luġat an-nassālīn* 'the languge of pickpockets', *luġat al-ḥarāmiyya* 'the language of thieves', *luġat an-naššābīn* 'the language of swindlers', *luġat al-muxaddirāt* 'the language of drugs', *luġat al-mutasawwilīn* 'the language of beggars', *luġat aṣ-ṣāġa* 'the language of goldsmiths', *luġat al-munajjidīn* 'the language of craftsmen who renew and restuff upholstery', and *luġat al-ʿawālim* 'the language of female entertainers'. It may seem that the *munajjidīn* are keeping strange company among groups that hide their activities, for one reason or another. As ʿĪsā points out, however, the work of the *munajjidīn* takes place in their clients' homes. Their jargon lets them communicate privately even when their clients are nearby (1988:99).

Divisions between and within groups of jargon users are to a certain extent linguistic as well as vocational. The jargon of the *munajjidīn* borrows some terms from Turkish (ʿĪsā 1988:99); the jargon of *ḥarāmiyyat al-baḥr* 'thieves who steal from harbors and airports' takes some lexical items from Greek (1988:22); and that of goldsmiths has many items from Hebrew (1988:92).

For all of the variety seen in the jargons of Alexandria, they share some features. A number of these jargons share what appears to be a reflexive, *xaḏa* or *xaḏat*, as in *juwar baʿīd ʿan xaḏatahu* 'get away from him!' (ʿĪsā 1988:49). The → gypsies (*al-ġajar*), whose possible Romani borrowings ʿĪsā does not discuss in detail, use instead *būs*, as in *bi-tšawri ʿala būsna* 'she is spying on us' (1988:22). The term

ʿarwa 'money' occurs in a number of jargons. Other shared items include *kūdiyāna* 'woman', *suhni* 'good', and *ʿasaf* 'to take, steal; to arrest' (perhaps from Classical Arabic *ʿasafa* 'to take by force').

Rowson (1983) describes the jargons of Cairo as equally complex. He, however, sees linguistic as well as vocational and social divisions between groups of jargon users. The two primary jargons are *sīm iṣ-ṣāġa* 'the jargon of goldsmiths' and *sīm il-fannanīn* 'the artists' jargon'. They are distinguished by their sources. The jargon of goldsmiths in Cairo, as in Alexandria, borrows from Hebrew, while Romani is the source of some lexical items in the artists' jargon. The artists' jargon also shows links to the medieval *luġa Banū Sassān*, as noted above. These two jargons, *sīm iṣ-ṣāġa* and *sīm il-fannanīn*, have given rise to others. One, the *sīm il-xartiya* 'the jargon of those who prey on tourists' derives from *sīm aṣ-ṣāġa*, augmented by borrowings from *sīm il-fannanīn* and *sīm il-ḥarāmiya* 'jargon of thieves'. Van Nieuwkerk (1995:102) and Rowson (1983) both see *sīm il-kawanīn* 'homosexuals' jargon' as the most productive and widely used of Cairo jargons. According to Rowson, the vitality of *sīm il-kawanīn* results from its role in building and maintaining group solidarity, as well as its importance in protecting the secrecy of what is being said (1983:20).

The jargons of various criminal groups, which make up a large part of ʿĪsā's material from Alexandria, hardly figure in Rowson's study. This is due to differences in data collection. Rowson relied mainly on social contacts (1983:14–15), where ʿĪsā conducted his research in a prison (1988:5). Rowson notes, however, that the kind and amount of data collected from criminal groups was not promising (1983:20). He says the same of the jargon of drugs in Cairo. Hashish was at that time so widespread that its terminology was hardly secret, although other drugs were relatively uncommon (Rowson 1983:22).

Two jargons, those of *il-fannanīn* (Rowson 1983) and *al-ʿawālim* (ʿĪsā 1988), allow us to compare the jargons used in Cairo and Alexandria. A number of lexical items are shared. They include *šalaf* 'bad', *kanwin* (ʿĪsā 1988:111) and *ikkanwin* (Rowson 1983:19) 'to participate in homosexual acts'. There are also borrowings from Italian that appear to date to 19th-century tours by commedia dell-l'arte troupes, such as *furti* 'hurry up!' (Rowson 1983:18). Van Nieuwkerk confirms that the jargon of *il-fannanīn* in Cairo and Alexandria share much vocabulary (1995:96).

3.2 *Jargon in North Africa*

Published research on jargon in North Africa to date surveys processes of word formation rather than groups of jargon users. Brief articles by Roux (1936), Pianel (1950), and Youssi (1977) consider Moroccan jargons in Berber as well as Arabic. Most of these are play languages, relying on phonological change. Pianel (1950) identifies artisans in Rabat and Marrakech, as well as women and children in Rabat, as users of play language. Youssi (1977) also cites a number of play languages as used by professional and social groups in Marrakech, Casablanca, and Oujda. In addition to play language, Roux (1936) gives examples of what he calls "argot métaphorique". These are words and phrases used as code in certain professions, such as the phrase *lmansəj xawi* 'the loom is empty [or unwarped]', used by weavers to describe a woman too thin for beauty (Roux 1936:1074). He also identifies two jargons, the Berber *taʿjmiyt*, used by poet-bards, and the Arabic *ġūṣ* or *ġawṣ*, used by singers and other professional groups. Each jargon makes use of a variety of processes to form words. Where the Berber *taʿjmiyt* borrows from Arabic, however, the Arabic *ġūṣ* or *ġawṣ* borrows from *llašuniya*, the Hebrew-derived jargon of the Moroccan Jewish community, discussed in Section 3.3, below. Roux points out that the Arabic *ġūṣ* or *ġawṣ* is identifiably Arabic, even when its lexicon makes it incomprehensible to outsiders, as in *f-sratu lfərza b-əzzaṛṛaṛ* 'he has a lot of money' (1936:1086).

In a longer article, Bencheneb (1942) examines processes of word formation in the jargon of criminals and prostitutes in Algiers. He catalogs a variety of processes for word formation: phonological, morphological, and semantic. In addition, words are borrowed from Classical Arabic, from the spoken Arabic of the surrounding countryside, and from Berber, Turkish, Persian, French, Spanish, Italian, and Judaeo-Arabic.

3.3 *Jargons of Arabic-speaking Jews*

The jargons of Arabic-speaking Jews vary as much as do the spoken Arabic varieties used by the larger populations within which Jews lived. They are all, however, characterized by the use of words and phrases derived from Hebrew. The origins of some of these jargons are recognized in their names. In Damascus, for example, the jargon used by Jewish goldsmiths is *lkalām ilʿəbrānī* 'the Hebrew language' (Barbot 1974:83). Libyan Jewish peddlers called their jargon *lašōn haqodēš* 'the holy tongue' or *lašōn haqodēš aṭ-ṭawāfa* 'the peddlars' holy tongue'; Libyan Muslims called it *ʿabrānī* or *ʿabriya* 'Hebrew' (Goldberg 1983:90). Chetrit (1994) records a number of names used by Jews to name their jargons in the cities and regions of Morocco. Muslims throughout Morocco, however, referred to these jargons as *llašuniya*, from the Hebrew word *lašon* 'language' (Chetrit 1994:523).

Published research on the jargons of Arabic-speaking Jews are mainly brief sketches. Khan (1995–1997) provides a short list of words and phrases collected from Karaite Jews from Cairo living in Israel (cf. Rosenbaum 2003). Goldberg (1983:90–91) does much the same for Libyan Jews. Roux's survey of Moroccan jargons touches on *lašuniya* in Rabat (1936:1086–1087). Chetrit (1994) examines two brief texts from Morocco. Barbot (1974:73–76) gives a list of specialist technical terms used by goldsmiths in Damascus. It begins with *ḥamḍ əlkibrīt* 'sulfuric acid' and ends with *ʾəṣbaʿ* 'ring-sizing rod [lit. finger]'. He follows up with a list of jargon terms (Barbot 1974:78–81) that includes workplace terms such as *madhobōt* or *mazhobōt* 'gold' and *bāṛ* 'alloy; imitation'. The list also includes words and phrases that conceal the speaker's meaning, such as *taštrīt* 'stupid' and *məġloyōt* 'expensive'.

3.4 *Jargons of Islamic scholars*

Jargons have been recorded as used by Qurʾānic students in Morocco (*ṭolba*) and in Nigeria. Their users aside, these jargons have little in common.

Perhaps the simplest of these jargons, in terms of use if not comprehension, is described by Pianel (1950:166–167). It is similar to the techniques used in play language. One version inserts the words of the intended utterance into the words of a given verse of the *Qurʾān*. Pianel cites the example of the question *škun huwa l-manjūr?* 'who is al-Manjūr?', inserted into Q. 2/7 *xatama llāhu ʿalā qulūbi-him* 'Allah hath sealed their hearts'. The result is *xatama llāhu škun ʿalā huwwa qulūbi-him al-manjūr?*. A variation of this technique inserts the root consonants of the words of the intended utterance into the words of a Qurʾānic verse.

Another jargon, briefly mentioned by Roux (1936:1074–1075), replaces frequently occurring words and phrases with proper names. The slave's name *Mbarek* means 'silver', while *xāli bə-lʿəššāb* 'my uncle Bə-lʿəššāb' means 'be quiet in front of this stranger!'. A third *ṭolba* jargon of Morocco is described by Berjaoui (1994). It spells out the intended word, but replaces the letters by the numeric values assigned to them by the *ʾabjadiyya*. In this jargon, the word *lxarəj* 'abroad' becomes *səttmya* [pause] *myatayn* [pause] *tlata* 'six hundred [pause] two hundred [pause] three', with deletion of the definite article.

The Nigerian *waris*, the jargon of students in the Qurʾānic schools of Borno State, is described in some detail by Owens and Hassan (2000). Of the processes this jargon uses to derive words and phrases, possibly the most unusual is that of association with Qurʾānic citations. This process links a word of Nigerian Arabic with a similar word contained in a phrase from the *Qurʾān*. A word such as *hūt* 'fish' is associated with the phrase *ka-ṣāḥibi l-ḥūt* 'like him of the fish' from Q. 68/48. Both words for 'fish', *hūt* and *l-ḥūt*, are displaced, leaving *ka-ṣāḥibi*. As a result, one could say *ʾakalna ka-ṣāḥibi aloom* 'we ate fish today' (Owens and Hassan 2000:228–229). A similar process of association and displacement is recorded by Roux (1936:1075) among *ṭolba* in Berber-speaking regions of Morocco.

It is frequently suggested that one of the primary purposes of jargon is secrecy. It allows group members to conceal their intended meaning from outsiders. If this is the case, *waris* is highly effective. When tested for their understanding of *waris* words and phrases, students as well as teachers from Qurʾānic schools had a success rate of 92 percent. This compares with a 22 percent rate for members of the larger community (Owens and Hassan 2000:250). Owens and Hassan suggest that *waris* may also serve another function for its users. It may aid learning: "[It] brings koranic language into contact with the everyday experience of the koranic students themselves" (2000:251).

Bibliographical references

Alkhuli, Muhammad Ali. 1982. *A dictionary of theoretical linguistics: English-Arabic with an* Arabic-*English glossary*. Beirut: Librairie du Liban.

Baʿalbakī, Munir. 1996. *Al-Mawrid: A modern English-Arabic dictionary*. Beirut: Dār al-ʿIlm li-l-Malāyīn.

Bākallā, Muḥammad Ḥasan. 1983. *Muʿjam muṣṭalaḥāt ʿilm al-luġa al-ḥadīṯ: ʿArabī-ʾinkilīzī wa-ʾinkilīzī-ʿarabī*. Beirut: Maktaba Lubnān.

——. 2000. "What is a secret language: A case from a Saudi Arabian dialect". *Perspectives on Arabic linguistics*, XIII–XIV, ed. Dilworth B. Parkinson and Elabbas Benmamoun, 171–183. Amsterdam and Philadelphia: J. Benjamins.

——. 2002. "*Al-Misfalawiyah*: A subdialect of the Makkan dialect of Arabic". *Aspects of the dialects of Arabic today: Proceedings of the 4th conference of the International Arabic Dialectology Association (AIDA) Marrakech, April 1–4, 2000: In honor of Professor David Cohen*, ed. Abderrahim Youssi a.o., 412–421. Rabat: AMAPATRIL.

Barbot, Michel. 1974. "Notes lexicographiques sur les orfèvres et bijoutiers de Damas (texte révisé)". *Arabica* 21:2.72–83.

Bencheneb, Rachid. 1942. "L'argot des arabes d'Alger". *Revue Africaine* 86.72–101.

Berjaoui, Nasser. 1994. "Preliminary notes on a secret language of *ṭəlbā* in Morocco". Caubet and Vanhove (1994:511–517).

Bosworth, Clifford Edmund. 1976. *The medieval Islamic underworld: The Banū Sāsān in Arabic society and literature*. I. *The Banū Sāsān in Arabic life and lore*. Leiden: E.J. Brill.

Caubet, Dominique and Martine Vanhove (eds.). 1994. *Actes des Premières Journées Internationales de Dialectologie Arabe de Paris*. Paris: Publications Langues'O.

Chetrit, Joseph. 1994. "Formes et structures du mixage linguistique dans les langues secrètes juives du Maroc". Caubet and Vanhove (1994:519–530).

Ducrot, Oswald and Tzvetan Todorov. 1979. *Encyclopedic dictionary of the sciences of language*. Trans. Catherine Porter. Baltimore and London: Johns Hopkins University Press.

Goldberg, Harvey E. 1983. "Language and culture of the Jews of Tripolitania: A preliminary view". *Mediterranean Language Review* 1.85–102.

Heath, Jeffrey. 1989. *From code-switching to borrowing: Foreign and diglossic mixing in Moroccan Arabic*. London and New York: Kegan Paul International.

ʿĪsā, ʿAlī. 1988. *al-Luġāt as-sirriyya*. [Alexandria], n. pub.

Khan, Geoffrey. 1995–1997. "A note on the trade argot of the Karaite goldsmiths of Cairo". *Mediterranean Language Review* 9.74–76.

Nieuwkerk, Karin van. 1995. *'A trade like any other': Female singers and dancers in Egypt*. Austin: University of Texas Press.

Owens, Jonathan and Jidda Hassan. 2000. "Making a fish of a friend: Waris, the secret language of Arabic Koran school students in Borno". *Arabic as a minority language*, ed. Jonathan Owens, 1–33. Berlin: Mouton de Gruyter.

Pianel, Georges. 1950. "Notes sur quelques argots arabes du Maroc". *Hespéris* 37.460–467.

Rosenbaum, Gabriel M. 2003. "Another Egyptian dialect: Spoken Jewish Arabic in Egypt in the twentieth century". *Association Internationale de Dialectologie Arabe (AIDA) 5th Conference Proceedings, Cádiz, September 2002*, ed. Ignacio Ferrando and Juan José Sánchez Sandoval, 545–560. Cadiz : University of Cadiz.

Roux, Arsène. 1936. "Quelques argots arabes et berbères du Maroc". *Revue Africaine* 79:II/2. 1067–1088.

Rowson, Everett. 1983. "Cant and argot in Cairo colloquial Arabic". *Newsletter of the American Research Center in Egypt* 122.13–24.

Youssi, Abderrahim. 1977. "Les parlers secrets au Maroc". *La Linguistique* 13:1.135–143.

Elizabeth M. Bergman (Georgetown University)

Jarr → ʾIʿrāb

Javanese

1. Introduction

Javanese is a member of the Austronesian family of languages, which includes the languages of the Philippines, Malaysia, Indonesia, Madagascar, and the Pacific. It is spoken on the island of Java, but not all who live on Java speak the language. It is spoken only in Central and East Java, with the Central Javanese variety being considered the standard. The Javanese constitute the largest ethnic group in the Indonesian population, and it is estimated that 40 percent of the Indonesian population speaks Javanese. In addition, due to transmigration, Javanese is also spoken on other islands, e.g. South Sumatra. Beyond Indonesia, Javanese is spoken in Surinam and New Caledonia.

Borrowing from Arabic must have begun during the period of East-Javanese Muslim literature (16th–17th centuries), following on the spread of Islam, through which the use of Arabic proliferated. It is important to note here that the phrase 'Arabic language' refers to Classical Arabic. As to Javanese, only words of the *ngoko* level are used as examples, since it is considered the neutral level, which, unlike the other levels, does not contain interpersonal meaning.

2. Phonological adaptation

There are 20 consonant phonemes in the Javanese phonological inventory: the bilabial stops /p, b/, the nasals /n, m/, the palato-nasal /ny/, the velar nasal /ng/, the sibilant /s/, the liquids /l, r/, the semivowels /w, y/, the velars /k, gh/, the retroflex palato-alveolars /ṭ, ḍ/, the dento-alveolars /t, d/, and the palatals /c, j/. There are six distinct vowel phonemes in Javanese: /a, o, e, u, i/ and a shwa vowel.

The following examples illustrate the phonological adaptation in the loanwords involving Arabic sounds that are nonexistent in Javanese:

i. The Arabic uvular fricatives /ḥ/, /x/, and the uvular plosive /q/ are commonly pronounced as [k] in Javanese. Except for the word *qurʾān*, *q* is normally represented by *k* in Javanese; *q* can occur initially, medially, or finally. As for *x*, it is sometimes also reflected as *h*.

ḥ as *h*	*haram* < *ḥarām*	‘forbidden’
	rahmat < *raḥma*	‘God’s mercy’
	sah < *ṣaḥḥ*	‘legal’
ḥ as *k*	*kurmat* < *ḥurma*	‘respect’
	mokal < *muḥāl*	‘impossible’
x as *k*	*kotbah* < *xuṭba*	‘sermon’
	makluk < *maxlūq*	‘creature’
	naskah < *nusxa*	‘writing; copy’
q as *k*	*kamus* < *qāmūs*	‘dictionary’
	akal < *ʿaql*	‘intelligence’
	mutlak < *muṭlaq*	‘absolute’

ii. The velar fricative *ġ* is predominantly pronounced as *g*, except for *bihal* < *biġāl* (pl. of *baġl*) ‘mule’, where it is pronounced as *h*. Some examples: *galib* < *ġālib* ‘usual’, *magrib* < *maġrib* ‘sunset’, *mubalig* ‘preacher’ < *muballiġ* ‘messenger’.

iii. The Arabic glottal plosive ʾ and the pharyngeal fricative ʿ are realized in various ways in Javanese. The most common way is to represent the sound as a fronted vowel in Javanese, as in *adil* < *ʿādil* ‘just’. In other cases, however, the sound is represented by the velar nasal *ng* in Javanese, e.g. *Ngahad* < *ʾaḥad* ‘Sunday’, *Dulkangidah* < *ḏū l-qaʿda* ‘the eleventh month’.

iv. The interdental *ḏ*, the velarized *ḍ*, and *ḏ̣* are most commonly realized as *d* in Javanese, but also as *l* in some cases, e.g. *dikir* < *ḏikr* ‘invocation of God’, *adan* < *ʾaḏān* ‘call to prayer’, *kadi* < *qāḍī* ‘judge’, *taklim* < *taʿḏ̣īm* ‘great respect’, *luhur* < *ḏ̣uhr* ‘midday’. Realizing *ḍ*/*ḏ̣* as *l* becomes compulsory when they occur in final position, as such sounds would not normally occur finally in Javanese, e.g. *lapal* < *lafḏ̣* ‘spoken word’, *lila* ‘sincere’ < *riḍā* ‘approval’, *aral* < *ʿarḍ* ‘hindrance’.

v. The post-alveolar sibilant *š* and the velarized sibilant *ṣ* merge with *s* or retain the fricative sound as *sy* in Javanese. The fricative interdental *ṯ* is also pronounced *s* in the few loanwords found:

š as *s*	*sahid* < *šāhid*	‘religious martyr’
	mesgul < *mašġūl*	‘sad’
	Kures < *qurayš*	‘name of tribe in Arabia’
ṣ as *s*	*sabar* < *ṣabr*	‘patience’
	tasawup < *taṣawwuf*	‘Sufism’
	kusus < *xuṣūṣ*	‘special’
ṯ as *s*	*Selasa* < *ṯalāṯāʾ*	‘Tuesday’
	isbat < *ʾiṯbāt*	‘to assert with parables’ (more specific meaning than its original, which means ‘assertion’)
	kadas < *ḥadaṯ*	‘ritual impurity’
š as *sy*	*masya Allah* < *mā šāʾallāh*	‘Good Lord!’
	syarif < *šarīf*	‘descendant of Muḥammad’

vi. The labiodental fricative *f* is commonly realized as *p*, except in a few words such as *fajar* < *fajr* ‘dawn’, *wafat* ‘to pass away’ < *wafāt* ‘demise, death’, *maaf* ‘to pardon’ < *maʿa ʿafw* ‘excuse me!’. Normally, no entry under *f* is found in Javanese dictionaries.

The above patterns of assimilation are rather simplified, showing mainly how nonexistent sounds are adapted in the borrowed word. In fact, even sounds that are shared by the two inventories are slightly different in pronunciation, for example, while the Arabic /h/ is glottal, it is velar in the Javanese sound system.

Vowels are also adapted in the loanwords. There are commonly two kinds of assimilation: by realizing the Arabic long vowels as short, as in *hurup* ‘alphabet’ < *ḥurūf* ‘letters’, or by pronouncing them as weak sounds. Even vowels that exist in Javanese may be changed in pronunciation, particularly when the back vowel *u* occurs in the first syllable of the Arabic word, where it is often realized as *o*, e.g. *kotbah* < *xuṭba* 'sermon', *sokur* < *šukr* 'thanks', *donga* < *duʿāʾ* 'prayer'.

3. SEMANTIC ADAPTATION

The largest part of the loanwords were borrowed directly without change of meaning or with a very slight change of semantic nuances. In the examples quoted above, words that were borrowed directly include *haram* 'forbidden', *rahmat* 'God's mercy', *mutlak* 'absolute'. Direct borrowing rarely involves a change of meaning, but meaning tends to be more specialized in the loanword. However, changes of meaning can occur along the singular-plural dimension of nouns. The change along this dimension is very common in a language such as Javanese, where plurality is not syntactic but lexical, through word reduplication. Likewise, the distinction between masculine and feminine is rarely indicated in Javanese, except in obvious cases such as in *mukmin/mukminat* 'believer [masc./fem.]'.

Indirect borrowing occurs mainly through translations of Arabic texts. There are three subtypes: loan mismatches (sometimes called loan shifts), loan blends, and loan transfer. Loan transfer involves more than translation proper since the loanword seems to have entered the lexicon through reinterpretation of the loanword in the light of the Javanese belief system.

Loan mismatches occur when the loanword has acquired a new meaning on the basis of the Javanese meaning. Loan mismatches here include not just changes of meaning but also of form. Loanwords in this category are mainly concerned with philosophical concepts and religious practices. In the examples of Table 1, the original Arabic meaning and the Javanese meaning are presented, to show the difference.

These examples demonstrate the specialized Javanese meaning compared to the original. The Islamic credo *šahāda* appears twice, with a very different meaning in Javanese. In fact, the same credo is also used as direct borrowing *sahadat* and *klimah sahadat*, with the latter usually glossed as 'phrase or expression containing the two Arabic phrases basic to Islam: believing in Allah and believing in Muḥammad as prophet'. This complex borrowing implies different chronological layers. The most probable explanation would be that the two loans in Table 1 are older than the direct borrowing. The older loans must have made their way into Javanese during the early years of Islamic civilization in Java. Both *kalimasada* and *sekaten* have Hindu-Buddhist elements in them and contain an element of syncretism central to the Javanese belief system.

Another important point to make concerns the doublet *slamat* and *slamet*. While the former was borrowed directly, the latter was not. The meaning of the latter changes dramatically. It is usually combined with another Javanese word to create loan blends, e.g. *slametan brokohan* 'christening ceremony at childbirth', *slametan jenang abang* 'ceremony celebrating a circumcision or wedding'. Cultural reasons such as this seem to be the most probable cause of loan mismatches.

Loan blends are formed in two ways: by taking some elements from Arabic and some from Javanese, or by adapting the Arabic word morphologically using Javanese affixes to form a new meaning. The main purpose is to make the meaning more specific for particular contexts.

Loan transfers occur when new words are constructed from Javanese forms parallel to Arabic models. Loanwords in this category seem to be generally related to spirituality and mysticism. In the examples in Table 3, they are presented as etymological doublets, even triplets, as they appear as entries in Horne's Javanese-English dictionary, along with the gloss.

The examples in Table 3 show doublets or triplets that have entered the lexicon; one is the loanword with original Arabic meaning and one is a loan transfer. In each case, the loan transfer retains an element of meaning from the loanword, e.g., the semantic element 'blessing' that is found in the loanword is retained in the loan transfer. The meaning of the newly formed words tends to be more specialized than their loan counterparts.

The loan transfers presented here reflect spiritual and mystical practices in Java that may be regarded as inherently Javanese. In these cases, Javanese has transferred the original Arabic loanword and redefined it in the context of the Javanese belief system and worldview.

Table 1. Semantic changes in loanwords

Loanword	Arabic meaning	Javanese meaning
isim	*ʿiṣma* 'to safeguard, protect'	'a written Arabic phrase used for warding off danger/illness'
kalimasada	*kalimat aš-šahāda* 'the Muslim credo'	'magical book having the power of resurrection'
sekaten	*šahādatayn* [dual] 'the Muslim credo'	1. 'an important court festival held during Mulud'; 2. 'the gong music accompanying the festival'
slamet(an)	*salāma* 'well-being; safety'	Dramatic changes in meaning with the verb-forming suffix *-an*: 1. 'to keep someone safe, well, secure'; 2. 'to save someone's life'; 3. 'to hold ceremony on someone's behalf'

Table 2. Loan blends

Loanword	Arabic meaning	Javanese meaning
abdi dalem	'one who serves in the royal palace/court'	Arabic *ʿabd* 'servant' + Javanese *dalem* 'inside'
kejiman	'being (inadvertently) possessed (by the genie)'	Arabic *jinn* 'genie' + Javanese affix *ke-an* meaning '[done] inadvertently'
ilmu kedjawen	'mystical Javanism'	Arabic *ʿilm* 'science; knowledge' + Javanese *kedjawen* 'Javanese-ism'
ratu adil	'the messiah'	Javanese *ratu* 'queen' + Arabic *ʿadl* 'justice'

Table 3. Examples of loan transfers

Loanword	Gloss	Loan transfer	Gloss
berkah	'blessing'	*berkat*	'food, blessed by a religious official, taken home from a ritual ceremony by the guests after they have eaten a portion of it'
sariah	'canon law; lawfulness'	*serengat*	'fulfilment of religious duty according to Muslim law'
tarekah	'Dervish order; spiritual path [esp. mystical]	*tarekat*	'religious acts that bring one closer to God, according to Islamic principle [e.g. meditation, devotion to the ethical principles]'
		tirakat	'to deny oneself food and sleep as a sacrificial act in order to be granted one's desire'

4. SEMANTIC FIELDS OF LOANWORDS

Following Jones (1984) in his list of Arabic loanwords in Indonesian, the loanwords are categorized into ten categories, with one additional category, ‘Food and drink’. Two categories constitute the biggest part: more than 60 percent of the loanwords belong to ‘Islamic religion and Arabic culture’ and to ‘Abstract and philosophical terms’. The least in number are loanwords in the categories of ‘Euphemisms’ and ‘Food and drink’, each 0.5 percent (see Table 4).

‘Abstract and philosophical terms’ is a cover term for words that cannot be listed under any other category. ‘Arabic culture’ has been added to ‘Islamic religion’ to cover words such as *cadhir* ‘veil’. Words in the ‘Food and drink’ category may have entered the language through Persian.

Table 4. Semantic fields

Semantic fields	Loanwords	
Islamic religion and Arabic culture	*donga*	< *duʿāʾ* ‘prayer’
	haram	< *ḥarām* ‘forbidden’
	pekih	< *fiqh* ‘fiqh’
	cadhir	< Persian *čādor* ‘veil’
Abstract and philosophical terms	*aib*	< *ʿayb* ‘shame’
	ajaib ‘miraculous’	< *ʿajāʾib* ‘miracles [pl.]’
	hakekat	< *ḥaqīqa* ‘essence’
Euphemisms	*wafat* ‘to pass away’	< *wafāt* ‘death’
	hamil	< *ḥāmil* ‘pregnant’
Political and military	*sultan*	< *sulṭān* ‘Islamic ruler’
	majelis	< *majlis*, pl. *majālis* ‘council’
	Kures	< *Qurayš* ‘an Arabian tribe’
	laskar	< *al-ʿaskar* ‘the soldiers’
Nautical and trade	*atlas*	< *ʾaṭlas* ‘atlas’
	dinar	< *dīnār* ‘gold coin’
	kutub	< *quṭb* ‘pole’
Botanical and zoological	*bihal*	< *baġl*, pl. *biġāl* ‘mule’
	hewan	< *ḥayawān* ‘animal’
	jaetun	< *zaytūn* ‘olive’
	jerapah	< *zirāfa* ‘giraffe’
Anatomy, medicine	*dakar*	< *ḏakar* ‘male genital’
	tabib ‘medicine man’	< *ṭabīb* ‘physician’
	preji	< *farj* ‘female genital’
	bawasir	< *bawāṣir* ‘hemorrhoids’
Times, dates, and numerals	*Ngahad*	< *ʾaḥad* ‘Sunday’
	Jimakir	< *jumādā l-ʾāxira* [name of month]
	Mulud	< *mawlid* ‘Muḥammad’s birthday’
Education, books, and writing	*abjad* ‘letter’	< *ʾabjad* ‘alphabet’
	ilmu	< *ʿilm* ‘knowledge; science’
	ijasah	< *ʾijāza* ‘diploma’
Cultural innovation	*sabun*	< *ṣābūn* ‘soap’
	merjan	< *marjān* ‘coral; beads’
	rebab	< *rabāb* ‘violin’
Food and drink	*juadah*	< *zuwāda* ‘snack; supply’
	arak	< *ʿaraq* ‘strong drink’
	serbat	< *širba* ‘a drink’

Bibliographical references

Campbell, Stuart. 1996. *The Arabic element in Indonesian: What do students need to know about it?* http://intranet.usc.edu.au/wacana/4/campbell.html. University of the Sunshine Coast, Queensland, Australia.

Echols, John M. and Hasan Shadily. 1989. *Kamus Indonesia-Inggris* [An Indonesian-English dictionary]. 3rd rev. ed. John U. Wolff and James T. Collins. Jakarta: Gramedia.

Geertz, Clifford. 1964. *The religion of Java*. New York: Free Press.

Horne, Elinor Clark. 1974. *Javanese-English dictionary*. New Haven, Conn.: Yale University Press.

Jones, Russell. 1984. "Loanwords in contemporary Indonesian". Verhaar (1984:1–38).

Kamus Besar Bahasa Indonesa (KBBI). 2001. 3rd ed. Jakarta: Balai Pustaka. [Indonesian monolingual dictionary published by the National Office of Education.]

Koentjaraningrat. 1984. *Javanese culture*. Singapore and New York: Oxford University Press.

Partaaatmadja. 1992. *Kamus Praktis Jawa-Indonesia* [A practical Javanese-Indonesian dictionary]. Surabaya: Djojobojo.

Poedjosoedarmo, Soepomo. 1982. "Javanese influence on Indonesian vocabulary". *Pacific Linguistics Series* 38.105–130.

Uhlenbeck, E.M. 1978. *Studies in Javanese morphology*. The Hague: Nijhoff.

Verhaar, John W.M. and Bambang Kaswanti Purwo (eds.). 1984. *Towards a description of contemporary Indonesian: Preliminary studies*, II. (= *NUSA Linguistic Studies in Indonesian and Languages in Indonesia*, XIX.) Jakarta: Universitas Atma Jaya.

Wehr, Hans. 1979. *A dictionary of Modern Written Arabic*. Ed. J. Milton Cowan. Wiesbaden: O. Harrassowitz.

Zoetmulder, Petrus Josephus. 1974. *Kalangwan: A survey of Old Javanese literature*. The Hague: Nijhoff.

Rochayah Machali
(University of New South Wales)

Jazāʾ

The term *jazāʾ* 'requital' is the most common term used by early medieval Arab grammarians in connection with conditionality. It can be interpreted as involvement of a condition. The conditional particles (*ḥurūf al-jazāʾ*) are those that trigger a conditional relation (*mā yujāzā bihi*). In addition to terminological differences in describing conditional structures, Arab grammarians throughout the ages held different opinions concerning the scope of these particles, and hence the structures that they described under this heading also differed.

1. Conditional structures in Classical and Modern Standard Arabic

The following description is only a very brief introduction to the basic types of conditional sentences in Arabic, which seem to have undergone a number of modifications in Modern Standard Arabic. For a detailed description of all the possible variations within these structures in Classical Arabic, see Trumpp (1881) and Peled (1992); for the Qurʾānic usage, see Tietz (1963); and for the changes that are apparent in Modern Standard Arabic, consult Badawi a.o. (2004: Chap. 8).

The basic type of conditional sentences in Arabic is a symmetrical structure made up of a protasis and an apodosis and introduced by a conditional particle. There are three main types of conditional sentences primarily differentiated by the conditional particle which introduces them:

i. *ʾiḏā* introduces possible (probable) conditions: *ʾiḏā duʿīnā ʾajabnā* 'if we are invited, we accept [the invitation]';
ii. *ʾin* (together with some compound particles ending in *-mā*: *mahmā* 'whatever', *ʾaynamā* 'wherever', etc.) introduces real conditions: *ʾin taʿjal tandam* 'if you hurry, you will be sorry';
iii. *law* introduces hypothetical (irreal) conditions; its apodosis is generally introduced by the emphasizer *la-* (→ asseverative particle): *law kunta mubṣiran la-raʾayta bašaratī an-nāṣiʿa* 'if you had been endowed with eyesight, you would have been able to see my white skin' (for a complete analysis of *law* clauses, see Kinberg 1977).

The primary word order is: particle (*ḥarf*) + protasis (*šarṭ*) + apodosis (*jawāb*). In Classical Arabic, in the cases of (i) and (iii) the two verbs are in the perfect, whereas in (ii) the primarily expected verb form is jussive + jussive, but perfect + perfect and other combinations are also accepted (see below). For the historical interpretation of these usages, see Bravmann (1953, esp. Secs. 103 and 105; 1977 and the bibliography cited there). If the apodosis is not the logical consequence of the protasis, or if it is not a verbal sentence, it should be introduced by *fa-*. The use of *fa-* also seems to be semantically determined (Peled 1985, esp. 224). Negation of the verbs in the protasis in all the three types happens with *lam*. The verb in the apodosis is negated by *lam* (+ jussive) in (i) and (ii), and by *mā* (+ perfect) in (iii). The verb *kāna* 'to be' is used to indicate that perfect or past meaning is intended: *ʾin kāna fahima ḏālika kāna xaṭiʾa* 'if he understood it that way, he was mistaken'.

In Modern Standard Arabic, where the inversion of apodosis and protasis is permissible (Badawi a.o. 2004:640–641, 649, 659), the function of *ʾin* – except in the highest register, and in the case of proverbial and set phrases – is being taken over on the one hand by *ʾiḏā* and on the other hand by *law* (cf. Badawi a.o. 2004:636ff.), although the interchangeability of *law* and *ʾin* had already been observed by medieval Arab grammarians (Ibn Yaʿīš, *Šarḥ* II, 1206).

2. TERMINOLOGY OF THE ARAB GRAMMARIANS FOR THE CONDITIONAL STRUCTURE

2.1 *Early grammarians*

Sībawayhi called the whole conditional period *jazāʾ*, the protasis *al-kalām al-ʾawwal* (*Kitāb* I. 387), and the apodosis *jawāb* or *jawāb al-jazāʾ* (*Kitāb* I, 40; III, 387). The name of the apodosis ('answer') refers to the fact that Sībawayhi saw similarities between the interrogative and the conditional sentences. He considers this similarity from several semantic and structural aspects at different places in his book: "The answer (*jawāb*) [given to question words] is like the answer [i.e. apodosis] of the conditional sentence (*jazāʾ*).... The question words, just like the *jazāʾ*, are not necessary [i.e., the verb following it indicates an action that is not a fact, but something that may or may not happen]" (*wa-jawābuhā ka-jawābihi...wa-hiya ġayr wājiba ka-l-jazāʾ*; *Kitāb* I, 40). The relationship between the protasis and the apodosis is similar to that between a question word and the question following it (*Kitāb* I, 385):

> The same can be said about the conditional sentence and the question. Don't you see that if you use a question word you do not make what follows it a → *ṣila*. The correct way is to say that the verb in the conditional sentence is not a *ṣila* to what stands before it [i.e. the conditional particle], and the same applies to the verb preceded by a question word. If you say: *ḥayṯumā takun ʾakun* 'wherever you will be I shall be', the verb is not a *ṣila* to what stands before it. Likewise, if you ask *ʾayna takūnu* 'where will you be?', the verb is not a *ṣila* to what stands before it.... And when you ask *man yaḍribuka* 'who beats you?', or when you make a conditional sentence *man yaḍribka ʾaḍribhu* 'if somebody beats you, I beat him', the verb in neither of them is a *ṣila* (*wa-lākinna l-qawl fīhi ka-l-qawl fī l-istifhām ʾa-lā tarā ʾannaka ʾiḏā stafhamta lam tajʿal mā baʿdahu ṣilatan fa-l-wajh ʾan taqūla al-fiʿl laysa fī l-jazāʾ bi-ṣila li-mā qablahu ka-mā ʾannahu fī ḥurūf al-istifhām laysa ṣilatan li-mā qablahu wa-ʾiḏā qulta ḥayṯumā takun ʾakun fa-laysa bi-ṣila li-mā qablahu ka-mā ʾannaka ʾiḏā qulta ʾayna takūnu wa-ʾanta tastafhimu...wa-taqūlu man yaḍribu-ka fī l-istifhām wa-fī l-jazāʾ man yaḍrib-ka ʾaḍrib-hu fa-l-fiʿl fīhimā ġayr ṣila.*)

Al-Farrāʾ uses the term *jazāʾ* to denote the following meanings (cf. Kinberg 1996:117–118): (i) the condition (i.e. the semantic content of conditional relationship, conditionality); (ii) the conditional sentence; (iii) the protasis; and (iv) the conditional particle as the indicator of a conditional sentence.

As for the apodosis, al-Farrāʾ uses two terms. On the one hand, like Sībawayhi, he usually employs the term *jawāb* or its extended form *jawāb al-jazāʾ* or more specifically *jawāb ʾin*, *jawāb law* (e.g. Farrāʾ, *Maʿānī* II, 506, 202). On the other hand, a special usage of the word *šarṭ* is encountered as well (e.g. *Maʿānī* I, 157; III, 45). In these instances, al-Farrāʾ uses the term *šarṭ* (and *šarṭ li-l-jazāʾ*) in the meaning of 'apodosis'. This usage is only found in the *Maʿānī*. The examples with this term always contain an imperative (*ʾamr*) in their protasis (*Maʿānī* I, 157; III, 45). Sībawayhi does not use this term at all, while later grammarians use it to denote the protasis (e.g. Zamaxšarī, *Mufaṣṣal* 151).

Like Sībawayhi, al-Farrāʾ uses the word *mujāzā* when speaking about the contents of a sentence: "As for *jazm*, it is because of the conditional relationship (*mujāzā*) induced by an imperative" (*Maʿānī* I, 160). Unlike Sībawayhi, however, al-Farrāʾ does not explain an ending from an *ʿāmil* (→ *ʿamal*) but rather from a semantic category. In the case of conditional sentences, this category is that of *mujāzā* (cf. Dévényi 1990:106).

2.2 *Later grammarians*

In later periods, the terminology went through some changes. For the protasis, the term *šarṭ* 'condition' became almost uniformly used, while the earlier term, *jazāʾ*, came to mean 'apodosis' and was used together with *jawāb*, e.g. by Ibn ʿAqīl (*Šarḥ* II, 377), "if there is a second verb after the apodosis" (*ʾiḏā waqaʿa baʿda jazāʾ aš-šarṭ fiʿl*).

Although both *jazāʾ* and *jawāb* were used for the apodosis, later grammarians still regard them as different to a certain extent: "*ʾiḏan* is answer (*jawāb*) and requital (*jazāʾ*) as well. Somebody says *ʾana ʾātīka*, then you say *ʾiḏan ʾukrimaka*. With this you have answered him and at the same time made your hearty welcome a requital for his coming" (Zamaxšarī, *Mufaṣṣal* 151).

The term *šarṭ* was given the original meaning of *jazāʾ* as well: "As for *ʾammā*, it has the role (*maʿnā*) of triggering conditional relations (*šarṭ*)" (Zamaxšarī, *Mufaṣṣal* 151).

3. THE CONDITIONAL PARTICLES

Sībawayhi defined the conditional particles basically from a formal point of view: "The conditional particles put the verbs into the *jazm* (jussive)" (*ḥurūf al-jazāʾ tajzimu l-ʾafʿāl*; *Kitāb* I, 386). Accordingly, he only included (e.g. *Kitāb* I, 384) in this group those particles and other nouns that trigger a conditional relation (*mā yujāzā bihi*), which corresponded to this definition. This means that the conditional particles par excellence are *ʾin* and those words that resemble it in taking two structures with *jazm* after them, while the particle *law* – which can only take a perfect (*māḍī*), and not *jazm* – is in no way connected to these particles in Sībawayhi's description.

On the other hand, constructions with *man*, *mā*, and *allaḏī*, which do trigger the jussive, belong to the conditional particles (Mosel 1975:I, 161ff.). Furthermore, since participles are sometimes equivalent to constructions with *allaḏī* (as in *as-sāriqu* in Q. 5/38 = *allaḏī saraqa*), these participles have conditional meaning (cf. Schöck 2006:79–88).

The grammarians' analysis went hand in hand with a total disregard of the actual usage of verb forms in conditional structures of contemporary prose, where *ʾin* was most often followed by two verbs in the *māḍī* (Amayreh 1983:311ff.), although the language of the *Qurʾān* and pre-Islamic poetry – the two main sources of the grammarians – seems to fit their description (Tietz 1963; Dévényi 1991:43; Peled 1992). It should also be pointed out that as part of the → *ʾiʿrāb* system, *jazm* – and not *māḍī*, which belongs to → *bināʾ* – was the form that called for a linguistic explanation.

Later grammarians – probably under the influence of Greek logic (see Versteegh 1991) – treated the conditional particles and the conditional structure quite differently. So much so that az-Zamaxšarī (*Mufaṣṣal* 150) defines *ʾin* and *law* as the two [primary] conditional particles (*ḥarfā š-šarṭ*).

4. THE CONDITIONAL STRUCTURES

The *Kitāb* contains a series of chapters on conditional structures (*jazāʾ*; *Kitāb* I, 384–401), which do not contain sentences with *law* at all, since these do not contain *jazm*, and so, viewed from a formal criterion, *law* cannot be treated together with *ʾin* and similar particles that entail a modification of *ʾiʿrāb*.

In the first chapter on *jazāʾ*, Sībawayhi defines the conditional structures in the following four rules (Sībawayhi, *Kitāb* I, 386; cf. Peled 1992:2–7):

(i) The conditional particles put the verbs into *jazm* and the [verb] of the apodosis is in *jazm* because of what precedes it (*ḥurūf al-jazāʾ tajzimu l-ʾafʿāl wa-yanjazimu l-jawāb bimā qablahu*).
(ii) The apodosis either is a verb or is introduced by *fa-* (*lā yakūnu jawāb al-jazāʾ ʾillā bi-fiʿl ʾaw bi-l-fāʾ*).
(iii) The basic structure of the apodosis is the verb (*al-jawāb bi-l-fāʾ fī mawḍiʿ al-fiʿl; ʾaṣl al-jazāʾ al-fiʿl*).
(iv) It is grammatically incorrect to have a verb in *jazm* in the protasis which is not being followed by another verb in *jazm* in the apodosis, for this is what corresponds to it in their usage if it [i.e. *ʾin*] is the *ʿāmil* (*qabuḥa fī l-kalām ʾan taʿmala ʾin ʾaw šayʾ min ḥurūf al-jazāʾ fī l-ʾafʿāl ḥattā tajzimahu fī l-lafẓ ṯumma lā yakūnu lahā jawāb yanjazimu bimā qablahu*... *lammā kānat ʾin al-ʿāmila lam yaḥsun ʾillā ʾan yakūna lahā jawāb yanjazimu bimā qablahu fa-hāḏā llaḏī yušākiluhu fī kalāmihim ʾiḏā ʿamilat*).

Tables 1–6 sum up the opinions of some of the major Arab grammarians concerning permissible verb forms in the two parts of the conditional structure introduced by *ʾin*. Parentheses mark the authors' negative opinion of the given structure or the fact that they allow it only under certain circumstances.

Tables 1–6. Opinions of the grammarians about the *jazāʾ*

1. Sībawayhi

1 2	*māḍī*	*jazm*
māḍī		
jazm	(+)	+
rafʿ	(+)	

2. al-Farrāʾ

1 2	*māḍī*	*jazm*
māḍī	+	+
jazm	+	+
rafʿ		

3. Ibn Jinnī

1 2	*māḍī*	*jazm*
māḍī		
jazm		+
rafʿ		

4. az-Zamaxšarī

1 2	*māḍī*	*jazm*
māḍī	+	?
jazm	+	+
rafʿ	+	

5. Ibn al-Ḥājib

1 2	*māḍī*	*jazm*
māḍī	+	
jazm	+	+
rafʿ	+	

6. Ibn Mālik

1 2	*māḍī*	*jazm*
māḍī	+	(+)
jazm	+	+
rafʿ	+	(+)

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Farrāʾ, *Maʿānī* = ʾAbū Zakariyyāʾ Yaḥyāʾ ibn Ziyād al-Farrāʾ, *Maʿānī l-Qurʾān*. Ed. Muḥammad ʿAlī an-Najjār. 3 vols. Cairo: ad-Dār al-Miṣriyya, 1955–1972.

Ibn ʿAqīl, *Šarḥ* = Bahāʾ ad-Dīn ʿAbdallāh Ibn ʿAqīl, *Šarḥ ʿalā l-ʾAlfiyya*. Ed. Muḥammad Muḥyī d-Dīn ʿAbd al-Ḥamīd. 2 vols. 14th ed. Beirut: Dār al-Kitāb al-ʿIlmī, 1964.

Ibn al-Ḥājib, *Kāfiya* = Jamāl ad-Dīn ʿUṯmān ibn ʿUmar Ibn Ḥājib, *al-Kāfiya fī n-naḥw*. Cawnpore: lithography, n.d.

Ibn Jinnī, *Lumaʿ* = ʾAbū l-Fath ʿUṯmān Ibn Jinnī, *Kitāb al-lumaʿ fī n-naḥw*. Ed. Hadi M. Kechrida. Uppsala, 1976.

Ibn Mālik, *ʾAlfiyya* = Jamāl ad-Dīn ʾAbū ʿAbdallāh Muḥammad ibn ʿAbdallāh Ibn Mālik, *al-ʾAlfiyya fī n-naḥw*. Ed. and trans. Antonin Goguyer. Beirut, 1889.

Ibn Yaʿīš, *Šarḥ* = ʾAbū l-Baqāʾ Muwaffaq ad-Dīn Yaʿīš ibn ʿAlī Ibn Yaʿīš, *Šarḥ Mufaṣṣal az-Zamaxšarī*. Ed. Gustav Jahn, *Ibn Jaʿîš Commentar zu Zamachšarî's Mufaṣṣal, nach den Handschriften…herausgegeben*. Leipzig: F.A. Brockhaus, 1876–1886.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *Kitāb Sībawayhi*. Ed. Hartwig Derenbourg, *Traité de grammaire arabe*. 2 vols. Paris: Imprimerie Nationale, 1881–1889.

Zamaxšarī, *Mufaṣṣal* = ʾAbū l-Qāsim Maḥmūd ibn ʿUmar al-Zamaxšarī, *Kitāb al-mufaṣṣal fī n-naḥw*. Ed. Jens Peter Broch. Christiania: W.C. Fabritius, 1859.

SECONDARY SOURCES

Amayreh, Ismail. 1983. *Das Verhältnis zwischen der Theorie der arabischen Nationalgrammatik und dem Textbefund dargestellt am Beispiel des Konditionalsatzes*. Ph.D. diss., University of Erlangen-Nürnberg.

Badawi, Elsaid, Michael G. Carter, and Adrian Gully. 2004. *Modern Written Arabic: A comprehensive grammar*. London and New York: Routledge.

Bravmann, Meir Max. 1953. *Studies in Arabic and general syntax*. Cairo: Institut Français d'Archéologie Orientale.

——. 1977. *Studies in Semitic philology*. Leiden: E.J. Brill. ["The use of the perfect tense in the conditional sentences of Arabic", pp. 563–571.]

Dévényi, Kinga. 1988. "The treatment of conditional sentences by the mediaeval Arabic grammarians". *The Arabist. Budapest Studies in Arabic* 1.11–42.

——. 1990. "On Farrāʾ's linguistic methods in his work *Maʿānī l-Qurʾān*". *Studies in the history of Arabic grammar*, II, ed. Michael G. Carter and Kees Versteegh, 101–110. Amsterdam and Philadelphia: J. Benjamins.

——. 1991. "The Arabic conditionals and 8th century Arab grammarians: Between theory and practice". *al-Karmil* 12.33–44.

Kinberg, Naphtali. 1977. *Hypothetical conditionals in Arabic: A study of* law *clauses*. Ph.D. diss., University of Michigan.

——. 1981–1982. "A study of *la-ʾin* clauses in early literary Arabic". *Jerusalem Studies in Arabic and Islam*. 3.203–222.

——. 1996. *A lexicon of al-Farrāʾ's terminology in his Qurʾān commentary*. Leiden: E.J. Brill.

Mosel, Ulrike. 1975. *Die syntaktische Terminologie bei Sībawaih*. Ph.D. diss., University of Munich.

Peled, Yishai. 1985. "On the obligatoriness of *fa-* in Classical Arabic *ʾin* conditional sentences". *Journal of Semitic Studies* 30.2.213–225.

——. 1992. *Conditional structures in Classical Arabic*. Wiesbaden: O. Harrassowitz.

Schöck, Cornelia. 2006. *Koranexegese, Grammatik und Logik: Zum Verhältnis von arabischer und aristotelischer Urteils-, Konsequenz- und Schlußlehre*. Leiden: E.J. Brill.

Tietz, Renate. 1963. *Bedingungssatz und Bedingungsausdruck im Koran*. Ph.D. diss., Eberhard-Karls-Universität, Tübingen.

Trumpp, Ernest. 1881. "Der Bedingungssatz im Arabischen". *Sitzungsberichte der königlichen bayerischen Akademie der Wissenschaften, Philosophisch-philologische Classe* II/4, 337–446.

Versteegh, Kees. 1991. "Two conceptions of irreality in Arabic grammar: Ibn Hišām and Ibn al-Ḥāǧib on the particle *law*". *Bulletin d'Etudes Orientales* 43.77–92.

KINGA DÉVÉNYI
(Corvinus University)

Jazm → ʾIʿrāb

Jerusalem Arabic

1. GENERAL

Jerusalem Arabic is the colloquial Arabic dialect spoken in Jerusalem. It includes three communal subdialects of the Muslims, Christians, and Jews, although the differences among them are not large (Blanc 1960). Jerusalem Arabic is an urban dialect within the larger dialect area of aš-Šām, which extends from Syria to Lebanon, Israel, and Jordan. The dialects of this region are part of the Eastern dialects within the Arabic-speaking world.

The number of Jerusalem's inhabitants is at present about 500,000 people (mainly Jewish since the establishment of Israel). During history, Jews, Christians, and Muslims inhabited the city and ruled it in turns. Due to its cultural and historical importance, the dialect has been rather well researched. Among these studies see Löhr (1905), Spoer and Haddad (1909), Bauer (1913), Rice and Said (1953), and Piamenta (1964, 1966).

Other textbooks (Kapliwatzky 1941; Ben-Zeʾev 1945, 1948; Piamenta 1968) present local urban colloquial Arabic, but do not usually refer specifically to Jerusalem Arabic. Levin (1994), though mainly based on previous literature, is the major recent grammar of Jerusalem Arabic. Piamenta's studies (1970, 1979, 1981, 1991, 1992) culminate in his book (2000). Dictionaries referring at least partly to Jerusalem Arabic include Barthélemy (1935–1954), Bauer (1957), Denizeau (1960), Elihay (1977), Piamenta (1979), the glossary in Piamenta (2000), and Othman and Neu (2002).

2. LINGUISTIC DESCRIPTION

2.1 *Phonology*

2.1.1 Consonants

Table 1 presents the inventory of consonants and vowels in Jerusalem Arabic. The consonants in brackets mark allophones or borrowed phonemes.

Uvular /q/ is usually pronounced as a glottal stop /ʾ/ (unlike rural and Bedouin dialects). It occurs as uvular mainly in Literary Arabic words, e.g. *al-qurʾān* 'the *Qurʾān*', *al-quds* 'Jerusalem [in Arabic], *qura* 'villages', etc. **k* remains and is not palatalized into /č/.

Table 1. Consonants and vowels in Jerusalem Arabic

feature	bila-bial	labio-dental	alveolar	palatal	velar	uvular	laryn-geal	glottal
stops	(p)b		t d ṭ ḍ		k(g)	(q)		ʾ
fricatives		ṯ ḏ	s z	š ž	x ġ		ḥ ʿ	h
		f (v)	ṣ ẓ					
affricate			(c)	(č) ǧ				
nasals	m		n					
trill			r					
lateral			l					
semivowel	w			y				
vowels	short	long	anaptyctic	diphthongs				
	i e a	ī ē ā	ə	iw aw				
	o u	ō ū						

The interdental phonemes /t̠/, /d̠/ are usually pronounced as /t̠/ or /t/, and as /z/ or /d/ respectively (e.g. *t̠āniye* ‘a second’, *tilmīz* ‘a pupil’). Likewise, /ḏ̣/ is usually merged into /ḍ/ in Jerusalem, e.g. *ḍall* ‘he stayed, remained’, *ḍuhr* ‘noon’, but words of the root *ḏ̣-b-t* have *ẓ*, such as *maẓbūṭ* ‘right, correct’, *ẓābeṭ* ‘officer’. See Al-Wer (2004) on these processes.

The phoneme **g* is articulated in Jerusalem Arabic as either /j/ or (rarely) as /ž/, as often elsewhere in this country, e.g. *jār* (~ *žār*) ‘neighbor’.

The glottal stop /ʾ/ is retained in the middle of certain non-monosyllabic words, e.g. *yuʾmor* ‘he commands’, *saʾal* ‘he asked’, *muʾmin* ‘believer’, but is often deleted at the head of a word, e.g. *ʾakal* > *akal* ‘he ate’; when in a prefixed unstressed syllable of the plural pattern aCCāC, e.g. *wlād* for *ʾawlād* ‘children, boys’; when it is the third consonant of a verb, e.g. *bada* ‘he began’; when it is part of the feminine noun ending /āʾ/, e.g. *sama* ‘sky’; in the imperative of *akal* ‘eat’ and *axad* ‘take, seize’, e.g. *kul* ‘eat!’, *xūd* ‘take!’.

Elsewhere /ʾ/ is substituted by other phonemes: (a) In the environment VʾC (forming a cluster), it is replaced by a long vowel, e.g. *raʾs* > *rās* ‘head’, *yaʾkul* > *yākul* ~ *yōkul* ‘he eats/will eat’, *mēkel* ‘having eaten’. (b) In the sequence ʾVʾ it is replaced by /h/, e.g. *huʾmur* ‘command!’, *huʾʿud* ‘sit!’, *haʾall* ‘less’. (c) At the head of some words /ʾ/ becomes /w/, e.g. *wadda* ‘hand over’. (d) Preceding the vowel /i/ in the middle of a word, /ʾ/ becomes /y/, e.g. *nāyim* ‘asleep’.

The phonemes /g, v, p, c, č/ are borrowed from Hebrew or from European languages, such as English, French, or Italian.

Emphasis (palatalization or velarization) is retained, but is phonetically somewhat weakened. Emphasis may spread to non-emphatic homophonic phonemes as in, for example, *mabṣūṭ* ‘satisfied, happy’, *ṣulṭān* ‘sultan’, and the numerals 11–19 (*ṭnaʿš* ‘twelve’, *xamesṭaʿš* ‘fifteen’). New emphatics occur only in a few words, such as *aḷḷāh* ‘God’, *ḅāḅa* ‘the pope’. In certain words emphatics have become de-emphacized, e.g. *saddaʾ* ‘to believe’, *sandūʾ* ‘box,’ *zġīr* ‘small’ (where *ṣ* > *s, z*).

The laryngeal /ʿ/ is also usually rather weak. A special case is the verb *mazzaʿ* (Literary Arabic *mazzaq*) ‘to tear up’.

Such phonetic changes do not weaken the phonemic status of the phonemes. Thus, for example, emphatics still affect adjacent vowel phonemes.

2.1.2 Vowels

The vowel system includes five short and five long vowels /i, ī, e, ē, a, ā, o, ō, u, ū/. The three short and long vowels /i, u, a/ are the main phonemes of the system, as in Literary Arabic.

The vowels /e, o/ are not always phonemically distinguished from /i, u/, e.g. *zerr* ~ *zirr* ‘button’, *bēthom* ~ *bēthum* ‘their house’. /i/ remains when it is followed by two consonants (iCC), e.g. *nzilt* ‘I ~ you went down’. /e/ may also be a raised allophone of /a/ in → *ʾimāla* of word-final feminine endings, e.g. *madrase* ‘school’, or an anaptyctic vowel, e.g. *isem* ‘name’.

The long vowels /ē, ō/ usually reflect a diphthong in Literary Arabic, e.g. *yawm* > *yōm* ‘day’, *zayt* > *zēt* ‘oil’. In *zitūn* ~ *zetūn* ‘olives’, the first vowel is also short. /ē, ō/ may also be the phonetic articulation of the parallel /ī, ū/ vowels in Literary Arabic, e.g. *rōḥ* ‘soul, spirit’ vs. *rūḥ* ‘go!. Exceptions are *sēʿa* ‘hour; watch’ (also *sāʿa*) and *embēriḥ* ‘yesterday’ with *ʾimāla*.

The vowel /a/ has fronted and back allophones, depending on the environment: labial and coronal consonants attract the forward allophone [æ], whereas laryngeal and pharyngeal consonants attract the back allophone [a]. Emphatic consonants tend to attract a still higher back allophone [ɑ].

A shwa [ə] serves as anaptyctic vowel, with various allophones, to prevent illegal consonant clusters within and between words (*laḥəm* ‘meat’, *bēt-əlwalad* ‘the boy’s house’).

The diphthong /iw/ occurs in roots with /w/ as first consonant, following the prefix *yi-*, as in *yiwṣal* ‘he arrives ~ will arrive’. The diphthong /aw/ is retained when /w/ is a root consonant, e.g. *mawʿūd* ‘there is’, *awsaʿ* ‘wider’.

2.1.3 The syllable

Syllable patterns in Jerusalem Arabic include the following patterns:

Cv, Cv̄, CvC, Cv̄C, CvCC, e.g. *wa-lad* ‘boy, child’; *ʿam-mi* ‘my paternal uncle’; *rāḥ* ‘he went’; *bint* ‘girl’, *min-dīl* ‘scarf, shawl’, *sā-far* ‘he traveled’, *ka-tabt* ‘I ~ you wrote’.

Usually consonant clusters are prevented (by anaptyctic vowels), but in some biconsonantal and triconsonantal clusters, mainly with /t/ or /d/ as second consonant, no anaptyctic vowel occurs, yielding the patterns CCv, CCvC, or CvCC, e.g. *šta*ǵ*al* 'he worked' or *stanna* 'he waited', *miftker* 'thinking [masc. sg.]', *ʿindkom* 'at our place'. Sometimes vowel shifts (reshufflings) occur, affecting word-internal syllable structures, e.g. *yuk-tu-bu* > *yu-kut-bu* 'they write/will write'.

Long vowels in certain syllable patterns may be phonetically shorter due to stress rules (see below).

2.1.4 Stress

Stress rules in Jerusalem Arabic include the following:

i. Syllables include short syllables (Cv), medium syllables (CvC, Cv̄), and long syllables (Cv̄C, CVCC).
ii. A long syllable in the word is stressed, e.g. *fallāḥ* 'farmer', *maktūb* 'written [Literary Arabic]', *tuffāḥ* 'apple', *muftāḥ* 'key'.
iii. If there is more than one long syllable in the word, the last one is stressed and the previous one is slightly shortened, e.g. *fallaḥīn* 'farmers', *tuffaḥāt* 'apples', *mafatīḥ* 'keys'.
iv. If there is no long syllable but there is a medium syllable, it is stressed unless it is at the end of the word. If the medium syllable is the last one in the word, the stress falls on the syllable before it, e.g. *katabtu* 'you [pl.] wrote', *safarna* 'we traveled', *madāres* 'schools', *kassar* 'he broke,' *maxzan* 'storeroom'.
v. If the word has only short syllables, the first syllable in the word is stressed, e.g. *ʾaja* 'he came', *mara* 'woman, wife', *zalame* 'man', *katabu* 'they wrote', *baqara* 'cow', *waladi* 'my child, my son'.

The definite article, many enclitics, and prefixes and anaptyctic vowels do not affect word stress, e.g. *el-wa-lad* 'the boy', *ʿa-l-bēt* 'to the house', *binet* 'girl, daughter,' *iš-ta-ǵal* (or *šta-ǵal*) 'he worked'. However, in some other environments anaptyctic vowels and/or suffixes change the structure of the last syllable and cause it to be stressed, e.g. in *ṣūritha* 'her picture', the negative *–š ~ –šš (mā) katabš* 'he did not write', *(mā) ʿindošš* 'he doesn't have', and the inflected *l-* 'to, for' following a verb, e.g. *kātib* > *katib-lak* 'has written to you'. Thus, syllable structure is dominant and not morphological structure.

2.1.5 Phonotactics

2.1.5.1 Assimilation

The major assimilation phenomenon in Jerusalem Arabic is assimilation of the definite article *l-* by 'sun letters'. Novel sun letters /j/ (and /č/) occur in Jerusalem Arabic.

Other assimilation patterns include spread of emphatization (see above); full assimilation of I*w*/*y* to /t/ of the infix *-ta-*, for example in *ittaṣal* 'he contacted', partial assimilation of the /t/ of the infix *-ta-* to an emphatic first consonant, e.g. *iḍṭarab* 'he was uneasy', *jh* > *š* in *wišš* 'face', and word-final /ḥ/ > /ʿ/ when followed by a word-initial /ʿ/ *rāḥ ʿa-lbēt* > *rāʿ ʿa-lbēt* 'he went home'.

2.1.5.2 Dissimilation

Some quadriradical words seem to have developed by dissimilation from triradical roots, e.g. *bi-l-imšabraḥ* 'explicitly' <*mšarraḥ* 'explained'.

2.1.5.3 Metathesis

marsaḥ for *masraḥ* 'theater' is found, though not frequently nowadays, in Jerusalem Arabic. Examples such as *maʿlaʾa* 'spoon' (cf. Literary Arabic *milʿaqa*), *naʿal* (for *laʿan*) 'to curse', or *ijer* (for *rijl* via **lijr*) 'leg' exist as well.

2.1.5.4 *ʾImāla*

Word-final → *ʾimāla*, as in *kalbe* 'bitch', is found in Jerusalem Arabic. The two words *sēʿa* 'hour; watch' and *embēriḥ* 'yesterday' are rare cases of internal *ʾimāla*.

2.1.6 Morphophonology

2.1.6.1 Vowel elision

Unstressed vowels elide under the following conditions:

i. Feminine form of active participles, e.g. *kātbe* 'having written; writer [fem.]' *mʿal(l)me* 'teacher [fem.]'
ii. An active participle singular to which a bound pronoun is suffixed, e.g. *kātbo* 'having written it', *samiʿto* 'she hears him'
iii. A short unstressed vowel preceding a stressed vowel in adjectives, e.g. *kbīr* 'big [sg. masc.]
iv. A verb in the 3rd person feminine singular, e.g. *katbat* 'she wrote'

v. The preformative of Form V, e.g. *tʿallam* ‘he studied’
vi. Imperfect and imperative forms of Form VIII verbs: *yištġil* ‘he works/will work’, *ištġil* ‘work!’

2.1.6.2 Vowel insertion

Shwa is inserted to break triconsonantal or quadriconsonantal clusters, especially when suffixes are added. These anaptyctic vowels usually do not change the stress pattern of the word or the phrase, e.g. *walad-i-kbīr* ‘a big boy’, *binet* ‘girl, daughter’, but *ṭāwle* ‘table’ > *ṭawilto* ‘his table’, *ṭawlitna* ‘our table’.

2.1.6.3 Construct state and suffixation

In construct state, *-t* (*tāʾ marbūṭa*) is suffixed to feminine word endings, e.g. *maliket jamāl* ‘a beauty queen’. In the bound personal pronoun, this added *t-* is the hinge of the added syllable, e.g. *ġuriftak* ‘your [masc. sg.] room’; *šayifto* ~ *šayfito* ‘she sees him’.

2.2 *Morphology*

2.2.1 Pronouns

2.2.1.1 Personal pronouns

Table 2 shows independent and bound pronouns in Jerusalem Arabic. As in other urban dialects of the area, gender difference is limited to the 2nd and 3rd person singular. The independent pronoun for the 3rd person singular and plural can be short or long: *hū* ~ *huwwe, hī* ~ *hiyye, hum* ~ *humme.*

When a word ends in a vowel, the 1st person singular bound pronoun begins with /y/ or /n/. For the 3rd person singular masculine, the final vowel of the word is lengthened. Sometimes, a weak /h/ can be heard following this vowel. In these words, 2nd and 3rd person suffixes begin with a consonant (/k/ or /h/, respectively).

Following a verb, the direct object suffix of the 1st person singular is *–ni* (not in the table).

2.2.1.2 Relative pronoun

The relative pronoun for both genders and numbers has only one uninflected form: *illi* ~ *elli.*

2.2.1.3 Demonstrative pronouns

Demonstrative pronouns are shown in Table 3 (allomorphs between brackets). Short forms exist for near objects in addition to long forms.

2.2.2 Nouns

Noun structures include root consonants only (e.g. *walad* ‘boy’, *dars* ‘lesson’) or root consonants with affixes (e.g. *maʾʿad* ‘seat’, *madrase* ‘school’, *muftaraʾ* ‘road junction’, *xaddām* ‘servant’, *ʿarabūn* ‘guarantee’).

The noun system distinguishes natural and grammatical gender (masc./fem.) usually by suffixing a feminine ending to the (unmarked) masculine form, e.g. *kalb/kalbe* ‘dog/bitch’. The number category distinguishes between singular, dual, and plural forms. Dual forms exist only for nouns. These distinctions are marked by suffixation of *-ēn* for the dual, *-ūn* for sound masculine plural, and *-āt* for sound feminine patterns. Different patterns take ‘broken plural’ forms, e.g. *dukkān/dakakīn* ‘shop/shops’, *yōm/ayyām* ‘day/days’.

Table 2. Independent and bound pronouns in Jerusalem Arabic

pronouns	independent pronouns	bound pronouns	
		after consonant	after vowel
3rd sg. masc.	*huwwe*	*-o*	-lengthened vowel
3rd sg. fem.	*hiyye*	*-(h)a*	*-ha*
3rd pl.	*humme*	*-hom*	*-hom*
2nd sg. masc.	*inte*	*-ak*	*-k*
2nd sg. fem.	*inti*	*-ek*	*-ki*
2nd pl.	*intu*	*-kom*	*-kom*
1st sg.	*ana*	*-i*	*-yy(i)/-ni*
1st pl.	*iḥna* ~ *niḥna*	*-na*	*-na*

Table 3. Demonstrative pronouns in Jerusalem Arabic

number/gender	near object (normal)	near object (short)	far object (normal)
sg. masc.	*hāda*	*hā, hāy*	*hadāk*
sg. fem.	*hādi*	*hā, hāy*	*hadīk*
pl. masc. fem.	*hadōl (hadōla, hadōli)*	*hā, hāy*	*hadulāk (hadulīk)*

2.2.3 Numerals

As in the rest of the region, 'one' and 'two' distinguish gender by separate forms for the masculine and feminine (*wāḥad/waḥade* 'one', *tnēn/tintēn* 'two'); the numbers 3 through 10 mark syntactic cases, i.e. independent (nominal) vs. construct state of the numeral: the short (fem. or base) form for the construct state and the long form (reflecting an old *tāʾ marbūṭa*) for the independent form, e.g. *talat iwlād/banāt* 'three boys/girls' vs. *talāte* 'three'. The numbers 11 through 19 distinguish the same syntactic cases: the independent form uses a short form of the suffix – *ʿaš* whereas the construct state form uses the full suffix – *ʿšar*, e.g. *tnaʿšar walad/bint* 'twelve boys/girls' and *tnaʿaš* 'twelve'. In certain cases, emphasis spreads in these numbers, e.g. *talaṭʿšar/talaṭṭaʿeš* 'thirteen', *xamesṭaʿeš/xamesṭaʿšar* 'fifteen', etc.

Tens get the suffix *-īn*, e.g. *ʿišrīn* 'twenty', *talatīn* 'thirty'. *miyye* is 'one hundred', *mitēn* 'two hundred', *talatmiyye* 'three hundred', etc.; *ʾalf* is 'one thousand,' *ʾalfēn* 'two thousand', '*talat-alāf*' 'three thousand', etc.

When the counted noun begins with a vowel (due to a deleted glottal stop), the numeral often begins with an affixed *t*-, as in *xams-t-alāf* 'five thousand'.

Ordinals use the CāCiC pattern, e.g. *tāni/tānye* 'second [masc./fem.]', *rābiʿ/rābʿa* 'fourth [masc./fem.]', *sādes/sādse* 'sixth [masc./fem.]', except for *ʾawwal* 'first [masc.]'/*ʾūla* 'first [fem.]'. The dissimilation of the double /t/ of *sitte* 'six' recurs in fractions, with the pattern CuCC, e.g. *tult* 'one third', *xums* 'one fifth', *suds* 'one sixth'.

2.2.4 Adjectives

Adjectives have singular and plural masculine and feminine forms. Many adjective patterns occur in CāCiC and maCCūC active and passive participle forms, respectively. Among the most frequent adjective patterns are CaCīC or CCīC (*ṭawīl* 'long', *kbīr* 'big'), CāCiC (*wāsiʿ* 'broad'), CaCCān (*taʿbān* 'tired'), CaCūr (*faxūr* 'proud'). The pattern aCCaC denotes colors and defects as well as elatives and superlatives, e.g. *aswad* 'black', *aʿma* 'blind', *akbar* 'bigger, biggest'. Most of the adjectives form the feminine as in nouns, by suffixing the ending *-a* ~ *-e* to the masculine form; adjectives indicating colors or defects use a different pattern, namely CaCCa, e.g. *zarʾa* 'blue [fem.]', *ʿamya* 'blind [fem.]'.

2.2.5 Verbs

2.2.5.1 Forms

Quadriradical roots are declined as verbs of Forms II and V, e.g. *tarjam* 'to translate', *tzaḥlaʾ* 'to slide'.

2.2.5.2 Form I

Form I has two subpatterns in the perfect inflection and three subpatterns in the imperfect conjugation. These differ by the vowel following the second root consonant (see Table 4, the conjugation of *katab* 'to write' and *nizel* 'to descend').

The CiCeC pattern in the perfect conjugation consists of two subgroups: 1st and 2nd person vs. 3rd person (sg. and pl.). This grouping recurs in the conjugation of the weak verbs and geminates.

In the imperfect 2nd person singular feminine, 2nd plural, and 3rd plural, a vowel reshuffling occurs, e.g. *tuktubi* > *tukutbi* 'you [sg. fem.] write'.

The imperfect conjugation presented here is used in the subjunctive mood; for the indicative, *b-* is prefixed to the verb forms, e.g. *btuktub* 'you [sg. masc.] write'.

The glottal stop of the 1st person prefix *ʾa-* is elided when *b-* is prefixed to it, e.g. *baktob* 'I

Table 4. Form I verbs in the perfect and imperfect conjugation in Jerusalem Arabic

	perfect			imperfect		
	3rd pers.	2nd pers.	1st pers.	3rd pers.	2nd pers.	1st pers.
sg. masc.	*katab* *nizel*	*katabt* *nzilt*	*katabt* *nzilt*	*yuktob* *yinzel* *yirkab*	*tuktob* *tinzel* *tirkab*	*ʾaktob* *ʾanzel* *ʾarkab*
sg. fem.	*katbat* *nizlet*	*katabti* *nzilti*		*tuktob* *tinzel* *tirkab*	*tukutbi* *tinizli* *tirkabi*	
pl.	*katabu* *nizlu*	*katabtu* *nziltu*	*katabna* *nzilna*	*yukutbu* *yinizlu* *yirkabu*	*tukutbu* *tinizlu* *tirkabu*	*nuktob* *ninzel* *nirkab*

write'; 3rd person *y-* can also be elided in the indicative, e.g. *buktob* 'he writes'. In the 1st person plural *b-* is often assimilated to the following nasal and is replaced by *m-*, e.g. *mnuktob* 'we write'. When this *b-* causes a consonant cluster, a prosthetic vowel can precede it, e.g. *ibtuktob* 'she writes', or follow it, e.g. *bitkūn* 'you [sg. masc.] are'.

Tables 5, 6, and 7 present the conjugation of the weak and geminate verbs in the perfect and imperfect tenses.

2.2.5.3 Derived forms

The derived verb forms vary by affixes: gemination of the 2nd consonant (Form II *ʿallam* 'to teach', V *tʿallam* 'to learn'); prefixation of *ʾa-* (Form IV, *ʾakram* 'respect') or of *ta-* ~ *t* (Form V *tmaddad* 'to stretch out', VI *tnāzal* 'to do without'), lengthened vowel following the first consonant (Form III *sāfar* 'to travel', VI *txānaʾ* 'to quarrel with'), prefixed *n-* (Form VII *inʾatal* 'to get killed') and *st-* (Form X *staʿmal* 'to use'), infixed – *t-* (Form VIII *ištarak* 'to participate'), and geminated third consonant (Form IX *iswadd* 'to become black'). See Table 8.

Jerusalem Arabic has no internal passive; thus, *wiled, xileʾ* 'he was born' are isolated residues. The passive and reflexive are expressed mainly by Form VII, e.g. *inmasak* 'he was caught', Form V, e.g. *tḥakkam* 'he was treated medically', and Form VIII *iltazam* 'to oblige oneself'.

Form II has practically taken over the causative function of Form IV; thus, only a few verbs or partial declensions in Form IV are found in Jerusalem Arabic. Similarly, Form IX, usually denoting the process or change of physical features and color, is hardly used and is replaced by *ṣār* 'to become,' e.g. *ṣār aswad* = *iswadd* 'he became black'.

2.2.5.4 Participle patterns

Active and passive participles are inflected only to gender and number. Thus, they are semi-verbal and semi-nominal. Table 9 presents the participles in Form I.

The active participle of Form I is CāCeC. However, the participle of *akal* 'to eat' and *axad* 'to take' adds a prefixed *m-*: *mēkel, mēxed.*

The pattern of the passive participle is maCCūC, e.g. *maktūb* 'written', *maftūḥ* 'open(ed)'. The patterns of active and passive participles in the derived forms are presented in Table 10, although the passive participle in derived forms is not frequent. The forms with asterisk in Table 10 are theoretical only.

2.2.5.5 Verbal nouns

Numerous forms are considered verbal nouns of Form I (e.g. *naʾl* 'transfer', *ziyāde* 'addition', *wuʾūf* 'standing manner, stance'). Weak roots yield variations, e.g. *talbiye* 'fulfillment [a wish]' (Form II). Form III has two verbal nouns. See Table 11 for examples (→ verbal nouns).

2.3 *Syntax*

2.3.1 The noun phrase

Nouns are defined by the prefixed definite article (*il-, el-*). Jerusalem Arabic has no indefinite article for indefinite nouns, although *wāḥad* (masc.), *waḥde* (fem.) 'one, someone' sometimes denotes the indefinite.

Table 5. Perfect and imperfect conjugation of II*w/y* roots in Jerusalem Arabic

gloss	to be	to occur	to fear	to be	to occur	to fear
		perfect			imperfect	
3rd sg. masc.	*kān*	*ṣār*	*xāf*	*ykūn*	*yṣīr*	*ynām*
3rd sg. fem.	*kānat*	*ṣārat*	*xāfat*	*tkūn*	*tṣīr*	*tnām*
3rd pl.	*kānu*	*ṣāru*	*xāfu*	*ykūnu*	*yṣīru*	*ynāmu*
2nd sg. masc.	*kunt*	*ṣirt*	*xuft*	*tkūn*	*tṣīr*	*tnām*
2nd sg. fem.	*kunti*	*ṣirti*	*xufti*	*tkūni*	*tṣīri*	*tnāmi*
2nd pl.	*kuntu*	*ṣirtu*	*xuftu*	*tkūnu*	*tṣīru*	*tnāmu*
1st sg.	*kunt*	*ṣirt*	*xuft*	*akūn*	*aṣīr*	*anām*
1st pl.	*kunna*	*ṣirna*	*xufna*	*nkūn*	*nṣīr*	*nnām*

Table 6. Perfect and imperfect conjugation of III*w/y* roots in Jerusalem Arabic

gloss	to throw	to forget	to read	to throw	to forget	to read
		perfect			imperfect	
3rd sg. masc.	*rama*	*nisi*	*ʾara*	*yirmi*	*yinsa*	*yiʾra*
3rd sg. fem.	*ramat*	*nisyet*	*ʾarat*	*tirmi*	*tinsa*	*tiʾra*
3rd pl.	*ramu*	*nisyu*	*ʾaru*	*yirmu*	*yinsu*	*yiʾru*
2nd sg. masc.	*ramēt*	*nsīt*	*ʾarēt*	*tirmi*	*tinsa*	*tiʾra*
2nd sg. fem.	*ramēti*	*nsīti*	*ʾarēti*	*tirmi*	*tinsi*	*tiʾri*
2nd pl.	*ramētu*	*nsītu*	*ʾarētu*	*tirmu*	*tinsu*	*tiʾru*
1st sg.	*ramēt*	*nsīt*	*ʾarēt*	*ʾarmi*	*ʾansa*	*ʾaʾra*
1st pl.	*ramēna*	*nsīna*	*ʾarēna*	*nirmi*	*ninsa*	*niʾra*

Table 7. Perfect and imperfect conjugation of geminate roots in Jerusalem Arabic

gloss	to knock	to destroy	to stay	to knock	to destroy	to stay
		perfect			imperfect	
3rd sg. masc.	*daʾʾ*	*hadd*	*ḍall*	*yduʾʾ*	*yhidd*	*yḍall*
3rd sg. fem.	*daʾʾat*	*haddat*	*ḍallat*	*tduʾʾ*	*thidd*	*tḍall*
3rd sg.	*daʾʾu*	*haddu*	*ḍallu*	*yduʾʾu*	*yhiddu*	*yḍallu*
2nd sg. masc.	*daʾʾēt*	*haddēt*	*ḍallēt*	*tduʾʾ*	*thidd*	*tḍall*
2nd sg. fem.	*daʾʾēti*	*haddēti*	*ḍallēti*	*tduʾʾi*	*thidd*	*tḍalli*
2nd pl.	*daʾʾētu*	*haddētu*	*ḍallētu*	*tduʾʾu*	*thiddu*	*tḍallu*
1st sg.	*daʾʾēt*	*haddēt*	*ḍallēt*	*ʾaduʾʾ*	*ʾahidd*	*ʾaḍall*
1st pl.	*daʾʾēna*	*haddēna*	*ḍallēna*	*nduʾʾ*	*nhidd*	*nḍall*

Table 8. The base form of derived verb forms in the perfect and imperfect in Jerusalem Arabic

Derived Form	perfect	imperfect	gloss
II	*sallam*	*yisallem*	‘to greet’
III	*sāfar*	*ysāfer*	‘to travel’
IV	*ʾarsal*	*yirsil*	‘to send’
V	*tʿallam*	*yitʿallam*	‘to learn’
VI	*tnāwal*	*yitnāwal*	‘to take’
VII	*inkasar*	*yinkser*	‘to be broken’
VIII	*iftakar*	*yiftker*	‘to think’
IX	*iswadd*	*yiswadd*	‘to get black’
X	*istaxdam*	*yistaxdem*	‘to use’

Table 9. Active and passive participle in Form I in Jerusalem Arabic

	sg. masc.	sg. fem.	pl. masc.	pl. fem.	gloss
active	*kāteb*	*kātbe*	*katbīn*	*katbāt*	‘having written’
active	*ʾāʿed*	*ʾāʿde*	*ʾāʿdīn*	*ʾāʿdāt*	‘sitting’
passive	*maṭlūb*	*maṭlūbe*	*maṭlubīn*	*maṭlubāt*	‘wanted’
passive	*maʿrūf*	*maʿrūfe*	*maʿrufīn*	*maʿrufāt*	‘known, famous’

Table 10. Active and passive participles of derived forms in Jersualem Arabic

Derived Form	active	passive	gloss
II	*mʿallem*	*mʿallam*	‘to teach’
III	*mkāteb*	*mkātab*	‘to correspond’
IV	*mirsil*	*mursal*	‘to send’
V	*mitʿawwed*	**mitʿawwad*	‘to get used to’
VI	*mitnāzel*	*mitnāzal*	‘to do without’
VII	*minikser*	**minkasar*	‘to be broken’
VIII	*mintʾil*	*muntaʾal*	‘to move [place]’
IX	*miswadd*	*miswadd*	‘to become black’
X	*mistaxdem*	*mustaxdam*	‘to use something’

Table 11. Verbal nouns in derived verb forms in Jerusalem Arabic

Derived Form	form	gloss
II	*taˁlīm*	teaching
III	*mkātabe,*	correspondence
	firāʾ	parting with someone
IV	*ikrām*	respect (for)
V	*taṣarruf*	behavior, attitude
VI	*tanāzul*	doing without
VII	*inkisār*	being broken
VIII	*ištirāk*	participation
IX	*iswidād*	becoming black
X	*istixdām*	use, using something

The nominal head may be free (*wlād* 'children'), have a bound pronoun suffixed to it (*wlādo* 'his children'), or may be part of a construct state structure (*wlād elmudīr* 'the manager's children').

Construct state structures often denote possession or belonging, e.g. *ktāb ilwalad* 'the boy's book', *ʾiben ˁamm* 'masc. cousin'. When the first (dependent) noun is feminine, *tāʾ marbūṭa* appears at its end, e.g. *sēˁat muˁāyane* 'reception hour [clinic]'. In a 'double' construct structure the preposition *l-* is added to the second noun in the construct structure, following the bound pronoun suffixed to the first noun in the structure, e.g. *bēto la-mūsa* 'Mūsa's house'.

Analytic possession structures are formed by inserting a preposition between the two nouns, e.g. *elktāb tabaˁ elwalad* 'the boy's book'. This occurs when it is (i) phonetically conditioned, e.g. when the dependent word ends with a vowel and cannot take the final *t-* (e.g. *irrādyo tabaˁ jārti* 'my neighbor's radio'), (ii) in constructs with more than one noun, e.g. *ilbēt u-ssayyāra tabaˁ axūy* 'my brother's house and car', (iii) when an adjective complements the first noun of the construct, e.g. *ilbēt ilikbīr tabaˁ ilmudīr* 'the manager's big house'.

Adjectives follow the nominal head and usually agree with it in number and gender. When the head is an inanimate plural, the adjective often takes the masculine plural form (instead of the fem. sg. form expected according to the Literary Arabic rule), e.g. *byūt kbār ~ kbīre* 'big houses'.

2.3.2 Quantifiers

Quantifiers usually precede the noun: *kull innās* 'all the people', *baˁḍ innās* 'some people'. However, 'many people' can be *ktīr nās, nās ktīr*, or *nās ktār*, the latter with the quantifier as adjective (inflected in gender and number). The preposition *min* 'from' is used as partitive; thus, *yōm min iliyyām* 'one day [lit. a day from the days]'. The preposition *baˁḍ* is used also in reciprocals, e.g. *baˁḍhom baˁḍ* 'each other, one another'.

2.3.3 The numeral phrase

The singular number 'one' follows the noun as an adjective, e.g. *zalame wāḥad* 'one man'. The number 'two' is usually substituted by the dual (e.g. *waladēn* 'two boys'). The number *tnēn* 'two [masc.]', *tintēn* 'two [fem.]' may follow the noun in the dual or plural to stress the number (*waladēn tnēn ~ wlād tnēn* 'two boys'). The numeral phrase uses the noun in the plural form for the numbers 3 through 10 (*talat iwlād* 'three boys,' *talat banāt* 'three girls'); above this number, the noun is only in the singular, e.g. *sitt-mīt dolār* 'six hundred dollars'.

2.3.4 Elatives and superlatives

The elative has only one uninflected (masc. sg.) form, e.g. *huwwe akbar min ʾuxto* 'he is bigger than his sister', *hum akbar min ʾuxto* 'they are bigger than his sister'. The same pattern is used for the superlative, usually in construct with the head noun, e.g. *maryam akbar bint fi ṣṣaff* 'Maryam (is) the biggest girl in the class'. The Literary Arabic structure [(definite article + noun) (definite article + aCCaC)] also occurs (with aCCaC only), e.g. *maryam hiyye lbint elʾakbar fi ṣṣaff* 'Maryam (is) the biggest girl in the class'.

2.3.5 Relative clauses

Relative clauses follow the uninflected relative particle *illi*, e.g. *ilwalad illi rāḥ hunāk huwwe axūk* 'the boy who went there is your brother'. The Literary Arabic rule concerning the deletion of the relative particle when the antecedent noun is indefinite is not always observed; thus, the following is possible: *šuft walad illi rāḥ fi ššāriˁ* 'I saw a boy who went in the street'.

2.3.6 The verbal phrase

Verbal phrases include intransitive and transitive verbs governing direct or indirect objects (with or without attached negative particles) and verbs governed by modal verbs or particles. Generally speaking, the system is similar to that in Literary Arabic and regional dialects. Only special structures are referred to here.

The preposition *iyya-* is used to indicate the suffixed accusative following an indirect object, e.g. *aʿṭāni iyyā* ‘he gave it to me’.

2.3.7 Functions of participles

The active participle indicates (a) ongoing actions for verbs of motion or space and verbs of sensory perception (e.g. ‘to go, sit, lie down, see, hear’); see also *lābes* ‘wearing, being dressed’; (b) actions that have already taken place and are still valid, similar to the present perfect tense in English (e.g. *huwwe msāfer* ‘he has traveled’, i.e. he is not here now; *huwwe mēkel* ‘he has eaten’, i.e. he is not hungry now).

The passive participle refers to a completed action, the result of an action, or a situation, e.g. *maktūb* ‘written’, *maftūḥ* ‘open(ed).’ However, active participles often replace passive participles of derived forms above Form II.

2.3.8 Verbal aspect, time, and tense

The imperfect conjugation takes a prefix *b-* for the indicative mood (*byifʿal*). This form is usually used to indicate (non-past) present, habitual, or future actions, but may also be used for past events in a lively narrative.

The prefix *b-* is deleted in the subjunctive mood, usually after certain modal verbs, certain prepositions, and other particles. Examples are given in Table 12.

To imply durativity the prefix *ʿam*, and less frequently also *ʿamma-*, or *ʿammāl* precedes the *b+yifʿal* form, e.g. *ʿam buktob* ‘he is writing’.

The form *rāḥ, rāyiḥ, ḥā-*, or *biddo* followed by the subjunctive indicates the future tense.

To imply imminence and processes, verbs such as *kān* ‘to be’, *ʾām* ‘to get up’, *ʾaʿad* ‘to sit down’, *naṭṭ* ‘to jump’, etc. are used, followed by verbs in the perfect or imperfect, e.g. *ʾām ʾāl* ‘he said’, *ʾaʿad katab* ‘he began writing’, *ʾaʿad yuktob* ‘he sat down to write, he was writing’.

The ‘narrative imperative’ indicates a lively action in the past. This structure, usually considered typical of Bedouin dialects, is also found in Jerusalem Arabic. Wishes can be expressed by perfect tense or by *yifʿal* verbs following a conditional particle such as *ya rēt* ‘would that…’, *kān* ‘if’, *kinno* ‘if’, or *law* ‘if’ (irrealis).

2.3.9 Negation

Negation particles are of three types: *mā, lā* (also *laʾ, laʾʾa*, and even *laʿ, lah*), and *miš ~ muš*.

Preceding perfect verb forms, *mā* negates the past; *mā* before *byifʿal* negates the imperfect; preceding *yifʿal, mā* indicates a negative command (i.e. a prohibition). The particle *mā* also serves to enhance activity (like the jussive), if the whole utterance has a different intonation (pitch and loudness rising at the end of the

Table 12. Examples of components governing the subjunctive verb in Jerusalem Arabic

modal element	subjunctive verb form	gloss
kān	+ *yifʿal*	‘he used to (do)’
ṣār	+ *yifʿal*	‘he began (doing)’
biddo	+ *yifʿal*	‘he wants/has to (do)’
lāzim	+ *yifʿal*	‘he must (do)’
byiʾdar	+ *yifʿal*	‘he can (do)’
mamnūʿ	+ *yifʿal*	‘it is forbidden (to do)’
ʾabel ma	+ *yifʿal*	‘before (doing)’
baʿed ma	+ *yifʿal*	‘after (doing)’
rāḥ, rāyiḥ	+ *yifʿal*	‘he is going to do, he intends to do’

utterance); *mā* also negates → pseudo-verbs such as *fī* 'there is' or *ʿind-, ʾil-, maʿ-* 'have something (by someone).'

The particle *lā* indicates a negative command (i.e. a prohibition) when preceding *yifʿal* verb forms. It also serves for the 'overall negation', i.e. implying negation in holophrastic utterances. *lā* can also appear in double negation with any tense, e.g. *lā ʾakal wa-la šireb* 'he neither ate nor drank'.

Usually a correlative suffix *-š* follows *mā* and *lā*, e.g. *mā katab-š* 'he did not write', *lā tuktob-š* 'don't write!'.

miš ~ *muš* negates nominal sentence constituents, including active and passive participles, nouns, adjectives, prepositions, and certain particles, e.g. *muš ana* 'not me', *muš ʿāref* 'don't ~ doesn't know', *muš fi lbēt* 'not at home,' *muš sāken hōn* 'not living here', *muš rāḥ yuktob* 'he is not going to write'.

2.3.10 Word order

The unmarked word order is SVO. VSO is somewhat more marked, when the subject is not a bound pronoun. It occurs almost as frequently as SVO.

Negative particles precede the negated word. In addition, a correlative *-š* is usually suffixed to verbs in the perfect or imperfect forms.

Interrogative particles usually head a sentence, e.g. *mīn hāda* 'who is this?' For pragmatic or thematic functions, the interrogative particle may sometimes follow the subject or predicate, e.g. *šuft ʾēš* 'you saw – what?'

2.3.11 Agreement/concord

Within a sentence, the predicate and complements agree to the subject or headword in gender and number. But when the subject or headword is an inanimate plural noun, adjectives, verbs, and other complements may take the plural form or the feminine singular form.

2.3.12 Interrogations

In addition to the usual interrogative particles, *iši* at the end of the question expresses doubt or vagueness, e.g. *elmuxtār ʿazamkom iši* 'has the village chief invited you [at all, possibly]?'. Rhetorical questions may begin with any of the words *ʾayy, hū*, or *miš*; e.g. *hū ana aṭraš* 'am I deaf?', *ʾayy hu biftaḥ tummo ʾuddām marato* 'will he open his mouth in front of his wife?', *miš ʾultillak hāy* 'haven't I told you this?'.

2.3.13 Subordinated clauses

Subordination is expressed with or without subordinating particles. Adverbial clauses include temporal and locational clauses, often headed by particles, although asyndetic subordination also occurs. Typical particles are *lamma* 'when', *ʾabel ma* 'before', *baʿad mā* 'after', *ma-dām* 'as long as', *ḥattā* 'until', *wēn-mā* 'wherever'.

→ *Ḥāl* (circumstantial) clauses begin with *w-* or *u-*, followed by the subject and the verb in the imperfect or active participle. *Ḥāl* clauses may precede the main clause.

Conditional clauses begin with *iza, kān*, or *kūn* 'if.' The clause may be verbal or nominal. The particle *ʾin* is used in fossilized formulae, e.g. *in-ša-lla* 'God willing'. For unreal conditions, *law* (irrealis) and *lōla* (negative irrealis) are used.

3. LEXICON

Jerusalem Arabic speakers belong to at least one of three large ethnic and religious groups: Muslim, Christian, and Jewish. The Jerusalem Arabic lexicon comprises community-specific elements, mainly those related to sociocultural domains. Piamenta (2000) describes the Jerusalem Judaeo-Arabic dialect and compares it with general Jerusalem Arabic. The present linguistic trend in Jerusalem is toward standardization, unification, and koineization, based on Modern Standard Arabic, Arabic communal dialects, and Modern Hebrew. This tendency has been going on in Jerusalem since 1948 (the establishment of Israel) and especially since 1967. Other sources of influence on the Jerusalem Arabic lexicon are Turkish, due to the Ottoman rule of the area until the end of World War I, and European languages, such as Italian, Spanish, and French. Ladino and Massoretic Hebrew influenced Jerusalem Judaeo-Arabic. Since the 20th century, English has also been lending lexical elements to Jerusalem Arabic. These borrowings fill cultural voids, enable contemporary emotive expressions, and form modern doublets with original colloquial expressions. This situation leads to the decay of specific features of Jerusalem Arabic. Some examples of Jerusalem Arabic vocabulary are given in Table 13. Most of them have cognates in the country and the surrounding regions.

Table 13. Examples of vocabulary used in Jerusalem Arabic

	lexeme	gloss	notes
nouns	*ḥrām*	‘[woolen] blanket’	
	maʾām	‘rank, position of dignity’	
	xazāne	‘wardrobe’	
	sufra	‘[dining] table, set table’	
	hawwāye	‘fan, fire fan’	
	lagan	‘wash basin’	< Persian
	nafs	‘appetite, wish to eat’	
	mšaxxeṣ	‘actor’	
	ṭābe, ṭabbe	‘ball [toy]’	
	luġa, luġġa	‘language, dialect’	
	šadde	‘(deck of) cards’	
verbs	*ḥass*	‘to feel, sense’	
	irtaʿab	‘to become frightened, alarmed’	
	nišef	‘to be or become dry’	
	ʿaṭa kilme	‘to give one’s word, promise’	
	thanna bi-ḥayāto ~ bi ʿīšto	‘to prosper, enjoy happiness in life’	
adjectives	*mahmūm*	‘worried’	
	mdāyaʾ	‘annoyed, oppressed; short of cash’	
	ʾādami, ʾiben nās	‘genteel, well-bred’	
	middayyen	‘pious, religious’	
	mʾangej	‘joining arms with someone’	< French
	mhandaz	‘arranged, adorned’	< Persian
adverbs	*zayy ennās*	‘decently, properly; normally’	
	ʿa lmōḍa	‘fashionable’	< French
	eṣṣuboḥ	‘in the morning’	
	ellēle	‘at night’	
	bukra	‘tomorrow’	
	lēl u nhār	‘day and night’	
	jumʿet ejjāy	‘next week’	
	ʿāmin-awwal	‘last year’	
prepositions	*min ~ minni*	‘from’	geminated before bound pronouns beginning with vowel
	ʾili – ʾilak	‘to have’	in the negative and before nouns, prosthetic *i* of *ʾili* etc. is deleted
	ʿindi – ʿindak		
	maʿi, maʿāy	‘with; to have’	two declension bases
	fiyyi, fīk, fi(h)	‘in, within’	
	wiyyāni, wiyyāk	‘with me, with you’	

Table 13 (*cont.*)

	lexeme	gloss	notes
interrogatives	*šū* *ʾēš*	‘what?’	only *ʾēš* after prepositions
	ʾayya, *ʾanu, ʾāni*	‘which?’	*ʾayya* for all numbers and genders; *ʾanu/ʾāni* for sg. masc./fem.
	ʾaddēš *ʾakam*	‘how much, how many?’	
	wēn, fēn	‘where?’	
	ʾēmta	‘when?’	
Expressions (greetings, exclamations)	*ṣaḥḥa w-ʿāfiya*	‘in health [at a meal]!’	
	ʾaḷḷa yxallīk	‘thank you [lit. may God preserve you]’	
	pardon	‘pardon!’	< French
	mersi	‘thanks!’	< French

Bibliographical references

Al-Wer, Enam. 2004. “Variability reproduced: A variationist view of the [ḏ]/[ḍ] opposition in modern Arabic dialects”. *Approaches to Arabic dialects: A collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday*, ed. Martine Haak, Rudolf de Jong, and Kees Versteegh, 21–31. Leiden: E.J. Brill.

Barthélemy, Adrien 1935–1954. *Dictionnaire arabe-français: Dialectes de Syrie: Alep, Damas, Liban, Jérusalem.* Paris: Geuthner.

Bauer, Leonhard. 1913. *Das palästinische Arabisch.* 3rd ed. Leipzig: J.C. Hinrichs’sche Buchhandlung.

——. 1957. *Wörterbuch der arabischen Umgangssprache: Deutsch-Arabisch.* 2nd ed. with Anton Spitaler. Wiesbaden: O. Harrassowitz.

Ben-Ze’ev, Y. 1945, 1948. *Colloquial Arabic.* (2 vols.) Jerusalem: Ahiassaf. [In Hebrew.]

Blanc, Haim. 1964. *Communal dialects in Baghdad.* Cambridge, Mass: Harvard University Press.

Denizeau, Claude. 1960. *Dictionnaire des parlers arabes de Syrie, Liban et Palestine.* Paris: Maisonneuve.

Elihay, Yohanan 1977. *Millon ʿivri-ʿaravi la-safa ha-ʿarvit ham-medubberet* [*Hebrew – Arabic Dictionary of Colloquial Arabic*]. Jerusalem: Yanets.

Kapliwatzky, Yohanan 1940, 1968. *Colloquial Arabic.* Jerusalem: R. Mass.

Levin, Aryeh. 1994. *A grammar of the Arabic dialects of Jerusalem.* Jerusalem: Magnes Press. [In Hebrew.]

Löhr, Max. 1905. *Der vulgärarabische Dialekt von Jerusalem nebst Texten und Wörterverzeichniss.* Giessen: Toepelmann.

Othman, O. and T. Neu. 2002. *Qamuus il-quds: A dictionary of the spoken Arabic of Jerusalem (English-Arabic).* Jerusalem: O. Othman.

Piamenta, Moshe. 1964. *The use of tenses, aspects, and moods in the Arabic of Jerusalem.* Jerusalem: Bureau of Adviser on Arab Affairs.

——. 1966. *Studies in the syntax of Palestinian Arabic: Simple verb forms in subordinate and main clauses of complex sentences.* Jerusalem: Israel Oriental Society.

——. 1968. *Dabber ʿArvit* [Speak Arabic]. Tel Aviv: Maʿariv. [In Hebrew.]

——. 1970. “The uses of fossilized and semi-fossilized verbs in the Jerusalem Arabic dialect”. *Memorial Adam-Noah in memory of Dr. A. Noah Brāʿōn*, 470–492 Jerusalem. [In Hebrew.]

——. 1979. “Jerusalem sub-standard Arabic: A linguistic analysis of an idiolect”. *Jerusalem Studies in Arabic and Islam* 1.263–285.

——. 1979. “Jerusalem Arabic lexicon.” *Arabica* 26.229–266.

——. 1981. “Selected syntactic phenomena of Jerusalem Arabic narrative style in 1900”. *Studies in Judaism and Islam presented to Shelomo Dov Goitein on the occasion of his 80th birthday*, ed. Shelomo Morag, Issachar Ben-Ami, and Norman A. Stillman, 203–230. Jerusalem: Magnes Press.

——. 1991. “Hypothetical sentences in Jerusalem Arabic”. *Semitic studies in honor of Wolf Leslau*, ed. Alan S. Kaye, II, 1203–1219. Wiesbaden: O. Harrassowitz.

——. 1992. “Note on the decay of Jerusalem Judeo-Arabic under the impact of sociopolitical transformation”. *Asian and African Studies* 26.1–88.

——. 2000. *Jewish life in Arabic language and Jerusalem Arabic in communal perspective.* Leiden: E.J. Brill.

Rice, Frank A. and Majed F. Said. 1953. *Jerusalem Arabic.* Washington, D.C.: Georgetown University Press.

Spoer, Hans H. and E. Nasrallah Haddad. 1909. *Manual of Palestinian Arabic for self instruction.* Jerusalem: Syrisches Waisenhaus.

Judith Rosenhouse
(Swantech Ltd.)

Jibbala → Modern South Arabian

Jīm

Jīm is the name of the fifth letter of the Arabic alphabet, usually transcribed as *ǧ* or *j* (in English publications), sometimes as *d̲j̲* (e.g. in the *Encyclopaedia of Islam*). The letter *jīm* in the Arabic alphabet is used for the phoneme that is pronounced in the tradition of 'correct' standard written Arabic, both in the Arab countries and in Arabist practice elsewhere, as [dʒ], i.e. a voiced lenis palato-alveolar affricate. There are several reflexes of the Standard Arabic phoneme in the modern dialects:

i. [dʒ] is standard in the majority of the dialects of Iraq, in rural dialects of Syria, Palestine, and Jordan, in many dialects of Yemen, and in several dialects of northern Algeria.
ii. This phoneme is realized as a post-alveolar/palato-alveolar fricative [ʒ], mainly in most urban dialects of Syria, with the exception of Aleppo; in Palestine Arabic, except the Muslim dialect of Jerusalem, where it is [dʒ]); partially in Jordan, the whole of Lebanon, except to the north of the Biqāʿ, and in most but not all Maghreb dialects.
iii. It is realized as a strongly palatalized [gʸ] or [dʸ], for instance in some Bedouin dialects of Arabia, in Middle and Upper Egypt (see Behnstedt and Woidich 1985–1999:II, maps 10–14), and in the Sudan; Gairdner (1925:23) believed that this might have been the original pronunciation.
iv. It is realized as a voiced velar plosive [g], mainly in Cairo Arabic and in the central and northeastern Delta, as well as in the area of Fayyūm and Bani Swēf; it is also realized this way elsewhere, for instance in various Bedouin dialects of Central Arabia and in some Yemenite dialects.
v. It is realized as a palatal semivowel or sonorant [j], mainly in many dialects of the Gulf, including southern Iraq, also in some dialects of North Arabia, in which it is partly lexically conditioned, with some irregular free variation of [j] and [dʒ], e.g. *diyāya/dayāya* for *dijāja/dajāja* 'hen', *yaryūr/jarjūr* 'shark' (→ Bahraini Arabic).
vi. Very seldom it is realized as [z], merging with /z/, for instance in Jewish Arabic dialects of Morocco and Algeria.
vii. It is realized as a voiceless affricate [tʃ], e.g. in Palmyra and in some villages in the Anti-Lebanon.
viii. It is realized as a dental-alveolar affricate [ts], e.g. in Sukhne and two regions north of Damascus in Syria.
ix. It is realized as a dental plosive [d], e.g. in some dialects of the western Nile Delta in front of liquids and nasals as well as in Upper Egypt, and originally in what used to be Nubia (Gairdner 1925:23–24; Behnstedt and Woidich 1985–1999:I, 70).

Since these variants are not phonemically distinctive (Cantineau 1946), apart from those dialects in which /q/ has shifted to [g], and apart from those dialects in which there has been a partial merger of /j/ with /y/, they also occur in the oral realization of modern written Arabic by native speakers of the particular dialects, who may apply the [dʒ] variant, usually after some kind of training, only for special purposes, for instance for the recitation of the *Qurʾān*. Therefore, there are variants like [gamal], [gʸamal], [dʸamal], [dʒamal], [tʃamal], and [jamal] all meaning 'camel'. In final position, /j/ is either voiced or voiceless, i.e., there is free variation.

Assimilations and dissimilations in which the *jīm* takes part include [dʒ] > [ʒ] > [ʃ] before /t/, e.g. *ijtamaʿū* > [iʃtamaʕuː] 'they gathered'; between two lexemes or in sandhi: *-j š-* > *-šš-*; *-j t-* > *-tt-*; *j z-* > *-zz-*; *-ḏ j-* > *-jj-*. Assimilation and dissimilation at a distance occur in *jazzār* > *žazzār* > *zazzār* and *dazzār*, also *gazzār*, 'butcher' in some Tunisian and Algerian dialects; *žūž* < *zūž* < *zūj* < *zawj* 'two' (Cantineau 1960:60–62). Another example is *ed-dzāir* < *al-jazāʾir* 'Algiers'. Typically Algerian is *bezzāf* < *bi-l-jizāf* 'a lot, much'. Remarkably, a shift *iyy* > *ijj* is reported about the pre-Islamic dialects of Ṭayyiʾ, Tamīm, and partially ʾAsad in the Arabian Peninsula, e.g. *iyyal* > *ijjal* 'mountain goat, stag'; this phenomenon is called *ʿajʿaja* (Fleisch 1990:I, 78).

It is usually taken for granted that the Proto-Semitic occlusive, postpalatal, voiced phoneme /g/ was pronounced as a voiced velar plosive [g], but as a matter of fact, /g/ may have had fronted or palatalized allophones, first of all [g^y], [d^y], and affricated [dʒ], which were phonemically nondistinctive (as they are today in modern Arabic dialects), and which could be typical of some Proto-Semitic dialects. There is no need to assume, and actually there is no proof, that the fronted allophones of /g/ appeared within Semitic for the first time in Proto-Arabic or in some pre-Classical Arabic dialects. There can be no doubt that different allophones, i.e. different realizations of the phoneme later written with the letter *jīm*, must have existed in some pre-Classical Arabic dialects. According to one interpretation (Fleisch 1990:I, 228) of the description given by Sībawayhi (*Kitāb* II, 453–454), the 'correct' pronunciation in his period was [g^y], i.e. a voiced medio-palatal dorsal plosive with obvious palatalization, rather than [d^y], [dʒ], or [ʒ], while the pronunciation as [g] was mentioned but disapproved of.

As far as the relative chronology of the historical phonology of Arabic is concerned, Semitic /g/ was fronted in Arabic before the occurrence of the shift /q/ > [g] (Blanc 1969), which took place in mainly Bedouin dialects (→ *qāf*). The particulars of the conditioning of the fronting of /g/ are not clear (see Cantineau 1960; Martinet 1959; Blanc 1969), but adjacent front vowels must have played at least some role, and the shift /q/ > [g] must have contributed as well. The evolution was as follows: /g/ > [g^y] > [d^y], and then in some dialects [d^y] > [dʒ], in another group [dʒ] > [ʒ], in still another group [d^y] > [y], which means a merger with the phoneme /y/, and finally [d^y] > [d], which means a merger with the phoneme /d/. It is impossible to accept Cowan's (1960:50) reconstruction of Proto-Colloquial Arabic *jīm* as [ʒ] (supported by Kaye 1972:62–63), since it is based on very incomplete data and is methodologically unacceptable. It has been suggested that the Cairene Arabic pronunciation as a voiced velar plosive [g], but often pronounced as a palatal plosive in the environment of /i/ and /ī/, does not continue the Proto-Semitic pronunciation but rather goes back to a secondary and rather recent (late 18th century, according to Blanc) reintroduction of [g] going back to [g^y]. Woidich (→ Cairene Arabic; see also Hary 1996) states that [g] must have been prevalent there in the Middle Ages. The phonetic development within the varieties of Arabic was similar to what is found, for instance, in Romance and Germanic languages, e.g. English *yard/garden* and German *Garten*; English *day* and German *Tag*; English *say* and German *sag-en*; Italian *gamba*, French *jambe* [ʒɑ̃b]; Italian *giardino* [dʒardino], French *jardin* [ʒardɛ̃] 'garden'; Latin *gens*, genitive *gentis*, Italian *gente* [dʒɛnte], French *gens* [ʒɑ̃], etc. Therefore, it is rather astonishing that some Arabists emphasized an alleged idiosyncrasy of Arabic in this case. The evolution of the pronunciation of /j/ (here it is better to use this symbol rather than /g/ because of the phonemic conflict within *gāl* dialects, i.e. the dialects in which /q/ resulted in [g], and because of the common pronunciation as [g] in such a sociolinguistically important and influential dialect as Cairene Arabic) is still going on due to internal dynamics of the particular phonemic systems, as well as to interference ('borrowing') between different dialects. In Alexandria, for instance, /q/ has shifted to [g], and rather recently, under the influence of Cairo prestige pronunciation, /j/ has shifted to [g], which means that the two phonemes have fully merged.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCE

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. Ed. Hartwig Derenbourg. 2 vols. Paris: Imprimerie Nationale, 1881–1889.

SECONDARY SOURCES

Al-Ani, Salman H. 1970. *Arabic phonology*. The Hague: Mouton.

Behnstedt, Peter and Manfred Woidich. 1985–1999. *Die ägyptisch-arabischen Dialekte*. 5 vols. Wiesbaden: L. Reichert.

Blanc, Haim. 1969. "The fronting of Semitic G and the *qāl - gāl* dialect split in Arabic". *Proceedings of the International Conference on Semitic Studies… Jerusalem 1965*, 7–37. Jerusalem: Israel Academy of Sciences.

Cantineau, Jean. 1941. *Cours de phonétique arabe*. Algiers. (Repr. with *Notions générales de phonétique et de phonologie*, Paris 1960; both included also in *Etudes de linguistique arabe: Mémorial Jean Cantineau*, 5–164. Paris: C. Klincksieck, 1960.)

——. 1946. "Esquisse d'une phonologie de l'arabe classique". *Bulletin de la Societé de Linguistique de Paris* 43.93–140. (Repr., *Etudes de linguistique arabe: Mémorial Jean Cantineau*, 165–294. Paris: C. Klincksieck, 1960.)

Cowan, William G. 1960. *A reconstruction of Proto-Colloquial Arabic*. Ph.D. diss., Cornell University.

Denz, Adolf. 1964. "Q > g, k > č und g > ǧ als Beispiele strukturbedingten Lautwandels im Irak-Arabischen". *Münchener Sprachwissenschaftliche Studien* 16.13–16.

Fleisch, Henri. 1990. *Traité de philologie arabe*. 2 vols. Repr., Beirut: Dar el-Machreq.

Gairdner, William Henry Temple. 1925. *The phonetics of Arabic*. Oxford: Oxford University Press.
Hary, Benjamin. 1996. "The *ǧīm/gīm* in colloquial urban Egyptian Arabic". *Israel Oriental Studies* 16.153–168.
Johnstone, T.M. "The sound change j > y in the Arabic dialects of peninsular Arabia". *Bulletin of the School of Oriental and African Studies* 28.233–241.
Kaye, Alan S. 1972. "Arabic /žiim/: A synchronic and diachronic study". *Linguistics* 79.31–72.
Martinet, André. 1959. "La palatalisation spontanée de g en arabe". *Bulletin de la Societé Linguistique de Paris* 54.90–102.
Marçais, William and Henri Fleisch. "*D̲j̲īm*". *Encyclopaedia of Islam*, 543–544. 2nd ed. Leiden: E.J. Brill.
Mitchell, Terence F. 1990–1993. *Pronouncing Arabic*. 2 vols. Oxford: Oxford University Press.

Andrzej Zaborski (University of Cracow)

Jinās

Jinās is one of the names Arab theorists use to designate a rhetorical figure based, in its most general meaning, on the repetition of an identical or similar sequence of sounds in the same verse or colon. Both definition and name vary according to author and period; moreover, Arab theorists have described *jinās* by introducing more and more complex and subtle subdivisions. The possible translations of these technical terms (*jinās, tajnīs, mujānasa, mujānas, tajānus*) are therefore equally different, e.g. '→ paronomasia' (the most general), 'alliteration', or 'pun'.

The kinds of *jinās* early authorities seem to have had in mind are pun and *figura etymologica*; these are combined by Qudāma (d. ca. 320/932) in his definition as follows: *maʿānin mutaġāyira qad ištarakat fī lafḏa wāḥida wa-ʾalfāḏ mutajānisa muštaqqa* 'different meanings joined in the same word, and cognate words from the point of view of their consonantal basis' (*Naqd* 92, s.v. *al-muṭābaq wa-l-mujānas*). *Figura etymologica* is represented, for instance, by the definition of *tajnīs* in Ibn al-Muʿtazz (d. 296/908) and the example he quotes: *wa-ʾaslamtu maʿa sulaymāna* (Q. 27/44; *Badīʿ* 25). Later on, *figura etymologica* constitutes a separate and minor category, called *jinās al-ištiqāq*. Pun is represented in Ṯaʿlab's (d. 291/904) definition: *takrīr al-lafḏ bi-maʿnayayn muxtalifayn* 'repetition of the [same] word in two different meanings' (*Qawāʿid* 64–67, s.v. *al-muṭābaq*).

The origin of the linguistic interest in this phenomenon seems to be connected with the lexicological side of early linguistic thinking, in this case the question of → *ištiqāq* (the possibility of producing words with similar meanings from the same consonantal basis) and the question of → *muštarak* (homonymous polysemic words). Ibn al-Muʿtazz (*Badīʿ* 25) and al-ʿAskarī (d. 395/1004; *Ṣin.* 330), in their definition of the *tajnīs*, quote as reference the *Kitāb al-ʾajnās* of al-ʾAṣmaʿī (d. 213/828), which, judging by the passage reproduced by as-Suyūṭī (d. 911/1505; *Muzhir* I, 372–373), must have been a collection of homonymous polysemic words.

Of these two aspects of *jinās*, the latter is well represented in ancient poetry and abundantly quoted, for instance, by Ibn al-Muʿtazz (*Badīʿ* 25–28) and al-ʿAskarī (*Ṣin.* 330–337), but ancient poetry also presents other kinds of phonetic repetitions (such as the often quoted hemistich of al-ʾAʿšā: *šāwin mišallun šalūlun šulšulun šawilu* '[a boy] who roasts [meat]; light, active, agile'. These words, which describe a drinking companion, were variously interpreted, for instance by al-ʿAskarī (*Ṣin.* 344), al-Ḥalabī (d. 725/1325; *Ḥusn* 199), and Ibn Ḥijja (d. 837/1434; *Xizāna* I, 378). Such phonetic repetitions, as Heinrichs says, "found their way into the later *tad̲j̲nīs* category only in part" (1998:68).

The former aspect, on the contrary, came to constitute the type par excellence of *jinās*. Ibn al-ʾAṯīr (d. 637/1249; *Maṯal* I, 56–58) emphasizes the importance of the knowledge of the 'common words' for the sake of their use in *jinās* (especially the *jinās tāmm*, which in his view is the only true *tajnīs*, cf. *Maṯal* I, 342). Ancient poetry presents examples of repetitions of the same word, but often with the same meaning; early authorities do not insist on a difference in meaning between the two terms (*rukn*) of a *jinās*, and the examples they present do not always have this characteristic (Jundī 1954:196). For later theorists, however, this feature becomes a means to determine whether a line presents this rhetorical figure (Ibn al-ʾAṯīr, *Maṯal* I, 342, III, 198; as-Suyūṭī, *Janā* 74, concerning Q. 30/55).

From the rhetorical point of view, *jinās* (s.v. *tajnīs*) belongs to the five kinds of *badīʿ* 'ornaments of style' listed by Ibn al-Muʿtazz, who states that modern poets multiplied its use but did not invent it. According to a tradition related by Ibn Rašīq (d. ca. 456/1063–1064; *ʿUmda* I, 331), the term *tajnīs* was unfamiliar to ancient poets, and Ibn al-Muʿtazz was the first to use it. The definitions one finds in various works are multifarious, and it was easy for later authors, such as aṣ-Ṣafadī (d. 764/1363; *Jinān* 36–42), to criticize the lack of accuracy, consistency, or completeness of earlier authors. Scholars like Ibn Rašīq (*ʿUmda* I, 321) or Ibn Ḥijja (*Xizāna* I, 384–385) prefer to define each subcategory separately rather than provide a general definition. The final setting distinguishes two major kinds of *jinās*: the 'complete one' (*tāmm*) and the rest.

In the *jinās tāmm* the two words are identical in consonantal and vocalic frame, and have different meanings. The article, the vowel, or the *sukūn* of the final consonant does not count. This kind receives different designations (e.g. aṣ-Ṣafadī, *Jinān* 45; as-Suyūṭī, *Janā* 73; cf. Jundī 1954:64). Additional subcategories take into account whether or not the two words belong to the same grammatical class (noun, verb, or preposition) as in the frequent instances in which one *rukn* is a verb and the other a proper noun: *yaḥyā* 'he lives' and *yaḥyā* as a proper name (Farḥāt [d. 1145/1732], *Bulūġ* 77). The *jinās tāmm* can be *murakkab* 'composed' when one of the two *rukn* consists of two distinct words, e.g. *ḏā hiba* 'having a gift' and *ḏāhiba* 'leaving' (as-Suyūṭī, *Janā* 126; aṣ-Ṣafadī, *Jinān* 53). A further distinction concerns the case in which the composed term consists of two words, or of one word and part of another; or in which both terms are composed.

In the 'non-*tāmm*' kind the two words are not identical and the subcategories are classified according to the kind of dissimilarity. In the *jinās muḥarraf* the difference is in the vowels: *qidr/qadr* (as-Suyūṭī, *Janā* 163) or in vowel and *sukūn: šiʿr/šaʿar* (as-Suyūṭī, *Janā* 171). In the *jinās nāqiṣ*, the difference concerns the number of consonants: one of the two words presents one or two additional consonants: *dāʾ/dawāʾ* (as-Suyūṭī, *Janā* 246), *ṣafā/ṣafāʾiḥ* (as-Suyūṭī, *Janā* 252), which can occur at the beginning, the middle, or the end of the word. In the *jinās maqlūb*, the difference lies in variations in the arrangement of consonants: *fatḥ/ḥatf* (as-Suyūṭī, *Janā* 198). In the *jinās muḍāriʿ*, it lies in the quality of the consonants: one of the consonants is not the same in the two words, but it may have the same point of articulation. Each kind and subcategory receives its own designation, which is often different from one author to another (cf. Jundī 1954; Heinrichs 1998).

Modern authors point out the euphonic effect of this rhetorical ornament (Jundī 1954:29, 31). Bencheikh (1975:186–197) shows convincingly its poetical impact on Abbasid poetry. But the medieval theorists who expressed their opinion were more concerned with the wordplay implied by *jinās* than with its sonorous effect. They considered the *jinās tāmm* the most perfect variety, and Ibn Ḥijja (*Xizāna* I, 390–391), who finds *jinās* a heavy and feeble stylistic device, argues that it is possible to lighten it by using *tawriya* 'double entendre', because in this way the double meaning is restricted to only one *rukn*. Moreover, later subdivisions classify among the types of *jinās* also *jinās al-xaṭṭ* (e.g. as-Suyūṭī, *Janā* 180; al-Ḥalabī, *Ḥusn* 192), in which the similarity between the two words lies in the writing, not in the sounds: *ṯāʾir/bātir* (as-Suyūṭī, *Janā* 187). In the 'virtual' (*maʿnawī*) *jinās* (as-Suyūṭī, *Janā* 277; al-Ḥalabī, *Ḥusn* 197), the author elides or hints at one of the two terms of the *jinās*, which is easily understood by the audience or the readers, for instance in the case of a segment of the proper name of a well-known person.

Jinās is a very popular device in Arabic rhetoric. Thus, all treatises dealing with poetical questions have a chapter on this topic; some monographs, too, are devoted to *jinās* (for a quite considerable list of these, see Heinrichs 1998).

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

ʿAskarī, *Ṣin.* = ʾAbū Hilāl al-Ḥasan ibn ʿAbdallāh al-ʿAskarī, *Kitāb aṣ-ṣināʿatayn*. Ed. ʿAlī Muḥammad al-Bijāwī and Muḥammad ʾAbū l-Faḍl ʾIbrāhīm. Cairo: ʿĪsā al-Bābī al-Ḥalabī, 1971.

Farḥāt, *Bulūġ* = Jarmānūs Farḥāt, *Bulūġ al-ʾarab fī ʿilm al-ʾadab: ʿIlm al-jinās*. Ed. ʾInʿām Fawwāl. Beirut: Dār aš-Šarq, 1990.

Ḥalabī, *Ḥusn* = Šihāb ad-Dīn Maḥmūd ibn Sulaymān al-Ḥalabī, *Ḥusn at-tawassul ʾilā ṣināʿat at-tarassul*. Ed. ʾAkram ʿUṯmān Yūsuf. Baghdad: Dār al-Ḥurriyya, 1980.

Ibn al-ʾAṯīr, *Maṯal* = ʾAbū l-Fatḥ Ḍiyāʾ ad-Dīn Naṣrallāh Ibn al-ʾAṯīr, *al-Maṯal as-sāʾir fī ʾadab al-kātib wa-š-šāʾir*. Ed. ʾAḥmad al-Ḥūfī and Badawī Ṭabāna. 4 vols. Cairo: Dār an-Nahḍa, n.d.

Ibn Ḥijja, *Xizāna* = ʾAbū Bakr ibn ʿAlī Ibn Ḥijja al-Ḥamawī, *Xizānat al-ʾadab wa-ġāyat al-ʾarab*. Ed. Kawkab Diyāb. 4 vols. Beirut: Dār Ṣādir, 2001.

Ibn al-Muʿtazz, *Badīʿ* = ʾAbū l-ʿAbbās ʿAbdallāh Ibn al-Muʿtazz, *Kitāb al-badīʿ*. Ed. Ignatius Kratchkovsky. London: Luzac, 1935.

Ibn Rašīq, *ʿUmda* = ʾAbū ʿAlī al-Ḥasan ibn Rašīq al-Qayrawānī, *al-ʿUmda fī maḥāsin aš-šiʿr wa-ʾādābi-hi wa-naqdi-hi*. Ed. Muḥammad Muḥyī d-Dīn ʿAbd al-Ḥamīd. 2 vols. Beirut: Dār al-Jīl, 1972.

Qudāma, *Naqd* = ʾAbū l-Faraj Qudāma ibn Jaʿfar al-Kātib al-Baġdādī, *Kitāb naqd aš-šiʿr*. Ed. Seeger Adrianus Bonebakker. Leiden: E.J. Brill, 1956.

Ṣafadī, *Jinān* = ʾAbū ṣ-Ṣafāʾ Xalīl ibn ʾAybak aṣ-Ṣafadī, *Jinān al-jinās*. Ed. Samīr Ḥusayn Ḥalabī. Beirut: Dār al-Kutub al-ʿIlmiyya, 1987.

Suyūṭī, *Janā* = Jalāl ad-Dīn ʾAbū l-Faḍl ʿAbd ar-Raḥmān ibn ʾAbī Bakr as-Suyūṭī, *Janā l-jinās*. Ed. Muḥammad ʿAlī Rizq al-Xafājī. N.p.: ad-Dār al-Fanniyya, n.d. [1986].

——, *Muzhir* = Jalāl ad-Dīn ʾAbū l-Faḍl ʿAbd ar-Raḥmān ibn ʾAbī Bakr as-Suyūṭī, *al-Muzhir fī ʿulūm al-luġa wa-ʾanwāʿi-hā*. Ed. Muḥammad ʾAḥmad Jād al-Mawlā, ʿAlī Muḥammad al-Bijāwī, and Muḥammad ʾAbū l-Faḍl ʾIbrāhīm. 2 vols. Cairo: ʿĪsā al-Bābī al-Ḥalabī, n.d.

Ṯaʿlab, *Qawāʾid* = ʾAbū l-ʿAbbās ʾAḥmad ibn Yaḥyā Ṯaʿlab, *Qawāʿid aš-šiʿr*. Ed. Ramaḍān ʿAbd at-Tawwāb. Cairo: Dār al-Maʿrifa, 1966.

Secondary sources

Bencheikh, Djamaleddin. 1975. *Poétique arabe*. Paris: Anthropos.

Heinrichs, Wolfhart P. 1998. "*Tad̲jnīs*". *Encyclopaedia of Islam* X, 67–70. 2nd ed. Leiden: E.J. Brill.

Jundī, ʿAlī al-. 1954. *Fann al-jinās: Balāġa ʾadab naqd*. Cairo: Dār al-Fikr al-ʿArabī, n.d. [1954].

Omar, Ahmed Mokhtar. 1993. "Early Arabic lexicons of homophonic words". *The Arabist: Budapest Studies in Arabic* 6–7.3–11.

Lidia Bettini (University of Firenze)

Jordan

1. Historical introduction

This entry refers to modern-day Jordan, as established in 1921 by the British. Initially it was a princedom designated to Emir ('prince'), later King, Abdullah. During that stage, between 1921 and 1946, it was known as Trans-Jordan. It was also called 'the East Bank', referring to the River Jordan. The second stage began after the annexation of the Palestinian territories, or 'the West Bank', that were left unoccupied by the State of Israel in 1948. In 1950, the name changed to the Hashemite Kingdom of Jordan. In the wake of the 1948 war, Palestinian refugees significantly increased the population of Trans-Jordan, especially after the annexation of the West Bank. The 1967 war resulted in the loss of the West Bank, causing more Palestinians to take refuge in Jordan. 'Trans-Jordan' is used in this entry to designate the period predating 1946; thereafter, this country is referred to as 'Jordan'.

The agreement between Emir Abdullah and Britain in 1923 recognized an autonomous administration in Trans-Jordan, separate from that of Palestine, despite the continued general supervision of the British high commissioner in Palestine (Vatikiotis 1967:xi, 45–48). Thus, as of the early 1920s, Palestine and Trans-Jordan began to develop geopolitically in rather different ways. The populations of these two regions have had separate social, political, and economic environments that have resulted in the development of relatively separate speech patterns. However, it would be extremely difficult to examine the linguistic situation in today's Jordan in isolation from its larger geographical, demographic, and linguistic Arab context.

In 1921, the population of Trans-Jordan was estimated by the two high-ranking British officers Somerset and Peake at 230,000 (Musa 1972:126). In 1946, the population figures were about 433,659 (Al-Madi and Musa 1959:448; Vatikiotis 1967:9; Patai 1958:10).

Ecologically, this population included three groups: nomadic and seminomadic tribes (Bedouin), sedentary villagers (*fallāḥīn* 'subsistence farmers'), and town dwellers (Glubb 1938; Epstein 1938; Dann 1984:4). Ethnically, the overwhelming majority of the population were Arab. However, at the turn of the 19th century and into the early 20th century, non-Arabic-speaking Circassian and Chechen refugees from the Caucasus were settled by the Ottoman sultan Abdul-Hamid in and around Amman, Jerash, Naʿīr, Ṣwayliḥ, Wadi al-Sīr, al-Azraq Oasis, Sukhnah, and al-Zarqa (Hourani 1947:59; Patai 1958:20–21; Kazziha 1972:24; Dann 1984:4; Jaimoukha 2001:106).

Because of high illiteracy among Trans-Jordanians in the early 20th century, their relative isolation, and the lack of effective means of communication, divergent Trans-Jordanian dialects resulted in different localities. Thus, the three ecological groups each developed their own characteristic 'social' dialects. Regional dialects also emerged in the north (Irbid region), in the central part of the country (aṣ-Ṣalṭ), and in the south (Maʿān, Karak, and Aqaba). Social dialects intersect with regional ones, producing linguistic varieties that combine these two parameters.

The emergence of Jordan in 1921 as an autonomous political and administrative entity created a need for skilled manpower. Consequently, about 10,000 skilled, educated civil servants from Palestine moved to Amman long before 1948 (Plascov 1981:33). Merchant families as well as professionals from Syria settled in Amman, Jerash, Maʿān, and Irbid in the 1920s and 1930s (Musa 1972:73; Aruri 1972:34). In addition, a number of Syrian nationalists took refuge in Jordan as a result of their resistance to the French domination of Syria in the early 1920s (Musa 1972:73). These new immigrants contributed to the diverse nature of the population and, consequently, to the linguistic make-up of the nascent Jordanian entity. It is reasonable to speculate that Trans-Jordanian dialects started to change in the 1920s and 1930s as a result of the arrival of Emir Abdullah in 1921 and his retinues from the Ḥijāz region in the western parts of the Arabian Peninsula, and the arrival of Palestinian and Syrian civil servants.

2. THE LINGUISTIC SITUATION IN JORDAN

Records of the linguistic situation in Trans-Jordan in the early part of the 20th century are not available. Consequently, it is hard to state with certainty which dialects dominated then: the indigenous rural Trans-Jordanian dialects of the small populations of Amman, Irbid, aṣ-Ṣalṭ, and Karak; or the Syrian and/or Palestinian dialects introduced by the socially prestigious newcomers; or the Ḥijāz dialect of the royal Hashemite family; or the dialects of the Bedouin, the monarch's favored groups, who served in the armed and security forces and in other governmental agencies. Four possible linguistic outcomes of this situation could have evolved. First, there might have been a linguistic shift toward the newly introduced Syrian and Palestinian dialects, the language varieties of the government employees with high social status. The indigenous Trans-Jordanian population might have emulated these prestigious linguistic varieties, thus starting a trend toward linguistic change. Second, there could have been a shift by the Syrian and Palestinian newcomers toward indigenous Trans-Jordanian linguistic forms as a way of identifying with the indigenous population. Third, the new Syrian and Palestinian civil servants could have adopted the dialect of the royal family, i.e. the Ḥijāz dialect spoken by Emir Abdullah and his retinue. Fourth, a polydialectal situation might have evolved.

It is important to consider women and their possible impact on the choice of a spoken language variety. Because of social restrictions imposed on females in Trans-Jordanian society at that time, women of the new Syrian and Palestinian arrivals most likely preserved their dialectal features. It is reasonable to hypothesize that in the nascent evolving society, Trans-Jordanian women who came in contact with Syrian and Palestinian women were inclined to adopt some dialectal features of these newcomers, due to their higher social status.

The annexation of the Palestinian areas in 1949 by the Trans-Jordanian monarch caused profound changes, not only in the population numbers, social structure, and political and economic life of the kingdom, but also in its linguistic composition. In addition, the merging of the two territories, Trans-Jordan and the West Bank, resulted in the increase of urban ratios and a decrease in nomadic life patterns (Aruri 1972:49; Patai 1958:50). The demographic composition of many Jordanian towns, such as Amman, Irbid, aṣ-Ṣalṭ, and Karak, started to change. The Palestinian population movement into these towns was accompanied by a voluntary migration from the Jordanian countryside to developing towns that offered better living conditions, economic advancement, new employment opportunities, and better educational facilities.

The majority of the population in present-day Jordan who identify themselves as Arabs are perhaps descendants of the waves of Arab tribes that migrated to the Greater Syria regions after the Islamic conquest of these lands in 637 C.E. Some are perhaps descendants of the Ghassanids, the indigenous Arabic-speaking tribes, or speakers of other Semitic languages such as Aramaic or Syriac who populated the southern parts of Greater Syria (today's Jordan) during the Byzantine control that preceded the Islamic conquest, thus maintaining continuity of Arabic in Jordan over these centuries.

The Arab majority aside, at the turn of the 19th century and in the early 20th century, the Ottoman Empire settled 3,500 Circassian and Chechen refugees in agricultural villages in various parts of Trans-Jordan. These refugees had fled their original homelands pursuant to the Russian Caucasian Wars that ended in 1864. Initially, these displaced non-Arabic-speaking groups lived separately from the indigenous Arab tribes populating those regions. As the numbers of such immigrants were small and they were then scattered in isolated communities, their linguistic impact on Jordanian Arabic was negligible. Over the years, however, these groups began to assimilate in their new adoptive country. Emir Abdullah was instrumental in integrating them into his new government, especially in senior positions in the army, the security forces, and other governmental agencies. The new and younger Circassian and Chechen generations assimilated into Jordanian society and adopted Arabic. The number of Circassians at present is estimated to range between 20,000 and 100,000. Modest efforts are being made to stimulate interest in Circassian language instruction in Jordan (Jaimoukha 2001:107–109). By way of contrast, the Chechens, estimated to number around 9,000 people, have had more success at maintaining their native language for 'different domains and situations'. This could be attributed to many factors: positive attitudes toward the Chechen language, social isolation of the community, and resistance to interethnic marriage (Dweik 2000:184).

Armenians, a much smaller minority, estimated at present to be around 4,000 people, is another group that sought refuge in Trans-Jordan, principally due to their forced migration after World War I as a result of the pogroms perpetrated against them by the Ottoman Turks. These new immigrants, especially the elderly, initially formed a close-knit community, maintaining their language at home, in churches, clubs, and so on. Employable and professional members of this minority group, however, had to learn Arabic to function in the new society, thus causing a shift toward Arabic. Eventually, this community's younger generations who continued to live in Jordan learned Arabic through educational institutions, causing a gradual loss of Armenian (Al-Khatib 2001:153–177). However, recent contacts with the Republic of Armenia have revived interest in ethnic affiliation and opened opportunities for higher education in Armenia for members of this community, thus fostering the maintenance of Armenian among Armenians in Jordan.

Few Turkish words entered into Jordanian Arabic, partly because Jordan was the backwater of the Ottoman territories, and partly because of the absence of major urban centers in the country to justify the presence of Turkish-speaking communities or institutions (→ Turkish loanwords). While the Ottomans maintained a limited police force and army garrisons and other administrative offices in some towns, Trans-Jordan was divided into sectors that were administratively subordinate to larger urban centers, *sanjak*s 'districts' and *vilayet*s 'provinces' in Palestine and Syria. The Turkish words that filtered into Jordanian Arabic came via Turkish-speaking administrators, the military, and the police, or those few who were schooled in Damascus, Jerusalem, or Istanbul, where the official language of the government was Turkish.

Turkish words in Jordanian Arabic are often either military in nature, e.g. *ṣawārī* 'cavalry, mounted policemen', *ṭubči* 'artillery officer'; or administrative, e.g. *qāʾim-maqām* 'district administrative officer; or names of culture items. These military and administrative words are no longer used officially; they have been replaced by Arabic substitutes in recent years. Instead of *ṣawārī* and *ṭubči*, Jordanian Arabic uses *fursān* and *madfaʿi*, respectively. The institution of mounted policemen is almost extinct in the modern police force in Jordan. The words, nonetheless, linger in the dialects. The term *qāʾim-maqām*, on the other hand, was replaced by *mudīr nāḥiya* 'district administrator'. Names of culture items that continue to be used in today's Jordanian Arabic include *māsūra* 'pipe', *kafkīr* 'ladle', (*čafčīr* in rural or Bedouin pronunciation), *ʾūzi* 'young lamb meat', and a few others.

The British, through their mandate over Trans-Jordan, had a brief presence in the country that started with the foundation of the princedom in 1921 and ended in the early 1950s. They maintained airfields and military camps in Amman, Mafraq, and Aqaba, in rather isolated quarters. British officers occupied high positions in civil and military positions in the Jordanian government (Butros 1963:80–81). As these personnel were small in numbers and tended to operate through Arabic-speaking subordinates, their impact on

Jordanian Arabic was limited to a few lexical items. However, as contacts with the English-speaking West increased, more English words entered Arabic in Jordan via other channels (→ English loanwords). English vocabulary items span two areas, those representing Western cultural items relating to dress, food, furniture, sports, and architecture, as in *jākēt* 'jacket', *blūzih* 'blouse', *sandwīš* 'sandwich', *ṣandal* 'sandals', *sōfa* 'sofa', *baranda* 'veranda', *tanis* 'tennis', etc.; and those representing modern Western inventions introduced into the region in the early part of the 20th century, as in *ṭumbīl* 'automobile', *bank* 'bank', *karāj* 'garage', *alamunyu* 'aluminum', etc. (Al-Khatib and Farghal 1999). It is worth mentioning, however, that a push toward preservation of the purity of Arabic displaced some of the English borrowings. The word *ṭumbīl*, for example, lost ground to the Arabic coinage *sayyārah*. On the other hand, *bank*, *talifōn*, and *rādyu*, among many others, endured longer survival in the Arabic of Jordan and are still in stiff competition with the Arabic *maṣraf*, *hātif*, and *miḏyāʿ*, respectively.

3. INVESTIGATION OF JORDANIAN DIALECTS

Western travelers recorded impressions about the Arabic dialects in the region. Perhaps the earliest specific study of linguistic matters in the regions of Palestine and Trans-Jordan is Christie (1901) on the dialects of rural people in the middle Galilee region in Palestine. Studies conducted by Friedrich Binder (1939) and Jean Cantineau (1939) are very broad. Cantineau lumped together diverse and varied pronunciations of sedentary dialects in the Syria-Lebanon-Palestine area. In a later study, however, Cantineau (1946) included descriptive statements about dialectal features used in villages in the Horan areas (southern Syria), some of which are now inside the northern borders of Jordan. Charles Ferguson (1962) provided a survey of dialect studies in Greater Syria produced before World War II and up to 1962.

Ray Cleveland's (1963) study attempts to classify the Arabic dialects of 'Jordan' using phonological, morphological, and syntactic parameters to categorize them. Cleveland (1963:56) recognizes the presence of "the astonishing diversity of dialects in the spoken Arabic of the Hashemite kingdom of Trans-Jordan", which, he says, contributes to the difficulty of categorizing these dialects satisfactorily in one article. Nonetheless, he goes on to present "the most general outlines of the situation illustrated by a very limited number of dialectal characteristics". He categorizes dialects in this region into four different groups corresponding "though not precisely, to social and economic stratification and geographical zones" (Cleveland 1963:56). Despite its generalizations, his work provides an overall picture of the linguistic situation in Jordan and captures many of its realities.

In the 1980s and 1990s there was a rise in the study of Jordanian Arabic, primarily by linguistics practitioners or Jordanian doctoral students in linguistics in Western countries, mainly the United States. The phenomenon of → diglossia in Arabic generated much interest among such scholars. Several articles, dissertations, and monographs were written on the subject, including, among others, Mitchell (1978, 1980), Mitchell and El-Hassan (1994), El-Hassan (1977), Zughoul (1980), Hussein (1980), and Suleiman (1985). Much work concentrated on sociolinguistics and phono-logy and, to a lesser degree, on syntax. The sociolinguistic studies include, among others, Abdel-Jawad (1981, 1986), Sawaie (1986, 1987, 1994a, 1994b), Farghal and Shakir (1994), Al-Khatib (1988), Migdadi (2003), Saidat (2003), and Al-Wer (2003). The phonological studies include, for example, Palva (1976, 1980, 1986), Hussein and El-Ali (1989), El-Yasin (1982), Bani-Yasin and Owens (1984, 1987), Irshied (1984), Al-Sughayer (1990), Sakarna (1999, 2002), and Abu Abbas (2003). Some of these studies do not restrict themselves to the Arabic of indigenous Jordanians. Cleveland (1963), for example, interpreted Jordan to encompass Jordanian Arabic native speakers and Palestinian speakers who acquired Jordanian citizenship after 1949.

In his study of phonological and lexical variation in Amman, Abdel-Jawad (1981:68ff.) speaks of a trichotomy of ecological groups: urban, rural, and Bedouin. In defining the rural group (*fallāḥīn*) in Amman, he restricts this usage to "those who came originally from the countryside of Palestine" (Abdel-Jawad 1981:72). Bedouin represent "those who came from a tribal origin or from various parts of the

East Bank of Trans-Jordan and the Southern parts of Palestine". Abdel-Jawad's nomenclature lumps together disparate dialectal groups from various regions of Jordan and Palestine that differ in terms of dialectal features. It is well-known that within villages of the same geographical zone, different dialectal features exist (Bauer 1913).

The example of the ancient Roman city of Jerash illustrates this complex linguistic situation. The demographic situation in this city has undergone many changes since the early part of the 20th century. In the 1920s and the 1930s its population comprised three groups: Circassians (settled at the turn of the 19th and early 20th century), Syrian mercantile families and professionals (settled in the 1920s and the 1930s), and indigenous Trans-Jordanians, originally from surrounding villages. In the 1950s, the demographic composition of Jerash, however, began to change. Some Palestinians, displaced in 1948, took residence in the city. Many other Palestinians who settled in two refugee camps outside Jerash after the 1967 Arab-Israeli war slowly began to move into the city, lured by employment opportunities. This migratory trend was paralleled by indigenous Jordanian migration from surrounding villages for economic opportunities. The influx of these two groups was countered by emigration of long-time residents, the Circassian and Syrian families who started to emigrate to the capital city of Amman for its more lucrative economic environment. If this ebb and flow of Jerash population resulted in social change, a linguistic change was likely to have occurred as well.

Jordan can be divided into discrete zones geographically, socioeconomically, and ecologically. Within each zone we find subdialects, not one uniform language variety. These dialects, however, are not exclusive in their features; overlapping of many features is an observable phenomenon in the dialects of the region. By applying the socioeconomic and ecological sets of variables, we can speak of city vs. village dialects, sedentary vs. Bedouin dialects, high vs. low socioeconomic class dialects, and so on. The geographical parameters produce northern, central, or southern dialects.

Dialectal categorization has inherent problems because of aggregating villages, towns, cities, and Bedouin settlements, each of which may have distinct dialectal features. Neighboring Jordanian villages each have their own dialect(s) and distinct features that identify speakers' local affiliation (ʿUzayzī 1973:I,8–9). Since there are no clearly defined linguistic boundaries to characterize each dialect group, intradialectal features suggest that there is no absolute homogeneity in dialectal divisions. Nonetheless, phonological, lexical, and morphological features distinguish Jordanian groups and make their speakers identifiable on the basis of their speech. The comments below are restricted to the indigenous dialects of Trans-Jordan (the East Bank; see also → Jordanian Arabic: Amman). A profile of Arabic in Palestine (the West Bank) is provided elsewhere in this work (→ Palestinian Arabic).

While Jordanian speakers undoubtedly have many phonological features in common, they reveal phonological differences from which one can readily identify their regional affiliation (northern, central, or southern) or ecological (Bedouin, rural, and urban). Rural and urban features tend to overlap at times.

Phonologically, Arabic in Jordan comprises vocalic as well as consonantal variations. Vocalically, Jordanian dialects are characterized by the presence of the diphthongs [ay] and [aw] as in *ṣayf* 'summer' and *mawt* 'death'. In rural Jordanian dialects the final vowel is [+back] as in *gamoḥ* 'wheat', *gabor* 'grave', etc. (ʿUzayzī 1973:I,5–6).

With respect to consonants, the rural features involve distinction between [g] as a variant of the phonological variable (q), the Classical Arabic /q/. All indigenous Jordanian speakers use [g] in words such as *galb* 'heart', and *gad aysh* 'how much?'. Only a small number of the population, the Druze in some northern border villages and in the eastern desert (Umm ul-Quṭṭayn, and al-Azraq Oasis, respectively) use [q] as a variant of the Jordanian [g]. No indigenous Jordanian rural dialect makes use of [ʔ] as a variant of [g]. The variant [ʔ] is an exclusively urban feature and an import by speakers from urban centers in Palestine (Jaffa, Jerusalem, and elsewhere) and Syria (Damascus). Some rural speakers, especially females, are likely to switch from the rural variety [g] to the city [ʔ] upon taking residence in urban centers. The variant [ʔ] generally tends to elicit value judgments from opposite sides: those who view its use negatively, associating it with 'effeminacy' and affectation, and, on the other side, those who view its use positively, associating it with sophistication, urbanism, and desire for upward mobility (Sawaie 1986, 1987).

Jordanian dialects tend to use some velarized consonants (→ *tafxīm* ‘emphasis, velarization’), instead of their plain counterparts. Examples include: /ḷ/ (i.e. ‘dark’ [ɫ], as in *xaḷaf* ‘a man’s name’, *baġuḷ* ‘a mule’, etc.). This use of emphatics extends also to the voiced interdental fricative /ḏ/ [ð], thus resulting in /ḏ̣/ [ð̴], most noticeably in the pronunciation of the demonstrative pronouns, especially in Bedouin and Fellahin speech: *hāḏ̣a* ‘this’, *haḏ̣ōla* ‘these’, *haḏ̣olāk* ‘those’, and *haḏ̣īč* ‘that [fem.]’. Note that /č/ [tʃ] in the last item is a variation of /k/ [k], commonly used by rural and Bedouin speakers, never by city speakers unless they use it jocularly (ʿUzayzī 1973:I,25). Instead of /ḏ/, city speakers use the interdental stop /d/ in the pronunciation of the demonstrative pronouns, as in *hāda* ‘this [masc.]’, *hadīk* ‘that [fem.]’, etc. Additionally, among Bedouin and Fellahin speakers the emphatic /ṣ/ [s] often replaces the plain sibilant [s], as in *ṣaxan* ‘to become warm’, or *ṣāxin* ‘warm’. The *hamza* ‘glottal stop’ is produced as a voiced pharyngeal sound /ʿ/ in the dialects of Bani Hamida and Bani Hasan, both Bedouin tribes, as in *masʿala* for *masʾala* ‘matter’ and *hayʿa* for *hayʾa* ‘shape’ (ʿUzayzī 1973:I,9).

It is probably accurate to examine lexical items broadly in terms of the ecological parameters of city vs. rural. This group-specific lexical categorization does not necessarily preclude the same lexical items in the repertory of a different region or in another group. Around the 1930s, about half of the Trans-Jordanian population lived in villages and were settled farmers (Aruri 1972:34; Epstein 1938:228); urban centers and towns were not at all developed. One can safely extrapolate that linguistic differences among Trans-Jordanians were two-tiered: rural and Bedouin. Bedouin lexical items in Trans-Jordanian dialects tend to be transparent, marked phonologically, and stigmatized by other speakers (Abdel-Jawad 1986). The use of some of these items is age- or generation-dependent; younger, educated speakers tend to use lexical items differently from outmoded ones commonly used by members of older generation(s).

Certain lexical items tend to be specific to rural Jordan, a situation that induces association of certain lexical items with certain national groups and leads to regional identifiability: *ḥača* ‘he said’, *inhazam* ‘he ran away’, *gunnayniyyi* ‘a bottle’, *zamm* ‘he carried’, *baddi* [*waddi* or *widdi* in Bedouin speech] ‘I want’, etc.

Additionally, Arabic in Jordan exhibits morphological variations. Speakers in Karak in the south, for example, use *ʾa-* as an aspectual form for the 1st person singular prefix in the aspectual form of imperfect verbs. Speakers in the Irbid region in the north, however, use *ba-* to mark the same feature. Thus, *ʾa-šrab* ‘I drink’ in Karak is realized as *ba-šrab* in Irbid.

In the numerals 3–10, no distinction is shown between masculine and feminine in Jordanian Arabic (ʿUzayzī 1973:I,32). Thus, one hears *ṯalaṯ zlām* ‘three men’ and *ṯalaṯ niswān* ‘three women’. However, when these numerals are followed by a noun beginning with a glottal stop, an epenthetical [t] is introduced, as in the following: *ṯalaṯ t-iyyām* ‘three days’.

In Jordanian dialects, the feminine adjectival and nominal suffix tends to have the suffix *-ah* to mark feminine nouns, as in *barakah* ‘blessing’, *samakah* ‘a fish’, *fašakah* ‘a bullet’. The suffix *-ih* is used, however, to indicate feminine nouns and/or adjectives in some lexical items, as in *tisʿih* ‘nine’, *malīḥih* ‘good’, *zaġīrih* or *zġayyrih* ‘small’, etc.

The preceding variations among the Bedouin, Fellahin, and urban dialects aside, the official language sanctioned in all spheres of formal public interaction is the language variety that has acquired the name of Modern Standard Arabic (MSA). This variety is taught in schools and is acquired by schoolchildren as a second language, along with their already developed regional dialects. Modern Standard Arabic is commonly used in formal and semiformal settings. In its spoken form, it is used in radio and television newscasts, public lectures, religious sermons (in churches as well as in mosques), classroom instruction, specialized historical films, and television and radio programs. Written Arabic in newspapers and official documents is always in Modern Standard Arabic, unless the material is folk literature. Whereas stories and novels are often written in Modern Standard Arabic, it is not unusual to encounter dialogues in the dialectal varieties. Personal letters can be either in Modern Standard Arabic or the dialect, depending on the writer and the recipient.

Attitudinal studies in the context of Arabic in Jordan tend to focus on the study of speakers’ attitudes toward various dialects (Sawaie 1986, 1987, 1994a). As far as we know, attitudinal studies regarding foreign languages have not been conducted to date in Jordan. English is the predominant foreign language in the country, introduced at an early stage in public schools and used at the university level in the teaching of sciences, engineering, and medicine. However, the Jordan Arabic Language Academy is attempting to Arabicize instruction in these disciplines. Translation of scientific books into Arabic for university instruction has already been undertaken, despite opposition in some quarters.

The prevalence of English can be explained on historical grounds due to the role of the British in founding the country and its various governmental institutions, and also due to the present predominance of English globally. As more people acquire university education, they aspire to further their studies in English-speaking countries in order to enhance their professional and personal lives.

On the other hand, lately there has been a trend to acquire other languages. French was introduced in recent years on an experimental basis in some governmental secondary schools, despite the fact that private parochial schools had engaged in teaching this language much earlier. As this experiment shows success, French instruction in secondary schools may receive wider interest. Some public universities have introduced degree programs in modern European languages (French, Spanish, and Italian), as well as degree programs in Eastern languages (Persian, Hebrew, and Turkish). However, the percentage of students pursuing studies in these languages is very small relative to those who major in English.

BIBLIOGRAPHICAL REFERENCES

Abdel-Jawad, Hassan R.E. 1981. *Lexical and phonological variation in spoken Arabic in Amman.* Ph.D. diss., University of Pennsylvania.

——. 1986. "The emergence of an urban, dialect in the Jordanian urban centers". *International Journal of Sociology of Language* 61.53–63.

Abu Abbas, Khaled Hasan. 2003. *Topics in the phonology of Jordanian Arabic: An optimality theory.* Ph.D. diss., University of Kansas.

Al-Khatib, Mahmoud A. 1988. *Sociolinguistic change in an expanding center: A case study of Irbid City, Jordan.* Ph.D. diss., University of Durham.

——. 2001. "Language shift among the Armenians of Jordan". *International Journal of the Sociology of Language* 152.153–177.

——. and Mohammed Farghal. 1999. "English borrowing in Jordanian Arabic: Distribution, function and attitudes". *Grazer Linguistische Studien* 52.1–18.

Al-Madi, Muneeb and Sulayman Musa. 1959. *Tārīx al-ʾUrdunn fī l-qarn al-ʿišrīn,* 1st ed. Amman.

Al-Sughayer, Khalil Ibrahim. 1990. *Aspects of comparative Jordanian and Modern Standard Arabic phonology.* Ph.D. diss., Michigan State University.

Al-Wer, Enam. 2003. *Variation and change in Jordanian Arabic.* New York and London: Routledge Curzon.

Aruri, Naseer H. 1972. *Jordan: A study in political development (1921–1965).* The Hague: M. Nijhoff.

Bani-Yasin, Raslan and Jonathan Owens. 1984. "The Bduul dialect of Jordan". *Anthropological Linguistics* 26.202–235.

——. 1987. "The phonology of a northern Jordanian Arabic dialect". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 137.297–331.

Bauer, Leonhard. 1913. *Das Palästinische Arabisch: Die Dialekte des Städters und des Fellachen.* 3rd ed. Leipzig: J.C. Hinrichs'sche Buchhandlung.

Bergsträsser, Gotthelf. 1915. *Sprachatlas von Syrien und Palästina.* Leipzig. [Also publ. in *Zeitschrift des Deutschen Palästinavereins* 38.169–222.]

Binder, Friedrich. 1939. "Zur Aussprache des 'k' und 'q' bei den Beduinen Arabiens". *Wiener Zeitschrift für die Kunde des Morgenlandes* 46.209–218.

Butros, Albert J. 1963. *English loanwords in the Colloquial Arabic of Palestine (1917–1948) and Jordan (1948–1962).* Ph.D. diss., Columbia University.

Cantineau, Jean. 1939. "Remarques sur les parlers de sédentaires syrio-libano-palestiniens". *Bulletin de la Sociéte de Linguistique de Paris* 40.80–88.

——. 1946. *Les parlers arabes du Horan.* Paris: Klincksieck.

Christie, W. 1901. "Der Dialekt der Landbevölkerung des mittleren Galiläa". *Zeitschrift des Deutschen Palästinavereins* 24.69–112.

Cleveland, Ray L. 1963. "A classification for the Arabic dialects of Jordan". *Bulletin of the American Schools of Oriental Research* 171.56–63.

Dann, Uriel. 1984. *Studies in the history of the Transjordan, 1920–1949: The making of a state.* Boulder, Colo. and London: Westview Press.

Dweik, Bader S. 2000. "Linguistic and cultural maintenance among the Chechens of Jordan". *Language, Culture and Curriculum* 13:2.184–195.

El-Hassan, Shahir. 1977. "Educated Spoken Arabic in Egypt and the Levant: A critical review of diglossia and related concepts". *Archivum Linguisticum* 7.112–132.

El-Yasin, Mohammed K. 1982. *Voice phenomena in Jordanian Arabic.* Ph.D. diss., Cornell University.

Epstein, Eliahu. 1938. "The Bedouin of TransJordan: The social and economic problem". *Journal of the Royal Central Asian Society* 25.228–236.

Farghal, Mohammed and Abdullah Shakir. 1994. "Kin terms and titles of address as relational honorifics in Jordanian Arabic". *Anthropological Linguistics* 36.240–253.

Ferguson, Charles A. 1962. *Arabic dialect studies.* Washington, D.C.

Glubb, John B. 1938. "The economic situation of the Transjordan tribes". *Journal of the Royal Central Asian Society* 25.448–459.
Hourani, Albert H. 1947. *Minorities in the Arab world*. London and New York: Oxford University Press.
Hussein, Riyad F. 1980. *The case of triglossia in Arabic (with special emphasis on Jordan)*. Ph.D. diss., State University of New York at Buffalo.
——. and Nasser El-Ali. 1989. "Sub-jective reactions of rural university students toward different varieties of Arabic". *al-ʿArabiyya* 22.37–54.
Irshied, Omar Musleh. 1984. *The phonology of Arabic: Bani Hassan, a Bedouin Jordanian dialect*. Ph.D. diss., University of Illinois.
Jaimoukha, Amjad. 2001. *The Circassians: A handbook*. New York: Palgrave.
Kazziha, Walid. 1972. *The social history of southern Syria (Trans-Jordan) in the 19th and early 20th century*. Beirut: Beirut Arab University.
Migdadi, Fathi H. 2003. *Complimenting in Jordanian Arabic: A socio-pragmatic analysis*. Ph.D. diss., Ball State University.
Mitchell, T.F. 1978. "Educated spoken Arabic in Egypt and the Levant, with special reference to participle and tense". *Journal of Linguistics* 14.227–258.
——. 1980. "Dimensions of style in grammar of educated spoken Arabic". *Archivum Linguisticum* 11.89–106.
——. and Shahir El-Hassan. 1994. *Modality, mood and aspect in spoken Arabic, with special reference to Egypt and the Levant*. London: Kegan Paul.
Mūsā, Sulaymān. 1972. *Taʾsīs al-ʾImāra al-ʾurdunniyya*: 1921–25 2nd ed. Amman.
Owens, Jonathan and Raslan Bani-Yasin. 1987. "The lexical basis of variation in Jordanian Arabic". *Linguistics* 25.705–738.
——. 1991. "Spoken Arabic and language mixture". *Bulletin d'Études Orientales* 43.17–32.
Palva, Heikki. 1976. *Studies in the Arabic dialect of the semi nomadic əl-ʿAjārma tribe* (*al-Balqāʾ district, Jordan*). Göteborg: Acta Universitatis Gothoburgensis.
——. 1980. "Characteristics of the dialect of Bani Ṣaxar tribe". *Orientalia Suecana* 29.112–137.
——. 1984–1986. "Characteristics of the dialect of the Ḥwēṭāt tribe". *Orientalia Suecana* 33–35. 295–312.
Patai, Rafael. 1958. *The Kingdom of Jordan*. Princeton, N.J.: Princeton University Press.
Peake, F.G. 1939. "Transjordan". *Journal of the Royal Central Asian Society* 26.375–96.
Philby, H. St. John. 1924. "Transjordan". *Journal of the Royal Central Asian Society* 11.296–312.
Plascov, Avi. 1981. *The Palestinian refugees in Jordan, 1948–1957*. London: Frank Cass.
Saidat, Ahmad Mahmoud. 2003. *A sociolinguistic comparison of the syntax of Modern Standard Arabic and Jordanian Arabic*. M.A. thesis, University of Texas at Arlington.
Sakarna, Ahmad Khalaf. 1999. *Phonological aspect of ʿabady Arabic, a Bedouin Jordanian dialect*. Ph.D. diss., University of Wisconsin.
——. 2000. "The Bedouin dialect of Al-Zawaida tribe, Southern Jordan". *al-ʿArabiyya* 35.61–86.
Sawaie, Mohammed. 1986. "A sociolinguistic study of classical and colloqial Arabic varieties: A preliminary investigation into some Arabic speakers' attitudes". *al-Lisān al-ʿArabī*. 26.1–19. (Repr., *Arab Journal of Language Studies* 6, 1988. 129–155.)
——. 1987. "Speakers' attitudes toward linguistic variation: A case study of some Arabic dialects". *Linguistische Berichte* 107.3–22.
——. 1994a. *Linguistic variation and speakers' attitudes: A sociolinguistic study of some Arabic dialects*. Damascus: Al-Jaffan and Al-Jabi Publishers.
——. 1994b. "Sociolinguistic factors and historical linguistic change: The case of /q/ and its reflexes in Arabic dialects". *Investigating Arabic: Linguistic, pedagogical and literary studies in honor of Ernest N. McCarus*, ed. Raji M. Rammuny and Dilworth Parkinson, 15–39. Columbus, Oh.: Greyden Press.
Suleiman, Saleh M. 1985. *Jordanian Arabic between diglossia and bilingualism: Linguistic analysis*. Amsterdam and Philadelphia: J. Benjamins.
Tristram, Henry Baker. 1873. *The land of Moab: Travels and discoveries of the east side of the Dead Sea and the Jordan*. London: J. Murray.
ʿUzayzī, Rukus ibn Zaʾid. 1973–1974. *Qāmūs al-ʿādāt, al-lahajāt wa-l-ʾawābid al-ʾurdunniyya*. 3 vols. Amman: Culture and Arts Department.
Vatikiotis, P.J. 1967. *Politics and the military in Jordan*. New York: F.A. Praeger.
Zughoul, Muhammad Raji. 1980. "Diglossia in Arabic: Investigating solutions". *Anthropological Linguistics* 22.201–217.

MOHAMMED SAWAIE
(University of Virginia)

Jordanian Arabic (Amman)

Jordan is the southernmost country of *Bilād aš-Šam*, and, like the other countries in the region, it came to exist as a separate political entity after the dismantlement of the Ottoman Empire. This part of the region had been the poorest in terms of natural and human resources, and it lacked large urban centers that could act as cultural and linguistic focal points for the local population. Throughout the first half of the 20th century, Jordanians looked outside their country's borders to cities such as Haifa, Jerusalem, Damascus, and Beirut for cultural refinement and education. The linguistic impact of this situation was that while in Syria, Lebanon, and Palestine, the course of linguistic developments in the local dialects has favored the emergence of national norms based on the dialects of the metropolis, in the Jordanian capital the linguistic situation remained largely diffuse, and, until recently, a label such as 'Ammani Arabic' has been rather ill-defined. The absence of a linguistic metropolis in the country also partly explains the obvious linguistic influence of urban Palestinian dialects on Jordanian Arabic, compared with the effect of these dialects on Syrian and Lebanese dialects. It is noticeable, for instance, that the Palestinians in Syria and in Lebanon readily accommodate to the local Syrian or Lebanese dialects, although the Palestinian dialects have a considerably larger demographic representation in Amman than in Damascus or Beirut.

Jordanian dialects are generally under-studied, and no comprehensive accounts of their grammars are available (→ Jordan). Among the valuable resources are a series of articles by Heikki Palva on certain tribal dialects from the Balqa region, and by the same author on the dialect of el-Karak. The area east of the River Jordan is included in Bergsträsser's atlas. In addition, Cleveland (1963) provides a rough classification of the dialects spoken in the country. Within the framework of Labovian sociolinguistics, there are a few studies that address mainly phonological varia-tion and change in some Jordanian localities, namely Amman (Al-Wer 2000, 2002, 2003; Abdel-Jawad 1981), Irbid (Al-Khatib 1988), and Sult, Ajloun, and Kerak (Al-Wer 1991, forthcoming).

According to Cleveland's classification, Jordanian Arabic is classified into two major dialect groups. Broadly speaking, the eastern and southern dialects are of the Najdī Arabic type and have fewer speakers than the northern and northwestern dialects. The northern dialects are classified as southern Levantine. Within this group, one can add a further distinction between the far northern dialects, which are clearly of the Ḥōrānī type (e.g. the dialect of the city of Ajloun), and Balgāwī (e.g. the dialect of the city of Sult and its surroundings). Balgāwī could well turn out to be essentially Ḥōrānī, but contemporarily the two types can be distinguished by a number of features.

1. INTRODUCTION: AMMAN ARABIC

Amman, Roman Philadelphia and Ammonite Rabbath Ammon, is a new city that has no traditional native Arabic dialect simply because it did not have a native population. At the turn of the 20th century it was a deserted Roman site that had been used as a summer watering ground by Bedouin tribes, particularly those of the nearby Balga region. By 1906, it became home to 5,000 or so Circassian settlers, who were at that time monolingual in dialects of Adyghe, a western branch of the North Caucasian language family. Amman was declared the capital city of the newly formed Emirate of Transjordan in 1921, and thus attracted migrants from other locations. By the 1930s, it had received an additional 5,000 migrants, mainly from the Balga and Ḥōrān regions in the north, from Kerak and Madaba in the south, and from the Palestinian cities of Haifa, Jaffa, Nablus, and Hebron, as well as a few merchant families from Damascus. Internal migration from other Jordanian and Palestinian towns and villages increased steadily in the following decades, but the most sudden population increases occurred as a result of the wars with Israel in 1948 and 1967. In the early 1950s, Amman had just over 100,000 people, and by the late 1990s the population had increased more than fifteen times, to reach 1.6 million, according to the most recent census. Including the earliest emigrants, Amman only has three generations of native inhabitants.

The population is mainly comprised of two groups, who will be called here the Jordanians and the Palestinians for simplicity, but without losing sight of the fact that neither group is homogeneous, and intermarriages between the two groups are commonplace. Under these labels, the Jordanians are the sector of the population whose families originally migrated from various localities within Jordan and spoke Jordanian dialects, and the Palestinians are those whose families originally came from Palestinian towns and cities and spoke Palestinian dialects. Unfortunately, no statistics are published with details of the breakdown of the city's inhabitants in terms of their origins. Informal reports indicate that those of Palestinian origin may comprise a majority of the city's population. The dialects spoken by both groups exhibit a high degree of variability, especially so in the case of the Palestinian dialects. The three major Palestinian dialect types, the urban (Madanī), the rural (Fallāḥī), and the Bedouin, are represented in Amman. Among the Palestinian dialects, the most important distinction is between the Madanī and the Fallāḥī groups of dialects, which can be readily distinguished by a number of linguistic features, such as the variants of (q), which are [ʔ] in Madanī and [k] in Fallāḥī. As widely documented, the trend of linguistic developments in Palestinian dialects in Palestine itself is overwhelmingly in favor of the Madanī dialects. In Amman, too, the Fallāḥī Palestinian dialects are certainly the losers, although their demographic representation in the city may be larger than the representation of the Madanī dialects. None of the linguistic features that have been focused in Amman or play an important role in the formation of its dialect are Fallāḥī Palestinian in origin. Thus, the competing linguistic features in Amman are really of the Madanī Palestinian type on the one hand, and the Jordanian East Bank type (linguistically Bedouin) on the other. In terms of numbers of speakers, these two types are roughly equally represented.

In three generations, the speech of Amman has undergone a considerable degree of focusing, to the extent that it is now possible to speak of a distinctive dialect in its own right, which the younger generation in the city call ʿAmmānī. This is really a textbook case of the process of new dialect formation, as outlined in Trudgill (1986, Chap. 3). Here, first-generation inhabitants speak dialects that can be easily identified with the original Jordanian or Palestinian towns from which they migrated. In the speech of the second generation, there is a mixture of both types and a complication in the sociolinguistic correlations. For instance, whereas in the speech of the first generation there is a straightforward correlation between origin of the speakers and linguistic usage, in the second generation gender emerges as an important factor, too. But the speech of the second generation can still be identified as either Jordanian or Palestinian, most clearly through vocalic features. The mixture is considerably reduced in the speech of the current generation; instead, we find a high degree of stability of usage of certain linguistic features, the use of fudged forms, and very many totally new features.

The extralinguistic aspects of the emergence and increased focusing of the new dialect are significant. For instance, youngsters in Amman call themselves ʿ*Ammāniyyīn*, which is a new derivation denoting the status of natives of the city. Their parents, on the other hand, even those who were born in Amman, most often refer to themselves as *sukkān ʿAmmān* 'inhabitants of Amman', while they affiliate themselves with the hometowns and villages of their forefathers. This development indicates that the city is acquiring a regional identity for the first time in its modern history. The symbols of this new identity are reflected in the youngsters' clear linguistic divergence from the dialects spoken by their parents and their grandparents, and the agreement and awareness among themselves as to what constitutes their dialect.

2. Linguistic description

2.1 *Phonology*

2.1.1 Consonants

The phonological inventory is given in Table 1. The interdentals (plain and emphatic) are currently in a state of variability with their stop counterparts. The origin of the variability in Amman is the contact between Jordanian dialects, which have interdentals, and Madanī Palestinian dialects, which do not. In the case of the plain interdentals, /ṯ/ and /ḏ/, many speakers (especially the male speakers) use them variably with /t/ and /d/. The emphatic interdental, /ḏ̣/, shows a considerably lower degree of variation with the stop variant /ḍ/ in the speech of the same speaker, i.e., the speakers tend to use either one or the other. The speakers who use the emphatic interdental are exclusively of Jordanian or Fallāḥī Palestinian origin and are most often male speakers, while those who use the stop counterpart can be from anywhere and of either sex. The inventory above lists both fricative *ž* [ʒ] and affricate *j* [dʒ]. These are not separate phonemes but rather variants of the same phonological unit. The variation between [ʒ] and [dʒ] is not conditioned phonologically. The occurrence of the fricative or the affricate correlates with independent variables, such as origin of the speaker and gender. The affricate variant occurs mainly in Jordanian and Fellāḥī Palestinian speech, especially in the speech of male speakers. The fricative can occur in the speech of all groups. Especially in the speech of the younger generation in Amman, there is evidence of a change in progress from interdental to stop and a change from affricate to fricative. If these processes go to completion, the system will become empty of interdental and affricate sounds altogether, and thus become similar to that found in the other major urban dialects in the region.

Table 1. Inventory of consonants in ʿAmmānī Arabic

	bilabial	labio-dental	dental	inter-dental	alveolar	postal-veolar	palatal	velar	pharyngeal	laryngeal
plosives										
voiceless, voiced	b		t, d					k, g		ʾ
emphatic			ṭ, ḍ							
nasal	m		n							
fricatives										
voiceless, voiced		f		ṯ, ḏ	s, z	š, ž		x, ġ	ḥ, ʿ	h
emphatic				ḏ̣	ṣ					
affricate						j				
trill or tap					r					
lateral					l					
glides	w						y			

The glottal stop /ʾ/ listed above occurs both as an independent phoneme, as in [saʔal] 'he has asked', and as a variant of /g/, as in [ʔaːl] 'he has said'. The [g] ~ [ʔ] variation is exceptionally salient in Amman. For many speakers, these sounds are used as labels to distinguish Jordanian (*gāl* dialect) from 'other' (*ʾāl* dialects), and male ([g]) from female ([ʔ]) speech. In Labov's terminology, they can be considered *stereotypes*. It is noticeable in this context that the Fallāḥī Palestinian variant [k] is completely leveled out in the emergent dialect. The competition is solely between [g] and [ʔ]. Generally speaking, the female speakers show a strong preference for [ʔ], and the female speakers from a Madanī Palestinian background do not use [g] at all. The highest rate of variation is found in the speech of male speakers from Madanī background, who often use [g] in public, especially when interacting with other male speakers. In practice, the speakers who alternate between the two sounds normally do so depending on the context, the participants, and the identities they wish to project in a particular setting. They utilize the full range of connotations associated with the use of these variants. It is difficult to predict on the basis of the available data whether one of the two sounds will oust the other, exactly because of the stereotypical nature of the variation between the two sounds. Theoretically speaking, [ʔ] should have a better chance of diffusing, given that it is the urban koine form in the region. On the other hand, [g] is strongly associated with a local Jordanian identity, which is occasionally called upon, especially at times of conflicts. It is very well possible that the variation in the use of [g] and [ʔ] will remain stable and that the dialect will have both sounds as a case of inherent variability. In the examples used here to demonstrate various features, these variations will be used alternately, with the implication that items such as *talāte* 'three', *gāl* 'he said' also occur as *ṯalāṯe* and *ʾāl*.

2.1.2 Vowels

Table 2. Inventory of vowels in ʿAmmānī Arabic

Short vowels	Long vowels
i u	ī ū
(o)	ē ō
a	ā

The phonetic property of the short high front vowel /i/ is normally [ɪ] or slightly lower, and that of the short high back vowel /u/ is [ʊ] or slightly lower. The 3rd person singular masculine clitic is phonetically [o] or slightly lower. The contrast between /o/ and /u/ can be found only word-finally, *ʾāmo* 'he removed it' vs. *ʾāmu* 'they removed', *šāfo* 'he saw him' vs. *šāfu* 'they saw'. Since the 3rd person singular masculine is phonetically [o] but the 3rd person plural inflection is [u], minimal pairs for the /o/, /u/ distinction in final position are plentiful. Similarly, /o/ and /i/ contrast phonemically in final position, in the capacity of *-i*, phonetically [i], being the 1st person singular possessive clitic form: *ṣāḥbo* 'his friend' vs. *ṣāḥbi* 'my friend', *ʿindo* 'he has' vs. *ʿindi* 'I have'. With such restricted position of occurrence, however, it is debatable whether a phonemic status can be established for /o/ and /e/ (on this point, see Garbell 1958:332). The status of /i/ and /u/ as separate phonemes is also precarious. Minimal pairs for this opposition are extremely hard to come by: *full* 'a type of flower', *fill* 'run away!', *ḥubb* 'love [noun]', *ḥibb* 'love! [verb]', thus, *ḥubbha lazīz* 'her love is sweet', *rūḥ ḥibbha wxallişna* 'for God's sake, go on love her and be done with it'. Some of the items mentioned in de Jong (2000:69) from North Sinai dialects also occur in ʿAmmānī: *fitt* 'break bread into little pieces!' vs. *futt* 'I dropped by', *xiff* (as in *xiff ižrak)* 'do something quickly!' vs. *xuff* 'hoof', and some of those mentioned in Blanc (1970; cited in de Jong 2000:69-70) also occur in ʿAmmānī: *gimt* ~ *ʾimt* 'I removed' vs. *gumt* ~ *ʾumt* 'I got up'; *ṭibb* 'medicine' vs. *ṭubb* 'arrive unannounced!'. The Amman dialect tends toward /u/ where other neighboring dialects, e.g. Damascus and Beirut, have /i/, e.g. *lubb* 'heart of the fruit', *ʾumm* 'mother', *ḥuṛṛ* 'free', *šurṭa* 'police', *kull* 'all'.

The short low /a/ has phonetic realizations that range from back [ɑ] (especially after /r/) to true front low [a] (cardinal vowel 4), and can be raised to [ɛ] (cardinal vowel 3). Three-way opposition between /a/, /i/, /u/ can be found in *fall* 'he ran away', *full, fill* 'run away!' (in addition to the items mentioned above for /i/ vs. /u/, e.g. *ḥabb* 'he loved/liked'). Two-way opposition between /a/ and /i/ or /u/ is plentiful: *kabb* 'he disposed [of something], *kubb* 'dispose of!', *ḥarr* 'spicy', *ḥurr* 'free', *lamm* 'he gathered', *limm* 'gather!'.

The mid long vowels /ē/ and /ō/ are reflexes of **ay*, and **aw: bēt* 'house', *zēt* 'oil', *ṣōt* 'noise; voice', *mōt* 'death'. The monophthongal realizations group the dialects of Amman, Jerusalem, and Damascus together, in opposition to the Beirut dialect, which has the diphthongal realizations. In ʿAmmānī, /ay/ and /aw/ are used before /y/ and /w/, respectively: *mayyil* 'drop by!', *ʾawwal* 'first'; they are also used in the comparative form of adjectives that begin with /y/ and /w/: *yābis/ʾaybas* 'dry/drier', *wāsiʿ/ʾawsaʿ* 'wide/wider'. For phonotactic reasons, the diphthongs are preserved in *law* 'if', *ʾaw* 'or', *fayy* 'shade', *ʾayy* 'any'. Phonetically, /ē/ and /ō/ are roughly of the quality of cardinal vowels 3 and 6 or slightly higher, respectively. In the speech of the younger generation, the mid and high front long vowels can develop an off-glide (on this process, see below).

The long low vowel /ā/ has a relatively high functional load and phonetic values that range from back [ɑː] (in the vicinity of /r/) through [ɑː] to [æː], and, especially in the speech of first- and second-generation speakers of originally Palestinian dialects, long /ā/ can be as close as long [eː]. Thus, the name of the city of Amman can occur as [ʕammaːn], [ʕammæːn], [ʕammɛːn], [ʕammeːn]. However, in the speech of the younger generation the closest realizations are leveled out. Focusing around [æː] is taking place. In the environment of /r/, where both Jordanian and Palestinian input dialects have a back /a/ (and in the case of the latter group, /a/ is extremely back with pharyngeal constriction), fronting to [a]-like quality takes place. From the perspective of traditional Jordanian dialects, these processes involve fronting and raising of long /a/, which may represent the onset of a chain shift. The movement of long /a/ in this fashion causes pressure in the phonological space of the long mid and long high vowels, which may explain the development of the off-glide mentioned above. Thus, *zēt* 'oil' becomes [zeːət], and *mīn* 'who' becomes [miːən] (see Al-Wer 2002).

2.1.3 Final *-a* ~ *-e*

The feminine ending *-a* is raised to [ɛ] everywhere, except after velarized, emphatic, back, and pharyngeal sounds, as in the other major urban Levantine dialects, e.g. *ḥilwe* 'pretty', *baṣale* 'onion', *kulliyye* 'college', but *bišʿa* 'ugly', *lamḥa* 'glance', *maraga* 'sauce'; the same is true for *huwwe* 'he', *hiyye* 'she'. The raised feminine marker in Amman is most often a half-open vowel, unlike, for instance, in Jerusalem, where it is considerably closer. A closer realization can be heard in the speech of older Palestinians.

2.1.4 Syllable structure and word stress

Open and closed syllables occur in ʿAmmānī. The possible syllable types are Cv, Cv̄, *bā.ba* 'daddy'; CCv, *sli.gī.ha* 'boil it!'; CCv̄, *ṣfī.ḥa* 'meat pie'; CvC, *saḥ.lab* 'milk pudding'; CCvC, *staʿ.baṭ* 'he acted foolishly'; Cv̄C, *ma.ka.tīb* 'letters'; CCv̄C, *krūm* 'vineyard'; CCvCC, *kriht* 'I hated'. Word stress falls on the ultimate, if super heavy, *ma.ka.tī́b*, otherwise, on a heavy penultimate, *ma.kā́.tib* 'offices'; and if there is no heavy penultimate, stress falls on the antepenultimate, *mádrase* 'school'. Stress is predictable in accordance with this rule also when the word inflects, or is attached to suffixes of various sorts: *kā́.tab, kā́t.bat, kā́t.*

ba.to, ka.ta.bū, ka.ta.bū.lo, ka.ta.būl.hum ‘he wrote, she wrote, she wrote it, they wrote it, they wrote to him, they wrote to them’; *yit.fā́l.saf* ‘he philosophizes’, *tit.fal.sá.fiš* ‘do not play smart!’, *tfal.sa.fī́.lo* ‘play smart with him!’, *tit.fal.saf.lī́š* ‘do not play smart with me!’; *tā́w.le* ‘table’, *tā.wíl.ti* ‘my table’, *tāw.lít.ha* ‘her table’.

2.2 *Morphology*

In the speech of the younger generation, gender distinction is maintained only in the 2nd and 3rd persons singular. Neutralization, as in *katabu* for *katabin* + *katabu,* is consistently in favor of the (originally) masculine forms. The same is true of gender neutralization in the pronouns and pronominal clitics, which is an important feature that distinguishes ʿAmmānī from the other major Levantine dialects. Evidence for the existence of the distinction in the plural forms is widely available in the speech of the first-generation speakers from Jordanian origins. Evidence is also available that the prevalent pattern (generalization of the masculine forms) is a local innovation by the third generation. The feminine forms occur as the generalized forms in the speech of the first- and some second-generation speakers of Palestinian origin. In the description which follows, only the forms occurring in the speech of the third generation have been included, on the basis that these forms are the features that are most likely to become characteristic of the dialect; where the emerging pattern is unclear, more than one form is listed.

The verb allows the suffixation of direct and indirect objects: *biʿzimha* ‘he invites her’, *ʾaʿṭētūna* ‘you gave us’, *katablo* ‘he wrote to him’. The nouns allows suffixation of possessive pronouns: *ktābo* ‘his book’. The verb also allows suffixation of the negative particle *-š* (whose use is optional with a preceding *ma*): *btiʿzimhūš* ‘she doesn’t invite him’.

2.2.1 Pronouns

2.2.1.1 Independent personal pronouns

Independent personal pronouns are set forth in Table 3.

Table 3. Independent personal pronouns in ʿAmmānī Arabic

	3rd pers.	2nd pers.	1st pers.
sg. masc.	*huwwe*	*inta*	*ana*
sg. fem.	*hiyye*	*inti*	
pl.	*humme*	*intu*	*niḥna/iḥna*

2.2.1.2 Possessive/object suffixes

There are two series of suffixes, which differ according to whether the word ends in a vowel or a consonant (including CC), as shown in Table 4.

The *-kum* of the 2nd person plural is used consistently by the third generation; older-generation Jordanians use *-ku* for both genders, or *-ku*/*-kin*, and older Palestinians use *-kon.*

2.2.1.3 Indirect object suffixes

A preceding -CC produces a third series, in addition to the two series after a vowel or after a consonant (Table 5).

Table 4. Possessive/object suffixes in ʿAmmānī Arabic

	after-v			after-C		
	axu- ‘brother’			*xāl* ‘uncle’		
	3rd pers.	2nd pers.	1st pers.	3rd pers.	2nd pers.	1st pers.
sg. masc.	'*-(h)*	*-k*	*-y*	*-o*	*-ak*	*-i*
sg. fem.	*-ha*	*-ki*		*-ha*	*-ik*	
pl.	*-hum*	*kum*	*-na*	*-hum*	*-kum*	*-na*

Table 5. Indirect object suffixes in ʿAmmānī Arabic

after -v			after -C			after -CC		
katabūlo ‘they wrote to him’			*katablo* ‘he wrote to him’			*katabtillo* ‘I wrote to him’		
-lu	*-lak*	*-li*	*-lo*	*-lak*	*-li*	*-illo*	*-illak*	*-illi*
-lha	*-lik*	*-li*	*-ilha*	*-lik*		*-ilha*	*-illik*	
-Ihum	*-lkum*	*-lna*	*-ilhum*	*-ilkum*	*-ilna*	*-ilhum*	*-ilkum*	*ilna*

2.2.1.4 Demonstratives

Table 6. Demonstratives in ʿAmmānī Arabic

	near deixis	far deixis
sg. masc.	*hād(a)*	*hadāk*
sg. fem.	*hāy ~ hādi*	*hadīk*
pl.	*hadōl*	*hadolāk*

The attributive demonstrative can occur before or after the noun: *hāda ilwalad ~ ilwalad hāda* 'this boy'. A contracted form of the near deixis forms is *hal-*, which is used in an invariant form with the noun, e.g. *halwalad, halbint, halbanāt.* The *-l-* element is the definite article. *Hal-* is usually used as a stylistic device, especially in expressions of astonishment and the like, e.g. *halwalad mā ʾašṭaro* 'how clever this boy is!'. To add emphasis, *hal-* can also occur together with a post-nominal *hāda/hāy/hadōl: halwalad hāda miš maʿ ʾūl šū tʾīl damm* 'it is incredible how insipid this boy is'. With indefinite nouns, and in similar expressions, *hēk* can be used, e.g. *hēk zalame willa balāš* 'this is what I call a man'.

2.2.1.5 Presentatives

Hayy functions as a presentative in the following forms: *hayyo, hayyha, hayyhum, hayyak, hayyik, hayykum, hayyni, hayyna: hayyni baktub* 'here I am, writing', *hayyo fahhmu* 'here he is, make him understand', *hayyhum žāyyīn* 'here they are, coming'.

2.2.1.6 Interrogatives

The interrogative pronouns are: *mīn* 'who?', placed nearest to the subject of inquiry, e.g. *mīn talfan* 'who called?', *maʿ mīn aḥki, aḥki maʿ mīn* 'whom shall I speak to?'; *anū* sg. masc. and pl. masc./fem., *anū waḥad* 'which one?', *anū karāsi* 'which chairs?'; *anī* sg. fem. and pl. fem. 'which?', e.g. *anī bint/banāt* 'which girl/girls?'. *Anū* can be generalized for use with both genders and all numbers. There are two interrogative particles, which can be used interchangeably: *ēš ~ šū* 'what?'. They, too, assume various positions depending on the subject of inquiry, e.g. *ēš ~ šū biddak tiʿmal* 'what do you want to do?', *ʿimlat ēš ~ šū* 'she did what?'.

2.2.2 Adverbs

Temporal: *ēmta* 'when', *lyōm* 'today', *bukra* 'tomorrow', *baʿid bukra* 'the day after tomorrow', *mbāriḥ* 'yesterday', *ʾawwal mbāriḥ* or *ʾabl imbāriḥ* 'the day before yesterday', *halla* 'now'. An interesting temporal adverbial expression which is losing ground in Amman, but is nevertheless worth mentioning, is *ṯāni yōm* 'the day after', e.g. *yaʿni bitsāfri ṯāni yōm ʿīd mīlādik* 'so, you travel the day after your birthday'. The younger generation in Amman do not generally parse this expression correctly (they may not have acquired its correct reference in the first place). They tend to understand the events mentioned in the example above, the act of traveling and the date of the birthday, the other way around, i.e., they interpret the date of the birthday to be after the day traveling takes place.

Local: *wēn* 'where?', *minwēn* 'from where?', *lawēn* 'where to?', *hōn* 'here', *hunāk* 'there'.

Manner: *kēf ~ kīf* 'how', *hēk* 'like this', *ktīr* 'very', *ʿalʾāxir* 'totally', *laʾalla* lit. 'to God', used as an intensifier as in *ʿindhum sanwišāt zākye laʾalla* 'they sell very delicious sandwiches'; *ʿādi* lit. 'ordinary', used in a range of situations, often in the sense of 'it makes no difference'.

Causal: *lēš* 'what for?' 'why?', e.g. *lēš ʿmilt hēk* 'what did you do this for?'; *ʿašān hēk* 'therefore'.

Number and mass: *ʾaddēš ~ ʾaddē* 'how many?', 'how much?'.

2.2.3 Particles

2.2.3.1 Article

The article is *il ~l*, as in *ilwalad* 'the boy', *liwlād* 'the boys'.

2.2.3.2 Genitive marker

The genitive marker is *tabaʿ: tabaʿo* masc., *tabaʿha* fem., *tabaʿhum* pl., or *tāʿ: tāʿo, tāʿha, tāʿhum, tāʿithum* pl. fem. (the latter two used more frequently by female younger speakers). In ʿAmmānī, however, the construct state is the unmarked option, and the analytical genitive tends to be the marked option, the use of which may be precipitated by pragmatic reasons, e.g., to add emphasis or to disambiguate an utterance. For instance, consider these exchanges:

vēlithum žanb innādi 'their villa is near the club'; *šū illi žanb innādi?* 'what is near the club?'; *ilvēlla tāʿithum* 'their villa'

laʾ laʾ issayyāra tabʿato bēḍa 'no, no, his car is white'

2.2.3.3 Negation

The negative particles are *mā, mū, mā… +š,…+š, miš*, and *la*. The particle *la* is used before imperatives: *rūḥ* ⇒ *la trūḥ* 'do not go!'. In this construction, *mā…+š* is also admissible: *mā trūḥ* ~ *trūḥiš*. In response to yes/no questions, an alveolar click, which may be repeated a number of times, can replace *laʾ* in informal situations.

miš is used to negate participles (active and passive), e.g. *miš rāyiḥ* 'I am not going', *miš maʿmūl* 'it is not done'. It is also used before prepositions: *miš bi lbēt* 'not in the house', *miš maʿi* 'not with me'; before adjectives: *miš ḥilu* 'not nice'; before nouns: *zēd miš daktōr* 'Zaid is not a doctor'; before quantifiers: *miš kull ilbanāt* 'not all the girls'; and for sentential negation in the jussive mood: *miš trūḥ tiḥkīlhum ilguṣṣa* 'don't you go telling them the story!'. The particle *mū* can replace *miš* in all of the above examples, but it is used less frequently.

mā, with or without +*š*, is used before verbs in the perfect, e.g. *mā šift ʾiši* ~ *ma šiftiš ʾiši* 'I haven't seen anything'; and in the imperfect: *mā baktub* ~ *ma baktubiš* 'I do not write'.

2.2.3.4 Prepositions

The usual set of prepositions are used in ʿAmmānī. Both *fi* and *bi* are used. However, when followed by the noun, although both can be heard, the younger generation prefer *bi: bi lbēt* rather than *fi lbēt* 'in the house'. Elsewhere, for instance, when pronominal suffixes are attached to these prepositions, only *fi* + suffix is permitted: *fiyyo* ~ *fī* 'in it'. The preposition *l-* is used as *ʾil* ~ *laʾil*, when attached to a pronominal suffix: *ʾili* ~ *laʾili, ʾilak* ~ *laʾilak, ʾilna* ~ *laʾilna* 'for me', 'for you', 'for us', and as *la* when followed by a noun: *la ʾummi* 'for my mother'.

2.2.3.5 Conjunctions

The conjunctions used include: *lamma* 'when'; *ta-* 'when', *tayiži iṣṣabi binṣalli ʿannabi* 'when the boy [newborn] arrives, we shall bless the Prophet'; *ʿašān* 'so that'; *ḥatta* ~ *ta-* (short for *ḥatta*) 'so that', *ʾihda tatifham* 'calm down so that you understand'; *liʾanno* 'because'; *bass* 'but'; *willa* 'or'; *ʾiza* ~ *ʾizan* 'if'; *law* 'if'; *la* 'lest'; *ʾinno* 'that'; *lamma* ~ *lamman* 'when'.

2.2.4 Nominal morphology

2.2.4.1 Gender

Feminine nouns without marking include: *ʾīd* 'hand; arm', *ʿēn* 'eye', *ʾižir* 'foot; leg'.

2.2.4.2 Productive patterns

For instruments: mu(i)CCāC *muftāḥ* 'key', *munfāx* or *minfāx* 'air pump', *mihbāš* 'a traditional coffee grinder'; maCaCC *mafakk* 'screwdriver', *magaṣṣ* 'scissors'; CaCCāCa(e) *massāḥa* 'windshield wiper', *šaffāṭa* 'plunger', *ʾaššāṭa* 'floor rubber wiper', *jarrāfe* 'digger', *maṣṣāṣa* 'drinking straw'; CuCCēCa(e), *ṣuffēra* 'whistle', *ḥummēle* 'hammock'. For professions, CaCCāC produces *laḥḥām* 'butcher', *ḥallāg* 'barber'; for this category noun + *ji* is productive: *kundarji* 'cobbler', *ḥalawanji* 'sweet maker'.

The → pseudo-dual is maintained in: *ʾidēn* 'hands; arms', *ʾižrēn* 'feet; legs', and for some speakers in *ʿinēn*. These items lose the *-n* when suffixed: *ʾidēha* 'her hands; her arms', *ʾižrayyē* 'my feet'; my legs', *ʿinē* ~ *ʿyūno* 'his eyes'.

2.2.5 Numerals

The cardinal numbers 1–10 are: *waḥad, tnēn* or *tintēn, talāte, ʾarbaʿa, xamse, sitte, sabʿa, tamanye, tisʿa, ʿašara*. When the numerals *waḥad* and *tnēn* are used in conjunction with the noun, they always assume a postnominal position and show gender agreement: *walad waḥad* 'one boy', *bint waḥde* 'one girl, *wlād tnēn* or *waladēn tnēn* 'two boys', *banāt tintēn* 'two girls'. When the numerals 3–10 are used with a following noun they are shortened as follows: *talat, arbaʿ, xamis, sitt, sabiʿ, taman, tisiʿ, ʿašar: talat banāt* 'three girls', *ʾarbaʿ iwlād* 'four boys'. If the noun begins with a vowel, *-t* can be added, and the vowel changes from /a/ to /i/, if it is followed by /a/: *ʾarbaʿ t-iyyām* 'four days', and from /a/ to /u/, if it is followed by /u/: *sitt-t-ušhur* 'six months'. For some speakers in the younger generation, however, these occur as *ʾarbaʿ ʾayyām, sitt ʾašhur*.

The numerals 11–19 are: *ḥdaʿš, ṭnaʿš, talaṭṭaʿš, ʾarbaʿṭaʿš, xamisṭaʿš, siṭṭaʿš, sabaʿṭaʿš, tamanṭaʿš, tisaʿṭaʿš*. If the noun follows, *-ar* is added to the numeral: *ḥdaʿšar walad* 'eleven boys'. The ordinal numbers are: *ʾawwal* 'first', *tāni* 'second' from 'third' upwards, they follow the CāCiC pattern: *tālit, rābiʿ, xāmis*, etc.

Table 7. Derived Forms in ʿAmmānī Arabic

I	II	III
daras/yudrus 'to study'	*darras/idarris* 'to teach'	*sāʿad/isāʿid* 'to help'
širib/yišrab 'to drink'	*kabbar/ikabbir* 'to enlarge'	*rāgab/irāgib* 'to watch'
kibir/yikbar 'to grow'		*sābag/isābig* 'to race'
	V (*t*-II)	VI (*t*-III)
	tdarras/yitdarras 'to be taught'	*tsāʿad/yitsāʿad* 'to be helped'
	tnaḍḍaf/yitnaḍḍaf 'to be cleaned'	*trāgab/yitrāgab* 'to be watched'
	tkabbar/yitkabbar 'to be enlarged'	*tṣāḥab/yitṣāḥab* 'to befriend'
VII (*n*-I)	**VIII (*t*-I)**	
ndaras/yindaris 'to be studied'	*xtalaf/yixtalif* 'to disagree'	
nšarab/yinšarib 'to be drunken'	*ftaraʾ/yiftariʾ* 'to part with someone'	
nṣaraf/yinṣarif 'to leave'		
X (*sta-*)	IX	
staslam/yistaslim 'to surrender'	*ḥmarr/yiḥmarr* 'to become red'	
staʿadd/yistaʿidd 'to prepare oneself'	*xḍarr/yixḍarr* 'to become green'	

2.2.6 Strong verb

2.2.6.1 Forms

Form I verbs are either CaCaC, which is always transitive and usually has an *u*- or *a*-imperfect, or CiCiC, which has an *a*-imperfect and can be transitive, e.g. *ʿimil/ yiʿmal* 'to do', or intransitive, e.g. *kibir/yikbar* 'to grow'.

Form II is always CaCCaC and has an *i* in the final syllable. It turns an intransitive verb into transitive, e.g. *barad/yubrud*, 'to feel cold', *barrad/ibarrid* 'to cool'; it is a particularly productive form and has a relatively high functional load conversationally, especially in storytelling.

Form III is CaaCaC and has an *i* in the final syllable. This form implies the involvement of more than one person/thing in the action, *ṣāḥab* 'he befriended'.

Forms VII (the *n*- prefix) derives the passive and also, in the case of some verbs, denotes acquiescence to an action which was precipitated by others, or by a situation, e.g. *nṣaraf/ yinṣarif* lit. 'to leave', but in ʿAmmānī (as in the Levant in general) its meaning is similar to British English 'to buzz off'. There is considerable variation in the imperfect of this Form among ʿAmmānīs. In the speech of the Jordanians, *n*- prefixed to Form I (both CaCaC and CiCiC) yields as imperfect either yinCaCiC, or yinCaCaC, *yinkatib ~ yinkatab, yinšarib ~ yinšarab*. For some speakers of Palestinian origin, resyllabification occurs, yielding yinCCiC, *yinktib, yinšrib*. It seems that the (originally) Palestinian yinCCiC is leveled out by the third generation, and that the (originally) Jordanian yinCaCiC ~ yinCaCaC is undergoing focusing.

Forms V and VI (*t*-prefix) also function to derive passives, and, in the case of Form VI, there is a sense of reciprocity. Form VIII involves an infix, and in the case of X with a *sta*-prefix, there is a change from /a/ to /i/ in the imperfect. The remaining Form IX is particularly productive for color associations, *swadd/ yiswadd* 'to become black or darker', *ṣfarr/yiṣfarr* 'to become yellow or pale'.

2.2.6.2 Inflections

2.2.6.2.1 Perfect

In the speech of the younger generation, gender distinction is maintained in the 3rd and 2nd person singular. The older generation, especially of Jordanian origin, occasionally maintain gender distinction in the plural forms, *katabtu/katabtin, katabu/katabin.*

2.2.6.2.2 Imperfect

Here, too, the younger generation show gender distinction only in the 3rd and 2nd person singular, while in the speech of the (Jordanian) older generation, gender distinction can be heard in the plural forms, *yudrusu ~ yudursu/ yudrusin ~ yudursin.*

Table 8. Inflection of the perfect in ʿAmmānī Arabic

	katab 'he wrote'			*kibir* 'he grew'		
	3rd pers.	2nd pers.	1st pers.	3rd pers.	2nd pers.	1st pers.
sg. masc.	*katab*	*katab(i)t*	*katab(i)t*	*kibir*	*kbirt*	*kbirt*
sg. fem.	*katbat*	*katabti*		*kibrat*	*kbirti*	
pl.	*katabu*	*katabtu*	*katabna*	*kibru*	*kbirtu*	*kbirna*

Table 9. Inflection of the imperfect in ʿAmmānī Arabic

	yudrus 'he studies'			*yifham* 'he understands'		
	3rd pers.	2nd pers	1st pers.	3rd pers.	2nd pers.	1st pers.
sg. masc.	*yudrus*	*tudrus*	*adrus*	*yifham*	*tifham*	*afham*
sg.fem.	*tudrus*	*tudrusi*		*tifham*	*tifhami*	
pl.	*yudrusu*	*tudrusu*	*nudrus*	*yifhamu*	*tifhamu*	*nifham*

Present and habitual actions are denoted by the use of *bi-*. In the case of the 3rd person masculine and 3rd person plural *bi-* either replaces the /y/ or is added to it. Thus, *budrus ~ byudrus, bifham ~ byifham, budrusu ~ byudrusu.* In the speech of first- and second-generation Ammanis, there is a straightforward correlation between origin and *y*-dropping: the Madanī Palestinians maintain /y/, while the Jordanians drop it. In the case of the third generation (the youth), the picture is more complicated. Here, gender becomes a factor, and the female speakers use /y/ more consistently, regardless of origin.

The future marker in ʿAmmānī Arabic is most often *raḥ ~ rāyiḥ; ḥa-* is also used but less frequently. The verb *bidd* 'to want' can also be used as a future marker, with or without 'volition' being implied, *biddi ʾarūḥ bukra* 'I will go tomorrow'.

2.2.6.2.3 Participles

Form I active and passive participles take the forms: CāCiC and maCCūC, respectively, *dāris, madrūs.* Forms II–VIII active participle derivations involve a prefix *m-, mi- ~ mu*, and /i/ in the final syllable: *mdarris, mrāgib, mindaris, mitdarris, mitrāgib, mixtalif, mistaslim.* The passive participle of these forms has a prefix *m-* and an /a/ in the final syllable: *mdarras, mrāgab, mindaras, mitrāgab, muštarak, mustaʿmal.* Form IX has one active and passive derivation with *mi-, miḥmarr, mixḍarr.*

2.2.7 Weak Verb

2.2.7.1 Geminated verbs

In verbs such as *ḥass/yḥiss* 'to feel', *šadd/yšidd* 'to tighten', the 2nd person singular masculine and the 1st person singular perfect inflect as: *ḥassēt, šaddēt.* In the active participle, they show variation between two forms: *ḥāss ~ ḥāsis, šādd ~ šādid*; the use of the latter pattern is on the increase, particularly in the speech of female speakers. Traditional Jordanian dialects, e.g. the dialect of Sult (18 km northwest of Amman), only have the former pattern, i.e. *ḥāss, šādd.*

2.2.7.2 Verbs Iʾ

Iʾ verbs show considerable variation in their imperfect (conjugations: with or without /y/ (for 3rd persons sg. masc. and pl.) and with low front /ā/ or raised back /ō/ as in the paradigm below (listed with *b-*):

The perfect conjugation is *ʾakal, ʾaklat, ʾakalu, ʾakalt, ʾakalti, ʾakaltu, ʾakalt, ʾakalna.* The participles are: *mākil, maʾkūl ~ mittākil*, and the imperative is *kul.*

2.2.7.3 Verbs I*w*

With *b-* prefix: *b(y)uwʾaf, btuwʾaf, bawʾaf, btiwʾaf, btiwʾafi, b(y)uwʾafu, btuwʾafu, bnuwʾaf.*

The imperative forms are: *waʾʾif, ʾuwsal.* The participles follow the pattern for strong verb Form I *wāʾif/mawʾūf.*

Table 10. Inflection of Iʾ verbs (imperfect) in ʿAmmānī Arabic

	yōkil ‘to eat’		
	3rd pers.	2nd pers.	1st pers.
sg. masc.	*b-ōkil* ~ *b-yōkil* ~ *b-yākul*	*b-tōkil* ~ *b-tākul*	*b-ākul*
sg. fem.	*b-tōkil* ~ *b-tākul*	*b-tōkli* ~ *b-tākli*	
pl.	*b-ōklu* ~ *b-yōklu* ~ *b-yāklu*	*b-tōklu* ~ *b-tāklu*	*b-nōkil* ~*b-nākul*

Table 11. Inflection of I*w* verbs in ʿAmmānī Arabic

	wiʾif ‘he stopped’			*wiṣil* ‘he arrived’		
	3rd pers.	2nd pers.	1st pers.	3rd pers.	2nd pers.	1st pers.
Perfect						
sg. masc.	*wiʾif*	*wʾif(i)t*	*wʾif(i)t*	*wiṣil*	*wṣil(i)t*	*wṣil(i)t*
sg. fem.	*wiʾfat*	*wʾifti*		*wiṣlat*	*wṣilti*	
pl.	*wiʾfu*	*wʾiftu*	*wʾifna*	*wiṣlu*	*wṣiltu*	*wṣilna*
Imperfect						
sg. masc.	*yuwʾaf*	*tuwʾaf*	*ʾawʾaf*	*yuwṣal*	*tuwṣal*	*ʾawṣal*
sg. fem.	*tiwʾaf*	*tiwʾafi*		*tuwṣal*	*tuwṣali*	
pl.	*yuwʾafu*	*tuwʾafu*	*nuwʾaf*	*yuwṣalu*	*tuwṣalu*	*nuwṣal*

Table 12. Inflection of II*w*/*y* verbs (perfect) in ʿAmmānī Arabic

	gāl ‘he said’			*bāʿ* ‘he sold’		
	3rd pers.	2nd pers.	1st pers.	3rd pers.	2nd pers.	1st pers.
sg. masc.	*gāl*	*gult*	*gult*	*bāʿ*	*biʿt*	*biʿt*
sg. fem.	*gālat*	*gulti*		*bāʿat*	*biʿti*	
pl.	*gālu*	*gultu*	*gulna*	*bāʿu*	*biʿtu*	*biʿna*

Table 13. Inflection of III*y* verbs (perfect) in ʿAmmārī Arabic

	rama ‘he walked’			*nisi* ‘he forgot’		
	3rd pers.	2nd pers.	1st pers.	3rd pers.	2nd pers.	1st pers.
sg. masc.	*rama*	*ramēt*	*ramēt*	*nisi*	*nsīt*	*nsīt*
sg. fem.	*ramat*	*ramēti*		*nisyat*	*nsīti*	
pl.	*ramu*	*ramētu*	*ramēna*	*nisyu* ~ *nisu*	*nsītu*	*nsīna*

2.2.7.4 Verbs II*w*/*y*

The imperfect of *ybīʿ* ‘to sell’, with *b-* prefix, conjugates as follows: *bibīʿ, bitbīʿ, babīʿ, bitbīʿ, bitbīʿi, bibīʿu, bitbīʿu, binbīʿ* (the /b/ before /n/ often assimilates to /m/ ~ *minbīʿ* ‘we sell’).

The imperative forms have long vowels: *gūl, bīʿ*. The participles follow the patterns CāyiC/maCyūC, *šāyil/mašyūl* ‘carrying/carried’, *bāyiʿ/mabyūʿ*.

2.2.7.5 Verbs III*y*

The imperative forms are *ʾinsa, ʾimši*. The participles follow the pattern CaCi/*ma*CCi *nāsi/mansi*.

The imperfect forms are *b+ yinsa, tinsa, insu, tinsa, tinsi, tinsu, ansa, ninsa; b+ yimši, timši, imšu, timši, timši, timšu, amši, nimši*.

2.3 *Syntax*

2.3.1 Noun phrase

The noun phrase has the usual constituents: nouns or pronouns + adjectives, prepositional phrases/adverbials. In addition to the definite article *al-* (or its derivatives, *il, l, i*), the numeral *waḥad/waḥde* is used to specify animate nouns, *tˁaraft ˁala waḥad ṣīni bištaġil bilmašrūˁ* 'I met a Chinese person who works on the project'. The following quantifiers are used: *kull* 'every', *kull walad* 'every boy'; *kull* 'all', *kull ilbanāt* 'all girls'; *kam/kammin/akammin ~ akkamman* 'a few', 'some', *ˁindi akammin mišwār* 'I have a few errands to run'; *šwayy* 'a little', *dilˁa zīdilha šwayyit miliḥ*, 'it is bland, add a little salt'; *kaza* 'a number of', *kaza marra ḥakētlo mā yitʾaxxar* 'a number of times I told him not to be late'. The noun phrase is negated by means of *miš or mū* positioned immediately before the negated element, e.g. *miš/mū daktōr* 'not a doctor', *miš ḥilu* 'not nice', *miš kull lbanāt* 'not all the girls'.

2.3.2 Verb phrase

2.3.2.1 Tense and aspect

Past tense is expressed through the perfect form, *daras* 'he studied'. The use of *xalaṣ* 'done', or *wxalaṣ* 'and done' can emphasize the completion of the action *xalaṣ daras, daras wxalaṣ* 'he has (done) studied'. The perfect form can combine with *kān* 'was' to denote modality (obligation) in the past, *kān darast mbāriḥ miš kān ʾaḥsan* 'you should have studied yesterday, would it not have been better?'

Like the other sedentary dialects in the region, ʿAmmānī utilizes the *b-* imperfect to express present tense as well as habitual action: *bištaġil* 'he works' or 'he is working'. The *bi*-imperfect combines with the markers *gāˁid* 'sitting', or *ˁam/ˁammāl* to express continuity, *gāˁid/ˁam bištaġil* 'he is working. These markers can be used with the bare form of the imperfect to express continuity: *gāˁid/ˁam/ˁammāl yištaġil.* Present continuous for some verbs can be expressed through the active participle *nāyim* 'he is sleeping', *gāˁid* 'he is sitting', *māši* 'I'm going/he is walking'. For other verbs, the active participle expresses perfect tense, *biḥki ʾinno mākil* 'he says he has eaten', *dāris* 'I have studied'. The verb *rāyiḥ* or its short form *raḥ* 'going' is used as a future marker, placed before the bare imperfect, *rāyiḥ/raḥ yudrus* 'he will study'. Some speakers use *ḥa-* as a future marker.

Kān placed before the *bi*-imperfect, or imperfect, expresses past continuous, *kān budrus/ yudrus* 'he was studying'; *kān* + imperfect denotes past habitual, *kān yudrus kull yōm* 'he was (in the habit of) studying every day'. *gāˁid/ ˁam/ˁammāl* + *kān* + *bi*-imperfect/imperfect also express past continuous: *kān gāˁid/ˁam/ˁammāl budrus/yidrus* 'he was studying'. *Kān* + *raḥ/ rāyiḥ* denotes past intention, *kunt raḥ amūt min ilǧūˁ* 'I was going to die of hunger', 'I was starving'.

2.3.3 Word order

Word order is most commonly SVO; VSO accompanied with rising intonation is used in interrogative sentences: *randa rāḥat ˁa ššuġul* 'Randa went to work', *rāḥat randa ˁa ššuġul*? 'did Randa go to work?'. Verb-initial sentences are also usually used when the subject is indefinite, *ʾižat waḥde ˁindhum btišbahik ktīr* 'a woman, looking very much like you, visited them'.

2.3.4 Conditional sentences

iza ~ izan, and *law* 'if' are used to introduce conditional sentences: *iza ḥaka maˁik ḥaki bahdilī ʾimsaḥi fī lʾarḍ* 'if he as much as talks to you, scold him (lit. wipe off the floor with him)', *alla ysāmḥak law talfanit kān ijitlak ˁala ṭūl* 'God forgive you. If you had called I would have come to you immediately'. Conditional meaning is implied by simply juxtaposing two clauses; the first clause usually starts with a verb in the imperative form, and the second clause begins with the verb in the *bi*-imperfect form: *ḥukkilli baḥukkillak* 'if you scratch my back, I'll scratch yours'; *dugg ˁa lbāb btismaˁ iljawāb* 'if you knock on the door you'll hear the answer', *nazzil rāsak bitšūf* 'if you lower your head, you'll see'.

BIBLIOGRAPHICAL REFERENCES

Abdel-Jawad, Hasan. 1981. *Lexical and phonological variation in spoken Arabic in Amman.* Ph.D. diss., University of Pennsylvania.

Al-Khatib, Mahmoud. 1988. *Sociolinguistic change in an expanding urban context: A case study of Irbid city, Jordan.* Ph.D. diss., University of Durham.

Al-Wer, Enam. 1991. *Phonological variation in the speech of women from three urban areas in Jordan.* Ph.D. diss., University of Essex.

——. 2000. “Raising of /a/ and related vocalic movements in the emerging dialect of Amman”. *Proceedings of the 3rd International Conference of the Association Internationale de Dialectologie Arabe*, ed. by Manwel Mifsud, 77–83. Malta: Salesian Press.

——. 2002. “Jordanian and Palestinian dialects in contact: Vowel raising in Amman”. *Language change: The interplay of internal, external and extra-linguistic factors*, ed. Mari C. Jones and Edith Esch, 63–79. Berlin: Mouton de Gruyter.

——. 2003. “New dialect formation: The focusing of *-kum* in Amman”. *Social dialectology in honour of Peter Trudgill*, ed. David Britain and Jenny Cheshire, 59–67. Amsterdam and Philadelphia: J. Benjamins.

——. Forthcoming. *Women, the vernacular and linguistic innovation: Variation and change in Jordanian Arabic.* London and New York: Routledge Curzon.

Bergsträsser, Gotthelf. 1915. “Sprachatlas von Syrien und Palästina”. *Zeitschrift des Deutschen Palästina Vereins* 38.169–222.

Cleveland, Ray L. 1963. “A classification of the Arabic dialects of Jordan”. *Bulletin of the American School of Oriental Research* 167.56–63.

Garbell, Irene. 1958. “Remarks on the historical phonology of an East Mediterranean Arabic dialect”. *Word* 14.303–37.

Jong, Rudolf E. de 2000. *A grammar of the Bedouin dialects of the Northern Sinai littoral: Bridging the linguistic gap between the eastern and western Arab world.* Leiden: E.J. Brill.

Palva, Heikki, 1984–1986. “Characteristics of the Arabic dialect of the Ḥwēṭāt tribe”. *Orientalia Suecana* 33–35.295–312.

——. 1989. “Linguistic sketch of the Arabic dialect of El-Karak”. *Studia linguistica et orientalia memoriae Haim Blanc dedicata,* ed. Paul Wexler, Alexander Borg, and Sasson Somekh, 225–251. Wiesbaden: O. Harrassowitz.

Trudgill, Peter. 1986. *Dialects in contact.* Oxford: Blackwell.

ENAM AL-WER
(University of Essex)

Juba Arabic

1. GENERAL

1.1 *Area, range, functions*

Juba Arabic is the name given to an Arabic-based variety spoken mainly in the southern part of the Sudan and more precisely in Juba, the capital city of the Equatoria region. There are indications that the Arabic varieties spoken in the other southern regions (Bahr al Ghazal and Upper Nile) are dialectally distinctive. Juba Arabic is spoken either as an interethnic → lingua franca or as a mother tongue among the members of the South Sudanese urban population who lost their ethnic native language (no number available). In Juba it is the dominant lingua franca and is used in various contexts such as administration, local courts, preaching, broadcasting, daily talk in the streets or with neighbors, and songs. Juba Arabic is a generic term applied to a wide range of social and geographical varieties that include pidgin/creole-like varieties up to decreolized or quasi-colloquial varieties. The pidgin/creole variety of Juba became ‘standardized’ and popularized in the 1970s, among other ways through the radio programs of the Sudan Council of Churches. The historical context of development (massive and quick language changes related to social upheavals in a multilingual context), the historical lack of contact with standard/colloquial Arabic, and the importance of its identity function are the main factors that explain the development and maintenance of this Arabic-based pidgin/creole as a language distinct from colloquial Arabic.

1.2 *History and development*

Contemporary Juba Arabic descends from an Arabic military pidgin, first known as Bimbashi Arabic (from Osmanli *bimbaši* ‘officer’) then Mangalla Arabic (a military garrison near contemporary Juba), that appeared in southern Sudan following the annexation of the Sudan by the Turkish-Egyptian government in 1820 (Prokosch 1986; pidginization; → creoles). The subsequent establishment of military and trade camps and the development of a large-scale slave trade between 1854 and 1889 led to major social upheavals and to the emergence of an Arabic pidgin in this previously non-Arabic-speaking area (Mahmud 1982; Owens 1997). The exact origin and form of this Arabic pidgin remains unknown. Its target language seems to have been a mixture and approximate form of Egyptian and Northern/Western Sudanese Colloquial Arabic spoken by the soldiers of the Turkish-Egyptian army. It is not certain whether a single variety spread all around southern Sudan or whether various varieties emerged simultaneously in the different camps. But the available historical and contemporary data indicate that this Arabic-based pidgin stabilized in a very short span of time, i.e. within approximately 45 years (1854–1889), and probably before the split between Juba Arabic and → Ki-Nubi.

Bimbashi Arabic continued to develop in southern Sudan, with minimal contact with standard Arabic, during the first decades of the 20th century, first in the military garrisons and urban centers and then as the main lingua franca in the heterogeneous rural areas. In spite of its policy of 'divide and rule', and its intention to stop the spread of Islam and Arabic in southern Sudan, the British condominium could not stop the development of Juba Arabic as the main lingua franca in the highly multilingual Equatorial Province. In the early 1970s, after the end of the first civil war (1956–1972) and the resettlement of the refugees, Juba Arabic became the dominant oral lingua franca for all types of communication. It thus became the mother tongue of a growing number of children in the southern urban areas due to mixed neighborhoods and interethnic marriages (Mahmud 1982). Since 1972 and more intensively since the 1980s, Juba Arabic speakers have been more exposed to Colloquial Sudanese Arabic and Modern Standard Arabic through mass media, schooling, urbanization, and migration/displacement to northern Sudan. Therefore, many Juba Arabic speakers can shift from a more creolized variety to a more colloquial one according to setting. However, the antagonistic relationship between the North and the South fosters the identity function of Juba Arabic, which is perceived as expressing and symbolizing an African identity. This is why Juba Arabic is not (yet?) threatened by a decreolization process. Today it can be heard in the displaced southern communities based in Khartoum, Cairo, or London, either in informal friendly or family settings or in more symbolic and formal settings such as artistic performances, special radio programs, and Christian religious celebrations (Miller 2002).

Juba Arabic is mainly a spoken language. It is not taught and has neither been standardized nor normalized through an official script. Its formal contexts of use include radio broadcasts, theatrical performances, songs, Christian religious preaching, some political speeches, and so on. Some prayer books are written in Juba Arabic using Latin script, while individuals educated in standard Arabic may use Arabic script to write Juba Arabic.

1.3 *State of research*

The most detailed sources for the historical context of the emergence of Juba Arabic in the 19th century are Mahmud (1979, 1982), Owens (1997), and Prokosch (1986). Many papers have discussed the linguistic origin and affiliation of Juba Arabic and Ki-Nubi and their degree of autonomy compared to other Arabic vernaculars (Owens 2001). The oldest description of the Arabic military variety is an unpublished manuscript by Jenkins 1908 analyzed by Kaye and Tosco (1993). Other sources are Thorburn (1925) and missionary books intended for the southern population. Linguistic sketches are provided by Nhial (1975), Yokwe (1985), and Watson (1984). A sociolinguistic description of Equatoria and a monographic description of Juba Arabic are given by Miller (1984). The verbal system is discussed by Mahmud (1979), Miller (1986), and Tosco (1995). Additional linguistic and sociolinguistic descriptions are found in Bureng (1986), Kaye (1994), and Miller (1993, 2001, 2003) and Manfredi (2005). Smith and Ama (1985) provide the first Juba Arabic dictionary, though without a clear distinction between Juba Arabic and Sudanese Colloquial. Hago (2001) collected a corpus of different levels of southern Arabic used in the special programs of Radio Omdurman. Most linguistic descriptions rely on data recorded in Juba or in Equatoria in the 1970s and 1980s or among displaced Equatorian speakers in Europe and Khartoum. No update information about the evolution of Juba Arabic in Juba since the 1980s is available. Much more data are needed.

2. LINGUISTIC DESCRIPTION

The following linguistic sketch presents the pidgin/creole level. It does not take into consideration social and individual variables found in the rural lingua franca or in the more decreolized varieties often used by urban speakers.

Table 1. Consonants of Juba Arabic

	bilabial	dental	alveolar	palatal	velar	glottal
plosives						
voiced	b	d			g	
voiceless	(p)	t			k	
affricates						
voiced				(j)		
voiceless				č		
fricatives						
voiced			(z)			
voiceless	(f)		s	(š)		
laterals		l	r			
nasal	m	n		ň	ŋ	
glides	w			y		

2.1 *Phonology*

2.1.1 Consonants

Comments:

i. /č/, /ň/, /ŋ/ are 'borrowed' consonants which appear in non-Arabic words: *korbač* 'whip', *ňerkúK* 'child', *ŋõ ŋõ* 'termites'.
ii. /p/ is a full borrowed phoneme in loanwords like *patrõ* 'rich', *koropo* 'leaves', but it is also a frequent variant of /f/: *filwata* ~ *palata* 'down'.
iii. /h/ is often dropped: *sehí* ~ *seí* 'right'.
iv. /z/ is often realized as an affricate /j/: *zówju* ~ *júju* 'to marry'.
v. /š/ is most often realized as /s/: *ašurubu* ~ *asurubu* 'drink'.
vi. Emphatics, pharyngeal, and velars are not preserved (**ṭalaˤ* > *tala* 'to go out', * *ġāba* > *gába* 'forest', **daxal* > *dákalu* 'to come in').
vii. Geminate consonants do not occur (**gassam* > *gésemu* 'to divide', **dagga* > *dúgu* 'to hit').
viii. Many consonants drop in final position or lose voice opposition and merge into archiphonemes. The inventory in final position includes four archiphonemes (P, T, S, K) and four other consonants (/m, n, l, r/). The archiphonemes erase the voiced/voiceless distinction and thus P = /b/ and /p/; T = /t/ and /d/, S = /s/ and /z/, and K = /g/ and /k/, as in *abyeT* 'white' (< *abyed*); *suK* 'market' (< *suug*), *keleP* 'dog' (< *keleb*).

The Juba Arabic phonological system is rather unstable. Two trends have been noticed. One, recorded among the Bari speakers of Juba region, neutralizes the plosive/affricate/fricative distinction and the alveolar/palatal distinction and gives a system with 17 consonants, close to the Bari phonological system (see Table 2).

Table 2. Phonological system of Bari speakers

b	d	j	g
p	t	s	k
	l	r	
m	n	ñ	ŋ
w		y	

The other trend, among speakers influenced by Sudanese Colloquial Arabic, maintains the alveolar-palatal distinction and keeps 20 consonants (see Table 3).

Table 3. Phonological system of Sudanese Arabic Speakers

b	d	z	j	g	(')
f	t	s	š	k	h
	l	r			
m	n		ñ	ŋ	
w			y		

2.1.2 Vowels

Juba Arabic has five phonemic vowels, /i/, /e/, /a/, /u/, /o/, and three allophones [ɛ, ?, ɔ]. These three allophones are particularly noticeable among Bari speakers and appear both in Bari words and in a number of words that reflect former phonological restructuring.

[ɛ] is the allophone of /e/ and appears in the following:

i. In Bari words such as *ñerkuK* “child”
ii. In words where Juba Arabic /e/ reflects an etymological **ē* in contact with former emphatic or pharyngeal consonants, as in *tɛr (ter < *ṭēr)* ‘bird’, *ɛna* (*ena* < **ˤēn*) ‘eye’, *sɛT* (*seT* < **ṣēd*) ‘to hunt’
iii. In words where Juba Arabic /e/ comes from an etymological **a* or **i* in contact with former emphatic, pharyngeal, or geminate consonants, as in *sɛli* (*seli* < **ṣalla*) ‘to pray’, *waɛT* (*waeT* < **wāḥid*) ‘one’
iv. In open syllables like *bɛlɛ* < *beleT* < **balad* ‘country, village’

[ɜ] is an unrounded low-mid centralized vowel, transcribed here as ö. It is the allophone of /e/ and /o/ in various syllabic contexts:

köbir	<	*kebir*	<	**kabīr*	‘big’
göliba	<	*geliba*	<	**galba*	‘heart’
köli	<	*keli*	<	**xalli*	‘to let’
köruP	<	*koruf*	<	*xarūf*	‘sheep’
wökit	<	*wokit*	<	*wakit*	‘time’

[ɔ] is the allophone of /o/ in open CvCv syllables like *kɔre* < *kore* ‘to shout’ and *bɔlis* < *bolis* ‘policeman’ and in a few Bari words, e.g. *kɔrɔpɔ/ korofo* ‘leaves’.

Long vowels are not preserved. The vowels /e/and /o/ have a wide distribution and correspond either to an etymological /ē/ or /ō/ or to an etymological /a/ and /u/ in numerous consonantal environments: *CaCīC > CeCiC **kabīr* > *kebír* ‘big’; *CacaC > CeCeC if the second or the third consonant is a lateral or a vibrant **jabal* > *jebel*; *CaCCac > CeCeC if CC is a geminate **gassam* > *gesemu* ‘to share’; *CaCuC or *CawCiC > CoCuC or CoCiC **xarūf* > *korúP* ‘sheep’, **ṭawīl* > *towíl* ‘tall’.

/ɛ/, /ö/, /ɔ/, seem to reflect the influence of an [+Advanced Tongue Root] [–Advanced Tongue Root] set common to many southern Sudanese languages (Watson 1984).

A tendency toward vowel harmony is well attested. Leaving out of account a word-final vowel, a word can have only front vowels /i/ or /e/ or back vowels /o/ or /u/. Low vowel /a/ occurs in both contexts (*ásurubu* ‘to drink’, *kátifu* ‘to write’), but in many contexts /a/ has raised to /e/ before /i/ and to /o/ before /u/ *(badri* > *bedíri* ‘early’, *majnūn* > *mojnún* ‘mad’).

Vocalic harmonization spreads to morphemes like particles *ta, fi, bi, ge: te ñerekú* ‘of the child’, *ta mára de* ‘of the woman’, *to rujál* ‘of the men’, *paláta* (> *fi-l-wáta)* ‘down’.

2.1.3 Syllables

The most common syllable types are:
v: *u* ‘and’; Cv: *fi* ‘in’; CvC: *bab* ‘door’.

Other syllable types are:
vC in initial position if the first consonant is R or L: *arda* ‘termite’, or in final position: *boíT* ‘far’ (but variant *boyiT*), CCv in final position if the second consonant is R or L: *bafra* ‘manioc’, CCvC in a few loanwords like *skul* ‘school’.

2.1.4 Stress

Juba Arabic is characterized by a lexically defined high tonal stress. It is placed on the first high tone syllable of the word. In Arabic-based words, the position of stress is often determined by the etymological syllabic structure, i.e., it falls on an etymological heavy syllable as in *sabá* (< *ṣabāḥ*) ‘morning’ vs. *sába* (< *sabˤa*) ‘seven’. Synchronically, this leads to a phonemic stress contrast. Stress also plays a distinctive grammatical role. It can help to distinguish between verbs and verbal nouns and between active verbal forms and passive verbal forms, especially for CvCvCv forms: *kéleP de ákalu* ‘the dog has eaten’ – *akúlu ta keleP* ‘the eating of the dog’ – *kéleP de akalú* ‘the dog has been eaten’.

2.2 *Morphology*

The phonological reduction leads to a complete restructuring of the Arabic lexical bases and patterns. It prohibits the productivity of a number of Arabic morphophonological processes. Juba Arabic is characterized by the use of invariable forms and independent grammatical markers. In many instances, the same form can function as a verb, a nominal, an adjective, an adverb, etc. The surrounding grammatical morphemes mark the morphological and syntactic functions. Juba Arabic has a sharply reduced morphology in comparison with standard Arabic or Arabic dialects.

2.2.1 Pronouns

2.2.1.1 Personal pronouns

Juba Arabic has two sets of personal pronouns: independent and affixed. The independent personal pronouns are given in Table 4.

Table 4. Independent personal pronouns

	singular	plural
3rd	*úwo (~úo~u)*	*úmon*
2nd	*íta (~éta)*	*ítakum*
1st	*ána*	*anína*

Independent pronouns function as subject and direct object and object of prepositions: *ána bi dúgu íta* 'I hit you', *íta bi dúgu ána* 'you hit me', *anína wóri l-ítakum* 'we showed to you'.

Possessive pronouns are suffixed to the genitive particle *bita ~ ta* and form another set, given in Table 5.

Table 5. Possessive pronouns

	singular	plural
3rd	*tó-u*	*tó-mon*
2nd	*tá-ki*	*tá-kum*
1st	*táy*	*tá-nína*

júa ta ána ~ júa táy 'my house', *dé bita itakum ~ dé tákum*? 'is it yours?'

2.2.1.2 Interrogatives

The interrogatives are *minú* 'who?', *yatú* 'who, which?', *sunú* 'what?', *malú* 'what?', *winú* 'where?', *kéf ~ kefín* 'how?', *mitín* 'when?', *kam* 'how much?', *asán senú* 'why?', *le* 'why'? Interrogatives occur sentence-initially or in the same place as corresponding non-interrogative elements: *minú báda ákalu* 'who started eating?', *ákalu táni minú* 'who else has been eating?', *grús ligó winú fi zamán zede* 'where is money to be found these days?'.

Interrogative clauses are very often marked by *seí: íta seí rúdu ána* 'do you love me?', *seí nás pojulu ge ákulu makáko* 'do Pojulu people eat monkey?'.

2.2.2 Adverbs

The most common adverbs are:

i. Temporal: *badín* 'after that', *taraó* 'then', *zamán* 'at that time', *badri* 'recently', *aléla* 'today', *ombáre* 'yesterday', *asáde* 'now', *kulu mayóm* 'every day'
ii. Local: *wára* 'behind', *fi daar* 'behind', *juwa* 'inside', *íne* 'here', *ináK* 'there', *boíT* (~ *boyiT*) 'far', *paláta* 'down', *bára* 'outside', *gidám* 'in front', *gerib* 'nearby, recently'
iii. Others: *seí seí* 'very', *ketír* 'many', *kamán* 'also', *kalás* 'only', *akér* 'better, finally', *sáwa* 'together', *swéya* 'little'

2.2.3 Particles

2.2.3.1 The particle *dé*

The particle *dé* functions as a single invariable determinative particle. It is always postposed, like all nominal modifiers, and takes the last position (see Sec. 2.2.4): *bíneya dé* 'the/this *girl*', *jerán batál dé* 'the/this/those bad neighbor(s)', *bágara táki dé* 'your cow', *keliP ta rájil dé* 'the dog of this man', *rájil kebír taláta dé* 'the three big men'.

dé functions as a clause closer in relative clauses, conditional clauses and time clauses, as in

ána	*áozu*	*kamís*	*taé*	*el*	*ána*	*kásulu*	*dé*
I	want	shirt	mine	REL	I	wash	this

'I want the shirt that I washed'

dé can function as a preposed presentative: *dé kalám kaáP* 'this is a serious issue' vs. *kalám dé kaáP* 'this issue is serious' ~ *kalám kaáP dé* 'this serious issue'.

Another presentative is the focus marker *yaú: yaú gúna kan gunaú mátara binázel* 'this is a song, if/when it is sung, rain falls'.

2.2.3.2 Genitive marker

The genitive marker is *ta: rás ta júa* 'the head of the house = the roof'. Synthetic construct states are only found in lexicalized compound words: *jena murkáka* 'grindstone'.

2.2.3.3 Negative marker

The negative marker is *má* for nominal and verbal clauses. *má* precedes the predicate: *wéle dé má suker* 'this boy is not small', *ána má bíyu le zól dé* 'I did not sell (it) to this man'.

Negative existential *máfí* can be postposed to a negative clause to emphasize negation: *ána má dúgu zól máfí* 'I did not hit this man at all!'.

In imperative clauses, the negative marker is *máta: máta gésimu badúm* 'don't divide between yourselves!'.

2.2.3.4 Existential particles

The main existential particle is *fí: fí nás ketír fi júba* 'there are many people in Juba'. Existential constructions can be used to express possession: *fí grús le éta* 'there is money for you = you have money'.

2.2.3.5 Prepositions

Prepositions are *fi* 'in, at, to', *le* 'to', *ma* 'with, by', *madúm* 'without', *foK* 'on', *min* 'from, through', *tat* (*tibit*) 'under', *lakadi* 'until', *wára* 'after, behind', *gidám* 'in front', *zay (~ze)* 'like', *ila* 'except'. A number of prepositions are made of two elements: *fi rás* 'on', *fi dahar* 'behind', *fi nús* 'in the middle', *fi júwa* 'inside', *fi tibit* 'under', *fi gidám* 'in front', *fi bátina* 'among'.

The distribution of *fí* and *li* after verbs of movement indicates a distinction between a human and a non-human destination: *ána ge fútu fi béT/ fi súK/fi kór* 'I go to the house/the market/ the river', *ána ge fútu le-íta/le binéya/ le mára de/* 'I go to you/the girl/the woman'.

2.2.3.6 Conjunctions

Conjunctions are *wa ~ u* 'and', *wala* 'or', *lakín* 'but', *asán keli* 'therefore', *asán kede* 'in order to', *kan* 'if, when', *wokit* 'when', *gal* '[dependent clause] that'.

2.2.3.7 Definite particles

The definite particle *el (~ al ~ ale)* introduces verbal or nominal expansions; e.g. *dé júju ale adíl* 'this is a regular marriage', *el gúm dé bára dé zeríya* 'the ones coming out are germs', *binéya dé el ge wónosu ma ítakum dé ísom tó senú* 'this girl, who comes to talk with you, what is her name?', *uwo bikúbu matára el kan uwo kútu fi búrsa tó dé* 'he pours the rainwater that he has put in his water-skin'.

Another definite particle is *abu: úo bisílu ay bíneya abu úo áozu* 'he takes any girl he wants' (see Sec. 2.2.4).

2.2.4 Noun

Nouns and verbs are distinguishable through three main features:

i. Word order. Nouns are characterized by postposed nominal modifiers with the following word order: {N. Gen}, {N. Poss}, {N. Adj} {N. Rel} {N. Det}, while verbs are characterized by preverbal modifiers.
ii. Stress (see Sec. 2.1.4)
iii. Word ending. Most nouns end in *-a* or *-i* or -C, while approximately 50 percent of verbs end in *-u*: *fíli* 'an elephant', *bokra* 'tomorrow', *dúgu* 'to hit', *ámusuku* 'to take', etc. Juba Arabic tends to develop a productive *-a/-u* inflectional opposition for trisyllabic words: *setíma/sétimu* 'an insult/to insult'.

Arabic nominal patterns are not productive, although a number of words reproduce inherited patterns such as localities (*máal* 'place', *matár* 'airport'), instruments (*málaga* 'spoon', *máfata* 'key'), and some professions (*múdir* 'director', *najár* 'carpenter'). Most nouns derive from etymological singular forms, but a set of nouns derive from etymological internal plural patterns which have lost their meaning: *jerán* 'neighbor(s)', *sunún* 'teeth/tooth'.

Analytical constructions render professions, localities, adjectives, etc.:

Arabic	Juba Arabic	
ṣayyād	*zól ta dugu lahám*	'hunter'
mazraʿa	*mahál ta kurúju*	'farm, field'
ʿanīf	*zól ta dosomán*	'violent; troublemaker'
midagg	*jéna fundùK*	'pestle'

There is no gender distinction. Plural is often marked for nouns, but number agreement is irregular for adjectives. Very few cases of internal plural have been recorded:

singular	plural	
mara	*nuswán*	'woman/women'
âku	*akwán*	'brother(s)'
askári	*asakér*	'soldier(s)'
kebír	*kubár*	'big'

When expressed, plural is marked by the suffixes *-át* for nouns and *-ín* for adjectives and a few nouns: *ñerkúK* 'child'/*ñerkukát* 'children', *mobsút/mobsútín* 'happy [sg.Pl.]', *-in* is also often suffixed to internal plural: *kebír* 'big'/ *kubár ~ kubárin* 'big [sg.Pl.]', *áku/akwán ~ akwánín* 'brother/brothers'.

The word *nás* (*nasi*) 'people' + noun expresses a general plural, a collective, or a type marker, as in *nás ardá* 'termites', *nás silá* 'weapons', *nás dupi* 'the class of the slaves', *nás dúra* 'sorghum', *nás móya* 'water'. With noncountable objects, it often carries the meaning of large quantities; e.g. *súK yambio dé kebír seí seí ibiyú fógo nási* *gomása nási dúra nási merísa nási senú senú* 'the market of Yambio is very big, they sell a large [quantity of] materials, sorghum, local beer, and many other things'. The diminutive is expressed by *sukér* 'small': *júa sukér* 'tiny house', *binéya sukér* 'small girl'.

Most adjectives are built on CaCCán, CeCíC, and mvCCvC patterns: *takyán* 'nervous', *kebír* 'big', *moksút (mobsút)* 'happy'. Adjectives do not form a specific word category. They can be nominalized with the definite particle *al: al báreT to* 'the cold part of it', or they can be preceded by the two verbal markers *bi* and *gi* (see Sec. 2.2.6): *úwo tabán* 'he is tired', *úwo gi tabán* 'he is suffering', *úwo bi tabán* 'he will become tired'.

The definite particle *abu* before nouns, adjectives, or numerals expresses an inalienable possessive state: *móz abu nedíf dé* 'bananas of good quality', *gumasát abu seretú* 'torn clothes, rags', *arági abu ta dūra* 'a sorghurn-based alcohol', *anína rúwa abu síta* 'we went to six people'. Bodily defects are mainly expressed with the *abu* construction: *abu éna* 'one-eyed', *abu sála* 'bold', *abu dáŋa* 'paraplegic', etc. Color adjectives have an inherited form: *abye* 'white', *aswáT* 'black', *amár* 'red', etc., but they are often introduced by *abu: bágara abu abye* 'the white cow', *gélib abu aswáT de* 'a black heart'.

Comparison and superlative are expressed by analytical constructions using the verb *fútu* 'to go; to overtake', or the topic marker *yaú*, e.g. *Juba fútu Malakal* 'Juba is bigger than Malakal', *úwo kebír fútu éta* 'he is bigger than you', *kebír tómon yaú Tomáya* 'Tomáya is the biggest'.

2.2.5 Numerals

Cardinal numerals are 1–10: *wáhid ~ wáe, itnín, taláta, árba, kámsa, síta, tamániya, tísa, ásara*; 11–19: *idáser, itnáser, talátaser,* etc. 20–90: *isrín, talatín,* etc. 100: *miya (mia)* 1,000: *alef.* Digits precedes tens: 21 = *wáe u asrín.*

2.2.6 Verbs

The majority of the basic verbal stems have a CvCv or CvCvCv or aCvCvCv syllabic pattern: *gáta* 'to cut', *kátifu* 'to write', *asurubu* 'to drink', etc.

The majority of the verbal stems seem to derive from an etymological 3rd (pl. or sg.) perfect or imperative Form I stem with further phonological restructuring. The basic verbal stem remains invariable but can be pronounced differently according to speakers: *katabu ~ katibu ~ katifu ~ kataP* 'to write'.

There is no morphological vowel alternation except in one case: *ákulu* 'to eat'/*ákeli* 'to feed' (< *ʾakkal*). The only attested derived verbal pattern is a limited set of verbs with an opposition between a basic CuCu form and an expended Cow(u)Cu form expressing a transitive/factitive meaning:

gúmu 'to stand up'	*gówmu (gowumu)* 'to raise, to make rising, to make germinate'
númu 'to sleep'	*nówmu (nowumu)* 'to put to sleep'
fúru 'to boil'	*fówru (fowuru)* 'to get to boil'

As mentioned in Section 2.1.4, the shift of stress expresses an active/passive-like distinction and a verb/verbal noun distinction, particularly with CvCvCv verbs. In passive-like constructions, the agent can be introduced by the particle *ma* 'with'. This type of construction does not function with a human agent: *júa al kan durubú ma sága de salaú* 'the house that was destroyed by the storm was rebuilt'.

Person and number are marked by external personal pronouns (see Sec. 2.2.1). Tense, Moods and Aspect are expressed by two core verbal particles (affixed to the bare form), *bi* and *gi*, and by a number of verbal semi-auxiliaries: *kán, áozu, já, rúwa, gúm, téde,* and a few adverbs: *baga, kalás.*

2.3 *Verbal phrase*

Juba Arabic has developed a Tense-Mood-Aspect system through the combination of the various particles. Juba Arabic divides its verbal lexicon between statives and nonstatives (Tosco 1995). Injunctive, factitive, intensive, iterative, and other meanings are expressed through analytical constructions.

The bare verbal stem is the general unmarked form. In some functions (perfect and imperative) it can alternate with a more specific Tense-Mood-Aspect marker. It expresses:

i. A past or perfect action with nonstative verbs in main sentences: *nás ákulu lahám de kúlu* 'the people ate/have eaten all the meat' vs. *nás ákulu lahám de kúlu kalás* 'the people have eaten all the meat'
ii. A present tense with stative verbs in main sentences: *ána ma áozu kalám ketír* 'I don't want trouble'

iii. An imperative: *ákalu lahám de* 'eat the meat!'; *ákalu tákum lahám de* 'eat [pl.] the meat!'

iv. An infinitive in subordinate sentences: *bókora ána birúwa fátisu dáwa* 'tomorrow I will look for medicine'

The particle *bi* functions as an irrealis marker and is used to express:

i. A general present or an habitual action: *nás bigúm sabá bádri* 'people get up early'

ii. A future or possible/potential/conditional action: *bókra anína bikúruju* 'tomorrow we will cultivate'; *úwo bikátalu áku tó kalám grús* 'he could kill his brother for money'; *kan nás rúwo fi gába úmon bikátalu nas gém ketír* 'if/when people go to the forest, they kill a lot of game'

The particle *ge* functions as a continuous marker and expresses a continuous, habitual, iterative action: *báhar de ge mála kuluyóm* 'every day the river gets a bit higher'; *fi keríf nás ge kúruju* 'during the rain season people cultivate'; *zamán fi wokít zedé nási ga ákulu lahám* 'before, at that time people used to eat meat'. *bi* and *ge* are found in a number of similar semantic contexts (e.g. habitual or general present), and they never combine.

kán marks a past tense with stative verbs and an anterior with nonstative verbs: *úwo kán mási u lígo tumsá kebír* 'he went and found a big crocodile'; *úmon kán ákulu ketír* 'they have been eating a lot'. *kan* + *bi* marks an anterior irrealis: *w-ána kan bi rákabu senú lakín* 'but what could I cook?'. *kan* + *gi* marks a continuous past action: *wókit ána já nás kan ge álabu* 'when I came, people were dancing'. Both can mark a past habitual: *zamán nás kán gi* (~ *kán bí*) *kúruju ketír* 'before, people used to cultivate a lot'.

Other verbal Tense-Mood-Aspect markers include the following semi-auxiliaries:

i. *rúwa* 'go' combined with *bi* or *gi* expresses a near future but still implies the notion of movement, as in *ána gi rúwa núm* 'I am going to sleep'; *íta biróa wódi senú li sabí táki* 'what are you going to give to your friend?'.

ii. *áozu* 'want' expresses a near future without movement: *úwo áozu mútu* 'he is about to die'; *ñerkukát áozu ámulu subián* 'the children are about to become teenagers'.

iii. *já* 'come' before a bare verb can express a perfective (resultative) and has the same meaning as *kalás: úwo já ába/ úwo ába kalás* 'he has refused'.

iv. *bága* ~ *bíga* 'then, so, after' is a (resultative) perfective and an inchoative and indicates a change of state or action: *u bága askut gidám ta sultán* 'so he kept silent in front of the chief'; *kán mátara názal nás bága bikuruju* 'if it rains, then people will cultivate'.

v. *gum* 'stand' and *bitede* ~ *tede* 'start' indicate the beginning of an action: *wókit ñerkukát del lisa sukerín azigu gúm mútu* 'the children were still young when Azigu [suddenly] died'; *nás bitede kúruju* 'if it rains, people start to cultivate'.

vi. *lisa* marks an ongoing action in affirmative sentences: *úwo lisa gi bíyu skot jedíd* 'she is still buying a new skirt' vs. *úwo lisa ma biyu skot jedíd* 'she has not yet bought a new skirt'.

vii. Injunctive/imperative/orders are expressed either by the bare form (see above) or by *tákum: rúwa tákum fi súK* 'go to the market!', or by *keli* 'let' as *in kel-ita áskut* 'keep quiet'; *kel-ina ábinu béleT tanína* 'let us build our country!'.

viii. *kéli* is often translated as an optative: *kéli nás rúa sáyidu sámaK* 'people should go fishing', *kéli* introduces a complement clause after verbs of request as in *ána bi ásalu úwo keli rákabu sámaga dé* 'I will ask her to cook the fish'; *ánina áwuju kéli jówju íta ma wélet dé* 'we would like to marry you with this boy'.

ix. *kéde* expresses a request: *kéde jíbu lána merísa* 'can you bring me a beer?'.

x. *lazim* expresses a need, an obligation: *ánina lázim ábini béle tanína* 'we should build our nation'.

Verbal reduplication expresses durative, iterative, or habitual meaning (Miller 2003), as in *bágara de ge mútu mútu* 'cows are dying one after the other'; *mára de ge ákulu ákulu fi júa* 'this woman always eats in her house [i.e., she is a selfish woman]'.

Intensive is expressed by the following construction: *ána fáta báb ma fata* 'I open the door by force', *ána jére ma jere* 'I am running [like a runner]'.

Reciprocity is expressed by *badum: rijál dúgu badúm* 'men fought between themselves'.

Reflexive is expressed by *nefisa: kéli báb gáfalu nefísa bitá* 'let the door close by itself'.

kutu + verb expresses a factitive: *ána kútu úwo ásurubu* 'I make him drink'.

Complement clause and adverbial clauses are introduced by *gale* 'say' and *keli* 'let', which have lost their semantic meaning and become grammaticalized (Miller 2002): *nás kúlu ge sáadu ána asán keli ána ábinu júa táe* 'all the people help me to build my house'; *úmon ge ába gále kéli nás ma kátulu ayawanát* 'they refuse to let people kill animals'.

3. Lexicon

Approximately 80–85 percent of the lexicon derives from Arabic roots; the remaining 20 percent derives from local African vernaculars (Bari, Moro, etc.), English, Swahili, Bangala, and so on. Compound words seem less developed than in Ki-Nubi. Analytical constructions combining a single verb and various verbal extensions are frequent, e.g. *dúgu* 'to hit'; *dúgu ma tín* 'to plaster', *dúgu ma bóya* 'to paint', *dúgu bonya* 'to box', *dúgu paláta* 'to demolish'. Many idiomatic expressions appear to be transfers of vernacular expressions, e.g. *túfu buzaK* 'to spit [i.e. to give one's blessing]', *gafúlu ta báb* 'closure of the door [i.e. engagement]'.

More research is needed to verify the influence of local vernacular languages at the syntactic and semantic levels. If the lexicon is dominantly Arabic-based, the grammar of Juba Arabic exhibits considerable restructuring and innovations.

Bibliographical references

Bureng, George Vincent. 1986. "Juba Arabic from a Bari perspective." *Current approaches to African linguistics*, III, ed. Gerrit J. Dimmendaal, 71–78. Dordrecht: Foris.

Hago, AbdelMoneim Nasir. 2001. *Le juba arabic au Nord Soudan*. Mémoire de DEA, University of Paris-3.

Kaye, Alan. 1994. "Peripheral dialectology and Arabic pidgins and creoles". *Actas del Congreso Internacional sobre Interferencias Linguisticas Árabo-Romances y Paralelos Extra-Iberos, 10–14 diciembre 1990*, Madrid, ed. Jordi Aguadé, Federico Corriente, and Manuela Marugan, 125–140. Madrid and Saragossa: Navarro y Navarro.

—— and Mauro Tosco. 1993. "Early East African Pidgin Arabic". *Sprache und Geschichte in Afrika* 14. 269–305.

Mahmud, Ushari. 1979. *Variation and change in the aspectual system of Juba Arabic*. Ph.D. diss., Georgetown University.

——. 1982. *Arabic in the southern Sudan: History and spread of a pidgin-creole*. Khartoum: FAL Advertising and Printing Cie.

Manfredi, Stefano. 2005. *Descrizione grammaticale dell'arabo Juba (Sudan Meridionale) con riferimenti sociolinguistici alla communità sudanese del Cairo*. Naples: Università di Napoli La Sapienza.

Miller, Catherine. 1984. *Etude socio-linguistique du développement de l'arabe au Sud Soudan*. Ph.D. diss., Sorbonne Nouvelle University of Paris-3.

——. 1986. "Un exemple d'évolution linguistique: Le cas de la particule *ge* en Juba-Arabic". *Matériaux Arabes et Sudarabiques*. 155–166.

——. 1993. "Restructuration morpho-syntaxique en Juba-Arabic et ki-nubi: À propos du débat universaux/superstrat/substrat dans les études créoles". *Matériaux Arabes et Sudarabiques Nouvelle Série* 5.37–174.

——. 2001. "Grammaticalisation du verbe dire et subordination en Juba arabic". *Leçons d'Afrique: Filiation, rupture et reconstitution des langues. Un hommage à G. Manessy,* ed. Robert Nicolaï, 455–482. Leuven: Peeters.

——. 2002. "Juba Arabic as a way of expressing a southern Sudanese identity in Khartoum". *Aspects of the dialects of Arabic today*, ed. Abderrahim Youssi a.o., 114–122. Rabat: Amapatril.

——. 2003. "Reduplication in Arabic-based language contact". *Twice as meaningful: Reduplication in pidgins, creoles and other contact languages*, ed. Silvia Kouwenberg, 289–299. London: Battlebridge.

Nhial, Abdon J. 1975. "Ki-Nubi and Juba-Arabic: A comparative study". *Directions in Sudanese linguistics and folklore,* ed. Herman Bell and Sayyid Hamid Hurreiz 81–93. Khartoum: Institute of African and Asian Studies.

Owens, Jonathan, 1997. "Arabic-based pidgins and creoles". *Contact languages: A wider perspective*, ed. Sarah G. Thomason, 125–172. Amsterdam and Philadelphia: J. Benjamins.

——. 2001. "Creole Arabic: The orphan of all orphans". *Anthropological Studies* 43. 348–378.

Prokosch, Erich. 1986. *Arabische Kontaktsprachen in Afrika*. Graz: Institut für Sprachwissenschaft, Universität Graz.

Smith, Ian and Morris Ama. 1985. *A dictionary of Juba Arabic and English*. Juba: Committee of the Juba Cheshire Home and Centre for Handicapped Children.

Thorburn, D.H. 1925. "Sudanese soldiers songs". *Journal of African Studies* 24.314–321.

Tosco, Mauro. 1995. "A pidgin verbal system: The case of Juba Arabic". *Anthropological Linguistics* 37.423–459.

Watson, Richard. 1984. *Juba Arabic for beginners*. Juba: Summer Institute of Linguistics.

Yokwe, Eluzai M. 1985. "The diversity of Juba-Arabic". *Studies in African Linguistics, Supplement* 9.323–328.

Catherine Miller
(IREMAM, University of the Provence)

Judaeo-Arabic

The term 'Judaeo-Arabic' refers to a type of Arabic that was used by Jews and was distinct in some way from other types of Arabic. It is by no means a uniform linguistic entity and is used to refer both to written forms of Arabic and also to spoken dialects.

The Arabic language was used by Jews in Arabia before the rise of Islam. Some of the pre-Islamic Arabic poets were Jewish, the most famous of whom was as-Samaw'al ibn ʿAdī. The surviving written works of those Jewish poets do not exhibit anything that distinguishes them from the equivalent works of their non-Jewish contemporaries, and so they are generally not referred to as Judaeo-Arabic. It is assumed that the Jewish communities in Arabia spoke Arabic as their vernacular language. Although we do not have any direct evidence of the nature of this spoken language, some scholars claim there are indirect indications that it differed from the vernacular of the non-Jews mainly by the presence of Hebrew and Aramaic lexical elements transferred, through cultural contact, to the non-Jewish population and appearing in the *Qur'ān* and other early Arabic literature.

After the Islamic conquests in the 7th century C.E., the Arabic language gradually spread throughout the Near East. It was initially restricted to the Arab invading armies but soon began to be used by the local population. This applied not only to converts to Islam but also to Jews and Christians, who maintained their religion and traditional communal life. The Arabicization took place most rapidly in the large urban centers, where the Arab armies had settled and established centers of administration. In the pre-Islamic period, the Jews of Iraq and Syria spoke Aramaic, whereas further west they used Berber or Romance as their vernacular language. These languages were largely replaced by Arabic. The Jewish communities in rural areas were much slower to adopt the Arabic language. Although the Jews of the urban centers in Iraq appear to have become Arabic-speaking by the 8th century C.E., there is evidence that the Jews in the countryside continued to speak Aramaic at least until the 10th century. Some Jewish communities living in the isolated mountainous areas of northern Iraq never fully adopted Arabic as a vernacular and continued to speak Aramaic until modern times. A similar pattern applied to the spread of Arabic elsewhere in the Islamic empire.

During the first three centuries of the Islamic period, the Jews in the Near East used the traditional rabbinic languages of Hebrew and Aramaic as their written languages, although many of the urban communities were no doubt using Arabic as their vernacular in this period. One factor that may explain the slowness of the Jews to use Arabic as a written literary language was that the main centers of Jewish learning, such as the academies of Sura and Pumbeditha, were situated in the Iraqi countryside, where Aramaic remained the spoken language for a longer period (Fenton 1990:464). The earliest surviving records of Judaeo-Arabic are datable to the 8th or 9th century C.E. They were written in Hebrew script, which became one of the most conspicuous distinctive features of written Judaeo-Arabic. Thereafter, Arabic in Hebrew script continued to be used by Jews in Arabic-speaking lands throughout the Middle Ages down to modern times. The term 'Judaeo-Arabic' is frequently used to refer to all such cases of Arabic written in Hebrew script. This is based on a descriptive criterion, namely its graphic representation, and also, by implication, one of communicative function, since anything written in Hebrew script would, one would assume, be addressed to a Jewish readership.

Judaeo-Arabic in this sense, i.e. any form of Arabic written in Hebrew script, is not a linguistically uniform phenomenon. It is generally categorized into three chronological periods corresponding to three major phases in its linguistic development, namely Early Judaeo-Arabic, Classical Judaeo-Arabic, and Late Judaeo-Arabic.

The term 'Early Judaeo-Arabic' is used to refer to Judaeo-Arabic that was written before the 10th century. This material has come to light only in the last few decades. It consists of private documents on papyrus and some manuscript fragments of literary texts. These texts are datable to at least the 9th century and some possibly earlier.

The period of Classical Judaeo-Arabic began in the 10th century. During this period, Judaeo-Arabic was used in a very wide range of texts. Many of the traditional texts of Judaism were translated into Judaeo-Arabic, including first and foremost the Hebrew Bible, but also other texts such as the Mishnah, Talmud, Midrashim, and liturgy. Many new genres of Arabic text were adopted by the Jews from the Muslim cultural environment and adapted to Judaism. This reflected a close rapprochement between the Jews and Muslim culture in the High Middle Ages (approximately 10th–13th centuries C.E.). The new genres of texts included works on biblical exegesis, grammar, systematically arranged handbooks of legal subjects, and works on theology and philosophy. Judaeo-Arabic was also used for a wide range of documentary material. Most letters were written in Judaeo-Arabic and also a large proportion of Jewish legal documents. Hebrew was still used as a learned language in letters by some Jewish intellectuals, such as the Geonim. It was also used by the leading Jewish poets in the Middle Ages, but many popular verses and songs were composed by Jews in Judaeo-Arabic.

In the Late Judaeo-Arabic period, the range of texts written in Judaeo-Arabic became more restricted. Among the factors that brought this about was a decrease in intellectual rapprochment between the Jewish communities and their Muslim environment. In many of the regions of the Near East, the beginning of this period can be located in the 15th or 16th centuries, when Spanish and Portuguese Jewish refugees from the expulsions and their descendants came to be among the leading intellectuals in the Arabic-speaking Jewish communities. As a result, Hebrew was used in these communities for the composition of many literary texts. Judaeo-Arabic became restricted largely to popular texts such as stories and songs or private letters. Another common type of Judaeo-Arabic text in this period was a literal translation of the Bible and a few other traditional Jewish texts known as *šarḥ*. This was a word-by-word gloss that generally could not be understood independently of the original Hebrew source text. The medieval Judaeo-Arabic Bible translations ceased to be used in most Arabic-speaking Jewish communities and were supplanted by the *šarḥ*, the language of which was much closer to the local vernacular spoken dialect. The printing press gave an impetus to new genres of Late Judaeo-Arabic. In the 19th century, for example, Judaeo-Arabic newspapers were produced in several Arabic-speaking Jewish communities. Avishur (1986:3) has proposed that the beginning of the Late Judaeo-Arabic period in Iraq should be located in the 13th or 14th century, after the devastations of the Mongol invasions. In the Jewish communities of Yemen, Classical Judaeo-Arabic texts continued to be copied and read down to modern times, and the division between classical and late periods of Judaeo-Arabic is not as appropriate.

One of the main distinctive linguistic features of Early Judaeo-Arabic is the orthography with which Arabic is represented. It is a phonetic spelling representing the way the writers pronounced the language based on the orthographic practices used for rabbinic Hebrew and Aramaic at that period, rather than those of Classical Muslim Arabic in Arabic script. This is particularly noticeable in the use of vowel letters, e.g. the defective spelling of long /ā/ (סלם = سلام 'greeting', עפיה = عافية 'health') and the *plene* spelling of short /i/ and /u/ (אלחיכמה = الحكمة 'wisdom', אילא = الى 'to', תהרוב = تهرب 'you flee'). The letters *ḍād* and *ḏ̣ā*ʾ, which had no direct equivalent in the Hebrew consonantal inventory, were represented by the letter *dalet̠*, the nearest phonetic equivalent, e.g. יקבדוה = يقبضه 'he will receive it', עדה = عظة 'admonition'. The *lām* of the definite article was not represented when it was assimilated to the following letter, e.g. אסלם = السلام 'the greeting'. *Tā*ʾ *marbūṭa* was represented by *taw* when it was pronounced /t/ in a word in an annexation construction, e.g. עידת אלחיאה عظة الحياة 'the admonition of life'. Examples are from Blau (2002:136–154).

In Classical Judaeo-Arabic, which was used in most Arabic-speaking Jewish communities from the 10th to approximately the 15th centuries, the spelling used was made to correspond to the orthographic conventions of Classical Arabic. Long vowels were regularly represented by vowel letters, whereas short vowels were spelled defectively without vowel letters, e.g. סלאם = سلام 'greeting', אלחכמה = الحكمة 'wisdom', תהרב = تهرب 'you flee'. Long /ā/ was generally spelled defectively in the small set of words where this was the norm in Classical Arabic orthography, ד̇לך = ذلك 'that'. Final long /ā/ was represented by *yod̠* where Classical Arabic orthography had *ʾalif maqṣūra* spelled with *yā*ʾ, e.g. אלי = الى 'to'. The *lām* of the definite article was regularly represented, including where it was assimilated to the following consonant, e.g. אלסלאם = السلام 'the greeting'. *Tā*ʾ *marbūṭa* was represented by *heh* in all contexts, including when pronounced [t] in annexation constructions, e.g. מערפה אלברהאן = معرفة البرهان ‘knowledge of the proof’. The Arabic letters *ḍād* and *ḏ̣ā’* were represented respectively by *ṣaḏe* and *ṭeṯ*, with a superscribed dot in imitation of the Arabic alphabet, e.g. יקבצ̇ה = يقبضه ‘he will receive it’, עט̇ה = عظة ‘admonition’. The Arabic alphabet was not, however, imitated where the sound existed in Hebrew. The Arabic letters *xā’* and *ġayn*, for example, were represented by Hebrew *kaf* and *gimel*, often with diacritical marks (כ̇, ג̣), rather than *ḥeṯ* and *ʿayin* with diacritical marks. This is because the pronunciation of the fricative allophones of the Hebrew letters *kaf* and *gimel* corresponded to that of the Arabic letters in question.

In most forms of Late Judaeo-Arabic, which began to be used roughly after the 15th century C.E., scribes abandoned a rigorous imitation of the orthography of Classical Muslim Arabic and, as in the Early Judaeo-Arabic period, employed many of the conventions of spelling that were used for rabbinic Hebrew and Aramaic. Short /i/ and /u/ vowels were frequently represented with vowel letters, e.g. = איבני ابنى ‘my son’, קולת = قلت ‘I said’. Where Classical Arabic orthography had *’alif maqṣūra* written with *yā’*, the late texts often used the Hebrew vowel letter *heh*, in conformity with Hebrew/Aramaic orthography, e.g. עלה = على ‘upon’. Some of the orthographic conventions of Classical Arabic do, however, appear in Late Judaeo-Arabic. These are likely to be vestiges of Classical Judaeo-Arabic usage rather than direct imitations of Muslim Classical Arabic. Late Judaeo-Arabic is a diverse corpus of material from different regions, and the degree to which such features are found is not uniform across all texts. The following is the situation that is found in texts from 17th- and 18th-century Egypt. In many of these texts long /a/ is regularly written with the vowel letter *’alep̄*, e.g. קאלו = قالوا ‘they said’, קאעיד = قاعد ‘[he is] sitting’. The letter *ḍād* is generally represented by *ṣaḏe* with an upper diacritic, e.g. יחצ̇ר = يحضر ‘he attends’. The *lām* of the definite article is regularly represented, even when it is assimilated to the following consonant in pronunciation, e.g. אל נאס = الناس ‘the people’, אל רחמאן الرحمن = ‘the Merciful’. Examples are from Khan (1992).

The use of Hebrew script for writing Arabic was taken over from the Hebrew and Aramaic literary tradition of the Jews. The language was changed, but the traditional script continued. The different degrees with which the orthography of Arabic was adopted reflect the different degrees of rapprochement between Jewish and Muslim literature and culture at the various periods. This rapprochement was at its greatest in the High Middle Ages (approximately 10th–13th centuries). Indeed, at this period, certain circles of Jewish scholars belonging to the Karaite movement of Judaism wrote some of their Arabic works in Arabic script, even when intended only for a Jewish readership. This reflects a particularly advanced degree of absorption into the Islamic cultural environment. In addition to being determined by external influences, orthographic practices in Judaeo-Arabic were affected also by literary models and traditions internal to Judaeo-Arabic. In the Middle Ages the widespread adoption of Classical Judaeo-Arabic orthography appears to have been stimulated by Saadya Gaon’s use of this type of orthography in his Arabic translation of the Pentateuch (Blau and Hopkins 1984). Late Judaeo-Arabic orthography, as remarked, carried over some vestiges from Classical Judaeo-Arabic. In Yemen the tradition of Classical Judaeo-Arabic remained robust in the later period, with the result that its orthography stayed largely unchanged down to modern times.

In addition to differences in orthography in these three periods, there were also differences in grammatical structure. The writers of the Judaeo-Arabic texts spoke Arabic dialects as their vernacular. The language of the texts, however, is not a direct representation of these dialects. During all periods the writers attempted to produce a literary form of language that was distinct from their vernacular. In the Middle Ages, in both Early and Classical Judaeo-Arabic texts, the standard that the writers aimed at was that of Classical Arabic, or at least the post-Classical Arabic that was the normal form of literary expression in the cultural environment in which they lived. In the Late Judaeo-Arabic texts, the literary model was generally that of contemporary substandard genres of writing, which tended to differ from region to region. In all periods of Judaeo-Arabic, however, vernacular dialectal features appear in the texts. In the Middle Ages this came about because the writers fell short of their goal of conforming to their literary standard, producing a form of language that is sometimes referred to as → Middle Arabic. In the later period, the literary model itself was a substandard form that contained a high proportion of local dialectal elements, so the presence of dialectal elements in Late Judaeo-Arabic is far greater than in the medieval texts. On account of this extensive local dialectal component of Late Judaeo-Arabic, the language of the texts differs in the various regions in which it was written, e.g. North Africa, Egypt, Syria, and Iraq, whereas such regional differentiation is not as pronounced, or indeed is sometimes entirely absent, in the medieval texts. In the Late Judaeo-Arabic texts, the local vernacular dialects usually form the base of the language, to which are added a limited number of nonvernacular elements to raise the register above that of pure vernacular. In medieval Judaeo-Arabic, on the other hand, the base of the language, or at least the intended base, can be said to be Classical Arabic, and deviations from this base occurred by interference from the spoken vernacular.

In the Middle Ages the extent to which the language deviated from Classical Arabic varies from text to text. The same author sometimes used different degrees of vernacular elements according to the readership. This is seen, for example, in the writings of Maimonides (1135–1204), whose extant letters to private individuals tend to contain more vernacular elements than his literary works, which were intended for an educated readership. This continuum of the degree of vernacular admixture relative to the literary standard language has been termed by Hary (1992) 'multiglossia'. It should be noted that vernacular elements have sometimes entered Judaeo-Arabic literary texts in the course of scribal transmission, and the extent of their occurrence sometimes differs among the manuscripts of a single work.

The degree of vernacular interference in the language of Classical Judaeo-Arabic texts is disguised somewhat by the orthography of the texts. As remarked, this was an imitation of the spelling practices of Classical Arabic and so did not in principle indicate deviations from the Classical Arabic relating to vowels and syllable structure. Several extant manuscripts datable to the medieval period that are supplied with Hebrew vocalization signs reveal numerous dialectal features that would not be apparent in unvocalized texts. This can be illustrated by a few examples taken from the Genizah manuscript (Cambridge University Library, T-S Ar. 8.3): וְאַנַא עַבְדַךְ *wa-ʾana ʿabdak* 'and I am your servant' (invariable 2nd pers. masc. sg. pronominal suffix *-ak*); עַלֵי עִבֵּאדַךְ *ʿalē ʿibēdak* 'over your servants' (invariable 2nd pers. masc. pronominal suffix *-ak* and *ʾimāla* of long /ā/ vowels); וּבַעַד מַוּתֵהּ *baʿad mawtu* 'after his death' (invariable 3rd pers. masc. sg. pronominal suffix *-u*); חַתֵּי יִפְתַח עַיְנֵהּ *ḥattē yiftaḥ ʿaynu* 'until he opens his eye' (/i/ in verbal prefix, lack of verbal mood ending, invariable 3rd pers. masc. sg. pronominal suffix *-u*); יִנְתִטר *yintiẓir* 'he waits' (/i/ vowels in prefix and verbal base). Some of these dialectal features are visible in the orthography of the Early Judaeo-Arabic texts, which indicated some of the short vowels by vowel letters and was generally more phonetically based than the orthography of Classical Judaeo-Arabic.

The standardized orthography of Classical Judaeo-Arabic could be read with a variety of different vernacular vocalisms, so many of the regional dialectal differences of the writers were not manifested in the texts. This facilitated its use as a literary koine language across all Arabic-speaking Jewish communities.

Despite the standardizing tendencies of the orthography, some dialectal phonetic processes are occasionally exhibited by the spelling of words in Classical Judaeo-Arabic texts. These relate mainly to changes in syllable structure and the shortening of long vowels. A prosthetic *ʾalep̄*, for example, indicates the elision of a short vowel in the following syllable, as in אדראהם = دراهم 'dirhams' and in the spelling of the perfect of Forms V and VI, אתפעל and אתפאעל respectively. The shortening of a long vowel in an unstressed syllable is reflected by the occasional omission of a vowel letter, as in אלדנניר = الدنانير 'dinars', אלגוואר = الجوارى 'maidservants' (Blau 1999:70ff.). Some dialectal features relating to the pronunciation of the emphatic consonants are indicated by the spelling in the manuscripts, as in אלכצארה = الخسارة 'the loss', which reflects the emphatic pronunciation of *rāʾ* and the spread of emphasis (→ *tafxīm*; Blau 1999:77). A variety of dialectal features of morphology and syntax are revealed by the texts, such as the leveling of the case distinctions of sound masculine plural and dual endings and the use of the oblique form as the common form (*-īn, -ēn*), the diminishing use of the internal passive, and the extension of the use of the particle *mā* to negate future and subordinate clauses and its replacement as an → interrogative pronoun by the constructions *ʾayy šayʾ*, *ʾēš*, *ʾāš* (Blau 1999:105ff.).

The features described in the preceding paragraph are common to a large number of dialects and not distinctive of one particular region. Occasionally, however, the texts contain features that are distinctive of the regional dialect of the writer. Some texts of North African origin, for example, contain the 1st person imperfect forms *nqtl* (1st pers. sg.)/*nqtlū* (1st pers. pl.), and texts of Iraqi origin sometimes use 3rd person plural and 2nd person plural imperfect verb forms with the ending *-ūn* in all contexts, irrespective of mood. A few texts written in Egypt attest to the demonstrative pronouns *dā* (masc. sg.), *dī* (fem. sg.), and *dōl* (pl.), which are often placed after the noun (Blau 1999:60ff.).

There appears to have been particular resistance to a few specific dialectal features. The medieval texts, for example, regularly use the literary form of the relative pronoun *allaḏī* in preference to the dialectal form *illi*.

The dialectal features described above correspond directly to features that are found in the modern spoken Arabic dialects. Although attested in medieval texts, they generally do not appear to represent forms that are at an earlier stage of diachronic development from their counterparts in the modern dialects. The Judaeo-Arabic texts attest to the existence of spoken dialects in the Middle Ages that are remarkably similar to the modern dialects. One would expect that the spoken dialects would have changed over the course of a thousand years, and this may well have been the case in a number of details. It is not easy, however, to establish with certainty the details of such historical development from the medieval texts. The main reason for this is that not all deviations from Classical Arabic should be identified as the reflection of genuine dialectal features. In some cases these deviations are pseudo-literary features, which arise due to the fact that the writer attempts to avoid a dialectal feature but produces a form that does not exist either in the writer's spoken dialect or the Classical Arabic literary language. Blau (1999:28–31) distinguishes between two types of pseudo-literary features: → hypercorrections, where the writer completely replaces a form that exists in the vernacular with a literary form, although the literary form is incorrect in this context, e.g., the sound oblique plural ending *-īna* may be replaced by *-ūna* even in an oblique context; and hypocorrections, where the writer only partially transforms a dialectal form into a literary form and produces a hybrid form that exists neither in the dialect nor in the literary language, e.g. הם באקיון *hum bāqiyūn* 'they remain', which is halfway between the dialect form *bāqīn* and the literary form *bāqūna*. Such hypocorrections may have a syntactic dimension, e.g., the dialectal relative pronoun *illi* is generally replaced by the literary form *allaḏī*, but in many texts *allaḏī* is used as an invariable form in all syntactic contexts, thus retaining the syntactic behavior of the dialect form. In general, therefore, it is difficult to distinguish between, on the one hand, a genuine dialectal form that is at an earlier stage of diachronic development than the corresponding form in the modern dialects and, on the other hand, forms that are pseudo-literary phenomena.

Many of the remarks made above regarding the dialectal elements in the medieval texts apply also to the language of Late Judaeo-Arabic texts. These texts have a much more extensive dialectal base and clearly reflect the regional dialect of the writer. The predominantly phonetic nature of the orthography, moreover, reflects many details of dialectal vocalism. The use of these texts as a source for the study of the diachronic development of the spoken regional dialects is, however, problematic. Forms and constructions differing from what is found in the corresponding modern dialect are often archaisms or pseudo-literary features. This can be illustrated by examining briefly the syntax of the demonstrative pronouns in Judaeo-Arabic texts from 17th- and 18th-century Egypt that have been preserved in the Genizah. These texts generally use the typically Egyptian forms of the demonstrative *dā, dī, dōl*, but they are regularly placed before the noun, rather than after the noun as in the modern Egyptian dialect. In the modern dialect, the demonstrative occurs before the noun in a few fossilized expressions, e.g. *dilwaʾti* 'now', *ya delxēba* 'what a pity!', which may suggest that the Early Judaeo-Arabic texts preserve an earlier stage in the development of the syntax in the dialect. In fact, the placement of the Egyptian demonstratives after the noun is attested already in medieval Judaeo-Arabic texts. Their occurrence before the noun in the 17th- and 18th-century texts is a pseudo-literary feature. Classical Arabic syntax is used with dialectal morphological forms of the pronouns. This phenomenon is found in several dialectal literary texts that have been preserved from Mamluk and Ottoman Egypt. It is likely to have entered Late Egyptian Judaeo-Arabic texts from this dialectal literature.

The orthography of Late Judaeo-Arabic, moreover, contains some vestiges of Classical Judaeo-Arabic practice. Original long /ā/ vowels that are shortened in the modern spoken dialect are frequently written with the vowel letter *ʾalep̄*, e.g. Cambridge University Library T-S 10J16.30: סנה כאמלה = سنة كاملة = Modern Egyptian Arabic *sana kamla* ‘a whole year’. The presence of the *ʾalep̄* should not, therefore, be interpreted as reflecting the preservation of the long /ā/ in the spoken dialect of the 17th and 18th centuries.

Some genres of Late Judaeo-Arabic literature have been transmitted down to modern times orally. This applies, for example, to the *šarḥ* literature of the North African communities and poetry and folk literature in Yemen. The language of such orally transmitted texts is free of the disguise of orthographic practices. In Yemen it also had a more dialectal base than the contemporary written Judaeo-Arabic. It is still, however, problematic as a source for reconstructing earlier phases of the dialects due to the presence of literary or pseudo-literary constructions or, in the case of the *šarḥ*, the imitation of the syntax of another language. Finally, it should be noted that some Late Judaeo-Arabic texts were direct copies, in Hebrew script, of Muslim dialectal literature and so cannot be used as a source for the study of the history of Jewish dialects. Such texts, for example, were produced and distributed in printed form in Tunisia during the 19th century (Chetrit 1993).

Another feature of written Judaeo-Arabic of all periods is the presence of Hebrew and Aramaic words in the language. These words occur mainly in the fields of rabbinical law and religious tradition. They are often adapted to the morphological structure of Arabic (Blau 1999:134ff.). Hebrew verbs are given Arabic verbal inflection, the derived Hebrew verbal forms being assimilated to the corresponding Arabic forms, e.g., the *hiṯpaʿel* verb התאבל ‘to mourn’ is adapted as an Arabic Form V verb תאבל. Hebrew nouns are given Arabic broken plurals, e.g. פסוק (*pāsūq*), pl. פואסיק (*pawāsīq*) ‘verse’. There is occasionally some phonological adaptation. A particularly interesting phenomenon is the conversion of Hebrew *šīn* into Arabic *sīn*, e.g. פרשה (*pārāšā*) ‘weekly scripture lesson’ > פראסה, שופר (*šōpār*) ‘horn’ > סאפור. This probably arose due to the equation of Hebrew *šīn* with Arabic *sīn* in cognate words such as Hebrew שבת (*šabbāṯ*) = Arabic سبت (*sabt*). There are a few cases of Hebrew and Aramaic influence on the syntax of Medieval Judaeo-Arabic, e.g. the use of an anticipatory object suffix preceding a direct object nominal introduced by the preposition *li-*, e.g. סמאה לישראל בני בכורי ‘He called Israel “my firstborn son”’ (Blau 1999:82).

We have been concerned so far almost exclusively with Judaeo-Arabic in its written form. The term Judaeo-Arabic, however, is also used to refer to the spoken vernacular of Jewish communities in the Arabic-speaking world. Most of the members of these communities have left their original places of residence, and many have settled in the State of Israel. There are still, however, remnants of Arabic-speaking Jewish communities in some parts of the Arab world, especially in North Africa.

The spoken Judaeo-Arabic dialects originated in the Arabic that was adopted by Jewish communities in various regions of the Middle East and North Africa after the Arab conquests. Most of these dialects now differ in some way from the dialects spoken by their Muslim neighbors and also, in certain regions, from those spoken by neighboring Christian communities. There is a linguistic justification, therefore, for designating such dialects as Judaeo-Arabic. They are far more diverse in their structure than the various literary forms of Judaeo-Arabic. Even Late Judaeo-Arabic, with its several regional variations, tended to reflect the dialect of the major Jewish community of the region and so functioned as a regional koine.

The Jewish communities, which were generally urban-based, adopted the Arabic speech of the Muslims who settled in the various towns throughout the Middle East. The differences between the Jewish and non-Jewish dialects developed due to the different historical circumstances experienced by the Jews and their non-Jewish neighbors. The Jewish dialects do not have a common origin.

The degree of difference between the Jewish dialects and those of their non-Jewish neighbors varied from region to region. The greatest differences are found in cases in which the Jewish dialect was the continuation of an old sedentary dialect, whereas the Muslim population had adopted a Bedouin type of speech. The best studied case of this was in Baghdad and Lower Iraq (Blanc 1964). Similar cleavages between a Jewish urban dialect and a Muslim Bedouin dialect existed in some North African cities such as Tripoli, Benghazi, and Oran and some smaller towns in the region of Algiers. Differences were found between the speech of Jews and non-Jews also in towns where the speech of the whole population belonged to the sedentary type. Considerable structural differences existed, for example, between the pre-Hilalian sedentary dialects of Jews and Muslims in some North African cities, such as Algiers, Fes, and Tlemcen. The Christians of Baghdad speak a sedentary type of dialect that exhibits numerous differences from the Jewish dialect. Until the middle of the 20th century, therefore, Baghdad had three distinct communal dialects, spoken by the Muslims, Jews (→ Baghdad Arabic, Jewish), and Christians respectively. In other regions, the differences between sedentary Jewish and adjacent non-Jewish dialects were of a lesser degree. As far as can be established in the present state of research, this applied, for example, to the dialects spoken by the Jewish communities of Upper Iraq, southeastern Turkey, Syria, Jerusalem, Egypt, and Yemen. In some of these dialects the divergence consisted of little more than the incorporation of Hebrew vocabulary into the speech of the Jews and suprasegmental phenomena, such as intonation patterns, as was the case in the Old City of Jerusalem. In others, there were also a few minor morphological differences. Finally, it should be noted that according to reports from the first half of the 20th century, certain members of the Jewish communities, for the sake of social prestige, made conscious attempts to avoid distinctive features of their Jewish dialects in their speech and replaced them with features of the adjacent Muslim dialect.

The conspicuous cleavages between Jewish and non-Jewish dialects have in most cases come about by different migration histories of the communities, compounded by social distance. In the case of towns where the Muslims speak a Bedouin type of dialect, the Jews are typically an old component of the population who resisted the linguistic influence of Muslim nomadic elements settling in the town. In the case of some of the towns in the region of Iraqi Kurdistan, the Muslim population ceased to speak Arabic altogether. In Arbīl, for example, only the Jewish community continued to speak Arabic down to modern times, whereas the Muslims are now Kurdish- or Turkish-speaking. In Kirkuk, the Arabic-speaking Jewish com–munity resided with predominantly Turkish-speaking Muslim neighbors.

Some differences between Jewish and non-Jewish sedentary dialects have arisen by the immigration of Jews from one town to another town with an existing Arabic-speaking Muslim population. This appears to have been the case, for example, with the Jewish communities of the Tunisian towns, which at some point in history migrated from a center in Qayrawān. Jewish immigrants in some cases joined an existing Jewish population, which may have affected the speech of the Jews of the town. Migrations of large numbers of Jews from North Africa into Egypt from the Middle Ages onward has resulted in the presence of certain Maghrebi features in the dialect of the Egyptian Jews, although some pecularities of the Jewish Cairene dialect may be conservative retentions of features that were present in Muslim speech at an earlier period. In North Africa itself, the Jewish communities received successive waves of Jewish refugees from Spain. Those fleeing from the Almohads in the 12th century would have been Arabic-speaking, although later waves of immigration in the 14th and 15th centuries from Christian Spain would have spoken Spanish, and some features of the modern Jewish dialects could have arisen due to a Spanish substrate.

An illustration of structural differences between a Jewish sedentary dialect and a Muslim Bedouin type of dialect can be given by adducing a few examples from the dialects of Baghdad. Following Blanc (1964), the Jewish dialect of Baghdad is said to belong to the *qəltu* group of dialects. These are the old sedentary dialects of the Mesopotamian region. The shibboleth for this group, *qəltu* 'I said', contains two conspicuous features that distinguish them from the Bedouin dialects, namely the unvoiced uvular pronunciation of the *qāf* and the *-tu* inflection of the 1st person singular perfect. The Bedouin dialects of Lower Iraq, including that of the Muslims of Baghdad, on the other hand, belong to the *gilit* group, in which Classical Arabic *qāf* is pronounced as /g/ and the 1st person singular inflection of the perfect verb is *-it*. Some other phonological differences between the Jewish and Muslim dialects of Baghdad are: Classical Arabic *kāf*: /k/ (Jewish) vs. /č/ (Muslim, in some circumstances); Classical Arabic *rāʾ*: /ġ/ (Jewish) vs. /r/ (Muslim); → *ʾimāla* of Classical Arabic long /ā/ (Jewish, in most cases conditioned by the presence of an adjacent /i/ or /ī/ vowel in Classical Arabic, e.g. *klīb* 'dogs') vs. lack of any *ʾimāla* in Muslim Arabic (*člāb*). There are differences in the verbal inflectional morphology of the dialects, the Jewish one being, on the whole, more conservative, e.g. the subject inflection of the strong verb in the perfect: (Jewish) 1st sg. *-tu*, 3rd fem. sg. *-ət*, 2nd pl. *-təm*, 3rd pl. *-u* vs. (Muslim) 1st sg. *-t*, 3rd fem. sg. *-at*, 2nd pl. *-tu*, 3rd pl. *-aw*; and in the 3rd masc. sg. possessive suffix: *-u* (Jewish), *-a* (Muslim). There are also lexical differences between the two dialects. The other Jewish dialects of Iraq and southeastern Turkey, all of which belong clearly to the sedentary *qəltu* group, share some of these distinctive features of the Jewish dialect of Baghdad, the closest being the Jewish dialects belonging to what Jastrow (1990a) calls the southern Kurdistan group (from Kirkuk to Khānaqīn). A common feature running through all the Jewish dialects is the /q/ phoneme and the *-tu* 1st person singular perfect suffix. The dialect of the Karaite Jews in the town of Hīt on the Euphrates, however, was not as resilient against Bedouin influence, and although a *qəltu* dialect in origin, it now exhibits numerous Bedouin features and mixed sedentary and Bedouin forms such as *qilit* 'I said'. It is noteworthy that already in the Middle Ages the Karaite Jews were particularly open to absorption into the surrounding culture, as shown, for example, by their use of Arabic script in many of their writings.

An illustration of some differences between the Jewish and non-Jewish sedentary dialects can be provided by a few examples from the Jewish and Muslim dialects of Fes in Morocco. The distinctive phonological features of Jewish Arabic are the pronunciation of Classical Arabic *qāf* as /ʾ/ vs. Muslim /q/ and the phonological merger of the sibilants /s/ with /š/ and /z/ with /ž/ (Classical Arabic *jīm*) vs. the lack of merger in the Muslim dialect. In verbal morphology the 3rd person feminine singular inflection of the perfect in the Jewish dialect has merged with that of the 1st person singular and 2nd person singular forms (*ktəbt*), whereas in the Muslim dialect, the 3rd person feminine singular form is distinct (*kətbət*). In geminate triliteral verbs, Jewish dialect lacks an augment element before suffixes beginning with a consonant (e.g. *ḥabbət* 'I loved'), whereas the augment is present in Muslim dialect (*ḥabbīt*). These distinctive features were found in various other Jewish dialects in Morocco (Heath 2002:132, 218, 222) and elsewhere in the Maghreb. There are also various lexical differences.

A case of only minor differentiation between Jewish and non-Jewish dialects was Cairo, where until the middle of the 20th century there were communities of Rabbanite Jews and Karaite Jews. The dialect spoken by the Karaite Jews was virtually identical to that of the Muslims, in conformity with the general tendency for the Karaites to be particularly open to influences from the non-Jewish environment. The Rabbanite Jewish dialect, on the other hand, exhibits a few differences, but these were not always consistent. One notable feature was the use of the forms *niqtil/niqtilu* for the 1st person singular and 1st person plural of the imperfect, which, although found in some Egyptian dialects, especially in the western delta, is not found in the standard Muslim Cairene dialect. There are also a few differences in individual verbs, e.g. Jewish *gātit* 'she came' vs. Muslim *gat*, and interrogative particles, e.g. Jewish *ʾēš* 'what' vs. Muslim *ʾē*, Jewish *kīf* 'how' vs. Muslim *ʾizzāy*.

As in written Judaeo-Arabic, the spoken dialects all contain a certain number of lexical items of Hebrew and, to a lesser extent, Aramaic origin. The degree with which they are used depends on the level of education of the speaker. They are often adapted to Arabic morphology, as in written Judaeo-Arabic, by, for example, forming broken plurals of nouns (Jewish Tlemcen *sifr* pl. *syafər* 'book') and adapting Hebrew roots to Arabic verbal morphology (Bar-Asher 1992:77ff.). In many communities, Jewish traders used a secret argot, which consisted largely of Hebrew and Aramaic lexical items with Arabic grammatical inflection.

With regard to the historical depth of the distinctive features of the Jewish spoken dialects mentioned above, many of them can be traced back to the Middle Ages. The pronunciation of *rāʾ* as /ġ/, which is distinctive of the Jewish dialect of Baghdad, is reflected in some Medieval Judaeo-Arabic texts of Iraqi origin in which the letters *reš* and *gimel* interchange (Blau 1999:252). The patterns of *ʾimāla* that are characteristic of Jewish Baghdad dialect correspond closely to the descriptions by the Arabic grammarians in the Abbasid period of the *ʾimāla* that existed in the speech of the general population of Baghdad in their time (Blanc 1964:48–49). Some medieval Judaeo-Arabic texts of North African origin exhibit an interchange of *šīn* and *sīn*, reflecting a phonological merger of sibilants (Blau 1999:251). Judaeo-Arabic texts emanating from medieval Egypt attest to the *niqtil/niqtilu* inflection of the verb and the 3rd person feminine singular form *gātit* 'she came' (Blau 1999:57, 1979:68).

FURTHER READING

An introduction to Classical Judaeo-Arabic and its linguistic background is given by Joshua Blau in his book *The emergence and linguistic background of Judaeo-Arabic*, originally published in 1966 but reissued twice in revised editions (the latest in 1999) that take into account more recent work in the field. The only systematic grammar of Medieval Judaeo-Arabic is Joshua Blau, *A grammar of Mediaeval Judaeo-Arabic* (2nd ed., 1980), which is written in Hebrew. Blau has compiled a dictionary of Medieval Judaeo-Arabic, which has still not appeared at the time of writing. Ratzaby (1985) has produced a helpful dictionary of the rare words occurring in Saadya's Bible translation. Studies of the Early Judaeo-Arabic texts have been published by Blau and Hopkins (1984, 1987). Analysis of some of these texts can also be found in Blau (2002). For the language of the Jews in pre-Islamic and early Islamic Arabic, see Newby (1971, 1988). A detailed introduction to Late Judaeo-Arabic is that of Hary (1992), which gives a general introduction to the background of post-Medieval Judaeo-Arabic and a detailed analysis of the language of a text from Egypt. See Khan (1991) and (1992) for studies of Late Egyptian Judaeo-Arabic documentary texts. Important studies on the Late Judaeo-Arabic of the North African communities have been made by Bar-Asher (1998, 2001) and Chetrit (1994). Avishur (1986) discusses some features of Late Judaeo-Arabic written in Iraq. For studies of Hebrew and Aramaic lexical elements in Late Judaeo-Arabic see Chetrit (1991), Hary (1999), and Bahat (2002).

For a general survey of the literature on spoken Judaeo-Arabic dialects with special attention to those of North Africa, see D. Cohen (1978) and Bar-Asher (1996). For a general survey of the Judaeo-Arabic dialects of Iraq, see Jastrow (1990a). The classic treatment of the phenomenon of communal dialectal divergence is Blanc (1964). A general description of the Jewish Baghdad dialect is given by Mansour (1991). Jastrow has published numerous studies of the Jewish dialects of Iraq and the adjacent region, including the dialects of ʿAqra and Arbīl (1990b), Nuṣaybīn and Qamišli (1989a), Moṣul (1989b, 1991a), and Sendor (1991b, 1993). The dialect of the Karaite Jews of Hīt is described in Khan (1997). Important studies on North African Jewish dialects are on those of Tunis (D. Cohen 1964, 1975), Fes (Brunot and Malka 1939, 1940), Moroccan dialects in general (Heath 2001), Sefrou (Stillman 1988), Algiers (M. Cohen 1912), Constantine (Tirosh-Becker 1988, 1989), and Tripoli (Yoda 2005). Some details of the Jewish dialect of Tlemcen can be found in Marçais (1902). A description of the Jewish dialect of Aleppo was made by Nevo (1991). The dialect of Egyptian Jews is studied by Blanc (1974) and more comprehensively by Rosenbaum (2006). Piamenta (2000) gives some details of the dialect of the Jews of Jerusalem in the first half of the 20th century. Some details on the Jewish dialects of Yemen can be found in Goitein (1932–1933, 1934, 1960), which are compared to the adjacent Muslim dialects by Diem (1973:33–34, 77, 111), Morag (1963), Tobi (1986), and Piamenta (1990). Studies of the Hebrew and Aramaic component in modern Jewish dialects include Goitein (1931), Bar-Asher (1992, 1998), and Avishur (1993). For the description of a trade argot used by Karaite Jews in Egypt, see Khan (1995–1997) and Rosenbaum (2002).

BIBLIOGRAPHICAL REFERENCES

Avishur, Yitzhak. 1986. "Changes in the Late Judaeo-Arabic of the Jews of Iraq". *Massorot* 2.1–17. (In Hebrew.)

——. 1993. “The Aramaic elements in the Arabic dialects of Iraqi Jews”. *Massorot* 7.1–24.

Bahat, Yaʿaqov. 2002. *The Hebrew component in the written Arabic of the Jews of Morocco*. Jerusalem: Bialiq.

Bar-Asher, Meir. 1992. *La composante hébraïque du judéo-arabe algérien*. Jerusalem: Magnes Press.

——. 1996. “La recherche sur les parlers judéo-arabes modernes du Maghreb: Etat de la question”. *Histoire Epistémologie Langage* 18.167–177.

——. 1998. *Linguistic traditions of North African Jews*. Jerusalem: Hebrew University. (In Hebrew.)

——. 2001. *Le commentaire biblique* Leshon Limmudim *de Rabbi Raphaël Berdugo*. 3 vols. Jerusalem: Hebrew University.

Blanc, Haim. 1964. *Communal dialects in Baghdad*. Cambridge, Mass.: Harvard University Press.

——. 1974. “The *nekteb-nektebu* imperfect in a variety of Cairene Arabic”. *Israel Oriental Studies* 4.206–226.

Blau, Joshua. 1979. *A grammar of Mediaeval Judaeo-Arabic*. 2nd ed. Jerusalem: Magnes Press. (In Hebrew.)

——. 1980. *Judaeo-Arabic literature: Selected texts*. Jerusalem: Magnes Press.

——. 1999. *The emergence and linguistic background of Judaeo-Arabic*. 3rd ed. Jerusalem: Ben Zwi.

——. 2002. *A handbook of Early Middle Arabic*. Jerusalem: Hebrew University of Jerusalem.

—— and Simon Hopkins. 1984. “On Early Judaeo-Arabic orthography”. *Zeitschrift für Arabische Linguistik* 12.9–27.

——. 1987. “Judaeo-Arabic papyri, collected, edited, translated and analysed”. *Jerusalem Studies in Arabic and Islam* 9.87–160.

Brunot, Louis and Elie Malka. 1939. *Textes judéo-arabes de Fès*. Rabat: Ecole du Livre.

——. 1940. *Glossaire judéo-arabe de Fès*. Rabat: Institut des Hautes Etudes Marocaines.

Chetrit, Joseph. 1991. “A socio-pragmatic study of the Hebrew component in the Judeo-Arabic poetry of North Africa: Textual aspects”. *Massorot* 5–6.251–311. (In Hebrew.)

——. 1993. “Changes in the discourse and Arabic language of the Jews of North Africa at the end of the nineteenth century”. *Peʿamim* 53.90–123.

——. 1994. *The written Judeo-Arabic poetry in North Africa*. Jerusalem: Misgav Yerušalayim.

Cohen, David. 1964. *Le parler arabe des Juifs de Tunis*. I. *Textes et documents linguistiques et ethnographiques*. The Hague and Paris: Mouton.

——. 1975. *Le parler arabe des Juifs de Tunis*. II. *Etude linguistique*. The Hague and Paris: Mouton.

——. 1978. “Judaeo-Arabic dialects”. *Encyclopaedia of Islam* VII, 299–302. 2nd ed. Leiden: E.J. Brill.

Cohen, Marcel. 1912. *Le parler arabe des Juifs d’Alger*. Paris: Champion.

Diem, Werner. 1973. *Skizzen Jemenitischer Dialekte*. Beirut and Wiesbaden: F. Steiner.

Goitein, Salomon Dov. 1931. “The Hebrew elements in the vernacular of the Yemenite Jews”. *Lĕšonénu* 3.356–380. (In Hebrew.)

——. 1932–1933. “Jemenische Geschichte”. *Zeitschrift für Assyriologie* 8.162–181, 9.19–43.

——. 1934. *Jemenica, Sprichwörter und Redensarten aus Central-Jemen mit zahlreichen Sach- und Worterläuterungen*. Leipzig: O. Harrassowitz.

——. 1960. “The language of al-Gades: The main characteristics of an Arabic dialect spoken in Lower Yemen”. *Le Muséon* 73.351–394.

Hary, Benjamin H. 1992. *Multiglossia in Judeo-Arabic with an edition, translation, and grammatical study of the Cairene Purim Scroll*. Leiden: E.J. Brill.

——. 1999. “Hebrew elements in Egyptian Judeo-Arabic texts”. *Vena Hebraica in Judaeorum Linguis*, ed. Shelomo Morag, Meir Bar-Asher, and M. Mayer-Modena, 67–91. Milan: Università degli studi di Milano, Dipartimento di scienze dell’antichità; Jerusalem: Hebrew University of Jerusalem.

Heath, Jeffrey. 2002. *Jewish and Muslim dialects of Moroccan Arabic*. London: RoutledgeCurzon.

Jastrow, Otto. 1989a. “The Judaeo-Arabic dialect of Nuṣaybīn/Qāmeşli”. *Studia linguistica et orientalia memoriae Haim Blanc dedicata*, ed. Paul Wexler, Alexander Borg, and Sasson Somekh, 156–169. Wiesbaden: O. Harrassowitz.

——. 1989b. “Notes on Jewish Maṣlāwi”. *Jerusalem Studies in Arabic and Islam* 12.282–293.

——. 1990a. “Die arabischen Dialekte der irakischen Juden”. *XXIV Deutscher Orientalistentag in Köln, 1988. Ausgewählte Vorträge*, ed. Werner Diem and Abdoldjavad Falaturi, 199–206. Stuttgart: F. Steiner.

——. 1990b. *Der arabische Dialekt der Juden von ʿAqra und Arbîl*. Wiesbaden: O. Harrassowitz.

——. 1991a. “Brotbacken: Ein Text im arabischen Dialekt der Juden von Mosul”. *Zeitschrift für Arabische Linguistik* 23.7–13.

——. 1991b. “Unheimliche Begebnisse in Sendor”. *Semitic studies in honor of Wolf Leslau on the occasion of his eighty-fifth birthday, November 14th, 1991*, ed. Alan S. Kaye, I, 773–784. Wiesbaden: O. Harrassowitz.

——. 1993. “Geschichten aus Səndōṛ”. *Zeitschrift für Arabische Linguistik* 25.161–177.

Khan, Geoffrey. 1991. “A linguistic analysis of the Judaeo-Arabic of late Genizah documents and its comparison with Classical Judaeo-Arabic”. *Sefunot* 20.223–34. (In Hebrew.)

——. 1992. “Notes on the grammar of a Late Judaeo-Arabic text”. *Jerusalem Studies in Arabic and Islam* 15.220–239.

——. 1995–1997. “A note on the trade argot of the Karaite goldsmiths of Cairo”. *Mediterranean Language Review* 9.74–76.

——. 1997. “The Arabic dialect of the Karaite Jews of Hīt”. *Zeitschrift für Arabische Linguistik* 34.53–102.

Mansour, Jacob. 1991. *The Jewish Baghdadi dialect: Studies and texts in the Judaeo-Arabic dialect of Baghdad*. Or-Yehuda: Babylonian Jewry Heritage Center.

Marçais, William. 1902. *Le dialecte arabe parlé à Tlemcen*. Paris: Faculté des Lettres d’Alger.

Morag, Shelomo. 1963. *The Hebrew language tradition of the Yemenite Jews*. Jerusalem: ha-ʾAqademyā la-Lašōn ha-ʿIvrīt. (In Hebrew.)

Nevo, M. 1991. *The Arabic dialect of the Jews of Aleppo: Phonology and morphology*. Ph.D. diss., Hebrew University of Jerusalem.

Newby, Gordon Darnell. 1971. “Observations about an Early Judaeo-Arabic”. *Jewish Quarterly Review* 61.212–221.

——. 1988. *A history of the Jews of Arabia from ancient times to their eclipse under Islam*. Columbia: University of South Carolina Press.

Piamenta, Moshe. 1990. *Dictionary of post-Classical Yemeni Arabic*. Leiden: E.J. Brill.

——. 2000. *Jewish life in Arabic language and Jerusalem Arabic in communal perspective*. Leiden: E.J. Brill.

Ratzaby, Yehudah. 1985. *A dictionary of Judaeo-Arabic in R. Saadya’s Tafsir*. Ramat-Gan: Bar-Ilan University.

Rosenbaum, Gabriel. 2002. “Hebrew words and Karaite goldsmiths’ secret language used by Jews and non-Jews in Modern Egypt”. *Peʿamim* 90.115–153. (In Hebrew.)

——. 2006. *The Arabic dialect of the Jews of Cairo*. Jerusalem.

Stillman, Norman A. 1988. *The language and culture of the Jews of Sefrou, Morocco: An ethnolinguistic study*. Manchester: University of Manchester.

Tirosh-Becker, Ofra. 1988. *The translation of the Psalms in the Judeo-Arabic of Constantine (Algeria): Phonology and morphological issues*. M.A. thesis, Hebrew University of Jerusalem. (In Hebrew.)

——. 1989. “A characterization of the Judeo-Arabic language of Constantine”. *Massorot* 3–4.285–312. (In Hebrew.)

Tobi, Y. 1986. “Quadriliteral verbs in the Judaeo-Arabic spoken in Sanaa”. *Massorot* 2.65–78. (In Hebrew.)

Yoda, Sumikazu. 2005. *The Arabic dialect of the Jews of Tripoli (Libya)*. Wiesbaden: O. Harrassowitz.

Geoffrey Khan
(University of Cambridge)

Jumla

The lexical meaning of the word *jumla* (pl. *jumal*) is ‘sum, total’; in the Western linguistic tradition it is usually translated with ‘clause, proposition’ or even ‘sentence’; in modern Arabic linguistics it is used for ‘sentence’. Originally, *jumla* was the opposite of *mufrad* lit. ‘single, individual’ which in linguistics meant ‘simple, consisting of one word or element’. Thus, *jumla* means approximately ‘consisting of more than one word or element’. In the medieval Arabic dictionaries, *jumla* is explained by *jamāʿa* ‘group, sum’, and as such, it contrasts with ‘individuals’ (*ʾafrād*) and ‘separation’ (*tafarruq*). This meaning is also found in the verb *ʾajmala*, as in the expression *ʾajmaltu lahu l-ḥisāba wa-l-kalāma* ‘I summed up for him the account and the speech’ (Ibn Manẓūr, *Lisān*, s.v. *j-m-l*).

The word *jumla* is found eight times in the first Arabic grammar (end of the 8th century C.E.), Sībawayhi’s *Kitāb*, but only as a common noun, not as a linguistic term (Troupeau 1976:61, who translates it with ‘ensemble’). *Jumla* occurs four times in al-Farrāʾ’s (d. 207/822) *Maʿānī l-Qurʾān* (cf. Kinberg 1996:133), once in the combination *jumlat al-kalām*. Although it does not seem to be used systematically, Talmon (1988:90–91) may be right when he regards it – together with its morphological variant *mujmal*, which has the same meaning – as the first real sign of the birth of a new linguistic term. It seems a little far-fetched to translate these occurrences as ‘clause’, but they certainly demonstrate that the original meaning of the word has become specialized and that it is used for a combination of words taking the place of a single word in a specific syntactic position.

The first occurrence of the word *jumla* as a real linguistic term may have been in al-Mubarrad’s (d. 285/898) *Muqtaḍab*. Al-Mubarrad, probably the first grammarian to use *jumla* consistently in a linguistic sense, defines the combination of verb and agent as a *jumla* because “it may be followed by silence” (*yaḥsunu ʿalayhi s-sukūt*). Talmon (1988), Versteegh (1995) and others believe that al-Mubarrad’s usage of *jumla* corresponds to the notion of ‘clause’, while the view of Bohas a.o. (1990:56) is that it is not clear whether al-Mubarrad intends by *jumla* something different from what Sībawayhi means by → *kalām*. The main argument raised to support the view that al-Mubarrad already uses the term in the sense of ‘clause’ or even ‘sentence’ is that he analyzes *jumla*, like *kalām*, as a unit having *fāʾida* ‘communicative value’, or simply as containing useful information for the listener.

It must be noted, however, that 20th-century descriptive grammar defines the ‘utterance’ as ‘a linguistic unit between two pauses’. This is in accordance with Carter’s (1973) view on the resemblance of medieval Arabic grammar, at least at its roots, to modern descriptive theory.

However, *jumla* does not seem to mean for al-Mubarrad more than its original everyday sense: ‘a sum; a group of words’, as opposed to ‘an individual word’. Only a combination of at least two elements, a governor (*ʿāmil*) and a governed word (*maʿmūl*) can have → *ʾiʿrāb*, and only complex utterances with *ʾiʿrāb* can serve as meaningful communications (*fāʾida*) for the listener (*muxāṭab* ‘the one who is addressed’). “It is not possible to delete the agent (*fāʿil*), because the verb (*fiʿl*) cannot stand alone without an agent” (*Muqtaḍab* I, 157 *wa-lam yajuz ḥaḏf al-fāʿil li-ʾanna l-fiʿla lā yakūnu ʾillā bi-fāʿilin*). Al-Mubarrad defines the combination of a verb and its agent as *jumla* in the chapter about the agent, saying that “the agent may only have the nominative [*ʾiʿrāb* ending *-u*], because together with the verb it forms a *jumla*, which may be followed by silence and is necessarily accompanied by a profit [for the hearer]. The agent and the verb have the same status as the topic (*mubtadaʾ*; → *ibtidāʾ*) and its comment (→ *xabar*)” (*Muqtaḍab* I, 146 *hāḏā bāb al-fāʿil wa-huwa rafʿun…wa-ʾinnamā kāna l-fāʿilu rafʿan li-ʾannahu huwa wa-l-fiʿlu jumlatun yaḥsunu ʿalayhā s-sukūt wa-tajibu bihā l-fāʾida li-l-muxāṭab fa-l-fāʿilu wa-l-fiʿlu bi-manzilati l-ibtidāʾi wa-l-xabar*). Although this may be considered the most important occurrence of the term *jumla*, which might almost serve as a definition, al-Mubarrad concentrates even here mainly on the *ʾiʿrāb* ending and its justification in terms of the theory of position and status (see Versteegh 1978).

Al-Mubarrad’s use of *jumla* becomes clearer when considered in connection with the chapter on the relative pronoun *allaḏī* serving as → *ṣila* ‘connection’ to the *mawṣūl* ‘connected’ (*Muqtaḍab* I, 156): “[In the sentence *aḍ-ḍāribu ʾaxāhu zaydun* ‘the one who beats his brother is Zayd’] ‘the one who beats his brother’ is the *mubtadaʾ* and ‘zayd’ is its *xabar*; they all [both the *mubtadaʾ* and the *xabar*] are in the connection (*ṣila*) of the connected [relative pronoun] *allaḏī*)” (*aḍ-ḍāribu ʾaxāhu mubtadaʾun wa-zaydun xabaruhu wa-humā jamīʿan fī ṣilati allaḏī*). Here, al-Mubarrad uses the word *jamīʿ* ‘all’ in the same sense he uses *jumla* elsewhere. A topic/nominal subject (*mubtadaʾ*) and its comment/predicate (*xabar*) would be analyzed as a relative clause in the European grammatical tradition, but it could not be maintained that *jamīʿ* means ‘clause’, even though it is used here as a synonym of *jumla*. In a later passage (*Muqtaḍab* I, 157), however, al-Mubarrad uses *jumla* and says: “*Allaḏī* may be considered a noun only by the strength of its second, complementary part, the connection (*ṣila*). The latter must be a self-sufficient utterance (*kalām*), consisting of two parts, the *mubtadaʾ* and its *xabar*, or a verb and its agent. These compounds (*jumal*) [or: clauses] can only be connections if there is something in them that refers back to the connected word” (*allaḏī lā yakūnu isman ʾillā bi-ṣilatin wa-lā takūnu ṣilatuhu ʾillā kalāman mustaġniyan naḥwa l-ibtidāʾu wa-l-xabaru wa-l-fiʿlu wa-l-fāʿilu…wa-lā takūnu hāḏihi l-jumalu ṣilatan lahu ʾillā wa-fīhā mā yarjiʿu ʾilayhi min ḏikrihi*).

In the 10th century, az-Zajjājī (d. 337/949) compiled a book under the title *al-Jumal*. The term *jumla* occurs twelve times in this book (*jumla* has seven; *jumal*, its plural, has five occurrences, and in addition the title of the book has the plural *jumal*). According to the editor, Ben Cheneb, the correct interpretation of the title is ‘précis’ (summary), and this meaning is adopted by Versteegh (1995:3) as well: the term in the title does not refer to ‘sentences’ but means something like ‘summary, résumé’. The title might also be interpreted quite differently, however: ‘compounds [i.e., words in a given position]’ or ‘collected [sc., linguistic expressions]’, since the plural *jumal* does not mean ‘précis’.

There are three types of environment in which *jumla* or its plural are mentioned in az-Zajjājī’s work: (i) three times as the *xabar* ‘predicate’ of the verb *kāna* ‘to be’; (ii) eight times as the contents of *ḥikāya* ‘imitation of someone’s phrase [i.e., a quotation from a not entirely authentic source]’, where it is opposed to the simple phrase (*mufrad*); and (iii) once as a *ṣila* (*Jumal* 342.5): “It is possible for *allaḏī* and its like to be connected with every *jumla*, which stands on its own, assuming that mention has been made of it within the *jumla*” (*yajūzu ʾan yūṣala allaḏī wa-ʾaxawātuhu bi-kulli jumlatin taqūmu bi-nafsihā ʾiḏā kāna fīhā ḏikrun yaʿūdu ʿalā allaḏī*).

The commentator of the *Jumal*, Ibn ʿUṣfūr (d. 670/1271), uses *jumla* in a passage in which he explains the notion of ‘quotation’ (*Šarḥ* II, 461): “A quotation is the expression of the words of the speaker as he has mentioned them in his speech, be they quoted as an individual word, or a compound” (*al-ḥikāya ʾīrād lafḏ̣ al-mutakallimi ʿalā ḥasbi mā ʾawradahu fī kalāmihi wa-lā yaxlū ʾan yakūna al-maḥkī mufradan ʾaw jumlatan*). But *jumla* has a special status as a compound of words: as a whole, it has underlyingly (*taqdīran*) an implicit declension (*ʾiʿrāb*), even though the endings of its constituent elements cannot be changed by governors (*ʿawāmil*), because they stand already in a government relation to each other. Therefore, in a quotation, the wording of the *jumla*, for instance the object of the word *qāla* 'he said', cannot be changed (Zajjājī, *Jumal* 323.2: *iʿlam ʾanna l-jumala lā tuġayyiruhā l-ʿawāmilu wa-hiya kullu kalāmin ʿamila baʿḍuhu fī baʿḍin fa-hiya tuḥkā ʿalā ʾalfāḏ̣ihā*). The *jumla* may be *muʿraba*, which here means 'well formed', or *malḥūna* 'having wrong endings' (*fa-ʾin kāna l-maḥkī jumla fa-lā yaxlū ʾan takūna l-jumla muʿrabatan ʾaw malḥūnatan*), while *kalām* may only be grammatically correct. Az-Zajjājī (*Jumal* 313.12) also writes that one may repeat a *jumla* as it was told, or repeat its sense, i.e. not verbatim, without the possible faults of *ʾiʿrāb* (*ʿalā ḥālihā* vs. *maʿnā l-kalām*). This also shows that the original meaning of *jumla* 'sum of words' is still felt. In one place (*Šarḥ* I, 95.4), Ibn ʿUṣfūr explains *jumla* simply as *majmūʿ* 'totality, sum' of the words in it (*al-jumla tadullu ʿalā majmūʿihimā*). The term *jumla* is essentially used in az-Zajjājī's *Jumal* just as in al-Mubarrad's work, in the sense of a bundle of words taking the place of a single element in the *ʾiʿrāb* (*Jumal* 313.8 *al-jumla fī mawḍiʿ naṣb*). Versteegh (1995:214) proposes 'core-sentence' as the correct translation of *jumla* in az-Zajjājī's other main work, the *ʾĪḍāḥ*.

At the end of the 10th century, al-Fārisī (d. 377/987) uses *jumla* in a wider sense in his commentary on Sībawayhi's *Kitāb*: "If the noun after it [sc. after *munḏu* 'since'] is put in the *-u* ending, as in 'I have not seen him for two years', then the speech consists of two compounds (*jumla*) 'I have not seen him' and 'for two years'" (*fa-ʾammā ʾiḏā rufiʿa l-ismu baʿdahā fī naḥwi lam ʾarahu munḏu ʿāmāni fa-l-kalām min jumlatayni lam ʾarahu jumlatun wa-munḏu ʿāmāni jumla ʾuxrā*; *Taʿlīqa* I, 23). This means that the temporal adverbial expression *munḏu ʿāmāni* is labeled here as *jumla*. In that case, however, it cannot be translated as 'clause', but it is still in accordance with the definition of *jumla* by the medieval Arab grammarians: a compound substituted for an individual word, having an implicit (*taqdīran*) *ʾiʿrāb*, in this case the nominative. This is a perfect demonstration that for al-Fārisī, and perhaps also for other grammarians, the term *jumla* may not have meant more than a 'group of words; phrase'.

In the same period, al-Fārisī's student Ibn Jinnī (d. 392/1002) seems to use the two terms interchangeably (*Xaṣāʾiṣ* I, 17): *kalām* means any semantically independent sequence of sounds: this is what grammarians call *jumal* ('compounds'; perhaps 'sentences'?). But then he adds (*Xaṣāʾiṣ* I, 19): *kalām* is independent in itself, and does not need anything else; this seems to exclude, at least implicitly, 'dependent compounds', i.e. the so-called subordinate clauses (cf. Bohas a.o. 1990:56). According to Méhiri (1973:353) the definition of 'sentence' – his translation of *jumla* – is mixed with that of *kalām* 'discourse'. Ibn Jinnī (*Xaṣāʾiṣ* I, 21) states that *jumla* is in essence every expression that exists by itself and does not need a complement (*mutammim*). Méhiri (1973:355) maintains on the basis of *Xaṣāʾiṣ* (II, 331) that *kalām* is instituted (*wuḍiʿa*) for bringing sense (*fāʾida*). Thus, it cannot result from one isolated word but only from 'sentences' (i.e. 'phrases' *jumal*). These remarks, according to Méhiri, have great importance, because they explicitly indicate that *jumla* constitutes for Ibn Jinnī the basic unit of meaningful speech (*kalām*).

In Ibn Jinnī's *Lumaʿ*, the term *jumla* occurs 16 times altogether. He also handles *jumla* 'group' as the opposite of *mufrad* 'individual' (*Lumaʿ* 72): "The comment/predicate (*xabar*) of the topic/nominal subject (*mubtadaʾ*) may be of two types: individual and group" (*xabaru l-mubtadaʾi*…*ʿalā ḍarbayni mufradun wa-jumla*). On the next page (*Lumaʿ* 73) he further defines the term, explaining *jumla* by the word *murakkab* 'composed': "As for *jumla*, it is every utterance (*kalām*) that is profitable [for the hearer] and self-sufficient. It may be of two types: a *jumla* [i.e. a group of words] consisting of a topic and its comment; and a *jumla* consisting of a verb and its agent" (*wa-ʾammā l-jumlatu fa-hiya kullu kalāmin mufīdin mustaqillin bi-nafsihi wa-hiya ʿalā ḍarbayni jumlatun murakkabatun min mubtadaʾin wa-xabarin wa-jumlatun murakkabatun min fiʿlin wa-fāʿil*). The individual predicate is the 'first', i.e. basic, constituent, and the 'group' is what takes its place (al-Xaṭīb at-Tibrīzī, *Šarḥ* 92).

In the 12th century C.E. az-Zamaxšarī (d. 538/1144) sometimes uses *jumla* as the structural counterpart of *kalām* (communicative and meaningful utterance); at other times he uses it simply in the same sense as *kalām*. The latter needs to be composed (*muʾallaf*) of two parts that are related to each other by → *ʾisnād* ‘predicativity’. If there is no *ʾisnād*, there is no *kalām*. The result of the combination is called either *jumla* or *kalām* (*ʾUnmūḏaj* 6–7). In his main grammatical work, the *Mufaṣṣal*, he becomes more explicit, stating that the *xabar* ‘predicate’ is the second part of the *jumla*, and that it may be of two kinds: *mufrad* ‘individual’ or *jumla* ‘group’ (*Mufaṣṣal* 13). Ibn Yaʿīš (d. 643/1245), the author of a large commentary on the *Mufaṣṣal*, analyzes *jumla* as the secondary (*farʿ*) counterpart of *mufrad*, taking its place. In his usage of *jumla*, he returns to an earlier period when he states that *jumla* is the comment/predicate (*xabar*) of the *mubtadaʾ* (topic/subject), as is the individual (*mufrad*; Ibn Yaʿīš, *Šarḥ* I, 88).

Characteristically, Ibn al-ʾAnbārī (d. 588/1181) does not use the term *jumla* when comparing the Basran and Kufan grammatical views in his *ʾInṣāf*. The same holds true for Ibn Hišām al-ʾAnṣārī (d. 761/1360) in his *Šarḥ Šuḏūr aḏ-ḏahab* and for introductory works like the *ʾĀjurrūmiyya* written by Ibn ʾĀjurrūm (d. 723/1323; for the latter, see Carter 1981). In a short compendium, concentrating on the *ʾiʿrāb* endings, there probably was no need or place for the term *jumla*; alternatively, the lack of a comparatively later term may suggest that these and similar works followed an ancient tradition which remained basically unaffected by later developments. However, in his large work, *Muġnī l-labīb*, Ibn Hišām does treat the notion of *jumla* in detail (cf. Owens 1988:38; Gully 1995). He opposes *jumla* to *kalām* (*Muġnī* 419); the latter is defined by him as “informative, purposeful utterance” (*al-qawl al-mufīd bi-l-qaṣd*), whereas *jumla* is “the expression of a verb and an agent or a topic and a comment” (*al-ʿibāra ʿan al-fiʿl wa-l-fāʿil wa-l-mubtadaʾ wa-l-xabar*). Ibn Hišām’s treatment is interesting for another reason as well, because he introduces the notion of a ‘small sentence’. In *zaydun ʾabūhu munṭaliqun* ‘Zayd, his father is leaving’, the component *ʾabūhu munṭaliqun*, which serves as the comment to the topic *zaydun* is called by him *al-jumla aṣ-ṣuġrā* ‘the smallest sentence’, whereas the entire sentence is called *al-jumla al-kubrā* ‘the largest sentence’ (Owens 1988:38).

Goldenberg (1988) and Talmon (1988) present an extensive analysis of the term *jumla*, in relation with the term *kalām*. The only modern author, however, who makes notice of the relevant connection of *jumla* ‘group; clause’ with *mufrad* ‘individual’ is Reckendorf (1921:1, n. 1): “Einzelwort (im Gegensatz zum Satz): *mufrad*”.

The problem of understanding and translating the term *jumla* is a problem common to almost all Arabic linguistic terms. Once it is translated one way or the other, the translation determines what it means, rather than its original sense. Rabin (1969:191) puts this as follows: “The adoption of the Indo-European terminology probably to no small extent derived from the implicit assumption that the structure of the European phrase was the ‘natural’, perhaps even only possible one.... [O]ur initial assumption should rather be one of complete difference of [surface] structure”.

Bibliographical references

Primary sources

Fārisī, *Taʿlīqa* = ʾAbū ʿAlī al-Ḥasan ibn ʾAḥmad al-Fārisī, *at-Taʿlīqa ʿalā Kitāb Sībawayhi*. 4 vols. Ed. ʿAwaḍ ben Ḥamad al-Qawzī. Cairo: Maṭābiʿ al-Ḥasanī, 1990–1996.

Farrāʾ, *Maʿānī* = ʾAbū Zakariyyāʾ Yaḥyāʾ ibn Ziyād al-Farrāʾ, *Maʿānī l-Qurʾān*. Ed. Muḥammad ʿAlī an-Najjār. 3 vols. Cairo: ad-Dār al-Miṣriyya, 1955–1972.

Ibn al-ʾAnbārī, *ʾInṣāf* = ʾAbū l-Barakāt Kamāl ad-Dīn ʿAbd ar-Raḥmān ibn Muḥammad al-ʾAnbārī, *al-ʾInṣāf fī masāʾil al-xilāf bayna n-naḥwiyyīna l-baṣriyyīna wa-l-kūfiyyīn*. Ed. Muḥammad Muḥyī d-Dīn ʿAbd al-Ḥamīd. 2 vols. Cairo: Dār al-Fikr, 1954.

Ibn Hišām, *Muġnī* = Jamāl ad-Dīn ʾAbū Muḥammad ʿAbdallāh ibn Yūsuf Ibn Hišām al-ʾAnṣārī, *Muġnī l-labīb ʿan kutub al-ʾaʿārīb*. Ed. Māzin al-Mubārak and Muḥammad ʿAlī Ḥamdallāh. Damascus: Dār al-Fikr, 1969.

——, *Šarḥ* = Jamāl ad-Dīn ʾAbū Muḥammad ʿAbdallāh Ibn Hišām al-ʾAnṣārī, *Šarḥ Šuḏūr aḏ-ḏahab fī maʿrifa kalām al-ʿArab*. Ed. ʿAbd al-Ġaniyy ad-Daqr. Damascus: aš-Šarika al-Muttaḥida li-t-Tawzīʿ, 1984.

Ibn Jinnī, *Lumaʿ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Lumaʿ fī l-ʿarabiyya*. Ed. Ḥāmid al-Muʾmin. Beirut: ʿĀlam al-Kutub, 1985.

——, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. 3 vols. Ed. Muḥammad ʿAlī an-Najjār. 3 vols. 2nd ed. Beirut: Dār al-Hudā, n.d. (1st. ed., Cairo: Dār al-Kutub al-Miṣriyya, 1952–1956.)

Ibn Manẓūr, *Lisān* = ʾAbū l-Faḍl Jamāl ad-Dīn Muḥammad ibn Mukarram Ibn Manẓūr, *Lisān al-ʿArab*. Ed. ʿAbdallāh ʿAlī al-Kabīr, Muḥammad ʾAḥmad Ḥasballāh, and Hāšim Muḥammad aš-Šāḏilī. 6 vols. Cairo: Dār al-Maʿārif, 1981.

Ibn ʿUṣfūr, *Šarḥ* = ʾAbū l-Ḥasan ʿAlī ibn Muʾmin Ibn ʿUṣfūr al-ʾIšbīlī, *Šarḥ Jumal az-Zajjājī*. 2 vols. Ed. Ṣāḥib ʾAbū Janāḥ. [Cairo], [1971].

Ibn Yaʿīš, *Šarḥ* = Muwaffaq ad-Dīn ʾAbū l-Baqāʾ Yaʿīš ibn ʿAlī Ibn Yaʿīš al-Mawṣilī, *Šarḥ al-Mufaṣṣal*. 2 vols. Cairo: Maktabat al-Mutanabbī, n.d.

Mubarrad, *Muqtaḍab* = ʾAbū l-ʿAbbās Muḥammad ibn Yazīd al-Mubarrad, *Kitāb al-muqtaḍab*. 4 vols. Ed. Muḥammad ʿAbd al-Xāliq ʿUḍayma. 2nd ed. Cairo: al-Majlis al-ʾAʿlā li-š-Šuʾūn al-ʾIslāmiyya, 1399/1979. (1st ed., 1965–1968.)

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. Ed. Hartwig Derenbourg. Paris: Imprimerie Nationale, 1881–1889.

al-Xaṭīb at-Tibrīzī, *Šarḥ* = ʾAbū Zakariyyāʾ Yaḥyā ibn ʿAlī al-Xaṭīb at-Tibrīzī, *Šarḥ al-Lumaʿ fī n-naḥw*. Ed. as-Sayyid Taqī ʿAbd as-Sayyid. al-Manṣūra: Dār Wālī al-ʾIslāmiyya, 1991.

Zajjājī, *Jumal* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān ibn ʾIsḥāq az-Zajjājī, *al-Jumal*. Ed. Mohammed Ben Cheneb. Paris: Klincksieck, 1957.

Zamaxšarī, *Mufaṣṣal* = ʾAbū l-Qāsim Maḥmūd ibn ʿUmar az-Zamaxšarī, *al-Mufaṣṣal fī n-naḥw*. Ed. Jens Peter Broch. Christiania: W.C. Fabritius, 1859.

——, *ʾUnmūḏaj* = ʾAbū l-Qāsim Maḥmūd ibn ʿUmar az-Zamaxšarī, *al-ʾUnmūḏaj fī n-naḥw*. Ed. Ḥusnī ʿAbd al-Jalīl Yūsuf. Cairo: Maktabat al-ʾĀdāb, 1990. [With the commentary of Jamāl ad-Dīn Muḥammad al-ʾArdabīlī.]

SECONDARY SOURCES

Bohas, Georges, Jean-Patrick Guillaume, and Djamel Eddine Kouloughli. 1990. *The Arabic linguistic tradition*. London: Routledge.

Carter, Michael G. 1973. “An Arab grammarian of the 8th century A.D.: A contribution to the history of linguistics.” *Journal of the American Oriental Society* 93.146–157.

——. 1981. *Arab linguistics: An introductory classical text with translation and notes*. Amsterdam: J. Benjamins.

Goldenberg, Gideon. 1988. “Subject and predicate in Arab grammatical tradition”. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 138. 39–73.

Gully, Adrian. 1995. *Grammar and semantics in medieval Arabic: A study of Ibn-Hisham’s* ‘Mughni l-labib’. London: Curzon.

Kinberg, Naphtali. 1996. *A lexicon of al-Farrāʾ’s terminology in his* Qurʾān *commentary*. Leiden: E.J. Brill.

Méhiri, Abdelkader. 1973. *Les théories grammaticales d’Ibn Jinnī*. Tunis: Université de Tunis.

Owens, Jonathan. 1988. *The foundations of grammar: An introduction to medieval Arabic grammatical theory*. Amsterdam and Philadelphia: J. Benjamins.

Rabin, Chaim. 1969. “The structure of the Semitic system of case endings”. *Proceedings of the International Conference on Semitic Studies held in Jerusalem, 19–23 July 1965*, 190–204. Jerusalem: Israel Academy of Sciences and Humanities.

Reckendorf, Hermann. 1921. *Arabische Syntax*. Heidelberg: C. Winter.

Talmon, Rafael. 1988. “*Al-kalām mā kāna muktafiyan bi-nafsihi wa-huwa l-jumla*: A study in the history of sentence-concept and the Sībawayhian legacy in Arabic grammar”. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 138.74–98.

Troupeau, Gérard. 1976. *Lexique-index du* Kitāb *de Sībawayhi*. Paris: Klincksieck.

Versteegh, Kees. 1978. “The Arabic terminology of syntactic position”. *Arabica* 25.261–281.

——. 1995. *The explanation of linguistic causes: Az-Zaǧǧāǧī’s theory of grammar*. Amsterdam and Philadelphia: J. Benjamins.

TAMÁS IVÁNYI
(Eötvös Loránd University)

Jussive → Mood

K

Kalām

The word *kalām* means ‘speech’; it is used for the pure, uncorrupted speech of the Bedouin Arabs, as a synonym of → *ʿarabiyya*. *Kalām*, however, also means any length of words uttered in a grammatically correct form; it is “a complete [series] of sounds, beneficial [for the hearer]” (*al-kalāmu llaḏī lā yakūnu ʾillā ʾaṣwātan tāmmatan mufīdatan*). Therefore, ‘the *Qurʾān* is *kalām Allāh* ‘God’s speech’, because it is complete, and self-sufficient (*muktafī bi-nafsihi*; Ibn Manẓūr, *Lisān k-l-m*; for the theological implications of speech as an attribute of God, see Peters 1976). In later linguistic compendia *kalām* is usually defined as “a compound expression, intentionally beneficial [for the hearer]” (*al-kalām huwa l-lafḏ̣ al-murakkab al-mufīd waḍʿan*; *ʾĀjurrūmiyya*, quoted by Carter 1981:8). In this definition, the speaker’s intention has become a significant factor: “If that is what you intend, it is *kalām*” (*fa-ʾin ʾaradta ḏālika kāna kalāman*; Sībawayhi, *Kitāb* I, 353.20). This is an argument in favor of the interpretation of *waḍʿan* as *qaṣdan* ‘intentionally’ in the above definition (Ramlī, *Šarḥ* 57; for an opposing view, see Carter 1981:8–9).

Kalām has to be differentiated from *qawl* ‘saying’ because the latter may or may not be part of the accepted (regular) speech, i.e. *kalām*, while *kalām* is actualized by different sayings (mainly quotations from poetry). Everything may be quoted, both (grammatically) correct sayings and unsound (*malḥūn*) ones. Therefore, a Qurʾānic quotation is introduced by *qawluhu taʿālā* ‘the Almighty said’, which is followed by a part of the *kalām Allāh*. The distinction between *kalām* and *qawl* sheds light on the nature of *kalām*: every *kalām* is *qawl*, but not every *qawl* can be regarded as *kalām*. The expression *qawl baʿḍ al-ʿArab* ‘the way of speaking of some Bedouin’, for instance, means not only that the individuals or tribes concerned offer spurious evidence but also that they are insignificant informants, who cannot even produce a line of poetry from some well-known poet containing a poetic license in support of their saying. In such a case, Sībawayhi may qualify their speech in the following way: “[Such-and-such an expression]... is not very frequent in the speech of all of them; only some of them use it in their speech” (*wa-laysat...ʾakṯara fī kalāmihim jamīʿan wa-ʾinnamā yatakallamu bihā baʿḍuhum*; *Kitāb* I, 13.3). Other qualifications include: “If you said... it would not be an [accepted] utterance” (*law qulta...lam yakun kalāman*; *Kitāb* I, 223.13); “the Bedouin say in their speech...” (*ʾinna l-ʿAraba taqūlu fī kalāmihā...*; *Kitāb* I, 173.3). Here, *kalām* may mean either the corpus or a rule derived on the basis of the corpus, while *qawl* is perhaps what the Arabs (i.e. the transmitters) would say, if asked, on the basis of their knowledge of the corpus of accepted poetry.

The term *kalām* first occurs in Sībawayhi’s *Kitāb*, appearing at least 1,144 times (the precise number varies in the different manuscripts and editions); it is the most frequently used and most important term in the *Kitāb* (cf. Iványi 2002). Mosel (1975:18) maintains that in Sībawayhi’s view *kalām* can best be defined by “what can be followed by silence” (*mā yaḥsunu s-sukūt ʿalayhi*) and “what has no need of something else” (*mā yastaġnī ʿan ġayrihi*), more or less identical with the way Zellig Harris defines ‘utterance’. Beeston (1976:650), however, regards *kalām* more as an equivalent of the notion of ‘sentence’, in the sense that it not only can be actual but also potential, i.e., the speakers could break it off if they wish. Contrary to both these views, Troupeau (1976:184) does not translate *kalām* by either ‘utterance’ or ‘sentence’ but distinguishes between four different translations: ‘langage, langue’, ‘énonciation, énoncé’, ‘mots, mot’, and ‘prose’. Carter (1972) translates *kalām* as ‘speech’. In his analysis of the evolution of the concept of ‘sentence’ in medieval Arabic grammar, Talmon demonstrates that Sībawayhi’s *Kitāb* still lacked a technical term for ‘sentence’. He concludes that *kalām* meant ‘speech unit’ for Sībawayhi (Talmon 1988:89).

Originally, the term *kalām* may have been identical with or very much similar to → *ʾiʿrāb*, or it may have meant any part of the language containing *ʾiʿrāb*: “Concerning *niʿma* and *biʾsa*, there is no *kalām* [i.e. *ʾiʿrāb* endings] in them because they cannot be changed” (*wa-ʾammā niʿma wa-biʾsa…fa-laysa fīhimā kalām li-ʾannahumā lā tuġayyarāni*; Sībawayhi, *Kitāb* II, 32.15). The most telling examples are found in Sībawayhi (*Kitāb* I, 5.11, 13), when he equates *kalām* with *ʾiʿrāb*: “the nouns that resemble the imperfect verb in *kalām* [i.e. with respect to *ʾiʿrāb*] and agree with it in → *bināʾ* [i.e. in having immutable endings]” (*mā ḍāraʿa l-fiʿla l-muḍāriʿa min al-ʾasmāʾi fī l-kalām wa-wāfaqahu fī l-bināʾ*). A somewhat later development, brought about by the need to explain the ‘changing endings’, may have identified *kalām* as ‘the result of the activity of an *ʿāmil*’: “Because it is *kalām*, some parts of it act upon other parts” (*li-ʾannahu kalām qad ʿamila baʿḍuhu fī baʿḍ*; Sībawayhi, *Kitāb* I, 176.7). Although the two, *ʾiʿrāb* and → *ʿamal*, seem to have been strongly interrelated, originally they represented two distinct phases of the Arab approach to linguistic forms – the recognition of the endings themselves, and a logical way to interpret them.

Since grammaticality depends on the correct use of *ʾiʿrāb*, *kalām* is qualified accordingly: “If you put it in *rafʿ*, the *kalām* ‘speech’ is good” (*ʾin rafaʿta fa-l-kalāmu ḥasanun*; Sībawayhi, *Kitāb* I, 400.17). There is at least one example to show that *kalām* and *ʾiʿrāb* are strongly interdependent – the weakness of the first is explained by the rarity of the second: “This is ugly, weak *kalām*…its *ʾiʿrāb* is light” (*fa-hāḏā kalāmun qabīḥun ḍaʿīfun…ʾiʿrābuhu yasīr*; Sībawayhi, *Kitāb* I, 238.4). Since we read that *kalām* may be accepted by wide circles of the informants or only by some (groups) of them, the term may have come to denote ‘any part of the [accepted] speech’, but only as far as it fits into a larger structure of so-called useful communication (*fāʾida*): “The verb needs a noun, otherwise, there is no *kalām* [accepted speech]” (*al-fiʿlu lā budda lahu min al-ismi wa-ʾillā lam yakun kalāman*; Sībawayhi, *Kitāb* I, 5.10); “*ʾana* ‘I’ is not *kalām* until you build something upon it” (*li-ʾanna ʾana lā yakūnu kalāman ḥattā yubnā ʿalayhi šayʾ*; Sībawayhi, *Kitāb* I, 395.3).

Eventually, *kalām* came to mean any length or any kind of linguistic text: “It is permitted to be between two *kalām*s [i.e. words]” (*wa-jāza ʾan yakūna bayna kalāmayni*; Sībawayhi, *Kitāb* I, 36.1); “if there is *kalām* after it” (*ʾiḏā kāna baʿdahu kalām*; Sībawayhi, *Kitāb* II, 303.17). Thus, the term *kalām* could be applied to a part of speech, a clause, or a complete sentence: “the part of speech [that serves as] topic” (*kalāmun mubtadaʾ*; Sībawayhi, *Kitāb* I, 415.2); “the part of speech/clause after *fa-*” (*al-kalāmu llaḏī baʿda l-fāʾ*; Sībawayhi, *Kitāb* I, 398.4). In this way, *kalām*, as the collection of all words with *ʾiʿrāb*, naturally came to mean ‘corpus’, i.e. the linguistic material discussed (but not necessarily collected) by the linguists. Sībawayhi says, for instance: “in the rest of the *kalām* [i.e. ‘collected speech material’]” (*fī sāʾiri l-kalām*; Sībawayhi, *Kitāb* I, 19.5); “the majority of [the utterances in] their speech is inflected [completely]” (*ʾakṯaru kalāmihim yanṣarifu*; Sībawayhi, *Kitāb* I, 5.21).

For the informants (*al-ʿArab*, i.e. the *rāwī*s), *kalām* was what their linguistic intuition and usage suggested to them. Sībawayhi identified *kalām* by his analysis of its grammatical structure and recognizing the rules governing it (cf. Dévényi 1990; see also Carter 1972). It was not his task to accept or refuse utterances. In conditional sentences with *ʾin*, for instance, although he himself always uses the perfect in the *Kitāb*, he nevertheless advocates the use of the jussive, because this is the rule based on the informants’ corpus (see Dévényi 1988:41, n. 28; cf. also Dévényi 1991; → *jazāʾ*). Yet, he does accept or refuse grammatical rules or analyses. Linguists argue with each other, not with the informants, about the grammaticality and acceptability of structures and explanations for the *ʾiʿrāb*: “It is *kalām* ‘accepted speech’ in questioning, but it is impossible in a [nominal] predicate” (*huwa kalāmun fī l-istifhāmi muḥālun fī l-xabar*; Sībawayhi, *Kitāb* I, 353.22).

The expression *kalāmuhum* may differ from *al-kalām* in some cases, the former being the authentic corpus (lines of poetry or responses of the informants to questions put to them by the linguists), while the latter is a generalization built upon these data about the linguistic competence of the informants. Linguists (or nonauthentic speakers, i.e. not *rāwī*s) may try to vary their data, generating forms not given in the corpus, but the result cannot be *kalām*, no matter how much it complies with their grammatical rules: “If you did not make the utterance agree with the second [part], you would say *ḍarabtu wa-ḍarabūnī qawmaka*, but the corpus contains only *ḍarabtu wa-ḍarabanī qawmuka*” (*wa-law lam taḥmil al-kalāma ʿalā l-ʾāxiri la-qulta ḍarabtu wa-ḍarabūnī qawmaka wa-ʾinnamā kalāmuhum ḍarabtu wa-ḍarabanī qawmuka*; Sībawayhi *Kitāb* I, 29.19). Since this corpus consists primarily of poetry, the example cited by Sībawayhi as evidence is derived from a line of poetry from al-Farazdaq (*Kitāb* I, 29.23).

The different ways in which *kalām* is qualified demonstrate that it is not always absolute. The relativity of the term *kalām* means that the scope of a rule may or may not extend to the whole grammar: “I traced it back to the analogy of *ʾafʿalu* and to the predominant usage in the speech of the Bedouin” (*fa-radadtuhu ʾilā qiyāsi ʾafʿalu wa-ʾilā l-ġālibi fī kalāmi l-ʿArab*; Sībawayhi *Kitāb* II, 112.18).

The difference between *al-kalām* and *kalāmuhum* sometimes seems to be that between ‘language’ and ‘informants’ sayings’, but at other times both terms may be used to indicate the corpus. If, however, some important informants support the rare occurrence of a linguistic phenomenon, it is qualified as partly acceptable. “We have heard the eloquent Bedouin [i.e. the transmitters] say…, but this is not in the speech of all of the Bedouin (*samiʿnā fuṣaḥāʾa l-ʿArabi yaqūlūna…wa-laysat fī kalāmi kulli l-ʿArab*; Sībawayhi, *Kitāb* I, 426.9). This is contrasted with other sources of linguistic evidence: “This is said by all of those in whose knowledge and recitation we trust” (*wa-hāḏā qawlu jamīʿi man naṯiqu bi-ʿilmihi wa-riwāyatihi*; Sībawayhi, *Kitāb* II, 4); “This is the speech of most of the Bedouin and the most eloquent ones, and it is the analogy [rule]” (*wa-huwa kalāmu ʾakṯari l-ʿArabi wa-ʾafṣaḥihim wa-huwa l-qiyās*; Sībawayhi, *Kitāb* I, 184.20).

Linguists analyze the corpus (*kalām*) with analogical methods (*qiyās*, perhaps inherited from *fiqh*; see Carter 1973, 1997), but their results have to be validated by the informants. If the results cannot find support from the transmitters from among the Bedouin (*al-ʿArab*), Sībawayhi rejects them. Linguistic forms that seem acceptable on the basis of grammatical rules deduced from the whole corpus may disagree in some cases with forms given by the informants. This lies behind the differentiation between → *ʾaṣl* and *kalām*, and is expressed by him in various ways, for instance: “The speech [corpus] contains *mittu tamūtu*, although *muttu tamūtu* conforms better to the rule (*wa-qad jāʾa fī l-kalāmi mittu tamūtu wa-muttu tamūtu ʾaqyas*; Sībawayhi, *Kitāb* II, 240.7).

There are some indications that *kalām* may be interpreted as ‘prose’ in contrast with ‘poetry’. This interpretation finds support in statements like: “In poetry is permitted what is not permitted in *kalām*” (*yajūzu fī š-šiʿri mā lā yajūzu fī l-kalām*; Sībawayhi, *Kitāb* I, 7.19). It is to be emphasized, however, that for the linguists poetry forms the basis of *kalām*: “You are allowed to elide *lā* because it is from the speech of the Bedouin, as [the poet] said” (*wa-qad yajūzu laka wa-huwa min kalāmi l-ʿArabi ʾan taḥḏifa lā…wa-qāla [š-šāʿir]*; Sībawayhi, *Kitāb* I, 404.1). Throughout the grammatical literature from the time of Sībawayhi onward, grammarians used in their examples seemingly everyday utterances with *zayd*, *marartu*, *qāʾiman* (which, however, are not classified by them as ‘prose’). Yet, these are nothing else but grammatical rules expressed in an easily memorizable and useful way. Even → poetic licenses (*ḍarāʾir aš-šiʿr*) formed part of *kalām* meaning ‘corpus’ (see Iványi 1991): “Similar to this speech is the saying of the poet in license” (*wa-miṯlu hāḏā l-kalāmi qawlu š-šāʿir ʾiḏā ḍṭurra*; Sībawayhi, *Kitāb* I, 302.4); “This is used in poetic license in the speech” (*wa-ḏālika ḥayṯu ḍṭurra fī l-kalām*; Sībawayhi, *Kitāb* II, 145.16). Even such late grammarians as Ibn

Hišām (d. 761/1360) include ḍarūra in kalām (*Taxlīṣ* 481): "It is permitted in speech to elide the feminine *t* in the perfect tense as the poet said (in the *basīṭ* meter)" (*yajūzu fī l-kalāmi ḥaḏfu tāʾi t-taʾnīṯi min al-fiʿli l-māḍī…ka-qawlihi [al-basīṭ]*). He goes on, stating explicitly that he is speaking about *ḍarūra* (*Taxlīṣ* 482): "It is permitted in poetry to elide the feminine *t* in the perfect tense as the poet said (in the *mutaqārib* meter)" (*wa-yajūzu fī š-šiʿri ḥaḏfu t-tāʾi min al-māḍī…ka-qawlihi [al-mutaqārib]*). 'Poetry' means 'license' in this relation, while *kalām* refers here to the regular linguistic usage (mainly in poetry): "This kind of speech [utterance] occurs most frequently in poetry" (*wa-miṯlu hāḏā l-kalāmi ʾakṯaru mā yakūnu fī š-šiʿr*; Sībawayhi, *Kitāb* I, 237.18). When a form was used with considerable frequency in the corpus, it was not considered license, even though it occurred exclusively in poetry and deviated from the basic rules. When contrasted with poetry, *kalām* is generally equivalent to 'main rule' (*ḥadd*; cf. Iványi 1995).

Summing up, the scope of *kalām* in Sībawayhi's *Kitāb* may be defined as follows: (i) material: corpus (speech acts or utterances) or any part of it (*fī baʿḍi l-kalām*; Sībawayhi, *Kitāb* I, 19.10); (ii) immaterial: language as defined by a set of grammatical rules or linguistic knowledge (competence) of the informants. In other words, there is a horizontal dimension consisting in the speakers' knowledge, as well as a vertical dimension consisting in the linguistic data with their rules. In this way, *kalām* means simply 'Arabic': the Arabic language, speaker of Arabic, Arabic nominal phrase, Arabic nominal subject, etc. In this connection, Arabic is defined horizontally by the 'Arabs', and vertically by Sībawayhi and his colleagues: "We have heard them [sc. the Bedouin] use it in the *kalām*" (*samiʿnāhum [al-ʿAraba] yatakallamūna bihi fī l-kalām*; Sībawayhi, *Kitāb* II, 330.10). This implies that they had another language, too, along with *kalām*!

Kalām remained the most comprehensive term to deal with the grammatical structure of the language, long after Greek logic had made its way into Arabic grammar. The term → *lisān* was usually reserved for foreign languages, whereas → *luġa*, which is sometimes used in logical writings for 'language', in linguistics usually means 'lexicon'. In his analysis of the usage of the term *kalām* in grammatical writings from al-Mubarrad's time on, Goldenberg (1988, Sec. III) concludes that in later grammatical writings *kalām*, though always self-sufficient, as a syntactic term overlaps with → *jumla*, without being its synonym. According to Versteegh (1995), the concept of *kalām* is more related to semantic considerations, whereas the *jumla* is a syntactic unit. Talmon (1988), Versteegh (1995:214), and others believe that al-Mubarrad's usage of *jumla* corresponds accordingly to the notion of 'clause', but Bohas a.o. (1990:56) believe that it is not clear whether al-Mubarrad intends by *jumla* something different from what Sībawayhi means by *kalām*.

According to Bohas a.o. (1990:56), a century after al-Mubarrad, Ibn Jinnī appears to have used the two terms interchangeably (*Xaṣāʾiṣ* I, 17): *kalām* means any semantically independent sequence of sounds, which is what grammarians call 'sentences' (*jumal*). But then Ibn Jinnī adds (*Xaṣāʾiṣ* I, 19): *kalām* are those sentences which are independent in themselves and do not need anything else. This seems to exclude, at least implicitly, 'dependent sentences', i.e. subordinate clauses. In Méhiri's view (1970:352–354), *kalām* for Ibn Jinnī is a collective noun, and just like the word *ʾinsān* 'man', it can be employed for the unit as well as for the totality. In connection with the first sentence of the *Kitāb Sībawayhi* (*hāḏā bāb ʿilm mā al-kalim min al-ʿarabiyya*), the commentator as-Sīrāfī deals at length with the difference between *kalim* and *kalām*; he states (*Šarḥ* I, 49.7) that *kalām* may be used for the whole of language or any part thereof, while *kalim* is the collective plural of → *kalima* 'word'. In his *Lumaʿ*, Ibn Jinnī uses the term *jumla* only for the clause, while calling the complete compound sentence *kalām*. Short school compendia, like the *ʾĀjurrūmiyya*, did not take over the term *jumla* at all but kept *kalām* as their central notion, the object of their study.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Ibn Hišām, *Taxlīṣ* = Jamāl ad-Dīn ʾAbū Muḥammad ʿAbdallāh Ibn Hišām al-ʾAnṣārī, *Taxlīṣ aš-šawāhid wa-talxīṣ al-fawāʾid*. Ed. ʿAbbās Muṣṭafā aṣ-Ṣāliḥī. Beirut: Dār al-Kitāb al-ʿArabī, n.d.

Ibn Jinnī, *Lumaʿ* = ʾAbū l-Fatḥ ʿUṯmān ibn ʿAbdallāh Ibn Jinnī, *al-Lumaʿ fī l-ʿarabiyya*. Ed. Ḥāmid al-Muʾmin. 2nd ed. Beirut: ʿĀlam al-Kutub, 1985.

——, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān ibn ʿAbdallāh Ibn Jinnī, *al-Xaṣāʾiṣ*. Ed. Muḥammad ʿAlī an-Najjār. Cairo, 1952–1956. (Repr., Beirut: Dār al-Hudā, n.d.)

Ibn Manẓūr, *Lisān* = ʾAbū l-Faḍl Jamāl ad-Dīn Muḥammad ibn Mukarram Ibn Manẓūr, *Lisān al-ʿArab*. Ed. ʿAbdallāh ʿAlī al-Kabīr, Muḥammad ʾAḥmad Ḥasballāh, and Hāšim Muḥammad aš-Šāḏilī. 6 vols. Cairo: Dār al-Maʿārif, 1981.

Mubarrad, *Muqtaḍab* = ʾAbū l-ʿAbbās Muḥammad ibn Yazīd al-Mubarrad, *Kitāb al-muqtaḍab*. 4 vols. Ed. Muḥammad ʿAbd al-Xāliq ʿUḍayma. Cairo: al-Majlis al-ʾAʿlā li-š-Šuʾūn al-ʾIslāmiyya, 1965–1968. (2nd ed., 1399/1979.)

Ramlī, *Šarḥ* = ʾAḥmad ibn ʿAlī ar-Ramlī, *Šarḥ al-ʾĀjurrūmiyya*. Ed. ʿAlī Mūsā aš-Šawmalī. Riyadh: Dār ʾUmayya, 1986.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. Ed. Hartwig Derenbourg. 2 vols. Paris: Imprimerie Nationale, 1881–1889.

Sīrāfī, *Šarḥ* = ʾAbū Saʿīd al-Ḥasan ibn ʿAbdallāh as-Sīrāfī, *Šarḥ Kitāb Sībawayhi*, I. Ed. Ramaḍān ʿAbd at-Tawwāb and Maḥmūd Fahmī Ḥigāzī. Cairo: al-Hayʾa al-Miṣriyya al-ʿĀmma li-l-Kitāb, 1986.

SECONDARY SOURCES

Beeston, Alfred Felix Landon. 1976. Review of Mosel (1975). *Bulletin of the School of Oriental and African Studies* 39.648–653.

Bohas, Georges, Jean-Patrick Guillaume, and Djamel Eddine Kouloughli. 1990. *The Arabic linguistic tradition*. London: Routledge.

Carter, Michael G. 1972. "Twenty dirhams in the *Kitāb* of Sībawayhi". *Bulletin of the School of Oriental and African Studies* 35.485–496.

——. 1973. "An Arab grammarian of the 8th century A.D.: A contribution to the history of linguistics". *Journal of the American Oriental Society* 93.146–157.

——. (ed.). 1981. *Arab linguistics: An introductory classical text with translation and notes*. Amsterdam and Philadelphia: J. Benjamins.

——. 1997. "Sībawayhi". *Encyclopaedia of Islam* IX, 524–531. 2nd ed. Leiden: E.J. Brill.

Dévényi, Kinga. 1988. "The treatment of conditional sentences by the mediaeval Arabic grammarians". *The Arabist* 1.11–42.

——. 1990. "Empiricism and speculation in early Arabic grammar: Struggle or mixture". *Occasional papers of the School of Abbasid Studies*, III, eds. D.E.P. Jackson a.o., 37–70.

——. 1991. "The Arab conditionals and 8th-century Arab grammarians: Between theory and practice". *al-Karmil* 12.33(e)–45(e).

Goldenberg, Gideon. 1988. "Subject and predicate in Arab grammatical tradition". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 138.39–73.

Iványi, Tamás. 1991. "*Qad yaǧūz fī š-šiʿr*: On the linguistic background of the so-called poetic licences in Sībawayhi's *Kitāb*". *The Arabist* 3–4. 201–216.

——. 1995. "The term *ḥadd* at Sībawayhi: A list and a contextual analysis". *The Arabist* 15–16.57–68.

——. 2002. "*Kalām* at Sībawayhi: Two indices with sixteen notes". *The Arabist* 24–25.17–36.

Méhiri, Abdelkader. 1970. *Les théories grammaticales d'Ibn Jinnī*. Tunis: Faculté des Lettres et des Sciences Humaines.

Mosel, Ulrike. 1975. *Die syntaktische Terminologie bei Sībawaih*. Ph.D. diss., University of Munich.

Peters, Johannes R.T.M. 1976. *God's created speech: A study in the speculative theology of the Muʿtazilī Qāḍī l-Quḍāt Abū l-Ḥasan ʿAbd al-Jabbār bn Aḥmad al-Hamaḏānī*. Leiden: E.J. Brill.

Talmon, Rafael. 1988. "*Al-kalām mā kāna muktafiyan bi-nafsihi wa-huwa l-ǧumla*: A study in the history of sentence-concept and the Sībawayhian legacy in Arabic grammar". *Zeitschrift der Deutschen Morgenländischen Gesellschaft*. 138.74–98.

Troupeau, Gérard. 1976. *Lexique-index du* Kitāb *de Sībawayhi*. Paris: Klinksieck.

Versteegh, Kees. 1995. *The explanation of linguistic causes: az-Zaǧǧāǧī's theory of grammar*. Amsterdam and Philadelphia: J. Benjamins.

TAMÁS IVÁNYI
(Eötvös Loránd University)

Kalima

1. INTRODUCTION

The form *kalima* (pl. *kalim*), commonly denoting 'a word', sometimes occurs as a grammatical term corresponding in sense to the modern linguistic term 'morpheme'. This sense of *kalima* is inferred from Sībawayhi (*Kitāb* II, 330.15–339.19), al-Mubarrad (*Muqtaḍab* I, 36–52), Ibn as-Sarrāj (*ʾUṣūl* III, 171.1–179.5), and Ibn Yaʿīš (*Šarḥ* I, 21.5–20 ed. Jahn; I, 18.29–19.15 Cairo ed.). The discussion of *kalima* by al-Mubarrad and Ibn as-Sarrāj resembles that of Sībawayhi. Ibn Yaʿīš's short discussion of this topic is mentioned by Fleischer (1888:III, 540).

2. DIVISION INTO PARTS OF SPEECH

Words (= *kalim*) are divided by the grammarians into three main parts of speech: → *ism* 'noun', → *fiʿl* 'verb', and → *ḥarf* 'particle' (Sībawayhi, *Kitāb* I, 1.1–8). These parts are divided into many subcategories (see, e.g., az-Zamaxšarī, *Mufaṣṣal* 51.3–158.5).

3. THE SOUNDS COMPOSING A *KALIMA*

The smallest phonetic units composing a *kalima* are called *ḥurūf* (sg. *ḥarf*) and *ḥarakāt* (sg. → *ḥaraka*). The term *ḥarf* as a phonetic unit must not be confused with the term *ḥarf* as a particle. As a phonetic term, the form *ḥarf* denotes a sound, but it refers only to those sounds represented in Arabic orthography by a letter.

Irrespective of that, it is attested by the text of the *Kitāb* that the distinction between a sound as a phonetic unit, and a letter of the alphabet as an orthographic unit, was clear to Sībawayhi (*Kitāb* II, 56.17–57.12).

Sībawayhi mentions in his phonetic description 35 *ḥurūf* occurring in the Old Arabic dialects. Seven of them are not accepted in the recitation of the *Qurʾān* and in poetry. The *ḥurūf* described by him include all the consonants, the semivowels *w* and *y*, and the long vowels *ū*, *ā*, *ī* (Sībawayhi, *Kitāb* II, 452.3–455.13). The short vowels *u*, *a*, *i* are called *ḥarakāt* (sg. *ḥaraka*; Sībawayhi, *Kitāb* II, 384.1–3), and they are not considered to be *ḥurūf* since they are not represented by a letter in the Arabic alphabet. Sībawayhi holds that the short vowels form a part of their long equivalents (*Kitāb* II, 342.21–24; Blanc 1967:297).

The sounds (= *ḥurūf*) included in nouns ending with a *tanwīn* or with a case marker, and in verb forms occurring without a pronoun suffix, are divided into three main categories:

i. Sounds called *ʾuṣūl* (sg. → *ʾaṣl*) are the radical sounds of the root from which a given noun or a given verb is derived (Ibn Jinnī, *Munṣif* I, 11.9–12.14).
ii. Sounds called *ḥurūf az-zawāʾid* (sg. *ḥarf zāʾid*) are sounds added to the radical sounds of the noun or the verb (Sībawayhi, *Kitāb* II, 338.20–340.21). These sounds are frequently called by the shortened name *zawāʾid* (Ibn Jinnī, *Munṣif* I, 11.9–12.14).
iii. The short vowels, called *ḥarakāt*, and the → *tanwīn* are classified as *zawāʾid* and not as *ḥurūf az-zawāʾid* (Sībawayhi, *Kitāb* II, 342.21–24, I, 79.7–8), since they are not represented in the Arabic alphabet by a letter.

This division is illustrated by the classification of the sounds contained in the active participle form *ḍāribun*: this form contains the *ʾuṣūl*, which are the radicals *ḍ*, *r*, *b*; the *ā*, which is a *ḥarf zāʾid*; and the short vowels *i, u* and the *tanwīn,* which are *zawāʾid.*

Some verbs and nouns include sounds that belong to the category of *ḥurūf al-badal* 'the sounds that are substituted for one of the radical sounds' (Sībawayhi, *Kitāb* II, 340.22–342.21; Levin 1978:182). For example, the *ā* contained in *qāla* 'he said' and in *al-māʾ* 'the water' is a *badal min al-wāw*, i.e., it is substituted for the radical *wāw*, while the *ā* in *bāʿa* 'he sold' and in *ʿāb* 'a fault' is a *badal min al-yāʾ* 'substituted for the radical *yāʾ*" (Sībawayhi, *Kitāb* II, 341.2–4).

The above classification of *ḥurūf* does not refer to those sounds contained in words belonging to the category of particle (= *ḥarf*) and to the subcategory of nouns which do not take a case ending (Ibn Jinnī, *Munṣif* I, 7.1–8.16).

All the *ḥurūf* that belong to the categories of *al-ʾuṣūl* or *al-badal* are meaningless elements. Some of the sounds that belong to the category of *ḥurūf az-zawāʾid*, like the *ā* in *ḥimār* 'a donkey', are also meaningless elements. In contrast, some of the *ḥurūf az-zawāʾid* are meaningful elements (Ibn Yaʿīš, *Šarḥ* II, 1337.7–14 ed. Jahn = IX, 143.13–144.14 Cairo ed. (see below, Sec. 4).

4. *KALIMA* AND MORPHEME

Hockett (1960:123) defines morphemes as "the smallest meaningful elements in the utterances of a language". According to this definition, an Arabic word such as *fī* 'in' is a morpheme, as it is impossible to divide it into smaller meaningful elements. Other words may contain more than one morpheme, e.g., *ḍarabtuhu* 'I hit him' contains three morphemes: *ḍarab+tu+hu.* The morpheme *ḍarab* denotes that someone performed the act of hitting. The nominative bound pronoun *-tu-* denotes the agent who performed the act expressed in the verb *ḍarab*, and the accusative bound pronoun *-hu* denotes the object suffering this act. The division of the form *ḍarabtuhu* into three morphemes, according to the principles of modern linguistics, accords with Sībawayhi's concept of *kalima.*

The term *kalima* refers in Chapter 508 of Sībawayhi's *Kitāb* (II, 330.15–339.19) to words that are free forms containing one morpheme, e.g. *huwa* 'he' (*Kitāb* II, 336.6) and *kul* 'eat!' (*Kitāb* II, 332.9); to morphemes that are bound forms, such as the suffixes of the perfect *-tu* in *ḏahabtu* 'I went away' (*Kitāb* II, 331.12–13) and the *-t* in *faʿalat* 'she did' (*Kitāb* II, 331.14); to some accusative and genitive pronominal suffixes, such as *-ka* in *raʾaytuka* 'I saw you' and *ġulāmuka* 'your servant' (*Kitāb* II, 331.11); to the *-ta* contained in *ʾanta* 'you' (*Kitāb* II, 331.12–14); and to some other suffixes (Levin 1986:426–427). The term *kalima* also refers there to some morphemes, occurring as bound forms preposed to nouns and verbs (Levin

1986:428–429), e.g. the conjunction *wa-* 'and' (Sībawayhi, *Kitāb*, II 330.15–18), the definite article *al-* in *al-qawmu* 'the people' (Sībawayhi, *Kitāb* II, 335.5), the future marker *sa-* in *sa-yafʿalu* 'he will do' (Sībawayhi, *Kitāb* II 331.9), and the preposition *bi-* 'with' as in the example *xarajtu bi-zaydin* 'I went out with Zayd' (Sībawayhi, *Kitāb* II, 331.6). This shows that when referring to free forms containing one morpheme, e.g. *huwa* and *kul*, and to morphemes that are bound forms, e.g. *-tu* and *sa-*, *kalima* corresponds to morpheme, as it refers in these cases to the smallest meaningful elements in the utterances of the language. Hence, it is inferred that in these cases the grammarians' view of *kalima* basically corresponds to the modern linguistic concept of morpheme (Levin 1986:429–431). This inference is supported by Ibn Yaʿīš's discussion of *kalima* (*Šarḥ* I, 21.5–20 ed. Jahn = I, 18.29–19.15 Cairo ed.). However, the correspondence between *kalima* and morpheme is not complete, since the grammarians do not conceive of certain linguistic units as *kalim* that in modern usage would be morphemes, because of theoretical morphological considerations. Thus, some bound forms that are morphemes according to modern linguistic concepts are not classified by the grammarians as *kalim*, although the linguists conceive of them as meaningful elements. The grammarians refer to these morphemes as *zawāʾid* or as *ḥurūf zawāʾid*, i.e. *ḥurūf* (= sounds) that are added to the radical sounds of the word (Levin 1986:436–443). Hence, they conceive of these morphemes as an integral part of the form in which they are included (Levin 1986:432–435). For example:

i. The prefixes of the imperfect verb, *ʾa-*, *ta-*, *ya-*, *na-*, as in *ʾafʿalu*, *tafʿalu*, *yafʿalu*, *nafʿalu*, are classified as *zawāʾid* because the grammarians believe that they form an integral part of the morphological pattern of the form in which they are included (Levin 1986:432). The status of these prefixes as *zawāʾid* with a grammatical task but forming an integral part of the pattern in which they are included is the same as that of the *ā* contained in the *fāʿilun* pattern denoting the active participle, and as that of the *ma-* and *-ū-* contained in the *mafʿūl* pattern denoting the passive participle (Ibn Jinnī, *Munṣif* I, 11.15–12.5; see also *Munṣif* I, 12.3–7).
ii. Sībawayhi believes that the *tanwīn* is an integral part of the noun, since it forms its final limit (Levin 1986:436–437). Hence, the *tanwīn* itself and the suffixes preceding it are classified as *zawāʾid* and not as *kalim*. This notion explains the classification of the case endings and the morphemes preceding them as *zawāʾid* and not as *kalim*. Thus, the feminine suffix *-at-* in *ṭalḥatun* 'an acacia tree', the feminine plural suffix *-āt-* in *munṭaliqātun* 'going away', and the suffix *-iyy-* denoting the relation in *tamīmiyyun* 'Tamimite, belonging to the tribe of Tamīm', are classified as *zawāʾid* (Sībawayhi, *Kitāb* II, 340.12–13; Mubarrad, *Muqtaḍab* I, 57.10, 60.3–4). The notion that the *tanwīn* is the final limit of the noun also led Sībawayhi to classify nouns like *farasun* 'a horse; a mare' (*Kitāb* I, 1.1–2), *yadun* 'a hand', and *damun* 'blood' as *kalim* (*Kitāb* II, 332.4–8), irrespective of the fact that they include the case marker and the *tanwīn* that belong to the category of the smallest meaningful elements.
iii. In Sībawayhi's view, the dual endings *-āni* and *-ayni*, as in *ar-rajulāni* and *ar-rajulayni*, and the plural suffixes *-ūna* and *-īna*, as in *al-muslimūna* and *al-muslimīna*, correspond to the combination of the case ending and the *tanwīn* in nouns like *zaydun*: the *ā/ay* included in the dual endings and the *ū/ī* in the plural suffixes correspond to a case ending, while the final *-ni* and *-na* correspond to the *tanwīn* (Sībawayhi, *Kitāb* I, 3.12–4.1; Levin 1986:432–435). It seems safe to assume that since the case endings and the *tanwīn* are classified as *zawāʾid*, the dual and plural endings corresponding to them are also classified as *zawāʾid*.
iv. For the considerations that led Sībawayhi to classify other morphemes, including the mood endings, as *zawāʾid*, see Levin (1986:435–446).

5. CONCLUSIONS

i. The form *kalima* frequently refers in the *Kitāb* to linguistic units that are composite forms containing more than one morpheme, e.g. *ʾaʿṭaytu* 'I gave' (Sībawayhi, *Kitāb* I, 342.12) and *at-taṣdīru* 'the act of putting the breast girth upon the camel' (Sībawayhi, *Kitāb* II, 476.20). When referring to such composite forms, *kalima* occurs in its usual sense of 'a word' and is not used as a technical term. In Ibn Yaʿīš's use, the term referring to such composite forms is *lafḏa* (*Šarḥ* I, 21.5–20 ed. Jahn = I, 18.29–19.15 Cairo ed.).

ii. As a technical term, *kalima* corresponds to 'morpheme' when referring to two types of forms: morphemes that are free forms, e.g. *xuḏ* 'take!', and morphemes that are bound forms, e.g. the definite article *al-* in *al-qawmu* 'the people' and the pronoun suffix *-tu* in *ḏahabtu* 'I went'.

iii. *Kalima* also occurs as a technical term referring to nominal and verbal composite forms containing meaningful *zawāʾid*, e.g. *farasun* 'a horse' and *yaḍribu* 'he beats' (Sībawayhi, *Kitāb* I, 1.1–5). In this case, *kalima* does not correspond to the modern linguistic term 'morpheme', since it includes more than one morpheme.

iv. The points of difference between *kalima* and morpheme derive from the fact that some morphological considerations interfere with the main criterion, which is a semantic one, in the determination of a *kalima*, while such morphological considerations do not interfere with the same semantic criterion for the determination of a morpheme. Hence, in Arabic grammatical theory, unlike in modern linguistics, the smallest meaningful elements of the Arabic language are divided into two groups: *kalim*, which correspond to morphemes, and *zawāʾid*, which, irrespective of their being meaningful elements, are conceived of as an integral part of a wider linguistic unit, also called *kalima*.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Ibn Jinnī, *Munṣif* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Munṣif, šarḥ li-Kitāb at-taṣrīf li-ʾAbī ʿUṯmān al-Māzinī*. Ed. ʾIbrāhīm Muṣṭafā and ʿAbdallāh ʾAmīn. 3 vols. Cairo: ʾIdāra ʾIḥyāʾ at-Turāṯ al-Qadīm, 1954–1960.

Ibn as-Sarrāj, *ʾUṣūl* = ʾAbū Bakr ibn as-Sarī Ibn as-Sarrāj, *Kitāb al-ʾuṣūl fī n-naḥw*. Ed. ʿAbd al-Ḥusayn al-Fatlī. 4th ed. 3 vols. Beirut: al-Resalah Publishers, 1999.

Ibn Yaʿīš, *Šarḥ* = Muwaffaq ad-Dīn ʾAbū l-Baqāʾ Yaʿīš Ibn Yaʿīš, *Šarḥ al-Mufaṣṣal*. Ed. Gustav Jahn, *Commentar zu Zamachšarī's* Mufaṣṣal. 2 vols. Leipzig, 1882–1886. Ed. Cairo, 10 vols. Cairo: Maktabat al-Mutanabbī, n.d.

Mubarrad, *Muqtaḍab* = ʾAbū l-ʿAbbās Muḥammad ibn Yazīd al-Mubarrad, *Kitāb al-muqtaḍab*. Ed. Muḥammad ʿAbd al-Xāliq ʿUḍayma. Cairo: Lajna ʾIḥyāʾ at-Turāṯ al-ʾIslāmī, 1385–1388 A.H.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. Ed. Hartwig Derenbourg, *Le livre de Sībawayhi, Traité de grammaire arabe*. 2 vols. Paris: Imprimerie Nationale, 1881–1889.

Zamaxšarī, *Mufaṣṣal* = ʾAbū l-Qāsim Maḥmūd ibn ʿUmar az-Zamxšarī, *al-Mufaṣṣal fī n-naḥw*. Ed. Jens Peter Broch. 2nd ed. Christiania: Libraria P.T. Mallingii, 1879.

SECONDARY SOURCES

Blanc, Haim. 1967. "The sonorous vs. muffled distinction in Old Arabic phonology". *To honor Roman Jakobson: Essays on the occasion of his seventieth birthday*, 295–308. The Hague and Paris: Mouton.

Fleischer, Heinrich Leberecht. 1885–1888. *Kleinere Schriften: Gesammelt, durchgesehen und vermehrt*. 3 vols. Leipzig: Hirzel. (Repr., Osnabrück: Biblio, 1968.)

Hockett, Charles F. 1958. *A course in modern linguistics*. New York: Macmillan Company.

Levin, Aryeh. 1978. "The *ʾimāla* of *ʾalif fāʿil* in Old Arabic". *Israel Oriental Studies* 8.174–203.

——. 1986. "The Medieval Arabic term *kalima* and the modern linguistic term morpheme: Similarities and differences". *Studies in Islamic history and civilization in honour of Professor David Ayalon*, ed. M. Sharon, 423–446. Jerusalem and Leiden: Cana and E.J. Brill.

ARYEH LEVIN (Hebrew University of Jerusalem)

Kāna wa-ʾaxawātuhā

1. INTRODUCTION

The expression *kāna wa-ʾaxawātuhā* lit. '*kāna* and its sisters' occurs as a grammatical technical term in the sense of *kāna* and the verbs that grammatically resemble *kāna*. This term refers to a category of verbs sharing with *kāna* the same grammatical qualities and occurring in the same syntactic constructions.

In their discussions of this category of verbs, the grammarians focus on two main kinds of *kāna*: *kāna at-tāmma* and *kāna an-nāqiṣa* (Levin 1979:185). The grammarians also briefly discuss two other marginal kinds of *kāna*: *kāna az-zāʾida* and *kāna allatī fīhā ḍamīr aš-šaʾn* (see Sec. 6 below).

2. *KĀNA AT-TĀMMA* AND *KĀNA AN-NĀQIṢA*

The distinction between these two kinds of *kāna* originates in the grammarian's notion that most verb forms denote two things: occurrence of an act (= *ḥadaṯ*) and time (= *zamān*). For example, the verb *ḏahaba* ‘he went away’ denotes the occurrence of the act of going away (= *aḏ-ḏahāb*) and the fact that this act took place in the past (Sībawayhi, *Kitāb* I, 11.5–15). The verb called *kāna at-tāmma* ‘the complete *kāna*’ is conceived of as a complete verb because it denotes both an act and time. In contrast, *kāna an-nāqiṣa* ‘the incomplete *kāna*’ is regarded as an incomplete verb because it does not denote an act but only time, thus lacking one of the elements expressed in most verb forms (Ibn Yaʿīš, *Šarḥ* II, 996.19–24, ed. Jahn = VII, 89.15–19, ed. Cairo); cf. Ibn as-Sarrāj, *ʾUṣūl* I, 81.17–82.14, 91.15–92.4). For a different explanation of these terms, which seems to be incorrect, see az-Zamaxšarī (*Mufaṣṣal* 119.2–6).

These two types of *kāna* occur in two different syntactic constructions, and they differ in sense:

i. *Kāna at-tāmma* is an intransitive verb that takes a subject in the nominative. It denotes the sense of *waqaʿa* ‘it happened’ or *xuliqa* ‘he was born’, as in the examples *qad kāna l-ʾamru* ‘the thing happened’ or *qad kāna ʿabdu llāhi* ‘Abdallah was born’ (Sībawayhi, *Kitāb* I, 17.1–2). This *kāna* denotes also the sense of *wujida* ‘he existed’ (Zamaxšarī, *Mufaṣṣal* 119.16).
ii. *Kāna an-nāqiṣa* occurs as a copula in the sense of ‘he was’. It takes a nominative and an accusative, as in the example *kāna ʿabdu llāhi ʾaxāka* ‘Abdallah was your brother’ (Sībawayhi, *Kitāb* I, 16.17). Both the nominative and the accusative are indispensable for the construction of the sentence. The grammarians believe that the primary and underlying construction of a sentence beginning with *kāna an-nāqiṣa* is a nominal sentence, such as *ʿabdu llāhi ʾaxūka. Kāna* is preposed to this nominal sentence in order to denote the time of the act expressed by it (Sībawayhi, *Kitāb* I, 6.14–20, 16.17–18; Ibn as-Sarrāj, *ʾUṣūl* I, 82.7–17). Thus, *kāna* becomes the *ʿāmil*, affecting the case endings of the two indispensable parts of the underlying and primary nominal sentence: it produces the nominative in its subject (*mubtadaʾ*), *ʿabdu llāhi*, and the accusative in its predicate, *ʾaxāka* (Ibn as-Sarrāj, *ʾUṣūl* I, 82.9–11). The nominative is called *ism kāna* ‘the subject in the sentence beginning with *kāna*’, while the accusative is called *xabar kāna* ‘the predicate in the sentence beginning with *kāna*’ (Ibn Yaʿīš, *Šarḥ* II, 999.2–6, ed. Jahn = VII, 91.16–21, ed. Cairo). The term *xabar kāna* refers to the predicate of the subject of the underlying and primary sentence *zaydun ʾaxūka*, i.e., it is the predicate of *ism kāna* and not the predicate of *kāna* itself. Ibn Yaʿīš argues that *xabar kāna* cannot be the predicate of *kāna* because *kāna* is a verb, and a verb cannot take a predicate (Levin 1979:203–205). He states explicitly that *xabar kāna* is an inaccurate term, used by the grammarians for the sake of convenience when teaching Arabic grammar to beginners.

3. *ʾAXAWĀT KĀNA*

The notion that a given verb form occurs in a certain syntactic construction as a *fiʿl tāmm*, denoting both an act and time, while in another syntactic construction it occurs as a *fiʿl nāqiṣ*, denoting time only, is implied by the grammarians with respect to some other verbs, known as *ʾaxawāt kāna* ‘verbs which grammatically resemble *kāna*’. These verbs are divided according to semantic criteria into three main groups:

Group 1: Verbs that, as *ʾafʿāl tāmma*, denote the sense of ‘he did something at a certain part of the day’ or ‘he entered upon a certain part of the day’, such as *ḏ̣alla* ‘he did during the whole day’, *bāta* ‘he did during the whole night’, *ʾaṣfara* ‘he did at daybreak’, *ʾaṣbaḥa*, *ġāda* ‘he did in the morning’, *ʾaḍḥā* ‘he did in the forenoon, *ʾamsā*, *rāḥa* ‘he did in the evening’.When occurring as *ʾafʿāl nāqiṣa*, these verbs denote the sense of ‘he became’ and do not express an act, since their literal sense is ‘beginning with a certain point of time he was’, e.g. *ʾaṣbaḥa zaydun ġaniyyan* ‘Zayd became rich’ (Zamaxšarī, *Mufaṣṣal* 120.10) lit. ‘beginning with a certain point of time Zayd was rich’, *ḏ̣alla wajhuhu muswaddan* ‘his face became black’ (Q. 16/58; Wright 1951:II, 102–103). Some grammarians say that when some of these verbs occur as *ʾafʿāl nāqiṣa*, they can also express the sense of ‘he came in a certain part of the day [in any particular state or condition]’. Thus, *ʾaṣbaḥa* means ‘he came in the morning [in any particular state or condition]’ (Lane 1863–1893:IV, 1641B), *ʾamsā* ‘he came in the evening…’ and *ʾaḍḥā* ‘he came in the forenoon…’ (Zamaxšarī, *Mufaṣṣal* 120.4–6, 120.13–15; Ibn Yaʿīš, *Šarḥ* II, 1007.24–1008.5, 1009.6–10, ed. Jahn = VII, 103.18–23, 105.8–106.2, ed. Cairo). For example, *ʾaṣbaḥa zaydun munṭaliqan* ‘Zayd came in the morning in a state of walking’ (Ibn as-Sarrāj, *ʾUṣūl* I, 82.14–15).

Group 2: Verbs that, as *ʾafʿāl tāmma*, denote an act of motion or rest. According to Wright (1951:II, 102–103), these verbs are *ṣāra* ‘he went’; *rajaʿa*, *ʿāda*, *ʾāḍa* ‘he came back’; *jāʾa*, *ʾatā* ‘he came’; *qaʿada* ‘he sat down’. As *ʾafʿāl nāqiṣa* these verbs denote the sense of ‘he became’, e.g. *ṣāra l-faqīru ġaniyyan* ‘the poor man became rich’ (Zamaxšarī, *Mufaṣṣal* 120.2–3) and *al-ʿaduwwu mā yarjiʿu ṣadīqan* ‘an enemy never becomes a friend’ and *iḏhabū bi-qamīṣī hāḏā fa-ʾalqūhu ʿalā wajhi ʾabī yaʾti baṣīran* ‘go with this shirt of mine and throw it on my father’s face, and as a result he shall become seeing’ (Q. 12/93; Wright 1951:II, 102–103). All grammarians agree that *ṣāra* is one of the *ʾaxawāt kāna*, occurring both as a *fiʿl tāmm* and a *fiʿl nāqiṣ*, but not all of them agree that the rest of the above verbs denoting an act of motion also belong to this category. Some grammarians do not mention these verbs at all in their discussions of *kāna wa-ʾaxawātuhā*, while others say that only some of them belong to this category (Zamaxšarī, *Mufaṣṣal* 119.6–10; ʾAbū Ḥayyān, *Manhaj* 53.10–21). Sībawayhi (*Kitāb* I, 18.22–19.5) says that the verb *jāʾa* corresponds to *kāna* known later as *kāna an-nāqiṣa* only in the expression *mā jāʾat ḥājataka* ‘What has become your need?’.

Group 3: Verbs that, as *ʾafʿāl tāmma*, denote the sense of ‘to continue’ or ‘to cease’. These verbs are *dāma* ‘he continued’ and *zāla*, *bariḥa*, *fatiʾa*, and *infakka* ‘he ceased’. These verbs occur as *ʾafʿāl nāqiṣa* only when they are used in expressions designating the sense of an adverb denoting duration:

i. *mā dāma* ‘as long as’. In this expression *dāma* is inflected only in the past tense, and it is always preceded by *mā* expressing duration, as in the example *lā ʾuġannī li-nafsī šayʾan mā dumtu ḥayyan* ‘I will never sing a melody that I myself composed as long as I live’ (ʾIṣbahānī, *ʾAġānī* IV, 170.3).
ii. *mā zāla*, *mā bariḥa*, *mā fatiʾa*, and *mā infakka*, denoting the sense of the adverbs ‘still, yet’. The inflected forms of these verbs occur in all tenses and are usually preceded by a negative particle, as in the example *mā ziltu qāʾiman* ‘I am still standing’. In some examples in the *Qurʾān* and in old poetry, the negative particle is dropped, but its sense is implied in the utterance, as in *tallāhi taftaʾu taḏkuru yūsufa ḥattā takūna ḥaraḍan* ‘By God! You will never cease to mention Yūsuf until you will be an old man’ (Q. 12/85; Wright 1951:II, 101–103).

To Group 3 one can add the verb *baqiya*, which as a *fiʿl tāmm* denotes the sense of ‘he remained’ and as a *fiʿl nāqiṣ* denotes duration (Wright 1951:II, 101–103), as in the example *lam yabqa ṭiflan* ‘he is no longer a child’ (Wehr 1961:69A).

In addition to Groups 1–3, the verb *laysa* forms a category by itself: this verb denoting the negation of the nexus between the subject and the predicate occurs only as a *fiʿl nāqiṣ* (Jurjānī, *Muqtaṣid* I, 400.19–401.5).

4. THE GRAMMARIANS’ VIEW OF *KĀNA WA-ʾAXAWĀTUHĀ* AS *ʾAFʿĀL NĀQIṢA*

Sībawayhi’s view of *kāna*, later called *kāna an-nāqiṣa*, is based on five notions:

i. Sībawayhi emphasizes that grammatically, *kāna* is a transitive verb like *ḍaraba*. This *kāna* takes an agent called *ism al-fāʿil* or *al-fāʿil*, and a direct object called *ism al-mafʿūl* or *al-mafʿūl*, and it is the *ʿāmil* producing the nominative in the agent and the accusative in the direct object (Levin 1979:186–190).
ii. The direct object is indispensable for the syntactic structure of the sentence where this kind of *kāna* occurs (Levin 1979:186–187, 189–192).
iii. The relation between the agent and the direct object corresponds to the relation between a subject (*mubtadaʾ*) and a predicate in a nominal sentence (Levin 1979:196–202).
iv. The underlying and primary construction of a sentence beginning with *kāna* is a nominal sentence (Sībawayhi, *Kitāb* I, 6.14–20; see Sec. 2 above).

v. The aim of preposing *kāna* that does not express an act to a nominal sentence is to denote the time of the act expressed in it (Sībawayhi, *Kitāb* I, 16.17–18; see Sec. 2 above).

Al-Mubarrad (*Muqtaḍab* III, 97.1–6, 189.12–15; IV, 86.1–15) and Ibn as-Sarrāj (*ʾUṣūl* I, 81.1–82.17) accepted these notions. The later grammarians, from the 10th century onward, also accepted them. However, the contents and the structure of the grammarians' discussions of *kāna an-nāqiṣa* show that the later grammarians differed from Sībawayhi, al-Mubarrad, and Ibn as-Sarrāj in their evaluation of the importance of the notion that *kāna* is a transitive verb like *ḍaraba*. While this notion is of great importance to Sībawayhi, who deals with it in detail (*Kitāb* I, 16.13–24), the later grammarians briefly refer to it in their discussions, saying that *ism kāna* and *xabar kāna* are likened to the *fāʿil* and the *mafʿūl* respectively (Ibn Jinnī, *Lumaʿ* 15.13–14; Ibn al-ʾAnbārī, *ʾAsrār* 57.10–12; Zamaxšarī, *Mufaṣṣal* 33.19–20). It seems that their wording in this respect derives from the text of Ibn as-Sarrāj (*ʾUṣūl* I, 82.7–11). The later grammarians focused in their discussions of *kāna an-nāqiṣa* on Sībawayhi's other notions, although they expressed them in a different way (Ibn Jinnī, *Lumaʿ* 15.9–16.7; Ibn al-ʾAnbārī, *ʾAsrār* 55.3–57.20; Zamaxšarī, *Mufaṣṣal* 119.1–16).

Some of the later grammarians, including al-Mubarrad and Ibn as-Sarrāj, hold that *kāna wa-ʾaxawātuhā* as *ʾafʿāl nāqiṣa* are improper verbs (*ʾafʿāl ġayr ḥaqīqiyya*) because they do not denote an act (Levin 1979:206; cf. Ibn as-Sarrāj, *ʾUṣūl* I, 82.2–9). Others call them *ʾafʿāl al-ʿibara* '[words classified as] verbs [because of the way they] are pronounced', i.e. words classified as verbs because of their form and not because of their grammatical qualities (Levin 1979:206, n. 126). Ibn as-Sarrāj (*ʾUṣūl* I, 74.4–7) calls them *ʾafʿāl fī l-lafḏ*. This expression denotes the same as *ʾafʿāl al-ʿibara*.

5. THE TERMINOLOGY REFERRING TO *KĀNA WA-ʾAXAWĀTUHĀ*

The terminology referring to *kāna wa-ʾaxawātuhā* originates in the grammatical notions discussed in Sections 2 and 4 above.

i. In Sībawayhi's terminology for the nominative and the accusative in sentences beginning with *kāna*, the nominative is called → *ism al-fāʿil* 'the noun that is the *fāʿil* [= agent]' , and the accusative is called *ism al-mafʿūl* 'the noun that is the *mafʿūl* [= direct object]' (*Kitāb* I, 16.12–13, 16.20). The nominative and the accusative are also called *al-fāʿil* and *al-mafʿūl*, respectively (*Kitāb* I, 16.12–13, 16.20, 17.1, 18.4, 63.4, 63.11–12). The accusative is called once by Sībawayhi *xabar kāna* (*Kitāb* I, 63.8) because in his terminology the term *xabar* refers to the predicate of a nominal sentence, occurring as the underlying and primary sentence of some syntactic structures (Levin 1979:193–196).

ii. Some of Sībawayhi's terminology appears in al-ʾAxfaš (*Maʿānī* I, 197.3–4; II, 670.2–3). The whole of Sībawayhi's terminology, including the term *xabar kāna*, appears in al-Mubarrad (*Muqtaḍab* III, 97.1–3, 117.3, 189.12; IV, 86.1–3) and in Ibn as-Sarrāj (*ʾUṣūl* I, 91.7; II, 277.4–5, 288.12–14, 290.3–4). Apart from Sībawayhi's terminology, al-Mubarrad (*Muqtaḍab* IV, 88.3, 88.14, 98.2) and Ibn as-Sarrāj (*ʾUṣūl* I, 82.1, 82.2, 83.11) also use the terms *ism kāna*, *kāna wa-ʾaxawātuhā*, and *ʾaxawāt kāna*. These terms, as well as the term *xabar kāna*, became the usual ones for the later grammarians from the 10th century onward. These later grammarians also use the terms *al-ʾafʿāl an-nāqiṣa* (Zamaxšarī, *Mufaṣṣal* 119.1–2), *kāna at-tāmma*, and *kāna an-nāqiṣa*. The terms *nāqiṣa* and *tāmma* are used to refer to *kāna* for the first time by az-Zajjājī (*Jumal* 48.11–49.5). Ibn as-Sarrāj (*ʾUṣūl* I, 91.15–92.4) was the first to use the term *ʾafʿāl tāmma*.

6. THE MARGINAL TYPES OF *KĀNA*

Additionally, there are two marginal types of *kāna*:

i. The type of *kāna* called *kāna az-zāʾida* only occurs as a word denoting emphasis. This *kāna* has no literal meaning, and it does not affect the case ending of any word in the sentence, as in the example *lam yūjad kāna miṯluhum* 'nobody like them existed [among the people of their tribe]' (Mubarrad, *Muqtaḍab* IV, 116.6–118.3). For references to this *kāna* see Levin (1979:185, n. 1); see also Ibn as-Sarrāj (*ʾUṣūl* I, 92.7–8).

ii. The grammarians say that *kāna allatī fīhā ḍamīr aš-šaʾn* occurs in examples such as *kāna zaydun qāʾimun*, where *kāna* is followed by two nominatives. They assert that this *kāna* contains an implicit pronoun of the 3rd person singular, referring to the noun *aš-šaʾn* or *al-ʾamr* 'the matter, the affair' or *al-ḥadīṯ* 'the story', which does not occur in the sentence. Hence, when saying *kāna zaydun qāʾimun*, what the speaker actually intends to say is *kāna š-šaʾnu wa-l-ḥadīṯu zaydun qāʾimun* 'the affair and the story was: Zayd is standing' (Ibn Jinnī, *Lumaʿ* 16.13–17). The grammarians believe that this implicit pronoun referring to *aš-šaʾn* is a subject (*mubtadaʾ*) and that the following clause *zaydun qāʾimun* is assigned to it as its predicate (Ibn Yaʿīš, *Šarḥ* I, 435.13–19, 1005.8–19, ed. Jahn = III, 116.5–6; VII, 101.14–102.3, ed. Cairo). The implicit pronoun is analyzed as *ism kāna* (Zajjājī, *Jumal* 49.9–10). The examples mentioned by the grammarians show that this *kāna* occurred mainly in poetry.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

ʾAbū Ḥayyān, *Manhaj* = ʾAbū Ḥayyān Muḥammad ibn Yūsuf al-Ġarnāṭī an-Naḥwī, *Manhaj as-sālik fī l-kalām ʿalā ʾAlfiyyat Ibn Mālik*. Ed. Sidney Glazer. Ann Arbor, Mich.: Edwards Brothers, 1947.

ʾAxfaš, *Maʿānī* = ʾAbū l-Ḥasan Saʿīd ibn Masʿada al-ʾAxfaš al-ʾAwsaṭ, *Maʿānī l-Qurʾān*. Ed. ʿAbd al-ʾAmīr Muḥammad ʾAmīn al-Ward. 2 vols. Beirut: ʿĀlam al-Kutub, 1986.

Ibn al-ʾAnbārī, *ʾAsrār* = ʾAbū l-Barakāt ʿAbd ar-Raḥmān ibn Muḥammad Ibn al-ʾAnbārī, *ʾAsrār al-ʿarabiyya*. Ed. Christian Friedrich Seybold. Leiden: E.J. Brill, 1886.

Ibn Jinnī, *Lumaʿ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Lumaʿ fī n-naḥw*. Ed. Hadi M. Kechrida. Stockholm: Almqvist and Wiksell, 1976.

Ibn as-Sarrāj, *ʾUṣūl* = ʾAbū Bakr ibn as-Sarī Ibn as-Sarrāj, *Kitāb al-ʾuṣūl fī n-naḥw*. Ed. ʿAbd al-Ḥusayn al-Fatlī. 4th ed. 3 vols. Beirut: al-Resalah, 1999.

Ibn Xaššāb, *Murtajil* = ʾAbū Muḥammad ʿAbdallāh ibn ʾAḥmad Ibn Xaššāb, *al-Murtajil*. Ed. ʿAlī Ḥaydar. Damascus, 1972.

Ibn Yaʿīš, *Šarḥ* = Muwaffaq ad-Dīn ʾAbū l-Baqāʾ Yaʿīš Ibn Yaʿīš, *Šarḥ al-Mufaṣṣal*. Ed. Gustav Jahn. 2 vols. Leipzig, 1882–1886. Ed. Cairo, 10 vols. Cairo: Maktabat al-Mutanabbī, Beirut: ʿĀlam al-Kutub, n.d.

ʾIṣbahānī, *ʾAġānī* = ʾAbū l-Faraj al-ʾIṣbahānī, *al-ʾAġānī*. 20 vols in 10. Beirut: Ṣalāḥ Yūsuf al-Xalīl and Dār al-Fikr li-l-Jamīʿ, 1970. (Repr. of Būlāq ed.)

Jurjānī, *Muqtaṣid* = ʾAbū Bakr ʿAbd al-Qāhir ibn ʿAbd ar-Raḥmān al-Jurjānī, *Kitāb al-muqtaṣid fī šarḥ al-ʾĪḍāḥ*. Ed. Kāẓim Baḥr al-Marjān. [Baghdad]: Dār ar-Rašīd li-n-Našr, 1982.

Mubarrad, *Muqtaḍab* = ʾAbū l-ʿAbbās Muḥammad ibn Yazīd al-Mubarrad, *Kitāb al-muqtaḍab*. Ed. ʿAbd al-Xāliq ʿUḍayma. Cairo: Lajna ʾIḥyāʾ at-Turāṯ al-ʾIslāmī, 1385–1388 A.H.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. Ed. Hartwig Derenbourg. Paris: Imprimerie Nationale, 1881–1889.

Zajjājī, *Jumal* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān ibn ʾIsḥāq az-Zajjājī, *Kitāb al-jumal fī n-naḥw*. Ed. ʿAlī Tawfīq al-Ḥamad. Beirut and Irbid: Muʾassasat ar-Risāla and Dār al-ʾAmal, 1984.

Zamaxšarī, *Mufaṣṣal* = ʾAbū l-Qāsim Maḥmūd ibn ʿUmar az-Zamaxšarī, *al-Mufaṣṣal fī n-naḥw*. Ed. Jens Peter Broch. 2nd ed. Christiania: Libraria P.T. Mallingii, 1879.

SECONDARY SOURCES

Lane, Edward William. 1863–1893. *Arabic-English lexicon*. 8 vols. London and Edinburgh: Williams and Norgate.

Levin, Aryeh. 1979. "Sībawayhi's view of the syntactical structure of *kāna waʾaxawātuhā*". *Jerusalem Studies in Arabic and Islam* 1.185–211.

Wehr, Hans. 1961. *A dictionary of Modern Written Arabic*. Ed. J. Milton Cowan. Wiesbaden: O. Harrassowitz.

Wright, William. 1951. *A grammar of the Arabic language*. 3rd ed. 2 vols. Cambridge: Cambridge University Press.

ARYEH LEVIN
(Hebrew University of Jerusalem)

Kanuri

1. KANURI AND ARABIC

The first contact between Islam and the empire of Kanem, situated near Lake Chad, was made through trade. Kanem had commercial links with Tripoli in North Africa via Kawar and the Fezzan. This trade "provided the gateway for Islam to enter Kanem" (Clarke 1982:67). In the second half of the 8th century, a more permanent Muslim presence was established on the Kanem-North African trade route with the establishment of the small states of Ajar Fazzan and Zawila; Zawila, further south and close to Kanem, was a center for Ibadite Islam. Kanem became Muslim at the beginning of the 12th century. According to some scholars, the first Muslim ruler in Kanem was Humai Jilme of the Saifawa dynasty (from ca. 1085 to 1097 C.E.).

Kanuri is the Nilo-Saharan language with the largest number of Arabic loans, second only to Hausa (see Baldi 2002). Its lexicon "might be described as a *layered* lexicon. It is layered chronologically in the sense that it is possible to recognize words in the language that are older as distinct from words that have not been in the language as long and also others that have entered the language from other language sources. Borrowed words in the lexicon as well are layered chronologically, since certain Arabic loan words have been part of the language for as much as a thousand years, whereas words from Hausa entered much later, and words from English only very recently" (Hutchison 1981:10).

It has been observed that one important subpart of the Kanuri lexicon is made up of historically derived words that are now recognizable as composed of word formatives that are no longer productive in the language (see Baldi 1998).

Words that have entered the Kanuri language from sources other than Arabic are insignificant in comparison to those from Arabic. Many have been greatly altered by the phonology of Kanuri.

2. PHONOLOGY

Kanuri seems to be a language with a tendency to erosion of the phonetic shape of the borrowed words. This phenomenon does not seem to correlate with the position of the syllable, consonant, or vowel eroded in the word. In fact, one glance at the loanwords shows that this erosion may take place in any part of the word: *xarāj* > *arájí* '[land] tax', *ṣadaqa* > *sádáa* 'alms', *ṯaman* > *táma* 'price'. Sometimes the two variants (i.e. with and without the phoneme assimilated) coexist: *ṣulḥ*> *sə́lfu/solô* 'reconciliation', *xarāj* > *harájí/arájí* '[land] tax', *xalq* > *álə́ga/áláa* 'creation', *ʿaql* > *hángal/ángal* 'intelligence'.

In Kanuri there is a tendency to preserve within loanwords the original Arabic geminates: *dābba* > *dábba* 'animal', *ḥujja* 'argument' > *hújja* 'reason', *ʾillā* > *íllá* 'except', *ʾammā* > *ammá* 'but', *janna* > *zánna* 'paradise'. There are only a few exceptions, probably due to colloquial use: *muʿallim* > *máləm* 'teacher', *muʿallima* > *málamá* 'female teacher'.

Individual Arabic phonemes are realized in Kanuri as follows (for more details see Baldi 2002):

/ʾ/ > Ø: *raʾy* > *râi* 'opinion'

/b/ > /b/: *rajab* > *Rázab/Rájab* 'seventh Islamic month'; in intervocalic and medial position /b/ > /w/: *xuṭba* > *kútúwa* 'Friday sermon'; in final position /b/ > /b/ or sometimes /p/: *tartīb* > *tartîp* 'order'

/t/ > /t/: *turjumān* > *turjimán* 'interpreter'; in two cases, /t/ > /d/: *ṯābit* 'firm, fixed; permanent' > *táwadə* 'definite, certain, sure', *sabt* > *sə̀bdə* 'Saturday'

/ṯ/ > /t/: *ṯaman* > *táma* 'price'; in two cases /ṯ/ > /s/: *ṯumn* > *súmmun* 'one-eighth', *ḥadīṯ* > *hadîs/hadîsə* 'Hadith, traditions about the Prophet'

/j/ > /j/: *jarīda* (via Hausa *jàrīdā̀*) > *jerída* 'newspaper'; sometimes in intervocalic or initial position /j/ > /z/: *ʿajal* > *azála* 'hurry, haste', *jumʿa* > *zə́ma* 'Friday'; in one case /j/ > /g/: *farj* > *fárgi* 'vagina'; in another case /j/ > /d/: *sarj* > *sə́rdə́* 'saddle [for horse]'

/ḥ/ > /h/: *baḥr* > *báhar* 'sea'; often, in initial, intervocalic, or final position /ḥ/ > Ø: *ḥarf* > *árawu* 'letter of the alphabet', *muḥarram* > *Márrəm* 'first month', *ḏabḥ* > *dúwa* 'slaughter'; in two cases /ḥ/ > /k/: /*ṣaḥn* 'dish, plate' > *sakân* 'kettle', *faṣīḥ* > *fásek* 'eloquent'; in another case /ḥ/ > /f/: /*ṣulḥ* > *sə́lfu* 'reconciliation'

/x/ > /h/: *xabar* > *hawâr* 'news'; often /x/ > Ø: *xādim* 'domestic servant' > *ádim* 'eunuch'; in two cases /x/ > /k/: *xuṭba* > *kútúwa* 'Friday sermon', *xums* > *kúmsa* 'sign marking end of every fifth verse of the *Qurʾān*

/d/ > /d/: *dars* > *dáras* 'lesson'; in one case /d/ > /t/: *labbād* > *labbât* 'felt'

/ḏ/ > /d/: *ḏabḥ* > *dúwa* 'slaughter'; /ḏ/ > /z/ only in a very few cases: *ḏakar* > *zákar* 'penis'

/r/ > /r/: *raṭl* > *rátal* 'a weight'; once with dissimilation: /r/ > /l/: *ḥarīr* > *harîl* 'silk'

/z/ > /z/: *zamān* > *zamân* 'time'; but often /z/ > /j/: *zabād* 'civet' > *jíbda* 'civet cat'; in one case /z/ > /s/: *ḥizb* > *ísəp* 'one-sixtieth part of the *Qurʾān*'

/s/ > /s/: *bāsūr* 'hemorrhoids' (? via Fulfulde *baksur*) > *boksûr* 'tumor; piles'; but sometimes /s/ > /š/ before front vowel: *masjid* > *máshídi* 'mosque', *as-sirr* > *ashîr* 'secret'

/š/ > /š/: *šahāda* > *sháda* 'creed formula'; but sometimes /š/ > /s/, under the influence of Fulfulde, where it is regular: *šukr* > *askə́ra* 'thanks', *muškil* 'equivocal, problematic' > *mískil* 'misunderstanding, altercation', *qurūš* [pl.] > *gúrsu* 'Maria Theresa dollar'

/ṣ/ > /s/: /*ṣandūq* > *sandúwu* 'box'; but sometimes /ṣ/ > /š/: *ṣābūn* > *shawûl* 'soap', *ṣafar*

> *Sháwur*/*Sáwur* ‘second month’, *muṣība* > *mashíwa* ‘misfortune’, *aṣ-ṣawm* > *ashâm* ‘the fast [of Ramadan]’, *maqāṣṣ* > *mówosə* ‘scissors’

/ḍ/ > /l/: *al-qāḍī* > *alkáli* ‘judge’; sometimes /ḍ/ > /d/: *al-arḍ* > *lárdə* ‘earth’, *ḥaḍara* ‘to be present’ > *ádári* ‘witnessing’, *ramaḍān* > *Ramadán* ‘ninth month’, *maraḍ* ‘disease, malady’ > *mərdí* ‘wrinkles or folds of the flesh’

/ṭ/ > /t/: *ṭāsa* (via Hausa *tāsā̀*) > *tása* ‘metal bowl’; in one case /ṭ/ > /l/, with assimilation to the preceding phoneme: *lūṭ* ‘Lot’ > *lúlu* ‘sodomy’

/ḏ̣/ > /z/: *naḏ̣m* > *názəmu* ‘verse’; in one case /ḏ̣/ > /s/: *ʾaḏ̣hār* (pl. of *ḏ̣uhr*) ‘midday prayer’ > *ásar* ‘late afternoon’

/ʿ/ > Ø: *naʿnaʿ* > *naná* ‘mint’; -C- + /ʿ/ > -CC-: *rubʿ* > *rə́bbu*/*rúbbu* ‘one-quarter’

/ġ/ > /g/: *ġanīma* (via Hausa *gànīmā̀*) > *ganíma* ‘booty, loot’; in one case /ġ/ > /i/: *maġrib* > *máirúwu* ‘prayer at sunset’

/f/ > /f/: *farḍ* > *fáral* ‘religious duty’; sometimes /f/ > /w/: *al-fūṭa* ‘serviette, towel’ > *alwúta* ‘handkerchief’

/q/ > /k/: *qabr* > *káwar* ‘grave’; in some cases /q/ > /w/: *sūq* > *kasúwu* ‘market’, */ṣandūq* > *sandúwu* ‘box’, *al-qurʾān* > *Luwurân* ‘the *Qurʾān*’; but in a very few cases /q/ > /g/, reflecting the pronunciation of some West African languages and local Arabic: *bunduqiyya* > *bə́ndə́g* ‘gun’, *xalq* > *álagə* ‘creature’, *xalq* > *álə́ga* ‘creation’, *qahwa* > *gáwa* ‘coffee’; in a few cases /q/ > Ø: *xalq* > *áláa* ‘creation’, *ṣadaqa* > *sádáa* ‘alms’, *ṭāqa* > *táa* ‘window’, *waqaʿa* > *waajîn*/*wakajîn* ‘to occur, happen’; in one case /q/ > /y/: *waṯīqa* ‘document, paper’ > *wotíya* ‘letter’

/k/ > /k/: *kāfūr* > *kawûr* ‘camphor’; in a few cases /k/ > /g/: *baraka* > *bárga* ‘blessing’; in a few cases /k/ > Ø: *ḥākam* > *háá* ‘fine’; in one case /k/ > /w/: *ḥukm* > *hówum* ‘judgment’

/l/ > /l/: *jumla* > *júmla* ‘total’

/m/ > /m/: *miṯāl* > *misâl* ‘example’; in one case /m/ > /w/: *al-maʿāš* ‘means of subsistence’ > *alwóshi* ‘salary’

/n/ > /n/: *qarn* > *kárnu* ‘century’; in one case /n/ > Ø in final position: *ṯaman* > *táma* ‘price’; in another case /n/ > /m/: *burnus* > *barmúsu* ‘burnous, hooded cloak’

/h/ > /h/: *hijra* > *híjíra* ‘the Hegira’; in a couple of cases /h/ > Ø: *jahannam* ‘hell’ > *jánəm* ‘fire’, *rahn* > *rân* ‘pawn; mortgage’

/w/ > /w/: *wahla* ‘fright, terror’ (via Hausa *wàhalā̀* ‘trouble’) > *wahála* ‘difficulties, troubles’; sometimes /w/ > Ø, especially in medial position: *jawāb* > *jaáwu*/*zaáwu* ‘answer’; in one case /w/ > /m/: *ṭāwūs* > *támus* ‘peacock’

/y/ > /y/: *qiyāma* > *kiyáma* ‘resurrection’

3. MORPHOLOGY, SYNTAX, AND SEMANTICS

In some cases, but not as frequently as in Hausa, the Arabic definite article *al*+ is retained as a prefix on the Kanuri word: *al-bišr* ‘joy’ > *albíshir* ‘good news’. In other cases, the *l* of the definite article has become the initial consonant of the loanword in Kanuri, followed in some cases by an epenthetic vowel: *al-qurʾān* > *Luwurân* ‘the *Qurʾān*’.

Apart from the case of the Arabic article retained in loans, there are cases where the *k*+ prefix has been applied as a word formative to the Arabic borrowing (*sūq* > *kasúwu* ‘market’). The Kanuri lexicon is made up of historically derived words that are now recognizable as composed of word formatives no longer productive in the language: *ʿadāla* > *nəmʾádal* ‘justice’, *al-ʾislām* > *kərmə́sələm* ‘Islam’.

Arabic loanwords were introduced in domains of the Kanuri cultural lexicon: jurisprudence (*sharâ* < *šarīʿa* ‘Muslim law’), trade (*shirgamá* ‘co-owner’ < *širka* ‘partnership’), religion (*kútúwa* < *xuṭba* ‘Muslim Friday sermon’), science and culture (*ílmu* < *ʿilm* ‘knowledge’, *dárasə* < *dars* ‘lesson’), names of some everyday objects (*shawûl* < *ṣābūn* ‘soap’).

Kanuri not only received Arabic loans but was also a donor language. Many languages acquired Arabic loans in West Africa through Kanuri: Arabic *sūq* ‘market’ > Kanuri *kasúwu* > Buduma *káhukū*; Hausa *kā̀suwā*; Kotoko *gásə̀gbí*; Musgu *kàskú*; Tubu *kasúgu*.

BIBLIOGRAPHICAL REFERENCES

Baldi, Sergio. 2002. “On Arabic loans in Kanuri”. *Studi Magrebini* 25.45–57.

——. 2003. “Some statistical analysis on Arabic loans in Kanuri”. *Lexical and structural diffusion: Interplay of internal and external factors of language development in the West African Sahel*, ed. Robert Nicolaï and Peter Zima, 9–15. Nice: Sophia Antipolis.

——. Forthcoming. “Dating Arabic loans in Kanuri”. *Trends in Nilo-Saharan linguistics: Proceedings of the 7th Nilo-Saharan Linguistics Conference, Vienna, September 1998*, ed. Norbert Cyffer. Cologne: R. Köppe.

Clarke, Peter B. 1982. *West Africa and Islam: A study of religious development from the 8th to the 20th century*. London: E. Arnold.

Cyffer, Norbert and John Hutchison (eds). 1990. *Dictionary of the Kanuri language*. Dordrecht: Foris.

Hutchison, John P. 1981. *The Kanuri language: A reference grammar*. Madison: University of Wisconsin.

Sergio Baldi
(University of Naples, L'Orientale)

Kaškaša and Kaskasa

The terms *kaškaša* and *kaskasa* refer to the phenomenon of using the suffixes /š/ and /s/, respectively, for the attached object pronoun of the 2nd person feminine singular (cf. Jindī 1983:I, 359–364). These suffixes were not used in Classical Arabic, but they occurred in some dialects in the Arabian Peninsula. The Arab linguists describe this phenomenon, although there is some inconsistency in their descriptions. Sībawayhi (*Kitāb* IV, 199–200), Ibn Jinnī (*Sirr aṣ-ṣināʿa* I, 219; *Xaṣāʾis*, II, 11–12), and Ibn Yaʿīš (*Šarḥ al-Mufaṣṣal* II, 9.48–49) recognize two groups of people in the Arabian Peninsula who use *kaškaša*. One group uses /š/ instead of Classical Arabic /ki/, as in *ʾaʿṭaytiš* 'I gave you [fem.]', and the other group adds /š/ to Classical Arabic /ki/, as in *ʾaʿṭaytikiš* 'I gave you [fem.]'. Both phenomena occur in pausa.

The explanation that Sībawayhi advances for this change is that the speakers wanted to make a distinction between males and females when addressing them. The Classical Arabic reference to a male person is /ka/, and both feminine /ki/ and masculine /ka/ lose their final vowel in pausa, so that the distinction between the two genders disappears there. Sībawayhi adds that the speakers select /š/ since it is a voiceless sound, just like /k/, e.g. *ʾinniš ḏāhiba* 'you [fem.] are leaving'; *māliš ḏāhiba* 'why are you [fem.] leaving?', instead of *ʾinnaki* and *mālaki*.

Kaskasa, on the other hand, is explained by Sībawayhi as the addition of /s/ to the feminine /ki/, as in *ʾaʿṭaytikis* 'I gave you [fem.]', by some speakers in pausa. Neither Sībawayhi nor Ibn Yaʿīš mentions the replacement of /ki/ by /s/, but Ibn Jinnī and other linguists do mention that /s/ replaces /ki/ in the speech of some speakers, just as in the case of *kaškaša*.

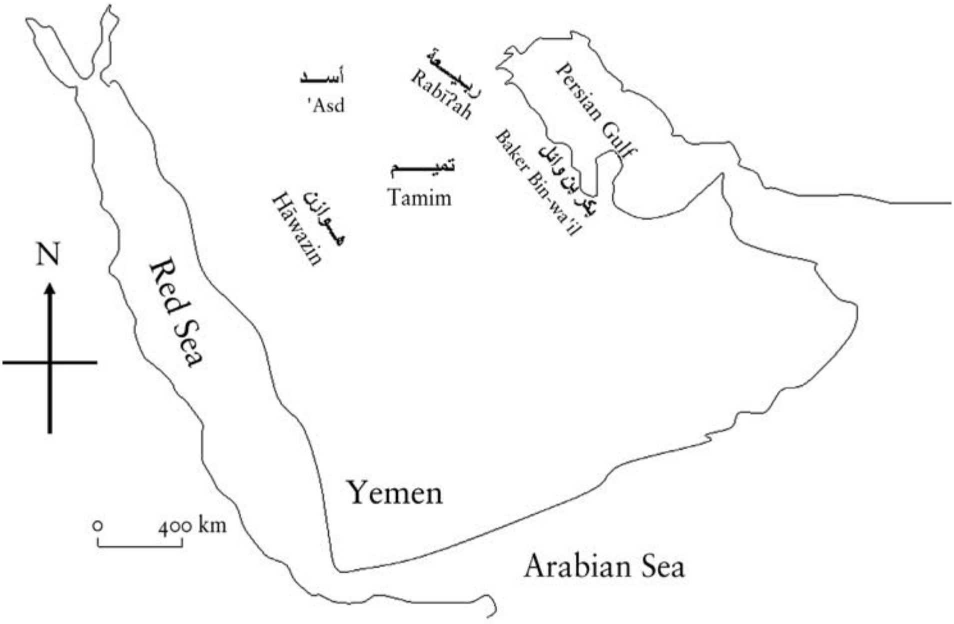
The Arabic grammatical literature mentions different tribes whose dialects exhibited these phenomena. *Kaškaša*, for instance, was said to be used among the Tamīm (Zamaxšarī, *Rabīʿ al-ʾabrār* IV, 273; Ibn Yaʿīš, *Šarḥ al-Mufaṣṣal* II, 49; Ṯaʿālibī, *Fiqh* 39). Ibn Manẓūr (*Lisān* XII, 232), as-Suyūṭī (*Muzhir* I, 221), and Ibn Jinnī (*al-Xaṣāʾiṣ* II, 11–12) conclude that *kaškaša* occurred among Rabīʿa. On the other hand, Ibn Yaʿīš (*Šarḥ, al-Mufaṣṣal* II, 49) and Ibn Fāris (*Ṣāḥibī* 33–36) conclude that Bakr ibn Wāʾil used to have *kaškaša*.

Kaskasa is mentioned for Bakr ibn Wāʾil as well (Ibn Manẓūr, *Lisān* XII, 80; Zamaxšarī, *Rabīʿ al-ʾabrār* IV, 273; Ibn Yaʿīš, *Šarḥ, al-Mufaṣṣal* II, 49; Ṯaʿālibī, *Fiqh* 39). Ibn Fāris (*Ṣāḥibī* 33–36) and as-Suyūṭī (*Muzhir* I, 221) conclude that *kaskasa* occurred in Rabīʿa, while Ibn Jinnī (*Xaṣāʾiṣ* II, 11–12) mentions that *kaskasa* occurred among the Hawāzin.

If one traces the occurrence of these phenomena among Arab tribes, it becomes obvious that there is a certain inconsistency in indicating the tribes. This inconsistency may have been caused by the huge size of some tribes. Rabīʿa, for example, was a large tribe that included several smaller tribes, such as ʾAsd and Bakr ibn Wāʾil, who settled in the north and east of the Arabian Peninsula. Another reason for the inconsistency may have been the continuous traveling of the Arabs, which led to a dispersal of dialectal features from one place to another through time. In the main, by tracing the locations of the old tribes using *kaškaša* and *kaskasa* on the map, one finds that all of them lived relatively close to each other, roaming from the middle to the north and east of the Arabian Peninsula (see Map 1). Incorrect description may have been another reason for the inconsistency in the descriptions (see Jindī 1983; ʾAnīs 1990).

The present-day situation in the Arabian Peninsula is slightly different from that described by the Arab linguists. In the east and northeast of the Arabian Peninsula, one finds that /č/, not /kiš/, is the form used to indicate the feminine attached pronoun. However, this change is not limited to the 2nd person attached feminine pronoun. Any /k/ may be replaced by /č/ in some positions, as in *čēf al-ḥāl* 'how are you?', *dīč* 'rooster', *fič* 'open!', instead of *kēf*, *dīk*, *fik*. In addition to the east and northeast of Saudi Arabia, Gulf countries such as Kuwait, Bahrain, Qatar, and the United Arab Emirates, Iraq, and the southern parts of Jordan and Syria have this feature as well (→ palatalization).

Map 1. Arabian Peninsula: tribes that used Kaškašah and kaskasah as mentioned by old Arab linguists



According to the description of the Arab linguists, /š/ for the 2nd person attached feminine pronoun is found in the province of Asir, which is located in the southwest of Saudi Arabia. We should note that there is another phenomenon mentioned by Sībawayhi and others, under the name of *šanšana*, which consists in the replacement of any /k/ by /š/ (cf. JindÌ 1983: I, 362–363). This phenomenon has survived in Yemeni Arabic (Watson 1992). The difference between the *šanšana* of Yemen and the *kaškaša* of Asir is that /š/ in Asir replaces only the attached pronoun of the 2nd person feminine singular. /s/ and /ć/ occur in the middle and some parts of the north of the Arabian Peninsula. Similarly to /č/, /s/ and /ć/ replace /k/ not only in the feminine attached pronoun but in other positions as well, as in *išlōnić* 'how are you [fem.]?', *ćalb* 'dog', *bāćir* 'tomorrow', instead of *išlōnik, kalb, bākir*.

This phenomenon is a feature of some Arabian peninsular dialects and exists until now. Nowadays, many speakers of these dialects are not aware that these features have a history in the pre-Islamic Arabic dialects and were discussed by the Arab linguists. They assume that it is a deviation from Classical Arabic and tend to shift to a neutral /k/ in an attempt to adopt the Classical Arabic form.

Bibliographical references

Primary sources

Ibn Fāris, *Ṣāḥibī* = ʾAbū l-Ḥusayn ʾAḥmad Ibn Fāris ibn Zakariyyāʾ, *aṣ-Ṣāḥibī*. Ed. ʾAḥmad Ṣaqr. Cairo: al-Babi Publisher, n.d.

Ibn Jinnī, *Sirr aṣ-ṣināʿa* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *Sirr ṣināʿāt al-ʾiʿrāb*. Ed. Muḥammad Ḥasan and ʾAḥmad Rušdī. Beirut: Dār al-Kutub al-ʿArabiyya, 2000.

——, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. Ed. Muḥammad an-Najjār. Beirut: Dār al-Kutub al-ʿArabiyya, 1957.

Ibn Manẓūr, *Lisān* = ʾAbū l-Faḍl Jamāl ad-Dīn Mukarram ibn Mukarram Ibn Manẓūr al-ʾAnṣārī, *Lisān al-ʿArab*. Cairo: ad-Dar al-Miṣriyya, n.d.

Ibn Yaʿīš, *Šarḥ al-Mufaṣṣal* = Muwaffaq ad-Dīn Yaʿīš ibn ʿAlī Ibn Yaʿīš, *Šarḥ al-Mufaṣṣal*. Beirut: ʿĀlam al-Kitāb, n.d.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān ibn Qanbar Sībawayhi, *Kitāb Sibawayhi*. Ed. ʿAbd as-Salām Muḥammad Hārūn. Cairo: al-Xānjī, 1982.

Suyūṭī, *Muzhir* = Jalāl ad-Dīn ʿAbd ar-Raḥman ibn ʾAbī Bakr as-Suyūṭī, *al-Muzhir fī ʿulūm al-luġa wa-ʾanwāʿihā*. Beirut: al-Maktaba al-ʿAṣriyya, 1987.

Ṯaʿālibī, *Fiqh* = ʾAbū Manṣūr ʿAbd al-Malik ibn Muḥammad aṯ-Ṯaʿālibī, *Fiqh al-luġa wa-sirr al-ʿarabiyya*. Lebanon: an-Nāširūn, 1997.

Zamaxšarī, *Rabīʿ al-ʾabrār* = ʾAbū l-Qāsim Maḥmūd ibn ʿUmar az-Zamaxšarī, *Rabīʿ al-ʾabrār*. Ed. Salīm an-Nuʿaymī, n.p., n.d.

Secondary sources

ʾAbū Xalīl, Šawqī. 2003. *ʾAṭlas as-sīra an-nabawiyya*. Damascus: Dār al-Fikr.

ʾAnīs, ʾIbrāhīm. 1990. *Fī l-lahajāt al-ʿarabiyya*. Cairo: Maktabat al-ʾAnglū al-Miṣriyya.

Ġunaym, Ṣālḥa al-. 1985. *Al-lahajāt fī l-Kitāb li-Sībawayhi ʾaṣwātan wa-bunya*. Jeddah: Dār al-Madanī.

Jindī, ʾAḥmad ʿAlam ad-Dīn al-. 1983. *al-Lahajāt al-ʿarabiyya fī t-turāṯ*. 2 vols. Tripoli: ad-Dār al-ʿArabiyya li-l-Kitāb.

Watson, Janet C.E. 1992. "*Kaškašah* with reference to modern Yemeni dialects". *Zeitschrift für Arabische Linguistik* 24.60–81.

Munira Ali Al-Azraqi
(King Faisal University)

Kazakh

Kazakh is a Central Asian Turkic language spoken by approximately 10 million people in Kazakhstan, Uzbekistan, Mongolia, and China. It belongs to the Qipchaq subgroup of the Turkic languages, within which it is in the first place closely related to Karakalpak, Bashkir, and Altay and secondly to Kyrghyz, Karachay-Balkhar, Kumyk, and Tatar (on the history of Kazakh, see Balaqaev and Sayrambaev 1997; Sïzdïqova 1993, 1994). Like most Turkic languages, the Kazakh lexicon includes a considerable number of Arabic loanwords, albeit fewer than, for instance, → Uzbek.

Arabic words entered Kazakh with the embracement of Islam by Turkic tribes in Central Asia, starting from the late 11th century. Through their nomadic lifestyle, which the Kazakhs managed to preserve up to the early 20th century, combined with the cultural and physical distance from the centers of Islamic culture and learning, they developed a nomadic variant of Islam. The Kazakhs retained their ancient shamanist focus on 'the spirits of the ancestors' (Kazakh *aruwah/aruwaq* < Arabic *ʾarwāḥ* [pl. of *rūḥ*] 'spirits') and on nature. The influence of Islam was reinforced in the 18th century by missionary activities of Tatar Muslims, which were directed by the Russian royal court in an attempt to control the peoples of the steppes (cf. Kreindler 1979). Most Arabic loans must have reached Kazakh, therefore, through a → Tatar filter.

The Kazakhs have put hardly anything in writing throughout their nomadic history. Although the first recordings of their ancient oral epics were made by Russian travelers in the 17th century, most date from the 19th century, when the second wave of Islamicization had already begun. The date of recording no doubt influenced the vocabulary of the epics (cf. Asanbayev a.o. 1995). It is therefore difficult to ascertain which words belong to the oldest layer of loans. An early loan is perhaps *mal*, the common word for 'cattle' (< Arabic *māl* 'possession'). An indicator might be the different treatment of the Arabic feminine ending *-a*, which is sometimes preserved as *-a* and sometimes as *-et*; these endings may belong to different layers in the borrowings (→ Persian).

Not surprisingly, the Arabic loans in Kazakh belong to the domains of religion and culture, e.g. Kazakh *bereket* < Arabic *baraka* 'blessing'; Kazakh *qalam* Arabic *qalam* 'pen'; Kazakh *tarïẖ* < Arabic *taʾrīx* 'history'; Kazakh *dünye* < Arabic *dunyā* 'earth'; Kazakh *momïn* < Arabic *muʾmin* 'believer'; Kazakh *mektep* < Arabic *maktab* 'school'. Yet, Kazakh *Qŭday* < Persian *Xuda* 'Lord' is the normal expression for 'God'. Sometimes, there is a slight shift in meaning, e.g. Kazakh *žanaza* 'corpse' < Arabic *janāza* 'funeral'.

Almost all Arabic loanwords in Kazakh are nouns; to form verbs from Arabic loanwords, Kazakh uses dummy verbs like *yetu* 'to do', e.g. *iyqpal yetu* 'to neglect' < Arabic *ʾihmāl* 'negligence', or *aytu* 'to say', e.g. *aytu salem* 'to greet' < Arabic *salām* 'greeting'; similar constructions are found in → Persian, → Turkish, and → Urdu.

Kazakh underwent a number of phonological shifts, some of which set it apart from other Turkic languages, such as the shift *e-* > *ye-*, e.g. *er* > Kazakh *yer* 'man'; *š* > *s*, e.g. *baš* > Kazakh *bas* 'head', followed by *č* > *š*, e.g. *čïq-* > Kazakh *šïq-* 'to leave'. The sequence of these shifts can to some extent be reconstructed, but due to the scarcity of written material, dating them is an extremely difficult enterprise. Even when ancient material is available, the Chagatay-based orthography in Arabic script in which Kazakh from the 17th century onward was written hardly reflects actual Kazakh phonology. Arabic loanwords too were subjected to a number of those phonological shifts, e.g. *Derbĭsèlĭ*, *derbĭs* + *èlĭ* < Arabic *darwīš ʿali* 'Darwish Ali'. Another such shift is *y-* > *ž-*, which Kazakh shares with Kyrghyz (e.g. *yol* > Kazakh *žol* 'road'); it is found in Arabic loanwords such as the Arabic proper name *Yūsuf* > Kazakh *Žüsĭp* 'Joseph', and Arabic *yatīm* 'orphan' > Kazakh *žetim*.

In the first decades of the 20th century, Kazakh was written in the Arabic, Latin, and Cyrillic scripts more or less specially adapted for Kazakh but none of which adequately met all aspects of Kazakh phonology (cf. Baldauf 1993). The Cyrillic alphabet, for example, still does not reflect the full vowel harmony, e.g. көзім (*közĭm*) for [közüm] 'my eye'.

Arabic loanwords have undergone shifts also encountered by Arabic loanwords in many other languages, such as the loss of emphasis in emphatic consonants, e.g. Arabic *ḍ* > *z*, Arabic *ḥāḍir* 'present' > Kazakh *azĭr/qazĭr* 'now'; *ṣ* > *s*, e.g. Arabic *ʾaṣl* > Kazakh *asĭl/èsĭl* 'base'; generally *-b* (and -C$^{\text{voiced}}$p) > *p*, e.g. Arabic *kitāb* > Kazakh *kĭtap* 'book'. Foreign phonemes readily change into a Kazakh one, e.g. Arabic *f* > Kazakh *p* (e.g. Arabic *farḍ* 'prescription' > Kazakh *parïz*; Arabic *kafīl* > Kazakh *kèpĭl*, as in *kèpĭldĭ* 'with warranty [i.e. 'guaranteed']'; Arabic *fāʾida* 'profit' > Kazakh *payda*); Arabic ʿ > Kazakh *ġ*/Ø (e.g. Arabic *ʿilm* 'science' > Kazakh *ġïlïm*; Arabic *ʿajāʾib*, pl. of *ʿajība*, 'strange things, miracles' > Kazakh *ġažayïp* 'splendid'; Arabic *ʿaql* 'common sense' > Kazakh *aqïl*); Arabic *ḥ* > Kazakh *q*/Ø (e.g. Arabic *ḥāja* 'need' > Kazakh *qažet/èžet* 'obligation'. Arabic *ḥāl* has two realizations: (i) Kazakh *qal* 'well-being; situation', as in *qalïŋ qalay?* 'how are you?', and (ii) Kazakh *èl* 'power', e.g. *èlsĭz* 'sick, powerless'; Arabic *ḥaqīqa* 'truth' > Kazakh *aqiqat*; Arabic *ḥuqūq* 'laws' [pl. of *ḥaqq*] > Kazakh *qŭqïq* 'law').

Since religion was officially prohibited during the seventy years of Soviet rule, many Arabic loanwords in the religious domain remained virtually invisible (on Islamic peoples in the Soviet Union, see Akiner 1986). The revival of Islam in the last decade of the 20th century has not only caused the religious terms to surface again, most of them—apart from the most evident (*aqĭret* 'the hereafter' < Arabic *ʾāxira; ažal* 'fate' < Arabic *ʾajal* 'fate'; *èwliye* 'saint' < Arabic *ʾawliyāʾ*, pl. of *wālī*)—not understood by the nonreligious part of the population, but it also brought reloans into the language, e.g. Kazakh *ġibadat* < Arabic *ʿibāda* 'prayer, adoration'. There must have been a certain reluctance against adapting the pronunciation of religious notions to Kazakh phonology; in this way Arabic *šayṭān* 'devil' has survived with *š-*, but in Kazakh dialects the regular Kazakh form *seytan* is still found. In 19th-century poetry, lines like the following can still be found: *Yekewi köp tilekpen dŭġa qïldï / Qŭdayïm dŭġalarïn qabïl qïldï* 'Both of them *prayed* with many intentions / My God *accepted* their prayers' (Qaliyasqarŭlï 1995:183).

One particular phenomenon of Kazakh is dissimilation of consonants, for instance in the plural suffix *-lar*, in which *l-* changes to *d-* (voiced) or *t-* (voiceless), e.g. *qol* 'hand' ~ *qol-lar* 'hand' > Kazakh *qoldar*. This also occurs in Arabic loanwords, e.g. Kazakh *žendet* < Arabic *janna* 'heaven'; *iyqpal (yetu)* 'to neglect' < Arabic *ʾihmāl* 'negligence' (with Kazakh *q* < Arabic *h* producing **iyqmal* through dissimilation of *qm* > *iyqpal*); Kazakh *molda* (as in the name *Moldabek*) in which *molda* is 'mullah' + *bek* 'lord' (which has now reentered the language as *molla*). A striking example of dissimilation is Arabic *Muḥammad*, which in Kazakh has different forms, e.g. *Membet, Maḫambet, Maḫanbet* (in which *mb/nb* < *mm* through dissimilation).

Abay's famous song is still full with Arabic loanwords: *Ayttïm salem, qalamqas / saġan qŭrban mal men bas* 'I greet you, o [girl] with your brows [thin like a] pencil / [I am] sacrificed to you [with all my] livestock and [my] head' (Qŭnanbayŭlï 1976:85).

BIBLIOGRAPHICAL REFERENCES

Akiner, Shirin. 1986. *Islamic peoples of the Soviet Union*. London and New York: Kegan Paul International.

Asanbayev, B. a.o. 1995. *Batïrlïq dastandar* [Heroic epics]. Almaty: Ġïlïm.

Balaqaev, M. and T. Sayrambaev. 1997. *Qazïrġï qazaq tĭlĭ* [The modern Kazakh language]. Almaty.

Baldauf, Ingeborg. 1993. *Schriftreform und Schriftwechsel bei den muslimischen Russland- und Sowjettürken (1850–1937): Ein Symptom ideengeschichtlicher und kulturpolitischer Entwicklungen*. Budapest: Akademiai Kiadó.

Bizaqov, S. 1997. *Tïldĭk norma zhène varianttïlïq* [Linguistic norm and variability]. Almaty.

Kreindler, Isabelle T. 1979. "Nikolai Il'minskii and language planning in nineteenth-century Russia". *International Journal for the Sociology of Language* 22.5–26.

Qaliyasqarŭdï, Qabidolda (ed.). 1995. *Ḫalïq žïrlarï*, I [People's narrative poems]. Ölgiy: Monġoliya Ġïlïm Akademiyasï [Academy of Sciences of Mongolia].

Qŭdayberdiev, Šèkerĭm. 1988. *Šïġarmalarï* [Articles]. Almaty: Žazušï.

Qŭnanbayŭlï, Ibrahim [Abay]. 1976. *Öleŋder* [Songs]. Almaty: Žazušï.

Sïzdïqova, Rèbiġa. 1993. *Qazaq èdebi tĭlĭnĭŋ tarĭhï* [A history of the Kazakh literary language], Almaty.

——. 1994. *Sözder söyleydĭ: Sözderdiŋ qoldanïlu*

tarixïnan [Words are speaking: From a semantic history of Kazakh]. Almaty.

Robert Ermers
('s Hertogenbosch, The Netherlands)

Kenya → East Africa

Khartoum Arabic

1. General

1.1 *Area, range*

The dialect referred to in this article as Khartoum Arabic is spoken in Greater Khartoum (Khartoum, Khartoum North, and Omdurman) and in other urban areas of central Sudan, roughly to the towns of Atbara in the north, Sennar on the Blue Nile, and Kosti on the White Nile.

1.2 *Speakers*

Sudan is a multilingual country, with Arabs making up around 40 percent of the population. Khartoum Arabic is the prestige Arabic dialect and has several million native speakers. Most of these are descendants of migrants into the cities of central Sudan during the 20th century.

1.3 *Position*

Sudan has a well-developed tradition of writing and performing in colloquial Arabic. The University of Khartoum Centre for Afro-Asian Studies has published collections of folk literature and oral histories in various Arabic dialects, a number of well-known literary figures have published collections of poetry in Sudanese Arabic, literary figures such as aṭ-Ṭayyib Ṣāliḥ have made use of colloquial Arabic in their writing, and there is a well-established tradition of colloquial Arabic broadcasting – plays, soap operas, and folkloric material, in particular.

1.4 *Linguistic type*

Khartoum Arabic is an Eastern-type Arabic dialect and seems to be more closely related to the dialects of Upper Egypt than to any other non-Sudanese dialect, although there are also Peninsular and North African influences. Historically, this reflects the major penetration route of Arabic speakers, from Upper Egypt and through Nubia (via the Islamization of the Nubian kingdoms) into central Sudan.

Within Sudan, Khartoum Arabic is closely related to the dialects of the Jaʾaliyyin tribal grouping, found to the north of Khartoum.

Table 1. Inventory of consonants of Khartoum Arabic

place of articulation	bilabial	labio-dental	apico-dental	apico-alveolar	dorso-prepalatal	dorso-palatal	post-dorso-velar	post-dorso post-velar	pharyngeal	glottal
manner of articulation										
plosives										
voiced										
plain	b		d		j		g			
emphatic				ḍ						
voiceless										
plain			t		č		k			
emphatic				ṭ						
fricatives										
voiced										
plain			z					ġ		ʾ
emphatic				ẓ					ʿ	
voiceless		f								
plain			s		š			x		h
emphatic				ṣ					ḥ	
trills										
plain				r						
emphatic				ṛ						
laterals										
plain				l						
emphatic				ḷ						
nasal	m			n	ñ					
glide	w					y				

The history of Khartoum Arabic is reflected in its lexicon. This includes old borrowings from Aramaic (including Syriac), Ancient Egyptian (via Nubian) and Coptic, and Persian (sometimes via Ottoman Turkish); more recent borrowings from Ottoman Turkish, English, French, Italian, Greek, Ethiopian languages, and Egyptian Arabic; as well as words from other Sudanese languages, particularly Nubian, Beja, and Fur.

1.5 *State of research*

Research into Sudanese Arabic began fairly early in the 20th century, impelled by the requirements of the British colonial administration and the personal interests of members of the British Sudan Service. Early works include Worsely (1925), Hillelson (1930), and Trimingham (1946).

More recent works include Kaye (1976) and Persson and Persson (1980). Collections of texts have been produced by Hillelson (1935), Yagi (1981), and Bergman (2002), which includes an account of the grammar (1–63). Particularly important is the Sudanese Arabic/Standard Arabic dictionary of Qāsim (2002).

Phonology is the only linguistic subdiscipline to have received detailed attention. There have been studies by Mustapha (1982; particularly recommended), Blair (1983), and Hamid (1984).

2. PHONOLOGY

2.1 *Phonemics*

/č/ and /ñ/ occur in only a small number of words and are not in the inventory of all speakers. Speakers who do not have /č/ and /ñ/ use /š/ and /n/ respectively, instead: thus *čarra* ~ *šarra* 'to gurgle [of water]', *ñarra* ~ *narra* 'to growl [of a dog]'.

Minimal /r/ vs. /ṛ/ and /l/ vs. /ḷ/ pairs include *jabar* 'to set [a broken bone]' vs. *jabaṛ* 'to force' and *galam* 'to prune' vs. *gaḷam* 'pen'. However, /ḷ/ never appears word-initially (cf. Mustapha 1982:61), and /ṛ/ is particularly subject to internal root allomorphy with /r/; thus *jaṛṛāḥ* 'surgeon' and *jariḥ* 'wound' belong to the same root.

/ʾ/ is positionally restricted, occurring only syllable-initially. Both /ˁ/ and /ʾ/ are more reasonably analyzed as fricatives than plosives, /ʾ/ typically being realized as creaky voice.

There are virtually no occurrences of the apico-alveolar emphatics /ḍ/, /ṭ/, /ẓ/, /ṣ/, /ṛ/, /ḷ/ in the environment of /ġ/ and /z/ (exceptions are *ṣāġ*, a rank in the Ottoman army, and some words borrowed from Standard Arabic). In the current state of the language, this is phonetically unmotivated since /ġ/ and /x/ do not themselves have emphatic-type realizations. At an earlier stage, however, /ġ/ and /x/ seem to have been emphatics and to have yielded de-emphasization of other emphatics in their environment.

Khartoum Arabic has developed a 4-way (2 × 2) emphatic/plain, voiced/unvoiced opposition out of the original Arabic 3-way emphatic, voiced (or unaspirated), unvoiced (or aspirated) opposition. The 2 × 2 opposition is apparent in its full form in the apico-dental and apico-alveolar series: /ḍ/, /ṭ/, /d/, /t/ and /ẓ/, /ṣ/, /z/, /s/, and in the guttural series: /ˁ/, /ḥ/, /ˀ/, and /h/, where /ˁ/ and /ḥ/ are the emphatic counterparts of /ˀ/ and /h/, one phonetic interpretation of this being that /ˁ/ and /ḥ/ are primarily glottals with secondary pharyngealization (cf. Jakobson 1957:112). The major phonetic correlate of emphasis in Khartoum Arabic is pharyngealization. In the case of /ḷ/ and /ṛ/, however, the phonetic correlate may be velarization.

Khartoum Arabic has three short vowels, /u/, /a/, /i/, and five long vowels, /ū/, /ā/, and /ī/, /ō/ and /ē/. It has four diphthongs. By far the most common are /aw/ and /ay/. However, /iw/ and /uy/ also occur, e.g. *siwsīw* 'chicken', *buyḍ* 'white [pl.]' in liaison before a following vowel (citation form *buyuḍ*).

2.2 *Syllable structure*

Khartoum Arabic has the following syllable types:

Type 1	Cv̲	e.g.	*da*	'this [masc.]'
Type 2	Cv̲C	e.g.	*bu̲n*	'coffee beans'
Type 3	Cv̲̄C	e.g.	*bē̲t*	'house'
Type 4	Cv̲CC	e.g.	*ba̲nk*	'bank [financial institution]'
Type 5	Cv̲̄CC	e.g.	*tā̲ks*	'taxi'

Khartoum Arabic also has a high tone, represented here as $^{\uparrow}$, and occurring most commonly in running text on 1st person singular pronoun suffixes, e.g. *ummi*$^{\uparrow}$ 'my mother', *lēy*$^{\uparrow}$ 'to me'.

2.3 *Consonant clusters*

Syllable-final consonant clusters are of two types:

i. Sonorant /m/, /n/, or /l/ followed by a fricative /f/, /s/, /z/ /š/, or by a stop, normally at an adjacent, but less commonly at the same point

Table 2. Independent personal pronouns

3rd sg. masc.	*hu*	2nd sg. masc.	*inta*	1st sg.	*ana*
3rd sg. fem.	*hi*	2nd sg. fem.	*inti*		
3rd masc. pl.	*hum*	2nd pl. masc.	*intu*	1st pl.	*niḥna*
[3rd pl. fem.	*hin*]	[2nd pl. fem.	*intan*]		

Table 3. Possessive/object suffixes

3rd sg. masc.	*-hu/-u*	2nd sg.	*-ak/-k*	1st sg.	*-y*$^{\uparrow}$/*i*$^{\uparrow}$/*-ni*$^{\uparrow}$
3rd sg. fem.	*-ha/-a/-a*$^{\uparrow}$/*i*$^{\uparrow}$	2nd sg. fem.	*-ik/-ki*		
3rd pl. masc.	*-hum/-um*	2nd pl. masc.	*-kum*	1st pl.	*-na*
[3rd pl. fem.	*-hin/-in*]	[2nd pl. fem.	*-kan*]		

of articulation. Sonorant-stop cluster forms include *jamb* 'beside', *asmant* 'cement' (but also *asmantiy* among noneducated speakers). Sonorant-fricative cluster forms include *hals* 'bad, corrupt'.

ii. Fricative /f/ or trill /r/ followed by a stop or, in the case of /r/ only, also a fricative, typically at an adjacent (but not identical) place of articulation. Fricative/trill-stop cluster forms include *zift* 'pitch, tar' (but also *zifti* among noneducated speakers), while fricative/trill-fricative cluster forms include *kurs* 'course [educational]'.

3. MORPHOLOGY

Feminine plural pronouns and verbs are considered a rural feature (and are placed in square brackets throughout this entry). Where a sound plural is used, there seems to be an increasing tendency for feminine plural human nouns to take masculine plural adjectives.

3.1 *Independent pronouns*

The independent personal pronouns are set out in Table 2.

3.2 *Possessive/object suffixes*

The possessive/object suffixes are: set out in Table 3.

Where the possessive/object suffix has two forms (allomorphs), the vowel-initial form is postconsonantal and the consonant-initial form postvocalic. The only exception is the 1st person singular suffix form *-ni*$^{\uparrow}$, which occurs as the object suffix on verbs and participles and also in *fīni*$^{\uparrow}$ 'in me'.

Where the suffixes *-a* 3rd person singular feminine, *-um* 3rd person plural masculine, and [*-in* 3rd pers. pl. fem.] are preceded by a short vowel, the main stress falls on the syllable immediately preceding the pronoun suffixes (thus *gaḷam* 'pen', but *gaḷáma* 'her pen'). This is a relic of a previous stage of the language, where these pronoun suffixes were here, as elsewhere, *-ha* 3rd person singular feminine, *-hum* 3rd person plural masculine, and *-hin* 3rd person plural feminine, a CvC suffix on this type yielding a syllable structure in which Sudanese standardly has stress on the penultimate syllable.

3.3 *Demonstratives*

The demonstratives are as follows:

Near: 'this', 'that', 'it [of inanimate objects]'

sg. masc.	*da*
sg. fem.	*di*
pl. com.	*dēl*

Far: 'that over there; that which is conceptually/discoursally distant'

sg. masc.	*dāk*
sg. fem.	*dīk*
pl. com.	*dēlak*; also *dēk*

3.4 *Interrogatives*

The main interrogatives are the following:

minu$^{\uparrow}$	'who?'
šinu	'what?'
yātu$^{\uparrow}$	'which?'
wēn	'where?'
mitēn	'when?'
kēf	'how?'

yātu↑ 'which?' occurs prenominally and postnominally. Prenominally, *yātu*↑ is invariable: *yātu*↑ *rājil* 'which man?', *yātu*↑ *mara* 'which woman?', *yātu*↑ *awlād* 'which children?'. Postnominally, *yātu*↑ takes different suffixes in agreement with the preceding noun (which occurs with the definite): *arrājil yātu*↑ 'which man?', *almara yātī*↑ 'which woman?', *alawlād yātum*↑ 'which children?'. Where interrogative pronouns take pronoun suffixes, the singular feminine suffix form is normally *-i*↑ rather than *-a*.

kēf 'who?' and *wēn* 'where?' optionally take pronoun suffixes; *kēfak, kēfik*, etc. are fairly common ways of saying 'how are you?'

Interrogatives may occur in the same place as corresponding noninterrogative elements: e.g. *māši wēn* 'where are you (etc.) [sg. masc.] going?' (e.g. *māši ssūg* 'I (etc.) [sg. masc.] am going to the market'), *jīt maᶜa minu*↑ 'who did you come with?' (e.g. *jīt maᶜa ḥmad* 'I came with Aḥmad').

Interrogatives may also occur sentence-initially; when they do, they are a major sentence element, e.g. main predicand, main predicate, or directly dependent on the verb. (The term 'predicand' is used in this entry as in Bohas a.o. 1990 and Watson 1993) to cover both the subject of a verb and the 'initial element' [*mubtadaʾ bi-hi*] of a verbless sentence.) Thus, *wēn māši* 'where are you (etc.) [sg. masc.] going?'. A postprepositional interrogative can only occur sentence-initially as a complete prepositional phrase, e.g. *maᶜa mnu*↑ *jīt* 'who did you come with?'. Sentence-initial interrogatives are somewhat emphatic.

3.5 *Adverbs*

3.5.1 Temporal adverbs

Temporal adverbs include:

hassaᶜ ~ hassi	'now; just now [past or future]'
yā dōb	'just now [past]'
gibēl	'recently'
garīb	'recently; soon'
baᶜdēn	'soon; afterward'
baᶜad dāk	'afterwards, after that'
allēla	'today'
amis	'yesterday'
bākir	'tomorrow'
bukra	'tomorrow' (possibly originally a borrowing from Egyptian, but extensively used)
yōm dāk	'a long time ago'
wakit dāk	'at that time [in the past]'
yōmáta	'at that time [in the past]', with impersonal use of the feminine singular *-a* pronoun suffix
badri	'a long time ago; early'
zamān	'a long time ago'

Forms with the demonstrative are also common to emphasize immediacy: *hassaᶜ da* 'right now' (*hassaᶜ* is masc.), *hassi di* 'right now' (*hassi* is fem.), *allēla di* 'this very day'.

3.5.2 Local adverbs

Local adverbs include:

hina	'here'; also *hina da* 'right here'
hinād	'there'
fōg	'up, upward'
tiḥit	'down, below'
šamāl ~ šimāl	'left; north'
yamīn	'right'

3.5.3 Other adverbs

These include:

kida (also *kadē*)	'so, thus'
xālis	'very'
jiddan	'very'
jadd	'very' (esp. with adjectives such as *ṣaᶜab* 'difficult' and *zaᶜlān* 'angry')
šiwēya	'a little'
ḥabba	'a little'

3.6 *Verbs*

3.6.1 Form I

There are two Form I patterns: CaCaC and, less commonly, CiCiC.

3.6.1.1 Form I perfects

3.6.1.1.1 CaCaC perfects

Prototypically, CaCaC verbs have agentive subjects:

katab	'he wrote'
daras	'he studied'

However, the subject may also be the patient:

wagaᶜ	'it fell'
barad	'it [objectively] got cold' (cf. *birid*, below)

Table 4. Conjugation of the perfect *daras* 'to study'

3rd sg. masc.	*daras*	2nd sg. masc.	*darasta*	1st sg.	*darasta*
3rd sg. fem.	*darasat*	2nd sg. fem.	*darasti*		
3rd pl. masc.	*darasu*	2nd pl. masc.	*darastu*	1st pl.	*darasna*
[3rd pl. fem.	*darasan*]	[2nd pl. fem.	*darastan*]		

Table 5. Conjugation of the perfect *simiʿ* 'to hear'

3rd sg. masc.	*simiʿ*	2nd sg. masc.	*simiʿta*	1st sg.	*simiʿta*
3rd sg. fem.	*simʿat*	2nd sg. fem.	*simiʿti*		
3rd pl. masc.	*simʿu*	2nd pl. masc.	*simiʿtu*	1st pl.	*simiʿna*
[3rd pl. fem.	*simʿan]*	[2nd pl. fem.	*simiʿtan*]		

Table 6. Conjugation of the perfect *ḥabba* 'to love'

3rd sg. masc.	*ḥabba*	2nd sg. masc.	*ḥabbēt*	1st sg.	*ḥabbēt*
3rd sg. fem.	*ḥabbat*	2nd sg. fem.	*ḥabbēti*		
3rd pl. masc.	*ḥabbu*	2nd pl. masc.	*ḥabbētu*	1st pl.	*ḥabbēna*
[3rd pl. fem.	*ḥabban*]	[2nd pl. fem.	*ḥabbētan]*		

Table 7. Conjugation of the perfect *ṣām* 'to fast' (*ā-u* alternation)

3rd sg. masc.	*ṣām*	2nd sg. masc.	*ṣumta*	1st sg.	*ṣumta*
3rd sg. fem.	*ṣāmat*	2nd sg. fem.	*ṣumti*		
3rd pl. masc.	*ṣāmu*	2nd pl. masc.	*ṣumtu*	1st. pl.	*ṣumna*
[3rd pl. fem.	*ṣāman*]	[2nd pl. fem.	*ṣumtan*]		

3.6.1.1.2 CiCiC perfects

The final /i/ of the CiCiC perfect pattern drops out before vowel-initial subject suffixes.

CiCiC verbs typically express the following notions:

becoming/getting into a state:	*biga* 'to become'; *kibir* 'to grow big, old'
involuntary activities:	*ʿirig* 'to sweat, perspire'; *wildat* 'she gave birth'
(loss of) senses:	*simiʿ* 'to hear'; *xiris* 'to go dumb'
subjective experience:	*girif* 'to get fed up'; *ziʿil* 'to get angry'; *birid* 'he [subjective] got/felt cold' cf. *barad*, above)

3.6.1.1.3 Perfect of weak verbs

Geminated verbs

IIIy CaCaC verbs have the same basic pattern as geminated verbs. Thus, *jarēt* 'I ran' (etc.), *jara* 'he ran', *jarat* 'she ran', *jaru* 'they ran'.

IIIy CiCiC verbs take the ending /ī/ before consonant-initial suffixes, thus from *biga* 'to become': *bigīt* 'I became'. Elsewhere, they take the same presuffix endings as final weak CaCaC verbs: *biga* 'he became', *bigat* 'she became', *bigu* 'they [pl. masc.] became'.

The great majority of II*w*/*y* verbs belong to one of two closely related perfect unsound patterns: *ā-u* alternation, or *ā-i* alternation.

Most II*w*/*y* verbs follow the same *ā-u* alternation pattern as *ṣām:* CāC before vowel-initial suffixes and CuC before consonant-initial suffixes. A somewhat smaller proportion have an *ā-i* alternation pattern, thus *ṭār* 'he flew', *ṭirna* 'we flew'.

The verb *ja* 'to come' (root *j-y-y*) is irregular, having the form *j-* before most consonant-initial suffixes (*jīt* 'I came'), *ja-* before almost all vowel-initial suffixes (*jāt* = /ja-at/) 'she came', and the monophthongized form *jō* 'they [pl. masc.] came' (=**jaw*).

3.6.1.2 Form I imperfects

The initial vowel of Form I imperfects is /a/, and there are three imperfect patterns: yaCCuC, yaCCiC, and yaCCaC.

Table 8. Conjugation of the imperfect *yadrus* 'he studies'

3rd sg. masc.	*yadrus*	2nd sg. masc.	*tadrus*	1st sg.	*adrus*
3rd sg. fem.	*tadrus*	2nd sg. fem.	*tadrusi*		
3rd pl. masc.	*yadrusu*	2nd pl. masc.	*tadrusu*	1st pl.	*nadrus*
[3rd pl. fem.	*yadrusan*]	[2nd pl. fem.	*tadrusan*]		

The *a-* imperfect is illustrated by *yasmaˁ* 'he hears', and the *i-* imperfect by *yaktib* 'he writes'. Apart from the vowel difference, both conjugate exactly like the *u-* imperfect.

All *u-* imperfects have *a-* perfects. Almost all *i-* imperfects have *a-* perfects. However, a small number have aCCaC (i.e. Form IV) perfects. Examples are *yasriˁ* 'he hurries' (perfect *asraˁ*); borrowings from Standard Arabic, such as *yasbit* 'he demonstrates' (perfect *asbat*); and the very common verb *yaddi* 'he gives' (perfect *adda*; root *d-d-y*).

Geminated verbs have either *u-* or *i-* imperfect: *yaḥibb* 'he loves', *yaġušš* 'he deceives' (perfect *ġašša*). The affixes are, in both cases, the same as for sound verbs.

II*w/y* verbs have either (i) medial /ū/, in almost all cases corresponding to perfect *ā-u* alternation, e.g. *yaṣūm* 'he fasts' (perfect *ṣām*), *yadūr* 'he goes round' (perfect *dār*); (ii) medial /ī/, e.g. *yaṭīr* 'he flies' (perfect *ṭār*), *yadīr* 'he administers' (derived from Standard Arabic Form IV); or (iii) very rarely, medial /ā/, e.g. *yaxāf* 'he fears' (perfect with *ā-u* alternation; *xāf* 'he feared', *xufta* 'I feared'). Subject affixes are the same as for sound verbs.

III*y* verbs have either final /i/, e.g. *yajri* 'he runs' (from *jara*) or /a/ e.g. *yabda* 'he begins' (from *bada*). The final /i/ or /a/ disappears before vowel-initial subject suffixes: e.g. *yajru* 'they [pl. masc] run', *tabdi* 'you [sg. fem.] begin'.

Verbs with medial and final weak radicals are medially sound but finally as for other weak final verbs: *yašwi* 'he grills' (perfect *šawa*), *yaˁya* 'he becomes ill' (perfect *ˁiya*).

I*w* verbs take various imperfect patterns. Some lose the /w/ of the perfect, e.g. *yagaˁ* 'he falls' (perfect *wagaˁ*); some have /ō/ as a reflex of root /w/, e.g. *yōjaˁ* 'it hurts' (perfect *wajaˁ*); and some retain the /w/, e.g. *yawzin* (but for some speakers *yōzin*) 'it weighs' (perfect *wazan*). The verb *wagaf* 'to stop, stand' is reinterpreted as a hollow verb in the imperfect: *yagīf* 'he stops, stands'.

3.6.2 Form II

Form II is produced by doubling of C_2. It is extremely common and has a very wide range of meaning correlates, the most important of which are the following:

Same meaning as Form I

kammal	'to finish [intrans.]' (= *kimil*; note that *kammal* is also used causatively)

Causative of Form I

The notion of causative covers a range of meanings from genuine causation to permission and enabling.

bagga	'to cause to become' (*biga* 'to become')
rabba	'to grow [e.g. a beard]' (*riba* 'to grow [intrans.]')

Causative of other forms

saffar	'to cause to travel' (Form III *sāfar* 'to travel')
šaġġal	'to cause to work' (Form VIII *ištaġal* 'to work) [intrans.]

Intensive

gaffal	'to close up [completely; intr. and trans.]'

Distributive

ḍabbaḥ	'to slaughter [lots of animals]' (*ḍabaḥ* 'to slaughter')

Accusational

sarrag	'to accuse of stealing, accuse of being a thief'
jahhal	'to consider ignorant, accuse of ignorance'

Having a disease/defect

nammal	'to get pins and needles'
ṣaddaˁ	'to get a headache'
ḥawwaṣ	'to go cross-eyed'

Becoming a color

zarrag	'to turn black; to make black' (root *z-r-g; azrag* 'black')
xaddar	'to turn brown/green; to make brown/green' (root *x-d-r; axdar* 'brown/green')

The imperfect of Form II is on the yiCaCCiC pattern: *darras* 'he taught', *yidarris* 'he teaches'.

3.6.3 Form III

Form III is produced by lengthening the initial /a/ vowel of the verb. Form III verbs typically express the following:

Action involving two people with subject as agent and object as patient

ġāmaz	'to wink at'
ʿālaj	'to treat, cure'
šāwar	'to consult'

Reciprocal relationship with discoursally foregrounded entity as subject

sāwa	'to be equal to'
bādal	'to exchange'
ḍārab	'to hit [someone who is hitting you]'

Other

ḥāwal	'to try'
sāfar	'to travel'
ʿāyan lē	'to look (at)'
bārak lē	'to congratulate'

The imperfect of Form III is on the yiCāCiC pattern: *šāfar* 'he traveled', *yisāfir* 'he travels'.

3.6.4 Form IV

Form IV is produced by adding an *a-* prefix in the perfect. Form IV verbs are rare, and most are recent borrowings from Standard Arabic. They include:

asmʿ	'to hurry'
adda	'to give' (root *d-d-y*)
aḍrab	'to go on strike' (from Standard Arabic)

The imperfect of Form IV is on the yaCCiC pattern: *yasriʿ* 'he hurries'.

3.6.5 Form V

Form V is produced by adding an *it-* prefix to Form II. The main meanings of Form V are the following:

Passive of Form II

itkassar	'to be smashed up' (Form II *kassar* 'to smash up')

Reflexive of Form II

itġatta	'to cover oneself' (root *ġ-t-y*)
itgaṭṭaʿ	'to become split up into' (root *g-ṭ-ʿ*)

Reciprocal of Form II

itwannas	'to chat (with one another)' (cf. *wannas* 'to chat to')

Reciprocality in this article is defined as subsuming both the notions of activity directed at one another (e.g. *itšākal* 'to quarrel with one another') and that of doing things together (e.g. *itnōna* 'to buzz together', not necessarily 'to buzz at one another').

Acting or pretense

itkabbar	'to act arrogantly, be arrogant'

The imperfect of Form V is on the yitCaCCaC pattern: *itkallam* 'he talked', *yitkallam* 'he talks'.

3.6.6 Form VI

Form V is produced by adding an *it-* prefix to Form III. The main meanings of Form VI are the following:

Reflexive

itgāwal	'to contract (oneself) to do' (root *g-w-l*)
itʿāhad	'to undertake [= get oneself to undertake] to do'

Reciprocal (esp. of Form III)

itšākal	'to quarrel with one another' (Form III *šākal* 'to quarrel with')
itgābal	'to meet one another' (Form III *gābal* 'to meet')

Acting as/pretense

itnāsa	'to pretend to forget'
itġāba	'to pretend to be an idiot'
itẓāhar (*bē*)	'to pretend to'

The imperfect of Form VI is on the yitCāCaC pattern: *itkāsal* 'he was (too) lazy (to do something), *yitkāsal* 'he is (too) lazy (to do something)'.

3.6.7 Form VII

Form VII is produced by adding an *in-* prefix to the root. Form VII gives a middle sense. It is unproductive in Khartoum Arabic and is confined to a small number of words, although in many Sudanese dialects it is standardly used to make the passive, where Khartoum Arabic uses the root prefix *it-*. Examples are:

inbasaṭ	'to become happy' (also *itbasaṭ*)
indahaš	'to become surprised, astonished'
inṭalag	'to become loose, dissolute'

That Form VII is older than the use of the root prefix *it-* is suggested by the frequent use of Form VII in proverbs, e.g. *šōkat ḥūt la tanbaliᶜ la tafūt* 'caught on the horns of a dilemma' (lit. 'a fishbone; it can't be swallowed and it won't go away').

The imperfect of Form VII is on the yanCaCiC pattern: *yanbasiṭ* 'he becomes pleased'. Some speakers use the yinCaCiC pattern (*yinbasiṭ*).

3.6.8 Form VIII

Form VIII is produced by adding a *-t-* post- C_1 infix. Form VIII often gives a residual reflexive sense, e.g. *ištaġal* 'to work', or reciprocal sense, e.g. *ijtamaᶜ* 'to meet (together)'.

The imperfect of Form VIII is on the yatCaCiC pattern: *yaštaġil* 'he works'. Some speakers use the yitCaCiC pattern (*yištaġil*).

3.6.9 Form IX

Form IX does not occur in Khartoum Arabic. Its functions – getting a disease/defect and becoming a color – are sometimes expressed by Form II verbs in Khartoum Arabic (see Sec. 3.6.2 above).

3.6.10 Form X

Form X is produced by adding an *ista-* prefix to the root. Form X gives the following main senses:

estimative

istaġrab	'to find strange/unusual, be surprised at'
istahwan	'to treat (someone) as unimportant, look down upon'

reflexive

istaᶜadda	'to get (oneself) ready for'
istaḥamma	'to take/give oneself a shower'

The imperfect of Form X is on the yistaCCaC pattern: *istafraġ* 'he vomited', *yistafraġ* 'he vomits'. Where a verb is a recent borrowing from Standard Arabic, the yastaCCiC form is sometimes found: *istawrad* 'he imported', *yastawrid* 'he imports'.

3.6.11 Quadriliteral verbs

Quadriliteral forms may be formed in Khartoum Arabic in a number of ways.

Quadriradicals

A few verbs are constructed from four root letters. Examples are:

hadrab	'to talk deliriously'
šalwaṭ	'to scald'

Twin-radical reduplicatives

The twin-radical reduplicative morpheme results in a quadriliteral stem on the form C_1-C_2-C_1-C_2 and occurs in the following contexts:

With biradicals (roots not attested in non-reduplicated form)

katkat	'to tremble, shiver'
maḍmaḍ	'to rinse out the mouth'

With doubled roots

jaṛjaṛ	'to pull backward and forward' (root *j-ṛ-ṛ*)
laflaf	'to wrap round and round, go round [and round]' (root *l-f-f*)

With final-weak roots

Table 9. Conjugation of the perfect *hadrab* 'to talk deliriously'

3rd sg. masc.	*hadrab*	2nd sg. masc.	*hadrabta*	1st sg.	*hadrabta*
3rd sg. fem.	*hadrabat*	2nd sg. fem.	*hadrabti*		
3rd pl. masc.	*hadrabu*	2nd pl. masc.	*hadrabtu*	1st pl.	*hadrabna*
[3rd pl. fem.	*hadraban*]	[2nd pl. fem.	*hadrabtan*]		

daldal	'to let hang down' (root *d-l-y* 'to let down')
lawlaw	'to twist round and round [of a rope]' (root *l-w-y* 'to twist')

With medial-weak roots

gargar	'to hollow out' (root *g-w-r* 'to hollow')

With sound triradicals having identical C_1 and C_3

galgal	'to disturb, not to let rest' (root *g-l-g* 'to disturb')

Twin-radical reduplication typically gives a sense of intensiveness, repetition of the action, and/or distributed action.

Single-radical reduplicatives

Single-radical reduplicatives occur only with sound and medial weak verbs and involve repetition of the initial root letter in post-C_2 position.

gargaš	'to eat bread [or similar] without broth [hence: 'to make a crunching sound']' (root *g-r-š* 'to crush, crunch (up)')
karkaʿ	'to drink with a gulping sound' (root *k-r-ʿ*)
lōlaḥ	'to wag [tail], move [of leaves, and similar]' (root *l-w-ḥ*)
ṭōṭaḥ	'to swing, sway' (root *ṭ-w-ḥ*)

Single-radical root reduplicatives share with twin-radical root reduplicatives the sense of repeated action, but not so strongly the sense of intensive or distributed action.

Pre-C_2 /ō/ infix

The pre-C_2 /ō/ infix is derived from the coalescence of an original infixed /w/ and the preceding /a/ of the perfect tense. Pre-C_2 /ō/ infix forms most commonly denote:

Noises (especially repeated noises)

kōrak	'to shout'
nōna	'to hum, buzz'
lōla	'to lull to sleep, sing a lullaby'

Other repeated action

gōlab	'to turn [liquid food] over [in pan]'
hōzaz	'to move slowly to and fro [trans.]'

Other

ṣōban	'to wash with soap' (*ṣābūn* 'soap')

Pre-C_2 /r/, /n/, and /l/ infixes

These typically add an intensive or repetitive sense to that of the root.

šarbak	'to complicate, ensnare, tangle' (root *š-b-k*)
ḥankal	'to trip (someone) up' (root *ḥ-k-l; ḥakal* – same meaning)
falṭaḥ	'to broaden' (root *f-ṭ-ḥ*)

Post-C_2 /b/, /m/, /w/ infix

This typically gives a repetitive or intensive sense:

xarbaš	'to scratch [skin]' (also Form I, *xaraš*)
ṭarbag	'to knock' (root' *ṭ-r-g*)
šarmaṭ	'to become a prostitute; to give (someone) over to prostitution; to make dry meat [*šarmūṭ*]' (root *š-r-ṭ* 'to slit')
kajwal	(also *kajal*) 'to cause (someone) to walk so that his legs obstruct one another [of paralyzing disease, etc.]'

Table 10. Conjugation of the imperfect *yihadrib* 'he talks deliriously'

3rd sg. masc.	*yihadrib*	2nd sg. masc.	*tihadrib*	1st sg.	*ahadrib*
3rd sg. fem.	*tihadrib*	2nd sg. fem.	*tihadribi*		
3rd pl. masc.	*yihadribu*	2nd pl. masc.	*tihadribu*	1st pl.	*nihadrib*
[3rd pl. fem.	*yihadriban*]	[2nd pl. fem.	*tihadriban*]		

Table 11. Conjugation of the imperfect *yitlaflaf* 'he goes [round and round]'

3rd sg. masc.	*yitlaflaf*	2nd sg. masc.	*titlaflaf*	1st sg.	*atlaflaf*
3rd sg. fem.	*titlaflaf*	2nd sg. fem.	*titlaflafi*		
3rd pl. masc.	*yitlaflafu*	2nd pl. masc.	*titlaflafu*	1st pl.	*nitlaflaf*
[3rd pl. fem.	*yitlaflafan*]	[2nd pl. fem.	*titlaflafan*]		

Forms derived from more basic nouns:

With initial *m-:*

maġrab	‘to get to sunset’ (cf. *muġrib* ‘sunset’)
maglab	‘to play a trick on’ (cf. *maglab* ‘trick’)

With final *-n:*

galban	‘to change [subtly or deviously]’ (*galban almawḍūˁ* ‘to shift the subject’)

Quadriliteral verbs take *a-i* internal imperfect forms with an /i/- vowel in the prefix.

3.6.12 Form II quadriliteral verbs

Form II quadriliteral verbs, i.e. quadriliteral verbs with an *it-* prefix, are used to express passive, reflexive, or reciprocal meaning or acting/pretending.

Passive

The *it-* prefix can be used to passivize virtually all transitive verbs with an active (nonrelational) meaning.

itṣōban	‘to be washed’ (root *ṣ-b-n* with /ō/ infix*;* cf. *ṣābūn* ‘soap’; imperfect *yitṣōban*)

Reflexive

Reflexive uses of the *it-* prefix are also common. They shade into passive uses and also into uses where the translation suggests a notion of pure becoming.

itˁarban (*min*)	‘to get an advance payment (from)’ (root *ˁ-r-b-n*)
itḥalḥal (*min*)	‘to get free of’ (root *ḥ-l-(l)*)

Reciprocal

itnōna	‘to buzz around/together [of flies, etc.]’

Acting as/pretense

itfalham	‘to pretend to knowledge’ (+ triradical *f-h-m* + pre-R2 /l/)
itšaxsan	‘to show off, pretend to be a big personality’

Form II quadriliteral verbs take *a-a* internal imperfect forms with an /i/ vowel in the prefix.

3.6.13 The *t-* prefix with Form I and Form IV verbs

The *t-* prefix also occurs in derivations from Form I, giving the following senses:

Passive
from Form I

itlaˁab	‘to be played’
itfaham	‘to be understood’

from Form IV

itʾadda	‘to be given’ (*alkitāb da* (*i*)*tʾadda lēy*$^{\uparrow}$ ‘that book was given to me’)

Reflexive
from Form I

itzagga	‘to slip into [e.g. a queue]’ (root *z-g-g*; *zagga* ‘to slip something in [e.g. a paper/ name]’)
itgaṭaˁ (*min*)	‘to stop (coming to see)’ (*gaṭaˁ* ‘to stop [someone else]’)

Form I verbs with *t-* prefix take *a-i* internal imperfect forms: *itlaˁab* ‘it was played’, *yatlaˁib* ‘it is played’. Some speakers have an initial /i/ instead of /a/ (*yitlaˁib*).

Form IV verbs with *t-* prefix take *a-a* internal imperfect forms: *itʾadda* ‘it was given’, *yitʾadda* ‘it is given’.

3.7 *Tense markers*

The imperfect occurs in the following tense forms:

bare imperfect
bi- + imperfect
gāˁid + imperfect
ḥa- + imperfect

3.7.1 Bare imperfect

The bare imperfect is used:

i. to express commands in the 1st and 3rd persons, and tentative commands in the 2nd person:

mā fi zōl yamši maˁāy$^{\uparrow}$	‘nobody is to go with me’
tadfaˁ baˁdēn	‘you must/should pay later’

ii. to express purpose in subordinate structures:

ana jīt ašūf šuġul	‘I came to find work’

iii. in other subordinate structures, most obviously after modals:

ana ʿāwiz naštaġil	'I want us to work'
ma btagdar tašūf	'she isn't able to see'

In many contexts, the bare imperfect has the same sense as the *bi-* + imperfect or *gāʿid* + imperfect.

3.7.2 *bi-* + imperfect

bi + imperfect expresses:

continuous present

bitsawwi šinu	'what are you doing?'

general present

bagūm assāʿa sitta	'I get up at six o'clock'
aššuġul mā bintahi	'work never finishes'

future

bajīkum bukra	'I'll come to [see] you [pl. masc.] tomorrow'

3.7.3 *gāʿid* + imperfect

gāʿid is the active participle form of *gaʿad* 'to sit; to remain, last', and has the forms *gāʿda* (sg. fem.), *gāʿdīn* (pl. masc.), and *gāʿdāt* (pl. fem.).

gāʿid expresses most basically:

continuous present

gāʿid tasawwi šinu	'what are you doing?'

general present

gāʿid agūm assāʿa sitta	'I get up at six o'clock'

gāʿid + imperfect is more emphatic than *bi-* + imperfect, giving greater focus and a greater sense of voluntariness to the action.

3.7.4 *ḥa-* + imperfect

ḥa- + imperfect seems to be a 20th-century borrowing from Egyptian that is now well established in Khartoum Arabic. It basically expresses future:

ḥaʾajīkum bukra	'I'll come to [see] you tomorrow'

4. SYNTAX

4.1 *Word order*

The interaction between subject and predicate and theme and rheme provides insights into the word order of most clauses that contain a verb phrase.

S-V(-O)

Subject-Verb(-Object) is the most common word order for clauses involving a verb, and is understood as follows. Subjects are typically definite and therefore thematic. Verb phrases are indefinite and therefore rhematic. Subjects thus come first in the clause and verb phrases last, since themes typically precede rhemes. Objects come after verbs, in line with the general tendency in Khartoum Arabic for the head (here the verb) to precede the modifier (here the object).

V(-O)-S

Verb(-Object)-Subject word order occurs either where the subject is indefinite and therefore rhematic (as in *gāmat nār* 'fire broke out'), or where the verb phrase is an initial rheme (as in *fihm almuškila arrājil da* 'that man understood the problem'). As noted above, the V-O word order embodies the general tendency toward head modifier sequencing.

Other word orders are also possible. Relatively infrequently one finds V-S-O word order, and somewhat more commonly the essentially syntactically identical V-S-C(omplement) word order with verbs such as *kān* 'to be' and *biga* 'to become':

šuft(a) inta azzōl da?	'have you seen that man?'
biga zzōl da diktōr	'that man became a doctor'

4.2 *The definite particle*

The definite article – or better, given its wide range of usage, the 'definite particle' – has the canonical form *al-*. The /l/ assimilates to the following letter when this is apico-dental, apico-alveolar, or dorso-prepalatal (cf. Table 1), and lacks initial /a/ following a preceding vowel.

The definite particle occurs before various elements, e.g. prenominally, *arrājil* 'the man', prejadjectivally *azzaʿlān* 'the angry one', preverbally *azziʿil* 'the one who got angry' [i.e. 'the-(he-)got-angry (one)'], preadverbially *alhassi* 'the one/ones who is/was [here] now', and before subject-predicate structures, e.g. *alana ziʿilta minnu* 'the [thing/person] I got angry with'.

Prenominally, the definite particle is co-referential with its following noun. Thus, *arrājil* 'the man' is both 'man' and definite (in respect of being [a] man). Preadjectivally, the definite particle is normally co-referential with the following adjective. However, if the adjective is modified by a subsequent prepositional phrase, the definite particle may be co-referential with some other element of the phrase. Thus, *[arrājil] azzaʿlān minnu* may mean either '[the man] who is angry with him' or '[the man] with whom he is angry'. The definite particle may be co-referential with a following verb (i.e. Verb-Subject), or with a verb-dependent element: *aššāfáta* can mean either 'the one [sg. fem.] whom she saw' or 'the one [sg. fem.] who saw her' (in the latter interpretation *aš-* [= *al-*] is co-referential with the Verb-Object 'her'). Correspondingly, with prepositions: *algiddāma* may mean either 'the one who is/the ones who are in front of her' or 'the one (fem.) whom he (etc.) is in front of'.

al- phrases (except those with a following noun) typically function as attributive adjectivals to preceding head nouns (e.g. *arrājil azziʿilta minnu* 'the man whom I got angry with'). However, they quite frequently occur without a head noun, and in this case may, like nouns, function as predicands, predicates, verb-objects, and preposition-objects; thus *algāl kida mnu*↑ 'who said that' (lit. '[the one who] said thus [was] who').

4.3 *Construct state: Types*

Khartoum Arabic has both the construct (synthetic genitive) and an analytical genitive with *bitāʿ* or *ḥagg*. These agree with the preceding noun in number and gender:

	masculine	feminine
singular	*bitāʿ, ḥagg*	*bitāʿ(a)t, ḥaggat*
plural	*bitāʿīn, ḥaggīn*	*bitāʿāt, ḥaggāt*

zōl bitāʿ/ḥagg mašākil	'a man of problems/ a problematic person'
alkutub bitāʿt/ḥaggat almadrasa	'the books of the school' (sg. fem. agreement with inanimate pl. noun)
nās bitāʿīn/ḥaggīn kalām katīr	'talkative people [people of much talking]'
attazākir ḥaggātu	'his tickets'

The internal structure of *bitāʿ/ḥagg* phrases is itself a construct, as suggested by the /t/ of the feminine singular *bitaʿ(a)t/ḥaggat*. Accordingly, *bitāʿ/ḥagg* phrases in which the following noun is definite are standardly definite (e.g. *bitāʿt/ḥaggat almadrasa*), while those in which the following noun is indefinite are indefinite (e.g. *bitāʿ/ḥagg mašākil*). *Bitāʿ/ḥagg* phrases are adjectival in nature; following a definite noun, the *bitāʿ/ḥagg* phrase is normally definite, and following an indefinite noun, the *bitāʿ/ḥagg* phrase is indefinite. Like indefinite adjectives, indefinite *bitāʿ/ḥagg* phrases cannot occur as objects of verbs or prepositions:

**bitaʿrif bitāʿ/ḥagg mašākil*	'do you know one who is problematic?'
**ja maʿa bitāʿ/ḥagg kalām katīr*	'he came with a talkative one/person'

Like definite adjectives, definite *ḥagg/bitāʿ* phrases can occur as objects of verbs or prepositions, typically with a following demonstrative:

bitaʿrif bitāʿ/ḥagg almašākil da
'do you know the problematic person?'
ja maʿa bitāʿ/ḥagg alkalām alkatīr da
'he came with a talkative one/person'

In most cases, the synthetic genitive and *bitāʿ/ḥagg* phrases are grammatically interchangeable. Inalienable possession, however, is expressed only through the synthetic genitive:

rijli↑	'my foot/leg' (not **arrijil bitāʿti*↑*/ ḥaggati*↑)

4.4 *Negation*

The normal negator is *ma*. This standardly occurs before predicates, e.g.

inta ma rājil/kwēyis/ fi lbēt	'you're not a man [prenominal]/nice [preadjectival]/in the house [prepositional]'
[inta] ma btafham ʿarabi	'you don't understand Arabic [preverbal]'

The distribution of *ma-* phrases is fairly similar to that of (indefinite) adjectives (e.g., both can occur as predicates, but not as objects of verbs or prepositions). The only common phrase involving *ma* with a suffixed negative *-š* is the adjectival *ma fīš* (sg. fem. *ma fīša*, pl. masc. *ma fīšīn*, pl. fem. *ma fīšāt* 'not there, absent').

BIBLIOGRAPHICAL REFERENCES

Bergman, Elizabeth. 2002. *Spoken Sudanese Arabic: Grammar, dialogues, and glossary*. Springfield, Va.: Dunwoody Press.

Blair, Philip. 1983. *The segmental phonemes of Sudanese Arabic*. Ph.D. diss., University of Exeter.

Bohas, Georges, Jean-Patrick Guillaume, and Djamel Eddine Kouloughli. 1990. *The Arabic linguistic tradition*. London: Kegan Paul International.

Hamid, Abdel Halim M. 1984. *A descriptive analysis of Sudanese Colloquial Arabic phonology*. Ph.D. diss., University of Illinois.

Hillelson, S. 1930. *Sudan Arabic: An English-Arabic vocabulary*. London: Sudan Government.

—— 1935. *Sudan Arabic texts*. Cambridge: Cambridge University Press.

Jakobson, Roman. 1957. "*Mufaxxama*: The 'emphatic' phonemes in Arabic". *Studies presented to Joshua Whatmough*, ed. Ernst Pulgram, 105–115. The Hague: Mouton.

Kaye, Alan S. (1976). *Chadian and Sudanese Arabic in the light of comparative dialectology*. The Hague: Mouton.

Mustapha, Abdel Rahman. 1982. *La phonologie de l'arabe soudanais (phonématique et accentuation*, 2 vols.) Ph.D. diss., Université de la Sorbonne Nouvelle, Paris.

Persson, Andrew and Janet Persson. 1980. *Sudanese Colloquial Arabic for beginners*. High Wycombe, England: Summer Institute of Linguistics.

Prokosch, Erich. 1983. *Osmanisches Wortgut im Sudan-Arabischen*. Berlin: K. Schwarz.

Qāsim, ʿAwn aš-Šarīf. 2002. *Qāmūs al-lahja al-ʿāmmiyya fī s-Sūdān*. 3rd ed. Khartoum: ad-Dār as-Sūdāniyya li-l-Kutub.

Trimingham, J. Spencer. 1946. *Sudanese Colloquial Arabic*. London and Oxford: Oxford University Press.

Watson, Janet C.E. 1993. *A syntax of Ṣanʿānī Arabic*. Wiesbaden: O. Harrasowitz.

Worsely, Allan. 1925. *Sudanese grammar*. London: Society for the Promotion of Christian Knowledge.

Yagi, Viviane A. 1981. *Contes d'Omdurman*. Paris: Publisud.

JAMES DICKINS
(University of Salford)

Khuzestan Arabic

1. GENERAL

The dialect of Khuzestan is of the Southern Mesopotamian or *gilit* type, to follow the terminology introduced by Haim Blanc. Specifically, it is similar to the dialect of the adjoining areas of Iraq, namely south of Kūt on the Tigris and Nāṣiriyyah on the Euphrates.

1.1 *Area*

Khuzestan is the area of southwestern Persia or Iran to the east of the Šaṭṭ al-ʿArab and generally up to the fringes of the Zagros Mountains. This area has been Arabic speaking from approximately the 17th century. The area of Khuzestan is named after the *Xūz*, the original inhabitants of the area at the time of the Arab invasions, and was similarly named by the Arabs Sūq al-ʾAhwāz 'market of the *Xūz*'. ʾAhwāz is the modern pronunciation, although in the earlier period it was often written ʾAḥwāz.

1.2 *Position*

The social organization of the population is by tribe, by far the largest in number and in area being the Kaʿb. Along the Šaṭṭ are the Muḥaysin, and further north around ʾAhwāz the Bani Ṭuruf, Bāwiyah, and Bani Lām are found, the latter spreading also into Iraq north of the Hōr al-Ḥuwayzah. All of these speak Arabic as their first language, but many could also speak Persian to some degree.

During the time of the shah the regional, non-Iranian languages of Iran were definitely suppressed, the idea being propagated that such languages were not languages but rather dialects, 'dialect' denoting a purely local form of speech of low prestige without a written literature.

At that time Persian was the official language and the main language of commerce in the towns. Outside the towns, however, the language was Arabic. Arab dress was worn in the distinctive North Gulf manner, usually, in summer, a light *ṯōb*, with a white head cloth called there *čaffiyyah*, without the head rope and with the long ends thrown back over the head. The weather was too hot and humid even for the very light Iraqi summer *bišt*, so it was carried neatly folded over one shoulder.

1.2.1 Dialect distribution

Three main dialects can be discerned. These can be described as *ḥaḍar*, *ʿarab*, and marshland dialects. The dialect distribution represents a continuum with that of neighboring southern Iraq as described below. The linguistic areas that can be perceived, again with a fair degree of gross generalization, are the following:

i. The *ḥaḍar* dialect along the Šaṭṭ al-ʿArab and Lower Kārūn and parts of the Euphrates

below Nāṣiriyyah and the banks of the Jarrāḥi around Šādigān, including the main old towns of the region, i.e. Basra and Khorramshahr.
ii. *The ʿarab* dialect in the *bādəyah* or open country between the rivers, embracing also some of the newer towns, such as ʾAhwāz.
iii. The marshlands dialect around the Hōr al-Ḥuwayzah and around ʿAmārah in Iraq and spreading into the Hōr al-Ḥammār.

The above conforms to a fairly well observed fact about linguistic geography, namely that rivers and water networks form the centers of linguistic regions, rather than dividing them. Thus, it would be unnatural to find that the Šaṭṭ al-ʿArab constituted a linguistic boundary.

1.3 *History*

The earliest firsthand account we have of the Arab population of Khuzestan is from the Portuguese Jew Pedro Teixera, who traveled up the Šaṭṭ al-ʿArab in 1604. At that time, he reported that all of the country to the east of the Šaṭṭ was ruled by Mubārak ibn Muṭlub, "an Arab chief who maintained a claim to Basrah and was perpetually at war with the Turks" (Lorimer 1915:iv, 2:1625). Mubārak was, according to *The Gazetteer*, one of the wālīs of Ḥuwayzah. These were descended from a family of Sayyids from Mecca who were chiefs of Wāsiṭ in southern Iraq and during the Safavid period extended their rule east to Ḥuwayzah in Khuzestan. This ushered in an era of Arab domination of the area, uninterrupted until the first half of the 20th century. The area was previously universally known as ʿArabistān, both by Arabs and Persians, until it was finally taken over again by Reza Shāh, when its name was officially changed to Khuzestan (Xūzistān). At the end of the 17th century the Arab tribe of the Bani Kaʿb (often referred to as the Chaub) moved into Khuzestan and took over the area from the Afshar Turks, becoming clients of the wālī of Ḥuwayzah and later themselves becoming rulers of the area. The Kaʿb seem to have been independent of the rule of Basra at most times and had settlements initially at Gubān at the head of Khōr Mūsa until 1747 and then later at Šādigān, or Fallāḥiyyah, to give it its Arabic name. The later rule of the Muḥaysin under Šayx Xazʿal in the late 19th and early 20th centuries moved the seat of power to Muḥammarah (Khorramshahr) on the lower Kārūn. Šayx Xazʿal built a flourishing trading economy at Muḥammarah and was a rival power to the neighboring ʾĀl Ṣubāḥ of Kuwait, to whom even the Bedouin of the Kuwait hinterland would come for help with local problems.

1.4 *State of research*

Available data on the use of Arabic in Khuzestan date from the mid-1970s before the upheavals resulting from the revolution and the Iran-Iraq war that devastated much of the area and caused much of the population to emigrate to neighboring cities, particularly Shiraz.

The main sources on the dialect are Ingham (1973, 1976, 1982) and D. Lorimer (n.d.) However, the dialect resembles in many ways the *gilit* dialects of Baghdad, which are more extensively described (→ Baghdad Arabic) than the dialect of Khuzestan. Useful historical background can be gleaned from Layard (1846), Loftus (1857), J. Lorimer (1915) and Stocqueler (1832).

2. LINGUISTIC DESCRIPTION

2.1 *Phonology*

The consonant inventory is conservative in having preserved the interdentals in such forms as *ṯələṯ* 'three' and *ḏāk* 'that'. It also shows the development of the fronted variants of velar /k/ and /g/ in fronting environments as /č/ and /j/, familiar in the area, as in *čaləb* 'dog' and *jəddām* 'before'. As with the Gulf litoral and southern Iraq, it regularly shows /y/ for /j/ as in *yāb* 'he brought', *yā* 'he came', *fəyəl* 'radish'. The marshland dialect here shows /ž/ as in *žīt* 'I came', *žība* 'bring it [sg. masc.].' The consonant /ʾ/ has disappeared except in borrowings from Classical Arabic, generally being elided or replaced by /y/ as in *gāyəl* 'having said' (Classical Arabic *qāʾil*), the verb *saʾal* 'to ask' being replaced by *nəšad*. The vowels follow the general Arabic pattern except that non-final /i/ and /u/ have merged as one phoneme, /ə/.

Initial consonant clusters can stand in the dialect, and new ones are also formed by the elision of short vowels. These can also show an anaptyctic vowel before the consonant cluster.

Thus, we find *əktāb* ~ *ktāb* ‘book’, *ṯnēn* ~ *əṯnēn* ‘two’.

Notice also that the *ʿarab* dialect shows words of the form CCvC- as in *ktəbat* ‘she wrote’, *xšəba* ‘piece of wood’, *nxala* ‘palm tree’, *ṣxala* ‘goat’, where the *ḥaḏ̣ar* dialect would show *kətbat*, *xəšba*, *naxla*, and *ṣaxla*. In both dialects the effect of the guttural group also produces elatives of the form CaCa(C) such as *xaḏ̣ar* ‘green’, *ʿaray* ‘lame’, *ʿaway* ‘crooked’, *ḥamar* ‘red’, *ʿama* ‘blind’ (→ *gahawa*-syndrome).

2.1.1 Consonants

Plosives: b, t, d, ṭ, k, g, q, ʾ
Affricates: č, j
Fricatives: f, ṯ, ḏ, ḏ̣, x, ġ, ḥ, ʿ, h
Sibilants: s, ṣ, z, š
Laterals and vibrants: l, r
Nasals: m, n
Semivowels/glides: w, y

/j/ is realized as a voiced palato-alveolar fricative [ʒ] in the marshland dialect, and /q/ mainly occurs in borrowings from Classical Arabic as an equivalent to /g/. /ṭ/ may be pronounced as pharyngealized voiceless or voiced dental plosive. /f/ is a voiceless labiodental fricative often pronounced rather laxly and sometimes voiced. /ġ/ is a voiced uvular fricative sometimes pronounced plosive [q] when initial. There is a degree of merger between Classical Arabic /ġ/ and /q/, yielding, for instance, *ġaṣər* (< *qaṣr*) ‘castle’, *ġalam* (< *qalam*) ‘pen’, *gālab* (< *qālab*) ‘mould’, *qēr* (< *ġayr*) ‘other than’. Note that the dialect has *y* for Classical Arabic *jīm* in all local words, as in *yəbal* ‘mountain’, *əyyi* ‘he comes’, *yəfal* ‘it shied’, *ʿayīn* ‘dough’, *ʿayūz* ‘old woman’, *ʿaray* ‘lame’.

2.1.2 Vowels

Short vowels: ə, a
Long vowels: ī, ū, ē, ō, ā

ə This vowel shows back rounded realizations [ʊ] in bilabial and velar environments and front spread realizations [ɪ] in dental and palatal or palato-alveolar environments, while becoming very indistinct and centralized with a shwa-like quality [ə] in neutral environments such as in the neighborhood of /r, h, ḥ, ʿ, x, ġ/.

a This vowel is generally central in quality, but it does show fronting to a sort of [ɛ] sound in front environments. In the environment of the pharyngeals /ḥ/ and /ʿ/, it has a rather more open position.

ē The typical realization in most front environments is a glide from lax half-close to mid central [ie], as in *zēn* [zien] ‘good’. In more neutral environments it has a pure vowel realization, as in *hēl* [he:l] ‘cardoman’, *ḥēl* [ħe:l] ‘very much’.

ō This vowel is generally a long half-close rounded vowel, as in *fōg* [fo:g] ‘above’, *yōm* [jo:m] ‘day’, *bōg* [bo:g] ‘stealth, trickery, dishonesty’.

ā The usual realization in all environments is a fully open, longish mid vowel.

Long /ī/ and /ū/ are shortened in word-final unstressed position; they are written here *-i, -u*.

2.1.3 Syllable structure

The dialect shows non-final syllables Cv, Cvv, CvC, and CvvC and final Cv, CvC, and CvvC. Examples are Cv- *nəšad* ‘he asked’, Cvv *šāyəl* ‘carrying [pl. masc]’. Non-final CvvC occurs only in active participle forms such as *šāyla* ‘carrying [sg. fem.]’ and *šāylīn* ‘carrying [pl. masc.]’. Final Cv occurs in *mara* ‘woman’, final CvC in *nəšad* ‘he asked’, and final CvvC in *ktāb* ‘book’. Note that underlying final CvCC does not occur, and an anaptyctic vowel *ə* separates the final cluster, which shows back rounded realizations [ʊ] in bilabial and velar environments and front spread realizations [ɪ] in dental and palatal or palato-alveolar environments, as in *čaləb* [tʃalɪb] ‘dog’, *gaḷəb* [gaɫʊb] ‘heart’. Note also that a final 3rd singular masculine suffix *-h* is often elided in rapid speech, resulting in final Cvv, as in *bī* from *bīh* ‘in it’, *šəfnā* from *šəfnāh* ‘we saw him’, etc.

Note that the *ʿarab* dialect shows verbal, nominal, and adjectival structures of the type CCvC-v as in *nšədat* ‘she asked’, *nxalah* ‘palm tree’, *ʿšara* ‘pregnant [of livestock]’, *bġadād* ‘Baghdad’, *nʿasat* ‘she dozed’.

2.2 *Morphology*

The dialect is in general progressive in its morphology, and particularly in the verb, merging verbal classes and leveling distinctions.

2.2.1 Pronouns and similar elements

2.2.1.1 Personal pronouns

Personal pronouns are unremarkable, though showing some idiosyncrasy in the 3rd person (Table 1).

Table 1. Personal pronouns in Khuzestan Arabic

	3rd pers.	2nd pers.	1st pers.
sg. masc.	*əhwa*	*ənta*	*āna, āni*
sg. fem.	*əhya*	*ənti*	
pl. masc.	*əhma*	*əntəm*	*əḥna*
pl. fem.	*əhna*	*əntan*	

2.2.1.2 Object pronoun suffixes

The forms of the suffixes are as follows: *-ni* 'me', *-(a)k* 'you [sg. masc.],' *-(ə)č* 'you [sg. fem.]', *-a(h)* 'him', *-ha* 'her', *-na* 'us', *-kəm* 'you [pl. masc.]', *-čan* 'you [pl. fem.]', *-həm* 'they [pl. masc.]', *-hən* 'they [pl. fem.]'.

The suffixes *-an, -at*, and *-tan* double the final consonant when followed by a vowel-initial object suffix, in *kətbatta* 'she wrote it [sg. masc.]', *kətbanna* 'they [pl. fem.] wrote it [sg. masc.]', *kətabtanna* 'you [pl. fem.] wrote it [sg. masc.]', while the vowel-final suffixes *-aw, -na, -ti*, and *-tu* lengthen the vowel, *-aw* showing *-ō*, as in *kətbōha* 'they [pl. masc.] wrote it [sg. fem.]', *kətabnāha* 'we wrote it [sg. fem.]', etc.

In the imperfect, where a vowel-initial object suffix follows stems without the number and gender suffixes, resyllabication occurs, as in *anəšdak* 'I ask you [sg. masc.]', *annəšdəč* 'we ask you [sg. fem.]', *ətkətba* 'you write it [sg.masc.]', *əykətla* 'he hits him', *nəšda* 'ask him!'

With the preposition *lə-* 'to, for', a particle *-əyya-* can occur as the bearer for the direct object pronoun, as in *gəltəlhəm-əyyā* 'I said it to them', *yābatənna-yyāha* 'she brought it [sg. fem.] for us'.

2.2.1.3 Demonstrative pronouns

The demonstrative pronouns distinguish near and far, singular and plural, and masculine and feminine. They also show variant forms in some cases, with or without final short vowels and with or without the element *hā-* (Table 2).

Table 2. Demonstrative pronouns in Khuzestan Arabic

		masculine	feminine
'this'	singular	*hāḏa, hāḏ*	*hāḏi, hāy*
	plural	*hāḏōla, hāḏōl* *ḏōla, ḏōl*	*hāḏanni, ḏanni*
'that'	singular	*ḏāka, ḏāk*	*hāḏīč, ḏīč*
	plural	*ḏōlāk, ḏōlāka*	*hāḏannīč, ḏannīč*

2.2.1.4 Interrogative pronouns

WH- questions show a specific type of intonation with high-level followed by low-level tone, the prominence coming early in the word, very often on the WH- element. The interrogative pronouns are the following:

'what?'	*sənu, š-*
'who?'	*yāhu, yāhaw, -man*
'which?	*yāhu, yāhaw, -man, yā-*
'why?'	*lēš, lawēš, ʿalēš*
'when?'	*yamta*
'whither?'	*ləwēn*
'where?'	*wēn*
'how'?	*šlōn*

The preposed element *yā* 'which?' occurs in combinations such as *yā walad* 'which boy?', *yā ṣōb* 'which direction?'. The forms *š-* 'what?' and *-man* 'who?' 'which?' occur, signaling the object of verbs and prepositions and the possessors of nouns. The stress on these combinations usually comes on the first syllable of the verb or noun or preposition as in *š-šawwēt* 'what did you do?', *š-gāl* 'what did he say?', *šəfət-man* 'who, which did you see?', *trīd-man* 'which, who do you want?', *ktāb-man* 'whose book?', *ʿalē-man* 'on, about whom, what?', *əl-man* 'belonging to whom; what for', *məm-man* 'from whom; what?', *ḥadər-man* 'underneath whom?', *māl-man* 'belonging to whom?'. Examples in context include *məmman əmsawwāya* 'what is it made of?', *lənsān ʿalēman əyʿīš* 'what does a man live on?', *ʿalēman əya* 'what did he come for?'.

The independent WH- elements usually occur initially in the sentence as in *yāhaw hāḏ* 'who is this?', *yāhaw ya* 'who has come?', *wēn rāyəḥ* 'where (are you) going?', *šlōn sawwēta* 'how did you do it?', *lēš mā yēt* 'why didn't you come?', *yamta təyi* 'when will you come?'.

2.2.2 Particles

Syntactic and grammatical elements, usually undeclinable, peculiar to the dialect occur both in the nominal and the verbal phrase and include the following:

i. *fad* → indefinite article: *fad walad* 'a boy', *fad waladēn* 'some two boys', *fad awlād* 'a few boys'.
ii. *māl* 'belonging to', general genitive particle also occurring in the forms *mālt-*, *mālīn,* and *mālāt: əlbēt māli* 'my house', *əlbēt māl əlwalad* 'the boy's house'
iii. *aku* 'there is' existential particle: *aku walad bəlbēt* 'there is a boy in the house'
iv. *hast, hassət* 'there is' existential particle: *māy hassət* 'there is water'
v. *ham, hammēna* 'also'

2.2.3 Noun

Nominal morphology does not differ from the general Arabic pattern except for the junction of nouns with possessive pronoun suffixes. In particular the anaptyctic vowel occurring between the two elements of a final cluster causes the following alternations: *čaləb* 'dog', *čalba* 'his dog', *čaləbha* 'her dog'; *galəb* 'heart', *galba* 'his heart', *galəbha* 'her heart'.

Nouns of the form CiCCa, when followed by vowel-initial possessive suffixes, resyllabify as follows: *ḥəjra* 'room', *ḥəjərti* 'my room', *ḥəjərta* 'his room', *ḥəjərtak* 'your [sg. masc.] room', *ḥəjərtəč* 'your [sg. fem.] room'.

Nouns of the form CaCaC delete the second *-a-* when a vowel follows the form, as in *balam* 'boat', *balmi* 'my boat'; *ġalam* 'pen', *ġalma* 'his pen'.

2.2.4 Verbs

The verb has leveled the distinction between the Classical Arabic transitive CaCaC and intransitive CaCiC or CaCuC types, having only CiCaC, representing the Classical Arabic CaCaC, although also showing CaCaC in the *ʿarab* dialect in guttural environments (see below 2.2.4.1).

2.2.4.1 Form I

The basic paradigm of the strong verb is given in Tables 3, 4, and 5.

Table 3. Perfect of the strong verb in Khuzestan Arabic

	3rd pers.	2nd pers.	1st pers.
sg. masc.	*kətab*	*kətabət*	*kətabət*
sg. fem.	*kətbat*	*kətabti*	
pl. masc.	*kətbaw*	*kətabtu*	*kətabna*
pl. fem.	*kətban*	*kətabtan*	

Note that in the *ʿarab* dialects the forms *ktəbat* 'she wrote', *ktəbaw* 'they [masc.] wrote' and *ktəban* 'they [pl. fem.] wrote' occur. Also, if C_1 or C_2 is one of the guttural group /h, ḥ, ʿ, x, ġ/ or if C_2 is one of the liquids /l, n, r/, the vowel of the first syllable of the stem is /a/ in accordance with the vowel-raising rule, as in the following examples: *ḥalaf* 'to swear', *xalaṭ* 'to mix', *ġalab* 'to conquer', *hamaz* 'to massage', *ʿabar* 'to cross', *zaʿal* 'to be angry', *naʿas* 'to doze', *saḥag* 'to crush', *naha* 'to forbid', *ṭalaʿ* 'to emerge', *šarab* 'to drink', *bana* 'to build'.

Table 4. Imperfect of the strong verb in Khuzestan Arabic

	3rd pers.	2nd pers.	1st pers.
sg. masc.	*yəktəb*	*təktəb*	*aktəb*
sg. fem.	*təktəb*	*ətkətbīn*	
pl. masc.	*ykətbūn*	*ətkətbūn*	*nəktəb*
pl. fem.	*ykətban*	*ətkətban*	

Note also the form *akətban* for 'I write' occurring optionally before object pronoun suffixes, as in *akətbanna* 'I write it [sg. masc.]'. The imperfect of verbs with one of the guttural group as C_1 shows resyllabication in the *ʿarab* dialect, as in *yḥaləf* 'he swears', *yxaləṭ* 'he mixes', *aġaləṭ* 'I make a mistake', *yġaləb* 'he defeats', *yharəb* 'he flees'.

Table 5. Imperative of the strong verb in Khuzestan Arabic

əktəb	'write [sg. masc.]!'
kətbi	'write [sg. fem.]!'
kətbu	' write [pl. masc.]!'
kətban	'write [pl. fem.]!'

Note the forms *əkətbi* 'write [sg. fem.]!' *əkətbu* 'write [pl. masc.]!', and *əkətban* 'write [pl.

fem.]!’, occurring as alternatives and as the most common forms in the ʿ*arab* dialect.

2.2.4.2 Internal passive

Two forms of the internal passive were recorded in narratives, namely *ġədər* ‘he was deceived’ and *čətəl* ‘he was killed’. These were recorded from speakers of the ʿ*arab* dialect, Kawāwila or Gypsies in the area of ʾAhwāz, who were professional storytellers and entertainers. Such forms were not recorded in informal speech. In general, Form VII *ənfəʿal* acts as the passive and is totally productive.

2.2.4.3 Derived forms

The dialect includes all the usual forms with the exception of Forms IV *afʿal* and IX *əfʿall.* These do not occur in the *ḥaḏ̣ar* dialect, but the former does occur in the ʿ*arab* dialect generally as verbs of movement, as in *adbar* ‘to go away’, *agfa* ‘to go away’, *ašmal* ‘to go north’, *agbal* ‘to approach’. Otherwise, the causative function is served by Form II. Note also that Form VII is the regular passive in this dialect. Examples of the regularly occurring forms are the following:

II	*laggaṭ*	‘to pick things up’ (repetitive)
	kassar	‘to shatter’ (intensive)
	gawwam	‘to raise’ (causative)
III	*ḥārab*	‘to fight with’ (comitative)
V	*tnaššad*	‘to inquire about’ (repetitive)
	tmašša	‘to stroll, walk’ (continuously)
VI	*tḥārab*	‘to fight together’ (reciprocal)
VII	*ənnəšad*	‘to be asked’ (passive)
VIII	*əštəra*	‘to buy’ (derived transitive)
	əftəham	‘to understand’ (derived transitive)
X	*əstafham*	‘to inquire about’ (reversed transitive)

New forms include CōCaC *sōlaf* ‘to talk’, tCōCaC *tsōlaf* ‘to talk, chat’, and tCēCaC *tnēšan* ‘to take aim’. These forms have no specific function. The derivation of *tsōlaf* is from the plural form *swāləf* ‘stories, talk’, and *tnēšan* is from the Persian loanword *nēšān* ‘aim’.

2.2.4.4 Geminate verbs

The geminate verb, as in most spoken dialects, introduces a vowel /ē/ before consonant-initial subject suffixes in the perfect, as in *raddēt* I/you [sg. masc.] answered’, *raddēti* ‘you [sg. fem.] answered’, *raddētu* ‘you [pl. masc.] answered’, *raddētan* ‘you [pl. fem.] answered’, *raddēna* ‘we answered’, in contrast to *radd* ‘he answered’, *raddat* ‘she answered’, *raddaw* ‘they [pl. masc.] answered’, *raddan* ‘they [pl. fem.] answered’. The 3rd person singular masculine perfect may also show a formative /ā/ before object pronoun suffixes, as in *šaddāha* ‘he tied it [sg. fem.]’.

The imperfect forms conform to the general Arabic pattern, showing *arədd* ‘I answer’, *əyrədd* ‘he answers’, *ətrədd* ‘she answers, you [pl. masc.] answer’, *ənrədd* (> *ərrədd*) ‘we answer’, *əyrəddūn* ‘they [pl. masc.] answer’, *əyrəddan* ‘they [pl. fem.] answer’, *ətrəddūn* ‘you [pl. masc.] answer’, *ətrəddan* ‘you [pl. fem.] answer’. In the ʿ*arab* dialect, in unsuffixed forms, the doubled consonant can be reduced and initial stress can occur, as in *yərəd, tərəd, nərəd*, and *arəd.* When followed by a vowel-initial suffix, however, the double consonant is heard, as in *əyrəddah* ‘he returns it [masc.]’ The *-an* alternative of the 1st person singular regularly occurs, as in *arəddan* ‘I answer’, *ašəddan* ‘I tie’.

2.2.4.5 Weak verbs

2.2.4.5.1 Initial weak verbs

The weakness of these verbs is apparent in the imperfect and imperative, showing initial ʾ- (hamza, *w*-, and *y*- types. Initial ʾ- shows two verbs, *axaḏ* (or *xaḏa*) ‘he took’, *yāxəḏ* ‘he takes’, *yāxḏūn* ‘they [pl. masc.] take’, etc.; and *akal* (or *kala*) ‘he ate’, *yākəl* ‘he eats’, *yāklūn* ‘they [masc.] eat’, etc. If the initial consonant is *w*, it has the exponent /ō/ as in *wəzan* ‘he weighed’, *yōzan* ‘he weighs’, *yōznūn* ‘they [pl. masc.] weigh’, *ōzan* ‘I weigh’. Initial *y*- has only one item, *yəbas* ‘it dried’. In the imperfect the exponent of *y*- is /ē/, as in *yēbas* ‘it dries’, etc.

The imperative has the forms *kəl* or *əkəl* ‘eat!’, *xəḏ* or *əxəḏ* ‘take!’, and *ōgaf* ‘stand!’. No examples of imperatives of *yəbas* were found.

2.2.4.5.2 *IIw/y* verbs

Here the medial *w* or *y* is realized vocalically or as zero, giving such forms as *šāl* ‘he took away', *šālat* 'she took away', *šǝlǝt* 'I took away', *ǝyšīl* 'he takes away', *yšīlūn* 'they [pl. masc.] take away', *šīl* 'take away!', *šīli* 'take away [sg. fem.]!', *šīlu* 'take away [pl. masc.]!' The *-an* alternative of the 1st person singular regularly occurs, as in *arūḥan* 'I go', *ašīlan* 'I carry', *agūlan* 'I say', *ašūfan* 'I see'.

2.2.4.5.3 *IIIw/y* verbs

The dialect only includes verbs in final *-y*, but it includes both *-a* and *-i* types, as in *yǝbči* 'he cries', *yǝlga* 'he finds'. The final *y* is realized vocalically or as zero in weak stems, as in *mǝša* 'he went', *mǝšat* 'she went', *mǝšēt* 'I went', *yamši* 'he goes', *yǝmšūn* 'they [pl. masc.] go', *yǝmšan* 'they [pl. fem.] go', *ǝmši* 'go!', *ǝmši* 'go [sg. fem.]!', *ǝmšu* 'go [pl. masc.]!', *ǝmšan* 'go [pl. fem.]!'. Note that the *ʿarab* dialect shows masculine singular imperatives with no final vowel, but an anaptyctic vowel is shown intervening between the resulting consonant cluster, as in *ǝmǝš* 'go!', *ǝḥǝč* 'speak!'

2.2.4.5.4 Generalization of the final weak type

The geminate type shows a formative *-ē-* in the perfect in *raddēt* 'I answered'. Some speakers generalize this to other classes, giving such forms as *kǝtbēt* 'I wrote', *kǝtbēna* 'we wrote', *wǝznēt* ' I weighed', *šālēt* 'I carried'.

2.2.5 Preverbal particles

Preverbal particles mark negation, tense, and mode. They are often reduced forms of verbs or other elements and include negators and tense and mode markers.

2.2.5.1 Negators

Negators include *mā, lā*, and, in the marshland dialect, *ʿēb-*, all of which are stressed. The element *mā* precedes indicatives, while *lā* precedes wishes and imperatives, as in *mā yā* 'he didn't come', *mā arīda* 'I don't want it', *lā trūḥ* 'don't go!', *lā ǝyyi* 'let him not come!', *lā yšūfna* 'let him not see us!'. The marshland element *ʿēb* occurs in indicatives, as in *ʿēb ǝyṭǝlʿan barra* 'they [the buffaloes] do not go out', *ʿēb nǝdri* 'do we not know?', *ʿēb bīhǝn īdām* 'they have no fodder'.

2.2.5.2 Tense and mode markers

Markers include *gāʿǝd* present continuous, *xall-* jussive, *raḥ-* future, *kūn-, wākūn-* 'must', *čān-* 'should have', as in *gāʿǝd ǝššūf* 'she is looking', *xall-nrūḥ* 'let's go!', *xal-ǝyyi* 'let him come!', *raḥ-arūḥan* 'I am going to go', *kūn ašūfannak* 'I must see you', *čān gǝtla* 'you should have told him'.

2.3 *Syntax*

The syntax of the dialect is unremarkable, as it follows the general Arabic pattern. Sentences can be of the form VSO or can front any noun phrase in a Topic/Comment construction. Examples are (VSO) *ḏǝbaḥ āfa rmǝḥi* or *ḏǝbaḥ rmǝḥi āfa* 'my lance killed a dragon', *jatlak mǝn šǝbb u šāyǝb* 'young and old came to you', (Topic/Comment) *abu zēd sawwōh ʿabǝd* 'they disguised Abu Zaid as a slave', *ṭāri aḏḏīb mā yābatta* 'she did not mention the wolf', *hāḏa ḏḏīb ʿabar* 'this wolf crossed over', *ǝzzanāti xalīfa ʿǝdda xandag* 'Zanāti Khalīfa had a ditch'.

3. LEXICON

The lexicon is mainly that of southern Iraq and is not very different from that of Baghdad, although there are some items that link it more to the Gulf Coast. Typical examples are *anṭa/yǝnṭi* 'to give', *bāryāw* 'flooded land', *bāwaʿ/ybāwǝʿ* 'to look at', *bǝ-* with the meaning 'in' and 'at' (*fī* does not occur), *farax* 'child', *harfi* 'early', *hēč* 'thus, like this', *hnā* 'here', *ḥadǝr* 'under', *ḥalg* 'mouth', *ǝntǝšal/yǝntǝšǝl* 'to catch a cold', *kaḏḏ/ykǝḏḏ* 'to grasp', *kǝtal/yǝktǝl* 'to hit, beat', *laʿad* 'so', *lǝban* 'yoghurt mixed with water', *lo* 'or', *mawwat/ymawwǝt* 'to kill', *nǝšad/yǝnšǝd* 'to ask', *rād/yrīd* 'want', *rōba* 'yoghurt', *sīda* 'straight ahead', *tāna/ytāni* 'to wait for', *xašǝm* 'nose'.

BIBLIOGRAPHICAL REFERENCES

Ingham, Bruce. 1973. "Urban and rural Arabic in Khūzistān". *Bulletin of the School of Oriental and African Studies* 36.633-53.

——. 1976. "Regional and social factors in the dialect geography of southern Iraq and Khūzistān". *Bulletin of the School of Oriental and African Studies* 39.23-5.

——. 1982. *North East Arabian dialects.* London: Kegan Paul International.

Layard, A.H. 1846. "Description of the province of Khuzistan". *Journal of the Royal Geographical Journal* 16.1-105.

Loftus, William Kenneth. 1857. *Travels and researches in Chaldaea and Susiana.* London.

Lorimer, D.L.R. (n.d.). *A notebook on the Arabic of Ahwāz.* Unpublished MS deposited in the library of the School of Oriental and African Studies, London University.

Lorimer, J.G. 1915. *Gazetteer of the Persian Gulf, Oman and Central Arabia*, 2 vols. Calcutta.
Stocqueler, J.H. 1832. *Fifteen months pilgrimage through untrodden tracts of Khuzistan and Persia*. 2 vols. London.

Bruce Ingham
(School of Oriental and African Studies)

Kināya

1. The concept of *kināya*

Unlike many Medieval Arabic rhetorical processes, such as *istiʿāra* 'metaphor', *kināya* cannot be matched appropriately with an English trope or figure. Yet, the rhetorical processes involved are of a general nature, supporting the idea that *kināya* is not only a specific trope of the Arabic linguistic tradition but also a general concept of rhetoric and speech analysis (Dichy 2003). Three arguments support this assumption and at the same time account for the difficulty in investigating the meaning of the term:

i. The notion that *kināya* represents a general rhetorical concept can only arise from a close analysis of lexical, rhetorical, and exegetic Arabic sources in a historical perspective. In Western Arabic studies, *kināya* is often translated as 'metonymy' (e.g. Pellat 1986:116) or 'periphrasis' (e.g. Lecomte 1965:296; Heinrichs 1977:31; Larkin 1995:75). These apparently contradictory translations correlate, at least partly, with specific texts, authors, or epochs. Medieval Arabic rhetoric underwent considerable development over the lengthy period between the 8th century and the 14th century C.E. Well-known definitions appearing in contemporary didactic publications (e.g. al-Jārim and ʾAmīn 1936) only report a late, stabilized 'state of the art' and do not account for either polysemy due to historical evolution or the possibility of referring this evolution to a general rhetorical process.
ii. The sense of *kināya* evolved from the beginning of Qurʾānic exegesis to later rhetorical treatises, in which it is defined (after Ibn al-ʾAṯīr [d. 637/1239], *Maṯal* II, 180ff.) as a device by which a word, phrase, or utterance may support both tropic and literal interpretations, or either one. *Kināya* thus appears as a dual rhetorical process to be interpreted beyond the mere notions of trope or figure (hence the difficulty in finding a corresponding term in English or French rhetoric).
iii. The original lexical meaning of *kināya* is 'indirect expression' (antonym: *taṣrīḥ* 'explicit expression'). In Islamic law, the word refers to a declaration of intent phrased in indirect terms (e.g. Ibn al-Mināwī [d. 1031/1621], *Tawqīf* 285; Ḥasb ʾAllāh 1997:223–226). In many rhetorical treatises *kināya* is introduced with *taʿrīḍ* 'implication, indirect intimation'. Both discourse devices – to which *tawriya* 'concealed expression' should be added – illustrate what can be called a "rhetoric of indirect wording" (Dichy 2004).

2. Early uses

Lexically, *kināya* is a nominalized infinitive form (*maṣdar*) of the verb *kanā*. Ibn Fāris (d. 385/1005) states that it denotes "the concealed expression (*tawriya*) of a denomination through [the use of] another. One says *kanaytu ʿan kaḏā* 'I avoided mentioning something' upon using another expression from which the first one can be inferred" (*Maqāyīs* V, 139). Such indirect wording may be due to modesty, where overt utterance of the concealed expression is deemed 'abominable, impudent, obscene' (*yustafḥaš;* Ibn Manẓūr [d. 710/1311], *Lisān*, root *k-n-y*). Lexicographers mention a second sense of the verb, with which another nominalized infinitive form, *kunyā* 'surname, agnomen', is connected. The *kunyā* consists of *ʾabū* 'father of' or *ʾumm* 'mother of' followed by the name of the son. It is still used either in order to avoid uttering in public someone's actual name (commonly that of a woman), or as an honorific or friendly term of address. Ibn Manẓūr (*Lisān,* root *k-n-y*) quotes the use of *kunyā* in wartime by competing fighters (*mubārizīn*), whose surnames are thus remembered. Ibn Fāris explicitly relates the second term to the first: "*Kināya* stands in opposition to 'explicit expression' (*muṣāraḥa*). This is why the surname is called *kunyā*, as if it were a 'concealed expression' (*tawriya*) of someone's name" (*Maqāyīs* V, 139).

Early metalinguistic uses of *kināya* are very much akin to the lexical description of the term, which is taken by Sībawayhi (d. ca. 175/791;

Kitāb II, 170, 415) to refer to pronouns (for discussion of *kunyā*, see *Kitāb* II, 93–101; also Ibn Fāris, *Ṣāḥibī* 439–443). In one of the oldest Qurʾānic commentaries to have reached us, ʾAbū ʿUbayda's (d. 210/825) *Majāz*, the term appears in the gloss of verses mentioning sexual intercourse indirectly (e.g. Q. 2/223, 4/43).

In al-Jāḥiẓ's (d. 255/868) famous rhetorical treatise and anthology, *kināya* occurs in the phrase *al-kināya wa-t-taʿrīḍ* 'inexplicit expression and indirect intimation' (*Bayān* I, 117). Historically, this is a crucial quotation for two reasons: first, *kināya* appears here in a definitely rhetorical sense, and second, in accordance with its original meaning, the word denotes a rhetorical device of 'indirect wording'. Other occurrences of the word in al-Jāḥiẓ's writings relate *kināya* to means of expression opposed to 'explicit utterance' (*taṣrīḥ*). The term describes a 'minimal expression', only partly related to understatement. Al-Jāḥiẓ gives the example of "declaring somebody moderate (*muqtaṣid*) as an inexplicit expression (*kināya*) of his being a miser" (*Bayān* I, 263). Gesture and silence are magnified along with *kināya* when they "bring forth what spoken utterance (*qawl*) is unable to [express]" (Jāḥiẓ, *Rasāʾil* I, 308). Al-Jāḥiẓ (*Bayān* I, 115–116) quotes Ibn al-Muqaffaʿ (d. ca. 142/749), who deems allusion (*lamḥa*) to be 'eloquent discourse' (*balāġa*) par excellence. Indirect wording and the various forms of minimal expression are, indeed, intimately related in Medieval Arabic culture, which is, of course, echoed by rhetorical treatises (e.g. al-ʿAskarī [d. 395/1005], *Ṣināʿatayn* 22ff., which includes a reference to Indian rhetoric).

3. THE TWO ASPECTS OF *KINĀYA*

Later definitions feature two fundamental aspects of the notion of *kināya*: direct vs. indirect reference to meaning (*taṣrīḥ* vs. *kināya* and *taʿrīḍ*), and 'tropic' vs. 'non-tropic expression' (→ *majāz* vs. *ḥaqīqa,* the English translation being a rough approximation). The first aspect is partly based on the original lexical meaning of *kināya*, the second on later substantial developments in the conceptual analysis of *majāz*. Both aspects remain interwoven throughout the 9th and 10th centuries C.E., before clear technical definitions of *majāz* appear. Ibn Qutayba (d. 276/889; *Taʾwīl* 15–16) mentions a number of subcategories of *majāz*, which he understands as "everything that goes beyond the strictly logical application of language, i.e., beyond being a true and simple copy of reality" (Heinrichs 1977:31). The notion includes for him both *kināya*, opposed to *ʾīḍāḥ* 'plain and clear designation', and *taʿrīḍ,* the antonym of which is *ʾifṣāḥ* 'overt utterance'. *Kināya* is discussed in relation to the question of avoiding mention (*taʿrīḍ*) of someone's name for the sake of discretion, as in Q. 25/28, where *fulān* 'so-and-so' denotes a blamable friend (Ibn Qutayba, *Taʾwīl* 202, with other examples). The point is significant for Ibn Qutayba's view of language, the nature of which includes rhetorical processes listed by him under *majāz*. Yet, he maintains a restrictive attitude toward the 'interpretation' (*taʾwīl*) of those Qurʾānic expressions that result from such processes (cf. ʿAbd al-Jabbār [d. 415/1024], *Muġnī* XVI, 272–275, on 'ambiguous verses'; for the Jewish tradition, see Fenton 1997:265).

One of the oldest works on Arabic poetic style, Ibn al-Muʿtazz's (d. 296/908) *Badīʿ* (64), also mentions *kināya* together with *taʿrīḍ*. Al-Mubarrid (d. 285/898; *Kāmil* II, 290–292) elaborates on the lexical definition and introduces some rhetorical categorization: 'obscuring' the name of the beloved or one's feelings for the sake of honor; 'covering up' an expression for the sake of modesty; 'honoring' a man by using his *kunyā*. In his more directly rhetorical treatise, *Naqd*, Qudāma ibn Jaʿfar (d. ca. 337/949) considers two tropes, the description of which is taken up by later scholars in the definition of *kināya* although he himself does not use the term. The first trope, *ʾirdāf* 'implication', consists of referring to a meaning (→ *maʿnā*) through an expression (→ *lafḍ*) 'implying' or 'entailing' it but not denoting it directly (*Naqd* 88). The second trope, *ʾišāra,* is defined as a general process by which an expression can take a number of meanings, through hinting and alluding (*ʾīmāʾ* and *lamḥa*). Meaning through allusion is, in addition, described, according to a very old tradition, as a distinctive feature of *balāġa* 'eloquent discourse' (Qudāma, *Naqd* 85; also Ibn al-Muqaffaʿ, quoted by al-Jaḥiẓ and al-ʿAskarī, *Ṣināʿatayn* 22ff.).

Al-ʿAskarī (d. 395/1005; *Ṣināʿatayn* 407–410) gives a single definition for *kināya* and *taʿrīḍ*. He opposes both to 'explicit expression' (*taṣrīḥ*), and analyzes *ʾišāra* in terms very close to those of Qudāma (*Ṣināʿatayn* 383–388).

The second aspect mentioned above (*majāz* vs. *ḥaqīqa*) only seems to emerge in the writings of ʿAbd al-Qāhir al-Jurjānī (d. ca. 474/1082), who elaborated on earlier treatments of tropic vs. non-tropic expression, particularly the one by ʿAbd al-Jabbār (d. 415/1024; *Muġnī* XV, 162; Larkin 1995:38). Al-Jurjānī (*Dalāʾil* 66) defines *kināya* as a rhetorical device in which an expression is used in a different meaning from its own, the 'other meaning' being inferred through a link that can be established 'in the existent world' (*fī l-wujūd*). This view can be related to Qudāma's notion of *ʾirdāf* and explains why *kināya* has so often been translated as 'metonymy'. Discussion on the literal and/or tropic nature of *kināya* continues with Faxr ad-Dīn ar-Rāzī (d. 606/1210; *Nihāya* 190–192) and as-Sakkākī (d. 626/1229; *Miftāḥ* 400–412).

4. Direct vs. indirect reference to meaning

4.1 *Zamaxšarī's denial*

The position of az-Zamaxšarī (d. 538/1144) deserves special mention: his great Qurʾānic commentary includes a substantial number of rhetorical analyses and presents a very technical view of *kināya*. The term noticeably appears in his commentary on 'anthropomorphic verses' in which 'the hand' of God and God's 'throne' are mentioned (e.g. Q. 48/10, 20/5). The interpretation of these verses plays an essential part in the development of Arabic rhetoric (an-Nuwīrī 2001:220ff.). Az-Zamaxšarī (*Kaššāf* I, 215) gives only a vague definition of *kināya*: "mentioning something by an expression other than its own", but he states, significantly, that "a speaker cannot denote, through a single expression, [both] non-tropic and tropic meanings" (III, 298; ʾAbū Mūsā 1988:545–563), which is a close application of what we would call the law of the excluded middle. Az-Zamaxšarī's denial of the conjunction of 'proper' and tropic meanings in *kināya* can be regarded as a forerunner of 7th/13th-century definitions of the term. It is related to his rationalistic Muʿtazilī position on the question of 'anthropomorphic verses' and to his rejection of al-ʾAšʿarī's doctrine of *bi-lā kayf*, according to which these verses are to be admitted by the believer 'without [asking] how' tropic and non-tropic interpretations relate.

4.2 *The turning point*

In the following century, Ibn al-ʾAṯīr (d. 637/1239; *Maṯal* II, 180–201) defines *kināya* as: "[an expression] drawn to either proper or tropic meaning, which can be interpreted from both sides", owing to a "descriptive feature comprehending (*waṣf jāmiʿ li-*) each of these senses" (II, 181–182). This definition brings – from a purely rhetorical standpoint – an elegant solution to az-Zamaxšarī's objection: the meaning of *kināya* is said to support either 'proper' or tropic interpretation, or both meanings, according to the case. In other words, Ibn al-ʾAṯir's move considerably increases the power of the rhetorical device by adding the possibility of disjunctive *kināya* ('either proper or tropic meaning') to that of conjunctive *kināya* ('both proper and tropic senses'). Az-Zamaxšarī, in the discussion quoted above, only considered disjunctive meaning.

Ibn al-ʾAṯīr furthermore includes in the process of *kināya* a sort of pendulum movement between conjunction and disjunction, which is expressed in the definition of the term by the verb 'to draw to' (*tajāḏaba*) and illustrated immediately thereafter by a revisiting of the traditional example "or [if] you touch women" (Q. 4/43; cf. ʾAbū ʿUbayda, Sec. 2 above). This phrase can be taken either in its proper meaning (no rhetorical process involved), or as a *kināya*, the tropic sense of which is sexual intercourse. The latter, Ibn al-ʾAṯīr points out, does include actual touching. The phrase is associated, in Q. 4/43, with the obligation of ritual washing, which can be entailed, according to the legal interpretation chosen, either by sexual intercourse including touching, i.e. conjunctively, by both 'proper' and tropic senses (only sexual intercourse renders ritual washing obligatory); or else, disjunctively, by either (or any) proper *or* tropic meanings (in which case, ritual washing is deemed obligatory for a man whenever touching a woman).

Treatises on Arabic rhetoric often quote al-Jurjānī's example of *naʾūm aḍ-ḍuḥā* 'a regular [or deep] morning sleeper', said, in a complimentary way, of a well-off woman. The tropic meaning to be inferred 'in the existent world' (*fī l-wujūd*) is that of affluence and the assistance of servants, so that the woman described does not have to wake up early in the morning (Jurjānī, *Dalāʾil* 66, 262; the traditionally unquoted source of the example is ʾImruʾ al-Qays, *Muʿallaqa*, l. 38). Ibn al-ʾAṯīr adds the idea that the 'comprehensive description' (*waṣf jāmiʿ*) virtually includes the non-tropic meaning, i.e. that this person effectively sleeps late, and that either sense – or both – can be borne by the expression *naʾūm aḍ-ḍuḥā*.

A century later, al-Qazwīnī (d. 739/1338) presents a revised definition, based on further progress in characterizing *majāz*: "*Kināya* refers to an expression indicating an implied meaning (*lāzim maʿnāhu*), together with the possibility of indicating its own meaning, here and then" (*ʾĪḍāḥ* V, 158). It is distinguished from *majāz* (*stricto sensu*) in that *majāz* is restricted to tropic meaning "by a contextual element (*qarīna*) prohibiting literal meaning" (*ʾĪḍāḥ* V, 12). 'Context' is to be taken here in a very broad meaning.

Both definition and restriction derive, together with some added technical discussion, from the treatise on which al-Qazwīnī's *ʾĪḍāḥ* is a commentary: as-Sakkākī's *Miftāḥ* (359, 402). The former includes, albeit implicitly, Ibn al-ʾAṯīr's conception. Restriction of *majāz* to tropic meaning through (broad) context had already been mentioned, in different terms, by al-Jurjānī (*ʾAsrār* 304), and the relation of 'implication' between the two meanings involved in a given *kināya* can be traced back to Qudāma's *ʾirdāf*. Al-Qazwīnī's definition is still widely reproduced today (al-Jārim and ʾAmīn 1979:125) and may be regarded as an essential synthesis of the tropic aspect of *kināya*.

On the other hand, the second fundamental aspect of the notion (explicit vs. indirect wording, cf. Qudāma's *ʾišāra*) falls out of focus. This not only obscures the relation between primary and later definitions of *kināya*, it also leaves unanswered the question of the rhetorical effect that explains the use of that speech process.

5. Tropic vs. non-tropic expression

Ibn al-ʾAṯīr's definition was considered crucial enough to be added almost verbatim to the original lexical meanings of *kināya* in al-Fayrūzābādī's (d. 816/1413) *Muḥīṭ* (IV, 386). Ibn al-ʾAṯīr introduced, in addition, an explicit distinction between *kināya* and *taʿrīḍ* 'indirect intimation, implication'. The fundamental difference is that *kināya* is based on tropic and non-tropic interpretations, and *taʿrīḍ* is "an expression that denotes something through induced meaning (*ʿan ṭarīq al-mafhūm*) and not by way of proper or tropic institution of meaning (*lā bi-l-waḍʿ al-ḥaqīqī wa-lā al-majāzī*)" (Ibn al-ʾAṯīr, *Maṯal* II, 186). He follows al-Jurjānī's definition of language 'institution' (*waḍʿ*), which may correspond either to *ḥaqīqa,* i.e. to an expression in which words directly reflect the world (cf. Ibn Qutayba in Sec. 3 above), or to *majāz,* i.e. to tropic meanings produced in specific utterances (Jurjānī, *ʾAsrār* 303–304). Ibn al-ʾAṯīr goes on: "If you say, for instance, to someone whom you expect to be beneficent and bestowing toward you without [any need for you] to ask, 'By God, I am in need, and have nothing in my hands' and 'I am naked and stung by cold', such expressions and other similar ones are indirect intimations (*taʿrīḍ*) of [your] asking". Contemporary pragmatic analyses describe such dialogical utterances as indirect speech acts (Searle 1979). But 'indirect intimation' can also be purely descriptive when included, for example, in a narrative: in the Qurʾānic story of Joseph, no expression directly refers to the young man's beauty (as in Genesis 39:6), but his beauty is expressed powerfully through *taʿrīḍ* in the description of women who had been invited to a light meal and who, upon catching sight of him, "made cuts to their hands" (Q. 12/31), presumably while they were peeling fruit. The two meanings inferred are that of ravishment and of the beauty that entailed it.

Both direct and indirect meanings are supported by the examples above. The person in Ibn al-ʾAṯīr's example *is* in need, and the women in rapture *did* make cuts to their hands in the Qurʾānic narrative. Likewise, in Searle's analysis both direct and indirect meaning are supported by indirect speech acts. The virtual conjunction of straightforward and inferred meanings is therefore common to *kināya* and *taʿrīḍ.* So are inference from context and indirect wording (as opposed to *taṣrīḥ* 'explicit expression'). The two main differences, according to Ibn al-ʾAṯīr, lie in the inference process in *taʿrīḍ:* it is – in modern terms – situational and pragmatic rather than based on the actual words of the utterance. The rhetorician significantly remarks that *taʿrīḍ* can only occur in compound utterances, whereas *kināya* can be restricted to a single word. Moreover, in *kināya*, as opposed to *taʿrīḍ*, meaning is produced by way of ‘proper’ and/or tropic meaning.

Taʿrīḍ and *kināya* are included in the same chapter by Ibn al-ʾAṯīr, and both belong to a “rhetoric of indirect wording” (Dichy 2003, 2004). The terms also share the meaning of ‘avoiding mention’ of something, when used with the privative preposition *ʿan* (*ʾaʿraḍa ʿan, kanā ʿan*). Inclusion of both rhetorical processes in the same general category links the old lexical meanings of *kināya* to its later sophisticated definitions, and also accounts partly for the not infrequent overlapping of *kināya* and *taʿrīḍ* in previous rhetorical works (e.g. those of Ibn Qutayba and al-ʿAskarī, as well as the compilation of examples by aṯ-Ṯaʿālibī [d. 430/1039] and the one by al-Qāḍī al-Jurjānī [d. 482/1089]). In Classical Arabic rhetoric, indirect wording has traditionally always been deemed ‘more eloquent’ (*ʾablaġ*) than direct speech.

Bibliographical references

Primary sources

ʿAbd al-Jabbār, *Muġnī* = ʾAbū l-Ḥasan ʿAbd al-Jabbār ibn ʾAḥmad al-Hamāḏānī, *al-Muġnī fī ʾabwāb at-tawḥīd wa-l-ʿadl*. Ed. Ṭāhā Ḥusayn and ʾIbrāhīm Madkūr. 20 vols. Cairo: Maṭbaʿa Dār al-Kutub, 1960–1969. (Not all volumes available to the public.)

ʾAbū ʿUbayda, *Majāz* = ʾAbū ʿUbayda Maʿmar ibn al-Muṯannā at-Taymi, *Majāz al-Qurʾān*. Ed. Muḥammad Fūʾād Sazgīn [Fuad Sezgin]. 2 vols. Cairo: Muḥammad Sāmī ʾAmīn al-Xānjī, 1954.

ʿAskarī, *Ṣināʿatayn* = ʾAbū Hilāl al-Ḥasan ibn ʿAbdallāh al-ʿAskarī, *Kitāb aṣ-ṣināʿatayn, al-kitāba wa-š-šiʿr*. Ed. Mufīd Qumayḥa. Beirut: Dār al-Kutub al-ʿIlmiyya, 1989.

Fayrūzābādī, *Muḥīṭ* = Majd ad-Dīn Muḥammad ibn Yaʿqūb al-Fayrūzābādī, *al-Qāmūs al-muḥīṭ*. 4 vols. Cairo: al-Bābī al-Ḥalabī, 1952.

Ibn al-ʾAṯīr, *Maṯal* = ʾAbū l-Fatḥ Ibn al-ʾAṯīr, *al-Maṯal al-sāʾir fī ʾadab al-kātib wa-š-šāʿir*. Ed. Muḥammad ʿAbd al-Ḥamīd. 2 vols. Beirut: al-Maktaba al-ʿAṣriyya, 1990.

Ibn Fāris, *Maqāyīs* = ʾAbū l-Ḥusayn ʾAḥmad Ibn Fāris, *Maqāyīs al-luġa*. Ed. ʿAbd as-Salām Hārūn. 6 vols. (Repr., Beirut: Dār al-Jīl, 1991.)

——, *Ṣāḥibī* = ʾAbū l-Ḥusayn ʾAḥmad Ibn Fāris, *aṣ-Ṣāḥibī fī fiqh al-luġa*. Ed. A.M. Ṣaqr. Cairo: al-Bābī al-Ḥalabī, 1977.

Ibn al-Munāwī, *Tawqīf* = Zayn ad-Dīn ʿAbd ar-Raʾūf ibn Tāj al-ʿĀrifīn ibn ʿAlī al-Munāwī, *at-Tawqīf ʿalā muhimmāt at-taʿārīf*. Ed. Muḥammad Riḍwān ad-Dāya. Beirut and Damascus: Dār al-Fikr, 1990.

Ibn Manẓūr, *Lisān al-ʿArab* = ʾAbū l-Faḍl Jamāl ad-Dīn Mukarram ibn Mukarram Ibn Manẓūr, *Lisān al-ʿArab*. Cairo: Dār al-Maʿārif, n.d.

Ibn al-Muʿtazz, *Badīʿ* = ʿAbdallāh Ibn al-Muʿtazz, *Kitāb al-badīʿ*. Ed. Ignatius Kračkovskij. London, 1935. (Repr., Beirut: Dār al-Masīra, 1982.)

Ibn Qutayba, *Taʾwīl* = ʿAbdallāh ibn Muslim Ibn Qutayba, *Taʾwīl muškil al-Qurʾān*. Ed. as-Sayyid ʾAḥmad Ṣaqr. Cairo: al-Bābī al-Ḥalabī, 1954.

Jāḥiẓ, *Bayān* = ʾAbū ʿUṯmān ʿAmr ibn Baḥr al-Jāḥiẓ, *al-Bayān wa-t-tabyīn*. Ed. ʿAbd as-Salām Hārūn. 4th ed. 4 vols. Cairo: Maktabat al-Xānjī, 1975.

——, *Rasāʾil* = ʾAbū ʿUṯmān ʿAmr ibn Baḥr al-Jāḥiẓ, *Rasāʾil*. 4 vols. 4th ed. (incomplete). Ed. ʿAbd as-Salām Hārūn. Cairo: Maktabat al-Xānjī, 1964.

Jurjānī, *ʾAsrār* = ʾAbū Bakr ʿAbd al-Qāhir ibn ʿAbd ar-Raḥmān al-Jurjānī, *ʾAsrār al-balāġa*. Ed. M. Rašīd Riḍā. (Repr., Beirut: Dār al-Maʿrifa, n.d.)

——, *Dalāʾil* = ʾAbū Bakr ʿAbd al-Qāhir ibn ʿAbd ar-Raḥmān al-Jurjānī, *Dalāʾil al-ʾiʿjāz*. Ed. M.M. Šākir. Cairo and Jeddah: al-Madanī, 1992.

Mubarrid, *Kāmil* = ʾAbū l- ʿAbbās Muḥammad ibn Yazīd al-Mubarrid, *al-Kāmil*. Ed. Muḥammad ʾAbū l-Faḍl ʾIbrāhīm and A. Šaḥāta. 4 vols. Cairo: Dār Nahḍa Miṣr, n.d.

Qāḍī al-Jurjānī, *Muntaxab* = al-Qāḍī al-Jurjānī, *al-Muntaxab min kināyāt al-ʾudabāʾ wa-ʾiršādāt al-bulaġāʾ* [followed by aṯ-Ṯaʿālibī, *al-Kināya wa-t-taʿrīḍ*]. Beirut: Dār al-Kutub al-ʿIlmiyya, 1984.

Qazwīnī, *ʾĪḍāḥ* = Jalāl ad-Dīn ʾAbū ʿAbdallāh Muḥammad ibn ʿAbd ar-Raḥmān al-Xaṭīb al-Qazwīnī, *al-ʾĪḍāḥ fī ʿulūm al-balāġa*. Ed. Muḥammad al-Xafājī. 6 vols. 3rd ed. Beirut: Dār al-Jīl, 1993.

Qudāma, *Naqd* = Qudāma ibn Jaʿfar, *Kitāb naqd aš-šiʿr*. Ed. Seeger A. Bonebaker. Leiden: E.J. Brill, 1965.

Rāzī, *Nihāya* = Faxr ad-Dīn Muḥammad ibn ʿUmar ar-Rāzī, *Nihāyat al-ʾījāz fī dirāyat al-ʾiʿjāz*. Ed. A.H. as-Saqqā. Cairo: al-Maktab aṯ-Ṯaqāfī, al-ʾAzhar, 1989.

Sakkākī, *Miftāḥ* = ʾAbū Yaʿqūb Yūsuf ibn ʾAbī Bakr as-Sakkākī, *Miftāḥ al-ʿulūm*. Ed. Nuʿaym Zarzūr. Beirut: Dār al-Kutub al-ʿIlmiyya, 1983.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. Ed. ʿAbd as-Salām Hārūn. 5 vols. Cairo: al-Hayʾa al-Miṣriyya al-ʿĀmma li-l-Kitāb, 1966–1977.

Ṯaʿālibī, *Nihāya* = ʾAbū Manṣūr ʿAbd al-Malik ibn Muḥammad aṯ-Ṯaʿālibī, *an-Nihāya fī l-Kināya* [known as *al-Kināya wa-t-taʿrīḍ*]. Ed. Faraj al-Ḥiwār. Sūsa and Tunis: Dār al-Maʿārif li-ṭ-Ṭibāʿa wa-n-Našr, 1995.

Zamaxšarī, *Kaššāf* = ʾAbū l-Qāsim Jār ʾAllāh Maḥmūd ibn ʿUmar az-Zamaxšarī, *al-Kaššāf*. Ed. Muṣṭafā Ḥusayn ʾAḥmad. 4 vols. Cairo: Maṭbaʿat al-Istiqāma, 1946.

Secondary sources

ʾAbū Mūsā, Muḥammad M. 1988. *al-Balāġa al-qurʾāniyya fī tafsīr az-Zamaxšarī*. 2nd ed. Cairo: Maktaba Wahba.

Dichy, Joseph. 2003. “*Kinâya*, a tropic device from Medieval Arabic rhetoric, and its impact on discourse theory”. *Proceedings of the 5th Conference of the International Society for the Study of Argumentation (ISSA), University of Amsterdam, 25–28 June 2002*, ed. Frans van Eemeren, J. Anthony Blair, Charles A. Willard, and A. Francisca Snoeck Nenkemans, 237–241. Amsterdam: Sit Sat.

——. 2004. “L’indicible beauté virile: Le mythe coranique de la tentation de Joseph”. *Le voyage et la langue: Mélanges offerts à Anouar Louca et André Roman*, ed. Joseph Dichy and Hassan Hamzé, 155–186. Damascus: Institut Français d’Etudes Arabes de Damas.

Fenton, Paul B. 1997. *Philosophie et exégèse dans* Le jardin de la métaphore *de Moïse Ibn ʿEzra, philosophe et poète andalou du XII^e^ siècle*. Leiden: E.J. Brill.

Ḥasb ʾAllāh, ʿAlī. 1997. *ʾUṣūl at-tašrīʿ al-ʾislāmī*. 7th ed. Cairo: Dār al-Fikr al-ʿArabī.

Heinrichs, Wolfhart. 1977. *The hand of the northwind: Opinions on metaphor and the early meaning of* ʾistiʿāra *in Arabic poetics*. Wiesbaden: F. Steiner.

Jārim, ʿAlī al- and Muṣṭafā ʾAmīn. 1936. *al-Balāġa al-wāḍiḥa*. 4th ed. (Repr., Cairo and Beirut: Dār al-Maʿārif, 1979.)

Larkin, Margaret. 1995. *The theology of meaning: ʿAbd al-Qāhir al-Jurjānī’s theory of discourse*. New Haven, Conn.: American Oriental Society.

Lecomte, Gérard. 1965. *Ibn Qutayba: L’homme, son oeuvre, ses idées*. Damascus: Institut Français de Damas.

Nuwīrī, Muḥammad an-. 2001. *ʿIlm al-kalām wa-n-naḏ̣ariyya l-balāġiyya ʿinda l-ʿArab*. Tunis and Sfax: Kulliyyat al-ʿUlūm al-ʾInsāniyya wa-l-Ijtimāʿiyya and Dār Muḥammad ʿAlī al-Ḥāmī.

Pellat, Charles. 1986. “*Kināya*”. *Encyclopaedia of Islam* V, 116–118. 2nd ed. Leiden: E.J. Brill.

Searle, John R. 1979. *Expression and meaning*. Cambridge: Cambridge University Press.

Joseph Dichy (Université Lumière-Lyon-2)

Kinship Terms

Kinship has always played a significant role in human societies on the political, economic, and social levels. Throughout history, people have sought refuge in biologically based bonds expressed socially through kinship relations. These relations have provided political, economic, and social security for both the individual and the group at varying levels in human culture (Farber 1968; Goody 1969). The impact of kinship bonds has generally diminished with the emergence of industrialization and urbanization as civil service structures and government institutions have replaced tribal affiliations in industrial societies. The change is less noticeable in underdeveloped, less industrialized societies, though, where tribal affiliations continue to perform many important functions.

Kinship was a very influential factor in Arab communities before Islam. As a result of tribal rivalry in the Arabian Peninsula and neighboring regions, tribal affiliation constituted a crucial factor in the allocation of political, economic, and social power among existing Arab tribes (Ḥuṣām ad-Dīn 1990; Smith 1990). With the advent of Islam and its subsequent large-scale conquests, it was proclaimed that tribalism had gone forever in compliance with deeply cherished Islamic teachings by Muslim Arabs. However, medieval Arab history shows hard evidence that Islam did not manage to eradicate tribal tendencies from the psychology of Arabs. On the contrary, kinship and tribal influence, rendered dormant only for a short period of time during the life of the prophet Muḥammad, reemerged on a more intense level after the Prophet’s death and, in effect, led to an everlasting rupture in the Arab-Muslim community, that is, the birth of the Sunni and Shiʿi sects.

Kinship bonds are still an all-pervasive phenomenon in Arab societies. Their impact can be easily felt in different walks of life, despite continuing, but mostly nominal, measures taken by modern Arab states to curb them. At the global political level, most, if not all, Arab dynasties and republics have evolved from or have evolved into family investments or bureaucracies that are overwhelmingly based on kinship orientation. At the societal level, the tribal paradigm may have lost some ground in urban centers, but it is still operating on full power in rural and Bedouin quarters. The popular Arabic proverbs *ʾana wa-ʾaxī ʿalā bn ʿammī wa-ʾana wa-bn ʿammī ʿalā l-ġarīb* ‘I stand by my brother against my (paternal) cousin and by my (paternal) cousin against the stranger’ and *unṣur ʾaxāka ḏ̣āliman ʾaw maḏ̣lūman* ‘stand by your brother whether he is oppressor or oppressed’ still represent cherished didactic morals for mainstream Arabs. In these encouraging sociocultural milieus, the sociolinguistics of Arabic kinship terms has developed enormously in terms of structure, scope, and function.

1. Morphological properties

Like most Arabic common nouns, kinship terms mainly mark gender distinctions derivationally by adding as a suffix the feminine marker *tāʾ marbūṭa* to the masculine kinship term. Examples include *ʿamm/ʿamma* ‘paternal uncle/paternal aunt’, *xāl/xāla* ‘maternal uncle/maternal aunt', *ḥafīd/ḥafīda* 'grandson/granddaughter', *ʾax/ʾuxt* 'brother/sister', *ibn/ibna* 'son/daughter', *zawj/zawja* 'husband/wife'. However, there are a few suppletive forms where the feminine term is not derived morphologically from its masculine counterpart; examples include *ʾab/ʾumm* 'father/mother' and *walad/bint* 'son/daughter'. Historically, many of the kinship terms that were not marked for gender in Classical Arabic now show such a distinction in the standard as well as the vernacular variety. For example, in the Classical Arabic variety, the terms *zawj* and *walad* used to mean 'husband/wife' and 'children', respectively, but they now mainly or exclusively denote the masculine member.

In terms of word structure, Arabic employs single lexemes plus gender marking to denote immediate (i.e. one stage removed) kinship bonds such as *ʾab/ʾumm* 'father/mother', *ʾax/ʾuxt* 'brother/sister', *zawj/zawja* 'husband/wife', *ʿamm/ʿamma* 'paternal uncle/paternal aunt', *xāl/xāla* 'maternal uncle/maternal aunt', *jadd/jadda* 'grandfather/grandmother', *ḥafīd/ḥafīda* 'grandson/granddaughter'. On the other hand, kinship terms referring to non-immediate bonds, i.e. more than one stage removed, are typically created by compounding. Examples of kinship compounds include *ibn ʾaxī/ʾuxtī* 'son of my brother/my sister [nephew]', *ibn xālī/xāltī* ' son of my maternal uncle/maternal aunt [cousin]', *zawjat ʾaxī* 'my brother's wife [sister-in-law]', *zawj ʾummī* 'my mother's husband [my stepfather]', *ʿamm ʾabī/ʾummī* 'paternal uncle of my father/my mother'. In creating these compounds, one kinship term is used to modify another. Further, more remote kinship relations can be designated by combinatory units featuring more than two kinship terms, such as *ibn bint ʾaxī* 'son of my brother's daughter', *ʾuxt ʾumm zawjatī* 'sister of my wife's mother', *ibn ibn xālti* 'son of my maternal aunt's son', and so on.

2. Semantic properties

Unlike other Arabic common nouns, kinship terms may not be used as one-place predicates because they are inherently relational, that is, they cannot function by themselves as predicating expressions for subjects. Contrast the acceptable sentences in (1) with the unacceptable ones in (2):

(1) a. *sāmī jundiyyun* 'Sami is a soldier'
b. *yazīdu muhandisun* 'Yazid is an engineer'

(2) a. ?*sāmī ʾabun* 'Sami is a father'
b. ?*yazīdu xālun* 'Yazid is a maternal uncle'

Kinship terms, unlike other common nouns in Arabic, cannot be predicated of subjects as shown in (2). They are two- rather than one-place predicates, as illustrated in (3).

(3) a. *sāmī ʾabun li-bnatayni* 'Sami is a father of two girls'
b. *yazīdu xālu fāṭimata* 'Yazid is Fatima's maternal uncle'

They can, however, be predicated of subjects if the predicating expression includes an embedded proposition in the form of a premodifier, as shown in (4).

(4) a. *sāmī ʾabun miṯāliyyun* 'Sami is an ideal father'
b. *yazīdu xālun sayyiʾun* 'Yazid is a bad maternal uncle'

In addition, Arabic kinship terms enter into several semantic relations with one another – relations such as synonymy, hyponymy, converseness, and incompatibility. Synonymous kinship terms are cognitive synonyms that usually differ in formality, as in the following pairs, where the second item in each pair is the formal Standard Arabic form: *zawja/ʿaqīla* 'wife', *ibn/najl* 'son', *ibna/karīma* 'daughter', *ʾab/wālid* 'father', *ʾumm/wālida* 'mother'. Similarly, Arabic vernaculars often use informal synonyms for some kinship terms, e.g. *mara* for *zawja* 'wife' and *jōz* for *zawj* in the Levant dialects, *mrātī* 'my wife' and *gōzī* 'my husband' in Egyptian Arabic, and *bāba* and *māma* for *ʾab* and *ʾumm* 'father and mother' in most urbanized Arabic vernaculars. Interestingly, vernacular forms may cause problems interdialectally. For instance, the Levantine *mara* 'wife' is pejorative in Egyptian Arabic, where it means 'a worthless woman', despite its phonological and etymological similarity to the Egyptian *mrātī* 'my wife'; both are vernacular versions based on the lexeme *imraʾa* 'woman' in Standard Arabic.

In some cases, the semantic relation is that of inclusion, where a kinship term is a hyponym of another. For example, *ʾab* and *ʾumm* ‘father and mother’ are co-hyponyms of *al-wālidāni* ‘parents’, *ʾabnāʾ* and *ʾaḥfād* ‘children and grandchildren’ are co-hyponyms of *an-nasl* or *al-xalaf* ‘descendants’, and *ibn* ‘son’, *rabīb* or *ibn az-zawja* ‘the wife’s son’, and *daˁiyy* ‘foster son’ are co-hyponyms of *ibn*. In a few cases, the superordinate is not specified for gender. The kinship term *ṣihr* ‘in-law’, for example, includes as co-hyponyms *ʾaxū z-zawj* ‘husband’s brother’ and *ʾuxt az-zawj* ‘husband’s sister’ with respect to the wife. Interestingly, there are few kinship terms that are stranded alone in terms of hyponymy. Examples include *ḍurra* ‘co-wife’ and *ˁadīl* ‘wife’s sister’s husband in relation to the husband’.

The other two semantic relations involve a kind of opposition. The first relates to converseness, where two kinship terms function as relational opposites, or converses of each other, e.g. *jadd/ḥafīd* ‘grandfather/grandson’, *zawj/zawja* ‘husband/wife’, *ʾab/ibn* ‘father/son’, *ʾumm/ibna* ‘mother/daughter’. The second concerns the semantic relation of incompatibility where a number of kinship terms stand in an opposite relation to each other. The kinship terms *ibn/ibna* ‘son/daughter’ and *ḥafīd/ḥafīda* ‘grandson/granddaughter’ constitute pairs of incompatibles.

Interestingly, some Arabic kinship terms have taken on new senses in communication. First, a few are used as ordinary adjectives in expressions like *mubādara ʾaxawiyyya* ‘a brotherly initiative’ and *dawla šaqīqa* ‘a sister country [i.e. an Arab country]’. Second, a few others have found their way into common interjections such as *ya māma/yammah* ‘oh, mother!’, meaning ‘gee!, goodness!’, and *ʾax* ‘brother!’, meaning ‘ouch!’. Third, some are employed in ritualistic expressions such as *ʾaxī l-ˁazīz/ʾuxtī l-ˁazīza* ‘dear brother/dear sister’, as address terms in informal letters, and *ʾayyuhā l-ʾixwatu wa-l-ʾaxawāt* ‘brothers and sisters’, as forms of address in formal speeches. These examples, which are only a few among a multitude, clearly show the far-reaching impact of kinship terms on the affective function of communication in Arabic.

In addition to the above semantic properties, Arabic kinship terms enjoy a noticeable presence in vocatives and in lamentation and/or distress calls. Standard kinship vocatives employ the vocative marker *yā* + kinship term, e.g. *yā ʾummī* ‘oh mother’, *yā ʾabī/ʾabatī* ‘oh father’, *yā ʾuxtī* ‘oh sister’, and *yā ˁammī* ‘oh paternal uncle’. These vocatives can be rendered more intimate by deleting the vocative marker and effecting a uniform morphological change in the kinship term, namely *ʾummāh*, *ʾabatāh*, *ʾuxtāh*, and *ˁammāh*, respectively. Lamentation forms are obtained by prefixing the intimate forms with the morpheme *wā*, namely *wāʾummāh*, *wāʾabatāh*, *wāʾuxtāh*, and *wāˁammāh*, respectively. For example, *wāʾummāh* ‘oh, mother! where are you?’ may be uttered upon the death of one’s mother. However, it may also be uttered to make a distress call by metaphorically lamenting the referent (e.g. mother), who is not coming to the rescue of her son or daughter. Hence, these forms effectively perform lamentation and call-for-help functions. Similar forms exist in vernacular Arabic. These include *wābayyāh* ‘oh, father! where are you?’, *wāxayyāh* ‘oh, brother! where are you?’, and *wāˁammāh* ‘oh, paternal uncle! where are you?’, as they are often heard in the Jordanian vernacular.

3. Blood vs. non-blood relations

Arabic kinship terms belong to two main categories: blood relations and non-blood relations. All immediate husband-wife-child blood relations, whether vertical, horizontal, or external, have non-blood counterparts. The following pairs and sets illustrate this: *ʾab/ˁar-rāb* or *zawj al-ʾumm/ʾab bi-t-tabannī/ḥamū/al-ʾab ar-rūḥī* ‘father/stepfather/foster father/father-in-law/godfather’, *ʾumm/ˁar-rābah* or *zawjat al-ʾab/ʾumm bi-t-tabannī/ḥamāt* ‘mother/stepmother/foster mother/mother-in-law’, *ibn/ar-rabīb* or *ibn az-zawja/daˁiyy* or *ibn bi-t-tabannī* ‘son/stepson/foster son’, *bint/ar-rabība* or *bint az-zawja/daˁiyya* or *bint bi-t-tabannī* ‘daughter/foster daughter/stepdaughter’, *ʾax/šaqīq/ʾax bi-t-tabannī/ʾax bi-r-riḍāˁā/ʾax az-zawja* ‘brother/stepbrother/foster brother/milk brother/brother-in-law’, and *ʾuxt/šaqīqa/ʾuxt bi-t-tabannī/ʾuxt bir-riḍāˁ/ʾuxt az-zawja* ‘sister/stepsister/foster sister/milk sister/sister-in-law’.

Excluding *šaqīq* and *šaqīqa*, only the first member of the above sets necessarily designates a blood relation. The other kinship relations may or may not designate a blood relation depending on the type of marriage, that is, whether or not the marriage in question is consanguineous. For example, in the case of cousin marriages, in-laws include blood relations. Conversely, in non-consanguineous marriages, in-laws will not involve blood relations. This partially contrasts with kinship relations in Western cultures where consanguineous marriages are not sanctioned, hence, there is no overlap between blood and non-blood relations as it exists in the Arab culture. In case of overlap, blood kinship relations take precedence in address forms. A Jordanian wife, for example, would call her father-in-law who happens to be her maternal uncle *xalōh* 'maternal uncle' rather than *ʿammōh* 'father-in-law' in Jordanian Arabic. The latter kinship term may, however, coincide with *ʿammōh* 'paternal uncle' if her husband's father happens to be her paternal uncle.

4. Teknonyms

Arabic teknonyms also employ kinship terms such as *ʾabū*, *ʾumm*, and *ibn* or *bin*, and they are commonly used in absolute titles of address (Yassin 1978). These constitute a popular and, probably, culture-specific trait of Arab culture. Foreigners interested in Arabic and Arab culture cannot miss the heavy presence of teknonyms in both spoken and written communication. Classic examples include *ʾAbū l-Qāsim* 'Prophet Muḥammad', *ʾAbū Bakr aṣ-Ṣiddīq* 'Abu Bakr [the first caliph in Islam]', *ʾAbū ṭ-Ṭayyib al-Mutanabbī* 'al-Mutanabbi [a medieval Arab poet]', *ʾUmm al-Muʾminīn* 'Mother of the Believers [i.e. the Prophet's wife Aysha]', *ʾumm al-maʿārik* 'mother of all battles [i.e. the 1991 Gulf War]', *ʾumm ad-dunyā* 'mother of the universe [i.e. Egypt]', *ʾumm al-qurā* 'mother of villages [i.e. Mecca]', *Ibn Rušd* 'Ibn Rušd [a medieval Arab philosopher]', *Ibn al-Muqaffaʿ* 'Ibn al-Muqaffaʿ [a medieval Arab writer]', *al-Xalīl ibn ʾAḥmad* 'al-Xalīl [a medieval Arab grammarian]', and *Bin Lādin* 'Bin Laden', among scores of other celebrity figures and expressions in Arab-Muslim culture.

Teknonyms still play an important role in the sociolinguistics of present-day Arabic. In the Levant and the Arabian Gulf, the combination *(ʾa)bū/ʾumm* + proper name is a common title of address among friends and in casual interaction, where a more formal title like *duktōr* 'medical doctor; university professor' or a less formal title like the addressee's first name would be avoided. Teknonyms, as they are applied to individuals, do not always reflect the reality of the individual's situation. In Jordan, for example, they do reflect that reality, whereby the eldest son of the addressee corresponds to the proper name in the formula. Terms such as *ʾabū yazīd* and *ʾumm yazīd* will be employed to address a couple whose eldest son's given name is Yazid. Alternatively, they may be used in anticipation of a son for an addressee who is still unmarried or who is married but only has daughters. Furthermore, the same formula is heavily present in Jordanian family names, where the second item in the teknonym is a common rather than proper noun, e.g. *ʾabū l-baṣal* 'father of onions', *ʾabū l-bandōra* 'father of tomatoes', *ʾabū r-rūz* 'father of rice', *ʾabū ḏ-ḏahab* 'father of gold'. Teknonyms can also be used relationally as praise formulas in expressions like *ʾabū l-karam* 'father of generosity' and *ʾumm al-mōḍa* 'mother of vogue' or as condemnation formulas as in *ʾabū l-muʾāmarāt* 'father of conspiracies' and *ʾumm al-mašākil* 'mother of troubles'. In both cases (praise or condemnation), it serves as an intensifier adding the meaning of 'very' or 'extremely'.

The other teknonym *ibn/bin* + proper name is still frequently used in personal names in North African Arab countries and Arabian Gulf states, e.g. *ʾAḥmad bin Billa* 'Ben Bellah [the first Algerian president after independence]', *Zēn il-ʿābidīn bin ʿAlī* 'Zeinilabidin [Tunisian president]', and *Qābūs bin Saʿīd* 'Qabus [sultan of Oman]'. A common variant of this teknonym in North Africa is *wild* 'son' + proper name, e.g. *Muʿāwiyah wild Tāyiʿ* 'Ould Tayi [ex-Mauritanian president]'. These teknonyms are only rarely used in the Levant and Egypt, except in the case of members of the ruling family in Jordan.

Of particular interest is the employment of analogous teknonyms featuring *ibn/bint/wild/ʾaxū* + common noun in Arabic imprecatives, considered by some a hallmark of Arabic swearwords. Examples often include obscenities such as *ibn iš-šarmūṭa* 'son of a prostitute', *bint il-manyūka* 'daughter of the fucked [fem.], *ʾibn/wild il-harām* 'son of the forbidden [i.e. bastard]', and *ʾaxu-l-qawwādi* 'brother of a pimp [fem.]'. Targets of these swearwords are the subject's kinswomen, who represent the important concept of *ʿirḍ*) 'honor' in the Arab culture (→ insults). The formulas, however, may also involve attributes that do not relate to honor, such as *ʾibn-il-ḥaywān* 'son of an animal' and *bint-il-majnūni* 'daughter of the crazy one [fem.]', among others. Furthermore, kinswomen often occur in popular imprecatives featuring *kus* 'pussy' + kinship term such as (*yilʿan*) *kus ʾummak* 'damn your mother's pussy!' and (*yilʿan*) *kus ʾuxtak* 'damn your sister's pussy!'. These imprecatives may exclude the obscene part but retain the kinship terms, as, for example, in the popular *yilʿan ʾummak* 'damn your mother!' and *yilʿan ʾabūk* 'damn your father!'.

5. RELATIONAL VS. ABSOLUTE KINSHIP TERMS

The above discussion has focused primarily on absolute kinship terms that denotationally designate family relations, such as *ʾab/ibna* 'father/daughter', *zawj/zawja* 'husband/wife', *ʾax-ʾuxt* 'brother/sister'. However, kinship terms in Standard Arabic as well as in different Arabic vernaculars are frequently used connotationally to maintain and enrich social interaction among both related and unrelated participants. These terms can be divided into distant and affectionate kinship terms, depending on their function (Levinson 1983; Farghal and Shakir 1994; Farghal 2002).

Distant kinship terms are commonly used to promote solidarity and politeness in casual summonses among strangers. The best way to get a stranger's attention on the street is to employ, among other honorifics, a kinship term. The summons *yā ʾax/yā ʾuxt* 'hey, brother/hey, sister [i.e., excuse me, sir/ma'am]!', for example, are frequently used, admittedly with some phonological variation, in the Levant, Egypt, and the Arabian Gulf. Other vernacular formulas include *yā xāl/xāla* 'hey, maternal uncle/maternal aunt!', *yā ʿamm/ʿammi* 'hey, paternal uncle/paternal aunt!', *yā jiddō/jiddi* 'hey, grandfather/grandmother!', *ʿammō* 'hey, paternal uncle!', *yā garāba* 'hey, relative!'. These and other distant kinship terms may vary from one Arabic vernacular to another and may be subject to subtle pragmatic constraints. Inter-regionally, for example, the summons formula *yalaxū* 'hey, brother' is frequently employed in Arabian Gulf vernaculars but may not be heard in urbanized areas in the Levant or Egypt, where the formula *yā ʾax* 'hey, brother!' would be employed. Similarly, age or social class constraints within the same region may instigate variation. For instance, *ʿammō* is a summons kinship term used by children and urbanized female youths when addressing adult male strangers in Levantine Arabic, but may not be used by a male youth. Rural female youths would employ *xayyō* 'hey, brother!' rather than *ʿammō* in this context, at least in Jordanian and Palestinian Arabic. Distant kinship terms often interact with Arabic greetings when addressing strangers in expressions such as *marḥaba yā garāba* 'hi, relative!' (Jordanian Arabic), *ʾizzayyak ya ʿamm* 'how are you, paternal uncle?' (Egyptian Arabic), and *šlōnak yalaxū* 'how are you, brother?' (Arabian Gulf or Bedouin Arabic).

Affectionate kinship terms, on the other hand, are usually employed among relatives and close friends to show intimacy. One of the main resources here is the use of morphology to indicate intimacy. Expressions such as *xayyō* 'brother', *xayyi* 'sister', *ʿammō* 'paternal uncle', *xālō* 'maternal uncle', *ʾibnayyī* 'my son', *ʾibnaytī* 'my daughter' are often used for this purpose in the Levant. Another important and common resource of intimacy manifests itself in the reversal of absolute kinship relations in Arabic, viz., *māma/yamma/yumma* 'mother' and *bāba/yāba/yūba* 'father' are overwhelmingly used by Arab mothers and fathers in addressing their children (Farghal and Shakir 1994; Rieschild 1998). Differences among the alternates are due to regional and/or social variation in Arabic vernaculars. The absolute kinship relation reversal is applicable to other kinship terms like *ʿammō* 'paternal uncle', *ʿammi* 'paternal aunt', *xālō* 'maternal uncle', *xāla* 'maternal aunt', *jiddō* 'grandfather', and *jiddi* 'grandmother'. This metaphorical application of kinship terms may obliterate natural or traditional social roles. It is customary, for example, in Egypt for educated husbands to address their wives by the kinship term *māma* 'mother' and for their wives to use *bāba* 'father' in addressing them. In addition to showing intimacy toward each other, this may be interpreted as a strategy to avoid the use of first names, which may sound a little direct and unaffectionate in this context. The avoidance strategy in the Levant vernaculars is not realized by twisting social roles but rather by using teknonyms such as *ʾabū ʾaḥmad* or *ʾumm ʾaḥmad* in husband-wife interaction as well as other non-kinship terms of endearment such as *ḥabibtī* 'my love [fem.]', *rūḥī* 'my soul', and *ʿumrī* 'my age/life'.

Bibliographical references

Farber, Bernard 1968. *Comparative kinship systems.* New York: John Wiley and Sons.

Farghal, Mohammed. 2002. "Situational and discoursal social honorifics in Jordan: An empirical study". *International Journal of the Sociology of Language* 158.163–181.

—— and Abdullah Shakir. 1994. "Kin terms and titles of address as relational social honorifics in Jordanian Arabic. *Anthropological Linguistics* 36:2.240–253.

Goody, Jack. 1969. *Comparative studies in kinship.* London: Routledge Kegan Paul.

Ḥuṣām ad-Dīn, K. Zaki. 1990. *al-Qarāba* [Kinship]. Cairo: Anglo-Egyptian Library.

Levinson, Steven 1983. *Pragmatics.* Cambridge: Cambridge University Press.

Rieschild, Robertson V. 1998. "Lebanese Arabic reverse role vocatives". *Anthropological Linguistics* 40:4.617–641.

Smith, W. Robertson. 1990. *Kinship and marriage in early Arabia*, ed. Stanley A. Cook. London: Darf Publishers.

Yassin, M. Aziz F. 1978. "Personal names of address in Kuwaiti Arabic". *Anthropological Linguistics* 20:2.53–63.

Mohammed Farghal (Kuwait University)

Ki-Nubi

1. General

About 25,000 (Ki-)Nubi speakers live scattered throughout the towns of Uganda and Kenya. Nubi is essentially a spoken language. Written literature hardly exists. Nubi, used as a lingua franca in the West Nile district in northern Uganda, has often been called an Arabic creole. (→ Creole Arabic).

The Nubi language most likely originated in the southern Sudan. Owens (1996) suggests that before 1820 a pidgin Arabic was already in use in sub-Saharan Africa, also called the Sudan. The events that fostered the development of Nubi, however, began around 1820 when Arabic speakers from the north moved southward for military operations and trading activities. Black Sudanese were taken as slaves, or they entered the military and trading camps deliberately. Around 1885, the trading and military activities came to an abrupt end as a result of the Mahdist revolt. Emin Pasha, governor of Equatoria province, withdrew to the Lake Albert area in present-day Uganda with his mixed Arab-black Sudanese troops. After a three-year period of seclusion, they left for the coast. However, Selîm Bey and Faḍl al-Mawlā were left behind with approximately 900 armed men and 10,300 followers, and were joined by many native Lendu and Lur. The group was met in 1891 by Captain Lugard, representative of the Imperial British East Africa Company, and later incorporated into the King's African Rifles. The group was divided into garrisons and scattered thoughout present-day Uganda and Kenya. The downfall of the Ugandan Nubi came soon after. Idi Amin brought them a period of revival, but in 1979 Amin fled and the Nubi were forced to seek exile in the surrounding countries. From 1986 onward, under Museveni, the Nubi gradually began to return and resettled throughout Uganda.

Presumably the Arabic lingua franca used for commercial purposes in the Sudan before 1820 found its way initially to the military training camps in and around Aswān, and later on in the southern Sudan through southward movements of the military and through the activities of merchants. The high-status Arabic-speaking officers and traders probably used a simplified Arabic when communicating with their southern Sudanese subordinates. The black slaves were thus addressed directly with a foreigner-talk variety of Arabic. Through contacts with the Arabic-speaking population they must have picked up some knowledge of Arabic native speech as well, however impeded they were by limited access to Arabic and by processes of imperfect language learning. The pidgin Arabic may have become a symbol of group membership for its speakers, differentiating them from their own tribal background and from their Arabic-speaking superiors.

By 1888, when Emin left for the coast, the pidgin Arabic had already evolved into a stable pidgin. It then experienced extensive input from speakers of local languages from the Lake Albert area, especially Lur and Lendu. It must still have taken many years before the number of newborn children in the group was large enough to bring about structural nativization or creolization of the language. By the time the groups were separated and the Nubi settlement in Nairobi was estab-

	bilabial	labio dental	dental	alveolar	post-alveolar	palatal	velar	uvular	pharyngeal	glottal
plosive	p b			t d			k g	(q)		(ʔ)
nasal	m	(-)		n		ɲ	(ŋ)			
trill				r						
flap				(ɾ)						
fricative		f v	(θ) (ð)	s z	ʃ		(x)		(ħ)	h
affricate					tʃ dʒ					
approx	w					j				
lateral approx				l						

[ʃ] = *š*; [tʃ] = *č*; [dʒ] = *j*; [ɲ] = *ny*; [w] = *w* or *u*; [j] = *y* or *i*; [θ] = *ṯ*; [ð] = *ḏ*; [x] = *x*, [ħ] = *ḥ*

lished in 1902, creolization must have taken place on a large-enough scale to explain the lack of major structural differences between the regional varieties. These were only affected marginally by substrate and adstrate influences. The extensive contacts between the Nubi people who live scattered nowadays in the larger towns of Kenya and Uganda, especially after 1979, when many Nubi went into exile, have affected the Nubi regional varieties and reduced their differences.

Owens (1977) and Heine (1982) describe the Nubi of Nairobi, while Wellens (2005) gives a detailed description of the Ugandan Nubi, including history and some texts. The latter is available on the Internet.

2. Linguistic description

2.1 *Phonology*

2.1.1 Inventory

2.1.1.1 Consonants

The consonants in parentheses have a marginal status and occur in Arabic and English loanwords or as the result of phonological processes. Nasal compounds may occur in borrowings from Bantu languages.

All consonants may occur in word-initial position. The consonants *t, d, k, g, ny, v, z, č, j*, and *h* do not normally occur in word-final position, except for a few words, like the nouns *ˈbit* and *ˈbint*, both meaning ‘girl’.

The consonants *s* – *š* and *y* – *z* are subjected to phonemic variation. The consonants *š* and *z* occur in the speech of older and/or northern Ugandan speakers, whereas *s* and *j* are restricted to young and/or southern speakers. The consonants *q, ṯ, ḏ, ḥ*, and *x*, may be used in Islamic expressions, or by speakers who have some knowledge of Arabic, alternatives for the more common Nubi *k, t, d*, and *h*.

2.1.1.2 Vowels

Nubi has five vowel phonemes: ***i*** ***u***
e ***o***
a

Allophonic variants: *a* [a; aː; ɐ; ɐː; ɑ; ɑː]
e [e; eː; ɛ; ɛː; æ; æː; ə; ɜ]
i [i; iː; ə]
o [o; oː; ɔ; ɔː]
u [u; uː; ə]

Vowel length is neither lexically nor grammatically distinctive in Nubi. Heine and Owens, however, mention at least one minimal pair each for Kenyan Nubi: *ˈbara* ‘outside [adverb])’ – *ˈbaara* ‘the outside [noun]’ (Heine 1982:25), *ˈsara* ‘to herd cattle’ – *ˈsaara* ‘to bewitch’ (Owens 1985:234). In Ugandan Nubi, this distinction is not made for the above words or for others. There is, however, a general tendency for vowels to be realized long in stressed syllables, while short vowels tend to occur in unstressed syllables, for instance in *ˈbasala* [ˈbaˑsala] ‘onion(s)’ and *biˈniya* [biˈniːja] ‘girl; daughter’.

Usually, front vowels do not co-occur with back vowels morpheme-internally (vowel harmony). The final vowel is exempted from this general tendency, as in *ˈberedu* ‘bath’. The

Table 1. Stress and tone in verbs (H = high tone, L = low tone; bold characters indicate stress)

	trisyllabic (**a**) '*kasulu* 'to wash'	disyllabic (**b**) '*fata* 'to open'	monosyllabic (**c**) '*so* 'to do'
bare verb	**H**LL	**H**L	**L**
gerund	L**H**L	**H**L	**L**
infinitive	H**H**L	H**H**	**H**
passive	LL**H**	L**H**	**H**
	(d) *ni'situ* 'to forget'	(e) *we'di* 'to give'	
bare verb	L**H**L	L**H**	
gerund	L**H**L	L**H**	
infinitive	H**H**L	H**H**	
passive	LL**H**	L**H**	
	(f) *fata'ran* 'to be' tired'		
bare verb	LL**H**		
gerund	LL**H**		
infinitive	HH**H**		
passive	LL**H**		

vowel *a* may co-occur with both groups of vowels, as in *anka'buti* 'spider'.

2.1.1.3 Syllable

There is a tendency toward a CV structure. However, syllables of the following types may occur:

V	'*ju-**a*** 'house', *ju-*'***a*** 'houses'
VC	'***am**-suku*, 'to grab, take'
C	***m***-'*ze* 'old man', *kele-*'***m*** 'it was said'
CV	'*ka-su-ru* 'to break'
CVC	*li-*'***fil*** 'elephant'
CVCC	'***bint*** 'girl'
CCVC	'***sten*** 'to wait for'

The last two types are far less common than the others. Disyllabic and trisyllabic words are more frequent than monosyllables. Words with more than three syllables are rare.

2.1.1.4 Stress, pitch, and tone

Stress is generally confined to one of the last three syllables in the word. There is a relation between stress and vowel length in Nubi. Vowels in stressed syllables are mainly articulated long, whereas vowels in unstressed syllables are generally short. Stress is also marked by loudness and high pitch. Word stress in Nubi largely depends on the stress patterns in the source languages of the Nubi words. The majority of words originally derive from the Arabic lexifier language. Although the words are subjected to phonemic changes, stress is retained, e.g. in Nubi *ge'ri* 'near(by)' from Sudanese Arabic *ga'rīb*. Stress is lexically distinctive in a few cases, as in the negator '*ma* vs. *ma* 'with; and'.

Stress is associated with grammatical meaning in some domains of grammar. Linked to vowel length and pitch, it distinguishes the predicatively used singular demonstrative from the one in attributive position: '*wede* and *we'de* 'this', respectively. In plural formation, stress is shifted to the final syllable, as in '*jua* 'house' >*ju'a*, '*bagara* 'cow' >*baga'ra*. Stress shift in verbs is shown in Table 1. The unmarked verb usually takes stress on the first syllable, with the exception of verbs in (d), (e), and (f). To form a gerund, stress is shifted to the penultimate syllable, while the passive takes stress on the last syllable. Tonal contrasts are also involved, as shown in the table. Pitch behaves independently of stress in marking verbal forms.

2.1.2 Phonotactics

Nubi nasal consonants tend to assimilate toward the articulation place of the following consonant, as in *kele'm neita* > [kɛlɛ'n'neˑta] 'you were told' and '*jengis* ['dʒɛngis] 'like'.

Consonants *k, d, h*, and *l* may be palatalized and are realized as *y* in the vicinity of front vowels, as in *la'kin* > *la'yin* 'but'. Voiced consonants are generally devoiced in prepausal position and before voiceless consonants, as

Table 2. Personal pronouns

person	singular		plural		
3rd	'*uo* (*'owo*)	‘he, she, it’	'*umon*	(*'omon*)	‘they’
2nd	'*ita* (*'ta*)	‘you’	'*itokum*	(*'itakum, 'tokum, 'tom*)	‘you [pl.]’
1st	'*ana* (*'an*)	T	'*ina*		‘we’

in '*kalabtu* > ['kaˑlaptu] ‘to mix’. Voiceless consonants may become voiced if followed by a voiced plosive, as in '*katif* '*buku* > [kaˑtiv 'buˑku] ‘to write a book’.

Regressive vowel assimilation is a common phenomenon in contemporary Nubi, especially in southern Uganda:

i/a > *o/u* before back vowels/bilabials, as in *fi 'bele* > *fu 'bele* ‘in the country’;
u/a > *e/i* before front vowels/*y*, as in *man'dil* > *men'dil* ‘handkerchief’;
r/l may cause an adjacent *a* to change into *e*, as in *tara'biya* > *tere'biya* ‘custom(s)’.

Unstressed *i* and *u* may be realized as a centralized vowel, as in *fi 'sokol* > [fə 'sɔkɔl] ‘in something’.

2.1.3 Morphophonology

Geminates are not common in Nubi; '*tenna* (an allomorphic variant of '*tenna* ‘our’), '*Allah* ‘God’, and '*yalla* ‘well’, ‘okay’ are probably the only instances. However, consonants can become geminated after processes of vowel deletion and/or consonant assimilation, as illustrated in *me'dida* > '*medda* ['mɛdda] ‘porridge’. The doubled consonant may, however, be degeminated in allegro forms, as in '*meda* ['mɛda].

Unstressed vowels are often elided in allegro forms. Deleted vowels are marked by round brackets, as in '*gez(i)ma* ‘shoe’, *(a)nka'buti* ‘spider’. As a result, the number of syllables in the word or word phrase may be reduced. Vowel elision occurs especially between homorganic consonants, as in '*badul(u) le'bis* ‘to change clothing’. Vowels, especially unstressed ones, may be elided before other vowels (fusion). In some cases, both vowels are absorbed and emerge as a new single vowel. This may occur across word boundaries, leading to new syllabic structures, as in '*ana li'go 'ita* > '*ana li'g-eta* ‘I met you’. Glide loss from the coda may take place in allegro forms, as in '*youm* > '*yom* ‘day’ and '*leil* > '*lel* ‘night’.

Vowels are added word-finally to avoid closed syllables, and particularly to avoid monosyllabic words. This is especially common in prepausal position. If the final consonant is alveolar and/or the preceding vowel is a front one, then *i* is attached, as shown in *ke'bir* > *ke'biri* ‘big’, *a'nas* > *a'nasi* ‘people’. However, if the word-final consonant is nonalveolar and/or the preceding vowel is a back vowel, then the added vowel is *u*, as in '*num* > '*numu* to sleep’, *a'jol* > *a'jolu* ‘person’.

2.2 *Morphology*

2.2.1 Pronouns

The Nubi pronominal system does not make a distinction for gender.

2.2.1.1 Personal pronouns

Nubi does not have a set of pronominal suffixes (verbal, prepositional) but rather uses the independent pronoun. Inanimates are normally not expressed pronominally:

'*itokum*	*bi-*	'*sten*	*helicopter.*
PRN2PL	FUT-	wait for	helicopter
bi-	'*tala*	*min*	*En'tebbe.*
FUT-	leave	from	Entebbe

‘You [pl.] will wait for the helicopter. [It] will leave from Entebbe’

2.1.1.2 Possessive pronouns and adjectives

Table 3. Possessive pronouns and adjectives

	singular		plural	
person				
3rd	'*to*	‘his, her, its’	'*toumon*	‘their’
2nd	'*taki*	‘your’	'*takum*	‘your [pl.]’
1st	'*tai/ta'yi*	‘my’	'*tena*	‘our’

2.2.1.3 Demonstratives

The core element of the Nubi demonstrative is '*de*. Plural is marked by '*dol*. The proximal and distal aspects are indicated by '*in* ‘here’ and '*na*' ‘there’.

Table 4. Demonstratives

	proximal	distal
singular	*(¸u)we'de* (ATTR) ~ (*u)'wede* (PRED) *'de* (ATTR ~ PRED)	*'na'de*
plural	*'dol'de, do'lin'de*	*'na'de, 'na 'dol'de*

The proximal singular demonstrative is composed of the 3rd person pronoun *'uo* and *'de*, interpreted as one form and subjected to some minor phonological changes (fronting of *o)*: *'uwe'de*. The independent predicative demonstrative is subjected to a stress shift to the penultimate syllable: *u'wede*. The particle *'de*, which is formally similar to demonstrative *'de*, is optionally attached to convey a certain emphasis.

2.2.2 Adverbs

Nubi adverbs often consist of a preposition or the genitive marker *ta* followed by a noun, like *ta 'tab* 'problematically'. The emphasizing element *'de* may be attached to the adverb, e.g. *'ase'de* 'now', even in reduplicated form, as in *'ase'de'de* 'now'.

2.2.3 Particles

2.2.3.1 Articles

The indefinite article *'wai* is derived from the numeral *'wai* 'one'. The article *'de* marks definiteness of singular and plural definite nouns.

2.2.3.2 Genitive marker

The genitive marker is *ta*, as in *'jua ta 'ragi 'de* 'the house of the man' and in the possessive pronouns (see 2.2.1.2).

2.2.3.3 Question words

The most common question words are *mu'nu?* 'who?', *su'nu?* 'what?', *(f)(u)'wen? ~ we'nu?* 'where?', *mi'ten?* 'when?', *ke'fin? ~ 'kef?* 'how?', *'le?* 'why?', *ma'lu?* 'why?', *ya'tu?* 'which?', 'what?', *'kam?* 'how many?'.

2.2.3.4 Prepositions

Nubi shows a variety of prepositions: single ones, like *fi* 'in', *min* 'from'; combined prepositions, such as *'ladi min* 'up to [spatial]', *fi gi'dam* 'in front of [spatial]'; and constructions like *fi 'batna (ta)* 'in the belly of', 'inside of' > 'inside', *fi 'te (ta)* 'at the bottom of' > 'under'.

2.2.3.5 Conjunctions

The conjuction *ma* 'with; and' is homophonous with the comitative marker. It generally joins together two noun phrases with similar functions. Occasionally, it joins together two sentences. The conjunction *wu, wa, u* 'and', on the other hand, generally joins together sentences, and only occasionally noun phrases.

2.2.3.6 Focus markers

The most frequent Nubi focus marker is *'ya,* which is homophonous with the vocative particle and with the conjunction *'ya* 'thus'.

2.2.4 Nouns

Gender is only indicated for persons and domestic animals. There either are separate names for masculine and feminine, like *bi'niya* 'girl' vs. *yo'wele* 'boy', or gender is expressed by juxtaposing *'marya* or *bi'niya* and *'ragi* or *yo'wele* respectively, as shown in *a'ku bi'niya* 'sister' vs. *a'ku yo'wele* 'brother'.

Plurals are formed in many ways:

i. by a shift of stress toward the final syllable. As a consequence of the heavy stress, the pitch on the last syllable becomes high, as in *gi'dida* 'chicken', pl. *gidi'da, 'bele* 'country'. pl. *be'le*;
ii. by suppletion, as in *'marya* 'woman; wife'; *pl. nus'wan*;
iii. by ablaut, as in *ke'bir* 'director' pl. *ku'bar*;
iv. by suffixation of *-ˀ(y)a*, as in *’seder* 'tree', pl. *sede'ra; -'iya*, as in *'asker* 'soldier', pl. *aske'riya; -'in*, as in *'tajir* 'rich person' pl. *taji'rin; -'an*, as in *'ter* 'bird', pl. *'te'ran; -'na*, as in *'sokol* 'thing', pl. *sokol'na; -'ka,* as in *nyere'ku* 'child', pl. *nyereku'ka*; and *-'u*, as in *'bab* 'door', pl. *ba'bu*. The word stress is shifted toward the suffix;
v. by a combination of plural markers, as in *ke'bir* 'director', pl. *ku'bar ~ kubari'na*.

Many words have more than one plural form, like *nyere'ku* 'child', pl. *nyereku'ka ~ 'yal - ya'la*. Nubi adjectives are not marked for gender, but they may be marked for number by suppletion, as in *sa'kar* 'small', pl. *du'ga ~ duga'ga*; by ablaut, as in *to'wil* 'long', pl. *'tu'wal*; by addition of a stressed suffix *-'in*, as in *a'sas* 'beautiful', pl. *asa'sin; -'ya,* as in *a'jusi* 'old', pl. *ajusi'ya; -'iya*, as in *fi'lan* 'certain', pl.

fila'niya. Like nouns, adjectives may combine plural markings, and they may have more than one plural form. Not all adjectives, however, have plural forms.

2.2.5 Numerals

The Nubi numeral system is a decimal one. Cardinals:

'wai 'one'	*i'dašar* 'eleven'	
ti'nin ~ ti'nen 'two'	*it'našar* 'twelve'	*iši'rin* 'twenty'
ta'lata 'three'	*tala'tašar* 'thirteen'	*tele'tin* 'thirty'
'arba ~ 'aruba 'four'	*arba'tašar* 'fourteen'	*ar'bein ~ arbe'yin* 'forty'
'kamsa 'five'	*kam(i)s-'tašar*	*kam'sin* 'fifty' 'fifteen'
'sita 'six'	*si'tašar* 'sixteen'	*si'tin* 'sixty'
'saba 'seven'	*saba'tašar* 'seventeen'	*se'bein ~ sebe'yin* 'seventy'
ta'maniya 'eight'	*taman'tašar* 'eighteen'	*tama'nin* *'eighty'*
'tisa 'nine'	*tisa'tašar* 'nineteen'	*ti'sein ~ tise'yin* 'ninety'
'ašara 'ten'		*'mia* 'hundred'

After 'twenty', the cardinal numerals are composed of the numeral followed by the tens and optionally linked together by *u, wu,* or *wa* 'and', as in *ta'lata w(u) ar'bein* 'forty-three'. An exception is *'wai* 'one', which is realized as *'waid*. It is joined with *iši'rin* by *u/i: 'waid (i) iši'rin* 'twenty-one'. Plurals of hundreds: numeral + *'mia,* as in *ta'lata 'mia* 'three hundred', with the exception of *mi'ten* 'two hundred'. Plurals of thousands are formed the other way around: *'elf* 'thousand' + numeral. Thousands and hundreds are optionally linked by *u, wu,* or *wa*; hundreds and tens are not, as in *'elf'arba (wu) 'kamsa 'mia 'sita (wu) ti'sein* 'four thousand five hundred and ninety-six'.

Ordinals are formed by the genitive particle *ta* and the cardinal. The ordinal 'first' is *ta aw'lan* or *ta 'wai.*

2.2.6 Verbs

The core element of the Nubi verbal system is the unmarked verb form (Ø). Verbs are not inflected morphologically, except for the passive and the gerundival verb forms.

2.2.6.1 Verb-final *u*

Most Nubi verbs end in a vowel, whether *-i, -e, -a, -o,* or *-u*. This feature corresponds to the Nubi tendency toward CV syllables. In Ugandan and Kenyan Nubi, about 57 percent and 45 percent of the verbs respectively end in *-u*.

'itokum	*'sulu*	*nyere'ku*	*'de*
PRON 2PL	take-Ø	child	DEF

'You took the child'

The function of *-u* is a topic for discussion. Some regard it as a verbal marker while others go much further, considering it a marker of high transitivity of the clause.

2.2.6.2 Passive, stative passive, and verb nominalizations

The passive is formed by a stress shift toward the final syllable, together with high pitch on that syllable. Passive monosyllabic verbs are marked by a high tone (see Table 1). Nubi has several verbs that can have both transitive and intransitive meanings without a change of form, like *'fata* 'to open' and 'to open something'.

There are two types of verb nominalization in Nubi. The stress pattern of the first type, called 'infinitive' corresponds to the stress pattern of the simple verb form. However, tone on the first and second syllables is high, irrespective of stress (see Table 1). Shifting the stress to the syllable preceding the last consonant forms the second type, called 'gerund'. The stressed syllable has higher pitch than the unstressed syllables (see Table 1). There are a few verbs where final *-u* is turned into *-a* in the gerund form, in addition to the usual stress changes, as in *'karabu* 'to destroy'—*ka'raba* 'destroying, destruction'.

2.3 *Syntax*

2.3.1 Noun phrase

2.3.1.1 Expression of number and gender

Nubi does not obligatorily mark number in the noun itself. The lower the noun is situated on the scheme of animacy hierarchy in Table 5, the less likely it is to be marked overtly on the noun for plurality:

Table 5. Number marking in the noun according to animacy hierarchy

Human beings: Kin, tribe, and/or gender terms>
Human beings in reference to their activities>
Domestic animals>
Non-domestic animals>
Inanimates, associated with the house and with housekeeping / *ka'lam, 'sokol* 'thing'
> Other inanimates

If not marked in the noun, plurality is inferred from the context, expressed periphrastically with separate quantifiers, such as *mi'lan, 'zaidi, ke'tir* 'many', numerals, and/or via the use of plural demonstratives. Adjectives expressing properties of human beings or animates are more frequently marked for number than others.

2.3.1.2 Definiteness and indefiniteness

A three-way division can be posited for the article use (Table 6; after Givón 1984, 1990).

Table 6. Three-way division of article use

	definite	indefinite
referential	*'de*	*'wai*
nonreferential		bare noun (Ø)

It is, however, possible for definite, referential noun phrases to appear without any overt marking because of their thematic centrality, and in certain contexts, such as nonindividuated ones, the unmarked noun phrase may occur instead of the indefinite article.

2.3.1.3 Possessive constructions

Possession is expressed by an analytic construction that binds the possessed to the possessor by means of the genitive marker *ta:* Possessed *ta* Possessor, e.g. *'kalwa ta 'šeik 'Musa* 'the religious school of Sheikh Musa'. Another possibility is to juxtapose possessed and possessor, e.g. *ba'kan be'redu* 'place used for bathing; bathroom'. In general, the N N-construction marks inalienable possession, which applies to kinship terms, body parts, some other part-whole relations, and goal. Alienable possessive relationships, such as typical owner-possessed relations, class, time, and location, are expressed by the possessed *ta* possessor-construction.

2.3.1.4 Quantifiers

The class of non-numeral quantifiers is small and includes *mi'lan* 'many, a lot of', *ke'tir* 'many, a lot of', *'sia* 'few, some', *'kulu* 'all of', *'kila* 'every'. Except for *'kulu* 'all', all quantifiers typically modify indefinite, often nonreferential nouns.

2.3.1.5 Numeral phrase

The Nubi numeral follows the noun. Nubi numerals do not agree with their head nouns. Similarly, Nubi head nouns have no obligatory agreement with numerals, as in *ku'baya ti'nin* 'two cups'.

2.3.1.6 Elative constructions

Comparison is expressed mainly by the preposition *'futu*, derived from the verb *'futu* 'to pass', and infrequently by the preposition *min.* The compared quality may be expressed by an adjective, quantifier, or noun and is often followed by a non-numeral quantifier or an adverb of degree.

'bei	*ta*	*lese'ri*	*'hari*	*'futu*
price	GEN	maize	high	in comparison with

'bei	*ta*	*'gwanda*
price	GEN	cassava

'The price of maize is high in comparison with the price of cassava'

An alternative is a clause with the verb *'futu/ 's(h)inda* 'to surpass': 'X (sur)passes Y, regarding quality Z'.

The superlative is realized with adverbs of degree following the adjective, as illustrated in:

'bele	*ke'bir*	*'sei-'sei*	*'de*
country	big	very	EMPH

'a very big country'

Superlative meaning may also be expressed by a relative clause (see 2.3.1.7) and by reduplication of adjectives.

2.3.1.7 Relative clauses

Nubi relative clauses consist of the relative marker (*'ali, a'li, al, 'abu, a'bu*, or *ab*) and the subordinate clause, and are placed directly after their head noun. Nubi distinguishes between nonrestrictive and restrictive relative clauses. The head noun of a nonrestrictive relative clause is optionally marked by a definite article. Often, nonrestrictive relative clauses are modified by the definite article or a demonstrative, as in:

'ahadi	*'na'de*	*al*	*ku'tu*
agreement	DEM DIS	REL	put-PASS-Ø

'nena	*'na'de*
for + PRN 1PL	DEM DIS

'that agreement, which was made for us'

Restrictive relative clauses are normally not marked by a determining element, and their head nouns are either definite or indefinite.

A resumptive pronoun is optional in the subordinate clause in object position and is obligatory in the prepositional phrase and possessive phrase. The co-referent of relative clauses related to place, time, manner, and purpose is indicated by *'fogo,* either compulsory (place, purpose) or optional (time, manner).

2.3.1.8 The focus marker *'ya*

The focus marker *'ya* functions as a contrastive device or as a highlighter of new or asserted information.

gi- 'ben	*je*	*'uo*
PROG-seem	like	PRN 3SG

'ya	*ma'ma*	*'taki*
FOC	mother	PRN POSS 2SG

'It looked as if **she** were your mother'

The focus marker may refer to any sentence constituent except the verb. It may, however, focus on the verbal predicate.

2.3.2 Verbal phrase

No typical analytic or synthetic devices exist to mark the direct or indirect object on the verb.

2.3.3 Verbal tense and aspect

2.3.3.1 Stative vs. nonstative verbs

Nubi distinguishes between stative and nonstative or dynamic verbs. Stative verbs include verbs expressing feeling and emotions, like *'aju* 'to wish', 'to want', 'to like', 'to need'; verbs reflecting a mental activity, like *'fahamu* 'to understand'; verbs semantically pertaining to state, like *'gen/'gai* 'to sit', 'to stay'; and verbs of (non-)possession, like *'endi(s)(i)* 'to have'. Stative and nonstative verbs behave differently, as seen below.

2.3.3.2 The unmarked verb vs. the *gi-* prefix

The unmarked verb form of nonstative verbs marks punctual aspect and realis. As long as no other time marking is available, the unmarked form of nonstative verbs refers to the past. Nonpunctuality of nonstative verbs is marked by the prefix *gi-* and essentially refers to present events. Stative verbs do not normally take a marker to express nonpunctuality, except when indicating inchoativeness or denoting habitual or generic aspect, and with the verbs *'gen, 'gai,* 'to sit; to stay' and *'ben* 'to look like', 'to seem' in the speech of a limited group of Nubi speakers. However, the unmarked verb form may express any tense or aspect, including continuation and repetition, on the condition that these are marked by other means, such as adverbs and adverbial phrases, the context, or other marked verbs, as illustrated in:

yo'wele,	*'kila*	*min 'sub,*	*'uo*
boy	every	(in) morning	PRN 3SG

'kub	*'moyo*	*'te*	*te*	*'maua*
pour-Ø	water	under	GEN	flowers

'The boy, every morning, he pours water under the flowers'.

The progressive marker *gi-* is equally neutral as regards tense, but it always denotes nonpunctuality.

2.3.3.3 The future marker *bi-*

The verbal prefix *bi-* marks unrealized future events involving volition, or strong expectation of future events. It may also mark habitual actions irrespective of time. In this respect, *bi-* and *gi-* are interchangeable.

'kweis,	*'ana*	*ka'man*
fine	PRON 1SING	also

bi-	*'wonusu*	*'sia...*
FUT-	talk	a little bit

'Fine, I will also talk a little bit (...)'

2.3.3.4 The anterior marker *'kan*

The marker *'kan* indicates that the event or state took place before the time in focus and is no longer in existence. Followed by a nonstative verb, *'kan* may also denote perfective aspect. With stative verbs, *'kan* often marks modality. The three core-markers *gi-, bi-*, and *'kan* may be combined in the order ANT FUT PROG.

ka'lam	*al*	*gi'bel*	*'kan rasu'l*
thing	REL	before	ANT send-PASS-Ø

'nana	*'ana*
to + PRON 1SING	PRON 1SING

'so	*ka'la*
do-Ø	already

'The thing that was sent to me before, I have done [it]'

2.3.3.5 Other auxiliaries and markers

Intention is marked by the auxiliary *'rua* 'to go', *gi-'rua* V and *gi-'ja* V are typically used to express immediate future, *bi-'rua* V and *bi-'ja* V mark an uncertain and remote future. The auxiliaries *'gum* 'to get up' and *'ja* 'to come' express inception, *'arija* V marks repetition, and the auxiliaries *'gen/'gai* 'to stay', 'to sit', and *'fi* 'to be there' typically mark duration. The marker *ka'la(s)(i)* signals completion. With *gi*-marked and stative verbs, *ka'las* emphasizes the current relevance of states or processes, *'lisa* marks that an action or state is still in progress.

'ase'de	*ka'las*	*'ja*	*la'siya.*	*'gu-*
now	COMPL	come-Ø	evening	PROG-

rwa	*'ja*	*'sa*	*ta*	*'num.*
go	come-Ø	hour	GEN	sleep-GER

'Now, it has already become evening. It is going to become the time of sleeping?

2.3.3.6 The imperative

The singular imperative consists of the unmarked verb form, e.g. *'jib* 'bring!'. Exceptions are *'rua/'ro* from *'rua* 'go', *'ja/'tal* from *'ja* 'come!, and *'lib* from *'alabu* -play!'. To form the plural imperative, the subject suffixes *-kum, -'tokum, or -'takum* are fixed to the unmarked verb, e.g. *'gum-kum* 'wake up [pl.]!',*'lebis-'takum* 'get dressed [pl.]!', and *'aruf-'tokum* 'know [pl.]!'.

2.3.3.7 The subjunctive

Subjunctive mood is expressed by the marker *ke'de* or *'ke* introducing a verbal clause. The verb is generally unmarked.

2.3.3.8 The stative passive

The stative passive is formed by prefixing *ma-* to the unmarked verb form, e.g. *'kasuru* 'to break – *ma-'kasuru* '(to be) broken'. The stative passive behaves like a predicative adjective and expresses a state resulting from a completed action.

2.3.3.9 Negation

Nubi employs two different devices for negating sentences or clauses. The marker *'ma* may take any position in the sentence, but sentence-final position is most frequent. The marker *'maf(i)* occurs generally in sentence-final position. Double negation exists, though rarely. For the negative imperative, the unmarked verb is preceded by *'mata* and *'matakum* for the negative singular and plural imperative respectively, e.g. *'mata 'fata* 'do not open!' and *'matakum ’wonus* 'do [pl.] not talk!'. Occasionally, the negative imperative consists of the positive imperative form followed by *'ma*, e.g. *'sul 'ma* 'do not take [it]!'.

2.3.4 Word order

Nubi is essentially an SVO language, in both the main and subordinate clauses. The word order may be reversed with focus-attracting devices, such as left dislocation, Y-movement, and strategies for introducing new information. Questions do not normally have a deviant word order. The question words *mu'nu?* 'who?', *su'nu?* 'what?', *ya'tu?* 'which?', *'kam?* 'how many?', *ke'fin?, 'kef?* 'how?', and *mi'ten?* 'when?' remain in situ depending on their function in the sentence. The question word *(f)(u) 'wen?, we'nu?* 'where?' usually takes sentence-final positions; *'le?* and *ma'lu?* 'why?' may take any position.

Within the noun phrase; the head noun is preceded by *'kila* 'every', *'aya* 'any', or collective marker *'nas* and is followed by the indefinite article *'wai* or a pronominal possessor, the adjective and/or the numeral, which may exchange positions, the relative clause, the definite article *'de*, and/or the demonstrative. The quantifier *'kulu* 'all' closes the noun phrase.

a'ku	*'to*	*'wai*
brother	PRN POSS 3SG	NUM

ke'bir	*'de*	*'na'de*
big	DEF	DEM DIS

'that one eldest brother of his'

2.3.5 Agreement

Number agreement within the noun phrase and agreement across the predicate are optional. The verb is not marked for gender or number, except for the imperative, which may take a plural suffix.

2.3.6 Existential sentences

Existence is indicated by the marker *'fi* 'there is, there are.' The marker *'fi* may be analyzed as a stative verb with a reduced tense/aspect system, restricted to the *gi-*, *bi-*, and *'kan* markers. To negate, the negative marker *'ma* is prefixed: *'ma'fi*, also produced as *'mafi*, *'maf*, or even *'ma*. Definite subjects generally precede the existential marker. Indefinite subjects follow it. The negated existential marker *'ma'fi* is generally preceded by its definite subject.

3. LEXICON

Approximately 90 percent of Nubi vocabulary is of Arabic origin. Wellens (2005) argues that the Nubi vocabulary indicates Egyptian Arabic sources, apart from its general Sudanese origin. Owens (1985, 1996), on the other hand, points to a western Sudanese Arabic origin. Nubi took borrowings especially from Swahili, Luganda, and Nilotic languages such as Bari, Luo, and Alur.

BIBLIOGRAPHICAL REFERENCES

Givón, Talmy. 1984, 1990. *Syntax: A functional-typological introduction*. 2 vols. Amsterdam and Philadelphia: J. Benjamins.
Heine, Bernd. 1982. *The Nubi Language of Kibera, an Arabic creole: Grammatical sketch and vocabulary*. [= *Language and dialect atlas of Kenya*, III.] Berlin: Reimer.
Owens, Jonathan. 1977. *Aspects of Nubi grammar*. Ph.D. diss., London School of Oriental and African Studies.
——. 1985. "The origins of East African Nubi." *Anthropological Linguistics* 27. 229–271.
——. 1996. "Arabic-based pidgins and creoles". *Contact languages: A wider perspective*, ed. Sarah G. Thomason, 125–172. Amsterdam and Philadelphia: J. Benjamins.
Wellens, Ineke. 2005. *The Nubi language of Uganda: An Arabic creole in Africa*. Leiden: E.J. Brill.

INEKE WELLENS (Ghent, Belgium)

Kitap- Afrikaans → South Africa

Koine → Dialect Koine; Poetic Koine

Kufic

Kufic (Arabic *kūfī*) is an angular style of script used in early Islamic times for monumental inscriptions and *Qur'ān* manuscripts. Derived from Kūfa, the city in southern Iraq renowned for its erudition in early Islamic times, the term 'Kufic' was introduced to Western scholarship in the late 18th century by Jakob George Christian Adler (1756–1834), a Lutheran cleric from Schleswig who was charged with cataloguing the Qur'ānic material in the Royal Library at Copenhagen. The collection comprised a mere five fragments, and Adler grouped them all under the rubric 'Kufic', a term he had found in the writings of the 13th-century biographer Ibn Xallikān and the 14th-century lexicographer al-Fīrūzābādī. The term gained currency and became widely accepted.

In the late 1980s, however, François Déroche, the French expert on Qur'ānic paleography, objected to the term 'Kufic', noting that it encompassed a wide variety of hands and that the connection to Kūfa was misleading because not all of the inscriptions or manuscripts written in an angular script were connected with that city. He proposed, as an alternative, 'old (or early) Abbasid style', a term chosen to contrast with the newer rounded style of script that came to the fore in Arabic manuscripts during the 9th and 10th centuries. To account for the many variants of the angular style, Déroche divided his 'early Abbasid style' into six groups (lettered A–F), further subdivided into categories identified by Roman numerals and subcategories identified by the addition of a lowercase letter (e.g. B.I.a).

In many ways, Déroche's choice of name was equally unfortunate. As he himself noted, this angular style was already used before the Abbasid period and, furthermore, was not necessarily connected with or limited to the Abbasid caliphate. His divisions, moreover, may seem subjective. They also assume a standardization readily obtainable only with mechanical means of dissemination, and they overlook a calligrapher's ability to adapt a style to different materials, sizes, and shapes.

The term 'Kufic' (French *coufique*) therefore should be retained. It should be taken, however, not as the name of a specific script used at a particular time or place, but as a general rubric for the angular style used in early Islamic times to inscribe monuments and transcribe the *Qurʾān*. Using this broad definition, we can trace the use of the term in texts, describe the salient characteristics of this style, enumerate its various forms, and chart its demise.

1. TEXTUAL EVIDENCE

The term 'Kufic' appears already in the *Fihrist* by Ibn an-Nadīm (d. 380/990), our main source for much of the information about books and copyists in the early Islamic period. The 10th-century encyclopedist uses the term (*Fihrist* 12) to designate the script written by a group of *Qurʾān* copyists in the time of the Abbasid caliph al-Muʿtaṣim (r. 833–842): ʾAbū Juday and, following him, Ibn ʾUmm Šaybān, al-Masḥūr, ʾAbū Xamīra, Ibn Ḥumayra, and finally ʾAbū l-Faraj, a calligrapher and *Qurʾān* reader who was active in Ibn an-Nadīm's own day and a friend of Ibn Šanabūḏ, the Baghdadi scholar condemned for his variant readings of the *Qurʾān*. No examples of these calligraphers' hands are known, nor do we know how Ibn an-Nadīm distinguished Kufic from similar scripts like *muḥaqqaq* and *mašq*, which he mentions in his following paragraph as also used to transcribe the *Qurʾān* in Abbasid times.

There is at least one instance in which a reference to Kufic in a late medieval text can be matched with an actual specimen. An inventory dated 693/1293–1294 of the library of the Great Mosque of Qayrawān refers to a large-format seven-part *Qurʾān* manuscript written on blue (*ʾakḥal*) parchment in gold ink in Kufic script with silver chapter headings and verse counts. Already in 1956 ʾIbrāhīm Šabbūh matched the description to the famous Blue *Qurʾān* (see Fig. 1), a manuscript now dispersed but with at least 75 folios once in the Great Mosque at Qayrawān (Déroche 1992, no. 42). A codex in seven sections (*manzil*), it is transcribed in gold ink and silver incidentals on large sheets of parchment dyed dark blue. Many sheets are trimmed, but on average they measure 30 × 35 cm and probably once measured 31 × 41 cm. Each page contains fifteen lines of bold angular script notable for its blocky and elongated letters.

The term 'Kufic' became widely accepted in later times, particularly by Persian authors who applied it to the angular style used before the codification of the round scripts known as the 'Six Pens' (Arabic *al-ʾaqlām as-sitta*; Persian *šīš qalam*). According to these sources, Kufic had one-sixth round strokes and five-sixths straight. This description is given, for example, by the Safavid calligrapher and commentator on calligraphy Qāḍī ʾAḥmad (1959:53–54), who contrasts Kufic with the completely straight *maʿqilī*, said to be named from the Maʿqil Canal near Baṣra. These two terms were repeated by many authors in the prefaces to the albums of calligraphy and painting composed at this time. Dūst Muḥammad, in the preface to the album he prepared for the Safavid prince Bahram Mirza in 951/1544 (Thackston 2001:7), for example, credits ʿAlī ibn ʾAbī Ṭālib with perfecting the Kufic used until the time of the master Ibn Muqla (d. 328/940). So do Mīr Sayyid ʾAḥmad (Thackston 2001:24) and Muḥammad Muḥsin (Thackston 2001:35). The Mughal chronicler Abū l-Faḍl ʿAllāmī repeats the same information. The Mamluk author aṭ-Ṭayyibī compiled a treatise in 908/1503, dedicated to the Mamluk sultan Qānṣawh al-Ġawrī, on the kinds of script in the tradition of Ibn al-Bawwāb. His is the first album of calligraphic scripts to survive from the Arab lands. Aṭ-Ṭayyibī, like other Mamluk authors, neither discusses nor illustrates the angular Kufic style because Kufic was regarded as predating the round scripts.

2. THE CLASSIC KUFIC STYLE

Kufic is written with thick uniform strokes and few diagonals and curves. The *ʾalif* is a straight vertical stroke that ends with a hook or foot to the right, and the upper strokes of other letters, such as *ṭāʾ*, are similarly vertical. In contrast to this verticality, the letter bodies are horizontal, posed on a flat (though usually unruled) baseline and connected directly with each other without the notch or indentation used in other scripts. Letters are aligned linearly; connected letters are usually not raised, even when joining to *jīm* or other letters that are traditionally begun from the top in other scripts. In Kufic, diacritical marks are not usually added to distinguish homographs or letters of similar shapes, such as *bāʾ, tāʾ*, and *nūn*, although they can be indicated by thin slashes in cases where important to avoid ambiguity. Vowels are not written, and the script is one of the *scriptiones defectivae* in which *ʾalif* is omitted in certain words, such as *samāwāt* 'heavens' or *qāla* 'he said', or hortative phrases with *yā*. To compensate for the unwritten vowels, in some Kufic *Qurān* manuscripts vowels are indicated by red dots, one above the line for *fatḥa*, one below the line for *kasra*, and one following the letter for *ḍamma*.

The spacing in Kufic is also distinctive, as the spaces between words are the same width as the spaces between the letters within words. That is, calligraphers writing Kufic willfully abandoned the spaces between words. In manuscripts written in Kufic, furthermore, both the right and left edges of the written area are justified. The alignment along the left edge is accomplished by stretching out the connecting strokes between letters or the horizontal bodies of letters such as *dāl, kāf*, and *ṣād*. Words can also be broken between lines. Such spacing, especially when combined with the lack of pointing, served deliberately to slow down reading and enhance the oral and rhetorical aspects of the text.

Angular script was used for the earliest monumental inscriptions in Arabic, such as the six lines scratched on a rock near Ṭāʾif in the Ḥijāz that record the construction of a dam by the Umayyad caliph Muʿāwiya in 58/677–678 (Blair 2006, Fig. 3.3). The script there includes a diagonal upstroke on *ṭāʾ*, but a more rectilinear style was regularized in the 70s/690s with the reforms by the Umayyad caliph ʿAbd al-Malik, who made Arabic the language of the chancery and administration. Arabic legends in Kufic were introduced on coins, culminating in 77/697–698 with the appearance of dinars, or gold coins, that are entirely epigraphic (Blair 2006, Fig. 3.4). Letters are posed flat and descend only slightly from a uniform baseline, with medial *jīm* a short stroke that bisects the baseline diagonally. Several traits show that the inscriptions were designed by a calligrapher using pen and ink. The *ʾalif*, for example, has a bent right foot; isolated *bāʾ* opens with a short curved stroke; *dāl* has a small hairline stroke at the top; final *mīm* ends with a short horizontal tail; and final *yāʾ* swings to the right below the word.

Many of the same features can be seen even more clearly in the mosaic inscription ringing the interior of the Dome of the Rock in Jerusalem begun in 72/692 (see Fig. 2). The text, which contains the first dated evidence of the writing down of the *Qurʾān*, is executed in gold mosaic against a blue-green ground. Diagonal strokes for pointing accompany at least 92 letters. Like the inscriptions on coins, the text for this inscription was drawn up by a calligrapher. Final *yāʾ*, for example, not only swings to the right but also tapers to a point. A similar but even more calligraphic style was used for the hammered copper plaques that once adorned the lintels over the inner doors of the east and north entrances to the building.

In addition to monumental inscriptions, Kufic was standard for transcribing *Qurʾān* manuscripts. The style seems to have been reserved for scripture, as distinct from the rounded scripts that scribes had used since the beginning of Islam for correspondence and non-Qurʾānic texts. Of the 1,300 manuscripts or fragments written in Kufic, only two do not belong to *Qurʾān* manuscripts: a fragmentary genealogical work (Paris, Bibliothèque Nationale, ms. arabe 2047, 13 folios; Berlin, Staatsbibliothek, ms. Or. 379, 2 folios); and a copy of the Acts of the Apostles transcribed by the monk Mūsā ar-Rāhib, probably in 9th-century Palestine (Mount Sinai, Monastery of St. Catherine; Paris, Bibliothèque Nationale, ms. arabe 6725; Guesdon and Vernay-Nouri 2001, no. 13). These two exceptions seem anecdotal, and Kufic remains a hallmark of *Qurʾān* manuscripts.

In format, most Kufic *Qurʾān* manuscripts are horizontal (landscape), although a few vertical (portrait) ones are known. We do not know why this oblong format was chosen, perhaps to distinguish scripture from other texts and certainly to take advantage of the elongations possible in Kufic script. Kufic *Qurʾān* manuscripts are transcribed on parchment, usually in a brownish-black ink whose iron tannates eat into the support. A few are copied in gold ink, as in a codex with fifteen lines to the page (Istanbul, Nurosmaniye Library ms. 27; Déroche 1992, no. 41) or another with five lines to the page (Tunis, ms. Rutbi 198; Déroche 1992, no. 19). As with these two gold examples, the text in most Kufic *Qurʾān* manuscripts ranges from a smaller and more crowded script written with fifteen to seventeen lines per page to larger and more spacious scripts with only three to five lines per page. The Blue *Qurʾān* exemplifies the former; a good example of the latter is the ʾAmājūr *Qurʾān* (see Fig. 3), so called because it was endowed by ʾAmājūr, Abbasid governor of Damascus, to a foundation in the port of Tyre in 262/876 (Déroche 1990–1991). None of the surviving *Qurʾān* manuscripts or fragments in Kufic is signed or dated, and such an endowment notice (*waqfiyya*) provides a convenient *terminus ad quem*, although we do not know how much earlier such a manuscript might have been copied.

3. The elaboration of kufic

To judge from examples in dated media such as coins, tombstones, and *ṭirāz* (the inscribed textiles issued by state manufactories), over the course of the 9th and 10th centuries inscriptions in Kufic became taller with more elongated stems to the letters. As the upper part of the band became increasingly large, designers gradually developed new devices to fill the empty space and balance the bodies of the letters in the lower part of the band. In addition to beveling the ends of the letters and carrying the tails of descending letters like *rāʾ* upward in sweeping strokes, designers added decorative devices such as barbs and palmettes. In this way, foliated Kufic slowly evolved into a floriated style in which flowers, tendrils, and scrolls seem to grow from the final or medial forms of the letters. Floriated Kufic was in full bloom by the mid-10th century, and magnificent bands of floriated Kufic sculpted in stone became a hallmark of the Fatimids, the wealthy and sophisticated rulers of Egypt from 969 to 1171 C.E. Some of the finest examples are found on the façade of the Aqmar Mosque, erected on the main street of Cairo in 519/1125–1126 (see Fig. 4). Such bands in floriated Kufic were also used for *sūra* headings in Kufic *Qurʾān* manuscripts to distinguish supplementary information from revelation (e.g. Déroche 1992, no. 68). In floriated Kufic, the decoration sometimes threatens to obscure the writing, which then devolves into a meaningless design. This type of script, called pseudo-Kufic or *kufesque*, was popular from the 10th century to the 12th, particularly in the Byzantine lands (e.g. Blair 1998, Fig. 7.32).

In the eastern Islamic lands artists developed another decorative device – interlacing – to meet the same demand of filling the upper zone of a Kufic inscription. Whereas in the western Islamic lands elaboration of the letter stems had led from beveling to foliation and then to floriation, in the East the tendency toward elongation and distortion of horizontal letters led to internal modifications and superimposed ornament. Like floriation, interlacing had already reached sophisticated levels by the 11th century, as shown by the tomb tower erected at Radkan in 411/1021, or the minaret erected at Tirmidh in 423/1031–1032 (Blair 1998, Fig. 7.33). The plethora of decorative devices threatened to overwhelm the readability of the words, and so artists increasingly moved the decoration to the upper zone above the bodies and even the stems of the letters. This type of bordered Kufic was particularly popular in the eastern Islamic lands in the 11th and 12th centuries, as with the foundation inscription that once graced the madrasa founded by Niẓām al-Mulk at Khargird and now in the Iranian National Museum (Blair 1998, Fig. 7.34).

Bricklayers also took advantage of the angularity of Kufic to develop the script known as *maʿqilī* (square), *bannāʾī* 'builder's [technique]', or square Kufic. In this script, words, phrases, and even sentences are inscribed within a square. Perhaps the most exotic examples are the stucco panels covering the interior of the shrine built in the opening decades of the 14th century for Pīr-i Bakrān at Linjān near Isfahan, with the Ninety-Nine Names of God and the names of the Four Orthodox Caliphs and the Fourteen Immaculate Ones beloved by Twelver Shiʿites (ʿAlī, his daughter Fāṭima, and the Twelve Imams). The texts are readable mainly because they are so well known.

4. The devolution of Kufic

From the late 9th or early 10th century, copyists began to replace Kufic with round scripts. The first surviving example of a *Qurʾān* manuscript penned in a round hand is a tiny dispersed copy with notes in Persian saying that it was corrected by one ʾAḥmad ibn ʾAbī l-Qāsim al-Xayqānī in 292/905 (Blair 2006, Fig. 5.2). The round book hand used in this manuscript is frankly unattractive, with wiggly lines and awkward spacing. It was clearly an experimental copy, as calligraphers were trying to regularize and monumentalize the round hands they traditionally used for copying documents and secular texts in order to transform them into a new style suitable for transcribing scripture. This new style, characterized by distinct diagonals and a marked difference between thick and thin strokes, is sometimes dubbed eastern (or occasionally, western) or broken Kufic, misnomers implying that it developed from Kufic. Rather, this stylized script, called broken cursive or 'new style' by Déroche (1992, nos. 75–98), should be seen as a stylization of round scripts.

Broken cursive came to the fore in the 10th century as shown by two fine manuscripts: an autographed copy of Muḥammad ibn ʿAbd al-Jabbār an-Niffārī's mystical reflections, *Mawāfiq*, dated 344/955–956 (Dublin, Chester Beatty Library ms. 4000), and a *Qurʾān* manuscript copied by ʿAlī ibn Šaḏān ar-Rāzī in 361/972 (Blair 2006, Fig. 5.3). This script was also used for notes added to earlier manuscripts, as in the *waqfiyya* added at the top of each folio of the ʾAmājūr *Qurʾān* (see Fig. 3). From the 11th century, *Qurʾān* manuscripts in broken cursive were increasingly supplanted by copies in round scripts, notably the Six Pens, as in the well-known manuscript transcribed by Ibn al-Bawwāb at Baghdad in 391/1000–1001 (Chester Beatty Library ms. 1431).

From the late 11th century, round scripts replaced Kufic in monumental epigraphy. At first they were used for religious inscriptions, such as the Qurʾānic verses inscribed across the façade of the south dome erected in the Friday Mosque at Isfahan in 479–480/1086–1087. They were then adopted for historical texts. While this was a gradual change in the East, it was more abrupt in the central Islamic lands. The last Kufic inscription in Cairo, for example, is found on the mosque of ʾAmīr Ṭalāʾī ibn Ruzziq built in 555/1160. With the arrival of Saladin in Egypt a dozen years later, Kufic was swiftly replaced by a round → *nasx*. The advent of the Zangid Nūr ad-Dīn (r. 1146–1174 C.E.) heralded a similar transition in Syria.

There are many reasons why the angular Kufic style was supplanted by round scripts. One is the nature of the support used in manuscripts. ʿAlī ibn Šaḏān's manuscript dated 361/972 is not only the first surviving *Qurʾān* in broken cursive, but it is also the first copied on paper. The more fluid round scripts took better advantage of the smoothness of the new support. The change to round scripts also heralds a change in the status of copyists: As opposed to the anonymous Kufic *Qurʾān* manuscripts, those in round scripts are signed by calligraphers and eventually by illuminators as well. This change was also part of the new grammar of legibility, in which oral recitation increasingly gave way to private reading. The abandonment of Kufic for monumental epigraphy also coincided with new propagandistic politics. Rulers like Saladin and Nūr ad-Dīn Zangī did not invent round scripts, but they certainly exploited them to distinguish their works from their predecessors' and to establish distinctive dynastic styles.

In succeeding centuries Kufic became a conscious archaism. It continued to be used sporadically for headings and incidentals in *Qurʾān* manuscripts, particularly those made in the 15th century for the Timurids. It was also revived occasionally for historical texts. For the foundation inscriptions dated 839/1435–1436 on the minarets of the mosque of Gawhār Šād at Herat, for example, the designers plaited the stems to form a interlaced band between the smaller letter bodies and the decorative terminals. Such Kufic was, however, rare, perhaps because it was difficult to read (and still is: it is sometimes reproduced upside down). Square Kufic was popular for wall revetments of matte and glazed bricks, in which the shining sacred names and phrases reverberated visually like the verbal repetitions in a Sufi *ḏikr*. Square Kufic was also used for puzzles and rebuses and often juxtaposed to the flowing round scripts, as in the double frontispiece to a volume of religious texts penned by the noted Ottoman calligrapher Ahmad Karahisari for Sultan Süleyman around 1550 (Istanbul, Türk ve Islam Museum 1443, Blair 2006, fig. 11.8).

Kufic remains a symbol of tradition. In 1846 Renard designed type fonts in Kufic and so-called Qarmathian Kufic (the new style of broken cursive) for the Imprimerie Royale in Paris. Kufic is still recognized as one of the traditional scripts in the international calligraphy competitions sponsored by IRCICA, the Research Center for Islamic History, Art and Culture founded in Istanbul in 1980. The gridded basis of square Kufic makes it a favorite for computer graphics, as in the designs composed by the Syrian-born American designer Mamoun

Sakkal. In most cases, these Kufic texts are readable only because they are well known, and Kufic is primarily used not to convey information but to evoke the past.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Ibn an-Nadīm, *Fihrist* = ʾAbū l-Faraj Muḥammad ibn ʾIsḥāq Ibn an-Nadīm, *al-Fihrist*. Ed. and trans. Bayard Dodge, *The* Fihrist *of al-Nadīm: A tenth-century survey of Muslim culture*. New York and London: Columbia University Press, 1970.

Qāḍī ʾAḥmad = Qāḍī ʾAḥmad, *Gulistān-i hunār*. Trans. Vladimir Minorsky, *Calligraphers and painters: A treatise by Qāḍī Aḥmad, son of Mīr-Munshī (ca. A.H. 1015/A.D. 1606)*. Washington, D.C.: Freer Gallery of Art Occasional Papers, 1959.

Ṭayyibī, *Jāmiʿ* = Muḥammad ibn Ḥasan at-Ṭayyibī, *Jāmiʿ maḥāsin kitābat al-kuttāb*. Ed. Ṣalāḥ ad-Dīn al-Munajjid. Beirut: Dār al-Kutub al-Jadīda, 1962.

SECONDARY SOURCES

Blair, Sheila S. 1998. *Islamic inscriptions*. Edinburgh: Edinburgh University Press.

——. 2006. *Islamic calligraphy*. Edinburgh: Edinburgh University Press.

Déroche, François. 1990–1991. "The Qurʾān of Amājūr". *Manuscripts of the Middle East* 5.59–66.

——. 1992. *The Abbasid tradition: Qurʾans of the 8th to the 10th centuries A.D.* (= *Nasser D. Khaili Collection of Islamic Art*, 1.) London: Nour Foundation (in association with Azimuth Editions and Oxford University Press).

Grohmann, Adolf. 1957. "The origin and development of floriated Kufic". *Ars Orientalis* 2.183–214.

Guesdon, Marie-Geneviève and Annie Vernay-Nouri. 2001. *L'art du livre arabe*. Paris: Bibliothèque Nationale de France.

Šabbūh, ʾIbrāhīm. 1956. "Sijill qadīm li-maktaba Jāmiʿ al-Qayrawān". *Revue de l'Institut des Manuscrits Arabes* 2:2.339–372.

Tabbaa, Yasser. 1991. "The transformation of Arabic writing. I. Qurʾānic calligraphy". *Ars Orientalis* 21.119–148.

——. 1994. "The transformation of Arabic writing. II. The public text". *Ars Orientalis* 24.119–148.

Thackston, Wheeler M. 2001. *Album prefaces and other documents on the history of calligraphers and painters*. Leiden: E.J. Brill.

SHEILA S. BLAIR
(Boston College)

Figure 1. Page from the Blue *Qurʾān*. Museum Associates/LACMA. The Nasli M. Heeramaneck Collection, gift of Joan Palevsky. 2005 © Museum Associates/LACMA.

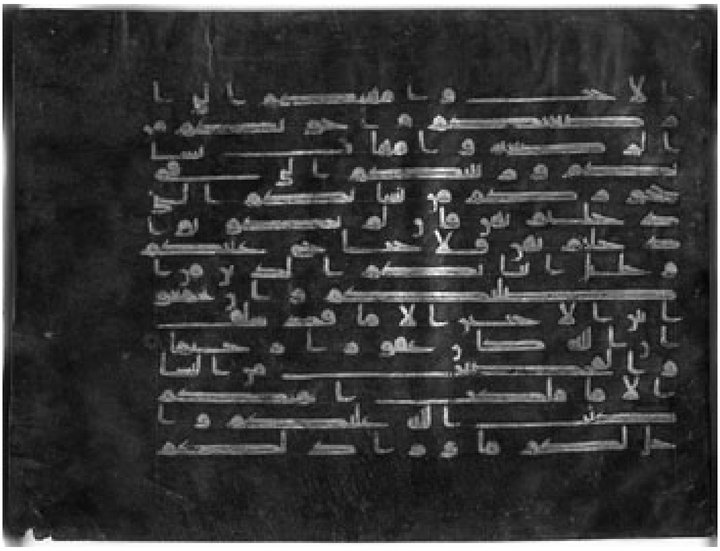


Figure 2. Dome of the Rock: Mosaic inscription. Photo Credit: Sheila Blair and Jonathan Bloom.

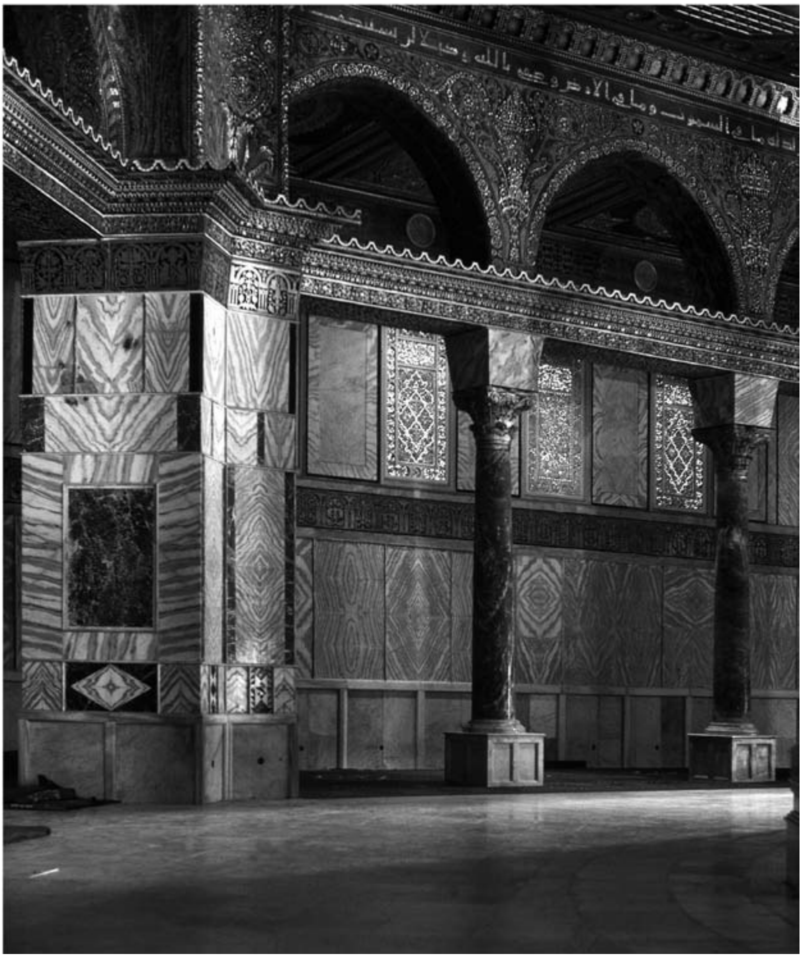


Figure 3. Amājūr *Qur'ān*: Damascus, 256–260H. In: *Encyclopaedia of Islam*, 1927.

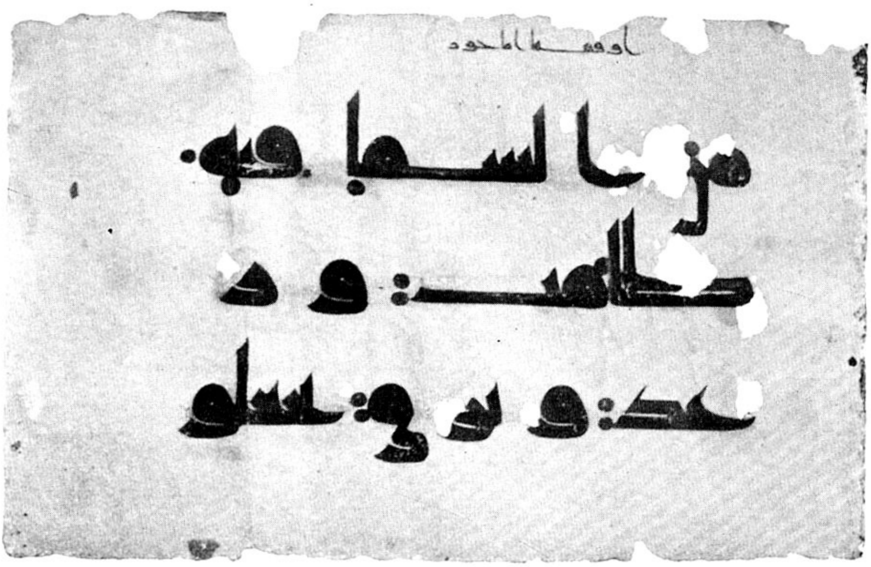
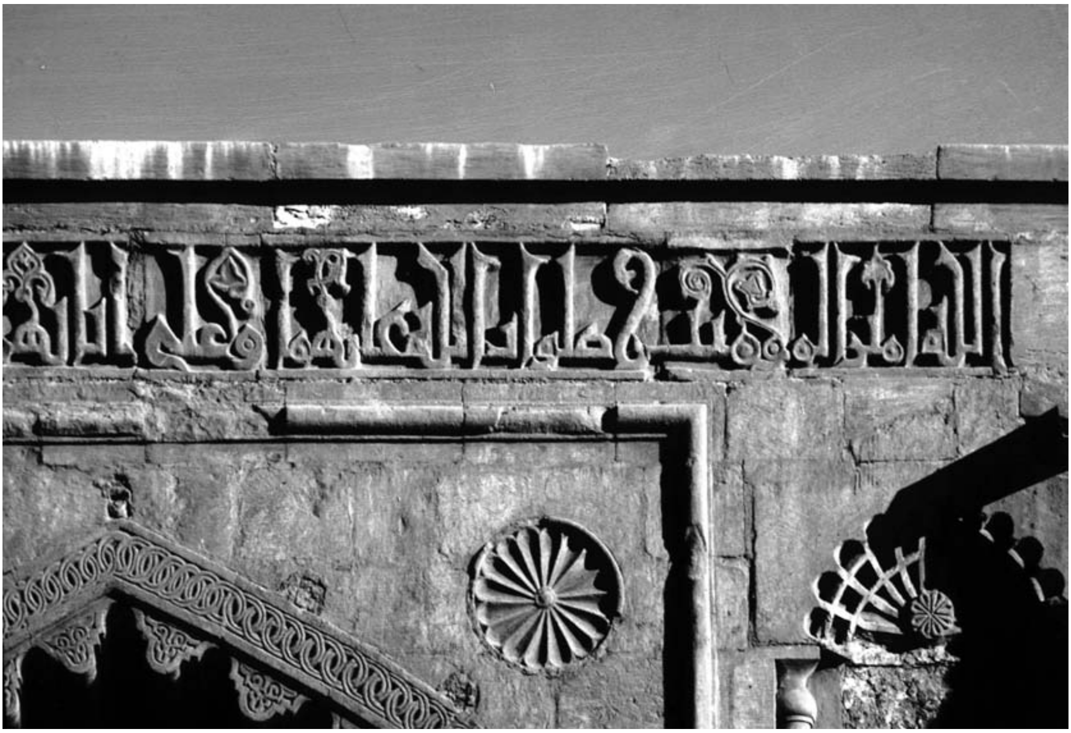


Figure 4. Floriated Kufic on the façade of the Aqmar mosque dated 519/1125–1126. Photo Credit: Sheila Blair and Jonathan Bloom.



Kurdish

1. Introduction

Kurdish is an Iranian language spoken in eastern Turkey, northern Syria, northern Iraq, and northwestern Iran, in a contiguous area known by the Kurds as Kurdistan. It is also found in pockets in the Caucasus, and even in Kazakhstan and Turkmenistan, as well as in central Anatolia and in Lebanon. The number of Kurds is difficult to estimate; the following numbers are sometimes quoted: 12 to 20 million Kurds in Turkey; 6.5 to 8 million Kurds in Iran; 4.5 to 5 million Kurds in Iraq; 1 to 2 million Kurds in Syria, but by now these numbers may be higher. However, the number of Kurds and the number of Kurdish speakers are not identical, particularly not in Turkey, where Kurdish is an endangered language.

The northern dialect of Kurdish, Kurmanji, is spoken by the Kurds of Turkey and Syria and the Caucasus, as well as in the villages in the hinterland of Urmia in northwestern Iran, and by half of the Kurds of Iraq, where it is known as Bahdînanî. The central dialect is known as Sorani and is spoken by half of Iraq's Kurds (in and around the towns of Sulaimania, Kirkuk, and Arbil), and by the majority of the Kurds of Iran (from Mahabad to Sanandaj [Sinneh]). The southern Kurdish dialects are spoken in the area of Kermanshah. Mention should also be made of the minority languages Zaza (Dimlî), spoken in Turkey, and Gorani (Hewramî), spoken along the Iraq-Iran border region.

Although contact with Semitic languages (Aramaic) predates the Islamic campaigns of the 7th and 8th centuries C.E., Arabic has had a great deal of influence on Kurdish ever since those momentous events. Arabic, and to a lesser extent Persian, have been the languages of learning, culture, and science as well as religion, and until recently it has been fashionable to pepper one's Kurdish with Arabic (and Persian) words. Consequently, early Kurdish classics such as Ahmed-i Khani's *Mem û Zîn* are so full of Arabic and Persian vocabulary that they cannot be used as models for modern Kurdish usage. Personal names also manifest Arabic influence. Only since the last decade of the 20th century has it become common practice to name Kurdish children after figures in Kurdish history and folklore, whereas before that period many people had Islamic names. This reflects a rise in

Kurdish national awareness. A common social phenomenon is the replacement of an original Arabic name with a new Kurdish name, so that someone originally named Ferhad is now known as Birûsk.

In transliterating Kurdish in this entry, the standard orthography is used, for which see: http://www.loc.gov/catdir/cpso/romanization/kurdish.pdf.

2. Phonetics

In the realm of phonetics, certain guttural and emphatic sounds are attributed to Arabic influence, such as the gutturals /ʿ/ and /ḧ/ in Kurdish. This is in fact an areal feature, also found in the Turkish of Diyarbakır, Van, etc., e.g. *ʿarabaya bin!* ‘get in the car!’, *raḧatsızlanmak* ‘to be uncomfortable’. It may well predate the advent of Arabic (i.e., it may be due to Aramaic influence). Several regional dialects are known for confusing /ʿ/ and /ḧ/, for instance the dialect of Yezidi speakers of Kurmanji, e.g. *siḧet* = *seʾat* ‘hour’ (also in Turkish: *saatler olsun* = *sıhhatler olsun*, said to someone who has just had his hair cut, or has just emerged from the hammam); likewise in the Sorani dialects of Arbil (Iraq) and Serdeşt and Şino [Ushnu] (Iran), e.g. *ʿAcî Ḧalî* = *Ḧacî ʿAlî*.

There are instances of hypercorrection, in which gutturals are added to Arabic loanwords originally lacking them, e.g. /ḧ/ in *deḧbe* < Arabic *dābba* ‘beast’; /ʿ/ in *ʿerd* < Arabic *ʾarḍ* ‘earth, ground, land’; and *ʿemir* < Arabic *ʾamr* ‘order, command’, in Kurdish a homonym with *ʿemir* ‘age’ < Arabic *ʿumr*. This phenomenon is not limited to words of Semitic origin, e.g. *ʿegît* ‘hero’ < Turkish *yiğit; ʿewr* ‘cloud’, cf. Persian *abr*. The /ʿ/ is unpredictable, having vanished in certain loanwords, such as *zeyf* ‘thin, weak’ < Arabic *ḍaʿīf, şayîr* ‘poet’ < Arabic *šāʿir*, but preserved in *şîʿir/şêʿr* ‘poetry’ < Arabic *šiʿr*.

Although Kurdish has the fricative /ġ/ (Arabic *ġayn*), spelled *ẍ* in Kurdish orthography (e.g. *aẍa* ‘agha, feudal lord’), in Arabic loanwords it is pronounced as unvoiced /x/ by many speakers, e.g. *xelet* ~ *ẍelet* ‘mistake, error’ < Arabic *ġalaṭ*, also Sorani *[heł]xełetandin* ‘to deceive, cheat, dupe’; *xeşîm* ~ *ẍeşîm* ‘naïve’ < Arabic *ġašīm; xulam* ~ *ẍulam* ‘servant boy’ < Arabic *ġulām; zext* ‘pressure’ < Arabic *ḍaġṭ*.

The emphatics /ṣ/ and /ṭ/ are of limited occurrence. Neither *xelet* nor *zext* (see preceding paragraph) carried over the Arabic /ṭ/. In addition to appearing in words – largely names – that were borrowed from Arabic, e.g. *S̱aliḧ*, *Ṯayyib*, these sounds also occur in words of Iranian origin, e.g. *şe* (Kurmanji)/*şeg* (Sorani) ‘dog’, *ṭa* (Kurmanji) ‘thread’, *ṭovik* ‘seed’ (cf. Persian *toxm* + *-ik*).

The sounds /q/ and /x/ (see discussion of /ġ/ above) are common in Kurdish, even in native Iranian words, e.g. *paqij* ‘clean’ (cf. Persian *pāk*), *qenc* ‘good’, *xwe* ‘oneself’ (cf. Persian *xvīš* and *xvud*), *xwar[d]in* ‘to eat’ (cf. Persian *xvurdan*).

Arabic /ḍ/, /ḏ̣/, and /ḏ/ all generally become *z* in Kurdish, e.g.

/ḍ/: *ʿerz* ‘honor’ < Arabic *ʿirḍ; razî* ‘willing, content’ < Arabic *rāḍī; Remezan* ‘Ramadan’ < Arabic *Ramaḍān; zext* ‘pressure’ < Arabic *ḍaġṭ; zirar/zerer* ‘damage’ < Arabic *ḍarar*

/ḏ̣/: *Mezher* ‘Mazhar [man’s name]’ < Arabic *maḏ̣har; weʾz* ‘sermon’ < Arabic *waʿḏ̣; wezîfe* ‘duty’ < Arabic *waḏ̣īfa; zerf* ‘envelope; adverb’ < Arabic *ḏ̣arf*

/ḏ/: *zikir* ‘recitation of the names of God’ < Arabic *ḏikr; zilfeqar* ‘trusty sword’ < Arabic *ḏū l-fiqār* ‘name of Ali’s sword’; *zur̄et* ‘offspring’ < Arabic *ḏurriyya*

Some of the above examples may have been borrowed through the intermediary of Turkish. There are also a few examples of all three being realized as *d*, which may indicate direct borrowing from spoken Arabic: *dêrandin* ‘to winnow’ < Arabic *ḏarā; derb* ~ *zerb* ‘blow, stroke’ < Arabic *ḍarb; ʿerd* ‘earth, ground, land’ < Arabic *ʾarḍ; xeyidîn* ‘to be angry’ < Arabic *ġayḏ̣*

In an article on Kurdish proverbs, MacKenzie (1970:110) mentions a phenomenon peculiar to Sorani, the pronunciation of Arabic /ḏ̣/ as velar *ł*, as an affectation of mullahs “who have been known to inject a fine emphatic Arabic [ẓ] into such plain Kurdish words as *minał* ‘child’, *bełam* ‘but’, *ełê* ‘says’, etc.”. He quotes a proverb containing the form *wełîfet*, i.e. *wezîfe[t]* ‘duty’.

Arabic /ṯ/ generally corresponds to *s*, e.g. *beḧs* in *beḧskirin* ‘to discuss, talk about’ < Arabic *baḥṯ* ‘studying’; *espab* ‘armor, equipment’ < Arabic *ʾaṯwāb*, pl. of *ṯawb* ‘garment’; *îsbat* in *îsbat kirin* ‘to prove’ < Arabic *ʾiṯbāt* ‘proving’; *mesel* ‘fable, tale’ < Arabic *maṯal* ‘parable; proverb’. As with /ḍ/, /ḏ̣/, and /ḏ/, this *ṯ* > *s* correspondence may also be through the intermediary of Turkish, whereas there are some examples of *ṯ* > *t* possibly from spoken

Arabic, e.g. *meteł* (Sorani) 'riddle' < Arabic *maṯal* 'proverb'; *mîratgîr* 'heir' < Arabic *mīrāṯ* 'inheritance' + Kurdish *-gîr* 'taker'.

The *tāʾ marbūṭa* (*-a*) is realized in some cases as *-e* and in others as *-et*/*-at*, as in → Persian and → Turkish. Examples of the former include: *feyde* 'benefit' < Arabic *fāʾida, ḧîle* 'ruse, trick' < Arabic *ḥīla, miẍare* 'cave' < Arabic *maġāra*. Examples of the latter include: *ḧekyat* (Kurmanji) 'folktale' < Arabic *ḥikāya* 'story'; *ḧikûmet* 'government' < Arabic *ḥukūma; seʾat* 'hour' < Arabic *sāʿa; siyaset* 'politics' < Arabic *siyāsa; welat* (Kurmanji)/*wiłat* (Sorani) 'country' < Arabic *wilāya* 'province'; *xîvet* (Kurmanji)/*xêwet* (Sorani) 'tent' < Arabic *xayma*. Rarely, both *-e* and *-et* occur as variants of the same word, such as *wezîfe* ~ *wełîfet* (affected Sorani form), or *tobe[t]* 'repentance' < Arabic *tawba*.

An original Arabic long *ā* is often realized in Kurdish as *ê*, e.g. *ḧisêb* 'account' < Arabic *ḥisāb; kitêb* 'book' < Arabic *kitāb; liḧêf* 'quilt' < Arabic *liḥāf; şêwirîn* 'to consult, deliberate' < Arabic *šāwara; xizêm* 'nose ring' < Arabic *xizām*.

An original Arabic *-m-* often becomes Kurdish *-v-/-w-*, e.g. *civat* ~ *cemaʾet* (Kurmanji) 'group' < Arabic *jamāʿa; ḧeravî* (Kurmanji) 'thief' < Arabic *ḥarāmī; silav* ~ *silam* (Kurmanji) 'greeting' < Arabic *salām[a]; siław* (Sorani) 'greeting' < Arabic *salām[a]; tewaw* (Sorani) 'complete, finished' < Arabic *tamām*. The same correspondence can be seen in native Iranian words, e.g. *hev* (Kurmanji)/*haw* (Sorani) = *ham* (Persian) 'each other'; *nav* (Kurmanji)/*naw* (Sorani) = *nām* (Persian) 'name'.

3. SEMANTICS

In the semantic realm, categories of borrowing from Arabic include:

i. Change in form, without change in meaning
cemaʾet (Kurmanji) 'group, assembly' < Arabic *jamāca* 'group, community'; *harsim* (Kurmanji) 'unripe grapes' < Arabic *ḥiṣrim; helbet* (Kurmanji) < Arabic *al-batta* 'absolutely'; *ḧekyat* (Kurmanji) 'folktale' < Arabic *ḥikāya* 'story'; *kitêb* < Arabic *kitāb* 'book'; *meşiyan* (Kurmanji) < Arabic *mašā* 'to walk'; *mizgeft* (Kurmanji)/*mizgewt* (Sorani) < Arabic *masjid* 'mosque'; *teqawît* (Sorani) < Arabic *taqāʿud* 'retirement'; *wext* < Arabic *waqt* 'time'

ii. Compound verbs
These consist of an Arabic noun and a Kurdish auxiliary verb, *kirin* (Kurmanji)/*kirdin* (Sorani) 'to do'. This is not limited to foreign borrowings (e.g. *nas kirin* 'to know a person', where *nas* is Kurdish). Examples with Arabic nouns include: *beyan kirin* 'to declare' < Arabic *bayān* 'statement'; *fêm/fehm/feʾm kirin* 'to understand' < Arabic *fahm* 'understanding'; *texmîn kirin* 'to suppose, presume' < Arabic *taxmīn* 'estimation, assessment'; *xilas kirin* 'to finish; to rescue, save' < Arabic *xalāṣ* 'rescue, salvation' and colloquial Arabic *xallaṣ* 'to finish'. The last example shows that both Classical Arabic and spoken Arabic dialects (Syrian and Iraqi in particular) are sources of borrowing.

iii. Kurdish singular meaning from original Arabic plural
cîran ~ *cînar* (Kurmanji) 'neighbor [sg. or pl.]' < Arabic *jīrān* [pl.]; *cemawer* (Sorani) 'crowd; group of people' < Arabic *jamāhīr* [pl.] 'multitudes, crowds'; *zilam* (Kurmanji)/*zelam* (Sorani) 'man, fellow' < Syrian Arabic *zilām* [pl.] 'men' (sg. *zalameh*). This phenomenon also occurs in Turkish, e.g. *talebe* 'student' < Arabic *ṭalaba* pl. of *ṭālib* 'student'; *tüccar* 'merchant' < Arabic *tujjār* pl. of *tājir* 'merchant'; *ukelâ* 'smart aleck' < Arabic *ʿuqalāʾ* pl. of *ʿāqil* 'wise man'.

iv. Change in meaning (Kurdish innovation)
berdêlî (Kurmanji) 'practice of marrying a brother and a sister of one family to the sister and brother of another family' < Arabic *badīl* 'substitute, replacement, stand-in'; *cahil/ciḧêl* (Kurmanji) 'young' < Arabic *jāhil* 'ignorant'; *cehş* (Kurmanji)/*caş* (Sorani) 'donkey foal; collaborator, Kurd who cooperates with the government against his own people' < Arabic *jaḥš* 'donkey foal'; *celeb* (Kurmanji)/*całeb* (Sorani) 'flock of sheep being led to market' < Arabic *jalaba* 'to bring, fetch, import'; *civat* (Kurmanji) 'society; evening social gathering of men' < Arabic *jamāʿa* 'group, community'; *deʾwat* (Kurmanji) 'wedding celebration' < Arabic *daʿwā* 'invitation'; *feqî* (Kurmanji) 'religious student' < Arabic *faqīh* 'jurisprudent'; *feqîr* 'nice, harmless' < Arabic *faqīr* 'poor, indigent'; *herikîn* (Kurmanji) 'to flow' < Arabic *ḥaraka* 'movement'; *hêrs* 'anger, fury, rage' < Arabic *ḥirṣ* 'greed'; *ḧez* in: *jê ḧez kirin* (Kurmanji) 'to love' < Arabic *ḥaḏ̣ḏ̣* 'luck, fortune'; *îsim* (Kurmanji) 'spell, incantation, charm' < Arabic *ism* 'name'; *kulfet* (Kurmanji) 'woman' < Arabic *kulfa*

'lady's maid; standing on ceremony; etc.'; *mamik* (Kurmanji) 'riddle' < Arabic *muʿammā* 'literary riddle'; *meteł* (Sorani) 'riddle' < Arabic *maṯal* 'proverb'; *micewir* (Sorani) 'mosque steward' < Arabic *mujāwir* 'neighboring, adjacent; student at al-Azhar University in Cairo'; *mitirp* ~ *mirtiv* (Kurmanji) 'Gypsy musician' < Arabic *muṭrib* 'singer, entertainer'; *qelp* (Kurmanji) 'counterfeit, fake, false' < Arabic *qalb* 'reversal, inversion, overthrow' [note that the Kurdish word is an adjective, while the Arabic original is a noun]; *qutabî* (Sorani) 'pupil, student' < Arabic *kuttāb* '*Qurʾān* school' + *-î; sekinîn* (Kurmanji) 'to stop, stand' < Arabic *sakana* 'to calm down, rest'; *sib[eh]ê* ~ *s̲ubaḧî* (Kurmanji)/ *sib[ḧ]eynê* (Sorani) 'tomorrow' < Arabic *ṣubḥ* 'morning'; *şer̄* 'war' < Arabic *šarr* 'evil, wickedness'; *şteẍilîn* (Kurmanji, dialectal) 'to speak' < Arabic *ištaġala* 'to work'; *taqî kirdinewe* (Sorani) 'to test, try out' < Arabic *taḥqīq* 'realization, assertion, verification'; *weḧş* (Kurmanji) [noun] 'pig' < Arabic *waḥš* [adj.] 'wild, untamed'; *xirab* (Kurmanji)/*xirap* (Sorani) [adj.] 'bad' < Arabic *xarāb* [noun] 'ruins'

iv. Calques from Arabic
pê rabûn (Kurmanji) 'to undertake, carry out' < Arabic *qāma bi-* lit. 'to stand up + in [preposition]'; *[kobûnewe] bestin* (Sorani) 'to hold [a meeting]' < Arabic *ʿaqada [ijtimāʿan]* (*bestin* = *ʿaqada* 'to tie')

v. In the realm of word formation, one remarkable phenomenon is the existence of predictable patterns in Kurmanji, based on Semitic triliteral roots. All examples quoted in Table 1 are from Arabic and Aramaic.

Table 1. Word formation patterns in Kurdish

noun	simple verb	causative verb
<FeʿaL>	*<FeʿiLîn>*	*<FeʿiLandin>*
ʿelam 'notification'	*ʿelimîn* 'to learn'	*ʿelimandin* 'to teach'
lebat 'movement'	*lebitîn* 'to move, stir'	*lebitandin* 'to move'
şewat 'burning'	*şewitîn* 'to be on fire'	*şewitandin* 'to burn'
xebat 'work'	*xebitîn* 'to work'	*xebitandin* 'to use'
xepar 'digging'	*xepirîn* 'to dig'	*xepirandin* 'to cause to dig'
zewac 'marriage'	*zewicîn* 'to be wed'	*zewicandin* 'to wed, marry off'

4. REGIONAL INFLUENCES

In the section on Kurdish language in the *Encyclopaedia of Islam*, MacKenzie (1954:497b) states that "Northern Kurdish appears to have been somewhat more open to the penetration of Arabic and especially Turkish loanwords". Arabic words may be found in Kurmanji (= Northern Kurdish) dialects as far afield as Armenia and Azerbaijan, and the same can be said for the neighboring Azerbaijani and Persian languages. Whereas Anatolian Turkish influence is prevalent throughout the Kurmanji subdialects, both in Turkey and beyond, the influence of other neighboring languages can also be discerned in particular regions. In the Northeast (the *Serhedan* or '[Ottoman and Persian] Border' dialects, spoken in Kars, Ağrı, the Republic of Armenia, and in the hinterland of Urmia in Northwest Iran), Azerbaijani influence can be seen in such words as *begem kirin* <*beyen-* 'to like', *qatix* <*gatig* 'yoghurt', *kʾok* < *kök* 'fat'. In the Northwest (Dêrsim/ Tunceli, Bingöl, Elâzığ), the Armenian influence is most strongly felt, e.g. *[h]avlik* < *avel* 'broom', *hêlî* < *hayeli* 'mirror', as well as the numbers 11–19 *deh-û-yek, deh-û-dido*, etc., patterned after Armenian *dasn-u-meg, dasn-u-ergu*, etc. Semitic influence is strongest to the south of this area: Aramaic influence is particularly noticeable in the Bahdinan region of Iraqi Kurdistan, both in vocabulary (*mişext* 'exiled, away from home', *xepirîn* 'to dig') and in the external marking of grammatical gender (*yê* masc., *ya* fem., *[yê]t* pl.) on adjectives and participles (Turkish *yê baş î* 'you are well [masc.]/*ya baş î* 'you are well [fem.]'). It should be noted that Kurdish influence on the Neo-Aramaic dialects is also quite strong. Arabic influence – in evidence throughout Kurdistan – is most strongly felt in Mardin province of Turkey and the Syrian dialects (the area known as *Binxet* 'beneath the line [drawn to separate Syria from Turkey]'), e.g. *ʿecibandi*n 'to be pleasing, to like' < Arabic *ʿajaba*; *ştexilîn* 'to speak' < Arabic *ištaġala* 'to work'; *neciḧîn* 'to succeed' < Arabic *najaḥa*. Arabic influence can also be seen in proverbs from this region, e.g. *meymûn çʾeʾvê dya xweda – xezale* (Dzhalil and Dzhalil 1972:219, no. 1154) 'a monkey is a gazelle in its mother's eyes', a well-known Arabic proverb (*il-qird bi-ʿayn immo ġazal*). Another example is *destê te nikanî geskira tê* *maçʾkî* (Dzhalil and Dzhalil 1972:95, no.329) ‘you will kiss the hand you cannot bite’ (Arabic: *īd il mā fīk tiʿaḍḍā būsā, widʿī ʿaleyā bilkasr* ‘kiss the hand you cannot bite, and pray for it to break’). By contrast, a version of this Kurdish proverb from the Mahabad region of Iranian Kurdistan resembles more closely the Turkish version, which is *kesemediğin eli öp* ‘kiss the hand you cannot cut off’, with ‘biting’ replacing ‘cutting off’.

Because the Sorani dialect (Central Kurdish), particularly the Sulaimania subdialect, boasts a literary tradition, Arabic influence on it is often minimized. However, many Sorani speakers use Arabic loanwords in their speech, particularly in Kirkuk and Arbil. There are also Sorani words of Arabic origin which have been changed beyond recognition, such as *cemawer* ‘crowd, group of people’ < Arabic *jamāhīr* [pl.] ‘multitudes, crowds’; *qutabî* ‘pupil, student’ < Arabic *kuttāb* ‘*Qurʾān* school’ + *-î; teqawît* < Arabic *taqāʿud* ‘retirement’; *taqî kirdinewe* ‘to test, try out’ < Arabic *taḥqīq* ‘realization, assertion, verification’.

5. KURDISH INFLUENCE ON ARABIC

With the exception of Arabic dialects spoken in the provinces of Mardin, Batman, Siirt, and Urfa in Kurdistan of Turkey (→ Anatolian Arabic; Vocke and Waldner 1982:xlv–li), where Kurdish is the dominant language, there are very few Kurdish borrowings into Arabic.

What few loanwords exist include: *jabas* (Arabic dialect of Aleppo, Syria) < Kurmanji *zebeş* ~ *şebeş* ~ *cebeş* ‘watermelon’; *sarqīn* < Kurmanji *sergîn* ‘dunghill’ < *sergo* ‘ dung’ < *ser* ‘head; on’ + *gû* ‘feces’; and possibly *kawm* ‘heap’ < Kurmanji *kom*/Sorani *ko* ~ *kom*.

There is a proverb shared by Arabic, Kurdish, and Aramaic: ‘He who knows, knows, and he who doesn’t know says, ‘It’s a handful of lentils’. It is unknown in Turkish and Persian. The Kurdish form exhibits rhyme: *yê zane zane, yê nizane – baqê nîskane*. The Arabic lacks rhyme (*il yidrī yidrī, il mā yidrī yiqūl kaff* [or *gaḍbit*] *ʿadas*), which suggests Kurdish origin. This proverb, implying that there is more to the incident at hand than meets the eye, is roughly the equivalent of English ‘if you only knew the half of it’.

BIBLIOGRAPHICAL REFERENCES

Chyet, Michael L. 2003. *Kurdish-English dictionary = Ferhenga Kurmanjiancî-Ingilîzî*. New Haven, Conn.: Yale University Press.

Dzhalil, Ordikhane and Dzhalile Dzhalil. 1972. *Mesele û met’elokê k’urda bi zimanê k’urdî û rûsî = Kurdskije poslovicy i pogovorki na kurdskom i russkom jazykax* [Proverbs and sayings in the Kurdish and Russian language]. Moscow: Glavnaja Redakcija Vostočnoj Literatury.

MacKenzie, D.N. 1954. “Kurds, Kurdistan. V. Language”. *Encyclopaedia of Islam*, V, 479–480. 2nd ed. Leiden: E.J. Brill.

——. 1970. “Some Kurdish proverbs”. *Iran, Journal of the British Institute of Persian Studies* 8.105–113.

Qazzaz, Shafiq. 2000. *The Sharezoor: Kurdish-English dictionary*. Erbil: Aras.

Vocke, Sibylle and Wolfram Waldner. 1982. *Der Wortschatz des anatolischen Arabisch*. Erlangen: University of Erlangen/Nürnberg.

MICHAEL CHYET
(Library of Congress)

Kurmanji → Kurdish

Kuwait Arabic → Gulf States

Kuwaiti Arabic

1. GENERAL

1.1 *Area*

The mainland of Kuwait is a wedge-shaped, almost perfectly flat expanse of desert (17,818 km^2) with borders to the north and northwest with Iraq and to the west and south with Saudi Arabia. It has a coastline approximately 195 km in length. There are three principal islands: Failaka, 20 km east of Kuwait City, which once had a small population but has been uninhabited since the Iraqis expelled the population in 1990, and the much bigger islands of Warbah and Bubiyan, at the head of the Gulf, which are uninhabited but on which the Kuwait government has plans to resettle up to 100,000 people. The population of Kuwait was ca. 2.2 million in 1999, of whom 798,000, or 35 percent, are Kuwaiti by nationality. The rest of the population is a shifting multinational mix of foreign workers.

1.2 *Society*

There is a sharp social distinction between Kuwaitis who have full citizenship, including the right to vote (if male and over 21) – urban, mainly mercantile, and descended from the Najdi tribes that arrived in the area from the mid-18th century – and other Kuwaitis, mostly recently sedentarized Bedouin who roamed the borderlands of Iraq and Saudi Arabia until the early 1950s, who have a lesser form of citizenship. There is an even sharper distinction between Kuwaitis and non-Kuwaitis, with certain residential areas of Kuwait City being reserved for Kuwaiti nationals only. Kuwait is an overwhelmingly urbanized society: official estimates put the urban population at 97 percent of the total, concentrated in Kuwait City and a few satellite townships such as Salmiyya, Fahahil, and Ahmadi.

1.3 *Regional context*

Kuwaiti Arabic is typologically similar to the Bedouin-descended dialects of other Gulf States, such as those of the ʿArab population of Bahrain, Qatar, and the United Arab Emirates, though containing some distinctive local features which ally it with the dialects of nearby southern Iraq. The tribal dialects that could once be distinguished in the Kuwait of the 1950s, described in Johnstone (1961) and (1964), have gradually coalesced into a dialect that, while still distinctively Kuwaiti, is now a leveled and somewhat classicized Gulf koine which has absorbed influences from written Arabic (85 percent of the population is literate) and other Arabic dialects, although the degree of influence in any individual's speech depends on age and level of education. The Kuwaiti population is approximately 75 percent Sunni and 25 percent Shiʿi, many of the latter being originally immigrants from other areas of the Gulf such as Bahrain and eastern Saudi Arabia, but there is no sect-based dialect cleavage.

Kuwait has a thriving television production industry that makes and exports soap operas in the local dialect to neighboring Gulf States. As in other Gulf States there is a thriving vernacular poetic tradition.

1.4 *Historical evidence*

There is no tradition of writing in the dialect, so little evidence of what the dialect was like before World War II, roughly the time when Western Arabists, principally with the aim of equipping Americans and others working in the oil industry with a knowledge of the local dialect, began to write pedagogical grammars such as that published by the Kuwait Oil Company (KOC) in 1951. Modern Kuwaiti appears to be a leveled local koine which has several ingredients: the dialects of the coastal population, originally Najdi in character but which, like those in the rest of eastern Arabia, had absorbed Persian, Indian, Portuguese, and English vocabulary over a long period; and the dialects of more recent migrant groups from northern Arabia, southern Iraq, and even southwestern Najd (Johnstone 1961).

1.5 *State of research*

Following Johnstone's (1961, 1964) early studies of the Dōsiri dialects of Kuwait and his general study of the Gulf littoral dialects (1967), which places Kuwait in a regional context, Maṭar (1969, 1970) carried out useful phonological and morphological studies that largely replicate Johnstone's work but provide many more examples. Ingham (1982) is a comparative study of the Bedouin and Bedouin-descended dialects of the whole northeastern region of the Arabian Peninsula, in which Kuwait is included. On the vocabulary side, Ḥanafī (1964) provides a study of the lexicon of the pre-oil culture of Kuwait, while Sabʿān (1983) is a brief study of the recent evolution of Kuwaiti neologisms and borrowings. There has been no more recent, in-depth, dialectological or sociolinguistic study of Kuwait City.

In what follows, the language level described is that of urban Kuwaitis speaking in relaxed conversational circumstances. Where appropriate, differences between educated and uneducated speakers are noted. In most details, the Kuwaiti dialect is similar to the A dialect of Bahrain (→ Bahraini Arabic).

2. Linguistic description

2.1 *Phonology*

2.1.1 Inventory

2.1.1.1 Consonants

Plosives: (p), b, t, d, ṭ, ḍ, k, g, q, ʾ
Affricates: č, j
Fricatives: f, ṯ, ḏ, ḏ̣, x, ġ, ḥ, ʿ, h

Sibilants: s, ṣ, z, š
Laterals and vibrants: l, r
Nasals: m, n
Semivowels/glides: w, y

Some speakers have /p/ in borrowings, e.g. *pēp* ‘pipe’, which others pronounce with /b/. /č/ and /g/ are common in borrowings, e.g. *čāy* ‘tea’, *rig* ‘oil rig’, and are indistinguishable phonetically from /č/ and /g/ resulting from internal phonological changes.

The Old Arabic interdentals, /ṯ, ḏ, ḏ̣/, are preserved in Kuwaiti, and words with etymological /ḍ/ are pronounced with /ḏ̣/, e.g. *ḏ̣arab* ‘he hit’. Normally /g/ < Old Arabic /q/, e.g. *gāl* ‘he said’, but in front vowel environments /g/ was fronted and affricated in many words to /j/, e.g. *bāji* ‘remainder’, *rīj* ‘saliva’. In similar environments, Old Arabic /k/ was fronted and affricated to /č/, e.g. *čibīr* ‘great, old’, *fačč* ‘to open’. Nowadays, these affricated variants in even slightly formalized speech are replaced by the corresponding stops /g/ and /k/; some educated speakers do not use the affricate variants at all. Old Arabic /j/ became /y/ categorically, e.g. *yār* ‘neighbor’, *daray* ‘step’, but /j/ in foreign borrowings is usually preserved, e.g. *jūti* (< Urdu) ‘shoes’, *jām* (< Persian) ‘pane of glass’ (although such old-fashioned borrowings may be avoided altogether by educated speakers). As in Bahrain, Old Arabic /ġ/ > [ɢ] or [q] as well as [ɣ]: [qeːr] ~ [ɢeːr] ~ [ɣeːr] are all possible realizations of /ġēr/ ‘other’. Modern Standard Arabic /q/ in neologisms is often pronounced [ɣ] or [ɢ], e.g. [taɣaddum] ~ [taɢaddum] ‘progress’, its allophones thus overlapping completely with those of dialectal /ġ/. Old Arabic /ʾ/ disappeared initially and finally, e.g. *xaḏ* ‘he took’, *gaḏa* ‘fate’, and was replaced medially by vowel length, e.g. *yīt* ‘I came’, *rās* ‘head’. /ʾ/ occurs even in uneducated speech in a few items which are Modern Standard Arabic-derived, e.g. *yisʾal* ‘he asks’ (alongside the dialect form *ysāyil*). /l/ and /r/ have velarized allophones in some words, especially when a labial is present, when the whole word may become velarized, e.g. [g̴ɑb̴ɪɫ] ‘before’.

2.1.1.2 Vowels

Kuwaiti has three short vowels, /a, i, u/, and five long, /ā, ē, ī, ō, ū/. Unstressed /i/ in non-final open syllables is often deleted (see Sec. 2.1.3.1). /ē/ and /ō/ are sometimes shortened in particular words, but the resulting short mid vowels are not phonemic.

2.1.1.2.1 Short vowels

Distribution and quality of short vowels: As in other eastern Arabian Bedouin-type dialects, /i/ occurred historically to the exclusion of /a/ in open, non-final syllables, except in the contiguity of guttural consonants /x, ġ, ḥ, ʿ, h/, or where the following consonant was /l/, /n/, or /r/ when at the same time the vowel of the following syllable was /a/ or /ā/ (Johnstone 1967:27). Thus, *kitab* ‘he wrote’ but *barad* ‘it got cold’, *tkallam* ‘he spoke’ but *tkallimaw* ‘they spoke’. However, there is now a tendency to replace /i/ with /a/ in such forms in educated speech. For some (perhaps less educated) speakers, the labials have a rounding and backing effect on /i/ (> /u/) whether /i/ is original or < /a/, especially if a velar, emphatic, or /l/ or /r/ are also present, e.g. *mukān* ‘place’, *buṣal* ‘onions’, *šrubat* ‘she drank’, but *xšiba* ‘piece of wood’, where /b/ is present but none of the other factors. For such speakers, /i/ and /u/ are virtually in complementary distribution in open syllables but not in closed ones. Final /ā/ and /āʾ/ are shortened; when not in contiguity with a guttural, emphatic, /l/, or /r/, and when preceded by an open syllable, they are both shortened and raised, e.g. *niši* ‘starch’, *simi* ‘sky’ (an exception is *māy* ‘water’) but *ḥamra* ‘red [fem.]’, *ramḏ̣a* ‘hot ground’.

/a/ is realized: as [æ] or [ɛ] where gutturals (excluding /h/) and emphatics are absent, e.g. [hæli] ‘my family’, [dæzː] ‘he pushed’; as [a] in guttural environments, e.g. [baʕad] ‘after’ [xalːeːt] ‘I/you allowed’; as [ɑ] with an emphatic, and often with labials, /r/ and /g/, e.g. [t̴ɑɫː] ‘mist’, [xɑm̴ɑr̴] ‘alcohol’, [g̴ɑb̴ɪɫ] ‘before’. Medial /i/ is retracted, e.g. [bɪnt] ‘girl’; in final position it is closer and more front, e.g. [rɪħti] ‘you [fem.] went’; with emphatics it is lowered, e.g. [yɪr̴t̴ʌn̴] ‘he gabbles’. /u/ is back and rounded [ʊ], e.g. [ʃrʊbæt] ‘she drank’, [gʊmɑr] ‘moon’, [s̴ʊbː] ‘pour!’

2.1.1.2.2 Long vowels

Generally, /ā/ has a more retracted articulation than /a/, especially when an emphatic or any of /x, ġ, r, g/ are present, e.g. [saːl] ‘it flowed’, but [rɑːħ] ‘he went’, [gɑːl] ‘he said’. /ī/ is a close, front vowel, but with the emphatics it is more centralized, e.g. [bɨ ː s] ‘keel of a boat’. /ē/ and /ō/, both pure mid vowels, correspond to the Old Arabic diphthongs /ay, aw/ and occur medially, e.g. *ġēm* ‘clouds’, *lōn* ‘color’.

2.1.1.3 Diphthongs

Where /aw/ occurs in the first syllable of a word which has a heavy syllable (Cv̄ or CvC) following, it is often reduced to /ā/, e.g. *mājūd* ‘existent’. In verb forms, final /aw/ > /ō/ when suffixed, e.g. *gālaw* ‘they said’, *gālōli* ‘they told me’. /ū/ is [ʊ ː], e.g. [ʃʊːf] ‘look!’

2.1.1.4 Syllable types

The following syllable types are possible.

Cv:	*ḏ̣arab* ‘he hit’ (Cv-CvC)
CvC:	*gafša* ‘spoon’ (CvC-Cv)
CvCC:	*gilt* ‘I said’, *maġarb* ‘evening’ (Cv-CvCC), *yisallimk* ‘[may God] preserve you’ (Cv-CvC-CvCC)
[Cv̄]:	*sōlaf* ‘he chatted’ (Cv̄-CvC)
[Cv̄C:]	*bāg* ‘he stole’

The above are the basic types. Cv̄CC also occurs, but only in one type of form (the active participle of geminate verbs), e.g. *rādd* ‘rejecting, replying’. The following also arise, partly as a result of the short vowel and resyllabication deletion rules operated by some speakers, partly as a result of borrowings.

CCv:	*drisat* ‘she studied’, *ghawa* ‘coffee’ (CCv-CvC)
CCv̄C:	*kwēt* ‘Kuwait’, *smīt* ‘cement’ (borrowing)
CCvCC:	*fhimt* ‘I understand’, *trinj* ‘citron’ (borrowing)

Distribution of syllable types

CCv is always word-initial: CvCvCv(C) is reduced by elision of the first vowel (see Sec. 2.1.3.1). A prosthetic vowel is inserted before the resulting consonant cluster after a word ending in a consonant, e.g. *l-ixšiba* ‘the lump of wood’.

CvCC is always word-final and, except for monosyllabic words like *šarg* ‘east’, arises in some words as a result of the so-called *ghawa*-syndrome (see Sec. 2.1.2.4) and in a few fixed phrases (e.g. *allah yisallimk*).

2.1.1.5 Consonant clusters

CCC clusters that arise via suffixation, where the first two consonants are the same, are usually reduced, e.g. *ḥagna* ‘our right, for us’ (< *ḥagg* + *na*), *ṭagha* ‘he hit her’ (< *ṭagg* + *ha*), and *kilkum, kilhum* ‘all of you/them’. In one common item, *ʿind*, where the first two consonants are different, there is reduction, e.g. *ʿidkum, ʿidna* (although ‘imported’ epenthesized forms such as *ʿindəna* are also heard). Otherwise, CCC clusters are stable, e.g. *bintkum* ‘your daughter’, *šiftkum* ‘I saw you’, although again, there is variation: some speakers, unpredictably, have forms of the *binitkum, šifitkum* type that are standard in Muslim Baghdadi and may be a reflection of Iraqi influence. Reduction in one high-frequency phrase is universal for all: *git* < *gilt*, ‘I/you [masc.] said’ in phrases like *git lik/lah* ‘I told you/him’.

The treatment of non-doubled final clusters in words of the structure CvCC in Old Arabic depends on the preceding vowel and the consonants:

Old Arabic CaCC

i. C_2 is /l, m, n, r/: the form is stable, e.g. *galb* ‘heart’, *ḏanb* ‘sin’, *warč* ‘thigh’, *šams* sun’.
ii. C_2 is a guttural: CaCaC is normal, e.g. *baḥar* ‘sea’, *šahar* ‘month’, *naxal* ‘palm trees’.
iii. In all other cases, the form is CaCiC, e.g. *ḥabil* ‘rope’, *xašim* ‘nose’, *gaṣir* ‘palace’, the /i/ having a central quality if a labial and an emphatic consonant (including /l/ and /r/) occur together as C_2 and C_3, e.g. in *ʿaḏ̣im* ‘bone’, *gabil* ‘before’, *tamir* ‘dates’.

Old Arabic CiCC and CuCC

i. C_2 is /l, n, r/: the form is stable, e.g. *bint* ‘girl’, *ʿilč* ‘chewing gum’, *ʿirs* ‘marriage’, *bunk* ‘essence’, *gurṣ* ‘flat piece of bread’.
ii. In other cases CvCvC, with vowel harmony, is the most common form, e.g. *ʿugub* ‘after’, *ḏ̣uhur* ‘noon’, *dihin* ‘oil, fat’, *mišiṭ* ‘comb’, but there are exceptions, e.g. *rizg* ‘sustenance’, *bišt* ‘man’s cloak’, *ṣubḥ* ‘morning’.

2.1.1.6 Stress

The rule is that the last syllable of a polysyllabic word is stressed if long (Cv̄C, CvCC);

otherwise, the penultimate is stressed, except in cases of a succession of three short Cv syllables, when the antepenultimate is stressed, e.g. *báraka* ‘blessing’, although some (uneducated) speakers reduce all such concatenations, verbal or nominal, to CCvCv and stress them accordingly, viz *bríka*. Stress is nondistinctive.

2.1.2 Phonotactics

2.1.2.1 Assimilation

The following are major loci:

i. The definite article: sun-letters and moon-letters behave as in Old Arabic.
ii. /l/ in the imperative *xall* ‘let’ > /n/ before the *-ni* and *-na* suffixes, e.g. *xanna* ‘let’s…’ For some speakers, this assimilation is general, e.g. *štaġanna* ‘we worked’, *ginna* ‘we said’.
iii. /ġ/ > /b/ in the imperfect forms of the verb *baġa* ‘to want, need’ e.g. *abbi, tabbi*, etc.
iv. /h/ in the 3rd person suffixes *-ha* and *-hum* is assimilated to the *-t* of the 3rd person singular feminine perfect verb and other feminine forms, e.g. *drisatta* ‘she studied it [fem.]’, *rgubattum* ‘their neck’.
v. /t/ in *ti-* verbal prefixes of various kinds is assimilated by some speakers to /t, ṭ, ṯ, d, ḏ, ḍ, ḏ̣, č, j, s, ṣ, š/ as a consequence of the deletion of unstressed /i/ in open syllables (see Sec. 2.1.3), e.g. *(i)ṣṣīr* ‘she becomes’, *(i)ṭṭawwar* ‘it developed’, *(i)ččiddūn* ‘you work for your money’.

2.1.2.2 Dissimilation

Certain dialectal quadriliteral verbs may have arisen historically via consonantal dissimilation, e.g. *xarbaṭ* ‘to get confused’ < *xabbaṭ* ‘to beat’.

2.1.2.3 Metathesis

Occasional examples noted *gaḏ̣b*, ‘grasping’ for *gabḏ̣*, but no regular and clear pattern was noted.

2.1.2.4 *Ghawa*-syndrome

The *ghawa*-syndrome (→ *gahawa*-syndrome) is the deletion of /a/ in CaC non-final syllables where C_2 is a guttural, and epenthesis of /a/ after C_2, e.g. *nxala* (< *naxla*) ‘palm tree’, *yiʿarf* (< *yiʿarif* < *yaʿrif*) ‘he knows’, *məġarb* (< *məġarib* < *maġrib*) ‘evening’. This rule is now moribund as an active phonological process, although its results survive in a few common words, especially colors, e.g. *xaḍar* ‘green [masc.]’, *ḥamar* ‘red [masc.]’, and some proper names, e.g. *ḥamad* (< *ʾaḥmad*).

2.1.2.5 Conditioned *ʾimāla*

See Section 2.1.1.2: vowel raising.

2.1.2.6 Spread of velarization

Velarization may spread, especially to /l/, /r/, /b/, /m/, e.g. *ṣabi* ‘boy’ (= [s̴ɑb̴i]), *ṭabil* ‘drum’ (= [t̴ɑb̴ɪɫ], *ṣabir* ‘patience’ (= [s̴ɑb̴ə̴r̴]. Combinations of velars, labials, and /l/ or /r/ may also become velarized when an emphatic consonant is not present, e.g. *ramil* ‘sand’ (= [r̴ɑm̴ɪɫ]), *gabil* ‘before’ (=[g̴ɑb̴ɪɫ]), *ixtarab*, ‘to go rotten’ (= [ixt̴ɑr̴ɑb̴]).

2.1.3 Morphophonology

2.1.3.1 Elision of vowels

Unstressed /i/ and /u/ in non-final open syllables are deleted in certain syllable structures (e.g. in CvCv̄C, CvCvCC, and as the first vowel in CvCvCv(C)), and a prosthetic vowel /i/ is inserted utterance-initially or if a consonant-final word precedes, e.g. *itgūl* ‘you say’, *l-ibyūt* ‘the houses’, *l-išyara* ‘the tree’ (citation forms *buyūt, šiyara*). Medially doubled consonants are reduced in certain verb forms as a result, e.g. *(i)tʿallim* ‘you [masc.] teach’ but *(i)tʿalmū́n* ‘you [pl.] teach’, *labbisaw* ‘they dressed’ but *labsṓha* ‘they dressed her’. Uneducated speakers tend to delete /a/ in the initial syllable of CvCvC(v/v̄) strings, e.g. *šarab* ‘he drank’ → *(i)šrubah* ‘he drank it’, *šrubaw* ‘they drank’.

2.1.3.2 Insertion of vowels

See Section 2.1.3.1 for prosthetic /i/. See Section 2.1.1.5 for epenthetic vowels, and Section 2.1.2.4 for the *ghawa*-syndrome.

2.1.3.3 Shortening and lengthening of vowels

Where two or more long vowels occur in a word, the unstressed first long vowel may be shortened, especially in rapid speech, e.g. *rayāyī́l* ‘men’ = [ræyæyiːl]. Final vowels that are short in unsuffixed form are lengthened and stressed when suffixed, e.g. *yxálli* ‘he lets’, *yxallī́ni* ‘he lets me’; *xálla* ‘he let’, *xallā́ha* ‘he let her’.

2.1.3.4 Clitics

There are a number of clitics:

i. Deictic *ha-* prefixed to the definite articles, e.g. *ha-l-bēt* ‘this house’.

ii. *b-* prefix indicating proximate intent, e.g. *b-inrūḥ bāčir* ‘we’ll go tomorrow’.
iii. *taw(w)-* is arguably a clitic, as it does not occur unsuffixed and signifies recently completed action, e.g. *taw-ni yāy* ‘I’ve just arrived’.
iv. *š-* ‘what?’, e.g. *š-gilt* ‘what did you say?’, *š-ḥaggah?* ‘what for?’.
v. Presentative *kā-*, e.g. *kā-hu-yāy!* ‘here he comes!’.
vi. Presentative *hā-*, e.g. *hā-č ilʕēš* ‘here’s the rice for you [fem.]!’.
vii. *mā*, inserted between the repetition of a word to indicate ‘and such like’, e.g. *fār mā-fār* ‘rats and things like that’.

2.1.3.5 Construct state

This is less used because of the development of an analytic genitive (Sec. 2.3.1.3). Plural and (for some speakers) dual nouns retain their endings when they are the head noun, e.g. *yūniyyat ʕēš* ‘a sack of rice’, *rōḥat issīf* ‘going to the seashore’, *mwaḏ̣ḏ̣afīn ilḥukūma* ‘government employees’, *īdēn irrayyāl* ‘the man’s [two] hands’ (but also *īdē lmara* ‘the woman’s hands).

2.1.3.6 Suffixation

(a) The /h/ of the *-ha* and *-hum* suffixes assimilates to the suffix /t/ of feminine forms (see Sec. 2.1.2.1). When vowel-initial suffixes are added to suffix-stem 3rd person verb forms, they are resyllabified, in line with the general rule which replaces /a/ in non-final open syllable with /i/ except in certain well-defined phonological environments (Sec. 2.1.1.2) and then deletes it, e.g. *(i)šrúbat* ‘she drank’ + *ah* > (*šrubitah*) > *(i)šrúbtah* ‘she drank it [masc.]’. In all suffixed plural forms, the final *-aw* becomes -*ō* and is stressed, e.g. *šrúbaw* ‘they drank’ + *ah* > *(i)šrubṓh* ‘they drank it’.

(b) Active participle + suffix forms:

Table 1. Participle *kātib* ‘writing’ + suffix in Kuwaiti Arabic

	vowel-initial suffix (*-ah*)	consonant-initial suffix (*-ha*)
sg. masc.	*kātbah*	*kātibha*
sg. fem.	*kātbitah*	*kātbatta*
pl.	*kātbīnah*	*kātbīnha*

The corresponding forms in II*w/y* and III*y* roots are *bānīh, bānyitah* and *čāyifha, čāyfitah*, etc.

(c) The treatment of suffixed CvCvC nouns is somewhat unpredictable. Some ‘core’ nouns such as *yimal* ‘camel’ become *ymalik* ‘your camel’, but others, seemingly neologisms, do not undergo this rule. Contrast *ḥarasah* ‘his defense force’, a neologism, with *ḥrisah* ‘he defended him’, both < *ḥaras* + *ah*.

2.2 *Morphology*

2.2.1 Pronouns

2.2.1.1 Personal independent pronouns

Table 2. Independent pronouns in Kuwaiti Arabic

3rd pers. masc. sg.	*hū, huwa*
3rd pers. fem. sg.	*hī, hiya*
3rd pers. pl.	*hum*
2nd pers. masc. sg.	*inta*
2nd pers. fem. sg.	*inti*
2nd pers. pl.	*intu*
1st pers. sg.	*ana*
1st pers. pl.	*iḥna, ḥinna*

2.2.1.2 Possessive/object suffixes

Table 3. Possessive/object suffixes in Kuwaiti Arabic

3rd pers. sg. masc.	*-ah*
3rd pers. sg. fem.	*-ha*
3rd pers. pl.	*-hum*
2nd pers. sg. masc.	*-ik*
2nd pers. sg. fem.	*-ič*
2nd pers. pl.	*-kum*
1st pers. sg.	*-i* (poss.); *-ni* (obj.)
1st pers. pl.	*-na*

There is some allomorphy in the 1st person with certain prepositions, e.g. with *l-: lī*; with *ila: ilī, ilayya*; with *fī: fiyya, fīni*; with *ʕala: ʕalayya*; with *bi: biyya*; with *wiyya: wiyyāy*.

2.2.1.3 Indirect object suffixes

With some verbs, the indirect object is suffixed directly to the verb and the direct object carried by the particle, e.g. *ʕaṭni iyyāh* ‘give it to me!; but in most cases of such constructions involving two pronoun objects, it is the direct object that is suffixed to the verb and the indirect object (= the embedded subject) to *iyya-*, e.g. *sammaʕōha iyyāy* ‘they made/let me listen to it’ (lit ‘they made it audible to me’).

2.2.1.4 Demonstratives

Table 4. Demonstratives in Kuwaiti Arabic

proximal sg. masc.	*(hā)ḏa*
proximal sg. fem.	*(hā)ḏi*
proximal pl.	*(hā)ḏōl*
distal sg. masc.	*(hā)ḏāk*
distal sg. fem.	*(hā)ḏīč*
distal pl.	*(hā)ḏōlāk*

When used as an adjective, the demonstrative precedes the noun; used contrastively, it follows. The proximal set of forms can be reduced to a clitic *ha-* prefixed to any defined noun, e.g. *ha-lbēt* 'this house', *ha-rrajājīl* 'these men'.

2.2.1.5 Presentatives

kā- may be prefixed to any suffixed pronoun, e.g. *kā-ni ihni!* 'here I am!' (see Sec. 2.1.3.4 (e)). It is also used with verbs, e.g. *kā-yat issayyāra!* 'here's the car!'. *hā-* followed by a 2nd person suffixed pronoun is also used, especially when handing over something, e.g. *hā-kum liktāb* 'here's the book for you [pl.]!'.

2.2.1.6 Relative pronouns

illi is used.

2.2.1.7 Interrogative pronouns

Typical examples: *šinhu* 'what?'; *min, minhu* 'who?'; *kēf ~ čēf* 'how?'; *čam* 'how much/many?'; *wēn* 'where?'; *mita ~ muta* 'when?'.

š- formations: *š-* + verb, 'what..?', e.g. *š-gilt?* 'what did you say?'; *š* + prep: *š-ḥaggah, š-ʿalēh, š-lēh, š-minnah* 'why, because of what?'; *š* + noun: *š-kiṯir, š-gadd* 'how much/many?'; *š-lōn* 'how?'; 'what kind?'; *š-fī* + pronoun 'what's wrong with...?; *yahu ~ yahi* and *ayhu, ayhi* 'which one?' Other formations borrowed from other dialects are also often heard, e.g. *lēš* 'why?'.

Interrogatives have no fixed sentence position: pragmatic factors determine their position.

2.2.2 Adverbs

Note that only unusual forms or local innovations are listed in this section.

Some adverbs, e.g. *ʿād* 'just, so, then, again', have very many diverse uses and are not easily classifiable.

2.2.2.1 Temporal

alḥīn 'now'; *ha-lḥazza, ḥazzat ḥāḏir* 'at this moment'; *bāčir* 'tomorrow'; *ʿugub bāčir* 'the day after tomorrow'; *ilbārḥa* 'yesterday'; *dōm* 'always'; *(min) iṣṣubḥ* 'in the early morning'; *iḏḏiḥa* 'in the forenoon'; *ilgāyla* 'at mid-day [i.e. approximately 12:00–2:00]; *ilʿaṣir* 'in the late afternoon'; *lawwal* 'in the old days'; *baʿad* 'still, yet'; *abdan* 'totally; ever, never'; *rāyiḥ* 'continuously'; *iljābla* 'tomorrow night'; *illābla* 'the day after tomorrow'; *mbaččir* 'early'; *mṣayyif* 'late'.

2.2.2.2 Local

ihni 'here', *ihnāk* 'there', *ḥadir* 'beneath', *ḥadir fōg* 'upside down', *sīda* 'straight ahead, directly', *ha-ṣṣōb* 'over here'.

2.2.2.3 Manner

ham 'also', *bass* 'only', *čiḏi* 'thus', *wājid ~ wāyid* 'much, very', *zēn* 'well', *killiš* 'completely', *zitāt* 'quickly', *falla* 'well, excellently', *wakād* 'certainly', *ʿasā-* 'hopefully'.

2.2.3 Particles

2.2.3.1 The definite article

The article is *(i)l-* and is assimilated by the sun-letters as in Old Arabic. In a few common phrases it is *al-*, e.g. *alḥīn* 'now'. When prefixed to nouns whose first syllable is an open, unstressed /i/ or /u/, especially if the following vowel is long, it normally become *li-*, e.g. *liktāb*.

2.2.3.2 Indefinite article

Some uses of *wāḥid* have the function of an indefinite article when it follows the noun, e.g. *ʿidna walad wāḥid yigāl lih* 'we once had a boy here called...'; *marra wāḥda iḥna fi ššwēx* 'one time, when we were in Shuwaykh...'.

2.2.3.3 Genitive markers

māl(at) and *ḥagg* are both used as genitive markers, e.g. *issayyāra mālat ʿammi* 'my uncle's car', *ilmuftāḥ ḥagg ilbāb* 'the key to the door', *māl* can be used with inalienable possession, e.g. *ilīdēn mālah* 'his hands'. Both markers have a wide range of other uses.

2.2.3.4 Negative particles

With indicative verbs the normal negative particle is *mā*. It is also used with 'dummy' verbs expressing possession (*ʿind*) and existence (*āku, fīh,* less commonly *hast*), e.g. *mā fīh šay fi lgūṭi* 'there's nothing in the can', *māku čara* 'it can't be helped'. Note also *mā min* in, e.g., *mā min mustašfa fi ḏāk ilwagt* 'there wasn't a hospital at that time'. *maḥḥad* or *mā…ḥad* is used for 'no one'.

With other parts of speech – nouns, adjectives, prepositions – *mu* or *muhu* are used, and (for some speakers) the following variants: *mub, muhūb, hūb, humb* (masc.) and *mahi, hīb, mahīb* (fem.). In negating negatives, e.g. 'I didn't not tell you', one uses *mu* or a variant, e.g. *mub mā aḥibbah, aḥibbah!* 'it's not that I don't like him, I do like him!'. In nominal sentences, and those with participial predicates, *mā* + independent pronoun is an alternative to pronoun + *mu*, etc., e.g. *inta mub rayyāl* or *minta (bi) rayyāl* 'you're not a [real] man', *ana mub rāyiḥ* or *māna bi rāyiḥ* 'I'm not going'.

lā is used for prohibitives and optatives, e.g. *lā yirkab fōg issaṭiḥ!* 'don't let him climb up on the roof!', *lā samaḥ allāh!* 'God forbid!'. It is also used in the sense of 'lest', e.g. *iṭlaʿ lā aḥibsik wasṭ iljilīb!* 'get out or I'll throw [= 'lest I throw'] you down the well!'.

2.2.3.5 Particles to introduce questions

The tag *mu čiḏi* 'isn't that so?' is used.

2.2.3.6 Existentials

āku is the local form, which Kuwait shares with southern Iraq. *fīh* is also common, but *hast* is heard less often than in Bahrain. All are negated with *mā*: see Section 2.2.3.4.

2.2.3.7 Prepositions

li 'to, for'; *bi* (or *ib*) 'with, by means of'; *fi* (or *if*) 'in, on'; *ḥagg* 'to, for'; *ila* 'to, toward'; *ʿala* 'on, against'; *min* 'from'; *ʿan* 'away from'; *fōg* 'on top of, above'; *taḥt* 'under, near'; *ḥadir* 'under, below'; *ṣōb* 'toward'; *wara* 'behind'; *ḥōl* 'around'; *jiddām* ~ *giddām* 'in front of'; *mjābil* ~ *mgābil* 'opposite'; *bēn* 'between'; *gabil* 'before'; *baʿad, ʿugub* 'after'; *bidūn, min dūn* 'without'; *ʿind* 'at, with, according to'; *māl* 'of, relating to'; *wiyya* 'with'; *yamm* 'beside'; *dūn* 'in contrast with, different from'; *gufa* 'dependent upon, incumbent on'; *dāyir madār* 'around'; *miṯil, šikil* 'like'.

2.2.3.8 Conjunctions

wi ~ *u* 'and'; *willa, aw* 'or'; *amma…aw* 'either…or'; *lākin, bass* 'but'; *innamā* 'only, except that'; *inn* 'that'; *yōm, min, lamma, lamman* 'when'; *lēn, ilēn, ilamma* 'whenever, until'; *liʾann* 'because'; *ḥatta, ḥagg* 'so that'; *ʿan lā* 'lest'; *gabil lā, gabil mā* 'before'; *baʿad mā, ʿugub mā* 'after'; *miṯil mā, šikil mā* 'like'; *wēn mā* 'wherever'; *mā dām* ~ *im dām* 'as long as'; *činn-* 'as if'; *mā…illa* 'as soon as, no sooner than'. Conditional conjunctions: *in, iḏa, ila, lēn, lō, (in) čān.*

2.2.3.9 Exclamations

Particles of affirmation and denial are:

'Yes!': *ē! ē naʿam! ajal, ʿajal! bala, mbala!*
'No!': *lā!*

yallah! 'come on!'; *yā rēt!* 'would that…!'; *bass!* 'stop! enough!'; *ʿajal* ~ *ʿayal* 'well…'; *ila…* 'lo and behold!'; *ʿād*, which has many uses, e.g. mild reproach, e.g. *ismaʿ ʿād!* 'just listen, will you!', cajoling, e.g. *rūḥi š-ʿād!* 'go, why don't you?' *ya!* 'what?!' for incredulity, often coupled with dismissal of a proposition; *ax!* used to express pain; *wēl!* to express woe, sorrow; *bwēl!* used to express surprise; *afa!* used to express disgust. *š* + diminutive noun is used to express admiration/surprise, e.g. *š-iḥlēla!* 'what a pretty little thing [she is]!'

2.2.4 Nouns

2.2.4.1 Gender

Feminine by usage: double parts of the body; notes denoting females, e.g. *umm* 'mother', *ʿarūs* 'bride'; names of countries; a few common nouns, e.g. *šams* 'sun', *arḏ̣* 'earth, land', *nār* and *ḏaww* 'fire', *rūḥ* 'soul, spirit', especially those which are reflexes of Old Arabic *-āʾ*, e.g. *kahraba* 'electricity'. *māy* 'water' may be of either gender.

2.2.4.2 Productive patterns

Common patterns are similar to those in Bahraini Arabic. Particularly productive are CaCCāC for occupations, e.g. *ʿayyār* 'swindler'; CaCCāCa for tools, e.g. *ṭaffāya* 'ashtray', *dabbāsa* 'stapler'; and miCCaC(a)/miCCāC, e.g. *mišxal* 'sieve', *midʿāb* 'sewer', *midwisa* 'treadle'.

A few nouns of very local reference (now obsolete) end in the suffix *-ō*, e.g. *ḥalwayō* 'a type of fish', *čaftō* 'keelson' (in maritime terminology).

The *-iyya* suffix is also productive, e.g. *yōmiyya* 'daily wage', *ʿaṣriyya* 'afternoon'.

2.2.4.3 External and internal plural

External plurals for human nouns are mainly formed with *-īn* (masc.) and *-āt* (fem.), the latter also being used for borrowings, e.g. *bāṣ/bāṣāt* 'bus'. *-a* is used for some human nouns of the CaCCāC form, e.g. *ḥaddāda* 'ironmongers'. The *-iyya* suffix is used for some male human nouns, e.g. *drēwil/drēwiliyya* 'driver', especially those ending in the relational *-i*, e.g. *baḥrēniyya* 'Bahrainis' (although the plural in *-īn* is now more common for this type of form). *abu* 'father' forms a plural *ubuhāt.*

Internal plurals are similar to those of the Gulf dialects as a whole, for which see Bahraini Arabic.

2.2.4.4 Diminutives

Diminutive forms are common, e.g. *klēb* 'little dog', *mrēx* 'mist', *ḥlēla* 'sweet little thing', *šnētir* 'chicken pox', *mrēkib* 'little ship', *dwērat ḥamām* 'pigeon loft', and are especially commonly encountered in the names of animals, birds, etc., e.g. *dxēxla* 'type of warbler', *ḏbēbna* (< *ḏubbāna*) 'little fly [the name of a children's game]'. The form with a relational *-i* suffix is also common in animal names, e.g. *zhēwi* 'cockroach', *ṯbēji* 'type of poisonous sea creature with tentacles', *wḏēḥi* 'oryx'.

CaCCūC is used for the diminutive of names, e.g. *ammūn < amīna, laṭṭūf < laṭīfa, xallūd < xālid.* The suffix *-ō* or *-(a)w* is also used for this purpose, e.g. *zambaw* (< *zēnabaw*), *fiḏḏaw* (< *fiḏḏa*).

2.2.4.5 Vocatives

There is a complex system of vocatives involving bipolar address forms, e.g. *yābū-k, yāxū-k, yāxt-ič, yumma-k, yā nasīb-ik*, etc., as well as *(yā) yubba, (yā) yumma.* See Yassin (1977).

2.2.4.6 Adjectives

The following are additional to the nominal patterns:

CaCCān: for example *taʿbān*, pl. *taʿāba* (also *taʿbānīn*) 'tired'.

CaCCāCi: a type of relative adjective, e.g. *rawwāsi* 'method of pearl diving in shallow water in which the diver dives headfirst', *ḏabbābi* 'type of fly-eating bird'.

CaCCūC: intensive pattern, e.g. *akkūl* 'greedy'.

CāCūC: intensive pattern, e.g. *šāġūl* 'hard-working'.

CaCCa: the feminine of some CaCīC and CaCCāC adjectives, e.g. *xanṯa* 'catamite' (< *xanīṯ*), *ġaḏba* 'angry' (< *ġaḏbān*).

CuCāC: plural of some adjectives, e.g. *kubār* 'old'.

CiCCaC: plural of some adjectives, e.g. *ʿittag* 'old', sg. *ʿatīj.*

2.2.4.7 Color and deficiency adjectives

aCCaC (masc.), CaCCa (fem.): e.g. *aswad, sōda* 'black'. The masculine forms of some words in this class are resyllabified because of the *ghawa*-syndrome, e.g. *xaḏar* 'green', *ġatam* 'dumb'. The plural is CuCC or CuCCān, the latter pattern usually designating animate beings: thus *ḥumrān* is (somewhat pejoratively) used to mean 'pink-skinned [people]' (i.e. Westerners and Persians).

2.2.4.8 Elatives

These are formed as in Classical Arabic. The aCCaC pattern has been extended, however, e.g. *ašwa* < *šway* 'better' (as in *ittōz ṣār ašwa* 'the dusty weather has abated'), *ašṭan* < *šēṭān* 'naughtier'.

2.2.5 Numerals

'One' is *wāḥid*, fem. *waḥda.* It precedes the noun as an indefinite article (Sec. 2.2.3.2), and follows it to signify 'one' (as opposed to 'two', etc.), e.g. *sayyāra waḥda* 'one (single) car', *marra waḥda* 'in one go, all at once'.

'Two' is *ṯnēn*, fem. *ṯintēn.* The dual or the plural followed by 'two' can be used, with no difference in meaning, e.g. *bintēn* or *banāt ṯintēn* 'two daughters'.

'Three' to 'ten': there are two systems in operation. The first (probably imported from non-Gulf dialects) uses the masculine form of the numeral with enumerated nouns of either gender, e.g. *xams awlād* 'five boys', *xams banāt* 'five girls'.

In the alternative system, which appears to be the original one, the gender polarity system operates as in Modern Standard Arabic, with no reanalysis of forms with *-t*, e.g. *xamsat awlād, xams banāt.*

Telling the time: the masculine form is normally used, except in *sāʿa ṯintēn* 'two o'clock'.

'Eleven' to 'nineteen': there is considerable variability, but the most common forms are those without the *-ar* ending (contrast this with Bahrain), whether there is an enumerated noun or not, and there is also some variability in the form of the numeral in the numbers 'eleven' and 'twelve', e.g. *idāš ~ iḥdāš ~ iḥdaʿaš sana* 'eleven years', *iṯnāš ~ iṯnaʿaš yōm* 'twelve days'.

'Hundred':	*miya* or *imya*, pl. *miyāt.*
'Thousand':	*alf*, pl. *ulūf. ulūf ilʾulūf* 'thousands upon thousands'
Larger numbers:	*lakk*, pl. *lkūk* signifies a large number (cf. English 'zillion'), for some speakers 10,000, for others 100,000!

Ordinal numbers are regular.

Count nouns: an *-a* suffix is added to the collective, e.g. *xōx* 'peaches', *xōxa* 'a peach', or to a verbal noun to form an instance noun, e.g. *nōm* 'sleep', *nōma* 'a nap, period of sleep'.

2.2.6 Verbs

2.2.6.1 Verbal forms

2.2.6.1.1 Form I

(a) Perfect stems

In verbs of the CaCaC- stem type in Old Arabic, the factor determining the first vowel was, as in the Bedouin-type Eastern Arabian dialects as a whole, originally phonological (see Sec. 2.1.1.2 (a)), thus *kitab, wuṣal* vs. *halak, daxal, ḏ̣arab.* Many of the dialectal CaCaC verbs, however, now have an alternative in CiCaC, e.g. *dixal, ḏ̣irab* as in southern Iraq and Muslim Baghdadi. Verbs of the Old Arabic CaCiC/CaCuC- stem type were originally voweled in the Kuwaiti dialect according to the same phonological rule as the CaCaC- type, thus *šarab, simaʿ*, but again CiCaC is an alternative for virtually any CaCaC- dialectal stem < Old Arabic CaCiC- /CaCuC-. Examples:

Old Arabic	Kuwaiti		
šariba	*šarab ~*	*širab*	'he drank'
zaʿila	*zaʿal ~*	*ziʿal*	'he got upset'
kaṯura	*kiṯar*		'it grew more numerous'
ḍaraba	*ḏ̣arab ~*	*ḏ̣irab*	'he hit'
kataba	*kitab*		'he wrote'
dafaʿa	*difaʿ*		'he pushed'

(b) Imperfect stems

As in the rest of Bedouin-type Eastern Arabic dialects, such as the Bahraini A dialect, if C_2 or C_3 is a guttural, the theme vowel is /a/ and the prefix vowel /i/, e.g. *yigʿad, yiṭbax.* If C_1 is a guttural, the stem vowel is /i/ and the prefix vowel /a/, the form then being resyllabified according to the rule already given, e.g. *yʿarf* < *yaʿrif, yxaṭib* < *yaxṭub.* Non-resyllabified forms (which apparently sound 'more educated') are now common in this dialect. In non-guttural stems, the dialect generally, like the rest of Bedouin-type Eastern Arabic dialects, follows the Old Arabic system: theme vowel /a/ for Old Arabic /a/, otherwise theme vowel /i/ or /u/ depending on the consonants in the C_2 and C_3 position, with the prefix vowel typically opposite in height from the theme vowel, e.g. *yilbas, yišrab, yigdar* but *yabriz, yaḏkir, yaṣbur, yaglub*, where in the last two examples the labial /b/ (as also /f/, /m/) has a backing effect.

2.2.6.1.2 Derived forms

Noteworthy semantic aspects are the following:

Form II is mainly causative and has largely supplanted Form IV, which survives only in a few verbs such as *aṣbaḥ/yiṣbiḥ* 'to get up in the morning'. There are some denominative Form II verbs, e.g. *ṣayyan* 'to stink' < *ṣyāna* 'stagnant dirty water'.

Forms V and VI: in the perfect, the vowel of the *t*-prefix is often elided; in the imperfect, Kuwaiti Arabic normally has *yit-, tit-, nit-* type forms. Form V is often reflexive or passive, e.g. *tbaṭṭaṭ* 'to burst open'. Form VI most often signifies reciprocal action, e.g. *thāwaš* 'to argue with one another', but also often implies gradation or repetition, e.g. *tgāṣar* 'to peter out', *tmāyal* 'to sway, incline'. Form VI has other values also, e.g. *tarāxaṣ* 'to buy cheaply'.

Form VII is the normal means of passivization but is also used for reflexives, e.g. *intaras* ‘to be filled’, *inšāf* ‘to be seen’, but also *inkamaš* ‘to shrink’, *inbaṭaḥ* ‘to collapse, throw oneself down’.

2.2.6.1.3 Internal passive

The internal passive occurs in normal speech only in a few fixed phrases, and usually in the imperfect, e.g. *wāḥid yigāl lih* ‘someone called…’, *wlidt* ‘I was born’, but more commonly in proverbs, e.g. *iḏa ṭigg ilxašim, ihmalat ilˁēn* ‘if the nose is struck, the eye weeps’.

2.2.6.2 Inflection

2.2.6.2.1 Imperfect

Table 5. Inflection of the imperfect in Kuwaiti Arabic

yaktib ‘he writes’	sg.	pl.
3rd pers. masc.	*yaktib*	*yikitbūn*
3rd pers. fem.	*taktib*	
2nd pers. masc.	*taktib*	*tikitbūn*
2nd pers. fem.	*tikitbīn*	
1st pers.	*aktib*	*naktib*
yilbas ‘he wears’	sg.	pl.
3rd pers. masc.	*yilbas*	*yilibsūn*
3rd pers. fem.	*tilbas*	
2nd pers. masc.	*tilbas*	*tilibsūn*
2nd pers. fem.	*tilibsīn*	
1st pers.	*albas*	*nilbas*

These are the canonical syllable structures for verbs that do not have a guttural in C_1 position in the original dialect, but forms like *yaktibūn, taktibīn* are also now freely variant with the forms noted. Verbs with C_1 = guttural routinely have resyllabified forms of the *yˁarf, yˁarfūn, tˁarfīn* type, but forms like *yaˁrif, yaˁrifūn* are also now freely variant with them. It seems that this is a tendency mainly found in educated speakers influenced by non-Gulf dialects and/or Modern Standard Arabic.

Like other eastern Arabian dialects, Kuwaiti prefixes *b-* to imperfect verbs to express proximate intent. *gāˁid* is used to express continuous or iterative processes.

The imperative: *iktib, kitbi* (or *kitbay*), *kitbu* (or *kitbaw*), the bracketed forms being more typical of older, less educated speakers, especially women. The negative imperative is *la tiktib/tikitbi/tikitbu*.

2.2.6.2.2 Perfect

Table 6. Inflection of the perfect in Kuwaiti Arabic

kitab ‘he wrote’	sg.	pl.
3rd pers. masc.	*kitab*	*ktibaw*
3rd pers. fem.	*ktibat*	
2nd pers. masc.	*kitabt*	*kitabtaw*
2nd pers. fem.	*kitabti ~ kitabtay*	
1st pers.	*kitabt*	*kitabna*

These syllable shapes apply to all Form I perfect verbs. In the 3rd person, forms of the *kitbat, kitbaw* type are also common and are syllabically similar to those of neighboring southern Iraq and Muslim Baghdadi.

2.2.6.3 Participles

The morphology of participles is similar to Old Arabic. CvCCān is productive for stative verbs, e.g. *ḏ̣amyān* ‘thirsty’, *bardān* ‘(feeling) cold’, *zaˁlān* ‘annoyed, upset’.

2.2.6.4 Verbal nouns

Form I: The patterns are similar to those found in other dialects (see → Bahraini Arabic for some Gulf examples).

Derived form verbal nouns that are original to the dialect (and attested in Old Arabic) but now obsolescent are:

Form II:	tvCCūC:	*taˁlūm* ‘teaching, education’
	tiCCāC:	*tisgām* ‘advance paid to pearl divers’
	CvCCāC:	*giffāl* ‘end of the pearling season’
Form III:	m(u)CāCaC:	*mˁāfar* ‘wrestling, throwing to the ground’, *mxāzar* ‘staring someone out’

Modern Standard Arabic-type verbal nouns of Forms II, III, IV, V, VI, VII, VIII, X occur, but less often in uneducated speech.

2.2.7 Weak verbs

2.2.7.1 Geminate verbs

Perfect forms with consonant-initial suffixes are on the usual pattern CaCCēC, e.g. *dazzēt* ‘I pushed, sent’. There is contraction in the active participle of Form I, e.g. *mārr* ‘passer-by’. But in Forms III and VI, there is no vowel contraction, e.g. *yitḥājaj* 'he gets into an argument'.

2.2.7.2 Verbs Iʾ

In the perfect of Form I, some verbs lose the initial /ʔ/, e.g. *akal, axaḏ*, and behave like strong verbs. *akal* and *axaḏ* have alternatives *kal* and *xaḏ* which behave like doubled verbs, although these forms are less common in educated speech. In all cases, the imperfect is of the form yāCvC, e.g. *yākil, yāxiḏ* and the active participle *mākil, māxiḏ* and the passive participle *mākūl, māxūḏ*. In Forms II and III, /ʔ/ has either been lost, e.g. *amman* 'to trust', or been replaced by a semivowel, e.g. *waxxar* 'to move back, get out of the way' (*axxar* has the Modern Standard Arabic sense of 'to postpone'). In the first case, in the imperfect, it reappears: *yiʾaḏḏin, yiʾakkil*. In Form VII it is preserved, e.g. *inʾaxaḏ, inʾakal*, but in Form X it is lost, e.g. *istānas* 'to be content' in the typically Kuwaiti greeting *mistānis?*

Form I imperatives: *ikil, ikli, iklaw*.

2.2.7.3 Verbs *Iw*

The perfect verb is formed as for the normal Form I, with 3rd person variants of the two types *wṣilat ~ wuṣlat*. The /w/ is preserved in the Form I imperfect, normally becoming *yō-* or *yū-* e.g. *yōguf, yōṣal, yūʿa*, and in the imperative, e.g. *ōguf!* 'stop!' However, some speakers have a *yā-* prefix (which also occurs in parts of Najd), as in the proverb *allāh yāgāh* 'may God protect him'.

2.2.7.4 Verbs II*w/y*

The perfect and imperfect paradigms are syllabically similar to those of most Arabic dialects. The vowel of the Form I imperative is always long: *gūl* 'say', etc. In the passive participle, /y/ and /w/ are treated as strong consonants (with /w/ > /y/), e.g. *madyūs* 'trodden on', *mabyūʿ* 'sold', *mašyūf* 'seen', *mabyūg* 'stolen'. The Form VIII perfect behaves as Form I, e.g. *iḥtijt* 'I needed', *gilt* 'I said'.

2.2.7.5 Verbs III*w/y*

There are two types: the /a/ type, e.g. *nisa, yinsa*, and the /i/ type, e.g. *miša, yimši*. Verbs which in Old Arabic had final *hamza* lost it and were absorbed into the /a/ type, e.g. *gara, yigra*, and Old Arabic III*w* verbs have been absorbed by the /i/ type, e.g. *tala/ yatli* 'to follow'. One verb, *baġa* 'to want', has alternative forms in the imperfect: *yabġi/ yabbi*.

Like the Bahraini dialects, the Kuwaiti dialect has masculine imperatives which lack a final *-i*, e.g. *ʿaṭ* 'give!', *saww* 'do!' *taġadd* 'eat lunch!'. It also, like many other Gulf and Najdi dialects, allows iCC imperative forms in form I, e.g. *ibg* 'stay!', *imš* 'go!', *isg* 'water!'. Other verbs, somewhat unpredictably, may have an epenthetic vowel, e.g. *imiš* (variant of *imš*) 'go!', *igir* 'read!', *irim* 'throw!', *išiw* 'roast!'.

In the imperfect, the endings of the suffixed forms of the imperfect in /a/ type verbs are *-ēn* and *-ōn*, e.g. *tinsēn, tinsōn*.

2.2.7.6 Irregular verbs

The verb *ja ~ ya* 'to come': this verb is nowadays often encountered with /j/ than with /y/, but /y/ is the 'original' form for Kuwait City:

Table 7. Inflection of the verb *ja ~ ya* 'to come' in Kuwaiti Arabic

singular	perfect	imperfect
3rd pers. masc.	*ya*	*iyi*
3rd pers. fem.	*yat ~ yāt*	*tyi*
2nd pers. masc.	*yīt*	*tyi*
2nd pers. fem.	*yīti*	*tyīn*
1st pers.	*yīt*	*ayi*
plural		
3rd pers.	*yaw*	*iyūn*
2nd pers.	*yītaw*	*tyūn*
1st pers.	*yīna*	*nyi*

2.2.8 Quadriliterals

Both Forms I and II are common. The main types: reduplicatives (some mimetic), e.g. *ġarġar* 'to gargle'; echoic, mimetic in which C_2 = /l, r, n, m/, e.g. *bambaʿ* 'to bleat'; C_2 = /w/: color verbs and bodily states (replacing Old Arabic Form IX), e.g. *bōyaḍ* 'to be whitish, go white', *fōšaḥ* 'to be bowlegged', and denominatives, e.g. *sōlaf* 'to chat' < *sālfa* 'matter, affair', with reciprocal pattern *tisōlaf* 'to chat to each other'; C_2 =/y/: with a *t-* prefix, denoting affectations of one kind or another, e.g. *tšēḥaṭ* 'to claim superiority, put on airs', *tṣēmax* 'to feign deafness', *tlēġab* 'to butt in on a conversation and ruin it' (cf. → Bahraini Arabic, → Omani Arabic); C_2 = /r, n/ inserted into a triliteral root, e.g. *šarbak* 'to ensnare' < *šabak* 'net', *txarbaṭ* 'to get confused' < *xabaṭ* 'to beat'. Others fall into no particular pattern: *barṭal* 'to bribe', *tgašmar* 'to joke, play tricks'. Apart from the above, there are denominatives, some formed from foreign borrowings, e.g. *kansal* 'to cancel', *tbančar* 'to get a puncture', both the latter from English (cancel, puncture).

Inflection, participle formation, and verbal nouns (where they exist) conform to the patterns already described, e.g. verbal nouns: *ġarġara* 'gargling', *mfōšaḥ* 'state of being bowlegged', *kilfāt* 'caulking' (< *kalfat* 'to caulk [a boat])'.

Bibliographical references

Ḥanafī, J. al- 1964. *Muʿjam al-ʾalfāḏ̣ al-kuwaytiyya.* Baghdad.

Ingham, Bruce. 1982. *Northeast Arabian dialects.* London: Kegan Paul International.

Johnstone, T.M. 1961. "Some characteristics of the Dōsiri dialect of Arabic as spoken in Kuwait". *Bulletin of the School of Oriental and African Studies* 24.249–297.

——. 1964. "Further studies on the Dōsiri dialect of Arabic as spoken in Kuwait". *Bulletin of the School of Oriental and African Studies* 27.77–113.

——. 1967. *Eastern Arabian dialect studies.* London: Oxford University Press.

Maṭar, ʿAbd al-ʿAzīz A. 1969. *Xaṣāʾiṣ al-lahja al-kuwaytiyya.* Kuwait.

——. 1970. *Min ʾasrār al-luġa al-kuwaytiyya.* Kuwait.

Sabʿān, L.K. as- 1983. *Taṭawwur al-lahja al-kuwaytiyya.* Kuwait.

Yassin, M. Aziz F. 1977. "Bi-polar terms of address in Kuwaiti Arabic". *Bulletin of the School of Oriental and African Studies* 40.297–330.

Clive Holes (University of Oxford)

L

Labiovelarization

Labiovelarization may be described as a rounding of the lips, accompanied by the raising of the back of the tongue; it may be analyzed as a case of secondary combined articulation or of co-articulation. This phenomenon is common to all → Afro-Asiatic (Hamito-Semitic) languages.

Cantineau (1960:30) mentions what he calls "l'affrication des labiales b, m, f", which can be uttered with "un w furtif, 'spirante de passage'". This appears mainly before vowels with "une ouverture maxima des lèvres", whereas the labials need "une fermeture des lèvres". The contradiction leads to a mechanic "son de passage", used as a transition.

In the Maghreb, one finds two types of labiovelarization, whose origins are probably quite different. One is a labialization of velars, which clearly comes from contact with Berber and is one of the discriminants used in the comparative dialectology of Maghrebi dialects (*sŭkkāṛ* vs. *səkk*^{w}*āṛ*). The other type, closer to what Cantineau described and used in a wider zone, corresponds to a velarization of labials, resulting from the contact of a labial or front consonant with the semivowel *w*, generally followed by a long vowel (*m* + *w*, *b* + *w*, *f* + *w*), with simultaneous velarization and doubling, together with rounding: *bwība* > *ḅḅ*^{w}*ība* (diminutive of *bāb* 'door'), *mwīma* > *ṃṃ*^{w}*īma* (diminutive of *mma* 'mother').

1. Labialization of velars

Labialization of velars is probably borrowed from Berber and thus remains a trait typical of the Maghreb. In the environment of velars, especially when the following phoneme is a front consonant or a vowel, there is a labialization of the velars.

In Morocco, this is mostly present in the southern dialects, such as that of Marrakesh and its region, which has a heavy Berber influence; it is also found commonly in Casablanca, where a number of southerners now live. In the northern region, this phenomenon is much scarcer. Morocco, for that matter, opposes two types of dialects: those with strong labiovelarization in the contact of velars, uvulars, and sometimes pharyngeals (/k, g, q, x, ġ, ḥ/, as described in Elmdari 1997:6 for Marrakesh), and those which have developed an opposition between two short vowels /ə/ : /ŭ/ (in the region of Rabat, Meknes, and Fes). This affects the phonological system of Moroccan Arabic, which varies from four to five vocalic phonemes, according to the presence or absence of a short /ŭ/ (to which must be added the long or medium phonemes /ā, ī, ū/).

Thus, regional differences appear explicitly in offering either an opposition between two short vowels /ə/ : /ŭ/ in northern dialects, with a long series of minimal pairs (see Table 1) or an opposition between only one short vowel, /ə/, and labiovelarization, as in Marrakesh (see Table 2).

Table 1. Opposition pairs /ə/ : /ŭ/ in northern dialects

/ə/	/ŭ/
rəkba 'the fact of riding'	*rŭkba* 'knee'
xəḍṛa 'green [fem.]'	*xŭḍṛa* 'vegetable'
dxəl 'he came in'	*dxŭl* 'the fact of coming in'
kḥəl 'black'	*kḥŭl* 'kohl [eye makeup]'

Table 2. Opposition pairs /ə/ : labiovelarization in Marrakesh

/ə/	labiovelarization
rəkba 'the fact of riding'	*rəkʷba* 'knee'
xəḍṛa 'green [fem.]'	*xʷəḍṛa* 'vegetable'
dxəl 'he came in'	*dxʷəl* 'the fact of coming in'
kḥəl 'black'	*kḥʷəl* 'kohl [eye makeup]'

Other pronunciations can appear without a phonemic opposition, denoting either a Marrakshi or rural accent (see Table 3):

Table 3. Regional variants [ə] : [ŭ]

Northern	Marrakshi/rural
sŭkkāṛ	*səkkʷāṛ* 'sugar'
ġzāl	*ġʷzāl* 'gazelle; handsome'
ṣġāṛ	*ṣġʷāṛ* 'small [pl.]'
ġṛāb	*ġʷṛāb* 'crows'
hakka	*hakkʷa* 'like this'
qtəl	*qʷtəl* 'he killed'
kbāṛ	*kʷbāṛ* 'big [pl.]'
xəbbīza	*xəbbʷīza* 'mallow [bot.]'
gbīla	*gʷbīla* 'before'
ṣḥāḥ	*ṣḥʷāḥ* 'strong; real [pl.]'
tqāl	*tqʷāl* 'heavy [pl.]'

2. VELARIZATION OF LABIALS

Labialization of velars seems to occur essentially in the Maghreb, with the presence of a Berber substratum. The second type, velarization of labials, occurs in all Arabophone areas. It takes place when labials such as /b, m, f/ encounter the labiovelar semiconsonant /w/, followed by a long vowel, which could be presented as a case of velarization of labials. This second type is what Cantineau (1960:30) describes for the nomadic dialects of North Africa and in the Orient.

In Morocco, velarization of labials can also distinguish between regional/dialectal accents (see Harrell 1962:10), when a word like *bwība/mwīma* is realized as *ḅḅʷība/ṃṃʷīma*, implying an important change at the level of the syllable, *bw* > *ḅḅʷ*: a gemination of the initial consonant, which is also pharyngealized and rounded (see Table 4).

Table 4. Velarization of labials

Northern	Marrakshi/rural
bwība	*ḅḅʷība* 'door [dim.]'
mwīma	*ṃṃʷīma* 'mother [dim.]'
mwīha	*ṃṃʷīha* 'a little water [dim.]'
mwāgən	*ṃṃʷāgən* 'wristwatches'
fwāṭ	*f̣f̣ʷāṭ* 'towels'
mwālīn (pl. of *mūl*)	*ṃṃʷālīn* 'owners'

Some words beginning with labials are always pronounced labiovelarized, especially names of parents, used as terms of address or in construct state, e.g. *ḅḅʷa* 'father, my father'; *ṃṃʷ-i* 'mother, my mother' (other dialects have *yəmma* instead); *ṃṃʷ-uk/-ha/-u* 'your/her/his mother' (can be used in insults); *ḅḅʷa-k/-h/-ha* 'your/his/her father' (used in insults).

As for other dialects, Cantineau (1946:92–95), indicates similar facts for the Horan for /b/ and /m/ (*ᵐmāˤīn*, *ᵐmʷāˤīn*). For the Fezzan region, as described by Ph. Marçais, the first type exists rarely, only for sedentaries, in Gorda (*ṣᵘġāṛ*; see Marçais 2001:6). For the second type, there are many occurrences: *fi-l-ᵘᵐmʷĕ́yya* (Marçais 2001:12), *ᵐmwāli* (Marçais 2001:32), *fī-ᵐmwēˤīn* (Marçais 2001:52), and for the nomads of Brak, the 3rd person plural pronoun is *hŭ́mmʷa*. D. Cohen (1963:4) describes velarization of labials for Ḥassāniyya, in *ḅḅʷēʸb* (the diminutive of *bāb*) and *ṃṃʷassax* 'dirty'.

To sum up the two cases, one (case 2, velarization of labials) has been described for the entire Arabophone area, the other (case 1, labialization of velars) seems to appear only on a Berber substratum. Both phenomena are variants inside an area or a country, pointing to a particular dialectal pronunciation.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Cantineau, Jean. 1946. *Les parlers arabes du Horan*. Paris: Klincksieck.

——. 1960. *Cours de phonétique arabe*. Paris: Klincksieck.

Caubet, Dominique. 2003. "Les parlers arabes nomades et sédentaires du *Fezzân*, d'après William et Philippe Marçais". *Approaches to Arabic dialects: A collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday*, ed. Martine Haak, Rudolf de Jong, and Kees Versteegh, 67-96. Leiden: E.J. Brill.

Cohen, David. 1963. *Le dialecte arabe ḥassānīya de Mauritanie*. Paris: Klincksieck.

Elmdari, Fouad. 1997. "La labiovélarisation dans le parler de Marrakech (Maroc), analyse en éléments". Paper presented at the 1st Congrès Chamito-sémitique de Fès. [Available at www.elmdari.recherche.free.fr.]

——. 1999. *Aspect phonologique et morphologique du parler de Marrakech*. Thèse de doctorat, Paris VIII. [Available at www.elmdari.recherche.free.fr.]

Harrell, Richard S. 1962. *A short reference grammar of Moroccan Arabic*. Washington, D.C.: Georgetown University Press.

Marçais, Philippe. 2001. *Parlers arabes du Fezzan*. Ed. Aubert Martin, Laurence Denooz, and Dominique Caubet. Liège: Université de Liège/INALCO.

DOMINIQUE CAUBET (INALCO)

Lafḏ

1. PRELIMINARY REMARKS

Etymologically, the term *lafḏ* is a *maṣdar*, i.e. a verbal noun. Originally it meant 'to spit, reject, vomit' and, in the specific context of speech activities, 'to emit words, to utter'. In practice, there has been a shift from action to result and hence from a verbal value of the *maṣdar* to a nominal one. Taken substantively, *lafḏ* is not to be understood as a singular noun but rather as a collective one. This collective and consequently generic meaning implies a globalizing comprehension of the utterance perceived as a whole, as opposed to the words that make it up. This means that *lafḏ* must not be understood as referring to an isolated word but rather to the global result of the uttering process. When single units making up this whole are intended, one uses the singulative form *lafḏa* 'term, word' and its plural *ʾalfāḏ*.

In its original meaning, nothing suggests that *lafḏ* only refers to the *signifier* side of utterances: on the contrary, it can be shown that the term → *maʿnā*, with which it is generally contrasted, did not refer primarily to the signifier side of an utterance but rather to the communicative intention which motivated it. This implies that *lafḏ* covered all specifically linguistic aspects of an utterance (*kalām*). As a matter of fact, one cannot understand the actual use of the word *lafḏ* in a great number of texts, linguistic or otherwise, unless it is understood as referring to a global linguistic entity, with its own rules, its own logic, and even its own 'meaning'.

In its original meaning, *lafḏ* does not refer to the concept of 'signifier', understood as the mere phonetic (or more generally, material) medium distinct from the semantic content which it conveys. Rather, it refers to the linguistic sign as a whole. This is why one finds in some ancient texts that grammar is defined as an 'art of *lafḏ*' (*ṣināʿa lafḏiyya*), i.e. a discipline which studies the global properties of the linguistic sign. It is from this point of view that this entry presents the main technical uses of the term *lafḏ* (but also, inevitably, of its conceptual counterpart *maʿnā*) in different fields of Arabic grammar.

2. *LAFḎ* (AND *MAʿNĀ*) IN MORPHOPHONOLOGY (*TAṢRĪF*)

Taṣrīf is one of the two basic components of traditional Arabic grammar, the other one being *naḥw*. The object of *taṣrīf* (→ *ṣarf*) is, on the one hand, the exhaustive description of the nominal and verbal base forms of the language, and on the other, the study of the changes which these bases may undergo in the course of inflectional or derivational processes. Traditionally, these two aspects of *taṣrīf* are carefully distinguished in a way which crucially involves the *lafḏ*/*maʿnā* distinction. The Andalusian grammarian Ibn ʿUṣfūr (d. 669/1270), for instance, presents this question in his great treatise *al-Mumtiʿ fī t-taṣrīf* as follows:

> *Taṣrīf* is divided into two parts: the first consists in modifying the form of a word in order to express different meanings, as in *ḍaraba*, *ḍarraba*, *taḍarraba, taḍāraba*, or *iḍṭaraba*: on the word formed on *ḍ*, *r*, and *b*, that is *ḍarb*, these different forms have been built to express different meanings.... The other part of *taṣrīf* [studies] the change in the base form of the word, without

> this change being an indication of a new meaning that would affect the word, like the change from *qawala* to *qāla*. Don't you see that this has not occurred in order to indicate a meaning other than the one *qawala*, which is the base form, would have indicated, had it been used? (*Mumtiʿ* 31–32)

After characterizing the first part of *taṣrīf*, which ultimately consists in making a complete inventory of all the nominal and verbal bases of the language, together with a semantic description thereof, Ibn ʿUṣfūr stipulates that this part is not generally the object of detailed studies in *taṣrīf* treatises and that his work will be no exception to this habit. This is tantamount to saying that in practice *taṣrīf* only deals with those changes in *lafḏ̣* which do not correlate with a change in *maʿnā*. As a matter of fact, *taṣrīf* treatises very seldom refer to semantic considerations, at least explicitly.

On closer scrutiny, however, the *lafḏ̣/maʿnā* distinction does have a methodological role, be it only implicit, in the study of *taṣrīf*. For example, in the above-cited text of Ibn ʿUṣfūr, the mere fact of stating that forms such as *ḍaraba*, *ḍarraba*, *taḍarraba*, *taḍāraba*, and *iḍṭaraba* are somehow 'derived' from the word *ḍarb*, and that each form has a different meaning, implies that all these words belong to the same semantic family and that each one differs from the others semantically in a specific way, as a result of its change in form. The invariable *maʿnā* which makes it possible to consider this series of words as related is linked with the invariable *lafḏ̣*, manifested by the persisting three radicals *ḍ*, *r*, and *b*, which show up in each word of this series. It is the "semantic load common to all words derived from the same root" according to Bohas (1984:27), who proposes to identify it as 'MAʿNĀ I'. As to the *maʿānī* (plural of *maʿnā*) which result from the changes in form of words of the same root, they are considered by the specialists of *taṣrīf* as semantic-grammatical meanings attached to the different verbal and nominal patterns, for instance *ḍaraba*, in which the verbal pattern expresses the 'past', or *maḍrib*, in which the nominal pattern expresses the 'name of place'. Bohas (1984:27) proposes to identify this level of semantic content as the 'MAʿNĀ II'.

Admittedly, this type of semantic analysis remains rather minimal in the works of the specialists of *taṣrīf*. It plays a crucial role, however, in some types of technical reasoning. Such is the case, for instance, when they say that a verb like *qāla* has *qawala* as its base and not a conceivable *qawula*, because it is transitive while all verbs with a FAʿULA pattern are intransitive (Bohas 1984:73ff.). Similarly, when a given *lafḏ̣* exhibits two *maʿnā*s normally related to two different patterns, as for example *muxtār*, which may be understood either as a name of agent (*ism fāʿil*) 'choosing' or as a name of patient (*ism mafʿūl*) 'chosen', they will posit at the base level (*ʾaṣl*) two different representations matching the normal constructions (in the above case, /muxtayir/ and /muxtayar/, respectively), and devise morphophonological rules applying to these basic representations in order to generate the ambiguous *lafḏ̣*. This clearly implies that there is a postulated regular correspondence between *lafḏ̣* and *maʿnā*. Sometimes they might even adduce that a given morphophonological process is blocked, precisely in order to avoid generating an ambiguous *lafḏ̣* (*li-manʿ al-iltibās*).

All this suggests that even in a field apparently indifferent to meaning, such as *taṣrīf*, semantic considerations, however minimal, do play an indispensable role.

3. *Lafḏ̣* (and *maʿnā*) in syntax (*naḥw*)

Sībawayhi's (d. ca. 177/793) *Kitāb*, the oldest Arabic grammatical treatise to have reached us, refers in its introduction (I, 24) to the *lafḏ̣/maʿnā* opposition, one of its chapters bearing the title "Chapter of [the relationship of] form to meanings" (*Bāb al-lafḏ̣ li-l-maʿānī*). As a matter of fact, this rather brief chapter only refers to ideas current at the time and according to which there may be differences in form expressing differences in meaning (the default case), differences in form without differences in meaning (i.e. synonymy), and differences in meaning without differences in form (i.e. homonymy). In spite of this, Sībawayhi does not seem to make any technical use of these distinctions in the rest of his book.

The chapter devoted, in the same introduction, to 'the changes in word endings in Arabic' (*majārī ʾawāxir al-kalim min al-ʿarabiyya*) is in this respect very different. This chapter presents a systematic and exhaustive description of the various endings, nominal and verbal, and proposes a different terminology for the final short vowels according to whether they are lexical vowels or case markers. This chapter somehow sets the tone for the rest of the book, where Sībawayhi continually discusses the case endings of inflected words in an attempt to justify them. And yet, the word *maʿnā* is far from absent in the *Kitāb*: according to Troupeau (1976) there are 891 occurrences of this word in the singular and 19 in the plural, not counting the multiple uses of the verb *ʿanā/yaʿnī*. But as Carter (1968) convincingly argues, almost all these uses refer to the notion of 'syntactic function' of a form or construction. A typical example from the *Kitāb*, in which the *lafḍ/maʿnā* pair is used as well, is the following. Sībawayhi says about doubly transitive verbs: "You may say *kasawtu zaydan ṯawban* 'I gave Zayd [acc.] a garment [acc.]', assigning a second object [to the verb], and you may say *kusiya zaydun ṯawban* 'Zayd [nom.] was given a garment [acc.]', without assigning [the verb] a second object, as the first [argument of the verb] has the status of an accusative, because the *maʿnā* is identical even if the *lafḍ* [of the verb's first argument] is in the nominative" (*Kitāb* I, 42). The point here is to show that in the doubly transitive verb in the objective (passive) form one cannot mention the agent, and therefore, the first object is promoted to the function of syntactic subject. It is clear that for Sībawayhi the *maʿnā* is here nothing more than the syntactic function, in this case of the verb's arguments.

Accounting for the assignment of cases remains the main concern of Arabic grammar at all steps of its development, and this largely justifies its traditional characterization as an 'art of *lafḍ*' (*ṣināʿa lafḍiyya*). This does not mean, however, that it has not paid any attention to questions relating to *maʿnā* and its relationship with *lafḍ*. But in fact, in the framework of *naḥw*, semantic considerations were always subordinated to the analysis of case assignment and were only taken into account in a sporadic and sometimes even contradictory manner, as illustrated below.

This is all the more striking as the writings of the first theoreticians of the fundamentals of grammar (*ʾuṣūl an-naḥw*) do present ideas that give the impression that they intended to give semantic considerations a real place in their systematization. For example, az-Zajjājī (d. 340/951), in one of the first treatises on the subject, the *ʾĪḍāḥ fī ʿilal an-naḥw*, explains that the reason why there are cases is that "inasmuch as nouns undergo differences in meaning so as to be sometimes subjects, and sometimes objects, specified or specifiers, while nothing indicates these meanings in their invariable forms, the case endings were made to indicate such meanings in them" (*ʾĪḍāḥ* 69). This was admittedly a good starting point for the development of a real 'grammatical semantics'. But the increasingly formalistic approach to the discipline in the course of its development led to a rejection of such semantic justifications, which were regarded as inadequate in the new framework of ideas. The new trend is well illustrated by Ibn al-ʾAnbārī (d. 577/1181), who argues in his celebrated book *al-ʾInṣāf* (81) that the idea of a correlation between case ending and semantic function is false since in the objective (passive) construction (*mabnī li-l-mafʿūl*) the semantic 'patient' occupies the syntactic position of the subject and takes the nominative. Such arguments led to a complete rejection of 'semantic-minded' interpretations of case endings and imposed the idea that their sole acceptable justification was that they were imposed by a 'case assigner' (*ʿāmil*), notwithstanding the semantic function of the element concerned.

The theory of case assignment (*naḍariyyat al-ʿamal*) has had a 'blocking effect' on other aspects of morphosyntactic analysis with respect to the integration of the semantic dimension (→ *ʿamal*). One obvious case has to do with the fact that this theory stipulates that only nouns are 'originally' (*fī l-ʾaṣl*) entitled to case marking. When it came to accounting for the verb (modal) endings of the imperfect, which Arabic grammar considers to be case endings, these 'verb cases' were explained by the formal 'resemblance' (*muḍāraʿa*) of the imperfect with certain nominal forms, hence the name of 'resembling verb' (*fiʿl muḍāriʿ*) given to this verbal paradigm. But the postulate according to which only nouns are rightfully entitled to declensional marks led to the idea that as far as verbs were concerned, these marks were purely formal (*lafḍiyya*) and devoid of any semantic content. This is why the great grammarian az-Zamaxšarī (d. 538/1144), when tackling the study of the imperfect verb endings in his well-known treatise *al-Mufaṣṣal*, states: "On the declensional forms of the resembling [verb]: these are the nominative, the accusative, and the apocopated; unlike the declensional forms of the noun, these forms indicate no meaning, because the verb has no primacy in case marking" (*Mufaṣṣal* 244–245). It is interesting to observe that this declaration of allegiance to the dogmas of the theory of → *ʿamal* does not prevent az-Zamaxšarī, as a true linguist, from perceiving and citing many cases in which a 'nominative' (i.e. an indicative) contrasts with an 'accusative' (i.e. a subjunctive) with a difference in semantic interpretation. Such is the case, for instance, in the contrasting pair *sirtu ḥattā ʾadxula-hā* 'I walked so as to enter it' and *sirtu ḥattā ʾadxulu-hā* 'I walked so that I entered it', the first utterance indicating the aim, the second its fulfilment. Surprisingly, he introduces the discussion of these cases by the following preamble: "It is not compulsory to put the verb in the accusative in such contexts. On the contrary, doing otherwise has an appreciable effect on the meaning that depends on the case marking". The least that can be said is that this contradicts the dogmatic statement introducing the study of verb endings.

There is another aspect of syntactic analysis where considerations articulating *lafḍ* and *maʿnā* came into play in Arabic grammar, but following rather different modalities. That is when the linguistic form of an utterance, its *lafḍ*, exhibits a case ending which cannot be attributed to a visible case assigner (*ʿāmil*). In such a case, it was necessary to postulate an underlying representation of that linguistic form, which reveals the reason for that case assignment. This operation, which played a fundamental part in the general workings of traditional Arabic grammar, was called → *taqdīr* 'postulation', as it consisted in postulating (*qaddara*) in the underlying representation of the sentence the presence of a case assigner, invisible in the phonetic form. In some cases, this operator was simply elided in the phonetic realization of the utterance and could, without difficulty, be restored even by a speaker with no particular grammatical sophistication. This is the case, for example, for certain 'emotional' utterances, in which one utters only a noun in the accusative, e.g. *al-kilāba!* 'the dogs [acc.]!', as a form of warning. Everyone would admit that the speaker had in mind something like *iḥḏar al-kilāba!* 'beware of the dogs!', and the grammarian did nothing more, in his *taqdīr*, than formalize this truism. In other cases, however, it was much more difficult to admit that the grammarian's *taqdīr* really reflected what went on in the speaker's head. In the classic example *zaydan ḍarabtu-hu* 'Zayd [acc.], I hit him', in which the noun bears an accusative ending notwithstanding the fact that the object position of the verb is saturated by an object pronoun, the theory stipulated that the underlying form of the utterance is *ḍarabtu zaydan ḍarabtu-hu* 'I hit Zayd [acc.], I hit him', and that the first occurrence of the verb, postulated to account for the accusative in the noun is later 'obligatorily elided'. In this case, the *taqdīr* no longer consists in restoring in the phonetic form something which has simply been elided in speech, but in positing in the underlying syntactic structure of the utterance something which is explicitly characterized as impossible to say. At this point of its development, traditional Arabic grammar may be said to have drifted away from common sense, so much so that as great a grammarian as Ibn Jinnī (d. 392/1002) even felt the need to warn the beginner against the temptation to cling too much to the meaning of an utterance when attempting to build a *taqdīr* to account for a case ending. For example, in an utterance like *ʾahla* [acc.]-*ka wa-l-layla* [acc.] 'go back to your family before the night', he warns against the appeal of such a *taqdīr* as *ilḥaq ʾahla-ka qabla l-layli* 'join your parents before the night', which, although quite near to the intended *maʿnā*, would leave unexplained the accusative present in the *lafḍ* and, even worse, could induce the use of a genitive instead of the accusative, which would be a gross violation of the linguistic usage of the Arabs. The expert grammarian will propose as the underlying representation of the utterance in question something like *ilḥaq ʾahla-ka wa-sābiq al-layla* 'join your parents and precede night', which alone can explain the structure of this utterance's *lafḍ*.

In practice, the technique of *taqdīr* not only took grammar away from the ordinary intuitions of the speakers, it also introduced, within that discipline itself, more or less deep divergences between grammarians, insofar as numerous 'postulations' proposed by one of them could differ from those advocated by another. It is not surprising then that many methodological writings bear on this question in an attempt at constraining *taqdīr*.

It is also necessary to mention another aspect of the ancient grammarians' method that proved harmful to a meticulous and methodical study of the relationship between *lafḏ* and *maʿnā*, namely the tendency, already present in Sībawayhi's *Kitāb*, to consider a given structure that is analyzed as 'equivalent' to another, better-known structure. A typical case is the well-known example from the *Qurʾān* (2/184) in which the string *wa-ʾan taṣūmū xayrun la-kum* 'and that you should fast is good for you' is said to be equivalent to *ṣiyāmu-kum xayrun la-kum* 'your fasting is good for you', so that the phrase '*ʾan* + verb' is equivalent in this context to a noun in the nominative. This analysis, which may have some validity at a very general level (comparable to the distributional method), has nevertheless the very negative effect of suggesting that two different *lafḏ*s have the same *maʿnā*. No wonder that al-Jurjānī (d. 471/1078), the founder of grammatical semantics, fought energetically against the implications of this kind of approach so widely used among grammarians. In this respect, he writes in his epoch-making book, *Dalāʾil al-ʾiʿjāz*:

> Do not be fooled by the fact that, when treating of questions bearing on topic (*mubtadaʾ* → *ibtidāʾ*) and comment (→ *xabar*), we postulate that the verb in such structures is the equivalent of a noun, as when we say that *zaydun yaqūmu* 'Zayd stands up' is equivalent to *zaydun qāʾimun* 'Zayd is standing up'. This by no means entails that the *maʿnā* is identical in the two structures, so that there would remain no difference. If this were so, you would not have a verb in one case and a noun in the other, but [you would have] in both either two verbs or two nouns. (*Dalāʾil* 136)

In another passage of the same book, he criticizes the grammarians' view according to which, if the two constituents of a nominal sentence are both defined, as in *zaydun al-munṭaliqu* 'Zayd is the one leaving', you may assume either word to be the *mubtadaʾ*, and treat *al-munṭaliqu zaydun* 'the one leaving is Zayd' as exactly equivalent to the first utterance. Al-Jurjānī explains that the semantic implications of the two utterances are completely different and that regarding them as equivalent entails a complete disregard of *maʿnā*. He even criticizes Sībawayhi in this respect, since all he said about word order in this sentence was a vague formula on the 'attention' (*ʿināya*) the speaker gives the different terms involved. He even goes so far as to say that this 'devil-inspired negligence' closed the doors of knowledge to grammarians!

One last aspect of the grammarians' handling of the *lafḏ/maʿnā* relationship must be mentioned here, albeit in a very cursory way: it has to do with the theory of 'expletive' elements. One example will help us define the concept better than a long theoretical exposition. Grammarians say, for instance, that *mā qāma min ʾaḥadin* is equivalent to *mā qāma ʾaḥadun* 'no one stood up' because in the first utterance the particle *min* is, as it were, superfluous (*zāʾida*) from a strictly grammatical point of view, even if it is generally conceded that it may slightly enhance the overall meaning. Some grammarians would even go so far as to say that "its presence or absence in the utterance makes no difference" (*duxūlu-hā fī l-kalām ka-xurūji-hā*). Again, such points of view are severely criticized by rhetoricians, mindful of giving a careful account of the actual semantic value of such utterances.

To conclude this cursory presentation of the place of semantics in Arabic grammar, it must be admitted that the characterization of *naḥw* as being essentially a *ṣināʿa lafḏiyya* is finally rather realistic. Admittedly, no minimally adequate grammar of a natural language can altogether neglect to take into account the semantic dimension, and traditional Arabic grammar is no exception to this principle. It would even be possible to conduct a systematic study of all the points where *naḥw* makes crucial use of semantic considerations. It remains true, however, that the permanent concern of Arab grammarians has been to account for the mechanisms of case assignment and that, in the course of this endeavor, the recourse to meaning has always remained essentially accessory.

Bibliographical references

Primary sources

Ibn al-ʾAnbārī , *ʾInṣāf* = ʾAbū l-Barakāt ʿAbd ar-Raḥmān ibn Muḥammad al-ʾAnbārī, *Kitāb al-ʾinṣāf fī masāʾil al-xilāf bayna n-naḥwiyyīna l-baṣriyyīna wa-l-kūfiyyīn*. Beirut: Dār al-Fikr, n.d.

Ibn Jinnī, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. Ed. Muḥammad ʿAlī an-Najjār. Cairo, 1952–1956.

Ibn ʿUṣfūr, *Mumtiʿ* = ʾAbū l-Ḥasan ʿAlī ibn Muʾmin Ibn ʿUṣfūr al-ʾIšbīlī, *al-Mumtiʿ fī t-taṣrīf*. Ed. Faxr ad-Dīn Qabāwa. Beirut: Dār al-ʾĀfāq al-Jadīda, 1978.

Jurjānī, *Dalāʾil* = ʾAbū Bakr ʿAbd al-Qāhir ibn ʿAbd ar-Raḥmān, *Dalāʾil al-ʾiʿjāz*. Ed. Muḥammad Rašīd Riḍā. Beirut: Dār al-Maʿārif, 1978.
Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. Ed. ʿAbd as-Salām Hārūn. Beirut: ʿĀlam al-Kutub, n.d.
Zajjājī, *ʾĪḍāḥ* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān ibn ʾIsḥāq az-Zajjājī, *al-ʾĪḍāḥ fī ʿilal an-naḥw*. Ed. Māzin al-Mubārak. Beirut: Dār al-Nafāʾis, 1973.
Zamaxšarī, *Mufaṣṣal* = ʾAbū l-Qāsim Maḥmūd ibn ʿUmar az-Zamaxšarī, *Kitāb al-mufaṣṣal fī n-naḥw*. Beirut: Dār al-Jīl, n.d.

SECONDARY SOURCES
Bohas, Georges. 1984. "Contribution à l'étude de la méthode des grammairiens arabes en morphologie et en phonologie d'après des grammairiens arabes 'tardifs'". Georges Bohas and Jean-Patrick Guillaume, *Etude des théories des grammairiens arabes*. I. *Morphologie et phonologie*, 1–333. Damascus: Institut Français d'Etudes Arabes.
Carter, Michael G. 1968. *A study of Sībawayhi's principles of grammatical analysis*. Ph.D. diss., University of Oxford.
Troupeau, Gérard. 1976. *Lexique-index du* Kitāb *de Sībawayhi*. Paris: Klincksieck.

DJAMEL EDDINE KOULOUGHLI
(CNRS)

Laḥn

In Medieval Arabic texts *laḥn* is the key term to refer to linguistic mistakes. According to Fück (1955:205), who studied the term in detail, it is first attested in this sense after the advent of Islam, at the end of the 1st century A.H., but the term itself is older. The common element in all archaic examples is 'leaning over; deviating' (*māla ʾilā* in Classical Arabic dictionaries). Another element in its meaning is the connection between *laḥn* and sound or voice. These two elements explain that in the earliest examples *laḥn* is used for any manner of speaking that deviates from the usual way, for instance a psalmody, a melody (a meaning still current), a way of speaking with outstanding eloquence (in a prophetic tradition), or an allusive way of speaking. The only Qurʾānic attestation of the term has this last meaning (*wa-la-taʿrifannahum fī laḥn al-qawl*, Q. 47/30*)*. According to the *Lisān al-ʿArab*, the expression *rajul laḥin* 'clever man' in a line by the poet Labīd (d. 40/660) is connected with this.

There is yet another element in the sense of the word which is significant for the history of language and culture. In the classical era, linguistic thought clearly distinguishes between *laḥn* and → *luġa*, for instance in a treatise by Ibn Hišām al-Laxmī (d. 577/1181): "On what has been transmitted by the Arabs with two *luġa* or even more. The common language has used the weakest *luġa*, sometimes the better established one, and sometimes it has diverged from the correct form and has used *laḥn*" (*Taqwīm* II, 99). In this passage, *luġa* refers to the dialectal variant of the *kalām al-ʿArab*, particular to a region or tribe. It represents legitimate linguistic variation, prior to the 'corruption of the language' that according to the sources appeared in the 1st century A.H. *Laḥn*, on the other hand, is illegitimate linguistic change, "the diverging [in speech] from the correct form" (Ibn Manẓur, *Lisān* 4013), as a result of 'corruption of the language'. But the *Lisān* (4013) also states that *laḥn*, in one of its meanings, is synonymous with *luġa*: "*Laḥana* can be said of a man who speaks his own *luġa*".

Thus, *laḥn* also means 'the proper way of speaking of a person or an ethnic group'. This meaning seems to be old: al-ʾAṣmaʿī (d. 216/831) says that it was used in a sentence attributed to the caliph ʿUmar: *taʿallamū l-farāʾiḍ wa-s-s-sunan wa-l-laḥn* (Zubaydī [d. 379/989], *Ṭabaqāt* 13; Ibn Manẓur, *Lisān* 4013–4014; Fück 1955:199). When reporting this sentence, Ibn al-ʾAṯīr (d. 606/1210) explains that ʿUmar meant, "Study the language of the Bedouin with their endings" (*Nihāya,* reported by ʾAbū l-Faḍl ʾIbrāhim in Zubaydī, *Ṭabaqāt* 13, n. 9). This meaning of *laḥn* is also found in the words of a Bedouin informant working for the scholar ʾAbū ʿAmr ibn al-ʿAlāʾ (d. 154/770). When the transmitter Xalaf al-ʾAḥmar (d. 180/796) quotes an unacceptable sentence to him, the informant says: *laysa hadā min laḥnī wa-lā min laḥni qawmī*. He then recites a few lines to the effect that he is determined not to speak Persian instead of his own idiom, using *laḥn* to designate both Arabic and Persian: "I'll not abandon my own idiom in order to speak correctly their idiom" (*wa-lā tārikan laḥnī li-ʾuḥsina laḥnahum*; Zubaydī, *Ṭabaqāt* 43–44; Suyūṭī, *Muzhir* I, 291).

All these usages of *laḥn* as a synonym of *luġa* seem to come from an archaic meaning of *laḥn* prior to the setting up of a linguistic norm. The norm is what differentiates between the classical use of *laḥn* and its more archaic use as a synonym of *luġa*. With the implementation of the norm, *laḥn*, which in its pre-classical acceptation meant a detour of speech in a positive sense, came to express a negative 'deviation', a speech error. The norm is also what differentiates *laḥn* from *luġa* in classical use: *luġa* is a way of speaking which does not go against the norm.

There is yet another difference between the classical and the pre-classical use of the term. Originally, *laḥn* seems to have denoted the wrong use of language in speaking, since it is linked to voice and sound. This is confirmed by the various stories (probably spurious), dating from the 1st century A.H., in which *laḥn* is one way to make puns and get laughs because of its play on the homophony of the language, which involves pronunciation (Versteegh 1984:11). Moreover, *laḥn* is associated with *xuṭba* and *qaṣīd* in the first attestation of the word to indicate linguistic errors (Fück 1955:26, 205); in other words, *laḥn* is associated with the oral use of the language, even in a literary context. Soon after, it came to refer to mistakes in the written language, demonstrating the status acquired over nearly a century by the → *ʿarabiyya* as a literary language, essentially linked to writing.

1. TREATISES ON *LAḤN*

Treatises listing linguistic errors are found as early as the 8th century C.E., under the generic term of *kutub laḥn al-ʿāmma* 'treatises on common language mistakes'. The first such treatise, attributed to al-Kisāʾī (d. 183/799), is contemporary with Sībawayhi's *Kitāb* (d. 177/793), thus confirming that *laḥn* and grammatical norm are corollaries. The genre is fairly productive. More than 50 treatises are listed, extending over more than eight centuries. The most commented upon are, in the East, *ʾIṣlāḥ al-manṭiq* by Ibn as-Sikkīt (d. 244/858), *ʾAdab al-kātib* by Ibn Qutayba (d. 276/889), *Kitāb al-faṣīḥ* by Ṯaʿlab (d. 291/904), and *Durrat al-ġawwāṣ* by al-Ḥarīrī (d. 516/1122), and in the West, *Kitāb mā yalḥanu fī-hi ʿawāmm al-ʾAndalus* by az-Zubaydī (d. 379/989), *Taṯqīf* by Ibn Makkī (d. 501/1108), and *Taqwīm* by Ibn Hišām al-Laxmī (see Thorbecke in the introduction to his edition of Ḥarīrī's *Durra*); Rizzitano 1956; Krotkoff 1957; Pellat 1960 for a full bibliography). After the 15th century C.E., these treatises were succeeded by treatises such as *at-Tanbīh ʿalā ġalaṭ al-jāhil wa-n-nabīh* by Ibn Kamāl Bāšā (d. 940/1533) and in modern times by treatises such as *Luġat al-jarāʾid* by ʾIbrāhīm al-Yāzijī (1847–1906) or by dictionaries dealing with common mistakes (*maʿājim al-ʾaxṭāʾ aš-šāʾiʿa).*

The aim of these treatises is puristic: the authors do not intend to understand the errors but rather to denounce them and to recall the *kalām faṣīḥ* (Versteegh 1984:7; Ayoub 2001:103–111). So, although many of the authors are grammarians, none of them use → *qiyās* or → *ʿilla* to understand the rationality of the mistakes. Their educational aim is clear. While assuming Qurʾānic expressions and style, al-Ḥarīrī states: "I wrote this book so that it might shed some light for those who want some clarification and something to resort to for those who wish to have the language ever present in their mind *(tabṣiratan li-man tabaṣṣara wa-taḏkiratan li-man ʾarāda ʾan yaḏḏakkara)*" (*Durra* 2). These works consist in catalogs of entries, without any systematic arrangement. The material is usually presented in identical ways: first the incorrect use is recorded, then the correct form. The word *laḥn* generally refers to the mistake, but words like *wahm, xaṭaʾ, ġalaṭ,* and *saqṭa* are also used. The correct forms are called *ṣawāb, wajh al-qawl, wajh al-kalām,* etc. In some treatises, the correct form is legitimized by a sentence from the *kalām al-ʿArab*: a Qurʾānic verse, a line of poetry. The proof is then a matter of transmission (*naql*). Grammarians' opinions may also be given as evidence. Sometimes, albeit rarely, treatises like the *Taqwīm* had recourse to the poets of the first three centuries, who either were great scholars themselves, like ʾAbū l-ʿAlāʾ al-Maʿarrī (*Taqwīm* II, 291), or who were never corrected by contemporary grammarians, like al-Mutanabbī.

In all texts (*Ṭabaqāt* and *Majālis*) dealing with language mistakes in the first century and in the early stages of grammar, *laḥn* often consists in an incorrect ending or an error in pronunciation. The symbolic value of the syntactic endings is clearly seen in the forceful rejection of such errors. Whatever the linguistic situation at the dawn of Islam and the status of the koine, the system of syntactic endings definitively collapsed with the conquests (Blau 1963, 1977; Fück 1955; Versteegh 1983, 1984; etc.) or had already collapsed (Rabin 1951; Corriente 1971; etc.). The manifest collapse of the inflectional system became the emblem of radical linguistic change in the texts. Such a view is not unreasonable since the collapse of the inflectional system brought about a chain reaction: the word order (Blau 1977), the verbal system (Blau 1977; Versteegh 1984; etc.), the negative system (Ayoub 1996), the complementary system (Blau 1977), etc., all of them linked to the system of syntactic endings. The endings gained a highly symbolic value, becoming the “finery of the language and the embroidery embellishing it” (Ibn Qutayba, *ʿUyūn* II, 172; Zubaydī, *Ṭabaqāt* 11; etc.) and one of the features of the Arabic language that made it superior to other languages (Ibn Fāris, *Ṣāḥibī* 40–41; Suyūṭī, *Muzhir* I, 327–328).

The treatises on *laḥn* are sometimes regarded as a ‘branch of lexicography’ (Pellat 1960: V, 609; Molan 1978). Numerous entries are indeed devoted to semantic shifts, confusions in meanings (*juniba* ‘to be altered by the southern wind’ for *ʾujniba* ‘to be spoiled by the flow of sperm’, Ḥarīrī, *Durra* 122), the use of one derived form for another (*xaṭiʾa* for *ʾaxṭaʾa*, Ḥarīrī, *Durra*, 113; *ʾarsat as-safīna* for *rasat*, Ibn Hišām, *Taqwīm* II, 171), etc. But the mistakes occur at all grammatical levels, whether phonetic or phonological: substitution of a consonant by another in some contexts (*tūṯ* for *tūt, Durra* 66; *barham* for *marham*, *Taqwīm* II, 260), weakening of the hamza (*maṭlūb bi-t-tār* for *maṭlūb bi-ṯ-ṯaʾr*, *Taqwīm* II, 184; *aš-šūm* for *aš-šuʾm*, Ḥarīrī, *Durra* 48), diphthongization, emphasis, palatalizations, etc., and the mistakes are morphological (*muṣān* from *ṣāna* for *maṣūn,* Ḥarīrī, *Durra* 58; *maġrafa* for *miġrafa,* Ibn Hišām, *Taqwīm* II, 174), syntactic (*hab-nī faʿaltu* for *hab ʾannī faʿaltu,* Ḥarīrī, *Durra* 111), morphosyntactic (agreement of the verb or the adjective: *qāmā r-rajulāni,* Ḥarīrī, *Durra* 108; etc.), and syntactico-semantic (value of *rubba* and *rubbamā,* value of *laʿalla*, etc.).

2. TREATISES ON *LAḤN* AND HISTORY OF THE LANGUAGE

The validity of these lexical data for a ‘reconstruction’ of the contemporary dialects is disputed. In fact, their existence is due to both the grammatical norm and the → diglossia. Quite often, the implicit influence of the spoken dialectal form is to be held responsible for the ‘deviation’. Thus, Molan (1978) attempted to reconstruct elements of the dialects in Sicily and Andalusia, starting from four Western treatises on *laḥn*, and Pérez Lázaro, in his edition of Ibn Hišām’s *Taqwīm* (1990:i), deduced some of the features of Hispano-Arabic from the *Taqwīm*’s data. However, this raises difficult epistemological problems. The first one concerns the status of the data and the identity of the speakers. For instance, the expression *ʿāmmat al-ʿArab* is used by Sībawayhi, and the expression *ʿawāmm al-ʾaṭibbāʾ* is found in Ibn Hišām (*Taqwīm* II, 180). Al-Jāḥiz (*Bayān* I, 137) clearly rules out the notion of a ‘mob’ and of allophone elements in *ʿāmma*. In the texts, other social categories, often of literate people, are mentioned: Ḥarīrī (*Durra* 199) mentions *al-xawāṣṣ*, *al-kuttāb*, *al-kubarāʾ*, and *al-ʾaʿ yān*; *an-nuḥāt* are mentioned in ʿUmar ibn Šabba’s book *Kitāb an-naḥw wa-man kāna yalḥanu min an-naḥwiyyīn*, quoted in all the sources; *al-ʿulamāʾ* is used, for instance, by Ibn Jinnī (*Xaṣāʾiṣ* III, 282–309), as are the native Bedouin (*Xaṣāʾiṣ* III, 273–282). Another question concerns the target group of the treatises. In his introduction, al-Ḥarīrī addresses the well-read and high-ranking persons of his time (*man tasannamū ʾasnimat ar-rutab*); Ibn Hišām addresses the *ṭālib al-luġa* ‘the student in lexicography’, who must first correct the altered Arabic words; and Ṯaʿlab is known to have written his book for the novice learner (Suyūṭī, *Muzhir* I, 204). The question remains whether these mistakes stem from previous catalogs or from personal observation. Pérez Lázaro notes that out of 1,906 words of the *Taqwīm*, 1,549 are characterized as *qawl ʿāmmat zamānina, qawl al-ʿawāmm, yaqulūna*, etc. But Ibn Hišām (*Taqwīm* II, 291) borrows most of al-Ḥarīrī’s entries before quoting his book. A more general question is which language these mistakes concern. It seems that the mistakes mentioned are ‘deviations’ from the literary language (*ʿarabiyya*, *fuṣḥā*) in its oral or written use. Actually, many errors concern writing, for instance *taḥrīf* and *taṣḥīf*, i.e. mistakes stemming from inversion of letters (*muḥazraq* instead of *muḥarzaq*; Ibn Jinnī, *Xaṣāʾiṣ* III, 283) and mistakes in reading or writing through badly placed diacritic marks. An example of the latter is the anecdote in as-Suyūṭī (*Muzhir* II, 393) about the *Kitāb at-taṣḥīf* by al-ʿAskarī (d. 382/993). According to him, the name of one poet was read by great scholars, including az-Zajjājī (d. 337/949) and Ibn al-ʾAnbārī (d.

271/885), as Maḥfaṣ, Maxfaḍ or Muḥayṣin, his real name being Muḥaffaḍ. Other data are difficult to explain. The observations in Ibn Hišām's *Taqwīm* (II, 326) about daily-life interjections, e.g. 'they say *ṣab* to the cat when they want it to go away' or interjections of pain (*'aḥ* for *'ax*), hardly concern the written register. It is, therefore, not certain that all data in these works have the same status.

The mistakes reported are generally made by learned people in their use of the *ʿarabiyya* and usually do not give any direct indication about the dialect spoken at the time of the author. Yet, sometimes they do. Thus, Molan (1978:3) notes that the alternation *nktb/nktbw* – a specific feature of Western dialects – is mentioned as a mistake only in the Western treatises, never in Eastern ones. Ibn Hišām (*Taqwīm* II, 195) denounces dialecticisms still current in Eastern dialects involving the interdentals: *ḏibbāna/ḏibbān* instead of *ḏubāba* 'fly', and *min barra* for *min barrin*. Therefore, these treatises can provide indirect indications about the author's dialect if the data are judiciously assessed, inasmuch as they allow us to formulate hypotheses on the dialectal forms that caused them.

Moreover, these treatises enable us to define the changes in the literary language in relation to the norm and indicate the interference by current dialectal forms, without providing a precise chronology of the changes (Versteegh 1997:97): the changes they denounce may be old (e.g. the agreement in number of the subject with the verb). It is necessary to take into consideration → hypercorrections (e.g. when inflectional vowels are assigned to the names of numbers or letters of the alphabet; Ḥarīrī, *Durra* 171–172) and questionable comparisons by the author, e.g. al-Ḥarīrī's (*Durra* 183) comparison of the 'Persianism' *ham* in the dialect of Baṣra with the form *'am* of the article used among the Ḥimyar. Yet, this Persian 'form' must stem from the spoken language and have existed at the time of al-'Axfaš (d. 221/835), 300 years earlier.

Even with these reservations in mind, the treatises are very instructive. They generally emphasize the kind of data produced by interference, as we know them from today's dialectal usage, modern Arabic writing, and the learning of the literary language by Arabic speakers. Thus, al-Ḥarīrī denounces the use of *infaʿala* forms for the apophonic passive form (*inḍāfa š-šay' 'ilay-hi*, *infasada l-'amr ʿalay-hi*; *Durra* 38). In morphology, he denounces the confusion between *faʿala, faʿila,* and *faʿula* (*ḥalā š-šay' fī ṣadrī wa-bi-ʿaynī* for *ḥaliya fī ʿaynī*, but *ḥalā fī famī*; *Durra* 166); the weakening of the causative value of Form I in favor of Form IV (*'uʿlifat ad-dābba* instead of *ʿulifat; Durra* 68); the formation of the *nomina instrumenti*; the formation of elatives and exclamatory words (*mā 'abyaḍa hāḏā ṯ-ṯawba*, *Durra* 30, already present in a Qur'ānic verse; *'ašarr* and *'axyar* for *šarr* and *xayr, Durra* 40; *al-'awwala* for *al-'ūlā,* Ibn Hišām, *Taqwīm* II, 261); errors linked to the dropping of the declensional endings (confusion between *-iyy* of the *nisba* and the *-ī* of the *manqūṣ: mukāriyy* for *mukārin*; **maʿ* for *maʿa*); hypercorrections (*dunyan mutʿiba*; *Durra* 70); the syntax of the numerals; phonological rules specific to the literary language (**al-muġaniyyīn, *al-musammūn* for *al-musammawn*); and last, classic morphological questions involving the diminutives, or the *nisba* (already a concern for grammarians before Sībawayhi; Versteegh 1983:142–148).

At times, a chronology of interference can be established. Thus, the form of the *nomina instrumenti* gave rise to errors as early as the 8th century, probably because it differed from the dialectal form. al-Kisā'ī (*Mā talḥanu* 114) already mentions this error, but unfortunately he only gives the corrected form: *miṯqab, minjal, mibrad, mijmara, mixadda,* etc. His data are taken up again by Ṯaʿlab in his *Faṣīḥ*, by Ibn Qutayba (*'Adab al-kātib* 583), and in many later works. Likewise, Ḥarīrī tells us that by the 11th century *qaṭṭu* had already lost the temporal value given by the norm, as in modern usage (Ayoub 2001:110). al-Ḥarīrī's (*Durra*, 13) contemporaries say *lā 'ukallimu-hu qaṭṭu*, a 'mistake' confirming the change in the verbal system: *qaṭṭu* in complementary distribution with *'abadan* is the norm, being used for the past, while *'abadan* is used for the present and the future. Unfortunately, al-Ḥarīrī does not say whether it is purely modal, as in present usage, nor does he speak about the real value of *'abadan*.

3. TREATISES ON *LAḤN* AND HISTORY OF THE CULTURE

The treatises that discuss *laḥn* are instructive not only linguistically but also from an anthopological and sociological point of view. They help us to define the representations relating to language in the classical era through the connotations of the terms used, the anecdotes scattered in the works, and so on. The analysis of the notions of → *faṣīḥ* and *laḥn* shows that correct speech is the condition for sincere and beautiful speech: aesthetics, ethics, and linguistic accuracy are closely linked (Ayoub 2001). Linguistic errors are moral errors, a wandering from the straight and narrow. On hearing a man misusing the language, the Prophet is said to have exclaimed: "Guide your brother [on the right path], for he has gone astray". In fact, the misuse of language threatens the continuity of the sacred text, which must be safeguarded and guaranteed out of concern for the beyond, but also for the affairs of this world and, first and foremost, for the laws determining social life – the sources of the law being based on the scriptures, first of all on the sacred text. In these works one finds quite remarkable and uncommon qualifying terms for *laḥn: laḥn fāḥiš wa-xaṭaʾ šāʾin* 'a horrible mistake, a shameful mistake' (Ḥarīrī, *Durra* 10); *wahm yašīn wa-xaṭaʾ mustabīn* 'a shameful mistake; an obvious mistake' (*Durra* 125); *mafāḍiḥu l-laḥn aš-šaniʿ* 'among the ugly mistakes and the disgraceful mistakes' (*Durra* 38); etc. These qualifying terms point at the special status of this language: language mistakes belong to the hideous, the horrible, the obscene, the dreadful, the indecent, the immoral, to turpitude.

Furthermore, linguistic errors degrade the beauty of the language. ʾAbū l-ʾAswad ad-Duʾalī (d. 62/681), the legendary founder of grammatical science, is quoted as saying, "I find that a language mistake is like the smell of bad meat" (Sīrāfī, *ʾAxbār* 14). *ʿArabiyya* is the language of *ʾadab*, based on pre-Islamic poetry, an essential linguistic and literary reference. To claim to distinguish oneself by one's *ʾadab* and then to misuse the language is to have a stain on one's honor and reputation. This is what al-Ḥarīrī asserts in his preface:

> I have seen (*fa-ʾinnī raʾaytu*) many people who have risen to the highest ranks in society wanting to distinguish themselves by their *ʾadab* but being nothing less than vulgar in the incorrectness of their speech (*ʾkalām*) and their written work (*ʾaqlām*). If these incorrect manners were to be known and become customary, they could diminish the respect due to distinguished people for their nobility (*al-ʿilya*), and tarnish the splendor of their merit. (*Durra* 2–3)

Thus, what is at stake in the language is social dignity and rank, people's *qadr* and *manzila*. In the passage quoted, some terms relate language to social rank. Countless anecdotes assert this. Numerous others illustrate the following idea: the *ʿarabiyya* is a source of material wealth and power. The treasure-language is worth its weight in gold: the learned man is honored by the greatest princes, the poor man is showered with riches.

The importance attached to linguistic correctness had tremendous consequences on the culture. First, it generated a grammatical tradition of high theoretical content. It also gave birth to a remarkable stylistic analysis, developed by great theoreticians. But it also led to purism. Language was regarded as a treasure that must be jealously protected from the vicissitudes of history, even more so when speakers lost their feel for the language, as early as the advent of the Arabic language in history. From then on, the language was perceived as an unattainable ideal. It enjoyed an extraordinary fortune and was the noble language of an empire, the language of tremendous poets and great minds, but it was never referred to as 'our language' in classical works, but rather as 'their language', the language of the Bedouin of Central Arabia, from whom the grammarians of the 8th century collected their data. Throughout its history, it was 'the language of the absent one' (Ayoub 2001:78–86).

The ill effects of purism were heavily felt in the history of the culture: in the name of the Arabic language, there was a lot of haggling and ostracizing as early as the first centuries. In the episode of *al-masʾala az-zunbūriyya* (Blau 1963; Talmon 1986; Ayoub 2003:52–55), many details are emblematic. The real grammatical argument was between Sībawayhi and Kisāʾī, focusing on data. It dealt with 'what can be said' and 'what cannot be said' in Arabic. The founder of the grammatical tradition, Sībawayhi, was being judged on his ability to speak correctly and not on the depth and the subtlety of his analysis: "You have made a linguistic mistake! This is not the way Arabs speak!". Linguistic mistakes were a matter of syntactic ending. Besides, 'the pure-speaking Arabs' (*fuṣaḥāʾ al-ʿArab*) were symbolically those ‘at the door’, not in the room. The critics judged between the two grammarians, but were absent from the debate.

Moreover, the views on language determined the views on evolution: linguistic change, a necessary consequence of the action of time, could only be a matter of decay and corruption. This view also affected innovation in language and literature. The fascination with the language, the magic of words, brought about tremendous masterpieces. But at the same time, the status of the written work was such that the spoken language was at risk. The views on works of art necessarily entailed a scrupulous respect of the Arabs’ speech, and on account of the status of the literary language, any mistake whatsoever became a mistake forever. These fears are expressed by al-Ḥarīrī in his conclusion of the *Maqāmāt* (IV, 272): “If I had been enlightened by the divine light and if I had looked at myself sympathetically, I would have buried my incompetence into oblivion forever”.

However, every law calls for its infringement, and every culture, even the most literate one, arranges a necessary breathing space. The horror of the mistake in Arab culture is linked with a certain fascination. *Laḥn* fills with horror, but also intrigues. Linguistic errors become fresh and stylish in anecdotes and witty remarks, as confirmed by great writers, from al-Jāḥiz (*Bayān* I, 145–146) to al-Ḥarīrī: “An anecdote is told as it is; a language mistake and coarse words remain unchanged. Didn’t some people say: all the charm of the anecdote lies in its faulty enunciation, and it is delightful when it is told with brio (*mulḥat an-nādira fī laḥni-hā wa-ḥalāwatu-hā fī ḥarāra maqṭaʿi-hā*)?” (Ḥarīrī, *Durra* 96).

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Ḥarīrī, *Durra* = ʾAbū Muḥammad al-Qāsim ibn ʿAlī al-Ḥarīrī, *Durrat al-ġawwāṣ fī ʾawhām al-xawāṣṣ*. Ed. H. Thorbecke. Leipzig, 1871. (Repr., Baghdad.)

——, *Maqāmāt* = ʾAbū Muḥammad al-Qāsim ibn ʿAlī al-Ḥarīrī, *al-Maqāmāt*. Ed. Muḥammad ʿAbd al-Munʿim Xafājī. Beirut: al-Maktaba al-Ṯaqāfiyya, n.d.

Ibn Fāris, *Ṣāḥibī* = ʾAbū l-Ḥusayn ʾAḥmad Ibn Fāris, *aṣ-Ṣāḥibī fī fiqh al-luġa wa-sunan al-ʿArab fī kalāmi-hā*. Ed. Muṣṭafā aš-Šuʾaymī. Beirut: Muʾassasa Badrān li-ṭ-Ṭibāʿa wa-n-Našr, 1964.

Ibn Hišām, *Taqwīm* = ʾAbū ʿAbdallāh Muḥammad ibn ʾAḥmad Ibn Hišām al-Laxmī, *al-Madxal ʾilā taqwīm al-lisān wa-taʿlīm al-bayān*. Ed. José Pérez Lázaro. Madrid: Instituto de Cooperación con el Mundo Árabe, 1990.

Ibn Jinnī, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. Ed. Muḥammad ʿAlī an-Najjār. Cairo, 1952–1956. (Repr., Beirut: Dār al-Hudā, n.d.)

Ibn Kamāl Bāšā, *Tanbīh* = Ibn Kamāl Bāšā, *at-Tanbīh ʿalā ġalaṭ al-jāhil wa-n-nabīh*. Ed. Muḥammad Sawaie. Damascus: Institut Français de Damas, 1994.

Ibn Makkī, *Taṯqīf* = ʿUmar ibn Xalaf Ibn Makkī, *Taṯqīf al-lisān wa-talqīḥ al-jinān*. Ed. ʿAbd al-ʿAzīz Maṭar. Cairo: Lajna ʾIḥyāʾ at-Turāṯ al-ʿIlmī, 1966.

Ibn Manẓūr, *Lisān* = Jamāl ad-Dīn ʾAbū l-Faḍl Muḥammad ibn Mukarram Ibn Manẓūr, *Lisān al-ʿArab*. Cairo: Dār al-Maʿārif, n.d.

Ibn Qutayba, *ʾAdab* = ʾAbū Muḥammad ʿAbdallāh ibn Muslim Ibn Qutayba, *ʾAdab al-kātib*. Ed. Muḥammad Muḥyī d-Dīn ʿAbd al-Ḥamīd. Cairo: Maṭbaʿat as-Saʿāda, 1963.

——, *ʿUyūn* = ʾAbū Muḥammad ʿAbdallāh ibn Muslim Ibn Qutayba, *ʿUyūn al-ʾaxbār*. Ed. Yūsuf ʿAlī aṭ-Ṭawīl. Beirut: Dār al-Kutub al-ʿIlmiyya, 1986.

Ibn as-Sikkīt, *ʾIṣlāḥ* = ʾAbū Yūsuf Yaʿqūb ibn ʾIsḥāq Ibn as-Sikkīt, *ʾIṣlāḥ al-manṭiq*. Ed. ʾAḥmad Muḥammad Šākir and ʿAbd as-Salām Hārūn. Cairo: Dār al-Maʿārif, 1987.

Jāḥiz, *Bayān* = ʾAbū ʿUṯmān ʿAmr ibn Baḥr al-Jāḥiz, *al-Bayān wa-t-tabyīn*. Ed. ʿAbd as-Salām Hārūn. Cairo: Maktabat al-Xānjī, 1985.

Kisāʾī, *Mā talḥanu* = ʾAbū l-Ḥasan ʿAlī ibn Ḥamza al-Kisāʾī, *Kitāb mā talḥanu fī-hi l-ʿāmma*. Ed. Ramaḍān ʿAbd at-Tawwāb. Cairo: Maktabat al-Xānjī and Dār ar-Rifāʿī Riyad, 1982.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān ibn Qanbar Sībawayhi, *al-Kitāb*. Ed. Hartwig Derenbourg. Paris: Imprimerie Nationale, 1881. (Repr., Hildesheim and New York: G. Olms, 1970.)

Sīrāfī, *ʾAxbār* = ʾAbū Saʿīd al-Ḥasan ibn ʿAbdallāh as-Sīrāfī, *ʾAxbār an-naḥwiyyīna l-baṣriyyīn*. Ed. Ṭāhā M. az-Zaynī and Muḥammad ʿAbd al-Munʿim Xafājī. Cairo: Muṣṭafā al-Bābī al-Ḥalabī, 1955.

Suyūṭī, *Muzhir* = Jalāl ad-Dīn ʿAbd ar-Raḥmān ibn ʾAbī Bakr as-Suyūṭī, *al-Muzhir fī ʿulūm al-luġa wa-ʾanwāʿi-hā*. Ed. ʾAbū l-Faḍl Ibrāhim a.o. Cairo: Dār ʾIḥyāʾ al-Kutub al-ʿArabiyya, n.d.

Ṯaʿlab, *Faṣīḥ* = ʾAbū l-ʿAbbās ʾAḥmad ibn Yaḥyā Ṯaʿlab, *Kitāb al-faṣīḥ*. Ed. Jacob Barth. Leipzig: J.C. Hinrich’sche Buchhandlung, 1876.

Yāzijī, *Jarāʾid* = ʾIbrāhīm al-Yāzijī, *Luġat al-jarāʾid*. Ed. Muṣṭafā Tawfīq al-Muʾayyadī. Cairo: al-Maktaba al-Maḥmūdiyya at-Tijāriyya, 1901.

Zubaydī, *Laḥn* = ʾAbū Bakr Muḥammad ibn al-Ḥasan az-Zubaydī, *Kitāb mā yalḥanu fī-hi ʿawāmm al-ʾAndalus*. Ed. ʿAbd al-ʿAzīz Maṭar. Cairo: Dār al-Maʿārif, 1981.

——, *Ṭabaqāt* = ʾAbū Bakr Muḥammad ibn al-Ḥasan az-Zubaydī, *Ṭabaqāt an-naḥwiyyīna wa-l-luġawiyyīn*. Ed. Muḥammad ʾAbū l-Faḍl ʾIbrāhīm. Cairo: Dār al-Maʿārif, 1973.

SECONDARY SOURCES

Ayoub, Georgine. 1996. *Prédicats, figures, catégories: La phrase nominale en arabe littéraire*. Lille: Septentrion.

——. 2001. "Le tout de la langue ou le malheur de l'infini: Une étude de *Durrat al-ġawwāṣ* de Ḥarīrī". *Paroles, signes, mythes: Mélanges J.-E. Bencheikh*, ed. Floréal Sanaugustin, 67–141. Damascus: Institut Français d'Etudes Arabes de Damas.

——. 2003. "Un idiome harmonieux et pur: Le trésor enseveli". *Cent titres à l'usage des bibliothécaires, libraires et amateurs: Poésie de langue arabe*, ed. Jean-Charles Depaule, 29–58. Marseille: Centre International de Poésie.

Blau, Joshua. 1963. "The role of the Bedouins as arbiters in linguistic questions and the *masʾala azzunbūriyya*". *Journal of Semitic Studies* 8.42–51. (Repr. in Blau 1988.)

——. 1977. "The beginnings of the Arabic diglossia: A study of the origins of Neo-Arabic". *Afroasiatic Linguistics* 4.1–28. (Repr. in Blau 1988.)

——. 1988. *Studies in Middle Arabic and its Judaeo-Arabic variety*. Jerusalem: Magnes Press.

Corriente, Federico C. 1971. "On the functional yield of some synthetic devices in Arabic and Semitic morphology". *Jewish Quarterly Review* 62.20–50.

Fück, Johann. 1955. *ʿArabīya: Recherches sur l'histoire de la langue et du style arabe*. Transl. Claude Denizeau. Paris: Librairie Marcel Didier.

Krotkoff, George. 1957. "The *Kitāb laḥn al-ʿawāmm* by Abū Bakr az-Zubaydī: Description of a lexicographical manuscript". *Majalla Kulliyyat al-ʾĀdāb wa-l-ʿUlūm* 12.183–196.

Maṭar, ʿAbd al-ʿAzīz. 1966. *Laḥn al-ʿāmma fī ḍawʾ ad-dirāsāt al-luġawiyya al-ḥadīṯa*. Cairo: Dār al-Kātib al-ʿArabī.

Molan, Peter D. 1978. *Medieval Western Arabic: Reconstructing elements of the dialects of al-Andalus, Sicily and North Africa from the* laḥn al-ʿāmma *literature*. Ph.D. diss., University of California.

Pellat, Charles. 1960. "*Laḥn al-ʿāmma*". *Encyclopaedia of Islam* V, 609–614. 2nd ed. Leiden: E.J. Brill.

Rabin, Chaim. 1951. *Ancient West-Arabian*. London: Taylor's Foreign Press.

Rizzitano, Umberto. 1956. "*Il Taṯqīf al-lisān wa-talqīḥ al-janān* di Abū Ḥafṣ ʿUmar b. Makkī". *Studia Orientalia* 5.193–213.

Talmon, Rafael. 1986. "Al-masʾala az-zunbūriyya: Dirāsa fī māhiyyat ixtilāf al-maḏhabayn an-naḥwiyyayn". *al-Karmil* 7.131–163.

Versteegh, Kees. 1983. "Arabic grammar and corruption of speech". *al-ʾAbḥāṯ* 31.117–138.

——. 1984. *Pidginization and creolization: The case of Arabic*. Amsterdam and Philadelphia: J. Benjamins.

——. 1997. *The Arabic language*. Edinburgh: Edinburgh University Press.

GEORGINE AYOUB (Institut des Langues et Civilisations Orientales)

Language Academies

1. ACADEMY PRECURSORS

The first 'academy' in the Arab world was established by the French during the Napoleonic occupation of Egypt (1798–1801); it came to an end with the French exit from Egypt in 1801. There were several attempts in the 19th century, especially in Egypt and Lebanon, by enlightened scholars concerned with intellectual issues in general, and language matters in particular, to establish similar organizations. A serious effort was made in 1892/1893 when the first meeting of *al-Majmaʿ al-Luġawī al-ʿArabī* 'Arabic Language Academy' was held at the home of Muḥammad Tawfīq al-Bakrī (1870–1933). Its goal was to coin terminology for newly introduced Western cultural items. In a subsequent meeting, more new coinages were suggested (Fahmī 1967:64–65). Despite the membership of leading scholars and writers, however, the *Majmaʿ* did not persist (ʿUbāda 1928:307; Madkūr 1981:22).

In 1907/1908 the *Nādī Dār al-ʿUlūm* 'Dār al-ʿUlūm Club' was founded by Ḥafnī Nāṣif (1855?-1919/1920?) in Cairo to coin new terminology and to address the dilemma of foreign words introduced into Arabic. During its brief life, the club proposed thousands of coinages in its journal (ʿUbāda 1928:307; Madkūr 1964:15). In 1916/1917, ʾAḥmad Luṭfī as-Sayyid (1872–1963), then director of the Egyptian National Library, established *al-Majmaʿ al-Luġawī al-ʾAwwal* 'First Arabic Language Academy', modeling it after the *Académie Française*. Its purpose was to compile a comprehensive dictionary of terminology in sciences, arts, and crafts. Like its predecessors, it disbanded without leaving a trace after the 1919 revolution (ʿUbāda 1928:308; Madkūr 1964:15). In 1921 some founding members convened to resurrect the *Majmaʿ*. Thus renewed, it held 40 sessions, the last one in December 1925. Eventually, it too became moribund, because of political divisions among members, lack of financial and moral support from the government, and linguistic polarization for philosophical reasons. The bone of contention between the two groups evolved around support of derivation vs. advocacy of Arabicization (ʿUbāda 1928:308–309). Certain members of the group, however, continued their untiring quest to establish an academy. Their efforts finally bore fruit in 1932 (ʿUbāda 1928:308–309).

Discussion follows of the language academies in Damascus, Cairo, Baghdad, and Amman, according to their chronological establishment, and of the Union of Arabic Language Academies, all established in the 20th century. *Al-Majmaʿ at-Tūnisī li-l-ʿUlūm wa-l-ʾĀdāb wa-l-Funūn – Bayt al-Ḥikma* in Tunis and *Markaz at-Taʿrīb* in Rabat will not be treated in this entry. Reports about recently established academies in Libya, Sudan, and Palestine (al-Faḥḥām 1996:28; ar-Rāwī 2002:78) have not been confirmed. Due to the paucity of information as to their existence or productivity, they are not discussed in this entry.

2. The Arabic Language Academy in Damascus

Following the dissolution of the Ottoman Empire after World War I, the newly established state of Syria faced new realities with respect to language. During the Ottoman rule (1516–1918), Turkish had been the official language. The Arab authorities that took control of the country in 1918 were bound to replace Turkish with Arabic; thus, for example, they established a special department for translation and book preparation (*Šuʿbat at-Tarjama wa-t-Taʾlīf*) in Damascus. This section later became the *Dāʾirat al-Maʿārif* 'Education Department' in 1919. Subsequently, it was replaced by *al-Majmaʿ al-ʿIlmī al-ʿArabī* 'Arabic Language Academy' on 8 June 1919. The newly established *Majmaʿ* held its first meeting on 30 June 1919. In 1920, it became a part of the Education Department, later coming under the aegis of the Syrian University, founded in 1923. In the early part of 1927, the *Majmaʿ* severed its relationship with the university and was placed again under the Ministry of [Public] Education (Kurd ʿAlī 1926:555–557). It finally acquired independent status in 1927.

The Damascus *Majmaʿ* was the first to be established in any Arab country and was in its turn modeled after the *Académie Française* (Kurd ʿAlī 1922:3). The *Majmaʿ al-ʿIlmī al-ʿArabī* merged with its Egyptian counterpart in 1960 following the political unification of Egypt and Syria in 1958, and acquired the name of *Majmaʿ al-Luġa al-ʿArabiyya bi-Dimašq*. Despite the collapse of that political unity in 1961, the Damascus *Majmaʿ* continues to carry this name (al-Faḥḥām 1996:27). When first established in June 1919, the *Majmaʿ* had only eight members in addition to its president, Muḥammad Kurd ʿAlī (1876–1953). Initially, it comprised two branches, one concerned with linguistic and literary matters and the other with scientific and technical issues. However, due to administrative reasons and financial difficulties, the work of the *Majmaʿ* was suspended in November 1919; only two members were retained to supervise the public library and the museum. Then, on 14 September 1920, the *Majmaʿ* was reconstituted and Kurd ʿAlī was reinstated as president.

In its meeting on 12 October 1920, the *Majmaʿ* elected Arab and European members, thereby swelling the ranks of its membership (al-Futayyiḥ 1956:233–237). The May 1947 legislative decree established two membership categories, active members (*ʿāmilūn*) and correspondent members (*murāsilūn*).

New statutes were issued in May 2001 to regulate the work of the *Majmaʿ*, effectively abrogating the previous statutes and redefining the membership categories. The active-member category was replaced by the category of *ʿuḍw Majmaʿ* 'Academy member'. This category's 25 members are required to be Syrian nationals and residents of Damascus. The correspondent-member category was maintained in the new regulations without restriction as to number of members, which may include Syrians, other Arabs, and non-Arabs. The 2001 law introduced the new category *ʿuḍw šaraf* 'honorary member' (*Majalla Majmaʿ al-Luġa al-ʿArabiyya bi-Dimašq* 78:1.109–136).

The statement of the *Majmaʿ*'s goals underwent several amendments. In 1919, the goals were determined by the needs of the nascent government and included dissemination of literary Arabic works to the public; resuscitation of Classical Arabic manuscripts and establishment of a national public library; coinage of technical and scientific idioms; and publication of a journal (al-Futayyiḥ 1956:11–14).

Kurd ʿAlī (1922:3–4) modified the goals of the *Majmaʿ* three years later to include reform of the Arabic language; coinage of new terminology for modern technical inventions; and support for research and encouragement of book publishing and translation, especially in the sciences.

The 2001 regulations redefined some of the *Majmaʿ*'s earlier objectives and added others. They stipulated, for example, that scientific, technical, and literary coinages follow set paradigms. Efforts should be made to standardize coinages and disseminate them in all Arab countries. Additionally, the regulations emphasized that the methods of teaching Arabic grammar and morphology must be simplified, along with standardization of orthography and spelling. Emphasis was, furthermore, placed on finding ways to limit the spread of dialects in all spheres of language use.

The *Majmaʿ* rendered valuable services to the new Syrian government and its administrative apparatus in three ways: by training employees in Arabic, by translating the Ottoman administrative terminology into Arabic, and by coining new terminology in *fuṣḥā* 'Standard Arabic' (Kurd ʿAlī 1921:43–46).

The transition into the new idiom took many forms. Ottoman Turkish administrative terminology was abolished and Arabic counterparts were adopted instead. Sometimes, Ottoman administrative words were slightly modified to bring out the vestiges of their Arabic elements, and certain terms were redefined to denote specialized meanings. To encourage use of the new technical terminology, the *Majmaʿ* urged the heads of governmental departments to enforce such usage in their transactions.

The *Majmaʿ* was concerned with serving the public at large. It initiated public lectures in 1921, some of which were designed for all-women audiences. Developed by *Majmaʿ* members and guest lecturers including women, the lectures covered topics in literature, language, history, geography, health, medicine, sociology, ethics, etc. Insofar as those lectures deviated from the original *Majmaʿ* goals, long and heated debates flared about its mission. Despite the rancor, however, the *Majmaʿ* pursued its commitment to lecturing as well as to publishing both in its journal and in separate monographs. The first monograph appeared in 1925 (al-Futayyiḥ 1956:75–76).

The *Majmaʿ* combated widespread linguistic errors common in the writing of journalists and others. Lists of corrections of errors were published in the *Majmaʿ* journal under the heading *ʿAṯarāt al-ʾaqlām* 'Writers' errors'. Errors in speech were also combated by the *Majmaʿ*, appropriately acquiring the title of *ʿaṯarāt al-ʾafmām* 'slips of the tongue.'

The journal *Majallat al-Majmaʿ al-ʿIlmī al-ʿArabī* ('*Majalla*') was first established in January 1921. Initially published monthly, in 1931 it became a bimonthly. Beginning with Volume 24 it became a quarterly, continuing in this format to the present (an-Naṣṣ 1966:105–116). Due to financial difficulties and administrative reasons, publication of the *Majalla* was interrupted twice, during 1932–1935 and 1938–1941. It resumed publication in 1941. The *Majalla* changed its name to *Majalla Majmaʿ al-Luġa al-ʿArabiyya bi-Dimašq* beginning with Volume 36, following the political union between Syria and Egypt in 1958 and subsequent to the merger of the Cairo and Damascus Academies in 1960. However, after the breakup of this political union in 1961, effective with Volumes 37, 38, 39, and 40, the *Majalla* reverted to its original name *Majallat al-Majmaʿ al-ʿIlmī al-ʿArabī*. From Volume 41 on, the name was once again *Majalla Majmaʿ al-Luġa al-ʿArabiyya bi-Dimašq* (Xaṭīb 1969:19; an-Naṣṣ 1996:114; Hamzaoui 1988:111–112).

The publishing and rehabilitation of Arabic manuscripts had been a primary *Majmaʿ* goal since the organization's foundation. Despite a limited budget, it published small pamphlets, extracts from larger classical works, and some of its public lectures. As of 1944 it focused on publishing works in literature, history, and language, especially Classical Arabic manuscripts from the Ẓāhiriyya Library collection.

The *Majmaʿ* publications, presented as gifts to universities, were also made available to the public at reasonable cost. Because early publications have gone out of print, the current *Majmaʿ* has arranged with commercial publishers to reproduce such items – possibly a financially unfeasible arrangement due to limited readership. The volume of publications has shrunk in recent years due to the modest royalties extended to authors or editors, compared to larger royalties and wider publicity offered by commercial publishers.

Demands were placed on the *Majmaʿ* to publish a dictionary comparable to that of the *Académie Française*. Kurd ʿAlī was resistant to these demands for two reasons: the unlikelihood of the young *Majmaʿ*'s being able to accomplish results comparable to those of centuries-old Western academies; and the fact that the necessary terminology for sciences and material objects was unavailable. A dictionary without such coinages, Kurd ʿAlī (1924:65) maintained, would amount to a mere duplication of Classical Arabic dictionaries. Nonetheless, the idea of a dictionary is still alive insofar as the 2001 statutes stipulate the compiling of modern dictionaries for the various scientific fields.

The *Majmaʿ* organized public commemoratives to honor leading classical poets, writers, and former members of the *Majmaʿ*. It also held ceremonies in honor of major figures in modern Arabic literature, including members of the *Majmaʿ*. In 1920 the *Majmaʿ* organized the first commemorative anniversary for its member Ṭāhir al-Jazāʾirī (1851–1920), credited with the founding of the public Ẓāhiriyya Library in Damascus. In August 1924 another commemorative was held to honor the Egyptian essayist and writer Muṣṭafā Luṭfī al-Manfalūṭī (1876–1924) and the Iraqi scholar Maḥmūd Šukrī al-ʾAlūsī (1857–1924) (al-Futayyiḥ 1956:85–86).

Commemorative ceremonies on a larger scale were held to honor poets and writers from the classical period. A week-long millennary memorial was held in 1936 to honor the 10th-century poet al-Mutanabbī. A similar ceremony was held in 1944 in honor of the poet, writer, and philosopher ʾAbū l-ʿAlāʾ al-Maʿarrī (d. 1058) on the millennary occasion of his birth (al-Futayyiḥ 1956:95–99). The *Majmaʿ* assumed the authority to defend Arabic against 'foreignisms' resulting from intensified contacts with Western languages. Three guidelines were established: existing Arabic words should be used to convey foreign notions; in the absence of existing words, new terminology that was coined from Arabic radicals must adhere to Arabic morphological and phonological patterns; and if appropriate Arabic words could not be devised, foreign words could be adopted provided they approximated those of Arabic, morphologically and phonologically.

One of the early attempts to coin technical and scientific terminology was made by a *Majmaʿ* member in 1935. ʿIzz ad-Dīn at-Tanūxī (1889–1966) published *Tašrīḥ ad-darrāja* 'Anatomy of the bicycle', an article in which he provided Arabic terms for the various parts of the bicycle, newly introduced to the region. At-Tanūxī based his terminology on the French equivalents by providing literal translations (at-Tanūxī 1935:363–368). He suggested, for example, *ad-dūlāb al-muwajjih* for the front wheel in imitation of the French *la roue directrice.* By the time he suggested this term, though, the popular term *ad-dūlāb al-ʾamāmī* 'the front wheel' had already gained currency. A complicating factor in the acceptance and currency of such coinages was the gap between the *Majmaʿ* and its members on the one hand, and society at large on the other. Another complicating factor was the multiplicity of suggested terms for the same referent.

Such technical and scientific lexical contributions were accomplished by Syrian scholars, scientists, and medical professionals. The *Majmaʿ* maintained, however, that other comparable Arab institutions must be involved for such terms to be adopted in educational and governmental institutions throughout the Arab countries. The *Majmaʿ* solicited views from specialists in other Arab countries with a view toward producing standard Arabic terminology and perhaps, eventually, a comprehensive modern dictionary.

3. Arabic Language Academy, Cairo

The 1932 decree of the Egyptian government to establish the *Majmaʿ al-Luġa al-ʿArabiyya al-Malakī* 'Royal Language Academy' was a result of demands for such an organization (Madkūr 1964:15). In 1934, the *Majmaʿ* started its activities, modeled after the *Académie Française* with respect to objectives and organization (Hamzaoui 1988:48–50, 107). In 1938 the name changed to *Majmaʿ Fūʾād al-ʾAwwal li-l-Luġa al-ʿArabiyya* '[King] Fouad I Academy for the Arabic Language'. In 1954 the name changed again to *Majmaʿ al-Luġa al-ʿArabiyya,* then, in 1960, the name became *Majmaʿ al-Luġa al-ʿArabiyya fī l-Qāhira* as a result of the 1958 political union between Egypt and Syria.

The Cairo Academy set the following objectives: preservation of the Arabic language and its development in ways expressive of modern sciences, arts, and society; inquiry into all means that could enhance the Arabic language; editing of Classical Arabic texts and manuscripts; compilation of a historical dictionary; and publication of a journal.

The journal was issued on an annual basis from 1934 until 1937. It was later suspended for eleven years for several reasons: the budget was too modest, the administrative and bureaucratic procedures were slow, and it was difficult to print the journal (Madkūr 1964:36). The fifth volume appeared in 1948; between 1948 and 1956 only four volumes were printed. Effective 1957, the journal resumed publication on an annual basis.

Membership is open to scholars recognized for their scholarship in the Arabic language without restriction to nationality or political or sectarian affiliation, thus giving the *Majmaʿ* the potential to be an international organization. The members are Egyptian, Arab, and non-Arab, membership being based on their qualifications and contributions to Arabic studies. They fall into three categories: active (*ʿāmilūn*), honorary (*faxriyyūn*), and correspondent (*murāsilūn*).

The Cairo Academy contributed the following advances: simplification of Arabic grammar and writing style; provision of scientific and cultural terminology through the mechanisms of → *ištiqāq* 'derivation', → *majāz* 'metaphors', *naql* 'translation', *naḥt* 'calque', and *taʿrīb* 'Arabicization'; refinement of Arabic language dictionaries; rehabilitation of Classical Arabic works; and compilation of a comprehensive historical dictionary (Madkūr 1964:34, 51ff.; Madkūr 1981:27–29; El-Khafaifi 1985:68–178; Sawaie 1983:58–61). In 1936, the Academy recommended support for the preparation of August Fischer's *al-Muʿjam al-luġawī at-tārīxī* 'Etymological historical Arabic dictionary' (Fischer 1967:31). It should be added that this project has never been realized.

The *Majmaʿ* published two dictionaries. *al-Muʿjam al-wasīṭ* appeared in 1960 in two volumes. In this dictionary, words are arranged alphabetically rather than morphologically. Verbs are introduced before nouns, triconsonantal verbs before augmented ones, and intransitive before transitive. With respect to nouns, concrete meanings are organized before abstract, real meanings before metaphorical. It also introduces items generated by analogy apart from that transmitted (Madkūr 1964:67). The second dictionary is *al-Muʿjam al-kabīr*, an endeavor begun in 1946, with the first volume, comprising part of the letter *hamza,* published in 1956 on an experimental basis. Five volumes have been published to date. Its organization is similar to *al-Muʿjam al-wasīṭ* in that it introduces Arabicized items alphabetically, with indication of their foreign sources. It also mentions the Semitic cognates where possible (Madkūr 1964:68–70; ʿUmar 2002:309–312).

The Cairo Academy faced various obstacles. In 1942 there was governmental interference in appointing members. Kurd ʿAlī mentions that the selection of members in the early stages of the Academy in Cairo was based on merit. However, membership appointment was at times compromised as a result of political considerations, and its workings were affected by partisanship (Kurd ʿAlī 1948:495–498). The situation may not have been dissimilar to that of the Damascus *Majmaʿ* (Kurd ʿAlī 1948:791–792; Hamzaoui 1975, 1988:58–71). Moreover, political instability and wars caused the suspension of *Majmaʿ* sessions and prevented members from attending meetings (Madkūr 1964:20; Kurd ʿAlī 1948:277–286).

Kurd ʿAlī provides a comparison between the Cairo and Damascus Academies, lauding Damascus Academy members for their successes while harshly criticizing members of the Cairo Academy for their ineffectiveness. Some members, he adds, never contributed or expressed an opinion to its journal (Kurd ʿAlī 1948:529–530).

4. The Iraqi Academy

The first attempt to establish an academy in Iraq was in 1921, through the establishment of *al-Maʿhad al-ʿIlmī fī Baġdād* 'Scientific Institute in Baghdad', whose main goals were the revitalization of the heritage of Islamic and Arab past, the abolishment of illiteracy, and the support of education (al-Jubūrī 1965:31–33). In 1934, *Nādī l-Qalam al-ʿIrāqī* 'Iraqi Pen Club' was founded to cement relations among authors and writers and to support research and Arabic literature in general. In 1945 *Lajnat at-Taʾlīf wa-n-Našr* 'Committee on Writing and Publishing' was established to support authors, translators, and publishers, and continued until 1947, when the Ministry of Public Education established *al-Majmaʿ al-ʿIlmī al-ʿIrāqī* 'the Iraqi Academy.'

In 1963 a presidential decree abolished the Iraqi Academy, along with its 1947 regulating guidelines and the 1949 amendments, and instituted new regulations (al-Jubūrī 1965:47). Thus reformed, the academy set up its own printing press and began an annual public lecture series by its active and honorary members as well as by other qualified scholars. These lectures were published in the Academy journal and also in a separate monograph in an effort to reach wider audiences (al-Jubūrī 1965:70–74). In addition to *al-Majmaᶜ al-ᶜIlmī al-ᶜIrāqī*, the Kurdish Language Academy and the Syriac Language Academy were established in 1968. All three academies were merged into one in 1978. This new unified academy was named *al-Majmaᶜ al-ᶜIlmī al-ᶜIrāqī fī Baġdād* (Maṭlūb 1983:161).

Al-Majmaᶜ al-ᶜIlmī al-ᶜIrāqī was modeled after the Damascus *Majmaᶜ* in its focus on language, sciences, literature, history, etc. (*Majallat al-Majmaᶜ al-ᶜIrāqī* 1950, 383–384). The Cairo Academy's primary interest, by contrast, was in language matters: grammar, morphology and coinages, etc. To unify the coinages in all Arab countries, the Iraqi Academy traditionally consulted scholars of Arabic and the proceedings of the Cairo Academy and its journal, in addition to the Damascus *Majmaᶜ* journal (Jawād 1955:121).

In 1995, a new law was passed regulating the Academy and naming it *al-Majmaᶜ al-ᶜIlmī*. The new academy was placed directly under the Presidential Office instead of the Ministry of Public Education or the Ministry of Higher Education (established 1968). The 1995 regulations reorganized the Academy into the following specialized sections, each charged with a special task: the Arabic language; Arabo-Islamic heritage; humanities; theoretical and applied sciences; terminology; and translation (al-ʾAlūsī 1997:62–63). In addition, two committees were established: a committee of the Kurdish language charged with all matters pertaining to the Kurdish language, culture, and heritage; and a committee for Syriac that was charged with the Syriac language, culture, and heritage. No updated information is available at this moment regarding the activities of the Kurdish and Syriac sections.

The goals of the Iraqi Academy were modified over the years. The 1947 goals were to enable the Arabic language to express the needs of the modern age in the sciences and arts; to publish, in Arabic, literature, history (especially Iraqi), languages, sciences, and civilization; to preserve rare manuscripts and archival materials; and to encourage translations of modern sciences and arts and promote scientific inquiry. In 1963, the goals were reformulated to include revitalization of scientific inquiry in Iraq to cope with the advancement of science; protection of the Arabic language from deterioration and promotion of ways to enrich it to express modern sciences and the arts; encouragement of original research, translation, and publishing in the sciences and the arts; and revitalization of Arabo-Islamic heritage in the sciences and arts. In 1995, the regulations set the following objectives: emphasis on and care of the preservation of Arabic; inquiry into modern sciences and technology; encouragement of translation and scientific inquiry; revitalization of the Arabo-Islamic heritage; and support for and revitalization of humanities, sciences, and the arts (al-ʾAlūsī 1997:61).

The 1947 regulations created four membership categories: active (*ᶜāmilūn*), assistant (*musāᶜidūn*), honorary (*faxriyyūn*), and correspondent (*murāsilūn*) (*Majallat al-Majmaᶜ* 1950, 4). New categories were established in 1949 as follows: active (*ᶜāmilūn*), supporting (*muʾāzirūn*), honorary (*faxriyyūn*), and correspondent (*murāsilūn*). In 1949, the 'assistant' category was replaced by the 'supporting' category. In the 1995 guidelines, members were divided into three categories: active (*ᶜāmilūn*), comprised of residents in Baghdad who formed the membership of the Academy Council; correspondent (*murāsilūn*), referred to at times as supporting, who were selected from among Iraqi scholars as well as various Arab, Islamic, and other foreign countries; and honorary (*faxriyyūn*), Iraqi nationals who were residing outside Iraq for extended periods of time. In addition, the Academy had the authority to select no more than ten non-Iraqi Arab scholars as active members to participate in Academy activities while residing in Iraq.

In 1950, the Academy established *Majallat al-Majmaᶜ* 'Journal of the Iraqi Academy'. Until its abolishment in 1963, the journal published ten volumes in twelve parts (al-Jubūrī 1965:75–83). A considerable amount of writing in these journals was devoted to topics pertaining to the Arabic language and coinages for various new inventions introduced in Iraq: the oil industry, the railway, air and space, public transportation, etc. In addition, the Academy published dictionaries, lists of coinages, and books on language, literature, and history and subsidized the publishing of other books (al-ʾAlūsī 1997:515).

Like the other academies, the Iraqi Academy faced several obstacles. Interference of politicians in the *Majmaʿ*'s affairs in various ways created difficulties, e.g. by decreasing the budget and by appointing unqualified members. The bureaucracy of the various governmental offices affected the operation of the Academy in its connection with those offices, and unilateral decisions by governmental bureaucrats were made on Academy matters (al-ʾAlūsī 1997:283; ʿAlī 1959:318–354). Due to the 2003 war on Iraq, no reliable information is available at this moment about the fate of the Iraqi *Majmaʿ*.

5. The Jordan Arabic Language Academy

The idea of establishing an Arabic language academy in Jordan dates back to the early years of the Arab Kingdom in 1921. At the end of April 1922, the government of Transjordan wished to establish an academy in Amman comparable to the one in Damascus. Prince Abdullah issued a decree in July 1923, naming the new academy's active members, whom he charged with electing honorary members from outside Jordan. The members held meetings, but in the end the academy idea was abandoned due to financial difficulties and lack of qualified members (Xaṭīb 1969:66–67; Kurd ʿAlī 1948:518–519).

However, the idea of an Arabic language academy never died. In 1961, the Jordan Committee on Arabization, Translation, and Publishing was founded. In 1976, authorization to establish *Majmaʿ al-Luġa al-ʿArabiyya al-ʾUrdunnī* 'Jordan Arabic Language Academy' was issued, and the Academy, consisting of five members, was inaugurated in 1977; it joined the Union of Arabic Language Academies that same year.

Majmaʿ al-Luġa al-ʿArabiyya al-ʾUrdunnī is headed by a president and vice president. There are three categories of members: active members from Jordan, honorary members from Jordan and elsewhere, and supporting members from Arab and foreign countries, as the Academy stated in its first annual decision in 1977.

The Academy set the following goals: preserving the Arabic language and enabling it to express modern sciences, arts, and technology; standardizing terminology of sciences and arts; and compiling a dictionary expressive of the needs of modern times. The Academy proposed the following means to accomplish its goals: inquiry into and research on Arabic; support of translation and publishing; publication of new coinages; publication of a journal (the first issue was published in 1978); and convening of conferences on language issues.

To that end, several practical steps were taken. Investigations were made into weakness in Arabic among speakers; foreign scientific terminology was Arabicized; university instruction was Arabicized; and, finally, steps were take to revive the Arabo-Islamic heritage. The following are among the Academy's accomplishments to date: translation of scientific books as part of all-Arabic instruction at the university level, including books in mathematics, physics, chemistry, biology, and geology; editing of Classical Arabic books and books on Arabic; cultural conferences; cataloguing of manuscripts; and coinages for metrology, commerce, economy, agriculture, and the military.

Despite its successes, the Jordan Academy has faced several problems, among them lack of financial support and lack of coordination among language academies, particularly with respect to standardization of new coinages.

6. Union of the Arabic Language Academies

The call for establishment of a union of Arabic language academies first appeared in October 1956, eventually bearing fruit in 1971. A constitution for the Union, which then included only the Academies of Damascus, Cairo, and Baghdad, and its council were established (*Majalla Majmaʿ al-Luġa al-ʿArabiyya bi-Dimašq* 1971, 593–598). The Jordan Arabic Language Academy joined the Union in 1977 as its fourth member. The Union has an executive council consisting of two members from each Academy, headed by a president who is elected from among the members, in addition to a secretary-general and two assistant secretaries-general.

The goals of the Union are to facilitate contacts among all the Arabic language academies, to coordinate their activities, and to standardize Arabic scientific, technical, and cultural terminology and its dissemination.

Several mechanisms were suggested for

coining terminology: the revivification of terminology from the Arab heritage, and the coining of new terminology by means of metaphors (*majāz*), derivation (*ištiqāq*), calque (*naḥt*), and Arabicization (*taʿrīb*). Emphasis was placed on the use of *fuṣḥā* 'Standard Arabic' terminology instead of Arabicized terms. Preference is to be accorded to terminology that yields itself to derivation.

The Union organized the following conferences to posit technical and scientific terminology on various topics: Legal Terminology (Damascus, 1972), Oil Terminology (Baghdad, 1973), Simplifying the Teaching of Arabic (Algeria, 1976), the Teaching of Arabic in the Last Quarter of the Century (Amman, 1978), Arabicization of Higher and University Education (Rabat, 1984), Unification of Scientific Symbols (Amman, 1987), Standardization of Arabic Medical Terminology (Tunis, 1992), Dictionary of Oil Terminology (Damascus, 1994), and Dictionary of Geology (Tunis, 1994) (Ḍayf 1996:92).

The Union of the Arabic Language Academies has faced obstacles, some quite severe at times. These include political disagreements among various Arab states that affect Union activities; lack of financial and moral support; and the failure of some Academies to pay their annual membership fees. In addition, there was a schedule conflict with the meetings of the Arabic Language Academy in Cairo, causing the meetings to lack seriousness; in general, some representatives failed to attend the meetings. There has also been a lack of support for Union decisions, both in the media and in institutions of higher learning. As a result, coinages have not filtered down easily to institutions of higher learning, academics, or research institutes. Finally, the Union has faced competition from other organizations engaged in coinages, such as the Arab League's Educational, Cultural, and Scientific Organization (ALECSO) (ar-Rāwī 2002:5–16).

BIBLIOGRAPHICAL REFERENCES

ʿAlī, Jawād. 1956. "Xulāṣa li-ʾaʿmāl al-Majmaʿ al-ʿIlmī al-ʿIrāqī". *Majallat al-Majmaʿ al-ʿIlmī al-ʿIrāqī* 4:2.738–739.

——. 1959. "al-Majāmiʿ al-ʿIlmiyya". *Majallat al-Majmaʿ al-ʿIlmī al-ʿIrāqī* 6.318–354.

ʾAlūsī, Sālim al-. 1997. *al-Majmaʿ al-ʿIlmī fī xamsīna ʿāman, 1947–1997*. Baghdad: Maṭbaʿat al-Majmaʿ al-ʿIlmī.

Ḍayf, Šawqī. 1996. "Ittiḥād al-majāmiʿ al-luġawiyya". *Majalla Majmaʿ al-Luġa al-ʿArabiyya bi-Dimašq* 71:1.90–97.

El-Khafaifi, Hussein M. 1985. *The role of the Cairo Academy in coining Arabic scientific terminology: An historical and linguistic evaluation.* Ph.D. diss., University of Utah.

Faḥḥām, Šākir al-. 1996. "*Majmaʿ* presidential speech". *Majalla Majmaʿ al-Luġa al-ʿArabiyya bi-Dimašq* 71:1.22–32.

Fahmī, Māhir Ḥasan. 1967. *Muḥammad Tawfīq al-Bakrī*. Cairo: Dār al-Kātib al-ʿArabī.

Fischer, A[ugust]. 1967. *al-Muʿjam al-luġawi at-tārixī*, I. Cairo: al-Hayʾa al-ʿĀmma li-Šuʾūn al-Maṭābiʿ al-ʾAmīriyya.

Futayyiḥ, ʾAḥmad al-. 1956. *Tārīx al-Majmaʿ al-ʿIlmī al-ʿArabī*. Damascus: Manšūrāt al-Majmaʿ al-ʿIlmī al-ʿArabī.

Hamzaoui, M. Rached. 1965. *L'Académie Arabe de Damas et le problème de la modernisation de la langue arabe*. Leiden: E.J. Brill.

——. 1975. *L'Académie de Langue Arabe du Caire: Histoire et oeuvre*. Tunis: Université de Tunis.

——. 1988. *ʾAʿmāl Majmaʿ al-Luġa al-ʿArabiyya bi-l-Qāhira*. Beirut: Dār al-Ġarb al-ʾIslāmī.

Jawād, Muṣṭafā. 1955. *al-Mabāḥiṯ al-luġawiyya fī l-ʿIrāq*. Cairo: Maʿhad ad-Dirāsāt al-ʿArabiyya al-ʿĀliya.

Jubūrī, ʿAbdallāh al-. 1965. *al-Majmaʿ al-ʿIlmī al-ʿIrāqī*. Baghdad: al-ʿĀnī Press.

Kurd ʿAlī, Muḥammad. 1921. "ʾIṣlāḥ luġat ad-dawāwīn". *Majallat al-Majmaʿ al-ʿIlmī al-ʿArabī* 1:2.43–46.

——. 1922. *ʾAʿmāl al-Majmaʿ al-ʿIlmī al-ʿArabī fī Dimašq, 1922, 1923 wa-1924*. Damascus: Orthodox Patriarchate Press.

——. 1926. "al-Jalsatāni l-ʾaxīratāni li-l-Majmaʿ al-ʿIlmī". *Majallat al-Majmaʿ al-ʿIlmī al-ʿArabī* 6:12. 555–557.

——. 1948. *al-Muḏakkirāt*. Damascus: Maṭbaʿat at-Taraqqī.

Madkūr, ʾIbrāhīm. 1962. *al-Majmaʿ bayna ʿām wa-ʿām*. Cairo: al-Hayʾa al-ʿĀmma li-Šuʾūn al-Maṭābiʿ al-ʾĀmīriyya.

——. 1963. *al-Majmaʿ bayna muʾtamarayn*. Cairo: al-Hayʾa al-ʿĀmma li-Šuʾūn al-Maṭābiʿ al-ʾĀmīriyya.

——. 1964. *Majmaʿ al-Luġa al-ʿArabiyya fī ṯalāṯīna ʿāman, 1933–1962: Māḍī-hi wa-ḥāḍiru-hu*. Cairo: al-Hayʾa al-ʿĀmma li-Šuʾūn al-Maṭābiʿ al-ʾAmīriyya.

——. 1981. *Majmaʿ al-Luġa al-ʿArabiyya fī ʿīdi-hi l-xamsīnī*. Cairo: al-Hayʾa al-ʿĀmma li-Šuʾūn al-Maṭābiʿ al-ʾAmīriyya.

Madxal ʾilā Majmaʿ al-Luġa al-ʿArabiyya al-ʾUrdunnī. 1988. Amman: Manšūrāt Majmaʿ al-Luġa al-ʿArabiyya al-ʾUrdunnī.

Majallat al-Majmaʿ al-ʿIrāqī. 1950. 383–384.

Majalla Majmaʿ al-Luġa al-ʿArabiyya bi-Dimašq, Qānūn al-Majmaʿ. 2003. 78:1.109–136.

Maṭlūb, ʾAḥmad. 1983. *Ḥarakat at-taʿrīb fī l-ʿIrāq*. Baghdad: al-Munaḏḏama al-ʿArabiyya li-t-Tarbiya wa-ṯ-Ṯaqāfa wa-l-ʿUlūm.

Naṣṣ, ʾIḥsān an-. 1996. "Majalla Majmaʿ al-Luġa al-ʿArabiyya". *Majalla Majmaʿ al-Luġa al-ʿArabiyya bi-Dimašq* 71:1.105–116.

Rāwī, Nājiḥ ar-. 2002. "Ittiḥād al-Majāmiʿ al-Luġawiyya al-ʿIlmiyya al-ʿArabiyya". *Majallat al-Majmaʿ al-ʿIlmī al-ʿArabī* 49:1.5–16.

Sawaie, Mohammed. 1983. "Arabic language academies as language planners". *Languages in the international perspective*, ed. Nancy Schweda-Nicholson, 56–65. Norwood, N.J.: Ablex Publishing Corporation.

Tanūxī, ʿIzz ad-Dīn at-. 1935. "Tašrīḥ ad-darrāja". *Majallat al-Majmaᶜ al-ʿIlmī al-ᶜArabī* 13:12.363–368.

ʿUbāda, ʿAbd al-Fattāḥ. 1928. "al-Majmaᶜ al-Luġawī wa-l-Majmaᶜ al-ʿIlmī". *al-Hilāl* 36:3.305–309.

ʿUmar, ʾAḥmad Muxtār. 2002. *ʾAna wa-l-luġa wa-l-Majmaᶜ*. Cairo: ᶜĀlam al-Kitāb.

Xaṭīb, ᶜAdnān. 1969. *al-Majmaᶜ al-ʿIlmī al-ᶜArabī: Majmaᶜ al-Luġa al-ᶜArabiyya bi-Dimašq fī xamsīna ᶜāman*. Damascus: Maṭbaᶜat at-Taraqqī.

Mohammed Sawaie (University of Virginia)

Language Acquisition → First Language Acquisition; Second Language Acquisition

Language and Gender

1. Introduction

The Arabic *fuṣḥā* has two gender-linked characteristics: it is not a mother tongue, and it entertains a diglossic (→ diglossia) relationship with the dialectal Arabic mother tongues with which it co-exists. Both characteristics make of Arabic a typically 'public' language in an overall patriarchal context where 'public' denotes 'male power', as opposed to 'private', which denotes 'women's realm' (El Saadawi 1980; Mernissi 1997; Sadiqi and Ennaji 2006). The study of Arabic from a gender perspective is still in its beginnings in spite of the fact that Arabic sociolinguistics has attracted the attention of scholars worldwide (Fück 1955; Cohen 1962; Al Ani 1978; Ibrahim 1986; Daher 1987; Ferguson 1987; Eid 1988; Suleiman 1994; Ennaji 1995; Holes 1995; Versteegh 1997; Boumans 1998; Haeri 2000; Owens 2001; Rouchdy 2002; Messaoudi 2003; Miller 2004; Caubet 2004). Some of this work uses the variable of 'sex' in deconstructing Arabic usage, but without paying significant attention to the use of gender as an analytical tool in deconstructing the men/women power relationship among Arabic users.

The interaction of Arabic and gender may be attested at two levels: the formal (grammatical) level and the sociolinguistic (relational) level. At the formal level, Arabic exhibits grammatical and semantic gender usages which may be qualified as 'androcentric' (male biased), and at the sociolinguistic level, Arabic is more often used in male-associated than female-associated contexts. It is important to note that although Arabic is androcentric, the claims made here should not be understood in strong Whorfian terms according to which language determines thought/culture, and culture determines language. Such claims would be too strong and at best unrealistic. Hudson (1996) makes clear that meaningful claims in Whorfian terms have to be carefully qualified and empirically established, which is not always feasible. Thus, the observations and claims made about the androcentric nature of Arabic are not related to Arab culture in a direct way. Whatever links exist in this respect must be mediated and indirect. Various factors are involved in this mediation such as the speaker's personal judgment and the general context of language use.

2. Formal androcentricity in Arabic

Scholarship on Arabic grammar contains extensive accounts of gender as a grammatical category (→ gender). Such grammatical accounts of Arabic gender have often been presented in androcentric terms. The term 'gender' was first used by grammarians; it is only in the mid-1970s that feminists took up the term and used it as an analytical tool to deconstruct the power relation between men and women in given societies and cultures. Thus, for example, ʾAbū Bakr Ibn al-ʾAnbārī (d. 328/940), a reputed medieval Arab grammarian, not only investigated the gender system of Arabic grammar in his *Kitāb al-mud̲akkar wa-l-muʾannat̲* in great and impressive detail, but he accompanied his investigation with typically androcentric explanatory comments on why things were the way they were so far as gender marking was concerned. According to this scholar, Arabic exhibits two types of gender markers: masculine and feminine (there is no morphological encoding of a neuter). These markers appear on verbs, nouns, adjectives, determiners, and quantifiers and may be used to signal grammatical agreement between these various categories. Thus, adjectives generally agree with the noun they modify in person, number, and gender, as in:

(1) *muʿallim-at-un*
schoolteacher-fs-Nom
ʾanīq-at-un
elegant-fs-Nom
'an elegant female schoolteacher'

But this pattern is far from being regular, as some adjectives may not show feminine gender agreement with the noun they modify, as in (2) and (3).

(2) *imraʾ-at-un* *ʿāšiq-un/-at-un*
woman-fs-Nom lover-fs-Nom
'a female lover'

(3) *baqar-at-un* *ḥalūb-un/-at-un*
cow-fs-Nom milky-fs-Nom
'a milky cow'

Such forms are not very frequent and could easily have been listed as exceptions to specific rules, a common feature of natural languages. What is interesting, however, is the explanation that Ibn al-ʾAnbārī gives to account for them. According to this author, the masculine forms *ʿāšiqun* and *ḥalūbun* in (2) and (3) are 'better' than the feminine forms *ʿāšiqatun* and *ḥalūbatun* because these forms denote "intensity and abundance", which are typically masculine qualities that are more associated with men and male attributes than with women and female attributes. Ibn al-ʾAnbārī adds in relation to *ʿāšiqun* 'lover [masc. sg.]' that 'being in love', which is closely related to 'courting', is a typically male state, as only men are supposed to 'show' or 'express' love.

The formal androcentricity attested in the comments on the examples above are far from being neutral; they stem from the subjective views of the grammarian that could not but creep into his supposedly scientific renderings of the gender system of Arabic. These comments also reflect the subordinate position in which women were (and still are) held in relation to men in the patriarchal Arab-Islamic societies and cultures. The force of such comments resides in the fact that they emanate from outstanding scholars who had considerable influence on their contemporaries' gender views.

Another example of grammatical androcentricity is attested in the fact that the masculine normally precedes the feminine in expressions and sentences, as in *rajulun wa-mraʾatun, ṭiflun wa-ṭiflatun*, etc. Again, this precedence would have seemed 'normal', except for the accompanying comments of grammarians such as Ibn al-ʾAnbārī (cited in Abu-Risha 1996: 31–32):

> The proof that the masculine precedes the feminine is that when you say *qāʾim* 'standing [3rd pers. masc. sg.]' and *qāʾima* 'standing [3rd pers. fem. sg.]' and *qāʿid* 'sitting [3rd pers. masc. sg.]' and *qāʿida* 'sitting [3rd pers. fem. sg.]' and *jālis* 'sitting [3rd pers. fem. sg.]' and *jālisa* 'sitting [3rd pers. fem sg.]', you find that the feminine contains additional material, and words in which something is added to the root are secondary. And when you see something from a distance and you do not know what it is, you say: 'a woman, an animal or something like that'.

Ibn al-ʾAnbārī's relegation of women to a secondary position in the above quote and his association of them with animals and unknown entities is reminiscent of the folk ideology still prevalent in the Arab-Islamic world whereby males are given precedence over females, as reflected, for instance, in the contemporary Moroccan proverb *lemṛa u leḥmaṛa ma ka-yḍayfuš* 'a woman and a donkey should not be treated as guests', i.e., they should not be served lest they become spoiled. Indeed, the comments of traditional Arab grammarians can be read only as a particular kind of language ideology, which often leads to stereotypical and sexist views in society at large.

Grammatical androcentricity is also attested in the derivation hypothesis on which a great part of the morphology of Arabic is based. The fact that in this language the feminine gender marker contains one sound or letter more than the masculine marker has been interpreted by traditional and modern grammarians as evidence that the feminine gender is historically derived from the masculine one. Grammar books on Arabic by grammarians like Sībawayhi, Ibn Jinnī, and Ibn al-ʾAnbārī are full of grammatical rules that formalize this derivation hypothesis.

The derivation hypothesis in the case of Arabic gender markers is, however, questionable. In fact, history has shown that human languages are naturally subject to economy (reduction at the level of form) as a result of being spoken. For example, so-called Modern English lost some of the inflection that characterized its ancestors, namely Middle English and Old English. Furthermore, the formalists' accounts of Arabic agreement data are different from the traditional grammarians' accounts; thus, some transformational generative accounts, for example, begin with the feminine form and derive the masculine form by a rule of deletion because it is 'less costly' on the theoretical level, and more predictable from the 'explanatory power' perspective than starting with the masculine and deriving the feminine form (Chomsky 1965).

Likewise, as languages are subject to the rule of economy, it is more logical to assume that the original form in Arabic is the feminine and that the masculine is obtained by the 'shrinking' process which languages undergo. These accounts show that the traditional Arab grammarians' derivation hypothesis was advanced and maintained mainly because it served sociocultural purposes and had a social meaning that fits within the overall Arab-Islamic patriarchy where women were subordinate to men, and hence men needed to 'grammatically' precede women. Additionally, the fact that the masculine usually doubles as the generic in grammatical forms or word choice makes it the social norm from which the feminine derives, hence excluding the feminine from the generic. Thus, the word *imraʾa* 'woman' in Arabic is said to derive from *marʾ* 'person', but only *marʾ* is used generically because it is grammatically masculine. Likewise, words like *al-muwāṭin* 'the citizen [masc. sg.]', *al-ʿāmil* 'the worker [masc. sg.]', or even *al-ʾustāḏ* 'the teacher [masc. sg.]' are used generically, although there are at least as many female as male citizens, workers, and teachers, and in spite of the fact that Arabic contains gender morphemes even in the dual form. This seems to be a general feature cross-linguistically: in many languages the masculine noun is used generically.

These facts show that the male social symbolic power is carried over to the grammatical masculine inflection. In the grammatical hierarchization process, as well as in the grammatical comments that explain this hierarchization, the masculine category is commonly taken to be bigger, unmarked, and higher than the feminine category because the social category 'male' has a bigger and higher status in Arab-Islamic societies and cultures than the 'female' category. As a result, the feminine category is socially viewed as the smaller, marked, and 'degenerate' version of the male category. A further supporting example of this view is the grammatical use of the term *al-ʾinsān* 'man/person'. Although this term may take both the masculine and the feminine gender markers in the singular, only the masculine appears when it is pluralized: *ʾinsān-at-un* 'person [fem. sg.]' but *ʾunās-un* 'people [masc. pl.]' and not **unās-ātun* 'people [fem. pl.]'. Furthermore, although the term *ʾinsān-un* 'man/person' is used generically, it cannot be used with feminine agreement markers: **qaraʾa-t al-ʾinsān* lit. 'read-she man/person', i.e., 'the person read' is ungrammatical. Likewise, although *ʾinsān-un* has a feminine counterpart, *ʾinsān-atun*, the term *fardun* 'individual' does not: **fard-atun* 'individual [fem. sg.]'. Indeed, the grammatical expression of individuality is predominantly male in Arab culture – except in the case of singulative nouns like *namlatun* 'ant', *ḏubābatun* 'fly', *dajājatun* 'chicken', where the feminine gender is used precisely for the individual. The expression *ʿabdu rabbihi* 'God's servant', a modest way of referring to oneself, has a feminine counterpart, *ʿabdatu rabbihā* 'God's (female) servant', which is never used. The reason is that the expression *ʿabdu rabbihi* was first used in all-male public formal domains.

At the syntactic level, full → agreement between the subject and the verb obtains in both VSO (Verb-Subject-Object) and SVO (Subject-Verb-Object) sentences and even in the dual form of nouns and verbs in Arabic. However, when the subject is plural, Arabic shows an interesting idiosyncrasy because agreement ceases to be symmetrical in this context, as in (4) and (5).

(4) *jāʾ-a* *l-ʾawlād-u*
came-3ms Def-boys-Nom
wa-l-banāt-u
and-Def-girls-Nom
'the boys and the girls came'

(5) *jāʾ-at* *al-banāt-u*
came-3fs Def-girls-Nom
wa-l-ʾawlād-u
and-Def-boys-Nom
'the girls and the boys came'

Examples (4) and (5) are characterized by the fact that the subject is formed of two coordinated nouns: a masculine and a feminine. In (4), the masculine gender agreement obtains between the verb and the subject because the nearest noun to the verb is masculine, and in (5) the gender agreement is feminine because the closest noun to the verb is feminine.

Grammatical gender in Arabic can be rather complex (Wright 1981; Benmamoun 1996; Harrell 2004, among many others). In fact, the appearance of the gender feature in Arabic does not always depend on the bipolar male/female opposition, as in (6).

(6) *qāla-t*	*al-ʾanbiyāʾ-u*
said-3fs	Def-prophets-Nom
‘the prophets said’	

In (6), the verb is in the singular form, whereas the subject is in the plural form, and interestingly, the agreement on the verb is feminine and singular. This type of agreement is referred to in the literature as ‘deflected agreement’ and appears on the verb only when the latter is initial in the sentence and when the subject is plural. If the subject precedes the verb, the feminine gender feature disappears: *al-ʾanbiyāʾu qāl-ū* ‘the prophets said’, where the morpheme *-ū* ‘they [masc. pl.]’ refers to *al-ʾanbiyāʾu* ‘the prophets’. The appearance of the feminine singular ending *-at* in (6) is, therefore, due to the syntax of the words *qāla-t* and *al-ʾanbiyāʾu*, i.e. their specific distribution in the sentence in (6).

As it does not match the gender of the agreeing constituents, this type of agreement may be termed ‘functional’ in the sense that it is not gender-dependent and pertains rather to the internal grammatical distribution of verbs and subjects and not to the relation of words with the outside world. Functional gender marking is not based on the male vs. female notion: in (6) above, the masculine noun *al-ʾanbiyāʾu* is functionally feminine because it agrees with a feminine verb, but it is not grammatically feminine, because its agreement is not based on the male/female opposition by virtue of the fact that all prophets were male.

Arab grammarians hypothesized that in instances such as (6) above, the gender morpheme *-t* refers to *majmūʿa* ‘group’, that is, to *majmūʿa min al-ʾanbiyāʾi* ‘a group of prophets’, where agreement takes place with *majmūʿa*, which has the feminine ending *-a*. However, here again, *majmūʿa min al-ʾanbiyāʾ* can never refer to a group of females, or even to a mixed-sex group of prophets because, again, all prophets were male. The explanation seems to be that the feminine ending on the verb is used with collectives, i.e. all broken plurals (which seem to function as collectives in Classical Arabic), but also cases like *al-ʾanbiyaʾu* or even *al-ʿArab*, which is almost always used with a 3rd person feminine singular, e.g. *kamā taqūlu l-ʿArab*’; in pre-Islamic Arabic, but even much later, tribal names (like Huḏayl or Ṭayyiʾ) are almost always treated as feminine. As a result, gender in Arabic may sometimes be purely functional (see Sadiqi 2003b for many more examples from the *Qurʾān*).

Overall, formal or grammatical androcentricity in Arabic is mainly due to the interpretations that Arab grammarians gave to grammatical phenomena. Such interpretations abound in grammar books, and it is high time for a new look at grammatical gender in Arabic. Two questions that arise at this juncture are, first, in what specific ways grammatical androcentricity in Arabic is related to the overall sociopolitical background in which this language is used nowadays. And second, what the general relationship of Arabic is to present-day Arab-Islamic women. Possible answers to these and similar questions bring us to the sociolinguistic androcentricity in Arabic.

3. Sociolinguistic androcentricity

Sociolinguistic androcentricity in Arabic can be understood only within the overall sociocultural framework within which it is created and perpetuated (Badran a.o. 2002; Sadiqi 2003b). Like most societies and cultures today, Arab-Islamic societies and cultures are patriarchal. However, patriarchy is far from being uniform across cultures; it differs from culture to culture. Arab-Islamic patriarchy is based on the notion of space dichotomy (El Saadawi 1980; Mernissi 1997): men are associated with the public space and women with the private space. This ‘space’ notion (*ḥudūd* ‘frontiers’) is not only spatial but also linguistic and symbolic. Thus, in addition to public places being associated with men and private places with women, public languages like Arabic are associated with men and mother tongues with women, and public rituals that are culturally symbolic, like Friday prayers, are associated with men, and those that are private, like birth rituals, are associated with women. Further, public spatial, linguistic and symbolic rituals are associated with the male attributes of rationality and reason. Arab-Islamic patriarchy is different from mainstream Western patriarchy in the sense that whereas the former is based on space, the latter is based on the power of 'image', which creates 'models' for men and women. Western women's emancipation has not been brought about by the church or through militancy, but mainly through the power of the great multinational companies, which kept 'guessing' at the needs of women and providing those needs through constant image creation.

The repercussions of the gendered space dichotomy are multifaceted and far-reaching: they not only associate the public space with the outside/exterior and the private space with the inside/interior, but they also imply that the outside is the place of power where the social norms are produced and the inside is the place where this power is exercised. These two spaces are strictly gender based and interact in a dynamic way in the sense that one does not exist without the other. It is true that women can be in some public spaces – for example on the street – but they are not encouraged to stay there as men are; rather, they must do their business and move on. And men do not generally spend time in the kitchen, for example, so the taboo works for them, too, though with very different consequences.

It is in these contacts that gender identities are constructed and power is negotiated. The private space is culturally associated with powerless people (women and children) and is subordinated to the public space, which is culturally associated with men, who dictate the law, conduct business, manage the state, and control the economy, both national and domestic. It is true that the strict public/private space dichotomy has been significantly disrupted ever since women started to take jobs outside their home from the 1960s onward. At least this was the case in North Africa; elsewhere, participation of women in work outside the home may have begun at an earlier date, for instance in Egypt. In rural areas, women have always worked on their families' farms, but it is also true that men insert themselves into the private space, where women may have real power (Davis 1983; Sadiqi 2003b), in order to satisfy their needs (food, rest, procreation), and some of men's most important life experiences, such as circumcision and marriage, take place in the private space. Thus, Arab-Islamic men have socially sanctioned power over both the public and private spaces which they direct and control. This control is supported by the various *Šarīʿa*-based family laws. The question to ask here is what the place of Arabic is in this overall space-based patriarchal system.

4. Arabic/gender interaction

Arabic has been very instrumental in creating and maintaining this gendered space dichotomy. In fact, although Arabic co-exists with a number of other indigenous and foreign languages in present-day Arab-Islamic societies and cultures, it has had a special social function ever since it became associated with Islam and was introduced as such by the Prophet Muḥammad in the year 622 C.E. This special function of Arabic made it a powerful tool in the hands of the rulers. Dominant groups in a society tend to achieve power through control of high languages, and it is through this control that they ensure the 'obedience' and 'allegiance' of 'subordinated' portions of the population, including women, as Mary Kaplan (1979) rightly puts it: "Refusal of access to public language is one of the major forms of the oppression of women within a social class as well as in trans-class situations."

Arabic/gender interaction is best perceived through the relationship between this language and the four sites of public power in the Arab-Islamic world: religion, politics, the law, and literacy.

4.1 *Arabic, religion, and gender*

As the language of the *Qurʾān* and the mosque, Arabic is more accessible to, and significant for, men than women. Although Arab-Islamic women strongly feel that they 'belong' to the official religion of their countries, and hence to Arabic as the medium through which this religion is expressed, they do not really participate in public religious practices because their culture does not encourage them to do so.

Most formal speeches involving women's issues are used as an opportunity to remind Arabs and Muslims that a woman's raison d'être is her home and children. Consequently, women's religious space in Arabic, through which religion is expressed, is rather limited and publicly constrained (Sadiqi 2003a). For example, in spite of the fact that many women are erudite in religious matters in the Arab-Islamic world, women's opinions in matters of religion lack authority and are not publicly sought. Even when some women venture to advance religious opinions in books, newspapers, and so on, they are never taken seriously and may even be severely rebuked or attacked, as the cases of Nawal El Saadawi and Amina Wadud attest. Many of Nawal El Saadawi's opinions on religious rituals, such as her view that circling the Kaʿba during the *ḥajj* is a *Jāhiliyya* (pre-Islamic) ritual, were severely attacked by the religious authorities in Egypt. Amina Wadud's leading of a mixed prayer stirred up very hostile reactions across the world. This overall negative attitude toward women's opinions on religious matters, especially those dealing with behavior, is explained by their lack of religious credibility in the eyes of society. As a reaction, many feminists (men and women) attribute this lack of religious authority more to the male-biased interpretations of the *Qurʾān* and the *Ḥadīṯ* (the Prophet's sayings) than to core teachings of Islam (El Saadawi 1980; Mernissi 1997; Wadud 1999).

Thus, women's religious space is more restricted than men's and never coincides with the latter as it is very different from it (Buitelaar 1993). Women often recite Qurʾānic verses in their prayers without understanding what they mean, and listen to official formal speeches on the radio or television without understanding them. Most Arab-Islamic women are not daily exposed to Arabic; unlike men, women, especially younger ones, do not usually attend the mosque and, thus, do not participate in the daily ritual of public prayers as frequently as men. Even when they attend the mosque, women are usually apologetic in this space. They pray in special places, where they may see men without being seen by them. As compensation, women visit tombs of saints and holy sanctuaries of ancestors more often than men in their search for *baraka* 'blessing', which ambiguously intermingles with religion in their minds (Gellner 1969; Doutté 1984). These sacred tombs are generally perceived by women as being associated with religious power. This is reinforced by the important place that religious sites have in Arab-Islamic culture; they are visited for a variety of reasons, which range from seeking to enter paradise after death to imploring God's help for bearing children, especially boys. Overall, women in general, and illiterate ones in particular, have a 'strange' relationship with Arabic: they venerate it but do not really feel spontaneously attached to it the way they are to their mother tongues (cf. Sadiqi 2003b, reporting on the results of a survey in 2002).

The fact that women in the Arab-Islamic world do not publicly announce prayers, pray aloud, or pronounce religious formulae that accompany important religious rites is often exploited by patriarchal ideology and taken as proof that women are not fit for public power. This explains the rather rare use of words like *ʾimāma* 'female leader of prayers', *faqīha* 'female religious consultant', *muftiya* 'female religious legislator', *muṣalliya* 'female leader of prayers', *muqriʾa* 'female reader of the *Qurʾān*', and *mujawwida* 'female reciter of the *Qurʾān*' in spite of the fact that the language does contain them.

4.2 *Arabic, politics, and gender*

Politically, the official standardization of Arabic was a direct consequence of the Arab countries' association with the Arab nation (*ʾumma*) after independence. This consciously constructed alliance was based on the 'one nation, one religion, one language' principle and was needed for cultural unity and cultural identity of newly independent countries in a specific historical era when such unity made genuine political sense.

It is basically for these reasons that Arabic is the official language of all Arab states which, just after independence, joined the Arab League, in which Arabic is the lingua franca. Indeed, on the eve of independence, women were not a priority on the political agenda of the newly formed Arab states in spite of the fact that almost all nationalists used Arabic and women's issues to promise more open and more egalitarian societies after independence.

Women were excluded from the political arena and had to fight for decades to gradually gain some public visibility. Because of their relative exclusion from politics, a general tendency to disqualify women as competent public speakers in Arab-Islamic societies has developed. This state of affairs has created an apparent paradox: women are perceived as 'conservative', in the sense that they preserve oral culture by speaking indigenous, often oral, languages and transmitting cultural values, and 'non-conservative' because they do not use the conservative means of public linguistic expression: Arabic. This paradox, however, makes sense politically in that it highlights the political status of oral and written mediums of language. It is true that both Arabic and indigenous oral languages (such as → Berber in Morocco) are socially defined as conservative, but that is so in very different ways: whereas Berber, for example, is perceived as 'conservative' because it expresses traditional oral literature and folklore in the Maghreb, Arabic is perceived as 'conservative' because it perpetuates traditional written literature, history, and poetry, in addition to the fact that it is the language of the *Qurʾān*, the Holy Book of all Muslims. Of course, the Arab Muslim world is not a uniform world, and all three categories (Arabs, Muslims, and women) are extremely variable, as sociolinguistic and other cultural studies have shown.

4.3 *Arabic, the law, and gender*

Legally, Arabic is the reference and vehicle of the law and its implementation. As the exercise of the law takes place in the public sphere, Arab-Islamic women, especially illiterate ones, do not generally understand the language of the law, and hence often fail to know their rights. A recent national survey of Leadership (a Moroccan NGO) has revealed that at least 87 percent of Moroccan women know nothing about the new 2004 Family Law. Although Arabo-Islamic laws regulating policy-making and the economy are based on liberal, modern, universal laws, those regulating the family and relations and behavior between men and women are still largely based on the *Šarīʿa*, which makes them more inaccessible to women.

It is important to note that Arab-Islamic feminist movements realize the legal power associated with Arabic. Politically aware feminists started to target the improvement of family law, using Arabic to enlarge their audiences and gain public credibility. In Morocco, for example, women feminist writers, activists, journalists, and others, who started by expressing their views in French during the 1960s and 1970s, skillfully switched to Arabic from the mid-1980s onward, especially when giving statements to the media, in an attempt to stop radical Islamists from using the language argument against them. By using Arabic in the public sphere, these women are also seeking a place in the powerful religious and legal spaces. They have succeeded in this respect, as the new Moroccan Family Law attests (cf. Sadiqi and Ennaji 2006). They have exhibited dexterity in the use of the *Qurʾān* and the Prophet's sayings on television to show that Islam as a religion and Arabic as a language are not the prerogative of men only. In so doing, these feminists highlight the fact that Islam preaches universal ideas about equality and tolerance between the sexes. This particular use of Arabic may be seen as being motivated by the wish to instigate ideological change and gain more credibility in society at large.

4.4 *Arabic, literacy, and gender*

Arabic is backed by a centuries-old documented history, literature, poetry, and prose; it is perceived as the language of literacy par excellence. Arabic poetry and literature have always been prestigious forms of symbolic language. The relatively greater number of male scholars and intellectuals both dramatizes the gap between literate and illiterate Arabs, and distances men from women. As Arabic is tightly linked to literacy (it can be learned only at school), the large number of Arab-Islamic women who are illiterate are excluded from using it. The rate of female illiteracy in the Arab-Islamic world varies from country to country. Morocco is one of the most affected countries in this regard: around 60 percent of Moroccan women are illiterate, according to the most recent official 2002 census. The rate is much higher in rural areas. This is one of the factors that make Berber (the indigenous language) and *dārija* (the Moroccan Arabic dialect) more accessible to women than *fuṣḥā*. In other words, the fact that Arabic is learned at school and not acquired during childhood puts it on a pedestal where men, not women, can use it and gain power through its use. Elsewhere in the Arabic-speaking world, the situation may be somewhat different, although one could probably say that the standard form of the language is usually felt to be more associated with men than with women. As a result, the Arabic language and Arabic writings have strong 'masculine' connotations and often result in the false view that thinking and rationalizing are 'male'. In contemporary times, Arab women's relation to literate knowledge is still ambiguous; it is generally believed that knowledge threatens women's 'femininity'. On a more general level, the scarcity of women writers in general is due, according to Kaplan (1979), to a prohibition at a deeper psychological level so far as women are concerned. The idea that poetry and literature are not a woman's domain is deeply internalized in women, according to this author.

Literate women have a less 'detached' attitude toward Arabic, but just like illiterate women, they are subject to a heavy patriarchy; they are not encouraged to be actors in the public sphere, and generally speaking they tend to use Arabic less than men do.

5. Conclusion

Overall, as the language of Islam, politics, the media, and written knowledge, Arabic is the 'recipient' of the dominating 'public' and 'high' culture that constitutes the male domain in Arab countries. It is the language of institutions where the gatekeepers of Arabic are most active. Only males have the right to recite the *Qurʾān* aloud in public, to lead the Friday prayers, to deliver Friday sermons, to slaughter animals while uttering specific religious formulae, to be present and participate orally during marriage and burial rites, to deliver 'important' political speeches, to debate 'serious' literary works. It should be added that there are degrees of exclusion. Women are excluded from leading the Friday prayers and delivering the Friday sermon in all Islamic countries, but in some countries, like Morocco, Egypt, and Lebanon, women do deliver political speeches and debate literary works, and they are seen doing all of this in the media, in abundance. Current research on the use of Arabic in the media, especially in Egypt, shows an increase in the use of the vernacular language (*ʿāmmiyya*) in contexts that used to be reserved for *fuṣḥā*. Some literary genres are written in Egyptian Arabic (poetry and plays); novelists and short story writers also incorporate elements of Egyptian Arabic into their texts. This, of course, may have a positive effect on the visibility of women in public life. Standard Arabic remains associated with formal, influential, and 'serious' language functions in which women's voices are often marginalized, at least in some countries. These are the main factors that mark Arabic as a male language. They do *not* make Arabic a men's language in the literal sense of the term; they simply mean that historically, more men may have been competent in the religious and literary-language-dependent professions, given their greater social opportunities.

Bibliographical references

Primary source

Ibn al-ʾAnbārī, *Muḏakkar* = ʾAbū Bakr Muḥammad ibn Qāsim Ibn al-ʾAnbārī, *Kitāb al-muḏakkar wa-l-muʾannaṯ*. Ed. Ṭāriq ʿAbd ʿAwn al-Janābī. Baghdad: Ministry of Waqf (Religious Affairs), 1978.

Secondary sources

Abu-Risha, Zuleikha. 1996. *al-Luġa al-ġāʾiba*. Amman: Center for Studies on Woman.

Al Ani, Salman. 1978. *Readings in Arabic linguistics*. Bloomington, Ind.: Indiana University Linguistic Club.

Badran, Margot, Fatima Sadiqi, and Linda Rashidi (eds). 2002. *Language and gender in the Arab world. Languages and Linguistics* 9 (special issue).

Benmamoun, Elabbas. 1996. "Agreement asymmetries and the PF interface". *School of Oriental and African Studies Working Papers in Linguistics* 6.106–128.

Boumans, Louis. 1998. *The syntax of codeswitching: Analysing Moroccan Arabic/Dutch conversation*. Tilburg: Tilburg University Press.

Buitelaar, Marjo. 1993. *Fasting and feasting in Morocco: Women's participation in Ramadan*. Oxford and Providence, R.I.: Berg.

Caubet, Dominique. 2004. *Les mots du bled, les artistes ont la parole: Création contemporaine et langues maternelles au Maghreb*. Paris: l'Harmattan.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Mass.: MIT Press.

Cohen, David. 1962. "Koinè, langues communes et dialectes arabes". *Arabica* 9.119–144.

Daher, Nazih. 1987. "Arabic sociolinguistics: State of the art". *al-Arabiyya* 20.125–159.

Davis, Susan Schaefer. 1983. *Patience and power: Women's lives in a Moroccan village*. Cambridge, Mass.: Schenkman.

Doutté, Edmond. 1984. *Magie et religion dans l'Afrique du Nord*. Paris: Maisonneuve and Geuthner.

Eid, Mushira. 1988. "Principles for code switching between Standard and Egyptian Arabic". *al-Arabiyya* 21.51–79.

El Saadawi, Nawal. 1980. *The hidden face of Eve: Women in the Arab world.* London: Zed Books.

Ennaji, Moha (ed.). 1995. *Sociolinguistics in Morocco. International Journal of the Sociology of Language*, 112.

——. 2001. "Women and development in North Africa". Paper presented at the Second Mediterranean Meeting. Florence, Italy, 21–25 March.

Ferguson, Charles A. 1959. "Diglossia". *Word* 15:2.325–340.

——. 1987. "'Come forth with a surah like it': Arabic as a measure of Arab society". *Perspectives on Arabic linguistics*, I, ed. Mushira Eid, 39–51. Amsterdam and Philadelphia: J. Benjamins.

Fück, Johann. 1955. *ʿArabīya: Recherches sur l'histoire de la langue et du style arabe.* Trans. Claude Denizeau. Paris: Didier. (1st ed., *Arabīya: Untersuchungen zur arabischen Sprach- und Stilgeschichte.* Berlin: Akademie-Verlag, 1950.)

Gellner, Ernest. 1969. *Saints of the Atlas.* Chicago: University of Chicago Press.

Haeri, Niloofar. 2000. "Form and ideology: Arabic sociolinguistics and beyond". *Annual Reviews of Anthropology* 29.61–87.

Harrell, Richard S. 2004. *A short reference grammar of Moroccan Arabic.* Washington, D.C.: Georgetown University Press.

Holes, Clive, 1995. "Community, dialect and urbanization in the Arabic-speaking Middle East". *Bulletin of the School of Oriental and African Studies* 58:2.270–287.

Hudson, Richard Anthony (ed.). 1996. *Sociolinguistics.* 2nd ed. Cambridge: Cambridge University Press. [Ch. 3: "Language, culture and thought", 70–105.]

Ibrahim, Muhammad. 1986. "Standard and prestige language: A problem in Arabic sociolinguistics". *Anthropological Linguistics* 28.115–126.

Kaplan, Mary. 1979. *The Jewish feminist movement in Germany: The campaigns of the Jüdischer Frauenbund, 1904–1938.* Westport, Conn.: Greenwood Press.

Mernissi, Fatema. 1997. *Rêves de femmes: Une enfance au harem.* Casablanca: Editions Le Fennec.

Messaoudi, Leila. 2003. *Etudes sociolinguistiques.* Kenitra: Faculté des Lettres et Sciences Humaines, Université Ibn Tofail.

Miller, Catherine. 2004. "Variation and changes in Arabic urban vernaculars". *Approaches to Arabic dialects: A collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday*, ed. Martine Haak, Rudolf de Jong, and Kees Versteegh, 177–206. Leiden: E.J. Brill.

Owens, Jonathan. 2001. "Arabic sociolinguistics". *Arabica* 48.419–469.

Rouchdy, Aleya (ed.) 2002. *Language contact and language conflict in Arabic.* New York: RoutledgeCurzon.

Sadiqi, Fatima. 2003a. "Women and linguistic space in Morocco". *Women and Language* 26:1.35–43.

——. 2003b. *Women, gender and language in Morocco.* Leiden: E.J. Brill.

—— and Moha Ennaji. 2006. "The feminization of public space: Women's activism, the family law, and social change in Morocco". *Journal of Middle East Women's Studies* 000.

Suleiman, Yassir (ed). 1994. *Arabic sociolinguistics: Issues and perspectives.* London: Curzon.

Tarrier, Jean Michel. 1991. "Sociolinguistique de l'arabe". *Bulletin des Etudes Orientales* 43.1–15.

Versteegh, Kees. 1997. *The Arabic language.* Edinburgh: Edinburgh University Press.

Wadud, Amina. 1999. Qurʾān *and woman: Rereading the sacred text from a woman's perspective.* Oxford: Oxford University Press.

Wright, William. 1981. *A grammar of the Arabic language.* 3rd ed. Repr. Cambridge: Cambridge University Press. (First publ. 1858–1862.)

FATIMA SADIQI (University of Fes)

Language Attitudes

Social psychologists define 'attitude' as "a psychological tendency…expressed by evaluating a particular entity with some degree of favor or disfavor" (Eagly and Chaiken 1988:269). The entity, or attitude object, may be anything from the color of a car – purple – to a class of behaviors – encouraging democracy – to specific individuals or groups – Oum Kalthoum or immigrants. When the entity is or involves a specific language, language variety, or language practice, social psychologists speak of 'language attitudes'. As Eagly and Chaiken have explained, because attitudes are internal states, they cannot be observed directly but must be inferred on the basis of observable behaviors, and the evaluative responses to which they give rise may be overt or covert – facts that make the study of any sort of attitudes particularly challenging. These psychologists likewise noted that evaluative responses have generally been seen as cognitive, affective, or behavioral in nature. Cognitive responses involve beliefs, and affective responses involve emotions, feelings, and sympathetic nervous system activity, while behavioral responses involve overt actions. Such responses are often in conflict. For example, people trained in linguistics agree that all language varieties are equally systematic and logical – a cognitive response – although they often continue to cringe on hearing a socially stigmatized dialect – an affective response reminding us of the visceral or embodied nature of this category of language attitude. Whether the cognitive or affective response wins out often influences overt behaviors, for example in assessing whether a speaker of the stigmatized variety is employable. Understandably, applied linguists – both students of second/foreign language acquisition and language planners – are likewise interested in questions of language attitudes because these researchers assume that a person's willingness to learn or use (or not learn or refuse to use) a language is linked to motivation, which grows at least partly out of attitudes that influence behavior. Indeed, much of the research related to language learning is couched in terms of motivation rather than language attitudes per se.

Anthropologists have likewise been interested in what psychologists term 'language attitudes' because "a group's beliefs about language, often unexamined beliefs at that, are typically at the heart of its sense of group identity" (Kroskrity 2004:511). Much of the recent work on this subject falls under the rubric of language ideology (Kroskrity 2004; cf. Eagly and Chaiken 1988 for social psychologists' understanding of attitudes and ideology), which can be defined as "the cultural system of ideas about social and linguistic relationships, together with their loading of moral and political interests" (Irvine 1989:255). Irvine's definition highlights several important aspects of language ideologies. They do not represent discrete, isolated attitudes but rather function as systems, linking aspects of language to aspects of social organization and to various sorts of positioned interests. Thus, members of a community come to believe that a particular language or language variety is appropriate for a specific function or context – reciting a sacred text, delivering a formal speech, talking with one's parents, or speaking on television. Similarly, they come to see the speakers of certain languages or language varieties as intelligent or not, as hard working or lazy, and so on. Likewise, political leaders often seek to persuade citizens that by speaking or not speaking certain languages, the society can achieve particular goals – the country can become modern or independent or be (or become) part of some larger entity, such as the Arab world or the European Union. In contrast to social psychologists, whose focus is individual or group differences, linguistic anthropologists are concerned with the ways that beliefs about language(s) structure and are structured by society, influencing everything from face-to-face interactions to local instantiations of abstractions like nationalism. Sociolinguists have drawn on all these strands of research in their analysis on language attitudes, combining them with research in quantitative sociolinguistics, which focuses on the unequal distribution and use across any speech community of various communicative resources. Generally, regardless of discipline, those using the phrase 'language attitude' align themselves with psychology and its methods, while those speaking of 'language ideology' are probably anthropologically informed.

Readers having any familiarity with the Arabic-speaking world should see the many questions about language attitudes that immediately arise in that connection. Likely every paper written about the sociolinguistics of the Arab world (and about many other topics relevant to the area) represents, in a nontrivial sense, a discussion of some aspect of language attitudes in the region.

Following a discussion of relevant research methodologies, this entry discusses five broad areas relating to language attitudes and attitude objects involving Arabic: the Arabic language itself; the relationship between the *fuṣḥā* and national varieties of Arabic; the relationships among various regional and social varieties of Arabic; the relationship between Arabic and other languages used in the region; and language practices like diglossic switching and code-switching.

1. METHOD IN THE STUDY OF LANGUAGE ATTITUDES

Because language attitudes are psychological states related in complex ways to larger abstract language ideologies, they cannot be inspected directly. Researchers use a range of direct or indirect methods in an effort to document their existence and nature. Borrowing from psychology and sociology, researchers using direct methods employ questionnaire-based studies or interviews wherein questions about language are posed. However, many find questionnaire-based studies suspect, contending that their findings are best taken as evidence of overt or imagined norms rather than actual behaviors. Additionally, it is very difficult to determine what speakers specifically have in mind when responding to questionnaire items. If they are asked their opinion about the *fuṣḥā* or French, for example, what kind of *fuṣḥā* or French do they 'hear' as they respond? The *fuṣḥā* or

French as spoken by whom and in what circumstances? Even more complex is the question of how speakers define a language variety like the *fuṣḥā* or why they might define it as they do (Dweik 1997). Perhaps more important, a clear fact about language attitudes and language behaviors more broadly is that language users are unable (and sometimes unwilling) to report accurately their own attitudes and behaviors or those of others. For example, speakers will consistently underreport the use of stigmatized varieties (e.g. dialectal Arabic in contrast to the *fuṣḥā*) or stigmatized practices (e.g. code-switching), and bilinguals are often unable to recall accurately the language(s) used in a particular interaction. In such cases, researchers do not assume that speakers are lying or have faulty memories. Rather, such reports are taken as evidence of the robust structuring power of language ideologies. Speakers generally believe they are reporting accurately on their behaviors and are surprised, shocked, or humiliated when presented with empirical evidence (such as tape recordings) to the contrary. A further problem with questionnaire-based studies is the quickness with which researchers and consumers of the research assume that the reported attitude (e.g. ranking Levantine Arabic as 'more beautiful' than Egyptian Arabic) or behavior (e.g. claiming never to code-switch between Arabic and English) can be taken as evidence of actual behavior – a very serious but nevertheless recurring logical fallacy. Similarly, researchers sometimes write as if the findings of a study done decades ago on a specific population represent current generalized attitudes, a risky assumption given the nature of social change in the Arab world and elsewhere.

To address the problem of the gap between speaker reports of attitudes and actual attitudes they might hold, in the 1960s social psychologists of language developed an especially powerful – and many would claim ingenious – indirect method, the matched-guise test, in work on bilingualism in Canada and elsewhere (Lambert a.o. 1960; Lambert a.o 1965; Lambert 1967). Lambert and colleagues initially recorded balanced English/French Canadian bilinguals reading the same passage twice, once in each language (or guise); they then created an experimental stimulus – an audiotape – on which the recordings, along with additional recordings of the same passages by monolinguals, were ordered so as to minimize the likelihood that experimental subjects would realize they were hearing the same voice more than once. Groups of Anglophone and Francophone subjects were then asked to evaluate each sample on the basis of a number of characteristics or personality traits (height, intelligence, religiousness, sociability, etc.), using a six-point scale of the sort pioneered by Osgood (Osgood a.o. 1957). Although the Francophone subjects expressed favorable attitudes toward French Canadians – their own group – on direct measures of attitudes, they evaluated the Anglophone guises in the matched-guise test more highly on most characteristics than the Francophone guises produced by the same speakers. In other words, there was a mismatch between attitudes elicited by direct and indirect measures. In these early studies, care was taken to employ both direct measures of attitudes toward various groups and indirect measures based on language sample, correlating the two. Over time, it has come to be assumed by some that the findings of the matched-guise test represent a group's 'real' attitudes toward a language or its speakers, attitudes that speakers themselves might not be consciously aware of. Innovations in the methodology have included the use of spoken-language samples, rather than samples read aloud, to make the stimuli more natural, and averaging responses to several stimuli from a specific style when balanced bilingual (or bidialectal) speakers are not available (e.g. El-Dash and Tucker 1975). Although each of these innovations ostensibly solves a methodological problem, it also results in less controlled experiments than the original methodology does, making comparison across studies difficult. Additionally, as Agheysisi and Fishman (1970) point out, designing completely parallel stimuli in diglossic situations is especially challenging because of the functional differentiation that defines appropriate use of each variety. Despite the many critiques of the matched-guise technique (see Fasold 1984:147–179; Bradac a.o. 2001), work on language attitudes continues to presume the use of the matched-guise methodology for many scholars.

Several important matters of interpretation arise with all these quantitative methods. One relates to the problem of the relationship of the sample – those responding to the questionnaire or participating in the experiment – to the population or sampling universe. As quickly becomes clear, most of the questionnaire-based research on Arabic has polled students or faculty, an elite and important group but hardly representative of society as a whole. As studies that seek a broader range of the population remind us, even samples that appear similar can, in fact, be quite different. For example, Bentahila investigated a number of topics related to language attitudes in Morocco, using several different subsamples, and concluded that his subsamples were not, in fact, comparable (1983:110). Additionally, the results of questionnaire-based studies conducted by researchers not trained as psychologists are often reported as percentages; at best, descriptive statistics, rather than inferential statistics, including tests of statistical significance, are used. Hence, generalizing from questionnaire-based studies presents a great challenge, one rarely addressed in research on language attitudes in Arabic (or other languages). A final methodological challenge results from the fact that most research on language attitudes in general remains in many ways locked in the past (cf. Bradac a.o. 2001). The research cited and models used date from several decades ago, and researchers have generally not kept abreast of empirical and theoretical work in sociolinguistics or social psychology and psychology more broadly, whether work on the nature of attitudes (Eagly and Chaiken 1998) or the complexity of constructing valid surveys. In research on language attitudes conducted by those who have not, for example, had training in social psychology, discussions of reliability and validity with respect to questionnaire items or methods are rarely found. The questions and methods used by earlier researchers are merely assumed to be reliable or valid and are recycled or adapted without comment. As Krauss and Chiu (1988:44) pointed out, research in social psychology demonstrates that surveys presenting questions about multiple attitude objects in distinctive contexts – e.g. within a single questionnaire – elicit different responses than those elicited when the objects are examined in different contexts. Sociolinguists would surely not be surprised to find this phenomenon relevant when the attitude objects are languages or language varieties. For example, expatriate and emigrant Tunisians living in the United States created a T-shirt with the slogan “Without Tunisian (Arabic), we’re not Tunisian”, transliterated into Roman script. Despite this overt expression of direct allegiance to Tunisian Arabic, one may expect all these Tunisians to rate the *fuṣḥā* more highly than the dialect on matched-guise tests that test attitudes toward the two language varieties in such a maximally contrastive context. In other words, attitudes reported are heavily influenced by the context in which they are reported, and the likelihood of isolating a group’s real or true attitude is probably limited because of the complex nature of attitudes. For all these reasons, then, research on language attitudes does not figure in theory building on attitudes; none of the references in Eagly and Chaiken’s 1998 survey, for example, discuss language attitudes in the sense treated here. Further, because the social psychology of language (e.g. Bradac a.o. 2001) is concerned with general processes, there has been little focus on the situations of specific languages or language communities, with notable exceptions such as Francophone Canada. Thus, there is no unified body of quantitative research treating attitudes involving Arabic.

Nonquantitative research methods take several distinct forms. Researchers often find evidence of language attitudes in a range of phenomena, including jokes, comedy routines, media and literary representations of language in use, metalinguistic comments in ongoing conversation, and public debates about language, as well as government policies and institutional practices that somehow involve language. Anthropologists in particular have favored long-term participant observation, where researchers use a combination of direct observation of community practices – including language practices – in context, ethnographic interviews, and other sorts of evidence to offer a ‘thick description’ of beliefs and ideologies in light of contemporary theorizing. Haeri (2003) represents such a perspective. The goal of such research is not prediction but rather interpretation.

2. The Arabic language as attitude object

Not surprisingly, there has been discussion of what can be characterized as the Arabic language as an attitude object. First, because definitions of ‘Arab’ often claim that an Arab is ‘one who speaks Arabic’, the language itself becomes an essential, nondetachable component of group membership – often the single such component. Likewise, because Arabic is spoken across a group of contiguous states that refer to themselves as 'the Arab world' (despite the presence now and in the past of various non-Arab groups), the language itself has become an important tool in theorizing and creating nationalism and pan-nationalism, often in exclusionary ways, as thinkers, groups, and polities have struggled to define the proper role and meaning of Arabic in defining nations or supranational entities (e.g. Suleiman 2003, 2004). Such theorizing has been markedly shaped by the experience of Ottoman and European colonization and their aftermath, when these countries were labeled 'underdeveloped' by the West. Finally, because the *Qurʾān* was delivered in Arabic, the language itself has become linked to Islam in ways that many believers, especially Arab believers, experience as essentialist, despite the presence of other faiths in the region and the fact that the majority of the world's Muslims are not speakers of Arabic (although many study it as a liturgical language). Part of the *Qurʾān*'s miraculous nature, for believers, lies in its language as well as the variety and style of Arabic in which it was delivered.

Understandably, then, the Arabic language itself has become a powerful signifier, a ready resource for those who wish to link Arabic and group, political, or religious identity, often collapsing these latter categories. Such uses are predicated on the indexical meanings of the language (on indexicality, see Ochs 1992; Myers-Scotton 1993). In recent years, linguistic anthropologists in particular have focused on the social processes that create links between language or language varieties and abstractions like the nation-state, social descriptors like 'religious' and 'masculine', or social categories like class or religious confession. Attitudes about the language and its nature are, as noted, a significant part of such ideologies. At least since Ferguson (1959a, 1959b), there has been discussion of Arabs' beliefs about the Arabic language. Ferguson used the potentially problematic label 'myths' to refer to such attitudes, and, as Eisele (2003) noted, Ferguson drew broad unqualified generalizations and did not contextualize the imputed attitudes discussed in a larger comparative framework. In particular, Ferguson discussed attitudes or beliefs in three general areas: the language's superiority resulting from what are perceived to be its beauty, logical structure, rich lexicon, and divine nature; speakers' frequent contention that their native dialect of Arabic is closest to the Classical variety; and the language's future, which is envisioned as a variety that is "unified, standardized, universal in the Arab world, used for both speaking and writing, and appropriate for all kinds of literature" (1959b:381). Elsewhere, Ferguson (1959a) discussed a fourth domain of language attitudes, those related to Arabic → diglossia, especially the stigmatized nature of dialectal varieties and the elevated status of the *fuṣḥā*. Ferguson's original framing of these issues has taken on a life of its own. Nader (1962), for example, expanded and critiqued the notion of 'prestige' assumed by Ferguson and, indeed, by many sociolinguists today, and Zeinab Ibrahim's survey-based attempts to test several of Ferguson's contentions led her to conclude that his "observations about Arabic language attitudes are still largely accurate" (2000:23). Eisele's (2003) comments on Ferguson's discussion of these 'myths' and reformulation of it are most promising.

3. The relationship between the *fuṣḥā* and various national varieties of Arabic

Likely no topic relating to attitudes about Arabic has spawned the volume or intensity of commentary as the complex relationship between the *fuṣḥā* and the various national varieties of Arabic, and the proper or possible roles of each. The functional differentiation between the two – often labeled High and Low, respectively – in terms of both modality (written/spoken) and context (formal/informal), which has held for centuries, is usually termed → diglossia. This functional differentiation is predicated on certain sets of attitudes or ideologies, a fact acknowledged by Ferguson (1959a) in his initial characterization of the phenomenon, even as he predicted the social circumstances that might give rise to its reconfiguration (Walters 2003). Ferguson (1959a) likewise explained that diglossic situations inevitably give rise to communicative tensions, which, under certain conditions, come to be constructed as social problems. As he also noted, the simple existence of spoken varieties of Arabic is seen as a problem – indeed, a crisis – by many commentators on the language. Bluntly put, “Arabs still revere *fuṣḥā*” (Z. Ibrahim 2000:23) and are quick to speak in pejorative terms of the dialects. Indeed, as Hussein and El-Ali (1989:41) insightfully pointed out, with regard to matched-guise experiments that include the *fuṣḥā*, “The colloquial varieties cannot withstand comparison with it”. In fact, such attitudes follow logically when one considers recent work in language ideology and the nature of diglossia. In her efforts to understand discrimination based on speakers’ accents in the United States, Lippi-Green (1994:166), building on the work of Milroy and Milroy (1985), posited the notion of standard language ideology, which she defined as “a bias toward an abstracted, idealized, homogeneous spoken language which is imposed from above, and which takes as its model the written language. The most salient feature is suppression of variation of all kinds”. In other words, the written language becomes the model for speaking (even though, as Milroy and Milroy [1985] amply demonstrated, there are many functional reasons to expect spoken and written language forms of any language to differ). We can predict that an ideology of the sort described by Lippi-Green, which is associated with the language variety accorded overt prestige, will be found in all speech communities where a superposed, consciously standardized variety of language exists (e.g. French, whether in Paris or Algiers). In many regards, we can claim that diglossia of the sort found in Arabic represents the most complete instantiation of standard language ideology. The *fuṣḥā* is, indeed, superposed; being no one’s native language, it is necessarily idealized and abstracted (though not ill-defined, *pace* Kaye 1972; see Walters 2003:104, n. 4). Because of its history, the *fuṣḥā* is inevitably and essentially associated with writing because the texts of the Islamic heritage, including the *Qurʾān*, and what many see as the zenith of Arab thought – a glorious past never to be equaled – are written in this variety. Similarly, it is the written version of the *fuṣḥā* that serves as the basis for its spoken variety: when speaking the *fuṣḥā*, one is, in some very real sense, talking like a book or, minimally, using the language variety historically associated with books and literacy to speak aloud. Given the symbolic loadings of these phenomena, the *fuṣḥā* has understandably come to be imbued with near-totemic power. There is persistent hand-wringing across the Arab world about → lexical variation in the *fuṣḥā* as used in various countries (especially in the domain of journalism) and about the language’s contamination because of lexical borrowings from other languages and about structural and stylistic influences from those same languages. This hand-wringing provides strong evidence for the desire to suppress variation of the sorts linguists expect to find in any living language, especially one in intense contact with other languages and other varieties of the same language. Lippi-Green’s (1994) description shares much with Ferguson’s (1959b) discussion of the idealized future of Arabic; in fact, they are identical in many regards.

In the same way that the circumstances of Arabic diglossia might lead us to predict the nature and strength of the ideology associated with the *fuṣḥā*, as well as the power granted it, they set up a polarized structural contrast with the dialects, which come to represent symbolically the absence of everything the *fuṣḥā* is claimed to be. Predictably, we might argue, some have come to the defense of the dialect as a way of overcoming the very real practical challenges Arabic diglossia presents, whether educating all a country’s citizens or creating a vibrant and vital literary language that reflects and instantiates daily life and contemporary national culture. Indeed, there have been proponents of elevating local varieties so that they could be used for functions traditionally reserved for the *fuṣḥā*, much as European vernaculars were elevated in status, ultimately replacing Latin as the language of learning. Suleiman (2004, Chap. 3) has offered an excellent survey and analysis of the history of these debates. As he aptly states, “Acting as a proxy for extralinguistic issues, the SA [Standard Arabic] versus dialects debate is used to signal metonymically the concern with identity, modernization, tradition, change, and globalization.... The fact that the main arguments in [this debate] tend to be repeated *ad nauseam* in each generation and, in the same generation, in different locations testifies to the perennial nature of the issues that animate them” (Suleiman 2004:93–94).

Likely the most provocative and fully developed recent such argument is that of Haeri

(2003), who analyzed the conflicts inherent in a situation in which 'ordinary people' are expected to use a 'sacred language' in their daily lives in the context of a nation-state in the modern world. Despite the richness of Haeri's insights in problematizing the Egyptian situation, she did not discuss the practical problems associated with elevating the status of the dialectal variety, including the challenge of standardizing Egyptian Arabic and the sorts of social inequalities any such standardization project would create or replicate. Nor did she engage the predictable (indeed, in many ways, ready-made) ideological critiques of any such suggestion. These critiques include the potential loss (or even destruction) of the 'unity' assumed to exist among Arab nations, a concern of those who claim that the *fuṣḥā* is the glue holding Arab culture and the Arab world together (given the essential role assigned to language in defining 'Arab'), and the potential loss of access to the texts of the past, especially a concern of those who link Islam and the *fuṣḥā*. Following Suleiman (2004), it may be predicted that Haeri's arguments will change the attitudes of few, if any, Egyptians or Arabs more broadly and that her status as non-Egyptian and non-Arab (though Muslim, trained in the West) will result in her being labeled a traitor (especially to Islam) in collaboration with modern-day Orientalists. That we can accurately predict such a response teaches important lessons about how language attitudes become part of larger social discourses. Such ideologies become part of the creation and perpetuation of social structures, granting power or hegemony to some but not others. Suleiman, for example, outlined the arguments made, on the one hand, by those who cast themselves as defenders of the *fuṣḥā* and simultaneously Islam, Arab nationalism, authenticity, purity, or any of several other values – positions that align them with a preexisting dominant ideology, and, on the other hand, by those who argue for modernizing the *fuṣḥā* or for expanding the roles in which the use of dialectal Arabic is legitimated, often in the name of democratizing the language or making it more useful in a globalized marketplace. Within such a context, the available positions to which one can stake claim are limited. Situations like this one demonstrate that it is not simply that we, as members of a society, choose to subscribe to particular ideologies, including language ideologies, but rather that, in a real sense, ideologies choose us, based on our position in the social order, our life experience, and our value commitments of various sorts.

From this perspective, ideologies can be said to beckon us, and we recognize ourselves in them much as we might respond to a stranger on the street hailing us with 'Sir' or 'Madame' or imagine that some piece of clothing we see in a shop window would suit us perfectly. An understanding of such structural configurations of attitudes and ideologies helps account for the findings of research on attitudes toward the *fuṣḥā*, the dialect(s), and projects like Arabization. Predictably, questionnaire-based surveys of attitudes, whether using direct questions or the matched guise in which the *fuṣḥā* appears alongside one or more dialects, yield results favoring the former. Thus, the sample of high school and university students studied by El-Dash and Tucker (1975) ranked the *fuṣḥā* high, in contrast to Egyptian Arabic and English spoken with three accents, on intelligence, likeability, religiousness, and leadership.

Hussein and El-Ali (1989) studied a sample of 303 university students enrolled at Yarmouk who ranked Modern Standard Arabic highest in status in contrast to three local dialects of spoken Arabic. Sawaie's (1994) study, which manipulated sociolinguistic variables in a matched-guise context, found that his sample of 321 Jordanians and Palestinians, mostly university students in their 20s, ranked the variants associated with the *fuṣḥā*, especially the sociolinguistic variant [q], as most elegant, in contrast to those associated with three local dialects of spoken Arabic. Al-Haq's (1998) sample of 211 faculty members at Yarmouk agreed that the *fuṣḥā* was superior to all other varieties, that it was beautiful, and that it marked a person as well educated; this sample expressed overwhelming support for Arabization of the curriculum. Respondents likewise agreed that Arabization represented 'a national and academic duty' (and thus, to raise any questions about Arabization is to set oneself up to be labeled unpatriotic). Such findings led Al-Haq to conclude that "using the colloquial does not entail attachment and loyalty to it, but rather is functional" (1998:21). Given the findings of other research, however, such an interpretation is likely too simplistic. Arabs generally prefer their own national dialect to those from elsewhere, clear evidence of attachment and loyalty in the ears of many. Further, Al-Kahtany's (1997) survey of 40 male university and postgraduate students from fourteen countries in the Arab world studying in the United States found that respondents understood and appreciated the functional differentiation of varieties associated with diglossia and did not perceive the situation to be a problem (cf. also Dweik 1997). In fact, Al-Kahtany's small sample of North Africans believed that local dialects could replace Modern Standard Arabic in schooling and the media, evidence of possible regional differences in perceptions of the imagined community of the Arab (and Arabic-speaking) world. Such a finding will predictably be troubling to those who fetishize the *fuṣḥā*. At the same time, it should not be surprising, given the history of North Africa, including the ways in which it is often marginalized in the Arab world.

In fact, we might say that the outcome of surveys of attitudes in which the *fuṣḥā* is placed in a context of contrast with other varieties is overdetermined: multiple factors that cannot easily be teased apart contribute to predictable responses. Additional challenges associated with any such survey – especially one that asks about the subject's ability to speak the *fuṣḥā* – are the question of using self-report data to assess abilities in such a value-laden attitude object (cf. Z. Ibrahim 2000) and the larger, more abstract issue of what speakers believe the *fuṣḥā* to be. Parkinson's work (e.g. 1991, 1993, 1994) is most instructive with respect to this latter question. All these factors help account for the apparent gap between expressed loyalty to and esteem for the *fuṣḥā*, even as speakers continue to use and prefer local varieties – their mother tongue – in their daily lives, if preference is assessed on the basis of practice. An especially interesting and robust finding of quantitative studies of sociolinguistic variation with respect to language attitudes between the dialects and the *fuṣḥā* involves behavior with respect to what Haeri (1996) has termed 'diglossic variables'. Such variables, like (q), which in Cairene Arabic has the variants [q], associated with the *fuṣḥā*, and [ʔ], associated with the dialect, show distributions unlike the sorts of patterns found in most Western speech communities. As Walters (1996) demonstrated, at the most general level these patterns fall out from the nature of diglossia, which presumes very different norms for speaking and writing (and hence predicts different distributions of data gathered from informal interviews and literacy-based prompts like paragraphs, word lists, and minimal-pair lists read aloud – data-gathering procedures most often used by quantitative sociolinguists). Especially relevant to this discussion, however, is the fact that in informal interviews, males as a group favor variants associated with what we might term the 'deterritorialized' written standardized variety, i.e. the *fuṣḥā*, more so than do females, who as a group use a higher percentage of the variants associated with the prestige variety of the local urban spoken dialect (→ language and gender). One way of interpreting this finding, which recurs across the Arabic-speaking world and contrasts markedly with findings in Western urban areas, is to claim that males and females have different attitudes toward these variants (i.e., they assign them different social meanings) and use them differently in constructing their respective gendered identities – masculinity or femininity – as these interact with their social class position. In other words, following Muhammad Ibrahim (1986), we can claim that in Arabic speech communities, because of diglossia we find a standard variety, associated with the *fuṣḥā*, and a local prestige variety (in contrast to Western speech communities, where the two are, with rare exception, the same) and that attitudes to these varieties can differ.

4. Relationships among various regional and social varieties of Arabic

A third relevant area of interest and research involves attitudes toward the various regional and social varieties of spoken Arabic (although national boundaries often do not correspond to regional dialect boundaries – the Arabic of northeastern Algeria is surely closer to that of northwestern Tunisia than it is to the dialect of Arabic spoken in southwestern Algeria, yet both are perceived and treated as Algerian Arabic). Natives of every country in the Arab world have strong attitudes about how people of various regions, faiths, and social groups within their country and across the region speak, and language and language attitudes become components of prototypes or stereotypes about these various groups. As Muhawi (1994) demonstrated, understanding the punch line to many jokes told in the Arab world presumes knowledge of just such linguistically based stereotypes. As nonnative speakers of Arabic learn the language, they internalize such attitudes or at least gain knowledge of them. Students in anthropology, folklore, and cultural studies remind us that in television programs around the world, we can expect the buffoons to speak socially stigmatized language varieties, a reflection of attitudes linking social hierarchy and language use. Interestingly, within a country, stigmatized urban and rural varieties of a language are generally treated differently. Again, around the world, researchers find that city folk are quick to laugh at the country bumpkin's speech, assuming him or her to be simple but trustworthy; in contrast, they are less charitable to the urban poor, whose speech is stigmatized in other ways, as Hussein and El-Ali (1989) and Sawaie (1994) illustrate for Jordan. Suleiman (2004, Chap. 4) has offered a different reading of the Jordanian situation, focusing on urban, Bedouin, and rural *fellahi* variants of the sociolinguistic variable (q) as well as the standard variant, which is also associated with the Druze and Syriacs in Jordan; he contextualized his findings in terms of the demographics of Jordanian society, especially the various waves of Palestinian immigrants to the country, linking the distribution of variants to historical events and the renegotiation of Jordanian nationalism in light of these immigrants. An interesting aspect of language attitudes in Arabic is the value attached to Bedouin varieties of Arabic (Ferguson 1959b; Nader 1962; Hussein and El-Ali 1989), which have been taken as purer or more authentic since the earliest study of the language (Versteegh 1997:57ff.).

Attitudes toward various national varieties of Arabic are, however, more complex. There has been no comprehensive study of this situation, for instance from the perspective of social psychology, one in which a representative random sample of raters from each of the Arab countries evaluates a well-constructed stimulus tape of samples from all the countries where Arabic is spoken natively, using the matched-guise technique. However, some outlines of the situation can be sketched, based on existing research. Herbolich's (1979) matched-guise experiment compared responses to Egyptian, Syrian, Saudi, and Libyan dialects by a sample of 80 female and male Egyptians, representing professionals, American University of Cairo students, national university students, and high school students. The sample ordered these dialects in the following way: Egyptian, Syrian, Saudi, and Libyan. (Thus, the study provides additional evidence of a speaker's preference for her or his own dialect.) As Herbolich demonstrated, however, raters often incorrectly identified speakers' nationality. Hence, it was the perceived nationality of the speaker and attitude toward the perceived nationality that influenced various rankings more than actual nationality or actual native dialect. (Similar findings are common in matched-guise studies in other languages, reminding us that raters are guided by their assumptions of what particular groups sound like rather than by how members of those groups in fact sound.) Al-Kahtany's matched-guise study of 40 Arab students in the United States included conversational stimuli representing six dialects, but his 1997 article focused only on attitudes toward Modern Standard Arabic and the Damascene dialect. With a sample of 70 male and female university graduates in Egypt and Morocco, Zeinab Ibrahim (2000) found that the Egyptians overwhelmingly ranked their dialect as most beautiful, while the Moroccans favored Egyptian (n=28) over Moroccan (n=25) and Levantine (n=17). (Responses to questions about which dialect subjects preferred males and females to speak showed that attitudes about national dialect were likewise linked in complex ways to the sex of the speaker. One might also question whether Ibrahim's own nationality, Egyptian, might not have influenced some Moroccans to rank Egyptian as they did.) As Ibrahim noted, Egyptians appear 'dialectally insulated' from the remainder of the Arab world, a finding she, like many, attributed to the role Cairo has long played in the entertainment industry – music, cinema, radio, and television. This situation results in Egyptian Arabic's being almost universally understood across the Arab world – especially the Cairene dialect of Egyptian Arabic. Ibrahim also pointed out that changing technology, especially the satellite dish and pan-Arab television stations, means that Arabs are (or can be) exposed to varieties of Arabic other than their native variety to a degree formerly not possible. Interestingly, she likewise contended that "Arabs usually do not feel the urge to learn another Arabic dialect", a claim called into question by S'hiri (2002), who examined Tunisians' understanding of their own accommodation to the Arabic of those she termed Middle Easterners. As S'hiri explained, because Middle Easterners often complain that they cannot understand Tunisian Arabic (and North African varieties of Arabic more broadly), Tunisians are encouraged, if not forced, to engage in asymmetrical accommodation, suppressing Tunisian forms while using forms more familiar to their interlocutors. For S'hiri, the causes of such a situation are both cultural and linguistic. Indeed, the differences in the verb morphology and stress, as well as lexis, between the Western and Eastern varieties of Arabic are great (Versteegh 1997, Chap. 10), and Easterners have little or no exposure to North African, or Maghrebi, varieties of Arabic. However, according to S'hiri's interviewees, Eastern Arabs often claim that because of the Maghreb's history of colonization by the French and the lexical influences on North African Arabic from Berber, Italian, Spanish, and French, they find the Maghrebi varieties incomprehensible. Not surprisingly, such attitudes reflect a much broader set of attitudes about the Mashriq, or Arab East, and about the Maghreb, or Arab West, a distinction replicated in many interesting ways. Even scholars writing about language as a unifier in the Arab world often focus on what they term 'the Middle East', noting that the North African situation is so distinct as to require separate treatment.

The relevant issue here is not the actual degree of difference or the need for separate treatment, but rather the observation that in focusing on the Middle East and so labeling it (instead of using a label like 'the eastern part of the Arab world', much closer to the meaning of the Arabic term *mašriq*), scholarly practice helps perpetuate and reify an East/West distinction, one in which North Africa, including its varieties of Arabic, is marginalized.

5. ARABIC AND OTHER LANGUAGES IN THE REGION

A fourth area of relevant research has involved attitudes toward Arabic as compared with those toward other languages in the region, whether indigenous (e.g. Berber and Kurdish) or superposed (English, French, Hebrew, Italian, and Spanish). With regard to indigenous languages, attitudes influence the behavior of the society and those in power with respect to the treatment of the language, especially its legitimated presence or absence in the public domain (e.g. education, media), and its speakers (e.g. whether those who continue to use the language are seen as backward or, more seriously, as not fully loyal to the nation-state). Within minority-language communities, the issues include whether the ethnic language is seen as a detachable or nondetachable part of group identity, whether speakers of these languages assimilate to the dominant language by choice or force, and how attitudes change across time, as they clearly have in Morocco and Algeria with respect to Berber since the 1990s. Because of the nature of Arab nationalism and nationalism in the Arab world, reliable data do not exist on minority-language speakers. Questions about such matters are not asked on censuses inasmuch as the national governments see such information as potentially divisive and threatening to national unity. Similarly, there is limited research on these languages or attitudes about them, with the exception of → Berber (→ language shift: Amazigh). Discussions of the situation of Berber in North Africa include Bentahila and Davies (1992), which also considers → Judaeo-Arabic, Ennaji (1997), Sadiqi (1997), Marley (2004), and Errihani (2006). There has been far less written on → Kurdish, but see Hassanpour (1992) on Kurdish in Iraq and elsewhere.

In contrast, there has been extensive discussion of superposed languages, notably French, English, and Hebrew. Discussions of French and English cannot, in most cases, be divorced from impassioned debates about Arabization, i.e. the use of some form of Arabic as the language of government, education (including higher education), and the media, in the postcolonial era and more recent discussions of competitiveness in times of globalization. (In this regard, Arabization represents a case of language spread: Arabic – or the standard variety of it – is to be used in contexts previously reserved partly or exclusively for the use of some other language.) Just as those defending the *fuṣḥā* against the dialects position themselves as defenders of specific cultural values, those arguing for Arabization and a reduced role for European languages do likewise. Bentahila (1983, Chap. 6) and Mouhssine (1995) have analyzed the ideologized contradictions such positions have represented in Morocco. Studies of language attitudes in North Africa have offered especially interesting findings with respect to language and what Bentahila (1983, Chap. 3) termed 'views of the world'. In sentence-completion tasks, his 87 respondents, aged 17–38, representing a range of professions and various regions of Morocco, demonstrated very different ideas, depending on whether they responded in Arabic or French, about what a person needed in order to have a successful life; the nature of marriage; one's duty to society; and one's own goals in life, among other views. Using a sample of 120 Tunisians aged 20–28, stratified by education, sex, class, and size of town or city of residence, Riguet (1981–1982a, 1981–1982b) used French- and Arabic-language versions of a questionnaire about attitudes regarding family, work, and society to investigate whether the language of the questionnaire influenced responses. (Thus, questions were not asked about language, but as in the study by Bentahila, language was an independent variable.) Riguet found that educated subjects reported different attitudes on some items depending on the language of the questionnaire, with the French-language version favoring certain modernist attitudes. At the same time, he did not find any simple or clear distinctions according to the language of the questionnaire; however, he found abundant evidence of a clear attachment to Islam, regardless of the questionnaire used. While questionnaires about language attitudes often demonstrate that students across the Arab world exhibit instrumental, rather than integrative, motivations for studying Western languages, reality is likely more complex. Lambert's (1967) instrumental/integrative dichotomy itself has been called into question by researchers in second-language acquisition. Further, one might contend that students who report wanting to master English in particular, because of its current global status – a motivation that might traditionally be seen as instrumental – simultaneously seek to integrate themselves into a globalized economy that uses English as its language and is much influenced by Anglo-American capitalist practices that currently may have little to do with American or British culture directly. Additionally, mastery of a Western language represents an important kind of embodied symbolic capital (Walters 1999b), a fact not lost on future elites across the Arab world. Knowledge of a Western language may be useful in getting a job or helping one's country modernize, but it also increases one's social status and may help women, in particular, in the marriage market. It does, however, appear that speakers from across the Arab world have different sets of attitudes about mastering second/foreign languages and particularly about speaking them 'without an accent', which, of course, means speaking them with an accent that resembles or is indistinguishable from that of a native speaker of the language.

North Africans and Lebanese, citizens of countries that were heavily colonized by the French (and where French is often considered a second, rather than foreign, language), appear more willing to adopt native-like accents when speaking French, English, or other languages, a practice those from other parts of the Arab world sometimes criticize, taking it as evidence of disloyalty to some construction of Arabness. Females, more so than males, seem willing to adopt native-like accents and are often expected to do so (e.g. Diab 2000). Representative discussions of attitudes toward English in the Arab world include El-Dash and Tucker (1975) for Egypt; Walters (1999a), Lawson and Sachdev (2000), and Daoud (2001) for Tunisia; Malallah (2002) for Kuwait; and Abu-Rabia (2003) for Israel. On French, see Bentahila (1983) and Marley (2004) for Morocco; Stevens (1983), Lawson and Sachdev (1997), and Daoud (2001) for Tunisia; Taleb Ibrahimi (1997) for Algeria; and Diab (2000) for Anglophone Lebanon.

Examining the status of Arabic in → Malta, a country whose linguistic situation does not fit comfortably with those of the rest of the region, Scirha (2001) sought to explain why secondary school children there showed little interest in studying the language although their native language is derived from Arabic and considered by many to be a variety of Arabic. For many reasons, the linguistic situations in Palestine and → Israel and attitudes toward → Hebrew among Arabs have been the subject of much research on language attitudes. Research on this area includes attitudes toward Hebrew and Arabic of Israeli Arabs, versus those of Palestinians in a divided community (Amara and Spolsky 2001) and of Israeli Arabs in contrast to Israeli Jews more broadly (Lambert a.o. 1965; Kraemer and Olshtain 1989; Kraemer and Birenbaum 1993, who also consider English; Ben-Rafael and Brosh 2001). Bentolila examined attitudes concerning → Judaeo-Arabic vs. Hebrew among Israeli Jews of Moroccan origin (Bentolila 2001). Suleiman (2004, Chap. 5) has offered a critique from a Palestinian perspective of much of the work done by Jewish Israeli scholars. Using the specific example of the languages of street signs and the larger issues of the naming or renaming of streets, he likewise documented the 'language conflict' in public spaces in Palestine and Israel. Al-Haq (2000) reported on changing attitudes toward Hebrew among Jordanian university students, who are now permitted to study that language.

6. LANGUAGE PRACTICES: DIGLOSSIC SWITCHING AND CODE-SWITCHING

Experience and research demonstrate that practices like diglossic switching between the *fuṣḥā* and a national variety (Walters 1996) and → code-switching between Arabic and another language evoke strong responses from speakers. Although there has been no empirical research on attitudes toward diglossic switching per se, given the ideologies associated with diglossia, it can be predicted that someone speaking the dialect can almost always use lexical resources from the *fuṣḥā* without criticism, whether single words or set expressions, especially where these fill lexical gaps, when interacting with a comparably educated person. Indeed, in many contexts, especially among educated speakers, one is expected to do so. As an early study of radio Arabic (Schulz 1981) demonstrated, speakers speaking extemporaneously, especially in group settings, often begin in the *fuṣḥā*, but the longer they talk, the more likely they are to move toward using the dialect, as what some researchers in code-switching would term the matrix or base language with frequent and extended embeddings from the *fuṣḥā*. Such language practices enable speakers to demonstrate simultaneously their loyalty to the local dialect and their education by using the *fuṣḥā*. It likewise enables them to overcome an especially complex interactional situation. Because the *fuṣḥā* indexes formality, because it is a superposed variety, and because of the ideologies of correctness associated with it, few speakers are comfortable speaking it for long periods of time – especially if they are expected to use all the case endings. Their switching – whether through using the dialect as a matrix or through using some form of what some would term → Educated (Spoken) Arabic (ESA) – permits them to transform the situation from one of formality to one of solidarity, as connoted by the use of dialectal Arabic. (As Suleiman [2004:8, n. 28] explains, Educated Spoken Arabic and diglossic switching, which he lumps together, have not become ideologized in the way varieties of Arabic like the dialect and the *fuṣḥā* have.) On the other hand, we can predict that if one is speaking the *fuṣḥā*, especially in a formal context, any switch to a dialect or another language must be rhetorically marked (e.g. as humor or solidarity more broadly); otherwise, the speaker is assumed to be unable to speak the *fuṣḥā* well.

There have been several studies of attitudes toward code-switching (e.g. Bentahila 1983, Lawson and Sachdev 2000 on Arabic and French; Hussein 1999 on Arabic and English). Bilinguals are sensitive – and often very sensitive – not only to whether code-switching occurs but also how often it occurs; which specific structural constituents are switched; which social contexts the switching occurs in; and what the relative social status of the person switching is vis-à-vis their own. Thus, it is not at all clear how to interpret the findings of these studies, especially those using matched-guise methodology. Clearly, such code-switching is stigmatized – and doggedly so – even by people who engage in it. The claim that code-switching constitutes a language variety (e.g. Lawson and Sachdev 2000) rather than, say, a communicative resource or language practice, remains problematic. For students of language structure, the claim of a new variety would require evidence of novel structures found in neither variety, structures that cannot be accounted for by existing models of code-switching. Likewise, given the nature of standard language ideologies and language purity, it should come as a surprise to no one that code-switching is often highly overtly stigmatized, especially when it involves two languages, each of which has undergone standardization. At the same time, speakers often willingly acknowledge that they engage in the practice and admit its potential usefulness in communicating.

7. Conclusion

Language, language varieties, and language practices are attitude objects in every speech community that has been investigated. Although the construct of language attitudes has traditionally been associated with social psychology, scholars from a wide range of disciplines are concerned with whatever it is that the label represents. Applied linguists, language teachers of native and nonnative languages, and language planners repeatedly demonstrate the significance of attitudes in their projects.

Anthropologists are quick to remind us that attitudes, including attitudes about language of one sort or another, become constitutive of communities and community membership: To speak a language or language variety (especially if one speaks it natively or speaks it well as a nonnative speaker) is to be aware of sets of beliefs associated with that language by in-group and out-group members, even if one seeks to reject them. Thus, language, or some aspect of it, becomes a proxy for other attitudes or beliefs, often justifying behaviors of many sorts. The Arabic-speaking world offers an especially rich and instructive context in which to examine attitudes associated with language in some way. Because of the language's intimate association with Islam, it presents an important example of the sorts of links believers can create between language and religious experience or practice, as well as the special, indeed, sacred, registers of language that can be created for religious purposes (Nelson 1985). Because of the nature of the Arab world – a group of contiguous states sharing a majority religion and a cultural past linked directly with the *fuṣḥā* but speaking different national dialects as mother tongues and having very different colonial and postcolonial trajectories – language varieties and practices predictably become ideologized. Attitudes toward them figure prominently in the available social and political positions with which individuals and groups can align themselves. Finally, an essential part of Arabic diglossia is attitudinal. It is not merely that, by convention, one writes the *fuṣḥā*, but that, in the minds of many, it is the *fuṣḥā* alone that should be written. For some, to seek to write the dialect or legitimate its use as a written variety is to engage in heresy or to favor national over pan-national interests, thereby playing into the hands of those who would destroy the Arab world.

As Gal and Irvine (2000) remind us, the processes such as those by which the conventional use of the *fuṣḥā* for writing gives rise to the array of attitudes and social, political, and moral arguments one finds in the Arab world merit close attention. Indeed, the complex chain of assumptions and reasoning that licenses such arguments, as one moves from the act of writing to debates about the sacred or nationalism (often linking the two), is by no means simple or transparent. That such links have become naturalized – or taken for granted as natural or commonsensical – in no way renders the Arab world or Arabic unusual; similar complex ideological processes recur across languages, cultures, and societies. At the same time, these links are very real; for certain Arabs, they are reality, not a mere representation of it, clear evidence of the importance and power of language attitudes.

Bibliographical references

Abu-Rabia, Salim. 2003. "Cognitive and social factors affecting Arab students learning English as a third language in Israel". *Educational Psychology* 23.347–360.

Agheyisi, Rebecca and Joshua A. Fishman. 1970. "Language attitude studies: A brief survey of methodological approaches". *Anthropological Linguistics* 12:5.137–157.

Al-Haq, Fawwaz Al-Abed. 1998. "Language attitude and the promotion of Standard Arabic and Arabization". *al-ʿArabiyya* 31.21–37.

——. 2000. "Changes in the attitudes of Jordanian students to the Hebrew language". *Journal of Sociolinguistics* 4.263–273.

Al-Kahtany, Abdallah Hady. 1997. "The 'problem' of diglossia in the Arab world: An attitudinal study of Modern Standard Arabic and the Arabic dialects". *al-ʿArabiyya* 30.1–30.

Amara, Muhammad Hasan and Bernard Spolsky. 2001. "The construction of identity in a divided Palestinian village: Sociolinguistic evidence". Herzog and Ben-Rafael (2001:273–288).

Ben-Rafael, Eliezer and Hezi Brosh. 2001. "Jews and Arabs in Israel: The cultural convergence of divergent identities". Herzog and Ben-Rafael (2001:289–310).

Bentahila, Abdelâli. 1983. *Language attitudes among Arabic-French bilinguals in* Morocco. Clevendon: Multilingual Matters.

—— and Eirlys E. Davies. 1992. "Convergence and divergence: Two cases of language shift in Morocco". *Maintenance and loss of minority languages*, ed. Willem Fase, Koen Jaspaert, and Sjaak Kroon, 197–210. Amsterdam and Philadelphia: J. Benjamins.

Bentolila, Yaakov. 2001. "Bilingualism in a Moroccan settlement in the south of Israel". Herzog and Ben-Rafael (2001:243–258).

Bradac, James J., Aaron Castelan Cargile, and Jennifer S. Hallett. 2001. "Language attitudes: Retrospect, conspect, and prospect". *The new handbook of language and social psychology*, ed. W. Peter Robinson and Howard Giles, 137–155. Chichester: J. Wiley.

Daoud, Mohammed. 2001. "The language situation in Tunisia". *Current Issues in Language Planning* 2.1–52.

Diab, Rula. 2000. *Lebanese students' beliefs about learning English and French: A study of university students in a multilingual context*. Ph.D. diss., University of Texas at Austin.

Dweik, Bader S. 1997. "Attitudes of Arab students towards *al-fuṣḥā wal-ʿāmmiyya*". *al-ʿArabiyya* 30.31–48.

Eagly, Alice H. and Shelly Chaiken. 1988. "Attitude structure and function". *The handbook of social psychology*, I, ed. Daniel T. Gilbert, Susan T. Fiske, and Gardener Lindzey, 269–322. New York: Oxford University Press.

Eisele, John. 2003. "Myth, values, and practice in the representation of Arabic". *International Journal of the Sociology of Language* 163.43–59.

El-Dash, Linda and G. Richard Tucker. 1975. "Subjective reactions to various speech styles in Egypt". *International Journal of the Sociology of Language* 6.33–54.

Ennaji, Moha. 1997. "The sociology of Berber: Change and continuity". *International Journal of the Sociology of Language* 123.23–40.

Errihani, Mohammed. 2006. "Language policy in Morocco: Problems and prospects of teaching Tamazight". *Journal of North African Studies* 11.2:143–154.

Fasold, Ralph. 1984. *The sociolinguistics of society*. Oxford: Blackwell.

Ferguson, Charles. 1959a. "Diglossia". *Word* 15.325–340.

——. 1959b. "Myths about Arabic". *Georgetown University Languages and Linguistics monograph series*, XII, ed. Richard S. Harrell, 75–82. Washington, D.C.: Georgetown University Press.

Gal, Susan and Judith Irvine. 2000. "Language ideology and linguistic differentiation". *Regimes of language: Ideologies, polities, and identities*, ed. Paul Kroskrity, 35–83. Santa Fe, N. Mex.: School of American Research Press.

Haeri, Niloofar. 1996. *The sociolinguistic market of Cairo: Gender, class, and education*. London: Kegan Paul International.

——. 2003. *Sacred language, ordinary people: Dilemmas of culture and politics in Egypt*. New York: Palgrave Macmillan.

Hassanpour, Amir. 1992. *Nationalism and language in Kurdistan, 1918–1985*. San Francisco, Calif.: Mellen Research University Press.

Herbolich, James B. 1979. "Attitudes of Egyptians toward various Arabic vernaculars". *Lingua* 47.301–321.

Herzog, Hanna and Eliezer Ben-Rafael. 2001. *Language and communication in Israel*. New Brunswick, N.J.: Transaction Publishers.

Hussein, Riyad F. 1999. "Code-alteration among Arab college students". *World Englishes* 18.281–289.

Hussein, Riyad F. and Nasser El-Ali. 1989. "Subjective reactions of rural university students toward different varieties of Arabic". *al-ʿArabiyya* 22.37–54.

Ibrahim, Muhammad. 1986. "Standard and prestige language: A problem in Arabic sociolinguistics". *Anthropological Linguistics* 28.115–126.

Ibrahim, Zeinab. 2000. "Myths about Arabic revisited". *al-ʿArabiyya* 33.13–28.

Irvine, Judith. 1989. "When talk isn't cheap: Language and political economy". *American Ethnologist* 16.248–267.

Kaye, Alan S. 1972. "Remarks on diglossia: Well-defined vs. ill-defined". *Linguistics* 81.30–48.

Kraemer, Roberta and Menucha Birenbaum. 1993. "Language attitudes and social group membership". *International Journal of Intercultural Relations* 17.437–449.

—— and Elite Olshtain. 1989. "Perceived ethnolinguistic vitality and language attitudes: The Israeli setting." *Journal of Multilingual and Multicultural Development* 10:3.197–212.

Krauss, Robert M. and Chi-Yue Chiu. 1998. "Language and social behavior". *The handbook of social psychology*, II, ed. Daniel T. Gilbert, Susan T. Fiske, and Gardener Lindzey, 41–87. New York: Oxford University Press.

Kroskrity, Paul V. 2004. "Language ideologies". *A companion to linguistic anthropology*, ed. Alessandro Duranti, 496–517. Malden, Mass.: Blackwell.

Lambert, Wallace E. 1967. "A social psychology of language". *Journal of Social Issues* 23:2.91–109.

——, Moshe Anisfeld, and Grace Yeni-Komshian. 1965. "Evaluation reactions of Jewish and Arab adolescents to dialect and language variations". *Journal of Personality and Social Psychology* 2.84–90.

Lambert, Wallace E., R.C. Hodgson, R.C. Gardener, and S. Fillenbaum. 1960. "Evaluational reactions to spoken languages". *Journal of Abnormal and Social Psychology* 60.44–51.

Lawson, Sarah and Itesh Sachdev. 1997. "L'alternance codique, stigmatisée ou valorisée?" *Peuples Méditerranéens* 79.101–119.

——. 2000. "Codeswitching in Tunisia: Attitudinal and behavioural dimensions". *Journal of Pragmatics* 23.1343–1361.

Lippi-Green, Rosina. 1994. "Accent, standard language ideology, and discriminatory pretext in the courts". *Language in Society* 23.163–198.

Malallah, Seham. 2000. "English in an Arabic environment: Current attitudes to English among Kuwait university students". *International Journal of Bilingual Education and Bilingualism* 3.19–43.

Marley, Dawn. 2004. "Language attitudes in Morocco following recent changes in language policy". *Language Policy* 3.25–46.

Milroy, Lesley and James Milroy. 1985. *Authority in language: Investigating language prescription and standardisation*. London: Routledge.

Mouhssine, Ouafae. 1995. "Ambivalence du discours sur l'arabisation". *International Journal of the Sociology of Language* 112.45–61.

Muhawi, Ibrahim. 1994. "The metalinguistic joke: Sociolinguistic dimensions of an Arabic folk genre". *Arabic sociolinguistics: Issues and perspectives*, ed. Yasir Suleiman, 155–176. Richmond, Surrey: Curzon.

Myers-Scotton, Carol. 1993. *Social motivations for codeswitching: Evidence from Africa*. Oxford: Oxford University Press.

Nader, Laura. 1962. "A note on attitudes and the use of language". *Anthropological Linguistics* 4:6.24–29.

Nelson, Kristina. 1985. *The art of reciting the* Qur'ān. Austin: University of Texas Press.

Ochs, Elinor. 1992. "Indexing gender". *Rethinking context: Language as an interactive phenomenon*, ed. Alessandro Duranti and Charles Goodwin, 335–358. Cambridge: Cambridge University Press.

Osgood, Charles, George Suci, and Percy Tannenbaum. 1957. *The measurement of meaning*. Urbana: University of Illinois Press.

Parkinson, Dilworth. 1991. "Searching for modern *fuṣḥā*: Real life Formal Arabic". *al-ʿArabiyya* 21.31–64.

——. 1993. "Knowing Standard Arabic: Testing Egyptians' MSA abilities". *Perspectives in Arabic linguistics*, V, ed. Mushira Eid and Clive Holes, 47–73. Amsterdam and Philadelphia: J. Benjamins.

——. 1994. "Speaking *fuṣḥā* in Cairo: The role of the ending vowels". *Arabic sociolinguistics: Issues and perspectives*, ed. Yasir Suleiman, 179–211. Richmond: Curzon.

Riguet, Maurice. 1981–1982a. "Variation dans l'opinion selon la langue en milieu tunisien, I". *Institut des Belles Lettres Arabes* 148.229–252.

——. 1981–1982b. "Variation dans l'opinion selon la langue en milieu tunisien, II". *Institut des Belles Lettres Arabes* 149.57–86.

S'hiri, Sonia. 2002. "'Speak Arabic please!': Tunisian Arabic speakers' linguistic accommodation to Middle Easterners". *Language contact and language conflict in Arabic*, ed. Aleya Rouchdy, 149–174. London: Routledge Curzon.

Sadiqi, Fatima. 1997. "The place of Berber in Morocco". *International Journal of the Sociology of Language* 123.7–22.

Sawaie, Mohammed. 1994. *Linguistic variation and speakers' attitudes: A sociolinguistic study of some Arabic dialects*. Damascus: Al-Jaffan and Al-Jabi Publishers (in cooperation with Institut Français d'Etudes Arabes de Damas).

Schulz, David E. 1981. *Diglossia and variation in formal Standard Arabic in Egypt*. Ph.D. diss., University of Wisconsin-Madison.

Scirha, Lydia. 2001. "Trilingualism in Malta: Social and educational perspectives". *International Journal of Bilingual Education and Bilingualism* 4.23–37.

Stevens, Paul B. 1983. "Ambivalence, modernization, and language attitudes: French and Arabic in Tunisia". *Journal of Multilingual and Multicultural Development* 4:2–3.101–111.

Suleiman, Yasir. 2003. *The Arabic language and national identity: A study in ideology*. Washington, D.C.: Georgetown University Press.

——. 2004. *A war of words: Language and conflict in the Middle East*. Cambridge: Cambridge University Press.

Taleb Ibrahimi, Khaoula. 1997. *Les Algériens et leur(s) langue(s): Eléments pour une approche sociolinguistique de la société algérienne*. Algiers: El Hikma.

Versteegh, Kees. 1997. *The Arabic language*. Edinburgh: Edinburgh University Press.

Walters, Keith. 1996. "Intrasentential codeswitching in diglossic settings and its implication for linguistic variation and language change". *Sociolinguistic variation: Data, theory, and analysis: Selected papers from NWAV23 at Stanford*, ed. Jennifer Arnold, Renée Blake, Brad Schwenter, and Julie Solomon, 401–416. Stanford, Calif.: Center for the Study of Language and Information.

——. 1999a. "'New Year Happy': Some sociolinguistic observations on the 'anglicization' of Tunisia". *English in North Africa*, ed. Mohamed Jabeur, Adal Manai, and Mongi Bahloul, 33–63. Tunis: TSAS and British Council.

——. 1999b. "'Opening the door of paradise a cubit': Educated Tunisian women, embodied linguistic practice, and theories of language and gender". *Reinventing identities: The gendered self in discourse*, ed. Mary Bucholtz, A.C. Liang, and Laurel A. Sutton, 200–217. New York: Oxford University Press.

——. 2003. "Fergie's prescience: The changing nature of diglossia in Tunisia". *International Journal of the Sociology of Language* 163.77–109.

KEITH WALTERS (Portland State University)

Language Contact

The study of contact situations involving Arabic permits a number of generalizations, most of which are unsurprising. More interesting, therefore, are the sometimes quite striking differences among the contact situations and their linguistic and sociolinguistic outcomes. In this entry, three sets of topics are surveyed, and some of their implications for general theories of contact-induced language change are discussed. Less attention is paid to the historical, political, and socioeconomic settings of the various contact situations, not because they are unimportant or uninteresting but because they vary so much: the one thing almost all of them have in common is the crucial role of Islam in the spread of Arabic throughout and beyond the Arabian Peninsula.

1. Topics and territories

By far the most prominent topic in virtually all studies on language contact involving Arabic is the adoption and adaptation of loanwords (see Sec. 2), primarily from Arabic into other lan-guages but also, in a small number of cases, from other languages into Arabic. The prominence of this topic comes as no surprise: with the exception of isolated minority communities of Arabic speakers in → Latin America, → North America, and → Europe, structural interference is only likely to have occurred within Arabophone areas; and most of the contacts with other languages – sub-Saharan African contacts, South and Southeast Asian contacts, and even European contacts – took place outside Arab lands. The most important subtopics within this area concern the semantic domains of loanwords, the particular word classes that are borrowed, and the phonological and morphological nativization, or lack thereof, of the loanwords. The question of structural interference is explicitly raised much less often, but is especially interesting when it does arise (see Sec. 3). Next comes a discussion of multilingualism and its manifestations, including language death and pidgins and creoles (Sec. 4). The final main section concerns language planning in a broad sense, with special emphases on the goals of teaching Arabic, the choice(s) of writing system(s), and language purism movements (Sec. 5). The sixth and final section is a brief conclusion.

The regions and languages covered in this entry can be roughly divided into five groups according to the nature of their contacts with Arabs and/or Arabic. First, the most intimate contacts are (or, in some historical cases, were) with languages spoken within or near Arabic-speaking territory, namely, the Near and Middle East and modern Turkey: → Modern South Arabian languages, → Aramaic/Syriac, Modern Hebrew or → Ivrit; → Coptic; → Berber languages; → Persian, and nearby → Tajik and → Kurdish (closely related to Persian); and → Turkish and other Turkic languages, such as → Tatar, → Kazakh, and → Uzbek. In these cases – with the possible exception of Coptic, because details of its contacts with Arabic are not well understood – influence between Arabic and the other language(s) has often been mutual and has involved a significant degree of bilingualism. Influence from other languages on Arabic all fall in this group, except for the influence of European languages like English, French, or Italian, which is a relatively recent phenomenon (→ English loanwords; → French loanwords; → Italian loanwords).

In the next three groups, influence has been entirely, or almost entirely, from Arabic into the other language(s), although of course this generalization is not meant to rule out the possibility of other-language influence on local varieties of spoken Arabic, where there were any (but usually there were not). The second group comprises languages of sub-Saharan Africa, primarily spoken in modern countries lying on or near the Arabs' ancient trade routes: → Nubian; → Somali; Tigrinya, Afar, and other languages of the Horn of Africa (→ Ethiopia); → Swahili and other East African languages; → Hausa, → Kanuri, → Songhay, → Wolof, → Yoruba, → Fulfulde, and → Bambara. In the third group are languages of Europe, where Arabic once had a dominant presence both in southern Spain (→ al-Andalus) and in → Sicily and (to a lesser extent) southern Italy, and from those regions influenced other European languages as well, including → Italian, the → Ibero-Romance languages, and → English. Fourth, and more distantly, Arabic has long been in contact (often indirect contact, via Persian) with languages of South and Southeast Asia, in the Indian subcontinent (in what are now Pakistan, India, and Bangladesh; → Indonesian/Malay) and in Thailand (→ Thai) and Indonesia (→ Urdu/Hindustani; → Bengali), especially the latter, with its largely Muslim population. In addition to its spread with Islam, Arabic reached the Indian subcontinent via trade, including trade settlements in Indian Ocean coastal areas. Finally, the fifth group is the case of the Arab 'diaspora', in the first place to → Latin America, where communities of 19th-century Arabic-speaking immigrants came into intimate contact with Spanish and Brazilian Portuguese; here, the pattern was Spanish or Brazilian Portuguese influence on Arabic, rather than vice versa. At a later stage, such communities arose in → North America; in → Europe, Arabic was brought by the large labor migration movement from the 1960s onward.

A methodological issue that arises frequently should be mentioned here. Although there exist deep and wide-ranging studies of the various phenomena for particular locations (for instance Arabic loanwords into Turkish), for other contact situations (for instance on Turkish loanwords into Arabic and Arabic loanwords into Nubian) no systematic studies have as yet been carried out, so that only preliminary analyses are possible. Among other things, this problem affects discussions of the numbers and variable phonological nativization of loanwords, and it also hinders attempts to discover structural interference in those (relatively few) cases where one might expect to find some. But the lack of definitive answers in most cases should not be seen as any kind of defect in existing research; rather, it is an indication that there is still much room for exciting new research on language contacts involving Arabic.

For the most part, sources on particular contact situations may be found in other entries, which are cross-referenced in the present entry (for a general survey somewhat similar to the one presented here, but with more limited access to data, see Versteegh 2001). For general literature on contact linguistics, see Thomason and Kaufman (1988) and Thomason (2001).

2. LOANWORDS

In all studies of contact between Arabic and other languages, loanwords are a central issue, often to the exclusion of other topics. Studies on loanwords from other languages into Arabic range from → English and → Italian loanwords in modern Arabic; → Turkish loanwords; → Ethiopian and → South Arabian loanwords in Classical Arabic; and → Ivrit (Modern Hebrew) loanwords in (specifically) Palestinian Arabic. One of the most interesting, but admittedly controversial, topics is that of the estimated 322 loanwords in the *Qur'ān* (Jeffery 1938), for instance from → South Arabian and from → Aramaic/Syriac. The near-universal focus on loanwords fits the general picture of Arabic in contact with other languages, especially but not only those spoken relatively far from Arabophone regions. Without the presence of communities of native speakers of Arabic, it is likely that linguistic interference will be largely or entirely limited to non-basic lexical items. The reason is that intimate contact, including significant degrees of bilingualism, is needed to provide an appropriate social setting for contact-induced language change that affects the receiving language more deeply than the adoption of loanwords (cf. Thomason and Kaufman 1988:47ff.). Accordingly, with the possible exception of highly formal Arabic-related contexts (such as religion and poetry), deeper influence is found only in Arabophone regions (see Sec. 3 for further discussion of this point).

Numerical estimates of Arabic loanwords in other languages are rare, but one figure that is given is startling: loanwords from Arabic and Persian once formed more than 80 percent of the vocabulary of written Ottoman Turkish; nonetheless, Ottoman Turkish is claimed never to have been a mixed language because the grammar and 'verbal core' remained Turkish (see → Turkish). For Modern → Persian, a much lower estimate is given: 8,000 loanwords in a dictionary count and considerably fewer, depending on the genre, in a count of text frequency. Since Persian, like Turkish, underwent a period of attempts to purge its vocabulary of loanwords (see Sec. 5), this figure is much lower than it would have been before the 1930s.

Calques (loan translations) are frequently highlighted in the consideration of loanwords. These are a subtype of lexical borrowing, involving morpheme-by-morpheme translation of words (involving only roots and derivational affixes, not inflection) and/or word-by-word translations of phrases. The most prominent of these discussions are in contact situations in the first group listed above (henceforth Group 1) – the regions in or near Arabophone territory, i.e. the Near and Middle East and modern Turkey. There are many calques from Hebrew in both written and spoken Arabic of the region (see → Ivrit); and conversely, there are many calques from Arabic in Ivrit. In the latter case, they come both from 'natural' sources – that is, from people's daily interactions with each other in speech and writing – and from planned calquing, especially by numerous innovators, beginning with Eliezer Ben-Yehuda and continued by the Hebrew Language Council and, later, by its successor, the Academy of the Hebrew Language. In → Turkish, calques were much harder to identify as being of foreign origin and were thus left as traces after the post-Ottoman Turkish language reform attempted to eliminate all Arabisms. And among the

Turkish loanwords in the Arabic dialects spoken by minority groups within Turkey there are many calques, but it is often impossible to tell in which direction the borrowing went because there are no detailed studies of phraseology in Arabic and Turkish (see → Turkish loanwords).

The major semantic domains in which loanwords cluster are roughly predictable, with one or two surprises, from the nature of the contacts. Words connected with Islamic religious practices and beliefs are prominent in almost all regions with a significant Muslim population, notably in sub-Saharan Africa and South and Southeast Asia; for → Hausa, for instance, over 50 percent of the Arabic loanwords are estimated to concern religion. Strikingly, however, and despite very large numbers of Arabic loanwords in other domains, → Persian has not borrowed heavily from Arabic in the domain of religion. This may be accounted for by the common-sense reasoning that a successful effort to convert people to a new religion will necessarily require that they understand what the missionaries are telling them, so that it makes sense for the missionaries to use the prospective converts' own words in the process (see → Persian). But this explanation does not work for other areas to which Islam spread, as indicated by the extent to which more distant cultures have adopted Arabic religious vocabulary along with Islam. This does not necessarily mean that the explanation is invalid for Persian, but it seems more likely that the explanation lies instead, or at least in part, in the prestige of Persian at the time of the conversions, rivaling that of Arabic and far higher than the prestige of most other languages spoken in regions in which Arabic loanwords predominate in the field of religion (→ Iran). It is worth noting in this context that the absence of lexical borrowing is well attested in even quite intense contact situations elsewhere in the world, for instance in Native American languages of the Pacific Northwest of the United States and Canada, some of which have borrowed almost no words from English in spite of more than 150 years of intimate contact and extreme cultural pressure. Clearly, therefore, cultural factors can and do influence the likelihood of lexical (and other) borrowing.

Other semantic domains are largely or wholly unconnected with religion, except insofar as Arabic-language administration and culture accompanied Islam. The most frequently mentioned domains, obtaining also in Europe and other regions that did not turn to Islam, are trade, science and technology, time, literacy and grammar, administration, maritime terminology, jurisprudence, food, items connected with daily life (curiously, 'soap' seems to be the item most frequently mentioned), and flora and fauna. Many authors emphasize the very wide semantic range of Arabic loanwords, but overall these domains, especially the first few, reflect the features of Arabic civilization (outside religion) that have been most important internationally: ancient trade routes around the Mediterranean and south to sub-Saharan Africa and the Indian Ocean, and world dominance in science – a dominance still visible in the large number of international words of Arabic origin in mathematics and astronomy. Items of trade (including 'soap'!) and food, as well as other everyday terms, speak to the influence of Arabic culture in near and distant regions.

When Arabic is the recipient language in a contact situation, religion is not a source of loanwords, which is hardly surprising. Instead, everyday items predominate; recent English loanwords in the domains of computers, cars, foods, and clothes provide one obvious example (obvious because English terms in these areas have spread around the world in recent decades). An interesting example of a more specialized set of loanwords is found in Sudanese Arabic, which has mainly borrowed agricultural terms from → Nubian.

Another topic that is covered in many studies of language contact involving Arabic concerns borrowed word classes. There is quite general agreement that nouns are by far the largest class of loanwords – no surprise there, as this is true all over the world. It is somewhat surprising, however, to find that in many contact situations involving Arabic, verb borrowing does occur, given the still widespread (though mistaken) view that verbs are rarely borrowed. True, finite verbs are very seldom borrowed from Arabic into Turkish; but even this means that at least a few verbs have been borrowed. Verbs have been borrowed into Arabic from (at least) Aramaic, Ivrit, and Turkish, and from Arabic into (at least) Neo-Aramaic, Modern South Arabian languages, Turkish (rarely), Tatar, Persian, Kurdish, Hausa, Yoruba, Fulfulde, Songhay, Swahili, Nubian, Urdu, Bengali, Telugu, Indonesian, and Malay. Sometimes, there are many such loanwords: by one count, there are 75 verbs in Syrian Arabic dialects from Turkish (→ Turkish loanwords). The actual mechanism of borrowing varies. As has been noted for other contact situations around the world, verb borrowing often involves the adoption of a nominalized verb (such as an infinitive or a participle) which is then combined with an auxiliary verb native to the borrowing language; this is found in some of the Arabic contact situations as well, for instance in → Persian, → Urdu/Hindustani, and → Bengali. But the language contact literature also has examples of direct borrowing of verbs as verbs, and that, too, happens in these contact situations, e.g. in → Fulfulde and → Songhay, where the Arabic imperfect is borrowed, and in the Nubian language Nobiin. By contrast to Nobiin, the closely related Nubian language Kenzi-Dongolawi borrows verbs by means of an auxiliary *-e(e)* 'to say' (see → Nubian), in a striking instance of different borrowing strategies employed in very similar systems. The overall picture of diverse means and rates of verb borrowing in contact situations involving Arabic is of considerable significance for general investigations of lexical transfer in language contact situations, because it arises from parallel studies of the phenomenon in situations involving a single language.

Two other frequently mentioned borrowed word classes are numerals and discourse markers (both particles and full words). Numerals have been borrowed in (at least) → Modern South Arabian languages, where most numerals above 10 are of Arabic origin; → Swahili, probably the best-known case, where the numerals for 6, 7, and 9 are of Arabic origin, the numerals for 11–19 are expressed by doublets, both Arabic terms and Bantu terms, and other numerals are Arabic; and in → Nubian, → Hausa, and → Javanese.

The borrowing of discourse markers is interesting from a syntactic and pragmatic viewpoint, as such morphemes tend to have structural functions, not (just) lexical ones; no details are available in the literature on syntactic implications of these borrowings, however. In some situations apparently all word classes can be borrowed, including adjectives, conjunctions, prepositions, and various kinds of particles, e.g. in → Modern South Arabian languages, → Hausa, and → Nubian, all borrowing from Arabic, as well as in the case of → Ivrit loanwords in Arabic. Other languages have borrowed mainly nouns, including some deverbal nominals, and also some adjectives (e.g. → Persian, → Indonesian/Malay, all from Arabic, and → Italian loanwords in Arabic).

One final point on loanword domains: → Modern South Arabian is said to have borrowed from Arabic a 1st person singular suffix. This isolated instance is worth noting because the borrowing of personal pronouns is rather rare in this and most other parts of the world. Pronoun borrowing is also mentioned in Turkish from Arabic (see → Turkish), but no details are given, so its significance cannot be assessed: it is only the borrowing of personal pronouns that is believed to be especially rare. Versteegh (2001:479) has noted that the Arabic pronouns *ane* 'I' and *ente* 'you' are used in Betawi Bahasa Indonesia "in order to avoid the complicated system of prestige pronouns that exist in many Indonesian languages"; pronoun borrowing is actually rather common, for this and other purposes, in Southeast Asia (see Thomason and Everett 2005 for discussion of similar cases).

The phonological, morphological, and syntactic nativization of loanwords is the norm in casual-contact situations, where typically only non-basic vocabulary items are borrowed (Thomason and Kaufman 1988:77–78). Only where there is enough knowledge of source-language structure within the borrowing-language speech community is there the possibility of borrowing structure along with words, including the possibility that loanwords will retain some source-language features that are new to the borrowing language (Thomason and Kaufman 1988:78–95). Since the lack of nativization is in fact structural interference, it is covered primarily in Section 3 below.

For contact situations involving Arabic, the phonological and, to a lesser extent, the morphological treatment of loanwords in the receiving languages is often mentioned. The main theme is indeed nativization, as one would expect in Arabic contact situations outside Group 1. In most varieties of → Hausa, for instance, Arabic sounds foreign to Hausa are replaced by the perceived closest equivalents in the native Hausa inventory, and Arabic loanwords with closed word-final syllables undergo either deletion of the final consonant(s) or epenthesis of a vowel (for instance the suffix *-i* for masculine nouns). The same seems to be true of → Bambara, in which (for instance) the Arabic glottal stop is either deleted or replaced by /w, y, h/, and of → Fulfulde, where, for instance, Arabic /q/ is replaced by /k/ or /g/ and consonant clusters are either broken up by an epenthetic vowel or simplified by consonant deletion. In most of the languages, emphatic consonants are replaced by non-emphatic counterparts or, when these are lacking, other (fairly) similar phonemes.

Nativization also often occurs when Arabic is the receiving language. In → Italian loanwords in Arabic, for instance, Italian /p/ is replaced by Arabic /b/, and only educated persons realize it as a voiceless stop; besides, initial Italian consonant clusters are broken up by epenthesis, e.g. Cairo Arabic *kirēma* 'cream' from Italian *crema*. In → Turkish loanwords non-Arabic sounds are usually replaced by native sounds; for instance, /p/ is almost always replaced by /b/.

Although, as seen in Section 3, lexical borrowing in these contact situations is sometimes accompanied by phonological interference, at least in Group 1 contact situations, morphological interference is much rarer and is almost entirely confined to a few borrowed plural endings in situations where Arabic is the donor language. That is, morphological nativization is pervasive in these contact situations. Telugu is a typical case. Although, like Arabic, → Telugu has a noun-class system that is semantically based partly on biological gender, the two languages differ sharply in their gender categories, and Arabic words are nativized into the Telugu gender system. Moreover, nominals are borrowed as uninflected singular forms and then used with Telugu plural suffixes – in other words, complete nativization (likewise in two other Dravidian languages, → Tamil and → Malayalam). Even in → Swahili, where phonological interference from Arabic is significant, Arabic loanwords are generally nativized morphologically (but see Sec. 3 for exceptions to this generalization). In most other languages that have borrowed from Arabic, morphological nativization is complete.

One final point should be emphasized here: the frequent pattern in which an Arabic noun is borrowed with the Arabic definite article *al-* (or one of its allomorphs) attached is not evidence of Arabic morphological interference in the receiving language. The reason is that the Arabic morpheme and the following nominal are borrowed as an unanalyzable whole. In none of the languages that have borrowed from Arabic is any productive use of the Arabic article mentioned as a separate morphosyntactic element; instead, it is a mere phonological part of the noun in the borrowing language, just as an English word like *alcohol*, also originally a borrowing from Arabic consisting of an incorporated Arabic article *al-* plus a nominal stem, is a single English morpheme. It has been suggested that the Arabic article *al-*, although not productive, has morphemic status in → Indonesian/Malay (cf. Verhaar 1984) – that is, it is identifiable as a morpheme. But there seems to be no evidence that nouns occur both with and without *al-*, which would be a requisite for identifying *al-* as a morpheme.

3. STRUCTURAL INTERFERENCE

In spite of the fact that most contacts involving Arabic are not intense enough to make extensive structural interference likely, the link between Arabic and one of the world's major religions raises the possibility of relatively minor kinds of structural interference in languages spoken in Muslim areas distant from Arabophone countries. This is especially likely in the writings of highly educated people who have learned Arabic as a second (or third or further) language. These phenomena will probably be confined mainly to features that enter the language attached to loanwords and are used only with loanwords, but the rather frequent reports of borrowing of conjunctions and discourse markers means that contact-induced syntactic change is also possible (although this possibility is rarely explored; for examples of syntactic interference in written Afrikaans, see → South Africa).

All these are the kinds of features that appear in (for instance) English as a result of borrowing from Latin, which once enjoyed the high level of prestige in Europe that Arabic has in the Muslim world. In addition to the many loanwords that English has adopted from Latin, and the many technical terms coined within English from Latin morphemes, English has a small number of Latin morphosyntactic features, e.g. a handful of Latin plurals such as *alumni* [masc.], *alumnae* [fem.], and *millennia* (compare the singular forms *alumnus, alumna*, and *millennium*). Typically, in such a situation, a language will borrow different forms of the same word separately and then, later on, speakers analyze the forms morphologically. In English, the singular/plural *-us/-i* pattern in particular has become modestly productive in loanwords, as seen in the innovative (non-borrowed, non-original) plural *octopi*: here, the singular, *octopus*, is ultimately from Greek, not Latin, and was originally a Greek compound *okto-pous*, eaning 'eight feet', which originally had the Greek plural *oktopodes*.

Similar kinds of minor structural incursions from Arabic are found in regions outside Arabophone territory that have been influenced by Arabic, for instance broken plurals in Arabic loanwords. Swahili, for instance, has a few of these, although they compete with native Bantu plural formations (→ East Africa); an example is *binti* 'daughter', variously pluralized as *mabinti* (with a Bantu plural class prefix) and as *banati* (with an Arabic broken plural formation). Ottoman Turkish, Tajik, and Persian all had broken plurals on some loanwords. Indeed, broken plurals were a target of the 20th-century movement to purge Turkish of foreign elements, and the broken plurals that remain in Turkish are lexicalized as singular forms (see → Turkish). In Ottoman Turkish, before the language reform, both Arabic *and Persian* loanwords were pluralized with Arabic formations (see → Turkish), a circumstance which, like the English plural *octopi*, indicates a cer-tain level of productivity of the Arabic plural patterns, even though they remained confined to loanwords. Elsewhere, too, Arabic broken plurals are borrowed intact but lexicalized as singular forms rather than as plurals (see e.g. → Tajik, → Persian, → Kurdish and → Indonesian/Malay). These lexicalized singular forms of course do not exemplify structural interference, since only the Arabic forms, and not their structure, have been adopted. It is worth noting that although borrowed nouns often have Arabic plural suffixes (and other Arabic features: see below), borrowed verbs are always nativized (Versteegh 2001:479).

Most examples of minor structural interference are reported for the phonology. For instance, some loanwords in → Wolof are pronounced with Arabic consonants not native to Wolof when the speakers have some acquaintance with Arabic – that is, these are learnèd loans. They display such Arabic features as a voiceless uvular stop /q/ and even, apparently, occasional pharyngeal consonants. Likewise, → Italian loanwords in Arabic are generally nativized, but the fact that educated people maintain the distinction between Italian /p/ and /b/ seems to reflect a similar pattern: people who know some Italian do not (always) nativize loanwords completely. And the varieties of → Hausa that are spoken in the Sudan differ from Hausa varieties spoken farther west: in the Sudan, where Hausa speakers have more exposure to Arabic, the speakers sometimes keep emphatic consonants in Arabic loanwords, even to the point of inserting them via hypercorrection where they were not present in the Arabic source word. Both for → Turkish and → Tatar, it is reported that Arabic loanwords often violate the two Turkic languages' vowel harmony rules.

The only hints of more extensive structural interference are found in contact situations in Group 1, namely regions in and near Arabophone territory – and also in Swahili, which is a special case. Versteegh (2001:495) reports, for instance, that in Ottoman Turkish one finds "much more productive" Arabic (and Persian) morphosyntax, including not only the broken plurals mentioned above but also such features as Arabic agreement rules in noun phrases. Citing Prokosch (1980:40), he also notes that such rules were used variably, not consistently, which might reflect a distinction between educated and less educated usage and/or between formal and less formal registers. And there is evidence of the use of some Ivrit loanwords in Palestinian Arabic with both Ivrit and Arabic suffixes (see → Ivrit loanwords); the choice of suffix often seems to reflect the speaker's level of education, such that highly educated people tend to use Ivrit suffixes in most contexts and words. There are some indications that the dialects of speakers of Arabic in Turkey are affected by interference from Turkish, not only in the lexicon but also to some degree in morphology and syntax (see → Turkish loanwords and → Uzbekistan Arabic), but the lack of details about this situation suggests that this is a topic that has not yet been systematically explored.

In the phonology, chronological layers of borrowing may be identified in some Jewish →

Neo-Aramaic dialects of northern Iraq according to the degree to which Arabic loanwords are nativized: in older borrowings, Arabic phonemes not native to Aramaic have been replaced by native Aramaic phonemes, but in more recent loans, certain Arabic phonemes are retained. Similarly, thanks to Nubian speakers' increasing proficiency in Arabic, Arabic loanwords in Nubian sometimes preserve non-native Arabic segments and structures, which leads to changes in the original Nubian phonological system. So, for instance, although emphatics are replaced by non-emphatic consonants, and Arabic consonant clusters are broken up by epenthetic vowels, a voiced alveolar fricative /z/, new to Nubian, is now found in loanwords only, and /l/ and /r/, previously barred from initial position in Nubian, now occur there in loanwords (see → Nubian). This picture resembles changing contact situations elsewhere in which an early period of casual contact, with little bilingualism among borrowing-language speakers, saw full phonological nativization of loanwords, while later on, when bilingualism had become widespread among borrowing-language speakers, loanwords were not nativized. A clear example is found in Siberian Yupik (Eskimo), where early Russian loanwords have only native Yupik sounds, but later Russian loanwords preserve (previously) foreign sounds and have thus changed the phonemic inventory of Siberian Yupik (Menovščikov 1969:124–130).

Swahili, with its huge number of Arabic loanwords and its establishment as the major lingua franca of East Africa by Arab traders, lies outside Group 1 regions but nevertheless seems to have undergone more extensive influence from Arabic than have other sub-Saharan African languages within the Arabic sphere of influence. In addition to the broken Arabic plurals on loanwords that were mentioned above, Swahili also has several new phonemes, confined to Arabic loanwords. These are used variably, but especially by those who have acquired some knowledge of Arabic and who feel that the use of Arabic phonemes adds to the prestige of their speech (see → Swahili). Non-Muslim Swahili speakers who have less (or zero) knowledge of Arabic nativize loanwords fully. Register matters: Arabic pronunciation may be heard in 'highly formal' speech but not in casual speech. The recurrent view that Swahili may have originated as a pidgin or creole, with Arabic as a major component, is rejected by modern scholars (see Nurse 1997). But there is little doubt that the flood of Arabic loanwords has had structural effects, and not only in semantic domains like the numeral system (where most native Bantu numerals have been replaced by Arabic numerals): Swahili is one of very few Bantu languages that have lost phonemic tone distinctions entirely, and the trigger for this development was surely the impact of all those tone-less Arabic loanwords.

Overall, then, the amount of structural interference reported in language-contact situations involving Arabic is nontrivial but also not deep. It is likely that further research, especially on Arabic contacts in Group 1 regions, will reveal considerably more structural interference, not only from Arabic into other languages, but also from other languages into Arabic-speaking minority groups.

4. Multilingualism and its effects

Not surprisingly, bilingualism/multilingualism in Arabic and other language(s) is concentrated in Group 1 regions that have a major Arab presence. Almost all speakers of → Modern South Arabian languages (except in Soqotra) speak their own language, one or two other Modern South Arabian languages, and also Arabic; and Hausa-speaking communities in Sudan, Libya, and Saudi Arabia tend to be bilingual in Arabic and → Hausa. The picture is often less clear for past eras – the question of when, and even whether, bilingualism obtained in → al-Andalus is highly controversial, for instance, but past bilingualism can be confidently discussed in, for instance, Aramaic-speaking regions (→ Aramaic/Syriac), → Sicily, Persia (→ Persian), and certain Indian towns whose residents were bilingual in Sindhi and Arabic (→ India; cf. Yusuf 1967:56). The ups and downs of bilingualism in Arabic and Portuguese (Brazil) and in Arabic and Spanish (especially Argentina) is particularly interesting, covering such factors as the speed of assimilation of Arabic-speaking immigrants and the wider community's negative or positive reactions to Islam (→ Latin America). The situation in Latin America also highlights, by contrast, the extent to which elucidation of the political and social conditions of past contact situations is hampered by incomplete information: even with all the documentation available on the former use of Arabic in Persia, Turkey, Andalusia, Sicily, and elsewhere, it remains impossible to carry out the kinds of fine-grained analyses that one can conduct on a live contact situation, and generally impossible to speak with confidence about the linguistic repertoires of illiterate folk, as opposed to the literate, educated, higher strata of society.

In some cases, there are reports of a recent decline in the use of Arabic in public life in regions outside Group 1, for instance on the Tanzanian mainland, where one no longer sees shop signs and other signage in Arabic in the towns (→ East Africa). But the instances of contact situations within Group 1 regions in which Arabic is replacing other languages – that is, in which language death is occurring – are more striking. Some → Modern South Arabian languages are being replaced by Arabic, and several → Nubian languages have already vanished as a result of Arabization.

A parallel topic, one that concerns the most dramatic things that happened to Arabic as it spread rather than what happened to the languages with which it came into contact, has to do with Arabic-lexifier pidgins and → creoles. The topic has been addressed by Owens in a number of articles (e.g. in two general survey articles, 1997, 2001), and it seems to provide at least some of the background for the situation in the → Horn of Africa where Arabic functions as a → lingua franca; it is also a topic that has received a significant amount of attention within Arabic studies over the past 20 years. The first major work to focus on pidgins and creoles in this domain was Versteegh (1984), an investigation of the possibility of → pidginization as a phenomenon accompanying the spread of Islam to what are now Arabophone territories; two years later, Prokosch (1986) surveyed Arabic-based pidgins and creoles in Africa, where almost all the ones reported in the literature are located. Proposals about Arabic-lexifier pidgins spoken early in the history of the spread of Arabic and Islam have been based on fragmentary documentation from as early as the 11th century C.E. (Thomason and Elgibali 1986, reporting on a passage in al-Bakrī) and on inferences about the kinds of contact situations that likely arose as Arabic spread (e.g. Versteegh 1984). This topic is not addressed in detail here because for the most part, though obviously related to the general subject of Arabic contacts with other languages, it does not lend itself to direct comparisons with other kinds of contact situations – namely those whose social contexts and linguistic results were less drastic.

5. Language planning

A final recurring set of topics in the literature on language contact involving Arabic falls under the general heading of language planning. Many authors discuss the teaching of Arabic in various regions, the choice(s) of a writing system for other languages in regions influenced by Arabic, and language reform movements designed to purge various languages of Arabic loanwords. These are all huge topics, of course, and space limitations make it impossible to do justice to them, especially as policies and practices have changed over time in a number of regions, sometimes more than once. This section, therefore, is merely a sketch of the issues that arise in this area.

The teaching of Arabic outside Arabophone regions has often, as one might expect, aimed at a reasonable level of competence in written or spoken Arabic, or both. But sometimes the goal is instead to teach the rudiments of the alphabet and the recitation of prayers, nothing more; comprehension of the Arabic prayers is neither required nor expected. There are various religious and political motivations for the latter practice, and it is tolerably widespread in Muslim regions; it is mentioned especially in South Asia (→ India; → Pakistan; and → Bangladesh). In → Pakistan, familiarity with the Arabic script has remained part of the traditional education, and even those who do not understand the text are taught to read the letters. Various other political and social motives have driven Arabic teaching in these and other regions, and there exists rich material for the study of the issues.

A related topic is the choice of one or more writing systems for languages in contact with, and under the influence of, Arabic and (usually) Islam (→ Arabic alphabet for other languages). This matter is mentioned for many contact situations involving Arabic, an indication of its central importance in language planning: it is a topic with profound political implications. Modern Persian, for instance, is written with a version of the Arabic alphabet, but Middle Persian was not – it was written in an Aramaic script (→ Iran). The change to the Arabic alphabet came with Islam by the mid-9th century C.E. Ottoman Turkish was written with Arabic letters, but Modern Turkish is written in the Latin alphabet; the change was made as part of the language reform undertaken by a secular government looking to the West for its future. The history of → Tatar writing has been unusually complex: in the mid-19th century, Tatar was written in the Arabic alphabet; in 1927, the Arabic alphabet was abandoned in favor of Latin letters; and in 1938, in accordance with Soviet policies regarding minority languages, a Cyrillic alphabet replaced the Latin one.

In → Malta, where the spoken Arabic variety lost its cultural connection to Classical Arabic very early, Maltese Arabic was never written in the Arabic alphabet; even the earliest texts, starting in ca. 1470 C.E., were written in Latin letters. In sub-Saharan Africa, the first writing systems were often Arabic (→ Swahili; → Wolof; → Fulfulde). Under European colonial rule, literate Swahili speakers generally knew only the Arabic alphabet, but the colonial government in Zanzibar persisted in publishing Swahili in the Latin alphabet (→ East Africa), a practice that guaranteed that almost no Swahili speakers would be able to read it. In South and Southeast Asia, Urdu, an official language of Pakistan that is also widely spoken in India, is written in Arabic script, while Hindi, an official language of India, is written in a Sanskrit-derived script; and yet Urdu and Hindi are so closely related as to be arguably dialects of the same language (→ Urdu/Hindustani). Farther east, a comparable split is found in Indonesia: until the end of the 19th century, Malay was (and sometimes still is) written in a version of the Arabic alphabet, while certain other Indonesian languages – notably → Javanese – still use a Sanskrit-derived writing system (→ Indonesia; → Indonesian/Malay).

This brief survey of alphabet choices necessarily omits a great many useful and important details in the discussions on the various language contact situations. It does, however, provide an indication of the spread of Arabic writing in addition to, and sometimes partly independently of, the spread of the Arabic language itself. A very recent development in Morocco underscores the cultural and political dimensions of such a choice. With the newly established policy of introducing Berber into some schools came the necessity for standardizing → Berber (specifically Tamazight Berber; → language shift: Amazigh), including establishing an official writing system. The planners rejected both the Latin alphabet, formerly a symbol of the French colonial government, and the Arabic alphabet, the writing system of the country's official language, Arabic; instead, they chose to revive the ancient Tifinagh writing system because it is unique to Berber (Hamid Ouali, p. c., 2005).

The move to establish Berber in Moroccan schools as a language independent of Arabic (and French) is just one of many instances in which speech communities have tried to free themselves of the cultural weight of Arabic (although in most cases they continue to revere Arabic as the language of the *Qur'ān*). Vigorous debates and policies on this issue took place in Persia/Iran and Turkey. The use of Persian itself in scholarly writing was controversial in Persia a thousand years ago; that changed in the 13th century, when Arabic ceased to be spoken in Persia (→ Iran), but it was not until the 1930s and 1940s, in Iran, that language reformers attempted to replace Arabic loanwords with native Persian words (see → Persian). They were less successful (or less radically inclined) than the language reformers in Turkey. There, during the transition from the Ottoman Empire to the modern Turkish state in the 1920s, the politically dominant radical purists wanted to eliminate all foreign elements, especially those from Arabic and Persian, from Turkish. As a result of their efforts, the percentage of Arabic (and Persian) loanwords in Standard Turkish was sharply reduced (see → Turkish). Similarly, → Tatar lost most of its Arabic and Persian loanwords as a result of Soviet policies, which led to a wholesale replacement of those words, not by native words as in Iran and Turkey, but by Russian words. Nowadays, half the words in a Standard Tatar-Russian bilingual dictionary are of Russian origin, although since perestroika some Arabic words have again been appearing in written Tatar. In old Bengal, an influential 18th-century grammarian, Halhed, regarded foreign elements as a 'pollution' of 'pure' → Bengali; but in the 19th century there was an influx of Perso-Arabic borrowings into Bengali in reaction to the British/Hindu Sanskritization of the language. Differences of opinion about the cultural mean-ing of loanwords remain, however.

A very different form of linguistic purism was inflicted on, and resisted by, the Maltese speech community (→ Malta). The British, while trying to replace Italian with English on the island, also promoted Maltese; but their efforts along these lines involved trying to make Maltese a 'purer' language by bringing it closer to Classical Arabic, an approach that met with strong negative reactions among the community's educated elite.

6. CONCLUSION

The study of Arabic in contact presents a picture of the rich variety of linguistic, social, and political settings. The picture is not complete, of course: most of the issues discussed above still need further systematic investigation, including both case studies of individual contact situations and comparative studies of partly similar contact situations. Comparative information is already available in the literature on some topics, for instance strategies of verb borrowing, patterns of phonological nativization of loanwords, and motivations for choosing writing systems. Other topics are barely touched on in the literature, perhaps most notably the issue of morphosyntactic structural interference; if its absence in the literature reflects the state of research on the topic, this is clearly a growth area for Arabic scholarship.

BIBLIOGRAPHICAL REFERENCES

Jeffery, Arthur. 1938. *The foreign vocabulary of the Qurʾān*. Baroda: Oriental Institute.

Menovščikov, G.A. 1969. "O nekotoryx social'nyx aspektax evoljucii jazyka" [On some social aspects of the evolution of language]. *Voprosy social'noj lingvistiki* [Problems of sociolinguistics], 110–134. Leningrad: Nauka.

Nurse, Derek. 1997. "Prior pidginization and creolization in Swahili?". Thomason (1997:271–294).

Owens, Jonathan. 1997. "Arabic-based pidgins and creoles". Thomason (1997:125–172).

——. 2001. "Creole Arabic: The orphan of all orphans". *Anthropological Linguistics* 43.348–378.

Prokosch, Erich. 1980. *Studien zur Grammatik des Osmanisch-Türkischen unter besonderer Berücksichtigung des Vulgärosmanisch-Türkischen*. Freiburg: K. Schwarz.

——. 1986. *Arabische Kontaktsprachen (Pidgin- und Kreolsprachen) in Afrika*. Graz: Institut für Sprachwissenschaft, Universität Graz.

Thomason, Sarah G. (ed.). 1997. *Contact languages: A wider perspective*. Amsterdam and Philadelphia: J. Benjamins.

——. 2001. *Language contact: An introduction*. Edinburgh and Washington, D.C.: Edinburgh University Press and Georgetown University Press.

—— and Alaa Elgibali. 1986. "Before the lingua franca: Pidginized Arabic in the eleventh century A.D.". *Lingua* 68.317–349.

Thomason, Sarah G. and Terrence Kaufman. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.

Thomason, Sarah G. and Daniel Everett. 2005. "Pronoun borrowing". *Berkeley Linguistics Society* 27.301–315.

Verhaar, John W.M. 1984. "Affixation in contemporary Indonesian". *Towards a description of contemporary Indonesian: Preliminary studies*, I, ed. Bambang Kaswanti Purwo, 1–26. Jakarta: Universitas Atma Jaya.

Versteegh, Kees. 1984. *Pidginization and creolization: The case of Arabic*. Amsterdam: J. Benjamins.

——. 2001. "Linguistic contacts between Arabic and other languages". *Arabica* 48.470–508.

Yusuf, S.M. 1967. "Arabic language and literature in the Indo-Pakistan subcontinent". *Iqbal* 16:1.56–66.

SARAH G. THOMASON (University of Michigan)

Language Impairment

The term 'language impairment' covers a wide array of language disorders that affect language abilities, including hearing, speaking, signing, reading, and writing across all levels of language structures and functions as well as processes of language comprehension and production. The field dedicated to the study of language impairments is called → language pathology.

1. CLASSIFICATION AND CAUSES

Language impairments can generally be classified into two main categories: acquired and developmental disorders. Acquired language impairments result from a variety of causes affecting the functioning of the brain and the nervous system, including disease (e.g. degenerative diseases, such as Alzheimer's and Parkinson's, and brain tumor), stroke or brain injury, and mental retardation due to genetic disorders such as Down Syndrome, Williams Syndrome, and Fragile X syndrome. The most widely and historically studied phenomenon of acquired language disorders is that related to stroke or brain injuries resulting in a language impairment, whether total or partial, called aphasia or dysphasia (→ language pathology).

Developmental language impairments are usually classified into two subcategories: (1) impairments associated with congenital mental retardation or brain damage (caused during pregnancy or at birth), such as cerebral palsy, autism (although the exact causes of autism are still not known), and verbal dyspraxia (also called developmental verbal dyspraxia); congenital physical abnormalities (caused before, during, or after birth) having to do with hearing (loss), abnormalities of the tongue, palate, lips, jaw, teeth, etc., or with early childhood illnesses such as otitis media 'middle ear infection' that result in temporary or permanent hearing loss; and (2) impairments that have no apparent causes and are detected during childhood. Developmental impairments of the latter type are also known as developmental aphasia, developmental dysphasia, developmental language disorder, developmental language delay, etc., but the term that is increasingly gaining currency and more widely used is 'specific language impairment' (SLI). Other language impairments that fall under neither of these categories include those associated with non-brain diseases, drug and alcohol abuse, vocal abuse, and swelling of the vocal cords, and result in many different voice disorders, the severest of which is aphonia or 'voice loss'. Other impairments (such as stuttering or abnormal dysfluency) are assumed to be caused not by a single factor but by a combination of factors: neurological, environmental, and developmental. This entry focuses on developmental language impairments, particularly non-brain acquired disorders (for acquired language impairments, see → language pathology).

2. DEFICITS OF SOME DEVELOPMENTAL LANGUAGE IMPAIRMENTS

Deficits of developmental language impairments that are caused by (congenital) physical disabilities are usually specific to the abnormality associated with the disorder and are usually accompanied with compensatory articulations or strategies that individuals develop to overcome or minimize the impairment deficit. What follows are deficit descriptions of most widely known and studied developmental language impairments that have been observed among speakers of Arabic.

2.1 *Hearing loss and speech of the hearing impaired*

Hearing loss directly impacts auditory and receptive processing abilities (of the underlying phonological, morphological, syntactic, semantic, pragmatic systems, etc. of language rules), which in turn impairs articulatory and expressive abilities. Children with hearing loss have no way of assessing the linguistic input or accessing a proper representation of the speech code nor of monitoring their own speech. Language impairment varies according to the degree of hearing loss and onset, whether pre- or post-lingually; hence the ability to develop compensatory skills varies accordingly (including use of speech, lip-reading, and signing; → sign languages). For example, a mild (sensory-neutral) hearing loss mainly affects the high frequencies, resulting in the reduced ability to hear consonantal information – consonants being higher pitched than vowels – important for comprehension of auditory information. To a lesser extent, high vowels may also be affected more so than low vowels, the former being relatively higher pitched than the latter. Speech of the hearing impaired is usually characterized by sound/syllable/word omission and/or distortion, high-pitched segments, and uncoordinated speech breathing patterns, resulting in an abnormal tone of voice. Additional characteristics include the use of simple sentence constructions, inflexible or stereotyped language style, overuse of content words (such as nouns and verbs), and underuse of function words (auxiliaries, conjunctions, and prepositions). Similarly, temporary hearing loss caused by an illness such as a history of otitis media usually results in similar impediments to language acquisition, since fluctuation in auditory output may impede the process of analyzing, categorizing, and organizing auditory information (see Gold 1980; Bamford and Saunders 1991).

2.2 *Cleft palate*

Cleft palate is a congenital physical defect in which the closure of the velo-pharyngeal port is incomplete. It mainly affects the production of vowels and oral obstruent phonemes (requiring intra-oral pressure) such as stops, fricatives, and affricates. Thus, vowels are nasalized and oral obstruents are produced with weak pressure. Additionally, nasal emission, weak or strong expiratory air, glottal replacement, glottal reinforcement, and backing are characteristic compensatory articulations of cleft palate impairment. Crosslinguistic data suggest that cleft palate deficits are universal and may persist even after repair, depending on many factors such as type of cleft and time of repair (e.g. Bronsted a.o. 1994; Bzoch 1997; Bernhardt and Stemberger 1998).

2.3 *Specific language impairment and phonological disorders*

Unlike the above conditions, specific language impairment (SLI) is not specific to a particular physical abnormality. Children with SLI exhibit no physical disability, mental retardation, or neurological damage. They display normal hearing and age-appropriate scores on nonverbal intelligence tests. The number of children with SLI seems to be significant. According to a U.S. survey, about 5–10 percent of children exhibit language delay, some of whom are able to catch up later, but some seem to continue to show persistent problems (Bishop and Leonard 2000:ix). SLI is generally characterized in terms of significant deficit or delay in language development. Thus, children whose language development lags behind other children in the same age group are considered to have SLI. Based on crosslinguistic evidence, SLI deficits include late emergence of first words, late emergence of word combinations, a protracted period of lexical development, problems with inflectional forms (e.g. grammatical morphemes for agreement and tense as in *he plays football* > *he play football*), and more difficulty with language production than with language comprehension (Leonard 2000). Additionally, children with SLI acquiring a language with rich morphology (what is referred to as the Sparse Morphology Hypothesis) are found to be generally less impaired than children with SLI acquiring a language with impoverished morphology. Errors in the production output of the former are usually characterized with substitutions but not omission. Function words seem to be problematic for children with SLI regardless of the nature of the morphological system of their language, due to the presence of additional production factors such as prosodic features (Leonard 2000). Finally, to be identified as exhibiting SLI, a child with phonological disorders needs also to perform poorly on other measures of language (e.g. exhibiting deficits in comprehension, production, and grammar); a child with only a phonological deficit is considered to exhibit simply ‘a language disorder’, not SLI (Leonard 1998:13–14). Children who are developing normally resort to naturally occurring phonological and phonetic processes as a means of simplifying or approximating mature adult speech that they cannot yet produce fully. Phonological disorders are identified in children when those processes persist longer than usual (see Sec. 5.3 for examples). Thus, usually this type of disorder is referred to in terms of delay (e.g. Ingram 1990).

3. MEDIEVAL ARABIC VIEWS ON LANGUAGE IMPAIRMENTS

The earliest recorded views of language impairments or speech defects are made, although in passing, by Sībawayhi (765–796 C.E.) in his *Kitāb* (II, 415), considered, to the present day, to be the most authoritative extant work on Arabic grammar. In explaining the close proximity of the *yāʾ* /y/ with respect to both the *rāʾ* /r/ and the *lām* /l/, Sībawayhi observes that *al-ʾalṯaġ* ‘one who lisps’ the *rāʾ* or the *lām* substitutes for this speech defect by producing the *yāʾ* [y], since it is the closest of all other sounds to the *rāʾ* and the *lām* with respect to their points of articulation. Al-Jāḥiẓ (780–869 C.E.), the well-known medieval Arab encyclopedic thinker, linguist, orator, rhetorician, and writer, takes more interest in the phenomenon. He identifies many different speakers with different types of speech defects in the introductory chapters of his *Kitāb al-bayān wa-t-tabyīn*, including *al-lajlāj* ‘the stutterer’; *at-tamtām* ‘the one who stutters with *tāʾ* [t]; *al-faʾfāʾ* ‘the one who stutters with the *fāʾ* [f]’; *al-ʾalṯaġ* ‘the lisper who mispronounces *qāf* [q] as *ṭāʾ* [ŧ], *sīn* [s] as *ṯāʾ* [θ], *lām* [l] as *yāʾ* [j] or *kāf* [k], and *rāʾ* [r] as *yāʾ* [j], *ġayn* [ɣ], *ḏāl* [ð] or *ẓāʾ* [ƶ]’; *ḏū l-ḥubsa* ‘the one whose speech is not heard’; *ḏū l-ḥukla* ‘the one whose speech is not clear’; *ḏū r-rutta* 'the one whose speech is too fast'; *ḏū l-lafaf* 'the one with intermingled speech'; and *al-ʾaʿlam* 'the one with a cleft lip'. He also discusses the necessity of the teeth and the nasal cavity for proper pronunciation of sounds and speech (Jāḥiẓ, *Bayān* 15–50). In addition, al-Jāḥiẓ considers such defects to be quite different from *al-xaras* 'muteness'. Although deaf individuals may or may not be able to use speech, depending on the extent of their residual hearing, the association between hearing and speaking was also observed by al-Jāḥiẓ. He states explicitly: "A mute person is language-impaired not due to defects with his tongue but rather for not having heard sounds, speech sounds or otherwise, and therefore he does not know their composition so as to emulate them" (*Ḥayawān* 404; see also Rockey and Johnstone 1979).

Ar-Rāzī/Rhazes (865–925 C.E.), the well-known medieval Arab scholar and medical doctor, seems to have been similarly interested in language impairments. He dedicates a sizable chapter in his book *al-Ḥāwī* to the subject. In particular, he points to different types of impairments attributable to congenital abnormalities of the tongue due to tongue size (*al-ʾaratt* 'the one with a too-big tongue', *al-ʾalkan* 'the one with a too-wide tongue', and *al-ʾalṯaġ* 'the one with a too-short tongue'), weakness of the muscle/nerve of the middle ear, and what is widely known today as *lingual frenum*. With respect to *lingual frenum*, ar-Rāzī states: "The more forward in the mouth the frenum is placed, the more speech becomes impeded" (*Ḥāwī* 207; see also Rockey and Johnstone 1979:233). In addition, he mentions at least three main types of stutterers: *al-faʾfāʾ* 'the one who stutters with the *fāʾ* [f]'; *al-ʾalṯaġ* 'the one who pronounces *sīn* [s] as *ṯāʾ* [θ] or *rāʾ* [r] as *ġayn* [ɣ], *lām* [l] or *yāʾ* [j]'; and *at-tamtām* 'the one who stutters with *tāʾ* [t] and *mīm* [m]' (*Ḥāwī* 208). He points out that *luṯġa* 'lisp' is caused by the short length of the frenulum, as a temporary phonological disorder that goes away with age, and briefly mentions developmental speech delay in children (*ʾabṭaʾa kalām al-ʾaṭfāl*; *Ḥāwī* 208). Ar-Rāzī also speculates on the brain as a possible source of speech impairments when accompanied with a disorder of the *ḥawāss* 'senses/sensory perception' (*Ḥāwī* 211; Rockey and Johnstone 1979:234), implying a brain-acquired language disorder (→ language pathology).

4. Data from recent studies on Arabic

Data reported in contemporary studies on developmental language impairment with respect to Arabic relate to three areas: speech of the hearing impaired, cleft palate disorders, and phonological disorders in SLI.

4.1 *Speech of the hearing impaired*

At least one study examines the speech of some hearing-impaired speakers of Egyptian Arabic. Ammar and Rifaat (1998) examined the acoustic characteristics of ten prelingually deaf females within the age range of 13–18 years. All the subjects were diagnosed as having severe to profound hearing loss. Ten normally hearing subjects within the same age range were included as a control group. The study focuses mainly on examining three widely attested features of deaf speech: high pitch, segment lengthening, and vowel reduction (in quality). The study therefore focuses on the three corresponding acoustic characteristics: fundamental frequency, duration frequencies, and formant frequencies, respectively. All participants were required to read four words containing long vowels flanked by similar consonants. The four words were *fāl* 'omen', *fīl* 'elephant', *fūl* 'beans', and *fār* 'mouse'. The words were randomized and presented to the participants in five lists. The main findings of the study reveal a statistically significant difference across the two groups with respect to fundamental frequency. In other words, the deaf participants exhibited significantly higher pitch in vowel production than did their hearing counterparts. In addition, whereas the hearing participants exhibited the widely attested phenomenon that high vowels have higher frequency in similar phonetic contexts than low vowels, the deaf participants did not.

A similar (statistically significant) difference with respect to duration frequencies was also found across the two groups. However, both groups showed (in line with a generally attested observation) that low vowels have higher duration values than high vowels. The two main results above are in line with crosslinguistic evidence (e.g. Plant and Hammarberg 1983; Oster 1990). An additional finding of Ammar and Rifaat's (1998) study is that fundamental frequency seems to be the better predictor of the two, since the results showed that while the subjects could not maintain the normal intrinsic fundamental frequency distinction between high and low vowels, they did maintain the normal intrinsic duration distinction between high and low vowels. As for the result of formant frequencies, it was found that high vowels of the deaf participants were significantly lower (in position) than those of their hearing counterparts. But no difference was found with respect to low vowels. In other words, only high vowels were found distorted in the deaf data; low vowels remained intact and kept within normal position. While this last finding is also observed in earlier studies, there is too much variability in the production of deaf vowel patterns for this observation to be universal.

4.2 *Cleft palate*

At least one study considers Arabic cleft palate data. Shahin (2002) analyzes data from three Palestinian-Arabic-speaking children: a girl, age 3¼ years, and two boys, ages 5 and 5½ years, all with complete (soft and hard) cleft palate. The data sets comprised 80 (mostly discrete) words produced mostly in picture-naming tasks during speech therapy sessions. The data show that all three subjects produced vowels as nasalized long vowels with weak expiratory air and oral obstruents as weak pressure consonants. Additionally, all three subjects exhibited devoicing of /b/ > [b̥]. Some exhibited devoicing of /d/ and /ḍ/ and backing of /t/ and /ṭ/, realizing them all as [k], or realizing /d/, /ṭ/, and /k/ as [q], e.g. *ḥaṭṭēthum* 'I put them' > *ḥakkēkhum*, *baṭāṭā* 'potato' > *baqāqā*. Some also exhibited glottal replacement of stops /b/, /t/, /d/, /k/, and /ṭ/, realizing them as [ʔ], as in *šurab* 'drink' > *šuraʾ*, *taḥt* 'under' > *ʾaḥʾ*. At least one subject exhibited nasal emission [b] > [m], e.g. *ʾabyaḍ* 'white' > *ʾamyaḍ*. Devoicing is explained as a strategy for increasing intra-oral pressure to "counteract velo-pharyngeal leak". More evidently, the exhibited backing and glottal replacement processes, compensating for the articulation of the affected phonemes, result in phoneme conflation or merger 'insensitive' of the phonemics of the language. Thus, the findings of the study support the crosslinguistic evidence and the general observation that the deficits resulting from cleft palate are universal.

4.3 *Specific language impairment and phonological disorders*

The area of specific language impairment has only recently begun to receive interest in studies investigating developmental language impairments in Arabic-speaking children. Ammar (1992) studies delayed speech in Arabic-speaking children but focuses exclusively on phonological deficits, hence, the conclusions reached are not sufficient to relate to SLI findings. The study examines the speech of 32 Egyptian children. The subjects belonged to two groups: 16 with normal speech and 16 with delayed speech. The age range of the former was 4 years one month–4 years 10 months; the age range of the latter was 4–9 years. The phonemic inventories and phonological processes exhibited in the output production of the groups were compared. The findings reveal that out of the 27 Egyptian Arabic consonants reported in Harrell (1957), 14 phonemes were acquired at the 'mastery' range level (at 90 percent correct criterion), and the other 13 were acquired at the 'customary' range (of over 50 percent correct criterion) by the normal group. The phonologically disordered/delayed group exhibited 10 of the Egyptian phonemes at the customary level (at 50–70 percent correct), 4 phonemes at below the customary level (at 25–50 percent correct), and 13 phonemes at much lower than the customary level (less than 25 percent correct). As for the phonological processes exhibited by the participants, 15 such processes were identified in both groups, including devoicing, deemphatization, sibilant fronting, [r] deviation, velar fronting, cluster simplification, syllabic simplification, diminutization, reduplication, metathesis, stopping, backing, consonant deletion, glottal replacement, and assimilation. The normal group exhibited a much lower number of processes in their speech samples with very low percentages (less than 25 percent), while the disordered group exhibited a larger number of such processes with considerably higher percentages than their normal counterparts.

A more recent study (Salameh 2003; Salameh a.o. 2003a, 2003b) investigated both phonological and grammatical disorders, thus falling more closely than Ammar's (1992) study within the scope of specific language impairments. Salameh a.o. (2003a) focus on the phonological disorders of the impaired participants of the study while Salameh a.o. (2003b) focus on the grammatical disorders. Salameh a.o. (2003a) produce similar findings to those of Ammar (1992). Salameh a.o. report on data from two groups of bilingual Swedish- and Arabic-speaking children with and without impairment in both languages. The two groups consisted of 20 children, 10 in each group. The two groups were matched for age (with age means of 4.96 and 5.03 for the normal and the delayed group, respectively), gender, and time of exposure to Swedish preschool and parental Arabic dialect at home. The language development of the children with impairment was found significantly lagging behind other areas of development, such as nonverbal intelligence and motor and socioemotional abilities, thus fitting inclusionary criteria of SLI. Moreover, they exhibited delay in both languages. The children came from different Arabic dialect backgrounds, including Iraqi, Lebanese, Palestinian, Gulf, and Syrian Arabic. Data from the two groups in the two languages were compared on the basis of a number of phonological simplification processes, similar to those examined in Ammar's (1992) study, but with more refined definitions of such terms as 'syntagmatic' or context-sensitive processes and 'paradigmatic' or context-free processes. The most frequent processes were found to be the same for both groups and both languages. Syntagmatic processes included consonant deletion (e.g. *ḏayl* 'tail' > *sei*), metathesis (e.g. *ṯalj* 'snow' > [θalʒd]), and assimilation (e.g. *ḍufdaˁ* 'frog' > *dufbaˁ*). Paradigmatic processes included devoicing (e.g. *dib* 'teddy bear' > [dub̥]), and fronting (e.g. *warde* 'flower' > *walib*). The main findings revealed that the children with impairment displayed a significantly higher number of both syntagmatic and paradigmatic processes in each of the two languages than their normal counterparts. There were also significantly fewer syntagmatic processes in the children with impairment if they had been additionally exposed to Swedish input for longer than 1.6 years. Thus, it is necessary to know both age and length of exposure to assess bilingual children. The general conclusion is that both groups seem to follow roughly the same developmental paths as monolingual children in each language, based on results from Nettelbladt (1983) and Hansson and Nettelbladt (2002) for Swedish, and Ammar (1992), Dyson and Amayreh (2000), and Amayreh and Dyson (2000) for Arabic. Children with language impairment, however, exhibited a higher degree of phonological simplifications typical of younger children.

Salameh a.o. (2003b) report on the grammatical disorders described in Salameh a.o. (2003a). They used Pienemann's (1998) processability model as a base line for language development by which morphological agreement forms are stipulated to be acquired along a set implicational sequence. The earliest stage of language development, stage 1, is characterized by the emergence of single words or unanalyzed chunks, followed by a 'category' stage 2, where, for example, tense and number suffixes are acquired, followed by a 'phrasal' stage 3, where grammatical-agreement morphemes within the same phrase, such as Noun-Adjective agreement, are acquired, then an 'inter-phrasal' stage 4, where grammatical-agreement morphemes across phrases, such as Subject-Verb agreement, are acquired, and a final stage 5, where subordinate clauses emerge. The Arabic results of Salameh a.o.'s (2003b) study were based on longitudinal data that examined the productive use of the feminine plural suffix /-āt/ as in *biss-e* 'female cat' vs. *biss-āt* 'female cats', 3rd person singular and plural suffixes of the imperfect and the perfect as in *yaktub* 'he writes' vs. *yaktub-ū* 'they write', and gender and number agreement in Noun-Adjective constructions as in *kalb kabīr* 'a big male dog' vs. *biss-e kabīr-a* 'a big female cat' for gender agreement and *al-walad al-kabīr* 'the big boy' vs. *al-ʾawlād al-kubār* 'the big boys' for number.

The main findings show that the children with impairment did in fact develop their two languages during the twelve months of observation but did so at a slower pace than their counterparts without impairment. With respect to the Arabic data, only three of the ten children with impairments progressed to stage 3 (Noun-Adjective agreement), while all children without impairment progressed to stage 3. Despite a number of methodological limitations of the study, including the limited Arabic structures examined, the mixed colloquial backgrounds of the children, and the use of Pienemann's model, which has been shown to be problematic in other studies examining Arabic acquisition data (Lykke-Nielsen 1997; Alhawary 1999, 2003), the general finding of Salameh a.o. (2003b) shows that the children with impairment exhibited problems with inflectional agreement forms at the phrasal level – an observation supported by crosslinguistic observations of children with SLI.

Another study (Abdalla 2002) examined SLI in Ḥijāzī-Arabic-speaking children. The children exhibited developmental language-learning difficulties (both receptive and expressive) without displaying nonverbal deficits (including hearing, cognitive, behavioral, neurological, or psychological deficits). Speech production of three groups of children, ten in each group, was compared. One group consisted of children diagnosed with SLI, while the other two groups consisted of children with typical development used as control groups matched with those of the SLI group by Mean Length of Utterance (MLU) and age. The age range of the groups is 4.0 to 5.3, 4.0 to 5.2, and 2.0 to 3.0 years of the SLI and the two control groups, respectively. The language level of the SLI group was found comparable to approximately that of 24- to 28-month-old normally-developing children. The target structures examined are Subject-Verb agreement, tense (present and past), determiners, and prepositions. The findings showed that the SLI group exhibited significantly more difficulty in all target forms than the control groups did. In particular, the performance of the SLI group is marked by more inconsistency than the control groups in the use of tense, substitution of 1st person in 3rd person masculine and feminine contexts in Subject-Verb agreement cases, and omission of determiners and prepositions. Such results, based on grammatical-agreement mismatches in the SLI group, lend further support for the general crosslinguistic observation that grammatical-agreement deficit is a characteristic of SLI. The study also provides support for another generally observed crosslinguistic tendency, expressed in the Sparse Morphology Hypothesis (see section 2.3 above). The errors of the children with SLI acquiring Arabic, an inflected language with rich morphology, are characterized as substitution rather than omission errors. More significantly, the correctness rate of one of the forms investigated, the past tense, is found at 82 percent, as opposed to, for example, the 22 percent found in Rice and Wexler's (1996) study of children with SLI speaking English, a language with impoverished morphology. Thus, children with SLI acquiring a language with rich morphology are found to be generally less impaired than children with SLI acquiring a language with impoverished morphology.

Language impairment continues to be of interest to Arabists, past and present. Studies of language impairment in Arabic conducted within current psycholinguistic models of analysis have been on the rise. Such studies are significant in that they add to the body of crosslinguistic evidence on the subject, providing support for crosslinguistic generalizations that often have theoretical implications as well (for more discussion of theoretical implications, → language pathology).

Bibliographical references

Primary sources

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. 2 vols. Būlāq, 1899.
Jāḥiẓ, *Bayān* = ʾAbū ʿUṯmān ʿAmr ibn Baḥr al-Jāḥiẓ, *Kitāb al-bayān wa-t-tabyīn*. Ed. Darwīš al-Juwaydī. Sidon: al-Maktaba al-ʿAṣriyya, 1999.
——, *Ḥayawān* = ʾAbū ʿUṯmān ʿAmr ibn Baḥr al-Jāḥiẓ, *Kitāb al-ḥayawān*. Ed. ʿAbd as-Salām Hārūn. Cairo: al-Bābī al-Ḥalabī, 1940.
Rāzī, *Ḥāwī* = ʾAbū Bakr Muḥammad ibn Zakariyyāʾ ar-Rāzī, *Kitāb al-ḥāwī fī ṭ-ṭibb*, III. Hyderabad: Osmania Oriental Publications Bureau, 1955.

Secondary sources

Abdalla, Fauzia. 2002. *Specific language impairment in Arabic-speaking children: Deficits and morphosyntax*. Ph.D. diss., McGill University.
Alhawary, Mohammad T. 1999. *Testing processability and effectiveness of computer-assisted language instruction: A longitudinal study of Arabic as a second/foreign language*. Ph.D. diss., Georgetown University.
——. 2003. "Processability theory: Counter-evidence from Arabic second language acquisition data". *al-ʿArabiyya* 36.107–166.
Amayreh, Mousa and Alice Dyson. 2000. "Phonetic inventories of young Arabic-speaking children". *Clinical Linguistics and Phonetics* 14:3.193–215.
Ammar, Wafaa. 1992. *Articulation disorders in Arabic-speaking children*. Ph.D. diss., University of Alexandria.
—— and Khaled Rifaat. 1998. "An acoustic study of vowels of deaf and normal-hearing thirteen to eighteen-year-old Egyptian females". *Jerash for Research and Studies* 2:2.29–38.
Bamford, John and Elaine Saunders. 1991. *Hearing impairment, auditory perception and language disability*. Baltimore, Md.: E. Arnold.
Bernhardt, Barbara H. and Joseph P. Stemberger. 1998. *Handbook of phonological development from the perspective of constraint-based nonlinear phonology*. San Diego, Calif.: Academic Press.
Bishop, Dorothy V.M. and Laurence B. Leonard (eds.). 2000. *Speech and language impairments in*

children: Causes, characteristics, intervention and outcome. Hove, East Sussex: Psychology Press.

Bronsted, K., P. Grunewell, G. Henningsson, K. Jansonius, J. Karling, M. Meijer, U. Ording, D. Sell, E. Vermeij-Zieverink, and R. Wyatt. 1994. "A phonetic framework for the cross-linguistic analysis of cleft palate speech". *Clinical Linguistics and Phonetics* 2.109–125.

Bzoch, Kenneth R. 1997. *Communicative disorders related to cleft lip and palate*. Austin, Tex.: Pro-Ed.

Dyson, Alice and Mousa Amayreh. 2000. "Phonological errors and sound changes in Arabic-speaking children". *Clinical Linguistics and Phonetics* 14:2.79–109.

Gold, Toni. 1980. "Speech production in hearing impaired children". *Journal of Communication Disorders* 13:6.395–418.

Hansson, Kristina and Ulrika Nettelbladt. 2002. "Assessment of specific language impairment in Swedish". *Logopedics, Phoniatrics, Vocology* 27: 4.146–154.

Harrell, Richard Slade. 1957. *The phonology of colloquial Egyptian Arabic*. New York: American Council of Learned Societies.

Ingram, David. 1990. *Phonological disabilities in children: Studies in disorders of communication*. London: Whurr.

Leonard, Laurence B. 1998. *Children with specific language impairment*. Cambridge, Mass.: MIT Press.

——. 2000. "Specific language impairment across languages". Bishop and Leonard (2000:115–129).

Lykke-Nielsen, Helle. 1997. "On acquisition order of agreement procedures in Arabic learner language". *al-ʿArabiyya* 30.49–94.

Nettelbladt, Ulrika. 1983. *Developmental studies of dysphonology in children*. Ph.D. diss., Lund University.

Oster, AnneMarie. 1990. "The effects of prosodic and segmental deviation on intelligibility of deaf speech". *Speech Transmission Laboratory, Quarterly Progress and Status Report* 2:3.79–88.

Pienemann, Manfred. 1998. *Language processing and second language development: Processability theory*. Amsterdam and Philadelphia: J. Benjamins.

Plant, Geoff and Björn Hammarberg. 1983. "Acoustic and perceptual analysis of the speech of the deafened". *Speech Transmission Laboratory, Quarterly Progress and Status Report* 2:3.85–107.

Rockey, Denyse and Penelope Johnstone. 1979. "Medieval Arabic views on speech disorders: al-Razi". *Journal of Communications Disorders* 12.229–243.

Salameh, Eva-Kristina. 2003. *Language impairment in Swedish bilingual children: Epidemiological and linguistic studies*. Ph.D. diss., Lund University.

——, Ulrika Nettelbladt, and Kjell Norlin. 2003a. "Assessing phonologies in bilingual Swedish-Arabic children with and without language impairment". *Child Language Teaching and Therapy* 19.338–364.

Salameh, Eva-Kristina, Gisela Håkansson, and Ulrika Nettelbladt. 2003b. "Developmental perspectives on bilingual Swedish-Arabic children with and without impairment: A longitudinal study". *International Journal of Language and Communication Disorders* 39.65–91.

Shahin, Kimary. 2002. "Remarks on the speech of Arabic-speaking children with cleft palate". *California Linguistics Notes* 27:1.1–10.

Mohammad T. Alhawary
(University of Oklahoma)

Language Loss

1. Introduction

Language loss refers to a variety of phenomena relating to the loss of a whole language or a portion thereof by an individual or a speech community (Freed 1982:1; Jaspaert a.o. 1986:38; Lambert and Freed 1982:6). Different definitions converge onto an understanding of language loss as partial or complete disappearance of one or more linguistic features from the grammar of a speaker. This feature can be the meaning of a lexical item (reception skills), the item itself (production skills), or a language rule (phonological, morphological, syntactic, or pragmatic).

A commonly accepted definition in the psycholinguistic tradition of language loss in general is that it "can be best characterized as difficulty in retrieval rather than total loss" (Hakuta and D'Andrea 1992) or that language loss is more of a "decrease in the ability to have immediate access to a word in production and perception" (Hulsen a.o. 2002:33), rather than a total loss of grammatical notions or lexical knowledge. The notion that inaccessibility rather than loss is at stake stems from research on memory systems. In essence, findings about retrieval and retention of knowledge in general are extended to retention of a specific type of knowledge, namely language. Loftus and Loftus qualify forgetting "much like being unable to find something that we have misplaced somewhere" (1976:78). In their view, forgetting is a matter of temporary inaccessibility, and once the right retrieval cue is provided, the information sought after could successfully be retrieved.

The literature also makes a distinction between pathological and nonpathological loss. As the terminology suggests, pathological loss occurs as a result of physical damage to the brain, as in certain types of aphasia (Broca's aphasics) or simply as a result of aging (→ language pathology). Nonpathological loss occurs in bilingual and multilingual situations where one language ceases to be used or becomes limited in the contexts in which it can be used. In the remainder of this entry, only nonpathological language loss is covered.

The term 'loss' does not imply that there is a total or complete loss of a language feature. To avoid this implication, some researchers use 'attrition', instead (e.g. Weltens and Grendel 1993).

2. DOMAINS OF LANGUAGE LOSS

In addition to the categorization of individuals who can be subject to language loss, another categorization has been made, this time in reference to the languages that can undergo language loss and the contexts in which this can happen. De Bot and Weltens (1985) provide a useful categorization of the languages and contexts where language loss can occur.

i. Loss of L1 in L1 environment (e.g. dialect loss within the dialect community)
ii. Loss of L1 in L2 environment (e.g. loss of native languages by migrant workers)
iii. Loss of L2 in L1 environment (e.g. foreign-language loss)
iv. Loss of L2 in L2 environment (e.g. second-language loss by aging migrants)

In this categorization, L2 does not necessarily stand for the second language a speaker has learned; it can refer to any language learned after or simultaneously with one's native language.

3. FACTORS DETERMINING DEGREE OF LANGUAGE LOSS

Weltens and Grendel (1993:144–148) list five factors that can explain variation in the degree of lexical attrition to which bilinguals are subjected. The length of the period of disuse is one of these factors, although it cannot explain all the variation between individuals. Some speakers end up retaining more language skills after a long period of disuse, compared to speakers with a shorter period of disuse. A second factor is the original proficiency level. The third factor is attitude and motivation, both important in explaining the amount of lexical loss. The fourth factor discussed by Weltens and Grendel is the language distance. It is generally assumed that the further apart the mother tongue is from a second or foreign language, the more difficult it will be for a speaker to retain lexical knowledge in the nonnative language. Finally, the language environment can also have tremendous effects on the degree of attrition. In situations where a second or foreign language is widely in use, attrition will not proceed as fast as it would under circumstances where the second or foreign language is not so widespread.

4. METHODOLOGICAL ISSUES IN LANGUAGE LOSS RESEARCH

Ideally, to determine whether speakers have lost a certain feature of their language, one has to establish first what is learned or known (original proficiency level), choose a time period over which to test whether this feature is still present in the language system of the speaker, and devise a method by which to test this. The first step is rather straightforward. A number of tests and experiments can indicate whether a speaker knows a certain feature of a language system.

Once the problem of determining the issue of 'knowing' a linguistic feature is solved, one can proceed to test whether this knowledge still exists after a certain period of time. In psycholinguistic literature, word-recognition experiments, for example, can tell us whether speakers know if a word X is part of language Y and whether they know the meaning of the word in question. Picture-naming tasks and word-recognition tasks are often used in psycholinguistic research on language attrition to test production skills and perception skills respectively. The advantage of this method is clear: instead of a wholesale testing of the knowledge of a word ('do you know what a book is called in Arabic?'), one can test this knowledge at different levels. Can speakers produce the linguistic form if shown a picture corresponding to that form? Do speakers know the meaning of a word they have identified as part of language Y? Do the speakers know the exact meaning of word X, or, as in our example of 'book', only that word X has to do with the semantic field of 'writing'? All these levels can be tested in an online method, which adds the extra dimension of measuring reaction times in the tasks at hand.

If speakers do not provide the meaning of a word, it does not follow that they do not know its meaning. The well-known 'tip of the tongue' phenomenon applies in a large number of cases where speakers almost certainly know the meaning of words but cannot get to them or access them in their mental lexicon.

In addition to the issue of the original proficiency level, there is also the problem of choosing a time interval after which a test or experiment can be run to check for language loss. Choosing a long period of disuse (e.g. ten years or longer) raises the problem of finding speakers who are willing to sit for an experiment or test and will also be available after such a period of time to re-sit for a similar experiment. Choosing shorter periods of time might have a test-retest effect. One way to get around this problem is to test immigrants who left their home country when they were adults and settled in a country with a different dominant language. Some studies in this vein are Hulsen (2000) and de Bot and Weltens (1992). An alternative method is to design a pseudolongitudinal research study that is synchronic in nature but uses different age groups. For example, in an immigrant community, one would test the language skills of the second generation in their home language and compare their skills to those of their parents. Obviously, such a comparison is flawed in that the researcher cannot know what the original language skills were like before the onset of language loss. The same problem would also arise when one compares similar age groups residing in different countries (El Aissati 1997). The purely longitudinal method is certainly the one with the most robust results; if there is a decrease in language skills, the decrease will be easy to identify since there is a clear reference point, namely the original proficiency level.

The following review, although not exhaustive, illustrates the studies in this field.

5. EXAMPLES OF CASES OF LANGUAGE LOSS

Research on language loss has uncovered a number of linguistic features susceptible to loss (see Andersen 1982 for a useful list of hypotheses about aspects of language susceptible to language loss). In a large number of publications on this issue, reference is made to attrition at the level of language and not at the level of the speaker. When we read that a given feature is not found in a language in an immigration context but is found in the 'same' language in the original country, we should not assume that a speaker has in fact lost this feature. A speaker not using the feature in question might simply never have learned it. With this in mind, we proceed to a brief survey of cases of language attrition.

5.1 *Lexical knowledge*

Knowing a word involves knowing its form, its position, its function, and its meaning (Nation 1990:31, cited in Weltens and Grendel 1993:141). In psycholinguistic tradition, this knowledge is measured by the speed with which it can be retrieved. Studies on attrition at the lexical level outnumber those on other aspects of language attrition because the lexicon is the domain most likely to be affected by the process of language attrition (Seliger 1985), but also because of the relative importance of vocabulary in language learning classes (Weltens and Grendel 1993:142). Some of the studies dealing with word processing in bilingual contexts, where language attrition is likely to occur, are reviewed in Weltens and Grendel (1993). The results of these different investigations conflict in a number of ways, for example in the amount of knowledge that was found to be subject to loss. Scherer (1957) and Weltens (1989) found no lexical attrition, while Bahrick (1984) and Verkaik and van der Wijst (1986) found some attrition, and others (Geoghegan 1950 and Schumans a.o. 1985) showed an improvement in the lexical knowledge of their subjects (Weltens and Grendel 1993:144). Next to this variation, in terms of the amount of lexical attrition, the authors also cite the rate of attrition as an important point of conflict. Thus, while Bahrick (1984) reports a 25 percent rate of attrition during the first year of disuse, Messelink and Verkuylen (1984) explain that attrition proceeds at a much lower rate in the first periods of disuse, and they speak of an 'initial plateau' before attrition can be evidenced.

At the level of word meaning, Seliger and Vago (1991:8) report that in Hungarian spoken in the United States, the distinction between the verb *tud* 'to know how to do something' and the verb *ismer* 'to know someone' has been lost as a result of → semantic extension under the influence of English. Another example is the category switch (Schmidt 1991), whereby a category is maintained at the conceptual level but is designated by a different category, for example by using prepositions to substitute for affixes. Further, loan translation or calquing has also been considered as a possible result of language loss. Seliger and Vago (1991:8–9) provide the example of an L2 phrase translated into L1, where it is ungrammatical. They explain that the German equivalent of the English *forget it!* is *lass es bleiben!* But the literal translation *vergiss es!* is attested in contexts of language loss.

5.2 *Morphosyntax*

Morphosyntactic constructions can also be subject to erosion or loss. For example, Maher (1991), drawing on minority languages like Finnish in Minnesota, Slavonic languages in the American Midwest, French in Louisiana, and Albanian in Greece, reports a tendency to replace synthetic forms by analytic forms or periphrastic constructions, and, as a result of disappearing inflections, a rigid word order replaces a previously more flexible one. Clyne (1977) reports a preference for SVO over SOV among the immigrant Dutch population in Australia.

A number of morphological changes instantiating language loss have been identified in the literature. Weinreich (1966:43) reports the loss of the dative/accusative case distinction among speakers of German in Texas. Maher (1991:71) cites Bhatia (1982) in reporting on an intergenerational loss of gender, number, and case distinctions in the NPs of young speakers of Trinidad Hindi. Romaine (1989:73) gives the example of gender distinctions in Asia Minor Greek, under the influence of Turkish, which does not have gender distinctions. Plural formation has also been documented as a case of language loss. Dorian (1978:601) reports that East Sutherland Gaelic has lost most of the variation it had in plural formation in the speech of semi-speakers (Dorian's terminology) of this language, who have resorted to the invariable use of the suffix *-en*. A further instance of morphological loss can be found in the Hebrew spoken as an L1 in a dominant L2 environment, where the verbal template iCaCe(C) has been found to be overgeneralized to other forms in a case study (Kaufman and Aronoff 1991:185).

5.3 *Phonology*

The phonological component has also been the subject of some research dealing with language loss. Weinreich (1966:18–19) provides a classic categorization of what could happen in a → language contact situation, i.e. in situations where language loss is very likely to be attested. Weinreich speaks of underdifferentiation of phonemes, overdifferentiation of phonemes, reinterpretation of distinctions, and phone substitution. An example of underdifferentiation is the merging of short /i/ and long /iː/ in English by French/English bilinguals who produce the two as a short vowel, since this distinction is not found in French (Romaine 1989). Overdifferentiation can be illustrated by speakers of Romansch as a native language who make a vowel-length distinction in Swiss German, where no such phonological distinction applies. Seliger and Vago (1991:9) report that Hungarian/English bilinguals growing up in the United States show evidence of attrition in intonational patterns of Hungarian as a minority language, under the influence of English as a dominant language.

6. LANGUAGE LOSS IN ARABIC

Studies on Arabic within the psycholinguistic tradition are in general very rare, and those specific to language loss are even more so. Kenny (1996) conducted research on fluency rates of adult Arab immigrants in the Detroit area, originating from Ramallah. Kenny used a combination of sociolinguistic and psycholinguistic insights in a 'Macro-Fluency Model' (1996:234), an approach that has as a goal "to provide information about the intensity of the overall breakdown of competence in L1, that is the loss of a speaker's ability to retrieve and process linguistic elements generally". His results indicate different degrees of processing difficulties among the informants, evidenced by a higher frequency of silent pauses in the speech of immigrants who had been living in the Detroit area for 21 years or more. The strength of this model, according to Kenny, is that it opens a window on the actual speech disturbances characteristic of speakers showing signs of language attrition, mainly pauses, errors, and false starts. These were correlated with external factors, like education level and length of stay. Data for Kenny's study were collected using a questionnaire for the sociolinguistic background and an interview for the speech part. In particular, silent pauses and verboseness were taken to be signs of language attrition in the study in question.

Kenny (1996) interprets his findings about the speech-delivery disturbances as being symptomatic of attrition at the competence level of language. This interpretation, however, goes against other findings on language attrition. Evidence of attrition at the level of competence has yet to be provided in studies on language attrition. The observed speech 'disfluencies' in Kenny's study might as well be indicative of troubles at the language-processing level and not at the competence level.

In a more recent article, Kenny focuses his search on the relationship between fluency in Arabic/English and English/Arabic → code-switching as an indicator of language attrition, and "a variety of temporal, social, and psychological variables" (2002:332). The findings of this study indicate a variety of correlations between code-switching disfluency and other speech disturbances, but they do not support a linear interpretation of the relationship of speech disturbances with length of stay, or language attrition with language disuse.

Another study on language attrition among speakers of Arabic is El Aissati (1997), which examines the loss of language skills in Moroccan Arabic of second-generation youngsters of Moroccan descent, born and/or living in the Netherlands. The areas examined in this study are phonology, morphology, and morphosyntactic sentence processing, at the production and perception levels.

6.1 *Phonology*

To test the production skills in Moroccan Arabic, participants in the study of El Aissati (1997) were asked to tell a narrative based on a picture book (Mayer 1969). The recordings of these narratives were checked for any anomalies, especially regarding the so-called marked phonemes or segments (like the uvular /q/ and the pharyngeals /ḥ/ and /ʕ/). Informants with the lowest language proficiency did show some difficulties in pronouncing some segments, especially /q/, /š/, and /ž/. None of these three segments is part of the phonological inventory of Dutch, which might be the reason for these difficulties. The uvular /q/ was substituted by the velar /k/ in a large number of cases. The palatals /š/ and /ž/ were substituted by the alveolars /s/ and /z/ respectively.

Next to substitution, the informants in question resorted to geminate reduction, which might be partly explained by the absence of lexical gemination in Dutch. Some of the lexical geminates that underwent reduction are /qq/ in *kayneqqez* 'it is jumping' and /dd/ in *dda* 'he took'. Geminates resulting from assimilation, especially word-initially, were more affected by the process of reduction. Some illustrative examples are /dd/ in *d-derri* 'the child', /nn/ in *nnmel* 'the ants', and /v/ in *ṣṣbaḥ* 'the morning'. Next to geminate reduction, a speaker with the lowest proficiency score in Moroccan Arabic resorted to resyllabification. The word *nnḥel* 'the bees' was produced as *neḥḥel* instead of *nnḥel*. The informant turned the monosyllabic word with the structure CCCVC into a bisyllabic one with the structure CVC.CVC. The initial geminate *nn* followed by *ḥ* results in a cluster that is not allowed by Dutch syllabification rules. This can be an explanation for the resyllabification carried out by the informant in question. While geminate reduction was observed in a large number of the informants, resyllabification was limited to one informant who was at a very advanced stage of language loss.

The lack of ability to produce a certain segment does not necessarily mean that the informant in question also cannot perceive this segment. The informants in El Aissati's (1997) study took part in an experiment that was set up to test the capacity of identifying the marked segments /ḥ/ /q/ /ʾ/ /ṣ/ and /ḍ/, none of which is part of the Dutch phonological inventory. The results indicate that the target group living in the Netherlands did not differ from the reference group in Morocco in identifying these segments. This confirms the general idea that production is more affected than perception in contexts of language loss.

6.2 *Morphology*

To test language loss in the area of morphology, a plural formation task was elaborated. Plural formation in Moroccan Arabic is a rather complex area of grammar, especially when compared with the plural morphology of Dutch.

In addition to the Classical Arabic distinction between broken plurals and sound plurals, Arabic has a large number of irregular forms. In his classic grammar manual for Moroccan Arabic, Harrell (1962) devotes 28 pages to plurals in this language.

The procedure of the morphology test was as follows. Informants were shown a picture of one object, and asked to say what it was; then, they were shown a picture of the same object drawn twice and asked to say what that was. The examiner noted the answers on a piece of paper.

The analysis of the results indicated large gaps in the knowledge of the speakers. The general tendency was that the speakers with the lowest proficiency levels opted for a limited number of preferred strategies, like the suffixation of *-in* or the suffixation of the Dutch plural marker *-s*, resulting in forms like *kursin* 'chairs', *kelbin* 'dogs', *žmilin* 'camels' in one speaker, and *dubbs* 'bears', *ṣenduqs* 'boxes', *ḥmars* 'donkeys' in another (El Aissati 1997:74).

6.3 *Morphosyntax*

Languages differ in how they use morphosyntactic cues in sentence processing (MacWhinney and Bates 1989). English, for example, relies heavily on word order to determine the role of agent in an utterance, while Dutch monolinguals rely more on case inflection. To examine whether second-generation speakers of Moroccan Arabic in the Netherlands differ from native speakers of Moroccan Arabic in Morocco in the way they process sentences in determining the agent function, a test was designed manipulating the following cues: word order, Subject-Verb number and gender agreement, animacy, and word stress (see El Aissati 1997:71-95). The results indicate that the group growing up in the Netherlands relied slightly more on agreement and a little less on animacy than the control group. These results can be at least partly ascribed to the influence of Dutch. Here, too, one can speak of signs of 'losing' processing strategies of the native language.

7. CONCLUSIONS

Research on language loss in Arabic is very scant. The limited amount of available data suggests that native speakers of Arabic as a minority language, as a consequence of restricted language input, show signs of language loss at the different levels of language. This does not necessarily imply a loss of linguistic features previously under the command of a given speaker. More research is certainly needed to uncover the processes that the different varieties of Arabic undergo when they are transplanted in an immigration context or when they are used as minority languages.

BIBLIOGRAPHICAL REFERENCES

Andersen, Roger W. 1982. "Determining the linguistic attributes of language attrition". Lambert and Freed (1982:83–118).

Bahrick, Harry P. 1984. "Fifty years of language attrition: Implications for programmatic research". *Modern Language Journal* 68.105–118.

Bhatia, Tej K. 1982. "Trinidad Hindi: Three generations of a transplanted variety". *Studies in the Linguistic Sciences* 11.135–150.

Bot, Kees de and Bert Weltens. 1985. "Taalverlies: Beschrijven versus verklaren" [Language loss: Description versus explanation]. *Handelingen van het 38e Nederlandse Filologencongres,* 51–61. Amsterdam and Maarsen: APA-Holland Universiteits Pers.

——. 1992. "Lexical aspects of language attrition and shift". *Immigrant languages in Europe*, ed. Guus Extra and Ludo Verhoeven, 309–322. Clevedon: Multilingual Matters.

Clyne, Michael. 1977. "Nieuw Hollands or Double Dutch". *Dutch Studies* 3.1–30.

Dorian, Nancy. 1978. "The fate of morphological complexity in language death: Evidence from East Sutherland Gaelic". *Language* 54.590–609.

El Aissati, Abderrahman. 1997. *Language loss among native speakers of Moroccan Arabic in the Netherlands*. Tilburg: Tilburg University Press. (Ph.D. diss., University of Nijmegen, 1996.)

Freed, Barbara F. 1982. "Language loss: Current thoughts and future directions". Lambert and Freed (1982:1–5).

Geoghegan, B. 1950. "The retention of certain secondary school subjects by high school students over the summer vacation period". Ph.D. diss., Fordham University.

Hakuta, Kenji and Daniel D'Andrea. 1992. "Some properties of bilingual maintenance and loss in Mexican background high-school students". *Applied Linguistics* 13:1.72–99.

Harrell, Richard. 1962. *A short reference grammar of Moroccan Arabic.* Washington, D.C.: Georgetown University Press.

Hulsen, Madeleine. 2000. *Language loss and language processing: Three generations of Dutch migrants in New Zealand*. Ph.D. diss., University of Nijmegen.

——, Kees de Bot, and Bert Weltens. 2002. "Between two worlds: Social networks, language shift, and language processing in three generations of Dutch migrants in New Zealand". *International Journal of the Sociology of Language* 153.27–52.

Jaspaert, Koen, Sjaak Kroon, and Roeland van Hout.

1986. "Points of reference in first-language loss research". *Language attrition in progress*, ed. Bert Weltens, Kees de Bot, and Theo van Els, 37–49. Dordrecht: Foris.

Kaufman, Dorit and Mark Aronoff. 1991. "Morphological disintegration and reconstruction in first language attrition". Seliger and Vago (1991:175–188)

Kenny, Keith Dallas. 1996. *Language loss and the crisis of cognition: Between socio- and psycholinguistics*. Berlin and New York: Mouton de Gruyter.

——. 2002. "Codeswitch fluency and language attrition in an Arab immigrant community". Rouchdy (2002:331–352).

Lambert, Richard D. and Barbara F. Freed (eds.). 1982. *The loss of language skills*. Rowley, Mass.: Newbury House.

Loftus, Geoffrey A. and Elizabeth F. Loftus. 1976. *Human memory: The processing of information*. Hillsdale, N.J.: Erlbaum.

MacWhinney, Brian and Elizabeth Bates (eds.). 1989. *The cross-linguistic study of sentence processing*. New York: Cambridge University Press.

Maher, Julianne. 1991. "A crosslinguistic study of language contact and language attrition". Seliger and Vago (1991:67–84).

Mayer, Mercer. 1969. *Frog: Where are you?* New York: Dial Books for Young Readers.

Messelink, J. and H. Verkuylen. 1984. *Het verlies van woordenkennis van het Frans* [The attrition of word knowledge in French]. M.A. thesis, University of Nijmegen.

Montfort, R. van. 1986. *Een onderzoek naar cue-validity bij vreemdetaalsprekers* [An investigation into cue-validity among speakers of a foreign language]. M.A. thesis, University of Nijmegen.

Nation, Paul I.S. 1990. *Teaching and learning vocabulary*. London: Newbury House.

Olshtain, Elite and Margaret Barzilay. 1991. "Lexical retrieval difficulties in adult language attrition". Seliger and Vago (1991:139–150).

Romaine, Suzanne. 1989. "Language maintenance, shift and death". *Bilingualism*, ed. Suzanne Romaine, 38–50. Oxford: Blackwell.

Rouchdy, Aleya (ed.). 2002. *Language contact and language conflict in Arabic: Variations on a sociolinguistic theme*. London and New York: Routledge Curzon.

Scherer, G.A.C. 1957. "The forgetting rate in learning German". *German Quarterly* 30.275–277.

Schmidt, Annette. 1991. "Language attrition in Boumaa Fijian and Dyirbal". Seliger and Vago (1991:113–124).

Schumans, J., Charles van Os, and Bert Weltens. 1985. "Vocabulairekennis in de vreemde taal na beëindiging van het onderwijs" [Foreign-language vocabulary knowledge after course completion]. *Toegepaste Taalwetenschap in Artikelen* 23.81–89.

Seliger, Herbert. W. 1985. "Primary language attrition in the context of other language loss and mixing". Unpublished ms., Queens College, City University New York.

—— and Robert M. Vago (eds.). 1991. *First language attrition: Structural and theoretical perspectives*. Cambridge: Cambridge University Press.

Soesman, Aviva. 1997. "An experimental study on native language attrition in Dutch adult immigrants in Israel". *Dutch overseas: Studies in maintenance and loss of Dutch as an immigrant language*, 181–194. (= *Studies in Multilingualism*, 9.) Tilburg: Tilburg University Press.

Verkaik, Paula and Per van der Wijst. 1986. "Taalverlies en woordherkenning in het Frans als vreemde taal" [Language attrition and word recognition in French as a foreign language]. M.A. thesis, University of Nijmegen.

Weinreich, Ulrich. 1966. *Languages in contact: Findings and problems*. Repr. The Hague: Mouton. (1st ed., 1953.)

Weltens, Bert. 1989. *The attrition of French as a foreign language*. Dordrecht and Providence: Foris.

—— and Marjon Grendel. 1993. "Attrition of vocabulary knowledge". *The bilingual lexicon*, ed. Robert Schreuder and Bert Weltens, 135–156. Amsterdam and Philadelphia: J. Benjamins.

Internet resources:

http://w3.byuh.edu/academics/lang/attritionbiblio/index.htm. Web site maintained by Professor Lynne Hansen.

Abderrahman El Aissati (Tilburg University)

Language Maintenance → Language Shift: Amazigh

Language Pathology

Language pathology' (also known as 'language disorder' or → 'language impairment') refers to deficiencies in language use (production or comprehension or both) due to a clear physical cause. Roman Jakobson (1971) was probably the first linguist to stress the insights that can be drawn from pathological deficits in linguistic performance. What he called 'experiments in nature', he argued, not only provide information on the nature of the deficit itself but also provide a testing ground for theoretical linguistic claims and psycholinguistic models of language acquisition and use. Over the past few decades, language deficits have indeed been used to argue for the psychological reality of constructs developed within linguistic theory. Likewise, both psycholinguists who follow the Chomskyan 'nativist', 'autonomous', 'modular' tradition and those who adhere to the more 'cognitivist', 'interactive', 'connectionist' position often turn to language-impaired populations as a source of neurolinguistic and psycholinguistic data to support their claims.

Language pathologies are distinguished from speech pathologies in that the former are concerned with dysfunctions of the symbolic linguistic system, including manifestations of the deficit in all modes of language use such as writing, reading, speech, hearing, and signing (e.g. aphasia, dyslexia, agraphia), while the latter are concerned with the use of sounds without any reference necessarily to meaning or phonology (e.g. aphonia, stuttering, or hearing impairment). The traditional classification of language pathologies is along the lines of production and comprehension. Several terms have been used to refer to this distinction, e.g. expressive/receptive, motor/sensory, decoding/encoding. Another classification within the field reflects whether the disorder affects the linguistic system during the course of its development (developmental language disorder), or after the system is realized (acquired language disorder).

1. Developmental language disorders

By the age of five or six most children will have mastered the sound system and the basic structure of their native language. Although there will be aspects of the language that they continue to refine, by six years old they will have an impressive vocabulary and a more impressive command of syntax. They can take part in conversations and use language appropriately in most social situations. There are children, however, who, in the absence of any organic etiology (e.g. cleft palate or hearing impairment), produce speech that exhibits a sound system that is deviant from adult language, have difficulties acquiring new words or using newly acquired words, or tend to acquire simpler, more frequently used and concrete vocabulary items limited to names of persons and objects. Likewise, there are other children who continue to use words and phrases stripped of grammatical morphemes, such as plural or past tense markers or articles, and there are those who produce short or incomplete utterances, often with incorrect word order. Complex structures such as passives or embedded questions may appear much later in development or not at all. These language-disordered children often show deficiencies in all aspects of language (phonological, morphological, syntactic, semantic, and pragmatic) that manifest themselves in production as well as comprehension.

Language disorders are often but not necessarily associated with well-defined physical handicaps such as cerebral palsy, Down syndrome, autism, brain injury, and hearing impairment. Normal hearing is essential for oral language development. Consequently, some language disorders can be attributed to hearing impairment or loss; the extent of the language disorder is positively correlated with the amount and time of onset of hearing loss (partial or total) the child suffers (→ language impairment). In many cases of developmental language disorders, however, it is not possible to specify the cause or causes of the language deficit.

One of the most studied syndromes that affect the developing sound system is known as Phonological Disorder. The speech production of children in this category exhibits a sound system that is deviant from the target or adult language and which continues beyond the norms of development. The disorder manifests itself in substitution errors such as substituting glides for liquids or stops for fricatives, or omission errors such as word-final stop deletion (see Dyson and Amayreh 2000). Stoel-Gammon and Dunn (1985) characterize the pattern of substitution and omission in phonological disorders within the framework of Natural Phonology. Assuming an adult-like underlying representation, the disorder then is the result of the application of natural simplifying phonetic processes such as fronting or gliding. Dinnsen and Elbert (1984) and Dinnsen (1993), however, attribute the errors to an underspecified underlying representation that varies among children (thus explaining the different patterns of substitution and omission observed in this disorder). For a full review, see Dinnsen (1996a, 1996b). Ingram (1989) argues that the disorder is one of delay rather than deviance, since the phonetic inventory of phonologically disordered systems resembles those of younger normally developing children. Ingram also argues that the reduced inventory is the result of merger between segments. Instrumental analyses, however, of the acoustical quality of these children's reduced inventory have revealed that the children indeed have some knowledge of the underlying target distinctions (see Maxwell and Weismer 1982; Hoffman a.o. 1983; Forrest a.o. 1990). Maxwell and Weismer (1982), for example, found that the child under investigation who collapsed or merged the voice contrast between /t/ and /d/ nevertheless used Voice Onset Time (VOT) for producing an adult-like [d] and for distinguishing [d] from other voiced obstruents and from voiceless obstruents. They argue that although the phonemic inventory is perceptually reduced, the children's underlying system retains the adult phonemic distinctions, albeit differently.

Developmental Aphasia (also known as congenital aphasia or childhood aphasia) traditionally referred to comprehension and production language disorders that are not attributed to postnatal neurological damage, hearing impairment, mental retardation, or social or physical deprivation (the term is controversial since it reflects diagnosis by exclusion rather than by confirmed neurological evidence). The child's language is characterized by limited vocabulary, short utterances, and a high degree of grammatical morpheme omission. Recently, the term Specific Language Impairment (SLI) has been used to refer to this syndrome. There is disagreement in the literature as to whether word order or phrase structure is affected in SLI. The consensus, however, is that SLI children suffer from damage to grammatical morphology, with inflectional morphology (free and bound grammatical morphemes encoding case, gender, number, person, etc.) being most impaired. SLI children either omit or use grammatical morphemes incorrectly. Abdalla (2002) investigated tense, Subject-Verb agreement, determiners, and prepositions in the production of children speaking Ḥijāzī Arabic with SLI. She found that the performance of SLI children was significantly different from that of matched-age and matched-language groups, with more errors of substitution than deletion. The SLI children in her study, for example, substituted 1st person for 3rd person and masculine for feminine, as well as the imperative verb form for the finite tense marking (see Safi-Stagni 1992 for similar findings on agrammatic aphasic speakers in the same dialect). Clearly, with highly inflectional languages like Arabic and Hebrew, children with SLI tend to have fewer omissions than in English. The linguistic accounts of the syndrome suggest that the normal development of the grammatical representations (the central grammatical module) is disrupted (see Clahsen 1991; Gopnik and Crago 1991; Leonard a.o. 1992). The psychological accounts, on the other hand, suggest a processing impairment, such as an auditory processing problem or a short-term memory deficit. Most of the empirical results, however, are consistent with a missing-agreement hypothesis where SLI children have problems establishing agreement relations in grammar, e.g. Subject-Verb agreement, agreement within noun phrases, and case and gender agreement (see Rice and Otting 1993; Clahsen and Rothweiler 1993). For a full review, see Bishop (1992, 1997) and Leonard (1998).

Children exhibiting a selective deficit to the processing of the written code (reading, writing, and/or spelling) are called dyslexic, and the deficit is known as Developmental Dyslexia. Dyslexia (which also includes dysgraphia) is characterized by systematic but peculiar reading and spelling errors (such as reversed or misordered letters), poor directional ability that often confuses left and right, difficulty in naming letters of the alphabet or correlating graphemes with phonemes, difficulty reading nonsense words, and better reading performance of concrete as opposed to abstract words (see Ellis 1993; Miles 1993).

Many researchers believe that language acquisition and language processing involve rules and representations that form a distinct cognitive subsystem. Others believe that the rules and representations developed within linguistic theory are epiphenomena of the more general and basic cognitive processes used in general learning. Since various components of grammatical knowledge develop in parallel, with normal development, it is difficult to determine whether any component is autonomous. Disorders of language during the early years of language development, such as SLI, can shed light on the nature of the developing representations and processes involved in normal acquisition. Developmental language pathologies attributed to genetics (autism or Down syndrome) have been used to argue for innate domain-relevant rather than domain-specific mechanisms that are crucial to language acquisition and development. See Elman a.o. (1996) and Karmiloff-Smith (1997, 2001) for a full discussion of the contribution the study of atypical language has to the question of innateness and to the bigger issues of nature vs. nurture or the role of gene vs. environment.

2. ACQUIRED LANGUAGE DISORDERS

The total or partial inability to understand or to use language in any one of its forms as a result of brain damage is known as 'aphasia'. The type and extent of the language disturbance is a function of the site and size of the associated brain lesion. Aphasias are qualitatively different from non-aphasic disorders of speech such as dysarthria (speech disorders due to paralysis or weakness of the speech muscle) or verbal apraxias (disorders of sequenced movements in the absence of speech muscle weakness or paralysis).

One of the first outcomes of the study of language pathologies, adult aphasia in particular, is the correlation of language functions or components with brain sites (the localizationist theory). In this theory, which was proposed back in 1885 with the Wernicke-Lichteim model, production and comprehension, reading and writing, inflectional morphology, or even regular vs. irregular morphology are localized in different brain regions such as Broca's area, Wernicke's area, the angular gyrus, the arcuate fasciculus, etc. Evidence, however, from patients who had lesions in Broca's or Wernicke's areas and who presented different language deficits, as well as patients who exhibited similar language difficulties as a result of lesions in different areas, has been used to advance the opposite view that claims that every region in the brain is equally involved in all language functions (the theory of equipotentiality). Despite the controversy, in the majority of cases where language was disrupted as a result of brain damage, the lesion is in or around the classic Broca's and Wernicke's areas. Additionally, data from brain-damaged adults and children have been relevant to issues of hemispheric lateralization or asymmetry and to questions of the roles of each hemisphere in language processing.

The classic taxonomy of aphasic syndromes was either neurologically (e.g. Broca's or Wernicke's aphasia) or behaviorally (e.g. anomia) motivated, or both (e.g. conduction aphasia). Broca's aphasia, associated with lesions to Broca's area, is also known as non-fluent aphasia, expressive aphasia, efferent aphasia, verbal aphasia, and encoding aphasia and is characterized by telegraphic speech (characteristically lacking grammatical markers), which is uttered slowly, with great effort and poor articulation and a relatively intact comprehension. Wernicke's aphasia is correlated with lesions to Wernicke's area and is also known as fluent aphasia, expressive aphasia, sensory aphasia, and decoding aphasia. Wernicke's aphasics produce well-articulated long structures with normal grammatical marking and correct prosody, which are nevertheless abnormal since they are devoid of content and are replete with circumlocutionary phrases or verbal or phonemic substitutions. The comprehension of Wernicke's aphasics is usually severely disrupted.

Goodglass and Kaplan (1972) and Goodglass (1982) offer a classification along the lines of comprehension and production, which is based on the nature of the language disturbance rather than lesion site. Auditory and visual comprehension can be disrupted jointly or independently. In the former, affected individuals fail to attach meaning to words or structures they hear, particularly if the words are presented in isolation or the structures are complex, while in the latter, affected individuals may not be able to recognize individual letters or attach meaning to symbols they see written on a page (a syndrome also known as alexia). The most prominent of production disorders are word-finding difficulty, paraphasia, and loss of grammar or syntax. These can affect oral as well as written production (a phenomena described as 'agraphia'). Difficulties in naming (things, objects, or persons) where the individuals know what they want to say but cannot 'find' the right word or cannot get the right word 'activated', are known as word-finding difficulty. Sometimes, affected individuals cannot find or retrieve a word but can talk about it. If this difficulty occurs in the absence of all other aphasic symptoms, the condition is known as anomia. Often, patients (typically Wernicke's aphasics) produce phonemic (sound-based) paraphasias (word substitutions) or verbal (meaning-based) paraphasias to substitute for words they cannot retrieve. At times, the patient substitutes the missing word with an invented form or a neologism. The loss of grammar or syntax is apparent in the production of short utterances that are either stripped of grammatical markers (a phenomenon also known as agrammatism) or carry the wrong markers (paragrammatism). For a full review, see Bates and Wulfeck (1989), Bates and MacWhinney (1991), and Crystal and Varley (1998).

Safi-Stagni (1991, 1992) describes the manifestation of aphasia in two agrammatic speakers of the Ḥijāzī dialect of Saudi Arabia. The agrammatic speakers in the study showed relatively intact auditory and reading comprehension with a severe naming deficit and a severe agraphia, being able to write only their names, addresses, and letters of the alphabet. Their oral production was laborious and replete with substitution errors. Both individuals substituted major content words and function words such as pronouns and clitics, and above all inflectional bound morphemes marking tense, gender, and number. The substitutions, however, were always within category, i.e. one clitic pronoun for another or one tense marker for another. One of the most pronounced difficulties these patients exhibited lay in the breakdown of agreement relations, specifically of those marking number and gender. The breakdown was sensitive to structural constraints, i.e., it was more evident across phrasal boundary (e.g. between NP and VP) than within phrasal category (e.g. between adjective and noun within an NP). Safi-Stagni argues that the 'manifestation' of agrammatism is language-dependent, insofar as the production errors of the affected individuals occur within the morphophonological constraints of the affected language. Thus, in Arabic, grammatical morphemes are substituted, as opposed to the deletion observed in Indo-European languages like English (the omission of grammatical markers in Arabic would result in unpronounceable three-consonant roots).

Mimouni a.o. (1995) describe an agrammatic Lebanese bilingual (Arabic-French) patient whose oral production was severely impaired but whose oral and written comprehension of single words was spared. Sentence comprehension, on the other hand, was severely affected in both Arabic and English. He showed patterns of semantic paraphasias and a plethora of errors mainly on morphological markers and prepositions. Mimouni a.o. argue that although the patient's written comprehension in general was better preserved in Arabic than in French, the locus of the selective 'functional' lesion was limited to the output lexicon but affected the same component in both Arabic and French. More recently, Friedmann (2001) presented findings of a syntactic analysis of 14 Hebrew and Palestinian Arabic agrammatic patients. Friedmann uses the Tree Pruning Hypothesis (TPH) to provide a unified account for the patterns of dissociations observed within and between languages (like English and Arabic), where structures that relate to high nodes of the tree (e.g. full relatives and embeddings as well as WH-questions) are impaired, whereas lower structures (e.g. agreement inflection, reduced relatives, non-finite verbs, and yes/no questions in Hebrew and Arabic) are spared. The underlying deficit, then, is in the inability to project the syntactic tree to its highest nodes of TP and CP. Friedmann uses these findings to argue for the psychological reality of the syntactic tree.

Within the domain of psycholinguistics, patterns of impaired and intact language functions are explained in terms of damage to one or more components of a theory or model of normal language processing in an attempt to draw conclusions about language processing in normal subjects. For example, the fact that the processing elements (or modules) of language (e.g. inflectional morphology) can be and are selectively impaired (e.g. in agrammatic and SLI cases) has been used to support hypotheses regarding the 'modularity' or autonomy of cognitive systems advocated by Chomsky (1980, 1999), Gardner (1983), and Fodor (1983). Data from language disorders are relevant to the understanding of the mental representations and computations involved in language use (production and comprehension) and acquisition. For example, using aphasic data, computational distinctions have been drawn between inflectional and derivational affixation (see Badecker and Caramazza 1989). Garrett (1980, 1984) develops a model of language production using evidence from both speech errors and aphasic patients. Safi-Stagni (1991, 1992) correlates the aphasic errors with speech errors of normal speakers of Ḥijāzī Arabic and attributes the aphasic deficit to a breakdown in the operations that specify the positional level in Garrett's model. Safi-Stagni (1995) uses the discrepancy in the performance of an Arabic-speaking patient in reading aloud simple vs. complex forms to argue for the decompositional view of processing complex forms. The substitution (almost always within class) of grammatical morphology shows that complex forms are stored decomposed and that speech production involves a process of assembly. In the same study, Safi-Stagni also reports an observed discrepancy in the processing of derivationally vs. inflectionally complex forms with a 50 percent success rate in reading the former as opposed to 2 percent in reading the latter. Additionally, a frequency effect was noted for derived but not for inflected forms.

In a study investigating aspects of the mental lexicon (representation, organization, and access), Mimouni a.o. (1998) describe two agrammatic speakers of Algerian Arabic. Both speakers in their study were bilingual (Arabic-French). Both individuals had lesions in the left hemisphere, which left them with a mild right hemiplegia. The female speaker exhibited a mild oral comprehension deficit of morphologically complex words and sentences, but a severe comprehension deficit when the same complex words and sentences were presented in writing. The second speaker (a male subject) had a relatively intact oral production with a severe impairment in oral and written comprehension of morphologically complex words and sentences. Mimouni a.o. used a lexical decision task of regular (sound) vs. irregular (broken) plural forms in Arabic to compare the performance of normal and language-disordered Arabic subjects. Their study revealed a discrepancy in the processing of irregular vs. regular forms of plurals in Arabic, with the irregular forms indicating a whole-word access strategy, while a decompositional strategy of word-plus-affix was used for the regular ones. They argue that although the overall pattern of lexical access for both normal and aphasic subjects was the same, the aphasic subjects in their study showed a higher Reaction Time (RT). Accordingly, the speakers' performance (on the lexical access) was simply slower than normal subjects rather than lexical access being unavailable. Recently, Prunet a.o. (2000) used production errors from an Arabic-speaking aphasic as well as normal subjects (speech errors and language games) as evidence for the independent psychological status of the consonantal → root as a computational unit used in production (a theoretical framework proposed by McCarthy [1981] and [1982] for Semitic languages).

Many researchers believe that reading and writing involve the mapping of written forms onto phonological representations and as such require prior activation of a phonological code, arguing that the initial stage of word recognition consists of extracting a pronunciation from sublexical units (also known as the phonological route to reading). Thus, the recognition of spoken and written words is the same. There are others who maintain that entire written words are recognized prior to being understood or transformed into sounds (also known as the lexical route to reading). The third group of researchers argue that reading and writing must involve both a phonologically mediated subword translation process (for reading novel forms) and a whole-word recognition and production process (for reading words with regular spelling). The relative contribution of the phonological as opposed to the lexical routes to processing written forms hinges on the nature of the orthographic system used, i.e. whether the language uses deep orthography where there is no one-to-one grapheme-to-phoneme correspondence, or whether the orthography is shallow where the relationship is transparent (Bentin and Frost 1987). Since Arabic has two types of orthographies (deep when written without diacritics and shallow when all the vowels are indicated with diacritics), it can be used to test the claim that individuals with Deep Dyslexia will present more semantic errors when confronted with deep orthographies. Beland and Mimouni (2001) describe a bilingual Arabic/French patient (ZT) who showed typical symptoms of Deep Dyslexia, such as semantic errors and morphological errors, concreteness effect when reading aloud, and impossibility of reading nonwords. The distribution of ZT's errors was significantly different in the two languages he spoke, with more semantic errors in French and more translation errors in Arabic. In addition, Beland and Mimouni (2001) show that the percentage of correct responses was significantly higher when ZT read Arabic stimuli presented in deep orthography than in shallow orthography. They also found that semantic errors (a key feature of Deep Dyslexia) occurred with the same proportion in both of the patient's languages. Interestingly, the concreteness effect, function word effect, and lexicality effect were found to be the same in both of his languages, whereas grammatical class effect was found only in Arabic (ZT had more difficulty reading verbs than nouns and adjectives). Beland and Mimouni attribute the discrepancy to 'the heavy morphological load characterizing Arabic verbs'.

The study of language pathologies has over the decades provided (and continues to provide) insights not only into the units of grammatical knowledge but also into the nature of the representations of these units and the mental computations involved in their acquisition and use (comprehension and production).

BIBLIOGRAPHICAL REFERENCES

Abdalla, Fauzia. 2002. *Specific language impairment in Arabic-speaking children: Deficits and morphosyntax.* Ph.D. diss., McGill University.

Badecker, William and Alfonso Caramazza. 1989. "Lexical distinctions between inflection and derivation". *Linguistic Inquiry* 20.108–116.

Bates, Elizabeth and Brian MacWhinney. 1991. "Cross-linguistic research in aphasia: An overview". *Brain and Language* 41.123–148.

Bates, Elizabeth and Beverly Wulfeck. 1989. "Comparative aphasiology: A cross-linguistic approach to language breakdown". *Aphasiology* 3.111–142.

Beland, Renee and Zohra Mimouni. 2001. "Deep dyslexia in the two languages of an Arabic/French bilingual patient". *Cognition* 82.77–126.

Bentin, Shlomo and Ram Frost. 1987. "Processing lexical ambiguity and visual word recognition in a deep orthography". *Memory and Cognition* 15:1.15–23.

Bishop, Dorothy. 1992. "The underlying source of specific language impairment". *Journal of Child Psychology and Psychiatry* 33:1.3–66.

——. 1997. *Uncommon understanding.* Hove: Psychology Press.

Chomsky, Noam. 1980. "Rules and representations". *Behavioral and Brain Sciences* 3.1–61.

——. 1999. "On the nature, use, and acquisition of language". *The handbook of child language acquisition*, ed. William C. Ritchie and Tej K. Bhatia, 33–54. New York: Academic Press.

Clahsen, Harald. 1991. *Child language and developmental dysphasia: Linguistic studies of the acquisition of German.* Amsterdam and Philadelphia: J. Benjamins.

—— and Monika Rothweiler. 1993. "Inflectional rules in children's grammars". *Yearbook of morphology 1992*, ed. Geert Booij and Jaap van Marle, 225–288. Dordrecht: Foris.

Crystal, David and Rosemary Varley. 1998. *Introduction to language pathology.* 4th ed. London: Whurr.

Dinnsen, Daniel. 1993. "Underspecification in phonological disorders". *Principles and predictions: The analysis of natural language. Papers in honor of Gerald Sanders*, ed. Mushira Eid and Gregory Iverson, 287–304. Amsterdam and Philadelphia: J. Benjamins.

——. 1996a. "Context effects in the acquisition of fricatives". *Proceedings of the UBC international conference on phonological acquisition*, ed. Barbara Bernhandt, David Ingram and John Gilbert, 136–148. Somerville, Mass.: Cascadilla Press.

——. 1996b. "Context-sensitive underspecification and the acquisition of phonemic contrasts". *Journal of Child Language* 23.57–79.

—— and Mary Elbert. 1984. "On the relationship between phonology and learning". *Phonological theory and the misarticulating child* (= *ASHA Monographs*, 22), ed. Mary Elbert, Daniel Dinnsen, and Gary Weismer, 59–68. Rock-ville, Md.: American Speech-Language-Hearing Association.

Dyson, Alice T. and Mousa M. Amayreh. 2000. "Phonological errors and sound changes in Arabic-speaking children". *Clinical Linguistics and Phonetics* 14.79–109.

Ellis, Andrew. 1993. *Reading, writing and dyslexia.* Hove: Psychology Press.

Elman, Jeffrey, Elizabeth Bates, Mark Johnson, Annette Karmiloff-Smith, Domenico Parisi, and Kim Plunkett. 1996. *Rethinking innateness: A connectionist perspective on development.* Cambridge, Mass.: MIT Press.

Fodor, Jerry. 1983. *The modularity of mind.* Cambridge, Mass.: MIT Press.

Forrest, Karen, Gary Weismer, Megan M. Hodge, Daniel Dinnsen, and Mary Elbert. 1990. "Statistical analysis of word-initial /k/ and /t/ produced by normal and phonologically disordered children". *Clinical Linguistics and Phonetics* 4.327–340.

Friedmann, Na'ama. 2001. "Agrammatism and the psychological reality of the syntactic tree". *Journal of Psycholinguistic Research* 30:1.71–90.

Gardner, Howard. 1983. *Frames of mind: The theory of multiple intelligences.* New York: Basic Books.

Garrett, Merrill. 1980. "Levels of processing in sentence production". *Language production*, ed. Brian Butterworth, 177–220. London: Academic Press.

——. 1984. "The organization of processing structure for language production". *Biological perspectives on language*, ed. David Caplan, André Roch Lecours, and Alan Smith, 172–193. Cambridge, Mass.: MIT Press.

Goodglass, Harry and Edith Kaplan. 1972. *The assessment of aphasia and related disorders.* Philadelphia: Lea and Febiger.

——. 1983. *The assessment of aphasia and related disorders.* 2nd ed. Philadelphia: Lea and Febiger.

Gopnik, Myrna and Martha Crago. 1991. "Familial aggregation of a developmental language disorder". *Cognition* 39:1.1–50.

Hoffman, Paul R., Sheila Stager, and Ray G. Daniloff. 1983. "Perception and production of misarticulated /r/". *Journal of Speech and Hearing Disorders* 48.210–215.

Ingram, David. 1989. *Phonological disability in children.* 2nd ed. London: Cole and Whurr.

Jakobson, Roman. 1971. *Studies in child language and aphasia.* The Hague: Mouton.

Karmiloff-Smith, Annette. 1997. "Crucial differences between developmental cognitive neuroscience and adult neuropsychology". *Developmental Neuropsychology.* 13.513–524.

——. 2001. "Development itself is the key to understanding developmental disorders". *Language development: The essential readings*, ed. Michael Tomasello and Elizabeth Bates, 331–350. Oxford: Blackwell.

Lichtheim, Ludwig. 1885. "On aphasia". *Brain* 7.433–484.

Leonard, Laurence. 1998. *Children with specific language impairment.* Cambridge, Mass.: MIT Press.

——, Umberta Bertolini, Maria C. Caselli, Karla McGregor, and Letizia Sabbadini. 1992. "Two accounts of morphological deficits in children with specific language impairment". *Language Acquisition* 2:2.151–179.

Maxwell, Edith M. and Gary Weismer. 1982. "The contribution of phonological, acoustic and perceptual techniques to the characterisation of a misarticulating child's voice contrast for stops". *Applied Psycholinguistics* 3.29–43.

Miles, Tim R. 1993. *Dyslexia: The pattern of difficulties*. 2nd ed. London: Whurr.

Mimouni, Zohra, Renee Beland, S. Danault, and Ali Idrisi. 1995. "Similar language disorders in Arabic and French in an early bilingual aphasic patient". *Brain and Language* 51.132–134.

Mimouni, Zohra, Eva Kehayia, and Gonia Jarema. 1998. "The mental representation of singular and plural nouns in Algerian Arabic as revealed through auditory priming in agrammatic aphasic patients". *Brain and Language* 61.63–87.

McCarthy, John. 1981. "A prosodic theory of nonconcatenative morphology". *Linguistic Inquiry* 12.373–418.

——. 1982. "Prosodic templates, morphemic templates, and morphemic tiers". *The structure of phonological representations*, II, ed. Harry van der Hulst and Norval Smith, 191–223. Foris: Dordrecht.

Prunet, Jean-Francois, Renee Beland, and Ali Idrissi. 2000. "The mental representation of Semitic words". *Linguistic Inquiry* 31:4.609–648.

Rice, Mabel and Janna Otting.1993. "Morphological deficits of SLI children: Evaluation of number marking and agreement". *Journal of Speech and Hearing Research* 36.1246–1257.

Safi-Stagni, Sabah. 1991. "Agrammatism in Arabic". *Perspectives on Arabic linguistics*, III, ed. Bernard Comrie and Mushira Eid, 251–270. Amsterdam and Philadelphia: J. Benjamins.

——. 1992. *Normal and pathological breakdown in Arabic*. Ph.D. diss., Louisiana State University.

——. 1995. "Morphological structure and lexical processing: Evidence from Arabic". *Perspectives on Arabic linguistics*, VII, ed. Mushira Eid, 93–106. Amsterdam and Philadelphia: J. Benjamins.

Stoel-Gammon, Carol and Carla Dunn. 1985. *Normal and disordered phonology in children*. Baltimore: University Park Press.

SABAH M.Z. SAFI (King Abdulaziz University)

Language Policies and Language Planning

1. INTRODUCTION

The place of Arabs in the modern world has been determined to a large extent by the fact that, beginning around the end of World War I and well into the post-World War II period, the majority of the countries of the Arab world were under British or French mandate, in accordance with the Sykes-Picot agreement of 1916. The development plans of these countries and their emergent political, economic, administrative, and educational systems were established during the colonial period and modeled largely after the French and British systems. As a result of the colonial hegemony at the time, the indigenous Arab linguistic, religious, and cultural traditions "tended to be swept under the carpet" (Findlow 2001). Immediately following independence – mainly as a reaction to years of deliberate suppression or marginalization of their native language(s) and culture (Haeri 2000) – Arab countries adopted a general policy of Arabization in all domains of societal life, especially in administration, media, political discourse, physical environment, and education. The expressed goal of plans, policies, and decrees promoting Arabization was the reestablishment and regaining of an autonomous Arab identity. In practical terms, this Arabization policy involved the use of the mother tongue, Arabic, in dispensing education to Arab children; in communication, self-expression, and interaction in all societal functions and domains; in the naming of streets, shops, companies, and other objects of the physical environment; and in all public and private oral and written communication in administrative offices and departments (Shaaban and Ghaith 1999; Daoud 2002; Marley 2004; Mostari 2004).

The first Arabization campaigns did not make significant gains in Arab societies due to both lack of adequate planning and underestimation of the degree of entrenchment of the colonial systems. However, later campaigns managed to overcome many of the shortcomings of earlier efforts and to achieve a great degree of success in pre-university education, in the physical environment, and in public administration (Maamouri 1998; Ḥammūd 2000; Mostari 2004). Despite its gains and successes, the Arabization movement fell short of fulfilling its promises and was marred by some questionable practices regarding indigenous languages other than Arabic. Thus, it is clear to any outsider coming into the Arab world that the colonial language (French or English) still occupies a prominent place in daily communication, in the physical environment, and in education, especially in private business and educational institutions. Furthermore, speakers of indigenous languages other than Arabic, such as Berbers, Kurds, Assyrians, Nubians, and southern Sudanese, still feel that their language, even when tolerated in communication and education, is a liability rather than an asset in the job market, government, and higher education.

The advent and crystallization of the phenomenon of globalization during the last few years of the 20th century, however, has forced the Arab countries to revise their original plans, for a variety of reasons: in order not to be left behind, to survive the assault of external pressures, to create social harmony among all indigenous communities, and to maintain global competitiveness. The response to globalization and to the pressures applied on the Arabs after the 9/11 attacks in the United States came in the form of a reevaluation of educational systems and curricula that were partly blamed for the development and promotion of a culture of violence and intolerance in the Arab Middle East (Chugtai 2004; Dokainish 2004). In addition to implementing major changes in religion curricula, the Arab countries opted for reviving and strengthening foreign-language education, as well as for maintaining and expanding the use of minority languages (Marley 2004; Benrabah 2004). Furthermore, serious work is underway throughout the Arab world to join and take advantage of the fruits of the worldwide revolution in information and communication technology (ICT) through the medium of the Arabic language, in order to empower young Arabs with tools and attitudes that help them function and prosper in a diversified, multilingual, and multicultural world and to provide them access to a continuing stream of knowledge coming through the Internet and multimedia satellite systems.

The challenges facing the Arab world in the area of language planning and language policy in the age of globalization are too many and too complicated to be covered adequately in one study. Therefore, it is the intent of the present entry to focus on the policies of Arabization and the issues associated with them. More specifically, this entry first presents a brief overview of the current language situation in the Arab world. Then it presents, analyzes, and discusses policies of Arabization and the related issues of standardization and modernization of Arabic, the choice of medium of instruction in schools and universities (→ first language teaching). It will also address the place of minority languages in the lives of their speakers and in the societies in which they live.

2. The language situation in the Arab world

Language in the Arab world is characterized by → diglossia and → multilingualism. The original definition of diglossia centered on the existence of two main varieties of the Arabic language, what Ferguson (1959) called High (H), which refers to Classical Arabic (CA), known in the Arab world as *fuṣḥā*; and Low (L), which refers to the spoken varieties of Arabic. The concept has been modified and revised by many linguists to emerge as the existence of a standard-vernacular, or H-L, continuum with many variations and overlaps in between (Bentahila and Davis 1991; Parkinson 1992; Owens 2001). Multilingualism refers to the existence of other languages, in addition to the native language, in the speech communities that serve various communicative functions.

All Arab countries have decreed in their constitutions that Arabic is the only official language of these countries, but the sociolinguistic situation remains far from being dominated by a unitary language. The sociolinguistic profile, a concept suggested by Ferguson (1971), shows that the Arab world exhibits a great degree of language diversity, where Arabic and one major foreign language, usually French or English, serve as major languages; indigenous languages serve as minor languages; and a variety of other languages serve as special-purpose languages. In diglossic terms, the *fuṣḥā* and the main foreign language constitute the high varieties, and the vernaculars, indigenous languages, and special-purpose languages constitute the low varieties.

The *fuṣḥā*, in its Classical Arabic (CA) and Modern Standard Arabic (MSA) forms, is used in formal written and spoken communication. It is "the literary dialect which is used in the *Qurʾān*; in most publications including books, magazines and newspapers; and in formal spoken discourse, including prayer, television news broadcasts and formal prepared speeches" (Warschauer a.o. 2002). This variety is mutually intelligible among all literate, and many illiterate, Arabs from various countries. It exhibits slight variations in stress patterns and word choice, but its syntax and morphology are virtually the same. The vernacular, on the other hand, appears mostly as a spoken variety and "is used [though not exclusively] in conversation, songs, films, and television soap operas. As for written forms, it is used in comic strips and, occasionally, in novels and short stories" (Warschauer a.o. 2002). Researchers and scholars have also identified other varieties that lie somewhere between the *fuṣḥā* and the vernaculars. Blanc (1960), Shaaban (1978), Sayah (2002), Boukos (1995), and Zughoul (1980) speak of an 'educated' form of Arabic (→ Educated Arabic) that draws heavily on Modern Standard Arabic for vocabulary and morphology but whose syntax and phonology are heavily affected by the vernaculars.

The role of foreign languages in Arab societies remains crucial, and in many countries these languages continue to compete with the native language in many societal, educational, and economic communicative functions. English and French remain the major languages of instruction at the tertiary level in most professional and scientific specializations in most Arab countries. In Lebanon and in some North African countries, the foreign language is the medium of instruction for mathematics and sciences in the secondary cycle and in some instances even earlier, especially in private schools (Shaaban 1997; Daoud 2002; Marley 2004). In fact, in countries where French is the dominant foreign language, we find that bilingualism is turning into trilingualism as a result of the desire on the part of Arab governments to be part of the wave of globalization. Thus, "the dynamism and complexity of the [language] situation are further enhanced by the promotion of several foreign languages, especially English as the language of science and technology, international trade and electronic communication" (Daoud 2002:207). Even in countries such as Egypt, Syria, and Sudan, where Arabic is the main language of instruction at the university level, English plays a dominant role in the education process because it is the major language in which research is disseminated worldwide. It is important to note that many Arab Gulf countries such as Saudi Arabia, Qatar, the United Arab Emirates (UAE), and Bahrain are planning to start teaching English as a foreign language in public schools in the first grade. Similar plans are underway in Palestine (Amara 2003).

Another impetus for the spread of English among Arabs is the introduction of information and communication technology (ICT) into the Arab world; in fact, "emphasis on ICT has tended to privilege the use of English over Arabic" (Warschauer 2003).

Minority languages, which have suffered as a result of policies of Arabization originally enacted "with the apparent aim of creating a monolingual nation" (Marley 2004:25), have managed to survive in the local communities to which they belong, serving as the vernaculars alongside spoken Arabic. Recently, however, these languages have acquired respectability as a result of being allowed to be used in communication, mass media, and education.

3. ARABIZATION

The terms 'Arabization' and 'Arabicization' have been used interchangeably in the literature on language planning in the Arab world to refer to the declared *taʿrīb* policy of Arabic-speaking countries, which decrees that Arabic is the official national language of these countries. Although the term 'Arabicization' seems to be more accurate as a reference to the concept, as it involves the language and not the ethnic group, 'Arabization' is the term most frequently used in the literature. Needless to say, this definition excludes its use in Arabic linguistic terminology to refer to loanwords that enter the language through "total assimilation of foreign vocabulary into Arabic" (Abu-Absi 1984:113). It also excludes its reference to the assimilation of non-Arab ethnic minorities and imposition of the Arab identity on them.

The policy of Arabization came as a natural move on the part of the newly independent Arab governments after years of linguistic and cultural submission to foreign colonizers during which Arabic served as the rallying point of opposition to the hegemony of the colonizer. It was a move to restore to Arabic all its natural rights and communicative functions as the native language of the population. The national constitutions that were drawn up after independence in all Arab countries stipulate that Arabic is to be adopted as the only official language.

In the countries of the Maghreb, which were under French rule, independence brought with it the launching of major Arabization campaigns, whose purpose was to replace French with Arabic in the various public sectors, most notably in administration, mass media, the physical environment, and education. Arabic was viewed by leaders of the Maghreb countries as the symbol of identity and of the rich cultural heritage that they share with the Arab countries of the Middle East (Marley 2000).

In the Middle East, on the other hand, the nationalist movements in the newly independent countries saw Arabic as the major force in their drive for a pan-Arab political and economic entity extending from the Atlantic Ocean to the Arabian Gulf (Barakat 1993). The attempts to Arabize administration, mass media, political and economic discourse, the physical environment, and education achieved varying degrees of success for a variety of reasons, including inadequate preparation and planning; paucity of resources; resistance by the elite who owed their privileges, especially in administrative positions, to their high level of proficiency in the language of the colonizer; and the deep resentment of the minority groups, who believed that their languages and cultures would be marginalized.

In Algeria, where French rule lasted from 1830 to 1962, the process of Arabization proved particularly difficult. The French left behind a heavy legacy in which French was the principal language of administration, education, and the environment. The vernaculars, Algerian Arabic and → Berber, were nurtured at the expense of the *fuṣḥā*, and Algerian leaders and civil servants of independent Algeria did not have the ability to express themselves in the *fuṣḥā* (Abu-Absi 1984; Mostari 2004; Benrabah 2004). "Several laws, decrees and ordinances aimed at implementing Classical Arabic and strengthening its position in all public domains were duly enacted, reinforced and applied" (Mostari 2004:27). The initial campaign faced many difficulties in the area of education, due to a shortage of teachers of Arabic in the country and the high cost of importing teachers from the Middle East who had to teach as well as train local teachers; a lack of linguistically and culturally appropriate instructional materials; and resistance by Algerian Francophones who fought to retain French, claiming that Arabic was incapable of coping with the demands of technology and modernity.

The Algerian government had to relaunch new Arabization campaigns, trying to avoid the pitfalls of the original campaign. In 1968, President Boumediene issued a decree which stated that "within a period of three years, functionaries should learn enough Arabic to work in the language" (Grandguillaume 1991:3). Although the implementation of such a decree was not very effective, the fact that it represented a serious step in the direction of Arabizing the administration made it possible to introduce other measures, such as the law known as Generalization of the Use of the Arabic Language, which stipulated that by July 1998, all administrative, financial, technical, and social interaction and communication in public institutions should be fully in Arabic (Mostari 2004). The Arabization of the environment was also carried out at a fast pace in compliance with "Article 3 of the Circular of July 1976... [which] stated its intent to Arabize totally all the external and internal signage of public administrations and companies and absolutely forbid any inscription in a foreign language" (Mostari 2004:28). The campaign was successful, but it did not entirely change the face of the country because many private enterprises opted to keep French names in their signs.

The Arabization of education was the most painful, especially in the first few years, owing to an attempt to achieve too much in a very short time. The ambitious plans to Arabize all levels of education at the same time backfired, creating many logistical and educational problems and forcing the government to maintain the teaching of science and mathematics through the medium of French. In an attempt to remedy the situation, the Algerian government introduced a number of legislative reforms in the 1970s, such as the Arabization of all teacher-training centers. Furthermore, the government established the tradition of holding periodic national conferences to assess the Arabization process and to move it in new directions. It is important to mention that although Arabization at elementary and secondary levels was achieved, Arabization at the university level failed to take hold. The humanities (Arabic language and literature, history, Islamic philosophy, and education) have been totally Arabized; the social sciences (economics, business administration, journalism, law, political science, psychology, and sociology) have been Arabized only partially; and the technical and scientific subjects (computer science, mathematics, physics, chemistry, biology, medicine, and engineering) are still taught in French (Mostari 2004).

In Morocco, initial plans for total Arabization, adopted after independence in 1956, were abandoned in favor of a gradual approach in order to avoid the problems of scarcity of instructional materials, shortage of teachers, and unneeded disruption in the educational system (Abu-Absi 1984). Another reason for the change in the pace of Arabization was the fear that it could lead to a rise in Islamic fundamentalism, similar to what had happened in Algeria (Marley 2004). One of the major moves in the direction of Arabization was the establishment in 1961 of the National Institute of Arabization Studies and Research, whose function was to assess and support the Arabization process and to address any problems that might arise (Abu-Absi 1984). Ḥammūd (2000:99) divides Morocco's Arabization efforts into five major phases. The first phase, which started in 1958, launched the process without affecting the role of French; the second phase, 1968–1972, saw the completion of Arabization of the elementary cycle; the third phase, 1973–1977, involved the Arabization of social studies in the secondary school; the fourth phase, 1978–1980, witnessed intensive efforts to Arabize all secondary school subjects; and the fifth phase, 1980 and beyond, relaunched a campaign for Arabizing scientific subjects in the secondary school to achieve total Arabization of pre-university education and civil service. The Arabization of university education has not been more successful than it was in neighboring Algeria. French remains the main medium of instruction in all technical and scientific fields, and English is entering the picture as an additional language of professional fields. Establishment of the English-medium, American-style Al-Akhawayn University in Ifrane in 1995 gave English a big boost in the country.

Moroccan authorities realized in the 1990s (and maybe even earlier) that the Arabization policy, despite its importance for the creation of a unified Moroccan identity, was in fact ignoring the reality of the long-established multilingual structure of the country, especially the vital role of French in Moroccan society and the language rights of the Berbers, who constitute more than one-third of the population. Therefore, in 2000 the government issued the Charter for Educational Reform, which highlights three main themes: "the reinforcement and improvement of Arabic teaching, diversification of languages for teaching science and technology and openness to Tamazight" (Marley 2004:31). The charter also provides for the establishment of an Arab Language Academy charged with modernization of the language. It is obvious that the new policy represented by the charter recognizes that Arabization has not achieved its goals and that the Arabic language needs to be developed to become fit for the teaching of scientific and professional subjects. Furthermore, it opens the door for restoring the role of French, as well as the possible introduction of English in the teaching of science and technology.

In Tunisia, the campaign for Arabization followed a gradual methodical approach in order to ensure a smooth transition from French to Arabic as the language of teaching in Tunisian schools. Arabic was added, one or two classes at a time, with a hiatus from time to time to allow for a natural assimilation of the changes. The Arabization campaigns, started after independence from France in 1956, were accompanied by fluctuations in the fortunes of French, which was delayed to the fourth grade, then brought back to second and third grades, then dropped again. Arabization of sciences, mathematics, and technical education up to the ninth grade was achieved in 1997 (Daoud 2002). In the year 1999, the Tunisian government confirmed its policy of Arabizing administration and set the year 2000 as the deadline for achieving that goal. It also emphasized the need to Arabize computer software and databases in all public institutions, and established committees to work on preparing specialized dictionaries in all fields of knowledge (Daoud 2002). Tunisia was more successful than the other Maghreb countries in the area of preparing teachers and instructional materials in Arabic. In fact, the textbooks are research-based and follow a communicative, functional approach that differs radically from traditional approaches to teaching Arabic (Habib and Shaaban 1983). University education in technical and scientific fields is mainly conducted in French, with a significant move toward English as well.

In the countries of the Middle East, Arabic was seriously challenged by the Turks during the rule of the Ottoman Empire (1517–1918), when Turkish was introduced as the language of instruction in schools as well as the language of government. However, in the 19th century, especially the second half, there was a literary, cultural, and journalistic renaissance spearheaded by Egypt, Lebanon, and Syria as well as by recent immigrants from the eastern Mediterranean region to the Americas. This literary and educational awakening started with Napoleon's campaign in Egypt and Syria in 1799 and the subsequent follow-up on what this campaign had started by Egypt's Muhammad Ali Pasha, whose rule extended beyond Egypt to include, for some time, Syria and Lebanon. The position of Arabic was strengthened by the opening of missionary schools in Lebanon, Syria, and Palestine (Amara 2003). These missionaries represented different Christian churches in the West. Their schools taught their own respective languages in addition to Arabic, the native language of the land: French or Italian by Catholics; English by British and American Presbyterians, Anglicans, and Quakers; German by Lutherans; and Russian by the Orthodox. Strong competition among these schools resulted in strengthening Arabic, the language they all stressed in order to attract students to their schools. The competition was especially fierce between the Jesuits and the Presbyterians; it culminated in the establishment of the French-medium Saint Joseph University and the American University of Beirut (AUB), the two most prominent universities in Lebanon and the Middle East. At AUB, Arabic was used for a short while as a medium of instruction in all fields, even medicine.

The defeat of Turkey in World War I put an end to the challenge of the Turkish language. However, a more serious challenge for Arabic emerged in the form of French and English, after the Sykes-Picot agreement of 1916 between Britain and France placed Lebanon and Syria under French mandate and Egypt, Sudan, Iraq, and Palestine under British mandate. French colonial authorities introduced new systems of education and administration modeled after the French system, and decreed that French and Arabic would be the two official languages in Syria and Lebanon, to be used in the physical environment, education, and administration. In reality, however, French dominated in education, as it was used as the medium of instruction in social studies, sciences, and mathematics (Shaaban and Ghaith 1999). In Egypt, Palestine, and Sudan, the British introduced English as an official language alongside Arabic.

In the case of Palestine, Hebrew was also treated as an official language (Amara 2003). In education, English was introduced as a second language, but it was not made to replace Arabic in schools, although many private schools were established emphasizing the use of English as a medium of instruction in areas of scientific subject matter.

Upon achieving independence, all of the Arab Middle Eastern governments proclaimed Arabic as the only official national language of their countries. The Arabization process in Egypt, Iraq, and Syria took place smoothly. In Syria, the Syrian Arab Science Academy "was established in 1919 with the goal of developing badly needed instructional materials and adopting the necessary scientific and technical vocabulary" (Abu-Absi 1984:116); in Egypt, the Arabic Language Academy was established in 1932; in Iraq, the Iraqi Science Academy was founded in 1947; and in Jordan, the Jordanian Arabic Language Academy was founded in 1976" (Elkhafaifi 2002; → language academies). It is important to mention here that the Permanent Bureau for Coordination of Arabization in the Arab World was established in 1962 by the Arab Educational, Cultural, and Scientific Organization (ALECSO) in Rabat to help coordinate all the various efforts of coining new terminology, especially in scientific and technical fields. These Arabic-language planning agencies have set for themselves four major goals: "the regeneration of Arabic as an effective communication medium for modern science and technology[;]...the preservation of the purity of the language[;]...collecting, editing, recording, and restoring manuscripts to preserve classical works, and to reprint and publish them for modern use[; and encouraging] new works by Arab scholars, including translations of foreign works into Arabic" (Elkhafaifi 2002:255–256).

The process of Arabization is considered most successful in Syria, mainly because it has covered all levels of education, including university education and scientific and technical fields such as engineering and medicine. Abu-Absi (1984:116) considers that "the Syrian experiment has thus set an example to the rest of the Arab world by demonstrating that Arabic can indeed be used as a language of science and technology".

The same is true in some Egyptian universities. However, all Arab countries have recently started to emphasize the knowledge of English as a necessity for students of scientific and technical fields, a very rational move in the age of globalization, in which English is the language of over 80 percent of scientific and technical research.

In Sudan, Arabic is the official language and the main language of education in the North. In the south, and after the Addis Ababa accord between north and south, a bilingual system has been installed with Arabic and English, and in some cases a minority language, used as media of instruction in schools. The Sudanese government remains committed to the Arabization of university education, after having successfully Arabized pre-university education. In 1991, Sudan established the Higher Authority of Arabicization (HAA) in order to implement a policy stating that Arabic should be the language of teaching and scientific curricula at Sudan's institutions of higher education. The HAA was entrusted with overseeing the processes of language standardization, including coining and unifying terminology in physical and applied sciences at all universities; establishing an Arabic-language scientific library; encouraging translation and scientific publication in Arabic; convening conferences on Arabization in Sudan; and coordinating with ALECSO and other Arab agencies for language planning.

The Arab Gulf States (Saudi Arabia, Kuwait, UAE, Oman, Qatar, and Bahrain) have always looked upon the Arabic language as the symbol of their identity and Islamic faith. In fact, these countries assign more hours in their curricula to the teaching of Arabic and Islamic studies than the other Arab countries (→ first language teaching). Furthermore, Arabic is the language of instruction of all school subjects; international schools operating in these countries are forced to teach Arabic. In the UAE, for example, "all subjects taught at government schools use Arabic language as a medium of instruction except for the English language classes. In technical education, English is used for technical and specialized subjects, but the rest of the subjects are taught in Arabic.... Some private schools for which Arabic is not the medium of instruction are to teach Arabic language to non-speakers of Arabic" (UAE 1996). At the university level, Arabic is officially the medium of instruction for all subjects, but some courses are taught in English. However, during the last decade, many private and public English-medium universities have been established in the Gulf countries, and they usually offer professional and scientific specializations.

4. Issues associated with Arabization

After gaining their independence, Arab countries in the Middle East and North Africa alike adopted a policy of Arabization in order to promote the use of Arabic in all aspects of Arab life. The "purpose has been to provide a modern standard language that enhances the transfer and growth of knowledge among speakers whose mother tongue is (colloquial) Arabic... thus enhancing the efficiency of education, and strengthening the cultural consciousness in Arab societies" (Meiering 2004:2). However, although the motivation was noble, actual implementation has been fraught with problems that have slowed down and continue to slow down the pace of educational and linguistic reform. Some of these problems include the need for modernizing and standardizing the language; the relation between the Arabic language and Islam; and limitations in the job market for those who are fluent only in Arabic.

4.1 *Standardization and modernization of Arabic*

Language standardization involves the creation or borrowing of new vocabulary and expressions, especially → technical terminology; the use of language in mass media, public spheres, government offices, courts, and political institutions; the use of language as a medium of instruction in schools and universities; and its use in economic, scientific, and political spoken and written discourse. Modernization, or "creating contemporaneity", as Haeri (2003) calls it, refers to the introduction of linguistic and educational reforms needed to make Arabic a more effective language of communication in the modern world. In particular, the emphasis has been on providing the tools that would turn Arabic into a suitable vehicle for mass media, modern science, and technology and making it easy to learn by both children and adults. Unfortunately, despite all the efforts put into standardizing and modernizing the language, little has been achieved, and little consensus has been reached regarding the need for these reforms. Thus, the debates that started at the time of independence about the place of Arabic and the need to modernize it continue years after the establishment of language academies and the publication of all kinds of lexicons and dictionaries and the holding of so many conferences on the issues related to Arabization. The calls by Arab scholars to modernize the language "by avoiding older or archaic vocabulary, accepting foreign borrowings, use of shared cognates with the nonclassical varieties, avoiding some of its syntactic constructions, and so on, proved to be highly contentious", and they remain so today (Haeri 2000:72).

Maamouri (1998) identifies three schools of thought regarding language modernization in the Arab world. The first is highly resistant to any change, claiming that Arabic is adequate for the functions that its speakers demand of it. The second school is striving to simplify some of the language rules without deviating radically from the traditional prescriptive approach. The third school demands radical changes in the rules, patterns, and uses. Maamouri himself suggests a few reforms based on experiments carried out in Egypt, Lebanon, Tunisia, and elsewhere in the Arab region: the selection of 'common Arabic' terms in reference to Modern Standard Arabic (MSA), terms that have their counterparts in the vernaculars, or using what Frayḥa (1955) called "cultivated *fuṣḥā*", which highlight forms closely related to the various colloquial varieties; doing away with grammatical declensions; secularization of Arabic grammar by taking it away from the hands of religious institutions and establishments, as proposed by Ṭāhā Ḥusayn (1957); first establishing → literacy in colloquial Arabic whereby "the incorporation of colloquial Arabic into the oral classroom discourse will reduce miscommunication and improve learning" (Maamouri 1998:59); and "the use of the diacritical marks and the use of only one letter form [which] should help reduce the complexities of the Arabic reading process and improve Arabic reading results" (Maamouri 1998:63).

Proponents of linguistic reform have identified four major areas that need to be addressed: orthography, syntax, technical vocabulary, and lexical expansion. Abu-Absi (1984, 1986) has discussed in great detail specific issues pertaining to each of these four areas. In the area of orthography, the main complaints are that Arabic does not have clear representation for short vowels, and that there are different shapes for consonants depending on their place in the word. The introduction of diacritics and the different proposals for having one or two shapes for each letter have reduced the acuteness of this problem. However, the increasing use of the Arabic language on the → Internet has highlighted some of the problems that need to be addressed before such use becomes more effective and widespread. Diab states that "the Arabic language is being increasingly used on the internet despite significant obstacles.... One of the biggest problems is the issue of multiple character sets that represent Arabic" (2003:17).

The area of grammar, which in the Arabic educational context refers to morphology and syntax, has traditionally been the most difficult part of learning Arabic, mainly because it is the grammar of Classical Arabic. Because Classical Arabic is not part of everyday communication in the Arab world, learning Arabic grammar is much like learning the grammar of a foreign language, with one major difference: Arab teachers avoid using foreign-language methodology in order not to be accused of treating the 'native tongue' as a foreign language. Calls for the simplification of Arabic grammar have been echoed all over the Arab world. The following reforms have been suggested: "(a) reduction of the most exceptional cases for maximum consistency; (b) a simplification of the declensional system; (c) freedom from multiplicity of forms; and most importantly, (d) selection of uses which have their counterparts in the colloquials" (Maamouri 1998:56). Similar reforms were advocated by Frayḥa (1955), but without much success because of the stand of traditionalists against such changes and accusations that such reforms are intended to weaken the Arabic language.

In the area of vocabulary expansion – especially the development of words and expressions in technical and scientific fields – the various language academies and the Arabization offices have been rather active in developing dictionaries in various fields to allow Arabic to express new concepts in Arabic. Thus, since the mid-1990s, the Bureau of Coordination of Arabization in Rabat and ALECSO have been producing what they call unified dictionaries of terms in different fields, such as meteorology (1999), hydrology (2000), information (1999), pedagogical techniques (1999), vocational and technical endeavors (1996), human science (1997), and geology (2000). However, how many of these dictionaries are in active use in the hands of practitioners in these fields and how much use is made of them is debatable. One of the rare studies to address the issue of the use of Arabized words in written Arabic discourse is that of Al-Qahtani (2000), who examined a corpus of 1,068,236 words compiled from Saudi newspapers to check the extent of use of 288 Arabized words. The findings of his study indicate that the selected words occur with reasonable frequency, especially in scientific discourse, and that words coined by morphological derivation are used more frequently than words coined by compounding. He also found that around 15 percent of the words were used in their original English form rather than in their Arabized form.

Yet, despite extensive efforts exerted in the area of developing technical terminology, the controversy continues, with traditionalists calling for tapping the language's inner resources, citing its use in the past as a language of science and technology, and warning of the danger that these foreign words might bring with them linguistic elements incompatible with the Semitic characteristics of the language (Maamouri 1998: 55). Modernists, on the other hand, call for assimilating foreign vocabulary into Arabic. Their argument rests on the actual assimilation by Arabic speakers of foreign terms as the need arises for using them, as in the case of computer science, in which new terms are added at a fast pace. They argue that users cannot wait for the academies to coin new terms for loanwords that have already been in active use for some time.

Finally, in the area of lexicographic reform, efforts have been exerted to modernize the way dictionaries are compiled, both in content and form (→ lexicography: monolingual dictionaries). The traditional Arab method of listing words by their roots rather than alphabetically has proven to be confusing to beginning learners of Arabic, who may have trouble recognizing the prefixes, suffixes, and internal changes that a root might undergo. At present, one can find dictionaries that arrange words by roots and others that arrange them alphabetically (Abu-Absi 1984). Maamouri (1998) calls for the compilation of an Arabic thematic dictionary that could blend Modern Standard Arabic with the colloquial to help build the confidence of Arab children by allowing them to be more articulate when they speak about familiar objects and actions in familiar language. One important development in the area of dictionary making in Arabic is the attempt by some researchers to base the dictionaries on language use, rather than on archaic and made-up examples. Thus, a team of researchers at the Catholic University of Leuven in Belgium, funded by the Dutch Language Union, has collected a database corpus of 3,000,000 words from various authentic written and spoken text materials to use as a base in compiling Modern Standard Arabic-Dutch and Dutch-Modern Standard Arabic dictionaries (Van Mol and Paulussen 2001). The database "contains a completely elaborate learner's dictionary consisting of 19,000 Arabic words translated in context and more than 10,000 illustrative sentences.... It is the first Arabic dictionary which is based on extensive corpus analysis" (Van Mol and Paulussen 2001:1).

4.2 *Language of instruction*

Teaching through the medium of the mother tongue is the preferred practice in most countries of the world; it is advocated by the United Nations as vital for the preservation of culture as well as for facilitating the process of learning (United Nations Foundation for Endangered Languages 2002). In fact, educators and researchers believe that educating children in their native language is necessary not only for their cognitive development but also for their sense of belonging and national identity, for "language is not a neutral tool of communication but is everywhere implicated in the ways in which we read and write the world, the ways in which knowledge is produced and legitimated, the ways in which a human subject is constructed as a complex set of identities based on, amongst other things, race, class, gender, education, age, nationality, [religion]" (Hanks 1997:241). In education, it is believed that "when…[classroom] communication takes place in a language known to the student, the chances of achieving understanding are high compared to when it takes place in a language with which the student is not familiar" (Ejieh 2004:73).

Given these facts, it looks and sounds strange to have a debate on which language to use – the native language or the ex-colonizer's language – in the Arab world, whose language, Arabic, is at the core of Arab identity and has a rich literary heritage and successful experience in serving as a language of science. But the fact of the matter is that this debate still goes on, and the educational practices reflect this dichotomy of opinion. In private schools in most countries of the Arab world, the foreign language is used in the teaching of science, mathematics, and technology as of the fourth or seventh grade; in Lebanon this phenomenon occurs as of the first grade. Arabic, on the other hand, is used in the teaching of all other subject matter in these schools and in all subjects in public schools.

In most Arab countries, educational authorities claim that because of the need for modernization and technology and its fruits, they cannot afford to wait for academies to coin new terms; thus, they argue for retaining a major role for foreign language in education. Furthermore, most Arab citizens themselves believe it is important to get education through the medium of international languages in order to stay competitive in the age of globalization. Parents who, for economic or ideological reasons, send their children to Arabic-medium educational institutions remain uncomfortable with their decision, as it becomes obvious to them that their children do not have the same competitive edge in the job market as children who have had their education in English- or French-medium schools. Babault and Caitucoli (1997) state this is the case in most former colonies, where the struggle between the national language and the colonial language continues even after independence, as is the case with former French colonies in many parts of Africa.

The expanding role of foreign languages, especially English, is more pronounced in tertiary education, where most professional, scientific, and technological fields are taught in French or English or a combination of both. Furthermore, many American-style, English-medium universities that teach all specializations in English are being established in many Arab countries in the East and in North Africa, and many British, American, and Australian universities are opening regular or distance-education branches. Another example of the increasing importance of foreign languages has been the production of literature in French and English by Arab writers in an attempt to achieve international recognition. This phenomenon began with Middle Eastern immigrants to Europe and America, such as Gibran Khalil Gibran, George Chehadeh, and Amin Rihani, and has grown phenomenally in the last four decades, with names like Taher Ben Jalloun, Kateb Yacine, Rashid Boujedra, Salah Stetie, Ahdaf Soueif, and Amin Maalouf leading a long list of lesser-known writers. Marley (2000) observes that "the use of French [and English] is still widespread in certain areas of cultural production, most notably literature, and in fact is actually expanding for a number of reasons".

Researchers have recently noticed that a new utilitarian attitude is becoming prevalent in most Arab countries (Shaaban and Ghaith 1999; Ḥammūd 2000; Daoud 2002). Very few Arabists are fighting for Arabization; similarly, very few French-educated elitists are insisting on French. Parents seem to be looking for what gives their children an edge in a competitive world of globalization. So, although the Arabists insist on their children acquiring a good basis in Arabic, and the others insist on a good French education, both parties are espousing English as a language of education in Arab societies, especially at the tertiary level. In this context, educators have warned of the danger of promoting bilingualism or multilingualism without having proper planning and adequate resources. They believe such a scenario would result in a situation of semilingualism and code-switching as a result of the absence of good school programs, especially in public schools, for any of the languages involved (Baġdādī 1998; Bayḍūn 1998).

Another issue raised by researchers is the fear that school graduates educated in Arabic-medium schools would not have a solid base in foreign languages and thus would not be able to compete in the job market. In extreme cases, such graduates are turning to Islamic fundamentalism, as is the case in Algeria (Marley 2000).

Many of the problems that plague Arab policies regarding the place of the Arabic language in education nowadays go far beyond the choice of medium of instruction. The main problem resides in the low quality of education in general and the absence of good language education programs in particular. Despite the various reforms, the most prevalent mode of teaching remains the lecture format. Students are not active participants in their classrooms; they concentrate mainly on memorizing handouts or summaries. Classroom interaction is didactic, supported by a set of textbooks in which knowledge is factual and by a national examination system that tests memorization and factual recall. 'Study skills', 'autonomous learning', and 'critical thinking skills' are terms thrown around in new curricula, but methods that embody the terms have not made their way into the actual classroom.

4.3 *Minority languages*

Minority languages in Arab societies have been, up to the very recent past, rather marginalized. Despite marginalization, these languages have had enough ethnolinguistic vitality to remain alive in their communities, serving as the vernaculars alongside spoken Arabic. Recently, however, there have been radical changes in the Arab world in this area. Minorities have continued the struggle to have their languages and cultures recognized inside their countries; activists have carried their cause to the international level; and Western countries, along with the United Nations, have applied pressure in the last decade, issuing calls emphasizing language rights and the preservation of endangered languages. UNESCO director-general Kochiro Matsuura has stated that "favoring the promotion of linguistic diversity and the development of multilingual education from an early age helps preserve cultural diversity and the conditions for international understanding, tolerance and mutual respect" (United Nations Foundation for Endangered Languages 2002). Thus, in 2000, Morocco issued its Charter for Educational Reform, which recognized the language diversity in the country and allowed "local authorities to use Tamazight or 'any local dialect' in order to facilitate…learning" (Marley 2004:32). Similarly, in 2001, the Algerian government amended the constitution to make Tamazight a national language, thus recognizing the right of the Kabylians (Algerian Berbers) to use their language in schools (→ language shift: Amazigh). The Algerian government also allowed the use of the language in university education at the University of Tizi Ouzou (Mostari 2004). The language of the Kurds of northern Iraq is now used in schools alongside Arabic, and it is also used in university education for humanities and social sciences (→ Kurdish). In Lebanon, the Armenians have full rights to use their language at all levels of education (Cobban 1985).

It is not enough, however, to merely refrain from suppressing or neglecting minority indigenous languages in favor of the national language. Governments must nurture these languages. Skutnabb-Kangas (2000:312) believes that "unsupported coexistence" will not be enough and will most likely lead to the extinction of minority languages. Even when policies are enacted, it is difficult to guarantee that such policies will have an impact on language behavior (Bourhis 2001). Therefore, advocates of minority languages have repeatedly stressed that "demographically weak languages need firm policies in order to survive and thrive" (Romaine 2002:7). Arab countries have taken the first positive steps in recognizing the right of minorities to their language and culture, and they certainly need to consolidate and develop this new approach.

5. CONCLUSION

The first United Nations Arab Human Development report, issued by UNDP in 2002, identified the major problems facing human development in the Arab world: deficiencies in knowledge and freedom; lack of empowerment of women; high levels of illiteracy; and lack of readers and reading material. The second report issued in 2003 reports that the same problems persist and are becoming more serious. Further, the brain drain of scientists and other intellectuals to the West is increasing as a result of the absence of government support for scientific research, as well as a lack of an environment conducive to generating science.

The report draws a vision for building an Arab knowledge society through encouraging freedom of expression; good and just governance; and quality education with priority for early childhood education. It also calls for building a genuinely open general-knowledge base through "promoting the Arabic language; invoking the Arab knowledge heritage in forging the knowledge society model; enriching, enhancing and appreciating cultural diversity; and opening up to other cultures through stimulating Arabization and translation from other languages" (Fergani 2003:10).

The UNDP reports address the same issues discussed in this entry in relation to the Arabic language. The ailments of the Arabic language are caused by societies and governments that have neither succeeded in generating and building a general-knowledge base, nor managed to create a culture of freedom of expression, good governance, and diversity that would be conducive to intellectual productivity. It is not reasonable, for example, to expect Arabic to serve as a vehicle for science and technology in schools and universities if Arab researchers (and they are not to be blamed) publish their work in English or French. However, it needs to be stressed that Arabic has in the past served as a language of science par excellence and has supplied the West with the base for its modern science (Versteegh 1997); there is no reason to claim that it cannot serve that function again if its speakers use it as a vehicle of scientific experiments.

It is also unrealistic to expect that any Arab country can on its own succeed in achieving the needed level of Arabization in its education, administration, economic and commercial trade, and wider communication. A coherent pan-Arab plan for economic and human development is needed, and within it one can factor in Arabization (Daoud 2002). Furthermore, it needs to be stressed that Arabization campaigns, no matter how intense, cannot succeed if they produce university graduates who are unemployable because they are not fluent in a foreign tongue.

Despite the abundance of problems and challenges facing Arabic, the language will always rule strong and supreme in the Arab world, not to be replaced or surpassed by a vernacular or foreign language. The domains of its use could well increase, but they will not as long as there are no serious concerted Arab efforts in the areas of economic development. The source of the Arabic language's strength and survivability is multifold. Thus, Arabic has a cultural dimension as the main identity base for 280 million people who use it as an official language and another 140 million people who use it as a first language (al-Ḥannāš 2004). Second, the religious dimension of Arabic as the language of the *Qur'an* helps rally around it 900 million Muslims worldwide. Third, Arabic is seen as an aspect of the struggle with the West, amounting to an ideology for the Arabs in the face of "the other", as Jacques Berque says (Shaaban 1993). In brief, there is no fear that Arabic will become extinct or be taken over by the dialects, as happened with Latin, although it could be marginalized in some domains.

The need for language planning in the Arab world, which has always been emphasized by concerned Arab linguists, educators, journalists, and researchers (Shaaban 1990; Maamouri 1998; Haeri 2000; Šaṭaṭ 2002; Elkhafaifi 2002), is much greater now in light of the linguistic and cultural challenges of globalization. Globalization makes it imperative that every individual Arab country and the Arab world as a whole come up with clear policies for language education and societal language use before Arabic is swept by globalization that might obliterate its linguistic and cultural distinctiveness (Findlow 2001; Chugtai 2004). The main challenge lies in the need to produce policies that draw a three-way balance among the need to maintain and strengthen the mother tongue, Arabic, as a language of social, cultural, and educational communication as well as a symbol of identity and culture; the need to master at least one international language in order to stay abreast of technological and scientific development; and the need to give indigenous languages and cultures a voice. Such policies will ensure that each language or language variety operating in Arab societies has clear-cut, well-defined communicative functions that could overlap in some cases but should not be allowed to create tensions and conflicts that could upset societal harmony, productivity, and development.

BIBLIOGRAPHICAL REFERENCES

Abu-Absi, Samir. 1984. "Language planning and education in the Arab world". *International Education Journal* 1:2.113–132.

——. 1986. "The modernization of Arabic: Problems and prospects". *Anthropological Linguistics* 28:3.337–348.

ʾAbū ʿAsalī, Dalāl. 1998. "at-Taʿaddud al-luġawī ṭabaqī wa-lā ṭāʾifī”. *an-Nahār*, 13 August 1998, 14.

Al-Qahtani, Saad. 2000. *Arabization in written discourse in Saudi Arabia.* Ph.D. diss., Ohio State University.

Amara, Muhammad H. 2003. “Recent foreign language education policies in Palestine”. *Language Problems and Language Planning* 27:3.217–232.

Babault, Sophie and Claude Caitucoli. 1997. “Linguistic policy and education in Francophone countries”. *Encyclopedia of language and education.* I. *Language policy and political issues in education*, ed. Ruth Wodak and David Corson, 159–167. Dordrecht: Kluwer Academic Publishers.

Baġdādī, Nādya. 1998. “Luġat ‘as-sukkān al-ʾaṣliyyīn’”. *an-Nahār*, 13 August 1998, 14.

Barakat, Halim. 1993. *The Arab world: Society, culture, and state.* Los Angeles: University of California Press.

Bayḍūn, ʾAḥmad. 1998. “Maʿrifat al-luġāt fī bilādinā ʾāxiḏa bi-t-tadannī”. *an-Nahār*, 13 August 1998, 14.

Benrabah, Mohamed. 2004. “Language and politics in Algeria”. *Nationalism and Ethnic Politics* 10:1.59–78.

Bentahila, Abdelali and Eirlys E. Davis. 1991. “One, two, or many”. *Indian Journal of Applied Linguistics* 17.69–88.

Blanc, Haim. 1960. “Stylistic variation in spoken Arabic: A sample of interdialectal conversation”. *Contributions to Arabic linguistics*, ed. Charles A. Ferguson, 81–151. Cambridge, Mass.: Harvard University.

Boukous, Ahmad. 1995. *Société, langues et cultures au Maroc: Enjeux symboliques.* Rabat: Publications de la Faculté des Lettres et Sciences Humaines.

Bourhis, Richard Y. 2001. “Reversing language shift in Quebec”. *Can threatened languages be saved? Reversing language shift revisited: A 21st century perspective*, ed. Joshua A. Fishman, 101–141. Clevedon: Multilingual Matters.

Chugtai, Shaheen. 2004. “The role of language in Arab reforms”. *Aljazeera.net.*June 20. http://English.aljazeera.net/NR/exeres/53E1EE06–C2BC-47–B28B-D49D609B89E (accessed 25 October 2004).

Cobban, Helen. 1985. *The making of modern Lebanon.* London: Hutchinson.

Daoud, Mohamed. 2002. “Language policy and planning in Tunisia: Accommodating language rivalry”, 206–224. http://www.miis.edu/docs/langpolicy/ch14.pdf.

Diab, Hassan. 2003. “Standardization related to Arabic language use in ICT”. Paper presented to the Western Asian Preparatory Conference for the World Summit on the Information Society (WSIS) held in Beirut, 4–6 February 2003.

Dokainish, Mona. 2004. “Educational reform is inevitable in the Arab world”. *Gulf News Online.* February 18. http://www.gulf-news.com/Articles/print.asp/ArticleID=111321.

Ejieh, Michael U.C. 2004. “Attitudes of student teachers towards teaching in mother tongue in Nigerian primary schools: Implications for planning”. *Language, Culture and Curriculum* 17:1:73–81.

Elkhafaifi, Hussein, M. 2002. “Arabic language planning in the age of globalization”. *Language Problems and Language Planning* 26:3.253–269.

Fergani, Nader. 2003. “Second Arab human development report: The need for a knowledge society”. *Newsletter of the Economic Research Forum for the Arab Countries, Iran and Turkey* 10:4.10.

Ferguson, Charles A. 1959. “Diglossia”. *Word* 15.325–340.

——. 1971. “National sociolinguistic profiles”. *Language structure and language use: Essays by Charles Ferguson*, ed. Anwar Dil, 157–184. Stanford, Calif.: Stanford University Press.

Findlow, Sally. 2001. “Global and local tensions in an Arab Gulf state: Conflicting values in UAE higher education”. Paper presented at the international conference “Traveling Policy/Local Space: Globalization, Identities, and Education Policy in Europe”, Keele University, 27–29 June.

Frayḥa, ʾAnīs. 1955. *Naḥwa ʿarabiyya muyassara.* Beirut: Dār aṯ-Ṯaqāfa.

Grandguillaume, Gilbert. 1991. “Arabisation et langues maternelles dans le contexte national au Maghreb”. *International Journal of the Sociology of Language* 87.45–54.

Ḥabīb, Mūnā and Qāsim Šaʿbān. 1983. *Tadrīs al-luġa al-ʿarabiyya fī l-marḥala al-ibtidāʾiyya fī l-bilād al-ʿarabiyya.* Beirut: Dār al-Kitāb al-Lubnānī.

Haeri, Niloofar. 2000. “Form and ideology: Arabic sociolinguistics and beyond”. *Annual Review of Anthropology* 29.61–87.

——. 2003. *Sacred language, ordinary people: Dilemmas of culture and politics in Egypt.* New York: Palgrave Macmillan.

Ḥammūd, Ṣalāḥ ad-Dīn. 2000. “Taʿrīb at-taʿlīm wa-s-siyāsa al-luġawiyya fī l-Mamlaka al-Maġribiyya bayna l-māḍī wa-l-ḥāḍir”. Shaaban (2000:91–120).

Hanks, Hilary. 1997. “Teaching language and power”. *Encyclopedia of language and education.* I. *Language policy and political issues in education*, ed. Ruth Wodak and David Corson, 241–251. Dordrecht: Kluwer Academic Publishers.

Ḥannāš, Muḥammad al-. 2004. “ʾAsbāb ḍuʿf al-luġa al-ʿarabiyya ʿinda ṭ-ṭullāb al-ʿArab”. *al-Xalīj*, cultural supplement.

Ḥusayn, Ṭāhā. 1957. “al-Luġa al-fuṣḥā wa-taʿlīm aš-šaʿb”. *Majalla al-Majmaʿ al-ʿIlmī al-ʿArabī* (Damascus) 32.44–56.

Maamouri, Mohamed. 1998. “Language education and human development: Arabic diglossia and its impact on the quality of education in the Arab region”. Paper presented at the Mediterranean Development Forum held in Marrakech, September 3–6.

Marley, Dawn. 2000. “Interactions between French and Islamic cultures in the Maghreb”. http://www.surrey.ac.uk/LIS/MNP/may2000/Marley.html (accessed 6 December 2004).

——. 2004. “Language attitudes in Morocco following recent changes in language policy”. *Language Policy* 3:1.25–46.

Meiering, Gregor. 2004. *Lost or found in translation: Translations’ support policies in the Arab world.* Study conducted by Thalassa Consulting for the Next Page Foundation.

Mostari, Hind A. 2004. “A sociolinguistic perspective on Arabization and language use in Algeria". *Language Problems and Language Planning* 28:1.25–43.

Owens, Jonathan. 2001. "Arabic sociolinguistics". *Arabica* 48.419–469.

Parkinson, Dilworth. 1992. "Good Arabic: Ability and ideology in the Egyptian Arabic speech community". *Language Research* 28.225–253.

Romaine, Suzanne. 2002. "The impact of language policy on endangered languages". *International Journal on Multicultural Societies* 4:2.194–212.

Sayah, Mansour. 2002. "Linguistic issues and policies in Tunisia". *French in and out of France: Language policies, intercultural antagonisms, and dialogue*, ed. Kamal Salhi, 411–431. Bern: P. Lang.

Shaaban, Kassim [Qāsim Šaʿbān]. 1978. "Code-switching in the speech of educated Arabs". *Journal of the Linguistic Association of the Southwest* 3:1.7–19.

——. 1990. "A sociolinguistic analysis of language-related issues in the Arab world". *al-Abhath* 38.17–35.

——. 1993. "Luġat at-taʿlīm wa-l-waḥda al-waṭaniyya". *al-Abḥāth* 41.123–159.

——. 1997. "Bilingual education in Lebanon". *Encyclopedia of language and education*. V. *Bilingual education*, eds. Jim Cummins and David Corson. Dordrecht, Boston, and London: Kluwer Academic Publishers.

——. (ed.). 2000. *al-Luġa wa-t-taʿlīm*. Beirut: al-Hay'a al-Lubnāniyya li-l-ʿUlūm at-Tarbawiyya.

—— and Ghazi Ghaith. 1999. "Lebanon's language-in-education policies: From bilingualism to trilingualism". *Language Problems and Language Planning* 23:1.1–16.

Šaṭaṭ, Yūnis. 2002. "Taʿlīm al-luġa wa-l-ʿawlama". *al-Xalīj*, 2 March 2002, 7.

Skutnabb-Kangas, Tove. 2000. *Linguistic genocide in education – or worldwide diversity and human rights*. Majhwah, N.J.: Erlbaum. http://www.ibe.unesco.org/International/Databanks? Dossiers/ruae. htm (accessed 25 October 2004).

United Arab Emirates. 1996. *National report of the United Arab Emirates on the development of education from 1993/1994–1995/1996*. National report presented at the 45th session of the International Conference on Education, Geneva, 30 September–5 October 1996.

United Nations Foundation for Endangered Languages. July 2002. *Newsletter 19*. Paris: UN Office for the Coordination of Humanitarian Affairs. http://www.ogmios.org/193.htm.

Van Mol, Mark and Hans Paulussen. 2001. "AraLat: A relational database for the development of bilingual Arabic dictionaries". *Proceedings of Asialex 2001, Asian Bilingualism and the Dictionary, Seoul, August 2001*, 206–211. Seoul.

Versteegh, Kees. 1997. *The Arabic language*. Edinburgh: Edinburgh University Press.

Warschauer, Mark. 2003. "The allures and illusions of modernity: Technology and educational reform in Egypt". *Education Policy Analysis Archives* 11(38). October 17. Arizona State University. http://epaa.asu.edu/epaa/v11n38/ (accessed 4 January 2005).

——, Ghada R. El Said, and Ayman Zohry. 2002. "Language choice online: Globalization and identity in Egypt". *Journal of Computer-Mediated Communication* 7:4. Indiana University. http://jcmc.indiana.edu/vol7/issue4/warschauer.html.

Zughoul, Muhammad Raji [Muḥammad Rājī Zuġūl]. 1980. "Diglossia in Arabic: Investigating solutions". *Anthropological Linguistics* 22.201–217.

——. 2000. "Izdiwājiyyat al-luġa: Ṭabīʿatuhā wa-muškilātuhā fī siyāq at-taʿlīm". Shaaban (2000:16–89).

Kassim Shaaban (American University of Beirut)

Language Shift: Amazigh

1. INTRODUCTION

The term 'language shift' refers to the change from the habitual use of one language to that of another. This implies that a community gives up a language in favor of another one. When shift has taken place, members of the community are said to have collectively chosen a new language instead of their native one (Fasold 1984). This is, in fact, a common result of extensive → language contact, occurring typically where there is a sharp difference in prestige and levels of official support for the languages concerned. Language shift is not the only possible outcome of language contact; in some circumstances, communities or groups are motivated to maintain their language (see below, Sec. 4).

This entry analyzes language shift and/or maintenance in three North African countries, Algeria, Morocco, and Tunisia, with a focus on the Moroccan situation, where the largest Amazigh population is found. The reference to the North African context is motivated by the features that the three communities share, especially at the historical and linguistic levels. They are Muslim states and former French colonies that have resorted to the process of 'Arabization' as a means to put an end to cultural colonization.

Language shift in the North African context in general and the Moroccan one in particular appears in the gradual disappearance of the Amazigh language from communities where it was widely spoken until the 1970s. The members of the Amazigh community seem to have collectively chosen an outside language (i.e. the colloquial or dialectal Arabic spoken in Algeria, Morocco, and Tunisia) where their native language, Amazigh, was formerly used.

This shift is the outcome of many factors, including language contact and bilingualism. In this context, Amazigh is shifted from because it is the language of the sociological minority group. The present entry explores the situation of Amazigh as a minority language in its contact with dialectal Arabic, a nonprestigious language when compared to Standard Arabic and French, but a more prestigious language when compared to Amazigh. The aim is to reveal the degree of shift the Amazigh language has undergone and to assess the outcome of the contact between Amazigh and dialectal Arabic, the two national mother tongues in North Africa. The objective is to determine the factors leading to shift, as well as those acting in favor of language maintenance in a community where orality predominates.

Given the nature of the topic, namely the outcome of the contact between the two ethnic groups and particularly the interplay between their languages, colloquial Arabic and Amazigh, a brief overview of the general linguistic situation in North Africa is needed to provide a sociolinguistic profile of North Africa, its language configurations, and the status of both Amazigh and dialectal Arabic in these countries. Although it is very hard to extend and generalize the facts in one country to the others, since each society has its own particularities, the overall situation is, to a great extent, similar.

2. The linguistic situation in North Africa: A brief overview

Amazighs, formerly referred to as Berbers, are the oldest known inhabitants of North Africa. Their language is also claimed to be the most ancient in the Maghreb. While the three countries (Algeria, Morocco, and Tunisia) share a number of features, each has its own particularities. Whereas Amazigh is moving toward extinction in Tunisia, it is fighting for revival and maintenance in Algeria and Morocco.

Algerian, Moroccan, and Tunisian varieties of Arabic are referred to as colloquial or dialectal Arabic. These varieties are the mother tongues of the large majority of the population in Tunisia (about 99 percent of whom are considered to be Arabs and speak Tunisian Arabic as their first language, while fewer than 1 percent of the population speak Amazigh), and of a great number of both Algerians and Moroccans, even those people who are of Amazigh origin. Algerian, Moroccan, and Tunisian Arabs use it in their everyday conversations. It can be considered to be the lingua franca (i.e. in areas of intensive language contact, a language adopted by speakers of different speech communities as their common medium of communication) of these three countries because it is used not only between the Arabs and Amazighs but also among Amazigh people speaking Amazigh varieties that are not mutually intelligible. It is originally an unwritten and noncodified language, standing in a diglossic relationship with Classical Arabic (Ferguson 1959; → diglossia), for they have different functions and a different status. Classical Arabic is the prestigious or 'High' variety, while colloquial Arabic is the 'Low' variety, not socially valued. Colloquial or dialectal Arabic is a mother tongue, the language of the home and intimacy and the vehicle of oral popular literature. It represents the most widely used language form in North Africa, both as concerns the number of speakers and the areas covered. Classical Arabic is learned only in a formal educational context. It is used for literature, newspapers, broadcasting, and religious ceremonies.

Colloquial Arabic is stigmatized because it is not codified. An individual who speaks only Algerian, Moroccan, or Tunisian Arabic is illiterate because these varieties are not taught at school. The prevailing attitude is that dialectal Arabic is a corrupt or incorrect form of Arabic, useless and inappropriate in important matters. For religious and historical reasons, the written medium is almost exclusively reserved for Classical Arabic.

The differences between Classical Arabic and dialectal Arabic concern all levels (lexical, phonological, morphological, syntactic, and semantic), to the degree that uneducated speakers of Algerian, Moroccan, and Tunisian Arabic cannot understand Classical Arabic, while fluent Classical Arabic speakers need some time to understand dialectal Arabic. The contrast between the two varieties is accentuated by the prestige of Classical Arabic as a sacred language, the language of the *Qur'ān*, used by God to give His message to the Prophet.

Amazigh is also known as → Berber. This name derives from the Latin term *barbarus*, which designated uncultivated people who lived outside the civilized world. The Amazigh people refer to their native language as Tamazight (*tamaziġt* or *ṯmaziġṯ*) and to themselves as *Imaziġen* (sg. *Amaziġ*). Most Amazigh researchers, especially in Morocco, have abandoned the term 'Berber' because of its negative connotations and adopted the term 'Amazigh' instead.

It is very difficult to estimate the number or percentage of Amazighs in North Africa, especially in the urban centers where the language is in clear regression because it is in strong competition with dialectal Arabic. There has never been any census for determining the exact figures of the Amazighs (or the Imazighen, to use the correct plural form) in any of the three countries because this has always been a very sensitive issue. Yet, the number is estimated to be about 40 percent of the overall population in Morocco, between 20 and 25 percent in Algeria, and less than 0.5 percent for Tunisia. Morocco has the largest Amazigh community. It also represents the largest number of Amazigh varieties. The discontinuous geographical distribution of Amazigh not only in Morocco but also in Algeria and Tunisia reflects the effect of rural exodus that North Africa has undergone especially since its independence from the French colonization. Today, Amazigh is spoken in various urban centers of the Maghreb such as Casablanca, Rabat, Agadir, and Nador in Morocco, and Algiers, Oran, and Tizi-Ouzzou in Algeria. In Tunisia, however, Amazigh is spoken only by a handful of people in Guellala on the island of Jerba and in a few southern villages in the regions of Mednine and Tataouine. In some remote and isolated areas, it is still possible to find monolingual Amazighs, mainly among women and children, especially those under the school age of seven years.

Amazigh is a spoken language with practically no written history. The Touareg dialects do have a writing system called Tifinagh. There is, however, no evidence that it has been used extensively, for no texts or documents exist. Nowadays, the language is essentially reserved for the domains of the home, family, and friends. The Amazigh literature is transmitted via oral tradition. Thus, the various Amazigh varieties are the vehicles of cultural traditions specific to the areas where they are spoken. With the growing interest in the Amazigh language and identity and the development of research in Morocco and Algeria, both the Arabic and the Latin alphabets have been used to write Amazigh. Recently, however, an adapted form of Tifinagh has been adopted as the official Amazigh alphabet by the Moroccan Royal Institute of the Amazigh Culture (IRCAM). This Amazigh alphabet has been acknowledged by the International Standardization Organization (ISO-UNICODE).

The areas where Amazigh is spoken are not continuous but are broken up into 'islands' surrounded by Arabic-speaking populations. The boundaries between the different areas are not well known because of the lack of a linguistic atlas. This discontinuity has brought the populations and their languages into contact. Given the inequality of the two languages and their speech communities, the majority language (i.e. dialectal Arabic) imposed itself as the language of intercommunication. Amazigh people had to learn dialectal Arabic for communication needs. Progressively, Arabic has gained in prestige and status among the Amazigh speech communities in North Africa. The Amazigh language has progressively lost its usefulness and a language shift is in progress in the three North African Amazigh communities. Amazighs learn Arabic not only for instrumental reasons, such as to achieve socioeconomic mobility, but also for integrative reasons, such as to become part of the Muslim community. Arabic is thought to be essential for understanding the *Qurʾān* and all the facets of Islam.

The shift from Amazigh to Arabic has been going on for centuries. It has been speeded up by recent sociopolitical developments. Many factors have, in fact, encouraged the spread of dialectal Arabic. Arabic/Amazigh bilingualism is progressively leading to the abandonment of Amazigh and the adoption of Arabic.

3. Language shift and language maintenance

There are a number of indications that a shift is in progress in the North African countries of Algeria, Morocco, and Tunisia. The extensive use of Algerian, Moroccan, or Tunisian Arabic in domains previously restricted to the Amazigh language is the most important evidence. Many factors have contributed either directly or indirectly to the weakening of the

Amazighs' linguistic and communicative competence, especially in contact areas. Amazighs have been submitted to a linguistic and cultural assimilation or accommodation process due to the stigmatization and exclusion of their native language and identity. Yet, recently language awareness and a cultural-identity consciousness have emerged with the aim to protect, maintain, and revive the Amazigh language, culture, and identity. The discussion is focused on causes of language shift as well as forces leading to maintenance or revival.

The factors that seem to have played – and are still playing – a role in the Amazigh language shift in North Africa in general and Morocco in particular are more or less the same as those associated with language shift in the literature. They include bilingualism or → multilingualism, language transmission and language attitudes, urbanization, migration, as well as intermarriage, the school system, and the media.

The situation of continuous contact between Amazigh and dialectal Arabic in North Africa makes it an ideal condition for the phenomenon of shift, although it is not a sufficient one. Most Amazighs are bilingual. The few monolingual people are those living in remote and isolated areas. A large number of the Amazigh people in Algeria, Morocco, and Tunisia had, in fact, to go through different linguistic stages. They moved from the exclusive use of their native language (i.e. Amazigh monolingualism) to the exclusive use of colloquial or dialectal Arabic (i.e. colloquial Arabic monolingualism). This was motivated by a number of factors such as migration, the school system, and the media.

Within the Moroccan context, El Kirat (2004) undertook extensive fieldwork to determine the degree of bilingualism among Amazigh speakers. The research showed that fluent bilingual speakers are in general past the age of 40. Younger people can still be fluent bilinguals depending on the area where they are (an urban contact area or a rural isolated area), the degree of contact they have with the language, and their attitudes toward it. Some people beyond forty can have low competence in the language given the lack of contact, diminished opportunities to use the language, and their attitudes toward it. People below the age of twenty have, in general, a reduced proficiency and a low degree of bilingualism. Unlike Morocco and even Algeria, however, no monolingual Amazigh speakers remain in Tunisia (Battenburg 1999). Pencheon (1968) reports the existence of a small number of elderly monolingual Amazigh females. Today all Amazighs (except for some preschool children) are at least bilingual, and many also speak other languages, including French (Battenburg 1999).

The degree of active bilingualism (i.e. the active use of more than one language in everyday interaction) is high only among the older generation (i.e. people aged fifty and above). It is very unstable among people aged between twenty and forty, and nearly absent among young people. The low degree of active bilingualism is mostly due to the reduction of the domains of use of the language and the decrease of the community's competence in its mother tongue. This is, indeed, clear evidence that a rapid shift is in progress among the young Amazigh generations.

The intergenerational dimension of language shift in Algeria, Morocco, and Tunisia should be emphasized. Language shift in this context should be seen as an intergenerational process rather than an individual phenomenon. An unmistakable sign of Amazigh language shift is the bilingual parents' transmission of only one language to their children, in this case Algerian, Moroccan, or Tunisian Arabic. As a result, the whole community is shifting to the use of dialectal Arabic because of the feeling that Amazigh, the language being shifted from, is not a prestigious language. The deliberate non-transmission of the Amazigh language to the younger generations is the main symptom not only that a language shift is in progress but also that the language is facing loss and even death.

Language shift in the North African bilingual community appears in the unstable form of bilingualism. Stability does not mean that the degree of use of the two languages, in this case Amazigh and dialectal Arabic, should be the same, but rather that the functional distribution should remain intact. The Amazigh community provides evidence for Boyd's (1986) claim that great domain overlap is a clear indication of language shift. As the domains of use of the Amazigh bilingual speakers overlap, the prospects for the maintenance of the Amazigh minority language are bleak. The domains of use of dialectal Arabic and Amazigh are so blurred that it is very difficult to establish any separation. The collective shift of the Amazigh community to a language other than its native language, dialectal Arabic here, has led – and is still leading – to the failure by some members of the Amazigh community to acquire the language or to partial acquisition.

The pattern of language use between the generations is also indicative of a language shift in progress, since larger proportions of older people use the language than the younger ones. This is mainly due to the fact that the use of Amazigh by the young people is vertical rather than horizontal (i.e., Amazigh is not the means of communication between the members of the same generation of young people). Interaction with age peers takes place in dialectal Arabic, while use of Amazigh is confined to interaction with the older generation.

At this stage, the degree of language shift in most of the Amazigh communities in North Africa has reached an advanced level. The members of the Amazigh community almost unavoidably shift to the use of dialectal Arabic, not only in their contacts and interactions with the dominant group (i.e. the Arab group), but also in their interactions within the Amazigh group. Indeed, if only interethnic shift had occurred in the Amazigh community, the situation would have evolved toward a form of stable bilingualism: Amazigh would have been used for communication within the group while Arabic would have been the language of communication in the other instances. What has led to language shift and even → language loss is the intra-ethnic shift that is widely spreading in the community. The choice of dialectal Arabic for communication within the group is a new form of behavior. This is central to the language shift and maintenance issue, for as long as there is an Amazigh minority group and as long as the minority group is not demographically broken up, the use of the Amazigh language should not disappear unless the norms for language use within the group are changed. The argument here is that the intra-ethnic shift has destabilized the situation to the extent that bilingualism has disappeared, or is disappearing, in the North African context in general and the Moroccan community in particular.

In the fieldwork in the Moroccan context (El Kirat 2001), people from different age groups and generations spoke about the phenomenon of shift and the lack of transmission of the Amazigh language in the community. A large number expressed no regret about the shift and loss of their native language. They all seemed to consider its nontransmission to children as an efficient solution to school failure. They also insisted on the disadvantages of not knowing Arabic. Many parents referred to their personal experience with the Amazigh language and how much they have psychologically endured. Many of them admitted that they have consciously and deliberately not transmitted their mother tongue to their children in order to spare them all the psychological pressures they went through themselves. They also insisted that Amazigh is of no use nowadays and that it represents a stigma that sets them apart and contributes to their estrangement from the dominant group, i.e. the Arabs. They also consider it a socially discrediting marker, a minus, and a handicap. Most of the informants pointed to the effects of school in lowering the value of the Amazigh language. Examples of cases of language shift and non-transmission in the North African community are found in abundance.

Urban life is not favorable to Amazigh because of the presence of another ethnic and linguistic group whose language is more highly valued and more widespread than Amazigh. This has encouraged Amazigh/dialectal Arabic bilingualism, which is, however, present only on the Amazigh side (i.e. unidirectional), because the Arabs have not felt the need to learn Amazigh.

The use of dialectal Arabic increased in North Africa with the increase of contact between the two ethnic groups. Arabic is used in all the domains of interaction between the two communities for communicative purposes because it imposes itself as the language of daily activities in the urban centers. This has been encouraged by the low official status of the Amazigh language and its institutional exclusion, as well as by intermarriage patterns. The two groups (i.e. the Amazighs and Arabs) have become so close that intermarriage between them is a very common phenomenon, invariably resulting in a failure to transmit the Amazigh language. It does not matter who is Amazigh: the father or the mother. Indeed, most of the communities where Amazigh is maintained (e.g. Souss and Rif for Morocco, Kabyle for Algeria) are opposed to intermarriage and marry mostly within their ethnic group.

The hard conditions of life in most Amazigh mountainous and rural areas, together with demographic pressures, have forced the Amazighs to migrate to the closest urban centers and even to some European countries (→ Europe). The 1960s were the starting point for an important wave of emigration toward the industrial European countries (mainly France and Italy) in search of work for the survival of the whole group. As it was easy then to find jobs because these countries needed cheap labor, all the male members of the same family used to emigrate, leaving behind only the elderly, the women, and the children. The men would soon return, either to buy a house in the city for the family, or to take their family members (i.e. wife and children) back with them, leaving behind in the homeland the elderly and those with no other alternatives. The dispersal of the Amazigh community together with its abandonment of the homeland made its language and culture easy prey to shift and loss, especially in the urban environment, which is the ideal setting for such a process.

The influence of dialectal Arabic as the dominant language (i.e. the most widely used) has increased exponentially with better communication and transport, and access to radio, television, and videos, which has extended to include even the remote areas. This has not only increased exposure to Arabic as a language and a culture, but also has enabled it to become the language of daily life.

The media is relevant here in terms of language broadcasts. It facilitates access to Arabic even in remote areas where there is no contact between the two communities. The Amazigh language has for a long time been excluded from the media and especially television. In Morocco, for instance, only very short programs in some varieties of Amazigh, mainly Tarifit, Tashelhit, and Tamazight, are broadcast. Until now, and even after the creation of the Royal Institute for the Amazigh Culture, there are still no Amazigh radio or television stations as such. The written press is available only for some Amazigh varieties, mainly in the Souss (i.e. southern Morocco). This is mainly due to the marginalization and exclusion of Amazigh and the lack of institutional support. In Algeria, the situation is a bit more favorable. Yet, this concerns mainly the Kabyle area. The school system has for long been an instrument of language suppression. In Morocco, Amazigh had no place in that system until September 2003. In Algeria, the regime created the Haut Commissariat à l'Amazighité (High Commission for Berber Identity) in 1995, describing Tamazight as a heritage language and promising to allow its use in education and the media. By 1997, the universities of both Tizi Ouzou and Bejaïa were offering Algeria's first-ever degree courses in Amazigh language and culture, and Algerian television offered a short daily newscast in Amazigh. In Tunisia, the language has always been completely excluded from both education and the media.

The educational system's exclusion of the Amazigh language has transmitted an indirect message to the Amazigh people about the relative worth of their mother tongue. School support for languages such as French and English has been taken to mean that these languages, spoken hundreds, if not thousands, of miles away, are worth the time and effort, while Amazigh is not worthy of any expenditure. The school can be seen to have played a negative role with regard to Amazigh at three levels: it has excluded Amazigh as a medium of communication; it has excluded Amazigh from the curriculum; and, in taking these actions, it has transmitted to the community at large and the Amazigh community in particular a low assessment of the value and utility of the Amazigh language. School is not singled out here to be held responsible for the stigmatization and derogation of Amazigh but rather to show that as a national institution in terms of the policies it sets and the practices and attitudes it adopts, it reflects the national rejection of Amazigh. Schools have a more persistent and penetrating effect in remote areas than most of the other institutions not only because they reach far more people directly but also because their impact is greater as it begins so early in the individual's life. Although Amazigh was not outlawed as a communicative medium within the school setting, as was the case for Gaelic in the United Kingdom or Breton in France, it was implicitly excluded by sociolinguistic rules existing in the community and assimilated by every Amazigh child before arriving at school. Although the Amazigh language has been introduced in the educational system – at the university level for Algeria and the primary school level for Morocco – this in no way means that the situation has changed and that the above facts do not hold anymore. Yet, it will surely reinforce the language awareness and cultural identity consciousness that have emerged recently, especially among the educated elite and young people.

Assimilation is the factor most often used by analysts to explain language shift and loss in minority language settings. As a universal phenomenon, it refers to the situation where a whole group takes on the characteristics of another, an out-group, in order to achieve equality with it. This is what has happened in the Amazigh language contact areas in Algeria, Tunisia, and mainly Morocco.

Consideration of the political, historical, economic, and linguistic realities of intergroup relations (i.e. Amazighs and Arabs) is essential for an understanding of the North African linguistic situation. Tajfel's (1974) theory of intergroup relations and social change, based on the understanding that individuals have a desire to belong to groups that provide them with satisfaction and pride, provides the ground for the discussion and explanation of the linguistic and cultural assimilation in the Amazigh community. The Amazigh people are not proud of, nor are they satisfied with, their social identity. This can, indeed, help account for their strong desire to change and their attempt to attain a more adequate and positive social identity through total assimilation into the mainstream society.

Fieldwork in the Moroccan Amazigh community (El Kirat 2001) has revealed not only the degree of linguistic and cultural assimilation but also the indifference and lack of concern the community shows towards the Amazigh identity and language. The research also shows that, despite their awareness of their inadequate social identity and low status as a group, the Amazigh people endure their negative social identity and do not undertake any collective attempts to change or improve it. The majority of Amazigh people (i.e. common people) consider the position of their own group vis-à-vis the out-group (i.e. the Arab group) as stable and legitimate. This leads them to attribute the blame for their low position in society internally to themselves as a group because of its inferior characteristics and to their language as a useless and primitive language. As they do nothing to change their group situation they adopt individualistic actions as a means of attaining a positive social identity and avoiding social stigmatization and exclusion. Thus, they attempt to pass into the dominant group so as to achieve a more positive identity. This involves assimilation along both linguistic and cultural lines, which leads to the abandonment of all those visible markers of 'groupness' or identity that might compromise their chances of success in the mainstream society.

In these social movements which lead to the erosion of the Amazigh group markers, one fact stands out: most Amazigh people, if not all, are animated by the desire for material well-being and advancement and are drawn toward the majority group's language, culture, and way of life. They do, indeed, welcome all the changes required for social mobility and access to a more positive social identity. The cost-benefit aspects seem to justify this enterprise. Throughout Moroccan society, Amazigh dress, ornamentation, and dance (and even songs and fairy tales in some Amazigh communities) have disappeared as ordinary markers of the group. They persist or are remerging only as symbolic markers. Their appearance is, indeed, limited to special festivals, 'days', and the like, linked to commercial interests (such as cultural manifestations and folklore shows for attracting tourists), and they have become available to anyone interested, group member or not.

4. Language maintenance

Language shift is not an automatic outcome of language contact. The same conditions that may lead one community to shift from its native language may motivate another one to maintain it. In cases of maintenance, the minority language shows not only retention of both use and proficiency but also resistance to replacement. In this situation, the minority language is used for communication within the group while the dominant group language is used in all other instances. The functional distribution of the two languages remains intact. Such a case resembles a diglossic situation. What governs the maintenance of a language is the community's choice of the minority language in intragroup communication. While Tunisia is claimed to be a nearly monolingual community because its Amazigh community has totally shifted to the use of Tunisian Arabic, Algeria and Morocco offer cases of Amazigh communities fighting for the revival and maintenance of their native language. This concerns mainly the Amazigh communities in noncontact areas such as the Rif, Souss, and the south in Morocco and the Kabyle area in Algeria. Thus, many factors have come into play for the maintenance of Amazigh. Among those having had a key role it is possible to distinguish between internal and external factors. The internal factors have to do with the isolation of the language and the nonintegration of some Amazigh communities. The external factors include language awareness, cultural action, scientific activities, and the authorities' change in attitude.

Internal factors refers to the causes that are directly related to the Amazigh community and language. Two major internal factors are identified: the isolation of the language and the nonintegration of the Amazigh community.

The isolation of Amazigh in rural and mountainous areas in the North African communities is a double-edged sword. It has, on one hand, led to its exclusion and stigmatization in these communities, especially in the urban context, but has, on the other hand, contributed to its protection and maintenance in these areas. Amazigh owes its continuity over centuries to its isolation. The use of the language in the rural areas far away from Moroccan social and political power has allowed the language to continue as the means of communication of large populations. Amazigh is a vehicular language used for all the needs of everyday communication, including oral, familial, and social needs. It has been maintained even in some urban centers where the Amazighs are the dominant group (i.e. Agadir and Nador in Morocco and Kabylie in Algeria). This isolation of the Amazigh communities and their limited contact with and exposure to the Arabic language has contributed to the maintenance of Amazigh as a major means of communication within these communities. The shift to Arabic occurs only for communicative reasons (i.e. in interethnic communication). Intra-ethnic communication is done in the native language. In the language contact areas with an Amazigh minority group, the language is mainly used within the family.

Another determining factor in the maintenance of Amazigh comes from the nonintegration of some Amazigh communities. This is mainly the case in noncontact areas (e.g. the Rif and the south of Morocco, and Kabyle in Algeria). In contact areas, the psychological pressure on individuals is such that they opt for assimilation and accommodation for social promotion and mobility and also for social integration. Indeed, this is what has happened not only to most of the Amazigh communities in Tunisia but also to some communities in Morocco (especially in the central and north-eastern parts) and Algeria (e.g. Chaouia).

They attribute the cause of their plight externally to the out-group's unfair advantage over them and to the fact that the intergroup situation can be changed. This motivates their search for distinctiveness. This is the group that articulates the attribution of blame away from oneself as an inferior individual group member to the dominant group as an agent of oppression. This awareness of the illegitimacy of their inferiority and their perception that change is possible in the status relationship between them and the dominant group has led them to undertake an action to achieve a positive social identity through collective group action.

One of the main external causes of the Amazigh language shift in North Africa was the authorities' negative attitudes toward the language and its people. For years, the language was excluded from all the formal domains (e.g. school and the media), and Amazigh was a political rather than a linguistic issue. It was taboo to speak of the Amazigh language or identity. These were perceived as dividing elements threatening the national unity. In Morocco, national attention has focused on the Amazigh language and culture only quite recently, especially with King Hasan II's 20 August 1994 speech in which he proclaimed Amazigh an important part of the Moroccan identity and announced the introduction of the language in primary school instruction. A spectacular shift also occurred in Algeria with the status of the Amazighs, their language, and culture when the Algerian government approved a new constitution in November 1996, the first in the Maghreb or even in North Africa that officially recognized its Amazigh population in a document stipulating that three fundamental features characterize the Algerian people: Muslim, Arab, and Amazigh. Given the very small number of Amazigh speakers in Tunisia, the Amazigh people and language did not receive any attention.

The shift in the authorities' attitudes in Algeria and Morocco is perceived as an official recognition for the Amazighs. This has made some people feel more at ease in their Amazigh identity. The introduction of Amazigh into the educational system has reinforced a positive feeling among young people. It has also led to more language awareness, especially in the language contact areas where a negative attitude was prevailing.

The cultural action undertaken at both the international and local levels by the Amazigh elite in the form of cultural associations has had a very positive impact on the status of the Amazigh language, identity, and culture. A large number of such associations are located in Rabat, Casablanca, Agadir, and Nador in Morocco, and Algiers, Tizi Ouzzou, Bejaia, and Gherdaya in Algeria. This has contributed to the development of language awareness and the valorization of the language and identity in these areas. The Summer Institute of the Agadir University (held from 1980 to 1991) has contributed to a deep reflection on both the language and culture. The publication of periodicals such as *Amud* and *Tasafut* in Morocco, and *Tafsut* and *Asalu* in Algeria, has enabled the diffusion of the Amazigh written literature. The associations allowed for cultural manifestations. The Fourth Moroccan Summer Institute of Agadir elaborated a charter with recommendations aiming at the juridical and legal recognition of Amazigh and its promotion at the level of public institutions, especially in education and the media. Algeria succeeded in gaining an institutional breakthrough with the creation of two departments of the Amazigh language and culture in Tizi Ouzzou and Bejaia.

Cultural production has contributed to an indirect normalization of the language. Many writers contributed to the coining of new terms. Two types of literature have emerged: an oral one (e.g. poetry, tales, proverbs, and riddles), connected with traditional architecture, songs, and jewelry; and a modern written one, including mainly written productions and modern instrumental songs. The song is the domain where the vitality of the Amazigh culture appears. It is the most efficient one, indeed. The production of audio- and videotapes testifies to the dynamism of the mass culture (Boukous 1995).

5. CONCLUSION

Language shift in the North African context, i.e. Algeria, Morocco, and Tunisia, appears to be the direct outcome of language contact between two unequal communities, the Arab and Amazigh communities. The members of the Amazigh community in the three North African countries seem to have collectively chosen an outside language (dialectal Arabic) where their mother tongue (Amazigh) was formerly used. Pragmatic considerations of power, social access, material advancements, and so on are of the utmost importance in understanding the patterns of language use and shift in these communities. Language shift in Algeria, Morocco, and Tunisia appears in the gradual disappearance of the Amazigh language from a community where it was widely used until the 1970s. The Amazigh language is being shifted from because it is the language of the sociological minority group.

Along with widespread bilingualism, another indication of language shift appears from the reduction or complete absence of domains in which the language is used. The failure in the transmission of the Amazigh language to the young people has allowed dialectal Arabic to invade all the domains where the Amazigh language was exclusively used, even at home. The ignorance of the mother tongue by the younger generation will progressively lead not only to language shift but to language loss and death. The analysis of such spheres of activities reveals that although Amazigh is still employed in Morocco and Algeria, its use is restricted to only a few domains. If nothing is done to maintain and revive the language, it will not exceed the limits of the home, as is the case in Tunisia.

BIBLIOGRAPHICAL REFERENCES

Battenburg, John. 1999. “The gradual death of the Berber language in Tunisia”. *International Journal of the Sociology of Language* 137.147–161.

Boukous, Ahmed. 1995. “La langue berbère: Maintien et changement”. *International Journal of the Sociology of Language* 112.9–28.

Boyd, Sally. 1986. “Using the present to predict the future in language contact: The case of immigrant minority languages in Sweden”. *Language attrition in progress*, ed. Theo van Els, Kees de Bot, and Bert Weltens, 99–115. Dordrecht: Foris Publications.

El Kirat, Yamina. 1997. “Some causes of the Beni Iznassen Berber language loss”. *Langues et stigmatisation sociale au Maghreb. Peuples Méditerranéens*, 79 (April–June), 35–53.

——. 2001. “The current status and future of the Amazigh language in the Beni Iznassen community”. *Languages and Linguistics* (Tilburg) 8.81–95.

——. 2004. *The lexical and morphological structure of the Beni Iznassen Amazigh language in a context of language loss*. Ph.D. diss., University Mohammed V, Rabat.

——. Forthcoming. "Language endangerment in Morocco: The case of the Beni Iznassen Amazigh language". Paper presented at the international expert meeting on UNESCO Program for the Safeguarding of Endangered Languages, UNESCO Headquarters, Paris, 10–12 March 2003. [To appear in the UNESCO Proceedings.]

Fasold, Ralph. 1984. *The sociolinguistics of society*. Oxford: Blackwell.

Ferguson, Charles A. 1959. "Diglossia". *Word* 15.325–340.

Pencheon, Tom. 1968. "La langue berbère en Tunisie et la scolarisation des enfants berbérophones". *Revue Tunisienne des Sciences Sociales* 13.173–186.

Tajfel, Henri. 1974. "Social identity and inter-group behaviour". *Social Science Information* 13.65–93.

Yamina El Kirat el Allame
(University Mohammed V, Morocco)

Language Teaching → First Language Teaching; Second Language Teaching

Language Variation → Variation

Encyclopedia of Arabic Language and Linguistics

Volume III

Lat-Pu

General Editor
Kees Versteegh
(University of Nijmegen)

Associate Editors
Mushira Eid
(University of Utah)
Alaa Elgibali
(University of Maryland)
Manfred Woidich
(University of Amsterdam)
Andrzej Zaborski
(University of Cracow)

Advisory Board
Ramzi Baalbaki *(American University of Beirut)*
Elsaid Badawi *(American University of Cairo)*
Dominique Caubet *(INALCO, France)*
Clive Holes *(University of Oxford)*
Manfred Kropp *(Orient-Institut Der Deutschen Morgenländischen Gesellschaft)*
Jérôme Lentin *(Institut National des Langues et Civilisations Orientales)*
John McCarthy *(University of Massachusetts)*
Jamal Ouhalla *(University College Dublin)*
Jan Retsö *(Göteborg University)*
Sabah Safi *(King Abdulaziz University)*

Copy Editor
Carolyn Russ

ENCYCLOPEDIA OF ARABIC LANGUAGE AND LINGUISTICS

VOLUME III
Lat-Pu

General Editor
Kees Versteegh

Associate Editors
Mushira Eid
Alaa Elgibali
Manfred Woidich
Andrzej Zaborski

BRILL
Leiden – Boston
2008

Cover Illustration:
Blackboard, Mali. Photo Courtesy: Dinie Bouwman.

Despite our efforts we have not been able to trace all rights holders to some copyrighted material. The publisher welcomes communications from copyrights holders, so that the appropriate acknowledgements can be made in future editions, and to settle other permission matters.

© Copyright 2008 by Koninklijke Brill NV, Leiden, The Netherlands
Koninklijke Brill NV incorporates the imprints Brill, Hotei Publishing, IDC Publishers, Martinus Nijhoff Publishers and VSP

All rights reserved. No part of this publication may be re-produced, translated, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without prior written permission from the publisher.

Authorization to photocopy items for internal or personal use is granted by Brill provided that the appropriate fees are paid directly to The Copyright Clearance Center, 222 Rosewood Drive, Suite 910, Danvers MA 01923, USA. Fees are subject to change.

PRINTED IN THE NETHERLANDS

ISBN: 978 90 04 14973 1 (Set)
ISBN: 978 90 04 14475 0 (Volume 3)

This book is printed on acid-free paper.

Cover design: BEELDVORM, Pijnacker

List of Contributors

Ahmed-Sokarno Abdel-Hafiz, South Valley University, Egypt
Mahasen Hasan Abu Mansour, Umm al-Qura University, Saudi Arabia
Amira Agameya, Cairo University and The American University in Cairo, Egypt
Jorge Aguadé, University of Cádiz, Spain
Salman H. Al-Ani, Indiana University, U.S.A.
Enam Al-Wer, University of Essex, U.K.
Nadia Anghelescu, University of Bucharest, Romania
David L. Appleyard, School of Oriental and African Studies, U.K.
Werner Arnold, University of Heidelberg, Germany
Asya Asbaghi, Berlin, Germany
R.E. Asher, University of Edinburgh, U.K.
Muhammad Hasan Bakalla, King Saud University, Saudi Arabia
Reem Bassiouney, University of Georgetown, U.S.A.
Kirk R. Belnap, Brigham Young University, U.S.A.
Thami Benkirane, University of Fes, Morocco
Abdelâli Bentahila, Abdelmalek Essaadi University, Morocco
Elizabeth M. Bergman, Georgetown University, U.S.A.
Lidia Bettini, University of Firenze, Italy
Georges Bohas, Ecole Normale Supérieure Lyon, France
Dinie Bouwman, University of Groningen, The Netherlands
Joseph M. Brincat, University of Malta, Malta
Ellen Broselow, Stony Brook University, USA
Vít Bubeník, Memorial University of Newfoundland, Canada
Ernesto Capello, University of Vermont, U.S.A.
Michael Carter, Sydney University, Australia
Dominique Caubet, INALCO, France
Bernard Comrie, Max Planck Institute for Evolutionary Anthropology, Germany
Janusz Danecki, University of Warsaw, Poland
Mihai Dat, Dijon, France
Eirlys E. Davies, King Fahd School of Translation, Morocco
Everhard Ditters, University of Nijmegen, The Netherlands
Jean N. Druel, Institut Dominicain d'Etudes Orientales, Egypt
Lutz Edzard, University of Oslo, Norway
Ragia M. Effat, Cairo University, Egypt
Mushira Eid, University of Utah, U.S.A.
Kerstin Eksell, University of Copenhagen, Denmark
Shahir El-Hassan, Yarmouk University, Jordan
Muhammad Al-Sharkawi, American University in Cairo, Egypt
Daniela Rodica Firanescu, Dalhousie University, Canada
Robert Fradkin, University of Maryland, College Park, U.S.A.
Stefan A. Frisch, University of South Florida, U.S.A.
Dmitry Frolov, University of Moscow, Russia
Adam Gacek, McGill University, Canada
Geert Jan van Gelder, University of Oxford, U.K.
René-Vincent du Grandlaunay, Institut Dominicain d'Etudes Orientales, Cairo, Egypt
Gottfried Hagen, University of Michigan, U.S.A.

Rebecca Hasselbach, University of Chicago, U.S.A.
Roni Henkin, Ben-Gurion University of the Negev, Israel
Robert D. Hoberman, Stony Brook University, U.S.A.
Clive Holes, University of Oxford, U.K.
Moheiddin Ali Homeidi, King Saud University, Saudi Arabia
Jan Hoogland, University of Nijmegen, The Netherlands
Frederick M. Hoyt, University of Texas at Austin, U.S.A.
John Hühnergard, Harvard University, U.S.A.
Elizabeth Hume, Ohio State University, U.S.A.
Zeinab Ibrahim, The American University in Cairo, Egypt
Bruce Ingham, The School of Oriental and African Studies, U.K.
Tamás Iványi, Eötvös Loránd University, Hungary
Angelika Jakobi, Bayreuth, Germany
Elie Kallas, University of Trieste, Italy
M.N. Karassery, Calicut, India
Geoffrey Khan, University of Cambridge, U.K.
Djamel Eddine Kouloughli, CNRS, France
Joost Kremers, University of Nijmegen, The Netherlands
Jérôme Lentin, Institut National des Langues et Civilisations Orientales, France
Xavier Luffin, Free University of Brussels, Belgium
Zouhair Maalej, University of Manouba-Tunis, Tunisia
Mohamed Maamouri, University of Pennsylvania, USA
Michael C.A. Macdonald, University of Oxford, U.K.
Miklós Maróth, Pázmány Péter Catholic University, Hungary
Nathalie Mazraani, Brussels, Belgium
John J. McCarthy, University of Massachusetts Amherst, U.S.A.
Ernest N. McCarus, University of Michigan, U.S.A.
Manwel Mifsud, Kunsill Nazzjonali ta' l-Ilsien Malti, Malta
Catherine Miller, IREMAM, France
Mustafa Mughazy, Western Michigan University, U.S.A.
Franz-Christoph Muth, University of Munich, Germany
Samia Naïm, LACITO-CNRS, France
ʿAbdulfattah Olayiwola, University of Maiduguri, Nigeria
Jamal Ouhalla, University College Dublin, Ireland
Jonathan Owens, University of Bayreuth, Germany & University of Maryland, U.S.A.
Heikki Palva, University of Helsinki, Finland
Bruno Paoli, University of Bordeaux-3, France
Christophe Pereira, CRÉAM-INALCO, France
John R. Perry, University of Chicago, U.S.A.
Stephan Procházka, University of Vienna, Austria
Tariq Rahman, Quaid-i-Azam University, Pakistan
Narivelo Rajaonarimanana, CEROI-INALCO, Paris, France
Robert R. Ratcliffe, Tokyo University of Foreign Studies, Japan
Siamak Rezaei, UNBC, Canada
Samuel Rosenthall, Rochester, MI, U.S.A.
Tilman Seidensticker, Jena University, Germany
Irfan Shahîd, Georgetown University, U.S.A.
Kimary N. Shahin, University of British Columbia, Canada
Muhammad al-Sharkawi, American University in Cairo, Egypt
Avihai Shivtiel, Leeds, U.K.
Petra M. Sijpesteijn, University of Oxford, U.K.
Udo Simon, University of Heidelberg, Germany
Reinhard Stempel, Universidad del País Vasco, Spain

MICHAEL P. STRECK, University of Leipzig, Germany
ZEINAB AHMED TAHA, The American University in Cairo, Egypt
AHMAD TAHER HASSANEIN, Misr University for Sciences and Technology, Egypt
CATHERINE TAINE-CHEIKH, CNRS, LACITO-UMR 7017, France
MAURO TOSCO, University of Naples, Italy
KEES VERSTEEGH, University of Nijmegen, The Netherlands
ANGELES VICENTE, University of Zaragoza, Spain
MARAT YAVRUMYAN, University of Bayreuth, Germany
MUNTHER A. YOUNES, Cornell University, U.S.A.
CHAKIR ZEROUAL, University Sidi Mohamed Ben-Abdellah, Morocco & CNRS, France

L

Latin America

1. Introduction

Arabic speakers have been part of Latin American society since the mid-17th century, when the Syrian Chaldean priest ʾIlyās al-Mawṣilī came to South America on a papal mission. The ties between Iberian and Arab culture date even further back, to the conquest of the Iberian Peninsula by the Moors, an event that left an indelible impression on Spanish and Portuguese culture, architecture, and language. So it should come as no surprise that Latin America, in particular the great economic powerhouses of Brazil and Argentina, formed one of the most important destinations for Middle Eastern emigration beginning in the late 19th century. The relative ease of acculturation, in particular on the part of the Syro-Lebanese, bespeaks the similarities of a common cultural heritage, as well as the ability of these migrants to adapt to shifting circumstances.

This entry examines the state of the Arabic language and the cultural space that Latin Americans of Arab descent have carved for themselves in the past century. It first presents a gloss of the historical background of Arab emigration to Latin America and then a discussion of the role and place of Arabic in the region. Overall, the discussion focuses on the two most important groups of Arab immigrants, those in Brazil and Argentina, although other countries with an important Arab presence, such as Chile, Ecuador, and Mexico, are touched upon.

2. The sociocultural background

The first Arabic speakers to emigrate en masse were Maghrebi Jews leaving Morocco in the wake of the Spanish-Moroccan war (1859–1860), a time when business shrank and xenophobia against Jews increased. Unlike later waves, the vast majority of these migrants left exclusively for South America, particularly Brazil, where by 1890 more than one thousand migrants worked as peddlers or middlemen in the great Amazonian rubber trade in Pará. Brazilian naturalization was readily available, an important factor for the many migrants who returned home to establish commercial ventures and did so with the protection of the Brazilian government against imprisonment or xenophobic violence (Lesser 1999:46–47).

The largest wave of emigrants, however, were the Syro-Lebanese, a group that began a massive migration at the close of the 19th century, driven largely by the economic flux in the Ottoman Empire and the dreams of a better future in America. (Because they historically formed a single province in the Ottoman Empire, the two nationalities are treated here as a single group.) Their entry during this time also led to their being labeled *turcos* by the local population, an appellation that the group has been unable fully to shake off despite the existence of independent Syria and Lebanon. Foreign and Lebanese companies attempted to profit from their migration and developed great advertising campaigns that brought many a landlubber onto their boats. Several embarked unaware of their final destination, assuming that America meant the United States and only realizing their mistake when it was too late. For Maronite Christians from Lebanon, the largest group emigrating, another factor was the increasing dominance of Muslims throughout the empire, following the large migrations from the Crimea, the Caucasus, and the Balkans in the wake of the Ottoman-Russian wars.

These early settlers navigated their new countries by tapping into the existing trade networks developed by the Moroccans. The Syro-Lebanese tended to congregate in large urban centers and formed a conduit for goods from the rural interior as well as for products obtained from overseas. The greatest populations were in large commercial centers such as São Paulo, Buenos Aires, Baranquilla, and Guayaquil. Arabic-speaking enclaves appeared, e.g. the neighborhood surrounding Rua 25 de Março in São Paulo, which rapidly developed into the urban heart of the local population. Many of these merchants swiftly acquired wealth and prestige, and by the early years of the 20th century they began to play a prominent role in local society.

At the same time, a backlash against the successful economic rise of the Syro-Lebanese began to be felt. While many Latin American countries had encouraged immigration from Europe in the 19th century, the influx of people from an unknown and unfamiliar land encouraged xenophobia. The settlement patterns of Arabic speakers in the large cities also displeased many elites in Brazil and Argentina, who had hoped immigrants would become agricultural workers. Violence fed by vitriolic anti-immigrant rhetoric rose against the Arabs in Buenos Aires in the 1910s, while restrictions on Middle Eastern entry were passed in Brazil and Colombia in the 1930s. Although efforts by the Ottoman commercial and diplomatic elite led some new-comers to venture into the hinterlands, the escalating Arab population (65,000 in Argentina by 1914, 107,000 in Brazil by the end of the 1930s) continued to congregate in the main centers (Klich 1993:199; Lesser 1999:49).

The concentration of the community helped to propel the Syro-Lebanese into their status as a powerful economic and political block in the 1920s, when they established banking concerns across the continent. These financial institutions featured generous credit policies for the community, fostering the creation of industrial endeavors such as the great Arab textile factories of São Paulo and Santiago de Chile. The most prosperous of these manufacturers began to assimilate into the upper strata of Latin American society. They moved into wealthy neighborhoods, and their children attended the best schools, becoming professionals and marrying into powerful families.

Although industrialists such as Juan Yarur in Chile had long used personal networks to organize manufacturing concerns, the second generation applied their connections to enter politics. In some areas, such as São Paulo and Guayaquil, the vast political web created by the Syro-Lebanese led to a situation that Oswaldo Truzzi (1995) has called *sobre-representaçao* 'overrepresentation'. Populists such as Ecuador's Assad Bucaram used the slums of Guayaquil to build the CFP party in the 1950s, while Paulo Jorge Mansur, owner of Radio Difusor in São Paulo, bankrolled a number of winning campaigns beginning in the 1960s (for Brazil, see Truzzi 1995:31–42; for Ecuador, see Almeida 1996:106–108). By the 1990s, the consolidation of the Arab position in Latin American politics had become obvious as presidents of Arab descent were elected in Argentina, Colombia, Ecuador (two), and Venezuela.

3. The role and place of Arabic in Latin American society

The spectacular ascent of the Syro-Lebanese during the 20th century was largely predicated upon their initial bond as members of the same culture. While the first generation of immigrants remained largely endogamous, subsequent generations have mostly assimilated into the local culture. Their common linguistic bond has suffered throughout the region, though mitigated by local circumstances. In general, the presence of Arabic is much higher in Brazil than in Spanish Latin America, except among the latter's largely bilingual Muslim community.

The cosmopolitanism of the early settlers and their numerical strength contributed to this quick assimilation. As merchants with extensive ties across the Mediterranean, the Maghrebi and Syro-Lebanese immigrants were already familiar with the French and Spanish languages before coming to the Americas. In countries like Mexico and Ecuador, where there were rather small communities initially, this cosmopolitan identity aided their absorption into the local culture. The lack of a large population also diminished the possibility of setting up Arabic-language schools, as happened in Argentina and Brazil. The second generation, by necessity, had to take advantage of the national public school system, which quickly eroded their use of Arabic in daily life to the point where the language is little used in these countries today.

In Argentina the situation was rather different. The large Arab immigrant population led to the development of schools conducted completely in Arabic, particularly among the Aleppine community. Arabic-language newspapers also circulated in the main cities and continue to do so, even though most contemporary Arabic newspapers are bilingual. However, the intensity of the negative reaction against a language that symbolized Islam in the early 20th century fed the desire to blend in, particularly among Christian circles. Over the years, the use of Arabic declined until it was largely relegated to the Muslim community, who often converted to Christianity in order to conduct business with greater ease and achieve prosperity. The case of Argentine ex-president Carlos Menem is a case in point. Like so many others, Menem's childhood in the La Rioja district in the north of the country was plagued by the anti-Muslim sentiment dominating the conservative Catholic province. His conversion to Catholicism in 1963 was a prelude to his rise in politics.

Among the Muslim community, Arabic continued to flourish in the private and religious spheres while diminishing in the public sphere as later generations became increasingly bilingual. A study by Estela Biondi Assali (1989) reviewed the correlation between social codes of conduct and the use of Arabic or Spanish among Muslims of the first, second, and third generations in Tucuman. Her research revealed that a highly stratified system of usage existed among Arab Argentines, who tended to equate the language with the intimacy of the home, the mosque, and the family, while Spanish constituted the language of public life. The degree of Arabic usage differed across generations. For the first immigrants, familial and pan-Islamic communication remained tied to Arabic, while both languages were used in their professional life. Reading and writing also tended to be mostly confined to Arabic. In the second and third generations, bilingualism increased to include both the home and the mosque, while reading and writing tended to be conducted in Spanish. The exceptions to these norms occurred at moments of intense emotion, when Arabic would become the dominant tongue, or in writing letters to relatives either in the Middle East or other parts of Argentina. In effect, the language was relegated to secondary status among a group striving to restrict their identity in order to prosper in a society intolerant of their religion, with which the language remained indelibly linked.

This situation has begun to change in recent years as Argentine society has become more tolerant of Islam, largely because of Menem's presidency and in particular because of the shocking death in 1995 of his son, who had reclaimed the faith of his forefathers. The presence of Islamic funeral rites on Argentine radio and television directly led to a revalorization of the religion by a great number of Argentines as well as the onset of an extension of the Islamic infrastructure in Argentina (Jozami 1996:67–68). To wit, the government provided support for the construction of a major cultural center in the Palermo Mosque in Buenos Aires. Built on land donated by the government, the corner-stone was laid on 7 December 1998. The structure now forms a major cultural complex, complete with religious instruction as well as both a primary and secondary school, each of which has about three hundred students. Independent cultural centers have also appeared since the early 1990s, as the younger generation seeks to understand their heritage. These centers often celebrate the trappings of Arab culture like music, food, and dance; at times they also offer Arabic lessons.

However, the overall state of Arabic in Argentina continues to be relatively under-developed, with little presence of the language outside of the religious and private spheres. There are no major universities, for instance, that offer courses in the language. This situation is rather distinct from Brazil, the other main center of Latin America's Arab population.

While Argentina's Arab community was forced to contend with a strong degree of xenophobia from the onset of their tenure in their adopted country, those who came to Brazil managed to formulate a prouder sense of identity by adopting Brazilian traits without shedding their Arab past. Part of the reason for their success concerned the racial differences prevalent in Brazil, a country whose population is largely of European and African descent. Arabic speakers entering Brazil did not stand out racially to the same degree as those who settled in Spanish and Italian Argentina. In addition, idealized notions of Arab culture had existed as part of Brazilian 19th-century romanticism, as in Teófilo Braga's lauding of the Mozárabs. Although many native Brazilians rebuked the growing Syro-Lebanese population during the strongly nationalist Vargas regime of the 1930s, nativist rhetoric paled in comparison to Argentina, and the *turco* continued to captivate the imagination of Brazilian intellectuals as in the central character Nassim of Jorge Amado's celebrated novel *Gabriela, clove and cinammon.*

Some Arab Brazilians felt pressured to assimilate by abandoning highly visible markers and adopting Portuguese-sounding names. One of the primary strategies, however, was to praise the notion of a special relationship between the two cultures in works such as Tanus Jorge Bastani's *O libano e os libaneses* (1945), which extols Middle Eastern culture while focusing on its historic connections with Brazil. One of the most noticeable signs of this appreciation for the synchronicity between the groups came in 1922 with the erection of a statue called *Amizade sirio-libanesa* 'Syro-Lebanese Friendship' that celebrated the centennial of Brazilian independence. The base of the statue represents Syro-Lebanese contributions to world culture, while the central figures depict the Brazilian Republic embracing an indigenous warrior and a Syrian maiden. The ceremonies dedicating the statue also included a poetry contest in which contestants presented material describing the life of the émigré in Brazil. The victor, ʾIlyās Ḥabīb Farḥāt, wrote in Arabic to underscore the hyphenated identity of the immigrant, at once thankful for Brazilian acceptance and confident that the new Lebanese presence expanded Brazil's Christian heritage.

The presence of Arabic poetry at this important event was due to the rapid inauguration of Arabic-language educational institutions, beginning in 1912 with the Gimnásio Oriental. Three others in São Paulo alone were founded by 1922, helping to transform the city into the center of South American Arabic intellectual production. Over the course of the next couple of decades, Brazil became the cornerstone of the southern *mahjar* literature. One of the early Brazilian poets was Fawzī al-Maʿlūf, whose poem *ʿAlā bisāṭ ar-rīḥ*, published in 1929, paints the world of the immigrant alone in a vast wilderness. Others, such as Farḥāt, Michel Maʿlūf, and Rašīd Salīm al-Xūrī (al-Khouri) celebrated the vast riches of Arab Brazilian culture when they founded the literary society *al-ʾUšra al-ʾAndalusiyya* as a place in which a 'second Andalusia' would be built. Al-Xūrī's early work celebrates pastoral themes.

The development of Arab-Brazilian literature continued through the next several decades, significantly after Jorge Salīm Safady founded a publishing house that catered to Arabic-language texts. Arabic poetry, histories of Brazil in Arabic, and translations of Brazilian literature formed the bulk of his publishing work, illustrating the degree to which the wish to formulate a simultaneously Brazilian and Arab identity dominated the community. Jose Xūrī added to this possibility when he wrote a Portuguese/Arabic dictionary in the 1950s. Others, such as Jorge Suleiman Yāzijī, crafted Arabic manuals for Brazilians using Latin letters to make the language more familiar.

As a result of these endeavors, Arabic flowered as an important aspect of Arab-Brazilian life and, indeed, of Brazilian culture in general. The University of São Paulo recognized this fact when it incorporated Arabic literary and language studies into its program of Oriental literatures in 1970. The program, which also offered instruction in Armenian, Chinese, Hebrew, Japanese, Russian, and Sanskrit, formed a joint department with the linguistics department until 1986, when the two were split into separate entities. Today, students can take an eight-term sequence in Arabic language and six courses on Arabic literature. Included in the latter is extensive study of immigrant literature (see www.usp.br).

Brazilian Arabic thus continues to flourish. It has also incorporated a number of Portuguese loanwords, illustrating its validity to contemporary Arab-Brazilian life. In a recent study by Neuza Neif Nabhan (1994), several levels of interference are identified. As in Argentina, Portuguese dominates the public sphere and so the presence of Portuguese loanwords is most highly noted in terms of work-related subjects. Some examples from Nabhan's study (1994:223–225) include:

bi-maskiet	< *mascatear* 'to peddle, to sell'
bi-marik	< *marcar* 'to mark'
bulūza ṣūf, pl. *bluzāt ṣūf*	< *blusa de lã* 'woolen sweater'
gravata, pl. *gravatāt*	< *gravata* 'tie'
maḥall mubilia	< *loja de mobilias* 'furniture store'

Other domains in which loanwords are common are domestic and social life, particularly involving relationships, food, and education (Nabhan 1994:235–238):

brimu, primu	< *primo* 'cousin'
namurado	< *namorado* 'boyfriend'
fōrn	< *forno* 'oven'
māmun	< *mamão* 'papaya'
bil-brimariu	< *no primário* 'in nursery school'

Some of these words represent the incorporation of words from Portuguese that do not have exact equivalents in Arabic, such as *māmun* and *namorado*. Others, however, represent a simpler lexeme than that contained in Arabic, in which *brimu* (< *primo* 'cousin') replaces *ibn xāl* 'mother's brother's son' or *ibn ʿamm* 'father's brother's son'. In cases where there is a complex lexeme in both languages, there are three ways to organize the resulting loanword (Nabhan 1994:238–240), either an Arabic/Portuguese syntagm (*maḥall mubilia* 'furniture store' < Portuguese *loja de mobília* + Arabic *maḥall ʾaṯāṯ*), a Portuguese/Arabic syntagm (*bulūza ṣūf* 'woolen sweater' < Portuguese *blusa de lã* + Arabic *qamīṣ ṣūf*), or a Portuguese/Portuguese syntagm (*bēti finu* 'fine-tooth comb' < *pente fino*).

Nabhan's study also showed that this phenomenon was common both in established descendants of earlier immigrants and in the smaller groups of newer émigrés who began appearing in the 1970s as trouble roiled the Middle East, particularly during the Lebanese civil war. In effect, Arabic continues to evolve as a living language in Brazil even as the Arab-Brazilian population grows.

Throughout the region, then, it is possible to say that Arabic has a definite presence in the countries with a larger Arab population, such as Brazil and Argentina, while it has largely faded from use among the descendants of Arabic speakers in countries with a smaller population such as Ecuador, Colombia, and Mexico. In Argentina, the resurgence of Islam in recent years is perhaps leading toward a renewed use of Arabic outside of religious circles, although that waits to be seen. In Brazil, with a strong tradition of Arabic literature and a transition relatively free of intense persecution, Arabic continues to hold a key place in Arab-Brazilian society as a method of communication, artistic expression, and academic study.

Bibliographical references

Almeida, Mónica. 1996. "Phoenicians of the Pacific: Lebanese and other Middle Easterners in Ecuador". *The Americas* 53:1.87–112.

Biondi Assali, Estela. 1989. "Alternancia de los códigos español-árabe entre los bilingües de Tucumán, Argentina". *Caravelle* 52.33–55.

Jozami, Gladys. 1996. "The manifestation of Islam in the Americas". *The Americas* 53:1.67–86.

Karpat, Kemal H. 1985. "The Ottoman emigration to America, 1860–1914". *International Journal of Middle East Studies* 17:2.175–209.

Klich, Ignacio. 1992. "*Criollos* and Arabic speakers in Argentina: An uneasy *pas de deux*, 1888–1914". *The Lebanese in the world: A century of emigration*, ed. Albert Hourani and Nadim Shehadi, 243–284. London: Center for Lebanese Studies and I.B. Tauris.

——. 1993. "Argentine-Ottoman relations and their impact on immigrants from the Middle East: A history of unfulfilled expectations, 1910–1915". *The Americas* 50:2.177–205.

—— and Jeffrey Lesser. 1998. *Arab and Jewish immigrants in Latin America: Images and realities.* London and Portland, Ore.: F. Cass.

—— (eds.). 1996. Special issue of *The Americas* 53:1.

Lesser, Jeffrey. 1999. *Negotiating national identity: Immigrants, minorities, and the struggle for ethnicity in Brazil.* Durham, N.C., and London: Duke University Press.

Mûsili, Elias al-[ʾIlyās al-Mawṣilī]. 2003. *An Arab's journey to colonial Spanish America: The travels of Elias al-Mûsili in the seventeenth century.* Transl. and ed. Caesar E. Farah. Syracuse, N.Y.: Syracuse University Press.

Nabhan, Neuza Neif. 1994. "The speech of the Arabic-Lebanese in Brazil: A lexical study". *Indian Journal of Applied Linguistics* 20.229–247.

Nagrib Sabbagh, Alphonse. 1978. *O meio ambiente na literatura árabe escrita no Brasil.* Ph.D. diss., Universidade Federal do Rio de Janeiro.

Ramírez Carrillo, Luis A. 1994. *Secretos de familia: Libaneses y élites empresariales en Yucatán.* México, D.F.: Consejo Nacional para la Cultura y las Artes, Dirección General de Publicaciones.

Roberts, Lois J. 2000. *The Lebanese immigrants in*

Ecuador: A history of emerging leadership. Boulder, Colo.: Westview.
Tasso, Alberto. 1989. *Aventura, trabajo y poder: Sirios y libaneses en Santiago del Estero, 1880–1980*. Buenos Aires: Ediciones Índice.
Truzzi, Oswaldo Mário Serra. 1992. *De mascates a doutores: Sírios e libaneses em São Paulo*. São Paulo: Editora Sumaré.
——. 1995. "Sírios e libaneses em São Paulo: A anatomia da sobre-representaçao". *Imigração e política em São Paulo*, ed. Boris Fausto, Oswaldo Truzzi, Roberto Grün, and Célia Sakurai, 27–69. São Paulo: IDESP.
——. 1997. *Patrícios: Sírios e libaneses em São Paulo*. São Paulo: Editora Hucitec.

Ernesto Capello (University of Vermont)

Latin Loanwords

Latin loanwords entered the Arabic language during the seven centuries or so which elapsed between the Roman conquest of Bilād aš-Šām in 63 B.C.E. and its subsequent conquest by the Arabs in the 630s C.E. The area extended from the Euphrates River or the Amanus Mountains in the north to the Sinai Peninsula in the south. Roman rule in this area was preceded by some three centuries of a Macedonian presence, initiated by the conquests of Alexander the Great. During that time the Greek language spread in this region and became its language of cultural dominance, even as Aramaic/Syriac had become dominant among the Semitic population of the region. Consequently, many Latin loanwords reached Arabic through the filter of Greek or Aramaic (→ Greek loanwords; → Aramaic/Syriac loanwords).

Greek contributed more loanwords than Latin, but some loanwords from Latin, as relatively few as they are, have become part and parcel of the Arabic language, owing to the strong Roman military and administrative presence in the region. Others had an ephemeral life in Arabic and have survived only in medieval historical texts that refer to the distant past, sometimes used only once *hapax legomena*. *Sirāṭ* < *strata* 'paved Roman road' is a key Qurʾānic term meaning 'path' or 'way' (Q. 1/5, 6). *Qaṣr* < *castrum,* in the sense of both 'castle' and 'palatial mansion' (Q. 22/45, 7/74), nowadays only means 'palace'. *Qasṭal* < *castellum* 'forti-fied post; tank, cistern' still appears in Modern Arabic as a toponym deriving from the first signification and as 'water pipe, water main', deriving from the second meaning (Dozy 1927:II, 352–353). *Fusṭāṭ* < *fossatum* 'camp; ditch' (Shahîd 2002:64) survives until the present day as the name of Old Cairo. *ʿAskar* < *exercitus* 'army' is a more plausible etymology than Persian *laškar* (Fraenkel 1886:239). *Mīl* < *miliarium* 'milestone' and 'measure of distance, mile' were both used in Classical Arabic (Maqrīzī, *Xiṭaṭ* 199), but the word has survived only in the second signification. *Sijill* < *sigillum* 'signet or its impression; a document; an imperial edict' in Late Antiquity is used in Arabic only as 'document, official or juridical' (Q. 21/104; De Blois 1997). *Sijn* < *signum* 'signet, signet's impression; prison [in Late Antique usage]' is still used in this sense in Arabic (Niehoff-Panagiotidis 1996; Schneider 1997). *ʾIṣṭabl* derives from *stabulum* 'stable' (Fraenkel 1886:124; Viré 1978). *Barīd* < *veredus* was applied in classical times to the post horse, the courier, and the stage and is used nowadays only for the postal service and the mail (Sourdel 1960b). *Hury*, pl. *ʾahrāʾ* derives from *horreum* 'storehouse, granary' (Fraenkel 1886:136). *Furn* derives from *furnus* 'oven, bakehouse' (Dozy 1927:II, 262). *Balāṭ* < *palatium* 'imperial residence' means in Arabic 'royal court'; in the sense of 'flagstone, paved way', it derives from Greek πλατεῖα (Sourdel 1960a). The Latin term *cohors*, pl. *cohortes* 'division of the Roman army; police' appears in Arabic in its two significations as *kurdūs*, *kurdūsa*, pl. *karādīs* 'a company, a subdivision in the army' (Fraenkel 1886:239) and as *šurṭa* 'police' (Brockelmann 1961:122; Nielsen 1997). *Qayṣar* < *Caesar*, the generic title for the head of the Roman/Byzantine state (Shahîd a.o. 1978), also survives as a personal name among Christian Arabs. The name *ar-Rūm* < *Rhōmaîoi* was applied to both the Romans of Rome and those of Constantinople, the Byzantines (Q. 30/2), although now the former are referred to as *Rūmān* < *Romani*.

Less common words from the Roman military and administrative establishment are *biṭrīq* < *patricius* 'honorary title of high-ranking dignitary, unrelated to any specific function', by far the most common of all Roman military terms in the sources, sometimes applied indiscriminately to a Persian commander (Xwārizmī, *Mafātīḥ* 128; Shahîd 1960); *fīqār* < *vicarius* 'nonmilitary officer endowed with various administrative functions' (Ṭabarī, *Taʾrīx*, prima series IV, 2099); *qusṭār* < *quaestor*, a financial officer (Fraenkel 1886:187; Kunitzsch and Ullmann 1992, nos. 391, 392); *qubiqlār* < *cubicularius* '*praepositus sacri cubiculi*, the high-ranking grand chamberlain' (Ṭabarī, *Taʾrīx*, prima series IV, 2099, 2100); *dumustuq* < *domesticus*, a term applied to ecclesiastical, military, and civil officers, but in Arabic only to military commanders (Xwārizmī, *Mafātīḥ* 129); *qawmas* < *comes* 'companion [honorary title for officers in various functions]' (Xwārizmī, *Mafātīḥ* 129; Dozy 1927:II, 436); *ʾibraṭuryūn* < *praetorium* 'the tent or headquarters of a Roman general', hence 'palace' (Ullmann 2002); *baqṭ* < *pactum* originally 'agreement or compact' but later 'tribute', which is what it signified in Arabic (Maqrīzī, *Xiṭaṭ* 199–200).

Certain words pertaining to the Roman monetary system have come from Latin, such as *dīnār* < *denarius* 'gold coin' (Q. 3/75; Balāḏurī, *Futūḥ* 573; Miles 1965); *fils* < *follis* 'copper coin piece' (Udovitch 1965); *nummiyya* < *nummus* 'coin of the lowest denomination' (Maʿarrī, *Ġufrān* 583). From Latin come four words for weights: *qinṭār* < *centenarium* 'weight of a hundred pounds' (Q. 3/14, 75; Xwārizmī, *Mafātīḥ* 179), *ʾūqiyya* < *uncia* 'twelfth part of a pound' (Xwārizmī, *Mafātīḥ* 178), *mudd* 'measuring vessel' < *modius* (*Oxford Latin Dictionary* 1123; Arabic sources in Fraenkel 1886:206–207), *ʾīṭālīqūs* < *italicus* 'weight of sixteen uqiyyas' (Xwārizmī, *Mafātīḥ* 178). Both *nummiyya* and *ʾīṭālīqūs* are now extinct.

Words related to everyday life have also come from Latin, e.g. *ṣābūn* < *sapo* 'soap' (Kühnel 1995; Dietrich 1995); *qindīl* < *candela* 'wax light, Arabic lamp' (Fraenkel 1886:95); *mindīl* < *mantelum* 'hand towel; napkin', in Arabic, 'handkerchief; head kerchief' (Rosenthal 1992); *dals* < *dolus* 'fraud in commercial transactions' (Schacht 1964:9; Fraenkel 1886:188). Now extinct are *muṣṭār* ~ *musṭār* < *mustum* 'new wine' and *mustarium* 'new wine's container' (Nöldeke 1961:54.12), *siqlāṭūn* < *sigillatus* 'textile covered with figures and patterns in relief' (Masʿūdī, *Murūj* I, 230; Colin 1930).

In modern times Arabic has picked up from Latin two terms, possibly three: *qunṣul* < *consul* 'functionary in an embassy' (Spuler 1965); *ʾimbarāṭūr* < *imperator,* never used for the Roman or Byzantine ruler in medieval texts but used in modern times for heads of states that still have emperors. *Qubṭān*, the captain of a ship from Late Latin *capitaneus*, is not attested in the Classical Arabic lexica but is used in post-Medieval and Modern Standard Arabic, mediated through Ottoman Turkish (Dozy 1927:II, 310).

Latin etymologies have been suggested for certain loanwords in Arabic, but questions still surround them: *ʿatīq* < *antiquus* 'old, ancient' (Q. 22/33; Fraenkel 1886:210–211); *qamīṣ* < *camisia* 'shirt or mantle' (Fraenkel 1886:45; Brockelmann 1961:169; Teubner 1997); *balad* < *palatium* 'town, inhabited area' (Fraenkel 1886:28); *qisṭ* < *iustitia* 'justice' (Q. 3/18, 21) or < *sextarius* 'measure of fluid or dry material' (Xwārizmī, *Mafātīḥ* 179); *qisṭās* < 'balance' (Q. 17/35; Fraenkel 1886:198), related to *constans* as used of the *libra* 'pair of scales, balance'.

Latin loanwords in Arabic have never been the topic of a monograph as Greek, Aramaic, and Persian have been. In medieval times, Muslim scholars were aware of the problem of foreign words in the *Qurʾān*, including Latin, which they called *Rūmiyya*, often confused with Greek, while others strongly denied this since it ran counter to the orthodox Islamic dogma of *ʿarabiyyat al-Qurʾān*, the Arabness of the *Qurʾān*. Of these medieval authors, Jalāl ad-Dīn as-Sūyūṭī (d. 916/1610–1611) was the only one who listed Latin words from what he called *ar-Rūmiyya* (Latin and Greek) in his work *al-Mutawakkilī* (91–100), which he arranged according to the various languages from which he thought the foreign vocabulary words were derived. In spite of the fact that modern research has shown many inaccuracies in his etymological attributions, some terms he listed as Latin are correct, such as *ṣirāṭ* and *qinṭār*, from *strata* and *centenarium*. The first scientific monograph on the foreign vocabulary of the *Qurʾān* including Latin was written in 1938 by Arthur Jeffery, whose work, arranged alphabetically, is still indispensable. Like as-Sūyūṭī, he limits the Latin loanwords in his work to those in the *Qurʾān* (Jeffery 1938:82, 133, 146–147, 163–174, 211, 243–244; cf. Rippin 2002); hence, the present entry is the first attempt to list Latin loanwords found not only in the *Qurʾān* but also in the Arabic language generally.

Bibliographical references

Primary sources

Balāḏurī, *Futūḥ* = ʾAbū l-Ḥasan ʾAḥmad ibn Yaḥyā al-Balāḏurī, *Futūḥ al-buldān*. Ed. S. al-Munajjid. Cairo: Maktabat an-Nahḍa, 1965.

Jawālīqī, *Muʿarrab* = ʾAbū Manṣūr Mawhūb ibn ʾAḥmad al-Jawālīqī, *al-Muʿarrab*. Ed. A.M. Šākir. Tehran, 1966.

Maʿarrī, *Ġufrān* = ʾAbū l-ʿAlāʾ ʾAḥmad ibn ʿAbdallāh al-Maʿarrī, *Risālat al-ġufrān*. Ed. ʿĀʾiša ʿAbd ar-Raḥmān ibn ʿAbdallāh. Cairo: Dār al-Maʿārif, 1963.

Maqrīẓī, *Xiṭaṭ* = Taqī d-Dīn ʾAbū l-ʿAbbās ʾAḥmad al-Maqrīẓī, *al-Mawāʿiḏ wa-l-iʿtibār fī ḏikr al-xiṭaṭ wa-l-ʾāṯār*. Bulaq, 1270 A.H. (Repr., Baghdad, 1970.)

Masʿūdī, *Murūj* = ʾAbū l-Ḥasan ʿAlī ibn al-Ḥusayn al-Masʿūdī, *Murūj aḏ-ḏahab*. Ed. Charles Pellat. Beirut: Publications of the Lebanese University, 1966.

Suyūṭī, *Mutawakkilī* = Jalāl ad-Dīn ʾAbū l-Faḍl ʿAbd ar-Raḥmān ibn ʾAbī Bakr as-Suyūṭī, *al-Mutawakkilī*. Ed. ʿAbd al-Karīm az-Zubaydī. Beirut: Dār al-Balāġa, 1988.

Ṭabarī, *Taʾrīx* = ʾAbū Jaʿfar Muḥammad ibn Jarīr aṭ-Ṭabarī, *Taʾrīx ar-rusul wa-l-mulūk*. Ed. M.J. de Goeje. Repr., Leiden: E.J. Brill, 1964.

Xwārizmī, *Mafātīḥ* = ʾAbū ʿAbdallāh Muḥammad ibn ʾAḥmad al-Xwārizmī, *Mafātīḥ al-ʿulūm*. Ed. Gerlof van Vloten. Repr., Leiden: E.J. Brill, 1968.

Secondary sources

Brockelmann, Carl. 1961. *Grundriss der vergleichenden Grammatik der semitischen Sprachen*. Repr. Hildesheim: G. Olms.

Colin, Georges S. 1930. “*Sigillatus*”. *Romania*, April 1930: 178–190, July 1930: 418.

De Blois, F.C. 1997. “*Sid̲j̲ill*”. *Encyclopaedia of Islam* IX, 538–545. 2nd ed. Leiden: E.J. Brill.

Dietrich, A. 1995. “*Ṣābūn*”. *Encyclopaedia of Islam* VIII, 693. 2nd ed. Leiden: E.J. Brill.

Dozy, Reinhart. 1927. *Supplément aux dictionnaires arabes*. Repr. Leiden and Paris: E.J. Brill.

Fraenkel, Siegmund. 1886. *Die aramäischen Fremdworter im Arabischen*. Leiden: E.J. Brill. (Repr., Hildesheim: G. Olms, 1962.)

Jeffery, Arthur. 1938. *The foreign vocabulary of the* Qurʾān. Barooda: Oriental Institute.

Kühnel, H. 1995. “Seife”. *Lexikon des Mitlelalters* VII/1, 1710–1711. Munich and Zurich: Artemis und Winkler Verlag.

Kunitzsch, Paul and Manfred Ullmann. 1992. *Die Plejaden in den Vergleichen der arabischen Dichtung. Sitzungsberichte der Bayerischen Akademie der Wissenschaften: Philosophisch-Historische Klasse*, Heft 4.

Miles, George. 1965. “*Dīnār*”. *Encyclopaedia of Islam* II, 297–299. 2nd ed. Leiden: E.J. Brill.

Niehoff-Panagiotidis, Jannis. 1996. “Romania Graeco-Arabica: Lat. *signum* > σίγνον > arab *sign*”. *Romania Arabica: Festschrift Reinhold Kontzi*, 1–19. Tübingen: G. Narr.

Nielsen, J.S. 1997. “*S̲h̲urṭa*”. *Encyclopaedia of Islam* IX, 510. 2nd ed. Leiden: E.J. Brill.

Nöldeke, Theodor. 1961. *Delectus veterum carminum arabicorum*. Repr. Wiesbaden: O. Harrassowitz.

Rippin, Andrew. 2002. “Foreign vocabulary”. *Encyclopaedia of the Qurʾān* II, 226–237. Leiden: E.J. Brill.

Rosenthal, Franz. 1992. “*Mandīl*”. *Encyclopaedia of Islam* VI, 402–403. 2nd ed. Leiden: E.J. Brill.

Schacht, Joseph. 1964. *Introduction to Islamic Law*. Oxford: Clarendon Press.

Schneider, Irene. 1997. “*Sid̲j̲n*”. *Encyclopaedia of Islam* IX, 547–548. 2nd ed. Leiden: E.J. Brill.

Shahîd, Irfan. 1960. “*Biṭrīḳ*”. *Encyclopaedia of Islam* I, 1249–1250. 2nd ed. Leiden: E.J. Brill.

——. 2002. *Byzantium and the Arabs in the sixth century*. Washington, D.C.: Dumbarton Oaks.

Shahîd, Irfan a.o. 1978. “*Ḳayṣar*”. *Encyclopaedia of Islam* IV, 839–840. 2nd ed. Leiden: E.J. Brill.

Sourdel, Dominique. 1960a. “*Balāṭ*”. *Encyclopaedia of Islam* I, 987. 2nd ed. Leiden: E.J. Brill.

——. 1960b. “*Barīd*”. *Encyclopaedia of Islam* I, 1045–1046. 2nd ed. Leiden: E.J. Brill.

Spuler, Berthold. 1965. “*Ḳunṣul*”. *Encyclopaedia of Islam* II, 60–61. 2nd ed. Leiden: E.J. Brill.

Teubner, J.K. 1997. “Chwaresmisch Kancik und das eurasisch-afrikanische Hemd/Camisia/Kanzu”. *Zeitschrift der Deutschen Morgenländischen Gesellschaft*, Suppl. III, 2.1084–1087.

Udovitch, Avraham. 1965. “*Fals*”. *Encyclopaedia of Islam* II, 768–769. 2nd ed. Leiden: E.J. Brill.

Ullmann, Manfred. 1997. “Zur Geschichte des Wortes *Barīd* (Post)”. *Sitzungsberichte der Bayerischen Akademie der Wissenschaften, Philosophisch-Historische Klasse* 1.1–77.

——. 2002. “*Praetorium*”. *Wörterbuch zu den griechisch-arabischen Übersetzungen*, 553. Wiesbaden: O. Harrassowitz.

Viré, F. 1978. “*ʾIṣṭabl*”. *Encyclopaedia of Islam* IV, 213–215. 2nd ed. Leiden: E.J. Brill.

IRFAN SHAHÎD (Georgetown University)

Leveling

1. DEFINITION

Leveling is defined by Blanc (1960:62) as a process that occurs in “inter-dialectal contact”. In such contacts, speakers may replace some features from their own dialect with those of another dialect that carries more prestige. The different dialect is not necessarily that of the listener. Blanc cites the example of villagers in central Palestine who may try to use the dialect of Jerusalem, or of non-Muslim Baghdadis who may try to move toward linguistic features of Muslim Baghdadis.

Leveling does not necessarily mean that the speakers will abandon their own dialect in favor of another one. They may, for example, select features which are more urban and abandon those which are more rural – and therefore more difficult for the listener to understand.

Thus, leveling, as defined by Blanc, is not just accommodation to a different dialect but also includes a process of unselecting local and rural features in favor of others which are easier to understand because they are more common.

This entry provides an in-depth explanation of the process of leveling. It includes examples of leveling, discusses factors that influence the occurrence of leveling, and explains the relation between leveling and language change.

According to Holes (1995:39), leveling is limited to the "elimination of very localized dialectal features in favour of more regionally general ones". Standardization (or 'classicization', as it is called by Blanc), on the other hand, is the elimination of local features in favor of standard ones. Standard or classical features are used not only to facilitate conversation but for other purposes as well. It is worth noting, however, that this distinction is not always maintained. Sometimes the term 'leveling' is used to refer to both standardization and classicization.

Versteegh (1993:65) defines leveling as a general process in which the differences between varieties of speakers that comprise a speech community have almost disappeared. This new variety of speech, which is the result of leveling, is different from all the specific existing varieties. He also uses the term 'koineization' as a term synonymous with leveling, and maintains that the process of koineization is in most cases connected with situations in which groups of speakers have been thrown together by accident. Thus, leveling, in his opinion, is usually an unplanned process. However, he mentions one case in which leveling was a planned process, namely that involving settlers in the new polders of North Holland that were reclaimed in the first half of the 20th century (Versteegh 1993:65).

Note that Versteegh prefers the term 'leveling' to that of 'koineization', which has diachronic connotations. The term 'leveling', on the other hand, is a sociolinguistic notion with synchronic connotations in a certain speech continuum. From an historical perspective, leveling could be used to describe the process of convergence between the early Arabic dialects, and it could also be used in connection with regional standards – in particular in the western part of the Arab world. The term 'leveling' refers to a process by which a mixture of dialectal varieties in a region take on the function of a prestigious substandard, often taking over from the existing varieties (Versteegh 1993:69). According to Versteegh, the emergence of a regional standard in Arab countries is connected with the predominance of the dialect of the capital, which imposes its variant on the surrounding areas. An example of this is the 'Moroccan dialect', meaning the dialect of Rabat, or the 'Egyptian dialect', meaning the dialect of Cairo. Again, according to Versteegh, this concept of the emergence of a dialect as a standard in a specific community is related to a process of leveling and may be its direct outcome.

2. Examples of leveling

In 1960, Blanc offered an example of leveling involving a conversation between Arabs from different countries. The speakers were discussing the future of the Arabic language and the differences between the various dialects. They then used leveling in order to facilitate their conversation.

According to Holes (1995:294), leveling can affect all linguistic levels: semantic, syntactic, phonological, etc. He gives the example of a conversation among educated Arabs from the Gulf, Baghdad, Cairo, and Jerusalem, and examines how these people express the existential 'there', as in 'there are people'. These speakers of Arabic had at least three dialectal ways of expressing existential 'there': Gulf: *hast*, *aku*, Baghdad: *aku*, Cairo: *fī*, Jerusalem: *fī*.

In this case, *fī* is the feature likely to be used by all of the speakers, since it has no association with any particular area and represents the nearest thing to a dialectal common meeting point for this specific group. Speakers in a "heterogeneous group tend to level their speech in the direction of a pan-Arab dialectal form". In such a case, Iraqis, for example, would resort to using a form that they do not have in their dialect.

Some linguists, like Versteegh, tend to study the process of leveling in relation to language change. Versteegh (2001:103) gives the historical example of the Arab armies. These armies consisted of a mix of different tribes, in which the existing differences between pre-Islamic dialects were leveled out. He posits that the new dialects in the conquered territories must have been the result of local independent evolution.

Continuing on this historical line, Versteegh

(2001:149) explains cases of leveling within the same dialect. He believes that in the Arabian Peninsula, the "nomadic sedentary dichotomy does not function in the same way as outside". This is because many tribes have settled members with whom there is frequent interaction, both economically and socially. Therefore, all dialects – including the sedentary ones – exhibit Bedouin features.

Versteegh (1993:70) also cites Egypt as an example of the process of leveling between speakers of the same country. He speaks about the influence of Cairene Arabic in the Nile Delta and mentions the example of the "isoglosses of the realisation of *q* and *j*". In Cairene Arabic, the reflex of Classical Arabic /q/ is /ʾ/ and that of Classical Arabic /j/ is /g/; in the Delta the reflex of /q/ is /g/ and of /j/ it is /j/. Versteegh thinks that there is a "formidable clustering of isoglosses" in Egypt. Many Egyptians will admit to modifying their dialect once they come in touch with the speech of the capital, and with time, they may give up their speech habits altogether. Note that Versteegh also mentions that Cairene Arabic is the language of the media, movies, and songs. Therefore, Cairene Arabic could become the future 'koineized' variety of the Arabic-speaking world. This is doubtful, however, because the influence of media is only one of a number of factors that influence leveling.

Versteegh (1993:72–75) also gives an example of leveling in relation to language change, in which there is an emergence of a regional standard. The example he offers is that of the development of → Juba Arabic in the southern Sudan. This example is significant, because Juba Arabic dialect displays ongoing decreolizing change in the development of aspectual and agreement marking of the verb. Versteegh refers to a thesis on Juba Arabic (Mahmoud 1979), which predicts that in the future, the linguistic variety or varieties spoken in Juba will become more and more similar to the Arabic spoken in Khartoum. Depending on the political situation, which determines the amount of exposure to standard Arabic, Juba Arabic may undergo the same equalizing influence of standard Arabic that all Arabic dialects undergo. This could result in a situation in which Juba Arabic would be nothing more than a regional variety of general Sudanese Arabic, without any trace of its creole origins. Versteegh concludes that if it is possible for a creolized variety to acquire through a process of semantic change features that pertain to normal dialects, one could conclude that the only way to distinguish between a decreolized and a normal dialect would be by an analysis of the historical facts connected with those varieties, since the linguistic structure does not offer us any clue to the genetic origins.

Versteegh's conclusion leads the linguist to wonder about the real origins of present-day regional varieties. He concludes that the leveling which takes place in Sudan and the rest of the Arab world proves that the emergence of a regional standard, when it occurs, is identical with the leveling process resulting from the influence of a prestigious variety of speech forms. For example, a creolized dialect may become decreolized to such a degree that it looks like a normal dialect.

3. FACTORS THAT INFLUENCE LEVELING

Because of the → diglossia in Arab countries, leveling is different there than in other countries (Ferguson 1959). Two language varieties in Arab countries exist side by side, each with its own function: standard Arabic and the vernacular of the individual country (as well as Classical Arabic, which is the language of the *Qurʾān* and pre-Islamic poetry). The position of standard Arabic is very strong, and it is difficult for any vernacular to replace it (Versteegh 2001:71). In most Arabic-speaking countries, Modern Standard Arabic (MSA) is the language of individual constitutions. Leveling in these countries may be different from leveling in other communities in which no language or variety has been given a special status. Versteegh believes that as a result of the special status of Modern Standard Arabic, inter-Arabic conversation in dialect will not converge in the direction of a regional dialectal variety but rather exhibit an increasing use of Modern Standard Arabic features. That is to say, leveling is not likely to lead to the ultimate disappearance of Modern Standard Arabic in favor of any vernacular.

Holes (1995:294) argues that leveling is a reaction to the dialectal differences between speakers whose aim is to emphasize shared elements and eliminate local ones. In addition, interaction between dialects – both economically and socially – also encourages the use of leveling (cf. Versteegh 2001).

Leveling is influenced by the following factors:

i. Stigmatization
Stigmatization can play a role in the direction of leveling. When there is interaction among different groups and one group's variety is stigmatized, leveling will take elements from the nonstigmatized variety. The following example clarifies this point. Woidich (1994) mentions that the Cairene dialect of today is a mixed dialect formed in the second half of the 19th century, when many people from the countryside moved to Cairo. A number of features became stigmatized as a result of being associated with low-prestige rural dialects. According to Versteegh (2001), this process of stigmatization led to the disappearance of rural forms and the emergence of new forms as a result of overgeneralization and leveling, as exemplified by the loss of the pausal → *ʾimāla*.

ii. Political and social issues
Al-Wer (2002b:45) discusses why pressures toward regional koineization will not lead the dialects of Beirut, Damascus, Jerusalem, and Amman to become identical. The reason she offers is that the countries of the Levant are separate political entities, with different political and social attitudes. Thus, political factors are important in the process of leveling.

iii. The issue of identity
The concept of identity also plays a major role in the process of leveling. Al-Wer (2002b:45; 2003) cites the example of Amman, where leveling has occurred as the country has acquired a new population – and, therefore, a new identity (→ Jordan Arabic). Among older population groups in Amman, the dialects can be clearly recognized as either Jordanian or Palestinian; however, among the new generation, leveling plays a role in decreasing regional differences and constructing a new dialect. Leveling here is used to increase localized and marked features, as well as to mark a symbol of a new identity.

iv. Markedness and simplification
Al-Wer (1997) mentions markedness and simplification as factors which influence the occurrence of a specific variable. She argues that a marked variable, which is associated with indigenous dialects, is not often used because people tend to simplify. This is connected to Myers-Scotton's (1997) idea of minimizing costs and maximizing rewards. She states that people try to facilitate the process of communication with the least effort on their part, at the same time trying to have a greater impact on the listener.

v. Context
Holes (1995:40) mentions that context plays a major role in the process of leveling. He asserts that in a supranational speech context, national dialects may be used, but with leveling and standardization. Holes adds that this kind of leveling may occur in certain contexts, such as a pan-Arab discussion on educational cooperation, depending on contextual factors (including the subject itself, which may have more or less strong associations with Modern Standard Arabic). Al-Wer (1997) also offers an interesting example of Palestinian and Jordanian dialects in contact. She claims that although the contact between the dialects is nothing new, the context for the interaction is new and different. Both groups form together the population of a new city, Amman, and thus share a new political entity. This context, Al-Wer argues, plays an important role in leveling.

vi. Media and education
Gibson (2002) mentions the fact that increased mobility and education influence leveling in Tunisia. He speculates that during the 20th century, nonstandard dialects became closer to the standard variety in many languages, including English. Gibson attributes this to the spread of media – print, television, and radio. He examines whether the same is true in the Arabic-speaking world, with reference to ongoing phonological and morphological changes in Tunisian Arabic. He admits that although there is a great influence of Modern Standard Arabic on the vocabulary of Arabic dialects – including Tunisian – the same may not be true for phonology and morphology. Gibson examines the assumption that because of the prestige of Modern Standard Arabic, as well as the spread

of education and media, change favors Modern Standard Arabic. He studies four Tunisian variables, including the increased use of /q/ instead of /g/, which is used in Bedouin dialects. One should bear in mind that the realization of this variable is shared between the urban dialect of Tunis and Modern Standard Arabic. Another variable Gibson examines is the treatment of the final vowel in weak verbs. He concludes that the direction of many language changes is toward the modern-day dialect of Tunis (Gibson 2002:28). In fact, in the case of the conjugation of weak verbs, changes are moving away from Modern Standard Arabic-like forms. This may be related to the spread of education and media.

4. Leveling in relation to language change

In drawing conclusions about leveling in relation to language change, an essential question must be considered, i.e., whether leveling has played a major role in explaining the process of decreolization and the formation of different dialects of Arabic. According to Holes (1986:221), leveling cannot explain the large difference between modern Arabic pidgins or creoles and mainstream dialects. He believes that similarities between modern dialects must stem from the main input, comprised of a range of slightly different dialects sharing similarities to Arabic, as opposed to being a unified form of Arabic. Versteegh (2004:352), on the other hand, stresses that the influence of the standard language should not be ruled out completely. Throughout the modern Arab world, one can find examples of illiterate speakers who can still produce standard forms (Palva 1969). This shows that the standard language may still have played a major role in current Arabic dialects. He also adds that at all levels, hybrid forms such as *b-tuktab* 'it [fem.] is being written' are heard (cf. Bassiouney 2003). The influence of Modern Standard Arabic is not just related to the spread of education and media but, according to Versteegh (2004), has religious significance as well, because it is used in mosques. The difference between Classical Arabic and Modern Standard Arabic is not discussed in detail here, because it does not contribute to the main argument; see Bassiouney (2003) for a full discussion of this difference.

5. Conclusion

A number of linguists who have studied leveling in Arabic argue that leveling is not necessarily in the direction of Modern Standard Arabic. Ibrahim (1986), Abdel Jawad (1987), Al-Wer (1997), and Gibson (2002) explain that Modern Standard Arabic is not a spoken variety. Thus, leveling does not necessarily have to be toward Modern Standard Arabic but could instead be directed toward the prestigious vernacular of different countries. Gibson also rejects the term 'prestige' when discussing leveling, since there is an overt and a covert prestige (cf. Trudgill 1974). Holes (1983) cites as an example to support Gibson the fact that despite increased literacy and urbanization in Bahrain, the local language has not moved toward Classical Arabic features. According to Al-Wer (2002b:46), linguistic change in the Arab world is determined by the status of the native varieties, which is in turn determined by the status of the speaker rather than the status of Modern Standard Arabic. Similarly, a higher level of education does not necessarily mean that speakers will use more standard Arabic forms.

Finally, it should be noted that leveling may occur in differing degrees. For example, Egyptians in an interdialectal context tend to accommodate their speech to others and use leveling less than Gulf Arabic speakers (→ speech accommodation). As stated above, leveling is not necessarily toward Egyptian Arabic in spite of the prestige of this dialect. This is because leveling is not only dependent on prestige but on a combination of factors. Versteegh (2001) offers another example to prove that there are degrees of leveling. He posits that the Bedouin dialects in the Arabian Peninsula are more conservative than those outside that region, the most conservative of them being Najdi Arabic. Therefore, they do not allow leveling to the same extent as other dialects outside the Peninsula do.

According to Al-Wer (2002b), the most important feature in leveling which may lead to language change is not education but rather frequency of interaction. Meanwhile, Versteegh (2004:355) argues that what is needed is "a much more detailed and fine-grained analysis of the demographic, cultural and social circumstances of the early period of Arabization". Since it is difficult to study change in progress (Gibson 2002:38), it is also difficult to study leveling in relation to language change. In fact, relatively few detailed studies on leveling in Arabic have been conducted, and there remains an urgent need to study this phenomenon from different perspectives and in relation to different variables, such as education, gender, social class, and country.

BIBLIOGRAPHICAL REFERENCES

Abdel Jawad, Hassan. 1987. "Cross-dialectal variation in Arabic: Competing prestigious forms". *Language in Society* 16.359–368.

Al-Wer, Enam. 1997. "Arabic between reality and ideology". *International Journal of Applied Linguistics* 7:2.251–265.

——. 2002a. "Jordanian and Palestinian dialects in contact: Vowel raising in Amman". *Language change: The interplay of internal, external and extra-linguistics factors*, ed. Mari C. Jones and Edith Esch, 63–79. Berlin: Mouton de Gruyter.

——. 2002b. "Education as a speaker variable". Rouchdy (2002:41–53).

——. 2003. "New dialect formation: The focusing of *-kum* in Amman". *Social dialectology in honour of Peter Trudgill,* ed. David Britain and Jessy Cheshire, 59–67. Amsterdam and Philadelphia: J. Benjamins.

Bassiouney, Reem. 2003. "Theories of code-switching in the light of empirical data from Egypt". *Perspectives on Arabic Linguistics* XV, ed. Dilworth B. Parkinson and Samira Farwaneh, 19–40. Amsterdam and Philadelphia: J. Benjamins.

Blanc, Haim. 1960. "Style variations in Arabic: A sample of inter-dialectal conversation". *Contributions to Arabic linguistics*, ed. Charles A. Ferguson, 81–156. Cambridge, Mass.: Harvard University.

Ferguson, Charles A. 1959. "Diglossia". *Word* 15:2.325–340.

Gibson, Maik. 2002. "Dialect levelling in Tunisian Arabic: Towards a new spoken standard". Rouchdy (2002:24–40).

Haak, Martine, Rudolf de Jong, and Kees Versteegh (eds). 2004. *Approaches to Arabic dialects: A collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday.* Leiden: E.J. Brill.

Holes, Clive. 1983. "Bahraini dialects: Sectarian differences exemplified through texts". *Zeitschrift für Arabische Linguistik* 13.27–67.

——. 1986. Review of Versteegh (1984). *Bibliotheca Orientalis* 43.218–222.

——. 1995. *Modern Arabic: Structures, functions and varieties.* London: Longman.

Ibrahim, Mohammad. 1986. "Standard and prestige language: A problem in Arabic sociolinguistics". *Anthropological Linguistics* 28.115–126.

Mahmoud, Ushari. 1979. *Variation and change in the aspectual system of Juba Arabic.* Ph.D. diss., Georgetown University.

Myers-Scotton, Carol. 1997. *Duelling languages.* Oxford: Oxford University Press.

Palva, Heikki. 1969. *Notes on classicisation in modern colloquial Arabic.* Helsinki: Snell-Manink.

Rouchdy, Aleya (ed.). 2002. *Language contact and language conflict in Arabic: Variations on a sociolinguistic theme.* Richmond: Curzon.

Trudgill, Peter. 1974. *The social differentiation of English in Norwich.* Cambridge: Cambridge University Press.

Versteegh, Kees. 1984. *Pidginization and creolization: The case of Arabic.* Amsterdam: J. Benjamins.

——. 1993. "Levelling in the Sudan: From Arabic creole to Arabic dialect". *International Journal of the Sociology of Language* 99.65–79.

——. 2001. *The Arabic language.* 2nd ed. Edinburgh: University of Edinburgh Press. (1st ed., 1997.)

——. 2004. "Pidginization and creolization revisited: The case of Arabic". Haak a.o. (2004:343–358).

Woidich, Manfred. 1994. "Cairo Arabic and the Egyptian dialects". *Actes des Premières Journées Internationales de Dialectologie Arabe de Paris*, ed. Dominique Caubet and Martine Vanhove, 493–507. Paris: INALCO.

REEM BASSIOUNEY (Georgetown University)

Lexical Variation: Modern Standard Arabic

1. LEXICAL VARIATION IN ARABIC

Lexical → variation is an area of sociolinguistics which studies differences in certain lexical items used by various speech communities. This phenomenon can be attributed to many variables, such as the nature of the language itself, geography, social status, individual preferences, topic, hierarchy, language academies, media, etc. → Modern Standard Arabic reflects the variability of one language still in use. It can be traced back centuries to the pre- and early-Islamic periods and found in the lexical variability among the numerous Arab tribes.

Lexical variation continues to exist in Modern Standard Arabic because it is of the used language variety and thus cannot remain static. A certain lexical item may be used in one Arab country while another is used in other Arab countries. In some cases, these lexical variations are simply a different form of the same root, while in other cases they are a completely different word. There are several elements contributing to this phenomenon, such as lexical preferences, → diglossia, the activities of the

Arabic → language academies and the media, and the influence of → Educated Spoken Arabic (Ibrahim 1997).

Modern Standard Arabic has been defined as "the mutual official language in the Arab world today" (ʿAbd al-ʿAzīz 1992:11), and as "the language uniting the Arab countries" (aṣ-Ṣāyiġ 1990:20). It is sometimes referred to as *fuṣḥā l-ʿaṣr* and defined as occurring "in all the subjects that deal with our contemporary life, and in which *fuṣḥā* Arabic is used. It is used to broadcast political commentaries on the radio, and found in prepared scientific speeches that are read from a written paper. Since this level is wide and comprehensive, it includes different types such as the scientific, political, literary, social…etc." (Badawi 1973:90). Modern Standard Arabic is also defined as "the variety of Arabic that is used in newspapers, magazines, textbooks, academic books, novels, short stories and other 'serious' writing. It is used orally in some university contexts, in political and other 'read' speeches and in the delivery of the news on radio and television" (Parkinson 1991:32).

Lexical variation in Modern Standard Arabic is not a new phenomenon in the Arabic language; on the contrary, it has existed for centuries. The morphological system of Modern Standard Arabic holds inherent characteristics which enable the formation of different lexical items derived from the same root. It also allows for the existence of different words carrying exactly the same meaning. The diglossic nature of the Arabic language influences the choice of words in Modern Standard Arabic. Personal preferences have always been and will continue to be a factor in the process of word choice.

2. Variation in Classical Arabic

Lexical variation has existed in the Arabic language as early as the pre-Islamic period (→ pre-Islamic Arabic). At the time of the Islamic Revelation, there was a certain degree of dialectal variation between the tribal groups living in the Ḥijāz and Najd (Holes 1995:14), leading Rabin (1951:1–2) to conclude that the Eastern dialects (Tamīm, ʿUqayl, and others) "must have sounded to the Arab from Najd like a foreign language". Linguists such as Rabin attribute the lexical variation to dialectal variation between tribes. Others, such as Versteegh (1997), attribute it to stylistic variation.

Numerous anecdotes are reported about lexical variation in the pre-Islamic period. For example, it is said that when a knife fell from the Prophet Muḥammad, he told ʾAbū Hurayra to hand it to him, using the word *sikkīn* for 'knife'. ʾAbū Hurayra did not understand, so he asked the Prophet to repeat his words three times. Finally, ʾAbu Hurayra said, "Is it the *mudya* 'knife' that you want?", and the Prophet answered "yes". Then ʾAbū Hurayra asked, "Is it called *sikkīn* among your tribe?". And he continued, "I had never heard it until that day" (ʾAnīs 1973:138–139).

Another story is told about ʿAlī ibn ʾAbī Ṭālib, cousin and son-in-law of the Prophet Muḥammad, who heard him speaking to a delegation from the tribe of Nahd (Banū Nahd). ʿAlī said to the Prophet, "We hear you speaking to the Arabs, and we are from the same father [i.e., 'we speak the same language'], but we do not understand much of what you say" (ʿAṭṭār 1958:26).

Lexical variation is also referred to in the words of the Prophet, "The *Qurʾān* was revealed in seven letters" (*nazala l-Qurʾān ʿalā sabʿati ʾaḥruf*; Naṣṣār 1956:79). The seven 'letters' are the seven different ways of reading the *Qurʾān* (→ *qirāʾāt*). The variations are due to dialectal and pronunciation differences among the tribes (Naṣṣār 1956:79). The *Qurʾān* includes dialectal words from the various tribes (al-Munjid 1946:8). For example, in sura *ʾĀl ʿImrān* (Q. 3/39), there is the following example of a lexical item from different tribes: *ṣabūr* means 'a patient man' in the tribe of ʿAmīr, but in the tribe of Kināna it means 'a man who does not need women'. There are also phonetic variations in the realization of the word *massakum qarḥ* 'you were wounded', as the Tamīm tribes pronounced it *qurḥ* and the Ḥijāzī tribes pronounced it *qarḥ*.

Examples of phonological variation are found in the variation in short vowels, e.g. *nastaʿīn* and *nistaʿīn* 'we seek help', one with a *fatḥa* in the prefix and the other with a *kasra*. According to al-Farrāʾ (Naṣṣār 1956:77–78), it was pronounced with a *fatḥa* in the language of the tribes of ʾAsad and Qurayš (the tribe of the Prophet Muḥammad), while other tribes say it with a *kasra* (→ *taltala*). Another example is found in cases of deletion, e.g. *maʿakum* 'with you' and *maʿkum*, and in *istaḥyaytu* 'I was shy' and *istaḥaytu*, with deletion of the first *y*.

Examples of sound substitution are *ʾulāʾika* 'those' and *ʾulālika; ʾammā* 'as for' and *ʾāmma*, where the *hamza* is substituted with a long vowel /ā/; deglottalization of the *hamza*, as in *mustahziʾūn* 'those who ridicule' and *mustahzūn*; → metathesis, as in *ṣāʿiqa* 'thunderbolt' and *ṣāqiʿa*; → assimilation, as in *muhtadūn* 'those rightly guided' and *muhdawn*; lengthening of the vowel, as in *unḏ̣ur* 'look!' and *unḏ̣ūr*.

An example of grammatical variation is that some Arabs say *hāḏihi l-baqar* 'these [fem.] cows', while others say *hāḏā l-baqar* 'these [masc.] cows'. An example of lexical variation in the plural form is *ʾasrā* 'prisoners of war' and *ʾasārā*. Variation in verbs involves the famous story about the imperative *ṯib* 'rise!' or 'jump!', which in → Ḥimyaritic meant 'sit down!' (Naṣṣār 1956:77–78).

Ibn Jinnī (d. 392/1002), who has been described in the *Encyclopaedia of Islam* (III, 754) as the founder of the science of etymology (→ *ištiqāq*), wrote a chapter about variation in his *Xaṣāʾiṣ* (I, 370), with the title *Fī l-faṣīḥ yajtamiʿu fī kalāmihi luġatāni fa-ṣāʿidan* 'About eloquent persons in whose speech there are two or more dialectal forms'. He wrote that if one says *baġdād* 'Baghdad' or *baġdān*, or even *miġdān*, all these variations are correct. Likewise, one can say both *liḥya* or *ʾaym* for 'beard', both being correct.

A case of lexical variation in poetry is reported by Ibn Jinnī (*Xaṣāʾiṣ* II, 467), about a poet using *zabn* 'narrowness, tightness, closeness' in one verse, and then reciting *ḍīq* in another line. A friend of the poet remarked, "This is not how you said it before, you used *zabn* before". The poet answered, "Don't you know that *zabn* and *ḍīq* are one [i.e. 'have the same meaning']?".

Variation of this kind existed in pre-Islamic times and has continued since then. Dialectal variation among the Arab tribes led to the existence of many words with the same meaning but differing in form (ʾAnīs 1973:138). Obviously, lexical variation is not a new phenomenon in the Arabic language.

3. LEXICAL VARIATION IN MODERN STANDARD ARABIC

3.1 *Lexical items from the same root*

The following are examples of words which come from the same root yet in some instances carry two or more meanings. In some Arab countries one derived form of the verb is used, while in other countries another form is used. The following examples illustrate specific cases of preference on the part of some Arabs for a certain lexical item or form over another, despite the fact that they are derived from the same root.

i. *Taḏ̣āhurāt* and *muḏ̣āharāt* 'demonstrations' both come from the root *ḏ̣-h-r* 'to be or become visible, clear, apparent' (Wehr 1980:583). Form III of this root, *ḏ̣āhara*, means 'to help, assist, aid, support', verbal noun *muḏ̣āhara* (Wehr 1980:584). In Form VI (*tafāʿala*), the root means 'to manifest, display, show, pretend', verbal noun *taḏ̣āhur* (Wehr 1980:584). Both verbal nouns are used to mean 'demonstrations'. Some speakers choose to use *muḏ̣āharāt* in order to avoid confusion with the word meaning 'to pretend'. In the sense of 'demonstrations, riots', *muḏ̣āharāt* is used in Egypt, while *taḏ̣āhurāt* is used in this sense in Lebanon; the latter word refers in Egypt to art exhibitions, musical concerts, or film festivals, just as it does in Morocco.
ii. *Istixbārāt* and *muxābarāt* 'intelligence' both come from the root *x-b-r* 'to try, test, experience, know by experience, have experience' (Wehr 1980:225). In Form X (*istafʿala*), this root means 'to inquire, ask about', verbal noun *istixbār* (Wehr 1980:225). In Form III (*fāʿala*), the root means 'to contact in writing, negotiate, treat', verbal noun *muxābara* (Wehr 1980:225). Both verbal nouns are used in the plural to mean 'intelligence'. In Lebanon and Morocco, *istixbārāt* is used, while *muxābarāt* is used in Egypt.
iii. *Taxṣīṣa, xaṣxaṣa*, and *xawṣaṣa* are all three used for 'privatization'. The first and third word come from the same root, *x-ṣ-ṣ* 'to particularize, designate, allocate to' (Wehr 1980:240), *taxṣīṣa* being the verbal noun of Form II of the verb, while *xawṣaṣa* is from the form *fawʿala*. The second word, *xaṣxaṣa*, is from a quadrilateral root. In Egypt, *xaṣxaṣa* is preferred, while in Lebanon one uses *taxṣīṣ* and in Morocco *xawṣaṣa*.
iv. *Makman* and *kamīn* 'ambush'. The root is *k-m-n*, used in Form I as *kamana* or *kamina* 'to hide; to be hidden, concealed, latent; to ambush' (Wehr 1980:841). *Kamīn* means 'hidden, lying in ambush, ambush, secret attack' (Wehr 1980:841), while *makman* means 'place where something is hidden,

ambuscade, ambush hiding place' (Wehr 1980:841). The pattern of *makman* is *mafʿal*, which is used to form the nouns of places. In Lebanon, *makman* is used, while *kamīn* is used in Egypt.

3.2 *Lexical items from different roots*

The examples provided below are cases of lexical preferences between synonyms.

i. *Istanfara* and *istadʿā* 'to put on alert'. The verb *istanfara* (Form X) means 'to be frightened away, to call upon, to fight, to call out' (Wehr 1980:984). The verb *istadʿā* (Form X) means 'to call or send, to summon, to recall [e.g. a diplomatic envoy]' (Wehr 1980:283). Thus both not only share the same meaning, they also share the same Form, *istafʿala* (Form X). In Egypt, *istiʿdād* (or *taʾahhub*) used to be the term for 'alert', often in combination with the adjective *quṣwā* 'ultimate', while in Lebanon and Morocco *istinfār* was used. However, from the year 2000 onward, Egyptian newspapers have switched to *istinfār*.

ii. *Humūm* and *mašākil* 'concerns; problems'. *Humūm* is the plural of *hamm* 'anxiety, concern, solicitude, worry' (Wehr 1980:1033), while *mašākil* is the plural of *muškila* 'problem, unsolved question, issue, difficulty' (Wehr 1980:483). In Lebanon, *humūm* is preferred, while in Egypt *mašākil* is more common in this sense.

iii. *Tatimma* and *baqiyya* are used in newspapers and magazines in the sense of 'continuation'. The literal meaning of *tatimma* is 'completion' (Wehr 1980:97); this is the current term in Lebanon and Morocco. The literal meaning of *baqiyya* is 'remainder, rest' (Wehr 1980: 69); this is the current term in Egypt.

iv. *Mašāġil* and *wiraš* 'workshops', the former being the plural of *mašġal* (Wehr 1980:476), the latter the plural of *warša* (Wehr 1980:1061). *Mašāġil* is used in Lebanon, while *wiraš* is used in Egypt.

4. CAUSES OF VARIATION

4.1 *Influence of dialect*

Diglossia plays a decisive role in the choices illustrated in the following examples. Speakers of a dialect sometimes prefer to restrict one particular lexical item to the dialect and another to Modern Standard Arabic to avoid confusion.

i. *ʾIjāza* and *šahāda* 'university degree'. The verb *ʾajāza* means 'to permit, allow; to license; to approve' (Wehr 1980:147). In Lebanese and Moroccan Modern Standard Arabic, *ʾijāza* is used for 'degree', while in Egyptian *šahāda* is used in this sense; *ʾajāza* is a false cognate because in Egyptian colloquial it means 'holiday, vacation'. Since the short vowels /i/ and /a/ do not occur in printing, Egyptians would read the word as *ʾagāza* (the /ž/ sound in Lebanese corresponds to /g/ in Egyptian). The word *šahāda* 'certificate' may come from the verb *šahada* 'to certify', which is usually written in contexts such as 'The Ministry of... certifies...'. *ʾIjāza* appears to be a translation from the French word *license*, meaning a bachelor of arts degree, a term still used in Lebanon.

ii. The oldest Arabic dictionary, al-Xalīl's *Kitāb al-ʿayn* states that "the two thighs (*wirkāni*) are above the two legs (*faxḏāni*)" (V, 403). This difference has become obsolete in Modern Standard Arabic, and both *wirk* and *faxḏ* are now used to refer to the same body part, 'thigh'. However, in written Lebanese Modern Standard Arabic, only *wirk* is used, while *faxḏ* is the corresponding word in written Egyptian Modern Standard Arabic, *wirk* being reserved for the colloquial.

iii. *Manāšīr* and *manšūrāt* 'leaflets' are two plurals for the singular *manšūr*, from the root *n-š-r* 'to spread out, unfold, publicize, publish' (Wehr 1980:965). However, in the Egyptian dialect, *manāšīr* is used as the plural of *minšār* 'saw [carpenter's tool]', leading Egyptians to prefer *manšūrāt* as the plural of *manšūr*.

iv. *Kulfa* and *taklifa* 'expenditure'. *Kulfa* means 'expenditure; trouble, fuss [for the benefit of guests]; trimming, garniture [of a dress or shirt]' (Wehr 1986:750). Egyptians limit the usage of the word *kulfa* to the dialect when it refers to the trimming of a dress, and prefer *taklifa* to refer to 'expenses' in both Modern Standard Arabic and dialect.

v. Both *ḥaṣala* and *ḥadaṯa* mean 'to happen'. Since some dialects use the verb *ḥaṣala* in their colloquial, they tend to use *ḥadaṯa* in written Modern Standard Arabic.

4.2 *Choosing different meanings of the same word*

The following examples illustrate the process of word choice between two meanings of a word. One group of speakers chooses a certain word, while another group chooses a different meaning.

i. The verbal noun *taʿāṭī* comes from the verb *taʿāṭā* (Form VI) 'to take, swallow, take medicine; to be occupied or busy with, undertake'. *Taʿāṭī* is, therefore, used to indicate two things, either 'to be occupied with, to be busy with' (Wehr 1980:622), or 'to take medicine'. In the former sense it is used in Lebanon and Morocco, in the latter in Egypt.
ii. *Takrīs* 'dedication, devotion', the verbal noun of Form II *karras* (Wehr 1980:820). However, it is used in Egypt to mean 'dedication', while in Morocco and Lebanon it means 'stabilization' or 'continuity'.
iii. *Jahawiyya* is used in Morocco for 'side, direction, region, part, section, area, district' (Wehr 1980:1052), while in Egypt it only means 'side, direction'. In Egypt, *ʾiqlīmiyya* is used instead for 'district' (from *ʾiqlīm* 'climate, area, region, province'; Wehr 1980:210).

4.3 *Influence of translation*

As a result of the French occupation of Lebanon, the French language exerted linguistic influence on Lebanese Arabic. This influence is clear in the two words *tawqīf* and *ʾijāza*. Both words are used in Lebanon. There is evidence that the word *ʾijāza* is a translation of the French word *license*, and the same applies to the word *tawqīf* 'arrest'. *Tawqīf* is the verbal noun of the Form II verb *waqqafa*, and it means 'rising, apprehension, seizure, arrest, parking' (Wehr 1980:1093). It may be derived from the French word *arrête* 'stop'. The verbal idiom *qabaḍa ʿalā* means 'to arrest', as well.

The influence of French on the written realization of Modern Standard Arabic in the Maghreb was studied extensively by Kropfitsch (1977, 1980). He mentions obvious manifestations of French influence, as in the names of the months (e.g. *jwilya* 'July', *ʾūt* 'August', instead of the current forms in the Mashreq, *jūliyū*, *ʾaġustus*), the use of French loanwords, such as *barṭmān* (< French *appartement*) 'apartment' or *tambar* 'stamp' (< French *timbre*), and the use of names of countries with the article, such as *al-bīrū* (French *le Pérou*), *at-tšād* (French *le Tchad*), *al-viyetnām* (French *le Vietnam*). In addition, French influence is found in the → semantic extension of words, e.g. when the adjective *jihawī* is used in the sense of French *régional*; when *ʾijrāʾāt* is used not only in the sense of 'measures' but also in that of 'regulations' (cf. French *mésures*); when *ḥuqūq* is used in the sense of 'sees' (French *droits*; Mashreq *rusūm*); or when *ʾiṭārāt* is used for persons, just like French *cadres*. Loan formations from French in Standard Arabic as it is found in media Arabic in the Maghreb include *šarika majhūlat al-ism* 'company with limited liability' (French *société anonyme*; Mashreq *šarika ḏāt masʾūliyya maḥdūda*), *sāmī l-muwaḏ̣ḏ̣afīn* 'high-ranking officials' (French *hauts fonctionnaires*), and even verbal constructions such as *waḍaʿa fī l-istiʿmāl* 'to put to use' (French *mettre en usage*). Such lexical items contribute to the reader's intuitive feeling of a difference between the language of the press in the Maghreb and the Mashreq.

4.4 *Different usages of verbal nouns*

When verbal nouns are abstract and do not indicate gender, tense, or number (Ḥassān 1960:III, 183), they should not be plural, but if used as an ordinary noun, they can be plural (Ḥassān 1960:I, 24). To avoid confusion, in Egypt, for instance, the plural of such nouns is formed through the 'artificial verbal noun' (*al-maṣdar aṣ-ṣināʿī*), so that the plural of *iqtiṣād* becomes *iqtiṣādiyyāt*. This kind of verbal noun is formed by suffixing *-iyyāt* to the singular noun. *Ittifāq* means 'treaty, agreement', just like *ittifāqiyya*. The plural of the first is *ittifāqāt*, and the plural of the second is *ittifāqiyyāt*. Both words are derived from the same root, *w-f-q*. According to Badawi a.o. (2004:749), "There is now a massive number of new nouns with the *-iyya* suffix, exploiting a C[lassical] A[rabic] mechanism for creating abstract nouns by this means". Their examples include *iqtiṣādiyyāt* 'economics' (< *iqtiṣād* 'economy', which is both a noun and a verbal noun; as a noun it can occur in the plural, *iqtiṣādāt*); *siyāsiyyāt* 'political science' (< *siyāsa* 'politics, policy').

There is another type of noun or verbal noun, to which the 'artificial verbal noun' (*al-maṣdar aṣ-ṣināʿī*) cannot apply, e.g. *najāḥāt*, the plural of the verbal noun *najāḥ* 'success'; *insiḥābāt*, the plural of *insiḥāb* 'withdrawal'; and *ʾiʿdāmāt*, the plural of *ʾiʿdām* 'execution'.

4.5 *Differences in usage of verb preposition*

Some Arab countries, including Lebanon and Morocco, tend to use the verb *iltaqā* 'to meet' as a transitive verb, without a preposition; others use it with the preposition *bi-* 'with', or with the preposition *maʿa* (Wehr 1980:876). This again can be a result of translation from American English, as British English uses the preposition *with* with the verb *to meet*.

In Lebanon, the verb *baḥaṯa* 'to discuss' is used with the preposition *fī* 'in', while in Egypt it is used without a preposition, but *tabāḥaṯa* 'to discuss mutually' is preferred with the preposition *fī*. The opposite situation occurs in Lebanon, where other verbs, such as *iltazama* 'to adhere' and *taʿāhada* 'to promise', are used without a preposition, while these verbs are commonly used with the preposition *bi-* 'with' in Egypt.

5. THE PRESENT LINGUISTIC SITUATION

All of the above are examples of the lexical variation process that takes place in Arab countries. The → language academies and the interference of the colloquial language through the use of → Educated Spoken Arabic play an important role which is quite different from the influence of the language of the → media. Language academies were formed to discuss urgent issues of the Arabic language and to ensure that Arabic could be used in all scientific fields. There are numerous organizations in the Middle East, including the linguistic organization that was established in Beirut in 1882 by Faris Namir under the name of the Eastern Scientific Academy; language academies in Damascus, Cairo, Baghdad, Amman, and Jordan; and the Rabat-based Bureau of Arabization, affiliated with the Arab League.

Although the academies have played a major role since the end of the 19th century, several problems remain (→ terminology). First, the coining of a new word, or the translation of another, often must be done on the spot, and when the academies are not in session. For example, the words *sayyāra* 'car' and *ṭāʾira* 'plane' were innovations of the media, which used the process of *qiyās* 'analogy' (aṣ-Ṣāʾiġ 1992:186) to create new words. Sometimes, old pre-Islamic words were revived, such as *hātif* 'ghost voice whispering in the desert' for 'telephone' or *qiṭār* 'caravan, line of camels' for 'train'. In some cases, these attempts succeeded, but in others they did not (for instance, in spite of the efforts of the academies, the French loanword *fārāt* is almost universally used for the 'headlights' of a car). Second, these academies rarely achieve consensus among themselves concerning the choice of a new word. For example, 'television' is *telfāz* in Lebanon, while it is *tilivizyōn* in Egypt; the official term for 'radio' is *mawwāj* in Lebanon, while it is *miḏyāʿ* in Egypt; in Lebanon, a computer is officially called *ḥāsūb*, while the official term in Egypt is *ʿaql ʾiliktrūnī* (along with *kumbyūtar*, obviously). Finally, the impact of their decisions is slow, and, in many instances, their recommendations remain theoretical.

Educated Spoken Arabic is one level of the Arabic language in which features of Modern Standard Arabic and a dialect exist. Speaking and writing are separate yet connected traits, and educated speakers are influenced by these speaking traits when they write. Thus, regional variations that exist in speech may well extend to writing. Mitchell (1986:8) refers to "the interplay between written Arabic and vernacular Arabic(s) that creates and maintains Educated Spoken Arabic (ESA), both nationally and internationally. A shared standard arising from a modern literary tradition has to be supported by a wide educated public and appear in their speech as well as in their writing".

In spite of the differences between Educated Spoken Arabic and Modern Standard Arabic, they still share many similar aspects. Both are an elevated level of the language, not used by illiterate people. Lexical choices in both are of *fuṣḥā* words and not colloquial words. In addition, since both are used daily in the media, whether written or oral, they must deal with rapidly changing, everyday issues. Moreover, there is a continuous interchangeable effect between the two. Journalists are educated and are expected to speak Educated Spoken Arabic during formal occasions. They are also expected to write in Modern Standard Arabic. They face urgent and sudden demands to either translate or coin new words. These newly coined words are found in Modern Standard Arabic and are used immediately in Educated Spoken Arabic. For example, the word *barīd* 'mail' was coined by journalists (Šaraf 1980:153).

Journalists, as speakers of Educated Spoken Arabic and writers of Modern Standard Arabic, are extremely influential in the area of lexical variation. This continuum in the journalistic linguistic repertoire plays a very important role in illustrating the variation in lexical items in the Arabic language. Journalists are impacted by their regional affiliation. Consequently, the tendency to use lexical items that are used in one's country dominate written Modern Standard Arabic. For example, a Lebanese journalist would use *tawqīf* 'arrest' instead of *al-qabḍ ʿalā*, and *tatimma* 'continuation' rather than *baqiyya*, unlike his Egyptian counterpart. In newspaper writing, a mobile phone is referred to as *maḥmūl* (< *ḥamala* 'to carry') in Egypt, while it is *jawwāl* (< *jāla* 'to roam, wander about, move freely') in the Gulf and *xalyawī* (< *xaliyya* 'cell') in Lebanon.

Interestingly, many educated Arabs, when reading articles in a newspaper, immediately perceive the origin of the writer on the basis of the lexical items used. As anecdotes are reported from the early Islamic era, there is anecdotal evidence from the modern period as well. Thus, for instance, Moroccan telephone booths carry the sign *mixdaʿ hātifī*, where *mixdaʿ* originally means 'small room, chamber cabinet, bedchamber' (Wehr 1980:229), leading Egyptians and Lebanese to wonder why bedrooms are being used as telephone booths in Morocco, because in their lexicon the word is only used in this sense.

In fact, several of the examples of polysemy mentioned by Ibrahim (2005:51–64) for Arabic dialects can be extended to the lexical variation in Modern Standard Arabic, especially those in the domain of food. For instance, the two Egyptian terms *bāmiya* 'okra' and *muluxiyya* 'Jewish mellow' have a reversed meaning in Morocco. Likewise, a report about an Egyptian eating *lūbiyā* 'black-eyed peas' would be interpreted by a Lebanese as a report about *faṣulyā* 'cowpea' (Wehr 1980:882, 692).

The existence of large newspaper corpora makes it possible nowadays to supplement intuitive judgments by native speakers with statistical analysis. Thus, the example of the different meanings of *taḍāhurāt* in Lebanon and Egypt, mentioned in Section 3.1, can be validated by a count in the arabiCorpus of Brigham Young University: in the Egyptian *al-ʾAhrām*, the word occurred only 47 times in one year (1999), while in the Lebanese *al-Ḥayāt* it occurred no fewer than 714 times in one year (1977), almost always in the sense of 'demonstrations'.

6. CONCLUSION

At the beginning of the 1990s, Fasold (1990:II, 240) wrote that "geography becomes an issue when the community that supplies the new influence is some distance from the one receiving the innovations". However, lexical variation is not geographically limited to its place of birth but rather extends throughout the Arab world by means of the satellite television channels. These channels have helped the spoken word of a certain area to spread to other areas and then be used in the local written media, as shown by the examples of *istinfār* 'putting on alert', which is now current in Egyptian media, and the term *ḥāsūb* 'computer'.

One of the main factors in the spread of lexical items is the country where an event took place. News starts from the original country and spreads to other countries through the different news → media. Egypt always had the most influential role, linguistically speaking, of all Arab countries, due to its movie industry, in addition to its TV soap operas and songs, and, of course, its political role in the region and the many political events taking place there. After the emergence of satellite channels such as Al-Jazeera, though, this role is being eroded, in addition to the fact that nowadays, most of the political news comes from other Arab countries rather than from Egypt. This results in mutual linguistic influence among all Arab countries.

All of the above-mentioned elements, such as language academies, Educated Spoken Arabic, and media, in addition to diglossia, affect each other consciously or unconsciously. The innovations are continuing to appear and the spread is continuing as well, depending on the geographical source of the events. National identity, therefore, plays an important role in the choice of words, in addition to the role of the media and the language academies. Each Arab country coins its own words without prior consensus or planning among the different institutions. This point was stressed by Elkhafaifi, who states in his discussion of → language policy that "in effect, each country is developing its own terminologies (if indeed, development is taking place at all)" (2002:258).

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Ibn Jinnī, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. Ed. Muḥammad ʿAlī an-Najjār. 3 vols. Cairo: Dār al-Kitāb al-Miṣrī, 1952–1956.

Xalīl, *ʿAyn* = ʾAbū ʿAbd ar-Raḥman al-Xalīl ibn ʾAḥmad al-Farāhīdī, *Kitāb al-ʿayn*. Ed. Mahdī al-Maxzūmī and ʾIbrāhīm as-Sāmarrāʾī. 8 vols. Beirut: Muʾassasat al-ʾAʿlamī li-l-Maṭbūʿāt, 1988.

SECONDARY SOURCES

ʿAbd al-ʿAzīz, Muḥammad Ḥasan. 1992. *al-Waḍʿ al-luġawī fī l-fuṣḥā al-muʿāṣira*. Cairo: Dār al-Fikr al-ʿArabī.

ʾAnīs, ʾIbrāhīm. 1973. *Fī l-lahajāt al-ʿarabiyya*. 4th ed. Cairo: Maktabat al-ʾAnglū al-Miṣriyya.

Apte, M.L. 2001. "Language in society: Overview". *Concise encyclopedia of sociolinguistics*, ed. Rajend Mesthrie and R.E. Asher, 36–48. Amsterdam: Elsevier Science.

ʿAṭṭār, ʾAḥmad ʿAbd al-Ġafūr. 1958. aṣ-Ṣiḥāḥ *wa-madāris al-muʿjamiyya al-ʿarabiyya*. Cairo: Dār al-Kitāb al-ʿArabī.

Badawi, Elsaid [as-Saʿīd Badawī]. 1973. *Mustawayāt al-ʿarabiyya al-muʿāṣira*. Cairo: Dār al-Maʿārif.

——, Michael G. Carter, and Adrian Gully. 2004. *Modern Written Arabic: A comprehensive grammar*. London and New York: Routledge.

Bell, Roger. 1976. *Sociolinguistics: Goals, approaches and problems*. London: B.T. Batsford.

Bright, William and A.K. Ramanujan. 1972. "Sociolinguistic variation and language change". *Sociolinguistics*, ed. J.B. Pride and Janet Holmes, 157–166. Harmondsworth: Penguin Books.

Elkhafaifi, Hussein M. 2002. "Arabic language planning in the age of globalization". *Journal of Language Problems and Language Planning* 26:3.253–269.

Fasold, Ralph W. 1978. "Language variation and linguistic competence". *Linguistic variation: Models and methods*, ed. David Sankoff, 85–95. New York: Academic Press.

——. 1984. *The sociolinguistics of language*, I. Cambridge: Blackwell.

——. 1990. *The sociolinguistics of language*, II. Cambridge: Blackwell.

Gal, Susan. 1978. "Peasant men can't get wives: Language change and sex roles in a bilingual community". *Language in Society* 7:1.1–16.

Ḥassān, ʿAbbās. 1960. *an-Naḥw al-wāfī*. 4 vols. 10th ed. Cairo: Dār al-Maʿārif.

Holes, Clive. 1983. "Patterns of communal language variation in Bahrain". *Language in Society* 12:3.433–457.

——. 1995. *Modern Arabic: Structures, functions and varieties*. New York: Longman.

Ibrahim, Zeinab. 1997. *Egyptian and Lebanese Written Modern Standard Arabic: Are they one and the same?* Ph.D. diss., Georgetown University.

——. 2005. "Polysemy in the Arabic dialects". *Investigating Arabic*, ed. Alaa Elgibali, 51–64. Leiden: E.J. Brill.

Kropfitsch, Lorenz. 1977. "Der französische Einfluss auf die arabische Schriftsprache im Magrib". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 128.39–64.

——. 1980. "Semantische Tendenzen im Neuhocharabischen". *Zeitschrift für Arabische Linguistik* 5.118–136.

Labov, William. 1983. "The boundaries of words and their meanings". *Variation in the form and use of language*, ed. Ralph Fasold, 29–62. Washington, D.C.: Georgetown University Press.

Milroy, James. 1992. *Linguistic variation and change*. Oxford: Blackwell.

Milroy, Lesley. 1987. *Language and social networks*. Oxford: Blackwell.

Mitchell, T.F. 1986. "What is Educated Spoken Arabic?". *International Journal of the Sociology of Language* 61.7–32.

Mobbs, Michael C. 1981. "Two languages or one? The significance of the language names 'Hindi' and 'Urdu'". *Journal of Multilingual and Multicultural Development* 3.203–211.

Munjid, Ṣalāḥ ad-Dīn al-. 1946. *Kitāb al-luġāt fī l-Qurʾān*. Cairo: Maṭbaʿat ar-Risāla.

Nader, Laura. 1962. "A note on attitudes and the use of language". *Anthropological Linguistics* 4:6. 25–2.

Naṣṣār, Ḥusayn. 1956. *al-Muʿjam al-ʿarabī: Našʾatuhu wa-taṭawwuruhu*. Cairo: Dār al-Kitāb al-ʿArabī.

Nelson, John and Jacques Rebuffot. 1984. "A sociolinguistic study of Montreal". *Canadian Modern Language Review* 40:3.360–373.

Parkinson, Dilworth B. 1991. "Searching for modern *fusha*: Real-life formal Arabic". *al-ʿArabiyya* 24.31–64.

Preston, Dennis (ed.). 1993. *American dialect research*. Amsterdam and Philadelphia: J. Benjamins.

Rabin, Chaim. 1951. *Ancient West-Arabian*. London: Taylor's Foreign Press.

Romaine, Suzanne. 1982. *Socio-historical linguistics: Its status and methodology*. Cambridge: Cambridge University Press.

Ṣāʾiġ, Fāʾiz aṣ-. 1992. *al-Luġa wa-t-taʿrīb wa-dawr al-ʾiʿlām*. Damascus: Manšūrāt Dār Majallat aṯ-Ṯaqāfa.

Šaraf, ʿAbd al-ʿAzīz. 1980. *al-Luġa al-ʾiʿlāmiyya*. Cairo: al-Markaz aṯ-Ṯaqāfī al-Jāmiʿī.

Ṣāyiġ, Mājid al-. 1990. *al-ʾAxṭāʾ aš-šāʾiʿa wa-ʾaṯaruhā fī taṭawwur al-luġa al-ʿarabiyya*. Beirut: Dār al-Fikr al-ʿArabī.

Trudgill, Peter. 1974. *Sociolinguistics: An introduction*. New York: Penguin Books.

Versteegh, Kees. 1997. *The Arabic language*. New York: Columbia University Press.

Wehr, Hans. 1980. *A dictionary of Modern Written Arabic*. Trans. Milton Cowan. Repr. Beirut: Librairie du Liban.

Zeinab Ibrahim
(The American University in Cairo)

Lexicography: Bilingual Dictionaries

This entry deals with bilingual dictionaries which have Arabic as the source or target language. The first bilingual dictionaries developed within the Arabic lexicographical tradition from 11th-century monolingual dictionaries (→ lexicography: Classical Arabic). In the 17th century, European Orientalists started to take part in Arabic lexicography (→ Arabic studies in Europe). The earliest representatives of this European development followed the original Arabic tradition closely. Only at a later stage did they join the Western lexicographical tradition.

1. Historical overview

According to Haywood (1965), in the early period of Arabic lexicography, the practical needs that necessitated the compilation of Arabic dictionaries were twofold: the need to preserve the language of the Revelation, i.e. the *Qurʾān*, and the need to impart knowledge of the language of the state to its new subjects in the Islamicized territories. Initially, this led to the compilation of monolingual dictionaries, in which Arabic words were explained with definitions in Arabic and with quotations from texts, mainly poetry, in order to justify their inclusion in the dictionary.

It was only in the 11th century that bilingual dictionaries started to appear. Their aim was to assist the inhabitants of the newly conquered states to learn and understand the language of their victors and their new religion. It was at this stage that dictionaries containing Persian and Turkish appeared.

The early monolingual dictionaries need to be mentioned for yet another reason. Their compilers applied different systems of dictionary ordering, and the first bilingual dictionaries followed these traditions. First of all, there were three different alphabetical orders: the old Semitic *ʾabjada* ordering (*ʾalif*, *bāʾ*, *jīm*, *dāl*, etc.); the phonetic order, beginning with pharyngeals and laryngeals and ending with the labials; and, of course the 'normal' alphabetical order of Arabic (*ʾalif, bāʾ, tāʾ*, etc.). Then there were different systems of arrangement of the roots. Some dictionaries used a permutative system, in which all possible combinations of a set of radicals were grouped together in the dictionary under the highest letter. Another way of root arrangement was by ordering the roots by their last letter, i.e. a rhyming order. To make things even more complex, Carter (1990) points out that these different systems could be combined: permutative/phonetic (*Kitāb al-ʿayn* by al-Xalīl), permutative/alphabetical (*al-Jamhara* by Ibn Durayd); alphabetical by first radical (*Kitāb al-jīm* by aš-Šaybānī); alphabetical by last radical (*aṣ-Ṣiḥāḥ* by al-Jawharī).

Of the monolingual Arabic dictionaries, the following should be mentioned: *Kitāb al-ʿayn* (8th century), ascribed to al-Xalīl; *aṣ-Ṣiḥāḥ* (10th century) by al-Jawharī; the *Lisān al-ʿArab* by Ibn Manẓūr (13th century); *al-Qāmūs al-muḥīṭ* by al-Fīrūzābādī (14th century); and the *Tāj al-ʿarūs* by az-Zabīdī (18th century).

The *Kitāb al-ʿayn* is usually regarded as the first dictionary of Arabic, but the entries are presented according to the roots, and within certain roots not all existing words are included. Furthermore, the arrangement of the roots is very complicated. This may explain the fact that the *Kitāb al-ʿayn* was not used on a broad scale. According to Haywood (1965), the *Kitāb al-ʿayn* was not designed for popular use but rather for scholars. He also asserts that, while early lexicography was undoubtedly Arabic, it was not Arab, since Persians played an important role in it, not to mention natives of Transoxania, Armenia, and Spain. Al-Xalīl may well have been the first scholar to attempt to register the complete vocabulary content of any language. By this, he meant all the roots, rather than all the words (Talmon 1997).

In the *Ṣiḥāḥ*, al-Xalīl's principle of arrangement was modified into an arrangement by the last radical of a root. The work contains definitions, often identical to those in the *Kitāb al-ʿayn*, as well as illustrative examples from Bedouin poetry and speech. Since the title means 'correct' or 'pure', this indicates that the author tried to include only pure Arabic, but he has been criticized for his too strict interpretation.

The *Lisān al-ʿArab* may be regarded as the apogee of the early development in Arabic lexicography. It set a standard of comprehensiveness and systematic arrangement that stood unchallenged for a number of centuries. It was in fact, as Ibn Manẓūr himself stated, entirely derived from previous works. The *Lisān* was the first dictionary that was fully based on the system of ordering by root. The aforementioned systems of dictionary ordering became obsolete because of lack of standardization.

In the *Qāmūs,* al-Fīrūzābādī aimed at eliminating extraneous matter from the works of his predecessors and added economy by using abbreviations. The *Qāmūs* served as the basis for the early European bilingual dictionaries (see Sec. 3).

In the *Tāj al-ʿarūs,* az-Zabīdī restored all the information that had been discarded by al- Fīrūzābādī in his *Qāmūs*. At the same time, he added an enormous amount of information from other sources, among them the *Lisān al-ʿArab*. This work became the starting point for Lane's immense bilingual dictionary.

In terms of macrostructure (the number and type of words serving as entries), all three monolingual dictionaries mentioned above have very much in common. In microstructure (quantity and type of information within the entries), however, there are considerable differences, which will not be elaborated on here.

In general, Haywood (1965) concludes, the Arabic lexicographers contributed to keeping the written language static and in spreading its understanding, thus preventing the spoken dialects from developing into independent languages, as the Romance languages did.

2. EARLY BILINGUAL DICTIONARIES

Historically, bilingual Arabic dictionaries can be divided into two categories: (i) dictionaries in which the foreign language is one of the other languages from the region, such as Persian, Turkish, Coptic, etc., and (ii) dictionaries in which the foreign language is a European language (Latin remained the first language for a number of centuries and was then followed by English, German, French, and Spanish).

The first bilingual Arabic/foreign language dictionary to be mentioned here is the Arabic/Persian dictionary by az-Zamaxšarī (middle of the 11th century). It introduces the modern dictionary order in its entirety, listing words under their roots alphabetically. This dictionary was a pioneering work, and it certainly helped the people from Persia to understand Arabic better. As a matter of fact, many dictionaries in Arabic were compiled and collected by Persians. The first Arabic/Turkish dictionaries were translations of the *Ṣiḥāḥ* and the *Qāmūs*. Haywood (1965) also mentions a Turkish/Arabic dictionary (*Dīwān luġat at-Turk* by al-Kāšġarī; cf. Ermers 1999:16–20). Some Arabic/Coptic word lists were compiled in these early years as well (cf. Sidarus 2000).

Lexicography in these languages became influenced by principles from Arabic lexicography, such as the ordering of the entries. Since the Arabic/foreign language dictionaries were translations of the early monolingual dictionaries like the *Ṣiḥāḥ* and the *Qāmūs*, the compilers tended to adopt the Arabic ordering of the entries.

The first Arabic/foreign language dictionary in Europe was Golius' Arabic/Latin dictionary. Golius was a professor of Arabic at Leiden and set himself to compiling his *Lexicon Arabico-Latinum*, which was published in Leiden in 1653 in one volume. It was based mainly on the *Ṣiḥāḥ*, but rearranged in the modern dictionary order.

The second Arabic/foreign language dictionary, also from Europe, was an Arabic/Latin dictionary, compiled by Georg Wilhelm Freytag and published in Halle in four volumes between 1830 and 1837. Freytag based his work on the *Qāmūs*, rearranging the entries, but he claimed that he had also consulted the *Ṣiḥāḥ* and Golius' dictionary. Since the *Qāmūs* is a lexicon without illustrative examples and quotations, Freytag's dictionary does not contain any examples, either, and the different meanings of Arabic words are explained without quotations from Arabic texts.

In the early stages of bilingual lexicography, the Orientalists followed Classical Arabic models like the *Ṣiḥāḥ* and the *Qāmūs* very closely with respect to their contents, although they adopted the modern arrangement of ordering the roots according to the first radical.

Edward William Lane decided to fill the gap that Freytag had left. He started his work in 1842, and when he died in 1876, he had reached the root *q-ḍ-ḍ*. His dictionary (Lane 1863–1893) remains an outstanding lexicographical tool, but the information on the roots after *q-ḍ-ḍ* is incomplete. Later dictionaries, such as Dozy (1881) and Kazimirski (1860), went some way toward completing the work that Lane had started.

In the third decade of the 20th century, August Fischer believed that this practice was no longer adequate; what was needed was an Arabic lexicon based on historical principles, which should illustrate every word and meaning from actual use in literature. However, he was unable to finish this immense task before his death.

Arabic/foreign language dictionaries in those days were meant for passive use by Orientalists, rather than as active dictionaries for speakers of Arabic. The rather isolated position of the Arab world during these centuries may have contributed to this tendency. As for foreign language/Arabic dictionaries during this early period, there are fewer initiatives to be mentioned. Pedro de Alcalá's Spanish/Arabic dictionary (1505) is worth mentioning, although it contains the spoken Arabic of Granada, and so is the Latin/Arabic dictionary of Jean-Baptiste Du Val (1632).

3. Modern bilingual Arabic dictionaries

The dividing line between the early stages of Arabic lexicography and modern lexicography is difficult to draw, but on the basis of the following characteristics it is possible to make a distinction. Most early dictionaries can be characterized as follows: they contain Classical Arabic; they are mainly meant for passive use; they present examples (*šawāhid*) as evidence, rather than illustrating the use of a word; and they were compiled through a philological approach. All bilingual dictionaries mentioned in the previous sections share a number of these characteristics. For modern bilingual dictionaries, this is obviously not the case. The most complete list of Arabic dictionaries is the bibliography by Wagdy Rizk Ghali (1971) and its supplement (1974). For each combination of languages, Ghali makes a distinction between Arabic/foreign language and foreign language/Arabic dictionaries. No other distinctions are made, so that many different kinds of dictionary are mentioned without any further categorization. Even travelers' vocabulary lists are sometimes included, as well as colloquial Arabic dictionaries. Since 1974, no comprehensive bibliography of Arabic dictionaries has been published.

For the purpose of the present entry, only bilingual Arabic dictionaries published after 1900 have been consulted, and specialized dictionaries are not included. Only bilingual dictionaries with a European language are dealt with here, with the emphasis on dictionaries with English, French, or German as source or target language. Most of the other European languages have a reference dictionary serving primarily the academic community in that particular language area, e.g. Corriente (1988), Spanish/Arabic (30,000 entries); (1991) Arabic/Spanish (70,000 entries); Baldissera (2004), Italian/Arabic (24,000 entries) and Arabic/Italian (27,000 entries). Most East European languages (Czech, Polish, Rumanian) have dictionaries in the middle range (about 30,000 entries), e.g. for Polish, Danecki and Kozlowska (1996). For Russian, Baranov's (1976) Arabic/Russian dictionary (33,000 entries) and Borisov's (1981) Russian-Arabic dictionary remain the classic reference tools. For the Scandinavian languages, most bilingual dictionaries are of lesser size, and only Rasmussen's (1984) Cairo Arabic/Danish is a scholarly dictionary. Others, e.g. Statens Institut för Läromedel (1987) for Swedish/Arabic and Touma (1991) for Arabic/Swedish, are more intended for practical use by Arabic-speaking immigrants. A special case is the Arabic/Hebrew dictionary by Ayalon and Shinar (1960), because both languages are official languages of Israel.

The bibliography by Ghali (1971, 1974) also mentions bilingual dictionaries for the languages of non-Arabic-speaking Islamic countries, such as Persian, Turkish, Urdu, and Indonesian. The list contains 2 Persian/Arabic dictionaries and 13 Arabic/Persian; 4 Turkish/Arabic and 4 Arabic/Turkish; 28 Arabic/Urdu and 2 Urdu/Arabic. The number of Arabic/foreign language dictionaries is significantly higher than the reverse, which may be explained by the fact that Muslims from non-Arabic-speaking countries probably have a greater need for understanding Arabic than for producing Arabic. Besides, the language variety used in these dictionaries is more Classical than Modern Standard Arabic.

Table 1. Numbers of Arabic bilingual dictionaries

Language	Arabic/ foreign language	Foreign language/ Arabic	Both in one volume
English	18	25	5
French	14	20	2
German	5	7	–
Spanish	4	3	–
Italian	2	5	–
Russian	4	4	–
Greek	1	1	–
Turkish	4	4	–
Persian	13	2	–
Urdu	28	3	–
Kurdish	–	1	–
Coptic	–	4	–
Armenian	1	–	–
Syriac	1	–	–
Hebrew	2	3	–
Esperanto	–	1	–
Czech	–	1	–
Malay	3	1	–
Latin	7	2	–

Table 1 lists the number of general Arabic bilingual dictionaries mentioned in Ghali (1971, 1974).

4. Classification of modern Arabic bilingual dictionaries

In modern lexicography, dictionaries are usually classified according to the following criteria:

i. Language variety (in the case of Arabic: Classical Arabic, Modern Standard Arabic, colloquial Arabic);
ii. Type of language (general vocabulary or specialized terminology from specific professional or scientific fields);
iii. Purpose of the dictionary (active for encoding in the target language, or passive for decoding the source language).

The principle of classification by language variety is particularly relevant for Arabic because there are at least three language varieties – Classical, Modern Standard, and colloquial – which should be treated differently. Dictionaries for all three categories do exist, but they have very few characteristics in common. One might say that Classical Arabic is a static language; many words have very specific meanings which are no longer current. On the other hand, Modern Standard Arabic is a language which, although evolved from Classical Arabic, is fully involved in the developments of modern times, where the language is trying to cope with the need for a rapidly extending scientific and technical vocabulary. The third language variety of Arabic is colloquial Arabic. Given the fundamental differences between Modern Standard Arabic and colloquial Arabic, as a consequence of the presence of → diglossia throughout the Arab world, as well as the immense differences between the varieties of colloquial Arabic, there is no other solution than compiling separate dictionaries for Modern Standard Arabic and colloquial Arabic. Given these differences in colloquial Arabic, the existence of separate dictionaries for all different varieties of colloquial Arabic is justified. According to this principle, at least three categories of Arabic/foreign language dictionaries exist: Classical Arabic/foreign language, Modern Standard Arabic/foreign language, and colloquial Arabic/foreign language. Obviously, the same division can be made for the reverse direction.

Theoretically, bilingual dictionaries could exist with two varieties of Arabic (for example Modern Standard Arabic/colloquial Arabic or Classical Arabic/Modern Standard Arabic), but if they exist at all, they should be treated as monolingual dictionaries.

The second criterion for classifying dictionaries is the type of language in the dictionary: general vocabulary or specialized terminology, related to specific domains. Both types have their own characteristics. General dictionaries contain general vocabulary not related to specialized domains. They contain not only words and their translations but also expressions, collocations, etc. Furthermore, a general dictionary contains grammatical information about the source or target word, as well as other information to assist the user in using such a word.

Specialized dictionaries contain vocabulary related to specific professional or scientific fields. Specialized dictionaries do not normally contain additional information and in most cases can simply be regarded as vocabulary lists. Specialized dictionaries make up a considerable percentage of all existing Arabic dictionaries. Ghali (1971, 1974) mentions 855 items, of which 331 are specialized dictionaries, which are classified by him according to the subject field and, subsequently, according to the language.

The vast majority of specialized dictionaries are of the foreign language/Arabic type. This can probably be explained by the fact that technical and scientific terminology is not unified in the Arab world, in spite of the enormous efforts that have been made by Arabic → language academies, the Arabization Bureau of ALECSO, and other organizations (→ technical terminology).

The last criterion for classifying bilingual dictionaries is the target group for which they are compiled. Usually, active dictionaries (for encoding in the target language) are distinguished from passive dictionaries (for decoding the source language). Passive dictionaries should contain information about the source language that will enable the user to decode texts written or spoken in the foreign language, while active dictionaries should provide the user with all necessary information to encode texts in the target language as correctly as possible. There are dictionaries which combine the functions of passive and active dictionaries, the so-called bidirectional dictionaries, which contain all necessary information for both languages.

The three criteria for classifying dictionaries can be combined, so that theoretically, one could have twelve types of Arabic/foreign language dictionaries: i. Classical Arabic-general-active; ii. Classical Arabic-general-passive; iii. Classical Arabic-specialized-active; iv. Classical Arabic-specialized-passive; v. Modern Standard Arabic-general-active; vi. Modern Standard Arabic-general-passive; vii. Modern Standard Arabic-specialized-active; viii. Modern Standard Arabic-specialized-passive; ix. colloquial Arabic-general-active; x. colloquial Arabic-general-passive; xi. colloquial Arabic-specialized-active; xii. colloquial Arabic-specialized-passive.

Obviously, not all of these categories actually exist. In the case of colloquial Arabic active dictionaries, colloquial Arabic/foreign language may at first sight seem rather unusual, because a speaker of Arabic is unlikely to need a dictionary of this type for the production of speech or written text in another language. On the other hand, a passive dictionary of colloquial Arabic/foreign language for non-Arabs is very common. In general, one might say that Modern Standard Arabic/foreign language dictionaries exist in all four possible combinations (i.e. categories v, vi, vii, viii), whereas Classical Arabic/foreign language dictionaries are almost exclusively passive (categories ii and iv), and colloquial Arabic/foreign language dictionaries exist as general passive dictionaries only (category x).

The same classification in twelve categories may be applied to foreign language/Arabic dictionaries. But here, too, not all types actually exist. Foreign language/Classical Arabic dictionaries, for instance, seem to be rare, and specialized dictionaries of foreign language/colloquial Arabic are very unusual. The most widespread types appear to be the categories v (Modern Standard Arabic, general, active), vi (Modern Standard Arabic, general, passive), vii (Modern Standard Arabic, specialized, active), viii (Modern Standard Arabic, specialized, passive), and ix (colloquial Arabic, general, active). Type vii (foreign language/Modern Standard Arabic, specialized, active) is very frequent, while its reverse counterpart (Modern Standard Arabic/foreign language, specialized, passive) is rather uncommon, due to the already mentioned lack of unification in technical and scientific terminology in the Arab world.

One additional category of specialized foreign language/Modern Standard Arabic dictionary is that of the so-called encyclopedic dictionary, which contains not only translations of technical and scientific terms but also explanations in Arabic. One example of this category is Najjar a.o. (1983), with 7,000 entries.

According to this classification system, it would be possible to assign a category indication to existing bilingual Arabic dictionaries, and it would certainly be useful for students of the languages involved to have this kind of information. As a matter of fact, however, no list or bibliographical index containing this type of information is available.

5. EXISTING BILINGUAL ARABIC DICTIONARIES

All dictionaries in the sample have been consulted by reading the introductions and looking at the presentation of the data. Thus, it was possible to determine which target groups the dictionary compilers had in mind for their dictionaries, as well as the language varieties included, the methods used for collecting the data, and so on. Amazingly, it came to light that some dictionaries do not contain any introduction at all. This group includes Abcarius (1974), English/Arabic, 23,000 entries; and Wahrmund (1898), German/Arabic, 7,500 entries, Arabic/German, 50,000 entries. Others contain only a very brief introduction, e.g. Belot (1963), French/Arabic, 33,000 entries. Other dictionaries contain extensive introductions in which the compilers account for the policy applied in compiling the dictionary, make general statements about the Arabic language, etc.

Some dictionary compilers mention their sources, including other dictionaries, while others do not mention their sources; there is very little agreement on what categories of information a dictionary introduction should contain.

The language variety covered by a dictionary is usually indicated by its title. If not, this information is given either in the introduction or by the content itself. A mixture of Classical Arabic and Modern Standard Arabic in one single dictionary is possible. The inclusion of colloquial Arabic in Modern Standard Arabic dictionaries, on the other hand, is very rare and considered improper.

The introduction of a dictionary should also supply the user with information about the category to which the dictionary belongs, and its contents show at first glance if a dictionary is of the general or specialized type. Most existing bilingual Arabic dictionaries can be classified as general or specialized on the basis of their title, introduction, or content. General dictionaries, on the basis of their title, are, among others, Wehr's (1979) 'Modern Written Arabic', translated from German into English by Cowan, with 28,000 Arabic/German entries and 28,000 Arabic/English entries; and Madina's (1973) 'Modern Literary Language', with 27,000 Arabic/English entries. The absence of any indication of a specific subject field also indicates the generality of the dictionary. Examples of specialized foreign language/Arabic dictionaries are Faruqi (1986; 11,000 entries) for legal terminology, and Al-Khatib (1978; 66,000 entries) for scientific and technical terminology (both English/Arabic).

However, the question of whether a dictionary is an active or a passive dictionary is often difficult to answer. Many introductions to existing dictionaries simply do not make mention of the dictionary's target group (Wehr 1979; M. Baalbaki [1981], *al Mawrid*, English/Arabic, 56,000 entries). Still, there are other indications which throw a light on the intentions of the dictionary compiler or publisher. The first indication is the language in which the introduction is written. If an Arabic/foreign language dictionary has an introduction in Arabic, it is aiming at Arabic-speaking users, which means the dictionary is meant to be an active one, e.g. R. Baalbaki's (1996) *al Mawrid*, with 27,000 Arabic-English entries. On the other hand, if the introduction is in the foreign language, the Arabic/foreign language dictionary is intended to be passive (Wehr 1979, 1985; Reig [1983], 16,000 entries). Some dictionaries contain introductions in both languages involved, a clear indication that the dictionary is meant by the compiler or the publisher to be a bidirectional dictionary (Wahba 1989).

The type of information included in the entries of a dictionary also indicates the target group it addresses. An active dictionary should contain as much information as possible about the target language, whereas a passive dictionary should provide as much information as possible about the source language. Thus, a dictionary may conceivably be intended to serve both groups of users, while the type of information provided indicates it is actually a monodirectional dictionary (Doniach 1972; 18,000 entries).

Within the sample of Arabic/foreign language dictionaries, it is clear that there are very few active Arabic/English dictionaries for speakers of Arabic. Only two dictionaries, R. Baalbaki's (1996) Arabic/English *al Mawrid* and Elias (1960), seem to be intended as active dictionaries by their compilers. However, a more thorough look at these two dictionaries reveals that neither contains information about the English equivalents presented. There is no information about pronunciation or word stress in English, no grammatical information, etc. So, if the above-mentioned criteria for active use are applied to these two dictionaries, they cannot be considered active dictionaries for speakers of Arabic. This observation leads to the conclusion that no single modern active Arabic/English dictionary for speakers of Arabic is available. As for the other languages in the sample, there are a number of bidirectional dictionaries Arabic/German (Schregle 1981–1986), Arabic/French (Reig 1983), or Arabic/Dutch (Hoogland a.o. [2003], Dutch/Arabic, 37,000 entries, Arabic/Dutch, 25,000 entries). But obviously, no single Arabic/foreign language dictionary supplies enough information in the target language for active use by speakers of Arabic.

Given this lack of active dictionaries for speakers of Arabic, some Western dictionary compilers have made an effort to provide Arab users with some information about the target language in dictionaries that were intended to be passive dictionaries. Thus, one finds gender information for German in Krotkoff (1976; German/Arabic, 13,000 entries, Arabic/German, 12,000 entries), or an index of French words in order to facilitate the use of an Arabic/French dictionary in reverse direction (Reig 1983; in this particular case there is no information about the gender of the French equivalents).

Some dictionaries contain directions for use in Arabic in order to serve the Arabic-speaking users (Elias 1960), and some of them could even be categorized as suitable for limited bidirectional use, since the amount of information in the target language is limited. Very few dictionaries seem fully suitable for bidirectional use, containing all essential information in both source and target language (Hoogland a.o. 2003).

Hinds and Badawi's (1986) dictionary of Egyptian Arabic/English, with 25,000 entries, deserves special mention as a good example of lexicography that is primarily of a dialect, but which covers many Modern Standard Arabic lexical items as well and labels them appropriately. Given the fact that written forms of the dialects are becoming increasingly more common, this kind of mixed dictionary containing more than one variety of Arabic might become more widespread in future.

6. Macrostructure of bilingual Arabic dictionaries with Arabic as source language

In the literature, the lack of a modern and reliable monolingual Arabic dictionary to serve as a reference to determine which words should be included in Arabic/foreign language dictionaries has frequently been deplored (R. Baalbaki 1996; Kropfitsch 1981; Krahl 1988; → lexicography: monolingual dictionaries). Many dictionaries with Arabic as source language are limited in their macrostructure with regard to modern words and expressions. On the other hand, these dictionaries contain many words and expressions that are no longer current in everyday usage.

One reason for the lack of modern terminology is the fact that most dictionaries are rather outdated. Abou El Aazm (2003), in comparing the root *ᶜ-l-m* in the Arabic/French *As-Sabil* (Reig 1983) and the monolingual *al-Wasīṭ* (→ lexicography: monolingual dictionaries) remarks that *al-Wasīṭ* does not contain a number of frequently used words, such as *ᶜilmī* 'scientific', *ᶜālima* 'dancer', *maᶜlūm* 'known', *mutaᶜallim* 'educated', *ʾiᶜlām* 'notification', *maᶜlūma* 'data', *maᶜlūmātī* 'informatic', *taᶜlīm* 'education', *maᶜlama* 'contour, outline', *ʾiᶜlāmī* 'informative, informational', *istiᶜlām* 'inquiring', *istiᶜlāmāt* 'inquiries', *ᶜālamī* 'global', *ᶜālamiyyan* 'globally, worldwide', *ᶜālamiyya* 'internationalism', *ᶜalmana* 'to secularize'. But even in *al-Mawrid* Arabic/English (M. Baalbaki 2003), which seems to be reprinted regularly, recently formed, nonspecialized technical and scientific terms are missing. Of the words just mentioned in El Aazm's comparison, in the 2004 edition of *al-Mawrid* Arabic/English, the following words are missing: *maᶜlūmātī, taᶜlīmāt*, *maᶜlama, istiᶜlāmāt, ᶜālamiyyan*. Other modern words, such as *ᶜawlama* 'globalization', *fāks* 'fax', *mawqiᶜ* 'website', *ʾintarnīt* 'Internet', *mūdim* 'modem', *hātif maḥmūl/naqqāl* 'mobile telephone', are missing as well in this edition. Not only does this underline the need for modern dictionaries of Arabic, but it also shows the speed with which Arabic is expanding or modernizing its vocabulary.

Regarding the method of arrangement of Arabic/foreign language dictionaries, it is ironic to note that most Western lexicographers tend to retain the tradition of arrangement by root, while Arab lexicographers are tending toward the more 'Western' alphabetical arrangement (exceptions are Krotkoff [1976] and Kropfitsch [1996], which use alphabetical ordering).

Another phenomenon related to the macrostructure of bilingual Arabic dictionaries is the fact that some dictionaries, especially foreign language/Arabic dictionaries compiled by Arab authors, contain many neologisms and coinages in Arabic, made up as equivalents for foreign language words. Asfour (2003) studied a number of foreign language/Arabic dictionaries and states that it is not imperative for a translator or a dictionary maker to find or make up a single word for a foreign single word. He lists a number of examples from M. Baalbaki 1981 (*al-Mawrid* English/Arabic), which he describes as "horrible coinages". Idriss (1996; 50,000 entries) introduces many neologisms in Arabic, but marks them with an asterisk to make the dictionary user aware of the fact that the equivalent is a newly coined word. These neologisms are followed by an explanation between brackets, to supply the user with an extra means for understanding the meaning of both the foreign word and the Arabic neologism.

7. Microstructure of Arabic bilingual dictionaries

The microstructure of a dictionary consists of the information provided within the entries. Included are illustrative examples such as idiomatic expressions, → collocations (frequent combinations), and short sentences illustrating the use or meaning of words. Modern Arabic bilingual dictionaries contain very few of these examples.

According to Al-Kasimi (1977), the early Arab lexicographers provided quotations from prose and poetry as evidence that the word under discussion was found in the Arabic language, rather than as an illustration of its meaning. The fact that in modern lexicography this need for examples as evidence is no longer present might explain the fact that in modern Arabic lexicography the desire to provide examples seems to have faded.

The absence of collocations is signaled by Hoogland (1993), who searched a number of Arabic/foreign language and foreign language/Arabic dictionaries for a number of frequent combinations (collocations) in Arabic. Frequent combinations like *irtakaba jarīmatan* 'to commit a crime', *baḏala juhūdan* 'to make an effort', *ʾaḥraza taqadduman* 'to make progress' appear in very few dictionaries, despite their frequent co-occurrence in actual usage.

One example worth mentioning is Schregle's (1981–1986) Arabic/German dictionary. This dictionary has a unique depth of coverage in its microstructure but was never completed. Its last volume covers part of the letter *qāf*.

8. Arabic dictionaries in the computer era

Various dictionaries are available on electronic media these days. In 1996, the Sakhr company (Cairo) produced a CD-ROM called *Al Qamoos*, which was a compilation of existing printed bilingual dictionaries. The first version was bilingual Arabic/English, but a later version included other European languages and Turkish. Since the market for this kind of product is much more ephemeral than the market for printed dictionaries, it is of no use to report on the present state of affairs.

Another recent development is the availability of Internet pages containing dictionaries or vocabulary lists. Lexicool.com has a list of 119 dictionaries containing Arabic, of which 40 are available on-line. The website of the ALECSO Bureau of Arabization offers a database containing all unified thematic dictionaries published by the organization in Arabic, English, and French (www.arabization.org.ma/Dictionnaire.asp). International organizations such as the UN (http://unterm.un.org/) or the World Health Organization (www.emro.who.int/umd/) have online databases with specialized terminology available.

9. Concluding remarks

Not many dictionaries of Modern Standard Arabic have been published in recent years, especially not for speakers of Arabic. It seems that principles of modern lexicography have been applied in very few dictionaries. To conclude, we may quote Al-Ajmi (2002:130), who states that "bilingual lexicography in the Arab World is suffering from a lack of guided practice and is in dire need of radical changes in both design and approach". According to him, neither the existing English/Arabic nor the Arabic/English dictionaries are suitable for the purpose for which they were designed.

Bibliographical references

Abcarius, John. 1974. *An English-Arabic readers' dictionary*. Beirut: Librairie du Liban.

Abou El Aazm, Abdel Ghani. 2003. "al-Luġa al-ʿarabiyya al-ḥadīṯa fī muʿjam *as-Sabīl* muʿjam ṯunāʾī al-luġa". *Etudes Lexicographiques* 2.5–26.

Al-Ajmi, Hassan. 2002. "Which microstructural features of bilingual dictionaries affect users' look-up performance?". *International Journal of Lexicography* 15:2.119–131.

Al-Kasimi, Ali M. 1977. *Linguistics and bilingual dictionaries*. Leiden: E.J. Brill.

Al-Khatib, Ahmed Sh. 1978. *A new dictionary of scientific and technical terms: English-Arabic*. Beirut: Librairie du Liban.

Asfour, Mohammed. 2003. "Problems in modern English-Arabic lexicography". *Zeitschrift für Arabische Linguistik* 42.41–52.

Ayalon, David and Pessah Shinar. 1960. *Millōn, aravī/ivrī la-lašōn ha-aravīt ha-ḥadaša* [Arabic-Hebrew dictionary of Modern Arabic]. Jerusalem: Magnes Press. (New ed., with Moshe Brill, *Arabic-Hebrew dictionary of Modern Arabic*, 1996.)

Baalbaki, Munir [Munīr Baʿlabakkī]. 1981. *Al Mawrid: A modern English-Arabic dictionary*. Beirut: Dār al-ʿIlm li-l-Malāyīn. (1st ed. 1967.)

——. 2003. *Al Mawrid: A modern English-Arabic dictionary*. Beirut: Kazi Publications.

Baalbaki, Rohi. 1996. *Al-Mawrid: A modern Arabic-English dictionary*. Beirut: Dār al-ʿIlm li-l-Malāyīn.

Baldissera, Eros. 2004. *Il dizionario di arabo: Italiano-arabo arabo-italiano*. Bologna: Zanichelli.

Baranov, Charlampij Karpovič. 1976. *Arabsko-ruskij slovarʾ*. 2 vols. 3rd ed. Moscow: Russkij Jazyk. (1st ed., 1940–1946; 2nd ed., 1957.)

Belot, Jean-Baptiste. 1963. *Dictionnaire français arabe*. Beirut: Imprimerie Catholique.

Borisov, V.M. 1981. *Russko-arabskij slovarʾ*. 2nd ed. Moscow: Russkij Jazyk.

Carter, Michael G. 1990. "Arabic lexicography". *Religion, learning and science in the ʿAbbasid period,* ed. M.J.L. Young, J.D. Latham, and R.B. Serjeant, 106–117. Cambridge: Cambridge University Press.

Corriente, Federico. 1970. *Diccionario árabe-español*. Madrid: Instituto Hispano-Árabe de Cultura.

——. 1988. *Nuevo diccionario español-árabe*. Madrid: Instituto Hispano Árabe de Culture. (1st ed. 1970.)

——. 1991. *Diccionario árabe-español*. 4th ed. Barcelona: Herder. (1st ed., 1977.)

Danecki, Janusz and J. Kozlowska. 1996. *Słownik arabsko-polski*. Warsaw: Wiedza Powszechna.

Doniach, Nakdimon. 1972. *The Oxford English-Arabic dictionary of current usage*. Oxford: Oxford University Press.

Dozy, Reinhart Pieter Anne. 1881. *Supplément aux dictionnaires arabes*. Leiden. (Repr. Beirut: Librairie du Liban, 1968).

Elias, Antun Elias. 1960. *Elias' modern dictionary, Arabic-English*. Cairo: al-Maṭbaʿa al-ʿAṣriyya.

Ermers, Robert. 1999. *Arabic grammars of Turkic: The Arabic linguistic model applied to foreign languages and translation of ʾAbū Ḥayyān al-ʾAndalusī's* Kitāb al-ʾidrāk li-lisān al-ʾatrāk. Leiden: E.J. Brill.

Faruqi, Harith Suleiman. 1986. *Faruqi's law dictionary: Arabic-English*. Beirut: Librairie du Liban.

Ghali, Wagdy Rizk. 1971. *Muʿjam al-muʿjamāt al-ʿarabiyya/Arabic dictionaries: A comprehensive annotated bibliography of printed Arabic dictionaries*. Cairo: al-Hay'a al-Miṣriyya al-ʿĀmma.

——. 1974. "Arabic dictionaries: An annotated comprehensive bibliography. Supplement". *Mélanges de l'Institut Dominicain d'Etudes Orientales* 12.243–280.

Haywood, John A. 1965. *Arabic lexicography: Its history and its place in the general history of lexicography*. Leiden: E.J. Brill.

Hinds, Martin and El-Said Badawi. 1986. *A dictionary of Egyptian Arabic: Arabic-English*. Beirut: Librairie du Liban.

Hoogland, Jan. 1993. "Collocations in Arabic (MSA) and the treatment of collocations in Arabic dictionaries". *Proceedings of the Colloquium on Arabic Lexicology and Lexicography*, ed. Kinga Dévényi, Tamás Iványi, and Avihai Shivtiel, 75–93. Budapest: Eötvös Loránd University and Csoma de Kőrös Society.

——, Kees Versteegh, and Manfred Woidich. 2003. *Woordenboek Arabisch-Nederlands Nederlands-Arabisch*. 2 vols. Amsterdam: Bulaaq.

Idriss, Souheil. 1996. *Al Manhal dictionnaire français-arabe*. Beirut: Dar Al Adab.

Kazimirski, Alfred de Biberstein. 1860. *Dictionnaire arabe-français*. Repr., Paris: Maisonneuve. (1st ed., Paris: Théophile Barrois, 1846–1847.)

Krahl, Günther. 1964. *Deutsch-arabisches Wörterbuch*. Leipzig: VEB Verlag. [12,000 entries.]

——. 1988. Review of: Schregle (1981–1986). *Orientalistische Literaturzeitung* 83:2.183.

——. 2001. *Handwörterbuch Deutsch-Arabisch*. Munich: Langenscheidt.

—— and Gharieb Mohamed Gharieb. 1984. *Wörterbuch Arabisch-Deutsch*. Munich: Max Hueber. [25,000 entries.]

Kropfitsch, Lorenz. 1981. Review of: Corriente (1970). *Zeitschrift für Arabische Linguistik* 6.75–80.

——. 1996. *Arabisch-Deutsch. Langenscheidts Handwörterbuch*. Berlin and Munich: Langenscheidt. [45,000 entries.]

Krotkoff, Georg. 1976. *Langenscheidts Taschenwörterbuch der Arabischen und Deutschen Sprache*. Berlin and Munich: Langenscheidt.

Lane, Edward William. 1863–1893. *An Arabic-English lexicon, derived from the best and the most copious Eastern sources…* Book I, in 8 parts (all published). London and Edinburgh: Williams and Norgate. (Repr., Beirut: Librairie du Liban, 1980.)

Madina, Maan Z. 1973. *Arabic-English dictionary of the modern literary language*. New York: Pocket Books.

Najjar, Ibrahim, Ahmed Zaki Badaoui, and Youssef Chellalah. 1983. *Dictionnaire juridique français-arabe*. Beirut: Librairie du Liban.

Rasmussen, Stig T. 1984. *Arabisk-dansk ordbog over Cairodialekten*. Copenhagen: Akademisk Forlag.

Reig, Daniel. 1983. *Dictionnaire arabe français: As-Sabil*. Paris: Larousse.

Schregle, Götz. 1981–1986. *Arabisch-deutsches Wörterbuch*. Wiesbaden: F. Steiner.

Sidarus, Adel. 2000. "L'influence arabe sur la linguistique copte". *History of the language sciences*, ed. Sylvain Auroux, Konrad Koerner, Hans-Josef Niederehe, and Kees Versteegh, I, 321–325. Berlin: de Gruyter.

Statens Institut för Läromedel. 1987. *Svenskt-arabiskt lexikon/Muʿjam suwaydī-ʿarabī (språklexikon för invandrare)*. Stockholm: Statens Institut för Läromedel.

Talmon, Rafael. 1997. *Arabic gramnar in its forma-*

tive age: Kitāb al-ʿayn *and its attribution to Ḫalīl b. Aḥmad*. Leiden: E.J. Brill.

Touma, Pierre. 1991. *Toumas omfattande ordbok: Svensk-arabisk*. Göteborg.

Wahba, Magdi. 1989. Al-Mukhtar: *A concise English-Arabic dictionary*. Beirut: Librairie du Liban.

Wahrmund, Adolf. 1898. *Handwörterbuch der neu-arabischen und deutschen Sprache*. Giessen: Ricker.

Wehr, Hans. 1979. *A dictionary of Modern Written Arabic*. Trans. J. Milton Cowan. Wiesbaden: O. Harrassowitz.

——. 1985. *Arabisches Wörterbuch für die Schriftsprache der Gegenwart: Arabisch-Deutsch*. 5th ed. Wiesbaden: O. Harrassowitz. (1st ed., 1952.)

Jan Hoogland (University of Nijmegen)

Lexicography: Classical Arabic

1. RANGE AND IMPORTANCE

Within Classical Arabic literature, lexicographical writings form an extensive and multifaceted branch, having produced remarkable results in the period from the late 2nd/8th century to the 12th/18th century and flourishing particularly in the 4th/10th century. Along with dictionaries proper – i.e. books arranging all the elements of the Classical Arabic vocabulary in alphabetical order and explaining them – there were many other types of books. The aim of covering the whole lexicon only seems to have arisen two centuries after the appearance of the first lexicographical monographs. The contributions range in length from just a few pages in the case of specialized treatises to more than five thousand quarto pages in the case of the printed edition of Murtaḍā az-Zabīdī's monumental *Tāj al-ʿarūs* (one of the latest and largest traditional works, compiled toward the end of the 12th/18th century).

Medieval Arabic lexicography is important in several respects, not least for its great practical value in understanding Classical Arabic texts. European lexicographical activities were for a long time mainly limited to translating the indigenous medieval dictionaries (→ Arabic studies in Europe). This method was followed from the times of Antonius Giggeius (*Thesaurus linguae Arabicae*, Milan 1632) and Jacob Golius (*Lexicon arabico-latinum,* Leiden 1653) until the second half of the 19th century, when Edward William Lane published his *Arabic-English lexicon* (London 1863–1893). Modern scientific lexicography of Classical or post-Classical Arabic started in 1881 with the *Supplément aux dictionnaires arabes* by Reinhart Dozy (published in Leiden), when, for the first time, texts were used as the basis for dictionaries. This method has been taken up on a much larger scale since 1957 in the *Wörterbuch der Klassischen Arabischen Sprache*, but so far only two letters (*kāf* and *lām*) have been covered (published Wiesbaden 1957–2004, mainly thanks to the self-sacrificing efforts of Manfred Ullmann). Paradoxically, one can say that the richness of the indigenous lexica has impeded the development of modern scientific lexicography. Therefore, we are still forced to make use of the medieval dictionaries with all their deficits in range, exactness, and documentation. Despite these failings, classical dictionaries are quite important for the history of linguistic thought due to their different methods of arranging the roots and the various types of dictionaries and their intentions and linguistic foundations.

2. LINGUISTIC AND CULTURAL PRECONDITIONS

As early as the reign of the Umayyad dynasty (660–750 C.E.), the texts of the *Qurʾān* and the tradition (insofar as it was fixed) were canonized as reference points for jurisprudence, theology, ritual, and the public demonstration of political power. Consequently, a certain historical stage of the Arabic language attained the position of an enduring ideal. In the belles lettres and in learned circles (*majālis*), from about 750 C.E. onward, great importance is attached to ancient Arabic tradition, above all to pre-Islamic poetry. This interest is due to struggles within Arab society as well as to rivalries between Arabs and non-Arabs over cultural orientation, known as the *Šuʿūbiyya* struggle (see the references in Seidensticker 2002:149, n. 4). Pre-Islamic poetry or prose, the *Qurʾān*, and prophetic tradition could no longer easily be understood by Arabs of the 8th century because the language had changed considerably due to radical shifts in social and cultural life after the conquest of Egypt, Syria, Mesopotamia, and Iran. These changes affected morphology, phonetics, syntax, and, of course, vocabulary. These developments – canonization and change – inevitably brought about a need to preserve and systematize. Attention was paid to both religious and nonreligious texts, and this is reflected in the development of two parallel strands of Arabic lexicography, which later partly merged. A voluminous literature devoted to obscure lexemes in the *Qurʾān* and the prophetic tradition (*ġarīb al-Qurʾān/al-Ḥadīṯ*) stands alongside dictionaries of merely secular orientation, such as the *Kitāb al-jīm* by ʾAbū ʿAmr aš-Šaybānī (d. about 213/828), which mentions only two quotations from the *Qurʾān* against 4,300 lines of poetry. A further factor strengthening the development of Arabic lexicography is the growing importance of Arabic as the language of administration. The Umayyad caliph ʿAbd al-Malik (r. 685–705 C.E.) had installed Arabic as the official language, and by the middle of the 8th century the conquerors' language had gained recognition in the chancelleries even of the more peripheral provinces. From the early Abbasid times onward (i.e. after 750 C.E.), officials of whatever ethnic background were required to have an advanced mastery of Arabic style. Moreover, Arabic was employed as a medium of literary expression even in nonreligious fields in many parts of the Islamic world (the most important exception being Persia). This proliferation meant that a growing need was felt for dictionaries designed less for scholars than for the educated classes, to concentrate on the common vocabulary rather than on obscure or rare items. An example of this type of dictionary is the *Mujmal al-luġa* by ʾAḥmad ibn Fāris (d. 395/1005), who explicitly says in his introduction that he has disregarded obscure words. The popularity of his book is attested by the large number of manuscript copies in libraries all over the world.

This demand for lexicographical works was met by a great variety of dictionaries. The new interest in language gained a momentum of its own and thus added to the diversity of approaches. The most important systems of arrangement and their historical development are presented in the following sections, beginning with the semasiological lexica (starting from sign to meaning) with their three main types of arranging the roots (Secs. 3 to 5), followed by the onomasiological lexica (starting from meaning to sign, i.e. those which arrange the notions according to topic areas; Sec. 6), and finally the different types of specialized lexica (Sec. 7). As for the question of foreign influences, the hypothesis of influence by Greek lexicography is discussed on account of typological parallels (Rundgren 1973; Wild 1965:6–7; Weninger 1994), and in the case of the phonetic arrangement of the Arabic letters in al-Xalīl's order (see below, Sec. 3), the impact of Indian linguistic thought has been assumed (discussed in Law 1990). Yet, Arabic lexicography is a branch of literature which has its roots in Arabic-Islamic culture and was influenced only marginally by foreign cultures.

3. The phonetic-permutative system

The phonetic-permutative system of arranging the roots of the Arabic words, probably the most impractical system, is known to us from the earliest Arabic semasiological dictionary, the *Kitāb al-ʿayn*, which in its main traits goes back to al-Xalīl ibn ʾAḥmad (d. about 175/791). It is based neither on the so-called Abjad order of the letters (ʾ/b/j/d/h/w/z etc.) nor on the common Arabic alphabet (ʾ/b/t/ṯ/j/ḥ/x etc.), which was probably already known in the 8th century C.E., but instead uses a third method of ordering. The key factor in this arrangement of letters is the place of articulation in the mouth or pharynx. The sounds articulated at the back of the pharynx (i.e. the laryngeals) are first in this sequence; the labials, articulated in the front of the mouth, close the so-called 'Xalīlian order'. The resulting alphabet is ʿ/ḥ/h/x/ġ/q/k/j/š/ḍ/ṣ/s/z/ṭ/d/t/ḏ̣/ḏ/ṯ/r/l/n/f/b/m/w/ʾalif/y/ʾ. Much more momentous for practical purposes was al-Xalīl's decision not to arrange the roots strictly according to this new alphabet but to include in each lemma the other combinations of roots which result from all possible permutations of the radicals. This means, for example, that the first triliteral root treated in the *Kitāb al-ʿayn*, *ʿ-h-q*, is immediately followed by the root *h-q-ʿ* (which one would expect to be treated two letters later). A third complication is the fact that within each chapter devoted to one of the letters from /ʿ/ to /ʾ/, there are subchapters treating separately the biliteral, geminated, sound triliteral, simply weak triliteral, doubly weak triliteral, and quadriliteral roots containing the letter in question.

To be accurate, this complicated system of ordering should be called 'phonetic – root-classificatory – permutative'. It allowed Arab scholars to gain some interesting insights; for example, it showed that certain consonants are never combined in a true Arabic root. For everyday use, the system is distinctly impractical, but despite these drawbacks it survived until at least the 14th century C.E. Quite famous dictionaries arranged according to al-Xalīl's method are the *Tahḏīb al-luġa* by al-ʾAzharī (d. 370/980) and the *Muḥkam* by the blind Andalusian scholar Ibn Sīda (d. 458/1066). Even in some dictionaries which abandoned the Xalīlian order of letters, either the root-classificatory or the permutative system (or both of them) are retained. Like most other Arabic dictionaries, the *Kitāb al-ʿayn* gives numerous quotations, primarily taken from early poetry but also from the *Qurʾān* and from prophetic tradition (*Ḥadīṯ*).

4. The alphabetical system

The second method of ordering the roots is the alphabetical system, which is the one most familiar in Western lexicography. It is found in rudimentary form in the second oldest semasiological dictionary known to us, the *Kitāb al-jīm* by ʾAbū ʿAmr aš-Šaybānī (d. about 213/828). Its order is not fully alphabetical, as the author groups the roots only according to the first radical. Within the chapters *ʾalif, bāʾ, tāʾ*, etc., the roots are grouped by the informants who provided them or by mere association. This stage could be called one-third alphabetical, as only one of the radicals, usually three in number, is taken into account when ordering. Naturally, this was only a transitory stage, used primarily for specialized lexica on difficult words in the *Qurʾān*, in prophetic tradition, etc., from the 10th to the 12th centuries. Ibn Durayd's (d. 321/933) voluminous *Jamharat al-luġa* is admittedly not a specialized dictionary and also uses this system, but by mixing it with the other peculiarities of the *Kitāb al-ʿayn*, its method of ordering proved to be a dead end.

The next stage, ordering by the first *and* the second radical, is known from Kurāʿ an-Naml's (d. 310/922) *al-Mujarrad*, devoted to rare words. Another two-thirds alphabetical work is the voluminous *Kitāb šams al-ʿulūm* by Našwān al-Ḥimyarī (d. about 573/1178), which aims to encompass the whole Arabic lexicon. A peculiarity is Našwān's way of grouping the words within the single chapters, defined by the first and second radicals. Within a division into nouns and verbs, he arranges the derivations of a given root according to a fixed order of morphemes, thus providing information which is not generally provided by the Arabic script.

The final stage of this system, i.e. complete alphabetical arrangement, seems to have developed as early as the 9th century. One of the earliest philologists to use the fully developed alphabetical system was ʾAbū Ḥanīfa ad-Dīnawarī (d. 282/895) in the encyclopedic section of his *Kitāb an-nabāt*, a work devoted to botany in the widest possible sense. He himself points out the practical advantage of this way of ordering. Az-Zamaxšarī's (d. 538/1144) *ʾAsās al-balāġa*, devoted to metaphorical expressions, and al-Fayyūmī's (d. about 770/1368) *Kitāb al-miṣbāḥ*, treating the technical vocabulary of Islamic jurisprudence, are two famous examples from later centuries.

5. The rhyme system

The rhyme arrangement in its mature form was used right up to modern times and can be considered as the Arabic way of ordering roots par excellence. In fact, it is closely related to the alphabetical system. The single chapters or books of these dictionaries using the rhyme system contain all roots having a common *last* radical consonant; within these parts, one has to look up entries in the same way as in an alphabetical work. It seems that this system was first developed for compiling rhyme dictionaries. This is evidenced by the title of al-Bandanījī's (d. 284/897) *Kitāb at-taqfiya*, which means 'rhyme book'. As the title suggests, it gives rhyming words (and their respective meanings) and is thus a helpful tool when composing poetry. Al-Bandanījī's work may well have been the model for al-Fārābī's (d. about 350/961) *Dīwān al-ʾadab*, but he combines the rhyme arrangement with subtle classifications of roots, parts of speech, and morphemes.

The fact that these dictionaries could be used for finding rhymes was useful not only for poets but also for the wider sections of the educated classes who needed to express themselves in rhymed prose (→ *sajʿ*). But none of all the works mentioned so far, whatever their system for arranging roots, provided users with easily accessible information on the meaning of all the lexemes of the simpler as well as the more elevated Classical Arabic vocabulary. It was al-Fārābī's nephew al-Jawharī (d. about 400/1009) who filled this need, in his *Tāj al-luġa wa-ṣiḥāḥ al-ʿarabiyya* (commonly *aṣ-Ṣiḥāḥ*). It is arranged exclusively according to the radicals in rhyme order, and from al-Jawharī's times onward until the *Tāj al-ʿarūs* (late 18th century), most important dictionaries were arranged according to this prototype. Ibn Manẓūr's (d. 711/1311) *Lisān al ʿarab* is based on the *Ṣiḥāḥ* in its arrangement and material and also contains the data from four other works (namely, the books of al-ʾAzharī and Ibn Sīda mentioned above as well as two works by Ibn Barrī and Majd ad-Dīn ibn al-ʾAṯīr). In modern Western scholarship, the *Lisān* has gained preeminent status as a work of reference for Classical Arabic vocabulary. The *Qāmūs* of al-Fīrūzābādī (d. 817/1415), also composed according to the *Ṣiḥāḥ*, was held in similarly high esteem in Arab countries, its title *al-Qāmūs* having become the Arabic word for 'dictionary'.

6. The onomasiological dictionaries

From around the end of the 8th century C.E., the first so-called onomasiological dictionaries or thesauri were composed (i.e. those which supply the notions for certain topic areas). Depending on the thematic scope of the subject matter treated, two groups can be distinguished. The first group consists of monographs on narrow semantic fields, such as treatises on camels, horses, falcons, pigeons, sheep, goats, palm trees, grapevines, the sun and moon, clouds and rain, and weapons, and also on oaths and curses. A strong interest in pre-Islamic Bedouin life is evident. These specialized treatises were composed from the earliest days of Arabic lexicography until quite late times. By the middle of the 11th century, for example, about thirty books on the parts of the human body (*xalq al-ʾinsān*) had been written, and on the eve of the Ottoman invasion of Egypt, Jalāl ad-Dīn as-Suyūṭī (d. 911/1505) contributed to this branch of lexicography by writing his *Kitāb ġāyat al-ʾiḥsān fī xalq al-ʾinsān*.

The second group is made up of books not restricted to a narrowly defined subject matter but rather at least intending to cover the complete Arabic vocabulary. One of the earliest of these works was an-Naḍr ibn Šumayl's (d. 203/819) *Kitāb aṣ-ṣifāt*, the organization of which is known to us although the work itself is lost. The earliest extant book is *al-Ġarīb al-muṣannaf* by ʾAbū ʿUbayd (d. 224/838). The organization of the subject matter does not always seem very logical to us; animals, for example, are treated in three different places in the book. The onomasiological branch of lexicography reached its zenith in Islamic Spain in the 10th and 11th centuries, starting with Ibn Sīd al-ʾAndalusī's (d. 382/992) *Kitāb as-samāʾ wa-l-ʿālam* 'Book of heaven and the world', which is said to have run to forty or even one hundred volumes but is now mostly lost. Ibn Sīda's (d. 458/1066) similarly extensive *Kitāb al-muxaṣṣaṣ* has been preserved and printed. Judging by the extant part of it, the former work was less finely subdivided than the latter, but it is probable that it served as a model, as did ʾAbū ʿUbayd's *al-Ġarīb al-muṣannaf*. The Cairo printed edition of the *Muxaṣṣaṣ* runs to seventeen large volumes. The thematic organization in the first volumes is better thought-out than in ʾAbū ʿUbayd, although there are a number of quite arbitrary insertions. From a certain point onward, Ibn Sīda seems to have given up his attempts at intelligible organization and merely arranged short chapters at random. The latter part of the *Muxaṣṣaṣ*, from Volume 13 onward, is in any case organized according to morphology, the model again being ʾAbū ʿUbayd's book. Ibn Sīda explains that only the onomasiological arrangement allows the user to find a term he does not know. In fact, his book is extremely useful for the study of the historical development of the Arabic vocabulary.

Even lexica that made no attempt at a systematic arrangement of the individual chapters could be successful. As the great number of manuscript copies show, the *Kitāb al-ʾalfāḏ al-kitābiyya* by Abd ar-Raḥmān ibn ʿĪsā al-Hamaḏānī (d. 320/932) was highly esteemed. More than three hundred chapters, bearing titles such as "To prepare for something" or "To do something well or badly", list nouns, verbs, and whole phrases, the connection between the chapters being more or less arbitrary. Another popular book of this type was aṯ-Ṯaʿālibī's (d. 429/1038) *Fiqh al-luġa*.

7. SPECIALIZED DICTIONARIES

Books on *nawādir* contain the unorganized raw material of Arabic lexicography. They explain rare and obscure (*nādir*) expressions from ancient poetry and Bedouin speech. The beginnings of this type date back to the 8th century, and its heyday was in the 9th century. ʾAbū Zayd al-ʾAnṣārī's (d. 215/830) and ʾAbū Misḥal's (d. mid-3rd/9th century) *Kitāb an-nawādir* are preserved, and both have been printed.

Books on *ġarīb al-Qurʾān* and *ġarīb al-Ḥadīṯ* explain rare and difficult words from the *Qurʾān* and from the prophetic tradition. Normally they are devoted to one corpus or the other, but ʾAbū ʿUbayd al-Harawī (d. 401/1011), in his *Kitāb al-ġarībayn*, treats lexemes and phrases from both sources. The earliest preserved book on *ġarīb al-Qurʾān* is *Tafsīr ġarīb al-Qurʾān,* ascribed to Zayd ibn ʿAlī (d. 122/740), grandson of the caliph ʿAlī ibn ʾAbī Ṭālib. However, most of it is now considered apocryphal. We are not on firm ground until we reach Ibn Qutayba's (d. 276/889) *Tafsīr ġarīb al-Qurʾān* and ʾAbū ʿAbdallāh al-Yazīdī's (d. ca. 313/925) *Ġarīb al-Qurʾān*. Both books give the words they explain in the order of their occurrence in the *Qurʾān*, i.e. from Sura 1 to Sura 114. In ar-Rāġib al-ʾIṣfahānī's (d. 502/1108) *al-Mufradāt fī ġarīb al-Qurʾān*, the notion of *ġarīb* is given such a wide meaning that the book is in fact a concise dictionary of Qurʾānic language. It is arranged in full alphabetical order. The earlier works on *ġarīb al-Ḥadīṯ* follow the order in which prophetic traditions are arranged in the large *Ḥadīṯ* collections: either like a *musnad* collection (ʾAbū ʿUbayd al-Qāsim ibn Sallām [d. 224/838] and Ibn Qutayba [d. 276/889]) or like a *muṣannaf* collection (al-Ḥarbī [d. 285/898]). The first dictionaries to use a strictly alphabetical order date from comparatively late times (Ibn al-Jawzī [d. 597/1200]: *Ġarīb al-Ḥadīṯ*; Majd ad-Dīn ibn al-ʾAṯīr [d. 606/1210]: *an-Nihāya fī ġarīb al-Ḥadīṯ wa-l-ʾaṯar*).

Works on *ʾaḍdād* are devoted to homonyms with two meanings which in some way are opposed to each other (→ *ḍidd*). The great interest Arab philologists took in this phenomenon can perhaps be explained by the role the theory of *ʾaḍdād* played in the exegesis of the *Qurʾān* (references in Seidensticker 2002:158, n. 23). Books on *ʾaḍdād* were composed from the last decades of the 8th century onward. Their total number amounts to nearly two dozen, about half of which are preserved. Among them an early example is Quṭrub's (d. 206/821) *Kitāb al-ʾaḍdād*. The matter of the *ʾaḍdād* was intensely discussed within the larger framework of the so-called *Šuʿūbiyya* quarrel, i.e. the dispute about the merits of Arabic culture compared to, principally, the Iranian cultural tradition. The Arabs' opponents argued that the large number of such lexemes could only be the result of intellectual confusion. The Arab reaction to this charge in part denied the existence of contradictory meanings, and in part tried to qualify and explain the phenomenon. ʾAbū ṭ-Ṭayyib al-Luġawī (d. 351/962) adds an appendix to his *Kitāb al-ʾaḍdād*, listing 'pseudo-*ʾaḍdād*', and Ibn al-ʾAnbārī (d. 328/940), in his book of the same title, argues that the seemingly contradictory meanings have a common semantic origin and that the context normally provides clarity. Many books on *ʾaḍdād* did not order the words treated; al-Luġawī groups them according to the first radical; and aṣ-Ṣaġānī (d. 650/1252) uses a fully alphabetical arrangement.

Books on homonyms were composed from the beginning of the 9th century. An early instance is ʾAbū l-ʿAmayṯal's (d. 240/854) *Kitāb mā ttafaqa lafḏuhu wa-xtalafa maʿnāhu*. Here, no system of ordering is discernible at all. Al-Mubarrad (d. 285/898) devoted a small book to homonyms and ellipses in the *Qurʾān*; perhaps his apologetic mode of argumentation is due to the fact that the matter was also discussed in the *Šuʿūbiyya* quarrel. Ibn aš-Šajarī's (d. 542/1148) book, arranged according to the first radical, contains no fewer than 1,670 lemmata. Kurāʿ an-Naml (d. 310/922) preferred the onomasiological arrangement, which is quite surprising in the case of homonyms, because in theory every word ought to be included in at least two places. In fact, the author decided to enter each word only under the more common meaning and to give the other meanings under that heading.

The *Dīwān al-ʾadab* by al-Fārābī (d. 350/961), mentioned in Section 5, orders words according to parts of speech, simple or extended morphemes, and vocalization, on the level just below the root class. Because of this arrangement and the huge amount of material included, the *Dīwān al-ʾadab* is the most important Arabic dictionary with a morphological arrangement. In addition, the last volumes of Ibn Sīda's (d. 458/1066) *Muxaṣṣaṣ* may be counted among these works, as well as ʾAbū ʿUbayd's book, which Ibn Sīda uses as a model (see above, Sec. 6). From the 9th to the 13th centuries, several books were written to discuss the question of the agreement or difference in meaning of the verbal Forms I and IV. Aṣ-Ṣaġānī (d. 650/1252) composed some smaller works on the lexemes of the morpheme types *infiʿāl, faʿalān, faʿāli*, and *yafʿūl.*

The term used by Arab philologists for the phenomenon of pairs of synonymous lexemes which differ in just one of the radicals (e.g. *jadaṯ* and *jadaf* 'grave') was → *ʾibdāl*. Ibn as-Sikkīt's (d. about 246/860) *Kitāb al-qalb wa-l-ʾibdāl* treats these pairs in 36 chapters, each one devoted to one of the consonants which can replace each other. The most voluminous book in this area, ʾAbū ṭ-Ṭayyib al-Luġawī's (d. 351/962) *Kitāb al-ʾibdāl*, arranges the chapters in strictly alphabetical order.

Language change posed a special problem in Classical Arabic because the canonization of the pre-Abbasid educated standard language led to a markedly conservative attitude. Not surprisingly, the literature devoted to cleansing the language is quite extensive. Along with semantic 'errors', deviant vocalization and morphology were also denounced. The first extant work of this ilk is *Kitāb mā yalḥanu fīhi l-ʿawāmm* by al-Kisāʾī (d. 189/805), and many other books bear the same title or were called *Laḥn al-ʿawāmm*. Despite the term *ʿāmma/ʿawāmm* 'common people', the target of the criticism is not colloquial or dialectal Arabic but rather insufficient mastery of standard Arabic. Many books lack an alphabetical or onomasiological arrangement of subject matter, although in some cases a distinction is made between formal and semantic offenses. Ibn al-Jawzī's (d. 597/1201) *Taqwīm al-lisān* arranges its material according to the first radical. Particularly prominent are Ibn as-Sikkīt's (d. about 246/860) *ʾIṣlāḥ al-manṭiq* and Ṯaʿlab's (d. 291/904) *Kitāb al-faṣīḥ*, which both present the correct forms and usages, arranged according to morpheme and vocalization in large numbers of chapters (about one hundred and forty, respectively). The large number of manuscript copies, commentaries, and extracts shows that they were very popular. Some works devote special attention to pairs of lexemes which differ only in one consonant of similar pronunciation. As the titles suggest, the risk of error was especially great in the case of the phonemes *ḍād* and *ḏ̣āʾ*.

8. Characteristics

Arabic lexicography did not develop a theory of semantics or lexicology. A special branch of thought called *ʿilm al-waḍʿ* (→ *waḍʿ al-luġa*) touched upon questions of semantics and the philosophy of language, but as it did not emerge before the 14th century, it did not influence lexicographical practice, which had reached its final form centuries earlier.

From the very beginning, compilers of Arabic dictionaries attached great importance to quotations to illustrate the particular meaning of a lexeme. In some cases, data were collected by special field research. Several philologists of the 9th century are said to have traveled extensively in order to receive instruction from the Bedouin of central Arabia. For the modern user, it is important to be aware of the limitations of what was considered worth explaining. Apart from the *Qurʾān*, prophetic tradition, proverbs, and ancient Bedouin prose, only pre-Abbasid poetry prior to about 730 C.E. was considered worthy of treatment. Consequently, large areas of the Arabic lexicon were completely ignored by Arab lexicographers. Some new material was added in the 10th century, but from the beginning of the 11th century the predominant method was to recompile material from earlier dictionaries.

From the modern point of view, the definitions given in the ancient lexica have many deficiencies. Metaphorical and rare use stand indistinguishably side by side with literal and common meanings. In addition, meanings are wrongly deduced from the context or simply guessed at. In general, the formulation of an abstract lexical definition was not considered an aim of prime importance (on definitions in Arabic lexica, see the literature adduced in Seidensticker 2002:164, n. 42).

An interesting exception to this general rule is ʾAḥmad ibn Fāris's (d. 395/1005) *Muʿjam maqāyīs al-luġa*. As its aim is to trace back all derivations of a root to one or two primary meanings, the author does not direct his attention to single quotations or special usages but rather tries to find a common origin for several lexemes.

Not surprisingly, religious matters had repercussions for Arabic lexicography (cf. Kopf 1956). The notion of the divine origin of language certainly contributed much to the conservative attitude to the Arabic language because it did not allow for language change. Al-ʾAṣmaʿī (d. 213/828) is said to have kept aloof from certain philological problems in the *Qurʾān* and the prophetic tradition in order not to come into conflict with traditional exegesis. In using the ancient dictionaries, it is useful to know that in some cases meanings are given which have their origins in exegetic or dogmatic disputes (cf. Rippin 1983). A particularly delicate issue was the question of foreign words, especially in the *Qurʾān*. Many thought it hardly conceivable that there should be words of non-Arabic origin in a text which styled itself 'a clear Arabic book'. Some early and some late authorities did not take offense at that possibility, but others objected. ʾAbū ʿUbayd (d. 224/838) prudently argued that some foreign words dated from the pre-Islamic period. The famous jurist aš-Šāfiʿī (d. 204/820), the philologist ʾAbū ʿUbayda (d. about 213/828), and the historian and commentator on the *Qurʾān* aṭ-Ṭabarī (d. 310/923) denied that there were any such borrowings but rather asserted a coincidental similarity in the articulation of words with a similar meaning in two languages (cf. Kopf 1956, Sec. 3; Gilliot 1990, Chap. 4). A separate set of lexicographical monographs on the question developed only at a later date. The most famous representative is al-Jawālīqī's (d. 539/1114) *al-Muʿarrab*, which orders the words according to the first letter only.

9. Further reading

For more bibliographical references, see Seidensticker (2002). References to many printed editions of Arabic dictionaries can be found in Weipert (2002). An important work of reference for bio-bibliographical information for the time up to about 430/1038 is Sezgin (1982; Sezgin 1984:310–319 supp.; and Weipert 1989:228–246). A weighty contribution to the history of Classical Arabic lexicography, arranged chronologically, is Kraemer (1953). Haywood (1960) is the fullest monograph on the topic in a Western language but is outdated now. The most complete overview in a monograph in Arabic is Naṣṣār (1968). Important information far beyond the topic proper (al-Xalīl's *Kitāb al-ʿayn*) is given in Wild (1965).

Bibliographical references

Primary sources

ʾAbū ʿAmr, *Jīm* = ʾAbū ʿAmr aš-Šaybānī, *Kitāb al-jīm*. Ed. ʾIbrāhīm al-ʾIbyārī a.o. 3 vols. Cairo: al-Hayʾa al-ʿĀmma li-Šuʾūn al-Maṭābiʿ al-ʾAmīriyya, 1974–1975.

ʾAbū ʿUbayd, *al-Ġarīb al-muṣannaf* = ʾAbū ʿUbayd al-Qāsim ibn Sallām al-Harawī, *al-Ġarīb al-muṣannaf*. Ed. Muḥammad al-Muxtār al-ʿUbaydī. 3 vols. Carthage: al-Majmaʿ at-Tūnisī li-l-ʿUlūm wa-l-ʾĀdāb wa-l-Funūn; Tunis: Dār Saḥnūn li-n-Našr wa-t-Tawzīʿ, 1989–1996.

ʾAzharī, *Tahḏīb* = ʾAbū Manṣūr Muḥammad ibn ʾAḥmad al-ʾAzharī, *Kitāb tahḏīb al-luġa*. Ed. ʿAbd as-Salām Muḥammad Hārūn a.o. 15 vols. Cairo: al-Muʾassasa al-Miṣriyya al-ʿĀmma, 1964–1967.

Fārābī, *Dīwān* = ʾAbū ʾIbrāhīm ʾIsḥāq ibn ʾIbrāhīm al-Fārābī, *Dīwān al-ʾadab*. Ed. ʾAḥmad Muxtār ʿUmar. 4 vols. Cairo: al-Hayʾa al-ʿĀmma li-Šuʾūn al-Maṭābiʿ al-ʾAmīriyya, 1974–1979.

Fīrūzābādī, *Qāmūs* = Majd ad-Dīn ʾAbū ṭ-Ṭāhir Muḥammad ibn Yaʿqūb al-Fīrūzābādī, *al-Qāmūs al-muḥīṭ*. 4 vols. Cairo: al-Maṭbaʿa al-Miṣriyya, 1913.

Ibn Durayd, *Jamhara* = ʾAbū Bakr Muḥammad ibn al-Ḥasan ibn Durayd, *Kitāb jamharat al-luġa*. 4 vols. Hyderabad: Maṭbaʿa Majlis Dāʾirat al-Maʿārif al-ʿUṯmāniyya, 1925–1932 (and reprints).

Ibn Fāris, *Maqāyīs* = ʾAbū l-Ḥusayn ʾAḥmad ibn Fāris, *Muʿjam maqāyīs al-luġa*. Ed. Abd as-Salām Muḥammad Hārūn. 6 vols. Cairo: Dāʾira ʾIḥyāʾ al-Kutub al-ʿArabiyya, 1946–1952. (2nd ed. 1969–1972.)

——. *Mujmal* = ʾAbū l-Ḥusayn ʾAḥmad ibn Fāris, *Kitāb mujmal al-luġa*. Ed. Hādī Ḥasan Ḥammūdī. 5 vols. Kuwait: al-Munaḏḏama al-ʿArabiyya li-t-Tarbiya wa-ṯ-Ṯaqāfa wa-l-ʿUlūm, 1985.

Ibn Manẓūr, *Lisān* = Jamāl ad-Dīn ʾAbū l-Faḍl Muḥammad ibn Mukarram ibn Manẓūr, *Lisān al-ʿArab*. 15 vols. Beirut: Dār Ṣādir, 1955–1956 (and reprints).

Ibn Sīda, *Muḥkam* = ʾAbū l-Ḥasan ʿAlī ibn ʾIsmāʿīl ibn Sīda, *Kitāb al-muḥkam wa-l-muḥīṭ al-ʾaʿḏam fī l-luġa*. Ed. Muṣṭafā as-Saqqā a.o. 12 vols. Cairo: Maʿhad al-Maxṭūṭāt bi-Jāmiʿat ad-Duwal al-ʿArabiyya, 1958–1999.

——. *Muxaṣṣaṣ* = ʾAbū l-Ḥasan ʿAlī ibn ʾIsmāʿīl ibn Sīda, *Kitāb al-muxaṣṣaṣ fī l-luġa*. 17 vols. Cairo: al-Maṭbaʿa al-ʾAmīriyya, 1316–1321 A.H. (and reprints).

Ibn as-Sikkīt, *ʾIṣlāḥ* = ʾAbū Yūsuf Yaʿqūb ibn ʾIsḥāq as-Sikkīt, *Kitāb ʾiṣlāh al-manṭiq*. Ed. ʾAḥmad Muḥammad Šākir and ʿAbd as-Salām Muḥammad

Hārūn. Cairo: Dār al-Maʿārif, 1949. (2nd ed. 1956.)

Jawālīqī, *Muʿarrab* = ʾAbū Manṣūr Mawhūb ibn ʾAḥmad al-Jawālīqī, *Kitāb al-muʿarrab*. Ed. ʾAḥmad Muḥammad Šākir. Cairo: Maṭbaʿa Dār al-Kutub al-Miṣriyya, 1361 A.H. (Repr., Tehran 1966.)

Jawharī, *Ṣiḥāḥ* = ʾAbū Naṣr ʾIsmāʿīl ibn Ḥamad al-Jawharī, *Tāj al-luġa wa-ṣiḥāḥ al-ʿarabiyya*. Ed. ʾAḥmad ʿAbd al-Ġafūr ʿAṭṭār. 6 vols. Cairo: Dār al-Kitāb al-ʿArabī, 1376–1377 A.H.

Našwān, *Šams* = Našwān ibn Saʿīd al-Ḥimyarī, *Šams al-ʿulūm wa-dawāʾ kalām al-ʿArab min al-kulūm*. Ed. Ḥusayn ibn ʿAlī al-ʿAmrī a.o. 12 vols. Damascus: Dār al-Fikr, 1999.

Ṯaʿālibī, *Fiqh* = ʾAbū Manṣūr ʿAbd al-Malik ibn Muḥammad aṯ-Ṯaʿālibī, *Kitāb fiqh al-luġa wa-sirr al-ʿarabiyya*. Ed. Muṣṭafā as-Saqqā a.o. Cairo: Muṣṭafā al-Bābī al-Ḥalabī, 1972.

Ṯaʿlab, *Faṣīḥ* = ʾAbū l-ʿAbbās ʾAḥmad ibn Yaḥyā Ṯaʿlab, *Kitāb al-faṣīḥ*. Ed. Jacob Barth. Leipzig: Hinrichs, 1876. Ed. ʿĀṭif Madkūr. Cairo: Dār al-Maʿārif, 1984.

Xalīl, *ʿAyn* = al-Xalīl ibn ʾAḥmad al-Farāhīdī, *Kitāb al-ʿayn*. Ed. Mahdī al-Maxzūmī and ʾIbrāhīm as-Sāmarrāʾī. 8 vols. Baghdad: Dār aš-Šuʾūn aṯ-Ṯaqāfiyya al-ʿĀmma, 1980–1985.

Zabīdī, *Tāj* = ʾAbū l-Fayḍ Muḥammad Murtaḍā ibn Muḥammad az-Zabīdī, *Tāj al-ʿarūs min jawāhir al-Qāmūs*. 10 vols. Cairo: al-Maṭbaʿa al-Xayriyya, 1306–1307 A.H. (and reprints). Ed. ʿAbd as-Sattār ʾAḥmad Farrāj a.o. 38 vols. Kuwait, 1965ff.

Zamaxšarī, *ʾAsās* = Jār Allāh ʾAbū l-Qāsim Maḥmūd ibn ʿUmar az-Zamaxšarī, *Kitāb ʾasās al-balāġa*. Ed. ʿAbd ar-Raḥīm Maḥmūd. Cairo: Dār al-Kutub al-Miṣriyya, 1953.

Secondary sources

Gilliot, Claude. 1990. *Exégèse, langue et théologie en Islam*. Paris: Vrin.

Haywood, John A. 1960. *Arabic lexicography: Its history, and its place in the general history of lexicography*. Leiden: E.J. Brill. (2nd ed. 1965.)

Kopf, Lothar. 1956. "Religious influences on Medieval Arabic philology". *Studia Arabica* 5.33–59.

Kraemer, Jörg. 1953. "Studien zur altarabischen Lexikographie". *Oriens* 6.201–238.

Law, Vivien. 1990. "Indian influence on early Arab phonetics: Or coincidence?". *Studies in the history of Arab grammar*, II, ed. Michael G. Carter and Kees Versteegh, 215–227. Amsterdam: J. Benjamins.

Naṣṣār, Ḥusayn. 1968. *al-Muʿjam al-ʿarabī: Našʾatuhu wa-taṭawwuruhu*. 2 vols. 2nd ed. Cairo: Maktaba Miṣr.

Rundgren, Frithiof. 1973. "La lexicographie arabe". *Studies on Semitic lexicography*, ed. Pelio Fronzaroli, 145–159. Florence: Istituto di Linguistica e di Lingue Orientali, Università di Firenze.

Seidensticker, Tilman. 2002. "Die einheimische arabische Lexikographie: Ein Überblick". *Neue Beiträge zur Semitistik*, ed. Norbert Nebes, 147–166. Wiesbaden: O. Harrassowitz.

Sezgin, Fuat. 1982. *Geschichte des arabischen Schrifttums. VIII. Lexikographie bis ca. 430 h.* Leiden: E.J. Brill.

——. 1984. *Geschichte des arabischen Schrifttums. IX. Grammatik bis ca. 430 h.* Leiden: E.J. Brill.

Weipert, Reinhard. 1989. "Beiträge zur *Geschichte des arabischen Schrifttums. II. Lexikographie und Grammatik*". *Zeitschrift für Geschichte der Arabisch-islamischen Wissenschaften* 5.225–264.

——. 2002. *Classical Arabic philology and poetry: A bibliographical handbook of important editions from 1960 to 2000*. Leiden: E.J. Brill.

Weninger, Stefan. 1994. "Das 'Übersetzerbuch' des Elias von Nisibis (10./11. Jh.) im Zusammenhang der syrischen und arabischen Lexikographie". *The world in a list of words*, ed. Werner Hüllen, 55–66. Tübingen: Niemeyer.

Wild, Stefan. 1965. *Das Kitāb al-ʿAin und die arabische Lexikographie*. Wiesbaden: O. Harrassowitz.

Tilman Seidensticker (Jena University)

Lexicography: Monolingual Dictionaries

1. Introduction

In the first half of the 2nd century A.H. (8th century C.E.), speakers of Arabic encountered numerous communication problems due to the variability of their language at the time. Natural processes of language change, due primarily to the lack of standardization, were evident in the way people pronounced words, structured words morphologically, and structured sentences. This phenomenon spurred Sībawayhi (d. 168/784) to write the first comprehensive Arabic grammar, *al-Kitāb*. This work set forth rules for all aspects of grammar, including phonology, morphology, and syntax.

Much of the variability in grammar during this period was evident in how people used words to convey meaning. Speakers of Arabic used familiar words in novel ways, which created the need for a standardized dictionary of Arabic. This need was satisfied when al-Xalīl (d. 175/791) developed the *Kitāb al-ʿayn*, the first Arabic dictionary. Thus began the discipline of Arabic lexicography (ʿAṭṭār 1990:11; → lexicography: Classical Arabic).

This entry focuses primarily on the preparation and development of modern Arabic/Arabic dictionaries and how they compare with medieval Arabic/Arabic dictionaries. It also discusses the methodologies and techniques used in the creation of Arabic dictionaries and provides a critical analysis of these methods. The first section includes some preliminary notes and details on the history of Arabic dictionaries from both medieval and modern times. Section 2 discusses modern Arabic/Arabic dictionaries in greater length. Section 3 provides a critical analysis of the techniques used in the development of Arabic dictionaries. Finally, Section 4 provides a brief conclusion.

2. Medieval versus modern Arabic/Arabic dictionaries

The discipline of Arabic lexicography began with the development of the *Kitāb al-ʿayn* by al-Xalīl (→ lexicography: Classical Arabic). This first Arabic dictionary was organized according to place of articulation of the first sound of the word. Between the end of the 8th century, when the *Kitāb al-ʿayn* was written, and the end of the 18th century C.E., 43 Arabic dictionaries were created, only 14 of which can be considered 'general' Arabic/Arabic dictionaries, as the remainder included specialized or technical vocabulary and items from tribal vocabulary, and some dictionaries contained linguistically relevant information in the entries (Muḥammad 2001:36).

The term 'dictionary' in itself can have a number of meanings. A dictionary can be a list of words that correspond to the vocabulary used by a language community. Most often, a dictionary is a book that lists words along with corresponding meanings, which is helpful when dealing with obscure words or less frequently used definitions (Bahnasāwī 1990:8). Arabic dictionaries generally serve both purposes.

Many different strategies are used in writing definitions for Arabic dictionaries, and these strategies vary according to both the dictionary compiler and the words to be defined. One method for defining a word is to list synonyms. For example, *ʾasad* 'lion' is defined in some dictionaries by using the word *layṯ*, a similar but vaguer term. Associations are frequently used in definitions: *ḥarb* 'war' can be listed in the definition of *waġā* 'battle sound,' since they are closely related. Antonyms are also used in defining words, e.g., *ṭawīl* 'tall' can be defined in terms of its opposition to *qaṣīr* 'short' (al-Jilālī 1999:20–24).

The primary purpose of medieval Arabic dictionaries was similar to that of modern dictionaries: to introduce knowledge about words to the user. It is important to consider what type of knowledge was encoded in these works. In medieval Arabic dictionaries, the primary type of information that was conveyed to the user is termed → *maʿnā* (pl. *maʿānin*), which basically translates as 'meaning'. A distinction is made between *maʿnā* and *dilāla* (or *dalāla*) 'semantics'. *Maʿnā*, or more technically, lexical *maʿnā*, is an entry in a dictionary that contains one or more definitions as well as notes on spelling, pronunciation, and morphology (i.e. case endings and irregularities). Lexical meaning, therefore, consists of three primary aspects: denotation of the word; connotations associated with the word; and correspondences between the denotation and connotations. *Dilāla*, on the other hand, is specific to words used in context and is generally not included in entries in Arabic dictionaries.

One issue with defining *maʿnā* is the problem of circularity of reference – the classic problem of stating exactly what a definition is. Two facts from Arabic lexicography illustrate this phenomenon fairly well. First, Arabic words often have multiple meanings. Often, these meanings are given classifications, such as central, basic, marginal, additional, impressive, or stylistic. In the second place, Arabic lexicographers seldom agreed on a specific meaning for a word. As discussed above, when defining a word, medieval Arabic lexicographers relied on a number of techniques, such as providing additional semantic content and using synonyms and antonyms. Medieval Arabic lexicographers also relied on a number of techniques to bolster the entry. One major method used to clarify the definition of a dictionary entry was to provide the user with illustrative examples from literary and religious texts (*šawāhid*). In addition, authors offered comprehensive conceptual definitions and, for further support, traced the etymological development of words.

These techniques seem to be subjective and somewhat in conflict with one another (and thus suspect), if the distinction between *maʿnā* and *dilāla* is considered. This conflict is evidenced in the interaction between Arabic lexicographers and Arabic writers of poetry and prose. Arabic lexicographers generally assigned the most narrow, basic meanings to words, and literary use of vocabulary was constrained by those definitions. This greatly affected the poets of the time, as the figurative use of words was violently condemned by Arabic literary critics. The use of old or no longer used words was also fiercely condemned, even though these archaisms were present in the dictionaries and were often listed as synonyms in the entries for more common words. Colloquialisms (even those derived from Classical Arabic, yet unrecognizable to the general reader) and foreign words (which are clearly not within the scope of Arabic dictionaries) were strictly banned from medieval Arabic prose and poetry. Thus, a conspiracy arose (either intentional or latent) between literary critics and the authors of dictionaries, which propagated conservatism within the Arabic language. Definitions for dictionaries were written in a very narrow way, supported by literary citations. In turn, the literary critics used these dictionaries as part of the basis for their criticism, excluding new literary works on the grounds that they did not conform to the standards of the Arabic language. In this manner, a very traditional, conservative form of Arabic was maintained, while linguistic innovations were suppressed from literary culture (Hassanein 1977:210–212).

As an extension of the medieval tradition, modern Arabic/Arabic dictionaries generally resemble medieval Arabic dictionaries, with one major exception. In recent times, it has become acceptable to develop dictionaries with an encyclopedic bent, allowing entries detailing historical, political, and geographical facts such as the names of well-known figures, governmental agencies, historic battles, rivers, mountains, cities, countries, and the like.

3. The development of modern Arabic/Arabic dictionaries

3.1 *The modern age of Arabic lexicography*

The beginning of modern Arabic lexicography can be traced to the second half of the 18th century, with the development of *Muḥīṭ al-muḥīṭ* by Buṭrus al-Bustānī (d. 1301/1883), published in 1870. The modern age of Arabic lexicography continues till the present time; the *Qāmūs al-hādī*, published in Tripoli (Libya) in 1994 by Fāyiz Yūsuf Muḥammad, is the most recent dictionary to be produced.

During the modern era of Arabic lexicography, sixty Arabic/Arabic dictionaries have been produced. Thus, in only a century and a quarter, about a third more dictionaries were written than in the ten centuries of the medieval period (60 dictionaries were written between 1870 and the present, as against 43 between the 8th and 18th centuries). One of the factors contributing to this dramatic increase in the number of dictionaries published is modern printing technology, i.e. mass-produced paper and ink and the accessibility of printing presses and publishing houses. The efforts of Lebanese lexicographers should be noted as well; their extensive experience and novel techniques have proved invaluable to the field. The *Šāmī* authors (primarily Syrian, Lebanese, and Palestinian) have also contributed greatly, not only to the field of lexicography but also to the disciplines of translation, journalism, and the Arabic literary tradition (Hassanein 1983:17).

Modern Arabic/Arabic dictionaries have primarily been published in Beirut. However, others have been published in Damascus, Cairo, Morocco, and Libya. Many of the early dictionaries were written by individuals, while Arabic → language academies in Damascus and Cairo and UNESCO (United Nations Educational, Scientific, and Cultural Organization) in Tunisia recently sponsored the development of Arabic dictionaries. Individual efforts continue to this day, either in conjunction with, or independent of, their corporate sponsors.

3.2 *The early modern period*

In the modern age of Arabic lexicography, three distinct stages of development can be identified. The dictionaries developed in any one stage generally share some common features, and a natural trend is that the dictionaries developed in one stage assimilate the advances made in previous stages. The first stage of modern Arabic lexicography extends for sixty years, from 1870 to 1930. Four works from this period can be distinguished as being especially important in the history of Arabic lexicography:

i. *Muḥīṭ al-muḥīṭ*, by al-Bustānī (Beirut, 1870)
ii. *al-Munjid*, by al-ʾAb Lūwīs Maʿlūf al-Yasūʿī (Beirut, 1908)
iii. *ʾAqrab al-mawārid*, by Saʿīd aš-Šarṭūnī (Beirut, 1912)
iv. *al-Bustān*, by ʿAbdallāh al-Bustānī (Beirut, 1930)

These dictionaries share one common feature in that they list entries based on the trilateral → root (a distinguishing characteristic of most Semitic languages, such as Hebrew and Arabic). Technical terms of the day are included in these four works, as well as proverbs and idioms, terms related to Christianity, and some encyclopedic entries (the names of historical figures, cities, geographical points of interest, etc.). Terms related to sexual relations are strictly excluded.

The lexical data presented in these four dictionaries are derived from medieval lexicons such as the *Lisān al-ʿArab* by Ibn Manẓūr (d. 711/1311), *al-Qāmūs al-muḥīṭ* by al-Fayrūzābādī (d. 817/1414), and the *Tāj al-ʿarūs* by az-Zabīdī (d. 1205/1791; Daqqāq 1977:191–199).

Of these four dictionaries of the early modern period, the *ʾAqrab al-mawārid* by aš-Šarṭūnī is the most systematic, presenting linguistic data including verbal and nominal forms, verb tense forms, and internal vocalization for all entries. Repeated words are marked with a hyphen. The format of the dictionary is set up to facilitate research by the user, with each page divided into three columns and each column headed with a word title.

Al-Munjid is perhaps the best of the four in that it employs an advanced format and the subject matter is presented in a unique way, similar to modern European dictionaries.

3.3 *The middle modern period*

The middle period of modern Arabic lexicography runs from 1946 to 1989. It is marked by a general movement away from individual production and toward development sponsored by language academies, established seminars, and the Arab League Educational, Cultural and Scientific Organization (ALECSO). Individual efforts were still involved, but these individuals worked in conjunction with the staff at such institutions. In this period, many common Arabic dictionaries were produced. The pioneering work of this period is *al-Muʿjam al-kabīr*, issued by the Fūʾād al-ʾAwwal (King Fuad) Academy for Arabic. Its first edition was printed in 1946, and it was reprinted in 1956 (twice) and in 1970, after the academy changed its name to the Arabic Language Academy in Cairo. This work consists of one volume and is unique in the fact that it organizes data based on certain parameters: basic to subordinate; concrete to abstract; direct to metaphoric; and familiar to odd. Verbs are presented in one section, followed by nouns, and texts are quoted to support definitions. It compares Arabic words with historically related terms from other Semitic languages. It is meant to serve both traditionalists and modernists equally well (Darwīš 1956:147; also Daqqāq 1977:225).

Another important dictionary from the middle modern period is *Matn al-luġa*, developed by ʾAḥmad Riḍā and inspired by the instructions of the Language Academy in Syria. Riḍā supervised the development of *Matn al-luġa* and participated in the editing process in Damascus in 1958. The five-volume work was published in Beirut in 1958. One distinctive feature of this dictionary is that it excludes technical terms from the arts and sciences. Riḍā chose to omit these terms because he believed their production did not belong to the pure language. When compiling the definitions of entries, Riḍā also chose to exclude *istišhādāt* 'cited literary examples' and lengthy explanations. *Matn al-luġa* should be credited, however, with including new meanings for words employed by contemporary poets and writers.

In comparison to rival dictionaries of the same era, *Matn al-luġa* must be regarded as a unique lexicon. It should be noted that Riḍā was not completely free from traditional notions of Arabic lexicography in his development of this work. For example, he included an immense number of names of geographical locations that were either vague or had fallen into disuse in contemporary speech and had therefore become irrelevant to the users of his dictionary.

Following the publication of *Matn al-luġa*, language academies began to play a more direct role in the development of new dictionaries. The Arabic Language Academy in Cairo issued two good dictionaries, the first of them being *Muʿjam al-wasīṭ*, which was first edited in 1960–1961 in two 1,200-page volumes. Each page was divided into three columns, and the dictionary contained approximately thirty thousand entries. The second lexicon developed by the Academy in Cairo was an abridged version of the *Muʿjam al-wasīṭ*, entitled *Muʿjam al-wajīz* 'Concise dictionary' and published in 1980. Both works, but to a greater extent *al-Wajīz*, were well received by students and scholars.

These dictionaries are still in use by university professors and research students, as well as by students in primary and secondary schools.

The *Muʿjam al-wasīṭ* excels in several aspects, primarily due to the collaboration between scholars and researchers who possessed a great deal of expertise in the Arabic language and in lexicography. The compilers of this work were highly regarded in their fields of study, not only in Egypt but in other Arabic-speaking countries as well, and the Academy is well-recognized for the authentic criteria and impeccable standards used in the creation of *al-Wasīṭ*. This dictionary (and its abridged form) excludes entries for archaic terms that are irrelevant to modern usage. Words are arranged in alphabetical order, which is generally considered to be a good technique. Single-root derived forms are organized into categories, and related roots are listed at the beginning of each entry.

Additional supporting data adduced in *al-Wasīṭ* are quotations from the *Qurʾān*, prophetic traditions, common literary expressions and proverbs, and established rhetorical expressions. A problem that has traditionally plagued Arabic lexicographers has been how to express the middle vowel of a present tense verb or a verbal noun derived from Form I. The editors of *al-Wasīṭ* overcame this dilemma by placing a symbol that designated the proper pronunciation next to entries of this form. Illustrations were used for entries for types of birds, plants, and tools.

In spite of *al-Wasīṭ's* success, minor shortcomings remain evident. For example, some non-Arabic words and borrowings are included, under the assumption that non-borrowed equivalents in Arabic will be found at a later time. Another shortcoming – not limited to *al-Wasīṭ* but common among Arabic dictionaries as a whole – is a failure to introduce definitions in entries on the basis of frequency of use. In addition, grammatical notions are not exhaustively presented for all entries in *al-Wasīṭ*.

In 1965, through the individual efforts of Jibrān Masʿūd, *ar-Rāʾid* was published in Beirut in two volumes. Each volume consisted of 1,638 pages, and each page was divided into two columns. *Ar-Rāʾid* is unique in a number of ways. Entries are organized according to pronunciation, making it much more user friendly. In addition, lexical entries have greatly simplified definitions, many of which are accompanied by illustrations. More important, definitions are listed in the entries sequentially, so that the most common or most frequently used definitions appear first. Novel definitions for words, as well as newly coined terms and foreign borrowings, are present in *ar-Rāʾid*. Lebanese words are also included, making this dictionary more appropriate for an audience from that region.

The dictionary that marks the end of the middle modern period in Arabic lexicography is the *Muʿjam al-ʿarabī al-ʾasāsī*, produced by ALECSO in 1989. It is divided into about twenty-five thousand sections, each containing derived and conjugated forms based on a trilateral root. Each entry is well-explained and supported with examples from the *Qurʾān,* the *Ḥadīṯ*, proverbs, and expressions from contemporary Arabic. This dictionary is regarded by scholars as a reference for special topics and is therefore not accepted as a general-use dictionary. The importance of this work lies in the organization and functionality of the entries, and thus it is duly noted in this survey.

3.4 *The late modern period*

The late modern period of Arabic lexicography extends from 1991 to the present. Four works should be noted in particular:

i. *al-Hādī ʾilā luġat al-ʿArab*, by Ḥassān Saʿīd al-Karmī (1991)
ii. *al-Muḥīṭ*, by ʾAdīb al-Lujamī (1993)
iii. *Luġat al-ʿArab*, by George Mitrī ʿAbd al-Masīḥ (1993)
iv. *al-Kāfī*, by unknown author(s), printed by the Books Company, Beirut (1994)

Al-Hādī is the most extensive of these four works, consisting of 2,372 pages in four olumes. This dictionary has good print quality and colorful graphics; in addition, it classifies the data according to the trilateral root followed by other verb forms. It includes many terms from modern technology and science, politics, and industry. It also cites the names of many plants, animals, and gems and some cosmic phenomena, accompanied by their English equivalents. One criticism of this work is that it is physically large, weighing about six kilograms. Another complaint is that it contains a great many colloquialisms (Maʿtūq 1999:113–119).

Al-Muḥīṭ is more authentic than *al-Hādī*. It contains a great many color illustrations depicting parts of human anatomy, instruments, audiovisual equipment, and aspects related to the generation of electrical and nuclear power. Photographs of famous international sports figures and stadiums, farms, factories, and national flags accompany pertinent entries. In this manner, *al-Muḥīṭ* resembles an encyclopedia. One commonly praised aspect of this work is that it maintains many of the high usages from Classical Arabic, rendering it appropriate for advanced users (Maʿtūq 1999:125).

Luġat al-ʿArab, published in 1993, is a lengthy work of three volumes. It was developed using old and new lexicographical methodologies, and it includes both traditional and archaic language, as well as much modern usage. This work is arranged according to the trilateral root, and in most cases words are cross-referenced with other words. It is not uncommon for an entry to appear in more than one location. Definitions within entries are not listed according to frequency of use (Maʿtūq 1999:125–133).

The one-volume dictionary *al-Kāfī* was published in 1994. It contains 116 pages, divided into two columns per page, and is organized according to pronunciation. This dictionary has many shortcomings and appears not to have benefited from the advancements made in prior works (Maʿtūq 1999:133).

In addition to the historical stages of development of Arabic lexicography described in the previous four sections, there are about twenty works that are abridged or derived from previous dictionaries. Most of these dictionaries were published primarily to serve students. For example, al-Fayyūmī's (d. 770/1368) *al-Miṣbāḥ al-munīr* is an abridged version of ar-Rāfiʿī's (d. 623/1226) dictionary *Šarḥ al-Wajīz*, which as a commentary on al-Ġazzālī's *al-Wajīz* contained a large number of technical terms from *fiqh* (Naṣṣār 1988:I, 55; ʾAḥmad n.d.:129–130). Similarly, the *Muxtār aṣ-Ṣiḥāḥ* was published by al-Jawharī (d. 396/1005) at the end of the 13th century C.E. as a revision of ar-Rāfiʿī's dictionary. The *Muxtār al-Qāmūs al-muḥīṭ* by az-Zāwī is an abridged version of al-Fayrūzābādī's work, and *Quṭr al-Muḥīṭ* by al-Bustānī is an abridged version of al-Bustānī's original work.

Al-Wāfī or *Fākihat al-bustān* is a smaller edition of al-Bustānī's original work, first published in 1930. *Rāʾid aṭ-ṭālib* is an abridged edition of Jibrān's *ar-Rāʾid*, and many editions cite *al-Munjid* as their primary source, including *al-Munjid al-ʾabjadī* by Fūʾād ʾAfrām Bustānī, and *Munjid aṭ-ṭullāb* and *al-Munjid al-ʾiʿdādī*. Some authors regard these abridged versions as an independent phase in Arabic lexicography, as they appear to be simply repetitions and revisions of previously published works.

4. ANALYTICAL REMARKS

Now that the history of Arabic lexicography and the development of modern Arabic/Arabic dictionaries have been reviewed, a few critical remarks are in order. First, it is important to observe the similarity between medieval and modern Arabic/Arabic dictionaries in their strategy for defining words. Generally, dictionaries from neither the medieval nor from the modern age have been particularly effective at conveying word definitions, due to the method known as the explanation technique. This methodology relies upon roundabout descriptions in order to provide definitions for entries; *al-Munjid*, *al-Wasīṭ*, and *ar-Rāʾid* are three examples of works which employ this technique.

A good example of the explanation technique can be seen in the perfect tense verb *faraġa* 'to be finished with, etc.'. Among the modern Arabic/Arabic dictionaries, 14 different synonyms are provided for this verb: *xalaṣa* 'to be matriculated', *ʾatamma* 'to complete', *qaṣada* 'to go toward', *intahā* 'to reach', *ʾaqbala* 'to come', *māta* 'to die', *taxallā* 'to give up', *nafaḏa* 'to go through', *ittasaʿa* 'to become wide', *ʾasraʿa* 'to hurry', *ḏahaba* 'to go away', *ḏahaba hadran* 'to go in vain', *inṣabba* 'to be poured', and *kanaḏa qalaq* 'to be anxious'. It should be noted that not all modern Arabic/Arabic dictionaries contain each of these synonyms that serve as definitions (Muḥammad 2001:60). *Al-ʾAsās* mentions four of these definitions; *al-Wasīṭ* and *al-Munjid* both include five; *al-Jadīd* contains eight; and *al-Ḥadīṯ* employs nine of these meanings. Thus, we can see that Arabic lexicographers are not always consistent and exhaustive when assigning definitions to entries.

Another aspect of modern Arabic/Arabic dictionaries worthy of further discussion is their use of illustrative examples in supporting definitions. The authors of most modern Arabic/Arabic dictionaries make a clear distinction between the central meaning of a word and its peripheral meanings within the framework of conceptual semantics. The use of illustrative examples is often much more effective than a classical definition in conveying meaning, especially when peripheral or contextual meanings are involved. For example, the verbal noun *ḍarb* 'striking' can have a number of different meanings, depending on the context: 'shape', 'multiplication [in mathematics]', 'coverage', 'extension', or 'prosodic form at the end of the second half of a line of poetry'. The use of such examples can help to clarify and explain these alternate definitions (ʿUmar 1998:120).

An additional method used to support the meaning of entries of modern Arabic/Arabic dictionaries is the use of *šawāhid* 'cited literary examples', usually derived from Qurʾānic verses, texts from other prophetic traditions, traditional poetry, and proverbs and other common expressions. These cited examples serve to confirm the meaning(s) of words through their contextual connotation as well as their functional usage. These examples reinforce the definition, while manifesting common word usage.

Voltaire would argue that literary citations used to support meaning in dictionaries add very little and simply illustrate stale associations that display neither the rich rhetoric not the genuine value of a language. Voltaire's position, however, may not apply to Arabic. This is due to the fact that the many abstract meanings of Arabic words are well-served by support from literary examples, providing the user with meaningful background contexts. Thus, quotations from literature, whether drawn from medieval or modern texts, can greatly enhance a word definition (al-Jilālī 1999:205).

In an attempt to verify the infrequency with which cited examples are used in modern Arabic/Arabic dictionaries, a brief statistical study has been conducted to demonstrate how four dictionaries measure up with respect to this methodology. In an examination of the entries of the letters *B* and *Y* that contained quotations, the following numbers were found: *al-Wasīṭ* contained 161 examples; *al-Ḥadīṯ* 197; *al-Jadīd* 237; and *al-ʾAsāsī* 217 – out of a total of 812 cited examples.

Another notable feature in modern Arabic/Arabic dictionaries is the use of symbols and abbreviations. Modern Arabic dictionaries are not consistent in the number and use of symbols and abbreviations, but most employ them to some degree. For instance, *al-Munjid* uses 35 abbreviations, 26 of them referring to the classification of plants, animals, mathematics, astronomy, and fine arts, as well as morphological and grammatical terms. *Al-Muḥīṭ* contains 25 abbreviations; *al-Wasīṭ* includes 20; the *Muʿjam al-ʿarabī al-ʾasāsī* employs 13 abbreviations; and the *New lexicon for students* contains 12 (al-Jilālī 1999:264–271). In comparison with the French dictionary *Petit Robert*, which includes more than four hundred abbreviations and signs, Arabic lexicons employ relatively few of these instruments (al-Jilālī 1999:87).

The nature of the content of modern dictionaries is another issue to consider. Although modern Arabic/Arabic dictionaries are relatively faithful to medieval Arabic dictionaries in both comprehensiveness and encyclopedic nature, the modern dictionaries are not completely devoted to *matn al-luġa* 'the basics of pure language'. Most modern Arabic dictionaries include information that is not directly relevant to the language.

A major issue to address is the internal organization of entries in modern Arabic/Arabic dictionaries. Contrary to the tradition of modern dictionaries in other languages, such as the *Oxford English dictionary*, definitions in Arabic dictionaries are generally not arranged according to frequency of use. Similarly, most modern Arabic/Arabic dictionaries do not trace the meaning of words through their etymological development. An attempt was made by the German scholar August Fischer to develop an etymological dictionary of Arabic as part of a larger project, but this work was never completed. Arabic scholars look forward to the day when such a project is realized.

Despite the shortcomings of modern Arabic dictionaries, the efforts of Arabic lexicographers from both the medieval and modern eras have yielded many positive results. Generally, the most important definitions for entries precede peripheral ones (i.e. general, common definitions come before specialized ones, concrete definitions come before abstract ones, etc.). This shows that Arabic lexicographers have carefully considered their undertaking and devoted a great deal of planning to their works.

In organizing their entries, medieval Arabic dictionaries employ three primary methods. The first is organization by rhyme, whereby words are arranged on the basis of the final letter, be it consonant or vowel. This technique can be traced back to al-Jawharī (d. 396/1005). The second is organization by alphabetical order, and the third is by the first radical consonant. The latter two techniques dominate in modern Arabic/Arabic dictionaries, as organization by rhyming ultimately fell out of favor with both lexicographers and users.

Arabic dictionaries demonstrate a wide variety of titles, some more modest than others. Some of the more unassuming titles include *al-Muʿjam al-kabīr*, *al-Muʿjam al-wasīṭ*, and *al-Muʿjam al-wajīz*, all sponsored by the Arabic Language Academy in Cairo. These titles illustrate the care and effort with which they were prepared, in contrast with the ambitious titles of some commercially produced dictionaries. The latter bear titles such as *al-Mufīd* 'the useful', *al-Muʿtamad* 'the most authoritative or authentic', *al-Munjid* 'the safe', *al-Marjiʿ* 'the source code', and *al-Fayṣal* 'the last word (or decisive)'. Some dictionaries in each group are specialized, targeting specific users. These are typically designated as school lexicons or as lexicons for students.

In contrast to the effort and the quality of research that has gone into the development of modern Arabic/Arabic dictionaries, the printing quality of most of these works is mediocre, if not altogether poor. They are often large sized and contain an immense number of pages, making them heavy and unwieldy. Cheap paper, haphazard typesetting, and inordinately small fonts end up detracting from the overall user experience (Maʿtūq 1999:120).

5. CONCLUSION

In tracing the history of Arabic lexicography, it is evident that many advances have been made in the development of Arabic dictionaries. At the same time, many of the techniques (beneficial and otherwise) used in the medieval dictionaries have carried over to the modern ones. Due to the shortcomings still present in modern Arabic dictionaries, many Arab scholars look forward to the development of an advanced Arabic lexicon that will avoid the errors of the past and better respond to the needs of all users. Prospective developers of such a work will have to answer many questions: What types of items should be included? What manner of organization should be utilized? What techniques should be used in developing the entries, including the ordering of the definitions? Another consideration central to the development of a superior Arabic dictionary will be the development of a consistent and thorough technique for giving the meanings of words. In pursuing such an endeavor, compilers could learn much from the experience of European lexicographers and linguists. The ideal modern Arabic dictionary would simultaneously do two things: it would avoid the mistakes and shortcomings of previous dictionaries, and it would respond to the needs of modern users. The ideal dictionary would be clear, concise, user friendly, comprehensive, and manageable in terms of both size and price.

BIBLIOGRAPHICAL REFERENCES

ʾAḥmad, ʿAbd as-Samīʿ Muḥammad. n.d. *al-Maʿājim al-ʿarabiyya*. Cairo: Dār al-Fikr al-ʿArabī.

ʾAnīs, ʾIbrāhīm. 1958. *Dilālat al-ʾalfāḏ̣*. Cairo: al-Maktaba al-ʾAnglū al-Miṣriyya.

ʿAṭṭār, Buštī. 1990. *al-Maʿājim al-ʿarabiyya: Ruʾyā tārīxiyya wa-taqwīmiyya*. El Jadida, Morocco: Université Chouaïb Doukkali.

Bacharach, Jere L. 1960. *A Near Eastern studies handbook, 570–1974*. Seattle: University of Washington Press.

Bahnasāwī, Ḥusām al-Bahiyy ʿAlī al-. 1990. *Fī l-maktaba al-luġawiyya*. Cairo: Kulliyyat ad-Dirāsāt al-ʿArabiyya al-ʾIslāmiyya.

Daqqāq, ʿUmar. 1977. *Maṣādir at-turāṯ al-ʿarabī fī l-luġa wa-l-maʿājim wa-l-ʾadab wa-t-tarājim*. Aleppo: Manšūrāt Jāmiʿa Ḥalab.

Darwīš, ʿAbdallāh. 1956. *al-Maʿājim al-ʿarabiyya*. Cairo: al-Maktaba al-ʾAnglū al-Miṣriyya.

Ġālī, Wajdī Rizq. 1993. *Muʿjam al-muʿjamāt al-ʿarabiyya*. Beirut: Maktaba Lubnān.

Hassanein, Ahmed Taher [ʾAḥmad Ṭāhir Ḥasanayn]. 1977. *Linguistic criteria in medieval Arabic literary criticism*. Ph.D. diss., Princeton University.

——. 1979. "Naḥwa qirāʾatin naḥwiyyatin muyassara". *Majalla Majmaʿ al-Luġa al-ʿArabiyya* 44. (Repr., Hassanein 1987:319–339.)

——. 1983. *Dawr aš-Šāmiyyīn al-muhājirīn ʾilā Miṣr fi-n-nahḍa al-ʾadabiyya al-ḥadīṯa*. Damascus: Dār al-Waṯba.

——. 1987. *Naḏ̣ariyyat al-iktimāl al-luġawī ʿinda l-ʿArab*. Cairo: Dār Ḥajar.

Jilālī, Ḥulām al-. 1999. *Taqniyat at-taʿrīf fī l-maʿājim al-ʿarabiyya al-muʿāṣira*. Algiers: Manšūrāt Ittiḥād al-Kuttāb al-ʿArab.

Maʿtūq, ʾAḥmad Muḥammad al-. 1999. *al-Maʿājim al-luġawiyya al-ʿarabiyya*. Abu Dhabi: al-Mujammaʿ aṯ-Ṯaqāfī.

Muḥammad, Muḥammad Saʿd. 2001. *al-Maʿājim: Dirāsa luġawiyya taṭbīqiyya*. Alexandria: ʿĀlam Nūr li-ṭ-Ṭibāʿa.
Nāṣif, Muṣṭafā. 1981. *Naḏariyyat al-maʿnā fī n-naqd al-ʿarabī*. Beirut: Dār al-ʾAndalus.
Naṣṣār, Ḥusayn. 1988. *al-Muʿjam al-ʿarabī*. Cairo: Dār Miṣr li-ṭ-Ṭibāʿa.
ʿUmar, ʾAḥmad Muxtār. 1998. *Fann ṣināʿat al-muʿjam al-ḥadīṯ*. Cairo: ʿĀlam al-Kutub.

Ahmed Taher Hassanein
(Misr University for Sciences and Technology)

Lexicon: Matrix and Etymon Model

1. Introduction

The first comments by Arab grammarians on the relationship between sound and meaning (→ *lafḏ*; → *maʿnā*) were related to the imitation of the sounds made by insects and birds. Their observations correspond to the theory that onomatopoeia is at the origin of human language. The idea that the shape of a word is the copy of the *perceptible* form, and thus a way of grasping the *intelligible*, is an idea that appears with considerable insistence in Ibn Jinnī's writings. He was one of the grammarians who paid the most attention to matters pertaining to the philosophy of language. This grammarian accepted that "at the origin of all languages, one finds the sounds which can be heard: the roar of the wind, the rumbling of the thunder, the murmur of water, the whinnying of the horse, and the bark of the deer" (*Xaṣāʾiṣ* I, 47). For Ibn Jinnī, as well as for his predecessors and followers, sounds can be either 'light' or 'heavy', which, all things considered, means that the meanings these sounds generate by their qualities do not greatly vary. The same may be said of the common meanings he proposes for certain triconsonantal roots in the framework of what he calls 'the great derivation' (→ *al-ištiqāq al-ʾakbar*). Taking as his point of departure the 'small derivation' (*al-ištiqāq al-ʾaṣġar*), in which the meanings are grouped around what was called the → root, he posits a general meaning linked not to the root of the word but rather to the radical consonants, independently of their syntagmatic linearity. The consonants *s-l-m*, he informs us, are capable of expressing the idea of 'gentleness' in all six of the possible combinations or rather in all those combinations that can form part of the language. The consonants *k-l-m* express the idea of 'force', while *q-w-l* expresses the idea of 'haste' and 'lightness' (Ibn Jinnī, *Xaṣāʾiṣ* I, 133–139). Versteegh (1984) speaks of this as a kind of 'curiosity'. In his view, one is confronted by a theory that transgresses the two postulates underlying contemporary linguistics: the arbitrary nature of the sign (or more precisely of its signification) and its linear nature, the evidence of which is accepted without hesitation.

Despite the *a priori* character of his theory and in spite of certain excesses in the details of his etymologies, Ibn Jinnī may be considered the precursor of modern research on the role of → metathesis in the Semitic languages. Although his ideas did not rally all traditional Arab grammarians, they were nonetheless taken up by several scholars, including Faxr ad-Dīn ar-Rāzī. In his commentary on the *Qurʾān, Mafātīḥ al-ġayb*, Faxr ad-Dīn dedicated a judicious analysis to the principle of the 'great derivation'. Ibn Jinnī should be credited above all for having seen, or rather grasped, a certain linguistic mechanism that could describe both the primary and the secondary forms of the Arabic lexicon within the process of a dual motivation.

2. The structure of the Arabic lexicon

The organization of the Arabic lexicon is one of the oldest topics in the study of this language. In order to analyze the internal structure of the lexicon, it is not enough to draw up a list of the lexical items. Rather, it is important to set as a goal the explanation of the relationships among the words, such as the phenomena of synonymy, polysemy, homonymy, and antithetical polysemy. For a long time, specialists have agreed that the minimal unit of the lexicon for Arabic and other Semitic languages is a compound of consonants. However, they part company when it comes to the question of whether the compound is bi- or triconsonantal. The majority favor the triconsonantal approach: "The predominance of roots with three consonants in Semitic is extremely clear" (D. Cohen 1974–1975:269). As early as 1817, Gesenius had noticed (and he was not the first to do so) that "this regular uniformity of the radicals in the Hebrew language consists...of the well-known fact that, except for a relatively small number of exceptions, all of the root words (at least the nouns and verbs) are made up of three root consonants" (Gesenius 1817:181–186). He adds in the next paragraph, though: "Nevertheless, no matter how striking this uniformity might be, there exist nonetheless numerous phenomena that leave no doubt that this [uniformity] was not so widespread from the start, but rather it was established only a bit later". One of these phenomena is the fact that "in a large number of roots words, which admittedly have in the present state of the language three radical consonants, it is clear that there were but two that were original and essential, to which the third one was then added or created by duplicating the second, or by attaching one of the movable vocal letters (at the beginning, middle, or end) or by adding at the beginning the semivowel *nūn*".

During the first half of the 20th century, authors such as Brockelmann (1908, 1910), M. Cohen (1947), Fleisch (1947), and Cantineau (1950) denied the importance of this binary conception. They advocated an organization based on the triconsonantal root, which imposed itself as an axiom among many of those scholars working on Semitic languages. Several studies, of which the best known is that of Ehret (1989), tried nonetheless to develop a biconsonantal conception of the lexicon. Unfortunately, in Ehret's theory the two primitive consonants of the biradical root are the first two, so that his results are manifestly different from those presented here. (For a treatment of the Arabic lexicon within the framework of standard theories on the structure of the root, → biradicalism.)

3. THE THEORY OF MATRICES AND ETYMA

The demonstration that the triconsonantal root does not take into account the immediately accessible relationships between words at the semantic and phonetic levels is at the heart of the reflection which led to the elaboration of the matrix and etymon theory (Bohas 1997, 2000).

Take, for example, Paradigm 1 (here and elsewhere, unless cited otherwise, definitions are taken from Lane).

Paradigm 1. The etymon {**b, t**} 'cutting'

_**b**a**tt**a_	'he cut it off, severed it'
**b**a**t**ara	'he cut, cut off' 'he cut off his tail'
in**b**a**t**aʾa	'he was/became cut off' (definition borrowed from Ibn Manẓūr, *Lisān al-ʿArab*)
**b**a**t**aka	'he cut it, severed it' 'he plucked it out'
**b**a**t**ala	'he cut it off, severed it' 'he separated it'
**b**ala**t**a	'he cut it off' (*Lisān*)
_al-**b**ur**t**u/al-**b**ar**t**u_	'the axe' (*Lisān*)
_sa**b**a**t**a_	'he cut the thing, cut the thing off' 'he shaved off his hair, shaved his head'

It is clear that all these words include a *b* and a *t* (in bold type in the paradigm) and they refer to the same idea of 'cutting'. This basis *bt* will be called an etymon and will be marked as follows: ∈{*b, t*}. The verbs in Paradigm 1 can be derived from the etymon ∈{*b, t*} by duplicating the last consonant – as in the case of ***batta*** – or by adding a final (e.g. ***batara***), medial (e.g. ***balata***), or initial (e.g. ***sabata***) consonant, without affecting the 'common primordial signification' (Brockelmann 1910). The etymon allows us to see that all of these words have something in common in phonetic terms, i.e. the phonemes *b, t*, as well as in semantic terms, i.e. the meaning of 'to cut'. Obviously, a dictionary organized in triconsonantal roots cannot take into account such observations.

Additional data are given in Paradigm 2.

Paradigm 2. The matrix {[+labial] [+coronal]} 'to deliver a blow'

badaḥa	'he cut off, split' (definition taken from Fīrūzābādī, *al-Qāmūs al-muḥīṭ*)
baḏaḥa	'he split the camel's tongue' (Ibn Manẓūr, *Lisān*)
bazzun	'sword'
bazala	'he cleft it, split it, slit it'
baḍaʿa	'he cut it, cut a piece off from it, he cut into pieces'
baṭṭa	'he slit a wound or an ulcer'

baṭara	'he cut it, divided it, slit it, split it'
tabba	'he cut, cut off'
ḥaḏafa	'he cut it off'
ḏubāb	'point or extremity of the sword'
ʿaḍaba	'he cut, cut off'
ʿaḍiba	'she had her ear slit'
hadaba	'he cut it, cut it off'
faʾsun	'an axe'
fattun	'a fissure in a rock'
faraṯa	'he ripped the basket (*julla*) and then scattered its contents'
farasa	'he [a lion] bruised or crushed and broke it'
faraṣa	'he cut it, slit it, divided it lengthwise'
faraḍa	'he notched it, made a notch or an incision in it [a piece of wood]'
fasaʾa	'he slit [the garment]' (*Qāmūs*)
faṣama	'he separated, divided'
faṣṣa	'he separated something from something'
faṭara	'he cleft, split, slit, rent, cracked it'
faṭama	'he cut it, severed it'
sayfun	'a sword'
šafratun	'a large knife'
ṣafiḥatun	'a sword'

Semantically, these words have something in common. They all revolve around the notional invariant 'to cut', although it is obvious that one cannot reduce them to one common root or etymon. At this point it is important to take up another level of analysis, not by focusing on the phonemes but rather by considering the features composing them. Note that each of the words in the paradigm contains a *b* or an *f*, which constitutes a class characterized by the feature [+labial], and in addition contains one of the elements of the category: *t, d, ḏ, ṯ, ḍ, ṭ, s, z, š, ṣ*. This class is defined by the feature [+coronal], which characterizes all sounds produced by the constriction formed by the front of the tongue and the area between the upper incisors and the hard palate (dental, alveolar, palato-alveolar, etc.). The common phonetic properties of the terms in Paradigm 2 can thus be expressed in the form of a formal invariant, [+labial] and [+coronal]. It is this combination of a formal invariant linked to a notional invariant that constitutes the matrix. In Paradigm 2, the formal invariant is made up by the two phonetic vectors [+labial] and [+coronal], and the notional invariant by the idea 'to deliver a blow' with the added meaning of using an instrument, 'with a cutting object'. In the context of this theory, the matrices constitute the minimal unit for establishing meaning. They allow us to group together all the terms on the basis of their common phonetic and semantic properties, which the root obviously cannot do.

This formal and conceptual organization of the lexicon is not only based on the identification of the lexical link 'among the words' (simultaneous existence of a formal invariant and a conceptual similarity) but also on the existing relationship 'between the words and the world'. This connection is of a mimophonic (vocomimetic) kind, i.e., there is an analogy between the phonetic material of the matrix, its notional invariant, and its referent. The physiological foundations of this analogy are of three kinds: "acoustic, when sounds reproduce a noise; kinetic, when an articulation reproduces a movement; visual, when the facial appearance (lips, cheeks) is modified, something which involves kinetic elements" (Guiraud 1967:125). This relationship is particularly manifest in Paradigm 3.

Paradigm 3. The matrix {[+pharyngeal] [+labial]} 'constriction'

∈{***b***, ***ṣ***}	
ṣabara	'he confined him, held him in custody, detained, retained, restrained, withheld him'
ʿaṣaba	'he bound it, tied it'
∈{***b***, ***ḍ***}	
ḍabba	'he clung to the ground'
ʾibaḍun	'the cord or rope with which the pastern of a camel's foreleg is tied'
∈{***b***, ***ṭ***}	
ṭunubun	'a tent rope'
rabaṭa	'he tied, bound, made fast'
∈{***b***, ***ḥ***}	
ḥabasa	'he confined, restricted, retained, arrested'
ḥabaka	'he bound it, tied it and made it fast or firm'
ḥablun	'a rope, a cord'
∈{***b***, ***x***}	
xabala	'he restrained, withheld' 'he prevented or hindered someone from doing something'

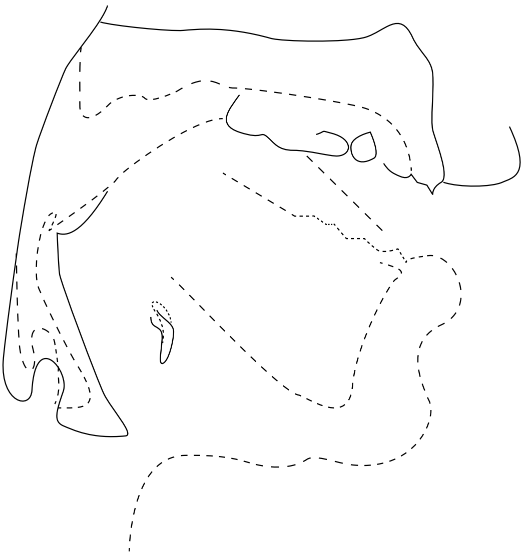
∈ {***b***, ʿ}	
ʿ*abala*	‘he hindered, prevented, impeded’
∈ {*f*, *ḍ*}	
ḍafrun	‘a camel’s girth’
ḍafana	‘to milk [a camel, cow, etc.] by hand’ (*Qāmūs*)
∈ {*f*, *ṭ*}	
ṭaffa	‘he bound all legs of the female camel’
∈ {*f*, *ḏ̣*}	
ḏ̣affa	‘he bound the legs of the camel’ (*Qāmūs*)
∈ {*f*, ʿ}	
ʿ*affa*	‘he abstained’
ʿ*afasa*	‘he restrained him’ (*Lisān*)

All terms of Paradigm 3 revolve around the notion of ‘constriction’. The etyma making it up are composed either of a labial (*b, f*) and a guttural (*x, h,* ʿ, *ḥ*) or a labial (*b, f*) and an emphatic (*ṣ, ṭ, ḍ, ḏ̣*). An analysis which only considers the phonemes cannot go any further. However, if one adopts the framework of the matrix and etymon model, one can go beyond this simple observation and detect the common phonetic denominator. Numerous studies have shown that the emphatics and the gutturals have the common feature [+pharyngeal], which refers to the constriction formed in the pharyngeal cavity, from the larynx to the uvula (Clements 1993:105). The matrix can thus be identified as: formal invariant: {[+labial], [+pharyngeal]}; notional invariant: ‘constriction’. All the terms in the paradigm are manifestations of this matrix. They include both constituent phonetic features and the notional invariant, which is the unifying element of all the words and refers directly or indirectly, metaphorically or metonymically, to the idea of ‘constriction’.

A mimophonic (vocomimetic) relationship between the notional invariant of ‘constriction’ and its developments on the one hand and an articulatory property of the pharyngeal is simple to demonstrate. One only has to consider Figure 1 taken from Ghazeli (1977:38) to realize that, for the articulation of pharyngeals, the pharynx is constricted. The relationship between the notional invariant of ‘constriction’ and the form of articulation then becomes clear.

A property which already appeared in the above paradigm (e.g. ***batta/tabba*** ‘to cut, cut off’) is the nonlinear ordering of the constituent parts of the matrix and the etyma descending from it. This property has been amply demonstrated for Arabic by Bohas (2000) and Bohas and Darfouf (1993), for Hebrew by Dat (2002), and for Moroccan dialects by Bohas (1997:135–138). Of all the constituent parts of the matrix and etymon model, the nonlinear ordering without a doubt poses the most problems for structuralist linguists, who are bound to point out that no other example of this phenomenon is known in the languages of the world. In fact, one must accept that, for the time being, no other linguistic group has been studied in this light. Although the search for other languages with this property should be pursued, there is always a theoretical possibility that the Semitic languages are the only ones with such a feature.

Figure 1. Articulation of a pharyngeal (Ghazeli 1977)



---- Shape of the lower pharynx during the articulation of the pharyngeal consonant [ʕ] in [ʕæll]

—— Shape of the pharynx before the initiation of the [ʕ] movement

Nonlinear ordering is a particularity which calls into question the linear principle of the linguistic sign, for each of its elements can be switched around freely without the semantic load being undermined. This fact expresses a characteristic of the Semitic lexicon, according to which meaning does not rely on the spatial configuration of the audible and/or graphic signs. The signification is based upon the constituent elements of the signifier irrespective of their layout (a+b: b+a) or their discontinuity (accretion of other elements, whatever their position in the chain, does not destroy the unity of the meaning). In terms of the lexemes, the reversibility within the biliteral units brings about → metathesis, conceptualized as a 'normal' phenomenon in Semitic languages (Lipiński 1997:192). In fact, if metathesis exists, it is because the etymon, by its very nature, transgresses the linear nature of the sign. The notion of nonlinear ordering is preferred here to that of metathesis, since it does not require one to set up an initial form (i.e. the one which undergoes the metathesis).

The general organization of the lexicon in the matrix and etymon model revolves around three levels of representation:

i. Matrix level: (μ) a combination, not arranged in linear fashion, of a pair of vectors of phonetic features, considered as a linguistic pre-sign or macro-sign and linked to a generic notion. It is at this level that the 'primordial meaning' is not linked to the sound or to the phoneme but rather to the phonetic feature which cannot be handled without the addition of supplementary phonetic material. At this level, the sounds appear as the translators of an articulation that evokes an object.
ii. Etymon level: (∈) the combination, not arranged in linear fashion, of phonemes containing these features and developing this generic notion. The etymon cannot be put on the same level as what is traditionally called the biconsonantal root; it is rather the element that is the basis of the multiconsonantal structures.
iii. Radical level: (R) etymon developed by the diffusion of the last consonant, or by infixation (word-initially, -medially, or -finally), and containing at least one vowel, recorded in the lexicon or provided by the morphological mechanisms of the language, and vectoring the notional invariant of the matrix and the etymon.

4. ADVANTAGES OF THE MATRIX AND ETYMON THEORY

Obviously, the model presented here is quite different from the organization of a lexicon by roots, i.e. from an organization in which the triconsonantal root is the primitive. The triconsonantal root is a hypothesis concerning the organization of the lexicon. Since this hypothesis cannot deal with a large number of regularities, the present model replaces it with a more explanatory model in terms of matrices and etyma. The latter takes into account the semantic and phonetic regularities existing between words (polysemy, homonymy, antithetical polysemy, etc.) that are not explained or even, in many cases, detected. In this context, one may resolve several lexical problems involving homonymy and polysemy (→ *muštarak*).

Homonymy may stem from two sources, the intersection of two etyma carrying two different semantic loads, or the actualization of two different matrices within the same etymon.

An example of homonymy by intersection is the verb *ġaraza*, which means both 'he inserted a needle into a thing; he stuck, fixed' and, more specifically, 'the locust stuck its tail into the ground to lay her eggs' and 'she [camel, sheep, or goat] had little milk; her milk became little'. In terms of the root, one cannot go further than this observation. With the help of the etyma, however, one observes that *ġazza* (∈{*ġ*, *z*}) means 'the female camel became scant of milk or deficient in milk (III)' and *razza* (∈{*r*, *z*}) means 'he pierced, stuck, or stabbed him and more particularly: the locust stuck her tail into the ground and laid her eggs'. This leads one to conclude that *ġaraza* results from the intersection of two etyma and this is why it combines meanings as different as 'the locust stuck her tail into the ground to lay her eggs' and 'deficiency of milk'.

Another example is that of the verb *xašafa*, which means both 'he entered into the thing' and 'he hastened, made haste'. In the context of the matrix and etymon model, one observes that *xašša* (∈{*x*, *š*}) means 'he entered into it' and *xaffa* (∈{*x*, *f*}) means 'he was light, active, prompt, and quick'. This leads one to conclude that *xašafa* results from the intersection of these two etyma and this is why it combines meanings as different as 'the entrance into the thing' and 'hurriedly, quickly'.

An example of homonymy stemming from an etymon which is the phonetic actualization of two different matrices is the word *ḍafrun*, which means (1) 'camel's girth', (2) 'a plait of hair', and (3) 'a great quantity of sand that has become heaped up on itself'. The *r* is an accretion, and the etymon ∈{***ḍ***, ***f***} is an actualization of the matrix mentioned above, {[labial], [pharyngeal]} notional invariant: 'constriction', while the second and third meanings are actualizations of the matrix {[labial], [dorsal]}, which is linked to the idea of 'curvature' viewed in diverse ways (Bohas 2000).

For antithetical polysemy (→ *ḍidd*), two types of solution can be found: the intersection of etyma of opposite meanings, and reverse interpretation of the matrix. The first case was already noted by Leguest (1858). The verb *šaʿaba* means both 'he collected, brought, gathered, drew together, united' and its opposite, 'he separated, put apart, divided, dispersed, scattered'. Within a matrix and etymon model, this can be explained by the intersection of the two etyma {***š***, ʿ} and {ʿ, ***b***}, the former being the carrier of the first meaning as it appears in its other actualizations, such as *šaʿita* 'the being separated or disunited' (*maṣdar* II and V), *šāʿa* 'he dispersed, scattered' (Ibn Manẓūr, *Lisān*), the latter being the carrier of the second meaning as it appears in its other actualizations, such as *ʿa**b**aʾa* 'he packed up goods or utensils, put goods or utensils one upon another', *waʿa**b**a* 'he collected, gathered together, congregated'.

An example of the reverse interpretation of the matrix is the term *jafjafun*, which means (1) 'raised ground', (2) 'hollow ground' (both Fīrūzābādī, *Qāmūs*). This term is a realization of the matrix {[labial], [dorsal]}, whose notional invariant is 'curvature', viewed from different perspectives. The fact that there are two contrary meanings can be explained unproblematically: in the first case the mimophonic meaning ∩ [convex] is apparent. In the second case, it is the mimophonic meaning ∪ [concave], with the same phonetic form. In this case it is possible to speak of antithetical polysemy of a polysemic type because there is a link between the two meanings, the notional invariant of the matrix. The same explanation goes for *jabʾun* ({***j***, ***b***}), which comes from the same matrix and means both 'a hill' and 'a hollow, a cavity in a mountain in which the water stagnates or collects'.

The elaboration of the theory of matrices was first based on the observation that the root was not able to explain such relationships. The matrix and etymon theory does not uphold the idea of a universal miming of form but rather the nonarbitrary character of the principle that directs the association of sound and meaning.

Without being absolutely universal, it is possible that the matrix level possesses a wider general nature. It is actualized in dialects so widely separated that one cannot posit a contact or a linguistic relationship (a large number of what may be called 'kinship terms', for instance, may be nothing more than the result of separate geneses to be explained by mimophony and not by genetic links). The perusal and systematic analysis of relative data allows one to extract a coherent system which assigns a notional value to the combinations of phonetic features. These combinations, linked to notional values, allow one to hypothesize both phonic and notional structuring characterizing the Semitic languages, even though some lexical zones may still appear opaque today. It all comes down to identifying a foundation process, one which has traversed the entire language and governs the structuring of the Semitic lexicon.

5. Objections against the theory

Various objections have been leveled against the model presented here. Some critics (e.g. Weipert 2000) have objected that the search for minimal units in the lexicon is useless, but clearly, if a particular theory enables us to discover and explain more semantic relationships between words than other theories, it constitutes a contribution to a better understanding of the linguistic structure of the lexicon.

It could be asked to what extent this theory is falsifiable. It seems that part of the theory, its very basis, is not falsifiable in the way this word is normally understood, 'to prove that a proposition is not founded'. Indeed, the principle that in Arabic meaning revolves around the matrices of the features, as assumed here, rather than around the phonemes, is not a theory but rather a 'discovery', in the sense that it provides a framework in which the data can be interpreted from a new angle. The same can be said in phonology: the axiom that the phoneme is not the fundamental, indivisible element of language is not a theory. That it can be broken down into features is a discovery. Admittedly, this discovery can be formulated in diverse ways. There are several theories on features, and one could always come up with others, in terms of elements rather than features, viewing the acoustic aspect rather than the articulatory one, etc. But the axiom that the phoneme can be broken down is not refutable.

There has also been an objection that the theory leaves too much freedom for the reconstruction of semantic relationships. For instance, the concept of 'family resemblances' appears to lend itself to all kinds of interpretations, for example linking up 'to walk, to run' with 'to divide up, to split off', which would be the consequence of 'to strike'. All work using the notion of family resemblances comes up sooner or later against this remark. Nevertheless, there are ways to limit the effects of arbitrary comparisons, e.g. by referring to parallel developments in other languages. The example mentioned above may be compared to similar developments in other languages. Indeed, it is precisely an area where the derivation of French, as Bohas has underlined, parallels closely that of Arabic and other Semitic languages. He remarks that "if one looks at the article **partir** in the dictionary of Bloch and von Wartburg (1932), one will note that it comes from the popular Latin, *partire*, itself derived from the Classical Latin, *partiri* and which "first meant 'to share', its everyday meaning until the XVIth century and [it] was preserved in the idiomatic *'avoir maille à'*" (Bohas 2002:101). In both cases, it is clear that 'to cut' is the concrete point of departure from which 'to leave; to set into motion' is derived. This is admittedly not the case in all languages, but for these two well-known languages, it is. This derivation is accepted by all in French, judged to be perfectly normal, banal, and expected, and no one thinks of treating Bloch and von Wartburg's (1932) approach as arbitrary.

In the context of the formal and notional fields organized around a matrix, lexical forms may have a common origin even if, for the time being, no element allows us to affirm this with certainty. Similarly, it is likely that numerous borrowed words are included in the matrix paradigm, whereas others, initially independent, would have been linked by paronymic attraction to the lexical grouping covered by the matrix in question. The morpho-phono-semantic structuring of the matrix generates this process and differentiates it from traditional etymological research.

The matrix and etymon model has been criticized, at least at the outset, for basing itself essentially on translated data. It is true that in the absence of a veritable etymological dictionary, it is sometimes hard to succeed in detecting the concrete, primitive meaning of a term, the etymon of a lexical form in the traditional sense of the word. The development of a theory should allow us to refine the lexical analyses progressively.

It should be emphasized that the theory sketched out here is not a return to the old biconsonantal theories. Its originality in theoretical terms stems from the fact that the matrix and etymon model shifts the discussion from the level of tangible structures to the more abstract level of phonology: the semantics of a given lexical form is supported not by the two consonantal elements, i.e. the constituent parts of the given lexical structure, but rather by a certain number of their constituent phonetic features.

It is clear that the matrix and etymon model is still under development. However, the empirical results obtained up to this point must be taken into account. The consequences stemming from it are not negligible, for the phenomenon touches upon the whole body of the lexicon. One should recognize that an analysis which takes into account the notion of matrix features is indeed possible and explanatory.

Bibliographical references

Primary sources

Fīrūzābādī, *Qāmūs* = Majd ad-Dīn Muḥammad ibn Yaʿqūb al-Fīrūzābādī aš-Šīrāzī, *al-Qāmūs al-muḥīṭ*. 4 vols. Beirut: Dār al-ʿIlm li-l-Jamīʿ.

Ibn Jinnī, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. Ed. Muḥammad ʿAlī an-Najjār. 3 vols. Cairo: Dār al-Kutub al-Miṣriyya, 1952–1956.

Ibn Manẓūr, *Lisān* = ʾAbū l-Faḍl Jamāl ad-Dīn Mukarram ibn Mukarram Ibn Manẓūr, *Lisān al-ʿArab*. 20 vols. Cairo: Dār al-Maʿārif.

Rāzī, *Mafātīḥ* = Faxr ad-Dīn ʾAbū ʿAbdallāh Muḥammad ibn ʿUmar ibn al-Ḥusayn ar-Rāzī, *Mafātīḥ al-ġayb* or *Kitāb at-Tafsīr al-kabīr*. Ed. Muḥammad Muḥyī d-Dīn. Cairo, n.d.

Secondary sources

Bloch, Oscar and Walter von Wartburg. 1932. *Dictionnaire étymologique de la langue française*. Paris: Presses Universitaires de France.

Bohas, Georges. 1997. *Matrices, étymons, racines: Eléments d'une théorie lexicologique du vocabulaire arabe*. Paris and Louvain: Peeters.

——. 2000. *Matrices et étymons: Développements de la théorie*. Lausanne: Editions du Zèbre.

——. 2002. "Du concret à l'abstrait sur les deux rives de la Méditerranée". *Langues et Littératures du Monde Arabe* 3.85–105.

—— and Noreddine Darfouf. 1993. "Contribution à la réorganisation du lexique de l'arabe: Les étymons non-ordonnés". *Linguistica Communicatio* 5:1–2.55–103.

Brockelmann, Carl. 1908. *Grundriss der vergleichenden Grammatik der semitischen Sprachen*. I. *Laut- und Formenlehre*. Berlin: Reuther und Reichard.

——. 1910. *Précis de linguistique sémitique*. Transl. (from German) William Marçais and Marcel Cohen. Paris: Geuthner.

Cantineau, Jean. 1950. "Racines et schèmes". *Mélanges William Marçais,* 119–124. Paris: Maisonneuve.

Clements, Georges Nick. 1993. "Lieu d'articulation des consonnes et des voyelles: Une théorie unifiée". *L'architecture des représentations phonologiques*, ed. Bernard Laks and Annie Rialland, 101–145. Paris: Centre Nationale de la Recherche Scientifique.

Cohen, David. 1974–1975. "Sémitique comparé". *Annuaire de l'Ecole Pratique des Hautes Etudes*, IV^{ème} section, 265–275.

Cohen, Marcel. 1947. *Essai comparatif sur le vocabulaire et la phonétique du chamito-sémitique*. Paris: Champion.

Dat, Mihai. 2002. *Matrices et étymons: Mimophonie lexicale en hébreu biblique*. Thèse de Doctorat, Ecole Normale Supérieure Lettres et Sciences Humaines, Lyon.

Ehret, Christopher. 1989. "The origin of third consonants in Semitic roots: An internal reconstruction (applied to Arabic)". *Journal of Afroasiatic Languages* 2.107–202.

Fleisch, Henri. 1947. *Introduction à l'étude des langues sémitiques: Eléments de bibliographie*. Paris: Maisonneuve.

Gesenius, Wilhelm. 1817. *Lehrgebäude der hebräischen Sprache mit Vergleichung der verwandten Dialekte*. Leipzig: F.C.W. Vogel.

Ghazeli, Salem. 1977. *Back consonants and backing coarticulation in Arabic*. Ph.D. diss., University of Texas at Austin.

Guiraud, Pierre. 1967. *Structures étymologiques du français*. Paris: Payot. (2nd ed. 1986.)

Lane, Edward William. 1863. *Arabic-English lexicon*. 8 vols. London: Williams and Norgate.

Leguest, Charles. 1858. *Etude sur la formation des racines sémitiques*. Paris: B. Duprat.

Lipiński, Edward. 1997. *Semitic languages: Outline of a comparative grammar*. Leuven: Peeters.

Versteegh, Kees. 1984. "La grande étymologie d'Ibn Ginni". *La linguistique fantastique*, ed. Sylvain Auroux a.o., 44–50. Paris: Klims-Denoel.

Weipert, Reinhard. 2000. Review of: Bohas (1997), *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 150.656–660.

Georges Bohas (Ecole Normale Supérieure Lyon)
Mihai Dat (Dijon)

Libya

1. History and Arabization

Before the Arab invasion in the 7th century C.E., Libyans, Jews, Phoenicians, Greeks, Romans, Vandals, and Byzantines peopled those countries which make up today's Libya. The Libyans were ancestors of the Berbers and spoke Libyan, from which derived the rest of the Berber dialects. Libyco-Berber (along with Egyptian, Couchitic, Chadian, and Semitic) is one of the five branches of the Hamito-Semitic or → Afro-Asiatic linguistic tree (Larcher 2001:43).

The Phoenicians, having first founded the Punic town of Carthage, built a coastal empire in Tripolitania, which lasted more than six hundred years (from the 8th to the 2nd century B.C.E.). They spoke Punic, the first Semitic language in North Africa, which was spoken in Tripolitania well after the Roman conquest, until the 2nd century C.E. Meanwhile, the Greeks settled in eastern Libya, as early as the 7th century B.C.E., founding Cyrene in 631, which would later give its name to the eastern part of Libya (Cyrenaica). For more than one thousand years, eastern Libya was a hotbed of Hellenistic culture. The Romans destroyed Carthage in 146 B.C.E. and invaded Libya shortly thereafter, making their home there until the 5th century C.E. With the Romans came Latin, but, neither replacing Greek in Cyrenaica nor Punic in Tripolitania, it was nothing more than an administrative language. The Vandals, heretofore established in Spain, invaded northern Africa in 430 C.E. and claimed Carthage for their own in 439, and then moved on to Tripolitania. The Byzantines were in Libya from 535 until the Arab invasion. Both occupations were brief and had little cultural effect on preceding civilizations.

The Arabization of North Africa, related to the Muslim conquest coming from the east, took place in two waves: first in the 7th century, and then much later in the 11th. These successive Arab invasions were responsible for its Arabo-Muslim features. The first conquest did not lead to a profound Arabization of Libya, nor of other North African countries, which remained essentially Berberophone. It was not until the 11th century – when the Bedouin tribes of the Banū Hilāl and the Banū Sulaym subsequently settled and left a strong influence on the second wave – that North Africa was significantly Arabized. Upon their arrival there, the first Arab conquerors were confined for 150 years in those countries which make up Tunisia, some of Tripolitania, and the eastern part of today's Algeria.

As for Libya, in 642 ʿAmr ibn al-ʿĀṣ, then governor of Egypt, took over Cyrenaica and then Fezzan; in 643 came Tripolitania, followed by Tripoli in 645. Shortly thereafter, between 669 and 675, the ʿUqba ibn Nāfiʿ expedition allowed the introduction of Arabs and of Islam into Tripolitania. As a result, several Berber tribes converted to Islam. Yet, it was not until 696–697 that the Ḥassān al-Ġassāni military expedition guaranteed the definitive settlement of oriental Arabs in Libya. Meanwhile, Kairouan, the first Arab city in North Africa and the center of Arab activities in North Africa, was founded in 689–690.

This first conquest (7th century) was not characterized by a systematic peopling by the conquerors, who were relatively few and only occupied cities and strategic points. Conversely, the Berber populations remained untouched in the countryside and high up in the mountains until the middle of the 11th century (Mantran 1975:21). Pre-Hilālian Arabic thus became the Arabic that was spoken in these first occupied and Arabized areas by Arabs in the 7th century.

The Bedouin arrived in North Africa in the middle of the 11th century. In just two years, the Banū Hilāl, followed by the Banū Sulaym, both banished by the Fatimid caliph, came from the east and reached North Africa; these tribes are mainly responsible for North Africa's Arabization. The Banū Hilāl traveled all across Libya, settled in Tripolitania and Tunisia, and also went all the way to Morocco, via northern Algeria (between the high plateaus and the Mediterranean Sea). The Banū Sulaym followed the path of the Banū Hilāl with some stopping along the way in Cyrenaica, others in Tripolitania, in the south of Tunisia and the southeast of Algeria. Another tribe was that of the Maʿqil, whose members took a more southern route (via the northern Sahara) and, reaching southern Morocco (Tafilalt), one of their branches, the Banū Ḥassān, Arabized Mauritania and even gave the name to the Arabic spoken there, → Ḥassāniyya (Caubet 2000–2001:78).

As for Libya, these same tribes penetrated into Cyrenaica in 1050–1051. While the Banū Sulaym stayed in Cyrenaica, the Banū Hilāl continued their journey toward the west. In Tripolitania, the Banū Zuġba tribe (a branch of the Banū Hilāl) occupied the whole region extending from Tripoli to Gabes, whereas the Banū Dabbāb peopled the eastern part of Tripolitania. Since the 12th century, Arabization and Islamization made considerable progress: Arab peopling was more significant in the coastal area and in the interior part of Tripolitania, especially during the Hafsid period, from 1230 onward.

During this period, Cyrenaica, under Egyptian rule, was mostly made up of nomadic Arabs. Fezzan was, at this time, an independent state, with some prosperous towns, and a station for caravaneer commerce, where Muslim merchants established commercial relations between the south (Sudan and Central Africa) and the Mediterranean. Bedouin Arabic thus became the Arabic spoken in these occupied and Arabized areas by Arabs in the 11th century.

2. ARABIC OF LIBYA

With regard to the dialects, Libya can be divided into three areas (Owens 1983): a western area (which includes Tripolitania and Fezzan), an eastern area (Cyrenaica), and a transitional zone (which includes Sebha in the Fezzan, Misurata in Tripolitania, going until Cyrenaica, which also includes the Syrt region and the Jufra area, i.e. the region of the Sokna, Houn, and Waddan oases).

Libyan dialects belong to the Maghrebi group, characterized by the prefix *n-* for the 1st person singular and by the prefix *n-* with the suffix *-u* for the 1st person plural, in the imperfect form of the verb; the difference in 1st person verbal forms, in the imperfect, is considered to be the main discriminant between Maghrebi and Levantine dialects. Philippe Marçais (1977:IX) wrote: "La Libye se présente comme un ensemble relativement homogène. Elle est caractérisée par des traits bédouins marqués au coin d'un conservatisme assez remarquable. Certains des rares centres urbains qui s'y trouvent (Tripoli notamment) usent de parlers sédentaires, mais ils ont parfois subi une forte influence des parlers bédouins". However, concerning certain features, Eastern Libyan dialects are clearly opposed to Western Libyan dialects. As for the dialects of the transitional zone, they show features of these two zones at the same time, but they also have their own features. Owens (1984:242) states that Sebha shares many phonological features with the Eastern Libyan dialects, but with regard to the lexicon, its features are closer to the Western Libyan ones. As for the dialect of Misurata, it has many common features with the dialects of Tripolitania, but it also has its own specific features.

Indeed, as a whole, the Libyan Arabic dialects are of the Bedouin type. This can be observed in phonetics, morphology, syntax, and lexicon. In phonetics, **q* corresponds to a voiced plosive [g], a sound revealing the Bedouin origins of the dialect. Examples: *gəddīd* 'dried meat', *gəlˁāwi* 'yellow melon', *gərfa* 'cinnamon', *gəmla* 'louse', *lge* 'he found', *nāga* 'she-camel', *ryāga* 'saliva', *dgīga* 'minute', *ṭuggāša* 'spark', *mṣəggəˁ* 'cold', *nəggəz* 'he jumped', *nəggəl* 'he copied', *šrəg* 'he choked himself', *fəyyāg* 'alarm clock', *lāsəg* 'glued', *məštāg* 'nostalgic', *ḥəgg* 'price; truth', *dəgg* 'he knocked at the door'. The phonemes /ē/ and /ō/ represent the reduction of the diphthongs /ay/ and /aw/, respectively, cf. for /ē/ *žēb* (< **žayb*) 'pocket', *kēf* (< **kayf*) 'enjoyment', *ṣēf* (< **ṣayf*) 'summer', *dēl* (< **dayl*) 'tail', *šēn* (< **šayn*) 'ugly', *šēṭān* (< **šayṭān*) 'devil', *bēn* (< **bayn*) 'between', *wēn* (**w-ʾayna*) 'where', *xšēm* (< **xšaym*) 'kiss', *nwēma* (< **nwayma*) 'nap', *žwēw* (< **žwayw*) 'a quite good atmosphere', *ḍwēw* (< **ḍwayw*) 'a little bit of light', *wāldēn* (< **wāldayn*) 'parents', *wədnēn* (< **wədnayn*) '(two) ears', *bābēn* (< **bābayn*) 'two doors', *ktābēn* (< **ktābayn*) 'two books', *xəššēt* (< **xəššayt*) 'I entered'; and for /ō/: *bōsa* (< **bawsa*) 'kiss', *mōt* (< **mawt*) 'death', *yōm* (< **yawm*) 'day', *lōn* (< **lawn*) 'color', *ḥōš* (< **ḥawš*) 'house', *lōm* (< **lawm*) 'blame', *dōm* (< **dawm*) 'long time', *nōm* (< **nawm*) 'sleep', *xōf* (< **xawf*) 'fear', *ṣōt* (< **ṣawt*) 'voice', *zōz* (< **žawz*) 'two', *šōg* (< **šawg*) 'nostalgia', *lōḥ* (< **lawḥ*) 'wood', *dōg* (< **dawg*) 'taste', *lōza* (< **lawza*) 'almond', *šōka* (< **šawka*) 'thorn', *kōsa* (< **kawsa*) 'courgette, zucchini', *fōg* (< **fawg*) 'on, above'. The reduction of the diphthongs /ay/ and /aw/ to /ē/ and /ō/, respectively, is a characteristic to be found in certain other North African Bedouin dialects.

As to morphophonemics, when a suffix with an initial vowel is added to the 3rd person feminine singular of the perfect form of the verb, the final vowel of the verb ending is lengthened from *-ət* to *-āt*, e.g. *ḍəṛbət* 'she hit' + *-ək* 'you' > *ḍəṛbāt-ək* 'she hit you', *fəhmət* 'she understood' + *-əh* 'him' > *fəhmāt-əh* 'she understood him'. This feature is found in the Bedouin and rural Maghrebi Arabic dialects.

In the verbal morphology, in the 2nd person singular, there is a gender distinction for the independent personal pronoun (masc. *ánta* vs. fem. *ánti*), but also for the verbal inflexion (masc. *klēt* 'you ate' vs. fem. *klēti*; masc. *tžīb* 'you bring' vs. fem. *tžībi*; masc. *dīr* vs. fem. *dīri* 'do!'). For verbs III*y*, the vowels /ā/ and /ī/ of the basis are elided with inflectional suffixes, e.g. the 2nd person feminine singular *telgi* 'you [fem. sg.] find', in the plural, *nešru* 'we buy', *telgu* 'you [masc. pl.] find' and *yebdu* 'they [masc. pl.] begin'. The same holds for the perfect: for the 3rd person plural *šru* 'they [masc. pl.] bought', for the 3rd person feminine singular Libyan Arabic exhibits the form C_1C_2ət, e.g. *mšət* 'she went away'. There are no forms with retention of stem final /ā/, e.g. *telgāy* 'you [fem. sg.] find', *yebdāw* 'they [masc. pl.] begin', *šrāw* 'they [masc. pl.] bought', *mšāt* 'she went away' (featuring the long vowel *ā*) common in other North African dialects.

As to pronominal morphology, the 3rd person masculine singular personal pronoun suffix is *-a* with a pausal form *-əh*, and never *-u* as in the pre-Hilālian dialects, e.g. *əlxāl f-xēṛ lēn yukbuṛ l-a wuld uxt-əh* 'the uncle is fine until his nephew has grown up'; *-a* is lengthened when a suffix is added, e.g. *mā gāl lāš* 'he did not tell him'.

As to nominal morphology, diminutives of nouns based on a triliteral root are formed on the model $C_1C_2\bar{e}C_3$; note that this form is characteristic of Bedouin dialects as opposed to the model C_1C_2*əyyə*C_3 of sedentary dialects. Examples: *kəlb* 'dog' and *klēb* 'puppy'; *marṣā* 'port' and *mrēṣā* 'small port'; *gubba* 'dome' and *gbēba* 'small dome'; *ḥəlu* 'sweet' and *ḥlēw* 'a little bit sweet'; *bāb* 'door' and *bwēb* 'small door'; *ḥōš* 'house' and *ḥwēš* 'small house'; *ṛās* 'head' and *ṛwēs* 'small head'. Quadriliteral nouns of the model $C_1vC_2C3\bar{a}C_4$ (with a long vowel between the third and fourth radical consonants) form their diminutives on the model $C_1C_2\bar{e}C_3\bar{\imath}C_4$, e.g. *məftāḥ* 'key' and *mfētīḥ* 'small key'; *sərwāl* 'trousers' and *srēwīl* 'small trousers'; *šəbbāk* 'window grill' and *šbēbīk* 'small window grill'; *dəllāˁa* 'watermelon' and *dlēlīˁa* 'small watermelon'; *fərrūž* 'cock' and *frērīž* 'small cock'. Moreover, quadriliteral nouns of the $C_1vC_2C_3\bar{a}C_4$ pattern form plurals on the model $C_1C_2\bar{a}C_3\bar{\imath}C_4$, e.g. *meftāḥ* 'key' and *mfātīḥ* 'keys'; *məṣmāṛ* 'nail' and *mṣāmīr* 'nails'; *šəbbāk* 'window grill' and *šbābīk* 'window grills'; *bābūr* 'boat' and *bwābīr* 'boats'; *fərrūž* ‘cock’ and *frārīž* ‘cocks’; *səkkīn* ‘knife’ and *skākīn* ‘knifes’. This also brings Tripoli spoken Arabic closer to Bedouin dialects.

As to syntax, we notice that in Libyan Arabic dialects, the noun appears without any marker for indefiniteness, and unlike the urban dialects of Algeria and Morocco, they do not use the indefinite construction *wāḥəd əl* formed from the numeral *wāḥəd* ‘one’. Examples: *bīr* ‘a well’, *žṛāna* ‘a frog’, *kṛāsi* ‘chairs’ instead of *wāḥd əlbīr* etc., the latter being used only in a very limited way in Bedouin dialects.

The lexicon of the Libyan Arabic contains many items of Bedouin origin. Indeed, verbs such as *dār* ‘he did’, *xəšš* ‘he entered’, *ṭləˤ* ‘he went out’, *yəbbi* ‘he wants’, *šbəḥ* ‘he saw’, *nšəd* ‘he asked’, *ṛāža* ‘he waited’, *gəˤməz* ‘he sat’, *gˤəd* ‘he stayed’, *dfəˤ* ‘he paid’, *wəlla* ‘he came back’, *ṛkəb* ‘he went up’, *tṣubb əlmṭəṛ* ‘it’s raining’ are employed; the following nouns are also used: *mṭəṛ* ‘rain’, *dəḥya* ‘an egg’, *ḥōš* ‘house’, *bēt* ‘tent’, *yədd* ‘hand’, *ṣubḥ* ‘morning’, *rəžžāla* ‘men’; other words such as *humma* ‘they [masc. pl.]’, *āməs* ‘yesterday’ and increased forms such as *ġudwīka* ‘tomorrow’, *ġādīkāy(a)* ‘over there’, *hādūkāy(a)* ‘those’, that are found in other Bedouin dialects, are also employed in Libyan Arabic.

However, certain differences in phonetics, morphology, syntax, and lexicon oppose the various Libyan dialects. As to phonetics, the Eastern Libyan dialects (Benghazi, Tobrouk, and Kufra) display interdental fricatives /ṯ/, /ḏ/, and /ḏ̣/ (such as *ṯlāṯa* ‘three’, *ḏhəb* ‘gold’, *ḏ̣ull* ‘shade’). In the Western Libyan dialects (Tripoli, Darj, Misurata, Sorman, Garabulli, and Sebha), these merged with the corresponding dental stops /t/, /d/, and /ḍ/ (i.e. *tlāta, dhəb, ḍull*; see Owens 1983). Owens also mentions that in the Zawia dialect (west of Tripoli), some people use the interdentals and say *hāḏa* ‘this one’, and others replace them by the dental stops and say *hāda* ‘this one’. Philippe Marçais (2001) points out that, in the Fezzan, interdentals are found in the Bedouin dialects but not in the sedentary ones.

In final position, in Tripolitania (Pereira 2004) and in the Fezzan (Caubet 2004:73), → *ʾimāla* has an impact on the final /a/ (<**ā*) in pause, moving it to [e], e.g. *āne* ‘I’, *ḥne* ‘we’, *hne* ‘here’. Final /a/ of the perfect of verbs III*y* undergoes *ʾimāla* and is pronounced as [e], even if it follows a back or emphatic consonant, e.g. *že* ‘he came’, *kre* ‘he rented’, *ġle* ‘it became more expensive’, *kse* ‘he wore’, *bde* ‘he began’, *mše* ‘he went away’, *lge* ‘he found’, *sge* ‘he watered’, *kfe* ‘he covered’, *gṛe* ‘he studied’, *ngle* ‘it was fried’, *nṭfe* ‘it turned off’, *rtxe* ‘he relaxed’, *dwe* ‘he talked’, *stḥe* ‘he was ashamed’, *ġze* ‘he attacked’, *staġne* ‘he became rich’. The same holds true for nouns ending in /a/ (< **āʾ*): *me* ‘water’, *sme* ‘sky’, *ˤše* ‘dinner, *ġde* ‘lunch’, *nse* ‘women’, *ġṭe* ‘blanket’.

As to syllabic structure, short vowels in open unstressed syllables are found in the Eastern Libyan dialects, in the Fezzan, and in Misurata, whereas Tripolitanian dialects do not know this feature (e.g. *žibál* ‘mountain’, *muṭár* ‘rain’, *kitáb* ‘he wrote’, as opposed to *žbəl, mṭər, ktəb* in Tripoli).

As to morphology, unlike the Western Libyan dialects, Cyrenaica dialects and the dialects of the transitional zone (Fezzan, Misurata, and Jufra) make a gender distinction in the 2nd and the 3rd person plural of the independent personal pronouns (in the 2nd person masculine *antum* is distinguished from feminine *antən* ‘you’, and in the 3rd person masculine *hum* and *huṃṃa* ‘they’ are distinguished from feminine *hən* and *hənna*). For the pronominal suffixes, masculine *lēkum* ‘for you [pl.]’ is opposed to feminine *lēkən*, and masculine *mˤāhum* ‘with them [pl.]’ is opposed to feminine *mˤāhən*. The conjugation of the perfect of the verb shows masculine *ktabtu* ‘you [pl.] wrote’ vs. feminine *ktabtan*, and masculine *ktabu* ‘they [pl.] wrote’ vs. feminine *ktaban*; in the imperfect, masculine *tekt*^{e}*bu* ‘you [pl.] write’ vs. feminine *tekt*^{e}*ban*, and masculine *yekt*^{e}*bu* ‘they [pl.] write’ vs. feminine *yekt*^{e}*ban*; in the imperative, masculine *ekt*^{e}*bu* ‘write [pl.]!’ vs. feminine *ekt*^{e}*ban*); and in the demonstratives, *hādōl* ‘these [masc. pl.]’ vs. *hādēn* ‘these ones, [fem. pl.]’ and *hādōlōk* ‘those ones [masc. pl.]’ vs. *hādānāk* ‘those ones [fem. pl.]’.

With respect to syntax, in the Western Libyan dialects (Tripolitania and Fezzan), the preverbal marker *b-* is used to express the future of intention (i.e. *bənšūfək ġudwa* ‘I will see you tomorrow’) and the preverbal marker *ḥā-* is used to express a close/coming future (e.g. *ḥā-nastāḥəš lībya* ‘I am going to miss Libya’).

Lexical differences (see Owens 1983) include the verb *gəll* ‘to bring’, which is used in the Fezzan, whereas the verbs *žāb* and *rfaˤ* are used in Tripolitania and in Cyrenaica with the same meaning. In western Libya, *wgəf, yūgəf* is used for 'to stand', whereas *ṣabba* is used in the east. In Darj (in the extreme west of Libya, in Tripolitania) and also in Cyrenaica, *sannaṭ* means 'to hear', but in Kufra, in Tripolitania, and in the Fezzan the word *smaʕ/yasmaʕ* is used. In the Fezzan, *əlmṭər tži* means 'it's raining', but *əlmṭəṛ tṣubb* is used in the other parts of Libya. The verb *šāf* 'he saw' is used in the Eastern Libyan dialects, whereas *šbaḥ* is used in the Western Libyan ones. In Libya, *gaʕməz* means 'he sat', but Owens also mentions *yižlis* in Kufra (in Cyrenaica).

With respect to the nouns, the word *rušd* means 'onions' in the Fezzan, but the word *bṣəl* is used in the other parts of Libya. *ḥažra* means 'stone', but we also find *ḥēṭa* in the Eastern Libyan dialects and *ršāda* in the Western Libyan ones. In Cyrenaica, *ḥāfir* means 'nail', whereas the word used in Tripolitania and in the Fezzan for 'nail' is *ḍufᵘr*. In Tripolitania, the word used for 'teeth' is *sənn*, but in Darj the term employed is *ḍurs* (*ḍərs* in Tripoli means 'molar'). In Cyrenaica, *immayya* means 'water' (in Tripoli *ṃṃayya* and *ᵘṃṃwēya*, besides the less common *me*), but the word used in Kufra is *šurāb*. *suḥāb* is the word used in Benghazi for 'cloud' (*sḥāb* in Tripoli), but the word *mizin* is used in Kufra. With respect to the adverbs, *bukra* is the word used in Cyrenaica and in the Fezzan with the meaning of 'tomorrow', whereas *ġudwa* is the one used in the other parts of Libya, in Tripoli *ġədwīka* as well. *hālba* is the word used in Tripolitania with the meaning of 'a lot', but we also find *yāsər* in the Fezzan and in Darj and *wāžid* in Misurata and in Cyrenaica. Moreover, because of the Ottoman presence in Libya (from 1551 to 1911) and because of the Italian occupation (from 1911 to 1955), we find Turkish loanwords such as *ḅāya* (< *boya*) 'shoe polish', *kubri* (< *köprü*) 'bridge', *kāšīk* (< *kaşık*) 'spoon', *šīša* (< *şişe*) 'bottle', *kīsa* (< *kese*) 'bag', *šārīt* (< *şerit*) 'tape', and Iranian words such as *šīšma* (< *češme* چشمه) 'source; tap, faucet' which found their way to Libya via Turkish *çeşme* 'public fountain; old copper tap). → Italian loanwords in Libyan dialects include *bārkājo* (< *parcheggio*) 'parking', *bāku* (< *pacco*) 'pack', *bīro* 'pen' (< *biro*), *bəšklēṭa* 'bicycle' (< *bicicletta*), *maṛšābēdi* (< *marciapiedi*) 'pavement', *kūbārta* (< *coperta*) 'blanket', *jībōto* (< *giubotto*) 'lumberjacket, windbreaker', *kačafīti* (< *cacciavite*) 'screwdriver', and *simāforo* (< *semàforo*) 'traffic lights'. As a matter of course, many English words reflecting modern life have been adopted in recent times, e.g. *boṛṭābəl* (< *portable*) 'laptop', *mōbayl* (< *mobile*) 'mobile phone', *mīksər* (< *mixer*) 'mixer', *māsāj* (< *message*) 'message, SMS'. Other English words have been borrowed from English soldiers during the 1940s, such as *būfta* (< *poofter* 'homosexual'), *bōy* (< *boy* 'homosexual'), *sīlfər* (< *silver*) 'inox', *səyyəv* (< *save*) 'he saved'. *mākyāž* 'makeup' (< *maquillage*) is a French loan.

3. Other languages spoken in Libya

In Libya there are 14 percent of Berberophones. Berberophone islets are found today in Tripolitania (in Zwara, in the Djebel Nefoussa, in Yefren, in Nalout, in Sinaouen, and in Ghadames). Nowadays, there are no more Berberophones in Cyrenaica, but they were found there at the beginning of the 20th century in Awjila oasis and in Al-Jaġbūb. There may be some remaining Berberophones in the Joufra oasis, in Houn, in Waddân, and in Sokna (Larcher 2001:46–48; Souriau 1986:40). In the extreme southwest of Libya (in Ghat and Barakat area), and also in Oubari and in Morzouk in the Fezzan, Touaregs are still to be found; they are also found in Ghadames. There are about 17,000 Touaregs in Libya (Anthony 2002:29).

At least one Nilo-Saharan language, Toubou, is spoken in the southeast of Libya, in the Koufra area. There are about 2,600 Toubous in Libya (Ham 2002:30).

Some old persons can still speak Italian, English, and French in Libya. Libya was under Italian occupation from 1911 until 1955, British troops invaded Libya after the Second World War, and the United States had a dominant position in the oil industry in Libya; the French also occupied Libya in 1943 and stayed there until 1955.

4. Modern Standard Arabic in Libya

Modern Standard Arabic is the official language of Libya and of all members of the Arab League (to which Libya has belonged since 1953). It is the language of the *Qur'an* (in a more Classical register); of Modern Written Arabic (in use in schools, literature, the press, and correspondence, and on signs and billboards); and of spoken monologue (on television, radio, and in public speeches and conferences). Colloquial Arabic is used in the arena of spontaneous oral communication and in everyday conversation.

The current government of Libya adopted a significant policy of Arabization in the early 1970s intended to fortify the country's Arab-Islamic identity. Learning the *Qurʾan* (reading and memorizing it) is obligatory in Libya. Moreover, systematic translation to Standard Arabic was encouraged in all domains of public life. The Latin alphabet was abandoned to make way for the exclusive use of the Arabic script; thus, road signs in Libya are written only in Arabic, and foreign visa applicants must translate their passports into Arabic. These policies contrast with what can be observed in other Arab countries, where languages such as French or English are used in addition to Arabic (in civic, administrative, and educational settings). Moreover, in 1984, and for about a decade, foreign languages were no longer taught in Libya.

In line with these policies the Libyan government introduced in the late 1980s a new set of names for the months of the year, which are used by the Jamahiriyya News Agency (JANA) and the official press: *ʾayannār, ʾan-nuwwār, ar-rabīʿ, aṭ-ṭayr, al-māʾ, aṣ-ṣayf, nāṣir, hānībāl, al-fātiḥ, at-tumūr, al-ḥarṯ, kānūn* (cf. Azema 2000:20).

Libya has made → literacy another national priority. A long-term literacy plan was in effect between 1972 and 1992. In 1973, approximately 32 percent of men and 73 percent of women (over ten years old) could not read. Today, Libya is home to North Africa's most literate population: in 1995, 63 percent of women and 87.9 percent of men – a total of 76.2 percent of the population – were literate.

Due to a lack of native educators, Libya has to resort to foreigners (from other Arab countries), who are required to teach in Arabic and sometimes in English. Wherever possible, foreign teachers are replaced by Libyans. According to the National Report (2001) presented in Geneva in 2001, there are 2,814 Libyans and 2,714 foreigners teaching in Libyan universities; and 1,403 Libyans and 2,234 foreigners teaching in other institutions and higher vocational centers. There are, therefore, more foreigners than natives teaching in Libya.

Bibliographical references

Agostini, Enrico de. 1954. "Sulle popolazioni della Libia". *Libia* 2.2.3–13.

Albergoni, Gianni (ed.) 1975. *La Libye nouvelle: Rupture et continuité.* Paris: CNRS.

Azema, James. 2000. *Libya handbook.* Bath: Footprint.

Beguinot, Francesco. 1942. *Il Berbero Nefûsi di Fassāṭo.* Rome: Istituto per l'Oriente.

Bessis, Juliette. 1986. *La Libye contemporaine.* Paris: L'Harmattan.

Burgat, François and André Laronde. 2000. *La Libye.* (= *Que sais-je?*, 1634.) Paris: Presses Universitaires de France.

Caubet, Dominique. 2000–2001. "Questionnaire de dialectologie du Maghreb (d'après les travaux de W. Marçais, M. Cohen, G.S. Colin, J. Cantineau, D. Cohen, Ph. Marçais, S. Levy, etc.)". *Estudios de Dialectología Norteafricana y Andalusí* 5.73–92.

——. 2004. "Les parlers arabes nomades et sédentaires du Fezzān, d'après William et Philippe Marçais". *Approaches to Arabic dialects: A collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday*, ed. Martine Haak, Rudolf de Jong, and Kees Versteegh, 67–96. Leiden: E.J. Brill.

Cesàro, Antonio. 1939. *L'arabo parlato a Tripoli.* Milan: Mondadori.

Cohen, Mordekhai. 1930. *Gli Ebrei in Libia.* Roma: Sindacato Italiano Arti Grafiche.

Curotti, Torquato. 1933. *Il dialetto libico: Grammatica, conversazione, dizionario.* Tripoli: Maggi.

Djaziri, Moncef. 1996. *Etat et société en Libye.* Paris: L'Harmattan.

Elfitoury, Abubaker Abdalla. 1976. *A descriptive grammar of Libyan Arabic.* Ph.D. diss., Georgetown University.

Felice, Renzo de. 1978. *Ebrei in un paese arabo.* Bologna: Il Mulino.

Geist, Stephen. 1981. *La situation linguistique à Tripoli.* Mémoire de DEA, University of Paris-3.

Goldberg, Harvey E. 1983. "Language and culture of the Jews of Tripolitania: A preliminary view". *Mediterranean Language Review* 1.85–102.

Griffini, Eugenio. 1913. *L'arabo parlato della Libia.* Milan: Ulrico Hoepli.

Ham, Anthony. 2002. *Libya.* Melbourne: Lonely Planet.

Iannotta, Epidio. 1933. *L'arabo parlato in Cirenaica.* Benghazi.

Julien, Charles-André. 1994. *Histoire de l'Afrique du nord des origines à 1830.* Paris: Payot.

Larcher, Pierre. 2001. "Les langues de la Libye: Passé et présent". *Revue des Deux Rives* 2.43–51.

Lethielleux, J. 1946. "Le Fezzan: Le pays et son histoire". *Revue de l'Institut des Belles Lettres Arabes* 34.185–208. Tunis: Institut des Belles Lettres Arabes.

Mantran, Robert. 1975. "La Libye des origines à 1912". Albergoni a.o. (1975:15–32).

Marçais, Philippe. 1977. *Esquisse grammaticale de l'arabe maghrébin*. Paris: Adrien Maisonneuve.
——. 2001. *Parlers arabes du Fezzân: Textes, traductions et éléments de morphologie*. Ed. Dominique Caubet, Aubert Martin, and Laurence Denooz. Liège and Paris: Universiy of Liège and Institut National des Langues et Civilisations Orientales.
Marçais, William. 1950. "Les parlers arabes". *Initiation à la Tunisie*, 195–219. Paris: Adrien Maisonneuve.
——. 1961. "Comment l'Afrique du nord a été arabisée". William Marçais, *Articles et conférences*, 171–192. Paris: Adrien Maisonneuve.
Martel, André. 1991. *La Libye, 1835–1990: Essai de géopolitique historique*. Paris: Presses Universitaires de France.
Motylinski, Gustave-Adolfe de Calassanti. 1898. *Le Djébel Néfoussa: Transcription, traduction française et notes avec une étude grammaticale*. Paris: Ernest Leroux.
——. 1904. *Le dialecte berbère de R'edamès*. Paris: Ernest Leroux.
Nehlil. 1909. *Etude sur la dialecte de Ghât*. Paris: Ernest Leroux.
Nikolova, Viara. 2003. *Fonetični i morfologični osobenostina libiyskiya dialekt* [Phonetic and Morphological features of the Libyan dialect]. M.A. thesis, University of Saint Kliment Ohridski (Sofia, Bulgaria).
Owens, Jonathan. 1983. "Libyan Arabic dialects". *Orbis* 23.97–117.
——. 1984. *A short reference grammar of Eastern Libyan Arabic*. Wiesbaden: O. Harrassowitz.
Panetta, Ester. 1943. *L'arabo parlato a Bengasi*. I. *Testi con traduzione e note*. II. *Grammatica*. Rome: Libreria dello Stato.
Pereira, Christophe. 2003. "À propos du parler arabe de Tripoli (Libye): Un parler mixte". *AIDA, 5th Conference Proceedings, Cádiz, September 2002*, ed. Ignacio Ferrando and Juan J. Sanchez Sandoval, 431–443. Cádiz: Universidad de Cádiz.
——. 2004. "Le parler arabe de Tripoli (Libye): Etat des lieux (d'après les travaux de Hans Stumme, Antonio Cesàro, Eugenio Griffini)". *Estudios de Dialectología Norteafricana y Andalusí* 7.7–57.
——. 2006. "Eléments de description du parler arabe de Tripoli (Libye)". *L'arabe dialectal: Enquêtes, descriptions, interprétations: Actes d'AIDA 6*, ed. Salah Mejri, 371–385. Tunis: Centre d'Etudes et de Recherches Economiques et Sociales.
——. 2007. "The impact of migrations and social change on the Arabic dialect spoken in Tripoli (Libya)". *Arabic in the city: Studies in dialect contact and language variation*, ed. Catherine Miller, Enam Alwer, Dominique Caubet, and Janet Watson. London: Routledge – Taylor.
Rossi, Ettore. 1954. "Storia della Libia dalla conquista araba al 1911". *Libia* 2:1.3–43.
Souriau, Christiane. 1986. *Libye: L'économie des femmes*. Paris: L'Harmattan.
Stumme, Hans. 1894. *Tripolitanisch-tunesische Beduinenlieder*. Leipzig: J.C. Hinrich'sche Buchhandlung.
——. 1898. *Märchen und Gedichte aus der Stadt Tripolis in Nordafrika*. Leipzig: J.C. Hinrich'sche Buchhandlung.
Trombetti, Alfredo. 1912. *Manuale dell'arabo parlato a Tripoli*. Bologna: L. Beltrami.
Yoda, Sumikazu. 2005. *The Arabic of the Jews of Tripoli (Libya)*. Wiesbaden: O. Harrassowitz.

Christophe Pereira (CRÉAM-INALCO, Paris)

Liḥyanitic → Thamudic

Lingua Franca

1. Lingua franca: Functions and structures

The term 'lingua franca' (or vehicular language) designates any language used as a means of interethnic communication in a multilingual setting. It usually refers to the spoken levels rather than to written levels, although an oral form of a lingua franca can become a written and standardized language. Lingua francas have been known since early Antiquity (for instance, Akkadian and Aramaic in the Near East). The development of a lingua franca is usually associated with one of the following historical and socioeconomic factors: expansion of trade, military conquest, political and religious domination, migration, or urbanization within multilingual areas. Its communicative function has no direct correlated impact on the linguistic structures, although expansion of a vernacular language as a lingua franca often results in a number of linguistic processes associated with simplification and semantic transparency, for example, phonological reduction, rule generalization, development of analytic structures over synthetic structures, and use of autonomous grammatical markers. It does not necessarily lead to a drastic linguistic restructuring such as those implied in → pidginization processes.

The linguistic structures of a lingua franca vary according to the social and historical context of its development and its contact. A lingua franca can develop from various types of language (an indigenous language, an imported dominant language, a 'mixed' language, or a pidgin language).

Due to its long historical expansion throughout many parts of the world, Arabic provides many different cases of its use as a lingua franca. Classical Arabic and, in particular, Modern

Standard Arabic (MSA) can be considered kinds of lingua franca among Arab intellectuals speaking different colloquial Arabic vernaculars and resorting to Modern Standard Arabic for inter-Arab communication, including in the American diaspora (Rouchdy 2002). An Arabic colloquial variety can became a regional or national dialectal lingua franca in countries or areas with high dialectal diversity (e.g. Casablanca Arabic in Morocco). Dialects spoken as regional or national lingua francas among Arab speakers have usually undergone → leveling processes and are labeled 'koine' (→ dialect koine). However, the typical cases of Arabic lingua francas are those where Arabic was or is used as an interethnic language among the non-Arab or non-Arabic-speaking population (e.g. in the → Horn of Africa). Some examples are provided herein, but due to the scarcity of available data, much remains to be studied in this domain.

2. Arabic as a historical lingua franca

Before the advent of Islam, there is no mention of an Arabic variety spoken as a lingua franca by non-Arabs or between Arabs and non-Arabs. Other languages, like Aramaic or Greek, were the regional lingua francas of the Middle East, and Old Arabic was either a colloquial vernacular or a literary koine (→ poetic koine). Following the first waves of the Arab conquests in the 8th century, Arabic is believed to have been spoken as a lingua franca among the Arabs and the various non-Arab groups of the Arab Empire, especially in the cities and the military garrisons. The spread of Arabic as a lingua franca has been regarded as one of the factors that could explain the formation of the Neo-Arabic dialects, which are characterized by a number of phonological reduction and analytic features compared to Classical Arabic (→ dialect koine). Early Arab sources (cf. Fück 1955) state that the expansion of Arabic and its use as a lingua franca induced the 'corruption' of the 'pure' Arabic language.

Very few data on the oral use of either the Arabic- or non-Arabic-speaking population are available for this early period in the written sources. The only linguistic sources are the numerous anecdotes about the speech of the early non-Arab converts, the *mawālī*. These anecdotes do not document the actual colloquial speech of the new converts but rather their effort to adopt the standard language in certain formal situations (Versteegh 1997:97). It is, therefore, difficult to assess what types of Arabic were spoken as a lingua franca and by whom exactly, including in the most bilingual or multilingual areas of the Arab Empire, such as Romance-speaking Andalusia, Berber-speaking North Africa, and South Yemen.

The further expansion of Islam and Arab traders or Arab groups in Africa and East and Central Asia throughout the medieval and Ottoman periods fostered contact between Arabic and local languages. Traces of this process are attested by the high degree of Arabic loanwords (Classical and sometimes colloquial), and often by the use of Arabic script to write the local African and Asian languages. Classical Arabic or → Middle Arabic may have functioned as a religious lingua franca among the various Muslim converts. However, we have as yet very little evidence of an Arabic lingua franca spoken in these areas during the premodern period. Other languages, such as Manding, → Fulfulde, → Songhai, → Hausa, → Swahili, → Persian, → Turkish, → Urdu, and → Malay, are known to have functioned and still function as lingua francas in these areas.

The few available historical data refer to very rudimentary forms of Arabic, which have been considered as kinds of 'pidgin' or 'jargon' by the authors who analyzed them. One of the earliest records is Maridi Arabic, which has been adduced as possible evidence of the presence of an Arabic-based pidgin used as a lingua franca in Saharan Africa (Thomason and Elgibali 1986). It was mentioned by the Andalusian geographer al-Bakrī in the 11th century. The geographical area of this pidgin is not clear (Mauritania, North Sudan, Upper Egypt?). The document is a very short text of fifty words written in Arabic script and characterized by a reduced morphosyntax (1).

(1)	<rʾdwl	dwmʿ	lw	ʾsm	dml
	man	Jumuʿa	to-him	name	camel
	lw	ʾww	bn	lwʾw>	
	to-him	and	son	to-him	

'A man called Jumuʿa had a camel and a son'

In spite of a careful linguistic analysis by Thomason and Elgibali, the limited data are not sufficient to establish the exact nature of the language or to prove the existence of a stable pidgin acting as a trade language in the southern marches of the Arab Empire.

Another source has been found in the writings of two Italian travelers of the 15th/16th centuries, who provided some Arabic sentences written in Latin script (Contini 1994, 1996). The Arabic variety presented in these Italian writings includes two registers: a Classical Arabic one, used in religious formulas, and a vernacular register, characterized by a simplified grammar, as in (2).

(2) *incane* *inte* *mayrith*
if you not-want
ane (= *ʾin kàn inta mà yrìd ʾana*)
me
'if you don't want me'

It has been analyzed as possible evidence of the existence of an Arabic lingua franca spoken by traders all around the ports of the Mediterranean Sea, the Red Sea, and the Indian Ocean (Cifoletti 2004). If the existence of an Arabic trade lingua franca appears historically quite probable, it has so far not been attested by the available linguistic data, which are merely caricatures of a 'corrupted' form of Arabic.

3. ARABIC LINGUA FRANCA IN MODERN TIMES

Today, the use of Arabic as an interethnic lingua franca among non-Arab speakers is recorded mainly in some African countries and in the Arabian Peninsula. It is linked either to the spread of Arabic in former non-Arabic-speaking areas or to the presence of a high percentage of non-Arabic-speaking migrant groups within an Arab host country (Gulf countries). In the first case, a distinction must be drawn between two types of areas: in bilingual areas, where non-Arabic-speaking groups have been living in more or less close contact with Arabic-speaking groups in a dominant Arabic environment (e.g. South Arabian-speaking groups in South Yemen, Berber groups in North Africa, African-speaking groups in northern Sudan and Mauritania); and in areas where Arabic native speakers are in the minority but where, nevertheless, Arabic functions as a lingua franca (Chad, South Sudan, Ethiopia, Eritrea). Information about the various types of Arabic spoken as lingua franca are rather scarce, and linguistic description is difficult due to the high degree of individual and social variation. It is often difficult to isolate a discrete linguistic variety, and very few structural factors appear to be shared by all vehicular Arabic varieties. According to the context, the Arabic lingua franca tends to approximate the regional Arabic colloquial varieties, with individual variation according to the degree of proficiency/acquisition, or it exhibits significant linguistic restructuring. Another trend is the development of mixed languages or → code-switching among the youth living in ethnically mixed urban areas.

In Algeria and Morocco, the bilingual → Berber-speaking population (approximately 40 percent of the population in Morocco and 20 percent in Algeria) is said to speak local colloquial Arabic as a lingua franca with the Arabic-speaking groups and with the Berbers of other dialectal/language groups (Boukous 1995, 1997). The degree of colloquial Arabic proficiency among Berber-speaking groups varies according to place and style of life. There is no precise description of a vehicular kind of Arabic spoken by the Berbers. The fact that Berber and North African dialects have coexisted for a long time and tend to converge (El Medlaoui 1998; Maas 2000) may explain the nonspecificity of Berber vehicular Arabic. Use of Berber/Arabic/French code-switching is also mentioned for young urban Berbers but is not well documented.

In Egypt, Arabic has been spoken as a lingua franca among the Matoki and Fadicca → Nubian speakers since their relocation after the construction of the Aswan High Dam (Rouchdy 1980). Nubian Arabic presents some cases of phonological restructuring (lack of pharyngeals /ḥ/ and /ʿ/) and some morphological particularities due to Nubian substrate (systematic use of suffix *-a* to mark adjectives).

In → Mauritania, there is no description of a specific Arabic vehicular spoken by the non-Arab population, although a simplified form of → Ḥassāniyya is reported to be spoken between the African population and the Bidān (Taine-Cheikh 1997). The lingua francas of the African-speaking population, who usually cluster in specific urban quarters, have been → Wolof and French. Today the urban youth tend to mix the various African languages, incorporating some Ḥassāniyya words, as, for example, in (3).

(3) *hiyye* *m'jappiye* *nak*
[Ḥassāniyya Wolof Wolof
mais *no* *wiyete*
French Pular Pular]
'She is beautiful, but what is her name?'
(Dia 2004)

In → Chad, Arabs represent about 10 percent of the population and have been, since the 16th and 17th centuries, scattered in the central part of the country. Since 1978, Arabic is one of the two official languages, together with French. Arabic is the first vehicular of Chad and is spoken as an interethnic lingua franca by about 50 percent of the Chadian population (Jullien de Pommerol 1997). Vehicular Arabic is one of the broadcast Chadian national languages and is often used in official contexts instead of Classical Arabic. The dominance of Arabic is a rather recent phenomenon, linked to the social and political transformations of Chad in the last three decades of the 20th century, but the use of a simplified Arabic as a trade language is attested since the beginning of the 20th century. In 1900, a military Arabic pidgin known as Turku was imported from Bahr al Ghazal (southwest Sudan) into Chari-Logome (present southern Chad) by the fleeing Sudanese soldiers of Rabeh. Turku became one of the trade languages of southern Chad, Central Africa, and northern → Cameroon (Derendiger 1923; Muraz 1932; for a detailed updated linguistic and historical analysis, see Tosco and Owens 1993). Turku was characterized by a reduced phonetic inventory (21 phonemes) and the quasi absence of derivational and inflectional morphology, as in (4) and (5).

(4) *ana* *doro* *kutu* *fi*
I want put on
dabra *anaki* *dawa* *seme*
ulcer your medicine good
'I want to put good medicine on your ulcer'

(5) *nas* *mardan* *ana* *hille* *anaki*
people sick of village your
'sick people of your village'

Today, the term 'Turku' is no longer in use, but Hagège (1973) briefly mentions a variety of 'rudimentary Arabic vehicular' with the same characteristic pidgin features, without giving any indication of the extent of its use. Jullien de Pommerol (1997:65) mentions a vehicular variety, Bongor Arabic (→ pidgin Arabic: Bongor), which is spoken in the south and is very similar to Turku, as in (6).

(6) *amis* *ana* *kutulu* *kalib* *al*
yesterday I kill dog who
addu *wiled* *hanay*
bite boy mine
'Yesterday I killed the dog who bit my son'

Jullien de Pommerol distinguishes Bongor Arabic from the other levels or varieties of vehicular Arabic spoken in the central and western parts of the country, as well as in urban environments like Abbéché (Roth 1979) or N'djamena (Jullien de Pommerol 1997, 1999). In these areas, the Chadian Arabic vehicular reproduces the main grammatical features of the regional dialectal varieties (i.e. derivational and inflectional morphology), albeit with a number of phonological and morphological restructuring processes: lack of emphatics, /x/ often realized as [k], /š/ realized as [s], reduction of morphological derivation, larger use of analytical structures, and idiomatic or lexical particularities indicating the influence of the local African languages or of French, as in (7) and (8).

(7) *mūsa bufūtna kullina fi ilim*
'Musa is more learned than us'

(8) *hī gālat kadar mohammet yamši l lekkōl*
hī tamši tibi leyāh fangāsu
'She says that when Mohammed goes to school, she will buy him fritters'

It may be noted that vehicular Arabic shares a number of phonological features with the non-nomadic Chadian Arabic dialects: lack of pharyngeals /ḥ/ and /ʿ/, frequent omission of laryngeals /h/ and /ʾ/, presence of /č/, and /ŋ/ and /ñ/ in non-Arabic loanwords. It also tends to select the features common to all Chadian Arabian dialects and excludes some more specific features (the genitive particle *hana* is generalized, while *hint* tends to disappear). Vehicular Arabic is becoming the national Chadian Arabic koine through the influence of the media, which act as a unifying force.

Vehicular Arabic is also becoming the mother tongue of urban children in major cities such as N'Djamena.

In northern Sudan, where Arab groups represent 50 percent of the population, Arabic began to expand during the 16th/17th centuries (Thelwall 1978; Miller 1989). Today, the degree of Arabization and bilingualism varies from one group to another and from one individual to another. During the second half of the 20th century, Arabic spread either as a lingua franca or as a mother tongue (51 percent Arabic as mother tongue in 1956, and 74 percent in 1993, according to the national census), mainly because of internal migration, urbanization, and interethnic mixing. The Arabic lingua franca is usually based on the local colloquial varieties, more or less influenced by the dominant Khartoum Arabic. Compared to the local Sudanese Arabic colloquial varieties (like Baggara Arabic in western Sudan, Šukriyya Arabic in eastern Sudan, and → Khartoum Arabic in northern central Sudan), it is often characterized by the following phonological and morphological restructuring: lack of emphatics; lack of pharyngeals; realization of velar /ġ/ as [x] or [g]; irregular gender and number agreement; irregular use of definite article; predominance of analytical genitive constructions (with *hana* or *bta*); use of preposition fi after a verb of motion; etc.; see, for example, (9) and (10).

(9) *mara bitāk al kwayes*
'your good wife'

(10) *yarju fi lbalad*
'They go back to the country' (Miller and Abu Manga 1992)

However, the Arabic lingua franca does not exhibit major linguistic restructuring and is often rather similar to the vernacular Arabic varieties spoken by some Arabized/sedentary groups of western Sudan or Chad. In Khartoum, the Arabic vehicular variety of the migrant population tends to adopt some koine features, such as the genitive particle *btā/btaʿ* instead of Sudanese *ḥagg* or *hana*, but it seems to retain (or adopt) many western Sudanese features (e.g. 1st pers. sg. imperfect in *n-: namši nagod* 'I go and stay').

The situation is rather different in southern Sudan, where a pidgin Arabic has developed as a lingua franca since the middle of the 19th century, due to a particular historical and social context (→ Juba Arabic). Although Arabic in southern Sudan presents extremely varied forms, going from pidgin/creole varieties to approximate colloquial forms, Juba Arabic appears today as a symbolic standardized norm clearly distinct from any colloquial varieties and structurally close to the Arabic-based creole ki-Nubi (→ Ki-Nubi). Juba Arabic may also be spoken in Sudan among the displaced populations of the south.

In → Ethiopia, mention of an Arabic lingua franca has been made by Ferguson (1972), who distinguishes three Arabic varieties: (i) the dialect spoken by a very small minority of about ten thousand former Yemenis or Sudanese immigrants; (ii) the lingua franca, structurally close to the Yemeni and eastern Sudanese dialects, which was used as a communicative medium between the Ethiopian Muslim communities; and (iii) the trade jargon used between the Arab traders and the local non-Arab population, which Ferguson describes as a "rudimentary pidginized form of Arabic with the usual features of pidginized Arabic, such as the m.sg. for all persons of the verb, and so on". No updated information is available.

In Eritrea, Arabic became one of the official national languages in 1993 (together with Tigrinya and English). A vernacular colloquial Arabic related to Saudi dialects is spoken by the Rashaïda (1 percent of the Eritrean population). Arabic is also spoken as a lingua franca, particularly among the Muslim population of the Eritrean southern coasts (Simeone-Senelle 2000; → lingua franca: Horn of Africa). Various factors explain the past and present diffusion of Arabic in Eritrea: the historical links between the African Red Sea Coast and the Arabic Peninsula; the diffusion of Islam; and the war between Eritrea and Ethiopia (1962–1992), which displaced more than one million Eritreans to neighboring Sudan, Egypt, and Saudi Arabia and created many internal displacements. The Arabic lingua franca spoken by the Kushitic-speaking population shows interference with the local languages (Afar, Saho), particularly regarding word order, which is Subject-Object-Verb, with the auxiliary coming after the verb (11).

(11) *ḫarb isāwwū wu igéttəˁū kānū ennās*
'They were fighting and slaughtering the people'

Furthermore, it exhibits reduction at the phonological level (lack of or weak emphatization, velar /ġ/ and /x/ realized as [g] and [k], merger of /š/ and /s/ in [s], etc.); and irregular use of definite article and of gender and number agreement. However, it remains structurally close to Yemeni and eastern Sudanese dialects.

In southern Yemen, Arabic is spoken as a lingua franca by the Mehri speakers (a → Modern South Arabian language) to communicate with Arabic speakers and with other South Arabian speakers (Ḥobyōt and Soqoṭri speakers). The Arabic lingua franca spoken by the Mehri speakers is phonologically and morphologically close to the local Yemeni varieties (Simeone-Senelle 2002). Some structures indicate a possible Mehri influence (feminine gender assignation of some words such as *baḥr* 'sea' like the Mehri equivalent *ráwrəm*). It tends to be a leveled variety and excludes the very specific features of both Mehri (like the ejective realization of emphatics) and the local Yemeni Arabic dialects (the realization of /j/ as [y] in some dialects).

In the Gulf countries, the foreign migrant population has increased drastically since the late 1960s. In countries such as Kuwait and the United Arab Emirates, foreign migrants represent 60 to 90 percent of the total population. A pidgin Arabic has developed as a lingua franca between native Arabic speakers and foreign migrant workers, especially those from the Indian subcontinent. Gulf Pidgin Arabic is marked enough to be satirized in serial TV shows, songs, and other media (Smart 1990). There is evidence (Wiswall 2002) that the Gulf foreigners' talk used by native Arab speakers is different from the lingua franca spoken by the migrants (overuse of copula *fi* among native speakers in sentences like *anta fi fakkar* ~ *anta sawwi fakkar* 'you [sg.] think'). Gulf Pidgin Arabic is based on Gulf Arabic, but it presents a reduced phonological and morphological system and is characterized by a high degree of analytical structures and many foreign loanwords, as in (12)–(15).

(12) *ana kallim inta sīda*
'I address you [sg.] directly'

(13) *niḥna mā yifham*
'We don't understand'

(14) *sayāra mal axū barra sawwi nazif*
'Wash my brother's car outside'

(15) *yemen janūbī sīm sīm lubnān*
'South Yemen is the same as Lebanon'

4. CONCLUSION

At the linguistic level, it appears impossible to provide a unified description of Arabic as a lingua franca. Each vehicular variety is built upon the local vernacular Arabic dialects and exhibits various degrees of restructuring. Two phenomena can be noted, however. First, in most recorded places, the Arabic lingua franca tends to approximate the local varieties, and there are very few examples of the emergence of a specific stabilized Arabic pidgin variety. Second, some regular linguistic tendencies can be found at the phonological and grammatical levels, such as the tendency to drop emphatic and pharyngeal consonants, irregular use of gender and number agreement or generalization of affix plural markers *-īn* and *-āt*, irregular use of the definite article, predominance of analytical genitive construction, absence of specific marked local dialectal features, etc. Substrate influence may appear at the phonological and lexical levels and in a few morphosyntactic features. In this sense, one may note the similarity of trends between the linguistic processes which led to the emergence of Neo-Arabic dialects and the linguistic restructuring of contemporary Arabic lingua franca. It should be noted that Arabic may function as a lingua franca in many other places that, for lack of data, are not recorded here.

BIBLIOGRAPHICAL REFERENCES

Boukous, Ahmed. 1995. "La langue berbère: Maintien et changement". *International Journal of the Sociology of Language* 112.9–28.

——. 1997. "Situation sociolinguistique de l'amazighe". *International Journal of the Sociology of Language* 123.41–60.

Cifoletti, Guidi. 2004. *La lingua franca barbaresca*. Rome: University of Udine/Il Calamo.

Contini, Riccardo. 1994. "Pidgins arabi nella letteratura italiana del cinquecento?". *Quaderni di Studi Arabi* 12.65–86.

——. 1996. " 'Iaseti ane mao migenon delain': Postilla all'arabo vartemanio". *Operosa parva per Gianni Antonini*, 147–165. Verona: Edizioni Valdonega.

Derendiger, R. 1923. *Vocabulaire pratique du dialecte arabe centre-africain.* Paris.

Dia, Alassane. 2004. "Politique et réalités linguistique en Mauritanie: Parlers jeunes en milieu négro-africain de Nouakchott". *Parlers jeunes ici et là-bas: Pratiques et représentations*, ed. Dominique Caubet a.o., 17–31. Paris: Harmattan.

El Medlaoui, Mohamed. 1998. "Le substrat berbère en arabe marocain: Un système de contraintes". *Langues et littératures 16: Contact et évolution historique des langues au Maroc*, 137–166. Rabat: Faculté des Lettres et Sciences Humaines de Rabat.

Ferguson, Charles A. 1972. "The role of Arabic in Ethiopia: A sociolinguistic perspective". *Sociolinguistics*, ed. J.B. Pride and Janet Holmes, 112–124. Middlesex: Penguin Books.

Fück, Johann. 1955. *Arabiya: Untersuchungen zur arabischen Sprach und Stilgeschichte.* Trans. Claude Denizeau, *ʿArabīya: Recherches sur l'histoire de la langue et du style arabe.* Paris: Didier.

Hagège, Claude. 1973. *Profil d'un parler arabe du Tchad.* Paris: Geuthner.

Jullien de Pommerol, Patrice. 1997. *L'arabe tchadien: L'émergence d'une langue véhiculaire.* Paris: Karthala.

——. 1999. *Grammaire pratique de l'arabe tchadien.* Paris: Karthala.

Maas, Utz. 2000. "Moroccan: A language in emergence". *Arabic as a minority language*, ed. Jonathan Owens, 383–404. Berlin: Mouton de Gruyter.

Miller, Catherine. 1989. "Langues et identité". *Le Soudan contemporain*, ed. Marc Laverne, 87–112. Paris: Karthala.

—— and Al-Amin Abu Manga. 1992. *Language change and national integration: Rural migrants in Khartoum.* Reading: Garnett; Khartoum: Khartoum University Press.

Muraz, Gaston. 1932. *Vocabulaire du patois arabe tchadien ou 'tourkou' et des dialectes sara-madjinngaye et sara-mbaye.* Paris.

Roth, Arlette. 1979. *Esquisse grammaticale du parler arabe d'Abbéché.* Paris: Geuthner.

Rouchdy, Aleya. 1980. "Language in contact: Arabic Nubian". *Anthropological Linguistics* 22.334–344.

——. 2002. "Language conflict and identity: Arabic in the American diaspora". *Language contact and language conflict in Arabic*, ed. Aleya Rouchdy, 133–148. London and New York: Curzon Press.

Simeone-Senelle, Marie-Claude. 2000. "L'arabe véhiculaire parlé en Erythrée sur la côte de la mer Rouge, de Massawa à Rahayta". *Oriente Moderno* 19 (80).153–181.

——. 2002. "L'arabe parlé dans le Mahra (Yémen)". *60 Beiträge zur Semitistik: Festschrift für Otto Jastrow zum 60. Geburtstag*, ed. Werner Arnold and Hartmut Bobzin, 669–684. Wiesbaden: O. Harrassowitz.

Smart, J.R. 1990. "Pidginization in Gulf Arabic: A first report". *Anthropological Linguistics* 32.83–119.

Taine-Cheikh, Catherine. 1997. "Pratiques de l'arabe et processus identitaires en Mauritanie". *Colloque 'Plurilinguisme et identité(s): Le cas du Maghreb', Mont Saint Agnan (2–3 Mai 1996)*, 66–79. Rouen: Université de Rouen.

Thelwall, Robin (ed.). 1978. *Aspects of learning in the Sudan.* Belfast: New University of Ulster.

Thomason, Sarah and Alaa Elgibali. 1986. "Before the lingua-franca: Pidginized Arabic in the eleventh century A.D.". *Lingua* 68.317–349.

Tosco, Mauro and Jonathan Owens. 1993. "Turku: A descriptive and comparative study". *Sprache und Geschichte in Afrika* 14.177–268.

Versteegh, Kees. 1997. *The Arabic language.* Edinburgh: Edinburgh University Press.

Wiswall, Abdul-Qadir. 2002. "Gulf Pidgin: An expandedanalysis". http://oak.cats.ohiou.ed/~aq321500. (Unpublished research for Ohio University.)

Catherine Miller (Iremam, Aix en Provence)

Linguistics and Arabic

1. Introduction

This entry shows some of the key relations between data from the Arabic language and developments in general linguistics. Although the relevant aspects are sometimes described primarily as contributions from general linguistics to Arabic language studies and sometimes from Arabic to general linguistics, the crucial point throughout is the interaction between our understanding of the structure of the Arabic language and our understanding of language as a general phenomenon. In most, if not all, cases, the particular phenomenon that is referred to in Arabic is also found in some other languages of the world, in particular in other → Semitic or, more generally, → Afro-Asiatic languages. However, given the prestige of Arabic and the amount of scholarship that has been devoted to this language – including all of Classical Arabic, Modern Standard Arabic, and the modern vernaculars – it is usually Arabic, rather than any other relevant language, that has played the crucial role in the development of the interplay with general linguistics; only Hebrew competes seriously with Arabic in this respect. Nonetheless, on occasion it is useful to refer to other languages in addition to Arabic, in order to emphasize the position of Arabic with respect to its linguistic structure among the languages of the world.

Until recently, in order to compare Arabic typologically with other languages, it was necessary to rely on essentially intuitive assessments by linguists of the frequency of different structural types across the languages of the world. With the publication of Haspelmath a.o. (2005), a tool has been placed in linguists' hands that permits rather exact statistical evaluation of certain structural features of language across the world's languages, and reference to the contributions to Haspelmath a.o. therefore figures in the body of this article. However, for features not covered in Haspelmath a.o., reliance must still in general be placed on subjective assessment of the incidence of particular structures crosslinguistically.

2. Phonetics and phonology

While one's initial impression might be that it is the richness of the consonant inventory of Classical Arabic (and many other varieties of the language) that is most striking, including the presence of crosslinguistically rather unusual articulations such as uvulars and pharyngeals, it is worth noting initially that one of the most striking features of the phonology of most varieties of Arabic, including the Classical and Modern Standard languages, is an absence, namely that of the phoneme /p/. Maddieson (2005c), using a sample of 566 languages, notes that only 32 languages (including Arabic, here represented by Egyptian Arabic of Cairo), or 5.7 percent, lack /p/. What is perhaps even more surprising is that all but five of these languages are spoken in the northern half of Africa, and Maddieson (2005c:27) speculates that the prestige of Arabic may have been one (though surely not the only) factor in the high incidence of /p/-less languages in northern Africa.

In terms of crosslinguistically rare articulatory types attested in Arabic and discussed in the contributions to Haspelmath a.o. (2005), the two relevant ones are uvulars and pharyngeals. Maddieson (2005b) finds uvular consonants in 98 of the 566 languages of his sample, or 17.3 percent, but the relatively unusual position of Arabic increases somewhat when we note that it has both uvular stops and continuants, found in only 48 languages of Maddieson's sample, or 8.5 percent. Turning to pharyngeals, Maddieson (2005a) finds them in only 21 of the 566 languages of his sample, or 3.7 percent, with the only areas of widespread incidence of pharyngeals being Afro-Asiatic languages and the languages of the North Caucasus. Since Maddieson's (2005a) reference variety of Arabic is Egyptian Arabic of Cairo, one other rare phoneme of many other varieties of Arabic, including both Classical and Modern Standard Arabic, is missed, namely /ṯ/, which occurs in 40 of Maddieson's 566 languages, or 7.1 percent. Maddieson (2005a) goes on to note the number of languages that contain both pharyngeals and /ṯ/, with a grand total of 2 in his sample, putting varieties of Arabic that have both in very select company indeed.

Among other phonetic features of the Arabic consonant system, pride of place must of course go to 'emphasis', i.e. the phonetic correlate of the so-called emphatic consonants, a feature that is unfortunately not covered in Haspelmath a.o. (2005). This feature in Arabic is important first for its phonetic nature, which seems to comprise → velarization as the main secondary articulation – and not, incidentally, pharyngealization, since velarization (including 'emphasis' in Arabic) and pharyngealization have very different assimilatory effects on neighboring vowels, as can be seen perhaps most clearly in the back [ɑ] quality in an emphatic environment versus the front [a] quality in a pharyngeal environment. Second, 'emphasis' is important in Arabic for the fact that it characterizes not only the phonetic nature of its locus, i.e. the segmental phoneme that is phonemically emphatic, but also spreads, to different extents in different varieties of Arabic, to other segments. Although a few varieties of Arabic, such as Maltese, have lost the phonemic opposition of 'emphasis', the former presence of emphatics has played an important historical role in the enrichment of the Maltese vowel system. A simple example can be seen in the Maltese realizations of Classical Arabic *sayf* 'sword' and *ṣayf* 'summer', namely *sejf* and *sajf* respectively, i.e. with the front vowel /e/ in a historically nonemphatic environment and the back vowel /a/ in a historically emphatic environment. As noted by Schabert (1976:50–52), in Maltese the effects of former emphatics show up essentially only in the case of *r* (where the distinction of emphatic vs. nonemphatic is secondary but widespread across varieties of Arabic) and of words that contained a labial or *n*. In other words, in the example of *sayf* vs. *ṣayf*, 'emphasis' must first have spread across to the *f*, and it is this *f* that is then most directly responsible for the originally allophonic vowel quality distinction; contrast the Maltese counterpart *dejjaq* of Classical Arabic *ḍayyiq* 'narrow', where in the absence of a labial, *r*, or *n* there is no trace of the earlier emphatic.

As a final feature of phonetic interest, we might note the → stress system, which varies considerably from one variety of Arabic to another. Although there is a basic system that characterizes Arabic in general, with stress falling on the last heavy syllable (where a syllable is heavy if it contains a long vowel or diphthong followed by at least one consonant or a short vowel followed by at least two consonants), there are nonetheless unusual exceptions that can pose problems for an elegant description, such as the characteristic Egyptian stress pattern...VC'CVCV, e.g. *mudar'risa* 'female teacher', contrasting with equivalents of *mu'darrisa* in most other varieties. Reference may be made to the discussion of stress in Arabic varieties in Goedemans a.o. (forthcoming).

3. Morphological structure

Arabic is well known for its basic triconsonantal root pattern, whereby the → root of a family of words, carrying the basic meaning of that family of words, typically consists of three consonants, with vowels, as well as prefixes, suffixes, and infixes, serving to express derivational or inflectional relations among words and word forms from the same root, as illustrated in the following forms derived from the root *k-t-b* 'to write':

kataba	'he wrote'
kutiba	'it [masc.] was written'
kitāb	'book'
kutub	'books'
kātib	'writer'
maktab	'office'
maktaba	'library'
muktatib	'subscriber'

This feature of Arabic was one of the main driving forces behind the development of autosegmental phonology (→ morphology), with its use of distinct tiers to represent different parts of the phonological structure of a word. In Arabic, for instance, the similarities and differences between *kataba* and *kutiba* would be represented by having the invariable root *k-t-b* on one tier and the vowels, *a-a-a* or *u-i-a*, on another, with rules for combining the two into a single derived representation by interdigitating consonants and vowels. But the power of this approach to the interface of phonology and morphology and its value in the description of Arabic goes beyond simple examples of the kind presented above. For instance, the third verbal derivative (Form III) of the root *k-t-b* has the active perfect form *kātaba* 'he corresponded', with the corresponding passive *kūtiba*. In this pair, one notes not only the difference in quality of the first vowel, which parallels that found in the basic form active *kataba*, passive *kutiba*, but also the consonant length of that first vowel in Form III, irrespective of its quality. In other words, decomposing the phonological representation into different tiers enables one to give a more adequate characterization of the interface between phonology and morphology. See further McCarthy (1985).

4. Morphological categories

For some of the morphological categories of Arabic, detailed comparison with other languages of the world is facilitated by relevant contributions to Haspelmath a.o. (2005), and in some instances Arabic turns out to represent an interestingly rare type. For instance, Siewierska (2005) examines the distribution of → gender across grammatical persons in independent persons. In her sample of 378 languages, 254 have no gender opposition in independent pronouns, and 103 restrict the gender opposition to the 3rd person, while only 20 (5.3 percent of the total sample, or 16.1 percent of those that show some gender opposition) have a gender opposition in the 1st and/or 2nd person; one of these is, of course, Arabic – Siewierska's reference variety is Modern Standard Arabic – with a gender opposition in both 3rd and 2nd persons. (There is, incidentally, an even rarer type, with a gender opposition in the 1st and/or 2nd person but not in the 3rd person, represented by only two languages in Siewierska's sample.)

One might also wonder about the crosslinguistic distribution of 'conjugated adpositions', which would include the combinations of preposition and pronominal suffix in Arabic, such as *min-hā* 'from her'. It is important to note in this connection that Arabic does not extend this phenomenon to prepositions governing lexical noun phrases, i.e. Arabic has *min baġdāda* 'from Baghdad' and not **min-hā baġdāda*, a construction type that is found in some languages. Bakker (2005) provides relevant information. In his sample of 378 languages, 63 have no adpositions, and are thus irrelevant to the present question. Of the 315 languages that do have adpositions, 209 have no agreement, while 106 (33.7%) do, putting Arabic in a minority, but a sizable one. Incidentally, only 23 languages (included in the just mentioned 106) go further by having agreement even when the adposition governs a lexical noun phrase.

The interest of the Arabic dual lies not so much in the existence of this category as in its transformation in the passage from Classical Arabic to the vernaculars. In Classical Arabic, the dual is a full-fledged grammatical category, obligatory on nouns in reference to a group consisting of two entities, and with an almost complete set of parallel morphological forms to express agreement (e.g. in adjectives) or to index pronominal arguments, with the sole exception of the absence of 1st person dual distinct from plural, both in the independent pronouns and in the morphology of the verb. Although the form of the dual is continued into the vernaculars, there have been two kinds of changes. The first is a shift in semantics, whereby the originally dual (here often called the → 'pseudodual') is simply used in the sense of the plural, as with Maltese *id-ejn* 'hands', which can refer to any number of hands greater than one, not necessarily just two; the pseudodual is characteristic of entities that typically occur in pairs but is not restricted to the interpretation of a pair. The second shift is a restriction of true dual semantics to a small number of nouns, typically of a type that are frequently counted, such as time units, e.g. Moroccan *yum-ayn* 'two days', Maltese *sent-ejn* 'two years'. Interestingly, some vernaculars have slightly different forms in the case of duals and pseudoduals; Moroccan, for instance, has *-in* for the pseudodual (e.g. *yədd-in* 'hands'), contrasting with *-ayn* for the true dual. Syrian even has a minimal pair with dual *ʾid-t-ēn* 'two hands' vs. pseudodual *ʾid-ēn* 'hands', distinguished by the presence vs. absence of the originally feminine marker *-t*. In other words, the loss of productivity of the Classical Arabic dual – arguably a simplification – has led to innovations marked by unusual complications.

One unusual feature of Arabic noun morphology, though one shared by a number of other Semitic and, more broadly, Afro-Asiatic languages, is the construct form of the noun, with its accompanying syntax, as in the contrast between the independent form of the nominative masculine sound plural seen in *muʿallim-ūna* 'teachers' and the construct form seen in *muʿallim-ū l-walad-i* 'the boy's teachers'. One feature of this construction is that the possessive relation is marked not only by the genitive case on the possessor (the *-i* suffix of *walad-i*), but also by a special form of the head noun. While a number of languages mark the head noun of a possessive construction by means of a possessive affix agreeing with the possessor, as in Turkish *adam-ın kitab-ı* 'the man's book', where the possessor *adam* 'man' has a genitive suffix and the head noun *kitap* 'book' has a 3rd person singular possessive suffix, an otherwise special form of the head noun as in Arabic seems crosslinguistically very unusual, although no specific detailed typological study seems to exist. Nonetheless, the presence of a distinct construct form seems remarkably tenacious across Semitic and some other branches of Afro-Asiatic. While the Arabic vernaculars have in general lost the distinction between independent and construct plural and dual forms (though they may survive in frozen form before possessive suffixes, as in Egyptian *rigl-ay-ya* 'my legs'), a new kind of construct has developed with nouns ending in *tāʾ marbūṭa*, following the loss of any segmental realization in the independent form but the retention of a pronunciation as /t/ in the possessive construction, as in Egyptian *rukba* 'knee', *rukb-it ʿali* 'Ali's knee', *rukb-it-u* 'his knee'. For syntactic properties of the construct, see Section 5.

Turning now to the verb, a major focus of attention has been on the tense/aspect system. Arabic has a basic morphological distinction between two forms, the so-called perfect (*faʿala*) and imperfect (*yafʿulu*), with variation in the vocalization separating *ʿ* and *l*. In the analysis proposed, for instance, in Comrie (1976:78–81), the semantics of this opposition is a combination of tense and aspect, which renders otiose the question whether this is a → tense or an → aspect opposition; it is both. The perfect combines past time reference with perfective aspect. The imperfect covers all other possibilities, i.e. any situation that holds at or after the reference point (thus covering present and future time reference), as well as imperfective with past time reference. An added feature is that the reference point is not necessarily the present moment, so that in appropriate contexts, for instance, the perfect can indicate a situation that precedes a future reference point. The use of various → auxiliaries and particles, some of which are already present in Classical Arabic, others of which develop differentially in the various vernaculars, and also of participial forms of verbs, enables more specific time reference and other values to be made explicit. For instance, to ensure an interpretation of the imperfect that combines imperfective aspect with past time reference, the perfect of the auxiliary 'to be' can be used to indicate the past time reference along with the imperfect of the lexical verb to indicate imperfective aspect, as in *kāna yaktubu* 'he was writing, he used to write'; the combination of the imperfect of the auxiliary 'to be' – this form of 'to be' normally receives future time reference, present time reference being expressed by absence of 'to be' even in copular sentences – with the perfect of the lexical verb indicates a perfective situation prior to a reference point in the future, as in *yakūnu kataba* 'he will have written'. The possibility of combining a small number of distinct forms into a number of different combinations gives rise to a system capable of expressing a rich set of semantic distinctions, with the vernaculars in particular having developed a number of fine semantic oppositions; see, for instance, Brustad (2000:165–230).

5. SYNTAX

In terms of constituent order typology, Arabic, or at least those varieties of Arabic that have basic Verb-Subject-Object constituent order in the clause, are typical representatives of this constituent order type, combining this order of the major constituents of the clause with prepositions rather than postpositions, a possessive construction in which the possessor (genitive) follows its head noun, and an order in which the adjective also follows its head noun. And even with respect to varieties of Arabic that have Subject-Verb-Object basic order in the clause, or for which it is hard to decide between the two orders, the correlations still represent by far the majority type. In this respect, then, Arabic is a good representative of a type that is widespread among the languages of the world, but equally does not display any strikingly unusual features. (Data relating to crosslinguistic variation here can be gleaned from comparing the relevant maps in Haspelmath a.o. [2005]; several varieties of Arabic are included in the samples of all or some of these maps, e.g. Modern Standard Arabic, Egyptian Arabic, Syrian Arabic, Iraqi Arabic, Gulf Arabic, Moroccan Arabic.)

While the differentiation of the major word classes (parts of speech) in Arabic, e.g. verb, noun, pronoun, adverb, preposition, is in general rather straightforward, there is a class of words where morphology and syntax seem to go in separate directions, in that these items have at least many syntactic properties of verbs although morphologically they appear to be nouns, pronouns, adverbs, or prepositions. In Arabic grammar these have come to be called → 'pseudoverbs'; for the most part, they are original nonverbs that have acquired some properties of verbs, although the class also includes original verbs that have lost some verbal properties. In Syrian, for instance, *ʿand* is originally a preposition expressing location at, and can still be used in this sense, but it can also be used as the translation equivalent of English *to have*, in which case it has some verbal properties, for instance being negated by means of the preverbal particle *mā*, and even taking a pronominal direct object suffixed to *yā-* (→ locatives). While the historical pathway by which pseudoverbs arise is reasonably clear, the phenomenon nonetheless presents problems in synchronic description, in particular in defining the word classes.

Within noun phrase syntax, perhaps the most striking feature of Arabic is the → construct (→ *ʾiḍāfa*), the basic way of expressing possession within the noun phrase in Classical Arabic and still a frequent construction in most vernaculars. In Section 4, the morphological peculiarity of this construction was already noted, namely the fact that there is often special marking of the head of the possessive construction (in addition, in Classical Arabic, to the genitive case on the possessor). But the construction also has unusual syntactic properties, as documented, for instance, by Benmamoun (2000:140–155) and other references cited there. First, there are restrictions on definiteness: while the possessor can be either definite or indefinite, the possessum can only be interpreted as definite, although it can never take an actual definite article, as can be seen in *kitāb-u l-muʿallim-i* 'the teacher's book', *kitāb-u muʿallim-i-n* 'a teacher's book'. Thus far, the construction is perhaps not strikingly different, other than in constituent order, from the so-called Saxon genitive construction in English, where again only the possessor, not the possessum, can show a definiteness opposition (*the teacher's book = the book of the teacher; a teacher's book = the book of a teacher*). However, the close link between possessor and possessum in the Arabic construction that is reflected in such restrictions goes well beyond that found in English. In Arabic, for instance, an adjective that belongs with the possessum may not separate possessum and possessor but instead must follow the possessor, e.g. *kitāb-u l-muʿallim-i l-jadīd-u* 'the teacher's new book'; contrast *kitāb-u l-muʿallim-i l-jadīd-i* 'the new teacher's book'. (In the absence of case marking, or if the possessum is also in the genitive, for instance after a preposition, ambiguity results as to whether the adjective belongs semantically with the possessum or the possessor.) These restrictions have given rise to a rich literature, including, in particular, attempts to find formal descriptions of the construction that are both general and well motivated as well as accounting for what might seem to be a set of rather idiosyncratic restrictions, all of which are, however, in some sense related to the particular tight relation between possessum and possessor in this construction.

An area of particular complexity in Arabic, → agreement, has parallels in other Semitic and more broadly Afro-Asiatic languages. This can be seen, for instance, in the number of references to Arabic (both Modern Standard Arabic and various vernaculars) and other Afro-Asiatic languages in Corbett (2006). The complexities in Classical Arabic range from the rule whereby broken plurals (at least if not referring to male humans) take feminine singular agreement, via the rule whereby the verb preceding its subject agrees in gender but not in number (remaining in the singular), to the pièce de résistance whereby the → numerals three to ten disagree in gender with their noun (i.e. masculine nouns take a feminine numeral, and vice versa), for example, *ṯamāniya-t-u kutub-i-n* 'eight books', with feminine *-t*, while *kitāb* 'book' is masculine. And while one could reanalyze the last of these by saying that the apparently feminine numerals are really irregular masculine forms, likewise for the apparently masculine forms used with feminine nouns, it should be noted that 'gender polarity' is not unknown in other languages of the world, especially other Afro-Asiatic languages; see, for instance, Corbett (1991:195–197). But even without going to the idiosyncrasies that delight typologists, more mundane features of the agreement system give rise to problems that have attracted the attention of formal grammarians, as can be seen, for instance, in Benmamoun (2000:121–139) and other references cited there. Benmamoun (2000:133–136) provides a nice example from Moroccan Arabic relating to conjoined noun phrases of the type *marwan w karim* 'Marwan and Karim'. If such a noun phrase as subject precedes its verb, then the verb must be in the plural, as in *marwan w karim ža-w* 'Marwan and Karim came'. In the order where the verb precedes the conjoined subject, either singular or plural verb is possible, i.e. *ža marwan w karim* or *ža-w marwan w karim*. However, under certain circumstances, only the plural verb is possible even when it precedes its subject. This is the case, for instance, when the verb expresses a 'collective predicate', such as *tlaqa* 'to meet', in the sense of 'to meet together, meet with one another', so that Moroccan allows *tlaqa-w marwan w karim* 'Marwan and Karim met' but not **tlaqa marwan w karim*. (By contrast, Modern Standard Arabic would have a singular verb in this last example.)

Another area of complexity in Classical Arabic, → negation, has been considerably simplified in the vernaculars (and to some extent in Modern Standard Arabic, at least in less formal registers). In Classical Arabic, the most general negator is *lā*, which is simply preposed to the imperfect (e.g. *yaḍribu* 'he strikes', *lā yaḍribu* 'he does not strike'). It also negates the imperative, although here instead of the imperative form one finds the jussive (e.g. *iḍrib* 'strike!', *lā taḍrib* 'do not strike!'). The negative corresponding to the perfect is the particle *lam* followed by the jussive (e.g. *ḍaraba* 'he struck', *lam yaḍrib* 'he did not strike), while that corresponding to the future with the particle *sawfa* is *lan* followed by the subjunctive (e.g. *sawfa yaḍribu* 'he will strike', *lan yaḍriba* 'he will not strike'). The only such complication likely to survive to a vernacular is a different means of negating the imperative. For instance, Maltese normally negates by simply adding the preposed particle *ma* and the suffix *-x* (pronounced [ʃ]), but the imperative is negated by using the corresponding indicative verbal form with the suffix *-x*, and either no preposed particle or preposed *la*, e.g. *ikteb* 'write!', *(la) tiktib-x* 'do not write!'. The theoretical implications of the splitting of different pieces of verbal semantics between the lexical verb and the negative particles are examined by Benmamoun (2000:94–110).

Object suffixes are a characteristic of all varieties of Arabic, but particular problems arise in the case of ditransitive verbs like 'to give'. In Classical and Modern Standard Arabic, it is possible for such a verb to take two accusative suffixes, as in *wahaba-nī-hi* 'he gave it to me', although another possibility is to construct the direct object as object of the particle *ʾiyyā*, as in *ʾahdaytu-ka ʾiyyā-hu* 'I gave it to you', a possibility that is also found in some vernaculars, e.g. Syrian Arabic *ʿaṭā-ni yā-ha* 'he gave it [fem.] to me'. In the West, the form used is rather with a reflex of the preposition *li* 'to' to mark the indirect object pronoun, as in Moroccan *ʿṭit-hom-lek* 'I gave them to you'. This is also the pattern in Maltese, as in *tajt-hom-lu* 'I gave them to him'. However, as discussed in detail in Comrie and Borg (1985) for Maltese, if the only pronominal object of certain ditransitive verbs (in particular, *ta* 'to give') is the indirect object, then in place of the dative form the accusative form is used, as in *tajt-u l-kotba* 'she gave the books to him'. Such examples point to an area of unexpected complexity that merits further examination across the full range of Arabic vernaculars.

6. SOCIOLINGUISTICS

Surely the most famous contribution of Arabic to the sociolinguistic literature is the concept of → diglossia, most clearly brought to the attention of general linguists by Ferguson (1959). The term refers to the substantial distinction between the 'High' variety of Arabic (Modern Standard Arabic) and the 'Low' variety (the local vernacular), with differences at all levels of grammar and in vocabulary. For instance in Iraqi Arabic (Baghdad), all four Modern Standard Arabic 2nd person nonsingular forms – masculine dual *katabā*, feminine dual *katabatā*, masculine plural *katabū*, feminine plural *katabna* – merge as the one form *kitbaw*, historically continuing the masculine plural. In Egyptian Arabic (Cairo), Modern Standard Arabic *ʾanf* 'nose' is replaced by *manaxīr*, while the verb *ḏahaba* 'to go' appears as *rāḥ*. Although the sharp division suggested by the term 'diglossia' has been useful in pointing to the differences between Modern Standard and vernacular Arabic, and may have been reasonably accurate in the past, it is clearly an oversimplification of the present-day situation, where not only are many intermediate varieties between 'High' and 'Low' found, but also other parameters of prestige define relations among language varieties, such as the higher prestige attached in many Arabic-speaking countries to urban over rural varieties, even when both might in some sense be characterized as 'Low' in terms of the original concept of diglossia. For a summary of the current situation, see Holes (2004, esp. 46–50, 341–389).

BIBLIOGRAPHICAL REFERENCES

Bakker, Dik. 2005. "Person marking on adpositions". Haspelmath a.o. (2005:198–201).

Benmamoun, Elabbas. 2000. *The feature structure of functional categories: A comparative study of Arabic dialects*. Oxford: Oxford University Press.

Brustad, Kristen E. 2000. *The syntax of spoken Arabic: A comparative study of Moroccan, Egyptian, Syrian, and Kuwaiti dialects*. Washington, D.C.: Georgetown University Press.

Comrie, Bernard. 1976. *Aspect: An introduction to the study of verbal aspect and related problems*. Cambridge: Cambridge University Press.

——. and A.J. Borg. 1985. "Object diffuseness in Maltese". *Objects: Towards a theory of grammatical relations*, ed. Frans Plank, 109–126. London: Academic Press.

Corbett, Greville G. 1991. *Gender*. Cambridge: Cambridge University Press.

——. 2006. *Agreement*. Cambridge: Cambridge University Press.

Ferguson, Charles A. 1959. "Diglossia". *Word* 15. 325–340.

Goedemans, Rob, Harry van der Hulst, and Ellen van Zanten (eds.). Forthcoming. *A survey of word accentual patterns in the languages of the world*.

Haspelmath, Martin, Matthew Dryer, David Gil, and Bernard Comrie (eds.). 2005. *The world atlas of language structures*. Oxford: Oxford University Press.

Holes, Clive. 2004. *Modern Arabic: Structures, functions, and varieties*. Rev. ed. Washington, D.C.: Georgetown University Press.

Maddieson, Ian. 2005a. "Presence of uncommon consonants". Haspelmath a.o. (2005:82–85).

——. 2005b. "Uvular consonants". Haspelmath a.o. (2005:30–33).

——. 2005c. "Voicing and gaps in plosive systems". Haspelmath a.o. (2005:26–29).

McCarthy, John J. 1985. *Formal problems in Semitic phonology and morphology*. New York: Garland.

Schabert, Peter. 1976. *Laut- und Formenlehre des*

Maltesischen anhand zweier Mundarten. Erlangen: Palm und Enke.

Siewierska, Anna. 2005. "Gender distinctions in independent personal pronouns". Haspelmath a.o. (2005:182–185).

Bernard Comrie
(Max Planck Institute for Evolutionary Anthropology)

Lisān

The basic meaning of *lisān* is 'tongue' (cf. al-Yāzijī 1954, s.v.). However, it has numerous other meanings. The data about this word can be gathered from texts in the *Qurʾān* and the *Ḥadīṯ* and from common literary expressions in both prose and poetry. The purpose of this entry is to clarify the sense in which this word is used in the Arabic heritage as well as in linguistics.

In the *Qurʾān, lisān* is used 25 times, 18 times in the singular, 7 times in the plural *ʾalsina* (ʿAbd al-Bāqī 1963:647, entry *l-s-n*). It means 'a part of the mouth' or 'an organ of speech and taste in human beings'. It is a blessing from God, for instance in Q. 90/9 "We have blessed him [sc. man] with two eyes, a tongue and two lips" (*ʾa-lam najʿal lahu ʿaynayni wa-lisānan wa-šafatayn*). Accordingly, it distinguishes mankind from animals: the tongues of animals are only for eating and swallowing, although they can be used as a means to produce various sounds. *Lisān* also means 'language': every apostle who receives a revelation is asked by God to address his people in their own tongue (Q. 14/4 *wa-mā ʾarsalnā min rasūlin ʾilla bi-lisāni qawmihi*). Through the *Qurʾān*, the Prophet Muḥammad was given a revelation specifically in the Arabic tongue, contrasting with those in non-Arabic tongues (e.g. Q. 26/195 *bi-lisānin ʿarabiyyin mubīn*; cf. Q. 16/3).

The word *lisān* is also used in the *Qurʾān* in other contexts. The 'faithful tongue' (Q. 26/84) is not only a blessing, it is also something that good people wish for. *Lisān* may also mean 'means of expression'. Moses first asks God to cure his speech defect or constraint (*ʿuqda lisānī*, Q. 20/27), then he describes Aaron (Hārūn) as being more fluent than he is (Q. 28/34 *huwa ʾafṣaḥ minnī lisānan*). Once, the singular *lisān* is mentioned as an instrument of *laʿn* 'cursing or damning [of unbelievers]' (Q. 5/78).

The plural form *ʾalsina* is mainly used for 'tongues', in various contexts. Muslims are warned not to employ their tongues as a means for satisfying their own desires (Q. 16/116). The creation of multifarious languages and colors for human beings is among the signs of the One God (Q. 30/22). *ʾAlsina* 'tongues' can utter false expressions (Q. 48/14), a characteristic of someone being a hypocrite (*munāfiq*): in the Hereafter, tongues will bear witness to the sayings and acts of mankind (Q. 34/24).

Other examples of the use of *lisān* are found in Arabic literature, both in prose and poetry. In prose, an example is a quotation from the Prophet Muḥammad: when asked about the meaning of *jamāl* 'beauty', he replied that it is the wording produced by the tongue (Jāḥiẓ, *Bayān* I, 19, 170). The tongue, though, the Prophet said, could in some cases be the cause of driving a person to hell (Jāḥiẓ, *Bayān* I, 194). Al-ʾAkṯam ibn Ṣayfī, an Arab thinker known for his wisdom, once preached to his people saying: "Be careful when you use your tongue in speech, because one word can be a direct cause of your death" (Faxūri 1960: 215). ʿAbdallāh ibn ʿAbbās is reported to have obtained his knowledge through two strategies, a wise heart and a frequently asking tongue (Jāḥiẓ, *Bayān* I, 85). The thing that needs to be kept permanently in jail is the tongue; once you release it, you will be charged for mistakes, and being silent in this case is much better than talking, for the tongue is a wild animal (*sabʿ ʿaqūr*; Jāḥiẓ, *Bayān* I, 194).

In popular wisdom, people still quote traditional statements such as these in various situations, or they use similar sentences in colloquial Arabic. A famous proverb says: *lisānak ḥuṣānak ʾin ṣunto ṣānak, wi ʾin hunto hānak* 'your tongue is as your thoroughbred; if you maintain it in good condition, it will do the same for you, and if you scorn it, it will scorn you'. The best and worst are contrasted in this respect, some people being characterized by having a tongue dripping with sugar or honey (*lisāno bi-ynaqqaṭ sukkar* or *ʿasal*), whereas others are accused of having an evil tongue (*lisāno zifir*), or their tongue is described as being as sharp as a file (*lisāno mabrad*), so that it should be cut off.

In Arabic poetry, the use of the tongue is frequently referred to. In many lines of poetry, it is said to be as sharp as a sword that brings victory (Jāḥiẓ, *Bayān* I, 156, 234). The tongue is the only defender of fame and reputation

(Jāḥiẓ, *Bayān* I, 159). An Arab jeweler-poet is quoted as saying that he is proud not of his skillful hands making jewelry but rather of his tongue composing poetry (Jāḥiẓ, *Bayān* I, 160). According to many poets, mankind is characterized by two things, either tongue and heart, or tongue and mind (Jāḥiẓ, *Bayān* I, 166). The *lisān* is always considered to be a tool that can be used to realize either good or evil. A body cut by a sword can be easily healed, but a hurt by a word of the tongue is never healed, remaining forever (Jāḥiẓ, *Bayān* I, 167, 170).

In his *Lisān al-ʿArab* (XIII, 385–387), Ibn Manẓūr (d. 711/1311) lists some of the Qurʾānic expressions in which *lisān* is a speech organ or a language, adding other meanings as well: 'a message', 'a piece of news', 'a word', 'speech in general', 'a lip', according to the context. Šarṭūnī (1989:II, 1141) also includes expressions in which the word *lisān* is used in genitive constructions (*ʾiḍāfa*). Some of the examples he cites are quoted and translated by Wehr (1971:866): 'silent language', 'mute expressions [as distinguished from the spoken word]', e.g. *lisān al-ḥāl* 'the language which things speak for themselves', *wa-lisān ḥālihi yaqūlu* 'while he seemed to say; with an expression as if he wanted to say', *ʿalā lisānihi* 'from this mouth; through him'; 'a newspaper', e.g. *ʿalā lisān aṣ-ṣuḥuf* 'through the medium of the press'; 'organ of a party or political movement', e.g. *lisān rasmī* 'official organ'; 'spokesman', e.g. *mutaḥaddiṯ bi-lisān wizārat al-xārijiyya* 'a spokesman of the foreign ministry'; *dāra ʿalā ʾalsinat al-xāṣṣa wa-l-ʿāmma* 'to be the talk of the town, to be on everyone's lips' (see also ʾAnīs a.o. 1972, s.v.).

The word *lisān* in addition is used in some compounds, especially names of plants, such as *lisān al-ḥamal* 'plantain [Plantago major]', *lisān al-ʿuṣfūr, al-ʿuṣfūr* 'common ash [Fraxinus excelsior]'. In popular speech, it is used for an object that resembles the tongue of a small bird, *lisān al-qufl* 'the tongue of the bolt of a lock'.

When people say about a person that he is *ḏū lisānayn*, this literally means that he is double-tongued, as a result of his being a hypocrite; hence, *malsūn* means 'a liar'. In addition to the entries of *lisān*/*ʾalsina*, there are two derivatives in the plural with a technical meaning, *lisāniyyāt* and *ʾalsuniyyāt*, both used to indicate a master's-level study in language or linguistics.

In the grammatical tradition, *lisān* is predominantly used for the tongue in a phonetic context. In the *Kitāb Sībawayhi*, the word *lisān* occurs 95 times, its plural *ʾalsina* 21 times (Troupeau 1976:189 translates with 'langue [phon.]'). Likewise, in al-Farrāʾ's *Maʿānī l-Qurʾān* the word is used for the articulatory activity of the tongue (Kinberg 1996:744). The only context in which *lisān* means 'language' is in connection with foreign languages. Al-Farrāʾ, for instance, mentions the *lisān al-Ḥabaša* 'language of the Ethiopians' (*Maʿānī* III, 206.7–8). In the early Qurʾānic commentaries, foreign languages are referred to both with *luġa* and with *lisān*, for instance, in references to the *lisān ʿAkka* and even to the *lisān Qurayš* (cf. Versteegh 1993:99–100); probably, this use was inspired by the Qurʾānic usage of *lisānun ʿarabiyyun mubīnun* 'a clear Arabic tongue'. In the meaning of '(foreign) language', it was replaced in Modern Standard Arabic by *luġa*, not unlike the development in English, where 'tongue' gave way to 'language'.

For dialectal varieties of Arabic, for instance of the Bedouin, → *luġa* was used. This also applies to the few descriptions of foreign languages in the Arabic linguistic tradition: ʾAbū Ḥayyān calls his book on Turkic *Kitāb al-ʾidrāk li-lisān al-ʾAtrāk* 'the Book of comprehending the language of the Turks', but when he mentions dialectal varieties in Turkic, he switches to *luġa* (Ermers 1999:292). Ibn Jinnī (*Xaṣāʾiṣ* I, 243), however, uses *al-luġa al-ʿajamiyya* even when he speaks about the Persian language in general, and az-Zajjājī (*ʾĪḍāḥ* 45.6) says that he checked the number of parts of speech "in a number of languages we got to know" (*fī ʿidda luġāt ʿarafnāhā*).

In the Coptic grammar in Arabic by Athanasius of Qūṣ (13th/14th century), Sahidic and Bohairic Coptic are said to form one *luġa* (Bauer 1972:234.5), but when Athanasius discusses the issue of the first language of mankind, he uses *lisān al-qawm al-ʾawwalīn* (Bauer 1972:244.11). Yet, *lisān* is not always used in discussions about the original language of mankind. The Qurʾānic phrase *wa-mā ʾarsalnā min rasūlin ʾillā bi-lisāni qawmihi* (Q. 14/4) is discussed by Ibn Ḥazm (*ʾIḥkām* I, 32–33) in a discussion about the original language of mankind, but in this discussion he consistently uses *luġa* to indicate the various languages that lay claim to this title, for instance when he quotes Galenos as saying that Greek is the best of all languages (*ʾinna luġat al-yūnāniyyīn ʾafḍal al-luġāt*).

In philosophical language, *lisān* and *ʾalsina* are used more frequently to indicate 'language', for instance by al-Fārābī (*Ḥurūf* 137.1), who speaks about the cause of the *ixtilāf ʾalsinat al-ʾumam* 'divergence of the languages of the nations'; in his enumeration of the sciences (*ʾIḥṣāʾ* 9.10ff.), he uses the term *ʿilm al-lisān* rather than *ʿilm al-luġa*. The predominance of *lisān* in philosophical usage is emphasized by Hadj-Salah (1986); he also states that *luġa* in the sense of *lisān* 'language' is unknown in grammatical treatises from the first centuries of Islam; it appears infrequently at the end of the 8th century, but usually *luġa* remains restricted to the naming of ethnic varieties.

The conclusion must be that *lisān* as a term for 'language as structure' was not common in linguistics but typically belonged to philosophical language. The usual term for 'language as structure' is → *kalām*, for instance in the frequent expression *fī l-kalām* 'in speech/in language' in Sībawayhi's *Kitāb*. *Lisān* was mainly used in a phonetic context for the tongue as an articulatory organ, which is involved in producing 16 consonants in Arabic out of the total number of 28. These 16 consonants are /g/, /š/, /y/, /ḍ/, /l/, /n/, /r/, /ṭ/, /d/, /t/, /ṣ/, /s/, /z/, /ḏ̣/, /ḏ/, /ṯ/ (Hassanein 1987:27, 28; Sībawayhi, *Kitāb* II, 404–405). In his description of the pronunciation of the Arabic consonants, Sībawayhi lists both the active and the passive articulator, the tongue being involved mostly as an active articulator, but he consistently derives the names of the categories of consonants from their place of articulation (*maxraj*). The action of the tongue is already mentioned in the *Kitāb al-ʿayn*, ascribed to al-Xalīl ibn ʾAḥmad. According to him (*ʿAyn* I, 51–52), the tongue only serves as an active articulator in the case of three consonants, /l/, /r/, and /n/ (*lā yanṭaliqu l-lisān ʾillā bi-r-rāʾ wa-l-lām wa-n-nūn wa-ʾammā sāʾir al-ḥurūf fa-ʾinnahā rtafaʿat fawqa ḏ̣ahr al-lisān min ladun bāṭin aṯ-ṯanāyā min ʿinda maxraj aš-šīn bayna l-ġār al-ʾaʿlā wa-bayna ḏ̣ahr al-lisān laysa li-l-lisān fīhinna ʿamal*). In its nomenclature, the *Kitāb al-ʿayn* does not distinguish between active and passive articulators, selecting either the one or the other to name the categories of consonants; the tongue is involved through its tip (*ʾasala*) in the *ʾasaliyya* consonants (/z/, /s/, /ṣ/) and through its apex (*ḏalaq*) in the *ḏalaqiyya* consonants (/r/, /l/, /n/; Al-Nassir 1993:14–16).

Bibliographical references

Primary sources

ʾAbū Ḥayyāṇ, *ʾIdrāk* = ʾAbū Ḥayyān Muḥammad ibn Yūsuf al-Ġarnāṭī an-Naḥwī, *Kitāb al-ʾidrāk li-lisān al-ʾAtrāk*. Ed. Ahmet Çaferoğlu. Istanbul: Evkaf Matbaası, 1931.

Fārābī, *Ḥurūf* = ʾAbū Naṣr Muḥammad ibn Muḥammad al-Fārābī, *Kitāb al-ḥurūf*. Ed. Muḥsin Mahdī. Beirut: Dār al-Mašriq, 1969.

——, *ʾIḥṣāʾ* = ʾAbū Naṣr Muḥammad ibn Muḥammad al-Fārābī, *ʾIḥṣāʾ al-ʿulūm*. Ed. and trans. Ángel Gonzalez Palencia. 2nd ed. Madrid and Granada: Consejo Superior de Investigaciones Científicas, 1953.

Farrāʾ, *Maʿānī* = ʾAbū Zakariyyāʾ Yaḥyā ibn Ziyād al-Farrāʾ, *Maʿānī l-Qurʾān*. Ed. Muḥammad ʿAlī an-Najjār. 3 vols. Cairo: ad-Dār al-Miṣriyya, 1966–1972.

Ibn Ḥazm, *ʾIḥkām* = ʾAbū Muḥammad ʿAlī Ibn Ḥazm, *al-ʾIḥkām fī ʾuṣūl al-ʾaḥkām*. Ed. ʾAḥmad Šākir. 8 vols. Cairo: Maṭbaʿat al-ʿĀṣima, n.d.

Ibn Jinnī, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. Ed. Muḥammad ʿAlī an-Najjār. 3 vols. Cairo, 1952–1956. (Repr., Beirut: Dār al-Hudā, n.d.)

Ibn Manẓūr, *Lisān* = ʾAbū l-Faḍl Jamāl ad-Dīn Mukarram ibn Mukarram Ibn Manẓūr, *Lisān al-ʿArab*. 3rd ed. Beirut: Dār Ṣādir, 1994.

Jāḥiẓ, *Bayān* = ʾAbū ʿUṯmān ʿAmr ibn Baḥr al-Jāḥiẓ, *al- Bayān wa-t-tabyīn*. Ed. ʿAbd as-Salām Muḥammad Hārūn. Cairo: Maktabat al-Xānjī, 1968.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. 2 vols. Bulaq, 1316 A.H. (Repr., Baghdad: Al-Muthanna Library, n.d.)

Xalīl, *ʿAyn* = ʾAbū ʿAbd ar-Raḥmān al-Xalīl ibn ʾAḥmad al-Farāhīdī, *Kitāb al-ʿayn*. Ed. Mahdī al-Maxzūmī and ʾIbrāhīm as-Sāmarrāʾī. 8 vols. Beirut: Muʾassasat al-ʾAʿlamī li-l-Maṭbūʿāt, 1988.

Zajjājī, *ʾĪḍāḥ* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān ibn ʾIsḥāq az-Zajjājī, *Kitāb al-ʾīḍāḥ fī ʿilal an-naḥw*. Ed. Māzin al-Mubārak. Cairo: Dār al-ʿUrūba, 1959.

Secondary sources

ʿAbd al-Bāqī, Muḥammad Fūʾād. 1963. *al-Muʿjam al-mufahras li-ʾalfāḏ̣ al-Qurʾān al-Karīm*. Beirut: Muʾassasa Jamāl li-n-Našr.

Al-Nassir [an-Nāṣir], ʿAbd al-Munʿim ʿAbd al-ʾAmīr. 1993. *Sibawayh the phonologist: A critical study of the phonetic and phonological theory of Sibawayh as presented in his treatise* al-Kitab. London and New York: Kegan Paul International.

ʾAnīs, ʾIbrāhīm a.o. 1972. *al-Muʿjam al-wasīṭ*. 2nd ed. Beirut: Dār al-Ḥadīṯ.

Bauer, Gertrud. 1972. *Athanasius von Qus* Qilādat at-taḥrīr fī ʿilm at-tafsīr*: Eine koptische Grammatik in arabischer Sprache aus dem 13./14. Jahrhundert*. Freiburg: K. Schwarz.

Ermers, Robert. 1999. *Arabic grammars of Turkic: The Arabic linguistic model applied to foreign* *languages and translation of ʾAbū Ḥayyān al-ʾAndalusī's* Kitāb al-ʾidrāk li-lisān al-ʾAtrāk. Leiden: E.J. Brill.

Fāxūrī, Ḥannā. 1960. *Tārīx al-ʾadab al-ʿarabī*. Beirut.

Hadj-Salah, A. 1986. "*Lug̲h̲a*". *Encyclopaedia of Islam* V, 803. 2nd ed. Leiden: E.J. Brill.

Hassanein, Ahmed Taher. 1987. *Naḏariyyat al-iktimāl al-luġawī ʿinda l-ʿArab*. Cairo: Dār Ḥajar.

Kinberg, Naphtali. 1996. *A lexicon of al-Farrāʾ's terminology in his* Qurʾān *commentary, with full definitions, English summaries, and extensive citations*. Leiden: E.J. Brill.

Šarṭūnī, Saʿīd al-Xūrī. 1989. *ʾAqrab al-mawārid fī fuṣaḥ al-ʿarabiyya wa-š-šawārid*. Beirut: Maṭbaʿa Marsalī al-Yasūʿiyya.

Troupeau, Gérard. 1976. *Lexique-index du* Kitāb *de Sībawayhi*. Paris: Klincksieck.

Versteegh, Kees. 1993. *Arabic grammar and Qurʾānic exegesis in early Islam*. Leiden: E.J. Brill.

Wehr, Hans. 1971. *A dictionary of Modern Written Arabic*. Ed. J. Milton Cowan. 3rd ed. Wiesbaden: O. Harrassowitz.

Yāzijī. Nāṣif al-. 1954. *Muʿjam al-qaṭīfa fī ʾasmāʾ ʾaʿḍāʾ al-ʾinsān wa-mā yataʿallaqu bihā wa-ṣ-ṣifāt al-jāriya ʿalayhi min al-ḥilā wa-l-ʿuyūb*. 1st ed. Beirut: Kulliyya Lubnān.

Ahmad Taher Hassanein
(Misr University for Sciences and Technology)

Literacy

This entry describes the definitional scope of Arabic literacy. Current literacy and educational statistics in the Arab region are presented and linked to the nature and complexities of Arabic reading. Some underlying linguistic reasons for the spread of illiteracy, such as → diglossia, → language policy and attitudes, and the Arabic writing system are then introduced and analyzed. Finally, a brief analytical review of current Arabic reading research and a short reference list are provided.

1. Definitional scope

Although not exactly a synonym of 'reading', the English term for 'literacy' has frequently been associated and often equated with 'reading' and 'reading achievement'. Literacy seems to refer to the basic knowledge of reading. Research summarized by the National Reading Panel (2000) describes literacy as a set of component skills, including phonemic awareness and decoding, fluency (speed and accuracy), vocabulary, and comprehension. Until recently, 'literacy' has frequently been understood and defined as a universal set of transferable reading and writing skills. This definition dominates much of the current policy and practice in literacy education and is opposed to the existence of different literacies.

Although the Arabic language has a term for 'reading' (*qirāʾa*), and even one for 'readability' (*inqirāʾiyya*), which relates to the ease with which a text can be deciphered and read, there is no word that translates the English term 'literacy' in Arabic. This latter concept is jointly covered by *ʾummiyya*, the term used for 'illiteracy', and *maḥw al-ʾummiyya*, which means 'eradication of illiteracy' or 'anti-illiteracy'. The frequent use of *ʾummiyya* when dealing with the literacy context gives a special social connotation to the meaning and place of the 'literacy' effort in the Arabic-speaking region. It also shifts the locus from the linguistic reality of the phenomenon to the social conditions and attitudes that are closely attached to it. This terminological void, which is detrimental to a clear understanding of the nature of the problem, would end with the coining and use of a new word such as *qirāʾiyya*, which would link Arabic literacy to its etymological source (Maamouri 1999).

Finally, 'literacy' and its opposite concept 'illiteracy' seem to relate to two different facets of the same reality. Stephen Pinker (1994:188) recently noted that illiteracy, the result of insufficient teaching, needs to be addressed within the defining framework of an incomplete and unsuccessful educational process – or the total lack of one – the reasons for which need to be studied and thoroughly analyzed.

In this entry, Arabic literacy is studied only within the linguistic perspectives of the Arabic language.

2. Literacy statistics in the Arab region

The number of illiterate persons in the 22 countries of the Arab region reached some 67 million in 2002, which accounts for 40 percent of the total population aged 15 years and over. A recent study conducted by UNESCO-Beirut in 2001 shows that Arab regional efforts have contributed greatly in reducing the levels of illiteracy from 48.7 percent in 1990 to 38.5 percent in 2001. Projections indicate that if these efforts continue successfully, 28 percent (about 75 million) of the region's population, estimated at 280 million, will still be illiterate. However, it is interesting to note that some experts believe that there must now be more than 100 million illiterate persons in the region because official literacy and educational statistics suffer from inadequate data collection and lack of accurate information.

According to UNESCO (UIS 2003), there is a noticeable regional discrepancy in country illiteracy statistics within and across the Arab states. Illiteracy rates vary widely in the region, ranging from 10.2 percent in Jordan to 59.8 percent in Mauritania. Five countries, namely Yemen (53.6%), Morocco (51.2%), Egypt (44.7%), Sudan (42.3%), and Algeria (33.3%) account for 49 out of the 67 million officially recognized illiterate adults in the region, while ten countries, namely Jordan, the United Arab Emirates, Bahrain, Djibouti, Oman, Qatar, Kuwait, Lebanon, Libya, and Mauritania, account for only 3.6 million illiterate adults.

The Arab region continues to show very alarming illiteracy rates among women and young girls, especially in the rural and underprivileged areas and sectors of society. Women's illiteracy is linked to other serious indicators of underdevelopment, such as infant mortality and family size. Although illiteracy rates for Arab women reached percentages which varied between 80 and 90 percent in the 1950s and 1960s, there has been a marked improvement in the education of girls in the past decades (Maamouri 1999). This improvement in education contributed to a drop in the average female illiteracy rates from 86.3 percent in 1970 to 49.4 percent by the year 2000, with a noticeable disparity which varies by age groups across and within most Arab countries. The breakdown of the illiteracy rates into age-specific rates shows that the highest proportion of female illiterates are in the 50-and-over age bracket. While illiterate older women are a feature common to all Arab states, with little or no exception, the illiteracy of young girls, who have traditionally been left out of the educational system, is highest in Saudi Arabia (44%), Sudan (50%), Morocco (56%), Yemen (66%), and Djibouti (69%).

3. Brief review of some Arab education statistics

The high rates of illiteracy that characterize the Arab region seem to indicate that the educational system is failing. While the educational crisis varies from country to country, all the Arab educational systems share the following negative characteristics: a questionable relevance, an unacceptably low quality level, and high repetition and drop-out rates, especially in poor rural and urban communities. There is a growing sense of inadequacy in the face of the deterioration of education in the Arab states.

Even though the Arab region registered a rapid expansion of its educational system, with enrollments increasing by 85 percent from 1975 to 1991, the proportion of school-age children who are left out of the system is still extremely high in Yemen, Morocco, and Sudan (almost 50% and higher). Between 1990 and 1995, enrollment grew by 5.2 million in the Arab states (from 30 to 35.2 million). As of 2000, school enrollment reached over 39 million. The 9 million school-age children (two-thirds of whom are girls) who are not yet enrolled represent 22 percent of the school-age population and are still a matter of great concern to the region.

The International Bureau of Education (IBE) – UNICEF statistics for primary school repetition based on 1990 figures supplied to UNESCO indicate that in ten studied Arab states, and with the exception of Jordan, the repetition trend appears to show a fall in the overall percentage and an increase in the actual number of repetitions. Analyzing the grade repetition phenomenon, the IBE study makes the following three points: (i) There is a significant link between repetition in the first grades of primary education and the learning of reading and writing; (ii) there is a need for signi-ficant changes in the teaching of reading and writing and for a thorough overhaul of the parameters and traditional practices usually applied to first literacy in formal and nonformal situa-tions; and (iii) there is a need for greater awareness of the impact of linguistic factors on school performance in general and on literacy in particular.

4. SOCIOLINGUISTIC CHALLENGES TO LITERACY

There is a marked differentiation between two related varieties of Arabic. On the one hand, *fuṣḥā* is mostly used for 'high' functions such as formal prayers, speeches, or lectures, and on the other hand, a number of Arabic dialects are usually used for 'low' functions, defined as home and family discourse, or trade and market conversations within and across diversified Arab societies. This situation, known as → diglossia, gives *fuṣḥā* special prestige valuation, as the language of the written Qur'ānic tradition, literary heritage, literacy, and education. It also creates a significant linguistic distance (Ibrahim 1983; Maamouri 1998) between the language of orality and the language of literacy, *fuṣḥā,* more commonly referred to as → Modern Standard Arabic (MSA). Modern Standard Arabic and *fuṣḥā* are used interchangeably from here on in.

The gap between *fuṣḥā*, the Arabic language of formal education and adult literacy, and the Arabic dialect or vernacular spoken at home and almost everywhere outside of school walls seems to be a major cause of low learning achievement in schools and low adult literacy levels everywhere in the Arab region. The mixture of language patterns in the classrooms (*fuṣḥā* and dialectal Arabic code-switching) is a cause of serious pedagogical problems, sometimes leading to a lack of adequate language competence, low linguistic self-confidence, and, consequently, social problems. *Fuṣḥā*, which is at the same time 'formal Arabic', the official language of all Arab states and a major key to socioeconomic promotion in the region, is difficult to learn and use because it is nobody's native language. The learning difficulties that relate to the common language of all Arabs stem from its lack of immediate relevancy to the learning process and to the environment of both child and adult learners.

The compartmentalization of the two major Arabic language varieties places *fuṣḥā*, the sole language of first literacy acquisition and educational learning, outside the immediate daily activities of the learners, whether children or adults. There is an important linguistic distance which separates *fuṣḥā* from the learners' personal experience, familiar topics, and concrete real-world materials. *Fuṣḥā* is thus disconnected from the reality of expressive functions, and its relevance and motivation for learners are significantly reduced, which in turn leads to serious educational and social consequences. The experience of learners with *fuṣḥā* literacy is that of an abstract and decontextualized language-learning situation, which brings with it 'linguistic insecurity' and often results in learner distress at error or failure to recall correct structures and patterns. *Fuṣḥā* is somewhat disconnected from the everyday reality of adult learner needs, and some literacy specialists are beginning to feel that it has now become somewhat urgent to look for new pedagogical approaches to literacy work.

Children come to the formal school setting with a great deal of knowledge about their oral language and with two to five thousand words which they comprehend aurally and can use grammatically to communicate. This is the foundation on which reading is usually built in other linguistic situations. Because they are rarely in contact with *fuṣḥā* in normal discourse situations with parents or friends or in real-life activities (at home or at play), Arab children's experience with their oral language does not serve as a satisfactory vehicle for drawing their attention to the features and conventions of Arabic reading. Instead, Arab children's proficiency in their mother tongue seems to create confusion and difficulty for the learning of connections between the diverse sounds of their oral language and the marks of the written language presented to them in the formal school setting. In special discourse events, in the classroom, or in play situations, Arab children learn to use oral *fuṣḥā,* but this does not usually happen without the appearance of artificiality and lack of spontaneity.

When learning to read, young and adult Arab readers cannot put their inherent native linguistic competence in colloquial Arabic to task. They cannot use their lexical familiarity with their native basic Arabic sounds, forms, structures, and syllabic and prosodic features because these are not necessarily identical with *fuṣḥā* forms and structures, even though they may show important and striking similarities. The linguistic relatedness which exists between *fuṣḥā* and the colloquial does not always provide helpful clues and does not necessarily contribute positively to successful reading, simply defined as easy and fluid word recognition and language comprehension (Perfetti 1986). In spite of their familiar etymological structure, *fuṣḥā* words are not necessarily easily understood because they show varying degrees of phonological and semantic differentiation.

5. Linguistic and orthographic challenges to literacy

The Arabic writing system is an alphabetic system, with 28 basic consonant letters. Most of these consonants show a very close resemblance in form, with only additional dots or strokes to distinguish them from each other. They are usually composed of one base form, and most of them have up to three or four distinct variant shapes. Graphemic variants differ depending on whether they occur independently (nonconnectors) or in initial, mid-, or final position within the word. The Arabic orthographic system is characterized by a plurality of letters (more than sixty base forms), which stems from the cursive nature of the Arabic script and its ample use of ligatures and letter combinations. The use of multiple letter forms leads to graphemic difficulty and becomes a significant learning problem and a considerable burden for the Arabic text-decoding process, which is vital for the acquisition of basic literacy skills.

The Arabic script uses diacritical forms (or diacritics) for vocalic representation (*a, i, u*). Four letters (*ʾalif* or → *ʾimāla, wāw, yāʾ*) are also used to represent vocalic length. One diacritical marking, the *šadda*, is used for lexical differentiation. Most of the grammatical functions at both the morphological and syntactic level are represented by the short vowels, which also represent mood and case endings in the Verb-Subject-Object *fuṣḥā* syntax. Thus, vocalic representation carries the weight of the whole grammatical system and is therefore extremely important in setting up functions leading to correct reading and acceptable text understanding. However, these short vowels are rarely present in everyday out-of-school writing, and they do not, as a rule, appear in most printed materials in the Arab region. Diacritical markings are rarely used in printed documents, and this generalized practice includes the *šadda* (consonantal length), as well as the → *hamza* (glottal stop).

The use of diacritics, which is restricted to primary school education and the sacred Qurʾānic text, seems to be limited to whatever length of time is considered sufficient for the learner to be initiated to reading without them – which generally amounts to between four and six years. The absence of vowels in the *fuṣḥā* Arabic text is an unnecessary and costly idiosyncrasy of the Arabic writing system. Nowadays, vocalized Arabic text seems to be used only in pure deference to the needs of young and inexperienced learners. In order to be able to read, everybody – even inexperienced neo- and low-literates – have to provide their own grammatical interpretations and bring to task considerable additional knowledge of syntax, vocabulary, and sometimes contextual interpretation in order to obtain correct and meaningful vocalizations, which allow them to reach acceptable word recognition and sense disambiguation. Because the Arabic reader needs to understand in order to read, the Arabic reading process seems to have completely reversed what is usually the norm in other languages, where people read in order to understand. The following examples show how complex and arduous the Arabic reading process is.

i. The bare unvocalized *fuṣḥā* form <k-t-b-t> has five readings and five corresponding semantic interpretations: (a) *katabtu* 'I wrote'; (b) *katabta* 'you [masc. sg.] wrote'; (c) *katabti* 'you [fem. sg.] wrote'; (d) *katabat* 'she wrote'; and (e*) kutibat* 'it [fem. sg.] was written'.

ii. Another important example of the reading complexities that are created by the above situation is found in the use of passive verb forms as sentence openers. These openers usually lead to interesting instances of 'garden path' sentences. In the bare/unvocalized Arabic sentence, graphemically represented by the consonantal strings <k-t-b + ʾ-l-k-t-ʾ-b>, one can, from the same graphemic form, start with the verb in the past tense, as in *kataba* 'he wrote', or choose the passive form, as in *kutiba* 'it was written'. Making either one of these two initial interpretations leads to specific and different reading paths. If the initial path is in the past tense, it should then be followed by the direct object, and the noun *kitāb* is in the accusative (*al-kitāba*). If the path is in the passive voice, then the noun *kitāb* is in the nominative (*al-kitābu*), and it is the agent of the passive verb. The comprehension monitoring required for a successful reading of the above sentence is usually difficult, as it sometimes takes a lengthy sequence of segments to reach the final clue that signals an incorrect initial interpretation and forces a startover in the reading process. Compared to English garden-path sentences cited by Pinker (1994:212), such as *Fat people eat accumulates* or *The man who hunts ducks out on weekends*, which are common to all languages, the above Arabic graphemic garden-path examples present a gratuitous and unnecessary obstacle to reading and would not exist if complete vocalic marking practices had been the norm in writing.

6. Review of current Arabic reading research

Surprisingly little scientific research has been conducted on Arabic reading acquisition and literacy, and even less in the Arab region itself. Most research on literacy relates to official and politically minded literacy statistics or statistical assessment of the performance of young or adult learners in reading skills in a formal or nonformal context – mostly done by UNESCO and affiliated education institutions. Little research appears to exist on the Arabic reading process as viewed from a cognitive or psycholinguistic perspective. Although scarce and not well-distributed in the Arab region, the currently available research mainly seems to address the contention that the linguistic duality which exists between the two varieties of Arabic might be related to some of the hardships that native Arabic beginning readers encounter and might even hinder their basic acquisition of basic academic skills (Ayari 1996; Maamouri 1998). The current research tries to provide some empirical backing to the role of diglossia in initial reading development and to the significant cognitive effect all diacritics and specifically vowels have on word recognition and reading comprehension.

An important body of existing research is represented in the longitudinal work done in the 1980s by Daniel A. Wagner and the University of Pennsylvania research team on the acquisition of literacy and Arabic-reading skills in Morocco. The researchers' objective was to provide a profile of the variability that exists in Arabic literacy acquisition in Morocco. Wagner (1993) gives a complete synopsis of the research project in his account "How to become literate in Morocco", as well as the two chapters he devotes to learning to read in Arabic and learning to read in a second language. The findings that orthographic features of Arabic are common stumbling blocks for word comprehension among young learners of Arabic (Wagner 1993:240) confirm the thesis presented above. The research also shows that knowledge in year one of Arabic letters, their graphemic variability, and pronunciation predicted more than 30 percent of the variance in reading achievement five years later. Early decoding skills at the single-word level explain an additional 14 percent of the same variance. Wagner's conclusion shows that there is "substantial reason to believe" that learning to read in Arabic necessitates an even greater reliance on decoding skills than in other languages. Wagner highlights the absence of vocalization diacritics as the main reason behind the growing difficulty of decoding for word recognition and paragraph comprehension, a difficulty which mars advanced Arabic reading stages and requires knowledge of appropriately correct inflectional endings and the ability to place full and correct diacritical marking. Wagner recognizes, however, that there is a great need for further empirical research to investigate the important question of the utility of diacritical marks for beginning versus proficient readers in Arabic.

In an empirical research study undertaken in Abu Dhabi on primary school reading errors and the role of diacritics for beginning readers, Rima Azzam (1990) examines the misreadings and misspellings that Arab primary school children make and identifies vocalization and its use of diacritical markings as the main culprit. Her research seems to suggest that diacritical markings are significantly important in the process of reading and comprehending written language at all levels of Arabic reading.

Salim Abu-Rabia (1998) investigated the effect of vowels on reading accuracy in Arabic orthography. Four kinds of written *fuṣḥā* Arabic texts (narrative, informative, poetic, and Qurʾānic) were administered to 64 native Arabic speakers. Three texts of each kind were presented in three reading conditions: correctly vocalized, unvocalized, and wrongly vocalized. The most important finding of this study is that vowels were found to significantly influence the reading of both poor and skilled readers in the four *fuṣḥā* writing styles in all three conditions. It was also found that both skilled and poor readers improved their reading accuracy in all writing styles when they read with vowels. This last study reinforces and supports similar previous findings obtained by Abu-Rabia (1996, 1997), in which it was demonstrated that the vowels and the sentence context were significant factors for word recognition for both skilled and poor *fuṣḥā* readers. Abu-Rabia (2000) investigated the contention that reading difficulties in Arabic in elementary school result from the diglossic situation of *fuṣḥā*, the language of books and school instruction, and its opposition to the spoken dialect of the home. Starting from the belief shared by educators, teachers, and parents that the exposure of young Arabic speakers to *fuṣḥā* in the preschool period is not useful and a burden to all, Abu-Rabia compared the reading comprehension performance of first- and second-grade children who had been experimentally exposed to literary Arabic throughout their preschooling period with the reading performance of a parallel control group only exposed to spoken Arabic during that period. He found, contrary to the commonly held belief, that the early exposure of Arab preschool children to *fuṣḥā* text (stories) enhanced their reading comprehension abilities and improved their performance in reading comprehension tests two years later. Finally, the following conclusions by Abu-Rabia (2000:155) are worth noting: (i) Policy makers may incorporate this pedagogy in all preschool years; (ii) kindergarten and elementary school teachers should be educated about diglossic issues; and (iii) “teachers at all levels [should] use literary Arabic as the language of instruction”.

Elinor Saiegh-Haddad (2003a) examined phonemic awareness and pseudo word decoding in kindergarten and first-grade Arabic native children. She hypothesized that because native speakers of Arabic first learn to read in *fuṣḥā* – a language structurally different from the local dialect they grow up speaking – the linguistic differences between the two Arabic language varieties would interfere with the acquisition of basic reading processes in *fuṣḥā*. Saiegh-Haddad studied the role of oral language in the acquisition of basic *fuṣḥā* reading processes for the purpose of researching the interface between exposure to *fuṣḥā* and top-level comprehension skill development, a vital issue for a theory of initial reading acquisition in diglossic or bidialectal settings. Going beyond the mere establishment of a possible causal link between exposure to *fuṣḥā* and the achievement of top-level comprehension reading skill development, Saiegh-Haddad (2003b) addressed some aspects of questions such as: Do diglossic variables or linguistic distance parameters interfere with the acquisition of basic reading processes in *fuṣḥā*? Which diglossic structures interfere with the acquisition of basic reading skills, the phonological, syntactic, morphosyntactic, or lexical? Finally, which reading skills (phonemic awareness, word decoding, reading fluency, or reading comprehension) are sensitive to diglossic variables? The study focused on phonemic awareness and pseudo word decoding because both are prerequisites to the acquisition of word reading. Its findings showed that although the first-grade children seemed to have benefited from the increased exposure to *fuṣḥā* structures that formal literacy instruction allowed, they still found the task of isolating standard phonological structures quite difficult. The study showed that diglossia and the phonological distance between the two varieties of Arabic are related to the native decoding ability of young Arab children.

7. Concluding remarks

An important part of the Arabic literacy problem is posed by the Arabic orthographic system and its failure to support easy and efficient reading. Orthography-related challenges, which usually result from centuries of use and misuse of the script, aggravate the linguistic problems described above. While it is difficult to deal with these linguistic issues, simple orthographic reforms could be introduced to improve reading and literacy in the Arab region. The only obstacle would be the mind-set of Arabs themselves and their adversity to – and reluctance toward – accepting any changes that relate to their most beloved language.

BIBLIOGRAPHICAL REFERENCES

Abu-Rabia, Salim. 1996. "The role of vowels and context in the reading of highly skilled native Arabic readers". *Journal of Psycholinguistic Research* 25.629–641.

——. 1997. "Reading in Arabic orthography: The effect of vowels and context on reading accuracy of poor and skilled native Arabic readers". *Reading and Writing: An Interdisciplinary Journal* 9.65–78.

——. 1998. "Reading Arabic texts: Effects of text type, reader type and vowelization". *Reading and Writing: An Interdisciplinary Journal* 10.105–119.

——. 2000. "Effects of exposure to literary Arabic on reading comprehension in a diglossic situation". *Reading and Writing: An Interdisciplinary Journal* 13.147–157.

Ayari, Salah. 1996. "Diglossia and illiteracy in the Arab world". *Language, Culture and Curriculum* 9.243–252.

Azzam, Rima. 1990. *The nature of Arabic reading and spelling errors of young children: A descriptive study.* Ph.D. diss., Teachers College, Columbia University.

Ibrahim, Muhammad Hassan. 1983. "Linguistic distance and literacy in Arabic". *Journal of Pragmatics* 7.507–515.

Maamouri, Mohamed. 1998. "Language education and human development: Arabic diglossia and its impact on the quality of education in the Arab region". Mediterranean Development Forum. Washington, D.C.: World Bank.

——. 1999. "Literacy in the Arab region: An overview". Wagner a.o. (1999:400–404).

National Reading Panel. 2000. *Report of the National Reading Panel: Teaching children to read.* Washington, D.C.: National Institute of Child Health and Human Development.

Perfetti, Charles A. 1986. *Reading ability.* New York: Oxford University Press.

Pinker, Stephen. 1994. *The language instinct: How the mind creates language.* New York: Harper Collins.

Saiegh-Haddad, Elinor. 2003a. "Linguistic distance and initial reading acquisition: The case of Arabic diglossia". *Applied Psycholinguistics* 24.431–451.

——. 2003b. "Bilingual oral reading fluency and reading comprehension: The case of Arabic/Hebrew (L1)-English (L2) readers". *Reading and Writing: An Interdisciplinary Journal* 16.717–736.

UNESCO Institute of Statistics (UIS). 2003. www.uis.unesco.org/en/stats/statistics/ed/g_lit_arab.jpg.

UNESCO-Beirut, Regional Office for Education in the Arab States. Unesco Institute for Education (UIE). 2003. *UNESCO Literacy and Adult Education in the Arab World, Regional Report for the CONFINTEA V Mid-Term Review Conference.* Bangkok: UNESCO Regional Headquarters.

Wagner, Daniel. A. 1993. *Literacy, culture, and development: Becoming literate in Morocco.* New York: Cambridge University Press.

——. 2004. "Literacy(ies), culture(s), and development(s): The ethnographic challenge". *Reading Research Quarterly* 39.234–241.

——, Richard L. Venesky, and Brian R. Street (eds.). 1999. *Literacy: An international handbook.* Boulder, Colo.: Westview Press.

MOHAMED MAAMOURI (University of Pennsylvania)

Literary Arabic → Modern Standard Arabic; Classical Arabic

Locatives

1. INTRODUCTION

In Arabic, as in many other languages, sentences with locative predicates, existentials, and possessives (*have*-type predicates) share many syntactic features. Crosslinguistic research shows that formal differences among these sentence types are restricted and highly predictable. Clark (1978), for example, groups them as 'locationals', and Freeze (1992) considers them part of a "universal locative paradigm", arguing that all three are derived from a single underlying structure in which a prepositional phrase is predicate and includes its own internal subject. Evidence from Arabic provides compelling evidence in support of this position.

Locatives are so called because their primary role is to "relate a referent to some point or location in space", a relation expressed formally through prepositions, case marking, or both; *existentials* differ in that they "indicate what may or may not exist", hence assert the existence of an item, or lack thereof, in a certain location; and *possessives* indicate "a relationship between someone who possesses something and the thing that they possess" (Matthews 1997:212, 121, and 288, respectively), with the possessor representing the location and the possessed the theme (→ possession). These semantic distinctions are at times associated with formal differences as well. The differences, as the English paradigm in (1)–(3) demonstrates, involve three areas: word order (predicate vs. theme), verb type (*be* in (1)–(2) vs. *have* in (3)), subject type (definite in (1) vs. indefinite in (2)), and the distribution of the expletive 'there' (present in existentials (2c), absent elsewhere).

(1) Locatives a. *The book is on the desk*
b. *On the desk is the book*

(2) Existentials a. *A book is on the desk*
b. *On the desk is a book*
c. *There is a book on the desk*

(3) Possessives a. *John has a book*

Substantial research has been done on these structures in the literature on syntax, semantics, pragmatics, and information structure (see the references at the end of this entry and others included therein).

The linguistics literature on Arabic, however, is curiously silent on the analysis of these structures. A brief survey of the literature shows that if they are discussed at all, the discussion is mostly within the context of a theoretical issue, such as grammatical relations and subjecthood (Comrie 1991; Mohammad 2000), relationship to other structures, typically under the rubric of equational (also equative, copular, nominal, or verbless) sentences (Anwar 1979; Eid 1991; Mohammad 2000), word order (Mohammad 2000), or some other syntactic phenomenon that is independent of their nature as locative, existential, or possessive constructions (Bahloul 1993; Belyayeva 1997; Eid 1993). The silence may be partly due to the extraordinary similarity of these structures in Arabic vis-à-vis such languages as English, where the similarity is at times not as obvious (see existentials and possessives in (2) and (3), respectively).

2. The Arabic paradigm

The Arabic paradigm, based on Egyptian in (4) and Modern Standard Arabic (henceforth "Standard Arabic") in (5), shows that they are all copular constructions consisting of a Theme (*kitāb*), a Locative Predicate (*ʿalā l-maktab*, *hina*, *ʿand sāmi* in (4)–(6) and their equivalents in (7)–(9)), and a Copula (*kān* 'to be', which is not lexical in these examples but appears in other tense-aspect configurations, as illustrated below in Sec. 5). Two major differences emerge between locatives, on the one hand, and existentials and possessives, on the other. First, the word order of the theme and its predicate is freer in locatives than it is in existentials and possessives: the subject can occur in pre- and post-predicate positions. Second, the existential marker *fī* and its equivalent *hunāka* appear in existentials and possessives but not in locatives. This patterning of possessives with existentials makes the two structures formally nondistinct, so much so that sometimes native speakers do not distinguish them. Anwar's (1979:118–122) discussion of existential 'expletive' *fīh*, for example, is illustrated by possessive sentences which include prepositional phrase predicates headed by the preposition *maʿa* 'with' and translated into English with the possessive verb 'to have', e.g. *maʿāya kitāb* (Anwar 1979:138) and *fīh maʿāya kitāb* (Anwar 1979:120), both translated as 'I have a book with me'. This may be explained on the basis that the possessive reading in Arabic is determined by the prepositional predicate head. Locatives and existentials allow a wide range of predicate locatives, which include adverbials such as *hina* 'here' and *hināk* 'there' (Standard Arabic *huna* and *hunāka*) and almost any preposition with a locative meaning, e.g. Egyptian Arabic *gamb* 'beside', *taḥt* 'below', *ʿand* 'at', and *fōʾ* 'above', and their Standard Arabic counterparts *bi-jānib*, *taḥta*, *ʿinda*, and *fawqa*, respectively. In possessives, however, the choice is limited to three: *ʿand/ʿinda* 'at', *maʿa* 'with', and *li* 'to', which are not always interchangeable and are subject to dialect variation. In Egyptian Arabic, *ʿand* is the most common, being applicable in more contexts than *maʿa* or *li*, which is the most restricted (→ possession, esp. Sec. 3). Some speakers indicate a preference for the possessive in (6d) over (6b–c) and for the existentials in (5c) over (5b). The structure in (6d) involves the movement of the locative NP (complement of the prepositional head) and co-indexation with its pronominal trace. This process is independently motivated in the language and is available to any definite NP.

(4) Locatives
a. *il-kitāb ʿala l-maktab/hina*
the-book on the-desk/here
'The book is on the desk/here'
b. *ʿala l-maktab/hina il-kitāb*
on the-desk/here the-book
'On the desk/here is the book'
c. **fī l-kitāb ʿala l-maktab/hina*

(5) Existentials
  a. **kitāb ʿala l-maktab/hina*
    ‘A book is on the desk/here’
  b. *ʿala l-maktab/hina kitāb*
    ‘On the desk/here is a book’
  c. *fī kitāb ʿala l-maktab/hina*
    there book on the-desk
    ‘There is a book on the desk/here’

(6) Possessives
  a. **kitāb ʿand sāmi*
    book at Sami
    ‘Sami has a book’
  b. *ʿand sāmi kitāb*
    at Sami book
    ‘Sami has a book’
  c. *fī kitāb ʿand sāmi*
    there book at Sami
    ‘Sami has a book with him’
  d. *sāmi*$_i$ *ʿand-u*$_i$ *kitāb*
    Sami at-him book
    ‘Sami has a book’

The Standard Arabic paradigm in (7)–(9) is identical to the Egyptian Arabic paradigm, except that the adverbial *hunāka* ‘there’ replaces *fī*. This use of *hunāka* as an existential marker may be an innovation in Modern Standard vis-à-vis Classical Arabic. (For comparable data from Palestinian and Standard Arabic, see Mohammad 2000:13–17.) In addition, the sentence-initial position occupied by *hunāka* in (8c) and (9c) can equally well be occupied by the verb *yūjadu* ‘is present’, an option not available in Egyptian (**mawgūd kitāb ʿala l-maktab*) without *fī* (*fī kitāb mawgūd ʿala l-maktab* ‘there’s a book present on the table’). Other distributional differences emerge between the two varieties due to the nature of the existential markers. Because the Standard Arabic *hunāka* is also an adverbial, it may occur pre- or postverbally. Hence, both orders, *yūjadu hunāka* and *hunāka yūjadu*, are possible, in which case the adverbial is interpreted as deictic, equivalent to English ‘over there, there is a book on the table’. In Egyptian Arabic, *mawgūd* cannot precede *fī*.

(7) Locatives
  a. *al-kitāb-u ʿala l-maktab-i*
    the-book-Nom on the-desk-Gen
    ‘The book is on the desk’
  b. *ʿala l-maktab-i l-kitāb-u*

(8) Existentials
  a. **kitāb-un ʿala l-maktab-i*
  b. *ʿala l-maktab-i kitāb-un*
    on the-desk-Gen book-Nom
    ‘On the desk is a book’
  c. *hunāka kitāb-un ʿala l-maktab-i*
    there book-Nom on the-desk-Gen
    ‘There is a book on the desk’

(9) Possessives
  a. **kitāb-un ʿinda sāmi*
  b. *ʿinda sāmi kitāb-un*
    at Sami book-Nom
    ‘Sami has a book’
  c. *hunāka kitāb-un ʿinda sāmi*
    there book-Nom at Sami
    ‘There’s a book at Sami’s’
  d. *sāmi*$_i$ *ʿinda-hu*$_i$ *kitāb-un*
    ‘Sami has a book with him’

There is a preference for the inversion strategy, illustrated in the (b) versions of (8) and (9), over the existential marker strategy in the (c) versions, in opposition to preferences expressed in the Egyptian Arabic paradigm, possibly supporting the view that the existential marker is an innovation in Modern Standard Arabic under the influence of the dialects. With the exception of the existential marker, syntactic differences among the three structures in Arabic are the result of processes independently motivated in the language and involve interactions of word order with subjecthood and (in)definiteness effects.

3. Word order and subjecthood

Variation in word order that distinguishes locatives from existentials or possessives is attributed to the (in)definiteness of the thematic subject and restrictions that apply to indefinite subjects. Mohammad (2000), for example, finds that there are no restrictions on the place that definite subjects can occupy in a sentence since they can occur before or after a verb in VS and SV positions (*nāmat al-bintu/al-bintu nāmat* ‘the girl slept’). But there are restrictions that apply to indefinite subjects, one being that nonspecific subjects cannot occupy the initial sentential slot (**waladun nāma/nāma waladun* ‘a boy slept’). The restriction is relaxed if the subject is specified; a specific subject is any NP that is modified, irrespective of its being definite or indefinite, by an adjective (*waladun ṭawīlun jāʾa* 'a tall boy came'), by another nominal (*rajulun ṭabībun jāʾa* 'a man doctor came'), and by being the first member of the construct state (*ibnu ṭabībin jāʾa* 'a son of a doctor came'). Examples are from Mohammad (2000:9–12), who also cites the Arabic grammatical tradition for the view that "the subject of an equative sentence can be any specific NP" (Mohammad 2000:13). For discussion of → specificity, see Anwar (1979:45–64) on Egyptian Arabic and Mohammad (2000) on Palestinian and Standard Arabic, among others; and for the analysis of the locationals in Arabic grammatical theory, → *mafʿūl fīhi*.

This restriction on the distribution of nonspecific subjects explains the ungrammaticality of the (a) versions of the existentials and possessives in the Arabic paradigms. The thematic NP (*kitāb*) is a nonspecific indefinite, hence it cannot occur in sentence-initial (pre-predicate) position. The inverted word order, illustrated in the (b) versions, is one of the strategies employed in the language to mitigate this restriction, allowing the predicate to appear in this sentence-initial position. Another strategy used to mitigate this word order restriction is the use of the locative marker *fīh/fī* 'there', as in the (c) versions, thereby licensing the word order nonspecific subject < predicate. As Mohammad (2000:17) correctly observes, word order becomes freer with *hunāka* and *fī* where the subject and predicate can occur in either position relative to each other.

The discussion above explains away the word order variation in the Arabic Locative Paradigm on the basis of a language-specific restriction on the distribution of definite and indefinite subjects, attributing it to language-specific constraints that motivate the two strategies: 'inversion' and 'proform'. Freeze (1992), however, finds that inverted word order, as in the (b) versions of the existential paradigms, is characteristic of existentials across the 35 languages he studied, which included Arabic. The 'pro-form' strategy, i.e. the use of a locative marker such as *there, fī*, and *hunāka*, is found only in some of them and is, therefore, restricted crosslinguistically.

4. THE EXISTENTIAL PRO-FORM

In the literature reviewed, two transcriptions appear for the existential dialectal marker *fīh* and *fī*, but there is no discussion of the difference, which may be coincidental but may also be related to an author's assumptions about the nature of this marker. The first transcription (Anwar 1979; Mohammad 2000; Mughazy 2006) is supported by the written form which appears in some texts as فيه with a final *hāʾ* [h]; the second (Eid 1991; Freeze 1992; → pseudoverbs; → possession) more clearly captures the prepositional source of this marker. The advantage of the first transcription is that it distinguishes this form, available only in the dialects, from the preposition *fī* 'in', which is assumed to be its source. Its drawback is the potential for an interpretation of the [h] as being the pronominal 3rd person masculine singular clitic *-h* that appears postvocalic, as in *ʿalē-h* 'on him', *fī-h* 'in him', and *wayyā-h* 'with him', which it is not. The distinction, however, is not formally necessary since the existential and prepositional forms are distinct in terms of their behavior and the positions in the syntactic structure in which they can occur. There is also a tendency to maintain vowel length in the pronunciation of the existential marker, which tends to produce aspiration, hence the final [h] pronunciation, although the vowel may be shortened as in a question like *fi ḥadd fi-l-bēt* 'is there someone in the house?', where word stress falls on *ḥadd,* allowing the vowel to shorten in accordance with Egyptian syllabification rules. A similar phenomenon occurs with question words such as *ʾēh/ʾē* 'what?' and *lēh/lē* 'why?', where the glottal [h] is pronounced when followed by a vowel, as in the famous song entitled *lēh ana baḥibbak* 'why do I love you?'. Syllabification is indicated on the first word where *lēh* 'why' and *ana* 'I' merge phonologically into one word [leːhana]. No such evidence is available for the existential marker, as the impossibility of **ma-fih-š walad fi l-bēt* with the meaning 'there is no boy in the house' demonstrates. Since the transcription issue has not been raised so far in the literature, it remains to be seen whether arguments will emerge in favor of one or the other transcription in the future and the extent to which such arguments would generate discussions of the nature of this marker.

Table 1. The existential marker in some Arabic dialects

A.	Ṣanʿānī Arabic	*bih*
	Egyptian Arabic	*fī(h)*
	Levantine (Jordanian, Lebanese, Palestinian, Syrian)	*fī(h)*
	Mardin Arabic (Anatolia)	*fīyu*
	Cypriot Maronite Arabic	*fia*
	Daragözü Arabic (Anatolia)	*kefī, kefī*
B.	Afghanistan/Uzbekistan Arabic	*hast* (< Persian *hast* 'it is')
	Khuzestan Arabic	*hasset*
	Baghdad Arabic	*aku*
	Moroccan Arabic	*kayen*
C.	Tunisian Arabic	*ṯamm, famm*
	Andalusian Arabic	*ṯám*
D.	Omani Arabic	*šay*

Crossdialectal studies of these structures may also provide interesting insights into the nature of this marker. While the majority of Arabic dialects use a version of *fī,* there is variation among some. The list in Table 1, which is not exhaustive nor systematically selected, illustrates this variation. The forms in groups A and B are related to the meaning of 'in' and 'be', respectively, the Omani Arabic form in D is related to 'thing', and those in C need further research.

There is substantial debate in the literature on the nature of the existential/locative marker 'there' and its equivalents across languages as being an argument, an expletive, or a pro-form. Freeze (1992) uses the term 'pro-form' and argues on the basis of its locative nature in his sample of 35 languages that it is not an argument and cannot be generated in argument position.

In the literature on Arabic, the term 'expletive' has been used by both Mohammad (2000) and Anwar (1979) to refer to dialectal *fī*; Mohammad uses it in relation to Standard Arabic as well. Mohammad (2000:10 n. 14) describes Standard Arabic *hunāka* as being exactly the same as English 'there' since *hunāka* can occur "both as subject [hunāka fi d-dār-i walad-un] and as a locative predicate [al-waladu hunāka]". It is possible to infer that Mohammad considers the Palestinian Arabic *fīh* as being an expletive as well. In another footnote (Mohammad 2000:39 n. 24), he includes 'the expletive *fīh*' as one of the 'non-specific' nominals that can amalgamate with negation and be licensed by it; others include pronouns (*ma-hu-š* 'he's not' and the nominal *ḥadd* 'someone' (*ma- ḥadda-š* 'no one'). Anwar (1979) and Freeze (1992), however, argue explicitly against the argument nature of this marker. Anwar considers it part of the predicate (1979:119), and Freeze includes it in INFL as a feature of verbal inflection, arguing that it is the spell-out of the feature [+Loc] in INFL. Eid (1991, 2004) also includes it as a feature in INFL on the basis of its behavior in relation to negation, agreement, and the copula *kān.*

Arguments in favor of the existential/locative marker being part of the predicate involve its ability to select its subject: it selects an indefinite thematic subject, allows –Def, +Spec, and excludes +Def, +Spec. Thus, it exhibits verblike (→ pseudoverbs) behavior in subject selection. Like verbs, it also licenses subject-initial word orders where they would not otherwise be licensed.

5. Copula visibility, negation, and agreement

Additional evidence in support of the nonargument status of the existential *fī* comes from its interaction with a visible copula verb (Eid 1991, 2004), negation (Eid 1993), and agreement patterns.

The pro-form *fī* can co-occur with the copula verb *kān* when the copula is visible (i.e. in past, future, and present progressive/habitual tense-aspect configurations), but it cannot precede the copula verb.

(10a) *kān fī walad fi l-bēt*
was there boy in the-house
'There was a boy in the house'
(10b) **fī kān walad fi l-bēt*

It also carries sentential negation, taking the discontinuous *ma-š* characteristic of verbs rather than *miš*-associated nominal negation, e.g. *miš ana* 'not I'. If the copula verb is visible, it carries negation, however. For the distribution of pronouns and negation, see Eid (1991) and Benmamoun (2000).

(11a) *ma-fī-š walad fi l-bēt*
Neg-there-Neg boy in the-house
'There isn't a boy in the house'

(11b) *ma-kan-š* *fī* *walad* *fi*
Neg-was-Neg there boy in
l-bēt
the-house
'There wasn't a boy in the house'

This behavior supports the analysis of *fī* as part of the verbal categories represented, in one theoretical model, as features included in INFL (Freeze 1992; Eid 1993). Finally, *fī* differs from the copula and other verbs in that it does not show any agreement features with its thematic subject. Eid (1993) shows this to be true of the verb *kān* in Egyptian Arabic as well. She finds, for example, no evidence of agreement with thematic subject in existential and possessive constructions, suggesting that they may be impersonal constructions; and if they are, then their subject would be an expletive (pro), which in Arabic is always 3rd person masculine singular. The examples in (12) taken from Eid (1993) illustrate the absence of gender and number agreement with feminine and plural themes.

(12a) *fī* *maktab-a/mudarris-īn* *fi*
there bookcase-f/teacher-mp in
'There is/are a bookcase/teachers in the room'
l-ʾōda
the-room
(12b) *kān fī maktab-a/mudarrisīn fi l-ʾōda*
'There was/were a bookcase/teachers in the room'

(13a) *ma-fī-š maktab-a/mudarris-īn fi l-ʾōda*
'There isn't/aren't a bookcase/teachers in the room'
(13b) *ma-kan-š fī maktab-a/mudarris-īn fi l-ʾōda*
'There wasn't/weren't a bookcase/teachers in the room'

In structures with → topicalization (14)–(15), pronominal co-indexing with the topicalized NP appears on the copula and on *fī*, a phenomenon independent of subject agreement.

(14a) *il-ʾōda*$_x$ *fī maktaba fī-ha*$_x$
(14b) *il-ʾōda*$_x$ *ma-fī-š maktaba fī-ha*$_x$
(14c) *il-ʾōda*$_x$ *ma-kan-š fī-ha*$_x$ *maktaba*

(15a) *ma-fi-hā-š* *ḥāga*
Neg-there-her-Neg something
'There's nothing in it'
(15b) *ma-fī-š* *ḥāga* *fī-ha*
Neg-there-Neg something in-her
'There isn't anything in it'
(15c) *ma-fī-š* *fī-ha* *ḥāga*
Neg-there-Neg in-her something
'There isn't anything in it'

(16a) *ma-ʿand-ī-š* *kitāb*
Neg-at-me-Neg book
'I don't have a book'
(16b) *fī* *ʿanda-ha* *kutub*
there at-her books
'She has books with her [in her possession]'

Comrie (1991) points to dialectal variation in patterns of agreement involving the copula verb *kān* in possessive constructions (cf. Stassen 1995 for crosslinguistic variation). According to Comrie, there is an across-the-board preference for the default 3rd person masculine singular, as illustrated in (12)–(13). But in some dialects, for instance those of Tunis and Meknes, the copula may agree with the possessor (i.e. location) or the possessed (i.e. theme). Agreement with the possessor is possible, as in (17) from Comrie (1991), but less preferred. In (17a), *kunt* carries marking for the 1st person singular, and in (17b), for the 3rd person feminine singular.

(17a) *kunt* *ʿand-i* *djāja*
was-I at-me chicken
'I had a chicken'
(17b) *faṭima* *kant* *ʿend-a* *le-ktuba*
Fatima was–3fs at-her the-books
'Fatima had the books'

These two dialects also distinguish between gender and number agreement. In (18a–b), the Tunis dialect, the verb *kān* may agree in gender with a feminine singular possessor (theme) but may not agree in number with a plural possessor (theme), and in (18c), the Meknes dialect, agreement in number is allowed.

(18a) *kān-it* *ʿand-ī* *djāja*
was–3fs at-me chicken
'I had a chicken'

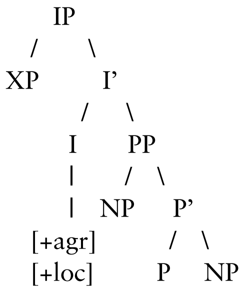
(18b)	**kān-u*	*ʿand-ī*	*wlād*	
	was–3p	at-me	boys	
(18c)	*faṭima*	*kan-u*	*ʿend-ha*	*le-ktuba*
	Fatima	was–3p	at-her	the-books
	‘Fatima had the books’			

6. Approaches and analyses

6.1 *Formal/syntax-based*

The literature on Arabic syntax does not explicitly address the derivation of these structures, perhaps because their syntax involves operations independently motivated in the literature. One proposal (Eid 1993) addresses the structure and derivation of locatives and possessives within the context of prepositional predicates and negation. The derivation proceeds in much the same way as suggested by Freeze (1992). The underlying structure from which all three constructions (locatives, existentials, and possessives) are derived is given in (19) and is based on Chomsky’s (1986) Barriers model. The subject position (Spec of IP) is empty and is not assigned a theta role; the predicate phrase is prepositional (PP). The theme argument is the Specifier of the predicate phrase and the location is the complement.

(19)



The analysis is reminiscent of the predicate-internal subject analysis; it essentially extends it to say that certain predicates contain all their arguments, i.e. the PP predicate phrase is a complete functional complex (CFC) in the sense of Chomsky (1985:168ff.). The copula arises in INFL and consists of morphological and syntactic features which, in the locative paradigm, include a locative feature and is given phonological form at the Phonological Form (PF).

The movement of the theme or the location to the subject position is governed by the [+/-definite] feature of the theme: a definite theme may move to the subject position, yielding the predicate locative. Alternatively, an indefinite theme may stay in place while the location moves to the subject position, yielding the existential. Given the arguments Theme and Location within a predicate PP, Predicate Locatives are the result of the theme moving to [Spec, IP], and Existentials, including Possessives, the result of locative phrase movement to [Spec, IP] (Freeze 1992:558–559).

According to this analysis, the locatives in (4a) and (7a) are the result of the movement of the definite theme *al-kitāb* to Spec of IP. The existentials in (5b) and (8b) are derived by moving the locative phrase – *ʿala l-maktab*, *ʿand sāmi* and their Standard Arabic equivalents – to Spec of IP. In (5c) and (8c), no movement takes place, and the pro-form is lexicalized as the locative feature in INFL, *hunāka* in Standard Arabic, and *fī* in Egyptian Arabic. The pro-form, being a feature in INFL, is simply a spell-out of that feature and has no syntactic relevance. As part of INFL, the pro-form is not an argument and cannot occur in argument position, contrary to some analyses of English where the pro-form *there* is placed in subject position.

6.2 *Discourse-based analyses*

Discrepancies have been reported in the literature on Arabic between (morpho)syntactic conditions for the realization of (in)definites, such as those discussed above, and results obtained from discourse-based analyses and pragmatic accounts. These typically involve notions of topicality and given/new information as explanations for the (in)definite effects. In concluding this entry, two such studies are briefly reviewed in order to provide as full a picture as possible of perspectives available in the literature on the subject.

Belyayeva (1997) proposes to account for certain realizations of definiteness in Palestinian Arabic discourse structure by adopting the Givenness Hierarchy (Gundel a.o. 1993), which utilizes six implicationally related cognitive statuses to explain the use of referring expressions in discourse, associating degree of topicality with cognitive status on the Givenness Hierarchy (20).

(20) Givenness Hierarchy
in focus > activated > familiar >
{pronouns} {that,this} {that N}
uniquely indentifiable > referential >
{the N} {indefinite this N}
type identifiable
{a N}

Belyayeva examines cases where nouns are not morphologically marked for definiteness but occur in environments associated with definiteness effects and from which indefinites are excluded, as, for example, modified indefinite nouns that occur in strictly definite environment (e.g. *walad zġīr ʾakal teffāḥa* ‘a small boy ate an apple’). The modification of the indefinite noun gives it a higher status on the hierarchy than nonidentifiable nouns, which as ‘type-identifiable’ have the lowest status on the hierarchy. Thus, the higher cognitive status acquired by a noun allows that noun to appear in definite environments.

Mughazy (2006) provides perhaps the only detailed discussion of existentials and (in)definiteness in Arabic from a pragmatic information structure perspective. He takes issue with analyses that treat indefiniteness as a prerequisite for grammaticality of the subjects in existential sentences and with pragmatic explanations that justify them. Such analyses, he argues, propose to explain the definiteness effect by arguing that only topical NPs occupy sentence-initial position because their referents are familiar to the addressee and represent old/given information. Since the use of indefinite NPs presupposes lack of familiarity with their referents (i.e., they present new information), they are usually unacceptable sentence-initially. To avoid sentence-initial indefiniteness, the argument continues, the existential marker is used.

Mughazy shows, on the basis of a corpus of Egyptian Arabic, that definite NPs occur in existential sentences with *fī*, a context ruled out by the syntax. See earlier discussion and the Egyptian Arabic paradigm in (3)–(5). His data include the examples in (21)–(23), where a definite NP occurs with the existential marker *fīh*: *il-badīl* ‘the alternative’, *il-muḥāmi* ‘the lawyer’, and *id-doktōr galāl* ‘Dr. Galal’. Mughazy’s transcription of the existential marker is retained in the examples.

(21) A: *ʿandu-ku id-dawa da?* (to a pharmacist)
‘Do you have this medicine?’
B: *lā wa-llāhi xalaṣ bass fīh il-badīl il-mistawrad*
‘I am afraid it is sold out, but there is the imported alternative’

(22) A: *ʿamalt ʾēh maʿa l-muḥāmi?*
‘What did you do with the lawyer?’
B: *kullu tamām, bass lissa fīh il-muškila iyyāha*
‘Everything is fine, but there is still that problem’

(23) A: *matiʿrafš mīn fi l-lagna di?*
‘Do you know who is on this committee?’
B: *fīh id-doktōr galāl w-id-doktōra samīḥa*
‘Dr. Galal is on it, and Dr. Samiha’

The felicity of existentials in Egyptian Arabic, Mughazy concludes, is not dependent on the indefiniteness or the hearer-new status of existential NPs, since they can be felicitous even if these NPs are definite and/or hearer-new, as in (21)–(23). He proposes that an existential sentence is felicitous only if it is informative, i.e. if the integration of the existential and the relevance propositions provides propositional information assumed to be new to the addressee. The existential proposition asserts or presupposes the existence of the NP referent, hence introduces it to the current universe of discourse. The relevance proposition relates that referent to the addressee’s background knowledge. This analysis, as Mughazy argues, differs from accounts based on hearer-new status in that it assumes information structure to be concerned with propositions rather than NPs or their referents.

Bibliographical references

Anwar, Mohamed. 1979. *“Be” and equational sentences in Egyptian Colloquial Arabic*. Amsterdam and Philadelphia: J. Benjamins.

Bach, Emmon. 1967. “*Have* and *be* in English syntax”. *Language* 43.2:462–485.

Bahloul, Maher. 1993. “The copula in Modern Standard Arabic”. *Perspectives on Arabic Linguistics*, V, ed. Mushira Eid and Clive Holes, 209–230. Amsterdam and Philadelphia: J. Benjamins.

Belyayeva, Dina. 1997. “Definiteness realization and function in Palestinian Arabic”. *Perspectives on Arabic linguistics*, X, ed. Mushira Eid and Robert Ratcliffe, 47–67. Amsterdam and Philadelphia: J. Benjamins.

Benmamoun, Elabbas. 2000. *The feature structure of functional categories*. Oxford: Oxford University Press.

Birner, Betty and Gregory Warad. 1998. *Information status and noncanonical word order in English*. Amsterdam and Philadelphia: J. Benjamins.

Chomsky, Noam. 1985. *Barriers*. Cambridge, Mass.: MIT Press.

Clark, Eve V. 1978. "Locationals: Existential, locative, and possessive constructions". *Universals of human language*. IV. *Syntax*, ed. Joseph Greenberg, 85–126. Stanford, Cal.: Stanford University Press.

Comrie, Bernard. 1991. "On the importance of Arabic for general linguistics". *Perspectives on Arabic linguistics*, III, ed. Bernard Comrie and Mushira Eid, 3–30. Amsterdam and Philadelphia: J. Benjamins.

Eid, Mushira. 1991. "Verbless sentences in Arabic and Hebrew". *Perspectives on Arabic linguistics*, III, ed. Bernard Comrie and Mushira Eid, 31–61. Amsterdam and Philadelphia: J. Benjamins.

——. 1993. "Predicates and their heads". *Principles and prediction: The analysis of natural language*, ed. Mushira Eid and Gregory Iverson, 135–152. Amsterdam and Philadelphia: J. Benjamins.

——. 2004. "Locatives and existentials". Paper presented at the *19th Annual Symposium on Arabic Linguistics*, 12 March 2004, University of Oklahoma, Norman.

Freeze, Ray. 1992. "Existentials and other locatives". *Language* 68:3.553–595.

Gundel, Jeannette K., Nancy Hedberg, and Ron Zacharski. 1993. "Cognitive status and the form of referring expressions in discourse". *Language* 69.247–307.

Hawkins, John. 1978. *Definiteness and indefiniteness: A study in reference and grammaticality prediction*. Atlantic Highlands, N.J.: Humanities Press.

Hoekstra, Teun and Rene Mulder. 1990. "Unergatives as copular verbs: Locational and existential predication". *Linguistic Review* 7.1–79.

Jelinek, Eloise. 1981. *On defining categories: AUX and predicate in Colloquial Egyptian Arabic*. Ph.D. diss., University of Arizona.

Lyons, Christopher. 1998. *Definiteness*. Cambridge and New York: Cambridge University Press.

Matthews, Peter. 1997. *The concise Oxford dictionary of linguistics*. Oxford: Oxford University Press.

Mohammad, Mohammad A. 2000. *Word order, agreement, and pronominalization in Standard and Palestinian Arabic*. Amsterdam and Philadelphia: J. Benjamins.

Mughazy, Mustafa. 2006. "The information structure of existential sentences in Egyptian Arabic". Paper presented at the *Conference on Communication and Information Structure in Spoken Arabic*, 8–10 June 2006, University of Maryland, College Park.

Reuland, Eric J. and Alice G.B. ter Meulen. 1989. *The representation of (in)definiteness*. Cambridge, Mass.: MIT Press.

Stassen, Leon. 1995. "Predicative possession". *The world atlas of language structures*, ed. Martin Haspelmath, Matthew S. Dryer, David Gil, and Bernard Comrie, 474–477. Oxford and New York: Oxford University Press.

Ward, Gregory and Betty Birner. 1995. "Definiteness and the English existential". *Language* 71.722–742.

Mushira Eid (University of Utah)

Loss → Language loss

Lowering → Vowel Backing

Luġa

1. The meaning of the term *luġa*

In the Arabic grammatical tradition, the term *luġa* (pl. *luġāt*) means (i) 'dialect', (ii) '(dialectal) word', (iii) 'word in a dictionary', and hence (iv) 'lexicography', and finally (v) 'language'. The term *luġa* was used in this latter meaning in the phrase → *waḍʿ al-luġa* 'the conventional nature of language' in speculations about the relationship between names and designation (see Versteegh 1987:168; Goldziher 1994:38–44; for a more detailed study, see Weiss 1974). The original meaning of the word may have been 'the way people [not in our tribe] speak', i.e. similar to *lahja* 'way of speaking', then later also 'dialect'. In Modern Standard Arabic, the word has come to mean '(foreign) language', for which in Classical Arabic the term → *lisān* was used more often.

Arabia before Islam can be divided into two environments, nearly independent from each other: the sedentary population in Mecca, Yathrib, the cities of Greater Yemen, Ḥīra in southern Iraq, and the Ġassānid cities in southern Syria; and the environment of the roaming Bedouin, who cannot be connected to fixed places. The different parts and groups of Arabia were not completely separated from each other, but they were separated enough to allow for the development of different tribal dialects. The important meeting places of the different tribal groups were the Meccan pilgrimage and the commercial marketplaces. These offered good opportunities for rivalry and debate among poets and reciters of various tribes, not only about the excellence of their poems but also about the superiority of their tribal dialect (aš-Šalqānī 1977:102–103). Rabin (1951:1) supposes that the two main groups of tribes may have had different origins in the family of Semitic languages, but this view is not shared by many scholars.

The oldest Arabic dictionary, the *Kitāb al-ʿayn*, attributed to al-Xalīl ibn ʾAḥmad al-Farāhīdī (d. 170/786), contains the following explanation: "*Luġa* is the difference of speech with the same meaning. The [verb] *laġā* means mingling of the speech with invalidity. In the *ḥadīṯ* we find: He who says on Friday: 'Harken!' has spoken (*laġā ʾay takallama*) [in Buxārī, *Jumʿa*: "If you tell your companion on Friday: 'Listen!', while the imam is still preaching, *laġawta*, i.e. 'you have spoken'"].... 'I deleted (*ʾalġaytu*) this word', i.e. 'I found it invalid and superfluous in speech, and redundant'" (*ʿAyn* IV, 449).

Ibn Jinnī (d. 392/1002) explains the meaning of the root *l-ġ-w* as follows: "As for its morphology and the knowledge of its consonants, it is of the pattern *fuʿla* from the [verb] *laġawtu*, that is, 'I spoke'. The origin of *luġa* is *luġwa*, similar to words of the same type, like *kura, qula, ṯuba*, each of which has *wāw* in the position of the *lām* [i.e. the third radical]. They say: *luġāt* – *luġūna*, like *kurāt* – *kurūna*. It is also said: *laġiya/yalġā* in the meaning 'to talk irrationally, to rave', its *maṣdar* being *al-laġā*... and also *al-laġw*. We find in the *Qurʾān* [25/72] 'And those who encounter foolish talk (*laġw*), encounter it with dignity'" (Ibn Jinnī, *Xaṣāʾiṣ* I, 33–34).

Ibn Manẓūr (d. 711/1311) in the *Lisān al-ʿArab* defines the verbal stem of *luġa* with the following expression: *laġā ʿan aṣ-ṣawāb wa-ʿan aṭ-ṭarīq* 'if he deviates from the right thing and the [right] way'. Then he quotes Ibn al-ʾAʿrābī, saying: "The word *luġa* is taken from this meaning because some people speak in a way (*kalām*) that deviates from the way (*luġa*) of other people" (*Lisān al-ʿArab* s.v. *l-ġ-w*).

The word *luġa* does not occur in pre-Islamic poetry, perhaps because its syllabic structure is not appropriate for the meter. Its first two occurrences are from the Umayyad period, in the *lāmiyya* of ʾAbū Dihbil al-Jumaḥī (d. 62/682) *hāḏā luġatun ʾunkiruhā* 'this is a dialect [or: way of speech] I dislike'; and in the *lāmiyya* of al-Kumayt ibn Zayd al-ʾAsadī (d. 126/743) *lahum luġatun tubayyinu man ʾabūhum* 'they possess a dialect which shows who their father is'. In the Abbasid age, the word occurs frequently, as well as in scientific prose.

Although *luġa* does not occur in the *Qurʾān*, different words from the root *l-ġ-w* occur eleven times in the forms *ilġaw, al-laġwu/a, laġwan, lāġiyatan*, all with a negative connotation. In Q. 5/89 it has the meaning of 'unintentional speech [oath]'.

2. SOURCES OF *LUĠA*

The basic sources of the Classical Arabic dialects (*luġāt*) are the following: (i) the variant readings (→ *qirāʾāt*) of the *Qurʾān*; (ii) the material recited by the 'Arabs' of the desert (in reality, by the specialist reciters, *rāwī*), mainly but not exclusively poetry; (iii) the Arabic proverbs (*ʾamṯāl al-ʿArab*). As a fourth source, the evidence from modern Arabic dialects may be included here, since according to many scholars (al-Jundī 1983:101–103, 213–233), these are the direct descendants of the Classical Arabic dialects. Although others hold that there is no link between the Classical Arabic and the modern dialects, Goldziher (1994:13) has noted that the data about the dialect of Tamīm show that this dialect already exhibited features later found in the colloquial Arabic of the 19th century.

In the 8th century, many longer or shorter treatises were compiled with the title *Kitāb al-luġāt fī l-Qurʾān* 'The book of the dialects in the *Qurʾān*'. Rabin (1951:6) mentions eleven works with this title, of which only the titles are known, with the only exception of ʾAbū ʿUbayd's (d. 224/838) *Risāla fīmā warada fī l-Qurʾān min luġāt al-qabāʾil* 'Epistle on what appears in the *Qurʾān* from the tribal dialects'.

The earliest extant Qurʾānic exegeses of the 8th century mention *luġa* or its plural *luġāt* (Versteegh 1990, 1993). Muqātil (d. 150/767), for instance, refers in his *Tafsīr* to the specific tribal dialects in order to explain the meaning of obscure words, or to confirm that a particular word is part of the Arabic language, e.g. the dialects of Ṭayyiʾ, ʾAzd Šanūʾa, ʿUmān, Qurayš, Banū Ġaṭīf, and Banū Mālik ibn Kināna (Versteegh 1990:217, 1993:91). Muḥammad al-Kalbī (d. 146/763; *Tafsīr* 131b27–28) refers to the *luġa* of Bal Ḥāriṯ ibn Kaʿb to explain the phrase *ʾinna hāḏāni* 'Verily, these two' (Q.

20/63), which he contrasts with the standard form *ʾinna hāḏayni* (Versteegh 1993:127). Versteegh (1993:99) believes that *luġa* may have been used here as a general term for pre-Islamic dialects, since elsewhere al-Kalbī refers to a special tribal dialect with the word → *lisān*.

3. *Luġa* in grammatical literature

As a rule, early Arab grammarians were not interested in preserving and describing dialectal material. They dealt with these data only when they were forced to do so because their informants accepted them. ʾAbū ʿAmr ibn al-ʿAlāʾ admits this freely (Ibn Fāris, *Ṣāḥibī* 49; also in Suyūṭī, *Muzhir* I, 260): "Once ʾAbū ʿAmr ibn al-ʿAlāʾ was asked about *ʿarabiyya*, whether they [the grammarians] put everything in it that they collected from the speech of the Arabs. He said 'no'. Answering a further question about how he decided what to choose from among a set of contradictory data, he said: 'I choose the most frequent ones (*al-ʾakṯar*) and I call the others dialect (*luġa*)'".

In the first Arabic grammar, Sībawayhi's *Kitāb*, the term *luġa* occurs 136 times (its plural 17 times), according to Troupeau (1976), who translates it as "manière de parler". As a rule, Sībawayhi uses the term *luġa* in cases where he cannot avoid mentioning variants to the forms preferred by him and his circle, because they are accepted by his informants or occur in well-known lines of poetry. He says, for instance, that there is an assimilation (*niʿimmā* < *niʿima mā*) in the dialect (*luġa*) of those who say *niʿima*, but not in the dialect of those who say *niʿma*. This is *luġa Huḏayl*. Or, with regard to sound variation, he speaks of "the velarized *ʾalif*, i.e. in the dialect of the people of the Ḥijāz" (*ʾalif at-tafxīm yaʿnī bi-luġa ʾahl al-Ḥijāz*; *Kitāb* II, 452; cf. Semaan 1968:40). He rejects, however, sound variation when it is not accepted by reliable speakers: "[The number of the consonants] is 42, when those that are not considered correct are included; not many of these consonants are used in the dialect (*luġa*) of those whose Arabic (*ʿarabiyya*) is approved of. These consonants are not considered correct in Qurʾānic reading, nor in poetry" (*Kitāb* II, 452). In many places, however, he does not use the word *luġa* but simply says, for instance, that the people of Mecca, according to some opinions, do not differentiate the two *tāʾ*s in words like *tatanājaw* (*Kitāb* II, 457). Elsewhere, he says: "All [imperfect verbs] that I have mentioned have the vowel *a* (*maftūḥ*) in the dialect (*luġa*) of the people of the Ḥijāz', and this is the basic rule (*ʾaṣl*). The basic rule (*ʾaṣl*) in *faʿiltu* is to have *-a-* [in the imperfect]: *yaftaḥ* according to the dialect (*luġa*) of the people of the Ḥijāz". Later, however, he omits the word *luġa*: "The people of the Ḥijāz say *yawjalu*" (*Kitāb* II, 276). Other dialects are not considered good enough to be considered correct by the grammarians. Sībawayhi says that the *kasra* in *yifʿal* is irregular and isolated (*šāḏḏ*; *Kitāb* II, 276), although we know from other sources that it was in use in a large part of Arabia. This bias, however, is understandable, since Sībawayhi and the other grammarians aimed at the description of a unified Arabic language. Once they decided on the 'basic' nature of the Ḥijāzī dialects, they had to regard the others as irregular, rare, or dialectal, and not acceptable as the point of departure for analogy. In connection with *kam*, Sībawayhi (*Kitāb* I, 253–254) states that "the first explanation for it is stronger, because it is not traced back to poetic license (*iḍṭirār*) or irregular (*šāḏḏ*) linguistic forms, and constitutes a good way [of explaining] (*wajh*) it". In another place (*Kitāb* I, 253), he says: "There are those among the Arabs [Bedouin], who put it in the accusative (*naṣb*)". These last two remarks refer to different tribal dialects – without calling them *luġa* – which are not rejected, although they are not as prestigious as others. Since Sībawayhi did not aim at describing different tribal dialects, he prefers the phrase "some Arabs" (*baʿḍ al-ʿArab*), sometimes meaning by it a whole tribe, sometimes only individuals. Speakers changing the *ʾiʿrāb* system are not recognized as speaking any truly Arabic dialect, but speak 'erroneously': "Some people from among the Arabs make errors and say *ʾinnahum ʾajmaʿūna ḏāhibūna* instead of *ʾajmaʿīna*" (*Kitāb* I, 250). Thus, we can only know from external evidence whether Sībawayhi meant by 'some Arabs' a whole tribe or only individual speakers. However, when a similar phenomenon (*aṣ-ṣābiʾūna*) occurs in the *Qurʾān*, he explains it by a syntactic reinterpretation, because no grammarian could possibly say that the *Qurʾān* contains faulty forms. Even though the word *luġa* occurs very often in Sībawayhi's *Kitāb*, it could have occurred much more frequently, considering the significance and quantity of the Arabic dialectal material preserved in later books. Sometimes Sībawayhi refers to individual choices that are easily explained by grammatical means, rather than to tribal differences. Speaking about the habit of reducing the feminine ending of nouns to *-a(h)*, he says: "Some people from among the Arabs maintain the [feminine ending] *hāʾ* and say *yā salamatu* and some say *yā salamata*. Those who elide (*ḥaḏf*) [this ending] in juncture, say in pause (*waqf*) *yā salamah*. It is a compulsory *hāʾ*, which cannot be omitted, since its function is to preserve the vowel (*ḥaraka*) *a* after the *mīm*" (*Kitāb* I, 287–288).

Al-Farrāʾ (d. 207/822) quotes the words of the Prophet's wife ʿĀʾiša on *laġw*, a close cognate of *luġa*, in a negative sense: "Verily, *laġw* is what runs in speech (*kalām*) without conclusion (*ʿaqd*), although it is similar to the speech of the Arabs" (*Maʿānī l-Qurʾān* I, 144). He also uses the word *luġa* many times in connection with various Qurʾānic readings, for instance with regard to Q. 2/235: "The Arabs [Bedouin] have in *ʾaknantu š-šayʾ* 'I concealed it' two dialectal forms (*luġatāni*): *kanantuhu wa-ʾaknantuhu*" (cf. Sallūm 1987:397 *k-n-n*; see also al-Jundī 1983:495). Then, he adds that the same is found in poetry, quoting two versions of a line: *takunnu, tukinnu* (al-Farrāʾ, *Maʿānī* I, 152). The use of Form IV instead of Form I must have been a widespread phenomenon among different tribal groups, so that understandably it occupied a significant place in early grammatical literature, and is reflected even in modern dictionaries. Like Sībawayhi, al-Farrāʾ does not always mention the word *luġa* in connection with variant readings, but sometimes only defines the geographical place where a specific linguistic phenomenon occurs in a reading. He says, for instance, that instead of *yaxṭafu* some readers of Medina have *yaxṭṭifu*, while in other (unspecified) places it is read as *yaxiṭṭifu, yaxaṭṭifu* (*Maʿānī* I, 18.1). These readings, however, are not explained by him as dialectal forms. Instead, he thinks that there are special linguistic subrules replacing the vowels of *ʾiʿrāb* according to the speaker's preference. In this, he follows Sībawayhi's approach to linguistic variants. All grammarians, even those from Kufa, who were said to be more lenient toward irregular forms, are apt to consider dialectal forms only if they can be explained by their linguistic rules: "In every *yāʾ* referring to the speaker [i.e. the suffixed pronoun of the 1st person singular], there are two [possible] linguistic forms (*luġatāni*): release [of a vowel, *ʾirsāl*] and vowellessness (*sukūn*), i.e. *niʿmatiya* and *niʿmatī*'" (*Maʿānī* I, 29, explaining Q. 2/40: *yā Banī ʾIsrāʾīla uḏkurū niʿmatiya llatī ʾanʿamtu ʿalaykum*). If this *yāʾ* was followed by the article (*al*), the Arabs favored that form (*luġa*) in which the *yāʾ* was supplied with a vowel (*-iya*) and rejected the other (*luġa*). It can be seen from this explanation that al-Farrāʾ's way of thinking strictly follows that of his Basran colleagues in that like them he does not accept tribal dialects and their variant forms in their own right but only inasmuch as they can be explained away by grammatical rules (*qiyās*). Therefore, he considers *luġa* in many cases only as a 'context-dependent variant'. On Q. 2/61 *wa-fūmihā* he says: "The word *fūm* is an ancient [dialectal] form (*luġa*).... The Arabs sometimes change the *fāʾ* into *ṯāʾ* and the other way round, ... and I heard many times speakers of the ʾAsad tribe use *al-maġāṯīr* instead of *al-maġāfīr*" (*Maʿānī* I, 41). In this way, *luġa* may have come to mean 'lexical entity'. On *luġa* in al-Farrāʾ's *Maʿānī*, see ʿAbd al-Karīm (1986); on the opinion of the Kufan grammarians concerning *luġa* in general, see Š. ar-Rājiḥī (2002).

Ibn Fāris (d. 395/1005) gives a list of some of the most conspicuous differences between the Arabic dialects (*Fiqh* 19–22; cf. Goldziher 1994:16–17): (i) *yafʿal* in the dialects of the Qurayš and the ʾAsad, *yifʿal* elsewhere; (ii) two open short syllables vs. one syllable with a consonantal ending: *maʿakum* vs. *maʿkum, kalima* vs. *kilma*, etc.; (iii) → *ʾibdāl al-ḥurūf*, the substitution of one consonant with another, e.g. *ʾanna* > *ʿanna* 'that [conj.]'; (iv) *talyīn al-hamza* 'softening of the *hamza*', i.e. using the so-called 'soft consonants', *ʾalif, wāw, yāʾ*, instead of the → *hamza*: *mustahzūna* < *mustahziʾūna* 'mockers'; (v) *qalb al-ḥurūf* '→ metathesis': *ṣāʿiqa* > *ṣāqiʿa* 'thunderbolt'; (vi) → *ḥaḏf al-ḥurūf* 'elision of consonants': *istaḥyaytu* > *istaḥaytu* 'I felt ashamed'; (vii) 'weakening of a strong consonant', i.e. the use of a *muʿtall* 'weak' consonant instead of a *ṣaḥīḥ* 'strong' consonant (→ *ʿilla*): *ʾammā* > *ʾaymā* 'as for'; (viii) → *ʾimāla* 'inclination [of the *ā* toward *ē* and *ī*]', and its opposite, → *tafxīm* 'intensification, emphatization' of the pronunciation of a long vowel: *qaḍā* > *qaḍē* 'he spent'; (ix) *taʾnīṯ* 'making a word feminine' instead of *taḏkīr* 'making a word masculine' and vice versa: *baqar* 'oxen, cows' and *naxl* 'palm trees' are either feminine or masculine in the dialects; (x) → *ʾidġām* 'assimilation' or its absence: *muhtadūna* > *muhaddūna* 'the rightly guided ones'; (xi) differences in the *ʾiʿrāb*: the negation *mā* may govern the *naṣb* 'accusative', instead of the more common *rafʿ* 'nominative': *mā zaydun qāʾiman* instead of *qāʾimun*; (xii) the differences in the *jamʿ at-takṯīr* 'broken plural' of a noun: the plural of *ʾasīr* 'prisoner' may be *ʾusrā* or *ʾusarā*; (xiii) *ixtilās* 'slurring, omission' vs. *taḥqīq* '[full] realization' of a short vowel: *yaʾmurukum* > *yaʾmurkum* 'he orders you [pl.]'; (xiv) the use of *-ah* (*hāʾ marbūṭa*) instead of *-at* (*tāʾ marbūṭa*) as feminine ending: *ʾama(h)* > *ʾamat(un)* 'slave girl'; (xv) the use of a long vowel instead of a short one (*taḥwīl al-ḥaraka ʾilā ḥarf*): *unḏ̣ur* > *unḏ̣ūr* 'look!'; (xvi) difference or even opposition in the meaning of words: *waṯaba* means 'to sit down', but in the → Ḥimyaritic dialect it means 'to jump'.

With the collection and transmission of the *qirāʾāt* of various *qurrāʾ* during the 8th and 9th centuries, some explanation was needed for the presence of apparently dialectal material in them. The mysterious expression that "the *Qurʾān* was revealed according to seven 'letters' (*ʾaḥruf*)" was sometimes interpreted as referring to the Arabic dialects. Ibn ʿAbbās related that the Prophet Muḥammad had said: "Gabriel recited to me [the Qurʾānic verse] according to one dialect (*ḥarf*), then I repeated it. But I did not cease to ask for more [dialects] and he gave more to me, until he reached the number of seven dialects (*sabʿa ʾaḥruf*)" (Ibn Ḥajar, *Fatḥ al-bārī* IX, 19; also mentioned by Ibn al-Jazarī, *Našr* I, 19). The books on the *maʿānī l-Qurʾān* are full of so-called irregular readings (*qirāʾāt šāḏḏa*), but their authors defend them as pertaining to the dialects of those Arabs who spoke correct, pure Arabic (*fuṣaḥāʾ*) but for one reason or another could not attain the same rank as others. Thus, their special linguistic usage was called 'rare' or 'irregular' (Ibn Jinnī, *Muḥtasib* I, 3; cf. al-Jundī 1983:107–108). ʾAbū Ḥayyān says: "The readings have come according to the Arabic dialects (*luġa*), the regular ones and the irregular ones as well" (*Baḥr* VIII, 493). In the treatise attributed to Ibn ʿAbbās, *al-Luġāt fī l-Qurʾān*, the names of 36 tribes are mentioned in this connection (Sībawayhi, *Kitāb*, 16 tribes; Sīrāfī, *Šarḥ Kitāb Sībawayhi*, 25 tribes; ʾAbū Ḥayyān, *Baḥr*, 64 Arabic dialects; see al-Jundī 1983:112–113).

There were and still are, however, great differences of opinion on how the word → *ḥarf* should be interpreted: as *luġa*, or *qirāʾa*, or yet something else. Ibn Qutayba (d. 276/889) says: "When you say: someone reads [the *Qurʾān*] according to the 'letter' (*ḥarf*) of ʾAbū ʿAmr or ʿĀṣim, this does not mean that there is a letter in the *Qurʾān* which can be read in seven different ways (*ʾawjuh*). The explanation of the *ḥadīṯ* that the *Qurʾān* was revealed according to seven 'letters' is that it was revealed in seven different dialects (*luġāt*), scattered in the Book. As the Prophet said: 'Read [it] as you like!'" (*Taʾwīl* 34). Then, Ibn Qutayba continues: "The meaning of *ḥarf* is 'all the letters in a word, the whole word or the whole speech [act]'" (*Taʾwīl* 35). Ibn Qutayba says explicitly: "It was an alleviation for His servants that He ordered His Prophet that every tribe should read [the *Qurʾān*] in its own dialect (*luġa*) and according to their customs. So the Huḏalī reads *ʿattā* instead of *ḥattā*, since he speaks in this way, and the ʾAsadī reads *tiʿlamūna, tiʿlam, tiswaddu wujūhun*, and the Tamīmī uses *hamza*, while the Qurašī does not" (*Taʾwīl* 39ff.; also Ibn al-Jazarī, *Našr* I, 22). Elsewhere, Ibn Qutayba, commenting on the different tribal readings, states: "Had every tribal group from among these [tribes] wanted to abandon its dialect (*luġa*) ... it would have become unbearable for them and it would have led to suffering among them" (*Taʾwīl* 39). This, however, is correct only in connection with the so-called accepted dialects, which are labeled by the grammarians *muṭṭarid* 'in widespread, general usage'. In the chapter in the *Taʾwīl* (50–64) about the accusation that the *Qurʾān* contains → *laḥn* 'faulty speech', Ibn Qutayba tries to prove that the alleged faulty speech is either accepted usage, which can be explained within grammatical theory, or it consists of errors made by copyists or readers (*qurrāʾ*). Thus, Ibn Qutayba did not connect the seven 'letters' with the seven readings, and did not, as a rule, accept their 'irregular' linguistic forms (*šāḏḏ*, pl. *šawāḏḏ*, the opposite of *muṭṭarid*) as correct Arabic, even when later sources reported their use by large tribes or groups of tribes.

Similar views about the presence of traces of tribal dialects in the *Qurʾān* are presented later by scholars like aṭ-Ṭabarī: "If they had to deviate from their dialect (*luġa*), and to depart from their tongue (*lisān*), that would have been such a burden upon them that they could not have borne it" (aṭ-Ṭabarī, *Tafsīr* I, 22). He adds that the Prophet asked his Lord to reveal the *Qurʾān* in more Arabic dialects as a relief (*taxfīfan*). ʾAbū Šāma (d. 665/1266–1267) says: "The Arabic *Qurʾān* contains all the dialects (*luġāt*) of the Arabs, because it has been revealed for all of them. So it was allowed to them to recite it in their own dialects. This is the cause of the differences between the Qurʾānic readings" (*ʾIbrāz* 487). For a comprehensive overview of the dialects in the Qurʾānic readings, see A. ar-Rājiḥī (1999).

4. MODERN VIEWS ON *LUĠA*

Nöldeke (1860) was the first to deal with the question of the *luġāt* vs. the *ʿarabiyya* (→ pre-Islamic Arabic). He rejected the customary Arab explanation of the *sabʿa ʾaḥruf* phrase as different dialects in the Qurʾānic text or in the variant readings, but took into account only the meanings 'letter' and 'way of reading' of → *ḥarf*, ignoring the meaning 'word'. Nöldeke (1961:1, 58–59) believed that *luġa* rarely replaced *ḥarf* in the tradition, and Nöldeke was the initiator (1904:2, 1910:4) of the traditional view in Western scholarship that the Arabic language was a unified, strongly standardized form of speech. Neither Nöldeke, however, nor any other scholar who accepted his views, did any research into the large amount of dialectal data. Their main argument, the mutual comprehensibility of the tribal poetry, is based on the mistaken view that tribal dialects were separate languages.

Goldziher (1994:10–18) was the first European scholar to deal with Classical Arabic dialects (*luġāt*) in a detailed way, giving a brief overview of the question on the basis of some medieval Arab authors. He states that the Arabs considered the dialect of the Meccan tribe of Qurayš, the tribe of Muḥammad, in which the *Qurʾān* had been revealed, the purest (*ʾafṣaḥ*) dialect, and that Arabic grammar was nothing else but the codification of this dialect. According to him, the medieval editors of poetry standardized the language, but they were corrected by commentators who adduced variant readings, which revealed dialectal peculiarities. According to Goldziher, lexicography also helped to preserve the fragments of Arabic dialects. Although he, like Nöldeke, did not believe that the most prominent Arabic dialects are represented in the *Qurʾān*, he recognized the significance of this theory for the efforts of collecting the dialects and assumed that the *sabʿa ʾaḥruf* tradition helped to preserve dialect material. Goldziher abundantly quotes from such late works as as-Suyūṭī's *ʾItqān* and *Muzhir*; Ibn Fāris' *Fiqh al-luġa*, which contains a vast corpus of data about the *luġāt*, was known to him only through the *Muzhir*. Although these works inform us about such characteristic dialectal features as → *kaškaša*, they helped to create a false image of the early linguistic situation, contrasting these dialectal features with an allegedly unified Arabic language. Goldziher quotes a long paragraph from the work of Ibn Fāris, listing the main differences between Arabic dialects of the 6th and 7th centuries.

Vollers (1906) made an important step toward a better understanding of the linguistic situation in the pre-Islamic period, positing the existence of a 'popular language' on the basis of noncanonical Qurʾānic readings. One of the criticisms of this view was that Vollers had collected his material on the ancient dialects from the *Muzhir* and similar 'late' works. His daring conclusions were not shared by most of the critics. Yet, the book contains many valuable ideas and is an important study of early Arabic. Classical Arabic dialectal research was for a long time hindered by the view that the *luġāt* are irrelevant for a better knowledge of the Classical Arabic language and grammar. Kofler (1940–1942) presents the hitherto fullest collection of data, assembling not only forms attributed to specific dialects but also unspecified material from the whole field of grammar, arranged under grammatical headings, without any attempt at geographical treatment. A further step was made by ʾIbrāhīm ʾAnīs (1946). He produced a critical study based on wide reading, although he was mainly interested in phonetic differences, minimizing the value of grammatical and syntactic data.

The first really significant contribution to the study of the Classical Arabic dialects was made by Rabin (1951). He shares Vollers' (1906) and Sarauw's (1908) hypothesis that the grammarians' schematization of all dialect differences into Ḥijāz and Tamīm was correct, and believes that this schema corresponds to a real division of the pre-Islamic dialects into Eastern and Western groups (Rabin 1951:1). For other dialect groups there is insufficient evidence. Rabin (1951:ix) assumes that the Eastern dialects are the result of comparatively recent linguistic developments and that they are basically the same as the Classical Arabic of the poets. In his view, Classical Arabic is based on one or several of the dialects of Najd, perhaps in an archaic form: “Najd was an area where East-Arabians and West-Arabians met and mingled. In the West of the region the Ġaṭafān and Hawāzin dialects were strongly West-Arabian, in the East those of Ġanī and ʿUqayl clearly Eastern” (Rabin 1951:3). He also holds the view that Western dialects possessed archaic features which set them apart from Classical Arabic. He classifies the following tribes as Western Arabian: Ṭayyiʾ, Saʿd, Ġaṭafān, Huḏayl, Kināna, Xaṯʿam, Maḏḥij, Ḥāriṯ, Ḥamdān, Murād, Ḥimyar, ʾAzd-Šanūʾa; and the following as Eastern Arabian: Bahrāʾ, Kalb, Namīr, Taġlib, Tamīx, Bakr, Ḥanīfa, Tamīm, ʿAbd al-Qays, Ġanī, ʿUqayl, Bāhila, Ḍabba, ʾAzd-ʿUmān. Finally, he states that there is a North-Western subgroup, not considered part of the Western group, consisting of Juhayna and Muzayna.

The most comprehensive and convincing presentation hitherto has been given by al-Jundī (1983). Based on an impressive corpus, he sheds new light on several problems and presents new explanations of many controversial issues, reasoning that both *šawāḏḏ* (so-called ‘irregular linguistic forms’) and *ḍarāʾir aš-šiʿr* → ‘poetic licenses’ are in reality dialectal phenomena not considered part of the standard (‘unified’) Arabic language. According to al-Jundī (1983:103–114), the *Qurʾān* and its variant readings provide the best sources for the investigation of the tribal dialects. This view is nowadays shared by many Arab scholars (see, e.g., ʿAbd al-Karīm 1986; Imam 1998). Al-Jundī regards ʾAbū Ḥayyān’s *Baḥr*, which contains vast material on the dialects, as one of the most important sources for study of the variant readings. This is why he criticizes Rabin for not having perused this voluminous Qurʾānic exegesis. He also rejects the so-called Eastern (Tamīmī) and Western (Ḥijāzī) dialectal division (al-Jundī 1983:55–56). According to him, we cannot base our judgment on this assumption, for the following reasons: (i) the Eastern and Western dialectal division is a vague notion, including units that are too vast and undefined, not being fine-grained enough, and unknown to the Arab geographers; (ii) the Bedouin tribes of Arabia have always been migrating, so that their homelands and centers were not fixed and cannot be incontrovertibly determined; (iii) even if their homeland can be determined with some certainty, the ancient geographical names were not reliable and the Arabian place names have always been used with great laxity. Medina, for example, was sometimes situated in the desert, sometimes in the Ḥijāz, and sometimes it was said to be in the Najd. Contrary to Rabin, he argues convincingly that many more dialects can be known and described than the Ḥijāzī and Tamīmī groups. Al-Jundī also demonstrates that even these two groups cannot be treated as two unified blocks, strictly separated from each other, since there are tribes and subtribes in both of them that exhibit linguistic features similar to those of the other group.

Mostly under the influence of ʾAnīs (1946) and al-Jundī (1965, 1983), a flourishing literature has been published on *luġa* in the latter half of the 20th century, mainly in Egypt, dealing not only with the dialects in general, but with specific dialects as well (see, e.g., Maṭlabī 1978).

Bibliographical references

Primary sources

ʾAbū Ḥayyān, *Baḥr* = ʾAṯīr ad-Dīn ʾAbū Ḥayyān al-ʾAndalusī, *al-Baḥr al-muḥīṭ*. Cairo: Maṭbaʿat as-Saʿāda, 1910.

ʾAbū Šāma, *ʾIbrāz* = ʿAbd ar-Raḥmān ibn ʾIsmāʿīl ʾAbū Šāma, *ʾIbrāz al-maʿānī*. Cairo: al-Ḥalabī, 1930.

Farrāʾ, *Maʿānī* = ʾAbū Zakariyyāʾ Yaḥyā ibn Ziyād al-Farrāʾ, *Maʿānī l-Qurʾān*. Ed. Muḥammad ʿAlī an-Najjār. 3 vols. Cairo: ad-Dār al-Miṣriyya, 1955–1972.

Ibn Fāris, *Ṣāḥibī* = ʾAbū l-Ḥusayn ʾAḥmad Ibn Fāris, *aṣ-Ṣāḥibī fī fiqh al-luġa*. Ed. Moustafa Chouémi [Muṣṭafā Šuʾaymī]. Beirut: A. Badran, 1964.

Ibn Ḥajar, *Fatḥ* = Šihāb ad-Dīn ʾAbū l-Faḍl ʾAḥmad Ibn Ḥajar al-ʿAsqalānī, *Fatḥ al-bāriʾ bi-šarḥ al-Buxārī*. Cairo: al-Maṭbaʿa al-Bahiyya al-Miṣriyya, 1929.

Ibn al-Jazarī, *Našr* = ʾAbū l-Xayr Muḥammad ad-Dimašqī Ibn al-Jazarī, *an-Našr fī l-qirāʾāt al-ʿašr*. Ed. ʿAlī Muḥammad aḍ-Ḍabbāġ. 2 vols. Beirut: Dār al-Kutub al-ʿIlmiyya, n.d.

Ibn Jinnī, *Muḥtasib* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *Kitāb al-muḥtasib* [*Kitāb al-muġtasab*]. Ed. Edgar

Pröbster. Leipzig: J.C. Hinrichs'sche Buchhandlung, 1904. (Repr., Leipzig: Zentralantiquariat der DDR, 1968.)

——. *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. Ed. Muḥammad ʿAlī an-Najjār. 2nd ed. 3 vols. Beirut: Dār al-Hudā, n.d. (1st. ed., Cairo: Dār al-Kutub al-Miṣriyya, 1852–1856.)

Ibn Manẓūr, *Lisān* = ʾAbū l-Faḍl Jamāl ad-Dīn Muḥammad ibn Mukarram Ibn Manẓūr, *Lisān al-ʿArab*. Ed. ʿAbdallāh ʿAlī al-Kabīr, Muḥammad ʾAḥmad Ḥasballāh, and Hāšim Muḥammad aš-Šāḏilī. 6 vols. Cairo: Dār al-Maʿārif, 1981.

Ibn Qutayba, *Taʾwīl* = ʾAbū Muḥammad ʿAbdallāh ibn Muslim Ibn Qutayba ad-Dīnawarī, *Taʾwīl muškil al-Qurʾān*. Ed. ʾAḥmad Ṣaqr. Cairo: Dār at-Turāṯ, 1973.

Muḥammad al-Kalbī, *Tafsīr* = ʾAbū n-Naḍr Muḥammad ibn as-Sāʾib al-Kalbī, *Tafsīr* [*Tafsīr Ibn ʿAbbās*]. Ms. Chester Beatty, no. 4224.

Muqātil, *Tafsīr* = ʾAbū l-Ḥasan Muqātil ibn Sulaymān, *at-Tafsīr*. Ed. ʿAbdallāh Maḥmūd Šaḥāta. 4 vols. Cairo: Hayʾat al-Kitāb al-Miṣrī al-ʿĀmma, 1980–1987.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. Ed. Hartwig Derenbourg. Paris: Imprimerie Nationale, 1881–1889.

Suyūṭī, *Muzhir* = Jalāl ad-Dīn ʿAbd ar-Raḥmān ibn ʾAbī Bakr as-Suyūṭī, *al-Muzhir fī ʿulūm al-luġa wa-ʾanwāʿihā*. Ed. Muḥammad ʾAḥmad Jād al-Mawlā, ʿAlī Muḥammad al-Bajāwī, and Muḥammad ʾAbū l-Faḍl ʾIbrāhīm. 2 vols. Cairo: Dār ʾIḥyāʾ al-Kutub al-ʿArabiyya, n.d.

Ṭabarī, *Tafsīr* = ʾAbū Jaʿfar Muḥammad ibn Jarīr aṭ-Ṭabarī, *Jāmiʿ al-bayān fī tafsīr al-Qurʾān*. Cairo: al-Maṭbaʿa al-ʾAmīriyya, n.d.

Ṯaʿlab, *Faṣīḥ* = ʾAbū l-ʿAbbās ʾAhmad ibn Yaḥyā Ṯaʿlab, *Kitāb al-faṣīḥ*. Ed. ʿĀṭif Madkūr. Cairo: Dār al-Maʿārif, 1984.

Xalīl, *ʿAyn* = ʾAbū ʿAbd ar-Raḥmān al-Xalīl ibn ʾAḥmad al-Farāhīdī, *Kitāb al-ʿayn*. Ed. Mahdī al-Maxzūmī and ʾIbrāhīm as-Sāmarrāʾī. Baghdad: Wizārat aṯ-Ṯaqāfa wa-l-ʾIʿlām, 1981.

Secondary sources

ʿAbd al-Karīm, Ṣubḥī ʿAbd al-Ḥamīd Muḥammad. 1986. *al-Lahajāt al-ʿarabiyya fī* Maʿānī l-Qurʾān *li-l-Farrāʾ*. Cairo: Dār aṭ-Ṭibāʿa al-Muḥammadiyya.

ʾAnīs, ʾIbrāhīm. 1946. *Fī l-lahajāt al-ʿarabiyya*. Cairo: Maktabat al-ʾAnglū al-Miṣriyya.

Goldziher, Ignaz. 1994. *On the history of grammar among the Arabs*. Trans. Kinga Dévényi and Tamás Iványi. Amsterdam and Philadelphia: J. Benjamins. (First published in Hungarian, *A nyelvtudomány története az araboknál*, Budapest: Franklin, 1877.)

Imam, Aḥmad ʿAli al-. 1998. *Variant readings of the* Qurʾan*: A critical study of their historical and linguistic origins*. Herndon, Va: International Institute of Islamic Thought.

Jundī, ʾAḥmad ʿAlam ad-Dīn al-. 1983. *al-Lahajāt al-ʿarabiyya fī t-turāṯ*. 2nd ed. 2 vols. Tripoli (Libya): ad-Dār al-ʿArabiyya li-l-Kitāb. (1st ed., Cairo, 1965.)

Kofler, Hans. 1940–1942. "Reste altarabischer Dialekte". *Wiener Zeitschrift für die Kunde des Morgenlandes* 47.60–130, 233–62; 48.52–88, 247–274; 49.15–30, 234–256.

Levin, Aryeh. "Sībawayhi's attitude to the spoken language". *Jerusalem Studies in Arabic and Islam* 17.204–243.

Maṭlabī, Ġālib Fāḍil al-. 1978. *Lahja Tamīm wa-ʾaṯaruhā fī l-ʿarabiyya al-muwaḥḥada*. Baghdad: Wizārat aṯ-Ṯaqāfa wa-l-Funūn.

Nöldeke, Theodor. 1860. *Geschichte des Qorans*. Göttingen: Verlag der Dieterich'schen Buchhandlung.

——. 1904. *Beiträge zur semitischen Sprachwissenschaft*. Strasbourg: Trübner.

——. 1910. *Neue Beiträge zur semitischen Sprachwissenschaft*. Strasbourg: Trübner.

——. 1961. *Geschichte des Qorans*. 2nd ed. rev. Friedrich Schwally (vols. I, II). Gotthelf Bergsträsser and Otto Pretzl (vol. III) 3 vols. Hildesheim: G. Olms. (Repr. 1970; originally published 1909–1938.)

Rabin, Chaim. 1951. *Ancient West-Arabian*. London: Taylor's Foreign Press.

Rājiḥī, ʿAbduh ar-. 1999. *al-Lahajāt al-ʿarabiyya fī l-qirāʾāt al-qurʾāniyya*. Alexandria: Dār al-Maʿrifa al-Jāmiʿiyya.

Rājiḥī, Šaraf ad-Dīn ʿAlī ar-. 2002. *Fī l-luġa ʿinda l-Kūfīyyīn*. Alexandria: Dār al-Maʿrifa al-Jāmiʿiyya.

Sallūm, Dāwūd. 1987. *al-Muʿjam al-kāmil fī lahajāt al-fuṣḥā*. Beirut: ʿĀlam al-Kutub.

Šalqānī, ʿAbd al-Ḥamīd aš-. 1977. *al-ʾAʿrāb ar-ruwāt: Ṣafaḥāt fī falsafat al-luġa wa-taʾrīxihā*. Cairo: Dār al-Maʿārif.

Sarauw, C. 1908. "Die altarabische Dialektspaltung". *Zeitschrift für Assyriologie* 21.31–49.

Semaan, Khalil I. 1968. *Linguistics in the Middle Ages: Phonetic studies in early Islam*. Leiden: E.J. Brill.

Troupeau, Gérard. 1976. *Lexique-index du* Kitāb *de Sībawayhi*. Paris: C. Klincksieck.

Versteegh, C.H.M. [Kees]. 1987. "Die arabische Sprachwissenschaft". *Grundriss der arabischen Philologie,* II, ed. Wolfdietrich Fischer and Helmut Gätje, 148–176. Wiesbaden: L. Reichert.

——. 1990. "Grammar and exegesis: The origins of Kufan grammar and the *Tafsīr Muqātil*". *Der Islam* 67.206–242.

——. 1993. *Arabic grammar and Qurʾānic exegesis in early Islam*. Leiden: E.J. Brill.

Vollers, Karl. 1906. *Volkssprache und Schriftsprache im alten Arabien*. Strasbourg: Trübner.

Weiss, Bernard G. 1974. "Medieval Muslim discussions of the origin of language". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 124.33–41.

Tamás Iványi (Eötvös Loránd University)

M

Mabnī → Bināʾ

Madagascar → Malagasy

Māḍī and Muḍāriʿ

Māḍī and *muḍāriʿ* are two terms for the verb in its apparent temporal or aspectual modes. They are part of the oldest vocabulary of Arabic grammar, occurring already in the opening pages of the *Kitāb* of Sībawayhi (d. ca. 180/796), where verbs are classified as *māḍī* lit. ‘having passed, elapsed’, or *muḍāriʿ* ‘resembling’, an abbreviation for *fiʿl muḍāriʿ li-sm al-fāʿil* ‘verb resembling the agent noun’. Since there is no firm evidence that they were in use much before Sibawayhi’s time, they probably arose in the period of his association with the pioneers of grammar, who were active from about the middle of the 2nd/8th century. The asymmetry of the terms is remarkable: the first refers to the completed state of the action, hence commonly, but misleadingly, translated as ‘past (tense)’, better ‘perfect’, the second to the morphosyntactic properties of the paradigm, seldom reproduced literally in English and usually replaced by the ad hoc label ‘imperfect’. In this entry they are not translated.

Of the various features which interested the Arab grammarians, there is space here to deal only with the time reference of verbs and the nature and consequences of the resemblance between the *muḍāriʿ* verb and the agent noun. Verb morphology in general, especially the numerous problems resulting from the clash of morphology and phonology, must be left out of account, although the larger texts devote scores if not hundreds of pages to this topic under the heading of *ṣarf* or → *taṣrīf*) ‘conjugation’. Because of the abundance of sources and wide variety of opinions, only the gist is offered here, without individual attribution. The reader is referred to the secondary works in the bibliography, especially Versteegh (1995), for a more detailed discussion.

In the earliest grammar, represented by Sībawayhi, there is no sign of the grammaticalization of time as verbal tenses. Consider the following selection from Sībawayhi’s vocabulary for time in relation to both verbs and nouns. The bracketed components are not always used but cumulatively produce a composite definition:

i. Past: *mā maḍā [min az-zamān, min ad-dahr]* ‘what has passed [of time]’, *mā tarāxā ʿanka* ‘what has gone away from you’ in time or space, *mā qad waqaʿa* ‘what has indeed happened’, *mā nqaṭaʿa* ‘what has ceased’
ii. Present: *kāʾinun lam yanqaṭiʿ* ‘being, uninterrupted’, *mā ʾanta fīhi* ‘what you are in’, *mā lam yamḍi* ‘what has not passed’, *al-ḥīnu llaḏī fīhi l-fiʿl* ‘the time in which the act (or verb) is’, *fī yawmika* ‘on this day of yours’, *[ittiṣālu fiʿlin] fī ḥīni wuqūʿihi [ġayri munqaṭiʿ]* ‘[the connection of an act (or verb)] at the time of its occurrence, [uninterrupted]’, *fī ḥāli l-fiʿl* ‘in the immediate circumstance of the act (or verb)’, *fiʿlun muttaṣilun fī ḥāli ḏikrika ʾiyyāhu* ‘an act (or verb) connected with the immediate circumstance of your mentioning it’
iii. Future: *mā yakūnu wa-lam yaqaʿ* ‘what will be and has not happened’, *mā lam yaʾti*

'what has not come', *mā yastaqbilu [min az-zamān, min ad-dahr]* 'what is in the future [of time]', *mā lam yakun baʿdu* 'what has not yet been', *mā sa-yaqaʿu* 'what will occur'

It is apparent that Sībawayhi's position is entirely pragmatic, that time reference is a matter of the speaker's perception and the context of the utterance. As shown elsewhere (→ parts of speech), Sībawayhi does not even define the future verb formally as the one prefixed with the future marker *sa-* or *sawfa*. On the contrary, having set up a category of verb for 'what will be and has not happened', he illustrates it with the unmarked *muḍāriʿ* form *yaḍribu*, ambiguously 'he strikes' and 'he will strike'. Moreover, his definitions of certain negative constructions involving time shifts seem deliberately framed to avoid mentioning tenses. Thus, *lam* 'not', which negates *muḍāriʿ* verbs but fixes their time reference in the past, is simply described as *nafyun li-qawlihi faʿala* 'negation of a person's statement 'he did'', scil. *lam yafʿal* 'he did not do'; likewise, *lan*, which again negates *muḍāriʿ* verbs but gives them a future meaning, is *nafyun li-qawlihi sa-yafʿalu* 'negation of a person's statement "he will do"' (here with the future marker!), scil. *lan yafʿala* 'he will not do'. It seems Sībawayhi would rather issue circular definitions than invoke the concept of tenses.

The closest Sībawayhi comes to a three-tense system is in his discussion of pseudoverbal interjections such as *ṣah* 'sh!'. These do not take the agent markers of what he calls *al-fiʿlu l-ḥādiṯu fīmā maḍā wa-fīmā yastaqbilu wa-fī yawmika* 'the act/verb happening in what has passed, in what is in the future, or on this day of yours' (*Kitāb* I, 102, ed. Derenbourg; the Bulaq edition, I, 123, reads *yustaqbalu*), but no verb paradigms are linked with these three times. More than once, Sībawayhi refers to verb patterns generically as *li-mā māḍā wa-mā lam yamḍi* 'denoting what has passed and what has not passed', suggesting that in fact there is no third category. For him, time reference lay in the pragmatic context and in the adverbial complements known collectively as the → *mafʿūl fīhi* 'what the act was done in', expressed as a noun signifying a place or time in the dependent (= adverbial) case, generically labeled the *ẓarf* 'container' of the action (= *mafʿūl fīhi*), e.g. *ʾamsi* 'yesterday' (uninflected), *al-yawma* 'today', *ġadan* 'tomorrow'.

In the two centuries after Sībawayhi, the grammarians found themselves having to superimpose the Greek concept of past, present, and future time on the indigenous bi-aspectual system, even though they were aware that it was a mismatch. To their credit, they did not try to replicate the Greek conjugations in Arabic (they would not have known what they were), and in the end all they could do was to paraphrase Sībawayhi's original classification, but now with the additional mention of time *zamān*, so *māḍī*, originally 'passed' (in time and space, as was pointed out) now came to mean 'past' in time, while the *muḍāriʿ* was acknowledged as ambiguously referring to both present and future time, the first metaphorically and the second literally, according to the theory that an act in present time is only the coming into being of a potential future and so is logically subordinate to the future. The imperative is sometimes included in the future category, exactly as in Sībawayhi, but there is no suggestion that it constitutes a third paradigm to make up a set of three tenses. It is hardly surprising that the terminology itself also vascillated. Although *māḍī* remained the most common term for 'passed/past', it is occasionally expressed as *munqaḍī* 'having come to an end'; the 'present' was either *ḥāḍir*, i.e. '[physically] present' or *ḥāl* 'the (current) situation', and in one context *muqīm* 'abiding', while 'future' was usually *mustaqbal* (which occurs occasionally in the *Kitāb*, mostly with reference to the real future rather than a verb), but also *muntaḏar* 'awaited' and *mutawaqqaʿ* 'expected'.

A different solution was found by the Kūfans, who distributed the three times as follows: *māḍī* verbs for the past, the *muḍāriʿ* verbs for the future *mustaqbal*, and between them the agent noun (→ *ism al-fāʿil*), under the name of the *fiʿl dāʾim* lit. 'ongoing action, lasting action' for the present *ḥāḍir*. It is not likely that the Kūfans were proposing this as a formal tense system, but rather as a method of exemplifying the three times, although the question deserves more study. The role of the active → participle in the modern colloquial Arabic tense system, of which the *fiʿl dāʾim* may be a distant relation, cannot be touched upon here, but Eisele (1999) is recommended for its coverage, not only of the Egyptian dialect but also of the whole topic of tense and aspect in Arabic. It would appear from this work (especially 1999:22ff.) that the Kūfans were close in spirit to Koschmieder's theory that the present cannot be a point of time but only an interval in time. In any case, their inclusion of the agent noun in the tense system was easily disqualified on the ground that it was not a verb at all. Furthermore, it has no specific time reference, as one grammarian was happy to point out with the pair *ʾana qāʾimun ġadan* 'I am/will be standing tomorrow' and *ʾana qāʾimun ʾamsi* 'I was standing yesterday'. It was well known that a time difference could be encoded syntactically with the agent noun, distinguishing *ʾana qātilun ġulāmaka* 'I will be killing your slave boy' (verbal syntax) from *ʾana qātilu ġulāmika* 'I am the killer of your slave boy' (nominal syntax, → *ism al-fāʿil*), but this is irrelevant to the issue of tense as a morphological property of verbs.

Mention of the agent noun brings us to the most interesting feature of the *muḍāriʿ*, from which it derives its full technical name, *al-fiʿl al-muḍāriʿ li-sm al-fāʿil* 'the verb resembling the agent noun'. This 'resemblance' (*muḍāraʿa*) is empirical, i.e., it is already observed in the language and not due to any synchronic assimilative tendencies by speakers (historically it may well be otherwise, of course). The features shared with the agent noun are: (i) both may function as predicates prefixed with the marker *la-* 'indeed', e.g. *ʾinnahu la-kātibun/la-yaktubu* 'verily he is indeed writing'; (ii) both may be prefixed with particularizing elements, viz. the definite article on the agent noun *al-kātibu* 'the writer' and the future prefix on the verb *sa-yaktubu* 'he will write'. The concept of 'resemblance' is well developed in Sībawayhi and was identified in a wide range of linguistic phenomena, but the idea was lost as grammar moved from the descriptive to the prescriptive, and the name *al-fiʿl al-muḍāriʿ* 'the resembling verb' is almost the only survival. Significantly, the resemblance is bidirectional: the inflection of *muḍāriʿ* verbs is due to their resemblance to agent nouns, but by the same token the verbal behavior of agent nouns (as in *ʾana qātilun ġulāmaka* 'I will be killing your slave boy', see above) is attributed to their resemblance to the verb. Sībawayhi was criticized for this apparent circularity, but in his defense Ibn Jinnī pointed out that such inconsistency is natural in language.

Because of this resemblance, the *muḍāriʿ* verb shares two out of the three nominal inflections, with a third unique to itself. The two common inflections (which could be called 'moods', but Arabic makes no terminological distinction between them and the nominal 'cases') are *yaktub-u* 'he writes' for syntactically independent verbs, similar to indicative, cf. *al-kātib-u* 'the writer' (independent case); and *yaktub-a* 'he may write' for subordinate verbs, similar to subjunctive, cf. *al-kātib-a* (dependent case). The third is a privative inflection, typically the loss of the final short vowel, e.g. *yaktub* 'he might write' used in a variety of contexts, negative, prohibitive, conditional, and indirect imperative, and having no analogue among nouns, which have their own unique case, the oblique (*al-kātib-i* 'of the writer'; → parts of speech).

In one respect, the resemblance is misleading: the dual and masculine plural suffixes on both nouns and verbs are superficially identical, but they are entirely different in segmentation. On nouns, the suffixes are *-āni* and *-ūna*, respectively, alternating with *-ayni* and *-īna* for the oblique cases (thus, *kātibūna/kātibīna* 'men writing'), with the final *-n-* also being dropped in possessive constructions (thus, *kātibū/kātibī l-madīnati* 'the writers of the city'). On *muḍāriʿ* verbs, they are also *-āni* and *-ūna*, but this time alternating with *-ā* and *-ū* for the nonindependent 'moods' (thus, *yaktubūna/yaktubū* 'they write/may write'), in addition to which there is a 2nd person feminine singular suffix *-īna/-ī* (*taktubīna/taktubī* 'you [fem. sg.] write/may write'). In the nouns, the long vowels and diphthongs are analyzed as case markers and the *-n-* as an allomorph of the → *tanwīn*, while in the verbal morphemes the long vowels are analyzed as agent pronouns and the alternation *-n-/-∅* as an inflectional (scil. mood) marker.

This has important consequences for the analysis of the verb, where the agent pronouns may be overt or latent. In the *māḍī*, the agent suffixes are mostly overt, e.g. *katab-tu* 'I wrote', *katab-nā* 'we wrote', etc., the exception being the 3rd person singular masculine and feminine, *kataba* and *katabat*. Given the agreement rules for verbal sentences (→ *ism al-fāʿil*), *kataba* and *katabat* are ambiguous until it is known whether an overt agent noun is present (usually immediately following), contrast *kataba r-rajulu* ‘the man wrote’ but also *kataba r-rajulāni/ r-rijālu* ‘the two men/the men [pl.] wrote’ with *kataba* alone, understood by default as ‘he wrote’. Under these circumstances the *-a* and *-at* of *kataba/katabat* cannot be analyzed as agent suffixes in parity with *katab-tu* etc., although this is a nearly universal practice in Western pedagogical and other treatments. In the Arab theory, since the verb is not marked for number when it precedes the overt agent, there is no visible element in the verb which can be identified as the agent marker, hence the agent is said to be ‘latent’ (*mustatir* lit. ‘hidden’) in the verb, and the sequences *-a/-at* mark only gender and otherwise are semantically empty and phonologically arbitrary (the *a* vowel is chosen because it is the ‘lightest’ *ʾaxaff*).

The same is true for the *muḍāriʿ*, but the situation is more complex. Consistent with the above, the invariable 3rd person masculine singular *yaktubu* in *yaktubu r-rajulu/r-rajulāni/ r-rijālu* ‘the man/two men/men [pl.] write’ is held to contain only a latent *mustatir* masculine agent pronoun, but the Arabs take it a stage further. All the prefixes of this paradigm are regarded as insufficiently distinctive to qualify as agent pronouns: in *taktubu* ‘you [masc. sg.] write’, for example, the *ta-* is not an agent pronoun but only a marker of the 2nd person. The reasoning for this is as follows: in the set *taktubu, taktubīna, taktubāni, taktubūna, taktubna* ‘you [masc./fem. sg., common dual, and masc./fem. pl.] write’, the agent pronoun is generally discernible as a suffix, mostly a long vowel followed by the mood marker *-na/-ni*, as in *taktubīna* ‘you [fem. sg.] write’, segmented as *ta-ktub-ī-na*, leaving the *ta-* prefix with no pronominal function, according to an axiom that no feature can be marked by more than one morpheme. By a similar logic, the *na-* of *naktubu* ‘we write’ is not an agent pronoun but only a marker of the 1st person plural. If its latent agent pronoun were to be expressed at all, it might be one of the suffixes of the 2nd and 3rd person plural, scil. **naktubūna*, although historically such a form is most unlikely, unattested in any ancient Semitic language. Interestingly, however, in certain dialects of Arabic, there is indeed a contrast between *niktib* ‘I write’ and *niktibū* ‘we write’, showing a two-way analogical spread. While this may seem confusing, it is at least as coherent as the Western habit of treating the prefixes and suffixes together as discontinuous agent morphemes.

Justice has not been done to the Arab grammarians in this short entry: their sensitivity to the aspectual and temporal features of the verbal system can perhaps best be appreciated by looking at the limits in which they had to operate, unlike English, which has, in addition to the two simple tenses ‘write’ and ‘wrote’, probably hundreds of compound verb forms, ranging from ‘he is writing’ to monsters such as ‘he would have had been going to have been writing’. Excluding participial constructions, Arabic has only four compound verb forms, one with a variant, all involving *kāna* ‘be’. These are, with ad hoc translations, (i) *māḍī* + *muḍāriʿ*, e.g. *kāna yaktubu* ‘he was writing’; (ii) *muḍāriʿ* + *muḍāriʿ*, e.g. *yakūnu yaktubu* ‘he will be writiing’; (iii) *māḍī* + *māḍī*, e.g. *kāna [qad] kataba* ‘he had written’; and (iv) *muḍāriʿ* + *māḍī*, e.g. *yakūnu [qad] kataba* ‘he will have written’. Type (i) has a variant with the future prefix, *kāna sa-yaktubu*, which might be translated ‘he was going to write’, although it is a mistake to try to map any of the Arabic structures unequivocally on to the possible English translations. The bare *māḍī* itself can have present and future time reference in performative, optative, and conditional uses, and the *muḍāriʿ* in its different environments can likewise be rendered by any of the three tenses in English (→ aspect).

In the end, Western scholars reveal much the same lack of unanimity as their medieval predecessors on the question of the relationship between the three times and the binary verb system of Arabic (and indeed of many other languages). Nebes (1982), for example, gathered three thousand tokens of the *kāna yaktubu* combination, and from a closer examination of a thousand or so, he confirmed that they denote some kind of imperfective with anteriority (*Vorzeitigkeit*, not *Vergangenheit* ‘past’) in an aspectual system. On the other hand, Aartun (1963) argued for a tense system, specifically an opposition of preterite and nonpreterite, but his still very valuable collection of data insidiously seems to demonstrate the contrary, that time reference is inferred from the context and the lexical properties of the verb, rather than from its conjugated form.

Bibliographical references

Primary source

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. Ed. Hartwig Derenbourg. Paris: Imprimerie Nationale, 1881–1889 (Repr., Hildesheim and New York: G. Olms, 1970); ed. Būlāq, 1898–1900 (Repr., Baghdad: Muṯannā, [1965]); ed. ʿAbd as-Salām Muḥammad Hārūn. Cairo: Dār al-Qalam etc., 1968–1977 (2nd ed., 1977; 3rd ed., 1983). [The first and last seven chapters of the *Kitāb* are published, with links to various editions, commentaries, translations, and manuscripts, by Michael G. Carter, A. Matveev, and Lutz Edzard at www.arts.usyd.edu.au/research_projects/sibawiki/homepage/, where there are also links to other Web editions.]

Secondary sources

Aartun, Kjell. 1963. *Zur Frage altarabischer Tempora.* Oslo: Universitetsforlaget.

Bohas, Georges, Jean-Patrick Guillaume, and Djamel Eddine Kouloughli. 1990. *The Arabic linguistic tradition.* London and New York: Routledge.

Carter, Michael G. (ed.). 1981. *Arab linguistics: An introductory classical text with translation and notes.* Amsterdam: J. Benjamins.

——. 1997. “The term *muḍāriʿ* in the *Kitāb* of Sībawayhi”. *Proceedings of the Arabic and Islamic Section of the 35th International Congress of Asian and North African Studies Budapest 1997.* I. *Linguistics, literature, history*, ed. Kinga Dévényi and Tamás Iványi, 3–14. Budapest: Eötvös Loránd University.

Eisele, John. 1999. *Arabic verbs in time: Tense and aspect in Cairene Arabic.* Wiesbaden: O. Harrassowitz.

Howell, Mortimer Sloper. 1886–1911. *A grammar of the Classical Arabic language translated and compiled from the works of the most approved native or naturalized authorities.* Allahabad: Northwestern Provinces and Oudh Govt. Press.

Jahn, Gustav. 1895–1900. *Sībawaihi's Buch über die Grammatik, übersetzt und erklärt.* Berlin: Reuther und Reichard. (Repr., Hildesheim: G. Olms, 1965.)

Nebes, Norbert. 1982. *Funktionsanalyse von* kāna yafʿalu: *Ein Beitrag zur Verbalsyntax des Althocharabischen mit besonderer Berücksichtigung der Tempus- and Aspektproblematik.* Hildesheim: G. Olms.

Owens, Jonathan. 1990. *Early Arabic grammatical theory: Heterogeneity and standardization*, Amsterdam and Philadelphia: J. Benjamins.

Troupeau, Gérard. 1960–1963. “La notion du temps chez Sibawayhi”. *Compte rendu du Groupe Linguistique des Etudes Chamito-Sémitiques* 9.44–46.

Versteegh, Kees [C.H.M.]. 1981. “La conception des ‘temps’ du verbe chez les grammairiens arabes”. *Analyses, Théorie* 3.47–68.

——. 1993. *Arabic grammar and Qurʾānic exegesis in early Islam.* Leiden: E.J. Brill.

——. 1995. *The explanation of linguistic causes: Az-Zağğāğī's theory of grammar.* Amsterdam and Philadelphia: J. Benjamins.

——. 1997. *The Arabic linguistic tradition.* London and New York: Routledge.

Michael G. Carter (Sydney University)

Mafʿūl

1. Definition

The word *mafʿūl* is derived from the Arabic root *f-ʿ-l* ‘to do, make’ and refers to something done or made. In the *Qurʾān*, the word *mafʿūl* occurs twice: “that Allah might conclude a thing that must be done” (Q. 8/42, 44; translation Pickthall 1938:262–263). In grammatical terminology, *mafʿūl* refers to the accusative noun/pronoun on which the act of the verb ‘falls’. This covers all the nominal complements of the verb, in particular the object, which was also called *mafʿūl bihi*. Both terms are usually translated with ‘object’, although a better equivalent for *mafʿūl* is ‘patient’ (Mosel 1975:247): *mafʿūl* indicates the part of the sentence whose referent is affected by the verbal action, just as → *fāʿil* indicates the one who performs the verbal action and is therefore best translated with ‘agent’. In later grammatical terminology, the term acquired a meaning that is closer to ‘object’, which is the translation used here.

An essential characteristic of the verb's complements is that they are additive, that is, they are external to the core proposition in a clause or sentence and thus constitute optional elements in the clause structure. Arabic grammarians referred to this optional status as *faḍla* or the ‘surplus’ part of a sentence, which does not form part of the kernel or core predication (Ryding 2005:276). According to them, the two basic constituents of the core verbal sentence are the verb and the agent, traditionally referred to as *ʿumda*. Any noun complement occurring after the inclusion of the verb and agent may be omitted or not mentioned, while the meaning of the sentence remains intact. Al-Mubarrad (*Muqtaḍab* III, 116) defines *faḍla* as “an element which adds to the meaning expressed by the verb and its agent and whose omission from speech does not affect the integrity of the text” (*ʾiḏā ḏakartahu zidta fī l-fāʾida wa-ʾiḏā ḥaḏaftahu lam tuxlil bi-l-kalām*; cf. Taha 1995:113–114).

The canonical case ending for optional nominal elements is the accusative case. All complements share this ending, but they differ in the role each plays. The Basran and Kufan grammarians disagreed about the number of *mafāʿīl*. According to the Basrans, there are five *mafāʿīl*, while the Kufans recognized only the direct object as *mafʿūl*. Al-Farrāʾ, for instance, uses the term *mafʿūl* infrequently, and without distinguishing between different types (Owens 1990:160).

The most common accusative complements (*mafāʿīl*) are the following (Mosel 1975:248, 253–257; Owens 1988:167–168; Ryding 2005).

- i. *Mafʿūl bihi* 'direct object [lit. 'that to which it is done']'; the direct object is the thing done. It is used with the transitive verb, which may have one, two, or three *mafʿūl bihi*.
- ii. *Mafʿūl muṭlaq* 'absolute object'; the absolute or cognate object emphasizes or enhances a previous statement by a verbal noun derived from the main verb (→ object, absolute).
- iii. → *Mafʿūl fīhi* 'locative object [lit. 'that in which it is done']'; the adverbials of time and place are locative objects, called in Arabic *ḏarf* 'container, vessel', which refers to either time or place.
- iv. *Mafʿūl maʿahu* 'object of accompaniment [lit. 'that with which it is done']'; an accompanying object denotes the thing with which the action is done; it must be prefixed with the particle *wa-*.
- v. *Mafʿūl li-ʾajlihi* 'object of cause [lit. 'that for which it is done']'; the accusative noun indicating the cause is the thing for (the sake of) which the action is done; it consists of an indefinite dependent noun expressing the reason for the action.
- vi. → *Tamyīz* 'specification'; the accusative noun of specification is the noun that specifies the content of a number between 11 and 19, or the term of a comparative or superlative structure; at the sentence level, this is an indefinite dependent noun which provides specific information about an antecedent, usually the agent.
- vii. → *Ḥāl* 'circumstance'; the accusative noun of circumstance is a dependent adjective or active participle which indicates the (temporary) condition of its antecedent; it answers the question of how the action was performed (Badawi a.o. 2004:161; Ryding 2005:276).

Some of these nominal complements function as adverbial elements denoting degree, manner, place, or time. Later summaries of the commentators contain up to fifteen different accusative complements (*manṣūbāt*). For instance, the one by aš-Širbīnī (Carter 1981:324–329) adds to the ones mentioned above the excepted element after *ʾillā* (→ *istiṯnāʾ*); the noun negated by *lā*; the vocative after *yā*; the predicate of *kāna* (→ *kāna wa-ʾaxawātuhā*); the topic after *ʾinna* (→ *ʾinna wa-ʾaxawātuhā*); the two objects of *ḏanna* and related verbs; and those constituents that derive their accusative from the word with which they agree in case ending (*tābiʿ*), such as the adjective (→ *ṣifa*) and the substitute noun (*badal*; → apposition). As-Sīrāfī and Ibn Manẓūr introduced two additional types of *mafāʿīl* (Xuḍayr 2003:30): the *mafʿūl minhu* and the *mafʿūl ʿalayhi* (cf. Suyūṭī, *ʾAšbāh* III, 59). The term *mafʿūl minhu* is used once by Sībawayhi (*Kitāb* I, 273; perhaps not part of the original text) to indicate the accusative *al-ʾasada* in the phrase *ʾiyyāka wa-l-ʾasada* 'beware of the lion!', which indicates the thing one is warned about (Mosel 1975:253–254).

2. TRANSITIVITY AND OBJECT

Within the standard theory of Arabic grammar, the governor of the object is the verb. Initially, there was a controversy between the Kufan and the Basran grammarians about whether the governor of the *mafʿūl bihi* is the verb alone, or the agent, or the verb and the agent together (Ibn al-ʾAnbārī, *ʾInṣāf* 40–42; → *ʿamal*). Eventually, the Basran view prevailed that it is the verb which governs the *mafʿūl bihi*, or one of the elements that resemble the verb, such as the verbal noun (→ *maṣdar*), the active participle, the passive participle, the *mubālaġa*, the *taʿajjub*, the *ṣifa mušabbaha*, or the → *ism fiʿl*. Of crucial importance in the debate were the concept of transitivity (→ *taʿaddin*) and the role of the verb denoting the action and the agent carrying it out. The Basran view, which regarded the verb as the sole governor of the direct object, was based on formal, syntactic grounds, and did not take into account the semantic reference of verbs such as *māta* in *māta zaydun* 'Zayd died', or the role of the agent, as in *ḍuriba zaydun* 'Zayd was hit': in neither case is Zayd the real agent of the action denoted by the verb (Ibn as-Sarraj, *ʾUṣūl* I, 169).

The *mafʿūl* closest to the concept of *mafʿūliyya* ‘objectness’ is the cognate accusative, which is called by many grammarians *ism al-ḥadaṯ* ‘noun of the event’ because it indicates and refers to the essence of the verb. Many grammatical treatises referred to the cognate object with the term *mafʿūl ḥaqīqī* ‘real object’. A late commentator on Ibn Mālik’s *ʾAlfiyya*, al-Xuḍarī (*Ḥāšiya* I, 167; cf. Xuḍayr 2003:34), says about the ranking of the *mafāʿīl*: “Rank their objects, first the cognate…then the direct, adverbial, cause, accompaniment” (*mafāʿīlahum rattib fa-ṣaddir bi-muṭlaqin…wa-ṯanni bihi, fīhi, lahu, maʿahu qad kamul*). Al-Mubarrad (*Muqtaḍab* II, 121, IV, 299) confirms that the real *mafʿūl* is the cognate object, since it indicates the action of the verb, and with it the agent becomes the carrier of the action.

Philosophically speaking, some grammarians, especially those who worked within a Muʿtazilite framework, argued that the cognate *mafʿūl* is the real object because it refers to the action produced by the agent. Thus, if one says *ḍaraba zaydan* ‘he hit Zayd’, the agent necessarily preceded his action, which only exists because of the action he carried out. The *mafʿūl bihi*, i.e. Zayd, in this sentence was not created by the agent, but by God, and may very well have existed before the agent, while the act of hitting only took place by the agent’s carrying it out (Zajjājī, *ʾĪḍāḥ* 83; cf. Versteegh 1995:136, 82–83).

Outside the philosophical context, Arabic grammarians assigned a special syntactic status to the *mafʿūl bihi*, distinguishing it from all the other *mafāʿīl*. The analysis of transitivity and the explanation of the role of the agent and the patient in completing the grammatical meaning of transitivity is apparent in most early grammatical treatments of this topic. Grammarians made it clear that transitive verbs (*ʾafʿāl mutaʿaddiya*) are those whose action goes beyond the agent to reach a direct object. Their treatment of transitivity did not contradict the basic belief that accusative complements are not ‘essential’ elements of the proposition. Yet, the direct object usually received special attention, because it is the only nominal complement implied by the definition of transitive verbs. It is not clear, therefore, whether it really is a *faḍla*.

Sībawayhi’s *Kitāb* contains many passages in which all nominal complements are referred to as *mafāʿīl*. From a purely structural point of view, they all form part of the verbal sentence, yet, Sībawayhi never regards them as essential for the meaning of the verbal sentence. Later grammarians, however, shifted toward a more verb-centered focus, together with an in-depth analysis of the verb’s domain and the role it plays with respect to doer and direct object (in the case of transitive verbs). As a result, the function of the direct object continued to be perceived as different from the other complements of the verb.

Al-Mubarrad suggests, albeit indirectly, that the *mafʿūl bihi* has a different status from other accusative complements of the verb. Apparently, he believes that the direct object is necessarily implied by the transitive verb (*Muqtaḍab* III, 116; Taha 1995:114; cf. Owens 1988:173–174). This means that it could not be an optional element (*faḍla*) of the sentence (*Muqtaḍab* IV, 335):

> If you say ‘Abdallah hit Zayd’, you could if you want, say ‘Abdallah hit’. You inform me [with this] that there was an [action of] hitting [performed] by Abdallah, and [the verb] will be equivalent to [the verb in] ‘Abdallah stood up’, except that you know that the [action of] hitting has gone beyond its doer to the one hit. [You also know that] your saying ‘he stood’ does not go beyond the doer. If you say ‘Abdallah hit Zayd’, you inform me who that patient was (*fa-ʾiḏā qulta ḍaraba ʿabdullāhi zaydan, fa-ʾin šiʾta qulta ḍaraba ʿabdullāhi fa-ʿarraftanī ʾannahu qad kāna minhu ḍarbun fa-ṣāra bi-manzila qāma ʿabdullāhi ʾillā ʾannaka taʿlamu ʾanna ḍ-ḍarb qad taʿaddā ʾilā maḍrūb fa-ʾin qulta ḍaraba ʿabdullāhi zaydan ʾaʿlamtanī man ḏālika l-mafʿūl*)

Although indications of the different status of the verb’s direct object are present in Sībawayhi’s *Kitāb* and al-Mubarrad’s *Muqtaḍab*, it was in Ibn as-Sarrāj’s *ʾUṣūl* that the status of the direct object, and ultimately the concept of *faḍla*, became distinct from the status of other surplus items. The term *faḍla* remains for Ibn as-Sarrāj a structural term referring to the fact that a verbal sentence can be represented on the surface structure level by the verb and the agent only. On the other hand, because the meaning of a transitive verb is only complete with the inclusion of its direct object, the semantic status of the object must not be considered that of a surplus (*ʾUṣūl* I, 412): “For the direct object of the verb [there exists] a share in the verb as [it is the case] for the agent…” (*li-l-mafʿūl ḥiṣṣa* *min al-fiʿl kamā li-l-fāʿil…*). This view is supported by his innovative classification of verbs (→ *taʿaddin*). Transitive real verbs are explained by Ibn as-Sarrāj as those which include a reference to their patients. Due to the verb's meaning, the action of experience denoted by it is carried out by the agent to affect a patient. This argument depends on semantic criteria rather than the traditional structural arguments held by his predecessors. Ibn as Sarrāj (*ʾUṣūl* I, 171) says, "These transitive verbs are not complete [as a semantic whole] nor do they exist except with the existence of the [direct] object, because if you say: 'I mentioned', and nothing is mentioned, [this] would be impossible…" (*wa-lā tatimmu hāḏihi l-ʾafʿāl al-mutaʿaddiya wa-lā tūjadu ʾillā bi-wujūd al-mafʿūl li-ʾannaka ʾiḏā qulta ḏakartu wa-lam yakun maḏkūr fa-huwa muḥāl*).

From the above, it is clear that while the initial argument presented by earlier grammarians confirms that the direct object is a structural surplus, Ibn as-Sarrāj presents a different argument, highlighting the existence of a semantic patient of the transitive verb. Such a semantic patient may be realized on the surface structure in the form of a direct object. The direct object of the transitive verb is never a semantic surplus, because it always exists in the deep structure of the verbal sentence. In his argumentation, it is clear that he looks into each verb's semantic field in order to determine the components that form it. The direct object is one of the components with which the semantic field of a transitive verb is completed.

Later grammarians, such as az-Zajjājī and al-Fārisī, did not pay much attention to the semantic classification of verbs proposed by Ibn as-Sarrāj. In their work, the treatment of *mafʿūl* shifts back to the structural explanation, which systematically uses the term *faḍla* to refer to any accusative complement, whether direct object or other complements. Modern grammarians often subsume the *mafʿūl bihi*, as well as the other accusative complements, under the category of *mafāʿīl*, in that they are all complementary items. Nevertheless, the terms *ẓarf, ḥāl, tamyīz*, and *mafʿūl muṭlaq* are retained to refer to specific noun complements, each according to its structural role in the sentence.

Sībawayhi employs the term *mutaʿaddin* to refer not only to the transitive verb but also to any verb taking any noun complement, such as the cognate accusative or the adverbial of time and place. The term *mutaʿaddin* in this sense should not be understood to mean 'transitive' in the sense of a verb having a direct object. Instead, it means that the verb's syntactic effect actually passes over and beyond the agent to an accusative complement (cf. Levin 1979). This accusative complement may be a cognate accusative or an adverbial of time or place. Along the same lines, the term *mutaʿaddin* is used to refer to the operation of verb-like elements such as the interjection (→ *ism al-fiʿl*).

The category of verbs that are connected with their nominal complement by means of a preposition remained problematic throughout the history of Arabic grammar. There are a few instances where Sībawayhi directly compares the direct object of the verb to other complements, whether in the accusative or the genitive case. In one instance (*Kitāb* I, 92, 94), he explains that the prepositional phrase *bi-zaydin* in the example *marartu bi-zaydin* 'I passed Zayd' equals both in meaning and in status a direct object *zaydan*, so that it is as if the speaker had said *marartu zaydan*. Sībawayhi states that such a verb 'reaches' or 'links with' (*yaṣilu*) the object noun by means of a preposition (*wāṣil bi-ḥarf al-ʾiḍāfa*). The verb in this case does not operate on the object noun, since the preposition is the operator and causes the object noun to be in the genitive (Mosel 1975:249). Thus, the noun governed by the preposition was regarded by him as being actually a direct object, which 'occupies an accusative slot', because the action of the verb either befalls the noun or reaches it in one way or another (cf. Ibn as-Sarrāj, *ʾUṣūl* I, 162; Owens 1988:176–177). Contemporary grammarians consider the genitive noun to be in the accusative slot but do not necessarily refer to it as the direct object of the verb.

A second problematic issue was the status of the agent of a passive verb, which functions syntactically as an agent, but semantically refers to something that is affected by the verb's action and is therefore a *mafʿūl fī l-maʿnā* 'object in meaning'. In Sībawayhi's *Kitāb*, a *mafʿūl* that is raised to become the 'agent' of a passive verb remains a *mafʿūl* (*Kitāb* I, 41ff.). In a sentence like *ḍuriba zaydun* 'Zayd was hit', *zaydun* is categorized by Sībawayhi (*Kitāb* I, 42.5) as 'an object/patiens whose verb does not go beyond it to an object/patiens' (*al-mafʿūl* *allaḏī lā yataʿaddāhu fiʿluhu ʾilā mafʿūl*; Mosel 1975:65–70). Sībawayhi does not say that this *mafʿūl* takes the place of an agent, a concept introduced by al-ʾAxfaš and used by al-Mubarrad, who states that the *mafʿūl* "takes the position of the agent" (*yaqūmu maqām al-fāʿil*; *Muqtaḍab* IV, 50; Owens 1990:155). In later grammatical terminology, such a *mafʿūl* was called 'a substitute of the agent' (*nāʾib al-fāʿil*; Owens 1988:180–185). Obviously, this is connected with the controversy about agency and the role of the agent in carrying out the action denoted by the verb. Note, however, that later grammarians, such as as-Sīrāfī (*Šarḥ* I, 262) and Ibn Jinnī (*Lumaʿ* 117), continued to regard the *nāʾib al-fāʿil* as a 'real' object.

Confusion between technical terms and semantic reference was common practice in early grammatical description. Yet, grammarians agreed on one thing: the direct object is the noun governed by the verb in that the transitive verb passes over the agent and reaches the noun that represents the entity affected by the action of the agent.

3. The status of the other accusative complements

In early grammatical treatises, several terms are used to refer to the accusative complements in the verbal sentence after the verb and its agent have been specified. Thus, the terms *mafʿūl* or even *mafʿūl bihi* are used indiscriminately to refer either to the direct object or to any accusative noun occupying the slot after the verb and agent. Yet, most Arabic grammarians regard the *mafʿūl bihi* as representing an entity different from the other accusative complements. They explain that the transitive verbs refer both to the agent (doer of the action) and to the direct object (the entity affected by such action). In addition, they argue that all verbs, whether transitive or intransitive, contain an indication of the essence of the act (expressed by the absolute object), the time (expressed by the adverb of time), the place (expressed by the adverb of place), the manner by which the action is carried out (expressed by the object of specification), the reason for carrying out the action (expressed by the object of cause), and other implications for the circumstance (expressed by the object of circumstance and of accompaniment). These indications constitute different elements pertaining to the meaning of the verb and illustrate what, how, when, where, and why a certain action takes place (Xuḍayr 2003:28).

With respect to most of the accusative complements other than the direct object, some of the earlier grammarians made an effort to analyze them as belonging to the category of the direct object, for instance the object of accompaniment. Other grammarians argued that the function of the particle *wa-* was to introduce the accompanying object in order to avoid confusion with the direct object. There was also disagreement about what governed this object, the particle *wa-* or the verb (Ibn al-ʾAnbārī, *ʾInṣāf* 110–113).

Some grammarians, including Sībawayhi (*Kitāb* I, 211), contemplated the possibility of explaining the accusative case endings of adverbials of time and place by their share of 'objectness' (*mafʿūliyya*). Many grammarians regarded them as 'objects' by expansion, i.e. *ʿalā s-saʿa*. Others argued that what makes them adverbials is the inclusion of the meaning of the preposition *fī* 'in, at' in their meaning (→ *mafʿūl fīhi*).

Along the same lines, some earlier grammarians considered the object of cause to be a direct object clarifying the cause of the action, especially when the preposition *bi-* or *li-* is omitted from the proposition. One argument in favor of this view was that the verb sometimes has a direct object in the genitive case governed by a preposition, in which case the verb is described as 'reaching' the accusative noun by means of a preposition (Xuḍayr 2003:46).

In the excepting construction (→ *istiṯnāʾ al-mufarraġ*), the accusative noun after the exceptive particle *ʾillā* is a direct object. After verbs such as *ḥāšā* and *ʿadā*, the noun is also a direct object. Several grammarians considered the noun after the exceptive verb to be a direct object, for it "comes after the completion of thought by mentioning the verb and its agent" (Ibn as-Sarrāj, *ʾUṣūl* I, 342; Xuḍayr 2003:53). One argument justifying the accusative case for the *mustaṯnā* is that the particle *ʾillā* carries out the role of the verb, and, thus, the noun following it is in the accusative, just as the direct object of a verb would be.

Many grammarians called the accusative of specification (→ *tamyīz*) a *mafʿūl*. There was disagreement between the Basran and the

Kufan grammarians as to what a *tamyīz* is (Ibn al-ʾAnbārī, *ʾInṣāf* 351–353). They all agreed, however, that it is an optional element (*faḍla*) added to the sentence after its indispensable constituents have been included, i.e. the verb and the agent (Xuḍayr 2003:57–59). Several grammarians identified the *tamyīz* with the *mafʿūl bihi* when it occurred after *kam* 'how much?', or after an elative verb (Mubarrad, *Muqtaḍab* IV, 299; Jurjānī, *Muqtaṣid* I, 671; Ibn as-Sarrāj, *ʾUṣūl* I, 258). The accusative of circumstance (→ *ḥāl*) was a similar case since most early grammarians considered the accusative noun following the verb and the agent to be a *mafʿūl*.

There was much terminological confusion in early grammatical treatises with respect to the relationship of verbs like *kāna* to their predicate (→ *kāna wa-ʾaxawātuhā*; → *nawāsix*). Sībawayhi regards the predicate of *kāna* in a sentence like *kāna zaydun munṭaliqan* 'Zayd was leaving' as its object because for him *kāna* is a transitive verb with an agent and an object (Mosel 1975:282–285; Levin 1979:186–190). Later grammarians posited an underlying nominal sentence (*zaydun munṭaliqun*) and accordingly called the underlying predicate the predicate of *kāna* (*xabar kāna*). Apparently, Ibn as-Sarrāj was the first grammarian to treat *kāna* as different from other verbs in that it is not a real verb whose action goes beyond the agent to a direct object (*ʾUṣūl* I, 74), Here, too, the discussion revolves around the question of whether the status of the object is determined by semantic or syntactic criteria.

The case of the topic after *ʾinna* and similar particles is less problematic (→ *ʾinna wa-ʾaxawātuhā*). Here, the grammarians emphasized (e.g. Ibn al-Warrāq, *ʿIlal* 333.7–8; Zajjājī, *ʾĪḍāḥ* 64.17) that the comparison of the topic to an object is purely formal and that, semantically, the topic governed in the accusative case by *ʾinna* has nothing to do with a direct object.

With respect to verbs like *kāda* (*ʾafʿāl al-muqāraba*), and for the verbs of perception (*ʾafʿāl aš-šuʿūr*) and the verbs of requiring (*ʾafʿāl ar-rajāʾ*), many early and modern grammarians consider the verbal noun or the clause composed with *ʾan* after these verbs to be in the place of the direct object of such verbs, rather than their predicate (Sībawayhi, *Kitāb* I, 51.3, 158.3, 160; Mubarrad, *Muqtaḍab* III, 70, 75; Ḍayf 1987:III, 231).

The disagreement among grammarians about what constitutes a 'real' direct object extended toward the category of the exclamatory verbs (*ʾafʿāl at-tafḍīl*). The accusative after these verbs, as in *mā ʾaḥsan-a zayd-an* 'how beautiful Zayd is!', is sometimes regarded as a direct object. Other grammarians consider it to be 'similar to the *mafʿūl*' because it comes after the elative verb, which is similar to verbs in that it governs an accusative noun (Ibn al-ʾAnbārī, *ʾInṣāf* 57–68). In this case, the discussion is more about the underlying structure of this construction. According to the Basran grammarians, the elative in this sentence is indeed a verb, and they analyze the underlying meaning as *šayʾun ʾaḥsana zaydan* 'something made Zayd beautiful'. For the Kufan grammarians, the question of the status of this object never arose because they analyzed *ʾaḥsana* as a noun, so that it could not have a direct object in any event.

Bibliographical references

Primary sources

Ibn al-ʾAnbārī, *ʾInṣāf* = ʾAbū l-Barakāt ʿAbd ar-Raḥmān ibn Muḥammad al-ʾAnbārī, *Kitāb al-ʾinṣāf fī masāʾil al-xilāf bayna n-naḥwiyyīna l-baṣriyyīna wa-l-kūfiyyīn*. Ed. Gotthold Weil. Leiden: E.J. Brill, 1913.

Ibn Jinnī, *Lumaʿ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Lumaʿ fī l-ʿarabiyya*. Ed. Ḥusayn Šaraf. Cairo: ʿĀlam al-Kutub, 1979.

——, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. Ed. Muḥammad ʿAlī an-Najjār. 3 vols. Cairo, 1952–1956. (Repr., Beirut: Dār al-Hudā, n.d.)

Ibn as-Sarrāj, *ʾUṣūl* = ʾAbū Bakr ibn as-Sarī Ibn as-Sarrāj, *Kitāb al-ʾuṣūl fī n-naḥw*. Ed. ʿAbd al-Ḥusayn al-Fatlī. 3 vols. Beirut: Muʾassasat ar-Risāla, 1985.

Ibn al-Warrāq, *ʿIlal* = ʾAbū l-Ḥasan Muḥammad ibn ʿAbdallāh al-Warrāq, *ʿIlal an-naḥw*. Ed. Maḥmūd Muḥammad Maḥmūd Naṣṣār. Beirut: Dār al-Kutub al-ʿIlmiyya, 2002.

Jurjānī, *Muqtaṣid* = ʾAbū Bakr ʿAbd al-Qāhir ibn ʿAbd ar-Raḥmān al-Jurjānī, *Kitāb al-muqtaṣid fī šarḥ al-ʾĪḍāḥ*. Ed. Kāẓim Baḥr al-Marjān. 2 vols. Baghdad: Wizārat aṯ-Ṯaqāfa wa-l-ʾIʿlām, 1982.

Mubarrad, *Muqtaḍab* = ʾAbū l-ʿAbbās Muḥammad ibn Yazīd al-Mubarrad, *al-Muqtaḍab*. Ed. Muḥammad ʿAbd al-Xāliq ʿUḍayma. 4 vols. Cairo: Dār at-Taḥrīr, 1965–1968.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. Ed. ʿAbd as- Salām Muḥammad Hārūn. 4 vols. Cairo: al-Hayʾa al-Miṣriyya li-l-Kitāb, 1977.

Sīrāfī, *Šarḥ* = ʾAbū Saʿīd al-Ḥasan ibn ʿAbdallāh as-Sīrāfī, *Šarḥ Kitāb Sībawayhi*. Ed. Muḥammad ʿAbd ad-Dāyim. Cairo, n.d.

Suyūṭī, *ʾAšbāh* = Jalāl ad-Dīn ʾAbū l-Faḍl ʿAbd ar-Raḥmān ibn ʾAbī Bakr as-Suyūṭī, *al-ʾAšbāh wa-n-*

naḏ̣āʾir. Ed. ʾIbrāhīm ʿAbdallāh. Damascus: Arabic Language Academy, 1986.
Xuḍarī, *Ḥāšiya* = Muḥammad ibn Muṣṭafā al-Xuḍarī aš-Šāfiʿī, *Ḥāsiyat al-Xuḍarī ʿalā Šarḥ Ibn ʿAqīl ʿalā ʾAlfiyyat Ibn Mālik*. Cairo: ʾAmīra aš-Šarafiyya Printshop, 1980.
Zajjājī, *ʾĪḍāḥ* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān Ibn ʾIsḥāq az-Zajjājī, *al-ʾĪḍāḥ fī ʿilal an-naḥw*. Ed. Māzin al-Mubārak. Cairo: Dār al-ʿUrūba, 1959.

Secondary sources
Badawi, Elsaid, Michael G. Carter, and Adrian Gully. 2004. *Modern Written Arabic: A comprehensive grammar*. London and New York: Routledge.
Carter, Michael G. 1981. *Arab linguistics: An introductory classical text with translation and notes*. Amsterdam: J. Benjamins.
Ḍayf, Šawqī. 1987. *ʾUṣūl al-luġa*. [Introduction to the *Majalla Majmaʿ al-Luġa al-ʿArabiyya*.] Cairo.
Levin, Aryeh. 1979. "The meaning of *taʿaddā l-fiʿl ilā* in Sībawayhi's *Kitāb*". *Studia orientalia D.H. Baneth dedicata*, 193–210. Jerusalem: Magnes Press.
Mosel, Ulrike. 1975. *Die syntaktische Terminologie bei Sībawaih*. Ph.D. diss., University of Munich.
Owens, Jonathan. 1988. *The foundations of grammar: An introduction to medieval Arabic grammatical theory*. Amsterdam and Philadelphia: J. Benjamins.
——. 1990. *Early Arabic grammatical theory: Heterogeneity and standardization*. Amsterdam and Philadelphia: J. Benjamins.
Pickthall, Mohammad M. 1938. *The meaning of the Glorious Qurʾan: An explanatory translation*. Hyderabad: Government Central Press. (Repr., London: Islamic Culture Centre.)
Ryding, Karin C. 2005. *A reference grammar of Modern Standard Arabic*. Cambridge University Press.
Taha, Zeinab Ahmed. 1995. *Issues of syntax and semantics: A comparative study of Sibawayhi, al-Mubarrad, and Ibn as-Sarraaj*. Ph.D. diss., Georgetown University.
Versteegh, Kees. 1995. *The explanation of linguistic causes: Az-Zaǧǧāǧī's theory of grammar*. Amsterdam and Philadelphia: J. Benjamins.
Xuḍayr, Muḥammad ʾAḥmad. 2003. *Qaḍāyā l-mafʿūl bihi ʿinda n-nuḥāt al-ʿArab*. Cairo: Al-Anglo Library.

Zeinab Ahmed Taha
(American University in Cairo)

Mafʿūl fīhi

In the canonical theory of the Arabic grammarians, the term *mafʿūl fīhi* indicates the adverbial adjunct of time and place as one of the complements of the verb. Adjuncts usually have an accusative ending, both in Classical and in Modern Standard Arabic, e.g. *ṣumtu ramaḍāna* 'I fasted during Ramadan', *sirtu farsaxayni* 'I walked two parasangs', *sa-ʾaḏhabu ġadan* 'I'm going away tomorrow', *qumtu xalfa-ka* 'I stood up behind you'. In the case of place adverbials, a prepositional phrase is preferred when the location is specified, e.g. *ṣallaytu fī masjidi n-nabiyyi* 'I prayed in the Prophet's mosque', not **ṣallaytu masjida n-nabiyyi* (for Classical Arabic, see Wright 1964:II, 109–112; for Modern Standard Arabic, see Badawi a.o. 2004:149–152; → locative).

In early Arabic grammar, the usual term for adverbial adjuncts was *ḏ̣arf* (pl. *ḏ̣urūf*), lit. 'container'. It has been suggested (Merx 1889:146; cf. Talmon 2000:248) that this is a Greek borrowing from the word *anggeíon* 'vessel, receptacle', used by Aristotle to indicate the temporal or local circumstances. In Sībawayhi's *Kitāb*, the term *ḏ̣arf* denotes both the extralinguistic reality of location and time, and the syntactic function (Mosel 1975:345–362). The extralinguistic reality is clearly what is meant by him when he says (*Kitāb* I, 201.8–9) at the beginning of the chapter on the adverbs of time and location:

> This is the chapter of those locations and times that receive an accusative; this is because they are containers in which things happen and exist; the reason they receive the accusative ending is that they are that in which something happens and in which something exists. (*hāḏā bāb mā yantaṣibu min al-ʾamākin wa-l-waqt wa-ḏāka li-ʾannahā ḏ̣urūfun taqaʿu fīhā l-ʾašyāʾ wa-takūnu fīhā fa-ntaṣaba li-ʾannahu mawqūʿun fīhā wa-makūnun fīhā*)

Other passages in the *Kitāb* give the impression that Sībawayhi treats the *ḏ̣urūf* as a category of words, for instance when he distinguishes (*Kitāb* I, 209.1ff.) between words that are neither *ism* nor *ḏ̣arf* (e.g. *bi-* 'in'), words that are *ḏ̣arf* (e.g. *xalfa* 'behind'), and words that are *ism*, but not *ḏ̣arf* (e.g. *ḥimār* 'donkey'; Mosel 1975:345). This implies that he posits a difference between those nouns that can function as *ḏ̣arf* and those that cannot. The words *jawf* and *dāxil*, for instance, which both mean 'interior', cannot be used in a *ḏ̣arf* construction: **huwa dāxila d-dāri*, whereas nouns like *xalfa* 'behind' and *ʾamāma* 'in front' can, e.g. *huwa xalfa d-dāri* 'he is behind the house'. The reason he gives is that the latter indicate a space that is contiguous to the noun (*talī l-ʾasmāʾ min ʾaqṭārihā*), whereas the former are presumably parts of the noun, just like *baṭn* 'belly' or *ḏ̣ahr* 'back' (*Kitāb* I, 204.10ff.; Mosel 1975:348–349).

Words that can function as adverbials of time and location continue to belong to the category of the nouns, as demonstrated by the fact that they can be governed by the particle *min*, e.g. *min xalfika* 'from behind you'. In this construction, however, they no longer function as *ḏ̣arf*. Their status as nouns also implies that they can become the topic of a nominal sentence. This point is emphasized by al-Mubarrad (*Muqtaḍab* IV, 341), who cites sentences like *xalfu-ka wāsiʿun* 'the area behind you is wide', in which the noun *xalf*, which usually functions as an adverbial of space, is the topic of the sentence (Owens 1990:148–149).

A different situation obtains when adverbial adjuncts of location are used as predicate (Mosel 1975:347), for instance *zaydun xalfa-ka* 'Zayd is behind you'. In such a construction, most grammarians posit an underlying verb *zaydun mustaqirrun xalfa-ka* 'Zayd is residing behind you' in order to fulfil the requirement that the topic and the predicate refer to the same entity (cf. Gully 1995:226). Time adverbials can only occur in this function if the topic of the sentence refers to an event, e.g. *al-laylata l-hilāl*, which is only allowed under the reading 'tonight the sighting of the new moon takes place' rather than 'tonight is the new moon' (cf. Ibn Jinnī, *Lumaʿ* 11–12).

In Sībawayhi's theory, the adjunct of time or place is not a direct object (cf. Owens 1990:111–115). A sentence like *sāra l-yawma* can therefore have two different readings: 'he went today' (adjunct of time) or 'he spent the day going' (direct object). The only way to find out which of the two readings is meant is to passivize the sentence; the direct object is promoted to the position of agent of the passive sentence (*nāʾib al-fāʿil*), as in *sīra l-yawmu*, whereas the adjunct remains in the accusative, *sīra l-yawma*. According to Sībawayhi (*Kitāb* I, 110), the former is an example of *saʿat al-kalām* or *ittisāʿ* (cf. Versteegh 1990), because the underlying sentence is *sīra sayru l-yawmi*, in which the word *sayr* is deleted (on *ittisāʿ* in adverbial expressions, see also Ibn Hišām, *Muġnī* II, 693–694; Gully 1995:227). The preferred reading determines which element is responsible for the accusative ending: if *al-yawma* is a direct object, it is governed by the verb; if it is an adjunct, its accusative is explained by the fact that it is separated from the verb (see below).

In the early stages of the Arabic grammatical tradition, an alternative terminology existed for the local and temporal adjuncts, in which they were regarded as → *ṣifāt* 'properties, attributes'. According to Talmon (2000:247–248), the Kufan way of looking at the adverbials as *ṣifāt* goes back to what he calls the period of Old Iraqi Grammar and is the result of borrowing from the Syriac and Greek tradition. The two terms occur together in al-Xalīl's *Kitāb al-ʿayn* (Talmon 1997:147) in the entry on the root *ḏ̣-r-f* (*ʿAyn* VIII, 157, cf. VII, 325), where *ḏ̣arf* is defined as follows:

> *Ḏ̣arf*: container of anything, just as a jug is a container of something in it; *ṣifāt* like *ʾamāma* and *quddāma* are called *ḏ̣urūf*; you say *xalfaka zaydun* 'behind you is Zayd' and it is put in the accusative because it is a container of what is in it and it is a place for something else. (*wa-ḏ̣-ḏ̣arf wiʿāʾ kull šayʾ ḥattā l-ʾibrīq ḏ̣arf limā fīhi wa-ṣ-ṣifāt naḥwa ʾamāma wa-quddāma tusammā ḏ̣urūfan taqūlu xalfaka zaydun ʾinnamā ntaṣaba li-ʾannahu ḏ̣arf limā fīhi wa-huwa mawḍiʿ lahu*)

This view of local and temporal adverbials as attributes of the noun was discontinued by Sībawayhi but retained by the Kufan grammarians: al-Farrāʾ never uses the word *ḏ̣arf* to describe the adjuncts (the one occurrence in the edition of the *Maʿānī l-Qurʾān* II, 95.3–4, is probably a printing error for *ṭaraf li-l-fiʿl*; cf. Kinberg 1996:448), but he calls them *ṣifāt*. In fact, he uses this term to indicate any locative, for instance when he says (*Maʿānī* II, 385.5) "*li-* or any other of the locatives" (*bi-l-lām ʾaw ġayrihā min aṣ-ṣifāt*; Owens 1990:145). Sometimes, he uses the term *maḥall* to refer to them (five times, according to Kinberg 1996:187–188). Ṯaʿlab, too, uses *ṣifa* and, like al-Farrāʾ, includes all prepositional phrases in this category (Owens 1990:149). In this approach, all expressions that are locative in meaning (i.e. both *yawma l-xamīsi* and *fī yawmi l-xamīsi* 'on Thursday') are brought together in one syntactic category, that of the *ṣifāt* (Owens 1990:147).

According to Ibn al-ʾAnbārī (*ʾInṣāf* 108–110, problem 29), the Kufans explained the accusative ending of the adjuncts by the fact that they are independent from the construction of the sentence, i.e. through the principle of *xilāf* or *ṣarf*, analyzed first by Carter (1972, 1973, 1981:355) and called by Owens (1990:111–115) the "separation and non-identity principle". In fact, Sībawayhi, too, employs this principle. It explains the accusative ending in constructions like *ʿišrūna dirhaman* 'twenty dirhams' or *kam dirhaman* 'how many dirhams?', which in Sībawayhi's view is caused by the fact that the ending *-ūna* separates the two words and prevents it from being governed by the preceding word. But he also holds this principle responsible for the accusative ending in adverbial adjuncts, when he explains the difference between *yā sāriqan al-laylata ʾahla d-dāri* and *yā sāriqa l-laylati ʾahla d-dāri* 'O you who steals during the night from the people in the house' (Sībawayhi, *Kitāb* I, 89–90; Owens 1990:112). In this pair of phrases, the first phrase contains the adjunct *al-laylata*, which is separated from the verb by the ending *-an*; in the second phrase, the *al-layla* is the direct object and is therefore governed by the participle (in this particular construction, it receives the genitive).

The later tradition of Arabic grammar retained the memory of the original difference in the form of a terminological difference between Basran and Kufan grammar. Aš-Širbīnī, for instance, in his chapter on the *ẓarf az-zamān wa-ẓarf al-makān* (Carter 1981:351–367) mentions that this class is also called *mafʿūl fīhi* and by the Kufans *ṣifa*. This demonstrates that Sībawayhi's Basran successors replaced the term *ẓarf* with *mafʿūl fīhi*, in line with a new explanatory framework for all accusative endings as objects. Al-Mubarrad (*Muqtaḍab* II, 120, IV, 328; Owens 1990:149), for instance, uses *mafʿūl fīhi* as the normal term for adverbials of time and place. The term *ẓarf* remained in use, however, as an alternative. When introducing the *mafʿūl fīhi* in his list of objects, az-Zamaxšarī (*Mufaṣṣal* 25.15) says, for instance, that it is represented by the "two adverbials of time and place" (*ẓarfā z-zamān wa-l-makān*).

The term *mafʿūl fīhi* was already employed by Sībawayhi (*Kitāb* I, 195.21, 260.12), in the chapter "On nouns that are in the accusative, and are neither attributes, nor masdars, because they are the situation in which the thing takes place, and it is in the accusative because it is that in which the action is performed" (*hāḏā bāb mā yantaṣibu min al-ʾasmāʾ allatī laysat bi-ṣifāt wa-lā maṣādir li-ʾannahu ḥāl yaqaʿu fīhi l-ʾamr fa-yantaṣibu li-ʾannahu mafʿūl fīhi*). This has nothing to do with the adverbial adjuncts of time and place but rather indicates the *ḥāl* in expressions like *hāḏā ʿabdullāhi munṭaliqan* 'this is Abdallah, while he is leaving'; *kallamtuhu fāhu ʾilā fiyya* 'I spoke with him directly [lit. 'his mouth toward my mouth']' (Mosel 1975:274–275).

In later grammatical theory (e.g. Zamaxšarī, *Mufaṣṣal* 25–26), all *mafʿūl*s are brought together in one framework of transitivity (→ *taʿaddin*). A verb may have several objects: the action of the verb itself (*mafʿūl muṭlaq* → absolute object); the direct object (*mafʿūl bihi* or simply → *mafʿūl*); the object of accompaniment (*mafʿūl maʿahu*), which covers constructions like *mā ṣanaʿta wa-ʾabāka* 'what have you done together with your father?'; the object of purpose (*mafʿūl lahu*), which covers constructions like *faʿaltu ḏālika ḥiḏāra š-šarri* 'I did this in order to avoid evil'; and the circumstantial (→ *ḥāl*).

The central thought behind analyzing the adverbial adjuncts as objects is that a verb by itself implies the existence of several things: an agent, an object, the action itself, the circumstances, and the time and location. Az-Zajjājī (*ʾĪḍāḥ* 100.16–101.2; Versteegh 1995: 178) discusses this in connection with his explanation of the heaviness of the verb as compared to the lightness of the noun. Some grammarians, he says, give the following explanation for this difference:

> The only reason why nouns are lighter is that they do not signify more than one underlying referent, whereas the heaviness of the verb is caused by the fact that they have to indicate an agent, as well as one, two, or three objects, an infinitive, the two adjuncts of time and place, the circumstantial, and other things like these. (*ʾinnamā xaffa l-ism li-ʾannahu lā yadullu ʾillā ʿalā l-musammā llaḏī taḥtahu wa-ṯaqula l-fiʿl li-dalālatihi ʿalā l-fāʿil wa-l-mafʿūl wa-l-mafʿūlayni wa-ṯ-ṯalāṯa wa-ḏ-ḏarfayni min az-zamān wa-l-makān wa-l-ḥāl wa-mā ʾašbaha ḏālika*)

This analysis of the arguments of the verb is used by the Basran grammarians as an explanatory framework for the presence of accusative endings in sentences as markers of the objects. They believe that each and every accusative ending is the result of the governance (→ *ʿamal*) of a verb, which implies that the adverbials of time and place have to be analyzed as objects on account of their accusative ending. Al-Mubarrad, for instance, states unequivocally (*Muqtaḍab* IV, 299; Owens 1990:121) that all accusative endings are objects, thus integrating all constructions mentioned above in one theory of transitivity.

The general problem the adverbial adjuncts posed for the Arabic grammarians is neatly summarized by Owens (1990:151):

> The Arabic grammarians were faced with two potentially contradictory classificatory criteria, the need to reconcile a mutually exclusive tri-partite word classification (= Sibawayh) with distributional facts brings two of these word classes, prepositions and nouns, into a common class as locatives (= Farra'). The resolution of this problem in this case genuinely did lead to an interesting synthesis of the two approaches, one originally represented by Sibawayh, the other by Farra'

In al-Farrāʾ's analysis, different word groups enter into the definition of locative, whereas in Sībawayhi's analysis the semantic relationship between adjuncts and prepositional phrases is lost. The synthesis consists in the supposition of an underlying preposition *fī* which operates on the distinction between specific (*muxtaṣṣ*) nouns like *dār* 'house' and vague (*mubham*) nouns like *ʾamām* 'front' and *xalf* 'behind' (Owens 1990:150–151). In this later synthesis, only vague nouns are allowed to be used as locatives. Another way of putting this is to state that locativity implies the preposition *fī*, whose use is obligatory in the case of vague nouns (e.g. Ibn Jinnī, *Lumaʿ* 138). Here, the locative has really become a *mafʿūl fīhi*.

There is one other distinction to be mentioned here, that between fully inflected (*mutaṣarrif*) and incompletely inflected (*ġayr mutaṣarrif*) adverbs; the latter category includes words like *qablu* and *qaṭṭu*, which are called by some grammarians *ġāyāt* (e.g. Zamaxšarī, *Mufaṣṣal* 67, who calls them part of the *ḏurūf*; Ibn al-ʾAnbārī, *Lumaʿ* 54.11ff.; cf. Carter 1981:367). This terminology goes back to Sībawayhi (*Kitāb* II, 47), and also occurs in al-Farrāʾ (once, according to Kinberg 1996:539, namely *Maʿānī* II, 320.11–13). The term occurs already in Ibn al-Muqaffaʿ's commentary on Aristotle's *Perì hermēneías*, in which he enumerates the eight parts of speech. The last part of speech is that of the *ġāyāt*, which he identifies as the prepositional phrase in a sentence like *fulān al-kātib fī d-dāri* 'someone is writing [or 'the writer'] in the house' (Ibn al-Muqaffaʿ, *Manṭiq* 26.24–25): "His words *fī d-dāri* are the limit that is reached by the complete speech, and it is that which he wishes to affirm" (*kāna qawluhu fī d-dāri huwa l-ġāya allatī yaṣīru ʾilayhā jamʿ kalāmihi wa-ʾiyyāhā ʾarāda ʾan yuṯbita*). Talmon (2000:247–248, 2003:166–167) concludes from this that the term is borrowed from Greek logic.

Bibliographical references

Primary sources

Farrāʾ, *Maʿānī* = ʾAbū Zakariyyāʾ Yaḥyā ibn Ziyād al-Farrāʾ, *Maʿānī l-Qurʾān*. Ed. Muḥammad ʿAlī an-Najjār. 3 vols. Cairo: ad-Dār al-Miṣriyya, 1966–1972.

Ibn al-ʾAnbārī, *ʾInṣāf* = ʾAbū l-Barakāt ʿAbd ar-Raḥmān ibn Muḥammad al-ʾAnbārī, *Kitāb al-ʾinṣāf fī masāʾil al-xilāf bayna n-naḥwiyyīna l-baṣriyyīna wa-l-kūfiyyīn*. Ed. Gotthold Weil. Leiden: E.J. Brill, 1913.

——, *Lumaʿ* = ʾAbū l-Barakāt ʿAbd ar-Raḥmān ibn Muḥammad al-ʾAnbārī, *Lumaʿ al-ʾadilla fī ʾuṣūl an-naḥw*. Ed. Hadi M. Kechrida. Stockholm: Almqvist and Wiksell, 1963.

Ibn Hišām, *Muġnī* = Jamāl ad-Dīn ʾAbū Muḥammad ʿAbdallāh ibn Yūsuf Ibn Hišām, *Muġnī l-labīb ʿan kutub al-ʾaʿārīb*. Ed. Māzin al-Mubārak and Muḥammad ʿAlī Ḥamdallāh. Damascus: Dār al-Fikr, 1969.

Ibn Jinnī, *Lumaʿ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *Kitāb al-lumaʿ fī n-naḥw*. Ed. Hadi M. Kechrida. Stockholm: Almqvist and Wiksell, 1976.

——, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. Ed. Muḥammad ʿAlī an-Najjār. 3 vols. Cairo, 1952–1956. (Repr., Beirut: Dār al-Hudā, n.d.)

Ibn al-Muqaffaʿ, *Manṭiq* = ʾAbū Muḥammad (ʾAbū ʿAmr) ʿAbdallāh Ibn al-Muqaffaʿ, *al-Manṭiq*. Ed. Muḥammad Taqī Danešpažuh. Tehran: Imperial Iranian Academy of Philosophy, 1398/1978.

Mubarrad, *Muqtaḍab* = ʾAbū l-ʿAbbās Muḥammad ibn Yazīd al-Mubarrad, *al-Muqtaḍab*. Ed. Muḥammad ʿAbd al-Xāliq ʿUḍayma. 4 vols. Cairo: Dār at-Taḥrīr, 1965–1968.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. 2 vols. Bulaq, 1316 A.H. (Repr., Baghdad: Al-Muthanna Library, n.d.)

Xalīl, *ʿAyn* = ʾAbū ʿAbd ar-Raḥmān al-Xalīl ibn ʾAḥmad al-Farāhīdī, *Kitāb al-ʿayn*. Ed. Mahdī al-Maxzūmī and ʾIbrāhīm as-Samarrāʾī. 8 vols. Beirut: Muʾassasat al-ʾAʿlamī li-l-Maṭbūʿāt, 1988.

Zajjājī, *ʾĪḍāḥ* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān Ibn ʾIsḥāq az-Zajjājī, *al-ʾĪḍāḥ fī ʿilal an-naḥw*. Ed. Māzin al-Mubārak. Cairo: Dār al-ʿUrūba, 1959.

Zamaxšarī, *Mufaṣṣal* = ʾAbū l-Qāsim Maḥmūd ibn ʿUmar az-Zamaxšarī, *Kitāb al-mufaṣṣal fī n-naḥw*. Ed. Jens Peter Broch. Christiania: Libraria P.T. Mallingii, 1870.

Secondary sources

Badawi, Elsaid, Michael G. Carter, and Adrian Gully. 2004. *Modern Written Arabic: A comprehensive grammar*. London and New York: Routledge.

Carter, Michael G. 1972. "Twenty dirhams in the *Kitāb* of Sībawaihi". *Bulletin of the School of Oriental and African Studies* 35.485–490.

——. 1973. "*Ṣarf* et *xilāf*: Contribution à la grammaire arabe". *Arabica* 20.292–304.
——. 1981. *Arab linguistics: An introductory classical text with translation and notes.* Amsterdam: J. Benjamins.
Gully, Adrian. 1995. *Grammar and semantics in medieval Arabic: A study of Ibn-Hisham's* 'Mughni l-Labib'. Richmond: Curzon Press.
Kinberg, Naphtali. 1996. *A lexicon of al-Farrāʾ's terminology in his* Qurʾān *commentary with full definitions, English summaries, and extensive citations.* Leiden: E.J. Brill.
Merx, Adalbertus. 1889. *Historia artis grammaticae apud Syros.* Leipzig. (Repr., Nendeln: Kraus, 1966.)
Mosel, Ulrike. 1975. *Die syntaktische Terminologie bei Sībawaih.* Ph.D. diss., University of Munich.
Owens, Jonathan. 1990. *Early Arabic grammatical theory: Heterogeneity and standardization.* Amsterdam and Philadelphia: J. Benjamins.
Talmon, Rafael. 1997. *Arabic grammar in its formative age:* Kitāb al-ʿayn *and its attribution to H̲alīl b. Aḥmad.* Leiden: E.J. Brill.
——. 2000. "The first beginnings of Arabic linguistics: The era of the Old Iraqi School". *History of the language sciences*, ed. Sylvain Auroux, Konrad Koerner, Hans-Josef Niederehe, and Kees Versteegh, I, 245–252. Berlin: Mouton de Gruyter.
——. 2003. *Eighth-century Iraqi grammar: A critical exploration of pre-H̲alīlian Arabic linguistics.* Winona Lake, Ind.: Eisenbrauns.
Versteegh, Kees. 1990. "Freedom of the speaker: The term *ittisāʿ* and related notions". *Studies in the history of Arabic grammar*, II, ed. Michael G. Carter and Kees Versteegh, 281–293. Amsterdam and Philadelphia: J. Benjamins.
——. 1995. *The explanation of linguistic causes: Az-Zaǧǧāǧī's theory of grammar.* Amsterdam and Philadelphia: J. Benjamins.
Wright, William. 1964. *A grammar of the Arabic language.* 3rd ed. rev. W. Robertson Smith and M.J. de Goeje. Repr. 2 vols. Cambridge: Cambridge University Press.

Kees Versteegh (University of Nijmegen)

Maġribī

The term *maġribī* is the generic name for a host of scripts or styles used in the western part of the Islamic world from Tunisia to Morocco (*ʾIfrīqiyā, Maġrib*), southern Spain (Andalusia), and sub-Saharan Africa, for the copying of books and for use in the state apparatus (principally the chancery), as well as for ordinary purposes of writing.

Although *maġribī* is easily identifiable as a group, there is still much research that needs to be done before we can attempt a comprehensive history of its development and its various styles. In a way, the problem here is similar to the situation with the → *nasx* script, the bookhand of the Islamic East (*Mašriq*).

There are a number of theories as to the origin of the *maġribī* scripts. The traditional view is given by Ibn Xaldūn (d. 808/1406), who was of the opinion that *maġribī* scripts developed from *ʾandalusī*, which eventually (after the collapse of Muslim rule in Spain) supplanted all the scripts of *ʾIfrīqiyā* (*al-xaṭṭ al-ʾifrīqī*) and, later, the Maghreb (Morocco). This view was originally opposed by Octave Houdas (1886), who pointed out that the inhabitants of the Maghreb made a clear distinction between the two scripts. However, recent research appears to support the view that the scripts and codicological practices first elaborated in Spain were then diffused through northwestern Africa (Stanley 1995:23).

Houdas was the first scholar to point out that the *maġribī* scripts had a common ancestor: 'cursive → Kufic' (Houdas 1886:95). Indeed, the latest research done by F. Déroche points to the development of *maġribī* from the Abbasid bookhand (Houdas' 'cursive Kufic'; other appellations include semi-Kufic, Eastern or Persian Kufic), which dates from the 3rd/9th century but has its roots in the 1st century of Islam in a handwriting used for documents, either in the chancery or for judicial purposes (Déroche 1994:77, 2004:75).

'Abbasid bookhand' (*écriture livresque abbasside*), a term coined by F. Déroche, is a generic term for a variety of scripts used originally for non-Qurʾānic texts. These scripts are known to have been used in the western part of the Arab world from the second half of the 3rd/9th century to the middle of the 5th/11th century (Déroche 1999:237); in other words, they coexisted with *maġribī*. Some of them appear to have been influenced by the Syriac *sertā* script. The *ʾalif* in these scripts often looks like a reversed <s> or a club, and can be found with or without a head-serif (seriffed or sans serif). Here we also encounter a *lām ʾalif* ligature known later as the *lām ʾalif al-warrāqiyya* (Déroche 1992:132). A dressed-up (stylized) version of the Abbasid bookhand is known as the New Abbasid Style or New Style (NS), used mainly, but not exclusively, for the copying of the *Qurʾān*, which had its golden age in the 4th/10th century. This term again was introduced by F. Déroche, although other contemporary scholars prefer to refer to it as 'broken cursive' (Blair 2006:151–160).

Maġribī comes into the domain of book production perhaps even before the first half of the 4th/10th century (Déroche 2004:75). The oldest known manuscript in *maġribī* script is *Kitāb maʿrifat al-bawl wa-ʾaqsāmih* by ʾIsḥāq ibn Sulaymān al-ʾIsrāʾīlī, dated 345/957. Two other prominent examples from a later period are *Kitāb siyar al-Fazarī*, dated 379/989–990, a copy of *al-Muwaṭṭaʾ* by Malik ibn ʾAnas, copied in 391/1001, and the earliest surviving *maġribī* Qurʾān, copied in 398/1008 (Déroche 2004:74–75, 1999:239–240).

In *maġribī* scripts, letters are not learned individually according to specified norms as is the case with the new 'proportioned' scripts of the Islamic East. Instead, writing is learned by imitating complete words. It is worth noting also that the expansion of the *maġribī* script is linked in a special way with the exclusive diffusion of the Malikite jurisprudence throughout that whole region (Houdas 1886:99, 101).

Maġribī is written using a calamus having a fine and soft point to the nib (Déroche 2004:75, 79–80), as opposed to the straight or obliquely cut nibs used for most of the eastern (*mašriqī*) scripts. This manner of cutting the nib was apparently connected with the type of reed available in the Maghreb, which was not as hard as the reeds used in the eastern part of the Islamic world (Houdas 1886:96, 98). This fact is of great importance since we know that the cut of the nib has a direct impact on the type of strokes produced by a pen. The pen strokes in *maġribī* are, therefore, almost always of the same thickness.

Two other major features are worth noting here: the *maġribī* alphabet (*hijāʾ*) has a different sequence from the one used in the eastern part of the Islamic world and its mnemotechnical arrangement (*ʾabjad*), and the numerical values are also different. In the *maġribī* sequence *zāy(zāʾ)* is followed by *ṭāʾ, ḏ̣āʾ, kāf, lām, mīm, nūn, ṣād, ḍād, ʿayn, ġayn, fāʾ, qāf, sīn, šīn, hāʾ, wāw*, and *yāʾ*.

Since the arrangement of the *ʾabjad* in groups of five, six, and eight also differs, the differences in terms of numerical values are substantial. Thus, *sīn* = 300, *šīn* = 1000, *ṣād* = 60, *ḍād* = 90, *ḏ̣āʾ* = 800, and *ġayn* = 900.

Other distinctive features of the *maġribī* scripts concern the pointing, vocalization, and orthography. Thus, the letter *fāʾ* had one point below and the *qāf* one above, while the two letters (as well as *nūn* and *yāʾ*) at the end of a word were rarely pointed (Houdas 1886:107). This way of pointing goes back to the first centuries of Islam. Also, the points of the *tāʾ marbūṭa* were very often suppressed, and their suppression was obligatory at the end of periods in rhymed prose.

Just as in the New Abbasid Style ('broken cursive') manuscripts of the *Qurʾān*, in *maġribī* manuscripts the vocalization and orthoepic signs were executed in colors (red, green, yellow, and blue). Red was used for vowels. *Šadda* was indicated by a semicircle in red: ◡ (with *fatḥa*) or ◠ (with *kasra*). *Hamzat al-qaṭʿ* was marked with a red or yellow dot, and *hamzat al-waṣl* with a green dot (Déroche 1991:231; Blair 2006:222, 223, 226). Many non-Qurʾānic manuscripts were also polychrome, using red, blue, and green colors for key words, chapter headings, and the like.

Although many *maġribī* manuscripts are associated with the square format, as far as we know, the square Qurʾāns appear only at end of the 5th/11th century. The earliest square-format codex was made at Malaga in 500/1106. Originally, *maġribī* manuscripts were written in codices of other formats such as the oblong format (Déroche 2001:593, 606, 611). Here, mention should also be made of the use of parchment in manuscripts as late as the 9th/15th century, as well as the use of colored paper (peach, red, purple, pale pink) from as early as the 7th/13th century (Blair 2006:393).

There are great differences in style; some manuscripts were written with very thin nibs, and others with thicker ones, and the scripts vary from very small to very large, often spiky in appearance. Moreover, the form of a letter is never characteristic of a style. On a single page, executed by the same hand, one may encounter up to four forms of the same letter. Looking at some manuscripts, one forms the impression of admiring a spiderweb-like composition.

Nevertheless, we find certain common characteristics, such as: the ascenders in *ʾalif, lām, ṭāʾ/ḏ̣āʾ* are rarely straight and are often curved; the loop of the *ṣād/ḍād* is identical with that of *ṭāʾ/ḏ̣āʾ* and has no 'tooth'; the initial *ʿayn/ġayn* has a very wide opening; some descenders (such as in the final or isolated *nūn*) have enormous but often irregular curves; the *ʾalif* of prolongation has a characteristic spur-like tail which descends below the baseline; and the distinct shape of *lām ʾalif al-warrāqiyya* (Houdas 1886:105–107; van den Boogert 1989:30).

Just as with the Abbasid bookhand and the New Abbasid Style, most *maġribī* scripts are seriffed, but there are some which are serifless (sans serif). The head-serif, when present, is left sloping and often in the shape of a dot, like a paunch (*kirš*; Gacek 2003:28).

Maġribī scripts are difficult to classify because of the great confusion of forms, stemming from the variety of scripts, itself the result of the lack of calligraphic standards seen in the Islamic East. Indeed, *maġribī* scribes tended to imitate the scripts of the manuscripts they copied, which could have been written in another region or country (Houdas 1886:100; Boogert 1989:31).

There are, at this stage, two classifications of the *maġribī* scripts. The first main classification of elegant scripts, suggested by O. Houdas, divides them according to regions or towns (centers of learning), e.g. *qayrawānī, ʾandalusī, fāsī*, and *sūdānī* (Houdas 1886:104; Abbott 1939:41–44).

The second classification, common in contemporary *maġribī* literature, divides the scripts into the following categories: *mabsūṭ, mujawhar, musnad* (also known as *zimāmī*), and *mašriqī*. According to Muḥammad al-Manūnī, these scripts established themselves during the Marinid dynasty (probably at the end of the 7th/13th century), after the *ʾandalusī* script became 'Maghribized'. *Mabsūṭ* script was used for Qurʾāns, *mujawhar* (the most common of *maġribī* scripts) for state documents and official correspondence, *musnad* for legal documents and personal use, and *mašriqī* for titles and headings in books and all types of other decorations, as well as in large inscriptions, e.g. in the Alhambra. Being an adaptation of → *ṯuluṯ*, the *mašriqī* is also referred to as *ṯuluṯ maġribī* (Gacek 2001; van den Boogert 1989:31–32).

The earliest mention of *ʾandalusī* script (as one of the twelve *kūfī* scripts) comes from ʾAbū Ḥayyān at-Tawḥīdī's work on calligraphy (Rosenthal 1971:24). Although some researchers, perhaps following Ibn Xaldūn's argumentation, have made a distinction between *ʾandalusī* and *maġribī* scripts (e.g. James 1992:87), this is perhaps an unwarranted distinction, since *ʾandalusī*, like the scripts of Qayrawān and Fās, certainly belongs to the large *maġribī* family of scripts.

This script has come to be seen as a "small, angular, archaic looking hand in which such letters as *kāf, ṣād* and *ḍād* are rather elongated" (James 1992:87), as opposed to a larger *maġribī*. We are told that it was a compact script usually found in small-format, square manuscripts of the *Qurʾān*. In this script, unlike other *maġribī* scripts, the final forms of such letters as *fāʾ, qāf, nūn*, and *yāʾ* were pointed (James 1992:87).

The distinction between the small (*ʾandalusī*) and larger (*maġribī*) is, however, confusing and perhaps even erroneous as it presupposes that the use of the small script was confined to *al-ʾAndalus*. It is likely, however, that both originated in Spain, and certainly both were later employed across the Straits of Gibraltar and in southern Spain. T. Stanley (1995:22–23) argues, "In fact, the two scripts can be seen as complementary, for in general the smaller type was used for single-volume copies of the Qurʾan, and the larger type for multi-volume copies".

The *sūdānī* scripts, employed in sub-Saharan Africa, are offshoots from *maġribī*, possibly coming from Tunisia (mainly Qayrawān), but perhaps also touched by influences coming from Egypt (Houdas 1886:102; Stanley 1999). The history of these scripts is not well known due to the fact that there are few dated manuscripts and not enough data regarding their provenance. Furthermore, they greatly differ in style; some are very coarse and heavy in appearance and others are more delicate and easy flowing. The strokes are often irregular in length and thickness (Abbott 1939:43). The oldest specimens of this script do not appear to be earlier than the 11th/17th century, and most come from the 12th/18th and 13th/19th centuries. The oldest manuscript we know of was copied as late as 1080/1669 (Déroche 2004:86–88; Bivar 1960:203).

Maġribī survived in its various forms into the early 14th/20th century thanks to printing by lithography. The well-known Fes lithographs are the best representation of the various styles used at that time. In movable type, however, *maġribī* was replaced by → *nasx*, and only recently have attempts been made to revive traditional *maġribī* calligraphy.

BIBLIOGRAPHICAL REFERENCES

Abbott, Nabia. 1939. *The rise of the North Arabic script and its Ḳur'ānic development, with a full description of the Ḳur'ān manuscripts in the Oriental Institute*. Chicago: University of Chicago Press.

Bivar, A.D.H. 1960. "A dated Kuran from Bornu". *Nigeria Magazine*, 199–205.

Blair, Sheila S. 2006. *Islamic calligraphy*. Edinburgh: Edinburgh University Press.

Boogert, Nico van den. 1989. "Some notes on Maghribi script". *Manuscripts of the Middle East* 4.30–43.

Déroche, François. 1991. "Deux fragments coraniques maghrébins anciens au Musée des arts turc et islamique d'Istanbul". *Revue des Etudes Islamiques* 59.229–235.

——. 1994. "O. Houdas et les écritures maghrébines". *al-Maxṭūṭ al-ʿarabī wa-ʿilm al-maxṭūṭāt (= Le manuscrit arabe et la codicologie)*, 75–81. Rabat.

——. 1999. "Tradition et innovation dans la pratique de l'écriture au Maghreb pendant les IVe/Xe et Ve/XIe siècles". *Numismatique, langues, écritures et arts du livre, spécificité des arts figurés: Actes du VIIe Colloque International sur l'Histoire et Archéologie de l'Afrique du Nord, réunis dans le cadre du 121e Congrès des Sociétés Historiques et Scientifiques, Nice, 21 au 31 octobre 1996*, ed. Serge Lancel, 233–247. Paris.

——. 2001. "Cercles et entrelacs: Format et décor des corans maghrébins médiévaux". *Académie des Inscriptions et Belles-Lettres: Comptes Rendus des Séances de l'Année 2001 Janvier-Mars*, 593–620.

——. 2004. *Le livre manuscrit arabe: Préludes à une histoire*. Paris: Bibliothèque Nationale de France.

Gacek, Adam. 2001. *The Arabic manuscript tradition: A glossary of technical terms and bibliography*. Leiden: E.J. Brill.

——. 2003. "The head-serif (*tarwīs*) and the typology of Arabic scripts: Preliminary observations". *Manuscripta Orientalia* 9:3.27–33.

Houdas, O. 1886. "Essai sur l'écriture maghrébine". *Nouveaux Mélanges Orientaux*, 85–112.

James, David. 1992. *The master scribes*: *Qur'ans of the 10th to 14th centuries A.D.* London and Oxford: Nour Foundation in association with Azimuth Editions and Oxford University Press.

Rosenthal, Franz. 1971. "Abū Ḥayyān al-Tawḥīdī on penmanship". *Four essays on art and literature in Islam*, 20–49. Leiden: E.J. Brill.

Stanley, Tim. 1995. *The* Qur'an *and calligraphy: A selection of fine manuscript material*. London. [Bernard Quaritch, catalogue 1213.]

——. 1999. "The Qur'anic script of Western Sudan: Maghribi or Ifriqi?". *The decorated word*: *Qur'ans of the 17th to 19th centuries*, ed. Manijeh Bayani, Anna Contadini, and Tim Stanley, I, 32–40. London and New York: Nour Foundation in association with Azimuth Editions and Oxford University Press.

ADAM GACEK (McGill University)

Maḥmūl

The term *maḥmūl* 'predicate' is part of Arabic philosophical terminology, equivalent to the Latin *praedicatum* (Georr 1948:217; Afnan 1969:80–81; Versteegh 1993:24–25). Its meaning in philosophical terminology corresponds to that of → *xabar* in linguistics (Elamrani-Jamal 1983:138–144; Fārābī, *Ḥurūf* 111.5–7), or → *ṣifa* as opposed to *mawṣūf* in theological terminology (Wolfson 1976:112–132; Ibn Sīnā, *Maqūlāt* 18–19). The composition of → *mawḍūʿ* 'subject' and *maḥmūl* can be brought about 'in the way of limitation' (*ʿalā naḥw at-taqyīd*) or 'in the way of report' (*ʿalā sabīl al-xabar*). Following the pattern of definitions, the first kind of composition has a nominal predicate: '[man is] a logical, mortal living being', i.e. '[man is] an animal that is logical and mortal' (*al-ḥayawān allaḏī huwa nāṭiq*), while the second one has a verbal predicate (Ibn Sīnā, *ʿIbāra* 30–34).

In speaking of predicates, the Arab philosophers take Aristotle's *Categories* (1 a 20–1 b 9) as their point of departure. On the basis of Aristotle's text, they usually distinguish between *yuqālu ʿalā* and *yuqālu fī*. As Ibn Rušd (d. 595/1198) puts it in his short commentary on Aristotle's *Categories* (*Talxīṣ* 12; cf. Ibn Sīnā, *ʿIbāra* 20), the names and definitions of secondary substances can be predicated of (*yuqālu ʿalā*) a subject, but accidents can be 'predicated in' (*yuqālu fī*) them.

A detailed presentation of a more elaborate version of Aristotle's theory was given by Ibn Sīnā (d. 428/1037). He distinguishes five types:

i. The subject of a sentence has a well-established essence and existence, e.g., the property by which it is characterized accedes 'from outside' as an accident or attribute (*ʿāriḍ, lāzim*): 'a human being is white' or 'laughing'.
ii. The subject of a given sentence is taken as something having a fixed essence, but the property is part of its constitution: 'a human being is an animal'.
iii. The subject is taken without having a well-established essence, and the property helps to establish its essence and existence, e. g. matter and form. Form always appears as a property of matter, not belonging to its

essence but contributing to the establishment of its actual subsistence.

iv. The subject is taken without having a well-established essence, and the property does not accede from outside, since it is part of its existence, e.g. 'substance' for 'body', being a predicate of 'animal'. 'Substance' is part of the definition of 'body'.

v. The subject does not have a well-established existence, and the property does not accede to its essence but rather to some inseparable attribute that gives it subsistence or to one of its first accidents.

'Human being' is the subject of 'animal' because 'animal' is a constituent part of 'human being'. 'Animal' is the subject of 'body' because 'body' is a constituent element of 'animal' and part of its existence. These are examples of the first type.

'Body' is the subject of 'whiteness' or 'blackness', without being constituted by these properties. A body is visible, consequently, it has color. Nevertheless, a body is not constituted by color, but 'color' is the subject of 'whiteness' or 'blackness'. This is the second type.

Consequently, a human being can be characterized by all these properties (Ibn Sīnā, *ʿIbāra* 18–19). Ibn Sīnā always refers here to the *Tabula Porphyriana* as being in the background of these five types (Maróth 1994:89–128).

In the logical approach to predication, there are problems which are unknown to linguists. The most interesting ones are the following:

i. A predicate can be expressed 'in agreement' (*bi-l-muwāḍaʿa: zaydun insānun* 'Zayd is a human being') or with 'derivation' (→ *ištiqāq: zaydun ʾabyaḍu* 'Zayd is white'; Aristotle, *De Interpretatione* 16 b 20; Fārābī, *Ḥurūf* 113.9–11; Ibn Sīnā, *Madxal* 28). A secondary substance can be predicated of (*yuqālu ʿalā*) a subject in agreement: 'a human being is an animal' (*al-ʾinsānu ḥayawānun*). The properties qualifying a subject (*yuqālu fī*) can be expressed by derivation: 'Zayd is white' (*zaydun ʾabyaḍu*), derived form *zaydun fīhi bayāḍun* 'in Zayd is whiteness'. If 'walking' (*ḏahābun*) is a property that can be found in Zayd (*yuqālu fī zayd*), then the following sentences can be derived from it: 'Zayd is walking' (*zaydun ḏāhibun, zaydun yaḏhabu*). 'Walking' (*ḏāhibun, yaḏhabu*) is derived from '(the act of) walking' (*ḏahābun*), which is the property being in Zayd. Consequently, all utterances have the same basic form: substance (the first one among Aristotle's categories) is connected with (or: is a bearer of) a property (one of the other nine categories). Thus, the ever-changing linguistic expression of this elementary structure is of secondary importance because *zaydun ḏāhibun, zaydun yaḏhabu*, and *yaḏhabu zaydun* are various linguistic representations of the same structure consisting of subject and predicate. All this means that accepting the theory of *ištiqāq*, the dichotomy of *jumla fiʿliyya/jumla ismiyya*, as well as the dichotomy of verbal/nominal predicates, becomes irrelevant because all of them go back to the same common logical pattern: 'whiteness is a property existing in Zayd', i.e. *zaydun ʾabyaḍu*, and 'walking is a property existing in Zayd', i.e. *zaydun ḏāhibun, zaydun yaḏhabu*, etc. The difference between *yuqālu ʿalā* and *yuqālu fī* was usually explained by philosophers in their commentaries written on Aristotle's *De interpretatione* (Fārābī, *ʿIbāra* 36–37; Ibn Sīnā, *Maqūlāt* 20, 27; etc.). Referring to Aristotle, Ibn Zurʿa (d. 398/1008) says that the predicate of a sentence always indicates an actual condition (*ḥāl*) of the subject (*Manṭiq* 34). Actual condition is an 'influence, impression' (*taʾṯīr*; the term corresponds to Greek *páthos* in the sense of 'quality, feature'), and because all qualities are of temporary character, and all words connected with time are considered verbs, the predicate (*maḥmūl*) is always a verb. What Ibn Zurʿa says corresponds to Ibn Sīnā's views; only their point of departure differs. Ibn Zurʿa's opinion tacitly assumes that a nominal predicate is always connected with an 'existential verb', i.e. a copula that expresses time.

ii. Philosophers realized that the lack of a → copula indicating present tense in sentences with nominal predicates was a basic difference between Arabic and other languages (Versteegh 1997:82–84; Jeremiás 2002:550–574; Fārābī, *Ḥurūf* 112; Ibn Sīnā, *ʿIbāra* 37–39). Even in this case there was an 'existential verb', if the connection between subject and predicate was made in the past or future tense. The Arab philosophers tried to express this connection by using the word *huwa*, for instance, in sentences like *zaydun huwa ʾabyaḍu* ‘Zayd *is* white’ (cf. *al-ḥayawānu llaḏī huwa nāṭiqun* cited above). In this case, *huwa* was not regarded as a pronoun but rather as a derived form from *huwiyya*, expressing existence. The idea of connection can be expressed by other ‘existential words’ (*kalim wujūdiyya*) as well: *zaydun mawjūdun ʿādilan* ‘Zayd is (being) just’. In the past and future tenses, the corresponding forms of ‘to be’ (*kāna, sa-yakūnu*) are used, as they are in other languages. The ‘existential words’ and the theory of their replacement by *huwa* as derived from *huwiyya*, expressing ‘existence’, were necessary, because verbal predicates themselves implied the existence of the subjects they were connected with (Fārābī, *ʿIbāra* 129.7–8.). Ibn Sīnā (*ʿIbāra* 39.13) says that *kāna* at the beginning of a sentence (*kāna llāhu ʿādilan ġafūran* ‘God is just and forgiving’) has a general validity for all three tenses, because *kāna* expresses here only a connection. The most important difference between verbs and existential verbs is that the former indicate their own connection with the subject, whereas the latter only establish a connection between two other things, one of them being a subject and the other one a predicate (Fārābī, *ʿIbāra* 36–37).

iii. In the logical analysis of language, a special problem is connected with quantification. A subject can be individual or general. This problem had been raised already by Aristotle (*De interpretatione* 20 b 13–18). If the subject is individual, should the predicate be individual as well? If its predicate, too, is individual, then there will be no similar individual in the outside world. In addition, two different individuals cannot be predicated of each other. In language this is admissible, especially in the case of names, but not in nature: ‘Zayd is ʾAbū l-Qāsim’. In all other cases, the situation is similar to the sentence ‘Zayd is white’. ‘Zayd’ is an individual being one can point at, but ‘white’ is a quality shared by many things. Consequently, in a sentence with both an individual and a general subject, the predicate will always be general (Ibn Sīnā, *Maqūlāt* 20–21; a similar problem is treated by al-Fārābī, *ʿIbāra* 141–142). In grammatical terms, Ibn Jinnī (d. 392/1002) gives what is basically the same rule (*Lumaʿ* 10.10–11): “If in a sentence a definite and an indefinite word co-occur, you make the definite word topic and the indefinite word comment” (*fa-ʾin ijtamaʿa fī l-kalām maʿrifa wa-nakira jaʿalta l-mubtadaʾ huwa l-maʿrifa wa-l-xabar huwa n-nakira*).

Bibliographical references

Primary sources

Aristotle, *De Interpretatione* = Aristoteles, *Perì hermēneías*. Ed. L. Minio-Paluello, *Aristotelis Categoriae et Liber de interpretatione*, 47–72. Oxford: Clarendon Press.

Fārābī, *Ḥurūf* = ʾAbū Naṣr Muḥammad ibn Muḥammad al-Fārābī. Ed. Muḥsin Mahdī, *Alfarabi’s Book of letters*. Beirut: Dar El-Machreq, 1970.

——, *ʿIbāra* = ʾAbū Naṣr Muḥammad ibn Muḥammad al-Fārābī, *Šarḥ Kitāb ʾArisṭūṭālīs fī l-ʿibāra*. Eds. William Kutsch and Stanley Marrow, *Alfarabi’s Commentary on Aristotle’s* Perì Hermeneias (De Interpretatione). 2nd ed. Beirut: Dar El-Machreq, 1986.

Ibn Jinnī, *Lumaʿ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *Kitāb al-lumaʿ fī n-naḥw*. Ed. Hadi M. Kechrida. Stockholm: Almqvist and Wiksell, 1976.

Ibn Rušd, *Talxīṣ* = ʾAbū l-Walīd Muḥammad ibn ʾAḥmad Ibn Rušd, *Talxīṣ manṭiq ʾArisṭū*, I. Ed. Gérard Jéhamy. Beirut, 1982.

Ibn Sīnā, *ʿIbāra* = ʾAbū ʿAlī al-Ḥusayn ibn ʿAbdallāh Ibn Sīnā, *Kitāb aš-šifāʾ*, *al-Manṭiq*. III. *al-ʿIbāra*. Ed. ʾIbrāhīm Madkūr. Cairo, 1959.

——, *Madxal* = ʾAbū ʿAlī al-Ḥusayn ibn ʿAbdallāh Ibn Sīnā, *Kitāb aš-šifāʾ*, *al-Madxal*. Ed. ʾIbrāhīm Madkūr. Cairo, 1952.

——, *Maqūlāt* = ʾAbū ʿAlī al-Ḥusayn ibn ʿAbdallāh Ibn Sīnā, *Kitāb aš-šifāʾ, al-Manṭiq*. II. *al-Maqūlāt*. Ed. ʾIbrāhīm Madkūr. Cairo, 1959.

Ibn Zurʿa, *Manṭiq* = ʾAbū ʿAlī ʿĪsā ibn ʾIsḥāq Ibn Zurʿa, *Manṭiq ibn Zurʿa*. Ed. Gérard Jehamy and Rafīq al-ʿAjam. Beirut, 1994.

Secondary sources

Afnan, Soheil M. 1969. *A philosophical lexicon in Persian and Arabic*. Beirut: Dar El-Machreq.

Elamrani-Jamal, Abdelali. 1983. *Logique aristotélicienne et grammaire arabe*. Paris: J. Vrin.

Georr, Khalil. 1948. *Les* Catégories *d’Aristote dans leurs versions syro-arabes*. Beirut: Institut Français de Damas.

Jeremiás, Éva M. 2002. “*Rābiṭa* in the classical Persian literary tradition: The impact of Arabic logic on Persian”. *Jerusalem Studies in Arabic and Islam* 27.550–574.

Maróth, Miklós. 1994. *Die Araber und die antike Wissenschaftstheorie*. Leiden: E.J. Brill.

Versteegh, C.H.M. (Kees). 1993. *Arabic grammar and Qurʾānic exegesis in early Islam*. Leiden: E.J. Brill.

——. 1997. *The Arabic linguistic tradition*. London and New York: Routledge.
Wolfson, Harry Austryn. 1976. *The philosophy of the* Kalam. Cambridge, Mass., and London: Harvard University Press.

Miklós Maróth
(Pázmány Péter Catolic University)

Mahmūsa → Majhūra/Mahmūsa

Maintenance → Language shift: Amazigh

Majāz

The term *majāz*, used both in Arabic linguistic thought and in the hermeneutics of the *Qurʾān*, is usually translated as 'trope', although its meaning is not completely congruent with the Western concept. *Majāz* is explained as a verbal noun of *jāza* 'to go beyond something', in the sense of a participle denoting *al-kalimatu al-jāʾizatu ʾay al-mutaʿaddiyatu makānahā l-ʾaṣliyya* 'a word that goes beyond its original place [i.e. its literal meaning in the language system]' (cf. Jurjānī, *ʾAsrār* 365; Mehren 1853:75). In the history of the term, a semantic specialization or narrowing from denoting all kinds of rhetorical features to some kinds of tropes can be observed. Whether *majāz* should be translated as 'trope' depends on the author examined.

Tašbīh and *majāz*, along with → *kināya*, are the main topics of the *ʿilm al-bayān*, the second branch of Arabic rhetoric as systematized in the 12th and 13th centuries, metaphor (*istiʿāra*) being the most important part of *majāz*. In time, metaphor became the predominant subject in the works of those who were engaged in building up a theoretical framework for Arabic rhetoric, and, thus, the study of *majāz* is essentially the study of metaphor (→ *istiʿāra*).

In Arabic language theory, tropical speech is closely linked with the notion of transference (*naql*) of meaning. Functioning as a unifying factor, this notion first complements and then almost replaces both the concepts of metaphor as a 'borrowing' of a name or a thing, and *majāz* as a licensed transgression. Another notion that *majāz* is connected with is its counterpart *ḥaqīqa*. Both are used in the discourse about the question of whether all passages of the *Qurʾān* have to be taken at face value. The idea of God sitting on a throne is true, of course, in a sense, but does it mean the same as in normal language? Given that it is the truth (*ḥaqīqa*), the question remains whether it is a veridical expression (*ḥaqīqa*) or a figurative one. Thus, in *kalām* a discourse developed about the connection between the ontological concept of *ḥaqīqa* and the modes of expressing reality in language. While *majāz* is rooted in the discussion of features of the language, *ḥaqīqa*, mostly denoting transcendental truth or the real nature of a thing, has an epistemological background. In his commentary on the *Qurʾān*, Muḥammad al-Kalbī (d. 146/763) mentions *ḥaqīqa* and *majāz* along with a number of other dichotomies (cf. Versteegh 1993:106, 122).

Both *ḥaqīqa* and *majāz* may denote the word as a formal unit (→ *lafḏ̣*) and the corresponding concept (→ *maʿnā*), and sometimes the referent as well. But, since the dichotomy form/meaning is firmly grounded in Arabic linguistic reasoning, the danger of confusing these levels is even less likely than in the Western tradition. Moreover, theorists from the 12th century were fairly aware of the fact that what words refer to are mental representations.

From the very beginning, grammarians of Arabic were conscious of deviant forms and rhetorical elements in language. Dealing with → poetic license, Kufan grammarians like al-Farrāʾ (d. 207/822) employ the term *ʾijāza*, a derivate of the same root *j-w-z* that *majāz* belongs to. Sībawayhi's (d. 180/793) expression *saʿat al-kalām* suggests an interpretation as a semantic expansion beyond the common function in the language system. Sībawayhi uses the term *ittisāʿ*, referring to cases like *isʾal-i l-qaryata* 'ask the village!', instead of *isʾal ʾahla l-qaryati* 'ask the people of the village!' (Q. 12/82), a standard example and constant companion throughout the history of Arabic linguistic thinking. This case is not an ordinary elision, as the sentence is in no way grammatically defective. One has to consider the meaning and the compatibility of components beyond formal aspects of the language. As-Sīrāfī, examining the passage in his commentary on Sībawayhi's *Kitāb*, uses the terms *ḥaqīqa* and *majāz* (cf. Versteegh 1990:281), a usage gradually becoming prevalent. Expansion (*tawassuʿ*) as a basis for metaphorical wording is still discussed in Islamic jurisprudence, even in our times (cf. Modarressi 1986:790).

More explicitly dedicated to linguistic exegesis than Ibn al-Kalbī, ʾAbū ʿUbayda (d. 210/825), following the order of the verses, writes his *Majāz al-Qurʾān* as a contribution to preserving the proper understanding of the Holy Scripture. This motivation prevails through the centuries and must be seen as the generally accepted background for research. ʾAbū ʿUbayda does not couple *majāz* with *ḥaqīqa*. At the beginning of his work, he lists 39 modes or means of expression, to be found in the *Qurʾān* and elsewhere. Several kinds of derivations, elisions, paraphrastic expressions, tropes, and the like are treated as *majāz*. Yet, there is no elaborate technical vocabulary in ʾAbū ʿUbayda's work. Without defining terms, he briefly characterizes the linguistic convention or means he wishes to present, or he limits himself to exemplifying it. As a result, later theorists often read his examples in a different way. *Majāz* in this context seems to stand for 'another way to say it', according to the standards of those who are learned in the Arabic tongue, or 'to say it in other words', using the means of the Arabic language to color speech, as applied in the *Qurʾān* and poetry, and also by the pure Arabs living in the time of the Prophet. *Majāz* is predominantly used for rhetorical devices. In some cases, however, it is the regular expression that is intended. On several occasions, the meaning of *majāz* comes close to *wujūh*. For ʾAbū ʿUbayda, the variety of expressions in the unsurpassable *Qurʾān* (→ *ʾiʿjāz*) is evidence of the possibilities and virtues of the Arabic language itself.

A number of his successors use the term *majāz* in a generic way. Al-Jāḥiẓ (d. 255/868) calls every use of words going beyond their conventional function in language *majāz*. It has a wider scope than metaphor and seems to cover everything that is not a veridical expression, and therefore forms a counterpart of *ḥaqīqa*. Ibn Qutayba (d. 276/889) employs *majāz* to cover various features, such as elision, and other means to make speech more concise or to extend it by paraphrasing. *Istiʿāra*, which has a wider meaning than in later times, is also included. Early writers in the field of literary theory, like Ṯaʿlab (d. 291/904), do not use the term (cf. Heinrichs 1991:271). Ibn al-Muʿtazz (d. 296/908) does not mention *majāz* in his *Kitāb al-badīʿ* among the figures of speech he presents. This may be an indication that it was not in use in literary circles, where the lines of the modern poets of his time were discussed. Qudāma ibn Jaʿfar (d. between 328/939 and 337/948) seems to apply the term in the same way that ʾAbū ʿUbayda did. The *majāz/ḥaqīqa* dichotomy is mentioned by Ibn Ṭabāṭabā (d. 322/934). ʾIsḥāq ibn ʾIbrāhīm (d. middle of the 4th/10th century) uses the term along with *istiʿāra* in the sense of figurative expression. Al-Ḥātimī (d. 388/998) finally comes to apply it to all kinds of deviant use of language found in poetry.

In the *Kitāb aṣ-ṣināʿatayn* of al-ʿAskarī (d. after 395/1004), no clear distinction is made between *istiʿāra* and *majāz*. The definition given for metaphor as a motivated transfer (*naql*) of an expression from its normal use in language to another seems to be intended to also cover *majāz* (*Ṣināʿatayn* 295). Most of the numerous examples are followed by some stereotyped formulas declaring metaphorical speech superior to veridical wording (*ʾablaġ min al-ḥaqīqa*).

In aṯ-Ṯaʿālibī's (d. 429/1038) *Sirr al-ʿarabiyya* (e.g. *Sirr* 397), a catalog of various features assigned to the Arabic tongue, another case of *majāz* used in a less systematic way is encountered. Quoting al-Jāḥiẓ, he states that the Bedouin have a linguistic habit that makes it easy for them to speak in tropes and the like, for they can always count on being understood. *Majāz* in this sense points at the creativity possible in speech, i.e. how far one can digress from veridical meaning, if only one can rely on the other participants' communicative competence to figure out the intended meaning. Ibn Rašīq (d. 456/1063 or 463/1070) is among those who have a broader concept of *majāz* that includes metonymy and even simile.

ʿAbd al-Qāhir al-Jurjānī's (d. 471/1078) main works, *ʾAsrār al-balāġa* and *Dalāʾil al-ʾiʿjāz*, are both milestones in the history of Arabic rhetoric, the former being an analysis of tropical speech while the latter is mainly dedicated to the idea of *naḏm*, i.e. generating semantic nuances by word combination and construction. He establishes the distinction between *majāz* as a kind of trope formed by a single word used in a sense other than its conventional meaning (*majāz luġawī*), on the one hand, and as a trope on the sentence level (*majāz ʿaqlī*), on the other. In the latter case, the words retain their original lexical meaning, and the trope lies in the affirmative predication (*ʾiṯbāt*) alone, not in the thing ascribed, e.g. *wa-šayyaba ʾayyāmu l-firāqi mafāriqī* 'the days of separation have made my head white'. Here, religious considerations come into play, for it is God who makes hair turn white. In cases like *fa-ʾaḥyaynā bihi l-ʾarḍa baʿda mawtihā* 'We made the earth living after it had been dead' (Q. 35/9), the trope is due to the thing ascribed. In the given verse, 'liveliness' stands for the freshness and bountifulness of the earth turning green, while the predication is a veridical one.

Simile, analogy, and metaphor are considered the roots of eloquence (Jurjānī, *ʾAsrār* 26). The implied comparison of two things on the word level forming an *istiʿāra*, e.g. 'I saw a lion', is paralleled by a comparison of two sets of things on the sentence level forming an analogy (*tamṯīl*), e.g. 'the horses and riding camels of youth's passion are unharnessed'. In contrast with the former example, here 'horses and riding camels' do not refer to any objects in the real world.

Al-Jurjānī's concept of *majāz*, and especially metaphor, is based on what he calls *mulāḥaḏa* (cf. *ʾAsrār* 325), a kind of immediate blending of the two mental images that the word evokes in conventional and metaphorical usage. This mental visualization may be strong and impressive, as in the lion example, or may evoke a less vivid association, such as 'hand' for 'benefit'. At any rate, in al-Jurjānī's conception this manner of indirect and abstract thinking is at the same time the most concrete one. Evoking a more concrete representation of meaning, *majāz* as a graphic description enters the hearer's mind in a more effective way.

Al-Jurjānī is not the first to interpret metaphor as an intensified comparison. He cites al-Qāḍī al-Jurjānī and the lexicographer Ibn Durayd as his predecessors. However, unlike his predecessors, al-Jurjānī points out that the concept of *majāz* is a broader one and not identical with metaphor. Every metaphor is a trope, but not every trope a metaphor. Al-Jurjānī suggests that dimensions other than similarity are to be excluded from the concept, thereby introducing a new aspect into the discussion. Consequently, two meanings connected on the basis of special property (*ixtiṣāṣ*), contact (*mulābasa*), or inextricable combination (*xalṭ*) cannot be taken as metaphor. However, this dictinction is not observed consistently.

The notorious disagreement about whether a metaphorical interpretation of the *Qurʾān* is admissible or not is paralleled by a thread of discussion in another field, namely that about the relation between poetry and lie. Some circles held that art would flourish best when fantasy is not limited by any restrictions, however far from reality mental pictures may be. This position found its expression in the dictum 'The finest poetry is the most untrue'. Others held that 'The finest poetry is the most truthful' and wished to keep poetry free from exaggeration and excessive use of rhetorical devices (*tajawwuz*). Al-Jurjānī tends toward the latter opinion but refuses a general decision. He states that metaphor has more in common with ellipsis than with the construction of fantastic worlds without any connection to reality. This could be seen in the *Qurʾān*, whose metaphors always have a substratum of reality. Again, reflections on language are related to religious arguments.

Majāz, as presented by al-Jurjānī, is a communicative strategy meant to go beyond the limits of conventional language and create new fields of associations. It operates on both the emotional and cognitive levels. Rhetorical communication is a process of encoding and decoding, with the words merely giving a vague indication and the listener being challenged to infer what is intended on the basis of his knowledge of context and situation.

The idea of a second meaning behind the original one is not restricted to the single-word level. Al-Jurjānī is very aware of the fact that not only single words but also whole classes of sentences can be used at odds with their normal function in the language system. Used as rhetorical questions, for example, → interrogative sentences may serve several communicative goals, such as ordering or blaming. Albeit not under the heading of *majāz*, al-Jurjānī deals in his *Dalāʾil* with questions that prepared the ground for what was later to be called *ʿilm al-maʿānī*. Known in his time mostly as a grammarian, he also discusses the functioning of *majāz* in terms of grammar.

Although in his works the methodological influence of the *mutakallimūn* has not yet completely superseded literary criticism and aesthetics, al-Jurjānī, an ʾAšʿarite, eventually promoted Muʿtazilite thinking. He laid down the foundations of a system upon which his successors built a firm edifice of theory.

Faxr ad-Dīn ar-Rāzī (d. 606/1209) dedicates the first of the two parts of his *Nihāyat al-ʾījāz* mainly to the linguistic unit of a single word and to the parts of speech, while the second deals with the combination of words on the sentence level. As a preliminary to simile and metaphor, he discusses *majāz* and finally adds → *kināya* as a figure in its own right.

Following al-Jurjānī, ar-Rāzī defines *majāz* as a word with a meaning other than its original one, because some relationship between the first and the second meaning is to be brought to one's eyes (*mulāḥaḏa nisba; Nihāya* 49). Ar-Rāzī discusses the relation between veridical and tropical speech and also the case of speech that is both simultaneously, and he makes some efforts to distinguish *majāz* from lying or talking nonsense. He stresses that *majāz* is the general term of which *istiʿāra*, being further subdivided according to various aspects, is one kind. The *qarya* example of Q. 12/82, again, corresponds to an elliptical *majāz* (*majāz bi-n-nuqṣān*), and its counterpart, *majāz bi-z-ziyāda*, too, is discussed not primarily as a grammatical phenomenon.

Ar-Rāzī contributes to the theory of *majāz* by introducing notions from logic and advocating a methodological approach that focuses on the way things are referred to by *dalāla waḍʿiyya* or *dalāla ʿaqliyya*. In his work, the hitherto achieved progress in rhetorical analysis meets a philosophically inspired theory of language that is in some respects of Greek origin (cf. Bauer 2005:292). Basic elements of this thinking found their way into the science of *ʾuṣūl* and became an accepted part of the framework. As one of the outstanding exponents of the *ʾuṣūl*, ar-Rāzī combined the methods of this science with the tradition of rhetoric represented by al-Jurjānī.

As-Sakkākī (d. 626/1229) goes a long way toward a unifying theory of language in his *Miftāḥ al-ʿulūm*. The author establishes a system of classes following the well-known scheme of *genus proximum* and *differentia specifica*. To start with, he defines *ḥaqīqa* as a word used in its original meaning without further interpretation, *al-kalimatu l-mustaʿmalatu fī-mā hiya mawḍūʿatun lahu min ġayri taʾwīlin fī l-waḍʿ* (*Miftāḥ* 358). A word may be used (i) according to its original meaning (*luġawī*); or (ii) as a legal term (*šarʿī*); or (iii) according to a special terminological or generally known and well-established convention (*ʿurfī*). Hence, the trope must be realized within one and the same mode, otherwise, every terminological use would constitute a *majāz*. This restriction is important because many words in Arabic, the language of science in the East, have several specific meanings, depending on the branch of learning referred to. If you call a brave man a lion, the meaning is transferred on the level of two words of nonterminological usage. If *ṣalāt*, which designates the obligatory ritual prayer, stands for *duʿāʾ*, i.e. private prayer, again, the transference remains on the same plane of defined terms of legal relevance. The third mode of expression refers to cases like the use of *dābba* '(creeping) animal' for a mule, being a shift from the lexical mode to a secondary conventional usage.

Then, as-Sakkākī defines *majāz* as *al-kalimatu l-mustaʿmalatu fī ġayri mā hiya mawḍūʿatun lahu bi-t-taḥqīqi istiʿmālan fī l-ġayri bi-n-nisbati ʾilā nawʿi ḥaqīqatihā maʿa qarīnatin māniʿatin ʿan ʾirādati maʿnāhā fī ḏālika n-nawʿi* 'a word that is used for something other than its original veridical meaning, both meanings pertaining to the same mode of expression [i.e. *luġawī, šarʿī*, or *ʿurfī*], while the context renders the word impossible to be intended in the literal sense'.

As-Sakkākī gives the following summary of the basic notions of his system (cf. Simon 1993:44): a word gives information (i) through its original meaning in the language system, or (ii) through an implication of that meaning. When using the word, one either intends (i) its lexical meaning alone, or (ii) something else, or (iii) both the one and the other. The first case is *ḥaqīqa*; the second is *majāz* as a single word figure, requiring a situational or contextual indication that blocks literal meaning; the third case is → *kināya*.

Both (i) and (iii) are considered veridical expressions, but they differ in the way they supply the information, either straightforwardly or in an indirect way. *Majāz* either substitutes an intensified comparison (*mubālaġa fi t-tašbīh*) for the literal meaning, thus being a metaphor (*istiʿāra*), or it does not, in which case it is a 'free' trope (*majāz mursal*). Metaphor is seen as an implied comparison mentioning either the thing compared, or the thing compared with, but not both at the same time. If the thing compared is named, it is the case of an *istiʿāra bi-l-kināya*, where the *mušabbah* (i.e. *mustaʿār lahu*) is given a characteristic trace of the *mušabbah bihi* (i.e. *mustaʿār minhu*), e.g. 'I saw heads that had grown ripe' in al-Ḥajjāj's famous inaugural speech in Kūfa. If the thing compared with is mentioned, as-Sakkākī speaks of an *istiʿāra bi-t-taṣrīḥ*, e.g. 'I saw a lion', either in the form of an *istiʿāra taḥqīqiyya* naming a real quality of the omitted compared thing, or of an *istiʿāra taxyīliyya* in cases where the thing compared with does not belong to the compared thing in reality.

The background for the whole system is a theory of referentiality similar to what ar-Rāzī taught. The intended meaning of a word may be the whole of its conventional meaning (*dalālat al-muṭābaqa*) or a part of it (*dalālat at-taḍammun*), or the word is used as a reference to something outside the literal meaning, yet connected to it (*dalālat al-iltizām*).

In as-Sakkākī's view, the division into *majāz luġawī* and *majāz ʿaqlī* is basically the wrong approach. Every *majāz* is *luġawī*, in his opinion. He advocates a semantic concept that comprises two meanings of a word. A wider range denotes all corporal and incorporeal qualities of the thing, while a reduced one stands for the incorporeal qualities alone. These abstract attributes might come to the speaker's mind when claiming an identity between the two things implicitly compared.

Kināya is different from *majāz* in two respects. First, it does not exclude the veridical communication and thus allows a double understanding; and second, the direction of inference is from consequence (*lāzim*) to origin (*malzūm*).

Although being more interested in building a system than in the detailed explanation of all facets of a given expression, as-Sakkākī holds that using *majāz* makes communication more impressive and effective (*ʾablaġ*), because it operates on different levels and a multiplicity of channels, and therefore bears more information than straightforward veridical speech.

The formation of a standard system of Arabic rhetoric theory did not exclude terminological and classificatory differences. For example, Ḍiyāʾ ad-Dīn Ibn al-ʾAṯīr (d. 637/1239) divides *majāz* into simile and expansion (*tawassuʿ*; cf. Modarressi 1986:789), and Badr ad-Dīn Ibn Mālik (686/1287) does not use the term *majāz mursal* in his *Miṣbāḥ* when dealing with *majāz*.

As-Sakkākī's commentator Jalāl ad-Dīn al-Qazwīnī (d. 739/1338), again, presents in his *ʾĪḍāḥ* a tripartite arrangement of the *ʿilm al-bayān*. The first section deals with simile, followed by a section dedicated to *ḥaqīqa* and *majāz*, and finally *kināya*. The second part contains discussions of (i) *majāz mursal*; (ii) *istiʿāra*, including the *istiʿāra bi-l-kināya* and *istiʿāra taxyīliyya*; and (iii) *majāz murakkab*, i.e. the compound trope focusing on the predication, along with a critique of as-Sakkākī's ideas. Al-Qazwīnī does not reject the idea of *majāz ʿaqlī* but assigns it – with constructions like the *qarya* example (Q. 12/82) in mind – to the field of rhetoric called *ʿilm al-maʿānī* rather than to *ʿilm al-bayān*.

Al-Qazwīnī points out that the 'free' trope is defined *per negationem*. Its common feature is simply that it is not built upon the relationship of similarity between the conventional and the tropical meaning of the word. Nine relations are listed as categories of *majāz mursal*: the synechdochical relations (i) *pars pro toto* and (ii) *totum pro parte*, (iii) the cause/consequence, and (iv) consequence/cause relation, (v) applying a specification that does not fit anymore, or (vi) will fit in future times, (vii) place/event that happens in the place relation, or (viii) vice versa, and finally (ix) the tool/purpose relation. At least fourteen more relations are added by commentators (cf. Mehren 1853:81). Principally, there is no limit to finding additional relations that might form a *majāz mursal*. It is worth noting that *majāz mursal* is subdivided mostly along the fundamental dimensions of time, space, quantity, and causality. These dimensions form the framework for mental operations, and speakers are used to relate things to each other within that framework. Thus, inferring hidden connections between two things along those lines has more in common with logical deduction than with the capability of association and fantasy. Compared with a true metaphor encoding a thought through metonymy or decoding, it is apparently seen as a less creative process.

With *majāz ʿaqlī* being finally reintegrated into the system, in aṭ-Ṭībī's (d. 743/1342) *Tibyān* a very clear structure has been established: *majāz* is subdivided into *luġawī*, comprising the 'informative' (*mufīd*) or 'not informative' free trope and metaphor, on the one hand, and *majāz ʿaqlī* on the other. Al-Jurjānī called an exchange of words with the same basic meaning, e.g. 'nose' and 'camel's nose', a 'not informative metaphor' (*istiʿāra ġayr mufīda*). Considering the status of the two parts of the predicative relation, *majāz ʿaqlī* is of four kinds: (i) both parts are *ḥaqīqī* and *waḍʿī*, e.g. *ʾanbata r-rabīʿu l-baqla* 'spring made the green grow'; or (ii) neither part is *ḥaqīqī* or *waḍʿī*, e.g. *ʾaḥyā l-ʾarḍa šabābu z-zamān* 'the youth of time made the earth alive'; or (iii) the predicate (*maḥkūm bihi*) is a *ḥaqīqa waḍʿiyya*, e.g. *ʾanbata l-baqla šabābu z-zamān* 'the youth of time made the green alive'; or (iv) vice versa, e.g. *ʾaḥyā r-rabīʿu l-ʾarḍa* 'spring made the earth alive' (*Tibyān* 259). Avoiding the term *majāz ʿaqlī*, as-Sakkākī had already classified predicative expressions according to the four combinations resulting from the fact that both the predicate (*maḥkūm bihi*) and the subject (*maḥkūm lahu*) of a predication can be veridical or tropical words.

Paradoxically, *majāz* sometimes happens to be treated on the same hierarchical level as other figures, while being used as a cover term at the same time. In a late source of the *badīʿ* tradition, *majāz* is subdivided into two classes: (i) applying the name of one object to another (*ḏikr aš-šayʾi bi-smi ġayrihi*), for which 'bottle' instead of 'a drink in a bottle', i.e. a case of *majāz mursal*, is cited as an example; and (ii) attributing to an object a characteristic belonging to another (*ʾiṯbāt mā li-ġayri š-šayʾi lahu*), e.g. personifying stars by assigning them qualities of living creatures (cf. Cachia 1998:65). This wording is reminiscent of early definitions of *istiʿāra*. Moreover, it reminds one of ar-Rāzī's definition of the two main categories of metaphor, although it lacks the addition "for the purpose of an intensified comparison".

A variety of terms are used to point at some specific qualities of the trope, e.g. a *majāz at-tašbīh* is a simile without a particle of comparison. One may find further diversification of the terminological vocabulary, such as *majāz ʾisnādī* or *majāz at-tarkīb* in the sense of *majāz ʿaqlī*, which may also be named *majāz ḥukmī* or *majāz fī l-ʾiṯbāt*.

Since the free trope comprises inferring reason from consequence, a partial overlapping of the categories *majāz mursal* and *kināya* seems to be a weak point of the system. As *kināya* is sufficiently defined by the two conditions mentioned above, the problem lies in the part of the free trope. However, *kināya* may imply a kind of creative reasoning over a chain of several intermediate consequences to the final deduction of the origin or antecedents. This may be the reason why al-Qazwīnī interprets an allusion to extraordinary hospitality like *zaydun kaṯīru r-ramādi* 'Zayd has a great deal of ash' as *kināya*, whereas 'wine' for 'grapes' is classed with *majāz mursal*.

It may be noted that in classical Western tradition, too, metaphor is usually interpreted as a result of transference (cf. Eggs 2001:1106). This corresponds to the notion of *naql* in Arabic theory. The Stoa knows three modes of transference: *similitudo*, *vicinitas*, and *contrarium*. Similarity covers metaphor, allegory, and the like, vicinity matches with antonomasia, metonymy, and synecdoche, and contrariness is associated with irony. While similarity and contiguity are parameters in the Arabic tradition as well, contrariness, being the underlying principle of irony, is not a parameter in classification. The strong connection between systematized rhetoric and Qurʾānic studies may be one of the reasons for this. Emerging from Qurʾānic studies, the *ḥaqīqa/majāz* problem has an ontological dimension which is absent in literary contexts, for it goes without saying that God's own word does not comprise any faults and mistakes, neither incorrect grammatical constructions nor tropical expressions of dubious quality.

As al-Jurjānī points out, there are two attitudes toward the use of *majāz*, both equally erroneous (cf. *ʾAsrār* 361). Some deny its existence as such and even refuse to take it into consideration. For them, it is a religious obligation to hang on to the literal meaning of the words. Hence, they take literally verses like *jāʾa rabbuka* 'your Lord comes' (Q. 89/22), or *ar-raḥmānu ʿalā l-ʿarši stawā* 'the Merciful sits down upon the throne' (Q. 20/5; cf. 7/54, 10/3, 13/2), despite the fact that movement in space and change of place can only be stated for a corporal entity and not for the incorporeal creator of space itself. Others mistake almost all expressions for figurative speech and make it a habit to go to extremes by applying far-fetched explanations to unambiguous passages. According to al-Jurjānī, these people are unaware of the fact that God would not have made the *Qurʾān* a miracle if the Holy Book were notoriously enigmatic, cryptic, incomprehensible, and unclear. On the contrary, the scripture has been revealed in clear Arabic language. In ʿAbd al-Qāhir's view, a thorough knowledge of the features and patterns of the Arabic language allows a valid interpretation of the text. Moreover, the methods of linguistic and hermeneutical analysis are not the secret knowledge or exclusive possession of some chosen ones but are accessible to everyone.

A disapproving attitude toward metaphorical interpretation is ascribed to Mālik ibn ʾAnas and *grosso modo* shared by Ibn Ḥanbal and al-ʾAšʿarī, the latter admitting the possibility of figurative language in the *Qurʾān* at least. This should not be confused with inclinations toward anthropomorphism, as the adherents of anthropomorphism who draw parallels between the human sphere and God (*tašbīh*) turn out not to have been influential in Islamic discourse.

On the contrary, any comparison with human attributes must be excluded. God has hands, in fact, but of course they are not the same as human hands. Every attempt to compare such attributes with whatever physical qualities must be banned. Thus, we do not know what kind of quality is ascribed. The idea of God sitting down on a throne is not at all absurd. According to these authors, we do not know exactly what it means but have to believe it without further asking (*bi-lā kayfa*). Reasoning about ambiguous (*mutašābih*) passages cannot go beyond speculation and is therefore worthless.

Thinkers of the opposite school, later to be called Muʿtazila, take the 'hands of God' as a symbol for his power and grace. They, too, reject anthropomorphism but see the concept of *majāz* as an interpretive tool to preclude it.

Even more frequently than in the *Qurʾān*, anthropomorphism occurs in *Ḥadīṯ* and in other sayings that give rise to the danger of denying God's absolute power, for instance when the Prophet Muḥammad is called "the lawgiver (*aš-šāriʿ*)".

However, both philosophers like Ibn Rušd and the adherents of Sufism hold that the concreteness and plasticity of language help those who are unable or unwilling to engage in abstract thinking. Moreover, metaphorical speech cannot be altered without suffering a loss of specific information. Some believe the use of tropes to be a bridge to the real nature of things (*al-majāzu qanṭaratu l-ḥaqīqa*). Extreme mystics hold that every word has an inner sense to be detected, the outer sense being a figurative expression for the real meaning.

Epistemological considerations from the *ʿilm al-kalām* helped in the formation of the methodological apparatus of the principles of religious law. The classical form of the *ʾuṣūl al-fiqh* contains a good deal of Muʿtazilī ideas, and even those who argue against it employ methods borrowed from them. The state of the discourse about *majāz* within the *ʿilm al-ʾuṣūl* around the middle of the 10th century is reflected in the work of al-Jaṣṣāṣ (d. 370/981). Compared to aš-Šāfiʿī's *Risāla*, the combination of linguistic analysis and hermeneutical exegesis has come to a new stage. In an extensive chapter on *majāz*, al-Jaṣṣāṣ enumerates six subcategories, such as ellipsis and metaphor, sometimes without having a clear term for the case (cf. Heinrichs 1991:258–270). Here, the idea of an indication (*dalīl*), later known as *qarīna*, that blocks the literal meaning and leads to the tropical one, is part of the explanation. The history of the *ʾuṣūl al-fiqh* documents the increasing influence of communication theory on the hermeneutical process. In ar-Rāzi's *Maḥṣūl*, the section about *ḥaqīqa* and *majāz* is mainly dedicated to semantics and the analysis of linguistic expression. He quotes a definition, most likely by ʾAbū l-Ḥusayn al-Fārisī, with whom he agrees, *mā ʾufīda bihi maʿnan muṣṭalaḥun ʿalayhi ġayru mā ṣṭuliḥa ʿalayhi fī ʾaṣli tilka l-muwāḍaʿati llatī waqaʿa t-taxāṭubu bihā li-ʿalāqatin baynahu wa-bayna l-ʾawwali* (*Maḥṣūl* I, 112).

Although a number of disciplines, such as grammar, literary criticism, and philosophy, have made contributions to the development of rhetoric as a system, its status as a semiotic theory comes from the decisive influence of hermeneutics of the *Qurʾān* and the methodology of religious law. Finally, a kind of regrammaticalization of rhetoric, already identifiable in ar-Rāzī's work, entered the handbooks. In modern times, Arabic linguists reconsider the classical heritage and try to integrate aspects of the old, and remarkably fresh, rhetorical theory into contemporary research.

Bibliographical references

Primary sources

ʾAbū ʿUbayda, *Majāz* = ʾAbū ʿUbayda Maʿmar ibn al-Muṯannā, *Majāz al-Qurʾān*. Ed. Fuat Sezgin. 2 vols. Beirut: Muʾassasat ar-Risāla, 1981.

ʿAskarī, *Ṣināʿatayn* = ʾAbū Hilāl al-Ḥasan ibn ʿAbdallāh ibn Sahl al-ʿAskarī, *Kitāb aṣ-ṣināʿatayn*. Ed. Mufīd Qumayḥa. Beirut: Dār al-Kutub al-ʿIlmiyya, 1984.

Jurjānī, *ʾAsrār* = Majd ad-Dīn ʾAbū Bakr ʿAbd al-Qāhir ibn ʿAbd ar-Raḥmān al-Jurjānī, *ʾAsrār al-Balāġa*. Ed. Helmut Ritter. Beirut: Dār al-Masīra, 1983. [German transl., Helmut Ritter, *Die Geheimnisse der Wortkunst des ʿAbdalqāhir al-Curcānī*. Wiesbaden: F. Steiner, 1959.]

——, *Dalāʾil* = Majd ad-Dīn ʾAbū Bakr ʿAbd al-Qāhir ibn ʿAbd ar-Raḥmān al-Jurjānī, *Dalāʾil al-ʾiʿjāz*. Ed. Muḥammad Rašīd Riḍā. Beirut: Dār al-Maʿrifa, 1978.

Qazwīnī, *ʾĪḍāḥ* = Jalāl ad-Dīn ʾAbū ʿAbdallāh Muḥammad ibn ʿAbd ar-Raḥmān al-Qazwīnī, *al-ʾĪḍāḥ fī ʿulūm al-balāġa*. Ed. Muḥammad al-Xafājī. Beirut: Dār al-Kitāb al-Lubnānī, 1985.

Rāzī, *Maḥṣūl* = Faxr ad-Dīn ʾAbū ʿAbdallāh Muḥammad ibn ʿUmar ar-Rāzī, *al-Maḥṣūl fī ʿilm ʾuṣūl al-fiqh*. 2 vols. Beirut: Dār al-Kutub al-ʿIlmiyya, 1988.

——, *Nihāya* = Faxr ad-Dīn ʾAbū ʿAbdallāh Muḥammad ibn ʿUmar ar-Rāzī, *Nihāyat al-ʾījāz fī dirāyat al-ʾiʿjāz*. Cairo: Maktabat al-ʾĀdāb, 1909.

Sakkākī, *Miftāḥ* = Sirāj ad-Dīn ʾAbū Yaʿqūb Yūsuf ibn ʾAbī Bakr as-Sakkākī, *Miftāḥ al-ʿulūm*. Ed. Naʿīm Zarzūr. Beirut: Dār al-Kutub al-ʿIlmiyya, 1983.

Ṯaʿālibī, *Sirr al-ʿarabiyya* = ʾAbū Manṣūr ʿAbdallāh ibn Muḥammad aṯ-Ṯaʿālibī, *Fiqh al-luġa wa-sirr al-ʿarabiyya*. Ed. Sulaymān Salīm al-Bawwāb. Damascus: Dār al-Ḥikma, 1984.

Ṭībī, *Tibyān* = Šaraf ad-Dīn Ḥusayn ibn Muḥammad aṭ-Ṭībī, *at-Tibyān fī ʿilm al-maʿānī wa-l-badīʿ wa-l-bayān*. Ed. H.ʿA.M. al-Hulālī. Beirut: ʿĀlam al-Kutub, 1987.

Secondary sources

Bauer, Thomas. 2005. "Rhetorik der arabischen Kultur". *Rhetorik: Begriff, Geschichte, Internationalität*, ed. Gert Ueding, 283–300. Tübingen: Niemeyer.

Cachia, Pierre. 1998. The arch rhetorician *or* The schemer's skimmer: *A handbook of late Arabic* badīʿ *drawn from ʿAbd al-Ghanī an-Nābulsī's* Nafahāt al-Azhār ʿalā Nasamāt al-Ashār. Wiesbaden: O. Harrassowitz.

Eggs, Ekkehard. 2001. "Metapher". *Historisches Wörterbuch der Rhetorik*, V, ed. Gert Ueding, 1099–1186. Tübingen: Niemeyer.

Heinrichs, Wolfhart. 1984. "On the genesis of the *ḥaqīqa–majāz* dichotomy". *Studia Islamica* 59.111–140.

——. 1991. "Contacts between scriptural hermeneutics and literary theory in Islam: The case of *majāz*". *Zeitschrift für Geschichte der Arabisch-islamischen Wissenschaften* 7.253–284.

Mehren, A.F.M. von. 1853. *Die Rhetorik der Araber*. Copenhagen and Vienna. (Repr., Hildesheim: G. Olms 1970.)

Modarressi, Hossein. 1986. "Some recent analyses of the concept of *majaz* in Islamic jurisprudence". *Journal of the American Oriental Society* 106:4.787–791.

Simon, Udo. 1993. *Mittelalterliche arabische Sprachbetrachtung zwischen Grammatik und Rhetorik:* ʿIlm al-maʿānī *bei as-Sakkākī*. Heidelberg: Heidelberger Orientverlag.

Versteegh, Kees. 1990. "Freedom of the speaker? The term *ittisāʿ* and related notions in Arabic grammar". *Studies in the history of Arabic grammar*, II, ed. Michael G. Carter and Kees Versteegh, 281–293. Amsterdam and Philadelphia: J. Benjamins.

——. 1993. *Arabic grammar and Qurʾānic exegesis in early Islam*. Leiden: E.J. Brill.

Udo Simon (University of Heidelberg)

Majhūra/Mahmūsa

The terms *majhūra/mahmūsa* denote a phonological correlation, generally held to correspond to the opposition between voiced/voiceless (e.g. Schaade 1911:13), or to that between lenis/fortis phonemes (for a discussion of the difference between these two oppositions, see Ladefoged and Maddieson 1996:95–99). Both terms were first used by Sībawayhi (d. 175/791) in his *Kitāb* (II, 405). He probably based his observations on earlier research by his predecessors. However, his teacher al-Xalīl ibn ʾAḥmad (d. 175/791) does not mention this classification in the *Kitab al-ʿayn*, although he does seem to know the terminology because he defines *hams* (*ʿAyn* IV, 10) as "the perception of sound in the mouth without mixing with sound from the chest, and without loudness in speaking, but it is a whispered sound in the mouth, like a secret" (*al-hams ḥiss aṣ-ṣawt fī l-fam mimmā lā ʾišrāba lahu min ṣawt aṣ-ṣadr wa-lā jahārata fī l-manṭiq wa-lākinnahu kalām mahmūs fī l-fam ka-s-sirr*). Nonetheless, he does not use it in his classification of sounds in the phonetic introduction (Talmon 1997:131). The terms are not used, either, in the Kufan tradition, for instance in al-Farrāʾ's *Maʿānī l-Qurʾān*.

It cannot be excluded that Indian phonetic theory influenced the formation of the Arabic description of the process (cf. Danecki 1985). The starting point of Sībawayhi's definition is the concept of *iʿtimād*, which may be understood as a phonatory effort or simply phonation, i.e. acoustic energy. This somewhat puzzling term was used only by him, and was dropped in later Arabic theory. The term may well derive from Indian phonetic theory, where phonation was called *prayatna* (but cf. Law 1990).

Sībawayhi defines *majhūra* as phonemes in the articulation point of which the phonatory effort is full (*ʾušbiʿa l-iʿtimād fi mawḍiʿihi; Kitāb* II, 405.19–20), while in *mahmūsa* it is reduced (*ʾuḍʿifa; Kitāb* II, 405.23). The following phonemes he regarded as *majhūra: ʾ, ʾalif, ʿ, ġ, q, j, y, ḍ, l, n, r, ṭ, d, z, ḏ̣, ḏ, b, m, w*. As *mahmūsa* Sībawayhi defines the following phonemes: *h, ḥ, x, k, š, ṣ, t, s, ṯ, f*.

Later grammarians unanimously repeat Sībawayhi's classification without any serious attempt to explain or even understand his criteria of classification. Al-Mubarrad (285/898) in his *Muqtaḍab* (I, 194.11) gives the following definition of the *mahmūsa* (Danecki 1990:91): "There are phonemes which, repeated on the tongue, cause the voice to flow with them" (*wa-minhā ḥurūf raddadtahā fī l-lisān jarā maʿahā ṣ-ṣawt wa-hiya mahmūsa*). In contrast, he describes the *majhūra* phonemes as follows: "If you repeat them, the voice will be stopped within them; these are voiced phonemes" (*ʾiḏā raddadtahā urtudiʿa ṣ-ṣawt fīhā wa-hiya majhūra; Muqtaḍab* I, 194.12). The notion of hindrance (*urtudiʿa*, elsewhere *muniʿa* 'is prevented') of the voice (*ṣawt*, elsewhere *nafas* 'breath'), which replaces Sībawayhi's *iʿtimād*, seems to be used by al-Mubarrad to indicate the phonatory process. Situating it 'on the tongue' introduces a confusion between articulatory and phonatory processes, connected with the fact that the function of the vocal cords was unknown in ancient phonetics.

One 'major' contribution of later grammarians was the invention of mnemotechnical phrases to learn the *mahmūsa* phonemes: *sa-tašḥaṯu-ka xaṣafatun* (Ibn Jinnī, *Sirr* I, 60) and *ḥaṯṯa-hā šaxṣun fa-sakata* (Ḥaydara, *Kašf* II, 281).

The status of the *majhūra* group remains unclear, since it includes phonemes which are usually regarded as voiceless, namely ʾ, *ṭ*, and *q*. The terms *jahr* and *hams* do not necessarily mean voicing and voicelessness. A number of interpretations were suggested to explain their inclusion in the group of *majhūra*. These interpretations either posited that *majhūra* are voiced phonemes and therefore strove to prove that the three phonemes in question were voiced, or aimed at explaining that *majhūra* defined phonemes that were different from voiced phonemes.

The first group of arguments aimed to prove that the three phonemes were voiced in Sībawayhi's time, rather than voiceless as they are realized in present-day Arabic (cf. Cantineau 1960:22).

i. From the fact that Sībawayhi (*Kitāb* II, 406.23) defines the phoneme *ṭ* as voiced by stating that without → *ʾiṭbāq* 'emphasis' *ṭ* (ط) would become *d* (د), one could draw the conclusion that *ṭ* (ط) was realized voiced, i.e. that it was pronounced as present-day *ḍ* (ض) (see, for instance, Cantineau 1960:32). But against this it might be argued that there are no borrowings from Arabic in which *ṭ* is rendered as *d*. Moreover, in this case the status of *ḍ* would become unclear, since *ṭ* and *ḍ* must have been different phonemes. If one considers that *ḍ* (ض) in modern dialects is frequently pronounced as interdental emphatic *ḏ̣*, then perhaps this pronunciation of *ḍ* was meant by Sībawayhi, unless, of course, it was realized as a special lateral(ized) phoneme, as seems to be implied by Sībawayhi's description of this phoneme (→ *ḍad*). A much later grammarian, Ibn Jinnī (d. 392/1002), regards ظ (*ḏ̣*), which is also classified as *majhūra*, as an emphatic equivalent of *ḏ* (ذ) (*Sirr* I, 61), i.e. *ḏ̣*, thus excluding the pronunciation of *ḍ* (ض) as *ḏ̣*. Moreover, he states that *ḍ* (ض) has no equivalent within nonemphatic consonants (*Sirr* I, 61). His description, however, does not necessarily reflect the original classification of Sībawayhi.

ii. The classification of *q* as *majhūra* may also be explained by facts from modern Arabic dialects, since *q* is pronounced as a voiced phoneme (emphatic *g*; → *qāf*). According to Cantineau (1960:67–71), the original pronunciation of *q* was voiced, as in the majority of the modern Bedouin dialects, whereas he believes that the voiceless pronunciation originated in the urban dialects.

iii. The consonant ʾ (→ *hamza*) is a pure vibration of vocal cords, and this is exactly what a phonatory process (Sībawayhi's *iʿtimād*) means, so it could have been classified as voiced. According to modern phoneticians (Ladefoged and Maddieson 1996:38, 74–76), glottal stops could not literally be voiced.

With the three above reservations, the distinction between *majhūra* and *mahmūsa* could be described as corresponding to the opposition between voiced and voiceless consonants.

The other set of arguments aims at showing that the *majhūra* represent a category defined differently than the modern notion of voiced phonemes (see, e.g., Bravmann 1934:21–25; Gairdner 1935:243–246; cf. Krotkoff 1960). In a lengthy discussion of earlier investigations of the problem, Blanc (1967:307) arrives at the conclusion that *mahmūsa* must be understood as lenes and *majhūra* as fortes. A further argument for this position is al-Ḥaydara's (d. 599/1202) definition (Danecki 1990:93–94), in which instead of Sībawayhi's *iʿtimād* the term *ittisāʿ* 'openness' appears. The *mahmūsa* are wide open at their articulatory points (*li-ttisāʿ maxrajihā*), while the articulation points of the *majhūra* are not wide open (*lam yattasiʿ maxrajuhā*; Ḥaydara, *Kašf* II, 281–282).

Bibliographical references

Primary sources

Ḥaydara, *Kašf* = ʾAbū l-Ḥasan ʿAlī ibn Sulaymān al-Ḥaydara, *Kašf al-muškil fī n-naḥw*. 2 vols. Ed. Hādī ʿAṭiyya Maṭar. Baghdad: Wizārat al-ʾAwqāf wa-š-Šuʾūn ad-Dīniyya, 1984.
Ibn Jinnī, *Sirr* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *Sirr ṣināʿat al-ʾiʿrāb*. Ed. Ḥasan Hindāwī. 2 vols. Damascus, 1993.
Mubarrad, *Muqtaḍab* = ʾAbū l-ʿAbbās Muḥammad ibn Yazīd al-Mubarrad, *al-Muqtaḍab*. Ed. Muḥammad ʿAbd al-Xāliq ʿUḍayma. 4 vols. Cairo: Dār at-Taḥrīr, 1965–1968.
Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. 2 vols. Bulaq, 1898.
Xalīl, *ʿAyn* = ʾAbū ʿAbd ar-Raḥmān al-Xalīl ibn ʾAḥmad al-Farāhīdī, *Kitāb al-ʿayn*. Ed. Mahdī al-Maxzūmī and ʾIbrāhīm as-Sāmarrāʾī. 8 vols. Beirut: Muʾassasat al-ʾAʿlamī li-l-Maṭbūʿāt, 1988.

Secondary sources

Blanc, Haim. 1967. "The 'sonorous' vs. 'muffled' distinction in Old Arabic phonology". *To honor Roman Jakobson*, 295–308. The Hague and Paris: Mouton.
Bravmann, Max. 1934. *Materialien und Untersuchungen zu den phonetischen Lehren der Araber*. Göttingen: Dieterichsche Universitäts-Buchdruckerei.
Cantineau, Jean. 1960. *Etudes de linguistique arabe*. Paris: Klincksieck.
Danecki, Janusz. 1985. "Indian phonetical theory and the Arab grammarians". *Rocznik Orientalistyczny* 44.127–134.
——. 1990. "The phonetical theory of al-Mubarrad". *Studies in the history of Arabic grammar*, II, ed. Michael G. Carter and Kees Versteegh, 91–99. Amsterdam and Philadelphia: J. Benjamins.
Fleisch, Henri. 1958. "*Majhūra, mahmūsa*: Examen critique". *Mélanges de l'Université Saint-Joseph* 35.193–210.
Gairdner, William Henry Temple. 1935. "The Arab phoneticians on the consonants and vowels". *Moslem World* 25.242–257.
Krotkoff, Georg. 1960. "Nochmals: *Majhūra, mahmūsa*". *Wiener Zeitschrift zur Kunde des Morgenlandes* 59–60.147–153.
Ladefoged, Peter and Ian Maddieson. 1996. *The sounds of the world's languages*. Oxford: Blackwell.
Law, Vivien. 1990. "Indian influence on early Arab phonetics: Or coincidence?". *Studies in the history of Arabic grammar*, II, ed. Michael G. Carter and Kees Versteegh, 215–227. Amsterdam and Philadelphia: J. Benjamins.
Schaade, Arthur. 1911. *Sībawaihi's Lautlehre*. Leiden: E.J. Brill.
Talmon, Rafael. 1997. *Arabic grammar in its formative age*: Kitāb al-ʿayn *and its attribution to* *Ḫalīl b. Aḥmad*. Leiden: E.J. Brill.
Troupeau, Gérard. 1958. "Le commentaire d'al-Sīrāfī sur le chapitre 565 du *Kitāb* de Sībawayhi". *Arabica* 5.168–172.

JANUSZ DANECKI (University of Warsaw)

Majrur → ʾIʿrab

Malagasy

Malagasy is a Malayo-Polynesian language spoken in Madagascar by approximately 13 million speakers. Its closest relative is Ma'anyan, a language spoken in southeast Kalimantan (cf. Dahl 1991). The language was probably brought to Madagascar around 400 C.E. Between the 11th and the 14th century, Islamic peoples migrated to the island, probably from the East African coast (Rajaonarimanana 1990:180), who became the ancestors of present-day clans like the Antemoro.

The medieval contacts between Madagascar, called by the Arab geographers *Jazīrat al-Qamar* (this term later came to designate the → Comoros), and the Arabo-Islamic world gave rise to a culture in the southeastern part of the country which preserved several traits of

Arab origin, without fully adopting Islam or the Arabic language, even after the reorientation of the maritime relations toward Europe as a result of the Portuguese colonization. Among these traces is the use of Arabic script, adapted to the structure of Malagasy, known as *sorabe* or *arabico-malgache*, which was used for recording esoteric knowledge in the field of astrology and geomancy (Munthe 1982; Rajaonarimanana 1990). Another trace is the existence of a number of Arabic loanwords that have become integrated into Malagasy.

The shape of the Arabic loanwords in Malagasy has to a large extent been determined by the existence of the *sorabe* script, which exhibits some particularities in the use of the Arabic alphabet, such as the reading of *ts* for ت, *k* for خ, *v* for و, and *z* for ي. The esoteric nature of these writings is clear from the semantic domain of some of these loanwords.

The *sorabe* writings are not the only way through which loanwords reached Malagasy. A second route was that between the northeastern coast of Madagascar and the East African coast, through the Arabo-Islamic culture of the Bantu languages → Swahili and Comorian (Shikomor). These languages have borrowed a large number of words from Arabic and transmitted several terms to Malagasy. In some cases, these may be distinguished from the direct loanwords by the different phonetic treatment. Thus, for instance, Arabic *xabar* 'news' is usually realized in Malagasy as *kabary* in accordance with the form attested in the *sorabe* literature, but regionally – in Majunga – this loanword has the form *habary*, in accordance with Swahili pronunciation.

The loanwords are not very numerous. Dez (1997) mentions about two hundred loanwords in his inventory, some of which belong to the basic lexicon and are frequently used. Most of the loanwords are found in special semantic domains:

i. Words belonging to commerce and traditional economic life. Some of these words remain current, e.g. *mizana* 'balance' (< *mīzān*) and *ariary* 'piastre, five-francs piece', today the name of the national currency (< *ar-riyāl*, itself a loan from Spanish *real*); *loso* 'half-piastre' (< *an-nuṣṣ* 'half') has dropped out of usage.

ii. Words referring to the measuring of time. The most coherent group of words refers to the measuring of time; apparently this group may be traced back to ancient esoteric practices connected to knowledge imported from the Arab world. To this field belong, in particular, the names of the days of the week, which may be traced back to *sorabe* writings, e.g. *alatsinainy* 'Monday', which represents Arabic *(yawm) al-iṯnayni; alakamisy* 'Thursday', which represents Arabic *(yawm) al-xamīs*, etc. To this category also belong the names of the signs of the Zodiac, which served as the names of the months in calendars that were originally used in divining, e.g. *alahamaly* 'Ram', which is still in use nowadays as the name of a traditional ritual in the ancestral cult (< *al-ḥamal*); *adaoro* or *asaoro* 'Taurus' (< *aṯ-ṯawr*), etc.

iii. Words in the political domain. Some words refer to the political domain, such as *kabary*, which has preserved the original meaning of Arabic *xabar* but is also used in Malagasy for 'official discourse, communications, affairs, process'; *jamà* 'traditional gathering [in the ancestral cult]' (< *jamāʿa* 'group'); *tale* (< *ṭāliʿ* 'ascendant [in astrology]'), originally a technical term in soothsaying, indicating the first box in the geomantic square, but today frequently used in the profane sense of 'director'.

iv. Greetings. A greeting term taken from Arabic is *salama* (< *salām*), which originally may have been used in the healer's jargon because it used to mean 'in good health'; it has become one of the most frequently used Arabic loanwords, equivalent to French *salut* or *ça va*; nowadays, it is a neutral way of greeting which avoids the hierarchical connotations of the more refined ways of greeting in Malagasy. Another expression is *arahaba* 'congratulations' (< *marḥaban* 'welcome!'), with the verb *miarahaba* 'to congratulate', which has probably entered Malagasy through the intermediation of Swahili *marahaba* 'hello; thank you [greeting when addressing someone inferior or dependent, or of lower rank]'.

v. Religious terms. Recently, a modest degree of conversion to Islam has taken place, which builds on the most ancient contacts, but this time in an urban environment (cf. Gueunier 1994). In this context, certain ritual or religious terms that used to be limited to internal use in Muslim communities have found their way into the common lexicon, e.g. *ramadany* '[the month of] Ramadan' (< *ramaḍān*), *mikosoaly* 'to pray' (< *ṣallā* + Swahili infinitive prefix *ku-* and Malagasy verbal prefix *mi-*). Other loanwords denote concepts that no doubt were first introduced in a religious context before passing into the common lexicon, e.g. *minia* 'to act intentionally' (< *niyya* 'intention'), *mino* (< *muʾmin* 'believer') 'to believe'.

vi. Urban jargon. Finally, some words have found their way into the urban jargon, e.g. *kaoatry* 'cafe' (< *qahwa* 'coffee'), perhaps through French argot *caoua; masikiny* 'poor, miserable' (< *miskīn*).

To some extent, the terms in the last category revive the old tradition of a secret jargon on the basis of Arabic, which used to be current in the restricted context of Anakara soothsayers on the southeast coast of Madagascar, who guarded it jealously (Rajaonarimanana 1990; Beaujard 1998). This secret jargon, called *kalamon'Antesitesy* 'language of the people of the sand' is of special interest for the history of Arabic because it seems to have preserved some archaic traits of Arabic in the form of loanwords (Versteegh 2001). Its precise origin is unknown, and it is unclear which functions the *kalamo* originally had, apart from ritual. According to Rajaonarimanana (1990:255–256), it was still used in the 1990s between members of the clan, possibly in conversations that were not intended to be understood by strangers.

The structure of this secret speech is entirely Malagasy, but about 70 percent of the lexicon is Arabic in origin. The shape of the Malagasy words in the *kalamo* is disguised by the kind of coding by insertion of meaningless syllables that is known from other → jargons; this procedure is found in a small number of Arabic loanwords as well, e.g. *kadamàfy* < *qadam* 'foot', and *bedàly* < *ʿabd* 'slave' (along with *abodo*). Most of the Arabic words have been borrowed, however, without formal changes. The Arabic words freely combine with Malagasy morphemes, e.g. *voahàky* 'satisfied' (< Malagasy prefix passive participle + *ḥaqq* 'right'); *midokòlo* 'to enter' (< *daxala* + Malagasy verbal prefix *mi-*), *mi-kòlo* 'to eat' (< *ʾakala*). As the last two examples show, the borrowed form of the Arabic verb may have been the Arabic imperative (*udxul, kul*).

Arabic nouns are sometimes borrowed with the article, e.g. *ladìky* 'rooster' (< *(ad-)dīk*), *lanofy* 'nose' (< *(al-)ʾanf*). There are cases when the article seems to be a morpheme in its own right, e.g. *alibètsy* 'house' (< *(al-)bayt*), which becomes *ambètsy* 'in the house' with the Malagasy preposition *am-*. Most borrowings are lexical, but there are a few grammatical items as well, e.g. *antà* 'you [sg.]' (< *ʾanta*), *alaikòmo* 'you [pl.]' (< *ʿalay-kum* 'on you'), as well as the Arabic prepositions *by* and *vy* 'in' in set expressions like *fizalàly* 'night' (< *fī l-layl* 'at night').

Some of the loanwords seem to point to learned transmission, for instance those words that preserve the Arabic feminine ending *-at*, e.g. *sanàntso* 'year' (< *sanat-*) and *maràtsy* 'woman' (< *marʾat-*). One interesting point is that some of the loanwords exhibit traces of a differential treatment of Arabic *ḏ̣* (> *z*) and *ḍ* (> *v*), e.g. *azohòra* < *(aḏ̣-)ḏ̣uhr* 'afternoon', as against *alibiàvy* < *(al-)ʾabyaḍ* 'white'.

The origin of the secret jargon is probably hybrid, since it seems to stem partly from a preexistent Arabic pidgin used by Arabic-speaking immigrants in contacts with the indigenous population, and partly from later borrowing, just like the loanwords in Malegasy. Whereas the other Antemoro and Antambahoaka groups of southeastern Madagascar preserved their Islamic heritage in the written literature of the *sorabe*, the clan of the Anakara was the only one to incorporate it in a secret jargon.

Bibliographical references

Beaujard, Philippe. 1998. *Le parler secret arabico-malgache du sud-est de Madagascar: Recherches étymologiques*. Paris: L'Harmattan.

Dahl, Otto Chr. 1991. *Migration from Kalimantan to Madagascar*. Oslo: Norwegian University Press.

Dez, Jacques. 1967. "De l'influence arabe à Madagascar à l'aide de faits de linguistique". *Taloha* 1.19–38.

Gueunier, Noël J. 1994. *Les chemins de l'Islam à Madagascar*. Paris: L'Harmattan.

Munthe, L. 1982. *La tradition arabico-malgache vue à travers le manuscrit A–6 d'Oslo et d'autres manuscrits disponibles*. Antananarivo: Trano Printy Loterana.

Rajaonarimanana, Narivelo. 1990. *Sorabe: Traités divinatoires et recettes médico-magiques de la tradition malgache antemoro*. 4 vols. Thèse de Doctorat, African studies, Institut National des Langues et Civilisations Orientales, Paris.

Vérin, Pierre and Narivelo Rajaonarimanana. 1991. "Divination in Madagascar: The Antemoro case and the diffusion of divination". *African divination systems: Ways of knowing*, ed. Philip M. Peek, 53–68. Bloomington: Indiana University Press.

Versteegh, Kees. 2001. "Arabic in Madagascar". *Bulletin of the School of Oriental and African Studies* 64.177–187.

Narivelo Rajaonarimanana
(CEROI-INALCO, Paris)

Malay → Indonesian/Malay

Malayalam

1. Historical background

Malayalam, a South Dravidian language (see Krishnamurti 2005), is the mother tongue of 96 percent of the population of the Indian state of Kerala. It is also the principal language of Lakshadweep, the Laccadive Islands, a chain of islands in the Arabian Sea running parallel to the coastal strip in the southwest of India that makes up Kerala. The total number of inhabitants recorded for Kerala in the 2001 census was 31,841,374, and for Lakshadweep 60,650. The history of Malayalam as a separate language goes back approximately twelve hundred years; for a comprehensive description of the Malayalam of today, see Asher and Kumari (1997).

The modern, linguistically homogeneous state of Kerala was formed in 1956, at the time of the major reorganization of states in India, by the merging of Malabar in the north, Cochin in the central part, and Travancore in the south. A year later, in 1957, the people of Kerala produced the world's first democratically elected communist government (that of E.M.S. Namboodiripad). Kerala, with an area of 38,863 km², is the most densely populated region of India, a fact that can be attributed to its fertile soil and abundant rainfall. The main occupation of its people is agriculture. There is also a developing tourist industry, resulting in part from its exceptional scenic beauty. The state's economy is now dependent to a considerable extent on money remitted by expatriates, numbering some millions, working in the United States of America, Germany, and the United Kingdom and in the Arabian/Persian Gulf countries.

External contacts with Kerala go back two millennia or more, with the spices that abound there being one of the attractions. The earliest contacts from distant lands were Greek and Roman. Next came Arab traders. The most widely held view of scholars is that this was before the time of the Prophet Muḥammad, although clear historical evidence that would allow anything like precise dating is not available. Ibn Baṭṭūṭa visited the region between 1342 and 1347; in his account of his voyages, he grouped Quilon and Calicut among the five greatest ports in the world (Ibn Baṭṭūṭa, *Riḥla* 46, 234) and gave the title *zamorin* to the rulers of the Malabar region during those days. The Chinese explorer Zheng He (a Muslim) sailed to places on the southwest coast of India – Quilon, Cochin, and Calicut – in 1409–1411. It was toward the end of the same century, in 1498, that Vasco da Gama landed at Calicut. Dutch influence in the region, which began when the Dutch East India Company sent Admiral van der Hagen there in 1603, ended a century and a half later with the defeat of the Dutch forces by Marthanda Varma of Travancore. British interest began with the arrival of Captain Keeling in Calicut in 1615, but firm control over the region as a whole was established only toward the end of the 18th century. Calicut, which is the port most frequently visited, was known to the Arabs as the capital of what they called Malabar 'the land of mountains'. Kozhikode (*kōẓikkōṭə*), as the town was called by the inhabitants, was Arabized to Kālikūt, which was further modified into Calicut by the Europeans.

Many of these contacts from overseas had an impact on the structure of the lexicon of Malayalam, with that of English being the most widespread and lasting. Equally important in this respect was the development of the religious scene over the centuries. The census of 2001 records 17,883,449 Hindus (56.2% of the total population of the state), 7,863,842 Muslims (24.7%), 6,057,427 Christians (19.0%), 4,528 Jains, 2,762 Sikhs, 2,027 Buddhists, and 2,256 others (including Jews, who many centuries ago settled in and around Cochin). Statistics from the same census show that, as has been the case since records were first kept more than a century ago, the level of literacy, for males and females alike, in Kerala is significantly higher than for any other part of India, at 94.20 percent and 87.86 percent respectively. There are some variations among the different religious communities, but the differences have become progressively less with each decennial census. For the three dominant groups, the percentage figures in 2001 were Hindus 90.2 and 86.7, Muslims 89.4 and 85.5, and Christians 94.8 and 93.5. These literacy figures are clearly relevant in discussing the spread of lexical items from one community to another, since much of the expansion of the understanding and use of new terms is through the written word.

The long history of Hinduism in southwest India, the beginnings of which have been traced back as far as 1000 B.C.E. (Sreedhara Menon 1967:94), had an impact on the structure of society and on language. The aspect of the caste system, which determined people's occupation on the basis of their place within society, was extended to the followers of other religions, such as Judaism, Christianity, and Islam, The borrowing into Malayalam of words from the language of Hinduism, namely Sanskrit, had the effect in due course of changing the phonological structure of the language by the introduction of sounds that were not present in the Dravidian base. This development was reflected in the writing system by the addition of a score of new symbols to make possible the representation of an increased number of phonological distinctions (see Asher and Kumari 1997:406–422). This greater range made it easier to accommodate later borrowings from Arabic and English without too great a departure from the sound patterns of the source language. It remains the case, of course, that some changes are necessary to make these loans fit into the phonology and orthography of Malayalam.

There is a belief among Syrian Christians, the oldest Christian group in Kerala, that Christianity was brought there by St. Thomas in 52 C.E., and among the Jews of Cochin that their ancestors took refuge there after the destruction of Jerusalem by the Romans in 70 C.E. (Spear 1961:73). Similarly, there is a belief among Muslims that Islam reached Kerala during the lifetime of the Prophet himself. There are reasons to believe that Arabs traveled to southwest India before the 7th century with a view both to conquest and trade, but Islam was introduced to Kerala not by the conquerors or the traders but by Mālik ibn Dīnār and his companions, who came with a specific interest in preaching the religion. Cheraman Perumal, a king of Kerala, is believed to have gone to Mecca to convert to Islam. Cheraman Masjid, located at Kodungallur, one of the port cities of Kerala, is the first mosque of India.

The oldest document available about the Muslim community of Kerala is the reference to a Muslim in the 849–850 C.E. record, the so-called Syrian Plates of Kottayam (*Tarisāppaḷḷi śāsanam*). A Muslim royal family by the name of Arakkal, which happens to be the only Muslim dynasty of Kerala, ruled a small portion of land at Kannur and in the Laccadive Islands (Lakshadweep) for some considerable time. As regards the beginning of the dynasty, which lost its power in the early 20th century, there are two very divergent opinions: it was in the 9th century or in the 17th.

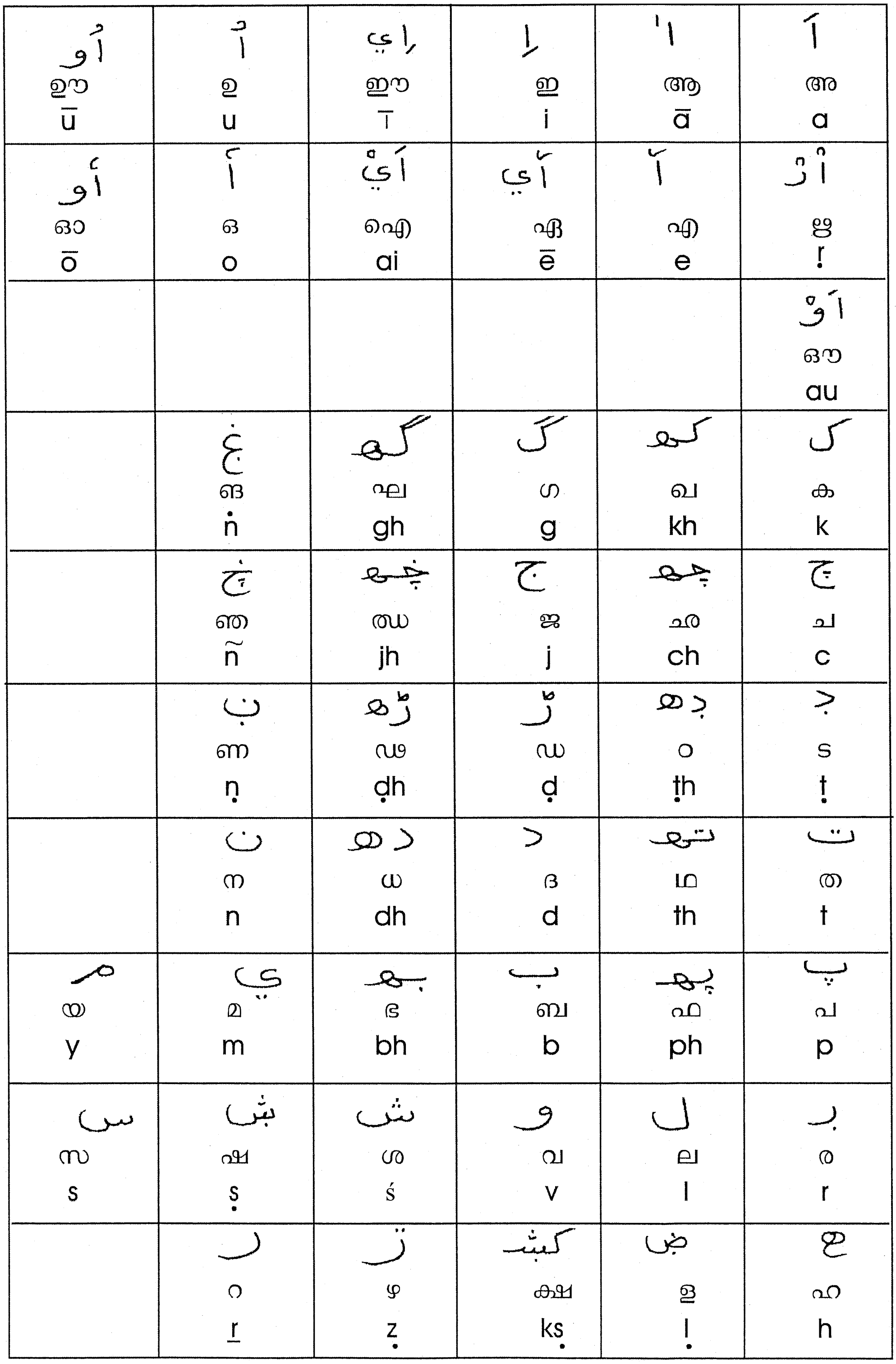
The social life of the Mappilas, as Muslims of Malabar are better known, indicates the presence of Arabic in the various spheres of the cultural life of Kerala. The medium of religious education of religious scholars and people who aspire to be priests (generally known as *mulla/maulavi/musliar*) is still Arabic. Most of their prayers are in Arabic. Mappilas respect Arabic as the language of heaven. Just as in the religious educational centers known as *madrassa*, *dars*, 'Islamiya college', or 'Arabic college', facilities are available for the study of Arabic in the secular educational centers as well. There are many Muslim students studying Arabic in Kerala, and the universities there offer postgraduate and doctoral courses in the subject. C.H. Mohammed Koya (1927–1983), who was the chief minister of Kerala for a few months in 1979 and minister of education from 1967 to 1973, took a special interest in the matter.

There are many books, including historical writings as well as creative literature in prose and verse, written in Arabic by Malayali writers. *Tuḥfat al-mujāhidīn*, an Arabic text written by Sheikh Zainuddin Maqdoom from the cultural center of Ponnani in the 16th century, is one of the seminal texts on the early history of Kerala. The Muslims of Kerala developed a system of writing Malayalam in modified Arabic script called Arabic-Malayalam, which is very similar to the Persian and Urdu scripts (see Fig. 1).

There is an important point to be noted here: just like Muslims all over the world, the Muslims of Kerala also consider Arabic to be

Fig. 1. Arabic Malayalam alphabet

This is a system of writing Malayalam in modified Arabic script. It imitates the alphabets of Persian and Urdu and has a history of at least 500 years. The first three lines of the table show the vowels and the other lines the consonants.



اَ അ a	اٰ ആ ā	اِ ഇ i	اِي ഈ ī	اُ ഉ u	اُو ഊ ū
اْرْ ഋ ṛ	اٚ എ e	اٚي ഏ ē	اَيْ ഐ ai	اٗ ഒ o	اٗو ഓ ō
اَوْ ഔ au					
ک ക k	کھ ഖ kh	گ ഗ g	گھ ഘ gh	غ ങ ṅ	
چ ച c	چھ ഛ ch	ج ജ j	جھ ഝ jh	ڃ ഞ ñ	
ڊ ട ṭ	ڊھ ഠ ṭh	ڑ ഡ ḍ	ڑھ ഢ ḍh	ڹ ണ ṇ	
ت ത t	تھ ഥ th	د ദ d	دھ ധ dh	ن ന n	
پ പ p	پھ ഫ ph	ب ബ b	بھ ഭ bh	ي മ m	م യ y
ڔ ര r	ل ല l	و വ v	ش ശ ś	ڜ ഷ ṣ	س സ s
ھ ഹ h	ض ള ḷ	کش ക്ഷ kṣ	ژ ഴ ẓ	ر റ ṟ	

the language of their religion. By and large, Muslims of India think of Urdu as their mother tongue, not the languages used in the region where they live. Muslims of Kerala are an exception to this phenomenon observed across India, in that they have accepted Malayalam as their mother tongue.

With a history of not less than five hundred years and numerous works of literature both in verse and prose, a category of literature called Arabic Malayalam literature has come into existence. Mappilappattu (*māppiḷappāṭṭu*), the verse branch of this Arabic Malayalam, is immensely popular with Malayalis at large. The most prominent among the writers of Mappilappattu is the poet Moyinkutty Waidyar (1852–1892). Prose writers such as Muhammad Abdurahiman (1898–1945) and the famous fighter for the national freedom of India, Vaikom Muhammad Basheer (1908–1994), are very much part of the mainstream of modern Malayalam literature. It is perhaps because of the success of Basheer as a novelist and short story writer, whose tales are often, though by no means exclusively, set in his own community, that Arabic words, particularly those relating to aspects of Islam, have increasingly become part of the everyday vocabulary of Malayalam (see, for example, the glossary at the end of Basheer 1980).

2. Words used in religious life

Arabic words used within Malayalam utterances can be seen as being of two types. The first are those connected with Islamic religion and used by Muslims in a religious context. These words are generally adapted into Malayalam without phonological or morphological modifications. For example, *Allāhu* is pronounced in the correct manner, in accordance with the phonological structure of Arabic. On the other hand, since Malayalam does not have the phone corresponding to the Arabic emphatic /ḷḷ/ in *Allāhu*, a Malayali non-Muslim would substitute what he felt to be the nearest Malayalam sound, i.e. retroflex [ɭɭ], and so, as far as a Muslim is concerned, mispronounce the word. Such words when used in a religious context are to be regarded at best as unassimilated loans, e.g. *ibādattə* 'prayer, rituals' (< Arabic *ʿibāda*), *jamāʾattə* 'mass' (< Arabic *jamāʿa*), *dīn* 'religion' (< Arabic *dīn*), *mad-hab* 'school of thought' (< Arabic *maḏhab*), *wahābi* 'a disciple of Muḥammad ibn ʿAbd al-Wahhāb', *-īmān* 'the [real] faith' (< Arabic *ʾīmān*), *hajj* 'pilgrimage' (< Arabic *ḥajj*), *iśā* 'night prayer' (< Arabic *ʿašāʾ* 'evening, night'), *ḷuhar* 'noontime prayer' (< Arabic *ḏ̣uhr* 'afternoon'), *phikh* 'jurisprudence' (< Arabic *fiqh*).

3. Secular domain

The second type of Arabic words to be heard within Malayalam utterances comprise those that are used in the secular domain and are fully assimilated into the phonology of Malayalam. In different domains there may be more than one thousand such Arabic words current in Malayalam. These words fall into two main classes: words relating to Islam used in a secular context, and words belonging to other domains. These words, and especially those in the latter class, have been adopted into the vocabulary of Malayalam so well that their Arabic origin is often not recognized. Most of them do not have parallel Malayalam words. A number of them have come through Persian, which was the official language of India during the Mughal period in the domains of law and administration, e.g. *sulttān* 'king' (< Persian, Arabic *sulṭān*), *kasaba* 'main town' (< Arabic *qaṣaba*), *jilla* 'district' (< *ḏilʿ*?), *tālukkā* 'a subdivision of a district' (Persian *taʾalloġ* or Arabic *taʿalluq* 'dependency'), *tāsīldār* 'head of a taluk' (< Persian *tahsīl-dār* 'tax collector' < Arabic *taḥṣīl* 'levying [of tax]'), *adālattə* 'court' (< Persian *adālat* < Arabic *ʿadāla* 'justice'), *amānattə* 'an amount deposited with the court for bail' (< Persian *amānat* < Arabic *ʾamāna* 'deposition in trust'), *harji* 'petition' (< *ʿarḍ*?), *osyattə* 'will' (< Persian *wasiyat* < Arabic *waṣiyya*), *japti* 'confiscation' (< Persian *zabt* < Arabic *ḍabṭ*), *vakkīl* 'advocate' (< Persian, Arabic *wakīl*), *hājar* 'state of being present' (< Persian *hāzir* 'present' < Arabic *ḥāḍir*). Some other important loanwords widely used in everyday language, both spoken and written, are given in Table 1.

Salām and *iṅkvilāb* need to be elaborated upon. *Salām* is abstracted from *as-salāmu ʿalaykum* 'peace be upon you!', and is used to convey one's regard. *Salām* has thus become a form of salutation. *Lālsalām* is used for 'red

Table 1. Arabic loanwords in Malayalam

Arabic	gloss	Malayalam	gloss
xaṭṭ	‘letter, handwriting’	*kattə*	‘letter’
bāqī	‘remnant’	*bākki*	‘remnant’
salām	‘peace’	*salām*	‘regards, salute’
maʿmūl	‘that which is done’	*māmūl*	‘tradition, custom’
dunyā	‘that which is near; physical world’	*duniyāvə*	‘the physical world’
inqilāb	‘change, revolt’	*iṅkvilāb*	‘the revolution’
badal	‘substitute’	*badal*	‘substitute’
ṣabūr	‘patient’	*sabūr*	‘to forgive’
šayṭān	‘Satan’	*cekuttān*	‘the devil’
takrār	‘repetition’	*takarār*	‘trouble, problem’
xalāṣ	‘salvation; settlement’	*khalāsi*	‘porter working in a shipyard’
musāfir	‘traveler’	*musāvari*	‘tourist bungalow’
xalāṣ	‘to save; the final’	*kalāśam*	‘the final’
qiṣṣa	‘story’	*kissa*	‘story’
ṭabla	‘drum’	*tabala*	‘a percussion instrument’
nabī	‘prophet’	*nabi*	‘prophet’
ʾiblīs	‘Satan’	*iblīs*	‘Satan’
ṣūfī	‘ascetic, Sufi’	*sūfi*	‘Muslim saint’
qabr	‘grave’	*khabar*	‘grave’
qalb	‘heart’	*khalb*	‘heart’
mulla	‘scholar’	*mulla*	‘Muslim priest’
maydān	‘playground’	*maitānam*	‘ground’
xālī	‘empty’	*kāli*	‘empty’
baraka	‘grace’	*varkkattə/ barkkattə*	‘grace’
ḥalaq(a)	‘earring’	*alikkattə*	‘a type of earring’
ʿaṭṭara	‘to emanate fragrance’	*attar*	‘aromatic spray’
jubba	‘long dress of men’	*jūba*	‘long dress of men’
kīs	‘pocket’	*kīśa*	‘pocket’
faṭīr	‘unleavened bread made of wheat’	*pattiri*	‘unleavened bread made from rice flour’
ʿaraq	‘juice’	*ṟākkə*	‘arrack’

salute’ among Communists (*lāl* being Hindi for ‘red’). *Iṅkvilāb*, from the Arabic word for a continuous change, is now a word commonly found in the rhetoric of politicians, especially of Communists, for ‘revolution’. Both of these loans have thus taken on new connotations in the Malayalam context.

4. PERSONAL NAMES

Generally, Muslims in Kerala, both men and women, use Arabic personal names just as Muslims elsewhere do. Most of them recall the Prophet Muḥammad, his family members, his companions, and Sufi saints. A large number of male names begin with the element *abd* (Arabic *ʿabd*) ‘servant [of God]’. God is referred to here by one of His one hundred holy names (*al-ʾasmāʾ al-ḥusnā*), such as Rahman (*ṟahman*), Raheem (*ṟahīm*), etc. Examples (given here first in the form in which they are usually romanized by Malayalis) of male names are Muhammad (*muhammad*), Abdullah (*abduḷḷa*), Abubakar (*abubakkar*), Umar (*umar*), Usman (*usmān*), Ali (*āli*), Abdul Khader (*abdulkhādar*), Abdurahiman (*abdurahimān*), Fakruddin (*phakṟuddīn*), Jamaluddin (*jamāluddīn*), Abdul Gafoor (*abdulgapūr*), Abdul Jabbar (*abduljabbār*); and of female names: Aysha (*āyiṣa*), Khadeeja (*khadīja*), Zainaba (*sainaba*), Jameela (*jamīla*), Safiya (*saphiya*), Maimoona (*maimūna*), Za-keena (*sakkīna*), Amina (*āmina*), Zuhara (*suhaṟa*), Laila (*lailā*).

These names are used in two different forms, either the original form as used in Arabic, or a changed form to suit the local phonological system. Examples of nativization include Muḥammad > Mammad (*mammad*), ʾAḥmad > Ammad (*ammad*), Muḥyī d-Dīn > Moideen (*moytīn*) ~ Maideen (*maytīn*). Syllabic reduction occurs in many ways, e.g. ʿAbd ar-Raḥmān > Rahman (*ṟahman*), Abdul (*abdul*), Abdu (*abdu*), Abdura (*abduṟa*), Adraman (*adṟamān*), Andraman (*antṟamān*). Sometimes Arabic names are used with the Malayalam prefix or suffix Kunhi (*kuññi*), Kutty (*kuṭṭi*), Unni (*uṇṇi*), or Kochu (*koccu*), all meaning 'child', e.g. Kunhi Muhammad (*kuññi muhammad*), Kunhi Amina (*kuññi āmina*); Muhammad Kunhi (*muhammad kuññi*), Ali Kunhi (*āli kuññi*); Kutty Ali (*kuṭṭi āli*), Kutty Ahammad (*kuṭṭi ahammad*); Muhammad Kutty (*muhammad kuṭṭi*), Ahammad Kutty (*ahammad kuṭṭi*); Unni Muhammad (*uṇṇi muhammad*), Unni Ali (*uṇṇi āli*); Muhammad Unni (*muhammad uṇṇi*), Ahammad Unni (*ahammad uṇṇi*); Kunhi Ali Kutty (*kuññi āli kuṭṭi*); Kutty Ahammad Kutty (*kuṭṭi ahammad kuṭṭi*); Kochu Muhammad (*koccu muhammad*), Kochu Zuhara (*koccu suhaṟa*). In rare cases, the Muslims of Kerala may use Malayalam (e.g. Marakkar, *marakkār*), Sanskrit (e.g. Sunita, *sunīta*), or Persian (e.g. Shajahan, *ṣājahān*) names as personal names.

5. PHONOLOGICAL CHANGES

The Semitic origin of Arabic and the Dravidian roots of Malayalam posit problems of being mutually exclusive in the case of certain phonemes. All the vowels in Arabic (*a, i, u*) are present in Malayalam as well, although Malayalam has 15 vowel sounds in addition to these. The Arabic consonants /b/, /t/, /j/, /d/, /r/, /s/, /š/, /k/, /l/, /m/, /n/, /h/, /w/, /y/ are present in the Malayalam inventory. Malayalam does not have sounds corresponding to all those of Arabic, as found in such words as *Allāhu* and *Muḥammad*, and because of the phonological structure of their language, Keralites are not able to pronounce such words properly. Instead, Malayalam uses the sounds that are closest to the original. *Allāhu*, for instance, is pronounced as [aɭɭɑːhu], and *Muḥammad* as [muhammad]. Muslims throughout Kerala make a conscious effort to pronounce words of religious importance such as *Allahu* properly, although other Arabic words are transformed phonologically to suit the framework of Malayalam. Non-Muslims routinely change the unfamiliar Arabic sounds to familiar Malayalam sounds. Table 2 lists the sounds used by Malayalam in place of the Arabic ones.

Table 2. Substitution of Arabic sounds by Malayalam speakers

Arabic	Malayalam
/ṯ/	/s/ (voiceless alveolar fricative)
/ḥ/	/h/ (voiceless glottal fricative)
/x/	/k/ (voiceless unaspirated velar plosive), /kh/ (voiceless aspirated velar plosive), /h/ (glottal fricative)
/ḏ/	/d/ (voiced dental plosive)
/z/	/s/ (voiceless alveolar fricative)
/ṣ/	/s/ (voiceless alveolar fricative)
/ḍ/	/ḷ/ (voiced retroflex lateral), /d/ (voiced dental plosive)
/ṭ/	/t/ (voiceless dental plosive)
/ḏ̣/	/ḷ/ (voiced retroflex lateral), /d/ (voiced dental plosive)
/ʿ/	/a/ (open central vowel)
/ġ/	/g/ (voiced velar plosive)
/f/	/ph/ (voiceless aspirated bilabial plosive), /p/ (voiceless unaspirated bilabial plosive)
/q/	/k/ (voiceless unaspirated velar plosive), /kh/ (voiceless aspirated velar plosive

The symbol used for representing the voiceless labiodental fricative /f/ of Arabic loans in Malayalam is the one originally used for the voiceless aspirated bilabial plosive – /ph/ – in Sanskrit loans. In modern Malayalam, its predominant use has come to be the representation of [f], as in the case of *phān* [faːn] and *phōn* [foːn] for the English loanwords *fan* and *phone*. Earlier loans from Arabic replaced Arabic /f/ by Malayalam /p/, e.g. *fāṭima* > *pāttummā*.

6. SEMANTIC CHANGES

Most Arabic loanwords in Malayalam are confined to the domains of administration, religion, and food, and it could be that the chances of these words undergoing semantic change are comparatively remote. But the sense of humor associated with the Malayali Muslim (what has come to be called *māppiḷaphalitam* < *māppiḷa* 'Malayali Muslim' + *phalitam* 'joke') gives a humorous twist to meanings in the case of words that are phonologically close. Arabic *xalaqa s-samawāti*, for instance, is a Qurʾānic expression meaning 'God has created the skies'; but in Malayalam, if someone says *kalakkassamavāti*, it means 'everything got messed up', 'turmoil', or 'a turbulent situation'. The word *mahšaṟa* (< Persian *mahzar* < Arabic *maḥḍar*), meaning 'assembly for the last trial after doomsday', is used with the meaning 'any place with a huge crowd' or 'any place that is in utter chaos'. In Malayalam, *phitna* (< Arabic *fitna* 'trouble') also means 'gossip' and 'backbiting'. The word *cakkāttə* is a modification of *zakāt*, which according to Islam is the mandatory distribution of a particular portion of one's property during Ramadan. *Cakkāttə* in Malayalam has a different connotation. It means what is given freely without any binding commitment. This word undergoes phonological as well as semantic changes. Arabic /z/ first became /s/, and this was later transformed into /c/. The development of the Arabic term *ṣāḥib* 'companion; comrade' was first used as an honorific term or suffix in referring to Muslim gentlemen, e.g. Muhammad Sahib (*muhammad sāhib*). Alongside this usage, the word underwent a number of phonological changes over time to become *sāyippə*, the Malayalam word for 'foreigner', and more specifically a Westerner. Now, both *sāhib* and *sāyippə* are used, but in different domains.

7. Morphological changes

In general, there was not much correspondence between Arabic and Malayalam in the domain of scholarship and knowledge. Those who were working in Arabic language and literature had no exposure to Malayalam literature. There could be such transactions in 'Arabic-Malayalam', but that was off the mainstream of Malayalam. Morphological changes are very few due to the lack of such exposure, but there are a few exceptions, e.g. Arabic *ʿālim* 'scholar', pl. *ʿulamāʾ*. Both *alīm* and *ulamā* are used in Malayalam, but Malayalis took *ulamā* as singular and pluralized it as *ulamākkaḷ*, with the Malayalam plural suffix. This somewhat unusual instance illustrates the normal treatment of words borrowed into Malayalam, in that they are fitted into the morphological structure of the language in the sense that the inflectional endings – such as markers of case – are added to them.

8. Syntactic changes

Grammatical peculiarities of Arabic have not affected Malayalam syntax. This might be partly because of the lack of contact between the two languages at the spoken level – the contact was mainly literary and cultural – and partly because, despite the considerable number of Muslims in southwest India, the members of the population having contact with Arabic were always in a minority.

9. Kinship terms

Of the kinship terms used by Malayali Muslims, only one is from Arabic, namely *umma* 'mother' (< Arabic *ʾumm*). The terms for father, *bāppa*/*vāppa*/*uppa*/*uppāva*/*bāppicci*/*vāppicci*/*vāyicci*, are variants of *bāp* in Hindi/Urdu. A few of the kinship terms of Malayali Muslims were borrowed from Persian (e.g. *kākka* 'elder brother' < Persian *kākā*), and the remaining ones are from Malayalam (e.g. *ammāvan* 'uncle'). It is nevertheless the case that there are marked differences among the kinship terms of Muslims, Christians, and Hindus (see Asher and Kumari 1997:451–454).

Bibliographical references

Primary source

Ibn Baṭṭūṭa, *Riḥla* = *Ibn Battuta, Travels in Asia and Africa, 1325–1354*. Trans. and selected H.A.R. Gibb, with an introduction and notes. London: Routledge, 1929.

Secondary sources

Asher, R.E. and T.C. Kumari. 1997. *Malayalam*. London and New York: Routledge.

Basheer, Vaikom Muhammad. 1980. *"Me Grandad 'ad an elephant!": Three stories of Muslim life in South India*. Trans. R.E. Asher and Achamma Coilparampil Chandersekaran. Edinburgh: Edinburgh University Press.

Joseph, P.M. 1984. *Malayāḷattile parakīya padaṅṅaḷ* [Loanwords in Malayalam]. Trivandrum: Kerala Bhasha Institute.

Karassery, M.N. 1995. "Arabic Malayalam". *Kerala Muslims: A historical perspective*, ed. Asghar Ali Engineer, 168–174. Delhi: Ajantha.

Krishnamurti, Bh. 2005. *The Dravidian languages.* Cambridge: Cambridge University Press.
Sreedhara Menon, A. 1967. *A survey of Kerala history.* Kottayam: National Book Stall.
Spear, Russell. 1961. *India: A modern history.* Ann Arbor: University of Michigan Press.

R.E. Asher (University of Edinburgh)
M.N. Karassery (Calicut)

Mali

In the contemporary Republic of Mali in West Africa, Arabic is not a native language of the largely Muslim population (80–90% Muslims), except for a minority of approximately 1 percent of Ḥassāniyya speakers (www.ethnologue.com). However, Standard Arabic plays an important role in Islamic education and scholarship in Mali, and in this role the language has a long history in the region. Arabic did not spread to the Sahel region of West Africa primarily as the language of Islam, however, but as a language of commercial communication. Trade routes connected West Africa across the Saharan desert to North Africa and the Middle East long before the rise of Islam (Hiskett 1994:92–93, 97; Saad 1983:4). It has even been suggested that these trade connections date from the introduction of the camel as a means of long-distance transport through the Sahara around the beginning of the Common Era (Levtzion 1973:6).

After the Islamic conquests and the subsequent spread of the Arabic language to Egypt and North Africa in the 7th and 8th centuries C.E., the language started to play an increasingly important role as a → lingua franca in commercial relations with the West African Sahel (Levtzion 1979:15). Unlike North Africa and the Middle East, however, West Africa was never conquered as part of the Islamic Empire. Instead, Islam was disseminated in West Africa through trade from as early as the 9th century (Hiskett 1994:97; Mbiti 1969:243), and it made its first documented appearance in the region in the first half of the 10th century C.E. (Amiji 1984:105). Muslim traders were the main agents who spread Islam, through trade networks that transcended family and village boundaries (Amiji 1984:105; Levtzion 1979:15–16). They were much respected for their wealth (Levtzion 1979:11). Moreover, they shared one law, the *šarīʿa*, which included a uniform system of weights and measures, and one language, Arabic (Hiskett 1994:100, 194; Levtzion 1979:15; Saad 1983:4). Arabic occupied a special position in the region, not only because it was an interregional lingua franca of trade but also because it was the first written language in West Africa. Muslims who used Arabic script were greatly respected for this knowledge (Levtzion 1979:11) and much sought after as scribes, advisors to local rulers, and faith healers (Amiji 1984:105). The appeal of the Arabic-writing Muslim traders from the north and east and the appeal of Islam as the first written and universal religion in the region (Amiji 1984:106; Goody 1971:461) resulted in conversions to Islam, first among local traders and rulers from the 10th century C.E. onward. An increasing number of conversions among larger parts of the population of West Africa took place from the late 13th century onward (Clarke 1982:47–48; Hiskett 1994:94). As a result, the role of Arabic expanded from that of a lingua franca of transregional trade to its present role as language of Islamic education and scholarship (Abdulaziz 1984:132; Amiji 1984:106).

In the 14th century, the famous Muslim ruler Kankan Musa of the ancient kingdom of Mali went on a pilgrimage to Mecca, also visiting Egypt on his way. He spent so much of his wealth in Egypt that his spending led to inflation there. His journey to the heartland of Islam left a strong impression on him. He became aware of the imperfections of Islamic practice in his kingdom. Upon his return, he therefore invested in Islamic education and scholarship by building mosques and sending local Islamic scholars to study in centers of Islamic scholarship in the Maghreb, above all in Fes. He firmly established Islamic education and scholarship in the West African Sahel, and the regional centers of trade Timbuktu and Djenné developed into centers of Islamic learning (Levtzion 1986b:185–186, 196–197).

In the 16th century, Timbuktu reached its peak as the center of Islamic scholarship for the entire region of West Africa. The city counted some 150 Qurʾānic schools and attracted students and scholars from West Africa, the Maghreb, and even the Middle East. Islamic sciences were taught on a large scale in the houses of scholars, Qur'ānic schools, and mosques, the main language of study and instruction being Arabic. From the surviving Arabic manuscripts of the 16th century, we know that West African scholars achieved a level of Standard Arabic that enabled them not only to study the classical corpus of Islamic literature but also to write commentaries and chronicles in Standard Arabic (Cissoko 1969:48, 59, 72, 1975:68; Hunwick 1966:25; Kaba 1984:248; Levtzion 1977:416–417, 1986a:12; Saad 1983: Chap.3). A large number of West African manuscripts written in Arabic in the 16th to 19th centuries are preserved and collected in various places, including libraries in Timbuktu such as the Centre de Recherche, Education et Documentation Ahmed Baba (CEDRAB) and the Mama Haidara Library. Some of the preserved Arabic texts, such as the *Nayl al-ibtihāj bi-taṭrīz ad-dībāj*, a bio-bibliographical dictionary of Maliki scholars by Ahmad Baba, a 16th/17th-century scholar from Timbuktu, and the 16th/17th-century Timbuktu chronicles *Ta'rīx as-Sūdān* and *Ta'rīx al-Fattāš*, contain further details on the system and content of Islamic education and scholarship at the time.

Students in Islamic education first learned to read and write, and sometimes memorized, the entire *Qur'ān*. In this way, they also learned the Arabic script. Students then proceeded to the second level of Islamic education, in which they studied books from the corpus of classical Islamic literature. These books were studied in Arabic, and, presumably, students learned Arabic in the first place through the study of these books, i.e. through the explanations their teacher gave of the Arabic text. They started with books on Maliki *fiqh*. Later on, they studied other Islamic sciences, such as theology (*tawḥīd* lit. '[dogma of the] unity of God'), *ḥadīṯ*, and exegesis (*tafsīr*). Next to the Islamic sciences, the study comprised subsidiary subjects, above all related to language, such as grammar and syntax (*naḥw*), inflection (*ṣarf*), logic (*manṭiq*), and rhetoric (*bayān*). Students could eventually reach a level where they started writing papers and commentaries on the classical works themselves. From the surviving works we know that at least some of them reached a level of Arabic and of Islamic scholarship that did not fall short of that of scholars in the Middle East and North Africa (Cissoko 1969:59–60; Levtzion 1977:417–418, 1986a:12; Saad 1983: Chap. 3; Kaba 1984:248; Bouwman 1996:19–24).

In the 16th century, scholars from the region of Timbuktu also traveled to other centers of Islamic learning in West Africa and the Maghreb, and even as far as Egypt and Arabia. They stayed abroad for years to study with the illustrious scholars of their time, especially at al-Azhar University in Cairo. Upon their return, they transmitted the knowledge thus acquired to their students. During their travels and studies in Muslim communities outside their native region, Arabic played an inevitable role as the language of communication and instruction. Timbuktu scholars also obtained *'ijāza*s, diplomas in which teachers authorized their advanced students to teach a specific work, from well-known Islamic scholars of the Maghreb and Middle East, and brought them to Timbuktu. The *'isnād*s, chains of authorized transmission of Islamic knowledge, were thus continued in West Africa (Hunwick 1966:25–26; Saad 1983:66; Bouwman 1996:24–28), and even today one can find Islamic scholars in Mali possessing *'ijāza*s with an *'isnād* that goes back to the medieval Middle East (Bouwman 2005:30).

The industrious intellectual life of Timbuktu came to a halt at the end of the 16th century C.E., when the Saʿdian forces from the Maghreb conquered the region of Timbuktu and killed or deported the greater part of the intellectual elite. Of some scholars, above all the famous Ahmad Baba, it is known that they continued their studies and teaching in the Maghreb after their deportation, but in general the educational and scholarly activities in the West African Sahel are said to have declined (Hunwick 1962; Levtzion 1977:414–416; Saad 1983:66–67). Islamic education and conversions reached a new peak in the wake of the 19th-century Jihadist movements in West Africa. However, unlike earlier periods, conversions now often took place under pressure and with force. Islamic education nevertheless started to flourish again, especially among the followers of the Jihadist movements, who established an Islamic state in part of the region of contemporary Mali. Islamic education was still the only formal education for children, and Arabic was still the only written language in the region (Amiji 1984:107; Brenner 1993:63–65; Mommersteeg 1996:21–22). Next to its predominant role in scholarship and trade, Arabic served diplomatic and documentary purposes in the precolonial period (Amiji 1984:106), and the Arabic script eventually was also used for writing West African languages (Abdulaziz 1984:133; Tamari 2002:99).

The situation changed drastically with the arrival of the French colonial powers in the region at the end of the 19th century. In their constant fear of anticolonial, pan-Arab, and pan-Islamic opposition, the French tried to counteract Islamic education and the predominant role of Arabic in West Africa. At first, they used Arabic in their dealings with West African rulers because it was the only possible language of communication, but in 1911 they banned Arabic from all official administration dealings and suppressed the circulation of written Arabic materials. They introduced secular education, with French as the language of instruction, and forced the children of the local elite into their schools (Brenner 1984:36–37, 1986:9; Harrison 1988:51–52; Mommersteeg 1996:22).

With several measures, the French tried to oppose the Qur'ānic schools and to limit the scope of Islamic education. In order to attract the local population to French education, they even set up a Franco-Arabic school in which Arabic and Islam were to be taught next to French and secular subjects, but with little success. The Islamic educational institutions of local scholars continued to exist alongside the official French system, and West African parents continued to resist enrolling their children in French schools and to opt for Islamic education instead. Parents of the elite who were forced by the French to enroll their children in French schools went so far as to enroll the children of their slaves, to prevent their own children from being influenced by the 'culture of the unbelievers' (Brenner 1984:36, 2000:39–54; Cissé 1992:84–85; Mommersteeg 1996:22).

Local initiatives in the first half of the 20th century added a third component to the educational landscape, the *madrasa*, which in the long run became a strong competitor of what henceforth came to be known as traditional Islamic education, but also of the secular, French education that had been established by the colonial administration and was continued by the independent government of Mali. This new educational system of the *madrasa* came up in the 1940s, first in Segou and Bamako, but soon similar institutions followed all over the region. The founder of the *madrasa* system in Segou was Saada Toure (b. ca. 1912). As a child, Toure received an Islamic education, without learning the Arabic language as such, as was common practice. Before he reached a higher level of proficiency, he was forced into French education by the French colonial administration (Brenner 1986, 2000:39–41, 54–84). Once in French school, he learned French with only four years of study. He thus became aware of the advantages of the French didactic system, but without being affected by its secular character and without giving up the path of Islam. Toure set up a new educational system in which he applied French didactic methods to the teaching of Arabic and Islam. Children in his school studied both Islamic and secular subjects and learned Arabic as a 'living language' from the start. Arabic was taught and employed as the language of instruction for all subjects, but Toure eventually added a limited teaching of French to his curriculum in order to prepare children for life in an increasingly Francophone society. The new school was called *madrasa*, after the Arabic word for school, and it was soon copied and spread throughout the region, together with the books Toure wrote on the teaching of Arabic (Brenner 1986, 2000:74–84).

Similar developments took place in Bamako, but these were influenced by another movement. In the first half of the 20th century, West African students of traditional Islamic education went to study in Egypt at al-Azhar University and came under the influence of new Islamic reformist movements, above all the Salafiyya, a movement that rejected secularism while at the same time relying fully on rational arguments in the interpretation of the Islamic sources. As part of this thinking, Muslims should study and know Arabic at a level that would enable them to read the Islamic sources in Arabic and take part in the global Islamic discourse. Students of Islam should become independent from the traditional religious leaders, above all from the Sufi orders that were dominant in West Africa and strongly opposed by the Salafists. They believed that a thorough knowledge of Arabic was needed to reach this goal, and therefore Islamic education in West Africa had to be reformed to achieve more efficient teaching of Arabic and less conservative teaching of the Islamic sciences. Upon their return to West Africa, the Azhar graduates set up a *madrasa* in Bamako similar to the one of Saada Toure in Segou, even though Toure was still closely associated with the Sufi orders and adverse to the Salafi movement and therefore more acceptable to the French. In the long run, Toure's initiative was the more influential one in the spreading of the *madrasa*s in the region (Brenner 1986:6, 2000:54–84; Kaba 1974:135–139, 154–166).

Independence in 1960 resulted in a Francophone, secular administration. The new Republic of Mali counted three parallel educational systems as a legacy of colonial and earlier times. The traditional Islamic education in the Qurʾānic schools continued to exist, especially in rural and remote areas. The French colonial school continued as the public school system of independent Mali, offering secular education with French as the language of instruction. The third factor in formal education remained the private *madrasa*s, with Arabic as the language of instruction and a limited teaching of French. The number of *madrasa*s continued to grow, and is still growing. The proportion of children enrolled in *madrasa*s nowadays is estimated at 30–40 percent of children being enrolled in formal education (Bouwman 2005:11–14, 50).

Government policy does not consider traditional Islamic education to be formal education; consequently, it is not included in official census figures. Together with the mosques, Qurʾānic schools fall under the Ministry of the Interior rather than the Ministry of Education, and as institutions of education, they are neglected by the government (Bouwman 2005:101). In spite of this, they still constitute an important component in the educational landscape of Mali and are still the only education many, if uncounted, children receive, especially influential in rural and remote areas that are not reached by the public educational programs or private *madrasa* initiatives. Many traditionally oriented parents still prefer to send one or two children to a Qurʾānic schoolteacher for years, to be educated as the Islamic expert of the family, while their other children do not receive any formal education at all. These parents are often opposed to the secular, French school system, which they perceive as foreign to their culture and traditions and as leading their children astray from the path of Islam. To many of these parents, even the Arabic Islamic *madrasa*s are too much influenced by the French educational system (Cissé 1992:149; Bouwman 2005:20–21, Chap. 6).

Children in traditional Islamic education still read the same books and study in much the same way as in 16th-century Timbuktu (Diakite 1991:34; Tamari 2002:104–111; Bouwman 1996:51). This system is today often criticized for the passive method of instruction and the limited knowledge students acquire. It is, moreover, criticized for harsh conditions; children are often severely beaten by their teacher, while living in their teacher's home for years under poor conditions and obliged to work and beg for their teacher (Dumestre 1997:39–43; Mommersteeg 1996:58–59; Tamari 2002:105; Bouwman 2005:11, 24–25, 44–46). Yet, many parents view this educational system as the only one fit to socialize their children and prepare them for the poor and harsh living conditions of the rural population in West Africa (Brenner 1991:64; Mommersteeg 1996:59; Bouwman 2005:24–25). In addition, proponents of traditional Islamic education stress that only in traditional Islamic education are the entire *Qurʾān* and works of Islamic sciences studied and their content memorized, so that the advanced students and scholars have instant access to the knowledge contained in the texts; according to them, *madrasa* students read only selected parts of the *Qurʾān* and Islamic literature without usually memorizing the content (Bouwman 2005:175–178).

Nevertheless, the *madrasa* system has also started to influence Qurʾānic education (Bouwman 2005:44–46). Traditionally, students were taught through individual tuition by the teacher, who taught them a new piece of text for ten to thirty minutes a day, which the students then repeated on their own for the rest of the day. In this way, students progressed according to their own abilities (Mommersteeg 1996, Chap. 3; Tamari 2002:112; Bouwman 2005:22–23, 29). Under the influence of the *madrasa* system, children in Qurʾānic schools are now increasingly taught in a class, with repetition of texts in chorus. Some Qurʾānic teachers, who were themselves educated in the *madrasa* system, even apply the teaching methods of the *madrasa*. In urban and semiurban regions, where the *madrasa* system is readily available, it has pushed traditional Islamic education even further to the background of the educational landscape. Children in cities usually attend Qur'ānic education only in the margin of formal education in a *madrasa* or a public French school, either before they enter this school or during the three months of summer vacation. In the Qur'ānic school, they usually only memorize a few suras that are relevant for prayer (Bouwman 2005:25–27).

Parents oriented toward Arabo-Islamic education who have access to a *madrasa* nowadays usually opt for the *madrasa* instead of traditional Islamic education. A *madrasa* is in many respects similar to the public, French school system. Whereas in Qur'ānic schools children often sit on the floor in the house or *sous hangar* of their teacher, *madrasa*s are genuine school buildings, in which children are taught in a class according to a set schedule, curriculum, and path of progress. Like the French system, the *madrasa* system is divided into three circles, after each of which students obtain an acknowledged diploma. The last diploma, the baccalaureate after twelve years, gives access to the secular study of Arabic at the public University of Mali – after an entrance exam – and to universities in Arab countries and some non-Arab Muslim countries (Brenner 2000:219ff.; Bouwman 2005:67–68, 89–92, 126–131).

The *madrasa* system, with its emphasis on the acquisition of a good working knowledge of Modern Standard Arabic, has a crucial influence on Islamic scholarly discourse in Mali and West Africa as a whole. For one thing, students come under the influence of different Islamic thoughts and new Islamic movements during their studies abroad; moreover, knowledge of Arabic has become an important factor in the Islamic knowledge hierarchy of Mali, in which scholars of traditional Islamic education increasingly fail to compete. In spite of the Francophone character of public life in contemporary Mali, the role of Arabic in society has not diminished but rather is growing, because those who want to participate fully in Islamic scholarly discourse in Mali have to know Arabic at a level enabling them to argue on the basis of Islamic sources they read independently in Arabic (Zappa 2004; Bouwman 2005:179–185).

Along with its role in Islamic scholarship and education, Arabic has yet another, very different role in Mali, as is the case in other West African societies. Throughout the centuries, Arabic was and still is employed in the field of magic, mysticism, and divination. People who need help with a specific problem, such as finding work or a marriage partner, or curing an illness, can address a so-called *marabout* (Bouwman 2005:166–167). According to Tamari (2002:110), many traditionally oriented scholars are involved in maraboutism, and Amiji (1984:112) observes that by "dispensing amulets, talismans, medicine and advice, [marabouts] eventually become indispensable to the community".

In order to solve people's problems, a marabout generally makes a gris-gris for them. A gris-gris is a material that has esoteric power attached to it for the benefit or harm of a person. The term is often translated as *amulette*, but it can also be, for instance, water one has to drink. For one type of gris-gris, the marabout writes words from the *Qur'ān* on a piece of paper which is then folded into an amulet. Another type requires the marabout to write words from the *Qur'ān* on a wooden board, then wash the words off; the supplicant ingests or pours that water over his or her body. In the making of a gris-gris, marabouts have to employ the Arabic script. However, apart from copying Arabic words from the *Qur'ān*, which they may have learned in traditional Islamic education, they need no further knowledge of Arabic (Mommersteeg 1996, Chap. 7; Bouwman 2005:167–168). Marabouts also offer benedictions, for which they have to recite from the *Qur'ān*, again without actually knowing the meaning of the words. As in the case of gris-gris, the ability to do benedictions depends on the knowledge of secrets rather than on knowledge of Arabic (Mommersteeg 1996, Chap. 5; Bouwman 2005:168).

Although the only knowledge of Arabic required for gris-gris and benedictions is the correct writing out and pronunciation of the Arabic words employed, a marabout who knows more Arabic is also considered more powerful in the secrets. Mommersteeg (1996:47) notes that the employment of Arabic in the making of gris-gris is directly related to the belief that Arabic is a revealed language and scripture. Clients believe that knowledge of Arabic puts a marabout in a special position, or in a closer contact to God, and that it provides him with special powers. Knowledge of Arabic is not a condition for his work, but it adds to the status and success of a marabout. On the other hand, marabouts without knowledge of Arabic can be considered even more powerful than marabouts with knowledge of Arabic if the former know more secrets (Bouwman 2005:168–169). As phrased by Kaba (1974:89), a marabout is first of all a person "who is initiated in mysticism and divination".

Bibliographical references

Abdulaziz, Mohamed H. 1984. "Language in Afro-Arab relations: The interaction between Arabic and African languages". *Historical and socio-cultural relations between Black Africa and the Arab world from 1935 to the present: Report and papers of the symposium organized by UNESCO in Paris from 25 to 27 July 1979*, 131–138. (= *The general history of Africa: Studies and documents*, VII.) Paris: UNESCO.

Amiji, Hatim M. 1984. "Religion in Afro-Arab relations: Islam and cultural change in modern Africa". *Historical and socio-cultural relations between Black Africa and the Arab world from 1935 to the present: Report and papers of the symposium organized by UNESCO in Paris from 25 to 27 July 1979*, 101–129. (= *The general history of Africa: Studies and documents*, VII.) Paris: UNESCO.

Bouwman, Dinie. 1996. *Geleerd, geroemd, gedocumenteerd: De godsdienstige geleerden uit Timbuktu in de negende en tiende Islamitische eeuw* [Learned, praised, and documented: The religious scholars of Timbuktu in the 9th and 10th centuries A.H.]. M.A. thesis, University of Groningen.

——. 1998. "De *ʿulamāʾ* van Timbuktu in de vijftiende en zestiende eeuw" [The *ʿulamāʾ* of Timbuktu in the 15th and 16th centuries). *S̲harqiyyāt* 10.85–104.

——. 2005. *Throwing stones at the moon: The role of Arabic in contemporary Mali*. Ph.D. diss., Leiden University.

Brenner, Louis. 1984. *West African Sufi: The religious heritage and spiritual search of Cerno Bokar Saalif Taal*. London: Hurst.

——. 1986. "Al-Hajj Sa'ad Umar Touré and the Islamic educational reform in Mali". *Language and education in Africa*, ed. Eckhard Breitinger and Reinhard Sander, 5–23. Bayreuth: African Studies.

——. 1991. "Medersas au Mali: Transformation d'une institution islamique". *L'enseignement islamique au Mali*, ed. Bintou Sanankoua and Louis Brenner, 63–85. Bamako: Editions Jamana.

——. 1993. "Constructing Muslim identities in Mali". *Muslim identity and social change in sub-Saharan Africa*, ed. Louis Brenner, 59–78. London: Hurst.

——. 2000. *Controlling knowledge: Religion, power and schooling in a West African Muslim society*. Bloomington: Indiana University Press.

Cissé, Seydou. 1975. *Tombouctou et l'empire Songhay: Epanouissement du Soudan nigérien aux XV^e–XVI^e siècles*. Dakar: Les Nouvelles Editions Africaines.

——. 1992. *L'enseignement islamique en Afrique Noire*. Paris: L'Harmattan.

Cissoko, Sékéné Mody. 1969. "L'intelligentsia de Tombouctou aux XV^e et XVI^e siècles". *Présence Africaine* 72.48–72.

Clarke, Peter B. 1982. *West Africa and Islam: A study of religious development from the 8th to the 20th century*. London: Arnold.

Diakite, Drissa. 1991. "Les fondements historiques de l'enseignement islamique au Mali". *L'enseignement islamique au Mali*, ed. Bintou Sanankoua and Louis Brenner, 25–44. Bamako: Editions Jamana.

Dumestre, Gérard. 1997. "De l'école au Mali". *Nordic Journal of African Studies* 6.31–55.

Goody, Jack. 1971. "The impact of Islamic writing on the oral cultures of West Africa". *Cahiers d'Etudes Africaines* 11.455–466.

Harrison, Christopher. 1988. *France and Islam in West Africa, 1860–1960*. Cambridge: Cambridge University Press.

Hiskett, Mervyn. 1994. *The course of Islam in Africa*. Edinburgh: Edinburgh University Press.

Hunwick, John O. 1962. "Aḥmad Bābā and the Moroccan invasion of the Sudan (1591)". *Journal of the Historical Society of Nigeria* 10.311–322.

——. 1966. "Further light on Aḥmad Bābā al-Tinbuktī". *Research Bulletin of the Centre of Arabic Documentation* (Ibadan) 1:2.19–31.

Kaba, Lansiné. 1974. *The* Wahhabiyya*: Islamic reform and politics in French West Africa*. Evanston, Ill.: Northwestern University Press.

——. 1984. "The pen, the sword, and the crown: Islam and revolution reconsidered, 1464–1493". *Journal of African History* 25.241–256.

Levtzion, Nehemia. 1973. *Ancient Ghana and Mali*. London: Methuen.

——. 1977. "The western Maghrib and the Sudan: The hegemony of the Songhay". *The Cambridge history of Islam* III, ed. J.D. Fage and Roland Olivier, 415–462. Cambridge: Cambridge University Press.

——. 1979. "Toward a comparative study of Islamization". *Conversion to Islam*, ed. Nehemia Levtzion, 1–23. New York: HM Publishers.

——. 1986a. "Rural and urban Islam in West Africa: An introductory essay". *Rural and urban Islam in West Africa*, ed. Nehemia Levtzion and Humphrey John Fisher, 7–26. Haifa: University of Haifa.

——. 1986b. "Mamlūk Egypt and Takrūr (West Africa)". *Studies in Arabic history and civilization in honour of David Ayalon*, ed. Moshe Sharon, 183–207. Leiden: E.J. Brill. (Repr., Nehemia Levtzion, *Islam in West Africa: Religion, society and politics to 1800*. Aldershot: Variorum, 1994.)

Mbiti, John S. 1969. "Islam in Africa". *African religions and philosophy*, 242–254. London: Heineman.

Mommersteeg, Geert. 1996. *Het domein van de mar-*

about: Koranleraren en magisch-religieuze specialisten in Djenné. Mali [The domain of the marabout: Qurʾānic teachers and specialists in religious magic in Djenné, Mali]. Amsterdam: Thesis Publishers.
Saad, Elias N. 1983. *Social history of Timbuktu: The role of Muslim scholars and notables, 1400–1900.* Cambridge: Cambridge University Press.
Tamari, Tal. 2002. "Islamic higher education in West Africa: Some examples from Mali". *Islam in Africa: Yearbook of the sociology of Islam* IV, ed. Thomas Bierschenk and Georg Stauth, 91–128. Münster: Lit Verlag.
Zappa, Francesco. 2004. *L'islamizzazione della lingua bambara in Mali tra pubblicistica scritta e letteratura orale.* Pisa and Rome: Istituti Editoriali e Poligrafici Internazionali.

Dinie Bouwman (University of Groningen)

Malta

Situated 93 km south of Sicily, 288 km east of Tunisia, and 355 km north of Libya, the island of Malta presents an intriguing linguistic situation. Although nearer to Sicily and culturally European, its inhabitants still speak a language that is basically a variety of Arabic, albeit a "highly deviant offshoot of vernacular Arabic" (Borg 1997:271, see also 245–247; Kaye and Rosenhouse 1997:263). The deviation derives from two parallel processes: while the original Arabic dialect lost contact with written and spoken Arabic in the 13th century, the urbanization of the harbor area in the 16th century introduced a koine which was heavily romanized by daily contacts with Sicilian and Italian. This variety was standardized in the 19th century and became modern Maltese. The disproportion between Malta's size (246 km^2) and its population (400,000 inhabitants), as well as the large number of Sicilian, Italian, and English surnames, shows the importance of immigration in Malta's linguistic development (Brincat 2004b).

The language spoken in prehistory might have been a Mediterranean language, according to the traditional theory, but since Renfrew (1987) linked the diffusion of the Indo-European linguistic family with the spread of agriculture, it seems likely that the temple builders, who came over from Sicily, spoke an Indo-European language. The first inscriptions found in Malta date back to the 6th century B.C.E. and are in Punic. The Romans introduced Latin in 218 B.C.E., but for two centuries three languages were in formal use, Punic, Greek, and Latin. However, St. Luke's definition of the islanders as "barbarians" (Acts 28:1–11) does not specify whether in 60 C.E. they spoke Punic or a local variety of vulgar Latin or Greek. The 600 years of Roman rule could have changed the language of a small community (under 10,000) and so would have the succeeding Byzantine period, 350 years long. Although isolation might have maintained Punic as the spoken language up to 870, the Punic origin of Maltese was a myth introduced by Jean Quintin (1536) and repeated by J.H. Majus (1718) before Punic script was deciphered by Barthélemy in 1758. The confusion in the early classificatory attempts is illustrated by Megiser's (1606:8) definition of Maltese: "im reden der Saracenischen, Mohrischen, oder Carthaginensischen oder der lingua Punica, welche ein art von der Arabischen ist und ihren Ursprung auss der Hebreischen hat". Yet, the Punic theory was fondly defended by Maltese scholars (De Soldanis 1750, Vassalli 1791, Magri 1907) up to the early 20th century because the connection with a glorious extinct civilization satisfied national pride and was exploited for political reasons by the imperialists (Strickland 1920). Nonetheless, Gian Francesco Abela (1647) was already aware of the Arabic origins of Maltese and knew that Arabic was widely spoken in Sicily under the Normans and in Pantelleria in his times. Although J.J. Bellermann opted for the Punic theory in 1809, W. Gesenius refuted it scientifically in 1810 (see Kontzi 1994).

The roots of the Maltese language date back to around 256/870, or perhaps to 440/1048. Arab geographers and historians gave little attention to Malta, and this is taken as proof of the island's insignificance during that period (Redjāla 1973). An exception is al-Ḥimyarī's description, based on sources almost contemporary with the events, which was unknown to Maltese historians until 1990 (Brincat 1995). The *Rawḍ al-miʿṭār* describes a ferocious raid in 255/870 that destroyed the Byzantine social structure and says that the island was subsequently only visited for wood, honey, and fishing until a new community settled there in 1048–1049. This explains the lack of a pre-Arabic substratum in present-day Maltese (see Borg 1996a): the descendants of the survivors of the 870 raid were too few to influence the new settlers' language. Distinguishing between 'Muslims' and their more numerous 'slaves', al-Ḥimyarī suggests a religiously mixed but homogeneously Arabophone community. This may have come from Sicily, where Norman raids had already begun, but deeper comparative studies are needed to provide conclusive evidence (see Agiùs 1996; Dessoulavy 1938; Kontzi 1993). Another intriguing point is that the conquerors of 870 might have introduced Berber, unlike the 1048 settlers who apparently spoke the variety of Arabic that had been forged in Sicily, untouched by the second wave of Arabicization brought by the Banū Hilāl into North Africa. According to Vanhove (1998:97–98), this accounts for the main differences between Maltese and the Maghreb dialects.

The Norman conquest of 1091 introduced contacts with Romance languages and dialects, but the use of Classical Arabic is witnessed by 12th-century tombstones, verses written at the court of Roger II by poets called 'al-Māliṭi', and by the Arabic version of the first known document in Latin addressed to the rulers of Malta, signed by Queen Constance in 1198. Although Christians were present on the island in 1091 and certainly comprised locals in 1150, the change in culture, religion, and population was gradual because in 1175 Bishop Burchardt described Malta as inhabited by Saracens, and in 1241 Giliberto Abbate reported the presence of 836 Muslim families. Thus, Malta was no different than Sicily, where Arabic was still widely spoken around Palermo and Agrigento (Varvaro 1981:150), and Christian rites were still held in Arabic in 1330 (Ludolphus 1851). After the Muslims' expulsion in 1224 and 1249, the linguistic history of the two islands diverged because in Malta total Christianization was not accompanied by full Latinization. Romance speakers increased under the Angevins, the Aragonese, and the Castillians, and many settled permanently, were slowly absorbed, and adopted Maltese speech. All documents were drawn up in Latin and Sicilian, while Arabic writing was only used by the Jewish community until 1492 (Wettinger 1985:174–204).

When the Maltese variety lost contact with Qurʾānic and Classical Arabic, it went through a process of phonological readjustment, morphological simplification, and constant lexical growth by Sicilian and Italian accretions (Cremona 1990; Krier 1976; Brincat 2004a:109–167, 344–366). Its speakers' perception was strikingly different from that of foreigners. While the latter tended to fit the local tongue into a genealogical or areal classification, calling it *parlata africana* (1536), *parlar saracino* (1558), *lingua degli Africani* (1567), or *un langage Arabe corrompu* (1694), the Maltese saw it from an autonomous point of view: *in lingua maltensi* (1436), *in lingua nostra maltensi* (1525), *in lingua melitea* (1540), or *in melivetana et vernacula lingua* (1554; see Cassola 1991–1992). Significantly, the first documented words since 1241, and the earliest known full text in Maltese, a Cantilena written about 1470 by Pietro Caxaro (Wettinger and Fsadni 1968; Brincat 1999), are in the Latin alphabet. The locals never wrote in Arabic script, although some scholars did propose the use of some Arabic letters in the 17th and 18th centuries when the first grammars and word lists were drawn up. The Knights introduced Italian as the language of administration and culture, but, being supranational, they never imposed a linguistic policy for the people, and thus Maltese survived. Moreover, the development of a dynamic urban society around Grand Harbour created the right conditions for the koine which became standard Maltese. Areal varieties were summarily described by Vassalli in 1796 and are still perceivable, but a comprehensive systematic study is lacking (see Schabert 1976; Aquilina and Isserlin 1981; Agiùs 1991; Kaye and Rosenhouse 1997; Vanhove 1999).

Due to Saracenic raids and the Ottoman threats that culminated in the Siege of 1565, the prestige of Arabic was low because it was only spoken by slaves. Although commercial contacts with North Africa must have existed, they were probably negotiated in the lingua franca (whose influence on Maltese deserves further investigation; see Borg 1996b). Surprisingly, interest in Arabic was revived by the Roman Catholic Church. In 1622, the Roman Congregazione de Propaganda Fide decided to establish the study of Arabic in Malta to prepare missionaries for spreading Christianity in Arab countries, and lessons started in 1632. One of the most important Maltese Arabists was Michele Antonio Vassalli, who furthered his studies in Rome and became a lecturer in Oriental languages at La Sapienza University. Vassalli considered Arabic useful for a deeper knowledge of Maltese and wrote a scientific grammar (1791) and dictionary (1796) of Maltese. He also drew up plans for teaching Maltese, Italian, and Arabic in the schools and abandoned the Punic myth in later works (1827).

Linguistic policies were launched by the French (1798) and the British (1813). The former were unsuccessful, and the British struggled to substitute English for Italian, but they promoted Maltese and encouraged the teaching of Arabic for academic reasons. However, in line with the naturalistic concepts of the comparativist school that privileged genealogical classification, they fostered a puristic approach, which would bring Maltese closer to its parent language, and provoked strong reactions among the educated classes, especially George P. Badger's suggestion that Arabic should spread easily in the Maltese islands and should become "in a short period the established language of the people" (1838:296–298). Academic prejudice against dialects threw Maltese into a precarious position. Diglossia, with Italian as the cultural language available to around 10 percent of the population and with spoken Maltese reigning supreme among illiterate monolinguals, was not ideal from the educational point of view, but it safeguarded the local idiom and ethnic identity. In the 1840s that balance was broken, and the defense of Italian, the introduction of English, and the confusion of Maltese with Arabic threatened the local language with extinction. Although John Hookham Frère had encouraged the study of Maltese since 1821 and established a chair at the university for Vassalli, Badger proposed teaching Arabic instead of Maltese in the schools because "the dialect is already corrupt" (1838) and insisted that in the schools Maltese should only be employed as a medium of instruction (1841). This attitude sparked off the Language Question which raged on for a hundred years. The battle was fought in all the institutions – the schools, the Church, the Law Courts, and the House of Representatives – where for seventy years members spoke in Italian or English, according to their stand for the nation or for the empire. Being mainly a question of identity, it was also fought in the literary field with interventions in and for English, Italian, or Maltese (Brincat 2001).

After the unification of Italy, the British twinned the building of coastal fortresses with stronger efforts to replace Italian with English. In the meantime, Maltese, which had been promoted by the locals throughout the 19th century as a literary medium, acquired the status of a 'language', according to the criteria of the times, and advanced steadily. Neither the purists, who would have stifled the language, nor the denigrators, who would have replaced it by Italian or English, had their way, and so the empirical approach prevailed. On the one hand, the high-quality verses of Dun Karm Psaila conferred on Maltese the prestige it formerly lacked. On the other, the rise of the political parties in a democratic environment made the leaders realize that their message would not reach the monolingual masses, whose support was now indispensable, unless it were delivered in their own tongue. The use of Maltese was allowed in parliamentary debates in 1921, in notary deeds in 1927, and in parliamentary records in 1933, and it obtained official status in 1934, next to English and Italian. In 1935 the university introduced the matriculation exam in Maltese, and the local channel of cable radio was set up and spread the standard variety to all the homes in Malta and Gozo. Italian was dropped in 1936, and public notices, street names, and Christian names were changed. English became compulsory for employment in the armed forces and the civil service, and the war dealt a decisive blow to the prestige Italian had enjoyed before 1939. Ironically, Anglicization spread faster in the 1950s when it no longer interested the British government. After World War II, Maltese society changed drastically: introduced in 1946, compulsory education strengthened the use of Maltese and English, while exposure to English increased with the cinema, pop music, and all things modern. In the meantime, Maltese kept pace in administrative domains and in higher education, thanks to the institution of a chair of Maltese (Joseph Aquilina in 1937). It became a compulsory subject for entry into the university (1945), and degree courses in Maltese language and literature were set up in 1948. Italian regained popularity in the late 1950s when television mended the image created by prewar propaganda. In 1964, a local station started broadcasting home-produced programs in Maltese and imported ones in English, but Italian channels continued to attract the majority of the audience up to 1996, when the political parties set up their own stations. Together with the introduction of satellite and cable television, they brought Italian viewership down to around 20 percent. In 1975, the study of Arabic was made compulsory at the secondary level, but the government's decision, aimed at developing commercial ties with North Africa, was not popular, and results were disappointing. It was therefore made optional in 1986, and now very few students choose it. Yet, at the university level, the quality of its teaching has improved, although numbers remain small, and, after a long break, which unfortunately included the notorious 18th-century forgeries of abate Giuseppe Vella (Agiùs 1990:28–30; Freller 2001), scientific contributions are being published (Zammit 2002).

The historical events in the last millennium have produced a community that has been exposed to different languages for many centuries. The census held in 1995 revealed that out of a population of 324,386 persons aged sixteen and over, 317,311 speak Maltese, 246,157 learned English well, and 118,213 know Italian, while languages studied only as school subjects follow at a certain distance: French (31,945), German (6,807), Arabic (5,955), and Spanish (1,955). The first result of this strong exposure to and widespread use of different languages is the ongoing development of the native language. Under the cumulative effect of nine hundred years of contacts, with foreign rulers and administrators as well as settlers at lower social levels, the stratification of the Maltese language evolved rapidly and substantially. The oldest layer, which is perceived as the main stratum (proving its Arabic origins), only provides 32.41 percent of the lexemes in Aquilina's *Maltese-English Dictionary* (1987–1990). Sicilian and Italian account for 52.46 percent, while English words make up 6.12 percent of the total of 41,000 lexemes. Consequently, words of non-Arabic origin form 60.23 percent of the lexicon, although the quantitative factor is offset by frequency. In fact, Arabic words prevail in any text because function words are repeated many times and basic terms occur more often than specialized terms. Although there are only 6 Romance words in the fundamental 100-word list, at the threshold level 52.8 percent of 1,585 words are of Romance origin. However, the lexical composition of texts varies greatly according to register or domain. For detailed morphosyntactic descriptions of Maltese, see Borg and Azzopardi-Alexander (1997); Mifsud (1995); Vanhove (1993); and → Maltese.

After more than fifty years of compulsory education where Maltese and English are not only taught as subjects but as the medium of instruction, and in spite of the practice of code-switching in informal conversation, there does not seem to be a real danger of English words eroding the core vocabulary, provided the two codes are kept apart by the schools and by social awareness. At present, Maltese is still spoken regularly by over 90 percent of the inhabitants, although most of them consider English indispensable in today's world. Interest in Italian is still alive, but it is no longer seen as a threat to Maltese. The latter has penetrated areas that a few decades ago were dominated by English, namely the written register and official use. Maltese is used regularly in Parliament, the Law Courts, the Church, all government offices, and in banks and private firms, and it has just been recognized as one of the official languages of the European Union. It is therefore the only "peripheral erstwhile Arabic dialect" (Borg 2001:422) to enjoy such a status. If official policy succeeds in achieving bilingualism with good competence in both languages, Maltese will survive so long as it is still considered important and efficient by its speakers. The community's knowledge of English may even ensure the survival of the local language, since the use of English for international communication will ease the pressures on Maltese to change and grow too rapidly.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Abela, Gio Francesco. 1647. *Della descrittione di Malta isola nel mare siciliano*. Malta. (Repr., Malta: Midsea Books, 1984.)

Agius De Soldanis, Giovan Pietro Francesco. 1750. *Della lingua punica presentemente usata da' Maltesi, etc.* Rome.

Badger, George Percy. 1838. *Description of Malta and Gozo*. Valletta.

——. 1841. *A letter on the eligibility of the Maltese dialect*. Malta.

Ludolphus de Suchem. 1851. *Ludolphi, Rectoris ecclesiae parochialis in Suchem, De Itinere Terrae Sanctae liber*, ed. F. Deucks. Stuttgart.

Majus, 1718. *Specimen linguae punicae in hodierna Melitensium superstitis*. Marburg.

Megiser Hieronymus. 1606. *Propugnaculum Europae*. Leipzig, ed. Albert Friggieri and Thomas Freller, *Malta: The bulwark of Europe*. Malta: Gutenberg Press, 1998.

Quintin, Jean. 1536. *The earliest description of Malta (Lyons 1536) by Jean Quintin d'Autun*, ed. Horatio Vella. Malta. 1980.

Vassalli, Michele Antonio. 1791. *Mylsen Phoenico-Punicum sive Grammatica Melitensis*, Rome.

——. 1796. *Lexicon Melitense-Latino-Italum (Ktyb yl Klym Malti)*. Rome.

——. 1827. *Grammatica della lingua maltese.* Malta.

SECONDARY SOURCES

Agiùs, Dionisius. 1990. *The study of Arabic in Malta, 1632 to 1915*. Louvain: Peeters.

——. 1991. "Morphological alternatives in the Gozitan dialects of Maltese". *Matériaux Arabes et Sudarabiques* 4.111–161.

——. 1996. *Siculo Arabic*. London and New York: Kegan Paul International.

Aquilina, Joseph. 1987–1990. *Maltese-English Dictionary*. 2 vols. Malta: Midsea Books.

——, and Benedict Isserlin. 1981. *A survey of contemporary dialectal Maltese*. I. *Gozo*. Leeds.

Borg, Albert and Marie Azzopardi-Alexander. 1997. *Maltese*. London: Routledge.

Borg, Alexander. 1978. *A historical and comparative phonology and morphology of Maltese*. Ph.D. diss., Hebrew University.

——. 1996a. "On some Levantine linguistic traits in Maltese". *Israel Oriental Studies: Studies in modern Semitic languages*, ed. Shlomo Izre'el and Shlomo Raz, 148–152. Leiden: E.J. Brill.

——. 1996b. "On some Mediterranean influences on the lexicon of Maltese". *Romania Arabica*, ed. Jens Lüdtke, 129–150. Tübingen: G. Narr.

——. 1997. "Maltese phonology". *Phonologies of Asia and Africa*, ed. Alan S. Kaye, I, 245–285. Winona Lake, Ind.: Eisenbrauns.

——. 2001. "Historical reflections on the indigenous grammar of Maltese". *Indigenous grammar across cultures*, ed. Hannes Kniffka, 421–432. Frankfurt a. Main: P. Lang.

Brincat, Joseph M. 1995. *Malta, 870–1054: Al-Ḥimyari's account and its linguistic implications*. Malta: Said International.

——. 1999. "The Cantilena. Vintura. Why? Who?". *Karissime Gotifride*, ed. P. Xuereb, 177–183. Malta: Malta University Press.

——. 2001. "The language question and education: A political controversy on a linguistic topic". *Yesterday's schools: Readings in Maltese educational history*, ed. Ronald Sultana, 137–158. Malta: PEG.

—— [Giuseppe]. 2004a. *Malta: Una storia linguistica*. Genova: Le Mani.

——. 2004b. "Surnames in Malta: What can they tell us?" *Sunday Times of Malta,* 3 February 2004, 48–49.

Cassola, Arnold. 1991–1992. "La lingua maltese nel Cinquecento attraverso i resoconti di scrittori e viaggiatori stranieri". *Journal of Maltese Studies* 21–22.79–90.

Cremona, Joseph. 1990. "The Romance element in Maltese: A review of the problems". *Transactions of the Philological Society* 88:2.163–199.

Dessoulavy, Charles Louis. 1938. *A Maltese-Arabic word-list*. London: Luzac.

Freller, Thomas. 2001. *The rise and fall of Abate Giuseppe Vella*. Malta: PIN.

Kaye, Alan and Judith Rosenhouse. 1997. "Arabic dialects and Maltese". *The Semitic languages*, ed. Robert Hetzron, 263–311. London and New York: Routledge.

Kontzi, Reinholdt. 1993. "Maltesisch-Maghrebinischer Sprachvergleich anhand von Bibelübersetzungen". *Miscellanea arabica et islamica: Disertationes in Academia Ultrajectiana prolatae anno MCMXC* (Orientalia Lovaniensa Analecta) 52.3–21.

——. 1994. "Der Beitrag deutscher Gelehrter zur Erforschung des Maltesischen seit Beginn des 17. Jahrhunderts bis zum Beginn des 20. Jahrhunderts". *Lingua et traditio: Geschichte der Sprachwissenschaft und der neueren Philologien*, ed. Richard Baum a.o., 231–257. Tübingen: G. Narr.

Krier, Fernande. 1976. *Le maltais au contact de l'italien*. Hamburg: H. Buske.

Magri, Emmanuel. 1907. "Précis de mythologie Maltaise". *Actes du XIV Congrès International des Orientalistes, Alger, Deuxième section*. Paris.

Mifsud, Manwel. 1995. *Loan verbs in Maltese: A descriptive and comparative study*. Leiden: E.J. Brill.

Prevaes, Mathias Hubertus. 1993. *The emergence of standard Maltese: The Arabic factor*. Ph.D. diss., University of Nijmegen.

Redjāla, Mbarek. 1973. "L'archipel maltais dans la littérature historico-géographique d'expression arabe à l'époque médiévale". *Proceedings I Congress on Mediterranean Studies of Arabo-Berber influence*, 203–208. Algiers.

Renfrew, Colin. 1987. *Archaeology and language: The puzzle of Indo-European origins*. London: Cape.

Schabert, Peter. 1976. *Laut- und Formenlehre des Maltesischen anhand zweier Mundarten*. Erlangen: Palm und Enke.

Strickland, Gerald. 1920. *Malta and the Phoenicians*. Malta.

Vanhove, Martine. 1993. *La langue maltaise: Etudes syntaxiques d'un dialecte arabe périphérique*. Wiesbaden: O. Harrassowitz.

——. 1998. "De quelques traits préhilaliens en Maltais". *Peuplement et arabisation au Maghreb occidental: Dialectologie et histoire*, ed. Jordi Aguadé, Patrice Cressier, and Ángeles Vicente, 97–108. Madrid: Casa de Velázquez, Universidad de Zaragoza.

——. 1999. "La dialectologie du Maltais et son histoire". *Revue d'ethnolinguistique (Cahiers du LACITO)* 8.171–191.

Varvaro, Alberto. 1981. *Lingua e storia in Sicilia*. Palermo: Sellerio.

Wettinger, Godfrey. 1985. *The Jews of Malta in the late Middle Ages*. Malta: Midsea Books.

——, and Michael Fsadni. 1968. *Peter Caxaro's Cantilena: A poem in Medieval Maltese*. Malta.

Zammit, Martin R. 2002. *A comparative lexical study of Qur'ānic Arabic*. Leiden: E.J. Brill.

JOSEPH M. BRINCAT (University of Malta)

Maltese

1. General

The affiliation of Maltese within other Arabic vernaculars is a controversial issue. It is likely that successive waves of impact reached the Maltese shores from different Arab stations and at different points in the island's history (for theories about a Phoenician origin of Maltese, → Malta). Most linguists (see, for example, Aquilina 1961, 1979) agree that typologically Maltese fits well into the general characteristics of Maghrebi dialects, including the most distinctive isoglosses such as the *n-* prefix for the 1st person plural of the imperfect (*niktbu* 'we write', *nimxu* 'we walk'). Some curious similarities with the Eastern dialects led others to posit the theory of an eastern extraction (cf. Stumme 1904) or, recently and more mildly, to point out suggestive Levantine traits in the language (Alex. Borg 1997). A recent theory developed from the former would regard Maltese as a brand of Sicilian Arabic implanted here at a later date (Agius 1996).

Today, Maltese is considered by its speech community to be a language on its own, and scholars of Arabic vernaculars tend to treat it separately (see, for example, Kaye and Rosenhouse 1997). The early loss of → diglossia with any form of Arabic standard and the massive influence of European languages, which it has subsequently undergone, are among the most important reasons to treat Maltese separately. Maltese has gone through a long and close contact and the Arabic dialect rooted on the island has gone through a fusion with the European languages which came in contact with it: medieval Sicilian, mainland Italian in its Tuscan form, and later English. This earned it the epithet of 'mixed language', a term which undoubtedly needs a clearer definition.

2. Linguistic description

2.1 *Phonology*

2.1.1 Consonants

The phonological inventory of standard Maltese has 24 consonantal phonemes (see Table 1).

Maltese has lost no fewer than eleven consonantal phonemes, seven of which are recorded in old Maltese documents. Positionally, five of the lost phonemes lie from the velum backward. Maltese has added seven new phonemes from non-Arabic sources, of which only one lies in this rear area. The changes have brought the general sound of the language closer to that of its languages of contact, from which only the glottal stop /ʔ/ now sets it apart.

The emphatic consonants of Maltese Arabic have merged with their nonemphatic counterparts (/ṣ/ > /s/, /ṭ/ > /t/, /ḍ/ + /ḏ̣/ > /d/), presumably at a very early stage, and there are no traces of their existence in any document of the language. Maltese has also merged the rest of the interdental series, which, on the contrary, are attested in old documents from the 16th to the 18th centuries (cf. H. Megiser, reproduced in Friggieri and Freller 1998, note: *veheb, fne*; Vassalli 1796:xxix n. 11); they are now represented by their alveolar counterparts (/ṯ/>/t/ and /ḏ/ > /d/). Of the velar and pharyngeal fricatives, only the voiceless pharyngeal /ḥ/ survives, although records (cf. Agius De Soldanis 1750; Wzzino 1752; Vassalli 1791) point to a distinction still made in the 18th century

Table 1. The consonantal phonemes in standard Maltese (phonemes in **bold** are subsequent additions to the Arabic repertory)

	Bilabial	Labio-dental	Alveolar	Post-alveolar	Palatal	Velar	Uvular	Pharyn-geal	Glottal
Plosive	**p** b		t d			k **g**			ʔ
Nasal	m		n						
Trill			r						
Affricate			**ʦ** **ʣ**	**tʃ** ʤ					
Fricative		f **v**	s z	ʃ **ʒ**				ħ	
Approximant	w				j				
Lateral			l						

between it and the voiceless velar fricative /x/, with which it is now merged. The latter may still be heard as a conditioned allophone of the former (Aquilina and Isserlin 1981:135). Although they are now indiscriminately represented by the digraph *għ* in the orthography, /ġ/ and /ʿ/ are both muted in standard Maltese (e.g. *xogħol* [ʃɔːl]; *xagħar* [ʃɐːr]) and are realized as /ḥ/ in definable morphophonemic contexts (e.g. *qlugħ* [ʔlʊːħ]) 'sails', *tefagħhom* [tɛˈfɐħːɔm]. However, the velar sound /ġ/ survives in the dialectal speech of a few speakers in remote villages such as Għarb in Gozo. The Arabic glottal stop /ʾ/ has been enfeebled to the point of losing its phonemic status, but the sound now exists in Maltese as the reflex of Arabic /q/. However, the uvular plosive survives in the speech of presumably few conservative speakers in Isla (Senglea). The voiceless glottal fricative /h/, though regularly represented in the orthography by *h*, has been muted in standard Maltese (e.g. *dahri* [ˈdɐːrɪ] 'my back'), but is pronounced as [ħ] in certain conditioned contexts, such as word-final position (*ikrah* [ˈɪkrɐħ] 'ugly'). Consonantal reflexes of /h/ may still be heard in the pronunciation of some older village speakers (e.g. *dahri* [ˈdɐħrɪ], *xahar* [ˈʃɐħɐr], *h(e)diet* [ħdɪːt]), irrespective of its position (for a survey of the mergers, see Table 2).

Table 2. Historical phonemic mergers

/t/	/t/
/ṭ/	
/t̠/	
/s/	/s/
/ṣ/	
/d/	/d/
/d̠/	
/ḍ/	
/ḍ̠/	

/x/		/ḥ/
/ḥ/		
/ʿ/		<għ>
/ġ/		
/q/	>	/ʾ/
/ʾ/	>	ø
/h/	>	<h>

Maltese has a rich dialectal situation which has yet to be fully documented before it disappears. The data available are still fragmentary and have been collected and analyzed at different times and with different aims and criteria (cf. Aquilina and Isserlin 1981).

The merging of a good number of phonemes and the muting of others inevitably resulted in the convergence of several pairs of root-bases originally differing only in one radical, less frequently in two, e.g. Arabic *ṣ-y-f* 'summer' + *s-y-f* 'sword' > Maltese *s-j-f* conveying both meanings; Arabic *š-ġ-l* 'work' + *š-ʿ-l* 'to light up' > Maltese *x-għ-l*. This often resulted in (i) homophones, e.g. [dɐːr] = *dar* 'a house' or *dahar* 'a back', [ˈɛːmɛs] = *hemeż* 'to attach' or *għemeż* 'to wink'; or (ii) homographs, e.g. *sħab* = 'clouds' (< *s-ḥ-b*) or 'partners' (< *ṣ-ḥ-b*), *għereq* = 'he sweated' (< *ʿ-r-q*) or 'he sank' (< *ġ-r-q*), *għażel* = 'he chose' (< *ʿ-z-l*) or 'he spun' (< *ġ-z-l*), *nstamat* = 'he was scalded' (< *s-m-ṭ*) or 'he was dumbfounded' (< *ṣ-m-t*).

To this impoverished consonantal repertory, Romance and English contributed through the accretion of no fewer than seven consonantal phonemes: /p/, /g/, /ts/, /dz/, /tš/, /v/, and /ž/. Only /ts/ and /dz/ create a new 'gap', although [ts] could have resulted from the adjacency of [t] or [d] to [s] in words of Arabic origin like *għatsa* [ˈɐːtsːɐ] 'a sneeze' and *qaddsu* [ˈʔɐtsːʊ] 'they sanctified'. The phoneme /ž/, presumably already present as the voiced version of /š/ in words like [ʒbɐːr] 'spans', is a relatively recent acquisition from English, and so far occurs in unconditioned environment only in a few recent loans such as *televixin* [tɛlɛˈvɪʒɪn] and *bex* [bɛːʒ]; both /š/ and /ž/ are represented by <x> in the orthography. Some linguists (e.g. Alb. Borg and Azzopardi-Alexander 1997:303) are cautious in accepting the phonemic status of this consonant in view of its rarity and the difficulty in finding minimal pairs which would definitely confirm this status.

There are in Maltese automatic constraints affecting the voicing feature of obstruents in definable phonological environments. Thus, a voiceless obstruent is normally voiced when it is immediately followed by a voiced segment (e.g. *jisbaħ* [ˈjɪzbɐħ] 'dawn breaks'), and a voiced obstruent is unvoiced when followed by a voiceless segment (e.g. *bsarna* [ˈpsɐrnɐ] 'we foresaw'). Moreover, a voiced obstruent is regularly devoiced in word-final position (e.g. *ħabib* [ħɐˈbiːp] 'friend', *tieġ* [tɪːtʃ] 'wedding'). This rule, however, seems to gradually become less obligatory in the Maltese pronunciation of loans from English, especially in bilingual speakers. On the other hand, the frequency of nondevoiced final voiced consonants in the orthography of old Maltese deeds (see, for example, *calb, tirag, ard* in P. Caxaro's medieval Cantilena) seems to suggest that it might be possible to date this obligatory devoicing rule with a proper examination of documents.

Maltese displays several cases of combinatory phenomena between phonemes which are adjacent or near to each other. The frequent cases of assimilation involve mainly the voicing feature and are sometimes sanctioned by diachrony (e.g. *gideb* 'to lie' < Arabic *kaḏab, ċaħad* 'to refuse' < *jaḥad, niggeż* 'to prick' < Arabic *nakaz*).

Cases of historical dissimilation involve mainly resonant segments and are found in both Arabic and Romance loans (e.g. *sinsla* 'backbone' < Arabic *silsila*, *ġunġlien* 'sesame seeds' < Arabic *juljulān, kurunell* 'colonel' < Italian *colonello*, *franella* 'flannel' < Italian *flanella*).

Diachronic cases of metathesis, especially from Romance, exist and are fully accepted and enshrined in the orthography (see, for example, *nofs* 'half' < Arabic *niṣf; ġibed* 'to pull' < Arabic *jaḏab; tema' (għ)* 'to feed' < Arabic *ʾaṭʿam; denfil* 'dolphin' < Italian *delfino; porvli* 'gunpowder' < Italian *polvere*, *korla* 'anger' < Italian *collera*).

2.1.2 Vowel nuclei

The phonology of Maltese operates with a system of eighteen vocalic nuclei, of which eleven are monophthongs (five short vowels and six long vowels) and seven are diphthongs (see Table 3).

Table 3. The vocalic nuclei of standard Maltese

Monophthongs		Diphthongs
	/iː/	/ɪʊ/
/ɪ/	/ɪː/	/ɛɪ/
/ɛ/	/ɛː/	/ɛʊ/
/ɐ/	/ɐː/	/ɐɪ/
/ɔ/	/ɔː/	/ɐʊ/
/ʊ/		/ɔɪ/
	/uː/	/ɔʊ/

The segments /ɪ/, /ɛ/, /ɐ/, /ɔ/, /ɪʊ/, /ɛɪ/, /ɛʊ/, /ɐɪ/, and /ɐʊ/ occur both in stressed and unstressed positions, while /iː/, /ɪː/, /ɛː/, /ɐː/, /ɔː/, /uː/ occur only in stressed position. Stressed /ʊ/ is rare.

With the exception of /ɔɪ/, which occurs only in loanwords, all the diphthongs are reflexes of Old Arabic /ay/ and /aw/, or secondary diphthongs developed within Maltese, mainly through the adjacency of *għ* to a following *ī or ū*. /ɪʊ/ is very rare, while the pairs /ɛɪ/ and /ɐɪ/, /ɐʊ/ and /ɔʊ/, as well as /ɛʊ/ and /ɔʊ/ in tonic position, often occur in free variation (e.g. *żgħir*: [zɛɪr] ~ [zɐɪr], *tiegħu*: [ˈtɪjɐʊ] ~ [ˈtɪjɔʊ], *għuda*: [ˈɛʊdɐ] ~ [ˈɔʊdɐ]).

Historical changes in the vocalism of Maltese have affected both the quantity and the quality of vowels. Like most North African vernaculars (*parlers non-différentiels*), Maltese has systematically lost Old Arabic short vowels in open syllables (Arabic *ṭarīq* > *triq,* Arabic *kilāb* > *klieb*, Arabic *katabt* > *ktibt*). Since similar vowel reduction occurs also with Romance Maltese material (Italian *popolo* > *poplu*, Italian *manico* > *manku*), it seems reasonable to believe that this process is attributable to the stress factor. The effect of this vowel loss is twofold: (i) the presence of many consonant clusters and the need to have frequent recourse to a euphonic vowel (usually *i*, referred to in Maltese grammar as *il-vokali tal-leħen*) to break unwieldy clusters made up of consonant + resonant (or *għ*) + consonant, or at word boundary, e.g. *jiksbu* 'they acquire', *joftqu* 'they unstitch', but *joħo̱rġu* 'they go out', *jifi̱rxu* 'they spread', *jixe̱għlu* 'they light up', *qalb i̱msaġar* 'in the middle of forests'; and (ii) the automatic development of stem allomorphs in verbal and nominal paradigms, e.g. *kiteb/ktib-t/kitb-et* 'he/I/she wrote', *nieżel/niżl-in* 'descending [masc. sg./pl.]'.

Final clusters made up of obstruent + resonant are broken up by inserting an epenthetic vowel, which now has phonemic status (e.g. **ħabl-* > *ħabe̱l*, **ġism-* > *ġise̱m*, **leħn-* > *leħe̱n*).

Maltese tends to merge short *a* and *i* into one phoneme, while *u* is retained. This change becomes very significant in cases where the merged vowels originally had distinctive features. Such is the case, for example, of a few residual Form II participles with both active and passive meanings (e.g. Maltese *mgħallem* 'teacher' and 'taught'; Maltese *mreddgħa* 'wet nurse' and 'breast-fed [fem.]'). These are the result of a merger between the Arabic active and passive participle forms, which were insufficiently marked by a short vowel (e.g. *muʿalli̱m* and *muʿalla̱m*).

The most important historical change in the vocalic content of Maltese is undoubtedly → *ʾimāla*, i.e. the fronting and raising of vowel /ɐ/ in normal conditions. Thus, the Arabic long vowel *ā* normally corresponds to a long high unglided vowel [ɪː], represented in the official orthography by the digraph <ie> (e.g. Arabic *salām* 'peace' > *sliem*, Arabic *xaddām* 'servant' > *ħaddiem* 'worker'). Short *a* also inflects to *e* and eventually to *i* if accented (Arabic *katab* 'he wrote' > *kiteb*, whence *kitibli* 'he wrote to me'). Conditions that inhibit *ʾimāla* include adjacency to back consonants /ḥ/ and /ʾ/ as well as to any historical emphatic consonant, e.g. Arabic *ṭalab* > *talab* 'he asked for', Arabic *naqaṣ* > *naqas* 'he decreased', Arabic *qatal* > *qatel* 'he killed', Arabic *saraq* > *seraq* 'he stole', Arabic *xajjāṭ* > *ħajjat* 'tailor').

One result of this change was a large variety of vowel sequences in both verbal and nominal forms. The 3rd person masculine singular of the perfect alone displays six different sequences (*a-a, a-e, e-a, e-e, i-e, o-o*) against three in Old Arabic (*a-a, a-i, a-u*). The fine distinction of meanings associated with the three classes of Arabic verbs could not hold good anymore, and phonology got here the better hand of morphology.

One of the most intriguing phenomena displayed by Maltese comes from the dialects in contrast with standard Maltese. In most of the Maltese dialect area, a long vowel *ā* regularly recedes and takes a color between *ō* and *ū*. This feature practically constitutes a line of demarcation between Maltese dialects as a group and standard Maltese, where it is totally absent. Cantineau (1960:100–101) compares this unconditioned → *tafxīm* of *ā* with the dialects of North Lebanon.

2.2 *Morphology*

The basic morphological structure of Maltese remains largely that of a typical dialect of Arabic, and one can find in it reflexes of most of the canonical forms of Old Arabic, albeit often supported by a small number of members. On the other hand, due to its severance from Arabic and its long and strong contact with languages with a different morphological type, Maltese may be the Arabic vernacular that has moved farthest from the original structure.

In its morphology, Maltese displays a tendency both to reduce the number of forms and to create innovative ones, often through the reorganization of existing elements. Besides, the productivity of the different Arabic canonical forms varies greatly and may be generally described as being in crisis, mainly due to the strong inroads which Romance and English have cut into the language through their eight-hundred-year-old contact. Due to its strong structural constraints, the root-based Arabic morphology of Maltese could not continue to cope with the full integration (i.e. *taʿrīb*) of Romance loans and eventually gave way to a hybridized schema of the stem-based type, incorporating both flexible elements of the native morphology and new features contracted from Sicilian and Italian (Mifsud 1995b). The morphology of Maltese is presently undergoing a slow but major typological shift from the Arabic root-centered structure to a more open concatenative morphology of the European type. This is leading, on one hand, to a constant search of paradigms with a stable stem rather than others which make use of stem allomorphs (hence, for example, the frequent use of verbal Form II at the expense of Form I and the neutralization of the intensive meaning associated with Form II), and on the other, to the frequent recourse to suffix morphemes in preference to internal or broken structures (e.g., the old Maltese plurals *sfuf* 'layers' and *iradi* 'lands' gave way to modern Maltese *saffi* and *artijiet*). This view of Maltese morphology finds support in the undeniable fact that hardly any of the hundreds of words that are added every year to the Maltese lexicon are being integrated according to the root-based patterns.

2.2.1 Pronouns

2.2.1.1 Personal pronouns

Personal pronouns are marked for person, gender, and number (see Table 4).

Just like some Maghrebi dialects, Maltese has lost the gender distinction in the 2nd person singular of the independent pronouns and uses the form with final *-i* for both genders. The gender distinction has been lost also in the plural forms, as is normally the case in sedentary dialects.

As in most forms of Arabic, the 1st person singular has a grammatically conditioned allomorph *-ni*, which occurs as the direct object of a verb after both consonant and vowel (*sawwatni* 'he beat me', *rani* 'he saw me').

There are two forms for each of the 1st, 2nd, and 3rd person masculine singular bound pronouns, one occurring after a consonant and the other after a vowel or semivowel (*sieqi* 'my foot' but *saqajja* 'my feet', *sabu* 'he found him'

Table 4. Personal pronouns

	Independent pronouns	Bound pronouns			
	(Subject)	(Direct object)			(Indirect object)
		after consonant	after vowel	followed by indirect object	
3rd sg. masc.	*hu/huwa*	*-u*	*-h*	*-hu-*	*-lu*
3rd sg. fem.	*hi/hija*	*-ha*		*-hie-*	*-lha*
2nd sg.	*int/inti*	*-ek/-ok*	*-k*	–	*-lek/-lok*
1st sg.	*jien/jiena*	*-i*	*-ja*	–	*-li*
3rd pl.	*huma*	*-hom*		*-hom-*	*-lhom*
2nd pl.	*intom*	*-kom*		–	*-lkom*
1st pl.	*aħna*	*-na*		–	*-lna*

but *rah* 'he saw him', *warajh* 'behind him').

The 2nd person singular has a phonologically conditioned allomorph -*ok* (indirect object -*lok*) which occurs after a syllable with vowel *o* (*ommok* 'your mother', *inħobbok* 'I love you', *ngħoddlok* 'I count for you').

The singular forms of the independent (subject) personal pronouns have short variants (*jien* etc.), which are used freely even within the speech of the same speaker. Since the subject is normally marked on the verb, the independent pronoun is optional, except in the case of emphasis or to avoid ambiguity.

The forms of the personal pronouns function also as a copula in nominal sentences, even when they are not strictly necessary, e.g. *it-tabib (huwa) ħija* 'the doctor is my brother'. When used as a copula, the independent pronouns (including the shorter variants) are negativized in the same manner as verbs, i.e. with *ma/m'*......*x* (see Table 5).

Table 5. The copula

	Copula	+ Negative
3rd sg. masc.	*hu/huwa*	*m'hux/m'huwiex*
3rd sg. fem.	*hi/hija*	*m'hix/m'hijiex*
2nd sg.	*int/inti*	*m'intx/m'intix*
1st sg.	*jien/jiena*	*m'iniex*
3rd pl.	*huma*	*m'humiex*
2nd pl.	*intom*	*m'intomx*
1st pl.	*aħna*	*m'aħniex*

Bound (or object) personal pronouns are suffixed forms denoting the possessor of a noun, the direct object of a verb, or the object of a preposition (e.g. *rasu* 'his head', *bagħtu* 'he sent him', *fuqu* 'upon him'). Preceded by *l* (harking back to the Arabic preposition *li* and Maltese *lil* 'to'), they also serve as the indirect object of a verb (*bagħatlu* 'he sent to him') and as such, may occur together with the direct object suffix (*bagħathulu* 'he sent it [masc.] to him').

With a transitive verb, the sequences in Table 6 are possible, in which every additional suffix can cause the stress (here marked by underlining) to recede.

Table 6. The receding stress

verb	*se̱raq*	'he stole'
verb + direct object	*sera̱qha*	'he stole her'
verb + indirect object	*sera̱qli*	'he stole from me'
verb + direct object (3rd pers. only) + indirect object	*seraqhi̱eli*	'he stole her from me'
verb + direct object + indirect object (+ negation)	*ma seraqhili̱x*	'he did not steal her from me'

Only the 3rd person object pronouns can stand between a verb and the indirect pronoun suffix (hence, *bagħathomli* 'he sent them to me', but not **bagħatkomli* 'he sent you to me', etc.).

Table 7. Demonstrative pronouns

	Near		Far
		+ article	
sg. masc.	*da/n(a)*	*dal-*	*dak(a)*
sg. fem.	*di/n(a)*	*dil-*	*dik(a)*
pl.	*daw/n(a)*	*dal-*	*dawk(a)*

2.2.1.2 Demonstrative pronouns

Maltese has two sets of demonstratives referring to entities that are relatively near to or relatively far from the speaker. They have the dual function of pronouns and adjectives and are marked for number and (in the singular) also for gender (see Table 7).

The longer forms (ending with *-a*) are less common and are encountered only in more formal speech or writing. Historically, Maltese had a variant set of demonstratives in which *he-* was prefixed to the above (e.g. *hedin, hedawk*); an example occurs in Caxaro's *Cantilena: Hactar min hedaun heme tred minne tamarra.*

Demonstrative adjectives always precede a definite noun. When used adjectivally, the near demonstrative can remain separate or it can be joined to a following definite article, e.g. *dal-kumment* 'this comment', *dit-tfajla* 'this young lady', *dal-ġrajjiet* 'these events'.

2.2.1.3 Relative pronouns

The relative pronoun is *li*, which alternates freely with a fuller form *illi* (*il-ġuvni li daħal* 'the young man who entered', *id-dar illi begħtu* 'the house you sold'). Both are invariable and regularly follow the head noun, but they may be omitted if followed by a nonfinite subordinate verb (e.g. *il-kotba magħżulin minnkom* 'the books chosen by you'). They may also introduce nominal clauses (e.g. *li għamiltu qabel ma jgħoddx* 'what you did before does not count', *għedtlek illi qed jidħak bina* 'I told you he was cheating us').

2.2.1.4 Interrogative pronouns

Maltese interrogative pronouns include *min* 'who', *xi* (or *x'* before a vowel or single consonant) 'what', and selective *liema* 'which, which one' (*min daħal?* 'who entered?', *xi rbaħt?* 'what did you win?', *x'kantajt?* 'what did you sing?' *liema trid?* 'which one do you want?').

As the object of a preposition, *xi* may take the form either of a suffix, *-iex*, or of a separate pronoun, *xiex* (e.g. *biex* 'with what?', *fiex* 'in what?', *għaliex* or *għal xiex* 'for what, why?', *mniex* or *minn xiex* 'from what?').

Interrogative pronouns normally introduce the question.

2.2.2 Adverbs

Some adverbs are loan translations into Arabic of their Sicilian or Italian counterparts, and, although they are composed of recognizable Arabic lexemes, they tend to fill the syntactic space occupied by the Romance model in the source language. Some (like *kultant* 'sometimes' < Arabic *kull* 'every'+ Romance *tant* 'so much') are clearly local hybrid formations. The Romance Maltese suffix *-ment* (< Italian *-mente*) is now a very productive tool for the formation of adverbs.

The following is a small sample of Maltese adverbs.

i. Adverbs of time: *qabel* 'before', *wara* 'after', *issa* 'now', *mbagħad* 'then', *llum* 'today', *għada* 'tomorrow', *pitgħada* 'the day after tomorrow', *pitpitgħada* 'two days hence', *lbieraħ* 'yesterday', *lbiraħtlula* 'the day before yesterday', *filgħodu* 'in the morning', *filgħaxija* 'in the evening', *kmieni* 'early', *tard* 'late';
ii. Adverbs of place: *hawn* 'here', *hemm* 'there', *taħt* 'underneath', *fuq* 'above', *quddiem* 'in front', *wara* 'behind', *kullimkien* 'everywhere', *ġewwa* 'inside', *barra* 'outside';
iii. Adverbs of manner: *sewwa* 'well', *tajjeb* 'well', *ħażin* 'badly', *malajr* 'quickly', *bil-mod* 'slowly', *ġentilment* 'gently';
iv. Adverbs of quantity: *ħafna* 'much', *ftit* 'a little', *iżjed* 'more', *iktar* 'more', *żżejjed* 'excessively', *bilkemm* 'hardly', *wisq* 'much, abundantly', *kważi* 'almost', *bil-wisq* 'exceedingly', *immensament* 'immensely';
v. Interrogative adverbs: *kif* 'how', *meta* 'when', *fejn* 'where', *min* 'who', *għaliex* 'why', *biex* 'with what', *fiex* 'in what';
vi. Other adverbs: *verament* 'really', *possibilment* 'possibly', *ċertament* 'certainly'.

2.2.3 Particles

2.2.3.1 The article

Nouns, and sometimes adjectives, are marked for definiteness with the article *l-* (< Arabic *al-*) preceding the word, but orthographically separated from it by a hyphen. The article is preceded by a euphonic vowel *i* when this is needed to break up the cluster C*l*C (*qara l-kotba* 'he read the books', *qrajt il-kotba* 'I read the books'). The euphonic vowel usually follows the article with nouns beginning with two consonants (e.g. *rajt l-imtieħen* 'I saw the windmills', *żar l-isptar* 'he visited the hospital', but *qrajt il-ktieb* 'I read the book').

The article assimilates to the 'sun letters', which in Maltese are: *ċ* [tʃ], *d* [d], *n* [n], *r* [r], *s* [s], *t* [t], voiceless *x* [ʃ], *ż* [z], *z* both voiceless [ʦ] and voiced [dz], but not *ġ* [ʤ].

There is a tendency to use the article with adjectives only when they are contrastive (e.g. *il-laqgħat importanti li attendejt* 'the important meetings which I attended', but *il-laqgħat ġenerali m'attendejthomx kollha imma attendejt il-laqgħat l-importanti* 'I did not attend all the general meetings but I did attend the important ones.'). In some cases, the article may even become unacceptable before the adjective (e.g. **il-katastrofi l-mondjali* 'the global catastrophe').

The lack of a definite article usually indicates indefiniteness, but occasionally *wieħed* (fem. *waħda*) 'one' or *ċertu* (with fem. *-a*, pl. *-i*, but also treated as invariable by many speakers) 'certain' may also precede the indefinite noun (*ħajjat*, *wieħed ħajjat*, *ċertu ħajjat* 'a (certain) tailor').

2.2.3.2 The genitive

In modern Maltese, the use of the construct state is largely limited to a closed list of nouns indicating inalienables (parts of the body, relatives, some personal belongings), as well as to a limited number of words that automatically trigger this structure, such as *nofs* 'middle', *ras* 'head, beginning', *tarf* 'last part', *tmiem* 'end', *bieb* 'door', *ktieb* 'book', *ġenb* 'side' (e.g. *ma' ġenb il-bieb* 'by the side of the door', *f'nofs it-triq* 'in the middle of the road').

The more productive formation today is the analytic structure with *ta'* + noun, in which *ta'* (*tâgħ-*) 'of' is the Maltese reflex of North African *mitāˁ*, *mtāˁ*, or *ntāˁ* , Egyptian *bitāˁ* (< Old Arabic *matāˁ* 'possession'). Both *mitāˁ* and *bitāˁ* are attested in old notarial deeds of the 15th century from Malta.

2.2.3.3 Negation

Positive statements are made negative, as in other Arabic vernaculars, by circumfixing *ma.....x* around the finite verb (e.g. *ma rebaħx* 'he did not win', *ma pparkjatx* 'she did not park').

Negative commands, however, just suffix.... *x* to the imperfect 2nd person forms (*tirkibx*, pl. *tirkbux* 'do not ride!'). In a stronger type of command, *la* (< Arabic *lā*) is placed before the imperative (e.g. *la tirkibx, la tirkbux*).

A small class of words originally having a strongly positive meaning (namely, *ħadd* originally 'someone', *qatt* 'ever', *mkien* 'somewhere, [lit.] a place', *xejn* 'something', and *ebda* '(any) at all') have developed an alternative meaning which is definitely negative (hence, *ħadd* 'someone' and 'no one'; *qatt* 'ever' and 'never'; *mkien* 'somewhere, [lit.] a place' and 'nowhere'; *xejn* 'something' and 'nothing'; *ebda* '(any) at all' and '(none) at all'), as a result of their association with negative statements, in which they were introduced for emphasis. They still retain a positive meaning in some common utterances (e.g. *qatt mort it-Tajlandja?* 'have you ever been to Thailand?'; *taf lil xi ħadd jismu Fidiel?* 'do you know anyone called Fidiel?').

2.2.4 Nouns

Among so-called primitive or underived nouns, the collective category has not only survived in Maltese but, within the semantic limits of genus or type, it is still very productive and has roped in many nouns from Romance and some from English, which produce derivatives in the normal way (e.g. Romance Maltese coll. *sardin* 'sardines' > *sardina* 'a sardine', *palm* 'palm trees' > *palma* 'a palm tree', *injam* 'wood' > *injama* 'a piece of wood', *sandli* 'a pair of sandals' > *sandlija* 'a sandal'; English Maltese *tofi* 'toffee' > *tofija* 'a sweet', *buz* '[pair of] boots' > *buza* 'a [loose] boot', *buziet* '[loose] boots', *bwiez* '[pairs of] boots'.

2.2.4.1 Nominal derivation

Much of the derivative machinery of Arabic Maltese is now practically unproductive and its use is limited to Arabic Maltese and early Arabicized loans (cf. Mifsud 1995a).

2.2.4.1.1 The noun of agent (form 1v22*ā*3) is common with Arabic and Arabicized root-bases and often assumes the role of noun of profession. As such, it is largely bound to and limited by a relatively small class of verbs indicating rudimentary crafts and handiwork (e.g. *bajjad* 'painter', *ħaddied* 'blacksmith', *reffiegħ* 'coffin bearer', *kaħħal* 'plasterer'). Many nouns with suffixes *-ar, -atur, -ant, -ist(a)* indicating sophisticated jobs were imported from Italo-Romance (e.g. *tapizzar* 'upholsterer', *arġentier* 'silversmith', *induratur* 'gilder', *awtur* 'author', *kantant* 'singer', *bandist(a)* 'musician, bandsman'). The same thing is happening today with English nouns ending in *-er* (e.g. *plamer* 'plumber', *welder* 'welder', *maniġer* 'manager') and others without a specific marker (e.g. *mekkanik* 'mechanic', *ners* 'nurse').

2.2.4.1.2 The mimated formations for the nouns of place and instrument, partly merged in Maltese, today are both practically unproductive. The formation for the noun of place survives in several place-names harking back at least to medieval times (e.g. *marsa* 'harbor', *miġra* 'watercourse', *mġarr* 'watercourse', *msida* 'fishing place' (cf. Wettinger 1983). Beyond that, it is limited to a few nouns of Arabic origin (e.g. *masġar* 'wood', *maqdes* 'temple', *moħba* 'hiding place', *mħadda* 'pillow'). A few learned neologisms from Arabic roots are Vassalli's *mylsen* 'grammar', Serracino Inglott's *miklem* 'lexicon', Psaila's *miftel* 'tepidarium', as well as the recently coined *mitjar* for 'airport'. Many Maltese nouns of place of Romance origin today end with the suffix *-erija* (e.g. *librerija* 'library', *spiżerija* 'pharmacy', *birrerija* 'brewery', *biljetterija* 'ticket office', and the recent coinage *fenkerija* 'rabbitry').

2.2.4.1.3 The two formations of the noun of instrument (types *mi*12*a*3 and *mi*12*ā*3) are found only with a number of names of traditional tools (e.g. *minfaħ* 'bellows', *magħsar* 'press', *mansab* 'bird trap', *mitħna* 'grinding mill', *mgħażqa* 'hoe', *muftieħ* 'key', *moqdief* 'oar', *mehrież* 'mortar', *moħriet* 'plow', *miżien* 'scales'), some of which seem to be heading toward extinction. The names of more sophisticated tools are of Romance origin (e.g. *ċana* 'plane', *martell* 'hammer', *furmatur* 'chisel'). The names for home appliances and more sophisticated technological or electronic equipment are almost exclusively derived from English, most of them ending in the suffix *-er* (e.g. *towster* 'toaster', *stejpler* 'stapler', *driler* 'electric drill', *ħuver* 'vacuum cleaner', *kompjuter* 'computer').

2.2.4.1.4 The noun of quality (types 12*ū*3*a* and 12v3*īja*) and the diminutive formations (12v*jj*v3 etc.) are all root-based formations limited to closed sets of words of Arabic origin (e.g. *qrusa* 'sourness', *bluha* 'foolishness'; *ħruxija* 'cruelty', *qlubija* 'courage'; *ktejjeb* 'booklet', *dwejra* 'small house'). Sporadic diminutive formations on nouns of Romance origin seem to be very old and now are rarely used; they include the affectionate form of some personal names (e.g. *ċrejjev* 'small deer', *kmajra* 'small room', *mnejka* 'small door-knocker'; *Stejfen* 'little Stephen', *Trejża* 'little Theresa'). Romance suffixes which have taken over the diminutive function in Maltese are -*īn(a)* (e.g. *festin* 'small party', *gallettina* 'small biscuit') and *-ett(a)* (e.g. *fjurett* 'small bunch of flowers', *kaxxetta* 'small box', *banketta* 'stool').

2.2.4.1.5 The relative adjective (→ *nisba*), formed by the suffix *-i* (Arabic *-ī*), has fared better. It is used to form adjectives from many nouns, including some canonical forms like the noun of agent 1v22*ā*3 (e.g. *raħli* 'related to or coming from the village', *belti* 'coming from the city', *dinji* 'worldly, international', *demmi* 'bloody', *kabbari* 'having a tendency to grow big', *għaġġieli* 'hasty', *qabbieżi* 'jerky'). Formerly, it was relatively successful with nouns/adjectives of Romance origin (e.g. *tali* 'certain', *qastni* 'chestnut color', *ġganti* 'gigantic', *bluni* 'bluish') and also with a few English adjectives (e.g. *buli* 'first rate' [< English *bully*], *tobi* 'tubby', *blondi* 'blonde', *ġinġri* 'redheaded'). The Arabic relative adjective formation, however, faces the strong competition of some very popular Romance adjectival suffixes such as *-iż*, *-uż*, *-iku*, and *-ali* (e.g. *Slimiż* 'from the town of Tas-Sliema', *gustuż* 'nice, charming', *demokratiku* 'democratic', *personali* 'personal').

2.2.4.2 Nominal inflection

2.2.4.2.1 Gender

The standard Arabic Maltese feminine marker *-a* (*kelb/kelba* 'dog/bitch') has converged with the Romance feminine suffix (*impjegat/a* 'employee'). Loan nouns from English are assigned a grammatical gender on the basis of diverse criteria (cf. Farrugia 2004) but are not marked for gender (e.g., *miter* 'meter, gauge' and *vann* 'van' are masculine, but *friżer* 'freezer' and *friġġ* 'fridge' are feminine).

2.2.4.2.2 Number

2.2.4.2.2.1 The dual

As in other dialects, the use of the dual, a suffix formation, is limited to names of measurements and to parts of the body (for the latter, → pseudodual). Hybrid formations (Romance Maltese + Arabic Maltese) include *spallejn* 'shoulders', *koxxtejn* 'thighs', *vjeġġejn* 'two trips', and *passejn* 'a short walk [lit. 'a couple of paces']'. Synchronically, the productivity of the dual is nil, and the analytic formation with *żewġ* + plural is used instead.

2.2.4.2.2.2 The broken plural

Broken (or internal) plural formations were very active both with Arabic material and with Romance loans. In Arabic Maltese, one plural pattern has sometimes been extended analogically at the expense of another (e.g. Arabic Maltese *fqar* 'poor', *ilħna* 'voices', *kotba* 'books', for Arabic *fuqarāʾ, ʾalḥān, kutub*). Borrowed nouns only rarely now receive a broken plural, and a suffix plural is being increasingly preferred to a broken plural whenever a competing pair occurs (e.g. *twapet* ~ *tapiti* 'carpets'; *laned* ~ *landi* 'tins'; *bnadar* ~ *bandieri* 'flags'; *tazez* ~ *tazzi* 'glasses'; *bolol* ~ *bolli* 'stamps').

2.2.4.2.2.3 The sound plural

The sound (or suffixed) plurals, on the contrary, have withstood the test of time much better. The suffix *-in*, as in other Arabic dialects, is largely limited to the plural of Arabic participles and *nisba* formations (e.g. *ħerġin* 'going out', *mħaddmin* 'employed', *Għawdxin* 'Gozitans, from the island of Gozo').

The Arabic collective plural with suffix *-a* is still productive, both as a variant of *-in* and with nouns of profession of Arabic or Romance origin, thanks to an accidental *appui* form in Sicilian (cf. Galante 1969:64; e.g. Arabic Maltese *kittieba* 'writers', *kelliema* 'speakers', *qaddiefa* 'rowers', *telliefa* 'losers', *għalliema* 'teachers'; Romance Maltese *furnara* 'bakers', *arluġġara* 'watch repairers', *nutara* 'notaries', *xufiera* 'drivers', *infermiera* 'nurses'). The names of modern jobs, however, are being introduced from English together with their plural form (e.g. *fiters* 'fitters', *elektrixins* 'electricians', *plamers* 'plumbers').

The suffix *-ijiet* (presumably derived from the plural of verbal nouns with a weak final radical, like *tiġrijiet* 'races', *tiswijiet* 'repairs') is very productive, not only with Arabic nouns (e.g. *ommijiet* 'mothers', *artijiet* 'lands', *isqfijiet* 'bishops'), but also with innumerable Romance loans (e.g. *bżonnijiet* 'needs', *sptarijiet* 'hospitals', *prinċpijiet* 'princes', *patrijiet* 'friars', *postijiet* 'places'). It is also applied to many English loans, where it is often in competition with the English plural *-s* (e.g. *kejkijiet* 'cakes', *tankijiet* 'tanks', *basktijiet* 'baskets', *garaxxijiet* 'garages'; *gowlijiet* ~ *gowls* 'goals', *ċekkijiet* ~ *ċekks* 'checks', *brejkijiet* ~ *brejks* 'brakes').

Two suffixes of European origin, Romance *-i* and English *-s*, are very productive. The former is limited to Romance Maltese nouns and adjectives (e.g. *kartieri* 'wallets', *anġli* 'angels'; *ċari* 'clear', *goffi* 'bulky') and sporadically to a handful of Arabic Maltese and English Maltese nouns (e.g. Arabic Maltese *saffi* 'layers', *wegħdi* 'promises', *nofs-ta'-nhari* 'half days'; *sammi* 'very hard'; English Maltese *jardi* 'yards', *frilli* 'frills'; *raffi* 'rough'). The English plural, although limited to loan nouns from English, is becoming more important as the number of such nouns, most of which take only this plural, is increasing rapidly (*plejers* 'players', *tajers* 'tires', *films* 'films').

2.2.5 Numerals

The Maltese numerical system (see Table 8) is largely derived from the Arabic system, with some influence from Italian and English. Some of the numbers, in fact, have correlates in Italian and English, which are used in particular sociolinguistic situations and sometimes in idiomatic expressions (e.g. *wieħed/unu* 'one'; *issa ġejna ottu!* lit. 'now we have taken the shape of an eight!', meaning 'we're in a fix'). English numbers are used extensively by speakers of Maltese in any mathematical context and generally when speaking out series of numbers,

Table 8. Numerals

	Cardinal			Ordinal	
	Alone	+Nominal		Alone	+Nominal
		+ C- or V-	+ CC-		
1	*wieħed* /fem. *waħda* 'one'	*art (waħda)* '(one) land'	*kelb (wieħed)* '(one) dog'	*l-ewwel* 'the first'	*l-ewwel (kelb)* 'the first dog'
2	*tnejn*	*żewġ (artijiet)*	*żewġt (iklieb)*	*it-tieni*	*it-tieni*
3	*tlieta*	*tliet*	*tlitt/tlett*	*it-tielet*	*it-tielet*
4	*erbgħa*	*erba'*	*erbat*	*ir-raba'*	*ir-raba'*
5	*ħamsa*	*ħames*	*ħamest*	*il-ħames*	*il-ħames*
6	*sitta*	*sitt*	*sitt*	*is-sitta*	*is-sitt*
7	*sebgħa*	*seba'*	*sebat*	*is-seba'*	*is-seba'*
8	*tmienja*	*tmien*	*tmint*	*it-tmienja*	*it-tmien*
9	*disgħa*	*disa'*	*disat*	*id-disa'*	*id-disa'*
10	*għaxra*	*għaxar*	*għaxart*	*l-għaxra*	*l-għaxar*
11	*ħdax*	*ħdax-il (art*, *kelb*)		*il-ħdax*	*il-ħdax-il*
12	*tnax*	*tnax-il*		*it-tnax*	*it-tnax-il*
13	*tlettax*	*tlettax-il*		*it-tlettax*	*it-tlettax-il*
14	*erbatax*	*erbatax-il*		*l-erbatax*	*l-erbatax-il*
15	*ħmistax*	*ħmistax-il*		*il-ħmistax*	*il-ħmistax-il*
16	*sittax*	*sittax-il*		*is-sittax*	*is-sittax-il*
17	*sbatax*	*sbatax-il*		*is-sbatax*	*is-sbatax-il*
18	*tmintax*	*tmintax-il*		*it-tmintax*	*it-tmintax-il*
19	*dsatax*	*dsatax-il*		*id-dsatax*	*id-dsatax-il*
20	*għoxrin*	*għoxrin*		*l-għoxrin*	*l-għoxrin*
50	*ħamsin*	*ħamsin*		*il-ħamsin*	*il-ħamsin*
100	*mija*	*mitt*		*il-mija*	*il-mitt*

such as telephone numbers. A likely reason for this would be the almost exclusive use of English in education.

Cardinal numbers have two forms, depending on whether they stand alone or precede a nominal. Of the independent forms, only *wieħed* 'one' has a feminine form *waħda*.

When cardinal numbers are followed by a nominal, they add a final *-t* (a residue from the Old Arabic numeral system) when followed by a plural noun beginning with two consonants (*ħamest ibliet* 'five cities', *tlett ikmamar* 'three rooms), but there are exceptions (e.g. *tliet snin* 'three years'). In such cases, an epenthetic vowel is normally added to the noun to break the consonant cluster. Note also the use of *żewġ/t* for 'two' followed by a nominal.

Numerals from 11 to 19 have two forms, a shorter independent one ending in *-ax* (e.g. *ħmistax* 'fifteen') and a longer one when followed by a noun (e.g. *ħmistax-il tarġa* 'fifteen steps'); *ax-il* is clearly the reflex of Arabic *ʿašar* (Maltese *għaxar*).

'Hundred' is *mija,* and *mitt* when followed by a noun. The conjunction *u* 'and' is placed between the units and the tens, and between the hundreds and the tens (e.g. *mija u erbgħa u sittin* 'one hundred and sixty-four').

2.2.6 The verb

Maltese has generally preserved the basic Arabic structure consisting of two indicative tenses (perfect and imperfect) and an imperative. These two tenses, together with other nonfinite verbal forms, such as the active participle, also support an aspectual system distinguishing between three different aspects of the action: completeness, progressivity, and habituality.

2.2.6.1 Verbal derivation

The vitality of the verbal Forms is reduced greatly, both formally and semantically (see

Table 9. Verbal Forms

Form		Example
I	1v2v3	*fired* 'to separate'
II	1v22v3	*fisser* 'to explain'
III	1*ā*2v3	*bierek* 'to bless'
V	*t*1v22v3	*tfisser* 'to be explained'
VI	*t*1*ā*2v3	*tbierek* 'to be blessed'
VIIa	*n*1v2v3	*nfired* 'to be separated'
b	*nt*1v2v3 (1=resonant)	*ntrebaħ* 'to be won'
c	*n*1*t*v2v3 (1=sibilant)	*nsteraq* 'to be stolen'
VIII	1*t*v2v3	*ntesa* 'to be forgotten'
IX	12*ā*3	*twal* 'to become long'
X a	*st*v12v3	*stagħġeb* 'to be surprised'
b	*st*1v22v3	*stħajjel* 'to imagine'
QI	1v23v4	*ħarbat* 'to destroy'
QII	*t*1v23v4	*tħarbat* 'to be destroyed'

Table 9). Four Forms (I, II, V, and VII) are represented by a large number of verbs, five (III, VI, VIII, IX, and X) are of very limited use, and one (IV) is defunct. At the semantic level, the rich array of meanings associated with the Arabic verbal Forms is practically reduced to a simple semantic schema in which Forms I, II, III, and quadriliteral I are active, corresponding to passive/reflexive Forms VII, V, VI, and quadriliteral II, respectively. Form IX stands alone as a small class of verbs with a clear definition: the inchoative.

The system of verbal Forms is only operative with verbs of Arabic origin (e.g. *ċekken* 'to reduce', *daqqas* 'to make proportional in size', *xejjen* 'to annihilate', *labbar* 'to fasten with pins'), and with some fully integrated Romance loans, probably dating from a time when Maltese, severed from the other dialects, was developing new words from its own resources.

Maltese dictionaries (e.g. Aquilina 1987–1990) record 373 fully integrated loan verbs which hark back to 247 different roots. Of these, 217 verbs (from 132 different roots) are still in use. These verbs, derived through the extraction of three or four root consonants from the stem, generally partake fully in Arabic Maltese morphology (cf. Mifsud 1995b:272–295).

In spite of this apparent productivity, the process of full integration by the extraction of roots has practically stopped today, except for sporadic coinages recorded in literature (e.g. *pallam* 'to plant with palms', *qanpen* 'to ring'). The fact that in the last two centuries there has been only one fully integrated verb from English (*fajjar* 'to hurl' < English *fire*) is an indication of the failing productivity of the system in modern Maltese. Besides, several of these loans are in competition with stem-based cognates, the latter being more often preferred. Some of the former are now obsolete (e.g. obsolete *mannas* ~ *mmansa* 'to tame', *barraġ* ~ *mborġja* 'to heap up', *rambel* ~ *rrombla* 'to roll', *senneg* ~ *ssingja* 'to draw lines'). Modern Maltese increases its verbal repertoire exclusively through loan stems coming mainly from English and preserving their consonantal and syllabic form practically intact (e.g. *startja* 'to start [a machine]', *pparkja* 'to park', *ffrejmja* 'to frame a person', *ddawnlowdja* 'to download').

The active participle (1*ā*2v3, fem. 1*ā*23*a*, pl. 1v23*īn*) is found almost exclusively with a group of verbs of Form I, mainly intransitive verbs indicating a state or a movement. When denoting a state, it tends to have an adjectival function (e.g. *biered* 'tepid', *ħieles* 'free', *ħiemed* 'silent'). It retains its original verbal sense only in verbs of motion, where it conveys the aspectual meaning of progressivity (e.g. *nieżel* 'going down', *ħiereġ* 'going out', *dieħel* 'entering').

The passive participle is generally found with all verbal Forms, but it is unusual with Forms that normally have a passive meaning (e.g. I*mv*12*ū*3,II*m*1v22v3,III*m*1*ā*2v3,V*mvt*1v22v3). Feminine forms add *-a* and the plural *-īn* or *-a*. The passive participle often acquires an adjectival sense (e.g. *magħlub* 'thin', *maħmuġ* 'dirty'). Alternatively, it occurs frequently as the last element of the passive construction after finite forms of the verbs *kien* 'to be' and *ġie* 'to come' (e.g. *kien imġiegħel jiġri* 'he was forced to run', *ġew maqtulin* 'they were killed').

Within the Arabic Maltese lexicon, verbal nouns display great vitality. The active Forms share the same verbal noun with their passive counterpart (e.g., III *bierek* 'to bless' and VI *tbierek* 'to be blessed' make use of the same verbal noun *tberik*).

Form I has at least six different patterns (1v23, 1v2v3, 12v̄3, 12*ā*3*a*, 1v23*ān*, *mv*123*a*).

Apart from the old *t*v12*ī*3 type, Form II has developed a new type *t*1v22*ī*3, which seems to be spreading at the expense of the former. The verbal nouns of all the other Forms normally keep the stem of the Form but carry a characteristic long vowel *-ī-* in their last syllable; some also have a *t-* prefix (e.g. III *bierek*: verbal noun *tberik,* X *stagħġeb*: verbal noun *stagħġib*).

The verbal noun and the passive participle are both very active with Arabic Maltese verbs, where they often form the basis for potential lexemes, i.e. word formations that may be created spontaneously by a productive process. These may eventually become established, especially if they acquire a specific technical sense (e.g. *titqiba* 'medical injection', *(ġie fis-) seħħ* '(to come in) force', *tisliba* 'crossword puzzle', *tisbiħ* 'embellishment', *xandir* 'broadcasting').

Since they are both closely bound by the root-system, both formations are only operative with fully naturalized loans (e.g. Romance Maltese *tbaxxija* 'act of lowering', *mbaxxi* 'lowered').

2.2.6.2 Verbal inflection

Like most sedentary dialects, Maltese has lost the gender distinction in the 2nd and 3rd person plural and, together with the Western dialects, also in the 2nd person singular. The result is a paradigm with just seven forms (Table 10).

Table 10. Conjugational affixes of the verb

	singular	plural
Indicative Perfect		
3rd masc.	*–Ø*	*–u*
3rd fem.	*–et*	
2nd	*–t*	*–tu*
1st	*–t*	*–na*
Indicative Imperfect		
3rd masc.	*j–*	*j–u*
3rd fem.	*t–*	
2nd	*t–*	*t–u*
1st	*n–*	*n–u*
Imperative		
2nd	v–	v–*u*

Verbal inflection is still fully productive and has been extended to all loan verbs. This was achieved through a paradigm made up of prefixes and a set of suffixes derived from that of weak-final verbs and consisting of the final weak radical *-j-* plus the normal suffixes (hence, *-ajt, -a, -āt; -ajna, -ajtu, -aw*). These affixes are applied to the imported stem, which is left unchanged. The following could be possible reasons for the preference of the weak-final paradigm to that of sound verbs: (i) these suffixes are all vowel-released, like those of the Romance paradigm; (ii) since they do not bring about allomorphic changes other than stress shift, they guarantee the formal integrity of the stem and hence its immediate recognizability; (iii) the weak-final conjugation has long been in the process of encroaching on other classes of Arabic verbs (doubled verbs, *hamza*-final verbs, verbs with final ʿ or *ġ*, and other defective or anomalous verbs; cf. Mifsud 1995b:296–318). As a result, the weak-final paradigm, originally a weak and marginal variant, has become the most important channel of integration for loan verbs, indeed its only productive category.

The new Maltese verbal paradigm has even adopted suffixes of Romance origin to complement those of the weak-final suffixes wherever Arabic Maltese had broken forms and could not offer a concatenative solution (e.g. passive participle *-āt, -ūt, -īt* instead of *m*v12*ū*3, *m*1v22v3, etc. and verbal noun *-ār* instead of 12*ī*3, *t*v12*ī*3, etc.).

In the process of integration, Romance verbs of the *-ere* and *-ire* conjugations joined the conjugation of weak verbs of the *jimxi/jimxu* type, while the larger group of Romance verbs of the *-are* conjugation fitted well with weak verbs of the *jaqra/jaqraw* type (e.g. Italian *pingere* > *jpinġ-i/jpinġ-u* 'he paints, they paint', like *jimx-i/jimx-u*, and Italian *protest-are* > *jipprotest-a/jipprotest-aw* 'he protests, they protest', like *jaqr-a/jaqr-aw*).

Table 11 gives examples of typical loan verbs integrated to different degrees, together with their Arabic Maltese models.

2.3 *Notes on syntax*

2.3.1 Maltese is basically an SVO language. The neutral order of the constituents of a sentence with a transitive verb is Subject – Verb – Direct object – Indirect object. Adverbial expressions follow in this order: Manner – Place – Time (cf. Alb. Borg and Azzopardi-Alexander 1997:57–58, e.g. *it-tabib kiteb ittra lil ħuh bil-għaġla fl-ispiżerija dalgħodu* 'the doctor wrote

Table 11. Conjugation of loan verbs in Maltese

	Arabic Maltese verbs with weak final		Loan verbs (< Italian and English) integrated in Arabic Maltese to different degrees			
Verb	*'ħalla* 'to leave'	*'qara* 'to read'	*'falla* 'to fail'	*'kanta* 'to sing'	*ppre'tenda* 'to pretend'	*'pparkja* 'to park'
Root	*ḥ-l-y*	*q-r-y*	*f-l-y*	*k-n-t-y*	–	–
Stem	–	–	*fall-*	*kant-*	*ppretend-*	*pparkj-*
Perfect						
3rd sg. masc.	*'ħall-a*	*qr-a*	*'fall-a*	*'kant-a*	*ppre'tend-a*	*'pparkj-a*
3rd sg. fem.	*ħall-'iet*	*qr-at*	*fall-'iet*	*kant-'at*	*ppretend-'iet*	*pparkj-'at*
2nd sg.	*ħall-'ejt*	*qr-ajt*	*fall-'ejt*	*kant-'ajt*	*ppretend-'ejt*	*pparkj-'ajt*
1st sg.	*ħall-'ejt*	*qr-ajt*	*fall-'ejt*	*kant-'ajt*	*ppretend-'ejt*	*pparkj-'ajt*
3rd pl.	*ħall-'ew*	*qr-'aw*	*fall-'ew*	*kant-'aw*	*ppretend-'ew*	*pparkj-'aw*
2nd pl.	*ħall-'ejtu*	*qr-'ajtu*	*fall-'ejtu*	*kant-'ajtu*	*ppretend-'ejtu*	*pparkj-'ajtu*
1st pl.	*ħall-'ejna*	*qr-'ajna*	*fall-'ejna*	*kant-'ajma*	*ppretend-'ejna*	*pparkj-'ajna*
Imperfect						
3rd sg. masc.	*j-'ħall-i*	*j-'aqr-a*	*j-'fall-i*	*j-'kant-a*	*ji-ppre'tend-i*	*ji-'pparkj-a*
3rd sg. fem.	*t-'ħall-i*	*t-'aqr-a*	*t-'fall-i*	*t-'kant-a*	*ti-ppre'tend-i*	*ti-'pparkj-a*
2nd sg.	*t-'ħall-i*	*t-'aqr-a*	*t-'fall-i*	*t-'kant-a*	*ti-ppre'tend-i*	*ti-'pparkj-a*
1st sg.	*n-'ħall-i*	*n-'aqr-a*	*n-'fall-i*	*n-'kant-a*	*ni-ppre'tend-i*	*ni-'pparkj-a*
3rd pl.	*j-'ħall-u*	*j-aqr-'aw*	*j-'fall-u*	*j-kant-'aw*	*ji-ppre'tend-u*	*ji-pparkj-'aw*
2nd pl.	*t-'ħall-u*	*t-aqr-'aw*	*t-'fall-u*	*t-kant-'aw*	*ti-ppre'tend-u*	*ti-pparkj-'aw*
1st pl.	*n-'ħall-u*	*n-aqr-'aw*	*n-'fall-u*	*n-kant-'aw*	*ni-ppre'tend-u*	*ni-pparkj-'aw*
Imperative						
2nd sg.	*'ħall-i*	*'aqr-a*	*'fall-i*	*'kant-a*	*ppre'tend-i*	*'pparkj-a*
2nd pl.	*'ħall-u*	*aqr-'aw*	*'fall-u*	*kant-'aw*	*ppre'tend-u*	*pparkj-'aw*
Passive participle						
	'mħoll-i	*'moqri*	*'mfall-i*			
			fall-'ut	*kant-'at*	*ppretend-'ut*	*pparkj-'at*
Verbal noun						
	tħo'llija	*'qari*	*tfa'llija*			
			fall-i'ment	*'kant*	*pretens-'joni*	*pparkj-'ar*

a letter to his brother in a hurry in the pharmacy this morning'.

VSO order also occurs, especially in subordinate clauses where VSO may be preferable in certain cases, e.g. *daħlu sħabu u sabuh minxur fl-art* 'his colleagues entered and found him spread out on the floor'; *it-tfal siktu f'daqqa meta daħal is-surmast* 'the children suddenly became silent when the headmaster entered'.

2.3.2 Unlike most Arabic vernaculars, Maltese has a specific object marker *lil*. This particle is traditionally associated with the preposition *lil* 'to' and in fact is subject to the same morphophonemic variation. It is obligatory only before personal names and stressed personal pronouns (cf. Alb. Borg and Comrie 1984), e.g. *is-seftura rat lil Fredu* 'the maid saw Alfred'; *Pawlu ħafer lil ħuh* 'Paul forgave his brother'; *is-surmast għażel lilna* 'the headmaster chose us'.

The use of the object marker is governed by a hierarchy of animacy based on the following criteria: expressions higher in the scale of animacy (such as personal names) regularly take an object marker, while expressions at the lower end, such as abstract nouns, or those referring to inanimates, regularly do not. Moreover, the animacy hierarchy is complemented by that of definiteness, so that definite common nouns referring to humans are more prone to be marked (e.g. *missierek sab lit-tifel* or *it-tifel* 'your father found the boy'; *missierek sab il-kelb* or, less frequently, *lill-kelb* 'your father found the dog') than those referring to indefinite animate ones or nonhumans (e.g. *il-karozza laqtet ġuvni* 'the car hit a young man', *is-seftura ferrgħet il-kafè* 'the maid poured the coffee'; *il-folla semgħet l-aħbar* 'the crowd heard the news').

Particle object marking seems to be totally absent in Western Arabic, and in Eastern Arabic it tends to be rare and occurs predominantly in conjunction with head marking. It is, however, attested in a vast area encompassed by Romance, from Portuguese and Spanish, through popular French to Italian and Sicilian, where it involves the use of the preposition *a* 'to' and is associated with animacy and definiteness (cf. Alb. Borg and Mifsud 2002). It remains to be decided whether the Maltese object marker *lil* represents a case of retention of an Aramaic substratal feature (cf. Sutcliffe 1936:171–172; Aquilina 1961:55; Alex. Borg 1994:41–67), or the result of adstratal Romance influence.

Bibliographical references

Primary sources

Agius De Soldanis, Giovan Pietro Francesco. *Della Lingua Punica*. Rome, 1750.
Vassalli, Michele Antonio. *Mylsen Phoenico-Punicum sive Grammatica Melitensis*. Rome, 1791.
——. *Lexicon Melitense-Latino-Italum (Ktŷb yl Klŷm Mâlti)*. Rome, 1796.
Wzzino, Dun Franzisk. *Taġhlim Nisrani*. Rome, 1752.

Secondary sources

Agius, Dionysius. 1996. *Siculo Arabic*. London and New York: Kegan Paul International.
Aquilina, Joseph. 1961. *Papers in Maltese linguistics*. Valletta: Royal University of Malta.
——. 1979. *Maltese-Arabic comparative grammar*. Msida: University of Malta.
——. 1987–1990. *Maltese-English dictionary*. 2 vols. Malta: Midsea Books.
——, and Benedict S.J. Isserlin. 1981. *A survey of contemporary dialectal Maltese*. I. *Gozo*. Leeds.
Borg, Albert and Marie Azzopardi-Alexander. 1997. *Maltese*. London: Routledge.
——, and Bernard Comrie. 1984. "Object diffuseness in Maltese". *Objects: Towards a theory of grammatical relations*, ed. Frans Plank, 109–126. London and Orlando: Academic Press.
——, and Manwel Mifsud. 2002. "Maltese object marking in a Mediterranean context". *Mediterranean Languages: Papers from the MedTyp Workshop, Tirrenia, 2000*, ed. Paolo Ramat, 33–46. Bochum: Brockmeyer.
Borg, Alexander. 1978. *A historical and comparative phonology and morphology of Maltese*. Ph.D. diss., The Hebrew University, Jerusalem.
——. 1994. "Some evolutionary parallels and divergences in Cypriot Arabic and Maltese". *Mediterranean Language Review* 8.41–67.
——. 1997. "On some Levantine linguistic traits in Maltese". *Studies in modern Semitic languages*, ed. Shlomo Izre'el and Shlomo Raz, 133–152. Leiden: E.J. Brill.
Cantineau, Jean. 1960. "Cours de phonétique arabe". Jean Cantineau, *Études de linguistique arabe*, 3–126. Paris: Klincksieck.
Farrugia, G. 2004. *Il-ġens grammatikali li jieħdu fil-Malti n-nomi Ingliżi*. Malta: Mireva.
Friggieri, Albert and Thomas Freller. 1998. *Malta: The bulwark of Europe*. Malta: Gutenberg Press.
Galante, P. 1969. *Grammatica storica della lingua siciliana*. Castellammare del Golfo.
Kaye, Alan S. and Judith Rosenhouse. 1997. "Arabic dialects and Maltese". *The Semitic languages*, ed. Robert Hetzron, 263–311. London and New York: Routledge.
Mifsud, Manwel. 1995a. "The productivity of Arabic in Maltese". *Proceedings of the 2nd International Conference of AIDA*, ed. Clive Holes, a.o. 141–150. Cambridge: University of Cambridge.
——. 1995b. *Loan verbs in Maltese: A descriptive and comparative study*. Leiden: E.J. Brill.
Stumme, Hans. 1904. *Maltesische Studien: Eine Sammlung prosaischer und poetischer Texte in maltesischer Sprache nebst Erläuterungen*. Leipzig: Hinrichs.
Sutcliffe, Edmund F. 1936. *A grammar of the Maltese language*. London: Oxford University Press.
Vanhove, Martine. 1993. *La langue maltaise: Études syntaxiques d'un dialecte arabe "périphérique"*. Wiesbaden: O. Harrassowitz.
Wettinger, Godfrey. 1983. "Some grammatical characteristics of the place-names of Malta and Gozo in early modern times". *Journal of Maltese Studies* 15.31–68.

Manwel Mifsud
(Kunsill Nazzjonali ta' l-Ilsien Malti)

Maʿnā

1. Introduction

In a modern Arabic-English dictionary, the term *maʿnā* is rendered by such words as 'sense', 'meaning', and 'signification', so that it forms with its conceptual counterpart → *lafḏ̣* a semantic pair quite akin to the signifier/signified couple familiar to contemporary linguistics. Yet, a close examination of a number of Arabic texts of different periods in which both these terms show up, separately or together, reveals that it has not always been so, and that it is only in the final stage of a long evolution that the *lafḏ/maʿnā* couple ended up with its present-day functional load. This evolution was not uniform but rather produced differing results in the various fields in which the two terms were widely used.

Consequently, it seems necessary, in order to properly understand how the term *maʿnā* was used in different domains and at different times, to start from its original conceptual content. This gives us an idea of how it was comprehended by speakers prior to any specialized use. The first section below shows that in a primitive stage of their use, the relationship between *lafḏ* and *maʿnā* was essentially conceived of as 'indeterminate', a given *lafḏ* being able to refer to different *maʿānī* (pl. of *maʿnā*), and vice versa. The section describes how the efforts of scholars who aimed to reduce this 'indetermination' of the *lafḏ/maʿnā* relationship led, at least in some of the disciplines having language as their object of study, to a one-to-one relation between the two words. Grammar (*naḥw*) is not included in this disciplinary excursus as it cannot really be regarded as a discipline in which semantic considerations play a fundamental role; grammarians themselves considered that grammar was essentially a 'formal art' (*ṣināʿa lafḏiyya*), aimed specifically at ensuring the quality of linguistic expression and in particular its correctness regarding case assignment (→ *lafḏ*). The third section describes the specific way in which Arabic grammar, in the course of the conflict which opposed the major Islamic science of grammar to logic, dealt with the *lafḏ/maʿnā* articulation, when this 'profane' discipline was introduced in the Islamic cultural field (for a more complete view, see Versteegh 1997).

2. THE ORIGINAL VIEW OF THINGS

The noun *maʿnā* is originally a verbal noun (*maṣdar*) formed on the verb *ʿanā/yaʿnī*, which properly means 'to aim at something'. *Maʿnā* consequently means, primarily, 'the act of aiming at something'. This primary sense is still perceived by Arab lexicographers. We find, for instance, in ʾAbū Hilāl al-ʿAskarī's (d. 395/1004) *Kitāb al-furūq fī l-luġa* the following statement: "The *maʿnā* is the intention which causes an utterance to be used in a certain way and not in another one. [But] in everyday use the *maʿnā* of an utterance may be that to which the intention refers". The last part of this definition clearly indicates how current usage passed from the original sense of 'act indicating a communicative intention' to that of 'content of this intention', and hence its 'meaning'. It is important, however, to bear in mind that the term *maʿnā* has, in many of its uses, retained something of its origin, referring more to the subjective goal of an act of communication than to the linguistic (or nonlinguistic) means used to reach this goal.

In its primary sense, *maʿnā* as a manifestation of a communicative intention has no privileged relationship with linguistic expression. The traditional view concerning this question is synthesized in a masterly way in the first work specifically devoted to this field of scholarship in Arab culture, Jāḥiẓ' (d. 256/869) *Book of expression and exposition* (*Kitāb al-bayān wa-t-tabyīn*). Jāḥiẓ stresses the relative independence of *maʿnā* from the means to convey it, listing five different means of expressing *maʿānī*: articulate language (*lafḏ*), indication (*ʾišāra*), numbers (*ʿuqad*), graphic representation (*xaṭṭ*), and situation (*ḥāl*; *Bayān* 55). It is true that the *Kitāb al-bayān wa-t-tabyīn* sets *lafḏ* in the lead of the means of expression, and that Jāḥiẓ continually extolls it under the category of eloquence (*bayān*). Yet, it is significant that he does not present it as the exclusive means of expressing *maʿānī*. In fact, and although *bayān* mostly means 'eloquence' in the *Kitāb al-bayān wa-t-tabyīn*, he insists that it may originally apply to all forms of expression, writing that "expression (*bayān*) is a term that encompasses everything, allowing you to disclose the *maʿnā* and lift the curtain on thought, so that the hearer may reach its essence and grasp its content, whichever it may be, no matter what this means of expression may be and whatever the nature of the sign [used]" (*Bayān* 54).

All this makes it possible to understand that, in the primordial Arab view of things, the term *lafḏ* globally referred to linguistic expression, i.e. both its signifier and signified side. It would consequently be a mistake to understand this word, in many of the ancient texts where this view of things was effective, as the mere formal aspect of linguistic expression. Much textual data supports this hypothesis. Jāḥiẓ, for example, in the general considerations on *lafḏ* and *maʿnā*, makes it quite clear that there exists between the two a difference of nature which precludes any possibility of setting them in a one-to-one relationship. He writes: "You must know, may God keep you, that the status of *maʿānī* is different from that of *ʾalfāḏ*, because *maʿānī* extend without limit and develop to the infinite, whereas the names that express them are limited, finite and fixed" (*Bayān* 55). This situation, in which *maʿānī* in infinite number can only be connected to *ʾalfāḏ* in finite number, i.e. in a one-to-many relationship (surjection), is very typical of the traditional Arab conception of the relations between *lafḏ* and *maʿnā*. It is, moreover, complicated by the fact that different *ʾalfāḏ* are considered to be able to refer to the same *maʿnā* (synonymy).

Another of Jāḥiẓ' important ideas on the relationship between *lafḏ* and *maʿnā* is presented in another book of this great thinker and writer, the *Book of animals* (*Kitāb al-ḥayawān*). Jāḥiẓ posits that *maʿānī* are universal and common to all human beings, whereas *ʾalfāḏ* as linguistic expressions are specific to a given language (*Ḥayawān* 132). From that position ensues the conclusion that all aesthetic judgments aiming at evaluating an author or a text can only and specifically bear on *lafḏ*. This is the real meaning of Jāḥiẓ' well-known saying that "*maʿānī* fill up the streets" (*al-maʿānī maṭrūḥa fī ṭ-ṭarīq*; *Ḥayawān* 131), which stresses the fact that a literary or poetic text must not be judged on its moral intentions or its content but rather on its qualities of formal expression, that is, on its *lafḏ*.

From all this we may conclude that in the traditional view, the relationship between *lafḏ* and *maʿnā* is basically "indeterminate", a given *lafḏ* being able to refer to different *maʿānī* and vice versa.

3. The confrontation with logic

When Greek logic came to be known in the Islamic world, around the 10th century C.E., it quickly exerted a deep fascination on scholars of all branches of knowledge, who saw in it the tool they needed to produce better founded and more systematic scientific works. At the same time, this foreign and profane discipline, with its potential claims to hegemony, aroused strong resistance from more traditional scholars. This was especially the case in the field of grammar, because logic pretends to have as one of its major fields of competence the correct use of language, a claim which could only be interpreted by grammarians as a threat to their specific fiefdom. In the course of the fierce debates which opposed the advocates of logic to those of traditional Arabic grammar, the idea emerged, among the moderate, that logic's domain was that of meaning (*maʿnā*), while grammar's domain was that of form (*lafḏ*). Thus, the famous polygrapher ʾAbū Ḥayyān at-Tawḥīdī (d. 400/1010), in his well-known book *al-Muqābasāt*, states, in a chapter dealing with the relationship of grammar and logic (Mahdi 1970), that "the gist of the logician's attention goes to meanings although he may not neglect forms which are, so to say, their containers and displayers, and the gist of the grammarian's attention goes to forms, although he may not neglect meanings which are, so to say, their content and essence" (*al-Muqābasāt* 22). However, this idea was not accepted by most grammarians, who considered that all fields of knowledge relative to the Arabic language fell within their competence. Logic's claims concerning its ability to deal with questions of meaning were generally swept away by showing the logician's failure to account for subtle semantic distinctions made in Arabic, such as the different values of the conjunctive particle *wa-* or those of topicalizing particles like *ʾinna* (→ grammatical tradition: history). It is important to point out that it is in the context of such discussions that the expression *maʿānī an-naḥw* 'grammatical semantics' made its first appearance, precisely to point to the subtle shades of meaning that specialized markers were able to express when aptly used by competent speakers of Arabic and which, grammarians claimed, completely eluded the gross tools of logical analysis. The detailed specification of the concept of *maʿānī an-naḥw* will be central in the 'science of meanings' (*ʿilm al-maʿānī*), whose foundations were laid by ʿAbd al-Qāhir al-Jurjānī in the 11th century C.E.

4. The debate around synonymy

Another field of research to which reference must be made in discussing the evolution of ideas on *lafḏ* and *maʿnā* is that of lexicography (→ lexicography: Classical Arabic). Arab philologists generally held the view that Arabic was the richest language in the world and that this was manifested, among other ways, by its richness in synonyms (→ *muštarak*; → *mutarādif*). Thus, Ibn Fāris (d. 395/1004) writes, in one of the standard works of Arabic philology:

> The Arabic language is the best of languages and the widest [...], for if we wanted to talk about the sabre and its characteristics in Persian, we could only do it with one word, whereas we have for the sabre in Arabic a multitude of qualifications, and the same goes for the lion, the horse, and still other things named by synonyms. How then could one draw a parallel between this and that? And whence could other languages have the abundance of Arabic? (*aṣ-Ṣāḥibī fī fiqh al-luġa* 44)

Comparable assertions are to be found in many other classical sources. The standard view that the abundance of synonyms is a sign of wealth in Arabic was not, however, held by all scholars. Some of them, admittedly a minority, maintained that synonyms, if they really existed, would not bear witness to the perfection of the language but, on the contrary, would indicate its lack of precision. ʾAbū Hilāl al-ʿAskarī (d. 395/1004), one of the most thorough proponents of this challenging view, writes, "If referring to a thing once makes it known, referring to it two or three times would be of no avail. Now the establisher of the language is too wise to have introduced in it non-informative elements" (*Kitāb al-furūq fī l-luġa* 13). ʾAbū Hilāl concludes that strict synonymy cannot exist in language. Only one word, he insists, directly refers to a given notion, and all supposed synonyms are only indirect designations, bringing secondary qualifications and connotations. In the introductory chapter of his *Book of differences* (cf. Kouloughli 1997), he proposes no fewer than eight different linguistic tests to demonstrate that there are no absolute synonyms in language, and the rest of the book is devoted to showing that the systematic use of these tests makes it possible to differentiate between words some people believe to be synonymous, like *ʿilm* 'science' and *maʿrifa* 'knowledge'.

The debate around synonymy went on for a long time, and it cannot be said that the more rigorous approach represented by ʾAbū Hilāl was victorious. The debate, however, played an undeniable role in the emergence of a clearer intuition of the functional relationship between *lafḏ* and *maʿnā*.

5. The theoretical elaboration of *lafḏ* and *maʿnā* in Arabic rhetoric

Theoretical debates on the relationship between *lafḏ* and *maʿnā* took place both in the context of the confrontation between grammar and logic and within the field of lexicography. It must be stressed, however, that it was neither in the field of Arabic grammar nor in that of lexicography that decisive advances were to be made toward the systematic study of the interrelations between form and meaning. Such advances actually emerged in a very different field of research, that of theological controversies, and more specifically the debate around the question of the exact nature of the *Qurʾān*'s inimitability (→ *ʾiʿjāz*). Briefly put (for more details, see Kouloughli 1983, 2002), two opposite orientations progressively emerged concerning the nature of *ʾiʿjāz*. The first one attributed it to the very nature of Qurʾānic themes: omniscience, wisdom, eschatological and moral content, etc., i.e. to the *maʿānī* conveyed by the revealed text. The second one, relying among other things on the many passages of the *Qurʾān* in which human beings are challenged to imitate the very expression of the divine text (for example Q. 2/23 or 10/38), considered *ʾiʿjāz* to be primarily a linguistic phenomenon and hence a matter of *lafḏ*. Furthermore, in the context of the latter orientation, two opposite approaches were proposed. Some considered that it was essentially in the rhetorical design of the text (metaphors, comparisons, and other figures) that *ʾiʿjāz* manifested itself, while for others, it was in the text as a whole, and more specifically in its 'textual organization' (*naḏm*) that the secret of *ʾiʿjāz* was to be sought.

The first significant breakthrough toward establishing that *ʾiʿjāz* was to be identified within the global textual structure of the Qurʾānic text was made by the Muʿtazilite ʾAbū Hāšim al-Jubbāʾī (d. 321/933). His ideas on the subject are known to us through the account which his disciple, the *qāḍī* ʿAbd al-Jabbār (d. 415/1024), gave of his thoughts in the bulky treatise known as the *Muġnī* and devoted to the defense of Muʿtazilite positions on questions of theology. The question of *ʾiʿjāz* seemed important enough to this thinker to dedicate an entire book (Vol. XVI) of his treatise to its elucidation. ʿAbd al-Jabbār writes the following concerning the teachings of al-Jubbāʾī: "Our master, ʾAbū Hāšim, said that discourse is eloquent only if it combines the elegance of expression with the beauty of content (*jazālatu lafḏ̣i-hi wa-ḥusnu maʿnā-hu*), the two being indispensable because a discourse whose expression is elegant but whose content is defective is not eloquent. [In order to be so] it must consequently associate both qualities" (*Muġnī* XVI, 197). At this point, he goes somewhat further than his master by showing that one cannot speak of eloquence concerning isolated words and that, consequently, one must take into account the modalities of composition (*ḍamm*) of words relative to one another. In this respect he writes: "You must know that eloquence cannot appear in isolated words but only in the composition of the utterance according to specified modalities (*bi-ḍ-ḍamm ʿalā ṭarīqatin maxṣūṣa*)" (*Muġnī* XVI, 199). He identifies three possible modalities for any composition: the choice of lexical units, their case marking, and their position relative to one another, making it clear that he thinks this analysis to be exhaustive, for he adds: "There is no fourth term after these three, for the word is either considered in itself, or regarding its case marking, or its position, and these considerations are necessary for each word, and the same goes for all other words when composed together [...]. Therefore, it is only according to these modalities that eloquence can appear and in no other way".

This methodological conclusion reached by the *qāḍī* ʿAbd al-Jabbār concerning the precise modalities to be taken into account in seeking the exact locus of eloquence in texts may be regarded as a decisive step towards the elaboration of operational procedures for text analysis, as it makes it possible to seek a connection between this emerging discipline and the more solidly established techniques of grammatical analysis. But although some passages of the *Muġnī* do suggest that the *qāḍī* perceived these logical repercussions of his analysis perfectly, the credit for actually elaborating its methodological conclusions and founding an effective technique for the semantic analysis of texts goes to the great ʾAšʿarite grammarian ʿAbd al-Qāhir al-Jurjānī (d. 471/1078). In his seminal book *Dalāʾil al-ʾiʿjāz*, in which he lays down the theoretical and methodological foundations of the new discipline, he blames his Muʿtazilite predecessors for having remained too vague when speaking of utterances as resulting from "composition according to specified modalities". He insists that, concretely, these modalities are nothing but the set of grammatical categories which the language puts at the disposal of all its users to express what they need to say but which are not used with the same felicity by all. Thus, what appeared as a mystery concerning the nature of *ʾiʿjāz* receives a clear and convincing explanation: the text of the *Qurʾān* is indeed inimitable, but all speakers of Arabic, inasmuch as they know and master the basic grammatical categories and rules of their language, can understand it and recognize its superiority.

The great contribution of al-Jurjānī in the debate around *ʾiʿjāz* is to have given to the term *naḏ̣m*, which had remained up to then a vague notion, a well-defined content likely to give rise to effective analytical procedures. He declares in this connection: "All I wanted to show you is that it is necessary, for any utterance which you find beautiful and any linguistic expression which you appreciate, that there be a recognized and rational foundation for this judgment and that we have means to express that and proofs to argue its validity" (*Dalāʾil* 33).

Al-Jurjānī strives to elaborate analytical procedures making it possible to characterize the precise way in which two different texts held to have the 'same *maʿnā*', in the traditional sense, actually differ from each other. His endeavors in this sense led him to his fundamental discovery, that of the existence of a strict correlation between a minimal variation of form and a minimal variation of meaning in utterances. This discovery, which establishes the principle of a strict functional relation between the form of an utterance and its semantic content, leads al-Jurjānī to 'displace' the sense of the word *maʿnā*: since a difference in *lafḏ̣*, however small, necessarily entails a difference in *maʿnā*, which will be perceived by the hearer, and since it must be admitted that a competent speaker always resorts deliberately to such differences, then, it must be admitted that the intention of the speakers (their *maʿnā*) is strictly correlated to the *lafḏ̣* they use to express it, and hence that the semantic value of that *lafḏ̣* is nothing but the *maʿnā* aimed at by the speaker. As a consequence, and, as far as we know, for the first time in the technical literature, the term *lafḏ* may be reduced to the univocal sense of signifier, i.e. as the purely material representative of *maʿnā*.

From this a complete reversal of the relationship between *lafḏ* and *maʿnā* follows: whereas, in the primitive view, two different texts could loosely be said to have the same *maʿnā*, and consequently any difference between them had to be seen as a pure difference in *lafḏ*, the new approach posits that the difference in form has at its root a difference in *maʿnā*, this difference being the necessary and sufficient reason why the two texts have a different *naḏm*, i.e. are two different texts. This radical reversal of perspective explains why al-Jurjānī so heavily insists on criticizing the traditional, deeply rooted view that two different *lafḏ*s may have the same *maʿnā*. This view, he says, "is absolutely untenable" (*fī ġāyat al-ʾiḥāla*; *Dalāʾil* 202).

Some superficial (or biased) readers have tried to argue that al-Jurjānī, as an ʾAšʿarite, was a 'champion of *maʿnā*' against the Muʿtazilites, who are supposed to have been 'defenders of *lafḏ*'. As a matter of fact, one would completely misinterpret al-Jurjānī's thought by trying to bring it down to one of the antagonistic terms of the traditional Arab conception of the *lafḏ/maʿnā* relationship. What makes him so important in the history of linguistic thought is precisely that he promoted a new methodology for the analysis of utterances resting on the thesis of a rigorous correlation of the two terms of the semiotic equation, so that any change in *lafḏ* necessarily entailed a change of *maʿnā* and vice versa.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

ʿAbd al-Jabbār, *Muġnī* = ʾAbū l-Ḥasan ʿAbd al-Jabbār ibn ʾAḥmad al-Hamāḏānī, *al-Muġnī fī ʾabwāb at-tawḥīd wa-l-ʿadl*. Ed. ʾAmīn al-Xūlī. Cairo, 1960.

Abū Hayyān, *Muqābasāt* = ʾAbū Ḥayyān ʿAlī ibn Muḥammad at-Tawḥīdī, *al-Muqābasāt*. Ed. Muḥammad Tawfīq Ḥusayn. Baghdad, 1970.

ʿAskarī, *Furūq* = ʾAbū Hilāl al-ʿAskarī, *Kitāb al-furūq fī l-luġa*. Beirut: Dār al-ʾĀfāq al-Jadīda, 1980.

Ibn Fāris, *Ṣāḥibī* = ʾAbū l-Ḥusayn ʾAḥmad Ibn Fāris, *aṣ-Ṣāḥibī fī fiqh al-luġa*. Ed. ʿUmar al-Tabbāʿ. Beirut: Maktabat al-Maʿārif, 1993.

Jāḥiẓ, *Bayān* = ʾAbū ʿUṯmān ʿAmr ibn Baḥr al-Jāḥiẓ, *Kitāb al-bayān wa-t-tabyīn*. Ed. Fawzī ʿAṭwī. Beirut: aš-Šarika al-Lubnāniyya li-l-Kitāb, 1968.

——, *Ḥayawān* = ʾAbū ʿUṯmān ʿAmr ibn Baḥr al-Jāḥiẓ, *Kitāb al-ḥayawān*. Ed. ʿAbd as-Salām Hārūn. Cairo: Maktabat al-Ḥalabī, 1938–1945.

Jurjānī, *Dalāʾil* = ʾAbū Bakr ʿAbd al-Qāhir ibn ʿAbd ar-Raḥmān al- Jurjānī, *Dalāʾil al-ʾiʿjāz*. Ed. Muḥammad Rašīd Riḍā. Beirut: Dār al-Maʿārif, 1978.

SECONDARY SOURCES

Kouloughli, Djamel Eddine. 1983. "À propos de *lafḏ* et *maʿnā*". *Bulletin des Etudes Orientales* 35.43–63.

——. 1997. "*Contre la synonymie*: Présentation et traduction française annotée du premier chapitre de l'ouvrage de ʾAbū Hilāl al-ʿAskarī (m. vers 1005): *Kitāb al-furūq fī l-luġa* [Le livre des différences dans la langue]". *Histoire Epistémologie Langage* 19:2.155–176.

——. 2002. "L'influence muʿtazilite sur la naissance et le développement de la rhétorique arabe". *Arabic Sciences and Philosophy* 12.217–239.

Mahdi, Muhsin. 1970. "Language and logic in Islam". *Logic in classical Islamic culture*, ed. Gustav E. von Grunebaum, 51–83. Wiesbaden: O. Harrassowitz.

Versteegh, Kees. 1997. "The Arabic tradition". Wout van Bekkum, Jan Houben, Ineke Sluiter, and Kees Versteegh, *The emergence of semantics in four linguistic traditions: Hebrew, Sanskrit, Greek, Arabic*, 227–284. Amsterdam and Philadelphia: J. Benjamins.

DJAMEL EDDINE KOULOUGHLI (CNRS)

Manṣub → ʾIʿrab

Marfūʿ → ʾIʿrab

Masculine → Gender

Maṣdar

1. INTRODUCTION

One of the first times a comprehensive discussion of the term *maṣdar* is encountered in the linguistic literature is in Sībawayhi's (d. ca. 177/793) grammatical treatise *al-Kitāb* 'the book'. His approach is presented here following Mosel (1974) by describing the category, form, and function of the *maṣdar* in Classical Arabic. Then, on the basis of Cantarino (1974–1975), the use of the *maṣdar* in Modern Literary Arabic is analyzed in order to allow for a comparison between past and present.

2. THE *MAṢDAR* IN CLASSICAL ARABIC

In the context of transitive verbs, Sībawayhi speaks about the ‘event’ and the ‘noun of the event’ from which the verb originates, hence its name *maṣdar* ‘origin, source’. In doing so he represents what will later be called the Baṣran point of view, as opposed to the Kūfan theory, which regarded the verb as the basis for the derivation of the noun (al-ʾAnbārī, *ʾInṣāf* I, *masʾala* 28).

Mosel (1974:185–189) follows Sībawayhi’s description of the *maṣdar* as an element of the category ‘noun’ but combining a noun-like behavior, as in example (6) below, with verb-like characteristics (cf. example (1)). Moreover, the *maṣdar* may realize different functions in the sentence which other elements of the category noun do not realize. On the other hand, Sībawayhi distinguishes between the occurrence of the *maṣdar* and the *ism al-maṣdar* ‘*maṣdar* noun’ (Sībawayhi, *Kitāb* I, 320.22–321/Būlāq I, 275.18):

Table 1. *Maṣdar* and *ism al-maṣdar*

maṣdar	translation	*ism al-maṣdar*	equivalent to	translation
wazn	weighing	*wazn*	*mawzūn*	weight
xalq	creating	*xalq*	*maxlūq*	creation
ḥalab	milking	*ḥalab*	*maḥlūb*	milk

3. THE REALIZATION OF THE *MAṢDAR* IN CLASSICAL ARABIC

Sībawayhi presents a number of examples of the use of the *maṣdar*:

(1) An indefinite *maṣdar* with an object accusative:
ʿajibtu min ḍarbin zaydan
was-surprised-I at beating zayd
(Sībawayhi, *Kitāb* I, 118.6/Būlāq I, 97.1)
‘A beating of Zayd surprised me’

(2) A definite *maṣdar* with an object accusative:
ʿajibtu mina ḍ-ḍarbi zaydan
was-surprised-I at the beating zayd
(Sībawayhi, *Kitāb* I, 120.8/Būlāq I, 99.1)
‘The beating of Zayd surprised me’

(3) An indefinite *maṣdar* with a nominative subject and an object accusative:
ʿajibtu min ḍarbin zaydun
was-surprised-I at beating zayd
ʿamran
ʿamr
(Sībawayhi, *Kitāb* I, 118.7/Būlāq I, 97.2)
‘Zayd beating ʿAmr surprised me’

(4) A *maṣdar* with a genitive subject and an object accusative:
ʿajibtu min ḍarbi-hi zaydan
was-surprised-I at beating-his zayd
(Sībawayhi, *Kitāb* I, 119.7/Būlāq I, 98.2)
‘His beating of Zayd surprised me’

(5) A *maṣdar* with a genitive object and a nominative subject:
ʿajibtu min ḍarbi-hi zaydun
was-surprised-I at beating-him zayd
(Sībawayhi, *Kitāb* I, 119.8/Būlāq I, 98.2)
‘Zayd’s beating of him surprised me’

(6) A *maṣdar* with two genitive objects (or subjects) in coordination:
ʿajibtu min ḍarbi zaydin
was-surprised-I at beating zayd
wa-ʿamrin
and-ʿamr
(Sībawayhi, *Kitāb* I, 120.3/Būlāq I, 98.8)
‘Zayd and ʿAmr being beaten surprised me’

(7) A *maṣdar* with a genitive object in coordination with an object accusative:
ʿajibtu min ḍarbi zaydin
was-surprised-I at beating zayd
wa-ʿamran
and-ʿamr
(Sībawayhi, *Kitāb* I, 120.4/Būlāq I, 98.9)
‘The beating of Zayd and ʿAmr surprised me’

4. THE FUNCTION OF THE *MAṢDAR* NOUN PHRASE IN CLASSICAL ARABIC

Mosel (1974:191–192) discusses the functions Sībawayhi assigns to the *maṣdar*. Her presentation is changed slightly in order to follow more closely Sībawayhi’s order of presentation and the selection of some of his alternative examples.

(8) A *maṣdar* construction instead of a finite verb of estimation (*ʾafʿāl al-qulūb*):
matā zaydun ḏ̣anna-ka ḏāhibun
when zayd opinion-your leaving
(Sībawayhi, *Kitāb* I, 80.8/Būlāq I, 63.8)
'When is Zayd, in your opinion, leaving?'

(9) A *maṣdar* as temporal adverbial in a sentence:
matā sīra ʿalay-hi? maqdama
when is-traveled upon-him? arriving
l-ḥājji
the-pilgrim
(Sībawayhi, *Kitāb* I, 137.2/Būlāq I, 114.1)
'When will he be used? When the pilgrims arrive'

(10) A *maṣdar* as absolute object but subject in a passive construction:
sīra ʿalay-hi sayrun šadīdun
is-traveled upon-him traveling intense
(Sībawayhi, *Kitāb* I, 140.7/Būlāq I, 117.7)
'He was used intensively for traveling'

(11) A *maṣdar* as substitute for a finite verb expressing a command or wish:
buʿdan
being-distant
(Sībawayhi, *Kitāb* I, 184.15/Būlāq I, 157.1)
'Take distance'

(12) A *maṣdar* as substitute for a finite verb expressing a habit:
ʾinna-mā ʾanta sayran sayran
however you moving moving
(Sībawayhi, *Kitāb* I, 197.17/Būlāq I, 168.10)
'How restless you are!'

(13) A *maṣdar* as a post-modifying comparison:
marartu bi-hi fa-ʾiḏā la-hu
passed-I by-him and-suddenly to-him
ṣawtun ṣawtu ḥimārin
sounding sounding donkey
(Sībawayhi, *Kitāb* I, 208.18; Būlāq I, 177.23)
'I overtook him and suddenly he produced a braying-like sound'

(14) A *maṣdar* in the function of object of reason:
faʿaltu ḏāka maxāfata fulānin
did-I that being-afraid someone
(Sībawayhi, *Kitāb* I, 216.6; Būlāq I, 184.11)
'I acted so for fear of somebody'

(15) A *maṣdar* describing the situation of the subject or object:
qataltu-hu ṣabran
killed-I-him fettering
(Sībawayhi, *Kitāb* I, 218.2; Būlāq I, 186.6)
'I killed him fettered'

(16) A *maṣdar* with an emphasizing sentence adverbial value:
hāḏā ʿabdullāhi ḥaqqan
this 'abdallah being-true
(Sībawayhi, *Kitāb* I, 221.23; Būlāq I,189.19)
'This really is ʿAbdallah'

5. THE *MAṢDAR* IN MODERN LITERARY ARABIC

In more recent literature, different terms are used to refer to the *maṣdar*, e.g. 'verbal noun', 'infinitive', 'event noun', '(deverbal) process nominal' as opposed to 'result nominal'. They all have in common their emphasis that the *maṣdar* belongs to the category 'noun' but may exhibit some characteristics of its corresponding verb. A standard definition speaks about the occurrence, in nominal form, of the corresponding verb without any reference to the aspect of time: "The nomina verbi, أسْمَاءُ الْفِعْلِ, are abstract substantives, which express the action, passion, or state indicated by the corresponding verbs, without any reference to object, subject, or time" (Wright 1974:I, 110). Here, the term *maṣdar* is used with the following definition: 'a noun, conditioned in its combinatorial behavior at phrase and sentence level by the semantic value and argument structure of the corresponding verb'.

6. THE LEXICAL FORM OF THE *MAṢDAR*

The root-and-pattern-based Arabic language system uses vowels (or their absence) and a limited number of auxiliary (semi)consonants in combination with the productive tri- and quadriliteral verbal root combinations to compose the most frequently used nominal representations of the abstract semantic value of the verbal root combination involved, as well as all possible regular *maṣdar* derivates within the Literary Arabic language system such as the *ism al-marra*, i.e. expressing a single realization of the finite verb.

Notwithstanding some phonologically and/or semantically based regularities, the *maṣdar* forms of stem I of triliteral root combinations are scarcely predictable. One needs to consult the dictionary to be sure about them. The lexical form of *maṣdar*s of the derived stems of the triliteral root and those of quadriliteral roots are very regular and far more predictable (→ verbal noun). Wright (1974:I, 110ff.) presents a rather comprehensive list of *maṣdar* forms.

7. THE FUNCTION OF THE *MAṢDAR* NOUN PHRASE

As a noun, a *maṣdar* can realize the functions any element belonging to the category 'noun' may fulfill in a sentence. Any exception to this statement is due to specific characteristics of the *maṣdar*.

(17) A *maṣdar* noun phrase as topic in a nominal sentence:
safku d-dimāʾi muḥarram
shedding the-blood forbidden
(Cantarino 1974–1975:I, 15.4)
'Shedding of blood is forbidden'

(18) A *maṣdar* noun phrase as comment in a nominal sentence:
as-sababu ṯ-ṯānī intiqālu ʿāṣimati
the-reason the-second transfer capital
l-xilāfa
the caliphate
(Cantarino 1974–1975:I, 18.6)
'The second reason is the transfer of the capital of the caliphate'

(19) A *maṣdar* noun phrase as subject in a verbal sentence:
qad kāna stixdāmu l-mawālī
certainly was-he employment the-clients
nādiran
rare
(Cantarino 1974–1975:I, 46.5)
'The appointment of clients (*mawālī*) was not common'

(20) A *maṣdar* noun phrase as direct object in a verbal sentence:
wa-lākinna-ka tafʿalu mā-lā ʾastaṭīʿu
and-but-you you-do what-not I-can
fiʿla-hu
doing-his
(Cantarino 1974–1975:II, 402.10)
'But you do things that I cannot do'

(21) A *maṣdar* noun phrase as prepositional object in a verbal sentence:
ʾiḏā šaʿartu bi-suqūṭi l-kitābi min
when felt-I by-falling the-book from
yad-ī
hand-my
(Cantarino 1974–1975:II, 402.1)
'when I felt the book falling from my hand'

(22) A *maṣdar* noun phrase as (temporal) sentence adverbial:
qubayla buzūġi š-šamsi
a-little-before rising the-sun
(Cantarino 1974–1975:II, 402.2)
'shortly before sunrise'

(23) A *maṣdar* noun phrase as verb phrase adverbial:
ṭaradtu-hā min hunā ṭarda
threw-I-her from here throwing
l-kilābi
the-dog
(Cantarino 1974–1975:II, 170.4)
'I threw her out as one would a dog'

8. THE STRUCTURE OF THE *MAṢDAR* NOUN PHRASE

A *maṣdar* noun phrase is a phrase in which a *maṣdar* realizes the head function. The semantic value of the head is defined in the lexicon and represents the abstract meaning of the corresponding verb in its base or derived stem. The number of arguments that can co-occur in a *maṣdar* noun phrase is lexically determined. The relationship between the head and its arguments may be governed by the characteristics of the head as belonging to the category 'noun' or be expressed by verbal characteristics of the corresponding verb.

The semantic value of the *maṣdar* blocks the occurrence of elements such as pre- or post-modifying demonstratives. It also precludes some other nominal processes such as the occurrence of a dual, plural, or diminutive form. On the other hand, a verbal modifier may occur (Fassi Fehri 1993:232–237).

The minimal structure of a *maṣdar* noun phrase is the occurrence of an obligatory head realization. Optional extensions are the occurrence of a pre- or post-determiner, possible complements, and post-modifying elements.

(24) Simple head realization:
zawāju-hu tazwīrun ḥayātu-hu
marriage-his falsifying live-his
tazwīrun
falsifying
(Cantarino 1974–1975:I, 18.7)
'His marriage was a fake and so was his life'

(25) A head realization with a pre-determiner:
li-māḏā l-bukāʾu
for-what the-crying
(Cantarino 1974–1975:I, 15.5)
'Why this crying?'

(26) A head realization with a post-modifier:
ḥattā sāfartu min al-qāhira safaran
until traveled-I from Cairo traveling
ṭawīlan
long
(Cantarino 1974–1975:II, 170.1)
'until I left Cairo on a long trip'

(27) A head realization with a post-determining genitive of the subject:
qabla majīʾ-ī hunā
before coming-my here
(Cantarino 1974–1975:II, 402.4)
'before my coming here'

(28) A head realization with a post-determining genitive of the object:
hamma bi-ʾidxāli yadi-hi
tried-he by-introducing hand-his
fī ṣ-ṣundūqi
in the-box
(Cantarino 1974–1975:II, 402.7)
'He tried to introduce his hand in the box'

(29) A head realization with a genitive of the subject and an accusative of the object:
min ṭalab-ī taḥwīla l-ḥurriyya
of asking-my granting the-freedom
li-n-nisāʾi
for-the-women
(Cantarino 1974–1975:II, 403.1)
'because I asked to grant freedom to women'

(30) A head realization with a genitive of the subject and prepositional object:
kāna yaʿrifu min hāʾulāʾi š-šabābi
was-he he knows of these the-boys
ḥubba-hum li-l-ʿilmi
loving-their for-the-science
(Cantarino 1974–1975:II, 404.6)
'He knew how much these young men loved to learn'

(31) A head realization with a genitive and an accusative of the object:
wa-lam yaktafi l-xūrī ʾilyās
and-not was-satisfied the-priest ilyas
bi-ʾiblāġi š-šayxi hāḏā l-xabara
with-bringing the-sheikh this the news
(Cantarino 1974–1975:II, 404.2)
'Father Ilyas, the priest, was not satisfied with just bringing such news to the sheikh'

9. COMPARISON AND CONCLUSION

Looking at the examples of *maṣdar* realization in Classical Arabic, one might say that a number of them (examples (1–3), (5), and (7)) seem to be theoretical constructs. In Modern Literary Arabic, a *maṣdar* with a nominative subject no longer occurs. The construction with a subject or object post-determiner (6) or with a subject post-determiner and object complement (4) seems to be the most common or most frequent realization in Modern Literary Arabic, but this has to be verified by a statistical corpus inventory (Badawi a.o. 2004: 237–241).

The *maṣdar* has an important function as source for the expansion of the lexicon, especially but not exclusively in the domains of science and technology. The notions 'event noun' and 'process nominal' are often used in describing the *maṣdar* , as opposed to a 'result nominal'. For a number of examples of 'result nominals' as *maṣdar* derivates, we refer to Monteil (1960:111ff.).

Bibliographical references

Primary sources

ʾAnbārī, *ʾInṣāf* = ʾAbū l-Barakāt ʿAbd ar-Raḥmān ibn Muḥammad al-ʾAnbārī, *al-ʾInṣāf fī masāʾil al-xilāf bayna n-naḥwiyyīna l-baṣriyyīn wa-l-kūfiyyīn*. Damascus: Dār al-Fikr.

Sībawayhi, *Kitāb* = ʾAbu Bišr ʿAmr ibn ʿUṯmān ibn Qanbar Sībawayhi, *Kitāb Sībawayhi*. Ed. Beirut: Muʾassasat al-ʾAʿlamī li-l-Maṭbūʿāt, 1967. Ed. Būlāq, 1316/1317 A.H.

Secondary sources

Badawi, Elsaid, Michael G. Carter and Adrian Gully. 2004. *Modern written Arabic: A comprehensive grammar*. London and New York: Routledge.

Cantarino, Vicente. 1974–1975. *Syntax of modern Arabic prose*. 3 vols. Bloomington and London: Indiana University Press.

Fassi Fehri, Abdelkader. 1993. *Issues in the structure of Arabic clauses and words*. Dordrecht: Kluwer Academic Publishers.

Monteil, Vincent. 1960. *L'arabe moderne*. Paris: C. Klincksieck.

Mosel, Ulrike. 1974. *Die syntaktische Terminologie bei Sibawaih*. Ph.D. diss., University of Munich.

Wright, William. 1974. *A grammar of the Arabic language*. Beirut: Librairie du Liban.

Everhard Ditters (University of Nijmegen)

Mauritania

In 1960, Mauritania, a French colony along the Atlantic Coast in West Africa, achieved independence under the name of the Islamic Republic of Mauritania. The name 'Mauritania', once used as the name of the Roman colonies of North Africa (*Mauretania*), was reintroduced by the colonial administration to designate this western part of the Saharo-Sahelian zone, which was called in Arabic literature *bilād Šinqīṭ* 'country of Šinqīṭ', *trāb al-bīḏ̣ān* 'land of the Whites' (Taine-Cheikh 1990), or *bilād as-sayba* 'country of anarchy'.

The borders of the country are as arbitrary as its name. Indeed, they do not follow at all the territorial limits of the Moorish Arabic-speaking country, which extended far beyond those borders, especially in the north, in the region of the Rio de Oro and of the Sagya el-Ḥamra, occupied by the Spanish. The Senegal River, chosen as the southern border, is inhabited on either side by the same black African populations, which have Pulaar, Soninke, or Wolof as their mother tongue.

Originally, Mauritania was conceived as a 'transition' country between 'arabity' and 'africanity' (Baduel 1990), hence the situation of Arabic is particularly complex but also, as in other places, subject to ceaseless change.

1. Historical background: The religious 'Arabization'

A variety of Arabic (→ Ḥassāniyya) is currently the mother tongue of a majority (usually estimated between 70 and 80 percent of a total of 2.9 million people) of the Mauritanian population. This is the consequence of the abandonment of the Berber language (Zenaga), at the end of a very long process of Arabization, which seems to have begun at the end of the 14th and beginning of the 15th century but has not yet been completed even today. The first contacts with the Arabic language took place at the end of the 1st millennium, through the Islamic religion, and concerned all the ethnic groups of the region.

For all societies concerned, Islamization represented a global cultural phenomenon, but its linguistic effects were variable. Although a perfect command of Classical Arabic is strongly recommended for every good Muslim, this perfect command was (and still is) often much more limited than expected. The ancient populations of Mauritania were no exception to this rule, although a relatively large number of people apparently became literate. This is probably due to the history of the region and the social organization of the different ethnic groups.

The Islamization of the Saharo-Sahelian part of West Africa was linked with the Almoravid movement, initiated in the 11th century by the Berber tribes of the region, according to local traditions, somewhere between the Adrar and the Senegal delta. This movement was to be very successful in Morocco and Spain (Norris 1972, 1986). The name 'Almoravid' is supposed to have come, through the Spanish language, from *murābiṭūn*; it probably does not signify 'those from the *ribāṭ* "fortified monastery"', as has been supposed for a long time, but rather 'those who wage a holy war' (participle of the verb *rābaṭa*).

Nevertheless, the penetration of Islam does not begin with the armed fighters of Yaḥyā ibn ʾIbrāhīm al-Gdalī. In fact, conversions had already taken place since the middle of the 8th century, especially among the Lemtuna Berbers and the Soninke of the Ghana empire. This peaceful Islamization was linked with trade (in particular with the Kharijite Mzab) but remained quite superficial. The local sources do not always distinguish it from the traditions concerning the Bafour, the mysterious population from the Adrar that is sometimes associated with the breeding of dogs. The preaching of Ibn Yāsīn, spread by the Almoravid troops, was a message of faith, but it also aimed at eradicating the strong influence exerted until then by Ibadite Kharijism in the Sahara and on its southern (and northern) borders. Henceforth, the aim was to attain a deeper knowledge of the sacred texts and a more orthodox practice of religion, in particular through the banishment of the pleasures of music and dance.

Apart from the Almoravid episode (and, later, some Peul Jihadist movements), the conquests did not play a very important role in the diffusion of Islam in sub-Saharan Africa. Yet, the numerous conflicts, between the communities as much as between the tribes, and even within the tribes, resulted among other things in social fissure and tribal or 'lineage' specializations. Some tribes chose to relinquish their arms and dedicate themselves entirely to the study and teaching of the religious sciences, particularly at the end of the Šurbubba war that in the 17th century opposed two tribal coalitions in southwest Mauritania. The 'learned' tribes are the *zwāya* among the Arabic speakers and the *guḏ̣ayən* (literally '*qadi*s') among the Zenaga speakers. In addition, among the neighboring Pulaar speakers, the influence of the Torobe group, representing the majority, relies partially on its religious status.

In all ethnic groups (including the Soninke and the Wolof), the learned men had at their disposal a unified corpus of reference, based on the triad of Malikism, Ashʿarism, and Sufi brotherhoods, and applied largely similar methods to transmit knowledge. These methods made a strong demand on memory and used all possible memorization techniques: repetitions and recitations (supported by rhythmic movement of the body), poetry (even about abstract topics such as grammar), and copying of texts. Writing was at the center of learning, but transmission took place from the master's voice to the student's ear. The performance was less an oralization, corresponding to a real command of Classical Arabic as a language of oral communication, than an auralization, meaning the recitation of literary Arabic (Taine-Cheikh 1998; Ould Cheikh 1998). Over the years, however, teaching has diversified and deepened thanks to a more frequent recourse to the mother tongue in order to explain the meaning of the text. Altogether, the level of Arabo-Islamic culture was significant, especially but not exclusively in the traditional schools (*maḥāḏ̣ər*) of the desert.

2. The dialectal Arabization

The four major Mauritanian caravan cities of the 2nd millennium (Wadān, Šingīṭi, Tišīt, and Walāta) were founded around the 12th and 13th centuries, as the decline of Āwdāġust (probably to be identified with modern Tägdāwəst, in southeast Mauritania) was already sealed. At that time, the Islamic religion had already reached an important breakthrough in the region, paving the way for a certain adaptation to the Arabic language. The only Arabic-speaking communities, though, still seem to have been made up by small groups of traders coming from the Maghreb. In all cities with a Berber majority, whose destiny was unquestionably linked to the road taken by the trans-Saharan trade, the Azer language – probably a variety of Soninke as spoken by Zenaga speakers – may have played a key role as lingua franca, despite the weakening and finally the disappearance of the Ghana empire.

The influence of Arabic-speaking groups only began to be felt in the Sahara from the 15th century onward. At the end of the 14th century, Ibn Xaldūn had pointed out the presence of the Banū Ḥassān in the Dra wadi (*wād darʿa*) in the south of Morocco. In constant rivalry with their cousins, the Šbanāt, they oppressed their neighboring Berber tribes (Ould Cheikh 1995:43). Ibn Xaldūn traced back the genealogy of their chief, Ḥassān, to a certain Maʿqil, but he did not specify the relationship between the latter and the Banū Hilāl. One should be careful not to take at face value a history of the Banū Maʿqil, even if certain authors tend to present them as a group distinct from both the Banū Hilāl and the Banū Sulaym.

The testimony of Arabic and Portuguese travelers provides some information about the migration of the Banū Ḥassān to the south and their slow penetration into the Sahara.

Nevertheless, there is, unfortunately, a lack of data for the reconstruction of the history of this 'dark age', during which the dominance of the Arabs over (part of) the local Berber populations was established. Although the traditions and the local denominations encourage simplistic equations – 'warlike tribes = Arabs' vs. 'maraboutic tribes = Berbers', or more recently, 'aristocratic = Arabs' and 'those who pay tribute = Berbers' – there is some evidence that history has given rise to a complex society whose culture probably achieved some kind of symbiosis between 'arabity' and 'berberity', even though generally speaking, only the Arabic part of the heritage is claimed.

As regards the language, the name Ḥassāniyya (or *klām ḥassān* lit. 'the language of Ḥassān'), assigned to the spoken Arabic of Mauritania, clearly suggests that this dialect is a legacy from the Ḥassān tribes. In view of what is known about Arabization in the Saharo-Sahelian zone, this identification is not surprising in itself. Because all Arabic-speaking groups that came to settle in this area claim to be of the same origin, it is not too far-fetched to think that the fundamental characteristics of the Mauritanian dialect were already present in the 15th century. This is all the more plausible since even today for Ḥassāniyya speakers mutual comprehension seems to be easiest with the Bedouin in the whole Arabic world, not only from the Maghreb but also from the Middle East (especially Jordanians). Besides, the Ḥassāniyya language shows a remarkable homogeneity from east to west and from north to south (and even beyond the Mauritanian borders). This fact would seem to support this theory, although it does not explain one of the rare important regional differences, the occlusive or fricative pronunciation of /ġ/.

Of course, even if the Ḥassāniyya language has retained many characteristics from the dialect once spoken by the Banū Ḥassān, this does not mean that it has gone through the centuries without changes. Even without mentioning the most recent evolutions, the numerous borrowings from local dialects, especially Zenaga, show the lexical enrichment produced by the contact with the Berber substrate language.

Various morphosyntactic innovations are particularly characteristic of the Ḥassāniyya language, especially the passive voice, the diminutive, and the elative (→ Ḥassāniyya Arabic). Despite certain similarities with Zenaga, the neologisms of the Ḥassāniyya language cannot be interpreted as a simple calque but must be regarded as an internal evolution of the Arabic dialect. Although it is even harder to understand the general uniformity of innovations than the permanency of conservatisms, the successful innovations must have corresponded to particularly important needs of expression for the Moorish society.

3. WRITTEN LITERATURE IN THE CLASSICAL AGE

The dispersal of the Banū Ḥassān and their settlement in Saharan Mauritania led to new relations between the Arabic speakers and the (former) Berber speakers. Between the 17th century and the first half of the 18th century, four emirates (Trarza, Brakna, Adrar, and Tagant) were established, which corresponded to the early stages of political concentration, when a family of warriors (coming from the Banū Ḥassān, except in the case of the Tagant) exerted their authority on the tribes of the region. Starting at the end of the 18th century and flourishing above all in the 19th century, an era ensued that seems to have been propitious for the development of culture and literature.

Inscriptions in *tifinaġ* characters seem to have ceased around the 15th/16th centuries. Between the arrival of the Banū Ḥassān and that of the European colonizers, virtually all writing was done in Arabic characters, probably most of it in Classical Arabic, because there are few traces in Mauritania of Berber manuscripts written in Arabic characters (did they disappear?), although there did exist literary productions in Soninke and Pulaar, some of which, such as the *beytol*, were written with an Arabic alphabet.

Broadly speaking, the concept of written literature is applicable to all intellectual productions recorded in the familial libraries. These were extremely numerous, even if they were often limited to the contents of a trunk. In the case of the literate Moors, the trunk was carried around on camels when the camp was moving. The Ḥassāniyya speakers of the western Sahara take a lot of pride in the fact that they were one of the rare nomadic societies in the Arab world to be strongly attached to the book and to the study of Muslim Arabic culture (Bonte and Claudot-Hawad 1998).

In the last few decades, some collections have been institutionalized, for both material and political reasons. The attention of the institutional power and the public was focused on this small number of libraries, particularly in the ancient cities of the Sahara. Such sedentary establishments may not have been the general rule, but the inventory of their contents gives an idea of the texts that were bought or copied most frequently.

Between Šingīṭi and Wadān, for example, twelve family libraries can be counted. The contents are variable in size, from several pages to hundreds of pages, and the 1,106 documents are unequally shared between the libraries. The library of the ʾAhl Ḥabät from Šingīṭi alone contains more than half of these documents, most of them purchased. Founded by Sīdi Muḥammad wəll Ḥabät in 1845 upon his return from his Mecca pilgrimage, the library is said to have contained up to three thousand books. As in most libraries, the great majority of the books date from the 19th century, but a considerable number date from the 17th and 18th centuries. Some are even more ancient; indeed, five copies of manuscripts made prior to the end of the 15th century are listed, with ten copies realized in the 16th century. The oldest document kept in Mauritania can also be found here, a copy (made in 480/1087–1088) of a commentary on the *Qurʾān* written by the Iraqi author ʾAbū Hilāl al-ʿAskarī (d. 395/1004–1005).

The books contained in the libraries of Šingīṭi and Wadān are mainly about religion and jurisprudence: about 40 percent on theology (*fiqh, ʾuṣūl, qawāʿid*, *nawāzil*) and almost 30 percent on the Qurʾānic sciences (copies of the *vulgate*, the exegesis, the words of the Prophet, and the hagiographical stories) and mysticism (*taṣawwuf*). Among the remaining 30 percent, mathematics (1.70%) and logic (2.78%) are relatively well represented, more so than history, astronomy, and medicine. But linguistic topics (*naḥw*, *ṣarf*, *luġa*, and *ʾadab*) are particularly popular (23.77%). The great lexicographical corpus *al-Qāmūs al-muḥīṭ*, for instance, gathered by the scholar from Širāz, Muḥammad ibn Yaʿqūb al-Fīrūzābādī (d. 817/1414–1415), was written in calligraphy for the library of the ʾAhl Ḥabät over the course of several years (the copy of the two first volumes dating from 1251/1835–1836 and of the last two from 1260/1844).

With regard to Mauritanian scholars, no writings are known before the 18th century (al-Bartalī 1981; Ould Bah 1981; Ḥāmidun 1990; Rebstock 2001). With the apparent exception of a Wadanian scholar of the 16th century who left a written commentary on the *Qurʾān*, the most ancient Moorish author known is the great *faqīh* of Šingīṭi, Muḥammad wəll al-Muxtār wəll Billaʿmaš (1625–1695). This major figure of the cultural history of the western Sahara wrote, among other things, a commentary on a book about the foundations of the dogma and one on astronomy.

It is precisely because the Moorish books are so recent that the middle of the 18th century appears as the beginning of a new era. However, in a strictly literary sense, Mauritanian cultural production was of variable value. Indeed, there is not a lot of prose literature, and it is often badly represented. Globally, works on *ʾadab* are rare. The only real prose writer seem to have been aš-Šayx Sīdi Muḥammad al-Kuntī (d. 1826), who wrote a biography of his father (the great mystic aš-Šayx Sīd al-Muxtār) and a number of treatises.

On the other hand, poetry is both abundant and of high quality, as demonstrated by ʾAḥmad ibn al-ʾAmīn aš-Šingīṭi. In Cairo, he wrote from memory a book on his country of origin, *al-Wasīṭ fī tarājim ʾudabāʾ Šinqīṭ* 'The best [book] on the work of poets and men of Šinqīṭ' (Miské 1970), including no fewer than 4,500 lines of verse. This anthology brings together 82 poets, divided into 18 tribes (all maraboutic), most of them from the 19th century, testifying to the vitality of the classical *qaṣīda* of the Moorish literate elite and of their high level in literary Arabic (Ould Bah 1971:26–48; Ṭulba 2000).

Since Classical Arabic poetry belongs to the domain of written (or 'auralized') literature, presumably the entire oral literature in Arabic is expressed in dialect. Indeed, despite the numerous isomorphisms between the written and the oral spheres, the separation between the two fields coincides almost exactly, at least until the 20th century, with the distinction between the two varieties of Arabic in use in Mauritania.

The first resemblance to note is the preeminence of poetry as literary genre (Martin-Granel a.o. 1992; Bariou a.o.1995). Even though Classical Arabic poetry is called → *šiʿr*, in Ḥassāniyya it has the generic name of *ġnä*. The metrical system of oral poetry (Taine-Cheikh 1985) presents many common points with the classical → meters in its general principles (regularity of the number of syllables per meter, quantitative distinction between short and long syllables, constant presence of a rhyme). More precisely, there may even be a fundamental common opposition between ascending rhythms (with a metrical accent on the long second syllable of the *watid majmūʿ* 'rising foot': short + long) and descending rhythms (with a metrical accent on the long first syllable of the *watid mafrūq* 'descending foot': long + short).

Despite the obvious resemblance between the two metrical systems, there are important differences as well. The first of these is the general tendency to simplify in the *ġnä*, through a reduction of the number of syllables per meter and through an increase of the proportion of short syllables compared to the long ones, the latter tending to remain only in rhyme. The second one is the adaptation of the principle of quantity to the vocalic system of the dialect. Open syllables of the Cv type having almost disappeared in Ḥassāniyya, closed syllables CvC or long vowels Cv are counted as short in contrast with the 'extra-long' syllables CvCC (twice closed) or Cv̄C (closed with a long vowel). The third difference concerns the use of rhyme in a verse unit that seems peculiar to the dialect, even if it shows some similarities with other forms of poetry expressed in dialect. On the one hand, the *gav* is made up of four hemistichs with alternating rhymes *ab-ab*, and on the other hand, the *ṭalʿa* is made up of six hemistichs *(aa-ab-ab)*, differing from the quatrain through the two first identical rhymes.

Until the 20th century, the great themes of the *ġnä* (Taine-Cheikh 1994) were very close to those of the *šiʿr*, in spite of the difference in name. Put simply, one might say that there were eulogies (*madīḥ* or *ṯanāʾ* for the *šiʿr, šəkr* for the *ġnä*) and satire (*šatm* or *hijāʾ* in Classical Arabic, *šätm* or *ʿayb* in Ḥassāniyya), on the one hand, and elegies and love poems (respectively *nasīb* and *ġazal*, although these apply mostly to *šiʿr*), on the other. The last two themes, very frequent, were practiced by authors (*mġannyīn*) who belonged more or less to all social classes. Many anonymous poems belong to common culture, even if they sing about a particular region, the one of the poet (Sīdi Brāhīm 1992). Others have well-known authors (Ould Zenagui 1994), sometimes very famous, including, in certain cases, those known for their knowledge or for their poetry in Classical Arabic. The verses may be purely elegiac or only about love, but very often they are both simultaneously. The expression of feelings of love always respect the laws of decency (Tauzin 1982, 1990) and often are limited to mention of places formerly frequented by the loved one (always a woman, as the men traditionally kept for themselves the right to compose *ġnä*, leaving to women only the minor, and historically more recent, form of the distich called *təbrāʿ*).

The writing of eulogies and criticism was subject to even more constraints. If the recitation of poetry took place in a context of rhymed exchanges between people of equivalent status, they assumed the form of sparring matches (the *gṭāʿ*, which imposed certain rules of meters and rhymes). As these matches often played a role in the rivalries between the tribes, certain warriors, including the chiefs, distinguished themselves in it. However, more generally (and with the exception of the very particular case of the *madḥ ən-näbi* 'the praise of the Prophet', which was composed and sung only by the former slaves and the *ḥrāṭīn*), eulogies as well as satire were inseparable from the very closed social group of the musician-singers (Norris 1968; Guignard 1975). It was their role not only to sing the *ġnä* (which is not necessarily sung, despite the meaning of its root *ġ-n-y*), but to do and undo reputations. According to certain local traditions (Ould Bah 1971:14), their ancestors, in the 18th century, were responsible for the most ancient verses known in Moorish poetry. The long poems with epic accents that some *griots* like Säddūm wəll Ndyartu or ʿAlī wəll Mānu composed in honor of their warrior chiefs constitute a particular genre (*thäydīn*) of Moorish poetical heritage, quite esoteric but also highly regarded.

Nonpoetic genres exist, of course, but not all are represented. In the field of the narrative forms, there are mainly fairy tales and stories. Of great variety, they are not limited to wisdom fairy tales and marvel stories intended for children (Tauzin 1993; Ould Mohamed Baba 2000–2001; Ould Ebnou n.d.). In the discursive field, mainly short forms can be noted. Among the dialogic, playful, and/or didactic forms, there is the one, quite common, of the riddle (*tḥāži*) – often around wordplay – and the more specific one of the pastoral enigma (Taine-Cheikh 1995) in which the playful rivalry between shepherds expresses itself (this is called *zärg*, like the riddles students of Classical Arabic posed to one another). Finally, proverbs and sayings (*'amṯāl*) are also found in large numbers (Ould Ebnou n.d.), which demonstrates the strong fondness of Ḥassāniyya speakers for gnomic speech.

4. Oral literature

Mauritanian Arabic literature peaked in the 19th century, but the seeds of change were already present before that time. The influence of France began to develop in the south in 1857 and soon led to a tight control of the whole Senegal River valley. The beginning of the colonial conquest itself took place at the beginning of the 20th century. As of 1920, Mauritania was officially regarded as a French colony, governed from Saint-Louis in Senegal. The effects of colonization were felt for a long time after the granting of independence.

For several decades, colonization, which was carried out essentially from the strategic perspective of pacifying the region, remained superficial. Its impact was especially weak on the nomadic world, which only experienced indirect administration (enlistment in the *goums*, particular groups of military nomads, taxes imposed on the tribes, etc.). During that time, the sedentary black Africans of the earlier colonized valley were already subject to conscription and scheduled taxes. The Moors offered a particularly tenacious resistance to the French education system, and the most aristocratic people did not hesitate to send the children of their slaves or their dependents to school instead of their own children, when pressure became irresistible. As the backwardness of Ḥassāniyya speakers increased, the colonial authorities agreed to open special schools, known as *medersa*s, for the sons from good Moorish families, in which Arabic was given an important place. These schools were abolished in the 1940s, but as compensation a few hours of Arabic were introduced in all schools attended by Ḥassāniyya speakers. This measure tended, however, to be withheld from the black Africans, under the pretext of making a distinction between the Arabic language as a 'language of culture' (reserved for speakers of Arabic) and as a 'language of religion' (excluded from the French state school system).

One of the effects of the colonial policy was that of modifying social relations, weakening the power of the warriors and supporting the position of the marabouts *(zwāya)*. It is no accident that the first president of Mauritania, Mokhtar Ould Daddah, and most of the Moorish executives of the young state belonged to the literate tribes, in particular those of the southwest, who had been schooled before the others. However, the securing of the latter's loyalty was realized to the detriment of traditional education, and, in the course of the 20th century, the influence of the *maḥāḏ̣ər* and the number of students attending them continued to diminish.

The nascent republic was qualified as 'Islamic', based on the idea that religion was the common denominator of the entire population of Mauritania, but the language of administration and the education system were French, even if at the time this concerned only a small minority of children. Very soon, this official predominance of a foreign language was denounced by the Moorish community.

Ever since its creation, Mauritania has belonged to various organizations uniting the black African countries formerly colonized by France. Its recognition by the other Arabic countries and its entry into the Arab League were less immediate because of the opposition voiced by Morocco. Only by the end of the 1960s did the Moroccan claims cease, opening the possibility for a readjustment between Arabic and the black world, more in accordance with the wishes of the Moorish community (Ould Cheikh 1995:32–33). However, the changes were fought by the black African communities of Mauritania, who regarded them as harmful to their vested benefits, inherited from colonization, and perceived them as contrary to their elementary political rights.

If one considers the education system of Mauritania, which is usually the first stumbling block in contacts between the ethnic groups, it becomes clear that the country has never ceased to Arabize since its independence. The first measures were directed at the whole student body: Arabic was imposed on all secondary school students, at least as a second foreign language, and before the primary cycle a first year was added, entirely in Arabic, under the pretext of introducing Qurʾānic Arabic. Soon, the influence of Arabic nationalist movements, affected by events in the Middle East, was reinforced by the massive settlement of nomadic people. Ruined by years of severe drought, Moorish cattle breeders were asking for schools for their children and possibilities of employment in the administration for those who were educated in the *maḥāḏər*. Under these circumstances the number of hours of Arabic were increased considerably, facilitating the integration into the state education system of teachers coming from the traditional system. One of the two curricula put in place – the one dominated by the Arabic language – was then invaded by students who were total beginners in the French language. In the 1980s, the contrast deepened between the 'Arab' course of study, compulsory for all Ḥassāniyya speakers and with a minor place for French, and the 'bilingual' one, leading generally only to a master's degree in French and attended by a majority of the black African population.

This system, which through two different courses of study was supposed to lead to Arabic/French bilingualism, was finally abandoned because of the costs involved and the inefficiency (Taine-Cheikh 2004). The balance of power became extremely unfavorable to the black Africans after the ethnic conflict of 1989, the most violent since independence. Consequently, the government decided to abolish the bilingual course of study. The effect of the measure was softened by the existence of a private educational system that was increasingly successful. Additionally, the French language was not removed entirely from the educational system, although the precedence of the Arabic language was affirmed. This was facilitated by the adoption of Arabic as the only official language as early as 1991, one of the consequences being a significant Arabization of toponyms, often to the detriment of French and Berber names, which had long been in use (Ould Cheikh 1995:33–34).

The 21st century seems to have opened a new era in which Arabization will once again experience considerable progress, probably being marked as much by education as by the Arabic media (in particular satellite television). The future will tell what the result will be for the practice of literary Arabic, of the Ḥassāniyya dialect, and of its 'modernized' version, the local version of standard Arabic (Taine-Cheikh 2002, 2004).

Bibliographical references

Primary sources

Šinqīṭī, ʾAḥmad ibn al-ʾAmīn aš-. 1329/1911. *al-Wasīṭ fī tarājim ʾudabāʾ Šinqīṭ*. Cairo: al-Xānjī. (Repr., Cairo and Casablanca: al-Xānjī, 1958, 1989.)

Ṭulba (Ibn aṭ-), M. 2000. *Dīwān*. Nouakchott: ʾAḥmad Sālim ibn ab-Būh.

Secondary sources

Baduel, Pierre-Robert (ed.). 1990. *Mauritanie entre arabité et africanité*. (= *Revue du Monde Musulman et de la Méditerranée* 54.) Aix-en-Provence: Edisud.

Bariou, Jacques, Ousmane Moussa Diagana, Marie-Clotilde Jacquey, and Catherine Taine-Cheikh (eds). 1995. *Notre Librairie ("Littérature mauritanienne")* 120–121.

Bartalī, Muḥammad ibn ʾAbī Bakr aṣ-Ṣiddīq al-. 1981. *Fatḥ aš-šakūr fī maʿrifat ʾaʿyān ʿulamāʾ at-Takrūr*. Beirut: Dār al-Ġarb al-ʾIslāmī.

Bonte, Pierre and Hélène Claudot-Hawad (eds). 1998. *Nomadic Peoples* (NS) 2:1–2. (Repr., *Les Cahiers de l'IREMAM* 13–14, 2000.)

Guignard, Michel. 1975. *Musique, honneur et plaisir au Sahara*. Paris: Geuthner.

Ḥāmidun (ibn), Muxtār. 1990. *al-Ḥayāt aṯ-ṯaqāfiyya fī Mūrītāniyā*. Tunis: ad-Dār al-ʿArabiyya li-l-Kitāb.

Martin-Granel, Nicolas, Idoumou Ould Mohamed Lemine, and Georges Voisset. 1992. *Anthologie de la littérature mauritanienne contemporaine*. Paris: L'Harmattan.

Miské, Ahmed-Bâba. 1970. *Al Wasît: Tableau de la Mauritanie au début du XX*$^{\text{è}}$ *siècle*. Paris: Klincksieck.

Norris, H.T. 1968. *S̲h̲inqīṭī folk literature and song*. Oxford: Clarendon Press.

——. 1972. *Saharan myth and saga*. Oxford: Clarendon Press.

——. 1986. *The Arab conquest of the western Sahara*. Beirut: Librairie du Liban.

Ould Bah, Mohamd El-Moktâr. 1971. "Introduction à la poésie mauritanienne (1650–1900)". *Arabica* 18.1–48.

——. 1981. *Littérature juridique et évolution du malékisme en Mauritanie*. Tunis: Université de Tunis.

Ould Cheikh, Abdel Wedoud. 1989. *Eléments d'histoire de la Mauritanie*. Nouakchott: Centre Culturel Français.

——. 1995. "La Mauritanie: Un pays qui descend?". Bariou a.o. (1995:22–35).

——. 1998. "Cherche élite, désespérément: Évolution du système éducatif et (dé)formation des 'élites' dans la société mauritanienne". Bonte and Claudot-Hawad (1998:235–252).

Ould Ebnou, Moussa (ed.). N.d.(a). *al-Ḥikāyāt wa-l-ʾasāṭīr aš-šaʿbiyya al-mūrītāniyya*. 2 vols. Nouakchott: Commission Nationale pour la Collecte et la Publication de la Culture Populaire.
——. (ed.). N.d.(b). *al-ʾAmṯāl wa-l-ḥikāya aš-šaʿbiyya al-mūrītāniyya*. Nouakchott: Commission Nationale pour la Collecte et la Publication de la Culture Populaire.
Ould Mohamed Baba, Ahmad-Salem. 2000–2001. "Cuentos de los Wlād Däymān: Sutil socarronería". *Estudios de Dialectología Norteafricana y Andalusí* 5.253–264.
Ould Zenagui, Cheikh El-Bou. 1994. *La poésie de Mohamed Ould Adebba*. Nouakchott: Institut Mauritanien de Recherche Scientifique (Projet TOTEM).
Rebstock, Ulrich. 2001. *Maurische Literaturgeschichte*. Wiesbaden: O. Harrassowitz.
Sīdi Brāhīm (ibn), Muḥammadun. 1992. *Tahḏīb al-ʾafkār fī ʾadab aš-šiʿr al-ḥassānī*. Nouakchott: Imprimerie Nouvelle.
Taine-Cheikh, Catherine. 1985. "Le pilier et la corde: Recherches sur la poésie maure". *Bulletin of the School of Oriental and African Studies* 48.516–535.
——. 1990. "La Mauritanie en noir et blanc: Petite promenade linguistique en *ḥassāniyya*". Baduel (1990:90–105).
——. 1994. "Pouvoir de la poésie et poésie du pouvoir: Le cas de la société maure". *Matériaux Arabes et Sudarabiques* 6.281–310.
——. 1995. "Quand les bergers maures se lancent des 'colles'". *Littérature orale arabo-berbère* 22–23. 173–204.
——. 1998. "Langues, savoirs et pouvoirs en milieu maure". Bonte and Claudot-Hawad (1998:235–252).
——. 2002. "De la variation linguistique dans le prêche populaire mauritanien". *Language contact and language conflict in Arabic: Variations on a sociolinguistic theme*, ed. Aleya Rouchdy, 177–203. London: Routledge Curzon.
——. 2004. "La Mauritanie: Vers une nouvelle politique linguistique?". *Revue d'Aménagement Linguistique ("Aménagement linguistique au Maghreb")* 107.205–226.
Tauzin, Aline. 1982. "Autour de la poésie amoureuse des Maures de la Mauritanie et du Mali". *Littérature orale arabo-berbère* 13.129–146.
——. 1990. "À haute voix: Poésie féminine contemporaine en Mauritanie". Baduel (1990:178–187).
——. 1993. *Contes arabes de Mauritanie*. Paris: Karthala.

Catherine Taine-Cheikh
(LACITO – UMR 7017 of the CNRS)

Mawḍūʿ

1. The meaning of the term *mawḍūʿ*

The term *mawḍūʿ* is used in two senses:

i. When derived from *waḍʿ*, which serves in the literature translated from Greek as the equivalent of the verb *títhēmi* 'to place', *mawḍūʿ* refers to 'mutual agreement, convention' (Pollak 1913:58; Georr 1948:249; Afnan 1969:315–316). *Mawḍūʿ* is used in this sense by az-Zajjājī (d. 339/949): "A noun is an invented sound with a conventional meaning, not specified by time" (*al-ism ṣawt mawḍūʿ dāll bi-ttifāq ʿalā maʿnan ġayr maqrūn bi-zamān*; Versteegh 1995:58). This term refers to the prevailing view in Arabic linguistics according to which the meaning of words can be derived from the conventions of human society rather than from nature (Talmon 1990:270; Versteegh 1995:58). This is a reaction to the general problem discussed in Greek philosophy, whether things are by convention (*nómōi/thései*) or nature (*phúsei*). In connection with the meaning of words, the question was raised by Plato in his *Kratylos*. The Arabs accepted Aristotle's view (*De interpretatione* 16 a 19–20) about the conventional origin of language (Versteegh 1977:139). The same term *mawḍūʿ* also served for the translation of *hupóthesis*, for instance, in Ibn aṭ-Ṭayyib, "Are [the species] found [in nature] or are they just posited in the imagination?" ([*al-ʾanwāʿ*] *hal hiya mawjūda ʾam ʾinnamā hiya mawḍūʿa fī l-ʾawhām*; *Tafsīr* 55), which refers to the axiomatic principles of sciences (Maróth 1994:129–153).
ii. In the Arabic translation literature, the Greek *(hupó)keimai* (Latin *(sub)iaceo/sterno*) 'to lie under' is rendered by *waḍʿ* as well. In this case, the word has either a metaphysical or a linguistic sense. In metaphysics, *hupokeímenon* corresponds to the Latin *substratum* and refers in philosophical terminology to the bearer of qualities; in linguistics, accordingly, it refers to the subject of a proposition (Pollak 1913:58; Georr 1948:249; Afnan 1969: 315–316; Versteegh 1997:123). Summing up the views of al-Fārābī and Ibn Sīnā, al-Ġazzālī (d. 505/1111) states that a categorical proposition (*qaḍiyya ḥamliyya*) consists of two parts, e.g. 'the world is created'

> (*al-ʿālam ḥādiṯ*). In this sentence, *al-ʿālam* is the 'underlying concept' or 'the bearer of the quality' (*al-mawḍūʿ*) of which the quality of 'being created' is reported (*muxbar ʿanhu*; al-Ġazzālī, *Maqāṣid* 54). In another passage, while expounding his own views, al-Ġazzālī says that a sentence consists of a 'report' (→ *xabar*) and 'that which is reported' (*muxbar ʿanhu*). "It is the habit of logicians (*al-manṭiqiyyūn*) to refer to the word 'reported of' by the term *mawḍūʿ*" (al-Ġazzālī, *Manṭiq* 110). These words prove that the term corresponds to the Latin term *subiectum* in medieval scholastic logic (about *mawḍūʿ* and its meaning in al-Fārābī's *ʾIḥṣāʾ al-ʿulūm*, see Versteegh 1997:87, 123).

The Andalusian Arab grammarian ʾAbū ʿAbdallāh al-Baṭalyūsī (d. 521/1127) writes that the more or less philosophical term *mawḍūʿ* corresponds to *mubtadaʾ* in linguistic terminology (Elamrani-Jamal 1983:137–141, 182–183, 1979:76–89; Endress 1986a:203; Versteegh 1993:138; → *ibtidāʾ*). ʾAbū Ḥayyān at-Tawḥīdī transmits a discussion between the logician ʾAbū Bišr Mattā ibn Yūnus and the grammarian ʾAbū Saʿīd as-Sīrāfī which took place in 320/932. The latter blamed the logicians for using Arab words in unknown meanings and for creating a new language within the traditional Arabic lexicon. One of his examples was the word *mawḍūʿ* (Tawḥīdī, *ʾImtāʿ* 107–108, esp. 122; Endress 1986b:260; Versteegh 1977:139). Accordingly, this term should be regarded as an innovation in the language of philosophy.

Thus, in philosophical terminology, the word *mawḍūʿ* denotes the subject of a judgment (*ḥukm*) in a categorical proposition, and in the linguistic tradition it corresponds to *mubtadaʾ* as used by Arab grammarians referring to the topic or grammatical subject of a sentence.

2. *MAWḌŪʿ* AND *MUBTADAʾ*

Although the logical term *mawḍūʿ* and the linguistic term *mubtadaʾ* share the meaning of 'subject', the connotation of the two terms is not exactly the same. All grammarians recognized two basic types of sentences: verbal (*jumla fiʿliyya*) and nominal (*jumla ismiyya*). These terms indicate that the sentence begins with a verb or a noun (Versteegh 1997:49), whereas in the philosophical tradition the basis of division was whether the predicate was a verb or a noun (→ *maḥmūl*). This implies that grammatical analysis was more concerned with formal criteria, whereas the philosophical tradition concentrated more on material aspects.

In sentences with a verbal predicate, the noun was necessarily the subject in every case, because verbs were necessarily analyzed as predicates. The situation was more complicated in sentences with a nominal predicate. Nouns refer to entities having qualities or distinctive features. Explained in terms of Aristotle's logic, subject always fell under the category of substance, physical or intellectual, of which the other nine categories were predicated. However, it always depends on the context which word can be explained as substance in a given sentence. In the sentences 'a human being is mortal' and 'Zayd is a human being', the same concept may be analyzed either as subject or as predicate. In the case of an inductive enumeration, even sentences like 'a human being is Zayd' admitted in logical analysis.

In the sentence 'Zayd is white', however, 'Zayd' is substance, and 'white' is a color falling under the category of quality. Nevertheless, 'whiteness is a color' complies with the above criteria, 'whiteness' being regarded as a substance. All of this means that the subject and predicate of a proposition were interchangeable in logic, if the nature and meaning of the concepts permitted so.

Some texts suggest that this was not the case in linguistics. As al-Baṭalyūsī states, there were discussions among grammarians concerning the sentence *šarru n-nisāʾi al-baḥātiru* 'the most malevolent women are those who are stout and short'. When analyzing this sentence, some held the view that *al-baḥātiru* is *mubtadaʾ*, consequently, *šarru n-nisāʾi* is predicate, while others held the opposite view. The author says that only the first option can be admitted (Elamrani-Jamal 1983:182).

The discussion quoted by al-Baṭalyūsī indicates that in grammar, in contradistinction to logic, one term was analyzed as natural subject while the other was analyzed as natural predicate within a sentence. In linguistic analysis, subject and predicate do not seem to be interchangeable. The different attitudes of grammarians and logicians point to the difference between the meaning of *mawḍūʿ* in philosophy and *mubtadaʾ* in grammar.

Referring to Aristotle, Ibn Zurʿa (d. 398/1008) states that the *mawḍūʿ* is always and necessarily a noun (Ibn Zurʿa, *Manṭiq* 35). This view relies on Aristotle's metaphysical doctrine, as indicated above, because nouns denote substances, and substances are bearers of the other nine categories being regarded as predicates.

A further theoretical difference concerned the fact that the definition of *mubtadaʾ* included an indication of grammatical government (Versteegh 1977:73–74). Its definition traditionally consisted of formal elements (Ibn Mālik, *ʾAlfiyya*, chapter on *ibtidāʾ*): "[*Mubtadaʾ* is] a noun devoid of overt governors without additions" (*ism mujarrad ʿan al-ʿawāmil al-lafḏiyya ġayr al-mazīda*), whereas this was not an official requirement in the case of *mawḍūʿ* (Versteegh 1997:87).

Ibn Sīnā (d. 428/1037) says that *mawḍūʿ* can be universal (*kullī*), particular (*juzʾī*), individual (*šaxṣī*), and devoid of quantification (*muhmal*). The sentence *zayd yamšī* 'Zayd walks' refers to an individual subject (*zayd*), but 'walks' (*yamšī*) is 'of more general occurrence' than *zayd*, but if his essence (*ḏāt*) is described (*mawṣūf*) by an attribute (*ṣifa*), it is universal (Ibn Sīnā, *Šifāʾ* 25). The connection of two equally individual terms can only express identity: *zayd huwa ʾabū l-qāsim* 'Zayd is ʾAbū l-Qāsim' (Ibn Sīnā, *Šifāʾ* 21). As Ibn Sīnā states, one can speak here of subject and predicate only from the point of view of linguistics, but this does not correspond to the nature of things.

All these considerations indicate that logic was always concerned with the nature of things and examined the linguistic expressions from the point of view of whether they adequately express the natural disposition of things, whereas linguistics analyzed mainly the form of sentences and the role that concepts play within the sentence.

The terminology in the passage quoted further proves that *mawḍūʿ* roughly corresponds to *mawṣūf* in *kalām* terminology (Ibn Sīnā, *Šifāʾ* 18–20).

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Aristotle, *De Interpretatione* = Aristoteles, *Perì hermēneías*. Ed. L. Minio-Paluello, *Aristotelis Categoriae et Liber de interpretatione*, 47–72. Oxford: Clarendon Press, 1949.

Fārābī, *ʾIḥṣāʾ* = ʾAbū Naṣr Muḥammad ibn Muḥammad al-Fārābī, *ʾIḥṣāʾ al-ʿulūm*. Ed. and trans. Ángel Gonzalez Palencia. 2nd ed. Madrid: Consejo Superior de Investigaciones Científicas, 1953.

Ġazzālī, *Manṭiq* = ʾAbū Ḥāmid Muḥammad ibn Muḥammad al-Ġazzālī, *Manṭiq tahāfut al-falāsifa al-musammā Miʿyār al-ʿilm*. Ed. Sulaymān Dunyā. Cairo, 1961.

——, *Maqāṣid* = ʾAbū Ḥāmid Muḥammad ibn Muḥammad al-Ġazzālī, *Maqāṣid al-falāsifa*. Ed. Sulaymān Dunyā. Cairo, 1960.

Ibn Mālik, *ʾAlfiyya* = Jamāl ad-Dīn ʾAbū ʿAbdallāh Muḥammad ibn ʿAbdallāh Ibn Mālik al-ʾAndalusī, *Matn al-ʾAlfiyya*. Cairo: Maṭbaʿa ʿĪsā al-Bābī al-Ḥalabī, n.d.

Ibn Sīnā, *Šifāʾ* = ʾAbū ʿAlī al-Ḥusayn ibn ʿAbdallāh Ibn Sīnā, *Kitāb aš-šifāʾ: al-Manṭiq, al-Maqūlāt*. Ed. ʾIbrāhīm Madkūr. Cairo, 1959.

Ibn aṭ-Ṭayyib, *Tafsīr* = ʾAbū l-Faraj ʿAbdallāh Ibn aṭ-Ṭayyib, *Tafsīr Kitāb ʾIsāġūjī li-Furfuriyūs*. Ed. Kwame Gyekye, *Ibn al-Ṭayyib's Commentary on Porphyry's* Eisagoge. Beirut: Dār al-Mašriq, 1986.

Ibn Zurʿa, *Manṭiq* = ʾAbū ʿAlī ʿĪsā ibn ʾIsḥāq Ibn Zurʿa, *Manṭiq ibn Zurʿa*. Ed. Gérard Jehamy and Rafīq al-ʿAjam. Beirut, 1994.

Tawḥīdī, *ʾImtāʿ* = ʾAbū Ḥayyān ʿAlī ibn Muḥammad at-Tawḥīdī, *Kitāb al-ʾimtāʿ wa-l-muʾānasa*. Ed. ʾAḥmad ʾAmīn and ʾAḥmad az-Zayn. Beirut, n.d.

SECONDARY SOURCES

Afnan, Soheil M. 1969. *A philosophical lexicon in Persian and Arabic*. Beirut: Dar El Mashreq.

Elamrani-Jamal, Abdelali. 1979. "Les rapports de la logique et de la grammaire d'après le *Kitāb al-Masāʾil* d'al-Baṭalyūsī". *Arabica* 26.76–89.

——. 1983. *Logique aristotélicienne et grammaire arabe: Etude et documents*. Paris: J. Vrin.

Endress, Gerhard. 1986a. "Grammatik und Logik: Arabische Philologie und girechische Philosophie im Widerstreit". *Sprachphilosophie in Antike und Mittelalter*, ed. Burkhard Mojsisch, 163–233. Amsterdam: B.R. Grüner.

——. 1986b. "Das Streitgespräch zwischen Abū Saʿīd as-Sīrāfī und Abū Bišr Mattā". *Sprachphilosophie in Antike und Mittelalter*, ed. Burkhard Mojsisch, 235–270. Amsterdam: B.R. Grüner.

Georr, Khalil. 1948. *Les* Catégories *d'Aristote dans leurs versions syro-Arabes*. Beirut: Institut Français de Damas.

Maróth, Miklos. 1994. *Die Araber und die antike Wissenschaftstheorie*. Leiden: E.J. Brill.

Pollak, Isidor. 1913. *Die Hermeneutik des Aristoteles in der arabischen Übersetzung des Ishaq ibn Honain*. Leipzig: Abhandlungen für die Kunde des Morgenlandes.

Talmon, Rafael. 1990. "The philosophizing Farrāʾ: An interpretation of an obscure saying attributed to the grammarian Ṯaʿlab". *Studies in the history of Arabic grammar*, II, ed. Michael G. Carter and Kees Versteegh, 265–279. Amsterdam and Philadelphia: J. Benjamins.

Versteegh, Kees. 1977. *Greek elements in Arabic linguistic thinking*. Leiden: E.J. Brill.

——. 1993. *Arabic grammar and Qurʾānic exegesis in early Islam*. Leiden: E.J. Brill.

——. 1995. *The explanation of linguistic causes: Az-Zaǧǧāǧī's theory of grammar*. Amsterdam and Philadelphia: J. Benjamins.

——. 1997. *The Arabic linguistic tradition*. London and New York: Routledge.

MIKLÓS MARÓTH
(Pázmány Péter Catolic University)

Meccan Arabic

1. General

Meccan Arabic refers to the variety of Arabic spoken mainly in the holy city of Mecca, Saudi Arabia. In addition, it is widely used and understood in both Jeddah and Madinah, and less so in Ta'if. While the majority of the native speakers of the dialect live in Mecca and its suburbs, it is difficult to arrive at an accurate number of the speakers due to the extremely cosmopolitan nature of the population of Mecca and its constant settlement by speakers of other dialects. The annual pilgrimage season and the continuous religious visits paid by Muslims bring different cultures and languages to the city. These have given Mecca its unique multicultural structure and left their influence on the vocabulary as well.

Meccan Arabic is basically a spoken language used in informal situations. Its use in recording literature depicting the area and its culture is restricted to a small number of books, pamphlets, and poems published sporadically in local newspapers.

Historically, Meccan Arabic belongs to the West Arabian (Ḥijāzī) group of dialects; however, it is substantially different from the Bedouin dialects of the Ḥijāz and Tihāma, and displays several characteristics of the sedentary dialects. It bears close linguistic affinity to the dialects spoken in Egypt, Sudan, and parts of the Levant (Ingham 1971). However, the structure of Meccan Arabic shows most of the features shared by the Arabic dialects in general. These include the changes that took place in the phonological system of these dialects and the reduction of morphological categories (Versteegh 2001:99–100).

Earlier studies of Meccan Arabic include Snouck Hurgronje's (1886) collection of Meccan proverbs and his (1888) description of life in Mecca in the last part of the 19th century. Satterthwait (1960) deals with the rate of morphemic decay in Meccan Arabic in relation to Classical Arabic.

The first serious description of the phonology, morphology, and syntax of the dialect is found in Schreiber (1971). This book also includes a glossary and the translation of selected texts into German. Ingham (1971) is a brief description of the phonology and morphology of the verb. Meccan Arabic has also been studied as part of the so-called Urban Ḥijāzī Arabic. Thus, Sieny (1978) is a syntactic description of the dialect conducted in the Tagmemic framework. Al-Sasi (1972) includes some Meccan proverbs and texts translated into German.

Bakalla (1973, 1979) present the first detailed transformational generative analysis of the morphology and phonology of the verb in Meccan. Abu-Mansour (1987) offers an extensive nonlinear analysis of syllable structure and syllable structure-related processes, while Kabrah (2004) is the most comprehensive optimality-theoretic account of syllable structure and stress phenomena. Word order in Meccan Arabic is

Table 1. Inventory of Consonants

	bilabial	labio-dental	alveolar	alveo-palatal	palatal	velar	uvular	pharyngeal	laryngeal
plosive									
voiceless, voiced	b		t, d			k, g			ʾ
emphatic			ṭ, ḍ						
nasal	m		n						
fricatives									
voiceless, voiced		f	s, z	š			x, ġ	ḥ, ʿ	h
emphatic			ṣ, ẓ						
affricates									
voiceless, voiced				j					
trill			r						
emphatic			ṛ						
lateral			l						
emphatic			ḷ						
glides	w				y				

described extensively in Abu-Mansour (1982). A glossary of Meccan terms can be found in Natto (1997).

2. Linguistic description

The following is a description of the major phonological, morphological, and syntactic properties of the dialect.

2.1 *Phonology*

2.1.1 Consonants

The two phonemes /x, ġ/ are uvular sounds, while /j, š, y/ are palatal. The only minimal pair that contrasts velarized /ḷ/ with clear /l/ is /waḷḷa/ ‘by God’ and /walla/ ‘or else’ (Bakalla 1979:460). Other occurrences of [ɫ] result from emphasis spread. Recent studies show that /h/ is voiced: *ʾabkam* > [ʔapkam] ‘mute’, but *ʾazhār* > [ʔazhaːr] ‘flowers’ (Abu-Mansour 1996:217).

Generally, interdentals are reflected by plosives in Meccan, e.g. *dahab* ‘gold’, *talāta* ‘three’, and by sibilants in loans from Standard Arabic, thus *zanb* ‘sin’ and *surayya* ‘chandelier’.

Meccan has four emphatic consonants, /ṭ, ḍ, ṣ, ẓ/; /r/ and /x/ behave like the emphatics in certain contexts. Standard Arabic interdental /ḏ̣/ is realized as the dental emphatic fricative /ẓ/, in a large number of lexical items, e.g. *ẓālim* ‘unjust’ and *ẓarīf* ‘charming’ for Standard Arabic *ḏ̣ālim*, and *ḏ̣arīf*. A few words have /ḍ/ instead: *ḍahar* ‘back’, *ḍalām* ‘darkness’.

Standard Arabic /q/ is reflected by /g/ in Meccan, e.g. *gamar* ‘moon’, except in loans from Standard Arabic such as *al-qurʾān* ‘the *Qurʾān*’ and *al-qāhira* ‘Cairo’.

2.1.2 Vowels

Table 2. Inventory of vowels

Short vowels			Long vowels		
i		u	ī		ū
			ē		ō
	a			ā	

The mid vowels /ā/ and /ō/ result from the historical change of **ay* and **aw*: *bēt* < **bayt* ‘house’, *mōt* < **mawt* ‘death’. The old diphthongs are preserved in the realization of morphological patterns: *aysar* ‘easier’ (aCCaC), mawwat ‘to cause to die’ (CvC_iC_ivC).

Nouns that end in CvCāʾ in Standard Arabic, e.g. *šitāʾ* ‘winter’, *ḥamrāʾ* ‘red’, are realized as *šita* and *ḥamra* in Meccan. Long vowels can only occur in open syllables except in final position, thus *fānūs, bēt*, but *fānūsaha* ‘her lantern’, *bētana* ‘our house’. Long /ā/ occurs in closed syllables in medial position only in the active participle CāCiCa pattern: *kātba* < *kātiba* ‘female writer’. Long *ī, ā, ō*, and *ū* in final position represent the pronominal subject and object: *katab+na* ‘we wrote’ vs. *katab+nā* ‘we wrote it [sg. masc.]’.

Vowels are strongly affected by adjacent consonants (Ingham 1971:275–276). First, /i/ and /a/ are both fronted and higher in the environment of /j/ and /y/, thus, [ˈʔa̠dʒi̠] < *aji*. Second, all vowels have a more open pronunciation in the environment of pharyngeals, with an off-glide in the case of tense vowels: [ˈjɛħlɪb] < *yiḥlib* ‘he milks’, [ˈʕæli̠] < *ʿali* ‘Ali’, and [riːᵊħ] < *rīḥ* ‘wind’. Third, an advanced realization of the vowels /u/ and /uː/ and a centralized realization of /a/ and /aː/ are found in the environment of plain consonants: [ˈgu̟ːdu̟] < *gūdu* ‘guide him, [ˈgʌtʌl] < *gatal* ‘he killed’. Fourth, emphatics are associated with a more retracted quality in the vowels with an on-glide accompanying tense vowels: [ˈɫɑbʕæn] < *ṭabʿan* ‘naturally’, [ɫu̠ːl] < *ṭūl* ‘length’.

2.1.3 Syllable

The possible phonological syllable types include Cv, Cv̄, CvC, Cv̄C, and CvCC. The last two occur only in word-final position. The consonant clusters in the CvCC class of nouns must follow the sonority hierarchy principle; otherwise, a vowel, /a/, /i/, or /u/, is inserted (Abu-Mansour 1987): *ẓarf* ‘envelope’, *bint* ‘girl’, but *guṭn* > *guṭun* ‘cotton’, *baḥr* > *baḥar* ‘sea’, *samn* > *samin* ‘ghee’.

2.1.4 Stress

The stress generalizations in Meccan are as follows (Kabrah 2004). First, in two-syllable words, stress lodges on the rightmost heavy syllable if there is one; otherwise, the penult is stressed: *nadḗt* ‘I called’, *katábt* ‘I wrote’, *kátab* ‘he wrote’, *xā́la* ‘maternal aunt’. Second, in three-syllable words, stress falls on the rightmost heavy syllable, else on the antepenult: *manādī́l* ‘handkerchiefs’, *ragabátha* ‘her neck’, but *sámaka* ‘fish’. Third, in four-syllable words, the rightmost heavy syllable within the last three syllables is stressed; otherwise, the penultimate syllable bears the stress: *mudarrisī́n* ‘teachers [masc.]’, *astalámna* ‘we received’, *kitā́bana* ‘our book’, *burtukā́na* ‘orange’, *ragabátu* ‘his neck’.

The feminine subject marker *-at* receives stress only before suffixes. Compare *ramátu* ‘she threw it [sg. masc.]’ and *šālátha* ‘she carried it [sg. fem.]’ to *rámat* ‘she threw’.

A light syllable is stressed despite the presence of a heavy syllable only in words that involve High Vowel Deletion and Epenthesis in initial position: *ṣāḥibátu* > *ṣāḥbátu* ‘his friend’, *ḥtáram* > *aḥtáram* ‘he respected’.

2.2 *Phonotactics*

Emphasis spread: The true emphatic consonants /ṭ, ḍ, ṣ, ẓ/ spread pharyngealization to their plain counterparts. /x/ spreads emphasis to /s/ only (Kabrah 2004), e.g. *tarāḍī* [t̴ɑ'r̴ɑːd̴i] ‘agreement’, *zēṭa* ['zeːt̴ɑ] ‘mess’, *ʾatwassax* [ʔɑt̴'wɑs̴s̴ɑx] ‘it became dirty’. Rightward spreading is weaker, as in *baṣmati* ['bɑs̴mati] ‘my fingerprint’. Emphasis spread is blocked by /i/ and the pharyngeal /ḥ/; however, a great deal of variation is involved (Abu-Mansour, forthcoming): *wisix* [wisix] ~ [wis̴ix] ‘dirty’, *tiḥtaḍir* [t̴iħt̴ɑd̴ir] ~ [tiħt̴ɑd̴ir] ‘she is dying’.

Voice assimilation: Medial and final biconsonantal clusters can be voiced, voiceless, or voiceless-voiced, e.g. *kibda* ‘liver’, *maska* ‘a catch’, and *ʾakbar* ‘older’. Voiced-voiceless clusters do not occur unless the first member of the cluster is a sonorant or a guttural sound, thus, *mabsam* ‘mouth’ and *adfa* ‘warmer’ are pronounced as [mapsam] and [ʔatfa], but *ʾams* [ʔams] ‘yesterday’ and *baʿtara* [baʕtara] ‘scattering’ (Abu-Mansour 1996). The voiceless gutturals /x/ and /ḥ/ trigger devoicing in a preceding consonant, as in *midxana* > [mitxana] ‘chimney’ and *ʾabḥar* > [ʔapħar] ‘he sailed’. The presence of voiceless-voiced clusters distinguishes Meccan from other dialects, such as Daragözü, Maltese, and Sudanese Arabic.

The prefix of the reflexive passive *at-* assimilates its /t/ to alveolar stops and sibilants but not to sonorants, velars, uvulars, or pharyngeals. Compare *addaffa* ‘he became warm’ to *atlattam* ‘he was masked’, *atgammar* ‘it got roasted’.

2.3 *Morphophonology*

Epenthesis: Medial Cv̄C and CvCC syllables require the insertion of a vowel /a/ when followed by consonant-initial suffixes. The final consonant of the CvCC or Cv̄C is always syllabified as an onset to the epenthetic vowel: *katabt+ha* > *katabtaha* ‘I wrote it [fem.]’, *kitābha* > *kitābaha* ‘her book’.

Vowel shortening: The vowels of hollow verbs are shortened in closed syllables before the clitics, *-l* ‘to/for’ and *-b* ‘with’. Compare *sībahum* < *sīb+hum* and *siblahum* < *sīb+l+hum* ‘leave for them’. Long vowels of the interrogative particles also shorten: *ʾišbaha* < *ʾēš-b-ha* ‘what is wrong with her?’. The vowel of a closed syllable does not shorten when followed by a pronominal object suffix only, instead, a vowel is inserted: *šāf+ha* > *šāfaha* ‘he saw her’ and *fēn+ha* > *fēnaha* ‘where is she?’.

Vowel deletion: Meccan is a differential dialect. Vowel deletion is restricted to the deletion of unstressed short high vowels from open syllables, provided that deletion will not result in medial CvCC (Abu-Mansour 1987): *kíbir+u* > *kíbru* ‘they grew old’, but *yidárrisu* < *yi-dár-ris-u* ‘they teach’. Cv̄C syllables are allowed as a result of high vowel deletion, e.g. *kātib+a* > *kātba* ‘a female writer’.

Gemination of the applicatives: Meccan is one of the dialects of Arabic that exhibit gemination of the two clitics *-l-* and *-b-* (Abu-Mansour 1987; Kabrah 2004). Obligatory gemination occurs when the clitic follows the subject morpheme *-t* and is followed by a consonant-initial object pronoun suffix with all types of roots, e.g. *ka.tab.tál.la.ha* ‘I/you wrote for/to her’, *ra.mē.táb.ba.ha* ‘I/you threw with it [fem.]’ not **ka.táb.ta.la.ha*,* *ra.mē.ta.ba.ha*. It also occurs in geminate consonant roots when the clitic follows the 3rd person singular subject pronoun realized as zero morpheme and followed by consonant-initial object pronoun, thus *rad.dál.la.kum* not **rad.dá.la.kum* ‘he returned to you [pl.]’. Optional gemination occurs when these clitics follow the subject morpheme *-t* and are followed by a vowel-initial object suffix: *ka.tab.tál.lu* ~ *ka.táb.ta.lu* ‘I wrote to him’, *mad.dē.tál.li* ~ *mad.dḗ.ta.li* ‘you stretched for me”, *gúl.ta.li* ~ *gul.tál.li* ‘you told me’.

2.4 *Morphology*

Meccan Arabic distinguishes between masculine and feminine only in the 2nd and 3rd person singular. Verbal phrases tend to be synthetic and may include several suffixes: *mā-kān-*

u-b-yi-n-ḍarb-u ‘they were not being beaten up’. Noun phrases, on the other hand, show analytical tendencies, such as the use of the genitive particle *ḥagg* instead of the construct structure.

2.4.1 Pronouns

2.4.1.1 Independent personal pronouns (Table 3)

Table 3. Independent personal pronouns

	3rd	2nd	1st
sg. masc.	*huwwa*	*inta*	*ana*
sg. fem.	*hiyya*	*inti*	
pl.	*humma*	*intu*	*iḥna~niḥna*

2.4.1.2 Possessive/object pronouns

Possessive and object pronouns are essentially the same with a few differences. They exhibit three forms depending on the final sound of the word. (Table 4)

Table 4. Possessive/object pronouns

after -v (e.g. *abu*)			after -v̄C (e.g. *kitāb*)		
-(h)	*-k*	*-ya*	*-u*	*-ak*	*-i*
-ha	*-ki*		*-aha*	*-ik*	
-hum	*-kum*	*-na*	*-ahum*	*-akum*	*-ana*

after -vCC (e.g. *bint*)

-u	*-ak*	*-i*
-aha	*-ik*	
-ahum	*-akum*	*-ana*

abū(h) ‘his father’, *kitābu* ‘his book’, *bintu* ‘his daughter’

2.4.1.3 Indirect object suffixes (Table 5)

Table 5. Indirect object suffixes

after -v

-lu	*-lak*	*-li*
-laha	*-lik*	
-lahum	*-lakum*	*-lana*

katabōlu ‘they wrote to him’

after -C

-lu	*-lak*	*-li*
-laha	*-lik*	
-lahum	*-lakum*	*-lana*

after -CC

-allu	*-allak*	*-alli*
-allaha	*-allik*	
-allahum	*-allakum*	*-allana*

katablu ‘he wrote to him’, *katabatlu* ‘she wrote to him’, *katabtallu* ‘I wrote to him’

2.4.1.4 Demonstratives (Table 6)

Table 6. Demonstratives

	close proximity	distant proximity
sg. masc.	*hāda*	*hadāk(a)*
sg. fem.	*hādi*	*hadīk(a)*
pl.	*hadōl*	*hadōlāk(a)*

These forms can be used attributively and might drop *hā-* in informal situations.

2.4.1.5 Relative pronouns

The relative pronoun *ʾilli* ‘who [masc./fem.], that, which, those’ is invariable and introduces either a verbal or a nominal relative clause.

2.4.1.6 Interrogative pronouns

The main interrogative particles are *mīn* ‘who?’, *ēš* ‘what?’, *ʾayyi* ‘which one?’. They can be used in pre-/post-verbal position: *mīn tiḥubb?* or *tiḥubb mīn?* ‘whom do you love?’. The particle *ʾayyi* must be used in a prenominal position, e.g. *ʾayyi kitāb tibġa* ‘which book do you want?’.

2.4.2 Adverbs

Temporal: *mita* ‘when’, *ʾams* ‘yesterday’, *ʾalyōm* ‘today’, *bukrah* ‘tomorrow’, *baʿad bukrah* ‘after tomorrow’, *sāʿāt ~ ʾaḥyānan* ‘sometimes’, *dāyman* ‘always’, *baʿdēn* ‘later’, and *alʿām* ‘last year’, e.g. *mātat alʿām* ‘she died last year’.

Place: *fēn* ‘where’, *hina* ‘here’, *hināk* ‘there’, *guddām* ‘in front of’, *wara* ‘behind’. Manner adverbs include *kēf* ‘how’, *zayy kida* ‘like this’, and *gawām* ‘quickly’, e.g. *taʿāl gawām* ‘come quickly!’ and may be formed by using the preposition *bi-* ‘by’: *biguwwah* ‘by force’, *bišwēš* ‘gently’.

2.4.3 Particles

2.4.3.1 Article

The definite article *al-* is prefixed to certain proper and common nouns, such as *annās* ‘the people’ and *alʾurdun* ‘Jordan’.

2.4.3.2 Genitive

The genitive marker is *ḥagg* [masc.], *ḥaggat* [fem.], *ḥaggōn* [pl.]; it can be used interchangeably with the construct structure except with inalienable nouns, thus *kitāb albint* or *alkitāb ḥagg albint*, but *rāsi* 'my head' and *ˀabūha* 'her father'.

2.4.3.3 Negation

There are four negation particles; *mā* 'be not' is used with verbs and pronouns, e.g. *māhum hina* 'they are not here'. Other parts of speech are negated by *mū*, a variant of *mā*, e.g. *mū kabīr* 'not big', *mū laha* 'not for her'. The particle *lā* 'do not' is used with negative commands or requests, e.g. *lā tinām* 'do not sleep'. *ˀiṣḥa* 'let…not' implies both warning and threats, e.g. *ˀiṣḥa tinsa* 'do not forget!'

2.4.3.4 Prepositions

Prepositions are followed by nouns, e.g. *fī makka* 'in Mecca', or suffixed pronouns, e.g. *minnaha* 'from her', *fīhum* 'in them'.

2.4.3.5 Conjunctions

Coordinators include *wu* ~ *w* ~ *u* 'and', *ˀaw* 'or' and *ˀamma…walla*, e.g. *ˀamma daḥḥīn walla baˁdēn* 'either now or later'. Subordinating conjunctions express time, place, manner, conditional, and purpose: *lamman* 'when', *ˀilēn* 'until', *yōmma* 'as soon as', *maḥalma* 'wherever', *zayy* 'as', *law* 'if', *lōla* 'had it not been', *ˁašān* 'because', *walawin* 'even though', *madām* 'as long as', and *aḥsan* ~ *laḥsan* 'lest that', e.g. *rūḥ aḥsan ˀaḍrubak* 'go!, otherwise, I will hit you'.

2.4.4 Nouns

2.4.4.1 Gender

Nouns without the feminine marker *-a* include names of parts of the body, places, and nouns that denote females: *yad* 'hand', *maṣur* 'Egypt', *šams* 'sun'. The gender of the noun governs the gender inflection of verbs, adjectives, and pronouns: *yad karīma(h)* 'a generous hand'.

2.4.4.2 Productive patterns

The majority of nouns are derived from verbs, adjectives, and other nouns: *katab/kitāba, baṭal/buṭūla* 'hero/heroism'. muCCāC, CaCCāCa are used for instruments, *muftāḥ* 'key', *wallāˁa* 'lightener'; maCCaC(a), maCCiC for location: *madrasa* 'school', *masjid* 'mosque'; and CaCCāC for occupation: *sabbāk* 'plumber'. The *nisba* suffix *-i* and the Turkish suffix *-ji* are also used: *kahrabāˀi* 'electrician', *gahwaji* 'coffee seller'.

2.4.4.3 Dual

The dual is marked by the morpheme *-ēn* added to a masculine noun: *maktabēn* 'two offices'. For phonological reasons, final *-t* of the feminine marker appears before the dual ending : *warda/wardatēn* 'a rose/two roses'. In the *nisba* nouns, and in nouns ending in *-ū*, the dual marker is *-yēn* and *-wēn: makki/makkiyēn* 'from Mecca/two Meccans', and *axu/axuwēn* 'a brother/two brothers'.

2.4.4.4 Diminutives

Diminutives are used for nicknames only. The most common patterns are *faˁˁūl(a), faˁlu: ḥasan > ḥassūn(a), suˁād > saˁdu* (Abu-Mansour 2000).

2.4.5 Numerals

1–2: *wāḥid* masc., *waḥdah* fem.: *walad wāḥid* 'one boy', *bint waḥdah* 'one girl'. In questions or negatives, *ˀaḥad* is used: *fī ˀaḥad?* 'anybody there?'. *itnēn* is invariant and used with dual or plural nouns to reinforce the dual meaning: *rijāl* (pl.)/*rijjālēn itnēn* 'two men'.

3–10: There is one form: *talāta, ˀarbaˁa, xamsa, sitta, sabˁa, tamanya, tisˁa, ˁašara*, e.g. *talāta rijāl*. When the number takes the definite article, the following noun has to be indefinite: *attamanyah banāt*.

11–19: The long forms keep the last syllable of *ˁašar* 'ten': *ˀiḥdaˁš/ˀiḥdaˁšar* 'eleven', *xamistaˁš/xamistaˁšar* 'fifteen'. Only the long forms can be used as a linked form and be followed by a singular noun: *xamistaˁšar bint* 'fifteen girls'.

100: *miyya, miyyatēn, tultumiyya, ˀurbuˁumiyya*, etc.

Ordinals from 2 to 10 follow the pattern CāCiC(a) and may be postposed: *ˁāšir bint* ~ *bint ˁāšra*. Irregular forms include *alˀawwal* (masc.), *alˀūla* (fem.), *alˀawāˀil* (pl.) 'first', *alˀāxir* (masc.), *alˀaxīra* (fem.), *alˀawāxir* (pl.) 'last'.

2.4.6 Verbs

2.4.6.1 Forms

Form I may have two morphological types only, CaCaC and CiCiC. The latter corresponds to both CaCiC and CaCuC in Standard Arabic (Table 7).

Table 7. Morphological types of Form I

Standard Arabic	Meccan Arabic	
kataba	*katab*	‘to write’
samiʿa	*simiʿ*	‘to hear’
kabura	*kibir*	‘to grow up’
ḥasuna	*ʾatḥassan*	‘to improve’

Form I is transitive in meaning and serves as the base of derivation for nine other forms.

Form II verbs are generally causative in meaning and can be derived from verbs, adjectives, or nouns, e.g. *wasīʿ* ‘wide’ > *wassaʿ* ‘to enlarge’, *xēma* ‘tent’ > *xayyam* ‘to put up a tent’. Some verbs denote intensity: *kasar* ‘to break’ > *kassar* ‘to smash’. The perfect is *kattab* and the imperfect *yikattib*.

Form III is usually reciprocal: *kātab-ni* ‘he corresponded with me’; a few are intransitive: *sāfar* ‘he traveled’. The perfect is *kātab*, the imperfect *yikātib*.

Form IV verbs have the prefix *a-* and are rare, e.g. *aʿṭa/yiʿṭi* ‘to give’, *aʿlan/yiʿlin* ‘to announce’. The causative meaning is expressed through the use of analytical expressions meaning ‘to make’: *xallā yiʿṭi* ‘he made him give’, or by Form II, cf. Standard Arabic *ʾajlas*, Meccan *jallas* ‘to make sit’.

Form V is derived by prefixing *at-* to Form II verbs. It expresses the reflexive meaning of Form II, *ʿallam* ‘to teach’ > *atʿallam* ‘to learn’, or the passive: *ṭannaš* > *aṭṭannaš* ‘to be ignored’.

Form VI is derived by prefixing *at-* to Form III verbs. They express reciprocity or pretense: *šāwar+na* ‘we consulted’ > *atšāwarna* ‘we consulted each other’, *mariḍ* > *atmāraḍ* ‘he pretended to be sick’.

Form VII verbs have replaced the internal passive of Standard Arabic: *ankatab addars* instead of *kutiba d-darsu* ‘the lesson was written’. *n-* is replaced by *t-* before /n/, /l/, /r/, /y/, and /w/, thus *atlasaʿ* ‘to get burned’, *atyassar* ‘to become easy’; before other consonants *n-* and *t-* alternate, *ankatab* ~ *atkatab, anḥabas* ~ *atḥabas* ‘to be detained’.

Form VIII verbs are derived from Form I by infixing *-t-* after the first radical. They are reflexive: *(a)htamm* ‘to become concerned’, *(a)ʿtaraf* ‘to confess’. This pattern is productive, although a few verbs have a passive meaning, *arṭabat* ‘to become obliged’.

Form IX is not productive and has been replaced by verbs of Form II. Thus, Meccan Arabic *ḥammar* and *bayyaḍ* are used instead of Standard Arabic *iḥmarra* ‘it turned red’ and *ibyaḍḍa* ‘it turned white’.

Form X verbs are common. They have the prefix *sta-* and denote the meaning of seeking for oneself: *ġafar* ‘to forgive’: *astaġfar* ‘to ask for forgiveness’.

2.4.6.2 Inflection of the verb

2.4.6.2.1 Imperfect

The base vowel can be /u/, /a/, or /i/. The prefix vowel is /i/, except for the 1st person singular (Table 8).

Table 8. Inflection of the imperfect

yiktub ‘he writes’

	3rd	2nd	1st
sg. masc.	*yiktub*	*tiktub*	*aktub*
sg. fem.	*tiktub*	*tiktubi*	
pl.	*yiktubu*	*tiktubu*	*niktub*

yišrab ‘he drinks’

	3rd	2nd	1st
sg. masc.	*yišrab*	*tišrab*	*ašrab*
sg. fem.	*tišrab*	*tišrabi*	
pl.	*yišrabu*	*tišrabu*	*nišrab*

yiʿrif ‘he knows’

	3rd	2nd	1st
sg. masc.	*yiʿrif*	*tiʿrif*	*aʿrif*
sg. fem.	*tiʿrif*	*tiʿrifi*	
pl.	*yiʿrifu*	*tiʿrifu*	*niʿrif*

The imperfect may express a modal meaning: *tijlis šuwayyah* ‘would you like to stay for a while?’.

2.4.6.2.2 Perfect

Verbs in the perfect tense are inflected by suffixes which show person, gender, and number (Table 9).

Table 9. Inflection of the perfect

katab 'he wrote'

	3rd	2nd	1st
sg. masc.	*katab*	*katabt*	*katabt*
sg. fem.	*katabat*	*katabti*	
pl.	*katabu*	*katabtu*	*katabna*

kibir 'he grew old'

	3rd	2nd	1st
sg. masc.	*kibir*	*kibirt*	*kibirt*
sg. fem.	*kibrat*	*kibirti*	
pl.	*kibru*	*kibirtu*	*kibirna*

2.4.6.3 Participles

The active and passive participles CāCiC, maCCūC can both express the perfective, continuous, or future aspect, e.g. *bēt mabyūʿ* 'a sold house', *annahr ajjāri* 'the running river', *ḥāḍir bukrah* 'he is coming tomorrow'. The participle of the verbal forms is derived by replacing the imperfect suffix *yi-* by *mi-*: *huwwa m(i)xalliṣ* 'he is finished'. The passive participle of Form I can be used as the participle of Form t-I, thus *atkasar* 'to be broken', but *maksūr* 'broken'.

2.4.7 Weak verbs

2.4.7.1 Geminate verbs

Geminate verbs, e.g. *ḥabb* 'to love', *madd* 'to stretch', have two allomorphs in the perfect, one before consonant-initial suffixes, the other before vowel-initial suffixes: *ḥabbē-na*, but *ḥabb-at*. The imperfect is invariant: *ʾaḥubb* 'I love', *yiḥubbu* 'they love'. The active and passive participles follow the patterns CāC$_{i}$iC$_{i}$ and maCC$_{i}$ūC$_{i}$: *ḥābib, maḥbūb*.

2.4.7.2 Verbs Iʾ

The initial glottal stop of *ʾakal* 'to eat', *ʾaxad* 'to take', and *ʾamar* 'to order' disappears after the imperfect prefixes, and the vowel lengthens: *ākul* 'I eat', *nākul* 'we eat', *yākul* 'he eats', *yāklu* 'they eat'. /ʾ/ is kept in Form II: *yiʾakkil* 'he feeds'. The imperative is *kul* and *xud*, the active participles *ʾākil, ʾāxid, ʾāmir*, the passive *maʾkūl, maʾxūd, maʾmūr*.

2.4.7.3 Verbs I*w*

I*w* verbs take the same inflection of the sound verb, e.g. *wigif, yiwgaf* 'to stand up'. The imperative is *awgaf, awgaf-i*, and the active and passive participles are *wāgif* 'standing up' and *mawgūf* 'detained'.

2.4.7.4 Verbs II*w/y*

II*w/y* verbs shorten their vowels before consonant-initial subject suffixes in the perfect. The conjugation of **gwm* 'to stand up' and **šyl* 'to carry' is: *gām, gām+u, šāl, šāl+at*, but *gum+t, gum+na* and *šil+t, šil+tu*. The vowels of the imperfect and imperative depend on the medial glide of the root, thus *yi-gūm, gūm* and *yi-šīl, šīl*. The reflexive and the passive of these verbs, too, follow this rule: *ʾatšāl, yitšāl* and *ʾatšal-na*.

2.4.7.5 Verbs III*w/y*

The *a-* type and the *i-* type of these verbs behave alike, except that in *-i* verbs /i/ changes to *-y-* before *-it* of the third person singular feminine and *-u* of the third person plural. The perfect and imperfect paradigms of *rama* 'to throw' and *nisi* 'to forget' are given in Table 10.

Table 10. Inflection of the verbs III*w/y*

perfect	3rd	2nd	1st
sg. masc.	*rama/* *nisi*	*ramēt/* *nisīt*	*ramēt/* *nisīt*
sg. fem.	*ramat/* *nisyat*	*ramēti/* *nisīti*	
pl.	*ramu/* *nisyu*	*ramētu/* *nisītu*	*ramēna/* *nisīna*
imperfect			
sg. masc.	*yirmi/* *yinsa*	*tirmi/* *tinsa*	*armi/* *ansa*
sg. fem.	*tirmi/* *tinsa*	*tirmi/* *tinsi*	
pl.	*yirmu/* *yinsu*	*tirmu/* *tinsu*	*nirmi/* *ninsa*

The imperative forms of the these verbs are *armi/ansa* (sg. masc.), *armi/ansi* (sg. fem.), and *armu/ansu* (pl.).

2.4.8 Quadriliteral verbs

Few of these verbs have four different radicals, e.g. *laxbaṭ* 'to mess up', *ḥarbag* 'to mingle'. Reduplicated quadriliterals, most of which have onomatopoetic meaning, are derived from Form I geminate verbs, e.g. *ʿaknan* 'to annoy' and *waswaš* 'to whisper'. Derivation from nouns is productive: *attaryag* < *taryaga* 'making fun'. The vowel of the second syllable is /a/ in the perfect and /i/ in the imperfect: *marmar/yimarmir* 'to make miserable'. Derived quadriliterals are similar in structure and meaning to Form V verbs, e.g. *atlaxbaṭ* 'to become confused'. The active and passive participles are derived by prefixing *mi-* and changing the final vowel to *i*, e.g. *milaxbiṭ* 'confusing', but *mitlaxbiṭ* 'confused.' Verbal nouns have the suffix *-a*: *laxbaṭa* 'confusion', *waswaša* 'whispering'.

2.5 *Syntax*

2.5.1 Noun phrase

The first item in a construct structure is indefinite if the second item is indefinite, e.g. *galam bint* 'a girl's pen'. Certain particles, such as *wāḥid/wāḥda* and *ʾayyi*, may express indefiniteness and specificity: *šuft wāḥid ʾamrīki* 'I saw an [a certain] American' is indefinite but specific, while *ʾākul ʾayyi šayyi* 'I will eat anything' is both indefinite and nonspecific.

2.5.2 Verbal aspect

Present tense is expressed by the simple imperfect or the *bi-* imperfect: *yikzib dayman* 'he lies', *biyinjaḥ* 'he always passes'. The past tense and the perfect aspect are rendered by the perfect: *xallaṣ* 'he is finished'. The future tense is expressed with the prefix *ḥa-* or the verb *rāḥ: ḥaktub* 'I will write'. The active participle may express the present or future tense depending on the adverb it is used with: *sāmiᶜ* 'I am listening', *rājiᶜ bukrah* 'he is returning tomorrow'.

Several prefixes express the progressive, either alone or in combination with one another: *biyiktub* 'he is writing', *kān* (*bi*)*yiktub* 'he was writing', and *kān gāᶜid ᶜammāl biyiktub* 'he was sitting writing'.

Continuation, durativity, and intensity may be expressed by a preverb particle *ᶜammāl* or by auxiliary verbs such as *fiḍil* and *gaᶜad* 'to remain': *ᶜammāl yinigg* 'he is continuously nagging', *fiḍil wāgif* 'he kept standing', and *gāᶜid yiktub* 'he kept on writing'. Intent and wishes are rendered by the use of *ibġa*, as in *ibġa yišrab* 'he wants to drink'. Verbs like *badaʾ, rijiᶜ*, and *gām* express resuming or starting an action: *badaʾ/rijiᶜ/gām yiᶜāriḍ* 'he started/resumed/initiated objecting'. Phrases like *mā ᶜād* express cessation of an action, e.g. *mā ᶜād tizūrana* 'she does not visit us any more.'

2.5.3 Word order

The basic word order is SVO; VSO is equally frequent. Both OVS and OSV occur in special contexts and require a resumptive pronoun that agrees with the preposed object in gender and number (Abu-Mansour 1982), e.g. *al-bint darras-ha mḥammad* 'Muḥammad taught the girl' and *al-awlād miḥammad darras-hum* 'Muḥammad taught the children'.

2.5.4 Existential sentences

The main existential particle is *fī* 'there is/are'. Other prepositions *ᶜind, lī* 'to have' and *minn* are used to form verblike constructions. The noun in such constructions is indefinite, e.g. *fī sayyārāt katīra* 'there are many cars', *bēt lū bābēn* 'a house with two doors', *mā minnu fāyda* 'it is useless'.

2.5.5 Conditional sentences

Conditional sentences are introduced by *iza* or *law* and less commonly by *inn: iza šuft-u gull-u* 'if you see him, tell him'. *inn* is used in unlikely conditionals when both verbs are in the perfect, e.g. *inn jā gultalu* 'if he came, I would tell him'. In counterfactual conditionals, *kān* 'to be' is obligatory in the main clause but optional in the if-clause, e.g. *law (kān) daras, kān najaḥ* 'if he had studied, he would have passed'.

2.5.6 *Ḥāl* sentences

Ḥāl sentences are used as temporal adverbial sentences. They are introduced by *w-* 'and'. They can be nominal, e.g. *zahamni w-ana xārij* 'he called me as I was leaving'; verbal, e.g. *kafašu (wu-huwwa) yiġušš* 'he caught him cheating'; or existential, e.g. *māt (wu-)ᶜindu flūs katīra* 'he died, having acquired a lot of money'. In nominal sentences, the conjunction *w-* can be deleted only if the *ḥāl* sentence is preposed, e.g. *ʾana xārij…zahamni* 'he called me as I was leaving'.

3. Lexicon

Some words can be traced back to old dialects, e.g. *madyūn* 'in debt', a form attributed to the Tamīm dialect, is used in Meccan Arabic instead of the Standard Arabic form *madīn*.

Other words came from Standard Arabic with some modifications in form or meaning: *zayy* 'like' < Standard Arabic *ziyy* 'a manner or fashion', *duwwērah* < *dāʾirah* 'circle'. The third group of words is strictly Meccan Arabic: *mirakkab* 'kitchen', *mabīt* 'a room on the top floor', *ṣuffa* 'living room', *xārja* 'roof', *judruwwah* 'walls', *barajōn* 'marbles', and *miḥramah* and *midawwarah* 'head scarves worn by old Meccan ladies'.

Some old loans are Turkish or Persian: *duġri* 'straight', *kurtah* 'dress', *badrōn* 'basement', *dandurma* 'ice cream', and *agzaxāna* 'pharmacy'. Others came from European languages such as French, English, and Italian: *kanaba* 'sofa', *taksi* 'taxi', *kamira* 'camera', *baranda* 'veranda', *winš* 'winch', *balakōna* 'balcony', *daraksōn* 'steering wheel'. Recent borrowings describe new concepts and inventions adopted by society; *kundēšan* 'air conditioner', *tilifōn, tilifizyōn, kumbuyūtar, intarnit*, and *dišš* 'satellite dish' are preferred to their Standard Arabic synonyms *mukayyif, hātif, rāʾī, ḥāsūb, šabakah ʿankabūtiyyah*, and *ṣaḥin faḍāʾī*.

Examples of blending include *bētalma* 'bathroom' < *bēt almāʾ* 'house of water', *daḥḥīn* 'now' < *hāda ʾal-ḥīn* 'this moment', and *daxīlak* 'I beg you' < *dāxil ʿala*, a phrase used by a person who enters a house and requests something.

Diachronic metonymy is exhibited in the unique use of *ṣabi/ṣabiyya* to mean 'male/female servant' instead of the original meaning 'boy/girl'. In synchronic metonymy, the original and extended meanings exist side by side, thus, *bint* 'girl, daughter' exists alongside *bint* 'virgin'.

Grammaticalization changes include the development of an auxiliary from a main verb *rāḥ* 'he went', occasionally reduced to *ḥa-* used to express the future tense; the use of a grammaticalized adverb *baʿdēn* 'later' as a future marker, e.g. *ākul baʿdēn* 'I will eat later'; and the use of a → serial verb construction to express the causative meaning, e.g. *axallīh yiktub* 'I make him write'. Grammaticalized *bi-*, a reduced form of *badaʾ* 'started', expresses the progressive, while the preposition *fī* is used in existential sentences, e.g. *fī ʾamal* 'there is hope'.

BIBLIOGRAPHICAL REFERENCES

Abu-Mansour, Mahasen Hasan. 1982. *Word order in Makkan Arabic*. M.A. thesis, University of Florida.

——. 1987. *A nonlinear analysis of Arabic syllabic phonology with special reference to Makkan*. Ph.D. diss., University of Florida, Gainesville.

——. 1996. "Voice as a privative feature: Assimilation in Arabic". *Perspectives on Arabic Linguistics* VIII, ed. Mushira Eid, 201–231. Amsterdam and Philadelphia: J. Benjamins.

——. 2000. "Hypocoristics in Arabic: An optimality-theoretic approach." Talk given at the Massachusetts Institute of Technology. Cambridge, Mass.

——. Forthcoming. "Emphasis spread in Makkan Arabic: The optimal analysis."

Al-Sasi, Omar. 1972. *Sprichwörter und andere volkskundliche Texte aus Mekka*. Ph.D. diss., University of Münster.

Bakalla, Muhammad Hasan. 1973. *The morphology and phonology of Meccan Arabic*. Ph.D. diss., University of London, School of Oriental and African Studies.

——. 1979. *The morphological and phonological components of the Arabic verb (Meccan Arabic)*. London: Longman.

Ingham, Bruce. 1971. "Some characteristics of Meccan speech". *Bulletin of the School of Oriental and African Studies* 34.273–297.

Kabrah, Rawiah S. 2004. *Opacity and transparency in the phonology of Makkan Arabic: A stratal optimality-theoretic approach*. Ph.D. diss., Boston University.

Natto, Ibrahim. 1997. *Makka al-Mukarrama: Mufradāt turāṯiyya fī l-lahja al-makkiyya*. Baḥrayn: al-Rifāʿ.

Satterthwait, Arnold. 1960. "Rate of morphemic decay in Meccan Arabic". *International Journal of Asian Linguistics* 26.256–261.

Schreiber, Giselher. 1971. *Der arabische Dialekt von Mekka: Abriß der Grammatik mit Texten und Glossar*. Freiburg: K. Schwarz.

Sieny, Mahmoud. 1978. *The syntax of urban Hijazi Arabic (Saʿudi Arabia)*. Beirut: Librairie du Liban.

Snouck Hurgronje, Christiaan. 1886. "Mekkanische Sprichwörter und Redensarten. *Bijdragen tot de Taal-, Land – en Volkenkunde van Nederlandsch Indie* 5:1.433–576.

——. 1888. *Mekka*. The Hague: Mouton.

Versteegh, Kees. 2001. *The Arabic language*. Edinburgh: Edinburgh University Press.

MAHASEN HASAN ABU-MANSOUR
(Umm al-Qura University)

Mechanisms of Linguistic Change

1. INTRODUCTION

Language change is a process at work in any language at any time, affecting all parts of its grammatical system. In an immediately perceptible way, the lexicon especially undergoes changes, either by introducing loanwords from other languages or by creating new words from material already existing in the language. The latter process, as well as phonological and morphological changes, takes place within the language system and, therefore, follows some generally valid lines of development, owing to the fact that, as shown by F. de Saussure in his *Cours de linguistique générale*, the linguistic sign is arbitrary and therefore conventional: Language is a system of symbols in which concepts are represented by sequences of sounds, and the relation between a concept and its acoustic symbol, the linguistic sign, is arbitrary. That is, there is no rule by which a certain concept might automatically evoke a certain sound chain, which is why the same concept may be symbolized by sound chains as different as Arabic /kalb/ and English /dog/. The arbitrariness of the linguistic sign, therefore, requires that it be conventional, i.e., a language community is constituted by a convention about which concept is symbolized by which sound chain. Linguistic change, then, implies either that a concept is symbolized by a different sound chain than before or that a sound chain is used to symbolize a different concept, and this, again, is subject to convention.

Although there is no interruption in the transmission of a language from one generation to the next, the accumulation of changes in the course of time are perceived as a new stage in the history of the language, leading to chronological distinctions like old, middle, and modern. Different changes in different regions and social groups lead to a regional or sociological diversification of the language into dialects and sociolects (→ variation).

On the other hand, linguistic change may also cause the unification of several dialects into a so-called koine (→ koineization). In this case, typically dialectal forms are substituted by forms that are used in a wider area. Any standard and literary language is characterized by this phenomenon to a greater or lesser degree, and the dialects may or may not survive beside the koine.

All these types of linguistic change can be observed in the history of Arabic, from its oldest documents through Qurʾānic and Classical Arabic to Modern Standard Arabic and the vast variety of dialects. They must equally be assumed for the prehistory of Arabic, i.e. for the time up to the first written records.

2. Language change and reconstruction

Although the different forms of Arabic, especially those which are remote from each other with respect to their chronological and spatial location, are no longer mutually understandable, they still are regarded as appearances of one and the same language because they are linked to each other by regular correspondences in the grammatical system and in the lexicon, and every dialect is characterized by certain sound changes and morphological developments with respect to Old and Classical Arabic. Again, what can be observed within the history of Arabic must be assumed as well for its prehistory. As is well known, Arabic shares regular correspondences in the grammatical system and in the lexicon with a number of other languages to such a degree that both accidental similarity and borrowing from one language to another must be excluded. The only viable explanation of these correspondences, then, is that all the languages involved go back to a common ancestor language and that the differences between them are the result of linguistic change – in other words, that these languages are genetically related, constituting the so-called → Semitic language family. Since the ancestor language, called Proto-Semitic or Common Semitic, is not attested, it must be reconstructed in order to show the different processes of change which led from it to the individual languages. This is why the reconstruction of a protolanguage is not an end in itself. It is, rather, a methodological requirement which cannot pretend to elaborate the ancestor language in its totality, but only as an abstract model containing all those features which can be claimed for it on the basis of the methods of comparative linguistics (cf. Stempel 1999:3–4; Schmidt 1973).

3. Archaisms and (common) innovations

Since from the point of view of the older language any change is an innovation, every more recent or dialectal form is characterized by certain typical innovations which, therefore, are the only relevant criteria for establishing it. Important innovations shared by two or more dialects may point to an intermediate stage in which these formed a single dialect or dialect area. Archaisms, on the other hand, are important for obtaining a more complete picture of the ancestor language and may point to a marginal position of the dialect in which they are found. Hence, the substitution of Old Arabic /ṯ ḏ ḏ̣/ by /t d ḍ / in the dialects of Egypt and parts of Syria and Lebanon, e.g. *talāta* 'three' instead of *ṯalaṯa*, must be taken into account as a common innovation when looking for greater dialect areas, while the preservation of the Old Arabic values in the Bedouin dialects, as well as those in Mesopotamia and Tunisia, only proves that this pronunciation must be assumed as the old one, not implying any closer connection between these dialects. With respect to the prehistory, a number of innovations shared by Arabic, on the one hand, and Epigraphic South Arabian and Ethiopic, on the other, point to a common South West Semitic stage (→ South Semitic languages), such as */p/ > /f/, development and expansion of the 'broken' plural (cf. below, Sec. 6), and certain verbal stem forms (cf. Stempel 1999:112ff., 116–118). On the other hand, the preservation in Arabic of 28 out of 29 consonants, 6 out of 6 vowels, case forms, and nasal endings (→ *tanwīn*) in the declension does help us to establish the grammar of Common Semitic, but it does not allow of any conclusion as to the position of Arabic within the Semitic languages.

4. Causes of linguistic change

One of the most important motives for linguistic change is → language contact. By this we understand a situation in which large parts of a speech community use, temporarily or permanently, one or more other languages regularly beside their own, thus transferring words, elements of word formation, syntactic patterns, and articulatory peculiarities from one language to the other. These influences or interferences may result in different types of strata: From a synchronic point of view, one language constitutes an adstratum to the other, e.g. Classical and Modern Standard Arabic to the vernaculars; the original language of the community may form a substratum within an adopted one, e.g. Berber dialects in Maghrebinian or Hispano-Latin in Andalusian Arabic; a second language used during a certain time may leave a superstratum in the original one, e.g. Arabic in Romance languages or Italian in Maltese (→ substrate).

Internal factors causing language change are to be sought in irregularities and inconsistencies within the grammatical system of the language. The latter are brought about by the fact that especially morphological change (see Sec. 6) does not mean the immediate replacement of one form by another but rather the coexistence of an older and a newer form side by side, one of which is often later given up in favor of the other.

Another motive is to be found in the tendency to reduce to a minimum the features which distinguish forms opposed to each other, the so-called linguistic economy. An example of this kind may be seen in the phonetic development of the sound known in Arabic as *ḍād*. On the basis of Arabic loanwords in Spanish, e.g. *alcalde* 'mayor' from *al-qāḍī*, and the results of the comparison with the other Semitic languages, this phoneme may be reconstructed as a lateral ejective affricate [tł'], opposed to the simple voiceless lateral affricate */ś/ [tł] (cf. Stempel 1999:56–60; → *ḍād*). After the development of the latter to Arabic /š/ (cf. Churchyard 1993), both the lateral and the affricate pronunciations are no longer distinctive, with the result that only the original point of articulation (dental) and the 'emphatic' pronunciation are preserved.

5. Phonetic-phonological change

A phonetic-phonological change occurs when a phone – either phoneme or allophone – is replaced by another one which may already exist in the language or be newly introduced by the very sound change in question. The substitution of phones is generally carried out by adding or giving up just one distinctive feature (cf. Hoenigswald 1960:72–73). In the case of */p/> /f/ already mentioned, the feature [+fricative] is introduced into the pronunciation, while all other features are retained. Similarly, the sound change /ṯ/ > /t/ means that [θ] has been replaced by [t] by giving up the feature [+fricative].

In the majority of cases, changes in the phonetic realization of phones affect the phonological system as a whole. In the case of */p/ > /f/, the old opposition between /p/ and /b/ as voiceless and voiced counterparts of a series of labial stops is given up. More important reorganizations of the system are brought about, however, by → phonological merger and → phonological split. The loss of the fricative pronunciation of /ṯ ḏ ḏ̣/ leads to the merger of these phonemes with /t d ḍ/ and hence to a complete loss of the opposition between dental stops and fricatives. The pronunciation of /k/ as [tʃ] before front vowels in the Bəgūl group of Syro-Palestinian (cf. Grotzfeld in Fischer and Jastrow 1980:174–175) means, at first, simply that the phoneme /k/ is realized by the two allophones [k] and [tʃ] which occur in complementary distribution, e.g. *baččīr* ‘early’, *kurr* ‘young donkey’; but the subsequent generalization of one of the two allophones through a whole lexeme, independently from the following vowel, produces the split of /k/ into two independent phonemes, namely, /k/ and /č/, which may now occur in the same phonetic contexts.

As to the prehistory, of the 29 consonant phonemes reconstructed for Common Semitic, 28 are preserved as distinct phonemes in Arabic, and only one merger can be observed, namely, Semitic */s/ and */š/ > Arabic /s/. This does not mean, however, that the phonetic values of Arabic may be assumed also for Common Semitic, as has often been and still is done, e.g. in the case of the (inter)dental spirants /ṯ/, /ḏ/, and /ḏ̣/, which may rather be reconstructed as former palatalized dentals */tʾ/, */dʾ/ and */tʾ/ or affricates */ts/, */dz/ and */ts’/ (cf. Stempel 1999:46–50).

Another phonetic feature typical of Arabic is the pharyngealized or velarized pronunciation of the ‘emphatic’ consonants. Apart from the question as to their original articulation in Common Semitic, it should be noted that only in Arabic are they continued both by voiceless and voiced consonants, while their correspondences in the other Semitic languages are always voiceless. The substitution of an earlier feature, most probably [+ejective] (cf. Stempel 1999:64–67), by [+velarized] gives way to a totally new system of oppositions in the consonantism, especially in the dental series: /t/ : /d/ : /ṭ/ : /ḍ/ and /s/ : /z/ : /ṣ/ : /ẓ/ (/ḏ̣/). The fact that /q/, the ‘emphatic’ member of the velar series which has no voiced counterpart, is pronounced voiced in some dialects also points to a rather recent development of voiced emphatics within Arabic.

The examples seen so far already indicate that sound change takes place in an absolutely regular manner, i.e., every occurrence of, for example, /ṯ/ in Old or Classical Arabic has been substituted by /t/ by the speakers of the Egyptian dialect. This is what in traditional historical linguistics is called a sound law without exception. Apparent deviations in which, in the case at hand, a sibilant is found instead of a dental stop are either later borrowings from Classical Arabic, e.g. *ẓarf* ‘envelope’ from Classical Arabic *ḏ̣arf* instead of **ḍarf,* which should be expected as the regular Egyptian outcome according to the sound laws (cf. Fischer and Jastrow 1980:50), or due to analogical influence within the paradigm, like the above-mentioned generalization of *k* or *č* regardless of the vowel that follows. Although in many cases the regular correspondences may be blurred by analogies no longer transparent, comparative linguistics cannot abstain from the principle that sound laws are basically exceptionless. It should be noted that even in the popular understanding of dialectal differences, the regular correspondences described here in the terms of sound laws are perceived as decisive.

6. Morphological change

Morphological change encompasses several types of change which follow basically the same mechanisms, independently from whether the morphological system as a whole is affected, i.e. whether a new grammatical category comes into being by the morphological change in question. The starting point is that any morph µ’ has a primary and a secondary function (or rather secondary functions), m_1 and m_2 respectively. The changing process begins when one of the two functions is expressed by another morph. Normally, a totally new morph µ’ is only used for the primary function, while for the secondary function a derivative formation already existing in the system may be introduced (cf. Kuryłowicz 1964:11). At this stage, the old and the new morph coexist side by side, the new one gradually taking over the secondary function(s) as well, so that in the end the old form may be given up. If, however, the coexistence of the two forms persists, the original morph splits into two independent ones, giving rise to a new grammatical opposition. An example of this type is the split of the Common Semitic prefix conjugation (cf. Stempel 1999:102–107). While **yaqtul* (μ), a so-called injunctive, is indifferent as to tense and mood, a new morph **yaqtulu* (μ') is introduced to express the primary function (m_1), called indicative, the old form being restricted to several secondary functions (m_2). This situation is still preserved in Arabic in the opposition between the indicative and the so-called jussive, while in the other West Semitic dialects, the new form has taken over, apart from a few residues, the secondary functions of the old one as well. Another possibly new category in this sense is the Arabic subjunctive (cf. Testen 1994).

From a morphological point of view, the source of a new morph is of no importance. It may result (i) from phonological change, e.g. by the loss of a final consonant, which is often and not quite correctly treated only among phonological changes; (ii) from → analogy within the morphological system, which we might call morphological change in the narrow sense; or (iii) from the use of a derivational form in a new syntactic context, which we might call morphosyntactic change.

The first source can be observed in the morphological exploitation of forms with and without final *-n* in the nominal declension of Arabic (→ *tanwīn*), which go back to Common Semitic endings in **-m* (→ mimation) and show the frequent development of final /m/ to /n/. The loss of the final nasal must have been a gradual development, most probably in accordance with the phonetic context, i.e., the final nasal may have been dropped before a consonant but preserved before a vowel; this implies at first the existence of the old and the new form side by side. Instead of generalizing the *n*-less form, however, Arabic preserves the old form, though regularly only in the literary language, in special syntactic contexts, thus constituting a new opposition between indefinite and definite forms. The fact that some nouns always show *tanwīn*, even though they are inherently definite, e.g. personal names like *muḥammadun*, proves that the nasal ending originally does not convey any special meaning but was exploited once forms with and without final nasal existed.

The second source of a new morph, a proportional analogy within the morphological system, is based upon the oppositional character of the linguistic sign, i.e. on the fact that every linguistic form stands in opposition to another. At first, there must exist a model in which one part of the opposition is perceived as marked by a certain feature absent in the other. The dual ending of the 2nd person of the prefix conjugation, not known from other Semitic languages, is evidently due to an analogy of this kind. The proportion between singular, plural, and dual in the 3rd person, *yaqtul : yaqtulū : yaqtulā*, is the model for the 2nd person: *taqtul : taqtulū : X; X = taqtulā*.

The third possibility of creating new morphs, the → grammaticalization of derivational forms, is best represented by the spread of the so-called broken plural (cf. Fischer 1987:51–58 as to the different formations) in South West Semitic. The starting point are collective formations characterized not by external morphemes (prefixes, infixes, or suffixes), but by certain apophonic types, e.g. *kutub* 'books' (sg. *kitāb*), *ʾayyām* 'days' (sg. *yawm*), which are known also from other Semitic languages, though not to the extent known from Arabic (→ apophony). These collective forms were evidently used with such frequency that they could take over the plural proper, i.e. the expression of several individuals. It goes without saying that only the nucleus of the 'broken' plurals is inherited from Common Semitic; the vast majority are developed within South West Semitic and Arabic itself, following the models presented by those first oppositions between a noun and its collective counterpart.

A quite similar replacement of regular forms by derivative, and hence less regular, ones, can be observed in the infinitive (*maṣdar*) of Form I, which by comparison with the other Semitic languages originally had the vocalization scheme *qatāl*, still preserved in Arabic outside the infinitive, cf. *salām*: Hebrew *šålōm*. Again, some regularity is achieved by combining certain verbal vocalizations with certain apophonic *maṣdar* types (cf. Fischer 1987:109–111).

7. OTHER DEVELOPMENTS

A further utilization of apophony is the → passive of the type *qutila* (: *qatala*). Although there are apophonic passives also outside Arabic, especially in Hebrew, their development must be independent (cf. Retsö 1989:6–31). The main difficulty is that there is no evident model on which the vocalization schemes may have been established.

With respect to deictic pronouns (→ deixis), Arabic shows a specific expansion by combining different pronominal stems and suffixes (cf. Fleisch 1968:139–151), e.g. *ḏālika* 'that there [by you]', which contains the old deictic *ḏā*; the second element *l* is known, from a comparative point of view, from the plural *ʾulā(ʾi)*, cf. Hebrew *ʾēl(leh)*, but of course also from the definite article of Arabic, and it is not quite clear how this combination exactly arose; the third element is simply the personal/possessive pronoun of the 2nd person and in the *Qurʾān* may still vary according to number and gender of the person addressed.

Another special development of Arabic is the emergence of the → relative pronoun *allaḏī*, which in its turn is also a combination of elements already known from deictic pronouns.

It is a point of controversy whether the article *(a)l* goes back to a Common Semitic formation or must be seen as a specific Arabic innovation, at least with respect to its form (cf. Zaborski 2000).

8. Concluding remarks

The changing processes which can be observed in the history of Arabic, from its Common Semitic ancestor language to the modern dialects, may in general be called 'normal' or 'expected' in comparison with similar developments known from, for example, Indo-European languages. Among these we find the loss of final consonants, palatalization of velars before front vowels, compensatory lengthening of vowels, and so on.

There is a remarkable tendency toward 'introflexion' (cf. Skalička 1979:340) or 'flexion interne' (cf. Fleisch 1968) by privileging apophonic or 'inner' formations to such with affixes. At first glance, this seems to contradict the general tendency of avoiding irregular formations in favor of regular paradigms, but this latter process takes effect as well by aligning certain apophonic oppositions into more regular patterns.

The position of Arabic within Semitic is certainly that of an archaic language in some respects, namely the phonological system and the preservation of the nominal declension. On the other hand, many of the changing processes have brought about a particularly complex language system with a multitude of new grammatical oppositions, so that on the whole it must be called 'modern' against the Common Semitic background.

Bibliographical references

Churchyard, Henry. 1993. "Early Arabic *siin* and *šiin* in light of the Proto-Semitic fricative-lateral hypothesis". *Perspectives on Arabic linguistics*, V, ed. Mushira Eid and Clive Holes, 313–342. Amsterdam and Philadelphia: J. Benjamins.

Fischer, Wolfdietrich. 1987. *Grammatik des klassischen Arabisch.* 2nd rev. ed. Wiesbaden: O. Harrassowitz.

——, and Otto Jastrow. 1980. *Handbuch der arabischen Dialekte.* Wiesbaden: O. Harrassowitz.

Fleisch, Henri. 1968. *L'arabe classique: Esquisse d'une structure linguistique.* 2nd rev. ed. Beirut: Dar El-Machreq.

Hoenigswald, Henri M. 1960. *Language change and linguistic reconstruction.* Chicago: University of Chicago Press.

Kuryłowicz, Jerzy. 1964. *The inflectional categories of Indo-European.* Heidelberg: C. Winter.

Retsö, Jan. 1989. *Diathesis in the Semitic languages: A comparative morphological study.* Leiden: E.J. Brill.

Schmidt, Karl Horst. 1973. "Grundlagen und Methoden der historischen Sprachvergleichung". *Perspektiven der Linguistik*, I, ed. Walter A. Koch, 105–133. Stuttgart: A. Kröner.

Skalička, Vladimir. 1979. *Typologische Studien*, ed. P. Hartmann. Wiesbaden: Vieweg.

Stempel, Reinhard. 1999. *Abriss einer historischen Grammatik der semitischen Sprachen.* Frankfurt a. Main: P. Lang.

Testen, David. 1994. "On the development of the Arabic subjunctive". *Perspectives on Arabic linguistics*, VI, ed. Mushira Eid, Vicente Cantarino, and Keith Walters, 151–166. Amsterdam and Philadelphia: J. Benjamins.

Zaborski, Andrzej. 2000. "Inflected article in Proto-Arabic and some other West Semitic languages". *Asian and African Studies* 9.24–35.

Reinhard Stempel (Universidad del País Vasco)

Media

'Media' (Arabic *wasāʾil al-ʾiʿlām*, pl. of *wasīlat al-ʾiʿlām*) is a term for vehicles for the widespread communication to and entertainment of an audience, including printed and electronic means. It is widely agreed that the media are not neutral impassive agencies that transmit news and views, but are themselves influential selectors, shapers, manufacturers, and even, on occasion, fabricators of news and views (McArthur 1998). The topic of this entry is the development of Arabic language media, specifically in Egypt; it discusses journalism, radio, cassette recorders, multichannel television, satellites, new media technology, and the Internet. For the language of the media, → Media Arabic.

1. Newspapers

Arabic journalism originated in Egypt in 1798. During the French campaign, Arabic pamphlets were distributed instead of the newspapers. The first official newspaper, issued in 1828, was *al-Waqāʾiʿ al-miṣriyya* ‘The Egyptian proceedings’, edited by Rifāʿ at-Taḥṭāwī. *Al-ʾAhrām* ‘The pyramids’ newspaper was issued by Salīm and Bušārā Taqla in 1875. It has been Egypt’s leading newspaper ever since.

At that time, the newspapers played an important role in politics. ʿAlī Yūsuf issued *al-Muʾayyid* ‘The supporter’; ʾAḥmad Luṭfī as-Sayyid issued *al-Garīda* ‘The newspaper’; Muṣṭafā Kāmil issued *al-Liwāʾ* ‘The banner’; and ʾAmīn ar-Rafʿī issued *al-ʾAxbār* ‘The news’. Humor magazines began to appear, such as *ʾAbū naḏḏāra* ‘The man with glasses’, issued by Yaʿqūb Ṣanūʿ, and *at-Tankīt wa-t-tabkīt* ‘Joking and rebuking’ by Nadīm, as a means of criticizing the negative aspects of society.

In 1884, Jamāl ad-Dīn al-ʾAfġāni and Muḥammad ʿAbdū founded *al-ʿUrwa al-wuṯqā* ‘The unbreakable bond’. Under British rule, journalism enjoyed a period of relative freedom. However, during World War I, strict censorship followed, especially after the event of Dinshaway (Al Gindi 1963). After the war, the political parties started to publish their own newspapers in an effort to promote their ideas and principles. The following are some of the more famous newspapers published at that time: *Ṣawt al-ʾumma* ‘The voice of the nation’, *al-ʾAsās* ‘The basis’, *Miṣr al-fattāḥ* ‘Victorious Egypt’, *al-ʾIštirākī* ‘The socialist’, *al-ʾIxwān al-muslimūn* ‘The Moslem Brothers’, and *al-Gamāhir* ‘The public’. The Wafd party owned *al-Balāġ al-miṣrī* ‘The Egyptian communiqué’. In 1944, *al-ʾAxbār* ‘The news’ and *ʾAxbār al-yawm* ‘The news of the day’ were published by Muṣṭafā and ʿAlī ʾAmīn.

In those years, newspapers enjoyed a period of freedom in which they could discuss political and social problems. It was during that time that journalism reached its golden age, especially since martial laws and censorship were abolished (Abdu 1982).

Nasser’s new regime was welcomed by *al-Miṣrī* ‘The Egyptian’ and *Rūz al-yūsuf*. However, it was not long before the new regime began to impose hard censorship and abolished the multiparty system. A private Socialist newspaper was established as a voice of the one-party system established by President Nasser. During his rule, Egyptian newspapers became the machine for international misinformation campaigns and falsified stories labeled by Egyptians as *kalām garāyed* ‘journalistic talk’ (Abdelfattah 1990). When nationalization was declared, all media outlets were placed under the control of the Arab Socialist Union, which set strict guidelines for the press. When multiple political parties were restored under Sadat’s regime, the number of daily newspapers and weeklies increased, and many Egyptians consider *al-ʾAhrām* to be the official newspaper of the country.

The important Egyptian newspapers and magazines are issued by a number of associations: Al Ahram Association issues the daily newspaper *al-ʾAhrām* and the periodicals *al-ʾAhrām al-iqtiṣādī* ‘The economic Ahram’, *aš-Šabāb wa-l-ʿulūm* ‘Youth and science’, *al-Mustaqbal* ‘The future’, *as-Siyāsa ad-dawliyya* ‘International politics’, *al-ʿArabī* ‘The Arab’, and others.

Axbar al-Yom Association issues *al-ʾAxbār* ‘The news’, *ʾAxbār al-yawm* ‘The news of the day’, and *ʾĀxir sāʿa* ‘The last hour’.

Dar al-Ma’aref issues *ʾÚktūbar*. Ros al-Yusuf issues *Rūz al-yūsuf* and *Ṣabāḥ al-xayr* ‘Good morning’. Dar al-Tahrir issues *al-Gumhūriyya* ‘The republic’, *al-Masāʾ* ‘The evening’, *ʿAqīdatī* ‘My belief’, and *Ḥurriyyatī* ‘My freedom’.

Dar al-Helal issues *al-Muṣawwar* ‘The illustrated’, *Hawāʾ* ‘Eve’, *al-Kawākib* ‘The stars’, and *al-Hilāl* ‘The crescent’. In addition, each political party issues its own newspaper: *al-ʾAḥrār* ‘The liberals’ (1977), *al-ʾAhālī* ‘The masses’ (1978), *Māyo* ‘May’ (1981), *aš-Šaʿb* ‘The people’ (1979), and *al-Wafd* ‘The delegation’ (1984).

According to Schleifer (1989), the first major impact of new satellite technologies on Arab media occurred in the 1980s, not the 1990s, with satellite daily newspapers, not television. First, the *aš-Šarq al-ʾAwsaṭ* newspaper and later the *al-Ḥayāt* newspaper began satellite transmissions from London to major population centers throughout the Arab world.

2. RADIO BROADCASTING

2.1 *Egyptian broadcasting*

The official opening date of the Marconi-operated Egyptian radio service was 31 May 1934. The Marconi contract was renewed in 1943. However, anti-British sentiment in the wake of World War II caused the government to cancel the Marconi contract, and the radio service became Egyptian-owned and operated (Youssef 1971). After the formation of the Egyptian Radio Television Federation, control over the electronic media alternated between the Ministry of Information and the Office of the President. Law No. 98 mandated that the language of the service be Arabic (Barrada 1970). However, the service has never broadcast domestically exclusively in Arabic.

Radio became the voice of the revolution. President Nasser's accomplishments were broadcast to all Arab countries. Nasser, a gifted public speaker, understood the power of the Arabic language and culture. Thus, he profited from this electronic medium to articulate the goals of the revolution. Village peasants and women found themselves sought after as important members of the listening audience. The radio service offered a variety of programs: Radio Cairo, 'the Main Program' broadcast, which expanded after July 1952 (Metwally n.d.), offered and continues to offer news, commentary, and various forms of entertainment, the most dominant of which has been drama. Residents of other countries have always been interested in what Egypt's Main Program is saying, especially during times of war. The Main Program is the most listened-to station in the Arab world, following the BBC and Radio Monte Carlo Middle East (RMCME; British Broadcasting Corporation yearly report, 1990). The second 'Program' provided the elite with intellectual programs. The Alexandria Local Service was established in 1954 as the first of a series of local services. The People's Program transmissions were intended for illiterate farmers, to promote agricultural advances, literacy training, population planning, and the concept of nationhood (Boyd 1999).

The Middle East Program (MEP) began in 1959. Nasser's presidential decree established new goals for Egyptian broadcasting: strengthening national feeling, reviving and spreading Arabic culture, encouraging talents, informing foreign countries about the United Arab Republic and the Arab world, and providing entertainment. Of particular importance, MEP allowed broadcast services to accept commercial advertising. This service quickly became popular and an important source of hard currency.

Egypt's religious service, the Holy Qur'an Broadcast, was established as an important means of emphasizing the fact that Egypt was an Islamic country, in spite of its close ties with the Soviet Union. It transmitted *Qurʾān* recitations and religious discussions.

Youth Broadcast started in 1975, intending to reach school-age audiences with educational, political, and social messages.

The Egyptian Palestine Program was devoted to the discussion of the Palestinian problem. After the Egyptian-Israeli Sinai agreement in 1975, the Palestine broadcasts were stopped.

The Voice of the Arabs is the best known and most widely listened-to regional Arabic radio service. Nasser used this medium to promote his views on Pan-Arabism (Boyd 1999). The Voice of the Arabs was the first major propaganda radio station in the Middle East to have a measurable impact on listeners. It was also used to broadcast to African countries south of the Sahara for the specific purpose of supporting liberation struggles. Since Sadat's assassination, the service has become Egypt's main regional service.

Radio Cairo's Main Program became domestic in orientation. In April 1981, Egyptian radio was reorganized around seven general network programs: the Main Network, the Local Network, the Qur'an Network, the Cultural Network, the Voice of the Arabs, the Communication Network, and the Overseas Network (Egyptian Radio and Television Union, 1988). In the 1990s, despite the name changes for some of the services, relatively few major programming changes took place. The new names are the General Program, the Local Network (comprising eleven stations), the Qur'an Network, the Cultural Network, the Voice of the Arabs, the Middle East Broadcasting Service, and the Foreign Language Service.

Radio remains an important medium of communication within the Arab world. The majority of the people in the Arab world still depend on radio as a source of entertainment, education, and information. Although the number of illiterate Arabs is decreasing, the overall rate of illiteracy remains very high. Today, Arab listeners no longer need to tune to the transmissions of other Arab countries for news and entertainment. It is also widely believed that television has decreased interest in radio broadc ast.

Egypt has expanded its international broadcasts. Programs in Arabic are transmitted to Southeast Asia, Latin America, East, Central, and South Africa, and North America (Egyptian Radio and Television Union 1988).

2.2 *International radio broadcasting to Arabic-speaking countries*

Eastern Europe and the Soviet Union have been significant contributors to the number of Arabic hours transmitted to the Middle East. However, the political and economic changes that have taken place in Eastern Europe since 1989 have negatively affected the number of services using Arabic. For example, Moscow's Radio Peace and Progress, a long-time Arabic broadcaster, ceased operations on 31 May 1991 (British Broadcasting Corporation yearly report, 1991). After English, Arabic is the world's most internationally broadcast language. The Iraqi invasion of Kuwait in August of 1990 and the subsequent Gulf War, along with the deterioration of the Eastern European countries, are two main factors that influenced transmission in Arabic. Since this period, electronic media in the Arab world have been operated by governments. Many Arab world radio listeners tune to foreign stations to learn the international perspective on news and current affairs and to hear programming not generally broadcast by government-operated stations.

The Arab world was the location of the first effort by the West to broadcast to a developing area for the purpose of attempting to influence people. Beginning in 1934, Italy, through its international radio service BARI, started broadcasting across the Mediterranean in Arabic (Radio Televisione Italiana 1979). Mussolini's radio broadcasts turned increasingly anti-British just after 1935. Britain began studying the possibility of an Arabic service. The BBC hired Egyptian announcers and tried to present appealing radio offerings. Competition arose between Radio BARI and the BBC. In 1939, the Soviet Union and France began broadcasting in Arabic. During the Second World War, the main international broadcasters to the Arab world were Germany and the United Kingdom. The BBC had popular announcers during the war. In his study in the early 1950s, Brunner (1953) mentions the importance of the coffeehouse as a place for listening to radio. There have been numerous major post-World War II events that have tended to increase Arabic broadcasts to the Middle East: the creation of Israel; the 1956 Suez War; the Arab-Israeli conflicts of 1967 and 1973; the increase of the oil price following the October 1973 war; the beginning of the Lebanese civil war in the mid-1970s; President Sadat's trip to Jerusalem in 1977 and his subsequent assassination; the 1982 Israeli invasion of Lebanon and the stationing of United States Marines in Beirut; and the Iraqi invasion of Kuwait in 1990 and the Gulf crisis.

Western Europe continues to be the leader in Arabic transmission hours. RMCME serves the eastern Arab states and Egypt. Since the 1980s, Radio France International (RFI) has broadcast an Arabic service. It was believed that a popular radio service competing with the Voice of America (VOA) and the BBC would produce closer Arab-French ties. The Radio Monte Carlo Middle East moved in the mid-1970s to the SOFIRAD building in Paris (Regnier 1980). The station has attained an enviable position among rival VOA and BBC broadcasts. Its pro-Arab political orientation influences its popularity.

Medi 1-Radio Mediterranée Internationale, a private venture between France and Morocco, began broadcasting in Arabic in 1980. Surveys indicate that among the major international services transmitting in Arabic are VOA, BBC, RMCME, Deutsche Welle (1959), and Radio Moscow.

Dutch Arabic programming by Radio Nederland Wereldomroep started in 1948. Austria, Greece, Italy, Malta, Spain, and Switzerland are minor Arabic broadcasters. Finally, the Austrian radio service started an Arabic service in 1989 (British Broadcasting Corporation yearly report, 1989).

Iranian Arabic transmission increased after the Iran-Iraq war and during the Iraqi invasion of Kuwait. Cyprus, an active Arabic broadcaster because of its geographic location, increased broadcasts in Arabic during the Lebanese civil war. India, which has a Muslim minority and many expatriate workers in the Arab world, began Arabic broadcasts in 1941. The People's Republic of China started an Arabic service after the 1949 revolution to increase its trade with Arab countries. The Muslim countries of Turkey, Pakistan, Indonesia, and Malaysia have increased their Arabic transmission hours. Four factors – history, trade, politics, and religion – are the major incentives to transmitting in Arabic. Presently, only four countries in North and South America broadcast to the Middle East in Arabic: the United States, Canada, Cuba, and Argentina. The VOA service ranks third, behind RMCME and the BBC, in terms of audience size in the Middle East. The newest North American Arabic service is that of Radio Canada International (RCI).

In 1964, Nigeria began transmitting in Arabic (Okesanya 1975). Ethiopia has a modest Arabic service, the Radio Voice of the Gospel (RVOG). Djibouti transmits in Arabic because of its Islamic orientation. We should also mention that religious Christian broadcasters wish to reach Christians whose native language is Arabic. New Jersey has the largest religious international radio broadcaster, and Arabic is an important language of Trans World Radio (TWR).

3. Egyptian television

In late 1959, a contract was signed between the United Arab Republic and the Radio Corporation of America (RCA) to provide a complete television service for Egypt. It was due to the vision of Nasser that 'Channel Five' began transmitting news and programs for development and education. The second program, 'Channel Nine', was designed to reach the urban areas. A third channel started in 1960 but was closed after 1967. The end of the war of 1967 saw a decrease of foreign programming. More importance was given to nationalistic, educational, and religious topics. However, the quality of Egyptian TV programs declined between 1967 and 1974. President Nasser signed a decree establishing the Egyptian Radio-Television Federation. Television in Egypt is owned by the state. The national television channels were extensions of the Ministry of Information.

More than any other mass media, television tended to reflect the changing international political orientation of the country. During the Egyptian-Israeli war of 1973, the Egyptian media took a very different attitude than they had during and after the 1967 war. Radio tended to be more honest and less confident about the victory when the Suez Canal was crossed. Television reflected the happiness of the Egyptians after 1973. With Sadat's declaration of an open-door policy, British and American programs increased on Egyptian television. The income of advertisement and commercials and the revenue from exporting Arabic programs to other countries increased significantly. Since the 1980s, the Egyptian television became the most important outlet to advertize a variety of goods and services (Egyptian Radio and Television Union 1988). The system of Egyptian Radio and Television Union (ERTU) is the most influential in the Arab world. Arab countries have generally relied on Egyptian media production for television programming such as drama series, variety programs, talk shows, and films. The 'Voice of Cairo' was established to become involved in production and distribution in the Arab world, and has become an important source of hard currency. After Egypt's peace treaty with Israel, some countries decreased or stopped altogether the purchase of Egyptian TV programs: Lebanon, South Yemen, Libya, Syria, and Iraq.

The weekly magazine *al-ʾIḏāʿa wa-t-tilivizyūn* was first issued in 1935 under the name of *ar-Rādiyū al-miṣrī* to provide articles about broadcasting figures, features of the various aspects of the broadcast media, and a detailed television and radio schedule. During the period from 1981 to 1992, several factors contributed to the development and expansion of Egyptian television: A third channel was established.

Currency restrictions on the import of goods were lifted. The death in October 1983 of President Anwar Sadat and the evolution of Egypt as part of the Arab world have contributed to the confidence building of Egyptian television and greatly added to the attractiveness of Egyptian video products. The Lebanese civil war ended Lebanon's competition with Egyptian program exports. President Mubarak gave more room for free expression in news, films, and television series.

Television has become the main source of entertainment over information at home or in the local coffee shop. For many lower-income people, a television set is more important than a refrigerator. With the spread of digital multichannel satellite platforms, the impact of television in this decade has exploded.

4. Videocassette recorders

Egyptians working outside the country brought home to their families television sets and videocassettes. Videocassette recorders are widely used in Egyptian urban and rural areas by viewers who want to have a choice in what they watch. Both Egyptian and Western films are widely available in rental stores and shops.

5. Satellite television

Satellite distribution has offered new opportunities to Egypt. The ARABSAT system has made it possible for television systems in the Middle East to rebroadcast Egyptian television live. Beginning in 1990, Egypt began offering via satellite connection rebroadcasts of the Egyptian main service. In late 1990, Bahrain permitted limited rebroadcasting of Egyptian television. For a time following the 1991 Gulf War, Egyptian television was used in place of Kuwaiti television. In 1992, the Egyptian government legalized the import and ownership of dishes.

Tapes of Egyptian television programming are part of virtually every Middle East television system. Satellite distribution has presented Egypt with new opportunities. Egypt's stock of films and television productions and ARABSAT-leased transponders make this technically possible. NILESAT was inaugurated in May 1995, providing specialized educational, cultural, family, children, religious, health, sports, drama, and other programming. It allows the rebroadcast of all Egyptian television and radio services, as well as national channels from other Arab countries. The activation of Egypt's NILESAT direct broadcast service makes it clear that the Egyptian government will continue to invest in television and radio production and delivery systems to maintain its historical dominance of Arab-world electronic media (Boyd 1999).

In 1989 CNN began to rebroadcast terrestrially, a pay-TV operation that came to be known as CNE (Cable Network Egypt) Cairo. MultiChoice Egypt now administers a subscription management service and holds a small minority share in CNE Cairo.

The Egyptian Space Net (ESN) began to broadcast thirteen hours of daily programming. After the Gulf War, dish ownership was further stimulated because of a decline in the price of satellite dishes, wider international programming, and an increase in the number of companies marketing, servicing, and even manufacturing dishes. Television broadcasting directly to receivers in and to the Arab world via satellite became possible. The first of these satellite systems was the Middle East Broadcasting Centre (MBC), which began transmission and production in London in September 1991. It broadcasts eighteen hours a day, providing a mix of news, sports, fashions, movies, and general entertainment.

The second private Arab satellite system – and the largest in reach and in Arabic programming – is Arab Radio and Television (ART), established by Sheikh Saleh Kamel. It began transmission via ARABSAT from Cairo in January 1994. ART started separate channels for movies, sports, and children's programming, as well as a general channel. Since then, ART has become a global platform for more than twenty channels, besides live and syndicated sports coverage. By mid-1997 ART offered six satellite program services. In its quest for quality, ART has increased the number of public shows which involve audience participation and tackle controversial and daring social issues, such as divorce, premarital sex, male impotence, and drug use (Schleifer 1998).

The most recent Arab satellite system to come onto the scene is Orbit, launched in May 1994 and transmitted from Rome. Orbit's Arabic programming is more or less confined to *at̠-Ṯāniya*, an Arabic-language general channel with some of the most successful entertainment and public affairs programming on the air. Orbit challenged MBC's strong card – an Arabic-language news service that provided field reporting at an international standard beyond comparison to propagandistic tendencies of the Arab world's national channels. Orbit made a contract with the BBC to produce a BBC Arabic World Television Service, which was offered exclusively in the Orbit package. It is an independent and credible news source in the Arab world. Orbit began transmitting an Arabic version of the BBC World Service with full editorial control remaining in the hands of the BBC – in other words, in the hands of non-Arabs.

When the BBC Arabic TV Service began in June 1994, it appeared to overwhelm MBC; however, the relative indifference to Arab cultural values on the part of many BBC executives, combined with often culturally irrelevant programming, troubled many Arab viewers. Thus, Orbit's management was obliged to break its contract with the BBC in 1996. MBC distinguished itself by following the international format; newsworthiness rather than government press releases determined the line-up and news stories were scripted to picture rather than to an anchor reading wire copy. MBC's style was also unique because, in contrast to the BBC Arabic service, it retained the more cautious approach (Schleifer 1998).

MBC's chief problem is that it is no longer alone. Al-Jazeera, the all-news Arabic-language channel transmitted via satellite from Qatar, employs many of the former Arabic-speaking broadcasters from the BBC venture. The other competitors include the Arab News Network (ANN), Abu Dhabi, Dubai, LBC, El Arabia, and the Egyptian channels.

As for Orbit, it has gained from the BBC experience to increase the amount of original Arabic programming by dramatically expanding its own production facilities in Cairo, Beirut, and the Gulf.

Arabic satellite programming – Arab in its cultural authenticity – has dramatically stimulated the sale of dishes. With the increasing availability of dishes, Islamic society in general, and Arab society in particular, is proud of the use of Classical Arabic and the media-prevalent Modern Standard Arabic. However, there is a concern about what is shocking to Arab conventions and against Islamic principles and ethical, moral, and social values. The first political danger is the fear of reaction of anti-Western, Islamic fundamentalism to the sudden easy availability of disorienting, subversive cultural materials of Western television programming via Direct Broadcast Satellite (DBS)/Direct To Home (DTH) satellite transmission. The second political concern has to do with government sensitivity to unfavorable news reporting by satellite stations. The third political concern is the possibility of satellite broadcasts from hostile countries.

By the mid-1990s, DTH television had become a fact of life, with widespread dish ownership throughout the region, particularly in the Gulf. DTH satellite transmission has become increasingly popular in Egypt. The increasing popularity of satellite television and the wide variety of choices have led some to question whether there should be some control over what is transmitted. Programs span all fields of knowledge: scientific, political, economic, cultural, religious, and educational, as well as entertainment programs.

Egypt has always been a pioneer in media production in the Middle East. Egypt is very keen to preserve the Arabic culture and to spread it to native and non-native speakers of Arabic. Satellite makes it possible for → Media Arabic to reach all parts of the world. Thus, the challenge for Egypt lies in upgrading and expanding its own television product, whether news or entertainment, so that it can ensure its cultural sovereignty in a globally competitive situation. Egypt, alone among the Arab states, has the depth of talent for such an undertaking. It has scores of actors, singers, dancers, musicians, comedians, journalists, producers, directors, and even *Qur'ān* reciters. In addition, Egyptian colloquial Arabic is universally understood throughout the Arab world due to the spread of Egyptian films.

Al-Mehwar, Nile News, and Nile TV are news channels that broadcast from Egypt. They provide something different than most of the news channels in the Gulf region. They represent the essence of Egyptian life from a news angle with a comprehensive concept of news, not just political news but also business news, cultural news, sports, and fashion. Egypt is a different culture, a different civilization, a different political situation. Egypt is the most populated country in the region, with the longest history, and many people across the region are interested in it.

6. CONCLUSION

Nowadays, the influence of the print media has declined in the face of technological advances. The televised media have now adapted the print media and blended it with visual and vocal material. In the 21st century, it is not the print news media but the space and terrestrial channels of the television, as well as the Internet, that have the greater influence in the trends of international and public opinion.

New media technology, from cassette recorders to multichannel television, mobile phones, and even the Internet, afford wider circulation to communications to a highly educated generation. The new generation has unprecedented opportunities for participation in media, and receptor models changed to be more interactive ones.

BIBLIOGRAPHICAL REFERENCES

Abdelfattah, Nabil. 1990. *Linguistic changes in journalistic language in Egypt, 1937–1997: A quantitative and comparative analysis*. Ph.D. diss., University of Texas at Austin.

Abdu, Ibrahim. 1982. *Egyptian press development*. Cairo: Association of Arab Register.

Abu-Lughod, Ibrahim. 1963. "The mass media and Egyptian village life". *Social Forces* 42.97–104.

Al Gindi, Anwar. 1963. *The Arabic language between supporters and opponents*. Cairo: Anglo Bookshop.

Arab States Broadcasting Union. 1969. "The United Arab Republic: TV". ASBU, Cairo. [Mimeographed.]

Badawī, as-Saʿīd. 1973. *Mustawayāt al-luġa al-ʿarabiyya al-muʿāṣira fī Miṣr*. Cairo: Dār al-Maʿārif.

Barrada, F. 1970. "The history of the U.A.R. radio since its establishment in 1934 till now". *Arab Broadcasts* 63–70. [Transl. from Arabic.]

Boyd, Douglas A. 1999. *Broadcasting in the Arab world: A survey of the electronic media in the Middle East*. Ames: Iowa State University Press.

British Broadcasting Corporation. 1980. *BBC handbook*. London: British Broadcasting Corporation.

Brunner, Edmund. 1953. "Rural communications: Behavior and attitudes in the Middle East". *Rural Sociology* 18.149–155.

Egyptian Radio and Television Union. 1988. *Egyptian Radio and Television Union in brief*. Cairo: Egyptian Radio and Television Union.

Glubb, John Bagot. 1959. *Britain and the Arabs*. London: Hodder and Stoughton.

Hamza, Abdellateef. 1961. *Literature and the press*. Cairo: Dār al-Fikr al-ʿArabī.

McArthur, Tom. 1998. *The concise Oxford companion to the English language*. Oxford: Oxford University Press.

Metwally, E.A. N.d. "Historical survey of the Egyptian Broadcast with background of the prerevolution period". Egyptian Radio Television Federation, Cairo. [Mimeographed.]

Okesanya, Ayo. 1975. *Nigerian Broadcasting Corporation*. Lagos.

Radio Televisione Italiana. 1979. Rome.

Regnier, Jean Pierre. 1980. *Radio Monte Carlo Middle East*. Paris.

Schleifer, Abdallah S. 1998. *Transnational broadcasting studies*. Cairo: Adham Center for Television and Journalism and The American University in Cairo.

Youssef, F. 1971. "What do you know about Egyptian broadcasting?" *Arab Broadcasts* 63–77. [Transl. from Arabic.]

RAGIA M. EFFAT (Cairo University)

Media Arabic

One of the most important aspects of Arab press and radio broadcasting is the Arabic language itself. The growth of Arab mass → media since the 1950s has greatly enhanced the wider use of → Modern Standard Arabic. This is the language of the newspapers and the electronic media, which is generally understood by the population of the Arab world.

Arabic itself is viewed as an important element in the effectiveness of a propaganda effort. The Arabic language, creatively employed and strongly delivered, will produce the intended reaction among listeners. Arabic is in many ways suited to radio broadcasts that are especially designed to influence others because of its rich grammar, repetitive style, and vagueness. The Arabic speaker who seeks to persuade others uses appeals that are more emotional than logical (Boyd 1999). The radio station Voice of the Arabs constituted a major part of the entire Egyptian propaganda effort, and its broadcasts were the most influential part of the propaganda campaign. As Glubb (1959) notes, "Broadcasting indeed appears to be ideally suited to the Egyptian mentality, with its eloquence, excitability and emotional appeal". In 1958, Nasser emphatically refused to stop the Voice of the Arabs broadcast.

Populations in urban and rural areas first became aware of the media through the spread of education. The language of the newspapers developed and reflected the reality of the contemporary language, in which the gap between Classical Arabic and Egyptian colloquial Arabic is narrowed (Badawī 1973). At the end of the 19th century, the Egyptian colloquial dialect began to be used with standard Arabic in a new trend known as 'Folkloristic Journalism'. The main target of such writing was the social situation in Egypt: "The conscious effort to downscale the complexity in Journalistic language was obviously successful, as evidenced when illiterate members of a family had their literate relatives read the newspapers for them, apparently with no difficulty of comprehension" (Abu-Lughod 1963).

In general, the use of dialect in the newspapers continues to be frowned upon, although it is occasionally used, for instance in reporting direct speech (cf. Diem 1974:91–95). Holes (1995:309–310) explains that when dialect is used in newspaper articles, it serves a special purpose; he quotes as an example an interview with a Lebanese politician who emphasizes the importance of the ordinary voters. In this instance, dialect is deemed to be appropriate. Nonetheless, the use of dialect continues to be something requiring a special excuse. An example is the speech President as-Sadat held the day before he was assassinated. This speech appeared two days later in the newspapers in its original form, i.e. in a mixed form of dialect and standard Arabic, and the editors of the newspaper apologized for not having been able to 'translate' it into standard Arabic because of the lack of time.

The use of dialect Arabic in an article is commonly regarded as a mistake. This concept throws together two categories: deviations from Standard Arabic as the result of code-mixing, and hypercorrections or grammatical mistakes. An interesting source for both kinds of 'mistakes' is constituted by the style books that some newspapers publish for their own journalists. These style books usually contain a chapter on 'frequent mistakes' (*ʾaxṭāʾ luġawiyya šāʾiʿa*) which illustrates the kind of phenomena one may find in newspaper Arabic. Thus, the stylebook of *al-ʾAhrām* (Nāfiʿ n.d.), along with mistakes in the orthography of the *hamza* and in the construction of the numerals, mentions such examples as *naḥnu l-miṣriyyūna* or *naḥnu ka-miṣriyyīna* instead of the correct *naḥnu l-miṣriyyīna* (Nāfiʿ n.d.:52), and the ubiquitous *laʿiba dawran kabīran* 'to play a big role' (Nāfiʿ n.d.:53), which is cited frequently as an example of nefarious foreign influence and is decried by the *al-ʾĀhrām* stylebook as wrong (their argument being that this expression is often used for serious matters that have nothing to do with play!).

The Modern Standard Arabic used in the newspapers differs in some respects from the language as it is used elsewhere. According to Ashtiany (1993), two general features characterize newspaper Arabic: variation and padding. Variation seems, for instance, to be at the root of the frequent use of synonyms for conjunctions. Of the simple conjunctions *wa-* and *fa-*, *wa-* continues to be used in Media Arabic, in particular in *ḥāl* constructions, but it is often reinforced or replaced by compound conjunctions, e.g. *kamā*, *(wa-)jadīrun bi-ḏ-ḏikri ʾanna*… (Ashtiany 1993:31); in correlative sentences it is often replaced by *ḥayṯu* or *ʾiḏ*. As for *fa-*, outside its use in the expression *ʾammā*…*fa-*… 'as for…', and in conditional sentences, it is not used frequently in Media Arabic (Ashtiany 1993:31). For other conjunctions, too, several new variants have become popular in newspaper Arabic (Girod 2000). Another example of variation is the shifting between prepositions, such as *bi-/fī*, *fī/ladā*, *ladā/ʿinda* (Ashtiany 1993:32).

For the purpose of padding, which in itself can be a means to achieve variation, the following devices are mentioned by Ashtiany (1993:28):

i. The use of *qāma bi-* as a synonym of active verbs, e.g. *qāma bi-ziyāra* instead of *zāra* 'to visit'; *tasallama r-raʾīs risālatan qāma bi-taslīmihā s-safīr* 'the president received a letter which was delivered by the ambassador' (1993:30).
ii. The use of *tamma bi-* as a synonym of passive verbs, e.g. *sa-yatimmu ġadan tawqīʿ ittifāqiyyatin tijāriyyatin* 'a commercial treaty will be signed tomorrow'; *tamma l-ittifāq ʿalā ʾirsāl wafdin ʾilā l-muʾtamar* 'it was agreed to send a delegation to the conference' (1993:30). Holes (1995:258–259) points out that this 'periphrastic passive' has an aspectual connotation as well, since it is typically used for durative or iterative actions. The same point is made by Girod

(2000), who has studied the use of this construction extensively. The verb *jarā* is used with the same meaning, e.g. *jarā xilāl al-ijtimāʿ tabādul al-ʾārāʾ* ‘during the meeting an exchange of opinions took place’ (Ashtiany 1993:44).

iii. The use of redundant expressions, such as *wa-ḏālika*, e.g. *qarrara l-wazīr bi-taʾjīl az-ziyāra wa-ḏālika li-ʾasbāb šaxṣiyya* ‘the minister decided to postpone the visit, [and this] for personal reasons’ (Ashtiany 1993:32), or *kull min*, to introduce a list of two or more names.

iv. A padding device that may have its source in the contemporary dialects is the increased use of *tābiʿ li-*, which sometimes indicates things that are under someone's control but may also be used as an expansion of *li-* without further connotations (Ashtiany 1993:122).

The effect of translation from Western languages on the language of the press should also be recognized. In the 19th century, Egyptian newspapers began to depend on foreign news agencies to obtain news (Hassan 1976). The language of the press must necessarily be up to date in order to meet the needs of the society and keep up with the political, economic, and technological changes that occurred within that society. Kanun (1983) confirms that thousands of new words and structures began to enter the language and enrich it, either through translation or metaphor. New loanwords and expressions were either translated or Arabized through derivational conventions adapted to the structure of the Arabic language. In fact, the press enriched Modern Standard Arabic with new vocabulary to express new ideas, concepts, and terms of technology by coining, adapting, or borrowing. In some cases, the introduction of new words can be pinpointed exactly, for instance the term *qamar ṣināʿī* ‘satellite’ (perhaps from English *man-made* or *artificial moon*), which was used for the first time in 1957 in an Arabic broadcast the day the first Sputnik was launched (Monteil 1960:191). This and similar words gradually became popular and an integral part of the language.

The influence of journalistic English and French was not limited to the lexicon but was also manifest in the style and → phraseology of the language of the Arab media. As far back as 1960, Monteil (1960:306–312), in his treatment of stylistic calques, especially in journalistic Arabic, referred to complaints by Arab intellectuals in the 1930s about foreign influence in expressions such as *qatala l-waqt* ‘to kill time’ and *sādat al-fawḍā* ‘anarchy reigned’. Monteil also refers to such stylistic calques in political speeches, for instance in one of President Nasser's speeches, *Miṣru kulluhā sa-tuqātilu li-ʾāxiri qaṭratin min dimāʾihā* ‘all of Egypt will fight till the last drop of blood’ (Monteil 1960:309). In modern newspaper Arabic, Ashtiany (1993:56, 61) mentions *rajul al-ʾaʿmāl* ‘*homme d'affaires*, businessman’, and a host of foreign metaphors and idioms are mentioned by Holes (1995:256), e.g. *jumhūriyyāt mawz* ‘banana republics’ (1995:275 n. 15). An interesting example is that of the use of *kubrā* in the sense of English *major*, e.g. *qaḍiyyatun kubrā* ‘a major issue’ or *al-xaṭāyā l-kubrā* ‘major errors’ (Badawi a.o. 2004:249). The influence of the foreign languages is also manifest in the use of new prepositional idioms, e.g. *iltaqā* with and without the preposition *maʿa* (Ashtiany 1993:26); *ʾaʿlana* and *ʾaʿlana ʿan* ‘to announce’; *waṣala* and *waṣala ʾilā* ‘to arrive’ (Ashtiany 1993:32).

Translation also necessitated the assimilation of grammatical structures. The use of introductory sentences and the increase of Subject-Verb-Object (SVO) sentences and prepositional phrases reflect the effect of translation. According to Parkinson (1981), SVO is the basic word order of Egyptian colloquial Arabic. His study also confirms the tendency to use SVO in the language of the newspapers. Verb-Subject-Object is considered to be the basic word order in Classical Arabic, while Subject-Verb-Object or prepositional phrases for the most part are used only for emphasis. Abdelfattah (1990) confirms the same trend toward the use of SVO sentences in Modern Standard Arabic, in *al-ʾAhrām* from 1935 to 1989. Such sentences are particularly frequent in headlines, often in the imperfect tense; in these headlines, the usual constraint against an indefinite subject (agent) is waived, e.g. *kātib ʾisbānī yuhdī l-ʾamīra ʾIlīnā ḥimārayn bi-munāsaba zafāfihā* ‘Spanish writer gives Princess Elena two donkeys on the occasion of her wedding’ (Watson 1999:170).

Note, however, that Badawi a.o. (2004:349) point out that in the body of the text the same sentence often recurs in the VSO order of Modern Standard Arabic, and this is also reported by Holes (1995:264). Nonetheless, the trend of using SVO word order brings Modern Standard Arabic closer to the colloquial dialects and to the more common structure of non-Semitic languages (Parkinson 1981; Abdelfattah 1996). According to El Hakim (1998), the noticeably high rate of nominal sentences, as well as the use of particles, conjunctions, and prepositional phrases could be considered a shift in the structure of journalistic Modern Standard Arabic in favor of a more universal structure.

In some cases, it is difficult to determine whether a certain change represents an independent innovation or a calque. This is the case with new functional expressions like *saʾala mā ʾiḏā*, which is sometimes regarded as a translation of English *whether*. About this construction Badawi a.o. (2004:721) remark "The compound is clearly not a direct calque of any Western conjunction but an indigenous innovation reproducing the semantic components of English 'whether'", as in *saʾaltuhu ʿammā ʾiḏā kuntu ʾastaṭīʿu ʾan ʾarā wālidahu* 'I asked him whether I could see his father' (Badawi a.o. 2004:721). The same may apply to the stereotypical use of *al-ʾamru llaḏī*, which has as its antecedent the entire preceding clause (Badawi a.o. 2004:513–514) and probably became popular as a handy device to translate English *which* in this function. A similar example is represented by the use of new expressions to translate the concept 'not only…but also': in Standard Arabic this should be expressed as *lā/laysa…fa-qaṭ, bal…*, but one often finds nowadays *laysa fa-qaṭ…wa-lākin (ʾayḍan)* (Badawi a.o. 2004:488). The tendency to express notions as one syntactic unit, which is apparent in this construction, may also be at the root of the use of *lā ʾaḥad* and *lā šayʾ* to translate *no one* and *nothing*. Within traditional grammatical analysis, these could be interpreted as instances of the absolute negator *lā* followed by a relative sentence, as in *min al-muʾakkad ʾan lā ʾaḥada yastaṭīʿu ʾan yalūmahu* 'it is certain that no one can blame him' instead of *ʾan lā yastaṭīʿu ʾaḥadun* (Badawi a.o. 2004:471). At a stretch, this might be interpreted as 'it is certain that there is no one who can blame him', but more likely, it is simply a translation of *no one* as one syntactic unit.

For the expression of the concept of reciprocity, newspaper Arabic increasingly uses expressions of the type *baʿḍuhum al-baʿḍ*, which are frowned upon by the official linguistic authorities. This, too, may very well be an instance of an expression that is not directly a calque but owes its popularity to the fact that it is a quick-and-ready translation of expressions like English *each other*. In the case of *ʾaḥaduhum…al-ʾāxar*, there can hardly be any doubt that it originated as a translation of French *l'un…l'autre*, e.g. *kilā l-qiṣṣatayni fī l-majmūʿ tukammilu ʾiḥdāhuma l-ʾuxrā* 'both stories in the collection complement each other' (Badawi a.o. 2004:394). Holes (1995:274 n. 9) claims that the use of *hunāka* as an existential in Modern Standard Arabic (→ locatives) was the result of massive translation from English and French.

A difficult case is that of the frequent occurrence of → passives with expressed agent in the language of the newspapers. According to the strict laws of Classical Arabic, in a passive sentence the agent cannot be expressed. Nonetheless, even in Classical Arabic, passives are sometimes connected with a preposition *bi-* or *min*, which comes very close to indicating the agent, e.g. (Q. 2/173) *man ʿufiya lahu min ʾaxīhi šayʾun* 'whoever is forgiven something by his brother' (Reckendorf 1921:233, 251–252). The frequency with which the agent of a passive verb is expressed in newspaper Arabic, however, seems to be connected with the influence of Western languages. When messages of the international news agencies are translated, these often contain an expressed agent with a passive, which tend to be translated with *min qibal* or *min ṭaraf* (cf. Holes 1995:259–260); Badawi a.o. (2004:385–386) mention alternative expressions like *bi-wāsiṭa*, *min jānib*, etc.

There are certainly innovations in newspaper Arabic that do not have their origin in foreign influence. An example is that of the use of *wa-* before a coordinated relative clause (Ashtiany 1993:32). This is one of the features of journalistic style that is mentioned explicitly in the *al-ʾAhrām* stylebook as an example of bad grammar (Nāfiʿ n.d.:64): instead of *Bank Miṣr wa-llaḏī yuʿtabaru rāʾidan fī l-iqtiṣād*, journalists should write *Bank Miṣr allaḏī yuʿtabaru* 'the Bank of Egypt, which is regarded as leading in the economy'. Other examples include the increasing use of *sawāʾan…ʾam* instead of *sawāʾan…ʾaw* ‘regardless if…or’ (Nāfiʿ n.d.:55), and the repetition of *kullamā* in expressions like *kullamā jāʾa fulān kullamā ʾakramahu* ‘each time someone came, he honored him’ (Nāfiʿ n.d.:54).

Perhaps the increased use of → compounds may also be counted as an innovation, although similar tendencies toward compounding already existed in the older language (Ashtiany 1993:58), like *radd fiʿl* ‘reaction’, and *raʾs māl* ‘capital’. Such compounds were often used in the creation of technical → terminology. Connected with this is the tendency to use certain constructions in order to translate the frequent prefixes in international technical terminology. Holes (1995:266–267) mentions the case of the prefix *re-*, which has led to numerous combinations of the verb *ʾaʿāda* ‘to repeat’ with a verbal noun, e.g. *ʾaʿāda tanḏ̣īm* ‘to reorganize’; the use of the substantive *ʿadam* with a verbal noun to translate Western *non-* or *in-*, e.g. *ʿadam al-istiqrār* ‘instability’; and the many combinations of *mutaʿaddid* to translate Western *poly-* or *multi-*, e.g. *mutaʿaddid al-jinsiyyāt* ‘multinational’.

The increased use of nominalizations in newspaper Arabic is in line with a general tendency in newspaper language all over the world to use nominal constructions. As Holes (1995:260) puts it: “Nominalisation, like passivisation, allows ‘unattributable’ claims to be made”, for instance in expressions like *hunāka iʿtiqād ʾanna…* ‘it is believed that…’, rather than an active sentence from which it is clear whose belief is at stake. Watson (1999:167) discusses the strategy of using noun phrases with nominalization in headlines, e.g. *iʿtiqāl ʿadad min ʾanṣār mawj fīʿadan* ‘arrest of a number of Mawj supporters in Aden’. Holes also mentions as possible motives for this increased use of nominal constructions the need for abstract reasoning in much of journalistic writing, and perhaps the wish to differentiate formal writing from dialect speech in which verbal phrases are more common than nominal ones (1995:261–262).

The Arabic of newspapers is a national and international pan-Arab medium of communication with limited regional variations, as opposed to the different regional dialects that impede rather than enhance the communication between them (Murgida 1993). Nonetheless, a certain amount of variation, especially in the lexical domain and in technical terminology, is apparent. Since terms in different countries may go back to different foreign languages, this may lead to → lexical variation, as in the example given by Ashtiany (1993:60) of the equivalents for ‘workforce’ in the Maghreb and the Mashreq, *al-quwā l-ʿāmila* (< English *workforce*) and *al-yad al-ʿāmila* (< French *main d’oeuvre*). The increased popularity of inter-Arab broadcasts, such as the immensely popular al-Jazeera, is difficult to assess for the moment, but a certain amount of → speech accommodation may be expected to result from this. Eventually, the popularity of such channels may lead to a more homogeneous terminology.

Although newspaper and media Arabic have been studied by several scholars, there is as yet no comprehensive account of this variety available. Two language courses focus on the variety of Arabic used in the media: Ashtiany (1993) is an exercise book teaching students how to write journalistic Arabic, whereas McCarus and Yacoub (1962) is a reader of texts taken from Arabic newspapers.

Bibliographical references

Abdelfattah, Nabil M.S. 1990. *Linguistic changes in journalistic language in Egypt, 1937–1997: A quantitative and comparative analysis.* Ph.D. diss., University of Texas at Austin.

——. 1996. “Reflections of the sociolinguistic force of journalism in the process of language development in Egypt”. *Understanding Arabic: Essays in contemporary Arabic linguistics in honor of El-Said Badawi*, ed. Alaa Elgibali, 129–136. Cairo: American University in Cairo Press.

Abu-Lughod, Ibrahim. 1963. “The mass media and Egyptian village life”. *Social Forces* 42.97–104.

Ashtiany, Julia. 1993. *Media Arabic.* Edinburgh: Edinburgh University Press.

Badawi, Elsaid [as-Saʿīd Badawī]. 1973. *Mustawayāt al-luġa al-ʿarabiyya al-muʿāṣira fī Miṣr.* Cairo: Dār al-Maʿārif.

——, Michael G. Carter, and Adrian Gully. 2004. *Modern Written Arabic: A comprehensive grammar.* London and New York: Routledge.

Diem, Werner. 19774. *Hochsprache und Dialekt: Untersuchungen zur heutigen arabischen Zweisprachigkeit.* Wiesbaden: F. Steiner.

Elhakim, Gehan H. 1998. *Sentence structure in Egyptian journalistic Arabic.* M.A. thesis, American University in Cairo.

Girod, Alain. 2000. *Faits d’évolution récents en arabe moderne à travers un corpus de presse égyptien.* Thèse de doctorat, University of Aix-en-Provence.

Glubb, John Bagot. 1959. *Britain and the Arabs.* London: Hodder and Stoughton.

Hassan, Kamel M. 1976. *Modern Standard Arabic.* Egypt: Dār al-Maʿārif.

Holes, Clive. 1995. *Modern Arabic: Structures, functions, and varieties.* 2nd ed. Washington, D.C.: Georgetown University Press. (2nd ed., 2005.)
Kanun, Abdallah. 1983. “The press and language development”. *Magazine of the Arab Academy.* Damascus.
McCarus, Ernest and Adil I. Yacoub. 1962. *Newspaper Arabic.* Ann Arbor: University of Michigan Press.
Murgida, Jacqueline. 1993. *Passive and passive-like expressions in journalistic Modern Standard Arabic.* Ph.D. diss., Georgetown University.
Nāfiʿ, ʾAḥmad. N.d. *Dalīl al-ʿamal aṣ-ṣuḥufī/Al-Ahram style book.* Cairo: [Al-Ahram].
Parkinson, Dilworth B. 1981. “VSO to SVO in Modern Standard Arabic: A study in diglossia syntax”. *Al-ʿArabiyya* 14.24–37.
Reckendorf, Hermann. 1921. *Arabische Syntax.* Heidelberg: C. Winter.
Watson, Janet. 1999. “The syntax of Arabic headlines and news summaries”. *Arabic grammar and linguistics*, ed. Yasir Suleiman, 161–181. Richmond: Curzon.

RAGIA M. EFFAT (Cairo University)
KEES VERSTEEGH (University of Nijmegen)

Mehri → Modern South Arabian

Merger → Phonological Merger

Mesopotamian Arabic → Iraq

Metaphor → Istiʿāra; Majāz

Metathesis

1. INTRODUCTION

‘Metathesis’ is a term used to describe a language sound pattern in which a sequence of sounds appears in one order in one context but in the opposite order in a related context. The transposition of sounds of this type is commonly observed in Arabic languages, as can be seen in Maltese, for example, by comparing verb forms such as [yokrob] and [yokorbu] ‘to groan [3rd pers. sg./pl. imperf.]’; the sequence [ro] occurs in the singular form of the word while the reverse order [or] is found in the plural.

In addition to the use of metathesis as a descriptive term, in generative linguistics (see e.g. Chomsky and Halle 1968), metathesis has taken on theoretical status as well. In this sense, metathesis is the term used to refer to a *rule* or *process* of reordering, i.e., the order of sounds in a word changes (or is derived) from an earlier (or underlying) form of the word that contains the opposite order. A sound change of this type is assumed for some varieties of French as a means of explaining different orders of consonants and vowels observed in words at different stages of the language’s development (Maze 1969; Spence 1990): Jersey French, *šɛvrɛt > šɛrvɛt ‘shrimp’; Le Hâvre French *fərme > frəme ‘closed, fermé’. It is likely that not all instances of metathesis are due to a process of reordering. Blevins and Garrett (1998), among others, argue that some cases of diachronic consonant-vowel metathesis are the result of syllable reduction, while others involve the insertion of a vowel identical to the preceding vowel with subsequent vowel deletion.

Metathesis is widely attested in the → Afro-Asiatic language family, of which Arabic is a member, as well as in many other language families, including Altaic, Austronesian, Salishan, Carib, Dravidian, Indo-European, Mayan, and Mixe-Zoque. While less common than processes such as → assimilation and deletion (→ elision), metathesis can be systematic within a language, and like more common sound changes, it is conditioned by phonetically natural properties, as well as by the phonological and lexical knowledge of language users.

2. MORPHOLOGICAL AND PHONOLOGICAL METATHESIS

In Arabic languages, transposition of consonants and vowels is commonly used as a means of distinguishing meaning in related words; in this sense, metathesis is *morphological* in nature. Recall the Maltese example from above: [yokrob] and [yokorbu] ‘to groan [3rd pers. sg./pl. imperf.]’. In these forms, a change in the order of the consonant and vowel is one means by which the singular form is distinguished from the corresponding plural. While the transposition of sounds in such forms is typically described as metathesis, theoretical linguists have argued that a *process* of metathesis is not required to explain the transposition. One alternative explanation involves vowel deletion and insertion (see e.g. McCarthy 1989; Hume 1994). For example, if we assume that the plural is derived from the singular form

/krob/ with the addition of the prefix /yo/ and suffix /u/, metathesis involving /ro/ can be explained with two processes: one that deletes the vowel in /krob/, giving /krb/, and a second process that inserts a vowel before the medial consonant, i.e. [korb]. In this analysis, the derivation of the plural form would involve the following changes: /yo+krob+u/ => yo+krb+u => yo+korb+u => [yokorbu].

Metathesis can also be *phonological*. In Modern Hebrew, as illustrated in (1), binyan 5 perfect verbs typically have the form /hit/ + verb, as shown in (1a) (the coronal stop /t/ agrees in voicing with an adjacent obstruent). However, when the stem-initial consonant is a sibilant, the coronal stop of the prefix /hit/ occurs to the stem-initial consonant's right, as in (b) (Bat-El 1988, 1989, 1992). Notice that unlike morphological metathesis, the transposition of sounds in (1) does not, in and of itself, result in a change in meaning. Rather, metathesis in this instance is phonological, conditioned by the quality of the consonant adjacent to the coronal stop.

(1) Modern Hebrew

a.	*hitnakem*	'he took revenge'
	hitraxetˢ	'he washed himself'
	hidbalet	'he became prominent'
	hitkabel	'it was accepted'
b.	*histader*	'he got organized'
	hizdaken	'he grew old'
	hištamer	'he preserved himself'

Bedouin Arabic provides a further example of phonological metathesis, as can be seen by comparing the verbal forms in (2a) and (2b). While both sets of verbs begin with a consonant, the order of the following two sounds differs. In (2a) we see the order vowel+consonant, e.g. *taktib* 'you write', while in (2b), the order is consonant+vowel, e.g. *nxaṭuf* 'we snatch'. The presence or absence of a guttural consonant, e.g. /ḥ/, /x/, /ˤ/, next to the initial vowel in each word is an important factor in explaining the different orders. In (2a), there is no guttural consonant flanking the first vowel, while in (2b) there is. A second key factor to understanding why the vowel follows, rather than precedes, the guttural consonant in (2b) is the observation that in Bedouin Arabic, a guttural consonant never occurs after the vowel [a] (Al-Mozainy 1981; Al-Mozainy a.o. 1985). Phonological theorists take this observation as evidence for a constraint in the language that prohibits the occurrence of a guttural consonant after /a/. Thus, while the verbal pattern in (2a) would lead us to expect the verbs in (2b) to show a word-initial CVC pattern, such a sequence is not possible since it would position the guttural consonant after /a/, a prohibited sequence. Thus, the order of consonant/vowel is reversed, creating a licit structure with the guttural preceding /a/ (→ *gahawa*-syndrome).

(2) Bedouin Arabic

a. CVCCVC	
taktib	'you write'
nasbaḥ	'we swim'
b. CCVCVC	
yḥakum	'he rules'
nxaṭuf	'we snatch'

3. The emergence of the process of metathesis

Understanding how and why metathesis occurs has interested linguists for a great many years (see e.g. Grammont 1933). The best-developed explanations are currently perceptual in nature (Blevins and Garrett 1998, 2004; Hock 1985; Hume 1998, 2004; Steriade 2001) and generally point to the importance of indeterminacy in the auditory signal as a factor necessary for metathesis to take place. Indeterminacy in this context relates to how well a listener is able to correctly identify the sounds in an utterance. The degree to which a sequence of sounds is indeterminate is influenced in part by the quality of the information in the speech signal: the better the quality, the more easily the sounds and their order can be recognized. The quality of information is, in turn, determined by the types of sounds involved and the context in which the sounds occur.

Pairs of sounds that undergo metathesis generally share one or more of the following properties. First, key phonetic cues that would aid in identifying at least one of the sounds are masked. The observation that stop consonants, as seen above for Modern Hebrew, commonly participate in metathesis relates to this point; phonetic cues such as vowel transitions and release bursts are crucial to the correct identification of the place and manner of articulation of stop consonants. When these cues are masked, as may happen when there is a following consonant, the quality of the stop consonant can be difficult to identify, thus creating indeterminacy in the signal. Second, the phonetic cues of at least one of the sounds are often relatively long in duration, as can be observed in vowels, liquids, glides, and fricatives (Blevins and Garrett 1998, 2004; Hume 2004). Cues of this nature tend to extend over a domain that may encompass adjacent sounds, leading to the overlap of important phonetic cues and resulting in potential ambiguity in a sound's onset and offset. The final property is auditory, or perceptual, similarity between the sounds involved; metathesis commonly involves sounds that are similar in terms of sonorancy, manner, or place of articulation. Similarity between sounds diminishes the distinctiveness of each of the sounds and, as a result, renders the sounds and their order less easily identifiable (Hume 1998, 2004). The observation that metathesis typically involves *adjacent* sounds also contributes to similarity given that proximity enhances similarity (Pierrehumbert 2002).

The sound patterns of an individual's language are also important in understanding the emergence of metathesis. This is particularly relevant when the speech signal is ambiguous. Indeed, the greater the indeterminacy, the more the listener will rely on familiarity with native sound patterns to infer the order of elements. This includes the particular sounds and types of sound sequences that occur in the language, as well as how often these elements are used. Hume (2004) makes two claims in this regard. First, the sound pattern resulting from metathesis must be an attested structure in the language, with the relevant structure being defined in terms of subsegmental properties of sounds (e.g. place, manner of articulation), or suprasegmental properties (e.g. syllable structure). Second, less frequent language structures are replaced by those that are more frequent within the particular language. Since languages differ in terms of their sound systems as well as in terms of the frequency with which the sound patterns are used, this approach correctly captures the observation that the result of metathesis can differ from language to language. In other words, either order of a given sound combination can emerge as the result of metathesis in some language. The case of metathesis in Modern Hebrew involving [t] + coronal sibilant (cf. Arabic *iftaʿala*) is consistent with this approach. Bat-El (1988) reports that this sequence is less common than the opposite order, which is restricted to tautomorphemic forms such as [hi-tsis] 'he fermented' and nonverbal forms like [tʃuva] 'reply'. Metatheses where the stop metathesizes to a position preceding a consonant are also attested, as discussed in Nakao (1986) for Old English and Silva (1973) for Lappish.

This view of metathesis also explains the observation that the sound sequence resulting from metathesis often has better phonetic cues than the original order. The explanation for this phenomenon draws on facts concerning the type of sounds that undergo metathesis, and the type of sequences that are most likely to influence how a language user interprets a speech signal. On the one hand, sounds with strong phonetic cues are not good candidates for metathesis since the language user would not generally have difficulty identifying the sounds or their order. On the other hand, sound sequences with weaker cues are prime canidates since there is a greater likelihood that their onsets and offsets will be ambiguous. Further, sounds with robust cues tend to be more frequent in a language system and will thus have a greater influence over how the speech signal is interpreted by the user. This then suggests that the reason why improved perceptual salience is a characteristic of so many cases of metathesis is simply an artifact of the nature of the sequences that undergo metathesis and those that influence how an indeterminate speech signal is interpreted by the language user.

There are a number of resources available to readers interested in learning more about metathesis in Arabic and other languages. A good starting point is the metathesis website at www.ling.osu.edu/~ehume/metathesis/, which contains information on about one hundred cases of metathesis, a metathesis bibliography, and other tools. Crosslinguistic surveys may also be helpful. These include Ultan (1978), Hock (1985), Blevins and Garrett (1998, 2004), Hume (1998, 2004); see the biography at the metathesis website for a more complete listing. In addition to the references throughout this entry, other sources which focus specifically on Arabic include Kenstowicz (1981), Kilani-Schoch and Dressler (1986), and Yoshida (1993).

BIBLIOGRAPHICAL REFERENCES

Al-Mozainy, Hamza Qublan. 1981. *Vowel alternations in a Bedouin Hijazi Arabic dialect: Abstractness and stress.* Ph.D. diss., University of Texas at Austin.

——, Robert Bley-Vroman, and John McCarthy. 1985. "Stress shift and metrical structure". *Linguistic Inquiry* 16:1.135–144.

Bat-El, Outi. 1988. "Remarks on tier conflation". *Linguistic Inquiry* 19:3.477–485.

——. 1989. *Phonology and word structure in Modern Hebrew.* Ph.D. diss., University of California at Los Angeles.

——. 1992. *Stem modification and cluster transfer in Modern Hebrew.* Ms., Tel-Aviv University.

Blevins, Juliette and Andrew Garrett. 1998. "The origin of consonant-vowel metathesis". *Language* 74:3.508–555.

——. 2004. "The evolution of metathesis". *Phonetically based phonology*, ed. Bruce Hayes, Robert Kirschner, and Donca Steriade, 117–156. Cambridge: Cambridge University Press.

Chomsky, Noam and Morris Halle. 1968. *The sound pattern of English.* New York: Harper and Row.

Grammont, Maurice. 1933. *Traité de phonétique.* Paris: Librairie Delagrave. (1st ed., 1905–1906.)

Hock, Hans Henrich. 1985. "Regular metathesis". *Linguistics* 23.529–546.

Hume, Elizabeth. 1994. *Front vowels, coronal consonants and their interaction in nonlinear phonology.* New York: Garland.

——. 1997. "Metathesis in phonological theory: The case of Leti". *Lingua* 104.147–186.

——. 1998. "The role of perceptibility in consonant/consonant metathesis". *Proceedings of the West Coast Conference on Formal Linguistics*, XVII, ed. Susan Blake, Eun-Sook Kim, and Kimary N. Shahin, 293–307. Stanford, Calif.: Stanford Linguistics Association, Stanford University.

——. 2004. "The indeterminacy/attestation model of metathesis". *Language* 80:2.203–237.

Kenstowicz, Michael. 1981. "Vowel harmony in Palestinian Arabic: A suprasegmental analysis". *Linguistics* 19.449–465.

Kilani-Schoch, Marianne and Wolfgang Dressler. 1986. "Métathèse et conversion morphologiques en arabe tunisien". *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 39.61–75.

Maze, Camille. 1969. *Étude sur le langage de la banlieue du Havre.* (= *Bibliothéque des dictionnaires patois de la France*, première série, X.) Geneva: Slatkine Reprints. (1st ed., Paris, 1903.)

McCarthy, John. 1989. "Linear order in phonological representation". *Linguistic Inquiry* 20.71–100.

Nakao, T. 1986. "Metathesis". *Linguistics across historical and geographical boundaries: In honour of Jacek Fisiak on the occasion of his fiftieth birthday*, ed. Dieter Kastovsky and Aleksander Szwedek, 547–556. New York: Mouton de Gruyter.

Pierrehumbert, Janet. 2002. "Dissimilarity in the Arabic verbal roots". *Proceedings of the 23rd Meeting of the Northeastern Linguistic Society*, 367–381. Amherst: Graduate Student Linguistic Association, University of Massachusetts at Amherst.

Silva, Clare M. 1973. "Metathesis of obstruent clusters". *Working Papers in Linguistics* 14.77–84.

Spence, Nicole. 1990. "'Sporadic' changes in Jersey French". *Variation and change in French: Essays presented to Rebecca Posner on the occasion of her sixtieth birthday*, ed. John N. Green and Wendy Ayres-Bennett, 210–225. New York and London: Routledge.

Steriade, Donca. 2001. "Directional asymmetries in assimilation". *The role of speech perception in phonology*, ed. Elizabeth Hume and Keith Johnson, 219–278. San Diego, Calif., and London: Academic Press.

Ultan, Russell. 1978. "A typological view of metathesis". *Universals of human language*, II, ed. Joseph Greenberg, 368–399. Stanford, Calif.: Stanford University Press.

Yoshida, Shohei. 1993. "Licensing of empty nuclei: The case of Palestinian vowel harmony". *Linguistic Review* 10.127–159.

ELIZABETH HUME (Ohio State University)

Meter

1. THE THEORY OF METER

Arabic versification (*ʿarūḍ*) is quantitative, a unique phenomenon among the Semitic languages, where accent verse systems dominate. Its theory was developed by al-Xalīl ibn ʾAḥmad (d. 175/791 or 170/786 or 160/776), who was also a prominent figure in the genesis of Arabic linguistic science. The classical theory postulates sixteen meters, although the number of meters used in poetic practice does not exceed a dozen.

The pioneers of the study of *ʿarūḍ* in Europe, such as William Jones (1777), Georg Wilhelm Freytag (1830), and Georg Ewald (1825–1854), saw it as a purely quantitative system, but their position was put to question later on. In spite of the fact that neither Arabic grammar, nor Arabic metrical theory, nor the science of Qurʾānic recitation (→ *tajwīd*) mention anything about accent, the question of the nature of Arabic versification became a controversial one. The most daring attempts to prove the accent nature of Arabic verse were made by

Stanislas Guyard (1877), who was followed in Russia by David Gincburg (1893–1897) and Gotthold Weil (1958). Still, in spite of their ingenious arguments, their theories did not hold. At the end of the 20th century, several scholars independently and almost simultaneously reaffirmed the quantitative character of Arabic metrics (Stoetzer 1989; Frolov 1991, 2000; Bohas and Paoli 1992, 1997).

A fitting start for an exposition of the system of *ʿarūḍ* is the following passage, in which it is explained succinctly and clearly: "The verse lines (*ʾabyāt*) are composed (*murakkaba*) of hemistichs (*maṣārīʿ*), the hemistichs are composed of feet (*tafāʿīl*), the feet are composed of *ʾasbāb*, *ʾawtād*, and *fawāṣil*, and the root (*ʾaṣl*) of all these are vocalized and unvocalized *ḥurūf*" (*Rasāʾil* I, 197).

2. Level one: *Ḥarf* (mora)

"You should know that the student of *ʿarūḍ* must begin with the knowledge of unvocalized and vocalized *ḥurūf*" (Ibn ʿAbd Rabbihi, *ʿIqd* V, 424; cf. also ʾAxfaš, *ʿArūḍ* 135; Ibn as-Sarrāj, *Miʿyār*).

The rhythmic core of Arabic meter is the alternation of longer and shorter prosodic segments. The central notion of Arabic metrical theory is that of → *ḥarf*, which is the most convenient instrument for the analysis of the structure of Arabic verse, by far surpassing in that capacity the traditional European notion of 'syllable'. The inadequate character of the notion of syllable, especially the short syllable, is stressed by Stoetzer (1989:121–123). The *ḥarf* has a twofold function in the metrical structure.

First, it is used as the means of the segmentation of speech into elementary prosodic units (EPUs) produced by the alternation of vocalized *ḥurūf* (*ḥarf mutaḥarrik*) and unvocalized *ḥurūf* (*ḥarf sākin*), the terms themselves conveying the idea of a prosodic inertia (→ *ḥaraka*) and a prosodic pause (*sukūn*; Sančes 1968). This goal is achieved by stops at each unvocalized *ḥarf*, and as a result, each EPU has an identical structural pattern: a number (one or more) of vocalized *ḥurūf* and a final unvocalized *ḥarf*. All linguistic phenomena occurring in ordinary language, such as consonant clusters, resulting from what is called in Arabic grammar *iltiqāʾ as-sākinayn* 'the meeting of two unvocalized *ḥurūf*', or → pausal word forms, and even speech pauses themselves, were eliminated altogether from the poetic diction because they might have distorted this unified pattern of the EPU. Partly because of this, a word like *šābb* 'youth, young man', which conveys a notion very common in the poetic imagery, is never used in poetry and is substituted by its synonym *fatā*.

Second, *ḥarf* is a unit of measurement of prosodic length, which may be regarded as the functional equivalent of the universal notion of 'mora'.

3. Level two: Metrical syllables (EPU)

"Then come *ʾasbāb* and *ʾawtād*, and feet (*ʾajzāʾ*) are composed of them" (Ibn as-Sarrāj, *Miʿyār*).

The theory of *ʿarūḍ* speaks of two-*ḥarf*, three-*ḥarf*, four-*ḥarf*, and five-*ḥarf* EPUs as building elements of the verse structure. The two-*ḥarf* element, which is the minimum for the EPU of the above pattern, is called *sabab* lit. 'rope'; the three-*ḥarf* element is called *watid* lit. 'peg'; and the four-*ḥarf* and five-*ḥarf* elements are called smaller and larger *fāṣila* lit. 'cut, section, segment'.

A mnemonic phrase which probably goes back to al-Xalīl himself helps one to memorize the whole set (each word stands for one type of EPU postulated by the theory): *lam ʾara ʿalā ḍahri jabalin samakatan* (or *šajaratan*) 'I have not seen a fish [or: a tree] on the slope of a mountain'.

Incidentally, the set of the EPU, which is the basis of the *ʿarūḍ* rhythm, coincides with the set of the Arabic root patterns (*ʾuṣūl*; → *ʾaṣl*), which can consist of two to five *ḥurūf* at the most. This coincidence is not surprising as both systems go back to the same person, al-Xalīl, as the *Kitāb al-ʿayn* attributed to him shows (Frolov 2000).

The similarity between the two sets of *ʾuṣūl*, linguistic and metrical, conveys traces of deliberate planning, because the longest EPU (the larger *fāṣila*) in fact never occurs in ideal metrical schemes but only as a transform, and has to be regarded not as a 'root' (*ʾaṣl*) or primary part of the structure but rather as a 'branch' (*farʿ*) or secondary part. Moreover, the theory introduces the prosodically impossible notion of the 'heavy *sabab*' (symbolized by *ʾara* in the above phrase), which consists of two vocalized *ḥurūf* and thus cannot be considered an autonomous element of the prosodic structure.

The distribution of this heavy *sabab* – only before the normal, 'light' *sabab* – leads to the conclusion that the four-*ḥarf* segment is to be treated not as an elementary but rather as a complex entity, made up of two *ʾasbāb* (2 + 2 = 4), with the eliminated prosodic stop between them, or a kind of 'syllabic cluster'. This conclusion was actually drawn by some Arabic theoreticians, who postulated that Arabic verse is made up only of *ʾasbāb* and *ʾawtād*, while all the rest are complex entities produced by their combination and belonging to the 'branch' or secondary part of the structure (cf. Ibn Rašīq, *ʿUmda* I, 138). This view is mirrored in a beautiful legend about an aged *šayx*, who taught a youth to compose poems by saying that all verse is made up of *naʿam* 'yes' (three-*ḥarf* segment, *watid*) and *lā* 'no' (two-*ḥarf* segment, *sabab*).

The last element invented by al-Xalīl in order to solve some theoretical difficulties is discussed later in this entry. It is a special kind of *watid*, the so-called *watid mafrūq* 'disjointed or dissociated *watid*' (symbolized by *ḍahri*), the name itself being significant. Although it consists of three *ḥurūf*, its structure – *lātu*, or 21 – is prosodically impossible because the prosodic borderline cuts it in half, and so it can be neither heard nor scanned. In other words, it is not a prosodic unity.

Frolov (1991, 2000) proposed to call the EPUs, which have no generic name in Arabic, 'metrical syllables' because of their function as the building elements of the *ʿarūḍ* feet. The term covers only those types which are prosodically real: two-, three-, and four-*ḥarf* segments of the normal structure, in other words mostly *ʾasbāb* and *ʾawtād*. In this case, the larger *fāṣila*, or both of them, can be called 'syllabic clusters'.

4. Level three: *ʿarūḍ* feet (Arabic *juzʾ* 'part' or *tafʿīla*)

Metrical theory distinguishes eight different types of feet, but only seven of them are real. Their symbolic representation is of two kinds. The first kind gives the scheme of the foot on the level of *ḥurūf* and uses two signs, one for the vocalized *ḥarf* (for instance, the symbol o), another for the unvocalized *ḥarf* (for instance, the symbol /). This representation is used mostly in the famous 'circles' (*dawāʾir*) of al-Xalīl (see below). The second kind gives the scheme of the foot on the level of 'metrical syllables' and uses artificial words derived from the root *f-ʿ-l*, similar to, but not identical with, the well-known morphological device for the representation of word models, both being the product of al-Xalīl's genius. This second representation, which has the benefit of making the rhythm heard by the mere repetition of the foot pattern, is the main one in the treatises on *ʿarūḍ*. It is used here with an additional numerical symbolic representation proposed by Sančes (1968): *sabab* – 2; *watid* – 3; *fāṣila* – 4.

The maximum length of the *ʿarūḍ* foot is seven *ḥurūf*, which is also the maximum length of the word models in Arabic, as al-Xalīl and his pupil Sībawayhi established. The feet are divided according to their length into five-*ḥarf* type (*xumāsī*) and seven-*ḥarf* type (*subāʿī*). The first type comprises two foot patterns: *faʿūlun*, or 32, and *fāʿilun*, or 23. The second type comprises six patterns, including a fictitious one. The five real feet are *mafāʿīlun*, or 322; *mufāʿalatun*, or 34; *fāʿilātun*, or 232; *mustafʿilun*, or 223; *mutafāʿilun*, or 43.

Note that each foot consists of one and only one *watid* and a variable number of *ʾasbāb*, one or two, or a *fāṣila*. The Arabic metrical scholars stress that the *watid* is the core of the rhythm, and as such it is not affected by any deformations in the real verse. It turns out that Weil (1958) was basically correct when he applied the terms *arsis* and *thesis* to the structure of Arabic verse and identified the place of the *arsis* with the *watid* segment. The difference is that the basis of the rhythm is purely quantitative, and *watid* is not the place of *ictus*, as Weil thought, but simply the longer segment of the two basic 'metrical syllables'.

Weil's idea of dividing rhythms of Arabic verse into rising and falling also proves very fruitful, provided it is carried out without any connection with the accent. Starting from the position of the *arsis* (*watid*) in the foot, three basic rhythms in the Arabic verse may be discerned: descending, either trochaic (32) or dactylic (322); ascending, either iambic (23) or anapestic (223); ambivalent, ascending-descending, or amphibrachic (232).

All three rhythms are found in the archaic, pre-*ʿarūḍ* stage of the development of the

Arabic verse and may be associated with different forms of ancient poetry, which were given names already in the time of the Jāhiliyya: descending rhythms were designated with the name *hazaj*, ascending rhythms with → *rajaz*, ambivalent rhythms with *ramal*. Two of them were originally associated with singing, *hazaj* with the autochthonous Arab tradition, now practically extinct, *ramal* with the Persian tradition imported via al-Ḥīra, *rajaz* with declamation (Frolov 2000).

The anomalous eighth foot is represented symbolically as *mafʿūlātu*, or 2221–. The first thing to notice is that it is prosodically incomplete and thus not autonomous, unlike the other foot patterns. Secondly, the foot does not seem to contain a *watid*, and so its structure is totally alien to the nature of Arabic verse. In short, it violates practically all rules of *ʿarūḍ* metrics.

In order to solve this complication, al-Xalīl invented the strictly conventional notion of the *watid mafrūq* 'disjointed or dissociated *watid*'. The Arabic metrical scholars stress that this foot has a very unusual distribution. Unlike all other feet, for instance, it cannot make a meter on its own, nor can it occur more than once in a verse.

Turning from theory to practice, one finds that the anomalous foot is entirely fictitious, as it practically never occurs in any real verse patterns. Of the three meters in which, according to the theory, this foot can be found, one (*sarīʿ*) never realizes it in its theoretical form and shows no anomalies, and the second (*muqtaḍab*) is a fictitious meter itself, never used by Arab poets. Only one meter (*munsariḥ*) exhibits metrical schemes that can be analyzed with the use of this strange foot; it does allow for an alternative interpretation, however, in conformity with the *ʿarūḍ* rules, except one. The middle part of the *munsariḥ* may exhibit a shift of the real *watid*, the core of the rhythm; if one takes this one step further, one has to admit for the prosodically oriented analysis of the structure the possibility of a nine-*ḥarf* foot (2223). This shift can occur in the verse only once, and only in the meter of the ascending, most archaic, rhythm. This explains the above statements by metrical scholars.

Al-Xalīl chose another solution. For him, the correspondence between the maximum length of the word model and verse foot (seven *ḥurūf*) was so important that he preferred to keep it untouched. He did this by introducing a whole set of theoretical notions that had no relation to the prosodic and rhythmic reality of the Arabic verse, and in fact make the theory rather cumbersome (for details see Frolov 2000).

5. Level four: Meters

Al-Xalīl himself distinguished a total of fifteen meters; the sixteenth meter, *mutadārik* (which also goes under different names), was added by his pupil and rival al-ʾAxfaš (d. 215/830?). This meter retained its place in theoretical expositions of *ʿarūḍ* but never became a reality, as Arab poets never used it.

The order of the meters in metrical treatises has been fixed once and for all, as well as the poetical *šawāhid* that exemplify them, and many seem to have been invented by al-Xalīl, or by others shortly after him. The meters are arranged in circles. The circle pattern always gives the longest possible, though not necessarily the most frequent or ancient form of the meter, which is considered to be the 'root' of all shorter forms. The ideal meter in the circle (often called *baḥr*) generates a family of concrete meters, characterized by length (number of feet) and the form of the final foot of the first hemistich (called *ʿarūḍ* like the name of the discipline) and the second hemistich (called *ḍarb*), which is the place of the → rhyme (*qāfiya*). These final feet may have a rhythm of their own, regarded as deformations of the basic rhythm, but in fact playing a role as end markers of the first hemistich and of the whole line, making the quantitative rhythm slightly less monotonous. In al-Xalīl's system, the term → *ʿilla* is used for this phenomenon, consisting of the deletion of an element. Bohas and Paoli (1997:156ff.) distinguish these deviations, which are confined to the last feet of the hemistich, from those called *ziḥāfa* in Arabic metrical theory, which only take place in the interior of a verse (*ḥašw*). The difference between them is that the *ʿilal* are obligatory in the sense that they apply to all verses of the poem, unlike the *ziḥāfa*; besides, the latter can only occur in the *sabab* of a verse, whereas the *ʿilal* may occur in both *sabab* and *watid*. This part of the theory is not dealt with here, but the initial pattern of the meter is presented with some additional basic information. Deviations of the sample *bayt* from the ideal scheme are not specially mentioned. For detailed analytical tables of the frequency of meters, see Frolov (2000:259–290).

The first circle

(1) *ṭawīl* 'long'. An eight-foot meter, and the only one which does not have shorter forms. It is the main meter of *qaṣīd* poetry, almost half of which is written in this meter.

Scheme

faʿūlun / mafāʿīlun / faʿūlun / mafāʿīlun //
32 322 32 33
faʿūlun / mafāʿīlun / faʿūlun / mafāʿīlun
32 322 32 322

Example (Ṭarafa)

ʾabā mun / ḏirin kānat / ġurūran / ṣaḥīfatī
fa-lam ʾuʿ / ṭikum fi-ṭ-ṭaw /ʿi mālī /wa-lā ʿirḍī

'Abū Mundhir, my face was arrogant,
And I did not voluntarily give you my wealth or my honor'

(2) *madīd* 'stretched, extended'. Theoretically an eight-foot meter, but used exclusively as a shorter six-foot variation. The frequency of *madīd*, absent from many poetic *dīwān*s, rarely exceeds 1 to 3 percent.

Scheme

fāʿilātun / fāʿilun / fāʿilātun //
232 23 232
fāʿilātun / fāʿilun / fāʿilātun
232 23 232

Example (Muhalhil)

yā la-bakrin / ʾanširū / lī kulayban
yā la-bakrin / ʾayna ʾay / na-l-firārū
'O Bakrites, bring Kulayb back to life for me!
O Bakrites, where, where to flee?'

(3) *basīṭ* 'outspread, unfolded'. An eight-foot meter, but with a shorter six-foot variation. It is one of the four main *qaṣīd* meters, whose frequency is 10 to 20 percent, rarely more.

Scheme

mustafʿilun / fāʿilun / mustafʿilun / fāʿilun //
223 23 223 23
mustafʿilun / fāʿilun / mustafʿilun / fāʿilun
223 23 223 23

Example (Zuhayr)

yā ḥāri lā / ʾurmayan / minkum bi-dā / hiyatin
lam yalqahā / sūqatun / qablī wa-lā / malikū
'O Ḥārith, let me not be hit from you by a blow,
Which never hit before either subjects or kings'

The second circle

(4) *wāfir* 'abundant, full'. A six-foot meter, which also has a rare four-foot variation. It is one of the four main *qaṣīd* meters, whose frequency is 7 to 20 percent.

Scheme

mufāʿalatun / mufāʿalatun / mufāʿalatun //
34 34 34
mufāʿalatun / mufāʿalatun / mufāʿalatun
34 34 34

Example (Imruʾulqays)

lanā ġanamun / nusawwiquhā / ġizārun
ka-ʾanna qurū / na jillatihā / ʿiṣiyyū
'We have numerous sheep for sale,
horns of the biggest are like poles'

(5) *kāmil* 'complete'. A six-foot meter, which also has a four-foot variation. It is one of the four main *qaṣīd* meters, whose frequency is 10 to 25 percent.

Scheme

mutafāʿilun / mutafāʿilun / mutafāʿilun //
43 43 43
mutafāʿilun / mutafāʿilun / mutafāʿilun
43 43 43

Example (ʿAntara)

wa-ʾiḏā ṣaḥaw / tu fa-mā ʾuqaṣ / ṣiru ʿan nadan
wa-kamā ʿalim / ti šamāʾilī / wa-takarrumī
'When I sober up, I do not stop being generous;
You know well my virtues and my nobleness'

The third circle

(6) *hazaj* 'quick vibration of sound'. Theoretically a six-foot meter, but used exclusively as a shorter, four-foot variation. It is a rare meter, rhythmically close to *wāfir*, whose frequency does not exceed 1 to 2 percent.

Scheme

mafāʿīlun / mafāʿīlun // mafāʿīlun / mafāʿīlun
322 322 322 322

Example (Ṭarafa)
ʿafā min ʾā / li layla-s-sah
bu fa-l-ʾamlā / ḥu fa-l-ġamrū
‘No trace of Laylā’s clan in as-Sah-
b or al-ʾAmlāḥ or al-Ġamr’

(7) *rajaz* ‘trembling sound, murmur’. One of the oldest meters, used from time immemorial exclusively as a short three-foot (or even a two-foot) variation; in the theory of *ʿarūḍ* it is present as a ‘normal’ six-foot meter, which emerged not earlier than the Abbasid times and never became popular. The short variation produced in the Umayyad period a quasi-*qaṣīda* type of poetry called *ʾurjūza*. The meter is not frequent in the poetic *dīwāns* but occurs very frequently in historical and linguistic pedagogical writings. Rhythmically, it is close to *kāmil*.

Scheme
mustafʿilun / mustafʿilun / mustafʿilun //
223 223 223
mustafʿilun / mustafʿilun / mustafʿilun
223 223 223

Example (anonymous)
dārun li-sal / mā ʾiḏ sulay / mā jāratun
qafrun tarā / ʾāyātihā / miṯla-z-zubur
‘A campsite of Salmā with Sulaymā in the neighborhood
Is empty, but its signs for you are like a book’

(8) *ramal* ‘woven cloth; sound of rain drops’. A six-foot meter, which also has a four-foot variation. Its frequency rarely exceeds 5 percent.

Scheme
fāʿilātun / fāʿilātun / fāʿilātun // fāʿilātun /
232 232 232 232
fāʿilātun / fāʿilātun
232 232

Example
miṭla saḥqi-l / -burdi ʿaffā / baʿdaka-l
-qaṭru maġnā / hū wa-taʾwī / bu-š-šīmālī
‘His dwelling place was made after you like a torn
Cloak by rains and blasts of northern winds’

The fourth circle

(9) *sarīʿ* ‘quick’. A six-foot meter, but also having a short, three-foot variation, which definitely shows its affinity to *rajaz*. In metrical theory, it is presented as having at the end the anomalous prosodically incomplete foot (see above), but in reality, it never has an ideal form, and the meter turns into a specific variation of *rajaz*, nothing more. Its frequency rarely exceeds 5 percent. Unlike previous meters, the deviation of the sample verse from the ideal scheme is shown here.

Scheme
mustafʿilun / mustafʿilun / mafʿūlātu- //
223 223 **2221–**
mustafʿilun / mustafʿilun / mafʿūlātu-
223 223 **2221–**

Example (anonymous)
ʾazmāna sal /mā lā yarā / miṯlaha-r //
223 223 **23**
-rāʾūna fī / šāmin wa-lā / fī ʿirāq
223 223 **23+1**
‘The time of Salmā! Nobody sees // a similar time, neither in Syria, nor in Iraq’

(10) *munsariḥ* ‘free, easygoing, unbound’. A six-foot meter, but having a short, two-foot variation, which also shows its affinity to *rajaz*. In metrical theory, it is presented as having the anomalous foot in the middle of the hemistich, but the prosodically relevant scanning shows that it is another specific variation of *rajaz* with a *watid* shifted one step further. Its frequency rarely exceeds 4 percent. Both ways of scanning are shown in this case.

Scheme
mustafʿilun / mafʿūlātu- /mustafʿilun //
223 2221– 223
223 2223 23
mustafʿilun / mafʿūlātu- /mustafʿilun
223 2221– 223 (theory)
223 2223 23 (reality)

Example (anonymous)
ʾinna-bna zay / din lā zāla / mustaʿmilan
li-l-xayri yuf / šī fī miṣri / hi-l-ʿurufā
‘Ibn Zayd is still a deputy
of good and spreads favors in his capital’

(11) *xafīf* 'light'. A six-foot meter, which also has a four-foot variation. The sequence of three *watids* in a row does not present a theoretical problem here, as they can be made part of different feet. Therefore, the interpretation with the help of the notion of the anomalous *watid* was not very popular among metrical scholars. Its frequency is about 10 percent, although with some poets it can rise up to 15 or even 20 percent.

Scheme
fāʿilātun / mus-tafʿi-lun [or simply
232 223
mustafʿilun] / *fāʿilātun // fāʿilātun /*
232 232
mus-tafʿi-lun / fāʿilātun
223 232

Example (al-ʾAʿšā)
ḫalla ʾahlī / mā bayna dur / nā fa-badaw
lā wa-ḫallat / ʿulwiyyatun / bi-s-sixālī
'My kin settled between Durnā and Badaw-lā and ʿUlwiyya settled in as-Sixāl'

(12) *muḍāriʿ* 'similar'. Theoretically a four-foot meter, but never used by poets. Again as in the case of *xafīf*, the sequence of three *watids* does not present a theoretical problem here, as they can be made part of different feet.

Scheme
mafāʿīlun / fāʿilātun // mafāʿīlun / fāʿilātun
322 232 322 232

Example (anonymous)
daʿānī ʾi / lā suʿādin // dawāʿī hawā suʿādī
'I was attracted to Suʿād // by my passion for Suʿād'

(13) *muqtaḍab* 'cut or torn off'. A four-foot meter, the third one for which al-Xalīl postulated the occurrence of the anomalous foot, which is placed in the beginning of each hemistich. Because it never existed in poetic practice, it is not necessary to enter into details here.

Scheme
mafʿūlātu-mustafʿilun // mafʿūlātu-
2221– 223 (or: 2223 23 // 2223 23)
mustafʿilun

Example (anonymous)
ʾaqbalat fa l-lāḥa lahā // ʿāriḍāni / ka-l-baradī
'She came up and showed // her teeth which were like hailstones'

(14) *mujtaṯṯ* 'cut or carved off'. A four-foot meter whose frequency oscillates around 1 percent. It presents no theoretical problems, as the sequence of the three *sababs* exists only in the structure of the circle and nowhere else.

Scheme
mustafʿilun / fāʿilātun // mustafʿilun /
223 232 223
fāʿilātun
232

Example (anonymous)
ʾal-baṭnu min / hā xamīṣun // wa-l-wajhu miṯ / lu-l-hilālī
'Her abdomen is hollow // And her face is the moon'

The fifth circle

(15) *mutaqārib* 'contracted or drawn near each other'. An eight-foot meter made only of five-*ḥarf* feet, which is unusual for Arabic verse, hence its name. Its frequency is 5 to 15 percent, and it is one of the six most popular meters of Classical Arabic poetry. Its short six-foot variation is much less frequent than the long one.

Scheme
faʿūlun / faʿūlun / faʿūlun / faʿūlun //
32 32 32 32
faʿūlun / faʿūlun / faʿūlun / faʿūlun
32 32 32 32

Example (Bišr ibn ʾAbī Xāzim)
fa-ʾammā / tamīmun / tamīmu-b / nu murrin
fa-ʾalfā / humu-l-qaw / mu rawbā / niyāmā
'As for the Tamīmites of Tamīm ibn Murr,
People found them drunk, sleepy'

(16) *mutadārik* ‘following closely each other’ or *muḥdaṯ* ‘newly invented’. Not originally included in the theory of al-Xalīl but added later by al-ʾAxfaš. It is postulated as an eight-foot meter, which is difficult to sustain or refute, as the meter was never practiced by the Arab poets.

Scheme
fāʿilun / fāʿilun / fāʿilun / fāʿilun // fāʿilun /
23 23 23 23 23
fāʿilun / fāʿilun / fāʿilun
23 23 23

Example (anonymous)
jāʾanā / ʿāmirun / sāliman / ṣāliḥan
baʿda mā / kāna mā / kāna min / ʿāmirī
‘ʿĀmir came to us healthy and safe
After all that happened to ʿĀmir’

6. Level five: Classification of meters and circles

Arabic philologists proposed several principles for the classification of meters. According to one of these, the meters are divided into two groups: homogeneous (*mufrada* or *basīṭa*), which are made of feet of only one type, and heterogeneous (*murakkaba*), which are made up of more than one type of feet (Ibn as-Sarrāj, *Miʿyār* 16; Jawharī, *ʿArūḍ* 11; Zamaxšarī, *Qisṭās* 75–77). This classification can be useful in comparative studies.

Another classification divides types (variations) of meters according to their length. Thus, as-Sakkākī (*Miftāḥ* 523–524) speaks of meters composed of eight, six, four, three, two, and one feet. The last three extrashort types are used only in *rajaz* (and in *sarīʿ* and *munsariḥ*, which are closely related to *rajaz*), which proves the very archaic nature of the *rajaz* verse. The developed structure of the *ʿarūḍ* verse is based on the extralong (eight-foot), long (six-foot), or short (four-foot) variations, the first type being a purely Arabic innovation, as other ancient Semitic versifications know only shorter forms.

European scholars added to these classifications of meters a classification according to the underlying rhythms. Hartmann (1896) considered Arabic poetry to be monorhythmic, based on one ascending (iambic) rhythm, which he correctly associated with *rajaz*. He attempted to derive all meters from this basic rhythm as its variations. Weil (1958) regarded the *watid* as the core of the verse rhythm and based his classification on the two types of *watids* (one of them described above as theoretical only). Accordingly, he postulated two basic rhythms in Arabic verse, rising and falling. In his early work, Bohas (1974) proposed a tripartite classification of Arabic meters according to the place of the *watid* at the beginning, the middle, or the end of the foot, in some respects similar to the classification proposed in Frolov (2000).

The classification of the three basic rhythms proposed above leads to the classification of the *ʿarūḍ* meters into three metrical families: that of the descending rhythm (*ṭawīl*, *wāfir*, *hazaj*, *mutaqārib*, *muḍāriʿ*), that of the ascending rhythm (*basīṭ*, *kāmil*, *rajaz*, *sarīʿ*, *munsariḥ*, *mujtaṯṯ*, *muqtaḍab*, *mutadārik*), and that of the intermediate, ascending-descending rhythm (*madīd*, *xafīf*, *ramal*). This classification proved very fruitful in the analysis of the genesis of the *ʿarūḍ* verse and the study of the metrical repertory of Arabic poetry from the classical period (Frolov 2000).

The original classification of meters incorporated in the *ʿarūḍ* theory by its author is that of the circles, which are often called ‘mysterious’. The division of the meters into the five circles has been shown above. The circle represents the ordered sequence of the *ʾawtād* and the *ʾasbāb* (or *fawāṣil*), which has neither a beginning nor an end, and which generates different meters according to the point of departure. This classification has nothing to do with rhythm, as every circle comprises meters of different rhythms. There are reasons to believe that the device of the circles apart from its evident mnemonic value (often combined with similar names of meters belonging to the same circle, with the exception of the anomalous fourth circle) was a means to fix the rhythmic law discovered by al-Xalīl, which is based on ordered alternation of the metrical elements (EPUs) of different length. The first three circles constitute the center of the system: the first two comprise all four main meters (*ṭawīl*, *basīṭ*, *wāfir*, *kāmil*), the third comprises three archaic forms of Arabic verse representing the system of the three basic rhythms which were made regular *ʿarūḍ* meters (*hazaj*, *rajaz*, *ramal*). The fourth and the fifth circles are the periphery of the system and deal with metrical schemes that are in some way anomalous, or fictitious. The system of the circles is, in a manner of speaking, the 'trademark' of al-Xalīl's system, since it can produce not only 'used' (*mustaʿmal*) or recognized metrical schemes, but also 'neglected' (*muhmal*) schemes, which are sometimes referred to in metrical treatises. In this respect, the circles resemble the device of root derivation that is practiced in the lexicon attributed to al-Xalīl, *Kitāb al-ʿayn*.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

ʾAxfaš, *ʿArūḍ* = ʾAbū l-Ḥasan Saʿīd ibn Masʿada al-ʾAxfaš al-ʾAwsaṭ, *Kitāb al-ʿarūḍ*. Ed. Sayyid al-Baḥrāwī. *Fuṣūl* 6 (1986): 125–160.

Ibn ʿAbd Rabbihi, *ʿIqd* = ʾAbū ʿUmar ibn Muḥammad Ibn ʿAbd Rabbihi, *al-ʿIqd al-farīd*. 3rd ed. 7 vols. Cairo, 1965.

Ibn Jinnī, *ʿArūḍ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *Kitāb al-ʿarūḍ*. Ed. Ḥasan Šāḏilī Farhūd. Beirut, 1972.

Ibn Rašīq, *ʿUmda* = ʾAbū ʿAlī al-Ḥasan Ibn Rašīq al-Qayrawānī, *al-ʿUmda fī maḥāsin aš-šiʿr wa-ʾādābihi wa-naqdihi*. Ed. Muḥammad Muḥyī d-Dīn ʿAbd al-Ḥamīd. 4th ed. 2 vols. Beirut: Dār al-Jīl, 1972.

Ibn as-Sarrāj, *Miʿyār* = ʾAbū Bakr ibn as-Sarī Ibn as-Sarrāj, *al-Miʿyār fī ʾawzān al-ʾašʿār*. Ed. Muḥammad Riḍwān ad-Dāya. 3rd ed. Beirut, 1979.

Jawharī, *ʿArūḍ* = ʾAbū Naṣr ʾIsmāʿīl al-Jawharī, *ʿArūḍ al-waraqa*. Casablanca, 1984.

Rasāʾil = *Rasāʾil ʾIxwān aṣ-ṣafā*. 4 vols. Beirut, n.d.

Sakkākī, *Miftāḥ* = ʾAbū Yaʿqūb Yūsuf ibn ʾAbī Bakr as-Sakkākī, *Miftāḥ al-ʿulūm*. Beirut, 1983.

Tibrīzī, *Kāfī* = ʾAbū Zakariyyāʾ Yaḥyā ibn ʿAlī al-Xaṭīb at-Tibrīzī, *al-Kāfī fī l-ʿarūḍ wa-l-qawāfī*. Ed. al-Ḥassānī Ḥasan ʿAbdallāh. Cairo. 1977.

Xalīl, *ʿAyn* = ʾAbū ʿAbd ar-Raḥmān al-Xalīl ibn ʾAḥmad al-Farāhīdī, *Kitāb al-ʿayn*, I. Baghdad, 1980.

Xazrajī, *Xazrajiyya* = ʿAbdallāh ibn ʿUṯmān al-Xazrajī, *al-Xazrajiyya*. Ed. René Basset, *La Khazrajiyah: Traité de métrique arabe par Ali el Khazraji*. Algiers, 1902.

Zamaxšarī, *Qisṭās* = ʾAbū l-Qāsim Maḥmūd ibn ʿUmar az-Zamaxšarī, *al-Qisṭās al-mustaqīm*. Baghdad, 1969.

SECONDARY SOURCES

Bohas, George. 1974. "La métrique arabe classique". *Linguistics* 140.58–68.

——. and Bruno Paoli. 1992. "Rythme et mètre en poésie arabe". *De la musique à la linguistique: Hommages à Nicolas Ruwet*, ed. Liliane Tasmowski and Anne Ribi-Hertz, 39–49. Ghent: Communication and Cognition.

——. and Bruno Paoli. 1997. *Aspects formels de la poésie arabe*. Toulouse: Amam.

Ewald, Georg. 1825–1854. *De metris carminum arabicorum libri duo*. Brunsvigae: Sumtibus Ludovici Lucius.

Freytag, Georg Wilhelm. 1830. *Darstellung der arabischen Verskunst*. Bonn: Cnobloch.

Frolov, Dmitry. 1991. *Klassičeskij arabskij stix: Istorija i teorija aruda* [Classical Arabic verse: History and theory of *ʿarūḍ*]. Moscow: Nauka.

——. 2000. *Classical Arabic verse: History and theory of* ʿarūḍ. Leiden: E.J. Brill.

Gincburg, David. 1893–1897. "Osnovy arabskogo stixosloženija [Fundamentals of Arabic versification]". *Zapiski Vostočnogo Otdelenija Imperatorskogo Russkogo Arxeologičeskogo Obščestva*, VII, IX, X.

Guyard, Stanislas. 1877. "Théorie nouvelle de la métrique arabe, précédée de considérations générales sur le rhythme naturel du langage". Paris: Imprimerie Nationale. [Extract from *Journal Asiatique*, 7th series, vols. VII, VIII, X.]

Hartmann, Martin. 1896. *Metrum und Rhythmus: Die Entstehung der arabischen Versmasse*. Giessen: Ricker.

Jones, William. 1777. *Poeseos Asiaticae commentariorum libri sex*. Lipsiae.

Sančes [Sánchez], Alcaén. 1968. "K voprosu o suščnosti sistemy arabskoj metriki [Concerning the essence of the system of Arabic metrics]". *Arabskaja filologija: Sbornik statej*, ed. Aleksandr Aleksandrovič Kovalev and Gračija Mikajelovič Gabučan, 86–95. Moscow: University Press.

Stoetzer, Willem. 1989. *Theory and practice in Arabic metrics*. Leiden: Nederlands Instituut voor het Nabije Oosten.

Weil, Gotthold. 1958. *Grundriss und System der altarabischen Metren*. Wiesbaden: O. Harrassowitz.

——. 1979. "*ʿArūḍ*". *Encyclopaedia of Islam* I, 667–677. 2nd ed. Leiden: E.J. Brill.

DMITRY FROLOV (University of Moscow)

Middle Arabic

1. DEFINITION

The very term 'Middle Arabic' is ambiguous because of the history of its use, the multiple meanings of the term 'middle' (historically middle, sociolinguistically intermediate, linguistically mixed – not to mention the middling quality of texts, in the opinion of some people), and the variety of views on the history of the Arabic language. Different definitions or characterizations have thus been proposed. As is the case for other languages, 'Middle' has been used to refer to a historical (post-Classical and modern) stage during which Standard Arabic itself had undergone deviations from the Ancient Arabic norm and innovations. The problem gets even more complicated, since it is often considered that Middle Arabic is structurally connected to Neo-Arabic dialects as regards linguistic typology and that, along with them, it breaks off from Ancient Arabic.

However, a thorough examination of the most ancient Middle Arabic texts that have reached us leads at least to a reconsideration of this point of view, since they suggest that Arabic → diglossia already existed in ancient times in a not substantially different way from the way it appears in later periods.

Consequently, there is no convincing reason why Middle Arabic should not start with the earliest documented period. In the same way, although it is customary to speak of Middle Arabic for the medieval period only (Blau 2002:4), there is no convincing reason for confining it to this period. As a matter of fact, Middle Arabic texts abound in later periods, at least until the *Nahḍa* (which seems to have put an end to its extensive written use, for reasons still to be elucidated), and sometimes until the beginning of the 20th century. For these periods, one also speaks of 'post-Classical Arabic'. This term is more explicitly chronological, but not unambiguous. Many post-Classical linguistic features are to be found in Middle Arabic, starting from the Classical period and even before (Blau 2002:15). One could thus consider using it only to qualify the language of the post-Classical period, when it is "almost [...] devoid of deviations in the field of orthography and morphology" (Blau 2001:4, n. 12), which actually means that it shares with Middle Arabic many syntactic, lexical, and stylistic features.

Whatever the precise chronological delimitation, one can agree to mean by Middle Arabic the language of numerous Arabic texts, distinguished by its linguistically (and therefore stylistically) mixed nature, as it combines standard and colloquial features with others of a third type, neither standard nor colloquial (for a noticeably different view, see e.g. Fischer 1982, 1991). To be more precise, Middle Arabic encompasses all the attested written layers of the language which can be defined as entirely belonging neither to Classical Arabic nor to colloquial Arabic, and as an intermediate, multiform variety, product of the interference of the two polar varieties on the continuum they bound, a variety that, for this very reason, has its own distinctive characteristics. Since the mixing is achieved to variable extents, one actually has to deal with a whole set of mixed varieties. One might try to draw up a typology of these varieties; Blau, for instance, has proposed the tentative categories of 'Literary Standard Middle Arabic' and 'Substandard Middle Arabic'. However, it should be remembered that, stylistic homogeneity not being the rule, different subvarieties often obtain in one and the same text.

If Middle Arabic refers to a variety (to varieties) of written Arabic, this is because there is obviously no direct testimony of any oral mixed Arabic variety available before modern times. But the nature of the linguistic mixing in Middle Arabic displays many similarities to what can currently be found in some oral mixed varieties (sometimes set down in writing, in the case of political speeches, debates in parliament, or in law courts). In this regard, nothing prevents us in theory, as far as the particular nature of oral and written language is taken into account, from regarding Middle Arabic written until precontemporary times as belonging to a large ensemble that could be labeled 'Mixed Arabic'.

2. The texts

Middle Arabic is the language of a very large number of texts: inscriptions, archival documents, public or private correspondences (the most ancient ones on papyri; see e.g. Diem 1997), travel accounts, historical chronicles, scientific (e.g. medical, physical) treatises, and juridical, religious, or mystical philosophical writings. Belles lettres belong here, too, not only 'popular' literature (*The Thousand and One Nights*) but also many texts of 'learned' literature (e.g. works by ʿIzz ad-Dīn at-Tanūxī, d. 384/994). One might reasonably think that the language of Classical Arabic texts (including celebrated ones like the *Kitāb al-ʾaġānī* of ʾAbū l-Faraj al-ʾIṣbahānī, d. 356/967; cf. Corriente 1975), often standardized by editors, is to some extent to be assigned to Middle Arabic, more than is usually supposed. In the later Middle Ages, some literary works were written in Middle Arabic, with an introduction and conclusion in literary Classical Arabic. The case of poetry is particularly complex (see Sec. 4).

This statement leads us to reconsider the accepted idea that Middle Arabic was the language of authors having insufficient command of Classical Arabic although their target was this linguistic level. One cannot but observe that many writers have left us works written both in faultless or even sophisticated Classical Arabic and works written in Middle Arabic. For those writers at least, one has to abandon the idea of their inadequacies in Classical Arabic. Moreover, it may be supposed that their choice to write some of their works in Middle Arabic was not arbitrary and was probably dictated, among other considerations, by the kind of audience they were writing for. All this points to the existence, beside the standard usage of Classical Arabic, of another standardized register available for those who wanted to use it. It may further be assumed that some authors, probably more trained in Middle Arabic than in literary Classical Arabic, a language with which they did not feel comfortable, actually gave their preference to Middle Arabic as their favorite or even only means of expression. From this point of view, it should be noticed that Middle Arabic facilitated for both writers and readers (and for the illiterate to whom texts were read aloud) access to written culture.

3. The norm of Middle Arabic

To talk about a standardized register means that the register conforms to a norm. Admittedly, this norm is neither institutionally codified nor explicitly recognized. More flexible than the classical norm, since it derived from different sources and hence admitted more variants, it is nevertheless present, agreed upon, and followed by all. One actually finds in all texts written in Middle Arabic, beyond their differences, a large amount of common features or common general processes, particularly regarding the peculiar (neither standard nor colloquial) features, the regular or systematic occurrence of which proves the existence of a norm to which anybody writing in Middle Arabic has to conform. Therefore, it is impossible to describe Middle Arabic as 'wild' use of the language, in which variation could be haphazard or even strongly idiolectal. On the contrary, one finds well-established usages, as well as stylistic hierarchies between variants. Notwithstanding a certain amount of freedom in its proportions, the mixing that obtains is all but random. This linguistic tool has been shaped and perfected through centuries in dealing with → diglossia, certainly perpetuating old traditions (including orthographical ones) that had been rejected by the classical norm, and certainly laying down new ones: Middle Arabic, being a living instrument of communication, grew richer in new features as it developed.

These considerations urge us to use the notion of pseudocorrection (hypocorrection and → hypercorrection; cf. Blau 1970) with great caution when analyzing the linguistic phenomena encountered in Middle Arabic. One cannot deny that some Middle Arabic features were in the beginning the result of incomplete or overzealous attempts to produce standard forms or constructions. Yet, their eventual systematic occurrence set them free from this original status and established them as genuine features, sometimes stylistically prestigious, of this second standard language.

In spite of the general conformity to the norm, one nevertheless notes differences according to space, time, and, to a lesser extent, communal affiliation.

i. Space. Dialectal diversity is ancient in the Arabic-speaking area, and it is only natural that it should exert an influence on Middle Arabic. Such an influence can be easily acknowledged in lexicon and, in a less direct and more subtle manner, in syntax. It is also obvious in phonetics and morphology and, more generally, in the frequent use of some linguistic devices, like demonstratives, conjunctions, adverbs, and auxiliary verbs, for which the Middle Arabic norm admits and sometimes prescribes the use of colloquial forms. True dialecticisms automatically end in differences between texts, according to the dialectal background, but the important fact is that dialecticisms are parallel. One must add that at least some dialecticisms also conform to a norm, in the sense that they are borrowed from prestigious dialects. Incidentally, this means that they cannot always be interpreted as deriving directly from the scribe's dialect, but rather from the dominant dialect in the surrounding sociopolitical context. Moreover, circulation of dialecticisms all over the Arabic-speaking area could occur, under the influence of dialects that were culturally prestigious or boasted a prestigious literary tradition (e.g. the Andalusian tradition in North African sung poetry), or even sometimes under the influence of particularly important individuals (like Saadia Gaon or Maimonides for Judaeo-Arabic). For all these reasons it is not always possible, especially in the case of only slightly colloquializing texts, to determine exactly the geographical origin of the texts and their authors (when no external element betrays it). It is also difficult to determine whether local Middle Arabic koines existed, like the Melkite lingua franca from the second half of the 1st millennium that Blau (1994) believes is identifiable.

ii. Time. Middle Arabic has undergone modifications throughout its long history, whether under the influence of the evolution of the dialects, or the evolution of the classical language. But these changes are still insufficiently studied and not well known. One should not underestimate the fact that because of its long tradition, Middle Arabic has accumulated, like any other literary language, including Classical Arabic, a large corpus. The history of Middle Arabic is incorporated in this corpus and is thereby at the disposal of the authors. Dating texts exclusively on the basis of linguistic criteria is therefore most of the time difficult, especially as supposedly recent phenomena often turn out to be rather ancient.

iii. Communities. It is often stated that Middle Arabic was more connected with Jews and Christians than with Muslims because, allegedly, these communities were less subjected ideologically to the requirement of excellence in language teaching and oral and written practice. This last point is questionable, as is the whole statement. Among other reasons, it might be the result of the development of Middle Arabic studies, which at first were devoted mainly to texts originating from these communities (such as translations or commentaries of religious texts from Greek and Syriac or Hebrew and Aramaic). It might also be due to the fact that many texts by Muslim authors are studied in artificially normalized editions. It remains true, however, that in some respects → Christian (Middle) Arabic and (Middle) → Judaeo-Arabic may be considered special varieties of Middle Arabic, at least if one considers the (quantitatively predominant) texts intended for the members of the communities concerned, which in practice were often accessible to them only, all the more so when they were written in non-Arabic scripts, which was the rule for Judaeo-Arabic (Hebrew characters) and Samaritan Arabic (Palaeo-Hebraic characters) and to a lesser degree for 'Christian Middle Arabic', which only in certain periods and areas and for certain communities was written in various scripts, among them an adapted version of the Syriac alphabet (*karšūnī*). These texts constitute largely autonomous literary traditions. Hebrew or Aramaic words and expressions are often found in Judaeo-Arabic, as well as calques of these two languages. Similarly, one finds in 'Christian Arabic' texts, at least in early ones, a clear influence of Greek or Syriac. Another element is that the influence of the various Neo-Aramaic dialects may have been stronger in the Arabic dialects spoken among Jewish or Christian communities, although their influence in the dialects spoken among the Muslims should not be underestimated (Blau 2000; → Neo-Aramaic loanwords).

Why is this Middle Arabic norm to be acknowledged through space and time and beyond communal affiliations? Or to put it differently, why did a unified Middle Arabic have such a long life? One can put forward two main reasons. First, in a manner comparable with literary Classical Arabic, Middle Arabic was an instrument of written communication in many fields of social and cultural life between the various societies concerned, and it must therefore have been standardized to a large extent. Second, as has been referred to above, one may think that Middle Arabic has inherited, through a long tradition, dating back maybe to pre-Islamic times, a transdialectal and transcommunal variety of Arabic, already standardized in actual fact, which continued living its parallel life after literary Arabic had been institutionally standardized in the 2nd and 3rd century of the Hijra, keeping alive ancient features eliminated from the 'learned' tradition.

4. The function and importance of Middle Arabic

As has been said, Middle Arabic embodies a whole set of mixed varieties. One may wonder whether a correlation exists between the various Middle Arabic subvarieties (more or less classicizing or colloquializing) and, for instance, the types or genres of texts, or maybe the types of audience aimed at or the types of written communication they set up. This is not unlikely but remains to be demonstrated. One might also wonder whether literary texts can be linguistically set apart from texts free from artistic ambition. This problem remains to be studied, but the answer could very well be affirmative, although literary texts, too, exhibit a large diversity of registers. Furthermore, one has to distinguish between literary prose and poetry. Popular literature comes first to mind. Among many others, the tales of *The Thousand and One Nights* are a good example. The language of the manuscripts edited by M. Mahdi (1984) is indisputably Middle Arabic. Compared with the language of nonliterary Middle Arabic texts, it makes moderate use of certain dialecticisms and of features proper to Middle Arabic. On the other hand, it makes extensive use, as can be expected, of stylistic and formulaic devices borrowed from literary Classical Arabic (often with modifications). The case of poetry is more complex, since it uses language even more for the purpose of artistic expression, by pressing it, playing with words and constructions, bringing it to submit to new formal constraints (with specific modifications or distortions). For the ancient period, the so-called → *rajaz* poetry (Ullmann 1966) is a good example of the use of a sophisticated variety of Middle Arabic. The same can be said, some centuries later, about the Andalusian *zajal*. In Maghrebi poetry, along with more popular genres, *malḥūn* poetry exhibits another type, equally sophisticated, archaistic, but far more colloquializing, to the extent that one has to be an expert to understand it properly. In this regard, one should probably abstain from characterizing its language as Middle Arabic: it is impossible to speak in this case of 'pan-Middle Arabic', but rather of a local (Moroccan and partly Algerian) literary colloquializing variety.

The importance of Middle Arabic should obviously not be underestimated. Quantitatively, it is the language of an important part of the written production in Arabic. Culturally, the alleged semiliteracy of writers must be reconsidered in many cases, even if some of them belonged to this category, as did a good part of their audience. As regards this audience, broadly speaking, it must be stressed that for centuries Middle Arabic was the main, if not the only, means of access to a part of learned culture, especially for the illiterate. In that sense, it can be labeled a cultural mediator. From a linguistic point of view, Middle Arabic is invaluable for the linguist and the historian of the Arabic language. It preserves very ancient features, it provides evidence of the dialects throughout their history, and at the same time it is the melting pot where new linguistic phenomena emerge, which eventually may spread to the literary language as well as to the colloquials. Incidentally, familiarity with Middle Arabic often helps to avoid misinterpretations of words and constructions in texts in 'Classical' Arabic.

To sum up, Middle Arabic has taken upon itself the glorious and obscure task of pragmatically dealing with Arabic diglossia, by filling the space of the linguistic continuum between both polar varieties, thus preventing it from becoming a 'no language's land' with all the consequences of such a situation. All things considered, this is perfectly natural, but in the case of Arabic, the function and importance of these intermediate varieties has often been neglected, both by Arabs and Arabists, mainly because of the outstanding status of Classical Arabic, and for reasons of purism and elitism. Yet, Middle Arabic is of fundamental importance for a correct understanding of the nature and history of the Arabic language.

5. Typical features

Below is given a short list, all but exhaustive, of linguistic features that can be considered typical of Middle Arabic. They have been selected for their universality (they can be found in various periods and areas) and for the representativeness of the phenomena or processes they exemplify. The list includes dialecticisms and features found only in Middle Arabic. It should be kept in mind that the occurrence of these features is never systematic, Middle Arabic texts being characterized by variation. However, the more the texts are colloquializing, the more colloquial features are likely to occur (in an order which is more or less fixed).

i. *Orthography and phonetics*

Hamza is written in various ways, some of them unknown in classical orthography; final *hamza* is generally absent. Interdentals are often replaced by the corresponding stops (*t*, *d*, *ḍ*), but *ḍ* and *ḏ̣* sometimes exchange, because of the merger of the two phonemes in dialects. Long vowels can be shortened in agreement with colloquial morphological patterns. Largely dependent on the classical norm, orthography also uses genuine notations (mostly generalizations of existing classical orthographic peculiarities) like *ʾalif* before an initial cluster of two consonants to indicate that it has to be pronounced *CC*-, or *ʾalif* after any final *u* or *o*. Consonantal assimilations, secondary 'emphatic' consonants (*ṣ* < *s* etc.) are often noted, as well as secondary loss of 'emphasis'. Enclitic morphemes are generally rightly attached to the word. In many texts, a syntactic distinction is made between the two allomorphs of the final feminine suffix, by using the letter *hāʾ* for *-a*(*h*) in absolute state and *tāʾ marbūṭa* for *-at* in construct state. → *ʾImāla* of long *ā* can be noted by *yāʾ*. *Scriptio plena* of short vowels occurs mainly in borrowings from Standard Arabic or foreign languages. Texts are generally unvocalized, only unpredictable short vowels being skillfully noted, as well as some others which would pass unnoticed or misread if not specified. But some texts are more or less fully vocalized, like the famous work of the Andalusian al-Ḥajarī (written in 1637) or some manuscripts of popular literature. Their vocalization sometimes appears rather strange, mainly because their technique has not yet been accurately studied or interpreted; indeed, they should provide us with important information about actual oralization of Middle Arabic.

ii. *Morphology and syntax*

Nouns: Gender is often colloquial; some plural patterns, colloquial or not, are largely used, like 'mixed' plural *CCūCāt*; the *n* of the sound plural ending *-īn* is often retained in construct state. A *t* can be added to plurals ending in **-āʾ* in construct state. Classical nominal declensional inflexion is normally absent, and occurs only in classicisms or pseudoclassicisms. Hence, the predominant 'sound plural' masculine ending is *-īn* (and for the dual *-ayn*), *-ūn*^{a} (and *-ān*^{i}) being left free for stylistically marked use, e.g. in numbers or collective proper names. The ending *-an* often appears as the mark of what could be labeled the 'emphatic state of the noun', whatever its syntactic function may be. The use of a final element *-an*, possibly historically related to → *tanwīn*, is largely attested; it mostly functions as a connective element, in a manner which recalls the connective *-Vn*, attested in present-day Bedouin dialects. In Judaeo-Arabic, it is often written as a separate word. As regards definiteness, mention can be made of the not-infrequent construction in which a noun qualified by an attributive adjective or relative clause has no determiner (*yawm at-tālī* 'the next day'); *kull* + noun is definite or indefinite; colloquial 'genitive' particles appear in most colloquializing texts but are otherwise generally avoided.

Pronouns: Independent personal pronouns are occasionally colloquial (especially in the 1st person plural). Interrogative pronoun 'what' is often colloquial (*ʾayš*, etc.). Very frequent is (*ʾa*)*ḥada*(*n*) 'somebody, someone'. Demonstrative pronouns and adjectives are often colloquial; besides their Classical Arabic use, *ḏālika* and *tilka* are stylistically elevated demonstratives, irrespective of gender. The reflexive pronoun can be colloquial (*ḥāl-*, *rūḥ-*, etc.). The relative adjective rather commonly appears as invariable *allaḏī*, which could very well be a conventional notation of colloquial *illi, ilḏi*, etc. The resumptive pronoun can be absent in clauses introduced by this relative.

Verbs: Verbal morphology is deeply influenced by the dialects. The 2nd person (common) plural of the perfect can be *-tū*; and verbs III*w* or IIIʾ generally shift to verbs III*y* (with active participles *CāCyīn* in the plural or 3rd person plural perfect *CVCyū*). In geminate verbs, the 1st and 2nd persons of the perfect are conjugated like verbs III*y;* and in the imperfect of I*w* verbs, *w* can be retained. In Andalusian and Maghrebi texts the colloquial 1st persons of the imperfect occur (singular *n-*, plural *n-ū*). In the 2nd and 3rd persons common plural of the imperfect, *-ū* is more frequent than *-ūna*. Mood inflection being absent, *-ūna* often appears as a variant occurring for prosodic reasons or as a stylistically or syntactically marked form, for instance after final conjunctions (*li-yamtalikūnahā* 'in order that they become the owner of it').

The use of derived verbal forms is often close to the colloquial pattern. In particular, Form VII is the prevailing means of expressing the personal and impersonal passive (although apophonic standard passives are not excluded); on Form X are formed → 'middle' verbs. 'Pseudoform IV' (= with the meaning of Form I) is frequent, in many cases apparently for phonetic reasons (presence of a 'weak' radical consonant), e.g. *ʾaḥkā* 'he spoke'. Colloquial or genuine → auxiliary verbs (e.g. inchoative or durative) are used frequently, e.g. *qāma*, *baqiya*, *istaqāma*, *ištaġala*, etc. Preverbal particles seem to be more frequent in the Mashreq (*b-*/*bi-*) than in the Maghreb. 'Future' particles occur, both colloquial (*ḥa-*, etc.) and typically Middle Arabic *muzmiʿ* and *ʿatīd*. Colloquial verbal nouns are not infrequent, e.g. for derived Form II *ti*CC*ā*C (Maghreb) and, if C_3 = *y*, *tə*CC*āy*(*a*) (Mashreq). The active participle (written *CāyiC* for Form I of II *w*/*y* verbs) is readily used with its colloquial constructions and aspectual values, among which the resultative perfect is particularly common for processive verbs. The 'neo-participle' *CaCCān* is of constant use in the Mashreq for many 'middle' Form I verbs.

Adverbs: Some adverbs are frequently used, e.g. *bi-l-kulliyya* 'entirely, completely'; some are colloquial, e.g. *barra* 'outside'; some are peculiar to Middle Arabic, e.g. *xāṣṣatan* 'exclusively'.

Prepositions: A frequent procedure is to generalize a partial correspondence between a dialectal and a standard preposition into total interchangeability (thus sometimes bringing about a pair of stylistically contrasting variants). The most typical example is *li/ʾilā* (hence the frequent *qāla ʾilayhi* 'he told him'); cf. also *fī/bi-*, *ʿalā/ʿan*. Other prepositions are: *wiyyā-* + bound personal pronoun 'with, in the company of'; *ka-maṯal/ka-miṯl* 'as'; *naḏ̣īr* 'like'; *min jihat* 'from, by, on behalf of'; *min ʾajl* 'because of', 'about'. Compound prepositions (with *li-*, pre- or postponed *min*) are frequent. Determinate direct object can be marked by *li-* (and sometimes *bi-/fī*). Various prepositions or prepositional phrases can introduce the agent of a passive verb.

Quantifiers: *jamīʿ* is extensively used as a substitute for *kull* 'all (the)' even with singular (determined) nouns; *sāʾir* also means 'the totality of'; *baʿḍ* is often used with nonclassical constructions; *xilāf* means 'other than' (= standard *ġayr*).

Negation: Verbal negation is predominantly with *mā* with perfect or imperfect (as in the dialects). *Lam* appears as stylistically marked and is often used with verbs in the perfect, or even as a nominal negation; this is a good example of a common procedure in Middle Arabic which consists in borrowing a linguistic tool from Standard Arabic and using it in a genuine manner without conforming to the syntactic constraints (and sometimes to the semantic limitations) attached to it. Classical *laysa* has become an invariable particle *lays* (*lēs, līs*). The second element *š(i)* of the bimorphematic colloquial negative particles is rather infrequent, except in strongly colloquializing texts or some types of poetry.

Subordinate clauses: Syndetic as well as asyndetic constructions abound. The latter are particularly frequent after modal verbs or → pseudoverbs ('to want to', 'to be able to', 'to have to'). The difference between *ʾan*, *ʾanna*, and *ʾinna* is blurred, *ʾan* (written <*ʾn*>) is the predominant subordination tool (Blau 1966–1967: § 402 and 2002: § 130). Many verbs (not only 'declarative' ones) govern *bi-ʾan*. Mood inflection is absent (except in classicizing passages or texts), as indicated for example by unapocopated imperfect forms of verbs II*w/y* or imperfect ending in *-ūn*^{a} after *lam*. On the whole, the system is close to that of the colloquials; distinctions parallel to classical ones can appear in texts in which the 'indicative' preverbal particles are used. Subjunctive particles are often colloquial. Typically Middle Arabic are *kawn* (and compound *li-kawn* and *min kawn*) 'because' (*kamā* is used with the same meaning, but can also mean 'as soon as') or *lammā ʾan* 'when'. It seems that to a certain extent, the shaping of these particles, often from a nominal to which *mā* is added, is free.

Concord/agreement: The system is mixed and complex. As in the colloquials, verbs and adjectives associated with nouns referring to nonhumans stand generally in the plural, as do those associated with nouns in the dual (referring to humans or nonhumans). Likewise, as in the colloquials, when the verb precedes its plural subject, it generally stands in the plural. As in the colloquials and in Ancient Arabic, some nouns referring to human groups have an agreement in the 'feminine singular'. But the concord patterns can vary in the same passage and even within the same sentence. This is due to the play between the colloquials and the standard rules, but also, it seems, to the fact that once an explicit agreement has been made, the sentence can go on with a less marked one, provided that the basic agreement is marked again whenever necessary.

Word order is not submitted to the constraints of classical use; as a matter of fact, it is close to the colloquials, for instance, regarding the position of the subject of the verb.

Lexicon, as can be expected, is extremely diversified and never refrains from making use of both the classical and the colloquial lexical store. The most usual words, nouns, and verbs referring to the most fundamental objects and operations, as well as basic adverbs or conjunctions, are most often colloquial. Examples of verbs: *(ʾi)jā* ‘to come’, *jāb* ‘to bring’, *šāf* ‘to see’ (also note Mashreq *bidd-* = Maghreb *ḥabb* or *bġā* ‘to want’; Mashreq *ṭalaʿ* ‘to go out/away’, Maghreb *nāḍ* ‘to stand up’); adverbs: *ḥattā* ‘also’ (Maghreb), *hallaq* or *hassaʿ* ‘now’ (Mashreq); *šwayya* ‘a little’. But for this basic vocabulary, Middle Arabic has also shaped or extensively developed the use of special words; they are often calques or transpositions of colloquial words, which can thus be avoided: *naḏ̣ar* [trans.] ‘to see’.

Stylistics: Clichés and formulas are frequent in narratives as well as in narrative passages of many texts (e.g. historical); passages in rhymed prose are not rare, especially in popular literature.

6. TEXTS IN MIDDLE ARABIC

For references about text editions see Blau (1981b); Blau (2002) gives large annotated extracts from early texts. Famous texts have been edited by Müller (1884), Zetterstéen (1919), Mahdi (1984), Wehr (1956), Bergsträsser (1914), Cachia (1960), Grand’Henry (1996, 2005), Hitti (1930), Shehadeh (1989), and Bengtsson (1995).

Among the numerous historical works in Middle Arabic from later periods, one could mention al-Ḥajarī, *Nāṣir ad-dīn* (al-Andalus, 17th century); *Riḥlat Ibn ʿĀbid* (970–1048 A.H.); Damurdašī, *Durra* (Egypt, 17th/18th century); Muḥammad aṣ-Ṣaġīr, *Mašraʿ* (Tunis, 18th century); Tāsāftī, *Riḥla* (Morocco, 18th century); *Ḥawliyyāt yamaniyya* (Yemen, 19th century); ʾIbrāhīm al-ʿAwra, *Tārīx* (Lebanon, early 19th century).

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

ʿAwra, *Tārīx* = ʾIbrāhīm al-ʿAwra, *Tārīx wilāyat Sulaymān Bāšā, 1804–1819*. Re-ed. A.B. Qīqānū. Beirut: Dār Laḥd Xāṭir, 1989.

Damurdašī, *Durra* = ʾAḥmad ad-Damurdašī, *ad-Durra al-muṣāna fī ʾaxbār al-Kināna*. Ed. D. Cresselius and ʿA. Bakr. Cairo: Dār az-Zahrāʾ li-n-Našr, 1992.

Ḥajarī, *Nāṣir ad-dīn* = ʾAḥmad ibn Qāsim al-Ḥajarī al-ʾAndalusī, *Nāṣir ad-dīn ʿalā l-qawm al-kāfirīn*. Ed. and transl. Pieter Sjoerd van Koningsveld, Qāsim as-Sāmarrāʾī, and Gerard A. Wiegers. Madrid, 1997.

Ḥawliyyāt Yamāniyya: al-Yaman fī l-qarn at-tāsiʿa ʿašara al-mīlādī. Ed. ʿAbdallāh Muḥammad al-Ḥibšī. Sanaa: Dār al-Ḥikma al-Yamāniyya, 1991.

Muḥammad aṣ-Ṣaġīr, *Mašraʿ* = Muḥammad aṣ-Ṣaġīr ibn Yūsuf, *al-Mašraʿ al-mulkī fī salṭanat ʾAwlād ʿAlīTurkī*, I. Ed. ʾAḥmad aṭ-Ṭuwaylī. Tunis, 1998.

Riḥlat Ibn ʿĀbid = *Riḥlat Ibn ʿĀbid al-Fāsī min al-Maġrib ʾilā Ḥaḍramawt*. Ed. ʾIbrāhīm as-Sāmarrāʾī and ʿAbdallāh Muḥammad al-Ḥibšī. Beirut: Dār al-Ġarb al-ʾIslāmī, 1993.

Tāsāftī, *Riḥla* = ʿAbdallāh ibn ʾIbrāhīm at-Tāsāftī, *Riḥlat al-Wāfid*. Ed. ʿAlī Ṣidqī ʾAzāykū. Kenitra: Ibn Tofail University, 1992.

SECONDARY SOURCES

Bengtsson, Per Å. (ed.). 1995. *Two Arabic versions of the Book of Ruth: Text edition and language studies*. Lund: Lund University Press.

Bergsträsser, Gotthelf (ed.). 1914. *Pseudogaleni In Hippocratis de septimanis commentarium ab Hunaino Q. F. arabice versum*. Leipzig and Berlin: Teubner.

Blau, Joshua. 1961. *Diqduq ha-ʿaravit ha-yĕhudit šel yeme ha-benayim* [A grammar of medieval Judaeo-Arabic]. Jerusalem: Magnes Press. (2nd ed., 1980.)

——. 1966–1967. *A grammar of Christian Arabic based mainly on South Palestinian texts from the first millennium*. 3 vols. (= *Corpus Scriptorum Christianorum Orientalium* 267, 276, 279; *Subsidia* volumes 27–29.) Louvain: Peeters.

——. 1970. *On pseudo-corrections in some Semitic languages*. Jerusalem: Israel Academy of Sciences and Humanities.

——. 1974. “Classical Arabic, Middle Arabic, Middle Arabic Literary Standard, Neo-Arabic, Judaeo-Arabic and related terms”. *Joshua Finkel Festschrift*, ed. S.B. Hoenig and L.D. Stitskin, 37–40. New York: Yeshiva University. (Repr., Blau 1988:255–259.)

——. 1981a. *The emergence and linguistic background of Judaeo-Arabic: A study of the origins of Middle Arabic*. 2nd enlarged ed. Jerusalem: Ben-Zvi Institute for the Study of Jewish Communities in the East. (3rd ed., 1999.)

——. 1981b. "The state of research in the field of the linguistic study of Middle Arabic". *Arabica* 28.187–203. (Repr., Blau 1988:118–134.)

——. 1982. "Das frühe Neuarabisch in mittelarabischen Texten". *Grundriss der arabischen Philologie.* I. *Sprachwissenschaft*, ed. Wolfdietrich Fischer, 96–109. Wiesbaden: L. Reichert.

——. 1983. "Vestiges of *Tanwīn -un* and the case ending *-u* as attested in Yemenite Judeo-Arabic texts from the seventeenth century". *Bulletin of the School of Oriental and African Studies* 46.529–531. (Repr., Blau 1988:285–287.)

——. 1984. "The contribution of Middle Arabic to the vocabulary of Modern Standard Arabic". *Logos Islamikos: Studia Islamica in Honorem Georgii Michaelis Wickens*, ed. R.M. Savory and Dionysius A. Agiùs, 9–20. Toronto: Pontifical Institute of Mediaeval Studies.

——. 1986. "On two works written in Middle Arabic Literary Standard". *Studies in Islamic history and civilization in honour of Prof. David Ayalon*, 447–471. Jerusalem: Magnes Press and Leiden: E.J. Brill.

——. 1988. *Studies in Middle Arabic and its Judaeo-Arabic variety.* Jerusalem: Magnes Press and Hebrew University Press.

——. 1994. "A Melkite Arabic *lingua franca* from the second half of the first millenium". *Bulletin of the School of Oriental and African Studies* 57.14–16.

——. 1999. "The status and linguistic structure of Middle Arabic". *Jerusalem Studies in Arabic and Islam* 23.221–227.

——. 2000. "Are Judaeo-Arabic and Christian Arabic misnomers indeed?". *Jerusalem Studies in Arabic and Islam* 24.49–57.

——. 2001. "The linguistic character of Saadia Gaon's translation of the Pentateuch". *Oriens* 36.1–9.

——. 2002. *A handbook of Early Middle Arabic.* Jerusalem: Hebrew University.

——, and Simon Hopkins. 1984. "On early Judaeo-Arabic orthography". *Zeitschrift für Arabische Linguistik* 12.9–27. (Repr., Blau 1988:381–400.)

Cachia, Pierre (ed.). 1960. *Eutychius of Alexandria: The Book of demonstration* (Kitāb al-burhān), I. (= *Corpus Scriptorum Christianorum Orientalium*, 1920, *Scriptores Arabici*, tomus XX.) Louvain: Peeters.

Corriente, Federico. 1975. "Marginalia on Arabic diglossia and evidence thereof in the *Kitâb al-aghânî*". *Journal of Semitic Studies* 20.38–61.

Diem, Werner. 1997. *Arabische Briefe des 7.–13. Jahrhunderts aus den Staatlichen Museen zu Berlin.* Textband. Tafelband. Wiesbaden: O. Harrassowitz.

Fischer, Wolfdietrich. 1982. "Frühe Zeugnisse des Neuarabischen". *Grundriss der arabischen Philologie.* I. *Sprachwissenschaft*, ed. Wolfdietrich Fischer, 83–93. Wiesbaden: L. Reichert.

——. 1991. "What is Middle Arabic?". *Semitic studies in honor of Wolf Leslau*, ed. Alan S. Kaye, I, 430–436. Wiesbaden: O. Harrassowitz.

Fück, Johann. 1955. *ʿArabīya: Recherches sur l'histoire de la langue et du style arabe.* Paris: Didier.

——. 1960. "ʿArabiyya. II/3. Le moyen arabe". *Encyclopédie de l'Islam*, I, 587–590. 2nd ed. Leiden: E.J. Brill.

Grand'Henry, Jacques. 1981. "Le moyen arabe occidental: Problèmes de caractérisation et de périodisation". *Proceedings of the ninth Congress of the Union Européenne des Arabisants et Islamisants (Amsterdam, 1–7 septembre 1978)*, ed. Rudolph Peters, 89–98. Leiden: E.J. Brill.

——. 1984. "Traits linguistiques de la version arabe du discours 24 de Grégoire de Nazianze". *Studi in onore di Francesco Gabrieli nel suo ottantesimo compleano*, ed. Renato Traini, I, 389–410. Rome.

——. 1996. *Sancti Gregorii Nazianzeni opera: Versio arabica antiqua*, I, *Oratio 21* (arab 20). (= *Corpus Scriptorum Christianorum Orientalium, series graeca*, 34, *Corpus Nazianzenum*, 4.) Turnhout: Brepols.

——. 2004. "Le moyen arabe de la version arabe du discours 40 de Grégoire de Nazianze, premiers éléments d'analyse". *Ultra Mare: Mélanges offerts à Aubert Martin*, ed. Frédéric Bauden, 1–9. Louvain: Peeters.

——. 2005. *Sancti Gregorii Nazianzeni opera: Versio arabica antiqua*, III, *Oratio* 40 (arab 4). (= *Corpus Scriptorum Christianorum Orientalium, series graeca*, 57, *Corpus Nazianzenum*, 19.) Turnhout: Brepols.

Grotzfeld, Heinz. 1992. "Schriftsprache, Mittelarabisch und Dialekt in 1001 Nacht". *Jerusalem Studies in Arabic and Islam* 15.171–185.

Hary, Benjamin H. 1989. "Middle Arabic: Proposals for new terminology". *Al-ʿArabiyya* 22.19–36.

——. 1992. *Multiglossia in Judeo-Arabic with an edition, translation and grammatical study of the Cairene Purim Scroll.* Leiden: E.J. Brill.

Hitti, Philip K. (ed.) 1930. *Usamah's Memoirs, entitled Kitab al-Iʿtibar by Usamah Ibn-Munqidh: Arabic text edited from the unique manuscript in the Escurial Library, Spain.* Princeton, N.J.: Princeton University Press. (Repr., Beirut: ad-Dār al-Muttaḥida li-n-Našr, 1981.)

Hopkins, Simon. 1984. *Studies in the grammar of Early Arabic, based upon papyri datable to before A.H. 300/A.D. 912.* Oxford: Oxford University Press.

Lebedev, Viktor V. 1977. *Pozdnij srednearabskij jazyk (XIII–XVIII v.v.)* [Late Middle Arabic (13th–18th century)]. Moscow: Nauka.

Lentin, Jérôme. 1997. *Recherches sur l'histoire de la langue arabe au Proche-Orient à l'époque moderne.* Thèse de Doctorat d'état, Université de Paris 3.

——. 2004. "La langue des manuscrits de Galand et la typologie du Moyen Arabe". *Les Mille et une nuits en partage*, ed. Aboubakr Chraïbi, 434–455. Paris: Sindbad-Actes Sud.

Mahdi, Muhsin. 1984. *The Thousand and One Nights (Alf layla wa-layla) from the earliest known sources: Arabic text edited with introduction and notes.* 3 vols. Leiden: E.J. Brill.

Müller, August. 1884. "Über Text und Sprachgebrauch von Ibn Abī Uṣeibiʿa's Geschichte der Ärzte". *Sitzungsberichte der philosophischen, philologischen und historischen Classe der kaiserlichen bayerischen Akademie der Wissenschaften* 5.853–977.

Schen, Ibrahim. 1972, 1973. “Usama ibn Munqidh’s Memoirs: Some further light on Muslim Middle Arabic, I, II”. *Journal of Semitic Studies* 17.218–236, 18.64–97.

Shehadeh, Haseeb (ed.). 1989. *The Arabic translation of the Samaritan Pentateuch edited from the manuscripts with an introductory volume.* I. *Genesis - Exodus*. Jerusalem: Israel Academy of Sciences and Humanities.

Spitaler, Anton. 1998. *Philologica: Beiträge zur Arabistik und Semitistik*, ed. Hartmut Bobzin. Wiesbaden: O. Harrassowitz.

Ullmann, Manfred. 1966. *Untersuchungen zur Raǧaz-poesie: Ein Beitrag zur arabischen Sprach- und Literaturwissenschaft*. Wiesbaden: O. Harrassowitz.

Versteegh, Kees. 1997. *The Arabic language.* Edinburgh: Edinburgh University Press. [Chap. 8, “Middle Arabic”.]

Wehr, Hans (ed.). 1956. *Das Buch der wunderbaren Erzählungen und seltsamen Geschichten* [*Kitāb al-ḥikāyāt al-ʿajība wa-l-ʾaxbār al-ġarība*]. Wiesbaden: F. Steiner. (Repr., Beirut: Dār al-Kitāb al-ʿArabī, n.d.; repr., Cologne: Al-Kamel Verlag, 1977.)

Zettersteen, Karl Vilhelm (ed). 1919. *Beiträge zur Geschichte der Mamlukensultane in den Jahren 690–741 der Hiǧra nach arabischen Handschriften.* Leiden: E.J. Brill. [Linguistic introduction, 1–33.]

Jérôme Lentin
(Institut National des Langues et Civilisations Orientales)

Middle Verbs

The transitivity system in Arabic includes two classes of verbs: *lāzim* lit. ‘stationary’, corresponding to the intransitive pattern in English, and *mutaʿaddin* lit. ‘crossing over [to an object]’, corresponding to the transitive pattern (→ *taʿaddin*). Intran-sitive verbs are self-sufficient, not requiring for the completion of their meaning more than the subject or → *fāʿil* lit. ‘the doer’, e.g. *ḏahaba zaydun* ‘Zayd left’. It is agreed that transitive verbs need for the completion of their meaning a *mafʿūl bihi* lit. ‘affected complement’, a direct object form. For a more comprehensive treatment of this, see Sībawayhi (*Kitāb* I), al-ʾAstarābāḏī (*Šarḥ* V), Ibn Yaʿīš (*Šarḥ* VI), Ibn ʿAqīl (*Šarḥ* I), al-ʾAnṭākī (n.d., I), an-Nādirī (1995), etc. Unlike English, which is a three-place language, Arabic is a four-place language, requiring two-, three- or four-place verbs. Illustrations of these are as follows: *ḍaraba zaydun ʿamran* ‘Zayd hit ʿAmr’, *ʾaʿṭā zaydun ʿamran nuqūdan* ‘Zayd gave ʿAmr money’, *ʾaʿlamtu ʿamran zaydan fāḍilan* ‘I informed ʿAmr that Zayd is honest’. In Quirk a.o.’s (1972) terminology, the first example is monotransitive, the second ditransitive, and the third complex transitive, with the difference that, unlike Arabic, English realizes the complex transitive with a place adverbial rather than a nominal, as in *I keep my money in a safe* or *I put my money in a bank*.

The assumption behind middle verbs (‘middles’, for short) is that they fill the middle ground between transitive and intransitive verbs in Arabic, without constituting a separate class of their own. In English, middles are more or less covered by what is known in the literature as the ‘unaccusative hypothesis’, which consists in regarding the subject of certain intransitive constructions as the underlying object (Matthews 1997:388), e.g. in *she died/she arrived*. The situation of these verbs in Arabic is more complex. Middles in Arabic do not totally overlap with unaccusatives in English. While unaccusatives occur only in intransitive constructions in English, middles in Arabic may occur as intransitives and transitives, and can be captured under what is known as *mušāraka* lit. ‘coparticipation’; as *al-mabnī li-l-majhūl* lit. ‘that which is built for the unknown’, or the apophonic → passive; and as *muṭāwaʿa* lit. ‘compliance’. Coparticipation consists in transitivizing the verb, while the apophonic passive and compliance detransitivize it. Thus, the former operates a → valence promotion to the verb, and the latter a valence reduction.

1. Coparticipation

Coparticipation is expressed in Arabic through the pattern *fāʿala*, as in *māšaytu ṣadīqī* ‘I had a walk with my friend’. The overt morphosyntax of the verb inscribes it as a transitive verb, where *ṣadīqī* ‘my friend’ is in the accusative form as a direct object to the process of walking. However, semantically the direct object is actually a subject, coparticipating in the process of walking with the noun subject. According to the framework of the present entry, a process of valence promotion is at play here, transitivizing what was intransitive. To understand the sentence, it must be paraphrased as follows: *ʾana wa-ṣadīqī mašaynā maʿan* ‘my friend and I walked together’. The likely motivation for this process of transitivization and its morphological correlate is to simplify the expression of accompaniment, which syntactically alleviates the structure of the sentence by shifting the morphology to the pattern of coparticipation, *fāʿala*. The other likely motive for attributing agency exclusively to the subject agent in overt morphology is probably linked to the subject being the initiator of the process of walking, which explains why the subject is thematized and the coparticipant is relegated to an accusative status in the rheme structure. Evidence that verbs of coparticipation are not transitive in meaning can be found in passivization. Since in Arabic only transitive verbs can be passivized – *māšaytu ṣadīqī* 'I had a walk with my friend' yields a passive, as in *mūšiya ṣadīqī* 'my friend was walked', although the form is pragmatically unnatural.

2. PASSIVIZATION

Whatever their grammatical persuasion, early Arab grammarians agreed that *al-mabnī lil-majhūl*, the apophonic passive in Modern Standard Arabic, deals with "a verb built for the object whose agent has not been named" (ʾAstarabāḏī, *Šarḥ* IV, 128). Arabic morphology is based on binyanim, which are templates that semantically generate Arabic verbs. The verb is assumed to consist of a discontinuous root tier (consonantism), penetrated by a prosodic tier or template, showing a skeletal sequence of consonants and vowels, and a melody tier (vocalism) marking the phonological melody a particular verb pattern takes (McCarthy 1981:399). The stem vocalism of the passive is *u-i*, which is infixed into the verb stem *k-t-b*, yielding *kutiba* 'was written'. Since the passive in Arabic only occurs with transitive verbs, it brings about a reduction of → valence, which gives the impression of intransitivity.

The subject of the apophonic passive in Arabic is obviously not the agent of the process but rather one who is affected by the process. The motives for adopting the apophonic passive are pragmatic, such as the obvious nature of the agent, ignorance of who the agent is, fear for oneself in mentioning the agent's name, or fear that something might happen to him or her by mentioning the agent's name (an-Nādirī 1995:503). Thus, transitive verbs of all patterns in Arabic have their valence reduced when they are shifted to the apophonic passive.

3. COMPLIANCE

Like the passive, the compliance form acts on all of the three types of transitive verbs, reducing the valence of four-place verbs to three, of three-place verbs to two, and of two-place verbs to one. One of the important motives for this reduction of valence as exemplified in cases of compliance is the fact that, as in the passive construction, the agent is too well known to mention, e.g. *ʾiḏā s-samāʾu nšaqqat* 'when the sky is rent asunder' (Q. 84/1), therefore, it is left implicit and its slot is left to the affected patient. It is implicit in this example that God is the agent of rending the sky asunder. This form is sometimes mistakenly taken to be a reflexive pattern, formed on the *infaʿala* template in Classical Arabic and Modern Standard Arabic, as in *inkasara l-kaʾs* 'the glass broke', where the sky and the glass act on themselves without causation. There is implicit causation underlying all forms of compliance.

In defense of the *jaʿliyya* 'causation' behind verbs of compliance, most of the verbs dealt with are change-of-state verbs, in which the grammatical subject has no physical responsibility for the change. For instance, in *inkasara l-kaʾs*, it is not the case that the glass as an inanimate object can break of itself, for instance by deliberation or intention. Therefore, it could only be that the agent of change-of-state from the absence of brokenness to brokenness is backgrounded under expression, with the caused object coming to occupy subject position. This is not, however, the opinion of Fassi Fehri (1997:97), who argues that, unlike English, there is no possible correlation in Arabic between compliance and causation. Another piece of evidence for the reality of causation in compliance can be studied through the particle *fa-* '(and) so…', which has a resultative value in Arabic. In examples such as *jamaʿtu l-ʾibila fa-jtamaʿat* 'I gathered the camels, so they gathered', the initiative for gathering is not incumbent on the subject of the compliant verb but rather on the subject of the first verb – the human agent, which counts as evidence that the two clauses stand in a cause-effect relation, with the subject of the first as the causal agent.

In a footnote, an-Nādiri (1995:353) explains compliance as follows:

> The complier is not necessarily the intransitive as many believe. Compliance is being influenced and accepting the influence of the verb, be it intransitive as in 'I opened the door and it opened', or transitive as in 'I taught him grammar and he learnt'. The complier in reality is the direct object which becomes a subject as in 'I pushed Khalil away, so he pulled away'. The complier is Khalil, but they called the verb that he performed as the one that does the compliance metaphorically. [*Laysa maʿnā l-muṭāwiʿ al-lāzim kamā yaḏ̣unnu baʿḍuhum wa-ʾinnamā l-muṭāwaʿa hiya t-taʾaṯṯur wa-qubūlu ʾaṯar al-fiʿl sawāʾan ʾa-kāna t-taʾaṯṯur lāziman naḥwa 'fataḥtu l-bāba fa-nfataḥa', ʾam mutaʿaddiyan ka-'ʿallamtuhu n-naḥwa fa-taʿallama' wa-l-muṭāwiʿ fī l-ḥaqīqa huwa l-mafʿūl bihi llaḏī ṣāra fāʿilan naḥwa 'bāʿadtu xalīlan fa-tabāʿada' al-muṭāwiʿ huwa xalīl lākinnahum sammaw fiʿlahu l-musnad ʾilayhi muṭāwiʿan majāzan*]

It is the direct object that complies with the logical agent, thus becoming the subject without becoming thereby the actual agent of the action. There is confirmation in the literature that "compliance is when you want something from something, and you get it" (Ibn ʿUṣfūr, *Mumtiʿ* 183). Thus, the volition of an implicit agent behind compliance is clear.

The verb patterns that express compliance in Arabic are the following (Ḥilwānī n.d.: 155–174):

i. *infaʿala*, as in *infataḥa* 'to open'. This is the pattern for compliance par excellence, occurring with most dynamic verbs. Importantly, Ibn as-Sarrāj (*ʾUṣūl* 138) divides verbs with this pattern into two types according to active participation of the subject, which is the agent of causation. He writes: "You would not say about 'I wrinkled' 'I did it', as in 'I broke it, so it broke'" (*wa-nkamašat lā taqūlu fīhi faʿaltuhu miṯla kasartuhu fa-nkasara*). Thus, in *inkamašat* 'it wrinkled', the speaker does not participate in the process of wrinkling, while in *inkasara* 'it broke', the speaker plays an active part in the process of breaking.
ii. *iftaʿala*, as in *ijtamaʿa* 'to gather'. This pattern also expresses compliance but with both dynamic and nondynamic verbs such as in *jamaʿtu l-ʾibila fa-jtamaʿat* 'I gathered the camels, so they gathered', and *ġamamtuhu fa-ġtamma* 'I saddened him, so he was filled with grief'.
iii. *tafaʿʿala*, as in *takassara* 'to break'. This pattern expresses compliance, as in *kassartu l-ʾaqlāma fa-takassarat* 'I broke the pencils, so they broke'. The difference between forms of compliance with *tafaʿʿala* and *infaʿala* is that the former derives compliance from *faʿʿala* (a four-consonant verb; note the doubling of the consonant), while the latter does so through the regular trilateral verb.
iv. *tafāʿala*, as in *tabāʿada* 'to keep away'. This pattern is not specialized exclusively in compliance; it is also shared with reciprocity as seen above. This pattern uses verbs deriving from *fāʿala*, which is a four-consonant verb, as in *bāʿadtuhu fa-tabāʿada* 'I kept him at a distance, so he kept away'.
v. *tafaʿlala*, as in *daḥraja* 'to roll'. This pattern is compliant with the pattern *faʿlala*, as in *daḥrajtu l-ʿaraba fa-tadaḥrajat* 'I rolled the car, so it rolled'.
vi. *ifʿanlala*, as in *iḥranjama* 'to gather'. This pattern is also compliant with the pattern of *faʿlala*, as in *ḥarjamtu l-ʾibila fa-ḥranjamat* 'I gathered the camels, so they gathered'. Verbs of this type are rarely used in contemporary Modern Standard Arabic.

It is interesting to note that not much remains of these patterns of compliance in the dialects of Arabic. In Tunisian Arabic, for instance, infixation is simplified to prefixation, with the compliance form being expressed predominantly by the prefix *t-*, as in *l-kas t-kissar* 'the glass broke'. Brahim (1996:43) reports, however, that the prefix *in-* is in complementary distribution with *t-* as a regular prefix in some southern subvarieties in Tunisia, where both the prefixed (*tkassar*) and infixed (*inkasar*) forms can still be heard (Zarzis, south of Tunisia).

The role of verbs of compliance in valence reduction is clear. The consequence of this reduction shows the verb overt morphology as if it were reflexivity, while its covert morphology is that of causation. The causation hypothesis can be used to account for why we can say *kasartuhu fa-nkasara* 'I broke, so it broke', but not **qaraʾtuhu fa-nqaraʾa* 'I read it, so it read'. There is a strategy of mystification at work, allowed by the overt morphology of verbs, which reduces the valence of the verb, thus hiding the agent of causation by raising the direct object to subject position.

Bibliographical references

Primary sources

ʾAstarābāḏī, *Šarḥ* = Raḍī d-Dīn Muḥammad ibn Ḥasan al-ʾAstarābāḏī, *Šarḥ al-Kāfiya*. 4 vols. 2nd ed. Benghazi: Gar Younes University, 1996.

Ibn ʿAqīl, *Šarḥ* = Bahāʾ ad-Dīn ʿAbdallāh ibn ʿAbd ar-Raḥmān Ibn ʿAqīl, *Šarḥ al-ʾAlfiyya*. 2 vols. Beirut: al-Maktaba al-Ḥadīṯa and Dār Ibn Zaydūn, n.d.

Ibn as-Sarrāj, *ʾUṣūl* = ʾAbū Bakr ibn as-Sarī Ibn as-Sarrāj, *al-ʾUṣūl fī n-naḥw*. 3 vols. Ed. ʿAbd al-Ḥusayn al-Fatlī. 3 vols. Beirut: Muʾassasat ar-Risāla, 1985.

Ibn ʿUṣfūr, *Mumtiʿ* = ʾAbū l-Ḥasan ʿAlī ibn Muʾmin Ibn ʿUṣfūr, 1979. *al-Mumtiʿ fī t-taṣrīf*. 4th ed. Beirut: Dār al-ʾĀfāq al-Jadīda, 1979.

Ibn Yaʿīš, *Šarḥ* = Muwaffaq ad-Dīn Yaʿīš Ibn Yaʿīš, *Šarḥ al-Mufaṣṣal*. 10 vols. Beirut: ʿĀlam al-Kutub, n.d.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. 5 vols. Ed. ʿAbd as-Salām Muḥammad Hārūn. Repr., Beirut: Dār al-Jīl.

Secondary sources

ʾAnṭākī, Muḥammad al-. N.d. *al-Muḥīṭ fī ʾaṣwāt al-ʿarabiyya wa-naḥwihā wa-ṣarfihā*. Beirut: Dār aš-Šarq al-ʿArabī.

Brahim, Ahmed. 1996. *L'occultif: Hypothèse pour un traṣtement trans-linguistique du "passif" et des structures apparentées*. Tunis: Association Tunisienne de Linguistique.

Fassi Fehri, Abdulkader. 1997. *al-Muʿjam wa t-tawsīṭ: Naḏarāt jadīda fī qaḍāyā l-luġa al-ʿarabiyya*. Beirut and Casablanca: Markaz aṯ-Ṯaqāfa al-ʿArabī.

Ḥilwānī, Muḥammad Xayr al-. N.d. *al-Muġnī l-jadīd fī ʿilm aṣ-ṣarf*. Beirut: Dar aš-Ṣarq al-ʿArabī.

Matthews, Peter H. 1997. *The concise Oxford dictionary of linguistics*. Oxford and New York: Oxford University Press.

McCarthy, John J. 1981. "A prosodic theory of nonconcatenative morphology". *Linguistic Inquiry* 12:3.373–418.

Nādirī, Muḥammad ʾAsʿad an-. 1995. *Naḥw al-luġa al-ʿarabiyya: Kitābun fī qawāʿid an-naḥw wa-ṣ-ṣarf*. Beirut: Al-Maktaba al-ʿAṣriyya.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1972. *A grammar of contemporary English*. London: Longman.

Zouhair Maalej (University of Manouba-Tunis)

Mimation

The term 'mimation' is used for the *-m* ending of the case suffixes in some Semitic languages. In Proto-Semitic, the singular of the noun had case suffixes with final *-m* (Diem 1975:243), as reflected in the most ancient attested languages: Akkadian and Amorite (both *-um, -im, -am*), Old Canaanite in Egyptian transcription (*-m*), and Sabaic (*-m*) (Brockelmann 1908:472–474; Moscati a.o. 1964:96–99; Diem 1975:241–242; Layton 1990:157–159; Lipiński 1997:272–273; Streck 2000:259–260). This final *-m* originally had no (in)determinate value. It cannot be assigned any other function, either, so that it seems to have been part of the case suffixes itself. Consequently, these have to be reconstructed as *-um*, *-im*, *-am*, rather than *-*u*, *-*i*, *-*a* (Diem 1975:246–249). Sabaic shows a secondary development in assigning nouns and proper names with final *-m* an 'absolute' value (*ṣlmm* 'a statue', ʾBDM /ʿAbdum/), as opposed to nouns ending in *-n* with determinate value (*ṣlmn* 'the statue') and nouns in the construct state (Beeston 1984:30). In addition to the singular, mimation is also found in the feminine plural of Akkadian and Sabaic (*-ātum*, *-ātim*, and *-m*) as well as in the broken plural of Sabaic (*-m*). According to Diem (1975:243–246), the *-m* in the dual and plural suffixes of Ugaritic (*-āmi, -ēma, -ūma, -īma*; cf. Tropper 2000:289–290, 293–294) and Hebrew (*-ayim*, *-īm*) replaced an original *-n* by analogy to the mimation of the singular.

North Arabian usually has nunation instead of mimation in the singular. This has been explained either as a phonological (Brockelmann 1908:472; Gelb 1969:140; Stempel 1999:92) or as an analogical (Diem 1975:243) process. Nevertheless, some pre-Classical North Arabian dialects show vestiges of mimation in personal names, e.g. Liḥyānic *ḥmrm* 'donkey', Ṣafāitic *ḥrbm* 'warrior', Amīritic *ʿbydm* 'slave' (Gelb 1930:255; Harding 1971:4; Müller 1982:27; Layton 1990:195), sometimes apparently with indeterminate (or 'absolute'?) value in opposition to the prefixed article *h-* (Moscati a.o. 1964:99; cf. the situation in Sabaic). Gelb (1930:255) also quotes some certainly archaic personal names with mimation in Classical Arabic, e.g. *fusḥum*. Some of these names have the article *al-*, e.g. *al-xaṣārum*, which shows that mimation here is fossilized.

Bibliographical references

Beeston, A.F.L. 1984. *Sabaic grammar*. Manchester: Journal of Semitic Studies, University of Manchester.

Brockelmann, Carl. 1908. *Grundriss der vergleichenden Grammatik der semitischen Sprachen. I. Laut- und Formenlehre*. Berlin: Reuther und Reichard.

Diem, Werner. 1975. "Gedanken zur Frage der Mimation und Nunation in den semitischen Sprachen". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 125.239–258.

Gelb, Ignace Jay. 1930. *La mimazione e la nunazione nelle lingue semitiche*. Rome: Scuola Orientale nella R. Università di Roma.

——. 1969. *Sequential reconstruction of Proto-Akkadian*. Chicago: University of Chicago Press.

Harding, G. Lankester. 1971. *An index and concordance of pre-Islamic Arabian names and inscriptions*. Toronto: University of Toronto Press.

Layton, Scott C. 1990. *Archaic features of Canaanite personal names in the Hebrew Bible*. Cambridge, Mass.: Harvard University.

Lipiński, Edward. 1997. *Semitic languages: Outline of a comparative grammar*. Leuven: Peeters.

Moscati, Sabatino (with Anton Spitaler, Edward Ullendorff, and Wolfram von Soden). 1964. *An introduction to the comparative grammar of the Semitic languages: Phonology and morphology*. Wiesbaden: O. Harrassowitz.

Müller, Walter W. 1982. "Das Frühnordarabische". *Grundriss der arabischen Philologie*. I. *Sprachwissenschaft*, ed. Wolfdietrich Fischer, 17–29. Wiesbaden: L. Reichert.

Stempel, Reinhard. 1999. *Abriß einer historischen Grammatik der semitischen Sprachen*. Frankfurt a. Main: P. Lang.

Streck, Michael P. 2000. *Das amurritische Onomastikon der altbabylonischen Zeit*. I. *Die Amurriter, die onomastische Forschung, Orthographie und Phonologie, Nominalmorphologie*. Münster: Ugarit-Verlag.

Tropper, Josef. 2000. *Ugaritische Grammatik*. Münster: Ugarit-Verlag.

Michael P. Streck (Leipzig)

Minimalism

1. The minimalist framework

Minimalism is the name of the predominant approach in generative linguistics today. It was first introduced by Chomsky in his work *The minimalist program* (1995) and has undergone several developments and changes since. The central idea of minimalism is that a linguistic theory should contain as few nonderived assumptions as possible. Many notions that had been developed in earlier generative theory, in particular the → Government and → Binding theory (GB), have been abandoned in an attempt to derive them from more basic concepts.

In Chomsky's (1995) view, minimalism is an implementation of the more general Principles and Parameters model. According to this language model, the human language capacity consists of a set of universal principles combined with a set of parameters. The principles are thought to be innate, which means that every language adheres to them. The parameters can be thought of as switches that can be set in two positions. A typical example of a parameter is the so-called → *pro-drop* (pronoun-drop) parameter, which specifies whether a language can drop its subject pronoun (as, for example, in Spanish, Italian and Arabic), or not (as in English and in Germanic languages in general). Learning the syntax of one's native language, according to Chomsky, is a matter of acquiring the correct parameter settings for the language.

Chomsky describes syntax as a cognitive system that connects two other cognitive systems: the conceptual-intentional system and the articulatory-perceptual system. Because syntax is linked to these two systems, the syntactic model defines two interface levels, one for each of them: Phonological Form (PF), the interface to the articulatory-perceptual system, and Logical Form (LF), the interface to the conceptual-intentional system.

The grammar model is built as follows. A clause is derived by selecting a set of lexical items from the lexicon. This set is called the numeration. The syntactic (tree) structure of the clause is built by taking words one by one out of the numeration and putting ('merging') them in the structure. At the same time, other syntactic operations can take place. The end result of this derivation is Logical Form, the interface form for the conceptual-intentional system. If, during the derivation of Logical Form, a principle of grammar is violated, the derivation is said to *crash*. The clause under consideration is then considered ungrammatical. At some point during the derivation of Logical Form, *spell-out* takes place. Spell-out refers to the process of deriving PF, the interface form to the articulatory-perceptual system. PF basically contains the phonological and prosodic features of the clause.

Apart from Logical Form and Phonological Form, the fundamental concepts in minimalist syntax are 'economy', the principle of 'Full Interpretation', and 'features'. These concepts are discussed here in order.

Economy means that operations of the syntax component (the computational system) must cost as little as possible. Cost is defined in terms of the number of operations that it takes to derive a certain structure, and sometimes also in the relative 'heaviness' of different operations, i.e., some operations can be more costly than others.

In the earlier years of minimalism, there was much discussion of economy. The idea was that the language model must be able to compare the cost of different derivations and then select the least costly one. However, this idea (often called 'global relative economy') was abandoned because it would require too much from the computational system: instead of deriving just one clause, two or even more would need to be derived. Furthermore, it turned out to be difficult to establish which derivations would be candidates for comparison of their 'cost'. Global relative economy was therefore replaced by local absolute economy, which basically means that at each step in the derivation, when there are different options for continuing, the least costly option would have to be selected, no matter what the cost of the entire derivation would be.

The adoption of local absolute economy has resulted in the disappearance of the notion of economy from most domains of the theory. The idea of economy was based on comparison of different options, but local absolute economy means that in each situation only one option is possible, and therefore there is no need for comparison. (The notion of 'economy' is still at the basis of what is often called 'minimality': the condition that the distance between two agreeing elements be as short as possible.)

Full Interpretation (FI) is a principle that is fundamental to minimalism. Full Interpretation states that the syntactic representation cannot contain elements that have no semantic relevance. Earlier generative theories had introduced a number of elements into the syntactic representation that existed solely to facilitate syntactic computation. On the basis of the principle of Full Interpretation, Chomsky (1995) claims that these elements should not be postulated: Full Interpretation has led to the abandonment of several core notions of Government and Binding theory, most notably → X-bar theory, indices, and category labels. (Many authors, however, still make use of indices and category labels and even of X-bar levels, because they make an analysis easier to read. Strictly speaking, however, they should be taken as mnemonic devices, not as part of the theoretical apparatus.)

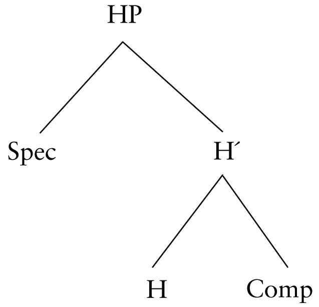
Features are properties of heads. Heads are the basic syntactic components, the elements with which the computational system builds structures. In principle, each word is a head, but elements such as affixes, determiners, and complementizers are also heads. Heads can be covert, i.e. void of phonological content. Technically speaking, a head can be described as a bundle of features. Typical features are gender, number, person, (in)definiteness, tense, aspect, and case, which are examples of so-called syntactic features. The grammar model also specifies semantic features, such as a quantifier feature, an interrogative feature, focus and topic features. Less important to syntax proper but nonetheless essential for language are phonological and prosodic features.

In earlier minimalist models, features were often considered binary. A head, for example, had a feature [Tense], which took the value [+Tense] or [-Tense]. In more recent developments, there is a move toward a valued feature system, in which features can have more than just plus or minus values. For example, a noun has a feature [Gender], which can have the value *masculine, feminine*, or *neuter*, and possibly others, depending on the language. The feature [Tense] now specifies tense (*past, present*, or *future*); its absence indicates an infinitival form (which in the old system was [-Tense]).

Features are at the heart of the syntax model of minimalism. All syntactic computation is done on the basis of features. The computational system, i.e. the component that builds the syntactic tree, has two operations: *Merge* and *Agree*. Merge is the operation that combines syntactic elements into larger structures. This is how trees are built. Agree is the operation that establishes a relation between two different elements in the syntactic structure through which feature values can be exchanged. This operation is responsible for agreement phenomena in language.

The approach sketched here can be considered strictly Chomskyan, because it is in essence the approach that Chomsky develops in his work. There are other approaches within generative grammar, however, that adopt minimalist principles in varying degrees and that at times may interpret certain principles differently, leading to different formalizations. For example, the operation Agree establishes

Figure 1. Basic structure of syntactic tree



an agreement relation between two elements. These two elements will obviously be in different locations in the structure, the higher one (which is always a head) being called the *probe*, the lower one (which is generally a phrase) the *goal*. In Chomsky's (1998) view, Agree can only be established from the probe to the goal, while other authors (e.g. Lasnik 1999) argue that it can be established in either direction.

The most influential alternative proposal within generative grammar has been developed by Kayne (1994). In a strict sense, his proposal, generally called *antisymmetry*, is not minimalist, because it is more of a continuation than a break with Government and Binding theory: some concepts that Chomsky (1995) abandons because they violate Full Interpretation, most notably X-bar theory, are retained by Kayne. In a broader interpretation, however, antisymmetry can be seen as a minimalist approach, because like minimalism it is an attempt to reduce concepts from Government and Binding to more basic notions.

The tree in Figure 1 may serve to illustrate differences between Chomsky's minimalism and Kayne's antisymmetry. It presents the basic structure of syntactic trees that generative grammar uses. H is the head (i.e. a bare verb, bare noun, preposition, etc.) which projects to H' (pronounced as *H-bar*) and HP. Here, the P stands for 'Phrase', while H can represent any head, such as N (noun), V (verb), A (adjective), P (preposition), and others. Thus are formed NP (Noun Phrase), VP (Verb Phrase), AP (Adjective Phrase), and PP (Prepositional Phrase) The head is combined with a *complement*, and the head-complement complex is further combined with a *specifier*. The first difference between Chomsky and Kayne is that Chomsky argues that if the complement is absent, there is no H', and similarly, if there is no specifier, HP is absent, (i.e., H or H' equals HP in that case). In Kayne's view, however, H' and HP are always present, even if specifier and/or complement are absent. Chomsky also argues that there can be more than one specifier, while Kayne maintains the stricter notion that there can be only one specifier. Furthermore, according to Kayne, the specifier always precedes the head, and the complement follows it, while Chomsky argues that there is no predefined order.

Other differences are more technical. Chomsky argues that the agreement relation is between a head (the probe) and an element further down in the tree (the goal). When an agreement relation is established, the goal can be moved up to the specifier position of the probe. In Chomsky's view, this is the only situation in which movement can occur: it must follow a successful Agree operation. Kayne, on the other hand, argues that Agree always takes place between a head and its specifier, and he argues that movement can always take place. In other words, for him, there is no relation between Agree and movement, as there is for Chomsky.

In short, *minimalism* is not a specific, well worked-out syntactic theory. Rather, it is a set of ideas on what a linguistic theory should look like. In principle, if one wished, it would be possible to distinguish several different syntactic theories based on minimalist ideas, but in practice it is very rare to find an author adhering strictly to one such theory. Most authors combine concepts from different sources. In other words, the field is still searching for a minimalist syntactic theory that suits both the data and the fundamental conceptual notions best.

One consequence of minimalist theory is that certain phenomena that were previously seen as syntactic are no longer considered as such, or at least it is questioned whether they should be. Generally, the derivation toward Logical Form as explained above is called *core syntax*. The operations that are available in core syntax are limited, which means that the type of phenomena that can be accounted for in core syntax is strictly defined. It turns out that there are systematic phenomena that have always been regarded as syntactic that cannot easily be described in core syntax within a minimalist framework. As a result, ideas have developed in which such phenomena are dealt with not in core syntax but in other domains, most notably in the derivation of Phonological Form.

2. MINIMALIST STUDIES IN ARABIC SYNTAX

In Arabic formal syntactic research, minimalism is the dominant approach (→ syntax; → interface studies). Although many different aspects are discussed (see, for example, the series *Perspectives on Arabic linguistics*), there are a few issues that receive more attention than others. One of the topics no doubt most often discussed is the VSO → word order pattern, the alternation with SVO order and the verbal agreement that appears in these cases. There are few authors who follow the idea of traditional Arabic grammar that SVO is basically a → topic/comment (*mubtadaʾ/xabar*) structure, although they would probably agree that it can sometimes be. The discussion therefore focuses on several points: what is the cause of the word order variation, i.e., what is the syntactic structure of both word orders, and how are the structures derived? How is the agreement relation established between the verb and the subject, both in SVO and in VSO orders? And why does SVO show full agreement, while VSO only shows agreement in gender?

There are several different proposals to solve these questions. A discussion of some of them can be found in Harbert and Bahloul (2002). One proposal says, for example, that there are actually two different types of → agreement relations in language (as opposed to the 'standard' assumption that there is only one type). SVO instantiates *spec/head agreement* (the standard type), while VSO instantiates *agreement under government*. The difference has to do with the structural relation that the agreeing element has to the element it agrees with. The alternative stipulates that there is only one type of agreement (*spec/head* agreement), but there are two independent agreement processes, which take place in different positions in the clause: one establishes agreement in gender, and the other establishes agreement in number. The idea is that the subject of a clause always first moves to the position where gender agreement occurs, which is the position directly following the verb. It is then possible but not obligatory to move the subject to a higher position preceding the verb, where number agreement takes place. If the verb moves to this position, SVO order results, together with agreement in both gender and number. If the verb does not move, VSO order results, and there is only agreement in gender. This type of proposal is somewhat problematic for minimalist theory, because a strict interpretation of minimalism does not allow optionality. Therefore, it would be necessary to establish why the subject moves to the highest (most frontal) position in some cases and not in others.

Another type of proposal is worked out by Ackema and Neeleman (2004). Their proposal is a good example of the development mentioned above that certain phenomena are no longer seen as part of core syntax. Ackema and Neeleman's proposal is formulated in another minimalist-based framework, called *flexible syntax* (Neeleman and Weerman 1999). The central idea of this framework is that certain types of operations take place not within syntactic domains but within prosodic domains. With this idea, Neeleman and Weerman account for word-order variation between various languages (VO vs. OV), and they link other phenomena to it (such as the possibility of *scrambling*: deviations from 'standard' word order for pragmatic reasons). Ackema and Neeleman build on this, arguing that there are more phenomena than just word order and scrambling that can be better accounted for in prosodic domains than within syntactic ones.

One example they discuss is so-called *split agreement*, the fact that certain languages show different subject-verb agreement paradigms in VS orders than they do in SV orders. The phenomenon occurs, for example, in some eastern Dutch dialects, and even in one form of present tense agreement in Standard Dutch. When the subject follows the verb, the subject and verb form a prosodic domain. The idea is that when this happens, a 'phonological' agreement process can take place, which has a different paradigm than 'syntactic' agreement. The latter occurs when the subject precedes the verb, in which case the two do not form a prosodic domain.

Ackema and Neeleman then develop the idea that something similar happens in Arabic. In an SVO order, subject and verb do not form a prosodic domain, and hence full agreement is required. In VSO order, however, subject and verb are in the same prosodic domain, which results in a weakening of the agreement. In other words, Ackema and Neeleman argue that there are two types of agreement, but unlike the proposal mentioned above, it is not the case that there are two types of syntactic agreement. Rather, the two types can be distinguished by the place where they occur: one type of agreement exists in core syntax, another type can occur in the derivation of Phonological Form.

The VSO nature of Arabic and the split agreement that it has are actually quite important to generative grammar and minimalism in particular. The most heavily studied languages in generative grammar are the Romance and Germanic languages, which are SVO with SOV alternations in some languages, and Japanese, which is SOV. Together with the Celtic languages, Arabic is probably the most widely studied VSO language in generative grammar, and as such it has much to contribute to the development of the general theory.

There are one or two other properties of Arabic that have a similar status, the → construct state being one of them. The first to discuss the construct state within a generative framework was Ritter (1991), who discusses the possessive structures of Modern Hebrew. As with the VSO/SVO alternation, there are several different proposals for the structure of the construct state. One proposal (e.g., Mohammad 1988), and perhaps the most common, is based on the observation that with deverbal nouns, the subject of the original verb becomes the genitive possessor. This prompts the suggestion that the possessor is located in the specifier position of the possessed noun, just like the subject of a verb is in the specifier location of the verb. Since this would lead to a surface order of possessor-possessed, which is incorrect for Arabic, it is suggested that the possessed noun moves to the position of the determiner. This movement would explain why the possessed noun in the construct state does not have a definite determiner at all: its canonical position (usually labeled D), is already occupied by the noun.

Another proposal is put forth by Benmamoun (2000), who argues that the lack of number agreement in VS structures is related to the absence of a determiner on the possessed noun in the construct state. The idea is that the verb and its subject, and likewise the possessed noun and its possessor, undergo some form of phonological 'merger' operation after syntax proper, that results in a single complex (phonological) word, in which the number feature and the definiteness feature can only be represented once.

A different analysis is developed by Kremers (2003). Kremers, whose analysis is made within a strictly minimalistic, Chomskyan approach, points out some similarities between the construct state and the so-called 'Saxon genitive' in English (the prenominal possessor construction such as *the man's car*) and argues that both constructions have basically the same structure. The obvious word order differences between the two constructions, which would normally follow from the structural analysis, are accounted for with an extrasyntactic linearization procedure.

This analysis shows that the concept of construct state has had some influence in generative syntax in general. Kremers is not the first to suggest that a construct state analysis could apply to possessive constructions in non-Semitic languages. Longobardi (1995) makes a similar suggestion for a certain type of possessive construction in Romance languages.

Another topic that may be considered to be of interest to generative grammar in general is → negation in Arabic. The fact that Standard Arabic has two types of negation, one that carries tense (*lā*, *lam*, and *lan*) and one that does not (*mā*) confirms the notion that both tense and negation are present on independent heads in the structure. Furthermore, Ouhalla (2003) argues that the negation circumfix that is common in several spoken varieties of Arabic confirms certain ideas on the logical representation of negation (i.e. its representation in Logical Form).

A final topic to be mentioned here is that of adjective ordering (and ordering on noun-phrase modifiers in general, i.e. demonstratives, numerals, quantifiers). Arabic has strictly postnominal adjectives, and as such contrasts with Germanic (which has prenominal adjectives) and with Romance (which usually has both prenominal and postnominal adjectives). Because of certain restrictions that the antisymmetric framework of Kayne (1994) mentioned above places on syntactic trees, a very specific type of analysis has to be developed to account for postnominal adjectives, which is attempted for example by Fassi Fehri (1999) and Shlonsky (2000). This analysis is challenged, however, by Kremers (2003), who develops an alternative account, which is compatible with Chomsky's minimalist approach rather than with Kayne's antisymmetry.

There are many other topics that are the object of investigation and discussion in Arabic syntax. Some good recent collections of papers can be found in Ouhalla and Shlonsky (2003) and in the series *Perspectives on Arabic linguistics: Papers from the annual symposia on Arabic linguistics*.

Because generative grammar aims at developing a universal syntax theory, it has a strong comparative nature. Structures that are studied in a Germanic or Romance language are often compared to equivalent structures in other Germanic or Romance languages, because the languages within these families are often very alike. As a result, there is some focus nowadays on *microvariation*, variation in details among languages and dialects that are by and large very similar. Arabic, with its large variety of spoken dialects, would lend itself quite well to this sort of research, but to date not much comparative work between different dialects within a generative framework has been done, one recent exception being Benmamoun (2000). Similarly, comparative studies between Standard Arabic and the spoken varieties are still not common.

A bit more comparative work is being done on Hebrew and Arabic, although mainly with Hebrew as the point of departure, as witnessed by several papers in Shlonsky (1997) and Ouhalla and Shlonsky (2002). Although the results of these studies are used, where appropriate, in studies focusing on Arabic (the construct state, for example, was first analyzed in Hebrew, and this analysis was adapted to Arabic), the results of the latter are usually not extended to Hebrew.

BIBLIOGRAPHICAL REFERENCES

Ackema, Peter and Ad Neeleman. 2004. *Beyond morphology: Interface conditions on word formation*. Oxford: Oxford University Press.

Benmamoun, Elabbas. 2000. *The feature structure of functional categories: A comparative study of Arabic dialects*. Oxford and New York: Oxford University Press.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, Mass.: MIT Press.

——. 1998. "Minimalist inquiries". *MIT occasional papers in linguistics*, 15. Cambridge: Massachusetts Institute of Technology.

——. 1999. "Derivation by phase". *MIT occasional papers in linguistics*, 18. Cambridge: Massachusetts Institute of Technology.

Fassi Fehri, Abdelkader. 1999. "Arabic modifying adjectives and DP structures". *Studia Linguistica* 53:2.105–154.

Harbert, Wayne and Maher Bahloul. 2002. "Postverbal subjects in Arabic and the theory of agreement". *Themes in Arabic and Hebrew syntax*, ed. Jamal Ouhalla and Ur Shlonsky, 45–70. Dordrecht, Boston, and London: Kluwer Academic Publishers.

Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, Mass.: MIT Press.

Kremers, Joost. 2003. *The Arabic noun phrase: A minimalist approach*. Ph.D. thesis, University of Nijmegen. (= *LOT publication series*, 79. Utrecht: Netherlands Graduate School of Linguistics [LOT].)

Lasnik, Howard. 1999. *Minimalist analysis*. Malden, Mass.: Blackwell.

Longobardi, Giuseppe. 1995. "A case of construct state in Romance". *Scritti linguistici e filologici in onore di Tristano Bolelli*, ed. Roberto Ajello and Saverio Sani, 293–329. Pisa: Pacini.

Mohammad, Mohammad A. 1988. "On the parallelism between IP and DP". *Proceedings of the West Coast Conference on Formal Linguistics*, 7, ed. Hagit Borer, 241–254. Chicago: University of Chicago Press.

Neeleman, Ad and Fred Weerman. 1999. *Flexible syntax*. Dordrecht: Kluwer Academic Publishers.

Ouhalla, Jamal. 2002. "Negative sentences in Arabic". *Themes in Arabic and Hebrew syntax*, ed. Jamal Ouhalla and Ur Shlonsky, 299–320. Dordrecht, Boston, and London: Kluwer Academic Publishers.

——, and Ur Shlonsky (eds.). 2003. *Themes in Arabic and Hebrew syntax*. Dordrecht, Boston, and London: Kluwer Academic Publishers.

Ritter, Elizabeth. 1991. "Two functional categories in the noun phrase: Evidence from Modern Hebrew". *Perspectives on phrase structure: Heads and licensing*, ed. Susan D. Rothstein, 37–62. San Diego: Academic Press.

Shlonsky, Ur. 1997. *Clause structure and word order in Hebrew and Arabic: An essay in comparative Semitic syntax*. New York and Oxford: Oxford University Press.

——. 2000. "The form of the Semitic noun phrase: An antisymmetric, non N-movement account". Ms., University of Geneva.

JOOST KREMERS (University of Nijmegen)

Modal Verbs

1. MODAL VERBS IN CLASSICAL ARABIC

The modal verbs in Arabic, as in other languages, are those verbs which through their meaning and function are linked to the category of → 'mood' (*modus* vs. *dictus*; cf. Ducrot and Schaeffer 1972), within a framework of 'modalization', borrowed from Aristotelian logic. The presence of a ‘modal’ verb in a sentence changes its content by offering information on the manner in which the speakers relate themselves to the respective content: whether they assume it to be true or false, known, necessary, impossible or possible, desired, good or bad, etc. The classification of the modalities adopted here was proposed by Anghelescu (2000:304) in a chapter about the modalities in Arabic:

i. Alethic (or ‘classic’) modalities: necessary, impossible, possible, etc.
ii. Temporal modalities: it is (was, will be) the case; always, sometimes, etc.
iii. Deontic modalities: obligatory, defended, permitted, etc.
iv. Evaluative modalities: good, bad, suitable, regrettable, etc.
v. Epistemic modalities: I know, I think that
vi. Boulomaic modalities (attitudes and states of mind): I hope, I wish, I am afraid

According to Anghelescu (1981:15–21, 2000:306–307), Classical Arabic grammarians had in mind an implicit theory of modalities when they spoke about the lexical-syntactical category of ‘modifying words’ (→ *nawāsix*). These words “have a formal ‘action’, which consists in the ‘modification’ of the case of the inchoative and predicative elements, and a ‘semantic’ one, related to the signification of each type [of modifying words] and, finally, of each word” (Anghelescu 1981:17).

The Arabic grammarians explain the inflectional endings of nouns and verbs by the action of an ‘operator’ (*ʿāmil*; → *ʿamal*), which is either expressed (*lafḏ̣ī*) or semantic (*maʿnawī*). Both the inchoative constituent (*mubtadaʾ*; → *ibtidāʾ*) and the predicative constituent (→ *xabar*) are affected by the ‘operator’, if it is present in the phrase. The ‘modifying words’, which express modalities, are a special category of such ‘operators’.

A large number of modal verbs and modal expressions with verbal value in Arabic traditional grammar are included in the category of ‘modifying words’. They are classified in the following (nonexhaustive) list, together with the modality they represent in Classical and Modern Literary Arabic:

i. Classes of ‘sisters’ (words with similar syntactic behavior), each having as ‘head’ a prototypical element:
   a. *kāna* ‘to be’ stands for temporal modalities (having various other values that will be indicated below); other verbs in the same class are: *ʾaṣbaḥa, ṣāra, ʾamsā, bāta, ʾaḍḥā, ġadā,* all in the sense of ‘to become’, which express temporal modalities and the transformation from one state or situation to another; *ḏ̣alla, dāma, baqiya, labiṯa, mā zāla, mā bariḥa, mā fatiʾa* ‘to continue’, *mā ʾāda* ‘to stop doing something’ express temporal modalities also related to the aspect of the action, namely its duration (→ *kāna wa-ʾaxawātuhā*).
   b. *ḏ̣anna* ‘to think, believe, suppose’, *ḥasiba, ʿadda, raʾā, wajada, laqiya, ʾalfā* in the sense of ‘to consider, find something in a specific way’; each of these has additional meanings, which are not relevant here; *zaʿama* ‘to pretend’, *ʿalima, darā* ‘to know, learn’, which stand for epistemic modalities related to the ‘mood’ of the assertion.
   c. *ʾinna*, an intensifying particle which, without being a verb, exercises the influence of a verb on the inchoative element; it may have the contextual meaning of ‘to be certain, convinced; to affirm, confirm’, or it may be translated by ‘verily, truly’, expressing an epistemic modality of certainty; to this class also belong particles that may have a verbal value because of their verbal origin, such as *ʿalla, laʿalla*, and *ʿasā* ‘it is possible; perhaps, maybe; who knows if…; it might/could be’, expressing the alethic modality of possibility; and *layta* ‘God willing!; if only…’, expressing the boulomaic modality of wishing or desiring (→ *ʾinna wa-ʾaxawātuhā*).
ii. The class of the ‘verbs of imminence’ (*ʾafʿāl al-muqāraba*), which are sometimes included in the expanded category of ‘modifying words’: *badaʾa, ʾaxaḏa, jaʿala, šaraʿa, ṭafiqa, ʾanšaʾa, qāma, rāḥa, habba* in the sense of ‘to begin, to start doing something’, expressing inchoativity, a temporal-aspectual modality. The same class includes also verbs like *ʾawšaka*, *kāda* ‘to be on the point of doing something, to be about to, to be close to’ and *mā labiṯa* (in the negative form) in the sense of ‘to keep doing something; it did not take long before he…’, which express temporal-aspectual modalities related to an ‘inchoative’ sense.

Other modal verbs and modal expressions

with verbal value which are not part of the category of ‘modifying words’ behave syntactically in different ways, which are not discussed here; some of them are classified according to semantic criteria.

iii. The class of ‘transformation verbs’ (*ʾafʿāl at-taḥwīl*): *ṣayyara* ‘to transform’, *jaʿala* ‘to make someone do something; to believe someone to be something; to take someone for something’, *ittaxaḏa* ‘to take someone or something as’, *radda* ‘to trace back to an origin; to transform’; *taraka* ‘to let someone do something’. They are used to express the ‘inchoative’ in the extended sense of transformation, change, modification of the situation, which may be regarded as varieties of the temporal-aspectual modalities. The Arabic grammarians classify them close to the ‘verbs of imminence’ of the type *badaʾa, ʾaxaḏa, jaʿala, šaraʿa, ṭafiqa, ʾanšaʾa, qāma, rāḥa*, etc., because of their general significance. Anghelescu (1981:115–116) notes that they do not express modalities because there are no nucleus phrases in the deep structure of the phrase marked by such a ‘transformation verb’.

iv. *wajaba* (almost always in the imperfect *yajibu), inbaġā* (frequently in the imperfect *yanbaġī*), *lā budda (min)*, *lazima*, *ʿalā* + affix or noun in the accusative, all of them expressing the alethic modality of necessity.

v. *ʾarāda, wadda* ‘to wish, desire’ (epistemic modality).

vi. *rajā* ‘to hope’ (boulomaic modality).

vii. *ʾamkana* (especially in the imperfect *yumkinu*) ‘to be possible’, *jāza* (in the imperfect *yajūzu*) ‘to be possible, admissible’ (alethic modality).

viii. *jadara* (in the imperfect *yajduru*) ‘to be suitable, worthy’, *ḥasuna*, ‘to be suitable, proper’, *niʿma* ‘what a wonderful . . . [he/it is]!’, expressing positive evaluation or approval, *biʾsa* ‘what an evil . . . [he/it is]!’, expressing negative evaluation or disapproval, *ṣaʿuba* (in the imperfect *yaṣʿubu*) ‘to be hard, difficult, unpleasant’, *ḥabbaḏā* ‘how nice/lovely [he/it] is . . . !’, *ḥabbaḏā law* ‘how nice it would be if . . . !’ (evaluative modalities).

ix. *lā šakka* ‘it/there is no doubt’ (epistemic modality related to the assertion, the expression of doubt).

x. *rubbamā* ‘many a time’ (alethic modality expressing possibility).

xi. *li* + affix +*ʾan* in the sense of ‘to be able to’ or expressing possibility, permission (deontic modality expressing the permission), etc.

The large morphological variety and the absence of any classificatory criteria of these verbs and particles are obvious. Syntactically, the verbs in class (i.a) have in common that they put the predicative element (*xabar*) in the accusative, while the inchoative element or subject (*mubtadaʾ*) is in the nominative (*kāna/ʾaṣbaḥa zaydun marīḍan* ‘Zayd is/became sick’). In the case of class (i.b) (*ḏ̣anna* and its ‘sisters’), the two elements of the nonmodal initial phrase *zaydun marīḍun* (*zaydun*: inchoative nominal element, subject; *marīḍun*: predicative element), are put in the accusative case in the corresponding modal phrase, by the ‘action’ of the modal verb *ḏ̣anna: ḏ̣anantu zaydan marīḍan* ‘I considered Zayd [to be] sick’. The modal particles in class (i.c) trigger the accusative for the inchoative nominal element and the nominative for the predicative element: *ʾinna zaydan marīḍun* ‘Zayd is [truly] sick’. Likewise, the particles *ʿalla*, *laʿalla*, and *layta* place the inchoative nominal element in the accusative, but in their case, the second element of the phrase is a verb in the imperfect: *layta š-šabāba yaʿūdu* ‘if only youth came back!’. The particle *ʿasā* is usually employed with a following verb in the subjunctive: *māḏā ʿasā ʾan ʾafʿala* ‘what should I do?’, but sometimes it is followed by a noun or affix in the accusative: *māḏā ʿasā-hu yaqūlu* ‘what could he possibly say?’. The verbs in class (ii) are followed by another verb in the imperfect: *ʾanšaʾa yaqūlu* ‘he began to speak’; *qāma/rāḥa yaquṣṣu* ‘he started to tell [a story]’.

Semantically, it may be remarked that each compact class (classes i, ii, iii) includes both modal verbs and expressions ‘specialized’ in expressing the same modalities or similar modal nuances (e.g. i.a *kāna* and its ‘sisters’). The semantic criteria must have been observed by the Arabic grammarians; in fact, when classifying these verbs, they actually classify modalities.

The other modal verbs and modal expressions (classes iii–xi), which are not part of the category of ‘modifying words’, vary in syntactic behavior, which is not discussed here in detail, but some of them are grouped on semantic criteria, as in the case of the verbs of ‘evaluation’ in class (viii).

Within this classification of the Arabic modal verbs and phrases, one subcategory, which is interesting from several points of view, is conspicuous, that of the ‘semi-auxiliaries’, to which verbs from the classes (i.a), (i.b), and (ii) belong. Messaoudi (1985:158) has drawn up a list of fifteen semi-auxiliary verbs from a large corpus (Naguib Maḥfūẓ’s novel *Zuqāq al-midaqq*), which she groups as follows (→ auxiliary):

i. Inchoative verbs: *kāda* ‘to be about to do something’, *jaʿala* ‘to start doing something; to let someone do something’, *ʾaxaḏa, badaʾa, ʾanšaʾa* ‘to start something’, *ʾawšaka* ‘to be on the point of doing something’, *(mā) infakka, (mā) fatiʾa* ‘not to cease being/doing something’;
ii. Motion verbs: *rāḥa, maḍā* ‘to go away, leave; to continue’, *ʿāda* ‘to come back; to continue’;
iii. Duration verbs: *labiṯa* ‘to remain, stay’, *(mā) zāla* ‘not to cease doing’, *ḍalla* ‘to spend [time]; to remain; to continue’, *bāta* ‘to spend the night; to continue’.

Messaoudi remarks that the verb *kāna* holds a special place among its ‘sisters’, which are semi-auxiliaries. The values that she assigns to this verb (1985:174, from which the examples are derived) are: autonomous verb (*lammā kānat al-laylatu ṯ-ṯāniyatu* ‘when the second night came/was’); copula (*kānat mašġūlatan* ‘she was busy’); auxiliary (*kāna yakrahu-hā* ‘he hated her’). She adds that, as an auxiliary, *kāna* renders itself void of meaning, becoming only a base of the temporal modality, while the semi-auxiliaries do not lose their lexical meaning completely; the lexical meaning is weakened but does not disappear completely (Messaoudi 1985:175). According to her, *kāna* as a perfect auxiliary enters in the structure of some ‘compound tenses’ (*kāna lā yakādu yafqahu* ‘he knew hardly anything’; *kānat lā tazālu muṣirratan* ‘she remained decided’).

According to Larcher (2003:143), *kāna* is not ‘a conjugation auxiliary’, i.e., it is not used to form new compound tenses. He translates a sentence like *kuntu qad qīla lī* not as *on m’avait déjà dit* ‘I had already been told’ , but as *je me trouvai [qu’]on m’ait dit* ‘I was in the situation that I had already been told’. At first sight, this sentence contains a compound tense, but Larcher asserts that the translation into languages with compound tenses has caused this misleading analysis. In the spirit of the Arab grammarians, he analyzes it as a nominal phrase to which the operator *kāna* is applied (Larcher 2003:143–144), and explains that in some contexts, *kāna* unquestionably has the value of introducing the modalities of possibility, necessity, and duration (2003:148, 153–154, 157). Therefore, it must be analyzed as a ‘modal exposant’ rather than an auxiliary. This point of view changes the view on this verb entirely, because it highlights its modal values. It also provides additional justification of the fact that the Arab grammarians introduce it without any hesitation in the list of the modal verbs as one of the modal operators.

The majority of the verbs and expressions listed here are still in use in Modern Literary Arabic, especially in literature, with the same values they have in Classical Arabic.

2. Modal verbs in modern dialects

Many recent studies in Arabic dialectology show that modal verbs from Classical Arabic, in the forms they have reached in various dialects, are affected by a process of → grammaticalization. This process has different stages for different items in each individual dialect, but almost always, it leads from a stage of full verbs to one of semi-auxiliaries, clitics, or even particles through decategorialization. In many cases, the full verb does not disappear but rather coexists with the decategorialized forms (for the Egyptian modals *rāḥ*, *ʾām/qām*, *ga*, *baʾa*, for example, see Woidich 1995; for the Egyptian modal *baʾa*: full verb > auxiliary > particle expressing various modalities, see Firanescu 2002).

The most frequently discussed modal verb is *kāna*, which in Classical Arabic expresses existence but many other values as well. In dialects, the complex nature of *kāna* is even more conspicuous because it has undergone a process of grammaticalization while preserving the value of 'full verb'. Numerous dialectological studies have dealt with this modal and with the various stages of grammaticalization it has undergone in the dialects. Grigore (1999:10–11) shows that in Mardin Arabic the verb *kān* is used with the following values: full verb (*kent fe-mardin//fe-mardin kent* 'I was in Mardin'); intermediate form between autonomous word and affixes (*kan naḥne we-l-akrad sadqan* 'we and the Kurds were friends'), functioning as a clitic without conjugation; inflectional affix *ka-*, as a result of grammaticalization (*lbent katerjef kama lwārāq* 'the girl was shaking like a leaf'). Grigore (1999:14–15) also indicates the modal value of conditional-optative for the particle *ka-*: *nuri kateyedrāb lay ʿalī be-l-xānjār, bas xāf* 'Nuri would have struck Ali with the sword, but he was afraid'. Grigore (2002:374–375) also mentions the existence of a particle *ku* in Mardin Arabic, whose origin may be the imperfect *yikūn* of the same *kān,* expressing the present continuous. Jastrow (1995:99) mentions the verbal modifier *kun* (< *kān*) for past habitual in Uzbekistan Arabic, which he compares with *ka*, *kan*, *kān* in some Anatolian dialects. According to Youssi and Zniber (2002:384), the 'modality' *ka-* has three values in Moroccan Arabic: declarative, durative, and repetitive. Sasse (1971:266) mentions that Blanc and Fischer recorded the existence of the 'prefix' *kən* in various Arabic varieties (Iraq, Uzbekistan), and reports that this prefix exists in the Mḥallamīye dialect (Turkey), where it has the same functions as in the extended area of Anatolian Arabic. The possibility modality of the prefix *ka-* is mentioned also by Ferrando (1998:198) for Andalusian Arabic (16th century).

Specialized studies deal with other modals in various dialects, for instance, Eksell (1995), who analyzes modal 'existential verbs' in the Syro-Palestinian dialects (*kān, baqa, ṣār, ʿād, ʿam, ḍall*) and 'verbs of motions' (*qām, rāḥ, rajaʿ, dār, ja, qaʿad*), showing that they "signify a structural renewal of the verbal system within the Arabic dialects studied". She considers that "the revitalization is still anchored in the ancient Semitic verbal system, exploring new possibilities within the old structural mould" (Eksell 1995:48). For the study of this development, a more detailed perusal of the Medieval Arabic historical documents is crucial, because they provide an important link between Classical Arabic and the newly emerging dialects, as Lentin (1995) suggests regarding an extended area (Syria, Lebanon, Palestine, Egypt). He lists a number of dialectal auxiliaries and 'mixed' forms (such as *rāyiḥ*, *ʿāwiz*, *bədd*, *rād*, *ʿāwed*, *baqa*, etc.), shared within a "trans-regional substandard koiné". By observing the stages of its diachronic development in dialects of the Bilād aš-Šām, Lentin is able to explain the values of the particle *ʿam (< ʿammāl)*, used for various temporal modalities and going back to Classical Arabic *ʿamila* 'to make' (factitive-causative, synonym of *jaʿala*).

Versteegh (1997:108) observes, as a general trend in all Arabic dialects, that "for non-modal aspects, the dialects have developed a new system of aspectual markers, originally *auxiliary verbs* or temporal adverbs, which became fossilised as part of the morphology of the verbal form". He mentions the following markers for continuous/habitual//future: *ʿam-, bi-//raḥ(a), laḥ(a)* (Syrian Arabic, Damascus); *bi-//ḥa-* (Egyptian Arabic, Cairo); *ka-//ġa-* (Moroccan Arabic, Rabat); *da-//raḥ* (Iraqi Arabic, Baghdad); *bi-//ʿa* (Yemeni Arabic, Sanaa) and states that "it seems to be the case that future markers often derive from verbs meaning 'to go' [...], whereas continuous markers derive from the verb *kāna*, or from participial forms meaning 'sitting', 'doing', standing' [...]" (Versteegh 1997:108). The incorporation of these markers ("originally *auxiliary verbs*") in the morphology is one aspect of the multifarious evolution of modals in dialects (→ auxiliary).

Another general trend is that some modal verbs and expressions acquire, both in peripheral and nonperipheral spoken Arabic, a surprisingly large number of various pragmatic values in expressing modalities (cf. Vanhove 1998 and Woidich 1995 on *xalli* > *xa* for permission, consecutive and cohortative particle, etc., in Maltese and Egyptian, respectively; Ingham 1994 and Firanescu 2000 on *bədd* for opposite modal meanings, from wish to obligation and necessity, etc., in Saudi and Syrian Arabic, respectively). Some recent studies have collected more and more examples of these phenomena in various dialects (cf. for instance Cuvalay-Haak 1997) and analyzed them within linguistic frameworks that are valid for other languages, as well.

Within the verbal system in dialects, the modals seem to be the most dynamic and 'creative' part; consequently, the study of modals in different dialects and their connection to those in Classical Arabic is important for the analysis of the contemporary verbal system of Arabic, the characteristics of the newly emerged dialects, and, not least, for understanding the semantic–pragmatic development through historical periods.

Bibliographical references

Anghelescu, Nadia. 1981. *Semantica modalităţilor în limba arabă* [The semantics of modalities in Arabic language]. Bucharest: Universitatea Bucureşti.
——. 2000. *Limba arabă în perspectivă tipologică* [The Arabic language in a typological perspective]. Bucharest: Univers Enciclopedic.
Caubet, Dominique and Martine Vanhove (eds). 1994. *Actes des Premières Journées Internationales de Dialectologie Arabe de Paris*. Paris: INALCO.
Cremona Joe, Clive Holes, and Geoffrey Khan (eds.). 1995. *Proceedings of the Third International Conference of the Association Internationale pour la Dialectologie Arabe, 10–14 September 1995*. Cambridge: University of Cambridge.
Cuvalay-Haak, Martine. 1997. *The verb in literary and colloquial Arabic*. Berlin and New York: Mouton de Gruyter.
Ducrot, Oswald and Jean-Marie Schaeffer. 1972. *Nouveau dictionnaire encyclopédique des sciences du langage*. Paris: Editions du Seuil. (Repr., 1995.)
Eksell, Kerstin. 1995. "Some punctual and durative auxiliaries in Syro-Palestinian dialects". Cremona a.o. (1995:41–49).
Ferrando, Ignacio. 1998. "Andalusī Arabic in Postislamic North of Spain: The language of Aragon's Mudejars and Moriscos". Mifsud (1998:195–200).
Firanescu, D. Rodica. 2000. "De *vouloir* à *devoir* en arabe syrien". *Analele Universităţii Bucureşti* 49.109–120.
——. 2002. "Valeurs discursives de *ba'a* en arabe du Caire". Youssi a.o. (2002:364–373).
Grigore, George. 1999. "*Ka-* a temporal prefix in Mardini Arabic derived from the verb *kana* (to be)". *Annals of the Faculty of Arts and Social Sciences-University of Balamand* 9.9–15.
——. 2002. "*Ku-* un préfixe temporel dans l'arabe mardinien". Youssi a.o. (2002:374–380).
Ingham, Bruce. 1994. "Modality in the Arabic dialect of Najd". Caubet and Vanhove (1994:185–200).
Jastrow, Otto. 1995. "Towards a reassessment of Uzbekistan Arabic". Cremona a.o. (1995:95–103).
Larcher, Pierre. 2003. *Le système verbal de l'arabe classique*. Aix-en-Provence: Université de Provence.
Lentin, Jérôme. 1995. "Existait-il un arabe levantin à l'époque ottomane? Traits dialectaux communs et koïné sous-standard transrégionale". Cremona a.o. (1995:133–139).
——. 1994. "*ʿAm yəktob, ʿam bi-yəktob*: Eléments pour une approche historique". Caubet and Vanhove (1994:297–316).
Messaoudi, Leïla. 1985. *Temps et aspect: Approche de la phrase simple en arabe écrit*. Paris: Librairie Orientaliste Paul Geuthner.
Mifsud, Manwel (ed.). 2000. *Proceedings of the third International Conference of the Association Internationale de Dialectologie Arabe, Malta, 29 March – 2 April 1998*. Malta: Salesian Press.
Sasse, Hans-Jürgen. 1971. *Linguistische Analyse des arabischen Dialekts der Mḥallamīye in der Provinz Mardin*. Berlin.
Vanhove, Martine. 1998. "Future, injunctive and purpose subordinating conjunctions: The case of Maltese *ħalli, ħa* and *biex*". Mifsud (2000:235–240).
Versteegh, Kees. 1997. *The Arabic language*. Edinburgh: Edinburgh University Press. (2nd ed., 2001.)
Wehr, Hans. 1980. *A dictionary of Modern Written Arabic (Arabic-English)*. Ed. J. Milton Cowan. Wiesbaden: O. Harrassowitz.
Woidich, Manfred. 1995. "Some cases of grammaticalization in Egyptian Arabic". Cremona a.o. (1995:259–268).
Youssi, Abderrahim and Mohamed Zniber. 2002. "Do auxiliaries really help in Moroccan Arabic?". Youssi a.o. (2002:381–393).
——, Benjelloun Fouzia, Mohammed Dahbi, and Iraqui-Sinaceur (eds.). 2002. *Aspects of the dialects of Arabic today: Proceedings of the 4th Conference of the International Arabic Dialectology Association (AIDA), Marrakesh, Apr. 1–4.2000*. Rabat: Amapatril.

Daniela Rodica Firanescu
(Dalhousie University)

Modern Standard Arabic

1. General

Arabic is a → Semitic language of the → Afro-Asiatic family of languages; it is the official language of the 22 nations of the Arab world, the area bounded by the Atlantic on the west, the Mediterranean on the north, the Sahara on the south, and the Arabian Peninsula and Iraq on the east; it is spoken by more than two hundred million people. Modern Standard Arabic is the High literary form of Arabic that goes back to the literary language of pre-Islamic Arabia (→ poetic koine; → Classical Arabic; see Ferguson 1956); it is learned in schools and is not the day-to-day language of any Arab population. It is used universally in formal writing and speaking, in professional meetings and conferences, in radio and TV news, and on other occasions where the aim is to communi-

Table 1. Modern Standard Arabic consonant phonemes

Type	labial	labio-dental	dental	alveo-palatal	alveolar	pharyn-gealized	velar	uvular	pharyn-geal	glottal
plosives	b		t, d			ṭ, ḍ	k	q		ʾ
affricates				j						
fricatives		f	ṯ, ḏ			ḏ̣	x, ġ		ḥ, ʿ	h
sibilants			s, z	š		ṣ				
liquids			l		r					
nasals	m		n							
glides	w			y						

Table 2. Modern Standard Arabic vowel phonemes

		front	central	back
High	close	ī		ū
	open	i		u
Mid			a	
Low			ā	

cate on specialized topics or with Arabs of different dialectal backgrounds (foreign-language soap operas and TV cartoons are dubbed in formal Modern Standard Arabic). The primary differences in Modern Standard Arabic as used across the Arab world are in the lexicon (→ lexical variation), because the structures are remarkably constant across space and time. As the language of Islam and the *Qurʾān*, it is used in speech and writing throughout the Muslim world (→ diglossia; see Badawi 1973; Versteegh 1997; Holes 2004).

2. PHONOLOGY

2.1 *Consonants and vowels*

Modern Standard Arabic has 28 consonants and 3 long-short pairs of vowels. They are presented in Tables 1 and 2.

Noteworthy among the consonants are the unaspirated voiceless uvular stop *q* (→ *qāf*); the phonemic glottal stop ʾ (→ *hamza*); the voiceless/voiced pair of faucalized pharyngeal fricatives *ḥ*, ʿ; and *ṭ, ḍ, ḏ̣, ṣ*, the pharyngealized counterparts of the dental *t, d, ḏ, s*. The term 'emphasis' from Semitic linguistics is often applied to these four pharyngealized consonants; they are articulated with tongue retraction, tongue tension, and lack of any aspiration (see also → velarization). The glottal fricative *h* is pronounced in all positions, including syllable-final; *j* is pronounced [g] in Cairo and the Nile Delta; [dʒ] in Upper Egypt (Egypt south of Cairo), Saudi Arabia, and Iraq; and [ʒ] in the Levant and North Africa.

Emphatic *ḷ*. On the basis of such minimal pairs as *waḷḷāhu* 'and God', where *ā* is low back [ɑː], and *wallāhu* 'he appointed him governor', where *ā* is low front [æː], *ḷ* is set up as a separate phoneme (Ferguson 1956); it does not occur after *i*, as in *li-llāhi* 'for God'. It is generally marked only in this word, since it is otherwise predictable from the semantic and phonetic environment.

The long vowels are twice as long in duration as their short counterparts in stressed open syllables; unstressed word-final long vowels are shortened to half-long. As there are only three phoneme types, there is much room for allophonic fluctuation; all vowels come in two main allophonic types, a front and a marked backed type. The front (unmarked) allophones are, broadly speaking, *ā* [æː], *a* [ɛ ~ æ], *ī* [iː], *i* [ɪ], *ū* [uː], *u* [ʊ]. The backed allophones, *ā* [ɑː], *a* [ɑ], *ī* [ᶦiː], *i* [ɪ], *ū* [uː], *u* [ʊ], occur contiguous to, and in some varieties of Modern Standard Arabic preceding in the same word, a pharyngealized consonant or following the uvular stop *q*; following pharyngeal *ḥ* or ʿ, *ā* has a backed low front allophone [aː]. Backing in general refers to centralizing the vowel and producing it with some tenseness. To illustrate with Lebanese Modern Standard Arabic, *ī* is preceded by a high central glide [ᶦiː], *i* is centralized [ɪ], *ū* and *u* are [+tense].

2.2 *Syllable structure*

There are two basic syllable structures, CV (consonant – vowel) and CVC, as in *fa.qaṭ*.

'only'. The vowel may be long, as in *kā.nā* 'they [masc. du.] were'; long vowels cannot occur before two nonidentical consonants: **ixtārta* > *ixtarta* 'you [masc. sg.] chose' but *hām.mun* 'important'. CVC syllables may occur word-medial, as in *ʾaḥ.san.tum* 'you [masc. pl.] have done well!; bravo!' (→ syllable structure). CVCC and C$\bar{V}$CC structures appear in pausal form: *ka.tabt* 'I wrote' and *hāmm* 'important'.

2.3 *Stress*

Whereas Modern Standard Arabic → intonation seems fairly uniform over regions, word → stress in Modern Standard Arabic reflects the local colloquial dialect. Egyptian and Lebanese seem to represent two major patterns of stress and are briefly summarized here. Reference is made to *strong units*: a vowel plus two consonants VCC or a long vowel plus a consonant $\bar{V}$C, independent of syllable structure. Egyptian word stress falls on the penult after a strong unit: *tuk**all**ímu*, *tak**āt**ábū*; otherwise, it falls on the vowel of the strong unit: *kat**ábt**a*. If there is no strong unit in the word, stress falls on the first vowel in the word but not further back than the antepenult: *mátā, kátaba, inkátaba*.

Lebanese word stress falls on the last strong unit in the word; otherwise, on the antepenult or earlier: *tuk**áll**imu*, *tak**ā́t**abū*, *kat**ábt**a*, *kátaba*, *mátā*, *ink**át**aba* or ***í**nkataba*.

2.4 *Pause*

Any interruption in speech – signaled in writing by → punctuation like commas, dashes, periods – is called *pause*. In formal Arabic, when all vowels at the ends of words are pronounced, it is customary not to pronounce short vowels (and → nunation) immediately preceding pause; this form of the word with its final vowel not pronounced is called the → 'pausal form', and the form with the vowel is called the 'contextual form'.

The pausal form of the feminine ending *tāʾ marbūṭa* is *-a*, as in (1).

(1) *ʾaʿṭī-nī* *furṣa, furṣatan*
give.Imper.2fs-me chance chance
wāḥida!
one
'Give me a chance, just one chance!'

2.5 *Morphophonology*

The inflectional and derivational morphology of Modern Standard Arabic is almost totally regular, given a set of morphonological rules (Brame 1970). The following informal rules are given to illustrate some of the morphophonological changes dealt with here (C = any consonant, G = glide, V = any short vowel, $\bar{V}$ = long vowel, period [.] denotes syllable boundary and plus [+] denotes a morpheme boundary).

i. Underlying Form
a. Glide deletion. VGV → VV, except *uwa, iya*
b. Vowel assimilation. aV → aa
 Vowel rewrite. $V_1 + V_1 \rightarrow \bar{V}_1$
c. Vowel shortening. $\bar{V}$ + C. → VC
Surface Form
huday+u+n 'guidance [indef. noun]'
a. **hudayun* > *hudaun*
b. *hudaan*
 hudān
c. *hudan*
 hudan 'guidance'

ii. Another instance of Vowel shortening is **ixtār-+tu* > *ixtartu* 'I chose'.

iii. Vowel Syncope: $C_1VC_1V \rightarrow C_1C_1V$ **madada* > *madda* 'he stretched'; **yaḥmariru* > *yaḥmarru* 'it turns red', but *yaḥmarirna* 'they [fem. pl.] turn red': no change.

3. LEXICON

The Modern Standard Arabic word is built on a root-and-pattern system; the 'root' is a series of typically three consonants, always occurring in a fixed sequence that has lexical identity. The consonants making up a root are called 'radicals' and may number as many as four (*TaRJaMa* 'to translate'), five (*SaFaRJaL* 'quince'), or six (*QaNDaLaFT* 'sexton'). Thus, the root *k-t-b* means 'to write; writing', *s-k-n* means 'to dwell, reside', and *ṭ-ʿ-m* means 'to eat, savor'. A 'pattern' is a fixed framework of consonants and vowels that likewise has lexical meaning, e.g. the pattern *maFʿaL* means 'noun of place'; the root is variable but the *ma* and the vowel *a* before the last radical are obligatory. Following the Arab tradition, the verb *faʿala* 'to do, make', with the root *f-ʿ-l*, serves to model all verbs. Substituting a root *x* in place of the model root *f-ʿ-l* produces a noun meaning 'place where *x* takes place', like *maktab* 'office; desk'; *maskan* 'dwelling place, residence'; and *maṭʿam* 'restaurant'. In principle, all native Modern Standard Arabic words except particles are subject to this analysis (→ root).

Since the → glides *w* and *y* are subject to morphophonological change, verbs containing a glide as a radical show differences in verb conjugation (→ weak verbs).

Word compounding is foreign to Semitic morphology, but Modern Standard Arabic has produced, under foreign language pressure, loan translations and calques, e.g. *al-huwa/ al-hū* 'the id', *ladāʾin* 'plastics', *raqmiyy* 'digital', *šarq-ʾawsaṭiyy* 'Middle Eastern', *lāmarkaziyya* 'decentralization', *lāʿuḍwiyy* 'inorganic'; and new forms, e.g. *ʿawlama* 'globalization' (from *ʿālam* 'world'), *taxaṣxaṣa* 'privatization' (*xāṣṣ* 'private'); and → compounds, e.g. *kahrūmaġnāṭīsiyy* 'electromagnetic'. Finally, there is outright borrowing, e.g. *dīktātūriyy* 'dictatorial', *sīkūlūjiyyan* 'psychologically', *jīyūsiyāsiyy* 'geopolitical'. Various → Arabic language academies publish lists of recommended new technical → terminology.

The relative adjective (gentilic, → *nisba*) suffix *-iyy-u-n* is added to nouns to make adjectives: *qamar* 'moon' – *qamariyy* 'lunar'; *al-ʿirāq* – *ʿirāqiyy* 'Iraqi'; *madīna* 'city' – *madaniyy* 'civil, civilian'. The feminine gentilic suffix *-iyyat-un* is an important word formant for creating new abstractions, e.g. *madaniyya* 'civilization'; *madyūn* 'indebted' – *madyūniyya* 'indebtedness'; *sahrān* 'sleepless, awake' – *sahrāniyya* 'vigil [in church]'; *ḥākim* 'ruler' – *ḥākimiyya* 'rule, dominion'. It is even added to abstract nouns, e.g. *istiqlāl* 'independence' – *istiqlāliyya* 'freedom of choice, personal freedom, privacy'. It is also used as a human collective plural (→ collective).

4. Morphology

4.1 *Parts of speech*

The traditional Arab classification of → parts of speech into → *ism* 'noun', → *fiʿl* 'verb', and → *ḥarf* 'particle' is an apt one: noun morphology fits all nominal forms well, including adjectives and pronouns; verbs have a unique set of inflections; and particles are indeclinable.

4.2 *Nouns*

'Nouns' include the subclass adjective, which is treated here with nouns, with degree, the main distinction between them, treated under 'Adjectives'. Nouns are inflected for case, number, gender, and determination.

4.2.1 Case

The noun has three cases, nominative, genitive, and accusative. In brief, the nominative is the citing case, naming the subject in the clause, the predicate in certain clauses, items in titles and lists, vocatives, etc. The genitive modifies other nouns, expressing possession, origin, measure, etc. And the accusative is the adverbial case, identifying verbal objects and modifying verbs, phrases, and other units (→ *tamyīz*). For examples see Section 5.

There are two noun declensions, 'triptote' with a three-case declension and 'diptote' with a two-case one. Triptotes have a distinctive vowel for each of the three cases, nominative *u*, genitive *i*, and accusative *a*, in both singular and plural; all nouns share the same dual declension (see Table 5). The illustration in Table 3 is with nunation *-n* (see below, 4.2.4); *rajulun* 'man' is masculine singular.

Diptotes have the inflection *-a* for both the accusative and genitive cases in both the singular and the plural; further, they do not take nunation (→ diptosis).

Diptotes made definite formally become triptotes, e.g. *min makātib-a jadīdat-in* 'from new offices', but *min makātib-i l-jāmiʿat-i* 'from the offices of the university' (defined by genitive construct), and *fī lubnān-a* 'in Lebanon', but *fī hāḏā l-lubnān-i* 'in this Lebanon' (with definite article). If proper noun diptotes are used as common nouns, they become triptotes with nunation, signifying a nonspecific referent, e.g. *hāḏihi ʿammān-u* 'this is Amman', but *hāḏihi ʿammān-u-n ṯāniyat-u-n* 'this is a different Amman'. (See Table 4)

Table 3. Triptote declension

	singular ‘a noble man’	plural ‘noble men’
nominative	*rajul-u-n karīm-u-n*	*rijāl-u-n kirām-u-n*
genitive	*rajul-i-n karīm-i-n*	*rijāl-i-n kirām-i-n*
accusative	*rajul-a-n karīm-a-n*	*rijāl-a-n kirām-a-n*

Table 4. Diptote declension

	singular ‘Lebanon’	plural ‘offices’
nominative	*lubnān-u*	*makātib-u*
genitive/ accusative	*lubnān-a*	*makātib-a*

Diptotes fall into the following semantic and morphological classes: foreign names, e.g. *ʾibrāhīm-u* ‘Abraham’, *landan-u* ‘London’, although European names are generally not vocalized in Modern Standard Arabic; feminine proper names ending in *tāʾ marbūṭa*, like *karīmat-u* ‘Karima’; words of specific patterns such as the elative pattern *ʾaFʿaL-u*, e.g. *ʾakbar-u* ‘greater’; and broken plural patterns containing four consonants (excluding gender and case markers), e.g. *makātib-u* ‘libraries’, *fanājīn-u* ‘cups’, *ʾasātiḏat-u* ‘professors’, *ʾakābir-u* ‘older; seniors’, *jarāʾid-u* ‘newspapers’. The feminine suffix *-āʾ-u* (آءُ -) is diptotic, as in *kibriyāʾ-u* ‘hubris’.

Nouns ending in long vowels are invariable, showing no change in case inflection; this includes ‘indeclinable’ nouns (not taking nunation), like those ending in the feminine suffix *ʾalif maqṣūra* (ى), e.g. *ḏikrā* ‘remembrance’, *hadāyā* ‘gifts’, and proper names like *sāmī* ‘Sami’; and ‘declinable’ nouns (accepting nunation), exemplified by defective-root nouns, e.g. هُدىً *hudan* (for **huday-u-n*) ’right guidance’.

4.2.2 Gender

→ ‘Gender’ in nouns, masculine and feminine, may be indicated by the presence or absence of feminine markers or by semantic content. Gender is determined in the first instance by the nature of the referent: if the noun has an animate referent, grammatical gender reflects natural sex; thus, masculine are: *rajul* ‘man’, *xalīfa* ‘Caliph’, *ḥiṣān* ‘horse’, and feminine: *rāqiṣa* ‘danseuse’, *ʾumm* ‘mother’, *faras* ‘mare’. Some unmarked nouns are of common gender: *ʿajūz* ‘old woman; old man’; *nāʾib* ‘deputy’ (masc. or fem.; *nāʾib-a* means ‘disaster’).

Names of cities are all feminine, reflecting the word *madīna* ‘city’, such as *bayrūt-u* ‘Beirut’, *bārīs* ‘Paris’. Country names are mostly feminine, agreeing with *dawla* ‘the state (of)’: *miṣr* ‘Egypt’, *al-kuwayt* ‘Kuwait’, except for a few with the definite article, which are masculine: *al-ʾurdunn* ‘Jordan’, *as-sūdān* ‘the Sudan’, *al-ʿirāq* ‘Iraq’, *al-maġrib* ‘Morocco’. The names *lubnān* ‘Lebanon’ and *al-yaman* ‘Yemen’ may be either gender.

Most parts of the body that come in pairs are feminine: *yad* ‘hand’, *ʿayn* ‘eye’; but *xadd* ‘cheek’ is masculine. A few unmarked common nouns are feminine by convention: *ʾarḍ* ‘earth, ground’, *ḥarb* ‘war’, *sūq* ‘marketplace’, *šams* ‘sun’.

The usual feminine marker is *-at-un* (pausal form *-a*), called *tāʾ marbūṭa* (ة), as in *malik* ‘king’ and *malika* ‘queen’. Other feminine markers are the suffix *-āʾ-u* (آءُ) on singular nouns or adjectives, e.g. *bayḍāʾu* ‘white’, and the suffix *ʾalif maqṣūra* (ى), as in *ḏikrā* ‘remembrance’.

4.2.3 Number

The bare noun is either singular or a → collective. The singular noun can be inflected for dual

Table 5. Dual inflections

	masculine		feminine	
	sayyid-u-n	‘gentleman’	*sayyidat-u-n*	‘lady’
nominative	*sayyid-ā-ni*	‘two gentlemen’	*sayyidat-ā-ni*	‘two ladies’
genitive accusative	*sayyid-ay-ni*		*sayyidat-ay-ni*	

Table 6. Case inflections: Sound plurals

	masculine		feminine	
	singular	plural	singular	plural
nominative	*-u-n*	*-ū-na*	*-at-u-n*	*-āt-u-n*
genitive	*-i-n*	*-ī-na*	*-at-i-n*	*-āt-i-n*
accusative	*-a-n*		*-at-a-n*	

or pluralized by means of suffixation (called the 'sound plural') or by internal vowel or pattern change ('broken plural'). With the dual and sound plural endings the genitive case has been generalized to include accusative case functions, resulting in two-case declensional systems.

The singular noun is dualized by the addition of the dual suffix *-ā-ni*, where the suffix *-ni* is a form of nunation in the dual (but without the meaning of indefinite): *sayyid-ā-ni* 'two gentlemen', *sayyidat-ā-ni* 'two ladies', *yad-ā-ni* 'two hands' (see Table 5).

The sound plural endings are lengthened forms of singular inflectional vowels. Thus, the masculine sound plurals *-ū-na* (nominative plus 'nunation') and *-ī-na* (genitive/accusative plus 'nunation') are lengthened forms of the masculine singular *-u-n* (nom.) and *-i-n* (gen.). The feminine sound plurals involve lengthening of the feminine suffix vowel itself, *-at-u-n* > *-āt-u-n*; here, too, the feminine genitive plural includes the accusative plural functions (see Table 6).

'Broken plurals' are ablaut forms (→ apophony) of the singular, e.g. *kitāb-un/kutub-un* 'book/books', *qalb-un/qulūb-un* 'heart/hearts'; note the polarity in vowel length in the two forms. Broken plurals may include prefixes, as in *kalb-un/ʾa-klāb-un* 'dog/dogs' or suffixes, as in *fāris-un/fursān-un* 'horseman/horsemen'. If the singular has four or more metric units (consonant or long vowel), a fixed broken plural pattern may be accessed: *fāris-un/fawāris-u* 'horseman/horsemen'. These patterns are largely predictable: briefly, any four-unit singular or a singular with five consonants takes a plural of the diptote type CaCāCiC-u: *maktab-un/makātib-u* 'office/offices', *safarjal-un/safārij-u* 'quince/quinces' (the fifth radical is dropped). A second-position long vowel is replaced by *w* and a third-position long vowel is replaced by *ʾ*: *šāriʿ-un/šawāriʿ-u* 'street/streets', *jarīdat-un/jarāʾid-u* 'newspaper/newspapers', *ʿajūz-un/ʿajāʾiz-u* 'old man/old men; old woman/old women'. A five-unit noun containing a long vowel takes the diptote pattern CaCāCīC-u, as in *finjān-un/fanājīn-u* 'cup/cups', and a five-unit noun with two long vowels takes the pattern CawāCīC-u, *mīzān-un/mawāzīn-u* 'scale/scales'. Specialized patterns preempt these rules: singular nouns of the pattern *FāʿiL-un*, referring to occupations or customary activities, for example, take the plural *FuʿʿāL-un*, e.g. *kātib-un/kuttāb-un* 'writer/writers'. and the pattern *ʾaFāʿiLat-u* is used for four- and five-unit human nouns, as in *ʾustāḏ-un/ʾasātiḏat-u* 'professor/professors', *mārūniyy-un/marāwinat-un* 'Maronites'.

The collective noun refers to a class of items or substances, and not to an individual member of such a class. Collective nouns are a kind of abstraction that includes the whole class of items and so cannot be counted or measured, but are in contrast with other types of object or material. They fall into three classes, material or smaller animals, human, and animal collectives.

Material collectives are the major type: nouns that are masculine singular in form but plural or unlimited in meaning: *šajar-u-n* 'trees', *laḥm-u-n* 'meat' *naml-u-n* 'ants'. A 'noun of unity' may be formed from material collectives by the addition of *-a* (*tāʾ marbūṭa*), which can then be pluralized and counted: *waraq-un* 'paper'/*waraq-at-un* 'a piece, sheet of paper': *xams-u waraq-āt-in* 'five sheets of paper'. The feminine plural of the noun of unity is not only countable, but, by contrast with the singular, it tends to have a specific referent, e.g. *al-mawz-u laḏīḏ-un* 'bananas are delicious' but *kānat-i-l-mawzātu l-latī ʾakalnā-hā l-yawma laḏīḏat-an* 'the bananas we ate today were delicious'. These collectives may also have broken plurals, which then have the general meaning of various types of object: *šajar-un* 'trees [coll.]/*šajarat-un* 'a tree'/*šajarāt-un* 'trees [countable]'/*ʾašjār-un* 'trees, shrubs'; *laḥm-u-n* 'meat'/*laḥmat-u-n* 'a piece of meat'; *al-laḥmāt-u* 'the meat [specific]'/ *luḥūm-u-n, liḥām-u-n* 'meats'.

Human collectives which refer to ethnic or national groups are masculine plural; the individual is indicated by the gentilic suffix *-iyy-un*, e.g. *ʾarmaniyy-un* [masc.], *ʾarmaniyyat-un* [fem.]/*ʾarman-un* 'Armenian/Armenians'. Some human collectives may also have broken plurals, e.g. *kurdiyy-un/kurd-un*, *ʾakrād-un* 'Kurd/Kurds'. The feminine singular of the gentilic suffix may also have a collective plural meaning, e.g. *al-kuwaytiyya* 'the Kuwaitis'.

Collectives referring to larger animals are feminine singular and have no noun of unity but may have a broken plural: *ʾibil-un* 'camels'; *xayl-un/xuyūl-un* 'horses'; *ġanam-un/ʾaġnām-un* 'sheep; sheep and goats'.

4.2.4 Determination

Common nouns are made definite with the definite article, *al-*, e.g. *al-kitāb-u* 'the book', and indefinite nouns receive → nunation (→ *tanwīn*), e.g. *kitāb-u-n* 'a book'. The functions of nunation in the dual and plural are quite different from those of the singular. The form of the nunation is *-n* in the singular, *-ni* in the dual, and *-na* in the plural. The dual and plural forms require a suffixed vowel, since by phonological rule a syllable cannot end in a long vowel and a consonant; the vowel is the opposite of the preceding long vowel in height, that is, low *ā* + high *-ni* and high *ī* + low *-na*. Every triptote noun in its citation form receives nunation, unless it is blocked by addition of the definite article, or if the noun becomes formally definite as the first term of an → *ʾiḍāfa*, as in *kitābu l-ʾaġānī* 'the Book of Songs', or receives a suffixed pronoun, *kitābu-hu* 'his book'.

4.2.5 Adjectives

→ 'Adjectives' are a subclass of nouns in that they share the inflectional features of nouns and may perform some of the functions of nouns; they are distinctive in that only adjectives may be inflected for comparative/superlative degree (the → 'elative'), and have some distinctive word patterns not shared by nouns.

The elative form proper is the diptote *ʾaFʿaL-u*, e.g. the elative of *kabīr-un* 'big, great' is *ʾakbar-u*. When indefinite it means 'bigger, greater', and the compared item is introduced by *min* 'than'; in this usage it is inflected for case but not for gender or number, e.g. *karīmat-u ʾakbar-u min karīm-in* 'Karima is older than Karim' and *ʾilā mudun-in ʾaqdam-a* 'to older cities'. When definite, whether by the definite article or as the first term of an *ʾiḍāfa*, it has superlative meaning, becoming a triptote and showing full agreement with the modified noun; the full nominative case paradigm, including the dual, is for the masculine *ʾafʿal-u – ʾafʿalā-ni – ʾafāʿil-u, ʾafʿal-ū-na*, and for the feminine *fuʿlā – fuʿlayā-ni – fuʿlayāt-un*.

The plural elative forms may serve as adjectives or as substantives, e.g. *al-ʾasmāʾ-u l-ḥusnā* 'the Exquisite Names [i.e. the 99 names of God]'; *ad-duwal-u l-kubrā* 'the great powers'; *quwā d-duwal-i* 'the most powerful [sg. or pl.] of the nations'; *maʿa ʾakābir-i l-jāliyat-i* 'with the leaders of the community'.

The masculine singular positive form and the masculine and feminine singular elative may be used attributively in fixed terms or official names, as in *lubnān-u l-kabīr-u* 'Greater Lebanon'; *aš-šarq-u l-ʾawsaṭ-u* 'the Middle East'; *bariṭāniyā l-ʿuḏ̣mā* 'Great Britain; *ad-duwal-u l-kubrā* 'the great powers'. This amounts to a kind of lexical capitalization.

Some basic adjective patterns are *faʿīl-un*, which indicates features by which something is characterized, e.g. *kabīr* 'big', *fahīm* 'intelligent'; and *faʿlān-un*, which indicates a resultant condition, e.g. *taʿbān-un* 'tired'.

The noun *ḏū* 'possessor of' plus a genitive

Table 7. Independent pronouns

person		singular	dual	plural
1 comm.		*ʾana* 'I'	—	*naḥnu* 'we'
2	masc.	*ʾanta* 'you'	*ʾantumā* 'you'	*ʾantum* 'you'
	fem.	*ʾanti* 'you'		*ʾantunna* 'you'
3	masc.	*huwa* 'he, it'	*humā* 'they'	*hum* 'they'
	fem.	*hiya* 'she, it'		*hunna* 'they'

Table 8. Suffixed pronouns

person		singular	dual	plural
3	masc.	-*hu* ‘him, it’	-*humā* ‘them’	-*hum* ‘them’
	fem.	-*hā* ‘her, it’		-*hunna* ‘them’
2	masc	-*ka* ‘you’	-*kumā* ‘you’	-*kum* ‘you’
	fem.	-*ki* ‘you’		-*kunna* ‘you’
1 comm.		-*ī* (gen.) ‘my’	—	
		-*nī* (acc.) ‘me’		-*nā* ‘us’

noun serves an adjectival function, e.g. *ḏū jadwā* ‘advantageous, beneficial’.

Some feminine plural adjectives may serve as neuter plural nouns, like *maʿlūm* ‘known’ – *maʿlūmāt* ‘known things, data’, from which comes the back formation *maʿlūma* ‘a piece of information, datum’.

4.2.6 Pronouns

→ Pronouns are independent (Table 7) or suffixed (Table 8) (→ clitics). They are inflected for person, number, and gender.

The dual pronouns consist of the plural pronoun plus -*ā* (*hum-ā*), and the feminine plurals are the masculine plural plus the feminine plural suffix -*na* (*hum* + *na* > *hunna*).

The independent pronouns are basically nominative in case and serve as (i) subjects in verbless sentences, as in *ʾana l-ʾustāḏa mahā* ‘I am Professor Maha’, or for contrastive focus or special highlighting of the pronoun, as in *wa-māḏā ʿamil-tu la-ka ʾana*? ‘and what did *I* do to you?’, where the 1st person agent is expressed in the verb (-*tu*) and in the pronoun (*ʾana*). The suffixed pronouns perform genitive and accusative functions, e.g. *ismu*-ka ‘your name’ (genitive as second term of an *ʾiḍāfa*); *māḏā qāla la*-ka ‘what did he say to you?’ (genitive as object of preposition); *la-qad zurnā*-ka *ʾalfa marra!* ‘we have visited you a thousand times!’ (accusative as direct object); *ʾinna*-ka *ʿārifun kulla ḏālika* ‘but you know all that!’ (accusative as subject after *ʾinna*); *ʾinna-hu ṣadīqu*-ka ʾanta ‘he’s *your* friend!’ (two pronominal references to one person).

Suffixed pronouns beginning in *hu*- change the *u* to *i* after *i, ī*, or *ay*: ‘his book’ is *kitābu-hu* (nom.) but *kitābi-hi* (gen.).

→ Relative pronouns are inflected for case, number, and gender (see Table 9).

The relative pronoun links a definite antecedent and the modifying clause, which must contain a pronominal or agreement reference to the antecedent (→ *ʿāʾid*). The relative pronoun does not enter into the syntax of the relative clause, as in (2) and (4). If the antecedent is indefinite, there is no relative pronoun (→ *ṣifa*), as in (3).

(2) *ʾaʿrifu* *l-kātiba* *l-laḏī*
I-know the-author who
kataba *ḏālika*
write.Perf.3ms that
‘I know *the* author who wrote that’

Table 9. Relative pronoun

	singular		dual		plural	
	masc.	fem.	masc.	fem.	masc.	fem.
nom.	*al-laḏī*	*al-latī*	*al-laḏānī*	*al-latānī*	*al-laḏīna*	*al-lātī,* *al-lawātī,* *al-lāʾī*
gen.			*al-laḏaynī*	*al-lataynī*		
acc.						

Table 10. Demonstrative pronoun *hāḏā*

	singular		dual		plural	
	masc.	fem.	masc.	fem.	masc.	fem.
nom.	*hāḏā*	*hāḏī,* *hāḏihi,* *hātihi*	*hāḏāni*	*hātāni*		*hāʾulāʾi*
gen./acc.			*hāḏayni*	*hātayni*		

Table 11. Perfect tense

	singular	dual	plural
3rd masc.	*katab-a* 'he wrote'	*katab-ā* 'they wrote'	*katab-ū* 'they wrote'
3rd fem.	*katab-at* 'she wrote'	*katab-atā* 'they wrote'	*katab-na* 'they wrote'
2nd masc.	*katab-ta* 'you wrote'	*katab-tumā* 'you wrote'	*katab-tum* 'you wrote'
2nd fem.	*katab-ti* 'you wrote'		*katab-tunna* 'you wrote'
1st comm.	*katab-tu* 'I wrote'	—	*katab-nā* 'we wrote'

(3) *ʾaʿrifu* *kātiban* *kataba*
I-know writer write.Perf.3ms
ḏālika
that
'I know *a* writer who wrote that'

(4) *man* *al-kātibu* *l-laḏī*
who the-writer who
katabta *ʿan-hu?*
write.Perf.2ms about-him
'Who is the writer about whom you wrote?'

→ Demonstrative pronouns are of two degrees, *hāḏā* (proximal: 1st person 'this' and 2nd person 'that'; see Table 10) and *ḏālika* (distant: 3rd person 'yon, that over there') (→ deixis).

The masculine singular demonstrative pronoun *ḏālika* 'that' has the feminine singular *tilka* and the common plural *ʾulāʾika* 'those'.

The demonstratives are built on *hā* 'this', *ḏā* 'that', and the demonstrative *l* (Wright 1951:I, 267). The first two often serve as an envelope for the independent pronouns, as in the word *hāʾanāḏā* 'behold, here I am'. The demonstrative pronouns are used as independent pronouns, as in (5), and attributively with a following noun having the definite article, as in (6).

(5) *sa-yakūnu* *ḏālika* *ḥasan-an*
Fut-be.Ind.3ms that good-Acc
'That will be fine'

(6) *ʾurīdu* *hāḏā* *l-kitāb-a*
want.1s this the-book-Acc
wa-tilk-a *l-waraqat-a*
and-that the-paper-Acc
'I want this book and that sheet of paper'

4.3 *Verbs*

4.3.1 Conjugation

The Arabic verb has two tenses, the perfect and the imperfect; it is inflected for person (1st, 2nd, 3rd), number (singular, dual, plural), and gender (masculine, feminine, common). The perfect tense paradigm is given in Table 11 using the verb *kataba* (*u*) 'to write' as a model; the imperfect stem vowel is given in parentheses.

The perfect tense is inflected solely with suffixes. The dual formative *-ā* is added to the 2nd person masculine plural for the 2nd person dual and to the 3rd person singular forms for the 3rd person duals.

The → defective verb *laysa* 'not be' is conjugated only in the perfect, albeit with present tense meaning: *laysa* 'he is not'. It has the stem *lays-* before vowel-initial endings (*lays-at* 'she is not', *lays-ū*, etc.) and the stem *las-* before consonant-initial endings (*las-tu* 'I am not', *las-tum*, etc.).

The imperfect indicative is given in Table 12; prefixes denote person and (in the 1st person) number, while suffixes denote number, gender, and mood.

Table 12. Imperfect indicative

	singular	dual	plural
3rd masc.	*ya-ktub-u* 'he writes'	*ya-ktub-ā-ni* 'they write'	*ya-ktub-ū-na* 'they write'
3rd fem.	*ta-ktub-u* 'she writes'	*ta-ktub-ā-ni* 'they write'	*ya-ktub-na* 'they write'
2nd masc.	*ta-ktub-u* 'you write'	*ta-ktub-ā-ni* 'you write'	*ta-ktub-ū-na* 'you write'
2nd fem.	*ta-ktub-ī-na* 'you write'		*ta-ktub-na* 'you write'
1st comm.	*ʾa-ktub-u* 'I write'	—	*na-ktub-u* 'we write'

The preformatives are subject markers: *ʾa-* and *na-* 1st person singular and plural respectively; *ta-* may denote any 2nd person verb as well as 3rd person feminine singular and dual; and *ya-* denotes all 3rd person masculine as well as 3rd feminine plural. The long vowels and the suffix *-na* contiguous to the stem are gender/number markers: *-ī-* is feminine singular, *-ā-* is dual, *-ū-* is masculine plural, and *-na-* is feminine plural. The indicative mood marker is *-u* when attached directly to the stem or *-ni*/*-na* after a long vowel. The feminine plurals are a special case, the indicative suffix *-na* originally found after the feminine plural suffix *-na* having been lost via haplology (**ya-ktub-na-na* > *ya-ktub-na*).

The verb *raʾā*/*yarā* 'to see' loses its middle radical glottal stop *ʾ* in the imperfect: *raʾaytu* 'I saw'/*ʾarā* 'I see'. For the conjugation of verbs with roots containing glides, → weak verbs.

4.3.2 Tense

The primary meaning of the → tenses is that of → 'aspect': the perfect, of perfective aspect, denotes an event, an occurrence with some kind of result or conclusion, e.g. *šaribat finjān-an min aš-šāy* 'she drank a cup of tea'; *ḥasunat al-ʾaḥwāl-u baʿda ḏālika* 'conditions got better after that'.

The imperfect tense is imperfective in aspect, denoting actions or states without a beginning or ending point: habitual, progressive, predictive ('future'), generic, stative, etc.: *yaʾkul-u* 'he eats, is eating, will eat/is going to eat'. Generic meaning is common in statements of common knowledge, e.g. *al-qamar yaṭlaʿu kulla masāʾ* 'the moon rises every evening', and proverbs, e.g. *al-kilāb al-latī taʿwī lā taʿaḍḍ-u* 'dogs that bark don't bite'. The other sub-aspects are denoted by adverbial contexts: *kulla marra* 'every time', and *dāʾiman* 'always', for example, signal habitual action. Punctual time expressions like *baʿda sāʿatayni* 'in a couple of hours' and *fī s-sāʿa ar-rābiʿa* 'at four o'clock' are predictions of future action.

It should be noted that, except for the verb *kāna* and its negative *laysa*, all verbs in the perfect denote an event – something which took place. Accordingly, 'stative' verbs in the perfect denote *a change of state*. The perfect *ʿaraftu* does not mean 'I knew' but 'I learned, found out' – an event. The equivalent of English 'I knew' is *kuntu ʾaʿrifu* – a state in past time.

The tenses differ also in 'timing': the perfect denotes an event that was completed before the moment of speaking, i.e. anteriority of action: *qaraʾtu-hā s-sanata l-latī fātat* 'I read it last year' (past time); *qaraʾtu-hā l-yawm-a* 'I read it today' (present time). The imperfect depicts an action or state co-occurring with the main verb of the utterance or, if it is the matrix verb, with the utterance itself, denoting simultaneity of action: *māḏā tadrus-u, yā salīm*? 'what are you studying, Salim?'; *wa-māḏā kunta tadrusu*? 'and what were you studying?'; *raʾaytu-hu wa-huwa yadrusu* 'I saw him as he was studying'. Time/tense is set by the adverbial context: 'today', 'yesterday', 'last year', etc. The verb *kāna* functions as a time switcher: it removes the timing from a present context to a past one, anterior to the time of the utterance. The verb *kāna* plus imperfect depicts habitual, progressive, etc., in past time (*kāna yadrusu* 'he was studying, used to study, studied', *kāna sa-yadrusu* 'he was going to study'), and *kāna* plus a perfect tense verb produces a past perfect tense, with *qad* preceding the main verb: *kāna l-ʾustāḏ qad nasiya d-dars* 'the professor had forgotten the lesson'. The imperfect of *kāna, yakūnu*, has only future meaning, ‘will be’; with *qad* and the perfect tense of the main verb, it provides future perfect tense: *sa-yakūnū qad taġaddaw* ‘they will have had lunch’.

The two tenses differ also in ‘function’: the perfect denotes events in a narrative, whereas the imperfect is situational, depicting background information against which the events of the perfect tense are foregrounded, as in

> *kāna dawman yasbaḥu maʿa rufaqāʾi-hi fī birka ṣaġīra laysa baʿīda jiddan ʿan bayti-hi wa-lākin ḏāta yawmin ʿalā ġayri ʿāda ḏahaba waḥda-hu ʾilā nahri dijla wa-sabaḥa bi-mufradi-hi* ‘he would always swim [imperf.] with his friends in a small pond not far from his house, but one day, contrary to his custom, he went [perf.] to the Tigris and swam [perf.] alone’

The perfect answers the question ‘What happened?’, and the imperfect, ‘How were things?’.

In addition to the primary functions above, each of the tenses has secondary functions in specific contexts. Both tenses may have → ‘performative’ force, i.e., an act is performed by the very uttering of the verb itself. For example, one cannot make a promise without saying ‘I promise’. Performatives are typically 1st person singular active voice; in some way or other, they create a change, whether in the hearer’s or the speaker’s own condition or behavior. There are five classes of performative. Two types are factive, asserting the truth value of the proposition. One, Assertives, commits the speaker to belief in or support of the proposition, such as *ʾuʾakkidu/ʾaʿtarifu ʾanna* ‘I assure/admit that…’. The other, Declarations, is a formal statement affecting the official status of a person or thing; it is typically uttered by an official representing an institution such as the government, church, etc. and sometimes occurs in the 3rd person, perfect tense, or passive, such as *ʾuʿlinu istiqālatī* ‘I [hereby] announce my resignation’; *ʿayyannā-ka raʾīsan li-š-šarika* ‘we hereby appoint you president of the company’; *tuʾammimu šarikatu l-qanāti* ‘the canal company is hereby nationalized’. Two classes are nonfactive and require changes in behavior, either on the part of the hearer – Directives – or of the speaker – Commissives. Unlike the first two classes, which take *ʾanna* clause complementation, these two take *ʾan* clauses. Some Directives are *ʾaṭlubu* ‘I request’ and *ʾamnaʿu* ‘I forbid’, and some Commissives are *ʾa ʿidu* ‘I promise’ and *ʾarfuḍu* ‘I reject’. The fifth class, Expressives, expresses the speaker’s psychological state concerning the propositional content, such as *ʾaškuru-ka* ‘I thank you’ and *ʾaʿtaḏiru ʿan* ‘I apologize for…’ (Khalil and McCarus 1999).

The imperfect is inherently situational in function, but the perfect may assume that function after the circumstantial particle *qad*, describing a resultant situation, as in *qad ʾataw* ‘they have come; they are here’. On the other hand, in fiction or other narrative contexts, such as events in progress, the imperfect often performs a perfect tense function, the ‘historical narrative’, wherein the writer can choose either the perfect or the imperfect to denote events in the context of the narrative: *yuwaddiʿu ʾaṣḥāba-hu wa-yarkabu t-tāksī* ‘he bids his friends goodbye and gets into the taxi’. Lastly, the perfect (negated by *lā*) has ‘optative’ force in the context of an oath: *lā ʾaflaḥa!* ‘may he not succeed!’.

The semantic structure of the verb is summarized in Table 13.

4.3.3 Mood

The imperfect distinguishes four → moods, the indicative, subjunctive, jussive, and imperative. The mood of the verb expresses the attitude of the speaker to the truth value of the statement: the indicative verb declares that the statement is a factual or truthful one, e.g. *tatakallamu l-ʿarabiyya wa-l-fārisiyya* ‘she speaks Arabic and Persian’. The subjunctive makes no claim about the actual realization of the denotation of the verb but instead presents the notion in the abstract; it deals with the unrealized act, not a fact, e.g. *yanbaġī ʿalay-ka ʾan tatakallama l-ʿarabiyya* ‘you ought to speak Arabic’. The jussive has two primary uses, (i) a relic function as a perfect tense after the negative particle *lam* ‘did not’, e.g. *lam tatakallam ʿarabiyya* ‘you did not speak Arabic’, and in both clauses of the conditional sentences with *ʾin* ‘if’, e.g. *ʾin tadrus tanjaḥ* ‘if you study you will succeed’; and (ii) as an indirect command after the exhortative particle *li-* (*l-* after *wa-*, *fa-*) ‘let, have…’, as well as a (iii) negative imperative after *lā*, as in *lā taqul ḏālika* ‘don’t say that!’. This indirect command is generally in the 3rd person, e.g. *li-yadxul ḥālan* ‘have him come in immediately!’, but it also occurs often in the 1st person plural, e.g. *li-našrab finjān-an min al-qahwa* ‘let’s have

Table 13. The Modern Standard Arabic tenses

	PERFECT TENSE	IMPERFECT TENSE
A. FUNCTION		
1. Primary	Narrational: *ʿaraftu ḏālika minka* ‘I learned that from you’	Situational: *ʾaʿrifu ḏālika* ‘I know that’
2. Secondary		
a. Performative	*qabiltu* ‘I accept’ [wedding ceremony]	*ʾaʿidu bi-ḏ-ḏahāb maʿa-kum* ‘I promise to go with you’
b. Situational	*qad waṣalat* ‘she has arrived; she’s here now’	
c. Narrational		*taqifu fajʾatan fa-taqūlu…* ‘she stops suddenly and says…’
d. Optative	*waffaqa-ka ḷ-ḷāhu* ‘may God grant you success’	
B. ASPECT		
1. Perfective	*šaribnā šāy* ‘we had tea’	
2. Imperfective		
a. Habitual		*yaktubu kulla yawm* ‘he writes every day’
b. Progressive		*yaktubu l-ʾāna* ‘he’s writing now’
c. Dispositional		*yaktubu jayyidan* ‘he writes well’
d. Predictive (Future)		*sa-yaktubu qarīban* ‘he’ll write soon’
e. Generic		*tušriqu š-šamsu min aš-šarq* ‘the sun rises in the east’
f. Stative		*yujīdu l-kitāba* ‘he is good at writing’
C. TIMING		
1. Anteriority	*kataba-hā ʾamsi* ‘he wrote it yesterday’	
2. Simultaneity		*šāhadtu-hu wa-huwa yaktubu-hā* ‘I saw him as he was writing it’

Table 14. Subjunctive verb

	singular	dual	plural
3 masc.	*ya-ktub-a* ‘that he write’	*ya-ktub-ā* ‘that they write’	*ya-ktub-ū* ‘that they write’
3 fem.	*ta-ktub-a* ‘that she write’	*ta-ktub-ā* ‘that they write’	*ya-ktub-na* ‘that they write’
2 masc.	*ta-ktub-a* ‘that you write’	*ta-ktub-ā* ‘that you write’	*ta-ktub-ū* ‘that you write’
2 fem.	*ta-ktub-ī* ‘that you write’		*ta-ktub-na* ‘that you write’
1 comm.	*ʾa-ktub-a* ‘that I write’	—	*na-ktub-a* ‘that we write’

a cup of coffee’, and, rarely, in the 2nd person, e.g. *fa-l-taʿrif ʾannanī…* ‘so know that I…, you should know that I…’. The imperative issues a direct command, e.g. *iḏhab ʾilā n-namlat-i yā kaslān-u!* ‘go to the ant, thou sluggard!’.

The subjunctive (Table 14) differs from the indicative by substituting *-a* for *-u* and deleting the indicative suffixes *-ni* and *-na*, which follow a long vowel.

The jussive (Table 15) is a shortened form of the subjunctive: the short vowel *-a* is deleted; in phrases a short *i* is suffixed before a word beginning in two consonants (→ epenthesis).

The imperative is like the jussive but without

Table 15. Jussive verb

	singular	dual	plural
3 masc.	*ta-ktub* ‘(that) she write’	*ta-ktub-ā* ‘(that) they write’	*ya-ktub-na* ‘(that) they write’
3 fem.	*ya-ktub* ‘(that) he write’	*ya-ktub-ā* ‘(that) they write’	*ya-ktub-ū* ‘(that) they write’
2 masc.	*ta-ktub-ī* ‘(that) you write’	*ta-ktub-ā* ‘(that) you write’	*ta-ktub-na* ‘(that) you write’
2 fem.	*ta-ktub* ‘(that) you write’		*ta-ktub-ū* ‘(that) you write’
1 comm.	*ʾa-ktub* ‘(that) I write’	—	*na-ktub* ‘(that) we write’

Table 16. Imperative verb

	singular	dual	plural
2 masc.	*u-ktub* ‘write!’	*u-ktub-ā* ‘write!’	*u-ktub-ū* ‘write!’
2 fem.	*u-ktub-ī* ‘write!’		*u-ktub-na* ‘write!’

the preformatives (see Table 16). If the stem begins with a two-consonant cluster, a helping vowel that may be elided is prefixed: *u* before stems with the stem vowel *u*, and *i* with the others. Thus, *i-qraʾ hāḏā wa-qraʾ hāḏā!* ‘read this and read this!’, *u-ktub lī qarīb-an* ‘write me soon!’.

The ‘energetic’ suffix, *-an* or *-anna* (→ energicus), may be added to the negative jussive or the imperative to give different degrees of force to commands; it is rare in Modern Standard Arabic and has a poetic or Qurʾānic flavor, e.g. *lā taqtulanna* ‘thou shalt not kill!’, *iḏhaban!* ‘now go!’.

4.3.4 Stem vowels

The ‘stem vowel’ is the vowel occurring before the last radical, e.g. ***i*** in *fah**i**ma* ‘he understood’, *mudarr**i**s* ‘instructor’, and *yastaqd**i**mūna* ‘they receive, they welcome’. The stem vowel has particular significance in the perfect tense verb as well as in other parts of speech. Here we discuss the ‘ablaut’ patterns of stem vowels in perfect and imperfect tense verbs (→ apophony).

Verbs may be classified as action verbs or qualitative verbs, with the latter further typed as temporary or permanent states or qualities. The perfect tense stem vowel *a* typically denotes an action, and *u* a permanent quality or state, while *i* often denotes an action or a temporary state; consider the following, which all share the semantic field of being in front of, spatially or temporally:

*qad**a**ma* ‘to precede [someone]’
*qad**i**ma ʾilā* ‘to arrive at [a place]’
*qad**u**ma* ‘to be old’

There are many intransitive/transitive pairs like *ġam**u**ra/yaġm**u**ru* ‘to be copious, abundant [water]’ – *ġam**a**ra/yaġm**u**ru* ‘to flood, inundate [something]’. Table 17 shows the relationship between the stem vowels of the two tenses; notice how the low vowel *a* and the high vowels *i* and *u* are always in opposition to each other; this polarity does not obtain with qualitative verbs with *u* in the perfect.

If an adjacent radical is a laryngeal (ʾ, *h*, *ḥ*, ʿ), the imperfect stem vowel *i* assimilates to *a*, as in *fat**a**ḥa/**yaft**i**ḥu > yaft**a**ḥu* ‘to open’.

4.3.5 Derived verb Forms

The base verb has a stem consisting of the three root consonants and one vowel or two, e.g. *katab-a/ya-ktub-u* ‘to write’. In addition, there are nine other verb Forms (plus five rare ones),

Table 17. Stem vowel ablaut patterns

perfect	imperfect	illustration
a	*i*	*jal**a**sa/yajl**i**su* ‘to sit’
	u	*dar**a**sa/yadr**u**su* ‘to study’
i	*a*	*ʿam**i**la/yaʿm**a**lu* ‘to work’
u	*u*	*kab**u**ra/yakb**u**ru* ‘to grow large’

which bear a patterned relationship to the base form in terms of both morphology and semantics. Table 18 presents verb Forms I–X.

Unlike Form I there are no stem vowel variations in the derived Forms: only *a* occurs in the perfect tense, except for epenthetic *i* (→ epenthesis), and the imperfect stem vowel is *i* in all derived Forms except for *a* in Forms V and VI, which have the reflexive prefix *ta-*. The preformative is *ya-* in all but Forms II, III, and IV, where it is *yu-*. The formatives are vowel lengthening (associative) and consonant gemination (intensive, causative) and the affixes *ta-* (reflexive) and *ʾa-*/*s-* (causative). Thus, Forms II and IV are causative of Form I; Form II is intensive of Form I (e.g. *kasara* 'to break' and *kassara* 'to smash'), and Form IX is intensive of natural traits; Form V is reflexive of Form II, Form VI is reciprocal of Form III, and Form VIII is reflexive/middle of Form I; Form VII is reflexive-passive of Form I; and Form X is middle of Form IV (= causative/middle of Form I). Causatives tend to become estimative (Form II *kaḏḏaba* 'to call someone a liar [*kāḏib*]', Form X *istaḥsana* 'to make someone out to be good [*ḥasan*], to approve of'), and reflexives tend to become intransitive or passive (Form I *kasara* 'to break [trans.]', Form VII *inkasara* 'to break [intrans.]; to be broken').

Quadriliterals are verbs based on four-radical roots, e.g. *sayṭara* 'to dominate, control'. Many are biradical reduplications, e.g. *zalzala* 'to quake', often with iterative meaning like *naṭṭa* 'to jump' and *naṭnaṭa* 'to jump up and down, skip', and others are augmented triliterals, e.g. *šamaxa* 'to be lofty, towering' and *mušmaxirr* 'lofty, towering'. Basic quadriliterals QI (*FaʿLaLa*/*yuFaʿLiLu*) are conjugated like triliteral Form II verbs, and have a derived conjugation QII in *ta-* (*taFaʿLaLa*/*yataFaʿLaLa*), which is conjugated like triliteral Form V. This derived form may be reflexive or intransitive of the basic form, e.g. *daḥraja* 'to roll [trans.]' and *tadaḥraja* 'to roll [intrans.]', or simulative ('to act like or claim to be'), e.g. *taʾamraka* 'to imitate the Americans, become Americanized', *tabaġdada* 'to act like a Baghdadi, swagger', *tamaskana* 'to feign poverty [*miskīn* 'poor']'. Quadriliteral III (*iFʿanLaLa*/*yaFʿanLiLu*) is represented by *islanṭaḥa*/*yaslanṭiḥu* 'to be broad' (cf. *saṭaḥa* 'to spread out') and IV (*iFʿaLaLLa*/*yaFʿaLiLLu*), with doubled fourth radical, by *iṭmaʾanna*/*yaṭmaʾinnu* 'to be reassured' (cf. *ṭamʾana* 'to reassure') (→ *ištiqāq*).

4.3.6 Voice

→ Voice is indicated primarily by ablaut patterns: → passive voice is indicated in all perfect tense verbs by the stem vowel *i* preceded by *u* in all preceding positions in the word, and in imperfect tense forms with the preformative vowel *u* followed by *a* in the rest of the stem:

> *kataba*/*yaktubu* 'to write' and ***kutiba*/*yuktabu*** 'to be written'
> *istaqdama*/*yastaqdimu* 'to welcome' and ***ustuqdima*/*yustaqdamu*** 'to be welcomed'

Passive may also be denoted by certain derived verb Forms with basically reflexive meaning (see Table 18).

The passive voice may have the meaning of potential action, e.g. *hāḏā lā yuṣaddaqu!* 'that is unbelievable!', *maqbūl* 'accepted; acceptable'.

Passive is expressed lexically by the verb *tamma*/*yatimmu* 'to be effected, accomplished', especially in scientific and other formal writing, e.g. *tamma ʾijrāʾu mubāḥaṯātin bayna r-raʾīsayn* 'discussions were held between the two presidents'.

4.3.7 Transitivity

The verb may be transitive, intransitive, or reflexive/middle in form (→ valency). In Form I, the stem vowel pattern /u-u/ is intransitive, and the patterns /a-i/ and /i-a/ tend to be transitive; Form VII is reflexive/middle and so intransitive; Forms V (reflexive) and VI (reciprocal) are essentially intransitive, and Form VIII when it has reflexive meaning is intransitive but transitive when the meaning is middle. Form IX, change of state, is exclusively intransitive. Forms II, III, IV, and X are largely transitive (see Table 18). Intransitive verbs may become transitive with preposition complements, e.g. *taraddada* 'to hesitate' and *taraddada ʿalā* 'to frequent [a place]'; most notable is the preposition *bi-*, which makes verbs of motion causative and transitive, as in *ʾatā bi-* 'to come with, bring' (see Sec. 4.3.8).

4.3.8 Semantic classes

Modern Standard Arabic verbs fall into four semantic groups on the basis of the presence or absence of progressive meaning in the imperfect

Table 18. Derived verbal Forms

Form	perfect	imperfect	example	meaning
I	*Faʿ aL-a*	*ya-Fʿ iL-u*	*ḍaraba/yaḍribu* ‘to hit’	action
	Faʿ aL-a	*ya-Fʿ uL-u*	*taraka/yatruku* ‘to leave’	
	Faʿ iL-a	*ya-Fʿ aL-u*	*fahima/yafhamu* ‘to understand’	(temporary) state or activity
	Faʿ uL-a	*ya-Fʿ uL-u*	*ḥasuna/yaḥsunu* ‘to be good’	permanent trait
			kabura/yakburu ‘to become big’	
II	*Faʿʿ aL-a*	*yu-Faʿʿ iL-u*	*fahhama/yufahhimu* ‘to cause to understand, instruct’	causative
III	*Fāʿ aL-a*	*yu-Fāʿ iL-u*	*kātaba/yukātibu* ‘to correspond with someone’	associative
IV	*ʾaFʿ aL-a*	*yu-Fʿ iL-u*	*ʾafhama/yufhimu* ‘to make someone understand, to explain something to someone’	causative
V	*taFaʿʿ aL-a*	*yataFaʿʿ aL-u*	*tafahhama/yatafahhamu* ‘to come to understand by degrees’	reflexive-causative: deliberate
VI	*taFāʿ aL-a*	*yataFāʿ aL-u*	*tafāhama/yatafāhamu* ‘to reach a mutual understanding’	reciprocal/associative
VII	*inFaʿ aL-a*	*yanFaʿ iL-u*	*inqaṭaʿ a/yanqaṭiʿ u* ‘to be cut off’	reflexive-passive
VIII	*iFtaʿ aL-a*	*yaFtaʿ iL-u*	*ijtamaʿ a/yajtamiʿ u* ‘to gather together, meet’	reflexive/middle
IX	*iFʿ aLL-a*	*yaFʿ aLL-u*	*iḥmarra/yaḥmarru* ‘to turn red’	inchoative: become (a color or a defect)
X	*istaFʿ aL-a*	*yastaFʿ iL-u*	*istafhama/yastafhimu* ‘to inquire’	causative-middle: requestative

indicative and on the meaning of the active participle; they are Statives, Activities, Acts, and Inchoatives (→ Aktionsart).

‘Stative’ verbs never have progressive meaning: the imperfect is essentially existential, denoting a state or condition with a particular semantic coloration, and never denotes an action; their active participles have only perfective meaning, i.e. ‘having entered the state of/having come to be…’, and are often translated as present tense verbs in English. Examples are *yaʿ rifu* ‘he knows’ and *ʿ ārif* ‘(having come to) know, having learned, knowing’; *ʾajāda* ‘to become good at, master’/*yujīdu* ‘he is good at’ and *mujīd* ‘having mastered, mastering, good at’.

There are two subclasses of statives, ‘qualitative’ verbs and → ‘impersonal verbs’. Qualitatives are existential statives that denote a quality or condition and for which the imperative is blocked; their participles are often noncanonical in shape. Examples are *ḥasuna* ‘to be good’/*yaḥsunu* ‘it is good’ and *ḥasan* ‘(having become) good’; *wasiʿ a* ‘to be wide, spacious’/*yasaʿ u* ‘it is wide, spacious’ and *wāsiʿ* ‘(having become) wide, spacious’.

Impersonals are qualitatives that take only clauses or verbal nouns as subject; they consequently occur only in the 3rd person singular. Examples are *wajaba ʿ alā* ‘to be necessary for’/*yajibu* ‘it is necessary’ and *wājib* ‘necessary’;

jadura bi- ‘to be worth (do)ing’/*yajduru bi-* ‘it is worth…-ing’ and *jadīr bi-* ‘worth…-ing’.

‘Activities’ are actions that take place over a period of time and so entail expressions of extent of time, such as ‘for (an hour)’; there is no result or achievement or change in action – the action is the same at every point in time from the beginning to the end of the activity. Thus, in ‘he waited for two hours’, there is no change in or result from the activity, but the activity was unvaried at every point in those two hours; we do not know whether he met the party in question. Arabic activity verbs have progressive meaning in both the imperfect indicative and the active participle, as in *šaġala* ‘to occupy’/*yašġalu* ‘he occupies, is occupying’ and *šāġil* ‘occupying’; *intaḏ̣ara* ‘to wait for’/*yantaḏ̣iru* ‘he waits, is waiting for’ and *muntaḏ̣ir* ‘waiting for’.

‘Acts’ are actions that entail punctual adverbial expressions like ‘in (an hour)’ and, unlike activities, may result in the successful realization of a goal, as in ‘he wrote the article in two hours’. Acts may have progressive meaning in the imperfect indicative, but the active participle has only future meaning with the element of positive intention, e.g., the active participle *kātib* means ‘going to write [without fail]’. Acts are of two types, (i) volitional, called ‘accomplishments’, like *kataba* ‘to write’/*yaktubu* ‘he writes, is writing’ and *kātib* ‘going to write, will write’; accomplishments are typically transitive; and (ii) intransitive nonvolitional change-of-state actions, called ‘achievements’, whose active participles have perfective meaning (‘having become’), such as *kabura* ‘he became big’/*yakburu* ‘he becomes, is becoming big’ and *kabīr* ‘(having become) big’; *taḥajjara* ‘he turned to stone’/*yataḥajjaru* ‘he turns, is turning to stone’ and *mutaḥajjir* ‘(having) turned to stone, ossified’. Note that as change-of-state qualitative verbs they may have noncanonical participles.

‘Inchoative’ verbs essentially denote the beginning of an activity or condition; they may be transitive or intransitive. They all lack progressive meaning in the imperfect indicative but are subdivided into three subsets on the basis of the meaning of the active participle:

i. ‘Developmental’ inchoative participles, like all other change-of-state participles, have perfective and not progressive meaning: *taʿiba* ‘to get tired’/*yatʿabu* ‘he gets tired’ and *taʿib, taʿbān* ‘tired’; *nasiya* ‘to forget’/*yansā* ‘he forgets’ and *nāsin* ‘having forgotten, forgetting’.

ii. ‘Inceptive’ verbs may denote either an activity or the commencement of that activity; they thus combine the features of both Activity verbs and Developmental verbs, e.g. *rakiba* ‘to ride; to mount, get on [horse, train, etc.]’/*yarkabu* ‘he rides; mounts, gets on’ and *rākib* ‘riding; mounted, on’. The participle has progressive meaning of the activity and perfective meaning of the inceptive act. Other examples are *nāma* ‘to sleep; to fall asleep’/*yanāmu* ‘he sleeps; falls asleep’ and *nāʾim* ‘sleeping; asleep’; *labisa* ‘to wear; to don, put on’/*yalbasu* ‘he wears; puts on’ and *lābis* ‘wearing; having put on, having on’.

iii. ‘Movement’ verbs are a subset of Inceptives that denote a change of location and whose participles may additionally denote predictive (‘future’) action; as with Inceptives the movement participle may have progressive meaning of the activity or perfective meaning of the state, but may further have predictive meaning of the activity, e.g. *sāfara* ‘to travel; to set out, depart’/*yusāfiru* ‘he travels; departs, leaves’ and *musāfir* ‘traveling; having departed, left; going to depart’; *ijtamaʿa* ‘to come together, assemble, meet’/*yajtamiʿu* ‘meets, is meeting’ and *mujtamiʿ* ‘meeting; having met, assembled; going to meet’. Included are zero members of this semantic class, verbs of staying: *bāqin* (*ʾilā*) ‘staying (until)’. These verb classes are summarized in Table 19.

‘Phrasal verbs’ consist of a verb plus a preposition which create a meaning different from that of the verb alone, e.g., *qāma* means ‘to rise, stand erect’; its meaning is altered as follows with the following prepositions governing the object: *qāma ʿalā* ‘to rise up against; to be based, founded on’; *qāma li-* ‘to rise in honor of’; *qāma ʾilā* ‘to go to’; *qāma bi-* ‘to undertake, begin, do’, as in *qāma bi-dirāsati s-sanskrītiyya* ‘he undertook the study of Sanskrit’, *qāmat bi-qatli ʾabīhi* ‘she killed his father’. Phrasal verbs also include verb plus noun, such as *ʾaṭlaqa sarāḥa* ‘to release [from imprisonment]’, consisting of *ʾaṭlaqa* ‘to unloosen’ and *sarāḥa* ‘release’ (→ collocations).

Table 19. Verbal semantic classes

Semantic class	Perfect tense: Event in narration	Imperfect tense: Attendant circumstance	Active participle: Resultant state: 'having become…'; 'having begun to…'
1. STATIVES	Event: Entered a state or condition	Existential: State, condition or quality	Perfective: 'having become…'
a. States	*ʿarafa* 'he learned'	*yaʿrif* 'he knows'	*ʿārif* 'knowing now'
b. Qualitatives	*ḥasuna* 'it became good'	*yaḥsunu* 'it is good'	*ḥasan* 'good'
c. Impersonals	*lazima* 'it became necessary'	*yalzamu* 'it is necessary'	*lāzim* 'necessary'
2. ACTIVITIES	*intaḏ̣ara* 'he waited'	*yantaḏ̣iru* 'he waits, is waiting'	*muntaḏ̣ir* 'waiting'
3. ACTS			
a. Accomplishments	*kataba* 'he wrote'	*yaktubu* 'he writes, is writing'	*kātib* 'going to write'
b. Achievements	*iḥmarra* 'he turned red'	*yaḥmarru* 'he turns red, is turning red'	*muḥmarr* 'reddened, reddish'
4. INCHOATIVES			
a. Developmental	*taʿiba* 'he got tired'	*yatʿabu* 'he gets tired'	*taʿib, taʿbān* 'tired'
b. Inceptive	*labisa* 'he wore; put on, donned'	*yalbasu* 'he wears; puts on'	*lābis* 'wearing; having on'
c. Movement	*sāfara* 'he traveled; he departed'	*yusāfiru* 'travels; departs'	*musāfir* 'traveling; departing; departed; going to depart/travel'

4.3.9 Participles

'Participles' are adjectives derived from verbs and retaining verbal force: they are active or passive in voice and transitive or intransitive; they denote a state resulting from a completed action (perfective) or one of having begun an action (progressive or future). Active participles from states, achievement, and developmental verbs (see Table 19) have perfective meaning; Activity participles have progressive meaning, and accomplishment verbs have [+volitional] future meaning; inceptive participles may have perfective or progressive meaning, and movement ones may have perfective, progressive, or future meaning. Passive participles denote 'having undergone the action of being…', e.g. *maktūb* '(having been) written', *ar-risāla al-maktūba* 'the letter that has been written, the written letter'. The timing of participles is present time, that is, co-occurring with the time of the matrix clause, e.g. *ʾa-lā tarā-hu rākiban al-ḥiṣān?* 'don't you see him riding/mounted on the horse?' and *ʾa-lam tara-hu wa-huwa rākibun al-ḥiṣān*? 'didn't you see him as he *was* riding/mounted on the horse?'. See also Section 4.3, Verbs. All participles take sound plurals: *lābis/lābisūn* and *lābisa/lābisāt* 'having donned; wearing'.

Participles often become nouns, in which case Form I participles may take broken plurals: *kātib/kātibūn* 'have written', but *kātib/kuttāb* 'writers'.

The passive participle denotes patient, e.g. *madrūs* '(having been) studied', *muḥaddad* 'limited'. Passive participles may have potential meaning, e.g. *maqbūl* 'accepted; acceptable', *masʾūl* '(having been) asked; responsible ['can be questioned']'. They may also be used as nouns of place, e.g. *muxayyam* 'camp, encampment', *mustaqbal* 'future'.

There are active and passive patterns specific to Form I and to the derived Forms. The active participle of Form I pattern is *FāʿiL*, and the passive pattern is *maFʿūL*: *kātib* and *maktūb*. There are also noncanonical Form I active participles, no doubt reflecting original parti-

Table 20. The derived participles

imperfect stem	active participle	passive participle	illustration
II -*Faʿʿ iL*-	*muFaʿʿ iL*	*muFaʿʿ aL*	*muḥaddid* 'limiting' *muḥaddad* 'limited'
III -*Fāʿ iL*-	*muFāʿ iL*	*muFāʿ aL*	*musāfir* 'traveling; traveler' *mubārak* 'blessed; fortunate'
IV -*Fʿ iL*-	*muFʿ iL*	*muFʿ aL*	*mursil* 'sending' *mursal* 'sent'
V -*taFaʿʿ aL*-	*mutaFaʿʿ iL*	*mutaFaʿʿ aL*	*mutakallim* 'speaking; spokesman' *mutaḥaqqaq* 'ascertained, certain'
VI -*taFāʿ aL*-	*mutaFāʿ iL*	*mutaFāʿ aL*	*mutaqābil* 'exchanging (with each other)' *mutaqābal* 'exchanged; reciprocal'
VII -*nFaʿ iL*-	*munFaʿ iL*	[*munFaʿ aL*]	*munsaḥib* 'withdrawing'
VIII -*Ftaʿ iL*-	*muFtaʿ iL*	*muFtaʿ aL*	*muštarik* 'participant' *muštarak* 'common, collective'
IX -*Fʿ aLL*-	*muFʿ aLL*	–	*muḥmarr* 'reddened, reddish'
X -*staFʿ iL*-	*mustaFʿ iL*	*mustaFʿ aL*	*mustaʿmil* 'using' *mustaʿmal* 'used'
QI -*FaʿLiL*-	*muFaʿLiL*	*muFaʿLaL*	*mutarjim* 'translator' *mutarjam* 'translated'

cipial forms: *FaʿL, FaʿaL, FaʿiL, FaʿīL*, e.g. *ṣaʿb* 'difficult', *ḥasan* 'good', *rajil* 'going on foot, walking', *kabīr* 'big'. For derived verbs the prefix *mu*- is added to the imperfect verb stem, and the stem vowel is changed to *i* for active participles and to *a* for passive participles (see Table 20). The stem vowel in Form IX is deleted by phonological rule (see Sec. 2.5).

The passive participle of Form VII does not occur as a participle, but, just like the other derived participles, it does occur as a noun of place, as in *munxafaḍ* 'low ground, depression'.

4.3.10 Verbal nouns

'Verbal nouns' are nouns (gerunds) derived from verbs on the basis of semantics (Form I) or by phonological rule (derived Forms); they name the denotation of the verb and retain some verbal force; they may be active or passive in voice. They are often lexicalized as concrete nouns resulting from the action of the verb, like *bināʾ* 'building [gerund]' and 'a building', e.g. *duhišnā min qatli xalīlin* 'we were astonished at the killing of Khalil/at Khalil's murder' and *duhišnā min qatli xalīlin jawādan* 'we were astonished at Khalil's killing of Jawad' (→ *maṣdar*; → verbal noun).

Form I verbal nouns show a high degree of semantic correlation with the verb. For example, transitive verbs with perfect stem vowel *a* or *i* tend to take the form *FaʿL*, e.g. *dars* 'studying'; intransitive verbs with perfect stem vowel *a* tend to take the pattern *FuʿūL* (*jalasa* 'to sit' – *julūs; waṣala* 'to arrive' – *wuṣūl* 'arrival') as do passives (*wujūd* 'to be found; existence'); hollow verbs of movement gravitate to *FaʿaLān* (*dawarān* 'turning, rotation' and *ṭayarān* 'flying; aviation'). See Wright (1951:I, 110) for a detailed listing of verbal nouns.

The basic rule for derived verbal nouns is to insert long *ā* before the final radical and to

Table 21. Verbal noun formation

Form	perfect stem	verbal noun	illustration
I	*Faʿ VL-*	*Faʿ L, Faʿ aL, etc.*	*fahm* 'understanding, comprehension'
II	*Faʿʿ aL-*	*taFʿīL, tiFʿāL, taFʿāL*	*taqsīm* 'division; partition' *tasmiya* 'naming'
III	*Fāʿ aLa-*	*muFāʿ aLa, Fiʿāl*	*mudāfaʿa, difāʿ* 'defense'
IV	*ʾaFʿaL-*	*ʾiFʿāL*	*ʾixrāj* 'extracting, extraction'
V	*taFaʿʿaL-*	*taFaʿʿuL*	*taʿajjub* 'amazement'
VI	*taFāʿaL-*	*taFāʿuL*	*tašāwur* 'joint consultation'
VII	*inFaʿaL-*	*inFiʿāL*	*insiḥāb* 'withdrawing, withdrawal'
VIII	*iFtaʿaL-*	*iFtiʿāL*	*intixāb* 'electing, election'
IX	*iFʿaLL-*	*iFʿiLāL*	*iḥmirār* 'reddening'
X	*istaFʿaL-*	*istiFʿāL*	*istiʿmāl* 'use'
QI	*FaʿLaL-*	*FaʿLaLa*	*handasa* 'engineering'
QII	*taFaʿLaL-*	*taFaʿLuL*	*tafalsuf* 'acting like a philosopher [*faylasūf*], philosophizing'

change all the preceding vowels to *i*. Exceptions are those of Form II, which fit the major pattern *taFʿīL* (*ʿallama* 'to teach' – *taʿlīm* 'teaching, instruction') and the intensive canonical pattern *tiFʿāL* or *taFʿāL* (*ḏakkara* 'to remind' – *tiḏkār* 'remembrance') (weak verbs have the special pattern *taFʿiya*); Form III with the patterns *muFāʿaLa* (*dāfaʿa ʿan* 'to defend' – *mudāfaʿa* 'defending, defense') and the canonical *FiʿāL* (*difāʿ* 'defense'); and Forms V and VI which change the stem vowel to *u*: *taFaʿʿuL*, *taFāʿuL* (*takallum* 'speaking [noun]', *taqābul* 'exchange') (see Table 21).

4.4 *Particles*

'Particles' are uninflected words, subclassified on the basis of function into → adverbs, → prepositions, → conjunctions, nominal particles, verbal particles, and → interjections.

4.4.1 Adverbs

Primary adverbs are a closed class, such as *ʾayḍan* 'also', *faqaṭ* 'only'. Most adverbials are indefinite accusative nouns or adjectives, e.g. *ruwaydan* 'slowly, gently', *laylan* 'at night', *qarīban* 'soon'; or accusative pairs without nunation, e.g. *layla nahāra* 'day and night'; or prepositional phrases, e.g. *ka-ḏālika* 'like that, thus, also', *bi-surʿatin* 'quickly', *ʿan qarībin* 'soon'. There is also a closed class of adverbs in *-u* of nominal origin complementing prepositions in *-a/i*, such as *baʿdu* 'afterward; yet, still', *fawqu* 'up, above', *qablu* 'earlier, previously'. The particles *-ʾiḏin* and *-ḏāka* 'at the time of' may be suffixed to accusative time nouns to form time adverbials, e.g. *waqta-ʾiḏin* and *ḥīna-ḏāka* 'at that time'. Expressions like *bi-šaklin*... and *bi-ṣūratin*... plus adjective are also a fertile source of adverbials, e.g. *bi-ṣūratin miṯāliyya* 'ideally'.

The 'cognate accusative' (→ object, absolute) is frequently used to strengthen, clarify, or modify the meaning of the verb; it is the accusative case form of the (cognate) verbal noun of the verb in question, as in English *I walloped him a walloping*. The cognate accusative may be unmodified, e.g. *šaraytu-hā širāʾan* 'I bought it [i.e., it wasn't a gift]', or modified, especially with a noun of quantity, e.g. *sāʿadū-nā musāʿadatan kabīran/kulla l-musāʿada* 'they helped us greatly/in every way'; *sakata sukūta l-mayyiti* 'he fell silent as a corpse'; *ʿāmala-hu muʿāmalata l-wālidi li-bni-hi* 'he treated him like a son [i.e. a father's treatment of his son]'.

The particle *ʾinna* 'indeed, verily' is a sentence-initial adverb that is followed by an accusative subject (→ *ʾinna wa-ʾaxawātuhā*).

4.4.2 Prepositions

Prepositions show the semantic or grammatical relationship between their objects and other portions of the sentence. Primary prepositions are a closed class, like *ʿalā* ‘on’, *ʾilā* ‘to’, *bi-* ‘in; by, with’, *fī* ‘in’, *li-* ‘for, to’, *ka-* ‘like, as’, *maʿa* ‘with’, *min* ‘from’. There are many noun-prepositions occurring in the accusative case, or genitive case when serving in a genitive function, e.g. *ʾamāma-ka* ‘in front of you’, but *min ʾamāmi-ka* ‘from in front of you’ (→ prepositions).

4.4.3 Conjunctions

Conjunctions connect words, phrases, clauses, and sentences. On the clause level they may be coordinating (→ parataxis) or subordinating (→ subordination). The conjunctions *wa-* ‘and’ and *fa-* ‘and, and then, and so’ occur on all of these levels. Of these two, *wa-* is a pure coordinator; *fa-* connotes subsequence or consequence, and in a dialogue may denote change of subject. Other coordinating connectors are *ṯumma* ‘then, and then’, *bal* ‘nay, indeed’, *wa-lākinna* ‘but’. Subordinating conjunctions are *ʾan* ‘[the concept] that’ (followed by subjunctive verb); *ʾanna* ‘[the fact] that’ (taking an indicative verb); *fa-* ‘in order that’, *li-(kay)/li-ʾan* ‘in order that’. Combinations exist, such as *bi-mā ʾanna* ‘inasmuch as, since’ (→ conjunctions; → causal clause; → concessive clause; → connectives).

4.4.4 Interjections

Interjections are words or phrases that express emotions; they do not enter into the syntactic structure of the sentence, but may serve as an entire utterance. Illustrations are *ʾajal* ‘yes’, *lā* ‘no’, *wā-…āh* ‘oh, what a…’ as in *wā-ʾasaf-āh!* ‘oh, what grief!’, *wā-farḥat-āh!* ‘what joy!’, and *wayla-ka!* ‘woe is you!’ (→ interjections).

4.4.5 Nominal particles

Nominal particles occur only with nouns; they include the vocative particles *yā* and *-āh* ‘o’, as in *yā rabb-ī* and *rabb-āh* ‘o, (my) Lord!’, and the annunciative particle *ʾidā (bi-)* ‘and then, all of a sudden there is …’, e.g. *ʾiḏā bi-ṣawtin yaqūlu* ‘all of a sudden there was a voice saying…’. The particle *wa-* with following genitive introduces an oath (*wa-ḷḷāhi* ‘by God!’) or may indicate quantity, ‘many a…’ (*wa-laylatin* ‘many a night’).

4.4.6 Verbal particles

Verbal particles include *qad* and *la-*. The circumstantial particle *qad* with a perfect verb denotes the resultant state of an event (present perfect); e.g. *qad ʿalimtu ʾanna-hu fī l-qāhira* ‘I have learned/I know now that he is in Cairo’. With the imperfect it denotes possibility: *qad yakūnu l-ʾāna fī l-qāhira* ‘he might be in Cairo now’. The → asseverative particle *la-* emphasizes the force of the verb: *la-qad ʿādat!* ‘she has returned!, she’s back!’.

4.4.7 Negation

The negative particle *lā* is the unmarked negative, serving to negate most parts of speech and structures (→ negation).

With indefinite accusative nouns, *lā* negates the class of items denoted by the noun, e.g. *lā ʾaḥad-a* ‘no one’, *lā šayʾ-a* ‘nothing’, *lā mafarr-a min-hu* ‘unavoidable, inevitable [lit. ‘there is no escape from it’]’, *lā nāqat-a lī fī l-ʾamr* ‘I have nothing to do with this [lit. “I have no camel in the matter”]’.

The noun *ġayr* followed by a genitive noun or adjective means ‘non-, in-’, e.g. *al-ʿarab-u wa-ġayr-u l-ʿarab-i* ‘Arabs and non-Arabs’, *ġayru ṣaḥīḥ-in* ‘incorrect’. The noun *ʿadam* ‘nothingness’ also negates a following genitive noun, e.g. *ʿadam at-tadaxxuli* ‘noninterference’, *ʿadam aṯ-ṯiqa* ‘[vote of] lack of confidence’.

With verbs the perfect tense is typically negated by *lam* ‘did not’/*lammā* ‘has not’ + jussive (*māta* ‘he died’ – *lam yamut* ‘he did not die’ – *lammā yamut* ‘he hasn’t died yet’). The auxiliary verbs → *kāna wa-ʾaxawātu-hā* are negated with *lam* + jussive or *mā* + perfect (*lam nakun hunāka* = *mā kunnā hunāka* ‘we weren’t there’, *lam yazal fī l-bayt* ‘he is still in the house’). The perfect tense negated with *lā* has optative force, e.g. *lā sāmaḥa-ka ḷ-ḷāhu yā najīb!* ‘may God not forgive you, Najib!’.

The imperfect indicative is negated by *lā*, occasionally by *mā* (*lā yafʿalu šayʾan* ‘he’s not doing anything’). Stative verbs may be negated by the verb *laysa* ‘not be’, producing a more literary style (*lastu ʾadrī* ‘I don’t know’ = *lā ʾaʿrifu*). Future tense (predictions) with *sawfa/sa-* are negated by *sawfa lā*, e.g. *sawfa lā ʾusāfiru ʾilā tūnisa ġadan* ‘I will not depart for Tunis tomorrow’.

The subjunctive with *lan* ‘will never, will not’ traditionally supplies a strong denial of future tense, but in contemporary usage it also serves as a negative of *sawfa/sa-*, e.g. *lan ʾaqūl-a la-ka māḏā faʿalnāhu* 'I will never tell you what we did'.

The 2nd person jussive with *lā* gives a negative command, e.g. *lā tansā-nā!* 'don't forget us!'. In the 1st and 3rd persons it produces an indirect command, e.g. *fa-lā naḏhab ʾilā l-maqhā l-yawma* 'let's not go to the coffee shop today'; *lā yadxul waḥda-hu* 'let him not enter alone'.

Correlative *lā…wa-lā* 'neither…nor' occurs with single word and with clauses, e.g. *lā hāḏā wa-lā ḏāka* 'neither this nor that'; *lā talfana wa-lā kataba* 'he neither phoned nor wrote'.

The defective verb *laysa* also negates phrases, e.g. *laysa bi-ṭ-ṭāʾira bal bi-s-sayyāra* 'not by plane but by car'.

Equative sentences are negated by the verb *laysa* 'not be' and occasionally by *mā*; a predicate nominal may be governed by *bi-*, e.g. *laysa l-mudīr bi-xayr* 'the director is not well'; *lastu qaliqan ʿalayhim* 'I am not worried about them'; *mā huwa bi-kāḏib-in* 'he is not lying'.

Clauses are negated by negating the verb.

5. SYNTAX

5.1 *Agreement*

Noun modifiers agree with human nouns in number, case, and gender, e.g. *maʿa ṭālibātin qaṭariyyātin* 'with Qatari students [fem. pl.]' (→ agreement). Nonhuman plurals, however, are syntactically feminine singular, as in (7):

(7) *hal ʾakalat hāḏihi l-kilābu*
Q eat.Perf.3fs this.fs the-dogs
l-ʾaʿḏuma kulla-hā?
the-bones all-her
'Did these dogs eat all of the bones?'

There is full agreement between a verb and a preceding subject, but number agreement is blocked if the subject follows, as in (8).

(8) *ʾatā l-ʾawlādu*
come.Perf.3ms the-boys
wa-laʿibū kurata l-qadam
and-play.Perf.3mp ball the-foot
'The boys came and played soccer'

For the semantic difference between the two word orders, → topic/comment and → theme/rheme.

5.2 *Nominal phrases*

Apposition is the juxtaposition of two nouns with the same referent to serve as a single syntactic unit, sharing the same case, e.g. *al-ʾānisatu hindun* 'Miss Hind'; *ʾaxūka salīmun* 'your brother Salim'.

Demonstrative phrases contain a demonstrative pronoun joined to a following noun by the definite article; they are analyzed as a subclass of appositives, the demonstrative agreeing with the noun in number and gender, while the noun agrees with the demonstrative in case, e.g. *ʿan hāḏihi l-masʾalati* 'concerning this question'. The demonstrative may alternatively, for stylistic reasons, follow the definite noun, as in *fī l-maqāli hāḏā* 'in this article'. If the noun is modified, the demonstrative must follow it, e.g. *al-ʿālamu l-jadīdu hāḏā* 'this new world', *fī madīnati-nā hāḏihi* 'in this city of ours'.

Noun Phrases consist of a noun head plus modifiers, which may be any other phrase. The 'genitive → construct' (→ *ʾiḍāfa*) consists of a noun in an appropriate case ('the first term') and a genitive noun ('the second term'). It may denote any of the following:

i. Possession: possessor – possessed, e.g. *sayyāratu z-zāʾiri* 'the car of the visitor, the visitor's car'
ii. Naming: feature – name, e.g. *madīnatu l-kuwayti* 'the city of Kuwait, Kuwait City'
iii. Composition: concrete noun – material, e.g. *finjānu faxxārin* 'an earthenware cup'
iv. Limitation: class – item, e.g. *finjānu qahwatin* 'a coffee cup' (cf. *finjānu qahwatin min al-māʾi* 'a coffee-cupful of water')
v. Content: container – contents, e.g. *finjānu qahwatin* 'a cup of coffee' = *finjānun min al-qahwati*

The first term never takes definite article or nunation but is definite by construction, e.g. *sāʿatu l-marʾati* 'the woman's watch' and *sāʿatu mraʾatin* 'a woman's watch'.

The following are specialized genitive constructs:

i. → Quantifier phrases have as first term a noun of quantity, such as *kull, jamīʿ* ‘all’; *ʾaġlab, ʾakṯar* ‘most’; *baʿḍ* ‘some’, e.g. *kullu n-nāsi* ‘all of the people’, *baʿḍu l-ʾaḥyāni* ‘some times’. Numerals 3 through 10 form a subclass of quantifier phrases, as in *xamsatu ʾayyāmin* ‘five days’ (→ numerals). The first term may also be an interrogative noun (*ʾayyu ṣafḥatin?* ‘what page?’), elative adjective (*ʾaḏkā ṭullābin* ‘the most intelligent students’; → elative), and ordinal numeral (*li-ʾawwali marratin* ‘for the first time’).

ii. Verbal noun phrases. If the first term is a verbal noun, the following constructions are possible:
State: existential, e.g. *wujūdu l-hāribi* ‘the whereabouts of the fugitive’
Action: agent or patient, e.g. *qatlu s-sāriqi* ‘the thief’s murdering (of someone); the thief’s murder’
Action: agent (gen.) + patient (acc.), e.g. *qatlu s-sāriq*-i *ṣāḥib*-a *l-maḥalli* ‘the thief’s murdering (of) the shopkeeper’ (Qafisheh 1968)

Adjectives may be modified by prepositional phrases, as in *mašhūr bi-ʾamwāli-hi* ‘famous for his wealth’. They may also be modified by a defined genitive noun, as in *midfaʿ baʿīd al-madā* ‘a long-range gun [lit. ‘gun long of range’]’. Attributively, it agrees with the noun in definiteness, e.g. *ar-rajul al-qāsī l-qalb-i* ‘the hard-hearted man’.

Participles have the force of verbs without expressing aspect, time, or agent but do have other verbal behaviors, e.g. *wāʿidan an-nāsa bi- l-ʿawda* ‘promising the people to return’ (participle with two objects); *qādimīna min dār bayḍāʾ* coming from Casablanca’ (intransitive participle; → participle).

5.3 *Sentences*

There are two basic types of sentence, the verbless sentence (→ nominal sentence; → locatives) and those with verbs (→ verbal sentence).

5.3.1 Verbless sentences

Verbless sentences contain a subject and a predicate nominative adjective or noun, or a prepositional phrase or clause; the predicate modifies or identifies the subject. Nominalized clauses may serve as subject. The unmarked word order is Subject Predicate (SP), but the order is reversed if the subject is indefinite or a clause. If the predicate begins with the definite article *al-*, a ‘pronoun of separation’ (a 3rd person pronoun typically in agreement with the subject) is inserted between the subject and predicate. The predicate may be an adjective, as in (9).

(9) *hāḏihi l-fikratu mumtāza*
this the-idea excellent
‘That’s an excellent idea!’

or a noun, as in (10):

(10) *al-ʾustāḏu ʿabdu l-qādir*
the-professor ʿAbd al-Qādir
mudarrisu-nā l-jadīd
instructor-our the-new
‘Professor ʿAbd al-Qādir is our new instructor’

or a prepositional phrase, as in (11):

(11) *yūsuf ḥāliyyan fī maktabi-hi*
Yusuf at.present in office-his
‘Yusuf is in his office at present’

or a noun with definite article, as in (12):

(12) *tawfīq huwa l-ḥakīm*
Tawfiq he the-doctor
‘Tawfiq is the doctor’

or a clause, as in (13):

(13) *ʾaxū-ki huwa l-laḏī*
brother-your he who
laʿiba dawr-a l-malik-i
play.Perf.3ms role.Acc the-king-Gen
fī t-tamṯīliyya
in the-play
‘Your brother is the one who played the role of the king in the play’

Clauses may also serve as subject, as in (14):

(14) *min l-mafrūḍ ʾan*
from the-assumed that
yazūrū-nā ṯāniyatan
visit.Subj.3mp-us again
‘It is assumed that they will visit us again’

In a sentence with an indefinite subject, the word order is Predicate Subject, as in (15):

(15) *ʿind-ī suʾālun*
with-me question.Nom.Indef
'I have a question'

Existential sentences are exemplified by the previous sentence, which contains an indefinite subject and a preposed predicate. Existence is also expressed by the passive verb *wujida/yūjadu* 'to be found', e.g. *hal yūjadu lawḥ-un fī l-ġurfa?* 'is there a blackboard in the room?'. Existence in past time is expressed with the insertion of the perfect tense of *kāna*, e.g. *kāna yawma-ḏāka luṣūṣun kaṯīrūn* 'there were at that time many thieves'.

5.3.2 Verbal sentences

In Verb Phrases, 'modal' or 'auxiliary' verbs may modify the meaning of the main verb in the clause, normally in the indicative mood; in addition to *kāna*, which switches the time frame of the main verb, there are auxiliary verbs that denote the commencing of an action or state, such as *badaʾa/yabdaʾu* 'to begin' and *ʾaxaḏa/yaʾxuḏu* 'to take', e.g. *badaʾū/ʾaxaḏū yadrusūna baʿda l-imtiḥān* 'they began to study after the exam'; auxiliary verbs of continuing, such as *istamarra* 'to continue', *ḍalla/yaḍallu* 'to remain' and *mā zāla/mā yazālu* 'not cease to', e.g. *lam tazal tatakallamu ʿan-ka* 'she still talks about you'; and miscellaneous verbs best translated as adverbs, such as *kāda/yakādu* 'almost', e.g. *kādat taqaʿu* 'she almost fell', *kidtu ʾuqtalu* 'I was almost killed' (→ auxiliary verbs; → modal).

The unmarked → word order in the verbal sentence is Verb Subject Object (VSO), with optional modifiers at any position. The verb preceding its subject is always singular, but otherwise agrees in person and gender (→ agreement; → verbal sentence), as in (16) and (17).

(16) *šahida r-raʾīsu*
witness.Perf.3ms the-president
l-munāwarāt-i l-baḥriyya
the-maneuvers-Acc the-naval
'The president witnessed the naval maneuvers'

(17) *ḥaḍarat al-mumarriḍātu ʾilā*
come.Perf.3fs the-nurses to
l-mustašfā fawran
the-hospital immediately
'The nurses came to the hospital immediately'

The word order SVO, however, is equally common with the imperfect tense, whose function is to describe the background situation rather than events, as in (18).

(18) *al-bayānu yuʾakkidu ḥaqq-a*
the-statement affirms right-Acc
š-šaʿb-i fī ntiqād-i
the-people-Acc in criticizing-Gen
l-ḥukūma
the-government
'The statement affirms the right of the people to criticize the government'

5.3.3 Focus

→ Focus extraposition of the topic occurs typically with *ʾammā…fa-* 'as for…then', as in (19) and (20):

(19) *ʾammā z-zuwwāru fa-lam*
as.for the-visitors *fa*-not
yašhadū-hā
witness.Juss.3mp-her
'As for the visitors, they did not witness them'

(20) *wa-l-jarḥā najaw*
and-the-wounded be.saved.Perf.3mp
kulla-hum
all-them
'(As for) the wounded, they were all saved'

The *ʾinna* group (→ *ʾinna wa-ʾaxawātuhā*) introduces clauses and is followed by an accusative subject, whether noun or pronoun, as in (21) and (22).

(21) *ʾinna l-qiṭār-a qad*
indeed the-train-Acc *qad*
ġādarat
depart.Perf.3fs
'The train has left'

(22) *naʿrifu ʾanna-hu*
know.Imperf.1p that-him
taxarraja min jāmiʿati-i
graduate.Perf.3ms from university-Gen
l-qāhira
Cairo
'We know that he graduated from Cairo University'

ʾinna does not trigger subject movement in verbless sentences, as in (23):

(23) *ʾinna li-d-dayri ḥuqūq-an*
verily for-the-monastery rights-Acc
muqaddasat-an
sacred-Acc
'The monastery does have sacred rights'

5.3.4 Relative clauses

The relative clause is a syntactically complete clause related to a definite antecedent by an agreeing form of the → relative pronoun *al-laḏī* 'who' (→ relative clause); if the antecedent is indefinite, the relative pronoun, which is definite, is blocked (→ *ṣifa*). In either case, there must be an agreement feature in the relative clause with the antecedent (→ *ʿāʾid*). The agreement features in (24) – (26) are in bold and italics.

(24) *hāḏā huwa l-muṭribu l-laḏī*
this he the-singer who
*takallamnā ʿan-**hu***
talk.Perf.1p about-*him*
'This is the singer that we talked about'

(25) *hāḏā huwa muṭribun takallamnā*
this he singer talk.Perf.1p
*ʿan-**hu***
about-*him*
'This is a singer that we talked about'

(26) *hāḏā muṭribun darasa fī*
this singer study.Perf.3ms in
ʾurubbā
Europe
'This is a singer who *studied* in Europe'

When there is no antecedent, *al-laḏī* or *man* 'whoever' serves as the subject of two consecutive clauses, as in (27) and (28).

(27) *al-laḏī qāla ḏālika*
who say.Perf.3ms that
ṣadaqa
speak.the.truth.Perf.3ms
'Whoever said that spoke the truth'

(28) *man jadda wajada*
who strive.Perf.3**ms** find.Perf.3**ms**
'He who tries hard will succeed' (proverb)

5.3.5 Circumstantial clause

The 'circumstantial clause' (*jumla ḥāliyya*; → *ḥāl*) modifies a noun describing an attendant circumstance (*ḥāl*) in which the antecedent finds itself. The basic structure of the clause is (i) *wa-* 'as, while' + a verbless sentence; or (ii) *wa-* 'as, while' + an independent pronoun referring to the antecedent + a verbal sentence; in every instance there is ideally a pronominal reference to the antecedent. An example of the former is *intaḏara qalīlan wa-zawjatu-hu fī l-maṭbax* 'he waited a while, while his wife was in the kitchen'; of the latter *ʿāda l-kāhin wa-huwa yurattilu bi-luġa qadīma ġarība* 'the priest returned as he was chanting in a strange ancient language'.

A perfect tense in circumstantial clauses is preceded by *wa-qad*, e.g. *ʿāda wa- qad bāraka ṭ-ṭifl* 'he returned having blessed the child'.

The *wa-* + pronoun is often omitted before an imperfect tense verb, e.g. *ʿāda yurattilu* 'he came back chanting', and the imperfect may be changed to an accusative indefinite participle, e.g. *ʿāda murattilan* 'he came back chanting'.

5.3.6 Conditional sentences

Conditional sentences consist of a condition clause (protasis) and a result clause (apodosis; → *jazāʾ*). There are three conditional particles, which indicate the degree of probability of the realization of the condition: *ʾin* 'if, if it should be that…' (pure hypothesis); *ʾiḏā* 'if, when' and *ʾiḏā mā* 'whenever'; and *law….la-* 'if it were the case that…' (contrary-to-fact condition). The verbs in both clauses have perfect tense but refer to present time; a form of *kāna* plus *qad* preceding the main verb switches it to a past-time context, and a form of *yakūnu* + *qad* switches it to future perfect. The jussive mood may replace the perfect tense in conditions with *ʾin*, e.g. *ʾin ḏahabta ḏahabtu* 'if you go, I'll go' (perfect tense) and *ʾin taḏhab ʾaḏhab* ‘if you go, I’ll go’ (jussive mood); *ʾin ixtarta ʾanta l-qahwa ixtartu-hā ʾanā kaḏālika* ‘if you choose coffee, I will, too’; *ʾiḏā ḏahabta ḏahabtu* ‘if/when you go, I’ll go’; *ʾayna-mā ḏahabta ḏahabtu* ‘wherever you go, I’ll go’; *law ḏahabta la-ḏahabtu* ‘if you were to go, I would go’; *law kunta qad ḥaḍarta l-jalsa la-kuntu qad raʾaytu-ka* ‘if you had attended the session, I would have seen you’. If the result clause verb does not have perfect tense, *fa-* must introduce the clause, as in (29).

(29) *ʾiḏā ʾataw* *maʿa-nā* *fa-sawfa*
if come.Perf.3mp with us *fa*-will
nazūru *l-ʾahrām*
visit.Imperf.1p the.pyramids
‘If they come with us, we will visit the Pyramids’

BIBLIOGRAPHICAL REFERENCES

For Arabic linguistic bibliography prior to 1983, see:
Bakalla, Muḥammad Ḥasan. 1983. *Arabic linguistics: An introduction and bibliography*. London: Mansell.
For Modern Standard Arabic dictionaries, → lexicography: Monolingual dictionaries; → lexicography: Bilingual dictionaries.

REFERENCE GRAMMARS

Badawi, Elsaid, Michael G. Carter, and Adrian Gully. 2004. *Modern Written Arabic: A comprehensive grammar*. London: Routledge.
Holes, Clive. 2004. *Modern Arabic: Structures, functions, and varieties*. 2nd ed. Washington, D.C.: Georgetown University Press.
Ryding, Karin C. 2005. *A reference grammar of Modern Standard Arabic*. Cambridge: Cambridge University Press.
Wright, William. 1951. *A grammar of the Arabic language. Translated from the German of Caspari and edited with numerous additions and corrections*. Cambridge: Cambridge University Press.

SECONDARY SOURCES

Badawi, ElSaid M. 1973. *Mustawayāt al-ʿarabiyya al-muʿāṣira fī Miṣr: Baḥṯ fī ʿalāqat al-luġa bi-l-ḥaḍāra*. Cairo: Dar al-ʿIlm bi-Miṣr.
Brame, Michael K. 1970. *Arabic phonology: Implications for phonological theory and historical Semitic*. Ph.D. diss., Massachusetts Institute of Technology.
Ferguson, Charles A. 1956. “The emphatic ‘l’ in Arabic”. *Language* 32.446–452.
Khalil, Aziz and Ernest McCarus. 1999. “Arabic performative verbs”. *Zeitschrift für Arabische Linguistik* 36.7–20.
Qafisheh, Hamdi A. 1968. *English pre-nominal modifiers and corresponding Modern Standard Arabic structures: A contrastive analysis*. Ph.D. diss., University of Michigan.
Versteegh, Kees. 1997. *The Arabic language*. New York: Columbia University Press.

ERNEST N. MCCARUS (University of Michigan)

Modifier → X-bar Syntax

Mood (Arabic Dialects)

1. INTRODUCTION

Mood in spoken Arabic is a complex grammatical category. Standard Arabic grammarians did not have a term for it and therefore made no explicit reference to it (→ mood (Standard Arabic)). Mood in spoken Arabic is complex for the following reasons: (i) There is no universally agreed definition of the term; (ii) spoken Arabic, including → Educated (Spoken) Arabic, is characterized by variation, which, albeit systematic, is manifested in seemingly messy realizations of forms and structures in the different Arab countries; and (iii) mood cuts across the grammatical categories of tense, aspect, modality, and person. An additional factor is that spoken Arabic is not codified; there are no definitive grammars to date nor comprehensive and exhaustive lexicons of any of the various Arabic vernaculars. These factors and the lack of respect for spoken Arabic among the Arabs, particularly among Arab grammarians, kept the term ‘mood’ in Arabic linguistics outside the interest and focus of researchers.

2. DEFINITION OF MOOD

A workable definition of mood may be attempted here at the risk of arousing objections among some linguists and students of language. A good starting point is Lyons’ definition (1968:307):

> Mood, like tense, is frequently realized by inflecting the verb or by modifying it by means of ‘auxiliaries’. It is best defined in relation to an ‘unmarked’ class of sentences which express simple statements of fact, unqualified with respect to the attitude of the speaker towards what he is saying.

This definition involves both verb inflections and speakers' attitudes toward the propositions in their utterances. In several world languages, including Arabic, the differences between statement, commands, and (to a lesser extent) questions are grammaticalized in verb inflections.

Mood, therefore, cannot be restricted to the study of the morphological forms of verbs: its nature and significance can only be revealed in reference to the semantico-pragmatic meaning of utterances in context. The various verb forms, in other words, have to be syntactically and pragmatically contextualized, so that statements of meaning can be attributed to them (cf. Lyons 1977). After all, the boundary between morphology and syntax is often hard to delineate. Obviously, the morphological forms by themselves fall short of revealing the diversity of propositional, functional, and interpersonal meanings expressed in contextualized language.

An anticipatory example from Jordanian/ Palestinian spoken Arabic illustrates this point. Consider the *b-* non-past and the zero non-past as two modal verb forms: *b(y)ilʿab* 'he plays/is playing', *yilʿab* 'he plays'. To say that the *b-* non-past form *b(y)ilʿab* is in the indicative mood, whereas the zero non-past *yilʿab* is in the so-called subjunctive mood, does not bring out the meaningful contrasts between them, unless these forms are contextualized. For instance, in answer to *wēn ilwalad* 'where is the boy?', the *b-* non-past *bilʿab* 'he is playing' is acceptable, whereas the zero non-past **yilʿab* is not. However, *ilwalad biḥibb yilʿab* 'the boy likes to play', with the zero non-past subjunctive form is acceptable, but *ilwalad biḥibb *bilʿab* with the *b-* non-past is not.

Now, compare the above-mentioned forms in subordinate clauses: *ʾulnālu bilʿab* (*b-* non-past) 'we told him (that) he is playing' and *ʾulnālu yilʿab* (zero non-past) 'we told him to play'. In the former, the accusative pronominal form *-u* and the agent of *bilʿab* are not coreferential, whereas in the latter they are. It is such contextualized contrasts that make linguistic statements meaningful, and therefore worthwhile.

3. Modal forms and structures in spoken Arabic

Arabic, being the language of literate people, comprises a continuum of several varieties including → Classical Arabic, → Modern Standard Arabic, → Educated (Spoken) Arabic, and colloquial vernaculars. The study of mood in spoken Arabic must take cognizance of variation as a salient feature of this medium of communication. The imperative form of the root *k-t-b* 'to write', for instance, is realized variously as *ktōb* in Syria, *ʾiktib* in Egypt, and *ʾuktub* in Jordan/ Palestine. But this kind of variation hardly interferes with mutual intelligibility.

In this entry, the grammatical category of mood is seen as comprising three main types (cf. Lyons 1977; Cowell 1964): the indicative mood, the imperative mood, and the subjunctive mood.

The indicative mood is mainly used in declarative sentences to make statements of fact, but the same verb form may also be used to ask questions and express surprise, provided the intonation is appropriate. For example, the *b-* non-past in the following utterances expresses a statement and a question/surprise, respectively: *bikdib ʿalēna* 'he lies/is lying to us', *bikdib ʿalēna?* 'is he lying to us?'.

The imperative mood is a means of expressing mands (i.e. commands, requests, directives, etc.).

The subjunctive mood is appropriate for several semantico-pragmatic functions, including the speech acts of suggesting, exhorting, praying (or invoking the aid of God), and similar performative illocutions (cf. Lyons 1977).

Each of these moods is discussed and illustrated herein. But it should be pointed out here that the following analysis pertains primarily to the non-past tense in spoken Arabic, since it is the non-past that exhibits the various inflections and proclitics of mood. The past tense in spoken Arabic is not subject to modal inflections, but it is capable of expressing factive (i.e. realized) propositions (e.g. *miši* 'he left'), where it is understood that the action asserted was unequivocally realized prior to the moment of speaking, and non-factive propositions (e.g. *ʾaftikir innu miši* 'I think he left'), whereby the speakers make no firm commitment to the truth or falsity of the proposition. They are simply saying that it is not a fact that the action was realized; it is only a possibility.

A further modality which the past tense can convey is what Mitchell and El-Hassan (1994: 7–8) call contrafactivity. Consider the Jordanian/ Palestinian *ya retni ʾbilt naṣīḥtak* 'I wish I had accepted your advice' and the Egyptian *law kutti smiʿti kalāmi kutti ksibt* 'if you had listened to me, you would have won'. These utterances are modal means of expressing contrafactivity, i.e. the unequivocal nonrealization of the propositions involved. Finally, the past tense cannot convey mands (commands, requests, directives), i.e. cannot be used in the imperative mood, simply because mands are addressed to a second person to perform an act subsequent to the moment of utterance (not prior to it). In the following sections, the focus of discussion is on the non-past, which exhibits the typical moods and diverse modal distinctions in Arabic.

4. MOOD AND THE NON-PAST TENSE

4.1 *The indicative mood*

The indicative mood typically expresses assertions in declarative sentences. The non-past tense compatible with this type of mood is fairly constantly marked by the prefix *bi-* or *byi-*, as in:

Egyptian *yūsif da bifakkar ibnafsu w bass*
'This Yusif thinks of himself only'
ʾinta ssabab fi kull illi byiḥṣal
'You are the cause of all that is happening'
huwwa lwaḥīd illi ḥabbētu ʾu lissa baḥibbu
'He is the only one I (ever) loved and still love'
da guzʾ min ilmifṣal ill ana barakkibu
'This is part of the joint that I fix'
di ʾāxir ḥāga baʾulhālak
'This is the last thing I say to you'

Jordanian/Palestinian *hāda rrijjāl bikdib ʿalēna*
'This man is lying to us'
buktub maktūb
'He writes/is writing a letter'
bimšu fillēl
'They leave/are leaving at night'
btishar ʿindhum
'She spends/is spending the evening with them'
ʾana baxtalif maʿku fi rraʾy
'I disagree with you'

Syrian *byākōl*
'He is eating [in answer to *wēnu yūsif* 'where's Yousif?']
wi btaʿrif tirsim
'And she can [lit. 'knows how to'] paint'
bisāʿid ʾimmu
'He helps his mother'
btikwi lġasīl
'She irons the laundry'
bišūfhun kill žimʿa
'He sees them every week'

Lebanese *hayda ši byidʿu lattafāʾul*
'This is something that prompts optimism'
bass niḥna binwēžih mustaʿmir žālim
'But we face/are facing an oppressive imperialist'
ʾilʿālam bitḥakkam fīna
'The world mistreats/is mistreating us'
byiržaʿ ʿa baladu
'He returns to his country'
ʾana bḥiss ḥāli ġarīb
'I feel I am an alien'

The *b-* prefix of the non-past may be optional in certain contexts and regions (Mitchell and El-Hassan 1994:22–23); this phenomenon is quite common in Egyptian but not Levantine Arabic, e.g. *ʾilmilḥ (bi)ydūb fi lmayya* 'salt dissolves in water', *(bi)yiʿgibik ilfustān da* 'do you like this dress?'. Sometimes the *b-* prefix is optionally replaced by an *m-* prefix, e.g. (Lebanese) *minʾūl* 'we say', *mnaʿrif* 'we know', *mnikrah ilʾistiʿmār* 'we hate imperialism'; (Jordanian/Palestinian) *minḥibb inniqāš* 'we like discussion/debate'. As these examples show, this process is grammatically conditioned in that the *m-* prefix is followed by the alveolar nasal of the 1st person plural prefix. Thus, both *binʾūl* and *minʾūl* 'we say' are attested.

One more contrastive note concerning the *b-* prefix of the non-past. The spoken Arabic of the Arabian Gulf rarely exhibits the *b-* prefix. Consider *ša gūl* 'what should I say?/by the way', *tigdar tamši* 'you can go', *talʿab b ʾaʿṣābi* 'you are vexing me [lit. 'playing on/with my nerves']', *ʾana xdim baladi* 'I serve/am serving my country’. Similarly, the *b-* non-past is rare in Bedouin Arabic in Jordan, Palestine, Syria, etc. Compare *ʾana šhad* ‘I bear witness’ (Jordanian/Palestinian *bašhad*), *ʾint tafham* ‘do you understand?’, *(ʾi)ššēx yirīdak* ‘the chief [of the tribe] wants you’, *yisīr yawlidi* ‘is that possible, sonny?’, *ʾagūl ya ḥamdān* ‘by the way, Hamdan’ (cf. Egyptian *baʾullak ēh*).

If the *b-* non-past is preceded by the negative particle *mā* ‘not’, usage varies between maintaining the *b-* prefix (e.g. in Jordan, Palestine, Lebanon, and Syria) and dropping it (in Egypt). Consider (Egyptian) *maʾdarš* ‘I can’t’, *maʿrafš* ‘I don’t know’, *matitḥamalš* ‘it can’t be carried’; (Jordanian/Palestinian) *mā baʾdar(iš)* ‘I can’t’, *mā baʿrif(iš)* ‘I don’t know’, *mā btinḥamil/btinḥamliš* ‘it can’t be carried’; (Syrian) *mā btinḥamil* ‘it can’t be carried’, *mā bḥākī* ‘I don’t/won’t talk to him’. A similar contrast between retention and dropping of *b-* is found in Egyptian *ʾilbāb yifawwit gamal* ‘the door is wide enough [lit. ‘admits a camel for those who wish to quit’]’ and the Jordanian/Palestinian equivalent *ʾilbāb bifawwit jamal*. Similarly, Egyptian Arabic has *tibāt nār tiẓbaḥ ramād* ‘it starts the night as fire/amber and in the morning it turns to ash’, where Jordanian/Palestinian Arabic has *bitbāt jamra btiṣbiḥ ramād*.

Two other prefixes and proclitics associated with the non-past are *ḥa-* and *ʿam-*. The former is attested in various forms, depending on the region. In Egypt, *ḥa-* is fairly regular but may be attested as *ha-* (with a glottal fricative instead of the pharyngeal fricative). In the Levant, the morphemic variants of *ha-* include *rāḥ-*, *raḥ(a)-*, *rāyiḥ*, and, particularly in Damascus, *laḥ(a)-*. The prefix *ʿam-* also exhibits some morphemic variation across the region, e.g. *ʿamma, ʿammāl*. Now, *ḥa-* typically expresses the speaker’s proximate intention; it is volitional when prefixed to a 1st person non-past, e.g. *ḥarūḥ* ‘I will/am going to go’, *ḥaʾullu(h)* ‘I will/am going to tell him’, *ḥanzurkum* ‘we will/are going to visit you’. The volitional modality signaled by this proclitic also implies reference to the near future. For the distant future, *ḥa-* normally precedes the verbal form *nibʾa*, as in *ḥanibʾa nzurkum* ‘we will [sometime] visit you’. Note that the future, which traditionally was regarded as one of the tenses, is rarely of purest alloy – it is typically associated with some modality or other, e.g. promising, intention, volition, possibility, prediction, etc. (Lyons 1968:310). With the 3rd person, e.g. *ḥayīgi* ‘he will/shall come’, *ḥayiwfi bwaʿdu(h)* ‘he will/shall fulfill his promise’, the speaker acts as a guarantor to assure the addressee that a third person will come, fulfill his promise, etc. Thus, the Jordanian/Palestinian *rāḥ yīji labētak u yiʿtaḏirlak* ‘he will/shall come to your house and apologize to you’ may in fact be said by someone with institutionalized authority, e.g. a parent, tribal chief, district commissioner, cleric, school headmaster, mediator. In a Jordan TV panel discussion on 17 December 2003, one of the panelists, a cabinet minister, said *ʾilbarlamān miš rāḥ yinḥall* ‘Parliament won’t/isn’t going to be dissolved’. Within a second, a caller from the public phoned the presenter of the program taking the minister to task for saying something not within the jurisdiction of his authority as a minister. As the caller rightly said, only His Majesty the King has that authority.

The Egyptian *ḥa-* does not inflect for any of the categories of person, gender, and number. The Levantine *rāyiḥ*, in contrast, is inflected for all of these categories: *rāyiḥ* (1st/3rd person masc. sg.), *rāyḥa(h)* (1st/3rd person fem. sg.), *rāyḥīn* (1st/3rd person pl.). Examples:

Egyptian	*ḥanšūfak tāni* ‘We will/are going to see you again’ *ḥakallimik baʿdēn* ‘I will/am going to call you later’
Syrian	*ʾaddēš laḥa tġīb?* ‘How long will/are you going to stay away?’ *mīn raḥ inaẓẓif ilʾōda* ‘Who will/is going to clean the room?’ *daxlak kīf laḥatḥill halmiškli* ‘Incidentally/by the way how will/are you going to solve this problem?’
Lebanese	*ḥašūfak ʿala bukra(h)* ‘I will/am going to see you tomorrow/in the morning’
Jordanian/Palestinian	*rāḥ ʾakalimmak baʿdēn* ‘I will/am going to call you later’ *miš rāyḥīn iwāfʾu* ‘They won’t/aren’t going to agree’

The proclitic *ʿam-* precedes the non-past to mark durative, progressive, or habitual aspect depending on context, particularly when accompanying time specifiers. Examples:

Jordanian/ Palestinian	*ʾiṭṭifil ʿam bibki*
	‘The child is crying’
	ʿam bīthassan yōm baʿid yōm
	‘He is improving day by day’
Egyptian	*ʿammāl yiḥarrak fīna zayy ilʿarāyis*
	‘He is moving [manipulating] us like dolls’
Lebanese	*ʿam yistažiwbū(h)*
	‘They are questioning him’
	šū ʿam iʾūl?
	‘What is he saying?’
	ʿam izūr bayrūt kil žimʿa(h)
	‘He visits/is visiting Beirut every week’
Syrian	*šu ʿam taʿmēl?*
	‘What are you doing?’

Note that *ʿam(māl)* typically occurs before the zero non-past, except in Jordanian/Palestinian Arabic, where the *b-* non-past is much more common. Compare Egyptian *ʿammāl yiḥarrak fīna zayy ilʿarāyis* ‘he is moving [manipulating] us like dolls’ with Jordanian/Palestinian *ʿammālu biḥarrik fīna zayy iḥjār iššaṭranj* ‘he is moving [manipulating] us like chessmen’.

4.2 *The imperative mood*

The imperative mood in spoken Arabic is formed from the *b-* non-past of unaugmented triliteral roots by dropping the *b-* prefix and any person prefixes, and then prefixing the glottal stop and adding the appropriate person suffix, e.g. *b(y)iktib* ‘he writes’, *ʾiktib* ‘write [masc. sg.]!’; *b(y)idris* ‘he studies’, *ʾidris* ‘study [masc. sg.]!’; *b(y)iftaḥ* ‘he opens’, *ʾiftaḥ* ‘open [masc. sg.]!’; *bitiftaḥ* ‘she opens’, *ʾiftaḥi* ‘open [fem. sg.]!’.

In Syrian Arabic, the glottal stop is often deleted along with the following vowel, and the second vowel is lengthened. Consider the following imperative forms addressed to the 2nd person masculine singular: *ftāḥ, drōs, ktōb.* If the glottal stop is the first of the three consonantal radicals of the root, the imperative simply drops the *b-* prefix and person prefixes, e.g. *ʾ-x-ḏ* ‘to take’: *xud* ‘take!’; *ʾ-k-l* ‘to eat’: *kul/kōl* ‘eat!’. In Jordan and Palestine, however, the usual variant with the glottal stop is also attested, e.g. *ʾuxud, ʾukul.*

The imperative in spoken Arabic expresses mands addressed to a second person, but unlike European languages, e.g. English, spoken Arabic imperatives are inflected for gender and number. Consider the following Jordanian/ Palestinian examples: *ʾimsik* ‘catch/hold [masc. sg.]!’, *ʾimsiki/ʾimiski* ‘catch/hold [fem. sg.]!’, *ʾimsiku/ʾimisku* ‘catch/hold [masc. pl.]!’, *ʾimiskin* ‘catch/hold [fem. pl.]!’. Note that in Egyptian and Syrian spoken Arabic the masculine and feminine plural imperatives have the same form: *msiku, ktibu.*

Evidently, these morphological inflections are also accompanied by vowel alterations including phonological processes like → anaptyxis (i.e. inserting short close vowels to obviate the occurrence of consonantal clusters) and → elision of certain vowels.

Mands addressed to two people usually take the plural form, and only rarely do they exhibit the dual form. Out of context, therefore, such imperatives as *ʾirḥalu* ‘depart!’ are ambiguous – they are equally addressable to two or more people. In fact, the dual in spoken Arabic mood may, for all practical purposes of synchronic description, be regarded as nonexistent.

Geminate triliteral verbs (e.g. *biymurr* ‘he passes’, *biyʿidd* ‘he counts’) form the imperative by removing the *b-* prefix along with the person prefixes and observing the required vowel alterations, if any, e.g. *biymurr* ⇒ *murr* ‘pass [masc. sg.]!’; *bitmurr* ⇒ *murri* ‘pass [fem. sg.]!’; *bitmurru* ⇒ *murru* ‘pass [pl.]!’. The corresponding imperatives of *biyʿidd* are *ʿidd, ʿiddi, ʿiddu.*

The imperative of verbs derived from non-triliteral roots as well as from augmented verbs of triliteral roots is made by dropping the *b-* prefix of the non-past as well as person prefixes thereof. Consider *bidaḥrij* ‘he rolls’ ⇒ *daḥrij* ‘roll [masc. sg.]!’, *bitdaḥrij* ‘she rolls’ ⇒ *daḥriji/ daḥirji* ‘roll [fem. sg.]!’; *bi(y)sāfir* ‘he travels’ ⇒ *sāfir* ‘travel [masc. sg.]!’; *bi(y)sakkru* ‘they close’ ⇒ *sakkru* ‘close [pl.]!’.

In sum, the imperative mood in spoken Arabic is formed from the *b-* non-past by: (i) dropping the *b-* prefix and person prefixes, (ii) allowing for vowel alterations where necessary, and (iii) adding the appropriate 2nd person suffixes.

Verbs derived from unaugmented triliteral roots also require the prefixing of the glottal stop (often absent from the surface structure of spoken Syrian). Here are some miscellaneous examples:

Egyptian: *biysafru* ‘they travel’, *safru* ‘travel [pl.]!’; *biyāxud* ‘he takes’, *xud* ‘take [masc. sg.]!’; *biyʾūlu* ‘they say’, *ʾūlu* ‘say [pl.]!’; *biyibtidi* ‘he begins’, *(ʾi)btidi* ‘begin [masc. sg.]!’; *bititʿallam* ‘she learns’, *tʿallami* ‘learn [fem. sg.]!’; *bitġanni* ‘she sings’, *ġanni* ‘sing [fem. sg.]!’; *biyigri* ‘he runs’, *ʾigri* ‘run [masc. sg.]!’.

Syrian: *b(i)yiktob* ‘he writes’, *ktōb* ‘write [masc. sg.]!’; *bitimsik* ‘she holds’, *msiki* ‘hold [fem. sg.]!’; *bitʿallamu* ‘they learn’, *tʿallamu* ‘learn [pl.]!’; *birūḥ* ‘he goes’, *rūḥ* ‘go [masc. sg.]!’.

Lebanese: *bixālif* ‘he disagrees’, *xālif* ‘oppose/disagree [masc. sg.]!’; *btiržaʿ* ‘she returns’, *(ʾi)ržaʿi* ‘return [fem. sg.]!’.

Jordanian/Palestinian: *bimsik* ‘he holds’, *ʾimsik* ‘hold [masc. sg.]!’; *bitdūr* ‘she/it turns’, *dūri* ‘turn [fem. sg.]!’; *byāklu* ‘they eat’, *kulu* ‘eat [pl.]!’; *bitʿallamu* ‘they learn’, *(ʾi)tʿallamu* ‘learn [pl.]!’.

Note that the imperative of augmented *b*-non-pasts of triliteral roots as well as quadriliteral roots involving the infix *-t-* in such forms as *bitqaddam* ‘he advances’, *btitʿallam* ‘she learns’, *bitsallamu* ‘they receive’ form the imperative by maintaining the infix *-t-*: *(ʾi)tgaddam* ‘advance [masc. sg.]!’, *(ʾi)tʾallami* ‘learn [fem. sg.]!’, *(ʾi)tsallamu* ‘receive [pl.]!’. The *-t-* infix must be distinguished from the *t-* pro-form in *btitʿallam*/*bititʿallam*, which is deleted along with the *b*-prefix in the imperative.

Finally, prohibitions semantically belong to mands except that the addressee is required not to carry out the speech act in question: *lā tišrab* ‘don’t drink [masc. sg.]!’, *lā tʾūmi* ‘don’t rise [fem. sg.]!’, *lā titkallamu* ‘don’t talk [pl.]!’. Similar prohibitions are also realized as *mā* + verb + *š*, particularly in spoken Egyptian, e.g. *ma tišrabš* ‘don’t drink!’, *ma tʾūmīš* ‘don’t rise!’, *ma titkallamūš* ‘don’t talk!’. In rural areas of Jordan and Palestine, prohibitions may be realized as *lā* + verb + *š*, e.g. *lā tišrabiš, lā tʾūmiš, lā titkallamūš*. Occasionally, the negative particle *lā* or *mā*, especially in rapid retorts, is left out, e.g. *tišrabiš, tʾūmīš, titkallamūš*.

Note that, at least in the spoken Arabic of Jordan and Palestine, the imperative may be addressed to a second person in absentia, so to speak. The form is imperative, but the function is advice or directive, often when it is too late. Thus, in a situation where someone has made a mistake or made a decision seen by others (e.g. relations, friends) as being unwise or harmful, they might, while discussing the issue, use the imperative in reference to the absent addressee with a patronizing, often ironic or pejorative, vocative, e.g. *willa fakkir, yā ġabi lbaʿīd* ‘alternatively, think, oh distant fool!’. This imperative is equivalent to ‘you [in absentia] should have thought [more carefully]!’. Other examples include *willa sībhum ʾu lā tiddaxxal bšuġulhum* ‘better leave them and don’t interfere with their business!’, meaning that the absent addressee should have left them and not interfered with their business; *yāxi blaš bḥālak* ‘my dear, mind your own business!’, *yā maskīn rūḥ laššēx wiškīlu(h)* ‘poor fellow, go to the chief and complain to him!’. This type of imperative is (semantically and pragmatically) different from, say, the equivalent of English *let him go to hell.*

In expressions with *xallīh* ‘let him’, e.g. *xallīh yišrab ilbaḥr* ‘let him drink the sea!’, the imperative *xallīh* ‘let him’ is often deleted, e.g. (Egyptian) *yišrab ilbaḥr*, (Lebanese) *ykūn fi maktabu ssēʿa tlēti* ‘let him/he is to be in his office at three o’clock’, (Jordanian/Palestinian) *yitḥarraku badri* ‘let them/they are to depart early’. Whether *xallī* is overt or covert, the zero non-past is the vehicle for conveying such instructions or directives to a third party.

4.3 *The subjunctive mood*

The subjunctive mood in spoken Arabic takes the form of the zero non-past (i.e. non-past without the *b*- prefix). Examples include *tišrab šāy?* ‘would you like to drink some tea?’, *ʾaʾūl liḥaḍritik ʿala ḥāga witsaddaʾīni?* ‘shall I tell you something, provided you believe me?’, *ʾilbāb yifawwit gamal* ‘the door is wide enough to admit a camel [i.e. those who don’t agree can leave]’, *ʾarjūk ya sīdi* ‘I beg you, sir’, *rabbina yāxdik (yaġdik)* ‘may our Lord take you!’, *ʾalla yʿīnak* ‘may God help you!’, *ʾalla yiḥmīkun* ‘may God protect you!’, *ʾaftaḥlak iḍḍaw(w)?* ‘shall I turn on the light for you?’.

As these examples show, the subjunctive is the vehicle for expression of suggestions, invitations (*tišrab šāy*), implorations (*ʾarjūk ya sīdi*), prayers and invocations (*ʾalla yiḥmīkun*, *rabbina yaġdik*); for seeking instructions (*ʾaftaḥlak ilbāb*, *ʾaftaḥlak iḍḍaw(w)*); and for assertions with implicit illocutionary force (*ʾilbāb yifawwit gamal*). The last-mentioned category is expressed in Jordanian/Palestinian by the *b-* non-past, *ʾilbāb bifawwit jamal.* Such assertions tend to be fossilized idioms. Here is another example: (Egyptian) *tibāt nār tiṣbaḥ (tiẓbaḥ) ramād*, (Jordanian/Palestinian) *bitbāt jamra btiṣbiḥ ramād.*

The subjunctive in spoken Arabic is attested in main clauses and in embedded (i.e. subordinate) clauses. All the above examples contain subjunctives in main clauses. Here are some more examples in main clauses: *ʾaḥkīlak raʾyi bṣrāḥa(h)* 'shall I tell you my opinion frankly?', compare *baḥkīlak raʾyi bṣarāḥa(h)* 'I am telling/will tell you my opinion frankly'. Here, the *b-* non-past is retained in contrast with the subjunctive: the former (i.e. the *b-* non-past) being assertive, the latter (i.e. the subjunctive) entailing a suggestion, waiting for the addressee's response or instructions; *ʾastaġfiru llāh* 'I seek God's forgiveness'.

Other examples include:

Egyptian	*nāxud baʿḍina w nimši* 'Let's go! [lit. 'we take ourselves and go]' *tišrab ʿaṣīr burtuʾān?* 'Would you like orange juice?' *tiẓbaḥ ʿala xēr* 'Good night!'
Syrian	*ʾalla ynažžḥak* 'May God make you successful' *tiṣbiḥ ʿala xēr* 'Good night!'
Lebanese	*nisʾīkun šāy willa (wella) ʾahwi?* 'Shall we offer you tea or coffee?' *tinṭurni hawn?* 'Would you wait for me here?'

The subjunctive in embedded clauses is typically introduced by the complementizer *ʾinna* 'that' overtly stated or covertly understood: *bẓinn innha tiʾbal* 'I think that she will accept/agree'; *ʾana nṣaḥ inniha tiʿmil ilʿamaliyya(h)* 'I recommend that she undergo the operation'; *law ʾaʾdar asaʿdak ḥasaʿdak* 'if I can help you, I will'; *ṭalab min ṣaḥbu yiktim issirr* 'he asked his friend to keep the secret'; *wʿidīni nnik mā tikizbi* 'promise me that you won't lie!'.

Subordination of this kind includes clauses introduced by the purposive phrase *ʿala šān/ʿašān* 'for/in order to' and the temporal *(la) ḥatta* 'until', e.g. *ʾana jāy ʿašān ʾaxabbirku* 'I've come to let you know'; *kōl mnīḥ ʿašān tiʾwa* 'eat well in order to be strong!'; *ʾana baʿmil kulli da ʿašān ʾaḥmīku(m)* 'I am doing all this in order to protect you'; *ʾintaẓir ḥatta ywaqqiʿ ilqarār* 'wait until he signs the decision/resolution'; *xallīk fi lbēt laḥatta yittaṣlu bīk* 'stay at home until they contact you!'.

Related to subordination is the syntactic process of catenation, where two (or more) verb forms succeed one another such that the second verb functions as a complement of the first verb (→ serial verbs), e.g. *žarrab yiʿzif ʿala lipyāno* 'he tried to play the piano'; *(ʾi)ttafaʾna nižžawwaz* 'we agreed to get married'; *mā ʾibil ināzilhōn (ināzilōn)* 'he did not agree to meet them in combat or compete with them'; *mā biḥsin ilʿab tanis* 'I can't play tennis'.

Modal auxiliaries in spoken Arabic can also be regarded as catenating in the sense that they are followed by the zero non-past, as in *lāzim tismaʿ kalāmi* 'you must listen to me', *mumkin yiwṣalu bukrah* 'they may arrive tomorrow', *ḍarūri tkallimha* 'it is necessary that you talk to her'. The affinity between the process of catenation and subordinate clauses is that most, if not all, such catenatives can be regarded as having the complementizer *ʾinna* in their structure. Compare *žarrab yiʿzif* with *žarrab innu yiʿzif, (ʾit)tafaʾna nižžawwaz* with *(ʾit)tafaʾna nnana nižžawwaz, lāzim tismaʿ kalāmi* with *lāzim innak tismaʿ kalāmi, ḍarūri tkallimha* with *ḍarūri nnak itkallimha.*

As these examples show, the complementizer *ʾinna* is inflected for person, e.g. *ʾinni* 'that I', *ʾinna(na)* 'that we', *ʾinnak* 'that you', *ʾinnha* 'that she'. Note that *ḍarūri tīji* is ambiguous: 'it is necessary that you/she come', whence the corresponding complementizer takes the forms *ḍarūri nnak/nnik tīji* or *ḍarūri nnaha/nnha tīji*, which are not ambiguous because the pronoun suffixed to *ʾinna* signals the intended person.

There is much more to be said about the subjunctive in spoken Arabic if only space would permit. But one last very common context of the subjunctive has to be added, namely where two verb phrases are joined by *walla* ~ *willa* 'or'. Consider *yīji willa mā yīji* 'whether he comes or not' in *yīji willa mā yīji ʾana mā līš ʿalāʾa* 'whether he comes or not is none of my business'. The same phrase *yīji willa mā yīji* can be used as a question calling for instructions, ‘is he to come or not?’, e.g. *yinžaḥ willa yifšal* ‘whether he passes or fails’; *tiˁūm walla ~ willa tiġraʾ* ‘whether you float/she floats, or drown/s’.

Note that the second verb phrase in these structures is the opposite/negative of the first, but this limitation is by no means mandative, as the following questions/suggestions show: *tiʾra willa ~ walla tnām?* ‘would you like to read or go to bed?’, *tākul willa tnām xafif?* ‘would you like to eat or sleep on a light stomach?’.

Finally, compare the zero non-past of the subjunctive with the *b-* non-past in *bīji willa mā bīji?* ‘does he come or not?’, *yīji willa mā yīji* ‘whether he comes or not’. The contrast is systematic and meaningful, as this entry has intended to show.

5. CONCLUSION

This entry addresses the category of mood in spoken Arabic from a morphological and syntactico-semantic perspective. Mood and modality are shown to be intertwined. The statement presented here is based on formal contrasts and the semantico-pragmatic meanings associated therewith. True, there is need for further elaboration to bring in regional variation and the impact of grammatical categories of the verb on the subject, but the limitations of space and the fact that this entry must not be too technical compel one to leave out any confounding details (cf. Mitchell and El-Hassan 1994).

BIBLIOGRAPHICAL REFERENCES

Comrie, Bernard. 1976. *Aspect.* Cambridge: Cambridge University Press.

Cowell, Mark W. 1964. *A reference grammar of Syrian Arabic.* Washington, D.C.: Georgetown University Press.

El-Hassan, Shahir. 1978. “Educated Spoken Arabic in Egypt and the Levant: A critical review of diglossia and related concepts”. *Archivum Linguisticum* 8:2.112–132.

Lyons, John. 1968. *Introduction to theoretical linguistics.* Cambridge: Cambridge University Press.

——. 1977. *Semantics*, II. Cambridge: Cambridge University Press.

Mitchell, T.F. and Shahir El-Hassan. 1994. *Modality, mood and aspect in spoken Arabic.* London and New York: Kegan Paul International.

SHAHIR EL-HASSAN (Yarmouk University)

Mood (Standard Arabic)

1. MOOD IN ARABIC GRAMMAR

The term ‘mood’ in Western grammars of Standard and Classical Arabic applies to the ‘imperfect’ or ‘prefixed’ verbal stem *-qtul-* and its four endings *-u*, *-a*, *-Ø* (zero), and *-an(na)*. (The full paradigms are given in Table 2, and the prefixes for grammatical person are discussed in Table 1.) The form *ya-qtul-u* ‘he kills/will kill’ (hyphens separate morphemes) serves as the present/future tense in direct opposition to the past or perfect tense stem *qatal-*. The same *ya-qtul-u* is called indicative or ‘realis’ mood in contrast to the two ‘irrealis’ moods, subjunctive *ya-qtul-a* ‘that he kill’ and jussive *ya-qtul* ‘let him kill’. These last two forms signal different kinds of non-narrative future senses. The subjunctive gives a ‘presumed’ or desired action in certain subordinate clauses, whether the action actually occurs or not, actions that the speaker offers as the purpose or motivation for the action of the main clause. The jussive (from the Latin verb *jubēre* ‘to command’) presents a wish, request, or order expressing the 3rd person imperative and 1st person imperative (*na-qtul* ‘let’s kill’). The jussive with no person prefix, *qtul*, is the imperative. (In utterance-initial position the consonant cluster requires a → prothetic vowel, thus *uqtul.*)

The European grammatical terms for these Arabic verb forms come from the tradition of Latin and Greek grammar and only approximate their syntactic and semantic range. The Arabic terms for these forms are partly syntactic and partly formal (see Fleisch 1979:123–132). The form *ya-qtul-a* is called *manṣūb* (meaning tautologically ‘pronounced with a final *-a*’) because it occurs only in subordinate clauses introduced by certain particles: *ʾan* ‘that’ (after verbs of wishing or wanting) and prefixed *li-* and separate *kay*, both meaning ‘so that, in order that’. Recent Western grammars of Modern Written Arabic tend to replace the term ‘subjunctive’ with the more syntactic term ‘dependent imperfect’ to reflect its limited syntax. The form *ya-qtul*, compared to *ya-qtul-u*, is observed to be *majzūm* ‘truncated, apocopated’ since it has, by comparison with *-u* and *-a*, no ending. Modern grammars parallel this formal assessment by calling it ‘apocopated imperfect’. Traditional Arabic grammar collects this set of forms under the term *muḍāriʿ* 'resembling' (→ *māḍī* and *muḍāriʿ*), a formal definition relating to the 'resemblance' of the mood endings *-u*, *-a*, -Ø to the three case endings of definite triptotic nouns nominative *-u*, accusative *-a*, genitive *-i*, although the formal resemblance is only partial. Genitive *-i* and jussive 'zero' do not resemble each other, and the moods have no parallel to the indefinite case endings, which add *-n* to the vowel ending, termed → *tanwīn*, e.g. *al-kitāb-a* 'the book-Acc.' vs. *kitāb-an* 'a book-Acc.'. By contrast, the Arabic term for the *qatal-* stem is *māḍī* 'past', a semantic definition.

A discussion of Arabic mood must also include the intensifying form called → energicus or energ(et)ic, *ya-qtul-an(na)* 'he will certainly kill'. (Older Arabic may also leave off final – *na*.) This fourth mood has no formal parallel in the series of case endings, nor is its syntax or semantics like the other moods. Nonetheless, it is a member of the set of non-past forms based on the imperfect stem *-qtul-*.

The primary opposition of the Arabic verbal system is *qatala/yaqtulu*, discussed extensively in the literature as either 'perfect/imperfect (completed/incompleted) aspect', 'past/non-past tense', a mixture of → tense and → aspect, or the lack of both. In the analysis of the moods, however, linguists do not explicitly raise the question of tense or aspect. (There is some difficulty with this in connection with the jussive negated with *lam*, as discussed below.)

2. Person-gender-number forms

The two stems *qatala/yaqtulu* encode the expression of three persons, two genders, and three numbers, but they distribute this information quite differently. The 'suffixed' stem *qatal-* uses thirteen synthetic endings, while the 'prefixed' stem *-qtul-* has both a prefix and a suffix. In general, the prefix encodes person, and the ending encodes mood; expression of gender and number is divided asymmetrically between the prefix and the suffix.

The stem *-qtul-* divides the three persons over four prefixes, but the mix with the gender (masculine, feminine), number (singular, dual, plural), and the four moods at issue here is not straightforward and illustrates some interesting principles of morphological economy. The 1st person distinguishes only singular and plural: ʾ- for 1st singular and *n-* for 1st plural. There are no forms for gender or dual, and the suffixes express mood only. The prefix *y-* marks only but not all 3rd persons. The prefix *t-* marks all 2nd persons and also 3rd feminine singular and dual. Table 1 illustrates the overlap of the prefix *t-* with 2nd person masculine and 3rd person feminine.

Table 1. Imperfect prefixes and person-gender-number marking

<table>
<tr><td colspan="2"></td><td>sg.</td><td>du.</td><td>pl.</td></tr>
<tr><td>1st</td><td>m.-f.</td><td>ʾ-</td><td colspan="2">n-</td></tr>
<tr><td rowspan="2">2nd</td><td>m</td><td colspan="3" rowspan="2">t-</td></tr>
<tr><td rowspan="2">f.</td></tr>
<tr><td rowspan="2">3rd</td><td colspan="2"></td><td></td></tr>
<tr><td>m.</td><td colspan="3">y-</td></tr>
</table>

Table 2 shows how the person-gender-number prefixes combine with gender-number-mood suffixes.

Four forms in each mood have short vowel endings, and these signal mood only: indicative *-u* (1st sg. *ʾa-qtul-u*, 1st pl. *na-qtul-u*, 2nd masc. sg./3rd fem. sg. *ta-qtul-u*, 3rd masc. sg. *yaqtul-u*), subjunctive *-a* (*ʾa-qtul-a, na-qtul-u, ta-qtul-a, ya-qtul-a*), and jussive zero (*ʾa-qtul, naqtul, ta-qtul, ya-qtul*). In Modern Literary Arabic, speech-final short vowels (after a single consonant) are usually deleted, leaving the semantic distinctions of the moods to the syntax. (In roots termed 'hollow', i.e. with an internal long vowel such as *-qūm-* 'to arise', the indicative and subjunctive are both *ya-qūm*, while the jussive shortens that vowel in *ya-qum* 'let him arise'.)

Five forms have endings beginning in a long vowel, and the indicative has the final syllable *-n(a)*: 2nd fem. sg. *ta-qtul-īn(a)*, 2nd masc. pl. *ta-qtul-ūn(a)*, 3rd masc. pl. *ya-qtul-ūn(a)*, 2nd du./3rd fem. du. *ta-qtul-ān(i)*, 3rd masc. du. *ya-qtul-ān(i)*. Both subjunctive and jussive of these forms have only the long vowel ending: *ta-qtul-ī, ta-qtul-ū, ya-qtul-ū, ta-qtul-ā, ya-qtul-ā*. The ending of the other two forms is the consonant-initial syllable *-na*, which never loses its vowel and is shared by all three moods: 2nd fem. pl. *ta-qtul-na* and 3rd fem. pl. *ya-qtul-na*. All five 2nd person jussives with zero prefix are the imperatives (see Table 3): masc. sg. *(u)-qtul*, fem. sg. *(u)-qtul-ī*, masc. pl. *(u)-qtul-ū*, fem. pl. *(u)-qtul-na*, dual *(u)-qtul-ā*.

Table 2. Combined imperfect paradigms with person prefixes (double-outlined sections show forms with prefix *ta-* that are 'shared' with the area just below it; parentheses enclose deletable elements)

			sg.	du.	pl.
1st	m.-f.	ind.	*ʾa-qtul-(u)*	*na-qtul-(u)*	
		subj.	*ʾa-qtul-(a)*	*na-qtul-(a)*	
		juss.	*ʾa-qtul*	*na-qtul*	
		erg.	*ʾa-qtul-an(na)*	*na-qtul-an(na)*	
2nd	f.	ind.	*ta-qtul-īn(a)*	*ta-qtul-ān(i)*	*ta-qtul-na*
d		subj.	*ta-qtul-ī*	*ta-qtul-ā*	
		juss.			
		erg.	*ta-qtul-in(na)*	*ta-qtul-ānni*	*ta-qtul-nānni*
	m.	ind.	*ta-qtul-(u)*	[shared with 2m.du., 3f.du.]	*ta-qtul-ūn(a)*
		subj.	*ta-qtul-(a)*		*ta-qtul-ū*
		juss.	*ta-qtul*		
		erg.	*ta-qtul-an(na)*		*ta-qtul-un(na)*
3rd	f.	ind.	shared with 2m.sg.		*ya-qtul-na*
		subj.			
		juss.			
		erg.			*ya-qtul-nānni*
	m.	ind.	*ya-qtul-(u)*	*ya-qtul-ān(i)*	*ya-qtul-ūn(a)*
		subj.	*ya-qtul-(a)*	*ya-qtul-ā*	*ya-qtul-ū*
		juss.	*ya-qtul*		
		erg.	*ya-qtul-an(na)*	*ya-qtul-ānni*	*ya-qtul-un(na)*

The energetic is always distinct in form from the other moods. The four forms discussed above with short vowel endings have a two-syllable energetic marker, *-anna*. (Shortened *-an* is more frequent in older Arabic than in Modern Literary Arabic.) Thus, one finds 1st sg. *ʾa-qtul-an(na)*, 1st pl. *na-qtul-an(na)*, 2nd masc. sg./3rd fem. sg. *ta-qtul-an(na)*, 3rd masc. sg. *ya-qtul-an(na)*. The indicative endings with long high vowels, *-īn(a)* and *-ūn(a)*, close the syllable by adding *-na* and shorten that vowel: 2nd fem. sg. *ta-qtul-in(na)*, 2nd masc. pl. *ta-qtul-un(na)*, 3rd masc. pl. *ya-qtulun(na)*. The duals have only one variant with long vowel and geminate consonant: 2nd masc./fem., 3rd fem. *ta-qtul-ānni*; 3rd masc. *ya-qtul-ānni*. The two feminine plural forms have similarly invariable -nānni in 2nd fem. pl. *ta-qtul-nānni*, 3rd fem. pl. *ya-qtul-nānni*. (These last four forms are highly specialized in several respects: the energetic is the most semantically specialized of the moods; the feminine plural and dual are the most grammatically specific, and the endings 'long vowel plus geminate consonant' are unusual in Arabic phonology.) The 2nd persons without a prefix are the energetic imperative (see Table 3): masc. sg. *(u)-qtul-an(na)*, fem. sg. *(u)-qtul-in(na)*, masc. pl. *(u)-qtul-un(na)*, fem. pl. *(u)-qtulnānni*, dual *(u)-qtul-ānni.*

3. SYNTAX

The indicative, energetic, imperative, and jussive can all occur as the verb in a main clause. The imperative occurs only in a main clause, while the other three and the subjunctive also occur in subordinate clauses usually governed by certain conjunctions or particles. The moods also have purely temporal uses to be treated below (cf. Cantarino 1974–1975:I, 77–83).

Table 3. Imperative

	imperative	energetic imperative
m. sg.	*(u)-qtul*	*(u)-qtul-an(na)*
f. sg.	*(u)-qtul-ī*	*(u)-qtul-in(na)*
m. pl.	*(u)-qtul-ū*	*(u)-qtul-un(na)*
f. pl.	*(u)-qtul-na*	*(u)-qtul-nānni*
du.	*(u)-qtul-ā*	*(u)-qtul-ānni*

The indicative opposes the perfect stem *qatal-* as the narrative present and future. It forms the negative *mā ya-qtul-u* or *lā ya-qtul-u*. The adverb *sa(wfa)* specifies the future in *sa-yaqtul-u* or *sawfa ya-qtul-u*, but only the latter can form the negative *sawfa lā ya-qtul-u*. The indicative also forms a past continuous with the past tense of *kān-a* 'to be' in *kān-a ya-qtul-u*. The particle *qad* in *qad ya-qtul-u* adds the modal sense of 'can kill, might kill'.

The jussive, often reinforced with the particle *li-*, gives the 1st and 3rd person indirect imperatives. The 1st person is called hortatory *li-na-qtul* 'let's kill', *li-ʾa-qtul* 'let me kill, I think I'll kill'; the 3rd person is *li-ya-qtul* 'let him (go ahead and) kill'. The negative is *lā: lā ʾa-qtul/lā ya-qtul* 'let me not kill/may he not kill'. The imperative is simply the 2nd person jussive without the prefix *t-*: *(u)qtul* 'kill!', but it can occur only in the positive. Its negative with *lā* requires the full jussive with the prefix: *lā ta-qtul* 'do not kill!'.

The topic of conditional sentences in Classical and Modern Standard Arabic requires a separate discussion (Cantarino 1974–1975:III, 309–371; Badawi a.o. 2004:636–670), but suffice it to say that in Classical Arabic both the jussive and the perfect functioned in hypothetical conditions after the particle *ʾin* 'if' of the type *ʾin ta-qtul, na-qtul; ʾin qatal-ta, qatal-nā* 'if you (were to) kill, we will/would kill'. In Modern Standard Arabic the perfect is taking over from the jussive in this type of condition and replacing *ʾin* with the particle *law*.

The subjunctive can occur only in a subordinate clause (for the syntax of subordinate clauses, see Badawi a.o. 2004:575–635). The complementizer *ʾan*, followed immediately by the subjunctive, forms noun clauses equivalent to an infinitive in English: *ʾarad-tu ʾan taqūm-a* 'wanted-I that you-rise-Subj'. The negative is also *lā*, which merges with *ʾan* to form *ʾallā: ʾarad-tu ʾallā ta-qūm-a* 'I-wanted that-not you-rise-Subj'. The conjunctions *li-* (prefixed) and *kay* (written as a separate word) both express a purpose, like English *in order that* (Badawi a.o. 2004:617–620), as in (1).

(1) *maddat* *yada-hā*
stretched3fs. hand-her
li-tusāfiḥ-a-nī
to-3fs-greet-Subj-me
'She stretched out her hand to greet me' (Cantarino 1975:III, 80)

In addition, *li-* can modify a noun like an English purposive infinitive, as in (2).

(2) *ʿind-ī* *kitābu-n* *li-ʾaqraʾa-hu*
with-me book-a to-1s-read-Subj-it
'I have a book to read'

The energetic has no unique syntactic distribution but can add force to almost any of the other moods, often, but not obligatorily, introduced by the particle *la-*. In place of the regular imperfect in a main clause one finds the energetic in (3).

(3) *fa-la-ʾafʿalanna* *bi-hi* *mā*
and-Part-1s-do-Energ with-him what
faʿal-tu *bi-ṣāḥibi-hi*
did-1s with-friend-his
'I will indeed do to him what I did to his friend' (Cantarino 1974–1975:I, 82)

Instead of a regular imperative or jussive, one finds the energetic in (4) and (5).

(4) *ʾiʿjabanna* *min*
be.surprisedImper-Energ from
sālimin *kayfa* *najā*
healthy.one how survived3ms
'Be (very) surprised that the healthy survived!' (Cantarino 1974–1975:I, 83)

(5) *lā* *ʾasmaʿanna* *ʾaḥadan* *yaqūlu*
Neg 1s-hear-Energ one.person 3ms-say
'Don't let me hear anyone say... !' (Cantarino 1974–1975:I, 83)

4. Negation, tense, mood

As demonstrated above, the mood forms *yaqtulu, yaqtula*, and *yaqtul* are negated with *lā*. The forms called subjunctive and jussive moods, however, also combine with special negative particles that occur nowhere else in Arabic to function as negative tenses (→ negation). The future-oriented subjunctive mood is also the negative future indicative in *lan ya-qtul-a*, the equivalent of *(sawfa) lā ya-qtul-u*. The normally futuristic jussive becomes the negative of the perfect in *lam yaqtul*, the equivalent of the negated perfect *mā qatal-a*. In Classical Arabic these special moods may have felt like 'strong' negations, but they are now the norm in Modern Written Arabic.

Bibliographical references
Badawi, Elsaid, Michael G. Carter, and Adrian Gully. 2004. *Modern Written Arabic: A comprehensive grammar*. London and New York: Routledge.
Cantarino, Vicente. 1974–1975. *Syntax of Modern Arabic prose*. 3 vols. Bloomington and London: Indiana University Press.
Fleisch, Henri. 1979. *Traité de philologie arabe*. II. *Pronoms, morphologie verbale, particules*. Beirut: Dar el-Machreq.

Robert Fradkin (University of Maryland)

Moroccan Arabic

1. General

This is a tentative description of the Moroccan koine, spoken in most urban areas of Morocco. It is important to note that this is not one of the old, prestigious, pre-Hilālian dialects like the ones spoken in Fes, Old Rabat, Salé, or Tetouan. The latter present too many peculiarities to become koineized. There are many regional varieties of Moroccan Arabic, but among a group of speakers from the Maghreb, one can always distinguish a Moroccan from an Algerian or a Tunisian after only a few words.

1.1 *Area, range*

The Moroccan koine serves as a linguistic vehicle in most large towns of Morocco. More than 55 percent of the population now live in towns (51.4 percent at the 1994 census, and 55.1 percent in 2004, according to the Haut Commissariat au Plan du Maroc). Morocco is a special country because of the importance of the presence of → Berber. Over 40 percent of the population is Berberophone; this is an estimate because there are no official figures. But one must add that the great majority of Berberophones are also fluent speakers of Moroccan Arabic, to the point that it can often be considered a second mother tongue. Another estimate states that probably only 8 percent of Berberophones are monolingual in Berber. Morocco is the only officially 'Arab country' that has such a range (in percentage and in absolute figures) of native speakers of a language other than the local Arabic.

Moroccan Arabic is spoken by nearly all Moroccans, and the koine is understood everywhere; it has been described as being the variety of Casablanca, the economic capital of the country.

Morocco has a very particular situation in that the dialect of the political capital, Rabat, is too peculiar (old pre-Hilālian dialect of Andalusi families) to be adopted as a koine; the intellectual capital, Fes, offers the same situation, and although the dialects of these old cities are prestigious in a way, they are not adopted nationwide.

This koine is a vehicular language in Morocco, used in the media and contaminating speakers of more excentric varieties, as in Oujda, where studies have noted that youngsters tend to adopt the koine and keep their local dialect for family use. It is also the language in which those Berberophones who speak varieties that are not mutually comprehensible (Tarifit, Tachelhit, and Tamazight) will communicate. The number of speakers equals the total population of Morocco, i.e. 30,850,000 (2005), from which one must deduce the monolingual Berberophones (estimated at 8 percent).

1.2 *Lifestyle*

The speakers live in towns and are, for the most part, of rural origin, so Moroccan Arabic may rightly be called the language of urbanized rurals.

1.3 *Position*

1.3.1 The variety described here is the koine, spoken by the younger generations in large towns, and in the media for spontaneous speech (interviews in the streets, debates, etc.). It was not used very much for writing until recently, and if it was written, it was mostly for literature or for the publication of poetry, theater plays, or the Moroccan patrimony, such as *melhoun*, proverbs, etc. It is also used for contemporary artistic creation (lyrics of songs, theater, humorists, dialogues in some novels, etc.).

Historically, it was written quite commonly, using various scripts: Jews wrote it in Hebrew script in text dating from the 15th century onward; Muslims wrote it in Arabic script; and, since the French colonization, Latin script was developed intensively, for toponyms, in the Records office, and in schoolbooks, and also for artistic purposes (cartoons). Since the mid 1980s, Arabic script is much more commonly used to represent Moroccan Arabic, due to the Arabization of teaching. Since the beginning of the 21st century, Moroccan Arabic is written frequently in Latin script on the Internet, during chat sessions, and for sending e-mails and SMS on mobile phones. In 2002, between 2 and 6 million SMS were sent daily, most of them in Moroccan Arabic; the figures were even higher in 2007, with over 17 million mobile phones in Morocco. The number of Internet connections was estimated at 1 million for 2003; there were over 4 million in 2006 (with 8,000 cyber-cafés), and 6 million in 2007 via cyber-cafés, although there were only 434,000 home connections.

1.3.2 There is another type of koine, not described here, which would be the Common North African literary language, the → diglossia of the people, who alternate between everyday and literary North African Arabic. This literary variety is probably common to the entire Maghreb (see Cohen 1994).

1.4 *Linguistic type*

Moroccan Arabic in the sense described above is a Western, North African type of Arabic, characterized by a mixture of Bedouin, rural, and urban features. Although it is spoken mostly by Muslims, a small Jewish minority of a few thousand individuals remain, essentially in Casablanca, among which young people speak exactly like Muslims, whereas older people have kept their old variety (mostly pre-Hilālī dialects).

1.5 *Historical evidence*

There are texts in Moroccan Arabic that date back to the 15th century, written by Jews (see Bar-Asher 1977), but the koine under review here is recent. It appeared when masses of rural people came to live in the main towns in the northwest of the country (Casablanca, Rabat-Salé, Fes, Tangiers, and also Marrakesh) and shows many traits that are pre-Hilālī, but also Bedouin influences.

1.6 *State of research*

Moroccan Arabic is a well-known dialect described in numerous descriptions in the first half of the 20th century and studied in a number of recent publications. There is detailed research on phonology, morphology, and syntax. Grammars: Harrell (1962); Caubet (1993); Youssi (1992). Dictionaries: Colin (1993). Dialect atlas: Heath (2002). Textbooks: Benjelloun (1998, 2000); Quitout (1999); Hoogland (1996); Herrero (1998, 2003). Chrestomathy: Colin (1951). For the main sources, see references in Caubet (1993, 2002).

2. LINGUISTIC DESCRIPTION

2.1 *Phonology*

Moroccan Arabic is characterized by innovation on the phonological level: there are some new phonemes, no interdentals, and four or five vowels only, according to the region.

2.1.1 Consonants

The inventory of the consonants is given in Table 1.

Table 1. Consonantal phonemes

	bilabial	labio-dental	dental	alveolar	post-alveolar	palatal	post-palatal	velar	uvular	pharyngeal	glottal
emphasis +/-	-E +E		-E +E	-E +E							
plosive unvoiced			t ṭ				k		q		’
plosive voiced	b ḅ		d ḍ				g				
nasal	m ṃ		n								
trill			r ṛ								
fricative unvoiced		f		s ṣ	š			x		ḥ	h
fricative voiced				z ẓ	ž			ġ		ʿ	
lateral			l ḷ								
semivowel	w					y					

Marginal consonants include /ẓ, ḷ, ṃ, ḅ/; borrowed consonants are /v, p/. There is no impact of Modern Standard Arabic; on the contrary, the standard interdentals are often replaced in Morocco by stops, as in *dahab* for *ḏahab* ‘he went’; in fact, it depends on the teacher, but news on public television is often read without the interdentals. There is some impact from Bedouin dialects, so that [g] may show up instead of [q], e.g. *gāl*, which is the common verb for ‘to say’, *bəgṛa* ‘cow’, *gəmṛa* ‘moon’ in the koine.

New phonemes include /ẓ, ḷ, ṃ, ḅ/, e.g. *ẓawž* ‘sparrow’, *ẓaẓa* ‘wailing’, *ẓga* ‘to shout’, *ṃŭṃṃu* ‘baby’, *ḅəḷḅəḷ* ‘to mumble’ vs. *bəlbəl* ‘to speak fluently’. /r/ can be depharyngealized from /ṛ/, mostly in the context of the vowel /i/; examples (minimal pairs) are *rāb/yrīb* ‘to fall apart’ vs. *ṛāb/yṛūb* ‘to curdle’; *dār/ydīr* ‘to do’ vs. *dāṛ/ydūṛ* ‘to turn’.

The sibilants /s, z/ tend to merge with the palato-alveolars, as in the root *z-w-ž* > *ž-w-ž*: *žūž* ‘two’, *žŭwwəž* ‘to marry’, *žwāž* (*zwāž*) ‘marriage’; *š-m-s* > *š-m-š*: *šəmš* ‘sun’, *tšəmməš* ‘to sunbathe’, *šmīša* (diminutive of *šəmš*) ‘nice and sunny’. There are cases of passage from /l/ to /n/: *sənsla* ‘chain’, *sənsəl* ‘to bind, to chain’. In some contexts, in the presence of liquids and sibilants, historical **j* may be reflected by [g] and lexicalized as such; this is another typically Moroccan trait: *gləs* ‘to sit down’, *gəbṣ* ‘gypsum’, *gnāza* ‘funeral’, *gəzzāṛ* ‘butcher’, *ngāṣ* ‘pear’.

Emphasis means phonetically pharyngealization or → velarization. It may spread forward, as in *ṛṭāḅ* ‘to become smooth’, where [ᵬ] is emphasized because of the presence of the previous /ṭ/ and /ṛ/; or *ḍləm* [ᵭɫəm] ‘to oppress’, where [ɫ] is emphasized because of /ḍ/; *šṛəb* [ʃᵲəᵬ] ‘to drink’ [ᵬ] because of /ṛ/. On the other hand, in *bəyyəḍ* [ᵬəyːəᵭ] ‘to whiten’, [ᵬ] is emphatic because of the following /ḍ/.

There are some articulatory pecularities present in the Moroccan koine: affrication of /t/ > [ts]. /ž/ is pronounced as [ʒ]: *žməl* [ʒməl] ‘camel’, *əžžrād* [əʒːraːd] ‘grasshoppers’.

2.1.2 Vowels

There are five vowels, three long or medium ones: /ā/, /ī/, /ū/, and two short or ultra-short ones: /ə/, /ŭ/. People from Marrakesh and the south only have /ə/, i.e. four vocalic phonemes.

Historically, the number of short vowels has been reduced: /a/ and /i/ merge into /ə/ which is opposed to /ŭ/: *skət* ‘he shut up’ versus *skŭt* ‘shut up!’.

The phonetic realization depends very much on the consonantal environment. /ə/ is realized as [a] in the context of emphatics and pharyngeals, e.g. in *ʿəlləm* ‘to teach’ [ʕalːəm]; [u] in the context of /w/ [wuld] ‘son’; [i] in the context of /y/: [biyːən] ‘to show’; otherwise [ə].

Sociolinguistically, there are two systems of short vowels. Speakers of southern dialects (Marrakesh) who now live in the north, in Casablanca, retain a system of four vowels /ā, ī, ū, ə/, with → labiovelarization, as borrowed from Berber. The word ‘sugar’ is pronounced [sukːaːɍ] in the northern part of the country, and [səkːʷaːɍ] by Marrakshi speakers and in the south.

2.1.3 Diphthongs

In most cases, /aw/ and /ay/ are reduced to /ī/ and /ū/, e.g. *bīt* ‘room’, *ḥūt* ‘fish’, *tūm* ‘garlic’, but the diphthong may remain in some words, e.g. *ḥayṭ* ‘wall’, *xayṭ* ‘thread’, *ḥawma* ‘neighborhood’, *fāyn* ‘where’, *mnāyn* ‘from where’ next to *ḥīṭ, xīṭ, ḥūma, fīn, mnīn*. The preservation of diphthongs is a Bedouin (rural) trait.

2.1.4 Syllable

There is a rule which states that there can be no short vowel in an open syllable. This affects the syllable structure. Vowel length is neutralized in final position.

Open syllables	Closed syllables
v̄: *ā.na* ‘I’, *ā.ži* ‘come!’	
Cv in final position: *mā.ši* ‘going’	CvC: *tək.təb* ‘you will write’
	CvCC: *šədd* ‘to grab’
Cv̄: *mā.ši* ‘going’	Cv̄C: *bān* ‘to show up’
	Cv̄CC: *šādd* ‘holding’
CCv in final position: *mša* ‘to go’	CCvC: *ktəb* ‘to write’
CCv̄: *kbī.ra* ‘big [fem.]’	CCv̄C: *ġlāḍ* ‘he grew thick’

2.1.5 Consonant clusters

In general, a cluster of two consonants is allowed, according to the syllable structure described above. The form #CC- is very common, and does not require a vowel between the two consonants; it mostly occurs in verbs, perfect or imperative: *ktəb* ‘to write’, *xrəž* ‘to go out’, *xrŭž* ‘go out!’, *xdəm* ‘to work’, *lbəs* ‘to dress’, *ġṛəs* ‘to plant’; in nominals: *ḥna* ‘we’, *ktāb* ‘book’, *ḥmāṛ* ‘donkey’. The form -CC# is also very common for nominals: *kəlb* ‘dog’, *ġəṛs* ‘plant’, *šəms* ‘sun’, *bənt* ‘girl’; in verbs it occurs in *šədd* ‘to grab’, *kŭbb* ‘pour!’.

2.1.6 Stress

Stress does not have a distinctive role in Moroccan Arabic. It is linked to intonation and expressivity.

2.1.7 Phonotactics

Sun letters including /ž/ assimilate the article: *əzzmān* ‘the time passed’, *əžžməl* ‘the camel’, *əṛṛāžəl* ‘the man’, *əssūq* ‘the market’, *əttūt* ‘the mulberry’. The → *ʾimāla* is very slight in Moroccan Arabic; it accompanies front consonants. For suprasegmental spread of pharyngealization, see section 2.1.1. In contact position, the imperfect prefix *t-* assimilates to the first consonant of the root, if it is voiced: *tdīr* > [dːiːr] ‘she will do’.

2.1.8 Morphophonology

When a vocalic suffix is added to the 3rd person singular feminine of the perfect, the /ə/ of the suffix is altered in both quality and length to /ā/: /ḍəṛbət + -ək/ >: *ḍəṛbātək*. This contrasts with some pre-Hilālī dialects, like in Fes, which lengthen the /t/ of the suffix: *dəṛbəttək*. Vocalic suffixes are added to the 3rd person plural without change of the ending: /xdāw + u/əh/ > *xdāwəh*.

In the construct state of the feminine noun, the *-a* is replaced by *-ət* and the /ə/ is elided subsequently when a vocalic possessive suffix is added: /fṣālat + u/ > *fṣāltu* ‘its shape’; /xālat + i/ > *xālti* ‘my aunt’, sometimes with lengthening of the vowel of the feminine suffix: /rukbat/ > *rŭkbāti* ‘my knee’. The active participle functions like a verbal form. When feminine, the suffix is directly added to the final *-a*, which is not replaced by *-ət* as with nouns: /mṣebbna + ha/ > *mṣəbbnāha* ‘I/you/she [sg. fem.] have washed it’.

In the → pseudodual, the final *-n* disappears when a suffix is added: → /yəddīn + i/ > *yəddīya* ‘my hands’; /yəddīn + k/ > *yəddīk* ‘your hands’.

Indirect object suffixes are added to the direct ones with the usual vowel lengthening: *gŭlthālu* ‘I told him so’, *gŭlnāhālu* ‘we told him so’.

2.2 *Morphology*

As a characteristic of Moroccan Arabic, there is no gender distinction in the 2nd person singular in the perfect, the historically feminine form serving for both persons: *xdəmti* 'you worked'. But in the koine pronouns, the imperative, and the 2nd person singular of the imperfect do distinguish between the two genders: *ənta* 'you [masc.]', *ənti* 'you [fem.]'; *nūḍ* 'get up [masc.]!', *nūḍi* 'get up [fem.]!'; *ka-təxdəm* 'you [masc.] work', *ka-txədmi* 'you [fem.] work'. There is no gender distinction in the plural.

The verbal system displays synthetic tendencies with several suffixes added one to the other: *ma-ṣīfəṭ-nā-hā-l-ū-š* 'we didn't send it to him'.

2.2.1 Pronouns

2.2.1.1 Personal independent pronouns

singular		plural	
3rd masc.	*hūwa*	3rd	*hūma*
3rd fem.	*hīya*		
2nd masc.	*ənta/əntāya*	2nd	*əntūma*
2nd fem.	*ənti/əntīya*		
1st	*āna*	1st	*ḥna*

2.2.1.2 Possessive/object suffix

singular		plural	
3rd masc.	*-u/-əh*	3rd	*-hŭm*
3rd fem.	*-ha*		
2nd	*-ək*	2nd	*-kŭm*
1st	*-i/-ni* (object)	1st	*-na*

2.2.1.3 Indirect object: *l-* or *li-* (more Bedouin, used in Casablanca)

singular		plural	
3rd masc.	*lu/līh*	3rd	*lhŭm/līhŭm*
3rd fem.	*lha/līha*		
2nd	*lək/līk*	2nd	*lkŭm/līkŭm*
1st	*li/līya*	1st	*lna/līna*

2.2.1.4 Demonstratives

Proximal demonstratives include *hāda* [masc.], *hādi* [fem.], *hādu* [pl.], which are used independently for 'this [one]'. As an innovation, with the disappearance of the gender distinction, invariable *hād-əl* is found in head position: *hād-əṛṛāžəl* 'this [masc.] man', *hād-əlbənt* 'this [fem.] girl', *hād-ənnās* 'these [pl.] people'.

As distal demonstratives *hādāk* [masc.], *hādīk* [fem.], *hādūk* are found. These occur without the *hā-* in head position: *dāk-əṛṛāžəl* 'that man [masc.]', *dīk-əlbənt* 'that girl [fem.]', *dūk-ənnās* 'those people [pl.]'.

2.2.1.5 Presentatives

There are two → presentatives: *ha-* and *ṛa-* (see Caubet 1992); their paradigms (Table 2) are complex, mixing independent pronoun and suffix.

Table 2. Presentatives

singular		plural	
3rd masc.	*hāhūwa*	3rd	*hāhūma*
3rd fem.	*hāhīya*		
2nd	*hāk/hānta*	2nd	*hākŭm/hāntūma*
1st	*hāni/hāna*	1st	*hāḥna*
singular		**plural**	
3rd masc.	*ṛāh*	3rd	*ṛāhŭm*
3rd fem.	*ṛāha*		
2nd	*ṛāk*	2nd	*ṛākŭm*
1st	*ṛāni*	1st	*ṛāḥna*

Whereas *ha-* is a simple presentative, as in *hāni žīt* 'I've come!', *ṛa-* is a means of reasserting or opposing what was said before; it is always linked to a previous utterance: *ṛā-ni mṛīḍ* means 'I am really ill [contrary to what you thought].'

2.2.1.6 Relative pronoun

The invariable *lli* serves as a relative pronoun.

2.2.2 Adverbs

Temporal adverbs include *dāba* 'now, right away'; *dīma* ~ *daʾimən* 'always'; *ġədda* 'tomorrow'; *əlbārəḥ* 'yesterday'; *əlyūm* 'today'.

Local adverbs: proximal *hna* 'here', *hnāya* 'here', *ha fāyn* 'here'; distal: *təmma* 'there', *təmmāk* 'over there', *təmmāya* 'over there', *hnāk* 'there', *ṛa fāyn* 'over there'. There is a third deixis which points to the not visible: *lhīh, dākəžžīh*. Other local adverbs are *əlfūq* 'above', *əltəḥt* 'under', *əllūṛ* 'at the back', *əldāxəl* 'inside', *əlgŭddām* 'forward', *mən hna* 'through here', *mən təmma* 'through there', *bəṛṛa* 'outside'.

The following manner adverbs may be noted:

wālu ‘nothing’; *bəzzāf* ‘a lot’; *šwīya* ‘little’; *gāᶜ* ‘completely’; *šḥāl* ‘how much/many’; *dəġya* ‘fast, quickly’; *yəmkən, wāqīla, tᶜāyli* ‘maybe’; *gāla, gālāk* ‘on the contrary’.

An innovative general question particle *wāš* is placed at the beginning of the sentence: *wāš žāw* ‘did they come?’, *wāš klītu* ‘did you eat?’, *wāš yžīw* ‘will they come?’ (see 2.2.3, v).

Interrogative adverbs, placed at the beginning of the sentence, include: *škūn* ‘who?’, *škūn ža* ‘who came?’; *šnu* ‘what?’, *šnu šrīti* ‘what did you buy?’; *šḥāl* ‘how much/many?’, *b šḥāl xditīh* ‘how much did you pay for it?’; *ᶜlāš* ‘why?’, *ᶜlāš žīti* ‘why did you come?’; *kīfāš* ‘how?’, *kīfaš lqītīh* ‘how did you find him?’; *fāyn/fīn* ‘where?’, *mnāyn/mnīn* ‘where from?’, *fāyn mšīti* ‘where did you go?’; *mnāyn žīti* ‘where did you come from?’; *fūqāš* ‘when?’ (< *f-wəqt-āš), *fūqāš žāw* ‘when did they arrive?’; *əmta* is used more in Casablanca and in the south; *mā lək* ‘what’s wrong with you?’

2.2.3 Particles

i. The definite article is *əl/lə-*, which assimilates to sun letters (see 2.1.7).

ii. Moroccan Arabic is extremely innovative in the use of the indefinite article, more than most other North African dialects (→ article, indefinite). There are two new quantifiers which mark some degree of indetermination, one formed on the numeral *wāḥəd* and the other on a grammaticalization of the word *šāy > ši ‘thing’, (see Caubet 1983a, 1983b, 1984).
The construction with the numeral is extremely common, using a construct state: *wāḥəd-əṛṛāžəl* ‘a man’, *wāḥəd-əlwəld* ‘a child’, *wāḥəd-əlbənt* ‘a girl’. The quantifier *ši* marks a certain kind of indetermination: *ši wəld* ‘a boy/some boy’, *ši nhāṛ* ‘one day/some day’.

iii. The most common genitive marker is *dyāl/d*, but there are more and more occurrences in the Moroccan koine of *mtāᶜ*, which originates from the Bedouin dialects. *dyāl* is mostly used with affix pronouns, *d* more in nominal constructions.

iv. Negational particles: Verbs, including the active participle, are negated by discontinuous *ma…š* (*ši/šāy*) (see Caubet 1986, 1990, 1993, 1996b), e.g. *ma-šŭftūš* ‘I did not see him, *ma-qālhālīš* ‘he didn’t tell it to me’, *ma-qāṛīš* ‘he hasn’t studied [he can’t read]’, *ma-mṣəbbnāš əlḥwāyəž* ‘I haven’t washed the clothes’, whereas continuous *māši* serves as nominal negation and is normally not combined with pronouns, e.g. *əlwəld māši mṛīḍ* ‘the boy is not ill’, *āna māši mṛīḍa* ‘I am not ill’, *āna māši ṭbība* ‘I am not a doctor’. If the speaker insists and one really wants to deny, one may say: *ma-nīš ṭbība bāš tgūlu*… ‘I am not a doctor for you to say…’.

v. Interrogative particles: The particle *wāš* introduces questions bearing on the whole clause: *wāš əlbənt šāfət xāha* ‘did the girl see her brother?’ (see Caubet 1993:73–76). There can also be a curious use of *ši*, after the predicate: *šəfti ši əlwəld* ‘have you seen the boy [at all]?’, *klīti ši* ‘have you eaten [at all]?’ (see Caubet 1983a, 1984).

vi. Existentials: *kāyn, ma kāyn š* ‘there is, there isn’t’ serve as existentials.

vii. Prepositions: Monoconsonantic prepositions are *f-/fi-* ‘in, at’: *f əlwād* ‘in the river’, *fīya* ‘in me’, *fīh* ‘in him’; *b-/bi-* ‘with’, indirect object *b əddārīža* ‘in Moroccan Arabic’, *bīya* ‘with me’, *bīha* ‘with her’; *l-/li-* ‘to, for’, *l-fās* ‘to Fes’, *li, līya* ‘to me’.
Other simple prepositions include *mən* ‘from’, *mən Fās* ‘from Fes’, *mənnu* ‘from him’, *mənha* ‘from her’; *ᶜla* ‘on’, *ᶜla lkŭrsi* ‘on the chair’, *ᶜlīya* ‘on me’, *ᶜlīhŭm* ‘on them’.

viii. Conjunctions include, *ū/w* ‘and, with’, *əlwəld w əlbənt* ‘the boy and the girl’; *wiyya* ‘with’, *āna wiyyāk* ‘you and me’; *wəlla, āw* ‘or’, *əlbənt wəlla lwəld* ‘the girl or the boy’, *əṣṣəḥḥ āw lla* ‘right or no?’; *walayənni* ‘but’.

ix. Vocative particles are the following: *a*, *ya*, to insist: *a mḥəmməd!* ‘hey Mohamed!’. If he doesn’t answer at first, one may insist: *yá mḥəmməd!* (see Caubet 1995b).

x. Exclamations: Both verbal and nominal sentences can be used for exclamations: *šḥāl kayəžri!* ‘how fast he runs!’, *ᶜla bənt!* ‘what a girl!’

2.2.4 Nouns

2.2.4.1 Gender

Some of the feminine nouns without feminine marker *-a(t)* are common in Semitic (Caubet 1990), while others are specific to the Maghreb under the influence of Berber. Body parts: *ržəl* ‘leg’, *wdən* ‘ear’, *yədd* ‘hand, arm’, *ʿīn* ‘eye’, *kərš* ‘belly’, *dāt* ‘body’, *sənn* ‘tooth’, and also *rūḥ* ‘soul’. Knives and swords: *sīf* ‘sword’, *mūs* ~ *səkkīn* ‘knife’. Others are *dāṛ* ‘house’, *bāb* ‘door’ (sometimes masc.), *ṭṛīq* ‘road’, *ṣūf* ‘wool’; *šəms* ‘sun’, *smən* ‘preserved butter’, *zīt* ‘oil’, *xātəm* ‘ring’, *xādəm* ‘woman slave’, *bənt* ‘girl’, *xətt* ‘sister’.

2.2.4.2 Productive nominal patterns

2.2.4.2.1 Instruments

For instruments mə-CCəC is used (Caubet 1993:100).

2.2.4.2.2 Professions

The intensive CəCCāC is used for professions, with a peculiar pattern of plurals. There are often two plurals, one for the tradesmen, one for the borough, in traditional towns like Fes: *nəžžāṛa* ‘carpenters’ as opposed to *nəžžāṛīn* ‘the carpenter quarter in Fes’. A pattern for names of profession or trades is borrowed from Berber ta-CəCCāC-t: *tanəžžārt* ‘carpentry’.

2.2.4.3 Plural patterns (see Caubet 1993; Harrell 1962)

2.2.4.3.1 External plurals

Common suffixes for external plural are *-īn*, which occurs mainly with participles (*məktūb*/*məktūbīn*), → *nisba* adjectives (*fāsi*/*fāsiyīn* ‘from Fes’), diminutives of adjectives (*kbībəṛ*/*kbībəṛīn* ‘biggish’, *zwīn*/*zwīnīn* ‘nice’); *-a*, which occurs with human groups, such as names of trades, ethnonyms (*zəllayži*/*zəllayžiya* ‘mosaicist’), the intensive pattern CəCCāC (*nəžžāṛ*/*nəžžāṛa* ‘carpenter’, *bənnāy*/*bənnāya* ‘mason’); *-āt*, which occurs with diminutives of nouns in both genders (*skīkən*/*skiknāt* ‘knife’, *ṣnīdəq*/*ṣnidqāt* ‘box, chest’); then *nisba* in *-iya* (*fūqīya*/*fūqīyāt* ‘overgarment’, *nāmūsīya*/*nāmūsīyāt* ‘bedstead’), the pattern CəCCāCa for females (*xəyyāṭa*/ *xəyyāṭāt* ‘dressmaker’), plurals of noun of unity (*nəmla*/*nəmlāt* ‘ant’, *bīḍa*/*bīḍāt* ‘egg’), new loanwords ending in *-u* (*kāḍu*/*kāḍuyāt* ‘present’, *rīḍu*/*rīḍuyāt* ‘curtain’, *kāmyu*/*kāmyāt*~*kamwāt* ‘lorry’); *-ān*, which occurs with a small number of nouns in CCəC > CəCCān (*drəʿ*/*dərʿān* ‘forearm’, *ḍfəṛ*/*ḍəfṛān* ‘nail’, as well as *ṭṛīq*/*ṭəṛqān* ‘road’, *xṛūf*/*xəṛfān* ‘lamb’, *ġzāl*/*ġəzlān* ‘gazelle’).

2.2.4.3.2 Internal patterns (see Caubet 1993: I; Harrell 1962)

Internal patterns include the insertion of /ā/ after the second consonant, as in CCāC, CCāC-a, CCāC-i, CCāC- īn, CWāC, CYāC, formed from singulars in CCəC, CəCC, CuCC: *žməl*/*žmāl* ‘camel’; *kəlb*/*klāb* ‘dog’; *ʿŭšš*/*ʿšāš* ‘nest’; adjectives in CCīC: *kbīr*/*kbār* ‘big; old’; *tqīl*/*tqāl* ‘heavy’; feminines in CəCCa: *bəlġa*/*blāġi* ‘babouche’; *žəṛḥa*/*žṛāḥi* ‘wound’.

The insertion of /ū/ after the second consonant, as in CCūC, CCūC-a, CCūC-āt, CwūC, CYūC, formed from singulars in CCəC, CəCC, CuCC: *gəṛn*/*gṛūn* ‘horn’, *qəlb*/*qlūb* ‘heart’; in CCāC: *ktāb*/*ktūb* ‘books’; in CīC, CāC: *ṭīṛ*/*tyūṛ* ‘bird’, *bīt*/*byūt* ‘room’, *dāṛ*/*dyūṛ* ‘house’, *dīb*/*dyūba* ‘wolf’.

In patterns such as CCəC, there is a shortening of the word: *ġāba*/*ġyəb* ‘forest’, *ṭbīb*/*ṭəbba* ‘doctor’.

CāC, CīCān represents a mixed pattern: *fāṛ*/*fīrān* ‘mouse; rat’, *bāb*/*bībān* ‘door’; even with loanwords, *kāṛ*/*kīrān* ‘coach’, and ethnonyms in *nisba* > CCāCa: *rīfi*/*ryāfa* ‘Rifian’, *tūnsi*/*twānsa* ‘Tunisian’, *sūsi*/*swāsa* ‘from the Sous region’, *žəbli*/*žbāla* ‘from the Jbala mountain region’.

CCūCāt marks a great amount in the case of uncountable nouns: *zīt*/*zyūtāt* ‘lots of oil; too much oil’; *dhəb*/*dhūbāt* ‘lots of gold; too much gold’.

2.2.4.4 → Pseudodual

The plurals of body parts are formed with a suffix *-īn*, like external plurals: *yədd*/*yəddīn* ‘hand’; *wdən*/*wədnīn* ‘ear’; *ržəl*/*rəžlīn* ‘leg’.

2.2.4.5 Diminutive

The diminutive is very productive and is used to express affectivity or close relations with the addressee, or to establish a climate of intimacy. Patterns are usually those of the adjectives of color, but there are differences according to the region. More urban: *kḥīḥəl* ‘blackish’, *ḥlīlu* ‘nice’, *ḥmīməṛ* ‘reddish’. More rural: *kḥiyyəl* ‘blackish’, *ṣġiyyəṛ* ‘smallish’, *kbiyyər* ‘biggish’.

2.2.4.6 Colors and deficiencies

The pattern is CCəC. The plural forms of these adjectives differ from one dialect to another: a more rural koine will have a plural in CəCCīn: *kəḥlīn* ‘black [pl.]’. A more urban one will use the pre-Hilālī plural form in CūCəC: *kūḥəl* ‘black [pl.]’, *ḥūməṛ* ‘red [pl.]’, *xūḍəṛ* ‘green [pl.]’.

2.2.4.7 Elative patterns

The pattern is generally CCəC: *kbīr* > *kbəṛ* ‘bigger’, *ktīr* > *ktəṛ* ‘more’, with emphatic /ṛ/ showing up in the absence of the /ī/: *ṣġīṛ* > *ṣġəṛ* ‘smaller’; *ṭwīl* > *ṭwəl* ‘longer’. It is generally followed by *mən* for comparison: *kbəṛ mənni* ‘older/bigger than me’, *ṭwəl mənnu* ‘taller than him’.

With the same meaning, the basic form of the adjective with the preposition *ˁla* may be used: *ṭwīl ˁlīya* ‘taller than him’, *kbīr ˁlīh* ‘older/bigger than me’. This is propably a calque from a Berber construction (see Aguadé and Vicente 1997).

2.2.5 Numerals

2.2.5.1 *wāḥəd, wəḥda, wəḥdá* is used as an adjective: *ˁəndu bənt wəḥdá* ‘he has only one daughter’.

2.2.5.2 Dual vs. *žūž* (*d-əl*)+ plural

The dual suffix is *-āyn* or *-āyən*. This has very restricted usage, in the sense that it is only used with nouns of measures of time, distance, and weight, of Arabic origin; for example, one finds from *ṛṭəl* ‘pound’, from *ṛəṭlāyn* ‘two pounds’; from *yūm, yūmāyn* ‘two days’; from *ˁām, ˁāmāyn* ‘two years’; from *šhəṛ, šəhṛāyn* ‘two months’; from *ūqīya, ūqīytāyn* ‘two ounces’, from *šbər, šəbrāyn* ‘two spans’. All metric measures are excluded from the use of the dual, so that *yiṭṛu* ‘liter’, *mīṭəṛ* ‘meter’, *kīlu* ‘kilo’ do not take a dual, but *žūž* is used instead: *žūž kīlu* ‘two kilos’, *žūž yiṭṛu* ‘two liters’.

2.2.5.3 From 3 to 10

There are two paradigms, one for nouns of measure, the other for all other nouns. The nouns of measure are those which take the dual suffix (see 2.2.5.2). These form a special paradigm using the reduced invariable form CəCC of the numeral. All other nouns use an analytic construction with the particle *d*, very similar to the one expressing possession (see Table 3).

Table 3. Numerals from 3 to 10

Noun	Short form (measure)	Analytic construction
tlāta	*təlt iyyām*	*tlāta d-əlktūb*
aṛbˁa ~ ṛəbˁa	*ṛəbˁ iyyām*	*ṛəbˁa d-əlktūb*
xəmsa	*xəms iyyām*	*xəmsa d-əlktūb*
sətta	*sətt iyyām*	*sətta d-əlktūb*
səbˁa	*səbˁ iyyām*	*səbˁa d-əlktūb*
tmənya	*təmn iyyām*	*tmənya d-əlktūb*
təsˁūd	*təsˁ iyyām*	*təsˁa d-əlktūb*
ˁəšṛa	*ˁəšṛ iyyām*	*ˁəšṛa d-əlktūb*

This analytic construction is peculiar to Moroccan Arabic. Another peculiarity is the taboo on the form *təsˁa* lit. ‘let him beg!’, which is replaced by a longer form *təsˁūd*, probably from the root *s-ˁ-d* ‘happy’.

2.2.5.4 From 11 to 19

The isolated form has lost the /ˁ/ and the final /ṛ/ (**xəmsət ˁašəṛ* > **xməstˁəšṛ* > *xməsṭāš* ‘fifteen’), the emphatic quality of which is still present in /ṭ/, which harks back to the feminine suffix **at*: *ḥḍāš, ṭnāš, təlṭāš, ṛbəˁṭāš, xməsṭāš, səṭṭaš, sbəˁṭāš, tmənṭāš, tsəˁṭāš*.

There are two tendencies: one is to expand this short form by *-əl*, the /l/ of which represents the former final /ṛ/ of *ˁəšṛ* (*-ən* and *-əṛ* can be found in this case elsewhere in the Maghreb), e.g. *ḥḍāšəl yūm* ‘eleven days’,

xməsṭāšəl yūm	‘fifteen days’
səṭṭāšəl yūm	‘sixteen days’

The other tendency is to use the same paradigm as for the numbers between 3 and 10, i.e., by connecting the above numerals with the plural form of the noun, which leads to construction such as *ṭnāš d-əl-ktūb*.

2.2.5.5 Higher numbers

20–99: The counted noun is in the singular; the nouns are *ˁəšrīn, təltīn, ṛəbˁīn, xəmsīn, səttīn, səbˁīn, tmānīn, təsˁīn* as in *ṛəbˁīn ktāb* ‘forty books’. Units precede the tens: *wāḥəd u təltīn* ‘thirty-one’, *xəmsa u xəmsīn ktāb* ‘fifty-five books’.

100: *mya* > *myāt* before a singular noun: *myāt ˁām* ‘a hundred years’, *myāt ktāb* ‘a hundred books’.

200 > 900: *mya* remains in the singular.

1,000 > 9,000: *ʾalf* is converted into the plural *alāf*.

The paradigms are constructed like nouns of measure (see 2.2.5.3); the counted noun is always in the singular: *təlt myāt ktāb*; *təlt alāf ktāb*.

Table 4. Hundreds and thousands

myatāyn	*alfāyn*
təlt mya	*təlt alāf*
ṛəbˁ mya	*ṛəbˁ alāf*
xəms mya	*xəms alāf*
sətt mya	*sətt alāf*
səbˁ mya	*səbˁ alāf*
təmn mya	*təmn alāf*
təsˁ mya	*təsˁ alāf*
ˁəšṛ mya	*ˁəšṛ alāf*

2.2.5.6 Ordinal numbers

These are usually formed with the pattern CāCəC, *tāni, tālət, sātət*, etc., which is applied to the numbers 11 and 12 as well: *ḥādəš* 'eleventh', *ṭānəš* 'twelfth'. For the rest of the paradigm, the isolated form is used: *xməsṭāš* 'fifteenth', *ˁəšrīn* 'twentieth'.

2.2.5.7 Count nouns

For collectives or small elements, *ḥəbba* is used to indicate the unit: *ḥəbba d-əžžəlbāna* 'a green pea', *ḥəbba d-əzzītūn* 'one olive', *ḥəbba d-əlmələḥ* 'one grain of salt'.

2.2.6 Verb

2.2.6.1 Forms

2.2.6.1.1 Form I

In Form I the two short vowels /ə/ and /ŭ/ of Moroccan Arabic often distinguish between perfect and imperfect, as in *skət/yəskŭt* 'to shut up', *šxəṛ/yəšxŭṛ* 'to snore'; on the other hand, one finds *ktəb/yəktəb* 'to write', *sxəf/yəsxəf* 'to pass out'.

2.2.6.1.2 Derived Forms

Form XI replaces Form IX and has an inchoative meaning; the pattern is CCāC, with some cases of CCəC: *ġlāḍ* 'to become thick', *ṣġāṛ* 'to become small', *smān* 'to become fat', *ṭwāl* 'to grow', but, on the other hand, *kbər* 'to get old'.

The derived Forms are given in Table 5.

Table 5. Derived Forms

II	CəCCəC/yCəCCəC:	*ˁəlləm/yˁəlləm* 'to teach'
III	CāCəC/yCāCəC:	*fāṛəq/yfāṛəq* 'to leave'
V	tCəCCəC/ yətCəCCəC:	*tˁəlləm/yətˁəlləm* 'to learn'
VI	tCāCəC/yətCāCəC:	*tfāṛəq/yətfāṛəq* 'to be separated'

The prefix for the passive Form VII is *t-* in the koine, but there are dialectal cases of *n-*, *tn-*, and *nt-* (see Aguadé a.o. 1995). This *t-* prefix becomes *ttə-* when the stem starts with a consonant cluster: *ttə-ktəb, t-kətbət* 'it was written'. There are only a few items of Form VIII with a *-t-* infix: *ḥtāž* 'to need', *htəmm* 'to be interested in'.

Form II has the usual factitive, intensive, and denominative functions; Form V is reflexive, Form VI reciprocal.

2.2.6.2 Inflection of imperfect and perfect

2.2.6.2.1 Imperfect

	singular	plural
3rd masc.	*yəktəb*	*ykətbu*
3rd fem.	*təktəb*	
2nd masc.	*təktəb*	*tkətbu*
2nd fem.	*tkətbi*	
1st	*nəktəb*	*nkətbu*

The prefixes *ka-* and/or *ta-* express for all persons habit, repetition, concomitance, generality: *ka-yəktəb* 'he is writing, he writes, usually or habitually', see 2.3.3.1.

	singular	plural
3rd masc.	*ka-yəktəb*	*ka-ykətbu*
3rd fem.	*ka-təktəb*	
2nd masc.	*ka-təktəb*	*ka-tkətbu*
2nd fem.	*ka-tkətbi*	
1st	*ka-nəktəb*	*ka-nkətbu*

2.2.6.2.2 Perfect

Moroccan Arabic does not make gender distinction in the 2nd person singular of the perfect.

	singular	plural
3rd masc.	*ktəb*	*kətbu*
3rd fem.	*kətbət*	
2nd masc./fem.	*ktəbti*	*ktəbtu* (pre-Hilālī -*tīw*)
1st	*ktəbt*	*ktəbna*

2.2.6.3 Participles

In general, the active participle of Form I is CāCəC *šāṛəb* ‘having drunk’, *fāhəm* ‘having understood’, but there are some exceptions where the pattern CəC$_2$C$_2$āC is used, e.g. *xəddām* ‘working’.

Derived Forms: The active and passive participles of Forms II and V, and Forms II and VI are the same, e.g. *mʿəlləm* ‘having taught, having learned’ (active and passive participle of the two verbs *ʿəlləm* ‘to teach’ and *tʿəlləm* ‘learn’); *mṣāləḥ* ‘having reconciled [trans. or reflex.]’: (active and passive participle of the two verbs *ṣāləḥ* and *tṣāləḥ*). In verbal usage, the feminine suffix is -*a*, and suffixes are added with lengthening of the -*a* as in *mṣəbbnāhŭm* ‘I/you [fem.] have washed them’.

2.2.6.4 Verbal nouns

Among the verbal nouns, one must distinguish between the abstract noun naming the activity, and the name of an individual action, usually marked by the suffix -*a*, e.g., both *təḥmīm* ‘the action of going to the public bath, of having a bath’ and *təḥmīma* ‘a bath, a session at the public bath’ are derived from the verb *tḥəmməm* ‘to go to the public bath, to have a bath’. Irregular is the verb *kla* ‘to eat’ with the verbal noun *mākla*.

2.2.6.4.1 Form I

Regular verbs: The most common types of verbal nouns derived from verbs CCəC are CəCC, CCīC, CCāC.

i. CəCC (CCəC-a for the nomen vicis). The position of the *ə* distinguishes between verb and nominal: *frəq* ‘to separate’ > *fərq* ‘difference’; *žṛəḥ* ‘to wound’ > *žəṛḥ* ‘wound’; *ṭṛə*z ‘to embroider’ > *ṭəṛz* ‘an embroidering’. For geminated verbs in CəC$_2$C$_2$, the verbal noun has the same form (*šəkk* ‘to doubt; doubt [noun]’, *hərr* ‘to tickle; tickle [noun]’).

ii. CCīC: *ġsəl* > *ġsīl* ‘washing’; *nbəḥ* > *nbīḥ* ‘barking’; *šṛəb* > *šṛīb* ‘drinking’.

iii. CCāC: *lbəs* > *lbās* ‘clothing, clothes’; *drəs* > *drās* ‘threshing’.

2.2.6.4.2 Derived Forms

Forms II and V: CəC$_2$C$_2$əC/tCəCCəC. The verbal noun has the pattern təCCīC or təCCāC; sometimes both forms exist: *bəddəl* > *təbdāl*/*təbdīl* ‘change’; *tḥəmməm* > *təḥmīm* ‘having a bath’; *ṣəbbən* > *təṣbīn* ‘laundry’; *fəttət* > *təftāt* ‘the act of crumbling’.

Forms III and VI: CāCəC/tCāCəC have a verbal noun with the pattern mCāCC-a: *gābəl* > *mgābla* ‘the act of taking care’, *ḥāwəl* > *mḥāwla* ‘an attempt’, *tʿānəq* > *mʿānqa* ‘a hug’, *lāqa* > *mlāqya* ‘a gathering’.

Forms IX or XI; CCāC have a verbal noun with the pattern CCīC: *ṣġāṛ* > *ṣġŭṛ* ‘the fact of being small, youth’; *tqāl* > *tqŭl* ‘the fact of being heavy; weight’.

2.2.7 Weak verbs

2.2.7.1 Geminated verbs

dəkk/ydəkk ‘to tramp down’ and *kəbb/ykŭbb* ‘to pour’

2.2.7.1.1 Imperfect of geminated verbs

	singular	plural
3rd masc.	*ydəkk/ykŭbb*	*ydəkku/ykŭbbu*
3rd fem.	*tdəkk/tkŭbb*	
2nd masc.	*tdəkk/tkŭbb*	*tdəkku/tkŭbbu*
2nd fem.	*tdəkki/tkŭbbi*	
1st	*ndəkk/nkŭbb*	*ndəkku/nkŭbbu*

2.2.7.1.2 Perfect of geminated verbs

	singular	plural
3rd masc.	*dəkk/kŭbb*	*dəkku/kŭbbu*
3rd fem.	*dəkkāt/kŭbbāt*	
2nd masc./fem.	*dəkkīti/kŭbbīti*	*dəkkītu/ kŭbbītu*
1st	*dəkkīt/kŭbbīt*	*dəkkīna/ kŭbbīna*

Imperative	sg.	*dəkk dəkki*	*kŭbb kŭbbi*
	pl.	*dəkku*	*kŭbbu*
Participles		*dākk*	*kābb*

2.2.7.2 Verbs Iʾ

The perfect of the verbs Iʾ has merged with that of the verbs III*y*, e.g. *kla/yākŭl* ‘to eat’, *xda/yāxŭd* ‘to take’, whereas in the active participle

/ʾ/ is replaced by /w/: *wākəl*. The imperative follows the paradigm of verbs IIw: *kūl* ‘eat!’. Like the participle, Form II replaces initial /ʾ/ by /w/: *wəkkəl/ywəkkəl* ‘to feed’, active participle *mwəkkəl*.

2.2.7.2.1 Imperfect of the verbs Iʾ

	singular	plural
3rd masc.	*yākŭl*	*yāklu*
3rd fem.	*tākŭl*	
2nd masc.	*tākŭl*	*tāklu*
2nd fem.	*tākʷli*	
1st	*nākŭl*	*nāklu*

2.2.7.2.2 Perfect of the verbs Iʾ

	singular	plural
3rd masc.	*kla*	*klāw*
3rd fem.	*klāt*	
2nd masc./fem.	*klīti*	*klītu*
1st	*klīt*	*klīna*

Imperative	sg.	*kūl kūli*	*xūd xūdi*
	pl.	*kūlu*	*xūdu*
Participles	sg.	*wākəl wākla*	*wāxəd wāxda*
	pl.	*wāklīn*	*wāxdīn*

2.2.7.3 Verbs I*w*

wqəf/yəwqəf~yūqəf ‘to stand, stop’

2.2.7.3.1 Imperfect of verbs I*w*

	singular	plural
3rd masc.	*yūqəf*	*yūqfu*
3rd fem.	*tūqəf*	
2nd masc.	*tūqəf*	*tūqfu*
2nd fem.	*tūqfi*	
1st	*nūqəf*	*nūqfu*

2.2.7.3.2 Perfect of verbs I*w*

	singular	plural
3rd masc.	*wqəf*	*wəqfu*
3rd fem.	*wəqfət*	
2nd masc./fem.	*wqəfti*	*wqəftu*
1st	*wqəft*	*wqəfna*

Imperative	*wqəf wəqfi wəqfu*
Participle	*wāqəf wāqfa wāqfīn*

2.2.7.4 Verbs II*w/y*

dār/ydīr ‘to do’, *gāl/ygūl* ‘to say’, *bāt/ybāt* ‘to spend the night’

2.2.7.4.1 Imperfect of verbs II*w/y*

	singular	plural
3rd masc.	*ydīr/ygūl/ybāt*	*ydīru/ygūlu/ybātu*
3rd fem.	*tdīr/tgūl/tbāt*	
2nd masc.	*tdīr/tgūl/tbāt*	*tdīru/tgūlu/tbātu*
2nd fem.	*tdīri/tgūli/tbāti*	
1st	*ndīr/ngūl/nbāt*	*ndīru/ngūlu/nbātu*

2.2.7.4.2 Perfect of verbs II*w/y*

	singular	plural
3rd masc.	*dār/gāl/bāt*	*dāru/gālu/bātu*
3rd fem.	*dārət/gālət/bātət*	
2nd masc./fem.	*dərti/gŭlti/bətti*	*dərtu/gŭltu/bəttu*
1st	*dərt/gŭlt/bətt*	*dərna/gŭlna/bətna*

Imperative	*dīr/gūl/bāt*
Participles	*dāyər/gāyəl/bāyət*

The derived forms take the endings of the verbs III*y*: 1.sg. VIII: *ṛtāḥīt* ‘I took a rest’.

2.2.7.5 Verbs III*y*

There are three types: *a*- type and *i*- type, which are very productive, and *u*- type, represented by only one verb (*ḥba/yəḥbu* ‘to crawl’). The paradigm has been entirely renovated: *nsa/yənsa* ‘to forget’, *mša/yəmši* ‘to leave’, *ḥba/yəḥbu* ‘to crawl’.

2.2.7.5.1 Imperfect of verbs III*y*

	singular	plural
3rd masc.	*yənsa/yəmši/yəḥbu*	*yənsāw/yəmšīw/yəḥbu*
3rd fem.	*tənsa/təmši/təḥbu*	
2nd masc.	*tənsa/təmši/təḥbu*	*tənsāw/təmšīw/təḥbu*
2nd fem.	*tənsāy/təmši/təḥbu*	
1st	*nənsa/nəmši/nəḥbu*	*nənsāw/nəmšīw/nəḥbu*

2.2.7.5.2 Perfect of verbs III*y*

	singular	plural
3rd masc.	*nsa/mša/ḥba*	*nsāw/mšāw/ḥbāw*
3rd fem.	*nsāt/mšāt/ḥbāt*	
2nd masc./fem.	*nsīti/mšīti/ḥbīti*	*nsītu/mšītu/ḥbītu*
1st	*nsīt/mšīt/ḥbīt*	*nsīna/mšīna/ḥbīna*

Imperative	*nsa, nsāy, nsāw* *sīr, sīri, sīru* is used instead of *mši, mšīw* *ḥbu*
Participles	*nāsi, nāsya, nāsyīn*

2.2.7.6 Irregular verbs

dda/yəddi 'to take away', *ža/yži* 'to come'

2.2.7.6.1 Imperfect

	singular	plural
3rd masc.	*yəddi/yži*	*yəddīw/yžīw*
3rd fem.	*təddi/tži*	
2nd masc./fem.	*təddi/tži*	*təddīw/tžīw*
1st	*nəddi/nži*	*nəddīw/nžīw*

2.2.7.6.2 Perfect

	singular	plural
3rd masc.	*dda/ža*	*ddāw/žāw*
3rd fem.	*ddāt/žāt*	
2nd masc./fem.	*ddīti/žīti*	*ddītu/žītu*
1st	*ddīt/žīt*	*ddīna/žīna*

Imperative	*ddi ddīw*	*aži, ažīw*
Participles	*ddāy ddāya* *ddāyīn, žāy žāya* *žāyīn*	

2.2.8 Quadriradical verbs

Reduplication: *kəṛkəṛ* 'to pack up hurriedly', *bəḷbəḷ* 'to mutter', *bəlbəl* 'to talk fluently', *zəˁzəˁ* 'to wobble', *ẓəˁẓəˁ* 'to yell'. The fourth radical is often *š*, *ṣ*, *ṭ*, *l*, *t*, e.g. *zəġṛəṭ/yzəġṛəṭ* 'to ululate'.

2.2.8.1 Imperfect of quadriradical verbs

	singular	plural
3rd masc.	*yzəˁzəˁ*	*yəzˁəzˁu*
3rd fem.	*tzəˁzəˁ*	
2nd masc.	*tzəˁzəˁ*	*təzˁəzˁu*
2nd fem.	*təzˁəzˁi*	
1st	*nzəˁzəˁ*	*nəzˁəzˁu*

2.2.8.2 Perfect of quadriradical verbs

	singular	plural
3rd masc.	*zəˁzəˁ*	*zˁəzˁu*
3rd fem.	*zˁəzˁāt ~ zˁəzˁət*	
2nd masc./fem.	*zəˁzəˁti*	*zəˁzəˁtu*
1st	*zəˁzəˁt*	*zəˁzəˁna*

Imperative	*zəˁzəˁ zˁəzˁi zˁəzˁu*
Participles	*mzəˁzəˁ mzəˁzˁa* *mzəˁzˁīn*

2.3 *Syntax*

2.3.1 Noun phrase

2.3.1.1 Definiteness

The article *əl-* expresses definiteness (see Caubet 1993, 1983).

2.3.1.2 Expressions of indefiniteness and specificity

Two quantifiers are very often used in Moroccan Arabic to increase the determination somewhat; one is derived from the numeral *wāḥəd*: *wāḥəd-əlwəld* 'a boy', the other from the word **šay* 'thing' > *ši wəld* 'a boy, some boy' (see 2.2.3); *hād-əlwəld* 'this boy'; *ḥāža* 'thing' > *ši ḥāža* 'something'.

2.3.1.3 Construct state: Types

The → construct state is restricted to certain domains (see Harrell 1962:191–200): constructs in which the first term is an adjective: *byəḍ əlwžəh* 'whitefaced'; body parts, used literally or metaphorically: *ṛās əššhəṛ* 'the first day of the month', *ˁŭnq əṛṛāžəl* 'the man's neck'; kinship terms: *bənt ˁammi* 'my cousin', *wlād ənnās* 'people's children'; expressions using *mūl* (lit. 'owner'): *mūl əddžāž* 'the chicken seller', *mūl əddāṛ* 'the master of the house', *mūl əlḥānūt* 'the shopkeeper'; and finally a series of set expressions, e.g.: *bāb əddāṛ* 'the main door', *klām ənnās* 'what people say'.

2.3.1.4 Quantifiers

'all, every' is expressed by *kŭll* as elsewhere; it may be postposed: *ənnās kŭllhŭm/kŭllha* 'all the people'.

'some': *ši nās* 'some people'
'little': *šwīya d-əlma* 'a little water'
'a lot': *bəzzāf d-əlma* 'a lot of water'

2.3.1.5 Relative clauses

When the antecedent is definite, *lli* usually introduces the relative clause, as in *əlktāb əlli šrīt* 'the book I bought'; with an indefinite antecedent, an asyndetic relative clause follows: *wāḥəd əlbənt šŭftha əlbāraḥ* 'a girl I saw yesterday'.

2.3.2 Verbal phrase

A direct object introduced by *fī* expresses duration, as in *əlbəgṛa katākŭl f ərrbīʿ* 'the cow is eating grass', but this is marked as more Bedouin speech. Pronominal indirect objects are suffixed to the verb by means of the preposition *l* or *li*: *gāl-l-u* or *gāl-lī-h* (Casablanca). Direct and indirect suffixes may be combined: *ma-ka-ygūl-hā-lhŭm-š* 'he does not say it to them'.

2.3.3 Verbal aspect; time and tense

2.3.3.1 Aspectualizers

ka- ~ *ta-* serve as aspectualizers with a global imperfect meaning, compatible with generality, repetition, and concomitance (see Caubet 1993: II). Concomitance can also be marked by *xəddām* ~ *gāləs*, as in *xəddām kayākŭl* 'he is busy eating' (Caubet 1996a). Some people tend to use only *ka-*. In Fes, the same speaker can alternate between *ka-* and *ta-*. Old families from Fes use mostly *ta-*.

2.3.3.2 Future-intent prefixes

The future-intent prefix is *ġādi*, as in *ġādi yži* 'he'll come', which can be reduced to *ġa* as in *ġa yži* 'he will come', *ma-ġa-yži-š* 'he won't come'.

2.3.3.3 Use of active participle

The meaning of the active participle depends on the semantics of the verb. In particular, verbs of movement and position of the body form a class on their own (Caubet 1991, 1993: II), the participle of which has a concomitant and prospective aspectual value: *žāy* 'he's coming', *fīn māšyīn ġədda* 'where are you going tomorrow?', *wəldək nāʿəs* 'your son is asleep'. With verbs not belonging to this group, the participle has resultative value (present perfect): *mṣəbbna əlḥwāyəj* 'I/you/she [fem.] have/has washed the clothes', *qāri* 'he can read and write', *wākla* 'I've [fem.] already eaten'.

2.3.3.4 Auxiliaries

Auxiliaries expressing durativity, intent, wishes, and so on include *bda* 'to start', *bqa* 'to go on', *ṛžəʿ* 'to do again', *mabqaš* 'to be no longer', *ma zāl* (fem. *ma zāla*, pl. *ma zālu*) 'to continue'.

2.3.3.5 Use of periphrastic narratives

The use of *ža* as a periphrastic narrative is common in Moroccan Arabic: *ža hūwa gāllu* lit. 'he came he told him' > 'then he told him' (see Caubet 1995a).

2.3.4 Word order

The normal word order is SVO for propositions, but VSO in subordinate clauses and with some interrogatives (see Caubet 1993: II). Interrogatives generally introduce the sentence: *fīn mšīti* 'where did you go?', *šnu bġīti* 'what do you want?', *fūqāš mša əlwəld* 'when did the boy leave?'. The grade adverb *bəzzāf* may follow a noun, as in *əlma bəzzāf* 'a lot of water', *hūwa qṣīr bəzzāf* 'he is very short'.

2.3.5 Agreement/concord

There is agreement in gender (sg.) or number (pl.). Examples (within the noun phrase) are: *əlbənt əṣṣġīṛa* 'the little girl'. There is generally no gender distinction for the plural of adjectives: *əlbnāt* ~ *lūlād əṣṣġāṛ* 'the little girls ~ boys'. Human and nonhuman take plural agreement: *əsskākən əlkbāṛ* 'the big knives', *əṭṭūmūbīlāt əlkbāṛ* 'the big cars', *ūlādək ma zālīn ṣġāṛ* 'your children are still small'. Within the verb phrase: *əṭṭūmūbīlāt əžždād kayžrīw ktīṛ mən əlqdām* 'the new cars go faster than the old ones'. Reference by pronouns: *ənnās lli ma kanʿərfūhŭmš* 'the people we do not know'. Bedouin dialects and Casablanca mark concord for plurals of human groups in the feminine singular: *ənnās kŭllha katəbġi əllḥəm* 'everybody likes meat'.

2.3.6 Existential sentences

kāyən and its negation *ma kāyənš* serve as markers of the existential; a definite subject is possible: *kāyən əlma* 'there is water'; *ma kāyənš əlma* 'there is no water'.

2.3.7 Conditional sentences

The marker for the realis is mainly *īla*, e.g. *īla ža dāba, dāba* 'if he comes right now'. For the irrealis there are many markers, all with the same usage: *lūkān, ūka kān, kūn, kūkān, kūn kān*, e.g. *lūkān ža* 'if he came', 'if he had come'.

2.3.8 *Ḥāl-* sentences

Ḥāl- sentences have the usual form *w* + subject + predicate, as in *hādi sāˁa w-āna kantsənna* 'I have been waiting for an hour', *hādi sāˁa wə-hwa nāˁəs* 'he's been asleep for an hour'.

3. Lexicon

There are many loanwords from Berber, Romance languages (probably very old), the *lingua franca*, Valencian, Catalan, Venetian, Sicilian, and others. (see Lévy 1990: V, VI). More recent loanwords come from French (→ French loanwords) and Spanish, especially for all technical terms. They are used for neologisms: French *prise* > *p/brīz*; French *téléphone* > *tilīfōn*; for all car parts, e.g. French *frein* > *fṛīn*, *pot d'échappement* > *šākma*, *code de la route* > *kūdlārūṭ* (see Heath 1989). When their pattern is compatible with Moroccan words, and when they are well integrated, they form a broken plural, e.g. *qāmīža* 'shirt', pl. *qwāməž; tənbəṛ* 'stamp', pl. *tnābəṛ*.

Bibliographical references

Aguadé, Jordi and Mohamed El Yaacoubi. 1995. *El dialecto árabe de Skūra (Marruecos)*. Madrid: CSIC.

Aguadé, Jordi and Ángeles Vicente. 1997. "Un calco semántico del beréber en árabe dialectal magrebí: El uso de la preposición *ˁla* en el comparativo". *Estudios de Dialectología Norteafricana y Andalusí* 2.225–240.

Bar-Asher, Shalom. 1977. *Les ordonnances rabbiniques des rabbins de Castille établis à Fès*. Jerusalem: Centre Zalman Chazar. [Ed. Abraham Ankawa, 1871, Livourne, Italie, under the title *Kerem ḥemre*.]

Benjelloun, Saïd. 1998. *El-luġa dyali: El-ˁarabiya el-maġribiya*, II. Tilburg and Oisterwijk: Babylon and Dutch University Press.

——. 2000. *El-luġa dyali: El-ˁarabiya el-maġribiya*, I. Tilburg and Oisterwijk: Babylon and Dutch University Press.

Caubet, Dominique. 1983a. *La détermination en arabe marocain*. (= *Collection ERA*, 642). Paris: Université Paris VII.

——. 1983b. "Quantification, négation, interrogation: Les emplois de la particule *ši* en arabe marocain". *Arabica* 30:3.227–245.

——. 1984. "À la recherche d'un invariant: Les emplois de la particule *ši* en arabe marocain". *Opérations de détermination*, II, 33–56 (= *Collection ERA*, 642). Paris: Université Paris VII.

——. 1986. "Les deux parfaits en arabe marocain". *Aspects, modalité: Problèmes de catégorisation grammaticale*, 71–102. (= *Collection ERA*, 1028). Paris: Université Paris VII.

——. (in collaboration with Marie-Claude Simeone-Senelle and Martine Vanhove). 1990. "Genre et accord dans quelques dialectes arabes". *Genre et langage, Actes du colloque tenu à Paris X-Nanterre les 14, 15, 16 Décembre 1988*, ed. Eliane Koskas and Danielle Leeman. *Linx* 21.39–66.

——. 1991. "The active participle as a means to renew the aspectual system: A comparative study in several dialects of Arabic". *Semitic studies in honor of Wolf Leslau*, ed. Alan Kaye, I, 209–224. Wiesbaden: O. Harrassowitz.

——. 1992. "Deixis, aspect et modalité: Les particules *hā-* et *ṛā-h* en arabe marocain". *La deixis* (*Colloque en Sorbonne, Juin 1990*), ed. Mary-Annick Morel and Laurent Danon-Boileau, 139–149. Paris: P.U.F.

——. 1993. *L'arabe marocain*, I. *Phonologie et morphosyntaxe*, II. *Syntaxe et catégories grammaticales, textes*. Louvain: Peeters.

——. 1995a. "*ža*, élément narratif dans le récit familier en arabe marocain". *Dialectologia arabica: A collection of articles in honour of the 60th birthday of Professor Heikki Palva*, 41–48. Helsinki: Finnish Oriental Society.

——. 1995b. "L'exclamation en arabe marocain". *Faits de Langue* 6.189–198.

——. 1996a. "*gāləs kayxdəm, xāyəḍ kayxdəm*: Approche sociolinguistique de l'expression de la concomitance en arabe marocain". *Estudios de Dialectologia Norteafricana y Andalusi* 1.87–100.

——. 1996b. "La négation en arabe maghrébin". *La négation en berbère et en arabe maghrébin*, ed. Salem Chaker and Dominique Caubet, 79–97. Paris: L'Harmattan.

——. 2002. "Questionnaire de dialectologie du Maghreb". *Estudios de Dialectologia Norteafricana y Andalusi* 5.73–92.

Cohen, David. 1994. Préface to *Actes des Premières Journées Internationales de Dialectologie de Paris*, ed. Dominique Caubet and Martine Vanhove. Paris: INALCO.

Colin, Georges S. 1951. *Chrestomathie marocaine*. Paris: Adrien-Maisonneuve.

——. 1993. *Le dictionnaire Colin d'arabe dialectal marocain*, ed. Zakia Iraqi Sinaceur. 8 vols. Rabat: Editions Al Manahil.

Harrell, Richard S. 1962. *A short reference grammar of Moroccan Arabic*. Washington, D.C.: Georgetown University Press.

Heath, Jeffrey. 1989. *From code-switching to bor-*

rowing: A case study of Moroccan Arabic. London and New York: Kegan Paul International.
——. 2002. *Jewish and Muslim dialects of Moroccan Arabic*. London and New York: RoutledgeCurzon.
Herrero Muñoz-Cobo, Bárbara. 1998. *Gramática de árabe marroquí para hispano-hablantes*. Madrid: Universidad de Almería.
——. 2003. *¡Habla árabe marroquí!: Método para principiantes*. Madrid: Universidad de Almería.
Hoogland, Jan. 1996. *Marokkaans Arabisch: Een cursus voor zelfstudie en klassikaal gebruik*. Amsterdam: Bulaaq.
Lévy, Simon. 1990. *Parlers arabes des Juifs du Maroc: Particularités et emprunts – Histoire, sociolinguistique et géographie dialectale*. 6 vols. Thèse d'état, University of Paris VIII.
Marçais, Philippe. 1977. *Esquisse grammaticale de l'arabe maghrébin*. Paris: Maisonneuve.
Quitout, Michel. 1999. *Initiation à l'arabe maghrébin*. Paris: L'Harmattan.
——. 2001. *Parlons l'arabe dialectal marocain*. Paris: L'Harmattan.
Youssi, Abderrahim. 1992. *Grammaire et lexique de l'arabe marocain moderne*. Casablanca: Wallada.

DOMINIQUE CAUBET (INALCO, Paris)

Morocco

1. STUDIES ON MOROCCAN DIALECTS

1.1 Modern research on Moroccan Arabic is more than two centuries old: the first grammar of a Moroccan dialect was published in the year 1800 in Vienna (Dombay 1800). But it was at the end of the 19th and the beginning of the 20th century that publications on Moroccan dialects increased (Lerchundi 1872, 1892; Marçais 1911; Kampffmeyer 1912). During the French-Spanish Protectorate (1912–1956), some of the most important works on Moroccan dialectology were published (see, for instance, Brunot 1931–1952; Colin 1921, 1955; Destaing 1937; Lévi-Provençal 1922; Loubignac 1952; Mercier 1951; Singer 1958a, and 1958b). In the years following independence, publications decreased rapidly (one exception being Harrell 1962, 1966), and only at the end of the 1980s did interest in Moroccan dialects and sociolinguistics begin to rise again, resulting in a considerable number of new studies (Stillman 1988; Youssi 1992; Caubet 1993; Durand 1994; Aguadé and Elyaacoubi 1995; Ishihara 2000; Vicente 2000; etc.).

A small but very useful dialectal atlas has been published by Heath (2002), and Peter Behnstedt is now working on a regional atlas for northern Morocco (Behnstedt 2002; Behnstedt and Benabbou 2005; Behnstedt 2005).

1.2 In spite of the existing body of publications, research on Moroccan dialects is still far from complete, and for large areas (mainly in eastern and southern Morocco), data are still very scarce.

2. LANGUAGES SPOKEN IN MOROCCO

2.1 The kingdom of Morocco has, according to the last census, an estimated population of 32 million inhabitants (July 2004).

In Morocco, there are two vernacular languages: Arabic and → Berber. No official data are available concerning the number of Berber speakers, but according to some scholars, they may represent between 35 percent and 40 percent of the whole population. In any case, due to the rapid increase of migration from rural to urban areas, Berber is in regression: its low social prestige leads Berber speakers to speak Moroccan Arabic more than their mother tongue, for integrative purposes (Ennaji 2005:78).

2.2 In outline, Berber is spoken in the Rif Mountains (Tarifit), in the Middle Atlas (Tamazight), and in a vast area between the High Atlas and the edges of the Sahara (Tashelhit: in the Sous, Draa and Dades Valley, Tafilalt, Antiatlas, with some Arabic language islands like Skura in the Dades or Oulad Tayma in the Sous).

3. ARABIC DIALECTS OF MOROCCO

3.1 Little is known about the linguistic situation in the country prior to the arrival of the Arabs. Berber was the mother tongue of the majority of the population, and in the north (in the former Mauretania Tingitana) a Romance language was also spoken in some major urban centers like Tangier, Ceuta, Sala, Volubilis, etc.

The Arabization of Morocco was the result of two waves: first, the Islamic conquest in the

7th century C.E., and second, the arrival of the Banū Hilāl in the 12th century.

From a diachronical point of view, there are therefore two different groups of Arabic dialects in Morocco: pre-Hilalian and Hilalian dialects.

3.2 To the pre-Hilalian group belong all dialects which arose after the Arab conquest in the 7th century: it seems that in the first centuries after the Islamic conquest the Arabization was quite superficial, and Arabic was spoken only in the most important cities; the rest of the country continued using Berber or a Romance language (Lévy 1998). Dialects belonging to this type are those spoken in the mountains of Jebala, in the traditional towns (like Tangier, Tetouan, Chaouen, Fes, Sefrou, Rabat, etc.), and all Jewish dialects.

3.3 The Hilalian dialects came to the country with the Banū Hilāl (Banū Sulaym and Banū Maʿqil) tribes, which emigrated from Arabia to Egypt and North Africa and eventually arrived in Morocco in the 12th century.

To the Hilalian type belongs → Ḥassāniyya, the dialect spoken in southern Morocco and in Mauritania.

3.4 A broad classification of Moroccan was made first by Colin (1986) and then completed, with more data, by Heath (2002:2–12). According to Heath, Moroccan dialects can be classified in the following categories:

i. The northern type (sedentary, pre-Hilalian): dialects of Tangier, Tetouan, Jbala, ancient medina of Rabat, Fes, Sefrou, Taza, etc.
ii. The central type (Hilalian, sedentary, rural): dialects of Oujda, Atlantic coast south of Rabat, Casablanca, rural dialects around Fes and Sidi Kasem, Atlantic coast south of Casablanca, El Jadida, Essaouira, Marrakech, Skura, etc. Due to the increasing influence of Casablanca (economic pole and biggest city in the kingdom) and migration of speakers of rural dialects to the most important urban centers, this type is now the most widespread in Morocco.
iii. The Saharan type (Hilalian, Bedouin but now sedentary): → Ḥassāniyya in Saharan towns like Mhamid, Tata, Goulimine, Bedouin dialects like that of the Zʿīr in the plains between Rabat and Casablanca.
iv. The Jewish dialects (pre-Hilalian, sedentary): dialects of the Jewish communities in the traditional Moroccan towns. Almost all Jewish dialects have now disappeared in Morocco due to emigration to Israel, Europe, and America.

4. Phonetics

4.1 *Vowels*

From a phonological point of view, Moroccan dialects show two vocalic systems.

4.1.1 By far the most common vocalic system in the Moroccan dialects (Hilalian and pre-Hilalian) shows three long and two short vowels:

/ā/ /ī/ /ū/
/ə/ /u/

4.1.2 In such dialects, [a, ɑ] are merely allophones of /ə/ in contact with opening phonemes like /ḥ/, /x/, /ʿ/, /q/, etc.: *xəmsa* [xamsa] ‘five’, *qəlb* [qɑlb] ‘heart’, *ʿənd* [ʕənd] ‘at, with, near’. In the same contexts /u/ has the allophone [ʊ, o]: *wəqt* [wʊqt] ‘time’. There is no /i/ in such dialects: [i, ɪ] are allophones of /i/ in contact with /y/: *zəyynu* [zijjnu] ‘they adorned’, *xəyyəṭ* [xɪjjəᵵ] ‘he sewed’, *yəktəb* [jiktəb] ‘he will write’.

4.1.3 Allophones of the three long vowels are:

/ā/: [æ:] in plain contexts, e.g. *šāfu* [ʃæːfu] ‘they saw’; [aː, ɑː] in contact with pharyngeal, uvular, and pharyngealized consonants, e.g. *xāla* [xaːla] ‘maternal aunt’, *bḥāl* [bħaːl] ‘like’, *qāḍi* [qɑːᵭɪ] ‘judge’.
/ī/: [iː] in plain contexts, e.g. *zīt* [ziːt] ‘oil’; [ɪː, eː] in contact with pharyngeal, uvular, and pharyngealized consonants, e.g. *ʿīn* [ʕeːn] ‘eye’, *ṣīf* [ᵴɪːf] ‘summer’.
/ū/: [uː] in plain contexts, e.g. *žūž* [ʒuːʒ] ‘two’; [ʊː, oː] in contact with pharyngeal, uvular, and pharyngealized consonants, e.g. *ṣūq* [ᵴʊːq] ‘market’, *qūq* e.g. [qoːq] ‘artichoke’.

4.1.4 Examples of vowel oppositions are:

/ə/ vs. /u/: *mədd* ‘hold out!’ vs. *mudd* ‘container used for measuring grain’, *ḥəbb* ‘he kissed’ vs. *ḥubb* ‘love’.
/ə/ (= [a]) vs. /ā/: *xəmsa* [xamsa] ‘five’ vs. *xāmsa* [xaːmsa] ‘fifth [fem.]’, *ḥməq* [ħmɑq] ‘crazy’ vs. *ḥmāq* [ħmɑːq] ‘he became crazy’, *kḥəl* [kħal] ‘black’ vs. *kḥāl* [kħaːl] ‘he turned black’.
/u/ vs. /ū/: *dxul* ‘enter!’ vs. *dxūl* ‘entrance’.

4.1.5 The second system has been documented only in some north-eastern Hilalian dialects (to the north of Debdou and south of Oujda), classified as types B and C by Behnstedt and Benabbou (2005:17–19).

It shows three long and three short vowels (whose phonetic realizations are the same as those described above):

/ā/	/ī/	/ū/
/a/	/ə/	/u/

4.1.6 Examples of vowel oppositions in these dialects are:

/a/ vs. /ə/: *ḥanna* ‘grandmother’ vs. *ḥənna* ‘henna’, *ḥažž* ‘he made the pilgrimage’ vs. *ḥəžž* ‘make the pilgrimage!’.
/a/ vs. /u/: *dagg* ‘he pulverized’ vs. *dugg* ‘pulverize!’, *ṛašš* ‘he sprinkled’ vs. *ṛušš* ‘sprinkle!’.
/ə/ vs. /u/: *skət* ‘he shut up’ vs. *skut* ‘shut up!’, *xrəž* ‘he got out’ vs. *xruž* ‘go out!’.

4.1.7 The short diphthongs **-aw* and **-ay* of Classical Arabic are generally preserved in northern and Jbala dialects as well as in some (Hilalian) Bedouin dialects (Zʕīr): examples from Anjra are **yawm* > *yawm* ‘day’, **nawba* > *nawba* ‘time, turn’, **lawḥa* > *lawḥa* ‘wooden tablet’ (Vicente 2000:34).

In most Moroccan dialects (for instance Rabat, Casablanca, Fes, Marrakech), though, such diphthongs are monophthongized.

4.1.8 Common to all Moroccan dialects is that short vowels do not occur in open syllables or word-finally: **daxaltu* > *dxəlt* ‘I entered’, **fahima* > *fhəm* ‘he understood’, **madīna* > *mdīna* ‘town’, **ṭarīq* > *ṭṛīq* ‘way’. In order to avoid the occurrence of a short vowel in an open syllable, vowel elision and metathesis take place: *ṣāfəṛ* ‘he traveled’ > *ṣāfṛu* ‘they traveled’, *kəmməl* ‘he finished’ > *kəmmlu* ‘they finished’, *ktəf* ‘shoulder’ > *kətfi* ‘my shoulder’, *šṛəb* ‘he drank’ > *šəṛbu* ‘they drank’ (Aguadé 2003a:95, 99).

4.1.9 Stress is not a distinctive feature: concerning its position some rules can be found for isolated words (see for instance Durand 1994:47–48), but in a phrase it is absolutely free.

4.1.10 The question about the existence of quantity opposition in Moroccan dialects is a controversial one: most scholars accept its existence, others, however, simply deny it. According to Heath, there is no length opposition in northern and Jbala dialects, but it is present in other, central and southern, dialects (Heath 2002:188–189).

In two recent articles, Behnstedt has defended, with strong arguments, quantity opposition in Moroccan, pointing out that long vowels in such dialects are as long as in Oriental dialects (Behnstedt and Benabbou 2002:62; Behnstedt 2004:53).

In any case, even if it is true that quantity opposition is not very functional in Moroccan dialects due to the general loss of vowels, an important number of minimal pairs can easily be found.

4.2 *Consonants*

4.2.1 **b*: reflexes of **b* are [ß] and *ḅ*. In northern dialects, [ß] appears in intervocalic position: *bībān* > *bīßān* ‘doors’, *bībi* > *bīßi* ‘turkey’ (Vicente 2000:38). *ḅ* appears in pharyngealized contexts: *ḍṛəb* > *ḍṛəḅ* ‘he beat’.

There are some examples of the shift **b* > *m*: *mašā* > *bša* ‘he went’ (dialect of the Zʕīr), *mʕa bən* ‘with whom?’, *dyāl bən* ‘whose?’ (dialect of Chaouen).

4.2.2 **ṯ*, **ḏ*, **ḏ̣*: reflexes of the interdentals **ṯ*, **ḏ*, and **ḏ̣* are respectively *t (ṭ)*, *d(ḍ)*, and *ḍ* in the majority of the Moroccan dialects: **ṯlāṯa* > *tlāta* ‘three’, **ṯaʕlab* > *təʕləb* ‘fox’, **burġūṯ* > *bəṛġūt* ‘flea’, **ṯawr* > *tūr* > *ṭūṛ* ‘bull’, **ḏahab* > *dhəb* ‘gold’, **hāḏā* > *hād* ‘this’, **haḏara* > *hdər* > *hḍəṛ* ‘he spoke’, **ḏura* > *dra* > *ḏṛa* ‘corn’, **ḏ̣ill* > *ḏ̣əll* ‘shadow’.

The interdentals of Classical Arabic are preserved only in Ḥassāniyya and some Bedouin dialects like that of the Zʿīr. Examples from the dialect of the Zʿīr are *_ḥarrāṯ_ > _ḥərrāṯ_ ‘farmer’, *_ʾakṯar_ > _kṯəṛ_ ‘more than’, *_ḏabaḥa_ > _ḏbəḥ_ ‘he slaughtered’, *_ḏ̣alla_ > _ḏ̣əll_ ‘he spent the day’, *_ġalīḏ̣_ > _ġlīḏ̣_ ‘thick’ (Loubignac 1952; Aguadé 1998:142).

In the dialect of the Zʿīr, the interdental *_ṯ_ has a reflex _f_ in cases like *_ṯamma_ > _fəmmāk_ ‘there’, _ṯānī_ > _fāni_ ‘also, again’ (Heath 2002:132).

Interdental phonemes in some northern pre-Hilalian dialects (Chaouen, for instance) are secondary (the result of the influence of Berber substratum), and they occur only in intervocalic or final position: *_qālat_ > _ʾāləṯ_ ‘she said’, *_zayt_ > _zīṯ_ ‘oil’, *_ṯlāṯa_ > _tlāta_ > _tlāṯa_ ‘three’, *_ibn ʾādam_ > _mnāḏəm_ ‘human being’, *_bilāḏ_ > _blāḏ_ ‘land’ (Moscoso 2003:39–40).

4.2.3 _t*_ is very often affricated and becomes [t^{s}]. In pharyngealized contexts it is realized as _ṭ_ [ᵵ].

4.2.4 *_j_: reflexes of *_j_ in contact with sibilants are deaffricated to _d_ or _g_, e.g.*_jaḥš_ > _dəḥš_ ‘young donkey’, *_jāza_ > _gāz_ ~ _dāz_ ‘he passed [by]’ (Heath 2002:136–138). In most dialects, *_j_ is pronounced as _ž_ (= [ʒ]). In northern and Jbala dialects, *_j_ is pronounced as [dʒ] if geminated or in contact with _n_ or _r_: _jjbəl_ ‘the mountain’, _jjīrān_ ‘the neighbors’, _šfənj_ ‘doughnut’, but in other positions it is realized as _ž_ (Vicente 2000:45). The shift _ž_ > _z_ is very common in Jewish dialects: _ībəl_ > _zbəl_ (Heath 2000:132–133).

4.2.5 *_l_: In some dialects of the Tafilalt valley (Igli/Igni, Zrigat), *_l_ is pronounced as _n_: _nxəl_ > _nxən_ ‘palm trees’, _tā-ngūlu_ > _tā-ngūnu_ ‘we say’ (probably Berber influence: Behnstedt 2004). In other dialects, the shift *_l_ > _n_ in cases like *_silsila_ > _sənsla_ ‘chain’ or _zilzāl_ > _zənzāl_ ‘earthquake’ is very common.

4.2.6 *_k_: is pronounced as _t_ in the Jewish dialect of Tafilalt: _lkəlb_ > _ltəlb_ ‘the dog’ (Heath 2002:140). Reflexes of *_k_ are [ç] and sometimes [x] in northern and Jbala dialects. Examples: *_yākul_ > _yāçul_ ‘he will eat’, *_ʾakaltu_ > _klīt_ > _çlīt_ ‘I ate’, *_kṯar_ > _xtəṛ_ ‘more’ (Heath 2002:140–141).

4.2.7 *_r_ is very often realized as a pharyngealized _ṛ_ (see below, 4.2.10). In some pre-Hilalian dialects (Fes, Tetouan, Chaouen, Taza), *_r_ is realized as _ġ_ (= [ɣ]) or [ʁ] (Aguadé 2003:78; Behnstedt 2003:165).

4.2.8 *_q_: in pre-Hilalian dialects *_q_ is realized either as _q_ or ʾ (= [ʔ]). The pronunciation _q_ is found for instance in Anjra, Tangier, Chaouen, or Tetouan: _qālu_ ‘they said’. The realization ʾ is found in Jewish dialects, Chaouen (female speakers), Fes, Taza, and Tetouan (especially among female speakers), but it is an archaic feature which now tends to disappear (Heath 2002:139, 141; Aguadé 2003:87–88): _ʾāl_ ‘he said’, _ʾbīḥ_ ‘bad, ugly’, _ʾāḍi_ ‘judge’.

In all Hilalian dialects the regular realization of *_q_ is _g_: _gālət_ ‘she said’, _dgīg_ ‘flour’, _bəgṛa_ ‘cow’. However, exceptions to the shift *_q_ > _g_ are numerous: _bqāt_ ‘she remained’, _qlīl_ ‘little’, _qdīm_ ‘old’.

4.2.9 In northern and Jbala dialects there is a phoneme _č_, especially in borrowings from Spanish: _kūčāra_ (< Spanish _cuchara_) ‘spoon’, _plānča_ (< Spanish _plancha_) ‘iron [for pressing]’. In some dialects (in Anjra, for instance), _šāf_ ‘he saw’ > _čāf_ resulting from a generalization of forms with _t_-prefix like _tšūf_ ‘she/you will see’ (Vicente 2000:44).

4.2.10 Pharyngealization of plain consonants is a very common feature in Moroccan (especially in all Hilalian dialects): *_dār_ > _ḍāṛ_ ‘house’, *_raʾs_ > _ṛāṣ_ ‘head’, *_fam_ > _fuṃṃ_ ‘mouth’, *_jarād_ > _žṛāḍ_ ‘grasshoppers’, *_sulṭān_ > _ṣəḷṭān_ ‘sultan’, *_zāra_ > _ẓāṛ_ ‘he visited’. Examples of a phonemic opposition between plain and pharyngealized consonants are _būla_ ‘urine’ vs. _ḅūḷa_ ‘light bulb’, _lbāba_ ‘crumb’ vs. _l-ḅāḅa_ ‘to daddy’, _lla_ ‘no’ (emphasizing) vs. _ḷḷa_ ‘God’.

4.2.11 Labialization of some consonants (_b, m, f, k, g, ġ,_ and _q_) is characteristic for Hilalian dialects: _ḅḅ^{w}a_ ‘daddy’, _ṃṃ^{w}i_ ‘mum’, _ġ^{w}bāṛ_ ‘dung’, _k^{w}bāṛ_ ‘old persons’, _x^{w}rəẓ_ ‘he got out’, _q^{w}təl_ ‘he killed’ (examples from Skura; → labiovelarization).

5. Personal pronouns

5.1 The independent personal pronouns in most dialects are:

	3rd	2nd	1st
masc.	*huwwa*	*nta*	*āna*
fem.	*hiyya*	*nti*	
pl.	*hūma*	*ntūma*	*ḥna*

According to the dialect, there may be some divergences in the singular, but never in the plural. The 1st person singular pronoun is *īna* or *āni* in the Tafilalt, *āni* in Debdou (Heath 2002:269; Behnstedt 2004:56). In northern dialects (Anjra, Tangier, Tetouan, Chaouen), there is no gender distinction for the 2nd person singular, and a common pronoun *ntīn* ~ *ntīna* is used.

A suffix *-ya* is sometimes added to the independent pronouns: *ānāya, ḥnāya*.

5.2 The suffixed pronouns are: Singular: 3rd person masculine *-u* ~ *-h*, 3rd person feminine *-ha* (*-a* in northern dialects), 2nd person *-(ə)k*, 1st person *-i* ~ *-y(a)* ~ *-ni*.

Plural: 3rd person *-hum* (*-həm, -um, -əm* in northern dialects), 2nd person *-kum*, 1st person *-na*. Gender distinction in the 2nd person singular is attested in the dialect of Igli: *šəftək* 'I saw you [masc.]' vs. *šəftki* 'I saw you [fem.]'. This dialect has *-ku* for the 2nd person plural and *-hu* for the 3rd person plural: *ʿəndku* 'you have', *ʿəndhu* 'they have' (Behnstedt 2004:57).

6. Indefinite article

Moroccan dialects have two indefinite articles, *ši* and *wāḥəd* (→ article, indefinite). The indefinite *ši* indicates vagueness or uncertainty, is invariable, and precedes singular or plural substantives: *ši bənt* 'some girl', *ši drāri* 'some children', *ši ṭbīb* 'some doctor', *ši nās* 'some people', *ši ḥāža* 'something'. The specifying article *wāḥəd* (sometimes *wāḥ*) is also invariable and precedes the definite substantive: *wāḥəd lbənt* 'a girl', *wāḥəd ṭṭbīb* 'a doctor', *wāḥəd ṣāḥbi* 'a friend of mine', *wāḥəd wəld ʿəmmi* 'a cousin of mine'.

7. Genitive

The most common genitive markers are *dyāl*/*d-* and *ntāʿ*/*tāʿ*; both markers sometimes have a feminine *dyālt, ntāʿt* and a plural *dyāwl, ntāwʿ* (Heath 2002:461462). The following examples are from Skūra: *žūž dyāl ʿīnīn* 'two eyes', *lwlād dyāwl zzāwya* 'the children of the *zāwya*', *ʿām əlḥəṛka dyālt ṣāġru* 'in the year of the Sagro campaign', *ssəlhām ntāʿ ḅḅwāh* 'the *səlhām* of his father', *lbībān ntāwʿ ḍḍāṛ* 'the doors of the house', *wāḥəd tāssūfra ntāʿt əlmāl* 'a little bag of money' (Aguadé and Elyaacoubi 1995:130).

8. Verbal morphology

8.1 Characteristic for all Moroccan dialects is the prefix *n-* for the 1st person singular of the imperfect (a feature common to all Maghrebi dialects, from Libya to Mauritania, including Maltese and Andalusi Arabic): *kā-nəgləs* 'I am sitting', *nšūf* 'I will see', *nəšri* 'I will buy'.

8.2 The perfect endings in most dialects are:

	3rd	2nd	1st
masc.	*-ø*	*-ti, -t*	*-t*
fem.	*-āt* ~ *ət, -t*		
pl.	*-u, -w*	*-tu* ~ *-tīw*	*-na*

Examples: *ktəb* 'he wrote', *šərbāt* ~ *šərbət* 'she drank', *šəfti* ~ *šəft* 'you [sg.] saw', *fhəmt* 'I understood', *fhəmtu* ~ *fhəmtīw* 'you [pl.] understood', *mṣāw* 'they went', *ktəbna* 'we wrote'.

8.3 The *-t* ending in the 3rd person feminine singular occurs in all dialects after a vowel: *mšāt* 'she went'. The ending *-ət* occurs in Fes, Anjra, Chaouen, Taza: *kətbət* 'she wrote'. The variant *-āt* is secondary (an analogy from the weak verbs) and occurs in Casablanca, Meknes, Marrakech.

8.4 In the 2nd person singular *-ti* occurs in Tangier, Rabat, Casablanca, Marrakech, Fes, Atlantic coast: *ktəbti* 'you [masc./fem.] wrote'. The ending *-t* occurs in Anjra, Chaouen, Jewish dialects: *ktəbt* 'you [masc./fem.] wrote'.

Some dialects, however, distinguish gender in the 2nd person singular: 2nd person masculine singular *-t*, 2nd person feminine singular *-ti*. This is the case in the Jewish dialect of Fes and in the Muslim dialects of the southern oases (Heath 2002:220).

8.5 As to the 2nd person in the plural, the ending *-tu* is the most common. The variant *-tīw* (a secondary form, plural of a singular *-ti*) occurs in Fes and in Jewish dialects of the Atlantic coast. Very interesting is the 2nd person plural ending *-tum* in Anjra: *ktəbtum* 'you wrote', which seems to be of Andalusi origin (Vicente 2000:61–62).

8.6 The most widespread imperfect paradigm is:

	3rd	2nd	1st
masc.	*y(ə)-…-ø*	*t(ə)-…-ø*	*n(ə)-…-ø*
fem.	*t(ə)-…-ø*	*t(ə)-…-i,-y*	
pl.	*y(ə)-…-u, -w*	*t(ə)-…-u, -w*	*n(ə)-… -u, -w*

Examples: *yəktəb* 'he will write', *tkətbi* 'you [fem.] will write', *ykətbu* 'they will write', *nəktəb* 'I will write'; *təqṛa* 'she will read/learn', *təqṛāy* 'you [fem.] will read/learn', *nəqṛāw* 'we will read/learn'; *tḥəll* 'she will open', *nḥəllu* 'we will open'.

An alternation between vowel *ə* in the perfect and *u* in the imperfect stem occurs often, especially in northern dialects: *šəmm* ~ *yšumm* 'to smell', *ḥəll* ~ *yḥull* 'to open', *skət* ~ *yəskut* 'to be silent'.

8.7 Some dialects (Anjra, Chaouen, Tangier, Jewish dialects) do not distinguish gender in the 2nd person singular: *təktəb* 'you (masc./fem.) will write'.

8.8 Stems with initial *ʾ, like **ʾakala* 'to eat' or *ʾaxaḏa* 'to take', are generally reshaped as weak, hollow, or even geminate verbs: *kla* ~ *kāl* ~ *kəll* 'he ate', *xda* 'he took'. Form like *kəl* (Skura) occur only in a very few dialects.

8.9 The passive voice is expressed either by the prefix *t(ə)-/tt(ə)-* or *n(ə)-*. The prefix *t(ə)-* ~ *tt(ə)-* is the most widespread and occurs in almost all Muslim dialects: *ttəktəb* ~ *yəttəktəb* 'to be written', *tfəšš* ~ *yətfəšš* 'to go flat', *tbān* ~ *yətbān* 'to be visible'. Forms with *n(ə)-* occur in eastern Jewish dialects and in some Jbala dialects (Anjra, Chaouen), where they coexist with *t(ə)-* ~ *tt(ə)-* forms: *nčāf* ~ *yənčāf* 'to be visible' (Heath 2002:356).

8.10 *Present and future markers*

kā- and *tā-* are the most common present markers: *kā-nži, tā-nži* 'I come, I am coming', *kā-tākul* 'she is eating', *tā-nšūfu* 'we are looking': in Casablanca it is not unusual for speakers to use indiscriminately both in the same phrase (Aguadé 2003b:304). The marker *dā-*occurs only in some Jewish dialects, *ā-* in dialects of Jbala (Heath 2002:210–211). The marker *lā-* occurs in Chaouen and Anjra (however, in both places *kā-*is also used): *lā-nəqqīwah f-əlma* 'we put it in the water' (Vicente 2000:103).

Future preverbs are *ġādi* (*ġād, ġa*) and *māši* (*māš*). The Jewish dialect of Sefrou uses *ā-*: *ā-nəsmaʿ* 'I will listen' (Stillman 1988:40).

9. ADVERBS

The most common adverbs for 'now' are *dāba* (Jewish, northern and Jbala dialects, Casablanca, Rabat) and *ḏṛūk* ~ *ḏṛūka* (< **hād lwuqt*: Marrakech, southern Atlantic towns, *ḏəṛwək* in Oujda, cf. Heath 2002:452).

For 'today' *lyūm* is predominant: in northern dialects *lyūma* occurs, and Chaouen has *llūma* ~ *llūm*.

'There' is *təmma* in almost all dialects (*fəmma* ~ *ṯəmma* among the Zʿīr); *təmmāk* (in Skūra sometimes *təmmākīna*) is also usual. Widespread through all dialects is *lhīh* 'there.'

The adverb for 'here' is *hna*, with the variant *hnāya*.

10. NEGATION

The most common nonverbal negation is *māši* or *ma…š(i)*: *lbīt māši kbīr, lbīt ma kbīr-š(i)* 'the room is not big', *hīyya māši hna* 'she is not here', *māši f-əḏḏāṛ* 'not in the house'. In Skura, *mīhi* and *mūhu* can be used (as well as *māši*): *mīhi* ~ *mūhu hād əlbāb* 'it is not this house' (Aguadé and Elyaacoubi 1995:147).

The verbal negation is *ma…š*: *ma təmši-š* 'do not go!', *ma kā-yākul-š* 'he is not eating', *ma žāw-š* 'they did not come'.

11. INTERROGATIVES

The interrogatives are *škūn* 'who?', *ăš* ~ *w-āš* 'what?', *āšnu* 'what?', *āš mən* ~ *īna* 'which?', *šḥāl* 'how much?', *kīfāš* 'how?', *fāyn* ~ *fīn* 'where?', *ʿlāšṣ* 'why?', *īmta/fūqāš/wəqtāš* (*fīwāx/fuyāx* in northern and Jbala dialects) 'when?'.

12. CONDITIONAL CLAUSES

Possible conditional clauses are introduced by *īla* (*īda* in some Jewish and northern dialects like Tangier, Tetouan, Anjra): *īla mšīti f-zzənqa* 'if you go by the street', *īla kān ʿəndi lflūs nəšri siyyāra* 'if I have money, I will buy a car' (examples from Casablanca).

Impossible conditional clauses are generally introduced by *kūn, lu kān*, or *lu kūn: kūn ma kānət-š rəžli mṛīḍa nəʿməlha* 'if my foot were not sore, I would do it' (Harrell 1962:169; Heath 2002:490–491).

13. LEXICON

In the lexical domain, the following terms are characteristic for Moroccan dialects: *sārūt* 'key', *məftāḥ* 'needle' (euphemism, to avoid *ībra*), *qniyya* 'rabbit', *ṭūḅḅa* 'rat', *žṛāna* 'frog', *bəbbūš* 'snail', *bībi* 'turkey', *bəllārəž* 'stork', *xīzzu* 'carrots', *ʿāfya* 'fire' (euphemism, to avoid *nāṛ*), *šəṛžəm* 'window', *gāna* 'patience, good humor', *dār/ydīr* 'to make', *dda/yiddi* 'to take away', *nʿəl* 'to curse', *ṣīfəṭ* 'to send', *žūž* 'two', *təsʿūd* 'nine' (euphemism, to avoid *təsʿa* 'you [masc.] will beg'), *lālla* 'lady, Madam', *nīšān* 'straight', *gūd* 'straight', *bəllāti* 'slowly', *bəzzāf* 'much, many, very', *wālu* 'nothing', *wāxxa* 'yes, okay, of course'.

14. BERBER SUBSTRATUM

14.1 About the influence and importance of Berber substratum in Moroccan Arabic, scholars have expressed completely opposite views, ranging from a general assumption of Berber influence to a general denial (Diem 1979:52–53). However, concerning this question it is important to take into account that we still lack comprehensive studies (→ Berber loanwords). One should keep in mind that matters may differ according to the dialect studied: substratal influence is surely less strong in a typical urban dialect like Fes than it is in Skūra, whose inhabitants are completely surrounded by Berber-speaking populations and where the influence of Berber substratum and adstratum is obvious (Aguadé and Elyaacoubi 1995).

14.2 In any case, there are in Moroccan Arabic some features whose Berber origin can hardly be questioned. Among the most common features which usually are explained as a result of Berber substratum, we can mention the following (Diem 1979:52–55; Colin 1986:1196; Lévy 1996:131–136):

i. Reduction of short vowels, especially in open syllables (a characteristic and striking feature of Moroccan Arabic)
ii. Spirantization of the occlusives *k, *t, and *d to the fricatives *x/ç*, *ṯ*, and *ḏ*
iii. Labialization of *k* and *g*
iv. Affrication of *t > t^s
v. Gender and number shifting in some nouns

14.3 Other clear cases of substratal influences are, for instance, the shift *l* > *n* in Igli (see above, sec. 4.2.5), as well as comparative sentences with the preposition *ʿla* instead of *mən*: *ttəlž byəḍ ʿla ṣṣūf* 'snow is whiter than wool' (Aguadé and Vicente 1997).

14.4 Of course it is in the vocabulary that the Berber influence is most evident, and among the most common Berber loanwords in Moroccan dialects maybe quoted: *āgwāl* 'large drum', *āġlāl* 'snail', *ārgān* 'argan tree', *āmrəd* 'crickets', *āṛəẓẓāy* ~ *ṛẓūẓi* 'wasp', *āšku* 'because', *āzāglu* 'yoke', *sārūt* 'key', *ṣīfəṭ* 'to send', *tākāwt* 'gall', *tāta* 'chameleon', *təllīs* (< Latin *trilix*) 'burlap sack, bag', *xīzzu* 'carrot'. Berber loanwords (and, by analogy, some European borrowings) never take the definite article: *ātāy* (< English *tea*) 'tea/the tea' (e.g. *šṛəbt ātāy* 'I drank the tea'), *tāta* 'chameleon/the chameleon' (e.g. *šəfti tāta lli kāyna f-əžžnān*? 'did you see the chameleon in the garden?').

15. SOCIOLINGUISTICS

15.1 Concerning → diglossia in Morocco, the situation is the same as described for → Egypt and other Arabophone countries, i.e., the coexistence of two varieties of the language, a High one for the written variety and a Low one for the spoken vernacular, each of them being used depending on the circumstances (the High variety predominates in formal, the Low one in informal situations). Shifting between registers is not unusual, however.

15.2 Classical Arabic is the only official language of the country, but recently the Moroccan government created a Royal Institute for Berber Studies (IRCAM in its French acronym; see www.ircam.ma) and in 2003 decided that Berber has also to be taught in the schools (this does not mean, however, that in Berber-speaking areas teaching should occur in this language: it is only intended that children learn Berber in the schools). According to the Moroccan Ministry of Education, in the near future, Berber should be taught in every school, not only in regions with a Berber population (Errihani 2006:143). Of course, it is too early to consider whether such a goal is feasible. But in any case, the problems facing such general implementation of Berber are enormous: lack of standardization, resources, and specialized teachers, teaching Berber in a new script (*tifinaġ*) that even teachers have difficulty with, the fact that children will have to learn three different scripts, Arabic, Latin, and Tifinagh, and, especially, lack of interest on the part of non-Berber speakers (Errihani 2006:152).

Bilingualism among Moroccan Berbers is now very common and generally they speak an Arabic dialect in addition to their mother tongue; trilingualism (Berber, Arabic, and French/Spanish) is common (→ language shift: Amazigh).

15.3 Soon after the beginning of the French-Spanish Protectorate in 1912, French became the first foreign language for Moroccans in the Southern Zone: everywhere in the country, new schools arose, with European curricula, and thus French became the key to a modern education and allowed the indigenous elites to acquire Western knowledge. Such schools coexisted with the traditional ones where students learned to read and write and memorize the *Qurʾān*.

In Berberophone areas, the colonial administration tried to strengthen the teaching of French, thereby weakening that of Arabic; in 1930 the *Dahir Berbère* created schools where only French and Berber were taught. The aim of this divide-and-rule policy was to encourage the separation of Arabs and Berbers in order to spread French values and thus thwart nationalist movements, which sought independence (Bentahila 1983:9–10). Moroccan nationalists countered the French educational policy by creating private schools where teaching in Classical Arabic and nationalist ideology was reinforced.

In spite of colonialism, the influence of the French language in Morocco was never as strong as in → Algeria because the Protectorate lasted only 44 years in Morocco (in some southern regions even less), and the number of French settlers and officials established there was far smaller than it was in the case of Algeria. It may seem paradoxical, but it was especially after independence in 1956 that French became widespread as a result of generalization of primary and secondary school teaching in the whole country; this was also stimulated through massive emigration to France, the Netherlands, and Belgium during the 1960s and 1970s. Today, almost all Moroccans are able to speak French, if sometimes only in a rudimentary fashion. Educated people switch immediately to this language when speaking about specialized topics.

There is an important Moroccan literature in French, with internationally known very popular authors like Noufissa Sbaï, Tahar Ben Jelloun, Driss Chraïbi, Lotfi Akalay, and Ahmed Sefrioui.

15.4 Prior to the Protectorate, Spanish was the mother tongue of the Jews and the majority of the Moriscos expelled from the Iberian Peninsula (the Jews in 1492, the Moriscos between 1492 and 1610), who took refuge in Morocco and settled mainly in towns like Tetouan, Fes, Meknes, and Salé. The Moriscos Arabized early on, but the Jewish communities retained their archaic Spanish until the 19th century.

With the establishment of the Protectorate, Spanish became the language of the administration in the Northern Zone as well as in the International Zone of Tangier and thus the second language of an important part of the native population. Spanish was taught in the schools, and some schoolboys went to Spain to continue their studies in Spanish universities. The Spanish administration (unlike the French one, and fearing the urban population less than the Rifian peasants) promoted the teaching of Classical Arabic, not only in the Arabic-speaking areas but also among Berbers in the Rif, who had previous revolted against Spanish colonial rule, and then even imported Arabic teachers from Lebanon.

After the end of the Protectorate and the independence of Morocco in the year 1956, the use and knowledge of Spanish decreased rapidly. However, in the last two decades of the 20th century Spanish partially recovered its former importance due to wide reception of Spanish television and radio channels (especially in the north but also in the south due to the proximity of the Canary Islands), emigration to Spain, and the increasing number of Spanish tourists who visit the country each year.

In northern Morocco there is today a small Spanish literature. Moroccan authors writing in Spanish are, for instance, Mohamed Lahchiri, Abderrahman El Fathi, Mohamed Sibari, and Mohamed Bouissef Rekab.

15.5 Today English and Spanish have become the preferred second languages in Morocco and are taught in secondary schools as well as in American or Spanish cultural centers and schools. In the year 1993, the Al Akhawayn University was founded in Ifrane. It is an elite academic institution where instruction is in English or Arabic. Wealthy Moroccans now prefer to study in the United States in order to have access to the latest knowledge and technology (Ennaji 2005:115). German is the third foreign language and is pursued mainly by Moroccans dealing with tourism as well as by students or scholars who intend to study in Germany or Austria.

There is a Hebrew school in Casablanca, belonging to the Moroccan Jewish minority, where Hebrew is taught to Jewish and Muslim children.

15.6 The Moroccan Radio and Television channels broadcast daily news and cultural programs (sometimes also films) in the three main Berber dialects. News is broadcast in French and Spanish as well. Arabic is the main language in all the Moroccan TV channels, but French continues to play an important role. According to official data, in 2005 the first channel of Moroccan TV (*at-Talfaza al-Maġribiyya*) broadcast 80 percent of its programs in Arabic (Classical as well as Moroccan, Egyptian, or other Arabic dialects), 16.4 percent in French, 1.6 percent in Berber, and 1.4 percent in Spanish. However, most Arabic/French bilingual Moroccans prefer the semiprivate second channel 2M, which broadcasts more than 75 percent in French, because of its better quality and more interesting programs (Ennaji 2005:104–105).

Due to the abundance of Egyptian films and songs on Moroccan TV or radio, Egyptian Arabic is now generally understood, especially by young people. Some Moroccan singers prefer to sing in Egyptian in order to become well known in other Arabic countries.

15.7 Concerning the press, Moroccan newspapers have to face serious competition with the French (*Le Figaro, Le Monde*) and the Saudi Arabian press (*aš-Šarq al-ʾAwsaṭ*). For decades the majority of the Moroccan newspapers, some of them published in both Arabic and French, belonged to political parties and generally had small print runs. In the last decade, a new independent press (newspapers, tabloids, and magazines) has begun to arise: the number of newspapers and tabloids in Arabic is slightly higher than those in French. On the other hand, the most important magazines are in French (Ennaji 2005:103–104).

15.8 Publications in dialect are very scarce in Morocco because almost all authors prefer to write either in Classical Arabic or in French. Only a few authors, e.g. Yūsəf Fāḍəl, əz-Zubīr Bən Būṣtā, ʿAbd əṣ-Ṣamād əl-Kənfāwī, and Yūsəf ʾAmīn əl-ʿAlamī, have written some small pieces (mainly theater) in dialect. However, in many novels written in Classical Arabic, the authors use the dialect in all the dialogues, looking for more realism.

Other works in vernacular are, of course, compilations of dialectal poetry (*məlḥūn*) and proverbs and sayings.

15.9 Immediately after independence in 1956, a very important effort was made to decrease the influence of French (Spanish in the Northern Zone) and to Arabize the educational system and administration. Concerning the attitude of political parties to the Arabization policy, two main groups can be distinguished: supporters of bilingual education, and followers of a more radical Arabization. Thus, the Moroccan policy in this field was marked by contradictions and inconsistencies depending on the changing influence of either tendency in each government.

The process of Arabization is even today far from being completed. It has been almost completed for primary and secondary schools, but in the universities scientific matters like medicine, biology, and engineering are still taught in French due to the lack of an adequate terminology. In business, the banking sector, and commercial transactions, French continues to be the preferred language: for many Moroccans, French is the language of social success (Ennaji 2005:41).

BIBLIOGRAPHICAL REFERENCES

Aguadé, Jordi. 1998. Un dialecte maʿqilien: Le parler des Zʿīr au Moroc. *Peuplement et arabisation au Maghreb occidental: Dialectologie et histoire*, ed. Jordi Aguadé, Patrice Cressier, and Ángeles Vicente, 141–150. Madrid and Saragossa: Casa de Velázquez and Universidad de Zaragoza.

——. 2003a. "Estudio descriptivo y comparativo de los fonemas del árabe dialectal marroquí". *Estudios de Dialectología Norteafricana y Andalusí* 7.59–109.

——. 2003b. "Notes on the Arabic dialect of Casablanca (Morocco)". *AIDA 5th Conference. Proceedings. Cádiz, September 2002*, 301–308. Cádiz: Universidad de Cádiz.

——, and Mohammad Elyaacoubi. 1995. *El dialecto árabe de Skūra* (*Marruecos*). Madrid: Consejo Superior de Investigaciones Científicas.

Aguadé, Jordi and Ángeles Vicente. 1997. "Un calco semántico del bereber en árabe dialectal magrebí: El uso de la preposición *ʕla* en el comparativo". *Estudios de Dialectología Norteafricana y Andalusí* 2.225–240.

Behnstedt, Peter. 2002. "La frontera entre el bereber y el árabe en el Rif". *Estudios de Dialectología Norteafricana y Andalusí* 6.7–18.

——. 2003. "Zwei Texte im Altstadt-Dialekt (T^{3}) von Taza (Marokko)". *Estudios de Dialectología Norteafricana y Andalusí* 7.165–171.

——. 2004. "Von an-ʾĀṣəṛ (al-Qaṣr) nach Īgni (īgli): Ein Vorbericht zu einigen arabischen Dialekten der Provinz əṛ-Rašīdīya (Maroko)". *Approaches to Arabic dialects: A collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday*, ed. Martine Haak, Rudolf de Jong, and Kees Versteegh, 47–65. Leiden: E.J. Brill.

——. 2005. "Materialien für einen Dialektatlas von Nordost-Marokko. I. Tiernamen, Teil 1: Einleitung, Transkription, Ortsnamenverzeichnis, Karten". *Estudios de Dialectología Norteafricana y Andalusí* 9.7–72.

——, and Mostafa Benabbou. 2002. "Zu den arabischen Dialekten der Gegend von Tāza (Nordmarokko)". "*Sprich doch mit deinen Knechten aramäisch, wir verstehen es!": 60 Beiträge zur Semitistik. Festschrift für Otto Jastrow zum 60. Geburtstag*, ed. Werner Arnold and Hartmut Bobzin, 53–72. Wiesbaden: O. Harrassowitz.

——, and Mostafa Benabbou. 2005. "Données nouvelles sur les parlers arabes du Nord-Est marocain". *Zeitschrift für Arabische Linguistik* 44.18–70.

Bentahila, Abdelâli. 1983. *Language attitudes among Arabic-French bilinguals in Morocco*. Clevedon: Multilingual Matters.

Brunot, Louis. 1931–1952. *Textes arabes de Rabat*. I. *Textes, transcription et traduction annotée*. II. *Glossaire*. Paris: Publications de l'Institut des Hautes Études Marocaines.

Caubet, Dominique. 1993. *L'arabe marocain*. 2 vols. Paris and Louvain: Peeters.

Colin, Georges Séraphin. 1921. "Notes sur le parler arabe du nord de la région de Taza". *Bulletin de l'Institut Français d'Archéologie Orientale* 18.33–119.

——. 1955. *Chrestomathie marocaine: Textes citadins en transcription latine*. Paris: Adrien–Maisonneuve.

——. 1986. "al-Maghrib. VII. Aperçu linguistique". *Encyclopédie de l'Islam*. V, 1193–1198. 2nd ed. Leiden and Paris: E.J. Brill.

Destaing, Edmond. 1937. *Textes arabes en parlers des Chleuhs du Sous (Maroc)*. Paris.

Diem, Werner. 1979. "Studien zur Frage des Substrats im Arabischen". *Der Islam* 56.12–80.

Dombay, Franz von. 1800. *Grammatica linguae Mauro-Arabicae juxta vernaculi idiomatis usum: Accessit vocabularium Latino-Mauro-Arabicum. Opera et studio Francisci de Dombay*. Vindobonae: Apud Camesina.

Durand, Olivier. 1994. *Profilo di arabo marocchino: Varietà urbane centromeridionali*. Rome: Università degli studi 'La Sapienza'.

Ennaji, Moha. 2005. *Multilingualism, cultural identity, and education in Morocco*. New York: Springer.

Errihani, Mohammed. 2006. "Language policy in Morocco: Problems and prospects of teaching Tamazight". *Journal of North African Studies*. 11:2.143–154.

Harrell, Richard Slade. 1962. *A short reference grammar of Moroccan Arabic*. Washington, D.C.: Georgetown University Press.

——. 1966. *A dictionary of Moroccan Arabic: Moroccan-English*. Washington, D.C.: Georgetown University Press.

Heath, Jeffrey. 2002. *Jewish and Muslim dialects of Moroccan Arabic*. London and New York: Routledge Curzon.

Ishihara, Tadayoshi. 2000. *Morokko arabiago: Kaiwa to bunpo* [Moroccan Arabic: Conversation and grammar]. Tokyo: Daigakushorin.

Kampffmeyer, Georg. 1912. *Marokkanisch-arabische Gespräche im Dialekt von Casablanca, mit Vergleichung des Dialekts von Tanger*. Berlin: Reimer.

Lerchundi, José de. 1872. *Rudimentos del árabe vulgar que se habla en el Imperio de Marruecos*. Madrid.

——. 1892. *Vocabulario español-arábigo del dialecto de Marruecos, con gran número de voces usadas en Oriente y en la Argelia*. Tangier: Misión Católica Española.

Lévi-Provençal, Evariste. 1922. *Textes arabes de l'Ouargha, dialecte des Jbala (Maroc septentrional).* Paris: Publications de l'Institut des Hautes Études Marocaines.

Lévy, Simon. 1996. "Repères pour une histoire linguistique du Maroc". *Estudios de Dialectología Norteafricana y Andalusí.* 1.127–137.

——. 1998. "Problématique historique du processus d'arabisation au Maroc: Pour une histoire linguistique du Maroc". *Peuplement et arabisation au Maghreb occidental: Dialectologie et histoire,* ed. Jordi Aguadé, Patrice Cressier, and Ángeles Vicente, 11–26. Madrid and Saragossa: Casa de Velázquez.

Loubignac, Victorien. 1952. *Textes arabes des Zaër: Transcription, traduction, notes, et lexique par Victorien Loubignac.* Paris: Librairie Orientale et Américaine.

Marçais, William. 1911. *Textes arabes de Tanger.* Paris: Bibliothéque de l'École des Langues Orientales Vivantes.

Mercier, Henry. 1951. *Dictionnaire arabe-français.* Rabat: Editions La Porte.

Moscoso, Francisco. 2003. *El dialecto árabe de Chauen (N. de Marruecos): Estudio lingüístico y textos.* Cádiz: Universidad de Cádiz.

Prémare, Alfred-Louis, de a.o. 1993–1999. *Dictionnaire arabe-français: Langue et culture marocaines.* 12 vols. Paris: l'Harmattan.

Singer, Hans Rudolf. 1958a. "Neuarabische Texte im Dialekt der Stadt Tetuan". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 108.106–125.

——. 1958b. "Grundzüge der Morphologie des arabischen Dialekts von Tetuan". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 108.229–265.

Stillman, Norman A. 1988. *The language and culture of the Jews of Sefrou, Morocco.* Manchester: University of Manchester.

Vicente, Ángeles. 2000. *El dialecto árabe de Anjra (norte de Marruecos): Estudio lingüístico y textos.* Saragossa: Universidad de Zaragoza.

Youssi, Abderrahim. 1992. *Grammaire et lexique de l'arabe marocain moderne.* Casablanca: Wallada.

JORGE AGUADÉ (University of Cádiz)

Morphology

Morphology is the study of word formation. Arabic has a rich morphological system with many interesting properties. This entry begins in Section 1 with a synopsis of some important morphological concepts. Sections 2 and 3 describe the morphology of Arabic verbs and nouns. Section 4 is an overview of recent research on morphology in Arabic and other Semitic languages.

1. MORPHOLOGICAL CONCEPTS

Words are composed of meaningful units called morphemes. The English word *obese* consists of a single morpheme, while *obesity* has two morphemes, the root *obese* and the affix *+ity*, which changes adjectives into nouns that refer to qualities. Every language has a large number of root morphemes, typically at least one thousand, and often many more. Furthermore, languages are always able to add to their stock of root morphemes by coining new ones or by borrowing words from other languages. On the other hand, languages typically have no more than about one hundred affixal morphemes, and often many fewer. Affixes are rarely borrowed from other languages.

Most affixes in most languages are prefixes or suffixes, like *pre+* in *preboard* or *+ity* in *obesity*. Infixes, which go inside a root, are rarer; an example is *+um+* in Tagalog *grumadwet* 'graduated'. Circumfixation is discontinous affixation, in which a prefix and suffix combine to mark a single morphological distinction (e.g. German *ge+sag+t* 'said'). → Clitics are a special type of affix; to the syntax, clitics seem like separate words, but to the phonology, they seem like affixes. (The contracted form of *not* in English *wasn't* is a clitic.) Many languages mark morphological distinctions by copying all or part of a word. This is called → reduplication, as in Ilokano *jan+jánitor* 'janitors'. Ablaut (or → apophony) is the term used for morphology that involves vowel changes, as in English *sing/sang/sung*. Occasionally, morphology is subtractive, removing part of a word to change its meaning or usage. An example is the formation of English nicknames, like *Benjamin* ⇒ *Ben* (→ truncation).

Morphological processes are traditionally classified as inflectional, like the *+s* suffix in English *gets*, or derivational, like the *+ity* suffix. This distinction is not uncontroversial, but it usefully describes a loose correlation of properties. Inflectional morphology is usually fully productive (→ inflection), but derivational morphology need not be (→ derivation). (Morphology is productive if it applies even to newly borrowed or coined words.) Inflectional morphology has compositional meaning, but derivational morphology often does not. Meaning is compositional when the meaning of the whole equals the sum of the meaning of its parts; an example of noncompositional meaning is English *electricity*, which does not mean 'the quality of being electric' (cf. *obesity*). Inflectional morphology marks syntactically relevant distinctions and never changes a word's part of speech, whereas derivational morphology marks lexically relevant distinctions and may change the part of speech (e.g., *obese* is an adjective and *obesity* is a noun).

For more information about morphology and morphological theory, Spencer (1991) or Spencer and Zwicky (1998) may be consulted.

2. VERB MORPHOLOGY

The Arabic → verb has a complex morphological structure. Its many parts are best understood by peeling back the layers of morphology from the outside in. The diagram in (1) is the starting point, using the word /wa+ya+ktub+ū+na+hu/ 'and they [masc.] are writing it [masc.]'.

(1) Schema for verb
/wa ya ktub
proclitic infl. prefix verb stem
ū+na hu/
infl. suffixes enclitic

In addition, grammatical voice is shown by changes in the vowels of the inflectional prefix and the verb stem. Like all other aspects of the verb's morphological structure, this is discussed in detail below.

By convention, proclitics are written as a single word with the following verb if they consist of a single consonant and a vowel, like /wa+/ 'and', /fa+/ 'then', and /li+/ 'for, that'. Their meanings are typical for conjunctions and prepositions. The enclitic system is reserved for pronominal objects of the verb, like /+hu/ 'him'. They are distinguished for person, number, and gender, e.g. /+ka/ 'you [masc. sg.]', /+ki/ 'you [fem. sg.]', /+kumā/ 'you [com. dual]', /+kum/ 'you [masc. pl.]', and /+kunna/ 'you [fem. pl.]'. Verbs that take both an indirect and a direct object can bear two enclitics in that order: /ʔa+ʕṭā+nā+hu/ 'he gave us [/+nā/] it [/+hu/]'. In many modern colloquial varieties of Arabic, this system has been extended to include enclitics that combine /+l+/ 'to, for' with a pronoun, e.g. Maltese /kitib+hu+l+na/ 'he wrote it [/+hu/] for [/+l+/] us [/+na/]'.

The inflectional prefixes are part of the system of subject → agreement. In general, the inflectional morphology of the verb is determined by the person, number, and gender of the verb's subject. The Arabic verb has two aspects, perfect, which refers to completed actions or events, and imperfect, which refers to incomplete actions or events. Different morphemes are used for subject agreement in the perfect and imperfect.

Subject agreement in the perfect involves suffixes only. There are thirteen different agreement suffixes in the perfect, ranging from 3rd person masculine singular /+a/ through 2nd person dual /+tumā/ to 1st person plural /+nā/. There are various partial resemblances among these suffixes, suggesting that fewer than thirteen actual morphemes are involved. For example, all of the 2nd person suffixes contain /t/, and all of the dual suffixes contain /ā/.

Subject agreement in the imperfect involves a kind of circumfixation: prefixes and suffixes combine to mark the relevant distinctions. For example, /ta+ktub+ī+na/ means 'you [fem. sg.] are writing', combining the /ta+/ prefix of the 2nd person with the /+ī+na/ suffix of the 2nd person feminine singular. The prefixing part of this system is simple: /ta+/ throughout the 2nd person and in the 3rd person feminine singular and dual, /ya+/ in all other 3rd person forms, and /ʔa+/ and /na+/ in the 1st person singular and plural, respectively. The suffixing part of the imperfect subject agreement system is more complex, however, because it is tied up with the marking of verbal → mood, which depends on how a verb is used in a sentence. The suffix /+ī+na/ consists of two morphemes: /+ī/ and /+na/. The /+ī/ part is found in all 2nd person feminine singular verbs, regardless of their mood, but the /+na/ part is limited to the indicative mood and is absent in the subjunctive and jussive moods: /ta+ktub+ī/ '(that) you [fem. sg.] are writing' or 'may you [fem. sg.] be writing!'. This special indicative morpheme appears only after the agreement suffixes that end in long vowels – /+ī+na/, /+ā+ni/ [dual], and /+ū+na/ [2nd, 3rd pers. masc. pl.] – and the height of its vowel is dissimilated from the height of the preceding vowel. (Other suffixes with /n/ mark the two energetic [→ energicus] moods of Classical Arabic: /ya+ktub+a+n/, /ya+ktub+a+nna/ 'he really is writing'.)

The jussive and imperative moods are related both functionally and formally. The imperative is used for commands and the jussive for exhortations. (The jussive is also required in certain negative and conditional contexts.) Like the subjunctive, the jussive lacks the /+na/ and /+ni/ suffixes of the indicative mood. The jussive also lacks the /+u/ suffix of the singular and 1st person plural indicative; for example, /ya+ktub/ 'may he write!' and /na+ktub/ 'let us write!' are typical jussives (cf. indicative /ya+ktub+u/, /na+ktub+u/). It is not quite accurate to say, however, that the /+u/ suffix is absent from the jussive. Rather, this suffix has been removed by a process of morphological subtraction (also called → truncation; for relevant discussion, see Aronoff 1976; Brame 1970; Levy 1971; Prince 1975; Weeda 1992).

The nature of this subtraction process is clear from the jussives of so-called → weak verbs, whose final consonant (originally /w/ or /y/) is absent for phonological reasons. For example, the indicative verb /ya+rmī/ 'he throws' has a corresponding jussive form /ya+rmi/ 'let him throw!', in which the final vowel has been shortened. Similarly, the jussive of /ya+rḍā/ 'he is satisfied' is /ya+rḍa/ 'may he be satisfied!'. In Arabic, as in many languages, there is a phonological equivalence between a single long vowel and two short vowels. Shortening a final long vowel, as in /ya+rḍā/ → /ya+rḍa/, is therefore equivalent to deleting a final short vowel, as in /ya+ktub+u/ → /ya+ktub/. This truncation process is the characteristic morphology of the jussive.

The imperative is identical to the 2nd person jussive, except that the inflectional prefixes are also truncated: /ktub/ 'write [masc. sg.]!', /ktub+ī/ 'write [fem. sg.]!', etc. (In Classical and Modern Standard Arabic, as well as some colloquial varieties, these initial clusters require an epenthetic vowel to be pronounceable: [ʔuktub].) That imperatives involve morphological subtraction and not simply the absence of an affix is shown by looking at weak verbs of another type, those that have /w/ as their initial consonant. The perfect verb preserves the /w/ – e.g. /waʿada/ 'he made a promise' – but the /w/ is deleted for phonological reasons after the imperfect prefixes: /ta+ʿid+u/ 'you [masc. sg.] make a promise'. The imperative is /ʿid/ 'promise [masc. sg.]!', with the /w/ also missing, so it must be derived from the imperfect verb by truncating the prefix. The imperative, then, involves double truncation, subtracting both the inflectional suffix /+u/, as in the jussive, and the inflectional prefix /ta+/.

We have now stripped back all of the outer layers of the Arabic verb in (1) to reveal its inner core, the → stem. The stem has considerable internal structure, and a good case can be made that several distinct morphemes are combined to form a verbal stem. The key observations are these:

i. In many languages, there are no limitations on the length or structure of stems except those that are required for pronounceability. In English, for instance, the noun *telephone* can also be used as a verb *to telephone* without difficulty. But when this same noun was borrowed into Arabic as /tilifūn/, it could only become a verb by adopting the disyllabic stem /talfan/.
ii. Arabic verbs can change their meaning by changing their syllable structure. For example, a verb stem like /katab/ 'to write' may take on a causative meaning by doubling the middle consonant (/kattab/ 'to make someone write') or a reciprocal meaning by lengthening the first vowel (/kātab/ 'to correspond with someone').
iii. As in many languages, the Arabic verb stem may also contain affixes. For example, prefixing /ta+/ to the /kātab/ form contributes a reflexive meaning: /ta+kātab+ū/ 'they corresponded with one another'. (This /ta+/ should not be confused with the homophonous prefix of the subject agreement system.)
iv. The vowels of the verb stem are not freely chosen. For example, causative verbs of the /kattab/ type always have the vowel /a/ in both stem syllables of the perfective active and the vowels /i/ and /u/ in the same syllables of the perfect passive: /kuttib/ 'to be made to write'.

Points (i)–(iii) fall naturally together and so they will be addressed first, with point (iv) taken up later.

Every verb fits into one of a small number of patterns or verbal Forms (called *ʾawzān* 'measures' [sg. *wazn*] in Arabic). A pattern or measure has a characteristic syllabic structure and may also include an affix. Most verbal Forms also

Table 1. Principal verb Forms

Pattern	Western number	Syllable structure	Affixes	Approximate meaning	Remarks
faʿal	I	CVCVC			The 'basic' pattern, with no characteristic meaning
faʿʿal	II	CVCCVC		Causative	Middle consonant is doubled
tafaʿʿal	V	ta+CVCCVC	Prefix *ta*+	Reflexive of II	
fāʿal	III	CV:CVC		Reciprocal	First vowel is lengthened
tafāʿal	VI	ta+CV:CVC	Prefix *ta*+	Reflexive of III	
ʾafʿal	IV	CVCCVC	Prefix ʾ+	Transitive	
nfaʿal	VII	CCVCVC	Prefix *n*+	Intransitive	
ftaʿal	VIII	CCVCVC	Infix +*t*+	Reflexive	Has infix +*t*+ after first consonant
stafʿal	X	CCVCCVC	Prefix *st*+	Reflexive	
fʿalal	IX	CCVCVC		Color or bodily defect	Final consonant is reduplicated

make a contribution to the meaning of the verb. The patterns /kattab/, /kātab/, and /ta+kātab/ exemplify three different verb Forms.

Any verb is in principle capable of assuming fifteen different Forms in Classical Arabic, although six of the Forms are significantly rarer than the others (Fleisch 1968:124–126). In the Western nomenclature, the patterns are identified by the Roman numerals I–XV; in the Arabic tradition, they are denoted by examples formed on the verb /faʿal/ 'to do'. As it happens, no verb is attested in all fifteen Forms – for example, /katab/ is found with seven different Forms, according to Wehr (1971). Nonetheless, it is standard practice to exemplify the Forms using the citation verb /faʿal/ as if it were actually attested in all of them. Table 1 provides information about all of the common Forms and the best-attested rare one, Form IX. (The column headed 'Syllable structure' is explained below.)

Form I of the verb is basic in two senses: it has no added material like affixes, doubling, or lengthening; and it contributes no characteristic meaning to the verb. It is not basic in any other sense; for example, it is not the case that every verb must appear in Form I; indeed, 16 percent of the verbs in Wehr (1971) are not found in Form I. Furthermore, it is Form II rather than Form I that is typically used for newly borrowed verbs like /barrak/ 'to park a car'.

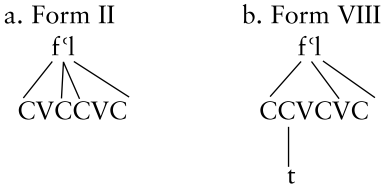
Forms II and V stand in an obvious morphological relationship to one another, as do Forms III and VI. Adding the derivational prefix /ta+/ to Forms II and III makes them reflexive. The /t/ of this prefix is related historically to the affixal /t/ of Forms VIII and X. In Form VIII, this /t/ is actually an infix, falling immediately after the first consonant of the root.

Form IX is uncommon, with just eighteen examples in Wehr (1971), but that is because it is limited to verbs that mean 'to be or become the color x' or 'to have or acquire the bodily defect y', such as /swadad/ 'to be or become black' or /ḥwalal/ 'to have a squint, be cross-eyed'.

The verbal Forms are a type of derivational rather than inflectional morphology. One reason for saying this is that they are not fully productive. Speakers of Arabic are not at liberty to impose any pattern on any verb, although typically they can recognize and understand in context a verb form they have never seen before if they are acquainted with the same verb in a different pattern. Another reason that the verb patterns must be derivational is that their meaning is not fully compositional. The brief explanations of meaning given in Table 1 are far from being comprehensive. For example, while Form II often conveys a causative meaning, it can also act as an intensifier, as in the pair /kasar/ 'to break' and /kassar/ 'to smash'. Often, a verb pattern will acquire some idiosyncrasies of meaning or usage, much like the English *electricity* example cited earlier.

From the perspective of formal linguistics, particularly phonological and morphological theory, the most interesting thing about the Arabic verb patterns is their fixed syllabic structure. The 'Syllable structure' column in Table 1 schematizes the syllable structure of each pattern using the symbols 'C' for consonant and 'V' for vowel. (The V symbol is doubled to indicate a long vowel.) It has been proposed that these schematizations are morphemes of Arabic (McCarthy 1979, 1981). Each verb pattern specifies one of these morphemes, called a CV-skeleton. The consonants of the verb are associated with C positions in the skeleton according to the principles of autosegmental phonology (see Goldsmith 1976a, 1976b, or consult any post-1980 phonology textbook). An affix, such as the *+t+* infix, may also be associated with a C-slot. The autosegmental diagrams in (2) show how Forms II and VIII are represented in this system:

(2)


a. Form II
b. Form VIII

Example (2a) represents the stem /faʿʿal/. The consonants of this verb are linked to C positions of the template, with the medial consonant /ʿ/ occupying two such positions – and therefore doubled in pronunciation. Example (2b) represents /ftaʿal/, with the infix /+t+/ separated from the rest of the consonants in the stem.

Verb stems with different patterns are related to one another if they have the same consonantal root, such as *f-ʿ-l* in (2) or *k-t-b* in the pair /katab/ – /kattab/. The mapping of a consonantal root onto a CV-skeleton, as in (2), is therefore the basis of Arabic verbal derivation. The roots are morphemes, as are the CV-skeleta, in what is known as root-and-pattern morphology. In terms of linguistic theory, root-and-pattern morphology is a special case of templatic morphology. A template is a specification of the syllabic structure of certain morphemes, stems, or words. The Arabic CV-skeleta are templates that apply to the entire stem; in Ilokano reduplication (recall the example *jan+janitor*), a CVC template applies to the reduplicative prefix (Marantz 1982).

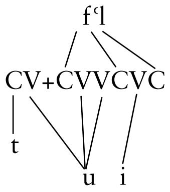
The consonantal → root is basic to the organization of the Arabic mental dictionary, called the lexicon (see Sec. 4). Words that have nothing in common with one another except a shared root will typically also share the meaning of that root. This is obviously true for verb patterns like /katab/ and /kattab/, but it is also true for nouns with very different patterns, such as /kitāb/ 'book', /kutayyib/ 'booklet', /makātib/ 'offices', /stiktāb/ 'dictation', /kātib/ 'writer', and so on. Words of diverse phonological shapes, sharing nothing except the consonant sequence *k-t-b*, all include the notion of 'writing' in their meaning.

In addition to triconsonantal or triliteral roots like *k-t-b* and *f-ʿ-l*, Arabic also has roots with four consonants, called quadriliterals. These roots fall into two classes, simple like /daḥraj/ 'to roll something' and reduplicated like /waswas/ 'to whisper; to tempt; to arouse scruples'. Quadriliteral roots are not common – they constitute just 6 percent of all the verbal roots in Wehr (1971), roughly evenly divided between plain and reduplicated – but the class of plain quadriliterals is clearly productive, since new borrowings readily enter it: /talfan/ 'to telephone', /talfaz/ 'to televise'. Quadriliterals have a detransitivizing pattern with prefixed /ta+/ like the triliteral Forms V and VI: /tadaḥraj/ 'to roll along', /tawaswas/ 'to have scruples'. They also have two other patterns, both rare. Only one is attested in Modern Standard Arabic, with eight exemplars in Wehr (1971): /dlahmam/ 'to be dark' from the root *d-l-h-m* (cf. Form IX, also with a reduplicated final consonant).

Besides the CV-skeleton or template, the consonantal root, and the affixes, there is a fourth piece of the verb stem: the vowels. In (2), the V positions are shown as empty, but in a more complete representation they must be associated with vowels. These vowels are supplied by morphemes. This is clearest in the → passive voice. In the perfect passive, the last vowel of the stem is /i/ and the preceding vowels are all

/u/ (long or short, depending on the pattern): /fuʕil/, /fuʕʕil/, /tufuʕʕil/, /fūʕil/, /tufūʕil/, and so on. In line with the overall root-and-pattern analysis, the vowel sequence /u-i/ is a morpheme that marks the perfect passive. The representation of /tufūʕil/ is given in (3).

(3) Form VI passive



Similarly, the imperfect passive is indicated by the morphemic vowel sequence /u-a/, with the /u/ realized on the imperfect prefix and the /a/ filling all of the V positions of the verb stem proper.

Except for Form I, the vowels of the perfect active are /a/ across the board, indicating that this vowel is a perfect active morpheme. The vowels of the imperfect active are not quite so predictable, although there are significant regularities. The patterns with prefixed /ta+/, namely Forms V and VI, have /a/ across the board in the imperfect as well. And the patterns whose skeleta begin with a CC sequence, namely Forms VII–X, select an /a-i/ imperfect active morpheme.

It is not uncommon for languages to exhibit irregular behavior only in the simplest and most common words; English strong verbs like *see/saw* and *sing/sang/sung* are typical examples. It is therefore no surprise that Form I in Arabic has certain idiosyncrasies in its vowels. All Form I active verbs have /a/ in the first syllable of both perfect and imperfect, but they differ in the vocalization of the second syllable. Table 2 summarizes the data.

Table 2. Ablaut in Form I of the verb

Example	Gloss	Frequency in Wehr (1971)	Remarks
katab/yaktub	'to write'	34%	
ḍarab/yaḍrib	'to beat'	28%	
šarib/yašrab	'to drink'	17%	
faʕal/yafʕal	'to do'	15%	C_2 or C_3 is a guttural consonant
balud/yablud	'to be stupid'	6%	Adjectival meaning

The two least common classes of verbs in Table 2 have no alternation in vowel quality between the perfect and imperfect. These two classes are also distinguished on phonological or semantic grounds: in /faʕal/ – /yafʕal/, the vowel remains fixed as /a/ for phonological reasons, because the adjoining second or third consonant of the root is an /a/-like guttural consonant; and in /balud/ – /yablud/, the verb has an adjectival or stative meaning.

As for the three most common classes of verbs in Table 2, it is generally impossible to predict which class a particular verb will belong to, so this information must be stored in the lexicon. There is, however, some predictability in the vowel alternation: a high vowel /u/ or /i/ in the imperfect is matched by the low vowel /a/ in the perfect, while low /a/ in the imperfect is matched by high /i/ in the perfect. This height-switching ablaut pattern has provoked discussion in the phonological literature (e.g. by Anderson and Browne 1973; Brame 1970; Chomsky and Halle 1968; Guerssel and Lowenstamm 1996; → apophony).

3. NOUN MORPHOLOGY

This section does for the noun what Section 2 did for the verb. Like the verb, the Arabic noun has an outer layer of affixing inflection and an inner core of templatic morphology. Affixation in the noun is much simpler than in the verb, but the templatic morphology of the noun presents complications that are not found in the verb. A particular focus of attention is the system of 'broken' plurals (→ number).

The basic schema for an Arabic noun is given in (4), which may be compared to the verb schema in (1).

(4) Schema for noun
/li walad i
proclitic noun stem infl. suffix
ka/
enclitic
'to your [masc. sg.] boy'

In addition, the feminine suffix /+at/ falls between the stem and the inflectional suffix: /li+malik+at+i+ka/ ‘to your [masc. sg.] queen’.

Among the proclitic elements are CV prepositions like /li+/, the definite article /al+/, and combinations of the two, as in /li+(a)l+waladi/ ‘to the boy’. The enclitics of the noun are possessive pronouns. They are identical to the pronominal enclitics of the verb, except that 1st person singular is rendered as /+nī/ in verbs and /+ī/ in nouns.

In Classical and Modern Standard Arabic, the inflectional suffixes of the noun mark case and sometimes definiteness, construct state, and number. Most nouns are ‘triptotes’, which means that they have three distinct case forms, nominative /walad+u/, genitive /walad+i/, and accusative /walad+a/. Triptotic nouns are marked for indefiniteness by adding the suffix /+n/, called → ‘nunation’ or → *tanwīn*, after the case ending: /walad+u+n/ ‘a boy [nom. indef.]’. For phonological and grammatical reasons, some nouns are diptotic, with syncretism of the genitive and accusative (→ diptosis), and a few are monoptotic or indeclinable: /ʿuṯmān+u/ ‘Othman [nom.]’, /ᶜuṯmān+a/ ‘Othman [gen./acc.]’; /dunyā/ ‘world [nom./gen./acc.]’.

Number inflection in Arabic nouns is a complex matter, and it has proven to be of great relevance to questions of phonological and morphological theory. There are two basic modes of plural formation, the suffixing or ‘sound’ plural and the templatic or ‘broken’ plural. (Whether a noun takes a sound or broken plural depends on its stem, explained below.) The sound feminine plural lengthens the vowel of the feminine suffix /+at/: /malik+at+u+n/ ‘queen [nom. indef.]’, /malik+āt+u+n/ ‘queens [nom. indef.]’. The suffixes of the sound masculine plural are /+ū/ in the nominative and /+ī/ in the genitive/accusative. To these suffixes /+na/ is added if the noun is not in the → construct state (i.e. if it is not immediately followed by a pronominal suffix or nominal possessor): /al-muʾmin+ū+na/ ‘the believers [nom.]’, /muʾmin+ū/ ‘believers of’. There is in addition a dual number that is marked by the suffixes /+ā+ni/ in the nominative and /+ay+ni/ in the genitive/accusative.

Before considering the broken plural, it is necessary to examine the system of noun stems in the language. Arabic noun stems fall into two basic classes that can be called canonical and noncanonical (Levy 1971; McCarthy and Prince 1990b). Canonical noun stems are analogous to verbs in that they are limited in their syllable structure: canonical nouns are never longer than two syllables and must begin with a CV sequence. Nearly all native, basic nouns are canonical, as are many borrowings and coinages; some examples are given in Table 3.

Table 3. Patterns of Arabic canonical nouns

Example	Gloss	Syllable structure	Approximate frequency in Wehr (1971)
baḥr	‘sea’	CVCC	33%
badal	‘substitute’	CVCVC	7%
ʾatān	‘she-ass’	CVCVVC	21%
kātib	‘writer’	CVVCVC	12%
jāmūs	‘buffalo’	CVVCVVC	2%
xanjar	‘dagger’	CVCCVC	14%
waswās	‘temptation’	CVCCVVC	11%

Noncanonical nouns are usually either borrowings that do not conform to the syllabic canons in Table 3, such as /tilifūn/ ‘telephone’ or /safarjal/ ‘quince’ (< Persian), or they are deverbal → participles or gerunds (→ *maṣdar*) acting as nouns, such as these derivatives of Form III /kātab/ ‘to correspond with’: active participle /mukātib/ ‘(the one) corresponding with’, /mukātab/ ‘(the one) corresponded with’, /mukātab+at/ ‘act of corresponding with’.

Canonical nouns from strong roots must conform to one of the syllable-structure patterns in Table 3. Moreover, some (perhaps most) canonical nouns are derived templatically in the same way that verbs are: they combine a productive or semiproductive CV-skeleton morpheme with a consonantal root, a vowel pattern, and possibly an affix. For example, the root *f-t-ḥ* not only forms verb stems like /fataḥ/ ‘to open’ but also noun stems like /miftāḥ/ ‘key’. The stem /miftāḥ/ exemplifies a template used to form instrumental nouns, as do /mibrad/ ‘file’, /mibḍaʿ/ ‘scalpel’, and so on. Similarly, the root *x-y-ṭ*, forms the verb stem /xāṭ/ ‘to sew’ as well as the noun /xayyāṭ/ ‘tailor’. This is an example of the template for occupational and habitual nouns like /ṣarrāf/ ‘money changer’ and /jarrār/ ‘potter’.

With few exceptions, all and only those nouns that conform to the canonical patterns in Table 3 form broken plurals. Broken plurals are marked by a change in syllable structure and vowel quality rather than suffixation. The broken plural system initially looks chaotic; in Wright's (1971) reckoning, there are 31 distinct plural patterns. Closer analysis shows that the system is much more orderly than it might seem. Arguably, all broken plurals can be assigned to one of four formal categories, and there are strong tendencies and even systematic regularities about which singular nouns map onto which plural patterns:

i. About 30 percent of singular nouns with the pattern CāCiC form their plural by doubling the middle root consonant and imposing the vowel pattern /u-a/, as in /kāfir/, pl. /kuffār/ 'infidel', or, less commonly, /kāfil/, pl. /kuffal/ 'breadwinner'.
ii. Several broken plural patterns have a monosyllabic stem, sometimes with a suffix added. Of these, only the CuCC plural is at all common, although it is limited to adjectives of color and bodily defect: /ˀaḥmar/, pl. /ḥumr/ 'red'.
iii. Plurals formed on a CVCVC template are not unusual. This is the most common way of forming plurals of feminine singular nouns with a CiCC or CuCC stem, such as /ḥikm+at/, pl. /ḥikam/ 'maxim', or /rukb+at/, pl. /rukab/ 'knee'. With different vocalization, this template is responsible for the plural of almost 60 percent of CVCVVC nouns, such as /wazīr/, pl. /wuzar+āˀ/ 'vizier'. (When the CVCVC template has /a/ in the first syllable, it undergoes → metathesis: /janāḥ/, pl. /ˀajniḥ+at/, from /janiḥ+at/ 'wing'.)
iv. Overall, the most common broken plural pattern begins with a CVCVV . . . sequence; it is called the iambic plural because this sequence recalls the short-long iambic foot (*watid*) of verse. Almost 95 percent of monosyllabic CVCC masculine singular nouns take a disyllabic iambic plural with one of three different vocalization patterns: /nafs/, pl. /nufūs/ 'soul'; /qidḥ/, pl. /qidāḥ/ 'arrow'; and /ḥukm/, pl. /ˀaḥkām/ (from iambic */ḥakām/ by metathesis) 'judgment'. Of CVCVC singular nouns, 92 percent form their plural in the same way: /ˀasad/, pl. /ˀusūd/ 'lion'. Nearly all members of the populous class of noun stems with four consonants form their plurals on the iambic pattern as well, though with a third syllable added: /jundub/, pl. /janādib/ 'locust'; /sulṭān/, pl. /salāṭīn/ 'sultan'. All nouns with CVCVVC+at singulars take iambic plurals (some have another mode of pluralization as well), as do all nouns with CVVCVVC singulars: /jazīr+at/, pl. /jazāʾir/ 'island'; /jāmūs/, pl. /jawāmīs/ 'buffalo'. The iambic plural pattern is also used with nearly all CaaCic+at nouns: /fākih+at/, pl. /fawākih/ 'fruits'.

Not only is the iambic pattern the most common way of pluralizing a noun in Arabic, but it is also used in the → diminutive. Compare the plurals cited above with the diminutives of the same nouns: /nufays/, /qudayḥ/, /ḥukaym/, /junaydib/, /sulayṭīn/, /juzayyir/, /juwaymīs/, /fuwaykih/. This diminutive pattern is used productively with all canonical nouns, even those that do not form an iambic plural. The only difference between the diminutive and the iambic plural is that the diminutive always has the vowel sequence /u-a-i/ and a /y/, rather than a long vowel, in the second syllable.

The iambic plural and diminutive have been approached from various theoretical perspectives (Hammond 1988; Idrissi 1997; Levy 1971; McCarthy 1983b, 2000; McCarthy and Prince 1988, 1990a; Ratcliffe 1990). The McCarthy and Prince (1990a) analysis is sketched in this entry. Plural and diminutive morphology have characteristic effects on word shape, vowel quality, and appearance of the consonants /w/, /y/, and /ˀ/. Descriptively, iambic plurals and diminutives have both invariant and varying properties of word shape. The invariant part is the initial iambic sequence, and the varying part is the rest. The data are shown in (5) with these two parts separated from one another.

(5) Singular	Plural		Diminutive	
nafs	*nufū*	*s*	*nufay*	*s*
fākih+at	*fawā*	*kih*	*fuway*	*kih*
jazīr+at	*jazā*	*ʾir*	*juzay*	*yir*
jāmūs	*jawā*	*mīs*	*juway*	*mīs*
jundub	*janā*	*dib*	*junay*	*dib*
sulṭān	*salā*	*ṭīn*	*sulay*	*ṭīn*

This division into invariant and varying portions is a consequence of 'prosodic circumscription'. The initial syllable of the singular noun is circumscribed and mapped onto an iambic template: /jun/ ⇒ /janā/. (For /jazīr+at/, which involves a further complication, see McCarthy and Prince 1990a.) The uncircumscribed part of the singular stem is then reattached to this template: /janā/ ⇒ /janādib/. In this way, the correct descriptive generalization is obtained: the plural and diminutive impose a fixed iambic template on the first syllable of the singular stem, but they do not affect properties like vowel length in the second syllable of the singular stem, as shown by the contrast between /janādib/ and /salāṭīn/, whose final syllables have the same length (though not the same vowel) as the corresponding singular nouns /jundub/ and /sulṭān/.

4. Recent research

Most research on Arabic and more generally Semitic morphology since about 1985 falls into one of four loose categories. There is a body of literature discussing the psychological implications of root-and-pattern morphology, using data from language games (McCarthy 1982, 1985), nicknames (Davis and Zawaydeh 1999; Zawaydeh and Davis 1999), aphasia (Prunet a.o. 2000; Safi-Stagni 1991, 1995), and psycholinguistic experiments (Berent a.o. 2001; Berent and Shimron 1997; Boudelaa and Marslen-Wilson 2001, 2004; Feldman a.o. 1995; Frost a.o. 1997). Many of the contributions to Shimron (2003) are also relevant to this topic. A number of works extend the templatic analysis to Arabic vernaculars like Moroccan (Boudlal 2001; Heath 1987) and Cairene (Watson 2002), as well as other Semitic languages like Chaha, which is spoken in Ethiopia (McCarthy 1983a; Prunet and Banksira 1996; Rose 1997). Some works address the issues surrounding consonant doubling in Form IX and geminate verbs like /madad/ 'to measure' (Gafos 1996, 1998; Hoberman 1988; McCarthy 1986; Rose 2000). By far the largest body of literature, however, is devoted to exploring the nature of and alternatives to the root-and-template system of morphology (Bat-El 1989, 1994, 1996; Farwaneh 1990; Gafos 2003; Hudson 1986; McCarthy 1993, 2005; McCarthy and Prince 1986/1996, 1990b; Ratcliffe 1997; Ussishkin 1999, 2000a, 2000b, 2003). This high level of ongoing research activity is proof that Arabic morphology continues to challenge and help advance the study of phonological and morphological theory.

Bibliographical references

Anderson, Stephen R. and Wayles Browne. 1973. "On keeping exchange rules in Czech". *Papers in Linguistics* 6.445–482.

Aronoff, Mark. 1976. *Word formation in generative grammar*. Cambridge, Mass.: MIT Press.

Bat-El, Outi. 1989. *Phonology and word structure in Modern Hebrew*. Ph.D. diss., University of California, Los Angeles.

——. 1994. "Stem modification and cluster transfer in Modern Hebrew". *Natural Language and Linguistic Theory* 12.571–593.

——. 1996. "Selecting the best of the worst: The grammar of Hebrew blends". *Phonology* 13.283–328.

Berent, Iris, Daniel Everett, and Joseph Shimron. 2001. "Do phonological representations specify formal variables? Evidence from the Obligatory Contour Principle". *Cognitive Psychology* 72.1–60.

Berent, Iris and Joseph Shimron. 1997. "The representation of Hebrew words: Evidence from the Obli-gatory Contour Principle". *Cognition* 64.39–72.

Boudelaa, Sami and William D. Marslen-Wilson. 2001. "Morphological units in the Arabic mental lexicon". *Cognition* 81.65–92.

——. 2004. "Abstract morphemes and lexical representation: The CV-skeleton in Arabic". *Cognition* 92.271–303.

Boudlal, Abdelaziz. 2001. *Constraint interaction in the phonology and morphology of Casablanca Moroccan Arabic*. Ph.D. diss., Mohammed V University, Rabat.

Brame, Michael. 1970. *Arabic phonology: Implications for phonological theory and general Semitic*. Ph.D. diss., Massachusetts Institute of Technology.

Chomsky, Noam and Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.

Davis, Stuart and Bushra Zawaydeh. 1999. "A descriptive analysis of hypocoristics in Colloquial Arabic". *Language and Linguistics* 3.83–98.

Eid, Mushira and John J. McCarthy (eds.). 1990. *Perspectives on Arabic linguistics*, II. Amsterdam and Philadelphia: J. Benjamins.

Eid, Mushira and Robert Ratcliffe (eds.). 1997. *Perspectives on Arabic linguistics*, X. Amsterdam and Philadelphia: J. Benjamins.

Farwaneh, Samira. 1990. "Well-formed associations in Arabic: Rule or condition?". Eid and McCarthy (1990:120–142).

Feldman, Laurie Beth, Ram Frost, and Tamar Pnini. 1995. "Decomposing words into their constituent morphemes: Evidence from English and Hebrew". *Journal of Experimental Psychology* 21.947–960.

Fleisch, Henri. 1968. *L'arabe classique: Esquisse d'une structure linguistique*. Beirut: Dar El-Machreq.

Frost, Ram, Avital Deutsch, and Kenneth I. Forster. 1997. "What can we learn from the morphology of Hebrew? A masked-priming investigation of morphological representation". *Journal of Experimental Psychology* 23.829–856.

Gafos, Adamantios. 1996. *The articulatory basis of locality in phonology*. Ph.D. diss., Johns Hopkins University.

——. 1998. "Eliminating long-distance consonantal spreading". *Natural Language and Linguistic Theory* 16.223–278.

——. 2003. “Greenberg’s asymmetry in Arabic: A consequence of stems in paradigms”. *Language* 79.317–355.

Goldsmith, John. 1976a. *Autosegmental phonology*. Ph.D. diss., Massachusetts Institute of Technology. [Published by Garland Press, New York, 1979.]

——. 1976b. “An overview of autosegmental phonology”. *Linguistic Analysis* 2.23–68.

Guerssel, Mohand and Jean Lowenstamm. 1996. “Ablaut in Classical Arabic measure I active verb forms”. Lecarme a.o. (1996:123–134).

Hammond, Michael. 1988. “Templatic transfer in Arabic broken plurals”. *Natural Language and Linguistic Theory* 6.247–270.

Heath, Jeffrey. 1987. *Ablaut and ambiguity: Phonology of a Moroccan Arabic dialect*. Albany: State University of New York Press.

Hoberman, Robert. 1988. “Local and long-distance spreading in Semitic morphology”. *Natural Language and Linguistic Theory* 6.541–549.

Hudson, Grover. 1986. “Arabic root and pattern morphology without tiers”. *Journal of Linguistics* 22.85–122.

Idrissi, Ali. 1997. “Plural formation in Arabic”. Eid and Ratcliffe (1997:123–145).

Lecarme, Jacqueline, Jean Lowenstamm, and Ur Shlonsky (eds.). 1996. *Studies in Afroasiatic grammar: Papers from the Second Conference on Afroasiatic Linguistics, Sophia Antipolis, 1994*. The Hague: Holland Academic Graphics.

Levy, Mary Mansnerus. 1971. *The plural of the noun in Modern Standard Arabic*. Ph.D. diss., University of Michigan, Ann Arbor.

Marantz, Alec. 1982. “Re reduplication”. *Linguistic Inquiry* 13.435–582.

McCarthy, John J. 1979. *Formal problems in Semitic phonology and morphology*. Ph.D. diss., Massachusetts Institute of Technology. [Published by Garland Press, New York, 1985.]

——. 1981. “A prosodic theory of nonconcatenative morphology”. *Linguistic Inquiry* 12.373–418.

——. 1982. “Prosodic templates, morphemic templates, and morphemic tiers”. *The structure of phonological representations*, ed. Harry van der Hulst and Norval Smith, 191–223. Dordrecht: Foris.

——. 1983a. “Consonantal morphology in the Chaha verb”. *The proceedings of the West Coast Conference on Formal Linguistics*, II, ed. M. Barlow, D. Flickinger, and M. Wescoat, 176–188. Stanford, Calif.: Stanford Linguistic Association.

——. 1983b. “A prosodic account of Arabic broken plurals”. *Current trends in African linguistics*, ed. Ivan Dihoff, 289–320. Dordrecht: Foris.

——. 1985. “Speech disguise and phonological representation in Amharic”. *Advances in nonlinear phonology*, ed. Harry van der Hulst and Norval Smith, 305–312. Dordrecht: Foris.

——. 1986. “OCP effects: Gemination and antigemination”. *Linguistic Inquiry* 17.207–263.

——. 1993. “Template form in prosodic morphology”. *Papers from the third annual Formal Linguistics Society of Midamerica Conference*, ed. L. Smith Stvan, 187–218. Bloomington, Ind.: IULC Publications.

——. 2000. “Faithfulness and prosodic circumscription”. *Optimality Theory: Phonology, syntax, and acquisition*, ed. Joost Dekkers, Frank van der Leeuw, and Jeroen van de Weijer, 151–189. Oxford: Oxford University Press.

——. 2005. “Optimal paradigms”. *Paradigms in phonological theory*, ed. Laura J. Downing, T.A. Hall, and Renate Raffelsiefen, 170–210. Oxford: Oxford University Press.

McCarthy, John J. and Alan Prince. 1986/1996. *Prosodic morphology, 1986*. Report. New Brunswick, N.J.: Rutgers University Center for Cognitive Science. [Available at http://ruccs.rutgers.edu/pub/papers/pm86all.pdf.]

——. 1988. “Quantitative transfer in reduplicative and templatic morphology”. *Linguistics in the morning calm*, ed. Linguistic Society of Korea, 3–35. Seoul: Hanshin Publishing Co.

——. 1990a. “Foot and word in prosodic morphology: The Arabic broken plural”. *Natural Language and Linguistic Theory* 8.209–283.

——. 1990b. “Prosodic morphology and templatic morphology”. Eid and McCarthy (1990:1–54).

Prince, Alan. 1975. *The phonology and morphology of Tiberian Hebrew*. Ph.D. diss., Massachusetts Institute of Technology.

Prunet, Jean-François and Degif Petros Banksira. 1996. “L’interaction entre schèmes et racines en chaha”. Lecarme a.o. (1996:302–336).

Prunet, Jean-François, Renée Béland, and Ali Idrissi. 2000. “The mental representation of Semitic words”. *Linguistic Inquiry* 31.609–648.

Ratcliffe, Robert. 1990. “Arabic broken plurals: Arguments for a two-fold classification of morphology”. Eid and McCarthy (1990:94–119).

——. 1997. “Prosodic templates in a word-based morphological analysis of Arabic”. Eid and Ratcliffe (1997:147–171).

Rose, Sharon. 1997. *Theoretical issues in comparative Ethio-Semitic phonology and morphology*. Ph.D. diss., McGill University, Montreal.

——. 2000. “Rethinking geminates, long-distance geminates and the OCP”. *Linguistic Inquiry* 21.85–122.

Safi-Stagni, Sabah. 1991. “Agrammatism in Arabic”. *Perspectives on Arabic linguistics*, III, ed. Bernard Comrie and Mushira Eid, 251–270. Amsterdam and Philadelphia: J. Benjamins.

——. 1995. “Morphological structure and lexical processing: Evidence from Arabic”. *Perspectives on Arabic linguistics*, VII, ed. Mushira Eid, 93–106. Amsterdam and Philadelphia: J. Benjamins.

Shimron, Joseph (ed.). 2003. *Language processing and acquisition in languages of Semitic, root-based, morphology*. Amsterdam and Philadelphia: J. Benjamins.

Spencer, Andrew. 1991. *Morphological theory*. Oxford: Blackwell.

——. and Arnold M. Zwicky (eds.). 1998. *The handbook of morphology*. Oxford: Blackwell.

Ussishkin, Adam. 1999. “The inadequacy of the consonantal root: Modern Hebrew denominal verbs and output-output correspondence”. *Phonology* 16.401–442.

——. 2000a. *The emergence of fixed prosody*. Ph.D. diss., University of California, Santa Cruz.

——. 2000b. “Root-and-pattern morphology without roots or patterns”. *Proceedings of NELS 30*, ed. Masako Hirotani, Andries Coetzee, Nancy Hall, and Ji-Yung Kim, 655–670. Amherst, Mass.: GLSA.

——. 2003. "Templatic effects as fixed prosody: The verbal system in Semitic". *Research in Afroasiatic grammar*, III, ed. Jacqueline Lecarme, Jean Lowenstamm, and Ur Shlonsky, 511–530. Amsterdam and Philadelphia: J. Benjamins.

Watson, Janet C.E. 2002. *The phonology and morphology of Arabic*. Oxford: Oxford University Press.

Weeda, Donald. 1992. *Word truncation in prosodic morphology*. Ph.D. diss., University of Texas, Austin.

Wehr, Hans. 1971. *A dictionary of Modern Written Arabic*. Ed. and trans. J. Milton Cowan. Ithaca, N.Y.: Spoken Language Service.

Wright, William. 1971. *A grammar of the Arabic language*. 3rd ed. rev. W. Robertson Smith and M.J. de Goeje. Repr. Cambridge: Cambridge University Press. (Originally published, 1896.)

Zawaydeh, Bushra and Stuart Davis. 1999. "Hypocoristic formation in Ammani-Jordanian Arabic". *Perspectives on Arabic linguistics*, XII, ed. Elabbas Benmamoun, 113–139. Amsterdam and Philadelphia: J. Benjamins.

JOHN J. MCCARTHY
(University of Massachusetts Amherst)

Motherese → Caretaker Talk

Muḍāriʿ → Māḍī and Muḍāriʿ

Muʿrab → ʾIʿrāb

Mubtadaʾ → Ibtidāʾ

Muḍāf → ʾIḍāfa

Mudġam → ʾIdġām

Muḍmar → ʾIḍmār

Mufaxxam → Tafxīm

Muḥaqqaq

Although the term *muḥaqqaq* as a script or style of handwriting appears in a number of early sources, it is far from clear what its salient features were. Moreover, some of these sources explore the term *muḥaqqaq* not as a particular script, but as a standard of handwriting. Thus, for instance, aṣ-Ṣūlī (d. 335/946), to begin with the earliest text, says that "the best looking of scripts is the delicate *muḥaqqaq*, with its rounded letters, its open (*maftūḥ*) *ṣ*'s and *ṭ*'s, and its slurred or curtained (*muxtalis*) *t*'s and *ḥ*'s" (Abbott 1939:29).

ʾAbū Ḥayyān at-Tawḥīdī (d. after 400/1009), the author of *Risāla fī ʿilm al-kitāba*, defines *taḥqīq* as one of the principles or forms of writing which involves accuracy, clarity, and elegant arrangement of letters (Rosenthal 1971:26). One could therefore say that perfection and accuracy of letter forms were the hallmark of this script. This idea is also explored by Ibn Xalaf, who, writing at the turn of the 4th/10th and the beginning of the 5th/11th century, states that script is divided into two categories: *muḥaqqaq* 'exact' and *muṭlaq* 'negligent; common'. The former is characterized by properly executed letters, and it is used for important matters such as appointments, registrations, and grants of property, while the latter, which allows for (unconventional) joining of letters, is used for urgent correspondence and everyday matters (Stern 1964:105–106; Abbott 1939:28–29).

Indeed, the letter forms that are termed by al-Qalqašandī (*Ṣubḥ* III) *muḥaqqaq* are all *muḏhar* 'with their elements clearly visible'; they stand in contrast to those termed *mudġam, maḥḏūf, muʿallaq*, and *muxfā*, i.e. 'assimilated/contracted/elliptical', a more common feature of the scripts belonging to the → *ṯuluṯ* family.

Muḥaqqaq as a script, however, features in the *Fihrist* of Ibn an-Nadīm (d. 380/990), who refers to scribes (*warrāqūn*) in the early Abbasid period who copied *Qurʾān*s (*maṣāḥif*) in the *muḥaqqaq* and *mašq* scripts (*xaṭṭ*), quoting an earlier unspecified source to the effect that *muḥaqqaq* was also known as *al-ʿirāqī* 'the script of ʿIrāq' and *al-warrāqī* 'the script of the scribes [as opposed to secretaries]' (Ibn an-Nadīm, *Fihrist* 15, 17).

Although Ibn an-Nadīm does not tell us what this ancient script looked like, some scholars have linked it to the New Abbasid Style or 'broken cursive' (known in the past by many names, including 'Kufic *nasxī*') supposedly canonized by Ibn Muqla (Blair 2006:180, 601). Whether this is true or not, this piece of information is important, as it places this script in the domain of scribes and, therefore, the copying of books.

This is also evident from a *Risāla fī l-kitāba al-mansūba*, which may tentatively be assigned to the early 5th/11th century. Its anonymous author states that Ibn ʾAsad, one of the teachers of Ibn al-Bawwāb (d. 413/1022), wrote poetry in → *nasx*, which approximated (*qarīb min*) *al-muḥaqqaq* (*Risāla fī l-kitāba*, 126). This characterization of *muḥaqqaq* also fits in with the later Mamluk tradition, which defines it as one of the two fundamental scripts (*ʾaṣl*, the other one being *ṯuluṯ*), having a rectilinear aspect/ductus and being rarely used in the chancery.

It is clear from this and other later statements by medieval Arab authors, as well as the numerous specimens of the script that have survived, that *muḥaqqaq*, as a well-defined formal script, is one of the new scripts that Arabic calligraphic tradition links with the reform of Ibn Muqla, initiated at the beginning of the 4th/10th century, but that are in reality associated at a later time with the figure of Ibn al-Bawwāb, whose life spans the second half of the 4th/10th and the beginning of the 5th/11th century.

Works on penmanship, principally from the Mamluk period, following the tradition (*ṭarīqa*) of Ibn al-Bawwāb, give us a rich picture of this script and its variants (Gacek 1987, 1989, 2003). According to this tradition, *muḥaqqaq* was established as the principal rectilinear script (*yābis, mabsūṭ*), in which only a small proportion of pen strokes (perhaps one-third) are curved or curvilinear (Soucek 1979:14). By the 7th/13th century, *muḥaqqaq* became a bookhand, more specifically a Qurʾānic hand/script, used for large-format Qurʾāns, although according to al-Qalqašandī (*Ṣubḥ* III, 48), it was also used in the chancery for the writing of *ṭuġra*s and for letters issuing from rulers (*ṭuġrāwāt wa-kutub al-qānāt*).

The script, as practiced in the eastern Islamic lands (from Egypt eastward), was seriffed with a right-sloping serif (*tarwīs*). Some head-serifs were very sharp and long, like barbs; others, especially in outlined letters, looked like short, thick beaks, wedges, or teardrops. Its *ʾalif* was straight, but tapered at its lower end (foot), and some sources indicate that its length was from seven to nine or even ten rhombic dots.

Another characteristic letter encountered in *muḥaqqaq* is the *lām ʾalif*. Although the *lām ʾalif al-muḥaqqaqa* (with a loop at its base, the *ʾalif* heavily tilted to the right, and the right-sloping serif on the *lām*) quite appropriately predominates, and in the manuscripts from the eastern Islamic lands appears to have been the only form used, we also encounter, here and there, especially in Mamluk Qurʾāns, the *lām ʾalif al-warrāqiyya* (لا), which has a triangular base, and the almost vertical upper part of the shaft (hast) of the *ʾalif*.

Other features include: open counters (‘eyes’ of letters) and flattened and tapered (*mabsūṭ*) endings of most of the descenders (sublinear strokes); indeed, the difference between the ascenders (supralinear upstrokes) and the descenders is very evident (Gacek 2003; Ṭayyibī, *Jāmiʿ* 67–72). The flattened and tapered descenders (*ʿarāqāt, muʿaqqafāt*) of not only *rāʾ/zāʾ, mīm, wāw* but also *jīm/ḥāʾ/xāʾ*, and *ʿayn/ġayn* contrast with the upward curve of these in the *ṯuluṯ* script, and are the best indication of the difference between these two scripts.

Furthermore, *muḥaqqaq*, because of its large size, was vocalized (*ḍabṭ*) with a different pen, and often in a different color, such as blue. Some deluxe *muḥaqqaq* Qurʾāns also have superscript *ʾalifs* of prolongation executed in red ink (Lings 2004).

According to various sources, *muḥaqqaq* had two varieties: large (*jalīl*) and small (*xafīf*). The smaller variants included *maṣāḥif* and *rayḥān* ‘sweet basil’. *Muḥaqqaq* was used from Egypt to Iran, and a type of *muḥaqqaq*, probably current before the Mongol invasion (5th/11th century), was later used in China (Stanley 1999: 12–21). Both *muḥaqqaq* and *rayḥān* were almost entirely discontinued in Ottoman Turkey and Iran after the 10th/16th century in favor of *nasx*.

Some of the greatest *muḥaqqaq* Qurʾāns were produced in the Il-Khanid, Mamluk, and Timurid periods (7th/13th–9th/15th centuries). Special mention should be made here of Qurʾāns associated with Sultan Ūlyāytū (Öljajtü) and Sultan Šaʿbān (James 1988), the superb Qurʾān penned by ʾAḥmad as-Suhrawardī, Yāqūt’s most famous pupil, between 701 and 708/1301–1308 (Blair 2006:248–253), the largest Qurʾān to survive from Mamluk times penned by Ibn aṣ-Ṣāʾiġ (d. 845/1441 or 1442), the author of *Tuḥfa ʾūlī l-ʾalbāb*, an important work on calligraphy (Blair 2006:323–325), and finally the gigantic Bāysunghur Qurʾān, measuring 177 x 101 cm, penned in the early 9th/15th century (Lings 2004:81; Blair 2006:265–268).

The *maṣāḥif* script (*qalam al-maṣḥif* lit. ‘script of the Qurʾāns’) is mentioned in the anonymous treatise *Risāla fī l-kitāba al-mansūba*. Here we are told that Ibn al-Bawwāb distinguished between two scripts: *matn* 'body of the text; block of text' (i.e. large → *nasx*) and *maṣāḥif* (*Risāla fī l-kitāba* 126). Another reference to this script comes from the *Muʿjam al-ʾudabāʾ*, whose author, Yāqūt al-Ḥamawī, mentions that ʿAlī ibn Ṭalḥa ar-Rāzī al-Baqašlan (d. 515/1121), who followed Ibn al-Bawwāb's method or 'school' (*ṭarīqa*), was particularly famous for writing in *qalam al-maṣāḥif* (Rice 1955:10). This script is later illustrated in a unique album from the end of the Mamluk period executed by aṭ-Ṭayyibī (*Jāmiʿ* 54–57). It emerges as a script similar in size to *nasx*, but much bigger than *rayḥān* (see below). One of the features of this script is that, just as with *muḥaqqaq*, it is seriffed, in contrast to *nasx* (*matn*), which is serif-less (sans serif). Furthermore, its appellation points to its usage, namely, the copying of single-volume Qurʾāns.

Probably the best example of this script (referred to in the original description by Rice as *nasxī*) is the codex of Ibn al-Bawwāb preserved in Chester Beatty Library (Dublin) and dated 391/1000 or 1001. It is worth noting here that the use of the head-serif in this copy is not systematic. A similar script (also seriffed) is seen in a Qurʾān falsely attributed to Ibn al-Bawwāb copied in 401/1010 (Rice 1955: 24–25).

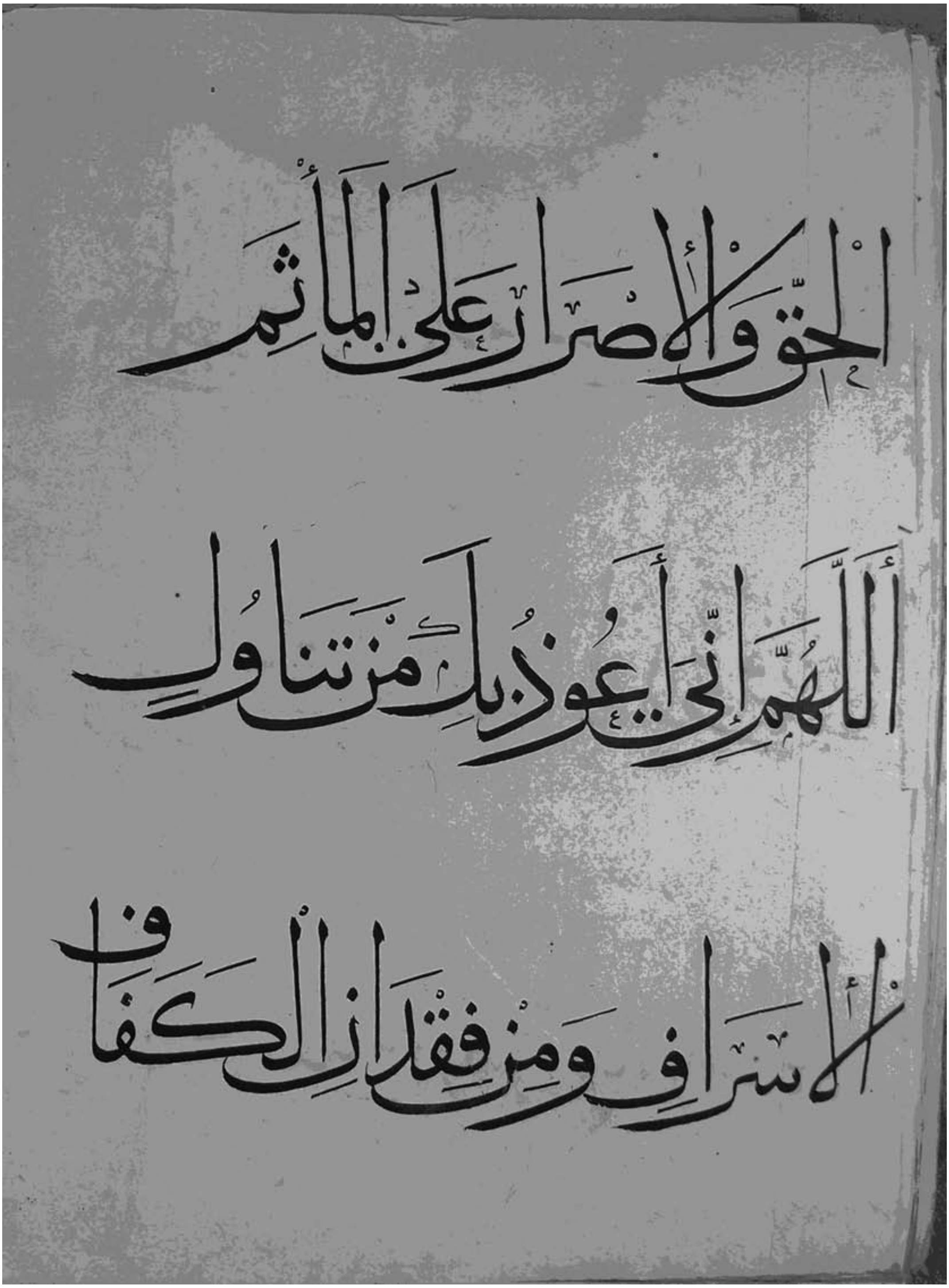
S.M. Stern, although speaking of the script of the Chester Beatty Qurʾān as being written in *nasx*, concedes in a footnote that "the specimen of the *maṣāḥif* script according to the manner of Ibn al-Bawwāb [...] corresponds fairly well to the script of the Chester Beatty MS". He goes on to say, "The name 'Koranic script' suggests that the *nasx* used by Ibn al-Bawwāb and his school for Korans was of a particular kind, distinguished from ordinary *nasx*" (Stern 1969:19). David James, even though he refers to this script as *nasx*, also concedes that "the semi-circular strokes are flattened out slightly in the manner of *muḥaqqaq*" (James 1988:18; for similar suggestions, see Blair 2006:167).

If Ibn al-Bawwāb really distinguished between *matn* and *maṣāḥif* scripts, as the text of the aforementioned *Risāla* has it, then it would stand to reason that he would have used this latter script, and not *nasx*, for the copying of Qurʾāns. *Maṣāḥif* script, however, seems to have had a short-lived existence as it was replaced, perhaps as early as the 6th/12th century, in favor of a large *muḥaqqaq* and a very small *rayḥān* as a favored pair of scripts for large Qurʾāns. A good example of an early Qurʾān using a combination of a large *muḥaqqaq* (top, middle, and bottom lines) and a small *rayḥān* (in the two blocks in between) is a Chester Beatty copy dated 582/1186 and executed by ʿAbd ar-Raḥmān al-Kātib al-Malikī (James 1980:35; see also Blair 2006:217–218)

Rayḥān script, also mentioned in the anonymous *Risāla* and other later works, had more or less the same relationship to *muḥaqqaq* as the *ḥawāšī* to *nasx* or the *ġubār* (the smallest of all scripts and a hybrid) to *riqāʿ* (→ *ṯuluṯ*), or, according to al-ʾĀṯārī, it was half the size of *muḥaqqaq*. All the letters in this script were the same as in *muḥaqqaq* but finer, and the script was also seriffed. Furthermore, unlike *muḥaqqaq*, its vocalization was done with the same pen, because of its small size (Gacek 1989:146; Ṭayyibī, *Jāmiʿ* 17, 73–77). One of the earliest specimens of this script can be found in a multi-script manuscript of the poetry of Salāma ibn Jandal, executed, according to the colophon, by Ibn al-Bawwāb (although this attribution is yet to be proven convincingly) and used for the text of the commentary (Blair 2006:172–173). A later excellent example of *rayḥān* as text script (as opposed to display script) is a Tehran Qurʾān attributed to Yāqut al-Mustaʿṣimī completed in 685/1286 (Blair 2006:243–245).

In the later Ottoman period and in Iran, the term *rayḥānī* was often used for, or as a variant of, *muḥaqqaq* (Soucek 1979:12; Gacek 2003). It is worth noting here that the same appellation (*ar-rayḥānī*) is mentioned by ʾAbū Ḥayyān at-Tawḥīdī as one of the twelve 'Kufic' styles, perhaps a reference to ʿAlī ibn ʿUbayda ar-Rayḥānī (d. 219/834), "a master of elegant writing and style" (Rosenthal 1971:24; Blair 2006:167).

Rayḥān does not seem to have been used widely as a script of the *Qurʾān* in Egypt and Syria in the Mamluk period. On the other hand, the combination *muḥaqqaq* and *rayḥān* on the same page was common in Iran and the Ottoman world (James 1988:20–21). One of the best examples of this kind of Qurʾān using both *muḥaqqaq* (first, median, and last line) and *rayḥān* on the same page, from the Ottoman



الحق والاصرار على المأثم

اللهم إني أعوذ بك من تناول

الاسراف ومن فقدان الكفاف

Fig. 1. *Jalīl al-muḥaqqaq* by Muḥammad aṭ-Ṭayyibī (Ṭībī), ca. 908/1502 (Istanbul, TSM K.882, f.35b).

period, is the large Qurʾān (615 × 425 mm) of ʾAḥmad Qarahḥisārī (Ahmed Karahisari), executed before 953/1546, preserved in the Topkapı Palace Library and recently reproduced in facsimile.

As far as is known, *muḥaqqaq* and its derivatives have not been used in printing (lithography or movable type), and even in modern calligraphy it does not occupy the important place it once had.

Bibliographical references

Primary sources

Risāla fī l-kitāba = (anon.), *Risāla fī l-kitāba al-mansūba*, ed. Xalīl Maḥmūd ʿAsākir, "Risāla fī l-kitāba al-mansūba". *Majalla Maʿhad al-Maxṭūṭāt al-ʿArabiyya* 1 (1955) 121–127.

Ibn an-Nadīm, *Fihrist* = Muḥammad ibn ʾIsḥāq ibn an-Nadīm al-Warrāq al-Baġdādī, *Kitāb al-fihrist*. Ed. Yūsuf ʿAlī Ṭawīl. Beirut: Dār al-Kutub al-ʿIlmiyya, 1996.

Qalqašandī, *Ṣubḥ* = ʾAḥmad ibn ʿAlī al-Qalqašandī, *Ṣubḥ al-ʾaʿšā fī ṣināʿat al-ʾinšāʾ*. 14 vols. Cairo: al-Muʾassasa al-Miṣriyya al-ʿĀmma, 1963–1965. (First publ., aṭ-Ṭabʿa al-ʾAmīriyya, 1913–1918.)

Ṭībī, *Jāmiʿ* = Muḥammad ibn Ḥasan aṭ-Ṭayyibī, *Jāmiʿ maḥāsin kitābat al-kuttāb*. Ed. Ṣalāḥ ad-Dīn al-Munajjid. Beirut: Dār al-Kitāb al-Jadīd, 1962.

Secondary sources

Abbott, Nabia. 1939. *The rise of the North Arabic script and its Ḳurʾānic development, with a full description of the Ḳurʾān manuscripts in the Oriental Institute*. Chicago: University of Chicago Press.

Blair, Sheila S. 2006. *Islamic calligraphy*. Edinburgh: Edinburgh University Press.

Gacek, Adam. 1987. "Al-Nuwayrī's classification of Arabic scripts". *Manuscripts of the Middle East* 2.126–130.

——. 1989. "Arabic scripts and their characteristics as seen through the eyes of Mamluk authors". *Manuscripts of the Middle East* 4.144–149.

——. 2003. "The head-serif (*tarwīs*) and the typology of Arabic scripts: Preliminary observations". *Manuscripta Orientalia* 9:3.27–33.

James, David. 1980. *Qurʾans and bindings from the Chester Beatty Library: A facsimile exhibition*. London: World of Islam Festival Trust.

——. 1988. *Qurʾāns of the Mamlūks*. London: Thames and Hudson.

Lings, Martin. 2004. *Splendours of Qurʾan calligraphy and illumination*. Stuttgart: Tradigital.

Rice, D.S. 1955. *The unique Ibn al-Bawwāb manuscript in the Chester Beatty Library, Dublin*. Dublin: Emery Walker.

Rosenthal, Franz. 1971. "Abū Ḥayyān al-Tawḥīdī on penmanship". Franz Rosenthal, *Four essays on art and literature in Islam*, 20–49. Leiden: E.J. Brill.

Soucek, Priscilla P. 1979. "The arts of calligraphy". *The arts of the book in Central Asia, 14th–16th centuries*, ed. Basil Gray, 7–33. Paris and London: Serinda.

Stanley, Tim. 1999. "China: Qurʾāns of the Ming period". *The decorated word: Qurʾāns of the 17th to 19th centuries*, ed. Manijeh Bayani a.o., 12–21. London and New York: Nour Foundation [in association with Azimuth Editions and Oxford University Press].

Stern, S.M. 1964. *Fāṭimid decrees: Original documents from the Fāṭimid chancery*. London: Faber and Faber.

——. 1969. "A manuscript from the Library of the Ghaznawid Amīr ʿAbd al-Rashīd". *Paintings from Islamic lands: Essays*, ed. R. Pinder-Wilson, 7–31. Oxford: Cassirer.

Adam Gacek (McGill University)

Multilingualism

Multilingualism can be defined either at the societal level, where two or more languages are used within a single community, or at the level of the individual, where the same person uses more than one language. These two phenomena do not always co-occur; within the Arab world, multilingual individuals may live within states that are regarded as monolingual, while in multilingual states many or even most individuals may be monolinguals. The multilingualism of individuals or communities may or may not receive official recognition within the constitution or legal system of a state. Attempts to influence and manipulate multilingual situations, with varying degrees of success, may come from state intervention or from activist movements.

1. Introduction

Discussions of language in the context of the Arab world have often focused on the status of Arabic as a cement holding together this community (for examples, see Suleiman 2003). Again and again it has been declared that Arabic is the key component of Arab identity, the heart and soul of the Arab nation (→ nationalism and language). According to al-Ḥuṣrī (1985:46), for instance, every individual belonging to an Arabic-speaking community is an Arab, whether or not he/she accepts this identity. Yet, these states, like the majority of other states in the world today, are far from monolithic, monolingual societies. Many of them are the sites of complex interactions among several languages.

The term 'Arabic' subsumes a large number of highly differentiated varieties which are not always mutually intelligible. States from Morocco to Oman are united in recognizing as an official language the formal written variety of Arabic (→ Classical or → Modern Standard) – a variety that is not the native language of any community but is acquired through formal education. Considerable differences between colloquial varieties can also be considered an obstacle to communication rather than a unifying force. An educated Algerian and a Lebanese may very well find it easier to understand each other if they both speak French, while a Sudanese and a Palestinian may find it more convenient to communicate via English!

This entry does not address differences across varieties of Arabic, although such differences have frequently resulted in their being treated in statistically oriented language research as separate languages (for instance by Gordon 2005 and Leclerc 2003). It instead focuses on cases where Arabic varieties are in contact with other languages on either the individual or societal level. The examples above simply argue against the tendency to assume that Arabic inevitably unites its community of speakers whereas other languages divide it. Furthermore, within the Arab world, linguistic divisions frequently fail to correlate with cultural differences. For instance, Arabic in the Middle East is used by both Christian and Muslim communities, whereas in North Africa Berber- and Arabic-speaking communities are distinguished by language and united by Islam. In Morocco, the Toshabim, a native Jewish community whose ancestors have been present there since before the advent of Islam, have for centuries also used Arabic and Berber varieties, depending on their geographical distribution. Thus, a shared language does not imply a shared religion, and vice versa.

2. SOURCES OF MULTILINGUALISM

Current instances of multilingualism involving Arabic can be traced to different sources. Many of the earliest large-scale contacts between Arabic and other languages are the result of the Arabs' movements out of Arabia following the founding of Islam in the 7th century C.E., which culminated in their conquest of the lands stretching from North Africa and Spain in the west to the Indus Valley in the east. As conquered peoples converted to Islam, they often adopted Arabic. The consequences, more than a thousand years later, are various. In some places, such as Tunisia, Arabic has almost totally displaced the indigenous languages; in others, such as Spain, the use of Arabic over centuries can now be attested only through traces identified in borrowings and place names, for example. In one remarkable case (→ Maltese), the Arabic spoken centuries ago has survived and remains the first language of the majority of the population, although it has changed its name and its alphabet; the Maltese language, written in the Roman alphabet and designated an official language of Malta along with English, is unmistakeably a variety of Arabic similar to those used in North African countries.

In those places where Arabic and the other languages still coexist side by side, a number of different types of situation can be found. There are many states now dominated by other languages where small Arabic-speaking communities still exist, as in the case of → Mali, Niger, → Nigeria, → Ethiopia, → Djibouti/Eritrea, → Turkey, and → Afghanistan. In states where Arabic is used on a large scale, there are quite sizable groups using other languages, such as the → Kurdish-speaking communities of Iraq and Syria and, in Morocco and Algeria, speakers of Amazigh (the name currently preferred in referring to the → Berber varieties spoken in this area). As discussed below, the assignment of official status to Arabic does not always correlate with the existence of an Arabic-speaking majority.

Alongside these old-established cases of bilingualism, there is a second, much more recent category, the result of the colonization of North Africa and the Middle East by European powers in the 19th and 20th centuries. Colonization led to the use of French in Morocco, Algeria, Tunisia, Syria, and Lebanon, Spanish in the north of Morocco and the western Sahara, Italian in Libya, and English in Egypt, Iraq, Kuwait and Palestine. Another recent development is the presence of Hebrew alongside Arabic in Israel.

In all these cases where two speech communities have come into contact, the extent to which individual members of each maintain their language, become bilingual, or abandon their original language may be determined by factors such as the amount of interaction between the two communities, relations of power and/or solidarity between them, group policies on providing or withholding access to the two languages, ideological positions, and perceptions of the value or usefulness of the other's language. For instance, the survival of Arabic-speaking communities in places like Nigeria may be related to its special status within Islam; the much greater use of French among Algerians compared to the use of Italian among Libyans may relate to degree and length of contact as well as colonizers' policies; and the fact that both Palestinians in the occupied territories and Israeli Arabs take advantage of opportunities to learn Hebrew (Spolsky 1997; Spolsky and Shohamy 1999) shows that ultimately pragmatic motives may dominate over ideological considerations.

Contacts between Arabic speakers and other groups have also continued as a result of migration, largely from the Arab world toward → Europe and → North and → Latin America. The first wave of emigration to the United States began in the 1870s, and current estimates of the number of Arab Americans vary between one and three million. In the 2000 U.S. census, 614,414 people claimed to use Arabic at home. The majority of Arabs in the United States are of Lebanese, Syrian, Palestinian, Egyptian, or Iraqi origin (for more details, see Rouchdy 1992; Samhan 2001; Suleiman 1999). On the other hand, large numbers of Arabic-speaking North Africans emigrated to France, Belgium, the Netherlands, and Italy; the Arabic-speaking population of France is currently estimated at between four and five million. Immigrants in most cases rapidly become bilingual, and there is typically a process of language shift over two or three generations, so that these bilingual communities may well prove transient. However, a recent Australian study found that the Arabic-speaking community of Melbourne was more resistant to language shift than the Chinese and Spanish-speaking communities (Clyne and Kipp 1999).

Other multilingual situations arise from the presence, within certain Arab states, of large numbers of foreign workers. For instance, in Kuwait, foreigners outnumber citizens and include speakers of Farsi, Hindi, Urdu, Bengali, and Filipino. Similar language communities are to be found in Saudi Arabia and several other Gulf states.

One final type of bilingualism should be mentioned – one which results from educational and cultural activity rather than from direct contact between different language communities. Citizens of Arab countries, like those of the rest of the world, are very conscious of the value of knowing an international language; many are willing to make a great deal of effort and sometimes financial investment in order to reach a high degree of proficiency in English, French, or Spanish. Likewise, Muslims all around the world are strongly motivated to acquire at least a minimal knowledge of Arabic. Knowledge of Classical Arabic is often transmitted through traditional Qur'ānic schools, particularly to those who may not use this language outside the religious domain.

The various historical factors have yielded several different types of language situation, ranging from multilingual communities where Arabic is very much a minority in terms of number of speakers to those where Arabic speakers predominate and where even those whose first language is not a variety of Arabic are almost bound to learn it as a second language. However, the official statements on language in the constitutions and laws of the various states frequently do not reflect the complex multilingualism prevalent among their peoples.

3. Official recognition of multilingualism

The states that are typically felt to constitute the Arab world are in fact unanimous in recognizing Arabic (understood to refer to the Classical or Modern Standard variety) as their sole official language. Among these states, relatively few, such as Kuwait and Saudi Arabia, can be described as linguistically homogeneous, in the sense that Arabic is the first language of (almost) all their citizens. In addition to these, there are states like Iraq, which includes communities speaking Kurdish, Azeri, Syriac, Farsi, Turkmen, Armenian, and Circassian, or Egypt, with its speakers of → Nubian, Greek, Kenuzi-Dongola, Siwi, and Domari. Many of these linguistic communities are spread across several states: Armenian speakers are found in Syria, Jordan, Lebanon, and Iraq, and speakers of varieties of Berber, known by different names, are found in Morocco, Algeria, Tunisia, Libya, Egypt, and Mauritania. As for European languages, Lebanon, with its 45 percent French-speaking population (according to a recent IPSOS survey), hosted the 2002 summit of Francophone countries. At the other extreme of the Arab world, French and Spanish are actually the first learned and home languages for many of the Jewish families remaining in Morocco (Bentahila and Davies 1992b). Linguistic diversity is most striking along the periphery of the Arab world – in → Mauritania, which recognizes four national languages besides its official one, and → Sudan, where about one hundred different languages are spoken.

The fact that, after independence, all these states chose to designate Arabic as their sole official language can easily be related to political as well as purely linguistic considerations. Identification with Arabic was often a highly symbolic move, representing a reaffirmation of the people's historical identity, internal unity, and wider solidarity with the other Arab states. For instance, in Morocco there was a tremendous counterreaction to the French colonizers' attempts to foster divisions between the Berber- and Arabic-speaking groups, through measures such as the *Dahir Berbère* (Berber decree) of 1930, which sought to provide the Berber-speaking areas with a separate education system using French and Berber but not Arabic (Bentahila 1983; Bidwell 1973). Once the colonizers left, national unity remained a priority; this was one of the considerations which led many states to focus on Arabic and offer little if any recognition of the other languages used by their citizens. Laws aimed at protecting Arabic from the dominance of the colonizers' language were passed, imposing the use of Arabic in domains such as education and administration; they also had the effect, coincidental or deliberately calculated, of excluding the use of indigenous minority languages from these domains.

The strategy of resorting to Arabic as a unifying symbol has not been equally successful across the entire region. For instance, although Arabic is the sole official language of Sudan, Arabic dialects are the first language of slightly more than half the population, concentrated in the North, while in the South around a hundred other languages are spoken, many by very small communities. The North/South linguistic division also coincides with religious and ethnic divisions, with the South harboring Christians and animists while the North is Muslim. These cultural and language divisions together with attempts by the North to impose its culture and language on the South are at the root of the civil war that has ravaged the country for more than forty years. Sudan represents a case where Arabization, far from being a unifying force, has torn a country apart.

In Mauritania, too, the relations between the Arabic-speaking North and the rest of the country are not entirely harmonious. Although the Constitution recognizes → Ḥassāniyya (an Arabic dialect), Fulfulde, Soninke, and Wolof as national languages, only Arabic has official status. Members of the Negro-African ethnic group of the South have perceived measures to promote the use of Arabic as another kind of colonialism and some have protested vociferously against what they see as a racist system which marginalizes users of other languages (see, for instance, Abou Sall 2001).

A second group of states consists of those that are officially bilingual, recognizing two official languages of which one is Arabic. In many of these cases, Arabic is actually spoken by only a small minority of the population. For instance, Arabic is one of the two official languages of the → Comoros (the other being French), yet 96.8 percent of the population speak Comorian, a Bantu language, as their first language. Arabic is also one of the official languages of Somalia, along with Somali. Yet, while the latter is spoken by 62 percent of the population, Arabic is spoken only by a tiny minority of 46,000. Similarly, in Djibouti, only 11 percent of the population are classified as Arabs, and in Chad only 12.3 percent, although Arabic is an official language in both of these states (all statistics are from Leclerc 2003).

The justification for giving Arabic official status in cases like these could be related to its symbolic value as the language of Islam. The vast majority of Comorians, for example, are Muslims, as are almost all Djiboutians. In communities where Islam predominates, Arabic tends to be highly valued even by those who have very little knowledge of it. A survey by Brahimi and Owens (2000) found that the non-Arabic-speaking Kanuri of northeast Nigeria would prefer to learn Arabic rather than English. Another factor favoring Arabic is its value as a language of much wider use than most of the indigenous languages of these states. In Chad, for instance, around 130 languages are spoken, but only eighteen of these have more than 50,000 speakers. In some cases, however, the strongest motive may have ultimately been mercenary. In the 1970s, states like Somalia and Chad adopted pro-Arabic policies in exchange for massive financial aid and political support from the Gulf States (Laitin 1992:90).

Assigning Arabic the status of an official language does not necessarily imply that it will be used in all official domains. Sometimes, its status is more symbolic than anything else, for where there is more than one official language there is not always equality of status between the two. Thus, in Chad, French remains the working language of the administration, and Arabic, though proclaimed as its co-official language, has a role which is very much subordinate to that of French. Likewise, although Arabic is one of the two official languages of Israel, where the Arabic-speaking community constitutes just 20 percent of the population, the true status of Arabic is more like that of a minority language. While it may be used in parliamentary debate and in courts of law, in practice most Arab deputies choose to address the parliament in Hebrew; and while the law stipulates that children in Jewish schools should study Arabic for a minimum of three or four years, in practice many of them are allowed to choose French instead (Spolsky 1997; Spolsky and Shohamy 1999, 2001; Talmon 2000).

Finally, Arabic-speaking minorities in states where Arabic has no official status have been treated in very different ways. For instance, in the Hatay province of Turkey, it is forbidden to teach through the medium of Arabic, to use Arabic personal names, or even to sing Arabic songs in public (Arnold 2000). On the other hand, in several European states, such as France and the Netherlands, fairly generous provisions are made within the state education system to provide Arabic language lessons for the children of North African immigrants (see Altena and Appel 1982; Bentahila and Davies 1991, 1992c; Extra and Gorter 2001, among others). Opinions have differed as to the usefulness of such provisions. In November 2002, President Chirac emphasized that French policy was to encourage the study of Arabic in French schools. In the Netherlands, however, the government abolished the state-financed system of Arabic language education in 2004.

4. Conflicts and rivalries

In situations where one language is perceived as threatening the position of another, movements aimed at protecting the supposedly threatened language often arise. In the immediate aftermath of colonization, most Arab governments were preoccupied with the need to restore Classical Arabic to what was perceived as its rightful position as the medium of government, administration, and education. Vast amounts of energy and money have since been expounded on campaigns for the protection and promotion of Arabic, subsumed here under the label of ‘Arabization programs’. States have varied in the zeal with which they have pursued these programs. Syria, for instance, has pursued a fairly aggressive policy, targeting what is essentially monolingualism by making Arabic practically the exclusive medium for instruction in the public education system and for written media. At the other extreme is Lebanon. Despite its having currently opted for Arabic as its sole official language, Lebanon maintains generous provisions for the teaching of other languages right from the first year of primary school and has adopted a noninterventionist approach, allowing considerable freedom to schools in both the public and private sectors that wish to use other languages as media of instruction (Ghaith and Shaaban 1996). Nevertheless, as a whole, Arabization programs can ultimately be seen as attempts to reduce multilingualism at the societal level, by removing languages other than Arabic from official domains, and at the individual level, by reducing the time devoted to teaching these other languages.

Probably the most powerful protection tool available to Arab states is the education system. Many states have pursued a policy of gradually removing the use of the colonizers' language as a medium of instruction, first in primary and then in secondary schools, and to varying degrees in higher education as well. Once relegated more or less to the status of a foreign language, its introduction into the curriculum is then delayed. This trend stands in sharp contrast to the more global movement toward introducing these same international languages, English and French, at earlier points in the curriculum, and to the general tendency to increase the time devoted to teaching foreign languages.

The other domain most readily manipulated by governments is administration. Strategies employed in this case have included laws designating Arabic as the only, or most prominent, language to be used in domains such as parliamentary debate, legal documents, and official reports. For instance, in Algeria, a 1968 law required all civil servants to be proficient in Arabic, while a 1991 law declared any document drawn in a language other than Arabic invalid and imposed fines of up to 5,000 Algerian dinars on anyone signing such a document. One of the most dramatic measures taken was Libya's 1973 decree requiring that persons entering the country must hold passports in which personal details were recorded in Arabic.

Quite a number of states also found it necessary to pass laws controlling the languages used on official signs and notices, business signs, and product labels. In Algeria, a 1976 circular required all company, administration, and road signs to be exclusively in Arabic, while the laws of 1991 and 1996 allowed bilingual labeling of products only if the Arabic version was prominent. Egypt passed a decree in 1986 banning monolingual signs in any language other than Arabic. More recent interventions have included the enforcement in Amman, Jordan, in 2001, of a municipal regulation requiring the removal of English shop signs, while a similar campaign in Tunis in 1999 required shopkeepers to remove the French parts of their signs. The continuing concern to defend Arabic against the use of other languages can also be seen in a symposium held in Cairo in 1999, which passed seventeen resolutions, one of which was implemented by the 1999 law banning businesses from using non-Arabic names, trademarks, or brand names.

Despite these often elaborate attempts to obtain a more homogeneous linguistic profile, multilingualism is still flourishing, even in those states that have pursued vigorous Arabization campaigns. Some of the measures taken failed to produce the intended effects; paradoxically, some have produced counterreactions. In many places, including Egypt and Morocco, the reduction in the use of English or French in the public education system has led to an increased recourse to private education as an alternative for extensive training in these languages, often to the detriment of Arabic. State policies have thus turned a high level of proficiency in French or English into something exclusive, for which parents are willing to pay. While the options open to the majority are often reduced by protectionist policies, the position of the elites able to bypass this policy is only strengthened. Laitin (1992:152) reports on a similar phenomenon, which he terms the "private subversion of a public good", in Somalia, where reforms making → Somali the dominant language in the national education system led the elite to send their children to Egypt to attain proficiency in Arabic.

Global trends in information technology and communications provide even stronger motivations for acquiring proficiency in just those languages weakened by Arabization programs, notably English. The World Wide Web has undoubtedly given new impetus to the study of English as a foreign language and reinforced its role as a global language (Crystal 1997). Some interesting recent statistics indicate that in 2001 Arabic accounted for only 0.04 percent of all web pages (68.39% being in English). When the numbers of pages in each language are ranked by the number of speakers of that language, Arabic comes right at the bottom of the list, with one Arabic web page for every 1,583.5 Arabic speakers, whereas English has one page per 1.5 English speakers (Carvin 2001). Figures like these would seem to represent a powerful motive for preserving bilingualism among Arabic-speakers.

In the administration, too, there is still a discrepancy between the public, official face and what goes on behind the scenes. For instance, official documents may be published only in Arabic, but the discussion in preparatory meetings and even preliminary drafts of the document may well involve the use of French or English. And when officials, businessmen, or intellectuals from different Arab countries meet, it is still common to find them opting to communicate, not in any variety of Arabic but rather in French or English.

Finally, attempts to regulate the languages used in street signs have rarely had the intended effects. Sometimes they were never fully enforced in the first place, having been applied only to certain cities, on a piecemeal basis, or only for a short period of time. Some were later withdrawn altogether. For instance, the 1999 Tunis campaign mentioned earlier initially led to the French parts of shop signs being hastily covered up with cardboard or plastic bags, practically overnight, in order to avoid hefty fines; but a few months later these temporary screens disappeared and the issue was dropped.

While governments have for decades been seeking to protect Classical Arabic from the excolonizers' languages, indigenous languages such as Kurdish in the Middle East and Berber varieties in North Africa have long suffered from a total lack of recognition. There are, however, notable signs of change here. For instance, the Kurds of North Iraq, thanks to the relative autonomy obtained following the 1991 Gulf War, are now able to use their language in education and adminstration. In contrast, in neighboring Syria, the use of Kurdish as a medium of instruction is not allowed. Similarly, when Morocco and Algeria obtained independence, no provision was made for the teaching of Amazigh in schools and no effort made to encourage writing in this language. A survey in Morocco by Bentahila and Davies (1992b) suggested that younger generations were abandoning the use of Amazigh, although Brahimi and Owens' (2000) study on Algeria did not find significant evidence for such a shift. As a result of activists' campaigns, such as the Berber Manifesto published in 2000, there have recently been significant changes in the status assigned to Amazigh, the year 2001 being a turning point. In October of that year, President Bouteflika promised to bestow official status upon Amazigh in Algeria; during that same month Morocco's Mohamed VI set up the Royal Institute of Amazigh Culture. Morocco now has a new policy of teaching Amazigh in primary schools, due to be implemented across the country by 2008. New materials are rapidly being prepared using the traditional Berber alphabet, Tifinagh, despite objections that opting for this writing system means that children will be required to master three separate alphabets (Tifinagh, Arabic, and Roman) in the early years of schooling (→ language shift: Amazigh).

New communication opportunities provided by satellite televisison and the Internet have been exploited effectively by promoters of minority languages. For instance, the Kurdish language television channel MED TV is accessible to the many Kurds living in exile, while web sites using Amazigh have proliferated in recent years (a recent compilation lists 288 such sites).

5. Multilingualism in everyday life

The multilingual situation prevalent in large parts of the Arab world means that many speakers of Arabic also use one or more other languages on a regular basis, and that many more who are not actually bilingual themselves are accustomed to being exposed to other languages in their daily environment. In some cases, languages are associated with quite separate domains, with fairly rigid divisions between the situations where each is likely to be used. A language may be closely associated with particular types of interlocutor, topic, and/or setting. For instance, a young Amazigh-speaking Moroccan may use this language only with parents and grandparents, speaking Arabic with his or her friends and peers and French with a doctor (for a detailed investigation of language choice among Moroccan bilinguals, see Bentahila 1983). An Egyptian may use Arabic in writing a letter about family news, but English when writing about a financial or business matter. A Tunisian out shopping may use Arabic at the butcher's but French at the bank. Furthermore, the emergence of new domains may lead to changes in the norms of usage for particular languages. For instance, colloquial varieties of Arabic have traditionally been used only orally, not in writing; yet, one interesting consequence of the new communication technology is that many people have taken to using colloquial Arabic, transliterated in the Roman alphabet, for e-mails and text messages as well as Internet chatting. This habit is now common among young people in Morocco. A recent study from Egypt (Warschauer a.o. 2002) reports that young professionals there tend to use English for formal e-mail correspondence and Egyptian Arabic for informal messages, suggesting that this may be a new type of → diglossia, from which Standard Arabic is being squeezed out.

In many cases, however, there is considerable overlap between the domains of the various languages. In domains where two languages seem usable, or in a situation involving a mixture of factors, some favoring one language and some another, speakers may find themselves frequently switching back and forth between the two languages. In many bilingual communities, varieties have emerged that exhibit recurrent patterns of → code-switching: French/Arabic in Algeria, Tunisia, and Morocco, as well as among immigrants in France (Belazi 1992; Bentahila and Davies 1983, 1992a, 1993, 1994, 1995, 1998; Boumans and Caubet 2000), Spanish/Arabic in the north of Morocco, English/Arabic by Egyptians and by immigrants to North America (Atawneh 1992; Eid 1992; Hussein and Shorrab 1993), Dutch/Arabic by Moroccans resident in the Netherlands (Nortier 1990), Hebrew/Arabic by Israeli Arabs, and so on. These code-switching varieties often serve as powerful in-group markers, allowing bilinguals to distinguish themselves from monolingual speakers of either language. The norms for where switches occur and what types of structure are typically involved may also vary among subgroups of a bilingual community; for instance, in Morocco, different types of code-switching have been associated with different generations and related to differences in educational background and experience with each language (Bentahila and Davies 1992a, 1995, 1998).

Naturally enough, where two or more languages are used within a community, each will acquire particular associations and values. Members of the community, whether or not they are themselves bilingual, may develop contrasting attitudes toward the different languages due to cultural, historical, and educational factors. Arabic speakers are known for their tendency to admire and respect Classical Arabic for its special religious and nationalist associations and for their feeling that colloquial Arabic varieties are inferior, inadequate forms of expression. Attitudes to other home languages may vary from a fierce attachment to the symbol of one's heritage among some Amazigh or Kurdish speakers, for example, to a feeling that such languages are not very useful or relevant to the modern world among others. Some studies have suggested that many people of Amazigh origin do not feel that this language is an essential marker of their identity (Bentahila and Davies 1992b). In a survey, many Moroccan bilinguals identified as their 'own language' a language other than the one they first learned or used at home (Davies and Bentahila 1989). As for the colonizers' languages, these have inspired very mixed feelings: some resent having to use them, seeing them as weapons of the Western aggressor; others enjoy them as keys to modernity and social advancement.

Interesting results have been obtained from matched guise tests where respondents are asked to listen to sound recordings and evaluate speakers for a list of personality traits without being told that they are hearing the same persons in two or more different guises (speaking different languages). In one such study carried out in Morocco, bilinguals were judged to be more modern, intelligent, important, and educated when using French than when using Moroccan Arabic; differences also emerged depending on the type of accent used in French (Bentahila 1983).

At present, then, multilingualism appears to be a normal, everyday part of people's lives throughout much of the Arab world. Moroccans, for instance, view national television channels where programs in Arabic alternate with ones in French and where the news is read each day in Standard Arabic, Amazigh, French, and Spanish, while on local radio stations speakers may switch freely between Arabic and French. They can listen to popular music of various genres whose lyrics exhibit code-switching between two languages (Bentahila and Davies 2002). National newspapers are available in Arabic, Amazigh, French, and Spanish; official forms and information signs are usually offered in both French and Arabic; and a stroll through any city center will confront the casual observer with shop and business signs using English, French, Arabic, and other languages in many different combinations. Trade names are sometimes formed by combining elements of two languages, while others use Arabic words transliterated in the Roman alphabet or French words transcribed in Arabic script. The overall effect is definitely of a place where multilingualism is the norm rather than a marked phenomenon.

6. FUTURE PERSPECTIVES

The coexistence of Arabic and other languages can be traced to contacts between different language communities, some very long-standing and others quite recent. Although firm predictions about future language situations are difficult to make, it is safe to say that languages will continue to be used as long as they are perceived to have a specific value for their users. This value may be pragmatic, as when a language is a key to employment, education, or social advancement; but it may equally be symbolic, as when a language is valued for what it represents: cultural heritage, religion, and minority group identity. Governments across the Arab world have made huge efforts to strengthen the position of Arabic in relation to both indigenous and colonizers' languages. Nevertheless, languages such as Kurdish and Amazigh have continued to defy their titanic neighbor, Arabic. There are also signs that numbers of users and/or domains of use are likely to increase for these languages in the future. On the other hand, despite determined attempts to root out English and French, the eradication of these languages from everyday life in the Arab world in the near future is very unlikely, due to their instrumental value in offering access to the West, hated or loved. The Arab world's currently multilingual societies will most likely maintain the use of other languages together with Arabic for the forseeable future.

Bibliographical references

Abou Sall, Ibrahima. 2001. "Maurétanie: L'épuration ethnique". Lecture given in Brussels, 28 April 2001. Available at http://members.lycos.co.uk/flamnet/IbrahimaSall_conf2804.html.

Altena, Nelleke and René Appel. 1982. "Mother tongue teaching and the acquisition of Dutch by Turkish and Moroccan immigrant workers' children". *Journal of Multilingual and Multicultural Development* 3.315–322.

Arnold, Werner. 2000. "The Arabic dialects in the Turkish province of Hatay and the Aramaic dialects in the Syrian mountains of Qalamun: Two minority languages compared". Owens (2000:347–370).

Atawneh, Ahmad. 1992. "Code mixing in Arabic-English bilinguals". *Perspectives on Arabic linguistics*, IV, ed. Ellen Broselow, Mushira Eid, and John McCarthy, 219–241. Amsterdam and Philadelphia: J. Benjamins.

Belazi, Hedi M. 1992. *Multilingualism in Tunisia and French/Arabic code switching among educated Tunisian bilinguals*. Ph.D. diss., Cornell University.

Bentahila, Abdelali. 1983. *Language attitudes among Arabic-French bilinguals in Morocco*. Clevedon: Multilingual Matters.

Bentahila, Abdelali and Eirlys E. Davies. 1983. "The syntax of Arabic-French code-switching". *Lingua* 59.301–330.

——. 1991. "On the evaluation of Arabic language and culture lessons: Another perspective". *Samenwijs* 11.186–187.

——. 1992a. "Code-switching and language dominance". *Cognitive processing in bilinguals*, ed. Richard J. Harris, 443–458. Amsterdam: North Holland.

——. 1992b. "Convergence and divergence: Two cases of language shift in Morocco". *Maintenance and loss of minority languages*, ed. Willem Fase, Koen Jaspaert, and Sjaak Kroon, 197–210. Amsterdam and Philadelphia: J. Benjamins.

——. 1992c. "Learning a 'home' language abroad: The case of Moroccan children in the Netherlands". *La recherche scientifique au service du développement: Actes de la Troisième Rencontre Universitaire Maroco-nerlandaise*, 69–84. (= Série Colloques et Séminaires No. 39.) Rabat: Université Mohamed V Publications.

——. 1993. "Code switching by Moroccan immigrant children: A stepping stone to fluency". *L'emigration maghrébine vers l'Europe: Cahiers du Centre d'Etudes sur les Mouvements Migratoires Maghrébins* 4.41–51.

——. 1994. "Two languages, three varieties: Code-switching patterns of bilingual children". *The cross-linguistic study of bilingual development*, ed. Guus Extra and Ludo Verhoeven, 113–128. Amsterdam: North Holland.

——. 1995. "Patterns of code switching and patterns of language contact". *Lingua* 96.75–93.

——. 1998. "Codeswitching: An unequal partnership?". *Codeswitching worldwide*, ed. Rodolfo Jacobson, 25–49. Berlin: Mouton de Gruyter.

——. 2002. "Language mixing in rai music: Localization or globalization?". *Language and Communication* 22.187–207.

Bidwell, R. 1973. *Morocco under colonial rule*. London: F. Cass.

Boumans, Louis and Dominique Caubet. 2000. "Modelling intrasentential codeswitching: A comparative study of Algerian/French in Algeria and Moroccan/Dutch in the Netherlands". Owens (2000:113–180).

Brahimi, Fadila and Jonathan Owens. 2000. "Language legitimization: Arabic in multiethnic contexts". Owens (2000:405–446).

Carvin, Andy. 2001. "Website language stats". *Papyrus News*, April 16, 2001. Available at www.gse.iuci.edu/markw/papyrus-news.html.

Clyne, Michael and Sandra Kipp. 1999. *Pluricentric languages in an immigrant context*. Berlin: Mouton de Gruyter.

Crystal, David. 1997. *English as a global language*. Cambridge: Cambridge University Press.

Davies, Eirlys E. and Abdelali Bentahila. 1989. "On mother and other tongues: The notion of possession of a language". *Lingua* 78.267–293.

Eid, Mushira. 1992. "Directionality in Arabic-English code-switching". *The Arabic language in America*, ed. Aleya Rouchdy, 50–71. Detroit: Wayne State University Press.

Extra, Guus and Durk Gorter (eds.). 2001. *The other languages of Europe: Demographic, sociolinguistic and educational perspectives*. Clevedon: Multilingual Matters.

Ghaith, Ghazi M. and Kassim A. Shaaban. 1996. "Language-in-education policy and planning: The case of Lebanon". *Mediterranean Journal of Educational Studies* 1:2.95–105.
Gordon, Raymond G. Jr. (ed.). 2005. *Ethnologue: Languages of the world.* 15th ed. Dallas: SIL. International online edition: www.ethnologue.com.
Ḥuṣrī, Sāṭi al-. 1985. *ʾArāʾ wa-ʾaḥādīṯ fī l-qawmiyya al-ʿarabiyya*, VII. Beirut: Markaz Dirāsāt al-Waḥda al-ʿArabiyya.
Hussein, Riyad Fayez and Ghazi A. Shorrab. 1993. "Syntactic constraints on the code switching of Arabic-English bilinguals". *International Review of Applied Linguistics in Language Teaching* 31.236–241.
Laitin, David D. 1992. *Language repertoires and state construction in Africa.* Cambridge: Cambridge University Press.
Leclerc, Jacques. 2003. *Aménagement linguistique dans le monde.* Available at www.tlfq.ulaval.ca/axl.
Nortier, Jacomine. 1990. *Dutch-Moroccan Arabic code switching among Moroccans in the Netherlands.* Dordrecht: Foris.
Owens, Jonathan (ed.). 2000. *Arabic as a minority language.* Berlin: Mouton de Gruyter.
Rouchdy, Aleya (ed.). 1992. *The Arabic language in America.* Detroit: Wayne State University Press.
Samhan, Helen. 2001. "Who are Arab Americans?" *Grolier's multimedia encyclopedia.* Available at www.grolier.com.
Spolsky, Bernard. 1997. "Multilingualism in Israel". *Annual Review of Applied Linguistics* 17.138–150.
—— and Elana Shohamy. 1999. "Language in Israeli society and education". *International Journal of the Sociology of Language* 137.93–114.
—— and Elana Shohamy. 2001. "Hebrew after a century of RLS efforts". *Can threatened languages be saved?*, ed. Joshua A. Fishman, 350–363. Clevedon: Multilingual Matters.
Suleiman, Michael W. (ed.). 1999. *Arabs in America: Building a new future.* Philadelphia: Temple University Press.
Suleiman, Yasir. 2003. *The Arabic language and national identity.* Washington, D.C.: Georgetown University Press.
Talmon, Rafael. 2000. "Arabic as a minority language in Israel". Owens (2000:199–220).
Warschauer, Mark, Ghada R. El Said, and Aiman Zohry. 2002. "Language choice online: Globalization and identity in Egypt". *Journal of Computer Mediated Communication* 7.4. Available at www.ascusc.org/jcmc/vol7/issue4/warschauer.html.

Abdelali Bentahila
(Abdelmalek Essaadi University)
Eirlys E. Davies
(King Fahd School of Translation)

Munṣarif → Ṣarf

Musnad → Isnād

Muštarak

Muštarak (or, currently, *al-muštarak al-lafḏī*) is used in Arabic rhetoric and grammar to indicate the 'homonymous polysemic word' (lit. 'the common one'). The question at the origin of the lexical category of *muštarak* is 'how the nouns apply to the *nominatum* (the named things)' (see, for instance, Ibn Fāris [d. 395/1004], *Ṣāḥibī* 114, *bāb al-ʾasmāʾ kayfa taqaʿu ʿalā l-musammayāt*). In the most general and common state of affairs, every thing gets its own noun, such as *rajul* 'man', *faras* 'horse', and so on (Ibn Fāris, *Ṣāḥibī* 327), but there are also single 'things' with many nouns, which is the case of the synonymous words (→ *mutarādif*), and different 'things' with one 'common' noun, which is the case of the *muštarak*. Ibn Fāris presents this situation as a matter of fact in Arabic, as Sībawayhi (d. 180/796) had already done before him. The latter does not use the technical term which was to become current, but simply states that *ittifāq al-lafḏayn wa-xtilāf al-maʿnayayn* 'the coincidence of the two sound groups and the divergence of the two meanings' is a part of the language of the Arabs (Sībawayhi, *Kitāb* I, 24). The early collections of homonymous polysemic words designate the phenomenon in this way.

The majority of the early scholars accepted as truth the existence in Arabic of homonymous polysemic words. The grammarian Ibn Jinnī (d. 392/1002; *Xaṣāʾiṣ* III, 110–111) notes that prepositions like *min* 'from' and particles like *lā* 'no; not' have more than one meaning, and in his opinion, so do nouns (e.g. *ṣadā* 'echo; corpse; the bird [which cries in the head of the slain when his blood has not been avenged]') and verbs. Ibn Fāris, when defining the phenomenon of *ištirāk* (*Ṣāḥibī* 456, *bāb al-ištirāk*, i.e. that a *lafḏa* can have more than one meaning) does not restrict himself to nouns but rather states that syntactic structures and sentences can be homonymous, too, for example, *ʾa-raʾayta* 'did you see?'. This is a *lafḏa* with more than one meaning, since it can be used to introduce a question, or alternatively to put someone on his guard (e.g. Q. 96/11–13).

Lexicons have a tendency to collect the largest possible number of examples of words with more than one meaning, as in the well-known case of the word *ʿajūz* 'old woman', for which al-Fīrūzābādī (d. 817/1415) in his *al-Qāmūs al-muḥīṭ* records about seventy meanings, e.g. 'wine; nail in the hilt of a sword; sword blade', and so on. Another frequently quoted item is *ʿayn* 'eye', for which authorities quote a large number of meanings (see e.g. Suyūṭī [d. 911/1505], *Muzhir* I, 372–375). Many works listing homonymous polysemic words are mentioned in the ancient historical and biographical repertories. Among those which have reached us, the earliest works are ʾAbū ʿUbayd's (d. 224/838) *ʾAjnās*, ʾAbū l-ʿAmayṯal's (d. 240/855) *Mā ttafaqa lafḏu-hu*, al-Mubarrad's (d. 286/900) *Mā ttafaqa lafḏu-hu wa-xtalafa maʿnā-hu*, Kurāʿ an-Naml's (d. 310/922) *al-Munajjad*. A more complete list of ancient and modern works is found in al-Munajjid (1998:23–26; cf. also Omar 1993, who examines in detail especially the treatise by Kurāʿ an-Naml).

Both 'homonymous polysemic words' and 'synonymous words' depart from the basic principle governing the invention of speech, namely *ʾibāna* 'clarification' (Ibn Durustawayhi [d. 346/957], as quoted by Suyūṭī, *Muzhir* I, 385). They are two parts of the same question and as such, they were dealt with by the same authors. Ibn Durustawayhi is among those who deny the possibility that a language can use the same word to signify different meanings because this would not lead to the intended clarity, which is the aim of language, but rather to obscurity and blindness. He holds that words which seem to have different meanings have in fact one meaning only. *Wajada* 'he found', for instance, which is a particularly authoritative example since it is the example of a homonymous word quoted by Sībawayhi (*wajadtu ʿalayhi* 'I was angry with him' and *wajadtu* 'I found', e.g. *aḍ-ḍālla* 'a lost [animal, or other thing]'), only has the meaning of 'hitting upon' or 'reaching' (*ʾiṣābatu-hu*) something, which may be good or evil (Ibn Durustawayhi, *Taṣḥīḥ* I, 364; Suyūṭī, *Muzhir* I, 384). In his view, homophonic words with a different meaning may exist either in two different languages, or in one language as the result of a process of ellipsis or elision, which led the two words to coalesce. In the same way, ʾAbū ʿAlī al-Fārisī (d. 395/1004) states that homonymy cannot be the aim nor the original condition of the invention of speech, although it can come about owing to contact between different languages (or of different dialects in the case of Arabic, *luġātin tadāxalat*), or to metaphorical use (Ibn Sīda [d. 458/1066], *Muxaṣṣaṣ* XIII, 258–259; and also Ibn as-Sarrāj [d. 316/928], *Ištiqāq* 32).

As-Suyūṭī (*Muzhir* I, 374–375) also remarks that almost all meanings of *ʿayn*, such as 'source; the choice, or best [of a thing]; money; spy', represent metaphorical or metonymical uses (*tašbīh*) of the basic meaning ('eye'), or can be derived from it. The same idea is found in works such as *Kitāb istiʿāra ʾaʿḍāʾ al-ʾinsān* by Ibn Fāris, which deals with the metaphorical meaning and use, especially in poetry, of terms denoting the parts of the human body.

Traditional legal theorists (*ʾuṣūliyyūn*), too, had to deal with the question of the *muštarak*, as they did with the *mutarādif* 'synonyms', because, as Weiss (1984:21) points out, they believed that "certain things must be established in the realm of language before one can proceed to interpret or to deduce law". Not surprisingly, their definition of this phenomenon is much more accurate and detailed than that of the philologists. The definition of *ʾahl al-ʾuṣūl* is the only one reported by as-Suyūṭī (*Muzhir* I, 369). In particular, they take care to exclude from their definitions the possibility to consider a word *muštarak*, if its two (or more) meanings comprise one proper meaning and in addition metaphorical ones, or if they do not refer to entities that differ in themselves. The word *lawn* 'color', for instance, which may refer to redness, blackness, or whiteness, cannot be considered *muštarak*, because it is applied to these meanings not in order to indicate entities that differ in themselves but because of their common character. This word is not a case of *muštarak* but rather of *mutawāṭiʾ* 'referring to different things in the same way' (see e.g. Ġazālī [d. 505/1111], *Mustaṣfā* 42–43).

The linguistic thinking of the traditional legal theorists on the question of the *muštarak* deals with several aspects: its definition, as we have seen; the possibility of its occurrence from a rational point of view; and its actual occurrence and the reasons for its occurrence. Those who maintain that the occurrence of the *muštarak* is rationally necessary argue that 'words' are finite (*mutanāhiya*), while 'meanings' are, on the contrary, infinite. They also believe that languages need some general words, such as 'thing'. Moreover, they point out that words that are *muštarak* do occur in fact, either because two identical words with different meanings in different Arabic dialects merged (an explanation often advanced by the philologists as well), or because one of the two meanings is metaphorical, or even because ambiguity may sometimes be the aim of language, when explicitness or completion can be a cause of evil (Rāzī [d. 606/1209], *Maḥṣūl* I, 97–99; Suyūṭī, *Muzhir* I, 369). In this respect, we owe to Ibn Durayd (d. 321/933) a book (*Malāḥin*) containing a list of 184 words which have both a well-known meaning and another, much more unusual or rare one. One may resort to the latter when one is forced to take an oath one does not want to take.

They also remark that if we take into consideration not only 'nouns' (*ʾasmāʾ*) but also particles and verbs, we find that the phenomenon of *ištirāk* 'sharing', i.e. the existence of homophonic polysemic words, is actually more widespread than *infirād* 'singleness', i.e. the existence of words with one meaning only (Rāzī, *Maḥṣūl* I, 106). The legal theorists believe that the basic condition is *infirād*, rather than *ištirāk*. The latter is a matter of probability (*iḥtimāl, rujḥān*), when it is impossible to ascertain, for instance in a passage of the *Qurʾān* or the *Sunna*, which single meaning the word concerned has.

Those who deny the possibility of the occurrence of *muštarak* words base their arguments mainly on the ambiguity which necessarily results from using homophonic polysemic words and which leads to misunderstandings and errors. This ambiguity cannot have been the aim of the 'inventor' of the language (Rāzī, *Maḥṣūl* I, 106–107).

The traditional legal theorists also dealt with the question of the possibility that a *muštarak* word could be used in all its meanings at the same time, as seems to happen in some Qurʾānic passages, for instance Q. 22/18, where *yasjudu* 'he bows down' refers not only to human beings but also to the moon, stars, mountains, trees, and animals. Ar-Rāzī's answer (*Maḥṣūl* I, 101–105) is that this would be inadmissible, because these words are invented for one of the two or more meanings separately, not for all of them together.

The linguistic thinking of the *ʾuṣūliyyūn*, which shows many similarities to that of logicians, has been analyzed in several studies; for a recent work, see al-Kafrāwī (2002).

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

ʾAbū l-ʿAmayṯal, *Mā ttafaqa lafḏu-hu* = ʿAbdallāh ibn Xulayd ʾAbū l-ʿAmayṯal, *Kitāb mā ttafaqa lafḏu-hu wa-xtalafa maʿnā-hu*. Ed. Fritz Krenkow. London, 1925.

ʾAbū ʿUbayd, *ʾAjnās* = ʾAbū ʿUbayd al-Qāsim ibn Sallām al-Harawī, *Kitāb al-ʾajnās min kalām al-ʿArab wa-mā štabaha fī l-lafḏ wa-xtalafa fī l-maʿnā*. Ed. Imtiyāz ʿAršī ar-Rāmfūrī. Bombay, 1938.

Fīrūzābādī, *al-Qāmūs al-muḥīṭ* = Majd ad-Dīn Muḥammad ibn Yaʿqūb al-Fīrūzābādī, *al-Qāmūs al-muḥīṭ*. 4 vols. Cairo: Muṣṭafā al-Bābī al-Ḥalabī, 1952.

Ġazālī, *Mustaṣfā* = ʾAbū Ḥāmid Muḥammad ibn Muḥammad al-Ġazālī, *al-Mustaṣfā min ʿilm al-ʾuṣūl*. Ed. Muḥammad Muṣṭafā ʾAbū l-ʿAlā. Cairo: Šarikat aṭ-Ṭibāʿa al-Fanniyya al-Muttaḥida, 1971.

Ibn Durayd, *Malāḥin* = ʾAbū Bakr Muḥammad ibn al-Ḥasan Ibn Durayd al-ʾAzdī, *Kitāb al-malāḥin*. Ed. Heinrich Thorbecke. Heidelberg, 1882; ed. ʿAbdallāh Nabhān. Damascus: Wizārat aṯ-Ṯaqāfa, 1992.

Ibn Durustawayhi, *Taṣḥīḥ* = Ibn Durustawayhi, *Taṣḥīḥ al-faṣīḥ*. Ed. ʿAbdallāh al-Jabbūrī. Baghdad, 1975.

Ibn Fāris, *Kitāb istiʿāra ʾaʿḍāʾ al-ʾinsān* = ʾAbū l-Ḥasan ʾAḥmad Ibn Fāris ibn Zakariyyā, *Kitāb istiʿāra ʾaʿḍāʾ al-ʾinsān*. Ed. ʾAḥmad Xān, *al-Mawrid* 12 (1983). 81–108.

——, *Ṣāḥibī* = ʾAbū l-Ḥusayn ʾAḥmad Ibn Fāris ibn Zakariyyā, *aṣ-Ṣāḥibī fī fiqh al-luġa*. Ed. ʾAḥmad Ṣaqr. Cairo: ʿĪsā al-Bābī al-Ḥalabī wa-Šurakāʾu-hu, 1977.

Ibn Jinnī, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. 3 vols. Ed. Muḥammad ʿAlī an-Najjār. Beirut: Dār al-Hudā li-ṭ-Ṭibāʿa wa-n-Našr, n.d.

Ibn as-Sarrāj, *Ištiqāq* = ʾAbū Bakr ibn as-Sarī Ibn as-Sarrāj, *Risālat al-ištiqāq*. Ed. Muḥammad ʿAlī ad-Darwīš and Muṣṭafā al-Ḥadrī. Damascus: Dār Majallat aṯ-Ṯaqāfa, 1973.

Ibn Sīda, *Muxaṣṣaṣ* = ʾAbū l-Ḥasan ʿAlī ibn ʾIsmāʿīl Ibn Sīda, *Kitāb al-Muxaṣṣaṣ*. 14 vols. Būlāq: al-Maṭbaʿa al-ʾAmīriyya, 1316–1321 A.H.

Kurāʿ an-Naml, *al-Munajjad* = ʾAbū l-Ḥasan ʿAlī ibn al-Ḥasan Kurāʿ an-Naml, *al-Munajjad fīmā ttafaqa lafḏu-hu wa-xtalafa maʿnā-hu*. Ed. ʾAḥmad Muxtār ʿUmar and Ḍāḥī ʿAbd al-Bāqī. Cairo, 1988.

Mubarrad, *Mā ttafaqa lafḏu-hu wa-xtalafa maʿnā-hu* = ʾAbū l-ʿAbbās Muḥammad ibn Yazīd al-Mubarrad, *Mā ttafaqa lafḏu-hu wa-xtalafa maʿnā-hu min al-Qurʾān al-majīd*. Ed. ʿAbd al-ʿAzīz al-Maymanī. Cairo: al-Maṭbaʿa as-Salafiyya, 1931.

Rāzī, *Maḥṣūl* = Faxr ad-Dīn Muḥammad ibn ʿUmar ibn al-Ḥusayn ar-Rāzī, *al-Maḥṣūl fī ʿilm ʾuṣūl al-fiqh*. 3 vols. Beirut: Dār al-Kutub al-ʿIlmiyya, 1988.

Sībawayhi, *al-Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān ibn Qanbar Sībawayhi, *al-Kitāb*. Ed. ʿAbd as-Salām Muḥammad Hārūn. 4 vols. Cairo: Maktabat al-Xānjī, 1968–1977.

Suyūṭī, *Muzhir* = Jalāl ad-Dīn ʾAbū l-Faḍl ʿAbd ar-Raḥmān ibn ʾAbī Bakr as-Suyūṭī, *al-Muzhir fī ʿulūm al-luġa*. Ed. Muḥammad ʾAḥmad Jād al-Mawlā,

ʿAlī Muḥammad al-Bigāwī, and Muḥammad ʾAbū l-Faḍl ʾIbrāhīm. 2 vols. Cairo: Maṭbaʿa ʿĪsā al-Ḥalabī, n.d.

Secondary sources

Kafrāwī, ʾAsʿad ʿAbd al-Ġanī as-Sayyid al-. 2002. *al-Istidlāl ʿinda l-ʾuṣūliyyīn*. Cairo: Dār as-Salām.

Munajjid, Muḥammad Nūr ad-Dīn al-. 1998. *al-Ištirāk al-lafḏī fī l-Qurʾān al-karīm bayna n-naḏariyya wa-t-taṭbīq*. Damascus: Dār al-Fikr.

Omar, Ahmed Mokhtar. 1993. “Early Arabic lexicons of homophonic words”. *Proceedings of the Colloquium on Arabic Lexicology and Lexicography*, ed. Kinga Dévényi, Tamás Iványi, and Avihai Shivtiel. *Budapest Studies in Arabic* 6-7.3–11.

Ramaḍān, ʿAbd at-Tawwāb. 1997. *Fuṣūl fī fiqh al-luġa*. Cairo: Maktabat al-Xānjī.

Šāhīn, Tawfīq Muḥammad. 1980. *al-Muštarak al-luġawī naḏariyyan wa-taṭbīqan*. Cairo.

Weiss, Bernard G. 1984. “Language and law: The linguistic premises of Islamic legal science”. *In quest of an Islamic humanism: Arabic and Islamic studies in memory of Mohamed al-Nowaihi*, ed. Arnold H. Green, 15–21. Cairo: American University Press.

Lidia Bettini (University of Firenze)

Mutaʿaddin → Taʿaddin

Muʿtall → ʿIlla

Mutamakkin → Tamakkun

Mutarādif

The term *mutarādif* means ‘synonym’; in addition to this technical term, other expressions are found in Arabic medieval works: *muwāfiq li-* ‘corresponding to’, *makāna* ‘in the place of’ (Gully 1994:38–39).

The question of synonymy was dealt with by medieval Muslim scholars from both a practical and a theoretical point of view. Of the former type are works such as al-ʾAṣmaʿī’s (d. 213/828) *Mā xtalafa ʾalfāḏu-hu wa-ttafaqat maʿānī-hi* ‘A collection of synonymous words and expressions’, a sort of mirror image of works bearing titles such as *Mā ttafaqa lafḏu-hu wa-xtalafa maʿnā-hu* ‘A collection of homonymous polysemic words’, for instance by al-Mubarrad (d. 286/900). Both treatises show a different side of the lexicographical activity of the Arab philologists, who made an effort to collect and classify the copious lexical heritage of the Classical Arabic language. In these works, homonymous and synonymous words are accompanied by poetical or Qurʾānic *šawāhid* ‘textual evidence’, and authors are not concerned with a theoretical definition of ‘synonymy’ (*tarāduf*). It is at any rate noteworthy, although not surprising, that one and the same author (e.g. Ṯaʿlab [d. 291/904]) does not admit theoretically the existence of synonymy, but at the same time draws up lists of synonymous words (e.g. Ṯaʿlab, *Majālis* 101; cf. Suyūṭī [d. 911/1505], *Muzhir* I, 411, 412). Lists of synonymous terms are found also in Ibn as-Sikkīt (d. *ca.* 244/858), ar-Rummānī (d. 384/994), and others (Schulz 1994b:248, 252).

Theoretical questions about synonymy occupied not only philologists but also theologians and literary critics. They all have the same point of departure, the notion of → *ism* ‘noun’, considered in its semantic-philosophical connection with its ontological correspondent, the *musammā* ‘nominatum’ (lit. ‘the named one’), or the *ḏāt* ‘essence [of the thing]’. According to the grammarian ʾAbū ʿUṯmān al-Māzinī (d. ca. 249/863; cf. ʿAskarī [d. 395/1004], *Furūq* 20), in fact, two kinds of ‘noun’ exist: the *ism maḥḍ* which indicates its nominatum ‘as a gesture’, and the *ism ṣifa* ‘descriptive epithet’, which connotes or signifies it. This important distinction influenced the thinking of all those who were interested, for various reasons, in defining synonymy rather than just showing a rich choice of similar words to be used for literary or lexicographical purposes.

According to Ibn Fāris (d. 395/1004; *Ṣāḥibī* 114–116) and as-Suyūṭī (*Muzhir* I, 402–407), two opposite views exist toward synonymy: the point of view of those who think that different words with the same meaning do exist, and the point of view of the others, who affirm, as Ibn Fāris himself does, that in fact every different word has a different meaning. The first and more general way of distinguishing between words appearing to have the same meaning is to apply the distinction between noun and attribute at a grammatical and formal level. For instance, for *sayf* ‘sword’, which can be designated by items such as *muhannad*, *ṣārim*, and others, the only *ism* is *sayf*, the other designations being *ṣifāt* (‘the Indian one‘, ‘the sharp one’, and so on; cf. Suyūṭī, *Muzhir* I, 405). Yet, at this level, the distinction barely touches the question of whether synonymy occurs in the Arabic language and it needs to be specified. The theologian al-Ġazālī (d. 505/1111; *Maqṣad* 18–24) deals with semantic questions in connection with the main issue, the relation between the essence of God and His attributes *(ṣifāt)*. He argues that two words are 'synonymous' only when they refer to 'one and the same thing', as for instance *layṯ* and *ʾasad* for 'lion' (*Maqṣad* 21). As for *ṣifāt*, al-Ġazālī distinguishes them from the *ism*, not from the point of view of their grammatical function but as far as their relation with the 'thing' they describe is concerned. When many epithets refer to the same 'noun' (as in the case of the attributes of God), al-Ġazālī locates the relation existing among them in the 'unity of the substrate' (*ittiḥād al-maḥall; Maqṣad* 22, 24). Just like Faxr ad-Dīn ar-Rāzī (d. 606/1209; cf. Suyūṭī, *Muzhir*, 402–403), al-Ġazālī introduces the distinction between *ism* and *ṣifa* in the definition of 'synonymous' nouns in the form of the 'point of view, way of considering' (*iʿtibār*). According to them, 'two nouns are synonymous when they denote the same thing from the same point of view'.

The approach of the philologists moves within an entirely linguistic framework, and the criteria they introduce for distinguishing among the meanings of words are not dependent on their relation to the ontological reality. For Ṯaʿlab (d. 291/904; cf. Ibn Fāris, *Ṣāḥibī* 114–115) and his pupil Ibn Fāris, two different words, no matter if they are *ism* or *ṣifa*, cannot have exactly the same meaning. These scholars also take into consideration verbs (e.g. *jalasa/qaʿada* 'to sit', or *ḏahaba/inṭalaqa* 'to go'). According to al-ʿAskarī (*Furūq* 13), al-Mubarrad does not count two words as synonymous if they can be joined by a conjunction: this is impossible unless the two words have a different meaning. Criteria for establishing whether two words have the same meaning were advanced also by Ibn as-Sarrāj (d. 316/928; *Ištiqāq*, 39–41): for instance, if the two words have the same antonyms, or if they can be qualified by the same attributes. Al-ʿAskarī's *Furūq* is entirely devoted to the 'differences' between similar words; he takes into account, for instance, the prepositions and the attributes which are used with the nouns in question, or their etymological meaning, or, for verbal nouns, the form (radical or derived) of the verb to which they belong (*Furūq* 13–19). Similar criteria (syntactic construction, antonym) are put forward by al-Xaṭṭābī (d. 388/998), a traditionist mindful of linguistic issues (*Bayān* 29–33).

Since semantic questions occupied the attention of traditional legal theorists (*ʾuṣūliyyūn)* a good deal (Weiss 1984:15–21), not surprisingly one finds in al-ʾĀmidī (d. 631/1233, *ʾIḥkām* I, 30–33) a detailed account of the arguments put forward by those who deny the existence of synonymy. The four arguments are as follows. First, it is useless to have more than one *ism* to designate a unique *musammā* 'nominatum'. Second, the prevalent condition being that multiple nouns correspond to multiple nominatums, it can be argued that this was the intention of the inventor of language. Moreover, what is contrary to the prevalent situation is contrary also to the basic rule. Third, to learn a single noun is less toilsome than to learn two or more of them. Finally, if nouns are multiple (for the same thing), it becomes necessary to learn either all of them or only some, at the risk that other people know the nouns one has neglected to learn. Al-ʾĀmidī refutes these arguments one by one and points out, as his senior contemporary Faxr ad-Dīn ar-Rāzī did, that nouns are not synonymous when they refer to the same object (*mawḍūʿ*) from the point of view (*iʿtibār*) of its *ṣifāt*, as for instance *sayf* and *ṣārim* (*ʾIḥkām* 33).

This kind of abstract argument contrasts with the 'sociolinguistic' considerations of the grammarian Ibn Jinnī (d. 392/1002) about the question of synonymy. Like other scholars, he points out the possibility that two 'synonyms' can originate from different nouns in the dialects of two different Arab tribes (cf. Suyūṭī, *Muzhir* I, 405–406), but he adds that they must be heard by the same hearer to be counted as 'synonyms', i.e., they must exist in the same linguistic variety at the same time (*Xaṣāʾiṣ* I, 373; on this and other aspects of Ibn Jinnī's views on synonymy, see Schulz 1994a, 1994b).

Bibliographical references

Primary sources

ʾĀmidī, *ʾIḥkām* = ʾAbū l-Ḥasan ʿAlī ibn ʾAbī ʿAlī al-ʾĀmidī, *al-ʾIḥkām fī ʾuṣūl al-ʾaḥkām*. 4 vols. Cairo: Maṭbaʿat al-Maʿārif, 1914.

ʿAskarī, *Furūq* = ʾAbū Hilāl al-Ḥasan ibn ʿAbdallāh al-ʿAskarī, *al-Furūq al-luġawiyya*. Ed. Lajna ʾIḥyāʾ at-Turāṯ al-ʿArabī. Beirut: Dār al-ʾĀfāq al-Jadīda, 1980.

ʾAṣmaʿī, *Mā xtalafa* = ʾAbū Saʿīd ʿAbd al-Malik ibn Qurayb al-ʾAṣmaʿī, *Mā xtalafa alfāḏu-hu wa-tta-*

faqat maʿānī-hi. Ed. Mājid aḏ-Ḏahabī. Damascus: Dār al-Fikr, 1986.

Ġazālī, *Maqṣad* = ʾAbū Ḥāmid Muḥammad ibn Muḥammad al-Ġazālī, *al-Maqṣad al-ʾasnā fī šarḥ maʿānī ʾasmāʾ Allāh al-ḥusnā*. Ed. Fadlou A. Shehadi. Beirut: Dar al-Machreq, 1971.

Ibn Fāris, *Ṣāḥibī* = ʾAbū l-Ḥusayn ʾAḥmad Ibn Fāris, *aṣ-Ṣāḥibī*. Ed. ʾAḥmad Ṣaqr. Cairo: ʿĪsā al-Bābī al-Ḥalabī, 1977.

Ibn Jinnī, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. 3 vols. Ed. Muḥammad ʿAlī an-Najjār. Beirut: Dār al-Hudā, n.d.

Ibn as-Sarrāj, *Ištiqāq* = ʾAbū Bakr Muḥammad ibn as-Sarī Ibn as-Sarrāj, *Risālat al-ištiqāq*. Ed. Muḥammad ʿAlī ad-Darwīš and Muṣṭafā al-Ḥadrī. Damascus: Dār Majallat aṯ-Ṯaqāfa, 1973.

Mubarrad, *Mā ttafaqa* = ʾAbū l-ʿAbbās Muḥammad ibn Yazīd al-Mubarrad, *Mā ttafaqa lafḏu-hu wa-xtalafa maʿnā-hu min al-Qurʾān al-majīd*. Ed. ʿAbd al-ʿAzīz al-Maymanī. Cairo: al-Maṭbaʿa as-Salafiyya, 1350 A.H.

Rummānī, *ʾAlfāḏ* = ʾAbū l-Ḥasan ʿAlī ibn ʿĪsā ar-Rummānī, *Kitāb al-ʾalfāḏ al-mutarādifa al-mutaqāribat al-maʿnā*. Ed. Fatḥallāh Ṣāliḥ. al-Manṣūra, 1987.

Suyūṭī, *Muzhir* = Jalāl ad-Dīn ʾAbū l-Faḍl ʿAbd ar-Raḥmān ibn ʾAbī Bakr as-Suyūṭī, *al-Muzhir fī ʿulūm al-luġa wa-ʾanwāʿi-hā*. 2 vols. Ed. Muḥammad ʾAḥmad Jād al-Mawlā, ʿAlī Muḥammad al-Bijāwī, and Muḥammad ʾAbū l-Faḍl ʾIbrāhīm. Cairo: ʿĪsā al-Bābī al-Ḥalabī, n.d.

Ṯaʿlab, *Majālis* = ʾAbū l-ʿAbbās ʾAḥmad ibn Yaḥyā Ṯaʿlab, *Majālis*. Ed. ʿAbd as-Salām Muḥammad Hārūn. Cairo: Dār al-Maʿārif, n.d.

Xaṭṭābī, *Bayān* = ʾAbū Sulaymān Ḥamd Muḥammad al-Xaṭṭābī, *Bayān ʾiʿjāz al-Qurʾān*: *Ṯalāṯ rasāʾil fī ʾiʿjāz al-Qurʾān*. Ed. Muḥammad Xalafallāh and Muḥammad Zaġlūl Sallām. Cairo: Dār al-Maʿārif, 1976.

SECONDARY SOURCES

Bettini, Lidia. 1983. "Sulla nozione di 'sinonimo' nella tradizione linguistica araba" [On the notion of synonymy in the Arabic linguistic tradition]. *Quaderni di Studi Arabi* 1.29–40.

Gully, Adrian. 1994. "Synonymy or not synonymy, that is the question: The case of the particle in medieval Arabic". *Zeitschrift für arabische Linguistik* 27.36–46.

Schulz, Eckehard. 1989. "Lexikalische Synonymie im Arabischen". *Studien zur arabischen Linguistik*, ed. Dieter Blohm, 69–80. Berlin: Akademie der Wissenschaften.

——. 1990. "Synonyme und Synonymie bei Abū Manṣūr ath-Thaʿālibī". *Orientalische Philologie und arabische Linguistik*, ed. Wolfgang Reuschel, 95–98. Berlin: Akademie-Verlag.

——. 1994a. "Die 'falschen Synonyme' im klassischen Arabisch". *XXV. Deutscher Orientalistentag, Vorträge*, ed. Cornelia Wunsch, 95–107. Stuttgart: F. Steiner.

——. 1994b. "Ibn Ǧinnīs *al-Ḫaṣāʾiṣ* und das Problem der Synonymie". *Gedenkschrift Wolfgang Reuschel*, ed. Dieter Bellmann, 247–254. Stuttgart: F. Steiner.

Weiss, Bernard G. 1984. "Language and law: The linguistic premises of Islamic legal science". *In quest of an Islamic humanism: Arabic and Islamic studies in memory of Mohamed al-Nowaihi*, ed. Arnold H. Green, 15–21. Cairo: American University Press.

LIDIA BETTINI (University of Firenze)

Muṭāwaʿa → Middle Verbs

Muṭbaq → ʾIṭbāq

N

Naḥt → Compound

Naḥw → Grammatical Tradition (Approach)

Najdi Arabic

1. GENERAL

This entry provides a description of the Arabic dialect of Najd in Central Arabia. Because this is a large area, showing considerable linguistic diversity, the description is based on the dialect of Sudayr in the central area, mentioning variation in the northern and southern areas where relevant.

1.1 *Area*

Geographically, the Najdi dialects span the following regions:

i. The speech of the sedentary population of the areas of Central Najd and of Qaṣīm and Jabal Šammar to the north and Najrān and Bīša to the south.
ii. The speech of the main Bedouin tribes of those regions, i.e. ʿAnizah, ʿUtaybah, Subayʿ, Suhūl, Bugūm, Dawāsir, Ḥarb, Muṭayr, ʿAwāzim, and Rašāyidah in the center, Šammar and Ḏ̣afīr in the north, and Ghaṭān, Āl Murrah, and ʿIjmān in the south and east.
iii. The speech of the emigré Bedouin tribes of the Syrian desert and the Jazīrah of Iraq of ʿAnizah and Šammar extraction.

The geographical core of the dialect area can be seen to be the sedentary speech of Central Najd and Jabal Šammar, the dialects of the Bedouin being an overspill into the surrounding area. These dialects can be divided into subgroups by linguistic criteria as follows:

i. Central Najdi. The dialects of Central Najd and associated Bedouin tribes, also the ʿAnizah of the Syrian desert.
ii. The Northern Najdi. The dialect of Jabal Šammar and of the Šammar tribes of northern Najd and the Jazīrah.
iii. The Mixed Northern-Central. The dialect of Qaṣīm and of the Ḏ̣afīr tribe.
iv. Southern. The dialect of Najrān and the Ghaṭān tribe of the south and of the Āl Murrah and ʿĀjmān tribes of the east.

The southern group is also linked to the dialects of Yemen by syntactic and lexical features.

1.2 *Linguistic type*

The dialects are archaic, as noted by earlier travelers such as Burckhardt, who remarks (1831:372–373): “The bedouins also agree in using, as common, many select words, which in the towns would be called ‘literal terms’…and in speaking with grammatical accuracy”. This impression arises from the preservation of nunation, of a particle *ǵid* or *gid* (a reflex of Old Arabic *qad*), and of the internal passive system.

1.3 *State of research*

General descriptions of dialects of this group have been available since the early 20th century. More recently, an interest in the oral literature of the area has added to our knowledge (see Sowayan 1992; Kurpershoek 1994).

2. LINGUISTIC DESCRIPTION

The main characteristics of the dialect are given below.

2.1 *Phonology*

The inventory is as follows:

2.1.1 Consonants

b	voiced bilabial plosive
w	voiced labiovelar continuant
f	voiceless labiodental fricative (often pronounced rather laxly and sometimes voiced)
ḏ	voiced dental fricative
ṯ	voiceless dental fricative
d	voiced dental plosive
t	voiceless dental plosive
z	voiced alveolar fricative
s	voiceless alveolar fricative
ǵ	voiced alveolar affricate
ć	voiceless alveolar affricate
ḏ̣	pharyngealized voiced interdental fricative
ṭ	pharyngealized voiceless or voiced dental plosive; often sounds like the realization of *ḍād* in other dialects
ṣ	pharyngealized voiceless alveolar fricative
y	voiced palatal continuant
j	voiced palatal plosive or palato-alveolar affricate
š	voiceless palato-alveolar fricative
g	voiced velar plosive
k	voiceless velar plosive
ġ	voiced uvular fricative sometimes pronounced plosive in the north, when initial
q	voiced uvular plosive occurring in Classical borrowings equivalent to *g* above
x	voiceless uvular fricative
ʿ	voiced pharyngeal continuant
ḥ	voiceless pharyngeal fricative
h	voiceless glottal fricative

2.1.2 Vowels

ī ū
i u
ē (or ay) a ō (or aw)
ā

The phonetic realization of the vowels is as follows:

ā	Does not show the marked → *ʾimāla* or 'fronting' in some environments familiar from some dialects.
a	Generally central in quality. In the environment of the pharyngeals /ḥ/ and /ʿ/, it has a closer pronunciation near to [ə], as shown in *ʿamm* [ʕəmː] 'paternal uncle' and *ḥamlin* [ħəmlin] 'a burden'.
i/u	May have a schwa-like quality in neutral environments, such as in the neighborhood of /r, h, ḥ, ʿ, x, ġ/, as in *xirzah* 'bead'.
ē/ō	Pure vowel realizations in most environments, but with a gliding pronunciation preceding a plosive: [leit] 'would that', [zoud] 'extra'. Contrast [ħeːl] 'very much', [zeːn] 'good', [zoːl] 'figure in the distance', [ʃoːr] 'advice'.

The distinction between final *-a* and *-ah*, which is lost in many dialects, is retained in Najdi, with the final *-h* being pronounced, as in *xirzah* 'bead'. Final *-a* may join with a following nonstressed particle or monosyllabic word, as in *mā linā-bih fāydah* 'we have no use for it', *mā lhā-dwa* 'there is no cure for it'. In the southern dialect of the Āl Murrah, the final *-a* is regularly long, as in *zargā* 'blue [fem.]', *ʿilwā* 'upward; southwesterly', *ḥadrā* 'downward; northeasterly'.

2.1.3 Phonological processes

Certain phonological processes are important in the derivation of affixed words. These are:

i. Anaptyxis
Anaptyctics are connected with (a) words ending in a consonant cluster and (b) the junction of a word with final long syllable CvvC- or CvCC- and consonant-

initial suffixes. Type (a) involves clusters in which the second element is one of the voiced continuants, i.e. *r, l, w, y*, and *n*: *maṣur* 'Egypt', *rajil* 'husband', *baduw* 'Bedouin', *ṭiliy* 'lamb', *firin* 'oven'. Contrast *barg* 'lightning', *ʿišb* 'grass', *ḥarb* 'war', *xabṭ* 'colliding with', *zand* 'bolt of a rifle'. Type (b) involves such examples as *galb-a-ha* 'her heart', *bēt-i-hum* 'their [masc.] house', *šāf-i-na* 'he saw us', *šifti-kum* 'I saw you [pl. masc.]'. These vowels are very unstable, especially if the consonants on either side are voiceless. In that case they are not always present, giving *šift-kum* 'I saw you [pl. masc.]', *bēt-hum* 'their [masc.] house'.

ii. The relationship of short vowels to syllable structure

The relationship of short vowels to syllable structure produces high short vowels /i/ or /u/ in nonfinal open syllables and a low vowel /a/ in closed syllables, as in *kitab* 'he wrote', *ktibat* 'she wrote', *glubat* 'she overturned', *gallab* 'he overturned repeatedly', *gallubat* 'she overturned repeatedly', *samʿ* 'he heard', *samʿat* 'she heard', *samʿitih* 'she heard him'. This general rule is constrained by the influence of the guttural consonants (see below, Sec. 2.1.3.iii), which, in most environments, require an /a/ vowel preceding or following them in open syllables, giving *ḥamal* 'he carried', not *ḥimal*, and by the influence of a group of apical voiced continuants, /l, n, r/ and also /w/. This gives such forms as *ḥalaf* 'he swore', *xalaṭ* 'he mixed', *ġamaz* 'he winked', *sarad* 'he related', *lawa* 'he twisted', *tala* 'he followed', *bana* 'he built'.

iii. The influence of the guttural group on syllable structure

The guttural group /x, ġ, ḥ, ʿ, h/ block the application of the vowel-raising rule mentioned above under ii. and do not occur in syllable-final position after /a/, giving *ḥafar* 'he dug' and *yḥafir* 'he digs'.

iv. The elision of the short high vowel /i, u/ in a sequence of open syllables

In forms where the addition of affixes to a stem results in a sequence of open syllables, one of these is elided; thus, *ktibat* 'she wrote' (< *kitab-+ at*), *ktibtih* 'she wrote it [masc.]' (< *kitab+at+ih*); also *glimi* 'my pen', *glimih* 'his pen' (< *galam+-i ,-ih*).

2.2 *Morphology*

The morphology is conservative. Many of the Old Arabic morphology and word class distinctions are maintained.

2.2.1 Pronouns and similar elements

2.2.1.1 Personal pronouns

Personal pronouns are as follows:

hu	'he'	*ant(a)*	'you [masc. sg.]'	*ana*	'I'
hi	'she'	*anti*	'you [fem. sg.]'		
hum	'they [masc.]'	*antum*	'you [masc. pl.]'	*ḥinna*	'we'
hin	'they [fem.]'	*antin*	'you [fem. pl.]'		

2.2.1.2 Object pronoun suffixes

The forms of the suffixes are *-ni* 'me', *-(i)k* 'you [masc. sg.]', *-(i)ć* 'you [fem. sg.]', *-(i)h* 'him', *-ha* 'her', *-na* 'us', *-kum* 'you [masc. pl.]', *-kin*, *-ćin* 'you [fem. pl.]', *-hum* 'they [masc.]', *-hin* 'they [fem.]'. The suffixes *-(i)k* and *-(i)ć* show the nonvocalized forms following a consonant, as in *nišadk* 'he asked you [masc. sg.]', *nišadć* 'he asked you [fem. sg.]', *nšidatk* 'she asked you [masc. sg.]'. The suffix *-(i)h* 'him' occurs as *-h* when followed by a vowel, as in *šālh azzaʿal* 'anger took him off'. Contrast *nšidtih* 'she asked him'. Examples:

šāfih	'he saw him'	*šāfihum*	'he saw them [masc.]
šāfaha	'he saw her'	*šāfihin*	'he saw them [fem.]'
šāfik	'he saw you [masc. sg.]'	*šāfikum*	'he saw you [masc. pl.]'
šāfić	'he saw you [fem sg.].'	*šāfićin*	'he saw you [fem. pl.]'
šāfini	'he saw me'	*šāfina*	'he saw us'

These same suffixes can be added to the prepositions *bi-* 'by' and *li-* 'for' and then suffixed to verb forms, as in *gālō-li* 'they said to me', *yittaṣlūn-bina* 'they get in touch with us', *gālih-li* 'he said it to me', *gālōhā-li* 'they said it to me', *giltih-lih* 'I said it to him'. Where two object pronoun suffixes occur, the suffix-bearing particle *-iyya-* may also be used, giving the alternatives

gilt-li-hum-iyyā-h	'I said it to them'
gilt-ih-li-hum	'I said it to them'

Note also that following the prepositions *min* 'from' and *ʿan* 'from, about' the form of the suffix *-(i)h* 'him' is *-h*, as in *minh* 'from him', *ʿanh* 'about him'.

The north Najdi dialects differ in showing *-an* 'me', *-ak* 'you [masc. sg.]', *uh/-w* 'him', *-ah/-h* 'her', *-kam* 'you [masc. pl.]', *-ham* 'them [masc.]'. Junction with stem and suffix also differs in such forms as *šāfitak* 'she saw you [masc. sg.]', *šifnāw* 'we saw him', *šifnāh* 'we saw her', *šāfwah* 'they [masc.] saw her'.

2.2.1.3 Demonstrative pronouns

The demonstrative pronouns show variant forms in some cases, with or without final short vowels and with or without the element *hā-*.

		masculine	feminine
'this'	singular	*hāḏa, ḏa, hā-*	*hāḏi, ḏi*
	plural	*haḏōla, haḏōl* *ḏōla, ḏōl*	*haḏōli*
'that'	singular	*ḏāk, haḏāk*	*ḏīć, hāḏīć*
	plural	*ḏōlāk, haḏōlāk*	*ḏōlīć,* *hāḏōlīć*

These may follow or precede the head noun, as in *yā zēn ḏa ššōf, min hi nūgih* 'what a beautiful sight, whose camels are they?', *mā lik ḥagg baṭṭurīgah ḏi* 'you have no right to behave in that way'. The unmarked demonstrative is *ha-* or *hā*, as in *halyōm* 'today', *halmukān* 'this place', *snitin min hā-ssnīn* 'one of these years'. In the north for the feminine plural, *hāḏalli, ḏallīć* or *ḏallinć* occur.

2.2.1.4 Interrogative words

These include *ćam* 'how many', *lēh, lēš* 'why', *min* 'who', *mita* 'when', *wēn* 'where', *wiš* 'what'. WH- questions where an object suffix is involved show a specific type of structure, with the WH-word occurring initially and the object suffix following the preposition or verb, as in *ćam hū-bih* 'how much is it for', *min-ant wiyyāh* 'who were you with', *min-ant ʿindih* 'who were you staying with?', *wiš kitābih* 'a book of what?', *wēn riḥt-lih* 'where did you go to?', *min nišadt ʿanh* 'who did you ask about?', *wiš ʿanh* 'what about?', *wēn jīt minh* 'where did you come from?', *min-ant minhum* 'who [what tribe] are you from?', *wēn antum mṭillīnin ʿalēh* 'where are you [masc. pl.] looking out onto?', *min hū-lih taktib, min taktib lammih* 'who are you writing to?', *min-ant šāryin-lih ḏa* 'who did you buy that for?', *min hi nūgih* 'whose camels are they?', *hāḏa wiš ṣōtih* 'what is that the sound of?'.

2.2.2 Particles

Included here are elements of a syntactic nature: *ćān* 'if', *hna* 'there is', *mā min* 'there is no…', *ḥadr* 'under, downward', *lēn* 'until', *mēd* 'rather, I mean', *mēr, mār* 'but', *wilā-, wilyā* 'behold', *wilā min-* 'when', *lōn* 'thing' (used in negative and interrogative sentences), *yamm* 'beside', *yōm* 'when [with past reference]', *lamm* 'toward', *wara* 'why'. Examples include *mā hna abrak min hannāgah* 'there is nothing better than that she-camel', *mā min nahār* 'there is no daylight left', *mā-lna lōn* 'we have nothing', *tabi lōn* 'do you want something?', *mā-bha lōn* 'there is nothing there'.

2.2.3 The noun

Nominal morphology does not differ from the general Arabic form except for a preference for broken plurals even in the participle, as in *sbūğ/sbāğ* 'fast, fleet-footed', *zēn/zyān* 'good', *wāǵif/waggāfi, wugūf* 'standing', *gāniṣ/gannāṣi* 'hunting', *jālis/jilūs* 'sitting', and for the form *fiʿlān* for human plural nouns, as in *šīxān* 'shaikhs', *šībān* 'old men', *šūyān* 'shepherds', *ʿurbān* 'nomads', *bidwān* 'Bedouin', *xuwwān* 'brothers', *šijʿān* 'brave men', *xiblān* 'mad people', *hiblān* 'silly people', *wiġdān* 'small boys'.

Certain of the phonological processes mentioned under 2.1.3 produce idiosyncratic forms in the following ways: (ii) relationship of short vowels to syllable structure: *ćitab* 'camel saddle'; (iii) the influence of the guttural group: *hal-* 'people, family', *nxalah* 'palm tree', *ghawah* 'coffee', *sḥama* 'dark [of camels],' *dhana* 'sand

Table 1. Form I of Action Verb

Action verb

Perfect

kitab	'he wrote'	*kitabta*	'you [masc. sg.] wrote'	*kitabt*	'I wrote'
ktibat	'she wrote'	*kitabti*	'you [fem. sg.] wrote'		
ktibaw	'they [masc.] wrote'	*kitabtu*	'you [masc. pl.] wrote'	*kitabna*	'we wrote'
ktiban	'they [fem.] wrote'	*kitabtin*	'you [fem. pl.] wrote'		

Imperfect

yaktib	'he writes'	*taktib*	'you [masc. sg.]'	*aktib*	'I write'
taktib	'she writes'	*taktibīn*	'you [fem. sg.] write'		
yaktibūn	'they [masc.] write'	*taktibūn*	'you [masc. pl.] write'	*naktib*	'we write'
yaktibin	'they [fem.] write'	*taktibin*	'you [fem. pl.] write'		

Imperative

iktib	'write [masc. sg.]!'
ikitbi	'write [fem. sg.]!'
ikitbu	'write [masc. pl.]!'
ikitbin	'write [fem. pl.]!'

desert'; (iv) elision of the short high vowel /i/u/ in a sequence of open syllables: *wlidin* 'a boy', *snitin* 'a year', *bduwi* 'Bedouin', *šjarah* 'tree', *šbićah* 'net, fence', *snimih* 'its hump [the best part of it]'.

The stress rules of the dialect also lead to stressing of the definite article *ál-* preceding nouns of the form CvC or CvCvC, as in *álbil* 'the camels', *álwalad* 'the boy', *álhawa* 'love'.

2.2.4 The verb

The basic morphological distinctions of the Classical Arabic verb are maintained with some further developments.

2.2.4.1 Form I

Form I exhibits Action and State types, illustrated in Table 1 and 2 by the Action verb *kitab* 'to write' and the State verb *samʿ* 'to hear'.

When followed by an object pronoun suffix, the *-n-* of *-in* and *-an* is doubled and the vowel in the suffixes *-i, -u, -aw, -ti, -tu*, and *-na* is lengthened: *ktibannih* 'they [fem.] wrote it [masc.]', *yaktibinnih* 'they [fem.] write it [masc.]', *ikitbīh* 'write [fem.] it [masc.]!', *ikitbūh* 'write [masc. pl.] it [masc.]!', *ktibōh* 'they [masc.] wrote it [masc.]', *kitabtīh* 'you [fem. sg.] wrote it [masc.]', *kitabtūh* 'you [masc. pl.] wrote it [masc.]'. With the suffix *-ta* of the 2nd person masculine singular, the vowel is elided, giving *kitabtih* 'I wrote it [masc. sg.]' or 'you [masc. sg.] wrote it [masc.]'. Often the inde pendent pronoun *ant* 'you' is postposed to the verb, as in *kitabtih ant* 'did you [masc. sg.] write it [masc.]?' The southern dialect of the Āl Murra, ʿIjmān, and Yām has taken this a stage further and shows a suffix *-hant* for 'you [masc. sg.]', as in *šifhant* 'you [masc. sg.] saw'.

The suffixes for the feminine singular and masculine plural in the imperative of the above are *-i* and *-u*. In most other verb classes and in the derived Forms from II onward, the forms *-ay* and *-aw* often occur, with, however, considerable regional variation.

The influence of the guttural group mentioned under 2.1.3 iii. above gives forms such as *yḥalif* 'he swears', *yxaluṭ* 'he mixes', *aġaluṭ* 'I make a mistake', *thalik* 'you [masc. sg.] perish', *yʿarif* 'he knows', *ḥalaf* 'he swore', *xalaṭ* 'he mixed', *ġalab* 'he conquered'.

The north Najdi dialects differ in showing the suffix *-tam* for 'you [masc. pl.]'.

Strong verbs in this conjugation include *xafḏ̣* 'to be reduced', *waǵf* 'to stand', *raxṣ* 'to

Table 2. Form I of State verb

State verb

Perfect

samʿ	‘he heard’	*simiʿta*	‘you [masc. sg.] heard’	*simiʿt*	‘I heard’
samʿat	‘she heard	*simiʿti*	‘you [fem. sg.] heard’		
samʿaw	‘they [masc.] heard’	*simiʿtu*	‘you [pl. masc.] heard’	*simiʿna*	‘we heard’
samʿan	‘they [fem.] heard’	*simiʿtin*	‘you [fem. pl.] heard’		

Imperfect

yismaʿ	‘he hears’	*tismaʿ*	‘you [masc. sg.] hear’	*asmaʿ*	‘I hear’
tismaʿ	‘she hears’	*tismaʿīn*	‘you [fem. sg.] hear’		
yismaʿōn	‘they [masc.] hear’	*tismaʿōn*	‘you [pl. masc.] hear’	*nismaʿ*	‘we hear’
yismaʿin	‘they [fem.] hear’	*tismaʿin*	‘you [pl. fem.] hear’		

Imperative

ismaʿ	‘hear [masc. sg.]!’
ismaʿi	‘hear [fem. sg.]!’
ismaʿu	‘hear [masc. pl.]!’
ismaʿin	‘hear [fem. pl.]!’

become cheap’, *samn* ‘to become fat’, *salm* ‘to survive’, *ġaluṭ* ‘to be mistaken’, *xarub* ‘to be spoiled’, *xasir* ‘to lose’, *rabḥ* ‘to gain’, *xaluṣ* ‘to finish’, *fahim* ‘to understand’, *ḏ̣aʿf* ‘to become weak’, *ḏ̣aḥć* ‘to laugh’, *laḥǵ* ‘to follow’, *ġanim* ‘to win’. The northern and southern dialects show CiCiC or CaCiC for these, i.e. *simiʿ*, *wiǵif*, *rixiṣ*, *ġaluṭ*, etc.

2.2.4.2 Derived patterns

The system of derived patterns is highly developed in the dialect, showing most of the Old Arabic Forms and some new developments. The system is as follows:

Form I	*giṭaʾ*	‘to cut’
Form II	*gaṭṭaʾ*	‘to cut in pieces’
Form III	*gāṭaʾ*	‘to boycott’
Form IV	*ašmal*	‘to go north’
Form V	*tifaṣṣax*	‘to undress’
Form VI	*tuwājah*	‘to meet with’
Form VII	*ingiṭaʾ*	‘to be cut off (from)’
Form VIII	*intišar*	‘to be distributed’
Form IX	*iswadd*	‘to become black’
Form X	*istankar*	‘to regard as strange’

Four other Forms occur in the dialect which are not known in Old Arabic. These are:

Form XI	*ṭōṭaḥ*	‘to totter’
Form XII	*tisōlaf*	‘to chat, talk’
Form XIII	*dēwar*	‘to turn around [car]’
Form XIV	*tihēbal*	‘to act stupidly’

In addition, the prefix *in-* of Form VII can be combined with Forms V and VI, giving a complex prefix *-int-*, producing items such as *yintuwaggaf* ‘he can be stopped’, *yintigaṭṭaʿ* ‘it can be cut up’, *yintixālaṭ wiyyāh* ‘he can be mixed with’, *mā yintisōlaf wiyyāh* ‘he cannot be spoken with’, *yintifahham* ‘he is capable of being enlightened’.

2.2.4.3 Geminated verbs

These are of the expected form, as in *radd* ‘he returned’, *raddat* ‘she returned’, *raddēt* ‘I returned’, *yirid(d)* ‘he returns’, *yriddūn* ‘they return’, *ridd* ‘return!’ Note, however, that in forms like *yirid(d)* the stress comes on the first syllable when unsuffixed, hence, the *-d(d)* represents a single consonant. When followed by a vowel-initial suffix, however, the double consonant is heard, as in *yiriddih* ‘he returns it [masc.]’

2.2.4.4 Weak verbs

2.2.4.4.1 Verbs I*w/y*

In the imperfect, the exponent of the first radical is *ā*, as in the following examples: *wizan/yāzan* ‘to weigh’, *amin /yāman* ‘to believe’. Some Iʾ verbs have assimilated to the I*w* type, giving *wamin* ‘he believed’ (but *yāman*), *waććad/ ywaććid* ‘assured’, *wkād* ‘surely’, etc.

2.2.4.4.2 Verbs II*w/y*

Here the forms are of the usual type, as in *šāl* 'he took away', *šālat* 'she took away', *šilt* 'I took away', *yšīl* 'he takes away', *yšīlūn* 'they [masc.] take away'. Note, however, the imperative *šil* 'take away [masc. sg.]!', *šīlay* 'take away [fem. sg.]!', *šīlaw* 'take away [masc. pl.]!', *šīlin* 'take away [fem. pl.]!'.

2.2.4.4.3 Verbs III*y*

Here both Action and State types appear.

Action: *miša* 'he went', *mišat* 'she went', *mišēt* 'I went', *yamši* 'he goes', *yamšūn* 'they [masc.] go', *yamšin* 'they [fem.] go'.
State: *nasi* 'he forgot', *nasyat* 'she forgot', *nisīt* 'I forgot', *yinsa* 'he forgets', *yinsōn* 'they [masc.] forget', *yinsan* 'they [fem.] forget'. Note also the form of the plain imperative, which shows no final vowel, giving *imš* and *ins*.

The verbs *axaḏ/yāxiḏ* 'to take' and *akal/yākil* 'to eat' have alternative perfect forms of the final weak type, as in *xaḏa* 'he took', *xaḏat* 'she took', *xaḏēt* 'I took', *kala* 'he ate', *kalat* 'she ate', *kalēna* 'we ate'. Note also the verb *baġa* 'to want', which shows imperfect forms with elision of the *ġ*: *yabi* 'he wants', *abi* 'I want', *yabūn* 'they [masc.] want', etc.

2.2.4.5 The internal passive

The Old Arabic internal passive is still operative in the dialect. It involves syllabic and internal vowel change and attendant velar fronting where relevant, as well as a difference in the vowel quality of the plural suffixes, giving the forms in Table 3.

Other verb classes have the expected type of form, as in the following: *simm* 'he was poisoned', *ysamm* 'he is poisoned', *rimi* 'it was shot', *yirma* 'he is shot', *šīl* 'it was carried', *yšāl* 'it is carried', *wixiḏ* 'it was taken', *yūxaḏ* 'it is taken'. The internal passsive is frequently used in the meaning of 'to be… – able' or 'to be in… – able', as in *ma yūkal* 'it is inedible', which is an alternative to *ma yinwikil.*

2.2.5 Preverbal particles

Preverbal particles mark negation, mode, and tense. The dialect has a particularly elaborate system of modal particles.

2.2.5.1 Negators

The verb is negated by the particles *mā* and *lā*. *Mā* may be combined with *bi-* as in Old Arabic, when preceding the participle. *Mā* precedes indicatives, while *lā* precedes wishes and imperatives, as in *lā tisāl al ʿirrīs lamman mā tidūr alḥōl* 'do not inquire of the bridegroom until the (first) year is over', *wara mā rićibtum min dīratkum* 'why did you not ride here from your tribal lands?', *mā hum bsāylīn* 'they are not asking [i.e. "not worried"]', *mana brāyiḥ* 'I am not going'.

2.2.5.2 Modals

The modals are often reduced forms of verbs or other elements.

i. *baġa/yabi* future intent/imminent future 'will'. Examples: *inšid ʿanh w akīd tabi tilgāh* 'ask about him and you will definitely find him', *tibi tilid* 'she is about to give birth', *baġēna nmūt* 'we almost died'.

Table 3. Internal passive

Perfect

srig/sirǵ	'he was robbed'	*sirigta*	'you [sg. masc.] were robbed'	*sirigt*	'I was robbed'
sirǵat	'she was robbed'	*sirigti*	'you [sg. fem.] were robbed'		
sirǵaw	'they [masc.] were robbed'	*sirigtu*	'you [pl. masc.] were robbed'	*sirigna*	'we were robbed'
sirǵan	'they [fem.] were robbed'	*sirigtin*	'you [pl. fem.] were robbed'		

Imperfect

yisrag	'he is robbed'	*tisrag*	'you [sg. masc.] are robbed'	*israg*	'I am robbed'
tisrag	'she is robbed'	*tisragēn*	'you [sg. fem.] are robbed'		
yisragōn	'they [masc.] are robbed'	*tisragōn*	'you [pl. masc.] are robbed'	*nisrag*	'we are robbed'
yisragan	'they [fem.] are robbed'	*tisragan*	'you are [pl. fem.] robbed'		

ii. *lēt-, ya lēt* desiderative/precative ‘would that, please’: *ya lētha ma njaḥat* ‘I wish she had not passed’, *ya lētik ḥāḏ̣ir* ‘if only you had been there’, *lētik tināwilni assammāʿah* ‘could you please pass me the microphone’.

iii. *ʿasa-* optative (1) ‘I hope’: *ʿasa ma xlāf* ‘I hope nothing is wrong’, *ʿasāhum wāṣlīn bissalāmah* ‘I hope they have arrived safely’.

iv. *jiʿil-* optative (2) ‘may’: *jiʿlik tislam ma gaṣṣart* ‘may you be well; you did well’.

v. *xal-, xall-* jussive (1) ‘let’: *xallik ʿindana* ‘stay with us!’, *xallih yijīna* ‘let him come to us!’, *xal nāxḏih* ‘let’s take it!’.

vi. *daʿ*-jussive (2) ‘let’: *daʿih yinbaḥ* ‘let him bark!’.

vii. *illa, kūd* (northern) necessitive ‘must’: *illa tifikk nāgiti* ‘you must release my she-camel’, *illa tiḏūg ma ḏāgaw* ‘you must suffer their fate’, *kūd yinćfūn wala yašrbūn illa mnah* ‘they must turn back; they have no other well to drink from’.

viii. *kūd-* (central), *aġadi-* (northern) optative resultant ‘mayhap, in case, perchance’: *abarūḥ kūd alga maʿhin zōd* ‘I will go [and look] and perchance may find more [coins] with them’, *kūd uxūy ṣalāḥ yigwāh* ‘perchance my brother Ṣalāḥ will be a match for him’, *nabi nistarxaṣ min alamīr aġadīh yasmaḥilna* ‘we will ask permission from the prince in the hope that he will permit us’.

ix. *yallah* qualificatory ‘just, barely’: *yallah tigdar tiʿīš ʿala harrātib* ‘you can only just survive on this salary’, *baʿad taʿab kiṯīr yallah inni gidart aji* ‘after much effort, I just managed to come’.

x. *ćān* obligative unfulfilled ‘should have’. This precedes the perfect only in this meaning: *wallah ćān gilt li* ‘by God you should have told me’.

xi. *ćān b-* unfulfilled past intent ‘was going to’: *ćān battaṣil bik* ‘I was going to telephone you’, *ćān abarūḥ* ‘I was going to go’.

xii. *widd-ak* (northern) obligative remote ‘ought to’: *widdak tijīna* ‘you ought to come to us [“please come to us”]’, *widdak yṣīr aḥsan* ‘it should be better’.

xiii. *ćinn-* speculative (1), polite inquiry ‘it seems’: *ćinnih ma yifham wājid* ‘he doesn’t seem to understand much’, *farasin ṭayyibah ma ćinnaha rićbat* ‘a good horse, as though it had not been ridden’.

xiv. *tigil* speculative (2) ‘it seems, it looks like, perhaps’: *tigil jayyhum ʿilm niḏir tigil jayyhum šayy* ‘it seems they received a warning or something’.

xv. *l-* jurative ‘I swear that’: *bjizūr inni laḏbaḥah* ‘I swear I will slaughter a fat camel [as a thanksgiving sacrifice]’.

2.2.5.3 The tense-marking particle *ǵid/gid*

When preceding the perfect, the tense-marking particle *ǵid/gid* produces the meaning ‘has/had done something’. It corresponds to what has been called the ‘experiential perfect,’ as in ‘he has/had been to America’: *hu ǵid rićib* ‘he has ridden [i.e., knows how to ride, but is not at this moment mounted on a horse]’, *int ǵid jītahum gabul* ‘you have visited them before [I think]’, *mā gid šiftih* ‘I have never seen him’. Where the general context is past, the meaning will be past perfect: *ana laḥagtkum lākin ǵid taʿaddētu l-kulliyyah* ‘I came after you, but you had already turned the corner of the college’, *mā ǵid jarrab rgād anniswān* ‘he had not experienced the love of women’.

2.3 *Syntax*

2.3.1 Noun phrase structure

The noun phrase structure is remarkable for its preservation of → nunation. This occurs particularly between noun and adjective and in phrases containing *li-* belonging to’, but it may also occur phrase-final. Examples include *snitin min hā ssnīn* ‘one of these years’, *alla brxuṣtin minh* ‘except by permission from him’, *rifīǵin-li* ‘a friend of mine’, *ligēna sāygin, wlidin ḥarbi* ‘we found a driver, a Ḥarbi lad’.

2.3.2 Sentence structure

Sentence structure does not differ markedly from Old Arabic. The unmarked word order is VSO in verbal sentences, while topic fronting can produce other orders as shown below:

VSO order

šāfat bint ilamīr lahā barg šimāl ‘the daughter of the prince saw lightning to the north’, *lēn jat alʿaṣur*…‘when evening came…’, *rikib tirki bin iḥmēd*… ‘Turki bin Humaid mounted up…’, *wi yšīl ibin garmalah min najd* ‘and Ibn Garmalah moved away from Najd’.

Topic fronting
ahalna xallēnāhum yamm aššifa 'we left our families at al-Shifa', *bini hlāl jālihum dōrin wagt šwayya* 'a slight drought came upon the Bani Hilāl',...*u bini hlāl tāli addīrah* 'and the Bani Hilāl took over the area'.

3. LEXICON

Certain characteristic lexical items are shown here: *abxaṣ* 'more knowledgeable', *amda/yimdi* 'to be time for' (used impersonally, as in *mā yimdīni* 'I have not time to..., cannot manage to...'), *baġa/yabi* 'to want', *bāg/ybūg* 'to steal, betray', *bēǵ* 'muffler for holding a coffeepot', *bḏa, fīḏa* 'here' (northern *bhāḏa*), *ḏarrab/yḏarrib* 'to do a thing properly', *ḏall/yḏill* 'to be afraid, act dishonorably', *tigahwa/ytigahwa* 'to take coffee', *haga/yhaga* 'to think, hold an opinion', *ḥaya* 'rain, grass', *igṭ* 'weys', *ʿilm/ʿlūm* 'news', *ḥaḏaf/yḥaḏif* 'to throw', *mizin* 'rain cloud', *nāḏ̣ar/ynāḏ̣ir* 'to look at', *nida* 'grass', *sām/ysūm* 'to make an offer', *sannaʿ/ysanniʿ* 'to do a thing properly', *sawwa/ysawwi* 'to do, make', *šanag* 'side', *ṯār/yṯūr* 'to stand up', *ṯawwar/yṯawwir* 'to shoot a gun', *zaham/yizham* 'to call', *zēn* 'good'.

BIBLIOGRAPHICAL REFERENCES

Burckhardt, John Lewis. 1831. *Notes on the Bedouins and Wahabys collected during travels in the East* 2 vols. London: Association for Promoting the Discovery of the Interior of Africa.
Ingham, Bruce. 1986. *Bedouin of northern Arabia: Traditions of the Al Dhafir.* London: Kegan Paul International.
——. 1994. *Najdi Arabic: Central Arabian.* Amsterdam and Philadelphia: J. Benjamins.
Johnstone, Thomas Muir. 1964. "Further studies on the Dōsiri dialect of Arabic as spoken in Kuwait". *Bulletin of the School of Oriental and African Studies*, 27.77–113.
Kurpershoek, Marcel. 1994. *The poetry of ad-Dindān: A Bedouin bard in Saudi Arabia.* Leiden: E.J. Brill.
Landberg, Carlo de. 1919. *Langue des bédouins ʿAnazeh.* Leiden: E.J. Brill.
Prochazka, Theodore. 1988. *Saudi Arabian dialects.* London: Kegan Paul International.
Sowayan, Saad Abdallah. 1992. *The Arabian oral historical narrative: An ethnographic and linguistic analysis.* Wiesbaden: O. Harrassowitz.

BRUCE INGHAM
(School of Oriental and African Studies, London)

Nasalization

In order to discuss nasalization, one has to give a brief account of nasality. Both nasality and nasalization are natural properties of language. Nasality in speech, as opposed to orality, is a reflection of the physical position of the soft palate, or the velum. In this position, the posterior part of the velum, or the uvula, is lowered so as to keep the nasal cavity or specifically the velopharyngeal port open while pronouncing the nasal sounds. There are two basic nasal consonants in most languages of the world, namely /m/ and /n/. In terms of place of articulation, the phonemes are stops at the bilabial and alveolar regions, respectively. The illustrations below display velic closure while articulating oral sounds (Fig. 1); velic opening while articulating nasal sounds (Fig. 2); and velic as well as oral opening during the articulation of nasalized sounds (Fig. 3).

Figure 1. Velic closure.

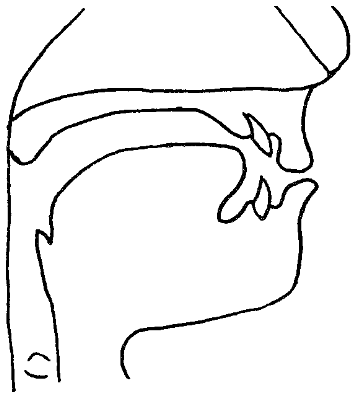


Figure 2. Velic opening.

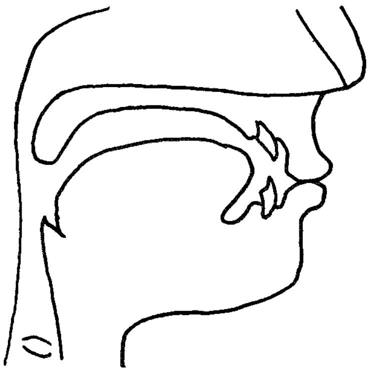
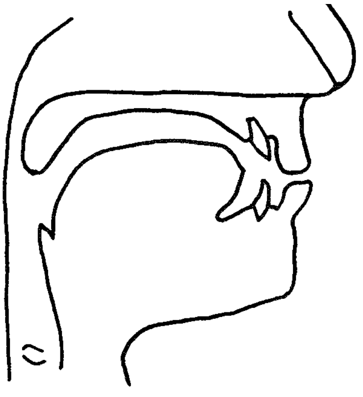


Figure 3. Velic and oral openings.



According to Sībawayhi (d. 180/796), the basic Arabic nasal consonantal phonemes are /m/ and /n/. He describes /m/ as a voiced (*majhūr*) bilabial (*min bayni š-šafatayn*) nasal (*min al-ʾanf*); and /n/ as a voiced alveolar (*fuwayqa ṯ-ṯanāyā*) nasal (Sībawayhi, *Kitāb* IV, 431ff.). In the same description, both /m/ and /n/ are analyzed as stop consonants (*ḥurūf šadīda*). Sībawayhi (*Kitāb* IV, 435) assigns the feature of nasality (*ġunna*) when the egressive air escapes through the nose and not the mouth. The term *ġunna*, which is derived from the root *ġ-n-n* 'to sing', indicates the nasal resonance resulting in the nose during the articulation of the nasal consonants. A nasal consonant is traditionally described as *ʾaġann* (Ibn Jinnī [d. 392/1001], *Sirr* II, 435). As an experiment, Sībawayhi (*Kitāb* IV, 434) adds, "If you held your nose tightly during the articulation of /n/ or /m/, the sound would not flow". In addition, he recognizes a group of nonbasic speech sounds resulting from the contact of /-n/ (*nūn sākina* 'unvoweled /n/'), whether in word-medial or word-final position, with certain consonants producing allophonic homorganic variations of /n/. The details about the *nūn sākina* and → nunation (→ *tanwīn*) are explicitly given in grammatical and phonetic works, under the sections on → *ʾidġām*, which largely involves assimilation (Sībawayhi, *Kitāb* IV, 445ff.; Ibn Jinnī, *Sirr* I, 421ff.).

The Arab and Muslim grammarians and → *tajwīd* (see below) specialists divide assimilation into two major types:

i. Assimilation with *ġunna*
   The rules state that /n/ assimilates to any of the immediately following sounds: /y/, /m/, /w/, or /n/, whether it is within a word or in word-final position. Thus, for instance (Sībawayhi, *Kitāb* IV, 452ff.):
   /n/ + /y/ → [ɲy], where /n/ is palatalized and /y/ may be partially nasalized
   /n/ + /w/ → [n^{w}], where /n/ is labialized and /w/ may be partially nasalized
ii. Assimilation without *ġunna*
   This class includes the combination of /n/ with /l/ or /r/. Thus (Bakalla 1982b:168):
   /n/ + /l/ → [ll], where /n/ may lose its nasality in the sequence
   /n/ + /r/ → [rr], where /n/ may lose its nasality

However, Sībawayhi (*Kitāb* IV, 452) mentions that nasality and/or nasalization largely occur in Arabic in such given sequences. It is apparent from the above statements and the above-given phonological rules that the nasal sounds basically underlie the spread of nasality, and, hence, nasalization may spread over a certain stretch of sounds.

It would be appropriate here to introduce the feature of nasalization, or what may be called *taġnīn*. Unlike nasality, in nasalization the vibrating air passes through both the nostrils and the mouth at the same time (see Fig. 3). In principle, this is a nonphonemic feature. Prosodically, the nasal consonants may spread their nasality over the neighboring sounds, whether they are vowels or consonants or, possibly, both. In some modern Arabic dialects, including the Meccan dialect, this process may affect the preceding and/or succeeding elements on either or both sides of the nasal consonants. Detailed analyses of the extent of, and constraints on, its spread have not yet been fully done.

Furthermore, in *tajwīd* (*ʿilm at-tajwīd* 'the science of Qurʾānic recitation'), both /m/ and /n/ are dealt with in a more extensive and elaborate way in chapters on the rules of /n/ and /m/. As for /n/, it is treated under four headings, one of which is assimilation as outlined above. A second heading is *ʾiḏ̣hār* (i.e. proper and clear pronunciation) when word-medial or word-final /n/ is followed by one of the following guttural sounds: *ʾ*, *h*, *ʿ*, *ḥ*, *ġ*, *x*. The third is *ʾiqlāb* or mutation, i.e. where a word-medial or word-final /n/ changes into [m] in the context of a following /b/; in modern terminology, this case would be treated as regressive assimilation, e.g. *ʾanbaʾa* > *ʾambaʾa*. The fourth case is *ʾixfāʾ* ‘hiding’ of the [n], which involves the pronunciation of /n/ as a homorganic nasal consonant. Here, /n/ assumes the pronunciation of one of fifteen consonants, such as /ḏ/ and /ṯ/; e.g. *munḏir* > *munḏir*, in which the alveolar [n] is articulated interdentally under the influence of the following interdental consonant. Regarding /m/, it is also studied under three subheadings, which include assimilation and doubling of the nasal /m/.

As for the frequency of occurrence of the Arabic nasals, the Arab and Muslim grammarians and phoneticians made some statements concerning doubling of nasal and non-nasal consonants in Arabic, which is not uncommon and occurs, in fact, frequently. Thus, a doubled (*mušaddad*) nasal [mm] or [nn] continues and lasts as long as two short vowels, in terms of duration, estimated as one second by modern standard (Taweel 2000:43).

The grammarians and phoneticians have also studied the duration nasality takes in pronouncing the nasal sounds. Different methods of measurement have been applied. One is by counting with the fingers: nasality lasts as long as it takes for one folding and one unfolding of the palm, in a style that is neither slow nor fast. Another method is by measuring nasality in terms of the short vowel duration: nasality takes the duration of what one, or more than one, short vowel takes in normal speech (Bakalla 1982a:402ff.).

A final note on nasality or nasalization concerns its presence or absence in some individuals’ speech. Cleft lip or cleft palate patients tend to nasalize throughout their speech. On the other hand, people who suffer from acute cold or flu experience partial or complete loss of nasality and nasalization throughout their utterances. The first case is known as hypernasality (*taxnīn*). Al-Kindī (d. ca. 256/868; *ʿUyūb* 530) lists it among the speech defects. The other one is currently known as hyponasality (Ball 1989:41).

To conclude, nasality and nasalization were studied by early Arab and Muslim grammarians and orthoepists in some detail. Their works, though fragmentary in some cases, deserve a close look and scientific assessment by modern Arabists and phoneticians as well.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Ibn Jinnī, *Sirr* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *Sirr ṣināʿat al-ʾiʿrāb*. 2 vols. Damascus: Dār al-Qalam, 1985.

Kindī, *ʿUyūb* = ʾAbū Yūsuf Yaʿqūb ibn ʾIsḥāq al-Kindī, *ʿUyūb al-kalām. Revue de l’Académie Arabe de Damas* 60 (1985): 515–532.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān ibn Qanbar Sībawayhi, *al-Kitāb*, IV. Ed. ʿAbd as-Salām Muḥammad Hārūn. Beirut: Dār al-Jīl, 1975.

SECONDARY SOURCES

Bakalla, Muhammad Hasan. 1982a. “Analysis of nasal elements by early Arab and Muslim phoneticians”. *Bulletin of the College of Arts* (Taipei, National Taiwan University) 395–414.

——. 1982b. *Ibn Jinni, an early Arab Muslim phonetician: An interpretive study of his life and contribution to linguistics.* London and Taipei: European Languages Publication.

Ball, Martin J. 1989. *Phonetics for speech pathology.* London: Taylor and Francis.

Taweel, Ahmad ibn Ahmad Muhammad Abdalla [Ṭawīl, ʾAḥmad ibn ʾAḥmad Muḥammad ʿAbdallāh aṭ-]. 2000. *Taysīr ʿilm at-tajwīd.* Riyadh: Dār Ibn Xuzayma.

MUHAMMAD HASAN BAKALLA (King Saud University)

Naṣb → ʾIʿrāb

Nāsix → Nawāsix

Nastaʿlīq

Nastaʿlīq is the Persian script par excellence, emerging in its definite form in Iran (Tabriz and Shiraz) in the late 8th/14th century (Richard 2001:77). This script, originally known as *nasx-taʿlīq* or *nasx-i taʿlīq* (Richard 2003b:77), implies a blend or derivation from both → *nasx* and *taʿlīq* ‘hanging, suspended’, the latter being a Persian chancery script which appears to have been derived principally from *tawqīʿ* script (→ *ṯuluṯ*) and which, although employed earlier (perhaps as early as the 5th/11th century), was practiced in its definitive form in the 7th/13th century (Richard 2003b:76–77). In the Arab world, *nastaʿlīq* is known as *al-xaṭṭ al-fārisī* ‘the Persian script’.

Taʿlīq had a characteristic aspect and ductus in which words descend onto the baseline and the end of the line curves upward. Just as with *tawqīʿ*, this script was found in its seriffed and serifless (sans serif) forms. When seriffed, the letters *ʾalif* and *lām* had their head-serifs left sloping (Gacek 2003).

The derivation from both *nasx* and *taʿlīq* has recently been challenged by Elaine Wright, who, through a meticulous analysis of specimens and textual support (notably a document drawn up by a Timurid scribe in the early 9th/15th century), showed that *nastaʿlīq* developed from *nasx* by modifying a number of letter forms, and that the name *nasx-i taʿlīq* should therefore be interpreted as 'hanging *nasx*' (Wright 2003; see also Blair 2006:274–276).

One of the first to use the appellation *nasx-i taʿlīq* was Yaʿqūb, known as Sirāj Šīrāzī (fl. 858/1454), the author of an important Persian treatise on calligraphy, *Tuḥfat al-muḥibbīn* (Richard 2003a:8). *Nastaʿlīq* was used initially for copying Persian poetry and romantic or mystical epics, but by the end of the 9th/15th century it had replaced *nasx* for the transcription of prose as well (Soucek 1979:18). It was rarely used for copying of the *Qurʾān*. To our knowledge, there are only three surviving copies of the *Qurʾān* penned in this script (Schimmel and Rivolta 1992:34). A superb copy of the *Qurʾān* in *nastaʿlīq* is preserved in Topkapı Palace Library, copied in 945/1538 by Šāh Maḥmūd an-Nīsābūrī (Lings and Safadi 1976:14; Blair 2006:433).

Other uses of *nastaʿlīq* included private purposes, inscriptions on coins, seals, inkpots, and other objects, as well as architectural inscriptions, albums of calligraphy (*muraqqaʿāt*), and interlinear translations and glosses in manuscripts of the *Qurʾān* (Bayani and Stanley 1999:125). It is thought that *nastaʿlīq* was particularly well suited to the grammatical structure of the Persian language because this language "has many verbal and nominal endings that require one of the rounded Arabic final letters" (Schimmel and Rivolta 1992:30). *Nastaʿlīq* became the script of choice in the Persianate world, i.e. the countries under Persian cultural influence, and it was widely used in its regional forms or variants in Mughal India and Ottoman Turkey.

According to Persian tradition, *nastaʿlīq* was given its definite shape by Mīr ʿAlī Tabrīzī (d. 850/1446) and was practiced in its two different styles: that of Mīrzā Jaʿfar Tabrīzī (9th/15th century), later known as the Xurasānī (or Eastern) style, and the style of ʿAbd ar-Raḥmān Xwārizmī (9th/15th century) and his sons (ʿAbd ar-Raḥīm and ʿAbd al-Karīm), known as the Western style. The Western style was regarded as less perfect and was eventually discarded in favor of the Eastern style (Hanaway and Spooner 1995:3; ʾAḥmad, *Gulzār* 1959:100–174). The main surviving manuscript by Mīr ʿAlī appears to be a copy of Niẓāmī's *Xusraw va Šīrīn*, preserved in the Freer Gallery of Art (Smithsonian Institution), Washington, D.C., penned ca. 813/1410. He signs his name as ʿAlī ibn Ḥasan as-Sulṭānī (Blair 2006:277–278).

In the late 9th/15th and the 10th/16th centuries, the main exponents of this style were such famous calligraphers as Sulṭān ʿAlī Mašhadī (d. 926/1519), the author of *Ṣirāṭ as-suṭūr* (ʾAḥmad 1959:106–125), Mīr ʿAlī Ḥusaynī Haravī (d. 951/1544), the author of *Midād al-xuṭūṭ*, Bābā Šāh Iṣfahānī (d. 996/1587), the author of *ʾĀdāb al-mašq*, and Mīr ʿImād al-Ḥasanī (d. 1024/1615) (Akimushkin 1996). Sulṭān ʿAlī Mašhadī, apart from copying books in Persian and Turkic, was responsible for designing inscriptions for buildings and tombstones (Soucek 1979:30).

In spite of the various differences in styles, *nastaʿlīq* on the whole has a characteristic ductus in which words descend onto the baseline, many horizontal lines are greatly elongated, and the last letter or word is often superscript. Unlike most types of *taʿlīq*, the script is serifless (sans serif). The *ʾalif* is rather short in comparison with the *nasx* of the same period, and the *nūn* looks like a semicircle (or bowl). The curves on the descenders of *qāf*, *ʿayn*, and *lām* are also deep and pronounced. Just as in medieval 'proportioned' scripts, the dimensions of the letters are determined by measuring them with the rhombic dot of the reed pen.

Nastaʿlīq, as seen in many calligraphic specimens, also has a characteristic system of pointing in which the letter *šīn* has three superscript dots arranged in the shape of a triangle, and where the letter *sīn* features three subscript dots executed also in a triangular fashion. There is usually no vocalization, due to its compact aspect, except sometimes for *kasra*, which can be vertical as opposed to inclined.

As a result of writing *nastaʿlīq* rapidly, a new style, known as *šikasta-nastaʿlīq* (*šikastah-yi nastaʿlīq*) and later *xaṭṭ-i šikasta* 'the broken' script, developed in the 11th/17th century. Combining the forms of *nastaʿlīq* and *taʿlīq*, *šikasta* appeared in the Safavid chancery in 1670 (Richard 2001:78). It was given its definitive form by Muhammad Šafīᶜ Haravī (d. 1081/1670–1671), and its most beautiful examples date from the early 12th/18th century. One of its greatest masters was ᶜAbd al-Majīd Ṭāliqānī (d. 1185/1771; Schimmel 1979:202; 1984:31). Just as with *nastaᶜlīq*, it is a serifless script.

Some of the salient features of *šikasta* include the free use of ligatures, assimilation/contraction of letters, and logographs, many of which have to be learned individually, as well as a scarcity and often misplacement of diacritical points. In terms of letter forms, there is, for instance, the characteristic final *nūn*, with its reversed (recurved) loop, and the long, uncurved final *yā*ʾ (Hanaway and Spooner 1995). *Šikasta* had its epistolary variants, known as *xaṭṭ-i tarassul* and *xaṭṭ-i taḥrīrī* (*xaṭṭ-i taḥrīr*; Gacek 2001: 30, 79).

By the 13th/19th century, *šikasta* script had become difficult to read, and therefore efforts were made to reform it. The result was a broad array of *nastaᶜlīq* styles with elements of *šikasta*, which came to be known as *šikasta-amīz* (Hanaway and Spooner 1995:3–4).

Nastaᶜlīq quickly spread to neighboring countries and regions such as Afghanistan, Indo-Pakistan, and Turkey. No rigorous study of the regional styles exists yet, however. In Ottoman Turkey, *nastaᶜlīq* was mostly known as *taᶜlīq* (*talik*), rarely as *nestalik*. It began to be used in Anatolia around the middle of the 9th/15th century for Turkish, Persian, and Arabic texts alike. It was not, however, until the 11th/17th century that this script became fully accepted and appreciated. This was due to the influence of the style of the Persian master Mīr ᶜImād al-Ḥasanī (d. 1024/1615). Muḥammad Asᶜad al-Yasarī (Yesari Esᶜad Efendi; d. 1213/1798) was one of the greatest practitioners of *nastaᶜlīq* in Ottoman Turkey, and his son Muṣṭafā ᶜIzzat Afandī (Izzet Efendi; d. 1265/1849) is regarded as the founder of a truly Turkish school of *talik*. The main characteristics of the Ottoman *talik* of Mustafa Izzet are the enlargement of letters and the establishment of fixed proportions between them using the rhombic dot of the pen (Alparslan 1973:277–278; Derman 1998:20).

Written rapidly and contrary to the rules, it was known as *taᶜlīq-qirmasī* (*talik kirmasi*). This type of less formal, though often quite elegant, style was often employed for major Arabic textbooks on jurisprudence, philosophy, biography, etc. Apart from being used as a bookhand, Ottoman *talik* was also used for architectural inscriptions, calligraphic specimens and panels and, in the office of the Šayx al-ʾIslām, for fatwas and endowment deeds.

After the end of the Timurid rule, many calligraphers and artists relocated to the Subcontinent to enjoy the patronage of the Mughal emperors, and *nastaᶜlīq* quickly became the favorite script in the Mughal court. Introduced probably toward the middle of the 9th/15th century, *nastaᶜlīq* in India was used widely only at the end of the 10th/16th century. The first known Indian specimen of *nastaᶜlīq* (somewhat clumsy and inelegant) is a copy of *Xamsa* by ʾAmīr Xusraw Dihlavī, datable to the first half of the 9th/15th century. It was only during the 10th/16th century that the use of *nastaᶜlīq* began to supplant *nasx*, which was still predominant in the royal manuscripts in the middle of that century (Brac de la Perrière 2003:91–92).

The Mughal period produced many great calligraphers of *nastaᶜlīq*. A number of names stand out in particular: Muḥammad Ḥusayn Kašmīrī, known as Zarīn-qalam ‘Golden Pen’ (d. 1020/1611 or 1612); ᶜAbd ar-Raḥīm, known as ᶜAnbarīn-qalam ‘Amber Pen’, one of Jahāngīr’s court calligraphers, whose portrait can be seen at the end of the colophon on his copy of Niẓāmī’s *Xamsa*, painted by the artist Dawlat (Blair 2006:538); Awrangzeb’s librarian Hidāyat Allāh, called Zarīn-qalam (d. 1118/1706 or 1707); and ᶜAbd ar-Rašīd Daylamī, a nephew and pupil of ᶜImād Ḥasanī, who became the court calligrapher of Šāh Jahān (Rahman 1979:74–92).

Although heavily influenced by the Persian style, *nastaᶜlīq* in Mughal India developed a number of characteristics of its own, including an increased frequency of elongated strokes and the regular width of spacing between words. It was used there not just as a bookhand for some of the greatest texts of Persian literature but also on coins and imperial seals (Blair 2006:554–557).

In the 13th/19th century, *nastaᶜlīq* was adopted in Iran for many texts printed by lithography, and in the Indo-Pakistani context it was chosen as the style for writing Urdu. It is still being used in Pakistan, even in newsprint.

Bibliographical references

Primary source

ʾAḥmad, *Gulzār* = Qāḍī ʾAḥmad ibn Mīr Munšī al-Ḥusaynī al-Qummī, *Gulzār-i hunar*. Transl. (from Persian) Vladimir Minorsky (with introduction by B.N. Zakhoder), transl. (from Russian) Tatiana Minorsky, *Calligraphers and painters: A treatise by Qāḍī Aḥmad, son of Mīr-Munshī (circa A.H. 1015/A.D. 1606)*. Washington, D.C.: Freer Gallery of Art, Smithsonian Institute, 1959.

Secondary sources

Akimushkin, Oleg. 1996. "The calligraphy of the St. Petersburg Album". *The St. Petersburg* muraqqaʿ: *Album of Indian and Persian miniatures from the 16th through the 18th century and specimens of Persian calligraphy by ʿImâd al-Hasanî*, 39–46. Lugano and Milan: Leonardo Arte.

Alparslan, Ali. 1973. "Ecoles calligraphiques turques". *Islam Tetkikleri Enstitüsü Dergisi* 5:1–4.265–278.

Bayani, Manijeh and Tim Stanley. 1999. "India and Iran: A complex relationship". *The decorated word: Qurʾans of the 17th to 19th centuries*, ed. Maniyeh Bayani and Tim Stanley, 200–227. London and New York: Nour Foundation, in association with Azimuth Editions and Oxford University Press.

Blair, Sheila S. 2006. *Islamic calligraphy*. Edinburgh: Edinburgh University Press.

Brac de la Perrière, Eloïse. 2003. "Bihârî et naskhî-dîwânî: Remarques sur deux calligraphies de l'Inde des sultanats". *Studia Islamica* 96.81–92 and ill. XVIII–XIV.

Derman, M. Uğur. 1998. *Letters in gold: Ottoman calligraphy from the Sakip Sabanci collection, Istanbul*. New York: Metropolitan Museum of Art.

Gacek, Adam. *The Arabic manuscript tradition: A glossary of technical terms and bibliography*. Leiden: E.J. Brill.

——. 2003. "The head-serif (*tarwīs*) and the typology of Arabic scripts: Preliminary observations". *Manuscripta Orientalia* 9:3.27–33.

Hanaway, William L. and Brian Spooner. 1995. *Reading* nastaʿliq*: Persian and Urdu hands from 1500 to the present*. Costa Mesa, Calif.: Mazda.

Lings, Martin and Yasin Hamid Safadi. 1976. *The* Qurʾān*: Catalogue of an exhibition of* Qurʾan *manuscripts at the British Library 3 April–15 August 1976*. London: World of Islam Festival Publishing Co.

Rahman, Parez Islam Syed Mustafizur. 1979. *Islamic calligraphy in medieval India*. Dacca: Dacca University Press.

Richard, Francis. 2001. "Écritures persanes". *L'art du livre arabe: Du manuscrit au livre d'artiste*, ed. Marie-Geneviève Guesdon and Annie Vernay-Nouri, 77–78. Paris: Bibliothèque de France.

——. 2003a. "Autour de la naissance du nastaʿlīq en perse: Les écritures de chancellerie et le foisonnement des styles durant les années 1350–1400". *Manuscripta Orientalia* 9:3.8–15.

——. 2003b. "Chancellerie et naissance de nouvelles écritures: La calligraphie persane". *Studia Islamica* 96.75–79.

Schimmel, Annemarie. 1979. "Poetry and calligraphy: Thoughts about their interaction in Persian culture". *Highlights of Persian art*, ed. Richard Ettinghausen and Ehsan Yarshater. Boulder, Colo.: Westview.

——. 1984. *Calligraphy and Islamic culture*. New York: New York University Press.

——, and Barbara Rivolta. 1992. "Islamic calligraphy". *Metropolitan Museum of Art Bulletin* [published independently].

Soucek, Priscilla P. 1979. "The arts of calligraphy". *The arts of the book in Central Asia, 14th–16th centuries*, ed. Basil Gray, 7–33. Paris and London: Serinda Pub.

Wright, Elaine. 2003. "The calligraphers of Shiraz and the development of the *nastaʿlīq* script". *Manuscripta Orientalia* 9:3.16–26.

Adam Gacek (McGill University)

Nasx

Nasx is a generic name for a variety of Arabic scripts used for many centuries, mostly for the copying of books and later for printing, from Egypt to China and Southeast Asia.

The root *n-s-x* of the word *nasx* is Nabataean in origin and appears to have been associated with copying and transcription from an early period of Islam. The term *nasxī* was originally introduced by Western Arabists to cover all round scripts of the earlier Muslim centuries (Abbott 1939:34, 37). Even though the original sources use the term *nasx*, the term *nasxī* has remained in use to cover both formal and informal (personal, idiosyncratic) scripts/hands, and may be taken as meaning *nasx*-related.

It is not known exactly when *nasx* was introduced as a formal script or what its original characteristics were. The name does not appear among the scripts mentioned by Ibn an-Nadīm, nor is it mentioned in any manual for secretaries (*kuttāb*) from or prior to the 4th/10th century. It is not found, either, in the treatise on calligraphy attributed to Ibn Muqla (d. 328/940) himself, but quotations referring to this script come from Mamluk sources, principally an-Nuwayrī (d. 733/1333) and al-Kātib ad-Dimašqī (fl. 781/1379; Gacek 1987:127; al-Kātib ad-Dimašqī, *Lamḥa* 43). These same sources report that *nasx* was practiced by ʿAbdallāh ibn Muqla, the brother of ʾAbū ʿAlī, the famous vizier (d. 328/940), who, in turn (surprisingly) is reported to have written in *darj* (a chancery script; Qalqašandī, *Ṣubḥ* III, 13).

Possibly the earliest direct reference to *nasx* as a script practiced in the 4th/10th century comes from an anonymous *Risāla fī l-kitāba al-mansūba* (ʿAsākir 1955:126). Interestingly, the author of the *Risāla* (which is likely to have been composed in the early 5th/11th century) mentions that Ibn ʾAsad (4th/10th century), one of the teachers of Ibn al-Bawwāb (d. 413/1022), who in turn was a pupil of Ibn Muqla, wrote poetry in a *nasx* that was close to (*qarīb min*) → *muḥaqqaq* script, because *nasx* in Mamluk texts on calligraphy is viewed precisely as belonging to this family of bookhands. Yet, the affiliation of *nasx* to *muḥaqqaq* in that period would indicate that this type of proto-*nasx*, as well as the *muḥaqqaq*, was probably quite different from the *nasx* and *muḥaqqaq* as elaborated and refined later by Ibn al-Bawwāb (Blair 2006:173–178).

There appears to be no doubt, however, that the use of 'modern' *nasx* (and the other new scripts) goes back to the end of the 4th/10th and the beginning of the 5th/11th century. Indeed, there is important evidence that a new family of scripts was in use toward the end of that period. Probably the best example here is the Chester Beatty Qurʾān, penned by Ibn al-Bawwāb in 391/1000 or 1001, using a number of new scripts. Although some have suggested that the main text of this copy was executed in *nasxī* or *nasx*, it is more likely that the script is in fact *maṣāḥif*, a relative of *nasx* used for medium-size Qurʾāns (→ *muḥaqqaq*).

The existence of this new family of 'proportioned' scripts, including *nasx*, is attested also by other 5th/11th-century manuscripts, some perhaps genuine, but some clearly falsely attributed to Ibn al-Bawwāb himself. A good example of this is a manuscript of the poetry of Salāma ibn Jandal, executed before 456/1064 and using a number of scripts, including → *ṯuluṯ* and *rayḥān*, and a remarkable very small Chester Beatty codex of the *Qurʾān* (K16(1)), executed in a *nasx* hand and dated 428/1037 (Rice 1955:19–22, 26; see also James 1992:22; Lings and Safadi 1976:43–46).

Although a type of *nasx* script was certainly used for the copying of small and middle-size Qurʾāns from the early 5th/11th century onward, *nasx* remained the principal bookhand for the copying of non-Qurʾānic texts, in the fields of *ḥadīṯ, tafsīr, fiqh, naḥw*, and the like (Gacek 1989:146).

Even though all subsequent calligraphers trace their pedigree to Ibn Muqla and Ibn al-Bawwāb, the 'school' or method (*ṭarīqa*) of Ibn al-Bawwāb was really preserved in the main Arab lands (Egypt and Syria). By contrast, Iran, Central Asia, and the later central Arab lands under the Ottomans appear to have followed the tradition of the third major figure in calligraphy, Yāqūt al-Mustaʿṣimī (d. 698/1298). It is interesting to note here that, throughout most of the Mamluk period, *nasx* was regarded as belonging to the rectilinear family of scripts (the *muḥaqqaq* family), whereas in the Yāqūtī tradition *nasx* was paired with *ṯuluṯ*, the curvilinear family.

A great number of distinct types of formal and informal *nasx* developed over the centuries in various regions and centers of learning in the Arabic-, Persian-, and Turkic-speaking lands. J.J. Witkam (1978:18), in the epilogue to his *Seven specimens of Arabic manuscripts* (copied mostly in Baghdad in the 5th/11th and 6th/12th centuries), made a remark that he could not properly label the various styles of *nasxī*, since no adequate criteria for their description and classification yet existed. The above statement, after some 28 years, is still valid, not just as regards informal hands but also, to a large extent, formal scripts. Many labels already assigned to various specimens may have to be changed and/or redefined in view of recent and future research and discoveries. The few general remarks which follow therefore have to be understood in this context.

One of the better known periods and geographical regions, thanks to the combination of extant literature and specimens, is the Mamluk period in Egypt and Syria (648/1250–923/1517). As mentioned above, *nasx* script, according to the Mamluk tradition, was regarded as belonging to the rectilinear family of scripts used mainly as bookhands. It was viewed by most calligraphers as one of the five or seven fundamental scripts (*al-ʾaqlām al-ʾuṣūl*). The earliest description of *nasx* comes from an-Nuwayrī, who states that it has larger (lit. 'thicker', *ġalīḏ*) and lighter (*xafīf*) versions: *qalam al-matn* 'body of the text; block of text' and *qalam al-ḥawāšī* 'glosses'. From *nasx* also stems the *manṯūr* script, which is characterized by large spacing between words. The *matn* script is referred to in other Mamluk sources as *an-nasx al-waḍḍāḥ* 'clear' or *an-nasx al-faḍḍāḥ* 'divulging secrets', while the counters ('eyes') of its letters are described as open (*fatḥ al-ʿuqad*; Gacek 1987:127, 1989:146).

The Mamluk *nasx* has a rather stiff (upright) aspect, with words firmly seated on the baseline. There is an almost total absence of head-serifs (*tarwīs, šaḏ̣iyya, minqār*), especially on such letters as *ʾalif* and *lām* of the definite article but also in such letters as the *dāl/ḏāl*, *ṭāʾ/ḏ̣āʾ*, and *nūn*. Often only the head-serif, in the form of a downward stroke, is visible on the letters *jīm/ḥāʾ/xāʾ* in their isolated forms. Another major feature is the presence of the *lām ʾalif al-warrāqiyya* (ﻻ), which is to be found in most types of *muḥaqqaq*-related scripts, although other types of this letter can be present (e.g. *lām ʾalif* with a rounded base, *al-muḥaqqaqa*; Gacek 1989, 2003).

In the 7th/13th century, *nasx* as a medieval bookhand became established as one of the so-called Six Pens (*al-ʾaqlām as-sitta*) in the eastern part of the Islamic world (principally Baghdad). According to the tradition, these six scripts (paired: *ṯuluṯ/nasx, muḥaqqaq/rayḥan, taqwīʿ/riqāʿ*) were canonized by Yāqūt al-Mustaʿṣimī, and by the early 9th/15th century, the Yāqūtī tradition had become firmly established in Iran and Central Asia (Roxburgh 2003:52).

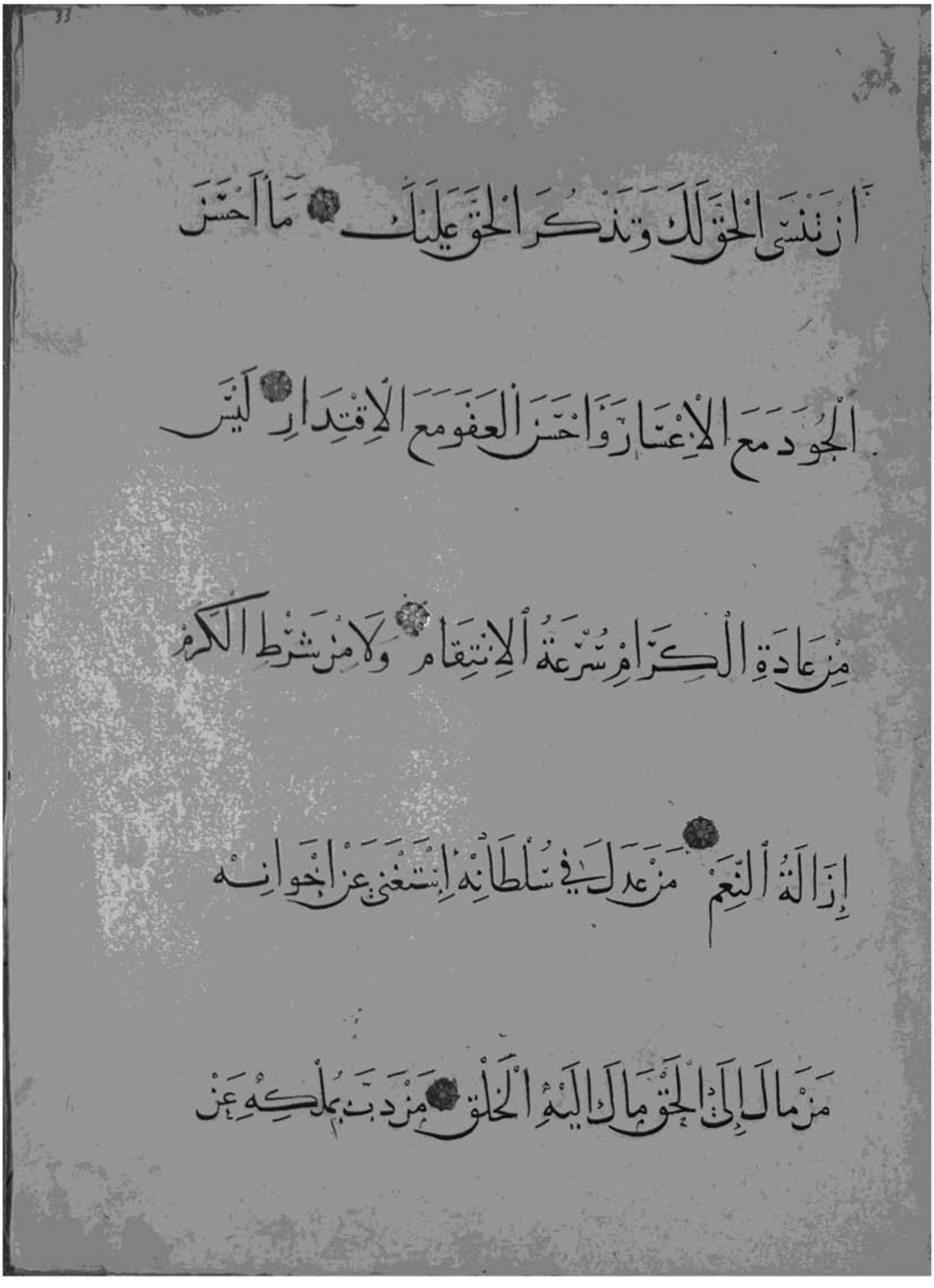
There is some evidence that Yāqūt may have penned, as early as 668/1269 or 1270, a manuscript in Persian in which he employed a type of *nasx* influenced by *ṯuluṯ* (Richard 2003:76). Certainly, the *nasx* in Iran was very different from the rather stiff Mamluk *nasx*. It was rounded but upright, with the letters extremely neatly drawn. A distinctive style also developed in Shiraz in the early 9th/15th century. It was characterized by its lightness and long (swooping) sublinear tails on the final forms of the letters *sīn, nūn*, and *yāʾ* (Blair 2006:263–264). It was exported to the east and west (Anatolia and India). This small hand is beautifully illustrated in Rūzbihān Šīrāzī's multiscript Qurʾān from the mid–10th/16th century (Blair 2006:419–420). The swooping tails are also visible in the *nasx* of Muḥammad al-Qūnyawī, who transcribed the *Maṯnavī* of Rūmī in 677/1278 (Blair 2006:367–369). It was from this type of *nasx* that → *nastaʿlīq* developed.

In Iran, in the early 12th/18th century, *nasx* acquired a particular character given to it by the calligrapher ʾAḥmad Nayrīzī, hence, it was popularly known as Nayrīzī *nasx*. Nayrīzī was a leading calligrapher at the court of Šāh Sulṭān Ḥusayn (Raby 1996). He specialized in *nasx*, which by his day had become associated with the *Qurʾān* and other explicitly religious texts in Arabic. On the whole, Nayrīzī's *nasx* is relatively large in size, with wide spacing of the lines of text, but smaller size was also practiced. Nayrīzī wrote with a very obliquely cut pen, and his script had a characteristic sharp left-sloping head-serif, especially on the *lām* of the definite article, a feature which is seen in many examples of Persian *nasx* from preceding centuries. On the other hand, the free-standing *ʾalif* was serifless (sans serif). In the 13th/19th century, the Nayrīzī style was continued by Muḥammad Šafīʿ, better known as Viṣāl-i Šīrāzī (Gacek 2003; Blair 2006:424–428).

Like their Iranian counterparts, Ottoman Turkish calligraphers refined the *nasx* used earlier in that region. In this period *ṯuluṯ* (*sülüs*) and *nasx* (*nesih*) emerged as the most popular pair of scripts, used both for text and display purposes with fine graceful letters which have a slight tilt to the left. The calligrapher responsible for the refinement of this style was Ḥamd Allāh al-ʾAmāsī (d. 926/1620), who studied under teachers from both Ibn al-Bawwāb's and Yāqūt's traditions. His *nasx*, after further refinements by later calligraphers (principally ʿUṯmān ibn ʿAlī, known as Ḥāfiḏ̣ al-Qurʾān, or Hafez Osman, d. 110/1698), became a model of beauty. Indeed, Ḥamd Allāh created what came to be known later as *aṭ-ṭarīqa al-ḥamdiyya* 'method/school of Ḥamd Allāh'.

Ḥamd Allāh's *nasx* is small and compact, with fine strokes of more or less the same thickness, short ascenders, and elegant sublinear loops (e.g. in *lām, nūn, ṣād*). The counters of many letters are closed, and the *lām ʾalif al-warrāqiyya* is used throughout. There is no head-serif on the free-standing *ʾalif*, but such letters as *ṭāʾ* and *lām* (in combination with *jīm* or *mīm*) are seriffed, but inconsistently with the head-serif right-sloping (Gacek 2003).

Nasx has survived in various forms and is practiced today by modern calligraphers. In the 13th/19th century and the early 14th/20th century, it was used in lithographic printing in such countries as Egypt, Turkey, Iran, and India, and in fact the books printed by lithography in



33

أن تنسى الحق لك وتذكر الحق عليك ● ما أحسن

الجود مع الإعسار وأحسن العفو مع الاقتدار ● ليس

من عادة الكرام سرعة الانتقام ● ولا من شرط الكرم

إزالة النعم ● من عدل في سلطانه استغنى عن إخوانه

من مال إلى الحق مال إليه الخلق ● من دب بملكه عن

Fig. 1. *Nasx al-faḍḍāḥ* by Muḥammad aṭ-Ṭayyībī (Ṭībī), ca. 908/1505 (Istanbul, TSM K.882, f.33a)

those countries represent the best recorded picture of its various regional styles (Gacek 1996). The small Ottoman Turkish *nasx* (perfected later by Hafez Osman) also became a favorite script for the copying and printing of medium and small-size Qurʾāns. In India (and later Pakistan), on the other hand, Manẓūrī *nasx*, introduced by the calligrapher Manẓūr Muḥammad (1866–1950), has been used principally for printing Qurʾāns (Ahmad 1984).

Bibliographical references

Primary sources

Risāla fī l-kitāba = (anon.), *Risāla fī l-kitāba al-mansūba*, ed. Xalīl Maḥmūd ʿAsākir, “Risāla fī l-kitāba al-mansūba”. *Majalla Maʿhad al-Maxṭūṭāt al-ʿArabiyya* 1 (1955) 121–127.

Ibn an-Nadīm, *Fihrist* = ʾAbū l-Faraj Muḥammad ibn ʾIsḥāq Ibn an-Nadīm al-Warrāq al-Baġdādī, *Kitāb al-fihrist*. Ed. Yūsuf ʿAlī Ṭawīl. Beirut: Dār al-Kutub al-ʿIlmiyya, 1996.

al-Kātib ad-Dimašqī, *Lamḥa* = Ḥusayn ibn Yāsīn ibn Muḥammad al-Kātib ad-Dimašqī, *Lamḥat al-muxtaṭif fī ṣināʿat al-xaṭṭ aṣ-ṣalif*, ed. Hayā Muḥammad Dawsarī. Kuwait: Muʾassasat al-Kuwayt li-t-Taqaddum al-ʿIlmī, ʾIdārat at-Taʾlīf wa-t-Tarjama wa-n-Našr, 1992.

Qalqašandī, *Ṣubḥ* = ʾAḥmad ibn ʿAlī al-Qalqašandī, *Ṣubḥ al-ʾaʿšā fī ṣināʿat al-ʾinšāʾ*. 14 vols. Cairo: al-Muʾassasa al-Miṣriyya al-ʿĀmma, 1963–1965. (First publ. aṭ-Ṭabʿa al-ʾAmīriyya, 1913–1918.)

Secondary sources

Abbott, Nabia. 1939. *The rise of the North Arabic script and its Ḳurʾānic development, with a full description of the Ḳurʾān manuscripts in the Oriental Institute*. Chicago: University of Chicago Press.

Ahmad, Syed Barakat. 1984. *Introduction to Qurʾanic script*. London: Curzon Press.

Gacek, Adam. 1987. “Al-Nuwayrī’s classification of Arabic scripts”. *Manuscripts of the Middle East* 2.126–130.

——. 1989. “Arabic scripts and their characteristics as seen through the eyes of Mamluk authors”. *Manuscripts of the Middle East* 4.144–149.

——. 1996. *Arabic lithographed books in the Islamic Studies Library, McGill University: Descriptive catalogue*. Montreal: McGill University Libraries.

——. 2003. “The head-serif (tarwīs) and the typology of Arabic scripts: Preliminary observations”. *Manuscripta Orientalia* 9:3.27–33.

Blair, Sheila S. 2006. *Islamic calligraphy*. Edinburgh: Edinburgh University Press.

James, David. 1992. *The master scribes: Qur’ans of the 11th to 14th centuries A.D.* London and Oxford: Nour Foundation, in association with Azimuth Editions and Oxford University Press.

Lings, Martin and Yasin Hamid Safadi. 1976. *The Qurʾān: Catalogue of an exhibition of Qur’an manuscripts at the British Library, 3 April – 15 August 1976*. London: World of Islam Festival Publishing Co.

Raby, Julian. 1996. “The Nayrizi tradition: Nasx in Safavid and Qajar Iran”. *The art of the pen: Calligraphy of the 14th to 20th centuries*, ed. Nabil F. Safwat, 212–227. London and New York: Nour Foundation, in association with Azimuth Editions and Oxford University Press.

Rice, D.S. 1955. *The unique Ibn al-Bawwāb manuscript in the Chester Beatty Library, Dublin*. Dublin: Emery Walker.

Richard, Francis. 2003. “Chancellerie et naissance de nouvelles écritures: La calligraphie persane”. *Studia Islamica* 96.75–79.

Roxburgh, David J. 2003. “On the transmission and reconstruction of Arabic calligraphy: Ibn al-Bawwab and history”. *Studia Islamica* 96.39–53.

Witkam, Jan Just. 1978. *Seven specimens of Arabic manuscripts preserved in the library of the University of Leiden*. Leiden: E.J. Brill.

Adam Gacek (McGill University)

Nationalism and Language

1. Definition

Language is a communication tool and a cultural vehicle, which implies that it is also a reference for identifying ‘ours’ and ‘theirs’, a content of loyalty and hostility, of social and cultural status. For nationalists, language is a tool that connects past and future, projecting a reconstructed centripetal unity out of the centrifugal reality of the present. Language is, therefore, one of the most visible symbols in the nationalistic museum, a symbol with a powerful legitimizing role (Fishman 1972).

The etymology of the word → *ʿArab* is still uncertain. It was connected with the desert in many Semitic inscriptions, and Arabic is claimed to have developed from a process of nomadization. As a revealed language, Arabic spread from the Arabian Peninsula. As an ethnic attribute, at the time of the Islamic conquest the term *ʿArab* indicated groups of Arabian conquerors, natives of the Arabian Peninsula. It became the ethnic attribute of those who acquired the Arabic language or are citizens of the member states of the Arab League. Through all its history, the most constant feature of this term has been and still is linguistic (Kallas 1999:17–30, 105–116).

Numerous levels of identification cohabit in the ‘Arab worlds’: religious, linguistic, regional, confessional, social, ethnic, and tribal (→ religion and language; → ethnicity). Sometimes two or more levels become integrated. Sometimes they are in conflict. This makes it hard to define the term 'Arab nation' and to understand how Arabic language(s) and Arab nationalism(s) have affected each other – a question whose elements are all subjective and symbolic since Arabic is neither one single language nor does it have a well-defined frame of reference (→ *ʿarabiyya*). A Palestinian may immediately identify a Jordanian speaker and sympathize more with a Christian Palestinian than with a Muslim Jordanian, even though he lives in Jordan. As for identity, it is empirically an individual experience; its collective formula weaves through many situational variables (profession, religion, etc.), and its national expression is an elitist construct. Hence, this article asks more questions than it provides answers.

All the aforementioned concepts oscillate between a minimal and a maximal definition, a completely positive or negative interpretation. At the minimal definition of Arabic stands only Standard Arabic, *al-fuṣḥā*, while the adjective 'Arab' includes all 'Arabic speakers', i.e. speakers of non-Standard Arabic, since *al-fuṣḥā* is nobody's mother tongue. As for the term 'nationalism', it wends between the clearly negative connotations of the Italian *nazionalismo* and the neatly positive implications of the Arabic correspondent *qawmiyya*. The former is biased by the fascist experience, the latter by the anticolonialist struggle for self-determination.

To translate nation and nationalism, Arabic uses at least two terms of different roots: *ʾumma,* which goes back to a Qurʾānic concept of 'community of faith' or 'nation of the Prophet' (Q. 3/110), and *qawmiyya*, which derives from the tribal term *qawm*, denoting a group of people claiming a common ancestor. Coined at the turn of the 20th century from *waṭan,* initially a place of residence, the concepts of 'citizen' (*muwāṭin*) and 'nationality/citizenship' (*muwāṭiniyya*) assume a connotation related to location. By adding the adjective *ʿarabiyya* to the term *ʾumma*, some pan-Arabists advocated to disassociate the *ʾumma* from its Islamic connotations, and by adding the same adjective to the term *qawmiyya* they meant to detach *qawmiyya* from its ethnic and tribal context, too. Some prefer a regional or national-state adjective, arguing that smaller entities are less generic.

Three types of nationalism are discussed in this article: pan-Arab, regional (e.g. pan-Syrian), and national-state (e.g. Egyptian, Tunisian, and Lebanese). Linguistic, cultural, and political features are examined, focusing on the role language plays in them, whether it be Standard Arabic or the language spoken in a specific state.

2. The excursus of Arab nationalisms

Bonaparte's expedition to Egypt (1798) was one of the most serious culture shocks the Ottoman Empire experienced. It also enhanced the debate over the capacity of Arabic to keep up with modernity and the compatibility of the latter with Islamic tradition. From the beginning of the 18th century, this debate was preceded by intensive missionary activity and the Christian revival (*nahḍa*) in Greater Syria, followed by the determined reforms of Muḥammad ʿAlī (1805–1848) in Egypt.

The drastic Turkification imposed upon the Arab provinces by the Young Turks (1908) forced nascent Arab nationalism to adopt more radical choices (al-ʾAfġānī 1962). The Arabic language emerged as that ingredient which distinguished Arabs from Turks. Arab nationalism elaborated from history and from Arabic literature a modern topic for Arabness (*ʿurūba*). It soon developed into a political movement, supported by Britain, aiming to weaken the Ottoman Empire. At the end of the First World War, no unified Arab nation was created, only colonized polities shared mainly between Britain and France. Except for Saudi Arabia, which was independent, political activism focused on resisting foreign occupation, and independence was negotiated by local partners often willing to safeguard colonial interests in exchange for their national sovereignty (Naṣṣār 1986; Zeine 1973).

The new Arab rulers established the Arab League in 1945, but continued to consolidate their local nation building. When, three years later (1948), they failed to halt the establishment of the State of Israel, overwhelming militant and populist Arab nationalism brought new Arab socialist regimes to power. Conflict between 'progressive' regimes and 'conservative' monarchies dominated the scene between the mid-1950s and 1967. The most active pan-Arab among the nationalists was President

Nasser of Egypt. Arab nationalism witnessed a euphoric period after he nationalized the Suez Canal Company, leading to a short-lived union between Egypt and Syria (the United Arab Republic, 1958–1961). The slogans were 'to restore Arab unity over every inch of Arab land from the Gulf to the Ocean' and 'Arab unity is the only means to liberate Palestine'. All hopes were, however, shattered when Israel (1967) single-handedly beat the armies of Egypt, Syria, and Jordan and seized what was left of Palestine.

Disappointed by socialism, Arabness found its way back to Islam with substantial financial support from the wealthy conservative rulers. The defeat of Arab nationalism brought new Arab regimes to power; Arab intellectuals are still debating the reasons for this defeat and prescribing remedial solutions.

Different forms of government are represented in the Arab world. Some of the countries are monarchies: Bahrain, Jordan, Kuwait, Morocco, Oman, Qatar, Saudi Arabia, and the United Arab Emirates. The other Arab countries are all republics, and their official names proclaim that they are democracies. In reality, in most Arab states, with the exception of Lebanon, the political power is concentrated in one monarch, president, or party.

3. PAN-ARABISM AND MINIMALIST LINGUISTIC ASSUMPTIONS

Arab national definitions of Arabness range between a minimalist and a maximalist assumption. Sāṭiʿ al-Ḥuṣrī (d. 1968), one of the principal exponents of pan-Arabism, sought to exorcize religious, racial, and linguistic superiority from the national idea, claiming secular and minimalist objective principles (history and language). For him, a member of an Arabic-speaking people is an Arab, by definition (al-Ḥuṣrī 1955). This idealistic and romantic nationalism had cultural, rather than civic, priorities in common with the German nationalistic model. It had many advocates between the two world wars. If for al-Ḥuṣrī the nation was a moral construct and an act of faith, for Sheikh ʿAbdallah al-ʿAlāylī and for the Christian Michel ʿAflaq, Arab self-identification is instinctive and mystical. The latter is the co-founder with Ṣalāḥ ad-Dīn al-Bīṭār of the Baʿth party (1940), which, in its Constitution (Art. 10), considers self-identification in Arabness a matter of language and self-ascription (Haim 1962:233–241).

4. PAN-ARABISM AND MAXIMALIST ASSUMPTIONS

The maximalist assumption singles out five fundamental features that have been uniting all 'Arabs': race (*ʿirq*), ancestors' homeland (*ʾarḍ al-ʾajdād*), history and historical memories (*tārīx wa-ḏikrayāt*), ethical principles and customs (*ʾaxlāq wa-ʿādāt*), and language (→ *luġa* or → *lisān*). In its maximal concept, nationalism is self-contradictory (being pan-Arab, pan-Slavic, or whatever), for any perfect combination between all these features is rather improbable or improper.

4.1 *Language and race*

Otherness is one of the most ancient contrastive tools of group self-identification; other languages were compared to the croaking of frogs, and common language meant common tribal origins. In the *Qurʾān*, non-Arabs were called *ʿajam* (Q. 16/103, 26/198, 41/44), referring to their *ʿujma*, i.e. their obscure way of speaking, and the adjective *ʿarabī* refers not to a race but rather to linguistic features.

Arab genealogists divide the Arabian Peninsula's populations into three main groups: (i) *al-ʿāriba* indicates the first and only pure Arabs, but it is also an extinct group of tribes; they descend from ʾIram and Lūṭ ibn Sām ibn Nūḥ, the first to speak Arabic after the confusion of tongues in Babel (*tabalbul al-luġāt*); (ii) *al-mutaʿarriba* are not pure-blooded; they descend from Qaḥṭān and live in southern Arabia; (iii) *al-mustaʿriba* 'the arabicized' descend from Maʿadd and ʿAdnān through ʾIsmāʿīl and inhabit the north. To one of these tribes (Qurayš), a tribe which learned Arabic by settling among the 'true' Arabs, the Prophet Muḥammad belonged.

A link between Arabic and the Arabs' superiority has been advocated by the most eminent grammarians, historians, and philosophers, such as Ibn as-Sarrāj (d. 316/928) and Ibn Jinnī (d. 392/1002). They all argue that shared and elevated natural and sociocultural values (*šiyam*) mold among Arabs a shared and superior language. Arabs are generous, ready to help and succor, proud and courageous, honorable and loyal, intuitive and patient. Hence, their language has a genealogical etymology, according to Ibn Durayd (d. 321/933), and a syntactic flexibility, for Ibn Fāris (d. 395/1004). When rebutting anyone who dared mention the Arabs' corrupted tongues and solecism (*laḥn*), the Arabic linguistic tradition was ready to swear that solecism had to be attributed to the Arabs' contact with the non-Arabs (*ʿajam*; Suleiman 2003:38–68).

These assumptions were reevoked in modern times by the pan-Arabist Zakī al-ʾArsūzī (1962) through 'acrobatic philology' aiming at 'restoring the Arab nation' and giving Arabs back their linguistic genius.

4.2 *Language, race, and religion*

According to some *ḥadīṯ*s, the Prophet affirmed several times that Arabness is not genetic but rather linguistic (Ibn Bābawayhi, *Maʿānī l-ʾaxbār* 199). As for the *Qurʾān*, it never refers by the term *ʿarabī* to an ethnic identity but rather to a clear (*mubīn*) language (Q. 16/103); but many of the verses in the *Qurʾān* reflect the problematic relationship with Jews and Christians. Sometimes, they are praised and granted forgiveness; at other times, they are blamed. Hence, they are to be treated as *ḏimmī* to whom hospitality and protection are accorded on condition of their submission to Islam and their paying a poll tax (*jizya*; Q. 9/24). This inferiority status of non-Muslims broke the bonds between claiming ethnic and linguistic Arabness and being Muslim, since many of the *ḏimmī* belonged to both categories.

Once Islam crossed the borders of the Arabian Peninsula, the conversion of many non-Arabs to Islam broke even more of these bonds. The Umayyad caliphs (41/661–132/750) gave institutional weight to Arabic and ethnic credentials to Arabness; ʿAbd al-Malik ibn Marwān, the fifth caliph of the Umayyad line (r. 65–86/685–705) proclaimed Arabic as the only language of administration (displacing Greek, Coptic, Persian, etc.). When the Abbasid dynasty overthrew the Umayyad caliphate, both language and ethnic affiliation became the target of non-Arab Muslims, most of whom were Persians, and claimed their merits even over Arabic grammar, denying any privileged position of the Arabs. This anti-Arab polemical movement is known as *šuʿūbiyya*. It appeared in the 2nd/8th century and reached its peak in the 3rd/9th century. In the 5th/11th century *šuʿūbiyya* appeared in al-Andalus, where anti-Arab polemics were practiced mainly by Berbers. In the Middle Ages, the term *šuʿūbiyya* fell out of use and reemerged at the time of Arab nationalism, when Arabness projected itself into the glorious past, calling modern *šuʿūbiyya* any call for the promotion of the vernaculars and any attack of Standard Arabic.

The elevated status of Arabic and its longevity are known to be attributed to the fact that it is the medium of the *Qurʾān* and God's mouthpiece (see Ferguson 1990). Its linguistic qualities are evoked several times in the *Qurʾān* (Q. 12/2) and its inimitability (→ *ʾiʿjāz*) and rhetorical perfection by the tradition (*ḥadīṯ*) and the exegesis (*tafsīr*), which suggest a tripartite affiliation between Islam, the Arabs, and the Arabic language. Since the Prophet was Arab, heaven's language is Arabic, and so is the language of the *Qurʾān*. This link was emphasized mainly by jurists and philologists, among whom stands out the famous imam aš-Šāfiʿī (d. 204/820). It was reaffirmed between the 19th and the 20th centuries by many Islamist reformers, e.g. Jamāl ad-Dīn al-ʾAfġānī (d. 1897), ʿAbd ar-Raḥmān al-Kawākibī (d. 1902), Muḥammad Riḍā (d. 1935), and Šakīb ʾArslān (d. 1946). Their vast rhetorical literature embraces with ease both Islam and Arabic, as if to say that the former is indissolubly connected with the latter (see Haim 1962).

The reconciliation process through language takes place, however, on both sides, Christian and Muslim. Like al-Ḥuṣrī, many Muslims praise the role the 'Christian Levantine' played in the revival of the Arabic language and the Arab national identity. Like ʾIbrāhīm al-Yāzijī (d. 1906), many Christians defend the modernization of Standard Arabic, urging Arabs to defend Arabic regardless of their faith. As for the incompatibility of the universalism of Islam and Arab nationalism, the latter seems to infect the former in the thought of some scholars and leaders, such as ʾAbū l-ʿAlāʾ Mawdūdī (d. 1979), who advocates an Islamic 'world-state', while other thinkers, such as al-Bazzāz, see no contradiction between nationalism and the universal character of Islam (Donohue and Esposito 1982).

4.3 *Language and education*

Language imposes cognitive categories that force an individual into a particular symbolic order in thinking, communicating, and the ordering of personal experience. Understanding the dynamics at work, Islamist and pan-Arabist leaders make the Arabization of the school system a primary goal. For nationalists, the educational field is crucial; the original circles that founded the Baʿthist Party in Syria, for instance, belonged to this same field.

Immediately after their independence, in the 20th century, the new Arab states urged the Arabization of school curricula. Some of them (e.g. Lebanon) kept their multilingual systems; others, such as Tunisia, tried Arabization but later backed down; and some others accomplished their transition to Arabic, Algeria for instance, where twelfth-graders graduated (in 1989) from a completely Arabic education system. As for the university system, while Syria managed to Arabize it, other Arab countries partially accomplished this, with the humanities and the social sciences largely or completely in Arabic and the scientific and technical sectors largely or completely in English or French. Scientific and technical instruction is conducted in English, even at Cairo's venerable Arab-Islamic institution al-Azhar.

Summing up the results for the Maghreb, Grandguillaume (1983:21) denounced the unsuitability of textbooks and pedagogy based on memorization, while for Algeria some sociologists like Abdelkader Yefsah (1990:381–382) state that the use of the Arabic language "leads straight to...the primacy of the religious". These researchers seem to say that Arabic and Islam are mutually reinforcing, for they partially share common cultural, cognitive, and symbolic thinking 'tool kits' (Moatassime 1992).

4.4 *Language and society*

Immediately after the death of the Prophet Muḥammad in 10/632, the Islamic conquests marked a turning point in the history of Arabic and the languages in use in the conquered empire. Later processes of contact and blending between the different Arabian superstrata (being a mixture of different tribes) and substratal languages into which they were imported generated 'New-Arabic' (or 'Neo-Arabic') types, the ancestors of the currently known spoken varieties. Since 750 C.E., thousands of grammarians have illustrated the *fuṣḥā* as a fully declined system (→ *faṣīḥ*). The disappearance of declension, if it ever was present in the conquerors' dialects, is an intrinsic dynamic factor in linguistic renewal. There is no global explanation for the evolution of individual dialects, since the historical sociolinguistic situation of each particular area is still terra incognita. The substratal influence and analytical nature of the spoken varieties are relevant, but they are not the only explanation for this historical development, or the only demarcation line between Standard Arabic and spoken varieties.

Since its standardization, the spread of Standard Arabic has been the exclusive task of the education system. For more than one thousand years, most of Greater Syria and the Nile Valley were under non-Arab rulers (Mamluks, Seljuqs, Ottomans, etc.). Schooling and literacy were not widespread. Arabic was operative in the religious and traditional education systems of the libraries (*maktaba*) and Islamic education institutes (*madrasa*), and only a small minority, most of whom were *ʿulamāʾ*, scholars who were versed in the Muslim sciences, had full access to it.

In the 19th century, the Arabic press proposed Arabic as a media tool. Arabic was seen as a uniting factor whereas neo-Arabic dialects were symbols of fragmentation. Arabic academies were supposed to play a central role in this linguistic revival (→ language academies), but the new Arab states failed to create a pan-Arab academy or to cooperate. After the foundation of the Academy of Damascus (1919), new academies were founded in Cairo (1932), Iraq (1947), and Jordan (1976). Before they found their way into dictionaries, new terms were preceded by long periods of heated discussion, and some were considered too artificial to be used after all (see Ali 1987). New terms were being created every day, sometimes at breakneck speed; dictionary compilation was thus assigned to individual experts in specific fields.

Despite all efforts, and more than a century and a half after the foundation of the School of Languages in Cairo (1837), Standard Arabic remains no one's mother tongue. People learn it at school for years and still make many mistakes in writing it. The idea of an absolute superiority of Arabic (practically Classical Arabic) has also been strengthened by the fact that Western Orientalists, with only a few exceptions, learned Arabic as a dead language (like Latin) and had, also with very few exceptions, no speaking proficiency, neither in Classical Arabic nor in any dialect. The failure of the traditional teaching of Arabic in Europe has greatly contributed to the idea widely spread among Arabs that Arabic is so difficult and marvelous that no European can learn it properly.

As for the prestige of Arabic, linguistic upgrading is not necessarily accomplished only through classicizing. "In the case of the Jordanian students, it is the urban dialect that is regarded by most speakers as the prestige variety" (Versteegh 1997:193).

4.5 *Language and/or dialects*

Among other linguistic situations, Arabic was described by Ferguson (1959) as 'diglossic', referring to Classical Arabic as a high variety (H) and Colloquial Arabic as a low variety (L). It is beyond the scope of the present article to discuss this topic (see Kallas 1999:59–97; → diglossia); the concern here is with the different solutions nationalism gave to this linguistic situation.

As mentioned earlier, pan-Arab nationalism considers Standard Arabic *fuṣḥā* as the most important link binding the Arab nation and the colloquial languages, and those who study the colloquial language as suspicious. All through her book dealing with the "call for adopting vernacular in Egypt, history and repercussions", Naffūsa Zakariyyā Saʿīd (e.g. 1964:9, 18, 37) accuses the West, mainly Britain, of encouraging Egyptians, through Spitta and Orientalism, to occupy themselves with their dialect, to be able to learn their vernacular and spy, divide, and subjugate the Egyptians. While this may not be totally wrong, the role vernacular poetry played in the social and anti-colonial struggles cannot be ignored (→ dialect literature). Vernacular poetry called for British withdrawal from Yemen (ʿAbdallāh ʾAḥmad ʿĀmir and ʾAḥmad Faḍl al-ʿAbdalī), Iraq (Mollā ʿAbbūd al-Karxī), and Egypt (ʿAbdallāh an-Nadīm and Fūʾād Nigm). It celebrated Nasser's revolution in Sudan (Ḥardallū) and criticized the French mandate over Lebanon (ʿUmar ez-Zeʿennī) and Syria (Salāma al-ʾArġawānī).

5. REGIONAL AND NATIONAL-STATE PATRIOTISM

5.1 *Syrian nationalism*

Before the crystallizing of pan-Arab nationalism, a Syrian *patrie* was present in the mind of the Christian Lebanese Buṭrus al-Bustānī (d. 1883), including a physical entity called 'Greater Syria'. A roughly similar entity called the 'Fertile Crescent' became a national-state dogma in the hands of another Christian Lebanese, ʾAnṭūn Saʿāda (killed in 1949). It included Iraq, Syria, Lebanon, Jordan, Palestine, and, for some reason, Cyprus. Needless to say, this entity did not remain a Lebanese political tendency. After the Anglo-Iraqi war of 1941, Nūrī s-Saʿīd, back in power, emphasized the creation of a 'Greater Syria' out of Syria, Lebanon, Jordan, and Palestine, with which Iraq would be associated. The role of Standard Arabic in this Greater Syria was subject to some debate. It was an essential requisite in al-Bustānī's mind, for example, but it was not a means of self-identification in Saʿāda's doctrine (Saʿāda 1951). To him, Arabic was one of several languages the Syrians have adopted through their history. If a national language were to characterize them, it would have to be Syrianized Arabic.

5.2 *Lebanese nationalism*

In September 1920, the League of Nations declared the newborn Lebanese state under a French mandate, which intensified the debate over pan-Arab, regional, and specifically Lebanese nationalism. The latter movement was led by Michel Chiha (d. 1954) and Jawād Būlus (d. 1982), calling for a liberal, multiethnic and multilingual Lebanese formula, with a touch of Phoenician ascendancy. Basically, this formula had little appeal to Muslim communities.

When the Lebanese Constitution was fixed, French was considered (Art. 11) an official language alongside Arabic. Many Catholics appreciated this solution because it conceptualized Lebanese national identity in a non-Islamic but Mediterranean culture. Some of them were Francophiles, using French as a medium of literary expression. Among them, Choucri Ghanem (d. 1930), Jacques Tabet (d. 1956), and other thinkers were convinced of their Phoenician past and their universal voca-tion; they were gathered (in 1920) by Charles Corm's (d. 1963) editorial initiative *La Revue Phénicienne*. The most famous of them were Elie Tyane (d. 1957), Hector Klat (d. 1976), and Michel Chiha (d. 1954). To this same trend belonged Saīd Akl (b. 1912), who was to lead the philo-Phoenicians into the 21st century.

After an exceptional literary career in Standard Arabic, Akl broke with the language and published his first collection of poems (*Yāra*, 1961) in the Lebanese vernacular, using his own version of the Roman alphabet. To promote this experience, Akl founded several publishing houses, launched the weekly newspaper *Lebnān* (in 1983), and established many prestigious prizes. The effects of his determination are still apparent in 2006, although it is still for the most part a one-man movement (Akl 1997; Płonka 2004).

Currently, nationalist Catholics are the keenest on differentiation through the use of French (Abou 1961). Some, however, opposed both Lebanese Francophiles and colloquialists. For example, Kamāl Yūsuf al-Ḥājj (killed in 1976) conceived of complete bilingualism (Arabic/French) as an impossibility that only serves to divide the Lebanese, and he saw the diglossia (*izdiwāj al-luġa; fuṣḥā* vs. Lebanese) as a universal duality of psyche and mind, essence and existence, a basic human need. For him, the solution was in simplifying Arabic (al-Ḥājj 1978).

Currently, globalization and economic priorities are taking over all sectarian ideologies. This is best shown in the new educational policies adopted by the formerly Arabized Sunnite al-Maqāṣid and the mostly Francophile University of Saint Joseph, both introducing more and more English in their curricula.

5.3 *Egyptian nationalism*

Egyptian nationalism goes back to the rise to power of Muḥammad ʿAlī in 1805 and his rivalry with the Ottoman sultan. He launched educational reforms, founded the School of Languages (Dār al-ʾAlsun), whose main task was to translate foreign works into Arabic, and appointed as its head Rifāʿa aṭ-Ṭahṭāwī (d. 1873), who promoted the vision of Egypt as a nation, culturally distinct, historically unique, and territorially well defined, aiming at being a powerful and modernizing state. To many Egyptian nationalists of the time, modernizing meant breaking with the Arab past. They basically agreed with aṭ-Ṭahṭāwī's vision. While all agree on the need for modernizing the language, variations existed as to strategies to be adopted in realizing this vision. Some, like Muḥammad Ḥusayn Haykal (d. 1956), were less radical toward the language issue; others, like Ṭāhā Ḥusayn (d. 1973), were opposed to the use of vernacular and called for linguistic reform; some others, like ʾAḥmad Luṭfī as-Sayyid (d. 1963), supported a modernizing linguistic strategy so as to develop a middle or Egyptianized Arabic (*tamṣīr al-luġa*); others, like Louis ʿAwwād and ʾIbrāhīm Jumʿa, maintained that Arabic reflects desert values and that it was imposed on Egypt, which has its own language that should be used in writing and literature. The most radical, however, was Salāma Mūsā (1945), who went so far as to call for the adoption of a Roman alphabet, although he never put his call to practice. By the end of the 1930s, the increasing role of Islam softened Egyptian nationalism, which eventually turned more in the direction of Arabism (Suleiman 1996, 2003:169–204).

5.4 *Tunisian nationalism*

The Tunisian president Habib Bourguiba cannot be considered a supporter of Arabization. He publicly supported the Tunisian vernacular (29 July 1968), declaring that Classical Arabic was not the language of the Tunisian people. The Tunisian language, he proclaimed, is the Tunisian vernacular. Tunisification instead of Arabization was debated at the National Assembly (December 1970). Among Tunisian intellectuals, Hedi Balegh was the most convinced (1973–1974) of the use of the Tunisian vernacular instead of what he called "deux langues aristocratiques" (Standard Arabic and French). In 1958, a monolingual Arabic section was created in the Tunisian scholastic system alongside the bilingual sections. Later, the Arabic section was closed, officially because of lack of interest on the part of the parents in sending their children to monolingual courses (Grandguillaume 1983:59–68).

6. ARABIZING AND THE 'NON-ARAB' MINORITIES

Many ethnolinguistic minorities live in Arab states (→ multilingualism): Armenian (Lebanon, Iraq, Syria, Jordan, Palestinian West Bank, and Gaza), Assyrian → Neo-Aramaic (Iraq, Syria), Azerbaijani (Iraq, Syria), Balochi (United Arab Emirates, Oman), Bathari (Yemen, Oman), → Berber (Morocco, Algeria, Libya, Tunisia, Egypt), Chaldean Neo-Aramaic (Iraq), Chechen (Syria, Jordan), Cushitic Bedawi (Sudan), Domari (Iraq, Syria, Jordan, Palestinian West Bank, and Gaza, Egypt, Sudan, Libya; → Gypsy Arabic), Farsi (United Arab Emirates, Oman, Qatar), Harsusi (Oman), Chadic → Hausa (Sudan), Kordofanian Acheron (Sudan), Kumzari (Oman), → Kurdish (Iraq, Syria, Lebanon), Mehri (Yemen, Oman, Kuwait), Niger-Congo Languages (Sudan), Nilo-Saharan Languages (Sudan), Pashto (United Arab Emirates), Shehri (Oman), Soqotri (Yemen), Turkish (Iraq, Syria), Turoyo (Syria), West Circassian Adygey (Iraq, Syria, Jordan), Western Neo-Aramaic (Syria), Nile → Nubian (Egypt), Tigré (Sudan). Only the linguistic rights of Berbers in Morocco and Algeria and Kurds in Iraq are discussed in this article.

6.1 *Berbers*

Descendants of the pre-Arab inhabitants of North Africa, the Berbers are scattered across Morocco, Algeria, Tunisia, Libya, Egypt, Mauritania, Mali, Niger, and Burkina Faso. They speak various languages belonging to the Afro-Asiatic language family. To write → Berber, three alphabets are used: Tifinagh, Arabic, and Latin-based scripts. Collectively, the Berbers refer to themselves simply as 'Imazighen' which has become a major indicator of Berber self-awareness and nationalism (→ language shift: Amazigh). In Libya, however, any suggestion that a Berber might be a non-Arab remains taboo.

Despite the Berber commitment to the cause of national unity and liberation in North Africa, both the Algerian and the Moroccan regimes have systematically pursued de-Berberization and Arabization policies since independence. However, centuries of social contacts, cultural borrowing, and intermarriage have made any purely ethnic distinction an arduous task. Therefore, what are known as Arabic-speaking groups are constituted to varying degrees by Arabized Berbers. Berber and Arab identities in these countries are generally defined by language rather than racial distinction.

When Morocco and Algeria achieved independence, their first constitutions made it clear that they were fully sovereign Islamic states whose official language was Arabic. Not a single mention was made of Berber, either as a language or as a part of the identity of the country.

The Berber demands originated in 1967, the year in which the Berber Academy was established in Paris to promote Berber culture and language. In the same year, a similar association, the Moroccan Association for Research and Cultural Exchange, was founded in Morocco. An event of paramount importance in the modern struggle of the Berber movement is the Berber Spring (*Tafsut n Imazighen*) of 1980. This collective awareness led to the tacit creation of university departments and university courses in Algiers, Tizi-Ouzou (1990), and Bejaia (1991), and to the politicization of the Berber movement. The Algerian reaction to these developments was the promulgation of Law 91–05 (16 January 1991), which considered null any official document written in a non-Arab language (Art. 29) and imposed heavy sanctions as punishment for any infringement of this law (Art. 31). Arabic was laid down as the only teaching language permitted, and a commitment was made fully to Arabize the education system by 5 July 1994 (Art. 37). Chairs for Berber studies were closed, and Berber singers were banned from singing in their own language. Language recognition demands were behind a school boycott (known as the schoolbag strike) throughout the Kabyle area from September 1994 to April 1995. In 1995, President Liemin Zeroual announced the creation of a High Commission for Amazighity. In the preamble of the revised Algerian Constitution (December 1996), Islam, Arabness, and Amazighity were considered fundamental components of Algerian identity, but only Arabic is declared the official language (Art. 3).

In Morocco, the Berber language question was publicly dealt with by the highest authorities on 20 August 1994. King Hassan II spoke of the importance of including Berber in the Moroccan education system. However, this was not achieved until September 2003.

In 2003, the Algerian authorities were compelled to recognize Tamazight as a national language, but Berbers there want it to have equal status as an official language alongside

Arabic. In 2004 Algerian Berbers threatened to boycott the presidential election over the language issue.

Recently both Algeria and Morocco have undertaken to integrate Berber into their education systems, although there are still some differences between their policies. The teaching of Berber in Morocco is intended to cover the whole kingdom, while in Algeria it is limited to the areas where there is a high concentration of Berber speakers. As to the alphabet chosen to teach Berber, activists in Morocco promote the Tifinagh alphabet, whereas Algerian Berbers prefer the Roman (Latin) script.

6.2 *Kurds*

For centuries, Kurds fought their conquerors' battles and wrote in their languages (Persian, Turkish, Arabic, Russian, etc.). A widespread conviction among Kurds was that their mother tongues were not a suitable medium for high literature (→ Kurdish). They live mainly in five different states (eastern Turkey, northern Iraq, northwestern Iran, northern Syria, and Armenia).

They use two main dialect groups closely related to Persian, and many subdialects (*Kermanšahī, Guranī, Zaza*, etc.). The northern group is called *Kurmànjī* (known as *Hawar* in Turkey), and the central group is called *Kurdī*, or *Sòrànī*. Their alphabet was Arabic-based until 1920. Later, Kurdish intellectuals used the Latin alphabet in Syria and Turkey to write their *Kurmànjī*, the Arabic alphabet in Iran and Iraq to write *Sòrànī*, and the Cyrillic alphabet in the Soviet Union (since 1939). When in 1918 British troops occupied what is now known as Iraqi Kurdistan, Kurdish replaced Turkish as an administrative language and made its entrance into school curricula. For a very brief period the Kurds enjoyed a form of limited autonomy within Iraq. In Syria, the French Mandate helped revitalize the Kurdish language and a Turkicized Latin-based alphabet was adopted in Kurdish periodicals. But when Syria became an independent Arab republic, Kurdish publications were banned, and many Kurdish intellectuals fled overseas. Under the Syrian Baʿthist regime, Kurds are required to feel that they are Arabs and act as Arabs. Under the Iraqi Baʿthist regime, Kurds had to act as Arabs as well. But in spite of Baghdad's iron fist, in Iraqi Kurdistan the Kurdish language and culture were better safeguarded.

The advent of the March Manifesto changed the general political atmosphere in the early 1970s and paved the way for wider autonomy. In Iraqi Kurdistan, school curricula allowed ethnic minorities to be educated in their mother tongues (Kurdish, Turkish, Neo-Aramaic), while in the rest of Iraq the only teaching language was Arabic. A Kurdish Academy was established in Baghdad (1971) and a Kurdish University in Suleimania (1968–1981, 1992–), but they were strictly controlled by the authorities. Currently (2006), Kurdish Universities are also present in Erbil and Dahok. Several publishing houses, cultural centers, newspapers, magazines, and radio and TV stations are flourishing in Iraqi Kurdistan, as well as in the Kurdish diaspora (mainly in Sweden, France, and Britain), where they were already active (Hassanpour 1992; Galletti 2002).

In the transitional Iraqi constitution, promulgated in March 2004, initial steps have been taken to restore linguistic pluralism. Article 9 defines both Arabic and Kurdish as the two official languages of Iraq, and undertakes to guarantee the "right of Iraqis to educate their children in their mother tongue, such as Turkmen, Syriac, or Armenian, in government educational institutions in accordance with educational guidelines, or in any other language in private educational institutions".

7. ARABIC, ARABNESS, AND ISLAM IN THE CONSTITUTION

A number of observations can be made based on the analysis of the constitutions of the 22 members of the Arab League, including the Iraqi Transitional Law (8 March 2004) and the Palestinian Draft Basic Law (June 1994).

First, Arabness oscillates between a minimal and a maximal level. At the minimal level, all constitutions declare Arabic the official and/or the national language. At the medium level, they define their states as Arab countries, or declare their people as belonging to a wider Arab nation, world, or family, the exceptions being the → Comoros, → Djibouti, → Somalia, and → Sudan. At the maximal level, some announce that the state will act to unify the

Arab nation (Bahrain, Egypt, Kuwait, Saudi Arabia, Syria, and Libya) or the Greater Arab Maghreb (Mauritania and Morocco).

Second, Islam oscillates between a minimal and a maximal level as well. At the minimal level, Islam is not mentioned at all (Lebanon, Palestine, and Djibouti). At the medium level, Islam is the religion of the remaining states. At the maximal level, 14 states declare themselves Islamic and/or declare Islamic law (*aš-šarīʿa*) to be the source of law (Bahrain, Egypt, Iraq, Kuwait, Libya, Mauritania, Morocco, Oman, Qatar, Somalia, Sudan, United Arab Emirates, Yemen, and Saudi Arabia).

Actually, what the constitutions of the Arab League states share most is the desire to consolidate their individual sovereignty and national identity. This principle is stressed in the first two articles of the Arab League charter as well.

8. Conclusion

A bird's eye view of the Mediterranean linguistic space reveals a contrast between the southern and the northern shores. The former seems to be homogeneous; the latter looks like the Tower of Babel, where the European Union pays an army of translators to keep up to date and to disseminate its institutional communications to an increasing number of member states. Nonetheless, while the latter moves toward an economic and political unity, the member states of the Arab League have no unity on their agenda, even though they share a common official language and despite their frequent assertions of brotherhood and commitment to a pan-Arab nation.

The Arab League defines an Arab as a person "whose language is Arabic, who lives in an Arabic speaking country, who is in sympathy with the aspirations of the Arabic-speaking peoples". However, Djibouti, Somalia, and the Comoros are all member states of the Arab League, even though their inhabitants are not predominantly Arabic-speaking. On the other hand, the → Maltese language is closely related to Tunisian Arabic, but Malta does not use Standard Arabic, and its inhabitants do not consider themselves Arabs. Chad, Eritrea, and Israel all recognize Standard Arabic as an official language, but none of them are members of the Arab League. Mali and Senegal recognize → Ḥassāniyya, the Arabic dialect of their Moorish minorities, as a national language, but grant no official status.

Many argue that the Arab states, born out of the disintegration of the Ottoman Empire, are artificial. However, they are now an established reality, and their authoritarian nature may lead to the crystallization of new ways of conceiving linguistically based ideologies.

Religious minorities (specifically Christians) promptly supported pan-Arab nationalism, for it gave them the illusion that through their linguistic status they could avoid religious discrimination and obtain equal civil rights. They were disappointed because their states' constitutions continued to consider "Islamic jurisprudence a major source of legislation". Linguistic minorities feared it because of its hegemonic tendency. They witnessed radical Islamic movements imposing Arabic and ignoring their linguistic rights.

The encounter with the West has led Arabs to ask themselves: 'Who are we?' 'Who are they?' Answers are mostly rhetorical. They emphasize subjective cultural factors and hardly touch on civil or political rights. The legitimacy of political entities depends more and more on their capacity to achieve unity in diversity and to pledge equal rights to all citizens. Arabic linguistic affiliation alone cannot build a modern nation.

Bibliographical references

Primary source

Ibn Bābawayhi, *Maʿānī l-ʾaxbār* = ʾAbū Jaʿfar Muḥammad ibn ʾAbū l-Ḥasan ʿAlī ibn Ḥusayn ibn Mūsā al-Qummī, *Maʿānī l-ʾaxbār*. Najaf: al-Ḥaydariyya Library and Press, 1971.

Secondary sources

Abou, Sélim. 1961. *Le bilinguisme arabe-français au Liban*. Paris: Presses Universitaires de France.

ʾAfġānī, Saʿīd al-. 1962. *Ḥāḍir al-luġa al-ʿarabiyya fī Bilād aš-Šām*. Cairo: Jāmiʿat ad-Duwal al-ʿArabiyya.

ʿAflaq, Michel. 1978. *Le point de départ*. Madrid: PBas.

Akl, Saīd. 1997. *Yaara: Inno alla donna*. Ed. Elie Kallas. Venice: Ca'Foscarina.

Ali, Abdul Sahib Mehdi. 1987. *A linguistic study of the development of scientific vocabulary in Standard Arabic*. London and New York: Kegan Paul International.

Antonius, George. 1938. *The Arab awakening*. Beirut: Khayats.

ʾArsūzī, Zakī al-. 1962. *ʿAbqariyyat al-ʿarabiyya fī lisānihā*. 2nd ed. Damascus: Dār al-Yaqḏa al-ʿArabiyya.

Donohue, John J. and John L. Esposito (eds.). 1982. *Islam in transition: Muslim perspectives.* New York: Oxford University Press.

Ferguson, Charles A. 1959. "Diglossia". *Word* 15.325–340.

——. 1990. "Come forth with a *surah* like it: Arabic as a measure of Arabic society". *Perspectives on Arabic linguistics*, I, ed. Mushira Eid, 39–51. Amsterdam and Philadelphia: J. Benjamins.

Fishman, Joshua A. 1972. *Language and nationalism: Two integrative essays.* Rowley, Mass.: Newbury House.

Galletti, Mirella. 2002. *Incontri con la società del Kurdistan.* Genoa: Name editore.

Grandguillaume, Gilbert. 1983. *Arabisation et politique linguistique au Maghreb.* Paris: Maisonneuve and Larose.

Ḥājj, Kamāl Yūsuf al-. 1978. *Fī falsafat al-luġa.* Beirut: Dār an-Nahār.

Ḥuṣrī, Sāṭiʿ al-. 1955. *al-ʿUrūba ʾawwalan.* Beirut: Dār al-ʿIlm li-l-Malāyīn.

Haim, Sylvia (ed.). 1962. *Arab nationalism: An anthology.* Berkeley: University of California Press.

Hassanpour, Amir. 1992. *Nationalism and language in Kurdistan (1918–1985).* San Francisco: Mellen Research University Press.

Kallas, Elie. 1996. "Rivendicare il futuro, costruire il passato e fingersi una nazione". *Dove va l'arca di Noè: Nazionalismo arabo-islamico, nazionalismo israeliano e le minoranze,* Futuribili 1, ed. Elie Kallas, 15–49. Milan: Franco Angeli.

——. 1999. *Qui est arabophone?* Gorizia: Istituto di Sociologia Internazionale di Gorizia (I.S.I.G.).

Moatassime, Ahmed. 1992. *Arabisation et langue française au Maghreb.* Paris: Presses Universitaires de France.

Mūsā, Salāma. 1945. *al-Balāġa al-ʿaṣriyya wa-l-luġa al-ʿarabiyya.* Cairo: al-Maṭbaʿa al-ʿAṣriyya.

Naṣṣār, Nāṣīf. 1986. *Taṣawwurāt al-ʾumma al-muʿāṣira.* Kuwait: Mu'assasat al-Kuwayt li-t-Taqaddum al-ʿIlmī.

Płonka, Arkadiusz. 2004. *L'idée de langue libanaise d'après Saʿīd ʿAql.* Paris: Geuthner.

Saʿāda, ʾAnṭūn. 1951. *Nušūʾ al-ʾumam.* 2nd ed. Damascus: A. Saʿāda Publisher.

Saʿīd, Naffūsa Zakariyyā. 1964. *Tārīx ad-daʿwā ʾilā l-ʿāmmiyya wa-ʾāṯāruhā fī Miṣr.* Cairo: Dār al-Maʿārif; al-ʾIskandariyya: Dār Našr aṯ-Ṯaqāfa bi-l-ʾIskandariyya.

Suleiman, Yasir. 1996. "Language and identity in Egyptian nationalism". *Language and identity in the Middle East and North Africa*, ed. Yasir Suleiman, 25–37. Richmond, Surrey: Curzon Press.

——. 2003. *The Arabic language and national identity: A study in ideology.* Edinburgh: Edinburgh University Press.

Vatikiotis, Panayotis J. 1991. *Islam and the state.* London and New York: Routledge.

Versteegh, Kees. 1997. *The Arabic language.* Edinburgh: Edinburgh University Press.

Yefsah, Abdelkader. 1990. *La question du pouvoir en Algérie.* Algiers: ENAP.

Zeine, Zeine N. 1973. *The emergence of Arab nationalism.* 3rd ed. Delmar: Caravan Books.

Elie Kallas (University of Triest)

Nawāsix

The term *nawāsix* (sg. *nāsixa*) 'cancelers' is borrowed from Muslim canonic law, where the verb from which it is derived, *nasaxa* 'to make something disappear', is used with the technical sense 'to abolish', with reference to a verse from the *Qurʾān* or a tradition about the Prophet which revokes another one. Used by the later Arab grammarians, probably not earlier than the time of ʾAbū Ḥayyān (d. 1344; see Carter 1981:207), to designate a category of 'operators' (*ʿawāmil*, sg. *ʿāmil*; → *ʿamal*) grouped by some previous grammarians under headings which indicate that these elements change the status of the basic members of the prototypical nominal sentence, the theme (*mubtadaʾ* 'inchoative' → *ibtidāʾ*) and the nominal predicate (→ *xabar* 'information', 'comment'): *bāb al-ʿawāmil ad-dāxila ʿalā l-mubtadaʾ wa-l-xabar* 'chapter on the operators which affect the theme and the predicate', or more specifically, *mā yarfaʿu l-mubtadaʾ/al-xabar wa-yanṣibu l-xabar/al-mubtadaʾ* '[elements] which assign the independent case [nominative] to the theme/predicate and which assign the dependent case [accusative] to the predicate/theme'. Words belonging to this category are subdivided into three main subcategories, in accordance with the kind of 'modification' they produce, and named after the element which appears as prototypical of the series of 'analogues' (*ʾaxawat* 'sisters'; see Carter 1981:206–238):

i. *kāna* 'to be' and its analogues, elements which assign nominative case to the theme and accusative case to the predicate: *ʾamsā* 'to be in the evening, to become'; *ʾaṣbaḥa* 'to be in the morning, to become'; *ʾaḍḥā* 'to be in the forenoon, to become'; *ḍalla* 'to remain'; *bāta* 'to be at night, to remain'; *ṣāra* 'to become'; *laysa* 'not to be'; *mā zāla* 'not to cease'; *mā nfakka* 'not to stop'; *mā fatiʾa* 'not to refrain'; *mā bariḥa* 'not to desist'; *mā dāma* 'as long as it remains' (→ *kāna wa-ʾaxawātuhā*);

ii. *ʾinna* 'verily' and its analogues, which assign accusative case to the theme and nominative case to the predicate: *ʾanna* 'that'; *lākinna* 'but'; *kaʾanna* 'as if'; *layta* 'would that' [particle of wishing]; *laʿalla* 'perhaps', to express hope or expectation (→ *ʾinna wa-ʾaxawātuhā*);

iii. *ḏ̣anantu* ‘I thought, I believed’ and its analogues, which assign accusative case to the theme and to the predicate: *ḥasibtu* ‘I reckoned’, *xiltu* ‘I imagined’, *zaʿamtu* ‘I asserted’, *raʾaytu* ‘I regarded’, *ʿalimtu* ‘I knew’, *wajadtu* ‘I found’. All these verbs are quoted in their 1st person singular past form instead of the conventional 3rd person singular masculine; all are used with the meaning ‘I find that, I consider’ (in the present). Their appearance here in the 1st person could be explained by the fact that they are, in Arabic as in other languages, verbs meant to convey the presence of the speaking subject in the utterance. Some grammars also add to the verbs mentioned in this class the verbs *ittaxaḏa* ‘to adopt’, *jaʿala* ‘to make’, *samiʿa* ‘to hear’, whose semantic relation to the aforementioned verbs is less transparent.

In addition to these three classes, some grammars also add to *an-nawāsix* the category of ‘verbs of imminence’ (*ʾafʿāl al-muqāraba*), whose prototype is *kāda*.

The concept shared by the category of functional elements subsumed under the name of *an-nawāsix* is part of the Arab grammarians’ view on case assignment (→ *ʾiʿrāb*), which posits the existence of an *abstract* case that may or may not have a phonological manifestation (for example, declension is virtual in the case of the predicate in the sentence *kāna zaydun yaktubu* ‘Zayd was writing’). The very ‘action’ of the ‘operators’ known as *an-nawāsix* is abstract (→ *ʿamal*). Theoretically, all *nawāsix* have two arguments, which they derive from the ‘modified’ version of the thematic structure (*jumla ismiyya* ‘nominal sentence’): *mubtadaʾ* ‘inchoative’, the theme, becomes ‘noun’ (*ism kāna* ‘the name [corresponding to] *kāna*’, for example), and *xabar* is the ‘predicate’ (*xabar kāna* ‘the predicate of *kāna*’, for example). The view of the Arab grammarians starts to make sense if we accept that all *nawāsix* belong to the class of modalities, as defined a long time ago in logic and taken over in linguistics in modern times. From a semantic point of view, all these elements belong to the class of modal expressions or modalities, i.e. expressions which introduce further qualifications to a given sentence (of the type ‘it is known, admitted, possible, desirable, etc.’ for *ʾinna* and its analogues; ‘it is always the case, it is sometimes the case’ for *kāna* and its analogues; of the type ‘I believe, I suppose, I find’ for *ḏ̣anna* and its analogues; etc.). Arab grammarians attempted to provide a semantic definition for at least one of the subcategories of this category (for instance, verbs in the third subcategory are called *ʾafʿāl aš-šakk wa-l-yaqīn* ‘verbs of doubt and certainty’), and they also discussed the ‘verbal force’ of particles of the category *ʾinna* (the latter paraphrased by *ʾuʾakkidu* ‘I affirm, I confirm’) in order to justify their ‘action’ (*ʿamal*) on a theoretical level through comparison with the verb (‘the strongest operator’) from a semantic and formal perspective. All the elements belonging to the class of *nawāsix* can be seen as general predicates, normally placed at the beginning of the sentence in order to signal that what is being uttered is not ‘objective’ but rather a statement made from the speaker’s point of view. The relation of this shift of perspective to a virtual change of case (neither ‘nominative’ nor ‘accusative’, but different), shows how one can theoretically conceive a relation between the categories of case and modalization starting from the material provided by the Arab grammarians.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Ibn ʿAqīl, *Šarḥ* = Bahāʾ ad-Dīn ʿAbdallāh ibn ʿAbd ar-Raḥmān Ibn ʿAqīl, *Šarḥ Ibn ʿAqīl ʿalā ʾAlfiyyat Ibn Mālik*. Ed. Muḥammad Muḥyī d-Dīn ʿAbd al-Ḥamīd. Cairo, 1931.

Ibn Hišām, *Muġnī* = Jamāl ad-Dīn ʾAbū Muḥammad ʿAbdallāh ibn Yūsuf Ibn Hišām, *Muġnī l-labīb ʿan kutub al-ʾaʿārīb*. Ed. Māzin al-Mubārak and Muḥammad ʿAlī Ḥamdallāh. 2 vols. Damascus: Dār al-Fikr, 1969.

Ibn ʾĀjurrūm, *ʾĀjurrūmiyya* = ʾAbū ʿAbdallāh Muḥammad ibn Muḥammad ibn Dāwūd aṣ-Ṣanhājī. *al-ʾĀjurrūmiyya*. Ed. and trans. (into French) L.-J. Bresnier. 2nd ed. Algiers, 1866.

SECONDARY SOURCES

Anghelescu, Nadia. 1973–1974. “Sur le rôle de *an-nawāsix* dans l’acte de la parole”. *Mélanges de l’Université Saint-Joseph* 48.291–302.

——. 1999a. “Eléments d’une théorie des modalités chez les grammariens arabes anciens”. *Revue Roumaine de Linguistique* 27:4.287–291.

——. 1999b. “Modalities and grammaticalization in Arabic”. *Arabic grammar and linguistics*, ed. Yasir Suleiman, 130–142. Richmond: Curzon.

Carter, Michael G. 1981. *Arab linguistics: An introductory classical text with translation and notes.* Amsterdam: J. Benjamins.

NADIA ANGHELESCU (University of Bucharest)

Negation

Negation refers to the phenomenon whereby a function word with the meaning 'not' (negation marker) is used to negate a sentence (sentence negation) or a given constituent of the sentence (constituent negation). Sentence negation markers can be part of the auxiliary system and interact fairly intimately with some of its constituents, notably tense. Constituent negation markers tend to relate to focus, the phenomenon whereby a given constituent of the sentence is highlighted for discourse considerations. Negation markers can consist of a single function word or a complex of two function words, one of which has a variable distribution determined by grammatical context and, when a clitic, by prosodic considerations as well.

1. SENTENCE NEGATION AND TENSE: *LĀ, LAM, LAN, LAYSA*

The Standard Arabic *lā/lam/lan* are used to negate sentences with a verbal predicate. The differences between them relate to the tense information they carry over and above their function as sentence negation markers. The marker *lam* expresses past tense (Neg+Past), *lan* future tense (Neg+Future), and *lā* present tense (Neg+Present):

(1) *lam* *ʾaktub*
Neg.Past 1s.write.Imperf
'I did not write'

(2) *lan* *ʾaktub-a*
Neg.Fut. 1s.write.Imperf-Suff
'I will not write'

(3) *lā* *ʾaktub-u*
Neg.Pres. 1s.write.Imperf-Suff
'I do not write'

The present tense *lā* is likely to be the variant closest to the citation or base form of the negation marker on the grounds that present tense is generally unmarked in Standard Arabic. The variant *lā* appears in sentences where tense is expressed independently of the negation marker, e.g. by the future modal *sawfa*:

(4) *sawfa* *lā* *ʾaḥḍur-u*
will Neg 1s.attend.Imperf-Suff
'I will not attend'

lā (and its variants) exert two major selectional restrictions on the predicate over and above the restriction that the predicate must be verbal. First, the verb must be in the imperfect form, widely thought to be unmarked for tense. This is expected in view of the fact that tense is expressed on Neg. The restriction extends to imperatives, which must include the person prefix of the imperfect form otherwise absent in non-negative imperatives, e.g. *ktub* 'you write!':

(5) *lā* *ta-ktub!*
Neg 2.write.Imperf.ms
'Don't (you) write!'

The second restriction is expressed in terms of vocalic suffixes on the verb, the nature of which is subject to debate. According to a widely held view, the suffixes express → mood distinctions which oppose the jussive, marked with a zero morpheme, to the subjunctive, marked with *-a*, to the indicative, marked with *-u*. Another view, arguably rooted in the Arabic linguistic tradition, takes the superficial similarity between some of these suffixes and the Case suffixes of nouns and adjectives (nominative *-u* and accusative *-a*) to be more than accidental. According to this view, the verbal suffixes are also Case endings, each of them governed by a different variant of *lā*. A recent defense of this view and the phenomenon of verbal Case can be found in Ellaty (1994).

laysa, often described as a negative copula, is used to negate sentences with a nonverbal predicate:

(6) *laysat* *l-bint-u*
be.Neg.3fs the-girl-Nom
mudīrat-an/ḥazīnat-an
director-Acc/sad-Acc
'The girl is not a director/sad'

The standard classification of *laysa* as a verb is due to at least three reasons. First, it triggers accusative Case on nominal and adjectival predicates, which otherwise carry nominative Case. Second, it inflects for tense-agreement and enters into agreement with the subject. Third, it occupies the initial position immediately before the subject usually reserved for the verb in the canonical order VSO. These properties are not inconsistent with an analysis which treats *laysa* as a compositional category consisting of *lā* and a copula 'to be' with a restricted distribution, the presence of which is necessary to satisfy the selection requirement of *lā*. If so, *laysa* does not constitute a separate sentence negation marker but rather is a variant (sister) of *lā*. One obvious problem with this classification is that *laysa* takes the tense agreement inflection of the perfect paradigm, suggesting that its *lā* constituent selects the perfect rather than the imperfect form of the verb. On the other hand, the perfect form of *laysa* is inconsistent with its present tense reading, suggesting that it does not have the compositional and transparent form that perfect verbs generally have. These apparently contradictory properties of *laysa* are not surprising in view of the fact that the copula is a historical remnant with the purely formal function of satisfying the selectional requirements of *lā*.

The selectional restrictions that it exerts on the predicate and the manner in which it interacts with tense indicate that *lā* (and its variants) is a clause-internal sentence negation marker (much like English *not*). Its structural and derivational aspects are explored in Benmamoun (2000), Fassi Fehri (1993), Ouhalla (1991, 1997, 2002), Shlonsky (1997), among other sources.

There is at least one context, cited in Moutaouakil (1993), where *lā* appears to function as a constituent negation marker with scope restricted to the constituent immediately following it:

(7) *lā* *rajul-a* *fī l-bayt*
Neg man-Acc in the-house
'There is no man in the house'

While it is not unusual for negation markers to function as both sentential and constituent negation markers, it is not clear how to reconcile the selectional restrictions of *lā* in the contexts that include both functions.

2. Constituent negation and focus: *mā*

Standard Arabic *mā* initially appears to overlap in function with *lā*, its distinctive property being that it has a wider and less restricted distribution. For example, it can occur in sentences with a nonverbal predicate as well as in sentences with a verbal predicate. Moreover, the verbal predicate can be in the perfect or the imperfect form:

(8) *al-bint-u* *mā* *ḥazīna*
the-girl-Nom Neg sad
'The girl is not sad'

(9) *mā* *kataba* *l-jawāb*
Neg write.Perf.3ms the-letter
'He has not written the letter'

(10) *mā* *yaktubu*
Neg 3ms.write.Imperf
'He doesn't write'

The noted properties show that *mā* does not exert any selectional restrictions on the predicate and does not interact with tense, at least not intimately. Other properties indicate further differences between it and *lā* and appear to show that *mā* is a constituent negation marker with scope restricted to the constituent immediately following it. The scope of *mā* can be reliably gauged through the use of *bal*-continuations, which pick out the negated constituent:

(11) *mā* *zayd* *fī* *l-bayt* *(bal ʿamr)*
Neg Zayd in the-house (but ʿAmr)
'It was not Zayd who was in the house. (It was ʿAmr.)'

The negated constituent can also be a preposed direct object ([*mā*+NP-Acc]):

(12) *mā* *zayd* *raʾay-tu* *(bal ʿamr)*
Neg Zayd see.Perf-1s (but ʿAmr)
'It was not Zayd I saw. (It was ʿAmr.)'

The constituent negation use does not necessarily exclude the possibility that *mā* could also be used as a sentence negation marker, e.g. in contexts (9) and (10). However, it seems plausible that *mā* is a constituent negation marker even in (9) and (10), with the negated constituent being the whole sentence [*mā*+S]. Sentence negation and constituent negation of a sentence (S) are logically equivalent, hence the impression that *mā* functions as a sentence negation marker in (8) and (9). The constituent negation use may itself be a function of a more fundamental property of *mā*, namely that it is a focus marker. In addition to helping gauge the scope of the negation marker, the *bal*-continuations bring out the focus reading of constituent negation. They spell out what the speaker believes is the more accurate value of the negated constituent. In other words, *mā* and the *bal*-continuation together have a discourse-related corrective function which can be described as contrastive focus. The focus properties of *mā* and related issues are explored in Moutaouakil (1989) and Ouhalla (1997), among other sources.

There is at least one context, cited in Ouhalla (1997), where the scope of *mā* appears to bypass the constituent immediately following it and affect a lower constituent which bears a prepositional constituent focus marker (FM):

(13) *mā ʾanā bi-šāʿir (bal rasūl...)*
Neg I FM-poet (but messenger...)
'I am not a POET (but a messenger...)'

This appears to be an instance of association focus seen in English sentences such as *John did not introduce MARY*, which has the constituent negation reading *[~ Mary, John introduced x]* (Jackendoff 1972). If this is indeed the correct analysis for (13), then *mā* has the function of a neutral sentence negation marker in this context. The focus reading is encoded by the prepositional constituent focus marker attached to the predicate rather than by *mā*.

3. Negation complex: *m...š*

The negation complex *m...š* is found in many spoken varieties of Arabic with largely similar patterns of distribution (see Mohamed and Ouhalla 1995; Benmamoun 2000). This presentation is restricted to data from Moroccan Arabic, where the negation complex has the form *ma-ši*.

The complex *ma-ši* is used to negate sentences with a verbal predicate as well as sentences with a nonverbal predicate. In sentences with a verbal predicate, *ma-* appears before the verb and *-ši* as an enclitic on the verb, following a clitic pronoun object if there is one. This is the case irrespective of whether the verb is in the perfect or imperfect form:

(14) *ma sarrədt-lum-ši*
Neg send.Perf.1s-to.them-Neg
l-kado
the-present
'I didn't send them the present'

(15) *ma ka-n-tkllǝm-ši*
Neg Prog-1s-play.Imperf-Neg
mʕa-hum
with-them
'I don't talk to them'

When the sentence includes the future modal *ġadi* or the auxiliary *kan* along with the main verb, *-ši* appears as an enclitic on the modal or the auxiliary:

(16) *ma ġadi-ši n-mši*
Neg will-Neg 1s-go.Imperf
'I won't go'

(17) *ma kan-u-ši*
Neg be-3pl-Neg
ka-yġanni-w
Prog-3ms-sing.Imperf-pl
'They were not singing'

However, when the future modal has the reduced form *ġa, -ši* appears as an enclitic on the main verb, suggesting that the distribution of *-ši* is subject to prosodic considerations in addition to structural ones:

(18) *ma ġa n-mši-ši*
Neg will 1s-go.Imperf-Neg
'I won't go'

As far as sentences with a nonverbal predicate are concerned, there are two patterns. The negation complex can appear on the left edge of the predicate:

(19) *samir ma-ši mudir/mriḍ/hna*
Samir Neg-Neg director/sick/here
'Samir is not a director/sick/here'

In the other pattern, *-ši* appears as an enclitic on the predicate, although this is generally restricted to situations where the predicate consists of a simple noun, adjective, or adverbial element:

(20) *samir ma mudir/mriḍ/hna -ši*
Samir Neg director/sick/here-Neg
'Samir is not a director/sick/here'

A complicating factor in their distribution is that the two members of the negation complex can appear on opposite sides of what appears to be the subject in nominal sentences with a pronominal subject:

(21) *ma huwwa-ši hna/mriḍ*
Neg he-Neg here/sick
'He is not here/sick'

In view of the fact that the distribution of the two members of the negation complex is generally determined relative to the predicate, it is not unreasonable to conclude that the pronoun in (21) corresponds to the inflectional, auxiliary system of the sentence linked to the predicate rather than to the subject. The subject is likely to be a null category the content of which is recoverable from the overt pronominal inflection (Eid 1983, 1991; Benmamoun 2000). Such sentences can have the more orthodox pattern for nominal sentences with the negation complex placed on the left edge of the predicate and the pronoun in the subject position:

(22) *huwwa ma-ši hna/mriḍ*
he Neg-Neg here/sick
'He is not here/sick'

The negation complex can also be used for the purpose of constituent negation with a contrastive focus reading. In this use, the two members of the complex invariably appear together on the left edge of the negated, focused constituent in the initial position of the clause:

(23) *ma-ši ktab šrit (majalla)*
Neg-Neg book buy.Perf.1s (magazine)
'It was not a book I bought. (It was a magazine.)'

(24) *ma-ši samir huwwa lli*
Neg-Neg Samir he who
šuft-u (rašid)
see.Perf.1s-him (Rashid)
'It was not Samir whom I saw. (It was Rashid.)'

The negated constituent can be an entire clause, including a negative clause:

(25) *ma-ši btasm (ḍḥək)*
Neg-Neg smile.3ms (laugh.3ms)
'He did not smile. (He laughed.)'

(26) *ma-ši ma tklləm-ši*
Neg-Neg Neg talk-3ms-Neg
'It is not the case/it is not true that he did not talk'

4. NEGATIVE POLARITY AND NEGATIVE CONCORD

Polarity expressions (or items) are expressions with no inherent, fixed meaning of their own but which can acquire a variety of readings from the grammatical contexts in which they occur. Good examples are the Standard Arabic expressions consisting of the determiner *ʾayy* and a bare noun [*ʾayy* N], which have a negative reading (negative polarity) when included in a negative sentence and a WH-word reading when included in a question:

(27) *lam ʾara ʾayy*
Neg.Past 1s.see.Juss any
ʾaḥad/šayʾ/kitāb
one/thing/book
'I did not see anyone/anything/any book'

(28) *ʾayy kitāb qaraʾ-ta?*
which book read.Perf.2ms
'Which book did you read?'

Negative polarity expressions are also found in some spoken varieties, including Moroccan Arabic, where they have the form [*ḥetta* N] (Benmamoun 1997). In non-negative contexts, these expressions have a non-negative reading which can be translated as 'also N' and 'even N', where N is a name:

(29) *ma šuft ḥətta waḥd/ḥaja/*
Neg see.Perf.1s anyone/anything/
ktab
any book
'I did not see anyone/anything/any book'

(30) *ḥətta samir ja*
also Samir come.Perf.3ms
'Samir also came'

Negative concord refers to the phenomenon whereby an expression that includes a negation element, and therefore has an inherent negative meaning, is used along with the sentence negation marker in a given sentence without yielding what is called a double negation reading. Negative concord is typically found in Levantine Arabic, where the negative expression has the form [*wala* N]. The examples below are from Lebanese Arabic (Lina Choueiri, p.c.):

(31) *ma* *šift* *wala* *ḥada/*
Neg see.Perf.1s no one/
ši/mʿallme
nothing/teacher
'I saw no one/nothing/no teacher'

A characteristic property of negative expressions is that they can mark negation in the absence of a sentence negation marker, although this is restricted to the subject function. Thus, while the Moroccan example (30) cannot have a negative reading, the Lebanese example (32) can:

(32) *wala* *walad* *ija*
no boy come.Perf.3ms
'No boy came'

There is at least one other test often used to distinguish between polarity and negative expressions. Polarity expressions cannot function as partial answers to WH-questions. In Moroccan Arabic, *ḥətta ʾaja* 'anything' is not a possible answer for the question *šnu baġi*? 'what do you want?'. In contrast, negative expressions can function as partial answers to WH-questions. In Lebanese Arabic, *wala ši* 'nothing' is a perfectly possible answer to the question *šu baddik?* 'what do you want?'

Despite the differences, polarity and negative expressions resemble each other in that they are in complementary distribution with *-š(i)* of the negation complex. All the examples above which include a polarity or a negative expression do not include the *-š(i)* of the negation complex. This suggests that these expressions and *-š(i)* have the same function, and that *-š(i)* performs this function on a default basis, i.e. in the absence of a polarity expression. These issues and their implications for the logical form of negative sentences are explored in Benmamoun (1997) and Ouhalla (2002).

5. CONCLUSION

Standard Arabic *lam*, *lan*, and *lā*, and arguably also *laysa*, are all variants of one sentence negation marker. The markers *lam* and *lan* are marked for tense, and *laysa* includes a copula. *lā* appears to be the unmarked form. The distribution of *lā* (and its variants) is consistent with it being a clause-internal sentence negation marker.

Standard Arabic *mā* has a distribution consistent with its being a constituent negation marker. This is confirmed by the contrastive focus reading to which it gives rise. To the extent that it also functions as a sentence negation marker, it occupies a peripheral position in the clause compared to *lā*.

The Moroccan Arabic negation complex *ma-ši* and its variants in other spoken varieties can function as a sentence negation marker or a constituent negation marker. The former function is typically associated with the distribution whereby *ma* appears on the left edge of the predicate and *-ši* is an enclitic on the predicate, although in nominal sentences the two members of the negation complex can be appear on the left edge of the predicate. This particular pattern of distribution is also characteristic of the constituent negation function.

Polarity and negative expressions can both be found in Arabic, the former in Standard and Moroccan Arabic and the latter in Levantine Arabic.

BIBLIOGRAPHICAL REFERENCES

Benmamoun, Elabbas. 1997. "Licensing of negative polarity items in Moroccan Arabic". *Natural Language and Linguistic Theory* 15.263–287.
——. 2000. *The feature structure of functional categories: A comparative study of Arabic dialects.* Oxford: Oxford University Press.
Eid, Mushira. 1983. "The copula function of pronouns". *Lingua* 59.197–207.
——. 1991. "Verbless sentences in Arabic and Hebrew". *Perspectives on Arabic linguistics*, III, ed. Bernard Comrie and Mushira Eid, 31–61. Amsterdam and Philadelphia: J. Benjamins.
Ellaty, Youḥanna. 1994. *Le cas verbal: Syntaxe et morphologie*. Ph.D. diss., University of Paris-8.
Fassi Fehri, Abdelkader. 1993. *Issues in the structure of Arabic words and clauses*. Dordrecht: Kluwer Academic Publishers.
Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar*. Cambridge, Mass.: MIT Press.
Mohamed, Ibrahim and Jamal Ouhalla. 1995. "Negation and modality in early child Arabic". *Perspectives on Arabic linguistics*, VII, ed. Mushira Eid, 69–90. Amsterdam and Philadelphia: J. Benjamins.

Moutaouakil, Ahmed. 1989. *Pragmatic functions in a functional grammar of Arabic*. Dordrecht: Foris.
——. 1993. *al-Waḏ̣īfa wa l-binya*. Casablanca: ʿOkāḏ.
Ouhalla, Jamal. 1991. *Functional categories and parametric variation*. London: Routledge.
——. 1997. "Remarks on focus in Standard Arabic". *Perspectives on Arabic linguistics*, X, ed. Mushira Eid and Robert R. Ratcliffe, 9–46. Amsterdam and Philadelphia: J. Benjamins.
——. 2002. "Negative sentences in Arabic". *Themes in Arabic and Hebrew syntax*, ed. Jamal Ouhalla and Ur Shlonsky, 299–320. Dordrecht: Kluwer Academic Publishers.
Shlonsky, Ur. 1997. *Clause structure and word order in Hebrew and Arabic: An essay in comparative Semitic syntax*. Oxford: Oxford University Press.

JAMAL OUHALLA (University College Dublin)

Negev Arabic

1. GENERAL

1.1 *Area and range*

Negev Arabic is spoken by some 150,000 Bedouin in the Negev (*án-Nagab* in Negev Arabic) desert, which occupies most of southern Israel. Today it is an isolated, peripheral minority dialect, not spoken and barely understood elsewhere within the contemporary political boundaries. Before 1948, however, the Negev belonged to a vast historical, economic, sociocultural, and linguistic continuum extending from the Ḥijāz to North Africa, along which nomadic and seminomadic tribes conducted raiding, trading, and smuggling relations with other nomadic tribes and with sedentary communities at the fringes of the desert. Arabian Bedouin culture in the peripheries was in constant contact with border communities, and this impact is clearly seen in Negev Arabic.

1.2 *Speakers*

About half of the contemporary population of the Negev Bedouin are descendants of the Tiyāha, Ṭarābīn, and ʿAzāzmih tribal confederations who migrated from Sinai in the 18th and 19th centuries (Bailey 1985). The rest are Bedouinized peasant families who migrated in waves throughout the 19th century, mostly from Egypt and the Gaza Strip. Fewer than a thousand are descendants of Palestinian villagers from the Hebron hills. Their dialect will be disregarded here, as will that of the Black Bedouin, descendants of Sudanese slaves.

Following the establishment of the State of Israel, the vast majority of the then seminomadic Bedouin population fled or were forced out. Most of the remaining tribes were relocated within a small area in the northern Negev. Today, about half the Negev Bedouin live in seven Bedouin towns. The remainder, reluctant to give up their disputed lands and move into the tribally heterogeneous towns, live in their tribal locations, 'the unrecognized villages'.

1.3 *Dialect type*

Negev Arabic comprises two major components: the original Bedouin Negev Arabic and the peasant or Fallāḥi Negev Arabic.

Together with → Sinai Arabic, Bedouin Negev Arabic constitutes the western subgroup of the → North West Arabian group (Palva 1991a), with maximal affinity to the neighboring dialects of the Judaean desert (Rosenhouse-Katz 1980) and the Ḥwēṭāt Najdi tribe (Palva 1986, but cf. de Jong 2000). In contrast, the northern Israeli Bedouin dialects belong to the North Arabian (→ Najdi) type. These two Israeli Bedouin 'islands' are separated by a block of sedentary → Palestinian dialects.

Since the Egyptian peasant migrants attached as hired laborers to the landowning tribes of the relatively fertile northern Negev, their dialects gradually accommodated, albeit only partially, to the socially prestigious Bedouin Negev Arabic of the majority. Mutual influence emerged between Bedouin and Fallāḥi-Bedouin groups whose interrelations were strong. As the Fallāḥi-Bedouin groups grew, though, their distinctive dialects developed separately, since Bedouin and Fallāḥi-Bedouin groups lived apart and did not intermarry, except for rare cases of Bedouin men taking Fallāḥi-Bedouin wives.

Today, some leveling is taking place, primarily in the towns and among the young and educated. The emerging leveled variety, influenced by the sedentary Palestinian koine, is drifting closer to Fallāḥi Negev Arabic. Finally, the dominant professional and business language in Israel, Hebrew, infiltrates the dialect of the young and educated bilingual generation, in some ways shielding it from further koineization.

Table 1. The consonants of Negev Arabic

	bilabial	labio-dental	interdental	alveolar	palatal	velar	pharyngeal	laryngeal
stops	b ḅ			t ṭ d		k g		ʾ
affricates					j			
fricatives		f f̣	ṯ ḏ ḏ̣			x ġ	ḥ ʿ	h
sibilants				s ṣ z	š ž			
liquids				l ḷ r ṛ				
nasals	m ṃ			n				
semivowel	w				y			

1.4 *State of research*

Blanc (1970) is a concise but amazingly accurate and encompassing description of Bedouin Negev Arabic. Other studies include Piamenta (1979, 1996), Borg (1996), and Henkin (1992–2002). An ongoing dialectal project initiated by the late Rafi Talmon will, it is hoped, make texts, a dictionary, a grammar, and sociolinguistic analyses available in the near future (Israel Science Foundation, grant no. 800/03).

Oral poetry from Sinai and the Negev is analyzed in a historical perspective in Bailey (1991).

2. LINGUISTIC DESCRIPTION

2.1 *Phonology*

2.1.1 Inventory of phonemes

2.1.1.1 Consonants (Table 1)

Interdentals are preserved, with some rare lexical exceptions in loans such as *dig*^{i}*n* 'beard', the elevated *zāt aliyyām* 'one day', and throughout the root *z-b-ṭ* 'to fix'. As elsewhere, the two historical velarized interdentals have merged; they are now /ḏ̣/, i.e. a velarized /ḏ/.

As typical of North West Arabian, there is no affrication of /k/ or /g/ (Old Arabic **q*).

/-áʾ/ with glottal stop is a characteristic pausal variant of /-á/, e.g. *lhaʾ#* 'to her', *mašáʾ#* 'he walked' [2.1.2]. Otherwise, /ʾ/ is usually elided in non-onset status, or changed to /w/ in new roots such as *w-k-d* (**ʾ-k-d*) 'to ascertain'; *w-l-f* (**ʾ-l-f*) 'to tame'; *w-n-s* (**ʾ-n-s*) 'to keep company'. The root *s-ʾ-l* 'to ask' is realized with a pharyngeal /ʿ/ by the elderly.

Secondary velarization of, for example, /l/, /m/, /b/, /r/, /f/ in lowback phonetic environments has created minimal pairs such as *xāḷ-ī* 'my maternal uncle' – *xāliy* 'empty'; *jāṛ-ī* 'my neighbor' – *jāriy* 'running'. Lexical velarization characterizes culturally significant names such as *Aḷḷāh*, *M(u)ḥaṃṃad* (in reference to the Prophet), and emotionally loaded items, such as *yuṃṃa*^{h} 'mum', *ʿaṣāk ṭayyib* 'hope you are well', *ʿaṣḷūj* 'young twig', *ixṣ!* 'pooh!' Similarly, labialization of /ṃ/ and /ḅ/ seems to be lexical in *aṃṃ*^{w}*ī* 'my mother' and *raḅḅ*^{w}*ī* 'my Lord'.

The affricate /j/ may be realized as a fricative [ʒ] in assimilation to /z/: *mijjawwiz židīd* 'newly married'; some speakers alternate [ʒ] with [dʒ] in phonetic distribution. /š/ dissimilates to /s/ in the vicinity of /j/ in Fallāḥi Negev Arabic *sájaṛ* 'trees', *sijāʿah* 'bravery'. Sonorant interchange occurs in *Ismāʿīn*, *Jubrīn*, *finjēl* 'cup', *ġayn* 'clouds', *balzīm* (< Hebrew *benzín*) 'gasoline', *ramzōn* (< Hebrew *ramzór*) 'traffic lights'. Metathesis occurs in the verbs *aʿṭa* ~ *aṭʿa* 'to give' and *aṭʿam* ~ *aʿṭam* 'to feed'.

2.1.1.2 Vowels

The three basic vocalic phonemes are high front /i/, high back /u/, low /a/, and their long counterparts /ī/, /ū/, /ā/, often shortened when unstressed. Notably, the systemic opposition of short vowels has been considerably reduced in unstressed open syllables, due to the elision of historical /i/ and /u/ (Blanc 1970:II3a), e.g. *jibāl* > *jbēl* 'mountains'. Moreover, historical /a/ in open unstressed syllables has assimilated to a following /u/ or, more often, to /i/, in a large-scale /a/ > /i/ shift, e.g. *kabīr* > *kibīr* 'big'. Today, however, this shift preceding an /a/ is partially reversed, so *jimál* 'camel' (Blanc 1970; Stewart 1990) is now predominantly *jamál* (Fallāḥi Negev Arabic *jámal*), *kitál* 'killed' is *katál* (Fallāḥi Negev Arabic *kátal*), etc.

A phonetically conditioned word-final *-á#* >

-íy# shift characterizes Bedouin Negev Arabic, so *ʿarjá* > *ʿarjíy* ‘lame [fem.]’, *ṭaršá* > *ṭaršíy* ‘deaf [fem.]’, *almāʾ* > *álmiy* ‘the water’; common to almost all Negev Arabic speakers are *hniy* ‘here’ (**hunā*), *kiḏíy* ‘so’ (**kaḏā*), *štiy* ‘winter’ (**šitāʾ*). This final → *ʾimāla*, as its name suggests, is annulled with suffixes, hence *ʿawjāk* ‘your crooked [fem.] one’, i.e. she-camel.

Internal *ʾimāla* (/ā/ > /ē/), generally considered nonphonemic in the Negev (Blanc 1970: III1), is today a major sociolinguistic differential, contrasting primarily Bedouin Negev Arabic and Fallāḥi Negev Arabic. It is typical of Bedouin Negev Arabic in patterns with a synchronic or historical front vowel, such as CāCiC (*gēʿid* ‘sitting’), *CiCāC (*rjēl* ‘men’), *yuCāCiC (*yḥērib* ‘he will fight’), *muCāCiC (*msēʿdih* ‘helper [fem.]’), unless blocked by emphatics or /x/ or /ġ/. The plural feminine morpheme also undergoes *ʾimāla* if the base fulfills the conditions: *kilmēt* ‘words’; *ṣāḥbēt* (**ṣāḥibāt*) ‘friends [pl. fem.]’ [2.1.3]. Minimal pairs include *jdēd* (**jidād*) ‘new [pl.]’ – *jdād* (**ajdād*) ‘forefathers’; *ʿēdiy* (**ʿādiy*) ‘attacking’ – *ʿādiy* (**ʿād+iy*) ‘regular’.

The new phonemes /ē/ and /ō/, resulting from either *ʾimāla* or monophthongization [2.1.1.3], are not stable (Blanc 1970:II4). Many speakers do not differentiate at all, or not consistently, and pronounce minimal pairs such as *dēn* ‘debt’ – *dīn* ‘religion’ or *dōṛ* ‘turn’ – *dūṛ* ‘houses’ identically.

2.1.1.3 Diphthongs

None of the diphthongs /ay/, /aw/, /iy/, /uw/ (Blanc 1970:II4) are stable. In Bedouin Negev Arabic, especially of the older generation, /ay/ and /aw/ are preserved, primarily in backing environments and in certain morphological patterns such as nominal CaCC (*ayš* ‘what’), verbal base-final (*ḥaṭṭayna* ‘we put’, *sawwaytiy* ‘you [sg. fem.] did’), and the plural ending (*gāḷaw* ‘they said’), as well as root-initial cases (*mawlūd* ~ *maylūd* ‘born’, *awsaʿ* ‘wider’, *awfa* ‘fulfilled’). However, in all except aCCaC and word-final position (*gāḷaw*), these diphthongs alternate with the new monophthongs /ē/ and /ō/. Moreover, due to the nondistinction of /ō ~ ū/ and /ē ~ ī/, multiple alternation is common; *jābawh* ~ *jābōh* ~ *jābūh* ‘they brought him’; *ʿalayk* ~ *ʿalēk* ~ *ʿalīk* ‘on you’. While certain lexical items, such as *lēl* ‘night’, and *ēš* ‘what’, seem to exclude /ī/, most items fluctuate.

The homorganic diphthongs, /iy/ and /uw/, tend to monophthongize and shorten in fast speech, so that *biynām* ‘he sleeps’ is actually pronounced *bīnām* ~ *binām*, and *yimšuw* ‘they walk’ is *yimšū* ~ *yimšu*. In this entry the full forms are retained.

2.1.1.4 Syllable

Syllables may be short (Cv), long (CvC, Cv̄), or extralong (Cv̄C *bāb* ‘door’, CvCC *bint* ‘girl’). Trochaic structures, such as *bābana* ‘our door’, are characteristic of poetic language, the quantitative prosody of which excludes extralong syllables in context.

2.1.1.5 Consonant clusters

Initial obstruent-sonorant and sonorant-sonorant clusters can occur (*klēb* ‘dogs’, *ḥmāṛ* ‘donkey’, *wlād* ‘children’). Sonorant-obstruent clusters need a helping vowel (*^{i}nfakkir* ‘we think’, *^{i}mʿaggid* ‘going’).

Word-final C_2-sonorant clusters are usually dissolved (*gaḅ^{a}l* ~ *gaḅ^{u}l* ~ *gaḅ^{i}l* ‘before’, *ṣuf^{u}r* ‘yellow [pl.]’, *ṛam^{i}l* ‘sand’, *taṃ^{i}r* ‘dates’, *dil^{i}m* ‘dounam’, *sam^{i}n* ‘sour cream’), as are voiced stop clusters (*kib^{i}d* ‘liver’, *ʿab^{i}d* ‘slave’, *ʿug^{i}b* ~ *ʿug^{i}ḅ* ~ *ʿug^{u}ḅ* ‘after’).

Medial CCC clusters containing *sk* are frequent (*usktuw* ‘shut up [pl. masc.]!’; *imskuw* ‘catch [pl. masc.]!’; *tusknuw* ‘you [pl. masc.] live’). Liquid-obstruent C_1C_2 help retain a CCC cluster (*urbṭiy* ‘tie [sg. fem.]!’, *yirkbih* ‘he lets him ride’); otherwise, syllable reshuffling (Blanc 1970:II3a) is predominant (*tik^{i}tbuw* ‘you [pl. masc.] write’, *nuḏ̣^{u}rbak* ‘we hit you’; *tum^{u}rguw* ‘you [pl. masc.] pass’, *tif^{i}gduw* ‘you [pl. masc.] lose’).

2.1.1.6 Stress

Stress (Blanc 1970:II5) falls on the last v̄C or vCC of the word, e.g. *ánḏabaḥ* ‘he was slaughtered’. In the absence of extralong syllables, the second syllable is stressed in Bedouin Negev Arabic (*walád* ‘boy’, *bagáṛah* ‘cow’ or ‘his cows’, *akálah* ‘he ate it’). Four-syllable words fluctuate: *madrasátak* ‘your school’, *akálatih* ~ *akalátih* ‘she ate it’. The article and other clitics are considered part of the content word with respect to stress, hence *álwalad* ‘the boy’, *álbil* ‘the camels’, *állḥiy* ‘the beards’, *mín-taḥat* ‘from beneath’, and sometimes *gāḷát-lak* ‘she told you’.

Fallāḥi Negev Arabic speakers tend to stress the initial syllable and disregard the article: *ákalah* ‘he ate it’, *ákalatih* ‘she ate it’, *(al)wálad* ‘(the) boy’.

Helping (anaptyctic) vowels do not change the underlying base, hence *áṣ*^{i}*l* ‘origin’ ~ *aṣíl* ‘I arrive’; due to such minimal pairs, stress is superficially distinctive. A notable exception, however, is *gaháwah* vowels [2.1.2].

The stress on the 1st person singular suffix *-(n)ī*, e.g. *bintī*$^{(h)}$ ‘my daughter’, *katalawnī*$^{(h)}$ ‘they killed me’, may be explained by the optional aspiration, especially in pause, which causes an extralong syllable. The stress of the *ḥamṛá* pattern seems historically determined (**ḥamṛāʾu*).

2.1.2 Phonotactics

Assimilation patterns are those known in other dialects: *mijjawwiz* ‘married’, *axatt* ‘I took’ (Fallāḥi Negev Arabic also *axaḏ*^{i}*t*), ^{i}*nná* (Fallāḥi Negev Arabic *ínna*) ‘to us’, *yissāʿalaw* ‘they ask around’, (*arʿha* ~) *arḥḥa* ‘there she is’, (*simiʿt* ~) *simiḥt* ‘I heard’.

The → *gaháwah* syndrome shifts a historical aXC structure (X being a back spirant) to aXáC (*gahwah* > *gaháwah* ‘coffee’). In Negev Arabic the following patterns undergo this shift (Blanc 1970:III2), wholly or partially:

CaCC(ah):	*šahár̩* ‘month’, *taḥát* ‘under’, *ahálak* ‘your family’, *baġúḷah* ‘mule [fem.]’
aCCaC:	*aḥámaṛ* ‘red’, *aʿáṛaj* ‘crooked’, *aġána* ‘richer’
maCCaC:	*maḥáṛam* ‘women’s quarters’, *maʿánad* ‘tent partition’
yaCCaC:	*yaʿárf* ‘he knows’, *yaġáziy* ‘he raids’

Notwithstanding the general rule that anaptyctic vowels are not stressed [2.1.1.6], the stabilized (hence unraised) *gaháwah* vowels are stressed in Bedouin Negev Arabic according to Bedouin stress rules; Fallāḥi Negev Arabic has hybrid variants, with vowel insertion but regular stress on the first syllable (*šáhaṛ; áḥamaṛ, máḥaṛam, yáġaziy*).

Thus, at least three variants coexist in Negev Arabic for certain patterns: koineized *maḥṛam*, Fallāḥi Negev Arabic *máḥaṛam*, Bedouin Negev Arabic *maḥáṛam*; *mḥáṛam*, however, is not heard in the Negev.

Of course, the presence of an extralong syllable in the word will prevent the stress from moving to the new vowel: *zahagān* ‘fed up’, *saʿadān* ‘monkey’, *maxalūg* ‘creature’, *ṣahabíy* ‘reddish-brown [fem.]’ (with several variants: *ṣahabá, ṣahába, ṣáhba, ṣahbá*).

The syndrome does not cross morpheme boundaries, hence *fataḥ-t* ‘I opened’, *maʿ-na* ‘with us’; Form IV *a-ġna* ‘he enriched (but *aġána* ‘richer’).

As in other dialects → velarization is a suprasegmental phenomenon. It spreads from a locus (primary or secondary) until blocked by a fronting element. In this entry just one velar per word is marked, because the spread of velarization is predictable.

Pausal phenomena are of several kinds, some of which are general, others particular to the dialect. Among the most general is devoicing of /b/, /d/, /g/: *zihig* > *zihík#* ‘he got bored’ (Blanc 1970:II2). Characteristic of Bedouin Negev Arabic is pausal glottalization of /-á/ > /-aʾ/ [2.1.1.1; 2.2.1.3].

2.1.3 Morphophonology

The ‘vanished vowels’ /u/, /i/ [2.1.1.2] leave traces with respect to secondary velarization and *ʾimāla*: historical /u/ is responsible for velarization in cases such as *tṛāb* ‘earth’ (**turāb*), *ṛkab* ‘knees’ (**rukab*); historical /i/ and /u/ explain *ʾimāla* in *blēd* ‘land’ (**bilād*) and *mēkil* (**ʾākil*) ‘having eaten’ as against its absence in *yākil* ‘he eats’ (**yaʾkul*).

2.2 *Morphology*

2.2.1 Pronouns (Blanc 1970:IV1)

2.2.1.1 Free subject pronouns

	singular regular	long	plural regular	long
3rd masc.	*hū*$^{(h)}$		*hum*	*huṃṃa*$^{(h)}$
3rd fem.	*hī*$^{(h)}$	*hiyyi*$^{(h)}$	*hin*	*hinni*$^{(h)}$
2nd masc.	*int*	*inti*$^{(h)}$	*intuw*	
2nd fem.	*intiy*		*intin*	
1st	*aná*	*anā*$^{(h)}$	*aḥna*	

Dialectal variants include Fallāḥi Negev Arabic *ána* and strengthened *anāhiy* limited to some Bedouin Negev Arabic groups.

2.2.1.2 Bound pronouns

Bound pronouns (possessive, object, and prepositional suffixes) are shown in the following inflections of ‘his sister and his brother love him’.

3rd person singular masculine *-ih* alternates with allomorphic *-ah* (*xāḷah* ‘his uncle’; Blanc 1970:IV2d); 3rd person feminine has a widespread dialectal variant *-hiy*. Strengthened suffixes (*lāgāhuṃṃa* ‘he met them’, *giddēmhuṃṃa* ‘in front of them’) characterize Fallāḥi Negev Arabic.

	singular	plural
3rd masc.	*uxtih w axūh yḥibbūh*	*uxthum w axū hum yḥibbūhum*
3rd fem.	*uxtha w axūha yḥibbūha*	*uxthin w axūhin yḥibbūhin*
2nd masc.	*uxtak w axūk yḥibbūk*	*uxtkuw w axūkuw yḥibbūkuw*
2nd fem.	*uxtkiy w axūkiy yḥibbūkiy*	*uxtkin w axūkin yḥibbūkin*
1st	*uxtī*$^{(h)}$ *w axūy yḥibbūnī*$^{(h)}$	*uxtna w axūna yḥibbūna*

2.2.1.3 Indirect (dative) object suffixes

	singular	long	plural	long
3rd masc.	*-lih*		*-lhum*	*-lhuṃṃa*$^{(h)}$
3rd fem.	*-lha~ -lhiy*	*-lhiyyi*$^{(h)}$	*-lhin*	*-lhinni*$^{(h)}$
2nd masc.	*-lak*		*-lkuw*	*-lkuwwa*$^{(h)}$
2nd fem.	*-lkiy*	*-lkiyyi*$^{(h)}$	*-lkin*	*-lkinni*$^{(h)}$
1st	*-lay*	*-layyī*$^{(h)}$	*-lna*	

The long forms, optionally aspirated, are often emphatic. The 3rd person masculine has an allomorph *lah* [cf. 2.2.1.2]. Pausal glottalization is frequent in *lhaʾ*, *lnaʾ* [2.1.1.1; 2.1.2]. A variant base *lī-* (*līkiy, līhiy, līna, līkuw*…) characterizes the Taṛābīn [2.2.3].

2.2.1.4 Demonstratives (Blanc 1970:VII2b)

proximate	sg. masc.	*hāḏa ~ hāḏ̣a*
	sg. fem.	*hēḏī*
proximate plural		
Bedouin Negev Arabic		*hawḏ̣al ~ hōḏ̣al(la)*
Fallāḥi Negev Arabic		*haḏ̣ōl(la) ~ haḏ̣ōla ~ haḏ̣alla*
distant	sg. masc.	*haḏāk(a) ~ haḏ̣āk(a)*
	sg. fem.	*hiḏīk(a) ~ hiḏīki*
distant plural		
Bedouin Negev Arabic		*hōḏ̣allāk(a) ~ hōḏ̣ullāk(a) ~ huḏ̣ullāk(a)*
Fallāḥi Negev Arabic		*haḏ̣ullāk(a)*

All the strengthened variants ending in a vowel can be optionally aspirated. Unvelarized singular masculine forms characterize some Bedouin Negev families, especially the elderly; plural forms are always velarized. Rare feminine variants *hāḏī* and *haḏīk* are marked as either elevated or Fallāḥi Negev Arabic. The plural distant forms have several vocalic variants, e.g. *hōḏ̣allāk*.

The long distant forms in adjectival use always follow the noun; the singular construct is characteristic of temporal phrases such as *hiḏīk(t) alḥīn* ‘at that time’ [2.2.2] but also *hiḏīkt algiṣīdih* ‘that poem’.

2.2.1.5 Presentatives

Negev Arabic is rich in → presentatives (Blanc 1970:VII2c), as is typical of an orally transmitted narrative tradition. Some primarily narrative particles are (*ġāṛ*) *win, wlin, illa w(in), willa, walla*, e.g. *willa w hū mēšiy biyfakkir win hálbil* ‘and there he was walking; he looked up and there were these camels.’ *illa lamminnih ja win hazzalámah* ‘no sooner had he arrived and there was this man’. The first two may be inflected: *ʿugḅ ⁱšwayyih winha mlawwḏih ʿalīhum hássaḥan* ‘after a short while there was this plate she had slipped in for them’. Unique to Bedouin Negev Arabic is the existential-presentative *wmāṛ* [2.2.2, 2.2.3] which may open a story: *wmāṛ hassammāk* ‘there was once this fisherman’.

Presentative particles include proximate *hay*, distant (often modal) *hawēn*, and *arʿ*, all of which may inflect nominally: *lōmin māl winhum hayhum allī maṛágaw ʿalēh albāriḥ winhum hayhum mkattalīn hum aṯṯnīn* ‘when he turned he found those who had passed by the day before, there they were, both killed’; *arʿ wāḥad minhum*, *arʿah ʿindī f-addāṛ* ‘there’s one of them, there he is at my house’. *arʿ*, originally an imperative ‘see!’ inflects as such (*arʿī*, *arʿuw*): *arʿīhum maʿ aṭṭirīg* ‘there they are along the road’.

The inflecting particle *iṯr*, primarily a modal-rhetorical evidential, has an additional presentative component: *algōṃ iṯṛātha jēyyih* ‘[we found that] the enemies had apparently come up’.

2.2.1.6 The relative pronoun

The relative pronoun *allī*, emphatically *hallī*, may shorten to the form of the article: *anā axaḏt al-amskih*… ‘I married him whom I grasp…’ It can function as a content complementizer: *alḥamd illēh allī ṛaddēna alḥalāl* ‘thank God we got the livestock back’.

2.2.1.7 Interrogatives

Interrogative pronouns and particles (Blanc 1970:VII2a) include *ayš* ~ *ēš*, *wēš* ~ *wiš* ‘what’; as a sentential element, *wiššū* (**wiš hū*) is often rhetorical: *wiššū alṃayy? hāḏ̣a wala išiy!* ‘what of it, just water? That’s nothing!’: Other particles include *z-īš* ‘in what way’; *lēš* ~ *līš* ~ *lēh* ‘why’; *wēn* ‘where’; *matá* ‘when’; *min* ‘who’; *yēt* ‘which’; *kēf* ~ *kīf* ‘how’; *ʿalāmak* ‘what’s up with you’. Sentence-final position (*hāḏ̣a min?* ‘who is it?’) is pragmatically marked.

Rhetorical yes-no questions may be introduced by *(ay) hū*: *ay hū albadáwiy ʿumṛah byíndimij l-algaryih almitḥaḏ̣ḏ̣rih?* ‘do you honestly believe a Bedouin could ever fit in a sedentary village?’.

2.2.2 Adverbs (Blanc 1970:VII3)

Temporal adverbs include *(h)alḥīn(iy)* ‘now’; *hiḏīk(t) alḥīn* ~ *alwagt* ~ *annahāṛ*; *ḥīnitha* ~ *sāʿitha* ‘then’; *yōm(it)ha* ‘that day’; *assēʿ* ~ *issēʿ* ‘still’, ‘not yet’.

The adverb *bēkir* may render ‘early in the morning’, as in *bēkir aljimʿih aljēyyih* ‘early the following Friday’, but is more frequent as deictic ‘tomorrow’, which, however, is predominantly *bukṛah* for the younger generation (no → *bukaṛa*-syndrome in the Negev). The narrative equivalents are the prepositional phrases *min bēkir*, *min bukṛah* ‘the following day’. A common narrative sequencer is *tēniy min yōm* ~ *nahāṛ* ‘another day’.

Sequencers include *ʿug*$^{(u)}$*ḅ* ~ *baʿ*$^{(a)}$*d* ^{i}*šwayyih* ‘after a while’ and the narrative particle *wmāṛ* typical of elderly Bedouin Negev Arabic speakers [2.2.2], *wmāṛ jiddī yaʿáṭs* ‘and then my grandfather sneezed’.

‘Yesterday’ or ‘the day before’ is *albēriḥ*; ‘the day before that’ or, more generally, ‘the other day’ is *awwal-ams*. As an adverb, *awwal* means ‘previously, formerly, in the old days’. For specific temporal designation, possessives like *yōm(it)ha* ‘that day’, *sanatha* ‘that year’ are common, while duration may be construed in constructs like *sanat zamān* ‘for a year’.

Local adverbs include the proximate deictics *hniy* ~ *hniyya*$^{(h)}$ ~ *hniyyān(iy)* ~ *hniyyāntiy*; *f-hāḏ̣a* ‘here’; and the distant equivalents *hnuh*, *ġād*, *f(ī) haḏ̣āk* ‘there’.

Manner adverbs include *kiḏíy* ~ *kiḏíyya*$^{(h)}$ ~ *kiḏiyyān(iy)* ~ *kiḏiyyāntiy* ‘so’; the short form often combines with the stressed particle *zay* ‘like’ (*záy-kiḏiy* ~ *zé-kiḏiy*).

2.2.3 Particles

The article *al-*, assimilating to sun letters and considered part of the lexeme for stress purposes [2.1.1.6], has a stable vowel which, in fast speech, may delete preceding vowels (Blanc 1970:VI3): *z-alyōm* ‘like today’, *f-álxaḷa* ‘in the wilderness’.

There are two analytic genitive markers: The marker *šuġ*^{u}*ḷ* (*šuġḷit-*, *šuġlīn*…) is specific to Bedouin Negev Arabic, while the sedentary *tabaʿ* is in general use. They are also used to denote characteristics, e.g. *māhum šuġlīn ḥiṣīdih* ‘they are not of the reaping type [i.e. ‘they are not proficient farmers’]’.

As elsewhere, the preposition *fī* functions as an existential, in which case it is realized in full and is often aspirated as *fī*^{h}, in contrast with prepositional use, where it tends to shorten, often leaving the consonant alone, as in *f-álxaḷa*, *f-hāḏ̣a* above.

Dialectal variation characterizes the inflection of some basic prepositions. The Taṛābīn, for example, use long variants of the extrashort bases *l-b-* (*lī-bī-*) and shortened bases of the longer *fī*, *ʿala (f-, ʿal-)*.

Taṛābīn	non-Taṛābīn
bīna, bīhiy	*bna, bha*
līkiy, līhum	*lkiy, lhum*
fak, fah	*fīk, fīh*
ʿálak	*ʿalayk ~ ʿalēk ~ ʿalīk ~ ʿilēk ~ ʿilīk*
ʿálah	*ʿalayh ~ ʿalēh ~ ʿalīh ~ ʿilēh ~ ʿilīh*

Major conjunctions include the temporal *ḥīn(t)-ma*, *yōmit-ma*, and *yōm(in)* or *lōm(in)* 'when', also inflected *yōminnak*, *lōminnak*, etc., as well as *laṃṃa(n)*, *lammin(nak)*, *lamma*, etc. for 'when, until'. As is often the case, temporality shades off into conditionality and causality, as in *yōminnak ṛaddēt ṛabʿak ʿinnih* 'since you warded your men off him [we pardon you too]'. *blōminnak* is unequivocally causal 'since you'.

The basic negator (Blanc 1970:VII1) is *mā*: *mā hāḏa hū* 'he's not the one'; *mā wāḥid ~ mā ḥad* 'nobody'; *māšiy* 'nothing', as in *wkaʾin māšiy sēyir* 'as if nothing has happened', often combined to *māš* as in *māš ṛawš* 'absolutely no bustle'; *māš jiṯṯah* lit. 'no body', 'featherweight'. Emphatic negators are several: *mā bīrīdna gēyim ~ xalāṣ ~ min maṛṛah ~ b-almaṛṛah* 'he doesn't like us at all'. Nominal negation is inflected:

	singular	plural
3rd masc.	*māhū ~ mūhū ~ mū*$^{(h)}$	*māhum, mūhum*
3rd fem.	*māhī ~ mīhī ~ mī*$^{(h)}$	*māhin, mīhin*
2nd masc.	*mant ~ manti*$^{(h)}$ *~ mint ~ minti*$^{(h)}$	*mantuw, mintuw*
2nd fem.	*mantiy ~ mintiy*	*mantin, mintin*
1st	*mānī*	*maḥna*

A negated future or participle of a passive verb encodes nonpotentiality, e.g. *mā byinḏāg* or *mū minḏāg* 'it is untastable', *mā byinfātin* or *mīhin minfātāt* 'they are unleavable.'

The sedentary *muš ~ miš* for nominal negation is relatively frequent, existential *ma-fīš* 'there is no' considerably less so. Otherwise, the *-š* suffix is highly marked as Fallāḥi Negev Arabic of Palestinian rather than Egyptian origin.

lā serves for negation of the imperative, for sentence negation, where it is more emphatically rendered *lah lah lah!* or *lallih!*, and in continued negation, where it may inflect as *walānī*, *walant*.

Particles discussed elsewhere include the relative particle [2.2.1.6], interrogatives [2.2.1.7], and vocatives [2.2.4].

2.2.4 Noun

The → pseudodual ending in Negev Arabic is *-ān*, distinct from the true dual *-ayn*; but since the former often undergoes *ʾimāla* in Bedouin Negev Arabic, while the latter undergoes monophthongization in Fallāḥi Negev Arabic, they often merge to *-ēn*. The pseudodual category comprises the core lexical items such as *īdān ~ īdēn* 'hands', *rijlān* 'legs' (but *ʿyūn* 'eyes', *aḏān* 'ears'). Additional items include *ḏuṛʿān* 'arms', *ḏ̣ubṭān* 'armpits', *jinḥān* 'wings'. These inflect as expected but may also retain the dual ending in suffixing: *īdāh ~ īdēh ~ īdānih ~ īdēnih* 'his hands' (also *adēh ~ adēnih* in some Bedouin Negev Arabic varieties).

Diminutives are frequent in nouns (*gīmih* from *gōṃ* 'enemy', *lgēmih* from *lugṃah* 'morsel'), personal names (*Xḷayyil* from *Xalīl*; *Ṛbayʿēt* are members of the *Abu Ṛabīʿah* clan), horse and camel names (*Zṛaygān*, fem. *Zrēgah* from *azṛag* 'dark') and adjectives (*azīrig*, fem. *zṛaygiy*, from *azrag; uʿáymiy*, fem. ^{i}*ʿmayíy* from *uʿáma*, fem. *ʿamyíy* 'blind'). Lexemes that have superseded their original forms include *ṣġayyir* 'small', *gḷayyil* 'little', *gṛayyib* 'near', *gṣayyir* 'short', *ḥṛayyim* 'women', and *ḏ̣ʿayfīn* (alongside *ḏ̣ʿūf*) 'children'.

A common pattern CaCūC for a characteristic quality includes *kuḏūb* 'liar', *luʿūb* 'one who likes to play', *ḏ̣uḥūk* 'one who likes to laugh', *šuṛūd* 'one who is always running away'.

Common vocatives, usually introduced by the particle *yā ~ ya*, include *yuṃṃa*^{h} 'mom' and *ībā*^{h} 'dad'; the vocatives for 'brother' and 'sister' are diminutive *ya xayy*, *ya xayyih*, respectively. Reversed kinship terms are most frequent in vocatives, hence *ya-xtak* or *yā xayytak*, lit. 'o your sister' (sister addressing brother), *yuṃṃa*^{h} (mother addressing son or daughter), *ya jiddkiy* (grandfather addressing granddaughter), etc.

2.2.5 Verb (Blanc 1970:V; all paradigms here represent Bedouin Negev Arabic unless otherwise stated)

2.2.5.1 Verbal forms

2.2.5.1.1 Form I

The CaCáC and CiCíC patterns (Fallāḥi Negev Arabic CáCaC and CíCiC respectively) display the semantic distribution typical of Semitic languages. CuCúC (Fallāḥi Negev Arabic CúCuC) is rare, e.g. *kuḅúr* 'grew up' (but *kabrit*, Fallāḥī Negev Arabic *kuḅrut*).

2.2.5.1.2 Derived Forms

Form IV	*aˤṭa ~ aṭˤa* 'to give'	
	Bedouin Negev Arabic	Fallāḥi Negev Arabic
imperfect	*yiˤṭiy ~ yiṭˤiy*	*yaˤṭiy*

Form V	*tajawwaz* 'to marry'	
	Bedouin Negev Arabic	Fallāḥi Negev Arabic
perfect	*tajawwaz*	*tjawwaz*
imperfect	*ytajawwaz*	*yitjawwaz*
2nd masc.	*ttajawwaz ~ tajawwa z*	*titjawwaz*

2.2.5.2 Inflection of Form I perfect + imperfect

2.2.5.2.1 Imperfect

Paradigm of the verbs 'to kill' (representing /u/ and /i/ patterns), and 'to understand' (/a/ pattern).

	sg.	pl.
3rd masc.	*yuktul*	*yuk*^{u}*tluw*
3rd fem.	*tuktul*	*yuk*^{u}*tlin*
2nd masc.	*tuktul*	*tuk*^{u}*tluw*
2nd fem.	*tuk*^{u}*tliy*	*tuk*^{u}*tlin*
1st	*aktul*	*nuktul*

	sg.	pl.
3rd masc.	*yafham*	*yafhamaw*
3rd fem.	*tafham*	*yafhaman*
2nd masc.	*tafham*	*tafhamaw*
2nd fem.	*tafhamay*	*tafhaman*
1st	*afham*	*nafham*

The indicative *b*- prefix is well established in Negev Arabic. Cohortative *xa*- (**xall*) 'let's' is rare; *d*- (**widdī*) 'I want to' is common for volition and futurity.

2.2.5.2.2 Perfect

	/a/ base	
	sg.	pl.
3rd masc.	*katál*	*katálaw*
3rd fem.	*katálat*	*katálan*
2nd masc.	*katalt*	*kataltuw*
2nd fem.	*kataltiy*	*kataltin*
1st	*katalt*	*katalna*
	/i/ base	
	sg.	pl.
3rd masc.	*fihím*	***fahmuw***
3rd fem.	***fahmit***	***fahmin***
2nd masc.	*fihimt*	*fihimtuw*
2nd fem.	*fihimtiy*	*fihimtin*
1st	*fihimt*	*fihimna*

Fallāḥi Negev Arabic is koineized both in stress (*kátal*, *fíhim*) and in the paradigm-leveling shift of the bolded items to *fihmit*, *fihmuw*, *fihmin*, in analogy to the other forms, whereas Bedouin Negev Arabic preserves the original base vowel (**fahima*). The 3rd person masculine form also preserves this form when cliticizing closes the syllable (*fahm-al-harj* 'he understood the talk'; Fallāḥi Negev Arabic *fihm-al-harj*).

2.2.6 Weak verbs

2.2.6.1 Iʾ and I guttural

The Iʾ verb 'to take' inflects *(a)xaḏ, yāxiḏ, xuḏ, xḏiy*, etc.

In the imperfect of I gutturals, the full *gaháwah*-syndrome (Blanc 1970:III2c) characterizes Bedouin Negev Arabic, while hybrid forms retaining the original stress characterize Fallāḥi Negev Arabic [2.1.2]. Both, however, are being replaced by koineized forms in the speech of the young generationm (Table 2).

Table 2. *gaháwah* syndrome in Negev Arabic

	Negev Bedouin	Fallāḥi Negev	koineized
'he reaps'	*yaḥáṣ(i)d*	*yáḥaṣid*	*yuḥṣud*
'he milks'	*yaḥál(i)b*	*yáḥalib*	*yiḥlib*
'he raids'	*yaġáza*	*yáġaza*	*yaġza*
'he knows'	*yaˤár(i)f*	*yáˤarif ~ yáˤaṛaf*	*yáˤṛaf yiˤṛaf*

2.2.6.2 I*w*

Bedouin Negev Arabic is characterized as conservative by deletion of the root-initial semivowel in the imperfect base (*yiṣíl* 'he arrives', *yigfuw* 'they stand'), while Fallāḥi Negev Arabic, in analogy to the sound verbs, has a monophthong in that position: (*yawṣal* >) *yōṣal*, *ōgaf* 'stop!'

2.2.6.3 II*w*

Some singular masculine imperatives have short alternants (*gūm* ~ *gum* 'get up!' and *gūl* ~ *gul* 'say!')

2.2.6.4 III*y*

Short imperatives, typical of elderly speakers, include Form I *irⁱm* 'throw!', Form II *lagg* 'go!', Form IV *iʿṭnī* 'give me!'; more rare are short imperfects (*tsaww gahávatak* 'you'll make your coffee').

2.2.6.5 Irregular verbs

Most prominent is *ja – yjiy*, *yjuw – taʿāl* 'come!'; the equivalent Fallāḥi Negev Arabic form *aja* occurs in the perfect, but the imperfect *yījiy* is extremely limited.

2.3 *Syntax*

2.3.1 Noun phrase

2.3.1.1 Construct state

The typical analytic genitive markers are listed above [2.2.3]. Certain fixed time expressions may occur in the construct state (*s(i)būʿ aljēy* 'next week').

A typical narrative pattern incorporates verbal noun constructs in presentatives [2.2.1.5] to denote vivid events: *win margit janāzih* 'then, a funeral procession passed by'; *win ṭallit hálbil w álġanaṃ* 'then suddenly, these camels and small livestock showed up'; *win jayyitna līna ʿijūz min kiḏiy* 'and suddenly, an old woman came up to us [out of nowhere]'.

2.3.1.2 Nunated nominals

Nunated nominals, typical of poetry, include noun/adjective phrases such as *bkāṛan mišēwīḥ* 'obedient camels', and other nominal phrases such as *badwin w fallāḥ* 'Bedouin and peasant'.

2.3.2 Verbal phrase

Punctual actions in the imperfect can be made durative by *fī* preceding the object: *hū byalbas fī hdūmih* 'he is getting dressed'.

2.3.3 Verbal aspect

2.3.3.1 Verbal prefixes and auxiliaries

Some verbal prefixes are listed above [2.2.5.2.1]; auxiliaries such as progressive *gēʿid* (lit. 'sitting') and future intent *ṛāyiḥ* (lit. 'going') are rare, while nominally inflected *widd-* (lit. 'to want') is the general future marker.

2.3.3.2 Active participle

In certain motion and emotion verbs the active participle is progressive; in all the rest it is resultative, as in many other dialects, and often functions modally as an evidential (Henkin 1992), e.g. *bētkuw…sēyrih ʿilēh fēlih. uṭāxxīn ʿalēhum ušērdīn ukiḏiyyāniy* 'your home…there seems to have been an attack on it. And they were apparently shot at and ran off and so on' (a host reporting rumors to his guest about the latter's home, from which he had been long absent).

2.3.3.3 Narrative imperative

The narrative imperative (Palva 1977, 1984) is an integral characteristic of Negev Arabic narrative style, especially frequent in men's stories (Henkin 1994). In a dramatic account of a warrior committing suicide on his horse: *mitrij maʿ halmutrāj w ugruṭ arrumḥ! yōm garaṭ arrumḥ w inthat alḥṣān ʿalēh, w siffih yā arrumḥ* 'he went down the slope and threw [lit. 'throw!'] the spear. When he threw the spear he drove his horse onto it and the spear sliced him [lit. 'and you, Bring! the horse onto it and Slice him, O spear!'].

2.3.3.4 Other narrative tenses

Other narrative tenses typical of men's narratives include motion verb complexes, often with the deictic dative suffix, namely *jāk/yjīk* + participle (Palva 1991b; Henkin 1996, 2000a, 2002), as in *yjīk migfiy* 'and off he went [lit. 'he comes to you going away']' and the sequencer *wmāṛ* + imperfect [2.2.2], as in *wmāṛ yjībinha wmāṛ yaṛkab ʿalēha wmāṛ ywajjih waṛa* 'and then they bring her [the mare] and he mounts her and turns around'. The more general auxiliary *gāṃ* (lit. 'got up') occurs in both men's and women's stories.

kān as a narrative auxiliary is rare: *kān yāxiḏ f-ṛās alⁱḥsān, wkān ykuḏḏ ʿa-lġalyūn, wkān ywinn* 'he grasped the horse's head, and sucked at his pipe and sighed'.

2.3.5 Existential sentences

The past marker *kān* is infrequent in existential structures at story openings, as narrative backgrounds are characteristically portrayed in the present (Henkin 1993), with simply the existential particle *fīʰ* [2.2.3] or *wmāṛ* [2.2.1.5].

2.3.6 Conditionals (Blanc 1970:VII4b; Henkin 2000b)

Potential conditions are introduced by *in*, *lin* + perfect, or *iḏa*, also *(i)la* and rarely *lan* in poetry. The prototypically counterfactual *law* may be simply hypothetical, as in *anā law agūl l-alwēḥid* 'suppose I say to one of them', or concessive, as in *law kān alkalb mā byanfaᶜ* 'even if the dog is no good'. As this last example shows, these particles are often followed by *kān*, fossilized or inflected nominally, as in *iḏa kānak ṣādig* 'if you are right', which can also act as an independent conditional: *mā bansāk kān amūt* 'I won't forget you, even if I die'. It is also frequent in the apodosis: *law ḏallayt mēšiy kān katalawnī* 'if I had gone on, they would have killed me'.

BIBLIOGRAPHICAL REFERENCES

Bailey, Clinton. 1985. "Dating the arrival of the Bedouin tribes in Sinai and the Negev". *Journal of the Economic and Social History of the Orient* 28:1.20–49.

——. 1991. *Bedouin poetry from Sinai and the Negev*. Oxford: Clarendon Press.

Blanc, Haim. 1970. "The Arabic dialect of the Negev Bedouins". *Proceedings of the Israel Academy of Sciences and Humanities* 4:7.112–150.

Borg, Alexander. 1996. "Linguistic and ethnographic observations on the color categories of the Negev Bedouin". *The language of color in the Mediterranean*, ed. Alexander Borg, 121–147. Stockholm: Almqvist and Wiksell.

Henkin, Roni. 1992. "The three faces of the Arabic participle in Negev Bedouin dialects: Continuous, resultative, and evidential". *Bulletin of the School of Oriental and African Studies* 55:3.433–444.

——. 1993. "'To be' or not in the Sinai Bedouin dialect of the Aḥaywāt". *Mediterranean Language Review* 6–7.173–197.

——. 1994. "The narrative imperative in colloquial Arabic and in Russian". *Journal of Semitic Studies* 39:2.245–283.

——. 1996. "Negev Bedouin vs. sedentary Palestinian narrative styles". *Israel Oriental Studies* 16.169–191.

——. 1998. "Narrative styles of Palestinian Bedouin adults and children". *Pragmatics* 8:1.47–78.

——. 2000a. "Narrative styles of Negev Bedouin men and women". *Oriente Moderno* 19 (80):1.59–81.

——. 2000b. "Pragmatics of conditionality in Negev Bedouin". *Proceedings of the Third International Conference of AIDA, Malta, March 29–April 2, 1998*, ed. Manwel Mifsud, 9–14. Malta: Salesian Press.

——. 2001. "Women's oral narrative styles in Palestinian Arabic: Bedouin and rural". *Proceedings of an International Conference on Middle Eastern Popular Culture Oxford August 17–21, 2000*, ed. Clive Holes, 56–65. Oxford.

——. 2002. "Tense-switching in narratives of Negev Bedouin men and women". *Aspects of the dialects of Arabic today: Proceedings of the Fourth International Conference of AIDA, Marrakesh, April 1–4, 2000*, ed. Abderrahim Youssi a.o., 288–298. Rabat: AMAPATRIL.

Jong, Rudolf E. de. 2000. *A grammar of the Bedouin dialects of the Northern Sinai littoral: Bridging the linguistic gap between the eastern and western Arab world*. Leiden: E.J. Brill.

Palva, Heikki. 1976. *Studies in the Arabic dialect of the semi-nomadic əl-ᶜAǧārma (al-Balqāʾ district, Jordan)*. Göteborg: University of Göteborg.

——. 1977. "The descriptive imperative of narrative style in spoken Arabic". *Folia Orientalia* 18.5–26.

——. 1978. *Narratives from Ḥesbān: Arabic texts recorded among the seminomadic əl-ᶜAǧārma tribe (al-Balqāʾ district, Jordan)*. Göteborg: University of Göteborg.

——. 1984. "Further notes on the descriptive imperative of narrative style in spoken Arabic". *Studia Orientalia* 55.379–391.

——. 1986. Characteristics of the Arabic dialect of the Ḥwēṭāt tribe. *Orientalia Suecana* 33–35.296–312.

——. 1991a. "Is there a North West Arabian dialect group?". *Festgabe für Hans-Rudolf Singer*, I, ed. Martin Forstner, 151–166. Frankfurt: P. Lang.

——. 1991b. "The form *jāk* in Bedouin narrative style." *Studia Orientalia* 67.55–64.

——. 1992. *Artistic colloquial Arabic*. Helsinki: Finnish Oriental Society.

Piamenta, Moshe. 1979. "Maᶜarexet lešonit šel nimusey ha-bedvim ba-Negev ve-teʾurah ha-lešoni" [The Negev Bedouin verbal etiquette system and its linguistic description]. *Studia orientalia memoriae D.H. Baneth dedicata*, ed. Joshua J. Blau a.o., 125–173. Jerusalem: Magnes Press.

——. 1996. "More on the Arabic dialect of the Negev Bedouins". *Quaderni di Studi Arabi* 14.123–136.

Rosenhouse, Judith. 1984. *The Bedouin Arabic dialects: General problems and a close analysis of North Israel Bedouin dialects*. Wiesbaden: O. Harrassowitz.

—— and Y. Katz. 1980. *Texts in the dialects of Bedouins in Israel*. (= Occasional Papers on the Middle East, N.S., III.) University of Haifa, Jewish-Arab Centre, Institute of Middle East Studies.

Stewart, Frank Henderson. 1990. *Texts in Sinai Bedouin law*, II. Wiesbaden: O. Harrassowitz.

RONI HENKIN
(Ben-Gurion University of the Negev)

Neo-Aramaic

1. THE NEO-ARAMAIC DIALECTS

The modern Aramaic dialects are the remnants of a wide variety of Old and Middle Aramaic dialects that dominated the Middle East in antiquity. The western variety of Aramaic survived only in three villages in the Qalamūn Mountains in Syria, 60 km north of Damascus, whereas the eastern variety survived until the beginning of the 20th century in large areas with hundreds of thousands of speakers in southeast Turkey, in northern Iraq, in Persian Azerbaijan, and in Persian Kurdistan. The massacres in eastern Turkey in 1915 exterminated not only Armenian but also Aramaic culture in the region, so that apart from single villages larger groups of Aramaic-speaking Christians survived only in the Ṭūr ʿAbdīn mountains until the 1960s. Their dialect, known as Ṭuroyo, is still spoken in Turkey, although the majority of the speakers fled from the civil war between Turks and Kurds in the last quarter of the 20th century to Europe. Virtually all the Aramaic-speaking Jews of Iraqi Kurdistan were forced to leave Iraq in 1951, together with the Arabic-speaking Jews of Iraq; the same thing happened to the Aramaic-speaking Jews of Persian Azerbaijan and Persian Kurdistan, although some left Iran only after the revolution of Ayatollah Khomeini. Many Christian Aramaeans, too, fled from this area after the seizure of power by Saddam Hussein in Iraq and the ayatollahs in Iran. A small minority of Aramaic-speaking Mandaeans survived in the city of Ahwaz in Khuzistan in spite of the Gulf wars.

Apart from the Mandaeans, the speakers of Eastern Neo-Aramaic are Christians and Jews, while Western Neo-Aramaic is spoken mainly by Muslims, and only the village of Maʿlūla has a Christian majority.

Most of the Aramaic settlements are situated in regions where the dominant languages are different dialects of Kurdish, Persian, or Turkish. Only the Aramaeans in the Mosul plain, the Mandaeans in Khuzistan, and the speakers of Western Neo-Aramaic live in regions where Arabic is the dominant language.

The fact that Aramaic and Arabic are very closely related languages, both with regard to phonology and morphology, facilitates the mutual adoption and assimilation of loans, but it also makes the identification of loanwords rather difficult. Sometimes, extralinguistic arguments are necessary to prove that a certain word is a loan.

2. DIRECT AND INDIRECT BORROWINGS FROM ARABIC

The Arabic elements in Western Neo-Aramaic and in the Eastern Neo-Aramaic of the Mosul plain are direct borrowings from Arabic, as Arabic is the current language of the region as well as the official language of the country. Khan (2002:xxiii) states that the dialects in the Mosul plain "are rapidly becoming overwhelmed by Arabic". The influence of Arabic along with the official Persian is also present in Modern Mandaic, as the spoken language of Khuzistan is Arabic.

The other Eastern Aramaic dialects received most of their Arabic borrowings indirectly through other intermediary languages, such as Kurdish, Persian, Turkish, and Azeri. Compared with direct borrowings these indirect borrowings show phonological and semantic differences, as is shown in the following examples.

Jewish Aramaic
Salamas (Shahpur): *silah* < (Turkish/Kurdish) *silah* < Arabic *silāḥ* 'weapon'

Christian Aramaic
Mlaḥsô: *hesir* 'refugee' < Kurdish *esîr* < Arabic *ʾasīr* 'prisoner of war'

Modern Mandaic
Ahwaz: *ǧovāb* < Persian *ǧovāb* < Arabic *jawāb* 'answer'

3. ARAMAIZATION OF ARABIC LOANS

In Western Neo-Aramaic, the adoption of Arabic loans without any assimilation to the Aramaic system of phonology and morphology is very rare. Arabic loans are normally integrated into Aramaic phonology and morphology. To a minor extent this is also true for the Eastern Neo-Aramaic dialects. Arabic *ā*, for example, appears regularly as *ō* in Western Neo-Aramaic and at least in older borrowings in the Eastern

Neo-Aramaic dialect of Ṭuroyo, where also ancient Aramaic *ā* shifted to long *ō*:

Ṭuroyo (Eastern Neo-Aramaic)	Maʿlūla (Western Neo-Aramaic)
ḥayōṭo	*ḥayyōṭa* < Old Aramaic *ḥayyāṭā* 'tailor'
naǧōro	*nažžōra* < Arabic *najjār* (itself < Aramaic) 'carpenter'

As in genuine Aramaic words, short vowels in open pretonic syllables are usually elided in words of Arabic origin in Western Neo-Aramaic and to some extent also in Eastern Neo-Aramaic:

Ṭuroyo (Eastern Neo-Aramaic)	Maʿlūla (Western Neo-Aramaic)
lḥēf	*lḥōfa* < *liḥāf* 'blanket'

Only in Western Neo-Aramaic are the speakers aware of the corresponding consonants in Arabic, so that they can easily replace the Arabic consonant by the corresponding Aramaic consonant, e.g. *t* > *č*: *ilčiḫōba* < *iltiḫāb* 'inflammation'; *d* > *t*: *warta* < *ward* 'rose'. In this way, the language is able to borrow nearly any Arabic word. Most of these words are no longer intelligible for speakers of Arabic.

In the field of morphology the Aramaization of Arabic words follows some simple rules, as illustrated by the following representative examples. Masculine singular nouns of Arabic origin receive the nominal ending -*a*, and -*o* in Ṭuroyo and Mlaḥsô. Feminine singular nouns receive an ending reflecting -*ṯā*/-*tā*, which appears in Western Neo-Aramaic as -*ṯa* or -*ča* and in Eastern Neo-Aramaic Ṭuroyo as -*ṯo* or -*to*; in many Jewish Neo-Aramaic dialects, it may even appear as -*la*:

Western Neo-Aramaic	*ḳesma* < *qism* 'part'
	šappṯa < *šābba* 'young lady'

Eastern Neo-Aramaic	
Ṭuroyo:	*ǧəddo* < *jidd* 'grandfather'
	ʾašərto < *ʿašīra* 'tribe'
Hertevin:	*mīra* < *ʾamīr* 'emir'
	dawelta < *dawla* 'state'
Salamas (Jews):	*naǧǧara* < *najjār* 'carpenter'
	šamala < *šamaʿa* 'candle'

In Western Neo-Aramaic, nouns of Arabic origin normally take an Aramaic plural ending. In Eastern Neo-Aramaic, also, some examples can be found:

Western Neo-Aramaic:	masc. pl.: -*ō*, *ḳismō* 'parts'
	fem. pl.: -*ōṯa*, *šappōṯa* 'young ladies'

Eastern Neo-Aramaic	
Hertevin:	masc. pl.: -*ane*, *mirane* 'emirs'
Ṭuroyo	fem. pl.: -*yōṯe*, *ʿašəryōṯe* 'tribes'

Arabic verbal roots are integrated into the Western Neo-Aramaic system of derived Forms (verbal stems) exactly like Aramaic roots within the traditional Aramaic system of derived Forms. This is the case with the following Arabic derived Forms:

	Arabic loan	Aramaic word
I Form	*iḍḥeḳ*	*išmeʿ*
	'he laughed'	'he heard'
II Form	*ḥammel*	*baššel*
	'he loaded'	'he cooked'
IV Form	*aġreḳ*	*arkeš*
	'he fell asleep'	'he woke up'
V Form	*čḥammal*	*čzappan*
	'he endured'	'he was sold'

All Arabic derived Forms can be incorporated into the Aramaic system of derived Forms. Those Arabic derived Forms that do not correspond to an Aramaic derived Form are converted in the following way:

	Arabic loan	Western Neo-Aramaic	
III Form	*šāraṭ*	> *šōreṭ*	'to bet'
VI Form	*tarāfaq*	> *črōfeḳ*	'to accompany'
VII Form	*infajar*	> *inᵊfžar*	'to explode'
VIII Form	*iftaḫam*	> *ifᵊčḫam*	'to understand'
X Form	*istaqbal*	> *sčaḳbel*	'to accept'

The integration of Arabic verbs into the Aramaic verbal system also occurs in the Eastern Aramaic dialects. In Ṭuroyo, for example, Arabic derived Forms I–IV are integrated into the three derived Forms of Ṭuroyo as follows:

I Form	*ṭləble*	< I *ṭ-l-b*	'to demand'
II Form	*mʿalaqle*	< II *ʿ-l-q*	'to hang'
II Form	*mʿawánle*	< III *ʿ-w-n*	'to help'
III Form	*manfaʿle*	< IV *n-f-ʿ*	'to benefit'

Sometimes also Arabic derived Form patterns are used in Eastern Neo-Aramaic with Aramaic inflections:

Qaraqosh	*muftəkə́rhen*	< VIII *f-k-r*	'they thought'
	məstaʿəmlíwa	< X *ʿ-m-l*	'they used to use'

4. Layers of Arabic loans in Neo-Aramaic

The divergence in phonemic representation of Arabic words must be attributed to different layers of borrowing. In some Jewish Neo-Aramaic dialects of northern Iraq, for instance, the Old Aramaic phonemes *f* and *ḥ* are usually reflected by *p* and *x*. This is also the case in older borrowings from Arabic, but in more recent loans the Arabic phonemes are preserved: (Zakho) *qapla* < *qāfila* 'caravan'; *safar* < *safar* 'journey'.

Speakers of Western Aramaic dialects began perhaps already in pre-Islamic times in the transition area between steppe and arable land to adopt loans from Arabian Bedouin, as far as they refer to the vocabulary of specifically Arabic culture. The beginning of language contact in the Aramaic villages of the Qalamūn Mountains, however, is not likely to have occurred before the time of the Umayyads, when the nearby city of Damascus became the capital city of the Islamic Empire.

Thus, we have a period of 1,400 years in which both Arabic and Aramaic underwent many changes. Sound shifts, in particular in Eastern Neo-Aramaic, enable us to distinguish between older Arabic loans and those of more recent times. At the same time, the Arabic loans afford us an insight into the history of the Arabic language in Syria.

The most noticeable sound shifts which occurred in Western Neo-Aramaic concern the so-called *begadkefat* consonants. In an earlier stage of the Aramaic language, the phonemes *b, g, d, k, p*, and *t* had the spirant allophones *v, ġ, ḏ, x, f*, and *ṯ*, which basically appear after vowels. In Neo-Aramaic, the acoustic difference between spirant and plosive pronunciation is preserved but has become lexically determined, so that the former allophones have become phonemes. In Western Neo-Aramaic, the spirants are widely preserved and have in addition spread to word-initial position, whereas the old voiced plosives were devoiced and the old voiceless plosives *k* and *t* were palatalized.

Spirantization is found in some loans of Arabic origin as well. These loans are convincing evidence that in the time when they were incorporated into the Aramaic language, the alternating pronunciation of the *begadkefat* consonants still existed. The oldest layer of Arabic loans in Western Neo-Aramaic, which belong exclusively to the vocabulary of everyday life and not to specifically Arabic culture, shows however that in a secluded mountain area such as the Antilebanon, the alternating pronunciation of the *begadkefat* consonants could survive for a longer period.

In words of Arabic origin we find, for instance, for the Arabic consonants /k/ and /j/ both *k* and *x* and *k* and *ġ* respectively. The Arabic loans in which Arabic /k/ appears as *x* belong to an earlier period in which the phonetic law of spirant pronunciation of *begadkefat* consonants after vowels and in (an older stage of) Western Neo-Aramaic word-initial position was effective. In later loans Arabic *k* remained unchanged in all positions, e.g. (Maʿlūla): old loan *xaffa* < *kafā* 'enough!'; later loan *kaffa* < *kaff* 'palm [of hand]'.

We might posit a pronunciation *g* for the Arabic consonant /j/ for the earliest time of language contact between Aramaic and Arabic in Syria; in any case, this Arabic /j/ was treated like Aramaic *g*. As in Aramaic words, Arabic /j/ shifted after consonants to the voiceless consonant *k* and in word-initial position and after vowels to the spirant *ġ*:

Initial position:	*ġmōʿča* < *jamāʿa*	'crowd'
After vowels:	*farraġ* < *farraj*	'he looked'
After consonants:	*mawkʿa* < *mawjaʿ*	'pain'

In all later borrowings, Arabic /j/ appears in all positions as *ž* (*žayša* 'army'). The above-mentioned word for 'pain' is of special interest, since it was borrowed twice in two different periods. The older loan *mawkʿa* is used now along with the later *mawžʿa*, and speakers are not aware that both words have the same Arabic origin. They rather believe that *mawkʿa* is an Aramaic word, whereas *mawžʿa* is an Arabic loan.

Modern loans may contain consonants which are not represented in the Aramaic phonological inventory, for example the consonants *ḥ* and *ḍ* in Modern Mandaic, e.g. (Ahwaz) *ṣafḥa* ‘page’, *xaḍrā* ‘vegetables’.

Because the Neo-Aramaic dialects are village dialects and normally have no literature, the Arabic loans also have their origin in the spoken Arabic dialects of the area. The Arabic loan *ǧihil* ‘child’ in Modern Mandaic is a word typical of the Arabic dialect of the Arabian Gulf. In more recent times, words from more prestigious city dialects and from Standard Arabic have been borrowed, too.

An Arabic word borrowed into Western Neo-Aramaic twice is *ḍarf*. It was adopted in earlier times from the Arabic dialects in the neighboring villages as *ḍarfa* with the meaning of ‘skin bag for butter’. In a later period, the word was borrowed with different pronunciation once more from modern Damascene Arabic as *ẓarfa* with the meaning of ‘envelope’.

BIBLIOGRAPHICAL REFERENCES

Arnold, Werner. 2002. “Zur Geschichte der arabischen Lehnwörter im Neuwestaramäischen”. *Neue Beiträge zur Semitistik: Erstes Arbeitstreffen der Arbeitsgemeinschaft Semitistik in der Deutschen Morgenländischen Gesellschaft vom 11. bis 13. September 2000 an der Friedrich-Schiller-Universität Jena*, ed. Norbert Nebes, 5–11. Wiesbaden: O. Harrassowitz.

Garbell, Irene. 1965. *The Jewish Neo-Aramaic dialect of Persian Azerbaijan: Linguistic analysis and folkloristic texts*. The Hague: Mouton.

Jastrow, Otto. 1988a. *Der neuaramäische Dialekt von Hertevin (Provinz Siirt)*. Wiesbaden: O. Harrassowitz.

——. 1988b. *Laut- und Formenlehre des neuaramäischen Dialekts von Mīdin im Ṭūr ʿAbdīn*. 4th ed. Wiesbaden: O. Harrassowitz.

——. 1993. *Der neuaramäische Dialekt von Mlaḥsô*. Wiesbaden: O. Harrassowitz.

Khan, Geoffrey. 2002. *The Neo-Aramaic dialect of Qaraqosh*. Leiden: E.J. Brill.

Sabar, Yona. 1984. “The Arabic elements in the Jewish-Neo-Aramaic texts of Nerwa and ʿAmādīya, Iraqi Kurdistan”. *Journal of the American Oriental Society* 104.201–211.

WERNER ARNOLD (University of Heidelberg)

Neologism → Terminology

Nigeria

1. THE INTRODUCTION OF ARABIC AND ISLAM TO NIGERIA

In the precolonial era, the West African country now known as Nigeria was referred to as one of the countries in the Western Sudan. This reference was adopted by early Arab writers such as al-Bakrī, Ibn Xaldūn, at-Timbuktī, and as-Saʿadī. Others identified it as Bilād Tukrūr. When the European colonists came, they imposed on it the name of Nigeria. Nigeria consists of several ethnic groups and tribes speaking approximately 396 vernaculars. A few of these tribes, the Hausa, the Yoruba, and the Igbo, have gained dominance over the others and are recognized as majority tribes. Others are recognized as minority tribes.

The relationship between the Arabic language and the Nigerian people began very early in the history of some of the major tribes inhabiting the country. Their historical affinity to Arabs and the Arabic language in Nigeria may be full or partial. Full affinity applies to an Arab tribe called the Shuwa. They live in the northeastern part of the country, constituting a part of the old Kanem Borno Empire. Today, they are found in the states of Borno and Yobe, along Lake Chad. The Shuwa speak a dialect of Arabic and exhibit several Arab customs in their daily life (Imam 2002). Several studies have been and are still being carried out on the tribe, its language Shuwa Arabic (→ sub-Saharan Arabic), and the works of some of its prominent scholars. A milestone in the history of Shuwa Arabs was the establishment of an Arabic village, as demanded by students of Arabic in Nigerian universities. The federal government built this village in the midst of the Shuwa tribe at Ngala, near Lake Chad.

A partial historical relationship with Arabic exists for some Nigerian tribes that claim a distant descent from Arabs. They include the Beriberi, the Fulani, the Hausa, and the Yoruba. These tribes also claim that their languages have been influenced by Arabic. In Yoruba, some scholars, for instance Al-Ilori (n.d. 34), go so far as to claim that the Yoruba words of Arabic origin constitute at least 50 percent of the Yoruba language, although this is not generally accepted (→ Yoruba). In addition, it has been observed that words of Arabic origin in Yoruba are found in all aspects of daily life, such as the political, economic, legislative, educational, and health spheres (Olayiwola 2002).

Some reports suggest that Arabic infiltrated into Nigeria even in pre-Islamic times through trans-Saharan trade routes. Mustapha (1987) suggests that even before the coming of Islam, Arabs and the Arabic language were not complete strangers on Nigerian soil. Yet, recent findings have shown that the reports of trans-Saharan trade have been embellished. According to these findings, war and raids characterized virtually all communities at that point in time, so it is unlikely that mutual trading activities could have flourished as asserted in those reports. Likewise, it may be noted that the nature of the trading activities, the commodities, the medium of exchange, and the successes which attended the trans-Saharan trade before the spread of Islam are obscure (Ajayi and Alagoa 1980:228). The conclusion must therefore be that it was only when Islam came that trans-Saharan trade received a boost and manifested itself (Olayiwola 2001). This implies that the impact of the trans-Saharan trade as a conductor of the Arabic language to Nigeria before the spread of Islam was either nonexistent or very weak.

Islam serves as the harbinger to Arabic, while Arabic in turn serves as the gateway to Islam. The connection between the two is so tight that various researchers have referred to Arabic and Islamic studies as twin studies. Abubakar (1972), Fafunwa (1974), Gbadamosi (1978), and Galadanci (1982) all emphasize the inseparable alliance between Arabic and Islam.

Arabic and Islamic studies are taught and administered in the same department in most Nigerian universities where these courses are available. Although experts are found teaching and specializing in either discipline, experience has shown that many lecturers in these institutions teach courses in the two disciplines interchangeably, as for instance Isaac Ogunbiyi, who is a scholar in the fields of both Arabic and Yoruba, being the editor of the first Yoruba folktales written in Arabic (Oseni 2002). Similarly, anyone who speaks Arabic is automatically attached to Islam. Azeez (2004) summarizes the situation as follows: "Right from its advent to the present time Islam has always been studied in Nigeria alongside Arabic and there has never been a time that the two are totally separated".

2. The development of the study of Arabic in Nigeria

The study of Arabic has passed through stages of development since the time of its introduction to Nigeria. Arabic was introduced in the period before the 10th century to the end of the 13th century. Islam demands a great deal of knowledge. The first revelation to the Prophet Muḥammad was an instruction on knowledge. It highlighted three basic elements involved in the process of acquiring knowledge, i.e. reading, teaching, and the pen (Q. 96/1–5). Therefore, no sooner had Islam penetrated into Borno than people understood the necessity of having a certain degree of knowledge of Arabic as a requisite for a successful practice of the new faith. The *kalimat aš-šahāda* has to be pronounced in Arabic and the ritual worship *ṣalāt* has to be performed five times daily through the medium of Arabic. For these and other ritual purposes, the whole of the *Qurʾān*, or at least some portions of it, have to be learned by heart. The desire to satisfy this religious need marks the beginning of the learning of Arabic.

In Nigeria, only Borno has had the singular privilege of being ruled by scholars, particularly at the time when Arabic was introduced, when many rulers memorized significant portions of the *Qurʾān* (Mustapha 1987). In the other regions, the position of Arabic was different. In Hausaland, individuals who converted to Islam or their children constituted the first group of scholars. After having studied the *Qurʾān*, they in turn taught it to others.

It was at about the end of this era that Islam and Arabic studies arrived in Yorubaland. The ruler, Oba, merely tolerated the scholars, allowing them to set up autonomous compounds in the suburbs where they practiced Islam and taught Arabic among themselves without being bothered (Gbadamosi 1978; Olayiwola 1990).

From the 14th to the 18th century, Arab scholars from North African countries, such as Muḥammad ibn Mānī and al-Maġīlī, visited what is now known as Nigeria. This age witnessed a tremendous growth of the study of Arabic in Nigeria. Arab scholars came and taught Nigerians the Arabic language, beginning with the *Qurʾān*. They also brought with them some books which were unknown before then. Thus, Muḥammad ibn Mānī introduced the *Risāla* of ʾAbū Zayd al-Qayrawānī to the palace of Borno.

In the last part of this period, Arabic had spread to such an extent that Borno rulers could establish strong links between their empire and the Arab lands, such as Tunisia, Morocco, Egypt, and even the Ḥijāz (Mustapha 1987). There are reports of written communication in Arabic to these Arab lands from the Borno rulers.

In Hausaland, Muḥammad ibn ʿAbd al-Karīm al-Maġīlī at-Tilimsānī visited Kano in the same period and taught its rulers Arabic (Galadanci 1982). This era also marked the influx of Malian traders and scholars into Yorubaland. Al-Ilori (1978) is of the opinion that those scholars *cum* traders taught and spread Islam and Arabic studies among people in Yorubaland during the reign of a Yoruba monarch called Alafin Obalokun.

In this period, too, a conscious effort was made to teach Arabic to children, and a curriculum for this type of education was developed. According to Al-Ilori (1978:53), this curriculum was similar to the one described for Morocco by Ibn Xaldūn in his *Muqaddima*.

Katātīb or elementary schools were held in three places: the mosque, the house of the imam or his assistant, or under shady trees in the compound. The imam, his deputy, or any one among the members of the immediate community volunteered to shoulder the responsibility of teaching the children. Even though this was a voluntary job, people who possessed the knowledge and were capable of transmitting it competed for it.

The teacher is known as *mallam* (< Arabic *muʿallim*), *modibbo* (< Arabic *muʾaddib*), *maalam*, and *aafaa* (connected with various Arabic etyma, e.g. < *xāʾifan* 'fearing', *ʿafāf* 'abstinence, asceticism', *ʾalfan* 'thousand') in Barbarchi, Fulfulde, Hausa, and Yoruba, respectively. At the earliest stage of study, children of about two years of age and above are gathered in groups and taught to recite the shortest chapters of the *Qurʾān*. This is read to them in short pieces, verses or parts of verses. The teacher reads while the pupils repeat in echo. Through this sing-song pattern, the short chapters of the *Qurʾān* are learned by heart without reading or writing.

Later, the pupils are introduced to writing on wooden slates called *allo* in Barbarchi and Hausa, *alloha* in Fulfulde (< Arabic *al-lawḥ*), or *wa:la:* in Yoruba. The pen, which is known as *alkaram, alkalami* (< Arabic *al-qalam*), *bindirgol*, or *kalamu*, respectively, is inserted in the ink, which is called *adowa, tawada, dawaa*, and *ta:da:* (< Arabic *dawā*) in Barbarchi, Hausa, Fulfulde, and Yoruba as well, and used to write on the slate.

Writing begins with the consonants of the Arabic alphabet and then the vowel signs. After the comprehension of each letter in its different position and shapes, the short chapters are taught in the form of spell-reading. This goes on simultaneously with further memorization of the *Qurʾān*. Through this method, the whole of the *Qurʾān* or a part of it is memorized, with the ability to read and write all of the parts.

After the study of the *Qurʾān*, competent and willing students begin the next level, which is generally known as *ilmi* (< Arabic *ʿilm* 'scholarship, knowledge'). Usually this level is attended by children of scholars. At this level, the beginning scholar is introduced to other books and disciplines, such as *ḥadīṯ, tafsīr*, and *fiqh* (Būsairī 2002). Graduation from this level is normally at adulthood. The graduate can then devote himself to teaching.

In rare cases, the graduate from this level proceeds to the next level, which may be called the level of professionalization. Complex subjects such as philosophy, astrology, poetry, mathematics, and world history are combined with deeper study of the *Qurʾān* and jurisprudence. Usually, scholars at this level do not graduate; rather, they shift to another teacher as they switch from one subject to the other.

During the period between 1804 and 1903, Sheikh ʿUṯmān Danfodio waged his *jihād* against the monarchy in the north and established what has become known as the Sokoto Caliphate. During this period, Arabic studies were pursued with unprecedented zeal. The language of the administration was Arabic. Administrators, judges, and other public officers were appointed on merit, one of the required qualities being literacy in Arabic.

Although the Sokoto Caliphate was not the first Islamic administration in Nigeria, its pursuit of Arabic knowledge was more zealous, and its application of the language affected the community more. Arabic manuscripts produced during this age are still being studied and referred to as evidence of the high quality of Arabic studies in Nigeria.

There was a retrogression in the study of Arabic during the age of colonization (1903–1960). The expanding Sokoto Caliphate was curtailed by the British colonists in 1903 (Hisket 1984). When the British colonists introduced English education, they met Muslim communities with their own type of education in Arabic. The Sokoto Caliphate, and the Borno Empire in particular, kept their records and conducted their administration in Arabic. Perceiving Arabic as a threat to their language and civilization, the British Administration took measures against Arabic and put hurdles on its track. They relegated Arabic to the background and encouraged and promoted the English language by elevating the status of their trainees in English in the society. Their trainees were offered paid jobs and appointed to leadership positions in all sectors of the administration.

Muslim scholars who were teaching Arabic lost their elevated position and respected status in society, and were not offered jobs. When the Muslims insisted that the twin subjects of Arabic and Islamic studies had to be taught in Western schools if Muslims were to contribute fully to the success of the administration, Arabic and Islamic studies were introduced in the missionary schools and some Muslims were employed. They were, however, discriminated against on the job since their salaries did not take into account their years of study and experience, as was the case with those under Western education; they had no on-the-job training opportunities; they were low-paid; and they were not monitored to put in their best (Galadanci 1982). All this resulted in the discouragement of Arabic in all Muslim societies. Some Muslims who gave in to this pressure and sent their children to the English schools had to accept that the majority of them converted to Christianity.

After independence in 1960, Muslims were able to improve the status of Arabic and its teaching. Due to the fact that the ritual prayer must be said in Arabic, Muslims have the obligation to learn sufficient Arabic to be able to perform this duty. The number of those who speak Arabic as a language of communication is much lower, however. Arabic schools are numerous in those parts of Nigeria where Muslims have a majority, i.e. in the northern and southwestern parts of the country. There are two types of Arabic students: those who learn the *Qurʾān* by heart (called *Almajirai*) and those who learn the *Qurʾān* along with other Islamic subjects under the Islamiyah system. Not much is known about the number of people in each category. The government of Kano State, which has a large Muslim majority, announced in 2004 that the number of its *Almajirai* alone is two million.

Institutions teaching Arabic and Islamic studies have proliferated (Galadanci 1982). Universities established on the platform of Western education have started to open departments for Arabic and Islamic studies. In 1964, for instance, the University of Ibadan started a certificate program, designed to assist students of private Arabic schools in their further development in the field (Abubakre 2002). At the present time, all federal universities, apart from those specifically dedicated to science or agriculture, teach Arabic and Islamic studies. Colleges of the federal government and some of the colleges owned by the states not only offer Arabic as a course but also recognize it as one of the official languages on their premises. Arabic has found its way into the broadcasting houses. *Ṣawt Nayjīriyā* broadcasts in Arabic from Lagos.

Although the electronic age with its computers and electronic devices did not start with the year 2000, this year marked a clear departure from the past. For instance, before this year very few computer centers were available in the different parts of Nigeria with Arabic Windows. From the year 2000, however, several centers where Arabic can be processed and printed have been established. Publishing houses with facilities for quality printing and publishing of Arabic materials are on the increase. Bilingual academic journals have increased in number in Nigerian universities. Arabic magazines, periodicals, and weekly newspapers have been founded, particularly in Lagos.

With the penetration of Islam into Nigeria, Arabic accompanied the faith as a tool necessary for its understanding. Since then, the two disciplines have maintained a close relationship. Between then and now, the study of Arabic has witnessed a tremendous growth. It has grown from the status of a discipline learned only by rote memorization, studied in private houses and under shade trees, into a huge complex field with different branches and ample opportunities, offered through all types of media.

BIBLIOGRAPHICAL REFERENCES

Abubakar, Ali. 1972. *aṯ-Ṯaqāfa al-ʿarabiyya fī Nayjīriyā*. Beirut: Mu'assasa ʿAbd al-Ḥafīẓ al-Bassāt.

Abubakre, Razaq Deremi. 2001. "Ḥāḍir al-luġa al-ʿarabiyya fī Nayjīriyā". *Journal of the Nigerian Association of Teachers of Arabic and Islamic studies* 6:1.199–212.

——. 2002. *Survival of Arabic in difficult terrains*. Ilorin: Library and Publication Committee, University of Ilorin.

Ajayi, J.F. Ade and E.J. Alagoa. 1980. "Nigeria before 1800: Aspects of economic developments and inter-group relations". *Ground work of Nigerian history*, ed. Obaro Ikime, 224–235. Ibadan: Historical Society of Nigeria.

Al-Ilori, Adam Abdallah. N.d. *ʾAṣl qabāʾil yawruba*. Agege: Maṭbaʿat aṯ-Ṯaqāfa al-ʾIslāmiyya.

——. 1978. *al-ʾIslām fī Nayjīriyā wa-š-šayx ʿUṯmān dan Fōdio*. Agege: Maṭbaʿat aṯ-Ṯaqāfa al-ʾIslāmiyya.

Azeez, Waheed O. 2004. "Arabic and Islamic education in Nigeria from rote learning to e-learning". Unpublished seminar paper on the place of Arabic and Islamic studies in the globalized world, delivered at the *22nd Conference of the Nigerian Association of Teachers of Arabic and Islamic studies*, Lagos, 29 December 2003–2 January 2004.

Būsayrī, Murtaḍā. 2002. "at-Taʿlīm al-ʿarabī wa-ʾamākinuhu qabla l-madāris al-ʿarabiyya an-niḍāmiyya fī Nayjīriyā ʿāmmatan wa-ġarbihā xāṣṣatan". *Journal of the Nigerian Association of Teachers of Arabic and Islamic Studies* 6:2.121–141.

Fafunwa, Aliyu Babs. 1974. *History of education in Nigeria*. London: Allen and Unwin.

Galadanci, Saʿīd. 1982. *Ḥarakat al-luġa al-ʿarabiyya wa-ʾādābihā fī Nayjīriyā*. Cairo: Dār al-Maʿārif.

Gbadamosi, T.G.O. 1978. *The growth of Islam among Yoruba, 1841–1908*. London: Longman.

Hisket, Mervin. 1984. *The development of Islam in Africa*. London: Longman.

Imam, Abdullah Muhammad. 2002. "al-Luġa al-ʿarabiyya wa-l-bīʾa an-nayjīriyya". *Journal of the Nigerian Association of Teachers of Arabic and Islamic Studie*s 6:2.97–108.

Mustapha, Abubakar. 1987. *The contribution of* saifawa ʿulama *to the study of Islam, 1086–1846*. Ph.D. diss., Bayero University Kano.

Olayiwola, Abdulfattah. 1990. *Effects of traditional culture on the practice of Islam in Iseyin*. M.A. thesis, University of Ibadan.

——. 2001. *The polytomy of Islam in Nigeria: Emanation implications and possible rectification*. Ph.D. diss., University of Jos.

——. 2002. "ʾAwjuh aš-šibha bayna l-luġatayn al-ʿarabiyya wa-l-yawruba". *al-ʾAqlām* 1.38–49.

Oseni, Zakariyau I. 2002. *Prose and drama in Nigerian literature in Arabic: The journey so far*. Ilorin: Library and Publication Committee, University of Ilorin.

ʿABDULFATTAH OLAYIWOLA
(University of Maiduguri)

Nigerian Arabic → Subsaharan Africa

Nisba

1. INTRODUCTION

The term *nisba* refers to an adjective which is derived from a noun. The *nisba* is sometimes called → *ʾiḍāfa* (Sībawayhi, *Kitāb* III, 335). It is first aimed at creating adjectives related to countries, tribes, cities, and persons (Ibn ʿAqīl, *Šarḥ* II, 471) and thus plays an important role in personal names (see Sec. 5).

There are many ways in Arabic to form the *nisba*: (i) insertion of the infix *–iyy–* between the noun and the declensional ending, which is by far the most common way; (ii) insertion of the infix *–āniyy–*, which denotes an intensive meaning and is used today mainly in the dialects; and (iii) the patterns *faʿʿāl, fāʿil, faʿil, faʿāl*^{in} (*al-faʿālī*), *mifʿāl*, and *mifʿīl*, which are no longer productive. The *nisba* is an adjective and is declined regularly, according to general adjectival patterns and grammatical rules.

For cases in which the formation of the *nisba* was doubtful or problematic, early grammarians relied on what they heard from the Bedouin Arabs or deduced the form by analogy (*qiyās*) with existing forms. Quite often, they would even give more than one possibility and indicate their preference. Their aim was not only to prescribe or describe but also to make clear the underlying structures (Versteegh 1997:74).

Reading the examples given by the grammarians, one might wonder about the purpose of forming a *nisba* from certain kinds of words: peculiar words ('little shark', 'young female slave'), words which already have an adjectival meaning ('long', 'good'), non-nominals ('if', 'how much'), numbers ('fifteen'), phrases ('the father of Bakr', 'sons of dogs'), complete sentences ('he carried evil'), and even words that are already morphologically *nisba*s (*aš-šāfiʿiyy*, name of the founder of a school of law). Actually, one should bear in mind that all of these are or could be personal names (particularly nicknames), and as such are in need of a *nisba*.

2. The infix *–iyy–*

The infix *–iyy–* is the most common way to form a *nisba*. It is inserted between the word and the declensional ending. Since the declensional endings tend to be left out in the oral realization of Modern Arabic, the masculine singular form of this infix is generally pronounced and transcribed *–ī*. In this article, it is transcribed fully in order to avoid confusion with the 1st person pronominal suffix *–ī*. Compare *manzilī* 'my home' (underlying structure /manzil–ī/) and *manzilī* 'domestic' (underlying structure /manzil–iyy–/).

The infix *–iyy–* is declined regularly. The feminine singular form is *–iyyat–*; the dual forms are *–iyyāni/–iyyayni*; and the plural forms are *–iyyūna/–iyyīna* and *–iyyāt–*.

The infix is generally added to the masculine singular form of the word. Thus, the six words *muslim, muslima, muslimāni, muslimatāni, muslimūna, muslimāt* 'Muslim [masc. sg., fem. sg., masc. du., fem. du., masc. pl., fem. pl]' share the same *nisba: muslimiyy.*

The feminine singular *nisba*, *–iyyat–*, has in more recent times been used to create neologisms and abstract nouns (Sublet 1993:54) such as *taʿbīr* 'expression': *taʿbīriyya* 'expressiveness'; *huwa* 'he': *huwiyya* 'identity'. The feminine plural *nisba*, *–iyyāt–*, is also used to form abstract nouns, e.g. *luġa* 'language': *luġawiyyāt* 'linguistics'; *ṣawt* 'sound': *ṣawtiyyāt* 'phonetics' (→ terminology).

In some cases, the addition of the infix may cause phonetic changes. If the word already ends in two *yāʾ*s, they are simply replaced by the infix *–iyy–* (*kursiyy* 'chair': *kursiyy*). If the word ends in an *ʾalif maqṣūra* or *ʾalif manqūṣa*, these are deleted (*ḥubārā* 'bustard': *ḥubāriyy*; *muʿtad*^{in} 'aggressor': *muʿtadiyy*). If the vowel before the last consonant is a *kasra*, it changes into a *fatḥa* (*malik* 'king': *malakiyy; namir* 'panther': *namariyy*). If the word contains the sequence *-yyi-* before the last consonant, this is replaced by *–y–*/ (*ṭayyib* 'good': *ṭaybiyy*). If the word is of the pattern *faʿīla* and its root is neither hollow nor contains a geminate, it changes to the pattern *faʿaliyy* (*ḥanīfa* 'orthodox': *ḥanafiyy*; but *ṭawīla* 'long': *ṭawīliyy*, because the root is hollow). If the word is of the pattern *fuʿayl* or *fuʿayla* and its root is neither hollow nor contains a geminate, it turns to the pattern *fuʿaliyy* (*qurayš* 'little shark': *qurašiyy; juhayna* 'little girl': *juhaniyy*; but *jalīla* 'venerable': *jalīliyy*, because here the root contains a geminate). In phrases and sentences, the second element is deleted unless it is the more meaningful, so that the first part is deleted (*xamsata ʿašar* 'fifteen': *xamsiyy*; *ʾabū bakr* 'the father of Bakr': *bakriyy; ibn az-zubayr* 'the son of az-Zubayr': *zubayriyy; ʿabd manāf* 'the slave of Manaf': *manāfiyy; taʾabbaṭa šarr*^{an} 'he carried evil': *taʾabbaṭiyy*). In compounds it depends on the case (*ḥaḍramawt* 'Hadramawt [a region in modern Yemen]': *ḥaḍramiyy*; but *baʿalbakk* 'Baalbek [a city in Lebanon]': *baʿliyy*). In biradical non-nominals, the second consonant may be doubled (*kam* 'how much?': *kammiyy; law* 'if': *lawwiyy*).

In some cases, a hidden root consonant may resurface in the *nisba*, and there is sometimes a discussion among grammarians about the nature of this hidden consonant, as in *ʾab* 'father': *ʾabawiyy*; *sana* 'year': *sanawiyy* or *sanahiyy; šafa* 'lip': *šafiyy* or *šafahiyy* or *šafawiyy; ḥir* 'female genitals': *ḥiraḥiyy* (*sic!*; cf. Sībawayhi, *Kitāb* III, 359.3–4)

There are exceptions to these general rules. For example, in some rare cases the *nisba* can be formed from the plural form of the noun, especially for personal names (*banū kilāb* 'sons of dogs': *kilābiyy; ʾanṣār* 'supporters [of Muḥammad in Medina]': *ʾanṣāriyy*), nouns that do not have a singular form (*ʾabābīl* 'ababil birds': *ʾabābīliyy*), or nouns that have a different meaning in the singular (*al-jazīra* 'the island': *jazīriyy*; but *al-jazāʾir* 'Algeria': *jazāʾiriyy* 'Algerian'). There are other notable exceptions, with no explanation, such as the *nisba* of *baṣra* (a city in modern Iraq): *biṣriyy* instead of **baṣriyy* (*sic!*; cf. Sībawayhi, *Kitāb* III, 336.2; Ibn ʿAqīl, *Šarḥ* II, 471.2–3); *marw* (modern Mary, a city in Turkmenistan): *marwaziyy*.

In many cases, especially in some dialects, /w/ is added before the infix *–iyy–*, which surfaces as *–awiyy–* or *–āwiyy–*. The word itself may undergo minor modifications. The addition of this *wāw* depends on the number of syllables in the word, as well as on the final phonemes of the word. It occurs after biradical words (*ibn* 'son': *banawiyy; dam* 'blood': *damawiyy; iṯnān* 'two': *ṯanawiyy*); after short words ending in *ʾalif maqṣūra* (*fatā* 'adolescent': *fatawiyy*; but *muṣṭafā* 'Mustafa': *muṣṭafiyy*, because it is a longer word); after short words ending in *ʾalif manqūṣa* (*šaj*^{in} 'worried': *šajawiyy*; but *muʿtad*^{in} 'aggressor': *muʿtadiyy*, because it is a longer word); after short words ending in double /yy/ (*nabiyy* 'prophet': *nabawiyy; ḥayy* 'alive': *ḥayawiyy*); after words with the feminine ending *-āʾ* (*ṣaḥrāʾ* 'desert': *ṣaḥrāwiyy*; but *kisāʾ* 'garment': *kisāʾiyy*, because here *-āʾ* is not a feminine ending); after short words with the feminine ending *Cyat*–, where C is any consonant (*zinya* 'last child': *zinawiyy*); after words whose root is weak (*quṣayy* [a male given name]: *quṣawiyy; ʾumayya* 'young female slave': *ʾumawiyy*). If the root is both assimilated and weak, the initial *wāw* is restored (*šiya* 'blotch': *wišawiyy*).

In some rare cases, a *hamza* is added and the infix surfaces as *–āʾiyy–*. This applies generally to words with the feminine ending *–āyat–*, as in *siqāya* 'irrigation': *siqāʾiyy*. Here, the ending and the infix combine into a new ending.

These are only a few examples, but the grammarians present many more detailed cases, and quite often they give two or three possible *nisba*s for one word. For example, *ʾāya* 'verse' may have the following *nisba*s: *ʾāʾiyy, ʾāyiyy*, and *ʾāwiyy*, both because it is a short word and because it ends with a radical consonant *yāʾ*; and *ḍabya* 'female gazelle' may have the following *nisba*s: *ḍabyiyy* and *ḍabawiyy*, according to usage.

There are also many exceptions to the preceding rules, such as *zakariyyāʾ* [a male given name]: *zakariyyāwiyy*, although the ending *āʾ* is not feminine; *qāḍ*^{in} 'judge': *qāḍiyy*, although it is a short word.

In modern reference grammars (ʾAfġānī 1971; Niʿma 1973), the morphological rules for the formation of the *–iyy– nisba* are very simplified, especially the rules for the addition of a consonant before the infix.

3. The infix *–āniyy–*

In some cases, the infix *–āniyy–* is used instead of *–iyy–* in order to indicate an 'intensive' *nisba*. In Classical Arabic, it was used only for physical traits (*liḥya* 'beard': *liḥyāniyy* 'wearing a long beard'; *raqaba* 'neck': *raqabāniyy* 'stiff-necked'; *šaʿr* 'hair': *šaʿrāniyy* 'having strong hair'). In many dialects, this intensive *nisba* is used more frequently than the regular *–iyy–* (compare *ʾamrīkiyy* and *ʾamrīkāniyy* 'American'; *taḥtiyy* and *taḥtāniyy* 'lower').

4. The patterns *faʿʿāl, fāʿil, faʿil, faʿāl*^{in} *(al-faʿālī), mifʿāl* and *mifʿīl*

When the *nisba* is used to indicate the profession of the person represented in the root noun, such as 'related to clothes', 'related to bread', 'related to camels', the pattern *faʿʿāl* is used, as in *ṯiyāb* 'clothes': *ṯawwāb* 'he who sells clothes'; *xubz* 'bread': *xabbāz* 'baker'; *ṣarf* 'expense': *ṣarrāf* 'money changer'.

When the *nisba* is used to indicate a relationship meaning 'possessor of', the patterns *fāʿil, faʿil, mifʿāl*, and *mifʿīl* are used, as in *dirʿ* 'armor': *dāriʿ* 'armored'; *nabl* 'arrow': *nābil* 'archer'; *ṭaʿām* 'food': *ṭaʿim* 'possessor of food'; *libās* 'clothes': *labis* 'possessor of clothes'; *ʿiṭr* 'perfume': *miʿṭār* 'possessor of perfume'; *ḥuḍur* 'speed': *miḥḍīr* 'running fast'.

The pattern *faʿāl*^{in} (*al-faʿālī*) also carries a *nisba* meaning, as in *yaman* 'Yemen': *yamān*^{in} 'Yemenite'; *aš-šaʾm* 'Syria': *šaʾām*^{in} 'Syrian'.

These patterns are no longer productive in modern Arabic to build new *nisba*s.

5. The *nisba* in the onomastic sequence

For general information about the use of the *nisba* in onomastics, see Caetani and Gabrieli (1915), Schimmel (1989), and Sublet (1991, 1993). The *nisba* 'name of relation' reveals the privileged participation of a personage in a determinate geographic space, as in the ethnonym *al-qāhirī*, or in a definite sociocultural area, as in *al-ʾanṣārī* (cf. Fierro 2004), *al-ḥanbalī, aṣ-ṣūfī*. (In this section, the *nisba* ending is transliterated by *ī* instead of *iyy*, according to conventional usage in transliterating names.) Whether the onomastic sequence by which a person is designated is short or expanded, the *nisba* is always placed at the end (→ proper names).

In fact, in contrast to the *ism* or the *kunya*, the *nisba* is not the primary component of the medieval Arabic name. Most of the persons mentioned by al-Jawzjānī (d. 259/873), for instance, in his *ʾAḥwāl ar-rijāl*, do not have any *nisba*. The *nisba*, however, supplies a complement of information often significant. Sometimes, it differentiates between homonyms, for example, one *kunya* from another in al-ʾAzdī

(d. 374/984; *ʾAsmāʾ* 29), who distinguishes between ʾAbū ʾUmāma al-Bāhilī, ʾAbū ʾUmāma al-ʾAnṣārī, and ʾAbū ʾUmāma al-Ḥāriṯī. The *nisba*'s special function is to reveal the sociocultural links of the person (properties, tribe, family, sect, clientele, teacher, legal *maḏhab*). This is above all the case in the great biographical repertories from the Mamluk period (Ayalon 1975). In aṣ-Ṣafadī (d. 764/1263) and al-Maqrīzī (d. 845/1441), as well as in Ibn Ḥajar al-ʿAsqalānī (d. 852/1449) and as-Saxāwī (d. 902/1497), the *nisba* can be aggregated differently, according to the wish of the author of the biographical notices. Different authors may present the same person with different chains of *nisba*s, giving the reader a very precise but nuanced idea of what his life was like.

Thus, the sequence composed by al-Maqrīzī about one of his contemporaries, ʾIsmāʿīl ibn ʾIbrāhīm ibn ʿAbd aṣ-Ṣamad, includes five *nisba*s: al-Hāšimī al-ʿAqīlī al-Jabartī aṣ-Ṣūfī aš-Šāfiʿī (Maqrīzī, *ʿUqūd* I, 404-406). Ibn Ḥajar al-ʿAsqalānī, in turn, composed a notice on this same person, whom he regarded as one of his teachers (Ibn Ḥajar, *Muʿjam* III, 85, no. 447), attributing to him only two *nisba*s: al-Jabartī ṯumma az-Zabīdī, this last *nisba* being the transformation of the expression *nazīl Zabīd*, which we find at the end of al-Maqrīzī's onomastic sequence. As for as-Saxāwī, while he picks up the nomenclature of al-Maqrīzī, he remains faithful to his teacher Ibn Ḥajar: al-Hāšimī al-ʿAqīlī al-Jabartī ṯumma az-Zabīdī aš-Šāfiʿī (Saxāwī, *Ḍawʾ* II, 282–283, no. 893). Here, the *nisba aṣ-ṣūfī*, which connects the person to the world of the mystics, has disappeared. Aš-Šawkānī (d. 1250/1834) follows his example, using precisely the same combination as as-Saxāwī (Šawkānī, *Badr* I, 139). Thus, the nisba can disappear from an onomastic sequence from one author to another, but it can also acquire additional specifications. By the expression *as-saxāwī al-ʾaṣl al-qāhirī al-mawlid* we understand that this person's family hailed from Saxā (a village in West Fusṭāṭ), but he himself was born in Cairo. The most relevant words are *al-ʾaṣl* (his origin, whether of the person himself or of one of his ancestors), *al-mawlid* (his birth), *al-manšaʾ* (where the person grew up, received his first education, or even more), *ad-dār* (his residence), *al-mawqiʿ* (his situation, especially with regard to the legal *maḏhab* he belonged to), and *al-wafāt* (his death).

A *nisba* may be equivocal for two main reasons: (i) defective script (without subscripted points, *al-jamalī* can be confused with *al-ḥamulī*); and (ii) homonymy (possible confusion between different *nisba*s, e.g. *al-ḥanafī*, formed from the tribal name of the Banū Ḥanīfa, and *al-ḥanafī*, formed from the name of the founder of the Ḥanafī *maḏhab*, al-ʾImām ʾAbū Ḥanīfa). These difficulties are amplified when the *nisba* reaches the level of the 'name of celebrity' (*šuhra*). In order to deal with these difficulties, medieval authors at an early age composed specific works collecting and sorting all the problematic *nisba*s. For problems with unvocalized spelling, al-ʾAzdī's *Muštabih* may be mentioned here, and for problems arising out of homonymy, al-Qaysarānī's *ʾAnsāb*.

The most important of the onomastic works is certainly the *Kitāb al-ʾansāb* 'Book of genealogies and proper names of relation' by ʾAbū Saʿd as-Samʿānī (d. 562/1166). This book was abridged and corrected by ʿIzz ad-Dīn Ibn al-ʾAṯīr (d. 630/1233; Ibn al-ʾAṯīr, *Lubāb*), which in its turn was abridged and revised by the Egyptian polygraph Jalāl ad-Dīn as-Suyūṭī (d. 911/1505; Suyūṭī, *Lubb al-lubāb*). Thus, the *nisba* was, in both scholarly and folk literature (Galley 2003), not only the way by which persons were described in their own personal history but also in their strong adherence to any number of group identities. This usage was abandoned in the course of the 20th century.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

ʾAzdī, *ʾAsmāʾ* = ʾAbū l-Fatḥ Muḥammad ibn al-Ḥusayn al-ʾAzdī al-Mawṣilī, *ʾAsmāʾ man yuʿrafu bi-kunyatihi min ʾaṣḥāb Rasūl Allāh*. Ed. ʾIqbāl ʾAḥmad ibn Muḥammad ʾIsḥāq Baskūharī. Bombay: Dār as-Salafiyya, 1989.

——, *Muštabih* = ʾAbū Muḥammad ʿAbd al-Ġanī ibn Saʿīd al-ʾAzdī, *Kitāb muštabih an-nisba*. Cairo: Maktabat aṯ-Ṯaqāfa ad-Dīniyya, 2001.

Ibn ʿAqīl, *Šarḥ* = Bahāʾ ad-Dīn ʾAbū Muḥammad ʿAbdallāh ibn ʿAbd ar-Raḥmān Ibn ʿAqīl al-ʿAqīlī al-Hamḏānī al-Miṣrī, *Šarḥ Ibn ʿAqīl ʿalā ʾAlfiyyat Ibn Mālik*. Ed. Ḥ. al-Fāxūrī. Beirut: Dār al-Jīl, 1997.

Ibn al-ʾAṯīr, *Lubāb* = ʿIzz ad-Dīn ʾAbū l-Ḥasan ʿAlī ibn Muḥammad Ibn al-ʾAṯīr aš-Šaybānī al-Jazarī aš-Šāfiʿī, *al-Lubāb fī tahḏīb al-ʾansāb*. 3 vols. Ed. Ḥusām ad-Dīn al-Qudsī. Cairo: Maktabat al-Qudsī, 1357–1369/1938–1949. (Repr. Beirut: Dār Ṣādir, n.d. [1972].)

Ibn Ḥajar, *Muʿjam* = Šihāb ad-Dīn ʾAḥmad ibn ʿAlī Ibn Ḥajar al-ʿAsqalānī, *al-Majmaʿ al-muʾassis*

li-l-muʿjam al-mufahris. Ed. Yūsuf ʿAbdallāh al-Marʿašlī. Beirut: Dār al-Maʿrifa, 1994.
Jawzjānī, *ʾAḥwāl* = ʾAbū ʾIsḥāq ʾIbrāhīm ibn Yaʿqūb al-Jawzjānī, *ʾAḥwāl ar-rijāl*. Ed. Ṣubḥī al-Badrī as-Sāmarrāʾī. Beirut: Muʾassasat ar-Risāla, 1985.
Maqrīzī, *ʿUqūd* = Taqī d-Dīn ʾAḥmad ibn ʿAlī al-Maqrīzī, *Durar al-ʿuqūd al-farīda fī tarājim al-ʾaʿyān al-mufīda*. Ed. Maḥmūd al-Jalīlī. Beirut: Dār al-Ġarb al-ʾIslāmī, 2002.
Qaysarānī, *ʾAnsāb* = ʾAbū l-Faḍl Muḥammad ibn Ṭāhir Ibn al-Qaysarānī, *al-ʾAnsāb al-muttafiqa. Homonyma inter nomina relativa, auctore Abu'l-Fadhl Mohammed b. Táhir al-Makdisí, vulgo dicto Ibno'l-Kaisarání, quae cum appendice Abu Musae Ispahanensis e codd. Leyd. et Berolin.* ed. Pieter de Jong, 2–166. Leiden: E.J. Brill, 1865. (Facsimile).
Ṣafadī, *Wāfī* = Ṣalāḥ ad-Dīn Xalīl ibn ʾAybak aṣ-Ṣafadī, *Kitāb al-wāfī bi-l-wafayāt*. Ed. Ramaḍān ʿAbd at-Tawwāb. Wiesbaden: F. Steiner, 1979.
Samʿānī, *ʾAnsāb* = Tāj ad-Dīn ʾAbū Saʿd ʿAbd al-Karīm ibn Muḥammad ibn al-Manṣūr at-Tamīmī al-Marūzī as-Samʿānī aš-Šāfiʿī, *al-ʾAnsāb*. 4 vols. Ed. ʿAbdallāh ʿUmar al-Bārūdī. Beirut: Dār al-Jinān, 1988.
Saxāwī, *Ḍawʾ* = Šams ad-Dīn Muḥammad ibn ʿAbd ar-Raḥmān as-Saxāwī, *aḍ-Ḍawʾ al-lāmiʿ li-ʾahl al-qarn at-tāsiʿ*. Beirut: Dar al-Jīl, 1992.
Šawkānī, *Badr* = Muḥammad ibn ʿAlī aš-Šawkānī, *al-Badr aṭ-ṭāliʿ bi-maḥāsin min baʿd al-qarn ar-rābiʿ*. Cairo: Maktabat Ibn Taymiyya, 1979.
Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān ibn Qanbar Sībawayhi, *al-Kitāb*, III. Ed. ʿAbd as-Salām Muḥammad Hārūn. Beirut: Dār al-Jīl, 1973.
Suyūṭī, *Lubb al-lubāb* = Jalāl ad-Dīn ʾAbū l-Faḍl ʿAbd ar-Raḥmān ibn ʾAbī Bakr al-Xuḍayrī aš-Šāfiʿī as-Suyūṭī, *Lubb al-lubāb fī taḥrīr al-ʾansāb*. Ed. Petrus Johannes Veth. Leiden: Luchtmans, 1840.

SECONDARY SOURCES
ʾAfġānī, Saʿīd. 1971. *al-Mūjaz fī qawāʿid al-luġa al-ʿarabiyya: an-Nisba wa-ʾaḥkāmuhā*. Available at: www.islamguiden.com/arabi/m_a_r_27.htm. (Retrieved 23 October 2006.)
Ayalon, David. 1975. "Names, titles and 'nisbas' of the Mamlūks". *Israel Oriental Studies* 5.189–232.
Caetani, Leone and Giuseppe Gabrieli. 1915. *Onomasticon arabicum*. Rome: Casa Editrice Italiana.
Fierro, Maribel. 2004. "La nisba al-ʾAnṣārī en al-Andalus y el cadí Mundir b. Saʿīd". *Al-Qanṭara* 25.233–237.
Galley, Micheline. 2003. "Note sur l'onomastique des narrations populaires au Maghreb: Le cas de la geste hilalienne". *Mélanges David Cohen*, ed. Jérôme Lentin and Antoine Lonnet, 255–263. Paris: Maisonneuve et Larose.
Garcin de Tassy, Joseph Héliodore. 1854. *Mémoire sur les noms propres et les titres musulmans*. Paris: Imprimerie Royale.
Gilliot, Claude. 2002. "Prosopography in Islam: An essay of classification". *Medieval Prosopography* (Special Issue, *Arabic-Islamic culture*, ed. Manuela Marín) 23.19–54.
Niʿma, Fuʾād. 1973. *Mulaxxaṣ qawāʿid al-luġa al-ʿarabiyya*. 18th ed. Cairo: al-Maktab al-ʿIlmī li-t-Taʾlīf wa-t-Tarjama.
Schimmel, Annemarie. 1989. *Islamic names*. Edinburgh: Edinburgh University Press.
Sublet, Jacqueline. 1991. *Le voile du nom: Essai sur le nom propre arabe*. Paris: Presses Universitaires de France.
—— 1993. "*Nisba*". *Encyclopaedia of Islam* VIII, 53–56. 2nd ed. Leiden: E.J. Brill.
—— 1999. *Ḥiṣn al-ism: Qirāʾāt fī l-ʾasmāʾ al-ʿarabiyya*. Trans. Salīm Muḥammad Barakāt. Damascus: al-Maʿhad al-ʿIlmī al-Firansī li-d-Dirāsāt al-ʿArabiyya.
Versteegh, Kees. 1997. *The Arabic linguistic tradition*. London and New York: Routledge.

JEAN N. DRUEL
RENÉ-VINCENT DU GRANDLAUNAY
(Institut Dominicain d'Etudes Orientales, Cairo)

Nominal Clauses

Nominal clauses in Arabic (*jumal ismiyya*) are clauses in which the first constituent is a nominal expression (*mubtadaʾ* 'that which is begun with, inchoative'; → *ibtidāʾ*), of which the remaining subconstituent of the clause (→ *xabar* 'news, announcement') is predicated. *Mubtadaʾ* and *xabar* are translated here as 'initial Noun Phrase (NP)' and 'report', respectively. The report constituent has two basic types. The first is a complete → 'verbal clause' containing a pronoun *rābiṭ* 'binder, connector' which 'resumes' or is bound by the initial NP, as illustrated in (1) and (8) below. The second type is headed by a nonverbal → predicate, such as an adjective (13a), participle (13b), preposition (13c), or noun phrase (13d). The initial NP must also be definite or 'specific'. Nominal clauses with verbal reports have two subtypes: those in which the initial NP is construed as the subject of the verb (1), and those in which the initial NP is construed as an object or oblique argument of the report (8), or possibly as an argument of a more deeply embedded clause (10).

Western grammarians have studied whether initial NPs in these different kinds of nominal clauses can be analyzed uniformly as either → subjects or → topics, or whether the different types require different analyses. The consensus seems to be that in some nominal clauses the initial NP is a topic, while in others it is a subject. However, certain recent studies suggest that at least some initial NPs that appear to be topics may also have subject properties.

1. Nominal clauses with verbal reports

The first major category of nominal clauses are those in which the report constituent is a verbal clause. There are two subtypes: those in which the initial NP is construed as the subject of the verb, and those in which it is construed as a more oblique argument.

1.1. *Nominal clauses with initial subjects*

The first subtype of verbal report takes the initial NP as its subject ('initial subject'):

(1) a. *al-ʾawlādu* *laʿabū*
the-boys.mp-Nom play.Perf.3mp
kurat-a *l-qadam*
ball-Acc the-foot
'The boys played football'
b. *al-banāt-u* *yadrusna*
the-girls.fs-Nom study.Imperf.3fp.Ind
l-qawāʿid-a
the-grammar
'The girls study grammar'

Some grammarians analyze initial subjects as topics occupying a *left-dislocated* position (Bakir 1980; Fassi-Fehri 1988, 1991, 1993; Ouhalla 1988, 1991, 1994; Demirdache 1991; Plunkett 1993, 1996). Others argue that initial subjects are not topics but rather preverbal subjects (Mohammad 1988, 1990, 2000; Bolotin 1995; Benmamoun 2000; Bahloul and Harbert 2002).

Four generalizations are at stake:

i. Like other initial NPs, the initial subject must be semantically definite or specific;
ii. The dependency between an initial subject and the subject position of the report is more local than are the dependencies between nonsubject initial NPs and their *rābiṭ*-pronouns;
iii. The initial subjects, like other initial NPs, must precede *fronted* constituents (→ fronting) such as question words or prepositional phrases;
iv. In Standard Arabic, initial subjects control full → agreement on the verb, while a postverbal subject controls only gender agreement.

The two approaches focus on different subsets of (i)–(iv). Treating initial subjects as left-dislocated topics accounts for generalizations (i), (iii), and (iv). According to this approach, the agreement marking on the verb, and in particular nonsingular number marking, is an incorporated pronoun which acts as the *rābiṭ* for the initial subject. This captures the distinction made in generalization (iv), a well-known contrast in the 3rd person paradigm between initial subjects and postverbal subjects (→ agreement). Postverbal subjects control only agreement in gender, the verb being marked in the singular (2), while initial subjects control full agreement in gender and number (3).

(2) a. *jāʾat* *al-banāt-u*
come.Perf.3fs the-girls.fp-Nom
'The girls came'
b. **jiʾna* *l-banāt-u*
come.Perf.3fp the-girls.fp-Nom

(3) a. *al-banāt-u* *jiʾna*
the-girls.fp-Nom come.Perf.3fp
'(As for) the girls, they came'
b. **al-banāt-u* *jāʾat*
the-girls.fp-Nom come.Perf.3fs

The analysis also captures the generalization in point (iii) by predicting data reported by Bakir (1980) and Fassi-Fehri (1982, cited by Plunkett 1993:243–244) which show that the initial subject must precede fronted elements (see also Lalami 1995 for Moroccan Arabic):

(4) a. *aṭ-ṭullābu,* *matā*
the-students.mp-Nom when
ḏahabū *ʾila l-ʿirāq?*
go.Perf.3mp to the-Iraq
'The students, when did they go to Iraq?' (Bakir 1980:128)
b. **matā ṭ-ṭullābu ḏahabū ʾila l-ʿirāq?*

If left-dislocated NPs occupy a position higher than and to the left of the position occupied by fronted constituents, then initial subjects are left-dislocated. Likewise, generalization (ii) is easily captured by the left-dislocation analysis, because a simple constraint can be stated for all initial NPs requiring them to be definite or specific (→ specificity).

The competing approach to initial subjects treats them as subjects proper because they behave more like English subjects than English topics. Mohammad (2000) argues that point (ii) – that initial subjects are in a more local dependency with their predicates than other initial NPs – shows that a different kind of dependency holds between an initial subject and the subject position in which it is construed than holds between a nonsubject initial NP and its *rābiṭ*.

In (5), for example, *al-ijtimāʕ* 'the meeting' is construed as the subject of the embedded verb *ḥaṣala* 'took place'. In (5a), it is an initial subject taking the clause headed by *ḥaṣala* as its report, while in (5b) it is the postverbal subject position. This complement clause is itself embedded within the interrogative clause headed by *tatasāʾalu* 'wonders':

(5) *qālat* *zaynab-u*
say.Perf.3fs Zainab.fs-Nom
ʾinna-hā *tatasāʾalu*…
that-cl3fs wonder.Imperf.3fs.Ind
'Zainab said that she wonders…'
a. [*man* *iddaʕā* *ʾanna*
who claim.Perf.3ms that
l-ijtimāʕ-a *ḥaṣala*
the-meeting.ms-Acc occur.Perf.3ms
fī *baġdād-a*]
in Baghdad-Gen
'…who claimed that the meeting took place in Baghdad'
b. [*man* *iddaʕā* *ʾanna* *fī*
who claim.Perf.3ms that in
baġdād-a *ḥaṣala*
Baghdad-Gen occur.Perf.3ms
l-ijtimāʕ-a]
the-meeting.ms-Acc

The examples in (6a, 6b) are variations on (5a, 5b), with left-dislocation of *al-ijtimāʕ*. (6a) is acceptable, while (6b) is not. The dependency in (6a) between *al-ijtimāʕ-u* and the *rābiṭ*-pronoun crosses the WH-island headed by *tatasāʾalu*. In contrast, the unacceptable (6b) associates *al-ijtimāʕ-u* with the postverbal subject position of *ḥaṣala*, the unacceptability being due to the dependency crossing the WH-island boundary:

(6) *al-ijtimāʕ-u*$_i$ *qālat*
the-meeting.ms-Nom say.Perf.3fs
zaynab-u *ʾinna-hā*
Zainab.fs-Nom that-cl3fs
tatasāʾalu
wonder.Imperf.3fs.Ind
a. [*man* *iddaʕā* *ʾanna-hu*
who claim.Perf.3ms that-cl3ms
ḥaṣala *fī* *baġdād-a*]
happen.Perf.3ms in Baghdad-Gen
b. [*man* *iddaʕā* *ʾanna*
who claim.Perf.3ms that
ḥaṣalat$_i$] *fī baġdād-a*
happen.Perf.3ms in Baghdad-Gen
'The meeting, Zeinab said that she wonders who claimed that it took place in Baghdad'

If agreement marking on the verb functions as a *rābiṭ*-pronoun, as claimed by the left-dislocation approach, then (6b) should be acceptable, because it should be the same kind of island-crossing dependency as seen in (6a). Mohammad (2000) takes the unacceptability of the example to indicate that initial subjects are not left-dislocated from the postverbal subject position but rather are associated with it by NP-raising.

This conclusion is supported by data from Bakir (1980:162–163), showing that a sentence with two initial NPs of the same number and gender can only be construed with the second initial NP as the subject of the verb:

(7) *fāṭimat-u* *hind-un*
Fatima.fs-Nom Hind.fs-Nom
raʾat-hā
see.Perf.3fs-cl3fs
'Fatima, Hind saw her'
*'Hind, Fatima saw her'
(Bakir 1980:163)

The potential ambiguity caused by the identical gender and number values for the two initial NPs is resolved by interpreting the second initial NP as a nondislocated subject.

1.2. *Topicalization and nominal clauses*

In the second subtype of verbal report constituents, the initial NP is associated with a resumptive pronoun in a more oblique position (→ resumption). Clauses of this type are often analyzed as involving left-dislocation of the initial NP. This means that the initial NP is 'dislocated' to a position outside of the clause where it fills a discourse role (such as → topic or → focus), rather than an argument or → thematic role (such as *subject, object*, or *oblique*). It is 'linked' or 'bound' (→ binding) to a *rābiṭ*-pronoun occupying an object or oblique position within the report constituent:

(8) a. *hind-un,* *samiʿa-hā*
Hind.fs-Nom hear.Perf.3ms-cl3fs
muḥammad-un
Mohammad.ms-Nom
'Hind, Mohammad heard her'
b. *aš-šāriʿ-u,* *qābaltu*
the-street.ms-nom meet.Perf.1s
sālim-an *fī-hi*
Salim.ms-Acc in-cl3ms
'The street, I meet Salim on it'
c. *fāṭimat-u,* *ištaraytu*
Fatima.fs-Nom buy.Perf.1s
kitāb-a-hā *l-ʾams-a*
book-Acc-cl3fs the-yesterday-Acc
'Fatima, I bought her book yesterday'
d. *al-muʿallimūna,* *ḏahabū*
the-teachers.mp.Nom go.Perf.3mp
ʾilā *buyūt-i-him*
to houses-Gen-cl3mp
'The teachers, they went to their houses'
(Bakir 1980:60–61)

Clauses can contain more than one left-dislocated NP, showing that one nominal clause can take another nominal clause as its report constituent:

(9) a. *hind-un,* *sālim-un,*
Hind.fs-Nom Salim.ms-Nom
taḍribu-hu
beat.Imperf.3fs-cl3ms
'Hind, Salim, she beats him'
(Bakir 1980:165)
b. *muḥammad-un,*
Muhammad.ms-Nom
as-sayyārat-u, *ʾuxt-u-hu,*
the-car.fs-Nom sister.fs-Nom-cl3ms
bāʿa-ha *la-hā*
sell.Perf.3ms-cl3fs to-cl3fs
'Mohammad, the car, his sister, he sold it to her'
(Bakir 1980:169)

Left-dislocation of the initial NP has been analyzed as insertion of the initial NP into the left-dislocated position (Bakir 1980; Ayoub 1981; Fassi-Fehri 1982, 1988, 1991, 1993; Plunkett 1993; Lalami 1995). The → binding relationship between the initial NP and its *rābiṭ*-pronoun is then established by a co-indexing rule. This is motivated by the insensitivity of the relationship between the initial NP and the *rābiṭ* to syntactic island constraints (Ross 1967), such as the *Complex-NP Constraint* (10a) or the WH-*Island Constraint* (10b):

(10) a. *zayd-un,* *ʾaḏ̣unnu*
Zaid.ms-Nom believe.Imperf.1s.Ind
r-rajul-a *[llaḏī*
the-man.ms-Nom Rel.ms
intaqada-hu
criticize.Perf.3ms-cl3ms
intaḥara]
commit:suicide.Perf.3ms
'(As for) Zaid, I believe that the man who criticized him committed suicide'
(Mohammad 2000:68)
b. *al-walad-u,* *tasāʾaltu*
the-boy.ms-Nom wonder.Perf.1s
[man *raʾā-hu]*
who see.Perf.3ms-cl3ms
'(As for) the boy, I wondered who saw him'
(Mohammad 2000:67)

Island-sensitivity is widely assumed to be diagnostic of syntactic movement, so the dependencies in (8) show that the initial NPs are not associated with their *rābiṭ*-pronouns by movement and so must be base generated in the left-dislocated position.

Evidence from Lebanese Arabic, provided by Aoun and Benmamoun (1998), Aoun, a.o. (2001), and Aoun and Li (2003), shows that initial NPs that are in island-violating dependencies should be distinguished from those that are not. This is because an initial NP in an island-compliant dependency can be interpreted in the position occupied by its *rābiṭ*-pronoun. To illustrate, in (11) the possessive pronoun inside the initial NP *talmīz-a* 'her student' is interpreted as bound by the quantificational NP *kəll mʿallme* 'every teacher'.

(11) *talmīz-a* *š-šiṭān*
student.ms-cl3fs the-naughty.ms
b-taʿrfo *ʾənno*
Ind-know.Imperf.2p that
kəll *mʿallmet*
every teacher.fs
ʾāʾṣət-o
punish.Perf.3fs-cl3ms

'Her naughty student, you know that every teacher$_i$ punished him'
'You know that every teacher punished her naughty student'
(Aoun and Benmamoun 1998:580)

This is an example of syntactic reconstruction, a process which is also taken to be diagnostic of syntactic movement.

By contrast, when the initial NP is separated from its *rābiṭ*-pronoun by a syntactic island, the reconstructed interpretation is not available and the pronoun must be assigned an interpretation from the context, as in (12).

(12) *talmīz-a* *š-šiṭān*
student.ms-cl3fs the-naughty.ms
fallayto *ʾablma* *kəll*
leave.Perf.2p before every
mʿallme *tʾāʾaṣət-o*
teacher.fs punish.Imperf.3fs-cl3ms
'Her$_j$/*$_i$ naughty student, you know that every teacher$_i$ punished him'
(Aoun and Benmamoun 1998:580)

The conclusions here are that (i) at least some initial NPs are moved into the initial position and (ii) the resumptive (→ resumption) pronouns they bind are not independent constituents but rather pronounced traces of movement.

2. COPULAR CLAUSES

The second major category of nominal clauses are those in which the report constituent is headed by nonverbal predicates, such as adjectives (13a), participles (13b), prepositions (13c), and nouns (13d).

(13) a. *ʿĪsā* *ṭawīlu* *ṣ-ṣabr*
Isa.ms long.ms-Nom the-patience
'Isa is very patient'
b. *xalīl-un*
Khalil.ms-Nom
nāʾim-un *fī*
sleep.Act.Part.ms-Nom in
l-bayt-i
the-house-Gen
'Khalil is sleeping in the house'
c. *jareš* *fī* *šamāl-i* *ġarb-i*
Jaresh in north-Gen west-Gen
l-ʾurdunn
the-Jordan
'Jaresh is in the northwest of Jordan'
d. *dayzi l-ʾamīr* *kātibat-un*
Daisy al-Amir writer.fs-Nom
ʿirāqiyyat-un
Iraqi.fs-Nom
'Daisy al-Amir is an Iraqi writer'

Clauses of this type are often called copular clauses.

In copular clauses, the report constituents are lexical or simplex predicates lacking a subject argument. For this reason, the initial NPs are usually analyzed as subjects because they saturate the predicate (Jelinek 1981, 1983; Fassi-Fehri 1982; Eid 1991; Plunkett 1993; Mohammad 1998; Eisele 1999, who treats them as left-dislocated). Furthermore, as noted by Fassi-Fehri (1982, cited by Plunkett 1993), unlike initial subjects with verbal report constituents, the initial NP in a nonverbal predication can follow a question word, as in (14).

(14) a. *kayfa* *l-jaww-u* *fī*
how the-weather.ms-Nom in
š-šitāʾ-i?
the-winter-Gen
'How is the weather in the winter?'
(Plunkett 1993:245)
b. **al-jaww-u kayfa fī š-šitāʾ-i?*

Nonetheless, initial NPs in copular clauses must still be definite or specific (→ specificity).

An important subtype of copular clauses are those in which the report is a definite noun phrase. These are referred to as equational clauses and require that the initial NP and the report be linked by a 'pronoun of separation' (*ḍamīr al-faṣl*), also known as a 'copular pronoun', as in (15).

(15) a. *marwān-un* *huwa*
Marwan.ms-Nom cl.3ms
ṭ-ṭawīl-u
the-tall.ms-Nom
'Marwan is the tall one'
b. *al-ḥaqīqat-u* *hiya* *ʾanna-nī*
the-truth.fs-Nom cl.fs that-cl1s
lam *ʾaqraʾ*
not.Past read.Juss.1s
al-wājib-a
the-assignment-Acc
'The truth is that I did not read the assignment'

The pronoun of separation is required to make it clear that the report constituent is not an adjectival or appositive modifier (16a) and that the whole string has clausal structure (16b).

(16) a. *ʾaḥmad-u* *ṭ-ṭawīl-u*
Ahmad.ms-Nom the-tall.ms-Nom
'Ahmad the tall, tall Ahmad'
b. *ʾaḥmad-u* *huwa*
Ahmad cl.3ms
ṭ-ṭawīl-u
the-tall.ms-Nom
'Ahmad is the tall one'

The copular pronoun has been analyzed as an auxiliary, in the same category as verbs like *kāna* 'to be' (Jelinek 1982; Eid 1983, 1991; Awwad 1987; Shlonsky 1997), and, in some dialects, as part of the → negation morpheme *miš* (Eid 1991, 1993). Eid (1983, 1991, 1993) and Awwad (1987), for example, note that copular pronouns can host the *ma-…-š* negation circumfix, which attaches primarily to verbs or auxiliaries. This suggests that the pronoun itself is an auxiliary. A copular pronoun hosting negation (called a 'pronoun of negation') also licenses → pro-drop, a property shared with verbal stems but lacking for the simplex negation *miš*, as in (17a, 17b) from Egyptian Arabic.

(17) a. *(ana)* *ma-nī-š* *sakna* *hina*
I.1s not-cl1s-Neg living.fs here
'I am not living here'
b. **(ana)* *miš* *sakna* *hina*
I.1s not living.fs here
'I am not living here'
Egyptian Arabic (Eid 1991:51)

Likewise, Eid (1991) shows that the copular pronoun can be used to contrast an auxiliary, as in (18).

(18) *ʿali* *ma-kan-š*
Ali.ms not-be.Perf.3ms-Neg
il-mudarris, *ʿali* *huwwa* *il-mudarris*
the-teacher Ali.ms he the-teacher
'Ali WASN'T the teacher, Ali IS the teacher'
(Eid 1991:49)

The first clause is in the past tense, and its truth value is being negated and corrected by the assertion of present tense in the second conjunct. Therefore, what is being contrasted is the information expressed by the tense-nodes in the clauses and, therefore, by the auxiliary element. Because the copular pronoun is providing this contrast, it appears to be in the auxiliary position.

3. Arabic as a topic-prominent language

Recently, it has been argued that initial NPs in non-island-compliant dependencies are in fact a kind of subject called 'broad subject' (Doron 1996; Doron and Heycock 1999; Heycock and Doron 2003; Alexopoulou a.o. 2004). A broad subject is the subject of a clause in which the predicate is itself a clause. The claim is that Arabic clausal predications can be based on both lexical predicates and derived, relative-clause-like predicates. It differs in this respect from languages like English, in which only lexical predicates may be the basis of a clausal predication.

The evidence for this claim is that broad subjects share four properties with the subjects of lexical predicates that distinguish them from initial NPs which are interpreted as topics. The first property is that a report constituent can be conjoined with a lexical predicate:

(19) *sayyārat-ī,* *[lawn-u-hā*
car.fs-cl1s color.ms-Nom-cl3fs
zahiyy-un] *wa-[maftūḥat-un*
bright.ms and-open.fs-Nom
min *al-ʾaʿlā]*
from the-above
'My car, its color is bright and [it is] a convertible'
'My car is brightly colored and a convertible'
(Doron and Heycock 1999:73)

The Coordinate Structure Constraint (Ross 1967) predicts that a constituent associated with a particular grammatical function in one conjunct must be associated with the same grammatical function in the other. Since the second conjunct in (19) is an incomplete predicate if it occurs by itself as in (20), the initial NP in (19) must be construed as its subject.

(20) a. *sayyārat-ī* *maftūḥat-un* *min*
car.fs-cl1s open.fs-Nom from
al-ʾaʿlā
the-above
'My car is open from above' or 'My car is a convertible'
b. **maftūḥat-un min al-ʾaʿlā*

According to the Coordinate Structure Constraint, if the initial NP (19) is the grammatical subject of the second conjunct, it must be the grammatical subject of the first conjunct as well.

The second property is that initial NPs can be semantically nonreferring, while topics must be referential. For example, the initial NP in (21) is a negative polarity item, and hence a nonreferring expression:

(21) *ʾayy-u* *muxrij-in*
any-Nom director.ms-Gen
ʾajnabiyy-in, *lā*
foreign.ms-Gen not
nuʿriḍu *ʾaflām-a-hu*
show.Imperf.1p.Ind films-Acc-cl3ms
min dūni *tarjamat-in*
without translation-Gen
'Any foreign director, we don't show his films without subtitles'
'We don't show any foreign director's films without subtitles'
(Doron and Heycock 1999:84)

The third property is that broad subjects appear to be used inside embedded clauses while topics cannot be:

(22) *ʾaḏ̣unn-u d-dār-a* *lawn-u-ha*
consider.Imperf.1s.Ind the-house.fs
zahiyy-un
color.ms-cl3fs bright.ms
'I consider the house to be brightly colored'

The last property is that the initial NP is not associated with a particular discourse function such as topic or focus. Instead, the initial NP can fill either of these functions, a point which Bakir (1980:129–130) also makes about initial NPs filling the subject function:

(23) Lebanese Arabic
a. *mīn* *šaʿar-a* *ṭawīl*
who hair.ms-cl3fs long.ms
'Whose hair is long?'
b. *rana* *šaʿar-a* *ṭawīl*
Rana.fs hair-cl3fs long.ms
'Rana's hair is long'
(Alexopoulou a.o. 2004:337)

Doron and Heycock (1999), Alexopoulou a.o. (2004), and Heycock and Doron (2003) conclude that Arabic is a discourse configurational language, meaning that it uses the subject function to encode discourse relations in addition to thematic relations (see also Brustad 2000, Chap. 10).

BIBLIOGRAPHICAL REFERENCES

Abdul-Raof, Hussein. 1999. *Subject, theme, and agent in Modern Standard Arabic.* Richmond: Curzon.

Alexopoulou, Theodora, Edit Doron, and Caroline Heycock. 2004. "Broad subjects and clitic left dislocation". *Peripheries: Syntactic edges and their effects*, ed. David Adger, Cécile de Cat, and George Tsoulas, 328–358. Oxford: Oxford University Press.

Aoun, Joseph and Benmamoun, Elabbas. 1998. "Minimality, reconstruction, and PF movement". *Linguistic Inquiry* 29.59–59.

Aoun, Joseph, Lina Choueiri, and Norbert Hornstein. 2001. "Resumption, movement, and derivational economy". *Linguistic Inquiry* 32.371–40.

Aoun, Joseph and Yen-hui Audrey Li. 2003. *Essays on the representational and derivational nature of grammar: The diversity of WH-constructions.* Cambridge, Mass.: MIT Press.

Awwad, Mohammad Amin. 1987. "Free and bound pronouns as verbs in rural Palestinian colloquial Arabic". *Journal of Arabic Linguistics* 16.108–118.

Ayoub, Georgine. 1981. *Structure de la phrase en arabe standard.* Ph.D. diss., University of Paris-3.

Badawi, Elsaid, Michael G. Carter, and Adrian Gully. 2004. *Modern Written Arabic: A comprehensive grammar.* London: Routledge.

Bahloul, Maher and Wayne Harbert. 2002. "Postverbal subjects in Arabic and the theory of agreement". *Themes in Arabic and Hebrew syntax*, ed. Jamal Ouhalla and Ur Shlonsky, 45–70. Dordrecht: Kluwer Academic Publishers.

Bakir, Murtada. 1980. *Aspects of clause structure in Arabic.* Bloomington: Indiana University Linguistics Club.

Bolotin, Naomi. 1995. "Arabic and parametric VSO agreement". Eid and Parkinson (1995:7–27).

Brustad, Kristen E. 2000. *The syntax of spoken Arabic: A comparative study of Moroccan, Egyptian, Syrian, and Kuwaiti dialects.* Washington, D.C.: Georgetown University Press.

Demirdache, Hamida. 1991. *Resumptive chains in restrictive relatives, appositives and dislocation structures.* Ph.D. diss., Massachusetts Institute of Technology.

Doron, Edit. 1996. "The predicate in Arabic". *Studies in Afroasiatic linguistics*, ed. Jacqueline LeCarme, Jean Lowenstamm, and Ur Shlonsky, 77–87. Leiden: Holland Academic Graphics.

——, and Caroline Heycock 1999. "Filling and licensing multiple specifiers". *Specifiers*, ed. David

Adger, Susan Pintzuk, Bernadette Plunkett, and George Tsoulas, 69–89. Oxford: Oxford University Press.

Eid, Mushira. 1983. “The copula function of pronouns”. *Lingua* 59.197–207.

——. 1991. “Verbless sentences in Arabic and Hebrew”. *Perspectives on Arabic linguistics*, III, ed. Bernard Comrie and Mushira Eid, 31-61. Amsterdam and Philadelpha: J. Benjamins.

——. 1993. “Negation and predicate heads”. *Principles and prediction: The analysis of natural language*, ed. Mushira Eid and Gregory K. Iverson, 135–152. Amsterdam and Philadelphia: J. Benjamins.

——, and Dilworth Parkinson (eds). 1995. *Perspectives on Arabic linguistics, IX*. Amsterdam and Philadelphia: J. Benjamins.

Eisele, John. 1999. *Arabic verbs in time: Tense and aspect in Cairene Arabic*. Wiesbaden: O. Harrassowitz.

Fassi-Fehri, Abdelkader. 1982. *Linguistique arabe: Forme et interprétation*. Rabat: Publications de la Faculté des Lettres et des Sciences Humaines.

——. 1988. “Agreement in Arabic, binding and coherence”. *Agreement in natural language*, ed. Michael Barlow and Charles Ferguson, 107–158. Palo Alto, Calif.: Center for the Study of Language and Information.

——. 1991. “Generalized IP structure, case, and VS word order”. *MIT working papers in linguistics, 10*, ed. Itziar Laka and Anoop Mahajan, 75–111. Cambridge, Mass.: MIT Working Papers in Linguistics.

——. 1993. *Issues in the structure of Arabic clauses and words*. Dordrecht: Kluwer Academic Publishers.

Heycock, Caroline and Edit Doron. 2003. “Categorical subjects”. *Gengo Kenkyu* 123.95–135.

Jelinek, Eloise. 1981. *On defining categories: AUX and PREDICATE in Egyptian colloquial Arabic*. Ph.D. diss., University of Arizona.

——. 1982. “Predicate as a universal syntactic category”. *Coyote papers: Studies in Arabic, Basque, English, Japanese, Navajo, and Papago*, ed. Thomas G. Larsen, 23–46. Tucson: University of Arizona Linguistics Circle.

——. 1983. “Person-subject marking in AUX in Egyptian Arabic”. *Linguistic categories: Auxiliaries and related puzzles*, ed. Richard Heny and Barry Richards. 29–49. Dordrecht: Reidel.

Lalami, Laïla. 1995. “Clitic left dislocation in Moroccan Arabic”. Eid and Parkinson (1995:115–129).

Mohammad, Mohammad. 1998. *The syntax of indefinite subjects in equative sentences in Palestinian Arabic*. Ms., University of Florida, Gainesville. Available from the Semitic Linguitics Research Archive: www.usc.edu/dept/LAS/linguistics/semitic/pdf/sisespa.pdf.

——. 2000. *Word order, agreement, and pronominalization in Standard and Palestinian Arabic*. Amsterdam and Philadelphia: J. Benjamins.

Ouhalla, Jamal. 1991. *Functional categories and parametric variation*. London: Routledge.

——. 1994. “Verb movement and word order in Arabic”. *Verb movement*, ed. David Lightfoot and Norbert Hornstein, 41–72. Cambridge: Cambridge University Press.

Plunkett, Bernadette. 1993. “The position of subjects in Modern Standard Arabic”. *Perspectives on Arabic linguistics*, V, ed. Mushira Eid and Clive Holes, 231–259. Amsterdam and Philadelphia: J. Benjamins.

——. 1996. “Binding in Arabic and the lexical clause”. *Perspectives on Arabic Linguistics*, VIII, ed. Mushira Eid, 85–98. Amsterdam and Philadelphia: J. Benjamins.

Ross, John R. 1967. *Constraints on variables in syntax*. Ph.D. diss., Massachusetts Institute of Technology.

Shlonsky, Ur. 1997. *Clause structure and word order in Hebrew and Arabic: An essay in comparative Semitic syntax*. New York: Oxford University Press.

FREDERICK HOYT
(University of Texas at Austin)

Nominalization

Event nominals, also known as verbal nouns, derived nominals, and action nominalizations, are nouns that refer to events, e.g. *suqūṭ* ‘falling’, *wuṣūl* ‘arrival’, *tadmīr* ‘destruction’ (→ *maṣdar*). This particular class of nouns has been subject to much discussion because it sheds doubt on the traditional distinction between nouns and verbs. The controversy stems from the fact that although event nominals display external syntactic properties typical of nouns, they seem to have internal verbal syntax. For example, event nominals occur in argument positions such as subjects and objects, yet they have argument structures, and their objects can be marked for the accusative case in Standard Arabic.

There are two main approaches to analyzing the mixed verbo-nominal properties of event nominals. The first is a derivational account based on the assumption that noun phrases headed by event nominals have underlying verbs, verb phrases, or complete clauses, and that there is a syntactic process of nominalization, namely V-raising to N, that provides the nominal output (Lees 1960; Wise 1975; Levi 1978; Fassi Fehri 1990, 1993; Hazout 1991, 1995; Borer 1995; Stepanov 1997; Rozwadowska 1997; Emonds 2000; Fu a.o. 2001; Ogawa 2001). The other approach is the Lexicalist Hypothesis, which treats event nominals as lexically specified nouns with argument structures that must be satisfied (Chomsky 1970; Lebeaux 1986; Zubizarreta 1987; Grimshaw 1990; Siloni 1997).

The motivation for treating event nominals as nouns is based on the observation that they have morphosyntactic and semantic properties that are typical of nouns. For example, they gram-matically occur in argument positions, e.g. as subjects, objects, and complements of prepositions, as in (1a), (1b), and (1c), respectively.

(1a) (Standard Arabic)
waqaʿa *infijār-un* *fī*
occurred-3ms explosion-Nom in
baġdād
Baghdad
'An explosion occurred in Baghdad'

(1b) (Standard Arabic)
intaḏ̣ar-tu *wuṣūl-a* *ḍ-ḍuyūf-i*
waited-1sg arrival-Acc the-guests-Gen
'I waited for the arrival of the guests'

(1c) (Standard Arabic)
waṣala-t *aš-šurṭa* *baʿda*
arrived-3fs the-police after
hurūb-i *l-liṣṣ-i*
escape-Gen the-thief-Gen
'The police arrived after the escape of the thief'

Occurring in argument positions is a necessary and sufficient characteristic of nouns, as they are the only forms that can be grammatically used in such positions. Moreover, event nominals carry case morphology according to their syntactic distribution, such as the nominative case in (1a), the accusative case in (1b), and the genitive case in (1c).

Other morphosyntactic properties that are restricted to noun phrases include the licensing of relative clauses, modification with adjectives, and heading → construct state constructions. Event nominals grammatically occur in these constructions, thus providing strong support for analyzing them as nouns. For example, the event nominal *al-hujūm* 'the attack' in (2a) is grammatically used as the head of a relative clause. In (2b), the nominal *ġiyāb* 'absence' is modified by the adjective *al-mutakarrir* 'repetitive' in a grammatical sequence. Finally, in (2c) the event nominal *ġurūb* 'setting' forms a construct state construction with the theme of the setting event.

(2a) (Standard Arabic)
al-hujūm *allaḏī* *tazāmana* *maʿa*
the-attack which coincided with
ziyārat-i *r-raʾīsi*
visit-Gen the-president-Gen
'The attack which coincided with the president's visit'

(2b) (Standard Arabic)
fuṣila *li-ġiyāb-i-hi*
fired.Pass.3ms for-absence-Gen-his
l-mutakarrir-i
the-repetitive-Gen
'He was fired because of his repeated absences'

(2c) (Standard Arabic)
taqābal-nā *baʿda* *ġurūb-i*
met-1pl after setting-Gen
š-šams-i
the-sun-Gen
'We met after the setting of the sun'

Among the semantic properties of event nominals that support analyzing them as nouns is the fact that they can be marked for definiteness, both on a specific reading, as in (3a), and on a generic reading, as in (3b). In (3a) the event nominal *iṭ-ṭalāʾ* 'the divorce' refers to a particular incident that was completed an hour before speech time, whereas in (3b) the same event nominal does not refer to any particular event but rather to the event type of getting divorced in general.

(3a) (Egyptian Arabic)
iṭ-ṭalāʾ *tamm* *min* *sāʿa*
the-divorce completed.3ms from hour
'The divorce was completed an hour ago'

(3b) (Egyptian Arabic)
iṭ-ṭalāʾ *b-yi-ʿmil* *mašākil*
the-divorce Imperf-3s-make problems
kitīra
many
'Divorce causes a lot of problems'

Finally, event nominals can be quantified, as in (4a), where the event nominal *taḥʾīʾ* 'interrogation' is bound by an existential quantifier, and in (4b), where the event nominal *iḥtirāq* 'burning' is bound by a universal quantifier.

(4a) (Egyptian Arabic)
fī *taḥʾīʾ*
there.is interrogation
šaġġāl *guwwa*
working.Act.Part inside
'There is an interrogation in progress inside'

(4b) (Standard Arabic)
kull *iḥtirāq* *ya-ntuju* *ʿan-hu*
every burning 3ms-result from-it
ṯāni ʾuksīd al-karbōn
carbon dioxide
'Every burning results in the release of carbon dioxide'

Although event nominals share the above-mentioned morphosyntactic and semantic features with other types of nouns, the existence of certain properties suggests analyzing them as a separate class. Generally, event nominals do not have broken plurals, and the ones that have regular feminine plurals do not co-occur with the complete set of the arguments associated with their corresponding verbs. Event nominals such as *suqūt* 'falling', *tadmīr* 'destruction', and *wuṣūl* 'arrival' have no broken or regular plural forms. In (5a), the event nominal *wuṣūl* 'arrival' denotes a single instance of arriving, and it licenses two arguments. The phrase in (5b), on the other hand, has two readings: a collective reading where the event nominal denotes a single event, if the guests arrive together, and a distributive reading where there are several events of arriving that occur at different times. Even on a distributive reading the event nominal cannot be pluralized, hence the ungrammaticality of (5c).

(5a) (Standard Arabic)
wuṣūl *munā* *l-bayt-a*
arrival Mona the-house-Acc
'Mona's arrival at the house'

(5b) (Standard Arabic)
wuṣūl *aḍ-ḍuyūf-i* *l-bayt-a*
arrival the-guests-Gen the-house-Acc
'the guests' arrival at the house'

(5c) (Standard Arabic)
**wuṣūl-āt* *aḍ-ḍuyūf-i*
arrival-pl the-guests-Gen
l-bayt-a
the-house-Acc
'the guests' arrival at the house'

Some Standard Arabic nouns with nominalization morphology have regular feminine plurals, e.g. *intixāb-āt* 'elections', *tarmīm-āt* 'renovations', *taʿzīz-āt* 'reinforcements'. These plural nouns do not refer to events, and they do not allow expressing the subject arguments of the verbs they are associated with. The singular event nominal *ʾiṣlāḥ* 'fixing' in (6a) refers to a particular event of fixing that took a long time, and it co-occurs with both the Agent and the Patient argument, the latter of which is licensed via the preposition *li-* 'for'. The plural form *ʾiṣlāḥāt* 'fixings' in (6b), on the other hand, is ungrammatical when used to refer to events, even though there are multiple agents and patients indicating that there were several events of fixing.

(6a) *istaġraqa* *ʾiṣlāḥ* *ʿalī* *li-s-sayyāra*
lasted.3ms fixing Ali for-the-car
waqt-an *ṭawīl-an*
time-Acc long-Acc
'Ali's fixing of the car took a long time'

(6b) **istaġraqa-t* *ʾiṣlāḥ-āt*
lasted-3fs fixing-pl.
al-fanniyy-īna *li-s-sayyārāt*
the-technicians-Gen for-the-cars
ʾawqāt-an *ṭawīl-at-an*
times-Acc long-fem.-Acc
'The technicians' fixing of the cars took long periods of time'

Another property that distinguishes event nominals from other types of nouns is the fact that event nominals do not occur in possessive constructions with *bitāʿ* in Egyptian Arabic or its equivalents in other dialects (→ analytic genitive). The participants of an event can be represented in construct state constructions headed by an event nominal, as in (7a) and (7b), but *bitāʿ* cannot be used regardless of which argument is represented, as indicated by the ungrammaticality of (8a) and (8b).

(7a) (Egyptian Arabic)
maḥaddiš *yi-ʿraf* *sabab*
nobody 3ms-know reason
intiḥār *rāgil* *il-ʾaʿmāl*
suicide man the-business
'Nobody knows the reason for the businessman's suicide'

(7b) *šurb* *il-ʾahwa* *ktīr* *miš*
drinking the-coffee a.lot not
kwayyis *ʿašān-ak*
good for-you
'Drinking a lot of coffee is not good for you'

(8a) **maḥaddiš* *yi-ʿraf* *sabab*
nobody 3ms-know reason
il-intiḥār *bitāʿ* *rāgil* *il-ʾaʿmāl*
the-suicide of man the-business

(8b) **iš-šurb* *bitāʿ* *il-ʾahwa* *ktīr*
the-drinking of the-coffee a.lot
miš *kwayyis* *ʿašān-ak*
not good for-you

Finally, event nominals have a special way of expressing constituent → negation that is not available for other types of noun phrases, namely the use of the negative marker *ʿadam* (< Classical Arabic *ʿadam* 'lack of'), as in the Egyptian Arabic example in (9a). Using this operator with nouns that denote things rather than events is ungrammatical, as in (9b).

(9a) (Egyptian Arabic)
ʿadam *dafʿ* *il-ʾigār*
not paying the-rent
f-il-maʿād *ʿamal-le-na*
in-the-appointment made-for-us
mašākil
problems
'Not paying the rent on time caused us problems'

(9b) **ʿadam* *il-filūs* *ʿamal-le-na*
not the-money made-for-us
mašākil
problems
'Not (having) money caused us problems'

What makes event nominals particularly interesting is that they have certain syntactic properties typically associated with verbs. For instance, they inherit the argument structures of the verbs they are derived from. In (10a), the event nominal *hurūb* 'escape' licenses the single Theme argument associated with the verb *haraba* 'to escape'. In (10b) the nominal *tadmīr* 'destruction' licenses an Agent and a Patient argument, just like the verb from which it is derived. In (10c), the event nominal *ḥuṣūl* 'obtaining' requires a prepositional complement, and in (10d) the event nominal *iqtirāḥ* 'suggestion' licenses a sentential complement.

(10a) (Standard Arabic)
hurūb *as-sajīn-i*
escape the-prisoner-Gen
'the prisoner's escape'

(10b) *tadmīr* *al-ʾaʿdāʾ-i*
destruction the-enemies-Gen
l-madīnat-a
the-city-Acc
'the enemies' destruction of the city'

(10c) *ḥuṣūl* *karīm* *ʿalā* *minḥa*
obtaining Karim on scholarship
'Karim's obtaining a scholarship'

(10d) *iqtirāḥ* *al-mudīr-i* *ʾan*
suggestion the-manager-Gen that
na-jtamiʿ-a *fī* *l-masāʾ*
1pl-meet-Subj in the-evening
'the manager's suggestion that we meet in the evening'

Another verbal property of event nominals is that they license manner, temporal, and instrument adverbials, as in the Egyptian examples in (11a)–(11c). However, sentential adverbs such as *dayman* 'always', *saʿāt* 'sometimes', and *ʾabadan* 'never' are excluded from phrases headed by event nominals, as indicated by the ungrammaticality of (11d).

(11a) (Egyptian Arabic)
kalām-i *maʿā-h*
talking-my with-him
bi-hidūʾ/bi-gad
with-quietness/with-seriousness
ʾaqnaʿ-u
convinced-him
'My talking with him quietly/seriously convinced him'

(11b) *wuṣūl-ak* *mitʾaxxar*
arrival-your late
ʾimbāriḥ/kull *yōm* *zaʿʿal*
yesterday/every day angered
il-mudīr
the-manager
'Your arriving late yesterday/every day angered the manager'

(11c) *fatḥ gōzit il-hind*
opening coconut
b-iš-šakūš ḥa-yi-ʿmil
with-the-hammer Fut-3ms-make
zarwaṭa
mess
'Your opening the coconut with a hammer will make a mess'

(11d) **taʾxīr-ak ʿan iš-šuġl*
being.late-your for the-work
saʿāt/dayman ha-yi-tsabbib
sometimes/always Fut-3ms-result
fi faṣl-ak
in firing-you
'Your being late for work sometimes/always will result in your getting fired'

The most significant verblike property of event nominals is that their object arguments can be marked for the accusative case in Standard Arabic, as in (12a).

(12a) (Standard Arabic)
sarra-nī jtiyāz-u
pleased-me passing-Nom
ṭ-ṭālib-i l-imtiḥān-a
the-student-Gen the-exam-Acc
'The student's passing the exam pleased me'

This pattern of case assignment is possible only if the nominalization phrase can be substituted with a complete clause headed by a complementizer, but only if the event nominal is not diminutive, dual, plural, or an instance noun (Ḥassān 1974; as-Sarrāj 1983; al-Ġalāyinī 2000). Interestingly, the Patient argument cannot be marked for the accusative case if the event is referred to by a pronoun, as in the second conjunct of the Standard Arabic example in (12b).

(12b) **ḥubb-ī l-ʾawṭān-a*
love-my the-homeland-Acc
ʿaḏ̣īm-un
great-Nom
wa-huwa bilād-an
and-it countries-Acc
ʾajnabiyyat-an ʾaqall-u
foreign-Acc less-Nom
'My love for the homeland is great, and it is less for foreign countries' (Ḥassān 1974:III, 215)

To account for the verbal properties of noun phrases headed by event nominals, derivational analyses posit an underlying verbal constituent or a larger structure that includes a verb. Although the syntactic derivation changes the categorical status of the underlying constituent, it maintains argument structure, case marking, and modifying adjuncts. Wise (1975) argues that event nominals in Egyptian Arabic are syntactically derived from complete sentences through a process of nominalization that neutralizes tense, aspect, and mood. This analysis is motivated by the meaning relation between event nominals, such as *ʾakl* 'eating' in (13a), and complete finite clauses with overt complementizers, such as *ʾinnu yākul* 'that he eat' in (13b).

(13a) (Egyptian Arabic)
šaggaʿ-t-u ʿala il-ʾakl
encouraged-1s-him on the-eating
'I encouraged him to eat'

(13b) (Egyptian Arabic)
šaggaʿ-t-u ʿala ʾinn-u
encouraged-1sg-him on that-he
yā-kul
3ms-eat
'I encouraged him to eat'

The major complication for analyzing an event nominal as the output of a syntactic process that takes a sentence as its input is that several sentential components are disallowed within noun phrases headed by event nominals. For example, sentential adverbials cannot be licensed by event nominals, as mentioned earlier. Moreover, sentential negation particles are ruled out within such phrases whether they are expressed continuously, as in (14a), or discontinuously, as in (14b).

(14a) (Egyptian Arabic)
**miš mirwāḥ-ak li-d-duktūr*
Neg going-your to-the-doctor
li-waḥd-ak kān ġalaṭ
on-own-your was mistake
'Your not going to the doctor on your own was a mistake'

(14b) (Egyptian Arabic)
**ma-mirwāḥ-ak-š* *li-d-duktūr*
Neg-going-your-Neg to-the-doctor
li-waḥd-ak *kān* *ġalaṭ*
on-own-your was mistake
'Your not going to the doctor on your own was a mistake'

Such an analysis assumes that the noun phrase headed by the event nominal *kawn* 'being' in the Standard Arabic example in (15a) is derived from a verbless sentence such as that in (15b). However, there is no explanation within this framework for why the copular auxiliary verb appears after the application of nominalization, but not before.

(15a) (Standard Arabic)
kawn-u *ʿalī* *muhājir*
being-Nom Ali immigrant
ḥadīṯ *yamnaʿu-hu* *min* *at-taṣwīt*
recent disallow-him from the-voting
'Ali's being a recent immigrant prevents him from voting'

(15b) *ʿalī* *muhājir* *ḥadīṯ*
Ali immigrant recent
'Ali is a recent immigrant'

Other derivational analyses avoid the complications of deriving event nominals from sentences by arguing that nominalizations are derived from verb phrases (Hazout 1995; Engelhardt 2000) or verbs (Fassi Fehri 1993; Gadalla 2000). However, other complications stem from the fact that there are numerous morphological templates for event nominals, as the basic *faʿala* verb form alone has forty templates for nominalizations (Gadalla 2000). Moreover, verbs can have multiple nominalizations, such as *laqiya* 'to meet' and *sāʾa* 'to turn bad', which have ten event nominals associated with each (Ḥalawānī 2000). Finally, some verbs do not have nominalizations at all, such as the Egyptian Arabic *xaff* 'to heal', *xallaṣ* 'to end', *baʾa* 'to become', and *ʾidda* 'to give'. Derivational analyses cannot explain how the morphological patterns of nominalizations are determined during the derivation or why some verbs do not have nominalizations, i.e., derivational accounts generate ungrammatical forms.

Lexicalist analyses argue that event nominals are lexically specified as nouns, rather than syntactically derived from verbs, and that they have argument structures that must be satisfied. The basic premise of the Lexicalist Hypothesis is that many nominalizations are ambiguous between a process reading that refers to events and a result reading that denotes things (Lebeaux 1986; Siloni 1997; Grimshaw 1990). For example, the nominal *bināʾ* 'building' can refer to a process of building, as in (16a), or to the outcome of such a process, as in (16b).

(16a) (Standard Arabic)
istaġraqa *bināʾ-u*
lasted.3ms building-Nom
s-sadd-i *ʿašr* *sanawāt*
the-dam-Gen ten years
'The building of the dam lasted ten years'

(16b) (Standard Arabic)
hāḏā *bināʾ* *min* *al-ʿaṣr*
this building from the-era
al-mamlūkī
the-Mamluk
'This is a building from the Mamluk era'

Result nominals can be preceded by demonstratives, have plural and dual forms, and carry diminutive morphology, while event nominals require the representation of their arguments, and their objects can be marked for the accusative case. Other diagnostics that distinguish process and result nominals include the observation that only event nominals can be used in control constructions, and that they co-occur with a wide range of modifiers, such as frequency and aspectual adverbials. It is important to note that event nominals do not always inherit the complete set of arguments associated with their corresponding verbs, as generic nominals that refer to types of events do not take arguments. Lexicalist analyses account for the variation in nominalization templates and their lexical gaps. However, they cannot explain the case-marking patterns associated with derived nominals, in particular those construct state constructions in which the subject argument is marked for nominative case, while the object argument is marked for the accusative case.

BIBLIOGRAPHICAL REFERENCES

Borer, Hagit. 1995. “The ups and downs of Hebrew verb movement”. *Natural Language and Linguistic Theory* 13.527–606.

Chomsky, Noam. 1970. “Remarks on nominalization”. *Readings in English transformational grammar,* ed. Roderick A. Jacobs and Peter S. Rosenbaum, 184–221. Waltham, Mass.: Ginn and Company.

Emonds, Joseph E. 2000. *Lexicon and grammar: The English syntacticon*. New York: Mouton de Gruyter.

Engelhardt, Miriam. 2000. “The projection of argument-taking nominals”. *Natural Language and Linguistic Theory* 18.41–88.

Fassi Fehri, Abdelkader. 1990. *al-Binā᾽ al-muwazzi῾: Naḏ̣ariyya fī binā᾽ al-kalima wa-binā᾽ al-jumla*. Casablanca: Dār Tubqal.

——. 1993. *Issues in the structure of Arabic clauses and words*. Dordrecht: Kluwer Academic Publishers.

Fu, Jingqi, Thomas Roeper, and Hagit Borer. 2001. “The VP within process nominals: Evidence from adverbs and the VP anaphor *do-so*”. *Natural Language and Linguistic Theory* 19.549–582.

Gadalla, Hassan. 2000. *Comparative morphology of Standard Egyptian Arabic*. Munich: Lincom Europe.

Ġalāyinī, Muṣṭafā al-. 2000. *Jāmi῾ ad-durūs al-῾arabiyya*. Beirut: Dār al-Kutub al-῾Ilmiyya.

Grimshaw, Jane. 1990. *Argument structure*. Cambridge, Mass.: MIT Press.

Ḥalawānī, Muḥammad Xayr. 2000. *al-Muġnī al-jadīd fī ῾ilm aṣ-ṣarf*. Beirut: Dār aš-Šarq al-῾Arabī.

Ḥassān, ῾Abbās. 1974. *an-Naḥw al-wāfī*. 3 vols. Cairo: Dār al-Ma῾ārif.

Hazout, Ilan. 1991. *Verbal nouns: Theta theoretic studies in Hebrew and Arabic*. Ph.D. diss., University of Massachusetts.

——. 1995. “Action nominalizations and the Lexicalist Hypothesis”. *Natural Language and Linguistic Theory* 13.355–404.

Lebeaux, David. 1986. “The interpretation of derived nominals”. *CLS Papers from the General Session at the Twenty-Second Regional Meeting of the Chicago Linguistic Society,* ed. Anne M. Farley, Peter T. Farley, and Karl-Erik McCullough, 231–247. Chicago: Chicago Linguistic Society.

Lees, Robert B. 1960. *The grammar of English nominalizations*. The Hague: Mouton.

Levi, Judith N. 1978. *The syntax and semantics of complex nominals*. New York: Academic Press.

Ogawa, Yoshiki. 2001. *A unified theory of verbal and nominal projections*. Oxford: Oxford University Press.

Rozwadowska, Bożena. 1997. *Toward a unified theory of nominalizations*. Wroclaw: Wydawnictwo Uniwersytetu Wrocławskiego.

Sarrāj, Muḥammad as-. 1983. *al-Lubāb fī qawā῾id al-luġa wa-᾽ālāt al-᾽adab*. Damascus: Dār al-Fikr.

Siloni, Tal. 1997. *Noun phrases and nominalization*. Dordrecht: Kluwer Academic Publishers.

Stepanov, Arthur. 1997. “Case properties of nominalization DPs”. *Perspectives on Arabic linguistics*, X, ed. Mushira Eid and Robert R. Ratcliffe, 69–90. Amsterdam and Philadelphia: J. Benjamins.

Wise, Hilary. 1975. *A transformational grammar of spoken Egyptian Arabic*. Oxford: Blackwell.

Zubizarreta, Maria Luisa. 1987. *Levels of representation in the lexicon and in the syntax*. Dordrecht: Foris.

MUSTAFA MUGHAZY
(Western Michigan University)

Nominative → Declension

Non-Concatenative → Morphology

North America

Arabic has roots in North America that extend at least as far back as when Muslim slaves were brought from West Africa. In the late 19th and early 20th century, Arab immigrants (largely Christians from Lebanon and Syria) established Arabic-speaking communities, which eventually assimilated. Unrest and difficult economic conditions have resulted in subsequent waves of immigrants arriving from Arabic-speaking countries, as well as even greater numbers of non-Arab Muslims. Arabic is also an important part of the religious life of millions of non-Arab Muslims. In addition, a comparatively small but steadily increasing number of non-Muslim, non-Arab North Americans are undertaking the study of this important world language.

1. ARABIC IN COLONIAL NORTH AMERICA

Relatively little is known of the West African slaves who first brought knowledge of Arabic to North America. More often than not, these slaves were educated and came from distinguished families. They had studied in Qur᾽ānic *madrasa*s before being enslaved, a few having studied in Timbuktu (→ Mali). Many were involved in trade and commerce. Their abilities seem to have been restricted largely to writing verses from the *Qur᾽ān* or simple business correspondence.

Extant remnants of the writing of these African Muslims are predominantly religious in nature. Versions of the *Fātiḥa* and other Qur᾽ānic verses were common and at times written as novelties for their masters. Some copies of the *Fātiḥa* are mislabeled 'the Lord's Prayer'. On the other hand, a document from the 1850s contains both the Lord's Prayer and Christian hymns in English but is written in Arabic script.

2. Buffalo, New York: An Arab immigrant community

Although Arabic-speaking communities existed first in New York City, as early as 1891, western New York was the location of the first Arabic-speaking community in the United States to have been extensively studied. Lebanese immigrants first came to Buffalo, New York, in 1888 (Dweik 1992). Searching for better economic, religious, and political conditions, most came to work at the Pan-American Exposition. The community was a mixture of literates and illiterates, the former opening shops and the latter peddling wares in the streets. Regardless of socioeconomic level, it was common for young men to immigrate first with their mothers, then send for other family members when enough money had been raised.

Like many other immigrant communities of the time, the Lebanese community held to their customs and traditions as a means of retaining their identity. Coffeehouses abounded, transforming into social clubs. Bards and musicians gave voice to the immigrants' experience and played at all major social functions of the community. It was to such a community in Boston that young Khalil Gibran came with his mother and three siblings in 1895. Young Lebanese immigrants like Gibran and Ameen Rihani became bicultural and went on to play significant roles in both American and Arab literary circles.

After World War II, the burgeoning Lebanese community in Buffalo (and elsewhere) dispersed, and family ties loosened. Integration, both culturally and linguistically, took precedence over preserving custom and tradition. A small school which had been established in the 1920s and taught Arabic to children closed. The notion of maintaining one's culture and language was regarded with much skepticism by the American public and politicians during this period. Arab Americans sought to disassociate themselves from their former culture and language. American xenophobia resulted in many Arabs anglicizing their names in order to more completely assimilate. Early Arab immigrants often identified with villages, not nationalities, which encouraged factionalism within the immigrant community, making it difficult to maintain a single ethnic identity. It was within this context that Arabic began to lose ground to English, which was perceived as having more prestige than Arabic.

3. Arab immigration to Canada and Mexico

The original patterns of Arab immigration in Canada (Hayani 1999) and Mexico (Marin-Guzman and Zeraoui 2003) are similar to those in the United States. Canada's first Arab immigrants were from Syria and Lebanon, arriving in Montreal in 1882. Most were Christians fleeing the Ottoman regime. Immigration was severely restricted until World War II, with only about seven thousand immigrants arriving in a sixty-year period. After World War II, however, restrictions were lifted, and the number of immigrants between 1946 and 1992 was approximately two hundred thousand. These later immigrants came primarily from Egypt and Morocco, in addition to the Levant, with roughly equal numbers of Muslims and Christians. As in the United States, the first generation of immigrants tends to have much stronger Arabic language abilities. Of those who have been in Canada less than ten years, 98.3 percent reported that they are able to speak Arabic and 96 percent read it; only 55 percent, though, of Canadian-born Arabs speak Arabic, and only 17 percent of them read it (Hayani 1999).

The first Arab immigrants in Mexico, small groups from Lebanon, arrived in 1878, in a period when Mexico had few restrictions on immigration. About five hundred Lebanese families settled in Mexico, mostly in the Yucatan Peninsula. There were also a few Palestinians, 122 between 1900 and 1909. The number of immigrants increased in the period just before World War II, with 245 registered Palestinian immigrants between 1920 and 1929, and an estimated 15,000 Lebanese living in Mexico by 1938.

The majority of Mexico's early Arab immigrants were Christians from small towns with agricultural economies, but once in Mexico they practiced commerce, often, as in the United States and Canada, working as peddlers and in small shops, selling trinkets and religious items. In 1933, although Arabs were only 4.23 percent of foreigners in Mexico, they owned 54.94 percent of the businesses registered to immigrants.

Many of the early immigrants took Mexican names in order to assimilate, but most either brought their wives and families from their homeland or married other Arab immigrants once they were in Mexico. In later generations, most of the immigrants have intermarried and are well integrated into Mexican society. Most do not speak Arabic, and many stand out in the intellectual, professional, and cultural life of Mexico, including Jaime Sabines, a poet; Jorge Hayeg Selu, an intellectual; Hector Azar Barbar, a playwright; and Emilio Chuayffet Chemor, a former governor and secretary of education.

4. Religious and cultural institutions

Religious and cultural institutions have played a key role in the preservation of identity, but not always of the Arabic language. Many early immigrants to the United States were Maronite Christians, who established an Eastern-rite church with services performed in Arabic and Syriac. It was not long before it became apparent that these services would be difficult to maintain, due to the lack of clergy. Attrition was inevitable. The clergy on hand adapted services to their congregation and its needs. Other factors also favored English, Spanish, or French over Arabic in North American liturgical settings. Among these were the notion that religion could be taught in any language; the death of older immigrants and lack of replacement immigrants; preference for English, Spanish, or French due to their prestige as the dominant language; and the changing ethnic composition of churches and mosques. The preference for the national language of their new home was an expression of immigrants' adaptation to their new environment and reflected their identity transformation away from the Old World.

North American Muslims come from varied backgrounds and countries. The language background of Muslims may stretch from English, Spanish, or French to Albanian, Amharic, and Urdu. English, French, or Spanish becomes the lingua franca for these worshippers. Given the role of Arabic in Islam, Muslims tend to maintain Arabic better than Christians, especially since the recent Islamic resurgence. This trend can also be seen in a rise in the trend of wearing head scarves among young girls, as well as one Dearborn imam's return to dressing like a Middle Eastern preacher. A prominent mosque in Dearborn that used to operate primarily in English hired a young Lebanese imam who speaks no English to take over for the retiring imam, and at the Islamic Institute in Dearborn, the sermons are given primarily in Arabic, to cater to many in the congregation who are new immigrants with poor English skills. The arrival of new immigrants is a key factor in creating the critical mass of Arabic speakers necessary to perpetuate or expand the use of Arabic.

The use of Arabic is in flux, highly dependent on the community and rhetorical needs of the priest or imam. Depending on situation and context, Arabic is either used or set aside for greatest impact. Sermons focused on Islamic theology or ethical matters are more likely to be in Arabic. However, where the thrust of the sermon aims to reach out to non-Arabic-speaking Muslims, the local language takes precedence. The continued maintenance of Arabic by religious institutions therefore depends on topic and the demands of the community and congregation, but above all the flow of new immigrants.

Cultural institutions based on ethnicity and language aided Arab immigrants in their transition to life in North America. That local culture and its maintenance were important to the immigrants and Arab community is emphasized by the fact that cultural institutions flourished in the early 20th century. During the Depression, when prejudice against foreigners increased, societies and centers formed to defend themselves and to redefine their role in society. In the period following World War II, Arabs sought greater assimilation over cultural and linguistic maintenance. This is not to suggest that Arabic disappeared entirely, but it was not until the 1960s that a reassertion of cultural identity allowed incorporation of Arabic into mainstream life for many Arabs.

5. Arabic media

Arabic-language newspapers sprang up shortly after the arrival of communities from the Arab world (on the development of Arabic media in North America, see Tayash and Ayoubi 1992). In the short span of time between 1892 and 1907, there were more than twenty such publications in the United States; by 1930, there were fifty. The establishment of so many newspapers was intimately tied to the village-mindedness of Arab-American communities. Each publication catered to the needs and views of a specific community or sector of a community. Supply outstripped demand, leading to the quick opening and closing of many publications.

The Arab-American press soon faced similar issues confronting religious institutions: what would the continued role of Arabic be? In order to cater to the full community, newspapers began to publish both Arabic and English articles. To provide coverage in only one language would effectively cut off half of the population and not contribute to the ideal of Arab unity through the press. As in the religious setting, context and topic are decisive in the use of Arabic or English. In bilingual papers, articles written in English deal primarily with issues related to life in North America. Publications dedicated to Islamic thought are primarily in Arabic. Student publications on university campuses fill the needs of expatriate students who share a common heritage and deal with U.S. and homeland issues. These highly specialized newspapers are typically funded by the home government in order to keep a connection between the expatriates and their country in an attempt to encourage students to return to their home countries after their studies. Representing a link between home and America, these papers offer coverage in Arabic and English.

In major metropolitan areas, radio and television broadcasts in Arabic have been available since the 1960s. Like its press counterpart, radio and television programming is designed to connect the immigrant community to its roots and serves to preserve Arab awareness and identity. Most programming is a mix of music, community news, and entertainment. Some programs advocate a political stance or ideology, but they do not dominate the airwaves. Where there are higher concentrations of Arabs there is naturally more programming available. By catering to different preferences, Arabic media are able to attract a wider audience. However, the diglossic nature of Arab speech communities exerts pressure on programming. A compromise between Modern Standard and vernacular Arabic is essential for appealing to the varied cross-sections of Arabs living in a given area. This is especially a difficult issue in terms of airing entertainment programs where humor is in large part linked to the dialect and turns of phrase.

While the media allow Arabs contact with their roots, they do not impede assimilation. Some have maintained that ethnic media reinforce the immigrant's culture at the expense of assimilation into the adopted country. This has not been the case with Arabic media. One study indicated that Arabic media in two Arab-dense centers (California and Michigan) attempt to impart assimilationist views to the community. This programming focuses primarily on American values and culture, addresses community issues, and motivates the Arab community to become involved with civic and local affairs. Arabic media are thus a transition zone between older customs and the new environment.

Programming available via Internet and satellite is transforming the Arabic media landscape, facilitating contact with the Arab world and Arabs elsewhere as never before. Two sizable television networks based in the United States cater to Arab audiences: the Arab Network of America (ANA; a subsidiary of the London-based but Saudi-funded Middle East Broadcasting Center) and TAC TV, a New Jersey-based private endeavor that prides itself on its independence. Internet resources also offer opportunities for Arabs living outside the Arab world to maintain connections previously unheard of to the homeland and to their native languages.

6. Arabic speakers in North America today

The Modern Language Association, using data from the 2000 U.S. Census (www.mla.org/map_main), estimates a total of 614,582 Arabic speakers in the United States, focused primarily in cities such as Los Angeles, Seattle, Chicago, Dearborn, and New York. The vast majority, 91 percent, of Arab immigrants live in urban areas, with 33 percent located in only three states: California, where Arabic is the 14th most spoken foreign language, New York, where it is the 13th, and Michigan, where it is second only to Spanish. However, the extent of Arabic's decline among Arab Americans is clear when the number of Arabic speakers is compared with the 1,189,731 people who reported Arab ancestry in the U.S. Census (this latter figure is smaller than many estimates, primarily because the U.S. Census form does not have a precise classification for Arab Americans).

There are also a number of Arabic-speaking communities in Canada, primarily in Ontario and Quebec. The 2001 Canadian Census (www.12.statcan.ca/english/census01/home/Index.cfm) reported 199,940 native speakers of Arabic, with 290,280 reporting some ability in speaking Arabic, making it the fifteenth most commonly spoken language. There are also nearly 10,000 speakers of Maltese, with 7,380 reporting it as their mother tongue. No parallel data on Arabic in Mexico was found.

7. The study of Arabic as a foreign language

The study of Arabic as a foreign language predates the American Revolution and has increased considerably from the 1950s onward. Recent world events have given particular impetus to its study.

Complementing the study of Hebrew and the Old Testament, Harvard began offering Classical Arabic in 1654, with Yale following suit in 1700. The study of Arabic at this time was connected with Biblical studies and interest in Semitic languages. Interest in the Arabic of the modern world came about as a result of World War II, which exposed the United States' lack of linguistic preparedness to deal with most of the world. Training manuals for various dialects of Arabic were quickly produced, and Arabic training programs soon appeared, created by the Foreign Service Institute for diplomats and the Defense Language Institute (DLI) for military personnel. Sputnik resulted in the creation of federally funded area studies centers at major U.S. universities. Both federal and other external funding resulted in American universities beginning to incorporate Arabic into their curriculum. In 1957, a Ford Foundation grant to a consortium of schools (Columbia, Harvard, Johns Hopkins, Michigan, and Princeton) sought to establish the regular offering of at least three of the major Middle Eastern languages each summer and to develop new pedagogical tools and to improve teaching methodology.

The trend toward teaching/learning Arabic with a greater focus on the modern world and with heightened emphasis on face-to-face communication (especially in the United States) was significantly accelerated by two major factors. Since the 1970s foreign language educators in general have become increasingly focused on helping students develop 'communicative competence', 'functionally useful foreign language skills'. Arabic was swept along with the 'proficiency movement', resulting in new methods and materials. Changing student demographics, particularly in the United States, also contributed to greater emphasis on oral skills. The typical student of the 1960s and before was a graduate student primarily interested in developing Arabic reading skills. This is no longer the case. Undergraduates now vastly outnumber graduate students, accounting for nearly 95 percent of all students of Arabic reported in the Modern Language Association's 2002 survey.

Increased emphasis on oral skills and on authenticity has resulted in more and more experimentation with introducing students to Arabic as it is actually used in the Arab world. For example, the DVDs that accompany the most widely used textbook series in North America (Brustad a.o. 2004) present monologues in Modern Standard Arabic but dialogues in Egyptian Arabic (efforts are underway to also make these available in Moroccan and Jordanian/Palestinian).

Significant growth in the number of heritage students in recent years has resulted in an even more varied pool of students. Some come to Arabic with considerable facility in spoken Arabic, looking to acquire literacy skills. Muslims interested in Arabic for purely religious reasons represent another type of heritage student. Accommodating the contrasting priorities of all of these students represents a significant challenge.

The events of 11 September 2001 again revealed an America unprepared to effectively interact with the world. Government needs and heightened media attention focused on the Middle East have spurred U.S. Arabic enrollments in higher education from 5,505 in 1998 to 10,584 in 2002, according to the Modern Language Association. No comprehensive national survey has been conducted since this time, but schools across the country have continued to report steady increases in enrollments, and more and more institutions that formally did not offer the language have placed record numbers of ads for instructors and professors of Arabic. This remarkable growth attests to the importance and prominence Americans attach to Arabic and its relevance to their lives and society.

Relatively few of the thousands of Americans that embark on the study of Arabic achieve high levels of proficiency. The Center for Arabic Study Abroad (CASA) hosted at the American University in Cairo is an outstanding exception. In spite of vacillating federal funding, this consortium of U.S. universities has managed to survive for decades and leads in producing Superior-level speakers. Demand for CASA is now at an all-time high. The new Flagship program funded by the National Security Education Program seeks to build national capacity and is opening up opportunities for students with government service plans to pursue advanced study of Arabic and other critical languages in the United States and abroad.

Opportunities for studying Arabic are largely restricted to major colleges and universities and programs for government employees. Public K–12 programs are rare but growing in number. Arabic instruction is common in private Islamic schools but typically limited in effectiveness beyond the earliest years. The lack of quality age-appropriate learning materials and an acute shortage of qualified teachers present formidable obstacles at present, to say nothing of the challenges of dealing with the decentralized U.S. educational system.

Students swelling the ranks of Arabic classrooms in U.S. colleges and universities are part of a trend that has been underway for decades. They are predominantly undergraduates with a primary interest in the modern Arab world. A recent survey of more than six hundred students enrolled in Arabic courses at 37 U.S. colleges and universities (Belnap 2006) reveals that most are serious about learning Arabic well so they can "function in it comfortably in [their] professional activities". They are learning Arabic for a variety of reasons. The most popular majors include Middle/Near Eastern studies (24.2%), political science (9.4%), international relations (8.4%), linguistics (6.3%), and history (4.2%). Most plan to work for the government (41%), for nongovernmental organizations (35%), in higher education (26%), or in business (20%). (These are nonexclusive choices, with 'other' chosen by 33 percent of the respondents.) Above all else, these students want to develop fluency in speaking. They want to travel to the Arab world. Most want to achieve a level of proficiency that will allow them to read the modern Arabic press and understand radio and TV broadcasts.

Heritage students (those of Muslim or Arab descent) account for less than 20 percent of the total number of students in the sample. Only 12 percent reported that they come from a home where one or both of their parents speak Arabic natively. The number of heritage speakers appears to be growing, but not keeping pace with the numbers of nonheritage students enrolling in the study of Arabic.

8. Arabic as a foreign language in Mexico and Canada

Although their statistics are less documented, universities teaching Arabic in Mexico and Canada have seen patterns of enrollment resembling those in the United States. In Canada, several universities, such as the University of Toronto, have large numbers of heritage students, especially from Egypt. Most have larger enrollments at the undergraduate than the graduate level, and all have seen increased enrollment since 2001. McGill University in Canada has seen a 20 percent increase in requests for enrollment, York University reports increased demand, and Universidad Regiomontana in Mexico reports offering Arabic classes for the first time ever, starting September 2006.

9. Conclusion

The presence of Arabic in North America is dynamic, responding equally to the world around it and to local conditions. Linguistic territory is contested space dependent on multiple factors. Language maintenance is always an uphill battle. Media outlets offer recent immigrants and long-time residents a connection to their culture and language. Arabic in America is not only the domain of Arabs. American students, acutely aware of the relevance of

Arabic to their lives, are enrolling nationwide in courses. This Arabic renaissance assures the language's continued place in America.

Bibliographical references

Austin, Allen. 1997. *African Muslims in antebellum America: Transatlantic stories and spiritual struggles*. New York: Routledge.

Belnap, R. Kirk. 2006. "A profile of students of Arabic in U.S. universities". *Handbook for Arabic language teaching professionals*, ed. Kassem M. Wahba, Zeinab A. Taha, and Liz England, 169–178. Mahwah, N.J.: Lawrence Erlbaum.

Brustad, Kristen, Mahmoud Al-Batal, and Abbas Al-Tonsi. 2004. *al-Kitaab fii Taʿallum al-ʿArabiyya with DVDs: A textbook for beginning Arabic*, I. 2nd ed. Washington, D.C.: Georgetown University Press.

Dweik, Bader S. 1992. "Lebanese Christians in Buffalo: Language maintenance and language shift". Rouchdy (1992:100–118).

Hayani, Ibrahim. 1999. "Arabs in Canada: Assimilation or integration?". *Arabs in America: Building a new future*, ed. Michael W. Suleiman, 284–303. Philadelphia: Temple University Press.

Marin-Guzman, Roberto and Zidane Zeraoui. 2003. *Arab immigration in Mexico in the nineteenth and twentieth centuries: Assimilation and Arab heritage*. Austin, Tex.: Augustine Press.

Rouchdy, Aleya (ed.). 1992. *The Arabic language in America*. Detroit: Wayne State University Press.

Suleiman, Michael W. 1994. "Los árabes en America Latina: Bibliografía preliminar". *Estudios Migratorios Latinoamericanos* 9.165–188.

——. 2006. *The Arab-American experience in the United States and Canada: A classified annotated bibliography*. Ann Arbor, Mich.: Pierian Press.

Tayash, Fahad A. and Kenneth K. Ayouby. 1992. "Arab-American media: Past and present". Rouchdy (1992:162–183).

R. Kirk Belnap (Brigham Young University)

North Arabian → Thamudic

Northwest Arabian Arabic

Northwest Arabian Arabic is a group of dialects spoken by the Bedouin population of the Sinai Peninsula, the Negev, southern Jordan, and the northwestern corner of Saudi Arabia, an area virtually identical with Arabia Petraea with its eastern and southern extensions. Culturally, the area is relatively homogeneous, representing Bedouin culture of seminomadic, or at times semisedentary, type. The society is based on a tribal system, and the most important means of livelihood are the tending of sheep and goats and, under favorable conditions, agriculture, palm cultivation – especially in the oases of southern Sinai – and also in a few cases, fishing. Lying at the crossroads between the Ḥijāz, Egypt, and Greater Syria, the area has traditionally been of vital importance for trade and pilgrim traffic. Most of the tribes living along the routes have been engaged in transport of goods, providing the food supply for pilgrims, and protecting the routes – activities implying lively contacts with the inhabitants of the adjacent sedentary areas.

The relatively homogeneous culture of the area is well illustrated by the uniform → poetic koine used by the Bedouin. As to the western branch of the Northwest Arabian dialects, consisting of those spoken in Sinai and the Negev, their use of a uniform poetic koine appears clearly from the collection of poems published by Bailey (1991), but, broadly speaking, this is true of the eastern branch spoken to the east of the Wādi ʿAraba as well. Linguistically, the area is less homogeneous, more as the result of language contact than of the migration history of the tribes.

1. The dialects of the area and the state of their study

The western branch of the Northwest Arabian dialects can be divided into four areas: the Negev (→ Negev Arabic), the northern Sinai littoral, the high plateau of at-Tīh in central Sinai, and the mountainous southern portion of the Sinai Peninsula (→ Sinai Arabic). The Negev dialects are well covered by linguistic investigations, with Blanc's (1970) concise monograph, chiefly based on the dialect of the Ḍullām, as the seminal pioneer study. Based on the material Blanc had collected from other Negev tribes, he found that only minor dialect differences exist between the Ḍullām and the rest. The dialect of the ʿAzāzmih, the Ḍullām's neighbors in the region of Beersheba, is virtually identical, as appears from the narrative recorded by Yehuda Katz and published by Palva (1984b) and, above all, from the narratives, jokes, and dialogues included in Henkin's (1985) monograph on the tense and aspect system in the dialect of the ʿAzāzmih. The pre-Islamic origins of the two seminomadic tribes can be traced back to the northern Ḥijāzī tribe of Bilī, but they have been living in the northern Negev or its immediate neighborhood for several centuries. Since the spread of the Tiyāha confederation and the Tarābīn to the area between Gaza and Beersheba during the 19th century, the Ḏullām and the ʿAzāzmih have been in close contact with them (Oppenheim 1943:122, 129; Bailey 1985, *passim*).

For the central Sinai dialects, an extensive text collection in plain colloquial is available, published by Stewart (1988, 1990). The majority of the speakers of these texts are members of the Aḥaywāt, whose tribal area comprises the southeastern part of the high plateau between Naxl and al-ʿAqaba. The Aḥaywāt probably came to their present habitat in the 17th century from the eastern littoral of the Gulf of ʿAqaba, where they had been attached to the Bani ʿAṭīye in the 16th century (Oppenheim 1943:149–151; Bailey 1985:48).

For the dialect of the Ṭawara (Ṭuwára) alliance living in southern Sinai, no descriptions have been published thus far, but for the small Jbāliyyah tribe living in the neighborhood of St. Catherine's Monastery and the littoral next to it, a basic vocabulary has been published by Nishio (1992). This tribe has lived in symbiosis with the monastery since the 6th century, and still today the servants of the monastery are members of the Jbāliyyah (Oppenheim 1943:165; Nishio 1992:ix–x). A more detailed description of the dialect of the Jbāliyyah and its western neighbor Garārša was published by de Jong in 2003, based on recordings and notes taken by Woidich in the 1960s. The Garārša have been attached to the Ṣawālḥa, who may have come from the eastern Delta, whereas the Mzēna and the ʿAlēgāt may be of Ḥijāzī origin. In any case, the Ṭawara tribes have lived in close alliance since the 17th century (Oppenheim 1943:156–157), and the earlier dialectal differences must have faded away long ago.

The dialects of the northern Sinai littoral have been investigated in detail, with special emphasis on dialect geography, by de Jong (2000). Among the tribes included in the study, the Bilī probably came from the Ḥijāz to northern Sinai already in the pre-Islamic era, and the Biyyāḏiyyah during the first Islamic centuries. Also, the Axārsah, Samāʿnah, and ʿAgāylah had moved here before the 13th century (Oppenheim 1943:140; Bailey 1985:47, 1991:51), while the rest arrived in the 16th century or later. In the typological classification of the dialects of the area, based on as many as 73 criteria, de Jong has joined together those spoken by the easternmost tribes, Rmēlāt, Sawārkah, and Bilī, into Group I, which displays important similarities with the dialects spoken in the Negev and central Sinai. The ʿAyāydah and Masāʿīd, although living close to the Šarqiyya province, also speak dialects of this group. Group III is composed of the dialects of the Biyyāḏiyyah and the Axārsah, which share many traits with those spoken in the eastern part of the Šarqiyya. The dialects of the Samāʿnah and ʿAgāylah constitute Group II, which occupies an intermediate position between Groups I and III (de Jong 2000:622–627). The dialect of the Dawāġrah, a relatively isolated pariah tribe living on fishing, shows a great number of typological deviations from the dialects of the adjacent area and is classified as a separate Group IV.

The eastern branch comprises the dialects spoken by the Ḥuwayṭāt (Ḥwēṭāt), Bani ʿAṭīye, and a number of small tribes attached to them. The Ḥwēṭāt are divided into three subtribes: Ḥwēṭāt Ibn Jāzi, mostly living in southern Jordan between the Šara mountains, Wādi Sirḥān, and Jabal Ṭubayq (Ṭubēg); Ḥwēṭāt Ibn Nijād (or the ʿAlāwīn), who live north and east of al-ʿAqaba (al-ʿAgabe); and Ḥwēṭāt at-Tahāma (or the Ḥwēṭāt Abu Ṭgēga), living in the littoral of the Red Sea between al-ʿAqaba and Wādi ad-Dāma, about 60 km south of Muwayliḥ. The tribal area of the Bani ʿAṭīye lies south of the Jordanian border and stretches on the Saudi Arabian side to the depression of al-Jaww, an important cultural and linguistic boundary between Ahl aš-Šimāl and Ahl Giblī (Oppenheim 1943:352), about 130 km south of Tabūk. In the east it extends as far as the western fringes of the Nufūd.

The dialect of the eastern branch is relatively homogeneous; in the words of the Ḥwēṭi and ʿAṭúwi tribesmen, these tribes speak one and the same dialect (*luġa waḥde*). The situation can plausibly be traced back to the common origins of the two tribes. According to tribal traditions, during the 16th and 17th centuries the Ḥwēṭāt first emerged from the Bani ʿAṭīye, then when they had expanded at the expense of the Bani ʿAṭīye, they forced them to move from all coastal areas and to extend their habitat to the highlands of Ḥisma in the south, where Tabūk became their most important center (Oppenheim 1943:291–292, 337–338). No extensive studies of these dialects have been published thus far. Their main characteristics are described by Palva (1984–1986), whose article was based on material recorded in al-Jafr (Ḥwēṭāt Ibn Jāzi), al-Badᶜ (Ḥwēṭāt at-Tahāma), and Tabūk with surroundings (Bani ᶜAṭīye), later complemented with material from ad-Dīse in Wādi Ramm (Ḥwēṭāt Ibn Jāzi; Palva 2004). Also, a text published by Werner (2003) was recorded in Wādi Ramm. The stylistic target level of the speaker of this text is Standard Arabic, but it is interesting to note which Ḥwēṭī features are not suppressed. The dialect of the Bdūl, who until recently used to live in the caves of Petra, is described in an article by Bani Yasin and Owens (1984), and that of the Nᶜēmāt living east of Petra by Yrttiaho (1988).

As far as the origins of the Ḥwēṭāt and the Bani ᶜAṭīye are concerned, only a few traditions claim that they – except some individual shaykhs or clans – have come from outside their present habitat. According to Rentz (1968:643), there are indications that the conversion of the bulk of the Ḥwēṭāt to nomadism took place fairly recently. In spite of their marked identity as formidable Bedouin raiders and soldiers, the Ḥwēṭāt, as well as the Bani ᶜAṭīye, probably are descendants of an old local population (*ahl ad-dīre*) (Musil 1926:20), whose culture for centuries has fluctuated between seminomadism and semisedentarism. The tradition that traces the origins of the tribe to the ᶜAneze confederation of northern Arabia (Werner 2003:62.2) can plausibly be attributed to the tendency of strengthening the identity of the Ḥwēṭāt as Bedouin of 'noble origin' (*aṣīl*). In view of the important role played by the tribe in the Arab Revolt in 1916–1918 and its resolute fight against the attacks of the Wahhabis in the 1920s and at the beginning of the 1930s (Lawrence 1935:256–287 and *passim*; Oppenheim 1943:180–181; Glubb 1948:70–88 and *passim*; Peake 1958:98–103), this kind of self-assertion is only natural. In the descriptions written before World War I, the image of the Ḥwēṭāt differs significantly from its later heroic identity. Thus, Wallin (1850:302) reports that they are "looked down upon by other tribes as mixed Bedooins sprung from Fellāḥs, not of pure nomadic origin". Doughty (1923:234–235), who on his way from Cairo to the Ḥijāz in 1875–1876 crossed the area, describes some of the Ḥwēṭāt at-Tahāma as "nomad herdsmen" and some as "husbandmen of palms and sowers of grain", while many of those living in the Ḥisma are "barley sowers". Doughty's common impression of the tribe is that they are "liker nomad fellahīn than Beduins" (Doughty 1923:29), and in his ears their speech "savours of peasantry, even in the mouths of those that live furthest in Arabia" (Doughty 1923:45). Typical of the early explorers' observations, no specific linguistic traits are given. In the impressionistic way common to these descriptions, Burckhardt (1831:211–212) also points out that "the language of the Nedjd Bedouin is as different from that of Sinai, as the dialect of the latter is from an Egyptian Bedouin's" (Palva 1997:228, 238).

2. TYPOLOGICAL CLASSIFICATION OF THE DIALECTS OF THE AREA

In a short article, the typological classification has to be made on the basis of only a few criteria. In the selection of the distinctive features no objective method exists. In the following sketch, a kind of balance between phonetic, phonemic, morphological, syntactic, and lexical features is aimed at. In the transcriptions, as well as in the names of the tribes, the orthography used in the cited studies is followed. The secondary emphatics have not been marked, however.

The Northwest Arabian dialects can unambiguously be classified as Bedouin dialects, justifiably so called with reference to the traditional cultural type of the speakers. They display several distinctive Bedouin hallmarks, which as a rule are not shared by the sedentary dialects of the adjacent areas, among the most significant of which are:

(a) The voiced reflex of **q*, a feature common to all Bedouin dialects. This trait is not restricted to Bedouin dialects alone, but appears in many sedentary dialects of Arabic as an integrated Bedouin feature (→ *qāf*). In the neighborhood of the Northwest Arabian dialect area, it is shared by the sedentary dialects spoken in the greater part of the Šarqiyya province in the eastern Nile Delta, in the villages of southern Palestine, the eastern part of the Plain of Jezreel, the Jordan valley and the settled areas to the east of the River Jordan. In all of these dialects the /g/ reflex of **q* is due to Bedouin influence, resulting from Bedouin predominance in the area in past centuries, as well as from sedentarization of Bedouin (Abul-Fadl 1961:259 and map 3; Behnstedt and Woidich 1985, map 7; Bergsträsser 1915, map 4; Palva 1984a:363–364, 1965:24–25, 156–160, 1989:228; for the language situation of modern Amman, see Abdel-Jawad 1981 and → Jordanian Arabic (Amman)).

(b) The *gahâwah* syndrome (→ *gahawa*-syndrome), a term coined by Blanc (1970:125–127), or guttural resyllabication rule, so called by Ingham (1994:19). It consists in the insertion of an anaptyctic *a* after X in most older (C)*a*XC(V) sequences in which X is any of the back spirants *h*, *ʿ*, *ḥ*, *ġ*, and *x*, e.g. *gahwa(h)* > *gahâwa(h)* 'coffee'; *naʿja(h)* > *naʿája(h)* 'ewe', *maḥl* > *maḥal* 'drought', *baġl* > *baġal* 'mule', *naxl* > *naxal* 'date palms [coll.]'.

This feature is not operative in all morphological categories. Two alternative diachronic explanations of the syndrome are given by de Jong (2000:110–111). In dialects influenced by neighboring sedentary dialects, the *gahâwah* syndrome is gradually losing currency; this is at present the situation in the dialects of the Biyyāḏ̣iyyah and the Axārsah in the western part of the northern Sinai littoral (Group III; de Jong 2000:351–352). Typical of the effort to speak Standard Arabic, the *gahâwah* syndrome is completely suppressed in the text from Wādi Ramm published by Werner (2003), whereas it occurs regularly in the plain colloquial texts from the same village, e.g. *nxabiz, nʿajin, nġazi, nḥaṣid, yġazu* (Palva 2004:202–204).

(c) Gender distinction in the 2nd and 3rd person plural in personal pronouns, pronominal suffixes, and finite verbal forms. This trait is also found in the rural dialects of central and southern Palestine, although at present on the decline, as well as in the sedentary dialects east of the River Jordan (Bauer 1926:18 and passim; Palva 1984a:367, 1989:232).

(d) Productivity of Form IV ($aC_1C_2aC_3$, $yiC_1C_2iC_3$). In the northern Sinai littoral it is productive in the dialects of the easternmost tribes only (Group I; de Jong 2000:218–221, 315–316, 391). In other dialects of the Northwest Arabian dialect area it is productive; Blanc (1970:135) even mentions that in the Negev it is in particularly frequent use.

(e) The definite article *al-* and the relative pronoun *alli*. The article is stressable as an integral part of the word, e.g. *álwalad, áljabal, álbaḥar*. The initial /a/ is stable enough to be preserved after *-ī* (*-iy*), which is dropped: *f-albēt, rāʿ-álġanam, fāḏ̣ albāl*, in some cases after *-ū* (*-uw*) as well: *ab-albint, ax-ajjōz* (Blanc 1970:143; Bani Yasin and Owens 1984:219; Palva 1984–1986:306; de Jong 2000:155–156). In the dialects of the Biyyāḏ̣iyyah and Axārsah, as well as in the southern Sinai dialects of the Jbāliyyah and Garāriša, the forms have initial /i/ (Group III; de Jong 2000:155; 2003:163), probably owing to language contact. In koineizing speech, initial /i/ is gaining ground in the whole area.

(f) A number of typical Bedouin lexical items (*gōṭar* 'to go', *sōlaf* 'to tell, narrate', *ṭabb* 'to arrive', *nišad* ~ *nišád* 'to ask', etc.).

(g) Frequent and productive use of diminutives, not only in lexicalized forms like *šwayyi(h)*, *kwayyis*, or *grayyib*, but productively as well, e.g. *glayyil* 'a little', *xbayz* 'bread', *aḥaymir* 'red(dish)' (Blanc 1970:140–141; de Jong 2000:278; Rosenhouse 1984:23–24).

The Northwest Arabian dialects share several salient features distinguishing them from the Bedouin dialects of the North Arabian type:

(a) Absence of the → *tanwīn* and its residues. The use of the *tanwīn* in these dialects is limited to poetry, proverbs, and formulaic expressions. The morpheme /-an/ occurring in commonly used words such as *ṭabʿan* 'naturally' and *tagrīban* 'about' is not an inherited feature but an integral part of a lexical loan from Standard Arabic, often even adopted in urban form, e.g. *masalan* 'for instance'. The word *abadan* 'never' probably belongs to the same category, but inherited forms still survive: *abdan* in central Sinai (Stewart 1990:69.53) and *abdan* ~ *abda* in the eastern branch (Palva 1984–1986:305). The absence of the *tanwīn* as a typologically distinctive feature was already noted by Wallin, who in 1846 wrote in a letter that compared them with the dialects spoken in Najd and the northern desert, the dialects of the Ṭawara, Aḥaywāt, and Ḥwēṭāt are an exception "as they do not make use of the *tanwīn*". This observation led him to assume that these tribes actually are not Najdi Bedouin (Elmgren 1864–1866: III 165; Palva 1997:231–232).

(b) Absence of affricated variants of /g/ and /k/. In the North Arabian dialects these phonemes have phonetically conditioned affricated variants *ǧ/ģ* and *č/ć: šarg, šarǧi; yarkab, bāčir*. In Northwest Arabian Arabic, affricated variants occur only exceptionally. A typical case is the pejorative use of *č* in the exclamations *čalb*! '[you] dog!' *čiḏib*! 'that's a lie!' (Bani ʿAṭīye; Palva, own observation).

(c) Absence of final /n/ in the imperfect, 2nd person feminine singular, 2nd person masculine plural, and 3rd person masculine plural, the only exception being the dialect of the fisherman tribe of the Dawāġrah (de Jong 2000:460). This feature is so well established that the Ḥwēṭī informant interviewed by Werner (2003) in Wādi Ramm, in spite of his effort to speak Standard Arabic, only once uses an imperfect form with final /n/.

(d) The pronominal suffix of the 2nd person masculine plural is *-ku* (*-kuw*), which contrasts with the North Arabian *-kom/-kam*.

(e) The use of the locative preposition *fi* (*fiy*), which in the majority of North Arabian dialects is only used in *fīh* 'there is'.

In spite of a number of internal differences, the Northwest Arabian dialects share several particular features:

(a) Stressed variants *-ī́* and *-nī́* of the pronominal suffix in the 1st person singular (Blanc 1970:131; Palva 1984–1986:297; de Jong 2000:167). This feature might be regressive; thus, in the Šarqiyya province it is reported to appear sporadically in the speech of the older generation (Abul-Fadl 1961:237; Behnstedt and Woidich 1985, map 150). On the other hand, in the Negev it often occurs even in koineized discourse: *muš ʿumrīh aḥsan min ʿumrak* 'my life is not better than yours' (Henkin 1985:11.130–12.131).

(b) Plural comm. forms *haḏalla, haḏallāk, hōḏ̣aḷḷa, hōḏ̣aḷḷāk* of the demonstrative pronoun. These forms are historically most interesting, since they display affinities to other Semitic languages (Rabin 1951:153; Fischer 1959:109).

(c) Initial /a/ in Forms VII, VIII, and X in the perfect, and stressed when in stressable position. This trait is shared, for instance, by the dialects of Mecca and the Šukriyya in northeast Sudan (Schreiber 1970:38; Reichmuth 1983:258–270). It is interesting to notice that Werner's Ḥwēṭī informant in Wādi Ramm, in his approximation of Standard Arabic, freely mixes these kinds of forms: *anḏ̣ammū, antaṣarū, astašhadū* (Werner 2003:63.25, 67.78, 67.80). For southern Sinai, forms of the type *infátaḥ* are reported (de Jong 2003:169).

(d) Initial /a/ in a number of irregular nouns (*amm, axt, axwān, adēn, afám*).

(e) The invariable pronominal suffix *-ki* of the 2nd person feminine singular.

The internal differences between the western and eastern branches of Northwest Arabian Arabic include the following:

(a) In the eastern branch of Northwest Arabian Arabic, the *b-* imperfect does not occur in plain colloquial, whereas in the entire western branch it is in regular use. In the Negev, it functions "much as among Palestinian and Syrian sedentaries" (Blanc 1970:139), that is, as indicative non-past. The only dialect in Sinai in which it is not used is that of the fisherman tribe of the Dawāġrah (de Jong 2000:32). In the eastern branch, it has been reported from the two small tribes living in the Šara mountains, viz. the Nʿēmāt (Yrttiaho 1988:148) and the Bdūl of Petra. Typical of a recently borrowed feature, in both cases the use of the *b-* imperfect is unsystematic. Thus, among the Bdūl it is always optional and said to have been adopted as late as the mid-20th century (Bani Yasin and Owens 1984:214–216). Today it is rapidly gaining ground as part of koineizing discourse in the whole of southern Jordan, especially among the younger generation (Werner 2003:61). In al-ʿAqaba and Tabūk with their mixed population, the *b-* imperfect is of common occurrence (Palva, own observations). It is worth noticing that the *b-* imperfect is also used in urban Ḥijāzī dialects in which it is reported to indicate progressive aspect (Sieny 1978:17).

(b) Unlike the dialects of the eastern branch, the western branch dialects make use of an analytic genitive, with *šuġl* (masc. sg.), *šuġlah* (fem.), *šuġlīn* (masc. pl.), *šuġlāt* (fem. pl.) as genitive markers (Blanc 1970:147; Stewart 1990, Glossary s.v. *šġl*). Interestingly, elsewhere this marker has been attested for Upper Egypt only, viz. the dialects spoken between Asyūṭ and the bend of the Nile (Behnstedt and Woidich 1985, map 189).

(c) The dialects of the western branch have → vowel harmony in the preformative of the active imperfect of Form I (Blanc 1970:136; de Jong 2000, map 53), whereas in the eastern branch the vowel of the preformative is mainly generalized /a/. In Ḥwēṭī, the ongoing dialect leveling is reflected by the difference between a female speaker of the older generation and the male speakers living in the same village. She uses systematically the generalized /a/, e.g. *narḥal, naglib, nʿajin, nġazil*, while the male speakers use the Jordanian koine pattern side by side with the older type, e.g. *yibḥaṯ, niṣḥa, yigṭaʿu; yašrabaw, yasgu, yġazu* (Palva, 2004:198). In this feature the Bdūl and the Nʿēmāt follow the western group (Bani Yasin and Owens 1984:211; Yrttiaho 1988:149).

(d) In the dialects of the eastern branch, as well as in the dialects spoken in southern Sinai, the reflexes of **aw* and **ay* are well-established monophthongs /ō/ and /ē/, usually after back consonants and emphatics as well (Palva 1984–1986:296; de Jong 2003:153). It is illustrative of the stability of the monophthongs that even the Ḥwēṭī speaker whose ambition it is to speak Standard Arabic, as a rule does not 'correct' the monophthongs, not only in common words such as *yōm, gōm, dōr, dōla, šēx, bēn, ġēr, ġēš, ʿēn*, but even in *aṯṯōra lʿarabiyya lkubra* (Werner 2003, *passim*). In most dialects of the western branch, **aw* and **ay* have been partially monophthongized, but the new monophthongs fluctuate with long phonemes: /ō/ ~ /ū/, /ē/ ~ /ī/. This trait also occurs in some dialects of the North Arabian dialect type, e.g. that of the ʿAjārma in the Balqa in central Jordan (Palva 1976:19–20). This implies that the phonemic status of the monophthongized diphthongs has not been established, and they tend to become merged into the older long vowel phonemes.

(e) Differences in the so-called *gaháwah* syndrome. In North Arabian Bedouin dialects, the old /a/ of the initial syllable is dropped, e.g. *ghawa, nxala, nʿaja* (Cantineau 1936:61–62, 1937:167–170). The eastern branch of the Northwest Arabian dialects follows this pattern, although not strictly (Palva 1984–1986:297); the variants *ghawa* and *gaháwa* appear freely. In Sinai and the Negev, the /a/ of the initial syllable is preserved, except in the salutation *hala* (< *ahála*) and in the dialect of the fisherman tribe of the Dawāġrah (Blanc 1970:125–126; de Jong 2000, map 19). Also in the dialect of the Bdūl, which in several other respects stands closer to the Negev dialect than the other dialects of the eastern branch, the vowel of the initial syllable is often dropped: *tḥat/taḥat, ḥamar/aḥamar, xawāl* (< *axwāl*), *ʿamām* (< *aʿmām*) (Bani Yasin and Owens 1984:203).

(f) There are differences in the synchronic resyllabification process of the C*a*C*a*CV sequence → CC*i*CV (*zaláma* → *zlima/zluma*) or CC*a*CV (*baṣála* → *bṣala*). In the dialects of the western branch – with the exception of the dialect of the Dawāġrah – it does not occur, as appears from forms such as *ragábatih* 'his neck', *zalámatih* 'his man' (Blanc 1970:133; de Jong 2000:461, map 17). In the dialect of the Bilī, traces of this feature are found (de Jong 2000:99). In the dialects of the eastern branch, resyllabified forms commonly occur, e.g. *ilgaḥa* 'pregnant she-camel', *ibduwi* 'Bedouin', perhaps as a result of language contact with speakers of dialects of North Arabian type, but forms like *bágara, waraga*, and *ʿaṭúwi/ʿáṭuwi*, earlier interpreted as K-forms, may be more "genuine", as suggested by *ragabatī́* (correction to Palva 1984a, 1984–1986, 1991; see de Jong 2000:48–50).

(g) The imperfect of the I*w* verbs in Form I are in the dialects of the western branch of the type *yawṣal, yōṣal*, whereas in the eastern branch it mainly follows the type *yāṣal* (Stewart 1990, Glossary s.v. *wṣl*; de Jong 2000, map 56; Palva 1984– 1986:300).

(h) The pronominal suffix of the 3rd person feminine singular is *–ha* throughout the area, with the exception of the Negev, in which it is *-ha/-hiy* (Blanc 1970:130). The masculine form in the dialects of the eastern branch is C*-ah*, in the western C*-ih/-ah*, as a rule phonetically conditioned; in southern Sinai *-u(h)* (de Jong 2000:165, 450; 2003:163; Bani Yasin and Owens 1984:217; Henkin 1985, *passim*; Stewart 1990, *passim*).

(i) In the eastern branch of Northwest Arabian Arabic, the 1st person plural personal pronoun has two variants, *ḥinna* and *iḥna*, whereas the western branch only has *iḥna* or *aḥna*.

(j) In the eastern branch of Northwest Arabian Arabic and in parts of Sinai, *-a* is the main reflex of *-ā(ʾ)* in neutral environments. In the Negev and the eastern part of the northern Sinai littoral, it is *-iy*, in back environments *-a* (Blanc 1970:123–124; de Jong 2000, maps 9, 10).

Differences (a) and (b) are obvious results of developments due to language contact, and consequently do not reflect deeper historical differences. A recent sedentary feature spreading especially in the Negev and among Group II speakers in the northern Sinai littoral, as well as in southern Sinai, is the split-morpheme negation, e.g. *biddiš iyyāk, ma-fišš, fiš mayye kifāye, muš ʿārif, muš mumkin* (Henkin 1985, *passim*); *ma-ǧūš, ma-txāfiš minhi(h), ma-šuftukš, ma-miʿīš* (de Jong 2000:317–318, 2003:174). The use of the analytic genitive markers *tabaʿ, btāʿ*, and *tāʿ* is a further trait of the same kind, e.g. *ilbyūt tabaʿna, aljamal tabaʿah*, even *álmarah tabaʿak, ilmarah tabaʿi* (Henkin 1985, *passim*); *btāʿ* (de Jong 2000:161–162, 281, 366, 449, map 29; 2003:171). An additional trait which has spread from Egypt is the use of the future markers *ḥa-* and *ha-* in Groups II and III (de Jong 2000:319, 394). Such ongoing developments can be expected to increase the typological gap between the western and the more conservative eastern branch. Analogous developments have naturally taken place in the past. The rapidly proceeding dialect leveling is not leading toward a more uniform Northwest Arabian dialect area. Rather, in the four states to which the area belongs, increased central power, sedentarization, and modern communications are bound to direct the development toward the linguistic centers inside the borders of each state.

The generalized *a* in the preformative of the active imperfect of Form I in the dialects of the eastern branch, although following the North Arabian pattern, is probably not due to dialect contact. Rather, it is a conservative feature still leaving a morphological slot for marking the active vs. passive contrast (/a/ vs. /i/). With the loss of productivity of the internally marked passive, the contrast has become neutralized, and the quality of the vowel of the preformative has lost its morphological function. This opens the way to phonetic changes, such as vowel harmony. Thus, the western branch probably represents a later stage of the development, and the eastern branch can be expected to proceed in the same direction.

On the other hand, some features of the dialects of the eastern branch, such as the use of *ḥinna* alongside *iḥna* and the C*a*C*a*CV ⇒ CCVCV resyllabification side by side with non-resyllabified C*a*C*a*CV sequences, can plausibly be attributed to the influence of the Bedouin dialects of the North Arabian type rather than regarded as inherited forms.

3. CONCLUSION

One might ask whether there are sufficient grounds for classifying the dialects of the eastern and western branches together as "Northwestern Arabian Arabic", as suggested by Palva (1991:165–166). In his thorough dialect-geographical study, de Jong (2000:626–630), referring to a relatively high number of typological differences between the two branches, with good reason comes to the conclusion that the existence of such a group is questionable and deserves reconsideration. One of his most weighty arguments is that the dialects of the western branch belong to the 'differential' dialect type, whereas the eastern branch dialects are 'nondifferential'. As referred to above, the dialects of the eastern branch probably are also basically 'differential'. Apart from this case, there are admittedly many typological differences between the two branches, and only on the basis of a detailed investigation of a great number of typologically pertinent criteria can bundles of isoglosses defining the boundaries of the dialect groups be identified.

The present stage of linguistic development in the area clearly shows how dialects change as the result of language contact. The fisherman tribe of the Dawāġrah is a good example of the effects of long isolation. As the only dialect in the Northwest Arabian Arabic area, it has preserved the /n/ in the *-īn* and *-ūn* – final forms in the imperfect, it resyllabifies the C*a*C*a*CV sequence, and it has not adopted the *b-* imperfect. As shown by de Jong, the Negev and the adjacent regions in Sinai constitute the core area of the western branch, while the dialects spoken west of this area, Group II in particular, occupy a middle position between the core area and the dialects of eastern Šarqiyya. In the eastern branch, the situation is not quite analogous to that in the west. The dialects of the Ḥwēṭāt and the Bani ʿAṭīye have certainly been influenced by neighboring North Arabian dialects, but so radical is the typological difference between them and their neighbors in the east and the south that they can scarcely be regarded as forming part of a transition to those dialects in the same sense.

Apart from the dialect of the Dawāġrah, which retains some traits of the North Arabian dialect type, it is virtually impossible to trace differences between the dialects of the different tribes in the Northwest Arabian dialect area on the basis of their different backgrounds or dates of arrival to their present habitats. This implies a long history of a relatively uniform language type in the area, a kind of proto-Northwest Arabian Arabic, which was dominant enough to assimilate the dialects of the newcomers from different directions. Among its typologically most prominent distinctive features, compared to the North Arabian Arabic dialect type, were the lack of the *tanwīn*, the lack of affricated variants of **q* and **k*, and the lack of the final /n/ in the imperfect, 2nd person feminine singular, 2nd person masculine plural, and 3rd person masculine plural. These belong to the most essential classification criteria, which to a significant degree correspond to major historical divisions of Arabic.

Bibliographical references

Abdel-Jawad, Hassan Rashid E. 1981. *Lexical and phonological variation in spoken Arabic in Amman.* Ph.D. diss., University of Pennsylvania.

Abul-Fadl, Fahmi. 1961. *Volkstümliche Texte in arabischen Bauerndialekten der ägyptischen Provinz Šarqiyya mit dialektgeographischen Untersuchungen zur Lautlehre.* Ph.D. diss., University of Munich.

Bailey, Clinton. 1985. "Dating the arrival of the Bedouin tribes in Sinai and the Negev." *Journal of the Economic and Social History of the Orient* 28.20–49.

——. 1991. *Bedouin poetry from Sinai and the Negev: Mirror of a culture.* Oxford: Clarendon Press.

Bani Yasin, Raslan and Jonathan Owens. 1984. "The Bduul dialect of Jordan". *Anthropological Linguistics* 26.202–232.

Bauer, Leonhard. 1926. *Das palästinische Arabisch: Die Dialekte des Städters und des Fellachen.* 4th ed. Leipzig: J.C. Hinrichs'sche Buchhandlung.

Behnstedt, Peter and Manfred Woidich. 1985. *Die ägyptisch-arabischen Dialekte.* I. *Einleitung und Anmerkungen zu den Texten.* II. *Sprachatlas von Ägypten.* Wiesbaden: L. Reichert.

Bergsträsser, Gotthelf. 1915. "Sprachatlas von Syrien und Palästina." *Zeitschrift des Deutschen Palästina-Vereins* 38.169–222 (+ maps XX–LXII).

Blanc, Haim. 1970. "The Arabic dialect of the Negev Bedouins". *The Israel Academy of Sciences and Humanities, Proceedings* 4.112–150.

Burckhardt, John Lewis. 1831. *Notes on the Bedouins and Wahabys, collected during his travels in the East.* London: H. Colburn and R. Bentley.

Cantineau, Jean. 1936. "Études sur quelques parlers de nomades arabes d'Orient". *Annales de l'Institut d'Études Orientales* 2.

——. 1937. "Études sur quelques parlers de nomades arabes d'Orient". *Annales de l'Institut d'Études Orientales* 3.

Doughty, Charles Montagu. 1923. *Travels in Arabia Deserta.* 2 vols. 3rd ed. New York: Boni and Liveright.

Elmgren, Sven Gabriel (ed.). 1864–1866. *Georg August Wallins reseanteckningar från Orienten, åren 1843–1849.* 4 vols. Helsingfors: J.C. Frenckell and Son.

Fischer, Wolfdietrich. 1959. *Die demonstrativen Bildungen der neuarabischen Dialekte: Ein Beitrag zur historischen Grammatik des Arabischen.* The Hague: Mouton.

Glubb, John Bagot. 1948. *The story of the Arab Legion.* London: Hodder and Stoughton.

Henkin, Roni. 1985. *The verbal system in Negev bedouin dialects: Tense and aspect.* Ph.D. diss., Tel-Aviv University. [In Hebrew.]

Ingham, Bruce. 1994. *Najdi Arabic: Central Arabian.* Amsterdam and Philadelphia: J. Benjamins.

Jong, Rudolf E. de 1996. "Examples of levelling and counterreactions in the dialects of the Bedouin tribes in northwestern Sinai". *Égypte/Monde arabe* 27–28.355–382.

——. 2000. *A grammar of the Bedouin dialects of the northern Sinai littoral: Bridging the linguistic gap between the eastern and western Arab world.* Leiden: E.J. Brill.

——. 2003. "Characteristics of Bedouin dialects in southern Sinai: Preliminary observations". *Approaches to Arabic dialects: A collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday*, ed. Martine Haak, Rudolf de Jong, and Kees Versteegh, 151–175. Leiden: E.J. Brill.

Lawrence, T.E. 1935. *Seven pillars of wisdom.* 2nd ed. London: J. Cape.

Musil, Alois. 1908. *Arabia Petraea.* III. *Ethnologischer Reisebericht.* Vienna: Kaiserliche Akademie der Wissenschaften.

——. 1926. *The Northern Ḥeğāz: A topographical itinerary.* New York: American Geographical Society.

Nishio, Tetsuo. 1992. *A basic vocabulary of the Bedouin Arabic dialect of the Jbāli tribe (southern Sinai).* Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.

Oppenheim, Max von. 1943. *Die Beduinen.* II. *Die Beduinenstämme in Palästina, Transjordanien, Sinai, Ḥedjāz.* Leipzig: O. Harrassowitz.

Palva, Heikki. 1965. *Lower Galilean Arabic: An analysis of its anaptyctic and prothetic vowels, with sample texts.* Helsinki.

——. 1976. *Studies in the Arabic dialect of the seminomadic əl-ʿAǧārma tribe (al-Balqāʾ district, Jordan).* Göteborg. Acta Universitatis Gothoburgensis.

——. 1980. "Characteristics of the Arabic dialect of the Bani Ṣaxar tribe". *Orientalia Suecana* 29.112–138.

——. 1984a. "A general classification for the Arabic dialects spoken in Palestine and Transjordan". *Studia Orientalia* 55.359–376.

——. 1984b. "Further notes on the descriptive imperative of narrative style in spoken Arabic". *Studia Orientalia* 55.377–391.

——. 1984–1986. "Characteristics of the Arabic dialect of the Ḥwēṭāt tribe". *Orientalia Suecana* 33–35.295–312.

——. 1989. "Linguistic sketch of the Arabic dialect of el-Karak". *Studia linguistica et orientalia memoriae Haim Blanc dedicata*, ed. Paul Wexler, Alexander Borg, and Sasson Somekh, 225–251. Wiesbaden: O. Harrassowitz.

——. 1991. "Is there a North West Arabian dialect group?". *Festgabe für Hans-Rudolf Singer, zum 65. Geburtstag am 6. April 1990, überreicht von seinen Freunden und Kollegen*, ed. Martin Forstner, I, 151–166. Frankfurt am Main: P. Lang.

——. 1997. "Linguistic observations of the explorers of Arabia in the 19th century". *Built on solid rock: Studies in honour of Professor Ebbe Egede Knudsen on the occasion of his 65th birthday April 11th 1997*, ed. Elie Wardini, 226–239. Oslo: Novus Forlag and Instituttet for Sammenlignende Kulturforskning.

——. 2004. "Remarks on the Arabic dialect of the Ḥwēṭāt tribe". *Jerusalem Studies of Arabic and Islam* 29 (2004), 195–209.

Peake, Frederick G. 1958. *History and tribes of Jordan.* Coral Gables, Fla.: University of Miami Press.

Piamenta, Moshe. 1996. "More on the Arabic dialect of the Negev Bedouins". *Quaderni di Studi Arabi* 14.123–136.

Rabin, Chaim. 1951. *Ancient West-Arabian.* London: Taylor's Foreign Press.

Reichmuth, Stefan. 1983. *Der arabische Dialekt der Šukriyya im Ostsudan.* Hildesheim: G. Olms.

Rentz, G. 1968. "Al-Ḥuwayṭāt". *Encyclopaedia of Islam*, 2nd ed. III, 642–644. Leiden: E.J. Brill.

Rosenhouse, Judith. 1984. *The Bedouin Arabic dialects: General problems and a close analysis of North Israel Bedouin dialects.* Wiesbaden: O. Harrassowitz.

Schreiber, Giselher. 1970. *Der arabische Dialekt von Mekka: Abriß der Grammatik mit Texten und Glossar.* Ph.D. diss., University of Münster.

Sieny, Mahmoud Esmaʿil. 1978. *The syntax of urban Hijazi Arabic (Saʿudi Arabia).* London and Beirut: Librairie du Liban and Longman.

Stewart, Frank Henderson. 1987. "A Bedouin narrative from Central Sinai". *Zeitschrift für Arabische Linguistik* 16.44–92.

——. 1988. *Texts in Sinai Bedouin law.* I. *The texts in English translation.* Wiesbaden: O. Harrassowitz.

——. 1990. *Texts in Sinai Bedouin law.* III. *The texts in Arabic.* Wiesbaden: O. Harrassowitz.

Wallin, George Augustus. 1854. "Narrative of a journey from Cairo to Medina and Mecca, by Suez, Arabá, Tawilá, al-Jauf, Jubbé, Háil, and Nejd, in 1845". *Journal of the Royal Geographical Society* 24.115–207.

Werner, Jürgen. 2003. "Ein arabischer Text aus dem Wadi Ram/Jordanien: Bestandsaufnahme alltäglicher sprachlicher Verwirrung am Beispiel eines beduinischen Sprechers". *Zeitschrift für Arabische Linguistik* 42.53–79.

Yrttiaho, Kaarlo. 1988. "Texts from Arabia Petraea in the dialect of the seminomadic an-Nʿēmāt tribe of the Shara mountains (Jordan)". *Studia Orientalia* 64.145–168.

HEIKKI PALVA (University of Helsinki)

Northwest Semitic Languages

1. THE NORTHWEST SEMITIC LANGUAGES

The term 'Northwest Semitic' is the traditional designation of a group of languages comprising Ugaritic, the Canaanite dialects, and the Aramaic dialects.

Ugaritic is the language of the ancient city of Ugarit (modern Rās Šamra, on the northeast coast of the Mediterranean in Syria). The roughly eleven hundred Ugaritic texts are written in an alphabetic cuneiform script on clay tablets; unlike other Semitic alphabets, the Ugaritic script reads from left to right. In addition to several important mythological texts, there are also some hundred letters, a few legal documents and treaties, and several hundred administrative texts. Most of the texts date to the 13th/12th centuries B.C.E., although some of the literary texts were probably written earlier. The alphabetic script indicates consonants only, although there are three signs for *aleph* (*hamza*), each of which indicates the glottal stop followed by a different vowel quality (e.g. <ŠMAL> for /šimʾālu/ 'left'; <ŠIL> for /šaʾila/ 'he asked'; <RPUM> for /rāpiʾūma/ 'healers' [nom.]). The standard reference grammar of Ugaritic is Tropper (2000).

The best-known form of Canaanite is Hebrew. The earliest Biblical Hebrew texts probably date to about the 12th century B.C.E., and the latest to the 2nd century B.C.E. There are also many Hebrew inscriptions, most of them quite short, beginning in the 10th century B.C.E.; the grammar of these is nearly identical to that of Biblical Hebrew, although a number of geographical variations are attested. Developments in the grammar over the centuries also occur in both biblical and inscriptional texts. From the 2nd century B.C.E. to the 5th century C.E. is the period of Mishnaic Hebrew, which reflects a spoken dialect not directly descended from Biblical Hebrew. Hebrew continued to be written in the medieval and early modern periods, and – a phenomenon unique among the world's languages – was revived as a spoken language in the 19th century (→ Ivrit). Hebrew vocalism is known from the use of vowel letters (*matres lectionis*), which began early in the history of the written language, and from vowel points that were added to consonantal texts beginning in the late 1st millennium C.E. Reference grammars of Biblical Hebrew include Gesenius a.o. (1910), Bauer and Leander (1922), Waltke and O'Connor (1990), and Joüon (1991).

Another well-known form of Canaanite is Phoenician, the language of the city-states of the eastern Mediterranean coast, such as Tyre, Sidon, and Byblos. The earliest texts, from Byblos, date to the beginning of the 10th century B.C.E., while the latest are from the 2nd century C.E. The dialect of the texts written in Phoenician colonies established around the Mediterranean and beyond, especially that of Carthage (Phoenician /qart ḥadašt/ 'new city') is referred to by scholars as Punic (from the 5th century B.C.E.). After the fall of Carthage, texts are said to be written in Neo-Punic (attested until the 5th century C.E.). The Phoenician alphabet is purely consonantal; there is no indication of vowels until late in the Punic period. The standard reference work on Phoenician is Friedrich, Röllig, and Amadasi Guzzo (1999).

Another Canaanite dialect is Moabite, attested almost solely in the long 9th-century inscription of the Moabite king Mešaʿ. Still other Canaanite dialects, attested in a few short inscriptions only, are Ammonite and Edomite (9th/6th centuries B.C.E.; see Parker 2002).

The earliest attestation of Canaanite is found in a group of several hundred letters sent by vassal rulers in cities in Syria-Palestine (including Byblos, Tyre, Jerusalem, Shechem, Gezer, Ashkelon) during the first half of the 14th century B.C.E., to their Egyptian suzerain. The texts were found in excavations at Akhetaten, a short-lived capital of the Egyptian Empire under King Akhenaten (Amenophis IV), modern el-ʿAmārna. While ostensibly written in Akkadian, the lingua franca of the period, the letters betray a great deal of the scribes' native Canaanite language, especially in the morphosyntax of the verb, which is strikingly similar to that of Classical Arabic. Glosses of actual Canaanite words are also frequently encountered in these texts. The still-standard critical edition of the texts is Knudtzon (1907–1915); the grammar is described in Rainey (1996).

Aramaic is first attested in inscriptions on stone monuments dating to the 9th century B.C.E. These early inscriptions, down to the 6th century B.C.E., are collectively referred to as Old Aramaic, an umbrella term, since the inscriptions reveal a variety of grammatical idiosyncrasies. The inscriptions are written in the 22-letter Phoenician alphabet; the usage of the consonants indicates that many of the consonantal mergers that characterize later Aramaic had not yet occurred in this early period. During the Achaemenid period, Aramaic became one of the official languages of the Persian chancery, a factor in its spread across the Near East as a lingua franca for much of the next millennium. Official (or Imperial) Aramaic, as this phase is called, comprises a large number of letters and legal documents, most found in Egypt, as well as the Aramaic of the Biblical book of Ezra. After the fall of the Persian Empire, in which Aramaic texts exhibit a relatively uniform standard language, Aramaic begins again to show dialectal diversity; in the so-called Middle Aramaic period (3rd century B.C.E.-2nd century C.E.), dialects include Palestinian (in the Biblical book of Daniel, in texts from Qumran, and in certain *targum*s, i.e. Aramaic translations of Biblical texts), Nabataean, Palmyrene, and Hatran. In Late Aramaic (3rd-9th centuries C.E.), a division into eastern and western forms of the language is evident. Western Late Aramaic includes the large corpus of Jewish *targum* and *talmud* texts, a significant number of Christian texts in a dialect known as Christian Palestinian, and the Aramaic of the small Samaritan sect. Eastern Late Aramaic includes the text of the Babylonian Talmud, as well as Mandaic. Syriac, the dialect of an enormous corpus of Christian texts, is variously considered to be an Eastern dialect or a separate branch. Aramaic continues to be spoken by more than one hundred thousand people, especially the many dialects known collectively as Northeastern → Neo-Aramaic (from Iran, Iraq, and northeastern Syria) but also Ṭūrōyo in Turkey, and the Western dialects of the villages of Maʿlūla and vicinity near Damascus. A comprehensive Aramaic grammar has not been written; for Old Aramaic, see Degen (1969) and Hug (1993); for Imperial Aramaic, Folmer (1995) and Muraoka and Porten (2003); for Biblical Aramaic, Bauer and Leander (1927) and Rosenthal (1995); for Nabataean, Cantineau (1930–1932); for Palmyrene, Cantineau (1935); for Mandaic, Macuch (1965); for Syriac, Nöldeke (1904); for other late varieties of Aramaic, see, for example, Dalman (1905), Müller-Kessler (1991), and Macuch (1982).

In Akkadian (and a few Egyptian) texts of the late 3rd and early 2nd millennia B.C.E. personal names are found that do not conform to the usual pattern of Akkadian personal names. The individuals bearing such names are often called Amorites. Many of the names exhibit features of the later Northwest Semitic languages (in phonology, e.g. in the name *ia-qa-rum* = /yaqarum/ 'esteemed', with initial *y*, vs. Arabic *waqara* and Akkadian *waqārum*; in morphology, e.g. preterite verbs with initial *ya* rather than *i* as in Akkadian, such as *ya-šu-ub-*^{d}*da-gan* = /yaθub-dagan/ '[the god] Dagan has returned'). In view of the great chronological and geographical spread of the relevant texts, it is likely that a number of early dialects are reflected in these Amorite names. A recent study is Streck (2000).

2. FEATURES OF NORTHWEST SEMITIC LANGUAGES

The following paragraphs review some of the linguistic features that distinguish and/or connect Arabic and the Northwest Semitic languages. In order to compare the Northwest Semitic languages with Arabic, we have to be aware of the linguistic features that were already present at a Proto-Northwest Semitic stage and those that ought to be considered innovations within the individual subbranches of Northwest Semitic. The following comparison is mainly based on Classical Arabic, although evidence from Ancient North Arabian, Old Arabic, and Arabic dialects is included where appropriate. For the relationship of these subgroups to Classical Arabic, see, for example, Rabin (1951:3) for Ancient North Arabian and Old Arabic, Fischer (1995) for modern Arabic dialects.

2.1. *Phonology*

The consonantal inventory of Proto-Northwest Semitic has to be reconstructed with all 29 Proto-West Semitic consonants.

In Ugaritic, many of the Proto-Semitic consonants remain distinct phonemes. Two unconditional mergers took place: **ś* and **s* merged to *š*; and *^{t}*ṣ* and **ṣ́* merged to *ṣ* (e.g. <ṢǴR> 'small'; <ARṢ> /ʾarṣu/ 'earth' vs. Classical Arabic *ʾarḍ*). **θ̣* was often preserved; it is transcribed *ẓ*, as in <ẒBY> /ẓabyu/ 'gazelle', although it sometimes merged with **ġ*, as in <NĠR> /naġara/ 'he guarded' (cf. Classical Arabic *naḏ̣ara* 'he watched'). The Proto-Semitic sound **ð* had merged with **d* to /d/ in most cases (e.g. <UDN> /ʾudnu/ 'ear'), although the alphabet still had a separate letter for /ð/ (e.g. <ḎRʿ> /ḏirāʿu/ 'arm').

In Hebrew (and Phoenician), **ð* and **z* merged to *z*: **ʾuðn-* > *ʾṓzen* 'ear'; **ḥ* and **x*

Table 1. Consonants in the Northwest Semitic languages

Proto-Semitic and Proto-Northwest Semitic	Ugaritic	Hebrew		Phoenician	Aramaic	Arabic
*ð	d / ḏ	z		z	d (OA <Z>)	ḏ
*ġ (γ)	ġ	ʿ		ʿ	ʿ	ġ
*ḥ (ħ)	ḥ	ḥ		ḥ	ḥ	ḥ
*ḫ (x)	ḫ	ś	ḥ	ḥ	ḥ	x
* ś ¢	š	ṣ	ś	š	s (OA <Š>)	š
*¢'	ṣ	š	s	ṣ	ʿ (OA <Q>)	ḍ
*s	š	s	š	š	š	s
*^{t}s	s	ṣ	s	s	s	s
*^{t}ṣ	ṣ	š	ṣ	ṣ	ṣ	ṣ
*θ	ṯ	ṣ	š	š	t (OA <Š>)	ṯ
*θ̣	ẓ / ġ	z	ṣ	ṣ	ṭ (OA <Ṣ>)	ḏ̣
*^{d}z	z		z	z	z	z

merged to *ḥ*: **ʾax- > ʾāḥ* ‘brother’; *ʿ and **ġ* merged to ʿ: *^{t}*ṣaġīr- > ṣāʿīr* ‘small’; **s* and **θ* merged to *š*: **s(i)m- > šēm* ‘name’; **θalāθ- > šālōš* ‘three’; **ṣ́*, *^{t}*ṣ*, and **θ̣* all merged to *ṣ*: **ʾarṣ́- > *ʾarṣ > ʾéreṣ* ‘earth’; *^{t}*ṣidq- > ṣéḏeq* ‘righteousness’; **naθ̣ara > nāṣar* ‘he guarded’. Common Semitic **ś* remained distinct for most of the early history of Hebrew, but under Aramaic influence it eventually merged with *s* (< *^{t}*s*). Since **ś* had merged with **s* and **θ* to *š* in Phoenician, from which Hebrew scribes borrowed their alphabet, there was no special letter to represent the still-distinct *ś*, and so the letter for *š* was used; later the two consonants were distinguished by diacritical dots.

The orthography of early inscriptions indicates that nearly all Proto-Semitic consonants remained distinct in most Old Aramaic dialects. Consonants that were later lost through mergers were represented with the closest letter available in the borrowed Phoenician alphabet. In Official Aramaic and in later Aramaic dialects, however, a large number of mergers occurred (examples from Syriac): *ð merged with **d* and **x* merged with **ḥ*: **ʾaxaða > ʾeḥad* (Old Aramaic <ʾḤZ>) ‘he seized’; **ś* merged with *^{t}*s* to *s*: **śāma > sām* (Old Aramaic <ŠM>) ‘he placed’; **θ* merged with *t*: **θalāθ- > tlāt* (Old Aramaic <ŠLŠ>) ‘three’; **ġ* merged with ʿ: **baġaya > bʿā* (Old Aramaic <BʿH> ‘he sought’); **θ̣* merged with *ṭ*: **naθ̣ara > nṭar* (Old Aramaic <NṢR>) ‘he guarded’; the reflex of **ṣ́* in Old Aramaic was written with <Q> but later merged with ʿ: Old Aramaic <MRQ>, Syriac *mraʿ* ‘he was sick’ (cf. Classical Arabic *mariḍa*).

Table 1 provides an overview of the main mergers and differences in the individual Northwest Semitic languages.

Classical Arabic and most Ancient North Arabian dialects had a consonantal inventory of 28 consonants, following the merger of Proto-Semitic *^{t}*s* and **s* to *s*. The Ancient North Arabian dialect of Tayman is exceptional in having preserved all three original voiceless fricatives (Macdonald 2004:499). Although Ancient North Arabian is not the direct ancestor of what is called ‘Arabic’ today, it nevertheless might be related to pre-Islamic West-Arabian (Rabin 1951:2). Should the dialect of Tayman be related to later ‘Arabic’, the preservation of **s*, *^{t}*s*, and **ś* could indicate that Proto-Arabic had all 29 Proto-West Semitic consonants.

A significant sound change that distinguishes Northwest Semitic from other Semitic languages is the shift of word-initial *w* to *y*, as in Hebrew *yéleḏ*, Syriac *yaldā*, Ugaritic <YLD> ‘child’, vs. Classical Arabic *walad*. In the Ancient North Arabian dialect Safaitic, we also find evidence for the occasional change of initial *w* > *y*, as in *yrḫ* ‘month’, which occurs as a biform of *wrḫ*. This change is rare, though, and not limited to word-initial position. Thus, it should not be evaluated as the same isogloss found in Northwest Semitic.

Another sound change that has sometimes been postulated for Northwest Semitic is the assimilation of *n* to a following consonant, which is attested in Ugaritic and Canaanite. Biblical Aramaic occasionally preserves original *n* before a consonant, as, for example, in *ʾant(h)* ‘you’ (2nd pers. masc. sg.) (Rosenthal 1995:23), while Syriac has reflexes of preconsonantal *n* in its orthography, although *n* is not pronounced in such cases. These attestations suggest a preservation of this phoneme in preconsonantal position until after the split of Aramaic from the rest of Northwest Semitic. Thus, *n* was probably not assimilated at the earliest stages of Northwest Semitic. Classical Arabic and most modern Arabic dialects, of course, do not assimilate *n* to a following consonant; in nearly all Ancient North Arabian dialects, however, with the exception of Ḥaṣaitic, *n* assimilates regularly (Macdonald 2004:501), as it does, less consistently, in the Old South Arabian dialects.

The reflex of Proto-Semitic **p* remains a stop, /p/, in Canaanite and Aramaic, although it has a postvocalic fricative allophone in both language groups, as do all nondoubled, nonemphatic stops: Hebrew *pāṯaḥ* and *yiftaḥ*, Syriac *pṯaḥ* and *neftaḥ* ‘he opened’, ‘he will open’, vs. Classical Arabic *fataḥa*, *yaftaḥu*.

The Proto-Northwest Semitic vowels can be reconstructed as in Proto-Semitic, viz. three basic short vowels, three long vowels, and two diphthongs, just as preserved in Classical Arabic:

**a, *i, *u*
**ā, *ī, *ū*
**aw, *ay*

The consonantal nature of the Ugaritic script precludes a detailed knowledge of the vowel system. Evidence of writings in syllabic cuneiform texts, and the use of the three *aleph* symbols, however, show a number of developments. The original diphthongs **aw* and **ay* contracted unconditionally to *ō* and *ē*, respectively: <MT> /mōtu/ 'death'; <BT> /bētu/ 'house'. There is vowel harmony around gutturals: /ṭuhūru/ < **ṭahūru* 'pure'; /tahāmatu/ < **tihāmatu* 'sea'; and in the patterns CvC-Cv̄C, as in Arabic: /ṣibbīru/ < **ṣabbīru* 'collective land'; <ULP> /ʾullūpu/ < **ʾallūpu* 'leader'. There is also evidence of the sporadic syncope of short, unaccented vowels: /nabakīma/ and /nabkīma/ 'springs'.

For Phoenician, because of the strictly consonantal orthography, we are dependent on transcriptions into other languages for any evidence concerning the vowel system. Such transcriptions show that in Phoenician, as in Ugaritic, the original diphthongs **aw* and **ay* contracted universally to /ō/ and /ē/ respectively: thus /mōt/ 'death' and /bēt/ 'house'. A characteristic of all Canaanite languages is the change of Semitic **ā* to *ō* (considered by some scholars to be dependent on stress, by others to be unconditioned): (Greek transcription) *ozer* for [ʿōzir] < **ʿāðir-* 'helper'. A specifically Phoenician development was the shift of an original short Semitic **a* to *o* in accented syllables: (Greek transcription) *labon* for [labon] < **laban-* 'white'.

Biblical Hebrew is characterized by a large number of developments in the vowel system. Diphthongs collapsed when unstressed but were triphthongized under the stress: *bēṯī* 'my house' but *bayiṯ* 'house' (< **bayt(-ī)*); *mōṯī* 'my death' but *māweṯ* 'death' (< **mawt(-ī)*). As in Phoenician and other Canaanite languages, original long **ā* became *ō*, as in *kōṯēḇ* < **kātib-* 'writer'. Short vowels underwent complex developments: lost word-finally; lowered or backed under the stress or in open syllables immediately before the stress: *dāḇār* < **dabar-* 'word', *yittēn* < **yittinu* 'he gives', *yiḵtōḇ* < **yiktubu* 'he writes'; reduced to *ə* or zero in open syllables otherwise, and generally in finite verbs: *dəḇārīm* < **dabarīma* 'words', *yittənū* < **yittinū* 'they give', *yiḵtəḇū* < **yiktubū* 'they write'. Words of the shape CvCC underwent anaptyxis: **ʾarṣ́ -* > **ʾarṣ* > *ʾéreṣ* 'earth'; **sipr-* > *sḗp̱er* 'book'; **quds-* > *qṓḏeš*. Still other developments occurred in the neighborhood of the guttural consonants (ʾ, *h*, *ḥ*, ʿ).

A number of vowel changes may be said to characterize Proto-Aramaic. As in Hebrew (and most Northwest Semitic languages), short final vowels were lost early, including the singular case vowels: **kātibu* > **kātib* (> *kāṯeḇ*) 'writer'; **kātibīna* > **kātibīn* (> *kāṯḇīn*) 'writers'; **tismaʿūna* > **tišmaʿūn* (> *tešmʿun*) 'you [masc. pl.] hear'. Resulting final consonant clusters were resolved by epenthesis: **katabtu* > **katabt* > **katabit* (> Targumic *kəṯáḇiṯ*; Syriac *keṯbeṯ*) 'I wrote'. A pervasive feature of Aramaic is the reduction of short vowels in open syllables: **kataba* > **katab* > *kəṯaḇ* 'he wrote'; **katab+ih* > *kaṯḇeh* 'he wrote it' (see also the preceding examples). Original long vowels generally remained unchanged.

2.2. *Morphology*

2.2.1. Pronouns

The Northwest Semitic system of personal independent pronouns and pronominal suffixes is similar to that of Classical Arabic, although only in Ugaritic are dual forms of the pronouns

Table 2. Independent personal pronouns

	PNWS	Ugaritic	Hebrew	Syriac	Arabic
1cs	**ʾanā(±ku)*	/ʔanāku/, <AN>	*ʾānōḵī*, *ʾᵃnī*	*ʾanā*	*ʾana*
2ms	**ʾanta*	<AT>	*ʾattā*	*ʾa(n)t*	*ʾanta*
2fs	**ʾanti*	<AT>	*ʾatt*	*ʾa(n)t(y)*	*ʾanti*
3ms	**huʾa*	/huwa/	*hū(ʾ)*	*hu(ʾ)*	*huwa*
3fs	**hiʾa*	<HY>	*hī(ʾ)*	*hi(ʾ)*	*hiya*
3cd	*?	<HM>	—	—	*humā*
1cp	**naḥnu*	?	*ʾᵃnaḥnū*	*ḥnan*	*naḥnu*
2mp	**ʾantum(±ū)*	<ATM>	*ʾattem*	*ʾa(n)ton*	*ʾantum*
2fp	**ʾantin(±na)*	?	*ʾatten(nā)*	*ʾa(n)ten*	*ʾantunna*
3mp	**hum(±ū)*	<HM>	*hēm(mā)*	*hennon*	*hum*
3fp	**hin(±na)*	<HN>	*hēnnā*	*hennen*	*hunna*

attested (vocalization unknown); Ugaritic even has a 1st person dual suffix (<-NY>, as in <BʿLNY> 'the lord of the two of us').

As the table indicates, the main differences lie in the 2nd person and 3rd person plural forms; while Ugaritic and Hebrew have *m* vs. *n* for the masculine/feminine contrast, as in Classical Arabic, the Hebrew forms exhibit an *i* vowel vs. the *u* of Classical Arabic. In most Aramaic dialects, the masculine/feminine contrast is in the vowels (*u/o* for masc. vs. *i/e* for fem.), the *n* having been leveled through all forms. In the 3rd person singular forms, it is likely that the *aleph* of the Canaanite and Aramaic forms is original, and developed into the glides *w* and *y* in Ugaritic, as in Classical Arabic, by assimilation to the preceding vowel.

The common Semitic longer biform of the independent 1st person singular pronoun, *ʾ*anāku* (cf. Akkadian *anāku*), is found in Ugaritic and in the Canaanite languages (e.g. Hebrew *ʾānōḵī*) alongside the shorter form, which alone has been preserved in Aramaic, as in Arabic.

Ugaritic and Phoenician also have independent 3rd person forms with final <-T>, such as 3rd person masculine singular <HWT> (/huwāti/??); in Ugaritic these are oblique (genitive-accusative) forms.

The Northwest Semitic pronominal suffixes were originally quite similar in form to those of Classical Arabic and were added to nouns (to indicate possession) and to verbs (to indicate objects) much as in Classical Arabic, i.e. simply attached to the bound form of a noun (or preposition) and to the end of finite verb forms. Phonological and analogical developments, however, have produced complex changes in the forms of the suffixes in both Canaanite and Aramaic.

It is possible to reconstruct a common Proto-Northwest Semitic determinative-relative pronoun that was declined for case, number, and gender, **ðV̄*. For Ugaritic this pronoun has feminine and plural forms with <-T> (vocalization unknown): masculine singular nominative <D> (/dū/), feminine singular <DT>, masculine plural <DT>. In Biblical Hebrew, the pronoun is only vestigially preserved in a few examples as indeclinable (old nominative) *zū* (this was generally replaced in Hebrew by the form *ʾᵃšer*, grammaticalized from a noun meaning 'place', cognate with Classical Arabic *ʾaṯar* 'trace'); a relative <Z> also occurs in some Phoenician dialects. In Aramaic, the reflex is the common indeclinable relative and genitive marker *dī* in Biblical Aramaic (the old genitive; written <ZY> in Old Aramaic inscriptions), which became *d(ə)-* in Syriac. There are still traces of the same fully declined determinative-relative pronoun **ðV̄* in Classical Arabic, although its semantic range shifted to the expression of possession. In Ancient North Arabian, this form of the relative pronoun is still regularly used in the form *ḏ* (masc. sg.) and *ḏʾt* (fem. sg.) (Macdonald 2004:508), while Yemenite has an undeclined form *ḏī* (Rabin 1951:39).

There is no common Proto-Northwest Semitic demonstrative pronoun, although all languages use the same basic constituents to form demonstratives. Ugaritic has a near deixis pronoun *hnd*, probably vocalized as /hanādu/, and, perhaps, a far deixis pronoun *hnk* /hunāka?/ (Tropper 2000:229–231). Especially the latter resembles the Classical Arabic word for 'there' *hunāka*, while the first resembles Classical Arabic *hunā* 'here' + demonstrative element *ðV̄. In Canaanite and Aramaic, near deixis is expressed by an element *ðV̄ in the singular, as in Hebrew *ze* 'this' (masc. sg.), *zō(ʾ)ṯ* (fem. sg.). Individual languages can add particles to this base, such as Targumic Aramaic *hāden* (masc. sg.) and *hādā* (fem. sg.), but these seem to be secondary additions, since, for example, Biblical Aramaic has a feminine singular without prefix, *dā* 'this'. Near deixis in the plural is expressed by a basic particle *ʾ*ill*V, as in Hebrew *ʾēlle* (comm. pl.) and Biblical Aramaic *ʾillēn* (comm. pl.). Classical Arabic employs the same basic particles as Hebrew and Aramaic, although in a slightly different distribution, for example *hāḏā* for the masculine singular and not, as in Targumic Aramaic, the feminine singular. The main difference between the Northwest Semitic and the Classical Arabic forms of the demonstrative pronoun is in the plural base. In contrast to the Hebrew and Aramaic base *ʾ*ill*V, with *i* and double *-ll-*, the Classical Arabic base has *u* and single *-l-*, as in *hāʾulāʾi*. Modern Arabic dialects either do not use this base (e.g. Moroccan, Tunisian), or they have a base that seems to correspond to Classical Arabic (Lebanese *hawlīk* < *hāʾulāʾi+k*). Since this base is not attested in Ugaritic, it is impossible to determine whether this vowel alternation is Proto-Northwest Semitic or not.

For far deixis, Biblical Aramaic uses the same base *ðV̄ with the addition of final *-k*: *dēk̲* 'that' (masc. sg.), *dāk̲* (fem. sg.), *ʾillēk̲* (comm. pl.). Other Aramaic dialects, such as Syriac, and Hebrew exclusively use the 3rd person pronouns as anaphoric elements to express far deixis, as in Hebrew *hāʾ-īš ha-hū(ʾ)* 'that man'. Classical Arabic is unusual among the Semitic languages in no longer employing the 3rd person pronouns anaphorically, using instead a full set of demonstratives, formed with the demonstrative elements for near deixis plus *-k*, as in *ḏālika* (masc. sg.), *ʾulāʾika* (comm. pl.). The feminine singular is an exception in that it has initial *t* instead of expected *ḏ*, *tilka*. Ancient North Arabian offers little evidence for demonstrative pronouns, but it seems that Dedanitic had a form *ḏ (h)* as well (Macdonald 2004:509; Müller 1982:20).

2.2.2. Nouns

Proto-Northwest Semitic undoubtedly had the same triptotic noun declension in the singular as Classical Arabic (see Table 3). The full declension, however, is preserved only in Ugaritic and in the Canaanite reflected in the Amarna Akkadian texts, later Canaanite and Aramaic dialects having lost case distinctions with the general loss of final short vowels (like later Arabic dialects). Certain nominal forms with pronominal suffixes in Hebrew and Aramaic still reflect the original case endings, such as Hebrew *dəb̲ārēk̲* 'your [fem. sg.] word' < **dabari-ki* [gen.], *malkō* 'his king' < **malku-hu* [nom.], Syriac *malkeh* 'his king' < **malkihi* (gen.; this last form is common Aramaic and reflects a frozen allomorph of the 3rd pers. masc. sg. suffix alternation *-hu/-hi* found in Classical Arabic). There is evidence in Ugaritic for a diptotic declension in the singular in some proper nouns (especially personal names ending in /-V̄n/; see Liverani 1963), but its status for Proto-Northwest Semitic is uncertain. The declension of the dual and plural was diptotic as in Classical Arabic (→ diptosis).

Hebrew and Ugaritic also preserve a directive ending *-h*, as in Ugaritic <ARṢH> 'to the ground', Hebrew *hāʿīrāh* 'to the city', which is probably related to the Akkadian terminative ending *-iš*. This morpheme is not attested in any form of Arabic.

All Northwest Semitic languages lost → mimation/→ nunation in the singular and in the (external) feminine plural. This loss might thus be reconstructed as Proto-Northwest Semitic (e.g. Proto-Semitic **ṭābum* > Proto-Northwest Semitic **ṭābu* 'good [masc. sg. nom.]'; Proto-Semitic **ṭābātum* > Proto-Northwest Semitic **ṭābātu* 'good [fem. pl. nom.]'), although the possibility that it was an independent development within the individual Northwest Semitic branches cannot be ruled out, since loss of mimation/nunation is a widespread phenomenon, which, for example, also occurred in post-Old Babylonian Akkadian and in Geʿez (and, of course, in most modern Arabic dialects).

In Northwest Semitic languages, the plural is normally indicated by external markers. For the masculine plural, corresponding to the final *-na* of Classical Arabic (as in *fallāḥūna*), some Northwest Semitic languages likewise exhibit *-n(a)* (e.g. Aramaic, some Canaanite dialects such as Moabite and Mishnaic Hebrew), while others exhibit *-m(a)* (Ugaritic, other Canaanite dialects such as Phoenician and Biblical Hebrew). Syllabic cuneiform evidence indicates that in Ugaritic the ending was /-ma/ on masculine plurals. Only Ugaritic retains case distinctions; as in Classical Arabic, these are nominative in /-ū/ and oblique (genitive-accusative) in /-ī/ (e.g. nom. /yāṣirūma/ 'potters', oblique /diprānīma/ 'junipers'). In the other Northwest Semitic languages, with the loss of case distinction in the singular (see above), the oblique form was generalized in the masculine plural; further, the final *a* of the ending **-na/-ma* also disappeared with the general loss of short final vowels; thus, we find, for example, Hebrew *ṭōb̲īm*, Aramaic *ṭābīn* 'good' < oblique **ṭāb-ī-ma/na*. Vestiges of broken plurals are rare, but they are found in nearly every subbranch of Northwest Semitic (Ratcliffe 1998:99). The most striking exception, and one of the major distinguishing isoglosses of Northwest Semitic, is the double marking of singular nouns of the pattern **qVtl-* in the plural with inserted *a* after the second radical in addition to external plural markers (Huehnergard 1991:284), as in Hebrew *məlāk̲īm* 'kings' < **malakīma* and Ugaritic /malakūma/ 'kings' from the singular base **malk-*. Vestiges of this *a*-insertion are also found in Syriac (Nöldeke 1904:63). In Classical Arabic double marking is also occasionally found, but it is not mandatory; cf. *laʿnat-* 'curse', pl. *laʿanāt-* vs. *ʾahl-*, pl. *ʾahlūna* 'people'.

The feminine singular is marked with either

-t or *-at*. While in Classical Arabic the latter ending has been generalized (except for a few words, such as *bint* and *ʾuxt*), in Hebrew, the choice of ending appears to be lexical: e.g. *deleṯ* < **dal-t-* 'door' vs. *ʾāmā* < **ʾam-at-* 'maidservant'. (This seems to be true in Ugaritic as well.) In Aramaic, the choice of ending was probably determined phonologically originally, **-at* appearing after bases ending in a consonant cluster (and sometimes after bases ending in v̄C), **-t* appearing otherwise. In both Hebrew and Aramaic, the original ending **-at* became *-ā* phrase-finally (i.e. when not in construct or followed by a pronominal suffix), as in Hebrew *malkā* 'queen', bound form *malkaṯ*, both < **malkat-* and *malkāṯī* 'my queen' < **malkat-ī*; cf. post-Classical Arabic dialects. (In Phoenician, however, the final *t* remained in all forms.) Other feminine endings that are found in Arabic, such as *-ā (-ay)*, are attested only vestigially in a few Northwest Semitic languages (e.g. Syriac *salway* 'quail').

The feminine plural is marked by the ending *-ātu* (nom.)/*-āti* (oblique), corresponding to Classical Arabic *-ātun*/*-ātin*, but again with loss of the final mimation/nunation. In Ugaritic and Amarna Canaanite, this ending is unchanged. In later Northwest Semitic languages, once again, the final case-vowel has been lost; thus, we find Aramaic *ṭāḇāṯ* 'good', while in Canaanite the change of **ā* to *ō* yields *-ōt*, as in Hebrew *ṭōḇōṯ*. Plurals of feminine nouns of the pattern *qVtlat* also exhibit *a*-insertion, as in Hebrew *məlāḵōṯ* 'queens' < **malakātu*.

The dual is fully productive in Ugaritic; the endings are nominative /-āma/ (also /-āmi/), oblique /-ēma/ (also /-ēmi/; from **-ayma/i*), as in nominative /mašlaḥāma/ 'two garments'. In Biblical Hebrew the dual is restricted to words denoting objects that naturally occur in pairs and a few time words; as in the masculine plural, the original oblique ending has been leveled, with **-aymV* > *-áyim*: *yāḏáyim* '(two) hands', *yōmáyim* 'two days'. In Aramaic the dual appears on a few forms in the early dialects (e.g. Biblical Aramaic *qarnáyin* '(two) horns' < **qarnaynV*), but in later dialects it has ceased to be used.

Genitive expressions are constructed as in Arabic, with the *nomen rectum* standing in a bound (construct) form. In Ugaritic, the bound form of singular nouns was probably not distinguished from the absolute form (the latter having lost the original final mimation). In Hebrew and Aramaic, different stress patterns in bound forms frequently resulted in allomorphism, as in Hebrew *bāqār*, bound form *bəqar* 'cattle' (cf. Classical Arabic *baqar*). As noted above, feminine nouns that originally ended in **-at* lost the *t* when not bound: *šānā*, bound form *šənaṯ* 'year' (cf. Classical Arabic *sana(tun)*). Masculine plural and dual bound forms in Ugaritic lost the final *-ma*/*-mi* (e.g. /maqqaḥā/ 'tongs of'), just as the final *-na*/*-ni* is lost in such forms in Arabic. Likewise the final *-n* or *-m* is lost in later Northwest Semitic languages in masculine plural and dual bound forms, as in Hebrew dual *ʿēnáyim*, bound form *ʿēnē* 'eyes' (< **ʿayn-ay(-mV)*). Curiously, in both Aramaic and Hebrew bound forms of masculine plurals, the expected final *-ī* is replaced by what appears to be the dual ending, *-ē*, as in Aramaic (Biblical) *ʾᵉlāhīn* 'god(s)', bound form *ʾᵉlāhē*.

Table 3. Proto-Northwest Semitic noun declension (masc. sg.)

	singular	dual	plural
nominative	**malku*	**malkān/ma/i*	**malakūn/ma*
genitive	**malki*	**malkayn/ma/i*	**malakīn/ma*
accusative	**malka*	**malkayn/ma/i*	**malakīn/ma*

It is unlikely that Proto-Northwest Semitic had a definite article, since Ugaritic does not contain evidence for such a morpheme. The different forms of the definite article in Canaanite and Aramaic also indicate that these two branches of Northwest Semitic underwent independent developments: in Canaanite, the definite article *ha-* is prefixed to the nominal base with gemination of the following consonant, as in Hebrew *ham-meleḵ* 'the king'; in Aramaic, the definite article *–ā* (originally *-aʾ*) is suffixed, as in *malkā* 'the king'. While all Northwest Semitic definite articles are perhaps to be derived from the same original particle, **han-* (Rubin 2004), Classical Arabic employs a different base for the definite article, *(ʾ)al-*, with the well-known variation *(ʾ)aC-* when followed by a coronal 'sun letter'. This article is already attested in the earliest evidence of Old Arabic (Macdonald 2000:50). Ancient North Arabian, on the other hand, employs the definite article *h(n)-*, a form that is much closer to the Canaanite definite article than to Old Arabic and Classical Arabic.

Yemenite has yet another form of the definite article, *am-*, which might be a phonemically altered form of *h(n)-* (Rabin 1951:34–36).

The syntax of the definite article in the Northwest Semitic languages that attest it is virtually identical to that of the Arabic article: the article may appear only on the last member of a genitive chain; it may not appear on a noun with a pronominal suffix; it does not appear on predicate adjectives.

2.2.3. Verbs

For the Proto-Northwest Semitic and Arabic finite verbal system, one suffix conjugation and several prefix conjugations may be reconstructed:

**qatala*	
**yaqtul*	pl. **yaqtulū*
**yaqtula*	pl. **yaqtulū*
**yaqtulu*	pl. **yaqtulūna*
‘energic’:	**yaqtulVn(n)a*

The suffix conjugation generally denotes the past tense in the Northwest Semitic languages, as in Arabic. In the earliest Northwest Semitic, such as Ugaritic and Amarna Canaanite, it seems to alternate with **yaqtul* for the past. The latter form, **yaqtul*, is the Proto-Semitic perfective form par excellence. As in Arabic, it is used as a jussive in all Northwest Semitic languages (though usually without a preposed *li-*). It also continues to be used as a past tense: in Ugaritic and Amarna Canaanite, as just noted; in Classical Hebrew with preposed conjunction *wa-*, as in *wayyiḵtōḇ* ‘and he wrote’ (also frequently in early poetry without the preposed *wa-*); similarly in a few examples in the earliest Aramaic inscriptions. The form **yaqtula* is used in Amarna Canaanite as a virtual equivalent of the jussive **yaqtul* (Amarna scholars usually term it the ‘volitive’); in Ugaritic its function is debated, but it seems to be roughly similar to its use in Amarna Canaanite; in Hebrew it has been reduced in range to 1st person forms, called the ‘cohortative’, likewise injunctive in meaning, as in *ʾeḵtəḇā* ‘let me write’; in Aramaic **yaqtula* has disappeared. The form **yaqtulu* is imperfective in Northwest Semitic, as in Arabic, used as a future, a present, a past habitual, and a circumstantial. As in Arabic, the masculine plural of **yaqtulu* in Amarna Canaanite and in Ugaritic ends in *-ūna* (although the prefix *y* has been replaced by *t* in those languages, i.e. *taqtulūna*). In Hebrew and Aramaic, the loss of final short vowels has resulted in the falling together of original **yaqtul* and **yaqtulu* in singular forms. Thus, Classical Hebrew *yiḵtōḇ* may be either jussive ‘let him write’ (< **yaktub*) or imperfect ‘he writes, will write’ (< **yaktubu*); a number of weak verb types, however, preserve the original distinction, such as *yāqōm* ‘let him stand’ < **yaqum* vs. *yāqūm* ‘he stands, will stand’ < **yaqūmu* and *yíḇen* ‘let him build’ < **yabni(y)* vs. *yiḇne* ‘he builds, will build’ < **yabniyu*. Early Aramaic preserves this distinction in a number of verbs as well, but it is lost in later dialects. The merger of **yaqtul* and **yaqtulu* in the singular made the distinction between plural **yaqtulū* and **yaqtulūna* redundant; Aramaic leveled the latter form (thus, e.g., Biblical Aramaic *yiḵtəḇūn*), while Hebrew leveled the former (*yiḵtəḇū*, although the latter, *yiḵtəḇūn*, continued to appear as a biform in imperfect usages).

All Northwest Semitic languages attest → ‘energic’ prefix conjugation forms, or at least vestiges of them. In Amarna Canaanite, the form is **yaqtuluna*, which occurs especially in the 1st person, and especially in emphatic questions, as in *mīna īpušuna* ‘what am I to do?’. In Ugaritic, energic forms written with both <N> (for /-an(n)a/?) and <NN> (for <-anVn(n)V/??) occur; their function is debated (see Tropper 2000:497–506). In Hebrew and Aramaic, the energic appears only before pronominal object suffixes, as in Hebrew *yiḵtəḇennū:*, Aramaic (Biblical) *yiḵtəḇinneh* ‘he will write it’; such forms replace **yaqtulu* + suffixes. (Both the Hebrew and the Aramaic suggest an earlier form **yaqtul-in-*, with *-i-* rather than the *a* of Arabic *yaqtulan(na)* [and Ugaritic?] and the *u* of Amarna *yaqtuluna*.) A recent treatment of energic forms in Arabic and Northwest Semitic is Zewi (1999).

The prefixes of the prefix conjugations in Proto-Northwest Semitic are similar to those in Classical Arabic except in the 3rd person plural, for which we have to assume an original heterogeneous distribution of **y-* (3rd pers. masc. pl.) and **t-* (3rd pers. fem. pl.) (see Table 4). The distribution of prefix consonants in the individual Northwest Semitic languages would otherwise be difficult to explain. Ugaritic and Amarna Canaanite leveled the feminine **t-*, so that the 3rd person masculine plural is most

Table 5. Derived verbal forms

	Classical Arabic	Hebrew	Aramaic (Biblical)	Ugaritic
D (II)	*faʿʿala/yufaʿʿil-*	*qittēl/yəqattēl*	*qattēl/yəqattēl*	*qattila/yVqattil-*
Caus. (IV)	*ʾafʿala/yufʿil-*	*hiqtīl/yaqtīl*	*haqtēl/yəhaqtēl* Syriac: *ʾaqtel/naqtel*	*šaqt^{i}/ala/yVšaqtil-*
N (VII)	*(i)nfaʿala/yanfaʿil-*	*niqtāl/yiqqātēl*	—	*naqtala/yiqqatil-*
L (III)	*fāʿala/yufāʿil-*	(*qōtēl/yəqōtēl*)	—	?
R (IX)	*(i)fʿalla/yafʿall-*	*qōmēm/yəqōmēm*	(*qāmēm/yəqāmēm*)	?*yukānin-*
tG (VIII)	*(i)ftaʿala/yaftaʿil-*	(*hiṯqātēl/yiṯqātēl*)	*hiṯqətēl/yiṯqətēl*	*ʾiqtat^{i}/ala/yiqtatil-*
tD (V)	*tafaʿʿala/yatafaʿʿal-*	*hiṯqattēl/yiṯqattēl*	*hiṯqattal/yiṯqattal*	*taqattala* or *ʾitqattila/ yVtqatt^{a}/il-*
Ct (X)	*(i)stafʿala/yastafʿil-*	(*hištaḥ^{a}wā/yištaḥ^{a}we*)	Syriac: *ʾeṯtaqtal/neṯtaqtal*	*ʾištaqt^{i}/al/yVštaqtil-*
tL (VI)	*tafāʿala/yatafāʿal-*	(*hiṯqōtēl/yiṯqōtēl*)	—	
tR		(*hiṯqōmēm/yiṯqōmēm*)		

often attested as *t-*, while Aramaic leveled the masculine **y-* (as did Akkadian and Ethiopic). Hebrew preserves the original distinction of 3rd person masculine plural **y-* and 3rd person feminine plural **t-*. Since even Ancient North Arabian and Old Arabic seem to have **y-* in both the masculine plural and feminine plural, this distribution might go back to the earliest stages of Arabic.

The prefix vowels in the Proto-Northwest Semitic basic stem were originally dependent on the theme vowel of the imperfect base. When the theme vowel was *i* or *u*, the prefix vowel was *a*, *yaqtul* and *yaqtil*, but it was *i* when the theme vowel was *a*, *yiqtal*. This distribution is known as the Barth-Ginsberg Law. Ugaritic preserves this vowel distribution, while Hebrew and Aramaic only have vestige forms (Barth 1894:4–5). Classical Arabic regularly has the prefix vowel *a* in the basic stem, but some ancient and modern dialects preserve reflexes of the Proto-Northwest Semitic distribution, indicating that the *a* of Classical Arabic is the result of leveling (Bloch 1967:22–25; Hetzron 1973–1974; Hasselbach 2004).

The feminine plural forms of the prefix conjugations exhibit final *-na* in Hebrew (and, probably, Ugaritic), e.g. 2nd person feminine plural *tiḵtōḇnā*, as in Arabic *taktubna*. This was probably also true in the earliest Aramaic dialects; later Aramaic, however, has replaced *-na* with *-ān*, as in *tiḵtəḇān* (Huehnergard 1987).

A few Ugaritic imperative forms, all feminine singular, exhibit a prothetic vowel, as in Classical Arabic imperative forms, e.g. <IBKY> for /ʾibkiyī/ 'weep!' (Tropper 2000:426–427). Most Ugaritic imperatives, however, have no prothetic vowel; rather, the first two root consonants are separated by an anaptyctic vowel (e.g. <RGM>, probably /r^{u}gum/ 'speak!'). The latter is the norm in all other Northwest Semitic languages (and the rest of Semitic), as in Hebrew and Aramaic *kəṯōḇ* 'write!'.

Table 4. Proto-Northwest Semitic prefixes

3ms	**yV-*	3cd	**yV-*	3mp	**yV-*
3fs	**tV-*			3fp	**tV-*
2ms	**tV-*	2cd	**tV-*	2mp	**tV-*
2fs	**tV-*			2fp	**tV-*
1cs	**ʾV-*			1cp	**nV-*

The active participle of the basic form of the verb may be reconstructed for Proto-Semitic as *qātil-*, and this remains unchanged in both Arabic and the Northwest Semitic languages (> *qōtēl* in Hebrew, with the change of **ā* > *ō*). A common passive participle, however, cannot be reconstructed. For Hebrew, the paradigmatic form is **qatūl*, as in *kāṯūḇ* 'written', whereas Aramaic regularized **qatīl*, as in *kəṯīḇ*. The pattern **maqtūl*, which was leveled as the paradigmatic passive participle in Arabic, as in *maktūb*, is not used as such in Northwest Semitic, with the possible exception of Ugaritic, where a few examples have been suggested, such as /mašnūʾu/? 'enemy' (i.e. 'hated') and <MDD> for /mōdūdu/? 'beloved', but other interpretations of such forms are equally possible (see Tropper 2000:476–477). A few **maqtūl* nouns in Hebrew seem to have passive semantics, e.g. *maspūnīm* 'treasures', i.e. 'hidden things'.

In contrast to the large set of patterns from which Arabic verbs may select their verbal nouns (*maṣdar*s), the Northwest Semitic languages exhibit only a small number of patterns. Most Aramaic dialects exhibit **miqtal* as the basic (G) infinitive, but the earliest inscriptions also attest forms without the initial *m* (vocalization uncertain). Biblical Hebrew has two paradigmatic infinitive forms: *qāṭōl* < **qaṭāl*, traditionally called the 'infinitive absolute'; and *qəṭōl* (perhaps also < **qaṭāl*, or perhaps < **quṭul*), called the 'infinitive construct'. There are vestiges of other patterns used as verbal nouns, however, such as **qitl* and **qitlat*. In Ugaritic it is likely that **qaṭāl* and **qitl* occurred as verbal nouns, and probably a few other patterns as well (Tropper 2000:480–490).

2.2.3.1. Derived verbal forms

Table 5 provides an overview of the derived forms in Classical Arabic and the main Northwest Semitic languages. Forms in parentheses are rare.

D-stem: The Proto-Central Semitic form of the D-stem was most likely **qattil-*. Arabic (Classical and dialects), and Geʿez for that matter, leveled the /a/-vowel, resulting in the attested form *qattala*, while Proto-Northwest Semitic preserved the form *qattila*, as reflected in Ugaritic and Aramaic. In Canaanite, another change occurred from **qattila* to **qittila* (Huehnergard 1992:219), resulting in the Hebrew form *qittēl*. Both Canaanite and Arabic reflect independent innovations, while Proto-Northwest Semitic preserved the original forms. For the meaning of the D-stem, which is similar in all Central Semitic languages, see Kouwenberg (1997).

Causative stem: The Ugaritic form of the causative indicates that the sibilant formative that is likewise attested in Akkadian, as in *ušapris*, was still preserved in the early stages of Northwest Semitic. For the reconstruction of an original sibilant, see Voigt (1988:60). The original sibilant of the causative is also still preserved in the Arabic Form X, *(i)stafʿala*. This sibilant changed to *h* in Canaanite and Biblical Aramaic, and further to ʾ in Classical Arabic and Syriac (Voigt 1988:57–59), i.e. **yusaqtil* > **yuhaqtil* > **yu(ʾ)aqtil* > **yaqtil* in Hebrew and Syriac, but *yuqtil* in Classical Arabic. Also note that Dadanitic still has a derived stem with prefixed *h-*, *hfʿl*, in addition to *ʾfʿl* (Macdonald 2004:512). The vowels of the suffix conjugation underwent the same changes as in the D-stem in both Canaanite and Arabic, i.e. **haqtila* > **hiqtila* in Caananite and **haqtila* > **haqtala* in Arabic.

N-stem: The N-stem is characterized by an /n/ that is prefixed to the verbal root in all languages that preserve it. In the suffix conjugations of Ugaritic, Phoenician, and Hebrew, the /n/ stands at the beginning of the word with a vowel between /n/ and R_1, nVR_1–, while in Arabic the /n/ is immediately prefixed to the verbal base and has a *hamzat al-waṣl* to resolve the initial consonant cluster, $(ʾi)nR_1$–. The vowel that follows /n/ in Northwest Semitic languages differs in that Hebrew has /i/, while Ugaritic has /a/. It is likely that Ugaritic preserves the original vowel of the perfect, while the Hebrew form is the result of an inner Hebrew development; cf. also the Akkadian N-verbal adjective *naprus*. In the imperfect, Arabic, Hebrew, and Ugaritic share the same basic form, except that Ugaritic and Hebrew assimilate the original /n/ to the following root consonant. Aramaic lost the N-stem completely. The N-stem is used as a middle/reflexive/passive stem in Arabic and in Ugaritic. When used as a passive, it usually relates to the G-stem. The function as passive to the basic stem is the one most commonly used in Hebrew and Phoenician, which might be explained by the loss of internal passives in these languages.

L-stem: Northwest Semitic languages do not have a productive L-stem as found in Arabic (Form III). There are, nevertheless, vestiges of such stems in Hebrew and perhaps also in Ugaritic (Tropper 2000:577–585). The rare attestations of the form *qōṭēl* for strong roots in Hebrew resemble the Classical Arabic Form III both formally – the long /ō/ between R_1 and R_2 derives from original **ā* – and semantically. As is well known, the L-stem of Arabic expresses the notion of 'having an action/somebody as goal', as in *qātala* 'to fight', i.e. 'having killing as a goal'. The same meaning is found in Hebrew, as in *məšōp̱ṭī* 'my adversary', i.e. 'somebody who would contend with me' (Gesenius a.o. 1910, § 55b). In Ugaritic, a similar stem seems to be used for geminate roots, but with the semantic range of the D-stem, as in *tʿzzk* /tuʿāzizūka/ 'may they make you strong', from the root *ʿ-z-z* 'to be strong'.

R-stem: The R-stem in Hebrew and Ugaritic is confined to roots II*w/y* for which it has the same semantic range as the D-stem, as in Hebrew *mōṯēṯ* 'to slay, kill' from the root **m-*

w-t 'to die' and Ugaritic *yrmm* /yurāmim/ 'he built/made high' from the root **r-y-m* 'to be high'. The Arabic R-stems (Forms IX and XI) are used for adjectival forms of the type *ʾafʿalu*, as in *(ʾi)ṣfarra* and *(ʾi)ṣfārra* 'to become yellow' from the adjective *ʾaṣfaru* 'yellow'. The different R-stems in Arabic and Northwest Semitic are probably not related since they differ both formally and semantically.

t-stems: The forms of the *t*-stems in the individual languages differ with regard to whether the *t*-preformative is prefixed or infixed. In Syriac, we only find prefixed forms, while Ugaritic and Hebrew have prefixed *t*-forms except in the Ct, where the *t* is infixed, but note that Hebrew only has one verbal root that occurs in the Ct. The infixation of the *t* in the causative is probably caused by the sibilant. In Hebrew, a metathesis of the prefixed *t* with a following sibilant regularly takes place in the Hithpaʿel, e.g. *hitpallēl* 'he prayed' vs. *hištammēr* 'he guarded himself'. In Arabic, the tD and tL are prefixed and the Gt and Ct infixed. For a reconstruction of these forms, see, for example, Diem (1982) and Testen (1999). In most of the languages, the *t*-stems are used as reflexive/middle, rarely as passive, although in Syriac, they are exclusively used for the passive. Note that the Proto-Semitic St stem (Arabic Form X) is vestigial in Hebrew (one verb, *yištaḥᵃwe* < **yistaḥwiyu* 'he prostrates himself') and was lost in Aramaic (though later reintroduced from Akkadian), where it was replaced by innovative *ʾittaqtal*. Also in Hebrew the Gt (Arabic Form VIII) was lost except for a few relic forms.

Other stems, such as Forms XII–XV, are missing or at best vestigial in Northwest Semitic.

2.3. *Syntax*

The basic word order in Arabic and the Northwest Semitic languages is the same, with Subject-Predicate in nominal sentences and Verb-Subject-Object in verbal sentences. Attributive adjectives usually follow the noun they modify. Variations to these basic rules are found when specific elements of a sentence are fronted for emphasis and topicalization (see Khan 1988). Classical Arabic differs from Northwest Semitic in that an indefinite subject cannot stand at the beginning of a sentence, in which case the predicate is fronted. This construction is not found in Northwest Semitic.

Predicate and subject usually agree in gender and number in both nominal and verbal sentences in Northwest Semitic languages, regardless of the position of the predicate in a sentence. (An exception is the dual in Canaanite and Aramaic, since the dual is only preserved in vestiges in these languages; dual subjects are therefore construed as plural, with the gender depending on the gender of the singular.) In Canaanite and Aramaic, as in Arabic, an attributive adjective also agrees in definiteness with the noun it modifies, while predicate adjectives remain indefinite. Exceptions to these agreement rules can occur. In Hebrew, plurals of names, animals, things, and abstracts are sometimes construed as feminine singular with verbal predicates. There also occur variations when the predicate precedes its subject, in which case the predicate can stand in the 3rd person masculine singular, independent of the number of the following subject, although agreement is more frequent. Interestingly, unlike Hebrew, Ugaritic does not seem to have cases of lack of agreement in verbal sentences (Tropper 2000:886).

Classical Arabic agreement rules differ from Northwest Semitic. Two major differences need to be addressed. First, a verbal predicate when it precedes its subject is always construed in the singular, whether the subject is singular or plural. Although this situation is sometimes found in Hebrew as well, it is not as strict as in Classical Arabic. In some colloquial Arabic dialects, a verbal predicate agrees in number and gender with its subject independent of its position (Rabin 1951:209). In Ancient North Arabian, verbs generally agree with their subjects in gender and number as well (Macdonald 2004:526). The second major difference is that inanimate plurals are generally construed as feminine singulars, likewise a situation that is found only sporadically in Northwest Semitic.

Another syntactic feature in which Arabic differs from Northwest Semitic languages is with regard to relative clauses. In Classical Arabic, relative clauses are not introduced by a relative pronoun when the nominal antecedent is indefinite, but they have to be introduced by a relative pronoun when the antecedent is definite. In Ṣafaitic, this distinction is not kept as strictly; although relative clauses may be introduced by the relative pronouns *ḏ* or *mn*, these pronouns can be absent even after a definite noun

(Macdonald 2004:528). In Northwest Semitic, no distinction of this kind is made in those languages that have a marker for definiteness. In Syriac, relative clauses are generally introduced by the relative pronoun *d-*. Exceptions to this rule are rare and often are Hebraisms (Nöldeke 1904, § 354). In Hebrew, relative clauses are usually introduced by the pronoun *ʾᵃšer*, but asyndetic relative clauses occur as well, especially in poetry. These asyndetic relative clauses, interestingly, are most commonly found after an indefinite antecedent, a situation that resembles Classical Arabic (Gesenius a.o. 1910, § 155d).

2.4. *Lexicon*

The Northwest Semitic languages and Arabic share a significant number of lexical items that are not attested in other Semitic languages, such as the prepositions **ʾil(ay)* 'to(ward)' and **ʿim/maʿ* 'with'; the nouns **ʿabd* 'servant', **kapp* 'palm of hand', **laban* 'white', and the extended form **ʾilāh* 'god'; and verbal roots such as *ʿ-ḏ-r* 'to help', *ḥ-g-g* 'to make a pilgrimage', *ḥ-z-y* 'to see', *n-ḥ-m* 'to console', *s-k-n* 'to dwell', *t-m-m* 'to be complete', and *ṭ-r-p* 'to pluck'. There are also, however, many items that are found in Arabic but not in Northwest Semitic, and vice versa.

3. SUBGROUPING

The most disputed aspect concerning the subgrouping of the Semitic languages is the position of Arabic. The traditional view is that Arabic belongs to a → South Semitic group that also includes Old South Arabian, Modern South Arabian, and Ethiopian Semitic (Nöldeke 1899:17; Brockelmann 1908:21). This view was challenged in several influential articles by Hetzron (1974, 1976), in which he argued for a subgroup labeled Central Semitic, consisting of Arabic and the Northwest Semitic languages, on the basis of certain shared morphological innovations in the verbal system, particularly the imperfect form *yaqtulu*. Hetzron's classification is widely accepted today (e.g. Faber 1980, 1997; Voigt 1987), although a number of scholars continue to argue in favor of a subgrouping of Arabic as South Semitic (Blau 1978; Diakonoff 1988; Diem 1980; Zaborski 1991, 1994; Ratcliffe 1998).

The debate centers on different evaluations of certain isoglosses that Arabic shares with Northwest Semitic on the one hand and with South Semitic on the other. Scholars who argue in favor of the subgrouping of Arabic with Northwest Semitic, in a Central Semitic subgrouping, usually consider the form of the prefix conjugation *yaqtulu*, underlying the Northwest Semitic and Arabic imperfect, to be a shared innovation of these languages (now known to be shared by the Old South Arabian languages as well; see Nebes 1994). South Semitic does not have a reflex of this verbal form but uses *yaqattal* instead, a shared retention from Proto-Semitic also found in Akkadian (Hetzron 1974:187, 1976:105).

Scholars who support the subgrouping of Arabic as a member of South Semitic explain similarities between Northwest Semitic languages and Arabic as results of language contact. The main two isoglosses quoted against a genealogical relationship with Northwest Semitic are the wide range of broken plural patterns used in Arabic and South Semitic (but see Huehnergard 2005), and the Arabic/South Semitic verbal stems with lengthened first vowel, *qātala* and *taqātala* (i.e. Arabic Forms III and VI), which are explained as shared innovations (Nöldeke 1899:17; Diem 1980:69; Zaborski 1991:370, 1994:399; Ratcliffe 1998:120). Rare occurrences of Hebrew *Poʿel* forms of strong roots resemble the Classical Arabic *qātala* both formally and semantically, however (Gesenius a.o. 1910, § 55b). Furthermore, this stem is attested outside Semitic, for example in Beja (Zaborski 1991:371), and may go back to an early stage of Afro-Asiatic; if so, it would not constitute a shared Arabic/South Semitic innovation.

BIBLIOGRAPHICAL REFERENCES

Barth, Jacob. 1894. "Zur vergleichenden semitischen Grammatik". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 48.1–21.

Bauer, Hans and Pontus Leander. 1922. *Historische Grammatik der hebräischen Sprache des Alten Testaments*. Halle: Niemeyer.

——. 1927. *Grammatik des Biblisch-Aramäischen*. Halle: Niemeyer.

Blau, Joshua. 1978. "Hebrew and North West Semitic: Reflections on the classification of the Semitic languages". *Hebrew Annual Review* 2.21–44.

Bloch, Ariel A. 1967. "The vowels of the imperfect preformatives in the old dialects of Arabic". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 117.22–29.

Brockelmann, Carl. 1908. *Grundriss der ver-*

gleichenden Grammatik der semitischen Sprachen, I. Berlin: Reuther und Reichard.

Cantineau, Jean. 1930–1932. *Le nabatéen*. 2 vols. Paris: Ernest Leroux.

——. 1935. *Grammaire du palmyrénien épigraphique*. Cairo: Institut Français d'Archéologie Orientale.

Dalman, Gustav. 1905. *Grammatik des jüdisch-palästinischen Aramäisch*. 2nd ed. Leipzig: J.C. Hinrichs.

Degen, Rainer. 1969. *Altaramäische Grammatik der Inschriften des 10.–8. Jh. v. Chr.* Wiesbaden: F. Steiner.

Diakonoff, Igor M. 1988. *Afrasian languages*. Moscow: Nauka.

Diem, Werner. 1980. "Die genealogische Stellung des Arabischen in den semitischen Sprachen: Ein ungelöstes Problem der Semitistik". *Studien aus Arabistik und Semitistik: Anton Spitaler zum siebzigsten Geburtstag*, ed. Werner Diem, 65–85. Wiesbaden: O. Harrassowitz.

——. 1982. "Die Entwicklung des Derivationsmorphems der *t*-Stämme im Semitischen". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 132.29–84.

Faber, Alice. 1980. *Genetic subgroupings of the Semitic languages*. Ph.D. diss., University of Texas at Austin.

——. 1997. "Genetic subgrouping of the Semitic languages". *The Semitic languages*, ed. Robert Hetzron, 3–15. London: Routledge.

Fischer, Wolfdietrich. 1995. "Zum Verhältnis der neuarabischen Dialekte zum Klassisch-Arabischen". *Studia Orientalia* 75.75–86.

Folmer, M.L. 1995. *The Aramaic language of the Achaemenid period: A study in linguistic variation*. Leuven: Peeters.

Friedrich, Johannes, Wolfgang Röllig, and Maria Giulia Amadasi Guzzo. 1999. *Phönizisch-Punische Grammatik*. 3rd ed. Rome: Pontificio Istituto Biblico.

Gesenius, Wilhelm, E. Kautzsch, and A.E. Cowley. 1910. *Gesenius' Hebrew grammar*. 2nd English ed. from the 26th German ed. Oxford: Clarendon.

Hasselbach, Rebecca. 2004. "The markers of person, gender and number in the prefixes of G-preformative conjugations in Semitic". *Journal of the American Oriental Society* 124.23–35.

Hetzron, Robert. 1973–1974. "The vocalization of prefixes in Semitic active and passive verbs". *Mélanges de l'Université Saint-Joseph* 48.35–48.

——. 1974. "La division des langues sémitiques". *Actes du Premier Congrès International de Linguistique Sémitique et Chamito-Sémitique*, ed. André Caquot, 181–194. The Hague: Mouton.

——. 1976. "Two principles of genetic reconstruction". *Lingua* 38.89–108.

Huehnergard, John. 1987. "The feminine plural jussive in Old Aramaic". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 137.265–277.

——. 1991. "Remarks on the classification of the Northwest Semitic languages". *The Balaam text from Deir ʿAlla re-evaluated*, ed. Jacob Hoftijzer, 283–293. Leiden: E.J. Brill.

——. 1992. "Historical phonology and the Hebrew piel". *Linguistics and Biblical Hebrew*, ed. Walter R. Bodine, 209–229. Winona Lake, Ind.: Eisenbrauns.

——. 2005. "Features of Central Semitic". *Biblical and Oriental essays in memory of William L. Moran*, ed. Agustinus Gianto, 155–203 (= *Biblica et Orientalia*, 48). Rome: Pontificio Istituto Biblico.

Hug, Volker. 1993. *Altaramäische Grammatik der Texte des 7. und 6. Jh.s v.Chr.* Heidelberg: Heidelberger Orientverlag.

Joüon, Paul. 1991. *A grammar of Biblical Hebrew*. Trans. and rev. Takamitsu Muraoka. 2 vols. Rome: Pontificio Istituto Biblico.

Khan, Geoffrey. 1988. *Studies in Semitic syntax*. Oxford: Oxford University Press.

Knudtzon, J.A. 1907–1915. *Die El-Amarna-Tafeln mit Einleitung und Erläuterungen*. Annotations and index by O. Weber and E. Ebeling. 2 vols. Leipzig: J.C. Hinrichs.

Kouwenberg, N.J.C. 1997. *Gemination in the Akkadian verb*. Assen: van Gorcum.

Liverani, Mario. 1963. "Antecedenti del diptotismo arabo nei testi accadici di Ugarit". *Rivista degli Studi Orientali* 38.131–160.

Macdonald, Michael C.A. 2000. "Reflections on the linguistic map of pre-Islamic Arabia". *Arabian Archeology and Epigraphy* 11.28–79.

——. 2004. "Ancient North Arabian". *The Cambridge encyclopedia of the world's ancient languages*, ed. Roger D. Woodard, 488–533. Cambridge: Cambridge University Press.

Macuch, Rudolf. 1965. *Handbook of classical and modern Mandaic*. Berlin: W. de Gruyter.

——. 1982. *Grammatik des samaritanischen Aramäisch*. Berlin: W. de Gruyter.

Müller, Walter W. 1982. "Das Altarabische und das Klassische Arabisch". *Grundriss der arabischen Philologie. I. Sprachwissenschaft*, ed. Wolfdietrich Fischer, 17–29. Wiesbaden: L. Reichert.

Müller-Kessler, C. 1991. *Grammatik des Christlich-Palästinisch-Aramäischen*. Hildesheim: G. Olms.

Muraoka, Takamitsu and Bezalel Porten. 2003. *A grammar of Egyptian Aramaic*. 2nd ed. Leiden: E.J. Brill.

Nebes, Norbert. 1994. "Zur Form der Imperfektbasis des unvermehrten Grundstammes im Altsüdarabischen". *Festschrift Ewald Wagner zum 65. Geburtstag. I. Semitische Studien unter besonderer Berücksichtigung der Südsemitistik*, ed. Wolfhart Heinrichs, 51–81. Stuttgart: F. Steiner.

Nöldeke, Theodor. 1899. *Die semitischen Sprachen*. Leipzig: Chr. Herm. Tauchnitz.

——. 1904. *Compendious Syriac grammar*. London: Williams and Norgate. (Repr., Winona Lake, Ind.: Eisenbrauns, 2001.)

Parker, Simon. 2002. "Ammonite, Edomite, and Moabite". *Beyond Babel: A handbook for Biblical Hebrew and related languges*, ed. John Kaltner and Steven L. McKenzie, 43–60. Atlanta, Ga.: Society of Biblical Literature.

Rabin, Chaim. 1951. *Ancient West-Arabian*. London: Taylor's Foreign Press.

Rainey, Anson F. 1996. *Canaanite in the Amarna tablets*. 4 vols. Leiden: E.J. Brill.

Ratcliffe, Robert R. 1998. "Defining morphological isoglosses: The 'broken' plural and Semitic

subclassification". *Journal of Near Eastern Studies* 57.81–123.
Rosenthal, Franz. 1995. *A grammar of Biblical Aramaic*. 6th ed. Wiesbaden: O. Harrassowitz.
Rubin, Aaron. 2004. *Studies in Semitic grammaticalization*. Ph.D. diss., Harvard University.
Streck, Michael. 2000. *Das amurritische Onomastikon der altbabylonischen Zeit*. I. *Die Amurriter: Die onomastische Forschung, Orthographie und Phonologie, Nominalmorphologie*. Münster: Ugarit-Verlag
Testen, David. 1999. "Arabic evidence for the formation of the verbal noun of the Semitic Gt-stem". *Journal of Semitic Studies* 44.1–16.
Tropper, Josef. 2000. *Ugaritische Grammatik*. Münster: Ugarit-Verlag.
Voigt, Rainer M. 1987. "The classification of Central Semitic". *Journal of Semitic Studies* 32.1–21.
——. 1988. "Die Personalpronomina der 3. Personen im Semitischen". *Welt des Orients* 18.49–63.
——. 1999. "Nominal (and verbal) nasalization and the nominal plural morphemes in Semitohamitic". *Afroasiatica Tergestina: Papers from the Ninth Italian Meeting of Afro-Asiatic (Hamito-Semitic) Linguistics, Trieste, April 23–24, 1998*, ed. Marcello Lamberti and Livia Tonelli, 11–22. Padua: Unipress.
Waltke, Bruce K. and M. O'Connor. 1990. *Introduction to Biblical Hebrew syntax*. Winona Lake, Ind.: Eisenbrauns.
Zaborski, Andrzej. 1991. "The position of Arabic within the Semitic dialect continuum". *Proceedings of the Colloquium on Arabic Grammar*, ed. Kinga Dévényi, 365–375. Budapest: Csoma de Kőrös Society of Islamic Studies.
——. 1994. "Problèmes de classification des dialectes sémitiques méridionaux". *Actes des Premières Journées Internationales de Dialectologie Arabe de Paris*, ed. Dominique Caubet, 400–411. Paris: Publications Langues'O.
Zewi, Tamar. 1999. *A syntactical study of verbal forms affixed by -n(n) endings in Classical Arabic, Biblical Hebrew, el-Amarna Akkadian and Ugaritic*. Münster: Ugarit-Verlag.

Rebecca Hasselbach (University of Chicago)
John Huehnergard (Harvard University)

Noun

1. Definition

Generally speaking, the term 'noun' can be used either as a synonym for 'substantive' and 'adjective', or as an umbrella term for all 'nominal' parts of the sentence, including adjectives as well as pronouns and numerals. In the context of Semitic and Arabic linguistics, 'noun' (→ *ism*) is always used in the first, narrower sense. The demarcation between substantive and → adjective poses a morphosyntactic problem in some Semitic languages (→ noun phrase; → adjective phrase) because from a typological point of view, substantive-adjective phrases could also be analyzed as two nouns in apposition, e.g. *ar-rajulu l-kabīru* 'the man the big (one)'. Moreover, the strict distinction between nominal patterns (including declensional features) specific to substantives and those specific to adjectives that is found, for instance, in Akkadian (Huehnergard 2000:607–609) does not exist in (Classical) Arabic (see, e.g., Fischer 1997:192). Therefore, the following notes do not exclude references to adjectives, because these can function regularly as substantives, which is especially evident in the case of the → participle.

From the native Arab grammarians' point of view, the definition of 'noun' is more complicated. On the one hand, → *ism* philosophically denotes everything that can be assigned a name. Much in the spirit of the Qurʾānic verse *wa-ʿallama ʾĀdama l-ʾasmāʾa kullahā* 'and He taught Adam all the names [sc. of things that can be assigned a name]' (Q. 2/31), Sībawayhi writes that *ʾiḏā qulta marartu bi-rajulin ʾinnnamā zaʿamta ʾannaka ʾinnamā mararta bi-wāḥidin mim-man yaqaʿu ʿalayhi hāḏā l-ism* 'by saying "I passed by a man" simply means that one passed by one of those to whom this name [i.e. 'man'] applies' (*Kitāb* I, 2201ff./187.18ff.). On the other hand, Sībawayhi also subsumes the demonstrative pronoun, the participle, the elative form *ʾafʿalu*, and certain indeclinable words under his concept of *ʾasmāʾ* (see Diem 1970–1971:316ff.). This broader definition is still reflected in Wright's grammar (1967:I, 104ff.), which lists six kinds of nouns: (1) *al-ism al mawṣūf* or a*l-manʿūt* 'the noun that can be qualified by an adjective', (2) *aṣ-ṣifa*, *al-waṣf*, or *an-naʿt* 'the adjective', (3) *ism al-ʿadad* 'the number', (4) *ism al-ʾišāra* 'the demonstrative pronoun', (5) *al-ism al-mawṣūl* 'the relative pronoun', and (6) *aḍ-ḍamīr* or *al-muḍmar* 'the pronoun'.

The concept 'noun' in the sense of a word denoting a person or a thing (cf. Sanskrit *nāman* or Greek *ónoma*) is discussed here with special attention to morphological notions; the entry does not, nevertheless, neglect semantic categories, because these are closely associated with noun patterns. 'Nominalized' clauses (cleft sentences and relative clauses), which can adopt the syntactic position of a noun, are excluded from further consideration. Questions regarding number, gender, and case (including the opposition of → diptosis vs. triptosis) are only touched upon briefly here (→ declension). Proper nouns as well as recent developments in the integration of foreign words into Arabic are taken into consideration below.

2. The noun in Classical Arabic

The formal apparatus of nonconcatenative template-based morphology is especially suitable to describe the interdigitation of a (prototypically triliteral) root morpheme and an inflectional morpheme, as is characteristic of the nominal and verbal realm of the lexicon in Arabic and in other Semitic languages (see McCarthy and Prince 1995 and Ratcliffe 1997, with copious further references for a formal treatment of this situation). The conceptual parallels between the native Arabic notion of *taṣrīf* 'derivation' (→ *ṣarf*) and modern morphophonological theory are striking (see, e.g., Owens 1984; Bohas a.o. 1990).

The most basic form of a noun is a $C_1vC_2(v)C_3$ structure. In addition, such nonconcatenative structures can be expanded by derivational affixes. This latter type of expansion is often defined in a way that includes 'inner' derivation, e.g. derivation resulting in the → diminutive $C_1uC_1ayC_3$ pattern.

As a basic notion it is important to keep in mind that the Arabic term → *ʾaṣl*, conventionally rendered as → 'root' in the Western Arabist tradition, might better be understood as the 'base' out of which one Arabic word can be derived from another (see, for instance, Larcher 1999 and the contributions in Shimron 2003). Nouns can be divided into two basic types, 'primitive' and 'derivative'. Whereas *rajul* 'man' and *māʾ* 'water' constitute examples of primitive nouns, *miftāḥ* 'key' and *taqsīm* 'division' are derivative nouns, derived from *fataḥa* 'to open' and *qasama* 'to divide', respectively. According to the traditional position in modern grammars of Arabic (notably Wright), derivative nouns can in turn be divided into deverbal ones, e.g. *miftāḥ* from *fataḥa*; denominal nouns, e.g. *maʾsada* 'a place abounding in lions' from *ʾasad* 'lion'; and nouns derived from particles, e.g. *ʾanāniyya* 'egoism' from *ʾana* 'I', or *kayfiyya* 'quality' from *kayfa* 'how' (Wright 1967: I, 106). Of these types, the category 'deverbal' might appear questionable, however, because it is difficult to determine whether the verbal realm or the nominal realm has any kind of linguistic 'priority' (see also the remarks regarding the *maṣdar*, below). Therefore, it might be more appropriate to divide nouns into those which are derivable from a root (e.g. *f-t-ḥ*) and those which are not (e.g. the ones derived from particles). With regard to their morphological structure, nouns are either 'simple' or 'extended' by preformatives and afformatives or infixes.

A true wealth of noun patterns can be identified in Classical Arabic (see Wright 1967:I, 110ff.; for a data collection in the framework of comparative Semitics, see Barth 1894; Brockelmann 1908:329–404; Lipiński 2001:215–235). The following survey of patterns that are relevant for Modern Standard Arabic is based on Badawi a.o. (2004:49); it does not take into account surface deviations in weak roots.

i. Minimal triradical patterns (i.e. without affixes but with possible feminine ending): $C_1aC_2C_3$, $C_1iC_2C_3$, $C_1uC_2C_3$, $C_1aC_2aC_3$, $C_1aC_2iC_3$, $C_1iC_2aC_3$, $C_1uC_2aC_3$, $C_1uC_2uC_3$;
ii. Extended triradical patterns (i.e. with internal lengthening of vowels and/or gemination of either the second or third radical; the full inventory of affixes in Classical Arabic consists of the prefixes *ʾa-*, *ʾi-*, *ʾu-*, *ma-*, *mi-*, *mu-*, *ta-*, *ti-*, *tu-*, *ya-* and the suffixes *-a(t)*, *-ā*, *-āʾ*, *-ān*, *-iyy(a)*, *-ūt*): $C_1\bar{a}C_2aC_3$, $C_1aC_2C_3\bar{a}$, $C_1aC_2\bar{a}C_3(a)$, $C_1aC_2\bar{a}C_3\bar{a}$, $C_1aC_2\bar{a}C_3\bar{\imath}$, $C_1uC_2C_3\bar{a}$, $C_1aC_2\bar{\imath}C_3(a)$, $C_1aC_2\bar{u}C_3(a)$, $C_1iC_2C_3\bar{a}$, $C_1iC_2\bar{a}C_3(a)$, $C_1uC_2\bar{u}C_3(a)$, $ʾaC_1C_2aC_3(a)$, $ʾuC_1C_2\bar{u}C_3a$, $maC_1C_2aC_3(a)$, $maC_1C_2iC_3(a)$, $maC_1C_2uC_3a$, $miC_1C_2aC_3(a)$, $miC_1C_2\bar{a}C_3$, $yaC_1C_2\bar{u}C_3$; *ta(/i/u)*-prefixes that are not part of a *maṣdar* of the Forms II, V, or VI are almost unattested in Modern Standard Arabic or are restricted to proper nouns; the pattern $yaC_1C_2\bar{u}C_3$ is rare, e.g. *yanbūʿ* 'well';
iii. Quadriradical (and longer) patterns: $C_1aC_2C_3aC_4(a)$, $C_1aC_2\bar{a}C_3C_4$, $C_1aC_2\bar{a}C_3\bar{\imath}C_4$ (in the native Arab system these are represented by adding another *l* to the *f-ʿ-l* skeleton, e.g. *faʿlala* for $C_1aC_2C_3aC_4a$, a principle which is unfortunate in a modern descriptive framework because it suggests gemination of the type that occurs, for instance, with the last radical in Form IX).

The most important types of deverbal and denominal nouns with their typical semantic load are the following: The C_1*a*C_2C_2*ā*C_3 pattern typically expresses intensity or professionality (e.g. *xayyāṭ* 'tailor'). The C_1*u*C_2*ay*C_3 pattern expresses the diminutive (e.g. *kulayb* 'little dog'). The name of place or time (*ism al-makān/ism az-zamān/ism aḏ-ḏarf*) has the pattern *ma*C_1C_2*i*C_3(*a*) for verbal roots with *i* or *u* as thematic vowel in the imperfective (e.g. *manzil* 'resident place' from *nazala/yanzilu* 'to descend') and *ma*C_1C_2*a*C_3(*a*) for such verbal roots with *a* as thematic vowel (e.g. *mašrab* 'drinking place' from *šariba/yašrabu* 'to drink'). The noun of instrument *ism al-ʾāla* is characterized by a *mi-* prefix (e.g. *miftāḥ* 'key'); in Classical Arabic this prefix can also denote intensity, e.g. *miḥrāb* 'very belligerent or *minṭīq* 'very articulate'.

Two other Classical Arabic patterns that are still productive are the noun of instance (*ism al-marra*), which indicates occurrences of an action (pattern C_1*a*C_2C_3*a*), and the noun of manner (*ism an-nawʿ*), which indicates the manner of an action (pattern C_1*i*C_2C_3*a*). Other types of nouns that are not denoted by a special term include the pattern C_1*u*C_2*ā*C_3, which denotes different kinds of illnesses, such as *buḥār* 'seasickness', *xumār* 'hangover', and *ṭuḥāl* 'infection of the spleen'.

The → *maṣdar*, properly meaning 'origin' and commonly referred to as the → 'verbal noun', is set on the borderline between noun and verb (see Wright 1967:I, 110–122) and is thus the prototype of a deverbal noun, corresponding to both an infinitive and a gerund in some European languages. The → participle (noun of agent/patient – *ism al-fāʿil/ism al-mafʿūl*) can likewise function as a substantive, an adjective, and – notably in the Arabic dialects – as a verbal continuous form or as a verbal form with resultative (perfective) function. For a synopsis relevant for both Classical Arabic and Modern Standard Arabic, see Wright (1967:I, 116ff.); for Forms XI–XV, see Fischer (1972:215); quadriliterals are only construed according to Forms II and V; Forms XI to XV are virtually unattested in Modern Standard Arabic. Retsö (1989) gives an overview of the diatheses in the Arabic dialects (for the semantic load of the individual forms/stems/diatheses, → aspect; → diathesis; → verb).

Interestingly, the question of directionality between the *maṣdar* and the finite verb forms also figures as one of the issues (*masāʾil*) discussed between the Basran and Kufan grammarians in Ibn al-ʾAnbārī's *ʾInṣāf* (102; no. 28 in the enumeration of the Leiden codex). The Kufan grammarians derived the *maṣdar* from the (finite) verb (*al-maṣdar muštaqq min al-fiʿl*), whereas the Basran grammarians derived the (finite) verb from the *maṣdar* (*al-fiʿl muštaqq min al-maṣdar*).

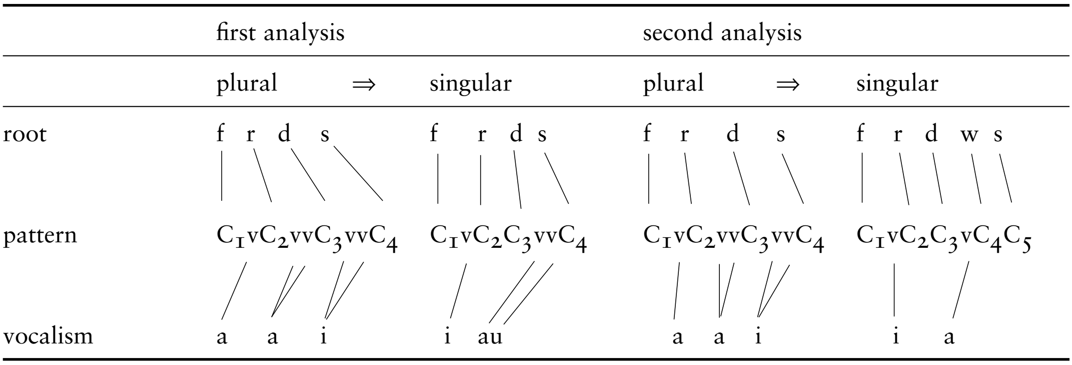
In modern times, many of the nouns of agent (active participles) designate not only persons but also instruments, e.g. *mukayyifa* (II) 'air conditioning', *mursila* (IV) 'transmitter'. The nouns of place or time of stems higher than the basic Form I use the nouns of patient (passive participle), e.g. *munṭalaq* 'point of departure (Form VII), *mustašfā* 'hospital (Form X) (see, e.g., Badawi a.o. 2004:89ff.; Holes 2004:151ff.).

In accordance with the terminology used in other Semitic languages, notably Akkadian (cf. Huehnergard 2000:24ff.), certain types of adjectives derived from the basic Form I are labeled 'verbal adjective' (see Wright 1967:I, 133ff. for the most comprehensive list). Noteworthy in a comparative Semitic perspective is, for instance, the pattern C_1*a*C_2*ī*C_3 of transitive verbs that has a passive meaning comparable to *ma*C_1C_2*ū*C_3, e.g. *qatīl* 'slain'. The pattern C_1*a*C_2*ī*C3 also has an active/stative meaning as 'participle' of verbs of the type C_1*a*C_2*u*C_3*a*, e.g. *jamīl* 'beautiful' (from *jamula*) and *ṯaqīl* 'heavy (from *ṯaqula*).

In certain cases, a noun of action has historically acquired the meaning of noun of actor associated with the action in question, notably *rasūl* 'message', which acquired the meaning of 'messenger' (cf. also Old Akkadian *šiprum* 'message' and later, 'messenger').

While the previous patterns usually are referred to as 'deverbal', there are also a number of notable 'denominal' patterns. Abstract nouns are regularly expressed by adding the feminine form of the → *nisba-* (gentilic) ending (e.g. *ʾimkāniyya* 'possibility'). Just like the noun of instrument (*ism al-ʾāla*, see above), the noun of vessel (*ism al-wiʿāʾ*) is expressed by a *mi-* prefix (e.g. *miʾbar* 'needle case' from *ʾibra* 'needle'). The noun of individuality (*ism al-waḥda*) designates one individual out of a species (e.g. *baqara* 'one cow' vs. *baqar* '[the species] cow').

Figure 1. Alternative analyses of *firdaws* ‘paradise’



An example of the noun of abundance (*ism al-kaṯra*) was given above (*ma'sada* ‘a place abounding in lions’).

While the root-pattern system in Arabic nominal (and verbal) morphology is highly regular, there are other synchronic and diachronic factors that determine the morphophonological analysis (and hence the place in the lexicon) of a number of nouns. This can be seen, for instance, when the different plural formations of the noun *madīna* ‘city’ are taken into consideration. While the older broken plural *madā'in* testifies to an analysis of this noun as a *ma*C_1C_2*a*C_3*a* (or *ma*C_1C_2*i*C_3*a*) pattern, the noun of place derived from the root *d-y-n*, the more recent broken plural *mudun* testifies to a synchronic reanalysis of the *ma*C_1C_2*v*C_3*a* pattern as C_1*a*C_2*ī*C_3*a*, as is clearly witnessed by the entry of this noun under the root *m-d-n* in the modern dictionaries (Wehr's dictionary actually has both entries, *d-y-n* and *m-d-n*). A similar analysis can be presented for the noun *taqwā* ‘faith’. While this noun synchronically must be analyzed as a C_1*a*C_2C_3*ā* pattern of the root *t-q-y*, a diachronic analysis will derive it as a back-formation of Form VIII *ittaqā* belonging to the root *w-q-y*. Such developments are by no means restricted to weak roots. The neologism *ta'aslum* ‘[passive] Islamization’, for instance, must be synchronically interpreted as a Form II *maṣdar* (corresponding to a Form V *maṣdar* of a triliteral root) of a quadriliteral root *'-s-l-m*, comparable to a root *'-m-r-k*, which yields *ta'amruk* ‘[passive] Americanization’. Only diachronically can one claim in this case that *'-s-l-m* reflects the Form IV of the root *s-l-m*, which has another meaning, namely ‘to surrender oneself [to God]’.

The extrapolation of consonants of non-Arabic lexical items that are incorporated into Arabic is a well-known process, attested as such in nearly all ancient and modern Semitic languages. One interesting example in point is the ‘retrograde’ derivation *firdaws* ‘paradise’ (< Greek *parádeisos*). *Parádeisos* is reanalyzed as the plural pattern C_1*a*C_2*ā*C_3*ī*C_4, to which a fictitious singular pattern C_1*i*C_2C_3*aw*C_4 would belong (thereby, *r* and *d* are being ‘mapped’ onto the C_2 and C_3 slots). Alternatively, *firdaws* could be analyzed as a quinquiliteral back-formation (such nouns are regularly ‘reduced’ to quadriliteral nouns in the plural, e.g. *ʿandalīb*/*ʿanādil* ‘nightingale’) of the attested pattern C_1*i*C_2C_3*a*C_4C_5, in which *w* is taken as a root consonant, and not as part of a diphthong. A formal representation of the two analyses is given in Figure 1.

While this lemma is not concerned with the noun phrase per se (→ noun phrase), there exist tendencies toward the formation of → compound nouns in Arabic that are of interest for morphology. Of the four types of *naḥt* (lit. ‘sculpture’) occurring in Arabic, the last three types affect the nominal realm (Stetkevych 1970:48–55):

i. *an-naḥt al-fiʿlī* (verbal): *samʿala* ‘to say *wa-s-salāmu ʿalay-kum*’
ii. *an-naḥt al-waṣfī* (adjectival): *ḍabaṭa* ‘to hold tight’ + *ḍabara* ‘to jump’ > *ḍibaṭr* ‘strong [said of a lion]’
iii. *an-naḥt al-ismī* (nominal): *jaluda* ‘to be strong’ + *jamada* ‘to become firm’ >*julmūd* ‘big rock’
iv. *an-naḥt an-nisbī* (referential): *ṭabarxazī* ‘belonging to Tabaristan and Khwarizm’

Compound formations (including the phenomenon of blends, where one or both of the constituting elements are shortened) are not unusual in Modern Standard Arabic (see, e.g., Brockelmann 1908:481–484; Monteil 1960:131–142; Blau 1981:172–174; El-Ayoubi a.o. 2001:49–51; Badawi a.o. 2004:58ff., 751ff.). One finds substantives and adjectives with prefixes, as well as structures with an internal appositional structure. (As the entry on → abbreviations shows, acronyms, which abound in modern Ivrit, are untypical of Modern Standard Arabic.)

lā-nihāʾiyya	‘infinity’
lā-sāmiyya	‘anti-Semitism’
ṣahyū-xūmayniyya	‘Zio-Khomeinism’
qāʾim-maqām	‘governor [Iraq]’
janūb-ʾifrīqī	‘South African’
šarq-ʾawsaṭī	‘Middle Eastern’
ʾafrū-ʾāsiyawī	‘Afro-Asiatic’
šibh-rasmī	‘semiofficial’
qab-tārīxī	‘prehistoric’
bar-māʾī	‘amphibic’
waṭanī-qawmī	‘ethnopolitical’
iqtiṣādī-ijtimāʿī	‘socioeconomic’

The phenomenon of *naḥt* (blending) is also productive in modern times (see, e.g., Versteegh 2001:181–183) e.g. *kahrabāʾ* ‘electricity’ + *maġnāṭīs* ‘magnet’ ⇒ *kahraṭas* ‘electromagnetism’. Such structures can in turn be the starting point of linear derivation with the *nisba*- and the feminine ending, e.g. *raʾsmāl* ‘capital’ ⇒ *raʾsmāliyya* ‘capitalism’; *lā-nihāʾī* ‘infinite’ ⇒ *lā-nihāʾiyya* ‘infinity’.

On the level of the noun phrase, annexations can be synchronically reanalyzed as compound nouns (or ‘quasi compounds’) in modern Arabic, e.g. *raʾs+māl* > *raʾsmāl* ‘capital’ ⇒ *ar-raʾsmāl* ‘the capital’ (?*raʾs al-māl*) (cf. Syriac *rēšmālā*); *ġayr(u) ḥaqīqi(yyin)* ‘untrue’; *ʿadam(u) tadaxxul(in)* ‘noninterference’.

Arabic nouns are either masculine (*muḏakkar*), feminine (*muʾannaṯ*), or of common gender (for details, see Wright 1967:I, 177ff.; → gender). Assuming masculine as the ‘default’ gender, feminine nouns may be feminine by virtue of denoting females, certain kinds of wind or fire, body parts occurring in pairs, or collective animal names. Alternatively, feminine nouns can be marked by the endings *tāʾ marbūṭa, ʾalif mamdūda* or *maqṣūra*, and *-āʾ* if that ending does not belong to the root (conversely, not all nouns with these endings are feminine, e.g. *xalīfa* ‘caliph’). A restricted number of Arabic nouns can take either gender, e.g. *ṭarīq* ‘path’.

Arabic nouns come in three numbers, the singular, the dual, and the plural (→ number). While the dual is always expressed by a suffix, there exist two kinds of plural, the sound (external) plural, *al-jamʿ al-muṣaḥḥaḥ*, and the broken (internal) plural, *al-jamʿ al-mukassar*.

Nouns designating persons, tribes, or places (→ proper nouns; → toponyms) have recourse to a wide variety of nominal patterns. Even verbal patterns (and whole verbal clauses) are attested, for instance *yazīd* ‘he augments’ or *taʾabbaṭa šarran* ‘he bore evil under the armpits’. Nominal patterns are often characterized by → diptosis, even when the respective noun as a common term would be triptotic (for details, see Wild 1982).

Loanwords denoting scientific vocabulary can either be fully integrated in the Arabic nominal system or retain (part of) their original structure. These principles can be followed back all the way to the transmitters of Greek terminology (via Syriac) into Arabic (see Schall 1982; Endreß 1992:14–23, for a typology of loan mechanisms; technical → terminology). A modern example of a fully integrated technical term is the neologism *raskala* ‘recycling’ ($C_1aC_2C_3aC_4a$). A modern example of a partially integrated technical term is the neologism *kibrītīd* ‘sulfide’, where a European-style suffix *-īd* is attached to the Arabic equivalent of ‘sulphur’, *kibrīt* (see Badawi a.o. 2004:741 for this example, and Ali 1987 in general on such issues).

3. The noun in the Arabic dialects

Noun patterns in the Arabic dialects feature exactly the same correspondence between form and semantic content that can be observed in Classical Arabic and Modern Standard Arabic. A comprehensive description of this situation can be found in Woidich (2006, Chap. 2.4 ‘Nomen’), one of the most detailed grammars of any Arabic dialect to date (here: → Cairene Arabic). Instead of the typology established above (including minimal triradical patterns, extended triradical patterns, and quadriradical or longer patterns), a typology based on syllable number, number of consonants, and featured preformatives and afformatives may be more appropriate to describe the situation in a modern Arabic dialect.

Many colloquial patterns can be derived from their 'underlying representation' in Classical Arabic, e.g. the diminutive pattern C$_1$(*u*)C$_2$*ē*C$_3$ from C$_1$*u*C$_2$*ay*C$_3$ (see, e.g., Broselow 1976 for Cairene Arabic). But new (or extended) types are found as well. A prominent example is given by the afformative *–ān*, which designates adjectives (e.g. *bardān* 'freezing', *šaʿbān* 'satiated'), certain animals (e.g. *tiʿbān* 'snake'), verbal nouns (e.g. *šukrān* 'thanking', *bunyān* 'building'), and certain professions (e.g. *rubbān* 'captain'), far beyond its scope in Classical Arabic. In Cairene Arabic (cf. Woidich 2006:10), the pattern C$_1$*a*C$_2$*a*C$_3$*ān* can express the verbal noun of virtually all verbs, e.g. *ġamaḏān* 'closing of the eyes', *ġalayān* 'cooking', *rawaḥān* 'going'.

In addition, Cairene Arabic features the following types of derivational suffixes, all except the first three representing innovations (cf. Woidich 2006: 105–109: *-a* (feminine, noun of individuality, etc.), *-i* (*nisba*), *-iyya* (certain devices, actions, abstract terms, etc.), *-āna* (derogatory), *-āwi* (provenance), *-āya* (diminutive, noun of individuality, etc.), *-āni* (body colors), *-āti* (certain professions), *-gi* (of Turkish origin: professions, certain characteristics, e.g. *bustāngi* 'gardener', *ixwāngi* 'Muslim brother'), *-angi* (negative characteristics), *-li* (relation), *-xāna* (building), *-u* (hypocoristic), *-āh* (regret), *-ti* (relation), *-ēra* (vessel).

In the incorporation of foreign words, a high degree of flexibility can be observed. Lower sociolects tend to follow indigenous noun patterns more closely in such processes (e.g. *garabuks* instead of *girbuks* 'gear box'). See Bolozky (1999) for an analysis of competing strategies in such cases of incorporation of loan vocabulary in modern Ivrit.

4. CONCLUSION

The formation of noun patterns has proven to be an open, productive process throughout the history of Arabic. The situation in the dialects reflects the precise mapping between morphological form and semantic content in the standard language. New patterns continue to emerge, e.g. the pattern C$_1$*aw*C$_2$*a*C$_3$*a* for denominal verbs from nouns with a long first syllable, e.g. *ʿawlama* 'globalization' from *ʿālam* 'world' (Badawi a.o. 2004:762). By coincidence, some patterns have become popular on the basis of → *qiyās* from one word, e.g. C$_1$*ā*C$_2$*ū*C$_3$ for technical instruments, a pattern which was triggered most likely by the noun *ḥāsūb* 'computer'. See also Kaye and Rosenhouse (1997:282) for the development of new verbal nouns in some Maghrebi dialects.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Ibn al-ʾAnbārī, *ʾInṣāf* = Kamāl ad-Dīn ʾAbū l-Barakāt ʿAbd ar-Raḥmān ibn ʾAbī Saʿīd al-ʾAnbārī, *Kitāb al-ʾinṣāf fī masāʾil al-xilāf bayna n-naḥwiyyīna l-baṣriyyīna wa-l-kūfiyyīn*. Ed. Gotthold Weil. Leiden: E.J. Brill, 1913.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. Ed. Būlāq, 1898–1900. (Repr., Baghdad: Muthanna, 1965.)

SECONDARY SOURCES

Ali, Abdul Sahib Mehdi. 1987. *A linguistic study of the development of scientific vocabulary in Standard Arabic*. London and New York: Kegan Paul International.

Badawi, El-Said, Michael G. Carter, and Adrian Gully. 2004. *Modern Written Arabic: A comprehensive grammar*. London: Routledge.

Barth, Johann. 1894. *Die Nominalbildung in den semitischen Sprachen*. Leipzig. (Repr., Hildesheim: G. Olms, 1967.)

Bauer, Laurie. 2000. "Word". *Morphologie: Ein internationales Handbuch zur Flexion und Wortbildung*, ed. Geert Booij, Christian Lehmann, and Joachim Mugdan, 247–257. Berlin and New York: Mouton de Gruyter.

Blau, Joshua. 1981. *The renaissance of Modern Hebrew and Modern Standard Arabic: Parallels and differences in the revival of two Semitic languages*. Berkeley and Los Angeles: University of California Press.

Bohas, Georges, Jean-Patrick Guillaume, and Djamel Eddine Kouloughli. 1990. *The Arabic linguistic tradition*. London and New York: Routledge.

Bolozky, Shmuel. 1999. *Measuring productivity in word formation: The case of Israeli Hebrew*. Leiden: E.J. Brill.

Brockelmann, Carl. 1908. *Grundriß der vergleichenden Grammatik der semitischen Sprachen*. I. *Laut- und Formenlehre*. Berlin: Reuther und Reichard. (Repr., Hildesheim: G. Olms, 1999.)

Broselow, Ellen. 1976. *The phonology of Egyptian Arabic*. Ph.D. diss., University of Massachusetts, Amherst.

Diem, Werner. 1970–1971. "Nomen, Substantiv und Adjektiv bei den arabischen Grammatikern". *Oriens* 23–24.312–332.

El-Ayoubi, Hashem, Wolfdietrich Fischer, and Michael Langer. 2001. *Syntax der arabischen Schriftsprache der Gegenwart*. I/1. *Das Nomen und sein Umfeld*. Wiesbaden: L. Reichert.

Endreß, Gerhard. 1992. "Die wissenschaftliche

Literatur". *Grundriß der arabischen Philologie. III. Supplement*, ed. Wolfdietrich Fischer, 3–152. Wiesbaden: L. Reichert.

Fischer, Wolfdietrich. 1972. *Grammatik des klassischen Arabisch*. Wiesbaden: O. Harrassowitz.

——. 1997. "Classical Arabic". *The Semitic languages*, ed. Robert Hetzron, 187–219. London: Routledge.

——, and Otto Jastrow (eds.). 1980. *Handbuch der arabischen Dialekte*. Wiesbaden: O. Harrassowitz.

Holes, Clive. 2004. *Modern Arabic: Structures, functions, and varieties*. 2nd ed. Washington, D.C.: Georgetown University Press.

Huehnergard, John. 2000. *A grammar of Akkadian*. 2nd ed. Winona Lake, Ind.: Eisenbrauns.

Kaye, Alan and Judith Rosenhouse. 1997. "Arabic dialects and Maltese". *The Semitic languages*, ed. Robert Hetzron, 263–311. London: Routledge.

Larcher, Pierre. 1999. "Vues 'nouvelles' sur la dérivation lexicale en arabe classique". *Tradition and innovation: Norm and deviation in Arabic and Semitic linguistics*, ed. Lutz Edzard and Mohammed Nekroumi, 103–123. Wiesbaden: O. Harrassowitz.

Lipiński, Edward. 2001. *Semitic languages: Outline of a comparative grammar*. 2nd ed. Leuven: Peeters.

McCarthy, John and Alan Prince. 1995. "Prosodic Morphology". *The handbook of phonological theory*, ed. John A. Goldsmith, 318–366. Cambridge, Mass.: Blackwell.

Monteil, Vincent. 1960. *L'arabe moderne*. Paris: C. Klincksieck.

Owens, Jonathan. 1984. "Structure, class and dependency: Modern linguistic theory and the Arabic grammatical tradition". *Lingua* 64.25–62.

Ratcliffe, Robert. 1997. "Prosodic templates in a word-based morphological analysis of Arabic". *Perspectives on Arabic linguistics*, X, ed. Mushira Eid and Robert R. Ratcliffe, 93–122. Amsterdam and Philadelphia: J. Benjamins.

——. 1998. *The 'broken' plural problem in Arabic and comparative Semitic: Allomorphy and analogy in non-concatenative morphology*. Amsterdam and Philadelphia: J. Benjamins.

Retsö, Jan. 1989. *Diathesis in the Semitic languages*. Leiden: E.J. Brill.

Schall, Anton. 1982. "Geschichte des arabischen Wortschatzes: Lehn- und Fremdwörter im Klassischen Arabisch". *Grundriß der Arabischen Philologie. I. Sprachwissenschaft*, ed. Wolfdietrich Fischer, 142–153. Wiesbaden: L. Reichert.

Shimron, Joseph. 2003. "Semitic languages: Are they really root-based?" *Language processing and acquisition in languages of Semitic, root-based morphology*, ed. Joseph Shimron, 1–28. Amsterdam and Philadelphia: J. Benjamins.

Stetkevych, Jaroslav. 1970. *The modern Arabic literary language: Lexical and stylistic developments*. Chicago: University of Chicago Press.

Versteegh, Kees. 2001. *The Arabic language*. 2nd ed. Edinburgh: Edinburgh University Press.

Wild, Stefan. 1982. "Arabische Eigennamen". *Grundriß der Arabischen Philologie. I. Sprachwissenschaft*, ed. Wolfdietrich Fischer, 154–164. Wiesbaden: L. Reichert.

Woidich, Manfred. 2006. *Das Kairenisch-Arabische: Eine Grammatik*. Wiesbaden: O. Harrassowitz.

Wright, William. 1967. *A grammar of the Arabic language*. Trans. from the German of Caspari. 3rd ed., rev. W. Robertson Smith and M.J. de Goeje. 2 vols. Cambridge: Cambridge University Press.

LUTZ EDZARD (University of Oslo)

Noun Phrase

1. OVERVIEW

The Arabic noun phrase (*tarkīb ismī*) is a syntactic constituent consisting of a noun (→ *ism*) or verbal noun (→ *maṣdar*) and its dependents or modifiers. The Arabic noun phrase has been one of the major preoccupations of researchers studying Arabic syntax from a variety of theoretical and methodological points of view, and a very extensive literature is dedicated to it.

General studies of the Arabic noun phrase have been done by Ayoub (1981), Ditters (1992), Fassi Fehri (1993), and Kremers (2003). The greater part of the literature on Arabic noun phrases focuses on more specific aspects of its grammar.

2. PROPERTIES OF NOUN PHRASES

Arabic nouns have several morphosyntactic and semantic properties. These include definiteness, case (→ declension), → gender, → number, → specificity, and individuation. Although some of these properties are expressed through noun-internal morphology, the categories they express affect the syntactic distribution and semantic interpretation of the noun phrase as a whole, and as such they are properties of noun phrases rather than just nouns.

2.1 *Definiteness*

Arabic noun phrases are either definite or indefinite. Definiteness is a morphosyntactic rather than semantic property, but it generally coincides with the semantic/pragmatic categories of familiarity (*taʿrīf*) and/or → specificity (*taxṣīṣ*). The definiteness or indefiniteness of a noun is expressed by the presence or absence of the definite article (→ article, definite), but either can be further augmented with various → determiners, such as quantificational determiners (*kull* ‘every’, *lā* ‘no’, *baʿḍ* ‘some’, etc.; → quantifiers), → demonstrative adjectives (*hāḏā* ‘this [masc. sg.]’, *hāḏihi* ‘this [fem. sg.]’, etc.), and, in some dialects, indefinite articles (Lebanese, Moroccan *ši* ‘some’, Iraqi *fad* ‘a; some’, *wāḥid* ‘one’).

A noun phrase as a whole is definite if it is headed by a noun marked with the definite article (*al-kitāb-u* ‘the book’), or if it is a → construct state possessive construction with a definite possessor (*muḍāf ʾilayhi*).

There is considerable variation in indefiniteness marking. In fully inflected Standard Arabic, most morphologically indefinite nouns that are declinable (*muʿrab*) are marked with a final *-n* following the case-vowel (→ *tanwīn* ‘addition of *-n*’), as in *kitāb-un* ‘a book’. In the dialects, degrees of indefiniteness are distinguished through different strategies. In the Najdi Bedouin dialects of Saudi Arabia (Abboud 1964; Ingham 1994a), as well as in some dialects of Central Asia (Fischer 1961; Ingham 1994b), the Sudan (Owens 1993), and Yemen (Behnstedt 1985), vestigial *tanwīn* is used on indefinite nouns to indicate degrees of specificity. Other dialects use forms of indefinite articles, such as *šī* ‘some’ in Moroccan Arabic (Harrell 1962; Harrell and Sobelman 1966) and some Levantine dialects (Feghali 1923; Cowell 1964), *fad* in Iraqi (Van Ess 1961; Wallace 1969), and derivatives of *wāḥid* ‘one’ in various dialects (see Brustad 2000, Chap. 1) to indicate various degrees of specificity.

In contemporary Western syntactic theory, the Arabic definite article has been analyzed as a nominal affix; the fact that the article assimilates to the place and manner of articulation of a stem-initial sun letter has been cited as evidence of this. In transformational approaches to syntax, the affixhood of the article has been analyzed by treating it as the head of a Determiner Phrase (DP), to which the nominal head raises and incorporates (Mohammad 1988; Fassi-Fehri 1993; Benmamoun 2000), as in (1).

(1a) *al-kitāb-u*
the-book-Nom

(1b) $[_{\mathrm{DP}}\ [_{\mathrm{D}}\ \text{al- kitāb-u}_{i}\ ]\ [_{\mathrm{NP}}\ t_{i}\ ]\ ]$

However, in some dialects and in Modern Standard Arabic, the definite article appears to be a phrasal affix or → clitic which attaches to the noun phrase constituent as a whole, rather than directly to the noun stem. Evidence for this are noun phrases in which numeral modifiers occur prenominally and host the definite article, as in (2)–(5) (Hoyt 2000, 2002; Shlonsky 2004; Borer 1996 adduces similar arguments for Hebrew).

(2) *haṯ-ṯalāṯ* *ḥabbāt*
these-three pills
‘these three pills’ (Palestinian Arabic: Schmidt and Kahle 1918:31.6)

(3) *bāʿ* *el-ʾarbaʿ* *ʾeḥṣān*
sell.Perf.3ms the-four horses.pl
el-mlāḥ
the-good.pl
‘He sold the four good horses’ (Lebanese Arabic: Feghali 1928:190)

(4) *al-xams-u* *qur-an*
the-five-Nom villages-Gen
‘the five villages’ (Wright 1898:II, 264)

(5) *aṯ-ṯalāṯat-u* *rijāl-in*
the-three-Nom men-Gen
‘the three men’ (Wright 1898:II, 264)

Accordingly, analyses have been proposed according to which → numerals occupy a position between the article and the noun stem (Hoyt 2000, 2002; Shlonsky 2004; see Ritter 1991 for similar constructions in Hebrew).

2.2 *Case*

Nominal case marking (→ *ʾiʿrāb* ‘declension’) is exclusive to more formal registers of Standard Arabic and has been entirely lost in the dialects (but see Owens 1998a, 1998b, for arguments that the dialects have never had case marking). In Standard Arabic with full *ʾiʿrāb*, nouns with regular (*muʿrab* ‘inflectable’) form are marked with one of three cases: the nominative (*marfūʿ* ‘raised’), the accusative (*manṣūb* ‘lifted’), and the genitive (*majrūr* ‘attracted’), marked with the suffixes *-u*, *-a*, and *-i*, respectively. Singular nouns with irregular (*mabnī* ‘fixed’) form are not marked for case (→ *bināʾ*).

2.3 *Gender*

Grammatical gender in Arabic corresponds to several different semantic categories, including biological feminine gender as well as inanimate plurality, individuation of mass nouns, and, in some dialects, unindividuated human plurals (see below). Gender is indicated morphologically only in specific classes of words, but it plays an important role in → agreement marking and other concordial relations.

Feminine gender is most often expressed with the 'bound *-t*' morpheme (*tāʾ marbūṭa*). The bound *-t* morpheme is used to express feminine sex (*ṭālib* 'student'/*ṭālib-a* 'female student'); individual instances of mass noun denotations (*baqar* 'cattle'/*baqar-a* '[a] cow'); collective plurals of sound plural nouns (*al-filasṭīniyyūn* 'the Palestinians'/*al-filasṭīniyy-a* 'the Palestinians [as a whole]'); devices used for specific purposes (*šams* 'sun', *šamsiyy-a* 'umbrella, parasol'); or more specific meanings of preexisting nouns (*maktab* 'office', *maktab-a* 'library'); and, in a few cases, nouns denoting biologically male entities (e.g. *xalīfa* 'caliph'; Levantine colloquial *zalami* 'guy'). Other feminine nouns are unmarked for gender and are feminine either because they denote biologically feminine persons (*bint* 'girl, daughter', *ʿarūs* 'bride'), or by convention (*ḥarb* 'war', *dār* 'house', *šams* 'sun').

Masculine gender is unmarked on singular nouns except for relative (→ *nisba*) adjectives and de-adjectival nouns derived from *nisba* adjectives (*ʿarab-iyy* 'Arabic man' vs. *ʿarab-iyy-a* 'Arabic woman'), and is only marked in plural nouns with sound plural forms (*mudarrisūn* '[male] teachers'/*mudarrisāt* 'female teachers').

2.4 *Number*

Number is explicitly marked only on singular nouns ending with the bound *-t* morpheme and relational adjectives. Arabic plural morphology is extremely complex and is divided into two paradigms: sound plurals and broken plurals. Sound plurals are marked with suffixal endings, while broken plurals are expressed with a range of roughly thirty different forms of varying degrees of productivity and semantic specificity (Levy 1971; McCarthy and Prince 1990). Mass nouns take two plural forms; the plural of plenty, usually expressed as a broken plural, expresses different varieties of the object in question, while the paucal plural, usually a sound plural, expresses a specific and limited number of objects (e.g. *šajar* 'trees', *ʾašjār* 'kinds of trees', *šajara* 'a tree', *šajarāt* '[some] trees.'

2.5 *Individuation*

Individuation (*tafrīd*) is the degree to which the members of the denotation of a plural or mass noun can be counted or identified as individuals. It is closely related to specificity, and so many of the devices used to express specificity imply individuation as well.

In some dialects, plural nouns with low individuation sometimes control feminine singular agreement even though they denote groups of human males (Sallam 1979; Belnap 1991; Brustad 2000). In (7), from a story in the Bir Zeit dialect of rural Palestinian Arabic, describing an attack by a group of Bedouin tribesmen on a Turkish governor's palace, two conjoined clauses show different agreement forms for this group.

(6) *lammin* *simʿu* *hāḏā*
when heard.Perf.3mp that
kāmat *il-ʿurbān*
rose.Perf.3fs the-Bedouin
tihḍim *ʿa-s-sarāya*
attack.Imperf.3fs upon-the-palace
'When they heard that, the Bedouin attacked the palace' (Schmidt and Kahle 1918: §15.3)

The sheikh of the tribe instructs his followers to listen for a pistol shot as the signal to attack, and goes inside to confront the governor. The first clause describes the followers' hearing the signal and shows the verb *simʿu* 'they heard' marked in the masculine plural. In contrast, the main clause describes them attacking the palace, and the verb sequence is marked in the feminine singular *kāmat tihḍim*, presumably indicating that they attacked en masse (and are therefore not individuated). Belnap (1991) notes similar alternations in Cairene Egyptian Arabic.

In dialects which distinguish gender in the plural, nonsentient nouns with plural denotations are marked as feminine in the plural but can control either singular or plural agreement, the former expressing less individuation and the latter more. To illustrate again with examples from Bir Zeit Palestinian Arabic, (7) shows the nonspecific indefinite *člāb* 'dogs' controlling feminine singular agreement on the verb, while (8) shows the specific-marked indefinite *hač-člāb* 'these dogs' controlling feminine plural agreement.

(7) *hāḏa lammin ṭiliʿ*
this.ms when left.3ms
twaršaʿat-e čläb
attacked.3fs-cl3ms dogs
'When he left, dogs attacked him' (Schmidt and Kahle 1918:§38.5)

(8) *ʿāwadat illa w-hač-člāb*
returned.3fs except and-these-dogs
b-ōčilin fi-č-čbēbāt
Ind-eat.3fp in-the-meatballs
min rakbat-e
from neck-cl3ms
'She went back, and there were these dogs eating the meatballs from his neck' (Schmidt and Kahle 1918: §30.11)

3. Adjectival modification

Arabic noun phrases can include modification by simplex → adjectives or → adjective phrases, as well as other modifiers, such as → relative clauses and prepositional phrases. Modifiers generally follow the nouns they modify, as in (9)–(11).

(9) *al-kitāb-u l-kabīr-u*
the-book.ms-Nom the-big.ms-Nom
'the big book'

(10) *ar-rajul-u llaḏī*
the-man.msg-Nom Rel.ms
yuḥibbu-hā
Ind-love.3ms-cl3fs
'the man who loves her'

(11) *madīnat-un fī miṣr-a*
city.fs-Nom in Egypt-Gen
'a city in Egypt'

However, → elative adjectives and ordinal numerals can precede the nouns they modify in a construct state, as in (12).

(12) *ʾakbar bayt-in*
biggest house-Gen
'[the] biggest house'

(13) *ʾawwal yawm-in*
first day-Gen
'[the] first day'

The preferred order of postnominal adjectives mirrors that of English adjectives, with more inherent properties following less inherent ones (Fassi Fehri 1998). In noun phrases containing both adjectival modifiers and other kinds, such as relative clauses or prepositional phrases, adjectival modifiers precede the others.

A particularly interesting form of adjectival modification common in Standard Arabic involves an adjectival construct or 'pseudo-construct' (*ʾiḍāfa ġayr ḥaqīqiyya*; see Siloni 2000, 2002 for the equivalent phenomenon in Hebrew). This consists of an adjective or participle 'possessed' by a noun (referred to here as the 'inner noun') specifying the degree or manner of the property expressed by the adjective. This constituent then modifies or is predicated of another noun (the 'outer noun'), as in (14) and (15).

(14) *ṭawīl-u ṣ-ṣabr-i*
long-Nom the-patience-Gen
'long of patience, very patient'

(15) *kabīr-u s-sinn-i*
big-Nom the-age-Gen
'great of age, elderly'

The adjective agrees with the outer noun, just as any other adjective does, in number and gender, as in (16) and (17), as well as in definiteness, as in (18) and (19).

(16) *ḥasan-un ṭawīl-u*
Hasan.ms-Nom long.msg-Nom
ṣ-ṣabr-i
the-patience-Gen
'Hasan is very patient'

(17) *laylā ṭawīlat-u*
Layla.fs-Nom long.fs-Nom
ṣ-ṣabr-i
the-patience-Gen
'Layla is very patient'

(18) *bint-un jamīlat-u*
girl.fs-Nom beautiful.fs-Nom
l-wajh-i
the-face-Gen
'a girl with a beautiful face'

(19) *al-bint-u l-jamīlat-u*
the-girl.fs-Nom the-beautiful.fs-Nom
l-wajh-i
the-face-Gen
'the girl with a beautiful face'

Note that, although the adjective agrees with the outer noun, its meaning is applied to the inner noun. For example, in (17) *ṭawīlatun* 'long' modifies *ṣabr* 'patience', a noun with masculine gender although it agrees with *laylā*, a feminine noun.

The meaning of the pseudoconstruct suggests that it is interpreted by means of abstraction over an understood possessor of the inner noun (where 'possessor' is construed very generally). For example, in (20) Hasan is understood as the possessor of the face referred to.

(20) *ḥasan-un* *jamīl-u*
Hasan.ms-Nom beautiful.ms-Nom
l-wajh-i
the-face-Gen
'Hasan has a beautiful face'

With respect to its interpretation, the adjectival construct resembles two other constructions: an adjective modified by the accusative of specification (→ *tamyīz*), as in (21), and the connected modifier (*tasbīb an-naˁt*; → *sabab*), as in (22).

(21) *ḥasan-un* *jamīl-un*
Hasan.ms-Nom beautiful.ms-Nom
wajhan
face-Acc
'Hasan has a beautiful face'

(22) *ḥasan-un* *al-jamīl-u*
Hasan-Nom the-beautiful.ms-Nom
wajh-u-hu
face-Nom-cl3ms
'Hasan, with his beautiful face'

In (21), *wajhan* 'face' is marked with the adverbial accusative and specifies the manner in which *jamīl* 'beautiful' can be an attribute of *ḥasan*. In (22), the possessor of the face is explicitly marked and is bound by *ḥasan*. The adjective phrases in (20), (21), and (22) are therefore interpreted as the property of having a beautiful face.

4. Possessive noun phrases

Perhaps the best known and most studied kind of noun phrase in Arabic is the → construct state possessive construction, a construction that Arabic shares with other Semitic languages and in particular Hebrew (Borer 1981, 1996; Hazout 1991, 1994, 1995, 2000; Siloni 1996, 1997a, 1997b, 2000, 2002). Theoretical analyses of the construct state have focused on parallels between its structure and the structure of full clauses, particularly in word order.

In Standard Arabic, the construct state expresses a wide range of possessive and partitive relationships. These include both material and inalienable → possession, location, part/whole relationships, measure or quantity, and comparison. In the dialects, the construct state is used with varying productivity and competes with the → analytic genitive for expressing the various genitive relationships (see Harning 1980; Mohammad 1999; Brustad 2000:70-88; Holes 2004:208–210).

The parallels between the construct state and a verbal clause are particularly clear in verbal constructs, and as such, much of the literature on the construct state focuses on examples in which the construct head is a *verbal noun* (→ *maṣdar* 'source') and the possessor its subject or object (Aoun 1978; Ayoub 1985; Aboudi 1985; Mohammad 1988, 1999; Fassi Fehri 1993; Benmamoun 2000; Kremers 2003; Shlonsky 2004), as in (23) and (24).

(23) *kitābat-u* *l-walad-i*
writing.fs-Nom the-boy-Gen
li-l-wājib-i
to-the-assignment
'the boy's writing of the assignment'

(24) *ʾixfāʾ-u* *l-māl-i* *ˁalā*
hiding-Nom the-money-Gen from
š-šurṭat-i
the-police-Gen
'the hiding of the money from the police'

Additional arguments can follow the inner NP either in the accusative case or marked with a preposition, and are subject to an ordering restriction which parallels the ordering restrictions on arguments in verbal clauses, namely that the possessor-NP must be the least oblique.

Most analyses of the verbal construct argue that the lexical phrase is a → verb phrase rather than a noun phrase. The verbal noun is the head of the verb phrase, and it raises and adjoins to the head position of DP, with the inner noun in a VP-internal subject position, and any objects in the complement of VP, as in the analysis of (25) in (26) and (27).

(25) *taksīr-u* *l-bint-i*
breaking-Nom the-girl-Gen
š-šubbāk-a
the-window-Acc
'the girl's breaking the window'

(26) [$_{\text{DP}}$ D [$_{\text{VP}}$ [$_{\text{DP}}$ l-bint-i] [$_{\text{V'}}$ taksīr-u
[$_{\text{DP}}$ š-šubbāk-a]]]]

(27) [$_{\text{DP}}$ [$_{\text{D}}$ D taksīr-u$_i$] [$_{\text{VP}}$ [$_{\text{DP}}$ l-bint-i]
[$_{\text{V'}}$ t_i [$_{\text{DP}}$ š-šubbāk-a]]]]

However, Fassi Fehri (1993) notes that if the object is marked with the accusative case, the verbal noun can be modified with adverbs (28a), while if it is marked with the preposition *li-*, the verbal noun must be modified with an adjective (28b):

(28a) *intiqād-u* *r-rajul-i*
criticizing-Nom the-man-Gen
bi-stimrār-in
with-persistence-Gen
al-mašrūˁ-a
the-project.ms.Acc
'the man's persistently criticizing the project'

(28b) *intiqād-u* *r-rajul-i*
criticism.ms-Nom the-man-Gen
l-mustamirr-u
the-persistent.msg-Nom
li-l-mašrūˁ-i
to-the-project-Gen
'the man's persistent criticism of the project'

This suggests that verbal nouns which mark accusative case on their objects are more 'verb-like' than those which do not. Fassi Fehri argues that more 'verb-like' construct states consist of a verb phrase which is dominated by the DP (29), while more 'noun-like' constructs consist of a noun phrase dominated by the DP (30).

(29) [$_{\text{DP}}$ [$_{\text{D}}$ D intiqād-u$_i$] [$_{\text{VP}}$ [$_{\text{DP}}$ r-rajul-i]
[$_{\text{V'}}$ t_i [$_{\text{DP}}$ al-mašrūˁ-a]]]

(30) [$_{\text{DP}}$ [$_{\text{D}}$ D intiqād-u$_i$] [$_{\text{NP}}$ [$_{\text{DP}}$ r-rajul-i]
[$_{\text{N'}}$ t_i [$_{\text{PP}}$ li-l-mašrūˁ-a]]]

The parallel with verbal clauses noted for the construct state has also been noted to hold for analytic genitive constructions. Mohammad (1999) proposes that the analytic possessive in Palestinian and Standard Arabic has a DP structure which mirrors the structure of clauses in SVO word order. The possessed noun is the 'subject' of the DP headed by the possessive particle, with which it agrees under a specifier/head relation. For example, the noun phrase *iṭ-ṭawla tabaˁit aḥmad* 'Ahmad's book' (Mohammad 1999) would have the structure in (31).

(31) [$_{\text{DP}}$ [$_{\text{DP}}$ iṭ-ṭawla] [$_{\text{D'}}$ tabaˁit [$_{\text{NP}}$ aḥmad]]]

Mohammad argues that this analysis supports the overall approach of assigning a parallel structure to Arabic noun phrases and clauses.

BIBLIOGRAPHICAL REFERENCES

Abboud, Peter F. 1964. *The syntax of Najdi Arabic.* Ph.D. diss., University of Michigan, Ann Arbor.

Aboudi, Fahmi B.S. 1985. *The syntax of* ʾiḍāfah *annexation in Arabic.* Ph.D. diss., Indiana University.

Aoun, Joseph. 1978. "Structure interne du groupe nominal en arabe: L'*ʾiḍāfa*". *Analyses Théorie* 2.1–40.

Ayoub, Georgine. 1981. *Structure de la phrase en arabe standard.* Ph.D. diss., Université de Paris–7.

Behnstedt, Peter. 1985. *Die nord-jemenitischen Dialekte. I. Atlas.* Wiesbaden: L. Reichert.

Belnap, Kirk R. 1991. *Grammatical agreement variation in Cairene Arabic.* Ph.D. diss., University of Pennsylvania.

Benmamoun, Elabbas. 2000. *The feature structure of functional categories: A comparative study of Arabic dialects.* Oxford: Oxford University Press.

Blau, Joshua. 1960. *Syntax des palästinensischen Bauerndialektes Bir Zeits.* Walldorf, Hessen: Verlag für Orientkunde.

Borer, Hagit. 1988. "On the morphological parallelism between compounds and constructs". *Yearbook of morphology*, I, ed. Geert Booij and Jaap van Marle, 45–65. Dordrecht: Foris.

——. 1996. "The construct in review". Lecarme a.o. (1996:30–61).

Brustad, Kristen. 2000. *The syntax of spoken Arabic: A comprehensive study of Moroccan, Egyptian, Syrian, and Kuwaiti dialects.* Washington, D.C.: Georgetown University Press.

Cowell, Mark. 1964. *A reference grammar of Syrian Arabic.* Washington, D.C.: Georgetown University Press.

Ditters, Everhard. 1992. *A formal approach to*

Arabic: The noun phrase and the verb phrase. Ph.D. diss., University of Nijmegen.

Fassi Fehri, Abdelkader. 1993. *Issues in Arabic clauses and words.* Dordrecht: Kluwer Academic Publishers.

——. 1998. "Layers in the distribution of Arabic adverbs and adjectives and their licensing". *Perspectives on Arabic linguistics*, XI, ed. Elabbas Benmamoun, Mushira Eid, and Niloofar Haeri, 1–45. Amsterdam and Philadelphia: J. Benjamins.

Feghali, Michel. 1928. *Syntaxe des parlers actuels du Liban.* Paris: Champion.

Fischer, Wolfdietrich. 1961. "Die Sprache der arabischen Sprachinsel in Uzbekistan". *Der Islam* 36.232–263.

Harning, Kerstin Eksell. 1980. *The analytical genitive in the modern Arabic dialects.* Stockholm: Acta Universitatis Gothoburgensis.

Harrell, Richard. 1962. *A short reference grammar of Moroccan Arabic.* Washington, D.C.: Georgetown University Press.

——, and H. Sobelman. 1966. *A dictionary of Moroccan Arabic.* Washington, D.C.: Georgetown University Press.

Hazout, Ilan. 1991. *Verbal nouns: Theta-theoretic studies in Hebrew and Arabic.* Ph.D. diss., University of Massachusetts at Amherst.

——. 1994. "Nominalizers in theta theory". *Linguistic Review* 11.5–48.

——. 1995. "Action nominalizations and linguistic theory". *Natural Language and Linguistic Theory* 13.355–404.

——. 2000. "Adjectival genitive constructions in Modern Hebrew: A case study in coanalysis". *The Linguistic Review* 17.29–52.

Holes, Clive. 2004. *Modern Arabic: Structures, functions, and varieties.* 2nd rev. ed. New York: Longman.

Hoyt, Frederick. 2000. *Word order, agreement, and specificity effects in rural Palestinian Arabic existential constructions.* M.A. thesis, Cornell University.

——. 2002. "Reduced agreement as a specificity effect in Rural Palestinian Arabic". *Perspectives on Arabic Linguistics*, XIII–XIV, ed. Elabbas Benmamoun, 111–141. Amsterdam and Philadelphia: J. Benjamins.

Ingham, Bruce. 1994a. *Najdi Arabic: Central Arabian.* Amsterdam: J. Benjamins.

——. 1994b. "The effect of language contact on the Arabic dialect of Afghanistan". *Actas del Congreso Internacional sobre Interferencias Linguísticas Árabe-Romances y Paralelos Extra-Iberos*, ed. Jordi Aguadé, Federico Corriente, and Marina Marugán, 105–118. Saragossa: Navarro y Navarro.

Kremers, Joost Merijn. 2003. *The Arabic noun phrase: A minimalist approach.* Ph.D. diss., University of Nijmegen.

Lecarme, Jacqueline, Jean Lowenstamm, and Ur Shlonsky (eds). 1996. *Studies in Afro-Asiatic languages.* The Hague: Holland Academic Graphics.

Levy, Mary. 1971. *The plural of the noun in Modern Standard Arabic.* Ph.D. diss., University of Michigan.

McCarthy, John J. and Alan S. Prince. 1990. "Foot and word in prosodic morphology: The Arabic broken plural". *Natural Language and Linguistic Theory* 8.209–283.

Mohammad, Mohammad. 1988. "On the parallelism between IP and DP". *Proceedings of the West Coast Conference on Formal Linguistics (WCCFL)*, VII, ed. Hagit Borer, 241–254. Stanford, Calif.: CSLI.

——. 1999. "Checking and licensing inside DP in Palestinian Arabic". *Perspectives on Arabic linguistics*, XII, ed. Elabbas Benmamoun, 27–44. Amsterdam and Philadelphia: J. Benjamins.

Ouhalla, Jamal and Ur Shlonsky (eds.). 2002. *Themes in Arabic and Hebrew syntax.* Dordrecht: Kluwer Academic Publishers.

Owens, Jonathan. 1993. "Nigerian Arabic in comparative perspective". *Sprache und Geschichte in Afrika* 14.85–175.

——. 1998a. "Case and Proto-Arabic, I". *Bulletin of the School of Oriental and African Studies* 61:1.51–73.

——. 1998b. "Case and Proto-Arabic, II". *Bulletin of the School of Oriental and African Studies* 61:2.215–227.

Ritter, Elizabeth. 1991. "Functional categories in noun phrases: Evidence from Modern Hebrew". *Perspectives on phrase structure*, ed. Susan Rothstein, 37–62. New York: Academic Press.

Sallam, A.M. 1979. "Concordial relations within the noun phrase in Educated Spoken Arabic (ESA)". *Archivum Linguisticum* 10.20–56.

Schmidt, Hans and Paul Kahle. 1918. *Volkserzählungen aus Palästina*, I. Göttingen: Vandenhoeck und Ruprecht.

——, and Paul Kahle. 1930. *Volkserzählungen aus Palästina*, II. Göttingen: Vandenhoeck und Ruprecht.

Shlonsky, Ur. 2004. "The form of Semitic noun phrases". *Lingua* 114.1465–1526.

Siloni, Tal. 1996. "Hebrew noun phrases: Generalized noun raising". *Parameters and functional heads: Essays in comparative syntax*, ed. Adrianna Belletti and Luigi Rizzi, 239–267. Oxford: Oxford University Press.

——. 1997a. "Event nominals and the construct state". *New comparative syntax*, ed. Liliane Haegeman, 165–188. London: Longman.

——. 1997b. *Noun phrases and nominalizations: The syntax of DPs.* Dordrecht: Kluwer Academic Publishers.

——. 2000. "Non-nominal constructs". *Research in Afroasiatic grammar*, II, ed. Jacqueline Lecarme, Jean Lowenstamm, and Ur Shlonsky, 481–510. Amsterdam and Philadelphia: J. Benjamins.

——. 2002. "Adjectival constructs and inalienable constructions". Ouhalla and Shlonsky (2002:161–187).

Van Ess, John. 1961. *The spoken Arabic of Iraq.* 2nd ed. London: Oxford University Press.

Wallace, Erwin. 1969. *A basic course in Iraqi Arabic.* Washington, D.C.: Georgetown University Press.

Wright, William. 1898. *A grammar of the Arabic language, translated from the German of Caspari and edited with numerous additions and corrections.* 2 vols. 3rd ed., rev. W. Robertson Smith and M.J. de Goeje. Cambridge: Cambridge University Press.

Frederick Hoyt (University of Texas at Austin)

Nubi → Ki-Nubi

Nubian

Loanwords are lexical items which are phonetically and morphologically integrated into a language other than the one where they originated. The prerequisite for such lexical items to be incorporated in another language is the presence of bilingual individuals. Arabic/Nubian bilingualism can be characterized as replacive in the sense that the Nubian languages are threatened by complete replacement by Arabic.

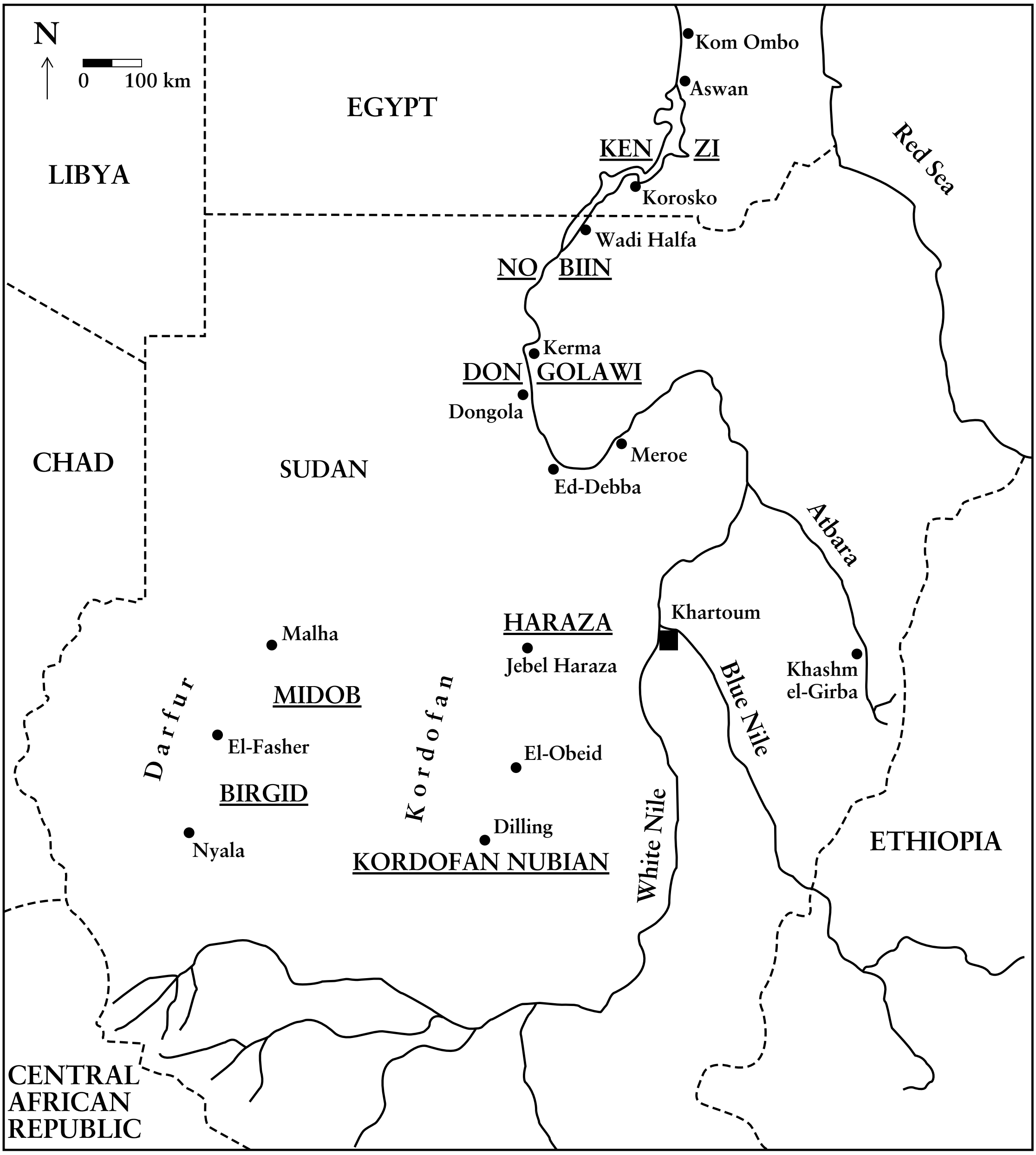
The Nubian languages are scattered today over a large area comprising both the northern half of the Republic of Sudan and southern Egypt. They form a language family that can be divided into three geographically defined branches (Map 1): Darfur Nubian in the western Sudan, comprising Midob and Birgid; Kordofan Nubian, a group of closely related dialects spoken in the northern Nuba Mountains; and Nile Nubian, comprising Kenzi-Dongolawi, Nobiin, and its medieval predecessor, Old Nubian. Before the successive building of the dams near Aswan, the Nile Nubian languages were spoken in the Middle Nile Valley between the first Nile cataract upstream to Debba at the great bend of the river. After the inundation of a large part of Nubia in the waters of Lake Nasser, a considerable part of the local Nubian population was displaced. The resettlement – outside of Nubia, near Kom Ombo in Egypt, and in the Khashm el-Girba scheme of the eastern Sudan – has contributed considerably to the decay of Nubian and the decreasing number of monolingual speakers.

Typologically, the Nubian languages are characterized by SOV word order and postpositions. The genitive usually precedes the noun, while the adjective follows it. Genetically, the Nubian language family forms a subgroup of Eastern Sudanic, which in turn is a subgroup of the Nilo-Saharan phylum (Greenberg 1963; Ehret 1989; Bender 1991).

Early contacts between Arabic and Nubian were probably established by Arab nomads and merchants who came from Arabia and entered the Nile Valley long before the coming of Islam (Ḥasan 1973:12). After the Islamic conquest of Egypt in 641 C.E., and throughout the era of the Christian Nubian kingdoms (6th–15th centuries), Arabic as a spoken language was gradually spread by Arabs engaging in trade between Egypt and the Middle Nile Valley. Arabic (along with Old Nubian, Coptic, and Greek) was also used as a written language in medieval Nubia, as attested by gravestones, graffiti, and letters of commerce (Shinnie 1974:46; Adams 1977:447ff.). The few Arabic loanwords in the medieval Old Nubian texts suggest that, despite the long presence of Arabic in the Nile Valley, its influence on Nubian was insignificant. It was only at the end of the Middle Ages, after the successive breakdown of the Christian kingdoms, that Arabization and Islamization gained momentum in Nubia and the Central Sudan (Adams 1977:556–563; Spaulding and Spaulding 1988). The immigration of Arab nomads, mainly from Egypt, and the intermarriage of the Arabs with the local Nubian population resulted in the spread of Arabic as a lingua franca. The intensified linguistic contacts appear to have triggered Nubian/Arabic bilingualism both among the Nubians and those Arab immigrants who became farmers. This can be inferred from the incorporation of Nubian loanwords, mainly terms from agriculture, into Sudanese Arabic (Gasim 1965).

Because of the impact of Arabization, several Nubian languages have disappeared, among them the Old Shaiqi language, spoken south of the fourth cataract until the beginning of the 19th century (Spaulding 1990), the Nubian languages of northern Kordofan, probably spoken until the end of the 19th century (MacMichael 1912:85; Newbold 1924; Bell 1975; Lea 1994:147), and, more recently, Birgid of eastern Darfur (Thelwall 1977; Idris 2004). Recent sociolinguistic studies (Rouchdy 1991; Patriarchi and Rottland 1993; Satti 2004) highlight the various factors that threaten to replace the Nubian languages with Arabic.

The integration of Arabic loanwords into the Nubian languages has not yet been systematically investigated, although most linguistic publications on Nubian languages dedicate some passages to them. Nor does the following discussion exhaust this topic. For example, Arabic stress patterns and their integration into the Nubian phonological system are an interesting but unconsidered problem. The present entry focuses on the Nile Nubian languages, drawing Kenzi data from Massenbach (1933)


N
0
100 km
LIBYA
EGYPT
CHAD
SUDAN
CENTRAL AFRICAN REPUBLIC
ETHIOPIA
Red Sea
Kom Ombo
Aswan
KEN ZI
Korosko
Wadi Halfa
NO BIIN
Kerma
DON GOLAWI
Dongola
Meroe
Ed-Debba
Atbara
Khartoum
HARAZA
Jebel Haraza
Malha
Khashm el-Girba
Blue Nile
Darfur
MIDOB
Kordofan
El-Fasher
El-Obeid
BIRGID
White Nile
Dilling
Nyala
KORDOFAN NUBIAN

Map 1. Geographical distribution of the Nubian languages.

and Kamil (1937), Dongolawi data from Armbruster (1960), and Nobiin data from Werner (1987).

Nobiin is a tone language, distinguishing between a high, a low, and a falling tone. The tones are marked on a vowel by an acute (*á*), grave (*à*), or circumflex (*â*) accent, respectively. In the case of long vowels, which in the Nubian items are represented by doubled vowel signs, only the first vowel receives a tone mark (*áa*). Long vowels in Arabic items, in contrast, are marked by a stroke (*ā*).

The source of Arabic loanwords in the various Nubian languages is usually the local Sudan Arabic dialect, rather than Modern Standard/Classical Arabic. This is attested by the following loanwords in Kenzi, which has adopted *huusa* < Sudan Arabic *xūsa*, rather than Classical Arabic *sikkīn* 'knife'; *angaree* < Sudan Arabic *ʿangarēb*, rather than *sarīr* 'bed'; *bit-ee* < northern Sudan Arabic *beyyat*, rather than *bāt* 'to spend the night'. Further evidence is provided by the following phonological characteristics of Sudan Arabic, which are also attested in loanwords integrated in the Nubian languages: (i) the opposition of Classical Arabic *t*/*ṯ* and *d*/*ḏ* is neutralized in favor of the dental plosives *t*/*d*, respectively; (ii) Classical Arabic *q* is realized as *g* or sometimes even as *ġ*; (iii) the voiced palatal affricate *j* may by replaced by the voiced alveolar plosive *d*, e.g. *deeš* < *jēš* 'army'; (iv) final consonant clusters in syllables of the type CVCC are often simplified by inserting an epenthetic vowel; (v) the short central low vowel *a* is realized as front mid vowel [ɛ] if it occurs in an open syllable and if the vowel of the following syllable is *i* or *ii* (Reichmuth 1983:58); in loanwords [ɛ] is therefore often rendered by *e*, *déris, dérs* < *dars* 'lesson'; *jediid* < *jadīd* 'new'.

Arabic loanwords are phonetically adapted to the Nubian phonological system. Because of the increasing knowledge of and proficiency in speaking Arabic, loanwords may preserve some of the characteristic Arabic segments or structures, thus enlarging and modifying the original Nubian phonological system. The following examples consider only those Sudan Arabic consonants which are foreign to the Nubian language, the data being drawn from Dongolawi. The emphatic consonants *ṭ, ḍ, ṣ, ḏ̣* (*ẓ*) are replaced by their nonemphatic counterparts *t, d, s, z: tabbaah* < *ṭabbāx* 'cook [noun]'; *fuuta* < *fūṭa* 'towel'; *daruuri* < *ḍarūrī* 'necessary'; *haadir* < *ḥāḍir* 'ready'; *seed* < *ṣēd* 'hunting'; *halaas* < *xalāṣ* 'finished'; *zarif* < *ḏ̣arf* 'envelope'; *naazir* < *nāḏ̣ir* 'superintendent'. The voiced alveolar fricative *z* is not an original consonant of the Nubian phoneme system, but, as the examples show, it is admitted in loanwords. The voiced velar fricative *ġ* is replaced by the voiced velar plosive *g*: *garib* < *ġarb* 'west', *gulgul* < *ġulġul* 'cottonseed'. The voiced pharyngeal fricative ʿ is usually deleted: *adu* < *ʿaduww* 'enemy'; *arbaiin* < *ʾarbaʿīn* 'forty'; *šema* < *šamʿ*, *šamaʿ* 'wax'. The voiceless velar and pharyngeal fricatives *x* and *ḥ* are replaced by the voiceless glottal *h*, which is a marginal consonant phoneme in the Nile Nubian languages: *habar* < *xabar* 'news', *washaan* < *wasxān* 'dirty', *haal* < *ḥāl* 'state', *ahsen* < *ʾaḥsan* 'better', *Saaleh* < *Ṣāliḥ* [proper name]. In final position, *ḥ* may be deleted: *faala* < *fāliḥ* 'skillful'. Originally, in the Nubian languages, the lateral *l* and the vibrant *r* were not admitted in initial position. In loanwords, however, they do occur there: *lijaam* < *lijām* 'bridle', *rukun* < *rukn* 'corner'.

The syllable structure in Nubian is CVC, i.e. consonant clusters in word-initial and final position are not admitted. This explains why the CVCC structure of (Sudan) Arabic words is changed to CVCCV or CVCVC when they are integrated into a Nubian language. In Kenzi the structure CVCC*i* is chosen if the Arabic CC cluster consists of geminates: *kummi* < *kumm* 'sleeve', *bunni* < *bunn* 'coffee', *hajji* < *ḥājj* 'pilgrim', *šarri* < *šarr* 'evil'. The structure CVCVC is chosen if the final Arabic CC cluster is represented by different consonants: *darub* < *darb* 'path, road', *fejir* < *fajr* 'dawn', *sahal* < *sahl* 'easy'. The change from Arabic CVCC > CVCVC is also attested in Nobiin. There are often two versions of the loanwords, however, which indicates that the Arabic CC cluster may be retained, although it does not conform to the canonical Nubian syllable structure: *áríš* < *ʿarš* 'roof'; *déris, dérs* < *dars* 'lesson'; *ésím, ésm* < *ism* 'name'; *másír, másr* < *maṣr* 'Egypt; Cairo'.

As for the morphological integration of Arabic loanwords, the data show that Nubian grammatical suffixes are attached to these loanwords. In Nobiin, for instance, consonant-final nouns take the low tone plural suffix *-ìi*, which lowers all tone(s) of the singular form: *báhár* [sg.], *bàhàrìi* [pl.] < *baḥr* 'river; Nile'; *déris* [sg.], *dèrsìi* [pl.] < *dars* 'lesson'; *dùkkáan* [sg.], *dùkkàanìi* [pl.] < *dukkān* 'shop'; *féjír* [sg.], *fèjrìi* [pl.] < *fajr* 'dawn'. Nouns ending in the Arabic feminine suffix *-a* are pluralized by the low-tone suffix *-ncìi*, which lengthens the final *-a* and lowers all preceding tones: *šèbèká* [sg.], *šèbèkàancìi* [pl.] < *šabaka* 'net'; *tàagyá* [sg.], *tàagyàancìi* [pl.] < *ṭāgiyya* 'skull cap'; *mèdrèsá* [sg.], *mèdrèsàancìi* [pl.] < *madrasa* 'school'. Less frequently, a loanword, whether consonant-final or vowel-final, takes the plural suffix *-gúu* or *-ríi*. These suffixes lower the preceding tone as well: *áadèm* [sg.], *àadèmríi* [pl.] < *ʾādamī* 'person'; *jèeb* [sg.], *jèeppúu* (< *jèeb-gúu*) [pl.] < *jīb* 'pocket'; *sèmàa* [sg.], *sèmàagúu* [pl.] 'sky' < *samāʾ*. The Dongolawi postposition *-gi* marks adverbs of time. It is attached to loanwords from Arabic: *baad-ki* < *baʿd* 'after', *kulyoom-gi* < *kull yōm* 'every day', *abadan-gi* < *ʾabadan* 'never'.

In Kenzi-Dongolawi, verbs based on loanwords from Arabic are formed by the auxiliary *-e, -ee* 'to say'. The loanword preceding *-e, -ee* is not necessarily a verb; its main function is to provide the lexical information, whereas *-e, -ee* is inflected, providing the grammatical information. This composite mode of verb formation is well known from the Saharan languages (Crass and Jakobi 2000:21), as well as from several Semitic and Cushitic languages of Ethiopia (Armbruster 1960, Secs. 226–233, 234–237). Examples from Kenzi: *barg-ee* 'to flash' < *barg* 'lightning', *lezm-ee* 'to be necessary' < *lāzim* 'necessary'; examples from Dongolawi: *harb-ee* < *xarab* 'to destroy', *gaabil-ee* < *gābal* 'to meet', *šekkir-ee* < *itšakkar* 'to thank'. The last items attest that the prefix *it-* is deleted when such an Arabic verb stem is borrowed into Dongolawi (Armbruster 1960, Sec. 3652). In Nobiin, however, this auxiliary construction does not exist. The data in Werner's vocabulary (1987:338–383) suggest that a loan verb functions as a base to which the inflection suffix *-ir, -il* is attached (only 1st pers. sg. present is attested). The loan verbs are integrated into one of the three tone classes typical of Nobiin verbs (Werner 1987:141–145): Class 1: low-low; Class 2: (high-) high-low; Class 3: (low-) low-high, realized in pausa as (low-) low-high-low. Examples of this are: *féttíš-ìr* (Class 2) < *fattaš* 'to look for'; *gárb-ìr* (Class 2) < *ġarbal* 'to sift'; *sàfàr-îr* (Class 3) < *sāfar* 'to travel'; *sàll-îr* (Class 3) < *ṣallā* 'to pray'. The loanwords exhibiting the stative extension *-fii* are morphologically treated like a Nubian verb, which is required to appear in the 'a-form' (Werner 1987:167–170): *sàkàn-á-fìir* < *sakan* 'to dwell', *sùum-á-fìil* < *ṣōm* 'fast [noun]'. The tone patterns of these verbs correspond to Class 3.

The preceding sections show that the loanwords come from a wide range of semantic domains. It should be pointed out that a substantial portion of the loanword vocabulary is borrowed from the Islamic domain, such as Dongolawi *jaama* < *jāmiʿ* 'mosque', *hamd-ee* < *ḥamad* 'to praise [God]', *halwa* < *xalwa* '*Qurʾān* school'. The loanwords comprise all word classes, such as nouns and verbs, adjectives, and numerals. To these may be added discourse markers (such as Dongolawi *bes* < *bas* 'only, just', *halaas* < *xalāṣ* 'finally') and conjunctions (*illa* < *ʾillā* 'except that', *laakin* < *lākin* 'but', *wala…wala* < *walā…walā* 'neither…nor'). The loanwords are not restricted to the cultural lexicon, but may exist side by side with or even replace Nubian core vocabulary, such as Dongolawi *eer* and *jediid* < *jadīd* 'new'; *kal* and *akil* < *ʾakl* 'food'; Nobiin *gèm* and *sènà* < *sana* 'year'; *shíráŋ* and *tùrbá* < *turba* 'grave'.

Bibliographical references

Adams, William Y. 1977. *Nubia: Corridor to Africa.* London: A. Lane.

Armbruster, Charles Hubert. 1960. *Dongolese Nubian: A grammar.* Cambridge: Cambridge University Press.

Bell, Herman. 1975. "Documentary evidence on the Ḥarāza Nubian language". *Sudan Notes and Records* 56.1–35.

Bender, M. Lionel 1991. "Sub-classification of Nilo-Saharan". *Proceedings of the Fourth Nilo-Saharan Conference Bayreuth, Aug. 30–Sep. 2, 1989,* ed. M. Lionel Bender, 1–35. Hamburg: H. Buske.

Crass, Joachim and Angelika Jakobi. 2000. "Der Kube-Dialekt des Beria (Zaghawa) im Tschad: Eine erste Skizze". *Afrika und Übersee* 83.3–46.

Gasim, Awn al-Sharif. 1965. "Some aspects of Sudanese colloquial Arabic". *Sudan Notes and Records* 46.40–49.

Greenberg, Joseph H. 1963. *The languages of Africa.* Bloomington: Indiana University; The Hague: Mouton.

Ehret, Christopher. 1989. "Subclassification of Nilo-Saharan: A proposal". *Topics in Nilo-Saharan linguistics*, ed. M. Lionel Bender, 35–49. Hamburg: H. Buske.

Ḥasan, Yūsuf Faḍl. 1973. *The Arabs and the Sudan.* Khartoum: Khartoum University Press.

Idris, Helene Fatima. 2004. "Documentation of the Birgid language". Paper presented to the Ninth Nilo-Saharan Linguistic Colloquium, 15–19 February 2004, University of Khartoum.

Kamil, Murad. 1937. “Arabischer Einfluß auf die nubische Sprache”. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 91.616–629.

Lea, Cyril A.E. 1994. *On trek in Kordofan: The diaries of a British district officer in the Sudan, 1931–1933*, ed. Martin W. Daly. Oxford: Oxford University Press for the British Academy.

MacMichael, Harold. 1912. *The tribes of northern and central Kordofán*. London: F. Cass.

Massenbach, Gertrud von. 1933. “Wörterbuch des nubischen Kunûzi-Dialektes”. *Mitteilungen des Seminars für orientalische Sprachen* 36.99–227.

Newbold, Douglas. 1924. “Some links with the Anag at Gebel Haraza”. *Sudan Notes and Records* 7.126–131.

Patriarchi, Christina and Franz Rottland. 1993. “Sprachkenntnis und Sprachgebrauch in den Nuba-Bergen”. *Die Sudanforschung in der Bundesrepublik Deutschland: Ergebnisse der Bremer Tagung 1993*, ed. Markus Waschkuhn and Karl Wohlmuth, 201–227. Münster: Lit.

Reichmuth, Stefan. 1983. *Der arabische Dialekt der Šukriyya im Ostsudan*. Hildesheim: G. Olms.

Rouchdy, Aleya. 1991. *Nubians and the Nubian language in contemporary Egypt: A case of cultural and linguistic contact*. Leiden: E.J. Brill.

Satti, Nasir. 2004. “The linguistic situation in Dongola”. Paper presented to the Ninth Nilo-Saharan Linguistic Colloquium, 15–19 February 2004, University of Khartoum.

Shinnie, Peter L. 1974. “Multilingualism in medieval Nubia”. *Studies in ancient languages of the Sudan*, ed. Abdelgadir Mahmoud Abdalla, 41–47. Khartoum: Khartoum University Press.

——, and Margaret Shinnie. 1956. “Bilingualism among the Mahass”. *Man* 102–105.100.

Spaulding, J.L. and Jay Spaulding. 1988. “The democratic philosophers of the medieval Sudan: The pig”. *Sprache und Geschichte in Afrika* 9.247–268.

Spaulding, Jay. 1990. “The Old Shaiqi language in historical perspective”. *History in Africa* 17.283–292.

Thelwall, Robin. 1977. “A Birgid vocabulary list and its links with Daju”. *Gedenkschrift Gustav Nachtigal 1874–1974*, ed. Herbert Ganslmayr and Herrmann Jungraithmayr, 197–210. Bremen: Übersee-Museum.

Werner, Roland. 1987. *Grammatik des Nobiin* (Nilnubisch): Phonologie, Tonologie und Morphologie. Hamburg: H. Buske.

Angelika Jakobi (Bayreuth)

Null Subject → Pro-Drop

Number

In Classical and Modern Standard Arabic, nouns, pronouns, adjectives, and verbs are morphologically marked for number. Three categories of number are recognized by both medieval and modern grammarians: singular (*mufrad*), dual (*muṯannā*), and plural (*jamʿ*). The singular in general refers to a single item in a class, the dual to two items of a class, the plural to three or more. An important exception to this generalization is that for a small set of nouns, termed → ‘collectives’, the unmarked singular form refers to a collection or group, and a singulative (*ism al-waḥda* lit. ‘noun of the unit’) referring to a single individual of the group can be derived from the singular-collective by suffixation. Within the category of ‘plural’, the Arab grammarians traditionally draw a distinction between plurals of paucity (*jamʿ al-qilla*), referring to from three to ten items and plurals of multiplicity (*jamʿ al-kaṯra*), referring to more than ten items. However, such a distinction is not productively exploited in Modern Standard Arabic, and to what extent it ever has been is unclear (Ferrando 2006). Number marking is not optional in Arabic. The plural is required in certain syntactic environments, for example after the numbers from three to ten. Inflectional number marking in the verbal conjugation refers exclusively to the number of the subject, although → agreement patterns are complex.

Aside from the singulatives, singular nouns are not overtly marked as such. The dual is marked by the suffixes *-āni* (nominative) and *-ayni* (oblique). For feminine nouns ending in *tāʾ marbūṭa*, the *-t-* is retained before the dual suffix (Table 1).

Table 1. Singular and dual endings

	sg.	du.
nom.	*ṭālib(un)*	*ṭālibāni*
gen./acc.	*ṭālib(in/an)*	*ṭālibayni*
	‘student [masc.]’	
nom.	*ṭāliba(tun)*	*ṭālibatāni*
gen./acc.	*ṭāliba(tin/tan)*	*ṭālibatayni*
	‘student [fem.]’	

Formally, there are two different methods of plural marking in the noun and adjective. These are traditionally termed ‘sound’ (*al-jamʿ as-sālim*) and ‘broken’ (*al-jamʿ al-maksūr* or *al-mukassar*) plurals. The sound plurals, like the duals, are characterized by a suffix (or more precisely by the lengthening of an existing suffix). The broken plurals are characterized by syllable and vowel patterns distinct from their corresponding singulars (or, in processual terms, by stem-internal modification).

The endings of the masculine sound plural (Table 2) are *-ūna* (nominative) and *-īna* (oblique) and of the feminine sound plural *-āt(un)* (nominative) and *-āt(in)* (oblique).

Table 2. Endings of the sound plural

	sg.	pl.
nom.	*mudarris(un)*	*mudarrisūna*
gen./acc.	*mudarris(in/an)*	*mudarrisīna*
	'teacher [masc.]'	
nom.	*mudarrisa(tun)*	*mudarrisāt(un)*
gen./acc.	*mudarrisa(tin/tan)*	*mudarrisāt(in)*
	'teacher [fem.]'	

Acknowledging that the genitive and accusative have collapsed into a single case, this derivation can be most economically analyzed as a lengthening of the vowel after the stem, i.e. the case-marking vowel in the case of the masculine, and the gender-marking vowel in the case of the feminine (Moscati a.o. 1964).

Dual and sound masculine plural formations interact with case and definiteness markings in a complex way. The final *-na* of the sound masculine plural and the final *-ni* of the dual are, etymologically, simply the indefinite marker *-n* followed by an epenthetic vowel required to break up the overlong -Cv̄C- syllable. The distribution of these sequences is slightly different from that of the singular indefinite marker *-n*, termed 'nunation' (→ *tanwīn*). They are deleted in the genitive construct construction (→ *ʾiḍāfa*) before a noun in the genitive, but not when the definite article is attached (Table 3).

The sound plural does not have an unrestricted distribution. It cannot simply be applied freely to any noun as the dual can be. It is restricted to productively derived adjectives and nouns (such as participles) and to some loanwords and to words which cannot be readily analyzed in terms of a → root and a pattern. The sound masculine plural is further restricted to masculine nouns referring to rational beings. Other masculine nouns not subject to broken plural formation take a sound feminine plural (*ḥayawān* 'animal', pl. *ḥayawanāt*).

Virtually all lexical (nonderived) nouns and adjectives having what McCarthy and Prince (1990) term one of the canonical syllable shapes – CvCC, CvCvC, Cv̄CvC, CvCv̄C, CvCCvC, CvCCv̄C – can only form a broken plural. This includes the vast bulk of the basic vocabulary of the language. The sound plural, ultimately the sound feminine plural, can nonetheless be defined as the default, meaning not the most frequent plural but rather the plural of last resort.

The broken plural exemplifies a morphological phenomenon typical of Arabic and Semitic languages: the plural is distinguished from its corresponding singular not by a specific affix but rather by a difference in syllable structure and vocalism. The broken plural system is characterized by a high degree of allomorphy. Allomorphy in this context naturally refers not to a variety of diverse affixes but rather to a variety of syllabic/vocalic patterns associated with the plural function. Older grammars traditionally list approximately thirty plural patterns (depending upon criteria employed for defining a pattern). Wright (1896), for example, lists 32 patterns, with 5 noted as rare. This approach to describing the system is misleading, however, since it gives the false impression that plural patterns are associated with singulars in an entirely arbitrary way. In fact, there are strong patterns of correlation between singular and plural forms, as the medieval grammarians recognized (ʾAbū s-Suʿūd 1971; ʿAbd al-ʿĀl 1977) and as modern statistical studies (Murtonen 1964; Levy 1971) have confirmed. For any given singular class, defined in terms of syllable structure, gender marking, and vocalism, no more than five allomorphs generally account for more than 90 percent of all plurals, and a single allomorph generally accounts for a majority. Phonological factors, such as presence of a glide, and semantic and word class features (adjective vs. noun) also influence the choice of plural. On the basis of these criteria,

Table 3. Endings in the genitive construct construction

	indefinite	with definite article	in *ʾiḍāfa*
sg.	*mudarris(un)*	*al-mudarris(u)*	*mudarrisu l-madrasa(ti)*
du.	*mudarrisāni*	*al-mudarrisāni*	*mudarrisā l-madrasa(ti)*
pl.	*mudarrisūna*	*al-mudarrisūna*	*mudarrisū l-madrasa(ti)*

Table 4. Class 1 plural patterns

		C*u*C*ū*C	ʾ*a*CC*u*C	ʾ*a*CC*ā*C	C*i*C*ā*C	
C*a*CC	*qalb*	*qulūb*				‘heart’
	najm	*nujūm*	*ʾanjum*			‘star’
	šahr	*šuhūr*	*ʾašhur*			‘month’
	kalb				*kilāb*	‘dog’
	bayt	*buyūt*				‘house’
	wazn			*ʾawzān*		‘measure’
C*i*CC	*ṯiql*			*ʾaṯqāl*		‘weight’
	jism	*jusūm*		*ʾajsām*		‘body’
	ḥizb			*ʾaḥzāb*		‘group’
C*u*CC	*ḥukm*			*ʾaḥkām*		‘decision’
	qufl	*qufūl*		*ʾaqfāl*		‘lock’
	ḥulm			*ʾaḥlām*		‘dream’
CvCvC	*qalam*			*ʾaqlām*		‘pen’
	ṣanam			*ʾaṣnām*		‘idol’
	jabal			*ʾajbāl*	*jibāl*	‘mountain

Ratcliffe (1998) groups the singular-plural patterns into seven classes as follows. (Statements about frequency are based on Levy’s statistics and refer to the Modern Standard language.)

Class 1: For underived masculine nouns with the syllabic patterns CvCC and CvCvC, the preferred plural patterns are ʾ*a*CC*ā*C, C*u*C*ū*C, C*i*C*ā*C, and ʾ*a*CC*u*C. C*u*C*ū*C is the most frequent plural for singulars C*a*CC with vowel *a* and no glide in the stem. ʾ*a*CC*ā*C is the most common for nouns with the vocalism C*i*CC, C*u*CC, CvCvC and for C*a*CC singulars containing a glide. Note that here, as elsewhere, some nouns may have more than one accepted plural (Table 4).

For nouns of this class with the pattern Cv̄C (usually analyzed as containing an underlying medial glide as the second root consonant), ʾ*a*CC*ā*C is the most common pattern, but two other patterns, C*i*CC*ān* and C*i*C*a*C*a*t are also frequent, the latter exclusively with CīC singulars, the former with singulars C*ū*C and C*ā*C.

		C*ī*C*ān*	C*i*C*a*C*at*	ʾ*a*CC*ā*C	C*u*C*ū*C	
C*ā*C	*jār*	*jīrān*				‘neighbor’
	nār	*nīrān*				‘fire’
C*ū*C	*ḥūt*	*ḥītān*		*ʾaḥwāt*		‘fish; whale’
C*ī*C	*dīk*		*diyakat*	*ʾadyāk*	*duyūk*	‘rooster’
	fīl		*fiyalat*	*ʾafyāl*	*fuyūl*	‘elephant

However, roughly 50 percent of singulars of the pattern C*ā*C resist analysis into a consonantal root and take a sound feminine plural.

There are six masculine nouns in Arabic that have only two consonants in the singular, having the shape CvC. They all take either the plural ʾ*a*CC*ā*C or C*i*C*ā*C. The interesting peculiarity of these nouns is that in order to fill out the plural pattern, an extra non-root default consonant, usually a glottal stop, is supplied in the plural.

		ʾ*a*CC*ā*C	C*i*C*ā*C	
CvC	*ḥam*	*ʾaḥmāʾ*		‘father-in-law’
	dam		*dimāʾ*	‘blood’

Class 2: For feminine singulars with the pattern C*i*CC*a(tun)*, the usual plural is C*i*C*a*C, and for singular C*u*CC*a(tun)*, the plural is C*u*C*a*C. The only other possible broken plural is C*i*C*ā*C.

		CvC*a*C	
CvCC*a*	*ġurfa*	*ġuraf*	‘room’
	xirqa	*xiraq*	‘rag’
	qīma	*qiyam*	‘value’

Feminine C*a*CC*a(tun)* singulars, with stem-vowel *a*, as well as CvCvC*at*, overwhelmingly prefer the sound feminine plural, although C*i*C*ā*C and rarely C*i*C*a*C and C*u*C*a*C are also possible plurals. Many feminine nouns have a mixed plural with an *a* or a copy of the preceding vowel between the second and third consonants and a sound feminine plural suffix.

		C*i*C*ā*C	C*a*C*a*C*āt*	
CvCvC*a*	*ʿaqaba*	*ʿiqāb*		‘steep road’
C*a*CC*a*	*farxat*	*firāx*	*faraxat*	‘hen’
	jabhat		*jabahāt*	‘face’

Class 3: Broken plural formation is quite regular for four-consonant nouns. For CvCCvC singulars, the only broken plural pattern is C*a*C*ā*C*i*C^{2} (superscript ‘2’ indicating → diptote inflection). For CvCCv̄C singulars, the broken plural is either C*a*C*ā*C*ī*C^{2}, or rarely C*a*C*ā*C*i*C*at*, with a feminine suffix ‘compensating’ for the loss of vowel length in the last stem syllable.

	C*a*C*ā*CiC	C*a*C*ā*C*ī*C	C*a*C*ā*C*i*C*at*	
CvCCVC	*maktab*	*makātib*		‘desk; office’
	ʿaqrab	*ʿaqārib*		‘scorpion’
CvCCv̄C	*sulṭān*	*salāṭīn*		‘sultan’
ʾ	*ustāḏ*		*ʾasātiḏat*	‘professor’

Feminine four-consonant singulars also either take one of these patterns or a sound feminine plural.

		C*a*C*ā*C*i*C	sound fem. pl.	
CvCCvC*a*	*madrasat*	*madāris*		‘school’
	maktabat	*makātib*	*maktabāt*	‘library’

Words of more than four consonants are difficult to integrate into the root-and-pattern system and normally do not take broken plurals. When they do, however, they also conform to the four-consonant C*a*C*ā*C*i*C^{2} plural pattern. This is accomplished by deleting or ignoring one of the singular consonants in the plural.

		C*a*C*ā*C*i*C	
CvCCvCvC	*barnāmij*	*barāmij*	‘program’
	zanbarak	*zanābik*	‘[metal] spring’
CvCCvCv̄C	*ʿankabūt*	*ʿanākib*	‘spider’

Class 4: The greatest degree of allomorphy is found with nouns having three consonants and a long vowel, syllable patterns Cv̄CvC and CvCv̄C. Feminine nouns with these patterns, Cv̄CvC*a(tun)* and CvCv̄C*a(tun)*, however, present little difficulty. Virtually all of them have a plural analogous to that of four-consonant plurals: C*awā*C*i*C^{2} for Cv̄C*v*C*at* singulars and C*a*C*āʾi*C^{2} for CvCv̄C*at* singulars. Nouns of this class which do not have this plural have a sound feminine plural.

		C*a*CāC*i*C	
CvCv̄C*a*	*risāla*	*rasāʾil*	‘letter’
	rakūba	*rakāʾib*	‘mount’
Cv̄CvC*(a)*	*ṣāʿiqa*	*ṣawāʿiq*	‘thunderbolt’
	ḥāmil	*ḥawāmil*	‘pregnant’

The few nouns with the pattern Cv̄Cv̄C, whether feminine or masculine, might also be grouped here, as they also form a plural on the four-consonant pattern C*awā*C*ī*C^{2}.

		C*a*C*ā*C*ī*C	
Cv̄Cv̄C	*jāmūs*	*jawāmīs*	‘buffalo’
	qāmūs	*qawāmīs*	‘dictionary’

Class 5: For masculine nouns and adjectives with the syllable structure Cv̄CvC, the frequent plural forms are C*awā*C*i*C (as in the previous case), C*u*CC*a*C, C*u*CC*ā*C, and C*a*C*a*C*a(tun)*. Almost all singulars of this class are lexicalized active participles with the vocalism C*ā*C*i*C. Singulars with other vowel patterns and some nonrational nouns C*ā*C*i*C have the plural C*awā*C*i*C^{2}:

		C*aw*C*i*C	
Cv̄CvC	*qālab*	*qawālib*	‘mold’
	ʿālam	*ʿawālim*	‘world’
	ʿāmil	*ʿawāmil*	‘factor’

For singulars of the pattern C*ā*C*i*C which retain an adjectival sense, C*u*CC*a*C is a frequent pattern.

		C*u*CC*a*C	
C*ā*C*i*C	*ṣāmix*	*ṣummax*	‘proud’
	sājid	*sujjad*	‘bowing down’

For nouns of the pattern C*ā*C*i*C, referring to human beings, the usual plural patterns are C*u*CC*ā*C and C*a*C*a*C*a(tun)*. For nouns of this class with a glide as the final root consonant, the plural is C*u*C*a*(C)*at*.

		C*u*CC*ā*C	C*a*C*a*C*at*	C*u*C*a*(C)*at*	
C*ā*C*i*C	*ḥākim*	*ḥukkām*			‘judge’
	ṭālib	*ṭullāb*	*ṭalabat*		‘student’
	ḥāmil		*ḥamalat*		‘porter’
	qāḍ(in)			*quḍāt*	‘judge’
	ġāz(in)			*ġuzāt*	‘raider’

Table 5. Class 6 plural patterns

		ʾ*a*CC*i*C*at*	C*u*C*u*C	C*u*CC*ān*	C*a*C*ā*ʾ*i*C	
CvC*ā*C	*janāḥ*	*ʾajniḥat*				‘wing’
	suʾāl	*ʾasʾilat*				‘question’
	silāḥ	*ʾasliḥat*				‘weapon’
	kitāb		*kutub*			‘book’
	bilād			*buldān*		‘country’
CvC*ū*C	*rasūl*		*rusul*			‘prophet’
CvC*ī*C	*qamīṣ*	*ʾaqmiṣat*	*qumuṣ*	*qumṣān*		‘shirt’
	ḍamīr				*ḍamāʾir*	‘pronoun’

Class 6: For masculine nouns and adjectives of the pattern CvCv̄C, the most frequent plural patterns are C*a*C*ā*ʾ*i*C^{2}, C*i*C*ā*C, C*u*C*u*C, ʾ*a*CC*i*C*at*, C*u*CC*ān*, C*u*C*a*C*ā*ʾ^{2}, ʾ*a*CC*i*C*ā*ʾ^{2} (Table 5). The pattern C*i*C*ā*C is restricted to adjectives with the vocalism C*a*C*ī*C, for which it is the most common broken plural pattern.

		C*i*C*ā*C	
C*a*C*ī*C	*ṭawīl*	*ṭiwāl*	‘long’
	kabīr	*kibār*	‘big’
	qaṣīr	*qiṣār*	‘short’

The patterns C*u*C*u*C and ʾ*a*CC*i*C*at* are the most frequent plurals for nonrational singulars with the pattern CvCv̄C. There is a marked skew based on vowel quality: nouns with long *ā* in the second syllable favor plural ʾ*a*CC*i*C*at*, and nouns with *ū* or *ī* favor C*u*C*u*C or more rarely C*u*CC*ān*. The pattern C*a*C*ā*ʾ*i*C^{2} is also possible, as with the feminines.

The most frequent plurals for C*a*C*ī*C nouns referring to human beings are C*u*C*a*C*ā*ʾ^{2} and ʾ*a*CC*i*C*ā*ʾ^{2} (geminate variant ʾ*a*C*i*CC*ā*ʾ^{2}). The latter is preferred for nouns containing a glide or geminate, the first elsewhere.

		ʾuCaCāʾ	*ʾaCCiCāʾ*	*ʾaCiCCāʾ*	
C*a*C*ī*C	*ʾamīr*	*ʾumarāʾ*			‘prince’
	wazīr	*wuzarāʾ*			‘vizier’
	baxīl	*buxalāʾ*			‘miser’
	ġaniyy		*ʾaġniyāʾ*		‘wealthy’
	ṭabīb			*ʾaṭibbāʾ*	‘doctor’

Class 7: Finally, there are a number of adjectives for which both gender and number are marked in an idiosyncratic fashion. Adjectives of color and bodily defect have the patterns ʾ*a*CC*a*C^{2} masculine singular, C*a*CC*ā*ʾ^{2} feminine singular, and C*u*CC or C*u*CC*ān*, common plural.

masc. sg.	fem. sg.	pl.		
ʾ*a*CC*a*C	C*a*CC*ā*ʾ	C*u*CC	C*u*CC*ān*	
ʾaḥmar	*ḥamrāʾ*	*ḥumr*		‘red’
ʾaswad	*sawdāʾ*	*sūd*	*sūdān*	‘black’
ʾaṭraš	*ṭaršāʾ*	*ṭurš*		‘deaf’
ʾaʿraj	*ʿarjāʾ*	*ʿurj*	*ʿurjān*	‘lame’

Adjectives of form C*a*CC*ān*^{2} have a feminine C*a*CC*ā*, and usually a common plural C*a*C*ā*C*ā*, formally a Class 3 four-consonant pattern from the feminine.

masc. sg.	fem. sg.	pl.		
C*a*CC*ān*	C*a*CC*ā*	C*a*C*ā*C*ā*		
ʿajlān	*ʿajlā*	*ʿajālā*	*ʿijāl*	‘quick’
kaslān	*kaslā*	*kasālā*	*kaslā*	‘lazy’

The pattern C*a*CC*ā* appears as a plural for adjectives with a variety of patterns (though principally CaC*ī*C) having a meaning of weakness or injury.

		C*a*CC*ā*	
C*a*C*ī*C	*jarīḥ*	*jarḥā*	‘wounded’
	zamīn	*zamnā*	‘chronically ill’
C*ā*C*i*C	*fāsid*	*fasdā*	‘decayed’

The plurals of the elative (comparative/superlative) might be included here, too, masculine singular ʾ*a*CC*a*C^{2}, plural ʾ*a*C*ā*C*i*C^{2}; feminine singular C*u*CC*ā*, plural C*u*C*a*C.

masc. sg.	masc. pl.	fem. sg.	fem. pl.	
ʾ*a*CC*a*C	C*a*C*ā*C*i*C	C*u*CC*ā*	C*u*C*a*C	
ʾakbar	*ʾakābir*	*kubrā*	*kubar*	‘bigger’, ‘biggest’
ʾaqdam	*ʾaqādim*	*qudmā*	*qudam*	‘older’, ‘oldest’

But the feminine elative plural is essentially identical with the Class 2 feminine plural. It only has to be specified that the feminine suffix (which is deleted in the plural) is in this case *-ā* rather than the usual *-a(tun)*. The

Table 6. The Arabic ‘broken’ plural system

i.	C*a*CC	>>	C*u*C*ū*C, ʾ*a*CC*ā*C, C*i*C*ā*C, [ʾ*a*CC*u*C], (C*ī*C*ān*)
	CvCC	>>	ʾ*a*CC*ā*C, C*u*C*ū*C, (C*i*G*a*C*at*)
	CvCvC	>>	ʾ*a*CC*ā*C
ii.	CvCC*at*	>>	CvC*a*C, CvC*(a)*C*āt*
	C*a*CC*at*	>>	C*a*C*(a)*C*āt*, C*i*C*ā*C
iii.	CvCCvC*(at)*	>>	C*a*C*ā*C*i*C
	CvCCv̄C*(at)*	>>	C*a*C*ā*C*ī*C, CaC*ā*C*i*C*at*
iv.	Cv̄CvC*at*	>>	C*awā*C*i*C
	CvCv̄C*at*	>>	C*a*C*āʾi*C
	Cv̄Cv̄C	>>	C*awā*C*ī*C
v.	*CāCiC* (n.)(-rat.)	>>	C*awā*C*i*C
	C*ā*C*i*C (n.)(+rat.)	>>	C*u*CC*ā*C, C*a*C*a*C*at*, (C*u*C*āt*)
	C*ā*C*i*C (adj.)	>>	C*u*CC*a*C
vi.	CvC*ā*C	>>	ʾ*a*CC*i*C*at*, C*u*C*u*C
	C*a*C*ū*C	>>	C*u*C*u*C, ʾ*a*CC*i*C*at*
	C*a*C*ī*C (n.)(-rat.)	>>	ʾ*a*CC*i*C*at*, C*u*C*u*C
	C*a*C*ī*C (n.)(+rat.)	>>	C*u*C*a*C*āʾ*, ʾ*a*CCiC*āʾ*
	C*a*C*ī*C (adj.)	>>	C*i*C*ā*C, [C*a*CC*ā*]
vii.	ʾ*a*CC*a*C	>>	C*u*CC, C*u*CC*ān*

masculine elative is formed on the Class 3 four-consonant pattern.

The preceding discussion is summarized in Table 6, which indicates all broken (and mixed) plurals representing 10 percent or more of the plurals of each singular. The forms in parentheses ( ) count for less than 10 percent of the class as a whole but more than 10 percent of words containing glide or geminate. The forms in brackets [ ] are statistically rare (less than 10 percent) plural forms strongly associated with a particular class.

This scheme covers more than 90 percent of the broken plural system, but there remain a number of exceptions and anomalies. The pattern C*a*C*ī*C, for example is quite rare as a plural, but it is found with some high-frequency items.

sg.	pl.					
ʿabd	*ʿabīd*	also	*ʿubdān*	*ʿibdān*	*ʿibād*	‘slave’
ḥimār	*ḥamīr*	also	*ḥumur*	*ʾaḥmirat*		‘donkey’

The same can be said for the pattern C*i*CC*at*.

sg.	pl.			
ġazal	*ġizlat*	also	*ġizlān*	‘gazelle’
ʾax	*ʾixwat*	also	*ʾixwān*	‘brother’

Other singulars have plurals not normally associated with words of their syllable class. These too include high-frequency items, e.g. the word *ṣāḥib* with the meaning ‘friend; owner; companion’, which has the plurals *ʾaṣḥāb, ṣaḥb, ṣaḥāba, ṣuḥbān*, and *ṣuḥbat*. Some seemingly anomalous plurals like *ʾaḥādīṯ*, pl. of *ḥadīṯ* ‘talk’ and *ʾaqārib*, pl. of *qarīb* in the sense ‘relative’ can be analyzed as plurals of plurals. (The expected plurals are *ʾaḥdiṯa* and *ʾaqribāʾ*; with the initial glottal stop incorporated, these become four-consonant stems, hence the plural ʾ*a*C*ā*C*i*C.) In addition, there are a few cases of complete suppletion: *imraʾa* ‘woman’, pl. *nisāʾ* ‘women’.

As is clear from the examples above, some words may have more than one accepted plural. Occasionally, the possibility of multiple plurals is exploited to make distinctions of meaning. For homonyms or polysemous words, different plurals may be used for the different singular senses. From singular *bayt* we find *buyūt* ‘houses’ and *ʾabyāt* ‘verses’; from *ʿayn* we find *ʿuyūn* and *ʾaʿyun* ‘eyes’ or ‘springs’, *ʾaʿyān* ‘notable; important person’ and *ʿaynāt* ‘ayns [letter of the alphabet]’. According to the grammarians, the patterns *ʾafʿul, ʾafʿāl, ʾafʿilat*, and *fiʿlat* are plurals of paucity, and the rest are plurals of multiplicity. As noted above, this distinction is not exploited in the Modern Standard language, although there may be some historical basis for it. Fischer (1980), developing a proposal first made by Ferguson (1959), suggests that the *ʾa-* prefix in most of these forms may have originated as a junctural feature after the numbers from three to ten.

Outside of the broken plural system proper, nouns of profession of the pattern C*a*CC*ā*C, relational nouns with the suffix *-iyy*, and some others occasionally form a plural by addition of the feminine *tāʾ marbūṭa* suffix, in lieu of the more usual sound masculine plural.

sg.	pl.		
xabbāz	*xabbāzūna*	*xabbāza*	'baker'
fallāḥ	*fallāḥūna*	*fallāḥa*	'farmer'

Collective nouns capable of deriving a singulative belong to a semantically restricted set and refer to plants, animals, and products derived from them. The marker of the singulative is the polyfunctional *tāʾ marbūṭa* termination characteristic of the feminine gender. Singulatives are construed as feminine singular, while the corresponding singular collective is masculine.

singular collective		singulative	
baqar	'cattle'	*baqara*	'a cow'
tuffāḥ	'apples'	*tuffāḥa*	'an apple'

From the singulative, duals and plurals can be derived.

baqaratāni	'two cows'	*baqarāt*	'several cows'
tuffāḥatāni	'two apples'	*tuffāḥāt*	'several apples'

A somewhat similar phenomenon is found with names of tribes and ethnic groups, which also refer to a group or collectivity. The attachment of the relational adjective suffix (→ *nisba*) yields a singular adjective, which can be construed as a noun.

ʿarab	'Arabs'	*ʿarabiyy*	'Arabic; an Arab'
turk	'Turks'	*turkiyy*	'Turkish; a Turk'

Technically, however, this is not primarily a number-based derivation but rather a noun-to-adjective derivation.

Within the context of the 'orthodox' theory of root-and-pattern morphology (represented most explicitly according to Bohas 1993 by Cantineau 1950a, 1950b), the broken plurals present something of a problem. If one thinks of root-and-pattern morphology as a rigorous matrix, where lexical meaning is expressed by the consonantal root and grammatical functions by syllabic vocalic patterns, it is paradoxical to have a great variety of patterns for a single function. However, when one considers that plurals are not plurals of an abstract consonantal root but rather are plurals of specific singular nouns or adjectives, and that under any given consonantal root there are usually several distinct words distinguished from each other by differences in syllable structure and vocalism, the functionality, indeed the necessity for having a variety of plural patterns becomes apparent. It is convenient to be able to distinguish *ʾanwār* 'lights' (pl. of *nūr*) from *nīrān* 'fires' (pl. of *nār*), and *ʾaḥkām* 'legal judgments' (pl. of *ḥukm*) from *ḥikam* 'wise sayings' (pl. of *ḥikma*), and *ḥukamāʾ* 'wise men' (pl. of *ḥakīm*) from *ḥukkām* 'rulers' (pl. of *ḥākim*); and to distinguish 'riders' *rukkāb* or *rukbān* (pl. of *rākib*) from the beasts they ride upon *rakāʾib* (pl. of *rakūba*) and *marākīb* (pl. of *markūb*), as well as from *rukab* 'knees' (pl. of *rukba*) and *marākib* 'boats' (pl. of *markab*).

The fact that plurals of different singular patterns are generally distinct, combined with the fact that in many cases features of the singular are maintained in the plural (the first vowel in the feminines, vowel length in the four-consonant stems), raises the issue of whether there might be derivational processes relating words in this system, and if so how to formalize them. McCarthy and Prince (1990) argue that the so-called 'productive' plurals (the most statistically prominent forms in Classes 1, 3, 4 in the schema above) are formed by essentially the same rule, which brackets off an initial heavy syllable as a locus for the intercalation of a long *-ā-*.

CvCC	<CvC>*ā*C	(going to *ʾa*CC*ā*C by later rule)
Cv̄CvC	<Cv̄>*ā*CvC	(yielding C*awā*C*i*C)
CvCv̄C	<CvC>*ā*v̄C	(yielding C*a*C*āʾi*C)
CvCCvC	<CvC>*ā*CvC	
CvCCv̄C	<CvC>*ā*Cv̄C	

The closest thing to verbal number as such in Arabic is perhaps to be found in the system of derived stems or *ʾawzān*. The intensive sense of the so-called Form II verbs (pattern C*a*C-C*a*C*a*), like *kassara* 'to break to pieces, shatter', has been analyzed as plurality of action or of object (Greenberg 1991). However, the most common use of these stems is to indicate valence, that is, to add or subtract from the number of arguments a verb can take. The Forms conventionally labeled II, III, and IV, add an argument to the Form I verb (thus making intransitives transitive and transitives causative, or in the case of Form III, raising a peripheral actor to the role of a direct object). The Forms labeled VIII, V, VI, and X, generally function as the → middle or intransitive of Forms I, II, III, and IV, respectively, and thus can be said to subtract an argument. There are notable formal parallels between this system and the system of nominal number marking. Two of the argument-adding patterns, C*a*C-C*a*C*a* II and C*ā*C*a*C*a* III, are marked by a length increase over the Form I C*a*CvC verb and a characteristic *-a-* vowel, as are the most frequent broken plural patterns (Benmamoun 2003). The argument-subtracting verbs *i*C*ta*-C*a*C*a* VIII, *ta*C*a*CC*a*C*a* V, *ta*C*ā*C*a*C*a* VI, and *ista*CC*a*C*a* X are each marked by a prefix or infix *-t*, in the same way the singulative, which represents a decrease in number, is marked by the *-t-* of the *tāʾ marbūṭa* suffix.

Conventionally, verbal number is understood as the verbal inflection indicating the number of the subject. There is no object agreement, although there are object clitics. Both dual and plural are indicated by inflectional affixes in Classical Arabic, although the dual paradigm is defective, marking only the 3rd person of both genders and common 2nd person. The rules of → agreement established for Modern Standard Arabic are as follows. Verbs must agree in number with their subject only when the subject is a noun referring to a rational being, and only when the subject precedes the verb. When the subject follows, only gender agreement is obligatory for rational nouns. For the plurals of nouns with nonrational referents, the verb is feminine singular, regardless of the gender of the noun. Adjectives agreeing with nonrational plural nouns, and pronouns referring to them, should also be feminine singular. In fact, this pattern of agreement represents a comparatively late development in written Arabic. The oldest Arabic texts (pre-Islamic up to the 10th century C.E.) often show plural verb agreement and frequently show plural adjective agreement with inanimate plural nouns (Belnap and Shabaneh 1992; Beeston 1975).

While the sound plural and dual uncontroversially derive from Proto-Semitic, the question of the historical origin of the broken plural system has generated an extensive debate stretching back to the 19th century. The older viewpoint (developed by Barth 1904 and Kuryłowicz 1962, 1973), reiterated in most handbooks of comparative Semitic (e.g. Moscati a.o. 1964), is that the patterns of the broken plural were Proto-Semitic deverbal abstract noun patterns which gradually shifted into the role of the plural. The motivation for this hypothesis is found in the fact that patterns like C*u*C*ū*C and C*i*C*ā*C appear both as plurals and as verbal nouns (*jihād*, *kitāb*, *julūs*). The principal difficulty is that the most productive plural patterns (*ʾa*CC*ā*C, CvC*a*C, C*a*C*ā*C*i*C, C*u*CC*ā*l, *ʾa*CC*i*C*at*, C*u*C*a*C*āʾ*, etc.) do not appear in the role of verbal nouns or any other nonplural function in Arabic or any other Semitic language. Greenberg's (1955) discovery that languages distantly related to Arabic in the Berber, Chadic, and Cushitic subfamilies of Afro-Asiatic have plurals formed by infixing an *a*/*ā* vowel between the second and third consonant of the stem threw the older consensus into question (although Greenberg did not explicitly draw a connection with the Arabic broken plural). Taking this evidence into account as well as newer evidence from the South Arabian and Ethiopian Semitic languages with similar plural systems, Ratcliffe (1992, 1998) proposes that the core of the broken plural system has developed from a Proto-Afro-Asiatic process of plural formation by *-ā-* infixation, through a set of sound changes and analogies.

In broad outline, the system of number described above remains intact in the modern spoken Arabic dialects. Case distinctions are lost in all dialects, and only the oblique forms of the sound masculine plural and dual are retained, generally as *-īn* and *-ēn*, respectively (→ pseudo-dual). The productivity of the dual decreases as one goes from east to west. In Moroccan Arabic and neighboring Maghreb dialects, it is restricted to a few words of time and measurement. In the dialects of Syria and Lebanon it is actively employed (Ferguson 1959; Blanc 1970). Some of the less frequent broken plural patterns are lost in the dialects. Some new plural forms are created by --> analogy (Ratcliffe 2002, 2003). Loanwords are freely adapted into the system of broken plurals (Holes 1995:139–140). In contrast to the written standard, all the dialects, to varying degrees, allow the possibility of plural verb and adjective agreement with plural nouns regardless of word order or referent type (Brustad 2000). Dual marking in the verb and pronoun are lost.

Bibliographical references

ʿAbd al-ʿĀl, ʿAbd al-Munʿim Sayyid. 1977. *Jumūʿ at-taṣḥīḥ wa-t-taksīr fī l-luġa al-ʿarabiyya*. Cairo: Maktabat al-Xānjī.

ʾAbū s-Suʿūd, ʿAbbās. 1971. *al-Fayṣal fī ʾalwān al-jumūʿ*. Cairo: Dār al-Maʿārif.

Barth, Jakob. 1904. "Beiträge zur Pluralbildung des Semitischen". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 58.431–446.

Beeston, Alfred F.L. 1975. "Some features of Modern Standard Arabic". *Journal of Semitic Studies* 20.62–68.

Belnap, Kirk and Osama Shabaneh. 1992. "Variable agreement and nonhuman plurals in Classical and Modern Standard Arabic". *Perspectives on Arabic linguistics*, IV, ed. Ellen Broselow, Mushira Eid, and John McCarthy, 245–262. Amsterdam and Philadelphia: J. Benjamins.

Benmamoun, Elabbas. 2003. "Reciprocals as plurals in Arabic". *Research in Afroasiatic grammar*, II, ed. Jacqueline Lecarme, 53–62. Amsterdam and Philadelphia: J. Benjamins.

Blanc, Haim. 1970. "Dual and pseudo-dual in the Arabic dialects". *Language* 46.42–57.

Bohas, Georges. 1993. "Diverses conceptions de la morphologie arabe". *Développements récents en linguistique arabe et sémitique*, ed. Georges Bohas, 45–56. Damascus: Institut Français de Damas.

Brustad, Kristen E. 2000. *The syntax of spoken Arabic: A comparative study of Moroccan, Egyptian, Syrian, and Kuwaiti dialects*. Washington, D.C.: Georgetown University Press.

Cantineau, Jean. 1950a. "Racines et schèmes". *Mélanges William Marçais*, 119–124. Paris: G.P. Maisonneuve.

——. 1950b. "La notion de schème et son altération dans diverses langues sémitiques". *Semitica* 3.73–83.

Ferguson, Charles A. 1959. "The Arabic koine". *Language* 35.616–630.

Ferrando, Ignacio. 2006. "The plural of paucity in Arabic and its actual scope: On two claims by Siibawayhi and Al-Farraa". *Perspectives on Arabic linguistics*, XVI, ed. Sami Boudelaa, 39–62. Amsterdam and Philadelphia: J. Benjamins.

Fischer, Wolfdietrich. 1980. "Die arabische Pluralbildung". *Zeitschrift für Arabische Linguistik* 5.70–88.

Greenberg, Joseph. 1955. "Internal *a*-plurals in Afroasiatic". *Afrikanistische Studien Diedrich Westermann zum 80. Geburtstag gewidmet*, ed. Johannes Lukas, 198–204. Berlin: Akademie Verlag.

——. 1991. "The Semitic intensive as verbal plurality". *Semitic studies in honor of Wolf Leslau*, ed. Alan S. Kaye, 577–587. Wiesbaden: O. Harrassowitz.

Holes, Clive. 1995. *Modern Arabic: Structures, functions and varieties*. London: Longman.

Kuryłowicz, Jerzy. 1962. *L'apophonie en sémitique*. The Hague: Mouton.

——. 1973. *Studies in Semitic grammar and metrics*. London: Curzon Press.

Levy, Mary M. 1971. *The plural of the noun in Modern Standard Arabic*. Ph.D. diss., University of Michigan, Ann Arbor.

McCarthy, John J. and Alan Prince. 1990. "Foot and word in prosodic morphology: The Arabic broken plural". *Natural Language and Linguistic Theory* 8.209–283.

Moscati, Sabatino, Anton Spitaler, Edward Ullendorf, and Wolfram von Soden. 1964. *An introduction to the comparative grammar of the Semitic languages*. Wiesbaden: O. Harrassowitz.

Murtonen, Aimo. 1964. *Broken plurals: The origin and development of the system*. Leiden: E.J. Brill.

Ratcliffe, Robert R. 1992. *The broken plural problem in Arabic, Semitic and Afroasiatic: A solution based on the diachronic application of prosodic analysis*. Ph.D. diss., Yale University.

——. 1998. *The 'broken' plural problem in Arabic and comparative Semitic: Allomorphy and analogy in non-concatenative morphology*. Amsterdam and Philadelphia: J. Benjamins.

——. 2002. "The broken plural system of Moroccan Arabic: Diachronic and cognitive perspectives". *Perspectives on Arabic linguistics*, XIII–XIV, ed. Dilworth Parkinson and Elabbas Benmamoun, 87–110. Amsterdam and Philadelphia: J. Benjamins.

——. 2003. "The historical dynamics of the Arabic plural system: Implications for the theory of morphology". *Research in Afroasiatic grammar*. II. *Selected papers from the Fifth Conference on Afroasiatic Languages, Paris, 2000*, ed. Jacqueline Lecarme, 339–362. Amsterdam and Philadelphia: J. Benjamins.

Wright, William. 1896. *A grammar of the Arabic language*. 3rd ed. London: Cambridge University Press.

Robert R. Ratcliffe
(Tokyo University of Foreign Studies)

Numerals

Numerals represent a very particular semantic field of the lexicon. They concern a very small number of roots because, at least originally, all numbers, apart from zero (named *ṣifr* in Classical Arabic), are formed by a combination of a limited series of twelve numbers (1–10, 100, and 1,000). These roots mainly belong to a common semantic background. In addition to the cardinals, the ordinals, and the fractions, they have produced a certain number of verbal and nominal lexemes in ancient Arabic.

1. Cardinals

Cardinals have several usages that correspond to different syntactic functions, in which they belong to different linguistic categories, such as noun, pronoun, and adjective. In Arabic, the situation is particularly complex because these categories can change, not only with the syntactic function, but also according to the definiteness of the noun and the appropriate numeral, with significant differences from one variety of Arabic to another. A detailed presentation of the first ten cardinals in modern dialects is found in Mörth (1997).

1.1 *Numbers used in isolation*

Numerals are used alone in order to name the numbers or to count in the abstract, or in reference to names. As pronouns, they often have complements.

1.1.1 ‘One’

In Classical Arabic, the number ‘one’ is expressed by *ʾaḥad*, fem. *ʾiḥdā* or by *wāḥid*, fem. *wāḥida*. In modern dialects, the most frequently used forms to count are *wāḥid/wāḥad*, fem. *wāḥ(i)da/waḥ(a)da/wiḥda*. Like those of *wāḥid*, the dialectal equivalents of Classical Arabic *ʾaḥad* (*aḥḥad/ḥadän/ḥad(d)* etc.) are often used as indefinite article (→ article, indefinite). It is the same for their negative counterpart *ma-ḥadd(iš)*.

1.1.2 ‘Two’

In Classical Arabic, ‘two’ is *iṯnān/iṯnayn* in counting. When it is used as a noun or pronoun, its feminine equivalent is *iṯnatān/iṯnatayn* (more rarely *ṯintāni*). In the main Maghrebi dialects, its dialectal equivalent is *zawj* ‘pair’ (under various forms, such as *žūž/zūz/žus*), which is used for ‘two’. In the Bedouin Maghrebi dialects and the Eastern Arabic dialects, reflexes of *iṯnān* are used, e.g. *ṯ(i)nayn/t(i)nēn/hintēn/sənīn*, etc. In the feminine *ṯantayn*, *ṯ* and *ay* are realized with a similar degree of variation, and sometimes (in Chad and especially in Yemen) assimilation of *nt* > *tt* takes place, as in *ṯittē(n)*. In certain Eastern dialects, the isolated form ends in *-a(t)*, probably under the influence of the numerals from three to ten, e.g. (Mekka) *itnēna*.

In Classical Arabic, a specific form *kilā-*, fem. *kiltā-* is employed with pronominal suffixes. In dialects, pronominal suffixes are annexed to the common form (or to a lengthened form in *-āt*): (Syria) *tnēnət-/tnēnāt-*.

1.1.3 From ‘three’ to ‘ten’

In Classical Arabic, forms in *-a(t)* are used in counting: *ṯalāṯa* ‘three’, *ʾarbaʿa* ‘four’, *xamsa* ‘five’, *sitta* ‘six’, *sabʿa* ‘seven’, *ṯamāniya* ‘eight’, *tisʿa* ‘nine’, *ʿašara* ‘ten’. As pronouns, these forms in *-a(t)* are used to refer to masculine nouns. For feminine nouns, forms devoid of the final *-a* are used: *ṯalāṯ*, *ʾarbaʿ*, *xams*, *sitt*, *sabʿ*, *ṯamānin*, *tisʿ*, *ʿašr*, e.g. *ar-rijālu ṯalāṯatun* ‘the men are three’, *an-nisāʾu ṯalāṯun* ‘the women are three’. This behavior, peculiar to numbers from three to ten, which seems to reverse the usual gender agreement (*-a(t)* for the masculine instead of the feminine), can be found in other Semitic languages and remains hard to understand. For certain Arabic grammarians, it is the isolated form (in *-a*) which is prior, and this would explain its use for the masculine.

Most of the dialects only have forms in *-a(t)* and use them both for counting and in reference to nominals, both feminine and masculine. Bloch's hypothesis (1971) about the prevalence of the polysyllabic forms in isolation is discussed by Cowan (1972). The following points may be noted:

i. Frequent final *ʾimāla* (*-a* > *-i*)
ii. Elision of the *l* of ‘three’ in certain *qəltu* dialects of Mesopotamia: Mardin *ṯāṯe*
iii. *r* > *ġ* and V*ġ* > V̄: in ‘four’: Mossul *ōbʿa*, ʿAqra (Jewish) *ābʿa*
iv. Diminutive form for ‘six’: Takroûna *stētä*
v. Elision of /ʿ/ in ‘seven’ and ‘eight’: Chad *saba* and *tise*

Euphemistic forms are attested for certain numbers, sometimes becoming the only forms in use, e.g. *ʿodd yəddək* (lit. ‘count [the fingers of] your hand’) for ‘five’ among the Jewish women of Tunis. This is particularly frequent with ‘nine’, where it is done in order to avoid saying *tisʿa*, which can also mean ‘you will beg’, e.g. *təsʿūd* (lit. ‘you'll be happy’) in Morocco or *təsʿad* in Lebanon.

Cardinals from three to ten regularly admit, as pronouns, a pronominal complement through suffixation. The masculine form ends in *-(a)t*: Classical Arabic *aṯ-ṯalāṯa* ‘the three [masc.]’ > *ṯalāṯat-kum* ‘your three’, (Damascus) *tlātətna* ‘the three of us’.

1.1.4 From ‘eleven’ to ‘nineteen’

In Classical Arabic, the cardinals from eleven to nineteen are compound nouns, ending in *-a*, and indeclinable (with the exception of ‘twelve’). In counting and in referring to masculine nouns, ‘ten’ is always in the masculine form;

in referring to feminine nouns, the ten has the feminine form. The units follow the same rule as for the first ten cardinals: agreeing in gender for eleven and twelve, polarity of gender for thirteen through nineteen. The element 'ten' has a particular vocalization in these cardinals (masc. *ʿašara*, fem. *ʿašrata*). The resulting numerals are as follows: masc. *ʾaḥada ʿašara*, fem. *ʾiḥdā ʿašrata* 'eleven'; masc. *ʾiṯnā ʿašara*, fem. *ʾiṯnatā/ṯintā ʿašrata* 'twelve'; masc. *ṯalāṯata ʿašara*, fem. *ṯalāṯa ʿašrata* 'thirteen'; masc. *ʾarbaʿata ʿašara*, fem. *ʾarbaʿa ʿašrata* 'fourteen'; masc. *xamsata ʿašara*, fem. *xamsa ʿašrata* 'fifteen'; masc. *sittata ʿašara*, fem. *sitta ʿašrata* 'sixteen'; masc. *sabʿata ʿašara*, fem. *sabʿa ʿašrata* 'seventeen'; masc. *ṯamāniyata ʿašara*, fem. *ṯamāniya ʿašrata* 'eighteen'; masc. *tisʿata ʿašara*, fem. *tisʿa ʿašrata* 'nineteen'.

In Arabic dialects, the gender distinction has disappeared almost everywhere. The form corresponding to the Classical Arabic one used for counting and for referring to masculine nouns is at the basis of the dialectal numerals, so for thirteen through nineteen, the units are in the feminine form (ending in *-t*), and the element 'ten' is in the masculine form. The cardinals eleven through nineteen have undergone frequent abbreviations and contractions, of which the following may be noted:

i. The phoneme /ʿ/ in the units has often elided, e.g. in 'seven', (Damascus) *sabaṭaʿš*, and the same applies to the /ʿ/ of the ten, which is often compensated by the lengthening of the vowel, (Jewish Tunis) *ṭnāš* 'twelve', (Cairo) *sabaʿṭāšar* 'seventeen'.
ii. In most dialects (except in the Arabian Peninsula, in Uzbekistan, and in the non-Bedouin dialects of the Chadian-Sudanese zone), the *-t-* of the feminine ending in the cardinals thirteen through nineteen has become emphatic. This emphasis is often extended to the precedent consonants, and by analogy to 'twelve' or even to 'eleven': (Iraq) *ṣiṭṭaʿaš* 'sixteen', *ṭnaʿaš* 'twelve'.
iii. The last consonant, *-(a)r*, of the element 'ten' has been elided in many dialects (Arabian Gulf) *sittāš* 'sixteen'; it was maintained only in Egypt, in Chad-Sudan, and in part of the Arabian Peninsula: Oman, Dubai, and Qatar.

Certain peripheral dialects of Africa do not use compound nouns: (Nigeria) *ʾašaṛa (ɦaw) xamsa* 'fifteen' (lit. 'ten + five').

1.1.5 Cardinal tens

There is no gender distinction for the tens from twenty through ninety. In Classical Arabic, they are derived from the units through the ending of the masculine plural form (*-ūna* in the nominative, *-īna* in the genitive/accusative), including 'twenty', for which a dual would have been expected (and a first vowel *-a-*): *ʿišrūna* 'twenty', *ṯalāṯūna* 'thirty', *ʾarbaʿūna* 'forty', *xamsūna* 'fifty', *sittūna* 'sixty', *sabʿūna* 'seventy', *ṯamānūna* 'eighty', *tisʿūna* 'ninety'.

In dialects, the nouns of the tens correspond to the oblique case of the Classical Arabic forms and are in *-īn/-in*.

1.1.6 Beyond 'ninety-nine'

There is no distinction of gender for the hundreds, thousands, and millions. The dual is regularly used for 'two hundred' and 'two thousand', even where the dual has become rare: (Morocco) *myatayn/mitin* 'two hundred'.

'Hundred' is *miʾa* (pl. *miʾūn/miʾāt*) in Classical Arabic, *miyya/mīya/myā* in dialects. For three hundred to nine hundred, *miʾa* is normally preceded by the unit noun in the feminine (without *-at*), but exceptionally it remains singular in Classical Arabic and in most dialects (but not in Hadramawt). In Classical Arabic (with *miʾa* in genitive): *ṯalāṯu miʾatin* 'three hundred', *ʾarbaʿu miʾatin* 'four hundred', *xamsu miʾatin* 'five hundred', *sittu miʾatin* 'six hundred', *sabʿu miʾatin* 'seven hundred', *ṯamānī miʾatin* 'eight hundred', *tisʿu miʾatin* 'nine hundred'.

In some dialects, the form of the unit noun is slightly different (less contracted and closer to the one in Classical Arabic) in front of the hundred than in front of the other nouns, thus (Damascus) *tlāt* (**tlǝtt*) *miyye* 'three hundred', *tmān* (**tmǝn*) *miyye* 'eight hundred'.

'Thousand' is *ʾalf* (pl. *ʾālāf*) in Classical Arabic; dialects generally have closely related forms. For three thousand to nine thousand, *ʾalf* is normally preceded by a unit noun in the long form (with *-at*) and itself becomes plural. In Classical Arabic (with *ʾālāf* in genitive): *ṯalāṯatu ʾālāfin* 'three thousand', *ʾarbaʿatu ʾālāfin* 'four thousand', etc.

The feminine ending *-(a)t* was exceptionally kept in front of *ālāf/alāf* in a number of dialects, especially Eastern ones, but also in Kormakiti and in Ḥassāniyya. These dialects do not follow gender polarity anymore, hence, after reinterpretation, a new plural beginning with *t-* appeared, e.g. (Iraq) **xamist ālāf* > *xamis tālāf* 'five thousand'. In the Maghrebi dialects, on the contrary, it is usually the form without *-t* that is used, e.g. (Morocco) *xems alāf* 'five thousand'.

'Million' is *malyūn* (pl. *malyūnāt/mālāyīn*) and 'billion' is *milyār* or *balyūn*.

1.1.7 Complex numbers

In numbers from 21 to 99 the units precede the tens, and both of them, connected by *wa-*, are declined. Numerals indicating numbers including thousand, hundred, tens, or units, can be constructed in two ways. They are arranged either in a growing order or in a descending order, except for the units that precede the tens. The elements are normally connected by *wa-*. In composition, for 'two' *ṯnayn* is used, even in those dialects that have the dialectal equivalent of 'pair' for the first decade.

1.2 *Numbers used in counting nouns*

When they are accompanied by the noun of the counted thing, numerals can still be in the independent form, but more frequently, especially if the numeral precedes the nominal and above all in the absence of a definite article, they are in a state of annexion.

1.2.1 'One'

To the question 'how much?', the answer in Classical Arabic is a singular noun (*kitāb* 'a book'), the adjectival use of *wāḥid* 'only, unique' being emphatic (*yawm wāḥid* 'a unique day', *al-kitāb al-wāḥid* 'the only book').

In some modern dialects, *wāḥid* (maybe also *ʾaḥad*) has produced an indefinite article (→ article, indefinite). This evolution is probably the result of the pronominal construction of *wāḥid* directly constructed (without *min* 'of'). *ʾaḥad* is often used in dialects only for 'Sunday, first day of the week' (Classical Arabic *(yawm) al-ʾaḥad*).

1.2.2 'Two'

To the question 'how much?', the answer in Classical Arabic is a dual (*kitāb-ā-ni* 'two books'). The use of the cardinal, agreeing in gender, is possible and has an emphatic meaning (*kitāb-ā-ni ṯnāni* 'two books', with the definite article, *al-kitāb-ā-ni l-iṯnāni* 'the two books').

The more conservative dialects (especially the Bedouin dialects) use the dual with all nominals, whether definite or not: (Aleppo) *ššahṛēn* 'the two months'. Some use the dual only with indeterminate nouns and the cardinal adjective in all other cases: (Ḥassāniyya) *ṛāžläyn* 'two men' but *əṛ-ṛažžālä əṯ-ṯäntäyn* 'the two men'. Other, especially Moroccan, urban dialects use the dual only for a more limited series of nouns (in particular for measure units of time, space, quantity; → number).

If it is the cardinal *zawj* (or one of its variants) that is used, it is normally in a construct state. This construction (with the meaning of 'the pair of') is found in dialects that have preserved *ṯnayn*.

1.2.3 From 'three' to 'ten'

With indefinite nouns, the numerals three through ten govern the *pluralis paucitatis* and, in Classical Arabic, the genitive. The Classical construction is that of the construct state, with the cardinal as a noun in the position of determined word. The fact that in Classical Arabic the form in *-at* is used with masculine nouns and the form without *-at* with feminine nouns, was interpreted by some Europeans as a rule of gender crossing, intended to emphasize the substantival rather than adjectival nature of the cardinals (cf. Wright 1898: Sec. 319, Rem. *a*). But many other hypotheses have been proposed to explain this 'gender polarity', which originally may have been more connected with the notion of plurality than with gender (cf. in particular Rundgren 1968 and Brugnatelli 1982).

Gender differentiation has been maintained in some dialects of the Arabian Peninsula (Ṣanʿāʾ *xamsah biyūt* 'five houses', *xams nisēʾ* 'five women'), but it seems that in some dialects, it is disappearing. In other dialects, the short form was generalized to the detriment of the masculine form in *-(a)t*. However, the latter has left traces in several masculine nouns, in particular in those whose plural begins with a vowel, following the example of *(t-)ālāf* 'thousands': *xamstiyām* 'five days', *xamsteš̌hur* 'five months' (already written *xams tiyām* and *xams tišhur* in Middle Arabic).

The form of the cardinals used in the construct state has generally been subjected to a certain contraction (particularly the *ā* of 'three' and 'eight'), so that in those dialects in which the state of annexion remains the usual construction (essentially Eastern and/or Bedouin dialects), this series of numerals is clearly different from the first series used to count (series in isolation or 'free').

This system, which seems to confirm the hypothesis of a koine (Ferguson 1959:624–625), represents, however, only one of the solutions chosen in the dialects (Cohen 1970; Kaye 1976:173–174; Taine-Cheikh 1994).

On the one hand, there is a tendency to use the free form in *-a* with the counted nouns, either directly in a direct construction as first or second noun, e.g. (Chad-Sudan) *tamāne bagar/bagar tamāne* 'eight cows' or (with a preposition) in an analytic construction (Morocco) *tlāta d-lə-ktūb* 'three books'. The abandonment of the state of annexion for an analytical construction (with the exception of a small number of nouns frequently used with cardinals) is very frequent in the Maghreb, especially in the western part. Sometimes the use of the free form can also be linked to certain morphological or semantic characteristics of the counted noun (loanword without plural, adjective with an internal plural designating persons, nouns of measurement, orderings, etc.; cf. Woidich 1989), e.g. (Cairo) *ʿašara g(e)rām* 'ten grams', *xamsa ʿumy* 'five blind people', *talāta dahab* 'three golden ones', *talāta ʾahwa* 'three coffees!'.

On the other hand, there is a tendency to use both variants (with or without *-a/t*) according to the initial of the second word. Thus, the ending *-(a)t* is used before a vowel, e.g. in Kfar ʿAbīda, instead of the short form (*ʾarbaʿt-enfos* 'four persons') and in Sudan, instead of the free form ending in *-a*. Conversely, the ending *-a* appears before a double consonant (Gabès) or before a simple consonant (Jewish Tunis) *šəbʿä* (but without the *-a* before the vowel in *šəbʿ ulēd* 'seven children').

With definite nouns, the regular construction in Classical Arabic is the apposition of the noun to the cardinal (defined by the article *al-*), with the same gender polarity as in the construction with the indefinite noun, e.g. *aš-šarikātu l-xamsu* 'the five companies'. Constructions with an annexion are, however, frequent. Generally, the article precedes the counted noun, e.g. *ṯalāṯatu l-kutubi* 'three books'. Two other constructions are more or less tolerated (particularly in written Modern Standard Arabic): *aṯ-ṯalāṯatu kutubin* and *aṯ-ṯalāṯatu l-kutubi.*

A certain diversity of constructions can also be found in modern dialects. In the case of the apposition (the most frequent construction in Classical Arabic, often in decline in the dialects), the free form in *-a* is used when the gender polarity has disappeared, e.g. (Ḥassāniyya) *l-əʿläyyāt əs-sittä* 'the six women'. In the constructions with a construct state, frequent in the Middle East, the article generally precedes the numeral, e.g. (Iraq) *l-xams isnīn* 'the five years', sometimes with a repetition (Ṣanʿāʾ) *as-sitt al-banāt* 'the six girls', a less frequent variant of *al-banāt as-sitt.* On the other hand, the 'normal' form of the construct state (with an article before the counted noun) does not seem to be used.

1.2.4 From 'eleven' to 'nineteen'

In Arabic, all cardinals above ten govern the singular (unit noun for the collectives). In Classical Arabic, the counted noun is in the accusative as a 'specifying complement' (→ *tamyīz*), always indefinite (in *-n*), *xamsūna dirhaman* 'fifty dirhams'.

With indefinite counted nouns, gender agreement in Literary Arabic is marked for the units and the tens, but always with the same gender polarity as for three through nine. Dialects with gender variation are rare. Some modern dialects use the same forms in isolation and with counted nouns. The *-ar* forms are attested in particular in Egypt, Sudan, Chad, and Dubai, e.g. (Cairo) *tamanṭāšar daraga* 'eighteen steps', and the forms without *-ar* are attested in Mesopotamia and in part of the Arabian Peninsula, e.g. (Ṣanʿāʾ) *iṯnʿāš bint* 'twelve girls'.

Other dialects (in the Middle East, e.g. Mecca, Damascus, Bahrain; in the Maghreb, e.g. Djidjelli, Zaër, Ḥassāniyya) have kept the final consonant *r/ṛ* in the series used in annexion, while in isolation truncated forms (without *-ar*) occur, (Gulf Arabic) *sittaʿšar ḥmīsa* 'sixteen turtles'. A variant of this system is found in some sedentary Maghrebi dialects in which the enlarged form is not in *-r* (except when assimilated to the following consonant), but in *-ən*, (Cherchell) *xəmṣṭāšən yōm* 'fifteen days', or in *-əl*, (Tlemcen) *xməṣṭāšəl mra* 'fifteen women', or it may even be assimilated to the first radical, just like the *l* of the definite article.

1.2.5 From 'twenty' to 'ninety-nine'

In Classical Arabic agreement in gender is only marked in the unit, but always with the gender polarity mentioned for three through nine. In the dialects, there is no difference between the form used in isolation and in compounding. The constructions appear to be always synthetic, even in the Maghrebi dialects, e.g. (Morocco) *xemsa u ʿešrin xeṭma* 'twenty-five steps', *tesʿud u tesʿin meṛṛa* 'ninety-nine times'.

In the cardinals 11–99, definiteness of the counted noun as *tamyīz* is expressed in Classical Arabic by the article in the numeral, e.g. *aṯ-ṯalāṯata ʿašara rajulan* 'the thirteen men'. It is this type of construction that is generally used in the dialects, e.g. (Gulf Arabic) *li-ḥdaʿšar biʿīr* 'the eleven camels', *l-ʿišrīn gubguba* 'the twenty crabs'. Sometimes, however, the article is repeated before the counted noun, e.g. (Ṣanʿāʾ) *al-ʾarbaʿīn yawm*/(less frequent) *al-ʾarbaʿīn al-yawm* 'the forty days'.

Sometimes, the dialects only use the 'adjectival' construction, normally reserved for the numerals three through ten, while the form used for eleven through nineteen is the independent one, e.g. (Ḥassāniyya) *l-lyāli l-axməsṭaʿš* 'the fifteen nights', *l-lyāli l-xamsīn* 'the fifty nights'.

1.2.6 After 'one hundred' and 'one thousand'

In Classical Arabic, the counted noun is in the genitive singular. In Arabic dialects, the numeral 'one hundred' is different from the others because it has the form associated with feminine nouns in the construct state, e.g. (Morocco) *myat neʿla* 'a hundred curses'.

In Classical Arabic, definiteness with 'one hundred' and 'one thousand' is the same as with 'three' through 'ten', the only syntactic difference lying in the use of the singular for the counted noun.

In dialects, the construction is generally the one used for 11–99, mostly with the numeral following the article, e.g. (Ṣanʿāʾ) *al-miyat fāris* 'the hundred horsemen'.

2. Ordinals

In Classical Arabic, 'first', as in many other languages, is isolated because of its radical and its form. It is an elative, whose root is *ʾ-w-l*: masc. *ʾawwal*, pl. *ʾawwalūn*, fem. *ʾūlā*, pl. *ʾuwal*.

The ordinal numerals from 'two' to 'ten' have the form *fāʿil* and behave like adjectives (with a feminine form in *-a(t)*): *ṯāni*, *ṯāliṯ*, *rābiʿ*, *xāmis*, *sādis*, *sābiʿ*, *ṯāmin*, *tāsiʿ*, *ʿāšir*.

Numerals from 'eleven' to 'nineteen' are invariable and always have the accusative ending; the tens stay in *-a*. Only the units take the *fāʿil* form (for 'eleven', the unit is *ḥādiya*, fem. *ḥādiyata*), but the two parts agree in gender. If the noun is definite, the article is only present once; the ordinals 11–19 function as compound nouns.

From 'twenty' onward the cardinals are used as ordinals, without gender modification, except concerning the units. With a definite noun, the article is repeated in front of each element.

3. Fractions

In Classical Arabic, 'half' is *niṣf*, and the fractions from three to ten are expressed through nouns of the pattern *fuʿl* or (more rarely) *fuʿul*: *ṯul(u)ṯ* 'one third', *rub(u)ʿ* 'one quarter'.

Bibliographical references

Bloch, Ariel. 1971. "Morphological doublets in Arabic dialects". *Journal of Semitic Studies* 16:1.53–73.

Brugnatelli, Vermondo. 1982. *Questioni di morfologia e sintassi dei numerali cardinali semitici*. Florence: La Nuova Italia Editrice.

Cohen, David. 1970. "'Koinè', langues communes et dialectes arabes". David Cohen, *Etudes de linguistique sémitique et arabe*, 105–125. The Hague and Paris: Mouton.

Cowan, William. 1972. "The historical syntax of the Arabic numbers". *Glossa* 6.131–146.

Ferguson, Charles A. 1959. "The Arabic koinè". *Language* 35.616–630.

Kaye, Alan S. 1976. *Chadian and Sudanese Arabic in the light of comparative Arabic dialectology*. The Hague and Paris: Mouton.

Mörth, Karlheinz. 1997. *Die Kardinalzählen von Eins bis Zehn in den neuarabischen Dialekten*. Vienna: W U V – Universitätsverlag.

Rundgren, Frithiof. 1968. "Die Konstruktion der arabischen Kardinalzahlen: Zur historischen Würdigung der komplementären Distribution". *Orientalia Suecana* 17.107–119.

Taine-Cheikh, Catherine. 1994. "Les numéraux cardinaux de 3 à 10 dans les dialectes arabes". *Actes des Premières Journées Internationales de*

Dialectologie Arabe de Paris, ed. Martine Vanhove and Dominique Caubet, 251–266. Paris: INALCO.

Woidich, Manfred. 1989. "Langform versus Kurzform: Die Kardinalzahlwörter von 3 bis 10 im Kairenischen". *Jerusalem Studies in Arabic and Islam* 12.199–232.

Wright, William. 1898. *A grammar of the Arabic language*, II. Cambridge: Cambridge University Press.

Catherine Taine-Cheikh
(CNRS, LACITO – UMR 7017)

Nunation

In Classical Arabic, the grammatical term 'nunation' (from Arabic → *tanwīn*) is defined as the usually unvocalized suffix *-n*, which is pronounced but not written at the end of nouns. At first, nunation was graphically represented by double-colored dots (Endress 1982:179; Lipiński 1997:163), and later by doubling the appropriate vowel signs. The masculine accusative ending *-an*, which sounded in pause *-ā* was written with an *ʾalif* as lengthening marker in combination with the double vowel marker. As a marker of indefiniteness (*tankīr*), nunation is applied to triptotic nouns without the definite article *al-*, many proper names (Fleisch 1961:270–273), and some adverbs. Nouns with only two endings, diptotic nouns (→ diptosis), not fully inflected nouns, and proper names when followed by the word *ibn* or *ibna* in a genealogical series (Wright 1896:I, 249B–250A; A. Fischer 1931:106–108) do not fall under this rule.

The genesis and derivation of the nunation, especially in connection with the so-called → mimation in other Semitic languages, e.g. Akkadian and South-Arabian, was disputed for a long time. In the Semitic context, the best summary for this topic from a comparative and diachronic point of view is given by Diem (1975:239–258), to which a few references may be added (Carter 1981:75, Par. 3.87; Edzard 1998:121–122; Kienast 2001:143–144, Par. 139; Lipiński 1997:272–273; Stempel 1999:92).

From the point of view of modern Western grammars, e.g. Wright (1896:I, 234D), Reckendorf (1898, Par. 87–88; 1921, Par. 110), Krahl a.o. (1999:72), and W. Fischer (2006, Par. 11–12), nunation in Classical Arabic primarily serves as a marker between definite and indefinite nouns.

By comparison, the Arab grammarians distinguish up to five, sometimes even ten important functions of the nunation (cf. Howell 1883–1911:II–III, 699–703, Chap. 23). The first and most important distinction is that between complete, i.e. triptotic (*munṣarif*), and incomplete, i.e. diptotic (*ġayr munṣarif*), declension. Called *at-tanwīn ad-dāll ʿalā l-makāna*, 'the nunation which shows that a noun is fully declinable', this *tanwīn at-tamkīn* or *tanwīn at-tamakkun* 'nunation of establishment' is found in the singular and in the broken plural, whether defined, e.g. *zaydun* 'Zayd', or undefined, e.g. *rajulun* 'man' and *rijālun* 'men' (Howell 1883–1911: II–III, 699; Wright 1896:I, 235B n.; Carter 1981:16–17; Versteegh 1995:171–172; Ayoub 1991:154–155; Kouloughli 2001:21). Next is the *tanwīn al-muqābala* 'nunation of correspondence', which refers to the plural feminine, e.g. *muslimātun* 'female Muslims' (Howell 1883–1911:II–III, 700; Wright 1896:I, 235B n.; Fleischer 1885:I/1, 325; Carter 1981:18–19; Ayoub 1991:153; Kouloughli 2001:22). The third type is the *tanwīn at-tankīr* 'the nunation which distinguishes [in the case of an indeclinable noun] between the definite (*maʿrifa*) and the indefinite (*nakira*)', for which examples are given by Howell (1883–1911:II–III, 699–700), Wright (1896:I, 235C n.), Carter (1981:18–19), and Versteegh (1995:172; cf. also Ayoub 1991:153–154; Kouloughli 2001:21). The fourth function is called *tanwīn al-ʿiwaḍ* 'nunation of compensation', which is divided further into three, clearly more marginal, subcategories, including the elision at the end of the plural of some words with a third weak radical, e.g. *jawārin* 'girls' (Howell 1883–1911:II–III, 700; Wright 1896:I, 235C–D n.; Carter 1981:18–19; Versteegh 1995:172; Ayoub 1991:154; Kouloughli 2001:21). Apart from these kinds of nunation, which are peculiar to the noun, some Arab grammarians added another six, connected with → poetic licenses and more or less artificial (see e.g. Howell 1883–1911:II–III, 701–703 Chap. 23, and the instructive example of two verses of an anonymous writer, quoted by aš-Širbīnī, in Carter 1981:20–21).

Of special usage and connected with poetic recitation is the *tanwīn at-tarannum* 'nunation of trilling'. Attached to nouns, verbs, and particles, this nasalization of the rhyming vowel occurred especially in the eastern dialects of the Banū Tamīm (Howell 1883–1911:II–III, 701; Birkeland 1940:10–18; Rabin 1951:36–37; Fleisch 1961:192–193; Guillaume 2000:193) and of the Banū Qays (Wright 1896:II, 369C remark). As a phenomenon of phonetics at the sentence level, not of morphology (Rabin 1951:36), the pausal vowel was written *-an* (ـَـن) instead of *-an* (ـًا) (Fleischer 1885:I/1, 323–325).

Although the case endings and with them the nunation have disappeared in all modern Arabic dialects, nunation still functions as an indefinite marker in some Bedouin dialects of the Bedouin in the Tihāma in Yemen (Versteegh 1997:149), among other things in the annexion of attributive complements, e.g. *bint ʿāmm-in lih* 'a niece of his' (W. Fischer 1982:84).

Bibliographical references

Ayoub, Georgine. 1991. "La nominalité du nom ou la question du *tanwīn*". *Arabica* 38.151–213.

Birkeland, Harris. 1940. *Altarabische Pausalformen.* Oslo: J. Dybwad.

Carter, Michael G. 1981. *Arab linguistics: An introductory classical text with translation and notes.* Amsterdam: J. Benjamins.

Diem, Werner. 1975. "Gedanken zur Frage der Mimation und Nunation in den semitischen Sprachen". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 125.239–258.

Edzard, Lutz. 1998. *Polygenesis, convergence, and entropy: An alternative model of linguistic evolution applied to Semitic languages.* Wiesbaden: O. Harrassowitz.

Endress, Gerhard. 1982. "Herkunft und Entwicklung der arabischen Schrift". *Grundriß der arabischen Philologie.* I. *Sprachwissenschaft*, ed. Wolfdietrich Fischer, 165–183, 190–197. Wiesbaden: L. Reichert.

Fischer, August. 1931. "Grammatische arabische Miszellen. I. Stück". *Islamica* 4.94–108.

Fischer, Wolfdietrich 1982. "Frühe Zeugnisse des Neuarabischen". *Grundriß der arabischen Philologie.* I. *Sprachwissenschaft*, ed. Wolfdietrich Fischer, 83–95. Wiesbaden: L. Reichert.

——. 2006. *Grammatik des klassischen Arabisch.* 4th rev. ed. Wiesbaden: O. Harrassowitz.

Fleisch, Henri. 1961. *Traité de philologie arabe.* I. *Préliminaires, phonétique, morphologie nominale.* Beirut: Imprimerie Catholique.

Fleischer, Heinrich Leberecht. 1885. *Kleinere Schriften, gesammelt, durchgesehen und vermehrt*, I/1. Leipzig: Hirzel.

Gätje, Helmut. 1970. "Zum Begriff der Determination und Indetermination im Arabischen". *Arabica* 17.225–251.

Guillaume, Jean-Patrick. 2000. "*Tanwīn*". *Encyclopaedia of Islam* X, 193. 2nd ed. Leiden: E.J. Brill.

Howell, Mortimer Sloper. 1883–1911. *A grammar of the Classical Arabic language, translated and compiled from the most approved native or naturalized authors.* II, III. *The verb and the particles.* Allahabad. (Repr., Delhi: Gian Publishing House, 1986.)

Kienast, Burkhart. 2001. *Historische Sprachwissenschaft mit Beiträgen von Erhart Gräfe (Altägyptisch) und Gene B. Gragg (Kuschitisch).* Wiesbaden: O. Harrassowitz.

Kouloughli, Djamel Eddine. 2001. "Sur le statut linguistique du *tanwīn*: Contribution à l'étude du système déterminatif de l'arabe". *Arabica* 48.20–50.

Krahl, Günther, Wolfgang Reuschel, and Eckehard Schulz. 1999. *Lehrbuch des modernen Arabisch.* 2nd ed. (with Monem Jumaili). Leipzig: Langenscheidt.

Lipiński, Edward. 1997. *Semitic languages: Outline of a comparative grammar.* Leuven: Peeters.

Rabin, Chaim. 1951. *Ancient West-Arabian.* London: Taylor's Foreign Press.

Reckendorf, Hermann. 1898. *Die syntaktischen Verhältnisse des Arabischen.* Leiden: E.J. Brill.

——. 1921. *Arabische Syntax.* Heidelberg: C. Winter.

Stempel, Reinhard. 1999. *Abriß einer historischen Grammatik der semitischen Sprachen.* Frankfurt a. Main: P. Lang.

Tegnér, Esaias. 1865. *De nunnatione arabica.* Lund: Berlingianis.

Versteegh, Kees. 1995. *The explanation of linguistic causes: az-Zağğāğī's theory of grammar, introduction, translation, commentary.* Amsterdam and Philadelphia: J. Benjamins.

——. 1997. *The Arabic language.* Edinburgh: Edinburgh University Press.

Wright, William. 1896. *A grammar of the Arabic language.* 2 vols. 3rd ed. rev. William Robertson Smith and Michael Jan de Goeje. Cambridge: Cambridge University Press.

Franz-Christoph Muth (University of Munich)

O

Object → Transitivity: Object

Object, Absolute

1. Definition

The absolute object (*al-mafʿūl al-muṭlaq*) is defined in the Arabic syntactic tradition as "an accusative noun phrase that takes the form of its *maṣdar* (*nomina verbi* or infinitives) or its substitute; it is used to emphasize the action of its governor (the verb or its substitutes), its kind or number" (ar-Rājiḥī 1988:277). The term *mafʿūl muṭlaq* is not used by Sībawayhi; he discusses the function of the *maṣdar* in the construction of the absolute object under the term *ism al-ḥadaṯān* or *al-ḥadaṯ* 'the event' (*Kitāb* I, 15.2–3; cf. I, 117–120 *Bāb mā yakūnu min al-maṣādir mafʿūlan* 'Chapter on those verbal nouns that are object'):

> Know that the verb that does not go beyond the agent [i.e. the intransitive verb] goes beyond it to the noun of the event [i.e. has the verbal noun as object], which is derived from it, because it is only mentioned in order to indicate the event (*wa-ʿlam ʾanna l-fiʿl allaḏī lā yataʿaddā l-fāʿil yataʿaddā ʾilā ism al-ḥadaṯān allaḏī ʾuxiḏat minhu li-ʾannahu ʾinnamā ḏukira li-yadulla ʿalā l-ḥadaṯ*)

The term *mafʿūl muṭlaq* appears for the first time in a grammatical treatise from the 3rd/9th century, Ibn as-Sarrāj's *Kitāb al-ʾuṣūl* (I, 190.15). For the theories of the Arabic grammarians about the *mafʿūl muṭlaq*, see az-Zamaxšarī (*Mufaṣṣal* 16–18; cf. Levin 1991).

The canonical form of the absolute object is the → *maṣdar* of the verb, which normally precedes it in the sentence. Wright (1964:53) states:

> All verbs, whether transitive or intransitive, active or passive, may take their own abstract nouns (*nomina verbi* or infinitives) as also the deverbal nouns of the classes *nomina vicis* and *nomina speciei* as objective complements in the accusative. This may be the case either when they have no other objective complement or complements, or when they have one or more; and the verbal noun may either stand alone, or it may be connected with an adjective or demonstrative pronoun, a noun or pronoun in the genitive, or a descriptive or relative clause.

The absolute object is sometimes called 'absolute accusative' (Cantarino 1975:II, 170). It is called 'absolute' because it is considered to be the only true object among the other kinds of objects in Arabic, i.e. *al-mafʿūl bihi* 'the direct object' (→ *mafʿūl*), *al-mafʿūl maʿahu* 'the comitative object', *al-mafʿūl li-ʾajlihi* 'the object of result', and → *al-mafʿūl fīhi* 'the adverbial object designating either place or time'.

(1) *šarib-nā* *l-bāriḥat-a*
drank-we the-yesterday-Acc
wa-xālid-an *finjān-a*
and-Khalid-Acc cup-Acc
qahwat-in ***šurb-an*** ***sarīʿan***
coffee-Gen drinking-Acc quick-Acc
'We drank a cup of coffee very quickly with Khalid yesterday'

In (1) there are four accusative noun phrases: *al-bāriḥata, xālidan, finjāna*, and *šurban sarīʿan*. According to traditional Arabic syntactic analysis, only *šurban* is the true object because it is the only object affected by the agent/subject 'we'. The other objects, i.e. *al-bāriḥata* (temporal object, designating the time in which the action was done), *xālidan* (the comitative object, with whom the drinking was done), *finjāna* (an object which we did not make ourselves but with which the action of drinking was performed), are objects not affected by the agent/subject of the sentence.

For the absolute object construction in Classical Arabic, see Wright (1964:53–57) and Reckendorf (1921:79–83); for the absolute object in Modern Standard Arabic, see Badawi a.o. (2004:145–149, 451–453). The *mafʿūl muṭlaq* in the *Qurʾān* is studied by Talmon (1999). In contemporary Arabic dialects, the construction with an absolute object, as an instance of → paronomasia, is not unknown, for instance in Egyptian Arabic, as in (2),

(2) *nāyim* *fi* *l-ʿasal* *nōm*
sleep.Partic in the-honey sleep
'He is sleeping soundly' (Woidich 2006:269)

or in Syrian Arabic, as in (3):

(3) *ṣādaf-t-o* *mṣādafe*
encounter-1s.Past-him encounter
'I ran across him by chance' (Cowell 1964:442)

2. Semantic functions of the absolute object

The main semantic functions of the absolute object are the following:

i. The absolute object is used to emphasize the verb, as in (4).

(4) *ḥaṭṭam-a* *ʾax-ī*
broke brother-Nom-my
luʿbata-hu ***taḥṭīm-an***
toy-Acc-his breaking-Acc
'My brother destroyed his toy completely'

The emphasis may be indicated by the repetition of the absolute object, as in (5).

(5) *zayd-un* ***sayr-an*** ***sayr-an***
Zayd-Nom walking-Acc walking Acc
'Zayd is walking continuously'

ii. If the absolute object is a *nomen vicis* (a noun that expresses the performing of an action once), it is used for enumeration, as in (6).

(6) *ḍarab-a* *ʾaḥmad-u*
hit-Past Ahmad-Nom
l-kurat-a ***ḍarbat-ayni***
the-ball-Acc hitting-du.Acc
'Ahmad hit the ball twice'

iii. If the absolute object is a 'noun of kind', it is used to indicate the manner of doing what is expressed by the verb, as in (7).

(7) *jalas-a* *zayd-un*
sat-Past Zayd-Nom
julūs-a ***al-qurfusāʾi***
sitting-Acc the-squatting-Gen
'Zayd sat squatting [i.e. with thighs against the stomach and arms enfolding the legs]'

A *maṣdar* as absolute object may also serve as a further specification of an earlier *maṣdar*, as in (8).

(8) *li-zayd-in* *ṣawt-un*
to-Zayd-Gen voice-Nom
ṣawt-a *l-bulbul-i*
voice-Acc the nightingale
'Zayd has a voice like a nightingale's'

It may also specify the speaker's attitude toward the meaning of the sentence, as in (9),

(9) *ḥaqqu-ka* *ʿalay-ya* ***iʿtirāf-an***
right-your on-me confession-Acc
'I owe this to you, I admit' (instead of *ḥaqqu-ka ʿalayya, ʾaʿtarifu*)

or to disambiguate or assert the truth of the statement, as in (10).

(10) *zayd-un* *ʾax-ī* ***ḥaqq- an***
Zayd-Nom brother-my truth-Acc
'I confirm that Zayd is my true [blood] brother'

In some cases, the absolute object serves to specify an earlier statement, as in (11).

(11) *sa-tadxul-ūna* *l-imtiḥān-a*
will-sit-2mp the-exam-Acc
fa-ʾimmā ***najāḥ-an***
then-either passing-Acc

wa-ʾimmā	***rusūb-an***
and-either	failing-Acc

‘You will sit for the exam and then you will either succeed or fail’

3. The absolute object as replacement of the verb

The absolute object can be used as a substitute for its verb, in which case the verb is deleted and the absolute object is left, as in (12a), for which the full form is (12b).

(12a)	***sayr-an***	*ʾilā*
	walking-Acc	to
	l-ʾamām-i	
	the-fore-Gen	

‘[You,] advance forward!’

(12b)	*sīr-ū*	***sayr-an***	*ʾilā*
	walk-Imper.2pl	walking-Acc	to
	l-ʾamāmi		
	the-fore-Gen		

There are some semantic conditions specifying when the *maṣdar* form of the verb, i.e. the absolute object, can be used as a substitute for its verb. This is possible in imperatives as in (12a) or in (13),

(13)	***rujūʿ-an***	*ʾilā*	*l-warāʾ-i*
	returning-Acc	to	the-back-Gen

‘[You,] go backward!’

or in statements expressing prohibition, as in (14),

(14)	***ṣabr-an***	*lā*	*jazāʿ-an*
	patience-Acc	not	being.bewildered-Acc

‘Be [you] patient, not bewildered’

or in statements expressing a prayer to God, a good wish, or a request, as in (15),

(15)	***buʿd-an***	*li-ḏ̣-ḏ̣ālim-īna*
	distance-Acc	for-the-oppressors-Gen

‘May God destroy the oppressors!’

or in statements expressing scolding, astonishment, or complaint, as in (16)–(18).

(16)	***ʾa-tahāwun-an***	*wa-qad*
	Q-negligence-Acc	and-already
	sabaqa-ka	*rifāq-u-ka*
	surpassed-you	friends-Nom-your

‘Are you still negligent, now that your friends have surpassed you?’

(17)	***ʾa-šawq-an***	*wa-lam*	*yamḍi*
	Q-yearning-Acc	and-not	passed
	ʿalā	*safar-ī*	
	on	traveling-my	
	siwā	*šahrin*	
	except	month-Gen	

‘Do I yearn for my homeland, having been away for one month only?’

(18)	*ʾa-**faqr-an***	*wa-**ḏ̣ulm-an***
	Q-poorness-Acc	and-oppression-Acc

‘Is it fair that I become poor and oppressed [on top of everything else]?’

Sentences of this last type are preceded by a question particle (in this case *ʾa*) and have interrogative intonation.

A special case is that of a group of words that occur exclusively as absolute objects, such as *labbayka* ‘here I am, at your service’, *saʿdayka* ‘good luck to you!’, *hanānayka* ‘how unfortunate for you!’, *dawālayka* ‘by turns, alternatively’, *maʿāḏa (llāh)* ‘God forbid!’, and *subḥāna (llāh)* ‘God be praised!’. These are analyzed as absolute objects by the Arabic grammarians. According to them, the element *-ay-* in these words, except the last two, is a dual ending. The duality is interpreted to mean repetition of meaning and not duality of the morphology; *labbayka* is then assumed to mean *ʾana muqīm ʿalā ṭāʿatika ʾiqāmatan baʿda ʾiqāmatin* ‘I am here at your service one time after another/ I am here at your service all the time’ (Ibn Manẓur, *Lisān* III, 214). Another interpretation of this element is that it is a similar vocalic element as in the preposition *ʿalā*, which becomes before suffixes *ʿalay-(ka)* (Fleisch 1979:348).

4. Form of the *maṣdar*

The *maṣdar* as absolute object is derived from the verb whose object it is. It may, however, be replaced by a *maṣdar* from another derived Form of that verb, e.g. *nabāt* instead of *ʾinbāt* from the verb *ʾanbata* ‘to plant’, or *kalām* instead of *takallum* from the verb *takallam* ‘to talk’, or of *taklīm* from *kallama* ‘to address’,as in (19), or *tabtīl* instead of *tabattul* from the verb *tabattala* ‘to devote oneself’, as in (20).

(19) *kallama-nī* *zayd-un*
talked-me Zayd-Nom
kalām-an *mufīd-an*
talking-Acc useful-Acc
'Zayd did indeed make sense to me'

(20) *wa-tabattal* *ʾilayh-i*
and-devote.yourself-Imper to-him
tabtīl-an (Q. 73/8)
devoting-Acc
'And devote yourself to Him completely/exclusively'

Instead of the *maṣdar* of a particular finite verb, that of another verb with the same meaning is sometimes employed, as in (21).

(21) *jalast-u* ***quʿūd-an***
sat-I sitting-Acc
'I took my seat' (instead of *julūsan*)

The *maṣdar* may be replaced by the instrument by which the action of the absolute object is performed, as in (22).

(22) *ḍarab-tu-hu* ***sawṭ-an***
beat-I-him whip-Acc
'I beat him with a whip' (instead of *ḍarabtuhu bi-sawṭin*)

The *maṣdar* may also be replaced by a word that indicates how the action was performed, as in (23).

(23) *jalast-u* ***l-qurfusāʿa***
sat-I the-squatting-Acc
'I sat squatting' (instead of *jalastu julūsa l-qurfusāʾi*)

The *maṣdar* may also be replaced by a word that denotes how many times it was done, as in (24).

(24) *zur-tu-ka* ***marrat-ayni***
visited-I-you time-du-Acc
'I visited you twice'

It may also be replaced by the interrogative *mā* as in (25), and with the conditional/relative *mā/mahmā* as in (26).

(25) **mā** *ʾakram-ta* *zayd-an*
what honor-2ms.Past Zayd-Acc
'How did you honor Zayd?'

(26) ***mā/mahmā*** *tajlis* *ʾajlis*
whatever you-sit I-sit
'I will sit the way you sit' (instead of *ʾajlisu kulla julūsin tajlisuhu*)

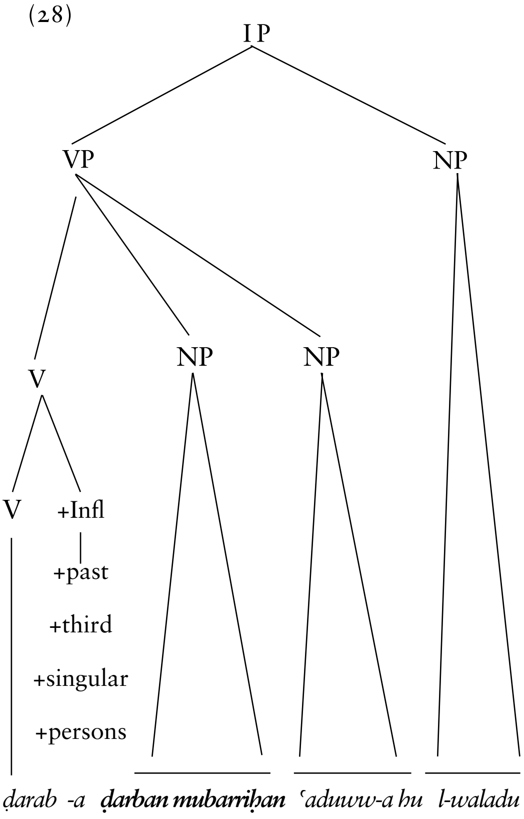
5. GOVERNMENT OF THE ABSOLUTE OBJECT

The typical absolute object in Arabic is always an accusative NP, and its typical governor is the verb or one of its substitutes, i.e. which can assign an accusative case, as in (27).

(27) *ḍaraba* *l-walad-u* *ʿaduww-a-hu*
hit3ms.Past the boy-Nom enemy-Acc-his
ḍarb-an ***mubarriḥan***
hitting-Acc violent-Acc
'The boy hit his enemy violently'

The verb *ḍaraba* is the governor, and the word in italics is the absolute object. Diagrammatically, the tree structure of (27) could be represented as in (28), (assuming that Arabic is a configurational language with a VP initially; see Homeidi 1987, 1994).

(28)



Other accusative case assigners to the absolute object in Arabic include the active participle (*ism al-fāʿil*) and the passive participle (*ism al-mafʿūl*), which are nominal forms derived from the same lexical root and which have the same governance as a verb, as in (29).

(29) *ʾinna l-qāriʾ-a kutub-a-hu*
indeed reading-Acc books-Acc-his
qirāʾatan muʿammaqat-an
reader-Acc deepened-Acc
sawfa yanjaḥ-u
Fut succeed-Ind
'The one who reads his books thoroughly will succeed'

The governor is the active participle *al-qāriʾ*, and the absolute object *qirāʾatan muʿammaqatan* is governed by it. Another example in which the passive participle is the governor and case assigner of the absolute object is (30).

(30) *hāḏā r-rajul-u*
this-Nom the-man-Nom
muḥtaram-un ***iḥtirām-an***
respected-Nom respect-Acc
šadīd-an *fī qawm-i-hi*
enormous-Acc in people-Gen-his
'This man is greatly respected among his people'

6. POSITION OF THE ABSOLUTE OBJECT

In canonical structures, the absolute object should come after its governor in the structure – although not necessarily immediately after it, especially when it emphasizes the action of the verb, as in (31).

(31) *daras-a ṭ-ṭullāb-u*
study.3ms.Past the-students-Nom
l-kitāb-a ***dirāsat-an***
the-book-Acc studying-Acc
wāfiyat-an
comprehensive-Acc
'The students studied the book comprehensively/very carefully'

However, the absolute object should be moved to the beginning of the sentence when it is substituted by a question or a conditional particle/word because these words occur sentence-initially in Arabic, as in (32) and (33).

(32) ***ʿayyat-a kitābat-in***
what-Acc writing-Gen
taktubu
you-write
'What a wonderful kind of writing are you writing?

(33) ***ʿayy-a julūs-in*** *tajlisu ʾajlisu*
what-Acc sitting-Gen you-sit I-sit
'I sit the way you sit'

In other cases, where the absolute object is used to clarify the way the action is performed, or when it is used for enumeration of the action, preposing the absolute object is optional, as in (34a, b) and (35a, b).

(34a) *rajaʿ-tu* ***l-qahqarā***
retired-I backward-Acc
'I retreated backward'

(34b) ***al-qahqarā*** *rajaʿ-tu*

(35a) *sāfar-tu* ***ʿišr-īna marrat-an***
traveled-I twenty-Acc time-Acc
'I traveled twenty times'

(35b) ***ʿišr-īna marrat-an*** *sāfar-tu*

7. MODIFIERS OF THE ABSOLUTE OBJECT

The *maṣdar* as absolute object may have an adjectival attribute, as in (36).

(36) *ḍaraba zayd-un ʿamr-an*
hit3ms.Past Zayd-Nom ʿAmr-Acc
ḍarb-an šadīd-an
hitting-Acc violent-Acc
'Zayd hit ʿAmr violently'

When the *maṣdar* is deleted, only its attribute remains, as in (37).

(37) *sāfar-tu* ***kaṯīr-an***
traveled-I much-Acc
'I traveled (too) much' (instead of *sāfartu safaran kaṯīran*)

The *maṣdar* as absolute object may also be specified by a demonstrative, as in (38).

(38) *jalast-u* ***hāḏā*** ***l-julūs-a***
sat-I Dem. sitting-Acc
'I sat in this way [lit. I sat this sitting]'

The absolute object may be specified by a relative sentence in which the resumptive pronoun refers to the *maṣdar*, as in (39).

(39) *katab-tu* ***kitābat-an*** ***lam***
wrote-I writing-Acc not
yaktub-hā ***ġayr-ī***
wrote-it other.than-me
'I wrote something not written by anyone else'

The absolute object may be a → construct state with a word such as *ʾayy* 'whichever', *kull* 'all', or *baʿḍ* 'some', to which the *maṣdar* is annexed, as in (40)–(42).

(40) *zayd-un* *yaʿmalu*
Zayd-nom work.3ms
kull-a ***l-ʿamal-i***
all-Acc the-work-Gen
'Zayd works very hard'

(41) *sahir-tu* ***baʿḍ-a***
pass.sleepless-1s.Past some-Acc
s-sahar-i
sleeplessness-Gen
'I spent some of the night sleeplessly'

(42) *ijtahad-tu* ***ʾayya*** ***jtihād-in***
exerted-I whichever exertion-Gen
'I worked very hard/I did my utmost'

A similar construction is that with the interrogative *ʾayy*, as in (43), and the conditional/relative *ʾayy* as in (44).

(43) ***ʾayyat-a*** ***kitābat-in***
what kind-Acc writing-Gen
taktub-u
write.2ms
'What kind of writing are you writing?'

(44) ***ʾayya*** ***julūs-in*** *ta-jlis* *ʾajlis*
any-Acc sitting-Gen you-sit I-sit
'I will sit the way you sit/whichever way you sit'

Words such as *miṯl* and the particle *ka-* and other words with similar meanings may be used in combination with the *maṣdar*, as in (45) and (46).

(45) *jalas-tu* ***miṯl-a*** ***julūs-i-ka***
sit-1s.Past like-Acc sitting-Gen-your
'I sat like you/I sat the way you sat'
(instead of *jalastu julūsika*)

(46) *jalas-tu* ***ka-mā*** *jalas-ta*
sit-1s.Past like-what sit-2ms.Past
'I sat the way you sat/I sat as you did'
(instead of *jalastu ka-julūsika*)

8. TRANSLATION OF ABSOLUTE OBJECT STRUCTURES INTO ENGLISH (AND PROBABLY OTHER LANGUAGES)

The absolute object in Arabic is usually to be translated in English with an → adverbial phrase, regardless of its governors (Homeidi 2000:218). A typical example was given in (27), where the absolute object *ḍarban mubarriḥan* was translated with 'violently'. Other examples were given in (29), where the absolute object *qirāʾatan muʿammaqatan* was translated with 'thoroughly', and (30), where the absolute object *iḥtirāman šadīdan* was translated with 'greatly, highly'. In (47), the word *marratayni* 'twice' replaces a *maṣdar*.

(47) *qaraʾ-tu* *l-qiṣṣat-a*
read-1s.Past the-story-Acc
marrat-ayni
time-du.Acc
'I read the story twice'

9. CONCLUSION

Arabic has many types of objects according to the Arabic syntactic tradition, i.e. *al-mafʿūl bihi* 'direct object', *al-mafʿūl maʿahu* 'comitative object', *al-mafʿūl fīhi* 'the object that designates time or place of the verb', *al-mafʿūl al-muṭlaq* 'absolute object', *al-mafʿūl li-ʾajlihi* 'object of result'. It seems, however, that these terms that designate different types of object are simply semantic terms or categories and do not reflect different syntactic structures. All these types of objects are complements of the verb; that is, the verb, or one of its substitutes, such as the verbal noun, is the governor.

The absolute object is an infinitive (verbal noun, *maṣdar*, *nomen verbi*) in the accusative, governed by the verb or any +verb, + noun word, derived from the same lexical root of its verb. It is used to express many different semantic functions, the most important of which relate to the manner in which the action is performed.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Ibn Manẓūr, *Lisān al-ʿArab* = ʾAbū l-Faḍl Jamāl ad-Dīn Mukarram ibn Mukarram Ibn Manẓūr, *Lisān al-ʿArab*. Cairo: Dār al-Maʿārif, n.d.

Ibn as-Sarrāj, *ʾUṣūl* = ʾAbū Bakr ibn as-Sarī Ibn as-Sarrāj, *Kitāb al-ʾuṣūl fī n-naḥw*. Ed. ʿAbd al-Ḥusayn al-Fatlī. 3 vols. Beirut: Muʾassasat ar-Risāla, 1985.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. 2 vols. Bulaq, 1316 A.H. (Repr., Baghdad: Al-Muthanna Library, n.d.)

Zamaxšarī, *Mufaṣṣal* = ʾAbū l-Qāsim Maḥmūd ibn ʿUmar az-Zamaxšarī, *Kitāb al-mufaṣṣal fī n-naḥw*. Ed. Jens Peter Broch. Christiania: Libraria P.T. Mallingii, 1870.

SECONDARY SOURCES

ʾAnṭākī, Muḥammad al-. 1982. *al-Muḥīṭ*, II. Aleppo: Dār aš-Šarq al-ʿArabī.

Badawi, Elsaid, Michael G. Carter, and Adrian Gully. 2004. *Modern Written Arabic: A comprehensive grammar*. London and New York: Routledge.

Cantarino, Vicente. 1975. *Syntax of Modern Arabic prose*, II. Bloomington: Indiana University Press.

Cowell, Mark W. 1964. *A reference grammar of Syrian Arabic*. Washington, D.C.: Georgetown University Press.

Fleisch, Henri. 1979. *Traité de philologie arabe*. II. *Pronoms, morphologie verbale, particules*. Beirut: Dar El-Machreq.

Haywood, J.A. and H.M. Nahmad. 1984. *A new Arabic grammar of the written language*. London: Lund Humphries.

Homeidi, Moheiddin Ali. 1987. *Modality in Government and Binding*. Ph.D. diss., University of Essex.

——. 1994. "Government and Binding and case assignment in Modern Standard Arabic". *Papers and Studies in Contrastive Linguistics* 29.123–140.

——. 2000. "Syntactic competence as a prerequisite for translation". *Babel* 46:3.211–227.

Levin, Aryeh. 1991. "What is meant by *al-mafʿūl al-muṭlaq*?". *Semitic studies in honour of Wolf Leslau on the occasion of his eighty-fifth birthday*, ed. Alan S. Kaye, II, 917–926. Wiesbaden: O. Harrassowitz.

Naser, T. 1967. *The structure of Arabic*. Beirut: Librairie du Liban.

Rājiḥī, ʿAbdallāh ar-. 1988. *at-Taṭbīq an-naḥwī*. Beirut: Dār an-Nahḍa al-ʿArabiyya.

Reckendorf, Hermann. 1921. *Arabische Syntax*. Heidelberg: C. Winter.

Talmon, Rafi. 1999. "The syntactic study of *Mafʿūl Muṭlaq*: A study of Qurʾānic syntax". *Arabic grammar and linguistics*, ed. Yasir Suleiman, 107–129. Richmond: Curzon.

Woidich, Manfred. 2006. *Das Kairenisch-Arabische: Eine Grammatik*. Wiesbaden: O. Harrassowitz.

Wright, William. 1964. *A grammar of the Arabic language*, II. 3rd ed. Cambridge: Cambridge University Press.

MOHEIDDIN ALI HOMEIDI (King Saud University)

Obligatory Contour Principle

The Obligatory Contour Principle (OCP), originally proposed to account for the distribution of tones in West African languages (see Leben 1973; Goldsmith 1976), has been extended to a wider range of phenomena, leading to McCarthy's formal definition of the principle: "At the melodic level, adjacent identical elements are not permitted" (1986:208). The study of Arabic phonology and morphology has had a profound influence in extending the role of the OCP in grammar. The OCP, as first proposed, is illustrated by the tonology of Mende, which has surface tone patterns LHH and HLL (H high tone; L low tone), but there are no words with tone patterns HHL and LLH. The OCP prohibits adjacent identical tones; therefore, the grammatical sequences LHH and HLL must be derived from LH and HL sequences, with doubling of the final tone as a result of autosegmental spreading (Goldsmith 1976). Since adjacent identical tones are not permitted, HHL and LLH surface tone patterns are ungrammatical, because autosegmental association (i.e. associating tones one-to-one, left to right) can only produce the grammatical HLL and LHH from HL and LH, respectively.

Although there are counterexamples to the OCP in the African tone phonology (see Odden 1986), the distribution of consonants and vowels in Arabic stems exhibits a similarity to the distribution of tones. McCarthy (1979) accounts for the distribution of consonants and vowels in Arabic by proposing that the OCP governs the Arabic lexicon. Arabic triliteral verb roots display a striking asymmetry: there are a number of verb roots where the second and the third radical are identical (called geminated verbs), e.g. *samam* 'poison', *jarar* 'pull', traditionally considered to be derived from the roots *s-m-m* and *j-r-r*. However, there are no triliteral verb roots with identical first and second radicals, that is, there are no verbs such as **sasam* or **jajar* from the roots **s-s-m* and **j-j-r*.

The asymmetry in Arabic roots, as McCarthy (1979) demonstrates, follows from the OCP's application to the Arabic lexicon. By extending the OCP to prohibit identical adjacent segments, McCarthy argues that all geminated verbs are derived from underlying biliteral roots, e.g. *s-m* and *j-r*. Autosegmental association (one-to-one, left-to-right) to the CVCVC template of Form I will never produce doubling of the first radical (see Fig. 1).

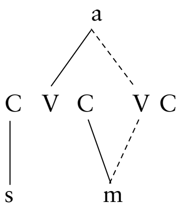


Figure 1. Template of Form I

Quadriliteral roots also show evidence of the OCP. There are no Form I stems like **dadraj* or **darraj*, because **d-d-r-j* and **d-r-r-j* are impossible roots by the OCP. Any doubling in a quadriliteral stem must be due to an underlying triliteral root, such as *d-r-j*, and to association to a CVCCVC template. This can only be mapped autosegmentally as *darjaj*.

The OCP is pervasive throughout Arabic grammar, as is also evident in the distribution of the vocalism of verb stems. Consider the distribution of vowels in the different morphological classes in (1) (from McCarthy 1979).

(1)				
	perfect active:		{a a, a a a, a a a a}	
	perfect passive:		{a, a a, a a a}	i
	imperfect passive:	u	{a a, a a a, a a a a}	
	active participle:	u	{a, a a, a a a}	i
	passive participle:	u	{a a, a a a, a a a a}	

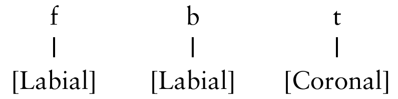
Perfect active stems can have two to four vowels, all of which are /a/, and perfect passive stems also have two to four vowels, with /i/ as the final vowel. Abstracting away from the final /i/ of the perfect passive and the active participle, the distribution of vowels in the stem can be accounted for autosegmentally. With the OCP, the underlying vocalism of the imperfect passive, for example, is /u a/, and autosegmental association ensures the second vowel spreads to all unassociated vowel slots. The underlying vocalism of the active participle is /u, a, i/, and it is /a, i/ in the perfect passive. The consistent appearance of /i/ as the final vowel of the stem is handled by a specific rule that associates it to the last vowel slot of the template before left-to-right association of the vowels (McCarthy 1979).

The OCP not only accounts for the absence of identical first and second radicals, but it can also account for the distribution of homorganic consonants in roots. Consonants (excluding glides) in the classes in (2) tend not to co-occur in roots (Greenberg 1950; McCarthy 1994).

(2)		
	Labial:	{f, b, m}
	Coronal stops:	{t, d, ṭ, ḍ}
	Dorsals:	{j, k, q, x, ġ}
	Coronal sonorants:	{l, r, n}
	Coronal fricatives:	{ṯ, ḏ, s, z, ṣ, ẓ, š}
	Gutturals:	{x, ġ, ḥ, ʿ, h, ʾ}

Leaving aside the issue of the dual class membership of /x, ġ/, the distribution of consonants in roots can be accounted for by applying the OCP to the place node of the feature-geometry (Yip 1989; McCarthy 1988, 1994; Padgett 1995). A root such as *f-b-t*, shown in (3), is ungrammatical because the adjacent labial nodes violate the OCP. OCP effects on place features have been noted in other languages (see Mester 1986; Yip 1989).

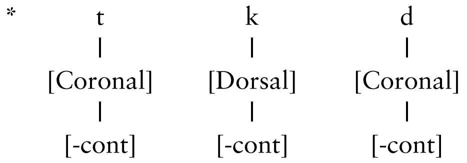
(3) *



The logic of feature-geometry predicts that nonadjacent homorganic consonants are also ungrammatical. Since the place features occupy different autosegmental tiers, the two labial nodes in a root such as *b-t-f* are adjacent, so the OCP is applicable. However, there are exceptions when the consonants are not adjacent.

The fact that coronal consonants with the same manner features tend not to co-occur is captured formally by augmenting the arrangement of the features in the feature-geometry. Padgett (1995) proposes that the manner features [continuant] and [sonorant] are dominated by the place features, as shown in (4). Hence, the absence of a root such as *t-k-d* is captured by the OCP's application to the manner features as well as the place features. As before, the OCP at the place node applies to nonadjacent homorganic consonants, although there are exceptions.

(4)



The OCP, as presented, is a categorical constraint, but this interpretation is too strong to apply in all cases in Arabic. The OCP properly accounts for the absence of identical consonants as first and second radical in roots, but it cannot properly account for the distribution of homorganic consonants because there are a number of exceptions when the homorganic consonants are not adjacent. As a result, the OCP is often considered to consist of two different constraints, with the OCP place constraint a weaker constraint (McCarthy 1988). It is difficult to formalize this notion of a weaker constraint (even in → Optimality Theory), since the co-occurrence restrictions on homorganic consonants are best described as tendencies and cannot be considered a set of categorical constraints that can be ranked in an Optimality-theoretic grammar. Furthermore, as noted by Pierrehumbert (1993), identical consonants are also homorganic, and so it is difficult to effectively separate these two uses of the OCP.

A statistical interpretation of OCP phenomena (Pierrehumbert 1993; Frisch a.o. 2004) provides an alternative to the categorical OCP. A statistical analysis of consonant co-occurrence in Arabic reveals that the OCP is a gradient constraint and that the basis for the OCP in Arabic is a tendency to avoid similar consonants in the root. The presence of different consonant combinations in the lexicon is measured by the O/E value, which is "the ratio of the observed number of occurring consonant pairs (O) to the number that would be expected if consonants combined at random (E)" (Frisch a.o. 2004:185). An O/E value of less than 1 indicates that there are fewer observed combinations of consonants than expected by chance. Lower values (approaching 0) indicate that there is a greater co-occurrence restriction on the consonant pair. Eliminating cases of identical consonants, the O/E values for homorganic consonants in roots are given in (5) (taken from Pierrehumbert 1993).

(5)	Class	O/E adjacent	O/E nonadjacent
	Labial	0.00	0.41
	Coronal sonorant	0.09	0.95
	Coronal obstruent	0.32	0.70
	Dorsal	0.04	0.36
	Guttural	0.07	0.69

The ratio of observed cases of adjacent homorganic consonants (either as first and second radical or as second and third radical) compared to expected combinations by chance is extremely low. This indicates strong avoidance of these consonants in the root, which is compatible with an OCP place constraint. Notice that in classes other than coronal obstruents there is a categorical OCP effect. However, in cases where the homorganic consonants are not adjacent (i.e. the first and third radical), there is a weakening of the OCP, as is evident by the greater O/E values. Nonadjacent, homorganic consonants are dispreferred but not categorically ungrammatical, as predicted by the OCP and feature-geometry. Even in cases in which the first and the third radical are identical, there is a weakening of the OCP, but there is still a strong tendency against co-occurrence (O/E = 0.14). Frisch a.o. (2004) note that this is problematic because the identical consonants are separated by intervening material on all relevant tiers, so a categorical interpretation of the OCP predicts there should be no co-occurrence restriction.

Frisch a.o. (2004) propose that the weakening of the OCP should be interpreted as avoidance of similar consonants in the same root. Similarity is measured as a function of the shared and nonshared natural classes of the homorganic consonants. Formally, similarity is equal to shared natural classes divided by the sum of shared and nonshared classes. Looking at the labial consonants, /f/ and /m/ have two shared natural classes and seven nonshared natural classes, whereas /f/ and /b/ have three shared natural classes, thus producing a higher similarity rating. Therefore, /f/ and /m/ are more likely to co-occur than /f/ and /b/. This is correct: of the 17 cases of nonadjacent homorganic labials in a root, only one contains two labial obstruents. The tendency for coronals to co-occur in a root is a product of the large number of natural classes of coronal consonants. Coronals /s/ and /n/ differ in a large number of natural classes (since there are many coronal distinctions). Hence, /s/ and /n/ are less similar to each other than any two labials and so the former pair of consonants are not subject to the co-occurrence restrictions. Similarity avoidance accounts for the subregularities found among coronals. For example, adjacent emphatic coronals do not co-occur (O/E = 0), but an emphatic coronal and other coronals are more likely to co-occur (O/E = 0.35) since these coronals are less similar. The effect of manner in the distribution of coronals (and in other classes) is directly accounted for by similarity avoidance. There is no need to separate coronals by manner features or to manipulate the feature-geometry.

By extending the OCP beyond tone distribution to the distribution of segments and features, the study of Arabic phonology and morphology has been instrumental in promoting principles in phonology. Furthermore, OCP effects on place of articulation and manner are used to justify the place node in feature-geometry and the organization of the tiers. The gradient nature of OCP place phenomena raises many interesting questions concerning the nature of gradient constraints in phonology, phonological principles in general, and questions concerning the interface between phonetics and phonology.

BIBLIOGRAPHICAL REFERENCES

Frisch, Stefan, Janet Pierrehumbert, and Michael Broe. 2004. “Similarity avoidance and the OCP”. *Natural Language and Linguistic Theory* 22.179–228.

Greenberg, Joseph. 1950. “The patterning of root morphemes in Semitic”. *Word* 5.162–181.

Goldsmith, John. 1976. *Autosegmental phonology*. Ph.D. diss., Massachusetts Institute of Technology.

Leben, William. 1973. *Suprasegmental phonology*. Ph.D. diss., Massachusetts Institute of Technology.

McCarthy, John J. 1979. *Formal problems in Semitic phonology and morphology*. Ph.D. diss., Massachusetts Institute of Technology.

——. 1986. “OCP effects: Gemination and antigemination”. *Linguistic Inquiry* 17.207–264.

——. 1988. “Feature geometry and dependency: A review”. *Phonetica* 43.84–108.

——. 1994. “The phonetics and phonology of Semitic pharyngeals”. *Phonological structure and phonetic form: Papers in laboratory phonology*, III, ed. Patricia Keating, 191–233. Cambridge: Cambridge University Press.

Mester, R. Armin. 1986. *Studies in tier structure*. Ph.D. diss., University of Massachusetts, Amherst.

Odden, David. 1986. “On the role of the obligatory contour principle in phonological theory”. *Language* 62.353–383.

Padgett, Jaye. 1995. *Stricture in feature geometry*. Stanford, Calif.: CSLI Publications.

Pierrehumbert, Janet. 1993. “Dissimilarity in the Arabic verbal roots”. *Proceedings of the 23rd Meeting of the Northeast Linguistic Society*, 367–381. Amherst: University of Massachusetts.

Yip, Moira. 1989. “Feature-geometry and cooccurrence restrictions”. *Phonology* 6.349–374.

SAMUAL ROSENTHALL
(Rochester, Mich.)

Old Arabic (Epigraphic)

1. THE EVIDENCE FOR OLD ARABIC

The term ‘Old Arabic’ (*vieil-arabe*, *Altarabisch*), like the terms ‘Old English’, ‘Old Aramaic’, etc., refers to the earliest surviving examples of the Arabic language, from which, it is assumed, the later forms evolved. This early stage almost certainly existed as a number of spoken dialects, used at various times in the pre-Islamic period in different parts of the Arabian Peninsula, the Levant, Mesopotamia, and Egypt (on the presence of peoples called ‘Arabs’ (→ ʿArab) in these regions from early times, see Macdonald 2003:313–316). However, several factors make it difficult to reconstruct the early stages of the language and complicate attempts to trace its development into the forms of Arabic known from the early Islamic period onward.

At present, we have no evidence that Arabic was widely written before the rise of Islam, or that any particular script was regularly associated with it, at least before the 5th century C.E. Although this may partly be due to accidents of survival and discovery, it is virtually certain that Arabic was a purely spoken, rather than a habitually written, language for most of its pre-Islamic history.

We have two sources for Arabic in the pre-Islamic period. The first may loosely be called the 'epigraphic evidence' and consists of a small number of documents expressed in the Arabic language but transcribed in scripts normally used to write other tongues (Sabaic, Nabataean Aramaic, Dadanitic, see Macdonald 2004:490, 492; Safaitic, see Macdonald 2004:492, see 2.1 below). In addition, there are some Arabic lexical items embedded in written texts in other languages. Some of these occur in a number of brief graffiti in Nabataean Aramaic recently discovered in northwest Arabia, which are written in transitional scripts between late Nabataean and what is recognizably an early Arabic script. When they can be dated, these appear to come from the 4th and 5th centuries C.E.

The second source for Arabic in the Jāhiliyya can loosely be called the 'literary evidence' since it consists of large numbers of texts from the pre-Islamic period (principally the *Qurʾān*, the pre-Islamic poetry, and the *ʾAyyām al-ʿArab*), which were originally transmitted orally and were not written down until the first two centuries of Islam.

These two types of source differ in several ways. First, while the epigraphic evidence has survived undisturbed since antiquity, the literary evidence has a transmission history. This is particularly relevant because the philologists of the early Islamic centuries took the language of the pre-Islamic poetry and the *Qurʾān* as the basis for the systematized form of the Arabic language, which became known as → *al-ʿarabiyya al-fuṣḥā* or → Classical Arabic. It is difficult to be sure whether, or to what extent, elements of the language of the pre-Islamic poetry and the *Qurʾān* may have been altered in order to exemplify the grammatical system of Classical Arabic (see, e.g., Rabin 1951:24, 1955:21–22; → pre-Islamic Arabic).

Second, with the exception of inscription no. VI, and possibly nos. V and VII, below, the documents which make up the epigraphic evidence are not literary works and did not call for a high or rhetorical style. By contrast, the literary evidence for Old Arabic is oral literature (only later recorded in writing) which is expressed in a high style and, in the view of many scholars, in a supratribal poetic dialect, often referred to as a koine (→ poetic koine). For a balanced description of the debate on this point, plus the essential references, see Versteegh (1997:46–52).

Finally, while the epigraphic evidence all comes from the western two-thirds of the peninsula, the poetic (and perhaps religious) koine is thought to be closer to the dialects of the tribes of eastern Arabia, although it was used in both areas for poetry and other forms of declamation.

The epigraphic evidence, with which this article is concerned, presents a number of problems. The documents in which it is contained date from periods when Arabic was not habitually written but instead coexisted with a number of other languages which were written (see Macdonald 2000:48–54, 2005:98).

None of the scripts used to transcribe Arabic were particularly suited to the task. Greek lacked a number of the consonants of Arabic, and, while it showed approximations to the quality of the vowels, by late antiquity it no longer showed vowel length. Nabataean Aramaic showed long vowels, except medial /ā/, but had only 21 signs to express the 28 consonants of Arabic (Aramaic *semkath* expressed a sound which, at that time, was not in the phonemic repertoire of Arabic and so was not used in writing Arabic; see Macdonald 2004:498–499, and the detailed discussion in Macdonald [in preparation]). The Sabaic and Ancient North Arabian scripts at least had a sufficient quota of signs to express Arabic's consonantal repertoire, but (with the exception of Dadanitic; see Macdonald 2004:495) normally expressed no vowels or diphthongs at all. Apart from Greek, none of these scripts showed short vowels, medial /ā/, or doubled letters, so, for instance, the Arabic verb Forms I, II, and III were indistinguishable from each other in these scripts, as were Forms V and VI. Thus, while we can be sure that the texts of these documents are original and have not been subjected to later interference, their inter-pretation is often uncertain, and the information they provide, though valuable, is incomplete.

Apart from the handful of texts composed intentionally in the Old Arabic language, there are texts where Old Arabic seems to have been recorded accidentally. The Arabic speakers who composed these inscriptions appear to have been trying to produce correct Sabaic or Nabataean Aramaic, for example, but because they had an imperfect grasp of these languages, they filled in the gaps in their knowledge with Arabic words or phrases. A parallel to this can be found in most → Middle Arabic texts, whose authors were trying to write Standard Arabic but often fell back on forms and structures from their spoken dialects (Blau 1981:189).

A particular difficulty is caused by the fact that from the second half of the 1st millennium B.C.E. until about the 4th century C.E., spoken Old Arabic coexisted with a number of written Ancient North Arabian dialects (→ Thamudic). These were distinct from Arabic in a number of ways, for instance by having a definite article *h-* (or no definite article), rather than *ʾl-* (see Macdonald 2000:41–50, 2004:488–493). Old Arabic and Ancient North Arabian can be seen as distinct but related dialect bundles which together formed the North Arabian linguistic group (Macdonald 2004:488). Given their closeness, however, and the deficiencies of the writing systems employed, it is often difficult to distinguish whether a particular text is in Old Arabic or in Ancient North Arabian (see Macdonald 2000:50–57). The language of texts where this is the case has therefore been classified as 'Undifferentiated North Arabian' (see Sec. 2.4) and is not really valid as evidence for Old Arabic specifically (Macdonald 2000:54–57, 61).

2. The material of epigraphic Old Arabic

2.1 *Early fragments*

Livingstone (1997) suggests that two North Arabian words (*nāqa* 'she-camel' and *bakr* 'young male camel') were included in the annals of the Assyrian king Tiglath-Pileser III (744–727 B.C.E.), in the forms *a–na–qa–a–te* and *ba–ak–ka–ri–ši–na*. The former bears the Assyrian feminine plural ending *-āte*, and the latter has the Assyrian 3rd person feminine plural pronominal suffix *-šina*. Livingstone suggests that the initial *a-* in *a–na–qa–a–te* represents the Arabic definite article *al-* (with the [l] assimilated before the [n] of *nāqa*) or the Ancient North Arabian definite article *ha-* (the [h] of which would not be represented in Akkadian). He considers that the Assyrian scribes may have accidentally included the article in the word when they took it over from the (spoken) language of the vanquished Arab tribes about whom they were writing. While this analysis is not undisputed, it certainly cannot be ruled out.

In the 5th century B.C.E., Herodotus reports that "the Arabs" (in this case in eastern Egypt) worshipped a sky goddess called Alilát (*Historiae* III.8; see also I.131). Alilát presumably represents **al-ʾilat* 'the goddess' (on the form of the word, see Macdonald [in preparation]). Interestingly, there is a roughly contemporary witness to the worship of 'the goddess' at a shrine at Tell al-Masxūṭa in the eastern Nile Delta, where she appears in the Ancient North Arabian form *hn-ʾlt* in inscriptions in Aramaic (Rabinowitz 1956, 1959).

The Nabataeans were regularly called 'Arabs' by their contemporaries, although it is unclear what exactly was meant by the term. It is widely supposed that, although they used Aramaic as their written language, most of the inhabitants of the Nabataean kingdom from the Ḥawrān to Ḥegrā, spoke forms of Arabic. While that may well be correct, at least for some parts of the kingdom, this view was until recently based on very little secure evidence (see the discussion in Macdonald 2000:46–48, and in preparation). Only a few Arabic loanwords had been identified in Nabataean Aramaic, and of these all but two were confined to texts in Arabia (O'Connor 1986; Macdonald 2000:47). Recently, however, a relatively large number of Arabic words have been identified in the Nabataean legal documents on papyrus found near the Dead Sea (Yadin a.o. 2002). This has more than doubled the number of Arabic loanwords so far recognized in Nabataean. Although they occur in legal documents, the words are taken from widely varied spheres of life, some being used as metaphors and others in their primary meaning. Moreover, Levine has shown how, in this archive, "the Jewish-Aramaic and the Nabataean-Aramaic documents differ in the way each proliferates legal formulas to embrace all sorts of conceivable, often redundant provisions". He reports that "scribes writing in Nabataean-Aramaic proliferated legal formulas by using Arabic equivalents of the Aramaic terms of reference, whereas scribes writing in Jewish Aramaic accomplished the same result with Hebrew and Aramaic synonyms" (Levine 2000:844–845; and see Yadin a.o. 2002:221–222, and 170–276 *passim*). If this comparison is exact, and Arabic had for the

Nabataeans anything like the status which Hebrew had for the Jews, then it implies that Arabic played a sophisticated cultural role in Nabataean society, even in regions where it might not be expected at that period (e.g. Moab [*P.Yadin* 1] and the southern end of the Dead Sea [*P.Yadin* 2 and 3]). For a detailed study, see Macdonald (in preparation). Some 250 years later, Epiphanius of Eleutheropolis records that the people of Petra, and Elusa in the Negev, sang hymns in the Arabic language (*Panarion* 51.22.11). While we do not know how long this had been going on, it could be another indication of the place of Arabic in the life of the Nabataeans, as is the ʿĒn ʿAvdat inscription, no. VI below.

The 6th-century Greek papyri found at Petra contain approximately 130 toponyms and names of buildings which are recognizably Arabic, 110 of them in two papyri, Inv 10 and Inv 98 verso, neither of which has yet been published. For a survey and discussion of these papyri, see Daniel (2001) and references there. A name or word *nāy(i)f* or *nāy(i)q* inscribed on a piece of wood in what would normally be regarded as a late form of the Arabic script using diacritical marks was apparently found in the same locus as some of the scrolls (see Ghul 2004). However, doubts must remain as to the authenticity of this piece given the relatively late form of the final letter, the arrangement of the diacritical dots under the *yāʾ*, and the difficulty of interpreting the word as anything other than the modern name *Nāyif*.

To this evidence should be added the substitution of the Arabic passive participle *mdkwr* (**madkūr*) for its Aramaic equivalent *dkyr* in three (out of almost four thousand published) Nabataean graffiti in Sinai (CIS ii 1331, 2768; Negev 1977:56, no. 219). However, apart from this one word, the language of these three graffiti (insofar as it is visible in texts consisting mostly of names) is Aramaic. Thus, Aramaic *bar* is used rather than Arabic *ibn* for 'son of' (although this is also a feature of the Old Arabic Namāra, Ḥarrān, Zebed, and Jabal ʾUsays inscriptions; see below), and the Aramaic phrase *b-ṭb* (**bi-ṭāb*) occurs in CIS ii 2768 and Negev no. 219, as it does in huge numbers of the graffiti from Sinai. There is also some scattered evidence for Arabic influence on the syntax of Nabataean Arabic (see al-Hamad 2005).

There is much less evidence of Arabic influence on the Palmyrene Aramaic language, with only a handful of Arabic loanwords appearing in the inscriptions known to date. See Müller (1982:31–32).

See Contini (1987:51–60) and Millar (1998) for excellent discussions of the complex problems of language use in Syria during the Roman and Byzantine periods.

2.2 *Old Arabic texts*

For a brief description and discussion of most of these texts, see Macdonald (2000:48–57, 61). For a full study of each of them and their contribution to our knowledge of the language, with a list of linguistic features and vocabulary, see Macdonald (in preparation).

2.2.1 In the Sabaic script

An Old Arabic text in the Sabaic script was found at Qaryat al-Fāw on the northwestern edge of the Empty Quarter, in Saudi Arabia. This was a thriving city on the caravan route between South Arabia and the Arabian/Persian Gulf, Mesopotamia, and the Levant. It had a wealthy and sophisticated culture between the 3rd/2nd century B.C.E. and perhaps the 5th century C.E. (al-ʾAnṣārī 1982). Although it was the 'capital' of the Arab tribes of Qaḥṭān, Maḏḥij, and Kinda, whose leaders called themselves 'kings', it was within the cultural orbit of the great South Arabian kingdom of Sabaʾ (Sheba), and thus the Sabaic language and script was the normal medium for written documents, even though Arabic was almost certainly the most widely spoken language.

I. The grave inscription of ʿIgl bn Hfʿm

A ten-line text in Old Arabic written in the Sabaic script, dated on palaeographical grounds (which are uncertain) to approximately the end of the 1st century B.C.E. See al-ʾAnṣārī (1979:6–8, Pl. 1); Beeston (1979:1–2); Robin (2001:548–549). Photographs al-ʾAnṣārī (1979, Pl. 1; 1982: 146).

2.2.2 In the Dadanitic script

Dadan (biblical Dedān, modern al-ʿUlā) was a large oasis in northwest Arabia, situated at a strategic point on the trade route between South Arabia and the Levant (Macdonald 1997:341–343). It was the seat of the kingdoms of Dadan and Liḥyān, whose subjects left numerous monumental inscriptions and graffiti in their own forms of the Ancient North Arabian alphabet and language: ‘Dadanitic’ (formerly called ‘Dedanite’ and ‘Lihyanite’; see Macdonald 2000:33). Unfortunately, these are at present impossible to date (Macdonald, in the press).

II. JSLih 384

An inscription of four lines in the Dadanitic script and orthography (Macdonald 2004:494–496), probably from al-Xereibah or al-ʿUlā and now in the Musée du Cinquantenaire, Brussels (registration no. 0.715). Published by Jaussen and Savignac (1909–1922:II, 532–534, Pl. 94). Recognized as Old Arabic by W.W. Müller (1982:32–33). For a discussion of its Old Arabic features, see Macdonald (2000:50).

III. JSLih 276

A single line in the Dadanitic script. Photograph in Jaussen and Savignac (1909–1922:II, Pl. XCII, facsimile from a squeeze Pl. CXXXV). For the reading, see Macdonald (2000:53), where it was treated as a ‘mixed text’. However, it is now considered to be pure Old Arabic (Macdonald, in preparation). It contains a formula which is also found in Dadanitic, but with the Arabic definite article – with the [l] assimilated before a following sibilant, rather than the Dadanitic definite article *h*.

Other inscriptions in the Dadanitic script which appear to contain the Arabic article have recently been discovered and will be studied in Macdonald (in preparation).

2.2.3 In the Safaitic script

Safaitic was an Ancient North Arabian alphabet used almost entirely by nomads to carve graffiti on the rocks of the deserts in what is now southern Syria, northeastern Jordan, and northern Saudi Arabia, between the 1st century B.C.E. and the 4th century C.E. (see Macdonald 1995, 2000:45–46, 2005:74–75).

IV. CIS v 5137 (not in Macdonald 2000)

A short inscription in the Safaitic script and orthography, known only from a hand copy (CIS v Pl. XLVI, no. 801). It was found in the basalt desert east of the Ḥawrān. It contains the author's genealogy followed by the statement *w ḥll ʾl-dr* (**wa-ḥallal ʾal-dār*) ‘and he camped here [lit. ‘in the place’]’, which parallels the very common Safaitic statement *w ḥll h-dr*. On the spelling *ḥll*, see Macdonald (2004:510). Note that in the inscription the *l* of the article is not assimilated to the *d* of *dr* (on this feature see Macdonald, in preparation). Naturally, this single feature in a text which otherwise consists only of names is by no means conclusive evidence that its author spoke ‘Old Arabic’, but given the rarity of *ʾl-* rather than *h-* in the Safaitic inscriptions, its occurrence here is worth noting.

A few other texts in the Safaitic script may contain a form of the Arabic article, this time assimilated before a sibilant (e.g. *ʾ–s*^{1}*nt* (‘the [this] year’ and *ʾ–s*^{1}*fr* ‘the [this] writing’; see Macdonald 2000:51–52). However, these should be treated as uncertain since the letters *ʾ* and *h* in the Safaitic script are differentiated by only one small sidestroke, which could have been added or omitted accidentally.

2.2.4 In the Hismaic script

Hismaic (see Macdonald 2000:44–45) is an Ancient North Arabian script used by nomads in southern Jordan and northwestern Saudi Arabia, and possibly by some settled people in central Jordan, ca. 1st century B.C.E. to 3rd century C.E.

V.

The language of two inscriptions in the Hismaic script from the region of Madaba, Jordan, has been identified as Old Arabic (Graf and Zwettler 2004). However, this identification is by no means universally accepted and will be examined in detail in Macdonald (in preparation).

2.2.5 In the Nabataean script

The Nabataeans used Aramaic as their written language and wrote it in a particular form of the Aramaic alphabet, the ‘Nabataean script’. This continued in use in several parts of the kingdom, particularly northwest Arabia, long after Rome's annexation of the kingdom, and eventually came to be used to write the Arabic, rather than the Aramaic, language (see inscriptions VI, VII, and VIII). A late form of the Nabataean alphabet is what we know as the ‘Arabic script’ (→ Arabic alphabet: origin).

VI. The ʿĒn ʿAvdat inscription (Negev a.o. 1986)

A graffito of six lines from ʿĒn ʿAvdat in the Negev in which lines 1, 2, 3, and 6 are in

Nabataean Aramaic and lines 4 and 5 in Old Arabic, written in the Nabataean script. It is undated (the dating suggested in Negev a.o. 1986:60 is speculative). The two lines of Arabic, which appear to be in praise of the deified Nabataean king Obodas (*ʿbdt*), have attracted a large number of interpretations, none of which has met with general agreement (see Lacerenza 2000 for most of the bibliography). However, the lines clearly have a rhetorical structure and it has even been suggested that they are the earliest example of Arabic verse (Noja 1989:192–193 [*rajaz*]; Bellamy 1990:78–79 [*ṭawīl*]; on both of which see Beeston 1994:239–243). For a photograph, see Negev a.o. (1986:Pl. 11B) or Kropp (1994:Pl. 1).

VII. The Namāra inscription
Discovered in 1901 by René Dussaud and Frédéric Macler in the desert of southern Syria near the Roman fort at a place known, both in antiquity and today, as Namāra. It is now in the Musée du Louvre (registration no. Antiquités Orientales 4083). The text consists of five long lines of Old Arabic in the Nabataean script, recounting the deeds of *Mrʾlqys* (often transcribed as 'Imruʾ al-Qays', like the Arab poet who lived two centuries later) *br ʿmrw, mlk ʾl-ʿrb kl-h* 'son of ʿAmr, king of all the Arabs' or 'king of the entire territory of *ʾl-ʿArab*' (Zwettler 1993:18). It is dated in *Kslwl* 223 in what is assumed to be the era of the Roman Province of Arabia (= November/December 328 C.E.). It has a large bibliography (see Kropp 1993:81–84), but for the most recent interpretation (by Christian Robin), together with an excellent new facsimile by Alain Desreumaux, see Bordreuil a.o. (1997), to which add the interesting discussions of the first line in Zwettler (1993, 2006) and the full-scale study of the inscription by Zwettler (in preparation). A recent detailed study of the original inscription by Macdonald has shown both that – with one or two small amendments (see Macdonald, in preparation) – Desreumaux's facsimile made from the original is extremely accurate and that none of the new readings proposed by Bellamy (1985) can be justified on the stone. The detailed interpretation of the text is, however, still a matter for lively debate among epigraphists, linguists, and historians.

2.2.6 In the transitional script between 'Nabataean' and 'Arabic'

VIII. Theeb 2002:146–148, 311: nos 132–133
A graffito from Jabal ʾUmm Jaḏāyiḏ, near al-ʿUlā, Saudi Arabia, dated to 250, which, if the era is that of the Roman Province of Arabia, would be 356 C.E. It is mostly in Nabataean Aramaic, but at the end it refers to *ʿmrw ʾl-mlk* (reading different from that of the editor, who reads the inscription as two texts). A number of other graffiti in Old Arabic, or with Old Arabic elements, written in this transitional script and apparently dating to the 5th century C.E., have been discovered recently by Saudi Arabian scholars and are being prepared for publication.

IX. The Jabal Ramm Arabic inscriptions
Three fragmentary lines in a transitional form of the Nabataean script as it developed into Arabic. They were incised in the plaster on an inner wall of the Nabataean temple at Jabal Ramm, in southern Jordan. On the same piece of plaster are two graffiti in the Ancient North Arabian scripts 'Thamudic D' and 'Thamudic B' (see Macdonald and King 2000:438; Macdonald 2000:33–35, 44). The original has disappeared. None of the texts is dated. They were presumably cut after the destruction of the temple, but there is at present no secure date for either the construction or destruction of the latter. The very speculative attempts at dating the texts (e.g. Grimme 1936:91) should be treated with caution. The best published photograph is in Grohmann (1967–1971: II, pl. I). The most recent discussion and bibliography are in Bellamy (1988:370–372), although no interpretation has found general acceptance. The right edge of the piece of plaster is broken and the beginnings of the three lines of Arabic are missing, so we have no way of knowing how much of each line has been lost. The second line seems to consist of the words… *yd-h b-ʾrm* (**yad-hu bi-ʾiram*) '…his hand, at Iram' (the premodern name of Ramm, known from Nabataean inscriptions and medieval Arab geographers), an identification made independently by Macdonald (2000:76, n. 171) and Mascitelli (2006:168).

2.2.7 In what is recognizable as the early 'Arabic' script

X. The Arabic inscription at Zebed
This inscription was found, together with one in Greek and another in Syriac, on the lintel of a martyrium at Zebed in northern Syria (ca. 60 kms southeast of Aleppo), and is now in the Near Eastern Department of the Musée du Cinquantenaire, Brussels (Inv. A.1308). It is not strictly speaking a trilingual (as it has been called) because the content of each inscription is different. The Greek and Syriac texts are dated to 823 of the Seleucid Era (= 512 C.E.). It should be noted that apart from a damaged word at the beginning, followed by the word *ʾlʾlh* (**ʾal-ʾil(ā)h*), the text in the Arabic script consists entirely of names, and the Aramaic word *br* 'son of' is used rather than Arabic *ibn*. Thus, although the script is recognizably 'Arabic', there is very little on which to identify the language. For bibliography see RCEA 1, no. 2; for excellent photographs see Grohmann (1967–1971:II, Pl. II.2); and for the most recent, and very interesting, treatment see Robin (2006:336–338).

XI. The Arabic graffito at Jabal ʾUsays
Four lines scratched on a rock at Jabal Says/ʾUsays in southern Syria. In situ. Dated to 423 in what is assumed to be the era of the Province of Arabia (= 528 C.E.). See Grohmann (1967–1971:II, 15–17, Pl. 1); and the new reading in Robin and Gorea (2002), based on a new photograph by Robert Hoyland (Hoyland 2001:202, Pl. 32b). The graffito is by a man who says he was sent to ʾUsays by *ʾl-ḥrṯ ʾl-mlk* (who is assumed to be the Jafnid [Ghassanid] king, al-Ḥāriṯ ibn Jabala, 529–569 C.E.).

XII. The Arabic inscription at Ḥarrān
An Arabic inscription with a parallel Greek text, dedicating a martyrium, now used as the lintel over the door of a house at Ḥarrān in the Lejā, southern Syria. Dated in both the Arabic and the Greek to 463 of the Province of Arabia and in the Greek to the first year of the indiction, thus 568 C.E. For bibliography see RCEA 1, no. 3; for the most recent discussion, see Robin (2006:332–336, and, briefly, 2001:544).

XIII. LPArab 1 from ʾUmm al-Jimāl, in northeastern Jordan
In situ. A five-line inscription on a rather rough piece of basalt found in a church at ʾUmm al-Jimāl (for a description, see Littmann 1949:1). The letters were originally filled with red paint. It is undated and may not be pre-Islamic, and it is uncertain whether it should really be included in the Old Arabic corpus. The most recent discussion and bibliography are in Bellamy (1988:372–377), although no interpretation has found general acceptance.

2.3 *Old Arabic mixed texts*

These are inscriptions in which elements of Old Arabic are mixed with those of the language usually associated with the script in which the text is written. The most natural explanation of this linguistic mix would seem to be that an Arabic speaker was trying to compose a text in a foreign (written) language and filled the gaps in his knowledge with words and phrases from his spoken language (Arabic), although this can, of course, be no more than a working hypothesis. The Old Arabic mixed texts can be divided according to the scripts in which they are written.

2.3.1 Dadano-Arabic

JSLih 71 is an honorific inscription incised on a rock face in the oasis of al-ʿUlā in northwestern Saudi Arabia. Photograph and squeeze in Jaussen and Saviganc (1909–1922: II, Pl. LXXXVI, facsimile Pl. CVIII). It recounts the exploits of a certain *ʿnzh bn ʾs*^{1} (*ʿAnazā ibn ʾAws) from the tribe of *ʾl-ʾḥnkt*, members of which have left inscriptions in various other parts of Arabia (see Robin 1992:54–55). The writer of the inscription incised the name, genealogy, and tribal affiliation of *ʿnzh bn ʾs*^{1} in the formal 'calligraphic' Dadanitic script (see Macdonald, in the press) and used the Dadanitic definite article *h(n)-* instead of the Arabic article *ʾl-* in front of the tribal name, thus *hn-ʾḥnkt* for *ʾl-ʾḥnkt*. However, in the lines which follow, the script becomes more and more informal, and the language appears to change to Old Arabic, written in the Dadanitic script and using Dadanitic orthography. See Beeston a.o. (1973); Macdonald (2000:52–53, and in preparation).

2.3.2 Nabataeo-Arabic

JSNab 17 is a funerary inscription at Ḥegrā (al-Ḥijr, modern Madāʾin Ṣāliḥ, some 20 km north of al-ʿUlā, in northwest Saudi Arabia). Photograph and facsimile in Healey and Smith (1989:Pl. 46). It is dated to the year 162, presumably in the era of the Province of Arabia (= 267 C.E.), and is carved in the Nabataean Aramaic script, with the name and patronym of the deceased repeated in a Thamudic D inscription carved vertically down the right-hand edge of the text. It is on a rock face between two monumental Nabataean tombs, with a rectangular burial slot beneath it (see Nehmé 2005:Fig. 20). It appears that the author was attempting to write a Nabataean Aramaic text, but like the author of JSLih 71, he filled in the gaps in his knowledge with Arabic words and phrases. For a detailed analysis, see Macdonald (in preparation). For a reading of the text as 'Arabic', see Healey and Smith (1989); for further discussion see Healey (2002), and Macdonald (2000:53).

2.3.3 Safaeo-Arabic

There are a few graffiti in the Ancient North Arabian script, in Safaitic, which seem to contain the Arabic definite article *ʾl-* rather than the Safaitic article *h-*, in texts which are otherwise in Safaitic. By contrast with the other mixed texts, these do not seem to have been attempts to write Safaitic by speakers of Old Arabic, since in each case, apart from the definite article, the text is indistinguishable from similar Safaitic graffiti. The clearest example is CIS v 2446, which contains the word *ʾl-nbṭy* ('the Nabataean'), referring to the killer of the author's brother, in an otherwise Safaitic text (see Macdonald 2000:51 for discussion).

2.4 *Undifferentiated North Arabian*

These are texts which are clearly in either an Old Arabic or an Ancient North Arabian dialect, but which do not contain sufficient diagnostic elements to distinguish which they represent. The majority of those identified so far come from the site of Qaryat al-Faw.

Ja 2122 (Jamme 1967:182–183, Pl. XLVII.1): Robin (1991:115; 2001:548, 560); Macdonald (2000:54). See Macdonald (in preparation) for a full analysis of the text.

The tomb inscription of Mʿwyt bn Rbʿt. Found at Qaryat al-Faw (al-ʾAnṣārī 1979:8). Photographs in al-ʾAnṣārī (1979:Pl.2) and Kropp (1992:65, Table 2), facsimile in al-ʾAnṣārī (1982:144/2). See the discussions in Macdonald (2000:54–55, and in preparation). Like the previous text, this could be either Old Arabic or Ancient North Arabian.

Al-ʾAnṣārī (1982:147/6) in the Sabaic script. See Macdonald (in preparation) for a reading and discussion. The language could be either Old Arabic or Ancient North Arabian. It should be noted that, despite its Arabic etymology, the divine name *ʾl-ʾḥwr*, which occurs in this inscription, can tell us nothing about the language of the text.

2.5 *Texts which cannot be treated as Old Arabic*

2.5.1 In the Sabaic script

Ghoneim (1980:323, Abb. 10). An inscription in the Sabaic script from Qaryat al-Faw, dated on palaeographical grounds (which are not particularly reliable) to the end of the 3rd or the beginning of the 2nd century B.C.E. (Robin 1991:114). Photograph Ghoneim (1980:323, Abb. 10); Kropp (1992:65, Table 1 [printed upside down]). For detailed discussions see Kropp (1992) and Macdonald (in preparation). The inscription is almost certainly in a mixture of Sabaic and Ancient North Arabian. However, Robin sees the language of the text as Old Arabic (1991:114, 2001:548, 550). For a counterargument, see Macdonald (2000:49).

Ja 2138 and 2142 (Jamme 1970:120–122, 137). Jacques Ryckmans (1980:197–198, 199 nn. 2–3) suggested that the language of these two fragmentary texts in Sabaic characters, also from Qaryat al-Faw, was North Arabian, rather than specifically Old Arabic. For arguments against this, see Macdonald (in preparation).

Al-ʾAnṣārī (1982:143/2), an isolated personal name *ʿbd-ʾl-ʿzy* (lit. 'servant of [the goddess] ʾl-ʿzy), carved in the Sabaic script into the plaster of a wall at Qaryat al-Faw. As explained above, a name is not evidence for the language of its bearer.

Haram 8, 10, 13, 33–36, 40, 53 (?), and 56. Robin (1992:33–34) suggested that these are in "un variété d'arabe avec un habillage morphologique inspiré du sabéen". However, for a discussion and counterargument, see Macdonald (2000:55–57), and for a fuller study, see Macdonald (in preparation).

2.5.2 In the Dadanitic script

JSLih 70/5–6. Werner Caskel (1954:120) claimed that these lines contain a statement in Arabic rhymed prose. This seems far-fetched, and indeed Caskel's interpretation of the whole inscription is unconvincing.

2.5.3 In the Nabataean script

Littmann and Meredith (1953:12, no. 23, Pl. I). It has been claimed (e.g. in Greenfield 1992:10) that this graffito from the eastern desert of Egypt, known only from an extremely bad hand copy, contains one line in Old Arabic. However, a glance at the copy shows that *any* interpretation can only be highly speculative, while that proposed (in desperation) by Littmann is more or less meaningless. It is therefore better to exclude this text from the Old Arabic corpus until a proper record of it is available.

2.5.4 In an eastern Arabian Aramaic script

Teixidor (1992). In Macdonald (2000:54, 61), it was suggested that a nine-line inscription on a bronze plaque from Mleiha [Mulayḥa] on the Oman Peninsula (Teixidor 1992) was 'Aramaeo-Arabic', i.e. an attempt to write a dialect of Aramaic by an Arabic speaker. However, this did not take into account Puech's rereading of the text with no Arabic elements (1998). It therefore seems safer to remove it from the list of evidence for Old Arabic.

2.5.5 In the Hebrew script

Jaussen and Savignac (1909–1922: II, 641–644, Pl. CXXI) published a handful of graffiti in the 'Hebrew' square script, from al-ʿUlā, and it has been suggested that at least some of these are in Judaeo-Arabic (Halper 1924–1925:243; Milik 1970:163; Noja 1979:308–313; Mascitelli 2006:149–151; Hopkins, in the press). However, these texts are undated, and there is no way of knowing whether or not they are pre-Islamic.

2.5.6 In the Greek script

In Macdonald (2000:50) and Macdonald (2005:100–103), it was suggested that the Arabic gloss written in Greek letters to the Septuagint version of part of Psalm 78 (LXX 77) on a parchment fragment published by Bruno Violet (1901) is in Old Arabic. Since then, good photographs of the parchment have been discovered, and it seems that it must be dated to the early Islamic period. See Mavroudi (in the press).

2.5.7 In what is considered to be the Arabic script

A mosaic in a church of the 6th century C.E. at Xirbat al-Muxayyiṭ in Jordan carries a name (Saōla) in Greek and a Semitic word, phrase, or name which the original editors interpreted as *bi-salām*, written in the Arabic script. However, this interpretation is palaeographically very difficult, and no convincing reading of the text has yet been offered (see Knauf 1984 for references).

3. CONCLUSION

It will be obvious from the above that our knowledge of epigraphic Old Arabic is at present extremely fragmentary, and it will probably always be so. However, in recent decades the pace of discovery and identification of texts has accelerated, and it is hoped that this will continue. Once their linguistic content is established, they need to be studied in the context of the other languages and dialects which coexisted with Old Arabic in pre-Islamic Arabia. At the same time, they need to be compared with the evidence from literary Old Arabic, the Classical Arabic language, and the dialects. The picture is becoming a little clearer, but there is still a long way to go.

ABBREVIATIONS

CIS II	*Corpus Inscriptionum Semiticarum*. II. *Inscriptiones aramaicas continens*. Paris: Imprimerie Nationale, 1889–1954.
CIS V	*Corpus Inscriptionum Semiticarum*. V. *Inscriptiones Saracenicae Continens*, Tomus 1, *Inscriptiones Safaiticae*. 2 vols. Paris: Imprimerie Nationale, 1950–1951.
Ja	Inscriptions in Jamme (1967)
JSHeb	Inscriptions in the Hebrew script in Jaussen and Savignac (1909–1922).

JSNab	Nabataean inscriptions in Jaussen and Savignac (1909–1922).
JSLih	Dadanitic (Lihyanite) inscriptions in Jaussen and Savignac (1909–1922).
LPArab	Arabic inscriptions in Littmann (1949).
RCEA.1	Arabic inscriptions in Combe, Sauvaget, and Wiet (1931).

Bibliographical references

ʾAnṣārī, ʿAbd ar-Raḥmān T. al- 1979 [1981]. "ʾAḍwāʾ jadīda ʿalā dawla Kinda min xilāl ʾāṯār Qariyyat al-Fāw wa-nuqūšihā". *Studies in the history of Arabia*. I. *Sources for the history of Arabia*, I, ed. ʿAbd ar-Raḥmān T. al-ʾAnṣārī, Abdelgadir M. Abdalla, Sami al-Sakkar, and Richard T. Mortel, 3–11, Pls. 1–3 and Map 1 [Arabic section]. Riyadh: University of Riyadh Press.

——. 1982. *Qaryat al-Fau: A portrait of pre-Islamic civilisation in Saudi Arabia*. London: Croom Helm; New York: St. Martin's Press.

Beeston, Alfred F.L. 1979. "Nemara and Faw". *Bulletin of the School of Oriental and African Studies* 42.1–6.

——. 1994. "Antecedents of Classical Arabic verse?". *Festschrift Ewald Wagner zum 65. Geburtstag*. I. *Semitische Studien unter besonderer Berücksichtigung der Südsemitistik*, ed. Wolfhart Heinrichs and Gregor Schoeler, 234–243. Stuttgart: F. Steiner.

——, Frederick V. Winnett, Jacques Ryckmans, and Maḥmūd al-Ghūl. 1973. "The inscription Jaussen-Savignac 71". *Proceedings of the Seminar for Arabian Studies* 3.69–72.

Bellamy, James A. 1985. "A new reading of the Namārah inscription". *Journal of the American Oriental Society* 105.31–51.

——. 1988. "Two pre-Islamic Arabic inscriptions revised: Jabal Ramm and Umm al-Jimāl". *Journal of the American Oriental Society* 108.369–378.

——. 1990. "Arabic verses from the first/second century: The inscription of ʿEn ʿAvdat". *Journal of Semitic Studies* 35.73–79.

Blau, Joshua. 1981. "The state of research in the field of the linguistic study of Middle Arabic". *Arabica* 28.187–203.

Bordreuil, Pierre, Alain Desreumaux, Christian J. Robin, and Javier Teixidor. 1997. "205. Linteau inscrit: AO 4083". *Arabie heureuse, Arabie déserte: Les antiquités arabiques du Musée du Louvre*, ed. Yves Calvet and Christian J. Robin, 265–269. Paris: Éditions de la Réunion des Musées Nationaux.

Caskel, Werner. 1954. *Lihyan und Lihyanisch*. Cologne: Westdeutscher Verlag.

Combe, Étienne, Jean Sauvaget, and Gaston Wiet. 1931. *Répertoire chronologique d'épigraphie arabe*, I. Cairo: Imprimerie de l'Institut Français d'Archéologie Orientale.

Contini, Ricardo. 1987 [1990]. "Il Hawrān preislamico: Ipotesi di storia linguistica". *Felix Ravenna* Fascicolo 1/2 –1987 (133–134).25–79.

Daniel, Robert W. 2001. "P.Petra Inv. 10 and its Arabic". *Atti del XXII Congresso Internazionale di Papirologia, Firenze, 23–29 agosto 1998*, ed. Isabella Andorlini, Guido Bastianini, Manfredo Manfredi, and Giovanna Menci, 331–341. Florence: Istituto Papirologico "G. Vitelli".

Ghoneim, Wafiq. 1980. "Saudi-Arabien". *Archiv für Orientforschung* 27.317–324.

Ghul, Omar al-. 2004. "An early Arabic inscription from Petra carrying diacritical marks". *Syria* 82.105–118.

Graf, David F. and Michael J. Zwettler. 2004. "The North Arabian 'Thamudic E' inscription from Uraynibah West." *Bulletin of the American Schools of Oriental Research* 335.53–89.

Greenfield, Jonas C. 1992. "Some Arabic loanwords in the Aramaic and Nabatean texts from Naḥal Ḥever". *Jerusalem Studies in Arabic and Islam* 15.10–21.

Grimme, Hubert. 1936. "À propos de quelques graffites du temple de Ramm". *Revue Biblique* 45.90–95.

Grohmann, Adolph. 1967–1971. *Arabische Paläographie*. 2 vols. Vienna: Hermann Böhlaus.

Halper, B. 1924–1925. "Recent Orientalia and Judaica". *Jewish Quarterly Review* 15.241–283.

Hamad, Muntasir al-. 2005. *Nabataean syntax, with special reference to other Aramaic dialects and Arabic*. Ph.D. diss., University of Manchester.

Healey, John F. 2002. "Nabataeao-Arabic: Jaussen-Savignac nab. 17 and 18". *Studies on Arabia in honour of Professor G. Rex Smith*, ed. John F. Healey and Venetia Porter, 81–90. Oxford: Oxford University Press.

—— and G. Rex Smith. 1989. "Jaussen-Savignac 17: The earliest dated Arabic document (A.D. 267)". *Atlal: The Journal of Saudi Arabian Archaeology* 12.77–84 (Pl. 46).

Hopkins, Simon. In the press. "Judaeo-Arabic inscriptions from northern Arabia". *Semitica in memoriam Alexandri*, ed. Werner Arnold, Michael Jursa, Walter W. Müller, and Stephan Prochazka. Wiesbaden: O. Harrassowitz.

Hoyland, Robert G. 2001. *Arabia and the Arabs from the Bronze Age to the coming of Islam*. London and New York: Routledge.

Jamme, Albert. 1967. "New Ḥasaean and Sabaean inscriptions from Saudi Arabia". *Oriens Antiquus* 6.181–187 (Pls. 47–49).

——. 1970. "The Pre-Islamic inscriptions of the Riyâdh Museum". *Oriens Antiquus* 9.115–139.

Jaussen, Antonin and Raphaël Savignac. 1909–1922. *Mission archéologique en Arabie*. 5 vols. Paris: Ernest Leroux and Paul Geuthner. (Repr., Cairo: Institut Français d'Archéologie Orientale, 1997).

Knauf, Ernst A. 1984. "Bemerkungen zur frühen Geschichte der arabischen Orthographie. I. Eine übersehene frühhocharabische Inschrift vom Nordostrand des Toten Meeres". *Orientalia* 53.456–457.

Kropp, Manfred. 1992. "The inscription Ghoneim *AfO* 27, 1980. Abb. 10: A fortunate error". *Proceedings of the Seminar for Arabian Studies* 22.55–67.

——. 1993. "Vassal, neither of Rome nor of Persia: Marʾ-al Qays the Great King of the Arabs". *Proceedings of the Seminar for Arabian Studies* 23.63–93.

——. 1994. "A puzzle of Old Arabic tenses and syntax: The inscription of ʿEn ʿAvdat". *Proceedings of the Seminar for Arabian Studies* 24.165–174 (Pl. 1).

Lacerenza, Giancarlo. 2000. “Appunti sull’iscrizione nabateo-araba di ʿAyn ʿAvdat”. *Studi Epigrafici e Linguistici sul Vicino Oriente Antico* 17.105–114.

Levine, Baruch A. 2000. “The various workings of the Aramaic legal tradition: Jews and Nabataeans in the Naḥal Ḥever archive”. *The Dead Sea Scrolls fifty years after their discovery*, ed. Lawrence H. Schiffman, Emanuel Tov, and James C. VanderKam, 836–851. Jerusalem: Israel Exploration Society and The Shrine of the Book, Israel Museum.

Littmann, Enno. 1949. *Arabic inscriptions.* (= *Syria. Publications of the Princeton University Archaeological Expeditions to Syria in 1904–1905 and 1909*, Division IV. Section D.) Leiden: E.J. Brill.

——. and David Meredith. 1953. “Nabataean inscriptions from Egypt, I”. *Bulletin of the School of Oriental and African Studies* 15.1–28 (Pls. 1–7).

Livingstone, Alasdair. 1997. “An early attestation of the Arabic definite article”. *Journal of Semitic Studies* 42.259–261.

Macdonald, Michael C.A. 1995. “Ṣafaitic”. *Encyclopaedia of Islam*, VIII, 760–762. 2nd ed. Leiden: E.J. Brill.

——. 1997. “Trade routes and trade goods at the northern end of the ‘Incense Road’ in the first millennium B.C.”. *Profumi d’Arabia. Atti del Convegno*, ed. Alessandra Avanzini, 333–349. Rome: “L’Erma” di Bretschneider.

——. 2000. “Reflections on the linguistic map of pre-Islamic Arabia”. *Arabian Archaeology and Epigraphy* 11.28–79.

——. 2003. “’Les arabes en Syrie’ or ‘la pénétration des arabes en Syrie’: A question of perceptions?”. *La Syrie hellénistique* (= *Topoi* Supplément, 4.), 303–318. Paris: de Boccard.

——. 2004. “Ancient North Arabian”. *Cambridge encyclopedia of the world’s ancient languages*, ed. Roger D. Woodard, 488–533. Cambridge: Cambridge University Press.

——. 2005. “Literacy in an oral environment”. *Writing and Ancient Near Eastern society: Papers in honour of Alan R. Millard*, ed. Piotr Bienkowski, Christopher B. Mee, and Elizabeth A. Slater. (= *Journal for the Study of the Old Testament*, Supp. Series 426.) New York and London: T&T Clark.

——. In the press. “Towards a re-assessment of the Ancient North Arabian alphabets used in the oasis of al-ʿUlā”. *Epigraphik und Archäologie des antiken Südarabien*, ed. Stefan Weninger. Wiesbaden: O. Harrassowitz.

——. In preparation. *Old Arabic and its legacy in the later language: Texts, linguistic features, scripts and letter-orders.*

—— and Geraldine M.H. King. 2000. “Thamudic”. *Encyclopaedia of Islam*, X, 436–438. 2nd ed. Leiden: E.J. Brill.

Mascitelli, Danieli. 2006. *L’arabo in epoca preislamica: Formazione di una lingua.* Rome: “L’Erma” di Bretschneider.

Mavroudi, Maria. In the press. “Arabic words in Greek letters. The Violet fragment and more.” *Proceedings of the First International Symposium on Middle Arabic and Mixed Arabic Throughout History, held in Louvain-la-Neuve, 11–14 May 2004*, ed. Jacques Grand’Henry and Jérôme Lentin (with Simon Hopkins and Han den Heijer). Louvain: Institut Orientaliste.

Milik, Joseph T. 1970. “Inscriptions hebraïques”. *Ancient records from North Arabia*, ed. Fred V. Winnett and William L. Reed, 163. Toronto: Toronto University Press.

Millar, Fergus G.B. 1998. “Il ruolo delle lingue semitiche nel vicino oriente tardo-romano (V–VI secolo)”. *Mediterraneo Antico* 1.71–94.

Müller, Walter W. 1982. “Das Altarabische der Inschriften aus vorislamischer Zeit”. *Grundriß der Arabischen Philologie.* I. *Sprachwissenschaft*, ed. Wolfdietrich Fischer, 30–36. Wiesbaden: L. Reichert.

Negev, Avraham. 1977. *The inscriptions of Wadi Haggag, Sinai.* Jerusalem: Institute of Archaeology, Hebrew University of Jerusalem.

——, Joseph Naveh, and Saul Shaked. 1986. “Obodas the God”. *Israel Exploration Journal* 36.56–60.

Nehmé, Laïla. 2005. “Towards an understanding of the urban space of Madāʾin Ṣāliḥ, ancient Ḥegrā, through epigraphic evidence”. *Proceedings of the Seminar for Arabian Studies* 35.155–175.

Noja, Sergio. 1979. “Testimonianze epigrafiche di Giudei nell’Arabia settentrionale”. *Bibbia e Oriente* 122.283–316.

——. 1989. “Über die älteste arabische Inschrift, die vor kurzem entdeckt wurde”. *Studia Semitica necnon Iranica Rudolpho Macuch septuagenario ab amicis et discipulis dedicata*, ed. Maria Macuch, Christa Müller-Kessler, and Bert G. Fragner, 187–194. Wiesbaden: O. Harrassowitz.

O’Connor, Michael. 1986. “The Arabic loanwords in Nabataean Aramaic”. *Journal of Near Eastern Studies* 45.213–229.

Puech, Émile. 1998. “Inscriptions araméennes du Golfe: Failaka, Qalaʿat al-Baḥreïn et Mulayḥa (ÉAU)”. *Transeuphratène* 16.31–55.

Rabin, Chaim. 1951. *Ancient West-Arabian.* London: Taylor’s Foreign Press.

——. 1955. “The beginnings of Classical Arabic”. *Studia Islamica* 4.19–37.

Rabinowitz, Isaac. 1956. “Aramaic inscriptions of the fifth Century B.C.E. from a North-Arab shrine in Egypt”. *Journal of Near Eastern Studies* 15.1–9.

——. 1959. “Another Aramaic record of the North-Arabian goddess Han-ʾIlat”. *Journal of Near Eastern Studies* 18.154–155.

Robin, Christian J. 1991. “L’Arabie antique de Karibʾîl à Mahomet: Nouvelles données sur l’histoire des arabes grâce aux inscriptions”. *Revue du Monde Musulman et de la Méditerranée* 61:3.9–166.

——. 1992. *Inventaire des Inscriptions Sudarabiques.* I. *Inabbaʾ, Haram, al-Kāfir, Kamna et al-Ḥarāshif.* Paris: de Boccard; Rome: Herder.

——. 2001. “Les inscriptions de l’Arabie antique et les études arabes”. *Arabica* 48.509–577.

——. 2006. “La réforme de l’écriture arabe à l’époque du califat médinois”. *Mélanges de l’Université Saint-Joseph* 59.319–364.

——. and Maria Gorea. 2002. “Un réexamen de l’inscription arabe préislamique du Ǧabal Usays (528–529 è. chr.)”. *Arabica* 49.505–510.

Ryckmans, Jacques. 1980. "ʿUzzā et Lāt dans les inscriptions sud-arabes: À propos de deux amulettes méconnues". *Journal of Semitic Studies* 25.193–204.

Teixidor, Javier. 1992. "Une inscription araméenne provenant de l'Émirat de Sharjah (Émirats Arabes Unis)". *Comptes Rendus de l'Académie des Inscriptions et Belles-Lettres*, 695–707.

Theeb [al-Ḏīyīb], Sulayman A. al-. 2002. *Nuqūš Jabal ʾUmm Jaḏāyiḏ an-nabaṭiyya: Dirāsa taḥlīliyya.* Riyadh: Maktabat al-Malik Fahd al-Waṭaniyya.

Versteegh, Kees. 1997. *The Arabic language.* Edinburgh: Edinburgh University Press.

Violet, Bruno. 1901. "Ein zweisprachiges Psalmfragment aus Damascus". *Orientalistische Literaturzeitung* 4.384–403, 425–441, 475–488.

Yadin, Yigael, Jonas C. Greenfield, Ada Yardeni, and Baruch A. Levine (eds). 2002. *The documents from the Bar Kokhba period in the Cave of Letters: Hebrew, Aramaic and Nabatean-Aramaic papyri.* Jerusalem: Israel Exploration Society and Institute of Archaeology, Hebrew University and Shrine of the Book, Israel Museum.

Zwettler, Michael. 1993. "Imra'alqays, son of ʿAmr: King of…???". *Literary heritage of Classical Islam: Arabic and Islamic studies in honor of James A. Bellamy*, ed. Mustansir Mir and Jarl E. Fossum, 3–37 (Pls. 1–5). Princeton, N.J.: Darwin Press.

——. 2006. "Binding on the crown". *Proceedings of the Seminar for Arabian Studies* 36.87–99.

——. In preparation. *The Namāra inscription.*

Michael C.A. Macdonald
(University of Oxford)

Fig. 1. The Tomb inscription of ʿIgl ibn Hafʿam at Qaryat al-Fāw, Saudi Arabia. The language is Old Arabic written in the Sabaic alphabet. (National Museum, Riyadh. Reproduced by kind permission of the Deputy Ministry of Antiquities and Museums, Kingdom of Saudi Arabia, and of the excavator, Professor ʿAbd ar-Raḥmān al-ʾAnṣārī).

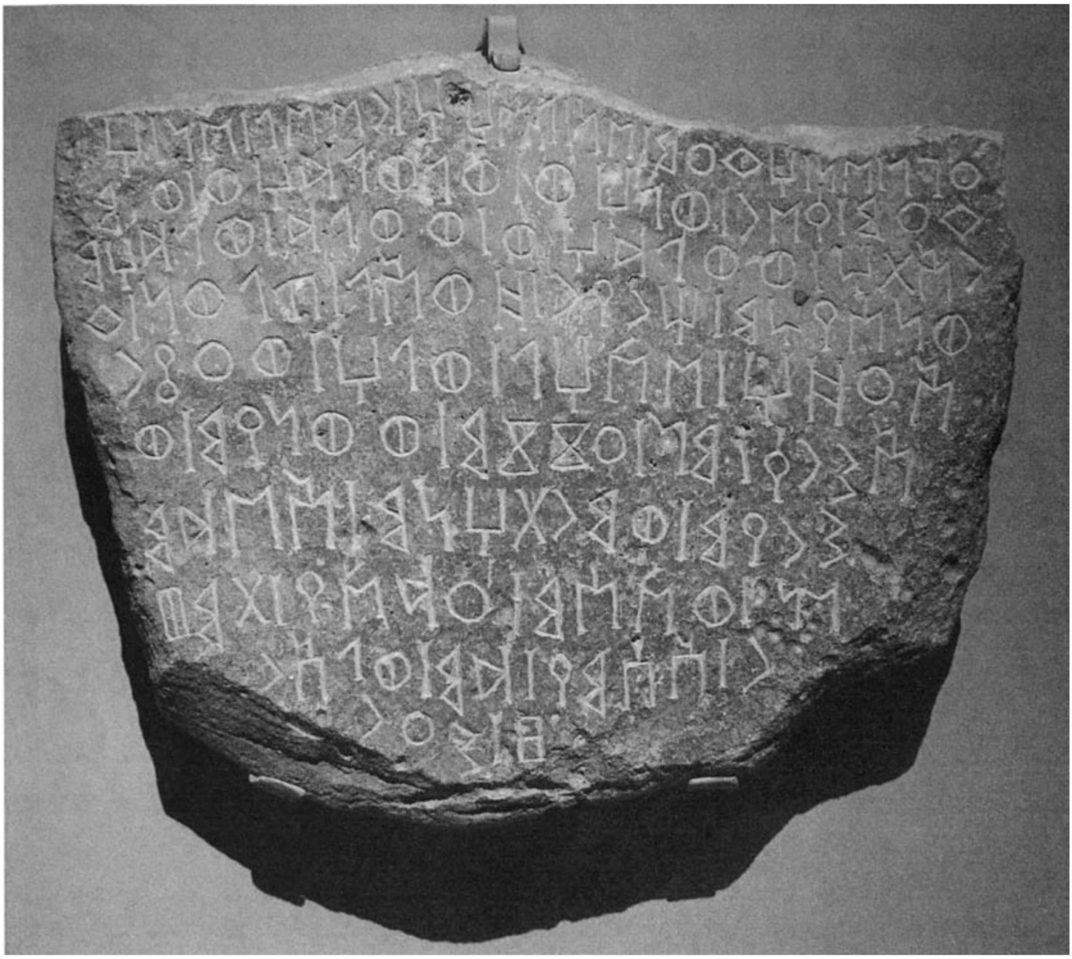


Fig. 2. The epitaph of king Mar' al-Qays ibn 'Amr from an-Namāra, southern Syria. The language is Old Arabic written in the Nabataean script. (Musée du Louvre A.O. 4083, reproduced by kind permission of the Musée du Louvre).

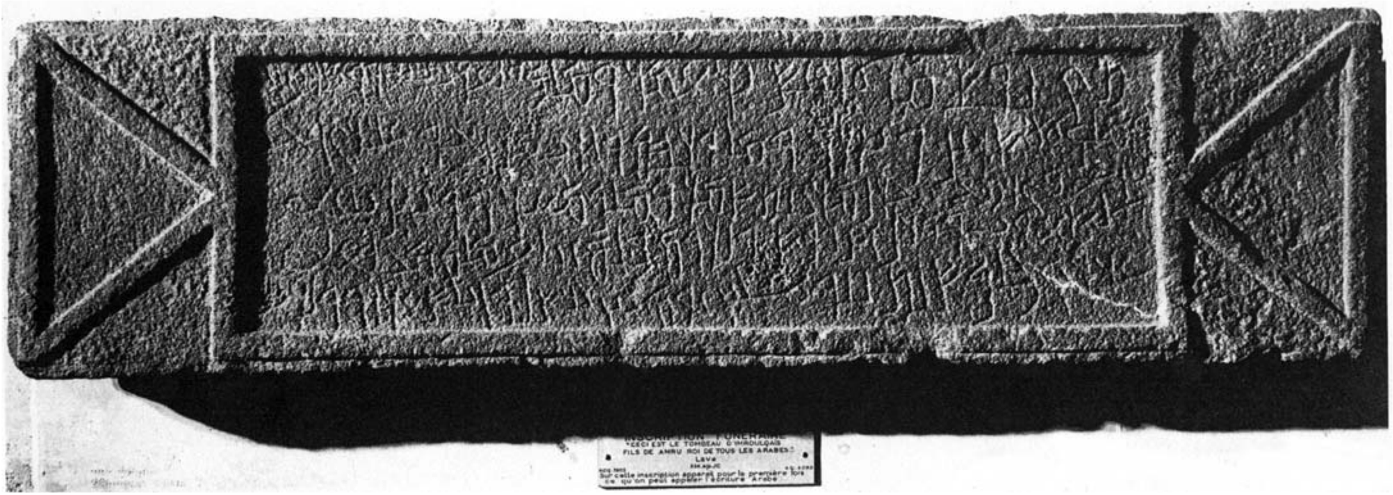


Fig. 3. The most recent facsimile of the Namāra inscription, drawn by Alain Desreumaux. (Reproduced, by kind permission of the authors, from *Arabie heureuse. Arabie déserte: Les antiquités arabiques du Musée du Louvre* by Yves Calvet and Christian Robin. Paris, 1997:265).

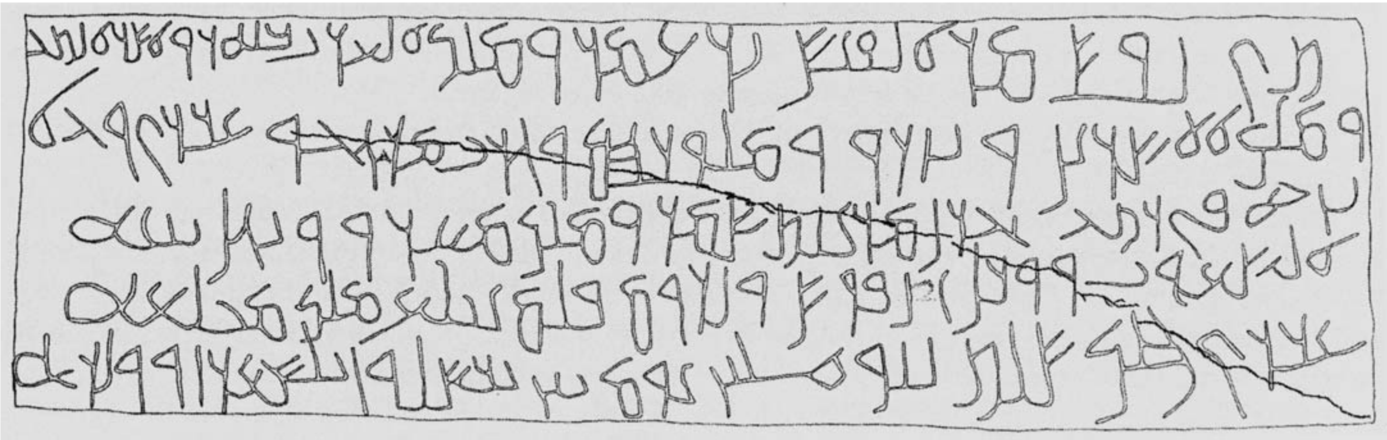


Fig. 4. A graffito of 528 C.E. at Jabal ʾUsays (Says). Old Arabic written in what is already recognizable as the Arabic script. (Photograph reproduced by kind permission of Robert Hoyland).

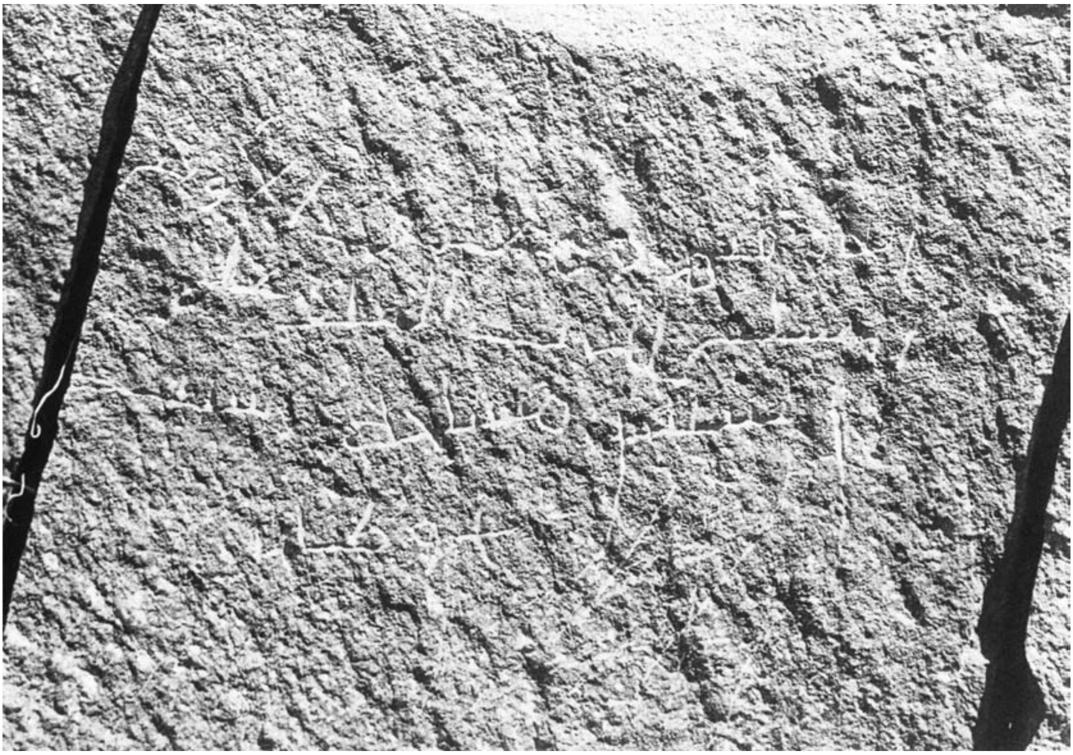
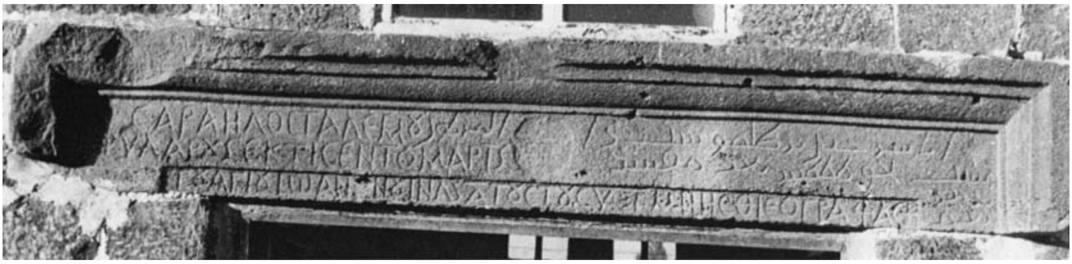


Fig. 5. The dedication of a martyrium at Ḥarrān, in the Lejā, southern Syria, in 568 C.E. A bilingual text in Greek and in Old Arabic written in what is recognizably the Arabic script. (Photograph reproduced by kind permission of Christian Julien Robin).



Old Church Slavonic → Slavonic languages

Omani Arabic

1. GENERAL

1.1 *Area*

The Sultanate of Oman is situated in the southeast corner of the Arabian Peninsula. In the north, a land corridor belonging to the United Arab Emirates divides the main land area from the Omani enclave of Musandam. Maṣīra and the Kuria Muria islands (now known as the Hallāniyyāt), both a few miles off the southern coast, also belong to Oman. Northern Oman is very mountainous, and the Bāṭina coastal strip runs between the northern mountains and the sea for approximately 300 km. Ṣalāla on the Indian Ocean is the main town of Ẓufār, the southern province of the country. The population of Oman was ca. 2.9 million in 2003, but this figure includes many foreign workers.

1.2 *Society*

Until the coup d'état that brought the present sultan to power in 1970, many families of Omani extraction lived outside the country, in particular in East Africa. With the change of government, there were many Omani 'returners' from East Africa who spoke Swahili as their first language. In the 1970s, the government made a major effort to reintegrate this group into Oman society: an adult Arabic-teaching program was launched, backed by UNESCO. In Muscat, Mutrah, and other coastal towns, there have for centuries been South Asian elements in the Omani population, in particular Baluchis. These elements still speak the languages of their ancestral communities, as well as Arabic. In Musandam, there are tribes, collectively known as Šiḥūḥ, who speak a variety of dialects (Jayakar 1902), some of which (e.g. Kumzārī) appear to be (Old-)Persian-based (Thomas 1930). The southern province of Ẓufār contains several tiny communities of speakers of the so-called → Modern South Arabian languages (Ḥarsūsī, Jibālī, Mehrī), although almost all those who speak them also speak Arabic. Smart (1990) provides data on the widespread → pidginization of Gulf Arabic that has occurred in the region as a whole as a consequence of the recent immigration of *Gastarbeiter* from the Indian subcontinent.

1.3 *Regional context*

Typologically, the dialects of Oman are divisible into two groups: the dialects of the mountainous interior, which are 'sedentary', and those of the deserts to the west, south, and southeast of the mountains, which are of the Bedouin type. The Bāṭina coast is a 'mixed' area where both types of dialect are encountered. The western Bedouin dialects are similar to those of southern Najd (Ingham 1986; → Najdi Arabic) and those of the northwest and southeast to the eastern Arabian dialects of the Gulf Coast. The sedentary dialects are similar to those described by Landberg (1920–1942) for the Ḥaḍramawt (→ Wādī Ḥaḍramawt Arabic), and by Holes for the Baḥārna of Bahrain (→ Bahraini Arabic). However, there is a bundle of features which virtually *all* Omani dialects, sedentary or Bedouin, share and which distinguishes them, as a group, from the dialects of neighboring areas:

(a) The interdentals /ṯ, ḏ ḏ̣/ are retained.
(b) Absence of the *ghawa* syndrome (→ *gahawa*-syndrome), except in the northwest Bedouin dialects.
(c) 2nd person feminine singular object/possessive suffix is *-iš*, except in the northwest Bedouin dialects (*-ič*) and the dialect of the Āl-Wahība (palatalized *-ik*; → *kaškaša*).
(d) There is an obligatory *-in(n)-* infix between an active participle having verbal force and an object suffix.
(e) Feminine plural verbs, adjectives, and pronouns occur regularly.
(f) The 'internal' passive is common.

1.4 *State of current research*

It is a century or more since the writing of the four most detailed studies of Omani Arabic we have: Jayakar (1889), Reinhardt (1894), and Rhodokanakis (1908, 1911). Galloway (1977) produced a mimeographed outline study of northern Omani sedentary dialects; Brockett (1985) is a useful study of the agricultural vocabulary of Xābūra on the Bāṭina coast; Webster (1991) provides notes on the Bedouin dialect of the Āl-Wahība of southeastern Oman. Holes (1989, 1996) attempts a typological study of the Omani dialects, placing them in an areal context. The present entry is based on Holes' own still largely unpublished materials, gathered during a two-year stay in the country (1985–1987), mainly in the north.

2. LINGUISTIC DESCRIPTION

2.1 *Phonology*

2.1.1 Inventory

2.1.1.1 Consonants

Bracketed consonants are marginal, found either in formal, educated speech or in a few foreign borrowings.

Plosives: *b, t, d, ṭ, (ḍ), k, g, q, (ʾ)*
Affricates: *č, j*
Fricatives: *f, (v), ṯ, ḏ, ḏ̣, x, ġ, ḥ, ʿ, h*
Sibilants: *s, ṣ, z, š*
Laterals and vibrants: *l, r*
Nasals: *m, n*
Semivowels/glides: *w, y*

Distribution:
All Omani dialects have the interdental fricatives as in Old Arabic, /ḏ̣/ being the normal realization of words which in Old Arabic had /ḍ/.

The most common sedentary Omani Arabic reflex of Old Arabic /j/ is a velar stop /g/. In the Bedouin dialects of the west and southeast, it is a glide, /y/. In parts of the Šarqiyya and in some parts of the Empty Quarter, it is an alveolar /j/. In this feature, as in so much else, the Bāṭina is a mixed area.

Old Arabic /k/ remained (a) a velar stop in the Capital Area and most of the Bāṭina, but was (b) palatalized in some mountain village dialects and affricated unconditionally to /č/ in others; and (c) affricated to /č/ in front vowel environments only in the Bedouin dialects on the border with the United Arab Emirates, in the northern Bāṭina, and in the port of Ṣūr and its hinterland, but not in the deserts of west and southeast Oman, where it is /k/.

Old Arabic /q/ developed as follows: (a) The Capital Area, most of the Bāṭina coast, the bigger towns on the mountain fringes, and some mountain villages have /q/; (b) some villages on the western and southern slopes of the Jabal Axḍar – the same ones that palatalize Old Arabic /k/ or unconditionally affricate it to /č/ – have /k/ < Old Arabic /q/; (c) the Bedouin dialects of the northern Bāṭina have /g/ with some evidence of affrication to /j/, as in the dialects of the neighboring United Arab Emirates; (d) the other Bedouin dialects – of the west, south, and southeast, including the ports of Ṣūr and Ṣalāla – have /g/, but, in the case of Ṣūr only, with some affrication to /j/.

Old Arabic /ʾ/ disappeared initially, e.g. *akal* 'he ate', or was replaced by /w/, e.g. *waxxar* 'he moved out of the way', or by /y/, e.g. *yāl bu saʿīd* 'ʾĀl Bu Saʿīd' [name of the ruling family]'; medially, it was mostly replaced by vowel length, e.g. *yākil~yūkil* 'he eats', or lost, e.g. *yinnis* 'it feels happy' < *yuʾannis* (Bedouin Empty Quarter example), but survives in a few words, e.g. *mitʾānis* 'healthy' (Bāṭina example); and finally it was either lost or replaced by /w/, e.g. *ḏ̣aww* 'fire', or by /y/, e.g. *qalāy* 'cooking'. In some Jabal Axḍar mountain dialects, however, /ʾ/ survives in particular roots, e.g. *l-əʾmūr* 'the matters, affairs', *yuʾmur* 'he ordered'.

/l/ and /r/, and /l/ and /n/ are for some (sedentary) speakers interchangeable in many words, e.g. *sulṭān~sirṭān* 'sultan', even *rəha~liha* 'for her' (Ṣūr); *dandal~daldal* 'it rocked, swung [something]'.

In a few foreign borrowings, /p/ has become /f/ or even /v/, a sound not otherwise heard, e.g. *firtugāl~virtugāl* 'orange' < *portugal*.

2.1.1.2 Vowels and diphthongs

The values given here are for the sedentary dialects of the north.

There are three short vowels, /a, i, u/, and five long /ā, ē, ī, ō, ū/.

/a/ and /ā/ are realized as [æ] or [ɛ] where gutturals (excluding /h/, but including /q/) and the emphatics are absent, e.g. [kælb] 'dog'; as [ə] in guttural environments including /q/, e.g. [qalam] 'pen'; as [ɑ] with an emphatic, e.g. [t̴ɑɫɑb̴] 'request'. There is, however, some unpredictable variation between front and back realizations of /ā/, e.g. [sæːr] and [saːr] 'he went'. Medial /i/ is retracted, e.g. [gɪld] 'skin, leather'; in final position it is closer and more front, e.g. [bɪnti] 'my daughter'; before /b, m, f, r, q/ and the emphatics, particularly when these are in final position, it is backed and (with the labials) rounded, e.g. [zaːhʊb] 'ready';

with the gutturals it is lowered and centralized, e.g. [jħəbːo] ‘he likes him’. /ī/ is a close, front vowel, but with the gutturals it is lowered, e.g. [jŧeːħ] ‘he falls’, and with the emphatics centralized, [bɨːs] ‘keel of a boat’. /u/ and /ū/ are back and round, e.g. [mʊrr] ‘bitter’, [sænnʊːra] ‘cat’, but lowered in emphatic environments, e.g. [ꞩɔbb] ‘pour!’. /ē/ and /ō/ correspond to the Old Arabic diphthongs /ay, aw/ and occur medially, e.g. *bēt* ‘house’, *bōš* ‘camels’.

2.1.1.3 Syllables types

Basic syllable types on which there is no restriction as to position:

Cv:
katab ‘he wrote’ (Cv-CvC)
CvC:
kambal ‘blanket’ (CvC-CvC)
CvCC:
šuftkum ‘I saw you’ (CvCC-CvC)
Cv̄:
gālis ‘sitting’ (Cv̄-CvC)
Cv̄C:
xāz ‘he moved out of the way’

The above are the basic types. Cv̄CC also occurs, but only in one type of form (the active participle of geminate verbs), e.g. *mādd* ‘extending’. The following types also frequently arise in word-initial position only as a result of the deletion of short unstressed vowels (see 2.1.3.1) and via borrowings:

CCv:
štarēt ‘I bought’ (CCv-Cv̄C)
CCv̄:
brābur ‘correct, right’ (CCv̄-CvC)
CCvC:
glast ‘I sat’
CCv̄C:
smīt ‘cement’ (borrowing)

When words ending in a Cv̄C have a consonant-initial suffix, no epenthetic vowel is inserted, e.g. *ṣōġha* ‘her jewelry’ (but some Jabal Axḍar dialects *sūqəna* ‘our market’). In many types of trisyllabic CvC-Cv-Cv(C) forms, in which stress falls on the penultimate syllable (see 2.1.1.5), there is a strong tendency to resyllabify to Cv-CvC-Cv, e.g. *yigílsu* ‘they sit’, *yišítġil* ‘he works’, the non-final short vowel often then being deleted, e.g. *ygilsu*, *yšitġil*. Suffixed CvC-Cv forms are transformed in the same way in all dialects, e.g. *ḥurma* ‘woman, wife’, *ḥurúmti~ḥrúmti* ‘my wife’.

2.1.1.4 Consonant clusters

CCC clusters

Where a CCC cluster is a consequence of the suffixation of doubled verbs, some speakers insert an epenthetic schwa, e.g. *šaqqəha* ‘he tore it’. Otherwise, CCC clusters are normally stable, e.g. *bintkin* ‘your [fem. pl.] daughter. The /l/ of *qult* in phrases of the *qult l*-type is often elided, e.g. *qut lak* ‘I told you [masc. sq.]’.

Initial CC- clusters arise frequently. Medial geminate clusters in forms of the CvC-CvCv type reduce the cluster, e.g. *tkalmu* ‘they spoke’ < *takallamu.*

The treatment of nondoubled final clusters in words of the structure CvCC in Old Arabic depends on the preceding vowel and the consonants:

(a) Old Arabic CaCC
   i. C_2 is a /l, n, r/: The form is stable, e.g. *qalb* ‘heart’, *gamb* (< *ganb*) ‘side’.
   ii. C_2 is a guttural: CaCaC is normal, e.g. *baḥar* ‘sea’.
   iii. In other cases, there is a strong tendency toward CaCvC, in which the v is a mid vowel with an [ə] coloring, or an [u] coloring in the presence of labials, e.g. *akil* ‘food’, *nahub* ‘robbery’. In such forms, some speakers delete the first vowel of the form, which is now in open syllable: *wgit* ‘time’ as in *l-əwgit hāḏi* ‘at that time’ (Ṣūr).

(b) Old Arabic CiCC and CuCC
   i. C_2 is /l, n, r/: The form is usually stable, e.g. *ḥilf* ‘oath’, *rinz* ‘rice’, *qirš* ‘Maria Theresa dollar’ (although in some Bedouin dialects, forms like *miliḥ* ‘salt’ occur).
   ii. In other cases, the form is usually stable, but if C_3 is /l, n, r/, some speakers have a CvCvC form with vowel harmony, in which the first vowel, now in open syllable, is liable to deletion, e.g. *šuġul~šġul* ‘work’, *gibin~gbin* ‘cheese’ (cf. the B dialects of Bahrain) with a shift of stress, e.g. *l-əḏ̣hur*.

2.1.1.5 Stress

The rules are:

i. Stress the long syllable (CV̄C, CvCC) in a word, e.g. *gā́lsa* 'sitting [fem. sg.]', *galást* 'I sat down'; if there are two, stress the second, *šāyfī́n* 'seeing [masc. pl.]'.
ii. If there are no long syllables, stress the penultimate syllable, except in cases of Cv-Cv-Cv, when the antepenultimate is stressed, e.g. *šárika*. Such forms are almost invariably reduced, however, to two-syllable CvC-Cv structures by the deletion of the second short vowel, viz. *šárka*. Stress is nondistinctive.

2.1.2 Phonotactics

2.1.2.1 Assimilation

The following are typical:

(a) The definite article: sun letters and moon letters behave as in Old Arabic.
(b) The *-n* of the preposition *min* before the definite article assimilates to the /l/ of the definite article (or the replacing sun letter), e.g. *mi lbēt, mi ššams*.
(c) The *-la* of the preposition *ʿala* is similarly assimilated, e.g. *ʿa ssāḥil*.
(d) The *b*-verbal prefix of proximal intent becomes /m/ before the *n*- of the 2nd person plural imperfect in some dialects, e.g. *minṣanʿuh* 'we make it' (Bahla example; cf. Bahrain B dialects).
(e) /ġ/ > /b/ in the imperfect forms of the verb *baġa* 'to want, need' *aba, tba, yba*, etc.
(f) /t/ in *ti*-verbal personal prefixes and the prefixed *t*- of Forms V and VI is assimilated to /t, ṭ, ṯ, d, ḏ, ḏ̣, č, j, s, ṣ, š/ as a secondary consequence of the deletion of unstressed short vowels in open syllable (see 2.1.3), e.g. *(i)ddūr* 'you [masc. sg.] search'. Doubling of *tāʾ marbūṭa* via the assimilation of the /h/ of suffixes is heard in some dialects, e.g. *qīmatta* 'its value' (< *qīmat* + *ha*) (Bāṭina example; cf. Bahraini A and Kuwait dialects; see also 2.1.3.6 (a) below).
(g) The forms *dōk, dōš, dōkum, dōkin* 'here you are!' may have arisen via the loss of intervocalic *n*- (cf. Old Arabic *dūnak*, etc., with the same meaning, although Reinhardt [1894:31] derives them from *ḏa* 'this').

/š/ is replaced by /h/ in some words, e.g. *dīhdāša* < *dišdāša* 'man's long shirt', *xamstaʿhar* < *xamstaʿšar* 'fifteen', and /ʿ/ devoiced to /ḥ/ if a voiceless consonant is in contiguity, e.g. *yaḥṭi* 'he gives' < *yaʿṭi* (all Ṣūr examples). The /d/ of *wāḥid* is regularly lost in many dialects: *wāḥi* 'one'.

2.1.2.2 Dissimilation

There are a few cases, such as *xast* < *xass* 'lettuce', *rinz* < *ruzz* 'rice'.

2.1.2.3 Metathesis

Examples are plentiful, e.g. *yixtifar* < *yiftixar* 'he is proud', *tinquṭ* < *tinṭuq* 'she speaks', *xamaš~šamax* 'he scratched'.

2.1.2.4 *Ghawa* syndrome

The *ghawa* syndrome, which is common to all the Bedouin-descended dialects of eastern Arabia, does not occur in the Omani sedentary dialects (cf. the Bahrain B dialects). Individual forms like *hal* 'family' which do occur are probably not to be interpreted as evidence of the syndrome but are rather the product of the twin tendency of CaCC forms to become CaCaC where C2 is a guttural, and /ʾ/ to be lost when in initial position (cf. *xit* 'sister', *ḥad* 'someone', *nṯa* 'female'). The syndrome occurs irregularly in the Omani Bedouin dialects as a whole, but more consistently in the northern border areas near Buraymi and ʿIbrī.

2.1.2.5 Conditioned → *ʾimāla*

Along the Bāṭina coast (Bedouin-type dialects) and in the towns of the seaward side of the mountains (sedentary dialects), there is a strong tendency to raise final *-a*, when not in contiguity with the emphatics, to [e] or even [i], e.g. *mistašfi* (Rustāq), as opposed to the speech of those originating from the mountains, who have [a] in such forms. In some inland locations in the Šarqiyya (e.g. Muḍaybi), very pronounced *ʾimāla* of *ā* in nonemphatic contexts was noted, e.g. *mīlḥa* (< *mālḥa*).

2.1.2.6 Spread of velarization

Velarization may spread, especially to /l/, /r/, /b/, /m/, e.g. *ṣabi* 'boy' (= [s̴ɒb̴i]), *ṭabil* 'drum' (= [t̴ɒb̴ɨɫ]). In a few roots, an originally nonemphatic consonant has become velarized, e.g. *ḏ̣bāb* 'flies'. This is particularly common with /s/, e.g. *ṣaṭḥ* 'roof', *ṣumṭ* 'woman's neck ornament', *ṣalaṭ* 'to swallow', where the presence of /ṭ/ seems to have been the cause.

2.1.3 Morphophonology

2.1.3.1 Elision of vowels

Unstressed short vowels of whatever quality in open, non-final syllables are liable to deletion in all kinds of forms, for example:

(a) Before a long (stressed) syllable:
glast < Cv-CvCC ‘I sat’
klāb < Cv-Cv̄C ‘dogs’
kalmāt < Cv-Cv-Cv̄C ‘words’
kātbāt < Cv̄-Cv-Cv̄C ‘writing [fem. pl.]’
mbārkīn < Cv-Cv̄-Cv-Cv̄C ‘blessed [masc. pl.]’

(b) Before a short (stressed) closed syllable:
rgaʿna < Cv-CvC-Cv ‘we returned’
dzawgu < Cv-CvC-Cv-Cv ‘they [masc. pl.] got married’

(c) In a succession of more than two short syllables:
kalma < Cv-Cv-Cv ‘word’
rguba < Cv-Cv-Cv (Bedouin dialects) ‘neck’
galsit < Cv-Cv-CvC (sedentary dialects) ‘she sat’
ylasit~ylisat < Cv-Cv-CvC (Bedouin dialects) ‘she sat’
mdarsa < Cv-CvC-Cv (regularly < CvC-Cv-Cv) ‘school’

(d) In a succession of two short syllables:
Generally the vowels are stable, but some speakers have forms of the type: *ḏ̣hur* < Cv-CvC ‘noon’

2.1.3.2 Insertion of vowels

See above, 2.1.1.3–4.

2.1.3.3 Shortening and lengthening of vowels

Where two or more long vowels occur in a word, the unstressed first (and second, if there is one) long vowel may be shortened, especially in rapid speech, e.g. *šāyfīn* ‘seeing [masc. pl.]’ = [ʃæyfiːn]. Final vowels that are short in unsuffixed form are lengthened and stressed when suffixed, e.g. *yxalli* ‘he lets’, *yxallīni* ‘he lets me’.

2.1.3.4 Clitics

There are a number of clitics:

(a) In the sedentary dialects, a question particle *ə*, after vowels *hə* or *yə*, suffixed to the questioned word or phrase, e.g. *bass-ak-ə*? ‘have you had enough?’ *fī ʿōk ši-hə*? ‘is there any pain?’ (cf. the B dialects of Bahrain).

(b) Deictics *ḏa/ḏi* prefixed to the definite article, e.g. *ḏa lbēt* ‘this house’, *ḏi lbint* ‘this girl’.

(c) The prefixes *b-* and *ha-* (or *ḥa-*), indicating volition or proximate intent, e.g. *b-itkahyu*? ‘would you like coffee?’ (Jabal Axḍar example), *ha-nsīr bākir* ‘we’ll go tomorrow’.

(d) *sa-/ṣa-* prefixed to a suffixed pronoun in subject position, e.g. *min ṣā-h ixtam minna, ṣā-h yrūḥ iskūn* ‘when he’s finished with learning the *Qurʾān*, he goes to [government] school’.

(e) Presentative *ha-* attached to the independent pronouns, e.g. *ha-hiyya ḥrumti* ‘this is my wife’.

(f) Presentative *dō-* (some dialects *hā-*) attached to suffixed pronouns, e.g. *hēn qalmi*? *dō-k*! ‘where’s my pen? there it is!’, *hā-č ṯumurtuh*! ‘here’s its fruit!’ (Jabal Axḍar example).

(g) *-in*, a vestige of Old Arabic → *tanwīn*, in certain phrase types, e.g. *ilinsān kill-in bi naḏ̣rītah* ‘everyone is entitled to his own opinion’.

(h) *mā*, inserted between the repetition of a word to indicate ‘and such like’, e.g. *ṭawi mā-ṭawi* ‘wells and such like’, *ṭlūʿ mā-ṭlūʿ* ‘climbing [palm trees] and things like that’ (→ Bahraini Arabic).

2.1.3.5 Construct state

This is less used because of the development of an → analytic genitive (2.2.3.3), e.g. *gōniyyit ʿēš* ‘a sack of rice’. The form of the feminine suffix in the sedentary dialect is *-it*, e.g. *šwayyit fijil* ‘a few radishes’. Plurals and duals that occur as head noun normally retain their final *-n*, e.g. *saḥḥārtēn ṭamāṭ* ‘two boxes of tomatoes’.

2.1.3.6 Suffixation

(a) When a vowel-initial suffix is added to verb forms ending in *-it* and *-an* (2nd person feminine plural, 3rd person feminine singular and plural of the perfect, 2nd and 3rd person feminine plural of the imperfect), the final consonant is doubled (and the stress is thereby shifted), e.g. *ḏ̣arbit* + *uh* → *ḏ̣arbíttuh* ‘she hit him’, *ḏ̣arban* + *uh* → *ḏ̣arbánnuh* ‘they [fem.] hit him’.

(b) There is an obligatory *-in(n)-* infix between the active participle, singular and plural, and an object suffix (cf. the B dialects of Bahrain). In the singular, the *-n* is doubled if the suffix is consonant initial (cf. (a) above); e.g., for *kātib* + *uh/ha* 'having written it', we have:

masc. sg.	*kātbínnuh*	*kātbínha*
fem. sg.	*kātbitínnuh*	*kātbitínha*
masc. pl.	*kātbī́nnuh*	*kātbī́nha*
fem. pl.	*kātbātínnuh*	*kātbātínha*

Some speakers have alternative forms of the type *kātbitnuh, kātbitnha* for the feminine singular forms. Weak and hollow roots follow the same pattern, e.g. *bāninnuh/bāninha*, etc. Some verbs form their active participle on the pattern CiCCāC, but behave in the same way, e.g. *nisyān* 'forgetting', *nisyānínnuh* 'having forgotten it'. Note also *ʿilmān* 'knowing' and *ʿilmāníbbuh* 'knowing about it' (< *ʿilmān* + *buh*), where the doubling process affects the governed preposition (both these are mountain dialect examples). Where an active participle has nominal force, it lacks the infix, so *huwwa mʿalmínnuh* 'he has taught him', but *huwwa mʿálmuh* 'he is his teacher'.

(c) Some speakers also insert an *-n(n)-* infix between imperfect verbs more generally, not just in the feminine forms specified in (a) above, and suffixed pronouns, e.g. *yšillinnah* 'he removes it', *yḏ̣aġṭinhum ḏ̣aġṭ* 'he oppresses them' (western Bedouin examples).

2.2 *Morphology*

2.2.1 Pronouns

2.2.1.1 Personal pronouns

3rd masc. sg.	*huwwa, hu*	*-uh, -ah*
3rd fem. sg.	*hiyya, hi*	*-ha, -a*
3rd masc. pl.	*hum*	*-hum, -him*
3rd fem. pl.	*hin*	*-hin*
2nd masc. sg.	*inta*	*-ak, -ik, -ič*
2nd fem. sg.	*inti*	*-iš, -ik, -ič*
2nd masc. pl.	*intu*	*-kum, -čim*
2nd fem. pl.	*intan*	*-kin, -čin*
1st sg.	*ana*	*-i* (poss.), *-ni* (obj.)
1st pl.	*naḥan, iḥna,* *naḥna,* *ḥanna, ḥan*	*-na*

Notes:

i. 3rd person forms: *-ah* is a Bedouin form, *-uh* sedentary.

ii. 2nd person forms: Feminine singular *-ik* is confined to the Āl-Wahība in southeastern Oman, and feminine singular *-ič* to areas in the north which border the United Arab Emirates. Both are Bedouin. Masculine singular *-ič*, masculine plural *-čim*, and feminine plural *-čin* are found only in certain central Jabal Axḍar villages (cf. Bahraini B village dialects). Some speakers have *anta, anti*, etc. rather than the forms listed.

iii. 1st person forms: *ḥanna* and *ḥan* are Bedouin.

iv. The independent pronouns are frequently 'suffixed' to a verb or noun without any particular emphatic sense, e.g. *aḥīdana* 'I know' (< *aḥīd* + *ana*), *fi zamannāḥna* 'in our [= my] time' (< *zamān* + *na* + *ḥna*), *ʿindīna* 'at my house' (< *ʿindī* + *ana*).

2.2.1.2 Indirect object suffixes

Indirect objects are often expressed by prepositional complements involving *li* 'to', which form part of the phonological word, e.g. *kātibílli* 'he has written to me' (cf. *kātbínnuh* 'he has written it', with a direct object, *kātibúbbuh* 'he has written with it', with an instrumental object). With some verbs, the indirect object is suffixed directly to the verb and the direct object carried by the particle, e.g. *ʿaṭni iyyāh* 'give it to me!'; in most cases of such constructions, it is the direct object which is suffixed to the verb and the indirect object to *iyya*, e.g. *xarribōha iyyāy* 'they've ruined it for me'.

2.2.1.3 Demonstratives

Proximal masc. sg.	*(hā)ḏa*
Proximal fem. sg.	*(hā)ḏi*
Proximal pl.	*(hā)ḏēla*
Distal masc. sg.	*(hā)ḏāk*
Distal fem. sg.	*(hā)ḏīk*
Distal pl.	*(hā)ḏēlāk*

The *hā-* element may be omitted, particularly when the demonstratives are used adjectivally, e.g. *ḏa lbēt* 'this house', *ḏāk iṣṣōb* 'that direction'. There are some variants, e.g. *hāḏiya* (masc. sg.) in Ṣūr.

2.2.1.4 Presentatives

hā- and *dō-* are used with 2nd person suffixes, e.g. *hāč ṯumurtuh!* 'here's its fruit [for you]!'

(Jabal Axḍar example), *dōk grābak*! ‘here’s your [masc. sg.] date sack!’. The latter form can also be used with a double suffix, e.g. *dōkhum man ʿumrəhum sittaʿašar sana wa mā fōq*…‘there are those who are over sixteen years old…’. *ha-* may be prefixed to 3rd person suffixes, e.g. *ha-hum ūlādi* ‘these are my sons’. The Bedouin dialects, like those of Najd, use the particle *hummalē-*, to which suffixed pronouns are added, to depict a sudden event or scene: *wila jaha ššta, hummalēha fi maḥallha illi hi rābb bah* ‘and when winter comes, there she [= oryx] is, back in her own territory’ (southern Empty Quarter example).

2.2.1.5 Relative pronouns

bu, illi, illaḏi, and *il* are all in use, often by the same speaker, although *bu* appears to be the main sedentary form, e.g. *ləḥbūb bu tistaʿmilīhin* ‘the pills that you [fem.] are taking’, and is less common in Bedouin dialects.

2.2.1.6 Interrogative pronouns

mu, mhu, hēš, eš, wēš ‘what?’; *min, minhu* ‘who?’; *hēn* ‘where?’; *mata* ‘when?; *kēf* ‘how?’; *kam/kam min, šgadd* (Bāṭina) ‘how many?’; *ḥāl min, māl min* ‘whose?; *min wēš, ʿala mu, ḥāl hēš, ḥāl mu* ‘why?’; *hēn min/mu min* ‘which of?’. In parts of the Bāṭina, *min* for ‘what?’, e.g. *min ismak* ‘what’s your name?’

Interrogatives have no fixed sentence position; pragmatic factors determine the position.

2.2.2 Adverbs

Note that only unusual forms or local innovations are listed in this section. Some adverbs, e.g. *ʿād* ‘just, so, then, again’, have very many uses and are not easily classifiable (see Brockett 1985:25–32 for details).

2.2.2.1 Temporal

taw, alḥīn, fi ṭṭāri, (fi) l-əwgit hāḏi ‘now, these days’; *nōb/nōba* ‘sometimes, also’; *mā ʿād* ‘no longer’; *dōm* ‘always’; *killah* ‘constantly’; *baʿad* ‘still, yet, also’; *ṯinna* ‘then’; *mbōn* + pronominal suffixes ‘originally’; *min qabil* ‘beforehand’; *xlāf, ʿuqub* ‘afterward’; *āxir* ‘finally’; *marra* ‘straightaway’; *abdan, marra* (with negation) ‘never, not at all’; *rāyiḥ* ‘continuously’; *awgāt* ‘sometimes’; *awwal bādi* ‘first of all’; *bākir* ‘tomorrow’; *ʿuqub bākir* ‘the day after tomorrow’; *ilbārḥa* ‘last evening’; (*min*) *iṣṣubḥ* ‘in the early morning’; *iḍ̣ḍaḥa* ‘in the forenoon’; *ilqāyla* ‘in the early afternoon’; *ilʿaṣir* ‘in the late afternoon’; *lēliyya* ‘all night’; *nhāriyya* ‘all day’; *awwal* ‘in the old days’; *fi lġādi* ‘in the future’.

2.2.2.2 Local

ihni, hnāha ‘here’, *ihnāk* ‘there’, *minnīka* ‘over there’, *ḥadir* ‘below’, *dāyir madār* ‘around’, *sīda* ‘straight ahead, directly’, *ḏāk ṣṣōb* ‘over there’, *ġād/ġādi/aġādi* ‘yonder’, *ajāy* ‘over here’, *šāwir* ‘outside’, *ḥādir* ‘seaward’, *sānid* ‘inland’.

2.2.2.3 Manner

bass ‘only’; *kida, kidāha, kidāk, čiha* ‘thus, like this’; *hāḏa nnamūna* ‘like this’; *hast, wāgid/wāyid* ‘much, very’; *zēn, hūdār, brābar* ‘well, properly’; *killiš* ‘completely’; *yigi, gōb/yōb, taqdīr* ‘approximately’; (*kill*) *rbāʿa* ‘together’; *mūl/mūliyya* ‘at all’ (used only negatively); *zitāt* ‘quickly’; *didīh* ‘quickly, directly, on the spot’; *balāš* ‘in vain’; *ʿasa* + pronominal suffixes ‘maybe, perhaps’; *gins* ‘like, such as’; *ʿala gōrt*…(with following noun) ‘like’.

2.2.3 Particles

2.2.3.1 The definite article

The article is *(i)l-* or *(a)l-* and is assimilated by the sun letters as in Old Arabic. When prefixed to nouns whose first syllable is an open, unstressed /i/ or /u/, especially if the following vowel is long, it normally becomes *lə-*, e.g. *ləktāb*.

2.2.3.2 Indefinite article

Some uses of *wāḥi* and *ḥad* ‘one’ have the function of indefinite article when they precede the noun, e.g. *māt wāḥi šāyib, baʿad ḥad bāqi minhum* ‘if an old man dies, another one of them still remains’. *šay* ‘thing’ is also used in this way when used with objects, e.g. *šay yūšib, u šay taw ṯāmir* ‘some [fruits] fattening up, some just coming out’. Dialectal *tanwīn* also fulfils this function in some phrase types, often in a ‘generic’ sense, e.g. *rajjāl-in ʿāqil* ‘a/the wise man’.

2.2.3.3 Genitive markers

māl, ḥāl, ḥagg, and *bu* (see 2.2.1.5) are all in use, e.g. *innabāt māl ilfaḥḥāla* ‘the pollen of the male palm tree’, *ġašmar ḥāl ḥīwān* ‘sorghum [feed] for farm animals’, *issēḥ ḥaqq ilbadu* ‘the steppe of the Bedouin’, *ilwayr bu lkarhab* ‘electrical wire’.

2.2.3.4 Negative particles

With all verbs: *mā*, e.g. *mā rumt asīr* ‘I couldn’t go’. ‘To be no longer’ is *mā ʿād*, e.g. *sbūʿ waḥda bass…taklīfah mā ʿād* ‘just a week, and the pain was no longer there’. The particle *mā* is also used with prepositional expressions of possession (*ʿind, maʿa, il-*), e.g. *mā maʿāy šay fi lbēt* ‘I have nothing in the house’, and, in most sedentary dialects, with adjectives and other prepositional phrases as well, e.g. *mā azyad* ‘not more’. Bedouin dialects in the Bāṭina normally have *mu, muhu* in all these cases, and in the Šarqiyya one hears *māb, mā hub*, e.g. *ʿadan māb zēna alḥīn* ‘Aden is no good now’, *ilhubūb mā hub zēn*…‘if the wind wasn’t blowing well…’. *mā* is also used in a variety of set expressions to negate existence: *mā šay* (or *mā ši*), *mā fīh, mā miš*, e.g. *mā šay maṣlūḥ* ‘there was no benefit’; *mā fīh ḥīwānāt kaṯīr* ‘there aren’t many [farm] animals’; *sayyāra mā miš* ‘there were no cars’. In western and southern Bedouin dialects, *mā bah* is used, e.g. *mā bah ibil* ‘there were no camels’. *mā ḥad* is normally ‘no one’, e.g. *mā ḥad yuṭri māl šōr* ‘no one mentioned about consultation’. In nominal sentences, *mā* + independent pronoun or pronoun + *mu* are both possible, e.g. *inta mu(b) rayyāl* or *minta (bi) raggāl* ‘you’re not a [real] man’, the former construction being more typical of Bedouin dialects. *šīšī* is used as an emphatic negative of existence, e.g. *šišī byūt* ‘there were no houses at all’.

lā is used for negative imperatives and optatives, and as the second negative particle in all coordinated negations: *ʿan lā* = ‘lest’, e.g. *mā axallīha trūḥ ilmadrasa ʿan lā titqaḥḥab* ‘I don’t let her go to school for fear that she’ll become immoral’.

2.2.3.5 Particles to introduce questions

Like the Bahraini B dialects, the Omani dialects attach a clitic *ə* to any word or phrase to create a yes-no question (see 2.1.3.4 (a) above). *šay/ši* may also be placed before or after the element to be questioned, e.g. *šī ahlik bākīn*? ‘are your parents still alive?’, *ilamma tšīli ḥāga, igi lwagaʿ akṯar šay*? ‘when you lift something, does the pain get worse?’. Some speakers of mountain dialects attach *lā* to the end of sentences as a kind of tag question, or an attention-maintaining device, e.g. *yaʿmluh bi rigluh, yikanzuh, lā*? ‘he makes it with his foot, he compresses it [= dates], right?’ (cf. Bahraini B dialects).

2.2.3.6 Existentials

fīh and *šay* are most commonly used, e.g. *zājra šay mbōnha* ‘there were ox-walk wells, originally’ (Jabal Axḏar example), both negated by *mā*, e.g. *mā ši fīha taʿab* ‘there was no effort involved in it’. *hast* (as in Bahrain) is also occasionally heard, e.g. *hast šay mā zēn* ‘there were some not very good ones’.

2.2.3.7 Prepositions

li ‘to, for’, *bi* (or *əb*) ‘with, by means of’, *fi* (or *əf*) ‘in, on’, *ḥaqq* ‘to, for’, *ila* ‘to, toward’ *ʿala* ‘on, against’, *min* ‘from’, *ʿan* ‘away from, instead of, than [in comparisons]’, *fōq* ‘on top of, above’, *taḥt* ‘under, near’, *ḥadir* ‘under, below’, *ʿind* ‘at, with, in the possession of’, *maʿa* ‘with, in the possession of’, *wiyya* ‘with’, *ṣōb* ‘toward’, *wara, qafa* ‘behind’, *quddām* ‘in front of, opposite’, *mqābil* ‘opposite’, *bēn, mā bēn* ‘between’, *qabil* ‘before’, *baʿad, xalf, xlāf, ʿuqub* ‘after’, *bilā, bidūn* ‘without’, *māl, ḥāl, bu* ‘of, relating to, concerning’, *kimā/čima* ‘like’, *miṯil, šarwa* ‘like’, *yamm, bi gamb* ‘beside’, *dūn* ‘in contrast with, different from’, *min šān* ‘for the purpose of, for the sake of’, *dāyir madār* ‘around’. *li* and *min* form compound prepositions with several of these, indicating movement toward or away from, e.g. *li wara, min fōq*.

2.2.3.8 Conjunctions

wa/wi/u ‘and’; *willa, aw, yallah* ‘or’; *amma…willa/amma…amma* ‘either…or’; *lākin/lāčin* and *bass* ‘but’; *inn* ‘that’; *liʾann* ‘because’; *kinn/činn* ‘as if’; *ġēr innuh* ‘except that’; *b wāsṭit innah* ‘because of the fact that’; *yōm, lamma*, ‘when’; *min* ‘when, since’; *lēn, ilēn, lēma* ‘until’; *ḥatta, ʿašān* ‘so that’; *ligil* ‘because, so that’; *ʿan lā* ‘lest’; *qabil lā, qabil mā* ‘before’; *baʿad mā, xalf mā, ʿuqub mā* ‘after’; *kill mā* ‘the more…the more; whenever’; *miṯil mā* ‘like’; *mā dām/imdām* ‘as long as’. Conditional conjunctions: *in, iḏa, ila, lō, (in) kān/čān*.

2.2.3.9 Vocative particles

yā is used when addressing people by name.

2.2.3.10 Exclamations

Particles of affirmation and denial:

All of the following are used for ‘yes!’: *ī! ē! kē! ila! hi! naʿam! ē naʿam!*

‘No!’ is *la!* or *hō!*

yallah! ‘come on!’; *yā lēt!* ‘would that…!’; *bass!* ‘stop! enough!’; *zēn!* ‘OK!’; *ila*…‘lo and behold!'; *ʿād*, which has many uses, for example reproach, e.g. *walla ʿād!* = 'clear off, will you!', cajoling, e.g. *ismaʿ ʿād*! 'listen, why don't you?'; *ax!* is used to express pain; *afa!* is used to express disgust; *hūd hūd!* is said by someone wishing to enter a house. *š*- and *wēš* + noun are used to express admiration/surprise, e.g. *š-ḥalāwatha!* 'how pretty she is!'; *yā mā*…'how often!', e.g. *yā mā qit lak*…'how often have I told you…' Also *amma: amma badar hiyya!* 'what a pretty girl [lit 'full moon'] she is!'

2.2.4 Nouns

2.2.4.1 Gender

Feminine by usage: most double parts of the body; nouns denoting females, e.g. *bībī* 'wife'; most names of countries; a few common nouns, e.g. *ṭawi* 'well', *arḍ/arḏ̣* 'earth, land', *yōm* 'day' (some dialects).

2.2.4.2 Productive patterns

Common patterns are similar to those found in Bahrain and Kuwait.

A few nouns of local reference end in the suffix -ō, e.g. *ḥalwayō* 'a type of fish [jack pomfret (?)]'.

The *-iyya* suffix is highly productive, e.g. *nhāriyya* 'the whole day, by day', *ḥārriyya* 'the hot part of the summer', *muxxiyya* 'brains, intelligence' (this form is also common in the Bahraini B dialects).

2.2.4.3 External and internal plural

Common patterns are similar to those found in Bahrain and Kuwait.

Examples of anomalous or rare patterns:

sinīn < *sana* 'year'
abwa < *abu* 'father'
xūt < *xit* 'sister'
fūda < *fwād* 'heart'

External plurals for human nouns are mainly formed with *-īn* [masc.] and *-āt* [fem.], the latter ending also being used for diminutives, e.g. *bwēbāt* 'little doors' < *bwēb*), and borrowings, e.g. *bāṣāt* < *bāṣ* 'bus'. Words ending in *-ā(ya)* also so pluralize, e.g. *ḥaṣayāt* 'stones' (< *ḥaṣa*). *-a* is used for some human nouns of the CaCCāC form, e.g. *bayyāʿa* 'sellers', *baḥḥāra* 'seamen'. The *-iyya* suffix is used for some male human nouns, e.g. *drēwiliyya* < *drēwil* 'driver', especially those ending in the relational *-i*, e.g. *ibāḏ̣iyya* < *ibāḏ̣i* 'Ibadi'.

2.2.4.4 Diminutives

The main patterns are:

CCēC(a): *qšēša* 'small [self-seeded] palm tree'
CCēCīC: *dkēkīn* 'small shop'
CCayCiC and CCayiC, common in (Bedouin) men's names: *zwayyid* < *zāyid, swaylim* < *sālim, slayyim* < *salīm*

The suffix *-ūna* is used to denote 'a bit' of something, e.g. *tiššūna* (< *tišša* 'a little'). The -*ō* suffix commonly added to personal names in the eastern Arabian dialects as a hypocoristic diminutive is also occasionally heard, e.g. *aḥmadō*.

2.2.4.5 Vocatives

The system of vocatives involving bipolar address forms (→ Bahraini Arabic, Kuwaiti Arabic) does not seem to be used in Oman. Address forms are on the pattern *yā būyi, yā xūyi*, etc. *yāk* + *wi* + noun is used in warnings, e.g. *yāk wi lliṣṣ!* 'watch out for the thief!'

2.2.4.6 Adjectives

CaCCān/CiCCāC: This is particularly common, and in the Omani reflexes of verbs that had CaCiC- stems in Old Arabic, it often replaces the active participle e.g. *samʿān* 'having heard', *nisyān* 'having forgotten', *kahsān* 'hard-working', *laḥmān* 'bogged down, run aground', *miḥlān* 'affected by drought'.

2.2.4.7 Color and deficiency adjectives

CCaC [masc.], CaCCa [fem.], e.g. *swad, sōda* 'black', *ḥmar, ḥamra* 'red'. Forms such as *ḥamar, xaḏ̣ar* are heard in the Bedouin dialects, although not consistently (see 2.1.2.4). The plural uses the suffixes *-īn* [masc.], *-āt* [fem.], e.g. *ḥamrīn, ḥamrāt*. Forms such as *aḥmar* are interpreted as elatives, i.e. 'redder'.

2.2.4.8 Elatives

These are formed as in Old Arabic. The aCCaC pattern has been extended, however, to include colors and physical defects (see 2.2.4.6) and other forms to which this pattern does not apply in Old Arabic, e.g. *argal* 'more manly' < *ragil* 'man', *astad* 'more expert' < *ustād* 'master, expert'.

2.2.5 Numerals

'One' is *wāḥi* ~ *wāḥid, waḥda* [fem.]. It may precede the noun (2.2.3.2), or follow it with no obvious difference in sense. Note also the phrase *kam min wāḥid/waḥda* ‘a few’.

‘Two’ is *t̲nēn*, feminine *t̲intēn, hintēn, t̲nēna*. The dual or the plural followed or preceded by ‘two’ can be used, with no difference in meaning; e.g. *ḥrumtēn, ḥrumtēn hintēn, ḥarīm hintēn*, and *hintēn ḥarīm*, for example can all be used to mean ‘two women’. In some mountain dialects, structures of the type *t̲nēn nafar* ‘two persons’ are used alongside these. The numeral in such structures is always masculine even where the noun is feminine, e.g. *t̲nēn yōm* ‘two days’ (note that *yōm* is feminine in these dialects), *t̲nēn sāʿa* ‘two hours’. Note also the common construction involving the number ‘two’ with suffixed pronouns = ‘both’, e.g. *sīran bi hintēn-kin!* ‘go, both of you [fem.]!’

‘Three’ to ‘ten’: In the sedentary dialects, denominations of money and dry weight are always singular, and the numeral follows the polarity principle. *yōm, sāʿa, šahar, sana* are also often singular, but here the numeral is always feminine e.g. *t̲alāt̲a yōm* ‘three days’. Otherwise, the sedentary dialects follow the normal gender polarity system. The Bedouin dialects treat money and weight denominations in the same way as the sedentary dialects, otherwise, they have plural forms with gender polarity marked on the numeral. But there are exceptions: ‘month’ is treated as feminine when in the plural, e.g. *t̲alāt̲ šuhūr*. When postposed, the numerals are normally masculine, e.g. *ṣibyān t̲alāt̲* ‘three boys’. Telling time: The feminine form is normally used in the sedentary dialects, and the masculine in the Bedouin dialects (except *sāʿa t̲intēn* ‘two o’clock’ for all).

‘Eleven’ to ‘nineteen’: When enumerating, all speakers use the long form, e.g *xamstaʿšar sana* ‘fifteen years’; in counting, and when no noun is mentioned, the Bedouin speakers have the short form without *-ar*.

‘Hundred’: *miya* or *imya*, pl. *miyāt*. ‘Thousand’: *alf* pl. *ālāf*.

Larger numbers: *lakk*, pl. *lkūk* signifies 100,000.

Ordinal numbers are regular.

Count nouns: Some speakers use the *-āya* suffix, e.g. *ṭamāṭāya* ‘a (single) tomato’; for animals, *rās* or *hayša* is used, e.g. *hayšat ġanam* ‘one goat’; for bananas, *qarn*, e.g. *arbaʿit qrūn mōz* ‘four bananas’; for other fruits, *šōb*, e.g. *šōb min d̲a zzaytūn* ‘one of those olives’; for flowers and plants on stalks, *ʿūd*, e.g. *ʿūd qatt* ‘a single stalk of lucerne grass’.

2.2.6 Verbs

2.2.6.1 Forms

2.2.6.1.1 Form I

(a) Perfect stems

In the sedentary dialects, the factor determining the first vowel in the perfect stem is morphological, broadly corresponding to the Classical Arabic split between roots with V_2 /a/, which have a dialectal *a-a* pattern, and those with Classical Arabic V_2 /i/ or /u/, which have a dialectal *i-i* or *u-u* pattern, the latter usually occurring in the presence of a labial in C_2 and C_3 position, e.g. *simiʿ, kubur*. In these *i-i* and *u-u* stems, the first vowel is often deleted, e.g. *smiʿ, kbur*. In the Bedouin dialects, the voweling of stems depends on phonological factors (cf. the Bahraini A dialects): roots in which C_1 or C_2 are gutturals have *a-a*, for example, and the rest have *i-a*, or *u-a* if labials are C_1 or C_2, There is, however, a great deal of variability, particularly in areas where speakers of the two dialect types live in close proximity.

(b) Imperfect stems

In all dialects, sedentary and Bedouin, voweling reflects both phonological and (diachronic) morphological factors.

i. If C_2 or C_3 is a guttural, the theme vowel is, virtually invariably, /a/ and the prefix vowel /i/. This same *i-a* vowel pattern is applied in both dialect types to any dialectal imperfect verbs that in Classical Arabic had CaCiC and CaCuC perfect stem voweling. If C_1 is a guttural, the sedentary and Bedouin dialects differ in their treatment. In the sedentary dialects, the stem vowel is /i/ or /u/ (depending on the dialect), and the prefix vowel is /a/ in some dialects, /u/ in others. In the Bedouin dialects, the stem vowel is /i/ and the prefix vowel is /a/, but in some northern and southeastern Bedouin dialects, the *ghawa* syndrome (2.1.2.4) applies to such verbs. Interestingly, where a verb has C_1 as a guttural and at the same time was in the Classical Arabic CaCiC-perfect-stem class, (i.e. where two different imperfect dialectal voweling systems, *a-i* and *i-a*, 'compete') a 'compromise' *a-a* form results, e.g. *yaʿmal, yaʿlam*.

ii. In nonguttural stems, there is a good deal of variation in all dialects, but the most frequent type of form has /i/ for both theme and prefix vowel. However, one also encounters forms with *u-u* in both dialect types.

2.2.6.1.2 Derived Forms

Noteworthy semantic aspects are as follows:

As in most dialects, Form II is mainly transitive and causative but can also be denominative, e.g. *šatta* 'to flower in winter' < *šita* 'winter', and intransitive, e.g. *rawwaḥ* 'to go home'. Like Form II verbs, Form III verbs can also be denominative, e.g. *ʿāyad* 'to go Eid visiting' < *ʿīd*, and intransitive, e.g. *xāfaš* 'to shrink, shrivel [fruit on a branch]'.

Form IV scarcely exists, having been replaced by Forms I and (mainly) II.

Forms V and VI: Form V is often reflexive or mediopassive, e.g. *tlaḥḥam* 'to cicatrize, heal', *tṣaffad* 'to get dressed up', but it can also describe actions which have an effect on the agent, e.g. *tšawwaf* 'to have a look'. As well as its 'reciprocal action' value, Form VI often implies gradation or repetition, e.g. *tihāwan* 'to get better from an illness'. In the perfect of Form V and VI verbs, the vowel of the *t-* prefix is often elided; in the imperfect some Bedouin dialects have *iyt-, tti-, nti-* type prefixes, whereas the sedentary have *yit-, tit-, nit-*.

Forms VII and VIII are frequently associated with reflexive or change-of-state verbs, e.g. *inqalad* 'to become fused together', *iḥtās* 'to become thick, coagulated'. Form VII is used to express the passive, but in all Oman dialects the normal means of passivization is internal voweling of the form (see 2.2.6.1.3). The *in-* prefix may be prefixed to Forms other than Form I to form a passive, e.g. Form II *sawwa* 'to do' ⇒ *insawwa* 'to be done [by human agency]' (as opposed to VIII *istawa* 'to turn out [of its own accord]'). In some sedentary dialects, there has been a → resyllabication of Form VIII, e.g. *gathad* 'to work hard' < *igtahad*.

Form IX is not used.

Form X is common, e.g. *istaḥmag* 'to get angry'.

2.2.6.1.3 Internal passive

As in a number of other Arabian dialects, the internal passive is a common dialect feature in Oman, especially in Forms I and II, but passives of Form V are also heard. The imperfect forms (yiCCaC/yiCaCCaC-type) seem more common than perfect ones (C(i)CiC/CiCCiC-type), and 3rd person forms are more common than 1st or 2nd person (see Holes 1998 for suggested explanations). In general, internal passives occur more often in rural sedentary dialects than Bedouin ones. In the Capital Area, prefixed inCaCaC-type forms, the normal means of passivization in the eastern Arabian dialects in general, seem to be taking over from the internal passive.

2.2.6.2 Inflection

2.2.6.2.1 Imperfect

Sedentary dialects:

yiglis 'he sits'	singular	plural
3rd masc.	*yiglis*	*ygilsu*
3rd fem.	*tiglis*	*ygilsan*
2nd masc.	*tiglis*	*tgilsu*
2nd fem.	*tgilsi*	*tgilsan*
1st	*aglis*	*niglis*

The above are the canonical syllable structures for all sedentary dialect imperfect forms. An exception noted is that in a few dialects (e.g. Qalhāt), in verbs in which C_2 or C_3 is a guttural, there is no resyllabication of plural forms, with stress as usual on the penultimate syllable, e.g. *yizráʿu*. The principles for determining prefix and theme vowels in unsuffixed forms are in 2.2.6.1.1 (b) above. In resyllabicated suffixed forms, the theme vowel is normally the same as in the unsuffixed form, e.g. *yiskin/yisiknu*, *yaḥzim/yḥizmu*, although there is a tendency for /a/ to replace /i/ in verbs which have a guttural as C_1, e.g., *yʿarfu* often occurs instead of *yʿirfu*.

Bedouin dialects:

yiylis 'he sits'	singular	plural
3rd masc.	*yaylis*	*iyalsu* or *iyalsūn*
3rd fem.	*taylis*	*iyalsan*
2nd masc.	*taylis*	*tyalsu* or *tyalsūn*
2nd fem.	*tyalsi* or *tyalsīn*	*tyalsan*
1st	*aylis*	*naylis*

In the Bedouin dialects of the Empty Quarter and in the dialect of the Āl-Wahība in the southeast, one regularly encounters plural forms in *-ūn*, typical of the dialects of eastern and central Arabia. In areas where there has been prolonged contact with other Bedouin dialects – the northern Bāṭina and Ṣūr – such forms are also encountered, though here mixed with sedentary forms lacking *-ūn*.

Aspect/mood prefixes: All dialects have *ba-* and *ha-* for proximate intent. *q/gāʿid* is used to express continuous or iterative processes.

The imperative: *glis, gilsi, gilsu, gilsan*. The negative imperative is *lā tiglis/tgilsi/tgilsu/tgilsan*.

2.2.6.2.2 Perfect

Sedentary dialects:

galas 'he sat'	singular	plural
3rd masc.	*galas*	*galsu*
3rd fem.	*galasit* or *galsit* or *galsat*	*galsan*
2nd masc.	*galast/glast*	*galastu/glastu*
2nd fem.	*galasti/glasti*	*galastan/glastan*
1st	*galast/glast*	*galasna/glasna*

Verbs with *i-i* voweling have similar forms, except that the initial /i/ is usually deleted in the 3rd person masculine singular, e.g. *kṯir* 'it grew abundant', and the 3rd person feminine singular is of the type *kiṯrit*. Where C_3 is a guttural, the vowel of the 3rd person feminine singular ending tends to be *-at*, rather than *-it*, e.g. *firḥat* 'she was happy', *simʿat* 'she heard'. In all forms with a Cv-CvCC syllable structure, the unstressed short vowel is often deleted, e.g. *glast*, etc.

Bedouin dialects:

yalas 'he sat'	singular	plural
3rd masc.	*ylas*	*ylasaw/u* or *ylisaw/u*
3rd fem.	*ylasit* or *ylisat*	*ylasan* or *ylisan*
2nd masc.	*ylast*	*ylastaw/tu*
2nd fem.	*ylasti*	*ylastan*
1st	*ylast*	*ylasna*

In some Bedouin dialects of the Empty Quarter, /j/ is the usual reflex of Old Arabic *jīm*, not /y/, thus *jlast*, etc.

2.2.6.3 Participles

The morphology of active and passive participles is similar to that of Old Arabic. CvCCān is a highly productive innovation and provides an alternative to CāCiC in many verbs that have *i-i* voweling in the perfect (i.e. those that had Old Arabic CaCiC- and CaCuC- perfect stems) – see 2.2.4.5 above. See 2.1.3.6 (b) for the shape of suffixed active participle forms.

2.2.6.4 Verbal nouns

Form I: These are similar to those in other Gulf dialects.

Derived form verbal nouns:

The unusual asterisked forms below were all recorded in the speech of uneducated speakers in interior villages in the Jabal Axḍar region, and from the Šarqiyya town of Muḍaybi. Several of them are attested for Old Arabic in Wright (1896: I, 115–116). Reinhardt (1894:168–169) gives many more such forms, now obsolescent.

II: taCCīC: *taṣfīd* 'tidying, cleaning preparing [coffee]'
tvCCūC (also Form V): *taqdūm* 'offering of food'
CaCCīCa/CaCCāCa: *ṭawwīra*/ṭawwāra** 'development'

V: tiCaCCīC(a): *tixarrīb** 'becoming corrupt', *tilabbīsa** 'dressing up', *tišawwīfa** 'having a look'
tiCvCCaCa: *tibiššara* 'season when dates first ripen'
tCuCCa: *trukba* 'fireplace of stone'

Modern Standard Arabic-type verbal nouns Forms V, VI, VII, VIII, X do occur, but not in uneducated speech.

2.2.7 Weak verbs

2.2.7.1 Geminates

Perfect forms with consonant-initial suffixes are on the usual pattern CaCCēC. There is contraction of the doubled consonant in the active participle of Form I, e.g. *hābb rīḥ* 'dextrous, expert', and in the Form III and VI paradigms, e.g. *tḥāggu* 'they argued with each other'.

2.2.7.2 Iʾ verbs

In the perfect of Form I, the initial syllable is lost in some verbs, e.g. *kal* 'to eat', *xaḏ* 'to take', and the verb is conjugated either like a strong verb (*kalt*, etc.) or like a doubled one (*kalēt*); if it has final weak radical, e.g. *ta* 'to come, bring', it behaves like a weak verb (*tēt*, etc.). In other verbs, /ʾ/ has simply been lost, e.g. *amar* 'to order', or been replaced by /w/, e.g. *wann* 'to groan'. In the sedentary dialects, the imperfect of *kal*, *xaḏ* has the prefix vowel /ū/ or /ō/, e.g. *yōkil*. In the Bedouin dialects, this vowel is /ā/: *yākil*. The verb *ta*, which is only found in the sedentary dialects, has a *yā*-imperfect prefix: *yāti* 'he comes, brings', passive *yāta* or *yūta* 'it is brought'. The active participle is of the māCCiC form: *mākil, māti*, passive participles *mākūl* (or *maykūl*), *matāy*.

In Form II, /ʾ/ has been lost in the perfect in some verbs, e.g. *aḏḏan* 'to call to prayer', but reappears intervocalically in the imperfect: *yiʾaḏḏin* (but Bedouin *yinnis*). In others, it has been replaced by /w/, e.g. *wakkad* 'to be certain of, know well'.

Form I imperatives are typically of the type *kil, kli, klu, klan*.

2.2.7.3 I*w* verbs

The /w/ is preserved in the Form I imperfect in all dialects, becoming *yō-* or *yū-*, e.g. *yūšab*, and in the imperative, e.g. *ōquf!* 'stop!' In the passive participle, /aw/ often becomes /ā/, e.g. *māgūd* 'existent'.

2.2.7.4 II*w/y* verbs

The vowel of the Form I imperative is usually long in the Bedouin and most sedentary dialects, e.g. *gūl* 'say!', etc. but short in a few sedentary ones: *sir* 'go!'. In the passive participle, /y/ and /w/ are treated as strong consonants (with /w/ > /y/), e.g. *maṣyūd* 'caught'. The VIII perfect behaves as Form I, e.g. *štuft* 'I saw'.

2.2.7.5 III*w/y* verbs

All III*w/y* verbs are treated as verbs with final /y/. There are two types: the imperfect /a/ type, e.g. *bqí, yibqa*, and the imperfect /i/ type, e.g. *dára, yidri*. Verbs which in Old Arabic had final *hamza* lost it, and their imperfects were absorbed into the /a/ type, and Old Arabic final-/w/ verbs have been absorbed by the /i/ type. One verb appears in both forms *baġa, yabġi/yibġa* 'to want'. In both types of weak verb, it is a characteristic of the Omani sedentary dialects that /y/ is always treated as a strong radical in all forms, e.g. *daryit* 'she knew', *tdaryan* 'you [fem. pl.] know'. The same is true of the derived forms. In the Bedouin dialects, forms without the /y/ are the rule, as in eastern Arabia.

In the sedentary dialects, another peculiarity is that the passive participle of final weak verbs is typically of the form maCCāy, e.g. *maqlāy* 'cooked'.

2.2.7.6 Irregular verbs

The verb *ga/ya* 'to come':

	sedentary		Bedouin	
singular				
3rd masc.	*ga*	*ygi*	*ya*	*iyi*
3rd fem.	*gayit*	*tgi*	*yat* or *yāt*	*tyi*
2nd masc.	*gīt*	*tgi*	*yīt*	*tyi*
2nd fem.	*gīti*	*tigyi* or *tgiyi*	*yīti*	*tyi* or *tyīn*
1st	*gīt*	*agi*	*yīt*	*ayi*
plural				
3rd masc.	*giyu* or *giyyu*	*yigyu* or *ygiyu*	*yaw*	*iyu* or *iyūn*
3rd fem.	*giyan* or *giyyan*	*yigyan* or *giyan*	*yān*	*iyan*
2nd masc.	*gītu*	*tigyu* or *tgiyu*	*yītaw* or *yītu*	*tyu* or *tyūn*
2nd fem.	*gītan*	*tigyan* or *tgiyan*	*yītan*	*tyan*
1st	*gīna*	*ngi*	*yīna*	*nyi*

In some Bedouin dialects of the Empty Quarter, as previously noted, this verb would have /j/, e.g. *jīt* etc. On suffixation with a consonant-initial suffix, there is, in the case of *-aw* often, and in the case of *-an* always, consonant doubling, e.g. *yawwōk* 'they [masc. pl.] came to you' (also: *yōk*), *yānnak* 'they [fem. pl.] came to you'.

There are a number of interesting variants for 'to come', all in the sedentary dialects. As already noted, *ta* 'to come; to bring' survives, used mainly in the imperfect, e.g. *tāti lih ṭaʿām* 'she brings him food'. It also occurs, as in Old Arabic, with *bi* in the sense 'bring' e.g. *yātā buh* 'it is brought', and in the 'amalgamated' form *tāb, yitīb* 'to bring', e.g. *tābu lī šway duwa* 'they brought me some medicine', alongside the more common *gāb, yigīb*. The variant *ada, yādi* 'to come; to bring' also occurs in some dialects as an alternative to *ta, yāti*.

2.2.8 Quadriliterals

Quadriliterals are common. The main types are reduplicatives, e.g. *taftaf* 'to tear, rip', and with *t*- prefix *ddabdab* 'to become covered in pustules'; echoic, mimetic in which C_2 often = /l, r, n/, e.g. *barbaq* 'to bubble'; C_2 = /w/ or /y/: *ṣōlaf* 'to chat' with reciprocal pattern *tiṣōlaf* 'to chat with each other' < *ṣālfa* 'matter'; *tfōša* 'to sit with parted legs'; *bēxal* 'to be miserly'. With a *t*- prefix, the latter denotes the adoption of bodily/mental states which are foreign to the individual concerned, e.g. *tṣēmax* 'to pretend to be deaf' (cf. Kuwaiti and Bahraini dialects). C_2 = /r, n/ may be an insertion or the result of dissimilation, e.g. *xarbaṭ* 'to mix up' < *xabaṭ* 'to beat', *tsansaḥ* 'to slide, slither' < *tsaḥsaḥ*. Others fall into no particular pattern, e.g. *lahwaz* 'to stain, make dirty', *tḥankar* 'to become turbid'. Apart from the above, there are denominatives formed from secondary forms, e.g. *tmargaḥ* 'to swing, sway' (< *margūḥa*). For more examples see Holes (2004).

Bibliographical references

Brockett, Adrian. 1985. *The spoken Arabic of Khābūra on the Bāṭina of Oman.* (= *Journal of Semitic Studies Monograph*, VII.) Manchester: University of Manchester.

Galloway, Douglas. 1977. *An introduction to the spoken Arabic of Oman.* Unpubl. mimeograph, University of St. Andrews.

Holes, Clive. 1989. "Towards a dialect geography of Oman". *Bulletin of the School of Oriental and African Studies* 52.446–462.

——. 1996. "The Arabic dialects of south-eastern Arabia in a sociohistorical perspective". *Zeitschrift für Arabische Linguistik* 31.34–56.

——. 1998. "Retention and loss of the passive verb in the Arabic dialects of northern Oman and eastern Arabia". *Journal of Semitic Studies* 43.347–362.

——. 2004. "Quadriliteral verbs in the Arabic dialects of eastern Arabia". *Approaches to Arabic dialects: A collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday*, ed. Martine Haak, Rudolf de Jong, and Kees Versteegh, 97–116. Leiden: E.J. Brill

Ingham, Bruce. 1986. "Notes on the Āl-Murra of eastern and southern Arabia". *Bulletin of the School of Oriental and African Studies* 49.271–291.

Jayakar, Atmaram. 1889. "The Omanee dialect of Arabic". *Journal of the Royal Asiatic Society* 649–687, 811–889.

——. 1902. "The Shihee dialect of Arabic". *Journal of the Royal Asiatic Society* 246–277.

Landberg, Carlo de. 1920–1942. *Glossaire dathînois.* Leiden: E.J. Brill.

Reinhardt, Carl. 1894. *Ein arabischer Dialekt gesprochen in Oman und Zanzibar.* Stuttgart and Berlin. (Repr., Amsterdam: APA Oriental Press, 1972.)

Rhodokanakis, Nikolaus 1908, 1911. *Der vulgärarabische Dialekt im Ḍofâr (Ẓfâr).* Vienna: Alfred Hölder.

Smart, Jack. 1990. "Pidginization in Gulf Arabic: A first report". *Anthropological Linguistics* 32.83–119.

Thomas, Bertram. 1930. "The Kumzari dialect of the Shihuh tribe, Arabia, and a vocabulary". *Journal of the Royal Asiatic Society* 785–854.

Webster, Roger. 1991. "Notes on the dialect and the way of life of the Āl-Wahība Bedouin of Oman". *Bulletin of the School of Oriental and African Studies* 54.473–485.

Wright, William. 1896. *A grammar of the Arabic language.* Cambridge: Cambridge University Press. (Repr., 1964).

Lutz Edzard (University of Oslo)

Onomastics → Proper Names

Onomatopoeia → Sound Symbolism

Optimality Theory

1. Introduction

Many phonetic and phonological observations can be conveniently recast in terms of theories of linguistic preference and natural generative phonology (cf. Hooper 1976), notably in terms of the approach of Vennemann (1983, 1988), which was applied to Semitic by Edzard (1991). Optimality Theory, originally proposed by Prince and Smolensky (1993), offers a formal means to capture the 'constraint ranking' that is implicit in the rejection of disallowed forms and the evaluation of competing forms ('candidates') of linguistic surface forms. While Optimality Theory nowadays also provides meaningful applications to other realms of grammar, notably syntax (for an application in Arabic historical morphology, see Kusters 2003), it continues to be most prominent as an explanatory device for linguistic features at the interface of phonology and morphology. Recent linguistic theory has paid so much attention to the phonotactics of Arabic dialects and other Semitic and Afro-Asiatic languages, notably varieties of Berber, that a summary of the most interesting applications of Optimality Theory to the morphophonology of different varieties of Arabic is warranted.

In Optimality Theory, the set of forms that may reasonably be assumed to be potential surface forms is often referred to as the 'richness of the base'. One can, in principle, use the same term in reference to the availability of simultaneously occurring forms, some of which may be true alternatives and some of which may be regional variants. In the framework of Optimality Theory, the co-occurring forms, or rather the forms that are subject to linguistic evaluation, are called 'candidates'.

One other central concept in Optimality Theory is markedness. Broadly speaking, 'marked' refers to forms that are unusual, rare, or harder to pronounce, whereas 'unmarked' refers to forms that are natural, more frequent, or easier to pronounce. The latter state is often called 'well-formedness'. On the segmental level, for example, velarized stops are considered 'marked', whereas plain (nonvelarized) stops are considered 'unmarked'. On the suprasegmental level, for example, the universally 'unmarked' syllable structure is CV: onset, nucleus, and no coda; other syllable structures (CVC, CVCC, CCVCC, etc.) are then considered 'marked'. And even within one and the same syllable type, there may be more or less marked specimens of different quality, depending on the internal structure of syllable onset and syllable coda.

2. Basic concepts and tenets of Optimality Theory

Optimality Theory, which always aims at singling out one 'optimal' form, has the potential to evaluate at least the following morphophonological parameters (cf. McCarthy and Prince 1994:335):

i. Segmental harmony (unmarkedness, itself consisting of various dimensions, some conflicting);
ii. Syllabic harmony (having an onset, lacking a coda);
iii. Faithfulness (identity between input and output);
iv. Alignment (coincidence of edges of morphological and phonological constituents);
v. Metrical parsing (satisfying constraints on exhaustivity and alignment of metrical feet);
vi. Template satisfaction (meeting shape or constituency requirements imposed on the reduplicated string);
vii. Exactness of copying relation; and
viii. Identity between the reduplicated string and the base to which it is attached.

While the references to segmental harmony and syllabic harmony are quite straightforward and unproblematic, the reference to faithfulness is interesting insofar as the concept of 'underlying representation' (i.e. the 'input'), which Optimality Theory purports to discard, is reintroduced, so to speak, via the backdoor.

It is important to note that usually not all of these parameters can be optimized in any given form. The principle underlying this circumstance is often called the 'fallacy of perfection' (cf. also Vennemann 1988:1ff.). For instance, words that are entirely made up of CV syllables – this being the 'optimal' syllable structure – may turn out to be lengthy or otherwise clumsy to pronounce.

One can narrow down Optimality Theory to five basic tenets (cf. McCarthy and Prince 1994:3):

i. Universality: U[niversal] G[rammar] provides a set {Con} of constraints that are universal and universally present in all grammars;
ii. Violability: constraints are violable, but violation is minimal;
iii. Ranking: the constraints of {Con} are ranked on a language-particular basis; the notion of minimal violation is defined in terms of this ranking, and a grammar is a ranking of the constraint set;
iv. Inclusiveness: the constraint hierarchy evaluates a set of candidate analyses that are admitted by very general considerations of structural well-formedness;
v. Parallelism: best-satisfaction of the constraint hierarchy is computed over the whole hierarchy and the whole candidate set; there is no serial derivation.

McCarthy and Prince (1994:335–336) conclude:

> The construction of a grammar in Optimality Theory is essentially a matter of determining the proper ranking of the set of constraints {Con}, and to that end the *constraint tableau* is a useful calculational

device. A typical constraint tableau, showing the domination of constraint B by constraint A, is the following:

(1) Example of a Constraint Tableau, A ⇒ B, /in$_k$/ ⇒ *k-cand*$_1$

Candidates	A	B
a. ☞ k-cand$_1$		*
b. k-cand$_2$	* !	

In this tableau, it is assumed that, given the input /in$_k$/, the generator *Gen* supplies at least the candidates k-cand$_1$ and k-cand$_2$. Constraints A and B disagree on these two candidates, and since the A-obeying k-cand$_1$ is optimal, constraint A must dominate constraint B. In this and other tableaux, constraints are shown in domination order, and violation-marks are indicated by '*'. The optimal candidate is called out by '☞', and fatal constraint violations are signalled by '!'. Below these fatal violations, cells are shaded to indicate their irrelevance to determining the outcome of the comparison at hand.

Preservation of faithfulness and preservation of markedness are the two basic competing constraints at the heart of Optimality Theory. Then, there are many other language-specific constraints that determine the morphophonological 'fine-tuning' in the language under observation. While constraints as such are supposed to be universal, their ordering is usually language-specific.

For an in-depth introduction to Optimality Theory, see Kager (1999) and McCarthy (2002). McCarthy (2004) is a reader covering the areas of prosody, segmental phonology, and interfaces between various levels of grammar. The Rutgers Optimality Archive (http://roa.rutgers.edu/index.php3) is an excellent resource for pdf versions of papers in the Optimality Theory framework. The OT archive can also be searched specifically for languages; as of this writing (May 2007), a search yields twelve papers dealing (among others) with Arabic, nine dealing with Modern Hebrew, and one dealing with Berber and Cushitic data.

3. Issues in Semitic and Afro-Asiatic linguistics as reflected in Optimality Theory

Already in their seminal paper on Optimality Theory, Prince and Smolensky (1993) place much weight on the analysis of syllabification in the Imdlawn Tashilhiyt dialect (variety) of Berber, based on data provided in Dell and Elmedlaoui (1985) and later studies (see the precise algorithm for syllabification in McCarthy 2004:7–17). In this language variety, almost any consonant apparently can serve as the nucleus of a syllable, but a nucleus with a higher degree of sonority is more natural. Consider the syllabification output of the form /ħaul-tn/ 'make them [masc.] plentiful!' (cf. Dell and Elmedlaoui 1985:110), which is *ħa.wL.tN* (syllable-nuclear consonants are represented in capitals). The relevant constraints for a preferred syllable structure in this context are the following:

> Ons (the Onset Constraint): Syllables must have onsets (except phrase-initially).
>
> Hnuc (the Nuclear Harmony Constraint): A higher sonority nucleus is more harmonic than one of lower sonority.

As the top priority is to provide syllables with onsets, the constraint Ons 'outranks' or 'dominates' the constraint Hnuc in this case. This circumstance is mirrored in Tableau 2.

(2) Constraint Tableau for the comparison of second syllable candidates for /ħaultn/ 'make them [masc.] plentiful!'

Candidates	Ons	Hnuc
a. ☞ ~.wL.~		|l|
b. ~.ul.~	* !	|u|

Having no consonantal onset, the candidate ~.ul.~ incurs a fatal violation (* !), as stipulated by the constraint Ons. As in Tableau 1, cells 'below' (i.e. to the right of) the fatal violation are shaded to indicate their irrelevance to determining the outcome of the comparison under investigation. In other words, the fact that |u| is a 'better' nucleus than |l| plays no role in terms of its place on the sonority scale.

Alderete (1997) analyzes dissimilation phenomena in the Cushitic language Oromo and in Tashilhiyt Berber, among other languages. As a result of a co-occurrence restriction in the latter variety of Berber, which rules out more than one labial in a word, derivational *m*-prefixes are dissimilated (or delabialized) in front of a root already containing a labial (/b, f, m/; cf. also Edzard 1992 for related processes in Akkadian and Ethio-Semitic). Again, we are looking at two competing constraints:

> ***Pl/Lab$^{2}_{\text{Stem}}$:** Ban any stem with two segments with independent Place specification [labial];
>
> **Ident[Place]:** Corresponding segments in input and output agree in [space] specification.

Tableau 3, using the reflexive form *nkaddab* 'he was considered a liar', captures the intrinsic hierarchy of these two constraints, i.e. the circumstance that *Pl/Lab$^{2}_{\text{Stem}}$ dominates Ident [Place].

(3) Constraint Tableau for delabialization as a result of dissimilation

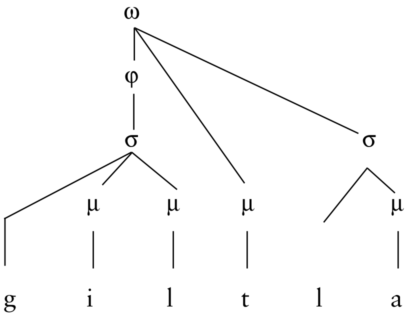
Input: m-kaddab	*Pl/Lab^{2}	Ident[place]
a. ☞ [n-kaddab]$_{\text{stem}}$		*
b. [m-kaddab]$_{\text{stem}}$	* !	

A number of recent papers (e.g. Graf 2000; Graf and Ussishkin 2002) analyze stress assignment and nominal reduplication in Modern Hebrew. Bat-El (1996) covers the morpho-phonology of Hebrew blends. Rose and Walker (2004) deal with co-occurrence restrictions in the Ethio-Semitic languages Amharic and Chaha and the Omotic language Gimira.

4. Case studies of the application of Optimality Theory to Arabic dialectology

The interaction of → stress on the one hand and vowel → epenthesis and deletion (→ syncope) on the other hand in Arabic dialects has been a crucial issue in recent linguistic theory (roughly since the late 1970s; see, e.g., Selkirk 1981; Broselow 1982). At issue are first and foremost the different patterns (or 'directionalities') of vowel epenthesis in various groups of Arabic dialects. Morphologically derived CCC clusters – one of the 'model forms' being /gilt-la/ 'I/you [masc. sg.] told him' – can be syllabified as *gilitla* ('VC-dialect'; this is how it appears in many Iraqi dialects), as *giltila* ('CV-dialect'; cf. Cairene *ʾultilu*), or (without change) as *giltla* ('C-dialect', cf. *qiltlu* in many Moroccan dialects). More recent research analyzes these different output forms in connection with constraints posed by prosodic theory and gives an account of the resulting surface forms in terms of constraint ranking (cf., e.g., Broselow 1992; Abu-Mansour (1995); Farwaneh 1995; Zawaydeh 1997). Kiparsky (2003) and Watson (forthcoming) additionally resort to the status of 'semisyllables' in order to describe the 'input' forms. The model form /giltla/, for instance, can be represented prosodically as in Figure 1 in VC- and C-dialects (ω = word; φ = foot; σ = syllable; μ = mora):

Figure 1. Prosodic representation of /giltla/



In a VC-dialect, the less-than-fortunate situation with the unaffiliated *t* is remedied in that an epenthetic *i* is inserted to its left side. Watson, making use of the constraint designations developed in Kiparsky (2003), provides the analysis in Tableau 4 for the postlexical level in VC-dialects. The crucial point here is the ranking of the constraints License-*μ* and Reduce. The former stipulates that each mora must be 'licensed', i.e. affiliated with the next higher prosodic level, in this case the syllable. The latter stipulates that the number of light (CV) syllables be minimal.

(4) Constraint Tableau for *gilitla* 'I/you [masc. sg.] told him' on the postlexical level in VC-dialects

Input:/gilt-la/	license-*μ*	Reduce
a. (gil)tμ.la	* !	*
b. ☞ (gi)lit.la		**

Tableau 4 illustrates that the candidate (gil)tμ.la incurs a fatal violation of the License-μ constraint (the *t* is not 'licensed'), which outranks (dominates) the other constraint Reduce. Beyond that example, Watson (forthcoming) shows that Kiparsky's (2003) typology of CV-, VC-, and C-dialects must indeed be refined and expanded to a four-type model in view of the complexity of the data.

In Table 1, the perfect paradigm of the verb *kitab* (Standard Arabic *kataba*) in the Arabic dialect spoken by educated Muslims in Baghdad (cf. Erwin 1963:84; Broselow forthcoming:16–20) is given.

Table 1. Perfect paradigm of *kitab* 'he wrote'

3.m.sg.	*kitab*
3.f.sg.	*kitbat*
2.m.sg.	*kitábit*
2.f.sg.	*kitabti*
1.c.sg.	*kitábit*
3.c.pl.	*kitbaw*
2.c.pl.	*kitabtu*
1.c.pl.	*kitabna*

Broselow is interested first and foremost in a contrastive analysis of the forms *kitbat* (3rd pers. fem. sg.) and *kitábit* (2nd pers. masc. sg./1st pers. comm. sg.), as the latter surface contradicts normal stress patterns in this and other Arabic dialects. Both forms can be derived synchronically as in Table 2.

Table 2. Synchronical derivation of *kitbat* and *kitábit*

	/kitab+at/	/kitab+t/
Syncope	kitbat	—
Syllabification	kit.bat	ki.tab.t
Stress assignment	kít.bat	ki.táb.t
Epenthesis	—	ki.táb.it
Resyllabification	—	ki.tá.bit
	[kítbat]	[kitábit]

Broselow's careful argument shows that the data do not lend themselves to a coherent phonological analysis in terms of serial derivation, especially if one takes into consideration corresponding forms of weak verbs. As an alternative within the Optimality Theory framework, she suggests an effect of morphological distinctness, contrasting the bases of [+3.ps.] and [–3.ps.] forms. This constraint is formulated as follows:

> [–3] Contrast: A base bearing a non-third person ([–3]) suffix must be distinct from the unmarked [+3] base in the identity of the stressed vowel.

This constraint is in turn outranked (dominated) by a constraint which outrules complex codas in this dialect (*ComplexCoda), but it dominates the other two other relevant constraints in this case, StressConstraints and Syncope. The Tableaux 5 and 6 illustrate the situation.

(5) Constraint Tableau for *kítbat* 'she wrote'

Input: /kitab+at/	Complex Coda	[–3] Contrast (kítab)	Stress Constraints	Syncope
a. kítabat				* !
b. ☞ kítbat				
c. kitábat			* !	

(6) Constraint Tableau for *kitábit* 'you [masc. sg.]/I wrote'

Input: /kitab+at/	Complex Coda	[–3] Contrast (kítab)	Stress Constraints	Syncope
a. kítabit		* !		
b. kítbit				
c. ☞ kitábit		* !	*	

In the case of (5), the candidates *kítabat* and *kitábat* incur fatal violations with respect to the constraints StressConstraints and Syncope, respectively. In the case of (6), the candidates *kítabit* and *kítbit* incur fatal violations with respect to the constraint [–3]Contrast, as the desire is to maximize contrast between the [+3] and [–3] forms. Hence, it is irrelevant here that the optimal candidate *kitábit* violates the constraint StressConstraints.

Another highly prominent issue in cyclic phonology, which by now has also witnessed an Optimality Theory treatment, concerns the 'underapplication' of → syncope in (most) VC-dialects, notably Levantine dialects, in which minimal pairs such as *fhímna* 'we understood' (Standard Arabic *fahimnā*) vs. *fihímna* 'he understood us' (Standard Arabic *fahima-nā*) arise; in both cases /fihimna/ is considered to be the 'underlying representation' or 'input' (cf. Brame 1974; Kenstowicz 1981; Mitchell 1993:156). Precisely this minimal pair has also caught the attention of scholars working in the framework of Optimality Theory (cf. Kenstowicz 1996; Kager 1999:278–293; Kiparsky 2003: 162–163). The traditional derivation of these forms in terms of a cyclical analysis ('extrinsic rule ordering') in Table 3 shows the interaction of stress assignment and vowel syncope, viz. the blocking of the latter in the base in front of the object (accusative) suffix (cf. Kager 1999:281).

Table 3. Derivation of *fhímna* 'we understood' and *fihímna* 'he understood us'

	we understood	he understood us
Input	[fihim-na]$_{\text{Subj}}$	[[fihim-]na]$_{\text{Acc}}$
Cycle 1		
Stress	fihim-na	fíhim
Cycle 2		
Stress	—	fihím-na
Postcyclic		
i-Syncope	fhím-na	*application blocked*
Destressing	*not applicable*	fihím-na
Output	fhímna	fihímna

Within the framework of Optimality Theory, which avoids derivations (and hence intermediary representations), it is possible to capture the emergence of the different output forms by claiming that the unstressed vowel [i] in a verb form with object suffix is protected from deletion if the verb form's base (without the object suffix) is indeed stressed, as is the case with *fíhim* 'he understood'. (The same holds *mutatis mutandis* for nouns with possessive suffixes.) Technically, this constraint can be formulated as follows:

> HeadMax-BA: Every segment in the base's prosodic head [i.e. in this case, the first syllable of the word ('base') *fíhim*] has a correspondent in the affixed form:
>
> [f i . h i m]
> | |
> [[f i . h i m] n a]

The other two necessary constraints in this context are:

> No [i]: /i/ is not allowed in light [open, unstressed] syllables;
>
> Max-IO: Every segment in the input has a correspondent in the output.

No [i] Is ranked higher (and thus represented more leftward in the tableau) since vowel deletion in forms such as *fhímna* occurs at the expense of the constraint Max-IO. In turn, HeadMax-BA is ranked higher than No [i], since vowel deletion is blocked, i.e. i-syncope is 'underapplied' whenever the base identity constraint is relevant, as in *fihímna*.

The interaction of these three constraints is shown in Tableaux 7 and 8.

(7) Constraint Tableau for *fhímna* 'we understood'

Input: fihim-na/ Base: none	HeadMax-BA	No [i]	Max-IO
a. [fi.hím.na]		* !	
b. ☞ [fhímna]			*

(8) Constraint Tableau for *fihímna* 'he understood us'

Input: fihim-na/ Base: none	HeadMax-BA	No [i]	Max-IO
a. ☞ [fi.hím.na]		*	
b. [fhímna]	*!		*

In the case of (7), the constraint No [i] 'outranks' the faithfulness constraint requiring that input segments have correspondents in the output. Therefore, the candidate [fi.hím.na] incurs a fatal violation. The constraint HeadMax-BA does not apply in this case, as there is no base. In the case of (8), where there *is* a base ([fi.him]), it is the candidate [fhímna] that incurs a fatal violation, since the highest ranking constraint HeadMax-BA is not upheld.

Kiparsky (2003:162ff.) deals with the same minimal pair based on his constraint-based elaboration of Lexical Phonology and Morphology and presents pertinent constraint tableaux for the lexical (as opposed to the postlexical) level of these output forms.

These examples of applications of Optimality Theory to Arabic (dialectology) are by no means exhaustive. Further applications include McCarthy (2005a), where the length of stem-final vowels in colloquial Arabic is accounted for, and McCarthy (2005b), where whole paradigms are ranked in terms of Optimality Theory.

Finally, one should not forget that there is also a historical dimension to Optimality Theory. Prince and Smolensky (1993; cited here after McCarthy 2004:29) acknowledge this by referring to the following statement as

"Pāṇini's Theorem on Constraint ranking": "the presence of a more general constraint in a superordinate position in a hierarchy will eliminate all opportunities for a more specialized constraint in a subordinate position to have any effects in the grammar", such as "Pāṇini's Theorem on Constraint ranking". Sībawayhi's radically descriptive approach to grammar (cf. Carter 1973:146 n.) lends itself especially well to Optimality Theory, as he often lists simultaneously occurring forms – sometimes associated with the speech of different tribes – which he then indeed ranks with value judgments such as *ḥasan* 'good', *ʾaḥsan* 'better', or simply *ʿarabī* 'Arabic' (e.g. in the very last sentence of Chap. 571, *Kitāb* IV, 485), according to certain parameters. The different *maṣādir* (sg. *maṣdar* 'infinitive') of the verb *watada* 'to pin' may serve as an illustration (cf. *Kitāb* IV, 474). According to Sībawayhi, there occurred an array of forms, ranging from *watd*/*watid* (in the Ḥijāz) and *wadd* (with the Tamīm) to *tida* (cf. Rabin 1951:1–5). While the *maṣdar* forms *watd* and *watid* are superior in terms of preserving the linguistic input (especially with respect to the root consonants), the assimilated *maṣdar* form *wadd* is superior in terms of linguistic 'naturalness', i.e., it is easier to pronounce and hence phonologically unmarked. The latter form has the disadvantage, though, of being identical with the *maṣdar* form of the verb *wadda* 'to love', which is, of course, not related to *watada*. The best ('optimal') *maṣdar* form is clearly *tida*, which meets the criteria of both faithfulness to the linguistic input (with respect to the second and third root consonants) and naturalness.

Sībawayhi's observations may be recast more technically in terms of Optimality Theory (cf. Edzard 2000). Recall that Optimality Theory is in principle designed to single out the optimal surviving candidate and to throw out the rest of the candidates. In contrast to the standard application of Optimality Theory, the *maṣdar* forms *watd*/*watid* (in the Ḥijāz) and *wadd* (with the Tamīm) are not disallowed, even though they are (were) regionally limited. The *maṣdar* form *tida* has the advantage of best preserving the root structure *w-t-d* in its (acoustic) output form, even though the first root consonant C_1 (= *w*) is lost. But in the case of verbs I*w* (C_1 = *w*) in Arabic, the clear articulatory preservation of the last two consonants appears to be crucial for an easy recognition of the form. This observation can be technically rephrased to the extent that the output has to match the input with respect to C_2 and C_3. Hence, this is the dominating constraint in this context, and *wadd* is clearly the worst candidate. Incidentally, *watd* is probably just as bad, because the surface pronunciation is almost certainly bound to be [watt]. The next constraint operating in this example is the circumstance that syllable codas with increasing sonority are universally disfavored for clear articulatory reasons. Such syllable codas are almost 'crying out' for an epenthetic vowel. The form *watid* may be considered the result of such an epenthetic process. Finally, there is the universal tendency to reduce the number of syllables with weak onsets. Tableau 9 simply demonstrates the mutually opposing forces ('constraints') of faithfulness between input and output on the one hand (Ident-IO-Root), and phonological naturalness (unmarkedness) on the other hand (*Incr-Son-Cod) – the opposing 'candidates' here are *watd* and *wadd*:

(9) Constraint Tableau, Ident-IO-Root >> *Incr-Son-Cod

Candidates	Ident-IO-Root	*Incr-Son-Cod
a. ☞ watd		*
b. wadd	* !	

Tableau 9 illustrates that neither of the two candidates *watd* and *wadd* is 'perfect'. Assuming, however, that preserving the root consonants is more important than having an easy-to-pronounce syllable coda, i.e. in technical terms, that Ident-IO-Root dominates *Incr-Son-Cod, the candidate *watd* emerges as the better and hence 'optimal' *maṣdar* form.

Tableau 10 summarizes the more complex situation involving all four *maṣdar* forms.

(10) Constraint Tableau, Ident-IO-C_2-C_3 >> *Incr-Son-Cod >> *Weak-Ons

Candidates	Ident-IO-C_2-C_3	*Incr-Son-Cod	*Weak-Ons
a. ☞ tida			
b. watid			*
c. watd		*	*
d. wadd		* !	

Tableau 10 clearly illustrates the ranking of the three constraints – here one might also call them 'tendencies' – that are operational in determining the 'quality' of the different *maṣdar* forms. Note that the form *wadd* is marked with an exclamation mark, which signals the fatal incurrence of the violation of the constraint Ident-IO-C_2-C_3. The forms *watd* and *watid* incur violation marks as well, but only on a lower constraint level.

5. SUMMARY

The central idea of Optimality Theory is that surface forms of language reflect resolutions of conflicts between competing constraints. A surface form is 'optimal' if it incurs the least serious violations of a set of constraints, taking into account their hierarchical ranking. Languages differ in the ranking of constraints, and any violations must be minimal. The data in the realms of Arabic phonetics, phonology, and elsewhere can be analyzed in harmony with these principles. This supports an explanatory approach to Arabic morphophonology in terms of naturalness and preference theory in general, and of Optimality Theory in particular.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCE

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. Ed. Muḥammad ʿAbd as-Salām Hārūn. 5 vols. Cairo, 1968–1977. (Repr., Cairo 1977.)

SECONDARY SOURCES

Abu-Mansour, Mahasen. 1995. "Optimality and conspiracy in the syllable structure of Arabic". *University of Massachusetts Occasional Papers in Linguistics*, XVIII: *Optimality Theory*, ed. Jill Beckmann, Laura Walsh Dicky, and Suzanne Urbanczyk, 1–20. Cambridge, Mass.: University of Massachusetts.

Alderete, John. 1997. "Dissimilation as local conjunction". *Proceedings of the North East Linguistic Society* XXVII, ed. Kiyomi Kusumoto, 17–32. Amherst, Mass.: GLSA Publications.

Bat-El, Outi. 1996. "Selecting the best of the worse: The grammar of Hebrew blends". *Phonology* 13.283–328.

Brame, Michael. 1974. "The cycle in phonology: Stress in Palestinian, Maltese and Spanish". *Linguistic Inquiry* 5.39–60.

Broselow, Ellen. 1982. "On predicting the interaction of stress and epenthesis". *Glossa* 16.115–132.

——. 1992. "Parametric variation in Arabic dialect phonology". *Perspectives on Arabic linguistics*, IV, ed. Mushira Eid and John McCarthy, 1–42. Amsterdam and Philadelphia: J. Benjamins.

——. Forthcoming. "Stress-epenthesis interactions in constraint-bases phonology". *Rules and constraints in contemporary phonology*, ed. Morris Halle and Bert Vaux. Oxford: Oxford University Press.

Carter, Michael G. 1973. "An Arab grammarian of the eighth century A.D.". *Journal of the American Oriental Society* 93.146–157.

Dell, François and Mohamed Elmedlaoui. 1985. "Syllabic consonants and syllabification in Imdlawn Tashilhiyt Berber". *Journal of African Languages and Linguistics* 7.105–130.

Edzard, Lutz. 1991. "Semitic phonology and preference laws for syllable structure". *Semitic studies in honor of Wolf Leslau*, ed. Alan Kaye, I, 397–410. Wiesbaden: O. Harrassowitz.

——. 1992. "The Obligatory Contour Principle and dissimilation in Afroasiatic". *Journal of Afroasiatic Linguistics* 3:3.151–171.

——. 2000. "Sībawayhi's observations on assimilatory processes and re-syllabification in the light of Optimality Theory". *Journal of Arabic and Islamic Studies* 3.48–65.

Erwin, Wallace. 1963. *A short reference grammar of Iraqi Arabic.* Washington, D.C.: Georgetown University Press.

Farwaneh, Samira. 1995. *Directionality effects in Arabic dialect syllable structure.* Ph.D. diss., University of Utah, Salt Lake City.

Graf, Dafna. 2000. "Stress assignment in the nominal system of Modern Hebrew". *Proceedings of the 15th Annual Conference of IATL*, ed. Adam Z. Wyner, 51–70. Haifa: University of Haifa.

——, and Adam Ussishkin. 2002. "Emergent iambs: Stress in modern Hebrew". *Lingua* 13:3.239–270. (Also available at Rutgers Optimality Archive ROA–446–0701, http://roa.rutgers.edu/index.php3.)

Hooper, Joan Bybee. 1976. *An introduction to Natural Generative Phonology.* New York: Academic Press.

Jastrow, Otto. 1980. "Das mesopotamische Arabisch". *Handbuch der arabischen Dialekte*, ed. Wolfdietrich Fischer and Otto Jastrow, 140–173. Wiesbaden: O. Harrassowitz.

Kager, René. 1999. *Optimality Theory.* Cambridge: Cambridge University Press.

Kenstowicz, Michael. 1981. "A note on cyclic structure in Levantine Arabic: Epenthesis and degenerate syllables in Cairene Arabic". *Theoretical issues in the grammar of Semitic languages.* (= *MIT Working Papers in Linguistics*, III), ed. Hagit Borer and Youssef Aoun 21–31. Cambridge, Mass.: MIT Press.

——. 1996. "Base-identity and uniform exponence: Alternatives to cyclicity". *Current trends in phonology: Models and methods*, ed. Jacques Durand and Bernard Laks, 363–393. Salford: European Studies Research Institute.

Kiparsky, Paul. 2003. "Syllables and moras in Arabic". *The syllable in Optimality Theory*, ed. Caroline Féry and Ruben van der Vijver, 147–182. Cambridge: Cambridge University Press.

Kusters, Wouter. 2003. *Linguistic complexity: The influence of social change on verbal inflection.* Utrecht: LOT Dissertation Series.

McCarthy, John J. 2002. *A thematic guide to Optimality Theory*. Cambridge: Cambridge University Press.

——. (ed.). 2004. *Optimality Theory in phonology: A reader*. Oxford: Blackwell.

——. 2005a. "The length of stem-final vowels in colloquial Arabic". *Perspectives on Arabic Linguistics*, XVII–XVIII, ed. Mohammad T. Alhawary and Elabbas Benmamoun, 1–26. Amsterdam and Philadelphia: J. Benjamins. (Also available at: Rutgers Optimality Archive ROA–616–0903, http://roa.rutgers.edu/index.php3.)

——. 2005b. "Optimal paradigms". *Paradigms in phonological theory*, ed. Laura Downing, Tracy Alan Hall, and Renate Raffelsiefen, 170–210. Oxford: Oxford University Press.

——, and Alan Prince. 1994. "The emergence of the unmarked: Optimality in prosodic morphology". *Proceedings of the North East Linguistic Society*, XXIV, ed. Mercè Gonzàlez, 333–379. Amherst, Mass.: GLSA Publications. (Also available at: Rutgers Optimality Archive ROA–13–0594, http://roa.rutgers.edu/index.php3.)

Mitchell, T.F. 1993. *Pronouncing Arabic*, II. Oxford: Clarendon Press.

Prince, Alan and Paul Smolensky. 1993. *Optimality Theory: Constraint interaction in generative grammar*. New Brunswick, N.J.: Rutgers University Center for Cognitive Science. (Also available at: Technical report RuCCS-TR–2. Rutgers Optimality Archive ROA–537, http://roa.rutgers.edu/index.php3.)

Rabin, Chaim 1951. *Ancient West-Arabian*. London: Taylor's Foreign Press.

Rose, Sharon and Rachel Walker. 2004. "A typology of consonant agreement as correspondence." *Language* 80.475–531. (Also available at: Rutgers Optimality Archive ROA–458–0801, http://roa.rutgers.edu/index.php3.)

Selkirk, Elizabeth. 1981. "Epenthesis and degenerate syllables in Cairene Arabic". *Theoretical issues in the grammar of Semitic languages* (= *MIT Working Papers in Linguistics*, III), ed. Hagit Borer and Youssef Aoun, 111–140. Cambridge, Mass.: MIT Press.

Vennemann, Theo. 1983. "Causality in language change: Theories of linguistic preference as a basis for linguistic explanations". *Folia Linguistica Historica* 6.5–26.

——. 1988. *Preference laws for syllable structure and the explanation of sound change*. Berlin: de Gruyter.

Watson, Janet C.E. 1999. "The syllable and syllabification in modern spoken Arabic (Ṣanʿānī and Cairene)". *The syllable: Views and facts*, ed. Harry van der Hulst and Nancy A. Ritter, 501–525. Berlin and New York: Mouton de Gruyter.

——. Forthcoming. "Syllabification patterns in Arabic dialects: Long segments and mora sharing". *Phonology*.

Zawaydeh, Bushra Adnan. 1997. "On an Optimality-Theoretic account of epenthesis and syncope in Arabic dialects". *Perspectives on Arabic linguistics*, X, ed. Mushira Eid and Robert Ratcliffe, 191–213. Amsterdam and Philadelphia: J. Benjamins.

LUTZ EDZARD (University of Oslo)

Orality

Orality refers to the ways in which information is processed and knowledge is transmitted in cultures that do not rely on the written word. The notion of orality derives primarily from work on oral-formulaic techniques of verse composition by Parry (1971) on Homeric Greek poetry and Lord (1960) on Serbo-Croatian verse epics (→ poetic koine). They describe oral texts as composed of standardized themes associated with formulaic phrases, all of which make up a poetic repertoire upon which the experienced poet-reciter can draw instead of relying on rote memorization (Ong 1982:23). Parry and Lord conclude that, although we know texts from preliterate cultures as written texts, they "may well have been, in reality, a written record of the words uttered during a single performance by a singer, performer, or narrator who was, at the same time, not reciting from memory but rather composing the work so taken down" (Zwettler 1978:4). Ong (1982) builds on work by Parry, Lord, and others to contrast orality with literacy, the more familiar way in which we process information and transmit knowledge in writing. He lists nine characteristics of orality and the oral style: it is additive rather than subordinative; aggregative rather than analytic; redundant or copious; conservative or traditionalist; close to the human lifeworld; agonistically toned; empathetic and participatory rather than objective; homeostatic; situational rather than abstract (Ong 1982:37–57).

Zwettler's (1978) study of the *Muʿallaqa* of Imruʾ al-Qays is the most detailed treatment of orality in Arabic to date. It considers the presence of formulaic techniques (1978:15–64), the infrequent occurrence of enjambment (1978:65–77), and the presence of conventional themes and thematic elements (1978:77–84) in the *Muʿallaqa*. It also examines the language of pre-Islamic poetry, → *al-ʿarabiyya*, to see whether oral-formulaic techniques can shed any light on the origins of *al-ʿarabiyya* (1978:97–188). Zwettler concludes that orality is an intrinsic aspect of classical Arabic poetry (1978:225) and of *al-ʿarabiyya* (1978:171). He admits, however, that this "does not, of itself, solve all the problems which that poetry presents" (1978:225). In an earlier work, Bateson (1970) investigates linguistic patterning and thematic divisions of five pre-Islamic *qaṣāʾid*. She finds a correspondence between language and theme but suggests that oral formulaic techniques may not tell the entire story of pre-Islamic Arabic poetry: "When called upon to recite, the poet might recite whole odes in which the passages had been carefully united to form a totality, or he might improvise long stretches at the interstices of the original design, to suit a mood or an audience" (Bateson 1970:123). A study by Monroe (1972) examines the use of formulas in samples of pre-Islamic poetry and compares it with the use of formulas by later, literate poets. He concludes that later poets used far fewer formulas than pre-Islamic poets, who relied on oral-formulaic techniques (1972:37).

Other researchers have considered the role played by orality in contemporary Arabic. Kurpershoek (1995) finds that oral transmission among Bedouin of the central Arabian Peninsula produces multiple versions of a poem in some cases. They result from variations in the order and number of verses or in vocabulary and phrasing (Kurpershoek 1995:12). In other cases, oral transmission results in a high degree of accuracy when oral texts are compared with versions recorded in writing (Kurpershoek 1995:11). Oral composition, however, relies on oral-formulaic techniques: "One might compare the Najdi oral culture to a 'do-it-yourself' store from which oral poets can draw freely in order to construct a house of their own liking" (Kurpershoek 1994:28). In fieldwork with professional epic-singers of the *Sīrat Banī Hilāl* in Egypt, Reynolds also observes techniques similar to those described by Parry and Lord: texts are composed in performance rather than sung from memory (1995:12). Sowayan portrays a similar process in a description of the *sālfih*, a genre of oral historical narrative in the northern Arabian Peninsula: "The narration of the *sālfih* is not a recital. It is a creative process" (Sowayan 1992:23). This creative process becomes more complex when the *sālfih* includes poetry. The presence of a *giṣīdih*, a line or more of poetry, within the *sālfih* may require commentary. Because the *giṣīdih* and the *sālfih* are so closely connected, a change in the wording of the *giṣīdih* often entails a change in the *sālfih* (Sowayan 1992:25–26).

Not all contemporary genres, however, present clear-cut examples of oral composition. Johnstone (1991) looks at repetition, parallelism, and paraphrase in written persuasive discourse by several authors. Although she finds in that discourse a number of linguistic features identified with orality, she declines to characterize the language and style used by well-educated, literate authors as oral (Johnstone 1991:114).

Even oral genres show variation in the degree to which oral composition occurs. Caton's work on oral poetry in Yemen examines several genres to describe a range of compositional conventions (1990:256–257). At one end is the *bālah*, a competition in which poets compose and recite before an audience that evaluates their verses (1990:80). At the other is the *qaṣīdah* in its distinctly Yemeni form, which is composed in advance of performance (1990:257) and may be recorded in writing or on tape to preserve and distribute it (1990:188). In a study of Nabaṭi poetry, oral poetry of Arabia, Sowayan finds that poets rely on "conventional themes, stock motifs, topics, and formulas" (1985:95). They do not, however, compose in performance. Performance is, instead, recitation of a memorized text (Sowayan 1985:111).

One aspect of orality that has been the subject of little research to date is 'secondary orality' in Arabic. Ong coined the term to refer to oral communication taking place through electronic technology, including telephone, radio, television, and sound recording (Ong 1982:136). It has, however, come to refer to all kinds of electronically mediated communication, whether that communication relies on sound or sight. Brief studies examine the orthography used by speakers of Arabic when communicating electronically (Berjaoui 2001; Palfreyman and al Khalil 2003). Another (Belnap and Bishop 2003) includes chat room communication as a type of personal correspondence in a consideration of the increasing use of spoken Arabic in domains formerly dominated by Modern Standard Arabic. In-depth research on the effects of secondary orality on electronically mediated Arabic language use, and especially on discourse, is still to come.

Bibliographical references

Bateson, Mary Catherine. 1970. *Structural continuity in poetry: A linguistic study of five pre-Islamic Arabic odes*. Paris and The Hague: Mouton.

Belnap, R. Kirk and Brian Bishop. 2003. "Arabic personal correspondence: A window on change in progress?". *International Journal of the Sociology of Language* 163.9–25.

Berjaoui, Nasser. 2001. "Aspects of the Moroccan Arabic orthography with preliminary insights from the Moroccan computer-mediated communication". *Chat-Kommunikation: Sprache, Interaktion, Sozialität und Identität in synchroner computervermittelter Kommunikation. Perspektiven auf ein interdisziplinäres Forschungsfeld*, ed. Michael Beißwenger, 431–465. Stuttgart: Ibidem.

Caton, Steven. 1990. *'Peaks of Yemen I summon': Poetry as a cultural practice in a North Yemeni tribe*. Berkeley: University of California Press.

Johnstone, Barbara. 1991. *Repetition in Arabic discourse: Paradigms, syntagms, and the ecology of language*. Amsterdam and Philadelphia: J. Benjamins.

Kurpershoek, P. Marcel. 1994. *Oral poetry and narratives from central Arabia*. I. *The poetry of Ad-Dindān: A Bedouin bard in southern Najd*. Leiden: E.J. Brill.

——. 1995. *Oral poetry and narratives from central Arabia*. II. *The story of a desert knight: The legend of Šlēwīḥ al-ʿAṭāwi and other ʿUtaybah heroes*. Leiden: E.J. Brill.

Lord, Alfred B. 1960. *The singer of tales*. Cambridge, Mass.: Harvard University Press.

Monroe, James. 1972. "Oral composition in pre-Islamic poetry". *Journal of Arabic Literature* 3.1–53.

Ong, Walter J. 1982. *Orality and literacy: The technologizing of the word*. London and New York: Methuen.

Palfreyman, David and Muhamed al Khalil. 2003. " 'A funky language for teenzz to use': Representing Gulf Arabic in instant messaging". *Journal of Computer-Mediated Communication* 9:1. http://jcmc.indiana.edu/vol9/issue1/palfreyman.html (accessed 13 June 2007).

Parry, Milman. 1971. *The making of Homeric verse: The collected papers of Milman Parry*, ed. Adam Parry. Oxford: Clarendon Press. (Originally published in 1928.)

Reynolds, Dwight Fletcher. 1995. *Heroic poets, poetic heroes: The ethnography of performance in an Arabic oral tradition*. (Myth and Poetics). Ithaca, N.Y., and London: Cornell University Press.

Sowayan, Saad Abdullah. 1985. *Nabaṭi poetry: The oral poetry of Arabia*. Berkeley, Los Angeles, and London: University of California Press.

——. 1992. *The Arabian oral historical narrative: An ethnographic and linguistic analysis*. Wiesbaden: O. Harrassowitz.

Zwettler, Michael. 1978. *The oral tradition of classical Arabic poetry: Its character and implications*. Columbus: Ohio State University Press.

ELIZABETH M. BERGMAN (Georgetown University)

Ottoman Empire

Arabic was one of the *elsine-i s̱elās̱e* 'the three languages', which constituted the basis of Ottoman elite culture. Thus, the Ottoman Empire can be described as part of a historical space-time characterized by the use of Arabic as a means of communication, which in analogy to Fragner's concept of 'Persophonie' can be named 'Arabophonia' (Fragner 1999). In the Ottoman Empire, in addition to its function as an ethnic language in the Arab provinces, Arabic was cultivated as the language of Islam and, more broadly, as a "written lingua franca in the Islamic world" (Richard Frye as quoted in Fragner 1999:32). Since Arabic in the Ottoman context has not been studied in terms of sociolinguistics and language ideologies, only a preliminary outline can be given.

The position of Anatolia and the Balkans within the Arabophonia can be divided into three phases: the emergence of Turkish as a literary language in competition with Arabic (and Persian) until the late 15th century, the period of Arabic as a transnational language of Islam until the early 19th century, and finally the recognition of Arabic as an ethnic language. The first phase is characterized by ongoing Islamization and the lack of a stable network of institutions and of Arabophone readers in the Ottoman and other principalities in Anatolia. The result was a certain fluidity between the spheres of Arabic and Turkish, due to which certain genres, such as legal and religious texts, were incorporated in the Turkish domain, which later became exclusively part of the Arabic sphere. The most prominent example is the translation of the *Qurʾān* made around 1405, entitled *Jewāhiru l-aṣdāf* (Zajączkowski 1937; Topaloğlu 1983). For other examples, see Adamovic (1990), Yavuz (1983), and İ. Fazlıoğlu (2003).

From the late 15th century onward, the Ottoman state acquired an increasingly Sunni-Hanafi character. The corresponding interest in the Islamic territories resulted in the conquest of Syria (1514) and Egypt (1517) and subsequent easy access to Arabic literature. Large numbers of manuscripts were transferred to Istanbul. Out of the 14,500 titles recorded in Ḥājjī Xalīfa's bibliographical dictionary, 95 percent are estimated to be in Arabic (Bilge and Yaltkaya 1941–1943:15; for a slightly different estimate, see Birnbaum 1994:154). At the same time, a unified system of higher education came into being as a venue for the study of a canon of scholarly texts in Arabic, and as a place to study Arabic. Subsequently, a stable division between the domains of Arabic and Turkish is characteristic for the second phase of Ottoman Arabophonia.

The hybrid character of Ottoman culture blurred the distinction between Arabic and Persian as distinct languages, on the one hand, and as components of Classical Ottoman, on the other, as is evident in a statement by the poet and historian Muṣṭafā ʿĀlī (d. 1600): "In fact, the astonishing language current in the state of Rūm, composed of four languages [West Turkish, Çağatay, Arabic, and Persian] is a pure gilded tongue which, in the speech of the literati, seems more difficult than any of these. If one were to equate speaking Arabic with a religious obligation [*farẓ*], and the use of Persian with a sanctioned tradition [*sünnet*], then the speaking of a Turkish made up of these sweetnesses [Arabic and Persian] becomes a meritorious act [*müsteḥabb*]" (Fleischer 1986:253ff.; Turkish text in İ. Fazlıoğlu 2003:160). The status of Arabic as a transnational language as opposed to Turkish as an ethnic language is illustrated by the following dialogue between a Turkish and an Iranian scholar in the course of a scholarly discussion of an Arabic text: "You don't know Turkish because you're not a Turk" – "Thank God for that!" (*ʾanta lā taʿrif bi-t-turkiyya li-ʾannaka lasta min at-Turk – al-ḥamdu li-llāhi ʿalā ḏālik*; İ. Fazlıoğlu 2003:181).

Lexicography reflected the twofold status, serving the understanding of Arabic texts as well as the expansion of the vocabulary of Ottoman Turkish. Two Turkish Qurʾānic glossaries from the 15th century prepared the ground for the second phase of Arabophonia (Muhtar 1993). Axtarī's (d. 968/1560) dictionary was an original achievement (*Türkiye Diyanet Vakfı İslam Ansiklopedisi*, s.v. Ahterî), and al-Jawharī's *Ṣiḥāḥ* was translated into Turkish several times, the most popular translation being that by Vānkulī in the late 10th/16th century. It was to be the only Ottoman book printed three times in the course of the 18th century (*İslam Ansiklopedisi*, s.v. Vankulu).

Especially since the classical period Ottomans not only consumed but also produced literature in Arabic. The option for one of the 'three languages' is not self-evident but requires analysis in context. Ottoman authors tended to justify their writing in Turkish (numerous examples in İ. Fazlıoğlu 2003), while apparently they regarded Arabic as the natural choice. Ottoman authors preferred Arabic for certain genres of texts, especially in prose. In keeping with the inherited canons of law, theology, philosophy, science, and Arabic philology as taught in the *medrese*s, Ottomans wrote such works preferably in Arabic (e.g. see Atsız 1966–1972 for a list of 183 works in Arabic by the *ʿālim* and historian Ibn Kemāl Paša [d. 940/1534]). Arabic was the preferred language for general as well as biographical, bibliographical, or geographical encyclopedias, such as works by Ṭašköprüzāde (d. 1561), Ḥājjī Xalīfa (d. 1657), and Sipāhīzāde (d. 1588). While chronicles of the Ottoman dynasty were written in Turkish, there are several prominent examples of world histories written in Arabic, including works by Jenābī (d. 1590), Ḥājjī Xalīfa, and Münejjimbašı (d. 1702). Ottomans studied poetry in Persian and Arabic and usually began writing poetry in imitation of canonical examples (Flemming 1977:16). Yet, while many Ottoman authors wrote substantial poetic works in Persian, the 465 verses in Arabic in Fuẓūlī's (d. 1555) *Dīwān* seem to have been an exception, and these verses were not regarded as highly as his other poetry or his Arabic prose (Demirel 1991). Deeds of larger pious foundations, as texts of legal importance as well as literary quality, were regularly drafted in Arabic (Kaleši 1972; Uzunçarşılı 1941; Schwarz and Winkelhane 1986) or existed in parallel versions (Singer 2002:44–46).

Modern scholars have offered explanations of the language choices of Ottoman writers, but only in limited contexts. Given the examples above (and below), geographical reasons cannot account for the choice of Arabic. Likewise, the status of Arabic as the language of the Islamic Revelation and the connection of Arabic to the sphere of religion do not suffice as explanations of language choice (see Ṭašköprüzāde, *Miftāḥ* I, 86, for the religious and literary merits of Arabic). Both the 15th-century *ṣūfī* Yazıjıoğlı Meḥmed and the 16th-century preacher Birgivī wrote their works in Arabic, before they prepared popular versions in Turkish. In the widespread genre of 'Forty *ḥadīṯ*' collections, it is possible to distinguish Ottoman works of theological and doctrinal content in Arabic, and didactic and edifying works in Turkish (see Karahan 1991). İ. Fazlıoğlu (2003:154) suggests that in the works of Ibn Sīnā, al-Ġazālī, Ṭūsī, and others, Arabic had been developed into a ‘semisymbolic’ language of formalized logic appropriate for the description of the exterior world and used as such in the Ottoman world as well. This concept may explain why many extant Turkish translations of Arabic texts do not simply transfer but paraphrase and interpret the original Arabic. It may be objected that the function of a symbolic language hinges on the vocabulary and thus could have been fulfilled by Classical Ottoman after its wholesale incorporation of Arabic vocabulary (Arabic loanwords in → Turkish). In his discussion of the works of Ḥājjī Xalīfa, Birnbaum (1994:150ff.) points out that works in Arabic could expect a wider audience in the Islamic world as a whole. However, there is no indication that Ḥājjī Xalīfa ever thought of an audience beyond the Ottoman elite of his time. Statements about prospective readership tend to be based on genre, which therefore should be the starting point for a general explanation. The examples gathered so far suggest that Ottomans wrote in Arabic about topics of universal validity and relevance, while issues of local or regional interest were dealt with in Turkish. Many such universal topics were typically dealt with in the *medrese*s, while little literature in Arabic emerged from the dervish lodges. This hypothesis would explain the widespread use of translations as the application of universal knowledge to local contexts. Numerous Ottoman authors wrote in Arabic first and then produced a Turkish version of the same text (in addition to Jenābī’s and Sipāhīzāde’s works mentioned above, see Ḥasan Kāfī Akḥiṣārī’s [d. 1616] treatise on ethics and politics).

Although Turkish works frequently claimed to have been written for an audience without knowledge of Arabic, it can be assumed that the Ottoman elite in general had at least reading proficiency, so that Arabic could not be instrumentalized to monopolize knowledge. Ottoman literacy was based on education in the *mekteb*, the elementary school, in which the recitation of the *Qurʾān* was taught as a ritual practice. However, proficiency in Arabic was not a goal of the *mekteb* (Ergin 1939; Somel 2001). Ṭašköprüzāde’s recommendation to use only Arabic in sermons was only theoretical, and allowed exceptions “in case of need” (*Miftāḥ* I, 56; see Flemming 1977:8). It was primarily the *medrese* as the institution of higher education which integrated Ottomans into Arabophonia. Instruction in the *medrese* was in Turkish, but the required texts were entirely in Arabic (lists of textbooks in İzgi 1997:I, 61–108; Ş. Fazlıoğlu 2003; Şeşen 1986). In the lower ranks of the *medrese*, the topics of → *ṣarf* ‘phonology/morphology’, *naḥw* ‘syntax’, and *maʿānī* ‘semantics’ made up a considerable part of the curriculum. According to Şeşen (1986:272), tables of morphological paradigms (often in anonymous works with contemporary commentaries) have been added since the mid-16th century (see the descriptions in Quiring-Zoche 1994, 2000). This would indicate an increasing consciousness of the difficulties of language education.

Training provided in the palace school for military and administrative officials was not substantially different than in the *medrese*. Ortaylı (1986:194) raises doubts regarding the proficiency of Ottoman officials. Mistakes in Ḥājjī Xalīfa’s Arabic, as noted by Bilge and Yaltkaya (1941–1943:I, 15), were due more to negligence than ignorance. On the other hand, a decree by Süleymān I in 1553 to provide Turkish summary translations of all previous sultans’ *waqfıye*s (Uzunçarşılı 1941:550) indicates that clerks in the imperial chancery had difficulty reading Arabic documents, the administrative language being Turkish. In exceptional cases, Arabic was used by the central chancery in internal correspondence (Veinstein 2001:427, pointing out that Arabic documents were registered in a different script in the *mühimme defterleri*). Local administration in the Arab provinces often involved the local language, as attested by extant court records (*Encyclopaedia of Islam*, 2nd ed., s.v. *Sid̲j̲ill*). While there is some evidence of translators being used in courts in different parts of the empire, so far translators in the service of governors and their employees (*beglerbegi, ṣanjaqbegi, ṣubašı*) have been found in the Arab provinces only (Veinstein 2001:434–437; see also *İslam Ansiklopedisi*, s.v. *Tercüman*). It is unknown whether this was for pragmatic reasons or for ideological reasons linked to the status of Arabic.

From the 12th/18th century onward, the Ottoman literate public extended more and more beyond the *medrese*-trained elite, as is indicated by the number of newly founded libraries. Thus, a growing section of Ottoman society remained outside Arabophonia. Numerous new translations from Arabic into Turkish, including theological works as well as Müterjim Aḥmed ʿĀṣim's translation of Fīrūzābādī's *Qāmūs*, entitled *el-Oqyānūs* (1815), have to be seen in this context.

The emphasis on ethnolinguistically defined identities among the subjects of the Ottoman Empire in the 19th century shaped the last phase of Ottoman Arabophonia in contradictory ways. Since the *Tanẓīmāt* era, the central government resorted to an ideology of a transnational Islamic Ottoman identity. Thus, it maintained the teaching of Arabic as a central part of the school curriculum throughout the numerous stages of the reform of public education (Somel 2001). In the Hamidian period (1876–1909), the Arabic element was consciously supported as the language of the caliphate and as the language of the largest non-Turkish ethnic group. On the other hand, attempts at centralization and tighter control intensified the contacts between authorities and subjects, and thus exacerbated the problem of ethnic languages. Government policies to promote the use of Turkish culminated in Article 18 of the Constitution of 1876, soon suspended but reinstated in 1908: "A prerequisite for Ottoman subjects' employment in State service is that they know Turkish, which is the official language of the State [*devletiñ lisān-i resmīsi*]" (quoted in Lewis 1999:16; see Prätor 1993b). Unification as a corollary of Ottoman modernization brought together the spheres of Arabic as an ethnic language and Arabic as a semisymbolic transnational language, and revealed the discrepancies between them. Typically contradicting attitudes are reflected in a debate in Parliament in 1916, during which a Turkish member insisted that Arabic was not a foreign language, while the Minister of Education sought to distinguish Eastern from Western foreign languages, subsuming Arabic together with Persian, Urdu, and Hindi under the former (Prätor 1993a:266). The Senate received numerous petitions in Arabic and debated about having them translated officially, but translators were available only for French (Prätor 1993c:217).

The concept of Arabic as a "ready-made Esperanto of the Orient", as a journalist expressed it in 1911, was not tenable: any use of Arabic had to be aligned with the standards of contemporary living Arabic. Numerous Turkish memoirs from the 19th century expressed frustration that the training in the *mekteb* did not provide communicative skills in Arabic (Prätor 1993a:264; see also Ergin 1939). Complaints about Turkish teachers of Arabic unable to speak and teach the language were common in the Arab provinces (Somel 2001:205, 218). Learning Arabic according to modern methods was promoted in Young Turk circles as well as by the government; one Young Turkish member of Parliament published a newspaper in Arabic in Istanbul (*al-ʿArab*; see Prätor 1993a).

Language policies of the central government caused tensions in the Arab provinces which were not resolved before the end of the Ottoman Empire. The requirement of Article 68 in the Constitution of 1876/1908 that members of Parliament had to be proficient in Turkish was never enforced (Prätor 1993b:133). On the other hand, ignorance of Arabic among Ottoman officials was a general complaint (Prätor 1993c:164–183). The unification of the law created a law code that was exclusively in Turkish (the so-called *mejelle*), but the law code pertaining to property (*ʾarāẓī qānūnnāmesi*) had been translated into Arabic, so that both versions were valid in court (Ortaylı 1986:195). Debates about Arabic as an official language in the Arab provinces occurred first in 1877 (Ortaylı 1986:196), but a directive in 1909 making Turkish the exclusive language of courts gave rise to widespread complaints about possible errors in rulings due to linguistic problems. Arabic was officially admitted as a language in court in 1913. The degree to which this was implemented is unknown. The same decree introduced Arabic as the language of instruction in the schools in the Arab provinces (Prätor 1993b).

While the Young Turk era had begun to reunite the different spheres of use of Arabic, it was the loss of the Arab provinces in World War I and Kemalist secularism that ultimately excluded the Ottoman world from Arabophonia.

Bibliographical references

Primary source

Ṭašköprüzāde, *Miftāḥ* = Ṭašköprüzāde Aḥmed, *Miftāḥ as-saʿāda wa-miṣbāḥ as-siyāda*. Hyderabad: Dāʾirat al-Maʿārif an-Niḏāmiyya, 1328–1357 A.H.

Secondary sources

Adamovic, Milan. 1990. *Die Rechtslehre des Imam an-Nasafi in türkischer Bearbeitung vom Jahre 1332*. Stuttgart: F. Steiner.

Atsız, [Hüseyn Nihal]. 1966–1972. “Kemalpaşaoğlu’nun eserleri [The works of Kemalpaşazâde]”. *Şarkiyat Mecmuası* 6.71–112, 7.83–135.

Bilge, Kilisli Rifat and Şerefettin Yaltkaya (eds.). 1941–1943. *Kâtib Çelebi: Keşf-el-Zunun*. Istanbul: Maarif Matbaası.

Birnbaum, Eleazar. 1994. “The questing mind: Kātib Chelebi, 1609–57: A chapter in Ottoman intellectual history”. *Corolla Torontonensis: Studies in honor of Ronald Morton Smith*, ed. Emmet Robins and Stella Sandahl, 133–158. Toronto: TSAR.

Demirel, Hamide. 1991. *The poet Fuzûli: His works, study of his Turkish, Persian and Arabic divans*. [Ankara]: Ministry of Culture.

Ergin, Osman. 1939. *İstanbul mektepleri ve ilim, terbiye ve san’at müesseseleri dolayısile Türkiye maarif tarihi* [A history of education in Turkey, dealing with primary schools and institutions of science, education, and art]. Istanbul: Osmanbey.

Fazlıoğlu, İhsan. 2003. “Osmanlı döneminde ‘bilim’ alanındaki Türkçe telif ve tercüme eserlerin Türkçe oluş nedenleri ve bu eserlerin dil bilincinin oluşmasındaki yeri ve önemi [Why certain scientific works were written in Turkish in the Ottoman period, and the importance of these works in the development of language consciousness]”. *Kutadgubilig: Felsefe-Bilim Araştırmaları* 3.151–184.

Fazlıoğlu, Şükran. 2003. “Nebī Efendi-zade’nin ‘Kaṣīde fî el-kutub el-meşhūre fî el-ʿulūm’ una göre bir medrese talebesinin ders ve kitab haritası [A *medrese* student’s classes and textbooks according to Nebī Efendi-zāde’s ‘Kaṣīde fî el-kutub el-meşhūre fî el-ʾulūm’]”. *Kutadgubilig: Felsefe-Bilim Araştırmaları* 3.191–221.

Fleischer, Cornell. 1986. *Bureaucrat and intellectual in the Ottoman Empire: The historian Mustafā Âlî (1541–1600)*. Princeton, N.J.: Princeton University Press.

Flemming, Barbara. 1977. *Die Sprachen der türkischen Dichter*. Leiden: E.J. Brill.

Fragner, Bert. 1999. *Die “Persophonie”: Regionalität, Identität und Sprachkontakt in der Geschichte Asiens*. Berlin: Das Arabische Buch.

İzgi, Cevat. 1997. *Osmanlı medreselerinde ilim* [The sciences in the Ottoman *medrese*]. Topkapı, Istanbul: İz.

Kaleši, Hasan. 1972. *Najstariji vakufski dokumenti u Jugoslaviji na arapskom jeziku* [The oldest *waqf* documents of Yugoslavia in Arabic]. Pristina.

Karahan, Abdülkadir. 1991. *İslâm-Türk edebiyatında kırk hadîs* [‘Forty *ḥadīṯ*’ works in Turkish-Islamic literature]. Ankara: Diyanet İşleri Başkanlığı.

Lewis, Geoffrey. 1999. *The Turkish language reform: A catastrophic success*. Oxford: Oxford University Press.

Muhtar, Cemal. 1993. *İki Kurʾan sözlüğü: Lugat-ı Ferişteoğlu (Abdullatif Ibn Melek) ve Lugat-ı Kânûn-ı İlâhi (Abdulmecid Ibn Ferişte)* [Two *Qurʾān* glossaries: Lugat-ı Ferişteoğlu (Abdullatif Ibn Melek) and Lugat-ı Kânûn-ı İlâhi (Abdulmecid Ibn Ferişte)]. Istanbul: Marmara Universitesi İlahiyat Fakültesi Vakfı.

Ortaylı, İlber. 1986. “The use of Arabic in Ottoman administration during the 19th century”. *Studies on Turkish-Arab Relations Annual* 193–197.

Prätor, Sabine. 1993a. “Sprache der Religion und Wissenschaft: Anmerkungen zum Arabischlernen im spätosmanischen Istanbul”. *Zeitschrift für Arabische Linguistik* 25.260–269.

——. 1993b. “Staatssprache Türkisch: Zu den Auswirkungen einer Verfassungsbestimmung in der ‘Zweiten Konstitution’”. *Sprach- und Kulturkontakte der türkischen Völker*, ed. Jens Peter Laut and Klaus Röhrborn, 133–138. Wiesbaden: O. Harrassowitz.

——. 1993c. *Der arabische Faktor in der jungtürkischen Politik: Eine Studie zum osmanischen Parlament der II. Konstitution (1908–1918)*. Berlin: K. Schwarz.

Quiring-Zoche, Rosemarie. 1994. *Arabische Handschriften*. III. *Verzeichnis der orientalischen Handschriften in Deutschland XVII, B, 3*. Stuttgart: F. Steiner.

——. 2000. *Arabische Handschriften*. V. *Verzeichnis der orientalischen Handschriften in Deutschland XVII, B, 5*. Stuttgart: F. Steiner.

Schwarz, Klaus and Gerd Winkelhane. 1986. *H̱oğa Sa’deddin, Staatsmann und Gelehrter (gest. 1599), und seine Stiftung aus dem Jahr 1614*. Bamberg: aku.

Şeşen, Ramazan. 1986. “Osmanlılar döneminde Arap dili ve edebiyatı öğretimi [The teaching of Arabic language and literature in the Ottoman period]”. *Studies on Turkish-Arab Relations Annual* 267–278.

Singer, Amy. 2002. *Constructing Ottoman beneficence: An imperial soup kitchen in Jerusalem*. Albany: State University of New York Press.

Somel, Selçuk Akşin. 2001. *The modernization of public education in the Ottoman Empire, 1839–1908: Islamization, autocracy and discipline*. Leiden: E.J. Brill.

Topaloğlu, Ahmed. 1983. “Kur’an-ı Kerîm’in ilk Türkçe tercümeleri ve Cevâhirü’l-Asdâf [The first Turkish translations of the *Qurʾān* and *Jawāhir al-ʾaṣdāf*]”. *Türk Dünyası Araştırmaları* 27.58–66.

Uzunçarşılı, İsmail Hakkı. 1941. “Çandarlı Zâde Ali Paşa Vakfiyesi [The *waqf* document of Çandarlızade Ali Paşa]”. *Belleten* 5.549–576.

Veinstein, Gilles. 2001. “L’administration Ottomane et le problème des interprètes”. *Pax Ottomana: Studies in memoriam Prof. Dr. Nejat Göyünç*, ed. Kemal Çiçek, 425–441. Haarlem: SOTA; Ankara: Yeni Türkiye.

Yavuz, Kemal. 1983. “XIII–XVI. asır dil yadigarlarının Anadolu sahasında Türkçe yazılış sebebi ve bu devir müelliflerinin Türkçe hakkında görüşleri [Reasons for writing language monuments of the 13th–16th century in Anatolia in Turkish, and opinions of authors of the period about Turkish]”. *Türk Dünyası Araştırmaları* 27.9–57.

Zajączkowski, Ananiasz. 1937. *Studia nad jezykem staroosmanskim*. II. *Wybrane rozdzialy z anatolijskotureckiego przekladu Koranu* [Studies on Ancient Ottoman. II. Selected chapters from the translation of the *Qurʾān* in Anatolian Turkish]. Cracow: Nakladem Polskiej Akademii Umietnośći, 1937.

Gottfried Hagen
(University of Michigan)

P

Pakistan

1. The history of Arabic in South Asia

Pakistan is a multilingual country with 6 major languages (see Table 1), and 69 languages in all (Grimes 2000:588–598). Urdu is the national language, but it is English, the ex-colonial language, which is used in the higher domains of power – government, military, higher education, judiciary, commerce, research, and media.

Table 1. Languages spoken in Pakistan

Language	Percentage of speakers	Number of speakers
Punjabi	44.15	66,225,000
Pashto	15.42	23,130,000
Sindhi	14.10	21,150,000
Siraiki	10.53	15,795,000
Urdu	7.57	11,355,000
Balochi	3.57	5,355,000
Others	4.66	6,990,000

Source: Census 2001, Table 2.7. (Population assumed to be 150 million in 2003 as it was 132,352,000 in 1998, the growth rate being 2.69%)

Arabic is not one of the indigenous languages of Pakistan, although some people claim Arabic as a mother tongue and others use it as an additional language (see Tables 2 and 3). Arabic periodicals are published, although they have a very small circulation (see Table 4). Since most Pakistanis (96.16%, according to Census 1998:107) are Muslim, Arabic is a symbol of religious identity and the liturgical language of the country. Because of its symbolic and affective significance, Arabic has always been an important part of the education of Pakistani and, indeed, South Asian Muslims.

Arabic came to South Asia with the Arab merchants who traded on the Malabar Coast and Sri Lanka, and even settled down there in due course, before the Arabs conquered Sindh in 711–712 C.E. (Nadwī 1972:69–71, 259–301; Kōken 1974). Some Arab historians claim that the languages of Mansura and Multan were Arabic and Sindhi (Ibn Ḥawqal, *Masālik*). It is possible that the Arabs who settled in Sindh and some parts of the Baluchistan coast after Muḥammad ibn Qāsim's conquest used Arabic, especially after the Umayyad caliph Walīd (r. 705–715) substituted Arabic for Greek in parts of his empire overlapping or adjacent to present-day Pakistan (Elliott 1867:I, 461). It is also possible that, because of frequent

Table 2. Speakers of Arabic in Pakistan (in millions)

	1951*	1961**	2004
Mother tongue	1,249	3,398	data not available
Additional language commonly spoken	63,794	250,522	data not available

N.B.: Detailed data on languages was not compiled after 1961.

* Census 1951, Tables 7 and 7-A (population 75,635,496)

** Census 1961, Statement 5.1 and 5.2 (population 90,282,674)

Table 3. Literacy in Arabic (in percentages)

Year	Population	Ability to read and write Arabic	Ability to speak Arabic fluently	Ability to read Qurʾānic Arabic without understanding	General literacy
1951	75,842,000	0.23 (out of population aged 5 and above)	0.09	10.5 (out of Muslims)	22 (out of population aged 5 and above)
1961	93,720,613	0.46 (out of population aged 5 and above)	0.28	6.29 (out of total population aged 5 and above)	19.2
1981	84,254,000	No information	No information	38.37 (out of Muslims aged 5 and above)	26.17 (aged 10 and above)
1998	132,352,000	No information	No information	55.35 (Muslims aged 5 and above)	43.92 (aged 10 and above)

Source: Census reports 1951, 1961, 1981, and 1998.
NB: 'Literacy' was defined in each census as follows:

Census 1951	The ability to read any language in clear print (even without understanding)
Census 1961	The ability to read with understanding a short statement on everyday life in any language
Census 1981	The ability to read and write with understanding
Census 1998	The ability to read a newspaper and write a simple letter

Table 4. Publications in Arabic

		Numbers			Circulation	
Year	Arabic newspapers and periodicals	Total	Percentage of Arabic out of total	Circulation of Arabic publications	Total circulation in all categories	Percentage of Arabic out of total
1999	2	1,571	0.13	1,750 (biweeklies)	7,310,986	0.092
				5,000 (monthlies)		
2000	4	815	0.49	1,750 (biweeklies)		
				5,000 (monthlies)	7,458,662	0.09
2001	4	763	0.52	5,100 (dailies)	7,589,136	0.067
2002	4	720	0.55	5,100 (dailies)	7,976,177	0.064
2003	4	945	0.42	5,350 (dailies)	8,250,635	0.065

Source: Audit Bureau of Circulation, Ministry of Information and Broadcasting, Govt. of Pakistan, May 2004.

intercourse with the Arabs, members of the elite picked up verses or bon mots from Arabic and could quote them on occasion (Mubarakpurī 1989:315–316).

The northern part of Pakistan was conquered by the Ghaznavid Turks in 1026, and there are references to Arabic being used by them in some domains of the state (Munawar 1972:27–28). It was taught in the *maktab*s (schools) and the *madrasa*s (colleges), and religious personalities wrote their works, including letters and the *Xilāfat nāma*s of mystics (*ṣūfī*s), in it (Shīranī 1929:83). Among the Arabic books taught during Ghias Uddin Balban's (r. 1266–1287) time were canonical works of grammar from the 13th century onward (see Table 5), some of which are still taught in South Asian *madrasa*s. The prose classic was the *Maqāmāt* of al-Ḥarīrī of Basra (1054–1122), which recounted the picaresque, somewhat risqué, adventures of ʾAbū Zayd from Sarūj, which was valued highly as a model of elegant Arabic prose. The *Maqāmāt* were often memorized, Sheikh Niẓām ad-Dīn Auliā (1234–1324) having memorized forty chapters (*maqāma*s; Hai 1947:I, 163). It was also emulated; Sayyid ʾAbū Bakr al-ʿAlawī wrote *al-Maqāmāt al-Hindiyya* (1715), in which the protagonist, Abul Zafar al-Hindi, plays out his adventures in South Asian cities such as Surat, Ahmadnagar, and Lahore (Ahmad 1946:225–228).

Even as writing in Arabic decreased, the Indian *ulema* continued to produce works in Arabic, as described in the *Nuzhat al-xawāṭir* (Hai 1947) and other sources (Ahmad 1946). For example, the famous Muslim reformer Shah Walīullāh (1703–1762) wrote his magnum opus, *Ḥujjat Allāh al-bāliġa* (published 1869), in Arabic. Even in the 18th century, when European learning appeared in India, Tafuzzul Husain (1727–1800) translated Newton's *Principia* (1687) and other scientific works into Arabic, on the assumption that Arabic would remain the language of science in the Islamic world including India (Khan 1998:274).

Arabic was a much more important part of Muslim education before the decline of the Mughal Empire than after it. Not only the *ulema* but also others, like Sher Khan (r. 1545–1553), who wrested away the Mughal Empire from Humayun (d. 1556), studied books on Arabic grammar (Sarwāni 1586:9). The emphasis on Arabic grammar and literature may have been more than pragmatic people realized; Aurangzeb Alamgir (r. 1658–1707), the most Islamic minded of the Mughal emperors, is said to have reprimanded his former teacher for having given far more attention to Arabic than to foreign languages (Bernier 1826:176–177). Later, the Arabic script (→ *nasx*) remained part of the traditional course of studies of a Muslim gentleman, and even women, while being denied literacy in other languages, were taught how to read the *Qurʾān*, though without understanding. The pupils merely learned to recognize the Arabic alphabet before going on to study Persian, a marker of elitist identity as well as the language of upward social mobility, the script of which (→ *nastaʿlīq*), was based on the Arabic script.

When the British arrived, they found elementary schools (*maktab*s) teaching basic literacy in Persian and Arabic, Persian schools teaching some Arabic and advanced Persian literature, and Arabic schools teaching religious subjects through Arabic with explanation in Persian. Initially, they followed the Orientalist policy of retaining the traditional system of study, and they learned Persian as well as Arabic themselves. Thus, the first educational institution established by the British in India in 1781, the Calcutta Madrassah, taught the usual texts in Arabic language and literature. However, the British added texts like the *ʾAlf layla wa-layla* 'The Arabian nights' and *Nafḥat al-Yaman*, the latter written by ʾAḥmad al-Yamanī (ca. 1820s). These texts were also taught in the secular educational institutions created by the British, where Arabic was an optional subject.

2. Arabic in the Islamic seminaries

The *madrasa*s, usually based on an endowment (*waqf*), had served primarily as Islamic seminaries since the 11th century C.E., when they were created to teach the tenets of Islam (Makdisi 1981:36–38). They taught Arabic grammar, literature, and rhetoric, again through canonical works and commentaries, all written in Arabic but sometimes with explanations in Persian. Different teachers taught different books (Sufi 1941:68–70), and there was no standardized curriculum until the Dars-i Nizāmī was created

Table 5. Arabic texts used in the *madrasa*s

Morphology *(ṣarf)*

Ḥāfiẓ Muḥammad ibn Barak Allāh Laxwī (d. 1893), *ʾAbwāb aṣ-ṣarf.*
Muḥammad ibn ʿAbd aṣ-Ṣamad al-Harīsī, known as Sheikh Bahāʾ ad-Dīn (1546–1621), *Ṣarf-i Bahāī.*
ʿAlī ʾAkbar ʾAllāhābādī (d. 1680), *Fuṣūl-i ʾAkbarī*. ʿAbd al-Karīm (n.d.), *Iršād aṣ-ṣarf.*
ʾAḥmad ibn ʿAlī Ibn Masʿūd (n.d.), *Marāḥ al-ʾarwāḥ*.
Mīzān aṣ-ṣarf. Author unknown (Sheikh Sāʿdi, Ḥamza Badāyūnī, and Muḥammad ibn Muṣṭafā [d. 1505–1506] are named in different sources).
Ḥamzā Badāyūnī, *Munšaʿib*, n.d. (Sirāj ad-Dīn al-ʾAwaḏī is also named in some sources as the author.)
Jamʿ al-Muqaddimāt, n.d. (Two works on *ṣarf, ʾAmṯila* and *Šarḥ ʾAmṯila*; taught in Shīʿī *madrasa*s.)
Panj Ganj. Probably by Sirāj ad-Dīn ʿUṯmānī al-ʾAwaḏī (d. 1356).
ʿAlī ibn Muḥammad Mīr Sayyid aš-Šarīf al-Jurjānī (1339–1413), *Ṣarf-i Mīr.*
Jamāl ad-Dīn ʾAbū ʿAmr ʿUṯmān ibn ʿUmar ibn al-Ḥājib (1174–1248), *aš-Šāfiya.*
ʿInāyat ʾAḥmad Kākōrvī (1812–1863), *ʿIlm aṣ-ṣīġa*, 1859.

Syntax *(naḥw)*

Bahāʾ ad-Din ʿAbdallāh ibn ʾAḥmad Ibn ʿAqīl (1298–1367), *Šarḥ Ibn ʿAqīl.*
Sirāj ad-Dīn al-ʾAwaḏī (or ʾAbū Ḥayyān al-ʾAndalūsī), *Hidāyat an-naḥw*, 14th c.
*Qāḍī Šihāb ad-Din al-Ġaznawī ad-Dawlatābādī (d. 1455), *al-ʾIršād fī n-naḥw.*
ʾAbū Barkat Nūr ad-Dīn ʿAbd ar-Raḥmān Jāmī (1414–1492), *Šarḥ al-Kāfiya (Mullā Jāmī)*, 15th c.
Ibn al-Ḥājib, *al-Kāfiya fī n-naḥw.*
ʿAbd al-Qāhir al-Jurjānī (d. 1078), *Miʾa ʿāmil.*
ʾAbū l-Fatḥ Nāṣir ibn as-Sayyid ʾAbū l-Mukarram ibn ʿAlī al-Xuwārizmī al-Muṭarrizī (1143–1209 or 1213), *al-Miṣbāḥ fī n-naḥw.*
Šarīf al-Jurjānī, *Naḥw-i Mīr*, 14th c.
Šarīf al-Jurjāni (or others?), *Šarḥ Miʾa ʿāmil*, 14th c.

Rhetoric/Poetics

Sirāj ad-Dīn ʾAbū Yaʿqūb Yūsuf ibn ʿAlī ibn Muḥammad as-Sakkākī (d. 1228), *Miftāḥ al-ʿulūm.
Saʾd ad-Dīn Masʿūd ibn ʿUmar as-Saʿd at-Taftazānī (1322–1389), *Muxtaṣar al-maʿānī*, 1355.
at-Taftazānī, *al-Muṭawwal*, 1347.
ʾAbū ʿAbdallāh Muḥammad ibn ʿAbd ar-Raḥmān Jalāl ad-Dīn Muḥammad Qazwīnī (1268–1338), *Talxīṣ al-Miftāḥ.*

Literature

*ʾAḥmad ibn Muḥammad aš-Širwānī al-Yamanī (d. 1840), *ʿAjab al-ʿujāb fīmā yufīd al-kuttāb.*
ʾAbū Tammām Ḥabīb (786–845), *Dīwān al-ḥamāsa.*
ʾAbū Muḥammad al-Qāsim al-Ḥarīrī (1054–1121 or 1122), *Maqāmāt al-Ḥarīrī.*
ʾAbū ṭ-Ṭayyib ʾAḥmad ibn Ḥusayn al-Mutanabbī (915–965), *Dīwān al-Mutanabbī.*
*ʾAḥmad al-Yamanī, *Nafḥat al-Yaman*, ca. 1820.
Muḥammad ʿAzīz ʿAlī (1883–1955), *Nafḥat al-ʿArab.*
ʿAlī ibn ʾAbī Ṭālib, *Nahj al-balāġa*, 10th c. (Mostly taught in Shīʿī *madrasa*s.)
Kaʿb ibn Zuhayr (d. 645), *Qaṣīdat al-Burda.*
Sabʿa Muʿallaqāt, by several pre-Islamic Arab poets, transmitted by ʾAbū l-Qāsim Ḥammād ibn Sābūr ibn al-Mubārak ar-Rāwiya (d. 772).

Code: In bold – books used in the Dars-i Nizami; with asterisk – books no longer used.

by Mullā Nizām Uddīn of Sihālī (d. 1748), a village near Lucknow (Robinson 2002:48–50). The texts used in the Dars-i Nizāmī are still used in the *madrasa*s of South Asia (Sufi 1941:73–75).

In present-day Pakistan there are about ten thousand *madrasa*s registered, with private boards of their own sects and subsects (ICG 2002:2). They teach in Urdu and sometimes in Pashto or Sindhi, but the qualifying examinations are held in Urdu or Arabic. Most students answer the questions in Urdu, although they use memorized passages from Arabic texts in their answers. In short, their understanding of Islam is predominantly in Urdu, despite their apparent ease with Arabic quotations.

The canonical texts of the Dars-ī Nizāmī still have explanations in Arabic and sometimes in Persian. However, most texts are available in Urdu translation, and the important ones have commentaries and explanations in Urdu. The grammatical texts are on grammar, morphology (*ṣarf*) and syntax (*naḥw*), literature, and rhetoric. The aim of grammatical studies is to preserve the language from change, which is seen as corruption (as by Ibn Xaldūn, *Muqaddima* 322; Shalaby 1954:44–47). The role of literature and rhetoric too is similar. The best writings of the classical period are meant to be exemplars. Since they are either in Classical Arabic or in a literary style considered elegant, they are still taught in the *madrasa*s. The objective is to conserve what is seen to be the identity-confirming icons from the past.

As the *madrasa*s saw themselves as the custodians of an increasingly beleaguered and besieged Islamic identity during colonial rule and after independence in 1947 when Pakistan came to be ruled by a secular elite, they remain defensive of the Dars-i Nizāmī and refuse to change the traditional texts.

In Pakistani *madrasa*s, however, Arabic grammar is actually taught through contemporary books which follow modern methods of language teaching (Rahman 2002:106–107). In Table 6, only the most commonly used textbooks are listed; otherwise, most Arabic texts have translations and explanations in Urdu.

In short, while the *ulema* preserve the Dars-ī Nizāmī for identity-related reasons, they are pragmatic enough to have introduced at least some modern means of teaching Arabic in their *madrasa*s. Despite this, students of *madrasa*s are generally unable to understand or speak modern Arabic. They are also unaware of modern Arabic literature or scholarship on Arabic outside their own texts.

3. Arabic and the Islamic identity in Pakistan

Among the Muslims of South Asia, Arabic has been seen more as a symbol of Islamic identity than as a language. Ordinary people begin their children's education in the traditional way by teaching them the rudiments of the Arabic script. Sometimes, however, the child merely recognizes Arabic letters in the *Qurʾān* without knowing Arabic or developing the ability to read languages in similar scripts. However, while in 1951 about 10.5 percent of Muslims could read the *Qurʾān*, in 1998 this

Table 6. Newer textbooks of Arabic of all the textbook boards of Pakistan, from classes 6 through 12

Charthāwalī, Muštāq Ahmad. N.d. *ʿIlm aṣ-ṣarf*. 3 vols. Rawalpindi: Raza Publications.
——. N.d. *ʿIlm an-naḥw*. Hazro: Maktabā-i Siddīqia.
——. N.d. *ʿArabī zubān kā āsān qāʿidā* [primer]. Hazro: Maktabā-i Siddīqiā.
——. 1962. *Rōzatul adab*. Multan: Maktabā Imdādia.
Misrī, Muhammad Amīn ul. N.d. *aṭ-Ṭarīqa al-jadīda fī taʿlīm al-ʿarabiyya*. 2 vols. Medina: al-Jāmiʿa al-ʾIslāmiyya bi-l-Madīna al-Munawwara.
Nadwī, ʿAbdul Majid. 1951–1952. *Muʿallim al-ʾinšāʾ*, I. Karachi: Majlis Našarāt-i Islām.
——. 1954–1955. *Muʿallim al-ʾinšāʾ*, II. Karachi: Majlis Našarāt-i Islām.
——. 1955. *Muʿallim al-ʾinšāʾ*, III. Karachi: Majlis Našarāt-i Islām.
Nadwī, Abul Hasan ʿAli. N.d. *al-Qirāʾāt ar-rāšida*. 3 vols.
——. N.d. *Muxtārāt min ʾadab al-ʿArab*. 2 vols.
——. N.d. *Qiṣaṣ an-nabiʾīn*.
Razzāk, ʿAbdul. 1980. *aṭ-Ṭarīqa al-ʿaṣriyya fī taʿlīm al-luġa al-ʿarabiyya*. Rawalpindi: Urdū Bazār. (Copy used in Pakistani *madrasa*s is dated 1980.)

figure had increased to 55.35 percent, reflecting both an increase in literacy and the increasing Islamization of Pakistani society (see Table 3).

The Pakistani state emphasized the Islamic identity to counter language-based ethnicity (Rahman 1996) and to differentiate the Pakistani identity from the Indian Other. Religious symbolism was employed, and Arabic was one major such symbol. Thus, Arabic roots were used to create new technical terms in Urdu (Durrānī 1993:446–504, Chap. 15; Rahman 1999:265–267). There was a proposal that all languages of Pakistan, including → Bengali, the language of 55.5 million Pakistanis from 1947 to 1971 when East Pakistan separated to become Bangladesh, were to be written in the Arabic script to create national cohesion (ABE 1949:9). The Bengalis opposed this, and the proposal was never implemented (Legislative Assembly Debates, Pakistan-P 02, March 1951:471–472). Similarly, the proposal for introducing Arabic as the national language of Pakistan was not implemented as it was considered impractical (Rahman 2002:92). The Council of Islamic Ideology, created in 1962 in order to Islamize Pakistani society and the state, recommended in 1971 that the teaching of Arabic should be encouraged at all levels and that it should be an "alternative compulsory language" (1982 report of the Council of Islamic Ideology, p. 13). Later, the 1973 constitution, though prepared by the Left-leaning prime minister Z.A. Bhutto (1928–1979), provided that the teaching of the *Qur'ān* should be compulsory and that the teaching of Arabic should be encouraged and facilitated (Article 31 [2] a). The proposal that Arabic should be taught in government schools was implemented, but the language was taught as part of Islamic studies.

General Ziā ul Haq (r. 1977–1988), who legitimized his usurpation of power by appeal to Islam, equated Arabic with the ideology of Pakistan and Islam (Edn. Pol. 1979: 48). The teaching of Arabic was increased, and it was taught through Islamic texts, making it an extension of Islamic studies. In 1982, Arabic was made compulsory for children in state schools (not for children in English-medium private schools for the elite) in classes 6 through 8 (Malik 1996:271). The Council of Islamic Ideology insisted on granting Arabic compulsory status from the secondary stage onward and requiring that it be taught to judges. As Islamization had strengthened the Islamists in the country, all subsequent governments continued with these policies.

4. Arabic in secular institutions in Pakistan

The secular educational institutions of Pakistan – schools, colleges, professional colleges, training institutions, and universities – prepare students for careers in the state or the private sector. Arabic is only of limited use in these sectors, and parents do not want to overburden their children with learning it. At the same time, parents as well as the state want the Islamic identity of the children to be preserved, and they therefore encourage the teaching of the rudiments of Islam, including the recognition of the words of the *Qur'ān*. Thus, the state's policy of making Arabic compulsory in government schools is not opposed despite the burden on the children. Private English-medium schools, catering to the elite, do not follow government curricula and have never made Arabic compulsory. In the colleges and universities, as well as on the competitive examinations for state services, students see Arabic as an easy option. It is taught through the translation method so that competence in the subject, even among university graduates, is very rare.

Modern methods and means of teaching Arabic, including the television, were used by the Allama Iqbal Open University (a distance-teaching institution) and the National Institute (now University) of Modern Languages. Help was received from Arab countries to expose Pakistani children to Arabic (Miṣrī 1984), and many secular institutions, like the Pakistan National Centres, started teaching everyday rather than traditional Arabic to students. Even Z.A. Bhutto's government disseminated the knowledge of Arabic during the 1970s on the grounds that many Pakistanis were emigrating to the Middle East for employment. As demand for teachers rose, a number of *madrasa* graduates joined state schools as teachers (Malik 1996:271–272).

The Islamic International University, established in Islamabad, uses Arabic and English as the media of instruction for all subjects. Because of the presence of students and faculty with Arabic as their mother tongue, the students of this institution get exposure to living Arabic. However, except as readers of Arabic news, interpreters, teachers of Arabic, and liaison persons with the Arab world, Pakistanis do not find Arabic of much utilitarian value. Thus, the number of those who learn it and gain real competence in it remains small.

To conclude, Arabic remains an iconic language for Pakistani Muslims as well as for the ruling elite of Pakistan, who legitimize themselves in the name of Islam whatever their actual policies and practices. For the religious forces, too, it is part of identity, an identity which is in confrontation with the secularizing trends of Pakistani mainstream education and perceived Western hegemony.

Bibliographical references

Primary sources

Census. 1951. *Census of Pakistan, 1951*. I. *Reports and tables*, E.H. Slade. Karachi: Manager of Publications, Government of Pakistan.

Census. 1961. *Census of Pakistan Population, 1961*. I. *Pakistan*, A. Rashid. Karachi: Ministry of Home and Kashmir Affairs, Government of Pakistan.

Census. 1998. *Census report of Pakistan*. Islamabad: Population Census Organization, Statistics Division, Government of Pakistan, 2001.

Ibn Ḥawqal, *Masālik* = Ibn Ḥawqal (d. 1193), *ʾAškāl al-bilād* or *Kitāb al-masālik wa-l-mamālik*. Ed. H.M. Elliott and John Dowson, *The history of India as told by its own historians: The Muhammadan period*, I, 37–40. London: Trubner, 1867. (Repr., Lahore: Islamic Book Service, 1976.)

Ibn Xaldūn, *Muqaddima* = ʾAbū Zayd ʿAbd ar-Raḥmān ibn Muḥammad Ibn Xaldūn, *al-Muqaddima*. Repr., Beirut: Dār al-Qalam, 1413/1992.

Sarwāni, ʿAbbās Xān. Ca. 1586. *The Tārīkh-i Shēr Shāhī*, II. Ed. and transl. from Persian S.M. ʾImām ad-Dīn. Dhaka: University of Dacca, 1964.

Secondary sources

Advisory Board of Education (ABE). 1949. *Proceedings of the Second Meeting of the Advisory Board of Education, 7–9 February 1949*. Karachi: Manager of Publications, Ministry of Education.

Ahmad, Zubaid. 1946. *The contribution of India to Arabic literature from ancient times to 1857*. Repr., Lahore: Sh. Muhammad Ashraf, 1967.

Bernier, Francis. 1826. *Travels in the Mogul Empire*. Trans. from the French by Irving Brock. 2 vols. London: William Pickering.

Durrānī, Atash. 1993. *Urdū istilāhāt nigārī* [Neologism in Urdu]. Islamabad: Muqtadrā Qaumī Zubān.

Edn. Pol. 1979. *National education policy and implementation programme*. Islamabad: Ministry of Education.

Elliot, H.M. and John Dowson. 1867. *The history of India as told by its own historians: The Muhammadan period*. 8 vols. London: Trubner. (Repr., Lahore: Islamic Book Service, 1976.)

Grimes, Barbara F. (ed). 2000. "Pakistan". *Ethnologue: Languages of the world*, 588–598. 14th ed. Dallas, Tex.: Summer Institute of Linguistics.

Hai, Abdul. 1947. *Nuzhat ul xawātir* [Biographical sketches of eminent Indian religious and other personalities]. 7 vols. Transl. from Arabic into Urdu by Abū Yahyā Imām Xān. Lahore: Maqbūl Academy, 1965.)

ICG. 2002. *Pakistan: Madrassas, extremism and the military*. Islamabad and Brussels: International Advisory Group Asia. (= Report no. 36, 29 July 2002.)

Khan, Gulfishan. 1998. *Indian Muslim perceptions of the West during the eighteenth century*. Karachi: Oxford University Press.

Kōken, Yusuf M. 1974. *Arabic and Persian in Carnatic, 1710–1960*. Madras: Armera.

Makdisi, George. 1981. *The rise of colleges: Institutions of learning in Islam and the West*. Edinburgh: Edinburgh University Press.

Malik, Jamal. 1996. *Colonialization of Islam: Dissolution of traditional institutions in Pakistan.*. Lahore: Vanguard Books (Pvt) Ltd.

Miṣrī, Muḥammad. 1984. *al-Luġa al-ʿarabiyya fī Bākistān*. Islamabad: Ministry of Education.

Mubarakpurī, Qazi Athar. 1989. *Hindustān mein Arbōn kī hukūmaten* [The states of the Arabs in India]. Lahore: Progressive Books.

Munawar, Muhammad. 1972. "Ghaznavī ahed: 998–1187" [The Ghaznavid era in the history of literature]. *Tārīx Adabiāti-e-Musalmānān-e-Pākistān-o-Hind*, II, 61–126. Lahore: Punjab University.

Nadwī, Syed Sulaiman. 1972. *ʿArab o Hind ke taʿaluqqāt* [The relations between Arabia and India]. Karachi: Kareem Sons Publishers.

Rahman, Tariq. 1996. *Language and politics in Pakistan*. Karachi: Oxford University Press.

——. 1999. *Language, education and culture*. Karachi: Oxford University Press.

——. 2002. *Language, ideology and power: Language-learning among the Muslims of Pakistan and North India*. Karachi: Oxford University Press.

Robinson, Francis. 2002. *The ulama of Farangi Mahall and Islamic culture in South Asia*. Lahore: Ferozsons (Pvt) Ltd.

Shalaby, Ahmad Q. 1954. *History of Muslim education*. Beirut: Dār al-Kaššāf.

Šīrānī, Ḥāfiẓ Maḥmūd. 1929. "Ā̱twīn aur navīn sadi hijrī kī Fārsī tālifāt sē Urdū zubān kē vujūd kā subūt" [Proof of the existence of Urdu from documents of Persian from the 8th and 9th centuries]. *Oriental College Magazine* (November). (Repr., *Muqālāt-i Ḥāfiẓ Maḥmūd Šīrānī* [The dissertations of Šīrānī] I, ed. Mazhar Maḥmūd Šīrānī. Lahore: Majlis Taraqqī-i Adab, 1966.)

Shalaby, Ahmad Q. 1954. *History of Muslim education*. Beirut: Dār al-Kaššāf.

Sufi, G.M.D. 1941. *Al-Minhaj, being the evolution of curriculum in the Muslim educational institutions of India*. Delhi: Idārāh-i Adabiyāt-i Dillī. (Repr., 1977.)

Tariq Rahman (Quaid-i-Azam University)

Palaeography

Palaeography is the science of deciphering and determining the date of ancient documents or systems of writing. Arabic palaeography is the study of the development of the Arabic script through time and place.

1. INTRODUCTION

The very first verses that were revealed to the Prophet Muḥammad, according to the tradition, symbolize the importance attached to writing in the society in which Islam arose: "Recite! Your Lord is the Most Bountiful One, who by this pen taught man what he did not know" (Q. 96/1–5). While only a minority in the world of medieval and premodern Islam (7th–19th centuries) would have been able to read and write, the written word was everywhere – in the form of administrative records, private and public legal documents, literary texts, monumental inscriptions, and graffiti. Arabic writing found its way onto almost every kind of surface, including textiles, bone, glass, marble, metal, wood, ostrich eggs, and ostraca, as well as those materials produced specifically for the purpose, such as leather, papyrus, parchment, and paper.

The central role of the *Qurʾān* and the related consecration of the Arabic language and script, combined with Islam's doctrinal aversion to depictions of animated objects, promoted the role of writing as the main vehicle of decoration, and with it the high valuation placed on penmanship and the art of writing in the Islamic world. Greatest veneration was reserved for the Arabic of the *Qurʾān*, and special rules governed how it was to be transcribed. But even profane texts, via their linguistic and formal associations with religious writings, attracted a certain degree of reverence, reinforced by the religious formulae frequently attested within them.

When studying Arabic palaeography, scholars – both Arabic and Western – have traditionally confined their attentions to formal, literary manuscripts and other documents of a self-consciously calligraphic nature. The drafts of literary texts, legal and administrative documents, and private letters, meanwhile, have tended to be excluded, even though they were composed, read, and used by, or contemporary with, the same individuals who wrote the literary manuscripts on which the discipline of palaeography is based.

Not surprisingly, parallels do exist between Arabic documents, the descriptions of scripts given in premodern sources, and the writing samples seen in manuscripts. Moreover, while maintaining a distinction between informal or nonprofessional texts (letters, bills, dockets, graffiti, and such like) and elaborate or calligraphic texts (albums, monumental Qurʾāns, and other manuscripts; Qalqašandī, *Ṣubḥ* III, 26; Déroche 2000:232), modern concepts of palaeographical classification are beginning to allow for, and even demand, the inclusion of documentary sources within the field's purview. This entry, therefore, takes into account the palaeography of decorative and monumental writings as well as mundane documents intended for everyday use. The focus is on writings executed with a pen, with occasional reference to those hewn in stone and other materials, because of the continuous reciprocal influence between these media and the links between the designers of documents and those of inscriptions (cf. Grohmann 1952:69, 1967–1971:1–2; Sourdel-Thomine 1978:1115).

2. TOWARD A HISTORY OF ARABIC PALAEOGRAPHY

That a self-conscious palaeographical awareness already existed in the medieval Islamic world is shown by the interest and deliberate experimentation in different writing styles, exemplified by manuscripts in which a text written in an older hand might, for example, be transcribed into a newer one, or a *maġribī* text 'translated' into an oriental one. Forgeries of old scripts made in antiquity are another indication of this awareness, as are documents written in an 'old' style. It is also clear from the presence of different names for scripts, as seen, to take but one example, in the medieval catalog of the library of Qayrawān (Déroche 2000:206–207), in which the script of each manuscript is carefully classified. The full extent of this awareness, however, is not easy to determine, and what the Qayrawān catalog's classifications actually signified at the time and in the place it was composed remains obscure. A similar problem exists with descriptions of Arabic scripts by medieval authors, who took no account of regional variations in script designation and were largely unaware of possible terminological changes in time. This issue of definition is further complicated by the replication of mistakes by copyists and deliberate changes in the script samples accompanying script descriptions.

Although there was an explosion in the 10th century of books describing the many different scripts in use, it is often difficult to match their descriptions with contemporary documents (Abbott 1941:86–89). Written for contemporaries, these studies were not intended for use by later generations, and the typologies they use are typically vague and imprecise. The script descriptions that Ibn an-Nadīm (d. after 377/987–988) gives in his *Kitāb al-fihrist* are entirely characteristic in this respect, and were much copied in later administrative handbooks. Rather than offering precise definitions tied to actual writing samples that could be used by working calligraphers, their focus is on collecting interesting snippets of disconnected information. They reflect the erudite but traditional outlook of the time, whose basic unit of understanding was the workshop, and which preferred to ascribe developments in script patterns to the innovations of individual master calligraphers rather than as the result of evolutionary adaptations to changing conditions or usages. The most important and extensive of these descriptions can be found in al-Qalqašandī's (d. 821/1418) *Ṣubḥ al-ʾaʿšā*. Attempts to assign terms and classifications from these medieval Arabic sources to book and document hands have therefore been largely ineffective, or have resulted in the unhelpful division of manuscripts under very generic labels such as 'Kufic' for essentially angular scripts or '*nasxī*' for essentially cursive ones. Another complicating consequence of the prestige of the Arabic written language is the emphasis placed on calligraphy as an art form (→ script and art), where subjective aesthetic value judgments dominate and individual artistry becomes the focus of analysis. While the lore of penmanship gave individual calligraphers enormous prestige, and compilations of biographies of calligraphers became very popular, making them among the best known professionals, the profession itself remains obscure.

Nor, in the absence of a robust Arabic typology, can typologies developed in Western palaeography, which are based mostly on studies of Latin manuscripts, be called upon to fill the gap. Cursive writing, for example, is defined in Western palaeography as that in which the letters of a word are connected by ligatures and which is used for documents and other mundane writings, in contradistinction to the noncursive, 'calligraphic' script used in manuscripts and literary texts. But this definition does not apply to Arabic script, which has a well-established, clearly defined system of ligatures. Cursive script in the Arabic tradition is better defined as that in which the scribe limits the movements of the pen and lifts it from the page as little as possible, using ligatures also in places where they do not otherwise appear, and which has a tendency to turn angles into curves and curves into straight lines (see below, Sec. 6). Arabic formal scripts have a horizontal baseline defined by the ligatures between the letters, and the writing is regular with a clear *ductus* (i.e. letter formation). Only a typology that is set within specifically Arabic parameters, with samples of securely dated and geographically defined texts grouped taxonomically, will suffice.

While recent Arabic palaeographical studies have tended to use this approach, taking more samples into account and basing their script descriptions on the hands found in the manuscripts themselves (e.g. Déroche 1992), several problems still remain. First and foremost is the lack of sound descriptions of book hands and document scripts (cf. Grohmann 1967–1971:32–65; Déroche 1998), the sole exception being the script of Qurʾānic manuscripts, which has been studied in detail (Déroche 2003, 1992; George 2007). While there are currently no handbooks of Arabic palaeography, a handful of palaeographical albums do exist, but these consist mainly of collections of photographs and facsimiles without proper analysis or discussion of the scripts depicted. Detailed and exact descriptions of securely datable manuscript and document collections according to standardized rules are only now slowly appearing.

The sheer complexity of the material also poses significant challenges. Writing styles did not, of course, change from one day to the next, and in practice it is very difficult to observe the transformation from one style to another and to determine clear chronological developments. Moreover, new scripts did not replace older ones, and different scripts could exist in mixed forms or separately at the same time and even in the same place. A long-lived scribe might employ a variety of scripts during his lifetime, or, if well traveled, take regional practices with him from place to place, such as a *maġribī* manuscript which was composed in Mecca in 706/1307 (Géhin 2005:98). All of this tends to frustrate attempts to draw up neatly defined script categories.

Turning to the scripts themselves, any meaningful system of classification must take into account, first, the shape and size of individual letters, with ligatures and the spacing between characters and words considered as an integral part of the script; and second, the page layout, with the number and size of lines, and the use of colored ink and other embellishments.

The practice of diacritical dots, which differed through time and place in the Islamic world, should be noted when distinguishing scripts (Géhin 2005:102). Dots, which in Arabic distinguish letters that have the same shape (*rasm*) are used, albeit occasionally, in the earliest known Arabic writings, including the first two dated Arabic papyri from 22/643 (see Fig. 1) and in the earliest inscription from Arabia dated 24/644–645 (Ghabban 2003). Dashes are used for the same purpose in some early Qurʾānic manuscripts, the Dome of the Rock Qurʾānic inscription (→ epigraphy), the Egyptian graffito dated 117/735 (Moritz 1905, plates 107–108), and some papyri (e.g. *P.Cair. Arab.* III 147.2; cf. Grohmann 1952:83 n. 292). Early papyri (7th–8th centuries) show the practice maintained in North African texts of writing the *fāʾ* with one dot under the letter (as opposed to the eastern practice of one dot over the letter) and the *qāf* with one dot over the letter as opposed to two. Other letters, too, are attested with a dot under instead of over the letter (Grohmann 1952:85–86).

Vowels are seldom encountered except in Qurʾānic manuscripts, and their presence generally indicates a manuscript on which particular care and attention has been lavished. In the earliest Qurʾāns, vowels are not used, while in the Abbasid period differently colored dots were applied, a practice continued in the Maghribi tradition. By the end of the 9th century, the system that is currently in use, with small lines drawn over and under the letters in the same color ink as the main text, had been introduced.

3. THE SOCIAL AND DEVELOPMENTAL CONTEXT

The shape and form of letters used to write a text affect not just its appearance but also its 'atmosphere', conveying important messages about the social and professional intentions and background of the scribe and reader beyond the content of the text itself.

Conversely, the function of the text and how readable it was required to be – whether it was a draft or a final text, a private letter or a petition, a monumental inscription or a graffito, a decorative band of (pseudo-) writing on a textile, an amulet or esoteric text – as well as the training and proficiency of the scribe and the time available to complete it all influenced the choice of script and the quality of execution. The relation between the content and form of a text is also important. For example, some texts, such as Qurʾāns, early Christian Arabic manuscripts, and chancellery documents, are more homogeneous in their style, while others, such as drafts of literary works or private letters, show more variation. While this variation should be distinguished from the historical development of the script, one wonders what influence changes in execution had on the emergence of new script types, and how scripts evolving from earlier practice were related to those introduced from different geographical areas or cultures, or developed by individual calligraphers.

The development of the Arabic script is also closely related to the question of literacy, the level and accessibility of schooling, the cost and availability of writing materials, and the use of public and private scribes. Different levels of expertise are clearly visible in witness statements at the end of legal documents. Throughout the premodern period, reading and writing were limited to a minority in society, but the production of texts fluctuated. The 9th century saw a sharp increase in Arabic writings, both in documentary and literary production, coinciding with the coming into use of a small, cursive script. The role of Arabization and Islamization in the development of the Arabic script, both processes becoming increasingly important at this time, also had a bearing, and, similarly, that role has not been adequately explored.

The sparse use of diacritical dots in early Qurʾāns has led to the suggestion that these texts functioned more as an aide-mémoire. Classical authors argued that diacritics were an insult to the reader and should only be used to prevent mistakes, and many documents and manuscripts in the medieval and early modern period continued to be written without full diacritics (Grohmann 1952:82–87). Similarly, technical texts, such as fiscal accounts, legal contracts, or petitions, used a specialized language and writing style, including abbreviations and ideograms, which required expert knowledge. On the basis of these observations, scholars have argued that the expertise required to write Arabic was limited to a highly educated, select minority whose writings were understood only by the initiated, well-trained few. Yet, these were specialized fields of activity; a better sense of everyday usage can be gained from the more mundane writings. Although used sparingly, diacritical dots in administrative letters seem to follow a definite pattern, suggesting that they were seen as an integral part of the text. Documents written by unskilled scribes and private documents, moreover, show a similar sporadic use of diacritics, implying that this was the normative way to write.

The 7th-century conquests greatly enlarged Arabic's field of operation, and, although non-Arabic languages continued to be used in the administration up to the end of the 8th century, local languages and scripts were to a certain extent sidelined. As Arab rulers at the end of the 7th and beginning of the 8th century became more involved in the day-to-day administration of their empire, and as the *dīwān* became increasingly Arabicized (decreed by ʿAbd al-Malik in 74/693), the demand for Arabic writers and the circulation of Arabic documents steadily rose. The link between the development of the Arabic script and the chancellery was also reinforced by the fact that many of those associated in the Arabic literary tradition with the development of Arabic scripts, or who described them, worked in the chancellery or had high positions in the court administration (Grohmann 1967–1971:13–32; Robin 2006:349–350). For the influence of chancellery writing on the development of Arabic script, see also below, Sections 6 and 7.

The use and spread of the *Qurʾān*, the most copied text in the Islamic world, and developments in its scripts were intimately connected to palaeographical trends in the wider Islamic world. As with other writings, certain scripts were exclusively reserved for the *Qurʾān*, while others were expressly prohibited, and copying the *Qurʾān* was a specialized profession in its own right. Nevertheless, interaction and cross-fertilization between Qurʾānic scripts, monumental inscriptions, government decrees, and chancellery scripts and more informal, 'lower' genres seems evident, not because the same scribes were often writing them but rather because of the emblematic and exemplary function that these writings served.

Little work has been done on education and schooling of scribes in the early medieval period. In the earliest bilingual Greek/Arabic papyri, the Greek and Arabic parts of the text were written by different scribes, with subtle differences in content. By the early 8th century, Arab and Greek scribes are still mentioned separately in the budget of the Muslim administration in Egypt. Scribal technique also differed, with Greek continuing to be written with a generally thinner cut reed than the Arabic, a practice which can be clearly observed in the (bilingual) papyri. Although the Arabic word for pen, *qalam*, was derived from the Greek *kálamos* (κάλαμος), we know of no examples of Arabic written with a 'Greek' reed pen and vice versa – as occurred in the Ptolemaic period in Egypt when Greek was occasionally written with an Egyptian rush – suggesting that there continued to be a separate educational tradition in the two languages in Egypt. From the end of the 7th century, however, bilingual Arabic/Greek papyri are literal translations of one text, implying a closer knowledge of both languages among the employees of the chancellery.

4. THE FIRST FIFTY YEARS: *ḤIJĀZĪ*

No complete or substantial manuscripts in Arabic, including Qurʾāns, Christian Arabic writings, and other texts, survive from the first two centuries of Islam. The oldest books preserved on paper, parchment, and papyrus date to the first half of the 9th century (Rāghib 1996:1–5). By contrast, there are thousands of papyrus and parchment documents and fragments of literary texts, including Qurʾānic ones, as well as inscriptions – including five pre-

Islamic Arabic inscriptions (→ Old Arabic) – datable to the 7th and 8th centuries with which a typology of the script of the earliest period can be tentatively established.

Within a century of the Prophet Muḥammad's death, the Arabic script had spread from Sogdiana to Upper Egypt, showing a well-developed and regular form from its first occurrence, although with a great variety of styles. It is not clear whether these variations were due to Arabic script still being in its developmental stage, or whether this was the result of the widespread increase in its use, unrestricted by any rules or canons of style. Variations also occurred as a result of the writing materials used, without there being, for example, recognizably different lapidary, documentary scripts, or book hands. There are, however, enough common characteristics for these scripts to be given the universal name of *ḥijāzī*, after the area in western Arabia where the cities Mecca and Medina are located. This script was used in Qurʾāns of the 7th and 8th centuries, on papyrus and parchment letters and documents, and in inscriptions (Abbott 1941:70; Gruendler 1993:131–132; Déroche 1992:27). Strikingly, it was used throughout the Islamic Empire, such as in a petition written on leather from Sogdiana (Tajikistan), dated 99–100/718–719; the earliest dated papyri (22/643) from Upper Egypt (Grohmann 1952:113–115; see Fig. 1); tax demand-notes, such as *P.Ness.* III 60 dated 54/674, and letters, all from Palestine (Grohmann 1963); pre-Islamic inscriptions from Arabia (Hoyland 2007); the oldest dated Arabic inscription, a graffito from Zuhayr in Arabia dated 24/644 (Ghabban 2003); the oldest Arabic epitaph from Egypt dated 31/652 (*RCEA* 6); Sasanian-Arab and Byzantine-Arab coins; and the earliest Qurʾānic fragments, found in Yemen, datable to the second half of the 7th century (cf. Robin 2006). Orthographic characteristics distinguish inscriptions and papyri from Qurʾānic texts, while the papyri have recognizable cursive tendencies of their own, due to their more rapid execution and a need to economize the efforts of the scribe.

Ḥijāzī writing is characterized by elongated ligatures between letters and long spaces between freestanding letters, sometimes extending over two lines, and not to be differentiated from the spaces between words. Individual letter shapes are cursive, and the writing is generally large with letters extending high while slanting. The only characteristic of this script identified in literary sources (by Ibn an-Nadīm) is the shape of the *ʾalif*, extending high above the other letters, slanting and bending to the right at the bottom, which is attested in documents and manuscripts (Abbott 1939, 1941:71).

5. Script standardization: The first geometric style

It is tempting to see the development of a more uniform and homogeneous Arabic script in the late 7th and early 8th centuries as the result of contemporary Umayyad administrative reforms aimed at the Arabization and Islamization of the government (Déroche 1992:34). Although it is impossible to trace the influence from the empire's center in Damascus outward, the introduction of a consistent script over the vast Umayyad Empire (660–750) required a well-functioning infrastructure, and government officials were surely involved in the development and propagation of this script. The Iraqi governor al-Ḥajjāj ibn Yūsuf (d. 95/714), for example, allegedly standardized the diacritical system. With the decision to use Arabic exclusively in the administration, as an expression of the caliphate's state-forming ambitions, the number of Arabic public writings, monumental inscriptions, and documents rose, and the distinction between 'private' and 'public' writings became more pronounced. Closely related variations of this script appear in administrative papyrus documents, such as the correspondence of the Egyptian governor Qurra ibn Šarīk, dating from 90–92/718–720 (*P.Heid.Arab.* I; *P.Cair.Arab.* III); treaties and protocols, especially those purely in Arabic (*CPR* III; *P.Cair.Arab.* I); milestones erected at the order of the caliph ʿAbd al-Malik (r. 64–86/685–705; *RCEA* 14–17), and other inscriptions; and in the earliest dated testimony of the *Qurʾān*, the mosaic inscription in the Dome of the Rock dating to 72/692 (Creswell 1969); as well as in Umayyad Qurʾānic manuscripts.

Based on this standardized *ḥijāzī* style, a form arose that has often been called → Kufic, but which François Déroche has suggested should be labeled instead "early Abbasid scripts" because of the many variations it encompasses (Déroche 1992). 'Kufic' Qurʾānic manuscripts can be grouped according to certain script characteristics (Déroche 1992), and similar classifications have been suggested for papyri (Abbott 1941:84; Gruendler 1993:132–137). Some documents and manuscripts show a mixture of the two types, but the relationship between the *ḥijāzī* and the early Kufic scripts has not been satisfactorily explained. Some common features of this script are the regular baseline, below which curves take the form of parts of perfect circles; thick lettering, with pronounced horizontal strokes, and short vertical ones; regular proportions for the text area; and identical spaces between groups of letters, creating a balanced layout.

The script of papyri dating to the first two Muslim centuries also shows quite a bit of variation, but it has clear commonalities with the writing of the *ḥijāzī* and Kūfic Qurʾānic manuscripts, as well as pre-Islamic and early Arabic inscriptions, coins and seals, and early papyrus protocols (Sijpesteijn, in press). Less thickly written than the Qurʾāns, the letters show similar angular features and proportioned style (for examples and an extensive description, see Khan 1992:27–39; cf. Déroche 1992:28). In spite of their cursive tendencies and great variations in script pattern, even within one document, papyrus letters and other documents were occasionally executed with great care, in a style comparable to contemporary Qurʾānic and other monumental manuscripts, e.g. the treaty with the Nubians dated 141/758, edited by Hinds and Sakkout (1981), and the letter dated 112/730, edited by Diem (1984:147; see Fig. 2). Hence, it is possible to mistake literary for documentary texts (Kister 1964). *Dāl/ḏāl* has a right-bending tip. The final and freestanding *ʾalif* extends below the line and often has a rightward bend at the bottom. The three dots of *šīn* are aligned horizontally. The *ṣād/ḍād* extends horizontally with straight parallel strokes. Initial and freestanding *ʿayn/ġayn* extends to the right horizontally. Medial and final *ʿayn/ġayn* is open at the top. The tail of the final *qāf* extends vertically downward before curving up. Initial and medial *kāf* has a hairpin shape. Final and freestanding *kāf* has an upward-moving and right-extending top. *Mīm* has a round head and a very short tail. Final *yāʾ* has a back-bending tail. The greeting formulas, blessings, and layout used in official and private letters of this period also show uniformity and are recognizably different from later documents.

The execution of Kufic differed due to the material used and the aims of the text, but the variants used under the Umayyads and early Abbasids were all recognizably one script family. Kufic remained popular and continued to be used for inscriptions and for headings and decorations in books.

6. Later styles

Administrative documents from the 8th and 9th centuries show how new developments in the Arabic script were common in the chancellery before they were introduced in book hands. They also indicate how in the 8th century scribal conventions spread from the eastern provinces, the heart of the Abbasid empire (750–1250), to the rest of the Islamic world. A recently discovered batch of 32 leather documents from Bactria (northeastern Afghanistan) dating to the early Abbasid period (138/755–160/777) allows us to examine these processes more closely. Most of the documents are tax receipts and other writings issued by the local administration (Khan 2007:201–202). They show a distinctively new style of writing with more cursive trends, a smaller script, and new administrative scribal practices, with different technical bureaucratic terms and formularies. These features appear several decades later in the Egyptian papyri, presumably introduced by Persian administrators appointed there in the Abbasid period (Khan 2007:207, 209–210). Equally, legal terminology attested first in documents from Central Asia appears later in the Egyptian papyri (Khan 1994).

Late–8th- and 9th-century papyri are written in a script clearly distinguishable through the more cursive round hand in which they are written, even though the change is not absolute and features from the early script continue to occur (Khan 1994:39–43; Sijpesteijn, in press see Fig. 3). The script is also noticeably smaller. Documents issued by the Egyptian chancellery and legal documents stand out by their uniformity (Abbott 1939:110). This style became popular also for letters and other private documents and by the mid- to late 9th century, book hands had taken it over (Sourdel-Thomine 1978:1121; see Fig. 4). In the beginning of the 11th century, these cursive scripts began to be used to copy the *Qurʾān*, first in Iran and Iraq and spreading from there throughout the rest of the Islamic world over the next two hundred years (James 1992:14; Abbott 1941:83); from the 12th century onward they also appear in inscriptions. 'Broken Kufic' (or *nasxī*, Eastern or Persian Kufic, or broken cursive), as this script was called, was the last script that became universally accepted throughout the whole Islamic Empire.

The success of this new script seems to be related to an increase in literary production to which the cursive, smaller script was more easily applied. Arabic private and official writings increased greatly in number in the 9th century, as witnessed by the large number of papyri dating to that period. Paper first introduced in the east had by this time started to spread through the rest of the Islamic Empire in this and the following century. The availability of this much cheaper writing material led to an increase in demand for books. Paper also became the preferred material for documents, replacing papyrus even in the main production center, Egypt, by the late 10th century. At this time, too, Arabic started to make serious headway as the dominant language also among non-Muslim communities, a process accelerated by large-scale conversion to Islam, all of which greatly enlarged the pool of Arabic users.

7. REGIONAL DIFFERENTIATIONS

By the middle of the 9th century, some three dozen scripts existed for secular use, many of which were related to each other (Abbott 1941:88–104). Of those scripts, six basic calligraphic forms (→ *nasx*, → *ṯuluṯ*, → *muḥaqqaq*, *rayḥān*, *tawqīʿ*, and → *ruqʿa*) and their variations remained in use in the central lands of the Islamic Empire. Of these, *tawqīʿ* and *ruqʿa* were mainly known as chancellery hands, and the other four were used for Qurʾānic manuscripts, while *nasx* especially had a wider usage. In Persia, → *nastaʿlīq*, which had become popular through its use in the chancellery, became the most frequently used script. It is possible to identify few documents that show characteristics of the calligraphic scripts as described in literary sources and that can be compared with writing samples from the calligraphic manuscript tradition (Abbott 1938:110, 1941:98–99; Grohmann 1952:69–82, 1963:xv–xxix).

Regional variations had always existed, but from the 10th century onward – doubtlessly related to the political situation of semi-independent separatist provinces – artistically diversified, regional scripts started to be used in books, documents, and inscriptions. By the end of the 10th century, the *maġribī* script had become dominant in the western part of the Islamic Empire, forming the basis of the *ʾandalusī* script. This script is characterized by the use of one dot over the *qāf* and one under the *fāʾ*, colored dots for vowels, and the very round form of its letters, as well as an idiosyncratic book format (van den Boogert 1989; James 1992:14). Sub-Saharan communities used another distinctive script, called *sūdānī*. Local variations also existed in more marginal areas of the Islamic world, such as China (*ṣīnī*), Indonesia, and India (*bīhārī*).

Political and religious motives always played a role in the dissemination and popularity of scripts. Influences moved from politically and economically important centers outward, while books reflected the relative importance of the areas in which they were produced. In the 14th century, Mamluk Egypt produced extraordinary manuscripts, but by the 15th century and continuing under Safavid rule (1502–1722), the center of book production lay in Persia, and its styles were imitated everywhere in the Islamic world. In the Ottoman lands and Mughal India, the two other political centers of the time, many different styles were in use, but by the 16th century *nasx* had become dominant following the practice in Persia (James 1992:10–12). In the 18th century, these empires all came under important European influence, which had an effect on cultural production, sometimes, as in the Ottoman Empire, blending foreign elements successfully with indigenous ones. Manuscripts continued to be copied by hand in the Islamic world until the early 20th century, and while the main scripts did not undergo the same dynamic development as in the first centuries of Islam, styles continued to develop and change. Cursive tendencies dominate documents of this period, with clear chronological developments and regional distinctions comparable to the developments in book hands. Scripts differed according to the function of the document and at times acquired such specialized styles that they were indecipherable, except by the practiced (see Fig. 5).

8. Conclusion

The art of the pen in the Islamic world, where the Arabic language is inseparably linked with the religion of Islam, became one of the most revered and highly regarded of professions. The special place accorded to writing in Islam resulted in a vast literature on scripts, penmanship, and famous calligraphers, which has not necessarily made it easier to follow the development of the many scripts in use during the almost fifteen hundred years of Arabic manuscript culture.

The shape and form of individual letters or a script and the connotations associated with them determine the scope and propagation of a text, its meaning, and value at a specific time and place. And as associations and ideas about their expressive or aesthetic effect changed, so too did their perceived utility and thereby their use. Arabic palaeography thus exhibits the changing political, religious, cultural, and artistic perceptions in the Islamic world.

Only by combining the information found in document and book samples with the extensive literature on Arabic palaeography can a typology of Arabic palaeography through time and place be established with which we can start to understand the historical changes in Arabic scripts.

Bibliographical references

Primary sources

CPR III = Adolf Grohmann, *Corpus papyrorum Raineri archiducis Austriae*: *Series Arabica*. I/1. *Allgemeine Einführung in die arabischen Papyri*. I/2. *Protokolle*. Vienna: F. Zöllner, 1923–1924.

Ibn an-Nadīm, *Fihrist* = Muḥammad ibn ʾIsḥāq Ibn an-Nadīm al-Warrāq al-Baġdādī, *Kitāb al-fihrist*. Ed. Yūsuf ʿAlī Ṭawīl. Beirut: Dār al-Kutub al-ʿIlmiyya, 1996.

P.Cair.Arab. III = Adolf Grohmann, *Arabic papyri in the Egyptian Library*. Cairo: Egyptian Library Press, 1955.

P.Heid.Arab. I = Carl Heinrich Becker, *Papyri Schott–Reinhardt*, I. (= *Veröffentlichungen aus der Heidelberger Papyrussammlung*, III.) Heidelberg: Carl Winter's Universitätsbuchhandlung, 1906.

P.Ness. III = Casper John Kraemer Jr., *Excavations at Nessana*. III. *Non-literary papyri*. Princeton, N.J.: Princeton University Press, 1958.

Qalqašandī, *Ṣubḥ* = Šihāb ad-Dīn ʾAbū l-ʿAbbās ʾAḥmad ibn ʿAlī al-Qalqašandī, *Ṣubḥ al-ʾaʿšā fī ṣināʿat al-ʾinšāʾ*. Cairo: Maṭbaʿa Dār al-Kutub al-Miṣriyya, 1913–1922.

RCEA = Étienne Combe, Jean Sauvaget, and Gaston Wiet (eds.), *Répertoire chronologique d'épigraphie arabe*. Cairo: Institut Français d'Archéologie Orientale, 1931.

Secondary sources

Abbott, Nabia. 1939. *The rise of the north Arabian script and its ḳurʾānic development with a full description of the Ḳurʾān manuscripts in the Oriental Institute*. Chicago: Oriental Institute.

——. 1941. "Arabic paleography". *Ars Islamica* 8.65–104.

Boogert, van den Nico. 1989. "Some notes on Maghribi script". *Manuscripts of the Middle East* 4.30–43.

Creswell, Keppel Archibald Cameron. 1969. *Early Muslim architecture*. I. *Umayyads, A.D. 622–750*. 2nd ed. Oxford: Clarendon Press.

Déroche, François. 1985. "L'écriture arabe". *Grundriss der arabischen Philologie*. I. *Sprachwissenschaft*, ed. Wolfdietrich Fischer, 325–329. Wiesbaden: L. Reichert.

——. 1992. *The Abbasid tradition: Qurʾans of the 8th to the 10th centuries AD*. London: Nour Foundation in association with Azimuth Editions and Oxford University Press.

——. 1998. "Les études de paléographie des écritures livresques arabes: Quelques observations". *Al-Qantara* 19.365–381.

——. 2000. *Manuel de codicologie des manuscrits en écriture arabe*. Paris: Bibliothèque Nationale de France. (Trans. D. Dusinberre and D. Radzinowicz. London: Al-Furqan, 2005.)

——. 2003. "Manuscripts of the *Qurʾān*". *Encyclopaedia of the Qurʾān*, ed. Jane Dammen McAuliffe, III, 254–275. Leiden: E.J. Brill.

——. 2004. *Le livre manuscrit arabe: Préludes à une histoire*. Paris: Bibliothèque Nationale de France.

——. 2006. "Written transmission". *The Blackwell companion to the Qurʾān*, ed. Andrew Rippin, 172–186. Malden, Mass.: Blackwell Publishers.

Diem, Werner. 1984. "Einige frühe amtliche Urkunden aus der Sammlung Papyrus Erzherzog Rainer (Wien)". *Le Muséon* 97.109–58.

Endress, Gerhard. 1982a. "Die arabische Schrift". *Grundriss der arabischen Philologie*. I. *Sprachwissenschaft*, ed. Wolfdietrich Fischer, 325–329. Wiesbaden: L. Reichert.

——. 1982b. "Paläographie der Handschriften". *Grundriss der arabischen Philologie*. I. *Sprachwissenschaft*, ed. Wolfdietrich Fischer, 278–284. Wiesbaden: L. Reichert.

Géhin, Paul (ed.). 2005. *Lire le manuscrit médiéval*. Paris: Armand Colin.

George, Alain. 2007. *The geometry of early Islamic calligraphy*. Ph.D. diss., Oxford University.

Ghabban, ʿAlī ibn ʾIbrāhīm. 2003. "Inscription de Zuhayr: La plus ancienne inscription arabe islamique date de l'année 24h./644–645 è chr.". *Arabia* 1.342–293. [In Arabic.]

Figure 1. Oldest dated Arabic papyrus, 22/643 (Austria National Library). Published Grohmann (1967–1971).

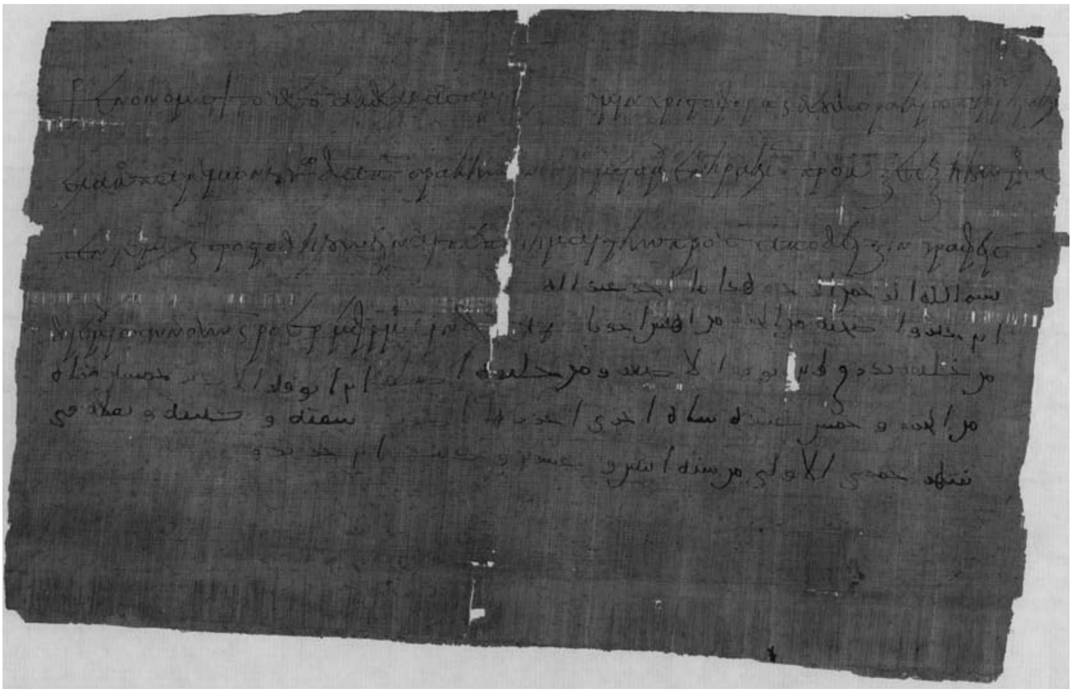


Figure 2. AP 359. Letter dated 112/730 (Austria National Library). Published Diem (1984).

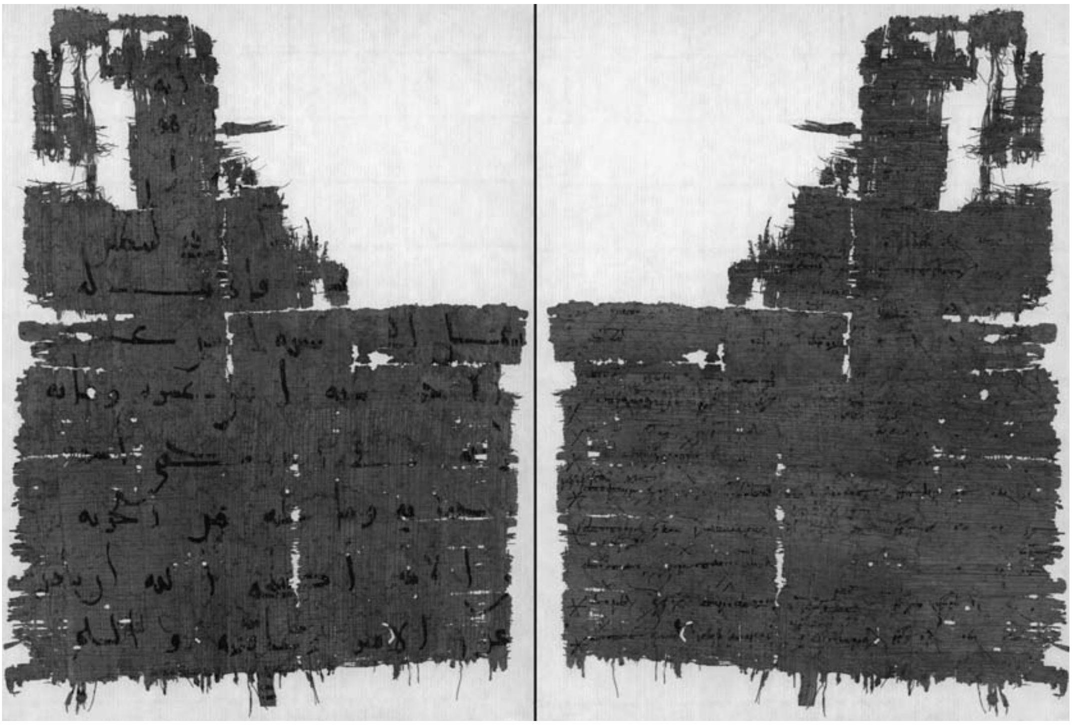


Figure 3. AP 9807. Papyrus with cursive tendencies (Austria National Library). Published *CPR* XVI, no. 6.

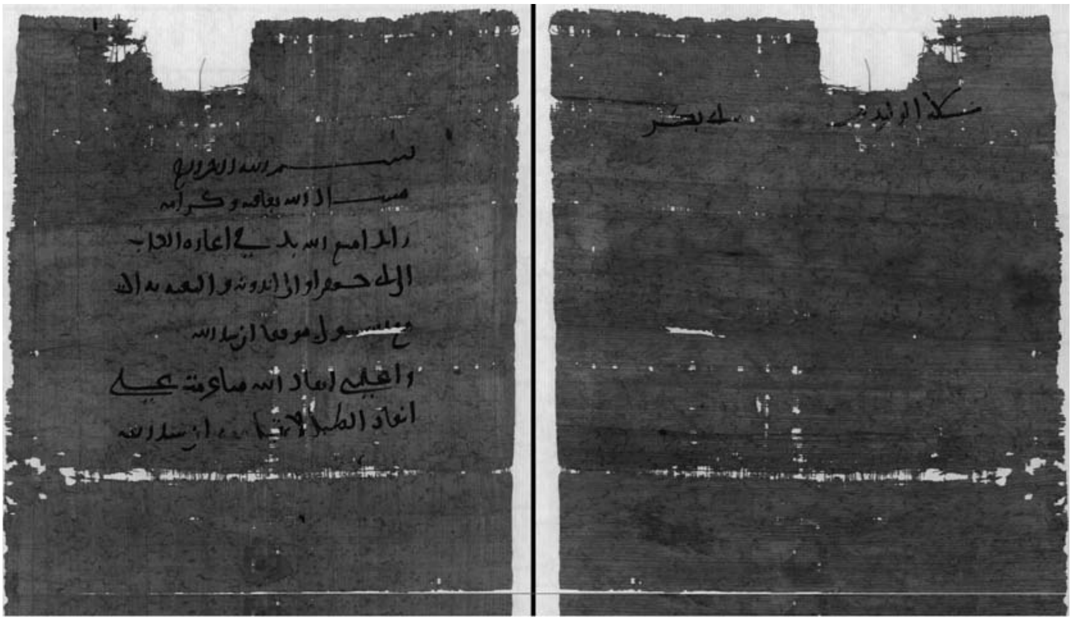


Figure 4. Book hand 9th century. Ms. Marsh 232, fol. 60v. (c) Bodleian Library, University of Oxford.

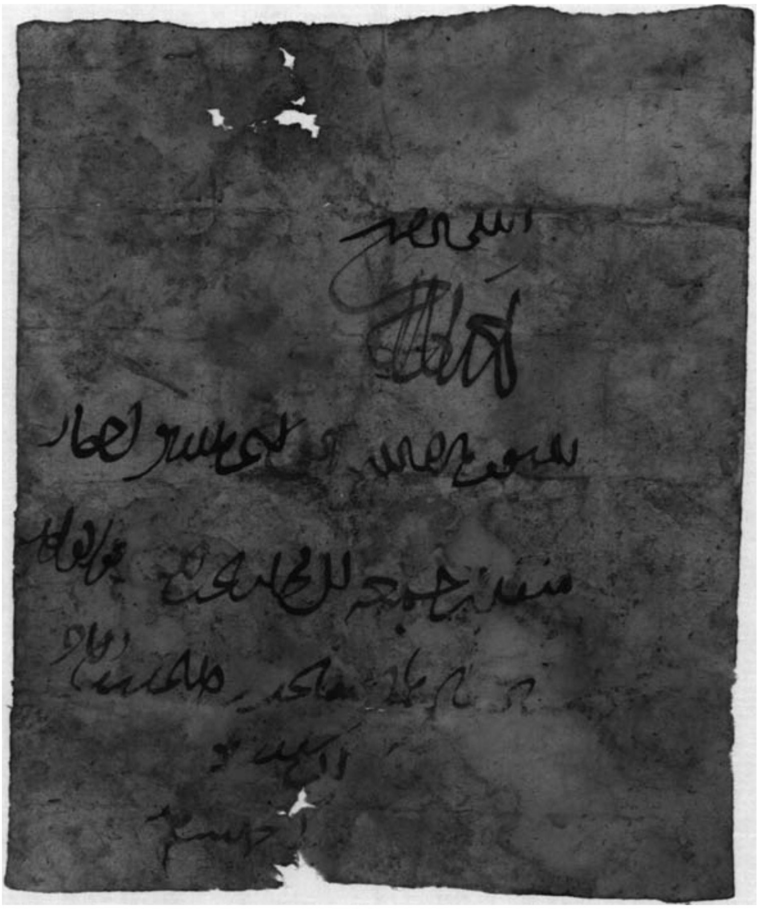
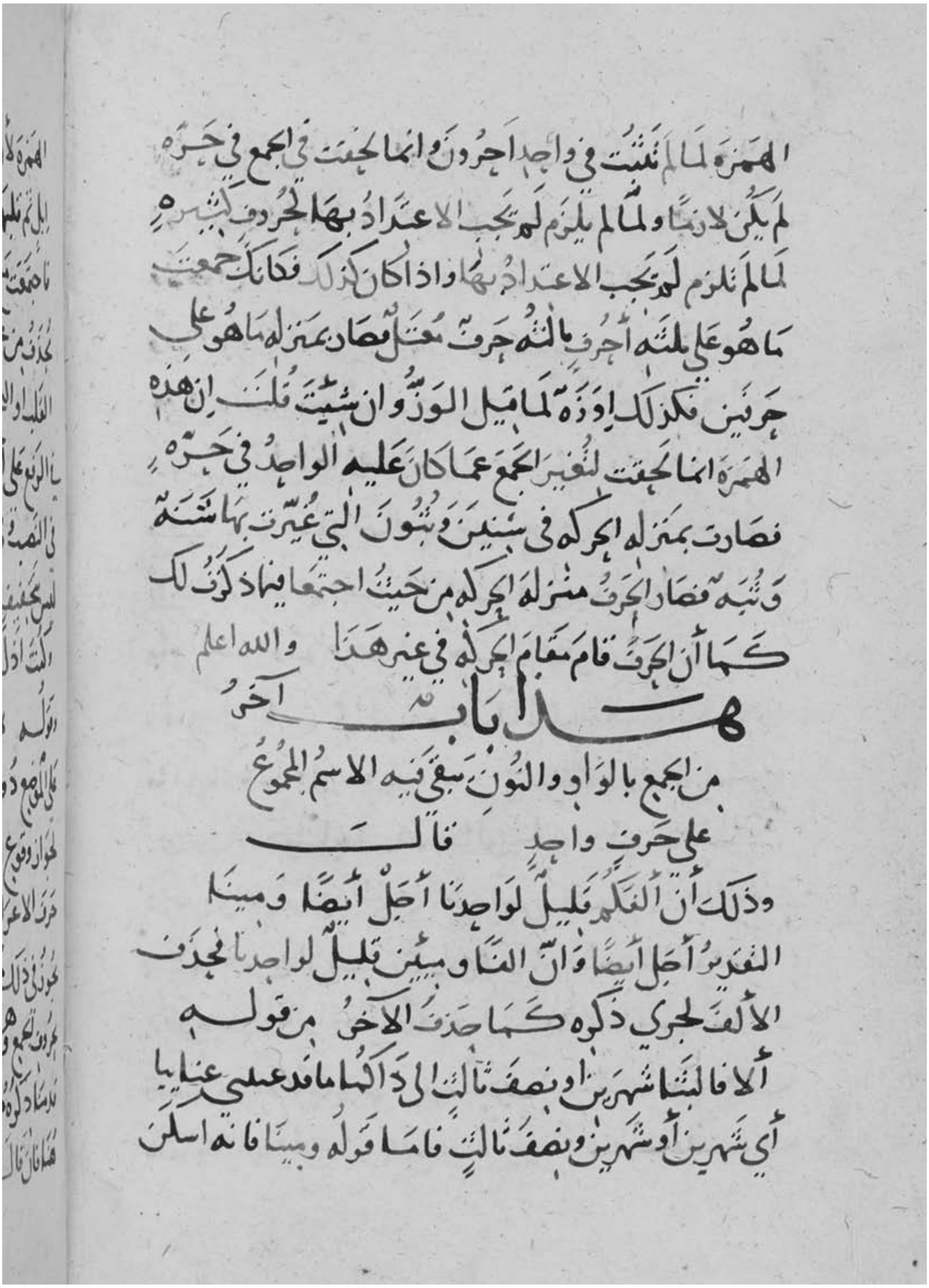


Fig. 5. Official Mamluk letter dated 698/1299–708/1309?) (ACh 12511) (Austria National Library). Published P. Vindob., no. 69.



الهمزة لما لم تثبت في واحد احرون وانما جيئت في الجمع في حرة
لم يكن لازما ولما لم يلزم لم يجب الاعتداد بها الحروف كثيرة
لما لم تلزم لم يجب الاعتداد بها واذا كان كذلك فكأنك جمعت
ما هو على ثلثة احرف بأنه حرف معتل فصار بمنزلة ما هو على
حرفين فكذلك اوزه لما قبل الوزّ وان شئت قلت ان هذه
الهمزة انما جيئت لتغيير الجمع عما كان عليه الواحد في حرة
فصارت بمنزلة الحركة في سنين وثبون التي غيرت بها سنة
وثبة فصار الحرف منزلة الحركة من حيث اجتمعا فيما ذكرت لك
كما ان الحرف قام مقام الحركة في غير هذا والله اعلم

هذا باب آخر

من الجمع بالواو والنون يبقى فيه الاسم المجموع
على حرف واحد قال

وذلك ان التكلم قليل لواحدنا اجل ايضا ومينا
التقدير اجل ايضا وان التا ومين قليل لواحدنا فحذف
الالف لجري ذكره كما حذف الاخر من قوله
ألا فالبثا شهرين او نصف ثالث الى ذاكما ما غيبتني غيابيا
اي شهرين او شهرين ونصف ثالث فاما قوله ومينا فانه اسكن

Grohmann, Adolf. 1952. *From the world of Arabic papyri.* Cairo: Al-Maaref Press.

——. 1958. "The problem of dating early Qur'āns". *Der Islam* 33.213–231.

——. 1963. *Arabic papyri from Ḫirbet el-Mird.* Leuven: Publications Universitaires.

——. 1967–1971. *Arabische Paläographie.* (= *Denkschriften Österreichische Akademie der Wissenschaften. Philosophisch-historische Klasse* 94.) Vienna: Böhlaus.

Gruendler, Beatrice. 1993. *The development of the Arabic scripts.* Atlanta, Ga.: Scholars Press.

Hinds, Martin and Hamdi Sakkout. 1981. "A letter from the governor of Egypt to the king of Nubia and Muqurra concerning Egyptian-Nubian relations in 141/758". *Studia arabica and islamica: Festschrift for Iḥsān ʿAbbās on his sixtieth birthday*, ed. Wadad al-Qadi, 209–229. Beirut: American University of Beirut.

Hoyland, Robert. 2007. "Epigraphy and the emergence of Arab identity". *From al-Andalus to Khurasan: Documents from the Medieval Muslim World*, ed. Petra M. Sijpesteijn, Lennart Sundelin, Sofía Torallas Tovar, and Amalia Zomeño, 219–142. Leiden: E.J. Brill.

James, David Lewis 1992. *The master scribes: Qur'ans of the 10th to 14th centuries A.D.* (= *The Nasser D. Khalili Collection of Islamic Art*, II.) London: Nour Foundation in association with Azimuth Editions and Oxford University Press.

Khan, Geoffrey. 1992. *Arabic papyri, selected material from the Khalili collection.* (= *Studies in the Khalili collection*, I.) Oxford: Nour Foundation in association with Azimuth Editions and Oxford University Press.

——. 1994. "The pre-Islamic background of Muslim legal formularies". *ARAM Periodical* 6.193–224.

——. 2007. "Newly discovered Arabic documents from early Abbasid Khurasan". *From al-Andalus to Khurasan: Documents from the medieval Muslim world*, ed. Petra M. Sijpesteijn, Lennart Sundelin, Sofía Torallas Tovar, and Amalia Zomeño, 201–215. Leiden: E.J. Brill.

Kister, Menachem J. 1964. "*'An Yadin*' (Qur'an, IX/29): An attempt at interpretation". *Arabica* 11.272–278.

Moritz, Bernard. 1905. *Arabic palaeography: A collection of Arabic texts from the first century of the hidjra till the year 1000.* (= *Publications of the Khedival Library*, XVI.) Cairo: Khedival Library.

Nöldeke, Theodor. 1909. *Geschichte des Qorāns.* 3 vols. Leipzig: Dieterich'sche Verlagsbuchhandlung Theodor Weicher.

Rāghib, Yūsuf. 1996. "Les plus anciens papyrus arabes". *Annales Islamologiques* 30.1–19.

Robin, Christian. 2006. "La réforme de l'écriture arabe à l'époque du califat médinois". *Mélanges de l'Université Saint-Joseph* 59.319–364.

Sourdel-Thomine, Janine. 1978. "*Khaṭṭ*". *Encyclopaedia of Islam*, IV, 1113–1120. 2nd ed. Leiden: E.J. Brill.

Sijpesteijn, Petra Marieke. In press. "Arabic papyrology and Islamic Egypt". *Oxford handbook of papyrology*, ed. R.S. Bagnall. Oxford: Oxford University Press.

Petra M. Sijpesteijn (University of Oxford)

Palatalization

1. INTRODUCTION

Palatalization characterizes a set of phonetic and/or phonological processes that have output segments produced with an articulation in or close to the palatal region (for a review, see Bhat 1978). This term has two common meanings: i. palatalization involves the shift of the articulation of a consonant toward the palatal region before front vowels, especially [i], and the glide [j] (e.g., /k/ > [tʃ] in Slavic and Arabic, and t > [tʃ] in Romance and Slavic); ii. palatalization adds a palatal articulation, generally considered as secondary, before front vocoids, without changing its primary articulation (e.g., /k/, /g/ > /kʲ/, /gʲ/ in Acadian French; see Clements and Hume 1995). In this second case, as the phonetic transcriptions show, an off-glide separates the palatalized consonant from the following vowel. Palatalization, as a secondary articulation, can also phonologically distinguish between palatalized vs. nonpalatalized consonants (example: Russian).

2. SOME PHONETIC CORRELATES OF PALATALIZATION

One frequent type of palatalization is that which is responsible for the derivation of a palato-alveolar affricate [tʃ] from a dorsal /k/. This alternation is often attributed to articulatory (Recasens 2003) and/or perceptual (Guion 1998; Flemming 1995) factors. Indeed, the dorsals /k g/ crosslinguistically undergo substantial fronting before front vocoids. In Romance, for example, different allophones (prepalatal, palatal, palato-alveolar, alveopalatal, or alveolar) can result from this fronting. These allophones vary with the Romance dialects and the speaker (Recasens and Espinosa 2003). This variability is due, according to Recasens (2003), to the difficulty of achieving a dorsopalatal closure. Recasens (2003) provides another argument in favor of this hypothesis, and mentions that palatographic observations show that plosive, nasal, and lateral 'palatal consonants', in many languages, are generally produced as alveopalatal ones (Recasens 1990). The alveopalatal (and palato-alveolar) stop possesses a burst that has acoustic properties similar to those of [ʃ], hence [tʃ] is often the final step in the fronting of /k/ before the front vocoids.

The conditioned or context-free palatalization that induces the realization of a segment with a palatal off-glide (Cʲ) is regarded by several phoneticians and phonologists as a vocalic gesture (secondary articulation) that is superimposed on the consonantal gesture (primary articulation). In fact, four types of secondary articulation are possible (labialization Cʷ, palatalization Cʲ, → velarization Cˠ, and pharyngealization Cˤ), and these articulations are analyzed in parallel. Two main arguments can justify the vocalic character of these secondary articulations. First, 'labialization + velarization', 'palatalization', and 'pharyngealization' share several articulatory and acoustic features with the vowels [u] (labiovelar), [i] (palatal), and [a] (pharyngeal) respectively. Second, the gesture of a secondary articulation is often coordinated with that of the primary articulation in such a way that the target of the former is closer to the syllable nucleus, i.e. to the vowel (Sproat and Fujimura 1993).

3. PALATALIZATION IN ARABIC

The least controversial palatalization example in Arabic concerns the alternation /k/ > [tʃ] present in some old and modern Arabic dialects. Indeed, the Arabic grammarians signaled the presence in some dialects of their time of an affricate [tʃ] that is the allophonic realization of /k/ followed by a front vowel (→ *kaškaša*). This special pronunciation of /k/ is regarded by the Arab grammarians as 'faulty' (Moscati 1980). An example of this same alternation is attested nowadays in Jordan, Kuwait, and Iraq. In other modern Arabic dialects (as in Baghdad, Kuwait), /g/, the realization of Classical Arabic /q/ in these dialects, is palatalized before front vowels and becomes [dʲ] (Mitchell 1993).

For the majority of modern linguists, Classical Arabic → *jīm* was generally pronounced [dʒ]: a palatoalveolar affricate (Moscati 1980; Mitchell 1993), but for Cantineau (1960), Classical Arabic *jīm* was [gʲ]: a 'palatalized dorso-palatal plosive'. Whereas the affricate [dʒ] is the more frequent modern reflex of Classical Arabic *jīm*, [gʲ] is attested only in a very few present dialects (Al-Nassir 1993).

Another palatalized consonant, [dj], is mentioned by modern linguists as a possible realization of *jīm* (e.g. in Sudan and Arabia; see Al-Nassir 1993). In some modern Arabic dialects (e.g. in Kuwait and Basra; see Mitchell 1993), we also have an alternation between [dʒ] and the palatal glide [j]; this latter appears in general in the context of a low vowel.

4. PHONOLOGICAL ANALYSES OF PALATALIZATION

This section is limited to two major models proposed for the representation of palatalization. The differences between these models are attributed to the hierarchical feature representations proposed for vocoids.

A first type of analysis of palatalization was initiated by the hierarchical model of Sagey (1986) that uses, for the representation of vowels, the same features proposed by Chomsky and Halle (1968), arranged under different articulatory nodes: [±back], [±high], [±low] dominated by [Dorsal], and [±round] dominated by [Labial]. Front vocoids are specified [Dorsal [-back]], and palatalization is represented by the spreading of [-back] and the insertion of the [Dorsal] node. Other rules and principles, generally language specific and often ad hoc (Lahiri and Evers 1991), are added to explain the different possible outputs of this process.

In the second type of analysis, inspired by the model of Clements (1989; see also Clements and Hume 1995), vowels and consonants are represented by identical place features, but these features are dominated by two different class nodes: [C-place] for the consonants and [V-place] for the vocoids. The front vocoids are specified [V-place [Coronal [-anterior]]]. The palatalization /k/ > [tʃ] is analyzed by the spreading of the complex [Coronal [-anterior]], from the front vocoid, to the [C-place] node of /k/, replacing its [Dorsal] node. The palatalization /k g/ > [kʲ gʲ] is treated as the spreading of [Coronal [-anterior]] and its association to [C-place] after the 'interpolation' of [V-place] (and [Vocalic]) node. This model formally represents, then, the hypothesis according to which the secondary articulations are vocalic gestures.

Even though Cantineau's hypothesis (Classical Arabic *jīm* was [gʲ]) might explain why *jīm* did not behave as a *šamsī* (or coronal) letter, other factors weaken this hypothesis. Indeed, it is an isolated hypothesis, and [gʲ] is present only in very few modern Arabic dialects. It is more probable that Classical Arabic *jīm* was phonetically a coronal consonant but continued to behave phonologically as its proto-Semitic cognate /g/.

BIBLIOGRAPHICAL REFERENCES

Al-Nassir, A.A. 1993. *Sibawayh the phonologist*. London and New York: Kegan Paul International.

Bhat, D.N. 1978. "A general study of palatalization". *Universals of human language*. II. *Phonology*, ed. Joseph Greenberg, 47–92. Stanford, Calif.: Stanford University Press.

Cantineau, Jean. 1960. *Cours de phonétique arabe*. Paris: Klincksieck.

Chomsky, Noam and Morris Halle. 1968. *The sound patterns of English*. New York: Harper and Row.

Clements, Georges N. 1989. "The representation of vowel height". Paper given at the Conference on Features and Underspecification, Massachusetts Institute of Technology. Unpublished ms. (preliminary version), Cornell University, Ithaca, N.Y.

——, and Elizabeth Hume. 1995. "The internal organization of speech sounds". *The handbook of phonological theory*, ed. John Goldsmith, 245–306. Oxford and Cambridge, Mass.: Blackwell.

Flemming, Edward. 1995. *Auditory representations in phonology*. Ph.D. diss., University of California at Los Angeles.

Guion, Susan G. 1998. "The role of perception in the sound change of velar palatalization". *Phonetica* 55.18–52.

Lahiri, Aditi and Vincent Evers. 1991. "Palatalization and coronality". *Phonetics and phonology*. II. *The special status of coronals: Internal and external evidence*, ed. Carole Paradis and J.-François Prunet, 79–100. San Diego, Calif.: Academic Press.

Mitchell, T.F. 1993. *Pronouncing Arabic*. Oxford: Clarendon Press.

Moscati, Sabatino (ed.). 1980. *An introduction to the comparative grammar of the Semitic languages: Phonology and morphology*. 3rd ed. Wiesbaden: O. Harrassowitz.

Recasens, Daniel. 1990. "The articulatory characteristics of palatal consonants". *Journal of Phonetics* 18.267–280.

——. 2003. "Articulation and sound change in Romance". *XVth International Congress of Phonetic Sciences*, Barcelona, 231–234.

——, and Aina Espinosa. 2003. "Articulatory motivation of velar softening". *Proceedings of the XVth International Congress of Phonetic Sciences, Barcelona*, ed. Maria Josep Solé, D. Recasens, and Joaquín Romero, 1903–1906. Barcelona: Futurgraphic.

Sagey, Elizabeth. 1986. *The representation of features and relations in nonlinear phonology*. Ph.D. diss., Massachusetts Institute of Technology.

Sproat, Richard and Osamu Fujimura. 1993. "Allophonic variation in English /l/ and its implications for phonetic implementation". *Journal of Phonetics* 21.291–311.

CHAKIR ZEROUAL
(University Sidi Mohamed Ben-Abdellah and CNRS)

Palestinian Arabic

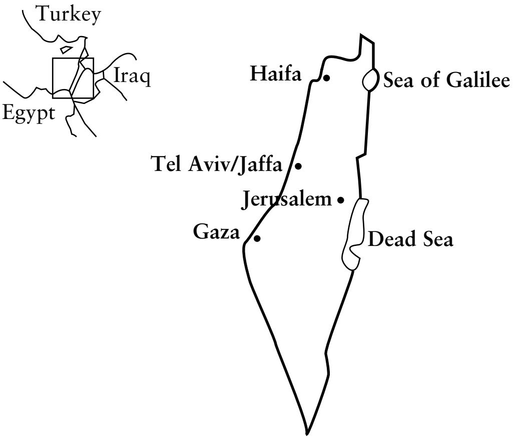
This entry describes Palestinian colloquial Arabic: its phonemic system, prosody, phonotactics, morphophonology, morphology, syntax, semantics, and lexicon. The description is based on the documentation of a sizeble body of literature since the early 1900s.

1. GENERAL

1.1 *Area*

Palestinian Arabic is spoken in Palestine (Israel, West Bank, and Gaza Strip; see Map 1). As more than 50 percent of Palestinians live elsewhere, it is also spoken around the world.

Map 1. Dialect area of Palestinian Arabic



1.2 *Speakers*

Palestinian Arabic is a native language to approximately 8.5 million people. The lifestyles in the dialect area are urban, rural, Bedouin, and Gypsy. In 1948 and 1967, when the State of Israel was formed and expanded on Palestinian land, many rural families resettled in towns and cities, so the number of speakers with an urban lifestyle has increased. Most speakers in the dialect area live under military occupation by the State of Israel, a large number in refugee camps. The others live inside Israel, subject to discrimination. Outside the dialect area, the speakers in refugee camps live in depressed conditions.

1.3 *Position and varieties*

Palestinian Arabic is a collection of mutually intelligible dialects which vary by social group and location. The social groups are urban, rural, Bedouin, and Gypsy. Urban, rural, and Bedouin speech differ in consonant phonemes. Urban speech frequently has → *ʾimāla* and more restricted ‘emphasis spread’. It uses *ma-* for verb negation, whereas rural speech also uses *ma-* and *-š*; Bedouin uses either system. Certain words are dialect specific, e.g. urban *hōn*(*i*) ‘here’ vs. rural/Bedouin *hān*(*a*). Not much is known of current Gypsy speech (→ Gypsy Arabic). Another socially based dialect is ‘intellectual’, which approximates the pronunciation and lexicon of Modern Standard Arabic (→ Educated Arabic). The locations are northern, southern, eastern, western, central, or coastal. Northern dialects share features with Lebanese Arabic. The Gaza dialect shares features with Egyptian. Southern Bedouin is similar to Jordanian; the two are sometimes grouped together. Certain cities have their own unique dialect, e.g. Nablus and Khalīl (Hebron). Differences exist based on religion; Christian speech might tend to be more urban. Due to interdialectal contact, the language on the street is now currently a mix of features. However, no general West Bank or Gaza dialect is emerging; the territorial cells enforced by the Israeli army help the nonhomogenization. Nonurban Palestinians are typically bidialectal, speaking their marked variety and an urban dialect. Schoolchildren are typically adopting urban speech, and rural varieties are possibly becoming more urban, so some marked dialects are endangered. Urban dialects have influenced Jordanian Arabic (most Jordan residents are Palestinian). Palestinian Arabic is not written or used in official media. In a computational setting, a system for writing Levantine Arabic has been developed.

1.4 *Historical evidence*

Early Palestinian texts date from 800 to 1000 C.E. Little documentation exists from 1000 C.E. until the Orientalist studies of the early 1900s.

1.5 *State of research*

There are several grammatical descriptions, dictionaries, and text collections (including Seeger 1998, with audio data on the Internet) and also theoretical linguistic studies. However, more work is needed in all areas, especially comprehensive documentation and comparison of the dialects, text collection, and semantics, in celebration of Palestinian culture and identity, of which the language is an integral part.

2. LINGUISTIC DESCRIPTION

2.1 *Phonology*

2.1.1 Inventory

2.1.1.1 Consonants

The consonant phonemes of urban, rural, and Bedouin dialects differ (see Fig. 2, Table 1). Some rural dialects realize *ġ* as velar [ɣ]. In rural dia-lects, pharyngealized *q* is realized between velar and uvular (transcribed as *ḳ*). *x* is realized always as a postvelar [χ]. Phonetic [p] occurs by *b*-devoicing; *ġ* and ʿ are approximants. ‘Emphasis’ is uvularization with pharyngealization.

Table 1. Consonant cognates of Palestinian Arabic dialects

urban	k	ʾ	ž	t	d or z	ḍ or ẓ
rural	č or k	k or ḳ	j	ṯ	ḏ	ḏ̣
Bedouin	k	g or G	j	ṯ	ḏ	ḏ̣

In the historical development of consonants (Fig. 3),*_q_ has disappeared (except in intellectual speech); urban dialects have also lost *_ṯ_ *_ḏ_ *_ḏ̣_. Innovations are _ṃ ḅ_; and _ž ẓ_ in urban, _č ḳ_ in rural, _g G_ in Bedouin speech. More study might reveal whether the reflexes of one consonant, e.g. urban _d z_ < *_ḏ_, are free variants, allophones, or dialectal variants. Mergers are: *_q_ *ʾ > ʾ, *_t_ *_ṯ_ > _t_, *_d_ *_ḏ_ > _d_, *_ḏ_ *_z_ > _z_, *_ḍ_ *_ḏ̣_ > _ḍ_ in urban, *_k_ *_q_ > _k_ in rural. Within root morphemes, consonants tend to differ in place of articulation. Phonemic ʾ is infrequent morpheme-initially. Emphatic _ṃ ḅ ḷ_ are infrequent and do not lexically contrast with their nonemphatic counterparts, unlike, for example, _ḏ̣ ṣ ṭ_, e.g. _ṭīn_ ‘mud’, _tīn_ ‘figs’.

2.1.1.2 Vowels

Environmental conditioning of the vowels (Fig. 4) yields backed _a ā_ (vs. front _ä ǟ_), lax short vowels, ʌ and ə. Lax vowels do not occur morpheme finally. Low-frequency short _e_ (_ẹ_) < *_a_ and _o_ < *_u_ are innovations in progress. (Some centenarians have them.) Final _a_ < *_āʾ_ occurs, e.g. _šita_ ‘winter, cold weather’. Some northern rural dialects have merger of *_u_ with *_i_ and _a_.

Table 2. Palestinian Arabic consonants

labial	labiodental	interdental	alveolar	post-alveolar	palatal	velar	uvular	pharyngeal	glottal
a. urban									
b ḅ			t ṭ d ḍ			k			ʾ
m ṃ			n						
			ṛ						
	f		s ṣ z ẓ	š ž			x ġ	ḥ ʿ	h
			l ḷ						
					y				
other consonant: w – voiced labiovelar approximant									
b. rural									
b ḅ			t ṭ d ḍ			k ḳ			ʾ
m ṃ			n						
			ṛ						
	f	ṯ ḏ ḏ̣	s ṣ z	š			x ġ	ḥ ʿ	h
			l ḷ						
					y				
other consonants: w; č – voiceless post-alveolar affricate; j – voiced post-alveolar affricate									
c. Bedouin									
b ḅ			t ṭ d ḍ			k g	G		ʾ
m ṃ			n						
			ṛ						
	f	ṯ ḏ ḏ̣	s ṣ z	š			x ġ	ḥ ʿ	h
			l ḷ						
					y				
other consonants: w; č (some dialects); j									

2.1.1.3 Diphthongs

The diphthongs *ay*, *aw* usually occur followed by long or short *y* and *w*, respectively, e.g. *ṃayy*(*a*) [m̴ɑijjː(a)] 'water', *ṭawla* [t̴ɑuwla] 'table'; *ay* has the variant *oy* in *šwoyy(a)* 'some, a bit'.

2.1.1.4 Syllables

Syllables are (C)(C)V(C)(C) or (C)(C)V̄(C). Phrase-initially, phonetic [ʔ] is inserted at the beginning of a vowel-initial syllable, e.g. [ʾ]*īd* 'hand'. Word-internal and word-final vowel-initial syllables, and word-final (C)CV̄ are infrequent. Word-internal (C)(C)V̄C occurs only before object/possessive pronouns, e.g. *šāf.-ha* 'he saw her', *bāb.-na* 'our door'. Nasals and *l* can be syllabic, e.g. *m.sax.xan* (type of main dish). Vowel hiatus is tolerated, e.g. *maaja* 'he didn't come'.

2.1.1.5 Consonant clusters

Word-initial CCC clusters occur rarely, with equal, then rising sonority, e.g. *ḅ̥tʿaṭīh* 'you give to him'. Word-internal CCC onsets and CCC codas do not occur. Onset CC occurs with equal or rising sonority, e.g. *fhimit* 'I understood', *kṛanīyya* [type of herb]; word-initially the first consonant can be a devoiced nasal, e.g. *ṃ̥sax.xan* [type of main dish]. Forms with initial CC vary freely with VC epenthetic forms, e.g. *ḥmāṛ* ~ *iḥ.māṛ* 'donkey'. CC codas occur word-internally with equal or falling sonority, e.g. *masakt.ha* 'I held it [fem.]', *kult.la* 'I said to him'. Word-final CC codas occur only phrase-internally, e.g. *šribt ḥalīb* 'I drank milk' vs. *šribit* 'I drank', unless the first consonant is *ṛ* or an approximant, e.g. *hāda iddars* 'that's the lesson', *bitaʿb* 'he gets tired'.

2.1.1.6 Stress

The primary stress weight scale is basically V̄(C) > VCC > VC > V. The heaviest rightmost syllable in the word is stressed. If all syllables are V, the leftmost, maximally antepenultimate syllable is stressed. Prefixal syllables are

Table 3. Historical development of Palestinian Arabic consonants

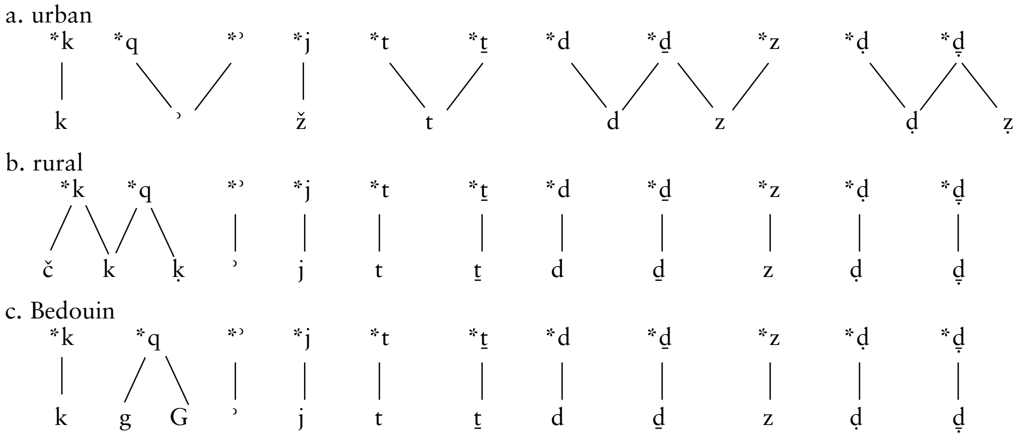

a. urban
*k *q *ʾ *j *t *ṯ *d *ḏ *z *ḍ *ḏ̣
k ʾ ž t d z ḍ ẓ
b. rural
*k *q *ʾ *j *t *ṯ *d *ḏ *z *ḍ *ḏ̣
č k ḳ ʾ j t ṯ d ḏ z ḍ ḏ̣
c. Bedouin
*k *q *ʾ *j *t *ṯ *d *ḏ *z *ḍ *ḏ̣
k g G ʾ j t ṯ d ḏ z ḍ ḏ̣

Table 4. Palestinian Arabic vowels

	front		back
high	i ī		u ū
mid	e ē		o ō
low		a ā	

not considered for stress assignment, and a word-final consonant never is, e.g. *ta-t-.ʿú.ṣuṛ* ‘when you [masc.] will squeeze’. Object and possessive suffixes shift primary stress to the stem-final syllable, e.g. *ṛa.mā́.-ha* ‘he threw it [fem.]’, *a.bū́.-ha* ‘her father’; direct object shift supercedes indirect object shift, e.g. *ṛa.ma-l.-hay.yā́-h* ‘he threw it [masc.] to her’. Stress shift occurs for the negative *-š*, superceding other shifts, e.g., *ma-ṛa.ma-l.-hā́.-š-iy.ya-h* ‘he didn’t throw it [masc.] to her’. Negative *ma-* can attract stress in urban dialects, e.g. *má-šuftha* ‘I didn’t see her’. The epenthetic vowel, *i*, is invisible to stress, e.g. *katábit* ‘I wrote’, although this is variable for shifted stress, e.g. *jábilha*, *jabílha* ‘he fetched for her’. In Bedouin speech, stress can fall exceptionally on the vowel of the definite article in *masálxēr* ‘good evening’. The stress system is trochaic and binary. Secondary stress is phonetic and rhythmic, on alternate syllables either side of the primary stress, e.g. *ma.-ṛà.ma-l.-hā́-.š-iy.yà-h* (Broselow a.o. 1997).

2.1.2 Phonotactics

Sun letters for total assimilation of the *l* of the article include *j*. Total assimilation occurs with *min* ‘from’ in *miš-šāni* ‘for me [lit. ‘from my sake’]’ and *yōm mil-layām* ‘one day [lit. ‘a day from the days’]’, and with *wēn* ‘where’ in *werrayiḥ* ‘where are you [masc. sg.] going?’, also in *ha-ʾ-ʾēta* (< *ha-l-ʾēta*) ‘now’ and *maḥ-ḥa* (< *maʿ-ha*) ‘with her’. Lower-sonority features tend to be preserved. Emphatic articulation spreads to other consonants, e.g. *ṭāwḷ-āṭ* ‘tables’, *ṇḍīf* ‘clean [masc. sg.]’; *š* blocks this, e.g. *ʿaṭšān* ‘thirsty [masc. sg.]’, as do other post-alveolars, *w*, *y*, and high vowels, depending on dialect (Davis 1995). The *ṛ* de-emphaticizes in certain environments. In rural dialects, *k* palatalizes (and becomes an affricate) in the environment of *i* or *ī*, e.g. *dīč* ‘rooster’. In some rural dialects, *č* has replaced *k* as phoneme; in others, *č* occurs outside the palatalization context in only some words, e.g. *čalb* ‘dog’, but *kursi* ‘chair’ (Naïm 1999). Nasal spread can occur, for instance affecting the imperfect *b-*, which becomes *m-* on account of the prefix *n-* in *m-in-ṛawwiḥ* ‘we’re going home’. Assimilatory voicing and devoicing occur, e.g. *ẓġīr* < *ṣġīr* ‘small [masc. sg.]’, *p-suṛaʿ* < *b-suṛaʿ* ‘quickly’.

Low *a ā* are backed in words with an emphatic consonant. In the numbers 11–19, backing occurs as a remnant of deleted *ṛ* of *ʿäšaṛa* ‘ten’ (Sec. 2.2.5). In a closed syllable, backed *a* raises to [ʌ]. Short vowels are laxed in a closed syllable and adjacent to a postvelar (emphatic, uvular, pharyngeal, or glottal) consonant; the laxing spreads throughout the word, e.g. *bịnịt* ‘girl’. Phonetically, *i u* lower in the context of a postvelar, e.g. *ḥẹlu* ‘pretty [masc. sg.]’, *mọṛṛ* ‘bitter [masc. sg.]’. Strong phonetic lowering of epenthetic *i* occurs in words like *ṣubaḥ* ‘morning’. The epenthetic vowel assimilates to root *u*, e.g. *kutub* ‘books’. Short *a* or *i* can occur as *u* conditioned by an emphatic or velar, e.g. *yunbuɣ* (not *yinbuɣ*) ‘he excels’. In northern and some urban dialects, *ʾimāla* raises word-final *a* to *e* (sometimes *ẹ*), e.g. *zalame* 'man'; this can be blocked by postvelar consonants, e.g. *šajaṛa* 'tree'. Rural dialects have unconditioned raising and fronting, e.g. *midrasa* 'school', *ṯimm* 'mouth' (cf. urban *madrasa*, *tumm*). Unstressed short *a* typically reduces to schwa, e.g. *šajəṛə* 'tree'.

Some words have historical metathesis, e.g. *jōz* 'husband' < *jōwz* < *zōwj*. Where affixation yields CVCCVCV, metathesized CVCVCCV occurs, e.g. *bikitb-u* 'they [masc.] write', via syncope, then epenthesis.

2.1.3 Morphophonology

Consonants occasionally elide morpheme-finally, e.g. *ha-l-kē* 'now' (~ *ha-l-kēt*). Morpheme-initial elision occurs in *min-ēn* (< *min wēn*) 'from where?'. Unstressed *i* and *u* elide in non-final open syllables, e.g. *fhímit* (not *fihímit*) 'I understood', and after geminates, e.g. *sakfu* (not *sakkifu*) 'clap [masc. pl.]!'. Syncope in the first context has cyclic effects, e.g. *fhímna* 'we understood', but opaque lack of syncope in *fihímna* 'he understood us'. Unstressed *a* can elide, e.g. *katb-at-ha* (not *katabat-ha*) 'she wrote it [fem.]'. On insertion of *ʾ*, see Section 2.1.1.4; *t* is inserted word-finally in the construct state. The vowel *i* is inserted for syllabification of consonant clusters. True geminates are never split by epenthesis, e.g. *biṣubb* (not *biṣubib*) 'he's pouring'.

Geminate consonants typically shorten before another consonant, e.g. *sakfu* 'clap [masc. pl.]!' (cf. *sakkif* 'clap [masc. sg.]!'). Long vowels shorten before another long vowel in the word, e.g. *jibī́-h* 'fetch [fem. sg.] it [masc. sg.]!' (cf. *jī́bi* 'fetch [fem. sg.]!'), and in a closed syllable before a consonant-initial subject suffix, e.g. *ṛuḥ.-na* 'we went' (cf. *b-a-ṛūḥ* 'I go'). Exceptionally, *ṛāḥ.-na* 'we went' can occur. The shortening can occur in forms with *-l-* 'for' and an object pronoun, e.g. *jáb-il-ha* 'he fetched for her', in which the epenthetic vowel shows its invisibility to vowel shortening (cf. *jāb* 'he fetched'). Lengthening occurs in *minn-i* 'from me', *minn-ina* 'from us'. Short vowels lengthen in an open syllable under shifted stress, e.g. *abū́-ha* 'her father' (cf. *ábu* 'father').

2.2 *Morphology*

In urban and northern dialects, masculine and feminine are not distinguished in the plural. In other dialects, they usually are. Gender can be distinguished for 1st person in noun phrases, e.g. *iḥna sakʿan-āt* 'we [fem.] are cold'. Many words are truncations, e.g. the numbers 11–19 (Sec. 2.2.5), *eš* 'what [nom.]?' (< *ay šayʾ* 'what thing?'), *šu(hu)* 'what?' (< *ay šayʾ hu* 'what thing is it [masc.]?'), *m-iš* 'not' (< *ma* 'what?', *šayʾ* 'thing'), *yiḻ̣ḻ̣a* 'let's go!' (< *ya aḻ̣ḻ̣ah* 'O God'), *animma* 'which [pl.]?' (< *anu himma* 'which are they [masc.]?')

2.2.1 Pronouns

2.2.1.1 Personal independent pronouns

With the subject pronouns (see Table 5), final *a* can become *e* by *ʾimāla*; *niḥna* 'we' occurs in Negev Bedouin Arabic. Independent possessive subject pronouns are *il-* with a possessive dependent pronoun, e.g. *il-ak* 'yours [masc. sg.]'.

Table 5. Subject independent pronouns

	singular	plural
3 masc.	*hu(wwa)*	urban/Bedouin: *hum(ma)* rural: *him(ma)*
3 fem.	*hi(yya)*	rural/Bedouin: *hin(na)*
2 masc.	*inta*	*intu*
2 fem.	*inti*	rural/Bedouin: *intin*
1	*ana*	*iḥna*

2.2.1.2 Possessive/object dependent pronouns

The clitic possessive/object pronouns for verbs, prepositions, and nouns (Table 6) serve for direct and indirect objects, e.g. *bi-ṭaʿmi-ha-yyā-ha* 'he's feeding it [fem.] to her [fem. sg.]'.

2.2.1.3 Demonstratives

The demonstratives are *hay* 'this', *hād(a)* [masc.], *hādi* [fem.] 'this', *hadāk* [masc.], *hadīk* [fem.] 'that', *hadōl* 'these [masc., fem.]', *hadolāk* 'those [masc., fem.]'. (Rural and Bedouin have *ḏ̣* in these forms, except for *hādi* and *hadīk*). Other demonstratives are *huta*, *hita* 'this, that [human, specific] [masc., fem.], and *ha-* 'such (a), that, this, these, those [masc., fem.]', which is the only demonstrative that occurs with the definite article, e.g. *ha-l walad* 'such a boy'.

2.2.1.4 Presentative

The presentative is *hay(ū)* (sometimes *hayā*) 'this is, here is'. It can be followed by a noun or object pronoun, e.g. *hay(ū)-ha* 'this is she',

Table 6. Possessive/object dependent pronouns

	singular	plural
1	for verbs: *-ni* for prepositions and nouns after a consonant: *-i* for prepositions and nouns after a vowel: *-i/(yy)a*	*-na*
2 masc.	after a consonant: *-ak* after a vowel: *-k*	urban: *-ko(m)/ku(m)* rural/Bedouin: *ču(m)*
fem.	after a consonant: urban: *-ik*; rural/Bedouin: *-ič* after a vowel: urban: *-ki*; rural/Bedouin: *-č*	rural/Bedouin: *čin*
3 masc.	after a consonant: urban: *-o(h)/-u(h)*; rural/Bedouin: *-a(h)* after a vowel: *-(h)*	urban/Bedouin: *-hum* rural: *-him*
3 fem.	*-ha*	rural/Bedouin: *-hin*

‘here she is’; *-ni* is used for 1st person singular object, e.g. *hay-ni* ‘here I am’.

2.2.1.5 Relative pronouns

The relative pronoun is *illi/alli* ‘which, whichever, who(m), who(m)ever’. Interrogative pronouns and *lemma* ‘when, whenever’ are used as relative pronouns, as are interrogative pronouns, e.g. *kultilha min inta* ‘I told her who you [masc. sg.] are’. In Negev Bedouin Arabic, *alli* can be shortened to *al-*, and demonstrative *ha-lli* ‘whoever, he who, whatever, that which’ (< *ha-alli*) occurs.

2.2.1.6 Interrogative pronouns

Interrogative pronouns are *eš* ‘what [nom.]?’; *ēš* ‘what [acc.]?’; *šu* ‘what?’; *šuhú*, *šu-hí* ‘what [specific]? [masc. sg., fem. sg.]’; *min* ‘who [nom]?’; *mīn* ‘whom?’; *an-ú*, *an-í*, *an-imma* ‘which?’ [masc. sg., fem. sg., pl.]; *lēš* ‘why?’; *waḳtēš*, *ēmta* ‘when?’; *wēn* ‘where?’; *kīf* ‘how?’ (*čēf* occurs in Gaza); *kam* ‘how many?’; *ḳaddēš* ‘how much?’. They are usually clause-initial. The form *wēn* can take an object suffix, e.g. *wēn-ak* ‘where are you [masc. sg.]?’. Frozen _*ma-_ ‘what?’ occurs in *ma-l-ak* ‘what’s [the matter] with you [masc. sg.]?’

2.2.1.7 Indefinite pronouns

The forms *iši* ‘something’ and *(wa)ḥad* ‘(some)one’ are indefinite pronouns.

2.2.2 Adverbs

Adverbs include *hōn(i)*, *hān(a)*, *hin(a)* ‘here’ (urban, rural/Bedouin, Gaza); *honāk*, *hanāk*, *hināk* ‘there’ (urban, rural/Bedouin, Gaza); *halkēt(a)*, *hellaʾ* ‘now’; *f-ilwaḳit hay* ‘then’; *ḳabil* ‘before’; *baʿden* ‘after’; *hēk*, *hēḏ* ‘(like) so’; *wēn makān* ‘anywhere’; *kull makān* ‘everywhere’; *wala maḥāl* ‘nowhere’; *dayman*, *dāʾiman* ‘always’; *maṛṛāt* ‘sometimes’; *wala maṛṛa*, *bi-l maṛṛa*, *abidan* ‘never’; *laḥāl-* ‘alone’ and *maʿ bāʿḍ* ‘together’, e.g. *ajīna laḥāl-na* ‘we came alone’, *ajīna maʿ bāʿḍ(-na)* ‘we came together’. Adverbs of time include *bukṛa* ‘tomorrow’, *ilyōm* ‘today’, *iṣṣubaḥ* ‘in the morning’. Adverbs of manner include š*woy(a)šwóy* ‘carefully’, *sawasáwa* ‘fairly, nicely’, and nouns prefaced with *b(i)-* ‘in, with’, e.g. *ṛāḥ ib-suṛaʿ* ‘he went quickly [lit. ‘with haste’]’. Sentential adverbs include *inš(a)aḷḷa* ‘God willing’; *ḥamdilla*, *al-ḥamdulillah* ‘thank God!’; *lāzim* ‘necessarily’; *mumkin*, *yimkin* ‘possibly’; and *maʿlūm* ‘of course’.

2.2.3 Particles

In addition to the definite article *l* ‘the’, particles include the prepositions *ʿa(la)* ‘to, at, for’; *la*, *-l-* ‘for, to’; *taḥt* ‘under, below’; *fōk* ‘over, above’; *f(i)-* ‘in, with, at’; *b(i)-* ‘in, with’; *waṛa* ‘behind’; *kuddām* ‘in front of’; *bēn* ‘between’; *ʿan* ‘away from’; *maʿ* ‘with’; *juwwa* ‘inside (of)’; *baṛṛa* ‘outside (of)’; *min* (*miṇ*) ‘from’; *ḥawalē* ‘around’; *jamb* ‘beside’; *ʿend* ‘with [French *chez*], to, in the possession of’; *la* occurs as genitive marker, e.g. *hādi la walīd* ‘this [fem.] is Walid’s’; *fi* takes *-ni* or *-yya* for 1st person singular object. The particle *fī(h)* ‘there is’ is an existential marker. It sometimes occurs as *f-*. Urban and some Bedouin dialects use *ma-* for negation, e.g. *ma-šufit* ‘I didn’t see’, *ma-ba-šūf* ‘I don’t see’. Other dialects use *ma- . . . -š* for a perfect tense verb, otherwise *-š*, e.g. *ma-šuft-iš* ‘I didn’t see’, *bašūf-iš* ‘I don’t see’. The prepositions *maʿ* and *ʿind*, existential *fī(h)*, and independent pronouns can be negated, e.g. *ma-maʿ-i*, *maʿ-ī-š* ‘I don’t have (any)’ (urban, rural/Bedouin); *ma-fī(h)*, *f-iš(i)* 'there isn't any' (urban, rural/Bedouin); *ma-hiyy binti*, *ma-hiyy-iš binti* 'she's not my daughter' (urban, rural/Bedouin). Other forms of negation use *m-iš* 'not', e.g. *m-iš taʿbān* 'not tired'. The form *wala* 'no, not a, not any' is used for nouns, e.g. *wala ġēma* 'not a cloud'; *laʾ* 'no' can negate an imperative, e.g. *laʾ t-īji* 'don't come [fem. sg.]!'. Conjunction particles include *u* 'and'; *willa*, *aw* 'or'; *bass*, *lākin* 'but'; *ʿa(la)šān* 'because'; *illa* 'except'; *zay ka* 'as if'; *ḥatta*, *ʿabēn* 'until'. The vocative particle is *ya*, e.g. *ya ṛabbi* 'o Lord'. Exclamations include *waḷḷa* 'really!', *ʿan jadd* 'truly', *waḷḷāhi* 'truly/I swear', *mašaḷḷa*, *maša(ʾ)aḷḷah* (expression of awe and praise), *yiḷḷa* 'let's go', and *ʿal yōm* 'if only!'.

2.2.4 Nouns

A noun is feminine if it is inherently female, is a body part of which there are two, or ends in *-a*. A masculine noun can be made feminine with the feminine singular suffix *-a*, e.g. *mʿallim-a* 'teacher [fem.]'. Dual is marked by *-ēn*, e.g. *bint-ēn* 'two girls'. Regular plurals are formed with *-īn* [masc. pl.], *-āt* [fem. pl.]. Irregular plurals are formed by an alteration of the segmental pattern of the singular, e.g. *malāḳit* 'pins, tongs' (sg. *milḳat*). Nouns are definite or indefinite, e.g. *ilwalad* 'the boy', *walad* 'a boy'. Partitives/diminutives are sometimes formed with feminine endings, e.g. *ruzz-a* 'bit of rice', *snan-āt* 'little teeth'. Endearment forms for some nouns use medial consonant gemination and *ū* vocalism, or a possessive suffix, e.g. *binnūt* '(dear) girl', *walad-i* 'my (dear) boy'. Nouns are primitive, e.g. *ḥṣān* 'horse', *ruzz* 'rice', or derived from a noun, e.g. *dirāsa* 'studies, lessons' (cf. *drūs* 'studies, lessons'), verb, or verb participle (Sec. 2.2.6.4). Some are reduplicative, e.g. *mišmiš* 'apricot'. Southern Bedouin dialects have the → '*gahawa* syndrome', inserting *a* between one of *x ġ ḥ ʿ h* and another consonant when they are preceded by *a*, e.g. *gahawa* 'coffee'. Adjectives are primitive, e.g. *mnīḥ* 'good', or derived from a noun or verb. Adjectives from a noun use the *-i(yya)* → '*nisba*' suffix, e.g. *balad-i, balad-iyya* 'local, domestic [masc. sg., fem. sg.]' (cf. *balad* 'land, country'), or *-ān*, e.g. *bard-ān* 'chilled [masc. sg.]' (cf. *bard* 'chill'). Participles can function as adjectives, e.g. *iši maʿṛūf* 'something known [masc. sg.]'. Colors and deficiencies use verbal Form IX (without the final geminate), e.g. *aḥmaṛ* 'red', *aṭruš* 'deaf', or *-i*, e.g. *lēlak-i* 'violet'. Comparatives and superlatives use the → elative form, e.g. *aḥsan* 'better, best'. Adjectives are inflected for gender and number like nouns; sometimes the plural is irregular. They agree in gender, number, and definiteness with the noun they modify; masculine plural sometimes replaces feminine plural, e.g. *bināt šaṭrīn* 'clever girls'.

2.2.5 Numerals

Cardinal numbers from 1 to 21 are *waḥad*, *itnēn* ~ *tintēn*, *talāt(a)*, *aṛbaʿ*, *xamsa*, *sitta*, *sabaʿ*, *tamania*, *tisaʿ*, *ʿašaṛa*, *ḥdāš*, *tnāš*, *talattāš*, *aṛbʿatāš*, *xamstāš*, *sittāš*, *sabʿatāš*, *tamantāš*, *tisʿatāš*, *ʿašṛīn*, *waḥad u ʿašṛīn*. Cardinals 11–19 are truncations of the phrase 'X *il ʿašaṛa* with the construct state *t*, e.g. *aṛbʿatāš* < *aṛbʿa-t ilʿašaṛa*. Other cardinals are *talatīn* 'thirty', *aṛbʿīn* 'forty', *xamsīn* 'fifty', *sittīn* 'sixty', *sabʿīn* 'seventy', *tamanīn* 'eighty', *tisʿīn* 'ninety', *miyya* 'one hundred', *waḥadu miyya* 'one hundred and one', *miytēn* 'two hundred', *talāt miyya* 'three hundred', *alf* 'one thousand', *malyōn* 'one million', *malyāṛ* 'one billion'. Cardinals 3–10 do not have agreement polarity. The ordinal numbers from first to thirteenth are *awwal* ~ *ūwwal* (fem. *ūwl-a*), *tāni*, *tālit*, *ṛābaʿ*, *xāmis*, *sādis*, *sābaʿ*, *tāmin*, *tāsaʿ*, *ʿāšiṛ*, *ḥdāšiṛ*, *tnāšiṛ*, *talattāšiṛ*; others are *ilʿašṛīn* 'twentieth', *ilmiyya* 'hundredth', *ilʾalf* 'thousandth'.

2.2.6 Verb

Denominative verbs include *malliḥ* 'to make salty' (cf. *maliḥ* 'salt'). This section presents indicative and imperative moods; for subjunctive, see Section 2.3.3.

2.2.6.1 Verbal Forms (patterns)

2.2.6.1.1 Form I

In Form I, strong verbs in the imperfect tense have *i* in the first syllable and *i* or *a* in the second syllable, e.g. *iktib* 'write', although *u* occurs in forms with an emphatic, e.g., *uḍrub* 'hit'.

2.2.6.1.2 Derived

Several secondary verb forms are used productively (see Table 7, from Munther Younes p.c.). Productive forms comprise an active set and a set of their passive/reflexive counterparts. Other forms occur nonproductively: Form IV

Table 7. Productive secondary verb forms

Form	Meaning of the Form	Example imperfect	perfect	meaning
II: $C_1VC_2C_2VC_3$	intensive/extensive	*jammiʿ*	*jammaʿ*	'to gather'
	causative	*farriz*	*farraz*	'to freeze'
	declarative/estimative	*saddiḳ*	*saddaḳ*	'to believe'
	denominative	*malliḥ*	*mallaḥ*	'to make salty' (cf. *maliḥ* 'salt')
III: CVVCVC	reciprocal (participative)	*kātil*	*kātal*	'to fight'
V: *t*-$C_1VC_2C_2VC_3$	reflexive of II	*tfarriz*	*tfarraz*	'to be frozen'
	denominative	*tġannim*	*tġannam*	'to become a goat owner' (cf. *ġinim* 'goats')
VI: *t*-CVVCVC	reflexive of III	*tkātil*	*tkātal*	'to fight with each other'
VII: *n*-CVCVC	passive/reflexive of I	*nḥaṛiḳ*	*nḥaṛaḳ*	'to burn oneself' (cf. *ḥaṛiḳ* 'burn')
VIII: *i*-C*t*VCVC	passive/reflexive of I	*iktašif*	*iktašaf*	'to discover [i.e. 'to uncover for oneself']' (cf. *kašif* 'uncover')

(*a*CCVC, causative or denominative), e.g. *atlij* 'to snow'; Form IX (*i*-$C_1C_2VC_3C_3$, becoming [color or quality]), e.g. *iḥmaṛṛ* 'to become red'; and Form X (*ista*-CCVC, denominative [becoming or thinking of something or oneself as]), e.g. *istaẓġir* 'to think of oneself as small'.

2.2.6.1.3 Passive

Forms VII and VIII express passive voice, e.g. *nkasaṛ* 'it [masc.] was broken', *intafax* 'it [masc.] was inflated'.

2.2.6.2 Inflection of imperfect and perfect

2.2.6.2.1 Imperfect

Imperfect tense is formed with *b*- and subject inflectional affixes (see Table 8, for strong verbs). Some urban and southern dialects use *y*- for 3rd person masculine singular and 3rd person plural e.g. *b-y-iktib* 'he's writing'. The positive imperative is the imperfect stem, e.g. *iktib* 'write [masc. sg.]!'. The negative imperative has 2nd person subject inflections, e.g. *ma-t-iktib*, *t-iktib-iš* 'don't write [masc. sg.]!' (urban, rural/Bedouin).

Table 8. Imperfect conjugation of strong verbs (*iktib* 'he writes')

	singular	plural
3 masc.	*b-iktib*	*b-ikitb-u*
3 fem.	*b-t-iktib*	*b-ikitb-in*
2 masc.	*b-t-iktib*	*b-t-ikitb-u*
2 fem.	*b-t-ikitb-i*	*b-t-ikitb-in*
1	*b-a-ktib*	*b-n-iktib*

2.2.6.2.2 Perfect

Perfect tense inflection has *a* vocalism (see Table 9, for strong verbs).

Table 9. Perfect conjugation of strong verbs (*katab* 'he wrote')

	singular	plural
3 masc.	*katab*	*katab-u*
3 fem.	*katb-at*	*katab-in*
2 masc.	*katab-it*	*katab-tu*
2 fem.	*katab-ti*	*katab-tin*
1	*katab-it*	*katab-na*

2.2.6.3 Participle forms

The active participle of Form I is CāCiC, e.g. *kātib* 'writing'. The passive participle is *ma*-CC*ū*C, e.g. *maktūb* 'written'. The active participles of productive secondary verb Forms are *m*-C*a*CC*i*C (II), e.g. *mkammil* 'finishing'; *m*-C*ā*C*i*C (III), *m-sāfir* 'traveling'; *mi-t*-C*a*CC*i*C (V), e.g. *mitfarriz* 'being frozen'; *mi-t-* C*ā*C*i*C (VI), e.g. *mitkatl-īn* 'fighting with each other [masc. pl.]'; *mu-n*-C*a*C*i*C (VII), e.g. *munḥaṛiḳ* 'burning oneself'; *m-i*-C*t*VCVC (VIII), e.g. *miktašif* 'discovering'. (Other participles are *mu*-CC*i*C (IV), *m-i*-$C_1C_2VC_3C_3$ (IX), and *m-ista*-CCVC (X).) The passive participles are the same as the active, except for *a* vocalism. Participles are inflected for number and gender like adjectives.

2.2.6.4 Verbal nouns

Verbal nouns are usually *mu*-CC*ā*C for instrumentals and *mu*-CV̄C*i*C for professions, e.g. *muftāḥ* 'key', *muṛāsil* 'reporter', and *ma*CC*a*C for places, e.g. *maktab* ‘office’. (Adjustments depend on verb Form, e.g. *m*-C*a*CC*i*C for a profession based on Form II, e.g. *mʿallim* ‘teacher’.) The plural is *ma*-CVC*ī*C, e.g. *mafatīḥ* ‘keys’, or external, e.g. *mʿallim-īn* ‘teachers [masc.]’. Others nouns are CVCV̄C, e.g. *kitāb* ‘book’, or use the participle, e.g. *msaxxan* (type of main dish).

2.2.7 Weak verbs

Initial weak verbs with primary *y*, e.g. *ības* ‘to dry up, become hard’, are rare. Imperfect vocalism depends on the verb type. Perfect vocalism is *a*, although *i* and *u* sometimes occur in 1st and 2nd persons. The imperative is formed as for strong verbs.

2.2.7.1 Geminated verbs

In verbs with a two-consonant root, the second consonant is a geminate (Table 10).

Table 10. Conjugation of biliteral verbs (*ḥaṭṭ* ‘to put’)

	imperfect		perfect	
	sg	pl	sg	pl
3 masc.	*b-ḥuṭṭ*	*b-ḥuṭṭ-u*	*ḥaṭṭ*	*ḥaṭṭ-u*
3 fem.	*b-t-ḥuṭṭ*	*b-ḥuṭṭ-in*	*ḥaṭṭ-at*	*ḥaṭṭ-in*
2 masc.	*b-t-ḥuṭṭ*	*b-t-ḥuṭṭ-u*	*ḥaṭṭ-ē-t*	*ḥaṭṭ-ē-tu*
2 fem.	*b-t-ḥuṭṭ-i*	*b-t-ḥuṭṭ-in*	*ḥaṭṭ-ē-ti*	*ḥaṭṭ-ē-tin*
1	*b-a-ḥuṭṭ*	*b-n-ḥuṭṭ*	*ḥaṭṭ-ē-t*	*ḥaṭṭ-ē-na*

Active participle: *ḥāṭiṭ*
Passive participle: *maḥṭūṭ*

2.2.7.2 Initial weak verbs: primary ʾ

A second type of weak verb has primary ʾ (Table 11).

Table 11. Conjugation of initial weak verbs: Primary ʾ (*ōkil* ‘to eat’)

	imperfect		perfect	
	sg	pl	sg	pl
1	*b-ōkil*	*b-n-ōkil*	*akel-it*	*akel-na*
2 masc.	*b-t-ōkil*	*b-t-ōkl-u*	*akel-it*	*akel-tu*
2 fem.	*b-t-ōkl-i*	*b-t-ōkl-in*	*akel-ti*	*akel-tin*
3 masc.	*b-ōkil*	*b-ōkl-u*	*akel*	*akl-u*
3 fem.	*b-t-ōkil*	*b-ōkl-in*	*akl-at*	*akl-in*

Active participle: *mēkil*
Passive participle: *makūl*

2.2.7.3 Initial weak verbs: Verbs I*w*

In I*w* (Table 12), *w* is realized in the imperfect as *u* vocalism.

Table 12. Conjugation of initial weak verbs: Verbs I*w* (*wakif* ‘to stand (up)’)

	imperfect		perfect	
	sg	pl	sg	pl
3 masc.	*b-ūkaf*	*b-ūkaf-u*	*wakif*	*wakaf-u*
3 fem.	*b-t-ūkaf*	*b-ūkaf-in*	*wakf-at*	*wakaf-in*
2 masc.	*b-t-ūkaf*	*b-t-ūkaf-u*	*wakif-it*	*wakaf-tu*
2 fem.	*b-t-ūkaf-i*	*b-t-ūkaf-in*	*wakif-ti*	*wakaf-tin*
1	*b-a-wkaf*	*b-n-ūkaf*	*wakif-it*	*wakaf-na*

Active participle: *wākif*
Passive participle: *mawkūf*

2.2.7.4 Medial weak verbs: Medial *w*, *y*, or *ā*

A ‘hollow’ verb (Table 13, verbs II*w*) with medial *w* has it in the imperfect, while verbs with medial *y* have *ī* in the imperfect, e.g. *ḍāʿ* ‘to lose oneself”; medial *ā* has *ā*, e.g. *nām* ‘to sleep’. The participles of all hollow verbs have medial *y*. In the perfect for 1st and 2nd persons, *ā* is shortened and raised to *u* for verbs II*w*, to *i* for verbs II*y* and verbs with medial *ā*, e.g. *ḍiʿt* ‘I lost myself’, *nimit* ‘I slept’.

Table 13. Conjugation of medial weak verbs (*kām* ‘to get up’)

	imperfect		perfect	
	sg	pl	sg	pl
3 masc.	*b-ikūm*	*b-ikūm-u*	*kām*	*kām-u*
3 fem.	*b-it-kūm*	*b-ikūm-in*	*kām-at*	*kām-in*
2 masc.	*b-it-kūm*	*b-it-kūm-u*	*kum-it*	*kum-tu*
2 fem.	*b-it-kūm-i*	*b-it-kūm-in*	*kum-ti*	*kum-tin*
1	*b-a-kūm*	*b-n-kūm*	*kum-it*	*kum-na*

Active participle: *kayim*

2.2.7.5 Final weak verbs: Tertiary *y* or *ā*

Final weak verbs have tertiary *y* or *ā* (see Table 14).

2.2.7.6 Irregular verbs

Some verbs are double (initial and final) weak, e.g. *aja* ‘to come’ (Table 15).

2.2.8 Quadriliteral verbs

The primary form of verbs with a four-consonant root is Q1 (CVCCVC) (Table 16). Secondary forms are Q2 (*t*-CVCCVC, reflexive of Q1 or denominative) and Q4 (i-$C_1C_2VC_3C_4VC_4$ becoming [state]). Q2 is productive, e.g., *txarbiṭ* ‘to be confused’. The participles for Q1 are *m*C*a*CC*i*C and *m*C*a*C-C*a*C. The active participles for Q2 and Q4 are *m-i*-$C_1C_2VC_3C_4VC_4$, and *m-i*-$C_1C_2VC_3C_4VC_4$, respectively. Quadriliteral verbs are frequently reduplicative, denoting repetitive action or action all over the place, e.g. *nuṭ-niṭṭ* ‘to jump repeatedly, jump all over the place’ (cf. *naṭṭ* ‘to jump’). Borrowed words are sometimes realized as quadriliteral verbs, e.g. *fantaz* ‘to fantasize, imagine’ < English *fantasy*.

Table 14. Conjugation of final weak verbs

i. Tertiary *y* (*nisi* ‘to forget’)

	imperfect		perfect	
	sg	pl	sg	pl
3 masc.	*b-insa*	*b-ins-u*	*nisi*	*nis-u*
3 fem.	*b-t-insa*	*b-ins-in*	*nis-it*	*nis-in*
2 masc.	*b-t-insa*	*b-t-ins-u*	*n.s-ī-t*	*n.s-ī-tu*
2 fem.	*b-t-ins-i*	*b-t-ins-in*	*n.s-ī-ti*	*n.s-ī-tin*
1	*b-a-nsa*	*b-n-insa*	*n.s-ī-t*	*n.s-ī-na*

Active participle: *nāsi*
Passive participle: *mansi*

ii. Tertiary *ā* (*bana* ‘to build’)

	imperfect		perfect	
	sg	pl	sg	pl
3 masc.	*b-ibni*	*b-ibn-u*	*bana*	*ban-u*
3 fem.	*b-t-ibni*	*b-ibn-in*	*ban-at*	*ban-in*
2 masc.	*b-t-ibni*	*b-t-ibn-u*	*ban-ē-t*	*ban-ē-tu*
2 fem.	*b-t-ibn-i*	*b-t-ibn-in*	*ban-ē-ti*	*ban-ē-tin*
1	*b-a-bni*	*b-n-ibni*	*ban-ē-t*	*ban-ē-na*

Active participle: *bāni*
Passive participle: *mabni*

Table 15. Conjugation of double weak *aja* ‘to come’

	imperfect		perfect	
	sg	pl	sg	pl
3 masc.	*b-īji*	*b-īj-u*	*aja*	*aj-u*
3 fem.	*b-t-īji*	*b-īj-in*	*aj-at*	*aj-in*
2 masc.	*b-t-īji*	*b-t-īj-u*	*aj-ī-t*	*aj-ī-tu*
2 fem.	*b-t-īj-i*	*b-t-īj-in*	*aj-ī-ti*	*aj-ī-tin*
1	*b-ā-ji*	*b-n-īji*	*aj-ī-t*	*aj-ī-na*

Active participle: *jāy*

Table 16. Conjugation of quadriliteral verbs (*xaṛbaṭ* ‘ to confuse’)

	imperfect		perfect	
	sg	pl	sg	pl
3 masc.	*b-xaṛbiṭ*	*b-xaṛbiṭ-u*	*xaṛbaṭ*	*xaṛbaṭ-u*
3 fem.	*b-t-xaṛbiṭ*	*b-xaṛbiṭ-in*	*xaṛbaṭ-at*	*xaṛbaṭ-in*
2 masc.	*b-t-xaṛbiṭ*	*b-t-xaṛbiṭ-u*	*xaṛbaṭ-it*	*xaṛbaṭ-tu*
2 fem.	*b-t-xaṛbiṭ-i*	*b-t-xaṛbiṭ-in*	*xaṛbaṭ-ti*	*xaṛbaṭ-tin*
1	*b-a-xaṛbiṭ*	*b-n-xaṛbiṭ*	*xaṛbaṭ-it*	*xaṛbaṭ-na*

Active participle: *mxaṛbiṭ*
Passive participle: *mxaṛbaṭ*

2.3 *Syntax*

2.3.1 Noun phrase

2.3.1.1 Definiteness

The noun phrase is definite if the head noun is definite.

2.3.1.2 Expression of indefiniteness with specificity

If prefaced by the definite article, *waḥad* ‘one, someone’ denotes a specific individual (out of a general group), e.g. *bitaˁb ilwaḥad* ‘one gets tired’.

2.3.1.3 Construct state

Characteristics of the → construct state are (i) the last member determines the definiteness; (ii) *-t* is affixed to a nonfinal feminine noun ending in *a*, e.g. *sāˁa* ‘hour’ in *sāˁa-t m-a-šūfak* ‘the hour when I see you’; (iii) any modifier, including a demonstrative pronoun, occurs only after both members, e.g. *ṭawla-t ilbinit hādi* ‘that table of the girl’, ‘that girl’s table’; and (iv) only the second member can have the definite article attached. The word orders of the construct state mimic those of the verb phrase. The constructions *imm/abu* X and *tabaˁ* X, e.g. *abu ilkitāba* ‘the best writer [masc.]’, *tabaˁ walīd* ‘Walid’s thing’, are similar to the construct state.

2.3.1.4 Quantifiers

Frequently used quantifiers are *kull* ‘all’ and *šwoyy(a)* ‘some, a bit’.

2.3.1.5 Numeral phrase

The number usually precedes the noun, e.g. *tnēn walād* ‘two boys’, except for *waḥad* ‘one’, e.g. *binit waḥad-a* ‘one girl’.

2.3.1.6 Adjectives

The adjective follows the noun or phrase it modifies, e.g., *iṭṭawla likbīra* ‘the big table’.

2.3.1.7 Elative constructions

Elative constructions use the pattern aCCaC, e.g. *m(a)-aḥsan-ak* ‘how good you [masc. sg.] are!’.

2.3.1.8 Relative clauses

Asyndetic relative clauses are rare. If the object is relativized, a redundant object pronoun occurs on the verb, e.g. *hāda zzalama illi šuft-o* ‘that’s the man whom I saw [lit. ‘whom I saw him’].’

2.3.2 Verbal phrase

2.3.2.1 Synthetic tendencies in the verb phrase

The pronominal verbal object can be doubled to increase its definiteness or for emphasis, e.g. *katab-t-hum himma* ‘I wrote them [masc.]’.

2.3.3 Verbal aspect

The structural imperfect expresses simple present, present continuous, present habitual, and future tenses, e.g. *b-iktib* ‘he writes’, ‘he’s writing’, ‘he will write’. The perfect expresses simple past, e.g. *katab* ‘he wrote’. Past continuous tense is formed with auxiliary *kūn* ‘to be’ (with *ibki* ‘to remain’ in rural dialects) in the perfect tense, and a main verb in the subjunctive; it expresses past continuous or past habitual, e.g. *kān y-iktib*, *baka y-iktib* ‘he was writing’, ‘he used to write’ (urban/Bedouin, rural). The continuous tenses can be expressed using the active participle (Section 2.3.3.4), and future tense by using an aspect prefix (Section 2.3.3.3). The subjunctive mood (Table 17) is used for contingent or possible action; for exhortations and wishes, e.g. *n-ṛūḥ* ‘let us go!’; in a complement clause e.g. *b-a-šūf-o y-iktib* ‘I see him write’; and for questions, e.g. *a-kūl* ‘should I say?’.

Table 17. Subjunctive conjugation (*katab* ‘to write’)

	sg	pl
3 masc.	*y-iktib*	*y-ikitb-u*
3 fem.	*t-iktib*	*y-ikitb-in*
2 masc.	*t-iktib*	*t-ikitb-u*
2 fem.	*t-ikitb-i*	*t-ikitb-in*
1	*a-ktib*	*n-iktib*

2.3.3.1 Innovations in the verbal system

Infixal *-iyya-* links indirect objects and direct objects, in that order. If *-l-* ‘for, to’ occurs, it precedes the indirect object, e.g. *ʿaṭ-e-t-il-ha-iyyā-him* ‘I gave them [masc.] for/to her’.

2.3.3.2 Indicative prefixes

The indicative prefix *b-* marks the imperfect.

2.3.3.3 Future-intent prefixes

Future is marked with *ta-* or *ḥa-* affixed to a subjunctive verb, e.g. *ta-y-iktib* ‘when he will write’, *ḥa-y-iktib* ‘he will write’.

2.3.3.4 Use of active participle

The active participle can express present or past continuous: *kātib hu* ‘he’s writing’, *šuf-t-o kātib* ‘I saw him [and he was] writing’.

2.3.3.5 Narrative imperative

The imperative is used instead of imperfect or perfect in Negev Bedouin Arabic for audience engagement or foregrounding of narrative content, e.g. *yōm mil-layām aja u šūf esmaʿ il awāmir* ‘one day he came and saw and heard what was going on’ [lit. ‘see! hear!’] (Henkin 1998).

2.3.3.6 Other means to express durativity, intent, wishes, etc.

Serial verbs express aspect, emphasis, purpose, or sequence of action, e.g. *ṛāḥ yimši* ‘he went and walked’. The verbs *ḍall* ‘to stay’, *kaʿad* ‘to sit’, and *ṭilaʿ* ‘to get up’ are used as auxiliaries to mark continuous action or to emphasize actuality, e.g. *ḍall y-ʿazzib-ni* ‘he kept getting me upset’; *kān kāʿad y-ōkil* ‘he was there just (carrying on) eating’; *ṭilaʿ y-ōkil* ‘he turned out to be eating’.

2.3.3.7 Use of periphrastic narratives

In narrative speech, auxiliary verbs as listed in Section 2.3.3.6 can be used for foregrounding information.

2.3.3.8 Negation

On negation of verbs, pronouns, and prepositions, see Section 2.2.3. The form *m-iš* ‘not’ negates a nonverbal predicate, e.g. *miš hōn* ‘not here’.

2.3.4 Word order

The default word order is Verb-Subject-Object. Several orders are permitted. In Subject-Verb, the subject is topicalized. Subject is frequently non-overt, e.g. *šafat* ‘she saw’ (cf. *šafat hi(yya)* ‘she saw’). Object is sometimes non-overt, e.g. *aḷḷah y-sahhil ʿalēk* ‘may God make it easy for you [masc. sg.]’. Grade adverbs, e.g. *aṣraʿ* ‘faster, fastest’, usually occur directly after the verb.

2.3.5 Agreement/concord

See Section 2.2.4 on adjective-noun agreement and Section 2.3.1.3 on the construct state. For a plural noun Subject in a Subject-Verb-Object sentence, the adjective is masculine. In constructions like *ilwalad illi maksūṛ ilijir* ‘the boy with the broken leg’, the adjective agrees with the noun modified by the relative clause, not with the noun it modifies within the relative clause.

2.3.6 Existential sentences

Existential sentences are equational, adjectival, or locative. The subject is definite unless the existential marker *fī(h)* ‘there is’ is used, e.g. *fī(h) walad iẓġīr* ‘there is a small boy’. An indefinite subject without *fī(h)* can occur if the predicate is negated, e.g. *maʿ-nā-š baṣala* ‘we have no onion [lit. ‘not with us an onion].’ An existential clause is interpreted as present tense; *kān* is inserted for past tense, yielding a verbal sentence, e.g. *kān walad iẓġīr*, *kān fīh walad iẓġīr* ‘there was a small boy’. A dependent existential clause can be interpreted with nonpresent tense (Sec. 2.3.9). The subject can be covert if it is available from the immediately preceding discourse, e.g. *šaṭṛīn* ‘they [masc.] are clever’, lit. ‘clever [masc. pl.]’.

2.3.7 Clause chaining and adverbial clauses

Clause coordination uses a coordinating conjunction. Subordination uses (i) a relative pronoun, (ii) a subordinating conjunction introducing an adverbial clause, e.g. *taʿāl ʿašān a-šūf-ak* ‘come [masc. sg.], so I can see you [masc. sg.]!’, or (iii) one of the complementizers: *(la)ʾinn-* ‘that’, *an* ‘so that, to’, or *ma-* ‘so that, e.g. *hāt an sauwwīhum* ‘let [masc. sg.] me do them [masc.]!’; *zay ma kultlak* ‘like I told you [masc. sg.]’ (Awad 1995). The complementizer *(la) ʾinn-* takes an object pronoun, e.g. *smiʿit (la-) ʾinn-hum iṣḥāb* ‘I heard that they [masc.] were friends’. An adverbial clause usually directly follows the verb.

2.3.8 Conditional sentences

The antecedent clause usually precedes the consequent in a conditional sentence. Conditional particles include *iza* (urban), *iḏa* (rural) for realis; and *law* ‘if, if only, were (that)’ and *kān* ‘if only’/had . . .’ for irrealis, with *kān* occurring only in a past counterfactual, e.g. *iza b-t-īji b-it-šuf-ī-ha* ‘if she comes, you [fem. sg.] will see her’; *law b-t-īji b-it-šuf-ī-ha* ‘were she to come, you [fem. sg.] would see her’; *law ij-it kān šuf-t-ī-ha* ‘had she come, you [fem. sg.] would have seen her’.

2.3.9 *Ḥāl* sentences

With the conjunction *u* ‘as, when’, an existential subordinate clause is interpreted with the tense of the main clause, e.g. *b-ā-ji u inta ġāyib* ‘I’m coming when you [masc. sg.] are absent’; *aj-ī-t u inta ġāyib* ‘I came when you [masc. sg.] were absent’. Asyndetic *ḥāl* sentences do not occur.

2.4 *Semantics*

The *imm/abu* X and *tabaʿ* X constructions (Sec. 2.3.1.3) are used with alienable possession. The same can be true of *ʿend* ‘with, *chez*, etc.’. Bare nominals have an existential interpretation. The complementizer *(la-)ʾinn-* induces uncertainty about the matrix proposition.

3. Lexicon

Many borrowings exist, e.g. *ḥrām* ‘blanket’ from Classical Arabic, *sabōn* ‘soap’ from French, ***ḅaḅa*** ‘dad’ from Italian, *nerviz* ‘nervous [masc. sg.]’, ‘to be nervous’ from English, *belifon* ‘cellular phone’ from Modern Israeli Hebrew. Metaphor and metonymy are common, e.g. *ṭawwil ṛōḥ-ak* ‘be patient [masc. sg.]!’ [lit. ‘lengthen your [masc. sg.] spirit!’, *mastūṛa* ‘married woman’, lit. ‘covered [fem. sg.]’ (Negev Bedouin Arabic) Onomatopoeic words include *ṭuxx* ‘shoot’, ʿ*aww* ‘bark [as a dog barks]’.

Bibliographical references

Awad, Maher. 1995. “Complementizers and modality”. *Papers from the Regional Meetings, Chicago Linguistic Society* 31.16–27.

Bauer, Leonhard. 1913. *Das palästinensische Arabisch: Die Dialekte des Städters und des Fellachen*. 3rd. ed. Leipzig: J.C. Hinrichs’sche Buchhandlung.

Blanc, Haim. 1953. *Studies in North Palestinian Arabic: Linguistic inquiries among the Druzes of western Galilee and Mt. Carmel*. Jerusalem: The Oriental Society.

Blau, Joshua. 1960. *Syntax des palästinensischen Bauerndialektes von Bir-Zêt*. Waldorf-Hessen: Verlag für Orientkunde.

Broselow, Ellen, Su-I Chen, and Marie Huffman. 1997. “Syllable weight: Convergence of phonology and phonetics”. *Phonology* 14.47–82.

Davis, Stuart. 1995. “Emphasis spread in Arabic and Grounded Phonology”. *Linguistic Inquiry* 26.465–498.

Elihai, Yohanan. 1985. *Dictionnaire de l’arabe parlé palestinien: Français-arabe*. Paris: Klincksieck.

Henkin, Roni. 1998. "Narrative styles of Palestinian Bedouin adults and children". *Pragmatics* 8.47–78.

Naïm, Samia. 1999. "Depalatalization and construction in the urban dialects of Palestinian Arabic". *La linguistique* 35.141–162.

Palva, Heikki. 1966. *Lower Galilean Arabic.* Helsinki: Societas Orientalis Fennica.

Seeger, Ulrich. 1998. "Arabische Dialekttexte aus il-Xalīl [Hebron]". *Mediterranean Language Review* 10.89–145.

Kimary N. Shahin
(University of British Columbia)

Parataxis → Sentence Coordination

Paronomasia

Commonly known as 'pun' (Redfern 1984), paronomasia is a figure of speech which is based on a play on meanings, either by using, in the same context, one homonymous or polysemic word, or by using two or more words with synonymous or closely related meanings, which are often alliterative or rhyming. Examples of this include situations where an important personality who has passed away is cynically described as *a grave man* (homonym), or where the expression *kith and kin* (alliterative idiom) is used.

Paronomasia is widely used in all genres of literature, as well as in daily communication, and is usually employed for a special, often humorous effect. Related English terms in stylistics include 'ambiguity', 'antanaclasis', 'catachresis', 'syllepsis', 'zeugma', and even 'jingles', 'quibbles', and 'polyptoton', while in other European languages we find, in addition, 'annomination', 'double entendre', 'calembour', and a few more terms, which come under the general notion of 'jeux sémantiques' in French, or 'Wortspiel' and 'doppelsinnig' in German.

Arab scholars of rhetoric (*ʿilm al-balāġa*), who postulate that rhetoric consists of three main branches, *ʿilm al-maʿānī* 'semantics', *ʿilm al-bayān* roughly 'stylistics', and *ʿilm al-badīʿ* 'the study of figures of speech', discuss paronomasia at length, and use different equivalent terms for it. The term *tawriya* (from the root *w-r-y* 'to hide, conceal') is used for cases where a word has two meanings (is 'ambiguous'), both of which may apply in a certain text or context, while only one, usually the less expected or 'far' meaning, is really intended, e.g. *huwa muḥammadun* 'he is Muhammad'/'he is praiseworthy', where *muḥammad* may mean a proper name or an attribute formed as a passive participle.

The term → *jinās* (or its equivalents *tajnīs*, *ʾīhām*, *ištirāk fī l-lafḏ̣*, or *ištirāk fī l-maʿnā*) refers to the use of two polysemic or homonymous words in one phrase or sentence, e.g. *xafaqat al-qulūb lammā ʾaxfaqa l-inqilāb* 'the hearts beat when the revolution failed'. Here, the play on the words *qalb* 'heart' and *inqilāb* 'revolution', both derived from the same root *q-l-b* 'to turn' and both making sense in combination with the root *x-f-q* 'to palpitate; to fail', is intentional. In addition, the syntactic phenomenon known as *mafʿūl muṭlaq* 'absolute object', e.g. *waṣafahu waṣfan daqīqan* 'he described him in detail', is regarded as *jinās* (→ object, absolute).

Bonebakker (1966), who has investigated the *tawriya* in classical Arabic poetry, has researched most aspects of it, as discussed by "Arab critics or literary theorists". According to his analysis, Arab scholars identify four (and possibly five) types of *tawriya*:

i. *Tawriya mujarrada* 'simple pun', when the *tawriya* word has no *lawāzim* 'attributes', and both meanings are possible.
ii. *Tawriya muraššaḥa* 'prepared *tawriya*', when the *tawriya* word is used in a context that contains 'attributes' of the meaning that is not intended by the writer. This is called *al-maʿnā al-muwarrā bihi* 'the hiding sense' or *al-maʿnā al-qarīb* 'the nearer sense'.
iii. *Tawriya mubayyana* 'explained *tawriya*', when the *tawriya* word is used in a context that contains 'attributes' of the intended meaning, known as *al-maʿnā al-muwarrā ʿanhu* 'the hidden sense' or *al-maʿnā al-baʿīd* 'the further sense'.
iv. *Tawriya muhayyaʾa* 'supported *tawriya*', when the context contains a word which clarifies the *tawriya* itself.
v. *Tawriya muqtarina* 'double *tawriya*', when two *tawriyas* may yield four possible meanings.

Three other terms are used in the study of *tawriya*:

i. *Istixdām*, when the double meaning of a word used as *tawriya* serves in two different morphological or syntactic constructions.
ii. *ʾIbhām* (*tawjīh*), when two opposite meanings are possible, often using contradictory senses, a positive and a negative, which may leave the context vague or ambiguous, or when no clear indication is given as to whether the literal or the metaphorical meaning is meant.
iii. *Tawhīm*, when the *tawriya* may be a mistake, though allowing a possible interpretation (Bonebakker 1966:9–23).

According to a modern work (ʾIsbīr and Junaydī 1981), there are several types of *tawriya* or *jinās*. *Tawriya* includes the four types mentioned above, whereas *jinās* includes *al-murakkab, al-muṭlaq, al-muštaqq, al-mulaffaq, al-muḏayyal, al-lāḥiq, al-muṭarraf, al-muṣaḥḥaf, al-muḥarraf, al-lafḏī, al-maqlūb,* and *al-maʿnawī.* In general, there are two categories. The first category is that of *al-jinās at-tāmm*, in which the two words are in complete harmony, i.e. they are homonyms or polysemes. The types found are *al-mumāṯil, al-mustawfī, jinās at-tarkīb, al-marfūw, al-mutašābih, al-mafrūq,* and *al-muḥarraf.* The second category is that of *al-jinās an-nāqiṣ*, in which there is some difference between the two words. To this category belong *al-muṭarraf, al-muḏayyal, al-lāḥiq*, *al-muḍāriʿ*, *jinās al-qalb* (*al-mujannaḥ* or *al-muzdawaj*), *al-muṭlaq, jinās al-ištiqāq, al-mulaffaq, al-muṣaḥḥaf al-lafḏī, al-ʿaksī,* and *al-maʿnawī.* The conditions are that either the two words have a common derivation (*ištiqāq wāḥid*), or the two words resemble each other morphologically (*mutašābihāni ištiqāqan*; ʾIsbīr and Junaydī 1981:369–370, 409–410).

The various classifications provided by Arab rhetoricians are, no doubt, too detailed, and according to Bonebakker, they are arbitrary and often vague. In his view, they are used inconsistently and leave room for speculations and different interpretations of what the writer or speaker has really meant.

ʿAbd al-Qāhir al-Jurjānī (d. 471/1078), who begins his book *ʾAsrār al-balāġa* with a long chapter on paronomasia, advocates that a good style is judged by 'the beauty of ideas' not 'the beauty of words'. Hence, using too complex paronomasia, not easily grasped by the intellect, may miss the point, and "it is like overdressing a bride with too many ornaments, that may make her look repulsive" (*ʾAsrār* 9). Al-Jurjānī goes on to quote Arab rhetoricians, such as al-Jāḥiẓ, and a number of famous poets to illustrate good and bad use of paronomasia.

Obviously, the richness of the Arabic language with its wealth of synonyms, antonyms, homonyms, and polysemes puts at the disposal of writers an inexhaustible supply of puns, which have from time immemorial been current both in classical and modern poetry and prose, orally and in writing (cf. Somekh 1992:78–79), usually reflecting the author's eloquence and aesthetic.

The following are but a few examples chosen randomly. The *Qurʾān* contains many instances of *tawriya* and *tajnīs*, e.g. *bismi llahi r-raḥmāni r-raḥīm* 'in the Name of Allah the Most Gracious, the Most Merciful' [the paronomasia involves the last two words, which are derived from the same root and are close in meaning]; *wa-s-samāʾu banaynāhā bi-ʾaydin* (Q. 51/47) 'We have constructed Heaven with power/by hands' [where the last word may be interpreted as 'with strength' from the noun *ʾayd* or as the plural of the word *yad* 'hand']; *wa-r-raḥmānu ʿalā l-ʿarši stawā* (Q. 20/5) 'and the Most Gracious has established Himself'. In the last example, according to some commentators, the 'far' meaning of *istawā* is *istawlā* 'to take control', and this is meant, rather than the 'close' meaning of 'to establish oneself, to get seated' (see, e.g., Bustānī 1987, s.v. *tawriya*).

Arabic poetry has innumerable examples of paronomasia, e.g. *wa-ʾiḏa mašat tarakat bi-ṣadrika ḍiʿfa mā bi-ḥuliyyihā min kaṯrat al-waswās* (ʾAbū Tammām) 'and when she walks, she leaves in your heart twice as many of her jewels, the manifold of her seductive stomping'. Here, the wordplay is on the word *al-waswās*, which is an onomatopoeia, meaning 'speaking under one's breath, whispering', but should also be associated with the Qurʾānic epithet for the 'seductive Devil' (Q. 114/4). Other examples include *ʾaqarra llāh ʿaynahu wa-kafāhu šarrahā wa-ʾajrā lahu ʿaḏbahā wa-ʾakṯara ladayhi tibrahā* (Ibn Hišām) 'may God please him, and may He protect him against its evil, and may He make its sweet water flow for him, and may He increase his gold [i.e. wealth]'; the poet is using as paronomasia four different meanings of the word *ʿayn*: the idiomatic use of ‘eye’ in the sense of ‘delight, pleasure’; ‘evil eye’; ‘spring’; ‘wealth’; and *min baḥri šiʿrika ʾaġtarif wa-bi-faḍli ʿilmika ʾaʿtarif* (ʾAbū Nuwās) ‘from the ocean of your verse I will draw and your abundance of knowledge I will acknowledge’; paronomasia is achieved here by the playing on the rhyming verbs *ʾaġtarif* and *ʾaʿtarif*. Even popular songs in Arabic employ paronomasia, e.g. *baḥibbak wa-ʾudārī u-ʾinta muš dārī* ‘I love you and conceal my [feelings] and you are not aware of it’, where the wordplay is based again on the morphological similarity between Forms I and III of the same root *d-r-y*.

Arabic prose offers many examples of paronomasia, e.g. *ma tarā raʾya mā narā* ‘you don’t see our point of view’; *ʾāna l-ʾawān li-jamʿ šamlinā* ‘the time for a reunion has come’; *kān wa-mā kān* ‘once upon a time’, a traditional beginning of children’s stories.

Many Arabic → proverbs and dicta are based on paronomasia, e.g. *kuff fakkayk wa-fukk kaffayk* lit. ‘stop your jaws and untie your hands’, i.e. ‘don’t talk, do!’; *ḥāmīha ḥarāmīhā* ‘her guards are those who stole her/it’; *al-ʾaqārib ʿaqārib* lit. ‘relatives are scorpions’, i.e. ‘those who are close to you are those who harm you’.

A large number of → collocations contain paronomasia, e.g. *junna junūnuhu* ‘to become frantic’; *layla laylāʾ* ‘dark night’; *ḥiṣn ḥaṣīn* ‘stronghold’.

Many → idioms are paronomastic, e.g. *lam yatruk bāban ʾillā wa-ṭaraqahu* lit. ‘he left no door without knocking on it’, i.e. ‘he left no stone unturned’; *šayʾan fa-šayʾan* lit. ‘a matter following a matter’, i.e. ‘gradually’; *dassa s-samm fī d-dasam* lit. ‘to put poison in the fat’, i.e. ‘to poison the good atmosphere’.

Many daily expressions (→ greetings, vocatives, interjections, etc.) contain paronomasia, e.g. *ʾahlan wa-sahlan* lit. ‘you are part of the family and an easy guest’, i.e. ‘you are welcome!’; *ʿāmirīn wa-ṯāmirīn* ‘[may God grant you] a long and fruitful union!’ as a congratulation for a newlywed couple; *hanīʾan marīʾan* ‘bon appétit’.

Titles of books, especially medieval works, were formed as paronomasia, e.g. *Fākihat al-xulafāʾ wa-mufākahat aḏ-ḏurafāʾ*) by Ibn ʿArabšāh (d. 1450); *Fawāt al-wafayāt* by Ibn Šākir al-Kutubī (d. 1362); *Xarīdat al-qaṣr wa-jarīdat al-ʿaṣr* by Muḥammad al-ʾIṣfahānī (d. 909).

Headlines often contain paronomasia, probably in order to attract the attention of the reader, e.g. *al-lājiʾūna ṣ-ṣumāliyyūn yahrubūna min nār bilādihim ʾilā jaḥīm al-yaman* ‘the Somali refugees escape from the inferno of their country to the hell of Yemen’; *ʾadabunā ʾilā l-ʿibriyya bawwāba li-maʿrifat al-bawāṭin al-ʿarabiyya* ‘our literature [translated] into Hebrew is the gate for the knowledge of the Arab hidden secrets’; *ʾiflās xazāʾin ḥamas* ‘Hamas’ insolvent treasures’.

Advertisements may also use paronomasia to attract the attention of the customer, e.g. *istaṯmir fī miṣr al-ḥabība balad al-ʾamn wa-l-ʾamān* ‘invest in beloved Egypt, the country of safety and security!’ [a government advertisement]; *taqniya jadīda maʿahā sa-taqūlu dāʾiman: kull šayʾ taḥta s-sayṭara* ‘a new technology with which you will always say: everything is under control’ [an advertisement by IBM for a new computer, alluding also to the control key on the keyboard]; *raḥāba fāʾiqa qiyāda salisa quwwat heimi* ‘exquisite space, smooth driving, Heimi power’ [an advertisement for a new car where the words *raḥāba* and *qiyāda* have a double meaning; the former indicates spaciousness but also alludes to generosity, while the latter means ‘driving’ but also ‘leadership’ and ‘control’].

Finally, a large number of humorous pieces (‘fun verses’, comic sketches, jokes, graffiti, riddles, tongue twisters, cartoons, mainly political, etc.) amuse because of wordplay i.e. paronomasia, e.g. *ṭaraqtu l-bāba ḥattā kalla matnī wa-lammā kalla matnī kallamatnī* ‘I knocked on the door until my strength was sapped and when it was sapped she spoke to me’ [a well-known ‘fun verse’]; *al-ʿilm nūr wa-l-jahl ʾanwar* ‘knowledge is light and ignorance is more illuminating’ [a political joke heard in Egypt, offering a wordplay on the first name of the late president Anwar Sadat]; *ḥubb lā tuḥārib* ‘love, don’t fight’ [a ‘mitigated’ version of the slogan ‘make love not war’, seen in graffiti; the Arabic word *ḥubb* simply means ‘love’ and not ‘make love’].

Bibliographical references

Primary sources

Ibn Qutayba, *ʾAdab* = ʾAbū Muḥammad ʿAbdallāh ibn Muslim Ibn Qutayba, *ʾAdab al-kātib*. Ed. Max Grünert. Leiden: E.J. Brill, 1967.

Jurjānī, *ʾAsrār* = ʾAbū Bakr ʿAbd al-Qāhir ibn ʿAbd ar-Raḥmān al-Jurjānī, *ʾAsrār al-balāġa*. Ed. Helmut Ritter. Istanbul: Government Press, 1954.

Secondary sources

Bonebakker, Seeger A. 1966. *Some early definitions of the* tawriya *and Ṣafadī's* Faḍḍ al xitām ʿan at-tawriya wa-ʾl-istixdām. The Hague and Paris: Mouton.

Bustānī, Buṭrus al-. 1987. *Muḥīṭ al-muḥīṭ*. Beirut: Maktaba Lubnān.

Douglas, Allen and Fedwa Malti-Douglas. 1994. *Arab comic strips: Politics of an emerging mass culture*. Bloomington and Indianapolis: Indiana University Press.

Garcin de Tassy, Joseph Héliodore Sagesse Vertu. 1848. *Rhétorique et prosodie des langues de l'Orient musulman*. Paris.

Hāšimī, ʾAḥmad al-. 1962. *Jawāhir al-ʾadab*. Cairo: al-Maktaba at-Tijāriyya al-Kubrā.

ʾIsbīr, Muḥammad Saʿīd and Bilāl Junaydī. 1981. *aš-Šāmil: Muʿjam fī ʿulūm al-luġa al-ʿarabiyya wa-muṣṭalaḥātihā*. Beirut: Dār al-ʿAwda.

Mehren, August Ferdinand Michael. 1853. *Die Rhetorik der Araber*. Copenhagen: O. Schwartz and Vienna: k.k. Hof- und Staatsdruckerei.

Monteil, Vincent. 1960. *L'arabe moderne*. Paris: C. Klincksieck.

Reckendorf, Hermann. 1909. *Über Paronomasie in den semitischen Sprachen*. Giessen: Töpelmann.

——. 1977. *Arabische Syntax*. Repr. Heidelberg: C. Winter. (1st ed. 1921.)

Redfern, Walter. 1984. *Puns*. Oxford: Blackwell.

Somekh, Sasson. 1992. "The neo-classical Arabic poets". *The Cambridge history of Arabic literature: Modern Arabic literature*, ed. Muhammad M. Badawi, 36–81. Cambridge: Cambridge University Press.

Xalīl, ʾAḥmad. 1968. *al-Madxal ʾilā dirāsāt al-balāġa al-ʿarabiyya*. Beirut: Dār an-Nahḍa al-ʿArabiyya.

Avihai Shivtiel (Leeds)

Participle

Both Classical Arabic and modern spoken varieties of Arabic have what are customarily termed active and passive participles. Although nearly identical in form in both varieties, the semantico-syntactic status of the active participle differs. This entry summarizes the commonalities, particularly the morphological, then deals separately with the two varieties.

1. Common structures

1.1 *Morphology*

Morphologically, both active and passive participles are regularly derived from a verb. The active/passive participles have the form *fāʿil/mafʿūl* in the basic form, and and in the derived forms they essentially have *m*V- + imperfect/perfect stem. Because both participles are inflected like adjectives, in Classical Arabic they take case endings. Furthermore, with a rare dialectal exception (see (22) and (23) below), like adjectives, they are not inflected for person. In Classical Arabic the masculine plural usually takes sound plural suffixes *-ūna/-īna/-āt*, although the active participle in particular may be associated with a broken plural (e.g., see (1)). Table 1 summarizes the basic set of forms, using spoken Arabic inflectional forms, further variants being noted below.

Table 1. Basic set of forms of the participle

	active		passive	
	masc.	fem.	masc.	fem.
sg.	*kātib*	*kātib-a(t)*	*maktūb*	*maktūb-a(t)*
pl.	*kātib-īn*	*kātib-āt*	*maktūb-īn*	*maktūb-āt*

Derived forms attach the prefix *mi-*, Classical Arabic *mu-*, to the respective imperfect (for the active participle) and perfect (for the passive participle) stems, e.g. *mi-ḥaddid/mu-ḥaddid* 'limiting', *mi-ḥaddad/mu-ḥaddad* 'limited'.

For the active participle, verbs I*w/y*, II*w/y*, III*w/y*, and geminated verbs all display irregularities of certain kinds. In spoken Arabic, geminated verbs have either the form *fāʿil*, e.g. *rādid* 'returning' (Levantine), or *fāʿl*, e.g. *rādd* (Cairene, North African, all Sudanic varieties, Arabian peninsular dialects); hollow verbs have *fāyil*, e.g. *šāyil* 'having picked up'; and weak verbs have *fāʿi*, e.g. *rāmi* 'having thrown, throwing'. Verbs Iʾ have either *m-* or *w-*, *mākil* or *wākil* 'having eaten'. In Classical Arabic, geminated verbs have *fāʿl*, e.g. *rādd*; verbs II*w/y* have *fāʾil*, e.g. *bāʾiʿ* 'buying', *qāʾil* 'saying'; and verbs III*w/y* have *fāʿī*, e.g. *rāmī* 'having thrown, throwing' (indefinite *rāmi-n* by phonological rule), *dāʿī* 'calling'. In some derived Forms (VI, VII, VIII), verbs II*w/y* and geminated verbs neutralize the difference between active and passive participle, in agreement with the perfect/imperfect verbs of the same classes, *muštaqq* 'deriving from/derived from' (cf. *ištaqqa/yaštaqqu*), *mubtāʿ* 'buying, bought' (*ibtāʿa/yabtāʿu*). In Forms V and VI, the active participle has the form *mutafaʿʿil/mutafāʿil* (cf. imperfect *yatafaʿʿal, yatafāʿal*).

Irregularities in the passive participle in Classical Arabic are generally explicable by morphophonological rule, the *-w-* of the participle infix being assimilated to a neighboring *y*, for instance; *mabīʿ* < **mabyuwʿ* ‘bought’, *mabnī* ‘built’ < **mabnuwy*. However, in Classical Arabic in particular, there are for both the active and the passive participle irregularities of various types, of which three can be mentioned:

i. Verbs in Form IV may take their passive participle from the basic stem, *ʾaḥabba* ‘to love, like’ but *maḥbūb* ‘loved’; *ʾajanna* ‘to make crazy’, *majnūn* ‘crazy’.
ii. A number of intensive (*mubāliġa*) nominal patterns are considered to behave like active participles and are treated as such, including the patterns *faʿʿāl, faʿūl, faʿīl*, and *mifʿāl*, e.g. *natūj* ‘bearing many [camels]’ < *ʾantaja, qadīr* ‘very able, potent’, *samīʿ* ‘listener, hearing, the All-hearing [of God]’.
iii. Participles can have irregular broken plurals, as in (1) and (2).

(1) *hunna ḥawājj-un bayta llāh*
‘They [fem.] are performing the pilgrimage to the Kaʿba’ (sg. *ḥājj*/*ḥājja*)

(2) *quṭṭān makkat-a*
‘those who dwell in Mecca’ (Sībawayhi, *Kitāb* I, 46; Zamaxšarī, *Mufaṣṣal* 229)

1.2 *Syntax*

In the Arabic linguistic tradition, the two participial forms are classified as subtypes of noun, the active participle being the noun of the agent (→ *ism al-fāʿil*) and the passive the noun of the patient (*ism al-mafʿūl*). This classification gives priority to their inflectional properties. Although active and passive participles lack an agent marker (see Ibn al-ʾAnbārī, *ʾInṣāf* I, 55–65, on the status of an inferred agent pronoun in participles), early linguists recognized that they replicate many properties of a verb. Like a verb, they allow fronting of a direct object and maintain the argument structure of the verb, for instance an active participle governing an object in the accusative, as in

(3) *zaydun ḍārib-un ġulām-a-hu*
‘Zayd is beating his boy’

Participles therefore stood between verbs and nouns (*ʾasad* ‘lion, lion-like’) in a cline from most verblike to most nounlike, a detailed classification of structures dependent on a complicated conjunction of lexical, syntactic, and semantic factors (Sībawayhi, *Kitāb* I, 194ff.).

In Classical Arabic, the precise rules of object governance become quite complex, with the case form of governed objects, definiteness, and number all being interacting factors (see Ibn as-Sarrāj, *ʾUṣūl* I, 122ff.; Wright 1977:II, 63ff.). In Classical Arabic, the passive participle as well has the governance capability of a verb, for instance governing a nominative subject, as in

(4) *zaydun maḍrūb-un ġulām-u-hu*
‘As for Zayd, his boy is beaten’ (Zamaxšarī, *Mufaṣṣal* 229)

In modern spoken varieties, the basic syntactic behavior of the passive participle mirrors that of Classical Arabic. However, the passive participle generally does not carry over the actants from a corresponding verb. In Nigerian Arabic, for instance, from an active verb *šarrab al-walad alme* ‘he made the boy drink water’, one cannot have **al-walad mišarrib alme* (where *mišarrib* is active or passive participle) ‘the boy was made to drink water’. Moreover, the passive participle itself may be lexically restricted. In Cairene Arabic, for instance, passive participles of derived verbs are comparatively rare.

In general, the major difference between Classical Arabic and spoken Arabic resides in tense/aspect, so at this point the varieties will bifurcate. The main point of difference is that in Classical Arabic the active participle is represented within the logic of basic tense categories. A standard explanation, for instance, is found in az-Zamaxšarī (*Mufaṣṣal* 228), that when an active participle governs an accusative, as in (3) above, the meaning is present or future (*ḥāl* or *mustaqbal*), and when a genitive, past (*māḍī*), as in (5) and (6).

(5) *zaydun ḍārib-u ġulām-i-hi*
‘Zayd has beaten his boy’

(6) *fa qālat ʾa-ʾanti bnatu qātil-i sayyid-i-hi*
‘And she said, are you the daughter of the one who killed his master?’ (Youssef 1990:205)

According to az-Zajjājī (*ʾĪḍāḥ* 86–87), the Kufan grammarians regarded the participle as the present tense, called *fiʿl dāʾim* 'permanent verb', along with the past and the future (Owens 1988:136–138). There are hints (e.g. Sībawayhi, *Kitāb* I, 194, summarizing Yūnus and ʿĪsā) that lexical aspect played a systematic role in Classical Arabic as well, although it was not a perspective elaborated upon in the Arabic linguistic tradition. Nonetheless, as is seen in the next section, aspect colors the meaning of the active participle in Classical Arabic.

2. THE ACTIVE PARTICIPLE IN STANDARD/CLASSICAL ARABIC: SYNTAX

An active participle must be contextualized to represent a tense, i.e., it has no inherent tense. One example of contextualization is the governance of the object (see (3) vs. (5) above). In the literature, the active participle is sometimes represented as a stylistic variant of either the imperfect or the perfect (Reuschel and Blöhm 1973:125ff.), as in (7) and (8).

(7) *naḥnu naṭlubu musāwāt tāmma*
'We demand complete equality'

(8) *wa māḏā ʾanta ṭālib-un ʾilayya*
'And what do you demand of me?'

Note that in this respect, the status of the active participle in Classical Arabic contrasts markedly with its role in spoken Arabic, where it nearly always contrasts in meaning with either an imperfect or perfect verb (see Sec. 3).

Lexical aspectual nuances are also relevant in Classical Arabic, however. In both punctual verbs, such as *jalasa* 'to sit down', *waqafa* 'to stand up', *labisa* 'to wear, dress', and *ittafaqa* 'to agree', and durative verbs, describing an enduring action, e.g. *mašā* 'to go, walk', *tazāyada* 'to increase', and *qāma* 'to stand', the active participle implies an action embarked upon, whose effects are still apparent.

(9) *huwa jālis*
'He is sitting' (has sat and is still sitting)

(10) *huwa muqīm-un bi-l-kūfa baʿdu lam yabraḥ*
'He is still in Kufa and hasn't left yet' (Youssef 1990:209)

With such verbs, the perfect tense may sometimes be paraphrased with an active participle (Youssef 1990:204ff.), as in (11) and (12).

(11) *allaḏī šariba al-xamra = šārib-u l-xamri*
'The one who drank the wine = the drinker (having drunk) of the wine'

(12) *qad faqada ʿaqlahu = fāqid-un ʿaqlahu* (Youssef 1990:204ff.)
'He has lost his senses = the one who has lost/having lost his senses'.

Active participles of verbs of motion either describe an action (durative) at the time of speaking, or they have a future meaning, often implying an impending action. Here, the time may be specified through an adverb.

(13) *ʾinnī musāfir-un baʿda yawmayn*
'I'll travel in two days' (Reuschel and Blöhm 1973:135)

(14) *ʾinnī ḏāhib-un … ibqa ʾanta ʾiḏā šiʾt*
'I'm going now; stay only if you want to' (Reuschel and Blöhm 1973:135)

(15) *ʾinnā mursil-ū n-nāqati fitnatan lahum*
'We have sent a camel to them as a temptation' (Q. 54/27; Youssef 1990:211)

With such verbs, the active participle, as opposed to the verb, is frequently intensified via particles (*ʾinna*, *ʾanna*), or via *mā/lā zāla*:

(16) *fa-lā yazāl-ū la-nā hāʾib-īna li-qatli-nā min-hum ʾamīr-a-hum*
'And they still fear us, because we killed their leader' (Youssef 1990:202)

(17) *ʾinnī rājiʿ-un al-ʾān ʾilā manzilī*
'I'm now on my way to my house'

The active participle has a fixed function in the so-called → *ḥāl* sentence, fixed as Sentence + *wa* + subject + active participle, in which the *ḥāl* describes an action or state coterminous with the preceding main clause.

(18) *fa māt-at wa-hiya rājiʿ-at-un ʾilā makka*
'And she died as she was returning to Mecca' (Youssef 1990:215)

Both active and passive participle in Classical Arabic often lexicalize into an adjective (often from intransitive verbs) or noun.

(19) *labisat ṯiyāban fāxirat-an*
'She wore a beautiful dress'

(20) *wa-lākinnahā ġayr lāʾiq-atin bika*
'But it doesn't suit you'

(21) *huwa fāris-un*
'He is a horseman'

3. THE PARTICIPLE IN SPOKEN ARABIC: SYNTAX, MORPHOLOGY

In spoken Arabic, the passive participle largely has the status of an adjective and therefore will not be treated further here (see (4)). The active participle, on the other hand, has a central role in the → aspect system of Arabic (see Sec. 3.1), and in this function it takes the pronoun object inflections of a verb, thus *-ni*, not *-i* in the 1st person singular. Over and above the basic features described in Section 1, the addition of a suffix may induce change in the active participle. Arabic dialects fall into three classes in this respect. The forms can be illustrated with the feminine singular of the active participle, because in this form the relevant changes are most obvious.

i. No change. An object suffix is simply added to the active participle + gender/number suffix: *kātib-a* + *ha* = *kātiba-ha* (= *katb-ā-ha* after phonological adjustment) 'she has written it [fem.]' (Cairene Arabic).
ii. Feminine *-it*. The feminine singular takes the construct form *-it*: *kātb-it-ha*; there are no other changes (Eastern Libyan Arabic).
iii. Intrusive *-in*. An intrusive *-in* is added before the suffix, *kātb-it-inno* 'she has written it', *kātb-in-no* 'he has written it' (Omani Arabic; Reinhardt 1972:139).

The first alternative is found, inter alia, in most of Egypt, the Sudan and Western Sudanic Arabic, and in most of Syria. The second one is found in much of the Arabian Peninsula, Algeria, Eastern Libyan Arabic, and in the Sinai. The third one is rare, found in the extreme eastern Arabian Pensinula, beginning in Bahrain and moving east and south through the Emirates into Oman, in one dialect of Western Sudanic Arabic, and in the isolated dialects of Khurasan and Uzbekistan.

There are many variations on these forms. In Omani Arabic, Shukriyya in eastern Sudan, and Nigerian Arabic, for instance, the intrusive *-in/an* is added to the feminine plural suffix *-āt*, *kātbāt-ann-a*. 'they [fem.] have written it [masc.]' (Reichmuth 1983; Owens 1993). The distinction between masculine/feminine plural occurs only in those dialects where the plural gender distinction is maintained. In Nigerian Arabic, the phonological contrast between active and participle is lost in derived forms, *mifaʿʿil* being used for both. In Uzbekistan, the active participle has been completely refunctionalized so that the first pronominal suffix on the active participle marks a subject. This turns the participle into a form morphologically marked for subject, since in the 3rd person the typical gender and number distinctions characteristic of the active participle (see Sec. 1) remain intact (Retsö 1988; Zimmermann 2002), as in (22) and (23):

(22) *zorb-in-nā-hum* (< *zōrib* < **zārib* < **ḍārib*)
beat-in-1p-3mp
'We have beaten them'

(23) *zorb-in-nā*
beaten-3mp-1p
'They have beaten us' (*-inn*-V < *-īn*-V by general rule)

3.1 *Function and syntax*

The use of the active participle is difficult to describe because it depends on structural and, more importantly, lexical factors (→ *Aktionsart*, i.e. lexical aspect), as Mitchell (1952) points out. As a verbal element, the active participle represents an event conceived of as a completed entity and having internal duration. It is aspectual, not temporal, and as in Classical Arabic, in and of themselves active participles do not specify a time or tense. Rather, the time is orientated according to the moment of speaking, or by the time implicit in a given textual context.

The active participle carries the same diathetic properties as does the verb it is derived from (→ diathesis), a ditransitive verb, for instance, remaining ditransitive in the active participle.

(24) *ana* *mʿalm-it-ha* *s-sawga*
I taught-fs-her Def-driving
'I [fem.] have taught her driving'
(Algerian Arabic)
cf. Perfect: *ʿallimitha s-sawga*
'I taught her driving'

Effectively, the active participle thus introduces a third temporal form, in addition to the perfect and imperfect, into the spoken Arabic verb system, although, as noted, the values are lexically sensitive.

Three basic categories of usage of the active participle can be distinguished, depending largely on the nature of the verb stem. The following examples are taken from Nigerian Arabic, Gulf Arabic, Egyptian Arabic, and Algerian Arabic; they are with a very few exceptions (e.g. Maltese; Mifsud 1995:70) applicable to nearly all varieties of modern spoken Arabic, as the geographical spread of the examples indicates.

i. Motion verbs. The active participle describes either a motion in progress, or one which has been decided upon.

(25) *jāy-īn waḥrān* 'they were coming to Oran' (Algerian Arabic)

(26) *ʾul-t ana rāyiḥ iskandariyya* 'I said I intended to go to Alexandria' (Egyptian Arabic)

ii. Action and stative verbs. The active participle describes a state which has been entered into. For an action verb, the difference between the active participle *yālsīn* and the imperfect verb *yīlisūn* (Gulf Arabic) is that between 'they are seated' and 'they are sitting down [now]', between an event whose effect is still visible and an incompleted event. Among statives, the active participle describes a state, e.g. *nāyim* 'asleep', whereas the imperfect describes a future or habitual event, *b-inām* 'he will sleep, he sleeps' (Egyptian Arabic). Looking at the perfect, the difference between the active participle –

(27) *ḫāṭ karāsi baṛṛa* (Gulf Arabic; *ṭṭ* > *ṭ*)

and the perfect –

(28) *ḫaṭṭ karāsi baṛṛa* (Gulf Arabic)

is that between 'he has put chairs out' (and you can see them there, perfective) and 'he put chairs out' (and they may or may not still be there, punctual).

iii. The third category, comparable to, e.g., (18)–(21) above in Classical Arabic, consists of various words having the form of an active participle, but which are lexicalized as adjectives, adverbs, or nouns.

Adjective
(29) *id-dōra ij-jāya*
'the coming week' (Nigerian Arabic)

Adverb
(30) *xārij al-bēt*
'outside of the house' (Nigerian Arabic)

(31) *taʾāl jāy*
'Come in this direction!' (lit. 'come coming'; Nigerian Arabic)

(32) *lamma yibʾa aswad xāliṣ*
'when it becomes completely black' (Egyptian Arabic)

Noun
(33) *is-sāyig həní*
'The driver is here' (Gulf Arabic)

3.2 *Further remarks*

Beyond these three basic categories, dialect-specific usages may further refine the aspectual range of the active participle, again according to lexical factors. In Cairene Arabic, for instance, *lissa* 'still' combines with what Eisele (1992) terms 'noninchoative verbs' to give a reading of either 'just' or 'still'.

(34) *hu lissa ʾāʿid fi kursi*
'He is still seated in a chair/he has just sat in a chair'

Noninchoatives, on the other hand, have only the 'just' reading' with *lissa*.

(35) *hu lissa kāwi il milābis*
'He has just ironed the clothes'

The active participle may also be realized in the form *faʿlān* (pl. *faʿʿāl*, or sound plural). In Nigerian Arabic, certain perception verbs take only this form.

(36) *hu fahmān-inn-a*
'He has understood it' (with intrusive participial *-in*; not **fāhim-a*)

(37) *n-rūḥ nəḥna ṣəyyām*
'We go fasting' (Gulf Arabic; instead of *ṣāym-īn*)

In Cairene Arabic, a two-way contrast may be found between *fāʿil* and *faʿlān* among statives, e.g. *bārid* 'cold [usually, weather]', *bardān* 'cold [feeling cold]'. In Algerian Arabic, on the other hand, *faʿlān* is not used at all as an active participle alternative to *fāʿil*.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Ibn al-ʾAnbārī, *ʾInṣāf* = ʾAbū l-Barakāt ʿAbd ar-Raḥmān ibn Muḥammad Ibn al-Anbārī, *Kitāb al-ʾinṣāf fī masāʾil al-xilāf bayna n-naḥwiyyīna l-baṣriyyīna wa-l-kūfiyyīn*. Ed. Muḥammad Muḥyī d-Dīn ʿAbd al-Ḥamīd. Repr., Beirut, n.d.

Ibn as-Sarrāj, *ʾUṣūl* = ʾAbū Bakr ibn as-Sarī Ibn as-Sarrāj, *al-ʾUṣūl fī n-naḥw*. Ed. ʿAbd al-Ḥusayn al-Fatlī. 3 vols. Beirut: Muʾassasat ar-Risāla, 1985.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān, *al-Kitāb*. Ed. Hartwig Derenbourg. Repr., Hildesheim: G. Olms, 1970.

Zajjājī, *ʾĪḍāḥ* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān ibn ʾIsḥāq az-Zajjājī, *al-ʾĪḍāḥ fī ʿilal an-naḥw*. Ed. Māzin al-Mubārak. Cairo: Dār al-ʿUrūba, 1959.

Zamaxšarī, *Mufaṣṣal* = ʾAbū l-Qāsim Maḥmūd ibn ʿUmar az-Zamaxšarī, *al-Mufaṣṣal fī ʿilm al-ʿarabiyya*. Ed. Muḥammad al-Ḥalabī. Beirut: Dār al-Jīl, n.d.

SECONDARY SOURCES

Blöhm, Dieter. 1981. "Zur Funktion des Partizips Aktiv im modernen Hocharabisch". *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 34.143–51.

Eisele, John. 1992. "Aspectual classification of verbs in Cairene Arabic". *Perspectives in Arabic linguistics*, II, ed. Mushira Eid, 192–233. Amsterdam and Philadelphia: J. Benjamins.

Holes, Clive. 1990. *Gulf Arabic*. London: Routledge.

Mifsud, Manwel. 1995. *Loan verbs in Maltese: A descriptive and comparative study*. Leiden: E.J. Brill.

Mitchell, T.F. 1952. "The active participle in the Arabic dialect of Cyrenaica". *Bulletin of the School of Oriental and African Studies* 14.11–33.

Owens, Jonathan. 1988. *The foundations of grammar: An introduction to medieval Arabic grammatical theory*. Amsterdam and Philadelphia: J. Benjamins.

——. 1993. *A reference grammar of Nigerian Arabic*. Wiesbaden: O. Harrassowitz.

Reichmuth, Stefan. 1983. *Der arabische Dialekt der Shukriyya im Ostsudan*. Hildesheim: G. Olms.

Reinhardt, Carl. 1972. *Ein arabischer Dialekt gesprochen in Oman und Zanzibar*. Amsterdam: Philo Press. (1st ed., Stuttgart, 1894.)

Retsö, Jan. 1988. "Pronominal suffixes with *-n(n)-* in Arabic dialects and other Semitic languages". *Zeitschrift für Arabische Linguistik* 18.77–94.

Reuschel, Wolfgang. 1996. *Aspekt und Tempus in der Sprache des Korans*. Frankfurt: P. Lang.

——, and Dieter Blöhm. 1973. "Das arabische Aktiv-Partizip als Prädikat". *Asien, Afrika, Lateinamerika: Beiträge zur Afrikanistik und Orientalistik*, Sonderheft 1.111–149.

Wright, William. 1977. *A grammar of the Arabic language*. 2 vols. Cambridge: Cambridge University Press. (Repr. of ed. 1896–1898.)

Youssef, Zafer. 1990. *Das Partizip im Arabischen*. Ph.D. diss., University Erlangen-Nürnberg.

Zimmermann, Gerit. 2002. *Das Arabische von Buchara zwischen alten Quellen und neuen Forschungsergebnissen*. M.A. thesis (Magisterarbeit), Bayreuth University.

JONATHAN OWENS
(University of Bayreuth, Germany,
University of Maryland),
MARAT YAVRUMYAN (University of Bayreuth)

Parts of Speech

The division of speech into its parts, or of words into their categories, is the first preoccupation of every significant grammatical text in Arabic, starting with the opening lines of the *Kitāb* of Sībawayhi (d. ca. 180/796): *hāḏā bābu ʿilmi mā l-kalimu min al-ʿarabiyya. fa-l-kalim: ismun, wa-fiʿlun, wa-ḥarfun jāʾa li-maʿnan laysa bi-smin wa-lā fiʿlin* (I, 1 ed. Derenbourg/I, 2 ed. Būlāq). This can be translated fairly literally as 'Chapter on the knowledge of what words are in Arabic. Those words are: name (→ *ism*), action (→ *fiʿl*), and a bit (→ *ḥarf*) which comes for some meaning and is neither a name nor an action'. The rest of this article, like Arabic grammar itself, is essentially a commentary on that statement.

It is at once striking that there are only three parts of speech, which will henceforth be referred to in their appropriate technical style as 'noun', 'verb', and 'particle', with the caveat that the literal meaning is never absent from discussion of these elements as linguistic entities. The question of where this threefold classification came from cannot be answered by direct evidence: a possible inspiration, it has been claimed (J. Fischer 1962–1963, 1963–1964; but see Carter 1972), is Aristotle's broad division in the *Poetics* into *ónoma* and *rhēma* and a miscellaneous group of items comprising *syllabē* 'syllable', *árthron* lit. 'joint', and *súndesmos* 'conjunction', all characterized as *phōnē ásēmos* 'a nonmeaningful sound'. Of these, however, only *ónoma* 'name' is a natural fit with *ism* 'name/noun', while *rhēma* '(predicate-) word' is very far from *fiʿl* 'action/verb', and *phōnē ásēmos* 'nonmeaningful sound' seems to be the opposite of 'bit which comes for some meaning'. (Merx [1889:142ff.] shows that this was a problem for the Greek commentators as well, one of whom, Ammonius, came to an interpretation similar to Sībawayhi's that these words are used to change the modalities and other aspects of sentences.) The post-Sībawayhian term *rābiṭ(a)* 'linking element' may look like a calque of the Greek *árthron* or *sundesmós*, but it refers to an entirely different category, a kind of anaphoric copula pronoun (see Carter 1997:40ff.). In the 4th/10th century, a literal rendering of *rhēma* as → *kalima* 'word' (Latin *verbum*!) appeared, but only in the works of logicians like al-Fārābī (see Zimmermann 1981, index s.v. *kalima*), and the grammarians rejected it as incompatible with their established terminology. Sībawayhi himself had no first-hand acquaintance with Greek thought, and he was either unaware of (or chose to ignore) the fact that Greek grammar divided speech into eight parts, information which he could easily have absorbed from his Syriac contemporaries (cf. Talmon 1991 for the parts of speech in Ibn al-Muqaffaʿ). Such was the influence of the *Kitāb* that the Greek system was never adopted, even though Sībawayhi's successors did import the classification of utterances into five types: 'imperative, request, entreaty, vocative, statement' (after Zimmermann 1981:43), which cannot be explored further here (see Versteegh 1977:145ff., 2004).

Sībawayhi hardly defined the parts of speech at all but simply enumerated them with a handful of examples. Nouns are illustrated by *rajul* 'man', *faras* 'horse', and, in a later manuscript tradition, *ḥāʾiṭ* 'wall'. Verbs are described as *ʾamṯilatun ʾuxiḏat min lafḏ ʾaḥdāṯi l-ʾasmāʾ*, freely translated 'patterns taken from the expression of nouns denoting events', a puzzlingly laconic formulation, which occasioned much disagreement. They are then subdivided by their form and time/aspect reference into an asymmetric scheme of three types, (i) 'built for what has passed' (*buniyat li-mā maḍā*), e.g. *ḏahaba* 'he went', *samiʿa* 'he heard', *makuṯa* 'he remained', *ḥumida* 'he was praised'; (ii) 'for what will be and has not happened' (*mā yakūnu wa-lam yaqaʿ*), e.g. the imperatives *iḏhab* 'go!', *uqtul* 'kill!', *iḍrib* 'strike!' and the predicatives *yaqtulu* 'he will kill', *yaḏhabu* 'he will go', *yuqtalu* 'he will be killed', *yuḍrabu* 'he will be struck' (we are left to infer that there is no passive imperative); (iii) 'for what is existing uninterrupted' (*mā huwa kāʾinun lam yanqaṭiʿ*), illustrated by the same predicative verbs as in the second category. Curiously, what we call the imperfect verb (→ *māḍī/muḍāri*) occurs twice in the list, once bundled with the imperative in a future meaning and once in a present meaning. Furthermore, even though there is an explicit future prefix, e.g. *sa-yaḏhabu* 'he will go', it is not invoked anywhere in this classification, nor in the definition of type (ii), where *yakūnu* 'will be' is used in its default future meaning. The particle is illustrated, without further comment, by *ṯumma* 'then', *sawfa* (another future marker!), *wa-* in swearing an oath (e.g. *wa-llāhi* 'by God!'), and the preposition *li-* 'of' etc. (the list varies slightly among manuscripts).

These three categories so tersely set out by Sībawayhi are perfectly adequate for his own analysis of Arabic, but the later grammarians soon increased the number and length of the definitions, which became topics of controversy between both individual grammarians and schools of grammar, as well as between the different disciplines which were by this time starting to assert their autonomy. What follows is condensed from a large body of material, which has been comprehensively surveyed by Versteegh (1995), and individual attributions are mostly dispensed with here.

The noun came to be defined with more precision, drawing heavily on Aristotle: it is a simple entity (*mufrad*, i.e., its constituent letters have no independent meaning), unrelated to time, denoting either a physical individual *šaxṣ* lit. 'person', i.e. concrete objects both animate and inanimate ('man', 'Bakr', 'stone') or nonphysical *ġayr šaxṣ*, i.e. abstract ('striking', 'eating', 'day', 'hour'). The second group of examples is intended to defuse the objection that some nouns are related to time or actions: the verbal noun (*maṣdar*) does indeed denote an action, but it has no time reference, while nouns which lexically refer to time do not denote an action.

Nouns are further defined by their morphology, although this is usually left implicit, in contrast with the comprehensive treatment of the distinctive morphology of verbs in the larger grammars. A set of markers unique to nouns was drawn up (although, like all such criteria, they were challenged, and exceptions were easily found), viz. the affixing of the morphemes of definiteness (*al-*) and indefiniteness (→ *tanwīn*, but see below), the oblique case ending in possessive and prepositional structures, and some negative collocations (neither *qad* 'already' nor the future marker *sawfa* may precede nouns, although there are counterexamples for *qad*), and, finally, only nouns may be pronominalized (excluding the interrogative nouns).

In addition, nouns are defined by their syntactic function: they alone may be both subjects (this includes agents) and predicates (e.g. Zajjājī, *ʾĪḍāḥ* 48). It did not go unnoticed that this borders on the logical definition of nouns as the subject of propositions, and some grammarians refused to accept it for that reason. From it arose the pedagogical simplification that a noun is any word which may be the agent of the verbs *ḍarranī* 'harmed me' and *nafaʿanī* 'benefited me' (attributed to al-ʾAxfaš al-ʾAwsaṭ [d. 215/830] by Zajjājī, *ʾĪḍāḥ* 49). A more sophisticated variant is the definition of nouns as those elements which function as what we might now call arguments of the verb, i.e. as agents and direct objects, or the equivalent, such as periphrastic agents and other kinds of complements. These comprise a number of subclasses of the noun which are identified more by their function than their form, such as the agent noun (→ *ism al-fāʿil*), patient noun, (*ism al-mafʿūl*), adjective (→ *ṣifa*), circumstantial qualifier (→ *ḥāl*; e.g. *rākiban* 'riding'), space/time qualifier (*ḏarf*, e.g. *laylan* 'by night', *maʿan* 'together'; → *mafʿūl fīhi*), and specifying element (→ *tamyīz*, e.g. *tafaqqaʾa šaḥman* 'he was bursting with fat').

Verbs, too, underwent a more rigorous process of definition in the centuries after Sībawayhi's death. They are now distinguished from nouns by their being semantically compound (*murakkab*), i.e. denoting an action and an agent (cf. ʾAstarābāḏī, *Šarḥ* I, 3ff.). The latter, being integral with the verb, can never be elided, unlike subjects and predicates, which reminds us that there is no infinitive or agentless verbal form in Arabic. Verbs are also contrasted with their agent nouns in that they cannot be made definite, dualized or pluralized, in other words the dual imperative *iḍribā* can only mean 'beat, you two!' and not 'beat twice!'. Sībawayhi's notion that verbs are denominative, i.e. derived from nouns denoting the action (called the → *maṣdar* or 'source' of the verb, the nearest to an infinitive in Arabic), led to a debate which polarized the grammarians into those who accepted his 'Baṣran' position and the 'Kūfans', who maintained that the *maṣdar* was in fact derived from the same radical consonants as the verb (see the discussion in Ibn al-ʾAnbārī, *ʾInṣāf* 102–107).

In parallel to the logical definition of the noun as subject, verbs were defined as inherently predicative. This is certainly a borrowing from Greek, but the predicative function of verbs was already axiomatic for Sībawayhi, which he may have arrived at inductively or by some intellectual osmosis which we cannot document. It is taken for granted in the first chapter of the *Kitāb*, where verbs are said to be used either in 'giving orders' (*ʾāmiran*) or 'giving information' (*muxbiran*), i.e. as predicates, and it is an important component in his explanation of the inflectional behavior of certain word patterns which, when they occur as adjectival qualifiers, do not take full nominal inflection (see below).

The time element is the most problematical feature of verbs, possibly because the Arabic verb does not fit into the three-tense system of the classical European languages, and it is clear that Sībawayhi had no specific terminology for past, present, and future tenses (→ *māḍī/muḍāriʿ*). Later grammarians soon filled this lacuna, though not with total consistency: the past is uniformly called *māḍī* 'passed, elapsed', but the present is either *ḥāḍir* '[physically] present' or *ḥāl* 'the [current] situation', and even *muqīm* 'abiding' in one context, and the future is mostly *mustaqbal* 'future' lit. 'what you are faced with', but sometimes *muntaḏar* 'awaited' and *mutawaqqaʿ* 'expected'. Perhaps because there are only two paradigms to spread over the three times, some grammarians were sympathetic to the sophism that the present does not really exist, while others, notably the Kūfans, created a virtual present continuous tense out of the active participle, which they labeled the *fiʿl dāʾim* 'ongoing, lasting action' (→ *ism al-fāʿil*). The borrowed philosophical term *zamān muḥaṣṣal* for the time reference of verbs (Versteegh 1995:58) has given rise to different interpretations, although it seems clear that it means simply a 'finite time', i.e. one to which limits can be set (cf. the synonymous alternative *zamān muʿayyan* lit. 'an appointed time'), thus escaping from the philosophical paradox of time being an infinite sequence of dimensionless points (a debate which is still going on; → *māḍī/muḍāriʿ*).

The particle is, by its negative definition, an amorphous category, and its name *ḥarf* lit. 'bit on the edge', expresses this. All the grammarians, Sībawayhi included, were perfectly aware that the same term was also applied to letters of the alphabet and their individual sounds, graphemes and phonemes respectively in our vocabulary, as well as to the etymological radical and augmentative consonants in the derivational system. These are not 'parts of speech' in the syntactic sense of 'morpheme' and are not dealt with here, but it is vital to give due weight to the key descriptive phrase *jāʾa li-maʿnan* 'comes for a meaning', which distinguishes the syntactic *ḥarf* and has been part of its definition from the very beginning. It is not enough to translate it as 'meaningful': the Arab grammarians themselves pointed out that even a letter of the alphabet has a meaning of sorts. For Sībawayhi, → *maʿnā* 'meaning' was primarily grammatical, i.e. it denoted a linguistic function, and he speaks of *maʿnā n-naṣb* 'the meaning of the dependent case', *maʿnā l-istifhām* 'the meaning of asking a question', *maʿnā n-nidāʾ* 'the meaning of calling', and many others. Nearly all these 'meanings' are effected through a particle, so when we read of a *ḥarf jāʾa li-maʿnan*, the statement is incomplete until we add the grammatical function mediated by that *ḥarf*. Thus, *ʾillā* 'except' is a *ḥarf istiṯnāʾ* 'a particle [which comes for the meaning] of making an exception'; the *wa-* in *wa-llāhi* 'by God' is a *ḥarf qasam* 'a particle [which comes for the meaning] of swearing an oath'; *lā* 'not' is a *ḥarf nafy* 'a particle [which comes for the meaning] of negating'; and so on for all the grammatical 'meanings' of which the particles are the exponent. This was the only way to identify the particle, and it is worth noting that an alternative name for it in this instrumental function was *ʾadā* 'tool', used once only by Sībawayhi (*Kitāb* II, 146 ed. Derenbourg/II, 143 ed. Būlāq) in tandem with *ḥarf*, but more frequently by others, and allegedly typical of the Kūfan technical vocabulary (see Kinberg 1996:8–10 for *ʾadā* in al-Farrāʾ).

Because it lacks a demonstrable external denotation, signifying rather a relationship between other elements or participants in the discourse, the *ḥarf* generated its own problems, both theoretical and pedagogical. It could not be a subject or predicate in natural language (but see below), so it was also devoid of the typical collocational features of nouns and verbs. Some grammarians seized on Sībawayhi's amplification of his original definition, where he added that the *ḥarf* occurs 'only' for a meaning (here contrasting the noninflectional endings of invariable words and the true case morphemes). Ibn Fāris (d. 395/1004) seems to have overinterpreted the negative aspects (*aṣ-Ṣāḥibī* 86): where Sībawayhi states that the *ḥarf* has a meaning but is not a noun or a verb, Ibn Fāris infers that it has a meaning which is not in a noun or a verb, and he cites as an example *zaydun munṭaliqun* 'Zayd is going away', which becomes a question when preceded by the interrogative particle *hal*, viz. *hal zaydun munṭaliqun* 'is Zayd going away?', now containing a meaning which is neither in *zaydun* nor *munṭaliqun*. This hardly seems to have been Sībawayhi's intention.

Needless to say, none of the definitions outlined above were accepted without question. In the period when the Islamic sciences were being formalized, there were two complementary pressures, the urge to bring the linguistic sciences into conformity with the principles laid down by Aristotle, and the desire for prestige and power within each science, which was most easily achieved by taking a position slightly different from one's predecessors and colleagues: in the words of a 4th/10th-century catchphrase, *xālif tuḏkar* 'disagree and be famous'. Over some two and a half centuries, from the 3rd/9th to the mid 5th/11th, there was an effervescence of public disputation, very lively and experimental (much of it recorded as controversies between the Baṣrans and the Kūfans), but by the 6th/12th century the debates, while still allowing dissent and interpretative latitude, were conducted in a purely scholastic framework, i.e. the Arabic data were now all agreed (and Arabic was technically no longer a living language), the analytical methods were fixed, and grammar was fully in the service of Islamic linguistic conservatism, its main purpose often expressly declared to be the correct understanding of the language of the *Qurʾān* and the sayings of the Prophet, from which both theology and law were derived by scholars, whose competence in this now artificial language was the prime source of their professional authority.

This does not mean the debates were sterile: indeed, they foreshadow familiar issues of modern linguistics, and the outcome was a reinforcement of methodology. There was, for example, disagreement on whether the 'parts of speech' should be considered as a division of the whole (*kalām* 'speech') into its parts (where the name of the whole does not apply to its separate parts, so we cannot say, for example, 'all nouns are speech'; cf. Širbīnī, Carter 1981:12–14), or of the universal (*kalima* 'word') into its particulars (where the name of the category applies to all its members, which can be tested by inversion: all nouns are words but not all words are nouns). This distinction has practical consequences for the nature of linguistic generalizations, not unlike that now recognized between *langue* and *parole*, and also influenced the arrangement of pedagogical grammars. A prominent symptom of the scholastic approach is the creation of hierarchies: nouns, verbs, and particles were ranked in that order because a valid utterance can be constructed of nouns alone, while verbs need nouns, and particles need both, a view already ascribed to al-Mubarrad (Bernards 1997:177–179). Another hierarchy gave priority to nouns because they are substances, then to verbs, which are accidents, with particles last because they are neither.

Even the number of three parts was contested. Needless to say, the Kūfans did not accept the pragmatic assumption of their Baṣran opponents that only three parts and no more were found after an exhaustive study of the language, and they put forward, among other words, *kallā* 'by no means' as a counterexample, which, they claimed, was neither noun, verb, nor particle (Carter 1981:13). The Kūfan challenge has deep implications: if they were right, then linguistic authority would be retained by the speaker in an open system, hence unpredictable and therefore uncontrollable, while the Baṣrans, well aware of this, made sure to close the corpus so that they could derive their authority deductively from it, exactly as their partners in the legal sciences were doing at this time. In the end, a degree of slippage was admitted, by conceding that not every element could be adequately defined, nor could every definition be watertight, especially when even the original Bedouin informants had been unsure, e.g. the syntactic behavior of the exceptive particles *ḥāšā, ʿadā, xalā* 'except', which fluctuates between that of particles and verbs. The so-called verb of admiration (*fiʿl at-taʿajjub*) is another good example: it certainly has verbal form and syntax (e.g. *mā ʾaḥsana zaydan* 'how handsome Zayd is!', with *zaydan* marked as a direct object), but not everyone agreed that it was a verb, especially as a diminutive (a nominal feature) is reported, e.g. *mā ʾumayliḥa zaydan* 'how rather sweet Zayd is!'.

It was recognized from the first, i.e. by Sībawayhi, that object language and metalanguage are different. For the parts of speech, this led to the rather flippant (though not at all irrelevant) observation that all words are nouns since they are things we say, an idea attributed to al-Mubarrad (d. 285/898) by az-Zajjājī (*ʾĪḍāḥ* 44). In metalanguage, this is precisely the case, as in, to quote a later example, *min ḥarfu jarrin* '[the word] *min* is a particle of obliqueness', where *min* 'from', being the subject, functions as a noun. At a higher level of abstraction, there was controversy about the nature of speech itself: if considered solely as an act of the speaker, the challenge was to justify analyzing this activity into 'noun', 'verb', and 'particle', in other words, where the boundary lies between the linguistic and the extralinguistic. For Sībawayhi, this boundary was deliberately obscured, as he treated language entirely as an activity of the speaker, hence *fiʿl* always meant for him both the physical 'action' and the linguistic entity 'verb'. He went further: not only the words but also the speaker and the circumstances are part of the grammatical structure, such that, if the situation is already clear, an utterance need not be articulated in full, hence an isolated noun in the dependent case could literally be the direct object of the situation and not of an elided verb. There are other issues here which cannot be pursued: speech is an attribute of God, and grammar has to reconcile the divine and secular linguistic planes of the chosen language, Arabic. There was particular sensitivity to the heretical speculation that human speech acts might be the result of free will and not predestination.

The threefold classification also had general consequences for the theory of inflection. After the major cleavage between variable/inflected (nouns, verbs) and invariable/uninflected words (particles), it remained to account for the overlapping inflections of some of the variable words. Nouns are morphophonologically 'lighter' (*ʾaxaff*) than verbs, i.e., they have a limited number of forms compared with the large range of the verbal paradigm, and for this reason they have a set of three case inflections (→ *ʾiʿrāb*) and, most important, a unique suffix *-n* (→ *tanwīn*) as the marker of their full nominal status (*tanwīn* also marks indefiniteness, but that is a secondary function). Verbs are by contrast 'heavier' (*ʾaṯqal*) than nouns, and large parts of the paradigm are therefore invariable (the whole of the *māḍī*), and the *tanwīn* suffix is totally absent. On this basis, verbs which have a functional resemblance to nouns (i.e. the *muḍāriʿ* form) also partake in nominal inflection (as a kind of mood marker), while conversely nouns and adjectives which have the same pattern as verbs (e.g. *ʾakbaru* 'greater/greatest', cf. *ʾaḏhabu* 'I go') may forfeit some of their nominal markers, notably the *tanwīn*, and one of the case endings under certain conditions, especially when used adjectivally, because here they assume the qualifying function which is the property of verbs (Carter 2004:115–119).

While the arguments about the parts of speech may appear to be pedantic, they are the foundation of a complex grammatical theory which not only had to satisfy its practitioners but also fit into the scheme of the Islamic sciences in maintaining the doctrines and law of Islam. Treatises of legal theory usually deal with the parts of speech very early in the work, and the first systematic semantics in Arabic arose out of legal hermeneutics. It goes without saying that Qurʾānic exegesis would be almost impossible without a stable grammar. Perhaps only a mystic could fail to see the importance of the three parts of speech, or rather, would see through them, like the *ṣūfī* who, when told by a grammarian that all speech is nouns, verbs, and particles, tore his robe and cried out in disappointment at the twenty years he had wasted looking for something more than that. What the grammarian did not know is that the *ṣūfī* was only teasing him, for he had heard the speech of God directly, but the less enlightened remain imprisoned in the *ism*, *fiʿl*, and *ḥarf* of the grammarians (van Ess 1991–1997:IV, 618).

Bibliographical references

Primary sources

ʾAstarābāḏī, *Šarḥ* = Raḍī d-Dīn Muḥammad ibn Ḥasan al-ʾAstarābāḏī, *Šarḥ Kāfiyat Ibn Hišām*. Ed. Istanbul: Maṭbaʿat al-ʾAmīra al-Maḥmiyya, 1275 A.H.

Ibn ʾĀjurrūm, *Muqaddima* = ʾAbū ʿAbdallāh Muḥammad ibn Muḥammad as-Sanhājī Ibn ʾĀjurrūm, *al-Muqaddima al-ʾĀjurrūmiyya*. Many editions. [See Carter (1981:465), for translations of this and other grammatical texts.]

Ibn al-ʾAnbārī, *ʾInṣāf* = ʾAbū l-Barakāt ʿAbd ar-Raḥmān ibn Muḥammad al-ʾAnbārī, *Kitāb al-ʾinṣāf fī masāʾil al-xilāf bayna n-naḥwiyyīna l-baṣriyyīn wa-l-kūfiyyīn*. Ed. Gotthold Weil. Leiden: E.J. Brill, 1913.

Ibn Fāris, *Ṣāḥibī* = ʾAbū l-Ḥusayn ʾAḥmad Ibn Fāris, *aṣ-Ṣāḥibī fī fiqh al-luġa*. Ed. Moustafa el-Chouémi. Beirut: A. Badran, 1964.

Ibn Hišām, *Qaṭr* = Jamāl ad-Dīn ʾAbū Muḥammad ʿAbdallāh ibn Yūsuf Ibn Hišām, *Qaṭr an-nadā wa-ball al-ġadā*. Many editions. (French transl., A. Goguyer, *La pluie de rosée, étanchement de soif, traité de flexion et syntaxe*, Leiden: E.J. Brill, 1887.)

Ibn Mālik, *ʾAlfiyya* = Jamāl ad-Dīn ʾAbū ʿAbdallāh Muḥammad ibn ʿAbdallāh Ibn Mālik, *al-Xulāṣa al-ʾAlfiyya*. Many editions. (French transl., L. Pinto, Constantine, 1887; A. Goguyer, Paris, 1888.)

Ibn as-Sarrāj, *ʾUṣūl* = ʾAbū Bakr ibn as-Sarī Ibn as-Sarrāj, *al-ʾUṣūl fī n-naḥw*. Ed. ʿAbd al-Ḥusayn al-Fatlī. Beirut: Muʾassasat ar-Risāla, 1985.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. Ed. Hartwig Derenbourg. Paris: Imprimerie Nationale, 1881–1889. (Repr., Hildesheim and New York: G. Olms, 1970.) Ed. Būlāq, 1898–1900. (Repr., Baghdad: Muṯannā [1965].) Ed. ʿAbd as-Salām Muḥammad Hārūn. Cairo: Dār al-Qalam, etc. 1968–1977. (2nd ed. 1977; 3rd ed. 1983.) [The first and last seven chapters of the *Kitāb* are published, with links to various editions, commentaries, translations, and manuscripts, by Michael G. Carter, A. Matveev, and Lutz Edzard at www.arts.usyd.edu.au/research_projects/sibawiki/homepage/, where there are also links to other Web editions.]

Širbīnī, *Nūr* = Muḥammad al-Xaṭīb aš-Širbīnī, *Nūr as-sajiyya fī ḥall ʾalfāḏ al-ʾĀjurrūmiyya*. Text, translation, and commentary in Carter (1981).

Zajjājī, *ʾĪḍāḥ* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān ibn ʾIsḥāq az-Zajjāji, *al-ʾĪḍāḥ fī ʿilal an-naḥw*. Ed. Māzin al-Mubārak. Cairo: Dār al-ʿUrūba, 1959.

——, *Jumal* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān ibn ʾIsḥāq az-Zajjāji, *al-Jumal fī n-naḥw*. Ed. M. Ben Cheneb. Paris: C. Klincksieck, 1957.

Zamaxšarī, *Mufaṣṣal* = ʾAbū l-Qasim Maḥmūd ibn ʿUmar az-Zamaxšarī, *al-Mufaṣṣal*. Ed. Jens Peter Broch. Christiania: Libraria P.T. Mallingii, 1879. (Rev. ed., several repr.)

Secondary sources

Anghelescu, Nadia. 2004. *La langue arabe dans une perspective typologique*. Bucharest: University of Bucharest.

Bernards, Monique. 1997. *Changing traditions: Al-Mubarrad's refutation of Sībawayhi and the subsequent reception of the* Kitāb. Leiden: E.J. Brill.

Bohas, Georges, Jean-Patrick Guillaume, and Djamel E. Kouloughli. 1990. *The Arabic linguistic tradition*. London and New York: Routledge.

Bohas, Georges, and Michael G. Carter. 2004. "Prolégomènes au *Kitāb* de Sībawayhi: Traduction". Bohas and Kouloughli (2004:43–59).

Bohas, Georges, and Djamel Eddine Kouloughli (eds.). 2004. "À propos du préambule au *Kitāb* de Sībawayhi". *Langues et Littératures du Monde Arabe* 5.7–79.

Carter, Michael G. 1972. "Les origines de la grammaire arabe". *Revue des Etudes Islamiques* 40.69–79.

——. (ed.). 1981. *Arab linguistics: An introductory classical text with translation and notes*. Amsterdam: J. Benjamins.

——. 1997. "Copula in Arabic grammar". *Proceedings of the 17th Conference of the Union of European Arabists and Islamicists*, ed. Wilferd Madelung, Yu. Petrosian, Hilary Waardenburg-Kilpatrick, A. Khalidov, and E. Rezvan, 37–45. St. Petersburg.

——. 2004. *Sībawayhi*. London and New York: I.B. Tauris.

Chairet, M. 1999. "*Ḫiffa*, *ṯiqal* et *tamakkun*: Régime d'incidence et classes des mots". *Langues et Littératures du Monde Arabe* 1.213–226.

Ess, Josef van. 1991–1997. *Theologie und Gesellschaft im 2. und 3. Jahrhundert Hidschra: Eine Geschichte des religiösen Denkens im frühen Islam*. Berlin and New York: de Gruyter.

Fischer, J.B. 1962–1963; 1963–1964. "The origin of tripartite division of speech in Semitic grammar". *Jewish Quarterly Review* N.S. 53.1–21; 54.132–160.

Fischer, Wolfdietrich. 1989. "Zur Herkunft des grammatischen Terminus *ḥarf*". *Jerusalem Studies in Arabic and Islam* 12.135–45.

Frank, Richard M. 1981. "Meanings are spoken of in many ways: The earlier Arab grammarians". *Le Muséon* 94.259–319. (Repr., Richard M. Frank, *Philosophy, theology and mysticism in medieval Islam: Texts and studies in the development and history of* kalām, I, ed. Dimitri Gutas, no. XII, Aldershot and Burlington: Ashgate, 2005.)

Howell, Mortimer Sloper. 1886–1911. *A grammar of the Classical Arabic language translated and compiled from the works of the most approved native or naturalized authorities*. Allahabad: Northwestern Provinces and Oudh Govt. Press.

Jahn, Gustav. 1895–1900. *Sībawaihi's Buch über die Grammatik, übersetzt und erklärt*. Berlin: Reuther und Reichard. (Repr., Hildesheim: G. Olms, 1965.)

Kinberg, Naphtali. 1996. *A lexicon of al-Farrāʾ's terminology in his* Qurʾān *commentary with full definitions, English summaries and extensive citations*. Leiden: E.J. Brill.

Kouloughli, Djamel Eddine. 1986–87. "Les particules ont-elles un sens? Autour d'une controverse dans la tradition grammaticale arabe". *Bulletin de Linguistique Générale et Appliquée, Université de Besançon* 13.235–237.

Owens, Jonathan. 1990. *Early Arabic grammatical theory: Heterogeneity and standardization*. Amsterdam and Philadelphia: J. Benjamins.

Merx, Adalbert. 1889. *Historia artis grammaticae apud Syros*. Leipzig: Deutsche Morgenländische Gesellschaft. (Repr., Nendeln: Kraus, 1966.)

Suleiman, Yasir. 1990. "Sībawayhi's 'parts of speech' according to al-Zajjājī: A new interpretation". *Journal of Semitic Studies* 35.245–63.

Talmon, Rafael. 1991. "Naḏ̣ra jadīda fī qaḍiyya ʾaqsām al-kalām: Dirāsa ḥawla kitāb Ibn al-Muqaffaʿ fī l-manṭiq". *al-Karmil* 12.43–67.

Troupeau, Gérard. 1981. "Les 'partes orationis' dans le *Kitāb al-Uṣūl* d'Ibn al-Sarrāj". *Historiographia Linguistica* 8.379–88.

Versteegh, Kees [C.H.M.]. 1977. *Greek elements in Arabic linguistic thinking*. Leiden: E.J. Brill.

——. 1993. *Arabic grammar and Qurʾānic exegesis in early Islam*. Leiden: E.J. Brill.

——. 1995. *The explanation of linguistic causes: Az-Zağğāğī's theory of grammar*. Amsterdam and Philadelphia: J. Benjamins.

——. 1997. *The Arabic linguistic tradition*. London and New York: Routledge.

——. 2004. "Meanings of speech: The category of sentential mood in Arabic grammar". *Le voyage et la langue: Mélanges en l'honneur d'Anouar Louca et d'André Roman*, ed. Joseph Dichy et Hassan Hamzé, 269–287. Damascus: Institut Français du Proche Orient.

Zimmermann, Fritz W. 1981. *Al-Farabi's commentary and short treatise on Aristotle's* De interpretatione*, translated with an introduction and notes*. London: Oxford University Press.

Michael G. Carter (Sydney University)

Passive

1. Morphology

The finite passive is formed two ways in Arabic: internally (the apophonic passive) and externally (formed by a prefix). The apophonic passive displays the vowel sequence *u – i* instead of *a – a* or *a – i* of its active counterpart in the perfect. In the imperfect, the apophonic passive uniformly displays the vowel *a* instead of *i/u* (as the second vowel), and all the forms are inflected with the *u*-series of the prefixes:

(1)		perfect		imperfect	
	I	*faʿala*	*fuʿila*	*yafʿa/i/ulu*	*yufʿalu*
	II	*faʿʿala*	*fuʿʿila*	*yufaʿʿilu*	*yufaʿʿalu*
	IV	*ʾafʿala*	*ʾufʿila*	*yufʿilu*	*yufʿalu*

The apophonic system 'leaks' in the formation of the causative imperfect (IV), which is homophonous with the basic imperfect (I). Hence, the ambiguity of the type *yuṭʿamu* 'he is eaten' (active *yaṭʿamu* 'he eats') and 'he is fed' (active *yuṭʿimu* 'he feeds'; Fischer 2002:120). Originally, however, the passivized causative imperfect displayed the characteristic structure of the causative stem **yuʾafʿalu*, which yielded the attested form by contraction *yufʿalu* (uncontracted causative forms are preserved in Akkadian *u-ša-pris* and Aramaic *yə-ha-ḵtiḇ*).

'Reflexive' (or rather 'mediopassive' or '→ middle voice' verbs) are passivizable in the same fashion as in (2).

(2)		
	perfect	
V	*tafaʿʿala*	*tufuʿʿila*
VIII	*iftaʿala*	*uftuʿila*
X	*istafʿala*	*ustufʿila*
	imperfect	
V	*yatafaʿʿalu*	*yutafaʿʿalu*
VIII	*yaftaʿilu*	*yuftaʿalu*
X	*yastafʿilu*	*yustafʿalu*

Some examples are provided in (3):

(3) *tawaffā-hu l-lāhu* 'God has taken him unto Him', passive *tuwuffiya* 'he was taken [by God]; he died'
iftahama-hu 'he understood him', passive *uftuhima* 'he was understood'
istaqāma-hu 'he made him stand straight', passive *ustuqīma* 'he was made to stand straight'

The verbs of Form VII, formed by the prefix *n-*, possess for the most part passive meaning; nevertheless, there are also many whose meaning can be described in the broadest sense as that of 'middle voice'. Therefore, it is customary to refer to this stem as 'reflexive-passive' (or rather 'mediopassive'). For instance, *inqāda*, Form VII of *qāda* 'to lead', means not only 'to be led' but also 'to follow, obey' (cf. Indo-European *verba deponentia*, Latin *sequor* 'to follow', Greek ἕπεσθαι 'to follow', Sanskrit *sacate* 'to follow', all of them displaying the mediopassive morphology). In addition, the passive meaning is found also with the Form VIII *iqtāda* 'to lead (for oneself)' and 'to be led'.

The same is true of modern vernaculars. For instance, Cowell (1964:238–239) gives several verbs distinguishing 'mediopassive' from 'true passive' in Syrian Arabic: *žtamaʿ* 'to meet, get together' vs. *nžamaʿ* 'to be brought together'; *mtadd* 'to extend, stretch' vs. 'to be extended, stretched'; *rtafaʿ* 'to rise' vs. *nrafaʿ* 'to be raised'; etc., but then he reminds the reader that mediopassive *rtafaʿ* can also mean 'to be raised', and, vice versa, passive *nmadd* can also mean 'to extend, stretch [intransitive]'.

The distribution of *n-* and *t-* forms across the spectrum of Arabic dialects is quite complex. According to the productive passive marker formation (in Form I), it is possible to distinguish (following Retsö 1983:164ff.) *n-* dialects, *t-* dialects (with *t-* prefixed), and *Gt-* dialects (with *-t-* infixed). Without going into their geography, it is important to realize that (i) there is a common stock of *nG* and *Gt* forms in all dialects, and (ii) Classical/Standard Arabic is an *nG* dialect. This classification also bears on the entire Semitic family (Hebrew and Akkadian are *n-* dialects, while Aramaic and Ethio-Semitic are *t-* dialects).

The polysemy of both *n-* and *t-* forms (passive ~ middle voice) undoubtedly contributed to the emergence of double marked passive forms, i.e. apophonic forms of Form VII and VIII, which are unambiguously passive in Classical Arabic, as in (4).

(4) VII *infaʿala unfuʿila*
VIII *iftaʿala uftuʿila*
unqīda < **unquyida)* ~ *inqīda* 'to be led'
uqtīda < **uqtuyida)* ~ *iqtīda* 'to be led'

Given the two synthetic passives and the adjectival nature of the passive participle, one does not expect an analytic (periphrastic) passive construction (familiar from Modern Germanic and Romance languages) in Modern Standard Arabic. Nevertheless, Cantarino (1975:417) quotes the example in (5) from al-Māzinī involving the passive participle of Form II:

(5) *wa-lākinnī kuntu muqayyadan bi-l-kitābi*
'But I was guided by the book'

Its synthetic counterpart would be *inqāttu* (Form VII) or *iqtāttu* (Form VIII) 'I was led' (or, perhaps, *unquyyittu* [passivized Form II] 'I was guided') with a 'disfigured' root due to the voicing assimilation.

In vernaculars, one finds passive perfect forms realized by the passive participle (without the verb 'to be'). Brustad (2000:184) gives example (6) from Moroccan Arabic:

(6) *ʾīwa l-ḥawli məšri*…
'Well, the sheep has been bought…'

There are also instances of an analytic gerundival construction (combining the verb 'to be' with the passive participle in its modal meaning of the future participle of necessity). The following example is from Nagīb Maḥfūẓ (*Zuqāq* 40.4):

(7) *ḥattā yaqḍiya l-lāhu ʾamran kāna mafʿūlan*
'until God determines what has to be done'
(cf. Latin…*rem faciendam* or *rem, quae facienda est*)

2. SYNTAX

In contemporary syntax, it is commonly assumed (as it was already by the Arab grammarians) that the passive construction is syntactically derived from its active counterpart. In morphosyntactic terms, active constructions containing transitive verbs and, to a certain degree, even intransitive verbs are passivizable. The following semantico-syntactic categories may be assigned the function of the grammatical subject in the passive construction: in Modern Standard Arabic, (i) direct objects and (ii) prepositional objects; in Classical Arabic, also certain (iii) expressions of place (distance) and time (duration), and (iv) cognate accusatives (if modified or specified). Under the functional approach, actives and passives differ in regard to subject assignment (Siewierska 1991:79ff.). In Classical Arabic, the functions of the agent and the goal are expressed by the nominative and the accusative, respectively. The GoalObject of the active sentence is realized as the GoalSubject in the passive sentence, which involves its movement into the immediate postverbal position, as in (8).

(8) *ḍaraba ʿumaru zaydan* (GoObj) ⇒ *ḍuriba zaydun* (GoSubj)*[min ʿumara]*
'Zayd hit Omar' ⇒ 'Zayd was hit (by Omar)'

This movement does not take place when the theme is realized as a subordinate object clause and the experiential phrase is added, as in (9).

(9) *ʿarafa zaydun ʾannahā min lubnāna* ⇒ *ʿurifa min zaydin ʾannahā min lubnāna*
'Zayd knew that she was from Lebanon' ⇒ 'It was known to Zayd that she was from Lebanon'

Arabic may assign the syntactic function of subject not only to the goal but also to the semantic function of recipient with three-argument verbs of 'giving' (cf. English *I was given a book*):

(10) a. *ʾaʿṭaytu zaydan* [Rec] *kitāban* [Go]
'I gave Zayd a book'
b. *ʾuʿṭiya kitābun* [Go, Subj] *zaydan* [Rec]
'The book was given to Zayd'
c. *ʾuʿṭiya zaydun* [Rec, Subj] *kitāban* [Go]
'Zayd was given a book'

The retained 'paradoxical' accusatives of Classical Arabic in (10b) and (10c) are abandoned in Modern Standard Arabic (the function of receiver is realized analytically by the prepositional phrase: *ʾaʿṭaytu l-kitāb li-zayd* ⇒ *ʾuʿṭiya l-kitāb li-zayd*, and *ʾuʿṭiya zayd al-kitāb* 'Zayd was given the book').

Prepositional objects (phrases) are treated in the same fashion from the point of view of the assignment of the subject, as in (11) and (12).

(11) *marartu bi-zayd* ⇒ *murra bi-zayd*
'I passed by Zayd' lit. 'it-was-passed by Zayd'
'Zayd was passed by'

(12) *nāma zayd fī l-firāš* ⇒ *nīma fī l-firāš*
'Zayd slept in the bed' lit. 'it-was-slept in the bed'
'This bed was slept in'

The passivized versions in (11) and (12) need a special context. If one wants to assure the police searching the bedroom for an escapee that 'nobody has slept in this bed', it is possible to predicate to 'bed': *mā nnām b-hat-taxᵊt* 'this bed hasn't been slept in'. Typologically, these passive constructions correspond to the 'pseudoreflexive' constructions of Slavic, Altaic, and other languages (e.g. Turkish *bu yol-dan gid-il-ir*, Czech *jde se touto cestou* lit. 'it-is-gone by this path', 'one goes by this path').

Classical Arabic may assign the function of subject even to expressions of place (distance) and time (duration). Saad (1982:31) provides examples (13) and (14):

(13) *sāra zaydun mīlayni* ⇒ *sīra mīlāni*
'Zayd walked two miles' 'Two miles were walked'

(14) *sāra zaydun sāˁatayni* ⇒ *sīrat sāˁatāni*
'Zayd walked two hours' 'Two hours were walked'

It should be observed that subject agreement in the passive versions indicates that we are dealing with local and temporal nouns that were indeed assigned the function of grammatical subject. This is impossible in Slavic pseudoreflexive constructions; in Czech, for instance, one cannot say **šla se hodina* [nom.] 'one hour was walked'; only the impersonal construction, which retains the accusative form of the temporal noun, is possible (*šli jsme hodinu* [acc.] 'we walked one hour' ⇒ *šlo se hodinu* [acc.]). It is not known whether the above constructions in Classical Arabic are found enlarged by the agentive phrase, e.g. *?sīra mīlāni min an-nāsi* 'people walked two miles'. Notice, however, that it is impossible to add the agentive phrase to their impersonal counterparts in Slavic (*Czech šlo se *od nás/*námi*), and that their German parallels are extremely unlikely (*oben wird *von uns getanzt*).

The same is true of the passivized expressions involving cognate accusatives (*jalasa zaydun julūsan ḥasanan* lit. 'Zayd sat down a nice sitting-down ⇒ *julisa julūsun ḥasanun* lit. 'a nice sitting down was sat'). Neither of the above two types is productive, and they do not exist in Modern Standard Arabic.

3. AGENTIVE PHRASE

Classical Arabic belongs to those languages which normally do not add the agentive phrase to the passive construction because the agent "was not known or ought not to be mentioned" (to quote Cantarino 1974:52). The former constraint is a general pragmatic restriction on the expressibility of the agent; the latter can be related to the specific cultural phenomenon of not naming God in numerous expressions involving the divine agency.

If the agent has to be mentioned, it is possible to augment the agentless passive construction with its active version, in which the direct object is replaced by the pronominal clitic. Cantarino (1974:53) quotes the following relevant example (15) from Ṭāhā Ḥusayn's *al-ʾAyyām* (II, 42.4):

(15) *qad muddat al-māʾidatu, maddat-hā kubrā ʾaxawāti ṣ-ṣabiyyi*
lit. 'the table was set, set-it the boy's oldest sister'
'The table was set by the boy's oldest sister'

Modern Standard Arabic can add the agentive phrase to the passive construction by means of a prepositional phrase involving the preposition *min* 'from' or, less frequently, *bi-* or *li-*. *Min* possesses two senses: partitive '[part] of' and ablative 'from'. As in Latin, the ablative sense lends itself to the agentive/instrumental usage (origin > cause > agent/instrument). Cantarino (1974:53) gives example (16) from Muḥammad Ḥaykal (*Ḥayāt* 36.21):

(16) *X. yūḥā ʾilayhi min al-lāhi ḏātihi*
'X. has been revealed to him by God himself'

The semantics of *bi-*, which has two notions, that of proximity ('in contact with') and that of instrumentality ('by means of'), makes this preposition less suitable for the expression of agentivity. Thus, the sentence *ḍuriba bi-l-ˁasā* 'he was struck with a stick', if augmented by another prepositional phrase featuring *bi-*, would be ambiguous between 'he was struck with a stick by Zayd' or 'he was struck with a stick instead of Zayd' (as in *ḍuriba bi-zaydin* 'he was struck instead of Zayd'). To guarantee the agentive reading, the agentive phrase would have to be introduced by *min* 'from' > 'by': *ḍuriba bi-l-ʿasā min zaydin.*

The examples for the putative agentive use of the preposition *li-* 'to', quoted by Fischer (2002:110, 157) and Cantarino (1974:53), are apparently limited to the *verba sentiendi* (but Reckendorf 1921:26 provides Classical Arabic examples of the type *al-laḏīna yusʾalūna lahu* 'those who were asked by him'). They are of the type *tuhālu lahu l-ʿaynu* lit. 'the eye is frightened by him', 'one is struck with terror by him'. Here, we are not dealing with (proto)typical action verbs, and their passive subjects are not patients in the ordinary sense (*ḍaraba-hu* 'he hit him') but rather experiencers (*ʾahāla-hu* 'he frightened him ⇒ *ʾuhīla* 'he was frightened') as in (17).

(17) *jaʿala ṣ-ṣabiyyu yurāʿu li-hāḏayni ṣ-ṣawtayni*
'The boy began to be frightened by these two voices' (Ḥusayn, *ʾAyyām* II, 42.4)

In vernaculars, one finds the agentive phrase in the passive construction introduced by the simple *min* (e.g. *l-ʾəttifāʾiyye lāzem tətsaddaʾ mən mažles ᵊš-šuyūx* 'the treaty has to be ratified by the senate') or, more commonly (?), by one of its several compound varieties, such as (Syrian) *mən qəbal* (*qibal* 'power, authority') and *mən ṭaraf* (*ṭaraf* 'side'). According to Cowell (1964:236), "such usage is limited to a rather pedantic classicizing style". Standard Arabic has several other compound agentive prepositions such as *ʿalā yad* lit. 'at the hand of', *min jānib* lit. 'from the side of', *min qibal* lit. 'from the authority of' (examples in Retsö 1983:26).

4. MODAL MEANING

A special modal meaning of possibility and necessity results from the agentless passive along the grammaticalization cline: 'X that is done' > 'X that CAN be done' > 'X that is WORTHWHILE doing' > 'X that SHOULD be done'. Examples in (18) are quoted by Cantarino (1974:56) from Tawfīq al-Ḥakīm and Muḥammad Mandūr, respectively:

(18) *laysa laka rūḥun tuqbaḍu*
lit. '...a spirit [that] is taken away' ⇒ 'you do not have a spirit that CAN be taken away'

lam yakun ladaynā minhu šayʾun yuḏkaru
lit. '...a thing [that] is mentioned' ⇒ 'we did not have any that are WORTH mentioning'

Many of these verb phrases have been lexicalized as expressions corresponding to Indo-European adjectives of possibility (and impossibility): *yuḏkaru šayʾun* 'a thing [that] is mentioned' > *šayʾun yuḏkaru* 'a thing [that] is mentioned' > 'a thing [that] is WORTH mentioning'; *lā yuṣaddaqu šayʾun* 'a thing [that] is not believed' > *šayʾun lā yuṣaddaqu* 'an unbelievable thing'. The adjectival status of these finite verbal phrases is indicated by their postnominal position (compare *šayʾun yuḏakkaru* 'a thing [that] is worth mentioning' with *šayʾun maḏkūrun* 'a memorable thing').

Also, passive participles of certain verbs can be used as adjectives of possibility ('X that can be done'), as in (19).

(19) *maʿqūl* 'understood' ⇒ 'conceivable'
masʾūl (ʿan) 'asked (about)' ⇒ responsible'
muʿtamad (ʿalā) 'relied (on)' ⇒ 'dependable'

An example of the necessitative reading of the finite passive participle from ar-Rayḥānī is in (20).

(20) *...wa-l-ḥaqqu yuqālu...*
lit. 'and the truth is told' ⇒ 'and the truth SHOULD be told' (Rayḥānī, *Mulūk* 25.23)

Several passive participles acquired the meaning of the future participle of necessity (corresponding to the Latin gerundive): e.g. *al-xarūf al-maḏbūḥ* lit. 'the slaughtered lamb' ⇒ 'the lamb TO BE slaughtered' (cf. Latin *agnus immolandus*).

Thus, the same participle may perform a triple duty of passive participle, adjective of possiblity, and future participle of necessity (corresponding to the Indo-European gerundive); e.g. *madkūr* 'remembered', 'memorable', and 'to be remembered' (contrast this polysemy of a single form with the three-way morphology of Latin: *memor-ātus, memor-ābilis*, and *memor-andus*).

5. PRAGMATICS

While the passive voice occurs in spoken Arabic, in many cases speakers prefer a topicalized construction of the type in (21)

(21) *farīʾna ġalabo n-nādi r-riyāḍi* (Syrian)
'Our team, the Athletic Club beat it'

to its passive counterpart in (22)

(22) *nġalab farīʾna mən (ṭaraf) n-nādi r-riyāḍi*
'Our team was beaten by the Athletic Club'

In negative terms, the use of the passive construction is practically ruled out in the illocutionary acts of commanding, questioning, and responding to questions (dialogue). For instance, if one is asked *man fataḥa š-šabābīk* 'who opened the windows?' (structuring the question in the passive is only conceivable as an echo question 'the windows were opened by WHOM?'), the response is going to be in the active of the interrogator: *zayd fataḥa-hā* 'Zayd opened them'. The participant in this dialogue is not going to switch to the passive voice **infataḥat aš-šabābīk min zayd*. This option, however, is claimed to be available to 'subject-prominent' languages such as English ('Who opened the windows?' 'The windows were opened by John'). The overwhelming proclivity of Arabic to use the topicalized version in the active voice derives from one of its salient typological features, namely that of a 'topic-prominent' language (cf. Bubeník 1979; Brustad 2000). The → topic (*mubtadaʾ* lit. 'beginning'; → *ibtidāʾ*) of the Arab grammarians is the noun or noun phrase which introduces the comment (→ *xabar* lit. 'message') and delimits its scope (individual, spatial, temporal). It is possible for the topic and comment not to be linked morphologically (of the type 'the eggs, the dozen is at fifty piasters'). However, the construction exemplified in (21), featuring the pronominal clitic whose antecedent is the topic, is by far most common (→ topicalization).

Finally, one may ask the question in positive terms: in which circumstances may the speaker of Arabic prefer the passive construction to its active counterpart? Xrakovskij's (1975) detailed inquiry into this matter has been summarized by Retsö (1983:183–186) in three points: (i) the passive construction is preferred if the agent is referred to by another argument whose semantic role is different (location, possession, instrument), e,g. *bə-ṣ-ṣēf ʾəḥtəreʾ bəš-šams* 'in the summer I want to get burnt by the sun' (the agent is identical with the instrument); (ii) the agent not being mentioned results in a context-conditioned finite passive voice, which serves as a stylistic means of binding sentences together; and (iii) the agent is unspecified (according to Xrakovskij, there are different degrees of non-specification, ranging from the unknown actor to a vague specification of professional agent). It goes without saying that one does not expect to encounter the agent with verbs designating 'events' that may take place without any agent (such as expressions of 'state' and 'change of state' realized typically by the mediopassive forms of Form VIII, e.g. *irtafaʿ mustawā l-māʾ fī n-nahr* 'the river rose').

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Ḥaykal, *Ḥayāt* = Muḥammad Ḥusayn Ḥaykal, *Ḥayāt Muḥammad*. 5th ed. Cairo, 1952.
Ḥusayn, *al-ʾAyyām* = Ṭāhā Ḥusayn, *al-ʾAyyām*. Cairo, 1st vol. n.d., 2nd vol. 1958.
Maḥfūẓ, *Zuqāq* = Nagīb Maḥfūẓ, *Zuqāq al-midaqq*. 4th ed. (?). Cairo, 1961.
Rayḥānī, *Mulūk* = ʾAmīn ar-Rayḥānī, *Mulūk al-ʿArab*. 3rd ed. Beirut, 1951.

SECONDARY SOURCES

Brustad, Kristen E. 2000. *The syntax of spoken Arabic*. Washington, D.C.: Georgetown University Press.
Bubeník, Vít. 1979. "Thematization and passivization in Arabic". *Lingua* 49.295–313.
Cantarino, Vicente. 1974, 1975. *Syntax of Modern Arabic prose*. I. *The simple sentence*. II. *The expanded sentence*. Bloomington: Indiana University Press.
Cowell, Mark W. 1964. *A reference grammar of Syrian Arabic*. Washington, D.C.: Georgetown University Press.
Fischer, Wolfdietrich. 2002. *A grammar of Classical Arabic*. 3rd rev. ed., transl. from German by Jonathan Rodgers. New Haven, Conn.: Yale University Press.
Reckendorf, Herman. 1921/1977. *Arabische Syntax*. Heidelberg: C. Winter.

Retsö, Jan. 1983. *The finite passive voice in modern Arabic dialects*. Göteborg: Acta Universitatis Gothoburgensis.
Saad, George Nehmeh. 1982. *Transitivity, causativization and passivization*. London: Kegan Paul International.
Siewierska, Anna. 1991. *Functional grammar*. London: Routledge.
Xrakovskij, V.S. 1975. "Formal'naja xarakteristika i upotreblenije pasivnyx konstrukcij v literaturnom arabskom jazyke". *Folia Orientalia* 16.13–36.

Vít Bubeník
(Memorial University of Newfoundland)

Passive (Syntax)

The passive in Arabic is a sentence structure in which the semantic subject or agent, i.e. the performer of or person/thing responsible for an action, is suppressed and, in fact, cannot be mentioned. This renders the passive in Arabic an impersonal structure. In the passive, the understood object of the active verb is the subject of the passive sentence and is marked for this role by nominative case in the Classical/Standard Arabic variety. The verb changes into the passive by either changing the vowels in the stem and tense prefix or by the insertion of a prefix (→ passive).

1. STRUCTURAL PROPERTIES OF THE PASSIVE

An active sentence changes into the passive by undergoing a number of structural changes. First, the subject of the sentence is deleted. Second, the object becomes the subject of the sentence and receives nominative case. Third, the active verb changes into the passive by changing its vowels, the change being dependent upon the tense or type of the verb, as described below. Fourth, the verb agrees in person and gender with the new subject in the Verb-Subject order and in person, gender, and number in the Subject-Verb order.

(1) *našara* *l-kātib-u*
published.3ms the-writer-Nom
l-maqāl-at-a
the-article-fs-Acc
'The writer published the article'

(2) *nušira-t* *al-maqāl-at-u*
Pass.published-3fs the.article-fs-Nom
'The article was published'

(3) *al-maqālatu nuširat*

Like active sentences in Arabic, the passive sentence can be subjectless, i.e. with an implicit subject represented by the agreement features expressed on the verb.

(4) *nušira-t*
Pass.published-3fs
'It was published'

The same rules apply in spoken Arabic, except that the verb is marked for the passive by the addition of a prefix rather than vowel change (see Sec. 4).

Because of their reliance on both structural and semantic properties of the passive in their account of it, the name given by the Arab grammarians to the passive indicates its most outstanding property – the absence of the subject. Thus, the passive voice is called 'that whose subject is not named' or *ṣīġat al-mafʿūl* 'the form of the patient' (Wright 1967:I, 50). To indicate that the subject of the passive is the semantic object/patient, the subject of the passive is also called *nāʾib al-fāʿil* 'the subject substitute' (Gaballa 1986:244), 'supplying the place of the agent' (Wright 1967:I, 50), or 'the deputy or representative of the doer' (Haywood and Nahmad 1965:144).

The subject of the passive is the object of the verb in the active counterpart. Once the morphology of the verb has changed to the passive, the object receives nominative case and becomes the structural subject of the sentence. Semantically, however, it retains its understood patient function in relation to the verb.

In terms of constituency, the subject of a passive sentence can be a noun phrase, an adverb, e.g. 'the month Ramadan' in *ṣīma ramaḍānu* 'Ramadan was fasted' (Wright 1967:II, 270), a → *maṣdar*, e.g. *ṭuʿina ṭaʿnatan* 'he was stabbed once', or a prepositional phrase (Babtī 1992).

2. FORM OF THE PASSIVE VERB

Passivization is a phenomenon normally associated with transitive verbs. Transitive verbs in Arabic, in both their primitive and derivative forms, can occur in the passive. The verb in Classical and Modern Standard Arabic changes into the passive by changing the vowels in the active verb. Overall, the change is regular for each verb type, and the form it takes hinges upon the tense of the verb. The perfect verb is changed into the passive by changing the vowel in the first syllable into a *ḍamma* /u/ and the vowel before the last radical into a *kasra* /i/ – thus taking the pattern *fuʿila*, e.g. *kusira* 'was broken' for a trilateral and *fuʿlila* for a quadrilateral verb, e.g. *turjima* 'was translated'. Some verbs in Form VII of trilateral verbs, marked by the addition of the prefix *in-*, have a passive meaning, e.g. *inhazama* 'was defeated'. Significantly, this is the prefix that has become the predominant passive-forming method in spoken Arabic, together with its phonological variant *it-*.

In the imperfect verb, the vocalic change indicating passivization is realized differently. The *ḍamma* /u/ replaces the tense and agreement prefix vowel; the vowel in the second syllable changes to a *fatḥa* /a/.

(5) *ʾurīdu* *ʾan* *tunšara*
Imp.1s.want that Imp.3fs.Pass.publish
l-maqāl-at-u
the-article.fs-Nom
'I desire/want the article to be published'

In spoken Arabic, the prefix *in-/it-* is attached to the active verb to change it into the passive. The voweling internal to the stem does not change. Thus, the passive verb has the form illustrated in the following structures:

Imperfect: tense + gender + Passive + verb stem

(6) *il-kitāb* *biy**it**bāʿ* *fi* *maʿraḍ*
the-book Pres.3ms.Pass.sell in fair
il-kitāb
the-book
'The book is on sale at the book fair'

Perfect: Passive + verb + tense + gender

(7) *il-kitāb* ***it**bāʿ* *fi* *maʿraḍ*
the-book PerfPass.3ms.sold in fair
il-kitāb
the-book
'The book was on sale at the book fair'

In Modern Written Arabic orthography, the marks indicating short vowels are hardly ever used, making the passive verb in its written form identical to its active counterpart. Badawi a.o. (2004:383), however, point out that readers are rarely confused, as they use the absence of the agent as a syntactic cue to interpret the structure as a passive one. They also point out that it is only when there is potential ambiguity that short vowels are used.

3. VERBS THAT PASSIVIZE

As indicated above, the passive in Arabic is associated with transitive verbs. Verbs taking two objects, i.e. ditransitive verbs, allow both the direct and indirect objects to become the subject of a passive sentence.

(8) *sallama* *ʾax-ī* *r-rajul-a*
handed.3ms brother-my the-man-Acc
r-risāl-at-a
the-message-fs-Acc
'My brother handed the man the message'

(9) *sullimat* *ar-risāl-at-u*
Pass.handed.3fs the-message-fs-Nom
li-r-rajul-i
to-the-man-Gen
'The message was handed to the man'

(10) *sullima* *r-rajul-u*
Pass.handed.3ms the-man-Nom
r-risāl-at-a
the-message-fs-message-fs-Acc
'The man was handed the message'

(11) *sullima* *r-risāl-at-a*
Pass.handed.3ms the-message-fs-Acc
'He was handed the message'

In each case, the verb must agree in gender with the subject of the passive clause.

An interesting property of Arabic is that verbs with prepositional complements can passivize. In a discussion of these verbs, Haywood and Nahmad (1965:253–254) point out that transitivity in Arabic is possible through a preposition, especially with verbs of motion. One of the verbs that passivize in this manner is *jāʾa* 'to come'. It occurs in two patterns, and its meaning changes depending on the pattern in which it occurs. In *jāʾa ʿumaru* 'Omar came', for instance, the verb does not take any complements. It can optionally take an adverb or a prepositional phrase with an adverbial function, i.e. an adjunct as in *jāʾa ʿumaru fī l-masāʾi* 'Omar came in the evening'. In this pattern, the verb cannot occur in the passive **jīʾa fī l-masāʾi* ‘he was come in the evening’. But in the other pattern, where the verb is followed by a prepositional complement, it can passivize.

(12) *jāʾa* *l-ḥurrās-u*
came the-guards.mp-Nom
bi-l-muttaham-īna
with-the-defendants.mp-Gen
‘The guards brought the defendants’

(13) *jīʾa* *bi-l-muttaham-īna*
Pass.came.3ms with-the-defendants.mp-Gen
ʾila *l-maḥkam-at-i*
to the-court-fs-Gen
‘The defendants were brought to the courthouse’

Notably, the verb *jīʾa* in (13) changes its meaning from ‘come’ to ‘bring’.

In traditional treatments of prepositional verb passives, the prepositional phrase is taken to be the subject or substitute of the subject. In the active voice, Arab grammarians call these prepositional complements ‘impure objects’ (*mafʿūl ġayr ṣarīḥ*), indicating that their being prepositional gives them a different status from noun phrases or ‘pure’ objects. In modern analyses, the passive in (13) is assumed to have an implicit, expletive pronoun in the subject position (Agameya 2001; Mohammad 1987). This is based on the evidence that the verb carries default 3rd person singular masculine agreement in prepositional passives, regardless of the nature of the features of the object of the preposition, even in the Subject-Verb order, which suggests that the subject is perhaps a nonovert, expletive element (equivalent to the overt impersonal subjects *it* and *there* in English; Mohammad 1987).

(14) *al-muttaham-ūna*$_i$ *jīʾa*
the-defendants.mp-Nom Pass.brought.3ms
*bi-**him***$_i$
with-them
‘The defendants were brought’

The lack of full agreement between the subject and verb in (14) suggests that the subject is in → topic position. A striking syntactic property of this sentence is the resumptive pronoun (→ resumption) that is attached to the preposition and which co-refers to the subject *al-muttahamūna*. The resumptive pronoun is a requirement of Arabic syntax, which does not allow prepositions to strand.

In spoken varieties, the range of prepositional verbs allowing passivization is much narrower than that in the standard variety. Example (13) does not have a grammatical counterpart in spoken Egyptian Arabic, for instance. The equivalent would use an impersonal structure, with an active verb (15).

(15) *gābu* *l-muttahamīn*
brought.3mp the defendants.mp
‘They brought the defendants’

The following are examples of a prepositional verb in spoken Arabic occurring in the passive in both word orders.

(16) *it-ʾabaḍ* *ʿala*
Pass.arrested.3ms on
l-mugrimīn
the-criminals.mp
‘The criminals have been arrested’

(17) *il-mugrimīn*$_i$ *it-ʾabaḍ*
the-criminals.mp Pass.arrested.3ms
ʿalē-hum$_i$
on-them
‘The criminals were arrested’

Of these two examples, the second is marked phonologically, as the fronted object of the preposition receives higher stress, indicating its status as a topicalized element (Agameya 2001).

4. THE AGENT IN THE ARABIC PASSIVE

The property of Classical Arabic passives that the agent cannot be mentioned makes it impossible for certain passive structures requiring the obligatory presence of the agent phrase in other languages, e.g. the ‘by-phrase’ in English, to be expressed in Arabic using the passive. It follows that a sentence like *Oliver Twist was written by Dickens*, in which the agent phrase *by Dickens* cannot be omitted without destroying the meaning, must be translated into Arabic using the active voice.

In Modern Standard Arabic, however, the agent phrase is sometimes expressed. Badawi a.o. (2004) observe, based on a corpus of current written Arabic, that the agent can be expressed following a number of prepositions. They ascribe the emergence of the agent prepositional phrase in modern Arabic passives to the influence of European languages. The following examples are adapted from Badawi a.o. (2004:385).

(18) *ʾuṭliqat* *ʿalayhi* *n-nāru* ***min***
Pass.shot.3fs on-him the-fire from
qibali ***ʿiṣābāti*** ***l-māfya***
direction gangs the-mafia
'He was fired on by the Mafia gangs'

(19) *ʾudīna* ***bi-wāsiṭati***
Pass.sentenced.3ms by-means.of
l-maḥākim *al-ʿaskariyya*
the-courts the-martial
'He was sentenced by the court martial'

In both the standard and spoken varieties, a prepositional phrase containing the instrument can be used. This, though, is not to be confused with the agent prepositional phrase.

(20) *qutila* *bi-s-sikkīn*
Pass.killed.3ms by the-knife
'He was killed with a knife'

In spoken Arabic, the use of the agent phrase is more limited both in the range of prepositions used and in its meaning (see Sec. 5).

5. Uses of the passive

The reasons for using the passive voice in Arabic are purely pragmatic. It is used mainly when the agent responsible for performing the act is not known or when the speaker chooses not to mention it. The passive is also used to place emphasis on the object, i.e. the patient. When the subject of the passive is absent, the structure is said to be impersonal. One interpretation of impersonal passives, identified by Wright (1967:II, 50), is when the agent is understood to be God.

(21) *kutiba* *ʿalay-kumu*
Pass.written.3ms on-you
ṣ-ṣiyām (Q. 2/183)
the-fasting
'You are required to fast'

When the agent/subject of a verb is not known or the speaker does not wish to mention it, there are two options for expressing the proposition: a personal or impersonal structure (Wright 1967:II, 266–267). The 'personal' way essentially uses the active voice without an explicitly stated subject pronoun, i.e. → pro-drop, with 3rd person masculine plural agreement on the active verb.

(22) *yaqūlūna* *ʾanna* *l-ʾasʿāra*
Imp.3.say.mp that the-prices
sa-tartafiʿ
Fut.rise
'They say that prices will go up'

The second option, the impersonal, employs the passive voice with 3rd person masculine singular agreement.

(23) *yuqālu* *ʾanna* *l-ʾasʿāra*
3.Pass.say.ms that the-prices
sa-tartafiʿ
Fut-rise
'It is said that the prices will go up'

In spoken Egyptian Arabic, both personal and impersonal forms are used.

(24) *bi-yʾūlu* *inn* *il-ʾasʿār* *ḥa-tirtifiʿ*
Imp-3.say.mp that the-prices Fut-rise

(25) *bi-yuqāl* *inn* *il-ʾasʿār*
Imp-3.Pass.say.ms that the-prices
ḥa-tirtifiʿ
Fut.rise

The last example is interesting in that the form of the verb combines the colloquial imperfect tense prefix *bi-* and the typical Standard Arabic passive form. This form is limited in its use to educated speech and formal contexts.

The passive is sometimes used in written Arabic with an imperative sense "in labels and instructions for use" (Badawi a.o. 2004:389).

In spoken Egyptian Arabic, it is sometimes used as an imperative, where the speaker assumes absolute authority over someone else perceived to be of lower status.

(26) *il-ʿarabiyya* *titnaḍḍaf* *kuwayyis*
the-car Pass.clean.3fs well
'The car is to be cleaned well'

To convey the same meaning with a toned-down effect, obliterating the speaker's absolute authority in (26), an impersonal active structure is used.

(27) *il-ʿarabiyya* *miḥtāga* *tanḍīf*
the-car needs cleaning
'The car needs to be cleaned'

In spoken Arabic, an agent phrase realized as a prepositional phrase with *min* 'from' in the passive has a special function when it occurs with verbs denoting some damage being done. Its use identifies the person responsible for the act and significantly exonerates this person from having done it deliberately or as an act of negligence. The effect is that the agent, the '*min* phrase', inadvertently caused the damage.

(28) *il-kubbāya* *it-kasarit*
the-glass.fs Pass-broke.3fs
min-ni/min yūsif
from-me/from Yousef
'The glass was broken by me/Yousef'

The use of an agent phrase is not very frequent in Egyptian Arabic, however. It can be used in contexts where the agent has a general, rather than specific, reference.

(29) *il-ʾiqtirāḥ* *da* *mitʾayyid*
the-proposal this Pass-supported
min *kull* *il-ʾaʿḍāʾ*
from all the-members
'This proposal is backed by all the members'

(30) **il-ʾiqtirāḥ* *da* *mit-ʾayyid*
the-proposal this Pass-supported
min *ʾaḥmad*
from Ahmed
'This proposal is backed by Ahmed'

The ungrammaticality of (30) is due to the fact that the prepositional phrase *min ʾaḥmad* refers to a specific individual.

6. INTERPRETATION OF THE PASSIVE

In some treatments of the passive in different languages within the framework of generative grammar, it has been argued that the passive morpheme stands for the subject, i.e. that the structurally absent subject is in fact implicitly present in the interpretation of the structure. The empirical test employed to support this argument is the use of certain adverbs that modify the manner in which the subject performs the action in active sentences. There is evidence that the Arabic passive has an implied subject with indefinite reference (Agameya 2001:21). The implicit subject in passive structures can be modified by certain manner adverbs.

(31) *il-ʾustāz* *laġa*
the-professor canceled.3ms
l-muḥaḍra *ʿamdan*
the-lecture deliberately

(32) *il-muḥaḍra* *itlaġit*
the-lecture Pass.3fs.canceled
ʿamdan
deliberately
'The lecture was deliberately canceled'

The adverb *ʿamdan* 'deliberately' modifies the way the agent/subject performed the act expressed by the verb. It cannot be used with expletive subjects, since these subjects lack referents, as, for example, in **ʾamṭarat as-samāʾu ʿamdan* 'it rained deliberately'. The same test demonstrates that impersonal passives of prepositional verbs do not have implicit subjects.

(33) **itʾabaḍ* *ʿala* *l-barīʾ*
PerfPass.arrest.3ms on the-innocent
ʿamdan
deliberately
'The innocent person was arrested deliberately'

(34) **il-barīʾ* *itʾabaḍ*
the-innocent PerfPass.arrest.3ms
ʿalēh *ʿamdan*
on-him deliberately

Similarly, other types of impersonal passives cannot be modified by this type of adverb – indicating that no agent is involved in the interpretation of these structures.

(35) **biyuqāl* ʿ*amdan*
ImpPass.say.3ms deliberately
ʾ*inn il-barīʾ* *itʾabaḍ*
that the-innocent PerfPass.arrest.3ms
ʿ*al-ēh*
on-him
‘It is said deliberately that the innocent was arrested’

Again, this suggests that the pleonastic implicit subject in these sentences lacks reference.

7. PASSIVE-RELATED STRUCTURES

In some structures, often called ‘unaccusatives’ (→ middle verbs), the subject of the sentence, like that of the passive, is the semantic object of the verb. The verb, however, does not undergo passivization – the morphological change associated with the passive. An intrinsic difference between passives and unaccusatives is that unlike the case with passives, where the agent is implied even though it is phonologically absent, no agent is implied in these structures. The unaccusative clause can only contain an adverbial that modifies the verb; it can never allow one that modifies the absent subject. The following examples from Egyptian Arabic have an unaccusative structure (see Agameya 2001). Examples from Standard Arabic demonstrate the same structural property.

(36) *il-maḥall* *fataḥ/biyiftaḥ* *is-sāʿa*
the-store opened/opens.3ms the-hour
tisʿa
nine
‘The store opened/opens at nine o’clock’

(37) **il-maḥall* *fataḥ* ʿ*amdan*
the-store opened deliberately

(38) *il-maḥall* *fataḥ/biyiftaḥ*
the-store opened/opens.3ms
badri
early

In unaccusatives, the fact that the morphology of the verb maintains its active form prohibits an implicit interpretation of the absent subject, as the ungrammaticality of (37) demonstrates. This fact is used as evidence that the passive morpheme stands for the absent subject in passives.

Classical/Standard Arabic, furthermore, has the interesting property of changing the morphology of the verb to derive new forms expressing different, but semantically related, ideas. For instance, there is a group of verbs called ‘effective’, whose semantic patient (object) becomes the subject of the sentence, as is the case with the passive (Wright 1967). These verbs are derived from the → causative form of some verbs, marked by the doubling of the second radical. In these structures, no agent is implied.

(39) ʿ*allama* *l-ʾustāḏu*
Perf.taught.3ms the-teacher
t-tilmīḏa *l-qirāʾata*
the-pupil the-reading
‘The teacher taught the pupil to read/the art of reading’

(40) *taʿallama* *t-tilmīḏu* *l-qirāʾata*
Perf.learned.3ms the-pupil the-reading
‘The pupil learned to read’

The class of effective verbs in Egyptian Arabic is much smaller than that in Standard Arabic. Hence, the counterpart of (40) in spoken Arabic must take the passive morpheme, rather than be used in its active form.

(41) *il-ʾustāz* ʿ*allim*
the-teacher Perf.taught.3ms
it-tilmīz *il-ʾirāya*
the-pupil the-reading

(42) *it-tilmīz* ***it***ʿ*allim*
the-pupil PerfPass.learned.3ms
il-ʾir āya
the-reading

Causative verbs, however, can be used unaccusatively in spoken Arabic.

(43) *il-ḥarr* *sayyaḥ* *it-talg*
the-heat Perf.Caus.-melted.3ms the ice
‘The heat caused the ice to melt/the heat melted the ice’

(44) *it-talg* *sāḥ*
the-ice Perf.melted.3ms
‘The ice melted’

Bibliographical references

Agameya, Amira. 2001. "The passive and its properties". *Bulletin of the Faculty of Arts, Cairo University* 61.7–30.

Babtī, ʿAzīza. 1992. *al-Muʿjam al-mufīd fī n-naḥw al-ʿarabī*, II. Beirut: Dār al-Kutub al-ʿIlmiyya.

Badawi, Elsaid, Michael G. Carter, and Adrian Gully. 2004. *Modern Written Arabic: A comprehensive grammar*. London and New York: Routledge.

Gaballa, Hassan. 1986. *Case-marking and Classical Arabic: A semantic and historiographic study*. Ph.D. diss., University of Delaware.

Haywood, John A. and H.M. Nahmad. 1965. *A new Arabic grammar of the written language*. 2nd ed. London: Lund Humphries. (Repr., 1970.)

Mohammad, Mohammad A. 1987. "The problem of subject-verb agreement in Arabic: Towards a solution". *Perspectives on Arabic linguistics*, I, ed. Mushira Eid, 127–154. Amsterdam and Philadelphia: J. Benjamins.

Wright, William. 1967. *A grammar of the Arabic language*. 3rd ed. Cambridge: Cambridge University Press. (Repr., 1977.)

Amira Agameya
(Cairo University and The American
University in Cairo)

Pathology → Language Pathology

Pausal Forms

1. Introduction

A pausal form is the form a word has at the end of a sentence or major phrase or before a pause or stop in the speech flow (*waqf*), if that is different from the form it takes in the beginning or middle of a phrase. In Classical and Modern Standard Arabic, most words have different pausal and medial forms. Phonetic pausal phenomena probably occur in all languages, although they may differ from language to language, but morphologically conditioned pausal changes are much rarer, and they are the ones most often referred to when discussing pausal forms in Arabic. The morphological pausal phenomena of Arabic prose are of a single general type: the pausal form is produced by subtracting from the base form of the word a final short vowel and the final consonant of certain suffixes. The pausal form is thus shorter than the medial form, and grammatical features (case and mood) that in a full form are marked by suffixes consisting of a short vowel are absent from the pausal form. Therefore, the medial or context form of a word may be appropriately called its full form or basic form, in that the pausal form can be deduced from it but it cannot be derived unambiguously from the pausal form. When mentioning an Arabic word in isolation, one usually cites the pausal form, for two reasons: a word in isolation is, in effect, in pausal position; and the Standard Arabic pausal forms are, on the whole, more similar than the full forms are to the way the words are pronounced in the modern vernacular dialects of Arabic. Thus, if one asks what the Holy Book of Islam is called, or how to say 'city', the answer is given in the pausal form, *al-qurʾān* or *madīna*, not the full form *al-qurʾānu* or *madīnatun*.

The most detailed description of the pausal forms in a Western language is by Fleisch (1990: 172–197). Wright (1898:368–373) gives a concise but full statement of the facts, Retsö (1994) presents a lucid synthesis of them, Birkeland (1940) focuses on the historical development of the system, and Roman (1982:493–554) attempts to reconstruct the phonetics and phonology behind it.

2. Pausal forms in Standard Arabic

Pausal forms are derived from the basic, full forms by (i) deleting final short vowels (so the pausal form of *kataba* 'he wrote' is *katab*); (ii) deleting a final suffix *-n* (so the pausal form of *kitābun* 'book' is *kitāb*); (iii) replacing the suffix *-at-* with *-ah* (*kitāb-at-u-n* 'writing' becomes *kitābah*). This Standard Arabic morphological alternation between full and pausal forms is absent from the modern vernacular dialects and was lost in the medieval period, if not earlier; the vernacular form of a word is generally derived from the Standard Arabic pausal form, not the full form.

The deletion of final short vowels is integrally related to the fact that, in both Old Arabic and the modern vernacular dialects, with few exceptions, the phonemic opposition of vowel and consonant quantity is neutralized in pausal position. (In some modern dialects this applies only to unstressed vowels.) Thus, the final consonants of the words *yad* 'hand' and *radd* 'response' are identical, when not followed by another word (both may be pronounced with a longer or a shorter *d*). The same is true with vowels. In the Damascus vernacular, for example, "If a final vowel is…unaccented, it varies between long and short depending on the phrasing and intonation. Thus…the *i* in *xədi* ['take:IMPER:FEM:SG'] is unaccented (i.e. *xə́di*), but is sometimes actually long" (Cowell 1964:19), so the difference in the vowels between *xə́di* 'take!' and *xədī́-ha* 'take her!' is that in *xədī́-ha* the stressed *ī* must be long, but in *xə́di* the unstressed final *i* may be long or short. Something similar was likely the case in Old Arabic (although stress is not known to have been a factor), and the same is true of Modern Standard Arabic, where the final vowels of *ʾanti* 'you [fem. sg.]' and *bintī* 'my daughter' are pronounced identically.

A word that ends in a long vowel is unchanged in pausal position, but when a word that basically ends in a short vowel appears in pausal position, it either loses that vowel, lengthens the vowel, or adds *h* (Retsö 1994). Loss of the vowel is the norm for final short vowels that are suffixes or part of suffixes, so the pausal form of *al-bayt-u* 'the house [Art-house-Nom]', *muʿallim-ūna* 'teachers [teacher-Nom.p]', *bayt-u-ka* 'your house [house-Nom-2ms]', *daras-a* 'he studied [studied-3ms]', *daras-nā-hu* 'we studied it [studied-1p-3ms]', *ʿan-hu* 'from it' are *al-bayt, muʿallimūn, baytuk, daras*, *darasnāh, ʿanh* respectively. Lengthening is frequent at the ends of lines in poetry. Final short vowels that are not part of suffixes are (in normative Classical Arabic) followed in pause by *-h* (called by the grammarians *hāʾ as-sakt*, cf. Fleisch 1990:185–186), so the pausal forms of the jussive *yaqi* 'he protects', the imperative *ra* 'see!', and *kayfa* 'how' are *yaqih*, *rah*, *kayfah*.

The suffix *-n*, marking the absolute state of nouns and adjectives or the energetic mood of verbs, is also deleted in pause; thus, the pausal forms of *bayt-u-n* 'a house [house-Nom-Abs]' is *bayt*. However, for words ending in *a-n*, the *-n* is deleted, but the *-a* (which may be the marker of accusative case or part of the stem) is not dropped but rather lengthened, so the pausal forms of *bayt-a-n* 'a house [house-Acc-Abs]', *fata-n* 'a boy [boy-Abs]' are *baytā*, *fatā*.

Pausal forms are based on the corresponding medial, full forms, and not directly on the abstract underlying form. For example, *fatan* 'boy' is derived from an underlying form /fatay-u-n/ [boy-Nom-Abs], via an intermediate stage *fatā-n*; if the pausal form were derived directly from /fatay-u-n/, deletion of the final *-u-n* would yield the incorrect form **fatay* (which is, however, attested in certain ancient dialects, cf. Rabin 1951:116). Rather, the pausal form must be derived from the full form *fata-n* (or an intermediate form *fatā-n*), yielding *fatā*. Similarly, in both the indicative and jussive forms of the verb 'he stands': *yaqūmu* and *yaqum*, the stem vowel is underlyingly long (indicative /yaqūm-u/, jussive /yaqūm/), but in the jussive the /ū/ becomes *u* by the general rule that shortens vowels in closed syllables. The difference of vowel length between the indicative *yaqūmu* and the jussive *yaqum* is maintained in their pausal forms, which are respectively *yaqūm* and *yaqum*. The vowel in the pausal indicative *yaqūm* does not shorten. A special case concerns words like *qāḍi-n* [judge-Nom/Gen-Abs], which is derived from /qāḍī-n/ by the same vowel-shortening rule. In pause, where the *-n* is deleted, the word may have the form *qāḍī*, *qāḍ*, or *qāḍi* (Carter 1990). Two of these are problematic: *qāḍ* is not acknowledged by the foremost grammarian, Sībawayhi, and probably is not used in oral Modern Standard Arabic; and *qāḍi* ends in a short vowel, which is unexpected in pausal position (where in any case it is not phonemically distinct from a long *ī*).

A special rule applies to the suffix *-at-*, which marks several different morphosyntactic features on nouns and adjectives, most often feminine gender but also some masculines and plurals. Regardless of function, *-at-* has the pausal form *-ah*, so for the full forms *muʿallim-at-u-n* [teacher-fs-Nom-Abs], *xalīf-at-a-n* [caliph-*at*-Acc-Abs], *al-ḥarāmiyy-at-u* [Art-thieves-at-Nom], the corresponding pausal forms are *muʿallimah*, *xalīfah*, *al-ḥarāmiyyah*. (This does not apply to the suffix *-at* which marks the 3rd person feminine singular on verbs; *katab-at* 'she wrote' is unchanged in pause.) In many dialects, the suffix is *-a* rather than *-ah*, and for this reason the suffix is often represented in conventional transcriptions simply as *-a*. That the suffix was *-ah* in Old Arabic is clear from the facts that some modern dialects preserve the *h* and that in classical poetry it rhymes with stems ending in *ah*, and not with final *ā*. One might suppose that the formation of this pausal *-ah* from *-at-* is a two-step process, first deleting the *-t* specifically in this suffix and then epenthesizing *-h* after the final short vowel by the general process. This might well have been the historical sequence of events, but this does not account for *ḥayāh*, the pausal form of *ḥayātun (ḥaya-at-u-n)* ‘life’, where deletion of the final *t-u-n* would leave *ḥayā*, which does not end in a short vowel and so would not get an epenthetic *h*. In Modern Standard Arabic, there is an alternative pausal form *ḥayāt*, a back-formation from suffixed forms like *ḥayāt-ī* ‘my life’.

Arabic orthography does not normally indicate the difference between pausal and full forms. The spelling is based on the pausal forms rather than full forms, and a word is spelled identically, whether in medial position or in pause. To be precise, the basic spelling, composed of letters of the alphabet, represents the pausal form, even in medial position, while the optional diacritics that augment the basic spelling with additional phonological information, including short vowels, represent the full form, even at the end of a sentence. This practice brought about the creation of two orthographic features that exist specifically to represent the full-pausal alternation. One is the representation of the absolute-state suffix *-n*, which is deleted in pausal forms. Rather than being written with the letter <N> (*nūn*), it is indicated by doubling the diacritical sign that represents the short vowel preceding it. For example, *dār-i-n* [house-Gen-Abs] is written دَارٍ <DaARii> (transliterating Arabic letters with roman capitals, and optional diacritics with lowercase letters); and this spelling indicates a full pronunciation *dārin* and at the same time a pausal pronunciation *dār*. The sequence *-a-n*, which becomes *-ā* in pause, is written with the letter *ʾalif* that normally represents the sound *ā*, so دَارًا <DaARaaA> or دَارًا <DaArAaa> ‘house’ (accusative, absolute) represents the full form *dāran* and the pausal form *dārā*. The second orthographic feature specific to a pausal phenomenon is the representation of the suffix *-at-*, which is *-ah* in pause. This is spelled with the symbol ة, known as *tāʾ marbūṭa* ‘tied T’, which is a hybrid of two letters: it has the shape of ه <H> but the dots of ت <T>. Thus, a spelling like مَدِينَةٌ <MaDiYNaH^{t}uu> ‘city [nom., absolute]’ (representing the dots of <T> with a superscript) simultaneously indicates the full pronunciation *madīnatun* and the pausal pronunciation *madīnah*. When *-at-* is followed by a suffix that contains a long vowel or a consonant other than the suffix *-n*, so that the *-at-* is not final and cannot change to *-ah* in pause, as in the word *madīnatī* ‘my city’, the *tāʾ marbūṭa* is replaced by a regular <T>: مَدِينَتِي <MaDiYNaTiY>.

The indigenous medieval grammarians describe several other phonetic features of Old Arabic in connection with the pausal phenomena (cf. Owens 2006:21–23, 230–234). In some ancient dialects or recitation traditions, a final short vowel might be neither deleted nor lengthened but rather shortened (this is referred to as *rawm*), or in place of a final *u* the lips might be inaudibly rounded after the preceding consonant (*ʾišmām*), and perhaps something analogous could occur with *i*. A final consonant might be lengthened (*taḍʿīf* ‘doubling’), so that *ʾaḥmad* (the name Aḥmad) could sound like *ʾaḥmadd* in pause, and this is quite audible in some modern dialects. If deletion of a final vowel would leave a word ending in a consonant cluster, metathesis (*naql*) might take place instead, so that *bakr-u-n* [Bakr-Nom-Abs] would become *bakur* (Rabin 1951:39).

In Modern Standard Arabic, speakers follow the same three basic rules: deletion of final short vowels, of the suffix *-n*, and of the *t* of the suffix *-at-* (usually pronouncing this as *-a* rather than *-ah*). Vowel reduction and metathesis are obsolete. However, the sequence *-a-n*, which in Classical Arabic should be *-aa* in pause, is more often retained as *-an* when reading, and deleted in spontaneous speech.

All the pausal phenomena that occur in prose also appear in poetry, but in poetry there is another possibility, which in fact occurs more frequently: a final short vowel may either be deleted or lengthened, as required for the → rhyme. Although *-i* and *-u* rhyme with *-ī*, *-ū* and presumably they were to be pronounced as long, they are often written as short, but final *-a* in such cases is always written long. Thus, *ʾanti* might rhyme with *bintī*, and be pronounced *ʾantī*, although both *ʾanti* and *bintī* might be written with a long or short final vowel. A similar option exists for the suffix *-at-*: a word like *madīn-at-u-n* may appear in pause as either *madīnah* or *madīnatū* (but not as *madīnatun*).

3. WHERE DO PAUSAL FORMS OCCUR?

What kinds of actual pauses, or the ends of what kinds of phrases, trigger the appearance of pausal forms? Because the technical term in

Arabic is *waqf* 'stopping, standing', it is usually assumed that pausal forms should appear before actual pauses. However, in speech, intonation contours signaling the end of a clause or other major syntactic constituent often are followed without delay by further sound, and the contrary is also true: one may pause to think, swallow, cough, or correct oneself in the middle of a phrase without applying a final intonation contour. Rather than viewing pausal forms as, ideally, an automatic consequence of a following actual pause, or silence, it is more realistic to think of pausal forms as signaling the end of a major syntactic constituent (Fleisch 1990:196–197). What is formalized for Qur'ānic recitation applies to all sorts of utterances: "The types of pauses are characterized by the syntactic and semantic completeness or incompleteness of the preceding phrase and determine whether the reciter is to stop, to continue with what follows, or back up to bridge a break in meaning or syntax" (Nelson 1985:19). It is the syntax that determines the potential for the use of pausal forms as well as for appropriate actual pauses in speech.

For classical Arabic prose we do not know where pausal forms occur, because the spelling does not differentiate between full and pausal forms. For classical poetry, the rhyme makes it clear that words at the end of a line, which always coincides with the end of a major syntactic constituent, are to be read in pausal form, and the meter shows that full forms appear in midline. In the *Qur'ān*, too, rhyme frequently shows where pausal forms are called for. For example, in sura 1, *al-Fātiḥa*, the words *ar-raḥīmi, al-ʿālamīna, ad-dīni, nastaʿīnu, mustaqīma*, and *aḍ-ḍāllīna* nearly rhyme as *ar-raḥīm, al-ʿālamīn, ad-dīn, nastaʿīn, mustaqīm, aḍ-ḍāllīn*. According to the rules of → *tajwīd*, which prescribe the phonetics of Qur'ānic recitation, pausal forms are used at the end of a verse and at certain spots in midverse, which are indicated by special symbols in the standard editions of the *Qur'ān*, but reciters have some discretion as to which of the indicated pausal locations they actually apply. An enlightening discussion of the rules and esthetics is found in Nelson (1985:27–31). Pause is obligatory following "semantic and syntactic independence of what precedes or follows.... [where] the phrase preceding the obligatory pause is self-contained and takes the form of an epigram or summation and most commonly signals the end of a subject matter" (1985:28), while, if there is "syntactic and semantic dependence on what follows, yet it may be a complete phrase... [p]ause... is permitted here, especially if the phrase is the first half of a conditional sentence, or similarly syntactically dependent on the following phrase" (1985:29).

In the oral use of Modern Standard Arabic, the practice is complex and variable (Meiseles 1977; Schulz 1981; Diem 1974:36–37 and passim; Holes 2004:63–68). Since the Standard Arabic system of cases and moods, which are marked for most words by short-vowel suffixes, is not part of any modern vernacular dialect, Arabs must expend significant effort as part of their schooling to learn it, and, like all school subjects, some individuals master it thoroughly and others less so; some can apply it 'on the fly', while others need to think before deciding on the appropriate form. Using pausal forms eliminates the need to make case and mood distinctions in many types of words. In the most formal situations, i.e. the reading of a prepared script by a skilled reader such as a professional newscaster or some public speakers, pausal forms are relatively few and found almost exclusively before actual pauses, i.e. silences or phrase-final intonation contours at the ends of clauses and sentences. At the opposite end of the spectrum, speaking spontaneously, all or nearly all words may be in their pausal forms. Thus, full forms connote formality, and pausal forms in medial position connote informality (Diem 1974:37), and for this reason full forms occasionally occur even before actual pauses. The grammatical function of the word may also influence whether a pausal or a full form is used. Holes (2004:63–68) observes the following tendencies in the retention of full forms in pausal position, in recent news broadcasts: nouns but not adjectives tend to keep the *-i-n* marking the genitive absolute; and the accusative absolute *-a-n* is generally retained in adverbs like *taqrīb-a-n*, may be retained or dropped in masculine nouns and adjectives, but is dropped in feminine nouns after the *-at-*, which becomes *-a(h)*. Extensive transcripts of speech in different registers, showing varying degrees of use of full and pausal forms, are given by Diem (1974).

4. FROM OLD ARABIC TO THE MODERN VERNACULAR DIALECTS

Most scholars accept that in Old Arabic full and pausal forms alternated just as the indigenous grammarians described, and that over the course of time the full forms (and with them the case and mood distinctions, represented predominantly by short final vowels) were lost from ordinary speech, surviving only in the learned use of Classical Arabic, so that the forms that appear in the modern vernaculars are derived from the Old Arabic pausal forms. Diem (1991) shows that the high degree of redundancy of the case and mood marking in Classical Arabic, followed by the loss of the syntactic-semantic categories of case and mood, must have preceded, and been conducive to, the elision of the final short vowels, that is, the generalization of pausal forms to medial position. Just when the full forms dropped out of colloquial use is a matter of dispute, with opinions ranging from a couple of centuries before the time of Muḥammad to a couple of centuries after. The issue of timing has profound implications for the understanding of how natural or artificial the language of early Arabic poetry was, and hence to what extent that poetry should be relied on as a basis for interpreting the *Qurʾān*. As important as the question of timing is, it does not affect the general picture of the history of the language, for it is clear that the short-vowel case endings were features of Proto-Semitic, having cognates in other Semitic languages, and that the pausal forms correspond closely to the forms in the modern dialects; the transition from the former type to the latter in Arabic could not have been abrupt, so there must have been a period of variation.

However, Owens (1998; 2006) has challenged this picture, suggesting that varieties of Arabic with full forms, case, and mood, and varieties without them, coexisted for many centuries, and in fact that such coexistence may have predated the rise of Arabic as a distinct Semitic language. If so, they must have remained in close contact and shared in the many innovations of grammar and vocabulary that characterize Arabic as a whole vis-à-vis other Semitic languages. In this view, Classical Arabic represents the variety that retained case and mood endings, while the modern dialects descend from a form that lacked them, and not from Classical Arabic at all.

There are, however, good reasons to believe that pausal and full forms coexisted within a single variety of Arabic and that such a variety is the ancestor of the modern dialects (Hoberman 1995:162–164). The grammarians, especially Sībawayhi, took pains to describe the language as they observed it in use, and would not have invented as pervasive a phenomenon as the pausal/full alternation; Classical Arabic poetry can be parsed for meter and rhyme only if words in midline are in their full forms but at line-ends are in pausal forms; there is also evidence from non-Classical spellings in the early Islamic period and from a transcription of an Arabic text in Greek letters (Hopkins 1984). Another sort of evidence comes from relics of the alternations that survive in the modern dialects. These are of two kinds.

The first involves the suffix *-at-* (pausal *-ah*). Nouns with this suffix have two forms in modern vernacular Arabic, a form ending in *-t* that appears when suffixed or followed immediately by another noun in a 'construct phrase' (→ *ʾiḍāfa*), as in *sayyārt-i* 'my car' or *sayyārit iṭ-ṭabīb* 'the doctor's car', and the other form, without *-t*, that appears everywhere else: *sayyāra*. The structure of *sayyārit iṭ-ṭabīb* derives from an Old Arabic genitive structure like *sayyārat-u ṭ-ṭabīb-i* 'the doctor's car [car-Nom Art-doctor-Gen]'. When final short vowels were lost, Old Arabic full forms (like *sayyāratu*) were replaced by pausal forms (*sayyārah*) wherever possible, i.e. at the end of every phrase, but the *t* was retained in a construct because the first element of a construct phrase is not a noun phrase but a single noun. Furthermore, the *t* has survived as a fossil in two types of words: as a 'connective *t*' in a few vowel-initial noun plurals after numerals, such as *xams-t-iyyām* 'five days', from Old Arabic *xams-at-u ʾayyām-i-n*, and as *ṭ* in the numerals from 13 through 19, for instance *xamsṭaˁš(ar)* from *xams-at-a ˁašar-a*. These examples are from Syrian Arabic (Cowell 1964:170–171), but the forms are similar in many dialects.

Alternations like *sayyāra* ~ *sayyārit* are found in all modern dialects. The second relic exists in only a few (Blau 1965:187–202; Fischer and Jastrow 1980:120–121; Owens 1998:215–217). In these dialects, an indefinite noun may be linked to a following adjective, prepositional phrase, or relative clause by a suffix consisting of a short vowel plus *n*. This is evidently a relic of the Old Arabic absolute state suffix *-n*, which is deleted in pausal forms (→ nunation). Thus, an Old Arabic structure like *nās-u-n kaṯīr-u-n* 'many people [people-Nom-Abs many-Nom-Abs]', or in pause *nās-u-n kaṯīr*, became *nās-en čeṯīr* [people-Linker many], where the *-n* survives internally but not at the end of the construction. Furthermore, this accounts for the *n* in such relic forms as Baghdadi *šinu* 'what', from *(ʾayyu) šayʾ-i-n huwa* [(which) thing-Gen-Abs it].

The survival of the *t* and *n* medially, but not finally, in these two modern productive alternations (*sayyāra* ~ *sayyārit* and *nās* ~ *nās-en*) and in the fossilized relic forms shows that at some time pausal and full forms coexisted in the same variety of Arabic, in synchronic alternation very much as the grammarians described it.

5. PAUSAL PHENOMENA IN MODERN VERNACULAR DIALECTS

The ends of phrases in modern Arabic dialects, as in all other languages, are marked by slight phonetic modifications of the sounds of words. In the simplest case, there might be a prolongation of some sounds. Thus, in Damascus Arabic, "the end of a phrase is often signaled by drawling out what comes *after* the accent.... With certain kinds of intonation – in questions, for instance – the phrase-end drawl is often exaggerated so that a post-tonic short vowel is as long as or longer than a true *long* vowel in other positions. In the question *kīf ḥālak?* 'how are you?', the last *a* may actually be longer than the *ā* in the preceding syllable" (Cowell 1964:17). In some dialects, the phrase-end effects are greater. In the dialect of the Negev Bedouin, in pause, stressed *a* is followed by a glottal stop (medial *mišá*, pausal *mišáʾ* 'he went'), and stressed long vowels may be followed by [h] (medial *dammí̄*, pausal *dammí̄h* 'my blood'), among other effects (Blanc 1970:119, 122–123). Final *ī* and *ū* become diphthongs in some dialects of Maltese and Lebanese and Palestinian Arabic (Borg 1977). In parts of Egypt (including Cairo until the end of the 19th century), final *a* is replaced by *e* in pause (Blanc 1973–1974). Some of these effects are strikingly reminiscent of those described for Classical Arabic.

BIBLIOGRAPHICAL REFERENCES

Birkeland, Harris. 1940. *Altarabische Pausalformen.* Oslo: Dybwad.

Blanc, Haim. 1970. "The Arabic dialect of the Negev Bedouins". *Proceedings of the Israel Academy of Sciences and Humanities* 4:7.112–150.

——. 1973–1974. "La perte d'une forme pausale dans le parler arabe du Caire". *Mélanges de l'Université Saint-Joseph* 48.375–390.

Blau, Joshua. 1965. *The emergence and linguistic background of Judaeo-Arabic: A study of the origins of Middle Arabic.* Oxford: Oxford University Press.

Borg, Alexander. 1977. "Reflexes of pausal forms in Maltese rural dialects?" Goldenberg (1977:211–225).

Carter, Michael G. 1990. "*Qāḍī, qāḍi, qāḍ*: Which is the odd man out?". *Studies in the history of Arabic grammar*, II: *Proceedings of the 2nd Symposium on the History of Arabic Grammar, Nijmegen, 27 April–1 May 1987*, ed. Kees Versteegh and Michael G. Carter, 73–90. Amsterdam and Philadelphia: J. Benjamins.

Cowell, Mark W. 1964. *A reference grammar of Syrian Arabic (based on the dialect of Damascus).* Washington, D.C.: Georgetown University Press.

Diem, Werner. 1974. *Hochsprache und Dialekt im Arabischen: Untersuchungen zur heutigen arabischen Zweisprachigkeit.* Wiesbaden: F. Steiner.

——. 1991. "Vom Altarabischen zum Neuarabischen: Ein neuer Ansatz". *Semitic studies in honor of Wolf Leslau*, ed. Alan S. Kaye, 297–308. Wiesbaden: O. Harrassowitz.

Fischer, Wolfdietrich and Otto Jastrow. 1980. *Handbuch der arabischen Dialekte.* Wiesbaden: O. Harrassowitz.

Fleisch, Henri. 1990. *Traité de philologie arabe.* I. *Préliminaires, phonétique, morphologie nominal.* 2nd ed. Beirut: Dar El-Machreq.

Goldenberg, Gideon (ed.). 1977. *Israel Oriental studies*, VII. Tel Aviv: Tel Aviv University.

Hoberman, Robert D. 1995. "Subtractive morphology and morpheme identity in Arabic pausal forms". *Yearbook of morphology, 1995*, ed. Geert Booij and Jaap van Marle, 161–174. Dordrecht: Kluwer Academic Publishers.

Holes, Clive. 2004. *Modern Arabic: Structures, functions, and varieties.* Washington, D.C.: Georgetown University Press.

Hopkins, Simon. 1984. *Studies in the grammar of Early Arabic.* Oxford: Oxford University Press.

Meiseles, Gustav. 1977. "Restitution of 'word-endings' in Modern Literary Arabic". Goldenberg (1977:173–195).

Nelson, Kristina. 1985. *The art of reciting the* Qurʾan. Austin: University of Texas Press. (Repr., Cairo: American University in Cairo Press, 2001.)

Owens, Jonathan. 1998. "Case and Proto-Arabic". *Bulletin of the School of Oriental and African Studies* 61.51–73, 215–227.

——. 2006. *A linguistic history of Arabic.* Oxford: Oxford University Press.

Rabin, Chaim. 1951. *Ancient West-Arabian.* London: Taylor's Foreign Press.

Retsö, Jan. 1994. "The treatment of final syllables in the classical Arabic metres: The linguistic background". *Arabic prosody and its applications in Muslim poetry,* ed. Lars Johanson and Bo Utas, 99–106. Uppsala: Swedish Research Institute in Istanbul.

Roman, André. 1982. *Etude de la phonologie et de la morphologie de la koinè arabe.* Marseille: Université de Provence.
Schulz, David E. 1981. *Diglossia and variation in Formal Spoken Arabic in Egypt.* Ph.D. thesis, University of Wisconsin-Madison.
Wright, W. 1898. *A grammar of the Arabic language*, II. 3rd ed. Cambridge: Cambridge University Press. (Repr., 1971.)

Robert D. Hoberman (Stony Brook University)

Perfect → Tense

Performatives

Performatives are utterances that are unambiguous with regard to their illocutionary force (Austin 1962, 1979). For example, the performative utterance *I hereby promise that I will come at noon* can only be interpreted as a promise, whereas the nonperformative statement *I will come at noon* can be intended and/or interpreted as a promise, a threat, a prediction, or a warning, among other possible illocutionary forces. A performative utterance is issued to perform only a particular type of communicative act, and it has no other interpretations regardless of the nature of the context.

The motivation for the distinction between performatives and statements is that the former constitute events rather than descriptions of events or states of affairs (Searle 1969, 1971). In other words, the act of uttering a performative is itself the action purported by the speaker. For example, the performative utterance in (1), when issued by a lawyer in the context of a court session, amounts to registering an objection rather than reporting or describing an event of objecting taking place at speech time.

(1) *ʾana* *ʾa-ʿtariḍ*
I 1s-object
'I object!'

The major difference between performative and nonperformative utterances is that performatives do not have truth conditions (Stampe 1975); they cannot be judged as true or false. Rather, they have contextually and culturally specified felicity conditions that must be met for a performative to take effect. For example, the performative utterance in (2a) presupposes that the speaker is the bride in a wedding ceremony and that she satisfies the cultural and legal requirements that give her the institutional capacity to issue a marriage performative, e.g., she is above a certain age, sober, and willing.

(2a) *zawwaj-tu-ka* *nafs-ī*
married-1s-you.2ms self-my
'I married you myself' = 'I hereby take you as a husband'

The performative utterance in (2a), and hence the marriage, is void if any of the felicity conditions are violated. For instance, if the speaker does not willfully intend to get married, say, because she is an actress performing a scene in a play, or if the performative is issued under duress, the act of marrying is not accomplished.

Certain performatives require uptakes. These are other performative utterances where the speaker acknowledges and validates a preceding performative (Lyon 1977). For example, the performative in (2a) by itself is not enough to institute a marriage. Rather, the other participant in the ceremony has to utter the uptake in (2b), which amounts to accepting and performing the marriage ritual.

(2b) *qabil-t-u* *z-zawāj* *min-ki...*
agreed-1s the-marriage from-you.2fs
'I accepted marrying you' = 'I hereby accept to marry you'

Arabic devices indicating illocutionary force, which are linguistic forms or constructions that ensure the successful recognition of the speaker's motivations for issuing a performative, fall into three types: explicit performative verbs, performative particles, and certain frozen formulaic expressions. Explicit performative verbs definitively specify the action purported in producing an utterance. Searle (1979) classifies explicit performative verbs into five categories representing the basic types of actions accomplished through speech, viz. representatives, directives, commissives, expressives, and declarations.

The pragmatic function of a representative performative is to commit the speaker to the truthfulness of the propositional content of the clause embedded under an explicit performative verb (Levinson 1983). In (3a), for example, the writer asserts that it is the case that he/she received the mentioned sum, and in (3b) the speaker affirms his/her belief in the truthfulness of the embedded clause proclaiming the oneness of God. Other verbs that can be used as explicit representative performatives include *ʾuqsimu* 'I swear', *ʾuʾakkidu* 'I assure', and *ʾaʿtarifu* 'I admit'.

(3a)	*ʾu-qirr-u*	*ʾana*	*al-muwaqqiʿ*
	1s-affirm-Ind	I	the-signing.ActPart
	ʾadnā-h	*ʾan-ni*	*tasallamt-u*
	under-it	that-I	received-1s
	miʾa	*junayh*	
	hundred	pound	

'I, the undersigned, affirm that I received the sum of one hundred pounds'

(3b)	*ʾa-šhad-u*	*ʾanna*	*lā*
	1s-testify-Ind	that	no
	ʾilāha	*ʾillā*	*llāh*
	god	except	Allah

'I testify that there is no god but Allah'

Directive performatives constitute attempts by the speaker to get the addressee to do something. In (4a), the speaker tries to get the addressee to provide assistance, and in (4b), the speaker attempts to receive redemption by directly asking for it.

(4a)	*ʾa-rjū-ka*	*ʾan*
	1s-beseech.Ind-you.2ms	that
	tu-sāʿid-a-nī	
	2ms-help.Subj-me	

'I beseech you to help me'

(4b)	*ʾallāh-umma*	*ʾa-sʾal-u-ka*
	God-Voc	1s-ask-Ind-you.2ms
	l-ʿafw	
	the-forgiveness	

'O God! I ask your forgiveness'

Directive performatives vary in terms of the power relations they presuppose. For example, demanding and ordering require the speaker to have social or institutional power sufficient to legitimize attempting to control the actions of others. On the other hand, beseeching and seeking forgiveness or permission require the speaker to assume that he/she is at a lower status than the addressee and that the addressee is capable of fulfilling the directives. Other explicit directive verbs include *ʾāmuru-ka* 'I order you', *ʾastaġfiru-ka* 'I request your forgiveness', and *ʾastaʾḏinu-ka* 'I request your permission'.

By uttering commissive performatives, speakers commit themselves to the completion of future actions that they will perform or cause. For example, in (5a) and (5b), the speakers hold themselves responsible for the achievement of the future events of increasing the salaries and paying off the debts.

(5a)	*ʾa-ʿid-u-kum*	*bi-ziyād-at-i*
	1s-promise-Ind-you.2mp	with-increasing-f-Gen
	r-rawātib	
	the-salaries	

'I promise to increase the salaries'

(5b)	*ʾa-taʿahhad-u*	*ʾan*
	1s-pledge-Ind	that
	ʾadfaʿ-a	*diyūn-ī*
	1s-pay-Subj	debts-my

'I pledge to pay my debts'

For a commissive performative to be felicitous, the speaker has to be willing and able to perform the expected action directly, or indirectly by getting others to execute it. Other explicit commissive performatives include *ʾaḍmanu* 'I guarantee' and *ʾurāhinu-ka* 'I bet you'.

Expressive performatives emote psychological states of mind resulting from preceding events. For example, the utterance in (6a) constitutes expressing gratitude and that in (6b) is used to register an apology. Other expressive performatives include *ʾuraḥḥibu bikum* 'I welcome you', *ʾuhanniʾu-ka* 'I congratulate you', and *ʾuḥayyī-kum* 'I salute you'.

(6a)	*ʾa-škur-u-ki*
	1s-thank-Ind-you.2fs

'I thank you'

(6b)	*ʾa-ʿtaḏir-u*	*ʿam-mā*	*faʿal-tu*
	1s-apologize-Ind	for-what	do-1s

'I apologize for what I have done'

Finally, performative declarations bring about immediate changes in the institutional status of the individuals involved in the speech context or its character. For example, uttering the performative in (7a) results in the addressee's being a nominee for chairing the committee, whereas uttering (7b) results in the auction's being open.

(7a) *ʾu-raššiḥ-u-ka* *raʾīs-an*
1s-nominate-Ind-you.2ms chairperson-Acc
li-l-lajn-a
for-the-committee-f
‘I nominate you to be the committee chairperson’

(7b) *ʾu-ʿlin-u* *fatḥ-a*
1s-announce-Ind opening-Acc
l-mazād
the-auction
‘I announce the auction open’

Declarations require the speaker to have particular institutional qualifications. For example, only an official in a position to nominate a committee chairperson can effectively perform (7a), and only an auctioneer on duty can open the auction, as in (7b). Other verbs that can be used in performative declarations include *ʾattahimu* ‘I accuse’ and *ʾaxtatimu* ‘I close/conclude’.

Although many performative utterances involve using explicit performative verbs, performativity is a pragmatic property of utterances, not of verbs. There are performative utterances that involve nonverbal predicates such as active and passive participles, as in the Egyptian Arabic examples in (8a) and (8b), respectively.

(8a) *ʾinti* *ṭāliʾ*
you.fs divorced.PassPart
‘You are divorced!’

(8b) *ʾinta* *marfūd*
you.ms fired.PassPart
‘You are fired!’

Issuing a performative with a participial predicate accomplishes the action denoted by the verb from which the participle is derived, such as divorcing in (8a) and firing in (8b). Note that the sentences *ʾaṭallaʾ-ik* ‘I divorce you’ and *ʾarfid-ak* ‘I fire you’ do not have performative effects. Moreover, some idiomatic expressions are conventionally associated with particular speech acts, such as the Egyptian Arabic *ḥissak ʿēnak* ‘your voice, your eye’, which exclusively signals threats. Besides, there are many Standard Arabic verbs that denote communicative acts but cannot be used as explicit performatives, including:

ḥawlaqa	‘to say *lā ḥawla wa-lā quwwata ʾillā billāh*’
basmala	‘to say *bismi llāhi r-raḥmāni r-raḥīm*’
kabbara	‘to say *allāhu ʾakbar*’
hallala	‘to say *lā ʾilāha ʾillā llāh*’
šammata	‘to say *yarḥamukum allāh*’

These verbs are not performative, because, for example, a person who says *ʾubasmilu* does not perform the act of naming God.

Arabic performatives vary significantly with regard to their linguistic structure, as they allow both 1st and 3rd person subjects, as in (9a) and (9b).

(9a) *ʾa-nṣaḥ-u-ka* *bi-l-ʾiqlāʿ*
1s-advise-Ind-you.2ms with-the-quitting
ʿan *at-tadxīn*
from the-smoking
‘I advise you to quit smoking’

(9b) *tu-ʿlin-u* *šarikat* *miṣr*
3fs-announce-Ind company Egypt
li-ṭ-ṭayarān *ʿan*
for-the-aviation about
qiyām *riḥlati-hā* *ʾila* *rōma*
departure flight-its to Rome
‘Egypt Air announces the departure of its flight to Rome’

Performatives marked for the 3rd person are used when the speaker acts on behalf of the intended agent of the communicative act, presupposing that the speaker is not the source of the performative and hence not responsible for its consequences.

Arabic performative utterances allow verbs marked for the active voice, as in all preceding examples, as well as the passive voice as in (10).

(10) *yu-stadʿā* *l-maḏkūr*
3ms-summon.Pass the-mentioned
li-t-taḥqīq *maʿa-h*
for-the-questioning with-him
‘The mentioned individual is to be summoned for questioning’

Performatives that involve the impersonal passive are usually directives where it is not the speaker or the addressee who will carry out the requested action. Rather, the speaker issues an order or a recommendation for others to act upon, usually for official rather than personal motivations. For example, by uttering (10), the speaker, if a police officer, officially issues a subpoena even though he/she is not the one who will actually carry out the action of summoning that individual.

The fact that issuing a performative is itself an event rather than a description of an event motivates analyzing performatives as tenseless constructions. There are two observations supporting this argument. First, Egyptian Arabic verbs that are used as explicit performatives do not allow the imperfective prefix *bi-*, which marks the progressive and habitual aspects, as in (11a) and (11b).

(11a)

ʾa-ʾaddim		*lu-kum*
1s-introduce		to-you.2mp
ʿali	*ʾibn*	*ʿamm-i*
Ali	son	uncle-my

‘I introduce to you my cousin Ali’

(11b)

ʾa-ḥlif-la-k		*ʾinn-i*
1s-swear-to-you.2ms		that-I
ma	*saraʾ-t*	*il-filūs*
Neg	stole-1s	the-money

‘I swear to you that I didn’t steal the money’

If these verbs describe events that are taking place at speech time, it is expected that they would be grammatical only if the imperfective *bi-* forms were used. However, the progressive/habitual marker inhibits performative interpretations.

The other observation is that in Standard Arabic perfect verb forms are restricted to declarations (Khalil and McCarus 1999). For example, the utterance (12) constitutes sentencing the defendant to a year in prison, provided that all felicity conditions are satisfied. These conditions include the requirement that the speaker be a judge on duty in a court session. However, if the same utterance is issued by a reporter as a newspaper headline, i.e. if the felicity conditions are not satisfied, the utterance is a mere description of a past event of sentencing.

(12)

ḥakam-at	*al-maḥkam-at-u*	*ʿalā*
sentenced-3fs	the-court-fem.-Gen	on
l-muttaham	*bi-sana*	*sijn*
the-defendant	with-year	prison

‘The court sentenced the defendant to a year in prison’

Another way of unambiguously signaling the communicative intent of an utterance is to use a performative particle. These particles are conventionally associated with particular speech acts such as *wā* (*wāw an-nadba*), which is used for lamenting, the verbal suffix *-nna* (*nūn at-tawkīd*), which is used to mark representative performatives (→ energicus), and instigative *hallā* and *ʿalā* (*ʾadawāt at-taḥḍīḍ*), among many other such particles.

BIBLIOGRAPHICAL REFERENCES

Austin, John. 1962. *How to do things with words*. Cambridge, Mass.: Harvard University Press.

——. 1979. *Philosophical papers*. Oxford: Oxford University Press.

Khalil, Aziz and Ernest McCarus. 1999. “Arabic performative verbs”. *Zeitschrift für Arabische Linguistik* 36.7–22.

Levinson, Stephen. 1983. *Pragmatics*. Cambridge: Cambridge University Press.

Lyon, John. 1977. *Semantics*. Cambridge: Cambridge University Press.

Searle, John. 1969. *Speech acts: An essay in the philosophy of language*. Cambridge: Cambridge University Press.

——. 1971. “What is a speech act?”. *The philosophy of language*, ed. John Searle, 39–53. Oxford: Oxford University Press.

——. 1979. *Expression and meaning*. Cambridge: Cambridge University Press.

Stampe, Dennis. 1975. “Meaning and truth in the theory of speech acts”. *Syntax and semantics*, III, ed. Peter Cole and Jerry Morgan, 1–39. New York: Academic Press.

MUSTAFA MUGHAZY
(Western Michigan University)

Persia → Iran

Persian

This entry includes an account of the incorporation of Arabic vocabulary into literary New Persian in general, and some discussion of the Arabic vocabulary in Standard Persian, the written and spoken language of educated

Iranians today. The topic is pertinent also for other Iranian, Turkic, and Indic literary languages of the region, such as Chaghatay and → Ottoman, Pashto, → Urdu, and Bangla (→ Bengali), which received their Arabic vocabulary through the medium of Classical Persian.

1. History and evolution

With the Arab conquest of Iran in the 7th century and the conversion of a majority of the population to Islam, Arabic came to exert a profound influence on the Persian language. The form of Persian affected was not literary Middle Persian (*pårsīk*, commonly called Pahlavi), which was identified with Zoroastrian religious literature and written in a form of the Aramaic script, but rather the related vernacular of the court milieu of Seleucia-Ctesiphon (Madāʾin) and other parts of the Persian Empire, called *Dari* by Ibn al-Muqaffaʿ (see Lazard 1990). By the middle of the 9th century, this spoken form of Persian had become a poetical koine, presented in Arabic script; within the following century, it systematically incorporated Arabic loanwords and became established at the quasi-independent Samanid court of Bukhara as a literary language, New Persian. Arabic works such as aṭ-Ṭabarī's *Tārīx* and *Tafsīr* were translated into New Persian, and a tradition of original prose *ʾadab* developed.

Arabic script was quite well adapted to transcribing Persian. Iranian bureaucrats continued in office even after the language of administration was changed to Arabic; and from the outset many Iranians became bilingual in the two tongues. Persian preserves traces of this vernacular component in a few early Arabic borrowings morphophonologically assimilated to Persian, and which have survived subsequent orthographic normalization: e.g. *mosalmån* 'Muslim' (by metathesis and, probably, modification of a plural, < Arabic *muslim*[-*ūn*]); the onomastic *bu* (Arabic *ʾabū* 'father [of]'); *mir* (Arabic *ʾamīr*) 'commander, prince' and its compounds such as *mir-åb* 'official in charge of water distribution', *mir-åxor* 'head groom', *mir-zå* 'born of a prince', which parallel the apheresis in Dari reflexes of Pahlavi words at this time; cf. *(a)yår* 'friend', *(a)bå* 'with', *(a)nåhid* 'Anahita', etc. Thereafter, the bulk of Arabic loanwords entered Persian as *mots savants* in the writings of bilingual poets and scholars, mostly from the 10th to the 12th centuries, and trickled down into spoken usage (Telegdi 1973).

2. Phonology and orthography

With very few exceptions, Arabic loanwords in Persian are written exactly as in Arabic. A number of Arabic characters represent consonants alien to Persian, which are therefore assimilated to the closest Persian phonemes: *s*, *ṯ*, and *ṣ* are all realized as /s/; *z*, *d*, *ḏ̣*, and *ẓ* as /z/; *t* and *ṭ* as /t/; *h* and *ḥ* as /h/. Sounds of Persian not found in Arabic (*p*, *č*, *ž*, *g*) were written with letters representing similar sounds (*b*, *j*, *z*, *k*), distinguished by diacritics (→ Arabic alphabet for other languages).

The glottal stop of Arabic (written as *hamza*) is pronounced after a consonant, but generally it is realized before a consonant as a prolongation of the (short) vowel, and between vowels as a glide: /sowål/ 'question' (Arabic *suʾāl*). The peculiarly Arabic sound of *ʿayn* is ignored in initial (and, colloquially, in final) position; it is realized between vowels as a glide or a glottal stop, and before a consonant as a prolongation of the vowel: /ba:d/ for *baʿd* 'after'; in Persian of Afghanistan, the quality of the vowel is also changed, as /bå:d/. The sounds of *qāf* and *ġayn* are pronounced alike in Standard Persian, as a voiced velar affricate or fricative: Arabic *waqt* > Persian /vaġt/ > /vaxt/ (after devoicing before a voiceless consonant). Thus, Arabic *ġiḏāʾ* 'food' and *qaḍāʾ* 'judgment, destiny', as loanwords, become homophones: /ġazå/. Arabic /w/ is realized as labiodental /v/ in Standard Persian. The other Arabic consonants have close Persian counterparts.

The Standard Persian articulation of vowels and diphthongs in Arabic loanwords is shown in Table 1. The equivalence of quantity between Arabic and Persian short and long vowels is fully preserved only in poetry, where the Arabic prosodic system of *ʿarūḍ* has been adapted to Persian prosody. In spoken Standard Persian the phonemic distinction is one of quality; but there is also a division between 'stable' vowels (corresponding to the Arabic long vowels; first three cells) and 'unstable' vowels (corresponding to the Arabic short vowels; next three cells). The stable vowels do not change appreciably in quality or length in response to the phonetic environment, whereas the unstable vowels may shift allophonically in quality and become shorter when unstressed, or assimilate to adjacent vowels or semivowels (see the diphthongs in Table 1). Consistent in current Standard Persian is the raising of word-final /a/ > /e/, which affects all Arabic loans in the feminine ending that do not have final *-t* in Persian (see Sec. 7): *qabḍa* > /ġabzé/ 'handful; handle'. The low back vowel /å/ is generally unrounded (though more rounded in Eastern dialects). Stress in Persian nominals (to which class belong most polysyllabic loans from Arabic) is word-final.

Table 1. Correspondence of Arabic and Persian vowels

	Long vowels			Short vowels			Diphthongs	
Arabic	ā	ī	ū	a	i	u	ay	aw
Persian	å	i	u	a	e	o	ey	ow

Stable vowel Unstable vowel ________

Arabic short/Persian unstable vowels in loanwords are also subject to assimilation, dissimilation, and syncope in certain environments, and to analogical changes. Thus, Arabic *nahār* 'noon' > Persian *nåhår* 'lunch' (one of very few such changes to be reflected in the spelling); *ṣadā* > /sedå/ 'sound' (/a/ is raised in proximity to a sibilant); *ḥaraka(t)* > /harekat/ 'movement', but *šarika(t)* > /šerkat/ 'partnership'. The vowel shift *musāfara(t)* > /mosåferat/ 'journey' (/a/ is raised in an open penultimate syllable), which applies to the whole form class of about 150 *mufāʿala* loans in all dialects of Persian, may result from morphological analogy rather than phonetic law, i.e. by contamination with the corresponding participial loanword, as in Persian *mosåfer* 'passenger', *mobårez* 'fighter', *monåseb* 'suitable', etc. This tendency to harmonize transparent cognates on familiar (Turco-Persian) principles of suffixation instead of the alien nonsegmental morphology of Arabic can clearly be seen in the Persian pronunciation /šojåat/ 'bravery' (Arabic *šajāʿa*[*t*]), by analogy with the borrowed adjective *šojåʿ* 'brave'.

Arabic posed a challenge to the phonotactics of Persian by introducing a number of alien word-final consonant clusters, as in *rabṭ, fiqh, ʿadl*. In eastern dialects of Persian (and most Turkic and Indic languages), the difficulty of articulation is solved by inserting an epenthetic vowel (generally schwa), as /húkᵊm/ for *ḥukm* 'decree' or /qábᵊl/ for *qabl* 'before'; Standard Persian prefers to delete one of the two consonants, as /vaxt/ or /vax/ for *waqt* 'time', /so:b/ for *ṣobḥ* 'morning' (with compensatory vowel lengthening).

3. LEXICAL STATISTICS

A dictionary-based sample yields an inventory of approximately eight thousand Arabic loanwords in current use (Rāzī 1987), or about 40 percent of an everyday literary vocabulary of twenty thousand words (excluding compounds and derivatives). Corpus-based inventories, and frequency of use of Arabic vocabulary per text, vary with date of composition, stylistic register, individual author, and topic of discourse. Thus, a sample from the versified national epic, the *Šåhnåme* of Ferdowsi (completed ca. 400/1010), yields an Arabic vocabulary of only 8.8 percent and a frequency of 2.4 percent (Moïnfar 1970:61–66); the eulogies of Ferdowsi's younger contemporary Onsori (ʿUnṣurī), modeled on the Arabic *qaṣīda*, yield approximately 32 percent and 17 percent, respectively (Osmanov 1970). In a sample of Sufi verse from about the 14th century, these proportions rise to 51.8 percent and 24.3 percent respectively (Utas 1977:75–102); and in prose fiction from the 1950s, they drop to 46.5 percent and 19.7 percent respectively (Koppe 1959–1960:90–93; see also Perry 1991a:203–205).

Since Arabic lexical morphology is highly systematic, certain prefixed and suffixed formatives of Arabic are salient in the Persian dictionary, as are certain assonant word patterns. Thus the letter *mīm*, the initial of three highly productive Arabic prefixes, accounts for about eighteen hundred loanwords, or almost a quarter of the Arabic inventory in modern Persian (including, e.g., seventy words of the pattern *mafʿala*); the overall inventory of initial *m-* is inflated to twice the size of the average letter.

4. LOANWORD CLASSES

Arabic loanwords in Persian are almost entirely nominal in origin: nouns, deverbal nominals (action nouns and participles), adjectives, and adverbs. With the exception of *tanwīn* adverbs (see Sec. 5) and feminine-ending loans (see Sec. 6), Arabic nominals are inducted into Persian in their bare stem form, without inflection or other modification. To this form may be juxtaposed all appropriate Persian affixes and enclitics: ***ketåb**-hå-i* 'some books'; *bi-**vafå**-i* 'disloyalty'. Verbs are not borrowed in inflected forms, but Arabic action nouns (*maṣdar*) and other deverbal nominals may form Persian verbs, in one of two ways:

i. Synthetically, by suffixation of the Persian past stem and infinitive, as *fahm-idan* 'to understand' (the original way of forming denominal verbs in Persian, e.g. *nåm-idan* 'to name'); this stratagem was favored in earlier Classical Persian.
ii. Analytically, by combining with a dummy auxiliary such as *kardan* 'to do, make' or *šodan* 'to become, be done': *ġat kardan* 'to sever, interrupt' (*qaṭʿ* 'cutting', action noun), *måneʿ šodan* 'to hinder, prevent' (*māni ʿ* 'hindering [part.]'). This way is preferred in Modern Persian. The meaning may be refined by use of an auxiliary with some semantic weight: *ġabul kardan* 'to accept, receive', *ġabul dåštan* 'to agree, concur; to take for granted' (Arabic *qabūl* 'acceptance'; Persian *dåštan* 'to have, hold'; here, 'to hold to be, consider as').

The incorporation of *maṣdar*s was recorded, regulated, and probably accelerated by the compilation of Arabic-Persian dictionaries devoted to them, called *maṣādir*, produced in Iran and India from the 11th to the 15th centuries. Other Arabic nominals were glossed in dictionaries called *ʾasmāʾ* or *ʾasāmī* 'nouns' (Perry 1993).

Hundreds of action nouns and participles from ten of the thematic extensions of the Arabic verb have been borrowed into Persian and commonly form compound verbs of the above type; these are the action nouns of Form II, both types of Form III, and those of IV–VIII and X, and the *maṣdar mīmī* (for a convenient overview, largely in tabular form, see Elwell-Sutton 1963:157–167). Thus, from the triliteral root *ṣ-l-ḥ* '(being) right, fit, proper, harmonious', the following Arabic derivatives appear in Persian, often as verbs or verbal idioms: *solh* 'peace' (Arabic *ṣulḥ*), *salåh* 'honesty, propriety, fitness' (Arabic *ṣalāḥ*), *salåh dånestan* 'to deem appropriate, see fit' (*dånestan* 'to know, acknowledge'), *eslåh kardan* 'to improve, correct, edit; to shave' (Arabic *ʾiṣlāḥ*), *mosålehat* 'reconciliation' (Arabic *muṣālaḥa*), *estelåh* (Arabic *iṣṭilāḥ*) and *mostalah* (pl. *-åt*; Arabic *muṣṭalaḥāt*) '(technical) term, idiom', *maslehat* 'interest, expediency' (Arabic *maṣlaḥa*), *maslehat didan* 'to deem prudent' (*didan* 'to see [as]'), the plural *masåleh* 'benefits, interests' (in Indo-Persian, and hence Hindi-Urdu, > *masāla* 'materials, ingredients, spices'), the adjective *såleh* 'wholesome, beneficial' (Arabic *ṣāliḥ*), the compounds *salåh-kår* 'charitable [good-doer]' and *eslåh-nå-pazir* 'irremediable [reform-not-accepting]'.

There are many other such multiple root-cognates in the Persian lexicon, conditioning educated readers by alliteration to the connection of a particular consonant combination with a certain semantic field, even though they may not know Arabic as such. Other patterns supplying loans are nouns of place, as *madrase* 'school' (place of teaching; cf. the cognate loan *dars* 'lesson'); of instrument, as *mezråb* 'dulcimer hammer' (Arabic *miḍrāb*; cf. *zarbat* 'blow, beat', Arabic *ḍarba*); of occupation, as *raġġås* 'dancer' (Arabic *raqqāṣ*; cf. *raġs*, Arabic *raqṣ* 'dance', from which are derived both *raġs-idan* and *raġs kardan* 'to dance'); several sorts of adjective (*šarīf* 'noble', *faʾål* 'active' < Arabic *faʿʿāl*) and quality nouns from adjectives (*nejåsat* 'impurity', cf. *najes* 'impure'). A few patterns, notably the elative and diminutive, do not normally appear as loanwords except as names (*Akbar, Hoseyn*).

Apart from participles (from 18 Arabic participial patterns, active and passive: see Elwell-Sutton 1963:162–163), the largest class of morphologically salient Arabic adjectives in Persian comprises the derivatives with the → *nisba* or relative suffix *-i* (< *-iyy*^{un}), e.g. *makki* 'Meccan', *šaxsi* 'personal' (Arabic *šaxṣī*). This suffix coincides in form and meaning with the Persian *-i* (< Middle Persian *-īk*), as in *širåzi* 'of Shiraz', *xåki* 'earthen, brown'. Highly productive, this hybrid may be suffixed directly to any class of nouns, including assimilated Arabic loanwords: e.g. *tejårat-i* 'commercial', *qahve-i* 'coffee-colored, dark brown' (where the orthography shows that this is not an Arabic form). In many cases, however, it is not obvious whether an adjective in *-i* represents an integral Arabic borrowing or a Persian derivative. The coincidence also results in homographs such as *dudi* 'smoky, smoked' (Persian *dud* 'smoke' +

-*ī*[*k*]) and *dūdī* 'wormlike, peristaltic' (< Arabic *dūd* 'worm' + *-iyy*^{un}).

5. PSEUDOLOANS AND RÜCKWANDERER

The degree to which not only individual loanwords but also their characteristic patterns entered Persian consciousness is shown in a number of common Persian words coined on Arabic morphological patterns from a native Persian or other lexical base: thus, *kaffåš* 'cobbler' (< Persian *kafš* 'shoe'), *nezåkat* 'daintiness' (< Persian *nåzok* 'dainty'). The *tanwīn* accusative adverb, commonly borrowed into Persian (*rasman* 'officially', *nesbatan* 'relatively', etc.) has remained productive, even forming adverbs from Persian nouns and adjectives: *jånan* 'wholeheartedly', *nåčåran* 'willy-nilly' (*nå-čår* 'having no recourse').

Arabicized forms of Persian words borrowed into Arabic were also accepted back, such as *fehrest* 'list, register', Arabic *fihrist*, originally Middle Persian *pahrist*. The form *Fårs* 'Pars province' is in origin a convention of the Arabic geographers (several of whom were ethnic Persians), but *fårsi* 'Persian (language)' is a gratuitous Arabicization, virtually a blend of Middle Persian *pårsīk* and Arabic *fārisiyy*^{un}; *fil* 'elephant' is a *Rückwanderer* from Persian *pil*, loaned to Arabic and returned with the enhanced prestige of its inclusion in the *Qur'ān*. Some Persian writers in all periods have preferred to use *pårs, pårsi*, and *pil*, thereby making a politico-cultural statement.

6. LOANWORDS WITH THE FEMININE ENDING

The grammatically feminine marker in Arabic is realized phonetically as either /-at/ (in pre-juncture position) or /-a/ (pausal form), according to the contextual syntax of Arabic, but written with a single graph (the *tā' marbūṭa*). This syntactically determined variation in Arabic was irrelevant to Persian, where these loans needed to be lexicalized systematically with or without final *t*: accordingly, some were written with regular final *t* (e.g. *hekmat* 'wisdom, philosphy' < *ḥikma*) and others with nonlinking final *h* to represent the open final syllable /a/ (later /e/; as in *xerġa, xerġe* 'rag; dervish's cloak' < *xirqa*). Historically, many have 'shifted' from the originally preponderant *-at* inventory to *-a*, so that in current Standard Persian there are more than 800 words ending in *-at*, and approximately 640 in *-a* (now realized phonetically as /-e/); this includes some forty pairs of doublets lexicalized with both endings. Analysis of the rationales behind this dichotomy and the associated shift affords some insight into the process of loanword incorporation from Arabic to Persian.

Distribution between *-at* and *-e* in the modern inventory is determined primarily by semantic features, and additionally by stylistic register or historical evolution of the words (Perry 1991a:195–224). Thus, nouns with more abstract and intangible, or less imageable and countable, referents tend to end in *-at: roxsat* 'permission, leave' (Arabic *ruxṣa*), *xošunat* 'asperity, roughness' (Arabic *xušūna*), *mojånebat* 'avoidance, nonintervention' (Arabic *mujānaba*); nouns with more concrete, tangible, imageable, and countable referents (more likely to be pluralized) tend to end in *-e: nosxe* 'text, prescription' (Arabic *nusxa*), *vasiġe* 'bond, security' (Arabic *waṯīqa* 'document'), *mahalle* 'place, neighborhood'. Even *maṣdar* forms ending in *-e* are more likely than those in *-at* to form common compound verbs in Persian and to have evolved count-noun referents: *estefåde kardan* 'to use', *estefåde-hå* 'uses' (Arabic *istifāda*); *ešåre kardan* 'to point out, indicate', *ešåre-hå* 'indications' (Arabic *'išāra*). These processes are even more apparent in the doublets: Arabic *quwwa* > *ġovvat* 'strength, power' (general, intangible), vs. *ġovve* '(military) force, (industrial) energy' (pl. *ġovve-hå, ġovå*); *erådat* 'wish, goodwill', *eråde* 'resolution, edict' (Arabic *'irāda*), etc. (Perry 1995).

In historical perspective, this process can be seen as one of 'exaptation', the adaptation of a redundant feature to a productive new purpose. The loans that were adopted in the form *-a* were often morphologically assimilated with the large class of native substantives in *-a* (such as *dåna* 'grain', *barnåma* 'program', and the active and passive participles), a lexical class which at the time of the Arab conquest had already lost (in spoken Dari) the final velar of Pahlavi (still to be seen in earlier Arabic borrowings from Persian, such as *dānaq* and *barnāmaj*): by analogy with the Persian class, this consonant (which still surfaces in derivatives of Persian words in final *-e*: e.g. *gorosne-gi* 'hunger' < *gorosne* 'hungry') was supplied to many of the Arabic loans before Persian suffixes, as in *xebre-gån* 'experts' (Arabic *xibra* 'expertise'), *bi-saliqe-gi* 'uncouthness' (Arabic *salīqa* 'refinement', also borrowed). This expanding class was thus marked as vernacular and productive.

Conversely, those loans adopted in *-at* stood out as unproductive *Fremdwörter*, since by this time virtually all instances of final *t* in Persian had been voiced (*t* > *d*). In the course of the next several centuries, hundreds of the *-at* class shifted to the *-a* class, some leaving behind traces as doublets in *-at*. In other words, an existing vernacular form in *-a* was acknowledged as useful and incorporated in the written language, either to replace the *-at* reflex or to supplement it (Perry 1991a:189–195). Some of the resulting doublets were mobilized to expand the technical lexicon; e.g., al-Ġazālī uses *ġovvat* to mean 'power, potential', but *ġovva* (pl. *ġovå*; see above) for a particular physiological or mental faculty. In general, the resulting *-e* words are semantically more specialized and/or more firmly established in the vernacular.

The loss of *-t* often corresponds additionally to a change of register, from literary to vernacular: thus Standard Persian *hekåyat* '(literary) anecdote' (< Arabic *ḥikāya*) has remained more of a literary word (in contrast with *ġesse* 'tale, story' (< Arabic *qiṣṣa*), whereas in modern → Tajik and → Turkish it has dropped final *-t* orthographically and functions as an everyday word, respectively *hikoya* and *hikâye* 'tale, story'.

These rules for binary sorting in Persian were transmitted to Turkish, Urdu, etc., together with the Arabic loans that they incorporated via Persian, and were slightly expanded or modified (Perry 1991a:139–188).

Salient among the earliest loanword classes (coined in Arabic during the philosophical-scientific heyday of Islam in the 3rd/9th and 4th/10th centuries) was the *nisba* subset of the feminine substantives, incorporated as *-iyat/ -iya*, e.g. *ensåniyat* 'humanity', *zojåjiya* 'crystalline lens'. Then during the 19th century, a wave of Arabic (and artificial Arabicate) neologisms, many calqued on French and originating in Ottoman Turkish, supplemented the technical and legal-administrative lexicon of Persian; these, too, included a large *nisba*-noun component, such as *melliyat* (Turkish *milliyet*) 'nationalism', *aġalliyat* 'minority' (Arabic *ʾaqalliyya*), *ehzåriye* '(writ of) summons, subpoena' (< Arabic *ʾiḥḍār* 'summons'), *bahriye* 'navy' (< Arabic *baḥr* 'sea'; see Faršidvard 1969:61–63). Some of these coinages, paradoxically, were circulating in Turkish and Persian before they were incorporated into Arabic.

7. CHANGE OF CATEGORY

Semantic change, especially vernacularization, in a loanword may entail a shift of grammatical categories. An early example is *tamiz* 'clean' (adjective, with orthographic modification, < Arabic *tamyīz* 'discernment, distinction'). Several quality nouns of Arabic origin are now used primarily as adjectives in Persian, e.g. *xalvat* 'private, quiet' (Arabic *xalwa* 'seclusion'), *råhat* 'easy, comfortable' (Arabic *rāḥa* 'ease, rest'), *salåmat* 'safe, well' (Arabic *salāma* 'safety'); the change was presumably achieved by way of a reanalysis of the word as predicate (*in råhat nist* 'this is not [my idea of] comfort' ⇒ 'not comfortable'). Most such words may now be used attributively (an exception is *šohre* 'famous; a byword', a doublet of *šohrat* 'fame; surname', both < Arabic *šuhra*); they may usually derive a new quality noun by suffixing *-i*: *salåmati* 'health', etc. Other nouns have become adverbs: *xolåsa* 'gist; in short' (Arabic *xulāṣa*). Others were produced by deletion of the head noun in a phrase: *kolfat* 'maid, cleaning woman' < *zan-e kolfat* 'woman for chores' (Arabic *kulfa* 'inconvenience, chore'); *balad* 'familiar (with), knowledgeable (about)' < ?(*ahl-e*) *balad* 'local, native' (Arabic collocation *ʾahl al-balad* 'belonging to the locale').

8. GRAMMATICAL ELEMENTS

Arabic plurals may sometimes be used instead of Persian plurals: *ketåb-hå* or *kotob* 'books', *moallem-ån/-hå/-in* 'teachers' (Arabic *muʿallimīn*). The choice is usually stylistic, but some plural loans have been lexicalized with a singular meaning (*arbåb* 'landlord; boss'; the singular *rabb* 'Lord' refers only to God); or the choice of plural is lexicalized, each form denoting a part of the semantic range of the singular, e.g. Arabic *ṣāḥib* > *såheb-ån* 'owners', *sahåbe* 'the Companions (of the Prophet)', *ashåb-e X* 'people characterized by X'; Arabic *ḥarf* > *harf-hå* '(spoken) words, utterance', *horuf* 'letters (of the alphabet), written characters'. Arabic broken plurals have occasionally been applied to Persian and other non-Arabic nouns, some still current: *banåder* ‘the lower Gulf littoral’ < Persian *bandar* ‘harbor’.

The few nouns in which the Arabic definite article *al-* is incorporated in Persian function as interjections or adverbs: *al-ʾamån* ‘mercy!’, *al-vedå* ‘farewell’ (Arabic *al-wadāʿ*), *al-ʾån* ‘now’, *al-batte* ‘of course’. Arabic nominal collocations (adverbial and noun phrases), frozen and lexicalized, play a larger role: *bel-ʾaks* ‘vice versa’ (Arabic *bi-l-ʿaks*; also Persianized as *bar aks*). In recent centuries, macaronic collocations such as *hasab al-farmån* ‘in accordance with decree’ (modeled on Arabic *ḥasab al-ʾamr*) were in vogue. Some Persian compounds originated as Arabic collocations: *kotob-xåne* ‘library’, originally in the form of a direct calque on Arabic *dār al-kutub* ‘house of [the] books’, was soon normalized as *ketåb-xåne*, with the modifier as a generic singular (although the original form survives in Turkish). Arabic sentences and verb phrases (interpreted as reduced relative clauses) also serve as adjectives: *lå-yanfakk* ‘inseparable’ (Arabic ‘[it, etc.] is not detached’), or as nouns: *må-jarå* ‘adventure, affair’ (Arabic ‘what transpired’; see Jazayery 1970).

9. SEMANTIC DISTRIBUTION

Studies of the Arabic component of specific semantic and experiential fields in Persian are as yet few and limited (see Asbaghi 1997). Arabic loans in Persian appear to comprise a greater proportion of abstract and intangible referents than of entities and other tangible and countable referents – i.e. more intellectual and high-cultural vocabulary (see Koppe 1959–1960; Perry 1991a: 206–208). In one experiment, comparing a random sample of Arabic loans in four languages, the vocabulary of material culture in Spanish was 52 percent of the Arabic loan inventory, while in Persian the total was 14 percent; the Arabic vocabulary of intellectual life was 8 percent in Spanish, 24 percent in Persian. This result for Persian is consistent with a sophisticated society actively involved in the ideological and intellectual aspects of the superimposed culture through its language (Perry 1991b). On the other hand, it is evident that many Arabic loanwords have (and had) everyday synonyms in Persian (e.g. Arabic *marīḍ*, Persian *bimår* ‘sick, ill’), and were probably borrowed for reasons of prestige or literary variety rather than need.

The field of (Islamic) religion is not (and never was) dominated by Arabic loanwords. Scores of Persian words, from *åxund* ‘cleric’ to *zendiq* ‘heretic’ (the latter in Arabicized form), are Persian, including the everyday terms for God, prophet, prayer, prayer leader, fasting, angel, creation, creator, heaven, hell, soul, sin, to worship, to repent, to forgive, etc. This is not surprising, since the process of conversion depends for its early success on comprehension, achieved by translation into, analogy with, and use of the language of the target population.

10. PURISM AND RELEXIFICATION

The debate over the use of the vernacular in scholarly writing was already lively in the 11th century. Both Ibn Sīnā (d. 428/1037) and al-Ġazālī (d. 505/1111), known for their works in Arabic, wrote their less specialized treatises in Persian and consciously resurrected or invented terminology in Persian; but al-Bīrūnī (d. after 442/1050) argued that Persian should be confined to popular literature and only Arabic was lexically sophisticated enough for scientific writing (Lazard 1975:631). After the Mongol conquest of Baghdad (656/1258), Arabic became a dead language in the Turco-Iranian world, and this source of loan vocabulary soon dried up; Persian, with its existing stock of Arabisms, became a medium of scientific writing in virtually all fields except mathematics.

With the language purism movement of the 1930s–1940s in → Iran, Arabic vocabulary was targeted for replacement by Persian neologisms. Although this reform was not as drastically implemented as in Turkey, many of the more recent technical terms were replaced, and officially sanctioned lexical policy ever since has preferred coinage of Persian terms or toleration of European loanwords (Perry 1985). The Islamic Revolution of 1979 has not fundamentally affected these trends. A few ideologically inspired Arabisms have been introduced, such as *mostazʾaf* (with Persian pl. *-ån* or Arabic pl. *-īn*) ‘dispossessed, underprivileged’ (*mustaḍʿaf*); but vocabulary is still being expanded primarily by appeal to native Persian words and morphs (supplemented in the spoken language by borrowings from English). Arabic is no longer a live lexical source for Persian.

Bibliographical references

Asbaghi, Asya. 1997. *Die semantische Entwicklung arabischer Wörter im Persischen*. Stuttgart: F. Steiner.

Elwell-Sutton, Lawrence P. 1963. *Elementary Persian grammar*. Cambridge: Cambridge University Press.

Faršidvard, Xosrow. 1969. *ʾArabi dar fārsi* [Arabic in Persian]. Tehran: Ahuri.

Jazayery, Mohammad Ali. 1970. "The Arabic element in Persian grammar: A preliminary report". *Iran* 8.115–124.

Koppe, Reiner. 1959–1960. "Statistik und Semantik der arabischen Lehnwörter in der Sprache ʿAlawīs". *Wissenschaftliche Zeitschrift der Humboldt-Universität zu Berlin* 9.585–619.

Lazard, Gilbert. 1975. "The rise of the New Persian language". *Cambridge history of Iran*, IV, 595–632. Cambridge: Cambridge University Press.

——. 1990. "*Pârsi* et *dari*: Nouvelles remarques". (= *Bulletin of the Asia Institute* N.S., 4.) *Aspects of Iranian culture, in honor of Richard Nelson Frye*, ed. O. Sjærvø and A. Sh. Shahbazi, II, 239–242. Ames, Iowa.

Moïnfar, Mohammad Djafar. 1970. *Le vocabulaire arabe dans le Livre des rois de Firdausi*. Wiesbaden: O. Harrassowitz.

Osmanov, M.N. 1970. *Častotnyj slovarʾ Unsuri* [Statistical glossary of ʿUnṣurī]. Moscow: Nauka.

Perry, John R. 1985. "Language reform in Turkey and Iran." *International Journal of Middle Eastern Studies* 17.295–311. (Repr. with corrections, *Men of order: Authoritarian modernization in Turkey and Iran*, ed. Touraj Atabaki and Erik-Jan Zürcher, 238–259. London and New York: I.B. Tauris, 2003.)

——. 1991a. *Form and meaning in Persian vocabulary: The Arabic feminine ending*. Costa Mesa, Calif.: Mazda.

——. 1991b. "Arabic loan vocabulary in Persian, Turkish, Urdu, Hindi, Spanish: Comparative indices". Paper delivered at the 201st Annual Meeting of the American Oriental Society, University of California, Berkeley.

——. 1993. "Early Arabic-Persian lexicography: The *asāmī* and *maṣādir* genres". *Proceedings of the Colloquium on Arabic Lexicology and Lexicography*, I, ed. Kinga Dévényi, Tamás Iványi, and Avihai Shivtiel, 247–260. Budapest: Eötvös Loránd University.

——. 1995. "Lexical doublets as a derivational device in Persian: The Arabic feminine ending". *Acta Orientalia Hungarica* 48.127–153.

Rāzī, Farīda. 1366/1987. *Farhang-e ʾarabi dar fārsi-i moʾāser* [Arabic vocabulary in contemporary Persian]. Tehran: Markaz.

Telegdi, Zsigmond. 1973. "Remarques sur les emprunts arabes en persan". *Acta Linguistica* (Budapest) 23:1–2.51–58.

Utas, Bo. 1977. *A Persian Sufi poem: Vocabulary and terminology*. London and Malmö: Curzon.

John R. Perry (University of Chicago)

Persian Loanwords

The history of mankind is characterized by an ongoing cultural exchange between different nations and cultures. Generally speaking, older cultures exert a much greater influence on the younger ones because they possess a greater number of achievements with regard to civilization. This process also has linguistic implications. Because the Arabs lived in a relatively remote area before the advent of Islam and had only limited cultural exchange with other peoples, they did not possess words for things uncommon in their cultural and ecological environment. Even in the pre-Islamic period, the Arabs borrowed words from the Persian language, but the process increased after the spread of Islam throughout the Middle East, which united peoples of different cultural backgrounds.

In the pre-Islamic period, Arabs and Persians had some contact in border areas of the Arabian Peninsula. There were, for example, the Lakhmids, who were in the service of Sassanian Persia and secured the border against invasions from Bedouin tribes from the desert. Almost half a century before the advent of Islam, Yemen came under the rule of Sassanians and Persian governors, who ruled there even after the advent of Islam. These contacts had linguistic implications, and we find a comparatively large number of Persian loanwords in the Arabic language already before the advent of Islam, e.g. *ʾibrīq* < Middle Persian *ābrēz* 'jug'; *ṣanj* < Middle Persian *čang* 'harp'; *tāj* < Middle Persian *tāj* 'crown'.

After Islam became the dominant religion throughout the Middle East, including Iran, this process changed to a certain degree in the opposite direction. However, since the new Arab rulers had no experience in the administration of a great empire and were dependent on Iranian experts, Iranian culture still exerted a very remarkable influence. That influence was felt, for example, in the translation of Middle Persian treatises on administration into Arabic and the adaptation of such Middle Persian administrative terminology as *daftar* 'register, account book'; *dīwān* < Middle Persian *dēwān* 'archive, collected writings'; *jizya* 'head tax, tax' < Middle Persian *gazīdag* 'poll tax'; *firmān* 'edict, decree' < Middle Persian *framān* 'order, command'.

Administration was not, however, the only field in which the Arab language derived its terminology from Persian. Because their natural environment differed from other, much richer cultural areas, the Arabs had to borrow names for plants not familiar to them from other languages. Many of these were taken from Persian, as in the following examples: *banafsaj* < Middle Persian *wanafšag* 'violet'; *yāsmīn* < Middle Persian *yāsaman* 'jasmine'; *narjis* < Middle Persian *nargis* 'narcissus'; *sūsan* (or *sawsan*) 'lily' < Middle Persian *sōsan; warda* 'rose' < Old Persian *varəδa* 'rose, flower'. Not only were many flowers unknown to the pre-Islamic Arabs, but also some kinds of fruits. Their names were taken from Persian as well as the names of vegetables, herbs, spices, and various kinds of nuts, as in the following examples: *turunj* 'a kind of citrus' < Persian *torang; zard ālūj* 'apricot' < Persian *zardālū* < Middle Persian *zardālūg; bāḏimjān* (or *bāḏinjān* or *bāḏiljān*) 'eggplant' < Persian *bādemjān; šāhdānaj* 'hemp seed' < Middle Persian *šāhdānag; šāhsbaram* (or *šāhsbarham* or *šāhsfarham*) 'basil royal' < Middle Persian *šāhesprahm; isfanāj* (or *isfānāj* or *isbanāj*) 'spinach' < Persian *äspanāx; zaʿfarān* 'saffron' < Persian *zarparān* 'with gold leaves'; *bistaj* (or *bastaq* or *fustuk*) 'pistachio' < Persian *peste* < Middle Persian *pistag*.

The Arabs also borrowed expressions for herbs from the Persian language, which the Iranians themselves had learned from other peoples through, for example, trade with Central Asia and India, e.g. Arabic *dāraṣīnī* 'cinnamon' < Persian *dārčīn*. The natural environment of the Arabian Peninsula was also less diverse compared to other areas of the Middle East, and it is not surprising that the various names of animals not known to the Arabs in their ancestral country were taken from Persian, e.g. *barastūk* or rather *farastūk* 'swallow' < Persian *parastū* < Middle Persian *paristog; ṭayhūj* 'small gray partridge' < Middle Persian *tihōg*.

The adaptation of words in the realm of culture generally results from the fact that they are taken from another language by people who do not have in their culture any equivalent for the phenomena denoted by these words. This proves especially true for the adaptation of the names for Iranian festivals, which were not widely observed by Arabs and for which they had no name of their own. Some expressions were connected with Old Iranian religion, e.g. *majūs* 'magician' < Middle Persian *magupat* 'the chief of the Magi, i.e. main priest of the Zoroastrian clergy' < Old Persian *magu*/ Awestian *moγu; zindīq* (pl. *zanādiqa*) 'dualist, Manichaean' < Middle Persian *zandīk* 'heretic, Manichaean'; *ʾāḏarjašn* < Persian *āzarjašn* < Middle Persian *ādūr* 'fire' + *jašn* 'celebration, festival'; *mahrajān* 'festivity, celebration' < Persian *mihrgān* < Middle Persian *mihr* 'the Aryan god Mithra' < Avestan *miθrā*, Mithra's birthday and the day in which he fought and defeated the dragon *Dahak*.

In the field of material culture, the Arabs were acquainted with some Persian achievements, which explains why they adopted the terms from the Persian language. This may be seen by expressions in the field of architecture denoting types of buildings not typical of the Arabs, such as *ʾīwān* 'palace, columned hall, gallery'; *balkūna* 'balcony' < Persian *bālkon* < *bālā* + *xāne* < Middle Persian *bālā* 'top, upstairs' < Old Persian *bardista* < Avestan *barzišta* + Middle Persian *xānag* 'house'; *dihlīz* 'corridor, columned hall' < Middle Persian *dahlīz* 'portico'.

There were differences in Arab and Iranian clothing, as may be seen in certain expressions that the Arabs borrowed from the Persian language, e.g. *sirbal* or, more frequently, *sirwāl* (pl. *sarāwīl*) 'trousers, pants' < Middle Persian *šalwār*; *bābūj* (or *bābūs* or *bābūš*) 'shoe' < Persian *pāpūš* < Middle Persian *pad* 'foot' + *pūš* 'to wear'.

The Arabs were familiar with some Persian musical instruments, e.g. *ṭanbūr* < Middle Persian *tambūr* 'zither, lute'; *sanj* ' harp' < Persian *čang* < Middle Persian *čang; tumbak* 'drum' < Middle Persian *tumbag*.

The natural sciences also drew from Persian expressions that found their way into Arabic. This may be seen in astronomy, e.g. *ʾasfahr* < Middle Persian *spihr* 'sphere, sky, firmament; fate'; *kaywān* < Middle Persian *kēvān* 'Saturn' or *bahrām* < Middle Persian *vahrām* 'name for the god of war'.

On the other hand, in the field of religion, the Arabic language left its traces in Persian as numerous terms and expressions were adapted by it (→ Persian), e.g. *qurbān* 'sacrifice', *ḥarām* 'prohibited', *ḥalāl* 'allowed, permitted', *ḥajj* 'pilgrimage'. It should be noted, however, that sometimes Persian words were also used to express certain Islamic ritual practices or concepts, such as the Persian *namāz* instead of the Arabic *ṣalāt* 'prayer', or Persian *rūze* instead of Arabic *ṣawm* 'fast'. The influence of the Arabic language on Persian was certainly remarkable, and even today there exists a very great admiration for it in Iran. In popular consciousness, most people are convinced that more than 80 percent of the Persian language consists of Arabic components. Many words, however, are erroneously regarded as Arabic, and it seems more appropriate to speak of Arabicized words than originally Arabic ones. Classical Arabic encyclopedias such as the *Lisān al-ᶜArab, Tāj al-ᶜarūs*, or *Muḥīṭ al-muḥīṭ* contain many expressions classified as *fārisiyya*. This is the case for words borrowed from Middle Persian which later, under the influence of the Arabic language, were adapted by Modern Persian in their Arabicized form, e.g. *fihrist* < Middle Persian *pat-rāst* 'register', *ʾamrūd* < Middle Persian *urmōd* 'pear'.

Persian words transferred to Arabic had to be adapted to the Arabic sound system. This process occurred in the following ways:

i. Although Arabic has a much richer sound system than Persian, some Persian sounds do not exist in Arabic. Since the Persian consonants *p, č, ž*, and *g* have no equivalents in Arabic, they were replaced by certain Arabic ones:

   a. The Persian plosive *p* was replaced either by the Arabic fricative *f*, as in *firdaws* 'paradise' < Persian *pardis* < Middle Persian *pardis* and *pālīz* < Avestan *pairidaeza*, or by the Arabic plosive *b*, as in *bābūs, bābūj, bābūš* 'slipper' < Persian *pāpūš* < Middle Persian *pad* 'foot' + *pūš* 'to wear' < Avestan *pād*.

   b. Persian *č* was replaced either by Arabic *š*, e.g. *šādur* 'veil' < Persian *čādor*, or by the emphatic Arabic sibilant *ṣ*, e.g. *ṣakk* 'court file' < Persian *ček*; another example is *ṣawljān* 'polo game' < Persian *čogān* < Middle Persian *čopgān*.

   c. The Persian sibilant *ž* was replaced either by the Arabic affricate *j*, e.g. *bāj* 'tribute' < Middle Persian *bāž* < Old Persian *baji*, or by the Arabic sibilant *z*, e.g. *qazz* 'crude silk' < Persian *kaž* < Middle Persian *kač*.

   d. Persian *g* was replaced by Arabic *j*, seen in the example of Persian *gohar* < Middle Persian *gōhr* 'substance, essence, nature, jewel', which changed in Arabic to *jawhar* and even developed the lexical plural form *jawāhir*. In this case, the form *gohar* means 'origin', *johar* 'ink', and *jawāher* 'jewel'. It could also be replaced by Arabic *q*, as in *sūq* 'market' < Middle Persian *sōg*.

   e. A special case is the Persian *v*, which changed in Arabic to the semivowel *w*, as may be seen in *wazīr* 'minister' < Persian *vazīr* < Middle Persian *vicīr*. Both sounds are realized by the same character in script.

ii. Some Persian consonants, although they have equivalents in the Arabic sound system, were also occasionally replaced by others, as may be seen in the following list:

   a. The Persian plosive *b* was sometimes replaced either by the Arabic fricative *f*, as in Arabic *ʿifrīz* 'frieze' < Persian *ābriz*, or by the Arabic semivowel *w*, as in Arabic *darwand* 'lock' < Persian *darband* < Middle Persian *dar-band*.

   b. The Persian plosive *t* sometimes changed to the Arabic emphatic plosive *ṭ*, as in Arabic *ṭāzaj* (or *ṭāza*) 'new, fresh' < Persian *tāze*.

   c. The Persian fricative *x* may change in Arabic in three different ways: it may be replaced (1) by the affricate *j*, as in Arabic *ʾisfānaj* 'spinach' < Persian *äspanax*; (2) by the plosive *k*, as in Arabic *kisrā* 'a title for the Persian kings' < Persian *xosrow* < Middle Persian *hu-sraw* 'famous, of good repute'; (3) by the fricative *ġ*, as in the Arabic *bādġays, bādġayš* 'name of a district in Xorāsān' < Persian *bādxīz* lit 'windy'.

   d. The Persian plosive *d* may change into the Arabic emphatic plosive *ṭ*, e.g. *bāṭiyya* 'jug, bowl' < Persian *bādiya*.

   e. The Persian vibrant *r* may change into the Arabic lateral *l*, as in Arabic *salband* 'headgear' < Persian *sarband*.

   f. The Persian sibilant *z* may change in Arabic (1) into the affricate *j*, as in *ʾābrij* 'milking' < Persian *ābriz*, (2) into the emphatic plosive *q*, as in Arabic *ʾibrīq* 'watering can' < Persian *ābriz*, (3) into the fricative *ḏ*, e.g. *ḏaryāb* 'gold' < Persian *zar-e nāb* 'pure gold', (4) into the emphatic sibilant *ṣ*, as in *raṣāṣ* 'lead, pewter' < Persian *arziz* < Middle Persian *arčič*.

g. The Persian sibilant *s* is replaced (1) by the emphatic sibilant *ṣ*, e.g. *ṣanja* 'weight' < Persian *sanje* < Middle Persian *sanj* 'weight', (2) by the sibilant *z*, e.g. *razdaq* 'row' < Persian *raste* < Middle Persian *rastak*, (3) by the sibilant *š*, e.g. *šarbūš* or *šarbūs* 'headgear' < Persian *sarpuš*.

h. The Persian sibilant *š* sometimes changed into the Arabic sibilant *s*, as may be seen in *ʾibrīsam* (or *ʾibrīsim*) 'silk' < Persian *abrišam*.

i. The long Persian vowel *ā* sometimes changed into the Arabic fricative *ʿ*, as in *ʿabqarī* 'a kind of carpet' < Persian *ābkārī* < Persian *āb* 'gleam, shine' + *kārī* 'working'. Other examples are *ʿifrīt* 'demon' < Persian *āfarīd* 'creature' < Middle Persian *afrīt; ʿaynak* 'glasses' < *darī āynak*, a diminutive from Persian *āyne* 'mirror'.

j. The Persian short vowel *a* also sometimes changed into the Arabic fricative *ʿ*, e.g. *ʿayyār* 'tramp, vagabond' < Persian *yār ʿayyār* 'friend' < Middle Persian *ayār*.

k. The Persian fricative *ġ* sometimes changed into the affricate *j*, as in *ʾurjuwān* 'purple' < Persian *arġavān* < Middle Persian *arġavān*.

l. The Persian plosive *k* changed into the Arabic emphatic plosive *q*, as in *qahrimān* 'major' < Persian *kohramān* < Middle Persian *kof* 'mountain, hill'.

m. The Persian lateral *l* changed into the Arabic vibrant *r*, as in *jarjandum* 'wheat' < Persian *golgandom*.

n. The Persian fricative *v* changed into the Arabic plosive *b*, as in *ʾasbārān* 'chevalier' < Persian *savār* < Middle Persian *asvār*.

iii. Some expressions were taken directly from Middle Persian into Arabic. For this reason, the Arabic forms sometimes look much more similar to Middle Persian than to the Persian forms of the word, as in the following:

a. Middle Persian words ending with the suffix *-ak* changed in Arabic into *-aj*; these words appeared in Persian with the suffix *-e*, as in Arabic *barnāmaj* 'program' < Middle Persian *bārnāmak*, Persian *bārnāme*; another example is Arabic *rāhnāmaj* 'nautical chart' < Middle Persian *rāh* 'way, road' + *nāmak* 'book', Persian *rāhnāme* 'guide'.

b. This characteristic also applies to words which contain the fricative *h* at the beginning in their Arabic as well as in their Middle Persian form; the Persian form generally lacks the *h* at the beginning of the word. This may be seen in Arabic *handāz* 'scale' < Middle Persian *handāz* 'to plan, allot', Persian *andāz*. The secondarily developed Arabic *handasa* 'geometry', with its derivations, e.g. *muhandis* 'engineer', found its way back into the Persian language.

c. Middle Persian words having the long vowel *ā* at the beginning changed in Arabic to the glottal stop *ʾ*. In the New Persian form, the vowel at the beginning disappears, as in Arabic *ʾusṭuwāna* 'cylinder' < Middle Persian *āstūn* 'column, pillar, mast' < Old Persian *stūnā*. Persian has both forms, but with different meanings: *sūtūn* 'pillar' and *ostovāne* 'cylinder'.

d. In some words, the affricate *č* was replaced in Arabic by the sibilant *z*, e.g. *wazīr* 'minister' < Middle Persian *vičir* < Old Persian *vicira*.

e. The fricative *x* was replaced in Arabic by the plosive emphatic *q*, as in *waqt* 'time' < Middle Persian *vextan* 'to flow'; New Persian adopted the Arabic form of the word, although the pronunciation changed to *vaġt*. Unlike Persian, Kurdish has *vaxt*, retaining the consonant *x*.

f. The plosive *k* was replaced in Arabic by the emphatic plosive *q*, as in *furāniq* 'guide, leader' < Middle Persian *parwānak* > Persian *parwāne*.

g. The affricate *h* was replaced in Arabic by the fricative *ḥ*, as in *junāḥ* 'sin, crime' < Persian *gonāh* 'sin' < Middle Persian *wināh*.

iv. There are even some words in Arabic that seem to have been borrowed directly from Old Persian and have no equivalents in Middle Persian or New Persian, e.g. *qaraš* 'change' < Old Persian *karša* 'scale', Arabic *qaṣr* 'palace' < Old Persian *tacara* 'palace' (an alternative etymoloy derives this word from Latin *castra*).

v. There are different versions of some Persian words in Arabic, which differ by the use of either a long or a short vowel, e.g. *sūsan* and *sūsān* < Persian *sūsan* < Middle Persian *sōsan* 'lily'.

vi. Metathesis: some word forms were developed by the transposition of two phonemes, as in *janzīr*, which is derived from Persian *zanjīr* 'chain'.

vii. Some forms came into use by confusion of diacritics, e.g. Persian *xoškpān* 'a generic term for some kinds of sweets, which are filled with nuts, e.g. almonds', which exists in Arabic as *xoškbān* as well as *xošknān*.

viii. Some versions came into use by different oral realizations, as in Persian *golgandom* 'knapweed', which in Arabic became realized as *joljandom*, *jorjandom*, and *jarjandum*.

ix. For some Persian words, a broken plural was formed, e.g. Arabic *ʾasātīḏ* and *ʾasātiḏa* from Persian *ostāḏ* 'master'.

x. Some Arabic verbs are derived from Persian nouns, e.g. *tawwaja* 'to crown', from Middle Persian *tāj* 'crown'.

Bibliographical references

Detailed references in:

Asbaghi, Asya. 1987. *Die semantische Entwicklung arabischer Wörter im Persischen*. Stuttgart: F. Steiner.

——. 1988. *Persische Lehnwörter im Arabischen*. Wiesbaden: O. Harrassowitz.

Further information in:

Bahar, Malek aš-šo'ara. 1984. "Javanmardi". *Ayin-i javanmardi*, ed. Henry Corbin and Ehsan Naraqi, 109–120. Tehran.

Tafazzoli, Ahmad and Jaleh Amouzegar. 1999. *Tarix-e adabiyat-e Iran piš az Eslam* [The history of Iranian literature before Islam]. Tehran: Soxan.

Asya Asbaghi (Berlin)

Personal Pronoun (Arabic Dialects)

The pronominal system of Arabic dialects has one characteristic in common with other varieties of the Arabic language and other Semitic languages: it consists of a series of independent pronouns and a series of suffixed pronouns, whose form may vary from region to region for various reasons, among which are the influence of the substratum (e.g. the Aramaic substratum in the Syrian-Lebanese area; see Diem 1971) and its own internal evolution, or even a mixture of both (see Behnstedt 1991). These pronoun series play different syntactic roles that may also vary depending on geographical factors.

1. Independent personal pronouns

The various paradigms may be grouped into three categories, according to types of dialects: Bedouin dialects, without geographical distinction, and sedentary dialects, both Eastern and Western.

The forms of the independent personal pronouns for each of the three categories are given in Table 1.

Table 1. Independent personal pronouns in three types of dialects: Bedouin (Rosenhouse 1984:17–18), Cairo (Jomier and Khouzam 1977:36), and Moroccan koine (Caubet 1993:I, 159)

Free pronouns	Bedouin	Eastern sedentary	Western sedentary
1 comm. sg.	*ani, āni, ane*	*ana*	*āna, ānāya*
2 masc. sg.	*inta, inte, ənte, ənta*	*enta*	*ənta, əntāya, əntīna*
2 fem. sg.	*inti, ənti*	*enti*	*nti, əntīya, əntīna*
3 masc. sg	*huwwa, hū*	*howwa*	*hūwa*
3 fem. sg.	*hiyya, hī*	*heyya*	*hīya*
1 comm. pl.	*iḥna, əḥna, naḥna, ləḥna*	*eḥna*	*ḥna, ḥnāya*
2 masc. pl.	*entum, entam, entu*	*entu, entum*	*ntūma*
2 fem. pl.	*intan, intin*	*entu, entum*	*ntūma*
3 masc. pl.	*humma, hum*	*homma*	*hūma*
3 fem. pl	*hinna, hin*	*homma*	*hūma*

These forms are the most dialectally 'neutral', and many variants are found from dialect to dialect. Generally speaking, and with regard to number, a cross-dialectally uniform characteristic is the lack of the dual pronoun, continuously receding in Neo-Arabic.

With reference to gender, the 1st person is common in both singular and plural generally in all dialects. The 2nd person singular usually differentiates gender, although there are some exceptions (an example is the form *əntˢīna* [comm.], characteristic of the north of Morocco), whereas in the plural forms there is a distinction between feminine and masculine in some Bedouin dialects only, for example, in Negev Arabic (see Blanc 1970:130) and in Najdi Arabic (see Ingham 1982:74), as well as in other areas of the Arabian Peninsula, such as the dialect of the United Arab Emirates (see Isaksson 1991:121). In the 3rd person, the distinction between genders in the singular always occurs, while in the plural it is only preserved in some dialects of the Bedouin type (in the same cases as in the 2nd person plural). A stronger influence of the sedentary dialects on the Bedouin ones in the western zone can therefore be affirmed, while in the eastern region it is less pronounced. As a consequence, various dialects lack gender distinction in the 2nd and 3rd person singular, and there is an almost total lack of such distinction in the plural persons.

Independent personal pronouns have the following grammatical functions: subject, copula, and the expression of emphasis. As subject, they usually appear in verbless clauses, e.g. *hu sākin taḥti* 'he lives right next to me' (Gulf Arabic; Holes 1984:67), although sometimes they also clarify the ambiguity of some verbal forms. As copulas, they are the 3rd person pronouns in verbless predicates with this function (Eid 1991), e.g. *āva lwalad ūwe bōš malī* 'this boy is very good' (Siirt Arabic; see Jastrow 1978:132). The third role of the independent personal pronouns is to emphasize some element that can either be subject of a clause or have another syntactic function. They are usually located in clause-initial position, e.g. *hūwa ygūl hād l-kəlma hakka* 'he says this word like this' (Moroccan Arabic).

Along with these standard forms and these basic syntactic functions, there are some forms and functions that are particular to each dialect. Most of them are due to different phonetic realizations, or to the appearance of 'secondary forms', i.e. forms to which an external element has been added, usually in the form of a suffix. As an example of phonetic differentiation, there is the case of the presence of the → *ʾimāla*, whose effect on the personal pronouns varies depending on the dialect, and which is marked both in pausal allomorphs and in context. A notable example of variation is the pronunciation of the pronoun of the 1st person singular *āni* in the dialect of the Muslims from Baghdad, which must also have been an ancient form in the western zone, since it already existed in the Andalusian dialect of Granada (see Corriente 1977:97) and still exists today in some Tunisian dialects (e.g. Isaksson 1999:59). Some Arabic dialects, for example those of Yemen and the Lower Gulf, have *ani* as a feminine variant of the 1st person singular (cf. Holes 1984:55), and it is generally accepted as a dialectal innovation because gender in the 1st person is not distinguished anywhere else. According to one theory, however, it may be a survival of an archaic pre-Islamic form rather than a dialectal innovation (see Zaborski 1995:293). Instances of the *ʾimāla* are also found in other pronouns, for example the Syrian-Lebanese forms *hūwe* and *hīye*.

Regarding the 'secondary forms' or dialectal innovations, we have the case of *-ya* (for example in *ānāya, əntāya*, and *ḥnāya*) or of *-k*. Both appear in the western zone, the former as an emphatic variation (Isaksson 1999:61), and both among the sedentary and the Bedouin dialects, for example in the dialect of Fez (Caubet 1993:I, 159) or in Ḥassāniyya in Mauritania (Cohen 1963:146). The second suffix is more typical of the rural dialects, distinguishing them from the urban ones, for instance in forms like *ntˢīnăk* for the 2nd person (comm.) singular and *ntˢūmăk* (comm.) plural (Vicente 2000:136).

With regard to the syntactic function of pronouns, there are also some dialectal innovations. They appear to be pronouns, but their syntactic behavior suggests they may be something else. Thus, the pronoun of the 3rd person may act as 'question marker', although dialects vary with respect to this use, which occurs only clause-initially and in matrix clauses. The pronoun agrees with the nearest substantive, but if the subject is a personal pronoun, such agreement is not necessary (Eid 1992). The dialects that make use of the pronouns in this way are dialects in which there are no question markers, such as Iraqi, Palestinian, and Egyptian Arabic, e.g. in Egyptian Arabic: *(huwwa) ʿali katab ilgawāb?* 'did Ali write the letter?', *(hiyya) nadya ḥa-tuxrug?* 'will Nadya leave?', *(humma) ilmudarrisīn katabu ilgawāb?* 'did the teachers write the letter?'.

There are also special pronouns to indicate nominal negation, which are known as 'negative pronouns'. They always agree with their subjects and are suffixed to the first part of the negative morpheme *ma-*. Sometimes reduced forms of the independent pronouns are used, except in the 1st person singular. Such structures are found, for instance, in Moroccan dialects (Table 2).

Table 2. Nominal negation with independent personal pronouns, in Moroccan Arabic (Caubet 1993:I, 167)

1 comm. sg.	*mā-ni ši, m-āna ši*	1 comm. pl.	*mā-ḥna ši*
2 masc. sg.	*mā-nta ši*	2 comm. pl.	*mā-ntūma ši*
2 fem. sg.	*mā-nti ši*		
3 masc. sg.	*mā-hūwa ši*	3 comm. pl.	*mā-hūma ši*
3 fem. pl.	*mā-hīya ši*		

2. SUFFIXED PERSONAL PRONOUNS

The forms of suffixed personal pronouns can be added to verbal, nominal, or prepositional forms (see Table 3).

The 1st person singular is the only one that has different forms depending on the function of the suffix: *-ni* when added to verbal forms, but *-i* (with allomorphs) after consonant and *-ya* (with allomorphs) after vowel to nominal forms and to prepositions.

The rest of the pronouns normally add a short vowel when the word to which they are added ends in a consonant. Nevertheless, the pronoun for the 3rd person masculine singular has two forms: *-u* (with various phonetic realizations), which is *-ah* when the word ends in a vowel or a diphthong, and another form *-ah* which is reduced to *-h* in front of a vowel or a diphthong. The former is the most common, and it appears in all types of dialects, while the second only exists in dialects of the Bedouin type, both Eastern and Western. According to Cantineau (1939:96–97), this double form is due to its different origin, since each comes from a different pronoun of ancient Arabic.

As for the other pronouns of the 3rd person, in some dialects their articulation is very weak and the *h* even disappears before a consonant, both in the feminine (e.g. *nkətba* 'we write it'),

Table 3. Bound or suffixed pronouns in various Arabic dialects: Bedouin (Rosenhouse 1984:18), Cairo (Jomier and Khouzam 1977:41), and Moroccan koine (Caubet 1993:I, 160)

Bound pronouns	Bedouin	Eastern sedentary	Western sedentary
1 comm. sg.	*-ni* *-i, -ya, -yi*	*-ni* *-i, -ya,* *-yya*	*-ni* *-i, -y,* *-iyya*
2 masc. sg.	*-k, -ak, -ek*	*-ak, -k*	*-ək*
2 fem. sg.	*-ek, -ik*	*-ek, -ki*	*-ək*
3 masc. sg	*-u(h), -o(h)* *-a(h), -i(h)*	*-o, -h*	*-u, -əh*
3 fem. pl.	*-ha, -he, -hi*	*-ha*	*-ha*
1 comm. pl.	*-na, -ne*	*-na*	*-na*
2 masc. pl.	*-kam, -kom,* *-kum, -kām*	*-kom*	*-kum*
2 fem. pl.	*-kin, -ken*	*-kom*	*-kum*
3 masc. pl.	*-hum, -ham,* *-hām*	*-hom*	*-hum*
3 fem. pl	*-hin, -hen*	*-hom*	*-hum*

and in the plural forms (e.g. *yəddŭm* 'their hands'), a feature which is common in northern Moroccan dialects (cf. Vicente 2000:138). These suffixed pronouns replace the independent pronouns in the following three syntactic functions:

i. Verb complements can be direct or indirect objects, when added to a verb.
ii. Noun complements act as possessives when added to a noun. This use of the pronoun is more widespread in the Bedouin-type dialects, since in the sedentary ones it is limited to certain noun categories; in these dialects, possession is expressed by means of an analytic construction (→ analytic genitive).
iii. Preposition complements can be added to prepositions. Combinations with the preposition *li-* to express an indirect object (or dative) are a well-known feature of Arabic dialects. There are differences, both in form and in use: some dialects distinguish two ways of annexing the suffixed pronoun to the preposition, e.g. in Egyptian Arabic, the series *li, lak, lek, lo, laha, lena, lokum, lohum* is used when the preposition complements to a verb, whereas the series *leyya, līk, līki, līh, līha, līna, līkom, līhom* has another semantic nuance, that of emphasizing → possession (see Jomier and Khouzam 1977:52–53).

Another distinguishing feature between dialects is the use of the combination of *li-* with a pronominal suffix, as an enclitic element of the verbal form or as a separate element. In the former case, the pronoun is cliticized to the verb, e.g. *katab-lak* 'he has written to you'. This most commonly occurs in Eastern dialects, particularly in the case of the Egyptian dialects and in the Syrian-Lebanese region (see Durand 1999:95), e.g. (Egyptian Arabic) *warri-hū-lo* 'show it to him!', *ma-ʾulti-lī-š ḥāga* 'you have not said to me anything at all'. In some dialects, however, cliticization is optional, as in the case of North African dialects. According to Durand (1999:96), originally pronouns were not cliticized in the Arabic of this area, and the cases that are found nowadays are due to an eastern innovation. For this reason, cliticization in this type of dialect occurs less systematically, as one moves westward. Thus, the descriptions of the North African dialects show that the verb and the preposition complement can be both separated or cliticized, the latter option being almost universal in Libyan Arabic, probably due to its geographical proximity to the Eastern dialects, which supports the theory of eastern influence, e.g. *šnu gāl lək?* 'what has he said to you?' (Moroccan Arabic; Caubet 1993:I, 207), *gŭlt-lhum* 'I have said to them' (Fezzan Arabic; Marçais 2001:176), *qāl-ha* 'he said to him' (Algiers dialect; Boucherit 2002:140).

These pronouns may also be suffixed to another type of particle, adding different semantic nuances. For instance, when they are suffixed to a genitive exponent, they serve to express possession, e.g. *dyāli* 'my/mine' (Moroccan dialect), *bitāʿak* 'your/yours' (Egyptian dialect), and added to a noun meaning 'self' or 'soul', they express reflexivity: *b-ṛāṣu* 'himself' (Ḥassāniyya Arabic), *nafsak* 'yourself' (Egyptian Arabic).

Another syntactic function of suffixed pronouns is to act as a referent in those situations where the relative pronoun is not the subject of the clause. In such cases, the use of the personal pronoun indicates the function of the relative pronoun, e.g. *il ʾōḍa illi ṣaḥbi sākin fīha* 'the room that my friend lives in'.

Dialectal differences in this case are usually due to different phonetic pronunciations. For example, in some Bedouin dialects the element *k* in the pronoun of the 2nd person is affricated, producing the bound pronouns *-ič* sg./*-čin* pl., for example in Najdi Arabic (cf. Ingham 1982:74). In the case of the Bedouin dialects of the Gulf, the different phonetic realizations have become grammaticized to mark the distinction of gender: 2nd person masculine singular *-ik, -ak*/2nd person feminine singular *-ič, -ač* (Ingham 1982:88; → *kaškaša/kaskasa*).

Finally, in relation with the suffixed pronouns, the secondary forms or dialectal innovations usually consist in the insertion of a linking element between the noun or verb and the suffix. This element can be a vocalic phoneme, like the vowel *ă*, which occurs in plural pronouns of some sedentary Western dialects, especially in the rural ones, e.g. *nəšrīwăh* 'he buys it' (Vicente 2000:138). It can also be a connective particle, as in the case of *-nn-, -n-*, which are derived from deictic elements, to which pronominal suffixes are added; they may mark both subject and object. This construction, which appears in other Semitic languages, occurs in some Arabic dialects of Oman (Daṯīna) and Uzbekistan and in other dialects in Central Asia and of the Gulf, although its use and distribution are not uniform in all of them. According to Retsö (1988:89), this linking element originally occurred with the suffixes of the 1st and 3rd person singular, and later it spread to the other persons. A final example is the case of the particle *-iyya*, which is used to introduce the direct object pronoun, e.g. *gilt li hum-iyyāh* 'I said it to them'. It occurs in northern dialects of Syria and Iraq, and in some of the Gulf dialects (Ingham 1994:30).

BIBLIOGRAPHICAL REFERENCES

Behnstedt, Peter. 1991. "Noch einmal zum Problem der Personalpronomina *hənne* (3. Pl.), *-kon* (2. Pl.) und *-hon* (3. Pl.) in den syrisch-libanesischen Dialekten". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 141.235–252.

Blanc, Haim. 1970. "The Arabic dialect of the Negev Bedouins". *Proceedings of the Israel Academy of Sciences and Humanities* 4:7.112–150.

Boucherit, Aziza. 2002. *L'arabe parlé à Alger: Aspects sociolinguistiques et énonciatifs.* Paris and Louvain: Peeters.

Cantineau, Jean. 1939. "Le pronom suffixe de 3e personne singulier masculin en arabe classique et dans les parlers arabes modernes". *Bulletin de la Société de Linguistique de Paris* 40.89–97.

Caubet, Dominique. 1993. *L'arabe marocain.* I. *Phonologie et morphosyntaxe.* II. *Syntaxe et catégories grammaticales, textes.* Paris and Louvain: Peeters.

Cohen, David. 1963. *Le dialecte arabe ḥassānīya de Mauritanie (parler de la Gəbla).* Paris: C. Klincksieck.

Corriente, Federico. 1977. *A grammatical sketch of the Spanish Arabic dialect bundle.* Madrid: Instituto Hispano-Árabe de Cultura.

Diem, Werner. 1971. "Zum Problem der Personalpronomina *hənne* (3. Pl.), *-kon* (2. Pl.) und *-hon* (3. Pl.) in den syrisch-libanesischen Dialekten". *Zeitschrift der Deutschen Morgenlandischen Gesellschaft* 121.223–230.

Durand, Olivier. 1999. "La série pronominale indirecte en arabe dialectal". *Estudios de Dialectología Norteafricana y Andalusí* 4.89–98.

Eid, Mushira. 1991. "The copula pronoun in Arabic and Hebrew". *Perspectives on Arabic linguistics*, III, ed. Bernard Comrie and Mushira Eid, 31–61. Amsterdam and Philadelphia: J. Benjamins.

——. 1992. "Pronouns, questions and agreement". *Perspectives on Arabic linguistics*, IV, ed. Ellen Broselow, Mushira Eid, and John McCarthy, 107–142. Amsterdam and Philadelphia: J. Benjamins.

Holes, Clive. 1984. *Colloquial Arabic of the Gulf and Saudi Arabia.* London: Routledge and Kegan Paul.

Ingham, Bruce. 1982. *North East Arabian dialect.* London and Boston: Kegan Paul International.

——. 1994. *Najdi Arabic: Central Arabian.* Amsterdam and Philadelphia: J. Benjamins.

Isaksson, Bo. 1991. "The personal markers in the modern Arabic dialects of the Arabian Peninsula". *Orientalia Suecana* 40.117–145.

——. 1999. "The non-standard first person singular pronoun in the modern Arabic dialects". *Zeitschrift für Arabische Linguistik* 37.54–83.

Jastrow, Otto. 1978. *Die mesopotamisch-arabischen* qəltu-*Dialekte.* I. *Phonologie und morphologie.* Wiesbaden.

Jomier, Jacques and Joseph Khouzam. 1977. *Manuel d'arabe égyptien.* Paris: Klincksieck.

Marçais, Philippe. 2001. *Parlers arabes du Fezzân: Textes, traductions et élements de morphologie rassemblés et présentés par D. Caubet, A. Martin et L. Denooz.* Liège: Diffusion.

Retsö, Jan. 1988. "Pronominal suffixes with *-n(n)-* in Arabic dialects and other Semitic languages". *Zeitschrift für Arabische Linguistik* 18.77–94.

Rosenhouse, Judith. 1984. *The Bedouin Arabic dialects: General problems and a close analysis of North Israel Bedouin dialects.* Wiesbaden: O. Harrassowitz.

Vicente, Ángeles. 2000. *El dialecto árabe de Anjra (norte de Marruecos): Estudio lingüístico y textos.* Saragossa: Universidad de Zaragoza.

Zaborski, Andrzej. 1995a. "First person pronouns in Arabic in the light of Arabic and Hamito-Semitic dialectology". *Dialectologia arabica: A collection of articles in honour of the sixtieth birthday of Professor Heikki Palva*, 289–294. Helsinki: Finnish Oriental Society.

——. 1995b. "Archaisms and innovations in Arabic independent pronouns". *Proceedings of the 2nd International Conference of l'Association Internationale pour la Dialectologie Arabe*, 269–274. Cambridge.

ÁNGELES VICENTE (Universidad de Zaragoza)

Personal Pronoun (Standard Arabic)

A pronoun may be defined as a part of speech belonging to the broad class of nominals that stand or substitute for a noun. Unlike nouns, pronouns form small, closed classes, whose reference can only be determined from the context of the utterance in which they are used. In inflected languages like Arabic, pronouns often have heterogeneous inflectional patterns which differ from those of nouns themselves, and indeed Arabic is no exception here. Pronouns are also subject to → agreement with their referents in terms of gender and number, but in Arabic they do not exhibit case marking. Pronouns are usually divided into several subgroups according to syntactic and semantic considerations, such as personal, demonstrative, relative, indefinite, and interrogative pronouns. In some grammatical descriptions of Arabic, the term ‘pronoun’ is used to refer only to the first of these, i.e. personal pronouns (Badawi a.o. 2004:44–46), and the same practice is followed here.

Pronouns in Arabic exhibit the usual three persons: 1st person or speaker(s), 2nd person or addressee(s), and 3rd person or the subject of discourse, i.e. the person(s) or thing(s) spoken about. In Arabic, the 1st person has only two number forms, a singular and a nonsingular (i.e. dual and plural), and does not distinguish gender. The other persons all distinguish three numbers, singular, dual, and plural, and two genders, masculine and feminine, except in the dual. This gives Standard Arabic an inventory of twelve personal categories, as in the tables below. Pronouns are either free or bound: free pronouns, also called independent pronouns (see Table 1), normally occur as topics or predicates, but they may also be used in apposition to a noun or bound pronoun, typically as a mark of emphasis. Bound pronouns (see Table 2) occur added to finite verbs as objects, or added to nouns as possessives and prepositions as objects. The forms in both functions are identical except in the 1st person singular.

1. Independent pronouns

Table 1. Independent pronouns

	singular	dual	plural
1st pers.	*ʾana*	*naḥnu*	
2nd pers. masc.	*ʾanta*	*ʾantumā*	*ʾantum*
2nd pers. fem.	*ʾanti*		*ʾantunna*
3rd pers. masc.	*huwa*	*humā*	*hum*
3rd pers. fem.	*hiya*		*hunna*

Independent pronouns typically occur as topics or subjects with nonverbal predicates (i.e. nominals, including participles, and adverbials):

> *hum ʾawlād ṣiġār* ‘they are little boys’
> *hiya mutakabbira* ‘she is haughty’
> *ʾana maʿakum* ‘I am with you’

Independent pronouns are not usually used as subjects of verbal predicates, except in comment clauses with the topicalizing construction *ʾammā…fa-* ‘as for’, e.g.

> *ʾammā t-talāmiḏa fa-hum yaḥfaḏūna mā yuqaddamu lahum* ‘as for the students, they memorize what is given to them’
> *ʾammā l-ḥakīm fa-huwa rajul qawiyy* ‘as for the governor, he is a strong man’

In such constructions as the latter, the use of the independent pronoun is optional, so *ʾammā l-ḥakīm fa-rajul qawiyy* is also possible.

The independent pronoun is further used as a marker of emphasis following a nominal or a bound pronoun and is in agreement with its referent, e.g.

> *kānat ṣanʿāʾ hiya hājisahu l-jadīd* ‘Sanaa was his new concern’
> *hāḏā laysa kalāmī ʾana* ‘these are not my own words’
> *ʾinnahā hiya ʾummuhā tuwāsīnī ʾana l-ġarība* ‘she, her mother, was consoling me, the stranger’

In the last example, independent pronouns are employed to emphasize both the focused subject with *ʾinna, ʾinnahā* ‘she’, and the object ‘me’, denoted by the bound pronoun *-nī* on the verb.

In a nominal sentence not introduced by *ʾinna*, when both the subject and the predicate are definite, including where the subject is a demonstrative, the two are normally linked by the appropriate independent pronoun, which is usually in agreement with the subject. This use of the independent pronoun is what is called *ḍamīr al-faṣl* ‘the pronoun of connection’. In nominal sentences where, for instance, the predicate is not definite, a linking pronoun is not used, e.g.

> *tilka hiya l-ḥaqīqa* ‘that is the truth’
> *dīn ad-dawla huwa l-ʾislām* ‘the religion of the state is Islam’
> *al-hadaf huwa fahm maṣādirihi* ‘the goal is the understanding of his sources’
> *al-bašar hum hum lā yataġayyarūna* ‘people are what they are, unchanging’

In this last example, the pronoun *hum* occurs first as a *ḍamīr al-faṣl* and then again as a predicate. But with an indefinite predicate, no linking pronoun is normally used, e.g.

> *az-zawāj masʾūliyya* ‘marriage is a responsibility’
> *al-junayna mawḍiʿ jamīl* ‘the garden is a beautiful place’

However, sometimes an independent pronoun is used in such contexts with resumptive or anaphoric function:

> *al-mawqif allaḏī yattaxiḏuhu r-raʾīsān... huwa mawqif ʾījābī* 'the position that the two presidents are taking... is a positive position'

In *ḍamīr al-faṣl* constructions, the pronoun is sometimes likened to a → copula, i.e. an equivalent of the verb 'to be' in equational sentences. This is not, however, an accurate description of its function, which is rather in origin to separate or distinguish the subject from the predicate. In Modern Standard Arabic, as may be seen from the examples above, this function has been somewhat extended to emphasize the predication itself.

Lastly, the independent pronoun is used after *hā* 'here [it] is!' in presentative sentences, e.g.

> *hā naḥnu nafī bi-waʿdinā* 'here we are, keeping our promise!'
> *wa-hā hiya l-ḥaṣra taskunu fī ʾaʿmāqī* 'and here is despair, taking up residence in my depths!'

2. Bound pronouns

Table 2. Bound pronouns

	singular	dual	plural
1st	*-ī/-ya* *-nī*	*-nā*	
2nd masc.	*-ka*	*-kumā*	*-kum*
2nd fem.	*-ki*		*-kunna*
3rd masc.	*-hu/-hi*	*-humā/-himā*	*-hum/-him*
3rd fem.	*-hā*		*-hunna/-hinna*

Only the 1st person singular bound pronoun has different forms according to whether it is added to nominals, including prepositions, when it has the form *-ī/-ya*, or to verbs, when it has the form *-nī*. All other persons use the same forms in both positions. The other variations in the suffixes (see Table 2) are governed by phonological context. Thus, the 1st person singular *-ī* replaces a final short vowel, while the variant *-ya* occurs after long vowels or a diphthong, e.g. *kitābī* 'my book' (all cases) as against *kitābuka, kitābaka, kitābika* 'your book' (according to case), or *lī* 'to me' as against *lahu* 'to him', but *dunyāya* 'my world', *fīya* 'in me', and *ʿalayya* 'on me'. However, the three kinship nouns *ʾab* 'father', *ʾax* 'brother', and *ḥam* 'father-in-law', which show lengthened case vowels before other bound pronouns, keep *-ī* here, e.g. *ʾabī* 'my father' but *ʾabūka* 'your father'. Also, the 3rd person suffixes other than the feminine singular have a short *i* vowel when the immediately preceding syllable is open and contains either *i*, short or long, or the diphthong *ay*, e.g. *fīhi* 'in him', *ʿalayhi* 'on him', *fī baytihim* 'in their house', etc., but *lahu* 'to him', *baytuhum ʾakbaru min baytī* 'their house is bigger than my house'. In Arabic, the underlying or original forms are those with the vowel *u*, and the variants with *i* are due to vowel assimilation across the relatively weak syllable boundary in *h* (→ vowel harmony). Any additional constituent to the syllable boundary prevents the vocalic assimilation, as in *minhu* 'from him' as against *fīhi* 'in him'.

Bound pronouns in the function of possessive pronouns are added to the noun in the annexed state, i.e. without either the article or → *tanwīn*. The addition of the bound pronoun, however, prevents the loss of case-marking short vowels that occurs in the pausal forms, as well as the apocope of the feminine ending orthographically indicated by *tāʾ marbūṭa*. To this extent, the relationship of the possessive bound pronoun to its annexed noun is very similar to that of a genitive noun to its head, as can be seen by comparing *al-fallāḥūn* 'the peasants' with *fallāḥū miṣr* 'the peasants of Egypt' and *fallāḥuhā* 'her peasants':

> *maʿa l-mudarris* 'with the teacher'
> *maʿa mudarrisikum* 'with your teacher'
>
> *aṣ-ṣuḥufiyyūn* 'the journalists'
> *ṣuḥufiyyūhā* 'its (her) journalists'
>
> *fī l-jāmiʿa* 'at the university'
> *fī jāmiʿatinā* 'in our university'

The same set of bound pronouns is also added to prepositions, including those of overtly nominal origin, of course. The processes of combining the bound pronoun suffix with a preposition are the same as outlined above, with the following exceptions: the prepositions *min* 'from' and *ʿan* 'about' double their final consonant before adding the 1st person singular *-ī*: *minnī* 'from me', *ʿannī* 'about me'. In addition, several prepositions have different stems when combined with bound pronouns than when occurring with a following noun, most notably those ending in *-ā* (*ʾalif maqṣūra*), which becomes *-ay*: *ʿalā s-sarīr* 'on the bed', but

ʿalaynā ‘on us’; note also *li-l-muslimīn* ‘for the Muslims’ but *lahum* ‘for them’.

A number of other noninflected words or particles may also have bound pronouns added to them, such as the focalizer or opening particle of equational sentences *ʾinna*, the complementizer *ʾanna* ‘that’, the adversative particle *lākin[na]* ‘but’, the dubitative particle *laʿalla* ‘perhaps’, and the optative particle *layta* ‘would that’. The first two of these, *ʾinna* and *ʾanna*, and compound particles such as *lākinna* and *laʿalla* have variant forms with the 1st person singular and plural suffixes: *ʾinnanī* or *ʾinnī* ‘[indeed] I’, *lākinnanā* or *lākinnā* ‘but we’, etc. One further particle occurring with bound pronouns is *ʾiyyā*, which mostly marks a preposed pronominal direct object, used instead of the bound pronoun object suffix added to the verb either for emphasis or rhetorical purposes. It is also used where an object pronoun suffix may not be added to the verb, for instance when the latter is a verbal noun or active participle, where any bound pronoun suffix usually refers to the subject or agent (→ transitivity). Modern Standard Arabic also tends to prefer to use *ʾiyyā* + bound pronoun to mark a second pronoun object of doubly transitive verbs, while in Classical Arabic the addition of two bound object pronouns to a finite verb is possible. Examples in Modern Standard Arabic are:

> *ʾinnahumā najmatāni lāmiʿatāni* ‘[indeed] they are two shining stars’
> *ʾinnī jāʾiʿ* ‘[indeed] I am hungry’
> *samiʿnā ʾannahu ḍaraba waladahu* ‘we have heard that he struck his child’
> *lākinnahā btasamat bi-riqqa* ‘but she smiled gently’
> *laʿallī* [or *laʿallanī*] *btasamtu lahā* ‘I may perhaps have smiled at her’
> *laytanā nuwājihu l-ʾumūr bi-wuḍūḥ* ‘if only we would face matters clearly’
> *ʾiyyāka naṭlubu* ‘we want you’, ‘you we want’
> *qatluhum ʾiyyāhum* ‘his killing them’
> *ʾaʿṭaytuka ʾiyyāhu* ‘I gave it to you’, instead of *ʾaʿṭaytukahu*

Bound pronouns are also added to finite verbs to mark a pronominal direct object, in which case the 1st person singular has the form *-nī*. In all other persons the bound pronouns have the same forms as in possessive function. In Modern Standard Arabic, the preference is not to have more than one bound object pronoun added to a verb, although in Classical Arabic it is possible to add two to doubly transitive verbs as long as the hierarchical sequence of object referents is maintained, i.e. that 1st person precedes 2nd person, which in turn precedes 3rd person (*ʾaʿṭaytukahu* ‘gave you it’).

> *taqūdunī* ‘she leads me’
> *ḍarabtihi* ‘you [fem.] struck him’
> *ġasalathu l-banāt* ‘the girls washed it’ (e.g. *qamīṣī* ‘my shirt’)

3. INTERNAL STRUCTURE OF THE PRONOUNS

Pronouns are not marked for case, although it is evident that in many functions the independent pronoun occurs in positions where a nominative case would be used, and the bound pronouns generally have functions equivalent to accusative and genitive case nouns. This division of functions between the two sets of pronouns is an inherited feature from Proto-Semitic. It is also the case that in most → Semitic languages the two sets are formally distinguished by being independent and bound forms. However, in Akkadian, pronouns formally related to what we see in Arabic as the bound pronouns do occur as independent pronouns with accusative, genitive, and dative functions. Comparison with other → Afro-Asiatic languages further suggests that the restriction of the non-nominative set of pronouns to bound position is an innovation that occurred in most of Semitic and that the independent forms of Akkadian are archaisms.

There are some formal connections between the Arabic independent and bound sets of pronouns, although aside from the 3rd persons these are not immediately apparent and it is therefore the normal practice not to link the two. The internal structure of Arabic and Semitic pronoun sets becomes clearer when we examine at the same time person marking in the finite verb, i.e. the suffixes of the perfect and the prefixes of the imperfect. So, the independent pronoun of the 1st and 2nd persons, except for the 1st person plural, are constructed on the base *ʾan*, followed by a personal morpheme that is most closely related to the corresponding markers of person in the verb: the suffixes of the perfect in the 2nd persons and the prefix of the imperfect in the 1st person singular (see Table 3). See further Fleisch (1968a, 1968b: 136–139) for a detailed discussion of the structure of Arabic pronouns.

Table 3. Structure of the independent 1st and 2nd person pronouns

	singular	dual	plural
1st pers.	*ʾan + ʾa > ʾana*	(*naḥnu*)	
verbal prefix	*ʾa-/ʾu-*		
2nd pers. masc.	*ʾan + ta*	*ʾan + tumā*	*ʾan + tum*
verbal suffix	*-ta*	*-tumā*	*-tum*
2nd pers. fem.	*ʾan + ti*		*ʾan + tunna*
verbal suffix	*-ti*		*-tunna*

Some other Semitic languages, e.g. Akkadian and Hebrew, have a longer form of the 1st person singular: Akkadian *anāku*, Hebrew *ʾānōki*, and elsewhere in Afro-Asiatic we may likewise note Ancient Egyptian *jnk* (Coptic *anak/anok*) and Berber forms, such as Tuareg *näk*. These longer forms include an element that is clearly related to the corresponding verbal ending of the perfect (or stative in Akkadian), which was originally, in Proto-Semitic, something like *-aku, substituted by *-tu in Central Semitic (including Arabic). While Classical Arabic and hence Modern Standard Arabic show only final *-ā* (sometimes shortened to *-a* in the former), there are various Arabic dialect forms with final *-i*, long or short. These may be due to analogy with the bound pronoun *-ī/-nī*, or may perhaps be archaisms, as the Hebrew variant *ʾănī* might suggest, unless this, too, is an analogical formation. Incidentally, the use of *ʾana* for the masculine and *ʾani* for the feminine in the Jabal ed-Drūz in Syria and in Yemen is clearly an innovation, as gender distinction in the 1st person is not a general Semitic feature (Zaborski 1995:291).

The 1st person plural appears at first sight to be outside this system. While several Arabic dialects, including notably Early Christian Arabic, have forms in final *-a/-ā* (compare the bound form *-nā*) alongside Classical and Modern Standard Arabic *naḥnu*, and some dialects further show variation in the vowel of the first syllable as well, for instance in a number of Gulf dialects which have *niḥna*, no form of Arabic has variants in initial *ʾan-*, as for instance seen in Hebrew *ʾănaḥnū*. Some other Arabic dialects do, however, have forms without any initial *n-*, e.g. Cairene Arabic *ʾiḥna*, Moroccan Arabic *ḥna*. It is often suggested that such forms are due to dissimilation or haplology, but similar varieties elsewhere in Semitic, such as Syriac *ḥnan* (alongside *ʾenaḥnan*) or Amharic *əɲɲa* (< *ḥən[n]ā), Tigre *ḥəna*, and in Afro-Asiatic, such as Beja *hinin*, could suggest that the variant without initial *n-* is very archaic.

The 2nd persons of the independent pronouns all have the person marker *-t-* following the initial *ʾan* identical to the person marker in the finite verb. The bound pronouns of the 2nd person all commence with the person marker *-k-*. The indicators of gender and number in each case are the same in the independent (or free) and in the bound pronouns. Furthermore, these same indicators are also found in the dual and plural forms of the 3rd person pronouns, both free and bound (see Table 4).

The 3rd person independent pronouns are clearly of the same origin and structure as the corresponding bound set. These all begin in *h-* in Arabic, as they do in other Central Semitic languages, Sabaean, and (as reconstructed) in Ethiopian Semitic. In Akkadian

Table 4. 2nd and 3rd person pronoun structure

		singular	dual	plural
2nd pers. masc.	free	*ʾan + t-a*	*ʾan + t-umā*	*ʾan + t-um*
	bound	*-k-a*	*-k-umā*	*-k-um*
2nd pers. fem.	free	*ʾan + t-i*		*ʾan + t-unna*
	bound	*-k-i*		*-k-unna*
3rd pers. masc.	free	*h-uwa*	*h-umā*	*h-um*
	bound	*-h-u*	*-h-umā*	*-h-um*
3rd pers. fem.	free	*h-iya*		*h-unna*
	bound	*-h-ā*		*-h-unna*

and non-Sabaean Ancient South Arabian, the same pronouns show a sibilant: Akkadian 3rd pers. masc. sg. *šū*, 3rd pers. fem. sg. *šī*, 3rd pers. masc. pl. *šunu*, 3rd pers. fem. pl. *šina*; Qatabanian 3rd pers. masc. sg. *-s*^{1}*[ww]*, 3rd pers. fem. sg. *-s*^{1}*[yw]*, 3rd pers. masc. pl. *-s*^{1}*m*, 3rd pers. fem. *-s*^{1}*n* bound forms). The sibilant forms are the older, as comparative evidence from other Afro-Asiatic languages shows, and the shift of **š* > **h* is an innovation here (Voigt 1994). Interestingly, in Modern South Arabian we see a similar but partial shift in some languages, e.g. in Mehri: 3rd pers. masc. sg. *ha[h]*, 3rd pers. fem. sg. *sē[h]*, 3rd pers. masc. pl. *hēm*, 3rd pers. fem. pl. *sēn*, while Jibbāli maintains sibilants throughout: 3rd pers. masc. sg. *šɛ*, 3rd pers. fem. sg. *sɛ*, 3rd pers, masc. pl. *šum*, 3rd pers. fem. pl. *sɛn*. In Arabic, all 3rd person pronouns, free and bound, except for the feminine singular, have the vowel *-u-* following the initial *h-*. We have seen how under the influence of a preceding *i*-vowel or diphthong *ay*, this is shifted to *i* in Arabic. However, comparative evidence, as for instance from Akkadian or Modern South Arabian, suggests that the *u*-vowel was probably originally an indicator of the masculine, while *i* was the vowel of the feminine (in the plural, at least): e.g. Proto-Semitic 3rd pers. masc. pl. **šumu* or perhaps better **šunu*, 3rd pers. fem. pl. **šina*. The same can be reconstructed for the 2nd person plural forms, **ʾantumu/*ʾantunu, *ʾantina; *-kumu/*-kunu, *-kina*. Then, the feminine plural forms in Arabic may have acquired their double *-nn-* by addition of the verbal suffix *-na* after the dropping of the final short vowel. The feminine singular forms of Arabic continue the vocalization of their Proto-Semitic antecedents **šiya*, **-šā*.

Classical Arabic shows metrical variant forms of the 3rd person masculine singular bound pronoun with a long vowel, *-hū/-hī*, when added to an open syllable with a short vowel: *dāruhū* 'his house', but *ramāhu* 'he threw him'. Since long vowel forms also occur elsewhere in Semitic, a short/long variation may not be an Arabic innovation but rather an archaism. The 3rd person plural masculine suffixes further have variants *-humū/-himū*, and similarly 2nd person plural masculine *-kumū*, which occur in Classical Arabic poetry (Fischer 1972:126), and which comparison with other Semitic languages also shows to be archaisms.

Lastly, the phonologically controlled variation between the possessive bound 1st singular forms *-ī* and *-ya* also is an inherited feature in Arabic and goes back to Proto-Semitic, as does the separate object form added to finite verbs, *-nī*.

BIBLIOGRAPHICAL REFERENCES

Badawi, Elsaid, Michael G. Carter, and Adrian Gully. 2004. *Modern Written Arabic: A comprehensive grammar*. London and New York: Routledge.
Fischer, Wolfdietrich. 1972. *Grammatik des klassischen Arabisch*. Wiesbaden: O. Harrassowitz.
Fleisch, Henri. 1968a. "Sur les pronoms personnels en arabe classique". *Mélanges de l'Université Saint-Joseph* 44.65–73.
——. 1968b. *L'arabe classique: Esquisse d'une structure linguistique*. Beirut: Dar el-Machreq.
Voigt, Rainer M. 1994. "Der Lautwandel S^1 > H in wurzellosen Morphemen des Alt- und Neusüdarabischen". *Semitic and Cushitic studies*, ed. Gideon Goldenberg and Shlomo Raz, 19–28. Wiesbaden: O. Harrassowitz.
Zaborski, Andrzej. 1995a. "First person pronouns in Arabic in the light of Arabic and Hamito-Semitic dialectology". *Dialectologia arabica: A collection of articles in honour of the sixtieth birthday of Professor Heikki Palva*, 289–294. Helsinki: Finnish Oriental Society.
——. 1995b. "Archaisms and innovations in Arabic independent pronouns". *Proceedings of the 2nd International Conference of l'Association Internationale pour la Dialectologie Arabe*, 269–274. Cambridge.

DAVID L. APPLEYARD
(School of Oriental and African Studies)

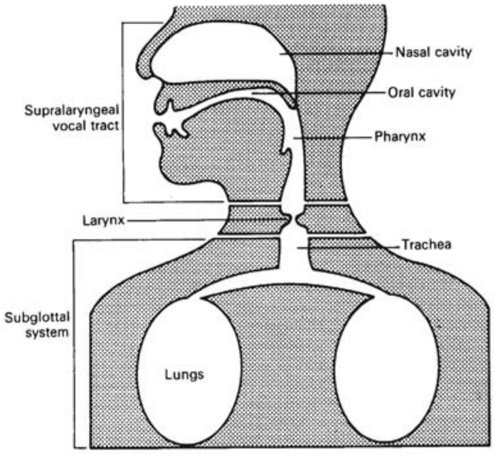
Pharyngealization → Velarization

Phonetics

1. INTRODUCTION

Phonetics is the science of describing the speech sounds of a language. The focus in the following description of Arabic speech sounds is on articulatory and acoustic phonetics. The phonetic characteristics of the vowels, consonants, and syllables of the Arabic language are introduced and described. A brief discussion of the vocal organs follows; these are responsible for the articulation of speech, which is produced by the pulmonic airflow that is transmitted from the lungs to the larynx and then to the oral tract. Literary Arabic (*al-fuṣḥā*), with occasional and brief reference to Arabic dialects, is

Figure 1. This figure illustrates the three physiological components of human speech production (adapted from Lieberman and Blumstein 1988:4).



the focus of this article (practical manuals for the pronunciation of Arabic are Kästner 1981 and Mitchell 1990). A brief reference to the development of Arabic phonetics is included.

2. Vocal organs of speech

Three basic physiological components are involved in the production and articulation of speech sounds (see Fig. 1).

The first component is the lungs, which supply the necessary energy through the mechanism of the respiratory system. The stream of air coming from the lungs is called the pulmonic egressive air. It is essential for the production and transmission of speech sounds. After passing through the lungs, this stream of air then passes through the windpipe and into the larynx. This is the initiation phase in the speech production process.

The larynx, where the vocal folds are located, is the second most important phase in the act of performing speech. The stream of air from the lungs, when it passes through the larynx, activates the vocal folds, which generate an audible acoustic energy. The vocal folds are protected and covered by the epiglottis, which is a flap of tissue that attaches to the front of the larynx. The pulmonic air passes between the vocal folds and flows unobstructed when breathing out; however, when speaking, the vocal folds are engaged and create at least three different states: voiced, voiceless, and whispered. When the vocal folds form a narrow passage between them, the outgoing egressive airflow causes them to vibrate, producing voiced sounds such as [z] in [zaːdɐ] 'he increased'. Voiceless sounds, however, are produced when the vocal folds are held wide apart, such as in the sound [s] as in [saːdɐ] 'he dominated'. The movements and vibrations of the vocal folds happen at a high rate of speed in the process of the production of sounds. The vocal folds may be closed at the glottis. This condition produces the → *hamza* 'glottal stop' [ʔ] in Arabic.

The region above the larynx that extends to the lips is called the vocal tract. This is the third component in the production and articulation of speech sounds. The vocal tract is divided into the oral cavity and the pharynx within the mouth, and the nasal cavity within the nose. The stream of air that passes through the larynx produces audible acoustic energy that encounters various shades of modifications when passing through the oral and nasal cavities. The articulators in the vocal tract form different sounds. The lower articulators, such as the different parts of the tongue and lower lip, move to make different modes of contact with the upper stationary articulators, for instance with the hard palate.

3. The description of Arabic vowels

Normally, Arabic vowels are produced with a relatively free passage in the vocal tract, while the Arabic consonants are produced with a partial or complete obstruction. The vowels are primarily described on the basis of auditory qualities and acoustic judgments, whereas the consonants are described in terms of their manner and points of articulation. The articulators involved in the production of the vowels do not come close to each other in order to make the outgoing air turbulent or to close the vocal tract. In the case of consonants, however, the articulators form a very narrow passage through the vocal tract in the production of the fricative segments and a complete closure in the production of stop segments.

In the production of vowels, two main articulators are involved: the tongue and the lips. The movements of the lower jaw are also important in the production of speech sounds. The tongue assumes a domed shape in the vocal tract during the production of the vowels. It is very versatile in its movements. Front vowels [i] and [iː] are produced when the front of the tongue is raised to form a closure in the front zone of the vocal tract. The airstream flows freely through this type of closure without any noticeable friction. The lips are not directly involved in the production of the front vowels. The back vowels [u] and [uː] are produced when the tongue is raised to the back part of the mouth. A similar closure may be formed that allows the airstream to flow freely without any obstruction. Unlike the front vowels, the lips are involved in the production of back vowels and assume a rounding shape. These back vowels are called back high rounded vowels. The vowels that are produced with the position of the tongue low and central and still assuming a relatively domed shape are called the unrounded central vowels. In the articulation of [a] and [aː], the lips assume a spreading position.

The vowel system of Literary Arabic contains three short vowels, /i/, /u/, and /a/, which contrast phonemically with their long counterparts /ī/, /ū/, and /ā/. The number of vowels varies considerably with the different forms of Arabic dialects. The Iraqi dialect vowel system, for instance, contains five short vowels: /i, u, a, e, o/. The vowels /e/ and /o/, however, have a limited distribution in comparison with the other three vowels, /i, u, a/. The Iraqi dialect also contains five long vowels: /ī, ū, ā, ē, ō/. The two long vowels /ē/ and /ō/, however, have limited distribution. They are primarily the result of a change in pronunciation of words in Literary Arabic that contain the diphthongs /ay/ and /aw/. The Literary Arabic word [bajt] becomes [beːt] 'house', and [lawn] becomes [loːn] 'color'.

What follows is a brief description of the vowels in terms of their auditory qualities, quantities, tongue shapes, and lip movements. The quality of the vowel is an indication of the position of the tongue in the vocal tract. It is the factor that distinguishes one vowel from another. The quantity, on the other hand, refers to the amount of time it takes to produce a vowel.

The /i/ is a short, high front unrounded vowel. It has two allophones: [ɪ], slightly centralized and low, occurring next to emphatic consonants, as in [s̴ɪjaːm] 'fasting'; and [i], occurring elsewhere as in [sinː] tooth.

The /ī/ is a long, high front unrounded vowel. It also has two allophones: [ɪː], slightly low and centralized, occurring adjacent to emphatic consonants, as in [s̴ɪːn] 'China'; and [iː], occurring elsewhere as in [siːn] 'the letter *s* in the Arabic alphabet'.

The /u/ is a short, high back rounded vowel. It has two allophones: a slightly lowered variant [ʊ], occurring next to emphatic consonants, as in [s̴ʊwɑr] 'pictures'; and [u], occurring elsewhere as in [sukuːn] 'tranquility'.

The /ū/ is a long, high back rounded vowel. It has two allophones: a lowered variant between [ʊː] and [ɤː], which occurs next to emphatic consonants, as in [s̴ʊːrɑh] 'picture'; and [uː], occurring elsewhere, as in [suːr] 'fence'.

The /a/ is a low, short central unrounded vowel. It has three allophones (Mitchell 1990: 72–82). The first, [ɑ], occurs next to emphatic consonants and also next to /q/ and /r/; it is centralized and slightly back, as in [s̴ɑlb] 'solid'. The second is a raised front vowel (→ *ʾimāla*), which can be as high as [æ] or even [ɛ] in dialect Arabic (Woidich 2006:8), but in the oral realization of Literary Arabic it is usually between [a] and [æ], as in [salb] 'looting'. The third is very near the cardinal vowel [a] and occurs next to the pharyngeals /ˤ/ and /ḥ/, as in [ħaliːb] 'milk' and [ʕajn] 'eye'. In the phonetic transcription, the last two allophones both appear here as [a]. Before a word boundary, /a/ is realized as [ɐ], except after emphatic consonants; this is not indicated in the transcription here.

The /ā/ is a low, long central unrounded vowel. It has three allophones. The first, [ɑː], occurs next to emphatic consonants and also next to /q/ and /r/; it is lower and further back, as in [t̴ɑːba] 'he became good'. The second allophone, [aː], close to the long cardinal vowel [a], occurs in the neighborhood of pharyngeals, as in [saːħa] 'it melted'. The third is a raised and fronted variant of the latter, between [aː] and [æː], which occurs elsewhere; for the sake of convenience it is transcribed here as [aː], as in [saːla] 'he flowed'.

The description of the vowels above is on the abstract phonemic level. Each vowel phoneme forms an auditory quality space continuum. Its domain embodies its allophones and the free variants. The environment usually determines the allophones and their distribution, while the free variants may occur or not occur depending on the nature of its context.

The vowels /i/ and /u/ differ from their long vowel counterpart /ī/ and /ū/ in quantity, whereas the vowel /a/ differs from its long counterpart /ā/ both in quantity and quality. Arabic vowels assume the traditional triangular shape in the mouth. The high front and back vowels occupy the front and back auditory quality zones, respectively. The low central vowels occupy the bottom auditory space of the triangle. Figure 2 illustrates the triangular shape of the Arabic vowels.

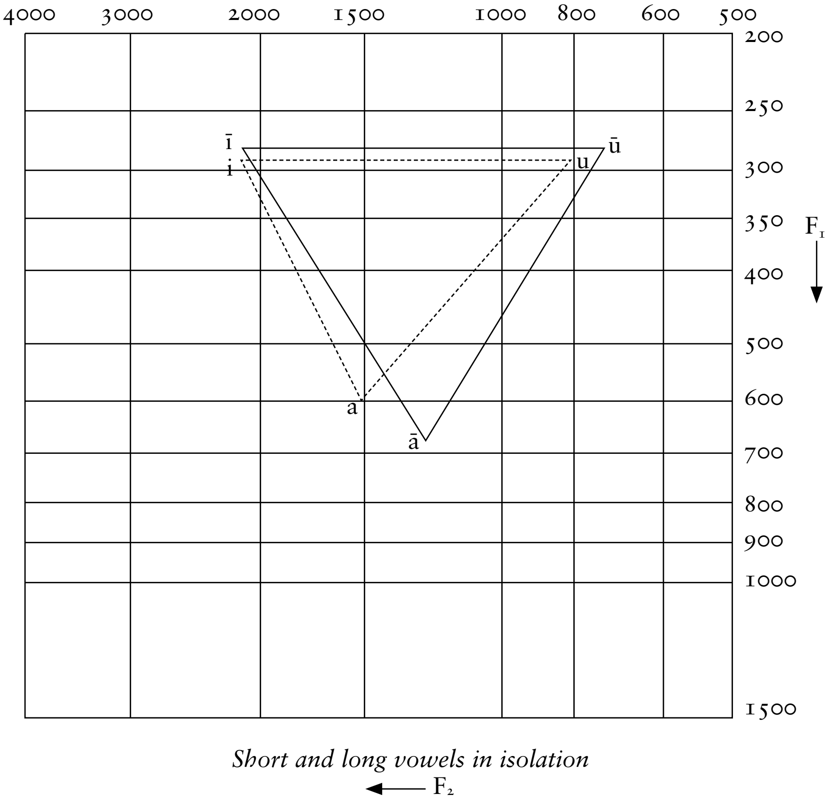
4. The description of Arabic consonants

The speech segments of Literary Arabic utterances consist of sequences of consonants and vowels that are usually arranged in a systematic manner. All Classical Arabic utterances, including syllables, words, phrases, and clauses, begin with a single consonant. This is not the case in most Arabic dialects. The majority of the dialects accept consonant clusters. In the Iraqi dialect, for instance, we find initial consonant clusters, as in [bnajːah] 'little girl'.

The degree of constriction of any consonant segment, when articulated in the vocal tract, is achieved in three phases: the onset, the medial, and the offset. The most prominent phase is the medial or the steady state of a segment. The onset and offset phases play a significant role in speech, especially with regard to the phonetic phenomenon of co-articulation. The phonetic description of Arabic speech sounds focuses on the medial phase of the speech segments.

There are three main groups of consonants: stops (or obstruents), fricatives, and resonants. They are classified on the basis of the degree of stricture in the vocal tract. Stops are articulated with a complete stricture that results in closure of the vocal tract. The fricatives are articulated with a small aperture in the vocal tract that allows the airstream to pass through the vocal tract, creating turbulent friction. The resonant sounds are articulated when the airstream passes through the vocal tract with an

Figure 2. This figure illustrates the articulation positions of long and short Arabic vowels. This triangular diagram of the Arabic vowels is a reproduction of the physiological tracings from an X-ray film made by Al-Ani. It also illustrates the acoustic plotting of formant one (F1) against formant two (F2) of the Arabic vowels. This diagram is adapted from Al-Ani (1970:25).



open aperture so that no turbulent friction is made (Laver 1994:133–134).

The Arabic stops, fricatives, and resonants are introduced and described below according to their place and manner of articulation (see Fig. 3). The term 'place of articulation' does not necessarily mean a specific point through the vocal tract but rather an area, especially in the case of the dento-alveolar consonants. The majority of Arabic consonants are produced in the vocal tract horizontally; however, a small group is produced vertically in the pharynx zone.

4.1 *Stop consonants*

Two main features characterize stops physiologically. The first phase is the formation of a closure within the vocal tract. The movement of one of the active articulators toward a certain point in the mouth forms a constriction with one of the stationary articulators. The second phase is the sudden release of pressure, which results from the driving pressure of the airstream. Oral stops are articulated in the mouth with a velic closure.

Consonant clusters do not occur initially; however, they may occur medially, e.g. [maktab] 'office', and finally, e.g. [bard] 'cold'. Geminates, identical consonant clusters, occur both medially, e.g. /ss/ in [kasːaɍɑ] 'he broke it', and finally, e.g. /dd/ in [sadː] 'dam'. Arabic words illustrating each consonant stop initially, medially, and finally, as well as geminated, are briefly described below. There are eight stops in Arabic. The emphatic consonants are introduced and explained later.

The /b/ is a voiced bilabial stop. The following are illustrative words of /b/: [baːb] 'door', [qublah] 'kiss', [sabːaba] 'he caused something to happen', [sabab] 'cause'. In the Iraqi dialect, the voiceless /p/ occurs in a few words, mostly foreign lexical items from Persian, as in [pɑɍda] 'curtain'.

The dento-alveolar area contains four stops. They contrast on a binary basis for the voiced /d/ and voiceless /t/. Emphatic /ṭ/ contrasts with nonemphatic /t/, and likewise /ḍ/ with /d/.

The /t/ is a voiceless dento-alveolar aspirated stop. Aspiration is a phonetic gesture that is manifested as a period of voicelessness of the articulation as shown on spectrograms. In narrow phonetic transcription, it may be symbolized by superscript [ʰ], as in [tʰaːba] 'he repented'. When released in the final position, it appears as a burst. Examples: [taɍɑka] 'he left', [kataba] 'he wrote', [katːaba] 'he dictated', [bajt] 'house'.

The /d/ is a voiced dento-alveolar unaspirated stop. Examples: [darb] 'road', [bard] 'cold', [madːa] 'he stretched', [bariːd] 'mail'.

Figure 3.

	Horizontal place of articulation								
	Bilabial	Labiodental	Interdental	Dental-alveolar	Palatal	Velar	Uvular		Vertical place of articulation
Stops voiceless				t ŧ		k	q		
Stops voiced	b			d đ	ʒ				
Fricatives voiceless		f	θ	s ꞩ	ʃ	x			
Fricatives voiced			ð ð̵	z			ɣ		
Nasals	m			n					
Trills				r					
Laterals				l ɬ					
Approximate	w				j				
								ħ ʕ	Pharyngeal
								h ʔ	Glottal

The /k/ is a velar voiceless stop that is aspirated in the same manner as /t/. Examples: [kalb] ‘dog’, [fikɍɑ] ‘idea’, [bakːaːhu] ‘he made him cry’, [malik] ‘king’.

The /q/ is a uvular voiceless unaspirated stop. Examples: [qalb] ‘heart’, [faqiːr] ‘poor’, [taqɑdːama] ‘he came forward’, [bɑɍq] ‘lightning’.

4.2 *Affricate*

There is only one voiced affricate consonant in Standard Arabic. However, in some Arabic dialects, there is a voiceless affricate. The affricate sound is defined as a stop followed by a homorganic fricative. This is the case in Iraqi speech, where the sound /k/ is often pronounced as /č/, so that the Standard Arabic word for dog [kalb] is pronounced as [tʃalib].

The /j/ is a voiced affricate that is a combination of two phones. The first is a stop, followed by a fricative. Affricate sounds are unique in this manner. The /j/ appears on the spectrogram as a stop followed by a fricative. This sound is pronounced as /g/ in the Egyptian dialect. Thus, the word [dʒamaːl] ‘name of a person’ is pronounced as [gamaːl]. More examples: [dʒabal] ‘mountain’, [dadʒaːdʒah] ‘chicken’, [zadʒːa] ‘he pushed, hurled something’, [zudʒaːdʒ] ‘glass’.

4.3 *Fricative consonants*

Fricative consonants are produced in the vocal tract by narrowing the distance between two articulators. This process makes a constriction that causes the airstream to be partially obstructed and to be consistently turbulent. Acoustically, fricatives, especially voiceless, possess a high random noise. The Arabic fricatives are listed with their descriptions and illustrative examples as follows:

The /f/ is a voiceless labiodental fricative. Labiodental articulation involves the lower and the upper teeth. There is no voiced counterpart /v/ in Arabic. Examples: [fam] ‘mouth’, [safɑɍ] ‘traveling’, [nafːasa] ‘he appeased’, [sajf] ‘sword’.

The /ṯ/ is a voiceless interdental fricative. Interdental articulation involves placing the tip of tongue between the upper and the lower teeth. Acoustically, it appears as a random noise. Examples: [θuːm] ‘garlic’, [kaθɍɑ] ‘abundance’, [kaθːafa] ‘he made thick’, [kaθː] ‘shabby; old’.

The /ḏ/ is a voiced interdental fricative. Acoustically, it appears as a random noise. Examples: [ðahab] ‘gold’, [qɑðafa] ‘he threw away’, [kaðːaːb] ‘liar’, [nufuːð] ‘influence’.

The /s/ is a voiceless dento-alveolar sibilant fricative. Examples: [safɑɍ] ‘traveling’, [fasada] ‘he became corrupt’, [nafas] ‘breath’, [kaːbuːs] ‘nightmare’.

The /š/ is a voiceless palato-alveolar fricative. Examples: [ʃadʒaɍɑ] ‘tree’, [muʃkila] ‘problem’, [baʃːaɍɑ] ‘he brought good news’, [ɍɑʃː] ‘sprinkling’, [kabʃ] ‘ram’.

The /x/ is a voiceless velar fricative. Examples: [xubz] ‘bread’, [naxlah] ‘palm tree’, [saxːaɍɑ] ‘he subjugated’, [manaːx] ‘weather’.

The /ġ/ is a voiced uvular fricative. Examples: [yɑɍb] ‘west’, [maɣrib] ‘evening’, [ɍɑyːaba] ‘he aroused someone’s interest’, [balaɣa] ‘to reach’, [fɑɍɑːɣ] ‘vacuum, empty space’.

4.4 *Nasal consonants*

The production of nasal consonants involves the nasal and oral cavities (→ nasalization). It is this unique combination that distinguishes nasals from other subgroups of consonants such as oral stops and fricatives. Nasality could be defined physiologically as the formation of a closure at the oral cavity while the air flows through the nose. The soft palate or velum position is involved, whether a nasal or an oral sound is produced. When the velum is raised, making a closure against the upper back of the pharynx, oral sounds are produced. Lowering of the velum, however, allows the airflow to travel through the nose, and as a result nasal sounds are made.

There are only two nasal sounds in Arabic: /m/ and /n/. The /m/ is a voiced bilabial nasal and is produced with a closure of the lips while the velum is lowered, thus allowing the airflow to pass through the nose. Examples: [malik] ‘king’, [samakah] ‘fish’, [damːaɍɑ] ‘he destroyed’, [nawm] ‘sleep’. The /n/ is a dento-alveolar nasal. Like the articulation of the /m/, the velum is lowered and the oral closure is made at the dento-alveolar zone. Both /m/ and /n/ occur initially, medially, and finally. They also occur geminated, medially, and finally. Examples: [naːma] ‘he slept’, [dʒanuːb] ‘south’, [qɑnːana] ‘he made a law’, [min] ‘from’. Acoustically, the /n/ possesses a weak resonance that appears as a formant structure along the baseline of a spectrogram. The first formant of both /m/ and /n/ is about 250 cps (cycles per second). The second formant of /m/ is around 1000 cps, whereas the second formant of /n/ is higher, around 1550 cps.

4.5 *Trill*

The /r/ is a voiced dento-alveolar trill. The trill articulation is achieved by holding the front part of the tongue loosely against the alveolar ridge so that the flow of air between them creates motion, alternately sucking them together and blowing them apart. Examples: [r̴ɑdʒul] ‘man’, [bard] ‘cold’, [barːada] ‘he made something cold’, [bɑr̴ɑda] ‘he filed’, [qɪdr] ‘cooking pot’.

4.6 *Laterals*

The /l/ is a voiced dento-alveolar lateral. In the articulation of lateral sounds, the airstream flows over the sides of the tongue. Acoustically, this sound possesses vowel-like formant structures. Examples: [lajl] ‘night’, [salaːm] ‘peace’, [salːama] ‘he submitted’, [dʒamal] ‘camel’.

4.7 *Semivowels and diphthongs*

The semivowel sounds are articulated in the same way as vowels; however, they do not form syllables on their own as the vowels do. There are two diphthongs in Arabic: /ay/ and /aw/, as in [kajfa] ‘how’, [law] ‘if’.

There are only two semivowels in Arabic. The /w/ is a voiced labiovelar semivowel. Acoustically, it possesses distinct vowel-like formant structures similar to those of the vowels /u/ and /ū/. The /w/ forms a diphthong with the vowel /a/, as in [lawm] ‘blame’. The /y/ is a voiced palatal semivowel. Acoustically, it possesses distinct vowel-like structures similar to those of the vowels /i/ and /ī/. It also forms a diphthong with the vowel /a/, as in [bajt] ‘house’.

4.8 *Pharyngeal and glottal consonants*

The pharyngeal and glottal consonants are distinguished from the rest of the Arabic consonants by having distinct vertical places of articulation. A vertical place of articulation is a set of anatomical locations from the palate to the glottis, and a horizontal place of articulation is from the lips to the uvula. The pharyngeal articulation usually involves the root of the tongue and the back wall of the pharynx, whereas glottal articulation involves the glottis. There are four consonants in this subgroup:

The /h/ is a voiceless oral fricative. Examples: [hilaːl] ‘crescent’, [dʒahl] ‘ignorance’, [sahːala] ‘he made easy’, [kurh] ‘hatred’.

The /ḥ/ is a voiceless pharyngeal constricted fricative. It is produced with a constriction formed at the dorsum of the tongue against the posterior wall of the pharynx. The movement of the pharyngeal muscles plays an important role in forming a narrow constriction passage that gives the /ḥ/ its distinct character. The constriction and the movements of muscles in the pharynx are considered the primary factors that differentiate /ḥ/ from /ʿ/. Examples: [ħarb] ‘war’, [saħar] ‘dawn’, [saħːaqɑ] ‘he crushed’, [fɑr̴ɑħ] ‘happiness’.

The /ʿ/ is a voiced pharyngeal fricative. In Iraq and the Arabian Peninsula, the /ʿ/ is a voiceless pharyngeal stop. This is especially the case when it occurs geminated. Examples: [ʕar̴ɑb] ‘an Arab’, [r̴ɑʕd] ‘thunder’, [zaʕːala] ‘he annoyed’, [zarʕ] ‘plantation’.

The /ʾ/ is a glottal stop (→ *hamza*). This sound is produced at the glottis with a very quick opening and closing action. Therefore, it cannot be considered voiced or voiceless. Examples: [ʔamal] ‘hope’, [saʔala] ‘he asked’, [taʔadːaba] ‘he was well educated’, [maːʔ] ‘water’.

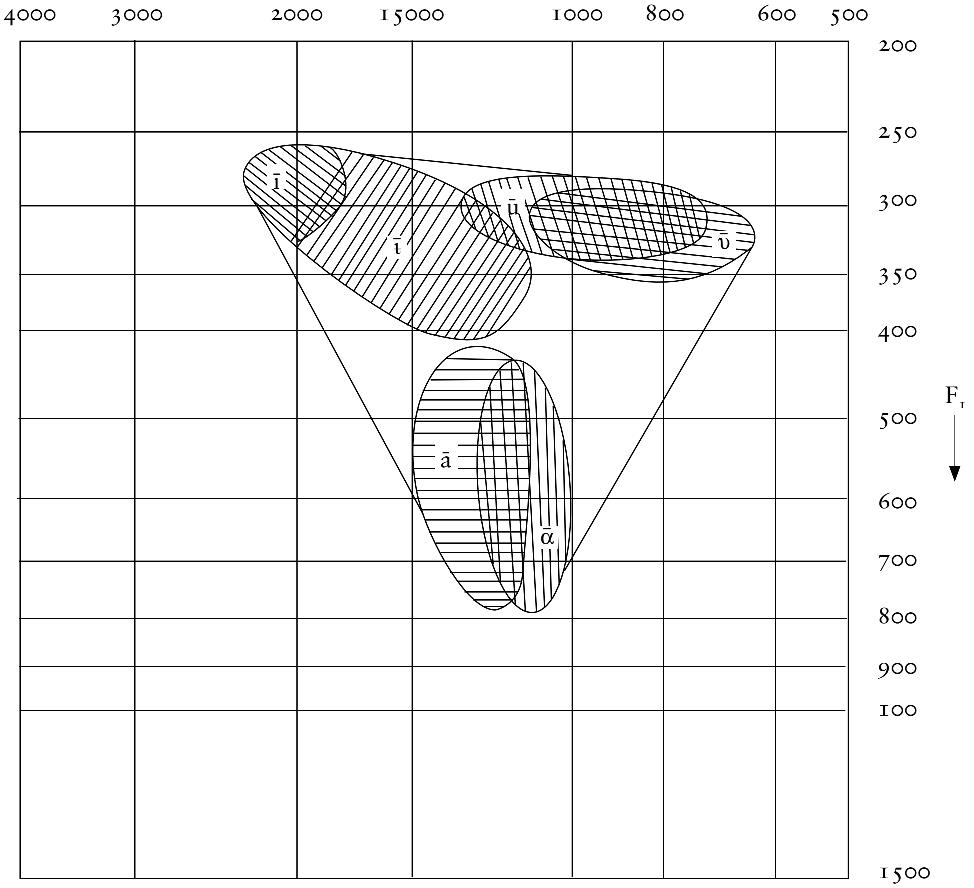
4.9 *Emphatic consonants*

The phenomenon of ‘emphasis’ is a phonetic feature characterized by having two points of articulation. The primary point is in the dento-alveolar area, and the secondary point engages the upper region of the pharynx. Several terms have been used to refer to these consonants, the most common being velarized, pharyngealized, retracted tongue root, and emphatic (→ velarization). Here, the term ‘emphatic’ is used to refer to the phenomenon of emphasis, which involves the manner of articulation of the traditional emphatic consonants /ṭ, ḍ, ṣ, ḏ̣/. They are transcribed here with a mid-tilde as velarized/pharyngealized [t̴, d̴, s̴, ð̴] rather than pharyngealized [tˤ, dˤ, sˤ, ðˤ] or velarized [tˠ, dˠ, sˠ, ðˠ] sounds. This subgroup of coronal consonants contrasts phonemically with the plain consonants /t, d, s, ḏ/. The secondary articulation feature gives the emphatics their unique characteristics. Secondary articulation is defined as “an articulation of a lesser degree of stricture accompanying a primary articulation of a higher degree" (Ladefoged and Maddieson 1996:354). It is important to emphasize the simultaneity of the occurrence of these two articulations.

Marginal emphatics may be found in most Arabic dialects. Their occurrence, number, and spreading depend, to a varying degree, on the social levels of speaking and the geographical region. The following marginal emphatics are the most common: /ḅ, ṃ, ṛ, ḷ/. These consonants occur only in a few words in Arabic. They are found adjacent to the low central vowels /a/ and /ā/. These marginal emphatics are not represented by alphabet symbols in the Arabic orthography. The /ḷ/ occurs in the word for God, [ʔɑɫːɑːh]. The phonemic contrast between the plain /l/ and the emphatic /ḷ/ is supported by the following minimal pair: [wɑɫːɑːh] 'by God' and [walːaːh] 'he appointed him'. The emphatic /ṃ/ contrasts with the plain consonants, as in [m̶ɑjː] 'water' and [maj] 'girl's name'. The emphatic consonants are known to condition neighboring sounds. This conditioning is often called spreading (→ *ʾiṭbāq*). There are two types of spreading: progressive, moving forward, and regressive, moving backward. The progressive spreading is the most common, whereas regressive is very rare. The conditioning of spreading covers the whole syllable and sometimes the whole word. The spreading condition occurs with both adjacent vowels and consonants. The acoustic effect on the adjacent sounds, especially the vowels, is to lower their frequencies (see Fig. 4).

The feature of emphasis attracted the attention of early Arab philologists as well as modern linguists. In 1957, Roman Jakobson wrote an article, "*Mufaxxama*: The 'emphatic' phonemes in Arabic". He presented the theoretical feature of emphasis as a distinctive feature in the features system. The emphasis is traditionally called → *tafxīm* 'thickness'. This was first mentioned in

Figure 4. This figure illustrates the long Arabic vowels adjacent to emphatic and nonemphatic consonants in overlapping positions (Al-Ani 1970:49).



the 8th century by the well-known grammarian Sībawayhi (d. 177/793–794), the author of the famous Arabic grammar *al-Kitāb*.

The four emphatics are described below:

The /ṭ/ is a voiceless emphatic unaspirated dento-alveolar stop. When released in final position, it appears as a burst. Examples: [ᵵabiːb] ‘doctor’, [baᵵal] ‘hero’, [baᵵːaːl] ‘unemployed’, [ᵳabᵵ] ‘tying’.

The /ḍ/ is a voiced emphatic interdental stop. Examples: [ᵭaᵳb] ‘beating’, [jaᵭribu] ‘he beats, strikes’, [faᵭːala] ‘he preferred’, [faāᵭ] ‘duty; decree’. The sound /ḍ/ is not pronounced as an emphatic stop in Iraq, the Arabian Peninsula region, or Bedouin and rural dialects, but rather as an emphatic interdental fricative. Also, it is pronounced as a dento-alveolar emphatic fricative in the urban areas in Egypt and the Levant. The Arabic language is called the language of the → *ḍād* [ᵭaːd], possibly because of the unique lateralized pronunciation this sound may have had in the earliest form of Classical Arabic, which is why the Arabs considered it to be the most distinguished and unique sound of the Arabic sound system.

The /ḏ̣/ is a voiced emphatic interdental fricative. Examples: [ð̴ʊhr] ‘noon’, [mað̴luːm] ‘unjustly treated’, [nað̴ːama] ‘to arrange’, [lafð̴] ‘expression’. It is usually realized as [ᵶ].

The /ṣ/ is a voiceless dento-alveolar emphatic sibilant fricative. Examples: [ᵴadr] ‘chest’, [qaᵴr] ‘palace’, [qaᵴːa] ‘to cut’, [nuᵴʊːᵴ] ‘texts’.

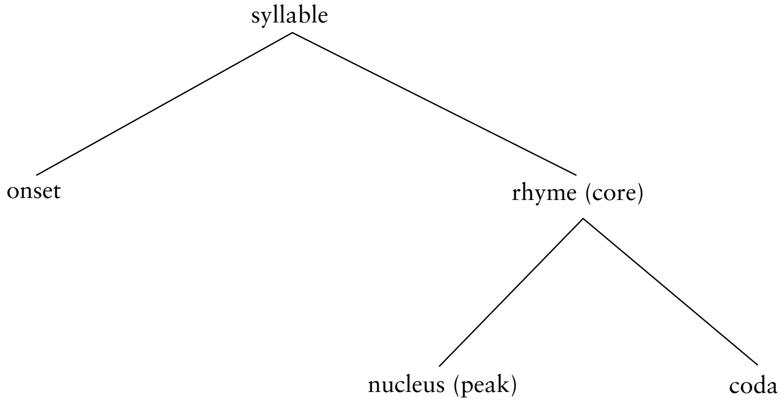
5. Syllable structure and types

Arabic utterances begin with a single consonant followed by short or long vowels. Consonant clusters occur medially and finally, and no vowel clusters are allowed anywhere in the sound structure of Arabic utterances. The → syllable structure may be expressed in the linear sequences of consonants and vowels. In the configuration of the structure of the syllable, the C stands for consonant and v for short vowels. Consonant clusters are indicated by CC and long vowels by V̄.

The structure of the syllable, expressed in Figure 5, consists of the onset and the rhyme. The onset of all the syllable types always consists of a single consonant. The rhyme encompasses the nucleus, which is always a vowel and any consonant(s). The nucleus is an obligatory component of the syllable structure. It may be followed by one, two, or no consonants. The prominence of the syllable is achieved through its weight. Syllable weight encompasses two kinds of syllables: light and heavy. A light syllable’s rhyme consists of a nucleus of a short vowel followed by a maximum of one consonant. A heavy syllable is any other type of syllable (Laver 1994:517–518).

The structure of Arabic consists of six syllable types: Cv, CvC, Cv̄, Cv̄C, CvCC, and Cv̄CC. They are classified into subgroups:

Figure 5. This figure illustrates the metrical structure of the syllables. It is divided into two basic components, the onset and the rhyme. The onsets of all Arabic syllable types and utterances begin with a single consonant. The configuration of the rhyme consists of the vowel and any consonant (Laver 1994:517–518).



short, medium, and long. The syllable Cv is short, e.g. [bi] ‘in’. There are two medium syllables: CvC [kam] ‘how many’ and Cv̄ [maː] ‘what’. The following syllables are considered long: Cv̄C [maːʔ] ‘water’, CvCC [bard] ‘cold’, Cv̄CC [maːdː] ‘stretching’.

The syllables may also be classified as open or closed. The syllables that end in a vowel are called open syllables, like Cv and Cv̄, whereas those that end in a consonant are closed, like CvC, CvCC, Cv̄C, and Cv̄CC.

The first three syllable types, Cv, CvC, Cv̄, are the most frequent and reflect the dynamic nature of the composition of Arabic words and phrases. The frequency and distribution of CvCC and Cv̄CC, in comparison to the rest of the other syllables, are restricted and occur only finally and in pause forms.

The most prominent elements of the syllable are the vowels, short or long. The vowels form the nucleus and the consonants the marginal elements in the configuration of the structure of the syllable. The number of syllables in Arabic utterances is equal to the number of vowels.

6. History of Arabic phonetics

Arab grammarians and orthoepists devoted special attention to Arabic phonetics in their writings (Gairdner 1925, 1935). Throughout the long history of the Arabic language, three scholars are credited for their creative and original contributions to the study of Arabic grammar and phonetics. The earliest work on Arabic phonetics is a treatise by the most celebrated Arab scholar and linguist, al-Xalīl ibn ʾAḥmad al-Farāhīdī (d. 175/791). He was followed by his faithful student Sībawayhi (d. 180/796), who wrote the most celebrated book on Arabic grammar, titled simply *al-Kitāb* ‘the book’. The last chapter of the book contains a meticulous and highly systematic phonetic description of the Arabic sounds (Al-Nassir 1993). Almost all the work on Arabic phonetics by later scholars can be considered a form of elaboration and explanation of Sībawayhi. The only exception is Ibn Sīnā (11th c.), whose treatise on Arabic phonetics, *Risāla fī ʾasbāb ḥudūṯ al-ḥurūf*, did not follow the traditional approach of Sībawayhi’s *Kitāb*. What follows is a brief mention of these three works.

The treatise on Arabic phonetics by al-Xalīl is actually an introduction to his Arabic dictionary, known as *Kitāb al-ʿayn*. This treatise is considered the earliest source on Arabic phonetics that provides a description of the Arabic sounds. The author arranged the lexical items of his dictionary on articulatory parameters. It started with the sound *ʿayn* [ʕ], articulated at the larynx. The rest of the sounds are arranged in the vocal tract, in measured steps toward the lips according to their points and manner of articulation. Al-Xalīl devised a special procedure of placing the letter *ʾalif* /ā/ before each sound and then proceeded to pronounce it, as in /āb/ and /āq/, and so on. In this manner he was able to specify the articulatory characteristics of the Arabic sounds.

Sībawayhi’s systematic phonetic description and classification of the Arabic sounds was completed just shortly after the death of his teacher. It is considered the most thorough and complete description. Nothing substantial has been added to Sībawayhi’s system. Most of the changes that have come down to us are in the form of modifications and interpretations.

In his description of the Arabic *ḥurūf* ‘letters; sounds’, Sībawayhi made a distinction between two major categories of the Arabic consonants: the → *majhūra* ‘voiced’ and *mahmūsa* ‘voiceless’. Modern linguists have been trying to explain these two concept terms. The voiced/voiceless dichotomy does not seem to fit exactly within Sībawayhi’s definition of the two terms. Several possible terms have been advanced as possible translations instead of ‘voiced’ and ‘voiceless’, for instance the terms → *majhūra* ‘breathed’ and *mahmūsa* ‘non-breathed’. The ten *mahmūsa* consonants are: /h, ḥ, x, k, š, s, t, ṣ, ṯ, f/, as described by Sībawayhi, which agrees with their present classification and description as being voiceless. He emphasized two distinguishing criteria for the description of the *mahmūsa*. The first is the amount of breath that flows with articulation of the sound, and the second is the ability to repeat the sound with the outgoing breath. The remaining nineteen consonants are described as *majhūra*, with the exception of three consonants. These are the → *hamza*, /q/, and /ṭ/. Today, the articulation and description of the sounds /q/ and /ṭ/ is voiceless and unaspirated. Aspiration is defined as a period of voicelessness after the release of an articulation. It is possible that the feature of the lack of aspiration led Sībawayhi to classify them as voiced.

The *Risāla* of Ibn Sīnā is considered unique in the Arabic phonetic tradition (Al-Ani 1993). This great scholar, physician, and philosopher presented and pursued a different approach for the description of the Arabic sounds. He classified the Arabic sounds as *mufrada* 'unitary' and *murakkaba* 'compound'. The former are characterized in their production by a *ḥabasa tāmma li-ṣ-ṣawt* 'complete closure' of the airstream. The second group consisted of continuant units. Ibn Sīnā did not make any reference to the more familiar phonetic terms *mahmūsa* and *majhūra* of Sībawayhi.

Ibn Sīnā explained the physics of speech and how sounds are produced in the physical sense. Also, he explained and analyzed the different parts of the larynx and its functions. Then he proceeded to describe the Arabic sounds one by one. He also compared the Arabic sounds with the sounds of other languages, such as Persian and Turkish. The treatise contains about forty pages and has been translated into several languages, including English and Persian.

Bibliographical references

Primary sources

Ibn Sīnā, *Risāla* = ʾAbū ʿAlī al-Ḥusayn ibn ʿAbdallāh Ibn Sīnā, *Risāla fī ʾasbāb ḥudūṯ al-ḥurūf*. Ed. Muḥammad Ḥassān aṭ-Ṭayyān and Yaḥyā Mīr ʿAlam. Damascus: Majmaʿ al-Luġa al-ʿArabiyya, 1983.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. Ed. ʿAbd as-Salām Muḥammad Hārūn. 5 vols. Repr. Beirut: ʿĀlam al-Kutub, 1975.

Xalīl, *ʿAyn* = ʾAbū ʿAbd ar-Raḥmān al-Xalīl ibn ʾAḥmad al-Farāhīdī, *Kitāb al-ʿayn*. Ed. Mahdī al-Maxzūmī and ʾIbrāhīm as-Sāmarrāʾī. 8 vols. Beirut: Muʾassasat al-ʾAʿlamī li-l-Maṭbūʿāt, 1988.

Secondary sources

Al-Ani, Salman H. 1970. *Arabic phonology: An acoustical and physiological investigation.* The Hague and Paris: Mouton.

——. (ed.). 1978. *Readings in Arabic linguistics*, I. Bloomington: Indiana University, Linguistics Club. (Repr., 1992.)

——. 1993. "The *Treatise* of Ibn Sina: An original contribution to Arabic phonetics". *Festschrift in honor of Ernest McCarus: Current issues in Arabic studies,* 132–153. Salt Lake City: University of Utah Press.

Al-Nassir, A.A. [ʿAbd al-Munʿim ʿAbd al-ʾAmīr an-Nāṣir]. 1993. *Sibawayh the phonologist: A critical study of the phonetic and phonological theory of Sibawayh as presented in his treatise* Al-Kitab. London and New York: Kegan Paul International.

Gairdner, William Henry Temple. 1925. *Phonetics of Arabic*. London: Oxford University Press.

——. 1935. "The Arab phoneticians on the consonants and the vowels". *Moslem World* 25.242–257. (Repr., Al-Ani 1978:187–202.)

Jakobson, Roman. 1957. "*Mufaxxama*: The 'emphatic' phonemes in Arabic". *Studies presented to Joshua Whatmough on his sixtieth birthday,* ed. Ernst Pulgram, 105–115. The Hague and Paris: Mouton.

Kästner, Hartmut. 1981. *Phonetik und Phonologie des modernen Hocharabisch*. Leipzig: Enzyklopädie. (2nd ed., 1987.)

Ladefoged, Peter. 2001a. *A course in phonetics*. 4th ed. New York: Harcourt College Publishers.

——. 2001b. *Vowels and consonants: An introduction to the sounds of languages*. Oxford: Blackwell.

——, and Ian Maddieson. 1996. *The sounds of the world's languages*. Oxford: Blackwell.

Laver, John. 1994. *Principles of phonetics*. New York: Cambridge University Press.

Lieberman, Philip and Sheila A. Blumstein. 1988. *Speech physiology, speech perception, and acoustic phonetics*. Cambridge: Cambridge University Press.

Mitchell, T.F. 1990. *Pronouncing Arabic*, I. Oxford: Clarendon Press.

Woidich, Manfred. 2006. *Das Kairenisch-Arabische: Eine Grammatik*. Wiesbaden: O. Harrassowitz.

Salman H. Al-Ani (Indiana University)

Phonological Merger

Phonological merger may be defined as a structural change in the sound system of a language by which previously distinct phonemes become one phoneme. Therefore, merger results in the elimination of distinctions, the reduction of distinctive word classes (or lexical sets), and the loss of information. Mergers are very common in human languages, and more common than their counterpart processes which preserve distinctions (e.g. chain shifts) or create distinctions (e.g. → phonological splits). An accepted property of mergers is that they always expand at the expense of distinctions, a principle which Labov (1994) calls 'Herzog's Principle'.

There are many debates as to the causation of mergers. Among the earliest proposals are those put forward by Martinet (1955), which stipulated that *functional load* was an important factor, such that phonemes with low functional load are likely to merge. The notion of 'functional load' refers to the relative importance of a distinction in the maintenance of a phonemic opposition. Although there is no agreed-upon way to measure functional load, it is usually determined by such factors as the number of minimal pairs that depend on the distinction, the extent to which the distinction depends on minimal pairs, and the number of phonetic features on which the opposition depends (see Labov 1994:327–329).

A good illustration of merger in Arabic is the change from fricative interdental to dental stop in many Mashreqi and Maghrebi dialects. In these dialects, the change from plain interdental to stop resulted in the merger of the following phonemes:

/t̲/ + /t/ → /t/, e.g. [θaːni] → [taːni] 'second'
/d̲/ + /d/ → /d/, e.g. [ðanab] → [danab] 'tail'

In relation to these mergers, it is first worth noting that the functional load of the interdental-stop opposition in Arabic is probably relatively low. For example, minimal pairs that depend on the opposition between these sounds are hard to come by. The mergers resulted in the disappearance of /t̲/ and /d̲/ and the creation of expanded lexical sets for /t/ and /d/. Historically, the interdental-stop merger is quite old and can be traced to the ancient Arabic dialects inside Arabia (see El-Gindi 1983:432–435). These mergers have continued to be operative in some of the modern dialects of Arabic. They have gone to completion in the dialects of Damascus, Beirut, Jerusalem, Cairo, Algiers, and Rabat (see Cantineau 1960), and are progressing vigorously in the major cities of Jordan (→ Jordanian Arabic; Al-Wer 2003).

Written Classical Arabic contains a third interdental phoneme, /ḏ̣/ [ð̴], which stands in phonemic contrast to stop /ḍ/ [d̴]. It is noticeable that this opposition does not exist in the spoken Arabic dialects. Instead, we find the following distribution:

(i) Dialects that contain the interdental plain sounds; these dialects have the emphatic interdental phoneme in words with etymological interdental, as well as in lexical items with original Arabic → *ḍād*; and
(ii) Dialects that do not have the interdental sounds and only have the phoneme /ḍ/, stop [d̴], in lexical items with original Arabic *ḍād*, as well as in those with original [ð̴].

Al-Wer (2003; see also Corriente 1978) suggests the following scenario of events for the developments that have affected the emphatic phonemes. First, original Arabic *ḍād*, which was an emphatic lateral fricative sound, shifted to an emphatic interdental and thus merged with /ḏ̣/. Secondly, the emphatic interdental in both sets of lexical items changed to its counterpart, the emphatic stop /ḍ/, but there was no merger involved in this development since the system did not contain /ḍ/. These events probably took place in pre-Islamic times, since there exist reports of the use of either the interdental or the stop for both lexical sets in the ancient *maṣāḥif* and in poetry (El-Gindi 1983:426–429).

Phonological merger is a type of sound change, and, therefore, the principles and parameters that govern sound change in general will be expected to apply to sound changes that lead to mergers. The issue of how sound change proceeds has been a central query in historical philology since the 19th century, as articulated in the Neogrammarians' doctrines. It continues to be a focal issue in modern linguistics, especially in historical linguistics, phonology, and variationist sociolinguistics.

The first and chronologically older view is rooted in the Neogrammarians' theory of sound change. It stipulates that sound change is phonetically gradual and lexically sudden. In this view, the change is exceptionless and subject only to phonetic conditioning. It is lexically sudden in the sense that all of the lexical items containing the sounds in question undergo the change at the same rate. Using the example of Arabic /t̲/, the Neogrammatical view would predict that once the (gradual) phonetic change has been completed, items such as /t̲āni/ 'second', /lit̲a/ 'tooth gum', /t̲awra/ 'revolution', and /mat̲alan/ 'for example' were all transferred to the lexical sets which contained /t/ or /s/, leaving no items containing /t̲/ unaffected.

The second view, associated with Wang (1969), is the lexical diffusion model, which holds that sound change is phonetically sudden but lexically gradual; lexical items are transferred word by word from one class to another, and sound change can affect only part of a lexical set, leaving items unaffected by the change. Within the framework of the lexical diffusion model, the Arabic items mentioned above would be transferred to the new class one at a time and at different rates, and residue items may remain in the old class. Kiparsky (1995) has argued that the Neogrammarians' position and the lexical diffusion model can be reconciled in terms of lexical phonology. He argues that lexical diffusion is a type of analogical change and that the process is "driven by rules of the lexical phonology" (see also Milroy 1992 for a review of these arguments).

Variationist sociolinguistics, as articulated by Labov (1994, 2001), Milroy (1992), and Chambers (1995), among others, have amassed considerable amounts of empirical data in favor of both the Neogrammarians' doctrine and lexical diffusion, although, quantitatively, the evidence in favor of the former is more substantial. The contribution of sociolinguistics, however, goes beyond substantiating or refuting existing positions. Most notably, the data available from variationist investigations identify two types of sound change: (i) regular sound change or change from below, which is "characteristic of the initial stages of a change that develops within the system, without lexical or grammatical conditioning or any degree of social awareness" (Labov 1994:542); and (ii) lexical diffusion or change from above, which is "the result of the abrupt substitution of one phoneme for another in words that contain that phoneme" (Labov 1994:542). With respect to sound changes that lead to mergers, variationists identify several types of mergers that seem to operate at different rates. These are (i) merger by approximation, which involves a gradual phonetic approximation of the two phonemes until they become nondistinct; (ii) merger by transfer, through which words move gradually from one lexical set into another; and (iii) merger by expansion, which is typical of cases where the two sounds are in proximity in phonological space. The outcome of such a merger is a phoneme that occupies the space of both sounds (see Labov 1994:321–323). An important further contribution of the sociolinguistic view of sound change is the incorporation of social factors in the trajectory of sound change across the community (see Milroy 1992, Chap. 7; Labov 1994:300–302).

It is an accepted and empirically tested principle in the study of language change that mergers are irreversible. What this means in practice is that it is conceptually impossible for native speakers to unmerge a merged word class. As an illustration, consider what would be involved in the unmerging of the /t/ lexical set for speakers of Arabic dialects for whom the set contains items with etymological /ṯ/ as well as items with etymological /t/. Given that there is no rule by which the /t/ and /ṯ/ items can be grouped into separate sets, the only way is to unlearn and relearn each item separately. The speaker will have to unlearn that [taːni] 'second' does not contain /t/, and to learn that it contains /ṯ/, and that [tamir] 'dates' contains /t/ and it does not contain /ṯ/ (in addition to learning a sound [θ] that is not part of the phonetic inventory). While it is conceivable for an individual to unmerge these sounds through active learning, it is inconceivable for an entire community of speakers to achieve this and thus to restore the original distribution.

BIBLIOGRAPHICAL REFERENCES

Al-Wer, Enam. 2003. "Variability reproduced: A variationist view of the /ḏ̣/-/ḍ/ opposition in modern Arabic dialects". *Approaches to Arabic dialects: A collection of articles presented to Manfred Woidich on his sixtieth birthday*, ed. Martine Haak, Rudolf de Jong, and Kees Versteegh, 21–31. Leiden: E.J. Brill.

Cantineau, Jean. 1960. *Cours de phonétique arabe*. Paris: C. Klincksieck.

Chambers, J.K. 1995. *Sociolinguistic theory*. Oxford: Blackwell.

Corriente, Frederico. 1978. "Ḍ-L doublets in Classical Arabic as evidence of the process of de-lateralisation of *ḍād* and development of its standard reflex". *Journal of Semitic Studies* 23:1.50–55.

El-Gindi, Ahmad A. 1983. *al-Lahajāt al-ʿarabiyya fī t-turāṯ*. Tripoli: ad-Dār al-ʿArabiyya li-l-Kitāb.

Kiparsky, Paul. 1995. "The phonological basis of sound change". *The handbook of phonological theory*, 640–670, ed. John A. Goldsmith. Oxford: Blackwell.

Labov, William. 1994. *Principles of linguistic change: Internal factors*. Oxford: Blackwell.

——. 2001. *Principles of linguistic change: Social factors*. Oxford and Malden: Blackwell.

Martinet, André. 1955. *Economie des changements phonétiques*. Bern: Francke.

Milroy, James. 1992. *Linguistic variation and change*. Oxford: Blackwell.

Wang, William. "Competing sound changes as a cause of residue". *Language* 45.9–25.

ENAM AL-WER (University of Essex)

Phonological Split

Phonological split is a sound change that leads to additions or alterations in the system of distinctions. In this process, one phoneme divides into two phonemes, which is precisely the opposite of → phonological merger. Strictly speaking, splits are the result of the phonemicization of preexisting allophonic variations in the system. It is this type of split which is usually referred to in the study of sound change.

The term 'split' is sometimes also used in cases where no clear phonemic distinction occurs, i.e. where the phonological structure of the language is not affected by the sound change. In these cases the term 'allophonic split' is used, which refers to a conditioned sound split. In the majority of cases, allophonic split is a precursor to phonological split. The latter happens when the conditions that caused the original split are broken and the two sounds can then occur in the same environment; they thus stand in phonemic opposition to one another.

A considerable number of phonological variations in Arabic dialects yield allophonic splits. A good example of such a split is the conditioned affrication of /k/ and /g/ in the environment of, mainly, front vowels. This feature is commonly known in Arabic philology as → *kaškaša/kaskasa*, and can be found in a wide range of Bedouin dialects in Arabia and the Levant (see Ingham 1982; Prochazka 1988; Behnstedt 1997). For example, in the Najdi dialect of al-Qāsim, /k/ splits into [k] and [ts], and /g/ splits into [g] and [dz]. The affricated sounds occur in the vicinity of front vowels, e.g. [tsiθiːr] 'much', [rifiːdz] 'companion' (Prochazka 1988:16–17). Elsewhere, the velar sounds are used. In the same dialect, however, /k/ and /ts/ behave like separate phonemes in the 2nd person clitic form /ik/. They occur in the same environment, and function to distinguish masculine gender [ik] from feminine gender [its]; thus we find [ʃaːfik] 'he saw you [masc.]', [xallik] 'stay! [masc.]', vs. [ʃaːfits] 'he saw you [fem.]', [xallits] 'stay! [fem.]' (examples taken from Al-Essa, forthcoming). Phonologically, these examples show a genuine split with a morphophonemic function. Notice, though, that the split is confined to this particular function and does not cause an unpredictable redistribution of the /k/ lexical set. Therefore, it may be described as subphonemic. Further analysis, using reconstruction, of this particular case is useful. The change from [k] to [ts] was a phonetically conditioned development, which yielded conditioned allophonic split in /k/ and a palatalized sound [ts]. As a process of assimilation, palatalization is usually governed by a hierarchy of conditioning environments, such that it is most likely to occur before /i/-type vowels and least likely before /a/-type vowels, i.e., front high vowels are the most likely vowels to induce palatalization. If palatalization is found before lower front vowels, it is normally also found before higher front vowels in the same language, i.e., there is an → implicational scale (on this, see Hock 1991:73–77). It follows then that /k/ was palatalized in the environment of /i/ before it was palatalized in the environment of /a/. Assuming that the 2nd person suffixes descend from masculine /ak/ and feminine /ik/, palatalization must have occurred before neutralization in the vowel quality. The feminine form would have undergone palatalization first, and thus we get /ak/ vs. /its/. It may be posited that the phonetic conditioning that spread palatalization to /k/ also before /a/ was blocked in the suffixes precisely to preserve gender information, yielding /k/ vs. /ts/. Especially in an unstressed syllable, the opposition between /k/ and /ts/ may have induced vowel neutralization (since vowel distinction became redundant), and thus we get masculine /ik/ vs. feminine /its/.

The type of split that is of most interest to linguists is where a preexisting lexical set is broken and redistributed between two phonemes along lines that cannot be entirely predictable, which is normally caused by an allophone developing an independent phonemic status. Phonological split of this type is typically very slow. A classic example of such a split from English is Middle English short /u/, which contained *put, full* as well as *cut, dull*, but split into /u/ and /ʌ/ in southern dialects (see Wells 1982). What looks like a clear example of phonological split is reported in Prochazka (1988). The sound [ʒ] is usually found as a variant of /j/ in many Arabic dialects, including some Arabian dialects. However, in three of the dialects described by Prochazka, Sarāt, ʿAbīda, and Najrān, the voiced apical [ʒ] occurs in items from the /z/ lexical set, e.g. /žyāra/ 'visit', /yžūr/ 'he visits', /žēn/ 'good' (Prochazka 1988:15). The same dialects have affricate [dʒ] for /j/. The description provided by Prochazka seems to suggest that a new phoneme, /ž/, which occurs in items from the /z/ lexical set, has been introduced, and thus a phonemic split in the /z/ lexical set has occurred.

A new phoneme can be introduced through borrowing from other languages or other varieties, and the assignment to the new phoneme of lexical items from preexisting sets. A good illustration of a distinction that may have been created in this way is the opposition between /q/ and /g/ in some varieties of Tunisian Arabic. Jabeur (1987:100–102) lists many minimal pairs which show that these sounds are treated as separate phonemes in the urban dialect of Tunis (which characteristically has [q] as the reflex of /q/), such as /gamra/ 'the moon', /qamra/ 'a traditional door handle', /dagg/ 'to prick', /daqq/ 'to hit', /dagdag/ 'to break', /daqdaq/ 'to knock on the door'. In the same dialect, [g] is used in a number of items which do not occur with [q], such as /gurṭ/ 'hay', /gʿūd/ 'a small camel'. Jabeur (1987:102) observes that a large proportion of this category of words relate to "different aspects of rural life". Following Cantineau (1960), he considers the origin of these items to be borrowings from rural dialects. By the same token, the rural Tunisian dialect covered in Jabeur's study, which has [g] as the normal reflex of /q/, contains both sounds as separate phonemes. Jabeur (1987:103) cites some examples from Sekik of this distinction: /qamqūma/ 'good', /gamgūma/ 'a mouthful of water', /ʿriqt/ 'I was ashamed of myself', /ʿrigt/ 'I sweated'. Cases of opposition between dialectal /ʾ/ and Classical Arabic /q/ are reported for a number of dialects; for instance, in Haeri's (1997) data /qawi/ 'strong' contrasts with /ʾawi/ 'very' in the dialect of Cairo.

A number of cases of vocalic splits, some of which are a result of the incorporation of learned loans from Classical Arabic, are cited in Versteegh (1997:153–154) from the Syro-Lebanese dialect group. In the dialect of Aleppo, which normally has → *ʾimāla* of /ā/ to /ē/ also after emphatic and guttural consonants, the following items occur: /ṭēleb/ 'striving' vs. /ṭāleb/ 'student', and /kēteb/ 'writing' vs. /kāteb/ 'writer. In the Lebanese dialect of Bišmizzīn, *ʾimāla* raises /ā/ to [æː], which can be blocked by a preceding emphatic sound which produces [ɒː], but in some contexts both variants can occur, e.g. [ʒɒː] 'to come' vs. [ʒæːb] 'to bring'. In Tripoli, the two allophones of /ā/ which result from *ʾimāla* can stand in phonemic contrast to the monophthongal variants of /ay/ and /aw/, /ē/ and /ō/, respectively, in closed syllables.

Bibliographical references

Al-Essa, Aziza. Forthcoming. "When Najd meets Hijaz: Dialect contact in Jeddah". *Proceedings of the AIDA conference, Vienna, 2006*, ed. Stefan Procházka.
Behnstedt, Peter. 1997. *Sprachatlas von Syrien.* Wiesbaden: O. Harrassowitz.
Cantineau, Jean. 1960. *Cours de phonétique arabe.* Paris.
Haeri, Niloofar. 1997. *The sociolinguistic market of Cairo: Gender, class and education.* London and New York: Kegan Paul International.
Hock, Hans H. 1991. *Principles of historical linguistics.* Berlin and New York: Mouton de Gruyter.
Ingham, Bruce. 1982. *North-east Arabian dialects.* London and Boston: Kegan Paul International.
Jabeur, Mohamed. 1987. *A sociolinguistic study in Tunisia: Rades.* Ph.D. diss., University of Reading.
Prochazka, Theodore. 1988. *Saudi Arabian dialects.* London and New York: Kegan Paul International.
Versteegh, Kees. 1997. *The Arabic language.* Edinburgh: Edinburgh University Press.
Wells, John. 1982. *Accents of English.* Cambridge: Cambridge University Press.

Enam Al-Wer (University of Essex)

Phonology

1. Definition

Every language has a system for mapping meaning to sound. The phonology of a language defines the set of sounds that signal meaning in a language (its phonemes), as well as the principles for combining these sounds into syllables, morphemes, words, and phrases. Two sounds are said to be separate phonemes of a language, or in contrast, when the substitution of one sound for the other can signal a change in meaning. The set of phonemes varies across languages; for example, the contrast between /p/ and /b/ is significant in English (as in *pat* vs. *bat*), but not in Arabic. In addition to defining the phoneme inventory of a language, the phonology defines the ways in which phonemes are pronounced in different contexts, the possible combinations of phonemes, and the characteristic stress and intonation patterns of the language.

Characterizing the phonology of Arabic is not a simple matter because the rubric 'Arabic' is used to include a large number of linguistic systems, including Classical Arabic, Modern Standard Arabic, and the wide range of spoken vernaculars associated with different regions and different social groups. While certain vernaculars are so different as to be largely mutually unintelligible, some features are shared by many or most varieties of Arabic. The following sections outline phonological features that are common to many different varieties and highlight some of the more systematic differences among the different vernaculars.

2. PHONEMES

One feature that distinguishes Arabic from many other languages is its relatively large inventory of consonant phonemes, particularly those produced at the posterior regions of the vocal tract. Eighth-century Classical Arabic employed 28 contrasting consonants (Watson 2002:13), and most modern varieties have to a large extent preserved a good proportion of these contrasts. In addition, in most varieties, consonants can be realized either as single or as geminate (also called doubled or long), as illustrated by the minimal pair *silim* 'he was saved' vs. *sillim* 'stairs' of colloquial Cairene Arabic, thus introducing another dimension of contrast. Table 1 shows the consonant inventory of Modern Standard Arabic.

Perhaps the most noteworthy feature of the consonant system of Standard Arabic is its extensive use of the back portions of the vocal tract. In addition to two laryngeal consonants, two pharyngeal consonants, and at least one uvular consonant (the Classical Arabic uvular fricatives corresponding in some modern varieties to velar fricatives), the so-called emphatic consonants are generally described as superimposition of a pharyngeal or uvular constriction on a primary coronal constriction (see McCarthy 1994 and Watson 2002 for review of relevant literature; → velarization; → *ʾiṭbaq*). Also of interest is the presence of interdental fricatives, which are relatively rare among the world's languages. A third feature is the lack of voiced/voiceless oppositions for certain obstruents; we find voiced labial stop /b/ but not its voiceless counterpart /p/; voiceless labiodental fricative /f/ but not its voiced counterpart /v/; and voiceless velar stop /k/ but not its voiced counterpart /g/ (although the affricate /j/ likely reflects older /g/, a pronunciation that persists in the pronunciation of speakers of Egyptian Arabic). We should note that the missing counterparts may occur as a result of allophonic (noncontrastive) variation in par-

Table 1. Modern Standard Arabic consonant phonemes (cf. Beeston 1970; Fischer and Jastrow 1980; phonetic symbols have been added between square brackets)

	labial	labio-dental	interdental	dental-alveolar	palatal	velar	uvular	pharyngeal	laryngeal
stop (plosive)									
voiceless				t		k	q		ʾ [ʔ]
voiced	b			d					
affricate, voiced					j [dʒ]				
fricative									
voiceless	f	ṯ [θ]		s	š [ʃ]	x [x]	or [χ]	ḥ [ħ]	h
voiced		ḏ [ð]		z		ġ [ɣ]	or [ʁ]	ʿ [ʕ]	
emphatic									
voiceless stop				ṭ [ᵵ]					
voiced stop/ fricative				ḍ [ᵭ]/ ḏ̣ [ð̴]					
voiceless fricative				ṣ [ᵴ]					
voiced fricative				ẓ [ᵶ]					
nasal	m			n					
lateral				l					
tap				r					
glide	w				y				

ticular phonetic contexts. For example, the final /b/ of the stem /katab/ 'to write [perf.]' may be pronounced as voiceless when it precedes a voiceless consonants, as in Cairene Arabic [katapt] 'I wrote' from /katab+t/, but the voiceless pronunciation is limited to these contexts in native vocabulary, and the contrast between /p/ and /b/ tends to be neutralized in assimilated loanwords (as in Egyptian /bilastik/ 'plastic').

In contrast to its profusion of consonant phonemes, Modern Standard Arabic exhibits a relatively restricted range of vowel contrasts. The Classical Arabic vowel system counterposes three short vowels, high front /i/, high back /u/, and low /a/, along with the corresponding long vowels /ii/, /uu/, and /aa/. Some modern Western Arabic vernaculars have reduced the number of contrasting short vowels to two or even one (Fischer and Jastrow 1980:33), and the contrasts among long vowels have been increased to five in some varieties (such as Cairene Arabic) with addition of the long mid vowels /ee/ and /oo/ (generally corresponding to Classical Arabic diphthongs /ay/ and /aw/).

While modern vernaculars have in general preserved the majority of the early Arabic consonantal contrasts, many differ in the pronunciation of individual consonants. For example, to the pronunciation of /j/ as [g] in Egypt can be added the frequent substitution of the palatal fricative [ʒ] (compare Iraqi [dʒidiːd], Egyptian [gidiːd], and Syrian [ʒidiːd] 'new'). Interdental fricatives are also subject to much variation, most often merging with dental/alveolar stops or fricatives, although Holes (1990:262) reports substitution of [f] for /ṯ/ by "socially defined groups" of Gulf Arabic speakers. Retention of the interdental articulation is typically associated with nomadic dialects, but because it is also associated with prestige varieties of a single dialect, variation may be found within a single region. Mitchell (1993) notes the possibility in Egyptian Arabic of three pronunciations of the word 'three' of decreasing formality: [θalaːθa, salaːsa, talaːta]. Original voiceless velar stop /k/ is realized as fronted and affricated [tʃ] in some dialects, either before front vowels or in all contexts (Watson 2002:16); compare Egyptian [kibiːr] and Kuwaiti [tʃ(i)biːr] 'big'. A number of reflexes of the Classical Arabic voiceless uvular stop /q/ are also found; for example, in Palestinian Arabic, Mitchell (1993:34) reports four possible pronunciations of 'he said': [qɑːl], [ʔaːl], [gɑːl], [kaːl] (with variation in vowel quality conditioned by the preceding consonant, as discussed in Sec. 4). Again, the choice among variants is associated with particular regional, stylistic, and social differences; for example, Cantineau (1939) lists the pronunciation of original /q/ as a principal diagnostic for sedentary vs. nomadic dialects. And the Classical Arabic distinction among four emphatics /ṣ, ṭ, ḍ, ḏ̣/ has in many varieties been reduced to a three-way distinction, either /ṣ, ṭ, ḍ/ or /ṣ, ṭ, ḏ̣/ (Mitchell 1993:31).

3. PHONEME CLASSES AND PHONEME COMBINATIONS

Examination of the morphophonological processes of Arabic reveals classes of phonemes sharing phonetic features that pattern together with respect to various processes. For example, in Standard Arabic, the /l/ of the definite → article assimilates to a following 'sun letter' but not to a 'moon letter' (/aš-šams/ 'the sun', /an-nahr/ 'the river', but /al-qamr/ 'the moon'). The sun letters are the set of consonants formed with the tongue blade, generally called coronal consonants, which includes interdentals, dental-alveolars, and alveopalatals. The one exception to this natural grouping is the alveopalatal affricate /j/, which (though coronal) fails to trigger assimilation of /l/ in Standard Arabic (→ *jīm*). A number of vernaculars also fail to assimilate the /l/ of the article either to /j/ or to its alternate form, the alveopalatal fricative /ž/ (Watson 2002:218). The reasons for this are no doubt historical, as /j/ and /ž/ are the reflexes of earlier noncoronal /g/. In Cairene Arabic, which has /g/ where other varieties have /j/, the /l/ of the article often assimilates to velar /g/ and to its voiceless counterpart /k/, as well as to coronal consonants.

The class of guttural consonants (uvulars, pharyngeals, and laryngeals; Sībawayhi's 'throat consonants') pattern similarly in many respects. For example, in various Bedouin vernaculars, a vowel is inserted to separate a guttural from a following consonant; compare Cairene /ʔahwa/ 'coffee' and Negev Bedouin /gahawah/ 'coffee' (Blanc 1970; McCarthy 1994; → gahawa-syndrome). This vowel insertion occurs only when the guttural is preceded by /a/, and in some varieties, the original /a/ is lost, with only the inserted vowel remaining (as in Bedouin Ḥijāzī

Arabic /ghawa/ 'coffee'; Irshied and Kenstowicz 1984; McCarthy 1994).

Another respect in which consonants with similar places of articulation pattern as a group is in terms of defining possible → roots. Like other Semitic languages, the Arabic lexicon is built on a system of consonantal roots and vocalic patterns, or templates, which define the shape of the word in which the roots appear. In Classical Arabic, certain of the logically possible combinations of consonants in a root are much less frequent than would be expected if consonants combined at random (Greenberg 1950; McCarthy 1994; Frisch, Pierrehumbert, and Broe 2004; → phonotactics). Generally, roots are unlikely to contain adjacent labial consonants (/b, f, m/). Adjacent coronals are avoided if they also share similar manners of articulation; thus, roots with adjacent coronal sonorants, coronal stops, or coronal fricatives are rare, and even combinations of a coronal stop and a coronal fricative are unlikely. In the posterior regions, combinations of velar and uvular consonants are avoided, as are combinations of guttural consonants (the uvular fricatives /x, ġ/, the pharyngeals /ḥ, ʿ/, and the laryngeals /h, ʾ/).

Roots with adjacent identical consonants, which are relatively frequent, pose a puzzle: why should a root like *s-m-m* (underlying, for example, /samamtu/ 'I poisoned') be allowed, while hypothetical **s-m-b* is not? McCarthy (1981) proposes that stems like /samam-/, with adjacent identical consonants, are actually derived from biliteral roots (*s-m*), with the second root consonant filling the second and third consonantal positions of the verb stem template (→ biradicals). Under this analysis, the constraint against adjacent homorganic roots consonants holds at the lexical level rather than the phonetic level. Biliteral roots persist in modern vernaculars, although the modern reflex of stems like /samam/ is a monosyllabic stem with a final geminate (cf. Cairene Arabic /samm-/) in all contexts (reflecting the shape taken by these stems before a vowel-initial suffix, as in Classical Arabic /samma/ 'he poisoned').

4. Emphasis

Emphatic consonants are those with a primary articulation in the oral cavity and a secondary articulation in the back of the vocal tract. Emphatic articulation was traditionally described as → velarization, although experimental studies have revealed that at least in most cases, emphatic sounds involve constriction in the upper pharynx (see, for example, Al-Ani 1970; Ghazeli 1977; McCarthy 1994), which may be accompanied by lip protrusion or lip rounding (Watson 2002:268). The emphatic consonants have been characterized by the feature [guttural] (McCarthy 1994), which predicts that they can pattern with primary guttural consonants. Thus, the raising (or → *ʾimāla*) of the feminine suffix from historically prior /a/ to /i/ or /e/ in many Eastern dialects is blocked when the suffix is preceded by a guttural or emphatic consonant (McCarthy 1994:219).

Vowels adjacent to uvular, pharyngeal, and emphatic consonants tend to be pronounced somewhat lower than in other contexts, and may be backed as well. This distinction is most salient with low vowels. For example, the Cairene pair /baat/ 'he spent the night' and /baaṭ/ 'armpit' differ at the phonemic level only in that the first has plain /t/ and the second emphatic /ṭ/, but the vowels of these words are quite distinct, particularly to the English ear, with the first vowel similar to the vowel of English *bet* and the second closer to English *bought*. With emphatic consonants, this coarticulatory effect may spread throughout the syllable or word, or sometimes into neighboring words, a phenomenon known as 'emphasis spread'. Dialects may differ in the domain of emphasis spread, the direction of emphasis spread, the set of consonants that trigger emphasis spread, and the set of segments that block emphasis spread (Watson 2002:273–275).

5. Syllables

In addition to restrictions on combinations of specific phonemes and phoneme classes, most varieties of Arabic have well-defined restrictions on possible combinations of consonants and vowels. Such restrictions are often analyzed in terms of possible syllables. Classical and Modern Standard Arabic allow syllables of the type CV, CVV, and CVC (where C = consonant, V= vowel, and VV = long vowel). In addition, CVVC and CVCC syllables are found (as well as CVVCC, where the final CC represents a geminate consonant), but this syllable type is restricted to word-final or utterance-final position (Al-Ani and May 1973).

Most analyses of the Arabic restrictions on syllable type posit two major intrasyllabic constituents, the onset (consonantal material preceding the vowel) and the rhyme (the vowel and any following consonants). Onsetless syllables are typically forbidden, leading to insertion of a glottal stop before a phrase-initial vowel, as, for example, before the definite article *al* in Ṣanʿānī Arabic and *il* in Cairene Arabic, and before vowel-initial loanwords such as /ʔiksibres/ 'express' (Watson 2002:66). Complex (multiconsonantal) onsets are also forbidden in Classical Arabic and in some (but not all) modern varieties (compare Cairene /kitaab/ and Syrian /ktaab/ 'book'); in Iraqi, complex onsets are possible only in phrase-initial position.

The syllable rhyme is generally assumed to consist of two components, the nucleus (vocalic material) and the coda (consonants following the nucleus). Many varieties limit the coda to a single consonant (with a special provision, in some cases, for phrase-final position). The prohibition on complex (multiconsonantal) codas, combined with the prohibition on complex onsets, restricts the number of possible consonants in a sequence to two. Many vernaculars insert anaptyctic or 'helping' vowels to break up clusters of three consonants (→ epenthesis). Both the position and the quality of the anaptyctic vowel may vary from dialect to dialect. For example, in Cairene Arabic, concatenation of the stem /katab/ 'to write' with two consonantal suffixes (/katab+t+l+u/) results in insertion of /i/ after the second of the three consonants: [katabtilu] 'I wrote to him'. In contrast, Iraqi speakers insert /i/ after the first of three consonants: [kitabitla] 'I wrote to him' (/kitab+t+l+a/). However, when morpheme concatenation creates a four-consonant sequence, the position of the inserted vowel is the same in both dialects, between the second and third consonants: cf. Cairene [katabtilha] (/katab+t+l+ha/) and Iraqi [kitabtilha] (/kitab+t+l+ha/) 'I wrote to her'. In all these cases, insertion of a vowel allows for the division of words into syllables that are maximally CVC. The Cairene/Iraqi contrast has been analyzed as a difference in a preference for 'stranded' consonants to serve as onsets vs. rimes (Selkirk 1981; Broselow 1992) or as a difference in the direction in which syllabification proceeds (Itô 1989; Farwaneh 1995; → resyllabication). Syrian Arabic is similar to Iraqi in inserting a vowel after the first of three consonants, but the inserted vowel is schwa rather than /i/ (Cowell 1964).

Another respect in which vernaculars may differ is in their tolerance of consonant sequences in utterance-final position. Some varieties forbid complex codas anywhere, while others will tolerate two consonants in utterance-final position; compare, for example, Cairene Arabic /bint/ 'girl' with Cyrenaican Bedouin /binit/ 'girl' (Mitchell 1993:69; Mitchell transcribes the inserted vowel as [ɨ]). However, even in Cairene, a complex coda is not tolerated when a consonant follows in the next word ([binti gamiːla] 'a pretty girl'). In many vernaculars, the possibility of final clusters depends on the identity of the final consonants, and insertion of a helping vowel may be optional (for example, Mitchell [1993: 70] reports both [banj] and [banɨj] 'inoculation' for Jordanian speakers). The most dramatic departures from the original restrictions on syllable shape are found in the Western dialects. Moroccan Arabic is notable for its sequences of consonants, and the analysis of Moroccan syllable structure is by no means straightforward, with researchers differing on whether consonants themselves may serve as the centers of 'vowelless' syllables, or whether transitions between consonants represent actual vowels (see, for example, Harrell 1962; Gafos 2002).

One somewhat unusual feature of many varieties of Arabic is the extent to which the grouping of segments into syllables may take place without regard for word boundaries. The phenomenon of vowel elision in Cairene Arabic provides a striking example of the close connection between words within a phrase. In Cairene, the vowel /i/ is normally deleted when unstressed in the context VC_CV, as illustrated in the verbal forms /širib/ 'he drank', /širibna/ 'we drank' (/širib+na/), /šir_bu/ 'they drank' (/širib+u/). This deletion is accompanied by reassignment of the consonant preceding the elided vowel (/r/ in /ši.rib/) to a preceding coda (/šir.bu/). The restriction of elision to VC_CV contexts, where the elided vowel is flanked by no more than a single consonant, ensures that the consonant that would have served as onset to the elided vowel can always be reassigned to another syllable without creating complex onsets or codas. Interestingly, elision may take place even when part of the VC_CV context is contained in a neighboring word, as in /ʔa.na #š._ribt/ 'I drank' (with overt subject pronoun), or in /šir._b# il.may.ya/ 'he drank the water' (where # indicates word juncture and a period indicates syllable juncture). Furthermore, the context for elision may include inserted vowels: the concatenation of /bint/ 'girl' and /kibiira/ 'big [fem.]' results in [binti k_biːra], with an anaptyctic vowel breaking up the cluster /nt#k/, which in turn provides the context for elision of the first vowel in /kibiira/. In the case of elision triggered by cross-word contexts, the result is a mismatch between word structure and syllable structure. In fact, the syllabic structure of [binti#kbiːra] 'a big girl' is quite similar to that of [bintik#biːha] 'your daughter is on it [fem.]', even though the word membership of the /k/ is different in the two phrases. The lack of isomorphism between word structure and syllable structure is a frequent source of misperceptions by English-speaking learners of Egyptian Arabic (Broselow 1984).

Not all dialects limit vowel elision to contexts in which consonants can be reassigned to neighboring syllables. In Syrian Arabic, for example, a vowel may be elided in an open syllable, even following two consonants, and such elision may necessitate insertion of a new vowel to make the output pronounceable according to syllable structure restrictions. Thus, when the singulative suffix is added to /məšmoš/ 'apricots', the /o/ is lost, but schwa is inserted in a different position to break up the resulting sequence of three consonants: /məšəmše/ 'an apricot' (Cowell 1964).

In addition to limiting the distribution of consonants, some varieties also limit the contexts in which long vowels can occur. In both Cairene and Meccan Arabic, a long vowel may be followed by a coda consonant in word-final position but not within a word. Thus, for example, when the possessive suffix /na/ 'our' is added to /xaal/ 'maternal uncle', the expected form /xaal+na/ would contain a word-internal CVVC syllable (since /ln/ is not a possible syllable onset). In Cairene, this problem is solved by shortening the long vowel in a closed syllable, yielding /xal.na/ 'our maternal uncle' (where the boundary between syllables is indicated by a period). In Meccan Arabic, however, a different strategy is used, inserting a new vowel with which the /l/ can syllabify, giving /xaa.la.na/ 'our maternal uncle'. Both alternatives allow speakers to avoid word-internal CVVC syllables.

The prohibition on CVVC syllables has generally been described in terms of syllable weight. The weight of a syllable is determined by the makeup of its rhyme, with each short vowel and each coda consonant associated with one weight unit, or mora, and long vowels with two moras. Therefore, CV syllables are monomoraic (light), while CVV and CVC syllables are bimoraic (heavy). A syllable containing both a long vowel and a coda consonant is 'superheavy', associated with three moras. The dispreference for CVVC syllables has been analyzed as the reflection of a bimoraic limit on the weight of Arabic syllables (Broselow 1992; Watson 2002:56). The possibility of CVVC syllables in word-final position has been described by positing that the right edge of a word confers special status: a single consonant at the right edge of a domain is not associated with the core syllable (see, for example, McCarthy 1979; Selkirk 1981; Angoujard 1990). Under this account, even word-final CVVC syllables obey the bimoraic limit, since the word-final consonant is not part of the same syllable as the long vowel.

Word-internal CVVC structures are tolerated in a number of varieties; for example, the Syrian pronunciation of 'our maternal uncle' is /xaal.na/. However, even this pronunciation has been argued to conform to the bimoraic maximum; Broselow a.o. (1997) measured the relative durations of vowels in CVVC, CVV, CVC, and CV syllables as produced by three speakers of Levantine dialects (one Jordanian, one Syrian [Damascus], and one Lebanese), and found that the vowel in a CVVC syllable was shorter than the corresponding vowel in a CVV syllable, though still longer than a short vowel. They therefore argue that the vowel and following consonant share a mora, thus avoiding a trimoraic structure.

6. Stress

A stressed (or accented) syllable is relatively more prominent than syllables surrounding it. Scholars disagree on the facts of word → stress in

Classical Arabic, although it is generally agreed that in the Classical Arabic system the position of a stressed syllable within a word was largely predictable with reference to the weight and position of syllables within the word. A word-final syllable was stressed if it was superheavy (CVVC or CVCC); stress otherwise fell on the rightmost nonfinal heavy syllable (CVV or CVC). In the absence of any nonfinal heavy syllables, it has been argued either that stress fell on the initial syllable of the word (Wright 1981) or that stress was limited to one of the last three syllables of the word (Angoujard 1990; Versteegh 1997). Stress in Modern Standard Arabic tends to depend on the native vernacular of the speaker; Mitchell (1993:194) reports four different stress patterns for the same word 'they both [fem.] wrote': /kátabataa/ (Upper Egypt), /katábataa/ (Jordan), /katabátaa/ (Cairene), and /katabatáa/ (Lebanon).

While the modern vernaculars may differ in the particulars of their stress systems, the Eastern varieties share many of the Classical Arabic features, most noticeably the intimate connection between stress and syllable weight. Ignoring for the moment speakers' pronunciations of Modern Standard forms, we find that in most modern vernaculars, as in Classical Arabic, a final superheavy syllable is stressed, as illustrated by Cairene /katábt/ 'I wrote' (vs. /kátab/ 'he wrote') or /waladéen/ 'two boys' vs. /wálad/ 'boy'. In the absence of a final CVV(C) or CVCC, stress in the vernaculars generally falls on the penultimate syllable if it is heavy (CVV or CVC), otherwise on the antepenultimate syllable: compare Cairene /katábna/ 'we wrote' and /katabnáaha/ 'we wrote it [fem.]', both of which have heavy (and stressed) penultimate syllables, with /kátabu/ 'they wrote', /inkásarit/ 'it [fem.] was broken', both of which have light (and unstressed) penultimate syllables. In Cairene, the connection between stress and weight is made even more salient by the characteristic shortening of long vowels that fall in an unstressed position: compare /ʾitnáaʾiš/ 'he discussed' with /ʾitnaʾíšna/ 'we discussed' (/ʾitnaaʾiš+na/), where the closed penultimate syllable attracts stress, resulting in shortening of the lexical long vowel.

The weight of the final syllable is determined at the lexical rather than the phonetic level. A final syllable closed in a geminate (long) consonant counts as superheavy, even though speakers of many different vernaculars shorten geminates in this context. For example, Harrell (1957:16) points out the Egyptian contrast between initially stressed /síkit/ 'he became silent' and finally stressed /sikít/ 'I became silent', where the stress contrast is clearly a reflection of the lexical differences between the two (/sikit/ vs. /sikit+t/). Final syllables ending in long vowels are also stressed in colloquial Cairene (/ḥayáa/ 'life', /gatóo/ 'cake'), patterning with final superheavy syllables. Final long vowels are rare in Cairene (due to the shortening of originally long final vowels in this vernacular), and given the lack of contrast between final VV and VVh, such syllables might be analyzed as ending in /h/ (cf. McCarthy 1979). If this analysis is accepted, then superheaviness is both a sufficient and a necessary condition for attracting stress in final position. An alternative is to consider word-final long consonants as extrasyllabic (Hayes 1995:68) and therefore as making no contribution to syllable weight. Under this assumption, word-final CVV, CVVC, and CVCC are all bimoraic, while word-final CVC and CV are both monomoraic. The assignment of stress to final syllables if and only if superheavy is nearly universal in the vernaculars. However, in Ṣanʿānī Arabic, spoken in Yemen, a superheavy final syllable may be left unstressed, stress falling instead on a syllable containing a long vowel or closed by a geminate consonant (Ṣanʿānī Arabic /maktúub/ 'letter' but /makáatiib/ 'letters'; Watson 2002:81). And in Negev Bedouin Arabic, a final syllable of any weight is stressed when preceded by a single light syllable: /biná/ 'he built', /jimál/ 'camel' (Blanc 1970; Kenstowicz 1983). The position of stress is frequently analyzed as a reflection of the grouping of syllables into rhythmic units called feet. The contrast between the Negev Bedouin final stress in forms like /biná/ 'he built' vs. corresponding Iraqi /bína/, Cairene /bána/ 'he built' can be analyzed as a difference in preferred foot type. In this view, feet may be either trochaic or left-dominant (with stress falling on the left-hand syllable of a bisyllabic foot), or iambic or right-dominant (with stress falling on the right-hand syllable of a bisyllabic foot). The Iraqi/Cairene pattern can be seen as reflecting a preference for trochaic feet and the Negev Bedouin pattern a preference for iambic feet (Kenstowicz 1983:21–211; Hayes 1995:226).

Another noticeable difference among vernaculars concerns the possibility of stress on a light penultimate syllable. In Cairene, stress falls on a light penult when it follows a heavy syllable, as in /madrása/ 'school', while in the Levantine dialects, and even in Upper Egypt, the antepenult is stressed in these cases (Syrian /mádrase/, Cowell 1984; Upper Egyptian /médrase/, Nishio 1994). One analysis of this difference (McCarthy 1979) asssumes that feet are constructed from the left edge of the word in Cairene and from the right edge in other dialects. A complete discussion of the formal analyses of the stress patterns of the Arabic vernaculars is beyond the scope of this article, as this topic has attracted a great deal of attention (see, for example, McCarthy 1979; Kenstowicz 1983; Angoujard 1990; Hayes 1995; Watson 2002; → stress). In many cases, these analyses of stress in the colloquial forms predict the stress patterns that emerge when speakers of a particular vernacular pronounce Modern Standard Arabic forms. Standard Arabic generally provides a wider range of possible syllable combinations than are found in the vocabularies of the vernaculars, in which the number of light syllables occurring in sequence is often severely limited by processes such as elision of vowels in open syllables.

The normal stress patterns of the vernaculars may be disrupted by the presence of particular grammatical morphemes. For example, in Iraqi Arabic, words with the dual suffix /-een/ are stressed not on the final superheavy syllable as expected, but on the initial syllable: /čálbeen/ 'two dogs' vs. /taˁbáan/ 'tired' (Erwin 1963:43). And in Cairene Arabic, where stress normally falls on the first of three light syllables, stress always falls on the 3rd person feminine singular suffix /-it/ when this suffix is followed by an object pronoun clitic, as in /banítu/ 'she built it [masc.]' vs. /kátabu/ 'he wrote it [masc.]'. This suffix is also exceptional with respect to the normal pattern of deletion of /i/ in the context VC_CV.

Another factor that may disrupt typical stress patterns is the presence of 'helping' vowels inserted to allow division of the word into allowable syllables. In Iraqi Arabic, syllable codas may contain no more than one consonant. A vowel inserted into a final consonant cluster is invisible for the purposes of stress, so that /kitab+t/ is stressed on the light penultimate syllable (/kitábit/ 'I wrote') rather than on the antepenultimate syllable as expected (Erwin 1963:87). However, when a vowel is inserted to break up a sequence of three consonants, fluctuation is found between the normal stress pattern [kitabítla] 'I wrote for her' (/kitab+t+l+a/), and the pattern in which stress is assigned as though the inserted vowel were not present: [kitábitla]. And where the inserted vowel breaks up a four-consonant sequence, it is always visible for the purposes of stress: [kitabtilna] 'I wrote for us' (/kitab+t+l+na/). This pattern is fairly common, being characteristic of many Levantine dialects. In Cairene Arabic, however, the inserted vowel behaves like underlying vowels with respect to stress, as in [katabtílu] 'I wrote for him' (/katab+t+l+u/).

Another area intimately connected to issues of stress and syllable weight concerns the restrictions on minimal word size. With only a few exceptions, content words (nouns, verbs, adjectives, and adverbs) in most vernaculars are either greater than one syllable or, if monosyllabic, are superheavy. This can be understood in terms of stress restrictions; most dialects do not stress a final syllable unless it is superheavy, so that a content word of the shape CV or CVC would not be sufficient to support a stress foot (Broselow 1984; McCarthy and Prince 1990) . Conversely, Negev Bedouin, which allows stress on a final nonsuperheavy syllable, also allows content words of CVC, such as /miy/ 'water' (Hayes 1995:228). These various aspects of the phonological system – the restrictions on syllable weight, stress, and word shape – are intricately connected in the varieties of Arabic.

Bibliographical references

Al-Ani, Salman H. 1970. *Arabic phonology*. Mouton: The Hague.

——, and Darlene R. May. 1973. "The phonological structure of the syllable in Arabic". *American Journal of Arabic Studies* 1.37–49.

Angoujard, Jean-Pierre. 1990. *Metrical structure of Arabic*. Dordrecht: Foris.

Beeston, A.F.L. 1970. *The Arabic language today*. London: Hutchinson University Library.

Blanc, Haim. 1964. *Communal dialects in Baghdad*. Cambridge, Mass: Harvard University.

——. 1970. "The Arabic dialect of the Negev Bedouins". *Proceedings of the Israeli Academy of Sciences and Humanities* 4.112–150.

Broselow, Ellen. 1984. “An investigation of transfer in second language phonology”. *International Review of Applied Linguistics* 22.253–269.

——. 1992. “Parametric variation in Arabic dialect phonology”. *Perspectives on Arabic linguistics*, IV, ed. Ellen Broselow, Mushira Eid, and John McCarthy, 7–45. Amsterdam and Philadelphia: J. Benjamins.

——, Su-I Chen, and Marie Huffman. 1997. “Syllable weight: Convergence of phonology and phonetics”. *Phonology* 14.47–82.

Cantineau, Jean. 1939. “Remarques sur les parlers de sédentaires syro-libano-palestiniens”. *Bulletin de la Societé de Linguistique de Paris* 40.80–88.

——. 1960. *Études de linguistique arabe*. Paris: Klincksieck.

Cowell, Mark W. 1964. *A reference grammar of Syrian Arabic*. Washington, D.C.: Georgetown University Press.

Erwin, Wallace. 1963. *A short reference grammar of Iraqi Arabic*. Washington, D.C.: Georgetown University Press.

Farwaneh, Samira. 1995. *Directionality effects in Arabic dialect syllable structure*. Ph.D. diss., University of Utah.

Fischer, Wolfdietrich and Otto Jastrow. 1980. *Handbuch der arabischen Dialekte*. Wiesbaden: O. Harrassowitz.

Frisch, Stefan, Janet Pierrehumbert, and Michael Broe. 2004. “Similarity avoidance and the OCP”. *Natural Language and Linguistic Theory* 22.179–228.

Gafos, Adamantios. 2002. “A grammar of gestural coordination”. *Natural Language and Linguistic Theory* 20.269–337.

Ghazeli, Salem. 1977. *Back consonants and backing coarticulation in Arabic*. Ph.D. diss., University of Texas at Austin.

Greenberg, Joseph H. 1950. “The patterning of root consonants in Semitic”. *Word* 6.162–181.

Harrell, Richard S. 1957. *The phonology of Egyptian colloquial Arabic*. New York: American Council of Learned Societies.

——. 1962. “Consonant, vowel and syllable in Moroccan Arabic”. *Proceedings of the 4th International Congress of Phonetic Sciences*, 643–647. The Hague: Mouton.

Hayes, Bruce. 1995. *Metrical stress theory: Principles and case studies*. Chicago: University of Chicago Press.

Holes, Clive. 1990. *Gulf Arabic*. London and New York: Routledge.

Irshied, Omar and Michael Kenstowicz. 1984. “Some phonological rules of Bani-Hassan Arabic: A Bedouin dialect”. *Studies in the Linguistic Sciences* 14:1.109–147.

Itô, Junko. 1989. “A prosodic theory of epenthesis”. *Natural Language and Linguistic Theory* 7.217–260.

Kenstowicz, Michael. 1983. “Parametric variation and accent in the Arabic dialects”. *Chicago Linguistics Society* 19.205–213.

McCarthy, John J. 1979. “On stress and syllabification”. *Linguistic Inquiry* 10.443–465.

——. 1981. “A prosodic theory of nonconcatenative morphology”. *Linguistic Inquiry* 12.373–418.

——. 1994. “The phonetics and phonology of Semitic pharyngeals”. *Phonological structure and phonetic form: Papers in laboratory phonology*, III, ed. Patricia A. Keating, 191–233. Cambridge: Cambridge University Press.

——, and Alan Prince. 1990. “Prosodic morphology and templatic morphology”. *Perspectives on Arabic linguistics*, II, ed. Mushira Eid and John J. McCarthy, 1–54. Amsterdam and Philadelphia: J. Benjamins.

Mitchell, T.F. 1993. *Pronouncing Arabic*, II. Oxford: Clarendon Press.

Nishio, Tetsuo. 1994. *The Arabic dialect of Qift (Upper Egypt): Grammar and classified vocabulary*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.

Selkirk, Elisabeth. 1981. “Epenthesis and degenerate syllables in Cairene Arabic”. *Theoretical issues in the grammar of Semitic languages* (= *MIT Working Papers in Linguistics*, 3), ed. Hagit Borer and Youssef Aoun, 209–232. Cambridge, Mass.: Department of Linguistics, Massachusetts Institute of Technology.

Versteegh, Kees. 1997. *The Arabic language*. Edinburgh: Edinburgh University Press.

Watson, Janet C.E. 2002. *The phonology and morphology of Arabic*. Oxford: Oxford University Press.

Wright, William. 1981. *A grammar of the Arabic language*. 3rd ed. Repr. Cambridge: Cambridge University Press.

Ellen Broselow (Stony Brook University)

Phonology: Metrical

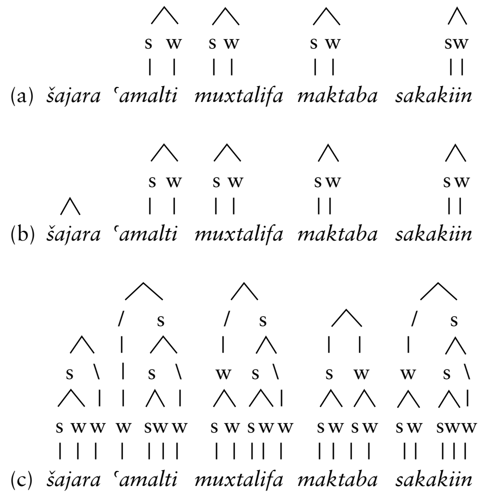
1. General framework

Metrical theory was founded by Liberman (1975) and elaborated on by Liberman and Prince (1977), Halle and Vergnaud (1978), Hayes (1980), and others as part of nonlinear phonology, in order to capture the hierarchical and rhythmical nature of stress in a representation of its own, in addition to the segmental matrix which contains other features. Although the theory has later been applied to other phonological phenomena, such as → vowel harmony, → syllable structure, deletion, and → epenthesis, word stress remained the central domain of metrical phonology (→ stress). The basic assumption of metrical phonology is that word-stress patterns universally rely on the underlying organization of words into a hierarchical structure of metrical constituents, and that word stress is the linguistic manifestation of this metrical structure, which receives two types of geometrical representations called ‘trees’ and ‘grids’. In the metrical tree, stress is represented as the result of a hierarchical binary branching structure, where all nodes are labeled strong-weak (SW) or weak-strong (WS). As for the grid, it is a representation of rhythmic structure, where the height of the grid columns represents the degree of prominence of stress-bearing units. Although early metrical theory derived the grid from the tree by a mapping rule, it soon became the main type of representation of rhythmical structure (Prince 1983; Halle and Vergnaud 1987a,b; Angoujard 1990; Hayes 1995). The shape of metrical feet and trees is governed by a number of basic parameters: boundedness (feet are either bounded, i.e. binary; or unbounded, i.e. enclosing all units at a certain level), foot dominance (feet are right-dominant, WS; or left-dominant, SW), quantity sensitivity (branching rhymes cannot be in a weak position), directionality (feet are constructed from right to left or from left to right), iterativity (only one foot or more than one foot is constructed), word-tree dominance (left or right), and extrametricality (an initial or final metrical unit is said to be invisible to metrical rules). Core grammars thus consist in a set of rule specifications, defined by values of the above universal parameters. A number of alternative models soon developed out of this basic theoretical mold, such as Prince’s (1983) Grid Theory, Halle and Vergnaud’s (1987a,b) Bracketed Grid Theory, and various Asymmetric Rhythmic Theories, such as Angoujard (1990) and Hayes (1995). Scholars of Arabic have always paid a great deal of attention to the various dialects, and study of some dialects, such as Cairene and Ḥijāzī Arabic, has played a major role in the elaboration and discussion of the theoretical models, as is shown briefly below.

2. Classical Metrical Theory

Most Arabic dialects share with Classical and Modern Arabic a number of basic properties concerning syllable structure. Arabic usually has three types of syllables, light (CV), heavy (open – CVV, where VV is a long vowel, or closed – CVC) and superheavy (CVVC or VVCC), which is only permitted, with a few exceptions, word-finally. Word-stress assignment relies on quantitative principles (in what follows, *aa* = *ā*, *ii* = *ī*, and *uu* = *ū*). Let us consider the case of Cairene Arabic, which was first analyzed within the framework of Classical Metrical Theory by McCarthy (1979). In this dialect, main stress is (i) on final superheavy syllables (CVVC, CVCC), e.g. *sakakíin* ‘knives’, *kitáab* ‘book’, or (ii) on heavy penults (CVV, CVC), e.g. *ʿamálti* ‘you [fem.] did’, *madáares* ‘schools’, or (iii) on the rightmost nonfinal odd-numbered light syllable counting from the nearest preceding heavy syllable or the initial syllable, e.g. *maktába* ‘library’, *šájara* ‘tree’, *muxtálifa* ‘different [fem.]’, *šajaratúhu* ‘his tree’, *šajaratuhúmaa* ‘their [du.] tree’, the last three examples being Classical Arabic words as pronounced by native Cairene Arabic speakers. McCarthy first analyzed superheavy syllables into a heavy syllable plus a degenerate syllable, which is the final consonant (CVV+C or CVC+C), so that cases of final stress can be considered special instances of penultimate stress. Following Liberman (1975), he assumed that “all stress information is encoded directly in the metrical trees” (McCarthy 1979:449). First of all, branching rhymes are projected onto the metrical tier (1a). Binary feet are then assigned from left to right to pairs of light syllables (1b). A right-branching superstructure gathers up all feet and stray syllables in the word and the entire tree is finally labeled according to the principle that a right node is strong (S) if and only if it branches (1c). The stressable unit which is only dominated by strong nodes (S) bears word stress (McCarthy 1979:449–450, 457):

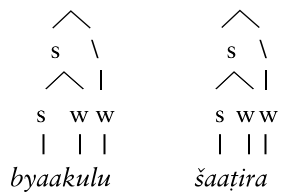
(1)


s w s w s w sw
(a) šajara ʿamalti muxtalifa maktaba sakakiin
s w s w sw sw
(b) šajara ʿamalti muxtalifa maktaba sakakiin
s s s
s s w s s w w s
s ww w sww s w sww sw sw sw sww
(c) šajara ʿamalti muxtalifa maktaba sakakiin

The Damascus dialect (McCarthy 1979:460) is subject to the same syllabification and labeling rules as Cairene Arabic. As for word-stress patterns, the ultima is stressed if it is superheavy (*darrást* 'I taught'), otherwise, the penult if heavy (*fátḥet* 'she opened', *madáares* 'schools'), otherwise, the antepenult (*dárasu* 'they studied', *mádrase* 'school', *muttáḥide* 'united [fem.]'). McCarthy proposes the following foot-assignment rule, applying from right to left: [[n1 n2] n3], where n1 and n2 do not branch, which means that a left-branching tree is constructed while preserving and using preexisting metrical structure (branching rhymes and pairs of light syllables).

Bohas and Kouloughli (1981) pointed out a number of words where McCarthy's rules failed to assign the right stress pattern, such as *byáaklu* 'they eat' and *šáṭra* 'talented' in Cairene Arabic, derived from /byaakulu/ and /šaaṭira/, respectively, where the short high vowels [i] and [u] in opened syllables are deleted and the long low vowel [aa] is shortened in closed syllables (*byaakulu* > *byaaklu* > *byaklu*). McCarthy's rules wrongly stressed the short penult of such words, which is further deleted. Bohas and Kouloughli suggested that open long syllables (CVV) do not have branching rhymes unless they are the final or penultimate syllable of the word. Word-stress patterns of the three words mentioned above would then be as in (2).

(2)



After publication of McCarthy's article, discussions arose in which Arabic played a major role – Harms (1981:430) says that "Cairo Arabic may be viewed as a sort of 'acid test' for any approach to stress assignment". In the discussions, two important facts were pointed out: first, that metrical structure should also predict or say something about processes of deletion and epenthesis; second, that the phonological substance of segmental units (the sonority of short vowels in Cairene Arabic) may also take part in the construction of metrical trees.

3. PURE GRID THEORY

Classic Tree Theory was soon challenged by Prince's (1983) Pure Grid Theory, based on the conviction that rhythmic notions such as alternation and clash are best represented in grids, and that metrical theory is simplified by eliminating constituency altogether. Grid Theory shares with Tree Theory basic parameters such as Quantity-Sensitivity and Extrametricality, other parameters being Perfect Grid (all even or uneven syllables are marked at the foot level) and End Rules (the first or last mark of the foot level is marked at the word level). The construction of the grid also depends on a few well-formedness conditions, such as the Continuous Column Constraint, according to which a grid containing a column with a mark on layer n+1 and no mark on level n is ill-formed; Avoid Clash, where clash is defined as the adjacency of two marks on layer n without an intervening mark on layer n–1; and Avoid Lapse, where a lapse is a sequence of marks on layer n, none of which has a corresponding mark on layer n+1. In order to fulfill these conditions, de-stressing processes and rhythmic stress shifts may occur, and these appear as simple operations of deletion (*Delete x*), insertion (*Insert x*), and movement (*Move x*). Let us see how word-stress patterns of Damascene and Cairene Arabic are represented within this alternative framework. Goldsmith (1990:197–202) first applies extrametricality to the last syllable of the word, with the exception of open syllables. The rule of Perfect Grid then applies, spreading from left to right, peak first, in Cairene Arabic (3a), and from right to left, trough first, in Damascene Arabic (3b). A grid mark that is placed on the second mora of a heavy syllable is automatically replaced on the first mora by Weak Mora Stress Correction. In both dialects, the End Rule (Final, Word) finally assigns word-level stress to the rightmost unit generated by Perfect Grid in the following examples, that of the words *kitáab* 'book', *katábti* 'you wrote [fem.]', *maktába/máktabe* 'library', and *muxtálifa* 'different [fem.]' (subsequent vowel deletions, which produce final *ktaab* and *məxtəlfe*, are not represented).

(3)

```
                       x                             x                   x                 x
      x x            x x           x x             x x              x   x             x   x
      x xx(x)        x xx(x)       x xx (x)        x xx(x)          xx  x (x)         xx  x x (x)
(a)   ki taa b >     ki taa b      ka tab ti >     ka tab ti        mak ta ba         mux ta li fa

        x              x                           x                    x
        x              x               x           x                x   x
      x xx(x)        x xx(x)       xx  x (x)       xx  x (x)        xx  x  x(x)
(b)   ki taa b       ka tab ti     mak ta be >     mak ta be        mux ta li fe
```

But these rules fail to assign the correct stress-pattern to the very same words that exposed the shortcomings of McCarthy's rules, as shown in (4) (these words have the same word-stress patterns in Cairene and Damascene Arabic):

(4)

```
          x                 x
     x    x            x    x
     xx   x (x)        xx   x(x)
     byaa ku lu        šaa ṭi ra
     /byaakúlu/        /šaaṭíra/
```

Goldsmith only applied the Quantity-Sensitivity Rule to Classical Arabic, where the supposed stress rules are as follows: (i) a final superheavy syllable is stressed; (ii) otherwise, the rightmost heavy syllable (that is not in final position) is stressed; and (iii) if there are no heavy syllables, then the first syllable is stressed. (Whether these principles are false or true is not discussed here.) After the Quantity-Sensitivity Rule has done its job, two end rules apply, to the left edge of the foot-level and to the right edge of the word-level, as shown in (5).

(5)

```
             x               x
     x       x           x   x
     xx x  xx(x)         x  xx x (x)
     kaa ti baa t        yu šaa ri ku

     x                   x
     x                   x
     xx  x  x  x(x)      x x (x)
     mam la ka tu n      ka ta ba
```

Paoli (1993) suggested that Quantity-Sensitivity should be the rule for Arabic because quantity appears to be a major feature of all varieties of the language, and, after Laks (1988), that, in the case of Damascene Arabic, Cairene Arabic, and many other – if not all – Arabic dialects, metrical rules should apply to a domain restricted to the last three syllables, for stress never goes back further. The case of Damascene Arabic would be as represented in (6).

(6)

```
           *             *               *
     *     *          *  *               *
     [** **<*>       [* ** *<*>          [** * *
     dek kaa n       ma daa re s         mad ra se

     *                   *
     *                   *
     [* *  *         ** [* **
     ka ta bu        mut ta ḥi de
```

Within the metrical domain (into square brackets), the final mora of a heavy (CVV or CVC) or superheavy (CVVC or CVCC) syllable is extrametrical. Grid marks are then placed on the first mora of all heavy syllables by the Quantity-Sensitivity Rule, and two end rules apply, to the left edge of the foot-level and to the right edge of the word-level, as is the case in Classical Arabic (Paoli 1993:191–192). The Pure Grid Theory was never fully exploited, except by Selkirk (1984) and Laks (1988, 1992, 1993, 1997). Prince himself finally adopted constituency as early as 1985, with an early Bracketed Grid Theory, and later on with Prosodic Morphology, which he founded in collaboration with McCarthy (McCarthy and Prince 1986, 1990).

4. Bracketed Grid Theory

In 1987, Halle and Vergnaud, following earlier proposals by Hammond (1984) and Prince (1985), proposed a model where constituents are directly built on the grid by means of bracketing at all levels and projection of heads of constituents at the upper level. Grids usually have three levels, or layers (stress-bearing units, feet, word). This theory relies on the same three basic parameters, namely Boundedness, Headedness (constituents are right- or left-headed, i.e. right- or left-dominant feet), and Directionality (left-to-right, right-to-left), to which a parameter that distinguishes head-terminal (SW or WS, i.e. binary) and non-head-terminal (WSW, i.e. ternary) bounded constituents is added. It should be noted that the use of ternary constituents is extremely rare and marginal. The construction of bounded constituents must also fulfill a number of conditions, such as the Exhaustivity Condition, which requires that all stressable elements, unless extrametrical, be head or domain of a constituent; and the Faithfulness Condition, which guarantees that heavy syllables are parsed as heads of line 0 constituents (they are inherently accented elements and the output metrical structure respects their distribution). This condition comes instead of the Quantity-Sensitivity parameter. Extrametricality also takes part in the model. Finally, in a language that has no secondary stress, a rule called Line Conflation eliminates all marks at the second level of the grid (L1), which is conflated with the first (L0). Halle and Kenstowicz (1991:485–488) proposed to apply the following set of rules to word-stress patterns of Damascus Arabic : (i) syllable heads are stress-bearing units; (ii) final syllables are extrametrical unless superheavy; (iii) assign line 1 asterisks to heavy syllables; (iv) on line 0, construct binary left-headed constituents from right to left and assign line 1 asterisks to the heads; (v) on line 1, construct unbounded right-headed constituents and assign a line 2 asterisk to the heads; and (vi) conflate lines 1 and 2. The result is as shown in (7) for the words *kitáab* 'book', *fataḥ* 'he opened', *kátabet* 'she wrote', *mádrase* 'school', *madáares* 'schools', and *muttáḥide* 'united [fem.]'.

(7)

```
     *     *        *           *
(*   *)   (*)      (* .)       (* .)
(*)(*)    (*)<*>   (*  *)<*>   (*  *)<*>
kitaab    fataḥ    katabet     madrase

  *                *
(*   *)         (*   * .)
(*)(*)<*>       (*) (*  *)<*>
madaares       muttaḥide
```

Halle and Vergnaud's theory is cyclic, which means that metrical rules may apply twice or more than twice, former material (level 1 head and the level 0 corresponding foot) being partly preserved. This is the case, for example, when a suffixal pronoun is added to a verb, such as *ʿallaméto* (*ʿallamet+o* 'she taught him') and *fátaḥo* (*fataḥ+o* 'he opened it [masc.]'). Metrical rules first apply to the verb alone (*ʿallam)* and then to the verb plus the pronoun (8a). But the case of *fataḥo* needs the addition of an ad hoc rule that de-stresses a light syllable under stress clash, i.e. if two degenerate feet are constructed over a pair of light syllables, the head of the second one is deleted (8b).

(8)

```
                                    *
      *             *              (*         *)
    (*  *)<*>      (*  *) *  <*>   (*  *)(*)<*>
(a) ʿallamet→      ʿallamet+o→     ʿallamet+o→

          *
    *  * (*)<*>
    ʿallamet+o

                                          *
      *              *              *   *
    (*)<*>          (*)  *<*>      (*)(*)<*>
(b) fataḥ →         fataḥ+o →      fataḥ+o →

      *             (*)             *
    (*)  *  <*>     (*)  *  <*>    (*)  *  <*>
    fataḥ+o →       fataḥ+o →      fataḥ+o
```

(9)

```
                                   *
                    *             (*)
(*  *)<*>       (*  *)<*>        (*  *)<*>
fataḥ+o→        fataḥ+o→         fataḥ+o→

  *
(*  *)<*>
fataḥ+o
```

This solution appeared quite unsatisfactory. Bohas (1986) had already proposed the same set of basic metrical rules, except for (iii), where he suggested assigning a line 1 asterisk to the sole branching rhymes that occur at the end of the metrical domain (final superheavy and penultimate heavy syllable followed by

(10)

*	*	*	*	word layer
(*)	(* .)	(*) (* .)	(* .)(* .)	foot layer
∪ –	∪ ∪ ∪	– ∪ ∪	∪∪∪∪	syllabic/moraic layer
tamaa<m>	*katabi<t>*	*madrasi<h>*	*katabitu*	segmental layer

an extrametrical heavy or light syllable). As for cyclic application of metrical rules, he proposed that no metrical information should be preserved from one cycle to the other except on an autonomous deletion tier, so that syllables which bear stress after the first cycle are prevented from deletion. Rules therefore applied to *fataḥo* as shown in (9) and no de-stressing rule had to be postulated.

5. Later Asymmetric Rhythmic Theories: Hayes (1995) and Angoujard (1990)

Two other theories are focused on here, that of Angoujard (1990), which can be considered the most complete metrical theory of Arabic phonology, and that of Hayes (1995), which was applied to Cairene (Egypt) and Ṣanʿānī (Yemen) Arabic by Watson (2000). Hayes' theory is in a way very similar to that of Halle and Vergnaud. Watson (2000:84–86), after Hayes (1995:2–3), defines the following basic principles: Metrical structure is represented by the bracketed grid, which is a hierarchy of rhythmic beats grouped into a hierarchy of constituents. The grid has four layers (moraic, syllable, foot, and word), plus a supraword phrase layer for sentences. As in Halle and Vergnaud (1987a,b), the grid columns are subject to the Continuous Column Constraint. The smallest metrical unit is the foot. Feet are either bounded or unbounded, and there are three common bounded foot types: the moraic trochee, the syllabic trochee (both left-headed), and the iamb (right-headed). The basis of the foot inventory is a principle called the Iambic/Trochaic Law, which forms part of the theory of rhythm and determines the set of possible feet and usual rhythmical patterns. According to this principle, and contrary to all previous theories, metrical structure creation is nonexhaustive (some syllables may be left unaccounted for), and degenerate feet are disallowed unless in strong position if specified. As for extrametricality, it may occur at a designated edge (left or right), and peripheral moras, as well as peripheral syllables and feet, may be considered extrametrical. In Cairene Arabic (Watson 2000:93–98), the last consonant is extrametrical (that is, the last is extrametrical if consonantal), and feet are moraic trochees constructed from left to right, degenerate feet being absolutely forbidden. Finally, a right end rule applies at the word level, as shown in (10) for words *tamáam* 'perfect, okay', *kátabit* 'she wrote', *madrásih* 'school', and *katabítu* 'she wrote it [masc.]'.

As for the crucial word *šaṭra* 'talented', Watson, as Halle and Vergnaud did, first assigns stress to the penult (11a), before the initial long vowel is subjected to unstressed long vowel shortening (11b). The resulting initial syllable can no longer support a foot. The foot parse applies again, locating the penultimate CV syllable in the weak element of the foot, so that it can be deleted (11c).

(11)

*	*	*
(*)(* .)	(* .)	(*)
– ∪ ∪	∪∪∪	– ∪
(a) *šaaṭira*	(b) *šaṭira*	(c) *šaṭra*

Stress assignment in Ṣanʿānī Arabic differs in two principal ways from that in Cairene Arabic (Watson 2000:98–106). First, stress falls on a heavy penult even if the ultima is a superheavy syllable (*jázzaar* 'butcher'). Second, as in Damascene Arabic, a heavy antepenult is stressed if the penult is light (*mádrasih* 'school'). The rules featured by Watson also differ from those of Cairene Arabic in two ways: first, degenerate feet are permitted in strong position; second, the last foot of the word is considered extrametrical.

Angoujard's (1990) theory, wrongly neglected by American scholars, remains the only one which adequately accounted for the whole range of Arabic phonological and metrical systems. Breaking with the classic approach, Angoujard suggested that the structure of binary feet relies on three universally defined hierarchies, of sonority (*a* > *i*, *u* > *e*, *o* > *ē* > Ø), of rhymes (VC > VV > V), and of feet ([HL] and [LH]

(12)

```
   *          *          *          *            *            *
 * *         * *       *   *      *     *      * * *        *   *   *
 * * *       * *       * * *      *   * *      * * *        *   * * *
 kitaab     fataḥ     katabet    madrase      madaares     muttaḥide

 L H H      L H       L L H      H L H        L H H        H L L H
 |  |  |    |  |      [+ –] |    [+ –] |      |  |  |      | [+ –] |
 |  |  |    |  |        V   |      V   |      |  |  |      |   V   |
 P  P  P    P P         P   P      P   P      P  P  P      P   P   P
```

> [LL], where H and L stand for 'light' and 'heavy' respectively). Whereas hierarchies of rhymes and feet are relevant in all Arabic dialects, that of sonority is not. Finally, extrametricality is rejected, and, on the contrary, all final syllables are considered heavy (as is the case in Arabic traditional phonetics as well as in Classical Arabic verse). According to Angoujard, the dialect of Damascus is one that establishes a clear difference between the low short vowel /a/ and all other vowels /i, u, e, o, ə/, so that feet are constructed according to the fact that light syllables with a short non-low vowel have priority to integrate into binary feet as weak positions. The final syllable is always in strong position, and feet are left-dominant ([+ –]). Stress is finally assigned to the strong position of the penultimate foot, as shown in (12).

Cairene Arabic also makes use of the sonority scale and of the distinction between strong /a/ and weak /i/ and /u/, which appears to be the only valid explanation for the contrast between the output of /šaaṭira/, i.e. *šáṭra*, where the unstressed short vowel [i] of the second syllable has been deleted, and the output of /ʕaalami/ 'my world', i.e. *ʕalámi*, where the second short low [a] bears word stress; and between *fíhmit* 'she understood' and *ḍárabit* 'she hit'.

6. THE DEBATE ABOUT CONSTITUENCY

According to Halle and Vergnaud, "The placement of stress reflects an organization of the sequence of stressable elements that is not concerned with the phonological or phonetic substance of these elements" (1987a:46). Stressable elements are considered mere positions, identified by their rank counted from right to left or from left to right, and constituency "has no direct or uniform correlates". On the contrary, Angoujard suggested that constituency, while remaining strictly binary, may sometimes rely on the phonological characteristics of the stress-bearing units, according to their relative ranking within the sonority scale. Whereas Angoujard assumed that constituency has external phonological correlates and is fully motivated by the phonological substance of stressable elements, Halle and Vergnaud suggested that "it can only be detected [and justified] indirectly through its effects on phonological rules", by means of purely internal evidence, i.e. by justifying the need for constituency by its explanatory power. Such arguments for metrical constituency were advanced, based on a number of supposed cases of deletion of stressed vowels, which do not result in the deletion of stress altogether but rather into its migration to an adjacent stress-bearing unit, so that constituency and foot-dominance predict what the direction of stress shift will be: if feet are left-dominant, stress shifts to the right, and if feet are right-dominant, stress shifts to the left. Such phenomena were described in Russian and Sanskrit (Halle and Vergnaud 1987a,b), as well as in Tiberian Hebrew (Prince 1985; Rappaport 1976) and some Arabic dialects (Kenstowicz 1983; al-Mozayni a.o. 1985; Hayes 1995). Al-Mozayni a.o. (1985), for instance, found that syncope in Bedouin Ḥijāzī Arabic leads to migrations of stress, whose direction depends on the shape of the metrical tree. In this dialect, as in Damascene Arabic, stress is on superheavy final syllables; if there is no superheavy final syllable, it falls on a heavy penult; if there is no heavy penult, it falls on the antepenult. According to Al-Mozayni a.o. (1985:136), a rule of Low Vowel Deletion deletes short /a/ in an open syllable if the following syllable is also open and contains short /a/ (a → Ø / C_ $[\mathrm{Ca}]_S$). This rule, they say, produces alternations such as *sáḥab* 'he pulled', *saḥábna* 'we pulled' vs. *sḥáʾbat* 'she pulled'. A particular interaction between stress and Low Vowel Deletion is revealed by alternations such as *ʾínkasar* 'he got broken' vs. *ʾinksárat* 'she got broken' (from /ʾinkasarat/). Stress assignment cannot follow Low Vowel Deletion, since this would wrongly produce *ʾínksarat. Stress assignment thus occurs before Low Vowel Deletion, and the deletion of the vowel in the head of a foot results in a rightward stress shift within the foot: *ʾinkásarat* > *ʾinksárat*. This process can be represented within Halle and Vergnaud's bracketed grid as in (13), where rules are the same as those of Damascene Arabic mentioned above in Section 4 (after Halle and Kenstowicz 1991:485–488):

(13)

	*	*	*
	(* .)	(* * .)	(* *)
	(* *)<*>	(*)(* *)<*>	(*)(*)<*>
(a)	*ʾinka sar*	(b) *ʾinkasa rat*	→ *ʾinksa rat*
	'he got broken'	'she got broken'	

Angoujard (1992), following Cantineau (1936) and Johnstone (1967), observed that Bedouin Ḥijāzī Arabic, as well as the Bani Ḥassān Jordanian dialect, is a 'differential dialect', which has two main types of regular past tense paradigms: /CaCaC/, which usually surfaces as [CiCaC] (*kítab* 'he wrote') or [CaCaC] in a guttural context (*sáḥab* 'he pulled'); and /CaCiC/, which surfaces as [CiCiC] (*šírib* 'he drank'). As for the feminine singular, the three verbs above have *ktíbat* 'she wrote', *ṣḥábat* 'she pulled', and *šírbat* 'she drank', respectively. As it seems, Al-Mozayni a.o. (1985), as well as Kenstowicz (1983), only considered the first of these two paradigms. The analysis of the complete facts leads Angoujard to suggest the following rules, which apply to both dialects: The three universally defined hierarchies are relevant; the hierarchy of sonority distinguishes between strong low vowels /a/ and weak high vowels. Feet are right-headed, and their construction is based on two principles: (i) [CV CV] are constructed first; and (ii) [CV CVC] are constructed at the end of the word, so that the metrical structure of the above crucial examples is as shown in (14), where vowels of light syllables in weak positions are subsequently deleted:

(14)

*	*	*	*	*
* *	* *	* *	* *	* * *
* *	* * *	* *	* * *	* * * *
saḥab	*saḥabat*	*šarib*	*šaribat*	*ʾinkasarat*
(L)(H)	(L L)(H)	(L)(H)	(L)(L H)	(H)(LL)(H)
	[– +]		[– +]	[– +]
	Ø		Ø	Ø
['ᵴɑħab]	['ᵴħabat]	['ʃirib]	['ʃirbat]	[ʔin'ksarat]

Angoujard has shown convincingly that word-stress patterns and vowel deletions of these Arabic dialects, as well as that of Tiberian Hebrew, can better be described without postulating elision of stressed vowels and 'migration' of stress, provided that constituency is motivated by significant phonological criteria. Paoli (1993) also argued that Pure Grid Theory could adequately and easily account for these differential dialects: the hierarchy of sonority operates within the metrical domain over sequences of light syllables which the Quantity-Sensitivity Rule failed to distinguish. Other cases of alleged stress shifts, in Sanskrit and Russian, have been extensively discussed by Laks (1997:123–138), who underlined the fact that traditional descriptions of stress and vowel deletion, in both languages, never suggested that stressed vowels could be deleted, and argued for a Pure Grid Theory, which allows an adequate and simple analysis of the empirical data under discussion where no stressed vowels need to be deleted.

It follows from what has been said that the argument based on the behavior of stress shifts accompanying alleged deletions of stressed vowels, as well as alleged cases of epenthesis of stressed vowels in Winnebago, which were refuted by Laks (1997:139–169), cannot be considered a valid or relevant argument for constituency.

7. Results and prospects

Metrical phonology has allowed great progress in our perception and understanding of stress phenomena and word-stress patterns of natural languages, and this is particularly true as far as Arabic is concerned. As we have seen, Arabic vernaculars received a great deal of attention within the various theoretical models which developed out of the original framework of metrical phonology. But the rules proposed often failed to account for all the facts described, mainly because of a very restrictive definition of constituency, strictly binary and totally independent from the phonetic and phonological substance of the stress-bearing units. In this regard, Angoujard's theory, in which constituency not only relies on quantity but also on the relative sonority of the stress-bearing units, appears to be much more suitable to account for the rhythmic structure and word-stress patterns of Arabic. As for Pure Grid Theory, though not fully exploited, it seems that it was able to account more simply for the facts under consideration, so that the debate about constituency should remain open, all the more so as the main arguments advanced to justify its relevance were shown to be invalid.

Bibliographical references

Al-Mozayni, Hamza Qublan, Robert Bley-Vroman, and John J. McCarthy. 1985. "Stress shift and metrical structure". *Linguistic Inquiry* 16.135–144.

Angoujard, Jean-Pierre. 1990. *Metrical structure of Arabic*. Dordrecht: Foris.

——. 1992. "Syncope et voyelles accentuées (parler du Hedjaz, Bani Hassan et hébreu biblique)". *Langues Orientales Anciennes, Philologie et Linguistique* 4.299–304.

Bohas, Georges. 1986. "Accentuation et effacement de schwa dans le parler de Damas". *Revue Québecoise de Linguistique* 16:1.39–60.

——. and Djamel Eddine Kouloughli. 1981. "Processus accentuels en arabe". *Analyses Théorie* 2.1–59.

Cantineau, Jean. 1936. "Etude sur quelques parlers nomades d'Orient". *Annales de l'Institut d'Etudes Orientales d'Alger* 2–3.1–121.

Goldsmith, John. 1990. *Autosegmental and metrical phonology*. Oxford: Blackwell.

Halle, Morris and Michael Kenstowicz. 1991. "The Free Element Condition and cyclic versus noncyclic stress". *Linguistic Inquiry* 22.457–501.

Halle, Morris and Jean-Roger Vergnaud. 1978. "Metrical structures in phonology". Unpubl. ms. Cambridge: Massachusetts Institute of Technology.

——. 1987a. "Stress and the cycle". *Linguistic Inquiry* 18.45–84.

——. 1987b. *An essay on stress*. Cambridge, Mass.: MIT Press.

Hammond, Michael. 1984. *Constraining metrical theory: A modular theory of rhythm and destressing*. Ph.D. diss., University of California, Los Angeles.

Harms, Robert. 1981. "A backwards metrical approach to Cairo Arabic stress". *Linguistic Analysis* 7:4.429–450.

Hayes, Bruce. 1980. *A metrical theory of stress rules*. Ph.D. diss., Massachusetts Institute of Technology.

——. 1995. *Metrical stress theory: Principles and case studies*. Chicago: University of Chicago Press.

Johnstone, T.M. 1967. *Eastern Arabian dialect studies*. London: Oxford University Press.

Kenstowicz, Michael. 1983. "Parametric variation and accent in the Arabic dialects". *Papers from the Chicago Linguistic Society*, ed. Amy Chukerman, Mitchell Marks, and John F. Richardson, 205–213. Chicago: Chicago Linguistic Society.

Laks, Bernard. 1988. "Des grilles et des arbres". *Recherches Linguistiques de Vincennes* 17.135–195.

——. 1992. "Constituance et métrique". *Architecture et géométrie des représentations phonologiques*, ed. Bernard Laks and Anne Rialland, 25–59. Paris: CNRS.

——. 1993. "La constituance revisitée". *De natura sonorum: Essais de phonologie*, ed. Bernard Laks and Marc Plénat, 173–220. Saint-Denis: Presses Universitaires de Vincennes.

——. 1997. *Phonologie accentuelle*. Paris: CNRS.

Liberman, Mark. 1975. *The intonational system of English*. Ph.D. diss., Massachusetts Institute of Technology.

——. and Alan Prince. 1977. "On stress and linguistic rhythm". *Linguistic Inquiry* 8.249–336.

McCarthy, John. 1979. "On stress and syllabification". *Linguistic Inquiry* 10.443–466.

——. and Alan Prince. 1986. *Prosodic morphology 1986. Report*. New Brunswick, N.J.: Rutgers University Center for Cognitive Science. (Available at http://ruccs.rutgers.edu/pub/papers/pm86all.pdf.)

——. and Alan Prince. 1990. "Foot and word in prosodic morphology: The Arabic broken plural". *Natural Language and Linguistic Theory* 8.209–283.

Paoli, Bruno. 1993. "Constituance ou domaine en phonologie métrique: Le cas de l'hébreu biblique et des parlers arabes de Damas, du Hijaz et de Jordanie (Beni Hasan)". *Linguistica Communicatio* 5:1/2.177–209.

Prince, Alan. 1983. "Relating to the Grid". *Linguistic Inquiry* 14.19–100.

——. 1985. "Improving Tree Theory". *Proceedings of the Berkeley Linguistics Society*, ed. M. Niepokuj, M. Van Clay, V. Nikiforidou, and D. Jeder, 471–490. Berkeley, Calif.: Berkeley Linguistics Society.

Rappaport, Malka. 1984. *Issues in the phonology of Tiberian Hebrew*. Ph.D. diss., Massachusetts Institute of Technology.

Selkirk, Elizabeth. 1984. *Phonology and syntax: The relation between sound and structure*. Cambridge, Mass.: MIT Press.

Watson, Janet. 2000. *The phonology and morphology of Arabic*. Oxford: Oxford University Press.

Bruno Paoli (University of Bordeaux-3)

Phonotactics

1. Definition

Phonotactics are the patterns of co-occurrence and avoidance between phonological units in a syllable, word, or phrase.

2. Consonant co-occurrence in Arabic

Arabic has phonotactic restrictions between consonants within the verbal roots that have played an important role in the development of phonological theory. Greenberg (1950) was the first systematic quantitative study of these patterns. However, Greenberg noted that these patterns of co-occurrence among consonants in the verbal roots were known to traditional Arabic grammarians long ago. Current quantitative descriptions of consonant co-occurrence in Arabic are generally based on dictionaries of Modern Standard Arabic, and so are based on a 'standard' Arabic consonant inventory. This inventory is an idealized inventory that roughly corresponds to the historical basis of the modern Arabic dialects (McCarthy 1994). Note that all descriptions of the co-occurrence restrictions for Arabic consonants assume that the restrictions are limited to the root consonant sequences in verbs; they do not apply within nouns or to co-occurrence between verbal root consonants and suffix, prefix, or infix consonants.

Arabic verbal roots are typically triliteral, containing three consonants. The first phonotactics observation that was made is that there are no verbal roots that repeat the same consonant in first and second position, e.g. **dadam*. Many verbs are found with identical consonants in the second and third positions of the root. Examples include *madad* 'to stretch' and *farar* 'to flee'. Some of these verbs appear to be cases where a historical suffix has been reanalyzed as part of the root (Elmedlaoui 1995).

More interestingly, Greenberg also examined the statistical rate of co-occurrence of consonant pairs relative to an estimate of expected co-occurrence based on chance. Chance co-occurrence is based on the frequency with which each of the consonants in the pair occurs in the dictionary. This analysis was further refined by Pierrehumbert (1993), who used the ratio of observed co-occurrence to the chance rate of co-occurrence as an index of phonotactic compatibility. Pierrehumbert's measure, known as O/E, provides a measure of co-occurrence frequency with clearly interpretable results. When O/E is near zero, there are very few observed pairs relative to the number of pairs expected by chance, suggesting that a phonotactic constraint may be present. When O/E is near 1 or larger, then combinations of the consonants are quite common, and there is no evidence for a constraint against their co-occurrence. If O/E is large, then there might be a constraint requiring co-occurrence, as in cases of assimilation or harmony.

A simple application of the O/E measure can be demonstrated by considering specific root pairs. For example, roots of the form /d t C/ (where C is any consonant) are not found. The expected frequency of this combination by chance is computed by multiplying the frequency of roots beginning in /d/ and the frequency of roots with /t/ in second position and dividing by the total size of the dictionary. For example, in the lexicon used by Frisch a.o. (2004), adapted from the Wehr dictionary (Cowan 1979), 2.3 such roots are expected (91 /d/ initial, 68 /t/ in second position, in a dictionary of 2,676 total roots: 91 × 68 ÷ 2,676 = 2.3). In this case, O/E = 0 ÷ 2.3 = 0, indicating strong underrepresentation. Two roots contain /d s C/, and 2.9 are expected, giving an O/E of 0.69 (weaker underrepresentation). There are 4 roots with /d g C/, and 3.3 are expected at random, giving an O/E of 1.21 (overrepresentation). The quantitative linguistic analysis of co-occurrence usually does not focus on such specific root pairs, as the number of occurrences is small, and also because linguistic generalizations are not expected to be so specific. For example, a study of the co-occurrence of /d/ and /t/ might consider the total observed vs. expected occurrence for all six combinations of /d/ with /t/ in computing the O/E ratio, i.e. /d t C/, /t d C/, /C d t/, /C t d/, /d C t/, and /t C d/. McCarthy (1994) grouped obstruents that only differed in voicing, and so for example aggregated all combinations of /t/ or /d/ with other pairs that do not differ by voicing, such as /s/ or /z/.

The major co-occurrence classes in the Arabic roots are shown with aggregate O/E ratios in Table 1. Roots are commonly found where each consonant comes from a different class

Table 1. Co-occurrence of consonant pairs in Standard Arabic, aggregated by major class and distance (major classes are shaded)

				Adjacent				
		Labial	Cor obs		Dorsal	Guttural		Cor son
		b f m	t d ṭ ḍ	ṯ ḏ s z ṣ ḏ̣ š	k g q	x ġ	ḥ ʿ h ʾ	l r n
Labial	b f m	0.00	1.37	1.31	1.15	1.35	1.17	1.18
Cor obs	t d ṭ ḍ		0.14	0.52	0.80	1.43	1.25	1.23
	ṯ ḏ s z ṣ ḏ̣ š			0.04	1.16	1.41	1.26	1.21
Dorsal	k g q				0.02	0.07	1.04	1.48
Guttural	x, ġ					0.00	0.07	1.39
	ḥ ʿ h ʾ						0.06	1.26
Cor son	l r n							0.06

				Nonadjacent				
		Labial	Cor obs		Dorsal	Guttural		Cor son
		b f m	t d ṭ ḍ	ṯ ḏ s z ṣ ḏ̣ š	k g q	x, ġ	ḥ ʿ h ʾ	l r n
Labial	b f m	0.30	1.08	1.02	1.26	1.25	1.28	1.11
Cor obs	t d ṭ ḍ		0.38	1.06	1.24	1.05	1.02	0.97
	ṯ ḏ s z ṣ ḏ̣ š			0.24	1.16	1.35	1.14	1.23
Dorsal	k g q				0.07	0.68	1.19	1.03
Guttural	x, ġ					0.25	0.12	1.10
	ḥ ʿ h ʾ						0.34	1.13
Cor son	l r n							0.67

(O/E > 1), but roots are rarely found with two consonants that come from the same class (O/E near 0). These generalizations are based on quantitative tendencies. In many cases, there are a few examples of roots that contain two consonants from the same class, as evidenced by the O/E near zero. However, there are far fewer examples of these combinations than would be expected if consonants could co-occur freely, at random. Greenberg noted that the classes are, for the most part, based on place of articulation.

A variety of additional subgeneralizations have been observed by later researchers that further refined Greenberg's place-based co-occurrence classes. McCarthy (1988, 1994) and Padgett (1995) observe that the coronal obstruents subdivided by manner of articulation as combinations of stops and fricatives (O/E = 0.14 and 0.04 for adjacent pairs) are more common than combinations of two stops or two fricatives (O/E = 0.52). Pierrehumbert (1993) showed that the liquids /l, r/ are more likely to be found in combination with a nasal than with one another. Frisch a.o. (2004) demonstrate that obstruents that differ in voicing are more likely to combine than obstruents that are the same in voicing. They also observed that the emphatic coronals (articulated with pharyngeal constriction) are less likely to co-occur with one another, and also less likely to occur with the dorsal and uvular stops.

In addition to these class-based co-occurrence restrictions, Greenberg (1950) observed that the co-occurrence constraints are most clearly seen in pairs of adjacent root consonants (C_1C_2 or C_2C_3). However, the same co-occurrence classes also restrict nonadjacent root consonants (C_2C_3), but not as strongly. In other words, more exceptions to the co-occurrence classes are found for nonadjacent consonants, but the classes are the same. Only in the case of the coronal stop and fricative pairs does the O/E for nonadjacent pairs become greater than 1, but this was also the weakest restriction for adjacent pairs.

In almost all cases, studies of consonant co-occurrence in Arabic have not differentiated the order of consonants combinations. However, Elmedlaoui (1995) noted some asymmetries in the co-occurrence patterns. In particular, combinations of stops followed by fricatives tend to be less common than combinations of fricatives followed by stops. Elmedlaoui (1995) claimed this was part of a more general pattern for roots to have a falling consonant sonority profile.

3. Autosegmental analysis of co-occurrence

McCarthy (1981, 19 86, 1988, 1994) provided the first formalization of the consonant co-occurrence restrictions in the Arabic verbal roots, using the notation of autosegmental phonology. An example of the autosegmental representation of Arabic verbal root morphology is given in (1). In this representation, the consonant and vowel phonemes are located on distinct autosegmental tiers, and they are interleaved based on their links to a skeletal tier. The lexical content of the root is contained within the consonantal tier. Grammatical content is contained in the vowel tier and CV-skeleton tier. Each of the verbal root classes or binyanim is derived by combining the lexical consonant root with distinct vowel and CV-skeleton tiers for the verb class.

```
(1) vowel tier:          u     i
                         |     |
    skeletal tier:    C  V  C  V  C
                      |     |     |
    consonant tier:   k     t     b
```

The multitiered representation in (2) provides a distinct domain of representation for the consonant sequence as a lexical root and so provides a domain over which the phonotactic constraint over consonant co-occurrence applies. McCarthy (1986) first proposed an account of the distribution of identical consonant pairs, e.g. *madad*, but **dadam* using this representation. Assuming association between tiers in Arabic proceeds from left to right, an underlying root form with just two consonants, e.g. /m d/, would result in multiple association from the consonant tier to the skeletal tier, as shown in (2).

```
(2) C V C V C
    |    \  /
    m     d
```

It is also theoretically possible to have a triliteral consonant tier with a repeated final segment or, for that matter, with a repeated initial segment. Following work on the phonology of tone, McCarthy proposed that the → Obligatory Contour Principle (OCP) applies to the Arabic root consonant tier, disallowing repeated items. Given the OCP, the only way that a repeated item can appear in the surface form is through multiple association, and so it will only occur with final consonants. A more recent proposal within → Optimality Theory follows in the spirit of McCarthy's analysis, taking these forms to involve copying through the process of → reduplication (Gafos 2003; → stem).

McCarthy (1988) proposed extending the use of autosegmental representation and the OCP to individual place feature tiers in order to account for the more general co-occurrence groups identified by Greenberg (1950). This constraint is referred to as OCP-Place. For example, a hypothetical root like */f t b/ would be represented as in (3). Individual features connected with place of articulation each occupy their own tier, and the form is ungrammatical because adjacent identical features on the labial tier violate OCP-Place.

```
(3) labial tier:   [lab]          [lab]
                     |              |
    skeletal tier:  *C       C      C
                             |
    coronal tier:          [cor]
```

In the original McCarthy (1988) analysis, the consonants are divided into co-occurrence classes by their place of articulation features. The presence of additional subclasses forced further refinements of the representation which tied into the development of feature geometry within phonological theory (McCarthy 1994; Padgett 1995). In this extension, manner features like [sonorant] and [continuant] are linked to the place tier, rather than directly to the skeletal tier. OCP-Place is assumed to be violated only if the entire subsidiary feature structures for two instances of a place of articulation feature are identical. As a result, the coronal obstruent/sonorant co-occurrence classes as well as the stop/fricative co-occurrence classes are differentiated. Overall, the co-occurrence patterns in Arabic played an important role in the development of autosegmental theory and feature geometry, as the patterns were taken as evidence for tier segregation of morphological and phonological information in the representation.

4. SIMILARITY ANALYSIS OF CO-OCCURRENCE

Pierrehumbert (1993) claimed that the autosegmental account, and any potential refinements of it, are incompatible with the gradient nature of the co-occurrence data. In the autosegmental account, both the original OCP and OCP-Place operate with the same formal mechanism. They both disallow identity at some level of representation. However, the original OCP is never violated, while OCP-Place is violated to various degrees, especially when nonadjacent consonants and the manner subclasses are considered. In order to explain the difference in degree between different OCP violations, a different mechanism for the co-occurrence restrictions is needed.

Pierrehumbert (1993) and Frisch a.o. (2004) link the gradience of the co-occurrence patterns to the perceived similarity of the consonants involved as they occur in the root. They claim the constraint is one of similarity avoidance and the gradience of the co-occurrence pattern follows from the gradience of similarity. For example, consonants that share manner features in addition to place features are more similar to one another than consonants that do not share manner features. The other subgeneralizations, involving coronal sonorants, emphatic consonants, and voicing, are all predicted by similarity between phonemes. In addition, the similarity of consonants that are adjacent to one another is more salient than the similarity of consonants that are separated by an intervening consonant, accounting for the effect of distance on the rate of co-occurrence. The diminishing restriction on co-occurrence over distance has also been demonstrated for quadrilateral roots in the Semitic language Tigrinya (Buckley 1997). In addition, the effects of multiple OCP-Place violations have been shown to be cumulative, so that roots that contain more than one violation are even less common than roots that contain only a single violation (Frisch 2000). Finally, quantitative OCP-Place effects have been found in a variety of languages other than Arabic (Berkley 2000; Frisch a.o. 2004; Padgett 1995). In general, these other languages show weaker effects than are found in Arabic, often with no absolute restriction on co-occurrence.

Taken together, these findings point toward the need for a gradient account of the phenomenon that goes beyond the previous representational solution. Frisch a.o. (2004) link co-occurrence to similarity stochastically using an S-shaped categorization function and a similarity metric based on features and natural classes. A second approach to gradient co-occurrence patterns uses the formalism of → Optimality Theory. Optimality Theory allows constraints to be violated when a constraint of higher importance needs to be satisfied. This account uses constraints against combinations of feature co-occurrence in a manner that replicates the similarity hierarchy symbolically (Alderete 1997; MacEachern 1999). In this type of account, constraints that prevent the combination of several features (e.g. both place and manner co-occurrences) are ranked more highly than constraints that prevent the combination of fewer features (e.g. place-only co-occurrence). The ranking of these constraints relative to faithfulness constraints (which require all underlying contrasts to be preserved) determines the degree of co-occurrence that is allowed or disallowed. In order to predict quantitative patterns of co-occurrence in the lexicon, an additional mechanism that permits variable ranking of constraints is needed (Hayes 2000). Even with these modifications, there are still some conceptual difficulties in using Optimality Theory to predict the distribution of lexical forms (Pierrehumbert 1999).

Note that the similarity approaches to consonant co-occurrence rule out roots that contain adjacent identical consonants and thus cannot predict the occurrence of *madad* type forms without an additional mechanism. Frisch a.o. (2004) adopt the autosegmental representation approach, applying the similarity to the underlying consonant root form, and assuming that the surface *madad* forms are derived through a process of spreading or reduplication. Accounts using Optimality Theory could adopt additional higher ranked constraints that take precedence over similarity avoidance (e.g. Rose 2000).

5. Functional basis for dissimilation

One avenue of recent research in phonological theory ties phonological patterns to functional phonetic patterns (e.g. Hayes and Steriade 2004).The gradient nature of the consonant co-occurrence constraints in Arabic and their analysis as a pattern of similarity avoidance contribute to this literature as well. In particular, there are functional reasons to have lexical items that avoid repeated phonemes or highly similar phonemes (Berg 1998; Frisch 2004). Forms without repetition are easier to produce, perceive, and hold in short-term memory. The patterns in Arabic are claimed to provide particularly compelling evidence for a functional basis due to their gradient nature (Frisch 2004). Occurrence in the lexicon is least frequent for forms with the highest level of similarity, which would be the forms that are most difficult to process. Forms with a lower level of similarity would be easier to process, and they occur more frequently in the lexicon.

Bibliographical references

Alderete, John. 1997. "Dissimilation as local conjunction". *Proceedings of the North East Linguistics Society* 27.17–32.

Berg, Thomas. 1998. *Linguistic structure and change: An explanation from language processing*. Oxford: Clarendon Press.

Berkley, Deborah Milam. 2000. *Gradient OCP effects*. Ph.D. diss., Northwestern University.

Buckley, Eugene. 1997. "Tigrinya root consonants and the OCP". *Penn Working Papers in Linguistics* 4:3.19–51.

Cowan, J Milton (trans.). 1979. *Hans Wehr: A dictionary of Modern Written Arabic*. Wiesbaden: O. Harrassowitz.

Elmedlaoui, Mohamed. 1995. "Géométrie des restrictions de cooccurrence de traits en sémitique et berbère". *Canadian Journal of Linguistics* 40:1.39–76.

Frisch, Stefan A. 2000. "Temporally organized representations as phonological units". *Papers in laboratory phonology*. V. *Acquisition and the lexicon*, ed. Michael B. Broe and Janet B. Pierrehumbert, 283–298. Cambridge: Cambridge University Press.

——. 2004. "Language processing and segmental OCP effects". *Phonetically based phonology*, ed. Bruce Hayes, Robert Kirchner, and Donca Steriade, 346–371. Cambridge: Cambridge University Press.

——, Janet B. Pierrehumbert, and Michael B. Broe. 2004. "Similarity avoidance and the OCP". *Natural Language and Linguistic Theory* 22.179–228.

Gafos, Adamantios. 2003. "Greenberg's asymmetry in Arabic: A consequence of stems in paradigms". *Language* 79.317–357.

Greenberg, Joseph H. 1950. "The patterning of root morphemes in Semitic". *Word* 5.162–181.

Hayes, Bruce. 2000. "Gradient well-formedness in Optimality Theory". *Optimality Theory: Phonology, syntax, and acquisition*, ed. Joost Dekkers, Frank van der Leeuw, and Jeroen van de Weijer, 88–120. Oxford: Oxford University Press.

——. and Donca Steriade. 2004. "Introduction: The phonetic bases of phonological markedness". *Phonetically based phonology*, ed. Bruce Hayes, Robert Kirchner, and Donca Steriade, 1–33. Cambridge: Cambridge University Press.

MacEachern, Margaret. 1999. *Laryngeal co-occurrence restrictions*. New York: Garland.

McCarthy, John J. 1981. "A prosodic theory of nonconcatenative morphology". *Linguistic Inquiry* 12.373–418.

——. 1986. "OCP effects: Gemination and antigemination". *Linguistic Inquiry* 17.207–263.

——. 1988. "Feature geometry and dependency: A review". *Phonetica* 43.84–108.

——. 1994. "The phonetics and phonology of Semitic pharyngeals". *Papers in laboratory phonology* III, ed. Patricia A. Keating, 191–283. Cambridge: Cambridge University Press.

Padgett, Jaye. 1995. *Stricture in feature geometry*. Stanford, Calif.: CSLI Publications.

Pierrehumbert, Janet B. 1993. "Dissimilarity in the Arabic verbal roots". *Proceedings of the North East Linguistics Society* 23.367–381.

——. 1999. "Formalizing functionalism". *Formalism and functionalism in linguistics*, ed. Mike Darnell, Edith A. Moravcsik, Frederick Newmeyer, Michael Noonan, and Kathleen M. Wheatley, I, 287–305. Amsterdam and Philadelphia: J. Benjamins.

Rose, Sharon. 2000. "Rethinking geminates, long-distance geminates and the OCP". *Linguistic Inquiry* 31.85–122.

Stefan A. Frisch (University of South Florida)

Phraseology

1. Introduction

Phraseology (also 'idiomaticity') is a branch of linguistics which deals with fixed combinations of words whose meaning cannot be deduced from the conjoined meanings of their components. Each combination is normally termed an 'idiom', e.g. *at sixes and sevens; to put up with; to kick the bucket*. It is to be distinguished from cognate terms, such as → 'collocation', i.e. a fixed combination whose components retain their literal meanings, e.g. *economic sanctions*; → 'compound', i.e. a phrase consisting of two or more words, rendering a new meaning, e.g. *breakfast*; 'blend' or 'portmanteau', i.e. an amalgamation of parts of two words or more yielding a new meaning, e.g. *breakfast* + *lunch* = *brunch*; → 'metaphor', i.e. a word that has undergone a semantic or stylistic shift of meaning, e.g. *a fox* when referring to a cunning person; and → 'proverb', i.e. a common expression that may be understood literally but also conveys a deeper idea, e.g. *learn to walk before you run*.

In spite of the limitations of the above definition (many idioms contain one idiomatic element; some are not completely fixed; several idioms are borderline cases because they may also be regarded as figures of speech, mainly metaphors, or proverbial sayings, which may also be classified as idioms), the conventional general definition is adhered to in the present discussion.

Although 'idiom' has several names in Arabic, such as *qawl maʾṯūr/sāʾir, ʿibāra maʾṯūra, taʿbīr xāṣṣ, lafḏ/tarkīb/taʿbīr maskūk, ʿibāra/ṣīġa maskūka*, and a few more, the majority of classical Arabic linguists and lexicographers treat the idiom as a → proverb (*maṯal*). Thus, medieval collections of proverbs, such as al-Maydānī's, contain a large number of idioms along with proverbs, in the same way that *maṯal* refers to 'adage', 'dictum', 'aphorism', 'apothegm', 'maxim', 'saw', 'saying', and the like. However, the majority of modern works prefer to name the idiom *taʿbīr iṣṭilāḥī* or *ʿibāra iṣṭilāḥiyya*.

2. Origin, source, and derivation of Arabic idioms

Although the coiners of the overwhelming majority of the Arabic idioms are anonymous, the sources of some of the idioms can be traced, while their derivations may easily be classified according to their subject matter. Thus, the idiom *al-yawm ḥamiya l-waṭīs* 'there was fierce fighting' is attributed to the Prophet Muḥammad, who coined it on the day of the Battle of Ḥunayn in 630 C.E. (cf. Ibn al-ʾAṯīr, *Maṯal* 49–50). The idiom *lā yamliku qiṭmīr* 'he does not own a thing' is based on the *Qurʾān* (Q. 35/13), while the idiom *istansara l-buġāṯ* lit. 'the sparrow has become like an eagle', i.e. 'the weak pretends to be or has become strong', is derived from the animal world. However, as in the case of proverbs, reliance cannot be automatically placed on the information contained in background stories provided by lexicographers, or echoed in the idiom itself about the inventor of the idiom or the circumstances of its coining. Thus, for the idiom *ʿalā ʾahlihā tajnī Barāqiš* 'Baraqish has brought disaster on her own people', the *Lisān al-ʿArab* gives five different versions of the circumstances in which the idiom was coined.

The following are the main sources of Arabic idioms:

i. The *Qurʾān*. Examples: *qurraʿayn* (Q. 25/74, 28/9, 32/17) 'consolation, comfort'; *ʿalā qāb qawsayn* (Q. 53/9) 'very close; imminent'; *mā nazzala/ʾanzala llāh bihā min sulṭān* (Q. 7/71 and passim) 'absurd; baseless'.
ii. The *Bible*. Although we may assume that Arabic has never borrowed any idiom from Biblical Hebrew directly, it is quite feasible that Christian Arabs have borrowed idioms from the Old and New Testament through the available translations of the Bible. Examples: *kabš al-fidāʾ* 'scapegoat' (Leviticus 16:10); *ṣayḥa fī wād* 'a cry in the wilderness' (Matthew 3:3, based on Isaiah 40:3).
iii. The *Ḥadiṯ*. Examples: *al-ġanīma al-bārida* 'easy prey' (based on a dictum attributed to the Prophet Muḥammad, who said *aṣ-ṣawm fī š-šitāʾ al-ġanīma al-bārida* 'fast in winter is easy'; *Lisān al-ʿArab* s.v. *ġ-n-m*).
iv. Poetry. Examples: *lā nāqa lahu wa-lā jamal* (*fī hāḏa l-ʾamr*) 'he has nothing to do with it' (part of a verse by Muʾayyid ad-Din aṭ-Ṭugrāʾī, 1061–1121).
v. Proverbs. Example: *jāʾa bi-qarnay ḥimār* 'to return empty-handed' (based on the proverb *ḏahaba l-ḥimār yaṭlubu qarnayn fa-ʿāda maṣlūm al-ʾuḏunayn* 'the donkey went to look for horns and returned with its ears cut off').
vi. Folklore. Example: *watad Juḥā* 'foothold' (based on an amusing folktale about the famous mythical person named Juḥa, who

was ‘emotionally’ attached to a peg in his old house).

vii. Historical events. Example: *baʿda xarāb al-Baṣra* lit. ‘after the destruction of Basra’, i.e. ‘after giving up hope’.

viii. Technical terms. Example: *hunā marbaṭ al-faras* lit. ‘this is the place where the mare is tied to’, i.e. ‘there lies the rub’.

ix. Colloquial Arabic. Example: *madda rijlahu bi-qadr kisāʾihi/madda rijlayhi bi-qadr liḥāfihi* ‘he stretched his feet to the size of his carpet/blanket’ (based on the proverbs *ʿala qadd bsāṭak midd ijrēk = mudd rijlaka ʿala qadr al-kisāʾ*).

x. Calque. Example: *ḏarra r-ramād fī l-ʿuyūn* ‘to throw dust in someone’s eyes’.

3. The morphology of idioms

Morphologically, Arabic idioms may be divided into six main categories:

i. One-word idioms which have become ordinary lexemes, usually as a result of ellipsis or amalgamation. These are divided into five categories:

   a. Aphetic; a single-word utterance originally constituting part of a phrase or sentence, which at some stage was shortened, e.g. *šaʾnaka* lit. ‘your matter’, i.e. ‘do as you like’, which was originally *iʿmal šaʾnaka bihi*.

   b. Substantival; when an adjective acquires the meaning of the substantive it qualifies and becomes an independent word, e.g. *ad-dunyā/al-ʾāxira* lit. ‘the near/the far’, instead of *al-ʾarḍ* (or *ad-dār*) *ad-dunyā/al-ʾāxira* ‘this world/the hereafter’; *al-muṣība* lit. ‘the hitting’, instead of *aḍ-ḍarba al-muṣība* lit. ‘the hitting blow’, i.e. ‘disaster’.

   c. The first part of an → *ʾiḍāfa* conveying the meaning of the combination, e.g. *ṭālib* ‘seeking’, short for *ṭālib al-ʿilm* lit. ‘seeker of knowledge’, i.e. ‘student’; *qaṭīʿa* ‘enmity among relatives’, short for *qaṭīʿat ar-raḥim* lit. ‘rupture of the womb’.

   d. → Compounds (*naḥt*); when two words amalgamate, yielding a new meaning, e.g. *talāšā* ‘to become nothing [*lā šayʾ*]; to be suppressed’; *rāʾsmāl* lit. ‘head of assets’, i.e. ‘capital’. Many compounds are the result of Turkish or Persian influence, e.g. *qāʾimmaqām* ‘commander’ (from *qāʾim* ‘standing’ and *maqām* ‘place, position’); *ʿarḍḥāl* ‘petition’.

   e. Idiomatic duals; one word in the dual instead of two words, yielding a new meaning, e.g. *al-ʾaḥmarāni* lit. ‘the two red ones’, i.e. ‘wine and meat’; *al-xāfiqāni* lit. ‘the two throbbing ones’, i.e. ‘East and West’.

ii. Phrasal or prepositional verbs. Widely common in English, in particular, there are several examples in Arabic too. They refer to verbs which, in combination with prepositions, yield a new meaning that cannot be deduced from the denotative meaning of the verb. The verb *ʾaxaḏa* ‘to take’, for instance, may be used with the following prepositions: *ʾaxaḏa bi-* ‘to grab, acquire, keep, observe, follow, initiate, admonish’; *ʾaxaḏa ʿalā* ‘to hold against, blame’; *ʾaxaḏa ʿan* ‘to learn, acquire knowledge’; *ʾaxaḏa fī* ‘to begin’; *ʾaxaḏa min* ‘to understand, infer, deduce’; *ʾaxaḏa bi-…ʾilā…* ‘to lead to’. The verb *ḥāla* may be used with the following prepositions: *ḥāla dūna* ‘to prevent’; *ḥāla bayna* ‘to come between’.

iii. Prepositional phrases. These are combinations consisting of prepositions followed by nouns, e.g. *bayna yadayhi* lit. ‘between his hands’, i.e. ‘in front of him’; *fī ʾaṯnāʾ* lit. ‘in the folds’, i.e. ‘during’.

iv. Phrasal particles. These are particles (prepositions, conjunctions, pronouns, etc.) which, when combined, yield a new meaning, e.g. *hāḏā ʾilā* ‘this in addition to…’; *ʿalā ʾan* ‘on condition that’.

v. Semi-idioms, i.e. combinations in which one component retains its denotative meaning, e.g. *rajaʿa bi-ṣifr al-yadayn* ‘to return empty handed’; *ṭaʿana fī s-sinn* ‘to be advanced in age’.

vi. Full idioms, i.e. combinations yielding a new meaning altogether, e.g. *ḍaraba ʾaxmās fī* or *li-ʾasdās* ‘to be at one’s wit’s end’; *šammaʿa l-fatla* ‘to slip away, abscond’.

4. The grammatical structure of idioms

Broadly speaking, idioms are structured as phrases (or part of clauses) or sentences.

i. Phrasal idioms may be structured mainly as:
  a. noun + adjective, e.g. *ʾāla ṣammāʾ* ‘willing tool’; *aṭ-ṭābūr al-xāmis* ‘fifth column’
  b. attribute/participle + substantive in annexation (*ʾiḍāfa*), e.g. *xafīf ad-dam* ‘charming, easygoing’; *ṭawīl al-bāl* ‘patient’; *ṭāhir aṯ-ṯiyāb* ‘with a flawless character’; *nāʿim aḏ̣-ḏ̣ufr* ‘young’; *maktūf al-ʾaydī* ‘helpless’
  c. verbal noun (→ *maṣdar*) + substantive in annexation, e.g. *raḥābat al-fanāʾ* ‘hospitality’; *jawalān al-yad* ‘embezzlement’
  d. nouns in annexation, e.g. *masqaṭ ar-raʾs* ‘birthplace’; *ḏāt al-bayn* ‘enmity’
  e. noun annexed to proper name, e.g. *mawāʿīd ʿUrqūb* ‘deceptive promises’; *jazāʾ Sinimmār* lit. ‘Sinimar’s award’, i.e. ‘to requite good with evil’
  f. elative + *min* + substantive/proper name, e.g. *ʾandar min al-ġurāb al-ʾaʿṣam* ‘rarer than a white-footed crow’; *ʾakḏab min Musaylima* ‘a greater liar than Musaylima’. These may be regarded as proverbs as well.
  g. Binomials
    1. With the same noun repeated: either without any addition, e.g. *baʿḍuhum baʿḍan* ‘each other’; or linked by a conjunction, e.g. *šayʾan fa-šayʾan* ‘bit by bit’; or linked by a preposition, e.g. *raʾs bi-raʾs* ‘alike’; or both linked by prepositions, e.g. *fī l-fatra baʿda l-fatra* ‘once in a while’;
    2. With different nouns or participles: either without any addition (*ʾitbāʿ*), e.g. *xarāb yabāb* ‘complete destruction’; or linked by a conjuction, e.g. *ʾaxḏ wa-radd* ‘discussion, debate’; or linked by a preposition, e.g. *raʾs ʿalā ʿaqib* ‘upside down’; or both linked by prepositions, e.g. *min ʾawwalihi ʾilā ʾāxirihi* ‘from A to Z’; or linked by a preposition and a conjunction, e.g. *bi-l-bāʿ wa-ḏ-ḏirāʿ* ‘with might and main’; or preceded by a negative particle and linked by a conjunction, e.g. *lā fatīl wa-lā naqīr* ‘nothing’; or preceded by a preposition and a negative particle and a conjunction followed by a negative particle, e.g. *bi-lā kitāb wa-lā sunna* ‘lawlessness’.

ii. Idiomatic sentences occur when an idiom constitutes an independent sentence or, more commonly, a clause in a compound or complex sentence. The main types of this category are:
  a. verb + subject, e.g. *ḏarra qarnuhu* ‘to begin to show’
  b. verb + direct object, e.g. *malaʾa l-ʿayn* ‘to satisfy completely’
  c. verb + indirect object, e.g. *ixtalaṭa l-ḥābil bi-n-nābil* ‘to be completely confused’
  d. verb + adverb, e.g. *tafarraqū ʾaydī sabā* ‘to scatter in all directions’
  e. verb + absolute accusative (*mafʿūl muṭlaq*), e.g. *ʾarsala l-kalām ʾirsālan* ‘to speak without restraint’
  f. verb + verb, e.g. *saraḥa wa-maraḥa* ‘to do as one likes’
  g. verb + verb + object, e.g. *ʾaqāma d-dunyā wa-ʾaqʿadahā* ‘to move heaven and earth’
  h. verb + object + object, e.g. *ḍaraba wajh al-ʾamr wa-ʿaynahu* ‘to touch on the very essence of a matter’
  i. verb + adverb + adverb, e.g. *ʿāda sāliman ġāniman* ‘to return safe and sound’
  j. nominal sentence, e.g. *ka-ʾanna ʿalā ruʾūsihim aṭ-ṭayr* ‘they are silent with awe’
  k. compound sentence, e.g. *yuqaddimu rijl wa-yuʾaxxiru ʾuxrā* ‘he hesitates’
  l. complex sentence, e.g. *yaʿlamu min ʾayna tuʾkalu l-katif* ‘to fall on one’s feet’

5. THE SEMANTICS OF IDIOMS

Semantically, Arabic idioms may be divided into five main categories:

i. Conceivable figurative idioms. The idioms which belong to this category get their intensity from the literal meaning of the expression which is behind the idiomatic meaning (‘metaphorization’). These may be subdivided into three main groups:
  a. Gestures, e.g. *basaṭa ḏirāʿayhi* lit. ‘to spread one’s arms’, i.e. ‘to welcome’
  b. Actions, e.g. *ʾaxaḏa ʿalā ʿātiqihi* lit. ‘to carry on the shoulder’, i.e. ‘to undertake a task’
  c. Hyperboles, e.g. *ḏābat ʾaḏ̣fāruhu* lit. ‘his fingernails melted’, i.e. ‘to strive in vain’

ii. Inconceivable figurative idioms. The idioms which belong to this category have no conceivable literal meaning, e.g. *raqqaṣa al-ḥanājir* lit. ‘it made throats dance’, i.e. ‘to provoke loud laughter’; *waqaʿat al-ʿayn ʿalā l-ʿayn* lit. ‘the eye fell over the eye’, i.e. ‘fighting broke out’.

iii. Allusive idioms. These are idioms which are part of the cultural bond of Arab society, alluding to religion, history, literature, folklore, and the like, e.g. *yaxluṭu bayna šaʿbān wa-ramaḍān* lit. ‘he confounds Šaʿbān with Ramaḍān’, i.e. ‘he does not know his right hand from his left’.

iv. Idioms containing obsolete words, which are usually used only in this idiom, e.g. *ʾatā/jāʾa šayʾan fariyyan* ‘to do something unheard of’; *ʾalqā l-xabar ʿalā rusaylātihi* ‘he did not take the matter seriously’.

v. Miscellaneous categories. These refer to special cases of idioms which may have intrinsic features or may belong to more than one category:

  a. Binomials (*ʾitbāʿ*) [see above for their morphological structure]. These may contain the same word repeated, e.g. *ʿāl al-ʿāl* ‘excellent’; the same root repeated, e.g. *ʾaḍʿāf muḍāʿafa* ‘manifold’; synonyms or cognates, e.g. *zift wa-quṭrān* ‘unpleasant, awkward’; two words with opposite meaning, e.g. *bayna ʿašiyya wa-ḍuḥāhā* ‘overnight, suddenly’; a word that is only added for rhyming purposes, e.g. *lā yaʿrifu l-ḥayy min al-layy* ‘he wouldn’t know a snake from a garden hose’.

  b. Common expressions. These include civilities, → greetings, wishes, swearing, exclamations, oaths, vocatives, and rhetorical and parenthetical expressions, e.g. *ʾahlan wa-sahlan* ‘welcome!’; *bayyaḍa llāhu wajhahu* ‘may God make him happy!’; *ʿalā ṭ-ṭāʾir al-maymūn* ‘bon voyage!’; *lā ʿalayka* ‘don’t worry!’; *yā salām* ‘good Lord!’; *wa-ḥayāt raʾsī* lit. ‘by the life of my head’, i.e. ‘I swear by my life’.

  c. Terminology. Various terms referring to different disciplines are also idiomatic, e.g. *lisān ḥāl* ‘organ [of a party, movement etc.]’; *radd fiʿl* ‘reaction’; *ʿalā ʾamwāj al-ʾaṯir* ‘on the radio’; *dāʾira maʿārif* ‘encyclopedia’.

6. Stylistic aspects of idioms

Generally, idioms contribute to the clarity and aesthetic of the utterance and therefore play a decisive role in shaping its style, often elevating it to a very high status of literary expression. Stylistically, the idiom may fulfill three main functions. In the first place, since no monolexemic equivalent exists, the idiom fills a lacuna in the language. The alternative would be a lengthy explanation, for instance for the expression *ʾadlā dalwahu bayna ad-dilāʾ* or *ʾadlā bi-dalwihi fī d-dilāʾ* ‘to make one’s contribution [together with others]; to add one’s touch’. Additionally, the idiom offers information that is not conveyed by the single lexeme, as in the expression *ʾalqā riḥālahu/ʾalqā ʿaṣā t-tirḥāl* ‘to stop [after a long journey]’, instead of the general verbs *waqafa/tawaqqafa* ‘to stop’. Finally, the main function of the idiom is to offer a more aesthetic and even more picturesque utterance than the common word, as in the expression *ṭalaʿa ʿalā bāb Allāh* lit. ‘to turn up before God’s door’, instead of *irtazaqa* ‘to earn one’s bread’.

Moreover, idioms are often used as figures of speech to intensify the rhetoric of the speaker or writer. The most common figurative idioms are:

i. Simile, e.g. *ka-l-māʾ al-jārī* lit. ‘like flowing water’, i.e. ‘fluently’; *wāḍiḥ wuḍūḥ aš-šams fī rābiʿat an-nahār* ‘clear as daylight’

ii. → Metaphor, e.g. *ʾaqāma wazn kabīr* ‘to attach great importance’; *dārat raḥā l-ḥarb* ‘the war broke out’

iii. Metonymy, e.g. *qaṣīr al-bāʿ* ‘powerless’; *laḥm wa-dam* ‘flesh and blood’

iv. Synecdoche, e.g. *ṭalaba yadahā* ‘to ask her hand in marriage’; *masqaṭ ar-raʾs* ‘birthplace’

v. Epithet, e.g. *rasūlu llāh* ‘the Prophet Muḥammad’; *an-nāṭiqūn bi-ḍ-ḍād* lit. ‘those who pronounce the → *ḍād*’, i.e. ‘the Arabs’

vi. → Euphemism, e.g. *intaqala ʾilā raḥmat Allāh* ‘he died’; *bayt al-ʾadab* ‘latrine’

vii. Hyperbole, e.g. *zawbaʿa fī finjān* ‘a storm in a teacup’; *hadaf tataqaṭṭaʿu dūnahu l-ʾaʿnāq* ‘an unattainable goal’

viii. Litotes, e.g. *lā baʾs bihi* ‘considerable; not bad’; *lā yuṣṭalā bi-nārihi* ‘invincible’

ix. Merismus, e.g. *ṣāba ʾam ʾaqlaʿa* ‘under all circumstances’; *al-qāṣī wa-d-dānī* ‘everybody’
x. Polyptoton, e.g. *nasaja nasjahu* ‘to imitate (someone)’; *kataba kitābahu* ‘to marry (someone)’

7. THE COHESION OF IDIOMS

By definition the idiom is made of a combination of words which together convey a new meaning. This suggests an ‘en bloc cluster’ characterized by semantic stability, which does not allow any morphological, syntactic, or semantic changes in the structure or the substituting of its components. Thus, the idiom *ʾaqāma d-dunyā wa-ʾaqʿadahā* ‘to move heaven and earth’ cannot be transposed to **rafaʿa l-ʿālam wa-ʾajlasahu*, etc. Nevertheless, a large number of idioms have undergone certain semantic changes, mainly by using synonyms, e.g. *kašafa n-niqāb/al-qināʿ, rafaʿa/ʾazāḥa s-sitār, ʾamāṭa l-liṯām*, all meaning ‘to reveal, disclose’; *zāda diġṯ ʿalā ʾibāla/zāda ṭ-ṭin ballan/zāda n-nār ḍirām* ‘to add fuel to the flames’.

Moreover, morphological changes which allow grammatical and syntactic variation may also occur, often affecting the meaning, e.g. *ʾinna l-buġāṯ bi-ʾarḍina yastansir* became *istansara l-buġāṯ* ‘the small bird became like an eagle’. The proverb *iḍrib al-ḥadīd mā dāma bārid* ‘strike while the iron is hot’ became an idiom: *ḍaraba fī ḥadīd bārid* ‘to take futile steps, beat the air’. The expression *hāḏā min ṯalāṯat al-mustaḥīlāt* ‘this is one of the three impossible things’ became *hāḏā rābiʿ al-mustaḥīlāt* ‘this is the fourth impossible thing’.

8. THE USE OF IDIOMS

Since idioms are more emotive than the equivalent single word and often carry more weight, it is not surprising that users of the language prefer to incorporate idioms into their written texts or in their speech for enriching and embellishing their style. Literature at large, the media, and Arab writers, essayists, orators, politicians, and intellectuals use scores of idioms, which have a great impact on the hearer or the reader. Thus, when Yusuf Idris describes the eagerness of one of his protagonists to visit Vienna, he prefers to tell the reader that Vienna was *ḍāllatahu l-manšūda* lit. ‘his sought-after lost sheep’, i.e. ‘the goal of his persistent search, the object of a long-cherished wish’. Such an idiom is certainly more expressive and emotional than the ordinary words for ‘target’ or ‘wish and desire’. Similarly, when describing the absent-mindedness of one of his protagonists, Najib Mahfuz tells us that *lā yaʿrifu raʾsahu min rijlayhi* lit. ‘he does not know his head from his legs’, i.e. ‘he does not know his right hand from his left’. Moreover, paraphrasing proverbs and idioms, a favorite pastime of writers, also carries an emotional impact, especially when the common phrase is in the background, as in *hal ʾaṣlaḥa l-ʿahd mā ʾafsadahu l-fāsidūn*? (*an-Nahār* 1.5. 1974, 2/1) ‘did time mend what the immoral people have corrupted?’. This is a paraphrase of the famous proverb *lā yuṣliḥu l-ʿaṭṭār mā ʾafsadahu d-dahr* ‘no perfumer can repair what time has destroyed’. Finally, daily expressions may carry more weight than a common word, as in the expression *bi-r-rifāʾ wa-l-banīn* ‘may you enjoy happiness and have children!’, which as a wish to a newlywed couple is more expressive than the general good wish *mabrūk* ‘be blessed!’.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCE

Ibn al-ʾAṯīr, *Maṯal* = Ḍiyāʾ ad-Dīn Naṣr Allāh Muḥammad Ibn al-ʾAṯīr, *al-Maṯal as-sāʾir fī ʾadab al-kātib wa-š-šāʿir*. 2 vols. Cairo, 1939.

SECONDARY SOURCES

Cowie, Antony P. (ed.). 1998. *Phraseology: Theory, analysis and application*. Oxford: Oxford University Press.
Ḥusām ad-Dīn, Karīm Zakī. 1985. *at-Taʿbīr al-iṣṭilāḥī*. Cairo: al-Maktaba al-ʾAnglū al-Miṣriyya.
Makkai, Adam. 1972. *Idiom structure in English*. The Hague: Mouton.
Moon, Rosamund. 1998. *Fixed expressions and idioms in English: A corpus-based approach*. Oxford: Clarendon Press.
Qāsimī, ʿAlī al-. 1979. “at-Taʿābīr al-iṣṭilāḥiyya wa-ṣ-ṣiyāġiyya wa-muʿjam ʿarabī lahā”. *al-Lisān al-ʿArabī* 17:1.17–34.
Shivtiel, Avihai. 1976. *Idioms in Literary Arabic, their structure, meaning and usage*. Ph.D. diss., University of Cambridge.
——. 2001. “On the origin of some current Arabic idioms”. *Linguistic and cultural studies on Arabic and Hebrew*, ed. Judith Rosenhouse and Ami Elad-Bouskila, 175–191. Wiesbaden: O. Harrassowitz.
Smith, Logan Pearsall. 1947. *Words and idioms*. London: Constable.
Ušakov, V.D. 1963. “Fixed word-groups in contemporary Literary Arabic”. *Kratkije Soobščenija Instituta Narodov Azii* (Moscow) 72.37–47. [In Russian.]

——. 1970. “Idiomatics as an inherent feature of language at various levels (in Arabic and other languages)”. *Narody Azii i Afriki* (Moscow) 3.120–127. [In Russian.]

Wehr, Hans. 1979. *A dictionary of Modern Written Arabic*. 4th ed. Wiesbaden: O. Harrassowitz.

Avihai Shivtiel (Leeds)

Pidgin Arabic → Ki-Nubi; Juba Arabic

Pidgin Arabic: Bongor Arabic

1. GENERAL

1.1 *Area*

Bongor Arabic, locally called *árab aná bóngor*, less often *túrku* or *túrgu*, is a pidginized form of Arabic that is spoken in southwestern Chad in the Mayo-Kebbi area, more specifically in Bongor, a city which is close to the border with Cameroon.

Bongor Arabic is used as a lingua franca between the Masa (or Masana) and the Tupuri (two populations who speak a Chadic language and a Niger-Congo language respectively) on the one hand, and Arabic-speaking traders from the north on the other. There are no data concerning the number of speakers.

Information concerning the actual development of Bongor Arabic is contradictory: some observers consider that Muslim Masa and Tupuri families tend to teach their children Bongor Arabic in addition to or instead of their own language, while others say that Bongor Arabic is more and more influenced by Chadian dialectal Arabic, which competes with it (→ Chad). If the former assertion is correct, it may lead to the formation of a creolized form of Bongor Arabic (see Sec. 1.2).

1.2 *Bongor Arabic as a pidgin*

A pidgin language is usually seen as a speech form without native speakers (the main feature distinguishing a pidgin and a creole) which is therefore primarily used as a means of communication among people who do not share a common language (→ pidginization). Pidgin is a simplified form of language compared to the lexifier language, especially with respect to the morphology and the lexicon. Pidgins are born and used in specific social situations: interethnic contacts, contacts among traders, contacts between colonial people and local workers, etc. (Arends a.o. 1994).

The lexifier of Bongor Arabic is obviously Chadian Arabic, which is seen as a dialectal group, not as one specific variety of Arabic spoken in Chad, since there is more than one Arabic dialect in the country (Jullien de Pommerol 1997). Chadian Arabic itself belongs to Western Sudanese Arabic (Kaye 1976).

The main features that render Bongor Arabic a pidginized form of Chadian Arabic are the following:

i. The long vowels are replaced mainly by stressed vowels.
ii. Some phonemes tend to disappear (/x/ > /k/ in *kalí* ‘to let’) or to be modified (/š/ > /s/ in *mási* ‘to go’, /f/ > /p/ in *pi* ‘in, to’).
iii. Lack of gender distinction.
iv. Lack of derivational morphology.
v. The article *al-* is replaced by the demonstrative *dá* and *dólda*.
vi. The genitive is almost always analytic.
vii. The inflection of the verb is marked by the independent pronoun alone.
viii. The negation particle stands at the end of the sentence.

1.3 *Historical background*

A pidginized form of Arabic, called Turku (Tourkou), was already in use in southern Chad, and even in Nigeria and what is now Central African Republic, in the last decades of the 19th century. At the beginning of the 20th century, Derendiger (1912) noted that various forms of Arabic were spoken in Chad, including a pidginized form. Turku was mentioned by Lethem (1920), describing a form of Arabic spoken in Bornu, Nigeria. He attributed its origin to Rabeh, a Nubian trader who established himself in Chad in 1879. This language was later described by a French officer, Muraz (1930), in a booklet which was analyzed by Tosco and Owens (1993).

Quite apart from the name of the language itself and the area in which it was spoken, there are many parallels in the phonology, morphosyntax, and lexicon of the two languages. However, more research is needed to understand the evolution from the pidgin described by Muraz to today’s Bongor Arabic.

1.4 *Bongor Arabic in relation to other Arabic pidgins and creoles*

At least two other Arabic Pidgins and Creoles are spoken nowadays: → Juba Arabic, which is spoken in southern Sudan, and → Ki-Nubi, which is spoken mainly in Kenya and Uganda. Here we present the main features that Bongor Arabic on the one hand, and Juba Arabic and Ki-Nubi on the other, have in common, as well as differences between them:

Common features:

i. Many phonological processes are alike.
ii. There is no consonant gemination.
iii. Long vowels are replaced by stressed vowels.
iv. The 2nd person plural independent pronoun is built on the singular.
v. Nominal and verbal derivational morphology is very poor.
vi. There is no gender distinction.
vii. The definite article has disappeared and is replaced by demonstrative pronouns.
viii. Verbal inflection is marked by Tense-Mood-Aspect markers.
ix. The negation stands at the end of the sentence (Ugandan Ki-Nubi).
x. A large part of the lexicon is common, especially the verbal roots, mainly built on the imperative forms of the lexifier.

These common features may be divided into two categories: those features that can be explained by the fact that the lexifiers of Bongor Arabic and other Arabic pidgins and creoles belong to the same group, the Sudanese/Chadian Arabic dialects, and those features that cannot be explained by this common origin but are found in many other processes of pidginization or creolization, such as the lack of gender distinction and the use of Tempus-Modus-Aspect markers.

However, there are also differences between Bongor Arabic and Ki-Nubi and Juba Arabic. The main ones are:

i. Some specific phonological processes, probably linked to the substratum, like /f/ > /p/ and /t/ > /d/.
ii. Tempus-Modus-Aspect markers: Where Juba Arabic and Ki-Nubi use Ø, *gi-*, and *bi-* as main Tempus-Modus-Aspect markers, Bongor Arabic uses Ø and *gáy-*.
iii. The lexicon. Many lexical items/dialectical lexifiers are different: *aná* (BA), *bitá* (JA), *tá* (KN) 'of'; *orú* (BA), and *kélem* (KN, JA) 'to say'; *yátu* (BA), *munú* (KN, JA) 'who'; *almé* (BA), *móyo* (KN, JA) 'water'; *antú* (BA), *wedí* (KN, J) 'to give', *dár* (BA), and *béle* (KN, JA) 'country'.
iv. The adstrate. While Bongor Arabic lexicon is influenced by French, Juba Arabic is influenced by English and Arabic, and Ki-Nubi by Swahili and English.

2. Linguistic Description

2.1 *Phonology*

2.1.1 Consonants

		labials	dentals	palatals
plosives	voiceless	p	t	č
	voiced	b	d	j
fricatives	voiceless	f	s	š
	voiced		z	
nasals		m	n	ny
vibrants			r	
laterals			l	
semivowels	w			y

velars	glottals
k	h
g	
ng	

There is no consonant gemination. Some speakers tend to replace /f/ by /p/ and /t/ by /d/.

2.1.2 Vowels

i u
e o
a

These vowels may also be realized long, but the vowel length is not distinctive.

2.1.3 The syllable

The patterns of the syllable are v, vC, Cv, and CvC, as in *á-na* 'I', *al* 'that, who' (relative pronoun), *dá* 'this' and *dár* 'country'.

2.1.4 Accent

There is a tonic accent (which in most cases corresponds diachronically to the long vowel or the stressed vowel in the lexifier: *kalám* 'word', *dérib* 'way'), which may be distinctive for some words, like *ána* 'I' and *aná* 'of'.

2.2 *Morphology*

2.2.1 Pronouns

Bongor Arabic distinguishes between independent and suffix pronouns, although the pronouns of the 3rd person are similar:

	independent pronouns	suffix pronouns
3rd sg.	*hú*	*-hú*
2nd sg.	*ínti/índi*	*-k(i)*
1st sg.	*ána*	*-(y)i*
3rd sg.	*(h)úman*	*-úman*
2nd sg.	*índukum*	*-kú(m)*
1st sg.	*anína*	*-na*

The independent pronouns are used in four cases:

i. Subject of nominal sentences: *úman sára dóldá* 'they are Sara'
ii. Subject of verbal sentences: *ána orú lé-hu* 'I told him'
iii. Object of verbs: *ána kúlu dóru orí ínti* 'me, too, I want to tell you'
iv. After a preposition: *hú god ma ána fi bóngor* 'he stayed with me in Bongor'

They may also be used as possessors: *akú aná ána* 'my brother'.

The suffix pronouns are used after the possessive marker *aná* and the preposition *le: usm aná-y kaltúma* 'my name is Kaltuma', *anína gáy-só lé-ku sókol dá* 'we will do this for you'.

The relative pronoun has a unique form, *al*, and demonstrative pronouns are *dá* (sg.) and *dólda* (pl.).

2.2.2 Adverbs

ówul 'before, formerly', *ása, ásada* 'now', *alyóm* 'today', *ambákir* 'tomorrow', *lahása* 'up to now', *dá* 'so, that way', *kúlu* 'too', *séme* 'well', *šíya* 'few', *ketír* 'very', *híni* 'here', *hinák* 'there', *ké ké* 'so-and-so, here and there', *batán* 'again, still', *ma…batán* 'not anymore', *wén* 'where?', *kikép* 'how?', *lé* 'why ?'.

2.2.3 Prepositions

min 'from', *ál* 'in, inside'.

2.2.4 Conjunctions

u, o 'and', *aw* 'or', *walá* (interrogative), *lákin* 'but', *ašán* 'because'.

2.2.5 Nouns and adjectives

There is no productive marker of word classes, the distribution of the words being their main indication.

2.2.5.1 Gender

There is no morphological gender distinction: *uléd sakár* 'a young boy', *bináya sakár* 'a young girl'.

2.2.5.2 Number

There is no dual (the suffix *-én* is still present in the numerals, but it is not productive: *tinén* 'two', *mitén* 'two hundred'). The majority of nouns have no morphological plural, plurality being marked by the demonstrative *dólda*, or by quantifiers like *katír* 'many' or *kul* 'all': *mer dólda* 'the mayors', *sára dólda* 'the Sara', *kalám katír katír* '(many) questions'. However, some nouns may bear the suffix *-ín* as a plural marker: *malán, malan-ín* 'full'; *wáy, wad-ín* 'some'; *gáid, gaid-ín* 'staying'. Other plural markers may occur, but these are less common: internal plural, e.g. *rájil* 'man'/*rujál* 'men', *kebír/kubár* 'big'; and suppletive plural, e.g. *nádum* 'person'/*nas* 'people', *mára* 'woman'/ *awín* 'women'.

2.2.6 Numerals

Etymologically, all the numerals (*wáy, tinén, taláta, árba, kámsa*, etc.) are related to the Arabic lexicon. However, the unit always follows the ten – a feature which also appears in other varieties of Chadian Arabic: *ásara u kámsa* 'fifteen'. The speaker may switch to French to express high numbers: *trente mille, quarante mille* 'thirty thousand, forty thousand'.

2.2.7 The verb

2.2.7.1 Verbal form

All the verbs have a unique form, which is the verbal stem. There is no derivational morphology, except for some compound verbs with *só* 'to do', like *só sarági* 'to steal', *só mer* 'to become a mayor', *só mersí* 'to thank'.

2.2.7.2 Inflection

The perfect is marked by the use of the non-inflected verb stem: *hú mási* 'he went', while the imperfect is marked by the prefix *gáy-: índukum gáy-mási wén?* 'where are you [pl.] going?'. The marker *kalás* may reinforce the perfect: *kalás úman máso* 'they are gone'. All markers are in preverbal position.

2.2.7.2.1 Perfect

3rd sg.	*húm ási*
2nd sg.	*índi mási*
1st sg.	*ána mási*
3rd pl.	*úman mási*
2nd pl.	*índukum mási*
1st pl.	*anína mási*

2.2.7.2.2 Imperfect

3rd sg.	*hú gáy-mási*
2nd sg.	*índi gáy-mási*
1st sg.	*ána gáy-mási*
3rd pl.	*úman gáy-mási*
2nd pl.	*índukum gáy-mási*
1st pl.	*anína gáy-mási*

From the point of view of tense, the noninflected verb mainly renders the past, while *gáy-* may be used for the present as well as the future: *índi gáy-árfu bóngor máfi?* 'don't you know Bongor?', *ambákir dá ána gáy mási súk* 'tomorrow I'll go to the market', *sána al gáy-já dá ána gáy-mási france* 'next year I'll go to France'.

Stative verbs may be used without *gáy-*, even in the present: *ána dóru árge* 'I want [to drink] alcohol'; *be kalám árab dá índi árpu máfi?* 'you don't know it in Arabic?'.

2.2.7.2.3 The imperative

The imperative is rendered by the use of the verbal root alone: *orí* 'say', *asérbi* 'drink'. For some verbs, there is a suppletive form of imperative: *ámsi* 'to go' (*mási* 'to go'). However, the verb is often preceded by the 2nd person pronoun: *kalám dá índi orú ána* 'explain to me the problem', *sopér, índi ágip* 'driver, stop'.

2.2.7.2.4 The subjunctive

The verb introduced by an auxiliary is always noninflected: *almé aná-ku al gáy-já kasurú bét aná-na* 'your [pl.] water which comes to destroy our houses', *ána gáy-dorú orú kalám dá* 'I want to say this'.

2.3 *Syntax*

2.3.1 Noun phrase

2.3.1.1 Definiteness

The demonstrative pronouns are used as definite markers: *dérib dá* 'the road', *journalistes dólda* 'the journalists'. Indefiniteness may be marked by the omission of the definite marker, as in *hú gáy-adulú lé-na dérib* 'he will build a road for us', or by the use of the numeral *wáy* 'one', as in *mára wáy* 'a woman' or *ána gáy-orú kalám wáy* 'I will say something'.

2.3.1.2 Modifiers

The noun modifier – demonstrative/definite marker, adjective, numeral – always stands after the noun: *iyál dá* 'the kids' or 'these kids', *bináya sakár* 'a young girl', *rujál tinén* 'two men'.

2.3.1.3 Number agreement

Since many words have no specific plural marker, the agreement is not systematic: *bakán ishirín* 'twenty places'; *anína ma masás* 'we are no witches'. However, *rujál tinén* 'two men' is also found.

2.3.1.4 Genitive

The genitive marker is *aná*. It may be followed by a suffix pronoun: *hasáy aná-y* 'my stick', *kalám aná-ki* 'your speech', *bet aná-hu* 'his house'; or by a noun: *kalám aná depité dá* 'the deputy's speech', *nádum aná sar* 'a man from Sarh', *gúrs aná petról* 'the money of the oil'. The use of synthetic genitive may also occur, but it is very rare and often renders specific expressions or compounds: *kalám nasára* 'French', *kalám árab* 'Arabic', *kášam bár* 'riverbank' etc.

2.3.1.5 Comparative

There is no elative form. The comparative is marked by *fútu* after the adjective: *mamát dá tawíl fútu ali* 'Mamat is taller than Ali'.

2.3.1.6 Relative clauses

The relative clause is introduced by *al*: *hú kalám al fí gélb aná-k* 'this is the thing that you keep in your heart', *dá almé aná-ku al gáy-já kásuru bét aná-na* 'this is your [pl.] water which comes and destroys our houses'.

2.3.1.7 Reduplication

Reduplication may be used to emphasize the meaning of an adjective or an adverb: *nas kubár kubár* 'very important persons', *lámpa dólda al fóg fóg dá* 'the lamps which are above', *kalám katír katír* '(many) questions'.

2.3.2 The verb phrase

2.3.2.1 Phrase order

The phrase order is SVO: *ána súfu sókol dá* 'I saw this', *hú gáy-adulú bét aná-hú* 'he builds his house'. However, topicalization is quite common: *dérib dá bes anína dóru* 'we want this road'.

2.3.2.2 Indirect discourse

After the verbs *gál* and *orú, orí* 'to say', the particle *ké* is often used: *índukúm bas hása gal ké potról pí dár aná-kum* 'now you say that there is oil in your country'.

2.3.2.3 Existential and possessive sentences

An existential sentence is usually marked by the use of *pí* or *fí* (*mapí* or *mafí* in a negative sentence): *opitál fí fi bóngor* 'there is a hospital in Bongor'. But *fí* may also be omitted: *índukúm bas hása gal ké petról pí dár aná-kum* 'now you [pl.] say that there is oil in your country'.

The possessive sentence is rendered by various expressions:

i. Possessor + *gáy be* (lit. 'to be with') + Possessed, as in *ána gáy be wadír jedíd* 'I have a new car'.
ii. Possessor + *be* + Possessed, as in *ána be akú aná-ye tinén* 'I have two brothers'.
iii. Possessor + Possessed + Existential word, as in *íntida livre wáy bas anáki fí sé* 'you have a book'.

In the last two expressions, possession is emphasized by *aná* + possessive pronoun.

2.3.3 Negation

The marker *mafí* or *mapí* is used to negate both nominal and verbal sentences. It occurs usually in sentence-final position: *ána lúti mapí*, 'I am not stupid', *ána gáy-gáta mapí* 'I don't cut', *ána árifu mafí* 'I don't know'.

Another marker, *má*, is sometimes used. It is placed before the predicate or the verb. However, it is rare and may be an influence of Chadian Arabic: *anína má masás* 'we are no witches', *ána má dorú* 'I don't want'. *Má…batán* expresses 'not anymore': *ána má uléd batán* 'I am not a child anymore'.

2.3.4 Depidginization

Bongor Arabic speakers may be influenced by their knowledge of other forms of Arabic – mainly Chadian dialect(s) – which leads to a partial depidginization of their speech. For instance, in sentences like *anína rikíb-na wotír dá sáwa* 'we took the car together', *ána kúlu gáy-ni-dóra ni-orí séi* 'me too, I would like to say something', *ána ma n-árfa* 'I don't know', *índi gáy-árifu úsum-i* 'you know my name', personal prefixes or suffixes are added to the verb, the negation marker *má* stands before the verb, and the synthetic genitive replaces the analytic genitive.

3. LEXICON

The lexicon is clearly derived from the vocabulary of Chadian Arabic: *orú* 'to say', *aná* 'of', *dóru* 'to want', *yátu* 'who', *almé* 'water', *dár* 'land, country', etc. However, many loanwords come from French (*maire* 'mayor', *député* 'deputy', *infirmier* 'nurse'…), which is the official language of Chad, a phenomenon which can be observed in Chadian Arabic as well. Depending on the speaker, these loanwords may be used as such, or may sometimes be phonologically modified (*fantalón* for *pantalon* 'trouser', *boté* for *voter* 'to vote', *apermé* for *infirmier* 'nurse'). A few other loanwords come from local languages (*bursdiya* 'Monday', supposedly from Masa).

BIBLIOGRAPHICAL REFERENCES

Arends, Jacques a.o. (eds.). 1994. *Pidgins and creoles: An introduction.* Amsterdam and Philadelphia: J. Benjamins.

Derendiger, R. 1912. "Notes sur le dialecte arabe du Tchad". *Revue Africaine* 56.341–358.

Jullien de Pommerol, Patrice. 1997. *L'arabe tchadien: Emergence d'une langue véhiculaire.* Paris: Karthala.

Kaye, Alan S. 1976. *Chadian and Sudanese Arabic in the light of comparative Arabic dialectology.* The Hague: Mouton.

Lethem, G.J. 1920. *Colloquial Arabic, Shuwa Arabic dialect of Bornu, Nigeria and of the region of Lake Chad.* London: Crown Agent for the Colonies.

Muraz, Gaston. 1930. *Vocabulaire du patois arabe tchadien ou 'tourkou' et des dialectes Sara-Madjinngaye et Sara Mbaye.* Nancy: Charles-Lavauzelle.

Tosco, Mauro and Jonathan Owens. 1993. "Turku: A descriptive and comparative study." *Sprache und Geschichte in Afrika* 14.177–267.

Xavier Luffin (Free University of Brussels)

Pidginization

1. Definitions and generalities

Pidginization is, strictly speaking, simply the process whereby a pidgin is brought about. Its definition hinges, therefore, upon a previous delimitation of the concept of pidgin. According to Holm (2000:5), a pidgin is "a reduced language that results from extended contact between groups of people with no language in common". Generally, but not always, the groups in contact have unequal social (economic, political, etc.) status: the language of the group in power acts as the target (the superstrate) for the other groups. Because full acquisition of the target language is limited or even precluded, the nonnative speakers produce imperfect approximations of the target. Such approximate versions of the superstrate language may then be adopted even by the superstrate speakers in dealing with the other groups. The languages in contact must display robust differences at all levels (structural, lexical, etc.): contact among sufficiently close varieties may result in koineization or dialect leveling (→ dialect koine), rather than in pidginization. (Holm actually prefers to stipulate that the languages in contact are not closely related, which seems to put undue emphasis on genetic relationship.) Nobody speaks a pidgin as his or her first or only language; by definition, once a pidgin has acquired a community of native speakers, it is called a creole (→ creole Arabic). A pidgin is also different from a → jargon because the latter is an ad hoc, individual solution, with no fixed norms; a pidgin, by contrast, is more stable and has certain norms, although variation still looms large because of the transfer of features from the speakers' first languages. In time, variation may be reduced and pidgins may become a more or less istitutionalized means of intergroup communication. One often speaks in such cases of expanded, or enriched, pidgins.

Although there is no unanimous definition of pidginization, nor an accepted theory of pidgin genesis (see Kaye and Tosco 2001 for an overview), a multilingual setting and no common language available seem to be the minimum common denominator among competing hypotheses.

2. Pidginization and the Arabic dialects

Pidgins and pidginization are germane to Arabic linguistics in two domains: first, pidginization has been invoked as a historical process in the history of the Arabic language and, in particular, in the genesis of the modern Arabic dialects (which it would perhaps be better to call 'native Arabic', as proposed by Owens [2001a:426, n. 9]). Second, in the field of Arabic dialectology, pidginization and creolization are indispensable concepts in the analysis of → Juba Arabic, → Ki-Nubi, and other varieties, most of them spoken (either in the past or still presently) in Africa (see also → pidgin: Bongor Arabic).

As a hypothesis concerning the history of Arabic, pidginization is associated with the name of Versteegh and his influential *Pidginization and creolization: The case of Arabic* (Versteegh 1984). Versteegh's hypothesis accounts for the rise of the Arabic dialects out of the pidginization of Classical Arabic (or a variety thereof), followed by creolization and, finally, decreolization. Swift incorporation within the newly formed Islamic Empire brought about for the invading Arab armies and the conquered peoples alike major communicative problems: the necessity to communicate with the new Arabic-speaking power holders and with other segments of the population in the rapidly growing new urban centers, coupled with the absence of a ready-made common medium, forced the non-Arabophones to acquire, as best and as fast as they could, some knowledge of Arabic, albeit simplified and deviant from native speakers' speech.

The result was an Arabic-based pidgin strongly influenced by the native language of its speakers, which soon spread as a result of interethnic marriages, population movement, and the general situation of a social melting pot of the early Islamic times. Originally limited in scope, it gradually expanded and was enriched, and was readily adopted as the native language of a new generation of urban dwellers, thereby becoming, at least for many speakers, a creole. The Arabs themselves soon found it useful to have recourse to this simplified form of Arabic in their contacts with the indigenous populations (→ foreigner talk). At the same time, and as far as it remained in close contact with the Classical Arabic of culture, religion, and administration, this Arabic-based pidgin/creole underwent a steady process of decreolization and alignment to the norms of 'good' Arabic – a process which even now is still far from completed.

Versteegh's hypothesis breaks away from the dichotomy between monogenetic theories on the emergence of the Arabic dialects, such as Ferguson's (1959) hypothesis of a military koine spoken in early Islamic times, and polygenetic accounts, whereby the colloquial varieties sprang up independently from each other and assumed their common characters only out of later contact (see also Cohen 1970). In Versteegh's view, an identical process (pidginization) was set in motion whenever the 'right' historical and social conditions applied (and from this process the dialects assumed their common characteristics), but was locally conditioned by the different → substrate (the local language spoken at the time of contact). On the other hand, as remarked by Owens (2001a:424), Versteegh's view is consistent with Ferguson's (1959) as well as with Blau's (1966–1967) theories, insofar as the dialects are seen as innovative with respect to the Classical language, rather than the endpoint of a parallel line of development.

Much more than simply another hypothesis on the history of Arabic, the origin of the dialects, and the rise of diglossia in the Arab world, Versteegh's hypothesis amounts to a radical change: rather than a slow process of progressive differentiation of the dialects from the Classical language (or any other variety of 'Old Arabic'), an abrupt deviation from the model of the classical language would be followed by a slow realignment toward it. At the same time, instead of a slow but steady, inexorable process of differentiation among the dialects, a period of maximum differentiation would be followed by gradual coalescence, whereby the most aberrant and stigmatized solutions are inexorably dropped off.

Four pieces of evidence are advanced by Versteegh in support of his thesis. First is the testimony of the Medieval Arab grammarians, who constantly opposed the 'corrupted' language of the townspeople in favor of the more conservative and pure Arabic of the Bedouin. On the other hand, it is important to note that the 'mistakes' the Arab grammarians stigmatized do not bear direct evidence of pidginized Arabic.

Second is the data from peripheral Arabic dialects, such as those of Central Asia and Anatolia, largely isolated from the mainstream Arab culture and therefore less amenable to superstrate influence from the classical language. Again, as interesting and insightful as the evidence stemming from 'fringe Arabic' may be, it is also inconclusive: peripheral Arabic dialects are generally spoken by bilinguals in a non-Arabic milieu and thus show typical traits of mixed languages, but no overall simplification as found instead in pidgins and creoles (see Owens 2001b:353–354).

Third is the evidence of present-day Arabic-based pidgins and creoles, such as those of the Sudanese area and their offsping in East Africa, as well as that of pidgins and creoles in general, in order to show how the developments one finds in the Arabic dialects are typologically compatible with pidginization and creolization. Of course, typical, present-day or historically attested Arabic-based pidgins and creoles look profoundly different from any modern or historically attested Arabic dialect, but, in principle, there is no reason why a centuries-long, steady process of decreolization toward the sacred language and carrier of immense prestige which is Classical Arabic should not result in a drastic reshaping of pidginized varieties. After all, this is the picture shown by decreolized, 'high' varieties of European-language-based creoles in the Antilles and elsewhere.

Fourth is the historical data on the ethnic composition and demography of the early Islamic world.

Pidginization and/or creolization have been proposed for the development of many different languages and language groups, from the Romance languages and Middle English to Swahili, Turkish, and many more. In no case have they gained general acceptance. Nor, in general, has Versteegh's hypothesis for Arabic. A more positive attitude is taken by Kaye

(1985:201), who welcomes Versteegh's book as "a real beginning" – but more generally concentrates his review on the role of pidgins and creoles in Arabic linguistics in general.

In his review, Hopkins remarks that "urban, village and Bedouin dialects throughout the Arab world are typologically rather similar, and this similarity requires an explication" (1988:98). In particular, Versteegh (1984:79) recognizes that certain Bedouin dialects, and especially those spoken in the Arabian Peninsula, fall outside the scope of his theory and can be regarded as more or less normal developments of ancient dialects. Yet, they share with all the other dialects many typical, innovative traits of 'modern' Arabic. Moreover, "the general trend of development within Arabic is matched by similar features in Hebrew, Aramaic and Ethiopian" (Hopkins 1988:99). Therefore, far from being the reflex of any 'abnormal transmission' of Arabic (as implied in pidginization), the innovative features of the dialects are part and parcel of the general 'drift' of Semitic: given such and such a state of language as found in 'classical' Semitic languages, the argument goes, any development cannot but proceed along the general lines displayed by the Arabic dialects, whose peculiarity is therefore downplayed. Reference is made here to such features of the Arabic dialects as the dropping of nominal cases and the emergence of analytic structures, the most widely mentioned and discussed among them probably being the → analytic genitive. Both features should probably be reexamined, since, on the one hand, Owens (1998) has questioned the very nature and existence of the case system in Arabic, while, on the other, the role and frequency of the analytic genitive in the dialects has probably been unduly emphasized.

From the point of view of creolistics, Goodman (1986) has criticized Versteegh's definition of pidginization as too broad and generic. Much of the discussion actually turns on what is meant by pidginization. Is it sufficient, as Versteegh (1984:xi) does, to define it rather mildly as an "incomplete process of second language acquisition in an untutored context of language learning"? That such was the case for the acquisition of Arabic is obviously true. Not only was Arabization unplanned for by government agencies of the modern type, but one might even argue that it was at times indirectly opposed by the rulers, inasmuch as it was often a by-product of the process of Islamization, which brought about well-known budgetary problems: Islamization implied for the treasury the loss of the *jizya,* the head tax imposed on free male non-Muslims, and as such it was at times actively discouraged, as in Iraq by the governor al-Ḥajjāj ibn Yūsuf (d. 95/714). As a result, one might surmise, Arabization, too, was slowed down.

In general, untutored acquisition is simply the normal way to acquire a foreign language. As remarked by Kaye, "It is important to note that the question of *formal* instruction is not really germane because people rarely learn a language for communication purposes in a classroom [....] [A]t all stages of human history, foreign languages tend to be learned in a natural environment without any explicit foreign-language instruction" (1985:206; emphasis in the original). It does not happen nowadays, nor did it happen in the early days of the Islamic Empire. Furthermore, when one takes into consideration the general inefficiency of governmental agencies, one may indeed question whether any attempt at formally teaching a language variety may be very successful, by itself. The effective eradication of language diversity and dialect variation is much more the result of a complex network of social, economic, and ideological forces as can be found only in modern nation-states. The latter, as political constructs based upon territoriality and sovereignty of power, emerged only much later and in a totally different historical context (the Late Middle Ages in Western Europe).

Moreover, Owens (1989) has convincingly shown that the relationships between Arabs and indigenous populations in the Islamic Empire were in general far different from those which provided the input for the development of the 'classical' European-based pidgins and creoles: there was no rigid ethnic separation, nor was social mobility formally barred. Furthermore, a high degree of multilingualism, with no clear-cut dominance pattern, is the most typical source of pidgins, while the common situation in the early Islamic world was one which opposed a local majority and an exogenous, dominant Arab minority. It is doubtful that third parties played any major role. Thus, the Middle East, apart from a certain limited superstrate usage of Greek, was basically Aramaic-speaking; in

Egypt, Coptic was the native tongue; and in North Africa, Berber and pre-Romance varieties dominated the linguistic scene.

Pidgins are not at all rare, exceptional outcomes of language contact: they are instead rather common, but they are also ephemeral and liable to meet an untimely death when the conditions which favored their rise disappear, or when their speakers gain access to fuller (and more socially acceptable) versions of the target language. In a way, most pidgins 'depidginize' before having the possibility to creolize. This, for example, is what happens with many impoverished versions of Western European languages among immigrants. Even when pidgins pass the stage of individual solutions (jargons) and become somehow stabilized within a social milieu, they rarely expand and creolize. Jargons and pidgins are probably exceedingly common phenomena whose life span is generally very short. It is therefore no surprise that their usual fate is to rise, grow, and fade away unnoticed.

The arguments in favor of pidginization have been defended in Versteegh (1997:108–112) and, more recently, in Versteegh (2004). The major problem that any theory of the rise and development of the modern dialects ought to solve, according to Versteegh, is the presence of "a general trend that occurs in all Arabic dialects, and an individual translation of this trend in each area" (Versteegh 1997:108). As the substrate was different in each area, it cannot explain the similarities, nor can later convergence among the dialects be appealed to in order to explain the differences, because in this case one would expect the borrowing of actual morphemes from one dialect into another, rather than the borrowing of a structure filled in by different exponents.

Taking pidginization *stricto sensu*, its role in the history of Arabic is as yet to be ascertained. If one accepts instead a broader view of pidginization, it becomes the normal result of any language contact situation leading to abrupt language shift, and it is probably safe to admit with Bailey (1973:33) that *all* languages have creolized histories (Versteegh [1984], too, inclines towards this position). There can hardly be any doubt that "the modern dialects originated in a context of second language learning" (Versteegh 2004: 354), and this fact must be recognized and taken into account. In this case, pidginization remains a valid heuristic concept in historical linguistics, but it cannot be *proved* (see Tosco 1991). As Versteegh (2004: 345) writes, "What we can do is demonstrate that the scenario is a plausible one, that it may have taken place".

3. Pidginized varieties of Arabic

Simplified varieties of Arabic used as trade languages existed probably as long ago as 1000 C.E. Thomason and Elgibali (1986) have brought to the attention of linguists a short folktale from the 11th century found in ʾAbū ʿUbayd al-Bakrī's geographical work *al-Masālik wa-l-mamālik*, which allegedly portrays a pidginized 'Maridi Arabic'. Thomason and Elgibali (1986) locate Maridi in present-day Mauritania; on the other hand, Owens (1989:97–101), remarking that the text, according to al-Bakrī himself, was provided by a trader from Aswan and that it contains typical Upper Egyptian traits, has put forward the opinion that the text does not represent a 'true' pidgin, but "eine Karikatur einer pidginisierten Sprache, die in Ägypten benutzt wurde" (Owens 1989:100). Whatever the true status of Maridi Arabic, it represents by far the very first pidgin in absolute terms known to us from historical records.

It is again in Africa, and more specifically in the Sudan (in the geographical sense of the word), that we find the 19th-century and contemporary Arabic-based pidgins and creoles (see Owens 1997 for an overview). The southern Sudan has been the cradle of the two most widely known varieties: an expanded pidgin and creole used across most of the Equatoria province of the southern Sudan, and generally known as → 'Juba Arabic' (from the name of the provincial capital), and Nubi (or → Ki-Nubi), a creole spoken by a Muslim and urban minority in East Africa (Uganda, Kenya, Tanzania). Watson and Ola (1985) is a practical introduction to Juba Arabic; Heine (1982), and especially Luffin (2005a) and Wellens (2005) are good descriptions of Ki-Nubi. Ki-Nubi texts are found in Luffin (2005b). The complex and fascinating history of Arabic and its offsprings in the southern Sudan has been dealt with by Mahmud (1983) and Owens (1985).

Another variety, known to us only through the work of the French officer G. Muraz (1932), and analyzed linguistically by Tosco and Owens (1993), is Turku (French *tourkou*), a 'military' pidgin used around the turn of the 20th century among African troops of the French army in the Chad area. Many other Arabic-based pidgins and creoles certainly existed in the past. The status of other Arabic varieties of central and southern Sudan is equally unclear.

Typical phonological correlates of the pidginization of Arabic are the dropping of 'emphatic' (pharyngealized) segments, pharyngeals, and velar and uvular fricatives. Interdentals, another common target of inventory simplification (as witnessed by English-based pidgins and creoles), could not be an input to Arabic pidginization as they had already shifted to dentals in the Arabic dialects serving as input to the pidgins and creoles (and were, if one accepts Versteegh's [1984] hypothesis, the result of a 'first' pidginization). Phonemic vowel length is likewise lost. On the other hand, stress is distinctive. Out of 28 consonants phonemes in Classical Arabic and 25 in Sudanese (Khartoum) Arabic, the inventory of Juba Arabic is starkly reduced to 17 or 19 elements, lying therefore even below the modal number of consonants (21) in Maddieson's (1984) worldwide database (cf. Kaye and Tosco 2001:80).

Morphology is reduced to a minimum, although it is safe to say that it never disappears completely, except perhaps in basilectal varieties. The verbal system (cf. Owens 1980; Miller 1984; Tosco 1995) shows typical 'creole' traits in the opposition between stative and nonstative verbs, and in the use of preverbal markers in order to express tense, mood, and aspect distinctions. The basic preverbal markers seem to be *bi-*, which encodes the future and the conditional and may be interpreted as a marker of the irrealis mood; *ge-* as a marker of the nonpunctual, continuous aspect; and *kan-* as a marker of anteriority. Probably as a consequence of the ongoing superstratal influence of Sudanese and Standard Arabic on Juba Arabic, though not on Ki-Nubi, the verbal systems of the two are growing apart. As in almost all pidgins and creoles, the basic word order is SVO.

The lexicon is overwhelmingly Arabic in origin and contains many analytic formations; modern Ki-Nubi has of course borrowed extensively from Swahili and also English; code-switching with one or both of these languages is widespread (X. Luffin, personal communication).

Substratal influence reveals above all a robust Bari (East Nilotic, Nilo-Saharan) component, Bari being the major language spoken around Juba.

The respective roles of the substrate languages, the superstrate (i.e. Arabic), and universal tendencies and 'laws' are, as always in creolistics, a matter of considerable debate. The role of the substrate has been emphasized by Boretzky (1988), Vincent (1986), and Miller (1989), while Tosco (1995) has favored a 'universalist' approach and Owens (1980, 1985, 1991, 2001b) has stressed the key role of the superstrate.

Finally, Arabistics and general linguistics will have to reckon with an increasing number of pidginized varieties of Arabic in the fast-developing economies of certain Arab countries. Smart (1990) is a preliminary account based upon the speech of immigrant labor in the Gulf. Again, and unsurprisingly, pidginization has resulted in simplification of the consonant inventory, drastic reduction and regularization of the morphology, etc.

Pidginized and creolized varieties of Arabic is a branch of studies which so far has received surprisingly little attention; their sad neglect has often been remarked on, most notably by Owens (see esp. 2001a:458, n. 44, 2001b), who calls creole Arabic "the orphan of all orphans". On the other hand, it is probably a truism to remember that creolistics, too, is still focused primarily on European-language-based varieties. It is on the basis of data from European-based pidgins and creoles that general hypotheses on pidginization and creolization have so far been formulated, tested, discussed, and, often, rejected. The study of Arabic-based jargons, pidgins, and creoles could therefore provide creolistics and general linguistics alike with a less-Eurocentric testing ground against which to measure at last their alleged universality.

Bibliographical references

Bailey, Charles-James N. 1973. *Variation and linguistic theory*. Arlington, Va.: Center for Applied Linguistics.

Blau, Joshua. 1966–1967. *A grammar of Christian Arabic based mainly on South Palestinian texts from the first millennium*. 3 vols. Louvain: Corpus Scriptorum Christianorum Orientalium.

Bureng, George Vincent. 1986. "Juba Arabic from a Bari perspective". *Current approaches to African linguistics,* III, ed. Gerrit J. Dimmendaal, 71–78. Dordrecht: Foris.

Cohen, David. 1970. "Koinè, langues communes et dialectes arabes". *Études de linguistique sémitique et arabe,* ed. David Cohen, 105–125. The Hague: Mouton.

Ferguson, Charles A. 1959. "The Arabic koine". *Language* 25.616–630.

Goodman, Morris. 1986. Review of Versteegh (1984). *Journal of Pidgin and Creole Languages* 1.165–170.

Heine, Bernd. 1982. *The Nubi language of Kibera: An Arabic creole.* Berlin: D. Reimer.

Holm, John. 2000. *An introduction to pidgin and creoles.* Cambridge: Cambridge University Press.

Hopkins, Simon. 1988. Review of Versteegh (1984). *Zeitschrift für Arabische Linguistik* 18.98–99.

Kaye, Alan S. 1985. "On the importance of pidgins and creoles for historical linguistics". *Diachronica* 2.201–230. [Review of Versteegh (1984)].

——. and Mauro Tosco. 2001. *Pidgins and creoles: A basic introduction.* Munich: Lincom Europa.

Luffin, Xavier. 2005a. *Un créole arabe: Le Kinubi de Mombasa, Kenya.* Munich: Lincom Europa.

——. 2005b. *Kinubi texts.* Munich: Lincom Europa.

Maddieson, Ian. 1984. *Patterns of sounds.* Cambridge: Cambridge University Press.

Mahmud, Ushari A. 1983. *Arabic in the southern Sudan: History and spread of a pidgin-creole.* Khartoum: Khartoum University Press.

Miller, Catherine. 1979–1984. "Aperçu du système verbal en Juba Arabic". *Comptes Rendus du GLECS* 24–28.295–315.

——. 1989. "Bari interference in Juba Arabic". Ms. CNRS-IREMAM (Institut de Recherches et d'Études sur le Monde Arabe et Musulman). Aix-en-Provence.

Muraz, Gaston. 1932. *Vocabulaire du patois arabe tchadien ou "tourkou" et des dialectes sara-madjinngayé et sara-m'bayé (S.O. du Tchad).* Paris: Ch. Lavauzelle.

Owens, Jonathan. 1980. "Monogenesis, the universal and the particular in creole studies". *Anthropological Linguistics* 22.97–117.

——. 1985. "The origins of East African Nubi". *Anthropological Linguistics* 27.229–271.

——. 1989. "Zur Pidginisierung und Kreolisierung im Arabischen". *Afrika und Übersee* 72.91–107.

——. 1990. "East African Nubi: Bioprogram vs. inheritance". *Diachronica* 7.217–250.

——. 1991. "Local and universal sources of a creole verbal construction". *Semitic studies in honor of Wolf Leslau,* II, ed. Alan S. Kaye, 1169–1179. Wiesbaden: O. Harrassowitz.

——. 1997. "Arabic-based pidgins and creoles". Thomason (1997:125–172).

——. 1998. "Case and Proto-Arabic". *Bulletin of the School of Oriental and African Studies* 61.51–73, 215–227.

——. 2001a. "Arabic sociolinguistics". *Arabica* 48.419–469.

——. 2001b. "Creole Arabic: The orphan of all orphans". *Anthropological Linguistics* 43.348–378.

Smart, J.R. 1990. "Pidginization in Gulf Arabic: A first report". *Anthropological Linguistics* 32.83–119.

Thomason, Sarah G. (ed.). 1997. *Contact languages: A wider perspective.* Amsterdam and Philadelphia: J. Benjamins.

——, and Alaa Elgibali. 1986. "Before the lingua franca: Pidginized Arabic in the eleventh century A.D.". *Lingua* 68.317–349.

Tosco, Mauro. 1991. Review of Versteegh (1984). *Language Problems and Language Planning* 15.99–104.

——. 1995. "A pidgin verbal system: The case of Juba Arabic." *Anthropological Linguistics* 37.423–459.

——, and Jonathan Owens. 1993. "Turku: A descriptive and comparative study". *Sprache und Geschichte in Afrika* 14.177–267.

Versteegh, Kees. 1984. *Pidginization and creolization: The case of Arabic.* Amsterdam and Philadelphia: J. Benjamins.

——. 1997. *The Arabic language.* Edinburgh: Edinburgh University Press.

——. 2004. "Pidginization and creolization revisited: The case of Arabic". *Approaches to Arabic dialects: A collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday,* ed. Martine Haak, Rudolf E. de Jong, and Kees Versteegh, 343–357. Leiden: E.J. Brill.

Watson, Richard L. and Louis Biajo Ola. 1985. *Juba Arabic for beginners.* Nairobi: Summer Institute of Linguistics.

Wellens, Ineke. 2005. *The Nubi language of Uganda: An Arabic creole in Africa.* Leiden: E.J. Brill.

Mauro Tosco (University of Naples)

Place names → Toponyms

Plural → Number

Poetic Koine

The term 'poetic koine' (also 'poetico-Qurʾānic koine') refers to a supratribal variety of Arabic which, according to some scholars, was the variety of Arabic used in pre-Islamic poetry. The linguistic situation in the pre-Islamic period is a controversial topic (→ history of Arabic). Opinions about this situation may be divided into two main theories. According to one theory, which was also that of the Arabic grammarians, the language of the Arab tribes in the pre-Islamic period was basically homogeneous: although there were minor differences between the dialects of the Arab tribes (→ pre-Islamic Arabic), the colloquial speech of these tribes was identical with the language used in poetry, and this was also the language of the *Qurʾān*. This view is followed by some modern scholars, who believe that in the pre-Islamic period there was no → diglossia of the kind that is found after the Arab conquests (Blau 1977). One of the arguments to back up this view is that the Prophet Muḥammad would not have been likely to speak a language that differed substantially from the language of the Revelation (an argument already advanced by Nöldeke 1904).

According to the second theory, the linguistic situation before the coming of Islam was already characterized by diglossia, because the colloquial speech of the tribes differed from the language of poetry and the *Qurʾān*. In this theory, the term 'poetic koine' is the usual term to indicate the high variety used in poetry. Because this variety is believed to be identical with the language of the *Qurʾān*, those who adhere to this model also speak of the 'poetico-Qurʾānic koine'. The term itself is not very appropriate because the notion of 'koine' usually refers to a spoken variety of speech (→ koineization). It was originally used for the variety of Greek that emerged in the regions conquered by Alexander the Great in the 4th century B.C.E. (cf. Bubeník 1993). In the case of pre-Islamic Arabic, the term is used on the contrary for a variety that was restricted to poetic usage.

The term 'poetic koine' seems to have been applied to Arabic for the first time by William Marçais (cf. Rabin 1951:17; Zwettler 1978:173 n. 11), when he compared the language of pre-Islamic poetry to the artificial idiom of the Homeric epic poems, but even before him, Brockelmann (1908:24) had spoken of a *Dichtersprache*, a poetic idiom in pre-Islamic Arabia that differed from everyday speech. The notion of a special poetic language has become associated chiefly with a book by Michael Zwettler that appeared in 1978 under the title *The oral tradition of classical Arabic poetry*. The main thesis of his book is that the language used by the pre-Islamic poets was an archaic language that was identical with the language of the *Qurʾān*.

The main characteristic of the language used by the poets (→ *šiʿr*) was its traditional character. Zwettler believes that part of the traditional knowledge of the poets and their *rāwī*s was the knowledge of the rules of declension (→ *ʾiʿrāb*), which had disappeared from colloquial speech but were preserved in the language of the poems (Zwettler 1978:109). The poet was the *šāʿir* 'he who knows', and during their apprenticeship, poets and their *rāwī*s learned how to handle the intricate rules of a language that was no longer used colloquially, in particular the declensional endings (1978:85–88). The poets were "traditional masters of language" (1978:128), and in this capacity were responsible for preserving in their poems the traces of an archaic language. According to Zwettler, these traces are found, for instance, in the special → pausal forms in poetry and in those linguistic phenomena in pre-Islamic poetry that later generations had to interpret as → poetic licenses.

The many variant forms in pre-Islamic poetry are further evidence of the supradialectal character of the poetic language, because they are regarded as borrowings from different dialects. An example is the variation in the demonstrative pronouns. According to the grammarians (Fleisch 1979:29–60), the forms of the demonstrative differed among the tribes. The Banū Tamīm, for instance, used forms like *(hā-)ḏā* for proximate deixis and *(hā-)ḏāka* for distant deixis, whereas the Ḥijāzī dialect had *hāḏā* and *ḏālika*. Both sets occur in poetry, as do many other variant forms (Rabin 1951:120, 152–154). Other examples are the many variants of broken plurals, and the large number of synonyms in the poetic lexicon, which could be regarded as the result of borrowing from different dialects (Zwettler 1978:111). Although the poetic koine is believed to have borrowed from all dialects, Rabin (1951:1, 96; for comments see Zwettler 1978:115) asserts that it bore a special resemblance to one group of dialects, the East Arabian dialects, which were more conservative than the Ḥijāzī dialect. At the time the poems were composed, the poetic koine had ceased to be any tribe's colloquial language, although Zwettler (1978:119) believes that in some layers of society in the Ḥijāz, the use of the poetic language for other purposes than literary production may have become a sociolinguistic marker.

Zwettler's ideas are inspired by the notion of → orality that has become associated with the names of Milman Parry (1930, 1931) and Albert Lord (1968). They studied the language of the Homeric epic poems and concluded that the Ancient Greek tradition of rhapsodes was based on an oral tradition. Lord (1965) compared the tradition of the Homeric rhapsodes with that of the Yugoslav *guslar* in modern times. According to Zwettler, the role of the *rāwī*s, the transmitters of the pre-Islamic poets, was similar to that of transmitters of poetry in other oral traditions. In his view (1978:147), the oral character of Arabic pre-Islamic poetry, just like that of the Homeric poems, is demonstrated by the presence of formulaic schemas. The recurrent use of these formulas and the presence of archaic language are easily explicable as the result of how the poets worked. They used a stock of formulas, which could not be replaced, not even when the colloquial language had changed, because the metrical requirements did not allow that kind of variation.

A second property of oral poetry highlighted by Parry's analysis of the Homeric language is the near absence of enjambment. In pre-Islamic poetry, enjambment is indeed found to a much lesser extent than in later Arabic poetry (Zwettler 1978:64–77): verses in pre-Islamic poetry rarely draw on syntactically to the next line. For an oral transmitter of poetry, verses that are self-contained units are much easier to exchange. Later classical Arabic literary critics regarded enjambment (*taḍmīn*) as a fault (cf. van Gelder 1982; Sanni 1989), probably because it occurred very infrequently in the canon of pre-Islamic poetry.

The third property of oral poetry is brought out by thematic analysis (Zwettler 1978:77–85). In Parry's and Lord's analysis of the Homeric poems, the use of recurrent themes and motifs is another effect of the orality of these poems. Building on Jacobi's (1971) analysis of the thematic unity of pre-Islamic and early Islamic *qaṣīda*s, Zwettler suggests that the recurrence of these themes and motifs is consistent with the oral character of pre-Islamic poetry. It was not until much later that poets invented new themes in a tradition that was no longer oral.

As an example, the famous opening lines of Imruʾ al-Qays' *Muʿallaqa* may be quoted here:

qifā nabki min ḏikrā ḥabībin wa-manzilī
  bi-siqti l-liwā bayna d-Daxūli fa-Ḥawmalī
fa-Tūḍiḥa fa-l-Miqrāti lam yaʿfu rasmu-hā
  li-mā nasajat-hā min janūbin wa-šamʾalī

'Let us stop, oh friends, to weep in memory of a loved one and a dwelling place
  on the edge of the sandy desert between Dakhul and Hawmal
and Tudih and al-Miqra, whose traces have not been obliterated
  by [the pattern] woven by the winds from the south and north'

Here, Zwettler (1978:236–237) identifies as traditional formulas and themes the addressing of two companions, the weeping over places deserted by the beloved, the joining together of place names with *fa-*, and the recurrent rainstorms. Zwettler (1978:51–53) also points out that there are many parallel verses, which he calls 'formulaic elements', which have a common syntactic schema. One line in Imruʾ al-Qays' *qaṣīda*, *fa-ḏ̣alla l-ʿaḏārā yartamīna bi-laḥmi-hā*, is paralleled by verses with similar syntactic structure in ʿAlqama's poems (*fa-ḏ̣alla l-ʾakuffu yaxtalifna bi-ḥāniḏin*), in those of Ṭarafa (*fa-ḏ̣alla l-ʾimāʾu yamtalilna ḥuwāra-hā*), and others. Other examples include expressions with the *wāw rubba*: *wa-farʿin yazīnu l-matna* (Imruʾ al-Qays), *wa-baytin yafūḥu l-misku* (Imruʾ al-Qays), *wa-ġayṯin marat-hu r-rīḥu* (Imruʾ al-Qays), *wa-ṣadrin ʾarāḥa l-laylu* (Nābiġa); or with *bayna*: *bayna ḏirʿin wa-mijwalī* and *bayna ṯawrin wa-najʿalī* (both in Imruʾ al-Qays), *bayna dāmin wa-jālibī* (Nābiġa), *bayna xabtin wa-ʿarʿarī* (Labīd), etc.

Archaic language is the hallmark of the pre-Islamic poetic tradition, in which the oral features go hand in hand with formulaicity. This is what constitutes the poetic koine, not only in the pre-Islamic period but also in modern times. The term 'poetic koine' continues to be used for the language of traditional Bedouin poets in the Arabian Peninsula (Rabin 1951:17), such as Dindān, whose poems were recorded by Kurpershoek (1994). The language of this poetry is highly archaic and has the same properties as the oral poetry of pre-Islamic times. These traits are found in the poetic jargon of Bedouin tribes that have spread over a large region, from the Najd to the Sinai and the Negev (→ Northwest Arabian). Zwettler (1978:134) connects this modern tradition with the stories about the Bedouin in the early centuries of Islam; according to the grammarians, these Bedouin continued to speak the Classical → *ʿarabiyya*. Zwettler interprets this to mean that they preserved a tradition of poetry in an archaic language, including declensional endings, as it had been practiced by the pre-Islamic poets.

The concept of the pre-Islamic 'poetic koine' should not be confused with the theory of a military koine, connected with Ferguson's (1959) idea of a monogenetic origin of the modern Arabic dialects. In his view, after the Arab conquests the pre-Islamic dialects became koineized in the garrisons of the Arab armies, and it is from this so-called military koine that the modern dialects originate. The positing of the existence of such a koine is motivated by the presence of common traits in the modern dialects, which Ferguson feels can only be explained by positing the existence of a koine (→ history of Arabic).

BIBLIOGRAPHICAL REFERENCES

Blau, Joshua. 1977. *The beginnings of Arabic diglossia: A study of the origins of Neo-Arabic.* Malibu: Undena.

Brockelmann, Carl. 1908. *Grundriss der vergleichenden Grammatik der semitischen Sprachen*, I. Berlin. (Repr., Hildesheim: G. Olms, 1966.)

Bubeník, Vít. 1993. "Dialect contact and koineization: The case of Hellenistic Greek". *Koines and koineization*, ed. Jeff Siegel, 9–23. (= *International Journal of the Sociology of Language*, 99.) Berlin and New York: Mouton de Gruyter.

Ferguson, Charles A. 1959. "The Arabic koine". *Language* 25.616–630.

Fleisch, Henri. 1979. *Traité de philologie arabe.* II. *Pronoms, morphologie verbale, particules.* Beirut: Dar el-Machreq.

Gelder, Geert Jan van. 1982. "Breaking rules for fun/making lines that run/on: On enjambment in classical Arabic poetry". *The challenge of the Middle East: Middle Eastern studies at the University of Amsterdam*, ed. Ibrahim A. El-Sheikh, Aart van de Koppel, and Rudolph Peters, 25–31. Amsterdam: Institute for Modern Near Eastern Studies.

Jacobi, Renate. 1971. *Studien zur Poetik der altarabischen Qaṣīde.* Wiesbaden: F. Steiner.

Kurpershoek, P.M. 1994. *Oral poetry and narratives from Central Arabia.* I. *The poetry of ad-Dindān: A Bedouin bard in southern Najd.* Leiden: E.J. Brill.

Lord, Albert B. 1965. *The singer of tales.* Amsterdam: Atheneum Press.

——. 1968. "Homer as oral poet". *Harvard Studies in Classical Philology* 77.1–46.

Nöldeke, Theodor. 1904. "Das klassische Arabisch und die arabischen Dialekte". Theodor Nöldeke, *Beiträge zur semitischen Sprachwissenschaft*, 1–14. Strasbourg: K.J. Trübner.

Parry, Milman. 1930. "Studies in the epic technique of oral verse-making. I. Homer and Homeric style". *Harvard Studies in Classical Philology* 41.73–146.

——. 1931. "Studies in the epic technique of oral verse-making. II. The Homeric language as the language of an oral poetry". *Harvard Studies in Classical Philology* 42.1–50.

Rabin, Chaim. 1951. *Ancient West-Arabian.* London: Taylor's Foreign Press.

Sanni, Amidu. 1989. "On *taḍmīn* (enjambment) and structural coherence in classical Arabic poetry". *Bulletin of the School of Oriental and African Studies* 52.463–466.

Zwettler, Michael. 1978. *The oral tradition of classical Arabic poetry: Its character and implications.* Columbus: Ohio State University Press.

KEES VERSTEEGH (University of Nijmegen)

Poetic License

In a general sense, 'poetic license' is the freedom customarily given to poets to deviate from the normal rules of grammar, diction, or subject matter that are valid for prose, or even to depart from commonly accepted historical or scientific truth. Here, poetic license will be restricted to violations of the linguistic rules in the fields of morphology or syntax; other liberties of poets, such as being able to use far-fetched metaphors and imagery, to contradict themselves, to declare their love for women not their own or for boys, to celebrate the joys of drinking wine, to use anachronisms which in ordinary prose would be out of place, etc., belong to the domain of literary criticism and are not be discussed in this entry.

Whereas the English term implies liberty, the corresponding Arabic term most commonly used, *ḍarūra* 'necessity' (pl. *ḍarūrāt* or *ḍarāʾir*), stresses the lack of it, constrained as poets are by the exigencies of prosody: → meter and → rhyme; but the term *ruxṣa* (pl. *ruxaṣ*) 'license' is also found (e.g. ʿAskarī, *Ṣināʿatayn* 156; Ibn Rašīq, *ʿUmda* I, 269–280). It has been argued (Baġdādī, *Xizāna* I, 31–34) that the term *ḍarūra* is in fact a misnomer, since one way or another it should always be possible to avoid violating the rules. But if a poet is unwilling or unable to compromise between the rules of prosody, which are very strict in Arabic, and what he or she wants to say, either the prosody or the grammar has to give way. In rare cases, the rules of meter or rhyme are broken. The former may be seen in a line of the most famous Arabic poem of all times, the *Muʿallaqa* by the pre-Islamic poet Imraʾ al-Qays: *ʾAlā rubba yawmin laka minhunna ṣāliḥin*, where, extremely exceptionally, the seventh syllable is short instead of long. The latter is found, for instance, in a poem by a certain Jawwās ibn Huraym, who rhymes *ṣuduʿ* with *ṣuquʿ* (Marzubānī, *Muwaššaḥ*, 13). A breach of the prosodic rules is, however, a rarity and is not considered a poetic license, or *ḍarūra*, by the native tradition, where such phenomena are discussed as prosodic faults and blemishes. Normally, the reverse happens, and prosody prevails over grammar and accepted usage.

This posed a problem for the grammarians, since the codification of Classical Arabic grammar was primarily based on pre-Islamic and early Islamic poetry, the main source of *šawāhid* or linguistic evidence for 'correct' Arabic. The study of poetic license, therefore, dates from the very beginning of the study and codification of the language: near the beginning of his celebrated grammar, Sībawayhi discusses poetic license in a chapter entitled "What is tolerated in poetry" or, more precisely, "What poetry tolerates" (*Mā yaḥtamilu š-šiʿr*). He does not yet use the term *ḍarūra*, although he employs words derived from the same root. The chapter opens, "Know that in poetry is allowed (*yajūzu*) what is not allowed in (ordinary) speech (*kalām*)" (Sībawayhi, *Kitāb* I, 8–13); the last-mentioned word was taken as a synonym of prose (*naṯr*) by most commentators, even though there are indications that to Sībawayhi it had a more general sense, including poetry, which implies that the deviant and irregular 'licenses' are, somehow, not '(proper) speech' (Iványi 1991:211). Later scholars have written commentaries on this chapter or devoted monographs to the topic, among them al-Ḥasan ibn ʿAbdallāh as-Sīrāfī (d. 368/978), Muḥammad ibn Jaʿfar al-Qazzāz (d. 412/1021), and Ibn ʿUṣfūr al-ʾIšbīlī (d. 669/1270); poetic licenses are also treated in some works on poetry and literary criticism, such as *al-Muwaššaḥ* by al-Marzubānī (d. 384/993) and *al-ʿUmda* by Ibn Rašīq (d. ca. 456/1065).

The study of poetic license or 'necessity' should take into account that the language of poetry (→ *šiʿr* is not identical with that of everyday prose, or even literary prose. It is not always possible to distinguish between what is properly considered a license (and thus exceptionally allowed) and what is part and parcel of Arabic poetic diction. An example is the matter of diptote nouns: treating such nouns as triptotes by giving them a *tanwīn* (ending in *-n*) would be a gross error in prose, whereas in poetry it is extremely common and not considered a blemish at all, even though it may be described as *ḍarūra* by the commentators. Another example is the use of the particle *wa-* with a following indefinite noun in the genitive, which may be rendered as 'many a...' or 'I often think of...': this idiom, explained as a shortening of *wa-rubba* with following genitive, is restricted to poetry, but it could hardly be called a poetic license and is never discussed as such. There is no clear boundary between these very ordinary poetic features and licenses that are merely tolerated. Similarly, there is no hard-and-fast distinction between tolerated licenses and plain mistakes or blemishes that are condemned, which the grammarians and critics would call *ʾaxṭāʾ* 'errors', *ʾaġlāṭ* 'mistakes', or *maʾāxiḏ* 'reprehensible matters'. Moreover, the distinction between license and error may also depend on its provenance: What is accepted as a license in a pre-Islamic or an Umayyad poet would in some cases not be tolerated from a 'modern' (*muḥdaṯ*) poet of the Abbasid era. Matters are further complicated by the fact that poets, as is their habit everywhere, have often made their own rules and consciously flouted the rules of grammar. Thus, many oddities in early Islamic → *rajaz* poetry, which is very strongly represented in works on grammar and lexicography as serious evidentiary material, may well be explained as intentional violations of linguistic rules, not rarely intended humorously (see Ullmann 1966:214–217). Many features listed as licenses derive in fact from dialects. It should be mentioned here that the → *Qurʾān*, even though its language may strike us as 'poetic', is not considered poetry by Arabic or Muslim standards. Being traditionally taken to be God's literal speech, the words 'license' or *ḍarūra* can obviously not be applied to its numerous linguistic oddities (e.g. Sīrāfī, *Šarḥ* II, 119, 132; and see Ibn Rašīq, *ʿUmda* II, 277: in the *Qurʾān* such things are "eloquence and perfection rather than arbitrary change and necessity"). Nevertheless, some features found regularly in rhyme (e.g. *as-sabīlā* for *as-sabīla*), which are called *ḍarūra* in poetry, are declared identical with the same feature in Qurʾānic assonance or → *sajʿ* (e.g. Sīrāfī, *Šarḥ* II, 100; Ibn ʿUṣfūr, *Ḍarāʾir* 14).

"Poets are the commanders of speech: they shorten long vowels, lengthen short ones, they prepose and postpose" (Ibn Fāris, *Ṣāḥibī* 468; Suyūṭī, *Muzhir* II, 471, attributed to al-Xalīl in

Qarṭājannī, *Minhāj* 143). This statement seems to imply that poets are wholly free to change the rules of morphology and syntax. This extreme view was apparently held by ʾAbū ʿAlī al-Fārisī. However, Ibn Fāris continues, "It is no good saying that a poet may say things in his poetry that are not permitted [grammatically] in case of necessity (*ḍarūra*)", and he sets out his view more fully in his treatise *Ḏamm al-xaṭaʾ fī š-šiʿr* 'Reproving faults in poetry'. Al-ʿAskarī, in a handbook for would-be poets and prose writers, tells them to avoid "committing poetic licenses (*irtikāb aḍ-ḍarūrāt*)", as if it were some kind of crime; though linguists may call them 'license' (*ruxṣa*), they are ugly and spoil the beauty of one's speech. The ancient poets used them not knowing how ugly they were; they were the first, and being the first is full of pitfalls (ʿAskarī, *Ṣināʿatayn* 156). It is argued that ancient poetry was extemporized or made without much reflection, which provides an excuse for licenses. This is countered by pointing out that not all ancient poetry was made spontaneously and that much 'modern' poetry is extemporized (see the discussion in Ibn Jinnī, *Xaṣāʾiṣ* I, 323–327). One writer, al-Muẓaffar al-ʿAlawī (d. 656/1258), even finds that modern poets have a better excuse if they violate the rules, precisely because the ancient poets were naturally gifted, which made poetry easier to them (Muẓaffar, *Naḍra* 239); therefore, the moderns are permitted to adopt all licenses found in early poetry, with only a few exceptions. However, most grammarians and critics are less categorical and admit that some licenses are ugly and to be avoided, whereas others are less abhorrent. Very ugly, for instance, is giving the comparative form (*ʾafʿalu min*) a triptote declension (*ʾafdalun min*), even though other diptotes are regularly treated as triptotes in poetry; or lengthening a vowel which would result in nonexisting patterns, such as *ʾanḏ̣ūr* for *ʾanḏ̣ur* 'I see'. On the other hand, it is not ugly if the word-final morpheme *-āʾ* is used instead of *-ā* or vice versa (e.g. Ibn Rašīq, *ʿUmda* II, 269; Qarṭājannī, *Minhāj* 383; Suyūṭī, *Muzhir* I, 188–189).

The various kinds of licenses may be classified and categorized in several ways: on the basis of the linguistic levels involved (phonology, morphology, syntax); of acceptability and 'seriousness'; or of cause, intention, and stylistic or poetic effect (intended jesting, the demands of rhetoric or diction overriding the exigencies of prosody, etc.). Some linguists adopted a rather mechanical system and classified the *ḍarūrāt* on the basis of the transformational categories: addition (A > AB), deletion (AB > A), substitution (A > B), and permutation (AB > BA), or an extended version of this classification, but irrespective of the linguistic level or the stylistic effect. As-Sīrāfī, taking Sībawayhi as his starting point, recognizes seven categories: (i) *ziyāda* 'addition, augmentation'; (ii) *nuqṣān* 'diminishing, subtraction'; (iii) → *ḥaḏf* 'elision, deletion'; (iv) *taqdīm* 'pre-position, forward placement'; (v) *taʾxīr* 'post-position, backward placement'; (vi) → *ʾibdāl* 'substitution, interchange'; (vii) *taġyīr wajh min al-ʾiʿrāb* 'change of desinential inflection'. Categories (ii) and (iii) are closely related and are in fact discussed together, as are categories (iv) and (v). In the following (largely based on Sīrāfī, *Šarḥ* II, 95–256), some examples of each category are given. Surveys in Western languages, with examples, may also be found in Wright (1896–1898:II, 373–390) and Freytag (1830:466–527).

i. Addition (*ziyāda*). A short vowel may be lengthened. Since an Arabic verse cannot end in a short vowel or with *tanwīn*, words such as *aṯ-ṯiklu* or *ṯiklun* regularly become *(aṯ-)ṯiklū* in rhyme. This is a basic difference between prose and poetry and can hardly be called a poetic license, but it is mentioned by as-Sīrāfī for the sake of completeness and justified as providing euphony in reciting verse. If a diptote is treated as a triptote, this also implies an addition, e.g. *qaṣāʾidu* > *qaṣāʾidun, Makkata* > *Makkatin*. Very common in poetry of all ages, this is condoned with the somewhat dubious argument that the triptote declension is the basis or origin. Many authorities condemn this in the case of the comparative (*ʾafḍalun minka*), although the Basrians allow it (Sīrāfī, *Šarḥ* II, 102). Poets of all periods, certainly, seem to have avoided it.

A different kind of addition is the distortion of a word by lengthening ('doubling') a consonant (*Jaʿfarun* > *Jaʿfarrun*) or a vowel (*darāhim* > *darāhīm*). Sometimes these are merely lexical oddities that do not breach the rules of morphology (the pattern of *darāhīm* is, after all, a common one). However, additions may also result in the breaking of morphological rules by yielding nonexistent word patterns, or by flouting Arabic phonotactics, as when *rādd* appears as *rādid, al-ʾajalli* as *al-ʾajlali* (from a line by the *rajaz* poet ʾAbū n-Najm al-ʿIjlī), or *māḏin* as *māḏiyin* (from a line by Jarīr, which is read differently in other versions). Again, such forms are to some extent tolerated since they are said to have reverted to the 'original' forms.

Very common is treating the *hamzat al-waṣl* as a *hamzat al-qaṭʿ*: *jāwaza l-iṯnayni* becoming *jāwaza l-ʾiṯnayni*, or *bi-smihī* becoming *bi-ʾismihī*, adding a syllable.

ii. Elision (*ḥaḏf*). Names and nouns are found shortened, especially in vocative forms, e.g. *Fāṭimatu* to *Fāṭima, Ḥāriṯ* to *Ḥāri, yā ṣāḥibī* 'my friend!' to *yā ṣāḥi; ʾa-ʿāḏilatu* 'you [fem.] who reproaches [me]!' to *ʾa-ʿāḏila*, etc. This phenomenon, called *tarxīm* 'softening', seems to reflect the spoken language, and since this may be found in prose as well (→ truncation), it is dubious whether it should be called a poetic license. There are, however, some more startling cases of elision: *maṭā* for *maṭāyā* 'riding animals' and *al-ḥamā* (in rhyme position) for *al-ḥamāmi* 'of pigeons'. The former (Sīrāfī, *Šarḥ* II, 142) comes from an anonymous fragment which obviously belongs to light verse, and the latter is by the *rajaz* poet al-ʿAjjāj; one strongly suspects that neither poet was serious. More normal elisions are shortenings of the suffix *-āʾ* to *-ā*, the elision of *hamz* in forms such as *law-anna* (for *law ʾanna*), contractions such as *mil-* (for *mina l-* 'from the'), or the elision of the desinential ending in *hanki* (for *hanuki* 'your [fem.] thing', in an obscene sense), all apparently from ordinary speech.

iii. Substitution (*badal* or *ʾibdāl*). Some of the examples given are in fact shortenings: *ʾarānī, ḍafādī* for *ʾarānib* 'rabbits' and *ḍafādiʿ* 'frogs', but to the grammarians this counted as the substitution of one consonant (in this case *y*) for another (*b* and *ʿ*, respectively). The substitution may take place on the level of phonology (as in the examples given before); on the level of morphology, as in *ʿaṣaykā* (in rhyme position) for *ʿaṣayta* 'you were disobedient', the former apparently a dialectal form, or *ka-hā* (solecism for *miṯlahā* 'like her'); and on the level of lexicon and semantics, as when a poet uses *mašāfir* 'camel's lips' or *ʾaḏlāf* 'hooves' for the lips or ankles of a human being. Naturally, only the context can clarify whether this is a matter of 'necessity' and poetic license or of intended poetic effect.

iv. Pre- and post-position (*taqdīm wa-taʾxīr*). This concerns all deviations from the customary word order as determined by the syntactic norm. It became a standard topic in the scholastic study of eloquence (*ʿilm al-balāġa*), as part of the *ʿilm al-maʿānī*, which deals with the semantic and stylistic aspects of syntax. Of course, not every deviation is caused by the strictures of meter and rhyme. The interaction of prosody and syntax in early Arabic poetry, as compared with prose, has been investigated by Bloch (1946). An inflected language such as that of Classical Arabic poetry could be expected to tolerate a relatively greater freedom in the word order; nevertheless, in Arabic poetry it is much less free than, for instance, in Classical Greek or Latin verse. Among the examples discussed by the grammarians is the insertion of a word between the elements of a *status constructus* (→ *ʾiḍāfa*), normally not allowed: *bi-kaffi yawman yahūdiyyin* 'in the hand of, one day, a Jew', in a line by ʾAbū Ḥayya an-Numayrī). Critics and grammarians have few problems with this and similar cases. At the other extreme stands the notorious line by al-Farazdaq, often quoted as a deterrent example because of its contorted syntax: *Wa-mā miṯluhū fī n-nāsi ʾillā mumallakan/ ʾabū ʾummihī ḥayyun ʾabūhu yuqāribuh* 'there is no one like him among people, except a prince, whose mother's father is his father, who resembles him', meaning 'there is no one alive among people who is like him, except a prince whose mother's father is his father'.

v. Change of desinential inflection (*taġyīr al-ʾiʿrāb ʿan wajhihi*). Inflection, the correct use of case endings and verbal moods, is the main concern – one might say obsession – of the grammarians, since it is the hallmark and shibboleth of pure and correct Arabic. Therefore, a separate section is devoted to it, even though it could also be described under the heading 'substitution'; another final section deals with masculine and feminine forms being substituted for each other

> (*taʾnīṯ al-muḏakkar wa-taḏkīr al-muʾannaṯ*). Forced by the rhyme, a poet may use a subjunctive (*naṣb*) instead of an indicative imperfect (*muḍāriʿ*), an accusative (*naṣb*) instead of a nominative (*rafʿ*), or a feminine instead of a masculine form, etc. Whenever possible, the grammarians try to find a rationale other than mere ‘necessity’ caused by meter or rhyme. Thus, the genitive in a line by Imraʾ al-Qays: *kabīru ʾunāsin fī bijīdin muzammalī* (for *muzammalin*) ‘an elder of the people, in a striped cloak wrapped’, where a nominative *muzammalun* would be normal, is explained as an instance of attraction or ‘contiguity’ (*mujāwara*), whereby a word may take the case ending of an immediately preceding word. Since more cases of this phenomenon are found, it may be considered one of the ‘habits of the Arabs’. In the numerous commentaries on this line, the term *ḍarūra* is strikingly absent; but in at least one monograph on poetic licenses (Qazzāz, *Mā yajūzu* 146), it is given as a case of *ḍarūra*.

These ‘habits of the (pre-Islamic and early Islamic) Arabs’ are discussed at length in influential works by the two great contemporaries, the grammarians Ibn Jinnī (d. 392/1002), who uses the term *šajāʿat al-ʿarabiyya* ‘the boldness of the Arabic language’ (*Xaṣāʾiṣ* II, 360–441), and Ibn Fāris (d. 395/1004), who repeatedly uses the term *sunan al-ʿArab* in his *Ṣāḥibī* (much quoted by others, including as-Suyūṭī in his *Muzhir*). One also finds terms such as *ittisāʿ* or *tawassuʿāt* ‘expansions [of normal usage]’ – in fact various kinds of license. This ‘boldness’ and these ‘habits’ or ‘expansions’ include much of what could be termed poetic licenses: forms of elision, irregularities in the uses of grammatical persons, grammatical gender, number or tense, word order, etc. Ibn Rašīq, in his chapter on *ruxaṣ*, incorporates several of these topics. There is a difference, for in the discussion of the ‘habits of the Arabs’ both prose and poetry are considered, and the word ‘necessity’ (*ḍarūra*) cannot be applied, therefore, in spite of the close connections. Moreover, features that lie between linguistics, stylistics, and poetics are also included, such as figurative language (→ *majāz*) and metaphor (→ *istiʿāra*). In any case, the discussion of these ‘habits of the Arabs’, which focuses on precise linguistic and stylistic topics, is more satisfactory than the treatment by as-Sīrāfī, characterized by a mechanical subdivision that tends to ignore the boundaries between linguistic levels. Ibn ʿUṣfūr's monograph on *ḍarūra* is more systematic than as-Sīrāfī's, for he reduces the mutational categories to four (addition, deletion, permutation, and substitution), each subdivided as to the level (vowel, morpheme, or word). Al-Qazzāz abandons this system altogether and lists specific linguistic features in more than 140 short sections.

A broad interpretation of poetic license is offered by al-Muẓaffar al-ʿAlawī in his handbook for poets and critics, in his chapter “What a poet is allowed or is not allowed to use” (*Naḍra* 239–292). It incorporates not only what is normally discussed as *ḍarūrāt* but also the ‘habits of the Arabs’ and violations of meter and rhyme, thus bringing together topics often treated separately.

When poetic genres using the spoken language or at least a nonclassical form of the language, such as the *zajal*, became accepted by the literary critics, a new set of rules was developed in which the standard rules were in a sense reversed: The use of classical forms and *ʾiʿrāb* was now condemned or at most allowed as a license (Ḥillī, *ʿĀṭil*); it often serves the purposes of humor and parody. Modern poetry has so many forms that there is no longer any generally accepted set of rules regarding poetic license.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

ʿAskarī, *Ṣināʿatayn* = ʾAbū Hilāl al-Ḥasan ibn ʿAbdallāh al-ʿAskarī, *Kitāb aṣ-ṣināʿatayn al-kitāba wa-š-šiʿr*. Ed. ʿAlī Muḥammad al-Bijāwī and Muḥammad ʾAbū l-Faḍl ʾIbrāhīm. Cairo: ʿĪsā al-Bābī al-Ḥalabī, 1971.

Baġdādī, *Xizāna* = ʿAbd al-Qādir ibn ʿUmar al-Baġdādī, *Xizānat al-ʾadab wa-lubāb lisān al-ʿArab*. Ed. ʿAbd as-Salām Muḥammad Hārūn. Cairo: Dār al-Kātib al-ʿArabī and al-Hayʾa al-Miṣriyya al-ʿĀmma li-l-Kitāb, 1967–1986.

Ḥillī, *ʿĀṭil* = ʾAbū l-Maḥāsin ʿAbd al-ʿAzīz ibn Sarāyā Ṣafī d-Dīn al-Ḥillī, *al-ʿĀṭil al-ḥālī wa-l-muraxxaṣ al-ġālī*. Ed. Wilhelm Hoenerbach. Wiesbaden: F. Steiner, 1956.

Ibn Fāris, *Ḏamm al-xaṭaʾ* = ʾAbū l-Ḥusayn ʾAḥmad Ibn Fāris, *Ḏamm al-xaṭaʾ fī š-šiʿr*. Ed. Ramaḍān ʿAbd at-Tawwāb, *Majallat Maʿhad al-Maxṭūṭāt al-ʿArabiyya* 25 (1979) 29–59. [Also in Sanni (1993).]

——, *Ṣāḥibī* = ʾAbū l-Ḥusayn ʾAḥmad Ibn Fāris, *aṣ-Ṣāḥibī*. Ed. as-Sayyid ʾAḥmad Ṣaqr. Cairo: ʿĪsā al-Bābī al-Ḥalabī, [1977].

Ibn Jinnī, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. Ed. Muḥammad ʿAlī an-Najjār. Cairo: Dār al-Kutub al-Miṣriyya, 1952–1957.

Ibn Rašīq, *ʿUmda* = ʾAbū ʿAlī al-Ḥasan Ibn Rašīq al-Qayrawānī, *al-ʿUmda fī maḥāsin aš-šiʿr wa-ʾādābihi wa-naqdihi*. Ed. Muḥammad Muḥyī d-Dīn ʿAbd al-Ḥamīd. 2 vols. Repr., Beirut: Dār al-Jīl, 1972.

Ibn ʿUṣfūr, *Ḍarāʾir* = ʾAbū l-Ḥasan ʿAlī ibn Muʾmin Ibn ʿUṣfūr al-ʾIšbīlī, *Ḍarāʾir aš-šiʿr*. Ed. as-Sayyid ʾIbrāhīm Muḥammad. Beirut: Dār al-ʾAndalus, 1980.

Marzubānī, *Muwaššaḥ* = ʾAbū ʿUbaydallāh Muḥammad ibn ʿImrān al-Marzubānī, *al-Muwaššaḥ*. Ed. ʿAlī Muḥammad al-Bijāwī. Cairo: Nahḍa Miṣr, 1965.

Muẓaffar, *Naḍra* = al-Muẓaffar Ibn al-Faḍl al-ʿAlawī, *Naḍrat al-ʾiġrīḍ fī nuṣrat al-qarīḍ*. Ed. Nuhā ʿĀrif al-Ḥasan. Damascus: Majmaʿ al-Luġa al-ʿArabiyya, 1976.

Qarṭājannī, *Minhāj* = ʾAbū l-Ḥasan Ḥāzim ibn Muḥammad al-Qarṭājannī, *Minhāj al-bulaġāʾ wa-sirāj al-ʾudabāʾ*. Ed. Mohamed Habib Belkhodja [Muḥammad al-Ḥabīb Ibn al-Xawja]. Tunis: Dār al-Kutub aš-Šarqiyya, 1966.

Qazzāz, *Mā yajūzu* = ʾAbū ʿAbdallāh Muḥammad Ibn Jaʿfar al-Qazzāz al-Qayrawānī, *Mā yajūzu li-š-šāʿir fī ḍ-ḍarūra*. Ed. al-Munjī al-Kaʿbī. [Tunis: ad-Dār at-Tūnisiyya li-n-Našr,] 1971.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *Kitāb Sībawayhi*. 2 vols. Būlāq: al-Maṭbaʿa al-Kubrā al-ʾAmīriyya, 1316–1317 A.H.

Sīrāfī, *Mā yaḥtamilu* = ʾAbū Saʿīd al-Ḥasan ibn ʿAbdallāh as-Sīrāfī, *Mā yaḥtamilu š-šiʿr min aḍ-ḍarūra*. Ed. ʿAwaḍ ibn Ḥamd al-Qāzī. [Cairo: Dār al-Maʿārif,] 1991. [Separate ed. of the section on poetic license in as-Sīrāfī, *Šarḥ Kitāb Sībawayhi*, II, 95–132.]

——, *Šarḥ* = ʾAbū Saʿīd al-Ḥasan ibn ʿAbdallāh as-Sīrāfī, *Šarḥ Kitāb Sībawayhi*. Ed. Ramaḍān ʿAbd at-Tawwāb, Maḥmūd Fahmī Ḥijāzī, and Muḥammad Hāšim ʿAbd ad-Dāyim. 2 vols. Cairo: al-Hayʾa al-Miṣriyya al-ʿĀmma li-l-Kitāb, 1986.

Suyūṭī, *Muzhir* = Jalāl ad-Dīn ʾAbū l-Faḍl ʿAbd ar-Raḥmān ibn ʾAbī Bakr as-Suyūṭī, *al-Muzhir fī ʿulūm al-luġa wa-ʾanwāʿihā*. Ed. Muḥammad ʾAḥmad Jād al-Mawlā, ʿAlī Muḥammad al-Bijāwī, and Muḥammad ʾAbū l-Faḍl ʾIbrāhīm. Cairo: ʿĪsā al-Bābī al-Ḥalabī, n.d.

Secondary sources

Bloch, Alfred. 1946. *Vers und Sprache im Altarabischen*. Basel: Verlag für Recht und Gesellschaft.

Freytag, Georg Wilhelm. 1830. *Darstellung der arabischen Verskunst*. Bonn: C. Cnobloch. (Repr., Osnabrück: Biblio-Verlag, 1968.)

Iványi, Tamás. 1991. "*Qad yaǧūz fī š-šiʿr*: On the linguistic background of the so called poetic licenses in Sībawayhi's *Kitāb*". *The Arabist* (Budapest) 3-4.201–216.

Sanni, Amidu. 1992. "The exception that proves many rules: On poetic licences in Arabic literary theory". *Hamdard Islamicus* 15.29–42.

——. 1993. "A fourth century contribution to literary theory: Ibn Fāris's *Treatise on poetic licenses*". *Journal of Arabic Literature* 24.11–20.

Ullmann, Manfred. 1966. *Untersuchungen zur Raǧazpoesie: Ein Beitrag zur arabischen Sprach- und Literaturgeschichte*. Wiesbaden: O. Harrassowitz.

Wright, William. 1896–1898. *A grammar of the Arabic language*. 3rd ed., rev. W. Robertson Smith and M.J. de Goeje. Cambridge: Cambridge University Press.

Geert Jan van Gelder (University of Oxford)

Poetry, Language of → meter; poetic license; rhyme; rajaz; šiʿr

---

Polarity

The term 'polarity' is used for different meanings in linguistics. In connection with the Arabic system of → numerals, for example, it is used for gender polarity in the agreement between numerals and counted nouns. In the present entry, it is used for the contrast between negative and positive expressions in a language, whether these are syntactic or morphological. Negative polarity is a property of sentences modified with negative or downward monotonic operators. Negative-polarity items are expressions whose interpretation is affected by negative polarity, or which require the presence of a negative-polarity environment for felicitous use. Both Modern Standard Arabic and the Arabic dialects include a range of negative-polarity items, including nominal, adverbial, and idiomatic negative-polarity items. Several varieties of Arabic also show negative concord, the co-occurrence of multiple negative expressions in which only one negation is understood.

1. Introduction

In terms of classical logic, the polarity of a sentence describes whether it asserts a truth or a falsehood. Accordingly, sentences have either negative polarity, in the case of a negated meaning, or positive polarity, in the case of affirmative meaning. Many linguistic expressions are sensitive to negative polarity: their interpretation is affected by, or requires the presence of, a negative-polarity expression (referred to as the 'licensor'). Such expressions are referred to as 'negative-polarity items' (NPIs).

The study of negative-polarity items has been a major theme in modern general linguistics since Jespersen (1917). Important or influential studies include Klima (1964), Baker (1970), Ladusaw (1979), Linebarger (1987), Horn (1989), Progovac (1993), Fauconnier (1975), Kadmon and Landman (1993), Vallduví (1994), Krifka (1995), Lahiri (1997), van der Wouden (1997), Zwarts (1997), Giannikadou (1998), and Szabolsci (2004).

There has been as yet little investigation into negative polarity in Arabic. Work that has been done includes Woidich (1968), Ouhalla (1993, 1997b, 2002), Mohamed and Ouhalla (1995), and Benmamoun (1995, 1997). This entry provides a brief overview of negative-polarity items, insofar as they have been documented in Modern Standard Arabic and the Arabic dialects. Section 2 provides an overview of negative polarity as a general topic and negative-polarity items in both Standard Arabic and the dialects. Section 3 discusses negative concord.

2. NEGATIVE POLARITY

Traditionally, the term 'negative polarity' was used to describe sentences containing a negation morpheme. However, since Fauconnier (1975) and Ladusaw (1980), it has been recognized that many negative-polarity items are licensed in environments that lack a negation morpheme. Accordingly, the property that licenses them has been characterized as 'downward monotonicity' (Fauconnier 1975; Ladusaw 1980; Von Fintel 1999), 'pragmatic strengthening' (Kadmon and Landman 1993; Krifka 1995; Lahiri 1997), or as 'non-veridicality' (Giannikadou 1998, 1999, 2000).

Given the lack of theoretical work on negative-polarity items in Arabic, the remainder of this section is a brief survey of expressions in different varieties of Arabic that appear to be negative-polarity items (examples from Internet corpora are indicated with 'www').

2.1 *Nominal negative-polarity items*

The most common nominal negative-polarity items are cognates of Standard Arabic *ʾaḥad* 'one, anyone' and *šayʾ* 'thing, anything'. These resemble English *any*-nouns in having both negative and non-negative interpretations, as in (1).

(1a)	*mā-šuft-iš*	*iši*	*bass*	*ʿala*
	Neg-saw.1s-Neg	thing	but	on
	kull	*ḥāl*	*ṣuwar*	*hayfāʾ*
	every	case	pictures	Haifa
	ḥilw-a	*ktīr*		
	beautiful.f	much		

'I didn't see anything, but in any case the pictures of Haifa are very nice' (Palestinian Arabic)

(1b)	*ma-baġēt*	*ḥad*	*yrūḥ*	*maʿa-k*
	not-wanted.1s	one	go.3ms	with-cl.2s

'I didn't want anyone to go with you' (Moroccan Arabic)

A common negative-polarity determiner is *ʾayy* 'which, any', used both in Standard Arabic (2a) and in dialect (2b) to construct strongly contrastive negative-polarity items.

(2a)	*lam*	*yafham*	*ʾayy*
	not.Past	understand.3ms	any
	ʾaḥad	*ġayr-ī*	*ʾabadan*
	one.ms	other.than-cl.1s	ever

	ʾayy	*mawqif*	*min*	*mawāqifi-ki*
	any	opinion	from	opinions-cl.3fs

'There isn't anyone other than me who ever understood any one of your opinions' (Modern Standard Arabic)

(2b)	*bass*	*bi-ṣarāḥa*	*mā-fī*
	but	with-honesty	not-exist
	ʾayy	*ḥada*	*b-iqdar*
	any	one	be-able.3ms
	yiḥall	*maḥall*	*zīdān*
	occupy.3ms	place	Zidan

'...but honestly, there isn't any person who can take Zidan's place' (Levantine Arabic)

According to Erwin (1969), in Iraqi Arabic the quantifier *kullši* 'everything' has a negative-polarity interpretation in negative sentences (3).

(3) *la* *waḻ̣ḻ̣a,* *kullši* *ma-ʿind-i*
no by-God anything not-at-cl.1s
'I don't have anything' (Erwin 1969:88)

2.2 *Adverbial negative-polarity items*

Negative-polarity adverbs include *bi-l-marra* 'once, ever' and cognates of *ʿumr-* 'ever' (lit. 'life, age'). Like *ʾaḥad* and *šayʾ*, *bi-l-marra* is ambiguous between the interpretation as a negative-polarity item and a non-negative interpretation (4).

(4a) *ʾanta* *lam* *tūḍiḥ*
you.ms not.past clarify.2ms
šayʾan *bi-l-marra*
thing-Acc in-the-once
'You didn't ever clarify anything' (Modern Standard Arabic: www)

(4b) *ʾāsif* *yā* *ʾaxī,* *kalām-ak*
sorry oh brother-cl1s words-cl.2ms
ġēr *ṣaḥīḥ* *bi-l-marra*
other true in-the-once
'Sorry, Brother, your talk isn't ever true' (Egyptian Arabic: www)

The adverb *ʿumr-* 'ever' (which has cognates in most dialects) has only negative-polarity interpretations and must co-occur with a licensor, as in (5).

(5a) *ʾana* *ʿumr-i* *ma-šuft* *wāḥad*
I life-cl.1s not-saw.1s one
miṯl-u
like-cl.3ms
'I have never seen anyone like him' (Levantine Arabic: www)

(5b) *ḅḅa* *ʿammr-u*
father-cl.1s life-cl.3ms
ma-ka-išreb
not-Cont-drink.3ms
'My father, he never drinks' (Moroccan Arabic: Harrell 1966:260)

2.3 *Negative-polarity items as auxiliaries*

In many dialects, certain motion verbs are used as aspectual → auxiliaries or sentential connectives (Schmidt and Kahle 1918; Blau 1960; Hussein 1990; Mitchell and al-Hassan 1994; Khalaily 1997). Examples include *qām* 'to rise, stand' and cognates of *ʿād*, *ʿāwad*, or *rajaʿ* 'to return'. As auxiliaries, they express inchoative or iterative aspect, or, in the case of *qām* (→ serial verbs), a meaning very similar to that of Modern Standard Arabic *fa-* 'then, so' (6).

(6) *qām* *rajaʿ* *nām*
stood.3ms returned.3ms slept.3ms
'Then he went to sleep again' (Levantine Arabic: www)

As auxiliaries, they appear to have negative-polarity interpretations. For example, negating *qām* results in a meaning translatable with English *ever* (7).

(7) *ma-qām-iš* *yixallī-hin*
Neg-stood.3ms-Neg allowed.3ms-cl.3fp
yiṭlaʿin
go-out.3fp
'He didn't ever let them venture out' (Bir Zeit Palestinian Arabic: Schmidt and Kahle 1918, par. 46.1)

Similarly, negating *ʿād* or *ʿāwad* yields an interpretation translatable as English *anymore* (8).

(8) *ma-ʿadt-š* *tiʿrif* *iši*
Neg-returned.2ms.Neg know.2ms thing
'You don't know anything anymore' (Levantine Arabic: www)

In Moroccan Arabic, auxiliaries like *ʿad* 'anymore' and *gaʿ* 'at all' seem to have only the negative-polarity interpretation, as in (9).

(9a) *ʿad* *ma-yiji*
anymore not-come.3ms
'He doesn't come anymore'

(9b) *gaʿ* *ma-ta-txruž* *berra*
at.all not-Cont-go.out.3fs outside
'She doesn't go outside at all'

2.4 *Negative-polarity idioms*

Another category of negative-polarity items found in Arabic are → idioms that are felicitous only in the presence of a negative-polarity licensor. Stowasser and Ani (1964) give examples in Syrian Arabic such as *tʿallaq ʿala rijl-o* 'to hang from someone's shoe' (cf. English *couldn't care less*), *rafaʿ l-o qašše* 'to lift a match for someone' (10).

(10a) *mā-bitʿallaq* *ʿala* *rijl-i*
not-hang.3ms upon foot-cl.1s
šū *byaʿtaqed* *huwwe*
what thinks.3ms he
'What he says doesn't hang from my shoe' (Stowasser and Ani 1964:35)

(10b) *waḷḷa* *mā-barfaʿ-l-o* *qašše*
by-God not-lift.1s-to-cl.3ms match
'I won't lift a match for him' (Stowasser and Ani 1964:139)

3. NEGATIVE CONCORD

Negative concord describes sentences in which one or more N-words (words that can express negation in a fragment answer) occur with one or more other negative expressions (morphemes that express sentential negation) in a negative sentence, in which only one of them contributes negative meaning (Giannikadou 2000; Watanabe 2004). Negative concord is the subject of an extensive literature. Influential studies include Jespersen (1917), Labov (1972), Laka (1990), Ladusaw (1992), Progovac (1993), Vallduví (1994), van der Wouden (1997), Haegeman and Zanuttini (1996), Deprez (2000), Herburger (2001), Giannikadou (2000), and Watanabe (2004).

To illustrate, the Modern Standard Arabic adverb *muṭlaqan* 'completely, at all' can express negation in a fragment answer, as in (11).

(11) Q. *hal* *hāḏā* *l-maraḍu*
Query this.ms the-disease.ms
xaṭīrun?
dangerous.ms
A. *lā* *muṭlaqan*
no not.at.all
'Is this disease dangerous?' 'No, not at all' (Modern Standard Arabic: www)

Elsewhere, *muṭlaqan* appears along with a negation morpheme and does not itself express negation, indicating that it undergoes negative concord, as in (12).

(12) *aṣ-ṣīn* *lan* *tasmuḥa*
the-China not.Fut tolerate.3fs
muṭlaqan *bi-stiqlāli* *tāywān*
not.at.all with-independence Taiwan
'China will not tolerate Taiwan's independence at all' (Modern Standard Arabic: www)

As shown by data in published grammars, theoretical work, and text corpora (print, electronic, and broadcast) and elicited from native speakers, negative concord occurs in the Levantine dialects (Schmidt and Kahle 1918, 1930; Blau 1960; Cowell 1964), Egyptian Arabic (Woidich 1968), some Gulf dialects (Qafisheh 1977), North African dialects (Marçais 1977), and particularly in Moroccan Arabic (Harrell 1962, 1965, 1966; Ouhalla 1993, 1997b, 2002; Benmamoun 1995, 1997).

A common N-word in the Eastern dialects is the determiner *wala* 'not (even) one, not a single one'. Combined with a noun, it can be used to express negation, as in (13).

(13) Q. *šū* *qal-l-ak?*
what said.3ms-to-cl.2ms
A. *wala* *iši*
not.even thing

'What did he say to you?' 'Not a thing' (Palestinian Arabic, from the film *Paradise Now*)

However, when *wala* falls within the scope of sentential negation, it fails to express negation, as in (14).

(14) *ma-bansa* *wala* *wāḥid*
not-forget not.even one
min-nun. *wala* *wāḥid!*
from-cl3p not.even one
'I won't forget a single one of them. Not a single one' (Syrian Arabic, from the television program *Spotlight*)

This shows *wala* undergoing negative concord.

For some Levantine speakers, the negation morpheme itself can undergo negative concord with a preceding N-word. In the following example, the negation morpheme *mā-* follows *wala marra* 'not once' in a context that makes clear that only one negative meaning is expressed by the sentence, as in (15).

(15) *wala* *marra* *mā* *ʾinṭaradt*
not.even once not expelled.1s
min *iṣṣaff*
from the-class
'Not once have I been expelled from class'

The acceptability of such examples appears to be restricted to Palestinian.

Other N-words include adverbs like *ʾabadan* '(n)ever, (not) at all' and Modern Standard Arabic *muṭlaqan*, noted above. In Moroccan Arabic, *ḥətta* 'even' (Harrell 1962, 1965, 1966; Benmamoun 1997) can be used to express negation in fragment answers, satisfying the definition of an N-word:

> In English one would be more likely to say 'Not a one!' instead of 'He didn't take a single ticket!' in answer to 'How many tickets did he take?' Similarly in Moroccan one may say simply *ḥətta waḥda* instead of a full sentence such as, for example, *ma-ʿandu ḥətta werqa* 'He doesn't have a single ticket' (Harrell 1965:235)

This indicates that in Moroccan Arabic negative-polarity items with *ḥətta* undergo negative concord.

However, Moroccan negative-polarity items with *ḥətta* differ from those with *wala* in the Eastern dialects in that expressions with *ḥətta* never express negation in a full clause, regardless of their position with respect to the verb. Instead, they must be licensed by a negation morpheme in both pre-verbal and post-verbal positions, as in (16).

(16a) *ḥətta* *ḥadd* **(ma-)šāf-ni*
even one.ms not-saw.3ms-me
'Not even one person saw me'

(16b) **(ma-)šāf-ni* *ḥətta* *ḥadd*
not saw.3ms-me even one.ms
'Not even one person saw me'

Published grammars of other Arabic dialects include data that closely resemble the examples of negative concord given here but do not provide conclusive evidence of such. Without such evidence it is difficult to determine how widespread negative concord is in the Arabic dialects.

4. SUMMARY

Both Standard Arabic and various Arabic dialects exhibit a range of negative-polarity sensitive expressions and phenomena, including nominal, adverbial, and idiomatic negative-polarity items, as well as negative concord. Much work remains to be done both to inventory the kinds of negative-polarity items used in the dialects and to provide descriptive and theoretical analyses of their properties.

BIBLIOGRAPHICAL REFERENCES

Baker, Carl L. 1970. "Double negatives in English". *Linguistic Inquiry* 1.169–186.
Benmamoun, Elabbas. 1995. "Negative polarity and presupposition". *Perspectives on Arabic linguistics*, VII, ed. Mushira Eid, 47–66. Amsterdam and Philadelphia: J. Benjamins.
——. 1997. "Licensing of negative polarity items in Moroccan Arabic". *Natural Language and Linguistic Theory* 15.263–287.
——. 2000. *The feature structure of functional categories: A comparative study of Arabic dialects*. Oxford: Oxford University Press.

Blau, Joshua. 1960. *Syntax des palästinensischen Bauerndialektes Bir Zeits*. Walldort, Hessen: Verlag für Orientkunde.

Cowell, Mark. 1964. *A reference grammar of Syrian Arabic*. Washington, D.C.: Georgetown University Press.

Deprez, Vivianne. 2000. "Parallel (a)symmetries and the internal structure of negative expressions". *Natural Language and Linguistic Theory* 18.253–342.

Erwin, Wallace. 1969. *A basic course in Iraqi Arabic*. Washington, D.C.: Georgetown University Press.

Fauconnier, Giles. 1975. "Pragmatic scales and logical structure". *Linguistic Inquiry* 6.335–375.

Giannikadou, Anastasia. 1998. *Polarity sensitivity as (non-)veridical dependency*. Amsterdam and Philadelphia: J. Benjamins.

——. 1999. "Affective dependencies". *Linguistics and Philosophy* 22.367–421.

——. 2000. "Negative...concord?". *Natural Language and Linguistic Theory* 18.457–523.

Haegeman, Lilliane and Rafaella Zanuttini. 1996. "Negative concord in West Flemish". *Parameters and functional heads*, ed. Adriana Belleti and Luigi Rizzi, 117–179. Oxford: Oxford University Press.

Harrell, Richard S. 1962. *A short reference grammar of Moroccan Arabic*. Washington, D.C.: Georgetown University Press.

——. 1965. *A basic course in Moroccan Arabic*. Washington, D.C.: Georgetown University Press.

——. 1966. *A dictionary of Moroccan Arabic: Moroccan-English*. Washington, D.C.: Georgetown University Press.

Herburger, Elena. 2001. "Negative concord revisited". *Natural Language Semantics* 9.289–333.

Horn, Laurence. 1989. *A natural history of negation*. Palo Alto, Cal.: Center for the Study of Language and Information.

Hussein, Lutfi. 1990. "Serial verbs in colloquial Arabic". *When verbs collide: Papers from the 1990 Ohio State Mini-conference on Serial Verbs*, ed. Brian D. Joseph and Arnold M. Zwicky, 340–354. Columbus: Ohio State University.

Jespersen, Otto. 1917. *Negation in English and other languages*. Copenhagen: Host.

Kadmon, Nirit and Fred Landman. 1993. "Any". *Linguistics and Philosophy* 16.353–422.

Khalaily, Samir. 1997. *One syntax for all categories: Merging nominal atoms in multiple adjunction structures*. The Hague: Holland Institute of Generative Linguistics.

Klima, Edward. 1964. "Negation in English". *The structure of language*, ed. Jerry A. Fodor and Jerrold Katz, 246–323. Englewood Cliffs, N.J.: Prentice-Hall.

Krifka, Manfred. 1995. "The semantics and pragmatics of negative polarity items". *Linguistic Analysis* 25.209–357.

Labov, William. 1972. *Language in the inner city: Studies in the Black English vernacular*. Philadelphia: University of Pennsylvania Press.

Ladusaw, William. 1979. *Polarity sensitivity as inherent scope relations*. Ph.D. diss., University of Texas at Austin; publ. 1980, Indiana University.

——. 1992. "Expressing negation". *Proceedings of the Second Conference on Semantics and Linguistic Theory*, ed. Chris Barker and David Dowty, 237–259. Columbus: Ohio State University.

Lahiri, Utpal. 1997. "Focus and negative polarity in Hindi". *Natural Language Semantics* 6.57–123.

Laka, Itziar. 1990. *Negation in syntax: On the nature of functional categories and projections*. Ph.D. diss., Massachusetts Institute of Technology.

Linebarger, Marcia C. 1987. "Negative polarity and grammatical expression". *Linguistics and Philosophy* 10.325–387.

Marçais, Philip. 1977. *Esquisse grammatical de l'arabe maghrébin*. Paris: Librairie d'Amérique et d'Orient.

Mitchell, T.F. and S. al-Hassan. 1994. *Modality, mood, and aspect in spoken Arabic*. London: Kegan Paul International.

Mohamed, L. and Jamal Ouhalla. 1995. "Negation and modality in early childhood Arabic". *Perspectives on Arabic linguistics*, VII, ed. Mushira Eid, 69–90. Amsterdam and Philadelphia: J. Benjamins.

Mohammad, Mohammad A. 1998. "The syntax of indefinite subjects in equative sentences in Palestinian Arabic". Ms., University of Florida at Gainesville.

Ouhalla, Jamal. 1993. "Negation, focus, and tense: The Arabic *mā* and *lā*". *Rivista di Linguistica* 5.275–300.

——. 1997a. "The structure and logical form of negative sentences". *Linguistic Analysis* 27.220–244.

——. 1997b. "Remarks on focus in Standard Arabic". *Perspectives on Arabic linguistics*, X, ed. Mushira Eid and Robert Ratcliffe, 9–46. Amsterdam and Philadelphia: J. Benjamins.

——. 2002. "Negative sentences in Arabic". *Themes in Arabic and Hebrew syntax*, ed. Jamal Ouhalla and Ur Shlonsky, 299–320. Dordrecht: Kluwer Academic Publishers.

Progovac, Liljiane. 1993. "Negative polarity: Entailment and binding". *Linguistics and Philosophy* 23.101–139.

Qafisheh, Hamdi A. 1977. *A short reference grammar of Gulf Arabic*. Tucson: University of Arizona Press.

Rooth, Mats. 1992. "A theory of focus interpretation". *Natural Language and Linguistic Theory* 1.75–116.

Schmidt, Hans and Paul Kahle. 1918. *Volkserzählungen aus Palästina*, I. Göttingen: Vandenhoek und Ruprecht.

—— and Paul Kahle. 1930. *Volkserzählungen aus Palästina*, II. Göttingen: Vandenhoek und Ruprecht.

Stowasser, Karl and Moukhtar Ani. 1964. *A dictionary of Syrian Arabic: English-Arabic*. Washington, D.C.: Georgetown University Press.

Szabolsci, Anna. 2004. "Positive polarity – negative polarity". *Natural Language and Linguistic Theory* 22.409–452.

Vallduví, Enric. 1994. "Polarity items, n-words, and minimizers in Catalan and Spanish". *Probus* 6.263–294.

Von Fintel, Kai. 1999. "NPI licensing, Strawson entailment and context dependency". *Journal of Semantics* 16.97–148.

Watanabe, Akira. 2004. "The genesis of negative

concord: Syntax and morphology of negative doubling". *Linguistic Inquiry* 35.559–612.
Woidich, Manfred. 1968. *Negation und negative Sätze im Ägyptisch-Arabischen.* Ph.D. diss., Ludwig Maximilians University, Munich.
Wouden, Ton van der. 1997. *Negative contexts.* Ph.D. diss., Rijksuniversiteit Groningen.
Zwarts, Frans. 1998. "Three kinds of polarity items". *Plural quantification*, ed. Fritz Hamm and Erhard W. Hinrichs, 177–238. Dordrecht: Kluwer Academic Publishers.

Frederick M. Hoyt
(University of Texas at Austin)

Politeness

Politeness is a norm of social behavior that is expressed directly or indirectly by gestures or usage of common or personal expressions, either orally or in writing, showing endearment, respect, veneration; appreciation, esteem; awe; flattery; sycophancy or affection. Politeness may also indicate reservation or dissatisfaction in a way which does not bluntly offend or irritate another person (Brown and Levinson 1987). Thus, direct expression of politeness is the use of the words 'thank you' to indicate gratitude and appreciation, while the sentence 'this is an interesting dish' is a polite way to say that the food was not entirely to your taste.

Politeness, in Arabic *ʾadab* (less common: *kiyāsa*), is one of the highly praised qualities of man and is regarded as a vital ingredient in the 'relationship' between God and man. A dictum ascribed to the Prophet Muḥammad says *ḥusn al-ʾadab min al-ʾīmān* 'good behavior is part of the faith'. As in other languages, politeness is associated with a number of related concepts, such as *taṣarruf* and *sulūk* 'conduct, behavior', *tarbiya* 'bringing up', *tahḏīb* 'educating', *ṯaqāfa* 'culture', *xuluq/ʾaxlāq* 'morals; character', and *ḏawq* 'good taste; tact'. All these words indicate good nature, virtues, and high moral standards, reflected in the *mujāmalāt* 'etiquette, courtesy, civility' of the individual and the society. In view of the importance attached to the concepts of *ʿirḍ* 'honor', *karāma* 'good repute, self-respect', and *šaraf* 'dignity' in Arab society, politeness is no doubt an essential ingredient in daily contact and relationships between human beings.

Obviously, there are many differences, both geographical and social, in polite behavior in the Arab world. In this entry, a general picture is given that should not be regarded as universally valid. Because of the lack of studies on this topic, it is difficult to make fine-tuned distinctions. Some of the customs and terms given here are current in some layers of society and in some regions only. For general studies on Arabic politeness formulas, see Ferguson (1976), El-Sayed (1990), and Shivtiel (2005). A case study of politeness in Omani Arabic is Emery (2000). Piamenta (1979, 1983) is an inventory of religion-related expressions of politeness, especially in Jerusalem Arabic.

Among the customary gestures of politeness current in the Arab world, one finds the act of standing up in honor of a person or remaining standing as long as the other person is present; lowering of the voice while addressing a person (for instance in the relationships between children and parents, common people and dignitaries, young and old); shaking hands, bowing, cheek kissing (in some areas, nose rubbing), or touching the left side of one's chest with the right hand open (pointing at 'the heart') when greeting; thanking, swearing by God or on one's life, or indicating politely that one has had enough food; and in some areas even touching the beard of the addressee as a sign of friendly greeting or of swearing an oath. The custom of holding food with the right hand and leaving some food on the plate after finishing the meal are but two of the table manners current in the region that show politeness and good manners. Also, customs mainly current among nomadic society (i.e. the Bedouin), such as indicating to a guest who has stayed too long that he is expected to leave the tent by offering him a third cup of coffee, or shaking the empty cup of coffee to indicate that one does not wish to have more coffee, or by grabbing the hem of the host's cloak to hint that help is desperately needed, are all gestures of politeness, indicating a wish or desire without offending or pestering the addressee.

Verbally and in writing, Arabic has scores of phrases and idioms, demonstrating politeness and good manners by offering the user a variety of → terms of address, honorific titles, expressions of courtesy, and references to males or females, places, and situations. In addition, Arabic has different ways, often used as circumlocutions, to politely indicate dissatisfaction and dislike. These should by no means be regarded as hypocrisy but rather as evidence of self-control, patience, and forbearance (*ṣabr*).

As in other languages (e.g. French, German), the 2nd person plural is used when verbally addressing a dignitary or an old man or in official correspondence, e.g. *ʾarjū ʾan takūnū qad istalamtum risālatī*… 'I hope you [pl.] have received [pl.] my letter…'. However, the plural may also be used in the case of the 1st person instead of the singular as a polite way of expressing one's wish or requesting humbly, e.g. *narjū* 'we would like', *natašarrafu/yušarrifunā* 'we have the honor', *yusʿidunā/yasurrunā* 'we are delighted'.

For direct approach (→ greetings), the vocatives *yā* and rarely *ʾayyuhā* or *yā ʾayyuhā* 'oh!' may be used, though seldom, while the most common formulas are *as-sayyid/as-sayyida; al-ʾustāḏ; sayyidī/sayyidatī* 'Sir/Madam'; *sayyidātī wa-sādatī* 'Ladies and Gentlemen'; *mušahidīnā al-kirām/al-ʾaʿizzāʾ* 'our respected/dear viewers [in TV broadcasts]'; *ḥaḍrat*…; *ḥaḍrataka/ḥaḍrataki*, and in the plural *siyādatakum/saʿādatakum* 'Your Excellency; Sir/Madam'. One of the most popular formulas of address in the past was *yā mawlāya* 'Sir!'. In academic circles, the common approach is with the word *ad-duktūr/al-ʾustāḏ*…, followed by the surname, but among friends and colleagues by only the first name, e.g. Professor ʿAlī Ṭāhir will be approached officially as *ad-duktūr* or *al-ʾustāḏ Ṭāhir*, but by friends or colleagues as *ad-duktūr ʿAli*. Common usages which show politeness, respect, and affection between members of the family (→ kinship terms) and close friends are *al-wālid/al-wālida* 'father and mother'; *ʾaxī* 'my brother', *ʾuxtī* 'my sister'; *ʿazīzī/ʿazīzatī* 'my dear, darling'; *ḥabībī/ḥabībatī* 'my beloved'; *(yā) ʿēnī* 'my eye'; *(yā) rōḥi* 'my soul'. In archaic style, one may also find the expressions *ḥušāša qalbī/kab(i)dī* 'my last breath/my last spark of life', i.e. 'my dear'; *yā ʿuṣārat/suwaydāʾ qalbī; filḍa kabidī* 'my own flesh!'.

When approaching a specific holder of an official position or status, either orally or in writing, the following polite phrases are used: *jalālat al-malik* 'Your Majesty'; *sumuww al-ʾamīr* 'Your Royal Highness'; *faxāmat ar-raʾīs* 'Your Excellency, the President'; *dawla raʾīs al-ḥukūma/al-wuzarāʾ* 'Your Excellency, the Prime Minister'; *maʿālī l-wazīr* 'Your Excellency, the Minister'; *qadāsat al-bābā* 'Your Holiness, the Pope'; *samāḥat al-muftī* 'Your Eminence, the Mufti'; *ġibṭat al-baṭrīk* 'Your Eminence, the Patriarch'; *niyāfat al-muṭrān* 'Your Eminence, the Archbishop'; *faḍīlat aš-šayx* 'Your Grace, the Sheik'. Other titles are *janāb* 'Right Honorable' *saʿādat*…(e.g. to a judge); *ʿuṭūfat*…(to a pasha, sheik); *ṣāḥib al-ʿaḍama* (to a sultan). In the past, caliphs or other Muslim rulers were often addressed with the title *ʾamīr al-muʾminīn* 'Commander of the Believers'. Phrases of courtesy used by newspapers to refer to authors of articles in other newspapers are *al-ġarrāʾ, ar-raṣīfa*, or *ar-raṣīfa al-ġarrāʾ* 'noble colleague'.

For requests, the following are the most current polite expressions: *ʾarjū, ʾarjūka/ʾarjūki*, etc. 'I beg of you [masc./fem.]; please!'; *min faḍlika/min faḍliki* 'by your [masc./fem.] kindness', i.e. 'please'; *ismaḥ lī* 'pardon, forgive me'; *law samaḥt* 'with your permission' (also used to draw the attention of waiters, salespersons, etc.), *law takarramt/tafaḍḍalt* 'with your respect/grace', i.e. 'please, excuse me'; *maʿḏira* 'excuse me/forgive me/sorry'; *ʾarjū ʾan tatalaṭṭafa/tatafaḍḍala* 'please'; *ʾafandim* 'I beg your pardon?'; *ballāhi, tiʿmil maʿrūf* 'by God, will you do me a favor?'. When offering something, the common word is *tafaḍḍal* 'please!', while the polite answer is *ʾafḍalt* or simply *šukran* 'thanks'. A strong request, indicating helplessness and humble wish, is *daxīlak*, roughly 'I am under your protection', i.e. 'please, I beg of you'.

When addressing a man or a woman who has children, it is customary to use the *kunya*, i.e. 'father of…' or 'mother of…' followed by the name of the firstborn. Only mothers may normally have the names of either son or daughter; fathers are usually addressed with the name of a son, even if the son is not the firstborn child, e.g. *ʾabū Muḥammad* 'father of Muhammad', *ʾūmm Ḥamdān/Fāṭima* 'mother of Hamdan/Fatima'.

Within certain contexts, and social groups, a wife rarely approaches her husband or refers politely and affectionately to him by using his first name; she is more likely to use his *kunya* (*ʾabū Muḥammad*) or even *ibn ʿammī* 'my cousin', while the husband usually addresses his wife as *ʾumm Muḥammad* 'mother of Muhammad' or *bint ʿammī* 'my cousin'. The combination *ibn ʿammī/bint ammī* may also be politely used in the case of close friends addressing each other.

'Thank you' expressions include *šukran/šukran jazīlan/ʾaškuruka/mamnūn(īn)*. In colloquial Arabic, one also uses *kattar xērak* 'may God increase your goodness/well-being!' or *Allāh yirḍā ʿalēk* 'may God be pleased with you!'. The French word *mersī* or even *mersī ʾawī* 'strong *merci*!' is often used in Egypt, mainly by females. The common answer is, for example, *ʿafwan* or *al-ʿafū* 'don't mention it', and in more elaborate style *lā šukra ʿalā l-wājib* 'no need to thank for one's duty', *bāraka llāhu fīk* 'may God bless you!', *jazāk Allāh xayran/al-xayr* 'may God reward you with goodness!'.

Polite style in correspondence in Modern Standard Arabic includes opening the letter with *ḥaḍrat*…or *as-sayyid/as-sayyida*…'Sir/Madam', and with friends *al-ʾax/ʾaxī/ʾaxī wa-ʿazīzī/aš-šaqīq* (*al-fāḍil*) 'my dear/respected brother'; *al-ʾuxt/ʾuxtī* (*al-fāḍila*) 'my dear/respected sister'. In official correspondence, it is customary to place the name between two or more words denoting esteem, e.g. (*ḥaḍrat*) *al-ʾustāḏ al-fāḍil Maḥmūd an-Najjār al-muḥtaram/al-muwaqqar/al-fāḍil* 'the revered… the respected'. Closing formulas of official correspondence usually include *wa-tafaḍḍalū bi-qubūl fāʾiq al-iḥtirām* 'yours sincerely/faithfully' and/or in the case of friendly correspondence *al-muxliṣ* 'yours, yours sincerely'. In more orthodox or conservative circles, the formula *wa-s-salāmu ʿalaykum wa-raḥmatu llāhi wa-barakātuhu* 'may God's peace and His mercy be with you!' is commonly used. The same formula, but without the first conjunction, may also be used at the beginning of a letter immediately after the addressee's name (i.e. *as-salām ʿalaykum*…).

Another polite formula used in correspondence is *taḥiyya ṭayyiba* (or *taḥiyyāt ṭayyiba*) *wa-baʿd* or *taḥiyya ʿaṭira* (*wa-baʿd*) 'Greetings!'. These formulas were usually followed by the expression *ʾammā baʿdu* 'and after all this' (called in Arabic *faṣl al-xiṭāb*), which is less used in contemporary Arabic. Some formulas of address in correspondence in the past were lengthy and florid, e.g. *ʾaṭāla llāhu baqāʾakum wa-ʾadāma ʿizzakum wa-saʿādatakum wa-ʾatamma niʿmatahu ʿalaykum wa-kabata ʾaʿdāʾakum* 'may God give you long life and an everlasting might and happiness and bestow His grace upon you and subjugate your enemies!'.

When referring to or addressing people, some adjectives or attributes may be added, for instance when addressing one person, *šaxṣiyyatukum al-marmūqa* 'your notable personality', or when a politician addresses the whole nation, *ʾayyuhā š-šaʿb al-ʿarīq* lit. 'oh, deep-rooted nation', i.e. 'respectable nation'.

There are several polite expressions used in reference to women who are regarded as *ḥarīm* lit. 'those who are forbidden to other men' (from which the English word *harem* is derived), e.g. *rabbāt al-ḥijāl* 'the ladies'. A woman, and especially a married one, is supposed to be respectable, humble, and modest. Hence, when addressed or referred to, several expressions are current, stressing her merits, e.g. *ḥurma, qarīna, ʿaqīla*, *karīma* (also used for daughters), *ḥalīla, muḥsina, al-maṣūna, ṣāḥibat aṣ-ṣawn, ṣāḥibat aṣ-ṣawn wa-l-ʿiffa, ṣāḥibat al-ʿiṣma, bayḍat/rabbat al-xidr*, and *hānum* (from Turkish). A polite reference to someone else's wife, for instance when inquiring about her health, may be 'generalized' by using the word *al-ʾahl* 'the family' or *ʾumm al-ʾawlād* 'the mother of the children', and even *ʾumm al-ʾašbāl* 'the mother of the lion's cubs', i.e. 'the children', or *rabbat al-manzil* 'the lady of the house'. All these polite phrases suggest that the person who is referring to the woman has sincere and honest intentions. It is worth noting that while *imraʾa* (in the dialect *mara*) 'a woman' may be used generally to refer to a female, it may also have a derogatory sense in the Egyptian dialect, but not as a declined noun (*martī* 'my wife', etc.).

A man is politely referred to as *as-sayyid/al-ʾustāḏ* 'Mr.…'; *al-ʾax* 'the brother', *aṣ-ṣadīq* 'the friend', *az-zamīl/ar-raṣīf* 'the colleague'; and may be addressed with *(yā) ibn ḥalāl* lit. 'legitimate child', i.e. 'decent fellow', in some dialects also as *ya šāṭir* or *yā jadaʿ* roughly 'clever guy'.

Archaic words, mainly from Turkish, may still be heard, albeit less and less, such as *ʾafandī* (in Egypt also *ʾafandim*) 'Sir!'; the word *xawāja* 'Mr.…, Sir' may be used as a polite form of address for Westerners, although the word may have a derogatory meaning in some regions. In some dialects, the word *ʾusṭā* (a corrupt form for *ʾustāḏ* in which /t/ has been replaced by emphatic /ṭ/) is used as a form of polite address to a craftsman, etc. Other Turkish words are occasionally used, such as *bāšā* 'Pasha', *bey* 'Bey' (usually spelled *bek*), and *šāwīš* 'Sergeant', as well as the Arabic words *(yā) muʿallim* 'Teacher' and *yā bā* 'oh, Father', *yā šēx* 'oh, Sheik', and even *yā ḥājj* 'oh, Pilgrim'. All these are used in some regions as forms of polite address, showing affection, although they may also be used ironically.

It is also a matter of good manners for users of Arabic to use various eulogies when referring to certain personalities, such as the Prophet, adding after his name or his titles the wording *ṣallā llāhu ʿalayhi wa-sallam* 'may God bless him and grant him salvation!'. When referring to one of the first four caliphs (*al-xulafāʾ ar-rāšidūn*), one should add *raḍiya llāhu ʿanhu* 'may God be pleased with him!', while in the case of the other prophets mentioned in the *Qurʾān*, such as Moses, one should say *ʿalayhi s-salām* 'peace be upon him!'. Polite reference to the *Qurʾān* is *al-qurʾān al-karīm* and to the *Ḥadīṯ al-ḥadīṯ aš-šarīf* or *an-nabawī*.

Polite language involves the usage of various expressions at different situations and occasions, as postulated by the Arabic dictum *li-kull maqām maqāl* 'for every situation there is an appropriate saying' (→ greetings). Greeting a person is usually performed by using the formula *as-salāmu ʿalaykum*, or *marḥaba*, *ʾahlan*, or *ʾahlēn*, while the replies are *ʿalaykum as-salām, ʾahlan fīk, ʾahlan wa-sahlan*, or *marḥabtēn*. Customary formulas of inquiring about a person's well-being (English equivalent *how are you?*) are in Iraq and Syria *ʾāš lōnak* or *šlōnak* lit. 'what is your color?', in southern Syria, Lebanon, Jordan, Palestine, and the Arabian Peninsula *kīf ḥālak*, in Egypt *ʾizzayyak*, and in North Africa *lā bās?* The reply is usually *al-ḥamdu lillāh* 'praise be to God!', and in North Africa also *lā bās!*.

Other formulas of greetings associated with times or festivals are the morning greetings *ṣabāḥ al-xēr* 'morning of goodness', to which the typical answer is *ṣabāḥ an-nūr* 'morning of light'. For late morning, the formula *nahārak* 'have a happy day!' is used, to which the answer is *wa-nahārak (saʿīd u-mubārak)* 'and may your day (be happy and blessed)!'. For late afternoon or evening, the common formula is *masāʾ al-xēr* 'evening of goodness', to which the reply is *masāʾ an-nūr* 'evening of light'. For 'good night', the formula is *tiṣbaḥ ʿa-lxēr* 'wake up with goodness!', to which the reply is *wi-ʾinta min ʾahlo* 'and you too'. The common formulas during festivals are *kull sana wa-ʾanta bi-xēr* or *kull ʿām wa-ʾanta ṭayyib* 'be well every year!'. During the month of Ramadan, the common greeting is *ramaḍān mubārak* 'blessed Ramadan!', while at any other festival the common formula is *ʿīd mubārak* 'blessed festival'.

For an engagement, wedding, birth, graduation, or any new occasion, one is expected to use the felicitation *mabrūk* 'may it be blessed!', to which the answer is *Allāh yubārik fīk* 'may God bless you!'. A special formula to wish a newlywed couple well is *bi-r-rifāʾ wa-l-banīn* 'live in harmony and beget sons!'. When someone has taken a shower or has just had a haircut, it is customary to 'congratulate' the person with the formula *naʿīman* 'with grace', to which the reply is *Allāh yinʿam ʿalēk* 'may God bestow His grace upon you!'.

A person who sneezes should be 'blessed' with the formula *raḥimaka llāh* 'may God have mercy on you!', to which the reply is *turḥam* 'may you have God's mercy too!'. Welcome greetings are various, e.g. *ʾahlan wa-sahlan*, lit. 'you are part of the family and an easy guest', or *šarrifnā* 'honor us!', to which the reply is *ʾahlan fīkum* or *tašarrafnā*.

A polite way to express an apology for bothering a person is *ġallabnākum/taʿʿabnākum* or *bi-dūn taklīf/ʾizʿāj/ġalabe*, to which the common reply is *lā wallā* 'by God, not at all!', or even *ʾahlan wa-sahlan* or *taʿabkum rāḥa* 'the inconvenience you cause is easy', i.e. 'it's a pleasure'.

Coffee is not only a popular drink but also an institution which involves some formulas of politeness. When finishing drinking the cup, one should express one's satisfaction by using the word *dāʾiman* 'always', i.e. 'may coffee always be in this house!', while the host's answer is *bi-al-hanāʾ* 'with pleasure' or *bi-l-hanāʾ wa-š-šifāʾ* (colloquial *bi-l-hina wa-š-šifa*) 'with pleasure and health'. However, when one has coffee in the house of a bereaved family, the formula used is *ʿāmir* 'may this house be full', i.e. 'may no more people be missed!'.

A polite invitation to dinner is simply *tafaḍḍal* or *šarrifnā* 'honor us!'. A very peculiar formula of welcoming a person who arrives just in time for a meal is *ḥamātak bitḥibbak* lit. 'your mother-in-law loves you'. For 'bon appétit' the expressions *šahiyya ṭayyiba/hanīʾan* or *hanīʾan marīʾan/bi-l-hanāʾ* or the colloquialism *ṣaḥtēn* are customarily used, to which the host's reply is *ṣaḥḥa/ṣaḥḥa wa-ʿāfiya/ʿalā qalbak* 'with health'. At the end of a meal, it is customary to thank the host with the formula *as-sufra dāʾima* 'may this table always be full!' or even *bi-al-ʾafrāḥ* 'on happy occasions', to which the host's reply is *dumt/tadūm/ʿišt/taʿīš* 'may you live long!'. Another formula is *yusallim yadayk* 'may God bless your hands!', and if one wishes to thank the lady of the house in her absence, *yusallim yaday ʾumm*…, followed by the name of her son or daughter or generally *al-ʾawlād* 'the children'. In some circles, the formula used is *yusallim yaday al-madām*, to which the common reply is *tislam* 'be healthy!'.

One wishes success or good luck by the words *bi-n-najāḥ/bi-t-tawfīq/muwaffaq* 'with success!' or *Allāh yiʿṭīk il-ʿāfiya* 'may God give you health!'. 'Bravo!', 'well done!', etc. are expressed in Arabic *ʿafārim ʿalēk* (from Turkish), *ʾaḥsant, marḥa marḥa, bax bax*, and *mā šāʾa llāh* 'God wish'. The last expression is also used for expressing admiration or rapture, or when referring to successful children, a beautiful house, a nice car, and the like.

For wishing someone a safe journey one should say *bi-s-salāma, maʿa s-salāma* 'with safety', or *ʿalā ṭ-ṭāʾir al-maymūn* 'on the lucky bird'. Wishing someone a speedy recovery is expressed by the formula *salāmtak*, while satisfaction with someone's recovery is expressed by the formula *rāḥ aš-šarr* 'evil has gone by', or simply by *al-ḥamdu lillāh* 'praise be to God!', which is widely used as a general remark showing satisfaction and contentment or submission to God's wish.

Polite formulas expressing a request or intention to leave a party, house, etc. are *bi-xāṭirkum* or *ʿan ʾiḏn* 'with your permission', to which the common reply is *maʿa s-salāma* 'in good health'. The expression *ʿan ʾiḏn* is also used to interrupt politely a conversation or a speaker.

A polite formula to indicate one's positive response to a request or one's preparedness to fulfill a wish is *ḥāḍir* 'at your service'; other formulas are *ʿalā r-rās wa-l-ʿēn* or *ʿalā ʿēnī wa-rāsī* lit. 'on my eye and head', i.e. 'willingly'. The Classical Arabic formula was *samʿan wa-ṭāʿatan* lit. 'hearing and obeying', i.e. 'your wish is my command'.

A person's reaction to a compliment is usually *min luṭfak* 'you are very kind', while a polite reaction to hearing about the death of a person is the expression *ad-dāʾim Allāh* 'God is the only one who exists forever'. One may also say on such an occasion *al-ḥamdu lillāh* or *subḥāna Allāh* 'may God be blessed!', *lā ḥawla wa-lā quwwa ʾillā bi-llāh* 'no might and no power but with God', or *al-baqāʾ fī ḥayātkum* 'his memory will remain with you'. A polite reference to a dead person is *al-marḥūm/al-marḥūma* or *raḥimahu/raḥimahā llāh* or *Allāh yarḥamuhu/yarḥamuhā* 'God have mercy upon him/her'.

Polite expressions of sorrow or regret are *ʾāsif/mutaʾassif* 'sorry!', *(bi-)lā muʾāxaḏa* 'pardon me; no offense', while an expression of sympathy is *Allāh yisāʿdak* 'may God help you!' or *Allāh yifrijha ʿaleyk* 'may God give you salvation!'. When refusing to give a beggar alms, the polite way to refuse is *Allāh yiʿṭīk* 'may God give you!'.

However, politeness also involves 'refined' style, which is usually used in order to avoid offense or direct criticism. This can be achieved by circumlocutions, the use of antonyms, a witty remark, → euphemisms, the use of → proverbs, and the like. Thus, in a situation in which one disagrees with someone else's view, there are various options for expressing one's reservations, from a blunt objection to an indirect reservation. When disagreeing with a statement, such as *ʾuxtuhu jamīla* 'his sister is beautiful', one may politely react by using the word *yaʿnī* lit. 'this means', i.e. 'it's a question of taste', with 'lengthened' intonation, indicating a reservation.

The litotes *lā baʾs* 'not bad' may also be used in order to avoid a blunt negative reaction. The compliment *taqaddum la baʾs bihi* 'quite a good progress' may sound as truly satisfying progress but may also indicate a more polite criticism, implying that more could have been done or achieved.

Euphemism may also be used as a polite way of expression instead of a blunt or strong language, obscenities, or direct reference to sex (→ taboo), e.g. *baṣīr* 'one who is endowed with eyesight' instead of 'a blind person'; *hāḏa kalām kaḏib* 'this is a lie' may be 'mitigated' by *kalām lā ʾasās lahu min aṣ-ṣiḥḥa* or by using a Qurʾānic quotation (Q. 53/23): *mā ʾanzala llāhu bihā min sulṭān* 'for which God has not sent any authorization', i.e. 'baseless words'.

Euphemisms for the toilet are *ḥammām* 'bathroom' or *bayt ar-rāḥa* lit. 'house of rest', *bayt al-ʾadab* lit. 'house of polite upbringing', *bayt al-māʾ* lit. 'house of water', and also *tuwalēt*. The expression *al-hawā al-ʾaḥmar* 'red love/desire' may be used instead of *muḍājaʿa* 'sexual intercourse'.

Reference to dogs, donkeys, shoes, or toilet in conversation, even in passing, requires from the user the addition of a polite parenthetical remark such as *baʿīd ʿannak* 'far from you! or *ʾajallak* 'you are more exalted', suggesting that the hearer is highly regarded by the speaker above and beyond these 'degraded' objects, while a reference to a dead person or an evil man requires the addition of the word *al-baʿīd* '[who is] far from you'.

A polite rejection of an invitation or offering of help, especially when the person inviting or offering his assistance has already been generous many times, may be expressed by the proverb *law kān ṣadīqak min ʿasal lā talḥaso kullo* 'if your friend is made of honey, do not lick all of it'.

Bibliographical references

Brown, Penelope and Stephen D. Levinson. 1987. *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.

El-Sayed, A. 1990. "Politeness formulas in English and Arabic: A contrastive study". *Review of Applied Linguistics* 19.1–23.

El-Shalny, Hasan M. 1995. *Folk traditions of the Arab world: A guide to motif classification*. 2 vols. Bloomington and Indianapolis: Indiana University Press.

Emery, Peter G. 2000. "Greeting, congratulating and commiserating in Omani Arabic". *Language, Culture and Curriculum* 13:2.196–216.

Ferguson, Charles A. 1976. "Structure and use of politeness formulas". *Language and Society* 5.129–136.

Hinds, Martin and El-Said Badawi. 1986. *A dictionary of Egyptian Arabic*. Beirut: Librairie du Liban.

Piamenta, Moshe. 1979. *Islam in everyday Arabic speech*. Leiden: E.J. Brill.

——. 1983. *The Muslim conception of God and human welfare as reflected in everyday Arabic speech*. Leiden: E.J. Brill.

Shivtiel, Avihai. 2005. "*Takālīd*". *Encyclopaedia of Islam*, Suppl., 774–779. 2nd ed. Leiden: E.J. Brill.

Wehr, Hans. 1979. *A dictionary of Modern Written Arabic*, trans. J. Milton Cowan. 4th ed. Wiesbaden: O. Harrassowitz.

Avihai Shivtiel (Leeds)

Political Discourse and Language

This entry on Arabic and political discourse looks at language → variation from a sociolinguistic perspective, together with aspects of rhetoric. It presents a number of factors that characterize Arabic political speeches, including contextual factors, → register, → code-switching, and discourse strategies, to show how language forms relate to language functions. Such factors are illustrated through selected examples of linguistic combinatorial rules (within and between dialects) that accompany language levels, and through examples of rhetorical tactics.

1. Contextual factors

Sociolinguistic studies tend to start from quantitatively based analyses of language and move from the linguistic microlevel to the macrolevel (such as communication strategies and messages conveyed by the speaker). This entry, however, presents the macrolevel framework followed by a few examples of marked linguistic occurrences.

An analysis of political speeches could take into account the political context at the time of delivery and immediate contextual factors: audience, topic of discourse, context, place of delivery, and type of discourse – speech with no interruptions, debate, interview, a planned speech read out loud, extemporaneous discourse, etc. The analysis could also be based on the speaker's background (e.g. place of birth, education, religion), all the elements which would influence the speaker's discourse. One could also look at a community's attitudes toward a language variety and the 'values' attached to a particular dialect to understand the shifts between levels along the stylistic continuum, according to the communicative purposes.

2. Speech communities and individual speakers

Analysis of political discourse shares a number of common points with the study of speech communities: although individuals have linguistic idiosyncrasies, a group of people from the same society will share similar linguistic rules and patterns of use. Recurrent patterns can be found in the speech of individuals which reflect the speech of that community. Communities can be defined according to a number of criteria, and the number and relevance of these criteria differ from one community to another. The speech of a community can be influenced by several interacting social variables. Social variables such as education, gender, age, religion, occupation, social class, and regional factors are considered to be factors that correlate with differences in vocabulary, morphology, and phonology.

3. Arabic political discourse

The analysis of Arabic political discourse is at the crossroads of different disciplines, including pure dialectology, political science, anthropological linguistics, discourse analysis, and rhetoric. In fact, a number of studies analyzing political discourse combine various areas of study: Bengio (1998) bases her study of Saddam Hussein's discourse on the meaning of language, political culture, myth making, and symbolism, while integrating the wider context of political and historical developments in the country. Suleiman (2003), in his study on the Arabic language and national identity, aims at an understanding of nationalism in its language-related dimension (→ nationalism and language).

Studies of Arabic political discourse from a sociolinguistic perspective are, however, rather scarce. Holes (1993) looks at the use of variation in speeches from Nasser. Mazraani (1995, 1997) analyzes speeches from Nasser, Saddam Hussein, and Gaddafi and presents hypotheses as to the correlations between variation in Arabic and functions of the discourse. Taine-Cheikh (2002) looks at sermons from a well-known preacher in Mauritania: instead of the Modern Standard Arabic expected in such a setting, the author finds a mixed form of language, *arabe médian*, between Modern Standard Arabic and Ḥassāniyya dialect. This corresponds to one of the language functions noted in previous research (see Holes 1993; Mazraani 1995, 1997): intimacy is conveyed by the speaker through the dialectal variety – in this case Ḥassāniyya – whereas Modern Standard Arabic is used on more solemn occasions.

The correlation between language form and function is valid and perceptible, not only in the political arena but also in the areas of discourse and conversation in general: whether the intention is to give authority to a conversational point or to explain a concept, linguistic and extralinguistic features vary accordingly. Whatever similarities are revealed about the ways political leaders make use of level variation as a rhetorical strategy are very likely to reflect the 'rules of use' of the wider communities in which political leaders function, whether local or pan-Arab.

4. Register and code-switching

Politicians in public discourse tend to use Modern Standard Arabic linguistic forms as well as dialectal and mixed forms. Switches between passages in Modern Standard Arabic or dialectal or mixed levels can be described as changes in register, which is determined by three situational variables, referred to as field, tenor, and mode (what is going on, who is taking part, and what role the language is playing), and three semantic components, which are the ideational, interpersonal, and textual aspects (Halliday 1978).

Following the various developments on → diglossia (see Ferguson 1959, 1996), authors have analyzed the phenomenon of code-switching in considering variation between Arabic and European languages (Heath 1989; Auer 1999; Rouchdy 2002) or between different Arabic standards or varieties (Eid 1988; Holes 1995a). Code-switching is a recognition of linguistic reality and the complex relationship between different codes or registers. It relies on the meaningful juxtaposition of what speakers must consciously or subconsciously process as strings formed according to the internal rules of two distinct grammatical systems (Gumperz 1982:66) situated at each end of a continuum. In Eid's (1988) study, for instance, code-switching is analyzed mechanically, i.e. to see where and how it occurs, and 'switch positions' are determined by syntactic conditions or word order requirements. In conversations, code-switching emerges as one way of speaking among many others, for example formality vs. informality, politeness vs. impoliteness, sincerity vs. insincerity (Suleiman 1999).

5. DISCOURSE STRATEGIES

One can observe a number of functions and strategies directly associated with a particular language level. Abd El Jawad (1986:24) claims that politicians use code-switching to gain more favorable attitudes, fuller understanding, and more emotional support from their audience. Code-switching and style shifting serve functions such as group identification, solidarity, and intimacy, to signal unity and integration. Such messages are accompanied by linguistic realizations which vary according to the type of message presented; distancing, for instance, may be expressed through Modern Standard Arabic, while establishing solidarity or explaining an intricate concept may be uttered through the dialect.

Discourse-organizing strategies include storytelling, time-change markers, and reported speech. These devices are invariably accompanied by a code-switch from Modern Standard Arabic to dialect and a change in prosody; a rapid tempo, a conversation-like contour, for instance, will be the first indication to the audience that a change is occurring at other levels as well (a change at the linguistic level, or a change in the subject matter or the function of the language). Words such as *ʾinnaharda* 'today' and *dilwaʾti* 'now' in Cairene Arabic, produced at the beginning of a passage with a rising intonation and followed by a pause, are time-change markers (van Dijk 1981:181), indicative of a change occurring at different levels – changes in person, time, and subject matter. Reported speech (→ indirect speech) or constructed dialogue is a way of making the discourse livelier, more vivid, by imitating an everyday chat.

Speed of delivery, rhythm and intonation, and conversational vs. oratorical delivery will reinforce the audience's perception and classification of elements into various levels of formality. In cases of code-switching, such factors will indicate to the audience that a change has occurred and catch their attention.

6. LANGUAGE FORMS AND FUNCTIONS

The following illustrates the correlation that has been observed between the type of message the politician is conveying, the language level, and the linguistic realizations that are present in discourse.

Modern Standard Arabic is used whenever the speaker is constructing an abstract argument, recalling historical events, or expounding on new political ideas and axioms. General and abstract concepts are presented as if they were unquestionable text, as opposed to exegesis (Holes 1993). As speakers distance themselves from the audience, they tend to depersonalize the discourse; when they are instructing the audience, the tone is authoritative, and Modern Standard Arabic is used. The discourse is then characterized by abstractness, compactness, formality, detachment, planning, and integration, to borrow some of the terms used by Chafe (1982), Ochs (1979), and Tannen (1982).

In a sermon delivered in a mosque, for instance (Mawlāya Muḥammad, 26 May 1978; see Mazraani 1997), Gaddafi repeats some verses and extracts from the *Qurʾān* with various explanations. He distances himself from the audience, using an authoritative tone to guide the public into understanding the text. The sermon shows the discrepancy between, on the one hand, text, i.e. passages from the *Qurʾān*, which are fixed, unalterable, and in Classical Arabic, and, on the other hand, explanation of such passages, in dialect, and explanations that are given 'in other words' or at different levels of formality.

General analysis has shown that Modern Standard Arabic passages are characterized by:

i. Modern Standard Arabic phonemes such as /q/, /ṯ/, /ḏ/, /ḍ/.
ii. Modern Standard Arabic morphology of verbs, passive verbs, complementizer *ʾan* and subjunctive, *sa-/sawfa* followed by the verb to indicate future tense, and *ʾiʿrāb*, which could be seen as a Modern Standard Arabic extra marker since it adds a sense of formality to the discourse.
iii. Long sentences with subordinate clauses, long noun phrases, *maṣdar*s, word order (verb and subject), and Modern Standard Arabic negative particles.
iv. Abstract vocabulary (journalistic, political), frozen collocations, high lexical density (many content words), and lexical choices triggering Modern Standard Arabic noun-adjective agreement patterns.

Modern Standard Arabic passages tend to be accompanied by slow delivery; falling tonal nuclei and long pauses are employed to give full emphasis to some words and to signal the weight of the message.

On the other hand, dialect tends to be used when speakers explain and specify their political program, their conversations with various leaders, or their personal experience. Abstract concepts (in Modern Standard Arabic) are repeated and paraphrased in the dialect to clarify the message. This includes establishing a more personal rapport with the audience, and the discourse is characterized by speaker involvement, apparent spontaneity, and a greater degree of personalization; in general, these sequences sound unplanned and more natural. Typically, such moves involve a switch to dialect where the speaker wishes to establish a sense of solidarity or intimacy and get the audience involved. The tone is more familiar and 'we', 'you', and 'I' pronouns are used to create a sense of communion, a 'one-of-us' feeling of the speaker with the audience. The dialect is used in asides and narrative accounts and in recounting anecdotes. Interpersonal involvement is established through rhetorical questions, discourse organizing, time-change markers, reported speech, and storytelling with specific examples. The tone in such passages can display argument, emotion, and commitment. Specific examples can be used to elicit deep feelings of sympathy or pity from the audience and to unite behind the speaker's views.

Textual → cohesion is achieved in dialectal passages through imitation of an ordinary, conversational, dialectal prosody, with a variable pitch and a faster tempo. The dialectal intonation adds a feeling of spontaneity to the speech and gives the impression of sincerity. At times, an argumentative tone of voice or one of protest is used to display publicly the speaker's commitment and to unify the audience in its indignation at foreign criticism. A more friendly and conversational style is used to trigger solidarity and communion.

Such messages are characterized textually by the co-occurrence of the following dialectal elements. Phonologically, plosive /q/ and interdentals yield the most interesting results (in the dialectal data observed so far, through speeches of Saddam Hussein, Nasser, and Gaddafi).

i. In Cairene Arabic, /ʾ/ from Old Arabic /q/; /s/ and /t/ from /ṯ/; /z/ and /d/ from /ḏ/; and emphatic dental sibilant /ẓ/ and emphatic dental plosive /ḍ/ from the emphatic interdental sibilant /ḏ̣/;
ii. In Baghdadi Arabic, /g/ from Old Arabic /q/ (and a few cases of /č/ < Old Arabic /k/);
iii. In Tripoli Arabic, /g/ from Old Arabic /q/, /t/ < Old Arabic /ṯ/.

Morphologically, dialectal passages are characterized by preformative *yi-*, morphology of verbs, prefixes *bi-* and *ḥa-* followed by verbs, deletion of *hamza* (glottal stop), deletion of vowel in unstressed syllables, *-u*/*-ū* 3rd person plural ending instead of *-ūna*, and absence of *ʾiʿrāb*. Although verbs vary to a larger degree since they are formed of a greater number of components which are susceptible to variability, morphological variation is more limited in nouns and revolves around reflexes (*qimma* ~ *ʾimma*), or feminine endings of noun-pronoun, as in the case of *quwwatuhu* (Modern Standard Arabic), *quwwathu* (hybrid), and *quwwithu* (dialect) 'its force', as in the Libyan dialect.

Syntactically, Modern Standard Arabic passages are characterized by asyndetic verb strings, short sentences (hardly any subordinate clauses), subject-verb word order, colloquial negative, and interrogative particles.

Dialectal lexicon consists of everyday dialectal vocabulary, an absence of abstract nouns, and low lexical density (fewer content words).

A third level is also found which is neither Modern Standard Arabic nor a dialect but rather an intermediate combination of the two to some degree, depending on subject matter and communicative aims. It seems to be produced by Modern Standard Arabic and dialect combinatorial rules which limit particular rule combinations, although the resulting intermediate form of language cannot, yet, be defined comprehensively. In each country, it should be seen as part of a continuum with one end closer to Modern Standard Arabic and the other end closer to the local dialect. This third level is associated with relatively formal settings, e.g. political speeches and university lectures, but tempered by overriding contextual and interpersonal factors which tend to pull the language in the direction of the colloquial. Interestingly, it is by means of this third level that much of the purely Modern Standard Arabic discourse is organized, and through which involvement rather than detachment is signaled.

Items which are common to both Modern Standard Arabic and dialect often belong lexically to Modern Standard Arabic while their morphological realizations can be dialectal. Taine-Cheikh (2002) associates this intermediate level to a neutral code with hybrid forms.

This intermediate level also produces symbiotic and hybrid forms, i.e. combinations of Modern Standard Arabic and dialectal elements. Symbiotic verbs consist of aspectual *bi-* prefixed to a verb with a shared lexicosemantic status and a close-to-Modern Standard Arabic morphology, such as *binaḥtafil* 'we are celebrating' in Cairene Arabic and *bitudīru* 'they are managing' in Tripoli Arabic (Baghdadi Arabic does not have prefix *bi-*). Hybrid verbs consist generally of a dialectal element within a 'usually' Modern Standard Arabic structure. Examples, such as Tripoli Arabic resyllabication co-occurring with Modern Standard Arabic phoneme /q/ and Modern Standard Arabic morphology as in *taqublu* 'you [pl.] accept', between Modern Standard Arabic *taqbalūna* and Tripoli Arabic *tigiblu*. Hybrid forms are interesting as they show the persistence of some dialectal elements into higher levels of formality.

Besides Modern Standard Arabic and dialectal passages, further investigation is needed of this intermediate level which displays examples of combinatorial rules observed in the mixing of dialectal and Modern Standard Arabic features and which is in fact representative of what happens in the speech community (→ Educated Arabic).

7. COMBINATORIAL RULES

Each dialect has its own rules which differ from Modern Standard Arabic and from those of other countries. For example, in the Iraqi and Libyan dialects, we find cases of resyllabification, as in *yiḥimlūn* (Modern Standard Arabic *yaḥmilūna* 'they carry'), but not in Cairene; in Baghdadi Arabic, deletion and prosthetization of vowels, as in *ʾinḥawwal* 'we change'; and in Cairene and Baghdadi Arabic, deletion of unstressed syllables, as in *yiṣawwar* 'he imagines'. The aspectual non-past prefix *bi-* is found in Cairene and in Tripoli Arabic *binasmaʿ* 'we listen', but not in Baghdadi Arabic. As to verb endings and *ʾiʿrāb*, for the plural of the non-past tense form, Baghdadi has *-ūn* while Cairene and Tripoli Arabic have *-ū*. These dialectal endings have infiltrated so much into various levels of formality that they have become acceptable with Modern Standard Arabic forms. Besides the dialectal differences, discourse analysis offers grounds for investigating combinatorial rules which are set by the communicative competence of a speaker and belong to a speech community.

The communicative competence of a speaker consists of grammatical knowledge, sociocultural adaptability (i.e. knowing what level to use in what context), and individual skills of the speaker to use such knowledge to serve communicative purposes. Speakers have a knowledge of Modern Standard Arabic rules (which are formalized and consciously learned) and of dialectal rules (which are unconsciously controlled and not formalized). They also have a tacit knowledge of mixed, hybrid forms (which are nowhere formalized as such) and of the combinatorial rules as to how to combine elements from the same or different levels.

The paragraphs below use examples from the lexicon and from the phonology to illustrate multivalency and acceptability of some marked items in contrasting levels (reflexes /g/ or /ʾ/ in Modern Standard Arabic passages). Phonological realizations are usually less amenable to variation, and speakers normally retain phonological characteristics of their own native dialect even after being exposed to another language variety. Phonemes fill the condition of being nonreferential and of carrying no meaning in the language (Lavandera 1978).

In the case of individual phonological segments, there is generally a substantial difference between the Modern Standard Arabic realization and the dialectal one (see also Hary's [1996] study of variables along the language continuum). Two phenomena occur: either the dialectal realization is scarcely used at all in combination with Modern Standard Arabic elements, or there is an emergence of a third phoneme, intermediate between Modern Standard Arabic and the dialect, as in the case of interdentals. Reflexes of the Old Arabic interdental /ṯ/ in Cairene Arabic and in Tripoli Arabic, and of /ḏ/ and /ḏ̣/ in Egyptian, show the emergence of a third, intermediate phoneme: /s/ between

Modern Standard Arabic /ṯ/ and the dialectal /t/, e.g. *sawra* 'revolution' in Cairene Arabic and *mabsūs* 'scattered' in Tripoli Arabic); /z/ between Modern Standard Arabic /ḏ/ and dialectal /d/; and /ẓ/ (emphatic dental sibilant) between Modern Standard Arabic /ḏ̣/ (emphatic interdental sibilant) and dialectal /ḍ/ (emphatic dental plosive). The sibilants represent a slight downgrading from Modern Standard Arabic and are found occurring in mixed and what are otherwise purely Modern Standard Arabic passages.

In research on speeches in Baghdadi, Cairene, and Tripoli Arabic (Mazraani 1997), in the case of Old Arabic /q/, there was no third phoneme intermediate between Modern Standard Arabic and the dialect. The Baghdadi and Tripoli Arabic dialects very rarely have /g/, the dialectal realization of Old Arabic /q/, co-occurring with Modern Standard Arabic features, apart from the case of the item *gāl* 'he said', which is so widespread that it is accepted in more formal levels and in passages of reported speech. The reason for the scarcity of /g/ in more formal levels is that /g/ (as well as Baghdadi /č/ from Old Arabic /k/) is felt to be too local and dialectal and hence inappropriate for public speaking (see also Holes 1995a for a discussion on variation in intercommunal situations, of *gelet* and *qəltu* dialects in → Iraq).

Cairene Arabic on the other hand shows cases of /ʾ/ from Old Arabic /q/ co-occurring with Modern Standard Arabic elements. Reflex /ʾ/ is found in items such as *ḥaʾʾ* 'right', *waʾt* 'time', *mustaʾbal* 'future', *ʾāl* 'he said', and *ʾām* 'he stood up'. The higher frequency of Cairene /ʾ/ compared to Baghdadi and Tripoli /g/ might be explained by the fact that /ʾ/ has acquired pan-Arabic acceptability given the pervasiveness of Cairene Arabic through the media.

Phoneme /q/ in Modern Standard Arabic items is found overwhelmingly in public discourse. These items tend to be associated with the abstract, formal political vocabulary, but, because of their widespread use and the subject matter (a press conference, a political speech, a sermon in a mosque, etc.), they can also be found in dialectal passages of the speech, in country names, or in collocations, for instance. Modern Standard Arabic phoneme /q/ is invariably retained, and the whole item does not show any leveling to the dialect apart from the absence of *ʾiʿrāb*.

Political speeches sometimes offer us the possibility of comparing written and oral versions of the same discourse. What is 'accepted' in speech may sound too dialectal in print, and the written version is thus leveled up. This is seen in one of Nasser's speeches which has both written and audio versions (*Ḥadīṯ al-baṭal az-zaʿīm Gamāl ʿAbd an-Nāṣir ʾilā l-ʾumma* 'Speech of the leader Gamal Abdu Nasser to the nation', 1970). Some elements which clearly belong to the dialect were adapted: they were either replaced by items with a higher lexicosemantic status or are simply given the Modern Standard Arabic morphology. The following are some examples. In each case the audio version is given first, then the version in the written text:

ʿalašān	*likay*	'in order to'
ʾinnaharda	*al-yawm*	'today'

Dialectal prefix *bi-* is dropped in the written version, as in:

baʾullukum	*ʾaqūlu lakum*	'I say to you'

In some cases, a syntactic adjustment is made: verb strings are replaced by a verb and complementizer *ʾan* followed by a verb, as in:

ʾaʾdar ʾaʾūl	*ʾastaṭīʿ ʾan ʾaqūl*	'I can say'

Another adjustment is that negative particles in dialect are replaced by their equivalents in Modern Standard Arabic:

ma fakkarnāš	*lam nufakkir*	'we did not think'

Another interesting lexical aspect in political speeches is the recurrence of → collocations, i.e. fixed expressions of a political, journalistic nature. These include such cases as *quwwāt ʿarabiyya* 'Arab forces' and *al-qiyāda al-ʿāmma* 'the general leadership'. These collocations are found in the oral and written modes and do not necessarily have a formalizing effect on the passage in which they occur, given the widespread use of these terms in news bulletins and potentially in everyday conversation.

Items such as *ṯawra* 'revolution', *quwwāt* 'forces', *qimma* 'summit', and *qāʾid* 'leader', which are part of political terminology, display Modern Standard Arabic characteristics such as phonemes /q/ and /ṯ/, and occur in Modern

Standard Arabic, mixed, and dialectal passages. The word *ʾimma* (with /ʾ/ instead of /q/) exists as well in Egyptian, but with the meaning of ‘top’. Similarly, *quwwāt* ‘forces’ as part of a military collocation always has /q/, e.g. *quwwāt musallaḥa* ‘armed forces’ while *ʾiwwa* has the meaning of ‘strength’ in nonmilitary contexts and is found in nonliterary idioms.

Besides the concern with lexical realizations, some authors insist on the symbolic value attached to words such as *ṯawra* ‘revolution’, *ʾintifāḍa* ‘uprising’. In Baathist terminology, *ṯawra* “came to denote a value concept of the highest order, of almost sacred significance” (Bengio 1998:25). Suleiman (2003:3) wishes sociolinguistics would take into account the symbolic connotations of the language “because of the centrality of language in articulations of nationalism in the Arab Middle East”.

8. RHETORICAL TACTICS

Traditional Arabic public-speaking strategies, with features such as repetition, assonance, and → paronomasia, add an emotional dimension to the discourse. They are a way of fixing key elements onto the audience’s mind. They keep the attention of the listener and are highly appreciated in Arab culture; devices such as parallelism and repetition are ingrained in Arabic discourse (Johnstone 1991). Rhetorical ornamentation then induces involvement and emotion by arousing the audience’s feelings and poetic senses.

Rhetorical tactics in Arabic include devices found in other languages (Atkinson 1984). Such devices include not only the listing of three elements, contrasting pairs, and repetition, but also imagery and detail, and constructed dialogue with reported speech and dialogue. Organizing devices of this type maintain attention and often trigger applause, or at least positive audience feedback. Some of these devices are illustrated below.

i. Listing three elements. Arabic political discourse provides numerous examples of listing three elements. The device brings emotion to the discourse through the repetition and symmetry of the structure. The following example is from a recording of one of Nasser’s speeches (23 December 1957):

naḥtafil bi-ʿīd in-naṣr ʿalā siyāsati l-quwwa wa-ʿalā siyāsati l-ʿudwān wa-ʿalā siyāsati l-ġadr

‘We celebrate the anniversary of our victory over the policy of force, over the policy of aggression, and over the policy of treachery’

ii. Repetition of words or clauses. Repetition has been discussed by several authors. Holes (1995b) looks at the structure and function of parallelism and repetition in spoken Arabic, and Boumans (2002) looks at aspects of repetition in code-switching. Repetition of an item highlights it and helps its comprehension by providing textual redundancy and creating a less dense discourse. Rhythmically, repetition provides a musical aspect to language and may create a rhetorical crescendo and captures audience attention. Patterned repetition is a technique of persuasion. According to Koch (1983), the grammatical structure of Arabic makes repetition a strategy available especially to Arabic speakers and is the key to linguistic cohesion of many Arabic texts, and to understanding their rhetorical effectiveness.

iii. Imagery, memory, and details. Details make a concept understandable and the issue more explicit. Here is a short extract from an audio recording of one of Nasser’s speeches (*Xiṭba ʿAbd an-Nāṣir fī ʿīd iṯ-ṯawra* ‘Nasser’s speech on the celebration day of the revolution’, 23 July 1962):

ʾana ʾazkur min ʾawwil / ʾayyām iṭ-ṭufūla / kullama nikkallim fi ʾayyi ḥāga / yiʾullak bititkallim fi ʾēh Saʿd Bāša ʾāl mafīš fayda / fi hāza l-kalām

‘As I recall in my childhood days, whenever we were talking about any topic, they would say to you, what are you saying, Saad Pasha said that there is no use in that kind of talk’

Here, as the speaker recounts his personal memories, the level drops to the dialect: dialectal elements co-occur with the use of /z/ instead of /ḏ/ (as in *hāza* ‘this’), /ʾ/ instead of /q/ (as in *yiʾullak* ‘he tells you/he says’), vowel patterning and aspectual prefix *bi-* (*bititkallim* 'you talk'), discontinuous negative particle *ma- -š* (as in *mafīš*).

iv. Direct and reported speech. Speech contributes to making the discourse livelier and more conversational. The following is an extract from a press conference that Saddam Hussein held in Baghdad on 20 July 1980 (*Ḥadīṯ fī l-qaḍāya l-ʿarabiyya wa-d-dawliyya* 'Discussion on Arab and world issues'):

> *liʾannu bi-s-sābiq kānu yigulūn innu niḥin šū nigdar insawwi / wa tnāqašna / tnāqašna maʿa baʿḍ il-ʿarab / bimašākil mubāšir / gulnā l-kum lā tijibūn ilna l-ʾamrikān tunṭūnhum il-qawāʿid il-ʾajnabiyya*
>
> 'In the past, they used to say, what can we do? and we discussed urgent matters with some Arab [leaders]: we said to you: Don't bring here Americans and provide them with foreign bases'

This passage shows Baghdadi features such as /g/ from /q/ (*yigulūn*, *nigdar*), morphology of the pronoun *niḥin* for Modern Standard Arabic *naḥnu*, localism (*tunṭūn*), Baghdadi morphology of verb (*yigulūn*), verb string (*nigdar insawwi*), and lexical item *šū*.

The persuasive strategies used in political discourse are merely a particular subset of the common strategies used by speakers in everyday interactions. Political and conversational discourses are closely linked and interrelated: interlocutors who wish to communicate with one another use numerous speech functions: they persuade, criticize, show anger, tease, justify their actions, inform and advise, tell jokes, etc. A political speech is an act of persuasion of an audience, and, like a conversationalist, the politician mobilizes a certain array of communicative skills to get his message across and influence his listeners. Keeping the audience's attention is one motivation for varying rhetorical tactics, which results in → code-switching. Thus, Modern Standard Arabic and the dialect are used as the means for particular types of communication. Once politicians have finished appealing to the emotions of the audience and established solidarity with them through the dialect, they revert to speech functions associated to Modern Standard Arabic, so as to reestablish their authority and regain formality. The main motive behind this recurrent code-switching is to keep the audience's attention. This is an essential strategy in large gatherings and in long speeches, where the audience's attention can wander. A speech in Modern Standard Arabic requires a lot of concentration on the part of the public, for whom Modern Standard Arabic is not the native language and may sound monotonous. A speech in the dialect, on the other hand, would not fulfill the criteria for a 'serious' speech and would lose the politicians their credibility as authority figures. Hence, one device would be to avoid using long monotonous sentences and to vary functional strategies together with linguistic switches and accompanying paralinguistic features.

There is a need for more studies on political discourse with a concern to illustrate code-switching in public speaking, so as to deepen "our understanding of conventionalization of register and the change of register systems over time" (Ferguson 1996). Political debate is a rich avenue for analysis, particularly if one can access material-gathering speakers from different backgrounds and using various dialects and standards. This would provide further examples of combinatorial rules and analyze cases of leveling, borrowing, and convergence – a process by which speakers adapt their linguistic habits to those of the interlocutors in crossdialectal discourse (and in intercommunal situations), even if this process implies abandoning forms closer to Modern Standard Arabic to use linguistic forms closer to the standard or 'prestigious' dialect.

Other avenues of investigation concern the persuasive quality of a discourse, together with linguistic realizations, in different types of settings; analyzing presentations at professional conferences, religious sermons, and television debates on all kinds of topics whether formal or less formal, can help throw new light on variation in Arabic political discourse.

Bibliographical references

Primary Source

Ḥadīṯ al-baṭal az-zaʿīm Gamāl ʿAbdunnāṣir ʾilā l-ʾumma. Cairo: Dār at-Taḥrīr, 1970.

Secondary Sources

Abd El Jawad, Hassan. 1986. "Social functions of language variation". *al-Abḥāth* 34.21–37.

Atkinson, Max. 1984. *Our masters' voices: The language and body language of politics*. London: Methuen.

Auer, Peter. 1999. "From codeswitching via language mixing to fused lects: Towards a dynamic typology of bilingual speech". *International Journal of Bilingualism* 3.309–332.

Bengio, Ofra. 1998. *Saddam's word: Political discourse in Iraq*. Oxford and New York: Oxford University Press.

Boumans, Louis. 2002. "Repetition phenomena in insertional codeswitching". *Language contact and language conflict in Arabic*, ed. Aleya Rouchdy, 297–316. London: RoutledgeCurzon.

Chafe, Wallace. 1982. "Integration and involvement in speaking, writing and oral literature". *Spoken and written language: Exploring orality and literacy*, ed. Deborah Tannen, 35–53. Norwood, N.J.: Ablex.

Dijk, Teun A. van. 1981. "Episodes as units of discourse analysis". *Analysing discourse: Text and Talk*, ed. Deborah Tannen, 177–195. Washington, D.C.: Georgetown University Press.

Eid, Mushira. 1988. "Principles for code-switching between Standard and Egyptian Arabic". *al-ʿArabiyya* 21.51–79.

Elgibali, Alaa (ed.). 1996. *Understanding Arabic: Essays in contemporary Arabic linguistics in honor of El-Said Badawi*. Cairo: American University in Cairo Press.

Ferguson, Charles. 1959. "Diglossia". *Word* 15.325–40.

——. 1996. "Diglossia revisited". Elgibali (1996:49–67).

Gumperz, John J. 1982. *Discourse strategies*. Cambridge: Cambridge University Press.

Halliday, Michael A.K. 1978. *Language as a social semiotic*. London: Arnold.

Hary, Benjamin. 1996. "The importance of the language continuum in Arabic multiglossia". Elgibali (1996:69–90).

Heath, Jeffrey. 1989. *From code-switching to borrowing: Foreign and diglossic mixing in Moroccan Arabic*. London and New York: Kegan Paul International.

Holes, Clive. 1993. "The uses of variation: A study of the political speeches of Gamal Abd Al-Nasir". *Perspectives on Arabic linguistics*, V, ed. Mushira Eid and Clive Holes, 13–45. Amsterdam and Philadelphia: J. Benjamins.

——. 1995a. "Community, dialect and urbanization in the Arabic-speaking Middle East". *Bulletin of the School of Oriental Studies* 58.270–287.

——. 1995b. "The structure and function of parallelism and repetition in spoken Arabic: A sociolinguistic study". *Journal of Semitic Studies* 40.57–81.

Johnstone, Barbara. 1991. *Repetition in Arabic discourse: Paradigms, syntagms and the ecology of language*. Amsterdam and Philadelphia: J. Benjamins.

Koch, Barbara Johnstone. 1983. "Presentation as proof: The language of Arabic rhetoric". *Anthropological Linguistics* 25.47–60.

Lavandera, Beatriz R. 1978. "Where does the sociolinguistic variable stop?". *Language in Society* 7.171–182.

Mazraani, Nathalie. 1995. "Functions of Arabic political discourse: The case of Saddam Hussein's speeches". *Zeitschrift für Arabische Linguistik* 30.22–36.

——. 1997. *Aspects of language variation in Arabic political speech-making*. Richmond: Curzon Press.

Ochs, Elinor. 1979. "Planned and unplanned discourse". *Syntax and semantics,* ed. Talmy Givón, 51–80. New York: Academic Press.

Rouchdy, Aleya (ed.). 2002. *Language contact and language conflict in Arabic: Variations on a sociolinguistic theme*. London: RoutledgeCurzon.

Suleiman, Yasir. 1999. "Language and political conflict in the Middle East: A study in symbolic sociolinguistics". *Language and society in the Middle East and North Africa: Studies in variation and identity,* ed. Yasir Suleiman, 10–37. Richmond: Curzon Press.

——. 2003. The *Arabic language and national identity*. Edinburgh: Edinburgh University Press.

Taine-Cheikh, Catherine. 2002. "De la variation linguistique dans le prêche populaire mauritanien". Rouchdy (2002:177–202).

Tannen, Deborah. 1982. "Oral and literate strategies in spoken and written narratives". *Language* 58.1–21.

Nathalie Mazraani (Brussels)

Possession

Most of the modern dialects make use of two types of constructions, commonly referred to as synthetic vs. analytic, or direct vs. indirect, to encode relations of possession. The synthetic manner includes the processes of suffixation and juxtaposition (*status constructus*) following the word order *Possessed – Possessor* (Y – X). The analytic method makes use of a possession exponent (EXP). These different methods are bound by semantic constraints, mostly according to the notion of possession expressed, for example, alienable, inalienable, and abstract. The formal distinction in denoting inalienable relations dates back to ancient times. It was expressed in Classical and Middle Arabic and more generally in Semitic (Bravmann 1977; Blau 1965, 1979; Diem 1986). Alongside adnominal and prepositional constructions, some dialects have also developed forms of 'having', consisting of transitive constructions that are also semantically constrained.

1. Synthetic or direct constructions

In the process of suffixation, the Possessor (X) is represented by a pronoun suffixed to the Possessed element (Y–SUF). Suffixation is primarily compatible with so-called inalienable relations. The semantic domains covered by inalienability (+Time-stable and –control) are not the same in all dialects. This is one reason why suffixation appears to be more productive in some dialects than in others, contrary to claims made in the literature. McCarthy and Raffouli (1964), for instance, assert that there are no rules to predict when a particle is used and when simple suffixation is used, but they note that the use of the particle *māl* in colloquial Arabic of Baghdad is less frequent among educated speakers. Likewise, in a recent study on Cypriot Arabic, Borg (2004:78) writes: “This constitutes an original feature of CyA since most Arabic vernaculars attach pronominal suffixes to most native nouns in unrestricted fashion”.

Certain domains, however, appear to be inalienable in the majority of dialects, specifically those concerning partitive, parental, or neighborly relations, and body parts: *ḍyāf-hum* ‘their guests’ (Algeria), *kalb-ī* ‘my dog’ (Syria-Lebanon-Palestine), *ḥnāko* ‘his cheeks’ (Tangier, Morocco). Domestic animals are considered inalienable in the languages of the Syria-Lebanon-Palestine area, but not in other dialects. In a general manner, the dialects which favor analytic constructions in all contexts are those that are termed peripheral or isolated (Abbéché, Daragözü; see Sec. 2).

Other factors than the semantic one may intervene to prevent recourse to suffixation, depending on the dialect. These factors include (i) the nature of the final syllable: *wlad-hum* ~ *d-drari dyal-hum* ‘their children’ (Morocco), *əlkərsi ntaʿi* ‘my chair’ (Tunisia); (ii) borrowing from Standard Arabic or other languages, *wlad-hum* ~ *l-abnaʾ dyal-hum* ‘their children’ (Morocco), *il-baẓaburṭ ibtāʿak* ‘your passport’ (Egypt); nevertheless, borrowing is not submitted to the same process in all dialects: *sellulēre* ‘my cell phone’, *paspōre* ‘my passport’ (Syria-Lebanon-Palestine); and (iii) stylistic factors, such as → focus.

Juxtaposition is the process of → *ʾiḍāfa*, known as *status constructus* or → construct state; the Possessor (X), necessarily definite, is generally postposed to the possessed element (Y), usually indefinite (Y – X), and no other element, such as adjectives, may be inserted between the two. There are lacunae in the investigation of this type of construction, and it is difficult to establish any coherent groupings for the different uses (Bedouin, rural, urban). However, classifications proposed for the Maghreb (Marçais 1977) and for all dialects (Harning 1980) show certain tendencies: the construct state is more frequent in the Sahara Bedouin dialects; it forms one of the conservative features of the Bedouin dialects in the Arabian peninsula; and it is rarer in the peripheral or ‘isolated’ dialects.

In the dialects which have limited use of the construct state, it is mostly found in inalienable and abstract relations: *mart-xūya* ‘my brother’s wife’ (Maghreb); *ʿala rāg el-baḥar* ‘on the seashore’ (Iraqi *gilit* dialects); *zəġār əlmaḷḷa* ‘the molla’s children’ (Daragözü); *danab al-humār* ‘the donkey’s tail’ (Abbéché); *bin issultān* ‘the king’s son’ (Malta). Some dialects have constructions with a definite Possessed element: *al-wilēd ar-rāyil* ‘the son of the man’ (Abbéché). Others have constructions with inverse word order (X – Y-SUF), where the Possessor (X) is doubly present through a pronoun suffixed to the Possessed element: *amīr fulūs-u* ‘the emir’s money’ (Uzbekistan).

2. Analytic constructions

Most of the modern dialects exhibit both constructions, synthetic genitive (SG) and → analytic genitive (AG), with differing degrees of specialization for the synthetic genitive in inalienable possession. The analytic genitive appears to be more productive in sedentary, especially urban, speech, less frequent in rural, and practically absent from Bedouin dialects. In the dialects where both analytic and synthetic genitive are productive, the analytic genitive is found in inalienable contexts, where enunciative and pragmatic factors, such as specification and focus through dislocation, are present: *abu meli* ~ *abūy* ‘my father’ (Oman); *yšay əḥmil al-bēt ḥaqqu w-ənkullu yā ʾind bēt ʾabūh* ‘he wants me to transport his own house all the way to his father’s’ (insistence, Yemen); *haw j-jirān tabʿūlik* ‘those ones, your neighbors’ (exasperation and irony, Lebanon); *hāḏa l-*

Table 1. Variants of the genitive exponent in Arabic dialects

tabaʿ, tabʿūl, tāʿ, tēʿ, tūʿ, teʿūl	Syria-Lebanon-Palestine, Sudan, *qəltu* dialects
bətāʿ, btāʿ, bitūʿ, ntāʿ, tāʿ, mtāʿ	Egypt, Cairo, Sudan, Syria-Lebanon-Palestine, Algeria, Tunisia, Morocco
jna	Libya
dyāl, di, d-, eddi, elli	Algeria, Morocco
šī, šyāt, šēt, šīt, šayt, šayyet	Syria-Lebanon-Palestine, Cyprus
geyy, gī, gīt	Syria-Lebanon-Palestine, Iraq (central rural districts)
ḥagg, ḥaqq	Arabian Peninsula, Galilee Bedouin, Sudan
māl, mālt, mālāt	Iraq and Oman, *gilit* dialects (rural districts in central and lower Iraq), Baghdad, Basra, Persian Gulf
d̠īl, d̠īla, d̠ēl, d̠ēl	*qəltu* dialects (sedentary dialects of upper and lower Iraq, Anatolia), Syria-Lebanon-Palestine
lēl, lēlīl, līt, allīl	*qəltu* dialects, Daragözü, Sudan
šuġl	Negev Bedouin. Upper Egypt, Syria-Lebanon-Palestine
ʾihnīn	Upper Egypt
hana, hinē, hīl, hille	Chad, Nigeria, Sudan
hūl, hīl	Sudan
ḥāl	Oman, Zanzibar
ta, t-, tīʿ-	Malta
tel	Cyprus
la	epexegetic genitive: Lebanon-Syria-Palestine, Algeria, Morocco, Cilicia, Cyprus
ʿala	epexegetic genitive: Cilicia

bayt ḥaggī 'this house [this one] belongs to me' (specification, Yemen). Unfortunately, these factors are not always taken into account in available descriptions.

Furthermore, it is possible to solicit the analytic genitive for the syntactic possibilities it offers (control of each noun's definiteness, integration of an adjective between Y and X, formation of genitival phrases with more than two nouns), which are impossible with the synthetic genitive. Apart from these contexts, using the analytic genitive to encode inalienable relations is generally deemed unacceptable. This is why in Yemeni Arabic, for example, the exponent *ḥagg-* can operate in inverse order, making it possible to extract a part of a whole in order to attach it to a different entity: **dīwān-na* is unacceptable because *dīwān* belongs to the entity 'house', but *d-dīwān ḥagga-na* 'our diwan [the diwan of us]' is correct (Naïm 2004). Lastly, recourse to the analytic genitive makes it possible to avoid ambiguity due to homonymy: *bintik* 'your daughter' (inalienable) ~ *hayde l-bint tabʿītik* 'that one, your maid', *sitti* 'my grandmother' ~ *is-sitt ibtaʿti* 'my wife' (Cairo).

The coexistence of analytic and synthetic genitive in many dialects and the fact that in most of the cases the synthetic genitive remains an alternative to the analytic genitive, even though the latter was already observed in ancient times (Blau 1965; Versteegh 1984), has been explained by (i) the origin of some of the exponents that derive from nouns denoting 'possession' (the others deriving from demonstratives and relatives), which renders them unsuitable for encoding 'abstract' notions (Harning 1980); and (ii) the growing influence of Classical Arabic that has supposedly led to a 'decreolization' of the dialects rather than a stabilization of the analytic genitive (Versteegh 1984).

Generally speaking, the two terms of the possessive construction are definite. Depending on the dialect, the exponent placed between (Y) and (X) agrees with (Y), either partially (in number) or completely (gender and number, with or without taking into consideration the human/nonhuman distinction), or remains invariable, set in the singular or in the plural: *ad-dayma ḥagga-hum tāliʿ* ‘their kitchen is upstairs’ (Sanaa, Yemen); *is-sayyāra māltek* ‘your car’, *ḥaram bayt yaʿni ḥawiz mēl bayt* ‘the *ḥaram* of a house is the same as its *ḥawiz*’ (Oman); *l-xaddāmtēn mālāt-hum* ‘their two maids’ (Iraq); *əž-žnayne teʿūl ʾaḥmad* ‘Ahmad’s garden’ (Beirut); *elktub diyulu* ‘his books’ (Morocco); *al-kitāb an-nāfiʿ hūl wāḥid ḥakīm* ‘the useful book of a (certain) doctor’ (Sudan); *is-sawwāq b*^{e}*tāʿ it-taks* ‘the taxi driver’ (Cairo); *pagra šayt-i* ‘my cow’ (Cyprus). In most peripheral or isolated dialects, which do not have a synthetic genitive (apart from a few set phrases), the definiteness of (Y) and (X) does not seem to obey any fixed rules: *əbən lē ʿammi* ‘the son of my uncle’ (Daragözü); *[al] ahal hana mūsa* ‘Musa’s family’ (Abbéché).

The epexegetic genitive constitutes another type of analytic construction. Resembling transitive constructions introduced by a specific accusative marker such as *l-*, *ʿala*, or being identical to the analytic marker, the Possessor is here represented in an anticipatory manner by a pronoun suffixed to the Possessed element: (Y-SUF + EXP + X). These constructions have been explained in some cases as Aramaic substrate (Féghali 1928; Blau 1979; Borg 2004) and in others as Berber (Fischer 1907, 1909). The presence of a redundant suffix denoting intrinsic possession is attested at an earlier stage in Semitic and in Berber (Blau 1979). In modern dialects, these constructions are most often seen in kinship expressions: *mpratu l-ʿammi* ‘my paternal uncle’s wife’, *exla l-ʿarus* ‘the bride’s parents’ (Cyprus); *abūnu ssalmān* ‘Salman’s father’ (Baghdad, Mosul); *ʿamm le l-xūri* ‘the priest’s paternal uncle’ (Lebanon); *immu la-lmirḥūm* ‘the mother of the deceased’, *ibnu la-š-šayx* ‘the son of the sheikh’ (Cilicia, southern Turkey); *lūsetha ddi zīn*^{e}*b* ‘Zineb’s sister-in-law’ (Algeria, Djidjelli); *mamāhā de-sulṭāna* ‘Sultana’s mother’ (western Morocco).

The list of exponents presented in Table 1 uses the data given in Harning (1980), Shboul (1983), and Versteegh (1984), and provides additional examples. They are arranged in geographical order. Given that within any one dialect the agreement or rigidity of exponents is extremely unstable, and given the scarcity of information on this point for many dialects, morphological inflections have been ignored. Exponent variants are listed one after the other, but this does not necessarily mean that they are all present in all the countries listed in the opposite column.

In most of the dialects, the exponents carrying personal suffixes give rise to possessive pronouns which may serve as predicates in nominal sentences with a subject generally defined by a demonstrative.

3. TRANSITIVE POSSESSION OR ‘HAVING’

Three exponents, *li-*, *ʿind-*, and *maʿ-*, are found; on the semantic level, they refer to dative, location, and companion schemas, respectively, but all show the same morphosyntactic pattern, with (X) preposed to (Y), contrary to what is observed for these same schemas in the world’s languages (Heine 1997; Stassen 2001).

These preposition/exponents are found in Classical Arabic for encoding relations more or less closely associated with the notion of possession (Classical Arabic *ladā* has not followed the same path of evolution in the modern dialects). They do not have the same status in all dialects, some being more productive than others, and some being specialized in the encoding of a specific possessive notion. Unfortunately, data on the question are lacking for numerous dialects. For those dialects where information is accessible, the facts present themselves as follows:

i. *li-* (to X–Y) is rare in Syria-Lebanon-Palestine and Yemenite; it is only found in set expressions and for abstract possession. In Egyptian Arabic, it is found in inalienable contexts and with an inanimate Possessor. In the peripheral dialects of Cyprus and Malta, however, it is more productive and encodes the very general notion of ‘having’ and different notions of possession: *mā ʾəli nafs* ‘I have no appetite/I don’t feel like’ (Syria-Lebanon-Palestine); *mā l-ī šī* ‘I have nothing / I am well’ (Yemen); *liyya ʿamm*^{ə} *wāḥid bass* ‘I have only one uncle’, *il ʾōḍa lha bāb wi šibbakēn* ‘the room has a door and two

Table 2. The semantico-cognitive parameters subjacent to the expression of 'permanent' and 'temporary' possession

Notions	Exponents	Time stable	Control	Spatial contiguity
Permanent possession	*ʿind-*	+	±	–
Temporary possession	*maʿ-*	–	+	+

windows' (Egypt); *yám u xát kallíni ʿáskar maʿzumín* 'Sunday I had many invited people' (Cyprus); *kellu ktieb* 'he had a book' (Malta).

ii. *ʿind-* (at X–Y) is very productive for encoding 'having' in different Eastern, Western, and peripheral dialects. It is found in alienable (+Time stable and +control), inalienable (+Time stable and –control) and abstract contexts. Contrary to its meaning in Classical Arabic, it does not imply the notion of 'presence' or 'proximity' of the Possessor and the Possessed element: *ʿind-un televizyōn* 'they have TV', *ʿind-i bint w ṣabi* 'I have a girl and a boy' (Syria-Lebanon-Palestine); *ʿandi šuġl* 'I'm busy', *ʿand-ik awlād*? 'do you have any children?' (Cairo); *ʿnd-u bnt mzyana* 'he has a pretty daughter' (Fes); *m'għandux xobża* 'he doesn't have a loaf' (Malta).

iii. *maʿ-* is specific to 'temporary' possession and is characterized by the feature 'spatial contiguity'. Both *ʿind-* and *maʿ-* are compatible with the notion of abstract possession: *maʿāya ʿišrīn ginēh* 'I've got twenty pounds with me' (Egypt); *maʿa-ni- ya-xa* 'I have it, I'm holding it' (Cyprus); *maʿi məftēḥ* 'I've got a key with me', *maʿi/ʿindi ḥarāra* 'I have a fever' (Syria-Lebanon-Palestine). In the Yemen dialects, *ʿind-* is very rare and was possibly introduced through borrowing. On the other hand, *maʿ* (*maʾ-* in Zabid, Yemen) is more frequent and very productive (alienable, inalienable, and abstract contexts): *m-faqīr maʾāh ṯnēn banātu* 'the poor man had two daughters' (Zabid, Yemen); *mā mʿī la ʾumm wa la ʾabb* 'I have neither father nor mother', *maʿiš ḥagg* 'you are right' (Sanaa).

In these constructions, the Predicate precedes the Subject, which is usually indefinite. For focus, one finds the word order Subject + EXP-SUF + Y. Thus, possessive predication shows the same word order as the verbal phrase. But possessive constructions behave in a specific manner which distinguishes them from both nominal and verbal phrases:

i. The temporal exponent and the verbal modalities preceding the predicatoid (*lī-*, *ʿind-*, *maʿ*) remain invariable, set in the 3rd person singular regardless of the possessor: *jirān-nā ṣār ʿəndun televiziyōn* 'our neighbors now have a television' as against *l-wlēd ṣārū yʿayṭū* 'the children began shouting' (verbal predication); *l-banēt ṣārū kbār* 'the girls have grown up' (nominal predication, Beirut); *kān maʿāya ʿarabiyya* 'I had a car with me', *kān ʿandaha waʾt* 'she had time' (Cairo).

ii. Negation shows morphological marking of the verbal type: *mā maʿī məftēḥ* 'I don't have a key' ~ *l-məftēḥ məš maʿī* 'the key is not with me' (Syria-Lebanon-Palestine); *mā mʿayāš sagāyir* 'I haven't got any cigarettes with me' (Cairo); *pawlu m'għandux ktieb* 'Pawlu doesn't have a book' (Malta). With verbal predication, possessive constructions share the obligatory presence of an anaphoric pronoun representing the subject or possessor even when it is lexically present: *j-jirān ʿandun diš* lit. 'the neighbors, they have a dish antenna' (Syria-Lebanon-Palestine); *ʿammiti nabīla kān laha tnēn ixwāt* 'my aunt Nabila had two sisters' (Cairo); *pawlu sa ykollu ktieb* 'Pawlu will have a book' (Maltese). But these two types of construction differ as to the nature of the suffixed pronoun: the suffix representing the subject or possessor in the possessive construction is morphologically object marked (Comrie 1989) (→ pseudoverb).

Certain dialects show the particle *fī* at the beginning of constructions, namely in interrogative and negative ones. In this context, *fī* has assertive modality value: *fī ʿindi kətub* '[in effect, yes,] I have books', *mā fī maʿi maṣāri* '[no,] I don't have money' (Beirut); *fī ʿandi lamūn mixallil* 'I have some candied lemon' (Cairo). This situation may indicate the evolution of existential constructions, where the directional *ʿind* and the comitative *maʿ* govern the suffix form, toward possessive constructions, following formal processes of focus and topicalization (Naïm 2003).

Bibliographical references

Blau, Joshua. 1965. *The emergence and linguistic background of Judaeo-Arabic: A study of the origins of Middle Arabic.* Oxford: Clarendon Press.

——. 1979. “Redundant pronominal suffixes denoting intrinsic possession”. *Journal of the Ancient Near Eastern Society of Columbia University* 11.31–37.

Borg, Alexander. 2004. *A comparative glossary of Cypriot Maronite Arabic.* Leiden: E.J. Brill.

Boumans, Louis. 2002. “Possessive constructions in Morocco and in the Netherlands”. *Aspects of the dialects of Arabic today (Proceedings of the 4th Conference of the International Arabic Dialectology Association (AIDA)*, ed. Abderrahim Youssi a.o., 265–275. Rabat: Amapatril.

Bravmann, Max. 1977. “The idea of possession in linguistic expression”. Max Bravmann, *Studies in Semitic philology*, 357–373. Leiden: E.J. Brill.

Brockett, Adrian. 1985. *The spoken Arabic of Khābūra.* Louvain: Imprimerie Orientaliste.

Comrie, Bernard. 1989. *Language universals and linguistic typology.* Chicago: University of Chicago Press.

Diem, Werner. 1986. “Alienable und inalienable Possession im Semitischen”. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 136.227–291.

Féghali, Michel. 1928. *Syntaxe des parlers arabes actuels du Liban.* Paris: Geuthner.

Fischer, August. 1907, 1909. “Eine interessante algerisch-marokkanische Genitivumschreibung”. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 61.178–185, 63.825–827.

Fischer, Wolfdietrich and Otto Jastrow (eds.). 1980. *Handbuch der arabischen Dialekte.* Wiesbaden: O. Harrassowitz.

Harning, Kerstin Eksell. 1980. *The analytic genitive in the modern Arabic dialects.* Göteborg: Acta Universitatis Gothoburgensis.

Heine, Bernd. 1997. *Possession, cognitive sources, forces, and grammaticalisation.* Cambridge: Cambridge University Press.

Jastrow, Otto. 1973. *Daragözü: Eine arabische Mundart der Kozluk-Sason-Gruppe (Südostanatolien).* Nürnberg: H. Carl.

Marçais, Philippe. 1977. *Esquisse grammaticale de l'arabe maghrébin.* Paris: Adrien Maisonneuve.

McCarthy, R.J. and Faraj Raffouli. 1964. *Spoken Arabic of Baghdad.* Beirut: Librairie Orientale.

Naïm, Samia. 2003. “La prédication possessive et l'émergence de formes d'avoir en arabe oriental”. *Bulletin de la Société de Linguistique de Paris* 98:1.359–384.

——. 2004. “Le traitement syntaxique des relations inaliénables en arabe yéménite de Ṣanʿā”. *Proceedings of the Seminar for Arabian Studies* 34.209–219.

Palva, Heikki. 1976. *Studies in the Arabic of the semi-nomadic ʿAjarma tribe (al-Balqāʾ District, Jordan).* Stockholm: Acta Universitatis Gothoburgensis.

Procházka, Stephan. 2000. “Some morphological and syntactical characteristics of the Arabic dialects spoken in Cilicia (southern Turkey)”. *Proceedings of the Third International Conference AÏDA*, ed. Manwel Mifsud, 220–221. Malta: Salesian Press.

Roth, Arlette. 1979. *Esquisse grammaticale du parler arabe d'Abbéché (Tchad).* Paris: Geuthner.

Shboul, Ahmad M.H. 1983. “‘Having’ in Arabic”. *Zeitschrift für Arabische Linguistik* 11.24–47.

Stassen, Leon. 2001. “Predicative possession”. *Language typology and language universals*, ed. Martin Haspelmath, Ekkehard König, Wulf Oesterreicher, and Wolfgang Raible, 954–960. Berlin and New York: de Gruyter.

Versteegh, Kees. 1984. *Pidginization and creolization: The case of Arabic.* Amsterdam: J. Benjamins.

Woidich, Manfred. 2006. *Grammatik des Kairenisch-Arabischen.* Wiesbaden: O. Harrassowitz.

—— and Rabha Heinen-Nasr. 2004. Kullu tamām*: An introduction to Egyptian Colloquial Arabic.* Cairo: American University in Cairo Press.

Samia Naïm (LACITO-CNRS)

Pragmatics

Pragmatics, in its broadest sense, is “the study of action deliberately undertaken with the intention of causing the intended interpreter to re-assess his model of how things are, including his system of values and his model of the speaker's beliefs, attitudes, and intentions” (Green 1996:5). Pragmatics, according to this definition, encompasses all intentional communicative acts whether verbal or nonverbal. For example, a hand gesture that is characteristic of Egyptians involves holding the fingers of the right hand together and moving it up and down slowly. This gesture can be interpreted as a request for someone to slow down or wait, or as a threat (see El-Araby 1983 and Farghali 1997 for a description of the nonverbal communication patterns associated with Arabic). For viewers to successfully interpret such an ambiguous gesture and determine the goals and intentions behind it, they must recognize it as an intentional communicative act and utilize the available contextual cues and background knowledge according to universal and culture-specific pragmatic principles. Similarly, the Egyptian Arabic utterance *ʾistanna ʿalayya* ‘wait for me!’ can also be interpreted as a request or a threat, and the same pragmatic principles that are employed to interpret the equivalent gesture are employed to disambiguate the utterance. These principles and contextual factors are the subject matter of pragmatics.

Linguistic pragmatics involves the study of the mechanisms and principles that allow individuals to communicate successfully by recognizing the intentions and beliefs motivating linguistic acts, as well as the strategies speakers use to carry out such acts, i.e. the changes they achieve using language (Mey 1993; Green 1996). Levinson (1983) further narrows the scope of linguistic pragmatics by excluding those aspects of language use that are not reflected in the grammar, such as dialectal features and → code-switching, hence distinguishing pragmatics from sociolinguistics. Levinson's definition restricts the range of phenomena investigated within the framework of pragmatics to those context-dependent aspects of meaning that are neither structurally compositional (i.e. syntactic) nor truth conditional (i.e. semantic), such as → deixis (e.g. honorifics, demonstratives, and referential pronouns), conversational implicature, → speech acts, and presupposition, among other areas of research.

For example, the sentences in (1a) and (1b) have the same truth conditions and the same syntactic structure. In fact, these two sentences are semantically equivalent, as both of them are true if, and only if, both conjuncts are true relative to a given model.

(1a) *muna rāḥit* *is-sūʾ* *w-ištarēt*
Mona went the-market and-bought
gawāfa min *ʿand* *il-fakahāni*
guava from at the-fruit shop
'Mona went to the market and bought guava from the fruit shop'

(1b) *muna* *ištarēt* *gawāfa min ʿand*
Mona bought guava from at
il-fakahāni *w-rāḥit* *is-sūʾ*
the-fruit.shop and-went the-market
'Mona bought guava from the fruit shop and went to the market'

However, a speaker who utters the first sentence is usually understood to indicate that Mona went to the fruit shop at the market in order to buy guava, or that she went to the market first and later bought the guava from a fruit shop at a different location. The second sentence, on the other hand, suggests that Mona bought the guava first and then went to the market, with no causal interpretations. The syntax and the semantics of those two sentences assert only that Mona was the agent of two events: going to the market and buying guava, but the casual and sequential interpretations are only a matter of conversational implicature, triggered by the ordering of clauses. A speaker intentionally orders conjoined action sentences to implicate the sequence of events. Conversational implicature is not subject to truth conditions because it can be denied without affecting the truth value of the sentence. For example, the sentence in (1a) can be true even if Mona bought the guava before she went to the market.

The examples in (1a) and (b) demonstrate that conversational implicature is not a semantic or a syntactic phenomenon. However, sequencing implicature is semantically constrained so that it applies only to utterances describing events, as this implicature does not arise if the predicates in the conjoined clauses describe states, as in (2a) and (2b).

(2a) *nadya kānit* *taʿbāna wi-(kānit)* *zahʾāna*
Nadia was tired and-(was) bored
'Nadia was tired and (she was) bored'

(2b) *ʿali nāyim*
Ali fall.asleep.Act.Part
w-fātiḥ
and-open.Act.Part.
iš-šibbāk
the-window
'Ali is in a state of having fallen asleep and opened the window' =
'Ali fell asleep and opened the window'

The speaker of the sentence in (2a) does not imply that Nadia was tired first and then bored. The sentence describes her as being both tired and bored at the same time. Also, the sentence in (2b) does not describe Ali as having fallen asleep first and then opening the window. Rather, the sentence describes him as the theme of a state of having fallen asleep and having opened the window, both of which hold of him at speech time (Mughazy 2004). Therefore, conversational implicature is a pragmatic phenomenon that can be semantically and syntactically constrained.

One of the fundamental pragmatic principles is Grice's Cooperative Principle, which accounts for how language users calculate conversational implicature (Grice 1957, 1971, 1975, 1978, 1989). The basic premise of this principle is that communication is possible only because there are certain maxims that speakers expect their interlocutors to adhere to. These maxims are:

Quantity:

i. Make your contribution as informative as is required (for the current purposes of the exchange).
ii. Do not make your contribution more informative than is required.

Quality:

i. Do not say what you believe to be false.
ii. Do not say that for which you lack adequate evidence.

Relation:

i. Be relevant.

Manner:

i. Avoid obscurity of expression.
ii. Avoid ambiguity.
iii. Be brief (avoid unnecessary prolixity).
iv. Be orderly.

Grice's cooperative maxims are not rules for conversational etiquette within a particular culture. Rather, they are universal assumptions about how speakers execute their linguistic acts and interpret others' utterances. When individuals are faced with pragmatic ambiguity, they appeal to these maxims in an attempt to determine the speaker's intentions. For example, John and Bill are in a parking lot ready to leave, but Bill sees their acquaintance Mary, and the following exchange takes place:

Bill: There's Mary. I'll go say hi.
John: I don't have enough quarters.

To understand why John makes such a comment and what he means by it, other than the fact that he believes the quarters he has are not enough for some purpose, Bill needs to make a series of calculations. First, Bill needs to assume that John's utterance is related to his own initial assertion and the current context. For example, Bill assumes that the quarters are needed for the parking meter and that John believes that talking to Mary will take some time. Moreover, Bill needs to assume that John thinks his utterance provides enough clues for him to interpret it without explicitly asserting his intentions. Maybe John does not think it is appropriate to say what he thinks explicitly or he is concerned that Mary could overhear him. Finally, Bill has to believe that John is not telling a lie or trying to mislead him. These are only some of the contextual assumptions that Bill has to make to compute John's utterance.

However, John intentionally flouts the manner maxim, as his utterance is pragmatically ambiguous. For example, John could be understood to request that Bill pay the parking meter, to discourage him from talking to Mary, or to remind him that Mary is talkative. Bill needs to add to his calculations his beliefs about John's personal traits and his relationship with Mary. Once Bill puts the pieces together, he can make a calculated inference about the intentions behind John's utterance. Note that John must assume that Bill is capable of making these calculations and that Bill has sufficient background information about John and his relation to Mary; otherwise, he would fail to communicate his intentions.

Wierzbicka (1991) and Green (1996) note that Grice's maxims are prioritized differently across cultures, as, for example, members of some speech communities value the maxim of manner over quality, hence accepting exaggerations and 'white lies' for face-saving or rhetorical purposes, while other cultures value the maxim of quality most, hence displaying more directness in their discourse. Johnstone (1991), Omara (1993), Farghal (1995), Harris (1996), and Salih (2001) discuss discourse patterns in various spoken varieties of Arabic where speakers intentionally flout Grice's maxims to comply with the conventional cultural patterns associated with politeness and face relations. For example, in the Jordanian exchange in (3), where speaker A invites speaker B for food, speaker A violates the quantity maxim by providing value judgments, such as *mā fi ʾiši min wāžbak* 'there is nothing worthy of your status', and information, such as *mā žibna ʾiši min barra* 'we have not got any food from outside', that seem uncalled for in violation of the quantity maxim. Speaker B flouts the quality maxim by asserting that he does not want to eat and that he has just eaten.

(3) A:
ʾitfaḍḍal al-maysūr, mā fi ʾiši min wāžbak, ʾahlan wasahlan ʾakil min id-dār mā žibna ʾiši min barra
'Please help yourself to the available food; there is nothing that is worthy of your status; you are most welcome; we have not got any food from outside the house'

B:
maysūr ilġānmīn, kattir xeirku lākin mā bnafsī l-ʾakil/z-zād, hassaʾ wākil
'The food is that of the generous; I hope that your fortune will increase, but I do not want to eat because I have just eaten' (Salih 2001:92)

Successful interpretation of such utterances requires knowledge of the speakers' cultural norms. For example, speaker A flouts the quantity maxim to assure his addressee that his acceptance of the invitation does not constitute imposition, and speaker B violates the quality maxim because politeness norms in Jordanian Arabic require speakers initially to turn down invitations for food.

Conversational implicature concerns uses of language where a speaker explicitly asserts a proposition and implicates another at the same time. There are, however, many cases where a speaker utters a sentence to communicate its propositional truth-conditional content and at the same time to get the addressee to do something or to recognize that the speaker has done something in making that utterance (Searle 1969, 1979). For example, if a speaker says *it is cold in here*, addressees would understand the state of affairs described by the comment and act according to their interpretation of what the interlocutor intended for them to do by turning the heat up, or offering a blanket. Moreover, if a speaker asks *are you hungry?*, an addressee would interpret that utterance not only as a yes/no question but also as an invitation for food or as a request to go eat. These cases are examples of indirect speech acts.

Austin (1961, 1962) and Searle (1979) distinguish at least two types of communicative acts: locutionary acts, which constitute the uttering of a sentence or a phrase, and illocutionary acts, which constitute events completed by carrying out locutionary acts. For example, if someone asks *would you like some tea?* the speaker has carried out the locutionary act of making a question as well as the illocutionary act of offering. Speakers can choose how to execute their speech acts from various possible strategies depending on how they believe these strategies facilitate achieving their goals. For example, speakers can opt to use explicit performative verbs, as in (4a) and (4b) (→ performatives).

(4a) *ʾaqtariḥu* *taʾjīl* *al-iqtirāʿ*
1s.suggest postponing the-vote
ʾilā *jalsat* *al-xamīs*
to session Thursday
'I suggest postponing the voting to Thursday's session'

(4b) *ʾaškuru-kum* *ʿalā*
1s.thank-you.pl on
ḍiyāfati-kum
hospitality-your.pl
'I thank you for your hospitality'

Using performative verbs such as those in (4a) and (4b) ensures that the addressee(s) will immediately recognize the speaker's intentions because these verbs explicitly specify the illocutionary force of the utterance.

Another way of unambiguously signaling the intentions behind one's utterance is to use illocutionary force-indicating devices, which include specific phrases or syntactic constructions that are conventionally associated with particular speech acts (Stampe 1975). For example, Standard Arabic illocutionary force-indicating devices include the interjections or fixed expressions known as *ʾasmāʾ al-ʾafʿāl* (→ *ism al-fiʿl*), such as *ʾāmīn*, which has the illocutionary force of requesting acceptance; *ḥayya* (as in *ḥayya ʿalā ṣ-ṣalā*), which has the illocutionary force of summoning; and *hayya*, which is used to get someone to hurry. Kuwaiti illocutionary force-indicating devices include the phrases *hala biṭṭašš wirrašš*, which is only used to enthusiastically welcome an addressee, and the phrase *yā mʿawwad*, which is used to request help (Muḥammad 2000). Unlike performative verbs, these fixed expressions do not have propositional content, yet speakers recognize the motivation for using them by convention (El-Sayed 1990; Ghobrial 1993; M. Kamel 1994; Masliyah 1999). Standard Arabic illocutionary devices also include particles that are used with propositions, as is the case with instigation *halla* in (5a) and *hab*, which is used exclusively for hypothesizing, as in (5b).

(5a) *halla* *ʾawqaf-ta* *hāḏā* *ḍ-ḍajīj*
halla stopped-2masc.sg. this the-noise
'Will you stop this noise?'

(5b) *hab* *ʾanna-ka* *najaḥ-t*
suppose that-you succeeded-2ms
māḏā *tafʿal?*
what 2ms.do
'Suppose that you succeeded, what would you do?'

Illocutionary force can also be signaled by the use of constructions that are conventionally associated with particular speech acts. For example, the negative operator in Egyptian Arabic can be used with wide scope to issue polite requests, as in (6a). Requests can also be made with a lesser degree of politeness by using the polarity item *mā* preceding the utterance, as in (6b).

(6a) *miš* *tī-gi* *ti-sallim*
Neg 2ms-come 2ms-greet
ʿala *iḍ-ḍiyūf*
on the-guests
'Won't you come and greet the guests?'

(6b) *ma* *tī-gi* *ti-sallim*
ma 2ms-come 2ms-greet
ʿala *iḍ-ḍiyūf*
on the-guests
'Come and greet the guests!'

Metalinguistic negation is another example of how illocutionary acts can be associated with particular grammaticalized forms (Mughazy 2003). It is a specialized use of the negative operator where it is used as "a device for objecting to a previous utterance on any grounds whatever – including its conventional and conversational implicata, its morphology, its style or register, or its phonetic realization" (Horn 1989:121). Negation in Egyptian Arabic is expressed discontinuously in sentences with present and past verbal predicates as in (7a) and (7b), whereas metalinguistic negation is expressed continuously, as in speaker B's utterances in (8a) and (8b).

(7a) *ma-ruḥ-t-iš* *il-madrasa*
Neg-went-1s-Neg the-school
ʾimbāriḥ?
yesterday
'I did not go to school yesterday'

(7b) *muna* *šāy*
Mona tea
ma-b-ti-šrab-š
Neg-Imperf-3fs-drink-Neg
'Mona does not drink tea'

(8a) A:
ruḥti *l-marsaḥ* *ʾimbāriḥ?*
went.2ms the-theater yesterday
'Did you go to the theater yesterday?'

B:
ʾana *miš* *ruḥt* *il-marsaḥ*
I Neg went.1s the-theater
ʾana *ruḥt* *il-masraḥ?*
I went.1s the-theater
'I did not go to the *theater*; I went to the theater'

(8b) A:
ti-ftikir *maḥmūd*
2ms-think Mahmoud
bi-y-ḥibb *nadya*
Imperf-2ms-love Nadia
'Do you think Mahmoud loves Nadia?'

B:
da *miš* *bi-y-ḥibbā-ha*
that Neg Imperf-2ms-love-her
da *bi-y-mūt* *fīha*
that Imperf-2ms-die in-her
'He does not love her; he dotes on her'

Speaker B in (8a) is not providing false information followed by a contradicting statement. Rather, she is objecting to Speaker A's use of the metathesis in *marsaḥ* 'theater', which is associated with lower socioeconomic classes, and offers a rectification providing what she views as the appropriate form, namely *masraḥ*. In (8b), speaker B objects to speaker A's description of the relationship between Mahmoud and Nadia and offers what she believes to be a more accurate description. These examples demonstrate that metalinguistic negation is a grammaticalized form that is associated with the illocutionary force of objecting to nontruth conditional material such as pronunciation and conversational implicature.

Metalinguistic negation differs from sentential negation in that the former is not truth-functional and it does not change the truth values of sentences. Therefore, metalinguistic negation does not license negative → polarity items; it does not co-occur with contrastive *bass* 'but'; and it can take scope over a sentential negation operator resulting in what seems like double negation (Mughazy 2003). Negation is ambiguous between a metalinguistic and a sentential use in many languages, including English; continuous negation can also be used metalinguistically in Egyptian Arabic, especially in sentences with nonverbal predicates resulting in ambiguity. Speaker B in the previous examples chooses the unambiguous form in adherence to the manner maxim and provides a rectification to satisfy the quantity maxim. Although the use of metalinguistic negation does not affect the truth values of utterances, it provides a systematic way to facilitate recognizing speakers' intentions and goals.

Explicit performative verbs and illocutionary force-indicating devices facilitate recognizing speakers' intentions because they either state the kind of speech act intended by the speaker or because they are conventionally associated with particular speech acts. However, it is often the case that speakers opt for less obvious ways of expressing their intentions because of cultural norms that might discourage directness, as is the case with Arab cultures. For example, it is the norm in many Arab subcultures that when a guest decides to leave, the host is expected to insist that the guest stay longer. Therefore, a guest has to find a way to take leave without offending the host. There are many possible strategies, and speakers choose the ones that they believe best suit their goals. A guest can simply and directly express the need to leave, as in (9a), or make a direct request, as in (9b). A more appropriate strategy is to provide a reason for leaving so that the host cannot negotiate, as in (9c).

(9a) *lāzim* *ʾa-mši*
must 1s-leave
'I have to leave'

(9b) *mumkin* *ʾa-mši*
possible 1s-leave
'May I leave?'

(9c) *il-wilād* *zaman-hum*
the-children time-their
rigˁu *min* *il-madrasa*
returned.3p from the-school
'The children must have returned from school'

The same mechanisms used to interpret conversational implicature are employed to interpret indirect speech acts. For example, it is necessary for the addressee to presuppose that the children mentioned in (9c) are the guest's children or some other children that the guest is responsible for, that the guest needs to leave to go home, and that the guest needs to be home to do something important that involves these children. Again, such calculations are possible only if the host assumes that the guest is adhering to Grice's Cooperative Principle. The question now is why speakers choose less direct ways of carrying out their speech acts and which strategies are available to do so.

One of the major areas of research on Arabic pragmatics is indirect speech acts, such as compliments (Mursy and Wilson 2001; Migdadi 2003), compliment responses (Mughazy 2000; Farghal and Al-Khatib 2001), refusals (Stevens 1993; Nelson a.o. 2002a,b; Al-Issa 2003), apologies (Bataineh 2004), → greetings (Al-Nasser 1993; Hassanain 1994; Emery 2000), vows (Salih and Abdul-Fattah 1998; Abd el-Jawad 2000), cursing (Stewart 1997; → insults), and arguing (B. Kamel 1983; Arent 1998). These studies investigate particular types of speech acts, describing the communicative strategies used to carry them out as well as the cultural motivation for the ways they are used. For example, refusals are potentially face-threatening speech acts that can offend the addressee. Therefore, speakers design their refusals in ways that they believe to mitigate the potentially negative effects of their refusals, such as using apologies, making promises, and giving reasons for their acts. In other words, speech acts are rule-governed patterns of linguistic behavior that are subject to the constraints of grammar as well as social and cultural rules.

BIBLIOGRAPHICAL REFERENCES

Abd el-Jawad, Hassan. 2000. "A linguistic and sociopragmatic cultural study of swearing in Arabic". *Language, Culture and Curriculum* 13.217–240.

Al-Issa, Ahmad. 2003. "Sociocultural transfer in L2 speech behaviors: Evidence and motivating factors". *International Journal of Intercultural Relations* 27.581–601.

Al-Nasser, Abdulmunʿim. 1993. "The social function of greeting in Arabic". *Zeitschrift für Arabische Linguistik* 26.15–27.

Arent, Eugene R. 1998. *The pragmatics of cross-cultural bargaining in an Ammani* suq*: An exploration of language choice, discourse structure and pragmatic failure in discourse involving Arab and non-Arab participants.* Ph.D. diss., University of Minnesota, Minneapolis.

Austin, John L. 1961. *Philosophical papers.* Oxford: Oxford University Press.

——. 1962. *How to do things with words.* Cambridge, Mass.: Harvard University Press.

Bataineh, Rula F. 2004. *A cross-cultural study of the speech act of apology in American English and Jordanian Arabic.* Ph.D. diss., Indiana University of Pennsylvania.

Cole, Peter and Jerry Morgan (eds.). *Speech acts: Syntax and semantics*, III. New York: Academic Press.

El-Araby, Salah. 1983. *Intermediate Egyptian Arabic: An integrative approach.* Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.

El-Sayed, Ali. 1990. "Politeness formulas in English and Arabic: A contrastive study". *International Review of Applied Linguistics* 89.1–23.

Emery, Peter. 2000. "Greetings, congratulating, and commiserating in Omani Arabic". *Language, Culture and Curriculum* 13.196–216.

Farghal, Mohammed. 1995. "Euphemism in Arabic: A Gricean interpretation". *Anthropological Linguistics* 37.366–378.

——. and Mahmoud A. Al-Khatib. 2001. "Jordanian college students responses to compliments: A pilot study". *Journal of Pragmatics* 33.1485–1502.

Farghali, Ellen. 1997. "Arab cultural communication patterns". *Journal of International Relations* 21.345–378.

Ghobrial, Naguib. 1993. *Discourse markers in colloquial Cairene Egyptian Arabic: A pragmatic perspective.* Ph.D. diss., Boston University.

Green, Georgia M. 1996. *Pragmatics and natural language understanding.* Mahwah, N.J.: Lawrence Erlbaum.

Grice, Paul H. 1957. "Meaning". *Philosophical Review* 66.377–388.

——. 1971. "Utterer's meaning, sentence meaning, and word meaning". *The philosophy of language*, ed. John Searle, 45–70. Oxford: Oxford University Press.

——. 1975. "Logic and conversation". Cole and Morgan (1975:41–58).

——. 1978. "Further notes on logic and conversation". *Syntax and semantics*, IX, ed. Peter Cole, 113–128. New York: Academic Press.

——. 1989. *Studies in the ways of words.* Cambridge, Mass.: Harvard University Press.

Harris, Rachael. 1996. "Truthfulness, conversational maxims and interaction in an Egyptian village". *Transactions of the Philological Society* 94.31–55.

Hassanain, Khalid. 1994. "Saudi mode of greeting rituals: Their implications for teaching and learning English". *International Review of Applied Linguistics* 32.68–77.

Horn, Laurence. 1989. *A natural history of negation.* Chicago: University of Chicago Press.

Johnstone, Barbara. 1991. *Repetition in Arabic discourse: Paradigms, syntagms, and the ecology of language.* Amsterdam and Philadelphia: J. Benjamins.

Kamel, Boshra. 1983. *Arguing in English as a second language: A study of speech act performance of Arabic speakers.* Ph.D. diss., University of Illinois at Urbana-Champaign.

Kamel, Mustafa A. 1994. *A sociolinguistic analysis of formulaic expressions in Egyptian Arabic.* Ph.D. diss., Georgetown University.

Levinson, Stephen C. 1983. *Pragmatics.* Cambridge: Cambridge University Press.

Masliyah, Sadok. 1999. "A cross-cultural misunderstanding: The case of the Arabic expression *inshallah*". *Dialog on Language Instruction* 13.97–116.

Mey, Jacob L. 1993. *Pragmatics: An introduction.* Cambridge: Blackwell.

Migdadi, Fathi H. 2003. *Complimenting in Jordanian Arabic: A socio-pragmatic analysis.* Ph.D. diss., Ball State University.

Mughazy, Mustafa. 2000. "Pragmatics of the evil eye in Egyptian Arabic". *Studies in the Linguistic Sciences* 30.147–158.

——. 2003. "Metalinguistic negation and truth functions: The case of Egyptian Arabic". *Journal of Pragmatics* 35.1143–1160.

——. 2004. *Subatomic semantics and the active participle in Egyptian Arabic.* Ph.D. diss., University of Illinois at Urbana-Champaign.

Muḥammad, Xālid S. 2000. *Kināyāt wa-ʾaqwāl kuwaitiyya wa-ʾuṣūluhā al-luġawiyya.* Kuwait City: Khaytan.

Mursy, Ahmad A. and John Wilson. 2001. "Towards a definition of Egyptian complimenting". *Multilingua* 20.133–154.

Nelson, Gayle L., Joan Carson, Mahmoud Al Batal, and Waguida El Bakry. 2002a. "Cross-cultural pragmatics: Strategy use in Egyptian Arabic and American English refusals". *Applied Linguistics* 23.163–189.

——. 2002b. "Directness vs. indirectness: Egyptian Arabic and US English communication style". *International Journal of Intercultural Relations* 26.39–57.

Omara, Salma. 1993. *The comprehension of conversational implicatures: A cross-cultural study.* Ph.D. diss., Ball State University.

Salih, Mahmud H. 2001. "A pragmatic analysis of certain interactional units in Jordanian Arabic". *Zeitschrift für Arabische Linguistik* 39.90–96.

——. and Hussein S. Abdul-Fattah. 1998. "English and Arabic oath speech acts". *Interface* 12.113–124.

Searle, John R. 1969. *Speech acts: An essay in the philosophy of language.* Cambridge: Cambridge University Press.

——. 1979. *Expression and meaning: Studies in the theory of speech acts*. Cambridge: Cambridge University Press.

Stampe, Dennis W. 1975. "Meaning and truth in the theory of speech acts". Cole and Morgan (1975:1–38).

Stevens, Paul. 1993. "The pragmatics of 'no!'": Some strategies in English and Arabic". *Issues and Developments in English and Applied Linguistics (IDEAL)* 6.87–112.

Stewart, Devin. 1997. "Impolite formulae: The cognate curse in Egyptian Arabic". *Journal of Semitic Studies* 42.327–360.

Wierzbicka, Anna. 1991. *Cross-cultural pragmatics: The semantics of human interaction*. New York: Mouton de Gruyter.

Mustafa Mughazy (Western Michigan University)

Predicate

1. Introduction

The term 'predicate' (along with the associated terms → 'subject' and 'predication') has been used for centuries in the Western logical and grammatical traditions to describe the second portion of a bipartite division of a sentence into a subject and as predicate.

How predicate and predication are used in contemporary linguistics varies considerably between users and between theoretical frameworks. The parameters of variation include the following:

i. The domain of predication: whether predication is defined over sentences, over clauses, or over projections of lexical items;
ii. The nature of the predication relation: whether predication is purely grammatical (syntactic), purely semantic (or pragmatic), or some combination of these;
iii. The formal devices used to represent the predication relation: whether the relation between a subject and a predicate is defined in terms of phrase-structural position, or in terms of binding or argument application.

To the extent that there is a consensus about predication, it consists of the following:

i. Predication is primarily a grammatical or structural relationship;
ii. Predication requires a subject to be in a very 'close' or local relationship with a predicate to which it is linked, usually defined in terms of c-command.

This ambiguity in how subject and predicate are used in Western linguistics is reflected in the Arabic terms for which they have been used as translations:

i. The *mubtadaʾ* 'that which comes first, topic' (→ *ibtidāʾ*) and → *xabar* 'report, comment' in a → nominal clause (*jumla ismiyya*);
ii. The *musnad* 'attribute' and *musnad ʾilay-hi* 'that which is attributed to' standing in an → *ʾisnād* 'attribution' relation to one another (see Levin 1981 for a discussion of the use of *musnad* and *musnad ʾilayhi* by the Arab grammarians);
iii. The → *fāʿil* 'doer, agent' of an action or eventuality and the → *fiʿl* 'event, action' or its like.

A comparison between the way in which subject, predicate, and predication are used and the way in which the Arabic terms are used is beyond the scope of this article. Instead, the article focuses on how the different senses of predicate and predication are applied to Arabic grammar.

2. Predicate as a fundamental sentence division

According to its most traditional meaning, a predicate is the second part of a division of a sentence or clause into a subject and a predicate. This division corresponds, at least partially, to the division of the Arabic → nominal clause (*jumla ismiyya*) into an 'initial NP' (*mubtadaʾ* 'that which is begun with') and a 'report' (*xabar* 'report, comment'). The correspondence is imperfect because some nominal clauses are analyzed as topic/comment structures rather than as predications (Bakir 1980; Eisele 1999; Mohammad 2000), although this characterization continues to be debated (see Sec. 4 below; and → nominal clause).

In some cases, the term 'predicate' is used specifically in the description of equational or copular clauses (Bakir 1980; → copula), while for others it is a fundamental division applied to all sentence types with subject-initial word order (e.g. Eisele 1999:99–125; Badawi a.o. 2004:306–307).

In copular clauses, the predicate is a lexical category such as an → adjective phrase (1a), prepositional phrase (1b), → noun phrase (1c), or → verb phrase (1d) in the case of participial verb stems.

(1a) *muḥammad* [$_{\text{AP}}$ *ʿayyān]*
Mohammad sick.ms
'Mohammad [is] sick' (Egyptian Arabic; Eisele 1999:102)
(1b) *ʿumar* [$_{\text{PP}}$ *f-d-dār]* *daba*
Omar in-the-house now
'Omar is in the house now' (Moroccan Arabic; Benmamoun 2000:40)
(1c) *ʾaxū* [$_{\text{NP}}$ *ḥallāʾ],* *ʾəsm-o*
brother-cl3ms barber name-cl3ms
[$_{\text{NP}}$ *ḥasan]*
Hasan
'His brother is a barber; his name is Hasan' (Syrian Arabic; Cowell 1964:403)
(1d) *ma-ḥadā-š* [$_{\text{VP}}$ *qāyim*
not-one.ms stand.ActPart.ms
yidfaʿ *ʿan-ne]*
pay.Imperf.3ms for-cl3ms
'No one stood up to pay for him' (Bir Zeit Palestinian Arabic; Schmidt and Kahle 1918:§34.1)

Clauses headed by finite verbs in SV word order have been treated in two ways. Jelinek (1981, 2002), Diesing and Jelinek (1995), and Eisele (1999) subsume them under the copular clauses, differing in terms of the inner structure of the predicate constituent. For example, Eisele (1999:99–109) analyzes the predicate as an 'extended verb phrase' (EVP), consisting of a 'predicator' (a finite verb or auxiliary) and a complement (a main verb or nominal argument), as in (2).

(2a) *muḥammad* *kān*
Mohammad was.3ms
biyiktib *gawāb*
write.Imperf.3ms letter
'Mohammad was writing a letter' (Egyptian Arabic; Eisele 1999:100)
(2b) [$_{\text{S}}$ [$_{\text{NP}}$ muḥammad] [$_{\text{EVP}}$ [$_{\text{V}}$ [$_{\text{V}}$ kān] [$_{\text{V}}$ biyiktib]] [$_{\text{NP}}$ gawāb]]]

Proposals within the Principles and Parameters or Minimalist frameworks do not refer to the predicate per se. Nonetheless, they capture the intuition that the clause is split between its subject and a functional projection or a sequence of functional projections such as an IP, Tense Phrase (TP), or Agreement Phrase (AgrP), which roots the clause and mediates the predication relation between the subject and the lexical head of the clause (Benmamoun 2000; Mohammad 1999, 2000; Ouhalla 1991, 1994).

The relationship between the subject and the predicate is more complex in Modern Standard Arabic examples in which the predicate is a 'pseudoconstruct' (*ʾiḍāfa ġayr ḥaqīqiyya*) adjective phrase (3a), or an adjective modified by the specificational accusative (→ *tamyīz*) (3b).

(3a) *ḥasan-un* [$_{\text{AP}}$ *jamīl-u* *l-wajh-i]*
Hasan beautiful.ms-Nom the-face-Gen
'Hasan is beautiful of face [i.e., Hassan has a beautiful face]'
(3b) *ḥasan-un* [$_{\text{AP}}$ *jamīl-un* *wajh-an]*
Hasan beautiful.ms-Nom face-Acc
'Hasan is beautiful face-wise [i.e., Hassan has a beautiful face]'

In both (3a) and (3b), the subject does not bind an argument position in the lexical predicate *jamīl* 'beautiful' itself, but rather an implicit argument position in the possessor or the accusative-marked noun. For example, in (3a), *ḥasan* binds a possessor argument in the interpretation of *wajh* 'face'. This is schematized in (4) with the binding relation represented as lambda abstraction over the possessor argument in the interpretation of *wajh* 'face'.

(4) λP_{et}.[P(Hasan)](λx.[x's face is beautiful])

These examples show that, even in simple nominal clauses, the predication relation cannot be reduced simply to the application of a noun to a one-place lexical predicate. Instead, a co-indexing or abstraction mechanism is involved that takes a fully saturated syntactic category and returns it as a one-place function (see Williams 1980, 1983; Heycock 1991; Doron and Heycock 1999; Alexopoulou a.o. 2004; Heycock and Doron 2003). The (as yet) unanswered question is whether this abstraction mechanism is inherent in the predication relation, or is an independent part of the grammar of Arabic.

3. PREDICATION IN VERB-FIRST WORD ORDERS

Another sense of predicate has been used that does not describe a partition in the structure of a sentence but rather describes its semantic head. In this sense, ‘predicate’ is a hyperonym for lexical categories such as verbs, adjectives, and prepositions, which can be one-place, two-place, or three-place relations, of which the subject is the most salient or external argument (see in particular Williams 1980, 1983, 1984, 1994; with reference to Syrian Arabic, see Cowell 1964:380).

In practice, although different categories are referred to as predicates, the term ‘predication’ continues to refer to the relationship between a one-place predicate and its external argument (Rothstein 1983; Heycock 1991). Predication in this sense may correspond more closely to the relationship between the Arabic terms *fāʿil* and *fiʿl* or to *musnad* and *musnad ʾilayhi*.

Talking about predication in this way has allowed some theorists to describe even verb-initial clauses as involving predication. For example, according to the widely influential ‘predicate-internal subject hypothesis’ (Fukui and Speas 1986; Koopman and Sportiche 1991; Plunkett 1993; Mohammad 1999), → verb phrases and other lexical categories take subject arguments (see Stowell 1983), or are dominated by an abstract functional projection such as Bowers’ (1993) ‘Predication Phrase’ (PredP), which provides a subject position for a lexical predicate.

Mohammad (2000) argues that the underlying word order in both Modern Standard Arabic and Palestinian Arabic is SVO rather than VSO (see also → nominal clause and → verb phrase). This is because the structure of the verb phrase places the subject in a specifier position, c-commanding the verbal predicate. For example, (5a) has the underlying SVO structure in (5b), from which is derived the VSO order in (5c) by movement of the predicate head to the head position of the IP projection.

(5a) *kullu* *ʾumm-in*
every-Nom mother-Gen
tuḥibbu *ibna-hā*
love.Imperf.3fs son-cl3fs
‘Every mother loves her son’ (Modern Standard Arabic; Mohammad 2000:75)

(5b) $[_{\text{IP}}$ PRES $[_{\text{VP}}$ $[_{\text{NP}}$ kullu ʾumm-in] $[_{\text{V}}$’ tuḥibbu $[_{\text{NP}}$ ibna-hā]]]]

(5c) $[_{\text{IP}}$ $[_{\text{I}}$ PRES tuḥibbu$_{j}$] $[_{\text{VP}}$ $[_{\text{NP}}$ kullu ʾumm-in] $[_{\text{V}'}$ t$_{j}$ $[_{\text{NP}}$ ibna-hā]]]]

According to Mohammad’s and related analyses, then, predication is established within the VP projection, and the relation is then rendered opaque by the derivation of VS word order.

4. SECONDARY PREDICATES

A distinction is made in the literature between primary and secondary predication (Williams 1983, 1984, 1994; Rothstein 1983; Stowell 1983; Hornstein and Lightfoot 1987; Heim and Kratzer 1998, among others), where primary predication results in a full clause, while secondary predication predicates a subordinated constituent of an argument of a matrix or higher clause.

Examples of secondary predicates in Arabic include: circumstantial or → *ḥāl* predicates, used as complements of verbs of change or causation (6a–b), perception (6c), belief or knowledge (6d), or as adverbial modifiers (6e); and in the dialects, optative predicates (6f) and, according to some analyses, obligatory control complements (6g).

(6a) *ittaxaḏa-hā* *zawjat-an*
take.Perf.3ms-cl3fs wife-Acc
la-hu
to-cl3ms
‘He took her as his wife’ (Modern Standard Arabic; Badawi a.o. 2004:376)

(6b) *yajʿalu-hu* *yaṭmaʾinnu*
make.Imperf.3ms-cl3ms feel-confident.Imperf.3ms
ʾilā *l-mustaqbal-i*
to the-future-Gen
‘It makes him feel confident about the future’ (Modern Standard Arabic; Badawi a.o. 2004:377)

(6c) *šaʿarat* *bi-qalb-i-hā*
feel.Perf.3fs with-heart-Gen-cl3fs
yuḫfiqu *bi-šiddat-in*
beat.Imperf.3ms with-strength-Gen

‘She felt her heart beating strongly’ (Modern Standard Arabic; Badawi a.o. 2004:581)

(6d) *wajad-nā* *baʿḍ-a* *l-masājid-i*
find.Perf.1p some-Acc the-mosques-Gen
taṭlubu *t-tabarruʿ-a*
require.Imperf.3fs the-charity
‘We found some of the mosques in need of charity’ (Modern Standard Arabic; Badawi a.o. 2004:378)

(6e) *fa-yuhuzzu* *raʾs-a-hu*
and-shake.Imperf.3ms head-Acc-cl3ms
muṭmaʾinn-an *l-ī*
reassure.Act.Part.ms-Acc to-cl1s
‘…and he would shake his head at me reassuringly’ (Modern Standard Arabic; Badawi a.o. 2004:580)

(6f) *ʾaʿṭat-e* *čalb*
give.Perf.3fs-cl3ms dog
iḥāmi *ʿan-ne*
protect.Imperf.3ms for-cl3ms
‘She gave him a dog to protect him’ (Bir Zeit Palestinian; Schmidt and Kahle 1918:§42.16)

(6g) *bidd-u* *yḥāwil*
want.3ms try.Imperf.3ms
ytafāham *maʿā*
agree.Imperf.3ms with-cl3ms
‘He wants to try to reach an understanding with him’ (Jordanian; LDC2005S14: fsa_12424, 434.82–437.57)

The only detailed study to date of secondary predication in Arabic is Mohammad’s (1999) examination of asyndetic *ḥāl*-predicates in Modern Standard Arabic and Palestinian Arabic. Mohammad discusses constraints on the potential controller for the *ḥāl*-predicate, noting that, in Modern Standard Arabic, *ḥāl*-predicates headed by an active participle can be controlled either by the subject or object of the matrix clause.

For example, (7) can mean both that Zayd was riding the donkey when Ahmad saw him, or that Ahmad was riding the donkey when he saw Zayd.

(7) *šāhada* *ʾaḥmad-u* *zayd-an*
see.Perf.3ms Ahmad-Nom Zayd-Acc
rākib-an *ḥimār-an*
ride.Act.Part.ms-Acc donkey-Acc
‘Ahmad saw Zayd [while Zayd was] riding a donkey’
‘Ahmad saw Zayd [while Ahmad was] riding a donkey’

In contrast, *ḥāl*-predicates headed by imperfect verbs can only be predicated of the object of the main verb (8a), as shown by the unacceptable feminine agreement form marked on the *ḥāl*-verb in (8b).

(8a) *zaynab-u* *šāhadat*
Zaynab.fs-Acc see.Perf.3fs
al-walad-a *yarkabu*
the-boy.ms-Acc ride.Imperf.3ms
ḥimār-an
donkey-Acc
‘Zaynab saw the boy [while he was] riding a donkey’

(8b) **zaynab-u* *šāhadat*
Zaynab.fs-Acc see.Perf.3fs
al-walad-a *tarkabu*
the-boy.ms-Acc ride.Imperf.3fs
ḥimār-an
donkey-Acc

Mohammad argues that the *ḥāl*-predicate is in a complement position and must be controlled by the nearest c-commanding argument, namely the object in a transitive clause and the subject in an intransitive clause. This is supported by the acceptability of extraction out of it when it is predicated of the direct object of the verb (9a), but not when predicated of the subject (9b).

(9a) *mōna,* *māḏā* *šāhadat*
Mona.fs what see.Perf.3fs
ʾaḥmad-a *rākib-an?*
Ahmad.ms-Acc ride.Act.Part.ms-Acc
‘What did Mona see Ahmad riding?’

(9b) **mōna,* *māḏā* *šāhadat*
Mona.fs what see.Perf.3fs
ʾaḥmad-a *rākibat-an?*
Ahmad.ms-Acc ride.Act.Part.fs-Acc
‘Mona, what did she see Ahmad [while she was] riding?’

Mohammad postulates that a *ḥāl*-predicate consisting of an imperfect verb can only be realized as a complement, while a *ḥāl*-predicate consisting of an accusative-marked participle can be realized either as a complement or an adjunct. Palestinian Arabic, on the other hand, lacks accusative case morphology, so a participial *ḥāl* must be realized as a complement, just like a verbal *ḥāl*, as in (10a).

(10a) *šāf* *zayd* *aḥmad*
see.Perf.3ms Zayd Ahmad
rākib *iḥmār*
ride.Act.Part.ms donkey
'Zayd saw Ahmad [while Ahmad was] riding a donkey'
*'Zayd saw Ahmad [while Zayd was] riding a donkey'

The restricted interpretation of predicative *ḥāl* contrasts with the more flexible interpretation of a full *ḥāl*-clause introduced by *wa*-, which can be controlled by either the subject or the object because it does not stand in a predication relation with its controller, as in (10b).

(10b) *šāf* *aḥmad* *zayd*
see.Perf.3ms Ahmad Zayd
[wa-hu *rākib* *iḥmār]*
and-he ride.Act.Part.ms donkey
'Ahmad saw Zayd while [Zayd was] riding a donkey'
'Ahmad saw Zayd while [Ahmad was] riding a donkey'

The implication of Mohammad's argument is that predication can be established both by lexical specification in the case of complement predicates, and by syntactic specification in the case of adjunct predicates.

5. Predicates and broad subjects

As noted above, the term nominal clause (*jumla ismiyya*) appears to describe two different kinds of structure: a subject/predicate predication structure when the initial NP corresponds to the subject or agent of the report constituent, and a topic/comment dislocation structure when the initial NP is distinct from the subject or agent of the report constituent.

However, Doron (1996), Doron and Heycock (1999), Alexopoulou a.o. (2004), and Heycock and Doron (2003) show that a non-subject-initial NP can have grammatical properties associated with subjects. A non-subject-initial NP can be embedded under an 'exceptional case-marking' (ECM) verb or *ʾanna*, as in (11).

(11) *naʿtaqidu* *ʾanna* *ʾamrīkā*
believe.Imperf.1p that America
šaʿb-u-hā *ḏakiyy*
folk.ms-Nom-cl3fs clever
'We believe that the American people are clever' (Modern Standard Arabic; http://maxforums.net/showthread.php?t=96271)

In a copular clause with *kāna* 'to be', the initial NP (*ism kāna*) can be predicated of a fully saturated lexical predicate, as in (12).

(12) *kān* *il-bayt*
be.Perf.3ms the-house.ms
ʾalwān-o *fātḥa*
colors.fp-cl3ms bright.fs
'The house was brightly colored' (Lebanese Arabic; Alexopoulou a.o. 2004:333)

A clausal *xabar* can be conjoined with a lexical predicate that requires a subject. Hence, by the Coordinate Structure Constraint (Ross 1967), the initial NP must be a subject with respect to the clausal *xabar*, as in (13).

(13) *sayyart-i* *lōn-a* *fāteḥ*
car.fs-cl1s color-cl3fs bright.ms
u-maftūḥa *min* *fōʾ*
and-open.fs from above
'My car is brightly colored and a convertible' (Lebanese Arabic; Alexopoulou a.o. 2004:336)

The initial NP need not be referentially specific (unlike topics; → specificity), and can instead be a question word (14a) or a quantificational noun phrase (14b–c).

(14a) *mīn* *šaʿr-ha* *ʾaṭwal*
who hair.ms-cl3fs longer.ms
min *šaʿr* *il-banāt di?*
from hair the-girls these
'Whose hair is longer than these girls' hair?' (Egyptian Arabic; www.adwforums.com/archive/index.php/t–16652.html)
(14b) *iḥna,* *kull* *wāḥda*
we every one.fs
šaʿr-ha *la-nuṣṣ* *ḏ̣ahr-ha*
hair-cl3fs to-middle back-cl3fs

‘Each one of us has hair to the middle of her back’ (Levantine Arabic; www.rhalchat.net/vb/showthread.php?t=3858&page=4)

(14c)	*wala*	*wāḥda*
	not.even	one
	šaʕr-ha	*ḥalu*
	hair.ms-cl3fs	nice.ms

‘Not one of them has nice hair’ (Levantine Arabic; http://alfrasha.maktoob.com/archive/index.php/t–202674.html)

Doron and Heycock conclude that the initial NPs in examples of this kind are grammatical subjects even though they do not bind external arguments, and for this reason, they argue that Arabic allows a primary predication to be established between a subject and a ‘derived predicate’.

This recalls the observation in Section 2 that predication of a specified modifier requires a grammatical mechanism that abstracts over a pronoun or implicit argument inside the clause. For example, the primary predication in (13) would have a logical structure like (15c), derived by abstracting over the possessive pronoun in *lōn-a* ‘its color’ (15a) and conjoining this abstract with the interpretation of *maftūḥ min fōʾ* ‘open from above’ (15b).

(15a) [[lōn-a fāteḥ]] = λx.[x’s color is bright]
(15b) [[maftūḥa min fōʾ]] = λx.[x is open-topped]
(15c) λP.[P(my car)](λx.[x’s color is bright & x is open-topped])

The fact that both primary and secondary predicates in Arabic can be derived via an abstraction mechanism either implies that there is crosslinguistic variation in the grammatical properties of predication, or that languages like Arabic are able, for some reason, to use relative-clause-like predicates, where languages like English can only use lexical predicates.

BIBLIOGRAPHICAL REFERENCES

Alexopoulou, Theodora, Edit Doron, and Caroline Heycock. 2004. “Broad subjects and clitic left dislocation”. *Peripheries: Syntactic edges and their effects*, ed. David Adger, Cécile de Cat, and George Tsoulas, 329–358. Oxford: Oxford University Press.

Badawi, Elsaid, Michael G. Carter, and Adrian Gully. 2004. *Modern Written Arabic: A comprehensive grammar*. London: Routledge.

Bakir, Murtada. 1980. *Aspects of clause structure in Arabic*. Bloomington: Indiana University Linguistics Club.

Benmamoun, Elabbas. 2000. *The feature study of functional categories: A comparative study of Arabic dialects*. Oxford: Oxford University Press.

Bloomfield, Leonard. 1916. “Subject and Predicate”. *Transactions and Proceedings of the American Philological Society* 47.13–22.

——. 1933. *Language*. Chicago: University of Chicago Press.

Bowers, John. 1993. “Predication”. *Linguistic Inquiry* 24.591–656.

Cowell, Mark. 1964. *A reference grammar of Syrian Arabic*. Washington, D.C.: Georgetown University Press.

Culicover, Peter and Wendy Wilkins. 1986. “Control, PRO, and the projection principle”. *Language* 62.120–153.

Diesing, Molly and Eloise Jelinek. 1995. “Distributing arguments”. *Natural Language Semantics* 3.123–176.

Doron, Edit. 1996. “The predicate in Arabic”. *Studies in Afroasiatic linguistics*, ed. Jacqueline Lecarme, Jean Lowenstamm, and Ur Shlonsky, 77–87. Leiden: Holland Academic Graphics.

——, and Caroline Heycock. 1999. “Filling and licensing multiple specifiers”. *Specifiers*, ed. David Adger, Susan Pintzuk, Bernadette Plunkett, and George Tsoulas, 69–89. Oxford: Oxford University Press.

Eid, Mushira. 1991. “Verbless sentences in Arabic and Hebrew”. *Perspectives on Arabic linguistics*, III ed. Bernard Comrie and Mushira Eid, 31–61. Amsterdam and Philadelphia: J. Benjamins.

Eisele, John C. 1999. *Arabic verbs in time: Tense and aspect in Cairene Arabic*. Wiesbaden: O. Harrassowitz.

Fassi-Fehri, Abdelkader. 1993. *Issues in the structure of Arabic clauses and words*. Dordrecht: Kluwer Academic Publishers.

Fukui, Naoki and Margaret Speas. 1986. “Specifiers and projections”. *MIT Working Papers in Linguistics*, VIII, ed. Naoki Fukui, Tova Rappaport, and Elizabeth Sagey, 128–172. Cambridge, Mass.: MIT Working Papers in Linguistics.

Heim, Irene and Angelika Kratzer. 1998. *Semantics in generative grammar*. Boston: Blackwell.

Heycock, Caroline. 1991. *Layers of predication: The non-lexical syntax of clauses*. Ph.D. diss., University of Pennsylvania.

——, and Edit Doron. 2003. “Categorical subjects”. *Gengo Kenkyu* 123.95–135.

Hornstein, Norbert and David Lightfoot. 1987. “Predication and PRO”. *Language* 63.23–52.

Jelinek, Eloise. 1981. *On defining categories: AUX and PREDICATE in Colloquial Egyptian Arabic*. Ph.D. diss., University of Michigan.

——. 2002. “Agreement, clitics, and focus in Egyptian Arabic”. *Themes in Arabic and Hebrew Syntax*, ed. Jamal Ouhalla and Ur Shlonsky, 71–105. Dordrecht: Kluwer Academic Publishers.

Koopman, Hilda and Dominique Sportiche. 1991. “The position of subjects”. *Lingua* 85.211–258.

Levin, Aryeh. 1981. “The grammatical terms *al-musnad*, *al-musnad ʾilay-hi* and *al-ʾisnād*”. *Journal of the American Oriental Society* 101.145–165.

Mohammad, Mohammad. 1999. "Anaphoric agreement in non-finite clauses in Arabic". Ms., University of Florida at Gainesville. Available from the Semitic Linguistics Research Archive: www.usc.edu/schools/college/semitic/.

——. 2000. *Word order, agreement, and pronominalization in Standard and Palestinian Arabic.* Amsterdam and Philadelphia: J. Benjamins.

Ouhalla, Jamal. 1991. *Functional categories and parametric variation.* London: Routledge.

——. 1994. "Verb movement and word order in Arabic". *Verb movement*, ed. David Lightfoot and Norbert Hornstein, 41–72. Cambridge: Cambridge University Press.

Plunkett, Bernadette. 1993. "The positions of subjects in Modern Standard Arabic". *Perspectives on Arabic linguistics*, V, ed. Mushira Eid and Clive Holes, 231–260. Amsterdam and Philadelphia: J. Benjamins.

Ross, J.R. 1967. *Constraints on variables in syntax.* Ph.D. diss., Massachusetts Institute of Technology, Cambridge, Mass.

Rothstein, Susan. 1983. *The syntactic forms of predication.* Ph.D. diss., Massachusetts Institute of Technology.

Schmidt, Hans and Paul Kahle. 1918. *Palästinensische Volkserzählungen aus Palästina*, I. Göttingen: Vandenhoeck und Ruprecht.

Stowell, Timothy. 1983. "Subjects across categories". *Linguistic Review* 2.285–312.

Williams, Edwin. 1980. "Predication". *Linguistic Inquiry* 11.203–238.

——. 1983. "Against small clauses". *Linguistic Inquiry* 14.287–308.

——. 1984. "Grammatical relations". *Linguistic Inquiry* 15.639–673.

——. 1994. *Thematic structure in syntax.* Cambridge, Mass.: MIT Press.

Frederick M. Hoyt
(University of Texas at Austin)

Pre-Islamic Arabic

1. The sources for pre-Islamic Arabic

Pre-Islamic Arabic is the cover term for all varieties of Arabic spoken in the Arabian Peninsula until immediately after the Arab conquests in the 7th century C.E. Scholars disagree about the status of these varieties (Rabin 1955). Three different points of view stand out. Some scholars (Nöldeke 1904, 1910; Fück 1950; Blau 1965; Chejne 1969; Versteegh 1984) assume that the language of pre-Islamic poetry and the *Qurʾān* was similar, if not identical, to the varieties spoken in the Arabian Peninsula before the emergence of Islam. If differences existed, they concerned mainly stylistic and minor points of linguistic structure.

A second group of mainly Western scholars of Arabic (Vollers 1906; Fleisch 1947; Kahle 1948; Rabin 1951; Blachère 1950; Wehr 1952; Spitaler 1953; Rosenthal 1953; Fleisch 1964; Zwettler 1978; Holes 1995; Owens 1998; Sharkawi 2005) do not regard the variety in which the *Qurʾān* was revealed as a spoken variety of Arabic in the peninsula. Some of them (Zwettler 1978; Sharkawi 2005) go so far as to state that the function of the language of pre-Islamic poetry and the *Qurʾān* was limited to artistic expression and oral rendition (→ poetic koine). Others are not as clear about the functional load of this variety in pre-Islamic times.

A third group of scholars (Geyer 1909; Nöldeke 1904, 1910; Kahle 1948) assume that the variety of Arabic of pre-Islamic poetry and the *Qurʾān* was the variety spoken by Bedouin Arab tribes and nonsedentary Arabs, at least in the western parts of the peninsula where trade routes existed.

Some modern scholars of Arabic believe that the Classical Arabic grammarians held their view, that the language of pre-Islamic poetry and the *Qurʾān* was identical with at least the spoken varieties of some Arab tribes in the peninsula (Rabin 1955:21–22; Sharkawi 2005:5–6). A first reading of the grammatical texts seems to confirm that grammarians were quite aware of the existence of different language varieties in the Arabic-speaking sphere. They distinguished terminologically between → *luġa* 'dialect' and → *lisān* 'language' (ʾAnīs 1952:16–17; Naṣṣār 1988:58). Among several meanings of the word *luġa* is the technical meaning of a linguistic variety (Rabin 1951:9). As early as the 2nd century A.H., grammarians were aware of differences among the dialects. Among the earliest writers on tribal dialects were Yūnus ibn Ḥabīb (d. 182/798) and ʾAbū ʿAmr aš-Šaybānī (d. 213/828), the author of the *Kitāb al-jīm*, in which odd and archaic lexical items used in certain tribes are recorded. In the 3rd century A.H., several authors are said to have written books on tribal dialects, among them al-Farrāʾ (d. 207/822), ʾAbū ʿUbayda (d. 210/825), and ʾAbū Zayd al-ʾAnṣārī (d. 215/830). In addition to treatises on the dialects, there were those on the dialect words in the

Qurʾān. Among the earliest authors in this field was Ibn ʿAbbās (d. 68/687), to whom a treatise under the title *Kitāb al-luġāt fī l-Qurʾān* was ascribed (Rippin 1981). In this treatise, both dialect and foreign words were listed according to their order of appearance in the *Qurʾān*. Several other treatises were dedicated to the same subject in the 3rd century A.H. (Naṣṣār 1988:61–62). Ibn Fāris (d. 395/1005; *Ṣāḥibī* 19) identified differences among dialects, such as differences in vowels, assimilation, the addition of vowels, realizing the *hamza*, gender marking, and plurals.

However, grammarians were not interested in the study of *luġāt* as such. Had they really decided to study the dialects, they would have paid attention to variation and recorded phenomena with an abundance of details and examples. Rather, they chose to concentrate on the variety pertinent to their field of inquiry: their study of Arabic began as an ancillary to the study of the *Qurʾān*. Furthermore, if dialects had consumed the scholars' interest, they would have used Bedouin as informants in order to find out whether certain features existed in some dialects but not in others, or how certain dialects behaved in certain contexts. Instead, they chose to use Bedouin Arabs to emphasize how their dialects realized certain features of the Qurʾānic language, as arbiters in theoretical disputes, and as a means to verify the data.

A good example of the role of Bedouin is the famous story of the scholarly debate between Sībawayhi and the Kufan grammarian al-Kisāʾī (Zubaydī, *Ṭabaqāt* 68–71, in Bernards 1997:6). According to this story, at the court of the caliph ar-Rašīd, al-Kisāʾī and Sībawayhi could not settle a theoretical point, and they had to submit the case to Bedouin Arabs for arbitration. One Bedouin, who was waiting at the door, was admitted, and he favored al-Kisāʾī's judgment over that of his adversary.

The *Qurʾān* stresses that it was revealed in a 'clear' Arabic tongue (e.g. Q. 16/103, 26/195). Therefore, the *luġāt* in the text were marginal in comparison to pre-Islamic poetry, although this belonged to the same linguistic level as the *Qurʾān*. This may be the reason why the number of *šawāhid* 'evidentiary verses' from poetry outgrew the number of *šawāhid* taken from Bedouin speech in the grammarians' works. The majority of *šawāhid* in Sībawayhi's *Kitāb* were verses from the *Qurʾān* and lines of poetry; the grammarian al-Jarmī (d. 225/839) mentions a number of 1,050 lines of poetry in the *Kitāb* (*Xizāna* I, 8, quoted in Hārūn's introduction to the edition of the *Kitāb*). The purity of the Qurʾānic language served for grammarians as the touchstone for linguistic correctness, and the relative distance of dialects from this purity determined their functional value. The practical preference for certain Bedouin dialects over others in grammatical arbitration was due to the grammarians' focus on the similarities of these dialects to the *Qurʾān*.

There are four main sources for pre-Islamic Arabic. Books of a general nature from the Abbasid period, books of Classical Arabic grammarians, and the *Qurʾān* provide us with texts varying in length and usefulness; in addition, dictionaries contain a huge, albeit unorganized and unclassified, reservoir of lexical data on dialectal variation. Apart from the *Qurʾān*, these sources provide us with four types of data: speeches by pre-Islamic notables and famous orators; anecdotes from the utterances of seers and fortune tellers; proverbs (such as those in Jāḥiẓ, *Bayān* I, 184); and stories (such as in Ibn ʾIsḥāq, *Sīra* I, 321). From these sources some phonological, morphological, and syntactic variables stand out as indicators to the existence of dialects in the pre-Islamic Arabic sphere.

For grammarians in the first three centuries of the Islamic era (see Rabin 1951:6), the task of recording dialect features was marginal compared to the main target of codifying the features of Arabic. Thus, variable features and dialectal references to tribes occur rather at random in the books of grammar. However, Rabin (1951), ʾAnīs (1952), and al-Gindī (1983) collected references to tribal features in dialect sketches. In the next section these features are organized according to dialect.

2. The dialects of pre-Islamic Arabic

2.1 *The dialect of the Ḥijāz*

Ḥijāzī Arabic features appear in the grammarians' books more frequently than features of any other dialect. It is, therefore, a much better represented dialect in comparison to others, despite the fact that the region's geographical definition is not as clear. In pre-Islamic times, the Ḥijāz was the western part of the peninsula, between the Tihāma in the southwest and the Najd in the east. It included the Banū Sulaym and the Banū Hilāl. In the north was the territory of Bālī, and in the south that of Huḏayl. After the advent of Islam, the Tihāma was included in the Ḥijāz, thus the Bedouin tribes in the interior were sometimes included in the Ḥijāz. It seems that for the grammarians, Ḥijāz referred to regions defined according to the post-Islamic demarcation. In this way, the urban centers of Mecca, Medina, and Ṯaqīf were included in that region. The term *luġa ʾahl al-Ḥijāz* covers all differences that may have existed within this region.

Phonological features of this region include:

i. The pronunciation of /ʿ/ as *hamza*.
ii. The use of the full forms of vowels, without elision or vowel changes, e.g. *ʿunuq* 'neck' as against *ʿunq* in Eastern Arabian dialects, where short unstressed vowels were elided.
iii. The absence of vowel harmony, which was realized in Eastern dialects, e.g. Ḥijāzī *baʿīr* 'camel', corresponding to Eastern *biʿīr*. By the same token, uvular and pharyngeal consonants assimilated following vowels in the Eastern dialects, while in the Ḥijāz they rested immune, e.g. Ḥijāzī *ʿuqr* 'the main part of the house', corresponding to Eastern *ʿaqr*. In the neighborhood of uvulars and pharyngeals, the Ḥijāz had /u/, while the Eastern dialects had /a/.
iv. The tendency to shorten the long final vowels in pause positions
v. The elision of the *hamza*.

Morphological features of this dialect include:

i. The 3rd person suffix pronouns *-hu, -humā*, *-hum*, and *-hunna* did not change to the *-hi* form after *i* or *ī*.
ii. For the singular relative pronoun, the Ḥijāz used *allaḏī* rather than the Western and Yemenite *ḏī* and *ḏū*. For the feminine plural, the Ḥijāz used *allāʾī*. The same form may have been used for the masculine plural as well.
iii. The dual suffix in the Ḥijāz may have had a single form, *-āni*, for the nominative, accusative, and genitive cases alike. Ibn Hišām (*Muġnī* I, 37), in his explanation of the nominative case of the demonstrative pronoun *hāḏāni* 'these two' in the verse *ʾinna hāḏāni la-sāḥirāni* (Q. 20/63), claimed that in the dialect of the Ḥijāz, these demonstrative pronouns were indeclinable.
iv. The absence of → *taltala*.
v. The imperative of geminated verbs was conjugated as the strong verbs, e.g. *urdud* 'respond!'.

Syntactic features of this dialect include:

i. Some nouns were feminine in the Ḥijāz and masculine in the Najd and Tamīm. Some examples are *tamr* 'dates', *šaʿīr* 'barley', *ṣirāṭ* 'path'. The word *ṣirāṭ* appears in the first *sūra* of the *Qurʾān* (Q.1/6) followed by a masculine adjective (*ṣirāṭ mustaqīm*).
ii. In the Ḥijāz, the predicate of verbal sentences agreed in number with the head verb (known as the *luġa ʾakalūnī l-barāġīṯ*), unlike Standard Arabic, where the head verb is always in the singular.
iii. In the Ḥijāz, after the shortened forms *ʾin* and *ʾan*, the subject took an accusative case, while in Classical Arabic and in the east, shortened particles lost their effect on the following nominal clause.
iv. After the complementizer *ʾinna, ʾanna*, etc. (→ *ʾinna wa-ʾaxawātuhā*), the Ḥijāzī dialect put the subject and predicate of the sentence in the accusative case. Ibn Hišām (*Muġnī* I, 36) explains the agreement in case between the subject and predicate in a nominal sentence after *ʾinna* 'in one version of a *ḥadīṯ* (*ʾinna qaʿra jahannama sabʿīna xarīfan*) by saying that the Ḥijāz did not distinguish between the subject and predicate in case endings after *ʾinna*.
v. The predicate of *kāna* and other copulas (→ *kāna wa-ʾaxawātuhā*) was given a nominative case, while an accusative case is assigned to it in Classical Arabic.
vi. In the Ḥijāz, *mā, lā*, and *ʾin* had the same effect as the Classical Arabic *laysa* in assigning to the subject the nominative case and to the predicate the accusative case.
vii. Verbs in the indicative were used after *ʾan*. An example comes from Mujāhid (d. 104/722), who read the verse *li-man ʾarāda ʾan yutimma r-raḍāʿata* 'for those who want the suckling (period) to be completed' with an indicative ending, *yutimmu* (Q. 2/233).

2.2 *The dialect of ʾAzd*

The ʾAzd dialect is rarely mentioned in the literature. Whereas anecdotes and *šawāhid* from other Yemeni dialects are given, the dialect of ʾAzd receives little attention. More confusing still is the fact that there were two tribes by the name of ʾAzd, one in Oman and the other in the western part of Yemen. The two features that are mentioned, however, show the difference between this dialect and the rest of Yemen.

i. The retention of the nominal case endings *a*, *i*, and *u* in the pausal position.
ii. The retention of the vowel *a* in the prefixes of the imperfect, e.g. *yaktub* 'he writes' as against the → *taltala* in other dialects.

2.3 *The dialect of Huḏayl*

The tribe of Huḏayl was situated in the southeastern part of the Ḥijāz, to the north of Yemen and to the northeast of ʾAzd. Its location in the southeast of the Ḥijāz connected this tribe geographically to the Eastern dialect group, which earned the tribe its fame for speaking well-formed Arabic. Despite this connection with the east, the dialect of Huḏayl belonged mainly to the Western group and functioned as an intermediate zone between the Ḥijāz and northern Yemen (Rabin 1951:79). The evidence for this claim comes from the grammatical and lexical features it shared with the Western group. They shared, for instance, *ʾawwāb* 'obedient' and *jadaṯ* 'tomb' with Kināna. Other features mentioned by the grammarians include:

i. The insertion of short unstressed vowels in the middle of words, e.g. *ibin* 'son' instead of Classical Arabic *ibn*, and *jawazāt* 'nuts', sg. *jawza*. In Classical Arabic, words with a singular pattern *faʿla* receive an anaptyctic vowel *a* in the feminine plural, to become *faʿalāt*. This vowel is not added when the second radical in the root is *w* or *y*, but Huḏayl added an anaptyctic vowel to roots containing *w* and *y* as well.
ii. The absence of vowel harmony.
iii. The absence of the *hamza*.
iv. It is probable that in Huḏayl the final long vowels were shortened, as was the case in the Ḥijāz.
v. The change of the glides *wu* and *wi* into the long vowels *ū* and *ī*, respectively.
vi. Huḏayl used the relative pronoun *allaḏī*. The plural of this pronoun was *allaḏūna*, in all numbers and genders, in opposition to Classical Arabic, which uses *allaḏīna*.
vii. Concerning the *taltala* feature, Huḏayl was claimed to have used both forms: *-a-* imperfect like the Ḥijāz dialects, and *-i-* imperfect like the eastern tribes. This variation is also common in Ṭayyiʾ. Both tribes had contact with eastern tribes, which may explain the variation.

2.4 *The dialect of Ṭayyiʾ*

The Ṭayyiʾ tribe was situated in the north of the Najd. It occupied the southern frontiers of the Nufūd desert and was also situated toward the northeast of the Ḥijāz region. It shared with the tribes of the eastern part some linguistic features, such as the *taltala*. Rabin (1951:193) claims that such common features are suggestive of the connecting role this tribe played between the dialects of the eastern and western parts of the peninsula. The territory of Ṭayyiʾ during the early Islamic period was not the original habitat of the tribe. The tribe was traditionally known to have migrated from northern Yemen together with the tribes with which it shared some linguistic features. Features of this dialect include:

i. The weakening of the final syllable and elision of final nasals, laterals, *t*, and/or *y*.
ii. The absence of vowel harmony and vowel elision.
iii. The change of /ʿ/ into /ʾ/, e.g. *daʾ-nī* 'let me'; no other data about depharyngealization are available.
iv. The fate of *hamza* in this dialect is not known due to the absence of direct evidence.
v. The suffix pronoun of the 3rd person feminine in pause was *-ah* and *-hā* in context, which is in accordance with the Classical and Eastern Arabic weakening of final syllables.
vi. The form of the article was *am-*.
vii. The singular feminine demonstrative was *tā*, not *hāḏihi*.
viii. The relative pronoun was *ḏū*, which was used for the two genders and all numbers.
ix. The *-t* of the feminine plural was dropped in pause; again, this is in harmony with the weakening of final syllables.

x. az-Zajjājī (*Šarḥ* 152) claims that as in the Ḥijāz, the predicate of verbal sentences agreed in number with the head verb.

2.5 *The Arabic of Yemen*

The dialect of Yemen was very well represented in the writings of the grammarians because of the special interest it held for the scholars of the 3rd and 4th centuries A.H., especially for lexicographers like Ibn Durayd (d. 321/933) and Našwān (d. 573/1178). Although home to a host of South Arabian dialects, Yemen does not reflect much South Arabian influence, except for some lexical items that may be mere loanwords from that language. A good example is the word *baʿl* 'lord', which is still common in Mehri (Rabin 1951:25–27).

During the time of al-Hamadānī (d. after 360/971), the main source on Yemen, a dialect similar to the Central Arabian Bedouin dialects was spoken in the region east of Sarāt and in the extreme south. Al-Hamadānī describes these dialects as 'correct' Arabic. In the central and western regions of the Sarāt, different dialects were spoken. These dialects are characterized by al-Hamadānī as *mutawassiṭ* 'middle'. Rabin (1951:45) claims that this attribute must mean that they were mixtures of Arabic and → Ḥimyaritic. In the southern part of Sarāt and the mountains around Ṣanʿāʾ, the language showed strong traces of Ḥimyaritic. In the area to the west, a mixture of Arabic and Ḥimyaritic was spoken. In the villages, however, Ḥimyaritic was predominant. Outside the villages, in the nomadic areas, West Arabian dialects were spoken (Rabin 1951:45). Thus, there were two linguistic communities in Yemen, apart from the Bedouin in the east. The first was that of the settled farmer groups, which spoke a mixture of Ḥimyaritic and Arabic, while the other group consisted of the nomadic people who spoke West Arabian dialects. Although the Yemeni dialects spoken in this region were very similar to other Arabic dialects, Arabs considered them incomprehensible. There are several anecdotes in the literature showing that Arabs did not consider the dialects of Ḥimyar Arabic to be similar to their own. The attribute *ṭumṭumāniyya* was given in the literature to the Ḥimyaritic dialect as a form of mockery.

The northern Yemen region hosted tribes speaking dialects so similar to each other that they could be considered a defined group. This group was different from the rest of Yemen in the south and Huḏayl and the Ḥijāz in the north. Despite being distinct from both groups, the dialects of northern Yemen exhibited similarities with both. Rabin (1951:64) claims that because grammarians often ascribed Ḥijāzī dialect features to Kināna, this region can be considered as an extension to the West Arabian dialect group. Among the tribes that lived in this region were Kināna, Xaṯʿam, Hamadān, ʿAnbar, Zubayd, and Murād. The first four of these tribes are frequently mentioned in literature, but whenever a feature is mentioned as belonging to a certain tribe, it may have applied to the rest of the tribes as well. Rabin (1951:64) also assumes that whenever the grammarians mention the tribes of Yemen, they mean these tribes living in the northern part. Among the features mentioned for these dialects are the following:

i. The absence of → *ʾimāla*. Al-Hamadānī, however, states that the Bedouin tribe of Banū Ḥarb in the south realized *ʾimāla*.

ii. The realization of *hamza*. However, in some cases the original *hamza* of the word was changed into the glide *w*. An example is *ʾātaytu*/*wātaytu* 'I obeyed'. This feature is still heard in some modern dialects.

iii. In some Yemenite dialects, the feminine ending *-at* was generalized to pause positions. Yemenite dialect words may have received *tanwīn* even in the pause position.

iv. The definite article of the Yemenite dialect was *am-*. Unlike the Arabic definite article *al-*, it was not assimilated to dental and sibilant consonants. Words that received this article could also be given *tanwīn*. An example is found in al-Fīrūzābādī's *Muḥīṭ* (I, 37): *mani m-qāʾimun* 'who is standing?'

v. The dual suffix in northern Yemen, *-āni*, was suffixed to the noun. Although other tribes in the peninsula used a single dual ending as well, they coupled it with a different treatment of the final short vowel. They either used *-āna* as a fixed form or inflected the ending. This feature was ascribed to Dabba in the northwest of the Empty Quarter, which shows that this feature cut across dialect boundaries.

vi. There was a sentence-initial particle *ʾam* that was used with the verb in the imperfect (Rabin 1951:37).

vii. In southern Yemen, especially in Ḍufār, the demonstrative pronoun for both genders was *ḏī*, which followed the noun it modified, e.g. *iš-šuġl ḏī* 'this work' (Rabin 1951:75).

viii. The relative pronoun was *ḏī*, without distinction for gender or number. It was used in western Ḥaḍramawt and elsewhere. In other places of Yemen and as far north as Huḏayl, the Classical Arabic pronoun *allaḏī* was used, but without distinction for number or gender.

ix. The negative particle was *dū*. Another form, still used in Taʿizz, in the southernmost part of Yemen, is *daʾ*. This particle may stem from Ḥimyaritic, since a particle *daʾ* was found in some of the South Arabian inscriptions around the middle of the 6h century C.E.

x. The suffix of the 1st and 2nd persons of the verb in the perfect is *-k*, not *-t*. A good example is the saying of a woman: *raʾayku bi-ḥulm kawaladku ibnan min ṭīb* 'I saw in a dream that I gave birth to a son of gold'. The verbs *raʾayku* 'I saw' and *waladku* 'I gave birth' end in this suffix. The same use is still current in the Yemeni countryside (→ Yemen).

3. DIALECTAL DIFFERENCES AND LINGUISTIC CHANGE IN PRE-ISLAMIC ARABIC

The features of the pre-Islamic Arabic dialects that have been listed above show that the dialectal elements are random and inconclusive. However, the evidence suggests that some of the pre-Islamic dialects exhibited a tendency toward variation and that there was a certain dialect grouping. The dialects of the Ḥijāz and Yemen exhibit elements of agreement that group them together against the Eastern dialects of the Arabian Peninsula and Classical Arabic. On the phonological level, most of these dialects elided the *hamza*, except for parts of Yemen. Also, in the dialects of the Ḥijāz and Yemen, there was no *ʾimāla* or vowel harmony, and they share a tendency to change diphthongs into long vowels: northern Yemen changed /ay/ into /ā/, and Huḏayl changed /wu/ and /wi/ into /ū/ and /ī/.

In morphology, the dialects of the Ḥijāz and Yemen shared some similar tendencies with different realizations. With the exception of Ṭayyiʾ, all West Arabian dialects retained final morphemes unchanged in the pause position. In Yemen, the final *-t* of the feminine ending was not deleted in pause, and the nouns also retained *tanwīn* in pause. In ʾAzd, nouns retained case endings in pause. In the Ḥijāz, the final vowel at the end of the 2nd person singular pronoun was not elided in final pause position.

There are also features distinguishing dialects from one another. On the phonological level, the southern part of Yemen realized the *hamza*, as opposed to the rest of this group of dialects that elided it. The phoneme /ʿ/ was treated differently by each dialect. In Yemen, it may have been pronounced with a degree of nasalization. It was depharyngealized and pronounced as *hamza* in the Ḥijāz and Ṭayyiʾ.

In morphology, there was variation in the use of the demonstrative pronoun. In Yemen, the particle for both genders was *ḏī*, which was postpositioned to the definite noun. But in the Ḥijāz, each gender had its own demonstrative pronoun. The relative pronoun was another area of variation among the dialects. In southern Yemen and western Hadramawt, the relative pronoun was *ḏī*, without distinction of gender and number, whereas in northern Yemen, *allaḏī* was used without distinction in number and gender. Huḏayl, like northern Yemen, used *allaḏī* as a relative pronoun for the singular and *allaḏūna* for the plural. As was the case with Huḏayl, the Ḥijāz used *allaḏī* for the singular but had *allāʾī* for the feminine and, probably, the masculine plural as well.

The data show that there were certain tendencies toward language change. Especially remarkable were the sound changes, in both east and west Arabia. In Yemen, Huḏayl, the Ḥijāz, and Ṭayyiʾ, there was a tendency to change the pharyngeal sounds. In Yemen, /ʿ/ was changed into a *hamza*. The same change took place in both the Ḥijāz and Huḏayl. It is not clear from the data, however, whether there was the same conditioning for this change in the Ḥijāz and Huḏayl. However, /ʿ/ was changed into /ḥ/ in the same context by Saʿd ibn Bakr and in the area around Medina.

The phoneme /ḥ/ underwent lenition in the Ḥijāz, northern Yemen, and Huḏayl; it was almost completely devoid of pharyngeal friction and went in the direction of /h/. All the examples we have for this change in the Ḥijāz seem to involve the condition that for the change to take place, /ḥ/ must precede the open low short vowel /a/.

There is only one example of fortition in the data. In Yemen, the voiced palatal fricative consonant of Classical Arabic /j/ was realized as a voiced palatal stop consonant /g/.

There are also indications for → anaptyxis, whereby a vowel is inserted in a consonant cluster. Such short vowels were current in Huḏayl and the Ḥijāz. A good example is the word *ibin* 'son', with a vowel *i* between the consonants *b* and *n*. This phenomenon goes together with the general tendency in the western part of the peninsula to preserve short unstressed vowels in the middle of words and to prevent word-final consonant clusters. Dialects of the east, on the other hand, tended to delete unstressed high front and back vowels *i* and *u*.

The semivowels, in both parts of the peninsula, underwent changes when in the vicinity of vowels. The data indicates that the semivowel /w/ in Huḏayl was deleted when it preceded high vowels. This change may have caused the compensatory lengthening of the following vowel.

Change also extended to morphological and syntactic elements. On the syntactic level, there is a difference between the rules of Classical Arabic and the dialect of the Ḥijāz, in particular, and other dialects in the western part of the peninsula in general. Since the dialects of the east show a greater similarity to the standardized variety of Classical Arabic, we may assume they were more conservative than the Western dialects, because both seem to be more elaborate, especially in the field of morphosyntax. Certain syntactic developments in the line of uniformity and category reduction happened in Ḥijāz and the Western dialects. Among them were those directed toward altering the 'effect' (→ *ʿamal*) of certain 'operators' (*ʿawāmil*) on the nominal sentences they modify. According to the rules of Classical Arabic and the Eastern dialects, after the copular verb *kāna* 'to be', the subject of the following nominal sentence is in the nominative case, while the predicate is in the accusative (→ *kāna wa-ʾaxawātuhā*). In the Ḥijāz, however, both constituents of the sentence were in the nominative. The same generalization of case happened with the complementizer *ʾinna* (→ *ʾinna wa-ʾaxawātuhā*). In the standardized variety of Arabic, the subject of the nominal sentence governed by *ʾinna* is in the accusative while the predicate is in the nominative. In Ḥijāz, both constituents were put in the accusative. The same development toward overgeneralization of case endings affected verbs and verbal sentences. After *ʾan*, the Ḥijāzī dialect put the verb in the indicative, while the rules of the standardized variety prescribe a subjunctive here. In verbal sentences, verbs in the Ḥijāz agreed in number with their agents, as opposed to Classical Arabic, which limits the agreement between the verb and its agent to gender.

Within the general division of West Arabian dialects, some tribes shared features with the tribes of the eastern part of the peninsula. The clearest example was Huḏayl, which, like Eastern dialects, realized the *hamza*. Likewise, in Tihāma, elision affected short, unstressed vowels as in the dialects of the east, producing contracted combinations of preposition and article like *mil* and *ʿal* instead of *min al-* and *ʿalā al-*.

Some indications point to the progress of innovations in the peninsula. The data show that the Ḥijāz was at the center of innovations. On the phonemic level, the Western dialects were moving toward a more balanced system. Single voiceless sounds like the *hamza* were elided, and the pharyngeal /ʿ/ was moved from its place of articulation. It appears that the *hamza* was elided in the Ḥijāz in all environments but was retained in Yemen, except when it occurred before the long open vowel /ā/. If the *hamza* was fated to disappear from the old dialects of Arabic, this movement began in the Ḥijāz before Yemen, and before it was abandoned in all environments. Another instance of the innovative character of the Arabic spoken in the Ḥijāz is the articulation of the /ʿ/ phoneme. In the Ḥijāz, it was depharyngealized and shifted toward the *hamza*. In Yemen, the articulation of /ʿ/ was merely affected by nasalization. These examples suggest that innovations were born in the Ḥijāz in the north and traveled southward in the 7th century C.E.

There was a tendency toward generalizing a single relative pronoun in the → Northwest Arabian dialects. The pronoun *allaḏī* was used for masculine and feminine singular in the Ḥijāz and Huḏayl. In Ṭayyiʾ, a single relative pronoun *ḏū* was used for the two genders and all numbers. But in the south, Yemen used two relative pronouns *ḏī* and *ḏū*. In the south, then, there was more than one relative pronoun, while in all other dialects of the West Arabian group, there was only one single pronoun.

On the syntactic level, however, the dialects of the Ḥijāz and Yemen were on equal footing with respect to some innovations. There was a tendency to generalize one case ending for different sentence constituents under different effects. Both the dialects of Ḥijāz and Yemen generalized the use of one dual suffix for all cases, Yemen using *-āna* while the Ḥijāz used *-āni*. Another case of generalization is the use of *tanwīn* in Yemen, where words in pausal position retained the *tanwīn*.

The reduction of the declensional system may be used as an example of change and, probably, of a tendency toward an analytical type. This development was in opposition to the stability of this feature in the Qurʾānic variety. The case system was reduced before the period of the Arab conquests, if not abandoned altogether in some areas.

Corriente (1971:20–50) studied the functional load in the poetic language. He agrees with Fleisch (1947) and Blau (1965) that the case endings might have been a feature of Bedouin and urban vernaculars as well as the poetic language. Although case endings were fully operative in the poetic language and the *Qurʾān*, these texts show that there was a form of Arabic that did not realize the full case system, existing at the same time and in the same place as the full *ʾiʿrāb* form. This fact caused, Corriente assumes, the coexistence of two different evolutionary states in the development of Arabic (Corriente 1971:20–24). But, because the Arabic variety that realized case endings was a synthetic language which depended on the system for expressing syntactic relationships, Corriente conducted a survey of prose and verse texts from different periods. He discovered that the case endings, which characterized the poetic language and some vernaculars, had a very low functional load, since the meaning of the passages studied could be identified without the use of the case endings. Therefore, these cases became redundant in the dialects (Corriente 1971:25).

Blau (1988:261–262) opposes this reasoning, claiming that it is not possible to regard a redundant aspect of the language as a secondary set. Therefore, low functional yield and redundancy do not prove that case endings had been dropped in spoken vernacular. Blau (1988:262) states that nothing can be inferred concerning language use from the redundancy of case endings in the classical language, because this system in the Semitic languages is generally redundant. Corriente (1973:154–163) remarks that the case system is merely an indication of a changing variety.

Diem (1973:227–237), reviewing the body of Arabic proper names in the Aramaic Nabataean inscriptions, shows that the low functional load of case endings may have been the result of long processes of development. The significant aspect of the written forms of these proper nouns is that at the end of each, there were letters indicating vowels *u, a*, and *i*. These vowels resemble the case endings as preserved in Classical Arabic. He notes that 95 percent of the simple nouns ended in *w/u*, while the rest ended in *a or y/i* or had no ending. The *w/u* final endings seem to be the rule. Diem (1973:335) asserts that *w/u* was actually the nominative case ending in Arabic, added to the end of the noun to represent the sound that had long disappeared from pronunciation but lingered on in the conservative orthography. In the category of theophoric compound names, some inscriptions have them without ending, while in the majority they end with *y/i*, which is an echo of the once-pronounced *i* ending of the noun construct. As for nontheophoric names, like *ʿabd ʿamr*, which did not form a noun construct, the second noun was written with a final *w/u*. At other times, no vocalic ending was written. Since nontheophoric names developed later than the theophoric names, the simple noun part with its traditional *w/u* was simply annexed to the first part. Diem speculates that the forms with final *y/i* in compound nouns and those with final *w/u* in simple names belonged to a time when Nabataean Arabic had case endings. Thus, final *w* and *y* must have represented the nominative and the genitive respectively. But by the time of the writing of these inscriptions, Nabataean Arabic must have lost its inflections, and the proof is the difference in spelling between compound and simple nouns. The use of final *w/u* in nontheophoric names, while the *y/i* was used in the theophoric names, means that the old case system was no longer in use, except as a fossilized orthographic habit (Diem 1973:235).

Diem speculates that the occasional forms without vowel letters in final position may reflect the actual everyday use of the language during the time of the inscriptions, where case endings no longer existed. Regarding the spread of this change, Diem believes that if Nabataean Arabic lost its case system in the 1st century B.C.E., it is difficult to assume that the areas of central Arabia bordering the Nabataeans remained immune to this linguistic change until the 7th century C.E. It was the language of poetry that did not lose the case system. In addition, the relative importance of the Nabataeans until the 6th century C.E. may have enabled the change to spread into the Arabian heartland. However, the change did not creep into the poetic language because, apart from functional reasons, it was not a vernacular that was continuously checked by and subjected to fashion.

Diem's analysis identifies the locus of the beginning of change. The development toward a caseless language started in the peripheral area, where Arabic was only a vernacular. There are strong arguments that, due to the extensive contact between Arabs in the Arabian Peninsula and the Nabataean area in the period from the 1st to the 7th centuries C.E., the innovation probably moved to the northwestern and southwestern parts of Arabia, along the commerce lines and sedentary life in the peninsula. Because contacts between the Nabataean areas and the eastern and southeastern parts of the peninsula were minimal, these regions maintained usage of case endings for a longer time than Western dialects, hence the grammarians' admiration of the Eastern dialects. As evidence, there were more discrepancies between the dialects of the Ḥijāz and Yemen and Classical Arabic in the use of the case system than between the Eastern dialects and Classical Arabic. While Eastern nomadic dialects resembled Classical Arabic in their use of case endings, some other dialects retained only residues of the case endings.

The data from Huḏayl and ʾAzd show that they did not share in some of the innovations. ʾAzd preserved full case endings, and in Ṭayyiʾ, the *hamza* was replaced by *h*. This does not mean that where the case system was retained, it was not in a state of development. In ʾAzd, the cases were realized on the word in pause position, whereas, according to the rules of Classical Arabic, they must be deleted in final position.

Strong linguistic relationships between the Ḥijāz and Yemen were natural, due to the heavy influence of social and trade interests, which flourished after the signing of the treaty between Persia and the Byzantine Empire in 561 C.E. This treaty blocked trade routes in the north of the peninsula and forced caravans to use the West Arabian route between Mecca and Yemen (Shahid 1988:181–192). Trade moved between the urban centers in Yemen and their equivalents in the Ḥijāz. Along this route, linguistic innovations may have spread from the Ḥijāz to the southern parts of Yemen. But if this is true, what prevented the Bedouin tribes of Ṭayyiʾ, ʾAzd, and the Tihāma from sharing in linguistic innovations common between Yemen and the Ḥijāz, despite the tribes' positions along the route?

According to ʾAnīs (1952), it was natural for the dialects of the Tihāma and those of Huḏayl and Ṭayyiʾ to exhibit differences from those of the Ḥijāz and Yemen because the former were Bedouin tribes who shared some linguistic features with other, eastern Bedouin tribes. From the data in the Arab grammarians' books, ʾAnīs deduces that linguistic features were sometimes assigned to one sedentary tribe in the western part of the peninsula and at the same time to a Bedouin tribe in the eastern part. In other cases, opposite features were assigned to one and the same tribe. ʾAnīs explains these apparent contradictions by assuming that tribes could have both sedentary and Bedouin clans. Features typical of Bedouin speech are in the realm of phonology, for instance vowel harmony and *ʾimāla*. Therefore, when parts of Huḏayl were described by grammarians as having one of these features, they must have meant the Bedouin clans of that tribe, and one may further assume that they were the clans adjacent to the Najd. The same also applies to other Bedouin clans in the Ḥijāz and Yemen.

Yet, we do not know which parts of a tribe were Bedouin and which were not. More recently, al-Gindī (1983:36–38) accepts ʾAnīs's assumption but rejects the generalization that the majority of the inhabitants of the Ḥijāz were sedentary and the majority of the inhabitants of the eastern part of the peninsula, such as Tamīm, were Bedouin. He argues that the boundaries of the Ḥijāz and Tamīm, east and west, were not rigidly defined, and Bedouin clans were free to move from one geographical area to another without changing their identity and linguistic behavior. Moreover, Tamīm, ʾAsad, and Rabīʿa in the eastern part were large tribal alliances that included several tribes and may therefore have hosted different linguistic features.

It was natural for the Arab sedentary communities to gain and share in innovations more than Bedouin tribes did, since the former must have received the innovations through a constant line of communication with the source of innovation. If innovations that distinguished West Arabian dialects from Classical Arabic moved from the Ḥijāz to Yemen, they must have originated somewhere in the northwestern peripheral area. A case in point is the reduction of the declension. This may have started in the Nabataean Kingdom in the 1st century C.E. and spread later to the rest of west and southwest Arabia along the trade routes. Trade caravans between the Levant and Mecca and between Mecca and Yemen were responsible for transporting innovation. Such a trade line is unlikely to transport innovation to off-line tribes and clans that were not stations along the route.

All this leads one to assume that the sedentarized tribes along the western trade route were in the process of developing a special variety of Arabic, as opposed to the rest of the largely Bedouin dialects of Arabic. Although the data in the sources are random and few, it can be deduced that on the phonological level, these sedentary dialects were characterized by a change in the articulation of the pharyngeal /ʿ/, the elision of *hamza* in all or most environments, lack of vowel harmony, and absence of *ʾimāla*. On the morphological level, sedentary dialects were characterized by the reduction of linguistic categories. Finally, on the syntactic level, they were defined by the overgeneralized use of case endings.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Fīrūzābādī, *Muḥīṭ* = Majd ad-Dīn Muḥammad ibn Yaʿqūb al-Fīrūzābādī, *al-Qāmūs al-muḥīṭ*. 4 vols. Cairo, 1357 A.H.

Ibn Fāris, *Ṣāḥibī* = ʾAbū l-Ḥusayn ʾAḥmad Ibn Fāris, *aṣ-Ṣāḥibī fī fiqh al-luġa wa-sunan al-ʿArab fī kalāmihā*. Ed. Moustafa el-Chouémi. Cairo: Quṣūr aṯ-Ṯaqāfa, 1964.

Ibn Hišām, *Muġnī* = Jamāl ad-Dīn ʿAbū Muḥammad ʿAbdallāh ibn Yūsuf Ibn Hišām, *Muġnī l-labīb ʿan kutub al-ʾaʿārīb*. Ed. Māzin al-Mubārak and Muḥammad ʿAlī Ḥamdallāh. Damascus: Dār al-Fikr, 1969.

Ibn ʾIsḥāq, *Sīra* = ʾAbū Bakr Muḥammad Ibn ʾIsḥāq, *as-Sīra an-nabawiyya* (in the recension of ʾAbū Muḥammad ʿAbd al-Malik Ibn Hišām). Ed. Muṣṭafā as-Saqā, ʾIbrāhīm al-ʾIbyārī, and ʿAbd al-Ḥāfiẓ Šalabī. 4 vols. Cairo: Muṣṭafā al-Bābī al-Ḥalabī, 1936.

Jāḥiẓ, *Bayān* = ʾAbū ʿUṯmān ʿAmr ibn Baḥr Jāḥiẓ, *al-Bayān wa-t-tabyīn*. Ed. Ḥasan as-Sandūbī. 3 vols. Beirut: Dār al-Fikr, n.d.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. Ed. ʿAbd as-Salām Muḥammad Hārūn. 5 vols. Repr., Cairo: Maktabat al-Xānjī, 1982.

Zajjājī, *Šarḥ* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān ibn ʾIsḥāq, *Šarḥ Durrat an-nawāṣī*.

Zubaydī, *Ṭabaqāt* = ʾAbū Bakr Muḥammad ibn al-Ḥasan az-Zubaydī, *Ṭabaqāt an-naḥwiyyīna wa-l-luġawiyyīn*. Ed. Muḥammad ʾAbū l-Faḍl ʾIbrāhīm. Cairo: Dār al-Maʿārif, 1973.

SECONDARY SOURCES

ʾAnīs, ʾIbrāhīm. 1952. *Fī l-lahajāt al-ʿarabiyya*. Cairo: al-Maktaba al-ʾAnglū al-Miṣriyya.

Bernards, Monique. 1997. *Changing traditions: al-Mubarrad's refutation of Sībawayh and the subsequent reception of the* Kitāb. Leiden: E.J. Brill.

Blachère, Régis. 1950. "Les savants iraquiens et leurs informateurs bédouins aux IIe–IVe siècles de l'Hégire". *Mélanges offerts à William Marçais*, 17–48. Paris: G.-F. Maisonneuve.

Blau, Joshua. 1965. *The emergence and linguistic background of Judaeo-Arabic: A study of the origins of Middle Arabic*. Oxford: Oxford University Press.

——. 1988. "On the problem of the synthetic character of Classical Arabic as against Judaeo-Arabic (Middle Arabic)". Joshua Blau, *Studies in Middle Arabic and its Judaeo-Arabic variety*, 260–269. Jerusalem: Magnes Press, 1988.

Chejne, Anwar G. 1969. *The Arabic language: Its role in history*. Minneapolis: University of Minnesota Press.

Corriente, Federico. 1971. "On the functional yield of some synthetic devices in Arabic and Semitic morphology". *Jewish Quarterly Review* 62.20–50.

——. 1973. "Again on the functional yield of some synthetic devices in Arabic and Semitic morphology". *Jewish Quarterly Review* 64.154–163.

Diem, Werner. 1973. "Die nabatäischen Inschriften und die Frage der Kasusflexion im Altarabischen". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 123.227–237.

Fleisch, Henri. 1947. *Introduction à l'étude des langues sémitiques: Eléments de bibliographie*. Paris: Adrien-Maisonneuve.

——. 1964. "Arabe classique et arabe dialectal". *Travaux et Jours* 12.23–62.

Fück, Johann. 1950. *ʿArabiya: Untersuchungen zur arabischen Sprach- und Stilgeschichte*. Berlin: Akademie-Verlag.

Geyer, R. 1909. Review of Vollers (1906). *Göttingische Gelehrte Anzeigen* 171.

Gindī, A. al-. 1983. *al-Lahajāt al-ʿarabiyya fī t-turāṯ*. Beirut: ad-Dār al-ʿArabiyya li-l-Kitāb.

Holes, Clive. 1995. *Modern Arabic: Structures, functions and varieties*. London: Longman.

Kahle, Paul. 1948. "The *Qurʾān* and the *ʿarabiyyah*". *Ignaze Goldziher Memorial*, ed. S. Löwinger and J. Somogyi, I, 163–182. Budapest.

Naṣṣār, H. 1988. *al-Muʿjam al-ʿarabī: Našʾatuhu wa-taṭawwuruhu*. Cairo: Maktaba Miṣr.

Nöldeke, Theodor. 1904. *Beiträge zur semitischen Sprachwissenschaft*. Strasbourg: K. Trübner.

——. 1910. *Neue Beiträge zur semitischen Sprachwissenschaft*. Strasbourg: K. Trübner.

Owens, Jonathan. 1998. "Case and Proto-Arabic". *Bulletin of the School of Oriental and African Studies* 61:1.51–73.

Rabin, Chaim. 1951. *Ancient West Arabian*. London: Taylor's Foreign Press.

——. 1955. "The beginnings of Classical Arabic". *Studia Islamica* 4.19–37.

Rippin, Andrew. 1981. "Ibn ʿAbbās's *al-Lughāt fī l-Qurʾān*". *Bulletin of the School of Oriental and African Studies* 44.15–25.

Rosenthal, Franz. 1953. Review of Fück (1950). *Orientalia* 22.307–311.

Shahid, Irfan. 1988. *Byzantium and the Semitic Orient before the rise of Islam*. London: Variorum Reprints.

Sharkawi, Muhammad al-. 2005. *Arabicization: A case for second language acquisition*. Ph.D. diss., University of Nijmegen.

Spitaler, Anton. 1953. Review of Fück (1950). *Bibliotheca Orientalis* 10.144–150.

Versteegh, Kees. 1984. *Pidginization and creolization: The case of Arabic*. Amsterdam: J. Benjamins.

Vollers, Karl. 1906. *Volkssprache und Schriftsprache im alten Arabien*. Strasbourg: K. Trübner.

Wehr, Hans. 1952. Review of Fück (1950). *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 102.179–186.

Zwettler, Michael. 1978. *The oral tradition of Classical Arabic poetry: Its character and implications*. Princeton, N.J.: Princeton University Press.

Muhammad al-Sharkawi
(American University in Cairo)

Prepositions

Prepositions may be defined as function words indicating the relation of a noun or pronoun to other words in the clause. Thus, the study of prepositions includes aspects of morphology and syntax as well as the lexicon itself. This entry classifies the prepositions of Arabic, briefly describes their most important forms, and outlines their usage in Classical, Modern Standard, and modern dialectal Arabic.

All prepositions in Arabic, regardless of whether they are classified as 'primary prepositions' or 'secondary prepositions' (see discussion below), share several syntactic features. First, they always precede the noun they govern – in other words, in Arabic there are no postpositions. Second, all Arabic prepositions require the genitive case; hence, in Arabic they are called *ḥurūf al-jarr* 'particles of the genitive' (other terms used by the native grammarians for prepositions are *ḥurūf al-xafḍ, ḥurūf al-ʾiḍāfa*, and *al-jawārr*). This salient feature of Arabic prepositions applies to the Classical and Modern Standard forms of the language, but not to the dialects, because the latter have given up case marking. Third, if the dependent item is a pronoun rather than a noun, it is usually suffixed to the preposition. This rule requires the repetition of the preposition in coordinative phrases, since only one single suffix can be attached to it, e.g. *minnī wa-minka* 'from me and you'. There are very few prepositions that do not normally take suffixes, except in poetic language. Among them are *ḥattā* 'until', *muḏ ~ munḏu* 'since', *ka-* 'like', and the compound preposition *bi-lā* 'without'. The preposition *ka-*, however, is sometimes prefixed to an independent pronoun, e.g. *ka-huwa* 'like him'.

There is still no commonly accepted definition of which Arabic words are prepositions. This is mainly the result of the fact that many of the words which Western concepts of grammar consider prepositions are not regarded as such by the Classical Arabic grammarians. For the latter, only a very few words (the number ranges between eight and fifteen, depending upon the grammarian) are real prepositions (i.e. *ḥurūf al-jarr*) and thus belong to the word class of particles (→ *ḥarf*). In grammars of Arabic, these words are often called 'primary prepositions' or 'true prepositions'. Those words that are considered to be primary prepositions by Arab and Western grammarians alike are the following (a hyphen indicates that the word in question is attached to a following noun in Arabic script): *ʿalā* 'on'; *ʿan* 'away from'; *bi-* 'with, in'; *fī* 'in'; *ḥattā* 'until'; *ʾilā* 'to, toward'; *li-* 'for'; *ka-* 'like'; *min* 'from, of'; and *muḏ ~ munḏu* 'since' (< **min ḏū*; Arab grammarians give a sophisticated rule concerning *muḏ ~ munḏu* governing the nominative instead of the genitive; see Wright 1974:II, 173–174). In most Western grammars of Arabic, *maʿa* 'together with' and *ladā ~ ladun* (with many variants) 'at' are also regarded as primary prepositions, whereas Arab grammarians include the different oath particles (*wa-*, *bi-*, *ta-* 'by'), as well as other words such as *rubba* 'how much!' and *xalā*, *ʿadā*, *ḥāšā* 'except' among their list of prepositions.

The primary prepositions form a closed-list class and doubtless go back to a very early stage of the language. This is indicated by the fact that most of them are found not only in Arabic but also in other Semitic languages (especially *ʿalā*, *bi-*, *ka-*, *li-*, and *min*). Their three common characteristics are: they are invariable; they are normally used exclusively as prepositions (except *ḥattā* and *mundu*, which also serve as conjunctions); and they are – with the exception of *ʿalā* – at least on a synchronic level not related to any triliteral root and thus lie outside the derivational system.

The second class of prepositions are formally nouns in the construct state of the accusative (i.e. ending in *-a*). There have been several attempts in Western studies to coin terms other than 'preposition' to differentiate these forms from the 'true' prepositions listed above – for example 'secondary prepositions' (used, e.g., by Fischer 2002), 'locutions prépositives' (Fleisch 1979), and 'prepositionals' (Badawi a.o. 2004).

The words in this second class primarily express local and temporal relations, and their corresponding forms in other, especially Indo-European, languages are usually called 'prepositions'. Some examples are *fawqa* 'above', *ḥawla* 'around', *baʿda* 'after', and *qabla* 'before'. Many of these secondary prepositions are obviously derived from nouns, e.g. *bayna* 'between' < *bayn* 'interval', *xalfa* 'behind' < *xalf* 'backside', and *naḥwa* 'toward' < *naḥw* 'direction'. Nevertheless, it should be emphasized that, although most of these words are clearly related to triliteral roots, there is not always evidence that they have ever been used as nouns. Consequently, the nominal origin of several forms is likely but remains hypothetical. For instance, the two forms *qubālata* and *tujāha* 'opposite' are attested as prepositions only, and there is no evidence in any Semitic language for a noun **taḥt* 'bottom' that could be the underlying form of *taḥta* 'below.'

The nominal character of these secondary forms is evident when they are combined with primary prepositions, since in this case they are inflected like nouns, e.g. *taḥta* 'below', but *min taḥti* 'from beneath'; and *dūna* 'below, in front of' (for a detailed discussion of this preposition, see Ambros 2001), but *bi-dūni* 'without'. Furthermore, several of them can form a diminutive, specifically *baʿda* > *buʿayda* 'shortly after', *qabla* > *qubayla* 'shortly before', and *fawqa* > *fuwayqa* 'a little above'. Some of these prepositions can also be used as adverbs (e.g. *ʾamāman* 'forward'), among them a few that exhibit the peculiar invariable ending *-u*, e.g. *baʿdu* 'later'. Secondary prepositions represent an open word class, because both in Modern Standard Arabic and in the dialects new forms keep appearing.

A very small number of combinations consisting of a preposition and another word may be regarded as compound prepositions, since they have developed special meanings: *bayna yaday* 'in front of' (lit. 'between the two hands'); *bi-lā*, *min/bi-dūni* 'without'. These forms must be distinguished from the numerous instances where prepositional phrases or combinations of adverbs and prepositions are used in the same sense as the secondary preposition alone. The parallel usage of compound and single forms is already attested for the Arabic of the *Qurʾān*, but it has become very common in Modern Standard Arabic. Examples are: *mitla* ~ *ka-mitli* 'like'; *ṣadada* ~ *ʿalā/bi-ṣadadi* 'opposite'; *ḍimna* ~ *fī ḍimni* 'within, among'; *badala* ~ *badalan min* 'instead of'; and *raġma* ~ *raġman ʿan* 'in spite of'. Sometimes, however, the two forms are used differently, such as in Modern Standard Arabic *xilāla* (and *fī xilāli*) 'during' vs. *min xilāli* 'through, by way of'.

The main function of prepositions is to indicate local and temporal relations: position, direction, and concomitance. The primary prepositions possess a wider semantic range than the secondary forms because extensions of their basic meanings are frequently found and they are also often used in metaphorical senses. In Classical Arabic, the main forms indicating local position are *fī* and *bi-* 'in', *ʿalā* 'on, at', and *ladā/ladun* 'at'. The numerous secondary locational prepositions include *ʿinda* 'at, near', *fawqa* 'above', *taḥta* 'below', *ʾamāma* (and post-Classical *quddāma*) 'in front of', and *warāʾa* and *xalfa* 'behind'. Direction toward is expressed mainly by *ʾilā* 'to', *li-* 'to, for', *ḥattā* 'up to', and *naḥwa* 'toward', and direction from is indicated by *min* and *ʿan* 'from' (for details and the diachronic development of the latter two prepositions, see Ambros 1982).

Opposition is expressed by a number of prepositions, among them *ḥiḏāʾa*, *qubālata*, and *tilqāʾa* 'opposite'; interposition is indicated by *bayna* and *xilāla* 'between'. The most important temporal prepositions are *fī* and *bi-* 'at, in', *baʿda* 'after', *qabla* 'before', as well as *ʾilā* and *ḥattā* 'until'.

Association is particularly expressed by *bi-*, *maʿa* 'with', and, in a special construction, by *wa-* plus a following accusative (called *wāw al-maʿiyya* by the Arab grammarians), e.g. *al-ʾamīru wa-l-jayša* 'the commander with the army'. When combined with intransitive verbs of locomotion (esp. *ʾatā, jāʾa, ḏahaba, qāma*), *bi-* in the sense of 'with' imparts a transitive sense, e.g. *jāʾa bi-kitābin* 'he brought a book' (lit. 'he came with'; hence the verb *jāb* 'to bring' in many modern dialects) and *ḏahaba llāhu bi-nūrihim* 'God took away [lit. 'went away with'] their light' (Q. 2/17).

The most significant extended usages of prepositions include the indication of an instrument or a price by *bi-*, e.g. *bi-l-qalami* 'by means of the pen', *bi-dirhamayni* 'for two dirhams'. Contrast is expressed by *ʿalā* (e.g. *qatala ʾasadan ʿalā ṣiġari sinnihi* 'in spite of his youth he killed a lion') and *maʿa* (e.g. *maʿa šakkinā* 'in spite of our doubt'). Causal and final relations can be expressed with *ʿalā* and *li-*, e.g. *ʾaḥmaduhu ʿalā ʾālāʾihi* 'I praise him for his good deeds', *ʾa-jiʾtum li-qitālinā* 'have you come to fight us?' The preposition of the comparative is *min*, e.g. *laylatu l-qadri xayrun min ʾalfi šahrin* 'the Night of Power is better than a thousand months' (Q. 97/3). Both *min* (in modern Arabic mostly *min qibali*) and *li-* indicate the agent in a passive clause, e.g. *wa-mā ʾūtiya n-nabiyyūna min rabbihim* 'and what has been given to the prophets by their Lord' (Q. 2/136), *ḥattā mā ʾurāʿu lahu* 'until I am not frightened by him any longer'.

Four prepositions are used to indicate different aspects of → possession: *li-, ʿinda, ladā*, and *maʿa*. The first, *li-*, is used particularly with inalienable objects, e.g. *lahā bintun* 'she has a daughter'. Both *ʿinda* and *ladā* express mainly 'having at one's place' and 'owning'. They are often interchangeable, although *ladā*, at least in Classical Arabic, is applied more frequently to abstract than to concrete nouns. The fourth, *maʿa*, emphasizes possession at one's disposal, e.g. *kāna maʿa banī maʿnin kitābun min an-nabiyyi* 'the Banū Maʿn had a letter of the Prophet with them' (for more details concerning possessive prepositions, see Shboul 1983; Ambros 1984).

Primary prepositions, but seldom secondary prepositions, play a major role in connecting verbs with their indirect objects. Their main function is to specify the semantic range of the verb in question. Examples are *ʾalqā bi-* 'to throw something', *iḥtāja ʾilā* 'to be in need of something', and *tamakkana min* 'to master something'. Some verbs can even take opposite meanings when combined with two different prepositions, e.g. *daʿā li-* 'to bless someone' (lit. 'to pray for') vs. *daʿā ʿalā* 'to curse someone' (lit. 'to pray against'); *raġiba fī* 'to like something' vs. *raġiba ʿan* 'to dislike something'. Under certain circumstances, Classical Arabic prepositions are used instead of the genitive and accusative in analytical constructions. A genitive can be replaced by *li-* and – much more rarely – by *min* (see Reckendorf 1921:259–267), e.g. *man kāna ʿaduwwan li-jibrīla* 'whoever is an enemy to Gabriel' (Q. 2/97), *bi-mawtin min ḥalīliki* 'because of the death of your husband'. Rather than a direct object, a number of verbs have *bi-* (the so-called *bāʾ at-taʿdiya*; see Reckendorf 1921:236–237), e.g. *wa-baʿaṯa bi-l-ʾasrā* 'he sent the prisoners', while participles and verbal nouns usually have *li-* in this function, e.g. *kānat šāniʾatan lahu* 'she hated him' (→ transitivity; for *li-* as an object marker in the dialects, see below).

The differences between Classical and Modern Standard Arabic are mainly due to two developments: first, the ongoing creation of secondary prepositions (Modern Standard Arabic has about fifty more than Classical Arabic), and second, the semantic extension of old prepositions – caused to a large extent by the influence of European languages. Moreover, most forms have been standardized, i.e., morphological variants are no longer used or are used only to avoid synonymy (e.g., *ḥawla* is predominantly used in the sense of 'around', but *ḥawālā/ḥawālay* in the sense of 'approximately'). New prepositions expressing temporal and local relations include *ʾaṯnāʾa* 'during'; *ʾiṯra, ʿaqiba*, and *fawra* 'immediately after'; *ʿabra* 'by way of'; *ʾizāʾa, ṣawba* 'toward'; *qurba* 'near'; *ḍidda* 'against'; *dāxila* 'inside'; *xārija* 'outside'. Other new prepositions are *ḥasaba, ṭibqa*, and *wafqa* 'in accordance with'; *qaṣda* 'with intent to'; *qayda* 'subject to'; *jarāʾa* 'because of'; *ḥiyāla* ‘concerning’; *muqābila* and *naḏ̣īra* ‘in return for’.

Extended usage of prepositions as the result of translations from English and French is found with *ka-* ‘like’, which today occurs also in the sense of ‘as’, e.g. *yaštaġilu ka-ṭabībin* ‘he works as a doctor’. Another innovation is the tendency to combine *maʿa* with reciprocal verbs (see Wehr 1934:18), e.g. *tabārā maʿa* ‘to compete with’, *tanāfā maʿa* ‘to be incompatible with’, *ixtalafa maʿa* ‘to disagree with’. The usage of *taḥta* in expressions such as *taḥta riʿāyati* ‘under the patronage of’ or *taḥta ʿunwāni* ‘under the title of’ probably goes back to Western influences as well. The same must be true for *ḥawla* (actually ‘around’), which is used to indicate a topic, exactly like French *autour de* and English *about*.

The increasing use of *ladā* instead of *ʿinda*, however, seems to be the result of a negative dialectal interference: *ladā* has become the preferred preposition to indicate possession in Modern Standard Arabic precisely because it is *not* found in the dialects.

The prepositions used in the modern Arabic dialects display some striking differences from those of Classical Arabic. The following sketch is restricted to a description of the three main tendencies regarding prepositions found in the majority of the dialects: the emergence of new forms not attested in the old language; the simplification of the system, especially the decay of synonymous forms; and semantic change (see Procházka 1995). Three frequent prepositions of Classical Arabic are not reflected in any dialect, namely *ʾamāma* ‘in front of’, *ladā* ‘at’, and *munḏu* ‘since’.

Although some new prepositions are used over large regions, most of the new forms are locally restricted and thus often typical for the dialect of a given region. Nearly all Eastern dialects know the pair *barra/barrāt* ‘outside’ and *juwwa/juwwāt* ‘inside’ (the latter is used instead of *taḥt* for ‘below’ in the Iraqi dialects). The same is true of *jamb* (< *janb* ‘side’) ‘next to’. A reflex of Classical Arabic *wa-ʾiyyā-* (followed by a pronominal suffix) ‘with’ is the form *wiyya*, widely used in the sense of ‘together with’, particularly in the Iraqi dialects (which have given up the preposition *maʿa*).

Typical for most dialects of the Maghreb and Upper Egypt is the preposition *kīf/kēf* (shortened *kī-/kē-*) ‘like, as’; only in Morocco is the compound form *bḥāl* (< *bi-ḥāl* ‘in the state of’) used. Lower Egyptian and some Palestinian dialects use *zayy* (< *ziyy* ‘fashion of dress’) in this meaning. A large number of words express the concept of proximity, i.e. ‘at, near’. Among them are *jīht* < *jiha* ‘side’ in the dialects west of Libya; *ḥadd* < *ḥadd* ‘edge’ in Lebanon and Palestine; *rīḥ* < dialectal *rīḥ* ‘side’ in many rural dialects of Egypt; *šigg* < *šiqq* ‘side’ in Yemen and in Bedouin dialects of the Maghreb; *yamm* (of uncertain origin) in Iraq, Arabia, and parts of Syria.

In a few regions, designations for parts of the body have become prepositions. For example, reflexes of *qafā* ‘nape of the neck’ are used to express ‘behind’ in South Arabian, Anatolian, and Mauritanian dialects, e.g. Sanaa *gafa al-bayt* ‘behind the house’. A preposition meaning ‘after’ has emerged from the word *ʿaqib* ‘heel’ in many Mesopotamian, Arabian, and Bedouin dialects, e.g. Abu Dhabi *ʿugub bāčir* ‘after tomorrow’. In Upper Egypt and the Arab minorities of Central Asia, *xašm* ‘nose’ is used for ‘in front of’, e.g. Egypt *xašm id-dukkān* ‘in front of the shop’.

In several instances where Classical Arabic has two synonymous prepositions, either one form has not survived in the dialects, or their functions have been separated. This is illustrated by the two ablative prepositions *min* and *ʿan*: in virtually all North African and several Eastern dialects, reflexes of *ʿan* are unknown, but in those dialects which have both *min* and *ʿan,* the first is used in an ablative sense, while the latter usually marks a topic. The two locative prepositions *bi-* and *fī* have undergone three different types of development. First, the Iraqi dialects and many Arabian Bedouin dialects have reflexes of *bi-* only; second, in most Syro-Palestinian dialects, *fī* is used only with suffixes and *bi-* only with nouns; and finally, the North African dialects use *bi-* in an instrumental function and *fī* only to indicate a place.

Semantic change can be observed with *ʿalā*, which in many Syro-Palestinian and Mesopotamian dialects is used to indicate local direction, e.g. Damascus *rəḥna ʿala falasṭīn* ‘we went to Palestine’. Particularly in some dialects on the periphery of the Arab world, reflexes of the two local prepositions *quddāma* and *warāʾa* are also used in a temporal sense, e.g. Cilicia (Turkey) *quddām il-ġada* ‘before noon’, and Mauritania *wṛā l-maġrəb* ‘after sunset’ (temporal usage of *warāʾa* is, however, also attested for Qurʾānic Arabic).

One of the rare instances of substrate influence is the usage of *li-* as a direct object marker, which is always accompanied by a corresponding personal suffix attached to the verb. This typically Aramaic construction is found in many dialects of the Fertile Crescent, e.g. Lebanon *šift-o la-xayyi* 'I saw my brother', and Baghdad *bāʿ-a li-l-bēt* 'he sold the house' (→ transitivity).

Bibliographical references

Detailed information on prepositions in Classical Arabic can be found in Fischer (2002:134–148), Fleisch (1979:484–490), and Wright (1974:I, 278–282; II, 129–193). Qurʾānic usage of the primary prepositions is described exhaustively by Ghazoui (1994–1995). The best syntactic study for Classical Arabic remains Reckendorf (1921). The syntax of prepositions in Modern Arabic is discussed in detail by Cantarino (1974–1975, based on data from before the 1950s), and more recently by El-Ayoubi a.o. (2003) and Badawi a.o. (2004). For prepositions in the modern dialects, see the monograph by Procházka (1993).

Ambros, Arne A. 1982. "Beobachtungen zu den arabischen Präpositionen *min* und *ʿan*". *Wiener Zeitschrift für die Kunde des Morgenlandes* 74.95–102.

——. 1984. "Einige Bemerkungen zu 'Having in Arabic'". *Zeitschrift für Arabische Linguistik* 13.91–93.

——. 2001. "Die Funktionen von (*min*) *dūn* im Koran". *Zeitschrift für Arabische Linguistik* 39.7–19.

Badawi, Elsaid, Michael G. Carter, and Adrian Gully. 2004. *Modern Written Arabic: A comprehensive grammar*. London and New York: Routledge.

Cantarino, Vicente. 1974–1975. *Syntax of modern Arabic prose*. 3 vols. Bloomington and London: Indiana University Press.

El-Ayoubi, Hashem, Wolfdietrich Fischer, and Michael Langer. 2003. *Syntax der arabischen Schriftsprache der Gegenwart*, I/2. Wiesbaden: L. Reichert.

Fischer, Wolfdietrich. 2002. *Grammatik des Klassischen Arabisch*. 3rd ed. Wiesbaden: O. Harrassowitz.

Fleisch, Henri. 1979. *Traité de philologie arabe*. II. *Pronoms, morphologie verbale, particules*. Beirut: Imprimerie Catholique.

Ghazoui, Meriem. 1994–1995. *Les prépositions dans le Coran: Approche syntaxico-sémantique*. Paris: Septentrion.

Procházka, Stephan. 1993. *Die Präpositionen in den neuarabischen Dialekten*. Vienna: VWGÖ.

——. 1995. "Prepositions in modern Arabic dialects: Some striking changes compared to Classical Arabic." *Actas XVI Congreso de la Union Européenne d'Arabisants et d'Islamisants*, ed. Concepción Vázquez de Benito and Miguel Ángel Manzano Rodríguez, 417–424. Salamanca: CSCI/AECI.

Reckendorf, Hermann. 1921. *Arabische Syntax*. Heidelberg: C. Winter.

Shboul, Ahmad. 1983. "'Having' in Arabic". *Zeitschrift für Arabische Linguistik* 11.24–47.

Wehr, Hans. 1934. *Die Besonderheiten des heutigen Hocharabischen mit Berücksichtigung der Einwirkung der europäischen Sprachen*. Berlin: Reichsdruckerei.

Wright, William. 1974. *A grammar of the Arabic language*. 3rd ed. Repr. Beirut: Librairie du Liban.

STEPHAN PROCHÁZKA (University of Vienna)

Prepositional Clause → Mafʿūl fīhi

Presentatives

The term 'presentative' is used here to refer to a variety of constructions containing particles that have the function of drawing the attention of the hearer/reader. The particles draw attention either to a referent or to a proposition expressed by a clause. In Classical Arabic, this function is performed by demonstrative particles and *ʾiḏā (bi-)*.

When a presentative particle is used to draw attention to a referent, it forms a complete clausal unit, e.g. *hāḏā zaydun, ʾiḏā zaydun* 'here is Zayd'.

The presentative function of the demonstrative particles should be distinguished from their more usual function of identifying a referent ('this one, this man'; → deixis). The two functions are clearly related; in both cases, the particle points the attention of the hearer to a referent. The subtle difference lies in the fact that when the particle is used in the identifying function the referent is assumed to be perceptible by the hearer in the speech situation, whereas when it has the presentative function, the particle draws attention to the emergence of a referent into the speech situation and into the perception of the hearer. The presentative use of the demonstrative tends to be restricted to the near-deixis form (*hāḏā*) and is rarely attested with the far-deixis form (*ḏālika*). There is gender and number agreement between the demonstrative and the referent presented, e.g. *hāḏā zayd* 'here is Zayd', *hāḏihi zaynab* 'here is Zaynab'.

The noun after the presentative demonstrative may be qualified by a further description, which is typically a circumstantial phrase (→ *ḥāl*), e.g. *hāḏā zaydun munṭaliqan* 'here is Zayd, departing'. By a slight shift in structure, the circumstantial phrase may be construed as the predicate of the preceding noun, e.g. *hāḏā zaydun munṭaliqun* 'look, Zayd is departing'. In this case, the demonstrative presents the proposition expressed by this subject-predicate nexus, rather than the referent of the noun (for Sībawayhi's analysis of these two structures, see Mosel 1975:222, 251). It is likely that such proposition-centered presentative structures developed historically from the referent-centered construction with the circumstantial expansion. The close relationship between the two is demonstrated by the fact that the accusative and nominative case inflections of the descriptive element sometimes exist as variant readings of a text in manuscripts (*munṭaliqan* ~ *munṭaliqun*). When the descriptive element is not nominal, moreover, there is no inflectional distinction between the two constructions, e.g. *hāḏā zaydun yanṭaliqu* 'here is Zayd, departing' ~ 'look, Zayd is departing'. It is likely that the reanalysis took place in these structurally ambiguous environments.

The referent-centered presentative construction has an essentially uniform function of drawing attention to the emergence of a concrete entity into the speech situation. Since propositions are components of discourse, the basic overarching function of the proposition-centered presentative is to give prominence to the proposition within the discourse. The speaker/writer uses this prominence to perform a diverse range of functions. These include not only the function of drawing attention to an event unfolding in the present speech situation but also that of drawing attention to propositions that do not relate to the perceptible speech situation, yet which are deemed to be of particular importance in the discourse. The latter include propositions that supply an explanation or justification for some action or those that express an important point in an argument. In some cases, the construction is used in a clause that opens a narrative or narrative section.

In the proposition-centered constructions, the demonstrative sometimes agrees in gender and number with the subject referent, e.g. *hāḏihi zaynab munṭaliqatun* 'see, Zaynab is departing'. This may be regarded as a vestige of the original referent-centered construction. There is a tendency, however, for the demonstrative to remain uninflected in the default masculine singular form in proposition-centered constructions. This is attested mainly in Middle Arabic texts, e.g. *wa-hāḏā ʾāsiya wa-kull ad-dunyā tasjud lahā* 'and behold, Asia and all the world worship her' (Blau 1966–1967:464). The presentative demonstrative may also introduce a clause that does not have an initial subject component, in which case the demonstrative is in principle uninflected in the masculine singular form, e.g. *hāḏā ʿajiltum ʿani d-dahni* 'look, you have come so fast that I did not yet anoint myself' (Fischer 1959:161; Bloch 1986:60).

When the presented item is an independent pronoun, the construction is *hā* + pronoun + *ḏā*, in which the two components of the demonstrative particle *hāḏā* are split, e.g. *hā huwa ḏā* 'here he is'. The second component agrees in gender and number with the pronoun, e.g. *hā hiya ḏī* 'here she is', *hā naḥnu ʾulāʾi* 'here we are'. The first or third component may be omitted, e.g. *huwa ḏā* 'here he is', *hā ʾana* 'here I am'. In Middle Arabic texts, the invariable form *(hā) huwa ḏā* is sometimes used before nouns as a presentative particle, e.g. *huwa ḏā mraʾatuka* 'here is your wife'.

The Arabic dialects use numerous presentative particles that are formed from demonstrative elements, e.g. *hā, ḏā, hay, ʾā*. There is a particular tendency for these to be combined with pronominal suffixes, full paradigms being formed by attaching suffixes to the demonstrative base, e.g. *hā* + suffixes in various North African dialects: *hāni* (1st pers. sg.), *hāk* (2nd pers. masc. sg.), *hāki* (2nd pers. fem. sg.), *hāhu* (3rd pers. masc. sg.), *hāhi* (3rd pers. fem. sg.), *hāna* (1st pers. pl.), *hākum* (2nd pers. pl.), *hāhum* (3rd pers. pl.) (Ulād Brāhīm, Fischer 1959:165); *hay* + suffixes in Palestinian dialects: *heyni* (1st pers. sg.), *heyyak* (2nd pers. masc. sg.), *heyyo* (3rd pers. masc. sg.), *heyha* (3rd pers. fem. sg.), *heyna* (1st pers. pl.), *heykum* (2nd pers. pl.), *heyhun* (3rd pers. pl.) (Jerusalem, Fischer 1959:179). The form of the 1st person singular suffix *-ni* in these constructions demonstrates that these are object suffixes. In some cases, the paradigm is restricted to 3rd person suffixes, as in Egyptian dialects, e.g. *ʾaho* (3rd pers. masc. sg.), *ʾahe* (3rd pers. fem. sg.), *ʾahum* (3rd pers. pl.) (Cairo). In some cases, the 3rd person masculine singular form of these presentative paradigms is itself used as an invariable base of a secondary paradigm. The Palestinian dialects, for example, have the secondary paradigm *heyyūni, heyyūk, heyyūh, heyyūha, heyyūna*, etc., which is formed from the 3rd person masculine singular base *heyyo*. In Egyptian dialects, furthermore, the 3rd person masculine singular form *ʾaho* may be combined with independent 1st and 2nd person pronouns, e.g. *aho-na*, *aho-nta* (Mitchell 1956:56; Woidich 2006:48–49). When a full noun is used in a presentative construction, this is often combined with a pronominal presentative form, e.g. Egyptian Arabic *aho rrasmi* 'here is the plan', *ahum ilkutub* 'here are the books'.

The item expressing the referent that is presented by the Classical Arabic particle *ʾiḏā* is either in the nominative case, e.g. *ʾiḏā l-ḥimāru* 'there is the ass', or takes the preposition *bi-*, e.g. *ʾiḏā bi-ḏ̣ulmatin šadīdatin* 'behold a great darkness' (Reckendorf 1895–1898:312). As is the case with demonstrative presentatives, presentative constructions with *ʾiḏā* may be proposition-centered, e.g. *ʾiḏā bi-rajulin yuqālu lahu s-sayyidu barakatu qad ʾaqbala* 'behold, a man called *sayyid* Baraka came forward' (Wright 1974:II, 158).

Some Arabic dialects use presentative particles that are derived from the verb 'to see'. In North African dialects, for example, *rā* (in origin an imperative form) functions as a presentative particle with a distribution that is similar to that of the demonstrative *hā* (for Moroccan Arabic, see Caubet 1993:II, 25–26). In Bedouin dialects of the Syrian-Mesopotamian desert and dialects of the Arabian Peninsula, a variant form with a pharyngeal /ʿ/ is used, e.g. Daṯīna *raʿni* 'here I am' (Fischer 1959:194). A form derived from the 2nd person masculine singular of the imperfect *tarā* is widely used across the dialects as a presentative. Several dialects (e.g. Lebanon, Syria, Daṯīna) have a presentative particle with the base form *šaʿ*, which appears to be a shortened form of the imperative for the verb *qašaʿ* 'to see'.

The Classical Arabic particle *ʾinna*, typically used to express the speaker's certainty concerning an assertion, is likely to be presentative in origin. This is suggested not only by its etymological connection with presentative particles in other Semitic languages, such as Hebrew *hinne*, but also by examples of an apparently presentative function that are cited by the medieval Arabic grammarians, e.g. *ʾinna zaydan* 'there is Zayd (over there)' (→ *ʾinna wa-ʾaxawātuhā*). It is likely that the normal function of *ʾinna* to express the certainty of a proposition has developed from such presentative constructions in a way that is similar to the structural shifts of demonstrative presentative constructions described above. The accusative case of the noun following *ʾinna* can be interpreted as a vestige of the object inflection of the original referent-centered presentative structure. A reflex of *ʾinna* with presentative meaning is preserved in some modern, mostly Bedouin dialects, e.g. *waṣalna larās marqab wənn nabʿ əlmāʾ* 'we reached a high elevation and there was the source of water' (ʿAnaze, Syrian desert, Bloch 1986:133).

Bibliographical references

Blau, Joshua. 1966–1967. *A grammar of Christian Arabic*. Louvain: Secrétariat du CorpusSCO.

Bloch, Ariel. 1986. *Studies in Arabic syntax and semantics*. Wiesbaden: O. Harrassowitz. [In particular 54–136.]

Caubet, Dominique. 1993. *L'arabe marocain*. I. *Phonologie et morphosyntaxe*. II. *Syntaxe et catégories grammaticales, textes*. Paris and Louvain: Peeters.

Fischer, Wolfdietrich. 1959. *Die demonstrativen Bildungen der neuarabischen Dialekte*. The Hague: Mouton. [In particular, 157–201.]

Mitchell, T.F. 1956. *An introduction to Egyptian Colloquial Arabic*. Oxford: Oxford University Press.

Mosel, Ulrike. 1975. *Die syntaktische Terminologie bei Sībawaih*. Ph.D. diss., University of Munich.

Reckendorff, Herman. 1895–1898, *Die syntaktischen Verhältnisse des Arabischen*. Leiden: E.J. Brill.

Woidich, Manfred. 2006. *Das Kairenisch-Arabische: Eine Grammatik*. Wiesbaden: O. Harrassowitz.

Wright, William. 1974. *A grammar of the Arabic language*. Repr. Beirut: Librairie du Liban.

Geoffrey Khan (University of Cambridge)

Pro-drop

The term 'pro-drop' has been used since the early 1970s to refer to languages that do not require pronouns or nouns to appear in subject position, i.e., the subject is 'understood' but not lexically expressed. A significant majority of the world's languages, approximately 61 percent of the 674 languages sampled in Dryer (2005), are of this type. When first introduced (Perlmutter 1971, 1972), pro-drop was proposed as a 'free deletion' rule (cf. Pronoun Deletion in Hankamer 1972) that applies in the grammars of certain languages, Arabic and Spanish, for example, but not English and French, to optionally delete subject pronouns. The examples in (1) illustrate this difference.

(1) a. Arabic *ana/Ø šuf-t il-film da*
I saw–1s the-film this
b. Spanish *Juan/Ø vió ese film*
c. English *I/*Ø saw this film*
(Jaeggli and Safir 1989b:9)

Since the early 1980s, and particularly in Chomsky's Principles and Parameters Theory, the term 'pro-drop' has been used to refer to a parameter, the pro-drop parameter, that distinguishes languages in which verbs must have an overt subject in tensed clauses from those in which they need not. The distinction was also seen as linked, in universal grammar, to several other parameters, or clusters of properties (discussed below), that would obtain in a language that is positively set for this parameter. The pro-drop parameter is also known as the 'null subject parameter', reflecting shifts in theory and a research paradigm focused on empty/null categories following developments in Chomsky's (1981) Government and Binding Theory. The pronoun is no longer 'dropped', i.e. deleted as originally proposed. Instead, it is present in the structural subject position as *pro*, an empty (phonologically null) category whose overt counterpart is lexical pronouns (see Sec. 3), as per Chomsky's (1982:74) typology of empty categories. See Jaeggli and Safir (1989b) for an overview of null subject issues within parametric theory, Baker (2001) for an interesting discussion of parameters, and Lightfoot (1993) on the relationship between null subject, language acquisition, and language change; also → anaphora, → binding, → clitics, → minimalism, → pronouns, → resumption, and other research on empty categories, expletives, and null subjects.

1. PARAMETRIC VARIATION

The interest in the pro-drop phenomenon in the general linguistics literature stems from the contrast it establishes between languages that allow subjects of tensed clauses to be null and those that do not and the potential this contrast may have for theories of linguistic typology. The claim that the presence of the null subject property in a language tends to correlate with a number of other syntactic phenomena suggests that it may be possible to predict a set of other contrasts on the basis of this one property of grammar. The notion 'rich agreement', originally postulated as characteristic of pro-drop languages, was quickly found hard to define crosslinguistically and not able by itself to predict the class of languages that permit null subjects and those that do not. There are null subject languages, such as Chinese and Japanese, that have little if any agreement and others, like French and English, that have some agreement but do not exhibit null subject properties. Jaeggli and Safir (1989b:38) propose the notion of 'morphological uniformity', instead: "The class of languages that permit null subjects, or within languages, exactly those paradigms that permit null subjects, are those paradigms that are morphologically uniform". The uniformity property, they argue, provides a unified account of null subjects across Japanese, Chinese, and Irish as well as across the Romance, Germanic, and Semitic languages. While this may be true, the question remains as to why uniformity correlates with the null subject phenomena and 'mixed' morphology with non-null subject languages.

Most prominent among the set properties that tend to be associated with the pro-drop parameter are phonologically null subject pronouns as in (1a-b), free subject inversion whereby a subject can occur before or after the verb as in (2), and *that*-trace violations allowing → WH-movement out of *that*-type clauses, as in (3).

(2) a. Arabic (Egyptian)
il-bint nāmit/nāmit ilbint
b. English
*The girl slept/*slept the girl*

(3) a. Italian
chi hai detto che fuma? (Kenstowicz 1989:263)
b. English
**Who did you say that smokes?*

Research, initially by Rizzi on Italian (1982), suggests that parametric variation may also occur among these properties as a result of dialect variation, for example. The modern Arabic dialects provide rich material for such studies. Kenstowicz (1989) is one of the first to explore this venue.

Kenstowicz examines systematic variation among the properties of the pro-drop parameter in relation to two Arabic dialects: Levantine (viz. Levantine, Palestinian, and Jordanian, which exhibit the same behavior with respect to the pro-drop parameter) and Bani-Hassan, a Bedouin Arabic dialect spoken in Jordan. He finds that while both dialects permit null subjects and free inversion, admittedly with some restrictions applying in Levantine, the dialects differ in relation to *that*-trace violations. Bani-Hassan in (4) exhibits a positive setting for this property, allowing WH-extraction of the *innu*-clause subject, whereas Levantine in (5) does not. Otherwise, they both allow WH-movement of the *innu*-clause object (4b, 5b) and of its subject if *innu* is not present (4c, 5c). In the presence of *innu*, a resumptive pronoun is used (4d, 5d).

(4) Bani-Hassan
  a. _wayy binit$_i$ farīd gāl innu e$_i$ ištarat al-libās_
     'Which girl did Farid say that bought the dress?'
  b. _wayy libās$_i$ farīd gāl innu al-binit ištarat e$_i$_
     'Which dress did Farid say that the girl bought?'
  c. _wayy binit$_i$ farīd gāl e$_i$ ištarat al-libās_
     'Which girl did Farid say bought the dress?'
  d. _wayy binit$_i$ farīd gāl inn-ha$_i$ ištarat al-libās_
     'Which girl did Farid say that bought the dress?'

(5) Levantine
  a. *_ʾayy bint$_i$ farīd kāl innu e$_i$ ištarat l-fusṭān_
     'Which girl did Farid say that bought the dress?'
  b. _ʾayy fusṭān$_i$ farīd kāl innu l-bint ištarat e$_i$_
     'Which dress did Farid say that the girl bought?'
  c. _ʾayy bint$_i$ farīd kāl e$_i$ ištarat l- fusṭān_
     'Which girl did Farid say bought the dress?'
  d. _ʾayy bint$_i$ farīd kāl inn-ha$_i$ ištarat l-fusṭān_
     'Which girl did Farid say that (she) bought the dress?'

Levantine Arabic, according to Kenstowicz, exhibits negative setting for all three properties associated with pro-drop (i.e. no null subjects, free inversion, or WH-extraction) within *innu*-clauses. Other data from the Bani-Hassan dialect show that the setting of this parameter can be affected by the nature of the verbal inflection of the clause. If the predicate of the *innu*-clause is not fully inflected, as in the case of participles and other nonverbal predicates (see Sec. 2 for details), WH-extraction is not possible, and the resumptive pronoun is required as in (6).

(6) _wayy binit$_i$ farīd gāl *innu/inn-ha$_i$ mištarya al-libās_
  'Which girl did Farid say that (she) bought the dress?'

The Arabic data demonstrate, as Kenstowicz (1989:274) concludes, that the setting for this parameter can be a function of clause and predicate type rather than being "a gross feature of the grammar as a whole". Other research on Arabic, explored below, supports this conclusion. The variation itself, however, remains unexplained. Kenstowicz's article brings to focus an area of research combining dialect and parametric variation that may prove fruitful for future research.

2. Pro-drop and agreement

The bulk of the literature on pro-drop in Arabic has been focused on providing a principled and explanatory account for the distribution of pronominal and null subjects. Most if not all such accounts attribute the absence of subject pronouns to the 'rich' inflection and agreement features carried by the verb. The sample paradigm in (7) illustrates the perfect and imperfect conjugations of the verb *xarag* 'to go out' based on Egyptian Arabic, which can be viewed as representative of the underlying system shared by most varieties of Arabic including Standard. Full paradigmatic conjugations are available elsewhere in the literature, e.g. Benmamoun (2000:19–25) for Egyptian, Moroccan, and Standard Arabic; Fassi Fehri (1993:98–140) for an analysis of the Standard Arabic pronominal system based on incorporation; Eid (1996) for Egyptian within the context of bilingual grammars.

(7)

Pronoun	Perfect	Imperfect		Pronominal clitics	
(nominative)		Present	Future	(non-nominative)	
huwwa	*xarag-Ø*	*bi-yu-xrug*	*ḥa-yu-xrug*	*-u*	3ms
hiyya	*xarag-it*	*bi-tu-xrug*	*ḥa-tu-xrug*	*-ha*	3fs
humma	*xarag-u*	*bi-yu-xrug-u*	*ḥa-yu-xrug-u*	*-hum*	3p
inta	*xarag-t*	*bi-tu-xrug*	*ḥa-tu-xrug*	*-ak*	2ms
inti	*xarag-ti*	*bi-tu-xrug-i*	*ḥa-tu-xrug-i*	*-ik*	2fs
intu	*xarag-tu*	*bi-tu-xrug-u*	*ḥa-tu-xrug-u*	*-kum*	2p
ana	*xarag-t*	*bi-ʾa-xrug**	*ḥa-ʾa-xrug**	*-ni/i/ya*	1s
iḥna	*xarag-na*	*bi-nu-xrug*	*ḥa-nu-xrug*	*-na*	1p

The two forms in the conjugation of the 1st person singular *ana* marked with an asterisk are given in their underlying forms for the sake of paradigm regularity; they are pronounced as *baxrug* and *ḥaxrug* in accordance with Egyptian Arabic phonology.

The verb inflects for all pronominal features (person, gender, and number), except for the imperative, which inflects for gender and number but only has a 2nd person, as may be expected in imperatives: *ʾiktib/ʾiktib-i/ʾiktib-u* for 'you [masc. sg.]/you [fem. sg.]/you [pl.] write!', respectively. This inflection system makes the identification of the subject totally recoverable from the subject agreement inflection on the verb and thus somewhat redundant from a discourse perspective. Results from Parkinson (1987) suggest that the 'norm' in Arabic might be for subject pronouns to be missing (see Sec. 4). Parkinson's analysis of 3,200 clauses shows that the distribution significantly favors null ('unexpressed') subjects (47%) over nominal subjects (36%) and pronominal subjects (17%), with results combined for 1st, 2nd, and 3rd person pronouns.

In addition to verbs, other lexical categories can also serve as predicates. In copular constructions (*be-* and *have-* predicates), the copular verb *kān* is not lexical in certain tense-aspect configurations, in which case adjectives, nouns, and prepositional phrases serve as predicates, as in (8) (→ copula). Adjectives and nouns, however, inflect for gender and number only, and prepositions do not typically inflect for any pronominal features, although they take pronouns as their objects. In these contexts pro-drop is not possible, as shown in (8).

(8) Nouns a. **(huwwa) mudarris*
'He [is] a teacher'
b. **(inti) mudarris-a*
you.2fs teacher-fs
'You [are] a teacher'
Prepositions a. **(huwwa) fi l-bēt*
'He [is] in the house'
b. **(hiyya) gamb-ak*
'She [is] beside you'
Adjectives a. **(hiyya) ṭawīl-a*
she tall-fs
'She [is] tall'
b. **(iḥna) fahmīn*
we understanding-Part.pl
'We understand'

When the pronoun is absent, the structure cannot be understood as a sentence since its subject cannot be identified from within. Discourse conditions, however, may allow recoverability, hence identification, of the missing subject, as in the question-answer sequences in (9).

(9) Q: *šakl-aha ʾēh?* A: *ṭawīla*
'Q: What does she look like? A: Tall'
Q: *huwwa fēn?* A: *fi l-bēt*
'Q: Where is he? A: At home'

In addition to agreement features, the Arabic pronominal system distinguishes two sets of pronouns: independent and clitic (as in (7)). Independent pronouns appear in Specifier positions and are assigned nominative case: as subject in Spec of IP and as topic in Spec of CP; clitics appear in non-nominative configurations and receive their case assignment accordingly. Non-nominative pronouns attach to verbs (*šuft-ak* 'I saw you [masc. sg.]'), prepositions (*gamb-u* 'beside him'), to certain complementizers (*inn-ak* 'that you [masc. sg.]'), and to nouns in the construct state (*kitab-hum* 'their book'). They do not attach to adjectives, however. The examples in (10), when compared with (8), show that it is the presence of these pronominal clitics on the predicates (nouns and prepositions) that allows pro-drop.

(10) a. Nouns *(humma) nifs-uhum*
they desire-their
fi l-fūl
for beans
'They want beans'
b. Prepositions *(ana) warā-ya*
I behind-me
šuġl
work
'I have work [to do]'

Preposition-clitic combinations that can occur as predicates in copular constructions are unrestricted. The set of nouns that can do so, however, is very limited and subject to dialect variation. In Egyptian Arabic they are mostly psych nouns that identify a cognitive or emotional state such as desire or intention, which as predicates select thematic subjects that can experience such states (cf. Jelinek 2002). They also carry the sentential negation *ma-š* characteristic of verbs (→ negation).

(11) a. Verbs *(hiyya) ma-katab-it-š il-gawāb*
'She did not write the letter'
b. Nouns *(ana) ma-nifs-ī-š fi l-fūl*
'I don't want beans'
c. Prepositions *(ana) ma-wara-yā-š šuġl*
'I don't have work to do'

Thus, the presence of pronominal features on the predicate is a necessary condition for pro-drop in Arabic, as (12) clearly shows. The only difference between (12a) and (12b) is the presence of the lexical copula (a form of *kān*) in the former but not the latter.

(12) a. *(hiyya) ma-kanit-š fi l-bēt/mudarrisa/šaṭra/nayma*
'She was not at home/a teacher/clever/sleeping'
b. *(hiyya) ma-hiyyā-š fi l-bēt/mudarrisa/šaṭra/nayma*
'She is not at home/a teacher/clever/sleeping'

In the absence of a lexical copula, as in (12b), the pronominal features surface with the lexical negative as part of Inflection (I head of INFL) and are spelled out as 'negative copular pronouns' (Eid 1983a, 1991) or 'negative copula' (Jelinek 2002).

There is consensus in the literature that the inflectional feature [Person] must be present, however. Kenstowicz (1989:272), for example, argues that two features of verbal inflection play a crucial role in core grammar: whether the verb is tensed or not and whether the verb shows person agreement with the subject. Four possible combinations are predicted on the basis of these two features: [+Tense +Person] Finite, [-Tense +Person] Subjunctive, [+Tense -Person] Participle, [-Tense -Person] Infinitive. Arabic allows only the first three. Since it has no infinitive, the subjunctive form is used in this context, e.g. Egyptian Arabic *nadya*$_i$ *ʿawz-a [e*$_i$ *tu-xrug]* 'Nadia wants [fem. sg.] to go [fem. sg.]'. In analyses of English, the empty category that appears in subject position of infinitives is the anaphoric PRO. The nature of this empty subject position in Arabic has received little, if any, discussion in the literature on null subjects.

It is the presence of the feature [Person] then, together with [Tense] in inflection, that is characteristic of pro-drop environments. Fassi Fehri (1993) suggests a hierarchy of phi feature specification: Person < Number < Gender, whereby the presence of [Person] implies Number and Gender (see Eid 1983a, 1991; Mohammad 1988; Benmamoun 2000; Jelinek 2002; Fassi Fehri 1993; Kenstowicz 1989, among others).

3. NULL SUBJECTS

To satisfy the 'pro-drop' environment, a predicate phrase must be specified for the two features [+Tense] and [+Person]. According to the typology of empty categories, the empty pronoun, *pro*, fills the null subject position in such contexts; it is the non-overt counterpart of lexical pronouns. Null subjects also occur as a result of movement operations as, for example, WH-movement in (4) and (5), where a trace occurs in the *innu*-clause co-indexed with the WH-constituent.

In certain contexts, however, null subjects appear but the predicate does not show agreement as expected. Mohammad (1990, 2000) argues that Standard Arabic does not have NP raising since sentences like English *the girls seem to be happy* are not possible.

The examples in 13 (from Mohammad 1990) show that in *seem*-type sentences the verb *yabdū* ‘seems’ remains in its uninflected form, showing no agreement with the NP (*al-banāt*) that precedes it, from which he concludes that it cannot be the subject of *yabdū*.

(13) a. pro *yabdū* *ʾanna* *l-banāti*
seems–3ms that the-girls
sāfarna
departed.3fp
‘It seems that the girls have departed’
b. **al-banāt-u*$_i$ *yabdūna* *ʾanna*
the-girls$_i$ seem–3fs that
pro$_i$ *sāfarna*
pro$_i$ departed.3fp
c. *al-banātu* pro *yabdū*
the-girls seems–3ms
ʾanna-hunna *sāfarna*
that-they.fp departed.3fp

The same analysis is applied to sentences with modals, e.g. *yajibu ʾan yadrusa l-ʾawlād-u* ‘the boys must study’, *tough*-movement verbs, e.g. *sahl-un ʾan tanjaḥa l-bint-u* ‘it is easy for the girl to succeed’, and ‘subjectless’ passives, e.g. *yuxšā ʾan tahruba s-sajīnat-u* ‘it is feared that the prisoner will escape’. In each case, if the subject of the embedded *ʾan*-clause appears before the main clause verb, the verb does not inflect for agreement with that NP; it is always in the 3rd person masculine singular, which prompts Mohammad to suggest that the subject of these predicates is expletive *pro*. For an interesting discussion of pronouns and expletives, see also Fassi Fehri (1993:118–120), who argues that independent 3rd person pronouns in Arabic (Standard, in this case) are three-way ambiguous in that they can function as personal pronouns, as (pronominal) copulas, and as expletives.

Finally, Osman (1987) provides evidence showing that null (*pro*) and lexical pronominal subjects in Egyptian Arabic may need to be distinguished from each other. In certain contexts involving backward → pronominalization, the lexical pronoun cannot be co-indexed with a following antecedent, whereas the null pronoun can. Whereas the null pronoun in (14a) is free to take an antecedent (within or outside its clause), the lexical pronoun in (14b) is not.

(14) a [[il-fusṭān [illi pro$_{i/j}$ ištar-it-u]]
ʿāgib nadya$_i$
‘Nadia$_i$ likes the dress that
she$_{i/j}$ bought’
b. [[il-fusṭān [illi hiyya$_i$ ištar-it-u]] ʿāgib
Nadya$_i$
‘Nadia$_i$ likes the dress that she$_j$
bought’

Osman concludes that lexical pronouns in Egyptian Arabic, and possibly other null subject languages, may have to be treated as Referring expressions [+anaphoric, +pronominal], rather than as pronouns, and thus subject to Binding Condition (C).

4. PRO-DROP IN DISCOURSE CONTEXT

One of the earliest investigations into principles that govern the presence or absence of pronominal subjects in Arabic is found in Eid (1983b), who questions the account of pro-drop as a ‘free deletion’ rule that “does not require either another constituent to trigger the deletion or the statement of the environment in which it applies” (Perlmutter 1972:104). Eid argues that subject pronouns serve two communicative functions: they serve as anti-ambiguity devices and as indicators of subject switch. Her arguments are based on evidence from relative clauses, subordinate clauses, coordinate clauses, and question-answer discourses.

The anti-ambiguity function is illustrated in (15), where a pronominal subject has more than one possible antecedent within the sentential structure, hence more than one possible interpretation for the sentence.

(15) *ʿali* *kallim* *il-walad* *illi* *Ø*
Ali talked-to the-boy who
šatam-u *imbāriḥ*
insulted-him yesterday
a. ‘Ali talked to the boy who insulted him (Ali) yesterday’ = Favored
b. ‘Ali talked to the boy that he (Ali) insulted him yesterday’ = Less favored

The two NPs in the matrix clause (*ʿali* and *il-walad*) are possible antecedents for the subject and object NPs in the relative clause. The presence of the pronoun, as in (15c), forces the less-favored interpretation (where *ʿali* is subject in the relative clause).

c. *ʿali kallim il-walad illi huwwa šatam-u imbāriḥ*
'Ali talked to the boy that he (Ali) insulted (him) yesterday'

The subject switch function, illustrated in (16), differs in that no potential ambiguity is involved. The subject-to-subject reading in (16a) is the unmarked, hence expected, interpretation. The presence of the pronoun in (16b) indicates the opposite, i.e., the reading is different from the expected one, hence the switch to the object NP, the other antecedent available for the subject pronoun *huwwa* in the second clause.

(16) a. *ʿali$_x$ ḍarab samīr$_y$ wi Ø šatam-u$_y$*
'Ali hit Samir and insulted him'
b. *ʿali$_x$ ḍarab samīr$_y$ wi huwwa$_y$ šatam-u$_x$*
'Ali hit Samir and he (Samir) insulted him'

The analysis captures the idea that, unless otherwise specified, the norm or unmarked case is for the subject pronoun not to be present, consistent with the 'Avoid a pronoun' principle (Chomsky 1981). Its presence provides a signal to the hearer for an unexpected interpretation. In this sense, subject pronouns carry a communicative load worth further investigation, although some may argue it is still related to 'emphasis'.

Such an investigation is provided in Parkinson (1987), who conducted a quantitative study to test the notions proposed in Eid and others, such as Givon's (1983) 'Topic Accessibility'. Parkinson's study simulates the conversational set-up by using data collected from two Egyptian Arabic plays. The data consist of more than 3,200 clauses, each coded for several variables including subject switch as well as person of subject, type of clause, and type of predicate. The results and conclusions selected for presentation below are perhaps the most pertinent to the discussion above.

Parkinson's overall conclusion is that "the choice of a pronoun subject over an unexpressed subject is not simply a free variation, and that it probably cannot be accounted for by the anti-ambiguity principle, but that it clearly is a kind of constrained variation with the notions of subject switch (a kind of contrast), person, Lookback, predicate type and clause type all involved in constraining the variation" (1987:359).

The Anti-Ambiguity function of pronouns could not be confirmed because not enough examples appeared in the sample that fit this context (Parkinson 1987:353). Of the 218 relative clauses in the corpus (188 if only pronoun subjects and unexpressed subjects are included), only 14 had pronoun subjects, and of these only 2 were 3rd person and thus subject to the kind of ambiguous possibilities discussed in Eid. Over half have a participle or other equational sentence predicate, rather than a verb, which is sufficient to explain their presence independent of any anti-ambiguity. Only 5 percent of the relative clauses with verbs have pronoun subjects, while almost 54 percent of the relative clauses with participles have pronoun subjects. Predicate type turned out to be an important constraining factor in clauses other than relative clauses as well, for example in *lamma* 'when' clauses.

The Subject Switch function, however, was confirmed. The results of the analysis are significant and indicate that when the subject is the same as the preceding clause, less than one-fifth of the clauses have a pronoun subject, while over one-third of the clauses with a subject different from that of the preceding clause have pronoun subjects. Parkinson concludes that "Eid's notion of pronouns marking an 'unexpected' subject change is both a valid concept and one of an even higher degree of generality than she implied. 'Different' subjects are forcing a higher percentage of pronouns than would otherwise be expected" (1987:354). However, neither Eid's notion of 'Subject Switch' nor Givon's 'Topic Accessibility' (i.e. the longer it has been since a topic has been in the register, the more likely it is that a pronoun rather than an unexpressed subject would be used) are able to explain the entire variable process. Parkinson finds that they are clearly to be reckoned with as part of a final, overall explanation.

The results reported regarding the overall distribution of pronominal vs. null subjects are also interesting. Subject pronouns, for example, are 'omitted' approximately 74 percent of the time with a frequency consistent in both plays when analyzed separately. This finding supports the idea that in pro-drop languages the absence of the pronoun is the unmarked case and is consistent with the analysis attributing pro-drop to agreement and person inflection. Predicate type turned out to be a significant factor as well. Clauses with conjugated verbs had a very consistent 80 percent unexpressed subjects whereas equational sentences, with or without a predicate participle, had 50 percent each. Parkinson also found that 'Person' is an independent constraint affecting the outcome of pronoun frequency in all of its environments. For clauses with verbal predicates, the frequency of subject pronouns for each person (1st, 2nd, 3rd) is significantly lower than that of their unexpressed counterpart. This pattern is observed across persons as well, so that there are significantly more 1st person pronouns than there are 2nd person pronouns, and more 2nd person pronouns than 3rd person.

Bibliographical references

Baker, Mark. 2001. *The atoms of language*. New York: Basic Books.

Benmamoun, Elabbas. 2000. *The feature structure of functional categories*. Oxford: Oxford University Press.

——, and Heidi Lorimor. 2006. "Featureless expressions: When morphological markers are absent". *Linguistic Inquiry* 37:1.1–23.

Bolotin, Naomi. 1995. "Arabic and Parametric Agreement". *Perspectives on Arabic linguistics*, VII, ed. Mushira Eid, 7–27. Amsterdam and Philadelphia: J. Benjamins.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

——. 1982. *Some concepts and consequences of the theory of Government and Binding*. Cambridge, Mass.: MIT Press.

Dryer, Matthew S. 2005. "Expression of pronominal subjects". *The world atlas of language structures*, ed. Martin Haspelmath, Matthew S. Dryer, David Gil, and Bernard Comrie, 410–413. Oxford and New York: Oxford University Press.

Eid, Mushira. 1983a. "The copula function of pronouns". *Lingua* 59.197–207.

——. 1983b. "On the communicative function of subject pronouns in Arabic". *Journal of Linguistics* 19.287–303.

——. 1991. "Verbless sentences in Arabic and Hebrew". *Perspectives on Arabic linguistics*, III, ed. Bernard Comrie and Mushira Eid, 31–61. Amsterdam and Philadelphia: J. Benjamins.

——. 1993. "Predicates and their heads". *Principles and prediction: The analysis of natural language*, ed. Mushira Eid and Gregory Iverson, 135–152. Amsterdam and Philadelphia: J. Benjamins.

——. 1996. "Things are not what they seem: Pronoun doubling in monolingual and bilingual grammars". *Perspectives on Arabic linguistics*, VIII, ed. Mushira Eid, 7–31. Amsterdam and Philadelphia: J. Benjamins.

Fassi Fehri, Abdelkader. 1993. *Issues in the structure of Arabic clauses and words*. Dordrecht: Kluwer Academic Publishers.

Givon, Talmy. 1983. "Topic continuity in discourse: An introduction". *Topic continuity in discourse*, ed. Talmy Givon, 1–42. Amsterdam and Philadelphia: J. Benjamins.

Hankamer, Jorge. 1972. *Constraints on deletion in syntax*. Ph.D. diss., Yale University.

Jaeggli, Osvaldo and Kenneth Safir (eds.). 1989a. *The null subject parameter*. Dordrecht: Kluwer Academic Publishers.

——, and Kenneth Safir. 1989b. "The null subject parameter and parametric theory". Jaeggli and Safir (1989a:1–44).

Jelinek, Eloise. 2002. "Agreement, clitics and focus in Egyptian Arabic". *Themes in Arabic and Hebrew syntax*, ed. Jamal Ouhalla and Ur Shlonsky, 71–106. Dordrecht, Boston, and London: Kluwer Academic Publishers.

Kenstowicz, Michael. 1989. "The null subject parameter in modern Arabic dialects". Jaeggli and Safir (1989a:263–275).

LeTourneau, Mark. 1993. "Case-marking and binding of subject clitics in Arabic complement clauses". *Perspectives on Arabic linguistics*, V, ed. Mushira Eid and Clive Holes, 261–290. Amsterdam and Philadelphia: J. Benjamins.

Lightfoot, David. 1993. *How to set parameters: Arguments from language change*. Cambridge, Mass.: MIT Press.

Mohammad, Mohammad A. 1988. "Nominative case, I-subjects and subject-verb agreement". *Proceedings of the Chicago Linguistic Society Parasession on Agreement and Grammatical Theory*, ed. Diane Brentari, Gary Larson, and Lynn MacLeod, 223–235. Chicago: Chicago Linguistics Society.

——. 1990. "The problem of subject-verb agreement in Arabic: Towards a solution". *Perspectives on Arabic linguistics*, I, ed. Mushira Eid, 95–125. Amsterdam and Philadelphia: J. Benjamins.

——. 2000. *Word order, agreement and pronominalization in Standard and Palestinian Arabic*. Amsterdam and Philadelphia: J. Benjamins.

Osman, Mariam. 1987. "Bound anaphora in Egyptian Arabic". *Perspectives on Arabic linguistics*, I, ed. Mushira Eid, 156–172. Amsterdam and Philadelphia: J. Benjamins.

Parkinson, Dilworth. 1987. "Constraints on the presence/absence of 'optional' subject pronouns in Egyptian Arabic". *Variation in Language: NWAV-XV at Stanford*, ed. Keith M. Denning, Sharon Inkelas, Faye C. McNair-Knox, and John R. Rickford, 348–360. Stanford, Cal.: Department of Linguistics, Stanford University.

Perlmutter, David. 1971. *Deep and surface structure constraints in syntax*. New York: Holt, Rinehart, and Winston.

——. 1972. "Evidence for shadow pronouns in French relativization". *The Chicago which hunt*,

ed. Paul M. Peranteau, Judith N. Levi, and Gloria C. Phares, 73–105. Chicago: Chicago Linguistic Society.
Rizzi, Luigi. 1982. *Issues in Italian syntax*. Dordrecht: Foris.

Mushira Eid (University of Utah)

Pronominalization

1. Introduction

Pronominalization is a term that is used in different contexts (Crystal 2003:376). In traditional → transformational grammar, pronominalization is a rule replacing lexical items with a pronoun, whereas later approaches analyzed the pronouns as being generated in the base. In Government and → Binding theories, pronominals are a type of noun phrase with special government properties. On pronominalization in Standard and colloquial Arabic in this sense, see, for instance Mohammad (1999).

In this entry, the term 'pronominalization' is used in a text-linguistic context. Text linguists emphasize the function of pronominal or anaphorical reference (→ anaphora) as a device to avoid the repetition of lexical items and as a means to achieve → cohesion (Halliday and Hasan 1976). Languages differ as to the selection of lexical repetition and pronominalization. Levinson (quoted after Blum-Kulka) claims that "given the choice between lexical repetition and pronominalization, Hebrew writers tend to prefer the former while English writers tend to choose the latter" (cf. Blum-Kulka 2000:300). Berman (1978) argues that for both Hebrew and English writers, pronominalization is preferable, but lexical repetition is more common in Hebrew than in English. As for Standard Arabic, both Obeidat (1998) and Aziz (1998) claim that it prefers lexical repetition to pronominalization. Also, Baker (1992:207) has made a generalization to the effect that "Arabic prefers lexical repetition to variation". Yet, in Standard Arabic, pronominalization is an effective cohesive device that deserves more attention. Moreover, evidence drawn from the usage of free and bound pronominal forms shows that Standard Arabic is no different from languages such as English, in that it does not resist pronominalization.

2. Subject and object personal pronouns in Standard Arabic

Standard Arabic has two sets of pronominal forms: a set of free forms and "a set of bound pronominal forms which can be suffixed to verbs, nouns, prepositions, and particles of various types" (Holes 1995:145).

Standard Arabic has twelve subject pronominal forms which exhibit person, number, and gender distinction. In addition, it has a set of bound forms that are semantically and phonologically related to the independent forms, as in Table 1.

Standard Arabic has two 1st person forms: a singular form (*ʾana*) and a plural form (*naḥnu*), but there is no gender distinction in this subcategory. In the 2nd person, Standard Arabic has two singular forms that exhibit gender distinction: a singular masculine form (*ʾanta*) and a feminine form (*ʾanti*). It also has two plural forms exhibiting gender distinction: *ʾantum* (masc.) and *ʾantunna* (fem.). In addition, there is only one dual neutral form (*ʾantumā*). In the 3rd person, Standard Arabic has two singular forms: a masculine form (*huwa*) and a feminine form (*hiya*). It also has two plural forms, one masculine (*hum*), the other feminine (*hunna*), but only one neutral dual form (*humā*). Thus, Standard Arabic subject pronouns exhibit gender distinction only in the 2nd and 3rd person. The dual forms exhibit no gender distinction: the two forms (*ʾantumā* and *humā*) are gender-neutral.

Table 1. Standard Arabic subject pronouns and pronominal clitics

		sg	du	pl
1st		*ʾana (-tu)*		*naḥnu* (-*nā*)
2nd	masc.	*ʾanta* (-*ta*)	*ʾantumā* (-*tumā*)	*ʾantum* (-*tum*)
	fem.	*ʾanti* (-*ti*)		*ʾantunna* (-*tunna*)
3rd	masc.	*huwa* (Ø)	*humā* (-*ā*)	*hum* (-*ū*)
	fem.	*hiya* (-*t*)		*hunna* (-*na*)

The independent pronouns normally substitute for the nominative functions of nouns or noun phrases, as in (1).

(1a) *ʾana* *katab-tu* *d-dars-a*
I wrote-1s the-lesson-Acc
'I wrote the lesson'
(1b) *ʾanta* *katab-ta* *d-dars-a*
you.ms wrote-2ms the-lesson-Acc
'You wrote the lesson'

Table 1 also shows that each independent pronoun has a bound form with which it is associated. This has motivated Arab grammarians to analyze such verb forms as *ḍarabtu* "as combinations of a verb and a pronoun" (Versteegh 1997:84). Thus, the presence of a bound pronominal form on the verb can trigger dropping of the personal pronoun in the above sentences without affecting their semantic or syntactic structure (→ pro-drop). This can be done since the subject is already encoded on the verb via the agreement bound forms that are attached to the verb stem, i.e., the verb is inflected for person, gender, and number, as in (2).

(2a) *katab-tu* *d-dars-a*
wrote-1s the-lesson-Acc
'I wrote the lesson'
(2b) *katab-ta* *d-dars-a*
wrote-2ms the-lesson-Acc
'You wrote the lesson'
(2c) *katab-nā* *d-dars-a*
wrote-1p the-lesson-Acc
'We wrote the lesson'

Like subject pronouns, object pronouns can be bound or freestanding. Their form is the same, but the independent object pronouns have to be introduced by the object particle *ʾiyyā-*. They are distributed exactly like the independent subject pronouns along the following lines: person, number, and gender (Table 2).

Note that there is no gender distinction in the 1st person; there are only two common forms, one for the 1st person singular (*ʾiyyā-ya*) and the other for the 1st person plural (*ʾiyyā-nā*). In the 2nd person, the forms exhibit gender distinctions in the singular (*ʾiyyā-ka/ʾiyyā-ki*) and plural (*ʾiyyā-kum/ʾiyyā-kunna*). There is also a dual common form (*ʾiyyā-kumā*). In the 3rd person, gender is manifest in the singular and plural: *ʾiyyā-hu* and *ʾiyyā-hā* stand for a singular masculine and a singular feminine, respectively. Likewise, *ʾiyyā-hum* and *ʾiyyā-hunna* stand for a plural masculine and a plural feminine, respectively. In addition, there is a 3rd person dual common form (*ʾiyyā-humā*).

These independent pronouns stand for the object in a clause, as in (3).

(3) *ʾiyyā-ka* *naʿbudu*
Obj-you.2ms we-worship
wa-ʾiyyā-ka *nastaʿīn*
and-Obj.-you.2ms we-depend
'We worship you and we depend on you' (Q. 1/5)

The independent form is used for emphasis only. In other cases, the bound form is more common, as in (4).

(4) *ḍaraba-ka* *wa-ṭarada-ka*
hit.3ms-you.2ms and-cursed.3ms-you.2ms
'He hit and cursed you'

Arab grammarians do not make a distinction between subject and object dependent or bound forms. However, the difference between these two types is obvious: the forms that are associated with the subject must be used even if the subject appears in the sentence, as in (5a).

Table 2. Standard Arabic object pronouns and the pronominal clitics

		sg	du	pl
1st		*ʾiyyā-ya (-nī)*	*ʾiyyā-nā (-nā)*	
	masc.	*ʾiyyā-ka (-ka)*		*ʾiyyā-kum (-kum)*
2nd			*ʾiyyā-kumā (-kumā)*	
	fem.	*ʾiyyā-ki (-ki)*		*ʾiyyā-kunna (-kunna)*
	masc.	*ʾiyyā-hu (-hu)*		*ʾiyyā-hum (-hum)*
3rd			*ʾiyyā-humā (-humā)*	
	fem.	*ʾiyyā-hā (-hā)*		*ʾiyyā-hunna (-hunna)*

(5a) *ḍarab-nā* *naḥnu* *l-walad-a*
hit-1p we the-boy-Acc
'We hit the boy'
(5b) **ḍarab* *naḥnu* *l-walad-a*

In contrast, the object bound forms can be used even if there is a direct object in the sentence, as in (6a).

(6a) *ḍaraba-hum* *ar-rajul-u*
hit3ms-them the-man-Nom
'The man hit them'
(6b) **ḍaraba-hum* *ar-rajul-u*
hit3ms-them the-man-Nom
l-luṣūṣ-a
the-thieves-Acc
'The man hit the thieves'

This means that the object bound forms are → clitic object pronouns that replace the pronoun or noun they stand for, whereas the subject bound forms are agreement suffixes that can co-occur with the independent pronoun or noun to which they refer.

Moreover, the bound forms derived from the independent subject pronouns are attached to a verb stem only, as in (2). In contrast, the object clitics can be attached to verbs, nouns, prepositions, and particles of various types (Holes 1995:145), as in (7) and (8).

(7) *qaddama* *l-walad-u*
presented.3ms the-boy-Nom
l-kitab-a *la-hā*
the-book-Acc to-her
'He presented the book to her'
(8) *ʾaqāma* *fī* *manzili-hā*
stayed.3ms in house-her
'He stayed at her home'

3. DOES STANDARD ARABIC RESIST PRONOMINALIZATION?

Obeidat (1998) and Aziz (1998) argue that Standard Arabic does not prefer pronominalization, as shown by the difference between the English sentence *She was seated at her dressing table* (from Dickens' *Great expectations* 121) and its Arabic translation in (9).

(9) *kānat* *tajlisu* *ʿalā*
was.3fs 3fs.sit to
ṭāwilat-i *t-tajmīl-i*
table-Gen the-beauty-Gen

Such claims usually do not rely on text corpora to verify the generalization presented, and therefore do not reflect actual language use in an Arabic text. In Standard Arabic, a personal pronoun is often used as a cohesive device. For example, Nobel laureate Naguib Mahfouz uses personal reference as a cohesive device more than seven hundred times in his novel *al-Liṣṣ wa-l-kilāb* 'The thief and the dogs'. A longer passage is given as an example in (10).

(10)
تجنب الطريق الملاصق للثكنات، واخترق الصحراء نحو
مدفن الشهيد ليبلغه فى أقصر وقت. وكان كأنما يهتدى ببوصلة
مركبة فى رأسه لسابق درايته بصحراء العباسية. وعندما لاحت
له قبة المدفن الضخمة تحت ضوء النجوم راحت عيناه تفتشان
عن المكان الذى تنزوى فيه السيارة. ودار حول المدفن وهو يحد
بصره ولا يعثر على ضالته حتى بلغ ضلعه الجنوبى فتراءى له
شبح هيكلها راقدا على بعد. مضى نحوها مصمما، ثم ما لبث أن
أحنى ظهره حتى انخفض ظهره إلى مستوى ركب

'He knew this stretch of ground. Avoiding the road next to the barracks, he set out across the desert to reach the Martyr's Tomb in the shortest time possible, heading for it as if he had a compass built into his head. As soon as he saw the tomb's big dome in the starlight he began looking for the spot where the car would be tucked away. Walking around the tomb, he scanned the ground as sharply as he could, but it was only when he reached its southern wall that the shape at a little distance became visible. He made for it without another thought, keeping his head low...' (*Liṣṣ* 52; trans. 63)

This example demonstrates that the Arabic text does not resist pronominalization. The result of text counting in the source text is in keeping with Halliday and Hasan's prediction that "the third person forms constitute the most frequent single class of cohesive items" (1976:49).

Another evidence that Arabic does not resist pronominalization comes from the cataphoric use of the 3rd person pronoun (→ cataphora), as in the first sentences of Mahfouz's novel.

(11)

مرة أخرى يتنفس نسمة الحرية، ولكن الجو غبار خانق
وحر لا يطاق. وفى انتظاره وجد بدلته الزرقاء وحذاءه المطاط،
وسواهما لم يجد فى انتظاره أحدا.

> Once more he breathed in the air of freedom. But there was stifling dust in the air, almost unbearable heat, and no one was waiting for him, nothing but his blue suit and gym shoes> (*Liṣṣ* 7; trans. 13)

Note that the 3rd person pronoun does not anaphorically refer to any person since no person has been introduced. This cataphoric use of the 3rd person singular is intended to arouse the interest of the readers and to keep them in anticipation of the person who is being talked about.

In Standard Arabic, the 3rd person subject pronoun can be implicit, as in (12).

(12)

ولحق بهما كثيرون من الدكاكين على الجانبين، وارتفعت حرارة
التهانى، وسرعان ما وجد نفسه مطوقا من جميع الجهات بحشد من
أصدقاء غريمه ولا شك

> 'People came up to them from the shops on both sides of the street; voices were loud and warm in congratulation and Said found himself surrounded by a crowd – his enemy's friends, no doubt' (*Liṣṣ* 9; trans. 16)

In the above extract the implicit pronoun (*huwa*), which is implied by the verb *wajada* 'he found', is rendered as a proper noun (*Said*) in translation. There is nothing in the text which should hinder the use of a pronominal reference.

Even the presence of two referents does not sway writers from using pronominalization in Standard Arabic. In such a case, the use of pronominalization can result in ambiguity of reference, as in (13).

(13)

فقال له الشيخ إنه يطالبه بالبطاقة ليتأكد من أنه من الخاطئين
لأنه لا يحب المستقيمين فقدم له مسدسه وقال له ثمة قتيل وراء
كل رصاصة فى ماسورته

> 'When the Sheikh replied that he did not like the righteous and wanted to see Said's identity card to make sure that Said was really a sinner, Said handed him the revolver, explaining that every missing bullet meant a murder…' (*Liṣṣ* 64–65; trans. 77)

Note that the underlined parts of the original text given in (11) implicitly or explicitly refer to Said; the bold-faced forms refer to the Sheikh. The above excerpt shows that the pronominal form can refer to either Said or the Sheikh, which contributes to the ambiguity of reference confronted by the reader.

The conclusion must be that Standard Arabic is no different from languages such as English in that it does not resist pronominalization as Obeidat (1998) and Aziz (1998) claim.

Bibliographical references

Primary source

Mahfouz, *Liṣṣ* = Naguib Mahfouz, *al-Liṣṣ wa-l-kilāb*. Cairo: Maktaba Miṣr, 1977. (1st ed., 1961.) Trans. Trevor L. Gassick and M.M. Badawi, *The thief and the dogs*. Cairo: American University in Cairo, 1961.

Secondary sources

Abdel-Hafiz, Ahmed-Sokarno. 1991. "The development of agreement markers in Standard Arabic". *Bulletin of the Faculty of Arts* (Qena) 1:6.175–188.

—— 2003–2004. "Lexical cohesion in the translated novels of Naguib Mahfouz: The evidence from *The Thief and the Dogs*". *Occasional Papers in the Development of English Language Education* (Ain Shams University) 37.63–88. [Also available at: www.arbicwata.org/Arabic/The_WATA_Library/Research_Papers_and_Studies/Excerpts_from_Papers/English/index.html.]

Aziz, Y. 1998. "Translation and pragmatic meaning". *Issues in translation*, ed. A. Shunnaq, A.C. Dollerup, and S. Mohammad, 119–141. Irbid: Irbid National University and Jordanian Translators' Association.

Baker, Mona. 1992. *In other words: A coursebook in translation*. London: Routledge.

Berman, Ruth. 1978. *Modern Hebrew syntax*. Tel Aviv: University of Tel Aviv.

Blum-Kulka, Shoshana. 2000. "Shifts of cohesion and coherence in translation". *The translation studies reader*, ed. Lawrence Venuti, 298–313. London: Routledge.

Crystal, David. 2003. *A dictionary of linguistics and phonetics*. 5th ed. Oxford: Blackwell.

Halliday, Michael A.K. and Ruqaiya Hasan. 1976. *Cohesion in English*. London: Longman.

Holes, Clive. 1995. *Modern Arabic: Structures, functions and varieties*. London: Longman.

Mohammad, Mohammad A. 1999. *Word order, agreement, and pronominalization in Standard and Palestinian Arabic*. Amsterdam and Philadelphia: J. Benjamins.

Obeidat, Hussein A. 1998. *"Stylistic aspects in Arabic and English translated literary texts: A contrastive study". Meta 13:3.1–11. [Also available at:* www, erudit.org/revue/meta/1998/n3/003753ar.htm/.*]*
Versteegh, Kees. 1997. *The Arabic language*. Edinburgh: Edinburgh University Press.

Ahmed-Sokarno Abdel-Hafiz
(South Valley University)

Proper Names

Arabic proper or personal names (*ism*, pl. *ʾasmāʾ*, *ʿalam*; or *ism ʿalam*, pl. *ʾasmāʾ ʾaʿlām*) (Wright 1896:I, 107B), known from many sources and particularly abundant, are given for purposes of identification and for social and political interaction (Wild 1982:154). According to the rules of Arabic nomenclature, the full Arabic personal name is usually composed of the following elements: (i) the proper or personal name (*ism* or *ʿalam*); (ii) the lineage (*nasab*); (iii) the paternal or maternal name or agnomen (*kunya*); (iv) the relative name (→ *nisba*); and (v) the nickname (*laqab*) or a pejorative sobriquet (*nabaz*). Sometimes, the *nom de plume* (*taxalluṣ*) adopted by authors, poets, and artists in the Persian, Turkish, and Indo-Muslim worlds is added to this (Editors 1978:181; Bosworth 1986:619).

Because of the great variety of forms of names, many of which are quite difficult to read and to understand, the study of Arabic names already posed problems for Arab authors of premodern times. Examples are Ibn Jinnī (d. 392/1002), who in his *Kitāb al-mubhij fī tafsīr ʾasmāʾ šuʿarāʾ al-Ḥamāsa* explains the names of the poets quoted in ʾAbū Tamām's *al-Ḥamāsa* with regard to their form and etymology, and similarly Ibn Durayd (d. 321/933), who explains the names of tribes in his monographic handbook of genealogical etymology, the *Kitāb al-ištiqāq* (Senfft 1942:V). In classical times, numerous authors dealt further with names, though mainly in connection with the correct pronunciation of transmitters of Prophetic traditions, e.g. Ibn Xaṭīb ad-Dahša (d. 834/1431) in his alphabetically arranged *Tuḥfa* on the transmitters in al-Buxārī's and Muslim's *Ḥadīṯ*, as well as in Mālik ibn ʾAnas' *al-Muwaṭṭaʾ* (see Mann 1904). Extensive works on *kunya*s and *nisba*s and other parts of Arabic names were compiled. Ibn al-ʾAṯīr's (d. 630/1233) dictionary of *kunya*s, called *al-Muraṣṣaʿ*, gained wide currency, and as-Samʿānī's (d. 562/1166) *al-ʾAnsāb*, containing more than five thousand alphabetically arranged names, is the most elaborate dictionary. Like Ibn Durayd, aḏ-Ḏahabī (d. 748/1348) compiled his *Muštabih* in the 14th century C.E. as a pure name book in order to explain and to rectify rare or false transmitted names.

1. Research on personal names

Research in this area is sometimes of a more popular nature, e.g. Chahine (1987); usually the question of the 'meaning' of personal names is at the center of such publications. But a geographical and sociological study of personal names, as well as a statistical and historical survey of personal names based on the source materials, is essential for historical research on personal names as a social phenomenon (Lipiński 1997:570). An encyclopedia of personal names, sufficient for the needs of both philologists and historians, is still a desideratum for scholars working in this field. Most scholarship on Semitic und Arabic personal names (the relevant European literature is listed in Sezgin's bibliography (1991:III, 269–279, XI, 135) has concentrated on their etymology, categories, manners of origin, and method of development. There is a wealth of literature on different aspects of the topic. As for Arabic onomastics, a number of important studies have been published in this field since the end of the 19th century, e.g. Justi's still valid list of Iranian (and Arabic) names, published in 1895. Caetani initiated the project *Onomasticon Arabicum*, which was never completed. Only one volume, on the persons called *ʿAbdallāh*, with a total of 14,810 names, appeared in 1915. Fortunately, the introductory volume, edited by Caetani's co-author Gabrieli, offers an introduction to onomastics and contains a summary of the European and Arab literature up to 1915, as well as a most useful survey of types of Arabic names and their constituent parts with numerous examples. Before this, Nöldeke was one of the first European scholars to compare Hebrew and Arabic names in a single essay (Nöldeke 1861:806–810; see also Rosenhouse's study of modern Hebrew and Arabic names in Israel, 2002:97–114). Later, Nöldeke also published a short note, "Some groups of Semitic personal names", which also deals with a number of names of animals and, much rarer, of plants, mainly Arabic ones (1904:73–90). Other relevant studies include Bräu (1925) on selected Old North-Arabic ritual personal names (some of the hypotheses being rather far-fetched); and Senfft's dissertation, which deals with more than three thousand transmitted names of the Prophet's Companions (1942:VII, 1–150, with a long list containing 3,129 items). Similarly useful is the comprehensive survey of Arabic names by Walther (1966). Female names, which are of course in a patriarchal society much less numerous than male ones, are the topic of the Munich thesis by Gratzl (1906). He was the first to present (without, however, distinguishing between prose and poetry) a systematic list of six hundred female names (Gratzl 1906:63–84), which was supplemented with an additional hundred names by Geyer (1907:48–50). Names in the early classical period, until the 13th century, were the topic of Ringel's Erlanger dissertation, published in 1938. He listed about two hundred fifty female names in Arabic love poetry (Ringel 1938:72–121), among them seventy so-called 'Decknamen' or pseudonyms (1938:25–31, 121–126). Several papers deal with the names of Bedouin, some of which are animal and plant names in the plural, originally intended to denote groups. Worthy of mention are Hess with his academy paper on "Names of Bedouins in Central Arabia" (1912), and Littmann, who published a study of Bedouin and Druze names in the Syrian area of Ḥawrān (1921:6–20). Ten years later, Littmann offered a small, alphabetically arranged list containing about five hundred names from modern Egypt (1956:82–91). At an earlier date, Socin (1899) presented similar lists of Algerian names, aranged according to morphological pattern. The Biblical names, which are at the same time names that often appear in the *Qurʾān*, were listed separately by Horovitz in 1926. The relevant articles in several encyclopedias, reference books, and manuals of onomastics, e.g. Hasting's *Encyclopaedia of religion and ethics* (Margoliouth), the *Encyclopaedia of Islam* (Bosworth, Wensinck), the *Grundriß der arabischen Philologie* (Wild and Weipert), and the *International handbook of onomastics* (Fischer), allow a comprehensive survey of the traditional forms. For one of these name elements, the *kunya*, and the Islamic titulature, Bāšā gave a very useful historical survey with particular reference to Egypt, mainly from al-Qalqašandī's *Ṣubḥ al-ʾaʿšā*; in the second part, he gives an analytical list of honorific titles and forms of address. The wide ranging survey of Sāmarrāʾī (1964) is also very useful; it tackles fundamental issues and provides a detailed account on the various names from Iraq, Africa (Kenya, Niger, Maghreb), Central Asia, Iran, Indonesia, and the Balkan Peninsula. In a broader perspective, the *Dictionary of Arab names* (*Muʿjam ʾasmāʾ al-ʿArab*; az-Zubayr a.o. 1411–1412/1991) is a good tool for the researcher. Based upon a sociolinguistic field study of names used in many parts of the Arab world from Bahrain to Yemen, the dictionary contains more than eighteen thousand personal names which are most frequently used in the Arab world. Classified as male-only, female-only, and names that are both male and female, the entries are accompanied by authentication of their linguistic and etymological derivations, their historical development, and their social significance. *A treasury of Arab names* (*Sijill ʾasmāʾ al-ʿArab*) and two volumes on *Personalities in the history of Oman* (*Dalīl ʾaʿlām ʿUmān*) and a *Research guide to the study of Arab names* (*Manhaj al-baḥṯ fī ʾasmāʾ al-ʿArab*) complete this useful reference tool. At the same time, in 1991 Sublet published a serious study entitled *Le voile du nom*, which was also translated into Arabic under the title of *Ḥiṣn al-ism* (1999). Sublet, who worked for many years as the director of the project *Onomasticon Arabicum* and as editor of the *Cahiers d'onomastique arabe* at the CNRS in Paris (five volumes published in 1979, 1982, 1985, 1989, 1993), deals in particular with the typology and the function of Arabic names in the social context. By analyzing selected examples, she explains the use of the *kunya* as a polite formula, as well as the peculiar shaping of female names and the development of the titles of officials. A different approach is represented by Annemarie Schimmel's study *Islamic names* (1989), also published in German (*Von Ali bis Zahra*, 1993), which is a wide-ranging guide to common and less common Muslim names in many Arabic countries. The survey also includes names in Turkish, Persian, and Urdu.

2. Personal names

2.1 *Ism*

The proper or personal name, called *ism, ʿalam, ism al-ʿalam* (see, e.g., Wright 1896:I, 107B–D; Reckendorf 1921, par. 108; Wild 1982:154), is the main name, given to infants shortly after birth, usually by one parent or both. The naming may be the expression of a wish, it may relate to family events or to various circumstances of birth, sometimes of pregnancy, or it may aim at securing a favorable future for the bearer (*nomina boni augurii*). Names may also serve to avoid harm, which explains the use of antiphrastic, unpleasant, and apotropaic names (Wild 1982:155; Bosworth 1986:619; Fischer 1995:874; Lipiński 1997:568). Based on everyday vocabulary, the personal name, belonging mainly to the categories of nouns, adjectives, and verbs, can be of several types (Editors 1978:179): (i) ancient Arab names, mostly of pre-Islamic origin, partly altered and substituted (Margoliouth 1917:140; Stowasser-Freyer 1966:26ff.); (ii) Biblical names adapted to Arabic in the *Qurʾān* and enjoying great popularity, the most favored being *ʾIbrāhīm* (Abraham; see especially Luxenberg 2007:93), and the other Biblical figures revered as prophets in their Qurʾānic forms, such as *Mūsā* (Moses; see Luxenberg 2007:41), *ʾIsḥāq* (Isaac), *ʾIsmāʿīl* (Ishmael; see Dagorn 1981), *Yūsuf* (Joseph), *Dāwūd* (David), *ʿĪsā* (Jesus; see Fraenkel 1890:337ff.; Luxenberg 2007:41–43), *Maryam* (Mary; see Margoliouth 1917: 139; Horovitz 1926:78–155; Wild 1982:156; Fischer 1995:874; Endreß 2002:143); (iii) many genitive compound names, either using the proper and attributive names of God (*al-ʾasmāʾ al-ḥusnā*), such as the theophoric and very frequent name *ʿAbdallāh* 'Servant of God' (see, e.g., Fischer 1995:874; Endreß 2002:143; Lipiński 1997:568), or preceded by a construct substantive such as *Hibat Allāh* 'Gift of God'; (iv) foreign names, from Persian, Turkish, Coptic, Berber, and other languages (Editors 1978:179).

Most of the tribal names in the literature (e.g. Caskel 1966) are in the form of the names of individuals. Feminine in gender (Wright 1896:II, 292A; Weninger 2002:217ff.), these tribal names include a large number of animal, plant, and apotropaic names (Hess 1912:4; Wild 1982:156, 161). The notion held at one time by Robertson Smith (1885:186ff.) that these names are totem names was rejected by Nöldeke (1886:156–166; 1904:74, 93).

2.2 *Nasab*

Next to the personal name is the patronymic (*nasab*, pl. *ʾansāb*; Wild 1982:158), containing the name of the father, ending with the alleged ancestor of the tribe or clan. The lineage or pedigree is in form of a genealogical chain of the *ibn* – X pattern, each link being introduced by *ibn* 'son of' X, of Y, etc. The pattern *bint* – X is also used for females, and likewise, the pattern with *banū* – X, *bū* – X, or *ʾāl* – X is used for Arab tribal groups or clans and ruling houses of the Arab Bedouin (Hess 1912:8; Endreß 2002:144).

2.3 *Kunya*

The agnomen or paternal or maternal name is called *kunya*, referring to the person's eldest son or daughter (Wright 1896:I, 107D; Wild 1982:159). It is usually a compound form containing as the first element the noun *ʾabū* 'father of' or, in the case of a woman, *ʾumm* 'mother of', plus a name, in principle of the firstborn son or daughter, or sometimes a younger son or daughter. The *kunya* put in front of the *ism* may indicate a real relationship, but it can also be used purely metaphorically or allude to some desired quality, as, for example, *ʾAbū l-Faḍl* 'Father of merit' (Wensinck 1927:396; Wild 1982:169; Fischer 1995:874). A special case is the Prophet's *kunya*, *ʾAbū l-Qāsim*. According to a well-known *ḥadīṯ* (see Wensinck 1927:396), the name of the Prophet's son, *al-Qāsim*, should not be taken as a *kunya* by persons bearing the name *Muḥammad* or one of its synonyms. Yet, this rule was frequently violated, as demonstrated by Goldziher (1897:149/261; see also Margoliouth 1917:139).

As calling persons by their real name was usually considered to be impolite (→ politeness; → terms of address), the *kunya* as a mode of respectful address became very early on an element charged with expressivenes among friends in special circumstances, outside the sphere of private life. It was used in polite and direct speech and in respectful indirect reference. This pattern was also used sparingly for females (see Stowasser-Freyer 1966:31–38; Spitaler 1968:339–343/276–280; Endreß 2002:143).

2.4 *Nisba*

The relative name, → *nisba*, pl. *nisab* (see Wright 1896:I, 109D, 149D–165B; Wild 1982:160), is an adjective with the suffix *-ī*, always preceded by the definite article. Originally formed from the name of the tribe or clan or family of the individual, it may also denote the origin, place of birth or residence, sometimes the affiliation to a school of law or a political or religious group, and occasionally a profession or trade; see the examples given by Fischer (1995:875) and Endreß (2002:144). Often, the relative name is used as a *nom de plume* (*taxalluṣ*) by authors and artists in the Persian, Turkish, and Indo-Muslim worlds, for instance by the Persian poet *Firdawsī* 'the Paradisiacal' (Endreß 2002:145).

2.5 *Laqab*

Finally, the last and most colorful element in nomenclature is the sobriquet or nickname (*laqab*, pl. *ʾalqāb*). Originally intended as a description for a person, this name usually, though not exclusively, relates to some personal trait or characteristic or physical defect, e.g. *al-Jāḥiẓ* 'the Goggle-eyed' (Wild 1982:159; Fischer 1995:874), or it represents a honorific epithet, often a title. Poets and scholars often received such nicknames, which were usually placed after their *nisba* and often referred to striking lines from their poems or to their profession (see examples in Margoliouth 1917:137; Bosworth 1986:618; Endreß 2002:145). Honorific titles of famous men and rulers were often applied retroactively by later Arab historians, based on common practice in their own time, and they even attributed personal appellations and titles to the earliest periods (Margoliouth 1917:140; Bosworth 1986:620). These titles are found with greater regularity after the rise of the Abbasid caliphs. From this period, especially from the 4th/10th century onward, when the power of the caliphs receded and the empire became fragmented, the bestowing of honorific titles on high officers of state, semi-independent princelings, governors, and military leaders, as well as on scholars, mystics, and poets, became a political tool of the caliphs (Wild 1982:159ff.; Fischer 1995:874; Endreß 2002:146–150). The titles were usually compounds with the second element *ad-dawla* 'dynasty', from the 4th/10th century onward, especially in the eastern parts, with the second element *ad-dīn* 'religion'. A total of 94 compound titles were analyzed and listed by Kramers (1927:67), to which Dietrich (1961:45–53) added a further 92 titles. Recently, Sublet (1991:92ff.) has brought up additional titles, mainly drawn from Ibn Xallikān's biographical dictionary. Along with these genitive combinations, which were modeled on Iranian ceremonial and titulary practices (Kramers 1927:60ff.; Endreß 2002:148), similar compounds could also be formed with *al-mulk* 'kingdom' or *al-ʾislām*. Especially favored was the type of title containing a dual expression, e.g. *Ḏū r-Riyāsatayn* or the like, to which Goldziher (1899:326–329) has devoted a monograph (see also Caetani and Gabrieli 1915:I, 167–169, par. 154; Bosworth 1986:621; Sublet 1991:184–187). The most famous holders of such titles, for instance *Nāṣir ad-Dawla* 'Defender of the dynasty', belonged to the Daylamī dynasty of the Shiʿite Buyids and the Ḥamdanid and Ghaznawid families (see Spuler 1952:356–360; Busse 1969:159–184; Bosworth 1962:210–233).

3. Morphological classification of personal names

The classification of the forms of personal names given by the Arab grammarians is exhaustive (see, e.g., Howell 1883:5–12, chap. 3; Wright 1896:I, 107B–108D; Walther 1966:13 ff.). In their love for schematization, they divided the names into (i) nouns, nouns with suffixes, adjectives, derived nouns; (ii) compound elements, e.g. two words; and (iii) verbs in the perfect and imperfect.

Apparently, most Arabic personal names consist of a single semantic element, in general nouns (in the singular, as in *ʾAsad* 'Lion', or in the plural, as in *Kilāb* 'Dogs' or *ʾAḥlām* 'Dreams'). In many cases, these refer to natural and human, usually positive, characteristics or virtues. They also refer to animal, plant, and flower names. Such flower and jewelry names, regarded as euphemistic, were used properly for slaves, servants, and other persons of little importance, e.g. *Yāqūt* ‘Hyacinth’, ‘Ruby’ (see Wild 1982:156, 161; Bosworth 1986:619; Fischer 1995:874). Some names are derived from adjectives or participles, such as *Muḥammad* ‘Praised’, by far the most common and popular name, together with other variants of this root, e.g. *ʾAḥmad*, *Ḥāmid*, *Maḥmūd*.

Several naming patterns are two-word compounds. The compound may be made up either of a sentence, e.g. the name *Taʾabbaṭa šarran* ‘He carried mischief [i.e. his sword] under his armpit’ (Wright 1896:I, 108B-C; Margoliouth 1917:138; Wild 1982:155, 158), or of a word followed by another in the genitive case, e.g. *ʿAbdallāh* ‘Servant of God’. Among these theophoric names, restricted to male individuals, one finds typically God’s names or attributes and names of family members, mainly of the *ʿAbd*-X pattern.

In comparison to Biblical names, in which the deity is associated with a verbal predicate, verbal-sentence names are much rarer in the Arabic tradition. Imperfect names (Nöldeke 1861:807ff.; Reckendorf 1898, par. 90; Wild 1982:158; Lipiński 1997:570), built on the form *y/tfʿl*, sometimes with *ti*-prefix (*Tiḥyā*), and deviating from the morphological rules in Classical Arabic (e.g. *Yarḥum*), are treated by Weninger (2002), who collected some two hundred fifty imperfect names, which may express a wish, e.g. *Yaḥyā* ‘May he live’, or *Yazīd* ‘May he grow’.

The Arabic morphological system basically dictates gender distinction for male and female personal names in the morphological inflection and conjugation. Gender indication is made by adding the feminine suffix *-a(h)* to a masculine form. However, several names with feminine ending indicate male individuals, e.g. *Maslama*, *ʾUsāma*, and conversely, some female names do not have a feminine marker (Reckendorf 1898, par. 13). Furthermore, some names are used for both sexes. To know a priori whether names are diptotes or triptotes (e.g. *ʿUmar*, which is a diptote, and *ʿAmr*, which is a triptote) is not always easy (→ diptosis). As personal names are definite by nature, the tendency has always been to treat them as diptotes (see the lengthy treatment by Fleisch 1961:271–273; Wright 1896:I, 245, 247; Reckendorf 1898, par. 90; Wild 1982:155; Fischer 2006, par. 153d-f).

4. Modern trends of personal names

Partially following the custom among the upper classes in premodern times of adopting the name of a reputed ancestor or famous family member as a sort of family name (see Walther 1966:6ff.; Sellheim 1984–1986:377–379; Wild 1982:160) and especially accelerated by introduction of compulsory registration for the adoption of a Westernized way of naming in several Arab countries during the two last centuries (e.g. Algeria in 1882, Tunisia in 1925, Turkey in 1934, Iran in 1932, and Egypt in 1970; see Editors 1978:181; Schimmel 1989:80, 1993:176; Fischer 1995:873), many Arab countries have adapted modern Western naming conventions. Constructing a model of a first name and an additional name, instead of a full chain of names, the first name, which functions as a sort of personal name, and the second name, which is usually the father’s name, are simply juxtaposed (in Egypt, the official name, *ism ṯulāṯī*, adds the name of the grandfather). In most Arab countries, the word *ibn* or *bint* is omitted nowadays, as for instance in the name of the Egyptian writer *Ṭāhā Ḥusayn*, except in Arabia and the Maghreb, where it is replaced by the word *ʾabū* (Diem 1974:45, 1992:221; Wild 1982:158, 160ff.; Schimmel 1989:80, 1993:176; Fischer 1995:873, 875; Endreß 2002:144). Due to the impact of modern times and nationalistic trends, new Arab names have appeared, such as *Filasṭīn* (Palestine; see Rosenhouse 2002:106), as well as names like *Fuʾād*, *Ramaḍān*, and *Rajab* (Fischer 1995:874). The trend is to use religious names, neutral names drawn from the secular vocabulary, and European names side by side.

Bibliographical references

Primary sources

Ḏahabī, *Muštabih* = Šams ad-Dīn Muḥammad ibn ʾAḥmad aḏ-Ḏahabī, *Kitāb al-muštabih fī ʾasmāʾ ar-rijāl*. Ed. Paul de Jong. Leiden: E.J. Brill, 1881/ Ed. ʿAlī Muḥammad al-Bijāwī. Cairo: ʿĪsā al-Bābī al-Ḥalabī, 1962.

Ibn al-ʾAṯīr, *Lubāb* = ʿIzz ad-Dīn ʾAbū l-Hasan ʿAlī ibn Muḥammad Ibn al-ʾAṯīr, *al-Lubāb fī tahḏīb al-ʾansāb*. 3 vols. Cairo: Maktabat al-Quds, 1938–1950/1357–1369.

Ibn al-ʾAṯīr, *Muraṣṣaʿ* = Majd ad-Dīn ʾAbū s-Saʿādāt al-Mubārak Ibn al-ʾAṯīr, *Kitāb al-Muraṣṣaʿ fī l-*

ʾābāʾ wa-l-ʾummahāt wa-l-banīna wa-l-banāt wa-l-ʾaḏwāʾ wa-l-ʾaḏwāt. Ed. Christian Friedrich Seybold. Weimar: Felber 1896/Ed. ʾIbrāhīm as-Sāmarrāʾī. Baghdad: Riʾāsat Dīwān al-ʾAwqāf, 1391/1971.

Ibn Durayd, *Ištiqaq* = ʾAbū Bakr Muḥammad ibn al-Ḥasan Ibn Durayd al-ʾAzdī, *Kitāb al-ištiqāq*. Ed. Ferdinand Wüstenfeld. Göttingen: Dieterich, 1854/Ed. ʿAbd as-Salām Muḥammad Hārūn. Cairo: al-Xānjī, 1378/1958.

Ibn Ḥajar, *Tabṣīr* = ʾAḥmad ibn ʿAlī Ibn Ḥajar al-ʿAsqalānī, *Tabṣīr al-muntabih bi-taḥrīr al-muštabih*. 4 vols. Ed. ʿAlī Muḥammad al-Bijāwī and Muḥammad ibn ʿAlī an-Najjār. Cairo: ad-Dār al-Miṣriyya bi-t-Taʾlīf wa-t-Tarjama, 1383–1386/1964–1967.

Ibn Jinnī, *Mubhij* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Mubhij fī tafsīr ʾasmāʾ šuʿarāʾ al-Ḥamāsa*. Ed. Ḥasan Hindāwī. Damascus: Dār al-Qalam, Beirut: Dār al-Manāra, 1407/1987.

Ibn Xaṭīb, *Tuḥfa* = Ibn Xaṭīb ad-Dahša, *Tuḥfat ḏawī l-ʾirab fī muškil al-ʾasmāʾ wa-n-nisab*. Ed. Traugott Mann, *Tuḥfa dhawī-l-Arab* [sic!], *Über Namen und Nisben bei Boẖārī, Muslim, Mālik von Ibn H̱aṭīb al-Dahša*. Leiden: E.J. Brill, 1905.

Muqbil, *Durr* = Muḥammad Muqbil Bīk (Beg), *Kitāb ad-Durr aṯ-ṯamīn fī ʾasmāʾ al-banāt wa-l-banīn*. Cairo: Maṭbaʿat Wādī n-Nīl al-Miṣriyya, 1294 A.H.

Samʿānī, *ʾAnsāb* = ʾAbū Saʿd ʿAbd al-Karīm ibn Muḥammad as-Samʿānī, *al-ʾAnsāb*. 13 vols. Hyderabad: Dāʾirat al-Maʿārif al-ʿUṯmāniyya, 1382–1402/1962–1982.

Secondary sources

Ayalon, David. 1975. "Names, titles and *nisba*s of the Mamlūks". *Israel Oriental Studies* 5.189–232.

Barbier de Meynard, Casimir Adrien. 1907–1908. "Surnoms et sobriquets dans la littérature arabe". *Journal Asiatique*, 10th series, 9.173–244, 365–428; 10.55–118, 193–273.

Bāšā, Ḥasan al-. 1957. *al-ʾAlqāb al-ʾislāmiyya fī t-taʾrīx wa-l-waṯāʾiq wa-l-ʾāṯār*. Cairo: Dār an-Nahḍa al-ʿArabiyya.

Bosworth, Clifford Edmund. 1962. "The titulature of the early Ghaznavids". *Oriens* 15.210–233.

——. 1986. "*Laḳab*". *Encyclopaedia of Islam* V, 618–631. 2nd ed. Leiden: E.J. Brill.

Bräu, Hans Hermann. 1925. "Die altnordarabischen kultischen Personennamen". *Wiener Zeitschrift für die Kunde des Morgenlandes* 32.31–59, 85–115.

Busse, Heribert, 1969. *Chalif und Grosskönig: Die Buyiden im Iraq (945–1055)*. Beirut and Wiesbaden: F. Steiner.

Caetani, Leone and Giuseppe Gabrieli. 1915. *Onomasticon Arabicum ossia Repertorio alfabetico dei nomi di persona e di luogo contenuti nelle principali opere storiche, biografiche e geografiche, stampate e manoscritte, relative all'Islām*. I. *Fonti – Introduzione*. IIA. *Aʿābil-ʿAbdallah*. Rome: Casa Editrice Italiana.

Caskel, Werner. 1966. *Ǧamharat an-nasab: Das genealogische Werk des Hišām Ibn Muḥammad al-Kalbī*. 2 vols. Leiden: E.J. Brill.

Chahine, Jerome. 1987. *al-ʾAsmāʾ al-ʿarabiyya: Le mille et un noms arabes*. Paris: Editions Asfar.

Dagorn, René. 1981. *La geste d'Ismaël d'après l'onomastique et la tradition arabe*. Geneva: Droz.

Déclais, Jean-Louis. 2003. "Names of the Prophet". *Encyclopaedia of the Qurʾān*, ed. Jane Dammen McAuliffe, III, 501–505. Leiden: E.J. Brill.

Diem, Werner. 1974. "Personennamen". *Lexikon der islamischen Welt*, ed. Klaus Kreiser, Werner Diem, and Hans-Georg Majer, III, 43–45. Stuttgart: Kohlhammer. (Repr., *Lexikon der islamischen Welt*. Rev. ed. Klaus Kreiser and Rotraud Wielandt, 220–222. Stuttgart: Kohlhammer, 1992.)

Dietrich, Albert. 1961. "Zu den mit *ad-dīn* zusammengesetzten islamischen Personennamen". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 110.43–54.

Editors. 1978. "*Ism*". *Encyclopaedia of Islam*, IV, 179–181. 2nd ed. Leiden: E.J. Brill.

Endreß, Gerhard. 1997. *Der Islam: Eine Einführung in seine Geschichte*. 3rd ed. Munich: C.H. Beck.

——. 2002. *Islam: A historical introduction*. Trans. Carole Hillenbrand. 2nd ed. Edinburgh: Edinburgh University Press.

Fischer, Wolfdietrich. 1995. "Arabische Personennamen". *Namenforschung: Ein internationales Handbuch zur Onomastik*, ed. Ernst Eichler, Gerold Hilty, Heinrich Löffler, Hugo Steger, and Ladislav Zgusta, I, 873–875. Berlin and New York: de Gruyter.

——. 2006. *Grammatik des klassischen Arabisch*. 4th rev. ed. Wiesbaden: O. Harrassowitz.

Fleisch, Henri. 1961. *Traité de philologie arabe*. I. *Préliminaires, phonétique, morphologie nominale*. Beirut: Imprimerie Catholique.

Fraenkel, Siegmund. 1890. "Miscellen. A. Eigennamen". *Wiener Zeitschrift für die Kunde des Morgenlandes* 4.332–340.

Garcin de Tassy, Joseph Héliodore. 1854. "Mémoire sur les noms propres et les titres musulmans". *Journal Asiatique*, 5th series, 3.422–510.

Geyer, Rudolf. 1907. Review of Gratzl (1906). *Wiener Zeitschrift für die Kunde des Morgenlandes* 21.44–50.

Goldziher, Ignaz. 1897. "Gesetzliche Bestimmungen über Kunja-Namen im Islam". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 51.256–266. (Repr., Ignaz Goldziher, *Gesammelte Schriften*, ed. Joseph Desomogyi, IV, 143–154. Hildesheim: G. Olms, 1967–1973.)

——. 1899. "Ueber Dualtitel". *Wiener Zeitschrift für die Kunde des Morgenlandes* 13.321–329. (Repr., Ignaz Goldziher, *Gesammelte Schriften*, ed. Joseph Desomogyi, IV, 195–203. Hildesheim: G. Olms.)

——. 1928. "Verheimlichung des Namens". *Der Islam* 17.1–3.

Gratzl, Emil. 1906. *Die altarabischen Frauennamen*. Leipzig: W. Drugulin.

Hess, Jean-Jacques. 1912. *Beduinennamen aus Zentralarabien*. Heidelberg: C. Winter.

Horovitz, Josef. 1926. "Die koranischen Eigennamen". Josef Horovitz, *Koranische Untersuchungen*, 78–155, 156–165. Berlin and Leipzig: de Gruyter.

Howell, Mortimer Sloper. 1883. *A grammar of the Classical Arabic language translated and compiled from the most approved native or naturalized authorities*. The introduction: and Part I. *The noun*. Allahabad: North-Western Provinces and Oudh

Government Press. (Repr., Delhi: Gian Publishing House, 1986.)

Justi, Ferdinand. 1895. *Iranisches Namenbuch*. Marburg: Elwert. (Repr., Hildesheim: G. Olms, 1963.)

Kaḥḥāla, ʿUmar Riḍā. 1368/1949. *Muʿjam qabāʾil al-ʿArab al-qadīma wa-l-ḥadīṯa*. 3 vols. Damascus: al-Maṭbaʿa al-Hāšimiyya.

——. 1377–1379/1958–1959. *ʾAʿlām an-nisāʾ fī ʿālamay al-ʿArab wa-l-ʾIslām*. 4 vols. Damascus: al-Maṭbaʿa al-Hāšimiyya.

Khemiri, Tahir. 1942. “Die Formen *faʿūl* und *faʿʿūl*: Ein Beitrag zur arabischen Namenskunde”. *Der Islam* 26.159–160.

Kramers, Johannes Hendrik. 1927. “Les noms musulmans composés avec *Dīn*”. *Acta Orientalia* 5.53–67.

Lipiński, Edward. 1997. *Semitic languages: Outline of a comparative grammar*. Leuven: Peeters.

Littmann, Enno. 1921. “Beduinen- und Drusen-Namen aus dem Ḥaurân-Gebiet”. *Nachrichten der K. Gesellschaft der Wissenschaften zu Göttingen*, Philologisch-historische Klasse, 1–20.

——. 1956. “Eigennamen der heutigen Ägypter”. *Studi orientalistici in onore di Giorgio Levi della Vida*, II, 81–93. Rome: Istituto per l’Oriente.

Luxenberg, Christoph. 2007. *The Syro-Aramaic reading of the* Koran*: A contribution to the decoding of the language of the* Koran. Berlin: Schiler.

Mann, Traugott. 1904. *Beiträge zur Kenntnis arabischer Eigennamen*, I. Leiden: E.J. Brill.

Margoliouth, David Samuel. 1917. “Names (Arabic)”. *Hasting’s encyclopaedia of religion and ethics*, IX, 136–140. Edinburgh: T. and T. Clark.

Nöldeke, Theodor. 1861. “Bemerkungen über hebräische und arabische Eigennamen”. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 15.806–810.

——. 1886. Review of Robertson Smith (1885). *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 40.148–187.

——. 1892. “Kleinigkeiten zur semitischen Onomatologie”. *Wiener Zeitschrift für die Kunde des Morgenlandes* 6.307–316.

——. 1904. “Einige Gruppen semitischer Personennamen”. Theodor Nöldeke, *Beiträge zur semitischen Sprachwissenschaft*, 73–106, 138f. Strasbourg: Trübner.

Reckendorf, Hermann. 1898. *Die syntaktischen Verhältnisse des Arabischen*. Leiden: E.J. Brill. (Repr., 1967.)

——. 1921. *Arabische Syntax*. Heidelberg: C. Winter. (Repr., 1977.)

Ringel, Heinrich. 1938. *Die Frauennamen in der arabisch-islamischen Liebesdichtung: Ein Beitrag zum Problem der ost-westlichen Literaturübertragung*. Leipzig: G. Kreysing.

Robertson Smith, William. 1885. *Kinship and marriage in early Arabia*. Cambridge: Cambridge University Press. (Repr. and new ed. with additional notes by the author and by Ignaz Goldziher, Budapest. Ed. Stanley A. Cook. Oosterhout: Anthropological Publications. 1966.)

Rosenhouse, Judith. 2002. “Personal names in Hebrew and Arabic: Modern trends compared to the past”. *Journal of Semitic Studies* 47.97–114.

Sāmarrāʾī, ʾIbrāhīm as-. 1964. *al-ʾAʿlām al-ʿarabiyya: Dirāsa luġawiyya ijtimāʿiyya*. Baghdad: al-Maktaba al-ʾAhliyya.

Sauvaget, Jean. 1950. “Noms et surnoms de Mamlouks”. *Journal Asiatique* 238.31–58.

Schimmel, Annemarie. 1989. *Islamic (personal) names*. Edinburgh: Edinburgh University Press.

——. 1993. *Von Ali bis Zahra: Namen und Namengebung in der islamischen Welt*. Munich: Diederichs.

Sellheim, Rudolf. 1984–1986. “’Familiennamen’ im islamischen Mittelalter: Eine Skizze”. *Orientalia Suecana* 33–35.375–384.

Senfft, Joachim. 1942. *Beiträge zur frühislamischen Personennamenkunde*. Berlin: Universität.

Sezgin, Fuat (ed.). 1991. *Bibliographie der deutschsprachigen Arabistik und Islamkunde von den Anfängen bis 1986 nebst Literatur über die arabischen Länder der Gegenwart*. III. *Arabische Sprache, Philologie*. XI. *Nachträge, Inhaltsübersicht, Quellen*. Frankfurt a. Main: Institut für Geschichte der Arabisch-Islamischen Wissenschaften an der Johann Wolfgang Goethe-Universität. [In particular III, 269–279, XI, 135.]

Socin, Albert. 1899. “Die arabischen Eigennamen in Algier”. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 53.471–500.

Spitaler, Anton. 1968. “Beiträge zur Kenntnis der *Kunya*-Namengebung”. *Festschrift Werner Caskel zum siebzigsten Geburtstag 5. März 1966 gewidmet von Freunden und Schülern*, ed. Erwin Gräf, 336–350. Leiden: E.J. Brill. (Repr., Anton Spitaler, *Philologica: Beiträge zur Arabistik und Semitistik*, ed. Hartmut Bobzin, 273–287, 288–289. Wiesbaden: O. Harrassowitz, 1998.)

Spuler, Bertold. 1952. *Iran in früh-islamischer Zeit: Politik, Kultur, Verwaltung und öffentliches Leben zwischen der arabischen und der seldschukischen Eroberung 633 bis 1055*. Wiesbaden: F. Steiner.

Stowasser-Freyer, Barbara. 1966. “Formen des geselligen Umgangs und Eigentümlichkeiten des Sprachgebrauchs in der frühislamischen städtischen Gesellschaft Arabiens (nach Ibn Saʿd und Buḫārī)”. *Der Islam* 42.26–42.

Sublet, Jacqueline. 1991. *Le voile du nom: Essai sur le nom propre arabe*. Paris: Presses Universitaires de France.

——. 1999. *Ḥiṣn al-ism: Qirāʾāt fī l-ʾasmāʾ al-ʿarabiyya*. Trans. Sulaym Muḥammad Barakāt. Damascus: al-Maʿhad al-Faransī li-d-Dirāsāt al-ʿArabiyya bi-Dimašq.

Ullmann, Manfred. 1993. Review of Sublet (1991). *Die Welt des Orients* 24.191–193.

Walther, Wiebke. 1966. *Untersuchungen zu vor- und frühislamischen arabischen Personennamen*. Halle/Saale.

Weipert, Reinhard. 1992. “Ergänzungen zu den bibliographischen Teilen von Band I und Band II”. *Grundriß der Arabischen Philologie*. III. *Supplement*, ed. Wolfdietrich Fischer, 252. Wiesbaden: L. Reichert.

Weninger, Stefan. 2002. “Arabische Imperfektnamen”. *Altorientalische und semitische Onomastik*, ed. Michael P. Streck and Stefan Weninger, 209–226. Münster: Ugarit Verlag.

Wensinck, Arent Jan. 1927. “*Kunya*”. *Enzyklopaedie* *des Islām*, II, 1200. (Repr. with additions, *Encyclopaedia of Islam*, V, 395–396. 2nd. ed. Leiden: E.J. Brill, 1986.)

Wild, Stefan. 1982. “Arabische Eigennamen”. *Grundriß der Arabischen Philologie*. I. *Sprachwissenschaft*, ed. Wolfdietrich Fischer, 154–161, 164. Wiesbaden: L. Reichert.

Wright, William. 1896. *A grammar of the Arabic language*. 2 vols. 3rd ed. rev. William Robertson Smith and Michael Jan de Goeje. Cambridge: Cambridge University Press.

Ziriklī, Xayr ad-Dīn az-. 1372–1378/1954–1959. *al-ʾAʿlām: Qāmūs tarājim li-ʾašhar ar-rijāl wa-n-nisāʾ min al-ʿArab wa-l-mustaʿribīna wa-l-mustašriqīn*. 10 vols. Cairo: Kosta Tsumas.

Zubayr, Muḥammad az- (with as-Saʿīd Muḥammad Badawī, Fārūq Šūša, ʿAlāʾ ad-Dīn Hilāl, and Maḥmūd Fahmī Ḥijāzī). 1411–1412/1991. *Mawsūʿāt as-Sulṭān Qābūs li-ʾasmāʾ al-ʿArab*. I, 1–2. *Muʿjam ʾasmāʾ al-ʿArab* (*Dictionary of Arab names*). II, 1–4. *Sijill ʾasmāʾ al-ʿArab*. III. *Dalīl ʾaʿlām ʿUmān* (*Personalities in the history of Oman*). IV. *Manhaj al-baḥṯ fī ʾasmāʾ al-ʿArab* (*Research guide to the study of Arab names*). Masqat and Beirut: Jāmiʿat as-Sulṭān Qābūs and Maktabat Lubnān.

Franz-Christoph Muth (University of Munich)

Prosodic Template → Obligatory Contour principle; Morphology

Prosody

Etymologically, the Greek term *prosōidía* means ‘stress, quantity, in pronunciation’. We give it a broad sense which includes syllable, stress, pitch, intonation, rhythm, rate of speech, pause, etc. The American tradition uses the term ‘suprasegmental’ to express the idea that the domain of prosody is larger than a single segment and that prosodic phenomena are revealed by a comparison of items in sequence (Lehiste 1970).

The prosodic component of language plays a fundamental role in → first language acquisition. Several experimental studies have established that newborn children are particularly sensitive to prosody and that they are able to identify the rhythm of their mother tongue among other languages (Nazzi a.o. 1998).

In the domain of prosodic studies, there are two main tendencies. The first one gives priority to descriptive and theoretical investigations, while the second concentrates more on direct observation of speech signals, on measurement of the acoustic correlates of prosody, and on perceptual experiences (Cutler and Ladd 1983).

Prosodic units are two-sided: abstract and concrete. At the abstract level, they are organized into a hierarchy and can be subjected to phonological representation. At the physical level, their study has a close look at the acoustic parameters like fundamental frequency, intensity, and duration. At the functional level, prosody assumes a certain number of functions, which can be reduced to a dual scale (Bruce 1985; Rossi 1985): integration and organization into a hierarchy. The former is an indicator of cohesion between adjoining units in the utterance. The latter function is associated with the linguistic message: it organizes the ultimate constituents according to their informative weight.

The development of prosodic studies in several languages allows the establishment of typological comparisons between them. For that purpose, prosodic factors can help us to draw up similarities and differences among diverse idioms (Vaissière 1983; Hirst and DiCristo 1998; Barkat 2000).

On a narrow scale, prosodic processes can be manifested thanks to the overlapping effect of two adjacent sounds or to coarticulation. Such coarticulatory phenomena can vary across languages. In French, it is the vowel that affects the previous consonant. For example, the phoneme /t/ in the sequences [ti] and [ta] is realized with palatalization and pharyngealization, respectively. To show such an articulatory displacement of the coronal stop consonant, a transcription with Arabic characters of these two syllables requires two written forms: ت and ط. On the other hand, in Arabic, it is the consonant that influences the adjacent vowel. Thus, in an emphatic consonantal environment, vowels like /i, u, a/ become [e, o, ɑ] (→ vowel backing).

Another fact, related to Moroccan Arabic, concerns the supraliminal lengthening of a vowel followed by a final consonant: all things being equal, the duration value of the vocalic nucleus in items like [blæd] or [blæt] is significantly greater than the corresponding one in [blɑ]. This temporal increase seems to run counter to a widely held tendency observed in the world’s languages, namely Closed Syllable Vowel Shortening (CSVS) (Maddieson 1985:206). This contextual lengthening constitutes an important prosodic cue in the perception of Moroccan word boundaries and allows the hearer to disambiguate minimal pairs, as in (1) and (2) (Benkirane 2003):

(1) [ʃɾæ#tɾijɑ] 'He bought a chandelier'
(2) [ʃɾæːt#ɾijɑ] 'She bought lights [or: lungs]'

Unlike Moroccan Arabic, French follows the CSVS tendency. Thus, in examples (3) and (4), a duration increase of the vowel [i] is in favor of the V # C boundary interpretation (Nicaise and Bacri 1985):

(3) [anik#lasɛleʒã] 'Annick bored people'
(4) [aniː#klasɛleʒã] 'Annie sized people up'

In addition, the traditional study of the consonantal and vocalic relations reveals phonotactic information about the probabilities of transition between these two categories of sounds. Moreover, the computation of the proportion of vocalic/consonantal intervals in the utterance allows the capture of syllabic complexity and the deduction of the rhythm class of the language studied (Ramus and Mehler 1999). In particular, the degree of syllable complexity affects the rate of speech. For example, Western Arabic varieties exhibit the greatest degree of → vowel reduction in comparison with Eastern Arabic varieties and, as a consequence, disruption of the syllable's internal structure and a particular consonantal overload. These properties of Western Arabic languages have widely contributed to generate a faster rate of speech, which constitutes a strong cue for distinguishing them from Eastern Arabic languages (Barkat 2000; Hamdi 2002).

At this prospect, prosodic features can play an important role in the identification of the speaker's geographical origin (Benhallam and Dahbi 1990; Benali 2004). They are also a good indicator of the difficulties encountered during second language acquisition, and they can cue foreign accent (Kharrat 1994; Boula de Mareüil and Brahimi 2004). At a superior level of the prosodic organization, the syllable can be stressed or not. In the frame of the word, → stress is a prosodic feature that renders some syllables perceptually more prominent than others (Garde 1968). One or several prosodic parameters can underlie this particular prominence. Thus, one speaks of tonic stress to designate a culmination generated by a considerable increase of the fundamental frequency. But stress can result from an increase in intensity and/or duration, too. The role played by every prosodic parameter in the production and the perception of stress varies from one language to another and requires an organization into a hierarchy. In Russian, it is intensity that positively marks the stressed syllable. In English, it is fundamental frequency that takes precedence over the other two. In French and Moroccan Arabic (Benkirane 2000), stress is mainly associated with a significant increase in duration. In this Western variety of Arabic, in which the phonological opposition of vocalic quantity does not operate, the temporal contrast between short and long vowels is henceforth assumed by prosody (Maas 2002). All things being equal, Moroccan Arabic vowels tend to become much longer when stressed. This temporal contrast is based on the feature [± stress] which generates a particular rhythm.

From a rhythmic point of view, a distinction is made (Pike 1945; Abercrombie 1967; Ladefoged 1975) between syllable-timed languages (such as French, Spanish, Yoruba), stress-timed languages (such as English, Arabic, Russian) and moraïc-timed languages (such as Japanese, Tamil). This categorization is based on isochrony, which is a completely subjective hypothesis: equality of the syllables or morae in one case, return at regular intervals of stress in another (Dauer 1983).

Furthermore, two categories of languages can be distinguished according to the place of stress in the word:

i. When stress is independent of the word's boundary and its place varies because of morphological structure, it is called 'free stress' (German, English, Russian, and others). Thus, in English, the words *photograph, photographer, photographic, photography*, and *photogravure* have different stress patterns: ['fəʊtəgɾɑːf], [fə'tɑgɾəfə], [fəʊtə'gɾæfɪk], [fə'tɑgɾəfɪ], [fəʊtɑgɾə'vjʊə]. The stress shift, which is due to the stress properties of the morphemes, can assume a distinctive role. For example, in English, stress placement can determine the grammatical category of words like *import*: [ˈɪmpɔːt] (= noun) vs. [ɪmˈpɔːt] (= verb).

ii. When stress invariably falls on a determined place in the word, it is called 'fixed stress'. In French, stress falls on the last syllable of the word. In Czech, it is placed at the beginning of the word. This stress is word-boundary dependent and generally assumes a demarcative function. Within this second category, Garde (1968) proposes the almost-fixed stress label to designate languages where the place of stress is determined by the internal phonological structure of the word (Polish, Macedonian, Latin, Classical Arabic, and others). According to Cantineau (1960:119–120), the location of stress in Classical Arabic is governed by the following rules: "L'accent se place sur la première syllabe longue à compter de la fin du mot; si le mot ne comporte pas de syllabe longue, l'accent se place sur la première syllabe du mot; les longues finales ne reçoivent pas l'accent". For example: [ˈqɑːla], [ˈqultu], [almusaːfiˈruːna], [muˈsaːfirun], [ˈkataba], [ˈkatabataː].

These are roughly the same stress rules that prevail in varieties of Arabic spoken in Jordan and in Palestine (van de Vijver 1996). However, in the Jordanian Arabic of Amman, the final syllable can be stressed when it is heavy (De Jong and Zawaydeh 1999): [kilmiˈteːn], [daˈrast], [rɑːˈseːn]. The same is true of colloquial Lebanese Arabic (Chahal 1999).

The classification of Arabic as a stress language is now an established finding, judging by a nonexhaustive list of studies that have been dedicated to it (Birkeland 1954; Janssens 1972; Brame 1973; McCarthy 1980; Welden 1980; Selkirk 1981; Benkirane 1982; Lahlou 1982; Kenstowicz 1983; Zakaria 1984; Hammoumi 1988; Angoujard 1990; Benhallam 1990; Al-Ani 1992; Nejmi 1992; Chahal 1999; De Jong and Zawaydeh 1999).

On the other hand, concerning Standard or Classical Arabic, the problem of the presence of stress has not been completely determined. One reason behind this is that ancient Arab grammarians, as well as the authors of treatises on *Qurʾān* recitation (→ *tajwīd*), turned a deaf ear to the reality of this phenomenon. Another reason is that nowadays there are no ideal native speakers of this language. The situation is inevitable insofar as all speakers of Arabic are, first and foremost, people who speak their native language, whether this is colloquial Moroccan Arabic in Morocco or colloquial Yemeni Arabic in Yemen. In the Arab countries, Standard Arabic is learned at school. Accordingly, it is difficult to ignore the inescapable problems generated by interference in a situation of → diglossia between spoken Arabic and written Standard Arabic. In reality, it is the stress patterns of the dialects that determine the pronunciation of Standard Arabic. Such prosodic interference was not missed in a shrewd remark by Ferguson (1957:474): "The modern pronunciation of classical Arabic reflects stress patterns of the dialect of the speaker". This explains why the production of Standard Arabic by an Egyptian is relatively different from that of an Algerian (Barkat 2000).

In studies that attempt to determine the place of stress in Modern Standard Arabic, it is difficult to establish a consensus. According to Kouloughli (1975), in words of two syllables, stress falls on the penultimate regardless of the length of these syllables. But from a phonetic and perceptive survey of this question (Zahid 1990), it appears that the perception of stress is strongly related to syllabic weight. Nevertheless, in a word composed of two syllables with equal weight, like [ʕiːsa], stress would affect the last one. In the same way, according to Kaye (1997), the utterance [katabahu] is realized with stress on the antepenultimate by a Lebanese or by a Jordanian, but with stress on the penultimate by an Egyptian.

The determination of stress location in a given language constitutes a precondition for the study of → intonation. The domain of intonation is the utterance. To indicate a contrast or a prominence, stress and intonation make use of the same prosodic parameters: fundamental frequency, intensity, and duration. For this reason, the researcher must always make sure not to confound these two levels of analysis. Intonation is a linguistic form that assumes important traditional functions. In this way, the distinction between declarative, interrogative, imperative, and vocative modes is one of the most universal roles of intonation.

Experimental studies on the intonation of Arabic are not abundant. But the following studies provide accounts for some intonation patterns in varieties of Arabic: Abdallah (1960); Loi Corvetto (1982); Lahlou (1982); Haydar and Mrayati (1985); Norlin (1989); Rammuny (1989); Alharbi (1992); Rosenhouse (1995); Chahal (1999); Benkirane (1998, 2000); Kulk (2003); Kulk a.o. (2003); Mawhoub (2000, 2004); Yeou (2004).

BIBLIOGRAPHICAL REFERENCES

Abdallah, A.G. 1960. *An instrumental study of the intonation of Egyptian colloquial Arabic.* Ph.D. diss., University of Michigan, Ann Arbor.

Abercrombie, David. 1967. *Elements of general phonetics.* Chicago: Aldine.

Al-Ani, Salman 1992. "Lexical stress variation in Arabic: An acoustic spectrographic analysis". *Proceedings of the Colloquium on Arabic Grammar*, ed. Kinga Dévényi and Tamás Iványi, 9–27. Budapest: Eötvös Loránd University and Csoma de Kőrös Society.

Alharbi, Lafi M. 1992. *Formal analysis of intonation: The case of the Kuwaiti dialect of Arabic.* Ph. D. diss., Herriot-Watt University.

Angoujard, Jean-Paul. 1990. *Metrical structure of Arabic.* Dordrecht: Foris.

Barkat, Melissa. 2000. *Détermination d'indices acoustiques robustes pour l'identification automatique des parlers arabes.* Thèse de Doctorat, Université Lumière, Lyon 2.

Benali, Ismaïl. 2004. "Le rôle de la prosodie dans l'identification de deux parlers algériens: L'algérois et l'oranais". *Actes du colloque Modélisation pour l'identification des langues*, 127–132. Paris: ENST-Télécom.

Benhallam, A. 1990. "Native speaker intuitions about Moroccan Arabic". *La linguistique au Maghreb*, 91–109. Rabat: Okad.

——. and Mohamed Dahbi. 1990. "Accents of Moroccan Arabic: A preliminary study". *La linguistique au Maghreb*, 111–125. Rabat: Okad.

Benkirane, Thami. 1982. *Etude phonétique et fonctions de la syllabe en arabe marocain.* Thèse de 3ème cycle, Université de Provence.

——. 1998. "Intonation in Western Arabic (Morocco)". Hirst and Di Cristo (1998:345–359).

——. 2000. *Codage prosodique de l'énoncé en arabe marocain.* Thèse d'Etat, Université de Provence.

——. 2001. "Etude perceptive et prosodique de la pause dans un corpus spontané en arabe parlé marocain". *Revue Parole* (Université de Mons-Hainaut, Mons, Belgique) 17–18–19.1–52.

——. 2003. "Statut prosodique et phonologique des séquences CVC dans le parler arabe du Maroc". *Langues et Linguistique* (Fes) 12.1–12.

Birkeland, Harris. 1954. *Stress patterns in Arabic.* Oslo: Norske Videnskabs-Akademi.

Bohas, Georges and Djamel Eddine Kouloughli. 1981. "Processus accentuels en arabe (parlers du Caire, de Damas et arabe classique)". *Analyse, Théorie* 1.1–59.

Boula de Mareüil, Philippe and Belynda Brahimi. 2004. "Rôle du segmental et du suprasegmental dans la perception de l'accent maghrébin en français". *Actes des 14ème Journées d'Etude sur la Parole.* Fes. (Available at www.lpl.univ-aix.fr/jep-talno4/proceed/actes/jep2004/Boula-Brahimi.pdf)

Brame, Michael K. 1973. "Stress assignment in two Arabic dialects". *A festschrift for Morris Halle*, ed. Stephen R. Anderson and Paul Kiparsky, 14–25. New York: Holt, Rinehart and Winston.

Bruce, Gösta. 1985. "Structure and function of prosody". *Proceedings of the French Swedish seminar on speech*, ed. B. Guérin and R. Carré, 549–559. Grenoble: Université de Grenoble.

Cantineau, Jean. 1960. *Cours de phonétique arabe.* Paris : Klincksieck.

Chahal, Dana. 1999. "A preliminary analysis of Lebanese Arabic intonation". *Proceedings of the 1999 Conference of the Australian Linguistic Society.* (Available at www.linguistics.uwa.edu.au/research/als99/proceedings)

Cutler, Anne and D. Robert Ladd (eds.). 1983. *Prosody: Models and measurements.* Berlin: Springer-Verlag.

Dauer, Rebecca M. 1983. "Stress-timing and syllable-timing reanalysed". *Journal of Phonetics* 11.51–62.

De Jong, Kenneth and Bushra A. Zawaydeh. 1999. "Stress, duration, and intonation in Arabic word-level prosody". *Journal of Phonetics* 27.3–22.

Ferguson, Charles A. 1957. "Two problems in Arabic phonology". *Word* 13.461–479.

Garde, Paul. 1968. *L'accent.* Paris : Presses Universitaires de France.

Hamdi, Rym. 2002. "Variation rythmique des dialectes arabes". *Actes des XXIVème Journées d'Etude sur la Parole*, 68–71. Nancy.

Hammoumi, Abdeljebbar. 1988. *Syllabation, accentuation, effacement et abrègement dans le parler de Fès et autres parlers marocains.* Thèse de 3ème cycle, Université de Paris 3.

Haydar, Y. and M. Mrayati. 1985. "A study of intonation and the melodic curve of phrases in Standard Arabic". *Travaux de l'Institut de Phonétique de Strasbourg* 17.75–118.

Hirst, Daniel and Albert Di Cristo (eds.). 1998. *Intonation systems: A survey of twenty languages.* Cambridge: Cambridge University Press.

Janssens, Gerard. 1972. *Stress in Arabic and word structure in the modern Arabic dialects.* Leuven: Peeters.

Kaye, Alan S. 1997. "Arabic phonology". *Phonologies of Asia and Africa*, ed. Alan S. Kaye, I, 187–204. Winona Lake, Ind.: Eisenbrauns.

Kenstowicz, Michael. 1983. "Parametric variations and accent in the Arabic dialects". *Papers from the Regional Meetings, Chicago Linguistic Society* 19.205–213.

Kharrat, Muhammad. 1994. *A descriptive analysis of Standard Arabic intonation with implications for teaching English as a second language.* Ph.D. diss., University of Kansas.

Kouloughli, Djamel Eddine. 1975. "Contribution à l'étude de l'accent en arabe littéraire". *Annales de l'Université d'Abidjan* Série H Linguistique 115–130.

Kulk, Friso. 2003. *Intonatie in het Arabische dialect van Damascus.* M.A. thesis, University of Amsterdam.

——, Cecilia Odé, and Manfred Woidich. 2003.

“The intonation of colloquial Damascene Arabic: A pilot study”. *Proceedings of the Institute of Phonetic Sciences* 25.15–20.

Ladefoged, Peter. 1975. *A course in phonetics*. New York: Harcourt Brace Jovanovich.

Lahlou, Moncef. 1982. *Prolégomènes à l’étude de l’intonation de l’arabe marocain de Fès*. Thèse de 3$^{\text{ème}}$ cycle, Université de Provence.

Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge, Mass.: MIT Press.

Loi Corvetto, Ines. 1982. “L’intonazione nell’arabo siriano”. *Lingua e Estile* 17:3.371–393.

Maas, Utz. 2002. “L’union linguistique maghrébine”. *Aspects of the dialects of Arabic today*, ed. Abderrahim Youssi, Fouzia Benjelloun, Mohamed Dahbi, and Zakia Iraqui-Sinaceur, 211–222. Rabat: Amapatril.

Maddieson, Ian. 1985. “Phonetic cues to syllabification”. *Phonetic linguistics*, ed. Victoria A. Fromkin, 203–221. New York: Academic Press.

Mawhoub, Mourad. 2000. *L’intonation en arabe standard moderne, étude phonétique*. Thèse de doctorat d’état, Université Cadi Ayyad, Béni Mellal. [In Arabic.]

——. 2004. “Prosodie et ordre des constituants dans l’énoncé en arabe standard moderne”. *XXV$^{\text{ème}}$ Journées d’Etude sur la Parole, section Traitement automatique de l’arabe, Fès, Maroc*. Fes.

Mayer-Lambert. 1897. “De l’accent en arabe”. *Journal Asiatique*, 9th series, 10.402–413.

McCarthy, John J. 1980. “A note on the accentuation of Damascene Arabic”. *Studies in the Linguistic Sciences* 10:2.77–98.

Moutaouakil, Ahmed. 1984. “Le focus en arabe: Vers une analyse fonctionnelle”. *Lingua* 64.115–176.

Nazzi, Thierry, Josiane Bertoncini, and Jacques Mehler. 1998 “Language discrimination by newborns: Towards an understanding of the role of rhythm”. *Journal of Experimental Psychology: Human Perception and Performance* 24.756–766.

Nejmi, H. 1992. *Contribution à une étude des processus accentuels de l’arabe marocain de Casablanca*. Thèse de Doctorat, Université de Paris-3.

Nicaise, A. and N. Bacri. 1985. “Durée vocalique, variations de Fo et perception de frontières”. *Actes des 14èmes Journées d’Etude sur la Parole*. Paris.

Norlin, Kjell. 1989. “A preliminary description of Cairo Arabic intonation of statements and questions”. *Speech Transmission Laboratory Quarterly Progress and Status Report* 1.47–49.

Pike, Kenneth L. 1945. *The intonation of American English*. Ann Arbor: University of Michigan.

Rammuny, Raji. 1989. “Instrumental and auditory analysis of colloquial Jordanian Arabic: An update”. *Zeitschrift für Arabische Linguistik* 20.23–42.

Ramus, Franck and Jacques Mehler. 1999. “Language identification with suprasegmental cues: A study based on speech resynthesis”. *Journal of the Acoustical Society of America* 105.512–521.

Rosenhouse, Judith. 1995. “Features of intonation in Bedouin Arabic narratives of the Galilee (Northern Israel)”. *Dialectologia Arabica: A collection of articles in honour of the sixtieth birthday of Professor Heikki Palva*, 193–215. Helsinki: Finnish Oriental Society.

Rossi, Mario. 1985. “L’intonation et l’organisation de l’énoncé”. *Phonetica* 42:2–3.135–153.

Selkirk, Elizabeth. 1981. “Epenthesis and degenerate syllables in Cairene Arabic”. *Theoretical issues in the grammar of Semitic languages*, ed. Hagid Borer and Joseph Aoun, 3.209–232. Cambridge: Department of Linguistics, Massachusetts Institute of Technology.

Vaissière, Jacqueline. 1983. “Language-independent prosodic features”. Cutler and Ladd (1983:53–66).

Vijver, Ruben van de.1996. “Variation on an Arabic theme”. *Stress patterns of the world*, ed. Rob Goedemans, Harry van der Hulst, and Ellis Visch, 233–266. The Hague: Holland Institute of Generative Linguistics.

Welden, Ann. 1980. “Stress in Cairo Arabic”. *Studies in the Linguistic Sciences* 10:2.99–120.

Yeou, Mohamed. 2004. “Effects of focus, position and syllable structure on FO alignment patterns in Arabic”. *XXV$^{\text{èmes}}$ Journées d’Etude sur la Parole, Fès*. (Available at www.lpl.univ-aix.fr/jep-taln04/proceed/actes/arabe2004/PAMY11.pdf.)

Zahid, A. 1990. *L’accent en arabe standard moderne: Analyse acoustique, perceptive et articulatoire*. Thèse de Doctorat, Université de Paris-7.

Zakaria, A.A. 1984. *L’accent en arabe soudanais*. Thèse de 3$^{\text{ème}}$ cycle, Université de Franche-Comté, Besançon.

Thami Benkirane (University of Fes)

Prothetic Vowel

Prothesis (also prosthesis) is the addition of a segment, usually a vowel, to the beginning of a word (Crystal 1997:315; Trask 2000:266). Prothesis in Arabic involves the addition of a short vowel to prevent the occurrence of impermissible consonant clusters word-initially. There are differences between *fuṣḥā* and the modern spoken dialects in the application of the prothesis rule.

1. PROTHESIS IN *FUṢḤĀ*

Three basic syllable types (CV, CVC, and CVV) occur freely in the language; another three types (CVCC, CVVC, CVVCC) are limited to specific environments, mainly word-finally or before a pause (Bateson 1967:6–7; Holes 1995:62; Watson 2002:60; Badawi a.o. 2004:10). There is a general constraint against the occurrence of more than one consonant in syllable onsets or, as the Arab grammarians described it, ‘the meeting of two silent letters’ (*iltiqāʾ as-sākinayn*; Sībawayhi, *Kitāb* IV, 152; Zamaxšarī, *Mufaṣṣal* 349, 503). If a morphological process results in a potential CC onset, the prothetic vowel is added to become the nucleus of a new syllable, to which the first consonant is assigned as a coda. If the CC is at the absolute beginning of an utterance, a glottal stop, known in Arabic grammar as *hamzat al-waṣl*, serves as the onset of the syllable: [ʔiʃrab] 'drink!', [ʔiltiqaːʔ] 'meeting'.

The following are the environments in which a prothetic vowel is introduced (Ibn Jinnī, *Lumaʿ* 89; Wright 1996:19):

i. Before the definite article *l*, as in [ʔalmalik] 'the king'
ii. The imperatives of Form I verbs with strong first and second root consonants: [ʔismaʕ] 'hear, listen!', [ʔuktub] 'write!', [ʔimʃi] 'walk!'
iii. The perfect active, imperative, and *maṣdar* of Forms VII, VIII, IX, and X
iv. A handful of primary or nonderived nouns such as [(ʔi)bn] 'son', [(ʔi)sm] 'name', and [(ʔi)mraʔa] 'woman'

The unmarked realization of the prothetic vowel is *i* (*kasra*); the occurrence of *a* and *u* is restricted to specific environments (Sībawayhi, *Kitāb* IV, 152; Zamaxšarī, *Mufaṣṣal* 504; Bateson 1967:7). The vowel *a* occurs only before the definite article, and *u* in the imperative of Form I verbs with the stem vowel *u*, as in [ʔuktub] 'write!' and [ʔuxrudʒ] 'get out!'.

2. PROTHESIS IN THE MODERN SPOKEN DIALECTS

While the application of the prothesis rule as described by Sībawayhi, az-Zamaxšarī, and Ibn Jinnī is practically the same in all forms of *fuṣḥā*, old as well as modern (Bateson 1967:7; Badawi a.o. 2004:12), a number of phonological and morphosyntactic changes in the modern spoken dialects have resulted in more underlying consonant clusters and, thus, more prothesis environments than in *fuṣḥā*.

The first of these changes is the deletion of short vowels in unstressed open syllables (→ syncope). As a result of this rule, words like *nazalt(u)* 'I went down', *qirāʾa* 'reading', and *tazawwaj(a)* 'he got married' in *fuṣḥā* are pronounced *nzilt* (Cowell 1964:69), *qrāyæ* (Talmoudi 1980:60), and *tzawwaj* (Watson 2002:60) in Syrian, Tunisian, and Yemeni Arabic, respectively. Such initial two-consonant clusters occur in Iraqi Arabic (Erwin 1963:3) and Bahraini Arabic (Holes 1995:64), among others.

Another development that has affected the dialects is the loss of case and mood markers and other word-final short vowels, which has the effect of eliminating word-initial CC clusters in *fuṣḥā* in phrases like *kataba l-kitāb* 'he wrote the book', *katabtu l-kitāb* 'I wrote the book', *qismu l-luġāt* 'the department of languages', etc. The loss of these endings has resulted in sequences of three or four consonants at word boundaries, as the Syrian Arabic equivalents of these phrases show (before the application of the prothesis rule): *katab l-ktāb*, *katabt l-kitāb*, *ʾism l-luġāt*.

Syllabification in Arabic is a phrase-level process; it does not recognize word boundaries. A word-initial two-consonant cluster is consequently broken up by prothesis, and the prohibition on CC onsets is maintained. Consider, for example, the way the underlying initial CC cluster in the Syrian Arabic word *ktāb* 'book' is treated in a phrasal context. (Syllable boundaries are designated with quotation marks, and the prothetic vowel appears in italics.)

/katabu ktāb/
[ka.ta.bu k.taːb]
'They wrote a book'

/katbat ktāb/
[kat.ba.t *i* k.taːb]
'She wrote a book'

/katabt ktāb/
[ka.tab.t *i* k.taːb]
'I wrote a book'

The first example contrasts with the other two in that the word-final vowel provides a nucleus to eliminate the CC onset. In the second and third examples, the prothetic vowel serves the same function.

In general, the prothetic vowel takes the place of *fuṣḥā* final short vowels in the dialects. Compare the rendering of three examples in *fuṣḥā*, Egyptian, and Syrian:

fuṣḥā
[ka.ta.ba l.-ki.taːb]
ka.tab.tu l.-ki.taːb]
[qis.mu l.-lu.ɣaːt]

Egyptian
[ka.ta.b *i* l.-ki.taːb]
[ka.tab.t *i* l.-ki.taːb]
[ʔis.m *i* l.-lu.ɣaːt]

Syrian
[ka.tab. l*i*-k.taːb]
[ka.ta.bit. l*i*-k.taːb]
[ʔis.m *i* l.-lu.ɣaːt]

Note that the prothetic vowel rule has been simplified in the dialects: the *a* of the definite article and the final short vowels *a*, *u*, *i* of *fuṣḥā* have all been replaced by *i*, the unmarked variant of the prothetic vowel.

In conclusion, while the dialects differ from *fuṣḥā* in their lack of final short vowels and from one another in the placement of the prothetic vowel, the result of the application of the prothesis rule is virtually the same in all varieties of the language, i.e. preventing the occurrence of impermissible syllable types, particularly syllables with CC onsets, and maximizing the occurrence of the three basic syllable types CV, CVC, and CV shared by all these varieties.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Ibn Jinnī, *Lumaʿ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *Kitāb al-lumaʿ fī n-naḥw*. Ed. Hadi M. Kechrida. Uppsala: Almqvist and Wiksell, 1976.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān ibn Qanbar Sībawayhi, *al-Kitāb*. 4 vols. Ed. ʿAbd as-Salām Muḥammad Hārūn. Cairo: Dār al-Qalam, 1966–1975.

Zamaxšarī, *Mufaṣṣal*= ʾAbū l-Qāsim Maḥmūd ibn ʿUmar az-Zamaxšarī, *al-Mufaṣṣal fī ṣināʿat al-ʾiʿrāb*. Ed. Muḥammad Muḥammad ʿAbd al-Maqṣūd and Ḥasan Muḥammad ʿAbd al-Maqṣūd. Cairo: Dār al-Kitāb al-Miṣrī, 2001.

SECONDARY SOURCES

Badawi, El-Said, Michael Carter, and Adrian Gully. 2004. *Modern Written Arabic: A comprehensive grammar*. London and New York: Routledge.

Bateson, Mary Catherine. 1967. *Arabic language handbook*. Washington, D.C.: Center for Applied Linguistics.

Cowell, Mark. 1964. *A reference grammar of Syrian Arabic*. Washington, D.C.: Georgetown University Press.

Crystal, David. 1997. *A dictionary of linguistics and phonetics*. Oxford: Blackwell.

Erwin, Wallace. 1963. *A short reference grammar of Iraqi Arabic*. Washington, D.C.: Georgetown University Press.

Holes, Clive. 1995. *Modern Arabic: Structures, functions and varieties*. London: Longman.

Talmoudi, Fathi. 1980. *The Arabic dialect of Susa (Tunisia)*. Goteborg: Acta Universitatis Gothoburgensis.

Trask, R.L. 2000. *The dictionary of historical and comparative linguistics*. Chicago: Fitzroy Dearborn.

Watson, Janet C.E. 2002. *The phonology and morphology of Arabic*. Oxford: Oxford University Press.

Wright, William. 1996. *A grammar of the Arabic language*. Repr. Beirut: Librairie du Liban.

MUNTHER A. YOUNES (Cornell University)

Proverb

1. GENERAL OBSERVATION

A proverb is a common, pithy, and succinct statement which has been current in a language for generations and which sums up daily experiences as brief 'words of wisdom'. As a Dutch proverb says, "Proverbs are the daughters of daily experience" (*Spreekwoorden zijn de dochters van de dagelijkse ondervinding*). This genre is usually associated with the folklore and ethos of a certain society, although parallel proverbs are found in remote cultures. The dynamics of the proverb stems from the fact that it is easy to memorize and apply to similar circumstances and carries more weight than a lengthy statement, thus indicating a high level of rhetoric and eloquence. Moreover, the popularity of proverbs can be attested by the large number of collections that originated from ancient Egypt, Mesopotamia, Persia, Palestine, Greece, and Rome as well as India, China, and the Far East at large. Suffice it to mention, as examples, the Biblical books of *Proverbs* and *Ecclesiastes*, which are ascribed to King Solomon, the apocryphal book of *Ecclesiasticus*, ascribed to Jesus, the son of Sirach, or Publilius Syrus' *Sententiae*. However, since European languages have many cognate terms (see Sec. 2.2, below), it is often difficult, if not impossible, to distinguish between these terms. In general, the Greek *gnōmē* and the Latin *sententia* refer to sayings ascribed to famous personalities, while the Greek *paroimía* and the Latin *proverbium* denote a popular saying current in the language.

2. THE ARABIC PROVERB

2.1 *Introduction*

Works on the Arabic proverb (*maṯal*, pl. *ʾamṯāl*) and collections of *ʾamṯāl* which have come down to us over the years, as well as those collections which have, since the 19th century, been produced by Arab and non-Arab scholars, highlight the importance attached to this literary genre. This is not only because it is "one of the earliest forms of oral literature" (Serjeant 1983:115) but also because Arabic proverbs throw light on almost every aspect of pre-Islamic and Islamic daily life, in spite of the difficulties in interpreting many of them and the uncertainties regarding the origin of most of them (cf. Nicholson 1969:31).

Although it is impossible to establish the exact number of classical and modern collections of Arabic *ʾamṯāl*, a rough estimate of those which have come down to us and those which were lost but for which references are made in the relevant literature, as well as those which continue to appear to the present day, show the existence of hundreds of collections, containing thousands of proverbs of enormous popularity.

2.2 *The etymology of the word* maṯal

The Arabic word *maṯal*, which has equivalents based on the same root in some other Semitic languages, such as Hebrew, Syriac, and Aramaic, refers to a 'proverb' as well as to all other cognate terms, such as 'adage', 'aphorism', 'apophthegm', 'axiom', 'dictum', 'fable', 'gnome', 'maxim', 'parable', 'saw', 'saying', 'wisecrack', and a few more, and even to 'idiom' (→ phraseology); synonyms are *ḥikma* or *qawl* (*maʾṯūr/sāʾir/šāʾiʿ)*. Arab lexicographers claim that the root *m-ṯ-l* means 'to resemble, to be or look like', and also 'to stand straight' or 'to appear before'. In defining the term *maṯal*, al-Maydānī (*Majmaʿ* I, 5), quoting al-Mubarrad, says:

> The word *maṯal* is derived from the word *miṯāl*. It means a popular saying by which the condition of one [person, situation, etc.] is compared with another one. The etymology of the word pertains to 'resemblance'. Hence, when they say: 'He appeared before him', it means 'He stood before him'. That is to say: 'He appeared in the standing position', [and when we say] 'A person bears more resemblance than another', it means that he has more resembling features [in comparison with another person] (*al-maṯal maʾxūḏ min al-miṯāl, wa-huwa qawl sāʾir yušabbahu bihi ḥāl aṯ-ṯanī bi-l-ʾawwal, wa-l-ʾaṣl fīhi t-tašbīh fa-qawluhum ʿmaṯala bayna yadayhiʾ ʾiḏa ntaṣaba maʿnāhu ʾašbaha ṣ-ṣūra al-muntaṣiba wa-fulān ʾamṯal min fulān ʾay ʾašbahu bi-mā lahu [min] al faḍl*)

Although al-Maydānī's/al-Mubarrad's definition suggests a semantic expansion from 'resemblance' to 'proverbial saying', the fact that the same expansion occurs in other Semitic languages, such as Akkadian, Hebrew, and Aramaic, suggests that the root is polysemic, although in Phoenician and Hebrew it is a homonym, as it also means 'to rule, reign'.

2.3 *The origins of the* maṯal

Similarly to other languages, in the majority of cases it is unknown who coined the proverb or under what circumstances it was coined. Many proverbs suggest that they are in fact morals drawn from parables. However, even if the parable is well known, it is often impossible to establish when, where, and by whom the moral became a proverb. Moreover, even when a proverb is attributed to a certain person, it is quite possible that it was only ascribed to him in order to grant the proverb authority, for the same reason that some *ḥadīṯ*s and poems were attributed to famous personages. Furthermore, the fact that there exist many variations of proverbs (there are, for instance, twelve different proverbs beginning with *ʾaʿazz min…* 'mightier, dearer, rarer than…') obscures the origin, or as aṯ-Ṯaʿālibī (*Ṯimār*, passim) puts it, *fīhi ʾaqāwīl muxtalifa*.

Instead, one often finds vague formulas like *yuqālu* 'it is said', *yuḍrabu bihi l-maṯal* 'it is used as a proverb', *ḏahaba(t)/sāra(t) maṯalan* 'it became a proverb', *ʿalā raʾy al-maṯal* 'according to the proverb', and a few more. However, in view of the 'simple' message expressed by some proverbs, it may be assumed that a large number of them have undergone a transitional process, starting as a straightforward statement and gradually gaining popularity by wide circulation and eventually becoming a common proverb. Examples are *taḍarraʿ ʾilā ṭ-ṭabīb qabla ʾan tamraḍ* 'beseech the doctor before you fall ill', i.e. 'take precautions'; *al-xunfusāʾ ʾiḏa mussat natinat* 'when a beetle is touched, it stinks', i.e. 'don't get involved in an unpleasant matter'; *aš-šarr qadīm* 'there has always been evil', i.e. 'he has a history of being a rascal'.

On the other hand, as a great number of proverbs appear in ancient Arabic texts, we naturally tend to regard the writer as the coiner. Nevertheless, it is quite possible that the proverb had been in use before the composition of the text and that the author was using it in his text (*taḍmīn*) and even, to borrow Frayḥa's (1953:xvi) phrase, "retouching" a common utterance.

Generally, Arabic proverbs may have three major sources:

i. Literary Arabic texts, such as the *Qurʾān*, *Ḥadīṯ*, and poetry. Examples: *ʾinna maʿa al-ʿusr yusr* 'with every difficulty there is a relief' (Q. 94/5,6); *lā yuldaġ/yulsaʿ al-muʾmin min juḥr marratayn* 'the believer is not bitten twice from the same burrow' (ʾAbū Hurayra, 597–675); *ṯalāṯat al-mustaḥīlāt: al-ġūl wa-l-ʿanqāʾ wa-l-xill al-wafiyy* 'the impossible things are three: a demon, a phoenix, and a loyal friend' (attributed to the poet Ṣafī d-Dīn al-Ḥillī, 1277–1339).
ii. Oral proverbs, which are mainly in colloquial Arabic. Examples: *tītī tītī mitl mā ruḥti mitl mā jīti* 'you have come back the same way you went' (Palestinian); *rāḥ yuxṭubhā-lo ʾigawwizhā* 'he went to arrange the engagement on behalf of a friend and he married the girl himself' (Egyptian); *qīs gabl mā tġīṣ* 'measure the depth of the water before you dive' (Syrian).
iii. Borrowed proverbs are those which have been translated from a foreign language. Gibb (1966:38) believes that proverbs coined in pre-Islamic days are original and if parallels are found, they should be the result of parallel development. Post-Islamic proverbs include many sayings originating from Syriac, Persian, Greek, and Latin and, at a later stage, from other languages. This is often attested by their prosaic style. The Arabs usually call such borrowed proverbs *muwalladūn* 'half-caste'. Modern Arabic contains many examples of these, e.g. *al-ḥāja ʾumm al-ixtirāʿ* 'necessity is the mother of invention'.

It is worth mentioning that while all ancient collections of proverbs are mainly in Classical Arabic with sporadic references to a few colloquial examples (preceded by the formula *wa-l-ʿāmma taqūlu* 'the common people say'), the majority of modern collections are arranged by their geographical circulation and are therefore in colloquial Arabic, in accordance with the dialect spoken in the area where these proverbs are used.

2.4 *The structure of the* maṯal

Arabic proverbs may be divided into seven main syntactic patterns:

i. Ellipticals are those proverbs which are incomplete sentences. Example: *ṣaḥīfāt al-Mutalammis* lit. 'the letter of al-Mutalammis', i.e. 'he brings about his own destruction, he digs his own grave' (based on a story about a person named al-Mutalammis who carried a letter from the governor in which the latter ordered his execution).
ii. Nominals are those proverbs which begin with a noun. Example: *ad-dunyā yawm ʿasal wa-yawm baṣal* lit. 'the world is one day honey and another day onion', i.e. 'life consists of good days and bad days'; *ʾana ʾamīr wa-ʾanta ʾamīr wa-man yasūqu l-ḥamīr* lit. '[if] I am an emir and you are an emir, who is going to drive the donkeys?', i.e. 'we need to share unpleasant responsibilities'.
iii. A large number of proverbs denote a comparison between situations and characteristics involving human beings, animals, objects, places, or events. The most prolific are those based on the pattern *ʾafʿal min…* 'more than…', in which a person, animal, object, etc. is said to possess a property to a higher degree than others. Arab lexicographers estimate that there are well over one thousand of these in Arabic. This phenomenon may be regarded as a kind of a simile (see below). Hence, the proverb *ʾawfā min as-Samawʾal* suggests that the person in mind is more trustworthy than the pre-Islamic poet as-Samawʾal ibn ʿĀdiyāʾ. Another example, *ʾaʿqad min ḏanab aḍ-ḍabb*, means that the matter is more complicated than a lizard's tail.
iv. Verbals are those proverbs which, like most sentences in Literary Arabic, begin with a verb and have the word order VS(O). Example: *bāta maʿa d-dajāj ʾaṣbaḥa yaqūqu*

[in some dialects *yuqāqī*] ‘he spent the night with the chicken and woke up clucking’, i.e. ‘he quickly acquires habits’.

v. Imperatives are those proverbs which begin with an imperative, whether positive or negative (‘do!’ or ‘don’t!’). Examples: *isʾal ʿan al-jār qabla d-dār wa-ʿan ar-rafīq qabla aṭ-ṭarīq* ‘inquire about the neighbor before buying a house and about your companion before setting out on a journey’; *lā tanha ʿan xuluq wa-taʾtī miṯlahu* ‘don’t forbid people to do things and then do yourself what you have forbidden’.
vi. Many proverbs are formed as conditional sentences. Example: *ʾiḏa ʾaradta ʾan tuṭāʿ fa-sal mā yustaṭāʿ* ‘if you want people to obey you, ask yourself if what you ask them to do would be possible’.
vii. Many proverbs begin with an unspecified subject ‘whoever…’. Example: *man ḥafara ḥufra waqaʿa fīhā* ‘he who digs a pit [for someone else], will fall inside it himself’.

2.5 *The semantic/stylistic aspect of the* maṯal

Since proverbs are characterized by their conciseness and ability to condense and even conceal thoughts and ideas, it is not surprising that many of them possess deep meanings and have various connotations behind their ‘innocent’ literal meaning.

Thus, one may draw various conclusions from a proverb which on the surface refers to simple actions or facts. The proverb *ḍarabanī wa-bakā sabaqanī wa-štakā* ‘he beat me and cried and then rushed to complain about me’, for instance, suggests that the person who behaves in this way is a troublemaker, cunning, sly, a hypocrite, querulous, violent, unfair, a liar, malicious, and many more bad things, and that his complaints are baseless.

A large number of proverbs are figures of speech and therefore more colorful.

i. Simile. Example: *ka-l-ḥādī wa-laysa lahu baʿīr* ‘like a camel driver who has no camels’ (parallel to ‘who is worse shod than the shoemaker’s wife’).
ii. Metaphor. Example: *warāʾ al-ʾakama mā waraʿahā* lit. ‘there is something behind the hill’, i.e. ‘there is more to it than meets the eye’.
iii. Synecdoche. Example: *bi-yadayya lā bi-yadayka ʿAmr* ‘with my own hands, not yours, ʿAmr’, i.e. ‘it’s my decision, for better or worse’.
iv. Metonymy. Example: *ṣāḥib al-māl ṣāḥib al-kalima* ‘money talks’.
v. Hyperbole. Example: *bayḍat ad-dīk* ‘[as rare as] a cock’s egg’.
vi. → Euphemism. Example: *law ḏāt siwār laṭamatnī* lit. ‘I wish a woman wearing a bracelet had hit me’, i.e. ‘if only a respectable woman had slapped me [and not a woman of low status]’.
vii. Epithet. Example: *(ʾana) ibn jalā* ‘I am a man of honor’.
viii. Polyptoton. Example: *dur ad-dawra wa-law dārat xuḏ ʾaṣīla wa-law bārat* ‘take a circuitous route, even though it is longer [if it is safer], and marry a noble lady, even if she is a spinster’.
ix. Oxymoron. Example: *aš-šarr qalīluhu kaṯīr* ‘a little of evil is a lot’.
x. Zeugma. Example: *šahādat al-fiʿāl xayr min šahādat ar-rijāl* ‘the evidence shown by deeds is better than [a boasting] statement’. The word *šahāda* has here a double meaning, ‘evidence’ and ‘statement’.

Other popular features of proverbs include the fact that they have rhyme and therefore have more impact and are easy to memorize, e.g. *ʿinda l-imtiḥān yukramu l-marʾ ʾaw yuhān* ‘at the time of test, one is either praised or disgraced’. Many proverbs are structured as → paronomasia (i.e. puns), e.g. *man ʾaṭāʿa ġaḍabahu ʾaḍāʿa ʾadabahu* ‘he who obeys his anger loses his manners’. Some proverbs take the form of chiasmus, e.g. *kalb aš-šayx šayx al-kilāb* ‘the master’s dog is the dogs’ master’.

2.6 *The contents of the* maṯal

The themes of the *maṯal* are numerous and variegated, covering almost every sphere of daily life. They include human relations, beliefs, values, aspirations, wishes, and good qualities of people, as well as their disappointments, frustrations, shortcomings, and negative characteristics.

Many proverbs may be described as didactic since they give advice and guidance or warning and admonition, revealing the lessons learned from long experience, expressed with great firmness, yet with engaging simplicity, e.g. *ittaqi šarr man ʾaḥsanta ʾilayhi* ‘protect yourself against the evil of the one to whom you have been kind’; *lā tuʾaxxir ʿamal al-yawm li-ġad* ‘don’t put off till tomorrow what you should do today’.

Similar to fables, several proverbs have a background story whose moral or punch line has become a proverb. For instance, a person divorced his domineering wife. Another person who wished to marry her approached the former husband seeking information about her. The latter answered: *ʿiš rajaban tara ʿajaban* ‘live through the month of Rajab and you will see wonders’. The point is that since the month of Rajab was one of the months during which no wars were launched, the reply suggests that the wife will be obedient for a short period, after which problems will start again. Another interpretation is that although the husband hopes that they will live in harmony for some time, he should expect a war even in the month of Rajab.

2.7 *Variations of the* maṯal

Although proverbs are fairly cohesive, it is not unusual to find many parallel proverbs or different versions of the same proverb. For example, many proverbs referring to ‘stupidity’ begin with the word *ʾaḥmaq* ‘stupid, fool’, followed by names of different people or animals known for their folly or their foolish behavior, e.g. *ʾaḥmaq min Duġa/Habannaqa/Šaranbaṯ* (names of people who have acted foolishly, as recorded by some folktales), or *ʾaḥmaq min naʿāma* ‘ostrich’ or *ʾaḥmaq min al-qābiḍ ʿalā l-māʾ* ‘he who tries to hold water in his hand’, with the variant *ʾaḥmaq min nāṭiḥ al-māʾ/aṣ-ṣaxr* ‘he who butts the water/the rock’.

Likewise, variation occurs through the use of different words in certain proverbs, e.g. *al-ʾinsān yudabbir wa-llāh yuqaddir* and *al-ʾinsān bi-t-tafkīr wa-llāh bi-t-tadbīr* ‘man thinks/plans and God decides’ (parallel to ‘man proposes, God disposes’).

2.8 *Contradictory* ʾamṯāl

As in many other languages, there are some contradictory proverbs, which suggests that they were coined by different people or in different situations or circumstances, e.g. *ʾarsil ḥakīm wa-ʾawṣihi* ‘send a wise man and give him instructions’ vs. *ʾarsil ḥakīm wa-lā tūṣihi* ‘send a wise man and do not give him instructions’. Other examples include *rubba kalima salabat niʿma* ‘many a word has spoiled a favor’ vs. *rubba kalima ʾafādat niʿma* ‘many a word has helped a favor’; *al-ʾiʿāda ʾifāda* ‘repetition is beneficial’ vs. *at-tikrār li-l-ḥimār* ‘repetition is for donkeys’.

2.9 *Humor and the* maṯal

Many Arabic proverbs are characterized by humor, which is achieved by sarcasm, cynicism, irony, and even *Schadenfreude*. Although one may assume that some of them are based on real situations, familiarity with the situation is not necessary to understand the proverb in order to use it in similar circumstances, e.g. *rāḥat taʾxuḏ ṯaʾr ʾabīha wa-rajaʿat ḥublā* ‘she went out to avenge the death of her father and she returned pregnant’; *muš kull man ḥaṭṭ ʿalā badano maryūl ṭabbāx* ‘not everyone who wears an apron is a cook’; *il-faqīr ʾaxad il-faqīre u-jābu šaḥḥāde zġīre* ‘the poor [man] married a poor [woman] and they got a little beggar’.

2.10 *Parallel* ʾamṯāl *between literary and spoken Arabic*

An interesting phenomenon concerning the two levels of Arabic, known as → diglossia, is the existence of the same proverb in both strata. While one may assume a collateral development, some instances suggest the existence of a ‘translated version’ from one into the other (see Table 1).

Table 1. Literary and colloquial variants of proverbs

Literary (*fuṣḥā*)	Spoken (Levantine *ʿāmmiyya*)	Translation
wāfaqa Šann Ṭabaqa	*ṭanjara wa-lāqat ġiṭāha*	‘they suit each other’
mudd rijlak ʿalā qadr al-kisāʾ	*ʿalā qadd bsāṭak mudd ijrēk*	‘stretch your feet to the size of the carpet/cover’
li-l-ḥiṭān ʾāḏān	*il-ḥāyiṭ ʾilha ʾiḏān*	‘walls have ears’

2.11 *The use of the* maṯal

The fact that Arabic proverbs are one of the oldest genres of Arabic literature and that they play a key role in Arab lore and legacy demonstrates their popularity. Arabic proverbs have always been used in Classical Arabic texts, both in poetry and prose. This tendency has never ceased as modern Arab writers continue to incorporate proverbs into their narratives and verses. Moreover, the media at large are using proverbs widely in articles and in news bulletins and other programs. In addition, Arabic proverbs may appear as titles of books and articles, as well as headlines, and in concluding remarks, both verbal and written. Furthermore, a striking development in recent years shows that in modern texts, speeches, and conversation both literary and colloquial, proverbs are used indiscriminately, i.e. they cite proverbs from both strata of the language, the spoken and the literary, without indicating the different levels. This means that the traditional division between literary and colloquial proverbs, which was rigidly observed in the past, is no longer adhered to, either by writers or by speakers.

Bibliographical references

The bibliography provided by Sellheim (1991) covers most of the important collections published until 1989; the following list adds titles published since 1990 as well as some not mentioned by Sellheim.

I. General collections of proverbs in Classical Arabic

ʿAskarī, ʾAbu Hilāl al-, *Kitāb jamharat al-ʾamṯāl*. Ed. Muḥammad ʾAbū l-Faḍl ʾIbrāhīm and ʿAbd al-Majīd Qaṭāmiš. 2nd ed. 2 vols. Beirut: Dār al-Jīl, 1988.

ʾIṣbahānī, ʾAbū Muḥammad ʿAbdallāh ibn Muḥammad al-, *Kitāb al-ʾamṯāl fī l-ḥadīṯ an-nabawī*. Ed. ʿAbdallāh ʿAbd al-Ḥamīd Ḥamīd. Bombay: Al-Darussalafiah, 1987.

Maʾlūf, Lūwīs, *al-Munjid*. 20th ed. Beirut: Dār al-Mašriq, 1969.

Qabbiš, ʾAḥmad, *Majmaʿ al-ḥikam wa-l-ʾamṯāl fī š-šiʿr al-ʿarabī*. Damascus and Beirut: Dār ar-Rašīd, 1985.

Sieny, Mahmoud Isma'il, N.M. Abdul-'Aziz, and M.A. Suleiman. 1992. *Muʿjam al-ʾamṯāl al-ʿarabiyya*. Beirut: Librairie du Liban.

Zamaxšarī, ʾAbū l-Qāsim az-, *al-Mustaqṣā fī ʾamṯāl al-ʿArab*. 2 vols. Beirut: Dār al-Kutub al-ʿIlmiyya, 1987.

II. Collections of proverbs in various dialects

Since it is impossible to list here all the collections of colloquial Arabic proverbs published since the 19th century, the following list provides one representative item for each country.

Algeria – Ben Cheneb, M. 2003. *Proverbes de l'Algérie et du Maghreb*. Repr. Paris.

Arabia – al-Juhaymān, A. 1383 A.H. *al-ʾAmṯāl aš-šaʿbiyya fī qalb al-Jazīra al-ʿArabiyya*. 3 vols. Beirut.

Bahrain – al-ʿAbdallāh, J.M.H. 1984. *Min ʾamṯālinā š-šaʿbiyya*. Bahrain.

Emirates – as-Sabbāġ, E.R. 1978. *al-ʾAmṯāl aš-šaʿbiyya fī dawla ʾImārāt al-ʿArabiyya al-Muttaḥida*. United Arab Emirates.

Egypt – al-Baqlī, M.Q. 1986. *al-ʾAmṯāl aš-šaʿbiyya*. Cairo.

Iraq – at-Tikrītī, A. 1971. *Jamharat al-ʾamṯāl al-baġdādiyya*. 4 vols. Baghdad.

Jordan – al-ʿAmad, H. 1978. *al-ʾAmṯāl aš-šaʿbiyya al-ʾurdunniyya*. Amman.

Kuwait – Nūrī, A. 1981. *al-ʾAmṯāl ad-dārija fī l-Kuwait*. Kuwait.

Lebanon – Yaʿqūb, A.B. 1989. *Mawsūʿat al-ʾamṯāl al-lubnāniyya*. 3 vols. Beirut.

Libya – Ḥaqīq, M. 1982. *al-ʾAmṯāl aš-šaʿbiyya fī Lībiyā*. Tripoli.

Morocco – Westermarck, E. 1930. *Wit and wisdom in Morocco*. London.

Palestine – ʾAsqar, Š.ʿA. 1933. *Kitāb aṭ-ṭurfa al-bāhija fī l-ʾamṯāl wa-l-ḥikam al-ʿarabiyya ad-dārija*. Jerusalem.

Syria – Kayyāl, M. 1993. *Muʿjam durar al-kalām fī ʾamṯāl ʾahl aš-Šām*. N.p.

Yemen – al-ʾAkwaʿ, I. 1984. *al-ʾAmṯāl al-yamaniyya*. 2 vols. Beirut and Sanaa.

Primary sources

Ibn Manẓūr, *Lisān* = Jamāl ad-Din Mukarram ibn Mukarram Ibn Manẓūr, *Lisān al-ʿArab*. Beirut: Dār Ṣādir and Dār Bayrūt, 1955.

Maydānī, *Majmaʿ* = ʾAbū l-Faḍl ʾAḥmad ibn Muḥammad ibn ʾAḥmad an-Nīsābūrī al-Maydānī, *Majmaʿ al-ʾamṯāl*. 2 vols. Cairo, 1955.

Ṯaʿālibī, *Ṯimār* = ʾAbū Manṣūr ʿAbd al-Malik ibn Muḥammad aṯ-Ṯaʿālibī, *Kitāb ṯimār al-qulūb fī l-muḍāf wa-l-manṣūb*. Cairo, 1965.

Secondary sources

Al-Dubaib, Ahmad M. 1966. *A critical and comparative study of the ancient Arabic proverbs contained in al-Maidani's collection*. Ph.D. diss., Leeds University.

Frayḥa, ʾAnīs. 1953. *Modern Lebanese proverbs, collected at Rās al-Matn, Lebanon*. Jounié: Imprimerie Missionnaire Libanais.

Gibb, Hamilton A.R. 1966. "Introduction to the proverbs of Arabia". *Racial proverbs*, ed. Selwyn Gurney Champion, xxvii–xxxix, 329–344. London: Routledge. (1st ed., 1938.)

Goitein, Shelomo Dov. 1966. "The present-day Arabic proverb as a testimony to the social history of the Middle East". Shelomo Dov Goitein, *Studies in Islamic history and institutions*, 361–379. Leiden: E.J. Brill.

Huhne, F. Edward. 1902. *Proverb lore*. London: Elliot Stock.

Mieder, Wolfgang and Alan Dundes (eds.). 1981. *The wisdom of many: Essays on the proverb*. New York and London: Garland. (Repr., Madison: University of Wisconsin Press, 1994.)

Nicholson, Reynold A. 1969. *A literary history of the Arabs*. Cambridge: Cambridge University Press.

Ridout, Ronald and Clifford Witting. 1972. *English proverbs explained*. London: Pan Books.

Sellheim, Rudolf. 1954. *Die klassisch-arabischen Sprichwörtersammlungen*. The Hague: Mouton.

——. 1991. "*Mat̲al*". *Encyclopaedia of Islam*, VI, 815-825. 2nd ed. Leiden: E.J Brill.

Serjeant, Robert Bertram. 1983. "Early Arabic prose". *The Cambridge history of Arabic literature to the end of the Umayyad Period*, ed. A.F.L. Beeston, T.M. Johnstone, R.B. Serjeant, and G. Rex Smith, 114–153. Cambridge: Cambridge University Press.

Shivtiel, Avihai. 1994. "Some aspects of Arabic proverbs". *Proceedings of the Colloquium on Popular Customs and Monotheistic Religions in the Middle East and North Africa*, ed. Alexander Fodor and Avihai Shivtiel, 50–62. (= *Budapest Studies in Arabic* 9–10.) Budapest: Eötvös Loránd University.

——. 1996. "Women in Arabic proverbs from Yemen". *New Arabian Studies*, III, ed. G. Rex Smith, J.R. Smart, and B.R. Pridham, 164–175. Exeter: Exeter University Press.

——. 2000. " The image of the Jew in the Yemenite proverb". *Israel and Ishmael studies in Muslim-Jewish relations*, ed. Tudor Parfitt, 167–174. London: Curzon.

Smith, F.P. 1979. *The Oxford dictionary of English proverbs*. Oxford: Clarendon Press.

Taylor, Archer. 1931. *The proverb*. Cambridge, Mass.: Harvard University Press.

Tikrītī, ʿAbd ar-Raḥmān at-. 1984. *Dirāsāt fī l-mat̲al al-ʿarabī al-muqārin*. Baghdad: Maʿhad al-Buḥūt̲ wa-d-Dirāsāt al-ʿArabiyya.

Trench, Richard Chenevix. 1905. *Proverbs and their lessons*. Rev. ed. A. SmythePalmer. London: Kegan Paul. (1st ed., New York: Redfield, 1853.)

Avihai Shivtiel (Leeds)

Pseudo-correction → Hypercorrection

Pseudodual

The notion of 'pseudodual' indicates the use of a dual ending after a noun to mark the plural, or to mark a number greater than one without specifying whether it is dual or plural. The phenomenon, which is known in most Arabic dialects, is restricted to paired parts of the human body (and the word for 'parents') typically those of most frequent occurrence. An example of such a pseudodual is *arbaʿ žrēn* 'four feet'.

The definition and first comprehensive analysis of the phenomenon was given by Blanc, who wrote an extensive article about it in 1970. Before and after that, the phenomenon had been observed in several field studies. However, the problem has not yet been fully elucidated, and for some characteristics relating to pseudodual, a more detailed analysis is needed.

Pseudodual may be seen in the total perspective of the development of the dual (Fontinoy 1969). The cognitive origin of the dual as a category is debated. It may have developed by analogy out of natural dual, e.g. paired body parts, or it may have come about as a numerical category, simple dual. In older stages of Semitic, the dual is used extensively, both for natural and for simple dual. (Simple dual may also be referred to as 'occasional', 'true', or 'enumerative' dual.) The general development, in Semitic and other languages, is toward the weakening and final disappearance of the dual as a grammatical category. In Classical Arabic, the dual is a category very consistently expressed for all occurrences of 'twoness', be it natural or simple, and for all morphological classes. Any two items must be referred to by a dual ending suffixed to the term, verb, noun, adjective, or pronoun, and the morphology contains an almost complete set of dual forms for different morphological classes. In Arabic dialects – as in some other Semitic languages – the dual is considerably weakened. First, only the noun can be marked for dual, whereas verbs, adjectives, and other terms referring to a dual concept cannot be provided with a dual ending. Second, the use of dual as a cognitive category is reduced. Here, geographic and sociogeographical differences are at play. In Bedouin dialects, an extensive use of the dual as indicating any kind of twoness, simple or natural, may still occur. In Eastern urban dialects, for instance in Damascus, the simple dual does occur, but it is not obligatory. In Western North African urban dialects, the simple dual is not productive but is restricted to lexicalized items (often recognized by the ending *-ayn*), e.g. *saʿatäyn* 'two hours' (Caubet 1993). The natural dual, however, has survived in all dialects and social settings. Paired body parts are still regularly marked with a dual ending. The ending is always etymological -**ayn*, the Old Arabic oblique case, most often realized as *-ēn*, but sometimes, as in North African dialects, *-īn*. North African *-īn* is in opposition to lexicalized simple dual in -**ayn*, and similar morphological splits between simple dual and natural dual occur in single areas in the east as well.

The precise number of paired body parts marked morphologically for dual varies from dialect to dialect, from two up to about a dozen.

For some of these paired body parts, the dual ending -**ayn* is used regardless of whether it is thought of as a dual or as a plural, or, rather, it should be characterized as a plural rather than a dual. For these so-called pseudoduals, the actual information differs somewhat from dialect to dialect, or even between scholarly descriptions of the same dialect.

The most common paired body parts providing pseudoduals are probably those for 'eyes', 'ears', 'hands', and 'feet/legs'. These are, for instance, the ones given by Cowell (1964:367) for Damascus: *ʾəžrēn* 'feet/legs', *ʾīdēn* 'hands', *ʿēnēn* 'eyes', and *ʾadanēn* 'ears'; and for Morocco by Caubet (1993): sg. *ʿīn*, pl. *ʿīnīn* 'eyes'; sg. *yidd,* pl. *yiddīn* 'hands'; sg. *wdən*, pl. *wədnīn* 'ears'; sg. *ržel*, pl. *rəžlīn* 'feet'. However, the set may be smaller, as well as larger. Thus, Grotzfeld (1965) points out that the forms *rəžlēn* (from sg. *ʾəžr* ~ *rəžl*) and *ʾīdēn* are the ones typically used as plurals in Damascus. For Tunis, on the other hand, Singer (1984:453–454) reports a larger set of pseudoduals: *ʿaynīn* 'eyes', *wuḏnīn* 'ears', *korʿīn* 'paws', *ṣubʿīn* 'nails', *yiddīn* ~ *īdīn* 'hands', *fummīn* lit. 'mouths', *sāqīn* 'legs', *rukubtīn* 'knees', *xuddīn* 'cheeks', *sennīn* 'teeth, rows of teeth', *ḏrā̆ʿīn* 'arms', *karrōztīn* 'testicles', *wā̆ldīn* 'parents' (the last one is common all over the Arabic world; see Blanc 1970:47). There are areas with a much wider range of pseudoduals, e.g. Malta and Djidjelli (eastern Algeria); for the latter, the following body parts have been listed, in addition to the ones mentioned above: wings, shoulders, thighs, elbows, nostrils, eyebrows, fingers, testicles, palms, horns, flanks, cheekbones, udders, guts, molars, and incisors (Marçais, after Blanc 1970:47).

Pseudodual is usually treated in connection with three other grammatical features, which are considered to be related to it:

i. Parallel forms with a feminine ending *-t* affixed

When a real dual is meant, and not a plural, a special form with *-t* may have come into existence for those concepts that may form pseudoduals: *ʾəžərtēn* 'a pair of feet/legs', *ʾittēn* 'a pair of hands', *ʿēntēn* 'a pair of eyes', *ʾədəntēn* 'a pair of ears' (Cowell 1964, for Damascus). The same forms for Damascus are given by Grotzfeld (1965), here with suffix: *ʾīdtēni* 'my hands', *ʾəžərtēni* 'my legs', *ʿēntēni* 'my eyes', *ʾədntēni* 'ears' (all terms obviously designating pairs). The *t*-forms can be found all over the Syro-Mesopotamian area, e.g. in Palestine, Mosul, Anatolia, Aleppo, and Baghdad, Bedouin dialects excluded (Blanc 1970:47, n. 22). The *-t* forms are later than the pseudodual, because they have the new ending with *-n* and are in opposition to the pseudodual.

ii. Fluctuating principles of agreement

A pseudodual noun takes an attribute or a predicate in either the feminine singular or the 'broken' plural: *ʾīdēn kbīra* or *ʾīdēn kubār* 'big hands'. This may be compared to the common dual ('true dual'), which normally takes an attribute or a predicate in the plural: *bētēn kubār* 'two big houses'. Blanc (1970) devotes the most elaborate and speculative part of his article to the problem of concordance. Having gone through extensive material, mostly for the Cairo dialect, he is able to state that the norm is plural concordance for dual nouns (*bītēn kubār* 'two big houses'), as against either plural or feminine singular concordance for plural (*widān ṭawīla* or *ṭuwāl* 'long ears'), and either plural or feminine singular concordance for pseudoduals (*riglēn simīna* or *sumān* 'fat legs'). Informants from Palestine, Damascus, and Baghdad confirm this picture, as does textual evidence from Malta and Bukhara (Blanc 1970:51ff.). The examples found indicate that the plural concordance is chosen in connection with enumerative constructions, where the concept of plurality is more clearly felt. This would be the ordinary case with true duals, while pseudoduals may be conceived of as enumerative or nonenumerative.

iii. Dual endings -**ay* before pronominal suffixes, with the *-n* dropped

Those **ay*- forms, which correspond to Old Arabic *status pronominalis*, are a conservative feature, a residue from older language forms. The ordinary dual ending in the dialects is with *-n*, corresponding to the old *status absolutus*. For Damascus, the following four *-ē* forms are given: *ʾīdēk* 'your hands', *ʾəžrēk* 'your feet/legs', *ʿēnēk* 'your eyes', *ʾādānēk* 'your ears' (Grotzfeld 1965; Cowell 1964). According to Cowell (1964), these terms are identical with the ones occurring as pseudoduals. For Tunis, Singer

> (1984) lists the following as pseudoduals connecting pronominal suffixes with *-ī* forms: *ʿaynīk* 'your eyes', *wuḏnīk* 'your ears', *xuddīk* 'your cheeks', *sennīk* 'your teeth', *yiddīk* ~ *īdīk* 'your hands', *sāgīk* 'your legs', *wǟldīk* 'your parents', but the following terms, which according to Singer are also pseudoduals, cannot be used with *-ī* forms: *drǟʿīn* 'arms', *korʿīn* 'feet', *rukubtīn* 'knees'. In South Tunisian Bedouin dialects, forms like *rukbīk* may occur. These *-*ay* forms are recorded from many dialects. See also Grotzfeld (1965), where two forms are defined as pseudoduals, as against four occurring with *-ē* forms.

Although the dual nouns in question are given with plural suffixes in Blanc and Singer (e.g. *riglēhum* 'their (many) feet'), it is obvious that they usually occur with singular suffixes and in reality denote natural dual, one set of a paired body part. A paired body part is usually provided with a pronominal suffix, which acts as a determiner. The feature of short pronominal forms for certain paired body parts should thus be distinguished from the feature of pseudodual. Both are parallel developments of natural dual. They both demonstrate the high frequency of dual with body parts. The *-*ay* forms do not necessarily or mechanically coincide with the pseudodual forms, and they are not a secondary development dependent on the existence of pseudoduals.

The textual evidence demonstrates that the dual ending with paired body parts had a plural function already in early Arabic texts (9th-century Christian texts), clearly so in Andalusian Arabic (13th century), and that the *status pronominalis* forms were, as could be expected, of the old *-*ay* type. The agreement situation, however, is not the same for modern dialects. The medieval texts sometimes reflect dialectal usage and disregard the use of Classical Arabic dual forms for adjectives and verbs, but when this is the case, both plural and feminine singular agreement occur, regardless of the type of dual represented (Blanc 1970:54ff.)

Pseudoduals, or similar phenomena, are known in other Semitic languages (for the following, see Lipiński [1997:237ff.], unless otherwise stated). In Modern South Arabian, original plurals may be used as duals, e.g. *naṣfi* 'halves', from sg. *naṣf*. With paired body parts, dual endings used to designate plural are first known from Assyro-Babylonian, e.g. *erba šēpāšu* 'four are his feet' (von Soden 1995:93, here called *pluralis paucitatis*). The form *qatā(n)* came to mean 'hands [two or more]'. From Biblical Hebrew, there are examples like *bəkappēhäm* 'in their hands' (Blau 1976:66); *šeš kənāpayim* 'six wings'. In Punic, *pʾmm* refers to the paws of sacrificial animals. In all these three ancient languages, regular plural forms were used to designate the same words, which then acquired other lexical meanings.

With regard to the development of the dual, the pseudodual seems to be one of a group of features defining an intermediate stage in Semitic, also including common usage of natural dual, and a more or less rare or uncertain use of simple dual. To this stage belong modern Arabic dialects, Biblical Hebrew, and middle and late forms of Akkadian. Akkadian made extensive use of the dual in its early stages but replaced it gradually by plural, already from the Old Babylonian period onward. Biblical Hebrew used natural dual for body parts and a few other concepts, but only rarely for simple dual.

A difference with regard to the cognitive status of natural and simple dual, respectively, can be traced in the general development of Semitic and Arabic but is particularly well articulated in the Arabic dialects. The cognitive differentiation between natural dual and simple dual is reflected morphologically. Pseudodual, together with the related or closely concurrent features of the residual *status pronominalis* for similar concepts, and the reestablishment of a 'true' dual with a *-t* after the emergence of pseudodual, is an indication of the strength of the natural dual as a cognitive concept as opposed to the historically unstable enumerative or simple dual. This state of affairs is a synchronic one but may throw light on the (early) development of the dual in Arabic and Semitic. It is probably an ongoing process as well, and further stages of development may be expected.

Bibliographical references

Blau, Joshua. 1976. *A grammar of Biblical Hebrew*. Wiesbaden: O. Harrassowitz.

Blanc, Haim. 1970. "Dual and pseudo-dual in the Arabic dialects". *Language* 36.42–57.

Caubet, Dominique. 1993. *L'arabe marocain*. I. *Phonologie et morphosyntaxe*. Paris and Louvain: Peeters.

Cowell, Mark W. 1964. *A reference grammar of Syrian Arabic (based on the dialect of Damascus)*. Washington, D.C.: Georgetown University Press.
Fontinoy, Charles. 1969. *Le duel dans les langues sémitiques*. Louvain: Les Belles Lettres.
Grotzfeld, Heinz. 1965. *Syrisch-Arabische Grammatik*. Wiesbaden: O. Harrassowitz.
Lipiński, Edward. 1997. *Semitic languages: Outline of a comparative grammar*. Leuven: Peeters.
Singer, Hans-Rudolf. 1984. *Grammatik der arabischen Mundart der Medina von Tunis*. Berlin and New York: de Gruyter.
Soden, Wolfram von. 1995. *Grundriss der akkadischen Grammatik*. Rome: Pontificium Institutum Biblicum. (1st ed. 1952.)

Kerstin Eksell (University of Copenhagen)

Pseudoverb

In all varieties of Arabic, there is in general a clear distinction between verbs and other parts of speech. Morphologically, only verbs have an opposition between perfect and imperfect, with subject-verb agreement in person-gender-number shown by means of suffixes in the perfect and a combination of prefixes and suffixes in the imperfect. Even when a verb lacks certain forms, as when Classical/Modern Standard Arabic *laysa* has no imperfect forms (→ defective verb), the existing forms can be readily assigned to a position in the verbal paradigm. There are also more subtle differences, as when a 1st person singular object or possessor is shown by means of the suffix *-nī* (and its reflexes) with verbs, but *-ī/-ya* (and its reflexes) with prepositions and nouns. Finally, there are syntactic differences, perhaps most clearly in the case of negation, where both Classical/Modern Standard Arabic and the vernaculars distinguish negation of verbs from negation of other predicates: Classical/Modern Standard Arabic *lā/lam/lan* vs. *laysa*, Egyptian Arabic *mā…ši* vs. *miš*, Syrian Arabic *mā* vs. *mū*, etc. Nonetheless, there are some items that, though lacking the basic morphological properties of verbs, share at least some of their other properties; these items are known as pseudoverbs (thus Brustad 2000:151–157, although there are terminological variants, e.g. quasi-verbs, Cowell 1964:412–416).

Many pseudoverbs are originally nouns, prepositions, or adverbs, such as Syrian *bədd* 'to want' (originally 'requirement'), *ʿand* 'to have' (originally, and still, in other contexts, 'at'), *fī* 'to be able to' (originally, and still, in other contexts, 'in'), Maltese *hemm* 'there is' (originally, and still, in other contexts, 'there'). While most of these items retain an internal morphology that is completely nonverbal, there are occasional exceptions, so that the pseudoverb *fī* in Syrian usually has a 1st person form *fī-ni*, with the verbal suffix, whereas as a preposition it takes the nonverbal suffix: *fiy-yi* 'in me' (Cowell 1964:479). Usually, however, the most decisive reason for analyzing these items as being at least partially verbal is that they negate like verbs, either obligatorily or at least preferably, depending on the item and the particular vernacular under consideration. Thus, in Syrian Arabic, we find *mā bədd-ak* Neg want-2ms 'you do not want', *mā fī-hon* Neg in-3p 'they cannot' (Cowell 1964:412–416), and in Maltese, *m' hemm-x* Neg there-Neg 'there is not' (Comrie 1982; Borg and Azzopardi-Alexander 1997:89). (In applying negation as a criterion, it is important to test whether or not the variety of Arabic in question has extended its verbal negation more generally to nonverbal elements; if so, then negation cannot be used, at least not without great care, as a test for pseudoverbs; see the discussion in Brustad 2000:291–294.) Negation sometimes serves to distinguish between pseudoverb and other uses of the same item, as in Syrian Arabic *mā ʿand-i* 'I do not have' vs. *mū ʿand-i* 'is not at my place'.

The construction illustrated in Syrian Arabic, *ṣəhr-ak ʿand-o ržāl* 'your son-in-law has some men' poses interesting questions concerning the argument structure of the pseudopredicate. In the Classical Arabic ancestor of the construction, *ṣəhr-ak* would be considered a preposed topic, with the suffix on the preposition a resumptive pronoun and *ržāl* the subject of the clause, but there is evidence from different vernaculars of differing degrees of reinterpretation of this construction in the direction of subject-verb-object. In Syrian Arabic, for instance, the position of *ržāl* can be occupied by a direct object pronoun, as in *ʿand-i yā-ha* 'I have it [fem.]', *bədd-o yā-kon* 'he wants you [pl.]' (Cowell 1964:545). In many vernaculars, the item in this position is incapable of triggering subject-verb agreement on an auxiliary marking temporal reference, e.g. Syrian Arabic *kān ʿan-na ḍyūf* 'we had guests' does not allow the 3rd person plural *kān-u*. (Tunis Arabic, by contrast, does allow the auxiliary to agree with a 3rd person singular feminine possessum, but not with a plural possessum, even if feminine; cf. Comrie 1986.) Variation between different pseudoverbs can be seen in the fact that Syrian Arabic allows agreement between the auxiliary in this construction and the wisher with *bədd* 'to want', but not with the possessor with *ʿand* 'to have', e.g. *kān/kən-na bədd-na* 'we wanted', but *kān/*kənna ʿan-na* 'we had' (Cowell 1964:414).

In addition to nonverbs that have acquired some verbal properties, the set of pseudoverbs can also be taken to include original verbs that have lost some verbal properties. For instance, in several vernaculars, an earlier verb *ʿād* 'to return' has become an invariable particle meaning 'still, again, then', but still with verbal negation (Brustad 2000:160–161); in Syrian Arabic, *māʿād* 'no longer' not only still shows verbal negation but also permits optional subject-verb agreement, especially with 1st and 2nd person subjects (Brustad 2000:161).

One might also subsume under this category the use of personal pronouns as copulas with present time reference (Brustad 2000:157–158), especially in vernaculars where such pronouns take verbal negation to express a negative copula, e.g. Moroccan Arabic *m-anī-š mən-(h)na* 'I am not from here' (Brustad 2000:296–301). In most of the varieties considered by Brustad, this negative copula is pragmatically marked, but in Maltese, a construction like *intom m' intom-x magħ-na* 'you are not with us' is neutral, though existing alongside an alternative with invariable *mhux* (etymologically *m' hu-x*, i.e. negative of the 3rd person singular masculine pronoun) marking nonverbal negation: *intom mhux magħ-na* (Borg and Azzopardi-Alexander 1997:52–53). Finally, in some varieties, some participles used predicatively have taken on verbal characteristics, such as negation in Kuwaiti Arabic *mā gādir marra waḥda* 'I couldn't all of a sudden', lit. 'not being.able time one' (Brustad 2000:289–291), even to the extent of taking 2nd person singular feminine verbal inflection before a suffixed object in Syrian Arabic, e.g. *ʾənti kātəb-tī* 'are you the one who wrote it?', cf. verbal *katab-tī* 'you wrote it' (Cowell 1964:268–269).

Pseudoverbs in Arabic vernaculars provide interesting data for diachronic studies of word classes, with items intermediate between verb and nonverb arising as nonverbs are used verbally and acquire verbal characteristics or, less frequently, as verbs lose verbal characteristics as they are used nonverbally.

BIBLIOGRAPHICAL REFERENCES

Borg, Albert and Marie Azzopardi-Alexander. 1997. *Maltese*. London: Routledge.

Brustad, Kristen E. 2000. *The syntax of spoken Arabic*. Washington, D.C.: Georgetown University Press.

Comrie, Bernard. 1982. "Syntactic-morphological discrepancies in Maltese sentence structure". *Communication and Cognition* 15:281–306.

——. 1986. "The possessive predicate in North African vernacular Arabic". *The Fergusonian impact. I. From phonology to society*, ed. Joshua A. Fishman, Andrée Tabouret-Keller, Michael Clyne, Bh. Krishnamurti, and Mohamed Adulaziz, 197–210. Berlin: Mouton de Gruyter.

Cowell, Mark W. 1964. *A reference grammar of Syrian Arabic*. Washington D.C.: Georgetown University Press.

Eid, Mushira. 1991. "Verbless sentences in Arabic and Hebrew". *Perspectives on Arabic linguistics*, III, ed. Bernard Comrie and Mushira Eid, 31–61. Amsterdam and Philadelphia: J. Benjamins.

BERNARD COMRIE (Max Planck Institute for Evolutionary Anthropology)

Punctuation

Punctuation (*tarqīm*) is an important and cohesive device in written texts. Writers use punctuation to separate groups of words for meaning and emphasis; to signal the beginning or end of sentences, phrases or clauses; and to help avoid contextual ambiguity.

Many written languages use punctuation marks, but their function differs across languages. In the case of the Arabic language, punctuation marks (*ʿalāmāt at-tarqīm*) are added to the Arabic script. Hence, punctuation in Arabic texts is constrained by the characteristics of Arabic script (e.g. the absence of capital letters). The use of punctuation marks is a modern invention in Arabic writing; Classical Arabic used other means to set off or highlight parts of the text, for instance by the use of a stroke over titles of sections in manuscripts, while Qurʾānic manuscripts use a set of signs to indicate pauses in the recitation (*ʿalāmāt al-waqf*).

A short guide for the use of punctuation in → Media Arabic is given in the *ʾAhrām style book* (Nāfiʿ n.d.:42–47). For a fuller treatment, see Badawi a.o. (2004:21–25), who point out that punctuation marks tend to be used more frequently in Arabic writing than in Western writings (see also Khafaji 2001).

1. Punctuation marks in Arabic

Most Arabic punctuation marks have their formal equivalent marks in the English punctuation system. Examples of these marks:

i. The full stop or period (*nuqṭa*) terminates sentences that are neither interrogatory nor exclamatory; it is also used for making acronyms and abbreviations in Arabic.

ii. The comma (*fāṣila*) in Arabic has the form of an inverted *wāw* ، ; it signals a short pause and is used for separating words, phrases, or clauses in series; sometimes in Arabic, it is also used for delimiting sentences, and in such cases, the full stop functions as an end-of-paragraph marker; a double comma before and after a phrase introduces supplementary or explanatory material.

iii. The semicolon (*fāṣila manqūṣa*) in Arabic has the form ؛; it introduces a pause longer than a comma; the semicolon separates two clauses but at the same time indicates a linkage between the two; this linkage relation between the two clauses can be contrastive, additive, or emphatic. An example is given in the *ʾAhrām style book* (Nāfiʿ n.d.:44): عدت إلى المعاجم القديمة؛ لأتحقق من صحة النص. ‘I went back to the ancient sources in order to verify the correctness of the text’.

iv. The colon (*nuqṭatāni raʾsiyyatāni*) introduces a clause or phrase that explains, illustrates, amplifies, or restates what comes before it; colons can introduce direct quotations as well.

v. The question mark (*ʿalāmat al-istifhām*) in Arabic printing has the form ؟ ; it terminates a direct question; it tends to be used also after indirect questions (Badawi a.o. 2004:24).

vi. The exclamation mark (*ʿalāmat at-taʿajjub*) terminates an emphatic phrase or sentence that conveys happiness, sadness, surprise, or a prayer.

vii. Double quotation marks (*ʿalāmāt al-istišhād*) in Arabic usually have the form <<…>>; they enclose direct quotations; they are also used for highlighting a word or phrase in written texts; some writers use this mark to introduce proper nouns.

viii. The dash (*šarṭa* or *šarṭa ʾafaqiyya*) serves as equivalent of the phrase ‘from, up to’ when used between numbers (e.g. ١٦–٢٣ ‘16–23’); it is also used in enumerations after the numeral, and in the representation of conversations (Nāfiʿ n.d.:44). In literary texts it may also be used as a separator for adverbs like *ʾayḍan* and *kaḏālika* ‘also’ or *maṯalan* ‘for example’ (Badawi a.o. 2004:23).

ix. Double dash marks (*šarṭatāni*) enclose supplementary or explanatory material; the inclusion of such material marked by double dashes does not essentially alter the meaning of the sentence, for example: والدي – رحمه الله – كان تقيا – ‘my father – God rest his soul – was a God-fearing man’.

x. Dots (…, the ellipsis mark; *aṯ-ṯalāṯ nuqāṭ*) indicate the omission of one or more words in the sentence or the omission of one or more sentences. In the *ʾAhrām style book*, the use of the three dots for purposes other than the indication of an ellipsis is explicitly disapproved (Nāfiʿ n.d.:47), but in literary texts three (or sometimes more) dots are used quite frequently, sometimes even after the conjunction *wa-* ‘and’ to indicate a certain suspense (Badawi a.o. 2004:23). The use of two dots (*nuqṭatāni mutajāwiratayni*) in newspapers is allowed only for titles consisting of two parts, e.g. حما ية المستهلك ... وحماية المنتج ‘protection of the consumer; protection of the producer’ (Nāfiʿ n.d.:47).

xi. The diagonal line / (*aš-šarṭa al-māʾila*) is used for separating hour and minute in expressions for time (e.g. ٨/٣٠ 8/30 for half past 8); it is also used for separating day, month, and year in expressions for dates (e.g. ٢/١/٢٠٠١ 2/1/2001 for the second day of January 2001). In reference to Qurʾānic verses, the diagonal mark separates the number of the verse and the sura/chapter number in which the verse appears (e.g. ١٤/١٣ 14/13 for sura 14, verse 13); the diagonal also functions as the decimal point in Arabic numbers (e.g. ١/٥٠ 1/50 in Arabic represents one and a half). The diagonal mark also separates alternatives or combined notions, e.g. ‘the Egyptian-Moroccan treaty’ الاتقاق المصري/المفربي, and in physics notations like ‘kilowatt/hour’ كيلووات/ساعة. In imitation of the English use of *and/or*, it is also used in the expression و/أو *wa-/ʾaw*.

xii. Parentheses ( ) (*qawsāni*) introduce supplementary, parenthetic, or explanatory material into the sentence, and the introduced material complements the information of the sentence; parentheses are also used for citations both within and outside a sentence; references to publication sources appear in parentheses.

xiii. Brackets [ ] (*qawsāni murabbaʿāni* or *qawsāni bi-zawāyā qāʾima*) are used for inserting material that is not part of the original text in such cases as editorial interpolations.

2. Capitalization and abbreviations in Arabic

Arabic script does not use capitalization for marking proper nouns or marking the beginning of sentences. In contrast, capitalization in English is used for marking proper nouns (e.g. ‘Lebanon’), as well as the beginning of sentences and acronyms (e.g. ‘BBC’). English texts distinguish between capital and noncapital forms.

In Arabic script, proper nouns are not marked. Some Arabic newspapers use double quotations or parentheses for marking proper nouns, but this is not a general practice.

Arabic writers utilize different means for marking → abbreviations in Arabic script. One method is to use the dot notation (e.g. *bī.bī.sī.* بی . بی . سی . for ‘BBC’). For foreign abbreviations like ‘BBC’, the Latin letters are transcribed into their Arabic form and are separated from each other by period or space.

In some cases, no space or period appears in between, and because of the special characteristics of Arabic script, the spaces between letters are understood. This is possible because Arabic script distinguishes between the initial, medial, final, and isolated forms for letters of the alphabet. In *bībīsī* بي بي سي (BBC), for example, all the *ī* letters are in final forms and are not joined with the following letters. If *bībīsī* were not an abbreviation, the *ī* letters would have been joined with the next letters and would have appeared in their medial form.

In this example, the appearance of *ī* ی in its final form instead of the medial introduces a form of semispace in Arabic script, which may signal the presence of an abbreviation. This special form of capitalization in Arabic script by switching among initial/medial/final forms has been traditionally used as one of the mechanisms for introducing some of the abbreviations in Arabic, e.g. ع م (*ʿalayhi s-salām* ‘peace upon him’), ق ل (*al-Qilyūbī* in aš-Šāfiʿī’s book), ح ل (*al-Ḥalabī* in aš-Šāfiʿī’s book). There are also cases such as الشـ (*aš-šāriḥ* ‘the commentator’) and اهـ (*intahā* ‘it is finished’), in which the abbreviations end in letters with medial forms instead of final forms.

The other mechanism for forming abbreviations in Arabic is to join the letters of the abbreviation similar to a normal word, rather than switching between the different forms.

In general, in Arabic script, because of the possibility of final forms of letters that do not attach to the next letter, the use of space for separating words can be relaxed without affecting the readability of Arabic texts. This characteristic of the Arabic script does not exist in Latin scripts.

Bibliographical references

Badawi, Elsaid, Michael G. Carter, and Adrian Gully. 2004. *Modern Written Arabic: A comprehensive grammar.* New York: Routledge.

Khafaji, Rasoul. 2001. “Punctuation marks in original Arabic texts”. *Zeitschrift für Arabische Linguistik* 40.7–24.

Nāfiʿ, ʾAḥmad. N.d. *Dalīl al-ʿamal aṣ-ṣuḥufī/al-Ahram style book.* [Cairo: Al-Ahram.]

Siamak Rezaei (UNBC)

General Editor
Kees Versteegh
(University of Nijmegen)

Associate Editors
Mushira Eid
(University of Utah)
Alaa Elgibali
(University of Maryland)
Manfred Woidich
(University of Amsterdam)
Andrzej Zaborski
(University of Cracow)

Advisory Board
Ramzi Baalbaki *(American University of Beirut)*
Elsaid Badawi *(American University of Cairo)*
Dominique Caubet *(INALCO, France)*
Clive Holes *(University of Oxford)*
Manfred Kropp *(Orient-Institut Der Deutschen Morgenländischen Gesellschaft)*
Jérôme Lentin *(Institut National des Langues et Civilisations Orientales)*
John McCarthy *(University of Massachusetts)*
Jamal Ouhalla *(University College Dublin)*
Jan Retsö *(Göteborg University)*
Sabah Safi *(King Abdulaziz University)*

Copy Editor
Carolyn Russ

ENCYCLOPEDIA OF ARABIC LANGUAGE AND LINGUISTICS

VOLUME IV
Q-Z

General Editor
Kees Versteegh

Associate Editors
Mushira Eid
Alaa Elgibali
Manfred Woidich
Andrzej Zaborski

BRILL
Leiden – Boston
2009

Cover Illustration:
Fragment from the *Qurʾān* (Q. 110), attributed to Omar ibn Said (ca. 1770-1864), a Muslim from Futa Toro (Senegal), who was sold into slavery to North America. Used with kind permission of: Documenting the American South, The University of North Carolina at Chapel Hill Libraries.

Despite our efforts we have not been able to trace all rights holders to some copyrighted material. The publisher welcomes communications from copyrights holders, so that the appropriate acknowledgements can be made in future editions, and to settle other permission matters.

© Copyright 2009 by Koninklijke Brill NV, Leiden, The Netherlands
Koninklijke Brill NV incorporates the imprints Brill, Hotei Publishing, IDC Publishers, Martinus Nijhoff Publishers and VSP

All rights reserved. No part of this publication may be re-produced, translated, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without prior written permission from the publisher.

Authorization to photocopy items for internal or personal use is granted by Brill provided that the appropriate fees are paid directly to The Copyright Clearance Center, 222 Rosewood Drive, Suite 910, Danvers MA 01923, USA. Fees are subject to change.

PRINTED IN THE NETHERLANDS

ISBN: 978 90 04 14973 1 (Set)
ISBN: 978 90 04 14476 7 (Volume 4)

This book is printed on acid-free paper.

Cover design: BEELDVORM, Pijnacker

List of Contributors

Kamal Abu Deeb, School of Oriental and African Studies, London, U.K.
Al-Amin Abu-Manga, University of Khartoum, Sudan
Amira Agameya, Cairo University & The American University in Cairo, Egypt
Joyce Åkesson, Lund, Sweden
Mohammad T. Alhawary, University of Oklahoma, U.S.A.
Abdullah Hassan al-Saqqaf, Sultan Qaboos University, Oman
Enam Al-Wer, University of Essex, U.K.
Nadia Anghelescu, University of Bucharest, Romania
Joseph Aoun, Northeastern University, U.S.A.
Georgine Ayoub, Institut des Langues et Civilisations Orientales, Paris, France
Ramzi Baalbaki, American University of Beirut, Lebanon
Taïeb Baccouche, University of Tunis, Tunisia
Muhammad Hasan Bakalla, King Saud University, Saudi Arabia
Sergio Baldi, University of Naples, L'Orientale, Italy
Melissa Barkat, University of Lyon-2, France
Reem Bassiouney, The Georgetown University, Washington, D.C., U.S.A.
Wafaa Batran Wahba, Women's College, Ain Shams University, Egypt
Peter Behnstedt, Chipiona, Spain
Anna Gr. Belova, Moscow, Russia
Thami Benkirane, University of Fes, Morocco
Elabbas Benmamoun, University of Illinois at Urbana-Champaign, U.S.A.
Václav Blažek, Masaryk University of Brno, Czech Republic
Gert Borg, University of Nijmegen, the Netherlands
Michael G. Carter, University of Sydney, Australia
Ferid Chekili, University of Manouba, Tunisia
Lina Choueiri, American University of Beirut, Lebanon
Sven-Olof Dahlgren, Göteborg University, Sweden
Janusz Danecki, University of Warsaw, Poland
Stuart Davis, Indiana University, U.S.A.
Alain Désoulières, INALCO Urdu/South Asia and CERLOM
Joseph Dichy, Université Lumière-Lyon-2, France
James Dickins, University of Salford, United Kingdom
Anthony Diller, Australian Academy of the Humanities, Australia
Lutz Edzard, University of Oslo, Norway
John Eisele, The College of William and Mary, U.S.A.
Dina El Zarka, University of Graz, Austria
Moha Ennaji, Rutgers University, U.S.A.
Mohssen Esseesy, The George Washington University, U.S.A.
Said Faiq, American University of Sharjah, United Arab Emirates
Ali Farghaly, Oracle USA Corporation, U.S.A.
Daniela Rodica Firanescu, Dalhousie University, Canada
Stefan A. Frisch, University of South Florida, U.S.A.
Adam Gacek, McGill University, Canada
Adamantios I. Gafos, New York University, U.S.A.
Maik Gibson, Nairobi, Kenya
Jacques Grand'Henry, Catholic University of Leuven, Belgium

Muhammad Abdel Haleem, School of Oriental and African Studies, London, U.K.
Peter Hallman, McGill University, Canada
Jaakko Hämeen-Anttila, University of Helsinki, Finland
Jidda Hassan, Maiduguri, Nigeria
Bernadet Hendriks, Holy Land Institute for the Deaf, Jordan
Robert D. Hoberman, Stony Brook University, U.S.A.
Iris Hoedemaekers, University of Nijmegen, the Netherlands
Uri Horesh, Franklin and Marshall College, U.S.A.
Frederick M. Hoyt, University of Texas at Austin, U.S.A.
Bruce Ingham, School of Oriental and African Studies, London, U.K.
Maher Jesry, Kind Saud University, Saudi Arabia
Rudolf de Jong, University of Amsterdam, the Netherlands
René Kager, University of Utrecht, the Netherlands
Wilaiwan Khanittanan, Thammasat University, Thailand
Mark Kirchner, Giessen, Germany
Ernst Axel Knauf, University of Bern, Switzerland
Joost Kremers, University of Frankfurt, Germany
Wouter Kusters, Meertens Instituut, the Netherlands
Pierre Larcher, Université de Provence, France
Xavier Lelubre, Université Lumière-Lyon-2, France
Mark S. LeTourneau, Weber State University, U.S.A.
Antoine Lonnet, C.N.R.S., Paris, France
Heidi Lorimor, Bucknell University, U.S.A.
Zouhair Maalej, University of Manouba-Tunis, Tunisia
Soheir G. Mahfouz, Ain Shams University, Egypt
Miklós Maróth, Pázmány Péter Catholic University, Hungary
Fiona McLaughlin, Gainesville, U.S.A.
Alex Metcalfe, Lancaster University, U.K.
Philip Monahan, University of California at San Diego, U.S.A.
Mustafa Mughazy, Western Michigan University, U.S.A.
Mohamed Haji Mukhtar, Savannah State University, U.S.A.
Kristina Nelson, Cairo, Egypt
Fallou Ngom, Boston University, U.S.A.
Helle Lykke Nielsen, University of Southern Denmark, Denmark
Hans Nugteren, Dordrecht, the Netherlands
Isaac A. Ogunbiyi, Lagos State University, Nigeria
Jonathan Owens, University of Bayreuth, Germany & University of Maryland, U.S.A.
Dilworth B. Parkinson, Brigham Young University, U.S.A.
Yishai Peled, Tel Aviv University, Israel
Christophe Pereira, LACNAD-CRÉAM –INALCO, France
John R. Perry, University of Chicago, U.S.A.
Stephan Procházka, University of Vienna, Austria
Philipp Reichmuth, University of Halle, Germany
Stefan Reichmuth, Ruhr-Universität Bochum, Germany
Judith Rosenhouse, Swantech-Sound Waves Analysis and Technologies Ltd., Haifa, Israel
Valeriy Rybalkin, Institute of Oriental Studies, National Academy of Sciences of Ukraine
Sabah M.Z. Safi, King Abdulaziz University, Saudi Arabia
Mustafa A. Shah, School of Oriental and African Studies, London, U.K.
Sonia Shiri, University of California, Berkeley, U.S.A.
Avihai Shivtiel, University of Leeds, U.K.
Joan Smith Kocamahhul, University of Canterbury, New Zealand
Usama Soltan, Middlebury College, U.S.A.
V. Swarajya Lakshmi, Osmania University, India

Zeinab Taha, The American University in Cairo, Egypt
Ahmad Taher Hassanein, Misr University for Sciences and Technology, Egypt
David Teeple, University of California, Santa Cruz, U.S.A.
Torsten Tschacher, University of Heidelberg, Germany
Uwe Vagelpohl, Hampshire College, U.S.A.
Martine Vanhove, LLACAN – CNRS, INALCO, Université Paris-7, France
Kees Versteegh, University of Nijmegen, the Netherlands
Ángeles Vicente, University of Zaragoza, Spain
Rainer Voigt, Free University of Berlin, Germany
Janet C.E. Watson, University of Salford, U.K.
Bernard Weiss, University of Utah, U.S.A.
Suzanne Wertheim, Northwestern University, U.S.A.
Stefan Wild, University of Bonn, Germany
Abderrahim Youssi, Mohamed V University, Morocco
Andrzej Zaborski, University of Cracow, Poland
Martin R. Zammit, University of Malta, Malta
Petr Zemánek, Charles University, Prague, Czech Republic
Chakir Zeroual, Université Sidi Mohamed Ben-Abdellah, Morocco & CNRS, France
Ulrike Zeshan, International Centre for Sign Languages and Deaf Studies, University of Central Lancashire, Preston, U.K.
Gerit Zimmermann, Bayreuth, Germany

Note from the General Editor

With the publication of its fourth volume, the Encyclopedia of Arabic Language and Linguistics (EALL) is now complete; a comprehensive index to the entire encyclopedia will be published separately. The editors are grateful to all authors who contributed to the encyclopedia. They also wish to express their thanks to the copy-editors, Margaret Owen (vol. I) and Carolyn Russ (vols. II–IV), for their meticulous and painstaking correction of the text. The publication of this encyclopedia would have been impossible without the generous help of the Middle Eastern staff at Brill's. Throughout the project, Ingrid Heijckers has been a pillar of strength. She coordinated the project from the beginning in the most efficient way possible, and whenever there were complications, she could always be counted upon to come up with a creative solution.

Obviously, in a work of this scope, mistakes and omissions are inevitable. All users are invited to send their corrections and suggestions to the publisher. There is one mistake, however, for which we should like to make amends here: in volume III, the article on Omani Arabic by Clive Holes has inadvertently been ascribed to Lutz Edzard. We deeply regret this error and should like to offer our apologies both to Prof. Holes and to Dr. Edzard.

The *Encyclopedia of Arabic Language and Linguistics Online* will be launched in 2009. The online edition will contain all content of the print edition and new content will be added on a regular basis. New articles will be elaborations or updates of themes already discussed in the EALL, or will be new entries that are relevant to the field. Just like the print edition, the EALL Online will comprehensively cover all aspects of Arabic languages and linguistics. It will be interdisciplinary in scope and will represent different schools and approaches in order to be as objective and versatile as possible. The online edition will be cross-searchable, cross-referenced and regularly updated.

Kees Versteegh

Q

Qāf

1. *Qāf* in Arabic and Semitic

Qāf is the name of the 21st letter of the Arabic alphabet. In surveys of Modern Standard Arabic, /q/ is regularly described as a voiceless velar or uvular plosive/stop (cf., e.g., Kästner 1981:45–46). While this description reflects the pronunciation in Modern Standard Arabic and in some dialects, it is very likely that the uvular stop regularly transcribed as *q* in Arabic linguistics was a nonemphatic (nonvelarized) voiced counterpart to *k* (IPA [ɢ]) (cf., e.g., Versteegh 2001:21). As one of the *ḥurūf ʾaqṣā l-lisān* 'the sounds at the remotest part of the tongue', Ibn Jinnī groups *q* together with *k* and *j* in the context of discussing co-occurrence restrictions within the root (cf. Fleisch 1958a; Bakalla 1982:189). In a global survey of phoneme systems in modern times, Maddieson (1984:214) lists a long voiced velar plosive /gː/ as being specific for Arabic (dialects), Somali, Punjabi, and Shilha. Taking a wider Semiticist's perspective, the variety of pronunciations of /q/ has also been attested elsewhere, both diachronically and synchronically. In the Akkadian writing system, /q/ and /g/ were not systematically distinguished (cf. von Soden 1995:34; Lipiński 2001:144–145), and there was only one sign for the CV-sequences *ag*, *ak*, and *aq*. Rather than reflecting inherent 'weaknesses' of the Sumero-Akkadian syllabary, this circumstance may well testify to an early variety of pronunciation as is also present in the modern Arabic dialects. Whereas the velar quality of Arabic /q/, as pronounced in Modern Standard Arabic, corresponds to glottalization in modern Ethio-Semitic ([kˀ]), it is also the case that Arabic and Gəʿəz *qatala* evolved to *gäddalä* in modern Ethio-Semitic (unless the latter verb is associated with the root *g-d-l*, as in Hebrew; cf. Leslau 1987:452).

/q/ has straightforward correspondences across Semitic, with the aforementioned postglottalized variant [kˀ] in modern Ethio-Semitic. Old Aramaic /ʿ/ also has a variant /q/ in some lexical items, e.g. *ʾarʿā* and *ʾarqā* 'earth'. Putting this observation into context, Lipiński (2001:147) notes that a spectographic analysis shows that [q] is situated somewhere on a scale between [d] and [ʕ]. The letter <q> is also used to render <k> in (mainly Greek) loanwords in Hebrew and Aramaic so as to avoid postvocalic spirantization, as well as in loanwords that made it into Arabic via Syriac (e.g. *qānūn* < *kanṓn*). But <q> was also used in Arabic to render a number of words written with <g> in the source language, e.g. *qibṭī* 'Copt(ic)' < *aigúptios*.

2. *Qāf* in Classical Arabic

The fact that /q/ was an unaspirated stop with both voiced and voiceless variants at an early stage of the history of Arabic can be deduced from general linguistic considerations and from information provided by the Arab grammarians (on the importance of Sībawayhi's *Kitāb* for historical Arabic dialectology in general, see Levin 1999). Blanc (1969), taking up proposals made by Jean Cantineau and André Martinet, argues in terms of a linguistic push-chain mechanism. An early Semitic /g/ as part of a homorganic velar triad of phonemes /g-k-q/, was fronted to [g^j], [j], etc., as Semitic /q/ had developed mainly in a Bedouin milieu toward [ɢ], thus creating homophones, e.g. *faqr* 'poverty' vs. *fagr* 'dawn', which were to be avoided. Thus, a dialectal split between a *qāf ġayr maʿqūda* 'non-tied *q*', corresponding to [q], and a *qāf maʿqūda* 'tied *q*', corresponding to [ɢ] and sometimes even [k], emerged (Blanc 1969:22). Blanc (1969:29–30) sketches the following three-stage scenario, reflecting the assumed push-chain process within the velar phoneme inventory:

i. Proto-Arabic and Common Semitic

	q	
ġ	*g*	–
x	*k*	*š*

ii. Earliest Arabic

ġ	*q*	*g*	–
x	–	*k*	*š*

iii. Oldest Arabic (*ʿarabiyya*)

ġ	*q*	–	g^y
x	–	*k*	*š*

From the perspective of native Arab(ic) grammar, it is noteworthy that Sībawayhi (*Kitāb* §565) classifies /q/ among the → *majhūra* consonants, which are opposed to the *mahmūsa* consonants (cf., e.g., Al-Nassir 1993:36–41; Carter 2004:126). Except for the consonants /q/ and /ṭ/ in their modern pronunciation, Sībawayhi's opposition between *majhūra* vs. *mahmūsa* corresponds exactly to the modern opposition [voiced] vs. [devoiced] (cf., e.g., Fleisch 1958b; Odisho 1988), *pace* Lipiński (2001:144), who conceptualizes *majhūra* as 'fortis' vs. 'lenis'. Following Garbell (1958), Blanc (1967:306–307) suggested the terms 'nonbreathed' for *majhūr(a)*, and 'breathed' for *mahmūs(a)*, in an attempt to reconcile the terminology in Arabic grammatical sources with modern concepts.

Regarding the phonetic quality of /q/, Bravmann (1934:45) quotes from the *Tāj al-ʿarūs* (s.v. *qāf*): *wa-hiya ʾamtanu l-ḥurūfi wa-ʾaṣaḥḥuhā jarsan* 'it [sc. *q*] is the most solid of the sounds and the most real in terms of tone', a definition which in all its vagueness is compatible with both the characterization by the Arab grammarians and the evidence in modern times, pointing to a higher sonority of [q] in comparison with [k]. Beyond the evidence adduced above, there is also circumstantial evidence in written Arabic that compels us to view the voiced pronunciation /q/ as at least one statistically significant variant in the early stages of Arabic. Brockelmann (1908:121), referring to az-Zamaxšarī's *Mufaṣṣal* (§ 695b), adduces the Old Arabic variant *zaqar* of *saqar* 'fire in hell', which can be best explained by suprasegmental assimilation with respect to voicing (*s* < *z*, in this case presupposing a voiced [q]).

While /q/ is clearly not an 'emphatic' consonant (pronounced with → *ʾiṭbāq*, i.e. [+A(dvanced)T(ongue)R(oot)]), as evidenced by Form VIII verbs like *iqtaraba* 'to draw near', in which the *t*-infix is not partially assimilated with respect to emphasis (cf. Versteegh 2001:21), it does belong, together with /x/, /ġ/, and the four emphatics to the *ḥurūf mustaʿliya* 'ascending sounds'. These consonants have the effect of emphaticizing preceding nonemphatic consonants, as in *sabaqtu* > *ṣabaqtu* 'I preceded', due to their postvelar articulation (cf. Blanc 1969:19).

3. *Qāf* in the Arabic dialects

The array of different pronunciations of /q/, i.e. unvoiced, voiced, palatalized, or even plain [k], dates well back in history (cf. Rabin 1951:55–56, 125–126). Fischer and Jastrow (1980:52), as well as Kaye and Rosenhouse (1997:270), provide an overview of this broad scope of pronunciation. While the 'school pronunciation' of /q/ is indeed attested in some Syrian, Mesopotamian, and North African dialects, /q/ surfaces as [k] in those dialects (e.g. central Palestinian ones) which have palatalized original /k/. In many Bedouin dialects as well as in most of the Arabian Peninsula, /q/ surfaces as [ɢ], [dʒ], or even as [dz]. Fricative [ɣ] is attested in some Mesopotamian dialects (cf. Fischer and Jastrow 1980:143), where [q] and [ɣ] are 'switched' in pronunciation (cf. also Al-Nassir 1993:40). Interestingly, some dialects in Yemen and the central Najd have palatalized [dʒ] for /q/ and [tʃ] for /k/ (→ *kaškaša*), a parallelism that reflects again the assumed old voiced quality of /q/. A voiceless glottal stop [ʔ] is the most common reflex of /q/ in the cities around the Mediterranean as well as in most of Syria and Lebanon. This sound shift may be explained by the glottal co-occlusion that sometimes accompanies the pronunciation of voiceless [q] (cf. Blanc 1969:26). Some place-names, as well as cultural and religious terms, do not undergo this sound change, notably *al-qāhira* ‘Cairo’ and *al-qur*ʾ*ān* ‘the Qurʾān’ (‘Jerusalem’, however, winds up phonetically as [alʔuds]). Sociolinguistic factors can affect the pronunciation of /q/ as well. Blanc (1969:22) mentions a passage in Ibn Xaldūn’s *Muqaddima* where the sociolinguistic implications of the *q/g* split are discussed. Versteegh (2001:137–138) draws attention to the fact that the Muslim *gilit* dialect of the Baghdad area has had a higher prestige than *qəltu* dialects of the province, even though the pronunciation [q] is more closely associated with Classical Arabic. The phonetic surface [G] being one of the shibboleths of Bedouin pronunciation, this observation underlines once more the role of the Bedouin as arbiters in linguistic matters.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Ibn Jinnī, *Sirr* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *Sirr ṣināʿat al-ʾiʿrāb*. Ed. Ḥasan Hindāwī. 2 vols. Damascus: Dār al-Qalam, 1985.

Ibn Xaldūn, *Muqaddima* = Walī d-Dīn ʿAbd ar-Raḥmān ibn Muḥammad Ibn Xaldūn, *al-Muqaddima*. Ed. Darwīš al-Juwaydī. Beirut: al-Maktaba al-ʿAṣriyya, 1996.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān ibn Qanbar Sībawayhi, *al-Kitāb*. Ed. ʿAbd as-Salām Muḥammad Hārūn. 5 vols. Cairo: al-Hayʾa al-Miṣriyya al-ʿĀmma li-l-Kitāb, 1966–1977.

Zamaxšarī, *Mufaṣṣal* = ʾAbū l-Qāsim Maḥmūd ibn ʿUmar az-Zamaxšarī, *al-Mufaṣṣal fī n-naḥw*. Ed. Jens Peter Broch. Christiania: Libraria P.T. Mallingii, 1891.

SECONDARY SOURCES

Al-Nassir, Abdulmunim Abdalamir. 1993. *Sibawaih the phonologist: A critical study of the phonetic and phonological theory of Sibawaih in his treatise* Al-Kitab. London and New York: Kegan Paul International.

Bakalla, Muhammad H. 1982. *Ibn Jinnī, an early Muslim phonetician: An interpretative study of his life and contributions to linguistics*. London and Taipei: European Language Publications.

Blanc, Haim. 1967. “The ‘sonorous’ vs. ‘muffled’ opposition in Old Arabic”. *To honor Roman Jakobson* I, 295–308. The Hague: Mouton.

——. 1969. “The fronting of Semitic *g* and the *qāl-gāl* dialect split in Arabic”. *Proceedings of the International Conference on Semitic Studies held in Jerusalem, 19–23 July 1965*, 7–37. Jerusalem: Israel Academy of Sciences and Humanities.

Bravmann, Max (Meir). 1934. *Materialien und Untersuchungen zu den phonetischen Lehren der Araber*. Göttingen: Dieterich’sche Universitäts-Buchdruckerei.

Brockelmann, Carl. 1908. *Grundriß der vergleichenden Grammatik der semitischen Sprachen*. I. *Laut- und Formenlehre*. Berlin: Reuther und Reichard. (Repr., Hildesheim: G. Olms, 1999.)

Carter, Michael G. 2004. *Sībawayhi*. Oxford: I.B. Tauris.

Fischer, Wolfdietrich and Otto Jastrow (eds.). 1980. *Handbuch der arabischen Dialekte*. Wiesbaden: O. Harrassowitz.

Fleisch, Henri. 1958a. “La conception phonétique des Arabes d’après le *Sirr ṣināʿat al-iʿrāb* d’Ibn Ǧinnī”. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 108.74–105.

——. 1958b. “*Maǧhūra*, *mahmūsa*: Examen critique”. *Mélanges de l’Université Saint-Joseph* 35.193–210.

Garbell, Irene. 1958. “Remarks on the historical phonology of an East Mediterranean Arabic dialect”. *Word* 14.303–337.

Kästner, Hartmut. 1981. *Phonetik und Phonologie des modernen Hocharabisch*. Leipzig: Akademie-Verlag.

Kaye, Alan and Judith Rosenhouse. 1997. “Arabic dialects and Maltese”. *The Semitic languages*, ed. Robert Hetzron, 263–311. London: Routledge.

Leslau, Wolf. 1987. *Comparative dictionary of Geʿez*. Wiesbaden: O. Harrassowitz.

Levin, Aryeh. 1999. “The first book of Arabic dialectology: Sībawayhi’s *al-Kitāb*”. *Jerusalem Studies in Arabic and Islam* 23.208–220.

Lipiński, Edward. 2001. *Semitic languages: Outline of a comparative grammar*. 2nd ed. Leuven: Peeters.

Maddieson, Ian. 1984. *Patterns of sounds*. Cambridge: Cambridge University Press.

Mahadin, R.S. 1997. “Perspectives on the traditional Arab grammarian’s description of the /q/ sound”. *Zeitschrift für Arabische Linguistik* 34.32–52.

Odisho, Edward. 1988. “Sībawayhi˙s dichotomy of *majhūra/mahmūsa* revisited”. *al-ʿArabiyya* 21.81–91.

Rabin, Chaim. 1951. *Ancient West-Arabian*. London: Taylor’s Foreign Press.

Soden, Wolfram von. 1995. *Grundriß der akkadischen Grammatik*. 3rd ed. Rome: Pontificio Istituto Biblico.

Versteegh, Kees. 2001. *The Arabic language*. 2nd ed. Edinburgh: Edinburgh University Press.

Wright, William. 1967. *A grammar of the Arabic language*. Trans. from German. 3rd ed., rev. by W. Robertson Smith and M.J. de Goeje. 2 vols. Cambridge: Cambridge University Press.

LUTZ EDZARD (University of Oslo)

Qawl → Kalām

Qəltu Arabic → Iraq

Qirāʾāt

The *qirāʾāt* (sg. *qirāʾa*) represent the vast corpus of Qurʾānic readings that are traditionally linked to the textual transmission and recitation of Islam's sacred book. It is to the skeletal text (*rasm*) of the *Qurʾān* that all of these readings are ultimately bound, reflecting subtle variations in the linguistic features of the text. The nature of variance among these readings ranges from differences and distinctions which occur at the morphosyntactic and morphophonological levels of the Qurʾānic text and are seemingly of an infinitesimal countenance, to those in which the nature of variance is more pronounced and reflected in consonantal variants and manifest instances of exegetical interpolation. These readings were the subject of critical grammatical analysis and scrutiny by the earliest Arabic grammarians as they attempted to accommodate their grammatical features within the confines of a general theory of language. Naturally, the *qirāʾāt* not only serve as important sources for the linguistic situation in early Islam, but they also provide insights into attitudes toward the language of scripture and developments in grammatical thinking during the early periods. Modern scholarship has often referred to readings under the rubric Qurʾānic variants, although, theoretically speaking, many of them are not deviations from the standard text but rather encapsulate intrinsic facets of its articulation.

1. History of *qirāʾa* and the issues of canon

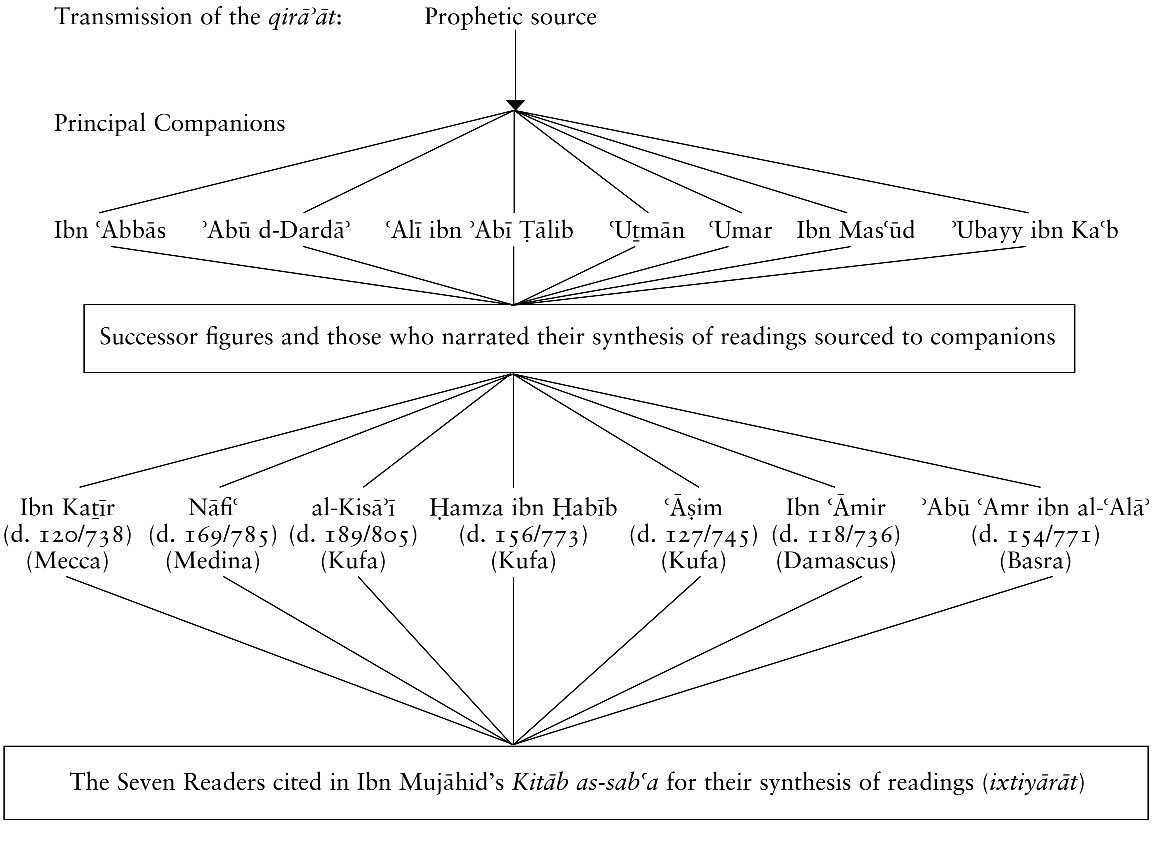
Classical Muslim sources relate that in the lifetime of the Prophet, the whole of the *Qurʾān* was not collected together in a single document (*ne varietur textus receptus*) but was partly preserved on sheets of parchment, the ribs and shoulder blades of animals, the stalks of palms, and above all memorized in the hearts of men. Despite the suggestion that following the Wars of Apostasy in 11/633 a collection of the *Qurʾān* was sanctioned by the second caliph ʾAbū Bakr, it is the third caliph, ʿUṯmān ibn ʿAffān (r. 23–35/644–656) who is formally credited with having commissioned an official collection of the *Qurʾān*. This version was imposed as the standard codex *(muṣḥaf)* throughout the territories of the state. It was in these regions that traditions of reciting and preserving the sacred text had been established by the Companions who settled there. Tradition states that differences and disagreements regarding the recitation of the sacred text led to ʿUṯmān's intervention. He appointed an editorial committee that was led by a scribe of the Prophet, Zayd ibn Ṯābit (d. 32/652–653). An official codex comprising the skeletal text of the *Qurʾān* was produced, and four recensions of this master copy were sent to major cities and garrison towns (*ʾamṣār*), including Mecca, Kufa, Basra, and Damascus; a further copy was retained in Medina. None of these original codices has survived, although genres of writing devoted to collating the orthographical features of indigenous codices do refer to instances of their being used as prototypes for the transcription of further copies (Danī, *Muqniʿ* 102.15–19; Cook 2004:103–104; Schoeler 1992:21–27).

Qurʾān readers associated with indigenous cities developed syntheses (*ixtiyārāt*) of readings which were sourced to luminaries among the Companions. They were identified as having derived their readings from the Prophet, defining a theoretical hierarchy of authority for the transmission of *qirāʾāt*. The term → *ḥarf* (pl. *ḥurūf* and *ʾaḥruf*) was used to designate a reader's specific *lectio* or reading. Minor variations among these readings were said to be sanctioned in a Prophetic tradition which refers to the *Qurʾān* being revealed in several modes or *ḥurūf*, and declares that each of these modes was liturgically valid. One reason given by traditional scholarship for the existence of so many Qurʾānic variants was that they were partially a reflection of the dialectal diversity of the indigenous Arab tribes, who were granted a measure of latitude in their recitation of the sacred text (→ pre-Islamic Arabic). Their syntactic, phonological, and morphological conventions and idiosyncrasies were enshrined within the corpus of readings (Ibn Qutayba, *Muškil* 39.1–12). The *textus receptus* or *ʾimām* distributed by ʿUṯmān was apparently transcribed on parchment in the so-called *scriptio defectiva*: certain long vowels were not physically represented in this script. A system of short vowel annotation had not yet been developed, and the use of diacritics to distinguish individual consonants was somewhat irregular. The Arabic script consisted of 15 basic graphemes which, through the addition of diacritic dots, produced the 28 characters required for its phonemic repertoire, allowing single homographs to represent more than one phoneme. There has been the suggestion that the proliferation of Qurʾānic readings was the result of ambiguities created by the incipient nature of the Arabic script; however, within the reading tradition it was always maintained that oral mechanisms for the transmission of readings retained overall hegemony, essentially governing the articulation of the written text, which served as a mnemonic aid.

Given that the vast majority of *qirāʾāt* reflected differences concerning vocalic values, consonantal variants, and the appendage of conjugational markers, the skeletal text promulgated by ʿUṯmān accommodated a large number of these readings. However, contraventions of the skeletal boundaries set by the ʿUṯmānic codices were not permitted, despite the fact that eminent Companions of the Prophet such as ʿAbdallāh Ibn Masʿūd (d. 32/652), ʾUbayy ibn Kaʿb (d. 20/641 or 22/643), ʾAbū Mūsā al-ʾAšʿarī (d. 42/662 or 52/672), and Ibn ʿAbbās (d. 68/687–688) all possessed personal codices which retained exegetical interpolations and consonantal variants inconsistent with the standardized text. A consensus of readings gradually developed, with different cities adopting *qirāʾāt* identified with individual readers who sourced their *lectiones* to earlier authorities. It is apparent that when Ibn Mujāhid (d. 324/936) composed a work collating seven Qurʾānic readings associated with seven distinguished readers, he was guided to a large extent by the preeminent status these readings had already acquired in their indigenous cities (see Fig. 1). Besides, earlier figures had already collated collections of readings which served as the principal sources for his work.

It is important to bear in mind the nature of variance among these *qirāʾāt*. The *Qurʾān* consists of some 6,236 separate verses, and given that the length of a verse and the units of semantically independent speech contained within verses varies, instances of differences among readings were frequently confined to the vocalic values of one or more lexemes within a single verse. For example, Ibn Mujāhid's

Figure 1. The Seven Readers selected by Ibn Mujāhid.


Transmission of the *qirāʾāt*:
Prophetic source
Principal Companions
Ibn ʿAbbās
ʾAbū d-Dardāʾ
ʿAlī ibn ʾAbī Ṭālib
ʿUṯmān
ʿUmar
Ibn Masʿūd
ʾUbayy ibn Kaʿb
Successor figures and those who narrated their synthesis of readings sourced to companions
Ibn Kaṯīr (d. 120/738) (Mecca)
Nāfiʿ (d. 169/785) (Medina)
al-Kisāʾī (d. 189/805) (Kufa)
Ḥamza ibn Ḥabīb (d. 156/773) (Kufa)
ʿĀṣim (d. 127/745) (Kufa)
Ibn ʿĀmir (d. 118/736) (Damascus)
ʾAbū ʿAmr ibn al-ʿAlāʾ (d. 154/771) (Basra)
The Seven Readers cited in Ibn Mujāhid's *Kitāb as-sabʿa* for their synthesis of readings (*ixtiyārāt*)

collection of seven readings is essentially an inventory of the documented differences among established readers and is consistently confined to the voweling of individual phonemes in addition to consonantal variants in respect of conjunctions, suffixes, and prefixes in selected verses of the *Qurʾān*. Many individual verses have no recorded differences. Critical to Ibn Mujāhid's survey of readings was an introduction to the approaches adopted by readers to sundry phonological phenomena, such as → *ʾimāla* 'fronting or inclination of the vowel *a*'; *taḥqīq al-hamza* 'giving the → *hamza* its full articulation'; → *ʾidġām* (*iddiġām*) 'assimilatory processes'; → *kināya* 'the articulation of pronouns'; and *yāʾāt al-ʾiḍāfa* 'the pronunciation of possessive suffixes formed in the 1st person singular'. Criteria for the acceptance of a reading included compatibility with the ʿUṯmānic codices, consistency with the conventions of the → *ʿarabiyya*, and a valid chain of authority. The principle that *qirāʾāt* had to be based on a legacy of defined precedents was accentuated within this arrangement, becoming one of the axioms of the reading tradition. Ibn Qutayba (213–276/829–889) made the point that, although the earliest generations of readers had exercised the license granted to them by the Prophetic statement concerning the *Qurʾān* being revealed in seven *ʾaḥruf*, later generations of readers were simply drawing from the pool of readings circumscribed by earlier luminaries. Thus, by the 3rd/9th century the corpus of canonical readings had effectively been determined. Any reading meeting these conditions was to be considered Qurʾānic in the strict sense of the word. Such readings could claim liturgical authority as representing the literal speech of God (*kalām Allāh*) and were deemed to be valid for devotional acts of worship in which the recitation of scripture was obligatory.

The nature of variance among readings led John Wansbrough to argue that such infinitesimal differences do not seem to have justified the imposition of the ʿUṯmānic codex, especially since minimal deviation from the canon could be justified by reference to the interpretation of the Prophetic tradition sanctioning the different modes or *ḥurūf* in which the *Qurʾān* was revealed (Wansbrough 1977:44–45). He was of the view that traditional discourse on this subject had an etiological function, aimed at creating the impression of the early existence of a canonical body of scripture. This view assumes that the *ʾaḥruf* doctrine was entirely arbitrary; however, the classical tradition seems to imply that it was equally governed by the strictures of precedent, and this appears evident in the readers' formulation of *ḥurūf* and *ixtiyārāt*. Given that the issue here is the devotional value of readings, 'infinitesimal differences' were deemed critical. Conversely, John Burton argued that variant readings were the conscious product of attempts to circumvent legal inconsistencies in the text of the *Qurʾān* (1977:141–146). Yet, this particular view does not account for the numerous readings which were univocal in nature.

Accepting the existence of written Qurʾānic materials in the early tradition, Gerhard Luling has propounded the theory that the text of the *Qurʾān* and its many readings were configured around a *Ur-text* consisting of pre-Islamic Christian strophic hymns; his argument is that the incipient nature of the Arabic script allowed Islamic scholars to rework and reinterpret these texts, reconciling them with the established Arabic vernacular that had hitherto gained ascendancy (Luling 2003:12–18). Employing a similar framework, Christoph Luxenberg (a pseudonym) contended that the lexical and syntactic structures of the *Qurʾān* were essentially Syro-Aramaic in origin. Accordingly, he suggested that the readings associated with the *Qurʾān* had been the product of the editorial endeavors of later Islamic scholarship. Such views dismiss the perspectives presented by traditional literature and reflect a belief that the authority of the oral tradition was contrived by later scholarship (Luxenberg 2007:22–32).

2. The *qirāʾāt* and the synthesis of early grammatical thought

The *Qurʾān* provided a preliminary framework for the development of Arabic linguistic thought, engendering activities in the field of orthographical improvements, conventions for the recitation of scripture, collating codices, and the lexical explication of the sacred text. However, it was never the intention that this sacred language should serve as the principal basis for a normative model of Arabic grammar (Carter 2004:48–49). Rather, the linguistic configuration and structure of the *Qurʾān* together with the extensive range of variations offered by the *qirāʾāt* provided grammarians with a profusion of data, which they used to give context and definition to their own derived grammatical constructs and theories.

Working toward a detailed description and study of the language of the Arabs, the earliest generations of grammarians explored a range of linguistic sources. These included the *Qurʾān* together with its *qirāʾāt*; the speech conventions of the Bedouin; the poetry of the Arabs; and proverbs and idiomatic expressions. The early grammarians generally accepted the sacrosanct status of the skeletal outline of the ʿUṯmānic text, adhering to the prevailing consensus regarding the authoritative status of established readings. Yet, because of their interest in the intrinsic theoretical value of such materials, they were prepared to defend grammatically those *qirāʾāt* whose canonical status was judged to be dubious. It has been argued that grammarians deliberately accepted an abstract distinction between the *Qurʾān* and the *qirāʾāt*, allowing them to be critical of the latter (Baalbaki 1985:31–32). Nevertheless, given the intimate nature of the relationship between the two sources, such an approach was theologically controversial, particularly when it came to commenting on the grammatical idiosyncrasies of readings whose canonical status was incontrovertible. The issue seemingly separated readers from grammarians, leaving its mark on the grammatical and exegetical literature of later periods.

Classical biographical literature is replete with references to theoretical discussions among early luminaries, which are inspired by attempts to explicate the grammatical features and idiosyncrasies of various *qirāʾāt*. A typical example of this type of discussion is preserved by Ibn Sallām (139–232/757–847) in his survey of the classes of ancient poets (*Ṭabaqāt* 32.16–20). Two rather prominent early Basran figures, ʾAbū ʿAmr ibn al-ʿAlāʾ (d. 154/771) and ʿĪsā ibn ʿUmar aṯ-Ṯaqafī (d. 149/766), are said to have expressed their disagreement regarding the effective cause (→ *ʿilla*) of the grammatical inflection of Q. 34/10, *yā jibālu ʾawwabī maʿahu wa-ṭ-ṭayr(a)* 'mountains, sing [God's] praises and you birds, too'. The issue was the cause of the accusative inflection of *aṭ-ṭayr(a)*. ʿĪsā ibn ʿUmar argued that a vocative agent governed its inflection; this was rejected by ʿAbū ʿAmr, who identified a process of ellipsis. Both readers agreed about the accusative inflection of the verse, but they disagreed as to its grammatical rationale. One has to bear in mind that this report occurs in a late biographical source. Nevertheless, the level of discourse appears to be somewhat advanced and commensurate with the technical treatment of *qirāʾāt* that one encounters in the *Kitāb* of Sībawayhi (d. 180/796), the earliest systematic grammatical text.

ʿĪsā ibn ʿUmar is linked with an early Meccan reader, Ibn Muḥayṣin (d. 123/740), who is recorded as having developed a synthesis of readings (*ixtiyār*) based on a system of *ʿarabiyya* (Ibn al-Jazarī, *Ġāya* II, 167.15–20; Ibn Mujāhid, *Sabʿa* 65.4–9). It was at variance with the general consensus on readings reached by the Meccans and contravened the *rasm* of the ʿUṯmānic codex. Such was the preoccupation with the grammatical features of readings among early 'reader-grammarians' that some scholars such as Kees Versteegh initially accepted that introspection of this kind seemingly provided the background for the development of grammatical thought; however, Versteegh subsequently concluded that exegetical frameworks formulated for the exigencies of grappling with the meanings of scripture realistically conferred a more feasible theoretical framework for the development of the grammatical tradition (Versteegh 1990:238–239).

In Sībawayhi's *Kitāb*, references to *qirāʾāt* are often in the context of demonstrating points of grammar, confirming that certain grammatical features of a particular reading could be reconciled and contextualized with the diction of the Arabs. Idiosyncratic readings which conflicted with consensus readings are sometimes defended. One such example is Sībawayhi's discussion of Q. 41/17 *wa-ʾammā Ṯamūd(a) fa-hadaynāhum* 'as for the people of Thamud, We guided them". This particular reading conflicted with the commonly accepted *lectio* favored by readers in which *Ṯamūd(u)* takes a nominative ending. Sībawayhi was aware of this fact because, having mentioned the accusative reading, he adds the caveat that one should not contravene the accepted reading, for it is "an established convention" (*Kitāb* I, 148.4–6). At a separate juncture in the *Kitāb*, Sībawayhi explains the syntactic rationale behind the nominative reading of *Ṯamūd*, referring to its inchoative status (*Kitāb* I, 95.6–8; 81.7–9). The whole point of this exercise was to demonstrate that the grammatical features of readings were consistent with the diction of the Arabs. His discussion of Q. 54/49, *ʾinnā kull(a) šayʾ(in) xalaqnāhu bi-qadar* 'and all things We created with divine decree', is an exceptional example. He equates such structures with the maxim *zayd(an) ḍarabtuhu* (*Kitāb* I, 148.4–6). Sībawayhi was not interested in the theological sensitivities of verses of this nature. He simply wanted to place such linguistic phenomena and constructions within the framework of a general theory of the Arabic language.

There are instances in the *Kitāb* when Sībawayhi refers to a reading as being 'infrequent' in a linguistic sense, using poetic citation or examples of Bedouin usage as his analogue (*Kitāb* I, 58.1–4). This has led to the contention that such approaches impinged upon the sacrosanct nature of readings; it is a charge leveled at Sībawayhi and indeed, over the centuries, against later Basran luminaries (Šalabī 1958:160–165). They were accused of indulging in the emendation of Qurʾānic readings, pursuing the hypothetical projection of grammatical constructions which contravened the ʿUṯmānic codex (Bernards 1997:24). The inference is that certain readings were hardly distinguished in this early period as being emblems of linguistic excellence, although perhaps such attitudes toward readings illustrate the very broad and sophisticated confines within which grammarians were able to operate and express their views candidly, while the use of profane sources such as poetry to justify the grammatical features of readings was always going to be contentious. The exegete Faxr ad-Dīn ar-Rāzī (d. 606/1209) was appalled that grammarians were prepared to adduce anonymous pieces of poetry to authenticate readings, retorting that the opposite should be the case (*Mafātīḥ* V, 169). One should, however, bear in mind the motives of the grammarians, for whom the *qirāʾāt* represented a source of linguistic data which, like other sources, had to be rationalized and placed within the abstract schema that was grammar. It was a fascination with language that spurred them on (Levin 2004). That such an extensive corpus of readings could be examined across a wide compass of grammatical topics and theories gives some indication of the sophistication of scholarship attained in these relatively early periods.

With reference to the historical existence of two conventional traditions of linguistic thought, namely the Basran and the Kufan schools, one recent study has argued that the former, beginning with Sībawayhi's efforts, extended its analyses to a much broader corpus of linguistic data. The suggestion is that the Kufans confined their linguistic endeavors to grammatical problems in the *Qurʾān* and its many readings and that they remained exponents of a tradition of grammatical thought in which the *qurrāʾ* or Qurʾānic readers were accepted as respected linguistic authorities (Versteegh 1993:178–179). The Kufan grammarian al-Farrāʾ (144–207/761–822) was the author of a *Maʿānī l-Qurʾān* text. It is structured around the critical grammatical exposition of selected verses of the *Qurʾān*, adhering to its traditional chapter order (Gilliot 2006:49). It adduces an array of *qirāʾāt*, both canonical and noncanonical, to flesh out sundry grammatical constructions. Poetic citation and the idiomatic expressions of Bedouin Arabs are frequently highlighted to illustrate underlying conventions and principles. It has been mentioned that individuals such as al-Farrāʾ resorted to ingenious ways of reconciling noncanonical readings, like those of Ibn Masʿūd, with the standardized text (Beck 1948:328; Versteegh 1993:39). This seemingly reflected the Kufan preoccupation with Qurʾānic variants and their receptivity to a broad and seemingly discursive corpus of linguistic data in their formulation of grammatical principles. However, at other junctures in the *Maʿānī*, al-Farrāʾ states that "adherence to the codex, if it can be related to an aspect of the speech of the Arabs and the readings of the *qurrāʾ*, is preferable to me than contradiction therein" (*Maʿānī* II, 293.14–15). Such statements give the impression that Kufans in general respected the sacrosanct nature of readings, although it did not temper the enthusiasm with which grammatical treatments of the sacred text were pursued; however, equally, it should be noted that Kufan luminaries such as al-Farrāʾ were prepared to countenance the rejection of Qurʾānic readings they deemed to be grammatically anomalous. Al-Farrāʾ's discussion of Q. 4/1 is indicative of this tendency (*Maʿānī* I, 252.7–12). This is despite the fact that the *qirāʾa* was accepted by readers as being canonically sound.

The grammatical justification and authentication of Qurʾānic readings is one of the underlying principles of the *maʿānī* works, presupposing the existence of a general theory of grammar within which the materials could be appraised. The *maʿānī l-Qurʾān* genre of writing provided the broad framework through which such forms of critical analysis were pursued among Kufan scholars, although it misleadingly created the impression that the *Qurʾān* and its readings formed the core of their tradition of language study. A survey of both references to the grammatical discourse ascribed to early Kufan luminaries together with the putative works that they composed betrays a much more extensive compass to their linguistic activities.

A contemporary of al-Farrāʾ, the Basran ʾAbū ʿUbayda (d. 210/825) was the author of a similar work, entitled *Majāz al-Qurʾān*. The contents of this work confirm that it belongs to the genre of *maʿānī* literature. ʾAbū ʿUbayda spoke of a linguistic symmetry between the *Qurʾān* and the language of the Arabs (*Majāz* I, 8.4–7). Additionally, the *Maʿānī l-Qurʾān* text attributed to al-ʾAxfaš al-ʾAwsaṭ (d. 215/830), a key contemporary of Sībawayhi, confirms that Basran grammarians also took an interest in this genre of writing. Qurʾānic readings continued to be the subject of the grammarians' interest. Later luminaries such as al-Mubarrad (210–285–286/815–898), Ibn as-Sarrāj (d. 316/928), az-Zajjāj (241–311/854–923), an-Naḥḥās (d. 338/949), and ʾAbū ʿAlī al-Fārisī (d. 377/987) all composed works in which canonical readings were grammatically defended (Shah 2004:94). The terms *iḥtijāj* and *ḥujja* were used to define this genre of writing. Even those readings which fell outside the confines of canonical material were the subject of grammatical apologia, with both Ibn Jinnī (d. 392/1002) and Ibn Xālawayhi (d. 370/980) being the authors of such texts. Classical scholarship categorized readings which enjoyed successive levels of multiple transmission as being *mutawātir*. Readings which did not enjoy prolific levels of recognition and reception, despite the fact that they met the criteria for acceptance associated with the imposition of the ʿUṯmānic codices, were initially designated as being *šāḏḏa* 'infrequent or exceptional'. Subsequently, the term *šāḏḏa* was used to denote readings which were in clear violation of the consonantal outline of the ʿUṯmānic codices or those without credible authority. That grammarians such as Ibn Jinnī were prepared to mount a grammatical defense of such readings underlines the objective attitude that grammarians had adopted toward *qirāʾāt* as a linguistic source. Intriguingly, the crystallization of formal schools of linguistic thought and the consolidation of a canonical model of grammar are viewed by Jonathan Owens as coinciding with the appearance of Ibn Mujāhid's catalog of seven readings. Such developments are perceived as indicators of the general trends toward polarization and homogeneity in the 3rd/9th century of the Islamic tradition (Owens 1990:219); however, such trends were already firmly in place within the reading tradition well before these periods.

3. Readings and the linguistic situation in the first centuries of Islam

The grammarians' treatment of the vast corpus of the *qirāʾāt* sheds light on the linguistic situation in the early periods of the Islamic tradition. Their willingness to level criticism at readings which conflicted with accepted norms confirms the rather prescriptive countenance of the grammatical models they developed. It also illustrates that a hierarchy of linguistic sources was already in the ascendancy in these early periods. The general consensus is that Eastern Arabian dialectal sources were critically used to cultivate grammatical models. Indeed, one view is that the Classical Arabic language or *fuṣḥā* with which the early grammarians were preoccupied was mainly referenced to Eastern Arabian sources (Carter 2004:41). Mention is made of the distinct differences between the 'Classical Arabic' idiom and the Ḥijāzī dialect (Versteegh 2001:39, 46–47). This is connected to the notion that the elevated literary diction in which the *Qurʾān* was composed reflected Eastern Arabian influences: it was seemingly modeled on a pre-Islamic → poetic koine (Zwettler 1978:109, 133–134; Corriente 1976:75). However, identifying the geographical origin of the literary koine together with its definitive substrate influence remains a rather speculative endeavor. The multifaceted nature of the relationship between the *Qurʾān* and the *qirāʾāt* intimates that attitudes toward the relative perception of this literary koine differed widely. The topic has obvious ramifications for the debate regarding the presence of declension in the spoken vernacular of the Arabs in the early periods and the whole issue of the traditional emphasis on the preeminence of the Qurašī dialect.

On the other hand, the frequent references in the *qirāʾāt* literature to the refined phonological features of Eastern Arabian or ‘Tamīmī’ dialects, such as *ʾimāla*, *taḥqīq al-hamza*, and *ʾidġām*, were genuinely redolent of the linguistic authority which Eastern Arabian dialects had enjoyed in these early periods. However, syntactic and phonological features of the Ḥijāzī dialect as preserved in the syntheses of *qirāʾāt* were noticeably accorded venerated status. Sībawayhi’s discussion of Q. 12/31 *mā haḏā bašar(an)* ‘this is no mortal being’, in which the particle *mā* operates in the same manner as *laysa*, with its predicate taking an accusative ending, is one such example (*Kitāb* I, 59.3–4; Ibn Jinnī, *Xaṣāʾiṣ* I, 125.1–7). He describes the fact that the verse exhibits a Ḥijāzī dialectal trait, adding that the tribes of Tamīm retain a nominative ending in such instances, except those aware of its transcription in the ʿUṯmānic codex, which would support only the accusative reading. The ensuing discussion provided by Sībawayhi and Ibn Jinnī verifies that the Tamīmī trait was judged to be more regular (*ʾaqyas*). Nonetheless, it is apparent that both individuals held the Ḥijāzī dialect in esteem: Ibn Jinnī stresses the point that the *Qurʾān* was revealed in its vernacular. It is also the case that Western Arabian phonological traits such as *fakk al-ʾidġām* ‘separating geminated consonants’, *tashīl al-hamz* ‘weakening or eliding the glottal stop at nonpausal junctures’, and *al-fatḥ* (or *at-tafxīm*) ‘opening the vowel *a*’, as recorded in the *qirāʾāt*, were clearly considered to be fine archaic features of the Ḥijāzī dialect (Sībawayhi, *Kitāb* IV, 424.8–11, IV, 120.12–16; Naḥḥās, *ʾIʿrāb* I, 250.7–14). Notwithstanding the nature of variance among readings and the modes of their transmission, the subtle range of syntactic, morphological, phonological, and phonetic properties which syntheses of the *qirāʾāt* encompassed serves not only as testimony to the refined measure of linguistic variety prevalent in the early periods, but also to the significance of the role the *lingua sacra* played in their preservation.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

ʾAbū ʿUbayda, *Majāz* = ʾAbū ʿUbayda Maʿmar ibn al-Muṯannā, *Majāz al-Qurʾān*. Ed. Fuʾād Sazkīn [Fuat Sezgin]. 2 vols. Beirut: Muʾassasat ar-Risāla, 1981.

ʾAxfaš, *Maʿānī* = ʾAbū l-Ḥasan Saʿīd ibn Masʿada al-ʾAxfaš al-ʾAwsaṭ, *Maʿānī l-Qurʾān*. Ed. Fāʾiz Fāris. 2 vols. Kuwait: Dār al-Bašīr, 1981.

Dānī, *Muqniʿ* = ʾAbū ʿAmr ʿUṯmān ibn Saʿīd ad-Dānī, *al-Muqniʿ fī maʿrifa marsūm maṣāḥif ʾahl al-ʾamṣār*. Ed. M. Dahmān. Damascus: Dār al-Fikr, 1983.

Farrāʾ, *Maʿānī* = ʾAbū Zakariyyāʾ Yaḥyā ibn Ziyād al-Farrāʾ, *Maʿānī l-Qurʾān*. Ed. Muḥammad ʿAlī an-Najjār. 3 vols. Cairo: ad-Dār al-Miṣriyya, 1966–1972.

Ibn al-Jazarī, *Ġāya* = Šams ad-Dīn ʾAbū l-Xayr Muḥammad ibn Muḥammad Ibn al-Jazarī ad-Dimašqī, *Ġāyat an-nihāya fī ṭabaqāt al-qurrāʾ*. Ed. Gotthelf Bergsträsser and Otto Pretzl. 3rd ed. 2 vols. Cairo: Maṭbaʿat as-Saʿāda, 1935.

——, *Našr* = Šams ad-Dīn ʾAbū l-Xayr Muḥammad ibn Muḥammad Ibn al-Jazarī ad-Dimašqī, *an-Našr fī l-qirāʾāt al-ʿašr*. 2 vols. Ed. ʿAlī Muḥammad aḍ-Ḍabbāʿ. Beirut: Dār al-Kutub al-ʿIlmiyya, n.d.

Ibn Jinnī, *Muḥtasab* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Muḥtasab fī tabyīn wujūh šawāḏḏ al-qirāʾāt wa-l-ʾīḍāḥ ʿanhā*. Ed. ʿAlī an-Najdī Nāṣif, ʿAbd al-Ḥalīm an-Najjār, and ʿAbd al-Fattāḥ Šalabī. 2 vols. Cairo: Lajna ʾIḥyāʾ at-Turāṯ al-ʿArabī, 1387 A.H.

——, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. Ed. Muḥammad ʿAlī an-Najjār. 3 vols. Cairo: Dār al-Kutub al-Miṣriyya, 1952–1956.

Ibn Mujāhid, *Sabʿa* = ʾAḥmad ibn Mūsā Ibn Mujāhid, *Kitāb as-sabʿa fī l-qirāʾāt*. Ed. Šawqī Ḍayf. Cairo: Dār al-Maʿārif, 1400 A.H.

Ibn Qutayba, *Taʾwīl* = Muḥammad ʿAbdallāh ibn Muslim Ibn Qutayba, *Taʾwīl muškil al-Qurʾān*. Ed. ʾAḥmad Ṣaqr. 2nd ed. Cairo: Dār at-Turāṯ, 1973.

Ibn Sallām, *Ṭabaqāt* = ʾAbū ʿAbdallāh Muḥammad Ibn Sallām al-Jumaḥī, *Ṭabaqāt fuḥūl aš-šuʿarāʾ*. Ed. Ṭāhā ʾAḥmad ʾIbrāhīm. Beirut: Dār al-Kutub al-ʿIlmiyya, 1988.

Kirmānī, *Mafātīḥ* = ʾAbū l-ʿAlāʾ Muḥammad ibn ʾAbī l-Maḥāsin al-Kirmānī, *Mafātīḥ al-ʾaġānī fī l-qirāʾāt wa-l-maʿānī*. Ed. ʿAbd al-Karīm Muṣṭafā Mudlij. [Beirut: Dār Ibn Ḥazm, 2001.]

Naḥḥās, *ʾIʿrāb* = ʾAbū Jaʿfar ʾAḥmad ibn Muḥammad an-Naḥḥās, *ʾIʿrāb al-Qurʾān*. Ed. Zuhayr Ġāzī Zāhid. 5 vols. Beirut: ʿĀlam al-Kutub, 1988.

Rāzī, *Mafātīḥ* = Faxr ad-Dīn Muḥammad ibn ʿUmar ar-Rāzī, *Mafātīḥ al-ġayb*. 16 vols. in 32 parts. Beirut: Dār al-Fikr, 1981–1983.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. Ed. ʿAbd as-Salām Hārūn. 5 vols. Beirut: Dār al-Jīl, 1991.

SECONDARY SOURCES

Baalbaki, Ramzi. 1985. “The treatment of *qirāʾāt* by the second and third century grammarians”. *Zeitschrift für Arabische Linguistik* 15.11–32.

Beck, Edmund. 1948. “Studien zur Geschichte der kūfischen Koranlesung in den beiden ersten Jahrhunderten”. *Orientalia* 17.326–355.

Bernards, Monique. 1997. “Al-Mubarrad as key-figure in the development of early Arab grammatical science”. *Zeitschrift für Arabische Linguistik* 34.7–27.

Brockett, Adrian. 1988. “*Qurʾān* readings in *Kitāb* Sībawayhi”. *Occasional Papers of the School of Abbasid Studies* (University of St. Andrews) 2.129–206.

Burton, John. 1977. *The collection of the* Qurʾān. Cambridge: Cambridge University Press.

Carter, Michael. 1994. "Writing the history of Arabic grammar". *Historiographia Linguistica* 21.385–414.

——. 2004. *Sībawayhi*. London: I.B. Taurus; New York: Oxford University Press.

Cook, Michael. 2004. "The stemma of the regional codices of the *Koran*". *Graeco-Arabica* 9–10.89–104.

Corriente, Federico. 1976. "From Old Arabic to Classical Arabic through the pre-Islamic koine: Some notes on the native grammarians' sources, attitudes, and goals". *Journal of Semitic Studies* 21.62–98.

Dévényi, Kinga. 1991. "Al-Farrāʾ and al-Kisāʾī: References to grammarians and *Qurʾān* readers in the *Maʿānī al-Qurʾān* of al-Farrāʾ". *Proceedings of the Colloquium on Arabic Grammar*, ed. Kinga Dévényi and Tamás Iványi, 159–76. Budapest: Eötvös Loránd University.

Gilliot, Claude. 2006. "Creation of a fixed text". *Cambridge companion to the* Qurʾān. ed. Jane Dammen McAuliffe, 41–57. Cambridge: Cambridge University Press.

Hopkins, Simon. 2003. Review of: Luxenberg (2000). *Jerusalem Studies in Arabic and Islam* 28.377–380.

Levin, Aryeh. 2004. "The status of the science of grammar among Islamic sciences". *Jerusalem Studies in Arabic and Islam* 29.1–16.

Luling, Günter. 2003. *A challenge to Islam for reformation: The rediscovery and reliable reconstruction of a comprehensive pre-Islamic Christian hymnal hidden in the* Koran: *Under earliest Islamic reinterpretations*. Delhi: Montilal Banarsidass.

Luxenberg, Christopher. 2000. *Die syro-aramäische Lesart des Koran: Ein Beitrag zur Entschlüsselung der Koransprache*. Berlin: Das Arabische Buch. (Rev. ed., Berlin: H. Schiler, 2004). [English trans. *The Syro-Aramaic reading of the* Qurʾān, Berlin: H. Schiler, 2007.]

Makram, ʿAbd as-Salām and ʾAḥmad Muxtār ʿUmar. 1982–1985. *Muʿjam al-qirāʾāt al-qurʾāniyya maʿa muqaddima fī l-qirāʾāt wa-ʾašhar al-qurrāʾ*. 8 vols. Kuwait: Ḏāt as-Salāsīl.

Motzki, Harald. 2001. "The collection of the *Qurʾān*: A reconsideration of Western views in light of recent methodological developments". *Der Islam* 8.2–34.

Owens, Jonathan. 1990. *Early Arabic grammatical theory: Heterogeneity and standardization*. Amsterdam and Philadelphia: J. Benjamins.

Rājiḥī, ʿAbduhu. 1969. *al-Lahajāt al-ʿarabiyya fī l-qirāʾāt al-qurʾāniyya*. Cairo: Dār al-Maʿārif.

Šalabī, ʿAbd al-Fattāḥ ʾIsmāʿīl. 1958. *ʿAbū ʿAlī al-Fārisī: Ḥayātuhu wa-makānatuhu bayna ʾaʾimmat al-luġa wa-ʾaṯaruhu fī l-qirāʾāt wa-n-naḥw*. Cairo: Maktabat an-Nahḍa.

Schoeler, Gregor. 1992. "Schreiben und Veröffentlichen: Zu Verwendung und Funktion der Schrift in den ersten islamischen Jahrhunderten. *Der Islam* 69.1–43. (Reprod. *The oral and the written in early Islam*. Trans. Uwe Vagelpohl, ed. James E. Montgomery. London: Routledge, 2006.)

Shah, Mustafa. 2004. "The early Arabic grammarians' contributions to the collection and authentication of Qurʾānic readings: The prelude to Ibn Mujāhid's *Kitāb al-Sabʿa*". *Journal of Qurʾānic Studies* 6.72–102.

——. 2005. "The quest for the origins of the *qurrāʾ* in the classical Islamic tradition". *Journal of Qurʾanic Studies* 7.1–35.

Talmon, Rafael. 1999. "From the history of the study of Qurʾānic syntax". *Tradition and innovation: Norm and deviation in Arabic and Semitic linguistics*, ed. Lutz Edzard and Mohammed Nekroumi, 71–77. Wiesbaden: O. Harrassowitz.

Versteegh, Kees. 1990. "Grammar and exegesis: The origins of Kūfan grammar and the *Tafsīr Muqātil*". *Der Islam* 67.206–242.

——. 1993. *Arabic grammar and Qurʾānic exegesis in early Islam*. Leiden: E.J. Brill.

——. 2001. *The Arabic language*. 2nd ed. Edinburgh: Edinburgh University Press.

Wansbrough, John. 1977. *Quranic studies: Sources and methods of scriptural interpretation*. Oxford: Oxford University Press.

Zwettler, Michael. 1978. *The oral tradition of Classical Arabic poetry*. Columbus: Ohio State University Press.

MUSTAFA SHAH
(School of Oriental and African Studies)

Qiyās

Qiyās is the term used for an argument that relies on the similarity between two things. It is a generally accepted opinion in Islam that the use of *qiyās* can be derived from the *Qurʾān* and the *Ḥadīṯ* (al-ʿUmarī 1987:19–53), although the word itself is not found in them. It has been demonstrated elsewhere (Lloyd 1966) that the observation of similarity is a starting point of the most elementary epistemological procedures in all cultures. In Classical antiquity, it is well attested, starting from Homer in the prephilosophical age right up to the closing years of the final period, when it was used in rhetoric argumentation. Both the Epicurean and the Sceptic schools of philosophy denied the validity of logic, but they admitted analogical inferences based on the similarity of two things. Consequently, analogical reasoning was the main methodological tool of scientific discovery in the schools committed to those disciplines (for instance the *Methodikoí* or *Empeirikoí* in the medical schools).

Analogical inferences were transmitted to the Islamic sciences, such as jurisprudence (*fiqh*) and rational theology (*kalām*), and to linguistics, from various sources through various channels. The definitions of the *qiyās* given by *mutakallimūn* have been collected recently in a lexicon containing the scientific vocabulary of rational theology (Majmaʿ al-Buḥūṯ al-ʾIslāmiyya 1995:282–283). The definitions listed there may serve as a general warning against the tendency to attribute invariably the same meaning to the term *qiyās* regardless of the time and locus of its occurrence. This tendency characterizes the treatment of the problem in some studies on the grammatical tradition (Fleisch 1961). Suleiman gives a generally valid description of *qiyās* throughout his book, but in some contexts he seems to exhibit the same tendency, for instance in the passages where he tries to make a difference between the use of *taʿlīl* in *qiyās* (1999:25–31). The core of the problem is that the use of *qiyās* in 8th-century grammatical theory preceded its use in jurisprudence, where its exact meaning and theory were fixed only at the beginning of the 9th century, as pointed out by Versteegh (1977:16, 90, 96, 101).

The fact that at the dawn of linguistic studies in Islam there were two competing schools of grammar, one in Basra, the other in Kufa, underlines the need for a differentiating approach (al-Maxzūmī 1958; Weil 1913). The two schools pursued investigations based on different methodological principles. The Basran school adhered to analogy, while their Kufan colleagues were partisans of tradition.

Sībawayhi (d. 177/793) was one of the most famous representatives of the Basran school. In the *Kitāb*, one notices that *qiyās* was not yet an established term of grammar. In the comparison Sībawayhi makes between two similar structures, the following analysis can be found (*Kitāb* I, 104): "*Al-yawm* and other adverbials are in the same category as *zayd* and *ʿabdallāh* when they are not used as adverbials; this occurs, for instance, in the expression 'is it on Friday that ʿAbdallāh leaves?', or 'is it ʿAmr who ʿAbdallāh talks about?'; and the expression 'is it on Friday that one will leave?', which is like 'is it Zayd that one will go away with?'" (*wa-l-yawmu wa-ḏ-ḏurūfu bi-manzilati zaydin wa-ʿabdillāhi, ʾiḏā lam yakunna ḏurūfan wa-ḏālika [qawlu-ka]: ʾa-yawma l-jumuʿati yanṭaliqu fīhi ʿabdullāhi; ka-qawli-ka: ʾa-ʿamran takallama fīhi ʿabdullāhi, wa-ʾa-yawmu l-jumuʿati yunṭalaqu fīhi, ka-qawlika: ʾa-zaydun yuḏhabu bihi*). In this passage, Sībawayhi examines various sentences with similar structures. The expression *bi-manzila* in this text indicates that some structures in certain individual sentences, which can be regarded as generally accepted (i.e. correct) utterances, are similar (or: analogous) to those in other individual sentences, where they appear to be dubious. He reaches the conclusion that the latter can also be accepted as being correct. He finds the structures comparable and, trying to identify the grammatically correct solution, he cites their similarity.

In another passage, Sībawayhi examines the sentences *ḍarabanī wa-ḍarabtu qawmuka* and *ḍarabūnī wa-ḍarabtu qawmaka* 'your people hit me and I hit them', where the word *qawm* refers to a group of human beings, but the corresponding verb is in the singular. In Sībawayhi's opinion, this sentence is correct, though not very beautiful. It is similar to *huwa ʾaḥsanu l-fityāni wa-ʾajmaluhu* 'he is the best and the most beautiful of the young men', where the singular personal suffix *-hu* refers to the plural *al-fityāni* (*Kitāb* I, 79). The latter sentence, he says, cannot be regarded as correct.

In al-ʾAxfaš' (d. 215/830?) opinion, the grammatical structures of these individual examples are not analogous; consequently, their comparison cannot be admitted. There is no real similarity between them, i.e., their comparison is a bad analogy (*radīʾ fī l-qiyās*) because if accepted, the above construction would permit sentences like *ʾaṣḥābuka jalasa*. The latter sentence is not analogical (*lā yuqāsu ʿalayhi*) to *huwa ʾaḏrafu l-fityāni wa-ʾajmaluhu*.

For all these cases, the valid definition is one of those mentioned in the lexicon referred to above, which says (in a wording that reflects the theory of *qiyās* as it was elaborated by later scholars, because *ḥukm* and → *ʿilla* are terms that were introduced in a later period): "[*Qiyās* is] the affirmation of a judgment (*ḥukm*) known from another case on account of the similarity of the reason (*ʿilla*) of the judgment" (Majmaʿ al-Buḥūṯ al-ʾIslāmiyya 1995:282). In these cases, it can be observed that the analysis relies on a loose concept of similarity expressed by various words: *bi-manzila, šubbiha bihi, ʾaqyas*, etc. The terminology has not been set as yet.

While examining the sentence *mā ʾaḥsana ʿabdallāhi*, Sībawayhi says in another passage that according to al-Xalīl, it is equal (*bi-manzila*) to *šayʾun ʾaḥsanahu ʿabdallāh* (*Kitāb* I, 72–73). He offers the following explanation: its structure is that of sentences beginning with *faʿala, faʿila, faʿula, afʿala*. In all these cases, "They created for it one example having the same course and they compared it with those which are not [derived] from a verb, such as *lāta* and *mā*" (*jaʿalū lahu miṯālan wāḥidan yajrī ʿalayhi, fa-šubbiha hāḏā bi-mā laysa min al-fiʿl naḥwa lāta wa-mā*).

In this context, the 'one example' is a general grammatical form that is valid in many individual cases. The expression 'having the same course' (*yajrī majrāhu*) implies that the individual cases follow the same pattern, i.e., they follow a general rule. Consequently, the second meaning of *qiyās* is 'a generally accepted rule, standard, measure, canon' (Versteegh 1980:7–30, 1993:26). This corresponds to one of the definitions in the lexicon referred to above (Majmaʿ al-Buḥūṯ al-ʾIslāmiyya 1995:282), "a judgment on the individual on the basis of the constancy of the same judgment in the universal" (*ʾan yuḥkama ʿalā l-juzʾī li-ṯubūt ḏālika l-ḥukmi fī l-kullī*), which reflects the essence of *qiyās* used in this sense. By adopting the sense of 'canon', *qiyās* became an explanatory principle in grammar (Versteegh 1977:104–106, 2001:75). According to Fleisch, "Le grand effort des grammairiens arabes a été précisément d'établir le norm: le *qiyās*" (1961:viii), and Endress (1986:176–177) explains:

> Dieses Verfahren (*qiyās*, das Massnehmen an einem 'Richtmass') ist also mehr als eine lose Entsprechung, welche Wahrscheinlichkeit begründet; es erhält bei den islamischen Gelehrten den strengen Charakter einer logischen Figur: die *ʿilla* ist der terminus medius…eines (hypothetischen) Syllogismus.

For ideological reasons, this interpretation of *qiyās* was favored also by those grammarians who adhered to the Muʿtazilite school of theology (Endress 1986:183).

The theory of the legal *qiyās* (*al-qiyās aš-šarʿī*) was established by jurisprudents: aš-Šāfiʿī (d. 204/820) was the first jurisprudent who applied *qiyās* in his legal works. (For a logic-based theory of *qiyās* in its historical setting, see an-Naššār 1947.) More than three centuries later, ʾAbū l-Barakāt Kamāl ad-Dīn Ibn al-ʾAnbārī (d. 577/1181) attempted to elaborate a new theory of grammar relying on that theory. His method is summarized by Attia Amer (*Lumaʿ*, introd., xiv) as follows:

> De toute façon, al-ʾAnbārī suivit, dans son ouvrage, la méthode des jurisconsultes et celle des savants de la tradition du Prophète, et il emprunta aussi leur terminologie: *al-ʾiǧmāʿ, an-naql, at-tawātur, al-ʾāḥād, al-qiyās, al-ʾistiḥsān, ʾahl-al-ʾahwāʾ*.

Later, he adds: "Il affirme être le premier qui eut écrit un ouvrage sur la méthodologie grammaticale, et qui eut inventé cette 'science'" (*Lumaʿ*, introd., xx). Ibn al-ʾAnbārī (*Lumaʿ* 42) adapted the concepts of legal reasoning to grammar and defined the meaning of *qiyās* as "an estimation of the branch on the basis of the judgment about the root.… Every *qiyās* necessarily consists of four things: root (*ʾaṣl*), branch (*farʿ*), reason (*ʿilla*), and judgment (*ḥukm*)" (*ʿibāra ʿan taqdīr al-farʿ bi-ḥukm al-ʾaṣl…wa-lā budda li-kulli qiyāsin min ʾarbaʿat ʾašyāʾ: ʾaṣl, farʿ, ʿilla, ḥukm*). These four terms belong to the terminology of *al-qiyās aš-šarʿī*, just like the transfer of the judgment from a well-known case (*ʾaṣl*) to an unknown case (*farʿ*) for a certain reason they have in common (*ʿilla*). This common reason is the basis for the similarity between the two cases compared – a typical way of reasoning in the theological sciences. In the two previous types of *al-qiyās*, there is no room for a common reason that could motivate a transfer of judgment. *ʿIlla* in this context cannot be identified with *ʿilla* in the grammatical theory of *taʿlīl*, where the term does not have the specific meaning it has in legal theory, and it is different from the *ʿilla* of earlier grammarians (Versteegh 1995:90–94, 98).

This third type of *qiyās* is described by the following definition (Majmaʿ al-Buḥūṯ al-ʾIslāmiyya 1995:283): "an attachment of the branch to the root in respect of a judgment by a comprehensive factor" (*huwa ʾilḥāq farʿin bi-ʾaṣlin fī ḥukmin bi-sabab jāmiʿin lahumā*).

By accepting this third meaning of *qiyās*, Ibn al-ʾAnbārī could not avoid introducing further means of the *kalām* methodology into grammar, as the second *ʾaṣl* from among the three *ʾuṣūl* of linguistics (*Lumaʿ* 27–28), first of all the heuristic method of *ṭard* and the method of refutation known as *ʿaks* (Suleiman 1999:128–134; *Lumaʿ* 58–65). Any attempt to explain the linguistic meaning of *qiyās* by deriving it from *fiqh* is only right with respect to Ibn al-ʾAnbārī and his school (Carter 1972:69–97, 1973:292–304); it fails with respect to the other two (earlier) types of *qiyās*.

One cannot but agree with Gotthold Weil (1913:27), who distinguishes three uses of *qiyās*:

i. Analogical explanation of certain forms (Versteegh 1997:47–48)
ii. Regular form, linguistic canon
iii. Rational analogical reasoning of theologians (Maróth 1995:101–108)

The explanation of *qiyās* as a kind of syllogism is rejected by some authors (Bohas a.o. 1990:22–26). They deny its deductive character, because to them it is more of a heuristic method. In that respect they are right, but unfortunately they do not see any difference between the various meanings of *qiyās*.

It may be added that in modern times, *qiyās* is often adduced as one of the most effective devices in the formation of new words (Versteegh 2003:180–181).

Bibliographical references

Primary sources

Ibn al-ʾAnbārī, *Lumaʿ* = Kamāl ad-Dīn ʾAbū l-Barakāt ʿAbd ar-Raḥmān ibn Muḥammad al-ʾAnbārī, *Lumaʿ al-ʾadilla fī ʾuṣūl an-naḥw*. Ed. Attia Amer. Stockholm: Almqvist and Wiksell, 1963.
Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. Ed. ʿAbd as-Salām Muḥammad Hārūn. 5 vols. Cairo: al-Xāngī, 1982–1988.

Secondary sources

Bohas, Georges, Jean-Patrick Guillaume, and Djamal Eddine Kouloughli. 1990. *The Arabic linguistic tradition*. London and New York: Routledge.
Carter, Michael G. 1972. "Les origines de la grammaire arabe". *Revue des Études Islamiques* 40.69–97.
——. 1973. "*Ṣarf* et *ḫilāf*: Contribution à l'histoire de la grammaire arabe". *Arabica* 20.292–304.
Endress, Gerhard. 1986. "Grammatik und Logik: Arabische Philologie und griechische Philosophie im Widerstreit". *Sprachphilosophie in Antike und Mittelalter*, ed. Burkhard Mojsisch, 163–299. Amsterdam: B.R. Grüner.
Fleisch, Henri. 1961. *Traité de philologie arabe*. I. *Préliminaires, phonétique, morphologie nominale*. Beirut: Imprimerie Catholique.
Lloyd, Geoffrey E.R. 1966. *Polarity and analogy: Two types of argumentation in early Greek thought*. Cambridge: Cambridge University Press.
Majmaʿ al-Buḥūṯ al-ʾIslāmiyya. 1995. *Šarḥ al-muṣṭalaḥāt al-kalāmiyya*. Mašhad.
Maróth, Miklós. 1995. "The role of *qiyās* in Arabic grammar". *Acta Orientalia Academiae Scientiarum Hungaricae*, 48.101–108.
Maxzūmī, Mahdī al-. 1958. *Madrasat al-Kūfa wa-manhajuhā fī dirāsat al-luġa wa-n-naḥw*. 2nd ed. Cairo: Muṣṭafā al-Bābī al-Ḥalabī.
Naššār, ʿAlī Sāmī an-. 1947. *Manāhij al-baḥṯ ʿinda mufakkirī l-ʾIslām*. [Cairo.]
Suleiman, Yasir. 1999. *The Arabic grammatical tradition: A study in* taʿlīl. Edinburgh: Edinburgh University Press.
ʿUmarī, Nādiya Šarīf al-. 1987. *Našʾat al-qiyās al-ʾuṣūlī wa-taṭawwuruhu*. Gīzah.
Versteegh, C.H.M. (Kees). 1977. *Greek elements in Arabic linguistic thinking*. Leiden: E.J. Brill.
——. 1980. "The origin of the term *qiyās* in Arabic grammar". *Zeitschrift für Arabische Linguistik* 4.7–30.
——. 1993. *Arabic grammar and Qurʾānic exegesis in early Islam*. Leiden: E.J. Brill.
——. 1995. *The explanation of linguistic causes: Az-Zaǧǧāǧī's theory of grammar*. Amsterdam and Philadelphia: J. Benjamins.
——. 1997. *The Arabic linguistic tradition*. London and New York: Routledge.
——. 2001. *The Arabic language*. 2nd ed. Edinburgh: Edinburgh University Press.
Weil, Gotthold. 1913. *Die grammatischen Schulen von Kufa und Basra*. Leiden: E.J. Brill.

Miklós Maróth
(Pázmány Péter Catholic University)

Qualifier → X-bar Syntax

Quantifier

1. What is a Quantifier?

This entry surveys quantifiers in Arabic and classifies them according to their morphosyntactic behavior. Quantifiers are terms that express quantificational relations between sets, where sets are expressed by predicates, e.g. noun phrases and verb phrases. For example, the quantifier *most* in (1) expresses a relation between *Egyptians* (a noun phrase (NP) denoting the set of Egyptians) and *love Umm Kulthoum* (a verb phrase (VP) denoting the set of individuals who love Umm Kulthoum).

(1) *Most Egyptians love Umm Kulthoum*

The relation that *most* expresses is true if and only if the intersection of the two sets (the individuals who are both Egyptian and love Umm Kulthoum) accounts for more than half of the first set (the Egyptians). The two set-denoting phrases *Egyptians* and *love Umm Kulthoum* are arguments of *most*. Similarly, *Every Egyptian loves Umm Kulthoum* is true if the Egyptians are a subset of the individuals who love Umm Kulthoum (i.e. if there is no individual in the *Egyptian* set who is not also in the *loves Umm Kulthoum* set). *Some Egyptian* *loves Umm Kulthoum* is true if the intersection of *Egyptians* and *loves Umm Kulthoum* is not empty (i.e. if there is at least one individual in both the *Egyptian* set and the *loves Umm Kulthoum* set). Keenan (1996) formalizes such relations as conditions on the truth of sentences that contain them, as in (2). Here, *A* and *B* are any sets, $|\alpha|$ stands for 'the cardinality of α', $I(\alpha)$ stands for 'the interpretation of α', and *T* stands for the truth value 'true'. (2a), then, which defines *most*, says for any sets *A* and *B*, the interpretation of *most (A, B)* (read 'most As are Bs') is 'true' if and only if the cardinality of the intersection of *A* and *B* is greater than the cardinality of the set of things that are in *A* but not *B* (the set *A-B*), that is, in the example at hand, if the Egyptians who love Umm Kulthoum outnumber the Egyptians who do not.

(2a) For any sets A and B, $I(\text{most}(A, B)) = T$ if and only if $|A \cap B| > |A - B|$.
(2b) For any sets A and B, $I(\text{every}(A, B)) = T$ if and only if $A \subseteq B$.
(2c) For any sets A and B, $I(\text{some}(A, B)) = T$ if and only if $A \cap B \neq \emptyset$.

The definitions in (2) follow the model of the generalization of the logical quantifiers *all* and *some* devised by Mostowski (1957). Such 'generalized quantifiers' have played an important role in the development of modern logic; see Lindström (1966), Montague (1970), Barwise and Cooper (1981), Higginbotham and May (1981), and Keenan and Stavi (1986).

A quantifier's first argument is its *restriction* (*R* in (3)); it restricts the universe to that portion that needs to be considered in order to evaluate whether the quantificational relation is true or not (non-Egyptians have no bearing on the truth of *Most Egyptians love Umm Kulthoum*). A quantifier's second argument is its *nuclear scope* (*NS* in (3)); it constitutes that portion of the scope of the quantifier not included in the restriction (Heim 1982).

(3) $[_{\text{Q}}$Most] $[_{\text{R}}$Egyptians] $[_{\text{NS}}$love Umm Kulthoum]

Note, though, that quantifiers may be more-than-two-place, as in *More students than teachers came to the party* (also see Sec. 4, below), which describes a relation between the set of students, the set of teachers, and the set of individuals who came to the party (Keenan 1996).

Morphological and syntactic criteria distinguish three classes of quantifiers in Arabic: (1) nominal quantifiers (e.g. *jamīʿ* '*all*'), (2) numerals (e.g. *ṯalāṯ* '*three*'), and (3) phrasal quantifiers (quantificational relationships expressed periphrastically). These are surveyed below.

2. Nominal Quantifiers

A defining characteristic of the nominal quantifiers is a parallelism to the syntactic expression of possession in Arabic. Possession is expressed through the → 'construct state', in which the NP denoting the possessee is the first term and the NP denoting the possessor the second. The first term of a construct state bears the case morphology corresponding to the grammatical function of the construct as a whole, while the second term bears genitive case. The first term may bear neither the definite article nor the → *tanwīn* (*-n* ending; glossed *Twn* below) that typically occurs on nouns in the absence of the article, and the two terms must be adjacent. See Ritter (1987; building on Abney 1987), Ritter (1988), Ouhalla (1988), Benmamoun (1992), Mohammad (1988, 1989), Fassi Fehri (1999), and Shlonsky (2004).

(4) *bāb-u* *l-ḥadīqat-i*
gate-Nom the-garden-Gen
'The gate of the garden; the garden's gate'

Nominal quantifiers occur as the first term of a construct state construction of which the second term is the NP denoting the quantifier's restriction. Here again, the first term, the quantifier, may bear neither the definite article nor the *tanwīn*.

(5a) *jamīʿ-u* *l-kutub-i*
all-Nom the-books-Gen
'all the books'
(5b) *muʿḏam-u l-kutub-i*
most-Nom the-books-Gen
'most of the books'
(5c) *baʿḍ-u* *l-kutub-i*
some-Nom the-books-Gen
'some of the books'

In addition to the morphosyntactic parallels between (4) and (5), quantificational and possessive constructions are parallel in the replaceability of the second term by a pronominal suffix.

(6a) *bāb-u-hā*
gate-Nom-3p
'their gates'
(6b) *jamīʿ-u-hā*
all-Nom-3p
'all of them; their entirety'
(6c) *muʿḏ̣am-u-hā*
most-Nom-3p
'most of them'
(6d) *baʿḍ-u-hā*
some-Nom-3p
'some of them'

The parallel between (4) and (5) and (6a) and (6b–d), in light of the fact that the first term of the construct state is a noun, suggests that there is something nominal about these quantifiers. The nominal character is also evident in morphological parallels between the quantifiers in (5) and other nouns. The quantifiers bear case morphology as nouns do, and, when the restriction is unexpressed, bear either the definite article or the *tanwīn* that occurs on nouns in the absence of the article.

(7a) *qaraʾ-tu* *l-jamīʿ-a*
read-1s the-all-Acc
'I read all of it/them' (i.e. 'I read it/their entirety')
(7b) *qaraʾ-tu* *baʿḍ-a-n*
read-1s some-Acc-Twn
'I read some/several'

Arabic differs in this respect from English, in which quantificational relationships are typically expressed by determiners (*every, most, some*, etc.). In Arabic, they are typically expressed by nouns.

Also like in possessive constructions, the second term of a construct state headed by a quantifier may be definite or indefinite, singular or plural. Interpretations for the quantifiers in (5) for the four possible definiteness/number combinations in the restriction are paraphrased in English as follows. Note that when the restriction of a nominal quantifier is indefinite, its interpretation typically must be restricted in some other way, as by an adjectival modifier and (the *b* and *d* examples, below).

(8a) *qaraʾ-tu* *jamīʿ-a* *l-kitāb-i*
read-1s all-Acc the-book-Gen
'I read all of the book'
(8b) *qaraʾ-tu* *jamīʿ-a* *kitāb-i-n*
read-1s all-Acc book-Gen-Twn
mamnūʿ-i-n
banned-Gen-Twn
'I read all of a banned book'
(8c) *qaraʾ-tu* *jamīʿ-a* *l-kutub-i*
read-1s all-Acc the-books-Gen
'I read all of the books'
(8d) *qaraʾ-tu* *jamīʿ-a* *kutub-i-n*
read-1s all-Acc books-Gen-Twn
mamnūʿ-at-i-n
banned-p-Gen-Twn
'I read all banned books'

(9a) *qaraʾ-tu* *muʿḏ̣am-a* *l-kitāb-i*
read-1s most-Acc the-book-Gen
'I read most of the book'
(9b) *qaraʾ-tu* *muʿḏ̣am-a* *kitāb-i-n*
read-1s most-Acc book-Gen-Twn
mamnūʿ-i-n
banned-Gen-Twn
'I read most of a banned book'
(9c) *qaraʾ-tu* *muʿḏ̣am-a* *l-kutub-i*
read-1s most-Acc the-books-Gen
'I read most of the books'
(9d) *qaraʾ-tu* *muʿḏ̣am-a* *kutub-i-n*
read-1s most-Acc books-Gen-Twn
mamnūʿ-at-i-n
banned-p-Gen-Twn
'I read most banned books'

(10a) *qaraʾ-tu* *baʿḍ-a* *l-kitāb-i*
read-1s some-Acc the-book-Gen
'I read part of the book'
(10b) *qaraʾ-tu* *baʿḍ-a* *kitāb-i-n*
read-1s some-Acc book-Gen-Twn
mumnūʿ-i-n
banned-Gen-Twn
'I read part of a banned book'
(10c) *qaraʾ-tu* *baʿḍ-a* *l-kutub-i*
read-1s some-Acc the-books-Gen
'I read some of the books'
(10d) *qaraʾ-tu* *baʿḍ-a* *kutub-i-n*
read-1s some-Acc books-Gen-Twn
mamnūʿ-at-i-n
banned-p-Gen-Twn
'I read some banned books'

These interpretations are largely as expected, assuming first that a singular noun denotes a set of subparts of a single individual (*book*, above), while a plural denotes a set of individuals, and second that the definite article in the restriction has the effect of relating the restriction to a discourse-salient set. Hence, the *a* and *c* examples (with a singular restriction) describe a relationship between subparts of a book and a property (here *that I read it*), while the *b* and *d* examples (with a plural restriction) describe a relationship between books and a property.

An exception to this very regular pattern is the quantifier *kull* (roughly, 'every'). *Kull* is special in that its interpretation with definite, indefinite, singular, and plural restrictions does not compose in the same manner as for the other quantifiers. In combination with a definite restriction, it is parallel in meaning to *jamīˁ* 'all' (compare (11a, c) with (8a, c)). However, in combination with a singular indefinite restriction, its restriction is interpreted not as a set of subparts of an individual but rather as a set of individuals (11b), just like a plural definite restriction (11c). And in combination with a plural indefinite restriction, it is ungrammatical altogether (11d). Note lastly that *kull* is also unlike *jamīˁ* in that a singular indefinite restriction for *kull* need not be further modified (compare (11b) with (8b)).

(11a) *qaraʾ-tu* *kull-a* *l-kitāb-i*
read-1s every-Acc the-book-Gen
'I read all of the book'

(11b) *qaraʾ-tu* *kull-a* *kitāb-i-n*
read-1s every-Acc book-Gen-Twn
'I read every book'

(11c) *qaraʾ-tu* *kull-a* *l-kutub-i*
read-1s every-Acc the-books-Gen
'I read all of the books'

(11d) **qaraʾ-tu* *kull-a* *kutub-i-n*
read-1s every-Acc book-Gen-Twn
(mamnūˁ-at-i-n)
banned-p-Gen-Twn

Kull alone has the property that it forces a set-of-individuals interpretation on a singular indefinite restriction, but even this quirk vanishes when the restriction is definite (compare (11b) with (11a)). As with other quantifiers, the restriction of *kull* can be pronominalized (12a) or null (12b, c) (compare (12) with (6) and (7)). Note that the fact that (12a) is interpreted on par with (11a), not (11b), indicates that pronouns are inherently definite in Arabic.

(12a) *qaraʾ-tu* *kull-a-hu*
read-1s every-Acc-3s
'I read all of it' (not: 'I read each one')

(12b) *qaraʾ-tu* *kull-a-n*
read-1s every-Acc-Twn
'I read each one' (not: 'I read all of it')

(12c) *qaraʾ-tu* *l-kull-a*
read-1s the-every-Acc
'I read all of it' (not: 'I read each one')

A null restriction may 'reappear' in a prepositional phrase.

(13) *qaraʾ-tu* *kull-a-n* *min*
read-1s every-Acc-Twn of
al-kutub-i
the-books-Gen
'I read each one of the books'

The possibility illustrated in (13) of separating the restriction out into a prepositional phrase avails itself generally in Arabic for most quantifiers of any type. In this respect, also, the syntactic relationship between the quantifier and its restriction parallels possession. (4) may also be paraphrased as in (14), although the head in this case is interpreted as indefinite (→ construct state).

(14) *bāb-u-n* *min* *al-ḥadīqat-i*
gate-Nom-Twn of the-garden-Gen
'a gate of the garden'

The quantifier *kilā* 'both' occurs with a dual restriction, which must be definite. It is the only nominal quantifier that agrees in gender with its restriction. The form *kilā* occurs with a masculine restriction (15a) and *kiltā* with a feminine (15b).

(15a) *qaraʾ-tu* *kilā* *l-kitāb-ayni*
read-1s both the-book-d/Acc
'I read both books'

(15b) *qaraʾ-tu* *kiltā* *r-risāl-at-ayni*
read-1s both/f the-letter-f-d/Acc
'I read both letters'

The nominal class also includes the proportional quantifiers (e.g. *one third of*). Like the other nominal quantifiers, these occur in construct with their restriction, which can be pronominalized or dropped, as shown in (17), and distribute like nouns.

(16a) *qaraʾ-tu* *ṯulṯ-a* *l-kitāb-i*
read-1s third-Acc the-book-Gen
'I read a third of the book'
(16b) *qaraʾ-tu* *ṯulṯ-a* *kitāb-i-n*
read-1s third-Acc book-Gen-Twn
mamnūʿ-i-n
banned-Gen-Twn
'I read a third of a banned book'
(16c) *qaraʾ-tu* *ṯulṯ-a* *l-kutub-i*
read-1s third-Acc the-books-Gen
'I read a third of the books'
(16d) *qaraʾ-tu* *ṯulṯ-a* *kutub-i-n*
read-1s third-Acc books-Gen-Twn
mamnūʿ-at-i-n
banned-f-Gen-Twn
'I read a third of banned books'

(17a) *qaraʾ-tu* *ṯulṯ-a-hu*
read-1s third-Acc-3s
'I read a third of it'
(17b) *qaraʾ-tu* *ṯ-ṯulṯ-a*
read-1s the-third-Acc
'I read the third'
(17c) *qaraʾ-tu* *ṯulṯ-a-n* *(min-hu)*
read-1s third-Acc-Twn (of-it)
'I read a third (of it)'

3. Numeric Quantifiers

Like other quantifiers, the cardinal numbers denote relations between sets (→ numerals). (18) asserts that the intersection of the set of students and the set of individuals who passed the exam has cardinality three.

(18) *najaḥ-a* *ṯalāṯ-at-u*
succeed-3s three-f-Nom
ṭullāb-i-n *fī* *l-imtiḥān-i*
students-Gen-Twn in the-exam-Gen
'Three students passed the test'

That is, *ṯalāṯ* has the truth conditions in (19).

(19) For any sets A and B, *I*(*ṯalāṯ*(A, B)) = *T* if and only if $|A \cap B| \geq 3$.

Like the nominal quantifiers, quantifiers formed from numerals may have any grammatical function. The numerals from 1 to 10, as well as 100, 1,000, and 1,000,000, occur in construct with their restriction, which, as usual, appears in the genitive case. Beyond these similarities, the numeric quantifiers differ from the nominal quantifiers in a number of respects that suggest they are taxonomically special.

First, the cardinal numbers mentioned above agree in gender with their restriction (except for *miʾat* '100', which is feminine and indeclinable), that is, the form of the numeral depends on the gender of the restriction, albeit in an unusual way. The numeral bears the opposite gender morphology of the noun that forms the restriction. Note that this gender 'polarity' effect, typical of the Semitic languages (Hetzron 1967, 1972), does not apply to the one agreeing nominal quantifier *kilā* 'both'.

(20a) *ṯalāṯ-at-u* *ṭullāb-i-n*
three-f-Nom students-Gen-Twn
'three (male) students'
(20b) *ṯalaṯ-u* *ṭālib-āt-i-n*
three-Nom student-f/p-Gen-Twn
'three (female) students'

Second, unlike the nominal quantifiers, the numerals may occur as adjectival modifiers of their restriction, whether definite or indefinite (Wright 1981:11, Sec. 321).

(21a) *najaḥ-a* *ṭullāb-u-n*
succeed-3s students-Nom-Twn
ṯalāṯ-at-u-n *fī* *l-imtiḥān-i*
three-f-Nom-Twn in the-exam-Gen
'Three students passed the test'
(21b) *najaḥ-a* *ṭ-ṭullāb-u*
succeed-3s the-students-Nom
ṯ-ṯalaṯ-at-u *fī* *l-imtiḥān-i*
the-three-f-Nom in the-exam-Gen
'The three students passed the test'

Here, the numeral functions as an adjective modifying *(aṭ)-ṭullābu-(n)*, agreeing with it in case, definiteness, and gender, as required of adjectival modifiers (although the gender polarity principle still applies here, and not to adjectival modification in general).

Third, although a numeral may occur in construct with a definite restriction or bear a pronominal suffix (Wright 1981:111, Sec. 107), the interpretation is not the expected one, given the pattern established by the contrast in (5) and (6).

(22a) *najaḥ-a ṯalāṯ-at-u*
succeed-3s three-f-Nom
ṭ-ṭullāb-i fī l-imtiḥān-i
the-students-Gen in the-exam-Gen
'The three students passed the test' (≠ 'Three of the students passed the test')
(22b) *najaḥ-a ṯalāṯ-at-u-hum*
succeed-3s three-f-Nom-3p
fī l-imtiḥān-i
in the-exam-Gen
'The three of them passed the test' (≠ 'Three of them passed the test')

The usual partitivity associated with the construct state does not carry over to numerals in construct with a definite noun. Semantically, *ṯalāṯ* has the function in (22) of an adnominal modifier (as in *they three*...), indeed, the same function as its adjectival counterpart in (21b). The partitive interpretation with a definite restriction is expressed with the restriction in a prepositional phrase dependent of the numeral.

(23) *najaḥ-a ṯalāṯ-at-u-n min*
succeed-3s three-f-Nom-Twn of
aṭ-ṭullāb-i fī l-imtiḥān-i
the-students-Gen in the-exam-Gen
'Three of the students passed the test'

In these respects, the numerals do not pattern together with the nominal quantifiers, nor does the one similarity between the numerals and the nominal quantifiers – their occurrence in the construct state – extend beyond the numerals mentioned previously. The numerals between 10 and 100 obligatorily precede their restriction, which is indefinite, accusative, and singular.

(24) *najaḥ-a ṯalāṯ-at-u-n wa-*
succeed-3s three-f-Nom-Twn and-
ʿišrū-na ṭālib-a-n
twenty-Twn student-Acc-Twn
'Twenty-three students passed'

See Ziadeh and Winder (1957) for a cogent discussion of additional properties of the Arabic numerals.

4. Phrasal quantifiers

In English, some quantificational relations are expressed by what one might call 'discontinuous determiners', such as *more*...*than*...(as in *More Egyptians than Iraqis love Umm Kulthoum), exactly as many*...*as*...(as in *Exactly as many Egyptians as Iraqis love Umm Kulthoum), three more*...*than*...(as in *Three more Egyptians than Iraqis love Umm Kulthoum), every*...*except*...(as in *Every Egyptian except Salma loves Umm Kulthoum*), etc. Such relations are expressed periphrastically in Arabic, as illustrated below.

(25) *ʿadad-u l-muḥām-īna*
number-Nom the-lawyer-p/Gen
l-laḏīna ya-qraʾ-ūna jarīdat-a
the-which 3-read-p newspaper-Acc
l-quds-i ʾakṯar-u min
the-Quds-Gen more-Nom than
ʿadad-i l-ʾaṭibbāʾ-i
number-Gen the-doctors-Gen
'More lawyers than doctors read Al-Quds' (lit. 'The number of lawyers who read Al-Quds is more than the number of doctors')

(26) *ʿadad-u l-ʾaṭibbāʾ-i*
number-Nom the-doctors-Gen
l-laḏīna ya-qraʾ-ūna jarīdat-a
the-which 3-read-p newspaper-Acc
l-quds-i yu-sāwī ʿadad-a
the-Quds-Gen 3-equal number-Acc
l-muḥām-īna
the-lawyer-p/Gen
'As many doctors as lawyers read Al-Quds' (lit. 'The number of doctors who read Al-Quds equals the number of lawyers')

(27) *ya-zīd-u ʿadad-u*
3-exceed-Ind number-Nom
l-ʾaṭibbāʾ-i l-laḏīna ya-qraʾ-ūna
the-doctors-Gen the-which 3-read-p
jarīdat-a al-quds-i
paper-Acc the-Quds-Gen
ṯalāṯ-at-a-n ʿan ʿadad-i
three-f-Acc-Twn on number-Gen

l-muḥām-īna
the-lawyer-p/Gen
‘Three more doctors than lawyers read Al-Quds’ (lit. ‘The number of doctors who read Al-Quds exceeds the number of lawyers by three’)

(28) *kull-u* *l-ʾaṭibbāʾ-i* *ya-qraʾ-ūna*
every-Nom the-doctors-Gen 3-read-p
jarīdat-a *l-quds-i*
newspaper-Acc the-Quds-Gen
ʾillā *salmā*
except Salma
‘All the doctors read Al-Quds except Salma’

Proportions named explicitly as percents are expressed as…*in 100* in Arabic, as in (29).

(29) *najaḥ-a* *ṯamānū-na* *fī*
succeed-3s eighty-Twn in
l-miʾat-i *min* *aṭ-ṭullāb-i*
the-hundred-Gen of the-students-Gen
‘Eighty percent of the students passed’

5. SUMMARY

Quantifiers denote relations between sets. Nominal quantifiers are morphosyntactically nouns and occur as the first term of a construct state construction of which the second term is the restriction, which may be definite or indefinite, singular or plural. *Kull* is a unique quantifier that may combine with a definite or indefinite restriction, but with an indefinite only in the singular, and with a unique interpretation vis-à-vis the other quantifiers. The numeric quantifiers are quasi-adjectival. They agree with their restriction. The basic ones occur either in construct with their restriction or as adjectival modifiers of it, and when the restriction is definite, they are semantically modificational, not partitive. Other kinds of quantificational relations are expressed altogether nonlexically.

BIBLIOGRAPHICAL REFERENCES

Abney, Steven. 1987. *The English noun phrase in its sentential aspect.* Ph.D. diss., Massachusetts Institute of Technology.

Barwise, Jon and Robin Cooper. 1981. “Generalized quantifiers and natural language”. *Linguistics and Philosophy* 4.159–218.

Benmamoun, Abbas. 1992. *Functional and inflectional morphology: Problems of projection, representation and derivation.* Ph.D. diss., University of Southern California.

Fassi Fehri, Abdelkader. 1999. “Arabic modifying adjectives and DP structures”. *Studia Linguistica* 53.105–154.

Heim, Irene. 1982. *The semantics of definite and indefinite noun phrases.* Ph.D. diss., University of Massachusetts, Amherst.

Hetzron, Robert. 1967. “Agaw numerals and incongruence in Semitic”. *Journal of Semitic Studies* 12.169–193.

——. 1972. “Phonology in syntax”. *Journal of Linguistics* 8.301–357.

Higginbotham, James and Robert May. 1981. “Questions, quantifiers and crossing”. *Linguistic Review* 1.41–80.

Keenan, Edward. 1996. “The semantics of determiners”. *The handbook of contemporary semantic theory*, ed. Shalom Lappin, 41–63. Oxford: Blackwell.

—— and Jonathan Stavi. 1986. “A semantic characterization of natural language determiners”. *Linguistics and Philosophy* 9.253–326.

Lindström, Per. 1966. “First order predicate logic with generalized quantifiers”. *Theoria* 32.186–195.

Mohammad, Mohammad. 1988. “On the parallelism between IP and DP”. *Proceedings of the Seventh West Coast Conference on Formal Linguistics*, ed. Hagit Borer, 241–254. Stanford, Cal.: CSLI Publications.

——. 1989. *The sentential structure of Arabic.* Ph.D. diss., University of Southern California.

Montague, Richard. 1970. “Universal grammar”. *Theoria* 36.373–398.

Mostowski, Andrzej. 1957. “On a generalization of quantifiers”. *Fundamenta Mathematicae* 44.12–36.

Ouhalla, Jamal. 1988. *The syntax of head movement: A study of Berber.* Ph.D. diss., University College, London.

Ritter, Elizabeth. 1987. “NSO orders in modern Hebrew”. *Proceedings of NELS 17*, ed. Joyce McDonough and Bernadette Plunkett, 521–537. Somerville, Mass.: Cascadilla Press.

——. 1988. “A head-movement approach to construct state noun phrases”. *Linguistics* 26.909–929.

Shlonsky, Ur. 2004. “The form of Semitic noun phrases”. *Lingua* 114.1465–1526.

Wright, William. 1981. *A grammar of the Arabic language.* 3rd ed. Repr. 2 vols. Beirut: Librairie du Liban.

Ziadeh, Farhat J. and R. Bayly Winder. 1957. *An introduction to Modern Arabic.* Princeton, N.J.: Princeton University Press.

PETER HALLMAN (McGill University)

Question → WH-Movement

Qurʾān

The *Qurʾān* uses the fact that it was revealed in Arabic to argue that the Arabs who hear it have no excuse for not understanding: "We have sent it down as an Arabic *Qurʾān* so that you [people] may understand" (Q. 12/2). As Sībawayhi (d. 177/793?) puts it, "They were addressed in their own speech, and the *Qurʾān* came to conform to their language and what they understood" (*Kitāb* I, 331–332). Similarly, aš-Šāṭibī (d. 790/1388) states: "The *Qurʾān* was revealed in Arabic, conforming to the way the Arabs spoke, making it easy for them to understand what God commands and prohibits" (*Muwāfaqāt* III, 346).

The *Qurʾān* conformed to Arab speech: it provided the reason for codifying Arabic grammar and stylistics and was used as a criterion for these disciplines. It was in order to make sure that all Muslims, especially non-Arabs, could read the *Qurʾān* accurately that Arabs began to think of writing down a grammar of their language, based on the text of the *Qurʾān* and on Arabic poetry and speech. Arabic phonetics originated in the quest to describe exactly the proper pronunciation and articulation of the *Qurʾān* and developed into the science of → *tajwīd* and the art of reciting the *Qurʾān*. Similarly, the study of Arabic rhetoric flourished for the service of the *Qurʾān*, culminating in the works of the outstanding scholar ʿAbd al-Qāhir al-Jurjānī (d. 471/1078), particularly his *Dalāʾil al-ʾiʿjāz* and *ʾAsrār al-balāġa*, which investigate the style and rhetoric of the *Qurʾān* in an attempt to identify the reasons for the inimitability (→ *ʾiʿjāz*) of its language. In his opinion, the central issue of Arabic grammar is *taʿlīq* 'syntactic relations', which is closely bound up with stylistics. In fact, it is only by studying these two disciplines together in the *Qurʾān* that one gains a true understanding of how they work.

Even in their study of the principles of literary criticism, Arab critics drew heavily on the text of the *Qurʾān*, as can be seen, for example, in the celebrated work of the critic Ibn al-ʾAṯīr (d. 637/1239) in his *al-Maṯal as-sāʾir fī ʾadab al-kātib wa-š-šāʿir* on the criticism of prose and poetry. As as-Suyūṭī (d. 911/1505), a writer on many subjects, including Qurʾānic studies and linguistics, states, the *Qurʾān* was in fact the source for all branches of Arab knowledge (*ʾItqān* IV, 33). In the Islamic educational system, grammar instruction is connected with the *Qurʾān*: in the first grammar lessons at the primary school of al-Azhar in Cairo (opened in 972), students are introduced to *al-mabādiʾ al-ʿašara* 'the ten introductory aspects', with which the grammar textbook begins, and which are necessary when embarking on the study of any branch of knowledge. They include the definition of that particular discipline, its name, the issues it deals with, its founder, its relation to other subjects, its status among other subjects, and how the *šarīʿa* views the learning of that subject. The status of grammar, students are told, is foremost because it ensures correct understanding and reading of the Arabic text of the *Qurʾān*, and as such, its study is incumbent on the Muslim community. In addition to any educational or secular purpose, learning grammar and linguistics in general is a religious duty in al-Azhar and similar traditional institutions of learning in the Muslim world.

Thus, in the Muslim world, the *Qurʾān* is the starting point of Arabic linguistic scholarship. Outside the Muslim world, another vigorous tradition of Arabic linguistic scholarship has developed in Western universities (especially since the 17th century), which from the beginning was driven in no small measure by its relevance to the *Qurʾān* and Islam (→ Arabic studies in Europe). This has led to the production of monumental studies of Arabic and Islam in leading European centers of learning, in various languages, such as Dutch, English, French, German, Italian, Russian, and Spanish.

1. GRAMMAR AND STYLE

It would be incorrect to talk about the grammar of the Arabic *Qurʾān* as being different from general Arabic grammar: no scholar, neither Arab nor non-Arab, has written as a separate study a comprehensive grammar of the *Qurʾān*. The *Qurʾān* conforms to the grammatical norms of Arabic and includes some dialectal features in the different readings, all of which were understood by the Arabs who first heard the Revelation. Around the Prophet, there were Muslims from all parts of Arabia, and there are no comments on record that any of the *Qurʾān*'s grammatical features were found to be foreign to the Arabic language people were used to (ʿUbāda 1980:I, 68–108).

When grammarians devised their rules based on what is general and regular in Arabic (*muṭṭarid*; see below, Sec. 2.2), and found that certain instances in the *Qurʾān*, notably various dialectal features, did not conform to the general pattern, they still recognized that these do conform to other recognized grammatical or stylistic patterns in the language, and thus, even if they were not included as part of the *muṭṭarid* rules of grammar (Ḥassān 1993:283) and were considered *šāḏḏ*, they were still correct (→ *faṣīḥ*) since they belonged to the language of the age of *iḥtijāj*. This age is normally considered to be the period up to the middle of the 2nd century A.H. in urban centers and the mid-4th century A.H. in Bedouin localities. The language of this era is regarded as correct and fit to be used as a linguistic model.

The Arabic language is broader than the *muṭṭarid* rules, and irregularity does not preclude correctness (*aš-šuḏūḏ lā yunāfī l-faṣāḥa*). Rather than speaking of the grammar of the *Qurʾān*, then, it would be more accurate to speak of the *Qurʾān*'s ways of using Arabic grammatical patterns in expressing its message and how these work in tandem with stylistic considerations. Various grammatical features in the *Qurʾān* are dealt with here. A scheme that may help the discussion to be more inclusive and systematic and that concentrates on the central issue of Arabic grammar (the *taʿlīq*) is to look at the cues of syntactic relations and how the *Qurʾān* deals with them, and then to look at some patterns the *Qurʾān* frequently uses.

2. Cues of syntactic relations

Like all other languages, Arabic offers a set of cues for determining syntactic relations between the different parts of a sentence. These can be either morphosyntactic or semantic. In Arabic, morphosyntactic cues have been categorized as follows (Ḥassān 1993:7–8):

i. Morphological structure (*binya*) of a word
ii. Declension (*ʾiʿrāb*)
iii. Binding (*rabṭ*)
iv. Order (*rutba*)
v. Syntagmatic requirement (*taḍāmm*)
vi. Verbal context (*siyāq*)

Semantic cues include such categories as predication, transitivity, and purpose, which are not dealt with here. These cues determine the grammatical function of words and make the meaning of the sentence clear. They act together in what is known in Arabic as *taḍāfur al-qarāʾin* 'interrelatedness of links'. This makes it sometimes possible to dispense with a given cue because other cues still make the grammatical function and meaning of the statement clear. The language of the *Qurʾān*, and Arabic in general, has norms for observing morphosyntactic cues or dispensing with them for stylistic purposes, in order to achieve certain effects that serve the intended message in any given situation.

2.1 *Morphology*

2.1.1 The morphological unit

A morphological unit can have multiple meanings. The definite article *al-* can refer to a class (*jinsiyya*), e.g. *al-ʾinsān* 'mankind/the human being' (Q. 70/19), or it can refer to something already mentioned or understood from the context (*ʿahdiyya*), e.g. *an-nās* 'the people of Mecca [mentioned earlier]' (Q. 17/93); *al-kitāb* 'the book [referring to the *Qurʾān*]' (Q. 5/48). *Al-* can also be used as a resumptive pronoun, e.g. Q. 9/26: "He punished the disbelievers – this is what the disbelievers [*al-kāfirīn* instead of 'they'] deserve", or it can be used for binding (*rabṭ*), e.g. *al-maʾwā* (Q. 79/39): "As for anyone who has transgressed and preferred the present life, hell will be home [*al-maʾwā*, meaning 'his home']".

This last usage serves to maintain rhythm and → rhyme, an important consideration in the *Qurʾān*, and gives the statement a generality that includes this person and any of his type. Generality of meaning (*taʿmim*) often has an important place in the *Qurʾān*.

2.1.2 Departure from regular morphology

i. Departure by omission

An instance of omission is that of the final *yāʾ* of the 1st person singular in *ʾujību daʿwata d-dāʿi ʾiḏā daʿāni*... 'When My servants ask you [Prophet] about Me, I am near, I respond to those who call Me...' (Q. 2/186). Here, the *yāʾ* has been omitted from *daʿāni* to leave only the *kasra*. Rhetorically, this serves the function of shortening the statement to express how close God is to anyone who calls Him, which is the point of this verse. This is enhanced by omitting the words 'Say to them' earlier in the sentence, as well as stated expressly in the words 'I am near'. The omission for stylistic effect does not impair the sense, as the *kasra* at the end of the two words indicates the original form of the word (the *yā'* is also omitted in *ad-dāʿi*). Again, in *rabbanā wa-taqabbal duʿā'i* 'Oh Lord, accept my request' (Q. 14/40), the final *yā'* is omitted in order to maintain the rhythm and rhyme (see also Q. 18/64, 66; 22/54; 40/15; 67/17).

ii. Departure by addition
Addition takes place, for instance in Q. 37/130: "Peace be to 'Ilyāsīn" instead of "'Ilyās" (although ar-Rāzī suggests that he is 'Ilyās ibn Yāsīn), where the addition of *īn* maintains the rhythm and rhyme for effect.

2.2 *Declension* ('iʿrāb)

'Iʿrāb has occupied Arab grammarians to a great extent, even though in many situations it is not necessary as a cue for syntactic relations, as in pronouns and indeclinable nouns, in verbs III*w/y* (*nāqiṣ*), and in nouns that do not allow final vowels to appear (*mabnī*), which are frequent in Arabic. The *Qur'ān* conforms to the general rules of *'iʿrāb*, but there are notable examples where the general rules are set aside and less common dialectal patterns are used for stylistic purposes. One well-known case is the occurrence of a sound masculine plural in a list of entities, where there is a departure from coordination for rhetorical purposes. In Q. 5/69, "The believers, the Jews, the Christians and the Sabians – all those who believe in God and the last day – there is no fear for them, nor will they grieve", all the nouns are governed by *'inna* in the accusative, but 'the Sabians' appears in the nominative (*wa-ṣ-Ṣābi'ūn*). This has been taken to be an example of *iltifāt*, that is, a departure from the general rule for the purpose of highlighting the Sabians, in order to emphasize that even they, who are less close to the scriptural tradition, will have no fear or cause to grieve (Abdel Haleem 1992:426). Another example of shift can be seen in Q. 2/177, in this case from the expected nominative *aṣ-ṣābirūn* 'those who are steadfast', which would be in coordination with the other nouns, to the accusative *aṣ-ṣābirīn*, so as to highlight the importance of this particular quality in the situation.

A number of occurrences of *iltifāt* have been described as "linguistic errors" by John Burton (1988:177). He regards, for instance, the occurrence of *al-muqīmīna ṣ-ṣalāt* 'those who perform the prayers' instead of *al-muqīmūna ṣ-ṣalāt*, in the middle of a list of types in Q. 4/162, as a linguistic error, although this is recognized as an admissible, familiar pattern in Arabic ('Abū ʿUbayda, *Majāz* I, 165, 173) and as a form of *iltifāt* (Zarkašī, *Burhān* III, 325). Another example that Burton considers to be an error is Q. 20/63 *'inna hāḏāni la-sāḥirāni* 'These are certainly two sorcerers', said about Moses and Aaron by Pharaoh's sorcerers. Normally, *hāḏāni* is governed by *'inna* and should have been rendered as *hāḏayni* (i.e. in the accusative). However, there is an Arabic dialect known as *luġa man yulzimu l-muṯannā l-'alif*, in which the dual invariably has the nominative ending (*hāḏāni*; 'Abū ʿUbayda, *Majāz* II, 21–22). The reason for this departure from the normal rules is therefore the observance of *ḥikāya* 'quoting what someone has said', regardless of the normal rules of inflection, a dialectal feature that has been used here for rhetorical effect. The statement quoted is made by sorcerers, and in sorcery, as observed even now in Arabic, sound effect is important as part of the incantation. Here, we have three successive words all with long *'alif*, preceded and succeeded by *fatḥa* followed by *nūn: hāḏāni, sāḥirāni, yurīdāni*, so to change *hāḏāni* into *hāḏayni* would break the pattern. The general grammatical rule has been suspended here for a rhetorical purpose, and as in all these examples, the suspension of *'iʿrāb* has not impaired the meaning, because other cues determine the grammatical functions of the individual words. Cues cooperate in the sense that the structure, word order, syntagmatic requirement, etc. all indicate the grammatical function, even if the general rule of *'iʿrāb* is suspended. Another occurrence of suspension of *'iʿrāb* occurs in Q. 12/11, when Joseph's brothers say to their father, "Why do you not trust us (*lā ta'mannā*) with Joseph?". The verb should have been in the indicative mood, *ta'manūna*, but this rule was set aside for the purpose of → *'idġām* 'gemination of the two similar *nūn*s', which makes the pronunciation lighter (*'axaff*) in Arabic – an intended phonetic effect.

2.3 *Binding* (rabṭ)

Binding is achieved by reference, concordance, and particles.

2.3.1 Reference (*ʾiḥāla*)

ʾIḥāla can be achieved through the use of the following devices: repetition of the same word or phrase; personal, demonstrative, or relative pronouns; and the definite article, *al-*; sometimes, the binding element may be omitted.

i. Repetition

There are numerous examples of repetition in the *Qurʾān*: "Do they not see that God brings life into being and reproduces it. Truly this is easy for God" (Q. 29/19). "It was not without purpose that we created the heavens and earth and everything in between. That may be what the disbelievers assume. Woe to the disbelievers from Hellfire!" (Q. 38/27). This type is known in *balāġa* as *wadʿ aḏ̣-ḏ̣āhir mawḍiʿ aḍ-ḍamīr* 'using a noun in place of a pronoun'. Repetition of the name of God, as above, in an independent statement to indicate His capabilities or qualities, in contrast to the other gods of the polytheists, is especially common. Repeating 'the disbelievers' instead of using the pronoun 'them' indicates that their disbelief was the cause of their opinion and their doom (Abdel Haleem 1992:428–429). It also creates, as in the first case, an independent statement, quotable on its own and effective in religious discourse.

Repetition may also be used to recapitulate and refresh the memory when a sentence has become long. Again, this type of binding is quite common in the *Qurʾān*: "When there came to them a scripture confirming what they already had; when…; when…; when there came to them…, they disbelieved" (Q. 2/89). "Why, when the soul of a dying man comes up to his throat, while you merely gaze on – We are nearer to him than you, though you do not see Us – why, if you are not subjected, do you not restore his soul to him?" (Q. 56/83–87).

ii. Pronouns

The personal pronoun must refer to a referent, normally the nearest suitable preceding referent, and it must agree with this both in form and in meaning. Sometimes, the referent is not stated but rather deduced (*mutaṣayyad*), for example, "But if anyone knows that the testator has made a mistake, or done wrong, and so puts things right between them, he will incur no sin" (Q. 2/182). 'Them' stands here for the parties involved. Also, "If God were to punish people [at once] for the wrong they have done, there would not be a single creature left on its [i.e. the earth's] surface" (Q. 35/45), and "everyone on it perishes" (Q. 55/26) where 'it' is generally understood to mean 'the earth'. This is in accordance with the rule according to which 'it is permissible to omit what is understood', and it has the benefit of conciseness.

The principle of proximity of the referent can be dispensed with if the relationship is understood without ambiguity, for example: "There are lessons in the story of Joseph and his brothers for those who seek them. They said, 'Although we are many, Joseph and his brother are dear…'" (Q. 12/7–8). Here, 'they said' does not refer to the closest referent ('those who seek') but rather to 'his brothers'.

As regards agreement between the pronoun and its referent, the *Qurʾān* departs from this in numerous situations. First and foremost, this is the case of *iltifāt* 'grammatical shift' for rhetorical purposes, a widespread feature of Qurʾānic style, of which hundreds of examples can be cited. In discussing some examples of this feature, Nöldeke (1910) remarks (without referring to the Arabic term) that "the grammatical persons change from time to time in the *Qurʾān* in an unusual and not beautiful way". Arab writers, in contrast, see the matter differently. The critic Ibn al-ʾAṯīr, for instance, after studying this stylistic feature, classes it among the "remarkable things and exquisite subtleties we have found in the glorious *Qurʾān*" (*Maṯal* II, 43–45). In rhetorical treatises, *iltifāt* is called *šajāʿat al-ʿarabiyya* because in their opinion it demonstrates the daring nature of the Arabic language. If any daring is to be attached to *iltifāt*, it should, above all, be the daring of the language of the *Qurʾān*, since it employs this feature according to effective patterns and for stylistic reasons, more extensively and in more variations than does Arabic poetry. Ibn al-ʾAṯīr, writing about rhetoric in prose and poetry, derives most of his examples from the *Qurʾān*. The hundreds of examples in the *Qurʾān* clearly show that stylistic considerations can overrule grammar, but always for a rhetorical purpose, without impairing the sense or causing any ambiguity.

Six types of *iltifāt* have been identified, but only those relating to shifts in person as reflected in the pronoun and its referent are discussed here. (The other types of *iltifāt* are change in number, between singular, dual, and plural; change in addressee; change in the tense of the verb; change in case marker; and use of nouns in place of pronouns [Abdel Haleem 1993:431]). Change in person (between 1st, 2nd, and 3rd) is the most common occurrence of this type. For example, in "Who created the heavens and earth? Who sends down water from the sky for you – with which We cause gardens of delight to grow: you have no power to make the trees grow in them – is it another god beside God?" (Q. 27/60), a shift from 1st to 3rd person in the *pluralis majestatis* occurs at a crucial point for the listeners (see also Q. 14/4). There is a shift from 1st to 3rd person in "Give, out of what We have provided for you, before death comes to one of you and he says, 'Reprieve me Lord', but Allah reprieves no soul when its time comes" (Q. 63/10–11). This shift makes the final statement independent and absolute; it also indicates the contrast with other gods who do not have such power.

An example of a shift from 3rd to 2nd person is: "Praise be to God.... You alone we worship, You alone we ask for help" (Q. 1/1–4). After establishing that praise is only due to the Creator, who has such attributes to make Him truly worthy of praise and the only true source of help, the worshipper turns to address God for the rest of the sura to ask for His help. A further example, this one from 2nd to 3rd person, can be found in "It is God who has given you spouses...and from them he has given you children and grandchildren. How can they believe in falsehood and deny God's favours!" (Q. 16/72). In this shift, God turns to call everyone to witness the ingratitude of the people in question.

Similarly, in examples of other types we witness departure from grammatical rules for specific rhetorical purposes. Without such rhetorical purpose, departure from the normal rules would be *mumtaniʿ* 'inadmissible' according to the rules of rhetoric, and there is always a proviso that the departure does not cause any *labs* 'confusion, obscurity'.

On some occasions, there are two preceding referents and only one is selected, as in *ʾarākum qawman tajhalūna* 'I can see you are foolish people' (Q. 11/29), in which the verb is in the 2nd person rather than the 3rd (*yajhalūna*), in agreement with *ʾarākum* (2nd person), rather than with *qawman*, which is 3rd person. In certain cases, the departure can also be in agreement in number, for example: "They swear by God to please you [believers]. But God and His messenger has more right that they should please Him" (Q. 9/62), in which use of the singular pronoun avoids referring to God and His messenger with a single pronoun, which would detract from *tawḥīd* 'monotheism' – and pleasing God would in any case please the Prophet.

Agreement in definite/indefinite status can also sometimes be departed from when the indefinite has been defined by an adjective, which brings it closer to being definite, as, for example, in *waylun li-kulli humazatin lumazatin alladhī jamaʿa mālan* 'Woe to every fault-finding backbiter, who amasses wealth' (Q. 104/1–2; see also 50/32–33; 57/23–24). Departure can also be from expected gender agreement, as in "If We had wished, We could have sent them down a sign from heaven at which their necks would have stayed bowed down in utter humility" (Q. 26/4), where *xāḍiʿin* 'in humility [masc. pl.]', which refers to their necks, should be feminine, but the use of the masculine plural shows more humility by implicitly including not just the necks of the disbelievers but also the rest of their persons, and also maintains the rhyme, which is an important consideration. Nor is the sense marred. Another example is "On the day when you [Prophet] see the believers, both men and women, with their light streaming out before them..." (Q. 57/12), in which *yasʿā nūruhum* 'their light streaming out' takes the masculine pronoun, which refers both to men and women according to the principle of *taġlīb* 'grammatical preponderance' (see also Q. 57/13, 64/14). Finally, a construction known as *murāwaḥa* 'alternation' between genders for a specific consideration can be seen in Q. 33/31: "Whoever of you [wives of the Prophet] is obedient (*wa-man yaqnut*) to God and his Messenger and does good (*taʿmal ṣāliḥan*) We will give her a double reward...". Here, the first verb, *yaqnut*, is masculine in agreement with the adjacent pronoun *man*, while the second verb, *taʿmal*, is feminine in agreement with the wife.

The demonstrative pronoun (*ḏālika, ʾulāʾika*, etc.) is frequently used as a binder in the *Qurʾān*: "...Whoever does good and believes, be it a man or a woman, these will enter paradise and be provided for without measure" (Q. 40/40); "The ones who lower their voices in the presence of God's Messenger, these (*ʾulāʾika*) are the ones whose hearts God has proved to be aware..." (Q. 49/3). The rhetorical purpose of using the demonstrative pronoun is to emphasize and highlight the class of persons who are being talked about. Similarly, the relative pronoun highlights, emphasizes, and singles out for praise *allaḏīna ʾāmanū* 'those who believe', as in "Those who believe and do good deeds, We do not let the reward of those who do good go to waste" (Q. 18/30), or disparagement *allaḏīna kafarū* 'those who disbelieve', as in "When Our revelations are recited to them, you [Prophet] will recognize the denial in the faces of those who disbelieve..." (Q. 22/72). Again, this usage can make the final clause independently quotable.

iii. The definite article *al-* used as a binder

"God is the light of the heavens and the earth. His light is like a niche in which there is a lamp, the lamp inside a glass, the glass is like a glittering star..." (Q. 23/35). The repetition (using the definite article *al-* here rather than a pronoun) is particularly effective for the image and the way it is built up one layer inside the other (see also Q. 33/10). Another example is "For him who feared the meeting with his Lord and restrained the soul [i.e. his soul] (*wa-nahā n-nafsa*) from base desires, Paradise will be the home [i.e. his home]" (Q. 79/40–41). This latter example is one of several in which the use of *al-* instead of the pronoun helps maintain rhyme (cf. Sec. 2.1.1).

iv. Omission of the binding element

In cases where there is no ambiguity, a binding element can be omitted in accordance with the general rule 'There is no omission without an indicator to what is omitted' (*lā ḥaḏfa ʾillā bi-dalīl*). Pronouns in the 3rd person in particular provide many examples of this in the *Qurʾān*. For example, in Q. 2/75, "God is not unaware of what you do" is rendered as *mā yaʿmalūna* rather than *yaʿmalūnahu*. The omitted pronoun is clear from the context, and in many cases its omission preserves the rhyme. In the following case the rhyme is not involved: "If you do not do, and you will not do..." (Q. 2/24), meaning 'If you do not do that', where the pronoun 'that' is omitted.

2.3.2 Concordance

Binding is also realized by concordance (*muṭābaqa*) in gender, number, *ʾiʿrāb*, definiteness/indefiniteness, and person (1st, 2nd, or 3rd). Concordance in these various elements binds words together and helps determine their grammatical functions. Concordance in number still pertains when a noun is singular in form but has the sense of plural, as in *wa-naḥnu ʿuṣbatun* 'while we are a group' (Q. 12/8), where *ʿuṣbatun* in the singular has a plural meaning and agrees with the plural pronoun *naḥnu* 'we'. Likewise, *ʾantum qalīlun* 'You were few' (Q. 8/26) and *naḥnu jamīʿun* 'we together' (Q. 54/44).

The relative pronoun *man* is grammatically singular and should, therefore, be followed by a singular pronoun even when the plural is intended, e.g. "Did he (Korah) not know that God had destroyed many generations before him, who had greater power than him (*man huwa ʾašaddu minhu*)?", where *huwa* is a singular pronoun meaning 'many'.

2.3.3 Particles (*ʾadawāt*)

Particles are very important in binding Arabic sentences. The Arabic sentence, whether declarative, conditional, or affective, with its numerous subdivisions, relies, in the overwhelming majority of cases, on the particle to bind its parts and express its grammatical category (Ḥassān 1973:243–246, 1993:56–57). In Q. 6/107, *law šāʾa llāhu mā ʾašrakū*...'If it had been God's will, they would not have joined other gods with Him...', without the conditional particle *law*, the negative particle *mā* would have become *maṣdariyya* (an infinitive particle) to turn the meaning of the statement to 'God willed their joining others with Him'. Even when particles introduce nouns, as in the case of prepositions, conjunctions, and exceptions, etc., the particle still acts as a binder. As is amply illustrated in works like Ibn Hišām's *Muġnī l-labīb* and by ʿAmāyra and as-Sayyid (1988), the *Qurʾān* employs particles extensively in its discourse for various stylistic purposes, such as the achievement of conciseness and cohesion and various other stylistic purposes.

2.4 *Word order*

There are some instances in which the order of various lexical items in the Arabic sentence is fixed, as is the case in the positioning of prepositions and the nouns they govern, and of the various particles denoting exception, conjunction, and interrogation, all of which must precede the noun. An example from the *Qur'ān* of another type of fixed-order sentence is *laᶜanahu llāhu* (Q. 4/118), where the object is an inseparable pronoun and so has to be attached to the verb and come before the subject. But there are also cases where the order is not fixed, such as the positioning of the object in relation to the verb and its subject, or that of the predicate of the subject of a nominal sentence. In a non-fixed-order sentence, the object may be introduced first to serve a stylistic purpose of restriction (*ḥaṣr*), as in *'iyyāka naᶜbudu* 'You alone we worship' (Q. 1/5). Departure from original word order for rhetorical purpose is very common in the *Qur'ān*. Take, for example, Q. 37/86: *'a-'ifkan 'ālihatan dūna llāhi turīdūna* 'How can you choose false gods below God?'. The rhetorical question, put by Abraham to his people, shows objection to their worship of false gods, and the level of objection is reflected in the arrangement of the word order of the question itself, in diminishing strength. The strongest objection is that they are false; the fact that they are worshipping false gods comes second, especially as these false gods are inferior to God, and the final cause for objection is the fact that this has all been done on their own volition. In this case, an alternative word order such as *'a-turīdūna 'ālihatan dūna llāhi 'ifkan* would weaken the impact of both the statement and the objection being made (Ḥassān 1993:95). Finally, in certain circumstances, the usually unfixed order has to be fixed, as in the verbal sentence *ḍaraba mūsā ᶜīsā* 'Mūsā hit ᶜĪsā', in which, if the order were unfixed, the lack of *'iᶜrāb* would leave the statement open to ambiguity, hence the first noun has to be the subject.

In a sentence where a number of attributes of various kinds occur, the *Qur'ān* tends to arrange them in order of length with the shortest first. For example, Q. 40/28: *wa-qāla rajulun mu'minun min 'āli firᶜawna yaktumu 'īmānahu*…'A secret believer from Pharaoh's family said…'. Here, we have a single adjective *mu'minun*, coming before a prepositional phrase *min 'āli firᶜawna*, and then we have the adjectival clause, *yaktumu 'īmānahu*. This particular order serves to balance the sentence and maintain a good, effective rhythm (see also Q. 2/68–69 and 23/117).

2.5 *Syntagmatic requirement*

Syntagmatic requirement (*taḍāmm*) is another wide area where grammar and style work together to serve the text of the *Qur'ān*, including such features as omission, addition, separation, parenthesis, and abnormal syntax.

2.5.1 Omission

Omission is a huge area in the *Qur'ān*, as can be seen from such books as *Majāz al-Qur'ān* by Ibn ᶜAbd as-Salām (e.g. *Majāz* 26–46, 261–471). It serves conciseness, an important stylistic feature of the *Qur'an*. As always with the principle of *lā ḥaḏfa 'illā bi-dalīl*, there has to be something to indicate the omitted part, such as a well-recognized pattern of the word, a cue, or the context. Here are a few examples.

i. Omission of the first member of the construct phrase (*'iḍāfa*), in particular, occupies the largest space in *Majāz al-Qur'ān*, For example: 'Ask the town' meaning 'ask the people of the town' (Q. 12/83).
ii. Omission of the final radical of the active participle, as in *ad-dāᶜi* instead of *ad-dāᶜī* (Q. 2/186), as explained earlier, or *hādi* < *hādī* (Q. 22/54).
iii. Omission of the interrogative particle, as when Moses asks Pharaoh, *wa-tilka niᶜmatun tamunnuhā ᶜalayya* 'Is this a favour you reproach me for?' (Q. 26/22), without the interrogative *'a*, since the intonation suffices.
iv. Omission of a clause: "If one of you is sick or on a journey, then other days later" (Q. 2/184) where "and breaks the fast" is omitted (see also Q. 23/10, 20).
v. Sometimes multiple sentences may be omitted, and there are many examples of this in the story of Joseph. For example, the king's cup-bearer says, "I can tell you the meaning of this dream. Send me…" (Q. 12/45), and here the scene cuts immediately to his interview with Joseph, omitting all the possible narrative in between.

2.5.2 Addition

An example of addition is found in Q. 27/5,… *wa-hum fī l-ʾāxirati humu l-ʾaxsarūna* ‘…and they, in the life to come, they will be the ones who will lose most’. The pronoun *hum* ‘they’ is added before the end for further emphasis.

2.5.3 Separation (grammatical)

An example of grammatical separation is found in Q. 62/1, *Yusabbiḥu li-llahi mā fī s-samāwāti wa-mā fī l-ʾarḍi l-maliki l-quddūsi l-ʿazīzi l-ḥakīm* ‘Everything in the heavens and earth glorifies God, the Controller, the Holy One, the Almighty, the Wise’, where the subject is placed between God and the nouns that qualify Him, to give the sentence rhythm, rhyme, and balance and keep the emphasis on the nouns at the end.

2.5.4 Parenthesis

Parenthesis is a very widespread feature of the *Qurʾān*. Many examples of separation are brought about by the insertion of parenthetical statements, for example: “And those who implore God’s forgiveness for their sin, when they do wrong – and who forgives sins but God? – and never knowingly persist in doing wrong” (Q. 3/136). Normally, when the *Qurʾān* mentions a view or a statement with which it does not agree, it interrupts the flow of the sentence with a comment (Šāṭibī, *Muwāfaqāt* III, 353–354), as in the following verse: “They apportion to God a share of the produce and the livestock He created, saying, ‘This is for God’ – so they claim! – ‘and this is for our idols’” (Q. 60/136).

2.5.5 Abnormal syntax

For example, in *wa-ʾinna kullan lammā la-yuwaffiyannahum rabbuka ʾaʿmālahum*…‘To each your Lord has not yet: He certainly will repay them for their deeds…’ (Q. 11/111), the particle *lammā* ‘not yet’ should, grammatically speaking, introduce an imperfect verb in the jussive mood (*yuwaffihim*), but instead it comes before ‘He will certainly’, which indicates what has been omitted: ‘He has not yet, but certainly will, repay them’. The omission is obvious; it avoids the unnecessary repetition that would weaken the statement, and the unusual syntax makes the statement more powerful and the threat more potent.

2.6 *Context* (siyāq)

Context is of crucial importance in determining meaning and syntactic relations. When other cues are departed from, which may result in ambiguity, context resolves this, and thus context has been rightly considered one of the key instruments in exegesis (Abdel Haleem 1993:71–98). There are numerous examples in the *Qurʾān*, of which only three will be cited here. In Q. 16/5, *wa-l-ʾanʿāma xalaqahā lakum fīhā difʾun wa-manāfiʿu*…‘And livestock – He created them for you too. You derive warmth and other benefits from them…’, grammatically, the phrase *lakum* may be governed by the verb *xalaqa*, where the meaning would be ‘He created the cattle for you, in them there is warmth and…’, or it could be the predicate of *difʾun*: ‘there is warmth…in them for you’. The context resolves this ambiguity, because the following verse has ‘and there is beauty in them for you’, which parallels ‘there is warmth in them for you’.

In Q. 13/11, *lahu muʿaqqibātun min bayni yadayhi wa-min xalfihi yaḥfaḏ̣ūnahu min ʾamri llāhi*…, ‘Each person has guardian angels before him and behind, watching over him, by God’s command…’, *min ʾamri llāhi* follows the verb *yaḥfaḏ̣ūnahu* and could be interpreted as grammatically governed by it, meaning ‘they protect him from God’s command’, but the correct reading is to see it as relating to *muʿaqqibātun*, even though it is distant from it, as becomes clear from the context at the end of the verse: ‘when God wills harm on a people, no one can ward it off’.

The above two examples work in a local context, but an example of the importance of global context can be seen in Q. 6/37–38, where the disbelievers have demanded that the Prophet should produce a miraculous sign, and he is instructed to say, “God is able to send down a sign”. Then comes the comment “All the creatures that crawl on the earth and those that fly with their wings are communities like yourselves. We missed nothing out in the book (*al-kitāb*) and in the end they will be gathered to their Lord”. Here, many exegetes have understood the word *al-kitāb* to be the *Qurʾān* and relate this to the preceding statement. They argue that this means that the *Qurʾān* contains everything, including all the animals and birds, and go on to say that these will be gathered before God, which is irrelevant to the context. Clearly there is a misreading here. The statement should stop at 'communities like yourselves'. Here they are reminded that all the creatures God created are miraculous signs. The rest of the verse is actually a retort by God, addressed to the intransigent disbelievers, and the book (*al-kitāb*) means the records of deeds. He warns that all that the disbelievers say is recorded against them, and they will eventually come up for judgment.

3. FREQUENTLY OCCURRING LINGUISTIC PATTERNS IN THE *QUR'ĀN*

3.1 *Affective sentences* (jumal ʾinšāʾiyya)

In addition to the declarative sentence (*xabariyya*), the *Qur'ān* frequently uses affective sentences. This serves to involve the reader or listener, a very important consideration in Qur'ānic discourse. This explains the frequent occurrence of the imperative, prohibition, interrogative, proposition, exhortation, wishing, hoping, supplication, exclamation, and oaths.

3.2 *Verbal sentences*

The *Qur'ān* employs the power of the verbal sentence, using the past tense for historical accounts in its argumentation and also when discussing the afterlife. This is effective in making such momentous events of the afterlife (mentioned directly or indirectly on almost every page of the *Qur'ān*) seem as if they are already here, a device crucial for Qur'ānic discourse and techniques of persuasion. This may involve *iltifāt* shift in tense as, for example, in Q. 20/125–126 and Q. 40/48–50.

3.3 *Generalization*

The *Qur'ān* frequently uses generalization, since it maintains that it is for all people. It classifies people, using such plurals as *al-mu'minūn* 'the believers', *al-muttaqūn* 'those who are mindful of God', *al-kāfirūn* 'the disbelievers', and *aḏ-ḏālimūn* 'evildoers', and employs conditional sentences with grammatical particles like *man* 'whoever', *mā* 'whatever', *ʾayy* 'whichever', *ḥayṯumā* and *ʾaynamā* 'wherever', and also the indefinite noun.

3.4 *Frequent use of adjectives*

Frequent use of adjectives is an important means of Qur'ānic persuasion and argument, noticeable from the very beginning: "Praise be to God, Lord of the Worlds, the Lord of Mercy, the Giver of Mercy..." (Q. 1/2–3). Because He has such attributes, He is worthy of praise and worship. The required path is the 'straight' one, the one 'whose followers are blessed and not the object of anger', or 'those who are astray', so qualified it is worthy of asking God's guidance to it. The believers are described in many ways (see, for instance, Q. 23/1–10, 70/22–29).

3.5 *Emphasis*

Because the *Qur'ān* addresses some people who have doubted or denied its messages, it frequently employs emphasis, using particles such as *ʾinna* and *la-*, and the suffix *-anna* with the imperfect verb (→ energicus), is widely used.

3.6 *Contrast*

A central feature of Qur'ānic style is contrast: between this world and the next (each occurring exactly 115 times); between believers and disbelievers; between paradise and hell. Remarkable patterns of contrast have been observed: angels and devils, life and death, secrecy and openness, and so on, occurring exactly the same number of times. Contrast comes naturally in a book that declares, "Say, [Prophet], 'Now the truth has come from your Lord. Let those who wish to believe in it do so and let those who wish to reject it do so'" (Q. 18/29). One of the linguistic characteristics of the *Qur'ān* is also to contrast two classes of a given thing, and their respective destinies. Grammatically, this contrast is achieved by such devices as *man*...*wa-man*...'those who...and those who...', as, for example, in Q. 4/123–124 and 92/5–8. Another device is *ʾammā*...*wa-ʾammā*, as in Q. 3/106–107: "On the day when some faces brighten and others darken, as for those with darkened faces it will be said...and as for those with brightened faces..." (see also Q. 79/37, 40). Sometimes the contrasted elements follow each other without any conjunction, which shows the contrast even more powerfully: for example, Q. 89/25–27: "On that Day, no one will punish as He punishes, and no one will bind as He binds. You soul at peace, return to your Lord, well pleased and well pleasing, go in among my servants and into my Garden".

3.7 *Dialogue and direct speech*

The *Qurʾān* frequently uses direct speech to bind each person by what he or she utters, rather than holding them responsible on the basis of reported speech. Arabic grammar allows shifts between direct and reported speech within a sentence after such verbs as *qāla* 'he said'. The fact that this verb occurs in the *Qurʾān* more than three hundred times is some indication of how frequently direct speech and dialogue are used.

3.8 *Rhyme*

Rhyme at the end of verses is a consistent stylistic feature in the *Qurʾān*, which has an aesthetic effect. It also gives finality to statements and accords with the general feature of classification and generalization, frequently using the plural endings *-ūn* and *īn*. The ending of the verse can be an integral part of it (as in sura 1) or a related comment, e.g. Q. 4/34–35, "…God is most high and great.… He is all knowing, all aware", but it is not just for embellishment (Omar 1999).

4. CONCLUSION

It is clear that the *Qurʾān* uses Arabic grammar and style together to serve its own purposes. Grammar may follow the normal rule (a process known as *istiṣḥāb al-ʾaṣl*). Considerations of style, however, can give priority to *ʿudūl ʿan al-ʾaṣl* 'departure from the original norm' or, as the scholars of *balāġa* say, *xurūj ʿalā muqtaḍā ḏ̣-ḏ̣āhir* 'departure from what is normally expected', but only 'for considerations required by the situation in certain contexts' (Hāšimī 1986:239). We have seen how the *Qurʾān* employs a feature like *iltifāt* more frequently than Arabic literature in general and how this was termed 'the boldness of the language of the *Qurʾān*'. The *Qurʾān* instructs the Prophet, "Speak to them about themselves, using penetrating words (*qawlan balīġan*)" (Q. 4/63).

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

ʾAbū ʿUbayda, *Majāz* = ʾAbū ʿUbayda Maʿmar ibn al-Muthannā, *Majāz al-Qurʾān*. Cairo, 1954–1962.

Ibn ʿAbd as-Salām, *Majāz* = ʿIzz ad-Dīn Ibn ʿAbd as-Salām, *Majāz al-Qurʾān*. London: Al-Furqan Foundation, 1999.

Ibn al-ʾAṯīr, *Maṯal* = Ḍiyāʾ ad-Dīn ʾAbū l-Fatḥ Naṣrallāh ibn Muḥammad Ibn al-ʾAṯīr, *al-Maṯal as-sāʾir fī ʾadab al-kātib wa-š-šāʿir*. 2 vols. Cairo, 1971.

Ibn Hišām, *Muġnī* = Jamāl ad-Dīn ʾAbū Muḥammad ʿAbdallāh ibn Yūsuf Ibn Hišām al-ʾAnṣārī, *Muġnī l-labīb*. 2 vols. Ed. Māzin al-Mubārak and Muḥammad ʿAlī Ḥamdallāh. Damascus: Dār al-Fikr, 1969 (and other editions).

Jurjānī, *ʾAsrār* = Majd ad-Dīn ʾAbū Bakr ʿAbd al-Qāhir ibn ʿAbd ar-Raḥmān al-Jurjānī, *ʾAsrār al-balāġa*. Ed. Helmut Ritter. Beirut: Dār al-Masīra, 1983. (1st ed. Istanbul, 1954. German trans., Hellmut Ritter, *Die Geheimnisse der Wortkunst des ʿAbdalqāhir al-Curcānī*. Wiesbaden: F. Steiner, 1959.)

——, *Dalāʾil* = Majd ad-Dīn ʾAbū Bakr ʿAbd al-Qāhir ibn ʿAbd ar-Raḥmān al-Jurjānī, *Dalāʾil al-ʾiʿjāz*. Ed. Muḥammad ʿAbd al-Munʿim Xafājī. Cairo: Maktabat al-Qāhira, 1969.

Šāṭibī, *Muwāfaqāt* = ʾAbū ʾIsḥāq ʾIbrāhīm ibn Mūsā aš-Šāṭibī, *al-Muwāfaqāt fī ʾuṣul aš-šarīʿa*. 3 vols. Beirut, 1975.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān ibn Qanbar Sībawayhi, *al-Kitāb*. Ed. ʿAbd as-Salām Muḥammad Hārūn. 5 vols. Cairo: Dār al-Qalam etc., 1988.

Suyūṭī, *ʾItqān* = Jalāl ad-Dīn ʾAbū l-Faḍl ʿAbd ar-Raḥmān ibn ʾAbī Bakr as-Suyūṭī, *al-ʾItqān fī ʿulūm al-Qurʾān*. 4 vols. Cairo, 1967.

Zarkašī, *Burhān* = Badr ad-Dīn Muḥammad ibn ʿAbdallāh az-Zarkašī, *al-Burhān fī ʿulūm al-Qurʾān*. 2 vols. Cairo: ʿĪsā al-Ḥalabī, 1957.

SECONDARY SOURCES

Abdel Haleem, Muhammad A.S. 1992. "Grammatical shift for rhetorical purposes: *Iltifat* and related features in the *Qurʾān*". *Bulletin of the School of Oriental and African Studies* 55:3.407–432.

——. 1993. "Context and internal relationships: Keys to Qurʾānic exegesis". *Approaches to the* Qurʾān, ed. Abdulkader A. Shareef and Gerald R. Hawting, 71–98. London: Routledge.

——. 2004. *The Qurʾan: A new translation (with introduction)*. Oxford: Oxford University Press.

ʿAmāyra, ʾIsmāʿīl ʾAḥmad and ʿAbd al-Ḥamīd Muṣṭafā as-Sayyid. 1988. *Muʿjam aḍ-ḍamāʾir wa-l-ʾadawāt fī l-Qurʾān al-karīm*. Beirut.

Burton, John. 1988. "Linguistic errors in the *Qurʾān*". *Journal of Semitic Studies* 33.181–196.

Hāšimī, ʾAḥmad al-. 1986. *Jawāhir al-balāġa*. Beirut.

Ḥassān, Tammām. 1973. *al-Luġa al-ʿarabiyya maʿnāhā wa-mabnāhā*. Cairo.

——. 1993. *al-Bayān fī rawāʾiʿ al-Qurʾān*. Cairo: ʿĀlam al-Kutub. (2nd ed., 2000.)

Nawfal, ʿAbd ar-Razzāq. 1976. *al-ʾIʿjāz al-ʿadadī fī l-Qurʾān al-karīm*. Cairo.

Nöldeke, Theodor. 1910. *Neue Beiträge zur semitischen Sprachwissenschaft*. Strasbourg: K. Trübner.

Omar, Ahmad Mukhtar. 1999. "The *fāṣila* in the *Qur'ān*: Word, context and meaning". *Journal of Qur'ānic Studies* 1:1.220–238.
Rabin, Chaim H. 1951. *Ancient West-Arabian*. London: Taylor's Foreign Press.
ʿUbāda, Muḥammad ʾIbrāhīm. 1980. *ʿUṣūr al-iḥtijāj fī n-naḥw al-ʿarabī*. Cairo: Dār al-Maʿārif.
Versteegh, Kees. 1993. *Arabic grammar and Qur'ānic exegesis in early Islam*. Leiden: E.J. Brill.
Wright, William. *A grammar of the Arabic language*. Repr., Mineola, N.Y.: Dover Publications, 2005.

MUHAMMAD ABDEL HALEEM
(School of Oriental and African Studies)

R

Rafʿ → ʾIʿrāb

Raising → Vowel Raising

Rajaz

The language of *rajaz* poetry is often complicated. This is due in part to the brevity of *rajaz* verses (usually eight to twelve syllables) and in part to the exigencies of → rhyme, which force the poet to repeat the same set of syllables every few words; in the most extreme case, as much as 25 percent of each verse may consist of rhyme. Not surprisingly, many of the anomalies of *rajaz* are found specifically in rhyme.

Grammatical and lexical irregularities and oddities are also found outside of rhyme. This made *rajaz* verses a favorite mine of *šawāhid* ‘evidentiary verses’ for lexicographers and grammarians. The → poetic licenses (*ḍarūrāt*) of *rajaz* should not be taken to represent a living language outside the genre itself without close examination of the examples. It is clear, though, that the existing potentialities of language did guide even poetic license. The limits of language were stretched in those directions which found some legitimacy in dialectal forms, existing anomalies, rare variants, and the like.

Rajaz falls into three distinct types, which in rough historical order of appearance are:

i. Short, presumably extemporized poems or, more properly, groups of verses. These are amply attested in historical literature, for example. They are usually devoid of any poetic merit and are written in relatively simple language, although they may occasionally contain linguistic oddities.
ii. Artistically developed *rajaz*, which flourished in Umayyad times and shortly thereafter. The last major *rājiz*, Ru'ba, died in 762 C.E., and his poems may thus be labeled Umayyad *rajaz*. A fully developed, polythematic *ʾurjūza* is typical of this group. In later periods, similar *rajaz* were mainly written on the theme of the hunt (*ṭardiyya*). This is the artistic variant of *rajaz*, the language of which is the main theme of this article, with some attention given to the preceding type.
iii. Didactic *rajaz*, often in *muzdawij* form, used for versified treatises (e.g. Ibn Sīnā's *ʾUrjūzat aṭ-ṭibb*). These fall outside our discussion here, the versified treatises containing little of linguistic interest.

The artistic tradition of *rajaz* retained a complex language even after the Umayyad heyday of the genre. After 800 A.H., *rajaz* of this type were mainly used in hunting poems, which retained the linguistic features developed in Umayyad *rajaz*. The influence of the genre on language may be seen in the complexity and sometimes even obscurity of the hunting poems by ʾAbū Nuwās, which are in stark contrast to the clarity of diction in his poems written in other meters. Thus, whereas it would not be appropriate to speak of the language of, say, *ṭawīl* or *kāmil*, it is reasonable to speak of a special language of *rajaz*.

This special language, however, does not reflect any one dialect or form of Arabic outside the genre itself, and the language used by *rajaz* poets does not coincide with their tribal background. Moreover, the poems of each poet exhibit heterogeneous features. The language of the *rajaz* is not a specific dialect but rather an artificial language created for the genre; or, to put it another way, it is standard poetic Arabic with artificial features. The use of such genre-related 'dialects', whether true or artificial, is well known in other literatures, too, such as Greek or Sumerian (where the EME.SAL was used specifically for certain genres).

The necessary complexity of longer *ʾurjūza*s and their linguistic peculiarities became a stylistic feature. Umayyad and later *rajaz* poets were supposed to use complicated language as a sign of their métier, making up words at will, changing morpheme patterns, and freely borrowing from foreign languages, especially Persian (see, e.g., Ruʾba, *Dīwān* 41, rhyming in *-aq*). Some of the foreign words may also have been used in contemporary colloquial language.

On the phonetic level, *rajaz* poems provide us with ample examples of both lengthening and shortening of vowels, as well as their elision. Some forms may go back to spoken language; others more probably reflect poetic license. Often, *metri gratia* forms were made acceptable by contemporary development of spoken language. Thus, for example, *tuʿṭi* (for *tuʿṭī*, see Ibn Manẓūr, *Lisān*, s.v. *l-y-q*) exhibits a shortening of the final vowel typical of spoken Arabic. Here, the acceptance of the deviation from standard grammar is further eased by the fact that the resulting form is morphologically correct, although its syntactic position is wrong: the sentence does not call for a jussive.

In *wa-t-tawan* (in rhyme) < *wa-t-tawānī* (infinitive Form VI; al-ʾAġlab, *Dīwān* 36.4, variant), the final long vowel is first dropped, and, subsequently, the resulting doubly long syllable is shortened, i.e., *tawānī* > **tawān* > *tawan* (as in Qurʾānic *dīn* for *dīnī*, in Q. 109/6, or in Syriac). Likewise, the irregular *qaranfūl* (in rhyme; Ibn al-ʾAnbārī, *ʾInṣāf* 24) can be understood as being caused by the fact that vowel length is almost neutralized in closed syllables: *qaranful/qaranfūl* is not a minimal pair, and the difference in length is unimportant.

The majority of cases of irregular vowel length seem to occur in closed syllables. Thus, forms like *qaranfūl* are more common than, for example, **qāranful*, which would run counter to phonetic reality. However, there inevitably are irregular cases caused by sheer metrical necessity, such as *nīḍāl* (for *niḍāl*; Ibn al-ʾAnbārī, *ʾInṣāf* 29). Here, one easily agrees with Ullmann (1966:72) in not regarding this as an archaic form of the infinitive of Form III. Before judging these cases as archaisms, one should first consider other possible explanations.

The freedom of *rajaz* poets is also seen in the ample use of quadriliteral verbs and quadri- and quinqueliteral nouns that are secondarily derived from triliteral roots and whose use is very often not dictated by metrical necessity. A good example of a secondarily quadriliteral verb is *darbaxa*, which is clearly derived by → dissimilation from *dabbaxa* (Ullmann 1966:124). The resulting verb is never metrically conditioned, as throughout the paradigm, *dabbaxa* and *darbaxa* share the same syllabic structures. The cases in which the irregularity is conditioned by neither rhyme nor meter deserve more attention as forms potentially derived from living usage.

What makes such *rajaz* formations possible and at the same time serves as the point of departure for new formations is the ability of Arabic to derive new lexical items from existing roots and morpheme patterns. The highly developed system of the root-and-morpheme pattern makes it easy to coin new words which are understood by the audience. One may compare this to such formations as the Greek/Latin *apokolokúntōsis* 'pumpkinification', which is not a 'real' Greek/Latin (or, for that matter, English) word, but whose meaning is easily understood. The word has hardly ever been used without a conscious reference to Seneca's work. Likewise, the extravagances of *rajaz* were understood, perhaps even used in learned language, but always with reference to *rajaz*, whether this was explicitly stated or not.

The parallel of pumpkinification is especially close to the often comically used verbs derived through denominalization, such as *taṯaʿlaba* (< *ṯaʿlab-*) or *dahqana* (< *dihqān-*; see Ullmann 1966:174). Such comic effects are also sought on the phonetic level, as in the famous poem in which *banī s-siʿlātī* rhymes with *an-nātī* (for *an-nāsi*; ʾAbū Zayd, *Nawādir* 345). The common word *an-nās* would undoubtedly have been familiar to the audience, and we can hardly imagine *an-nāt* to be a real dialectal form – at least, it does not find any support in either Semitic parallels or in Arabic dialects. Whether the comic effect came from an imitation of children's language (where dentals tend to appear before sibilants) or something else remains an open question, but the effect itself is obvious: This was not composed as serious poetry, nor was the word derived from any spoken or written variant of Arabic. A clear case of comic distortion of an originally serious poem is the famous poem by ʾAbū n-Najm (*Dīwān* 74; discussed in Hämeen-Anttila 1993), in which the bulk of the verses are transmitted as ending in *-amah* rhyme, but a group of four verses is usually transmitted as ending in *-amat* rhyme.

This freedom is taken advantage of on both the phonetic and morphological levels. When Ruʾba uses an expression such as *layl- lāʾil-* (*Dīwān* 45.260) or *kahf- kāhif-* (*Dīwān* 39.30), he does not tap any existing lexical resources but instead creates neologisms, examining the possibilities given by the linguistic structure of Arabic. In a similar manner, the phrase *a nightful night* may not be correct English, but it does convey something to the reader. To look for an 'exact' meaning to such words would be misguided. These words have meaning only within their own context; they are not lexical items to be defined. In a word, they are not actually lexemes but rather pegs for contextualized meanings.

It is also precarious to speculate on the exact shades of meanings of the unusual morpheme types that are only, or mainly, found in *rajaz*. Thus, for example, Ullmann (1966:64–65) is probably right in taking *fuʿālil-* to be merely a *rajaz* variant for *faʿlal-* and various other patterns. These rare *rajaz* morpheme types may well be taken as ad hoc patterns, and it is open to doubt whether they had any fixed or specialized semantic fields. Another good example of such an ad hoc formation is the seemingly archaic *al-ʾaġlalī* (*...li-ḏī l-ġalāli l-ʾaġlalī*, for *al-ʾaġalli*, in ʿAjjāj, *Dīwān* 29.68), whose form is identical with that of triliteral roots, *ʾafʿal* (and, thus, the supposed, though not real, historical form), and was certainly understandable for the audience (as *childs* would be to a modern speaker of English), even though it probably never occurred in standard usage.

Such forms were favored not only for metrical reasons but also for their expressivity. Words tend to become weakened by continuous use, and poets constantly search for words and expressions that startle readers and listeners and force them to listen attentively to the words. Rhetorical figures were one way to do this, but the *rajaz* poets found in the *curiosa* another way. Thus, when an anonymous poet uses the word *balandam-* (Ibn Manẓūr, *Lisān*, s.v. *b-l-d-m* = *balīd-*), he is not using an obsolete *-am* suffix, which might have some emphasizing meaning, but more probably he is taking advantage of the possibilities of Arabic in a playful mood.

To this point, we have considered the effects of meter and rhyme on the language, but in the earliest poems, the reverse has to be taken into account. Metrical irregularities of very early poems may lie behind oddities, which may sometimes 'correct' the meter or the rhyme at the expense of the language. It has to be remembered that the regular Xalīlian system of meters was born only after *rajaz* poetry had become a well-established genre.

The same concerns irregular rhyming. Some cases of *ʾikfāʾ* (as in *hayyin* rhyming with *ṭuʿayyim*; ʾAbū Zayd, *Nawādir* 400) actually coincided with Qurʾānic rhyming. In this case, the presumably original rhyming has been kept, but in many such cases the collectors of these poems may have preferred to resort to lexicographical oddities in order to preserve the rhyme.

Additionally, spoken language has sometimes clearly affected the *rajaz* conventions. Thus, for example, cases such as *lammā raʾā ʾan lā daʿah wa-lā šibaʿ* (instead of *daʿata*; Manẓūr, *Dīwān* 16.3) follow more the rhythms of spoken language than the grammatical rules concerning pausal phenomena.

Umayyad *rajaz* is full of neologisms. The ad hoc nature of most of these words may be seen from the fact that they often lack parallels and possible etymons both in other Semitic languages and in later Arabic dialects. In some cases, however, they may coincide with attested forms in other Semitic languages, and this has often induced scholars to regard them as archaisms. A famous case is *yuʾakram-*, attested in a poem by ʾAbū Ḥayyān al-Faqʿasī (quoted, for instance, by Ibn al-ʾAnbārī, *ʾInṣāf* 11) and repeated *ad nauseam* by various grammarians. This might seem an interesting parallel to the respective Aramaic stem (as well as some dialectal forms) and, moreover, might be taken to represent the etymological origin of the Arabic Form IV imperfect (**hafʿala, ʾafʿala*, imperfect **yuhafʿilu, yuʾafʿilu*). Yet, this is hardly the case, and Ullmann (1966:126) rightly dismisses the verse, referring to metrical necessities. In the vast majority of similar cases, it would be difficult to explain why a particular case should be differently explained in a system otherwise full of obvious ad hoc formations, and how such an archaic form could have been preserved. The existence of *hafʿala* is, incidentally, based on almost equally slight evidence and should not be taken as an archaism but rather as a neologism.

Most of these neologisms and other oddities are occasional and often confined to a single example, while other expressions and forms are recurrent and form a distinctive feature of the genre. One such grammatical curiosity is the use of a final *-a* in apocopates: *lam yuqdar-a, lam yaʿlam-ā* (in rhyme, for *yaʿlam-a*), *lā tajzaʿ-a, lā tuhlik-a* (Ullmann 1966:188). It is difficult to assess whether this was taken as an emphatic ending – comparable to the → energeticus *-an(na)* – or not, and what relation it might have with the poetic *-i/ī*, which may be suffixed, perhaps originally for prosodic reasons, to verse-final verbs ending in a consonant (e.g. *lam yafʿal-i*).

The several attestations of *maʿ* (for *maʿa*, see Ullmann 1966:96) are to be taken as a lexicalized variant, which, moreover, coincides with both modern tendencies (CvCv > CvC) and Semitic parallels (Hebrew *ʿim*) and may thus have been actually used in early Arabic, although the possible attestations in prose have been blurred by orthography.

Other *rajaz* peculiarities include the construction *bal* + genitive (often *bal baldatin*; see Ullmann 1966:190), where *bal* is used in the sense of *wāw rubba*, and the construction *min/fī ġayri lā* (Ullmann 1966:195), where *lā* is pleonastically added to *ġayr*.

Syntactic oddities are recurrent and gave much cause for speculation for medieval scholars, who tried to interpret them within the framework of Sībawayhian grammar. The word order is to a large extent conditioned by meter and rhyme, as is also the case in other genres of poetry, although the longer the line is, the easier it is to keep to grammatical rules concerning syntax.

A good example of syntactic freedom in *rajaz* is the verse *lam yajid yawman ʿalā man yattakil* (instead of **man yattakilu ʿalayhi*; Sībawayhi, *Kitāb* I, 394), which breaks the elementary rule of relative pronouns not taking prepositions. An even more striking case is *taḥta llaḏī xtāra lahū llāhu š-šajar* (instead of **taḥta š-šajari llaḏī xtārahū llāh*; ʿAjjāj, *Dīwān* 11.12).

Although neologisms may usually be understood from their context, Umayyad *rajaz* eventually overstepped its limits. As the poems grew longer and the language became ever more complicated, *rajaz* became virtually incomprehensible to courtly audiences, curiosities cropping up by the dozens. For the poet, the *rajaz* formations actually gave more freedom – with a little imagination one could make up scores of new words, fitting the rhyme and the meter – but this overtaxed the wit of his audience. The poems of Ruʾba verge on unintelligibility, and after a period of exploring its possibilities, the artistic *rajaz* was confined to a limited field, the *ṭardiyya*, the conventional expressions of which were easier to master without too much effort from the audience. Didactic and extemporized *rajaz*, on the other hand, avoided extreme curiosities and were satisfied with occasional oddities, often metrically conditioned.

Rajaz verses were culled by lexicographers in their quest for rare words and aberrant variants. In this way '*rajaz* words' found their way into lexicographical treatises and, from there, into standard dictionaries, such as *Tahḏīb al-luġa, Lisān al-ʿArab*, and *Tāj al-ʿarūs* (and into Western dictionaries), which contain much *rajaz* material. When the concomitant verse is given as a *šāhid*, one readily recognizes such *rajaz* words, but this is not always the case, and many words found their way into these dictionaries with no indication that they were *rajaz* words.

The lexicographical and grammatical importance of *rajaz* verses as *šawāhid* led to their fabrication, either by the philologists themselves or by their Bedouin informants, who were willing to provide the commodity that was expected of them. In some cases, such oddities may have been made bona fide; poems and sayings tend, in general, to develop aberrant variants (cf. the English *Parthian shot* > *parting shot*), and if these are highly valued, they have an even better chance of survival until they are eventually recorded. Philologists hunted for *nawādir* and wrote down whatever they found.

Philologists also seem to have often been actively searching *šawāhid* for phenomena they knew to exist, and at least sometimes they may have made up the missing *šawāhid* themselves. We have no reason to doubt the existence of

→ *kaškaša* (Ullmann 1966:226), but the verses given by philologists are suspect. The best examples are so full of these changes that it is highly unlikely that philologists just happened to come by these extremely liberal examples.

Sometimes, the hunt for *curiosa* is seen in the fact that the most bizarre variants are found only in philological literature, whereas the *Dīwān* contains (or contained) a standard form. Philologists preferred to circulate an aberrant variant (as in *al-ʾijjal* for ʾAbū n-Najm's *al-ʾiyyal*; see ʾAbū n-Najm, *Dīwān* 56.84) which, it may be suspected, they either made up to provide an example of *-iyy-* > *-ijj-* or, if we give more credence to their truthfulness, got it from the mouth of an informant who may have been – or imitated – a speaker of a dialect in which this phenomenon really existed.

In neither case should we speak of simple forgery. Philologists knew about these linguistic phenomena and just wanted to provide *loci classici* for them, as a kind of mnemonic device. What is important, though, is to realize that most of such variants quite obviously do not go back to the poets themselves. Indeed, the poems do not use any dialect consistently but exhibit a very heterogeneous collection of features, partly relying on actual spoken language but for the most part being playful inventions.

Rajaz has, however, been used, for instance by Kofler (1940–1942) and by Rabin (1951), to provide materials for ancient dialects. This is a precarious project, and the *rajaz* oddities cannot be taken as clear evidence for any attested form of language. There is, of course, reason to believe that what to the *rajaz* poets sounded realistic might actually have been attested in some way or another. There is, for example, a certain variation between /j/ and /y/ in both Classical Arabic and modern dialects, so that the ʾAbū n-Najm variant *al-ʾijjal* must have sounded credible, and it is not by accident that we have such a variant but not, say, **al-ʾimmal*, which would have been totally opaque to contemporaries. But one would be too credulous to treat this as a real attestation. In addition, it is totally useless to speculate, for example, on whether this was the dialect of Banū ʿIjl because the poet belonged to that tribe. The variant most probably does not go back to the poet himself, and even if it did, there would be no compelling reason to assume that he was using his own dialect.

On the other hand, the *rujjāz* used lexicographical materials for their own benefit, and this vicious circle created more curiosa. Cases such as Ruʾba's *šummaxzī* and *ḍummaxzī* (*Dīwān* 23.12–13) in a longer poem rhyming in /z/ raise the question to what extent the *rujjāz* used dictionaries or lexicographical texts for rare words. The variation /r/ ~ /z/ seems explicable only through misplaced diacritics. If this example occurred in a two-liner, it might be a joke, but as this is a full-length *ʾurjūza*, such a jocular misuse of words is probably out of the question. *Rajaz* poetry is thus both a veritable mine and a minefield, as the poems contain many ghost words of various types – ad hoc formations, comically used words, and obvious mistakes. In *rajaz*, there is a certain extemporizing play with words – in modern English literature one could compare this to Joyce's *Finnegan's Wake*. Words take on new, bizarre, yet associative forms and meanings, phonemes change their places or are replaced by cognate ones, and words take on new shades of meaning.

Rajaz neologisms hardly deserve a place in dictionaries, nor do they give a solid basis for historical linguistics, often being artificial formations. The basic rule for *rajaz* is 'anything goes'. Thus, in *qad marra yawmāni wa-hāḏā ṯ-ṯālī* 'two days have passed and this is the third' (Zamaxšarī, *Mufaṣṣal* 366), the last word (for *aṯ-ṯāliṯ* – although reading *at-tālī* is not, in this case, completely ruled out) may easily be understood in the light of the context. There is no underlying phonetic rule /ṯ/ > Ø before a pause, but enough information is given by the preceding dual and the first part of the word to allow the poet to arbitrarily shorten the last word. Similarly, *wa-š-šayxu ʿUṯmānu bnu ʿAf* (Ullmann 1966:100) contains enough information for the reader to be able to add the missing syllable (*ʿAffān*).

Thus, although *rajaz* poems do contain valuable material for historical linguistics, they have to be carefully analyzed before they can be used as linguistic material.

Bibliographical references

Primary sources

ʾAbū n-Najm, *Dīwān* = ʾAbū n-Najm al-Faḍl ibn Qudāma al-ʿIjlī, *Dīwān*. Ed. Jaakko Hämeen-Anttila, *Materials for the study of* rajaz *poetry*. I.

Dīwān *of Abū ʾn-Nağm*. (= *Studia Orientalia* 72.) Helsinki, 1993.

ʾAbū Zayd, *Nawādir* = ʾAbū Zayd Saʿīd ibn ʾAws al-ʾAnṣārī, *Kitāb an-nawādir fī l-luġa*. Ed. Muḥammad ʿAbd al-Qādir ʾAḥmad. Beirut: Dār aš-Šurūq, 1401 A.H.

ʾAġlab, *Dīwān* = al-ʾAġlab al-ʿIjlī, *Dīwān*. Ed. Jaakko Hämeen-Anttila, *Materials for the study of* rajaz *poetry*. II. *Five* rajaz *collections*. (= *Studia Orientalia* 76.) Helsinki, 1995.

ʿAjjāj, *Dīwān* = ʾAbū š-Šaʿṯāʾ ʿAbdallāh ibn Ruʾba al-ʿAjjāj, *Dīwān*. Ed. Wilhelm Ahlwardt, *Sammlungen alter arabischer Dichter*, II. Berlin: Reuther und Reichard, 1903.

Ibn al-ʾAnbārī, *ʾInṣāf* = ʾAbū l-Barakāt ʿAbd ar-Raḥmān ibn Muḥammad Ibn al-ʾAnbārī, *Kitāb al-ʾinṣāf fī masāʾil al-xilāf bayna n-naḥwiyyīna l-baṣriyyīna wa-l-kūfiyyīn*. Ed. Muḥammad Muḥyī d-Dīn ʿAbd al-Ḥamīd. 2 vols. N.p., n.d.

Ibn Manẓūr, *Lisān* = ʾAbū l-Faḍl Jamāl ad-Dīn Mukarram ibn Mukarram Ibn Manẓūr, *Lisān al-ʿArab*. Ed. ʿAlī Šīrī. 18 vols. Beirut: Dār al-ʾIḥyāʾ at-Turāṯ al-ʿArabī, 1408/1988.

Manẓūr, *Dīwān* = Manẓūr ibn Marṯad al-ʾAsadī, *Dīwān*. Ed. Jaakko Hämeen-Anttila, *Materials for the study of* rajaz *poetry*. III. *Minor* rajaz *collections*. (= *Studia Orientalia* 78.) Helsinki, 1996.

Ruʾba, *Dīwān* = Ruʾba ibn al-ʿAjjāj, *Dīwān*. Ed. Wilhelm Ahlwardt, *Sammlungen alter arabischer Dichter*, III. Berlin: Reuther und Reichard, 1903.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. Ed. Hartwig Derenbourg. 2 vols. Paris: Imprimerie Nationale, 1891–1899.

Zamaxšarī, *Mufaṣṣal* = ʾAbū l-Qāsim Maḥmūd ibn ʿUmar az-Zamaxšarī, *al-Mufaṣṣal fī ʿilm al-ʿarabiyya*. Beirut, n.d.

SECONDARY SOURCES

Hämeen-Anttila, Jaakko. 1993. "The lost *rajaz* of Abū ʾn-Nağm on the defeat of Yazīd ibn al-Muhallab by Maslama ibn ʿAbdalmalik". *Acta Orientalia* 54.46–52.

Kofler, Hans. 1940–1942. "Reste altarabischer Dialekte". *Wiener Zeitschrift für die Kunde des Morgenlandes* 47.61–130, 233–262; 48.52–88, 247–274; 49.15–30, 234–256.

Rabin, Chaim. 1951. *Ancient West Arabian*. London: Taylor's Foreign Press.

Ullmann, Manfred. 1966. *Untersuchungen zur Rağazpoesie: Ein Beitrag zur arabischen Sprach- und Literaturwissenschaft*. Wiesbaden: O. Harrassowitz.

JAAKKO HÄMEEN-ANTTILA (University of Helsinki)

Reanalysis

Reanalysis constitutes a major mechanism for synchronic and diachronic language change and is regarded as an indispensable catalyst for → grammaticalization (Hopper and Traugott 1993:32). The most elaborate study of this mechanism is found in Langacker (1977), who defines it as a "change in the structure of an expression or class of expressions that does not involve any immediate or intrinsic modification of its surface manifestation". It also is defined by Timberlake (1977) as "the formulation of a novel set of underlying relationships and rules". However, the definition and approach to the subject that seem to have been adopted by the great majority of works in this field are those of Langacker (1977). Predating Langacker, Bopp (1816) examined the evolution of the infinitive as a type of language change that may be termed reanalysis (see Harris and Campbell 1995:31 for a historical account of reanalysis).

Reanalysis is inextricably linked to (child's) language acquisition (Harris and Campbell 1995; Hopper and Traugott 1993), because children interpret syntactic relations of words and their constructions according to rules newer than those in the output of adults. Such rules are presumed to differ from the original ones used in their constructions. This mode of alternative construing of syntactic relations has been labeled 'abduction' and is said to lead to reanalysis.

1. TYPES OF REANALYSIS

The surface (phonetic) manifestation of the 3rd person masculine and feminine plurals of a subclass of Arabic verbs traditionally called 'defective' or → 'weak' verbs, like *yashūna* 'they [masc., fem.] neglect, omit', *yaġzūna* 'they [masc., fem.] strive, invade', is identical as a result of reanalysis and subsequent fusion along morpheme boundaries. The underlying masculine inflection is the morpheme *-ūna*, and the feminine is *-na*. Through phonological adjustment of the verb stem, *ya-ġzuw* drops the glide *-w-*, and when the personal suffixes, the masculine *-ūna* and the feminine *-na*, are attached to it, further vowel → truncation follows in accordance with the general tendency in the Arabic phonological rules to prohibit the occurrence of more than two vowels in a row. Therefore, in the case of the masculine and feminine, lengthening of the *-u-* takes place. Thus, the *-ū-* in the masculine is partially contributed by the stem and partially by the personal suffix, whereas in the feminine, it is fully contributed by the verb stem. The fusion of the stem with its affix in the masculine form is an instance of what Langacker calls 'boundary loss', a subtype of 'resegmentation', which triggers reanalysis: because of this fusion, masculine and feminine forms do not contrast in the plural conjugation, since the phonetic realization for both genders is the same.

When two independent elements are reanalyzed as a single unit, they may undergo boundary loss between them, which in turn may trigger subsequent semantic and syntactic change and further reanalysis. The loss of boundaries between *lā* 'not', used as preverbal negation particle of imperfect indicative verbs without change in the verbal mood inflection, and *ʾan* 'that', used as a subordinating particle triggering subjunctive mood inflection in the following verb, in the assumed bimorphemic structure *lāʾan* 'not that' (Ibn Manẓūr, *Lisān* V, 4082), has given rise to the particle *lan* 'not' used to negate the future in verbs, including verbs in the main clause – significantly different from its componential parts. As a negation particle of the future, *lan* was allowed to occur in constructions with fronted objects, such as *zaydan lan ʾaḍriba* 'Zayd, I will not strike' (Ibn Manẓūr, *Lisān* I, 4082), prior to the situation where the two particles were reanalyzed as monomorphemic. A single instance of reanalysis occurring in a localized construction may trigger a multiple series of reanalyses with far-reaching semantic and syntactic consequences and a change in the categorical status of the reanalyzed morphemes. In the sentence *jāʾa bi-s-sayyārati* 'he came by car', the preposition *bi-*, used to express an instrument, and cliticized to the following noun, was reanalyzed, or 'rebracketed' (for this notion, see Hopper and Traugott 1993:41, 82–83) as belonging to the verb rather than the following noun. This resulted in the idiomaticization of *jāʾa bi-* in Arabic as a 'verb preposition idiom'. In this case, the original semantic sense of the preposition *bi-* in the above construction has shifted from denoting an instrument, as in English *came by car* to *came with* > *brought the car*.

After reanalysis has taken place, an alternative interpretation of the above sentence *he brought the car* emerged. This new interpretation has given rise to the use of nouns denoting rational beings in the object of the preposition slot with reference to the outcome of a process, as in *jāʾa bi-l-waladi* 'he brought the boy', which was not permitted under the original interpretation. In some of the modern Arabic spoken varieties, the preposition in the periphrastic *jāʾa bi-* underwent further phonological erosion *bi* > *b*, with concomitant loss of semantic content. When the verb *jāʾa* underwent localized modification, entailing the loss of the glottal stop /ʾ/ in the word-final syllable, which is typical of spoken vernaculars, the eroded preposition *b*, having lost its categorical status, became an unrecognizable preposition, which facilitated a merger between the two components. In this case, after the boundary loss between the two elements, the erstwhile preposition was downgraded to a mere segment that was later attached to or incorporated with the verb stem, yielding *jāb* 'he brought'.

Reanalysis without triggering by morphophonological boundary shift, loss, or creation (identified by Langacker 1977) and without concomitant semantic change is also attested. This is the type that results in what Hopper and Traugott (1993) label 'rule change'. This type of reanalysis seems to have been familiar to medieval Arabic grammarians, such as Ibn Hišām (*Qaṭr* 165–170). He observes that the syntactic behavior of the negation particle *mā* 'not', which typically does not alter the case marking of its predicate, was reanalyzed in certain restricted contexts. It came to resemble the negation verb *laysa* 'not be', which marks its predicate in the accusative, as in the Qurʾānic verse *mā hāḏā bašaran* 'this is not a human being' (Q. 12/31), where *bašaran* 'human being' receives an accusative marking (see Harris and Campbell 1995:16 for this type of reanalysis). This type of change is additionally termed 'syntactic transfer' (Heine and Reh 1984:97).

Rule change resulting from earlier reanalysis is attested in the use of the plural suffix morpheme *-ūna/-īna*, which typically marks plurality for masculine participial forms having the feature [+ HUMAN], in the formation of Arabic decade numerals (e.g. *ʿišrūna* 'twenty', *ṯalāṯūna* 'thirty', etc.). In the case of these numerals, not only the plural morpheme has acquired a new, more precise semantic sense (i.e. 'ten'), but it has also become nonelidable when followed by another noun, signaling a different syntactic relation than before. When plural forms having such a morpheme are followed by other nouns, the general rule of Arabic grammar dictates the dropping of the final *-n* (e.g. *muwaḏ̣ḏ̣afū l-jāmiʿati* 'the employees of the university'; see Esseesy 2000 for further discussion).

2. Motivating factors

A given linguistic form or structure may have more than one semantic sense or function. For example, the word *ḥattā* (evolved from the noun *ḥatt* 'rubbing, peeling'; Ibn Manẓūr, *Lisān* II, 767–768) has the meaning 'until' in the time domain, in addition to a purposive sense, as in *sa-ʾuṣallī ḥattā ʾadxula l-janna* 'I will pray, so I may enter paradise/I will pray until I enter paradise'. The same form *ḥattā* can function as a preposition, as in *ʾakaltu s-samaka ḥattā raʾsihā* 'I ate the fish to its head' (Ibn Hišām, *Muġnī* I, 111); as a particle indicating unrealized goals and triggering the use of the subjunctive, as in *sa-ʾuṣallī ḥattā ʾadxula l-janna*; as a particle with the imperfect indicative, as in *ḥattā yaqūlu r-rasūlu* 'even the Prophet says…' (Ibn Hišām, *Qaṭr* 91); and as an exceptive particle, as in *fa-qātilū l-latī tabġī ḥattā tafīʾa ʾilā ʾamri l-lāhi*, 'so fight the [female] who covets except if she returns to God's command' (Q. 49/9; Ibn Hišām, *Muġnī* I, 112). These brief and sketchy remarks on the semantic senses and syntactic functions of *ḥattā* show that this word, deriving from a nominal source, has been reanalyzed as a preposition and as a particle, respectively (polygrammaticalization; see below, Sec. 3).

The effects of these reanalyses widen the range of syntactic functions in which *ḥattā* enters. Both the disambiguation of syntactic relations and syntactic functions in cases where there is more than one reading for one single linguistic form, and the principle of 'linguistic simplicity' aim at achieving optimality; both have been suggested as the chief motivating factors leading to reanalysis (Langacker 1977). From another related perspective, reanalysis is construed as fulfilling the potentials that a given syntactic construction or the semantic interpretation of a given form may have, which then may become amenable to a new analysis in the language (Harris and Campbell 1995:72). Nevertheless, it has been demonstrated empirically that ambiguity (Timberlake 1977) and structural opacity (Lightfoot 1988), while setting off reanalysis in some contexts, are neither necessary nor sufficient prerequisites in some attested cases of reanalysis (Harris and Campbell 1995:72).

3. Reanalysis and grammaticalization

In many cases, reanalysis leads to grammaticalization, especially when the unidirectionality principle, which is inherent in grammaticalization, obtains. In other cases, however, reanalysis need not lead to grammaticalization, and may even result in a reverse situation, as shown in Hopper and Traugott's examples of the English prepositions *up* and *down* (1993:49) becoming lexicalized as verbs, as in *up the ante*, or as nouns, as in *what a downer*. Such attested evidence led researchers (Heine a.o. 1991 and Hopper and Traugott 1993, among others) to reject the identification of reanalysis with grammaticalization and to dissociate the two phenomena of language change from each other, maintaining their conceptual separation. The two processes seem to affect one another in that a grammaticalized morpheme may lead to reanalysis of the construction in which it appears, and reanalysis may lead to the grammaticalization of a reanalyzed construction (see Heine a.o. 1991:215–220 for further discussion).

A conspicuous case where reanalysis has led to grammaticalization is the one concerning the evolution of the negation enclitic *-š* from the autonomous nominal source *šayʾun*, as in *mā bi-wuddī šayʾun* 'I do not desire/want a thing', where a series of reanalyses has taken place: the enclitic preposition *bi-* was reanalyzed as part of the verb stem, leading to a morphophonological adjustment that resulted in the form *biddī*, which assumes quasi-verbal function (→ pseudoverb). Most importantly, the reanalysis of *šayʾun* adds emphasis to the negation, hence, it becomes part of the negation domain. The effect of such a reanalysis is far reaching and includes morphophonological, syntactic, and semantic changes of the reanalyzed form. In the case just discussed, an originally optional element has become obligatory.

Reanalysis may also operate on grammaticalized words and result in moving already grammaticalized forms (i.e. grams) further along their grammaticalization track to reach maturity. The reanalysis of some Arabic interrogatives such as *man* 'who', *matā* 'when', *mā* 'what' as conditional particles and relative pronouns is a case in point. Disregarding the minor differences in mood inflections in Ibn Hišām's (*Muġnī* II, 18–19) examples *man yukrimnī*

ʾukrimhu ‘whoever honors me, I honor him’; *man yukrimunī ʾukrimuhu* ‘I honor whoever honors me’; *man yukrimunī ʾukrimuhu* ‘who honors me [so that] I honor him?’, a single morpheme *man* is reanalyzed in the same construction as an interrogative, as a relativizer, and as a conditional particle (polygrammaticalization). The consequence of such a reanalysis is a change in categorical status: the interrogative moves from the main clause to a subordinate clause, as in *kānat ʾiḏā man ittahama bihi mina l-jawārī* ‘she was one of the slave-girls whom he suspected’ (Fischer 2002:217); and in conditionals denoting timeless truth, as in *man jāla nāla* ‘he who roams will reach something’, it loses the sense of specific reference (Fischer 2002:227).

Renalysis of a single linguistic form may generate a series of reanalyses with far-reaching consequences, which create a new alternative conceptual interpretation of clausal relations and modify existing syntactic relations. The particle *mā*, and the various structures in which it appears, for instance, reveal the operation of reanalysis as a major mechanism in language change. This particle has a long and varied career in Arabic, Classical, Modern Standard, and spoken vernaculars alike. One of its assumed early usages is as an interrogative, as in *mā l-mawḍūʿātu l-latī tuhimmuk?* ‘what subjects interest you?’. Starting from its interrogative function, *mā* was used in constructions where the question was followed by an answer as an internal dialogue (within oneself), which was later construed as a relative clause. In such clauses. the analysis of *mā* resulted in its emergence as a relative pronoun, as in Bravmann’s (1977:254) example, *mā tajannabtu tannūran ḥāmiyan* ‘what I avoid is a hot furnace’, which he traces to the original *mā tajannabtu tannūran ḥāmiyan* ‘[as to] what I avoid, [I avoid] a hot furnace’. In these relative clause constructions, the interrogative structure is reanalyzed as sentential subject or topic, and the answer is reanalyzed as its predicate or comment.

The frequency of *mā* in combination with imperfect verbs, particularly with those inflected for the 3rd person masculine singular, may also have facilitated its analysis as a constituent with verbs, as in *mā yajīʾu* ‘that which comes/arrives (see Bravmann 1977 for extensive discussion). In this case, morphophonological adjustments followed the merger between the particle and the finite verb, giving rise to *majīʾun* ‘coming’, i.e. a verbal noun beginning with prefix *m-*, traditionally labeled *maṣdar mīmī*; participial forms like *muntaḏar* ‘expected’ from the base *mā yuntaḏaru* ‘that which is expected’; *nomina loci* (nouns of place), such as *majlis* ‘council’ from the base *mā yajlisu* ‘that which sits’. Reanalysis of syntactic structures containing *mā* have resulted, therefore, in the emergence of allomorphemes that are involved in several word formation processes in Arabic. Dissolution of the original syntactic relations and their replacement with morphological relations give further credence to Givón’s famous dictum “Today’s morphology is yesterday’s syntax” (1971:413).

In some instances, grammaticalization and reanalysis work in tandem, with the end result creating a new gram. In Classical Arabic, *rāḥa l-qawm* ‘people walked [at any given time]’, the verb was generalized from the original sense of *rāḥa* ‘to walk in the evening hours’ (Ibn Manẓūr, *Lisān* III, 1769). To be grammaticalized, it had to be reanalyzed as an auxiliary co-occurring with a verb in the imperfect. The process of grammaticalization entailed desemanticization of the verb of motion and its morphological splitting along the syllabic boundary, where the final syllable was preserved (in the inflection of the 3rd person masculine singular in the form of a future affix following the decomposition of the original lexical verb), creating *ḥa-yiʿmil* ‘he will do’ in some spoken Arabic dialects (Al-Najjar 1991:674).

Arabic script, stemming from a Semitic stock, often does not indicate the short vowels in word-final position, which may trigger conceptual reinterpretation. For example, in one reading of the verse *min šarri mā xalaq* ‘from evil that He created’ (Q. 113/2), the particle *mā* assumes the function of a relative pronoun; this is contrasted with another possible reading, where the same sequence is read *min šarrin mā xalaq* ‘from evil He did not create’, the *mā* particle being interpreted instead as the negative particle. Therefore, the orthography itself may play a role in creating syntactic and semantic ambiguity in certain structures, which causes reanalysis.

4. Reanalysis and actualization

Defined as “the gradual mapping out of the consequences of the reanalysis” (Timberlake 1977), actualization is distinct from, but dependent on, reanalysis. While reanalysis operates on the abstract level, relating to rules of grammar, actualization operates on the concrete level, the one where actual change materializes. The presence of actualization presupposes and implies the occurrence of reanalysis. An instance of actualization is found in Modern Standard Arabic, particularly in the language of the media, with the negation marker *lā*, which is typically used for negating indefinite nouns in the accusative, e.g. *lā salāma* 'there is no peace'. This *lā* of absolute negation is reanalyzed together with its negated noun as a single unit. This is motivated in part by loan translation (calque), mainly from English expressions having negation prefixes like *un-*, *in-*, *de-*, *non-*, and so on, as in *lā šuʿūr* 'unconsciousness'. The actualization stage, as delineated by Harris and Campbell (1995:81–82) and as applied to Arabic, shows *lā salāma* to be interpretable as 'there is no peace', or alternatively 'nonpeace'.

Consequences of this reanalysis include the inability of the definite article to attach to the noun stem directly since the negative *lā* intervenes between them; the loss of the accusative case on the negated noun; and the loss of the original interpretation, which signifies the categorical negation of what is denoted by the noun. Due to reanalysis of *lā* + *noun* as a single unit, this has led to a sequence DEFINITE ARTICLE-*lā*-NOUN STEM which is unattested in Classical Arabic. A continuation of actualization is found in the extension of such strings to adjectives in constructions where native words are used to express foreign concepts, e.g. *lā jadīda taḥta š-šams* 'there is nothing new under the sun' and *al-lā-markaziyya* 'decentralization' (→ compound). It is possible that such a process could be conventionalized and spread in the future, as the transfer of Western concepts increases exponentially as a result of intimate contact between the Arabic and the Western (mainly English) linguistic systems. If transfer becomes frequent and pervasive, this may lead to the limitation or elimination of the native linguistic process involving the genitive construction with *ʿadam* 'non-, lack', *ġayr* 'non-, un-', and the like, which would be expected in the formation of concepts of this type. If this prediction turns out to be correct, reanalysis and its subsequent actualization will lead to a gradual language change in the form of an innovation that is brought about in large measure by transfer of a foreign concept.

5. Reanalysis and analogy

Analogy has generally been characterized as a mechanism for language change that aims at homogenizing and eliminating markedness in a given class of linguistic structures or a paradigm. Analogy is contrasted with reanalysis in that the former interacts with the latter and makes language change explicit (see Hopper and Traugott 1993:61). When the plural *-ūna* morpheme in the formation of higher decades in Arabic numerals, such as *ṯalāṯūna* 'thirty' was reanalyzed into an inelidable morpheme with the interpretable meaning *ten*, it replaced the original dual suffix *-ā* of the numeral 'twenty' (Brockelmann 1982:490), likely through analogical extension. Reanalysis and analogy interacted in this case: after the suffix *-ūna* was reanalyzed as a suppletive form of the decimal base, the change of the dual suffix on the numeral 'twenty' to match that of the other higher decades brought regularity to the formation of the entire set of numerals. The consequence of this process of analogical extension is a compromise to the semantic interpretation of *ʿišrūna* 'twenty', since morphological change to that form made its semantic interpretation incompatible with its form; hence, loss in semantic transparency ensued, when compared with other numerals of the same set. Nevertheless, analogical extension in this case brought about a harmonizing effect to the model of formation of the decade numerals and eliminated inconsistencies in their morphosyntactic properties.

Another area where analogy as a mechanism for language change aims at reducing irregularity is in the derivation of various verb roots in Arabic. In spoken vernaculars, Arabic roots containing *-y* in their makeup, particularly in the final position, were construed as the model for derivation for other root types. Consonantal roots having *-w* as their final consonant (traditionally labeled 'defective' or → 'weak') via analogy were modeled in their derivation after the pattern in final *-y* verbs. Therefore, the verb root *r-m-y* generates *ramēt* 'I threw away', just as the root *š-k-w* generates *šakēt* 'I complained' (Versteegh 1997:100). Such cases obviously are the result of analogy without reanalysis, where change in the underlying structures of the verb root has a concomitant change in its surface manifestation.

6. Boundaries of reanalysis

Occurrence of reanalysis is not restricted to small linguistic stretches at the word and morpheme levels; larger structures are further amenable to this mechanism. The reanalysis of → topic-comment structures, for example, not infrequently leads to the emergence of relative clauses with embedded subordinate structures (Bravmann 1977:232–260: Lewkowicz 1971:810–825), and interrogative clauses (especially those beginning with *mā* 'what; whatever'; *matā* 'when; whenever'; *man* 'who; whoever') often emerge as conditionals.

Reanalysis is to be distinguished from other → mechanisms of linguistic change. A word-order change, such as the change from the preferred Verb-First position in both Classical and Modern Standard Arabic to the frequent Subject-First in spoken Arabic vernaculars, is often considered to be an instance of reanalysis. Such a rearrangement typically involves an immediate change in surface manifestation, and, strictly speaking, it should be ruled out as an instance of reanalysis. Nonetheless, a word-order change that results in minimizing variability is sometimes considered to be an instance of reanalysis, in particular when it leads to grammaticalization (Hopper and Traugott 1993:50).

A single order of constituents may become amenable to reanalysis, as in Fassi Fehri's example *al-ʾawlādu jāʾū*, which on one reading means 'the children, they came', when analyzed as a topic-comment sentence. On another reading, 'the children came', the preverbal noun is considered the subject rather than the topic (Fassi Fehri 1993:27–28).

The demarcation line between reanalysis and other language change phenomena, such as grammaticalization, is sometimes hard to draw. In *jāʾ-at al-banāt* 'the girls came', the *-at* attached to the verb functions as a (gender) agreement marker when compared with its use in *jāʾ-at* 'she came' as a pronoun (Fassi Fehri 1993:121). It is not at all clear whether the pronoun used as an agreement marker is due to reanalysis, to further grammaticalization, or to reanalysis that has resulted in grammaticalization.

Bibliographical references

Primary sources

Ibn Hišām, *Muġnī* = Jamāl ad-Dīn ʾAbū Muḥammad ʿAbdallāh ibn Yūsuf Ibn Hišām, *Muġnī l-labīb ʿan kutub al-ʾaʿārīb*. Ed. Muḥammad al-ʾAmīr. Cairo: Dār al-Kitāb al-Miṣrī, n.d.

——, *Qaṭr* = Jamāl ad-Dīn ʾAbū Muḥammad ʿAbdallāh ibn Yūsuf Ibn Hišām, *Qaṭr an-nadā wa-ball aṣ-ṣadā*. Ed. Muḥammad Muḥyī d-Dīn ʿAbd al-Ḥamīd. Beirut: al-Maktaba al-ʿAṣriyya, 2000.

Ibn Manẓūr, *Lisān* = ʾAbū l-Faḍl Jamāl ad-Dīn Mukarram ibn Mukarram Ibn Manẓūr, *Lisān al-ʿArab*. Ed. ʿAbdallāh ʿAlī al-Kabīr a.o. 3rd ed. Cairo: Dār al-Maʿārif, n.d.

Secondary sources

Al-Najjar, Balkees. 1991. "Grammaticalization of lexical markers in Kuwaiti Arabic". *Folia Linguistica* 25:3–4.665–675.

Bopp, Franz. 1816. *Über das Konjugationssystem der Sanskritsprache in Vergleichung mit jenem der griechischen, lateinischen, persischen und germanischen Sprache*. Frankfurt a. Main: In der Andraeischen Buchhandlung.

Bravmann, Max M. 1977. *Studies in Semitic philology*. Leiden: E.J. Brill.

Brockelmann, Carl. 1982. *Grundriß der vergleichenden Grammatik der semitischen Sprachen*, II. Repr., Hildesheim: G. Olms.

Esseesy, Mohssen. 2000. *Morphological and syntactic features of Arabic numerals as evidence of their diachronic evolution*. Ph.D. diss., Georgetown University.

Fassi Fehri, Abdelkader. 1993. *Issues in the structure of Arabic clauses and words*. Dordrecht and Boston: Kluwer Academic Publishers.

Fischer, Wolfdietrich. 2002. *A grammar of Classical Arabic*. Trans. Jonathan Rodgers. 3rd ed. New Haven, Conn., and London: Yale University Press.

Givón, Talmy. 1971. "Historical syntax and synchronic morphology: An archaeologist's field trip". *Chicago Linguistic Society* 7.394–415.

Harris, C. Alice and Lyle Campbell. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.

Heine, Bernd, Ulrike Claudi, and Friderike Hünnemeyer. 1991. *Grammaticalization: A conceptual framework*. Chicago and London: University of Chicago Press.

Heine, Bernd and Mechthild Reh. 1984. *Grammaticalization and reanalysis in African languages*. Hamburg: H. Buske.

Hopper, Paul and Elizabeth Closs Traugott. 1993. *Grammmaticalization*. Cambridge: Cambridge University Press. [2nd ed., 2003.]

Langacker, Ronald W. 1977. "Syntactic reanalysis". *Mechanisms of syntactic change*, ed. Charles Li, 57–139. Austin: University of Texas Press.

Lewkowicz, Nancy. 1971. "Topic-comment and relative clause in Arabic". *Language* 47.810–825.

Lightfoot, David W. 1988. "Syntactic change". *Linguistics: The Cambridge survey*. I. *Linguistic theory: Foundations*, ed. Frederick J. Newmeyer, 303–323. Cambridge: Cambridge University Press.
Timberlake, Alan. 1977. "Reanalysis and actualization in syntactic change". *Mechanisms of syntactic change*, ed. James Li, 141–177. Austin: University of Texas Press.
Versteegh, Kees. 1997. *The Arabic language*. New York: Columbia University Press.
Wright, William. 1982. *A grammar of the Arabic language*, II. 2nd ed. Beirut: Librairie du Liban.

Mohssen Esseesy
(The George Washington University)

Reciprocals

Reciprocals are expressions that indicate that the subject of the verb is at the same time the object. Unlike → reflexives, however, which have a similar function, reciprocals introduce a distributing element into the meaning. While a reflexive can be applied to a singular subject (x hits x), a reciprocal requires a plural subject, because it expresses the notion that the various members of the subject each perform the action described by the verb not on themselves, but on one or more of the *other* members of the subject: 'x and y hit each other' means 'x hits y and y hits x'.

Heim a.o. (1991) argue that in order to achieve this, a reciprocal expression such as *each other* is composed of two elements: a distributor (*each*) and a reciprocator (*other*). The distributor is associated with the subject (which in the theoretical framework that Heim a.o. use is expressed as a covert movement of the distributor to the subject) and thus turns a plural subject (*the men*) into a distributed subject (*the men each*). The reciprocator is in object position and, in a sense, directs the action back to the subject.

Arabic employs several different methods of expressing reciprocity. In Classical Arabic, one method is to use the word *baʿḍ* 'part, portion; some' in a correlative manner. This reciprocal expression fits well with Heim a.o.'s analysis: *baʿḍ* occurs twice, once associated with the subject and once associated with the object. (All examples from Modern Standard Arabic quoted here are from Kremers [1997], who collected them from a corpus of written Arabic.)

(1a) *wa-yuḏ̣āhiru* *baʿḍ-u-hum*
and-help.3ms some-Nom-them
baʿḍ-an
some-Acc.Indef
'and they help each other'

(1b) *la-ʾin* *inqasama* *n-nāširūna*
if be-divided.3ms the-publishers.Nom
baʿḍ-u-hum *ʿalā* *baʿḍ-in*
some-Nom-them on some-Gen.Indef
'if the publishers are divided among each other'

The *baʿḍ* associated with the subject can appear in subject position, as in (1a), in which case the verb agrees with it, taking 3rd person masculine singular form. The logical subject appears as a genitive (pro)noun on *baʿḍ*. If the subject is not pronominal, a different construction is often used, as in (1b). Here, the subject *baʿḍ* does not appear in subject position but rather stands in apposition to the subject *an-nāširūna* 'the publishers'. It takes a genitive suffix that refers back to the subject.

The *baʿḍ* associated with the object stands in object position and is declined: in (1a) it appears as a direct object with accusative case. (In ditransitive verbs it can also appear as the indirect object.) In (1b) it appears as the complement of the preposition *ʿalā* 'on', which is a prepositional object of the verb *inqasama* 'to be divided'. Here, *baʿḍ* takes genitive because all prepositions in Arabic assign genitive case. The object *baʿḍ* does not take any suffix and is often indefinite in Classical Arabic. In Modern Standard Arabic, it can also be definite, as in (2).

(2) *yulāqī* *baʿḍ-u-hum* *al-baʿḍ-a*
meet.3ms some-Nom-them the-some-Acc
'They meet each other'

Obviously, the logical subject of the verb can also be a 1st or 2nd person. In such cases, the verb may agree with *baʿḍ*, as in (3a), but it may also agree with the logical subject, as in (3b):

(3a) *wa-yajibu* *ʾan* *yastamiʿa*
and-must that listen.3ms
baʿḍ-u-nā *ʾilā* *baʿḍ-in*
some-Nom-us to some-Gen
'We must listen to each other'

(3b) *ʾan lā* *nusdiya* *l-ʾafʿāl-a*
that not 1p.confer the-benefits-Acc
baʿḍ-u-nā *li-* *baʿḍ-in*
some-Nom-us to- some-Gen
'…that we do not confer benefits on each other'

Example (3b) also shows an effect that is frequently found in reciprocals crosslinguistically: the distributor, here *baʿḍunā*, is not in subject position or even in apposition to it. Instead, it is positioned after the object. This position puts it close to the reciprocator, which is the complement of the preposition *li-* 'to'. This appears to be a common development in language: the distributor, although associated with the subject, has the tendency to remain close to the reciprocator, rather than appearing in or adjoined to subject position (Frajzyngier and Curl 2000). This development is also evident in English, where *each other* normally appears as one phrase (as opposed to constructions such as *they each hit the other* or *each of them hit the other*).

As seen in (3b), where the distributor follows the object, Modern Standard Arabic shows this tendency as well. It can also make the distributor and the reciprocator appear more or less as a fixed unit, which is clearly illustrated by prepositional objects, as in (4).

(4) *wa-hum* *yulawwiḥūna* *bi-l-ʾaydī*
and-they 3mp.wave with-the-hands
li-baʿḍ-i-him *al-baʿḍ*
to-some-Gen-them the-some
'and they wave their hands at each other'

In (4), the phrase *baʿḍihim al-baʿḍ* occurs more or less as a single element, as can be seen by the fact that the distributor *baʿḍuhum* now appears in the complement position of the preposition together with the reciprocator, rather than before the preposition, as in (3b). We may assume that *baʿḍihim* takes genitive case because of the preposition. It should be kept in mind, however, that this particular construction does not occur in Classical Arabic, which means that the case ending on the second *al-baʿḍ* cannot be established with certainty. Occasionally, the second *baʿḍ* is indefinite. When this happens, it is often (though not always) written with the indefinite accusative ending, which is one of the few case endings that *is* written, as in (5).

(5) *fa-qaʿadū* *yaštaġilūna*
and-3mp.sat 3mp.are.occupied
bi-baʿḍ-i-him *baʿḍ-an*
with-some-Gen-them some-Acc.Indef
'and they sat down, occupied with each other'

In (5), the second *baʿḍ* shows the indefinite accusative ending *-an*. However, because spoken Arabic has no case endings, we cannot assume that when a case ending appears in written form, it represents the intuition of the native speaker. According to Classical Arabic grammar, there would be no explanation for the occurrence of accusative case on the second *baʿḍ* in (5). Instead, it seems safe to assume that the ending here is a fossilized form.

In fact, it is not unlikely that the phrase *baʿḍuhum baʿḍan/al-baʿḍ* (with the appropriate suffix replacing *-hum* 'them') should be analyzed as a simplex (noncompound) form. This is exactly the final step of the development that draws together the distributor and the reciprocator, as can be seen in Dutch, for example, where the original *elk ander* 'each other' developed into Modern Dutch *elkaar*. That a similar development has taken place in Arabic is obvious from the fact that occasionally, only one occurrence of *baʿḍ* expresses the reciprocal in Modern Standard Arabic, as in (6).

(6a) *yuḥaddiṯūna* *baʿḍ-a-hum* *ʿan*
3mp.speak.to some-Acc-them about
ʾaxbār-i *l-qurā*
news-Gen the-villages
'They speak to each other about the news of the villages'

(6b) *tamassaknā* *bi-baʿḍ-i-nā*
1p.held with-some-Gen-us
wa-qtarabnā
and-1p.approached
'We held on to each other and went closer'

In both examples of (6), only one *baʿḍ* occurs, even though the meaning is clearly reciprocal. This *baʿḍ* occurs with a pronominal suffix and occurs as the logical object. In (6a), *baʿḍahum* is in direct object position, and in (6b) *baʿḍinā* is the complement of a preposition. This structure is most likely influenced by spoken Arabic. In spoken Arabic, the use of a single *baʿḍ* is quite common for expressing reciprocity, as in

(7a) from Egyptian Arabic, (7b) from Syrian Arabic, and (7c) from Moroccan Arabic.

(7a) *bi-yḥibbu* *baʿḍ*
Progr-3p.love some
‘They love each other’
(7b) *bə-tšūfu* *baʿəḍ-kon*
Progr-2p.see some-your
kəll *yōm*
every day
‘You see each other every day’
(7c) *ka-yətʿawnu* *bəʿḍiyat-hum*
Progr-3p.help some-them
‘They help each other’

In Egyptian Arabic, *baʿḍ* usually occurs without suffix, appearing, as expected, in object position. Syrian Arabic is similar, except that *baʿḍ* (here appearing with an epenthetic schwa) usually appears with a pronominal suffix referring back to the subject. Moroccan Arabic behaves like Syrian, the form of *baʿḍ* used here being *bəʿḍiya(t)*.

Classical Arabic represents the first stage in the development of *baʿḍ* as a reciprocal, while the dialect forms in (7) represent the end stage. In Classical Arabic, the reciprocal is expressed with a correlative use of *baʿḍ*: it occurs twice in the sentence, once in subject position, or as adjunct to the subject, and once in object position. The subject-*baʿḍ* takes a pronominal suffix that refers to the subject (or it takes the logical subject as possessive modifier). The object-*baʿḍ* is indefinite and does not take any suffix. In colloquial Arabic, the reciprocal is expressed with a single *baʿḍ* + suffix. Modern Standard Arabic shows a variety of methods, which are usually somewhere between the Classical and the colloquial expression.

Some dialects have chosen a different method. Iraqi Arabic, for example, uses a typical distributor-reciprocator structure, but not with the lexical element *baʿḍ*, as in (8).

(8) *yšūfūn* *wāḥid* *il-lāx* *kull* *yōm*
3p.see one the-other every day
‘They see each other every day’

In (8), the word *wāḥid* ‘one’ is used as the distributor, and *il-lāx* ‘the other’ as the reciprocator. The use of a word for ‘one’ as distributor is not unexpected, as evidenced by the English structure *they see one another every day*.

Apart from the method described here, Arabic has another way of expressing reciprocity, one which uses a derived verb stem. Form VI of Arabic verbs, which is formed with the pattern *ta*K*ā*T*a*B*a*, often has a reciprocal meaning, e.g., *ḍaraba* ‘to hit someone’, *taḍāraba* ‘to hit each other, to come to blows, to clash’; *ʿāwana* ‘to help someone’, *taʿāwana* ‘to help each other, to cooperate’; *kātaba* ‘to write to someone, to correspond with someone’, *takātaba* ‘to write to each other, to correspond’.

This process, however, is at best marginally productive. Many Form VI verbs have lexicalized meanings, as the examples show. Furthermore, there are Form VI verbs that do not have a reciprocal meaning: *xaḏala* ‘to leave, forsake’, *taxāḏala* ‘to wane, to decrease, to fade’; *bāraka* ‘to bless someone’, *tabāraka* ‘to be blessed’; *saqaṭa* ‘to fall’, *tasāqaṭa* ‘to fall down piece by piece, to collapse’. In fact, as Wright (1981: I, 38D) indicates, reciprocity is not the basic meaning of Form VI of the verb.

Bibliographical references

Frajzyngier, Zygmunt and Traci S. Curl. 2000. *Reciprocals: Forms and functions*. Amsterdam and Philadelphia: J. Benjamins.

Heim, Irene, Howard Lasnik, and Robert May. 1991. “Reciprocity and plurality”. *Linguistic Inquiry* 22:1.63-101.

Kremers, Joost. 1997. *How Arabs speak to each other about themselves: A study of* nafs *and* baʿḍ *in Modern Standard Arabic*. M.A. thesis, University of Nijmegen.

Wright, William. 1981. *A grammar of the Arabic language*. Beirut: Librarie du Liban. (Repr. of the 3rd ed., 1896.)

Joost Kremers (University of Frankfurt)

Redundancy

1. Definition

Redundancy is a universal property of human language. There is no language that does not have some degree of redundancy, and the Arabic language is no exception. In WordNet 2.1, Fellbaum (1998) gives four senses of redundancy:

i. Repetition of messages to reduce the probability of errors in transmission
ii. The attribute of being superfluous and unneeded with one direct hyponym

iii. Someone or something that is unwanted and unneeded
iv. Needless repetition of an act

The English/Arabic dictionary of theoretical linguistics (Al Khuli 1982) translates 'redundancy' with *ḥašw, ʾishāb*, and *ʾiṭnāb*.

Sentence (1) below represents an example of redundancy in Arabic.

(1) *hāḏā* *r-rajulu* *ʿāṭilun*
this the-man unemployed
ʿan *al-ʿamal*
from the-work
'This man is unemployed'

The phrase *ʿan al-ʿamal* 'from work' is redundant because the word *ʿāṭil* 'unemployed' conveys the same meaning. Sentence (2) gives the same meaning as (1), but to the Arabic ear, (1) sounds more idiomatic and more natural than the less ambiguous sentence in (2).

(2) *hāḏā* *r-rajulu* *ʿāṭilun*
this the-man unemployed
'This man is unemployed'

Another example of redundancy in Arabic is shown in (3). The word *dam* 'blood' in this sentence is redundant, because the verb *yanzifu* implies the loss of blood.

(3) *kāna* *l-marīḍu* *yanzifu* *daman*
was the-patient bleed.3ms blood
'The patient was bleeding blood'

The same meaning could be conveyed as in (4) without any loss of meaning.

(4) *kāna l-marīḍu yanzifu*
'The patient was bleeding'

2. Redundancy as a property of human language

The universality of redundancy suggests that it serves a purpose in human communication. Chomsky (1965) points out that linguistic performance is not always a true reflection of linguistic knowledge. He explains that the native speakers' production and perception of speech could be affected by lapses of memory, distraction, fatigue, etc. If human language did not exhibit a large degree of redundancy, important information would be lost when communication channels are 'noisy', i.e. when they are not optimal. Redundancy can be extremely useful in compensating for the loss of parts of the message in noisy channels. Secondly, redundancy allows for stylistic variations. There are situations that require very precise language, while in other situations a detailed, extensive, and very clear, or even repetitive style would be preferred. For example, Arabic speakers would prefer (1) to (2) above, in spite of the redundancy in (1).

3. Redundancy and predictability

Redundant words or phrases can also be predictable. Since their absence from the sentence or from discourse does not result in a loss of meaning, this implies that whatever meaning they represent is also present in the rest of the sentence or discourse. In languages like Italian, Hebrew, and Arabic, subject pronouns may be dropped from the sentence without any loss of meaning. These pronouns are regarded as redundant because the information they convey is also encoded in the verb. Consider sentence (5).

(5) *ʾana* *qābal-tu* *raʾīsa*
I met-1s president
l-jāmiʿati *ʾamsi*
the-university yesterday
'I met the president of the university yesterday'

The subject of the sentence, the pronoun *ʾana* 'I', can be easily deleted without any loss of meaning, as shown in (6).

(6) *qābaltu raʾīsa l-jāmiʿati ʾamsi*
'I met the president of the university yesterday'

The semantic features that make up the meaning of the Arabic word *ʾana* include {+subject, +1st person, +singular}. These same features are also encoded in the verb *qābal-tu* 'met-I'. Therefore, the subject pronoun can be predicted on the basis of the verb form. The relation between redundancy and predictability has also been pointed out by others, such as Pinker

(1994), who shows that it is because of redundancy that English speakers can understand the sentence in (8), even when each vowel in the sentence is replaced by the letter *x*, as shown in (7).

(7) *yxx cxn xndxerstxnd whxt x xm wrxtxng xvxn xf x rxplxcx xll thx vxwxls wxth xn 'x'*

(8) *You can understand what I am writing even if I replace all the vowels with an 'x'*

Pinker distinguishes between two types of predictability: syntactic and semantic predictability.

4. Syntactic predictability

Consider sentences (9) and (10) below. Native speakers of Arabic have no problem in recognizing that the missing word in (9) is the relative pronoun *allatī* 'who'. They recognize that this is a relative clause structure and that the lexical head of the relative clause is {+human, +female, +singular}, and that the relative clause that has these features is headed by *allatī* and not, for example, by *allaḏī*. Similarly, Arabic native speakers will have no difficulty in completing the sentence in (10) with the complementizer *ʾanna* 'that', because they also recognize that the verb *ʾaʿlana* 'announced' requires this particular complementizer. The knowledge that is used to give the correct prediction in (9) and (10) is the linguistic knowledge of the language, syntactic knowledge. This is not the case in (11) and (12) below.

(9) *hāḏihi* *hiya* *s-sayyida----*
this she the-lady----
qābal-tu-hā *ʾamsi*
met-1s-her yesterday
'This is the lady I met yesterday'

(10) *ʾaʿlana l-jayšu l-ʾamrīkī----quwwātu-hā ʿtaqalat ʿadadan min ʾaʿḍāʾi tanḏ̣īmi l-qāʿida*
'The American army announced----its forces arrested a number of Al-Qaida's members'

5. Semantic predictability

Consider the sentence in (11). A plausible completion of the sentence is the word *al-ḥayawān* 'the animal'. Native speakers would make such a prediction, because they know that the lion is an animal and that it is often referred to as the king of the jungle. This is not really linguistic knowledge specific to Arabic. The knowledge that the lion is an animal is common knowledge to everyone regardless of the language they speak. This is knowledge of the world. Likewise in (12), Arabic speakers would have no difficulty filling in the blank with the word *imraʾa*. They use their knowledge of the world, which tells them that Ahmed's wife cannot be other than a woman.

(11) *al-ʾasadu huwa----al-waḥīdu llaḏī yuʿtabaru malika l-ġāba*
'The lion is the only----that is regarded as the king of the jungle'

(12) *zawjatu ʾaḥmada----ḏakiyyatun jiddan*
'Ahmed's wife is a very intelligent----'

6. Telegraphese and the deletion of functional words

Evidence suggests that people recognize that human language is redundant. When people send telegrams, they economize by using minimal words to convey the message. They remove all words that are not necessary and keep only the words that unambiguously represent the meaning they want to express. They also combine both semantic and syntactic knowledge when they attempt to eliminate redundancy. Consider sentences (13) and (14) below.

(13) *sawfa ʾaṣilu ʾilā maṭāri l-qāhirati d-duwaliyyi fī yawmi l-ʾarbiʿāʾi l-qādimi l-muwāfiqi 23 ʾaġustus sanata 2006 fī tamāmi s-sāʿati r-rābiʿati baʿda ḏ̣-ḏ̣uhri ʿalā matni l-xuṭūṭi l-jawwiyyati l-bariṭāniyyati riḥlati raqma 572*
'I will be arriving at Cairo International Airport next Wednesday, August 23, 2006, exactly at 4:00 in the afternoon on British Airways, on flight number 572'

(14) *ʾaṣilu l-ʾarbiʿāʾ as-sāʿa 16.00 al-bariṭāniyya riḥla 572*
'Arriving Wednesday at 1600 British Airways flight 572'

Both sentences express the same information. However, (13) has 28 words, whereas (14) has only 7. In normal communication, it would sound absurd if people expressed their ideas in the kind of telegraphic language shown in (14). Only in very restricted contexts do people use telegraphese. A closer look at (13) and (14) reveals interesting insights into the implicit knowledge that speakers have not only of redundancy but also of the grammar and semantics of their language. For example, one notes the explicit use of the future tense marker *sawfa* (13). Time reference in many languages may be expressed through tense markers and time adverbials. In many cases it is expressed by both in the same sentence, as in (13), using the future tense marker *sawfa* and the time adverbial *yawm al-ʾarbiʿāʾi l-qādimi l-muwāfiqi 23 ʾaġustus sanata 2006*.

A strategy that speakers use when summarizing a text or sending a cable is to eliminate duplications in time reference. In (14), time reference is conveyed through the time adverbial *al-ʾarbiʿāʾ* 'Wednesday'. The absence of the future tense marker in (14) does not cause any loss of meaning expressed in (13), which shows that users of the language recognize such redundancy, and when the need arises, they know how to minimize it.

7. REDUNDANCY IN GRAMMAR

The grammar of many languages imposes some degree of redundancy on the users of the language. An often-quoted case is the requirement that speakers of English stick an *-s* at the end of the verb when the subject is 3rd person singular. Thus, (15) follows correct usage, whereas (16) does not and has to be asterisked.

(15) *Mary works as a waitress on weekends*

(16) **Mary work as a waitress on weekends*

The absence of the *-s* from the verb in (16) does not result in any loss of meaning, but its presence is required by the rules of the English grammar.

Consider how the number system works in Modern Standard Arabic. It is well known that Arabic requires that nouns following any of the numbers 3–10 be in the plural form, whereas a noun following numbers higher than 10 has to be in the singular form. This is a purely syntactic requirement, since we all know that semantically the nouns in both cases are plural. Consider (17).

(17) *xamsu mumarriḍātin*
'five nurses'

The plural in (17) is denoted by two markers: the number 'five', which is clearly plural, as well as the morpheme *-āt*, which indicates feminine plural. According to the rules of Arabic grammar, (18) is ungrammatical although Arabic speakers will have no problem understanding it with the same meaning as that of (17).

(18) **xamsu mumarriḍatin*
'five nurse'

Actually, the structure in (18) is perfectly acceptable in Modern Standard Arabic for numbers above 11, as shown in (19).

(19) *xamsūna* *mumarriḍatin*
fifty nurse
'fifty nurses'

It would be wrong, however, to use the structure in (17) for the equivalent sentence of (18), as in (20).

(20) **xamsūna* *mumarriḍātin*
fifty *nurses*

8. THE VALUE OF REDUNDANCY AS A PROPERTY OF LANGUAGE

Redundancy serves several purposes in human communication. First, it assists hearers in understanding, when communication channels are noisy. By 'noisy' we mean communication channels that are not perfect, such as bad and/or sloppy handwriting, spelling errors, loud interference from radio, television, traffic, distractions, etc. Redundancy makes up for any loss of any part of the message and for the temporary breakdown of communication channels. Second, there are specific discourses where repetition, paraphrasing, stating, and restating can be desirable. Consider a classroom situation where teachers would be commended for their ability to explain difficult concepts through detailed explanation. Good teachers often resort to explaining the same concept in different ways to make sure that students grasp it. They also tend to encourage students to look at the same phenomenon from different angles. Moreover, there are certain disciplines where learning involves acquiring a skill rather than comprehending a set of facts. A typical example is learning foreign languages, where students benefit from paraphrase and repetitions. Third, repetitions and elaborations are characteristic of religious discourse which is didactic by nature. A closer look at the *Qurʾān* and the Bible shows extensive repetitions and elaborations which appeal to the believers. Fourth, redundancy is a convenient way of characterizing differences among languages. While all human languages have redundancy as an intrinsic property, they vary in how much redundancy is the norm in each. Languages reflect the culture and the way of thinking of their speakers. Some speakers like flowery language and tolerate a high degree of redundancy, while others value precise language. The Arabic language has been perceived by non-Arabic speakers (Johnstone 1983) as allowing a very high degree of repetition and redundancy. However, for speakers of Arabic, this trait is perceived as one of the positive properties of the Arabic language. Fifth, redundancy is not an all-or-nothing feature. It can be present in various degrees. Thus, it allows speakers to express their individuality and their personal style. It is an integral part of the idiolect of the individual. Life would be very boring were redundancy to cease to exist. Imagine how life would be if we all spoke in such telegraphic language as presented in (12) above, which does not exhibit any degree of redundancy.

9. The communicative function of redundancy

It is often thought that redundant words or phrases do not have real communicative purpose. For example, in → pro-drop languages (Chomsky 1981), such as Italian, Spanish, and Arabic, the subject pronoun of a sentence may be dropped without any loss in the meaning of that sentence. Consider (21) and (22).

(21) *ʾana* *ḏahab-tu* *ʾilā*
I went-1s to
l-jāmiʿati *ʾamsi*
the-university yesterday
'I went to the university yesterday'

(22) *ḏahab-tu* *ʾilā* *l-jāmiʿati* *ʾamsi*
went-1s to the-university yesterday
'I went to the university yesterday'

In (21) the subject pronoun is present, whereas in (22) it is dropped without any loss of meaning. The reason there is no loss of meaning in (22) is that the subject pronoun *ʾana* 'I' is redundant in (21). The semantics of the independent pronoun in (21) is also encoded in the inflection of the verb. The subject pronoun could not be deleted if it was not redundant, as (23) and (24) show.

(23) *hiya fi l-maktab*
'She is in the office'

(24) *fī l-maktab*
'in the office'

(24) is a grammatical Arabic prepositional phrase but not really an independent sentence.

Both Eid (1980) and Farghaly (1982) show that, although the subject pronoun in Arabic seems redundant, it serves an important communicative function. Consider the examples (25) and (26), adapted from Farghaly (1982).

(25) *ḥaḍara ʿalī ṯumma nṣarafa*
'Ali arrived, then he left'

(26) *ḥaḍara ʿalī ṯumma nṣarafa huwa*
'Ali arrived, then he left'

In both (25) and (26), the subject of the second verb is a pro. In (25) it is dropped, whereas it is present in (26). The preferred reading for (25) is that the subject of the second verb is Ali. The preferred reading for (26) is that the subject of the second verb is someone other than Ali. So, although the subject pronoun in (26) is redundant and can be deleted, its presence or absence plays an important role in the interpretation of the sentence. Thus, redundancy plays here an important communicative role.

Bibliographical references

Al Khuli, Muhammad. 1982. *A dictionary of theoretical linguistics: English-Arabic with an Arabic-English glossary*. Beirut: Librairie du Liban.
Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Mass.: MIT Press.
——. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

——. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.
Eid, Mushira. 1980. "On the function of pronouns in Egyptian Arabic". Paper presented at the *Discourse Symposium*, University of Wisconsin, Milwaukee.
Farghaly, Ali. 1982. "Subject Pronoun Deletion rule in Egyptian Arabic". *Proceedings of the Second National Symposium on English Language Teaching in Egypt*, Cairo University, Egypt.
Fellbaum, Christiane (ed.). 1998. *WordNet 2.1: An electronic lexical database*. Cambridge, Mass.: MIT Press.
Johnstone, Barbara. 1983. "Presentation as proof: The language of Arabic rhetoric". *Anthropological Linguistics* 25.47–60.
Pinker, Steven. 1994. *The language instinct: How the mind creates language*. New York: W. Marrow.

Ali Farghaly (Oracle USA Corporation)

Reduplication

Reduplication is a special case of the more general device of repetition of linguistic material. It figures among the most prominent current issues in linguistic theory and mostly concentrates on morphological reduplication (cf. Rubino, 2005). Reduplication, however, can also occur on other linguistic levels, as already proposed by Pott (1862). But it is a matter of some debate where the line should be drawn between reduplication and other repetition phenomena. It is useful to distinguish between phonological reduplication and grammatically used reduplication, i.e. morphological and syntactic reduplication.

In formal terms, reduplication can be full (total) or partial. Full reduplication mostly iterates a whole word, e.g. *bḥal bḥal* and *kif kif* (Moroccan Arabic) 'alike'; in partial reduplication, a certain structure is only partly repeated, as in Classical Arabic *nāma nawman* 'he slept long/well' (→ object, absolute), where only the root consonants are repeated (→ paronomasia). Partial reduplication in Arabic is also a means of word formation.

The material reduplicated is mostly called 'base' and the copy of the base 'reduplicant'. The base of reduplication can be defined morphologically (e.g. root, stem, affix) and/or phonologically. The phonological material that is copied can frequently be defined in prosodic terms, possible bases typically having the size of a syllable or a foot. In linguistic theory, it is a matter of some debate if the repetition of single segments can also be counted as reduplication. Another formal issue is the position of reduplication, whether it is initial, final, or medial.

In some sense, reduplication is iconic, as more of the same form is also more of the same meaning. In functional terms, reduplication serves to signify different notions, among them plurality or intensity. Reduplicated nouns often denote plurals, but also seemingly contrary notions like diminution. With verbs, reduplication also often expresses plurality, be it iteration of the action or event, or the plurality or distribution of an argument, but also tense, aspect, attenuation, intensity, and transitivity.

Phonological reduplication is a characteristic feature of babbling and baby talk and most frequently exhibits CvCv-forms, like Syrian Arabic *wawa* 'hurt', *ninnī* 'sleep', or *zēze* 'breast; milk' (→ child language; Ferguson 1964).

One special kind of reduplication is that of echo words, such as Egyptian Arabic *šīkī-mīkī* 'fashionable' (derogative) or Moroccan Arabic *šaṭa-maṭa* 'boisterous' and *bih-fih* 'quickly, fast', where the first consonant of the second part is mostly a labial. In Egyptian Arabic, nicknames are formed by repeating the first consonant of the name plus *ī* or *ū*, e.g. *Fīfī* for *Faṭma* and *Zīzī* or *Zūzū* for *Zēnab*.

Reduplication in Arabic word formation has not figured prominently in the linguistic debate. The only words that quite obviously involve reduplication are quadriliterals of the type *waswasa* 'he whispered' and *zalzala* 'it trembled', but this type of Classical Arabic words is very frequently onomatopoetic and predominantly denotes sounds and movements or animal and plant names (→ sound symbolism; Procházka 1995). Despite their onomatopoetic origin, these words fit into the overall derivational system, and the root can be extracted and made the basis for the derivation of other forms. One common feature of the sound-symbolic words as well as the ones signifying movements is the iterativity of the event. This may have been the basis for the emergence of novel forms, especially in Neo-Arabic varieties, which can be traced back to an existing geminated root, e.g. *l-f-f* > *laflif* 'to wrap up', *l-m-m* > *lamlim* 'to gather up, gather together' (El Zarka, 2005), exhibiting verbal plurality. These verbs are somehow related to Form II verbs, as in the case of Egyptian Arabic *fattit* 'to crumble, break bread into pieces' and *fatfit* 'to crumble, break bread into tiny pieces'. This last pair of verbs is a clear example of phase-internal plurality, i.e., one event or action is divided into many parts; other examples are Egyptian Arabic *maṣmaṣ* 'to suck on, gnaw, nibble' < *maṣṣ* 'to suck'; *dašdiš* 'to reduce to fragments, smash, shatter' < *dašš* 'to mash, pound'. But this only seems to be possible with biconsonantal roots. With triconsonantal roots, only one consonant can be doubled, given the quadriliteral shape of the pattern that must not be exceeded.

Thus, a fair number of novel words in Neo-Arabic varieties are found with final and initial doubling. Final doubling is especially common among the Maghrebi varieties (Brockelmann 1908:518), but it has a predecessor in the obsolete verbal stems of Classical Arabic, whereas initial doubling apparently is a new feature of Neo-Arabic. According to Brockelmann (1908:517), final doubling in Semitic denotes a kind of iteration, in the sense that iteration is a prerequisite of becoming habitual. This is true in some sense for the Classical Arabic Forms IX and XI, where the doubled last radical forms a geminate (for gemination as reduplication, see below), and for Form XIV, all three of which denote a physical state pertaining to color (IX) or physical defect and its intensification (XI, but also XIV), e.g. *iswadda* (IX) and *iswādda* (XI) 'to be black', *isḥankaka* (XIV) 'to be dark'. With quadriliterals this form has long been used to derive denominative verbs like *jalbab* 'to put on or wear a *jilbāb* [a special kind of garment]'. But it also denotes plurality or distribution, like *šamlal* 'to gather ripe dates'. In Moroccan Arabic, there are many verbs that exhibit this form, but it is not always possible to trace them back to an existing simplex form or a root. While *bərgəg* 'to have an eye on' seems to be related to the Classical Arabic verb *baraqa* 'to shine [eye]', verbs like *gužəž* 'to gather together' and *hərnən* 'to grind one's teeth' are not obviously related to any root.

The same is true for initial doubling, which probably originates in the deconstruction of a full reduplicated form like **kabkab* > *karkib* (Egyptian Arabic) 'to throw into confusion' and > *kərkəb* (Moroccan Arabic) 'to roll', by the insertion of a sonorant consonant (*r, l, n,* or *w*) instead of the closing consonant of the first syllable. In Egyptian Arabic, there are many such verbs, e.g. *dardiš* 'to chat', *barbaʾ* 'to stare', *farfiš* 'to cheer up'. Neither final doubling nor initial doubling can be said to be productive rules in the language, and it is also difficult to differentiate between the various kinds of doubling in these quadriliteral verbs in terms of the change in semantics involved. Some sense of plurality or intensity seems to be common to most of them, while some also clearly exhibit a derogative connotation. The forms with full reduplication, though, seem to be much more productive, and they even compete with Form II verbs, the traditional derivational class for intensification (*kasar* 'to break' > *kassar* 'to break into many pieces', in case the root is biconsonantal, e.g. Egyptian Arabic *dašdiš* 'to reduce to fragments', not **daššiš*.

→ Gemination in Form II verbs has been viewed as an instance of reduplication, both in typological surveys (Pott 1862; Moravcsik 1978) and in language-specific studies (Ibrahim 1982; Greenberg 1991; El Zarka 2005), but also as the spreading of phonetic information (McCarthy 1982; Angoujard 1988). If gemination is classified as reduplication, it seems useful to subdivide it the way Pott (1862:17) did, into reduplication as the strengthening of segments, and prosodically defined reduplication or reduplication proper. Brockelmann (1908) already noted the close relationship of Form II with the other quadriliteral Forms. The older signification of plurality by this derivational class is superseded by the causative/factitive function in Modern Standard Arabic and Neo-Arabic, as may be seen in the loss of this meaning in verbs like *bakkā* 'to weep much', which in Modern Standard Arabic only means 'to make someone weep'; or *barraka* in *barraka l-jamal* 'to kneel down' (*barraka l-jamalu* 'the whole drove of camels kneeled down') or 'to make kneel down' (*barraka l-jamala* 'he made the camel kneel down'); the latter is the only possible interpretation nowadays.

In the nominal morphology of Classical, Modern Standard, and Neo-Arabic, reduplication of the second radical can signify intensity or habit, as in the adjective form *faʿʿāl*, e.g. *kāḏib* 'lying' > *kaḏḏāb* 'lying [habitual]'. Most other intensive adjectives are also formed by gemination of the second radical, for which the forms *fuʿʿāl, fiʿʿīl, fa/uʿʿīl,* and *fuʿʿal* are attested: *ḥussān* 'very handsome', *sikkīr* ‘drunken, drunkard’, *quddūs* ‘most holy’, *xullab* ‘deceitful’. As a noun, the form is used to signify professions, which also clearly involves repetition and habit: *ṭabbāx* ‘cook’, *ʿaṭṭār* ‘druggist’. A derived form is the instrumental noun, mostly with the feminine *-a(t)* suffix, to denote an instrument or a machine with which something is done regularly, e.g. the modern term *fattāḥa* ‘can opener’. The *fuʿʿāl* form is used to pluralize the lexicalized active participles, as opposed to the use of the sound plural: *kātib* ‘writing’ > *kātibūn* ‘writing [pl. masc.]’, but *kātib* ‘writer’ > *kuttāb* ‘writers’. Both functions are fairly productive. Apart from the last example, reduplication is not normally used to form plurals; the only minor instance is a handful of words in which a doubled second consonant appears instead of the expected semivowel: *dīnār/danānīr* ‘dinar’, not **dawānīr, dībāj/dabābīj* besides **dayābīj* ‘brocade’, thus giving rise to allomorphic alternation in the pattern. This drift, though, is more common in other Afro-Asiatic languages (Ratcliffe 1996).

The same tendency can be observed in → diminutive formation in Moroccan Arabic (Heath 1987:153; Ratcliffe 2001), e.g. the adjectives *zwen* > *zwiwən* ‘beautiful’, *smər* > *smimər* ‘of dark complexion’, and *nqi* > *nqiqi* ‘clean’, while others exhibit a default semivowel: *ṣġir* > *ṣġiwər* ‘little’. It is noteworthy that diminutives are often derived by reduplication in the languages of the world, but this is not the case in Classical Arabic. Whereas gemination definitely is an inherited process of word formation in Arabic (and Semitic generally), reduplication proper as a means of word formation seems to be an innovation.

The distinction between repetition and reduplication is not always clear-cut. Repetition is a very common feature of Arabic stylistics. Arabic written texts, in poetry and prose, are usually full of constructions that are characterized by the repetition of form, i.e. the prosodic pattern and the inflectional endings of the words involved. This kind of repetition is mostly reinforced semantically, as the doublets are most often synonyms or antonyms, for instance in the famous verse of the *Muʿallaqa* of Imruʾul-Qays *mikarrin mifarrin, muqbilin mudbirin maʿan* or in the popular saying *lā ʾaṣla lahu wa-lā faṣla* ‘he is not of noble descent’, but it seems that this strategy is mostly a matter of stylistics and thus should be considered outside the realm of reduplication. The same is probably true for syndetic reduplication, e.g. *ʾakṯar fa-ʾakṯar* ‘more and more’, where the connecting element indicates that repetition is involved. Furthermore, the construction could involve more than two copies of the word. The asyndetic variant, though, could more readily be considered as reduplication, such as the construction that occurs productively across varieties to signify a genuine instance of something, e.g. *ḥubb ḥubb* ‘real love’. The following Qurʾānic example exemplifies both possible meanings of such reduplications, intensity and distribution: *kallā ʾidā dukkat al-ʾarḍu dakkan dakkā / wa-jāʾa rabbuka wa-l-malaku ṣaffan ṣaffā* ‘Nay! When the earth is pounded to powder / And thy Lord cometh, and His angels, rank upon rank’ (Q. 89/21–22).

A frequent phenomenon of Arabic is root repetition, as in *maṭara l-maṭar* ‘it rained’ (→ paronomasia). Although the root is clearly doubled, this should probably not be regarded as reduplication, as it does not fulfill a special function or carry a special meaning. A more obvious case of syntactic reduplication is the modification of a verb by its → *maṣdar* in the *mafʿūl muṭlaq* construction (→ object, absolute; Maas 2005) to intensify the meaning of the verb. Here, the *maṣdar* is used instead of an adverb, which is not a proper category of Arabic, as in the Classical Arabic example above. This construction is reconstructed in Neo-Arabic varieties. Another case of syntactic reduplication might be pronoun reduplication (Bloch 1991), as in *raʾaytuhu huwa* ‘it is he I saw’, which is paralleled in other languages by such devices as clefting or accentuation of the pronoun. Syrian and Moroccan Arabic use reduplication with the demonstrative for focusing or specificity (Brustad 2000): *ha ṣ-ṣabi hāda* ‘this [specific] boy’ (Syrian) and *had l-mṛa hadi* ‘this [specific] woman’ (Moroccan).

BIBLIOGRAPHICAL REFERENCES

Angoujard, Jean-Pierre. 1988. «Gémination et redoublement». *Langues Orientales Anciennes – Philologie et Linguistique* 1.1–15.

Bloch, Ariel A. 1991. *Studies in Arabic syntax and semantics*. 2nd ed. Wiesbaden: O. Harrassowitz.

Brockelmann, Carl. 1908. *Grundriss der vergleichenden Grammatik der semitischen Sprache*. 2 vols.

Berlin: Reuther und Reichard. (Repr., Hildesheim: G. Olms, 1982.)
——. 1927. "Semitische Reimwortbildungen". *Zeitschrift für Semitistik und Verwandte Gebiete* 5.6–38.
Brustad, Kristen E. 2000. *The syntax of spoken Arabic: A comparative study of Moroccan, Egyptian, Syrian, and Kuwaiti dialects*. Washington, D.C.: Georgetown University Press.
El Zarka, Dina. 2005. "On the borderline of reduplication: Gemination and other consonant doubling in Arabic verbal morphology". Hurch (2005:369–394).
Ferguson, Charles. 1964. "Baby talk in six languages". *American Anthropologist* 66.6/2.103–114.
Greenberg, Joseph. 1991. "The Semitic 'intensive' as verbal plurality". *Semitic studies in honour of Wolf Leslau on the occasion of his 85th birthday*, ed. Alan S. Kaye, 576–587. Wiesbaden: O. Harrassowitz.
Heath, Jeffrey. 1987. *Ablaut and ambiguity: Phonology of a Moroccan Arabic dialect.* New York: State University of New York.
Hurch, Bernhard (ed.). 2005. *Studies on reduplication.* Berlin: Mouton de Gruyter.
Ibrahim, Muhammad. 1982. "Radical reduplication in Arabic". *Zeitschrift für Arabische Linguistik* 9.82–87.
Imouzaz, Said. 2002. *Interaction des contraintes dans la morphologie non-gabaritique de l'arabe marocain de Casablanca: Témoignages pour la théorie de l'optimalité.* Ph.D. diss., Hassan II University, Rabat.
Maas, Utz. 2005. "Syntactic reduplication in Arabic". Hurch (2005:395–429).
McCarthy, John. 1982. "Prosodic templates, morphemic templates, and morphemic tiers". *The structure of phonological representations*, I, ed. Harry van der Hulst and Norval Smith, 191–223. Dordrecht: Foris.
Moravcsik, Edith. 1978. "Reduplicative constructions" *Universals of human language.* III. *Word Structure*, ed. Joseph Greenberg, 297–334. Stanford, Calif.: Stanford University Press.
Pott, August F. 1862. *Doppelung (Reduplikation, Gemination) als eines der wichtigsten Bildungsmittel der Sprache, beleuchtet aus Sprachen aller Welttheile.* Lemgo und Detmold: Meyer.
Procházka, Stephan. 1995. "Semantische Funktionen der reduplizierten Wurzeln im Arabischen". *Archív Orientální* 63.39–70.
Ratcliffe, Robert R. 1996. "Drift and noun plural reduplication in Afroasiatic". *Bulletin of the School of Oriental and African Studies* 59:2.296–311.
——. 2001. "Analogy in Semitic morphology: Where do new roots and new patterns come from?". *New data and new methods in Afroasiatic linguistics: Robert Hetzron in memoriam,* ed. Andrzej Zaborski, 153–162. Wiesbaden: O. Harrasowitz.
Reckendorf, Hermann. 1909. *Über Paronomasie in den semitischen Sprachen.* Giessen: Töppelmann.
Rubino, Carl. 2005. "Reduplication: Form, function and distribution". Hurch (2005:11–29).

DINA EL ZARKA (University of Graz)

Reflexives

In a sentence such as *John sees himself*, the subject and the object of the verb are one and the same person, namely *John*. Crucially, languages generally do not allow such a meaning to be expressed with something like *John sees John*, or *John sees him*. Most, if not all, languages require a special pronominal form for the object. This form is generally called a 'reflexive'. In other words, a reflexive is a pronominal form that indicates that two arguments of a verb have the same referent. The syntactic expression with which the reflexive is coreferential is called the antecedent of the reflexive.

Reflexives can be simplex or complex. Simplex reflexives are reflexives like French or Spanish *se*, Italian *si*, Dutch *zich*, Norwegian *seg*, Finnish *itse*, etc. These reflexives have a form that cannot be decomposed. Complex reflexives, on the other hand, are composed of a pronominal element combined with some meaningful element such as *self, one's own, body, soul, limbs*, etc. Typical examples are English *himself*, Dutch *mijzelf* 'me-self', Hebrew *aṣm-o* lit. 'his bone', etc. A complex reflexive can also consist of a simplex reflexive combined with a *self*-type element, such as Dutch *zichzelf* or Italian *si stesso*.

The two types of reflexives show different syntactic behavior (Reinhart and Reuland 1991, 1993). The most conspicuous difference is that simplex reflexives always take the subject as the antecedent, while complex reflexives can take any argument of the verb as antecedent, as long as it is structurally higher. (There is a general argument hierarchy whose effects are visible in many languages: subject > indirect object > direct object > prepositional object.) The example in (1) illustrates this possibility (here and below, coreference is indicated with subscripted indices).

(1) *John$_i$ protected Mary$_j$ from herself$_j$*

In general, reflexives must take a co-argument as antecedent. It is generally not possible for the antecedent to be in a higher clause, as in (2).

(2) *John*$_i$ *thought that Mary loved him*$_i$*/ *himself*$_i$

The sentence in (2) is correct with *him* in the object position of the subclause. Using *himself* in this position, with the intention that it refer to *John*, is not possible. Reinhart and Reuland (1991) argue that complex reflexives function as reflexivizers: they turn the predicate of which they are an argument into a reflexive predicate.

There are, however, exceptions to this observation, some of which are systematic. For example, many languages have infinitival subclauses with subjects that are case-assigned in the embedding clause (phrases of this type are called Exceptional Case Marking, or ECM, in the generative tradition), as in (3).

(3) *Mary wanted John to win the race*

Example (3) consists of a finite main clause (*Mary wanted…*) and an infinitival subclause (*John to win the race*). *John* is semantically the subject of the infinitival verb, as it is the one doing (or supposed to do) the winning. Example (4) shows that the noun phrase in this position takes oblique (accusative) case. Only the oblique pronoun *him* is grammatical as the subject of constructions with Exceptional Case Marking; the nominative pronoun *he* is not.

(4) *Mary wanted him/*he to win race*

If the subject of the infinitival verb is coreferential with the subject of the main clause, the embedded subject is expressed as a reflexive and cannot be a pronoun, as shown in (5). Obviously, a pronoun is grammatical in this position if it is not coreferential with the subject of the main clause.

(5) *Mary*$_i$ *wanted herself*$_i$*/*her*$_i$ *to win the race*

Whether structures such as that in (5) do indeed contradict the observation that reflexive and antecedent must be in the same clause depends on the exact analysis one adopts for such structures: if one assumes that the embedded subject is in the main clause (cf. Postal 1974; Chomsky 2005), there is no contradiction.

More problematic are those cases in which a reflexive appears that does not have an antecedent. *Myself* in (6) would require a 1st person singular antecedent, but none is available.

(6) *This paper was written by Ann and myself*

Reflexives that lack an antecedent are sometimes called 'logophoric'. In the analysis presented by Reinhart and Reuland (1991), a complex reflexive can only function as a reflexivizer if it is an argument of a fully assigned predicate, i.e. a predicate of which all (potential) arguments are overtly realized. Because *myself* in (6) is not an argument itself but merely part of an argument (i.e. of *Ann and myself*), it cannot function as a reflexivizer. Reflexives that cannot function as reflexivizers can be used logophorically, if the language in question allows it.

Logophoric reflexives have the property that they can be replaced with a pronoun without changing the grammaticality or the meaning of the clause (although some expressiveness may be lost), as in (7).

(7) *This paper was written by Ann and me*

Arabic uses the noun *nafs* 'soul' to form reflexive expressions (as noted by Wright 1981:II, 280–282, other words are sometimes used in Classical Arabic, such as *ʿayn* 'eye', *wajh* 'face', *rūḥ* 'spirit', *ḏāt*, 'essence', *ḥāl* 'state'). This noun is combined with a possessive pronominal suffix, which makes it a typical example of a complex reflexive, as in (8) (most examples here are taken from Kremers 1997).

(8) *yuḥaddiṯu* *ʾaḥmad-u*$_i$ *nafs-a-hu*$_i$
speak.3ms Ahmad-Nom self-Acc-his
'Ahmad talks to himself'

The antecedent of the reflexive in (8) is the subject of the clause *ʾaḥmad*. The reflexive may also be a prepositional object, as in (9).

(9) *li-yudāfiʿa* *š-šaʿb-u*$_i$ *ʿan*
to defend.3ms the-people-Nom from
nafs-i-hi$_i$ *ḍidda…*
self-Gen-its against…
'so that the people can defend themselves against…'

Because the Arabic reflexive is a complex reflexive, one expects that non-subject arguments can function as antecedents, as in (10).

(10) *ḥattā* *takšifa-hu*$_i$ *l-baṭal-at-u*
so.that reveal.3fs-him the-heroine
ʾamāma *nafs-i-hi*$_i$
before self-Gen-him
'so that the heroine reveals him before himself'

In (10), the antecedent of the reflexive *nafs-i-hi* is the object of the verb, the suffix pronoun *-hu*. Note that with prepositional objects, there is some optionality in the use of reflexives. Especially prepositional phrases that tend toward adjuncthood may contain pronouns, even if such a pronoun is coreferential with the subject, as in (11).

(11) *yaznaqūna*$_i$ *ṣ-ṣiġār-a* *bayna-hum*$_i$
squeeze.3mp the-small-Acc between-them
'They squeezed the small ones between them'

In (11), the prepositional phrase *bayna-hum* contains a pronominal suffix that is coreferential with the subject of the clause. From the viewpoint of Reinhart and Reuland's (1991) theory, one can argue that *bayna-hum* is not an argument of *yaznaqūna* and that therefore it is not the case that the verb has two coreferential arguments, so that technically it is not a reflexive predicate.

The reflexive *nafs* is very common with → *maṣdar*s, as in (12).

(12) *naḥnu* *lā* *nurīdu* *ḥaṣr-a*
we not wish.1p restraining-Acc
ʾanfus-i-nā
selves-Gen-our
'We do not wish to restrain ourselves'

In (12), the antecedent of the reflexive is apparently the subject of the finite verb *nurīdu*. However, the data show that the *maṣdar* forms the domain of interpretation for the reflexive; that is, the antecedent of the reflexive must itself be an argument of the *maṣdar*. It is possible, however, that this argument is not expressed overtly. In other words, the reflexive in (12) signals that the subject of the *maṣdar* is *naḥnu*, as well. Example (13) illustrates this point more clearly.

(13) *lā* *yaḥṣulūna* *ʿalā* *mā*
not obtain.3mp on what
yakfī *li-ʾinqāḏ-i*
suffice.3ms to-saving-Gen
ʾanfus-i-him *min* *al-faqr-i*
selves-Gen from the-poverty-Gen
'They do not obtain enough to save themselves from poverty'

In (13), the *maṣdar* *ʾinqāḏ*, which has a reflexive argument, is the object of the verb *yakfī*. The subject of this verb is *mā*, which means that the clause containing the *maṣdar* does not contain an antecedent for the reflexive. Note, however, that the meaning of the sentence changes dramatically when the reflexive is replaced with a pronoun, as in (14).

(14) *lā* *yaḥṣulūna* *ʿalā* *mā*
not obtain.3mp on what
yakfī *li-ʾinqāḏ-i-him*
suffice.3ms to-saving-Gen-their
min *al-faqr-i*
from the-poverty-Gen
'They do not obtain enough to save them from poverty'

In (14), the non-overt, understood subject of *ʾinqāḏ* is different from the object: in contrast to (13), the ones doing the saving are not the ones being saved. Because of this difference in meaning, it cannot be argued that the reflexive in (13) is used logophorically. That is, (13) and (14) together show that *maṣdar*s are domains for the interpretation of reflexives and that a non-overt subject must be assumed.

One further example to illustrate this point is given in (15).

(15) *šadād-u* *yuwāṣilu*
Shadad continue.3ms
rtidāʾ-a-hu *li-malābis-i-hi*
putting.on-Acc-his of-clothes-Gen-his
wa-tahyiʾat-a *nafs-i-hi*
and-preparing-Acc self-Gen-his
'Shadad continues to put on his clothes and to prepare himself'

The example in (15) is a telling one, because it contains two *maṣdar*s that are both objects to the finite verb *yuwāṣilu* and both in construct state with a pronominal element. The first *maṣdar, irtidāʾ,* has a pronominal suffix, the second, *tahyiʾa*, a reflexive. Crucial here is the fact that the two pronominal elements are in the same structural relation with respect to their antecedent, *šadād.* Yet, one is (obligatorily) a pronoun, while the other is (obligatorily) a reflexive. The only way to account for these facts is to assume, as before, that the *maṣdar* is the domain of interpretation for the reflexive, where a non-overt subject is present.

A different issue arises with the so-called *ʾafʿāl al-qulūb* 'verbs of the heart' (cf. Wright 1981:II, 272). These verbs are similar to the constructions with Exceptional Case Marking mentioned above. The verbs of the heart are verbs that have a perceptual or cognitive meaning, such as *raʾā* 'to see, think, know', *wajada* 'to find, perceive', *ḏ̣anna* 'to think', *iʿtabara* 'to consider', etc. These verbs take two complements, the first a noun and the second either a nominal or a verbal predicate. The first complement is the subject of the predicate, as in (16a, b).

(16) a. *ḏ̣anantu* *zayd-an* *karīm-an*
thought.1s Zayd-Acc noble-Acc
'I thought Zayd noble'
b. *wajadtu-hu* *yašrabu* *xamr-an*
found.1s-him drink.3ms wine-Acc
'I found him drinking wine'

If the subject of the embedded clause is coreferential with the subject of the verb of the heart, Modern Standard Arabic generally uses a reflexive for the embedded subject, as in (17).

(17) *yaḏ̣unnu* *nafs-a-hu*
think.3ms self-Acc-his
markaz-a *l-kawn-i*
center-Acc the-universe-Gen
'He thinks himself the center of the universe'

A similar example, with a verbal predicate, is (18).

(18) *yaʿtabirūna* *ʾanfus-a-hum* *sabaqū*
consider.3mp selves-Acc-their lead.3mp
'They consider themselves to have been ahead'

Note that while these constructions superficially appear to be similar to the constructions with Exceptional Case Marking, they are in fact quite different. In English ECM-constructions, the embedded clause has an infinitival verb, while in Arabic, the verb of the heart may take a clause with a finite verb as its complement.

What this suggests is that verbs of the heart are verbs that take a nominal sentence (*jumla ismiyya*) as complement. The topic (*mubtadaʾ*) of this nominal sentence occupies the object position of the verb of the heart, and consequently is assigned accusative case.

In the theory of Reinhart and Reuland (1991), the fact that the embedded subject is in the object position of the verb of the heart means that this subject is, at least in syntactic terms, an argument of the verb of the heart. Therefore, if the embedded subject is coreferential with another argument of the verb of the heart (generally the subject), this verb is a reflexive predicate, and the embedded subject must be a reflexive. Given this theory, then, the occurrence of a reflexive in structures such as in (17) and (18) is expected.

Modern Standard Arabic allows logophoric use of *nafs*, although this use appears to be limited. Kremers (1997) only reports logophoric use of 1st person plural reflexives, as in (19).

(19) *xayr-un* *min* *rabb-i-nā*
good-Nom from lord-Gen-our
wa-šarr-un *min* *ʾanfus-i-nā*
and-evil-Nom from selves-Gen-our
'Good comes from our Lord, evil comes from ourselves'

The reflexive *ʾanfus-i-nā* in (19) has no antecedent: it would require a 1st person plural pronoun as antecedent, but none is present. Note that the reflexive in (19) could be replaced with a pronoun (*min-nā*) without affecting the grammaticality of the phrase. The expressiveness would be affected, of course.

Generally, reflexives in Arabic have a pronominal suffix. There are, however, instances of *nafs* that lack a pronominal suffix, although they must be interpreted as reflexives. In such cases, *nafs* appears with the definite article and has the meaning of an impersonal reflexive, as in (20a) and (20b).

(20) a. *lākin* *lā* *jadwā* *min*
but not use from
muxādaʿat-i *n-nafs-i*
deceiving-Gen the-self-Gen
'But there is no use in deceiving oneself'

b. *la-qad* *kānat* *furṣat-an*
Perf was.3fs opportunity
li-t-taʿarruf-i *ʿalā n-nafs-i*
to-the-knowing on the-self-Gen
'It was an opportunity to get to know oneself'

There are several lexicalized collocations containing an impersonal *nafs*, in which *nafs* has a meaning similar to English *self* in compounds: *aṯ-ṯiqa fī n-nafs* 'self-confidence', *ad-difāʿ ʿan an-nafs* 'self-defense', *ḍabṭ an-nafs* 'self-control'.

Another method of expressing reflexive verb meanings in Arabic should be mentioned here. Certain Form templates in Arabic have a *t*-affix. These *t*-Forms all correlate with 'non-*t*' Forms, and often have a reflexive (or in the case of Form VI, → reciprocal) meaning. Thus, Form V *takattaba* correlates with Form II *kattaba*, Form VI *takātaba* correlates with Form III *kātaba*, Form VIII *iktataba* correlates with the ground Form *kataba*, and Form X *istaktaba* correlates with Form IV *ʾaktaba* (which historically had the variant *saktaba*). Similarly, Form VII *inkataba* can have a reflexive meaning, related to the ground Form *kataba* (see Reckendorf 1967 for some discussion).

Often, however, verbs with a reflexive verb template have a meaning closer to English intransitive or even passive verbs. For example, *kataba* 'to write' has a *t*-Form *iktataba* 'to write oneself', i.e. 'to register', which can easily be considered reflexive, but Form VIII of the verb *rafaʿa* 'to lift, raise' does not mean 'to lift oneself' but rather 'to be lifted' or 'to rise'. Such a clustering of meanings is not uncommon in language: the simplex reflexive *se* in the Romance languages shows similar effects.

It should also be noted that even though many *t*-Form verbs indeed have a reflexive (or passive/intransitive) meaning, *t*-Form verbs do not represent a productive reflexivization strategy in Modern Standard Arabic (and were at best marginally productive in Classical Arabic). Verbs that show this pattern are lexicalized, not produced through productive derivational morphology.

BIBLIOGRAPHICAL REFERENCES

Chomsky, Noam. 2005. *On phases*. Ms., Massachusetts Institute of Technology.

Kremers, Joost. 1997. *How Arabs speak to each other about themselves: A study of* nafs *and* baʿḍ *in Modern Standard Arabic*. MA thesis, University of Nijmegen.

Postal, Paul. 1974. *On raising*. Cambridge, Mass.: MIT Press.

Reckendorf, Hermann. 1967. *Die syntaktischen Verhältnisse des Arabischen*. Leiden: E.J. Brill. (Photomechanical repr. of the 1895 edition.)

Reinhart, Tanya and Eric Reuland. 1991. "Anaphors and logophors: An argument structure perspective". *Long-distance anaphora*, ed. Jan Koster and Eric Reuland. Cambridge: Cambridge University Press.

——. and Eric Reuland. 1993. "Reflexivity". *Linguistic Inquiry* 24:4.657–720.

Wright, William. 1981. *A grammar of the Arabic language*. Beirut: Librarie du Liban. (Repr. of the 3rd ed. of 1896.)

JOOST KREMERS (University of Frankfurt)

Register

1. DEFINITION

'Register' refers to a variety of language defined according to its use in social situations, for example the register of scientific, religious, or formal English. In Hallidayan linguistics, the term is specifically opposed to those varieties of language which are defined according to the characteristics of the user's regional or class dialect (Crystal 1991:295). Ferguson also posits that "a communication situation that recurs regularly in a society (in terms of participants, setting, communicative functions, and so forth) will tend over time to develop identifying markers of language structure and language use, different from the language of other communication situations" (1994:295). That is to say, speakers who take part in the same situation regularly develop a similar vocabulary, similar phonological features (such as intonation), and similar structural and syntactic characteristics. Register markers are, for example, special terms for recurrent objects and events, formulaic sequences, or routines. These features are all used to facilitate conversation. According to Myers-Scotton (1997), people try to facilitate the process of communication with the least effort on their part and try to leave a great effect on the listener. Ferguson adds that the tendency to develop registers is inherent in man and starts very early on in children.

For different registers to be used in a community, there has to be what Hudson (1994:303) calls a 'presuppositional background' shared by the participants in a specific situation. He also thinks that casual utterances are more relaxed, and, therefore, forms that may seem inappropriate otherwise become appropriate, while in noncasual utterances, there is a fixed system (Hudson 1994:295). Indeed, there should be general agreement among persons in a specific community in evaluating the appropriateness of an utterance. This agreement is because of the 'presuppositional background' mentioned above.

A number of linguists, including Gumperz, differentiate between formal and informal types of interaction in relation to register. Note the difference in usage mentioned by Gumperz: in formal types of interaction, "modes of speaking are narrowly prescribed [...] while in more informal types of interaction, there is a loosening of the co-occurrence of restrictions, and forms which would not appear together in transactional encounters may now co-occur" (1964:10). This ties in with the concept of → diglossia mentioned by Ferguson (1959), which posited that in the Arab world, there are two varieties, a High one (Standard and Classical Arabic), which is used on formal occasions, and a Low one (the vernacular of different countries), which is used in informal situations.

2. FACTORS THAT INFLUENCE REGISTER

Register occurs in a specific community usually because of the growth of this speech community (Kay 1977:30). Kay adds that there are factors that encourage the occurrence of register, such as the emergence of occupational, regional, and other subcultures, and the division of labor within a community. These factors may result in the creation of special fields of knowledge, as well as a specialized terminology to be able to communicate about these fields. The general process of social evolution "produces speech communities in which situations calling for autonomous speech occur with increasing frequency" (Kay 1977:29). Kay also (1977:21–22) differentiates between 'autonomous' and 'nonautonomous' speech. Educated speakers in a formal academic context speak in a specific register, different from the one they speak in when playing baseball, quarreling, etc.

In addition, Bernstein posits that the more the intentions of the other participants in a verbal interaction are taken for granted, "the more likely that the structure of the speech will be simplified and the vocabulary drawn from a narrow range" (1964:60–61). This emphasizes the relation between register and presuppositional background knowledge between participants. The degree of knowledge taken for granted by participants will influence the register used for communication.

Hudson (1994:296) believes that formal occasions have a higher level of consistency, because structures keep recurring, and there tend to be no innovations from speakers. An example of this is ceremonies as opposed to conversations among friends.

Kay (1977:21–22), too, differentiates between formal and informal speech. He believes that in formal speech there is an increased use of longer and syntactically more complex sentences. There is also more explicit and varied vocabulary, and more edited delivery. Hudson (1994:300) mentions the difference in word order as a difference between formal and informal speech. There is also a difference in the use of introductory and connective particles (cf. Ferguson 1959). Whether these postulations about formal and informal registers are valid cannot be proven except with more studies. Meanwhile, in the next section some of the varied examples, which cover register, are given.

3. EXAMPLES

First, from a historical perspective, there are a number of useful studies carried out on English and other languages, and there is a need for similar studies to be carried out on Arabic. Ferguson (1994:20) mentions Kittredge (1982), who examines different registers, for instance the language of aviation hydraulics, cookbook recipes, regional weather forecasts, and stock market reports. He found that these differ lexically, phonologically, and syntactically. Culy (1987) produced a diachronic study of cookbook recipe register from the 15th century to modern times. He speculates that the language of cookbooks was not different from other written varieties of English, and mentions as an example the omission of definite object noun phrases.

Blau (1985:75) analyzes the use of different registers of → Judaeo-Arabic, concentrating on the writings of Maimonides. He notes that Judaeo-Arabic authors use a lower register, less Classical Arabic, which may reflect the gap between Classical Arabic and its uniformity in theory and its variability in practice. Blau (1985:76) concludes that the same author uses different registers, depending on the literary genre. He also mentions one of the problems of authors writing in Standard Arabic, namely dialogues. Even in the classical period there was a problem in writing dialogues in Standard Arabic, since people do not converse in the standard variety. Medieval authors faced this problem, too, and used a different register for conversations. Schippers (1991) studies style and register in Arabic, Hebrew, and Romance strophic poetry. He compares Hebrew and Arabic Andalusian poetry with Romance poetry, suggesting that Arabic and Hebrew poetry exhibits a similarity with Romance love poetry in their lexical, rhetorical, and syntactic possibilities.

Meanwhile, synchronic studies include a study published by Ferguson and Rice (1997), who analyze the register of Iraqi children's rhymes. They divide these into four groups: rhymes connected with folktales, counting-out rhymes, lullabies, and miscellaneous. They note, for example, the presence of echo words as a marker of this specific register, the register of children's rhymes. Ferguson (1997) also studied the register of Arabic baby talk, concentrating on Syrian Arabic baby talk (→ caretaker talk). Ferguson (1997:179) believes that the register of baby talk has special features, e.g. special intonation patterns, which are called in German *Ammenton*. His study, however, concentrates on lexical items. He notices that when comparing English and Arabic baby talk, one finds similarities. The presence of reduplication and diminutives, special sets of terms for parents, lack of inflection for specific words which are otherwise inflected, and frequent use of words as sentence words are all cases in point. This study is important because it emphasizes the element of universality, now of great significance in linguistic studies, which entails that all languages can function in a very similar way.

4. Conclusion

The term 'register' is quite an abstract and broad term. As shown above, examples covering register range from the stylistic analysis of poetry to the analysis of baby talk in modern-day Syria. It may therefore be useful to study register in relation to other, connected concepts, such as diglossia and language → variation. The diglossic situation in the Arab world in particular adds a new dimension to the study of register. Likewise, historical scientific texts in Arabic and the language of specialized professionals would provide a rewarding object of further study.

Bibliographical references

Belnap, Kirk and Niloofar Haeri (eds). 1997. *Structuralist studies in Arabic linguistics: Charles A. Ferguson's papers, 1954–1994*. Leiden: E.J. Brill.

Bernstein, Basil. 1964. "Elaborated and restricted codes: Their social origins and some consequences". *The ethnography of communication*, ed. John Gumperz and Dell Hymes, 55–69. (= *American Anthropologist* 66:6/2.) Menosha, Wisc.: American Anthropological Association.

Biber, Douglas and Edward Finegan (eds). 1994. *Sociolinguistic perspectives on register*. Oxford: Oxford University Press.

Blau, Joshua. 1985. "Notes on the use of different registers of Judeo-Arabic by one author". *Hebrew Annual Review* 9.75–78.

Corriente, Federico and Ángel Saenz-Badillos (eds.). 1991. *Poesía estrófica: Actas del Primer Congreso Internacional sobre Poesía Estrófica Árabe y Hebrea y sus Paralelos Romances*. Madrid: Universidad Complutense.

Crystal, David. 1991. *A dictionary of linguistics and phonetics*. 3rd ed. Oxford: Blackwell.

Culy, Cristopher. 1987. *Cookbook linguistics: Recipes and linguistic theory*. Ph.D. diss., Stanford University.

Ferguson, Charles. 1959. "Diglossia". *Word* 15.325–340.

——. 1994. "Dialect, register and genre: Working assumptions about conventionalisation". Biber and Finegan (1994:15–30).

——. 1997. "Arabic baby talk". Belnap and Haeri (1997:179–187). (1st publ., *For Roman Jakobson*, ed. Morris Halle, Horace G. Lunt, Hugh McLean, and Cornelis H. van Schooneveld, 121–128. The Hague: Mouton, 1956.)

—— and Frank Rice. 1997. "Iraqi children's rhymes". Belnap and Haeri (1997:188–197). (First publ., *Journal of the American Oriental Society* 80 [1960] 335–340.)

Givon, Talmy. 1979. "From discourse to syntax: Grammar as a processing strategy". *Discourse and syntax*, ed. Talmy Givon, 81–112. New York: Academic Press.

Gumperz, John. 1964. "Linguistic and social interaction in two communities". *The ethnography of communication*, ed. John Gumperz and Dell Hymes, 137–153. (= *American Anthropologist* 66: 6/2.) Menosha, Wisc.: American Anthropological Association.

Hudson, Alan. 1994. "Diglossia as a special case of register variation". Biber and Finegan (1994:294–314).

Kay, Paul. 1977. "Language evolution and speech style". *Sociocultural dimensions of language change*, ed. Ben G. Blount and Mary Sanches, 21–33. New York: Academic Press.

Kittredge, Richard. 1982. "Variation and homogeneity of sublanguages". *Sublanguage: Studies of language in restricted semantic domains*, ed. Richard Kittredge and John Lehrberger, 107–137. Berlin: de Gruyter.

Myers-Scotton, Carol. 1997. *Duelling languages*. Oxford: Oxford University Press.

Schippers, Arie. 1991. "Style and register in Arabic, Hebrew and Romance strophic poetry". Corriente and Saenz-Badillos (1991:311–325).

Reem Bassiouney (The Georgetown University)

Relative Clause

1. Subordination and relative clauses

→ Subordination is a syntactic configuration in which a clause functions as a constituent within the structure of another clause either as an argument (subject or complement) or as a modifier. What distinguishes subordinate clauses from independent (or main) clauses is that the former cannot be used in isolation. For example, the subordinate purpose clause *li-yaktuba* 'for him to write' and *ʾan yaktubū* 'that they write' cannot be used as independent sentences. Even a subordinate relative clause with no relative pronoun cannot be used independently because its anaphoric elements are uninterpretable without the antecedents in the main clause.

Subordinate clauses functioning as complements occur in the same positions that allow nominal complements, as both constructions are licensed by the same syntactic heads, including verbs, participles, nouns, and prepositions. Some verbs, such as *qāla* 'to say', *ʾadraka* 'to realize', *ʾarāda* 'to want', and *saʾala* 'to ask', select for clausal complements headed by a complementizer. In Standard Arabic, the choice of the complementizer depends on the selection restrictions of the licensing verb and the nature of the situation described in the subordinate clause. The complementizer *ʾanna* is used in clauses that report on actual events, while *ʾan* is used with subordinate clauses that describe irrealis situations. The structural properties of subordinate clauses are directly influenced by the complementizers; *ʾanna* requires the SVO word order with an accusative subject, while *ʾan* requires the VSO word order with a subjunctive verb. Subordinate clauses may occur in subject positions provided that they are used sentence-finally, as in *yajibu ʾan*... 'it is a necessity that...' and *min al-maʿrūfi ʾanna*... 'it is known that...', unless they are preceded by the topicalization marker *ʾammā* 'as for'.

Subordinate clauses that are used as modifiers include adverbial and relative clauses. Adverbial clauses are headed by temporal adverbs such as *lammā* 'when', *baynamā* 'while', and *ḥattā* 'until', or by purpose/reason conjunctions, such as *kay* 'in order to' and *fa-* 'so'. Conditional clauses are also adverbial in nature, and they have the same distribution patterns, but they are headed by the complementizers *ʾiḏā*, *ʾin*, and *law*, all of which mean 'if' (→ *jazāʾ*). What sets relative clauses apart from other types of subordinate clauses is that they modify noun phrases; hence, they have distinct syntactic and semantic properties.

Relative clauses are embedded adjunct constructions that modify noun phrases regardless of their syntactic positions, i.e., they can modify subjects, direct and indirect objects, complements of prepositions, and other noun phrases. Typically, a relative marker, such as *allaḏī* 'that', immediately follows the modified head noun phrase if it is definite, as in (1a) and (1b). Relative markers, however, are ruled out if the head noun is indefinite, as in (1c) (Galal 2004). The head noun cannot be lexically represented within the embedded clause, as indicated by the ungrammaticality of (2a). Rather, a coreferential pronominal element (a resumptive pronoun) that agrees with the head noun for number and gender is used in the corresponding position within the relative clause. In Standard Arabic, a resumptive pronoun can be left out only if the head noun is definite and it corresponds to the direct object of the embedded clause, as in (2b) (→ resumption).

(1a)	*waṣala*	*l-ʾustāḏ*	*allaḏī*
	arrived.3ms	the-teacher	Rel
	sawfa	*yulqī*	*l-muḥāḍara*
	Fut	give.3ms	the-lecture

'The teacher who will give the lecture arrived'

(1b) *waṣala l-ʾustāḏ sawfa
arrived.3ms the-teacher Fut
yulqī l-muḥāḍara
give.3ms the-lecture
*'The teacher will give the lecture arrived'

(1c) jāʾa rajul (*allaḏī) yaḥmilu
came.3ms man (*Rel) carry.3ms
la-ka risāla
for-you message
'A man carrying a message for you came'

(2a) *ʾarsaltu l-xiṭāb allaḏī katabtu
sent.1s the-letter Rel wrote.1s
l-xiṭāb
the-letter
*'I sent the letter which I wrote the letter'

(2b) ʾarsaltu l-xiṭāb allaḏī katabtu(-hu)
sent.1s the-letter Rel wrote.1s(-it)
'I sent the letter that I wrote'

2. Relative markers and their categorical status

Descriptive grammars of Modern Standard and Classical Arabic classify relative markers into two categories, specific and generic, based on the observation that members of the former class have morphologically represented agreement features for gender, number, and case, whereas the latter category includes only invariant relative markers (Ḥassān 1974; Fayyāḍ 1995; ʿAbd al-Ġanīy 2002). There are twelve relative markers with morphologically represented agreement features (→ relative pronoun): allaḏī (masc. sg.); allatī (fem. sg.); allaḏāni (masc. du. nom.); allaḏayni (masc. du. acc./gen.); allatāni (fem. du. nom.); allatayni (fem. du. acc./gen.); alladīna (masc. pl.); allātī, allāʾī, and allawātī (fem. pl.); and alʾulā and alʾūlāʾi (neutral pl.) (Wright 1967:I, 271). The invariant relative markers are *man* and *mā*. Even though these two have no morphologically represented agreement features, *man* is mostly used for human referents, while *mā* is usually used for nonhuman referents, as in (3a) and (3b).

(3a) wajadtu bayna l-ḥuḍūr man/*mā
found.1s among the-audience Rel
yušāriku-nī r-raʾy
share.3ms-cl1s the-view
'I found among the audience someone who shares my view'

(3b) nadima ʿalā mā/*man faʿala
regretted.3ms on Rel did.3ms
'He regretted what he did'

One distinction between the two categories of relative markers is that, unlike *man* and *mā*, *allaḏī* and its allomorphs are believed to be morphologically complex forms. Wright (1967) argues that *allaḏī*-type relativizers are made up of the definite article *al-*, the demonstrative *l*, and the masculine demonstrative *ḏā* or its feminine counterpart *tī*, in addition to number- and case-marking suffixes. Another distinction between the two categories is that *man* and *mā* can only be used in free relatives, while *allaḏī* and its variants can modify head nouns as well. The examples in (4a)–(5b) demonstrate that *allaḏī*-type relativizers as well as *man* and *mā* can be used in argument positions without head nouns, but *man* and *mā* cannot modify head nouns.

(4a) ḏakkarta-nī bi-mā/bi-llaḏī
reminded.2ms-cl1s with-Rel
qultu-hu
said.1s-cl3ms
'You reminded me of what I said'

(4b) tasāmartu maʿa man/allaḏīna
chatted.1s with Rel
kānū fī l-ḥafla
were.3mp in the-party
'I chatted with (those) who were at the party'

(5a) qaraʾtu l-kitāb allaḏī/*mā
read.1s the-book Rel
raššaḥta-hu
recommended.2ms-cl3ms
'I read the book that you recommended'

(5b) zāra karīm ṣadīqa-hu
visited.3ms Karim friend-his
llaḏī/*man kāna marīḍ
Rel was.3ms sick
'Karim visited his friend who was sick'

Ibn Hišām (*Qaṭr an-nadā* 99) and Ibn ʿAqīl (*Šarḥ al-ʾAlfiyya* I, 123) note that some old Arabian dialects had additional relative markers. For example, the relativizer *allaḏūna* (masc. pl. nom.) was used in the speech of Huḏayl and ʿAqīl in contrast with *allaḏīna* (masc. pl. acc./gen.), while the invariant *ḏū* was common in the dialect of Ṭayyiʾ. Other Classical Arabic relative markers include the invariant *ʾayy, al-*, and *ḏā*. The marker *ʾayy* is used only in possessive relatives, as in (6a); *al* is usually used with participial predicates, yet, there are some examples with verbal and prepositional predicates, as in (6b) and (6c). The categorical status of *ḏā* is rather problematic, especially since it is only used in clefted questions with WH-elements, as in (6d) (see Al-Bazi 1983).

(6a) *ḍaʿ* *al-kuswa* *wa-l-ʾiklīl*
put.Imper.2ms the-cloak and-the-garland
bayna *yaday* *ʾiraxt* *li-taʾxuḏ-a*
between hands Irakht for-take.2ms-Subj.
ʾayya-hā *šāʾat*
Rel-cl3fs wanted.3fs
'Present the cloak and the garland to Irakht to take whichever of them she wants' (Ibn al-Muqaffaʿ, *Kalīla wa-Dimna 205*)

(6b) *man* *lā* *yazālu* *šākiran*
Rel Neg cease.3ms thankful
ʿalā *al-maʿa-hu*
on Rel-with-him
lit. 'he who does not cease to be thankful for what is with him'
'he who is still thankful for what he has' (Wright 1967:I, 269)

(6c) *mā* *ʾanta* *bi-l-ḥakam*
Neg you.ms with-the-judge
al-turḍā *ḥukūmatu-hu*
Rel-accept.3fs.Pass sentence.Nom-his
'You are not the judge whose sentence is to be accepted' (Wright 1967:I, 269)

(6d) *man* *ḏā* *ʿalā* *tark*
who Rel on abandoning
aṣ-ṣalāt *yalūmu*
the-prayer blames.3ms
'Who is it that blames [others] for abandoning prayers?' (Ibn al-Jawzī, *ʾAxbār* 112)

Many contemporary spoken dialects have a single invariant relative marker used in all relativized contexts, such as *illi* in Egyptian and Baghdadi Arabic (Wise 1975; McCarthy and Raffouli 1965), *illi* in the dialects of Dubai and Riyadh (Hoffiz 1995; Feghali 2004), *alli* in the urban dialect of the Ḥijāz (Sieny 1978), *lli* in the Tunisian dialect of Sūsa (Talmoudi 1981), and *halli* and *yalli* in many Syrian and Lebanese varieties (Cowell 1964). Some dialects have more than one relative marker, such as *iladi* and *illi* in Bahrain (Holes 1983), *allaḏi* and *alli* in Ṣanʿāʾ, Yemen (Watson 1993), and *lli* and *aš* in Morocco (Harrell 1962). Unlike their counterparts in Standard Arabic, colloquial reflexes of *allaḏī* do not have morphological agreement features, as demonstrated in (7a) and (7b), where the relative marker does not agree in number or gender with the head noun. See Retsö (2004) for a survey of relative markers across Arabic dialects (→ relative pronoun).

(7a) *ʾuxt* *al-ḥarīw* *allaḏī* *kānat*
sister the-groom Rel was.3fs
tilbas *al-libs* *ar-ržžālī*
wear.3fs the-clothes the-manly
'the groom's sister who was wearing men's clothing' (Watson 1993:231)

(7b) *an-niswān* *allaḏī* *yžayn*
the-women Rel come.3p
'the women who came' (Watson 1993: 231)

The issue of whether the Arabic relative markers are relative pronouns (like English *who* and *whom*) or relative complementizers (like English *that*) reflects the diversity of these forms and their syntactic properties. For example, there are clear cases where Standard Arabic *allaḏī* and Egyptian Arabic *illi* are used as complementizers with no appropriate antecedents or embedded structures that include resumptive pronouns or missing constituents, as in (8a) and (8b). The same pattern is attested in Levantine (Mitchell and Hassan 1994; Brustad 2000) and Moroccan Arabic (Harrell 1962). However, the complementizer reading is not available in embedded contexts where verbs license sentential complements, as in (8c) and (8d).

(8a) A *qad* *suriqa* *ḥimāru-ka*
Perf stolen.3ms.Pass donkey-your
'Your donkey has been stolen'

B *al-ḥamdu* *li-llāh* *allaḏī*
the-praise to-God Comp
mā *kuntu* *ʿalay-hi*
Neg was.1s on-it
'Thank God I was not on top of it' (Ibn al-Jawzī, *ʾAxbār* 166)

(8b) _kuwayyis illi ʾulti-l-i_
good Comp told.2ms-to-me
'It is good that you told me'

(8c) *_samiʿtu allaḏī tazawwajta_
heard.1s Comp got.married.2ms
'I heard that you got married'

(8d) *_iftakart ill-o/illi huwwa_
thought.1s Comp-he/Comp he
sāfir
traveling.Act.Part
'I thought that he was traveling'

Suaieh (1980) and Galal (2004) argue that _allaḏī_ and its allomorphs are relative complementizers rather than relative pronouns. A relative pronoun copies the agreement features associated with the relativized noun phrase in the embedded clause, but the case marking of _allaḏī_ is dependent on the head noun instead. This pattern emerges clearly in sentences with dual relative markers, such as (9a), where _allaḏayni_ is marked for the accusative case in agreement with the head noun, even though the relativized phrase is the nominative subject of the embedded clause. In contrast, the relative marker _allaḏāni_ in (9b) is marked for the nominative case in agreement with the head noun, while the relativized phrase is an accusative object. This case assignment pattern indicates that _allaḏī_ and its allomorphs are not relative pronouns in embedded contexts.

(9a) _raʾaytu l-waladayni allaḏayni_
saw.1s the-boys.du.Acc Rel.du.Acc
kasarā š-šubbāk
broke.3md the-window
'I saw the two boys who broke the window'

(9b) _najaḥa ṭ-ṭālibāni_
succeeded.3ms the-students.du.Nom
allaḏāni sāʿadtu-humā
Rel.du.Nom helped.1s-cl3md
'The two students whom I helped succeeded'

True relative pronouns are grammatically used as complements of prepositions because they are nominal elements, as in _the friend to whom I wrote a letter_ vs. *_the friend to that I wrote a letter_ (McCawley 1988; Comrie 1999). The Standard Arabic relative marker _allaḏī_ and its colloquial reflexes, e.g. Egyptian Arabic _illi_, cannot be used as complements of prepositions in embedded contexts, as in (10a) and (10b). Moreover, nominal categories, such as relative pronouns, can be annexed in → construct state constructions. The ungrammaticality of annexing _allaḏī_ and _illi_, as in (11a) and (11b), demonstrates that they do not function as pronominal elements in embedded structures.

(10a) *_ṣadīq-ī ʾilā allaḏī katabtu risāla_
friend-my to Rel wrote.1s letter
'my friend to whom I wrote a letter'

(10b) *_ṣaḥbi li-lli katabt gawāb_
friend-my to-Rel wrote.1s letter
'my friend to whom I wrote a letter'

(11a) *_ar-rajul ʾumm allaḏī mātat_
the-man mother Rel died.3fs
'the man whose mother died'

(11b) *_ir-rāgil ʾumm illi mātat_
the-man mother Rel died.3fs
'the man whose mother died'

The annexation and complementation patterns of the Standard Arabic relativizers _man_ and _mā_ demonstrate that they are relative pronouns, as in (12a)-(12c).

(12a) _ʾilā man yuhimmu-hu_
to Rel concern.3ms-cl3ms
l-ʾamr
the-matter
'to whom it may concern'

(12b) _ʾaʿṭi-nī mim-mā_
give.Imper.2ms-cl1s from-Rel
ʾaʿṭāka llāh
gave.3ms God
'Give me from what God has given you!'

(12c) _hal hāḏā fiʿl man yaḥkī_
Q this doing Rel tell.3ms
ʿan-hu tilka l-ḥikāyāt
about-cl3ms these the-stories
'Is this the doing of the one who tells these stories about him?' (Ibn al-Jawzī, _ʾAxbār_ 51)

Interestingly, *allaḏī* and its colloquial reflexes are grammatical in these same contexts if used in free relative clauses, as in (13a)–(13d).

(13a) *ḥaddiṯ-nī* *bi-llaḏī*
tell.Imper.2ms-cl1s with-Rel
ʾafadta *l-yawm*
learned.2ms today
'Tell me what you have learned today!' (Tawḥīdi, *ʾImtāʿ* VI, 40)

(13b) *law* *maʿa-k-š* *ʾalam*
if with-you-Neg pen
xud *mi-lli* *ganb-ak*
take.Imper.2ms from-Rel next-you
'If you do not have a pen, take from the one who is next to you'

(13c) *lā* *tabluġu* *miqdār* *allaḏī*
Neg reach.3fs amount Rel
yuṭbax *waḥda-hu*
cook.Pass.3ms alone-cl3ms
'It does not reach the amount of what can be cooked by itself' (Jāḥiẓ, *Buxalāʾ* 24)

(13d) *ʾibn* *illi* *bāʿ* *arḍu-h*
son Rel sold.3ms land-his
'the son of the one who sold his land'

Haddad and Kenstowicz (1980) and Brustad (2000) demonstrate that the Lebanese relativizer *lli* and the Aleppan *il* can be annexed in construct state constructions in embedded contexts, as in (14a) and (14b). Evidence for this claim comes from the ungrammaticality of omitting the feminine marker *-t* from the head noun. Retsö (2004) demonstrates that this pattern is acceptable in several other dialects, including Iraqi and Bahraini.

(14a) *baṭṭ-et/*baṭṭ-a* *lli* *ʾakalnā-ha*
duck-f Rel ate.1p-cl3fs
ṭaybe
delicious
'The duck we ate is delicious' (Haddad and Kenstowicz 1980:144)

(14b) *manṭiʾ-it* *il* *ʾaxadna* *fī-ha*
area-f Rel took.1p in-it
ktīr *kwayyisa*
very nice
'The area in which we took [a house] is very nice' (Brustad 2000:101)

Choueiri (2002), Aoun and Li (2003), and Ouhalla (2004) propose that the relative marker *yalli* in Lebanese Arabic is a definite complementizer. This analysis is based on three observations. First, *yalli* can be annexed in construct state constructions as noted above. Second, *yalli* is grammatical only if the head noun is definite, indicating an (in)definiteness agreement relation. Third, the relative marker *yalli* can be bound by the universal quantifier *kəll*, as in (15).

(15) *kəll* *yalli* *baddon* *yəḥḍaro*
all Rel want.3p attend.3p
l-film *fallo*
the-movie left.3p
'All those who wanted to watch the movie left'

These patterns demonstrate the diversity of relativizers and their syntactic properties within and across dialects. For example, Standard Arabic *man* and *mā* are relative pronouns, as indicated by their grammaticality in construct state constructions and prepositional complement positions, yet *man* and *mā* are used only in free relatives. Standard Arabic *allaḏī* and Egyptian Arabic *illi* are relative complementizers, as they have complementizer functions, and they cannot be annexed or used as complements of prepositions in embedded contexts. In the absence of head nouns, *allaḏī* and *illi* have the same distribution patterns as *man* and *mā*, suggesting differences in syntactic structures across different types of relative clauses.

3. The structural properties of relative clauses

The distinction between restrictive relative clauses, where the embedded structure delimits the referential scope of the head noun, and nonrestrictive relatives is not formally encoded in Arabic (Wise 1975; Al-Bazi 1983; Drozdík 1999). Nevertheless, a nonrestrictive relative is phonologically marked with pauses that separate it from the rest of the sentence. For example, the Egyptian Arabic sentence in (16) is ambiguous between a nonrestrictive reading, where all the customers in the universe of discourse are happy, including those who returned the merchandise, and a restrictive reading, where only the customers who returned the merchandise are happy, triggering the presupposition that there are other customers.

(16) *iz-zabāyin illi raggaʿu*
the-customers Rel returned.3p
il-biḍāʿa mabsuṭīn
the-merchandise happy
'The customers, who returned the merchandise, are happy' (nonrestrictive)
'The customers who returned the merchandise are happy' (restrictive)

The discussion in the preceding section regarding the distribution patterns of relative markers indicates that there are three structurally distinct types of relative constructions in Arabic: (i) definite relatives, where an embedded clause with a relative complementizer modifies a definite noun phrase, as in (17a); (ii) indefinite relatives, where a subordinate clause lacking a relative marker is adjoined to an indefinite head noun, as in (17b); and (iii) free relatives with relativizers but no available antecedents, as in (17c).

(17a) *wajadtu l-qalam/*qalam-an allaḏī*
found.1s the-pen/*pen-Indef Rel
kuntu ʾabḥaṯu ʿan-hu
was.1s search.1s for-it
'I found the pen that I was looking for'

(17b) *wajadtu qalam-an (*allaḏī)*
found.1s pen-Indef (*Rel)
kuntu ʾabḥaṯu ʿan-hu
was.1s search.1s for-it
'I found a pen I was looking for'

(17c) *wajadtu allaḏī kuntu*
found.1s Rel was.1s
ʾabḥaṯu ʿan-hu
search.1s for-it
'I found what I was looking for'

Wright (1974) provides Classical Arabic examples where definite nouns are modified by relative clauses without relative markers, as in (18a). These cases do not necessarily obliterate the distinction between definite and indefinite relatives, because the head nouns in such instances are generic, i.e., they are not rigid designators that refer to particular entities (a specific donkey), but to their types. However, the definite/indefinite relative distinction becomes less clear in cases when indefinite nouns are modified by relatives with relativizers, as in the Kuwaiti example in (18b).

(18a) *ka-l-ḥimār yaḥmilu ʾasfār-an*
like-the-donkey carry.3ms books-Indef
'like a donkey carrying books'

(18b) *əndawwir-l-a bnayya lli*
look.for.1s-for-him girl Rel
tnāsib-l-a
suit.3fs-for-him
'We are looking for a girl for him that will suit him' (Brustad 2000:95)

Relative clauses involve a co-indexation relation between the available members of the triad: head noun, relative marker, and resumptive pronoun. In the case of definite relatives, the relative marker agrees with the head noun for case, number, and gender, if morphologically represented, while the resumptive pronoun agrees with the head noun for gender, number, and person, as in (19a), which illustrates person agreement. If the head noun corresponds to the subject of the embedded clause, the verbal predicate carries the same agreement features, as in (19b), where the modified head is a 2nd person masculine singular pronoun, and the verb in the embedded clause is marked accordingly.

(19a) *ʾana*$_i$ *allaḏī*$_i$ *tusammī-nī*$_i$
I Rel call.3fs-cl1s
n-nās aṣ-ṣāmit
the-people the-silent
'I am the one whom people call the silent one' (*ʾAlf layla* VI, 122)

(19b) *ʾanta allaḏī ṣanaʿta kitāb*
you.ms Rel made.2ms book
al-muʿallimīn
the-teachers
'You are the one who wrote the teachers' book' (Ibn al-Jawzī, *ʾAxbār* 137)

However, person agreement is sometimes violated in Standard Arabic, as in (20a) and (20b), where the pronominal head is marked differently from the verb in the embedded clause in terms of person. Schub (1991) analyzes these instances as a fourth category of relative constructions, namely cleft relatives, where the head is focalized.

(20a) *ʾanta llaḏī wahaba-hā*
you.ms Rel gave.3ms-cl3fs
l-ī al-yawm
to-cl1s the-day
'You are the one who gave it to me today' (Ibn al-Muqaffaʿ, *Kalīla wa-Dimna* 211)

(20b) *ʿalimat ʾan-nī ʾana llaḏī*
knew.3fs that-me I Rel
jaraḥa l-ʿabd
wounded.3ms the-slave
'She knew that it was I who wounded the slave' (*ʾAlf layla* VI, 33)

Relative clauses can be simple finite constructions, as in all the previous examples, but they can also be infinitival in Standard Arabic, as in (21a). Moreover, they can themselves be complex constructions, as in the Egyptian Arabic examples in (21b) and (21c), where adjunct and conditional clauses are embedded within relatives, and as in the Standard Arabic example in (21d), where the relative clause embeds a finite clause with a complementizer. Interestingly, Standard Arabic relatives can embed other relatives with their own relativizers, provided that the lowest clause is a free relative with conditional semantics, as in (21e). Finally, relative clauses can be coordinated constructions, as in (21f) and (21g).

(21a) *ittaqi l-ʾaḥmaq ʾan*
beware the-idiot Compl
taṣḥaba-hu
accompany.2ms.Subj-cl3ms
'Beware of accompanying the idiot' (Ibn al-Jawzī, *ʾAxbār* 33)

(21b) *da karīm illi lamma*
this Karim Rel when
tsallim ʿalē-h yezʿal
greet.2ms on-him get angry.3ms
'This is Karim, who, when you greet [him], gets angry'

(21c) *da z-zurār illi law*
this the-button Rel if
dust ʿal-ēh il-mumarriḍa
pressed.2ms on-it the-nurse
ḥa-tīgi
Fut-come.3fs
'This is the button which if you press [it] the nurse will come'

(21d) *ar-rajul allaḏī zaʿamat*
the-man Rel claimed.3fs
al-ʿulamāʾ ʾanna-hu ijtāza
the-scientists that-he crossed
bi-baʿḍ al-mafāwiz
by-some the-deserts
'the man whom the scientists claimed to have crossed some deserts' (Ibn al-Muqaffaʿ, *Kalīla wa-Dimna* 44)

(21e) *al-ʿaql allaḏī man ḥurima-hu*
the-mind Rel Rel deprived-it
fa-huwa ʾanqaṣ min kull faqīr
Conj-he less than every poor
lit. 'the mind which, whoever is deprived of it, is less than any poor person'
'Whoever is deprived of the mind is worse off than a poor person' (Tawḥīdī, *ʾImtāʿ* VI, 87)

(21f) *di is-sikirtēra illi il-mudīr*
this the-secretary Rel the-boss
ṭallaʾ mrāt-u
divorced.3ms wife-his
wi-itgawwiz-ha
and-married3ms-her
*'This is the secretary that the boss divorced his wife and married'

(21g) *ḏahaba l-ʾamr allaḏī*
went the-matter Rel
iḥtajta ʾila-yya wa-ḥtajtu
needed.2ms to-me and-needed.1s
ʾilay-ka fī-hi
to-you in-it
'The matter that you needed me and I needed you in is gone' (Ibn al-Muqaffaʿ, *Kalīla wa-Dimna* 182)

The internal structure of relative clauses varies significantly across Arabic dialects. For example, relativized prepositional complements in Standard, Egyptian, and Lebanese Arabic disallow pied piping, where a preposition precedes the relative marker, as in (22a) and

(22b). However, Moroccan and Algerian Arabic do allow pied piping, as in (22c) and (22d).

(22a) **il-bēt f-illi kānu saknīn*
the-house in-Rel were living
'the house in which they used to live'

(22b) **al-bayt fī llaḏī kānū*
the-house in Rel were.3mp
yaskunūna
live.3mp
'the house in which they used to live'

(22c) *s-sfina f-aš kanu rakbin*
the-ship in-Rel were.3p riding
'the ship in which they were riding' (Harrell 1962:164)

(22d) *əl-bənt mˁa mən takallamt*
the-girl with Rel talked.1s
'the girl with whom I talked' (Belkacemi 1998:148)

Resumption possibilities also vary across dialects, as Egyptian and Lebanese Arabic require resumption in all nonsubject positions (Wise 1975; Aoun and Li 2003), whereas Standard Arabic optionally allows resumption in the direct object position of definite relatives, as mentioned earlier. Syrian Arabic, on the other hand, allows direct object resumption optionally in both definite and indefinite relatives, as in (23a) and (23b) (Darrow 2003).

(23a) *dimašq hiyya l-mdina fi*
Damascus it the-city in
sūriyya yalli mā šuft-(ha)
Syria Rel Neg saw.1s-(cl3fs)
'Damascus is the city in Syria that I did not see [it]' (Darrow 2003:55)

(23b) *dimašq hiyya mdina fi*
Damascus it city in
sūriyya yalli mā šuft-(ha)
Syria Rel Neg saw.1s-(clfs)
'Damascus is a city in Syria that I did not see' (Darrow 2003:55)

4. SYNTACTIC ANALYSES OF THE RELATIVE CLAUSE

The syntax of relative clauses has been dominated by two main approaches within the Generativist Program: the raising/promotion hypothesis and the matching/adjunction analysis. Proponents of the matching analysis argue that the embedded clause is a modifying constituent adjoined to the head noun phrase, which is base-generated, i.e., it is structurally independent of the embedded clause. Moreover, the relative marker is generated in the position of the missing constituent (the trace) and moves into the clause-initial position (Chomsky 1977; Carlson 1977; Browning 1987). The raising analysis proposes that the embedded clause is a complement licensed by the determiner of the modified phrase, whereas the head noun is generated within the embedded clause and moved to the clause-initial position (Kayne 1994; Bianchi 1999; Bhatt 2002).

The main argument in support of the raising analysis is that the head noun can be interpreted as if it is in the gap position within the embedded clause. In other words, the head noun displays reconstruction effects that are reflected in idiom chunks, scope properties, and binding. For example, an idiom chunk is assumed to be generated as a unit, yet, the nominal part can be the head of a relative clause with the rest of the idiom stranded. With regard to scope properties, relative clauses allow → reflexive pronouns to be the heads while their logical antecedents are in the embedded clause. Finally, the head noun can be bound by → quantifiers in the embedded clause. The matching analysis, on the other hand, is supported by the fact that the embedded structure has a gap as indicated by sensitivity to island constraints.

Aoun and Choueiri (1997), Choueiri (2002), and Aoun and Li (2003) demonstrate that the structural patterns of Lebanese Arabic relatives do not follow the predictions of one analysis to the exclusion of the other. Rather, they propose an analysis where different derivation strategies are systematically used in different relative structures. Accordingly, definite relatives in nonisland constructions are derived by head raising, whereas definite relatives in island constructions and all indefinite relatives are base-generated. Support for this view comes from the observation that indefinite relatives do not allow reconstruction, as in (24a), where the pronoun in the head phrase cannot be bound by the quantifier in the embedded clause. Definite relatives, on the other hand, do allow reconstructions, as in (24b), provided that there is no barrier or island separating the resumptive pronoun from the head noun.

(24a) **šəft* *[ṣūra* *la-ʾəbn-a*$_i$*]*$_j$
saw.1s picture of-son-her
[kəll *mwaẓẓaf-e]*$_i$ *bədda*
every employee.f want.3fs
tˁalləʾ-a *bi-maktab-a*
hang.3fs-cl3fs in-office-her
'I saw a picture of her son every employee wants to hang in her office' (Aoun and Li 2003:128)

(24b) *ˁalēmit* *faḥṣ-o*$_i$ *yalli* *ʾarrarit*
grade.fs exam-his$_i$ Rel decided.3fs
l-mˁallme *ʾənno* *tfēžiʾ*
the-teacher.fs Comp surprise.3fs
[kəll *təlmīz]*$_i$ *fiy-a* *raḥ*
every student in-it Fut
təṭlaˁ *bukra*
come.3fs tomorrow
'The grading of his exam that the teacher decided to surprise every student with will be ready tomorrow' (Choueiri 2002: 138)

Galal (2004) develops a multiple strategy analysis for Modern Standard Arabic, arguing that the derivation of definite relatives with gaps involves movement, but not in the presence of resumptive pronouns, whereas indefinite relatives are base generated. This analysis is motivated by the observation that definite relatives with gaps are sensitive to island constraints, unlike definite relatives with resumptive pronouns, and indefinite relatives. For example, the sentences in (24a) and (24b) are grammatical even though the resumptive pronouns are within a WH-island and a complex noun phrase, i.e., these constructions are not sensitive to the island constraints. If the resumptive pronouns are missing, both sentences would be ungrammatical. Support for a base-generation analysis of indefinite relatives comes from the observation that sentences such as those in (25a) and (25b) are grammatical, even though they violate the island constraints, i.e., movement is not possible. Galal (2004) rejects head raising analyses for Modern Standard Arabic, demonstrating that reconstruction is not available, and therefore what moves cannot be the head.

(25a) *raʾaytu* *l-kitāb* *allaḏī* *saʾala*
saw.1s the-book Rel asked.3ms
ʾax-ī *ʾin* *kuntu*
brother-my if was.1s
*sa-ʾaštarī-hi/*ø*
Fut-buy.1s-cl3ms
'I saw the book that my brother asked if I was going to buy [it]' (Galal 2004:51)

(25b) *raʾaytu* *l-manzil* *llaḏī* *qābalnā*
saw.1s the-house Rel met.1p
l-muhandis *llaḏī* *ṣammama-hu/*ø*
the-architect Rel designed.3ms-cl3ms
'I saw the house of which we met the architect who designed it' (Galal 2004:52)

(26a) *raʾaytu* *kitāb* *saʾala* *ʾax-ī*
saw.1s book asked.3ms brother-my
ʾin *kuntu* *sa-ʾaštarī-hi*
if was.1s Fut-buy.1s-it
'I saw a book about which my brother asked if I were going to buy it' (Galal 2004:53)

(26b) *raʾaytu* *manzil* *qābalnā* *l-muhandis*
saw.1s house met.1p the-architect
allaḏī *ṣammama-hu*
Rel designed.3ms-cl3ms
'I saw a house of which we met the architect who designed it' (Galal 2004: 54)

Although multiple-strategy analyses agree that indefinite relatives are base-generated and that definite relatives involve movement, the nature of the moved element is still controversial. For example, Choueiri (2002) argues that the raised constituent is the head noun in Lebanese Arabic, while Galal (2004) adopts the view that a null relative operator is what moves. The basic assumption regarding the nature of the moved element is that if reconstruction is possible, then it is the head noun phrase that moves; otherwise, it is a relative operator. How-ever, Choueiri (2002), Galal (2004), and Darrow (2003) agree that reconstruction is quite problematic because it is subject to pragmatic factors.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

ʾAlf layla = *ʾAlf layla wa-layla*. Beirut: Dār al-Fikr al-ˁArabī, 1997.

Ibn ʿAqīl, *Šarḥ* = Bahāʾ ad-Dīn ʿAbdallāh Ibn ʿAqīl, *Šarḥ ʿalā ʾAlfiyyat Ibn Mālik*. Ed. Qāsim ar-Rifāʿi. Beirut: Dār al-Qalam, 1987.

Ibn Hišām, *Qaṭr an-nadā* = Jamāl ad-Dīn ʿAbdallāh ibn Yūsuf Ibn Hišām al-ʾAnṣārī, *Šarḥ Qaṭr an-nadā wa-ball aṣ-ṣadā*. Ed. H. Al-Fāxūrī. Beirut: Dār al-Jīl, 1996.

Ibn al-Jawzī, *ʾAxbār* = ʾAbū l-Faraj ʿAbd ar-Raḥmān ibn ʿAlī Ibn al-Jawzī, *ʾAxbār al-ḥumqā wa-l-muġaffalīn*. Ed. Muḥammad ʾAmīn Faršūx. Beirut: Dār al-Fikr al-ʿArabī, 1990.

Ibn al-Muqaffaʿ, *Kalīla wa-Dimna* = ʾAbū Muḥammad (ʾAbū ʿAmr) ʿAbdallāh Ibn al-Muqaffaʿ, *Kalīla wa-Dimna*. Ed. Muḥammad ʾAmīn Faršūx. Beirut: Dār al-Fikr al-ʿArabī, 1990.

Jāḥiẓ, *Buxalāʾ* = ʾAbū ʿUṯmān ʿAmr ibn Baḥr al-Jāḥiẓ, *Kitāb al-buxalāʾ*. Beirut: Dār Ṣādir, 2000.

Tawḥīdi, *ʾImtāʿ* = ʾAbū Ḥayyān ʿAlī ibn Muḥammad at-Tawḥīdī, *Kitāb al-ʾimtāʿ wa-l-muʾānasa*. Ed. ʾAḥmad ʾAmīn and ʾAḥmad az-Zayn. Cairo: Dār Maktabat al-Ḥayāt, 2004.

SECONDARY SOURCES

ʿAbd al-Ġanīy, ʾAyman. 2002. *an-Naḥw al-kāfī*, I. Cairo: Dār Ibn Xaldūn.

Al-Bazi, Mohammed. 1983. *Scopal properties of the Arabic relative clause*. Ph.D. diss., University of Washington.

Aoun, Joseph and Lina Choueiri. 1997. *Resumption and last resort*. Ms., University of Southern California, Los Angeles.

——. and Yen-hui Li. 2003. *Essays on the representational and derivational nature of grammar: The diversity of Wh-constructions*. Cambridge, Mass.: MIT Press.

Belkacemi, Manchester. 1998. "The relative marker: A definite marker substitute?" *Archív Orientální* 66.142–148.

Bhatt, Rajesh. 2002. "The raising analysis of relative clauses: Evidence from adjectival modification". *Natural Language Semantics* 10.43–90.

Bianchi, Valentina. 1999. *Consequences of antisymmetry: Headed relative clauses*. Berlin and New York: Mouton de Gruyter.

Browning, Marguerite. 1987. *Null operator constructions*. Ph.D. diss., Massachusetts Institute of Technology.

Brustad, Kristen. 2000. *The syntax of spoken Arabic: A comparative study of Moroccan, Egyptian, Syrian, and Kuwaiti dialects*. Washington, D.C.: Georgetown University Press.

Carlson, Greg. 1977. "Amount relatives". *Language* 53.520–542.

Chomsky, Noam. 1977. "On wh-movement". *Formal syntax*, ed. Peter Culicover, Thomas Wasow, and Adrian Akmajian, 71–132. New York: Academic Press.

Choueiri, Lina. 2002. *Issues in the syntax of resumption: Restrictive relatives in Lebanese Arabic*. Ph.D. diss., University of Southern California.

Comrie, Bernard. 1999. "Relative clauses: Structure and typology on the periphery of Standard English". *The clause in English: In honor of Rodney Huddleston*, ed. Peter Collins and David Lee, 81–92. Amsterdam and Philadelphia: J. Benjamins.

Cowell, Mark. 1964. *A reference grammar of Syrian Arabic*. Washington, D.C.: Georgetown University Press.

Darrow, James. 2003. "Relative clauses in Syrian Arabic: Two reconstruction problems". *Perspectives on Arabic linguistics*, XV, ed. Dilworth Parkinson and Samira Farwaneh, 53–83. Amsterdam and Philadelphia: J. Benjamins.

Drozdík, Ladislav. 1999. "Some aspects of relativization in Egyptian Colloquial Arabic". *Asian and African Studies* 8.74–82.

Fayyāḍ, Sulaymān. 1995. *an-Naḥw al-ʿaṣrī*. Cairo: Markaz al-ʾAhrām li-t-Tarjama wa-n-Našr.

Feghali, Habaka. 2004. *Gulf Arabic: The dialects of Riyadh and eastern Saudi Arabia*. Virginia: Dunwoody Press.

Galal, Mohammad. 2004. *A Minimalist approach to relative clauses in Modern Standard Arabic*. Ph.D. diss., University of Kansas.

Haddad, Ghassan and Michael Kenstowicz. 1980. "A note of the parallels between the definite article and the relative clause marker in Arabic". *Studies in the Linguistic Sciences* 10.141–149.

Harrell, Richard. 1962. *A short reference grammar of Moroccan Arabic*. Washington, D.C.: Georgetown University Press.

Ḥassān, ʿAbbās. 1974. *an-Naḥw al-wāfī*, I. Cairo: Dār al-Maʿārif.

Hoffiz, Benjamin. 1995. *Morphology of U.A.E. Arabic Dubai dialect*. Ph.D. diss., University of Arizona.

Holes, Clive. 1983. "Bahraini dialects: Sectarian differences exemplified through texts". *Zeitschrift für Arabische Linguistik* 13.26–67.

Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, Mass.: MIT Press.

McCarthy, Richard and Faraj Raffouli. 1965. *Spoken Arabic of Baghdad*. Beirut: Librairie Orientale.

McCawley, James. 1988. *The syntactic phenomena of English*, II. Chicago: University of Chicago Press.

Mitchell, T.F. and Shahir El-Hassan. 1994. *Modality, mood and aspect in Spoken Arabic with special reference to Egypt and the Levant*. New York: Kegan Paul International.

Ouhalla, Jamal. 2004. "Semitic relatives". *Linguistic Inquiry* 35.288–300.

Retsö, Jan. 2004. "Relative-clause marking in Arabic dialects: A preliminary survey". *Approaches to Arabic dialects: A collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday*, ed. Martine Haak, Rudolf de Jong, and Kees Versteegh, 263–273. Leiden: E.J. Brill.

Schub, Michael B. 1991. "Clefted direct relative clauses". *Zeitschrift für Arabische Linguistik* 23.105–106.

Sieny, Mahmoud. 1978. *The syntax of urban Ḥijāzī Arabic*. Beirut: Librairie du Liban.

Suaieh, Saadun. 1980. *Aspects of Arabic relative clauses: A study of the structure of relative clauses in Modern Written Arabic*. Ph.D. diss., University of Indiana.

Talmoudi, Fathi. 1981. *Texts in the Arabic dialect of Sūsa (Tunisia)*. Göteborg: Acta Universitatis Gothoburgensis.

Watson, Janet. 1993. *A syntax of Ṣanʿānī Arabic*. Wiesbaden: O. Harrassowitz.

Wise, Hilary. 1975. *A transformational grammar of Spoken Egyptian Arabic*. Oxford: Basil Blackwell.

Wright, William. 1967. *A grammar of the Arabic language*. 3rd ed. Repr. Cambridge: Cambridge University Press.

MUSTAFA MUGHAZY
(Western Michigan University)

Relative Pronoun (Arabic Dialects)

The relative pronoun in Arabic dialects is expressed by a form that is invariable in gender and number and may refer either to persons or objects. The most common form used in the majority of Arabic dialects is *(i)lli*, with variants *halli* and *yalli* (Retsö 2004:264–265). This form is found in the Egyptian dialects, the Levantine dialects, the *gilit* dialects from Iraq, most of the Arabian Peninsula, and the majority of the Maghrebi dialects. Examples are *əlli mā ḥməd qalīl mā yəḥməd kaṯīr* 'he who is not grateful for a little is not grateful for a lot' (Ḥassāniyya Arabic; Ould Mohamed Baba 2008:325); *ilbint illi gat* 'the girl who came' (Cairo Arabic; Woidich 2006:51).

According to David Cohen (1962:140), the explanation for this homogeneity could be the greater prevalence in Old Arabic of a form *illi*, whereas the form finally adopted and standardized in Classical Arabic, *allaḏī*, was less used in previous periods of the Arabic language. Both forms originally had a demonstrative function (Grand'Henry 1972:142).

There are other forms of the relative pronoun as well, such as *il* with variants *al, la, lə*, and *lē*, which mainly differs from the aforementioned forms by the lack of gemination in the phoneme *l* (about the origin of this gemination, see Brockelmann 1910:123). These are found in particular in *qəltu* dialects of northern Iraq and in Anatolia. In the latter region, the relative pronoun is *la* in Mardin and *lə-* in most dialects of the Mardin group, while in the dialects of Diyarbakır, Siirt, and Kozluk, the form *lē* is found. *Il* occurs in Central Asia, in some zones of the Syrian desert, in North Africa, and among the Shiʿites of Bahrain. However, the most frequent form of the relative in most of these cases is currently *(i)lli*, e.g. *santēn il dōxel fat walad kun ʿendu* 'he had a son who was two years old' (Central Asian Arabic; Zaborski 2008:419); *zalamət la aṛayna fə-d-dōlmūš mən mērdīn we* 'the man whom we saw in the minibus is from Mardin' (Mardin Arabic; Grigore 2007:237).

On some occasions, these relatives are phonetically assimilated to the following word, in the same way as the definite article in Arabic, e.g. *az-zlima aš šiftu* 'the man whom I saw' (Syro-Mesopotamian Bedouin dialects; Behnstedt 2008a:79); *sint iž-žāyi* 'the next year' (Bišmizzīn Arabic; Jiha 1964:172).

A much more limited group of dialects uses another particle with the same function, *ddi* and its allomorphs *d(i)-* and *da-*, which are mainly found in the most ancient Maghrebi dialects: rural dialects, for instance those spoken in the region of Jbala in northern Morocco, where it is typical of female speech (see Moscoso 2003:168), and dialects of the old medinas and the Jewish communities, for instance those spoken in Sefrou and Debdou, both in Morocco (see Heath 2002:494–495). Examples are *ər-rāžəl əddi ža* 'the man who came' (Djidjelli Arabic; P. Marçais 1956:493), *d ʿandu u d ma ʿandu ši* 'who has and who has not' (Moscoso 2003:169). This relative may co-occur with personal pronouns, which is typical of rural Maghrebi dialects, e.g. *dənnūma, dənni, dənnu* (Vicente 2000:142; Moscoso 2003:169). All of these forms are gradually being replaced by *(i)lli*.

In the eastern region, according to Rabin (1951:39), the form *ddi* used to exist in North Arabian dialects of the Ḥijāz, where it may be regarded as a trace of South Arabian influence since it is the current form in some areas of Yemen. Thus, even nowadays, in the southeast and the northernmost part of the country, the relative *ḏī* is typical, e.g. *im-marit ḏī kān maʿnā* 'the woman who was with us' (Behnstedt 2008b:116).

When the → relative clause is introduced by the two aforementioned groups of particles, *il* and *di*, it can sometimes be translated as a genitive construction, e.g. *ḥaṭab il mebīʿ* 'the seller of firewood' (Central Asian Arabic; Zaborski 2008:429); *əl-ktāb əddi xāy* 'my brother's book' (Tlemcen Arabic; W. Marçais 1902:272). This coincidence between relative and genitive forms is common in many of the oldest Arabic dialects, such as those spoken by Jews in the Maghrebi area, or the *qəltu* dialects of Mesopotamia and Anatolia. According to Retsö (2004:269–270), this cannot be put down to chance, since there are parallels in other Semitic languages.

In addition to the above-mentioned particles, another form of relative must have existed in Old Arabic which has survived in Neo-Arabic, the relative *allaḏī*. This form cannot be regarded as a loanword from Classical Arabic but must be the survival of an ancient form. This is the case of Andalusi dialect, whose standard variant had the invariable relative form *allaḏí*, with the allomorphs *allí, addí*, and *aḏḏí* in a low register, e.g. *arrijál allaḏi ḥamálu alqámḥ* 'the men who carried the wheat', *addí lak* 'the one that you have' (Corriente 2008:365). The same particle exists in contemporary Yemeni Arab, where *ʾallaḏī* is the most usual relative and it is used as an invariable form, along with the abbreviated forms *ʾallī, lī* (Behnstedt 2008b:116).

As regards its syntactic function, the relative pronoun is used in relative clauses that refer to a nominal antecedent, provided the antecedent is determined; in this case the relative clause has a qualifying function. When the relative pronoun is not the subject of the clause but is instead the direct or indirect object of a verb or a complement of the noun, a resumptive pronoun (→ resumption) is used, which agrees with it in gender and number and which expresses the function performed by the relative pronoun, e.g. *ikkitāb illi štarētu* 'the book that I bought' (Cairo Arabic; Woidich 2006:51); *ʿṭēni əs-stīlo lli ktəbt bīh əl-bra* 'give me the pen with which I wrote the letter' (Moroccan Arabic; Caubet 1993:I, 175).

There are other syntactic constructions where an independent pronoun appears after the relative particle; this is the case described for Egyptian Arabic by Mejdell (2003), who mentions two functions: the independent pronoun introduces a subject that is not coreferent with the antecedent, and it may function as a marker of nonrestrictedness, e.g. *il 'king' walla malik 'Harold' illi huwwa kān malik briṭānya fil-waʾti dā* 'the king or King Harold, who was king of Britain at that time' (Mejdell 2003:543).

In cases where the antecedent is an indefinite noun (a noun that is not determined by means of an article, by a construct state, or by nature), the relative pronoun does not appear and the relative clause is simply juxtaposed to the main sentence, e.g. *žāb līh ṭūmūbīl kānət xāsra* 'he brought him a broken car' (Skūra Arabic; Aguadé 1995:127). In other cases, when there is no overt antecedent, it clearly has a pronominal function, e.g. *lli fāt, māt* 'the past is the past' (Moroccan Arabic; Caubet 1993:I, 175), or the Egyptian version *lli fāt, fāt* (Cairo Arabic; Woidich 2006:51).

It is also possible to form the relative with the particles *ma* and *mīn*, which are less often used than the others; *ma* functions as a relative with included antecedent, e.g. *ma xalla li ma nākul* 'he did not leave for me something to eat' (Moroccan Arabic; Caubet 1993:I, 175); *ʾultilak ma fīh ilkifāya* 'I told you enough' (Cairo Arabic; Woidich 2006:201). A similar particle was also used in Andalusi Arabic, e.g. *mará man tukún qaríbatak* 'a woman who is your relative' (Corriente 2008:365). *Mīn* is found in Egyptian Arabic, e.g. *amma ʾana fa ʾalfi mīn yitmannāni* 'as for me, there are a thousand who would like to have me' (Cairo Arabic; Woidich 2006:201).

In several Arabic dialects, there are also compound relative forms, for example those made up by a preposition and the particle *-āš*, typical of Maghrebi dialects, e.g. *ṛāžəl lāš gult hād əš-ši* 'the man to whom you said this' (Skūra Arabic; Aguadé 1995:128). In this case, the antecedent can be a person or an object, and it can either be definite or nondefinite. It has the syntactic function of an adverbial complement, and because of this it can replace the *lli* + resumptive pronoun construction, e.g. *ha əd-dāṛ fāš kansəknu* 'this is the house in which we live' (Moroccan Arabic; Caubet 1993:I, 177), which could be rephrased as *ha əd-dāṛ lli kansəknu fīha*.

Another compound form is *mən*, which can only refer to persons and is accompanied by a preposition, e.g. *ši fqīh mʿa mən ṣāfəṛ* 'a lawyer with whom he traveled' (Skūra Arabic; Aguadé 1995:127).

BIBLIOGRAPHICAL REFERENCES

Aguadé, Jordi. 1995. *El dialecto árabe de Skūra (Marruecos)*. Madrid: Consejo Superior de Investigaciones Científicas.
Behnstedt, Peter. 2008a. "Árabe beduino (Península Arábiga y Oriente Próximo". Corriente and Vicente (2008:69–95).
——. 2008b. "Árabe yemení". Corriente and Vicente (2008:97–149).
Brockelmann, Carl. 1910. *Précis de linguistique sémitique*. Paris: Geuthner.
Caubet, Dominique. 1993. *L'arabe marocain*. 2 vols. Paris and Louvain: Peeters.
Cohen, David. 1962. "Koinè, langues communes et dialectes arabes". *Arabica* 9.119–144.

Corriente, Federico. 2008. "Árabe andalusí". Corriente and Vicente (2008:353–378).
——. and Ángeles Vicente (eds.). 2008. *Manual de dialectología neoárabe*. Saragossa: Instituto de Estudios Islámicos y del Oriente Próximo.
Grand'Henry, Jacques. 1972. *Le parler arabe de Cherchell (Algérie)*. Louvain-la-Neuve: Université Catholique de Louvain.
Grigore, George. 2007. *L'arabe parlé à Mardin*. Bucharest: Editura Universităţii din Bucureşti.
Heath, Jeffrey. 2002. *Jewish and Muslim dialects of Moroccan Arabic*. New York: Routledge Curzon.
Jiha, Michel. 1964. *Der arabische Dialekt von Bišmizzīn: Volkstümliche Texte aus einem libanesischen Dorf mit Grundzügen der Laut- und Formenlehre*. Beirut: Orient-Institut.
Marçais, Philippe. 1956. *Le parler arabe de Djidjelli (Nord constantinois Algérie)*. Paris: Maisonneuve.
Marçais, William. 1902. *Le dialecte arabe parlé à Tlemcen: Grammaire, textes et glossaire*. Paris: Leroux.
Mejdell, Gunvor. 2003. "Notes on (non-)restrictedness in relative clauses in Egyptian Arabic". *AIDA 5th Conference Proceedings*, ed. Jordi Aguadé and Ignacio Ferrando, 539–544. Cadiz: Universidad de Cádiz.
Moscoso, Francisco. 2003. El *dialecto árabe de Chauen (N. de Marruecos)*. Cadiz: Universidad de Cádiz.
Ould Mohamed Baba, Ahmed-Salem. 2008. "Árabe ḥassāniyya de Mauritania". Corriente and Vicente (2008:311–351).
Rabin, Chaim. 1951. *Ancient West-Arabian*. London: Taylor's Foreign Press.
Retsö, Jan. 2004. "Relative-clause marking in Arabic dialects: A preliminary survey". *Approaches to Arabic dialects: A collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday*, ed. Martine Haak, Rudolf de Jong, and Kees Versteegh, 263–273. Leiden: E.J. Brill.
Vicente, Ángeles. 2000. *El dialecto árabe de Anjra (Norte de Marruecos)*. Saragossa: Universidad de Zaragoza.
Woidich, Manfred. 2006. *Das Kairenisch-Arabische*. Wiesbaden: O. Harrassowitz.
Zaborski, Andrzej. 2008. "Árabe de Asia Central". Corriente and Vicente (2008:409–437).

Ángeles Vicente (University of Zaragoza)

Religion and Language

1. Introduction

It is difficult to analyze a form of language in which the topics that are dealt with are admittedly beyond ordinary experience. Religious language is generally regarded even by its most committed users as inadequate. The dispute on whether religious utterances are imbued with a deeper and fuller kind of meaning or whether they are basically meaningless has not come to an end (for general overviews of philosophical, theological, and linguistic approaches, see Kaempfert 1983; Stiver 1996). Despite the well-known ambiguities which make religious discourse an 'odd' language (I.T. Ramsay), it is generally admitted that it has been a crucial and productive sector within the development of many languages, and very often has provided linguistic models and patterns of literary expression that remained highly influential for centuries (Stark 1987; Samarin 1987; Moser 1964). This is particularly obvious for Judaism, Christianity, and Islam. A comparison between these three scriptural religions shows certain common aspects of their linguistic influence on their adherents. The emergence and conservation of linguistic registers that are regarded as sacred; frequent multilingualism; and an institutionalized transfer of religious texts and doctrines beyond linguistic borders can be found in all of them throughout their history. Sacred language was sometimes transformed into a medium of general cultural expression, or was even reconstructed as a national language (e.g. in the creation of Ivrit from Hebrew). Sacred texts and laws influenced the development of legal and political terminology. Quite often they became important symbols of social and political unity. Religious teaching and preaching also provided a framework for public discourse and became instrumental in the emergence of civic societies. The 20th century provided many examples of such developments among Jews, Christians, and Muslims in different parts of the world. The promising study of religious language, which combines approaches from general religious studies, philosophy, and sociology with those from different linguistic and literary disciplines, was only recently laid out in a first encyclopedic work (Sawyer and Simpson 2001).

2. Arabic as a sacred language

Arabic has arguably provided one of the most powerful examples of a sacred language in history. In the case of Judaism, Hebrew was kept as the language of scripture and of religious and legal literature, and it remained a stabilizing factor for religious and communal life of the different Jewish diasporas. But multilingualism was also essential for most Jewish groups throughout their history. Christianity, on the other hand, saw frequent changes in its sacred languages and language patterns. The New

Testament was first laid down in Greek, the cultural language and also the lingua franca of the Roman Empire, and the process of translation of the Christian scriptures already started in the course of the 2nd century with Latin, followed by Syriac and Coptic, and later by other languages. The case of Islam was again different. Perhaps no other linguistic culture ever blended the language of God Himself with that of human thought and culture on a comparable scale. Arabic combined the sacred character of Hebrew among the Jews with the wider scientific and literary functions reserved to Greek and Latin in the Christian world. → Pre-Islamic Arabic had already been very rich in oratory and poetic forms. The scriptural revelation gave an immense stimulus to writing and learning. The public style of political speech that developed within the caliphate was to a large extent derived from religious forms of preaching and exhortation. After the Umayyads had made Arabic the administrative language of their empire, a class of scribes emerged who developed Arabic into a refined literary tool, merging the religious and political with the old poetical dictions. From the late Umayyad period onward, notions and linguistic patterns derived from Greek, Indian, and Persian sources of philosophy, science, poetics, and literary prose were translated and absorbed into Arabic literary expression. After the decline of the caliphate and its political fragmentation, this language became one of the main factors of cultural unity within the still-expanding Islamic world, a unity which included a growing majority of non-native speakers of Arabic.

The linguistic merger of the divine and human spheres had ambiguous consequences for the Arabic language itself. One was its sacralization. Among Muslims, it was the "verbalization of the Sacred" (Habermas), which in the Islamic context came to prevail over its visual representation by hallowed images and symbols. This also frequently made it an expression of social sacrality, i.e. of the normative consensus prevailing within a given community, constantly renewed by ritual or linguistic performance (Durkheim). But the overwhelming presence of sacred registers and formulas in Arabic often also had the opposite effect, leading to their extension to nonreligious linguistic functions and to quite ordinary expressions of everyday life. Finally, sacred language became a crucial element of moral communication and political discourse, serving both as a prop to the exercise of established power and as an avenue of empowerment for aspiring leaders and groups.

3. INFLUENCE OF THE ISLAMIC CORPUS AND ITS IMPACT ON ARABIC LITERARY CULTURE

Islam is present in Arabic literary culture with an immense corpus of religious, theological, legal, historical, and poetical texts. Apart from the *Qurʾān, Ḥadīṯ*, and related commentaries, this includes a considerable variety of other literary genres (e.g. biography and historiography, and also mystical poetry and prose). With the spread of literacy and school education, these Islamic texts have now become widely accessible to an ever-increasing number of people, and their impact can often be noticed well outside the established religious institutions.

Apart from this textual presence, the historical significance of Islam for the development of the Arabic language went far beyond the religious domain as such. The sacralization of Arabic brought about by the Qurʾānic text had manifold consequences. It enhanced the development of Arabic grammar and lexicography, which, apart from their dealing with pre-Islamic poetry, also received a major impulse from Qurʾānic exegesis (→ grammatical tradition). Sacralization also came to be extended to pre-Islamic poetry itself. Despite its largely irreligious and this-worldly character, the linguistic and literary heritage of the pre-Islamic Arabs came to be regarded as a prime reference for the language of revelation. It was included in the sacralized canon of Arabic and came to be learned and studied even within highly religious institutions of learning. This created a humanist strand in Arabic literary culture, which was in a way comparable to the persisting importance of the pagan classics in medieval Christianity.

4. SACRED LANGUAGE AND POLITICAL RHETORIC

The pre-Islamic rhymed prose (→ *sajʿ*), used extensively in the *Qurʾān*, was adopted and greatly enhanced by Islam as a ceremonial form of public sermon and speech. It became a standard element of the Friday sermon and has even extended its use to extrareligious ceremonial functions in many parts of the Arab world. The exemplary Friday sermons recorded for the early caliphs and governors (for instance in al-Jāḥiẓ's *Bayān*) show a seamless interlacement of religious and political rhetoric.

The rise of written Arabic as a language of administration and general culture that occurred under the Umayyad and Abbasid rule further linked the religious and the literary patterns. With translations from Greek, Syriac, Persian, and Indian sources increasing from the 2nd century A.H. onward, the literary domain of Arabic was extended widely, and it developed into one of the leading scientific languages of the Middle Ages. The combination of religious and nonreligious ideals of expression and learning among the scribes and literary figures can be seen already in the works of the last head of the Umayyad chancery, the *mawlā* ʿAbd al-Ḥamīd al-Kātib (d. 750; al-Qāḍī 1993). His *rasāʾil* show an elegant and flexible Arabic style, based on the rhetoric of the *Qurʾān* and the *xuṭba* and blended with Hellenist and Persian influences. His use of notions and quotations from the *Qurʾān* shows particular virtuosity. His letters also provide many examples of the deliberate ideological use of the *Qurʾān* in the service of political power. Its religious terms and phrases are marshalled without scruples against the enemies of the Umayyads. Positive Qurʾānic terms like 'party of God' (*ḥizb Allāh*), 'friends of God' (*ʾawliyāʾ Allāh*), and 'those who trust [the word of God]' (*muṣaddiqūn*) are reserved for the Umayyads and their supporters. Their enemies are accused of 'sedition' (*fitna*), 'deviation' (*ḍalāla*), and 'rebellion' (*maʿṣiya*). They figure as 'party of the devil' (*ḥizb aš-šayṭān*) and as 'people who distort [religion]' (*muḥarrifūn*) and who will have to face 'God's wrath' (*ġaḍab Allāh*). Political and religious hostility are kept convertible in this discourse, a rhetorical strategy which can be observed throughout Islamic history and into the present in the articulation of social and political conflict, often by both parties concerned.

5. The Names of God: Fusion of religious symbols with social norms

The verbalization of the Sacred, which is so characteristic for Arabic, manifests itself with particular clarity in the Names of God and their use in prayer, sermon, and everyday language. Large parts of the *Qurʾān* can be understood at two levels. The narrative, admonitory, or legislative texts are constantly interspersed with statements on God's attributes and His ways of dealing with mankind and the cosmos. These statements comment upon and often conclude the narrative, placing the events and phenomena of the temporal world in the light of God's eternal will and sovereignty. As elementary components of Islamic theological discourse, the Names of God are also applied by the pious to the manifold situations of their lives, as parts of personal names, as praying formulas for specific occasions, and even as protective amulets.

The Names of God and their use also aptly illustrate the fusion of religious symbolism with general social norms, which is also quite common in Arabic. This is documented by Richard Antoun (1989:106–125) in his analysis of a sermon by a Jordanian preacher in his village. Calling for mutual solidarity and support in a diverse community of numerous clans which are only distantly related, the preacher uses the maternal symbol of the womb (*raḥim*) to stress the kinship of all. He then dwells upon the close resemblance between this word and the notion of mercy (*raḥma*), a key virtue of Islam, which is then linked to God's own name 'the Compassionate' (*ar-Raḥmān*). He thus joins three key terms sharing derivation from the same root (*r-ḥ-m*), which serve to fuse social, ethical, and religious notions and values: "Kinship (*raḥim*) is one of the indicators of mercy (*raḥma*), and it is derived from the name 'the Compassionate' (*ar-Raḥmān*)."

The close connection between Mercy, kinship, and God Himself is further illustrated by a *Ḥadīṯ*: "I heard the Messenger say, 'God said, to Him belongs glory and power: I am God; I am the Compassionate (*ar-Raḥmān*); I created the womb (*raḥim*) and [...] I split off to it one of my names. Whoever keeps to it, I keep to him and whoever cuts it off, I cut him off."

A merger between the human and divine spheres is thus attempted by this sermon: Whoever neglects mercy not only isolates himself from his fellow men but also from God and His Mercy.

6. Islamic phraseology in everyday life

Apart from use of the Names of God in religious sermons, written as well as spoken Arabic is generally interspersed with religious phrases which have often become conventional elements of general use. They also decorate the walls in private and public buildings, often in beautiful calligraphic designs. Religious phraseology is especially common for expressions of salutation and farewell, of apology, polite request, and gratitude. The rich material collected by Piamenta (1979) from Arabic dialects shows a wide range of situational functions of religious formulae in everyday Arabic speech. In reaction to negative as well as positive experience, to uncertainty and disappointment, but also to success and positive surprise, the resort to God and His Will is offered as the only source of protection and security:

> *yā ʿāliman bi-ḥālī ʿalayka ttikālī* ‘O Knower of my condition! In Thee I put my trust!’ (Piamenta 1979:27); *twakkal ʿalā ḷḷāh* ‘put your trust in God!’ (Piamenta 1979:27–30); *yā sātir* ‘O Protector!’ (Piamenta 1979:101–109); *yā laṭīf* ‘O Kind One!, O my God!’ (Piamenta 1979:10, 78, 107); *yā salām* ‘wonderful!, wow!, sorry!, oh dear!, goodness!’ (Piamenta 1979:58–64); *mā šāʾa ḷḷāh* ‘wonderful!, my goodness!’ (Piamenta 1979:199–202); *ʾaʿūzu billāh* ‘God forbid!’ (Piamenta 1979:100); *aḷḷāhu ʾaʿlam* ‘God knows best! [i.e. ‘I do not know for sure’]’ (Piamenta 1979:185ff.).

Oaths like *waḷḷa/waḷḷāh/waḷḷāhi* ‘by God!’ (Piamenta 1979:41ff.) are extremely common, very often found in utterances without any sizable religious content, and even as a simple preliminary to further speech, like English *well*,.... The reference to God’s Will with *ʾin šāʾa ḷḷāh* ‘if God wills’ (Piamenta 1979:203–220) accompanies any talk about one’s future, hopes, and intentions, and also expresses blessings, good wishes, and warnings. The speaker thus places his statement beyond his own personal sphere. Many speech situations require expressions of blessing and gratitude toward God (*al-ḥamdu li-llāh*; Piamenta 1979:77–80). Reproach and apology are generally expressed by asking for God’s forgiveness, which allows for saving one’s own face or that of one’s partner: *ʾastaġfiru ḷḷāh* ‘I ask for God’s forgiveness!’ (Piamenta 1979:135–140), *aḷḷāh ysāmḥak* ‘may God forgive you!’ (Piamenta 1979:143ff.).

Non-Muslim speakers of Arabic are also fully involved in this daily exchange of Islamic religious phrases which have become general expressions of politeness and respect and are sometimes used with a good deal of irony: *subḥāni l-ʿālim* ‘praise to God the Knower!, nobody knows!’ (Piamenta 1979:185). This heritage was also transmitted to other languages influenced by Islam. A trace of such a phrase might even be found in Spanish, where *ojalá* ‘I hope so, I wish’ seems to go back to *ʾin šāʾa ḷḷāh*, or perhaps rather to *wa-šāʾa ḷḷāh* ‘would God that...!’. The Arabic origin of *olé* ‘bravo!’, which is also often traced to *allāh*, seems to be more doubtful (cf. Corominas and Pascual 1981:IV, 268ff., s.v. *ojalá*, 278ff. s.v. *olé*; → Ibero-Romance).

7. Religious phraseology and Sufi literature

To a remarkable extent, religious phraseology also bridges the gap between Literary Arabic and the dialects. This can be shown in particular for religious and Sufi literature and poetry, which are often adapted and addressed to people of quite diverse educational backgrounds. Like the Arabic popular epic, Sufi literature often oscillates between a Literary Arabic interspersed with dialectal expression and dialectal texts with strong admixtures of literary religious terms (cf. the Egyptian popular religious poems edited and translated by Littmann 1950, 1951). The large biographical collection of the lives and legends of Sudanese scholars and saints which was written around 1805 (Ibn Ḍayfallāh’s *Ṭabaqāt*) regularly switches to Sudanese dialect when it reports oral speech. This device places the described personalities and their extraordinary feats in a very lively atmosphere. The strong separation of dialectal and literary levels cannot be maintained where the transmission of personal religious experience and emotion is at stake. In some cases, reports with strong dialectal elements were later purged and adapted to higher literary standards, changing the character of the story “from eye-witness report to stylised *karāma*” (Hofheinz 1990:22ff.).

The vocabulary and terminology of classical Islamic mysticism was based on the *Qurʾān*, *Ḥadīṯ*, and theology, but many notions from disciplines like grammar, astronomy, and medicine were also adapted and appropriated (Massignon 1954:46–52). Mystical exegesis moves from the exterior to the interior sense of the Qurʾānic text and further develops its own symbolic language of spiritual experience (Nwyia

1970). As in many other literary languages (see for German, Langen 1968), mystical language and terminology seem to have exerted considerable influence on the development of psychological, affective, and aesthetic expression in Arabic language and literature, leading in particular to an increased derivation of abstract nouns. Research in this field, however, has hardly begun.

8. Islamic terms and their political secularization in the 20th century

The political movements which have emerged in the Arab world since the 19th century have made strong use of the Islamic linguistic heritage (Rebhan 1986; Ayalon 1987). Although largely pursuing secularist aims and objectives, Arab nationalists have often resorted to religious terms and phrases in order to secure general acceptance and to enhance the dignity of their struggle (→ nationalism and language). This has led to a transfer of religious terms to other, mainly political, contexts, and to their generalization and secularization, generating a parareligious political language which can be compared to similar movements and political ideologies in other countries (Nietzsche provided important models for this shift; see Kaempfert 1971). Secularization of religious terms like *ʾumma* 'nation', *risāla* 'message', or *baʿṯ* 'resurrection > renaissance' can be observed in an exemplary way in the ideological language of the founders of the Baʿṯ party in Syria and Iraq and in the official Arab socialism of Jamāl ʿAbd an-Nāṣir in Egypt (for the former, see Salem Babikian 1977 and Glaß 1985; for the latter, see Rejwan 1974 and Kassian 1991).

8.1 *Michel ʿAflaq and the Baʿṯ party*

In the writings of Michel ʿAflaq (d. 1989; see Salem Babikian 1977), the founder of the Baʿṯ ideology, linguistic secularization results from his dialectical model of creative tension, which was heavily influenced by the French writer André Gide. Tension is built up and maintained between opposing tendencies that both confront and confirm each other. With respect to the Arabs, this tension is maintained between the eternal message and truth of the Arab essence (*ʿurūba*) and its temporal realizations, which oscillate between bloom and decay. Islam was for ʿAflaq the historical culmination of the 'Arab idea' as it had developed thus far, and he refuses to draw a clear-cut distinction between Islam and *ʿurūba*. But it is an ideal Islam, not its present and, in his view, quite miserable shape, on which ʿAflaq grounds his ideals of freedom, socialism, and national unity. An ethical and political revolution as a mobilizing force will bring the true *ʿurūba* back to its full realization. This will be the *baʿṯ* 'resurrection' that gave the party its name. Its slogan enshrines the sacred character of the Arab nation: *ʾumma ʿarabiyya wāḥida – ḏāt risāla xālida* 'One Arab Nation – bearer of an Eternal Message'. The Arab language holds a central place in Baʿṯ ideology. Al-ʾArsūzī, also a founding figure of the party, regarded it as the original language of mankind and as the only natural, nonconventional language, which preserved the initial congruence between meaning and its object (Carré 1980:199). By this, he gave the Arabs and their culture a unique place in history and a special message to other peoples, attempting to establish its universal significance even beyond the Qurʾānic revelation.

8.2 *ʿAbd an-Nāṣir and Arab socialism*

The Arab socialism of ʿAbd an-Nāṣir and his regime in Egypt had similar traits (Rejwan 1974; Kassian 1991). Nāṣirism also subordinated religion to national aims, put the religious institutions under close supervision, and tried to use them for an agenda of national mobilization. It regarded the early Islamic State as the first historical realization of socialism. In his speeches, ʿAbd an-Nāṣir not infrequently likened himself and his political reforms to the golden era of the Prophet, at the same time directing religious accusations against his opponents. The Charter of National Action (*mīṯāq al-ʿamal al-waṭanī*; published in 1962) was even hailed by the editor in chief of the journal of al-Azhar University with religious pathos, as "words from God", which nobody before ʿAbd an-Nāṣir had ever expressed (Kassian 1991:243). The Egyptian Revolution was presented as a *jihād* against the *jāhiliyya* of present-day society, and the revolutionary leaders were put on an equal scale with the early *mujāhidūn* (Kassian 1991:218). This strangely foreshadowed Sayyid Quṭb and his damnation of Egyptian government and society.

8.3 *ʾAnwār as-Sādāt*

The use of sacralized forms of political expression continued with ʿAbd an-Nāṣir's successor ʾAnwār as-Sādāt in the 1970s. In the famous speech he gave at the Knesset during his visit to Israel in 1977, the oscillation between religious and secular political notions is equally at work (Salem 1980). The speech is built on a number of key terms, like *ʾinsān* 'human being' and *ʾinsāniyya* 'humanity', *Allāh* 'God', *salām* 'peace', *masʾūliyya* 'responsibility', and *ʿadl* 'justice'. Religious metaphors are used by as-Sādāt in his attempt to present himself as a messenger of peace, with a message (*risāla*) from his people. He also stresses his destiny to shoulder the responsibility for both the Egyptian people and the Palestinians, which God dictated to him. God is further brought in as the creator of mankind and the guardian of equality among the peoples. As-Sādāt claims that he fulfills a religious task with his visit to Jerusalem, the land of God: to establish peace on a ground common to Muslims, Christians, and Jews – the worship of the One God. Apart from this transfer of religious rhetoric to international politics, he also made much use of it in Egypt itself, where he came to pose as *ar-raʾīs al-muʾmin* 'the Believer-President'. This did not prevent him from being branded as an apostate by his assassins in 1981.

8.4 *Modern Arabic literature*

In contrast to the sacralized pattern of legitimization common to a good part of the nationalist movements and regimes of the Arab world, even those with a socialist ideology, Arabic literature has largely been dominated since the 1950s by secular trends and ideologies and by a deep ambivalence with respect to the religious heritage. This ambivalence comes out with particular intensity in the writings of the Palestinian poets. Despite their leftist and nationalist inclinations, they often resort to religious symbols when expressing their most intense feelings of love, pain, hatred, and grief (Wild 1984). According to Stefan Wild (1984:294; see also Somekh 1991:79; Wild 2001), "Religion for these poets remains an ambivalent and oscillating symbol: an identity which has to be preserved, expression of deepest commitment whose sacred phrases easily overcome the loss of the established religion, and, on the other hand, a hindrance of progress, a prison, and a cause for revolt."

To a large degree this description would seem to apply also to other famous contemporary Arab poets, like ʿAbd al-Wahhāb al-Bayyātī (d. 1999) and ʾAdūnīs (b. 1930).

9. The Islamization of public life and the resacralization of Arabic political language

The gradual process of an "Islamization of Nationalism" (Leonard Binder) in the Arab world, which had been noted already in the 1960s (Aruri 1977:275), also had its impact on the political language. The process gained momentum during the 1970s, when Islamic movements like the Muslim Brotherhood and the Jamāʿa ʾIslāmiyya in Egypt came back to or newly appeared in the political arena and were able to reclaim the "terrain of social utopia" from the declining and persecuted leftist movements, parties, and trade unions (Kepel 1984:154ff.). A second process, less spectacular but perhaps even more important, was the gradual increase in religious orientation and practice that resulted from urbanization, and that also went along with a marked increase of interest in the *ḥajj* (see, for instance, for Iran and Jordan, Antoun 1989:97, 240). The move toward an Islamic political rhetoric gained in momentum after the spectacular events of the Iranian Revolution. These events demonstrated the political potential of the Islamic scholars, who until then had been regarded with condescension or even contempt by the ruling political elites; they also left their traces in the development of Arabic political language, which since the 1980s has gained an ever more intensive religious coloring.

9.1 *Political Islam*

The discourse of the Islamic movements themselves prepared the ground for a growing sacralization of politics. Ḥasan al-Bannā and Sayyid Quṭb provide classic examples, followed by many others. In their writings, the claim to reestablish God's sovereignty (*ḥākimiyya*) in human society and the totalitarian utopia of a unity of *dīn* and *dawla* goes along with a rhetoric of personal mission and dedicated action until death. Standard foe images are evoked and indefatigably impressed on the devout readers, directing them against Jews, Christian 'Crusaders' and missionaries, and communists, and against any form of secularism. The verdict of *jāhiliyya* condemns most of the present governments of the Arabic countries as unbelievers, and often contemporary Muslim societies as a whole are regarded by Islamic radical activists as systems of unbelief. At the same time, it must be stated that the political language of militant writers often builds upon the general ideals and values of their own societies, which they express in an Islamic terminology. Political aims and ideals like social justice, self-determination, and freedom from foreign influence and rule, put forward for a long time by Arab socialists, Baʿṯists, Marxists, and liberals, were more and more taken over by the Islamic movements. At the height of their activities, al-Jamāʿa al-ʾislāmiyya of Egypt presented a Charter of Islamic Action (*mīṯāq al-ʿamal al-ʾislāmī*) as their program. This was a calculated affront against the heritage of the Nāṣirist state, with its Charter of National Action already mentioned above (a picture of a poster with this charter appeared in *al-Majalla* 693, 19–25 May 1993, 24). These and other radical movements also made much use of the religious and legal genres of *xuṭba, fatwā*, and *risāla* to argue their political claims and views, and to give them a sound backing with quotations from the *Qurʾān* and the Prophetical Tradition and with the views of classical jurists, especially Ibn Taymiyya and Ibn al-Qayyim, but also many others (Jansen 1986).

9.2 Al-Qāʿida

The fusion of classical genres and formulas of the language of religion and Islamic law with the popular rhetoric of politics and the media fully comes out in the texts and video messages issued and distributed by *al-Qāʿida* and its supporters (cf. Kepel and Milelli 2005). Quotations from classical Arabic poetry, especially common in the writings of ʿAbdallāh ʿAẓẓām (d. 1989) and in the messages of ʾUsāma bin Lādin, provide a novel element which brings the rich Arabic stock of heroic poems into the focus of modern radical Islam. Martyr hagiographies are another genre that was revived and has become an equally important form of propaganda used in contemporary Islamic *daʿwā* (e.g. al-Quṭrī and al-Madanī 2003; *Katāʾib an-Nūr*). Research on this radical mobilization of Arabic rhetorical and literary resources is still in its beginnings. Only those traits of ʾUsāma bin Lādin's messages that are connected with language use are noted here.

The sacred aura of Bin Lādin's video messages is a product of both its performative form and its language, which also colors its political message. The general framework is a *xuṭba*, often introduced by formal praises of God and the Prophet, and by expressions of gratitude and even humility. Even the fiercest statements and threats are articulated in a soft voice and go along with mild religious expressions of thanks, hope, and gratitude, and the carefully planned attack of September 11 is presented as an act of God Himself:

> This is America, struck by God with one of her carnages. He destroyed her greatest building – praise and thanks to God. This is America, filled with horror from north to south, from east to west – praise and thanks to God. (*al-Ḥayāt*, 9 October 2001, 4)

The central message, repeated over and again, is an appeal to the Muslim *ʾumma* as a whole to become aware of the religious character of the conflict between the Muslims and America and the West, to join the camp of the righteous believers, and to support the *jihād*. The whole struggle is reduced to a religious war (*ḥarb dīniyya*), as prescribed by God in His book (Al-Jazeera, 7 November 2001). All political comments are geared toward that point. This goes along with the victim's discourse which pervades all statements: Muslims of different parts of the world, especially their children, are presented throughout as innocent victims of the suppression by Western powers and Israel which has been continuing unabated since the First World War and which is made worse by their allies in the Arab countries. The revolutionary term *mustaḍʿafūn*, already common in the Iranian Revolution, often comes in here. Humiliation has now been terminated by the mighty blows that struck America: "The wind of belief and the wind of change are now blowing in order to exterminate the evil!" (*habbat* *riyāḥ al-ʾīmān, habbat riyāḥ at-taġyīr li-ʾizālat al-bāṭil; al-Ḥayāt*, 9 October 2001, 4).

Frequent references to early Islamic figures and events place the whole of the present struggle in the pristine light of the Prophetic period. Already in his *Declaration of war against the Americans* of 1996, Bin Lādin mentions ʿAbd ar-Raḥmān ibn ʿAwf, the wealthy Companion of the Prophet, who during the Battle of Badr gave swords to two youths, who then went and killed ʾAbū Jahl, "the Pharao of this Umma" (MidEastWeb 2006). The role of Ibn ʿAwf, who advised and equipped the young fighters, might well illustrate the role which Bin Lādin saw – and perhaps still sees – for himself.

The suicidal martyrdom attacks are also justified with reference to a Prophetic tradition, a lengthy martyr legend attributed to the Prophet himself (Al-Jazeera, 27 December 2001). It illustrates the story of the People of the Trench (*ʾaṣḥāb al-ʾuxdūd*) in the *Qurʾān* (sura 85), and mentions two young believers who died as martyrs, one of them bringing about his own death by telling the pagan king how to kill him, the other encouraging his mother to jump into the burning trench (from Muslim, *Ṣaḥīḥ: Kitāb az-zuhd*). These two boys are presented as models for young believers, with their active quest for martyrdom and the encouragement which they provided to others. Here as elsewhere, the story bridges the gap between present-day politics and the heroic world of the early Muslims, enhancing the role of youth in the Islamic movements. The resacralization of Arabic political language has come full circle in these messages.

Bibliographical references

Primary sources

Aljazeera Net (*al-Jazīra*), *al-Maʿrifa.Taġṭiyāt xāṣṣa. Xiṭābāt wa-rasāʾil Bin Lādin wa-ẓ-Ẓawāhirī*. www.aljazeera.net/NR/exeres/392BC4D7-E9AA–4E5A–8F33–877EEB0C0781.htm (dated 14 May 2006, last checked 28 August 2006).

Azzam, Abdullah [ʿAbdallāh ʿAzzām]. 2001. *Join the caravan*. 2nd English ed. London: Azzam Publications. (Also available at http://islamistwatch.org/texts/azzam/caravan/caravan.html; accessed 27 August 2006.) [Trans. from *Ilḥaq bi-l-qāfila*, 1407/1987.]

Ibn Ḍayfallāh, *Kitāb aṭ-ṭabaqāt fī xuṣūṣ al-ʾawliyāʾ wa-ṣ-ṣāliḥīn wa-l-ʿulamāʾ wa-š-šuʿarāʾ*. Ed. Y.F. Ḥasan. 2nd ed. Khartoum: Khartoum University Press, 1974.

Jāḥiẓ, *al-Bayān wa-t-tabyīn*, ed. Ḥasan as-Sundūbī. Cairo: al-Maktaba al-tijāriyya, 1345/1927.

Katāʾib an-Nūr, *aṯ-Ṯamar al-mustaṭāb fī sīrat al-qāʿid Xāṭṭāb. http://saaid.net/Doat/index.htm (n.d., accessed 28* August 2006). [Hagiographic biography of the Saudi *mujāhid* in Chechnya.]

MidEastWeb. "Osama Bin Laden's jihad and text of fatwahs and declaration of War". www.mideastweb.org/osamabinladen1.htm (27 August 2006).

Secondary sources

Abou-Taam, Marwan, Ruth Bigalke, and Fiona Pappeler. 2006. *Die Reden des Osama bin Laden.* Munich: Diederichs.

Ammon, Ulrich, Norbert Dittmar, and Klaus J. Matteier (eds.). 1987. *Soziolinguistik: Ein internationales Handbuch zur Wissenschaft von Sprache und Gesellschaft.* (= *Handbücher zur Sprach- und Kommunikationswissenschaft*, III/1.) Berlin and New York: de Gruyter.

Antoun, Richard T. 1989. *Muslim preacher in the modern world: A Jordanian case study in comparative perspective.* Princeton, N.J.: Princeton University Press.

Aruri, Naseer. 1977. "Nationalism and religion in the Arab world: Allies or enemies". *Muslim World* 67.266–279.

Ayalon, Ami. 1987. *Language and change in the Arab Middle East: The evolution of modern political discourse.* New York: Oxford University Press.

Carré, Olivier. 1980. "Le mouvement idéologique baʿthiste". *La Syrie d'aujourd'hui,* ed. André Raymond, 185–224. Paris: Éditions du CNRS.

Corominas, Joan and José A. Pascual. 1981. *Diccionario crítico etimológico castellano e hispánico.* Madrid: Gredos.

Glaß, Dagmar. 1985. *Schlüsselwörter des arabischen Nationalismus: Zum Sprachgebrauch der Baʿṯisten um Michel ʿAflaq und Ṣalāḥ ad-Dīn al-Bīṭār*. Ph.D. diss., University of Leipzig.

Hofheinz, Albrecht. 1990. "Encounters with a saint: Al-Majdhūb, al-Mīrghanī and Ibn Idrīs as seen through the eyes of Ibrāhīm al-Rashīd". *Sudanic Africa: A Journal of Historical Sources* 1.19–59.

Jansen, Johannes J.G. 1986. *The neglected duty: The creed of Sadat's assassins and Islamic resurgence in the Middle East.* New York and London: Macmillan, 1986.

Kaempfert, Manfred. 1971. *Säkularisation und neue Heiligkeit: Religiöse und religionsbezogene Sprache bei Friedrich Nietzsche.* Berlin.

——. (ed.). 1983. *Probleme der religiösen Sprache.* Darmstadt: Wissenschaftliche Buchgesellschaft.

Kassian, Günter. 1991. *Die Orientierung an der frühislamischen Geschichte in der Ideologie des arabischen Sozialismus in Ägypten unter Nasser.* Ph.D. diss., University of Bonn.

Kepel, Gilles. 1985. *The Prophet and Pharaoh: Muslim extremism in Egypt.* London: Al Saqi. [Trans. from *Le Prophète et Pharaon*, Paris, 1984.]

——. and Jean-Pierre Milelli (eds.). 2005. *Al-Qaida dans le texte: Écrits d'Oussama ben Laden, Abdallah Azzam, Ayman al-Zawahiri et Abou Moussab al-Zarqawi.* Paris: Presses Universitaires de France.

Langen, August. 1968. *Der Wortschatz des deutschen Pietismus.* Tübingen: Niemeyer.

Lewis, Bernard. 1988. *The political language of Islam.* Chicago and London.

Littmann, Enno. 1950. *Aḥmed il-Bedawī: Ein Lied auf den ägyptischen Nationalheiligen*. Wiesbaden: F. Steiner.

——. 1951. *Islamisch-arabische Heiligenlieder*. Wiesbaden: F. Steiner.

Massignon, Louis. 1954. *Essai sur les origines du lexique technique de la mystique musulmane*. Paris: J. Vrin.

Moser, Hugo. 1964. *Sprache und Religion: Zur muttersprachlichen Erschließung des religiösen Bereichs*. Düsseldorf: Pädagogischer Verlag Schwann.

Nwyia, Paul. 1970. *Exégèse coranique et langue mystique*. Beirut: Dār al-Mašriq.

Piamenta, Moshe. 1979. *Islam in everyday speech*. Leiden: E.J. Brill.

Qāḍī, Wadād al-. 1993. "The impact of the *Qurʾān* on the epistology of ʿAbd al-Ḥamīd". *Approaches to the* Qurʾān, ed. Gerald R. Hawting and Abdul-Kader A. Shareef, 285–313. London and New York: Routledge.

Quṭrī, M. Ḥamad and Mājid al-Madanī. 2006. *Min qiṣaṣ aš-šuhadāʾ al-ʿarab*. Internet publication in Maktaba Ṣayd al-Fawāʾid al-ʾIslāmiyya, http://saaid.net/book/index.php. Ṣayd al-Fawāʾid (28 August 2006).

Rebhan, Helga. 1986. *Geschichte und Funktion einiger politischer Termini im Arabischen des 19. Jahrhunderts (1798–1882)*. Wiesbaden: O. Harrassowitz.

Rejwan, Nissim. 1974. *Nasserist ideology: Its exponents and critics*. New York, Toronto, and Jerusalem: Wiley and Sons.

Salem Babikian, Norma. 1977. "A partial reconstruction of Michel Aflaq's thought". *Muslim World* 67.280–294.

——. 1980. "The sacred and the profane: Sādāt's speech to the Knesset". *Middle East Journal* 34.13–24.

Samarin, William J. 1987. "The language of religion". Ammon a.o. (1987:85–91).

Sawyer, John F.A. and J.M.Y. Simpson (eds.). 2001. *Concise encyclopedia of language and religion*. Amsterdam and New York: Elsevier.

Somekh, Sasson. 1991. *Genre and language in modern Arabic literature*. Wiesbaden: O. Harrassowitz.

Stark, Werner. 1987. "Religion". Ammon a.o. (1987: 83–85).

Stiver, Dan R. 1996. *The philosophy of religious language: Sign, symbol and story*. Oxford and Cambridge, Mass.: Blackwell.

Wild, Stefan. 1984. "Judentum, Christentum und Islam in der palästinensischen Poesie". *Welt des Islams* 23–24.259–297.

——. 2001. "The *Koran* as subtext in modern Arabic poetry". *Representations of the divine in Arabic poetry*, ed. Gert Borg and Ed de Moor, 139–160. Amsterdam and Atlanta: Rodopi.

Stefan Reichmuth
(Ruhr–Universität Bochum)

Reported Speech → Indirect Speech

Resumption

Nominal elements in natural languages may appear in positions different from the ones in which they are interpreted. This is referred to as the displacement property of natural languages (Chomsky 1995). The examples in this entry are given in Lebanese Arabic.

(1a) *mīn fakkarto ʾənn-o sāmi*
who thought.2ms that Sami
ʿazam
invited.3ms
'Who did you think that Sami (has) invited?'

In sentence (1a), the interrogative WH-element (*mīn* 'who') is interpreted as the object of the embedded verb, even though it appears in the matrix clause. The displacement property, illustrated in (1b), is captured in generative transformational grammars by assuming that the WH-phrase has been generated in the embedded object position and then moved to the matrix position (→ WH-movement). This movement leaves a copy in the original position that is silent; this copy is referred to as trace (Chomsky 1981).

(1b) *mīn fakkarto ʾənn-o sāmi ʿazam*
who thought.2p that Sami invited
mīn
who
'Who did you think Sami invited?'

The movement in (1b) obeys stringent constraints. It cannot take place from within an adjunct clause (Adjunct Island Constraint [AIC]), as in (2a),

(2a) **mīn rəḥto min-dūn-ma tiʿəzmo*
who left.2p without invite.2p
*'Who did you depart without inviting him?'

nor from a complex noun phrase, such as a relativized clause (Complex NP Constraint [NPC]), as in (3a),

(3a) **mīn zaʿʿalto l-mara yalli*
who saddened.2p the-woman that

ʿazamit
invited.3fs
*'Who did you sadden the woman that invited?'

nor from a clause headed by a WH-element (WH-Island Constraint), as in (4a) (see Chomsky 1977).

(4a) **mīn* *baddkun* *taʿərfo* *mīn*
who want.2p know.2p who
ʿazam
invited.3ms
*'Who do you want to know who (has) invited?'

The island violations in (2a), (3a), and (4a) can be rescued by having a pronominal clitic resuming the WH-element in the matrix clause as in (2b), (3b), and (4b), respectively.

(2b) *mīn* *rəḥto* *min-dūn-ma*
who left.2p without
tiʿəzm-u
invited.2p-cl3ms
'Who did you depart without inviting him?'

(3b) *mīn* *zaʿʿalto* *l-mara* *yalli*
who saddened.2p the-woman that
ʿazamit-o
invited.3fs.-cl3ms
'Who did you sadden the woman that invited him?'

(4b) *mīn* *baddkun* *taʿərfo* *mīn*
who want.2p know.2p who
ʿazam-o
invited.3ms-cl3ms
'Who do you want to know who (has) invited him?'

The usual assumption, since Ross (1967), is that resumption in sentences (2b), (3b), and (4b) does not involve movement and thus is immune to the constraints applying to movement. In this respect, island constraints can be used as diagnostics for the existence of movement (see also Sells 1985).

Another diagnostic for movement is the phenomenon called 'reconstruction' (Chomsky 1995). Reconstruction refers to the fact that part of or the whole displaced element is interpreted in the position(s) from which it has been extracted. Bound pronouns are pronouns related to a quantificational antecedent or → quantifier phrase (such as *every* NP, *all* NP, and *no* NP) as in (5).

(5) *kill* *walad* *šēf* *mʿallimt-o*
every boy saw.3ms teacher.f-his
'Every boy saw his teacher'

The interpretation of bound pronouns varies with that of its quantificational antecedent. Sentence (5) is true if every boy (Sami, Karim...) saw his own teacher; there may be more than one teacher, one for each boy.

In order for a pronoun to be interpreted as a bound pronoun, it needs to be c-commanded by its quantificational antecedent (Higginbotham 1980). This condition is referred to as Condition on Pronominal Binding (CPB).

(6) *il-mudīra* *yalli* *ʿazamit*
the-principal.f that invited.3fs
kill *walad* *šēfit* *mʿallimt-o*
every boy saw.3fs teacher.of-his
'The principal who invited every boy saw his teacher'

In sentence (6), the pronoun cannot be bound by the quantificational antecedent because the quantifier phrase does not c-command it. (For the definition of c-command, → binding.)

Reconstruction can be used as evidence for the existence of a copy or trace left by movement. In (7), the pronoun contained in the displaced WH-phrase can be bound by the quantificational phrase, even though this quantificational phrase does not c-command it. This is because the movement leaves a full copy in the original position and the quantificational phrase c-commands the pronoun in the original position.

(7) *ʾayya* *tilmīz* *min* *tlemīza*
which student from students-her
fakkarto *ʾənno* *kill* *mʿallme*
thought.2p that every teacher
bitḥibb *tiʿəzm*
likes.3fs invite.3fs
'Which of her students did you think that every teacher likes to invite?'

(8) *ʾayya tilmīz min tlemīz-a*
which students from students-her
ʾənbasatto la-ʾənno kill mʿallme
pleased.2p to-that every teacher.f
bitḥibb tiʿəzm-o
likes.3fs invite.3fs-cl3ms
'Which of her students were you pleased with because every teacher likes to invite him?'

Sentence (8) contrasts with sentence (7) in that the WH-phrase could not have been extracted by movement from within the adjunct (AIC). The WH-phrase is directly generated in its surface position and does not leave a copy or trace. As such, the quantificational phrase in the adjunct clause does not c-command a copy of the pronoun contained in the WH-phrase, which is why the pronoun cannot be bound by the quantificational phrase.

In brief, reconstruction and island constraints can be used as diagnostics for the existence of movement. With this in mind, consider again the sentence in (1a) repeated here as (9):

(9) *mīn fakkarto ʾənno sāmi*
who thought.2p that Sami
ʿazam-o
invited.3ms-cl3ms
'Who did you think that Sami (has) invited (him)?'

In sentence (9), the WH-element in the matrix clause is related to a resumptive pronoun in the embedded clause. The presence of a resumptive pronoun is usually assumed to signal the absence of movement, since island violations are salvaged by resumption as discussed above (in relation to sentences (2b)–(4b)). This assumption, however, is not warranted. The discussion of examples (2b)–(4b) indicates that resumptive pronouns can be generated in nonmovement contexts. It does not establish, however, that a resumptive pronoun occurring in nonisland contexts such as (9) is not generated by movement.

Reconstruction, now used as diagnostic for the existence of movement, is available when the resumptive pronoun is not separated from its antecedent by an island. This is illustrated in (10), where the pronoun in the matrix WH-phrase can be bound by the quantificational phrase in the embedded clause.

(10) *ʾayya tilmīz min tlemīza*
which student from students-her
fakkarto ʾənno kill mʿallme
thought.2p that every teacher
bitḥibb tiʿəzm-o
likes.3fs invite.3fs-cl3ms
'Which of her students did you think that every teacher likes to invite (him)?'

The contrast between sentences (10) and (8) indicates that reconstruction is available with resumption when the antecedent and the resumptive elements are not separated by an island. In other words, movement is at work with resumption in nonisland contexts.

The above conclusion is based on the discussion of WH-constructions involving resumption (see Aoun a.o. 2001; Aoun and Choueiri 2000). A similar conclusion can be drawn from the discussion of left-dislocation constructions. Aoun and Benmamoun (1998) indicate that left-dislocated elements reconstruct only when the resumptive clitic is not separated from its antecedent by an island, as demonstrated by the availability of reconstruction in (11a) but not in (11b).

(11a) *tilmīz-a š-šāṭir fakkarto*
student-her the-good think.2p
ʾənno kill mʿallme
that every teacher
bitḥibb tiʿəzm-o
likes.3fs invite.3fs-cl3ms
'Her good student, you think that every teacher likes to invite him'

(11b) *tilmīz-a š-šāṭir ʾənbasatto*
student-her the-good pleased.2p
la-ʾənno kill mʿallme
to that every teacher of
bitḥibb tiʿəzm-o
likes.3fs invite.3fs-cl3ms
'Her good student, you were pleased because every teacher likes to invite him'

Thus, the generation of resumptive constructions is not uniform. The subsequent discussion is focused on what elements can be used as resumptives. In all the above examples, the resumptive element is a pronominal → clitic. In Lebanese Arabic, strong pronouns can be used as resumptives as well (Aoun and Choueiri 2000). In sentences (12a) and (12b), the dislocated element *sāmia* is resumed by a strong pronoun.

(12) *sāmia fakkaro ʾənno hiyye*
Samia thought.3p that she
bi-l-bēt
in-the-house
'Samia, they thought that she is in the house'

Interestingly, however, when the antecedent is quantificational, the distribution of strong pronouns is more limited. Strong pronouns can resume a quantificational antecedent only when they are separated from this antecedent by an island.

(13) No island
**kill muttahame ʿrəfto ʾənno*
each suspect.fs know.2p that
hiyye nḥabasit
she be.imprisoned.3fs
'Each suspect, you know that she was imprisoned'

(14a) Adjunct island
kill muttahame tfežaʿto
each suspect.fs surprised.2p
lamma/laʾənno ʿrəfto ʾənno
when/because know.2p that
hiyye nḥabasit
she be.imprisoned.3fs
'Each suspect, you were surprised when/ because you knew that she was imprisoned'
(14b) WH-island
kill muttahame baddkun taʿrfo
each suspect.fs want.2p know
mīn bifakkir ʾənno hiyye
who think.3ms that she
harabit
ran.away.3fs
'Each suspect, you want to know who thinks that she ran away'
(14c) Complex NP island
kill muttahame šēfto
each suspect.fs saw.2p
l-muḥāme
the-attorney.ms
yalli byaʿrif ʾənno hiyye
that know.3ms that she
harabit
ran-away.3fs
'Each suspect, you saw the attorney that knows that she ran away'

Epithet phrases in Lebanese Arabic, like strong pronouns, may function as resumptive elements, as illustrated in (15)-(16). See Aoun and Choueiri (2000),

(15) No island
ha-l-muttahame ʿrəfto ʾənno
this-the-suspect.fs know.2p that
ha-l-maždūbe nḥabasit
this-the-idiot.fs be.imprisoned.3fs
'This suspect, you know that this idiot was imprisoned'

As resumptive elements, epithet phrases like strong pronouns can be related to quantificational antecedents only when they are separated from these quantificational antecedents by an island. This is illustrated by the acceptability of the examples in (17) and the unacceptability of (16).

(16) No island
**kill muttahame ʿrəfto ʾənno*
each suspect.fs know.2p that
ha-l-maždūbe nḥabasit
this-the-idiot.fs be.imprisoned.3fs
'Each suspect, you know that this idiot was imprisoned'

(17a) Adjunct island
kill muttahame tfežaʿto
each suspect.fs surprised.2p
lamma/laʾənno ʿrəfto ʾənno
when/because know.2p that
ha-l-maždūbe nḥabasit
this-the-idiot.fs be.imprisoned.3fs
'Each suspect, you were surprised when/ because you knew that this idiot was imprisoned'
(17b) WH-island
kill muttahame baddkun taʿrfo
each suspect.sf want.2p know.2p
mīn bifakkir ʾənno
who think.3ms that
ha-l-maždūbe harabit
this-the-idiot.fs ran away.3fs
'Each suspect, you want to know who thinks that this idiot ran away'

(17c) Complex NP island
kill *muttahame* *šəft*
each suspect.fs saw.2s
il-muḥāme *yalli* *byaʕrif*
the-attorney.ms that know.3ms
ʾənno *ha-l-maždūbe* *harabit*
that this-the-idiot.fs ran.away.3fs
'Each suspect, you saw the attorney that knows that this idiot ran away'

(18) Obviation Requirement:
Strong pronouns and epithet phrases can resume a quantificational antecedent only when they are separated from this quantificational antecedent by an island.

An account for this generalization is found in Aoun a.o. (2001).

The general conclusions that emerge from this discussion of resumption in (Lebanese) Arabic are the following:

i. Resumption is a productive strategy. It is not solely used as a rescue device when movement is not available.
ii. Resumption is not generated uniformly. Movement is involved when the resumptive and its antecedent are not separated by an island.
iii. Pronominal clitics, strong pronouns, and epithets can be used as resumptive elements.
iv. The distribution of resumptive strong pronouns and epithets is more restricted than that of resumptive pronominal clitics. The former elements can only resume a quantificational antecedent when they are separated from it by an island.

In this entry, the linguistic theoretical apparatus is used to illuminate the behavior of resumption in (Lebanese) Arabic. The discussion of resumption has been restricted to constructions involving interrogatives. For discussion of resumption in relative clauses, see Ouhalla (2004) and Choueiri (2002) and references therein. The above discussion challenges some widely held assumptions regarding resumption as a last-resort strategy, at work when movement is not available. Other radical theoretical conclusions can be drawn from the study of resumption in (Lebanese) Arabic. Aoun and Li (2003), for example, show that constraints (such as Superiority), which have been viewed as regulating movement relations only, apply to resumptive constructions that do not involve movement. This finding leads them to assume that grammars in natural languages are not only derivational but also representational. That is, there are constraints that apply derivationally (bottom-up constraints) and constraints that apply to the resulting structure (top-down constraints). Space limitation prevents the discussion of these results in detail. It is sufficient here to highlight once more the dynamic nature of linguistic research in generative grammar. The theoretical apparatus can be used to illuminate the behavior of syntactic phenomena in a given language, in this instance (Lebanese) Arabic. In turn, the analysis of these syntactic phenomena can be used to refine, and even challenge, the theoretical framework. This is a healthy and fruitful interplay that advances linguistic science.

Bibliographical references

Aoun, Joseph and Elabbas Benmamoun. 1998. "Minimality, reconstruction, and PF movement". *Linguistic Inquiry* 29.569–597.

——. and Lina Choueiri. 2000. "Epithets". *Natural Language and Linguistic Theory* 18.1–39.

——, Lina Choueiri, and Norbert Hornstein. 2001. "Resumption, movement, and derivational economy". *Linguistic Inquiry* 32.371–404.

Aoun, Joseph and Yen-hui Audrey Li. 2003. *Essays on the representational and derivational nature of grammar: The diversity of Wh-constructions.* Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1977. "On *wh*-movement". *Formal syntax*, ed. Peter Culicover, Thomas Wasow, and Adrian Akmajian, 71–132. New York: Academic Press.

——. 1981. *Lectures on government and binding.* Dordrecht: Foris.

——. 1995. *The Minimalist Program.* Cambridge, Mass.: MIT Press.

Choueiri, Lina. 2002. *Issues in the syntax of resumption: Restrictive relatives in Lebanese Arabic.* Ph.D. diss., University of Southern California.

Higginbotham, James. 1980. "Pronouns and bound variables". *Linguistic Inquiry* 11.679–708.

Ouhalla, Jamal. 2004. "Semitic relatives". *Linguistic Inquiry* 35.288–300.

Ross, John R. 1967. *Constraints on variables in syntax.* Ph.D. diss., Massachusetts Institute of Technology.

Sells, Peter. 1985. "Restrictive and non-restrictive modification". *Report CSLI*, 85–28. Stanford, Cal.: Center for the Study of Language and Information, Stanford University.

Shlonsky, Ur. 1992. “Resumptive pronouns as a last resort”. *Linguistic Inquiry* 23.443–468.

Joseph Aoun (Northeastern University)

Resyllabification

Resyllabification involves adjustments of syllable structure across morpheme or word boundaries. In Arabic, there are two common reasons for resyllabification. The first is avoidance of clusters of three or more consonants, which typically triggers vowel → epenthesis. The second is satisfaction of the Onset Constraint: most Arabic dialects require onsets absolutely, which means that vowel-initial morphemes require either resyllabification of a preceding consonant as onset, or else epenthesis of the default consonant, which in Arabic is the glottal stop.

1. Cluster avoidance

According to the theory of directional syllabification (Selkirk 1981; Ito 1986, 1989; Farwaneh 1995), the algorithm which constructs syllables out of segmental material can proceed either from left to right or from right to left, with important consequences for the site of epenthesis. Famously, Cairene Arabic resolves a triconsonantal cluster by epenthesis of [i] between the second and third consonant (…VCCCV… →…VCC*i*CV…), while Iraqi Arabic resolves such clusters by epenthesis of [i] between the first and second consonant (…VCCCV… →…VC*i*CCV…) (Broselow 1980), as in (1).

(1) Triconsonantal clusters (Ito 1989:242)

a. Cairene Arabic

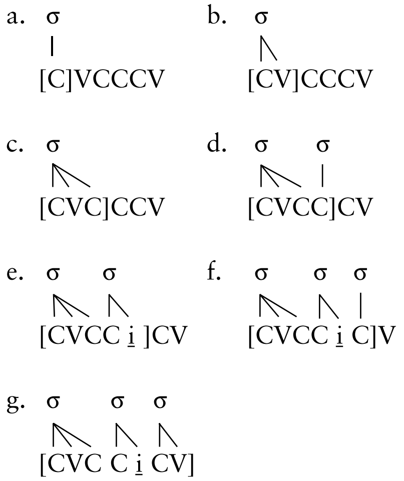
/ʔul-t-l-u/	> *ʾulti̲lu*	‘I said to him’
/katab-t-l-u/	> *katabti̲lu*	‘I wrote to him’
/katab-t dars/	> *katabti̲ dars*	‘you wrote a lesson’

b. Iraqi Arabic

/gil-t-l-a/	> *gili̲tla*	‘I said to him’
/trīd ktāb/	> *trīd i̲ktāb*	‘you [masc.] want a book’
/katab-t maktūb/	> *katabi̲t maktūb*	‘I wrote a letter’

The difference in choice of epenthesis site in triconsonantal clusters is due to a difference in syllabification algorithms: in Cairene Arabic, syllabification proceeds from left to right, while in Iraqi Arabic it proceeds from right to left. The process as it applies in Cairene Arabic is illustrated in (2). The maximal syllable template is CVC or CVV for nonfinal syllables and CVVC or CVCC for final syllables, and the direction of syllabification is left to right.

(2) Cairene Arabic (left to right)

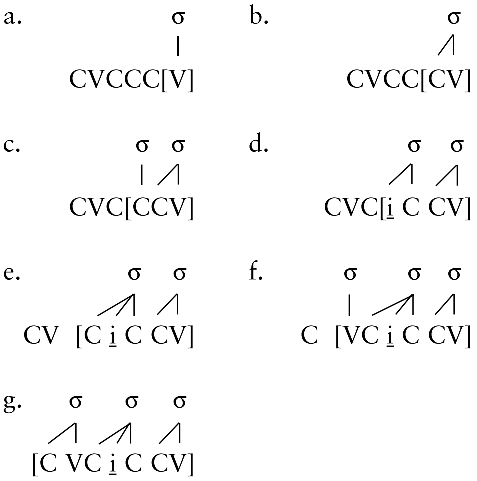


Steps (2a–c) illustrate stepwise construction of the first syllable, which fills out the syllable template completely. At this point, a new syllable must begin, as shown in (2d). But after (2d) there is a problem: complex onsets are not licensed by the syllable template. The remedy is to insert an epenthetic vowel, which in most cases is the default high front vowel, [i]; this is illustrated in (2e).

There appear to be two options for the next step, (2f): either incorporate the following C as a coda of the current syllable, or start a new syllable with that C as onset. Here the Onset Constraint comes into play. This constraint states that it is preferable for syllables to have onsets. The Onset Constraint is violated by the syllabification [CVC.V], and is satisfied by [CV.CV]. Hence, in step (2f) the C is made the onset of a new syllable. Lastly, in (2g), the final vowel is incorporated as a nucleus, completing the syllabification process. As desired, the epenthetic vowel appears between the second and third consonants of the cluster.

In contrast, the Iraqi syllabification algorithm proceeds from right to left, as shown in (3).

(3) Iraqi Arabic (right to left)



Steps (3a–b) are straightforward. At step (3c), the middle consonant of the triconsonantal cluster is syllabified as a coda, since the syllable template will allow maximally CVC. The difficulties begin at step (3d). The algorithm cannot build a complex coda, according to the syllable template, and hence vowel epenthesis is required. Then, in (3e), this syllable is completed by incorporation of an onset. The final syllable's construction is also straightforward (3f–g). And, also as desired, the output correctly shows epenthesis between the first and second consonants of the triconsonantal cluster.

2. HIATUS AVOIDANCE

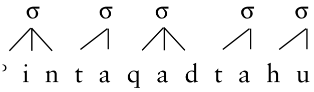
Most Arabic dialects absolutely require onsets (i.e., the Onset Constraint cannot be violated), and the glottal stop is usually the default option for epenthesis when an onset is lacking, as in (4).

(4) a.	*ʾintaqad-ta-hu*	‘you [masc.] criticized him’
	hal intaqad-ta-hu?	‘did you [masc.] criticize him?’
	(Classical Arabic)	
b.	*ʾil-kitāb*	‘the book’
	m-il-kitāb	‘from the book’
	(Cairene Arabic)	

Almost solely in utterance-initial position, a vowel-initial word triggers epenthesis of the glottal stop. The reason that glottal-stop epenthesis is not common in other positions is that a missing onset can almost always be supplied by resyllabifying a final consonant from a preceding word or morpheme. For instance, in (4a), the question particle *hal* contributes its final [l] as an onset for the following vowel: [.ha.lin.ta.qad.ta.hu.].

This sort of resyllabification occurs at least at morpheme, word, and phrase boundaries. Only when resyllabification is not a possible source of an onset does glottal stop epenthesis occur, as in (5).

(5)



Without epenthesis, this form would violate the Onset Constraint, and thus be ungrammatical.

BIBLIOGRAPHICAL REFERENCES

Broselow, Ellen. 1980. “Syllable structure in two Arabic dialects” . *Studies in the Linguistic Sciences* 10.13–24.

Farwaneh, Samira. 1995. *Directionality effects in Arabic dialect syllable structure*. Ph.D. diss., University of Utah.

Ito, Junko. 1986. *Syllable theory in prosodic phonology*. Ph.D. diss., University of Massachusetts, Amherst.

——. 1989. “A prosodic theory of epenthesis”. *Natural Language and Linguistic Theory* 7.217–260.

Selkirk, Elisabeth. 1981. “Epenthesis and degenerate syllables in Cairene Arabic”. *MIT Working Papers in Linguistics*, III. *Theoretical issues in the grammar of Semitic languages*, ed. Hagit Borer and Youssef Aoun, 209–232. Cambridge, Mass.: Dept. of Linguistics and Philosophy, Massachusetts Institute of Technology.

DAVID TEEPLE
(University of California, Santa Cruz)

Rheme → Theme/Rheme

Rhyme

Rhyme (*qāfiya*) is the recurrence of one sound or a number of sounds in the final position of poetic lines. Such recurrence is intrinsic to the music of verse. It plays the part of a recurrent refrain that lulls the ears of the listener and leaves the listener in a state of expectancy. The listener enjoys the recurrence of the refrain at regular temporal intervals, after a set of syllables that is determined by an intricate system of → meter (*wazn*).

Monorhyme characterizes Arabic poetry; that is to say, all lines have the same rhyme throughout the whole poem. Each line is composed of two hemistichs (*šaṭr*); the first one is called *ṣadr*, while the second is called *ʿajuz*. Generally speaking, the two hemistichs do not rhyme except occasionally; such a case is called *taṣrīʿ*, which can be represented graphically as:

________________a ________________a

A famous instance is a line by the Egyptian poet ʿAlī al-Gārim (d. 1949):

mālī futintu bi-laḥḏ̣iki l-fattāki
wa-salawtu kulla malīḥatin ʾillāki
'I wonder how I fell under the spell of your charming eyes; I can see no other beauty save you'

A poem with rhyming hemistichs throughout is rather a rare instance. It corresponds with the term *taṣrīʿ jāmiʿ ʾabyātihi* (Stoetzer 1989:63):

________________a ________________a
________________a ________________a

The following lines by the Egyptian poet al-Barūdī (d. 1904) can be taken as an example:

samiʿa l-xalliyyu taʾawwuhī fa-talaffatā
wa-ʾaṣābahu ʿajabun fa-qāla: "mani l-fatā"?
fa-ʾajabtuhu ʾinnī mruʾun laʿiba l-ʾasā
bi-fuʾādihi yawma n-nawā fa-tašattatā
'A carefree one heard me heave sighs of love.
He looked around and said:
"Lo, who are you, young man?". Came my answer: "I'm a lovestruck man, whose heart has been shattered by longing"

Another type of rhyme is called *muzdawaj*, whereby each line of the poem has a different rhyme, as an instance of *taṣrīʿ*. Such a case can be represented as:

________________a ________________a
________________b ________________b

ʾAḥmad Šawqī, Egyptian poet laureate (d. 1932), writes in this form:

yaḥkūna ʾanna rajulan kurdiyya
kāna ʿaḏ̣īma l-jismi hamšariyya
wa-kāna yulqī r-ruʿba fī l-qulūbi
wa-yukṯiru s-silāḥa fī l-juyūbi
'Once upon a time, there was a huge beastly Kurd, who used to carry heavy arms and scare all'

Mušaṭṭar is a different type of rhyme, in which each hemistich is an independent unit, governed by a unified rhyme. Several combinations of three, four, five, or six units were practiced in Arabic poetry, all encompassed by metricists under this title (ʾAnīs 1997:302). This rhyme is widely used in *muwaššaḥāt*, discussed in detail below.

Arabic metrical scholars have made several attempts to determine an exact definition of 'rhyme' (*qāfiya*) according to the number and the nature of the sounds used in the rhyme (cluster). The term that designates the minimally required number of sounds to be reiterated at the final position of lines is *rawiyy*. The *rawiyy* consists of a single sound, either a consonant or a vowel. The following lines are from Šawqī, given as an example of a *rawiyy* represented by a final /r/ preceded by a different short vowel:

rāḥilan fī miṯli ʾaʿmāri l-munā
ḏāhiban fī miṯli ʾājāli z-zaha
hāriban min sāḥati l-ʿayši wa-mā
šārafa l-ġamrata minhā wa-l-ġadur
'[He is] leaving in his prime, after a life as swift as a flower's. He has fled from a world in which he has experienced neither adversity nor deceit'

An example of a *rawiyy* represented by the long vowel /ā/, preceded by a different consonant, is taken from the Egyptian poet ʿAbbās al-ʿAqqād (d. 1964):

jānibu ṯ-ṯalji ʿalā n-nāri ṭaġā
ʿajabun ʾamruka yā haḏā ṯ-ṯarā
haḏihi d-dunyā l-latī naʿhaduhā
bidʿatun ʾam hākaḏā kullu d-dunā
qusimat ṯaljan wa-nāran fa-ʿtadā
jānibu ṯ-ṯalji ʿalayhā wa-ṭamā
'Ice has overwhelmed fire, Lo to you, O Dust, This world we live in, is it a novelty?, or this is how all the worlds go: Fire and Ice were partners, but Ice overpowered Fire and destroyed all'

A poem is usually named after its *rawiyy*; Šawqī's previous poem, for instance, is described as *rāʾiyyat Šawqī* (his poem rhyming in /r/). The /r/ sound is widely used in Arabic poetry, compared to other sounds. Sounds can roughly be classified according to their frequency (ʾAnīs 1997:248):

i. Most frequent sounds: /r/, /l/, /m/, /n/, /b/, /d/, /s/, /ʿ/
ii. Less frequent sounds: /q/, /k/, /ʾ/, / ḥ/, /f/, /y/, /g/
iii. Slightly frequent sounds: /t /, /ṯ /, / ṣ /, /h /, /ḍ/, /ṭ/
iv. Rarely used sounds: /ḏ/, /w /, /z /, /ḏ̣ /, / ġ/, /x /, /š/

Frequency of sounds in end rhyme is partly attributed to the sound production process, i.e. articulation. Some Arabic sounds, categories (iii) and (iv) above, require muscular effort, which explains their scarcity as *rawiyy*. On the other hand, certain sounds are widely current in word-final position in Arabic, and therefore the preferred choice of the poets. This has a considerable bearing on the frequency of sounds in the *rawiyy*.

Rhyme in Arabic poetry occurs in two forms according to its *rawiyy*. The first and the most common is the free rhyme, in which the *rawiyy* is influenced by a following short vowel, as in ʾAbū Nuwās' (d. 235/815) line:

ḥāmilu l-hawā taʿibu   yastaxiffuhu ṭ-ṭarabu
'A lover is afflicted by love, hence deeply moved by music'

The second is restricted rhyme, where the *rawiyy* is not followed by a vowel, as in this line by al-Buḥturī (d. 317/897):

qultu li-l-lāʾimi fī l-ḥubbi ʾafi
lā tuhawwin ṭaʿma šayʾin lam taḏuq
'I said to him who blames love: "Hark, do not undervalue what you have never experienced"'

Restricted rhyme represents a considerably smaller share of poems, not more than one-tenth of the whole poetic heritage (ʾAnīs 1997:260).

It is an established rule that in free rhyme the vowel that influences the *rawiyy* should invariably dominate the whole poem. Yet, there are instances where a poet would flout this rule. This phenomenon is called *ʾiqwāʾ*. An-Nābiġa aḏ-Ḏubyāni (d. 24/604) is reported to have committed this mistake in one of his poems, given here as an example:

min ʾāli mayyata rāʾiḥun ʾaw muġtadī
ʿajlāna ḏā zādin wa ġayra muzawwadi
zaʿima l-bawāriḥu ʾanna riḥlatanā ġadan
wa-bi-ḏāka ḥaddaṯanā l-ġurābu l-ʾaswadu
'Mayy's kinsmen are hastening to and fro, yet none of them brings me any news of her. My companions claimed that we depart tomorrow, and so heralded the black crow' (ʾAnìs 1997:261)

Ridf is the occurrence of a long vowel before the *rawiyy* in free rhyme. Ḥāfiẓ ʾIbrāhīm (d. 1932) wrote a full-length poem in this mode, using the long vowel /ā/:

rajiʿtu li-nafsī fa-ttahamtu ḥaṣātī
wa-nādaytu qawmī fa-ḥtasabtu ḥayātī
ramawnī bi-ʿuqmin fī š-šabābi wa-laytanī
ʿaqamtu fa-lam ʾajzaʿ li-qawli ʿidātī
'I questioned my senses and doubted my wisdom. I sought my kinsmen for help, but they let me down. My foes claimed I was sterile in my prime; I wish I were, for then I would grieve not at the claims of my foes'

Sinād at-tawjīh is another type of rhyming, where the poet flouts the rule of adhering to a specific short vowel preceding the *rawiyy* throughout the poem. In this mode, all three Arabic short vowels are used within the same poem, following no fixed order. A good example is found in three lines by Šawqī, selected randomly from his famous *Nahj al-burda* (190 lines), where /u/, /a/, and /i/ precede the free rhyme:

rīmun ʾalā l-qāʿi bayna l-bāni wa l-ʿalami
ʾaḥalla safka damī fī l-ʾašhuri l-ḥurumi
laqad ʾanaltuka ʾuḏnan ġayra wāʿiyatin
wa-rubba mustamiʿin wa-l-qalbu fī ṣamami
ṣalāḥu ʾamrika li-l-ʾaxlāqi marjiʿuhu
fa-qawwim an-nafsa bi-l-ʾaxlāqi tastaqimi
'[I met] a charming deer, between al-Ban and al-ʿAlam, where she unjustly sentenced me to death. / I have turned deaf ears toward you; a man would hear you, but his heart listens not. / Ethics governs your life, observe it closely if you wish to lead a happy life'

Three major metrical scholars had three different approaches to *sinād at-tawjīh*. Al-ʾAxfaš (d. 211/791) approved of it, saying it was not a defect. Kurāʿ an-Naml (d. 307/887?) rejected the vowel /i/ but approved the possibility of the co-occurrence of /u/ and /a/. The third scholar, al-Xalīl (d. 206/786), rejected the occurrence of /a/ in the same poem as a variation with /i/ and /u/ (ʾAnīs 1997:268).

Al-Xalīl's approach corresponds with modern theories of phonetics, which state that /i/ and /u/ belong to one category, i.e. closed vowels. In the articulation of these vowels, the tongue is raised to its highest position, toward the hard palate, leaving a very narrow passage for the air. On the other hand, /a/ belongs to another category, i.e. open vowels, where the tongue is lowered to the maximum in the mouth cavity, leaving a wide passage of air (Bišr 1975:144).

Arabic poets and metrical scholars have always realized the special auditory effect of the long vowel /ā/, which surpasses /ī/ and /ū/ in length and strength due to its special articulation. Accordingly, it is used frequently before the *rawiyy*, or even with an interceding consonant to add to the vigor of expression or to accentuate the lyrical effect of the lines. This mode is known as *ʾalif at-taʾsīs*. Al-ʿAqqād follows this mode in his lines:

> *lahijat bi-ḥusniki ʾalsunun wa-xawāṭiru*
> *wa-ṣabat ʾilayki jawāniḥun wa-nawāḏ̣iru*
> *wa-jarā ġarāmuki fī damī fa-twahhajat*
> *qatarātuhu fa-huwa l-ḥamīmu l-fāʾiru*
>
> 'Your charm has overwhelmed all eyes and hearts. Your love has streamed into my heart, erupting fire into my soul'

The perfect musical rhyme, according to ʾAbū al-ʿAlāʾ al-Maʿarrī (d. 477/1057), was composed of the maximum number of sounds in a rhyme cluster. The following examples come from his famous book *al-Luzūmiyyāt*, where he uses seven different sounds for the end rhyme:

> *ʾiḏā mā ʿarākum ḥādiṯun fa-taḥaddaṯū*
> *fa-ʾinna ḥadīṯa l-qawmi yunsī l-maṣāʾibā*
> *wa-ḥīdū ʿani l-ʾašyāʾi xīfata ġayyihā*
> *fa-lam tujʿali l-laḏḏātu ʾillā naṣāʾibā*
>
> 'If hardships befall you, relieve your souls by talking to your kin. Shield yourselves against the temptation of worldly pleasures, for your lots are justly distributed'

The line of Arabic poetry is an independent unit. This gives grounds for the double caesuras, or long pauses, within each line. The first one is obligatory at the end of the second hemistich (*ʿajuz*), while the second is arbitrary at the end of the first one (*ṣadr*). This latter caesura is usually observed in longer meters, probably to relieve the length of these verses. But in a run-on line, the last word of the first hemistich is split across the hemistich break. End rhyme is the necessary technical device required to integrate the structural unity of Arabic verse.

Al-Xalīl defined the 16 meters of Arabic verse, organizing previously unwritten rules into a coherent prosodic system. Each of the 16 meters dictates a line length which remains unvaried throughout the poem. The line is the basis of the musical unity of the traditional poem, and it is divided into two balanced parts, or hemistichs. Each hemistich is further divided into an equal number of syllables and accents. Patterns of repetition of these syllables and accents compose the poetic foot, or *tafʿīla*. Rhyme completes the musical unit at the end of each line and plays a significant role in sustaining the music from one line to another. In some cases, the rhyme serves as the binding agent to complete the idea.

Given the fact that in Arabic poetry each single verse is regarded as a structural unit of its own, rhyme serves to fulfill two important functions. First, it 'locks' each verse, thereby announcing its musical independence. Yet, it offers the only technical link between the structurally independent units. Each verse, as a closed musical unit, should exhaust its own content or meaning. The binding music of the rhyme confirms the total integration of theme and form. Secondly, rhyme serves to complete the recurring, identical pattern by which symmetry and equilibrium are sustained. Eagerly anticipated at the end of each verse, it adds to the sense of resolution effected by the cadence of regular meter.

The formal qualities of traditional Arabic poetry are demanding indeed, despite the license afforded by divisions and subdivisions of the meter. The aesthetic requirements of monorhyme, symmetry in form, and roundness of expression discourage any attempt to alter the long-established elegance of the poetic idiom.

The earliest objection to the traditional form of the Arabic poem appeared in the 2nd century

A.H. Frustration and dissatisfaction with the classical rigidity of form, coupled with an impetus to change, were represented by two major poets, ʾAbū Nuwās and ʾAbū Tammām (d. 265/845). They revolted against traditional meter and rhyme schemes and dismissed them as primary impediments to any will to change. They called for the adjustment of poetry to a changing environment and to a poet's own purpose.

In ʾAbū Tammām's poetry, the reader discerns a special technique of wordplay and an unusual usage and meaning of words as sources of beauty devised by the poet to present a modernism of a very high order. He also adopts a syntactic order of his own, where he omits points of detail and uses unfamiliar structures, thus creating hidden allusions and strange logical sequences. Together with ʾAbu Nuwās, he ventured further to effect a transformation in poetic form.

The drastic changes brought about by these two major poets were incorporated into the poetry of al-Mutanabbī (d. 385/965) and al-Maʿarrī at the height of the classical period. Yet, however much these figures were admired and emulated, their successors could not embrace the fundamental changes they introduced.

Modern Arabic poetry has undergone various attempts to change its forms, topics, and allusions. Movements in Arabic poetry have questioned the dictate of the fixed system of metrical forms, which had remained virtually unchanged from the pre-Islamic period to the 20th century. Despite ardent efforts on the part of Neoclassical and Romantic poets during the 1990s to introduce drastic changes, the grandiose style of the classical poem remained the cherished heart of the Arabic literary heritage. Treating the time-honored themes of love, pride, courage, heroism, war, laments for the traces of abandoned campsites, death, and nature, it had followed a great tradition deeply imprinted on the literary consciousness.

The oral nature of the traditional classical poem has helped to preserve the primary influences of music, established by regular rhymes and metrical patterns. But modern poets felt that the need was dire to invest the poem with stronger intellectual elements and new concepts of rhyme, meter, and form, in an age of accelerated scientific development. They aimed to merge the collective and personal experience in their poetry into contemporary life.

A new renaissance was brought about in the latter half of the 19th and the early-20th centuries by great poets in Egypt such as Maḥmūd Sāmi al-Bārūdi, Ḥāfiẓ ʾIbrāhīm, and ʾAḥmad Šawqī. They tried to revive the Arabic poem and restore its classical grandiose content and traditional techniques of meter and monorhyme.

Modern Arab poets of the mid-20th century suffered from the fact that much classical poetry abounds with words that hardly add to the meaning of the poem, yet are demanded by meter and rhyme. Modernists of the caliber of Badr Šākir as-Sayyāb (d. 1964) from Iraq, ʿAlī ʾAḥmad Saʿīd (ʾAdūnīs; b. 1930) from Syria, and ʿAlī Maḥmūd Ṭāhā (d. 1949) from Egypt are among the most notable leading figures. Their poetic practice did not abide by the laws of meter but tried instead to let prosody work to enhance the poem's music, sharpen its feelings, and intensify its images. A flow of internal music is heard in the following lines by ʾAdūnīs:

> *zamanun yajrī*
> *zamanun yahrubu miṯla l-māʾi*
> *wa-ʾana ʾajrī*
> *kullu nahārin sikkīnun fī ʾaḥšāʾī*
> *wa-l-laylu ḥirāb*
> 'Time runs, flows like a stream, and I run and run. Every morning cuts like a knife in my guts, and every night a spear'

The Dīwān school in Egypt, represented by al-ʿAqqād, ʾIbrāhīm al-Maznī (d. 1949), and ʿAbd ar-Raḥmān Šukrī (d. 1958), infused Arabic poetry with a new strain of Romanticism before the 1940s. These poets ushered in a new seriousness of subject, depicting the drama of human experience with a great share of sensitivity and with a notable absence of the techniques of form. Their poetic practice displayed daring prosodic experimentation. They used two *tafʿīla*s instead of three, or three instead of four, in a single line. They were also interested in quatrains in order to vary the symmetry of the meter and evade the monotony of the fixed rhyme scheme.

Inspired by the leading romantic poet Xalīl Muṭrān (d. 1949), the Dīwān school developed the musical element in the poem by substituting couplets or triplets for a more complicated rhyme scheme and, in some instances, by dispensing with the traditional two-hemistich line. Šukrī conducted the most extensive experiments with unrhymed verse (ʿAsfour 1992:25).

Šukrī wrote in the *dū-bayt* form, a Persian word meaning 'two lines'. This form consists of two rhyming lines, four hemistichs. But Šukrī used the *dū-bayt* without restrictions of rhyme, in an instance of *šiʿr mursal* 'blank verse':

> *yaqūlūna ṣ-ṣiḥābu ṯimāru ṣidqin*
> *wa-qad nablū l-marārata fī ṯ-ṯimāri*
> *šakawtu ʾilā z-zamāni banī ʾixāʾī*
> *fa-jāʾa bika z-zamānu kamā ʾurīdu*
>
> 'Friends are said to be the fruits of honesty; yet some fruits are bitter. Much have I suffered of my kinsmen, and you were my long-awaited reward'

The poets who were exposed to Western culture examined the development of its poetic form and content. They perceived that English romantic poetry, for example, had developed a language of its own. Their close examination of the poetic practice of such figures as Shelley, Wordsworth, and Coleridge was coupled with a strong rejection of the stringent traditional poetic form, the predictable rhyme scheme, and the drumming rhythms of the music. Henceforth, the poetic scene was set for the entrance of two schools: the Apollo school in Egypt, represented by ʾAḥmad Zakī ʾAbū Šādī (d. 1955), and the Mahjar 'Expatriates' poets in America, such as Mixāʾīl Nuʿayma (1889–1988), ʾIliyyā ʾAbū Māḍī (d. 1957), and Xalīl Jibrān (d. 1931), all from Lebanon. These poets benefited from a certain detachment that granted them an objective view of their native land, together with the habits and traditions that had shaped their thinking. Emigration provided them with a stronger impulse toward renewal. Their experiments with poetic form remain an uneven attempt.

The following lines are extracted from a poem by Jibrān, written in *dū-bayt* form, where all hemistichs save the third have the same rhyme:

> *ʾayyuhā š-šaḥrūru ġarrid*
> *fa-l-ġinā sirru l-wujūd*
> *laytanī miṯluka ḥurrun*
> *min sujūnin wa-quyūd*
>
> 'Sing, O dear thrush, for singing is the secret of life; I wish I were you, free from all bonds'

The poets of the Dīwān and the Apollo school believed that the poem's special music should carry psychological and philosophical undercurrents, and that the meter and rhyme scheme chosen should complement the content and emotional impact of the poem. By the late 1940s, a series of influential new literary magazines had appeared in Cairo, Damascus, Beirut, and Baghdad, presenting explorations of poetic form. Since then, a succession of broader attempts have been made to extend, rather than simply emulate, the tradition of the Arabic poem (ʿAsfour 1992:28).

During the 1940s, Arab poets grew increasingly dissatisfied with conventional concepts of poetic decorum, which imposed constraints on the poem's intellectual and thematic content and hindered the poets from bringing the full range of their religious, scientific, and philosophical experience to bear upon their work. A new concern with economy of expression, influenced strongly by the British and American imagists, accompanied the search for forms appropriate to subject matter and artistic purpose. Before the 1950s, the Arabic poem had never been much concerned with the variation of its musical structure.

Free verse, *aš-šiʿr al-ḥurr*, is attributed by some critics to the Egyptian initiator of the Apollo group, ʾAbū Šādī. Yet others assert that the first practitioners of free verse were as-Sayyāb and Nāzik al-Malāʾika (Iraq, b. 1923; ʿAsfour 1992:28).

With the decline of the traditional meter and rhyme, the musical consistency of the poem was altered, and the music of the 'new' poem could not be felt, unless read aloud. A number of poets ventured in free verse to introduce the *qaṣīdat an-naṯr* 'poem in prose', absolutely void of rhyme and meter. Yet, due to the enduring emphasis on the appealing musicality in Arabic poetry, a portion of the old metrical discipline has been preserved, manifesting itself in metrical free verse. The *tafʿīla* is the essential unit upon which this verse depends. The modern poem need not entirely dismiss rhyme but may instead create its own irregular inner rhyme, and the variations of the poetic foot (*tafʿīla*) affect the musical rhythmical structures. The poet may use a different *tafʿīla* in each line of the poem. The intricate scheme of the *tafʿīla* is liberated from the two-hemistich structure of the line, in an irregular metrical pattern that permits the line length to be determined by its content, often mixing two compatible meters, and permitting the use of run-on lines.

The free verse movement has generated experimentation on more than a prosodic level. The attributes of poetry have gradually gained importance. The length and shape of the poem, its point of view, and the suitability of dialect to the modern poetic idiom have been profoundly affected by the introduction of new forms. The long poem has made a comeback, but it seldom exceeds 150 lines, due to the limitations of the monorhyme scheme. Al-Malāʾika (Iraq) and Ṣalāḥ ʿAbd aṣ-Ṣabūr (Egypt, d. 1982) are foremost among the pioneers in this field. Similarly, the poetic play has enjoyed a revival at the hand of poets like ʿAbd aṣ-Ṣabūr, ʾAḥmad Suwaylam (b. 1942), and ʿAbd al-Muʿṭī Ḥijāzī (Egypt, b. 1935), now that it has been released from the restraints of the old meter. In the area of concrete poetry, several experiments have been made with the visual impact of the poem on the page.

Attempts to forge distinctive new forms have not all resulted in innovative poetry. The *muwaššaḥ* lit. 'ornamented scarf', for instance, is a verse form which was devised in Andalusia at the end of the 3rd century A.H., but which flourished in the 5th century A.H. It heavily depended on music and showed great variation of meter and rhyme schemes. Poets often mixed more than one meter in the same stanza. The *muwaššaḥ* is made up of several stanzas that follow an identical pattern in meter or meters, and in the rhyme scheme. It is written in two-hemistich lines of unequal length, and a case of *taṣrīʿ* is found in most of the lines.

Basically written to be sung, this exceptionally ornate form gives an atmosphere of playfulness and mirth. The influence of this form on free verse is limited to the fact that the revival of the art of the *muwaššaḥ* in modern times was a link in the chain of experimentation with the form of the Arabic poem, as it established the possibility of having hemistichs of different lengths in the same poem. Yet, it is not a free form of verse. One might say that it is even more restricted than the two-hemistich form in its rigid and elaborate composition. What at first sight would appear to have been a great revolution in poetic form was in fact nothing more than a clever attempt by the *muwaššaḥ* poet to apply Arabic words to certain musical tunes. Formal poetry remained untouched by this attempt, which could not breach its 'sacred' laws (Jayyusi 1977:556).

Some proponents of the new poetry continue to use an irregular blend of rhymed and unrhymed verse, while others use a pattern of long and short lines. Many are basically traditional poems modernized by simply cutting the traditional line into several ones on the page. Ḥijāzi's lines can be cited as an example:

> *ʾanaḏā yā zamana l-ḥurriyyah*
> *ʾašhadu mīlādaka fī ḏ̣-ḏ̣ulmah*
> *wa-ʾuġannī li-s-sārī fīhā*
> *li-l-mayyiti fī ʾaʿlā qimmah*
> 'O Time of freedom, here I witness your moment of birth in the dark; singing to him who departs, who dies at his peak'

Contemporary Arabic poets continue to experiment with great audacity, both in rhyme and rhythm. Yet, the Arab ear still cherishes an innate yearning for the music of the poetic heritage.

BIBLIOGRAPHICAL REFERENCES

ʾAnīs, ʾIbrāhīm. 1997. *Mūsīqā aš-šiʿr*. Cairo: Anglo Bookshop.

ʿAsfour, John Mikhail (ed. and trans.). 1992. *When the words burn: An anthology of modern Arabic poetry, 1945-1987*. 2nd rev. ed. Dunvegan, Ont.: Cormorant Books. (1st ed., 1988.)

Bišr, Kamāl Muḥammad. 1975. *ʿIlm al-luġa al-ʿāmm: al-ʾAṣwāt*. Cairo: Dār al-Maʿārif.

Jayyusi, Salma Khadra.1977. *Trends and movements in modern Arabic poetry*. Leiden: E.J. Brill.

Stoetzer, Willem F.G.J. 1989. *Theory and practice in Arabic metrics*. Leiden: Het Oosters Instituut.

Soheir G. Mahfouz
(Ain Shams University)

Rhyme Prose → Sajʿ

Riqāʿ → Ruqʿa

Romani → Gypsy Arabic

Root

1. THE NOTION OF ROOT

Root as an abstract linguistic unit (morpheme) may be seen as a common element in the nominal, verbal, adjectival, adverbial, and, to a certain degree, prepositional systems of Arabic. It is defined as an ordered set of consonants; vowels play a different role in the word derivation. Another role is played by the pattern, which represents the structure of the word itself (this structure is then filled in by the root and vocalization). Roots containing the phonemes *w, y*, and *ʾ* are considered weak (undergoing various morphonological changes). Roots can consist of two, three, four, five, and sometimes six consonants. Monoliteral roots are exclusively prepositions and other particles, which can also be bi- and triliteral. At the level of autosemantic words, nominal and verbal roots can be distinguished, the basic difference between the two being that the verbal roots take part in extensive word → derivation, while nominal roots are usually connected with a limited number of derived words (the derivative means being most commonly limited to the adjectival → *nisba -iyyun*). The basic form from which the derivatives are derived is sometimes difficult to reduce to the form of a consonantal root. Biliteral roots are mostly prepositions and other particles, as well as a small group of nouns, usually said to belong to the basic vocabulary, such as body organs, family members, etc. (→ biradicalism). According to Fleisch (1961:248, 252–254, with a full list), there are 37 biliteral words which can be identified with their root (e.g. *yadun* 'hand'). Triliteral and quadriliteral roots are verbal and nominal; roots with more than four radicals are almost exclusively nominal and often onomatopoetic in meaning. Autosemantic words have mostly three or more radicals, synsemantic words have one or two radicals, sometimes three, especially in the case of pseudoprepositions like *fawqa* 'above' (root *f-w-q*). The roots in which these pseudoprepositions appear usually have an inventory of derivational patterns (→ preposition).

The structure of the triliteral root is mostly 1–2–3 (e.g. *k-t-b* 'to write'), but a frequent type is also 1–2–2 (the so-called geminated roots, e.g. *m-r-r* 'to pass').

A great number of 1–2–1–2 types of roots (such as *w-s-w-s*) occur within the quadriliteral roots. These often carry descriptive (onomatopoeic) meaning (*waswasa* 'to whisper') or belong to loanwords (→ sound symbolism). Within the quadriliteral root, many elements like semivowels or coronal sonorants occur. Some authors regard these elements as taking part in the root extension.

In case the word consists of just one consonant, a root is usually not distinguished. But there are some words in Arabic apart from particles and prepositions that may be regarded as monoconsonantal, e.g. *fū* 'mouth'. For a list of possibly monoconsonantal words in Semitic, see Olmo-Lete (1998).

2. ROOT AS SEMANTIC BASIS

Roots (especially triliteral and verbal ones) are usually viewed as the semantic basis of a group of words derived from them. The consonants of the root form a skeleton on which the vowels are intercalated and affixes are added. By these means, words such as *KaTaBa* 'he wrote', *yaKTuBu* 'he is writing/he will write', *KiTāBun* 'book', *KuTuBun* 'books', *KiTāBatun* 'writing', *maKTaBatun* 'library' are formed from the root *k-t-b*. On average, about 17 forms are derived from a root, while the potential number is several times higher. Based on this phenomenon, the Semitic languages are sometimes described as introflective, i.e., → inflection takes place not only by (mainly desinential) affixes (→*ʾiʿrāb* 'desinential flexion') but also by restructuring the word in the manner indicated above. One could say that the root plays the role of the lexical morpheme, while the vocalization and affixes take part in the morphological → derivation from the semantic basis. In nonconcatenative (autosegmental) morphology (McCarthy 1986; → morphology), the set of morphemes (tiers) consists of the template (bare prosodic pattern), the root (consonantal melody), the vocalism (vocalic melody), and the affixes (consisting of both prosodic and melodic units). The specific tiers are put together by the tier conflation. An example of such an approach is given in Figure 1.

Figure 1. Tier conflation for the word *kutiba* ‘he/it was written’

```
consonantal tier   k    t    b      (root)
                   |    |    |
skeletal tier      C V  C V  C      (pattern)
                     |    |
vowel tier           u    i         (vocalism)
```

The OCP rule (see Sec. 4, below), as formulated by McCarthy (1986), does not allow repetition of adjacent consonants in the root. However, words with reduplicated final consonant are well attested in the vocabulary (e.g. *sababun* ‘reason, cause’). In his analysis, these are derived from underlyingly biconsonantal roots (*s-b*), and reduplicated consonants appear on the surface as a result of multiple association proceeding from left to right (see Fig. 2):

Figure 2. Reduplicated roots

```
s          b
|        /    \
C    V   C      V    C
```

The left-to-right direction prevents the system from deriving words like **sasabun*, a formation that is not attested in the lexicon.

The frequency of the different groups of roots (al-Bawwāb a.o. 1996) is given in Table 1.

Table 1. Frequency of groups of roots

	General	Verbal roots
Biliteral	115	0
Triliteral	7,198	5,590
Quadriliteral	3,739	1,830
Quinquiliteral	295	0
Total	11,347	7,420

These figures should be taken as relative and as suggesting only the proportions of the individual groups of roots. Greenberg (1950) collected from Classical Arabic dictionaries only 3,775 triliteral verbal roots. However, most of these figures seem to have been collected in a panhistoric way, reflecting the roots as they were used during the evolution of Arabic from the emergence of the Classical Arabic until today, without any chronological distinction of the various stages of the language.

According to Herdan (1962), there are statistical limits to the possible number of triliteral roots that can occur in Arabic. Taking into account all the possible constraints and the exclusion of *w* and *y*, the theoretical maximum number of all the possible combinations of consonants within the triliteral verbal roots is 6,332 (Herdan 1962:51–53).

3. Roots and derivation

The main role of the root is its function in word derivation (see, for example, Beesley 1999; Kiraz 1999). Beesley (1999) gives the following examples of templates:

DefineFormI: C V C V C
DefineFormII: C V C X V C
DefineFormIV: %’ V C C V C (the %’ represents the underlying glottal stop)
DefinePerfectActive: [a*]^V
Etc.;
[Ktb & FormI & PerfectActive] katab

The principal domain of the derivation based on the root is the verbal system. The entire system, with its opposition of perfect (*kataba*) vs. imperfect indicative (*yaktubu*), the verbal measures (derived verbal stems/forms, such as Form II *kattaba*; Form III *kātaba*; Form X *istaktaba*), and the system of participles and verbal substantives, is based on the root. Derivation is very stable and predictable for most of its members; for the verbal substantives, such stability is missing at the basic verbal pattern, or Form I (in this case, the number of possible patterns is relatively high; cf. Fleisch 1979:155–160; → verbal noun).

Another obvious domain where the root is employed is the system of the so-called broken (inner) plurals (→ number), which are created by means of intercalation of vowels and other affixes also (e.g., sg. *kitābun* > pl. *kutubun* ‘book > books’; sg. *ṣadīqun* > *ʾaṣdiqāʾu* ‘friend > friends’). For a detailed analysis, see, for instance, Ratcliffe (1998). Historically, these plurals can be seen as former → collectives (cf. Petráček 1960–1964). This may also be illustrated by several words in which the singular is a marked form (such as *ʿarabun* ‘Arabs’ vs. *ʿarabiyyun* ‘Arab, Arabic’).

Arabic also exhibits a tendency to organize the lexicon according to various forms of derivation. Thus, the form *miKTāBun* serves as the name of an instrument (*miFtāḤun* 'key'), the form *KaTTāBun* as the name of a profession (*najjārun* 'carpenter'), *maKTaBun* as the name of a place (*maKTaBun* 'a place/position where writing takes place: writing table, desk/office'), etc. The internal structure is also used for the formation of → diminutives (*KiTāBun* > *KuTayBun*) or → elatives (*ʾaKTaBu*), as well as other types of nouns. The highest stability of this type of derivation may be observed especially in the verbs, diminutives, and elatives. Other types of derivation are not as productive as the ones mentioned. However, the notion of the root and vocalization as separate layers with different roles can be applied to the vast majority of them.

There are, however, limits as to where such a type of derivation is employed. In fact, it is functional only in tri- and quadriliteral roots. There is no such derivation in monoliteral roots. The system of templates applies to biliteral roots only if the → weak verbs are considered to be based on biliteral roots (i.e. not containing a semivowel in the deep level of the root). The main domain of the system is triliteral and quadriliteral roots, where it functions with remarkable predictability. Outside these limits, there are various constraints on the occurrence of a derivation. When a quinquiliteral root, for instance, takes part in such a derivation, especially in the case of broken plurals, one of the root consonants is elided, as in *barnāmijun* > *barāmiju* 'program > programs', in order to fit into the quadriliteral derivational templates. In such cases, mainly liquids, laryngeals, and nasals on the 3rd, 4th, and 5th position are elided, i.e. mainly sounds that can be characterized as nonconsonantal and nonvocalic (for details, see Moujib 1998). This limit can also be observed in the incorporation of loanwords into the Arabic derivational system: root segmentation of loanwords and full integration of the loanword into the derivational system are possible only when a loanword has three or four consonants (e.g., *tilifizyōn* > *talfazatun/tilfāzun* 'television' and a possible verb *talfaza* 'to watch television'; *filmun/fīlmun* 'film, movie' > pl. *ʾaflāmun*; Tunisian Arabic: *dūš* 'shower' > verb *dawwaš* 'to take a shower', with insertion of a semivowel in a triliteral pattern).

4. Root at surface level

The root as an abstract unit appears only at the deep level; its manifestation at the surface level is indirect. Its existence can be proved by two arguments, the first being the OCP rule (→ Obligatory Contour Principle), the second one the mental representation of the root. The structure of the verbal roots has some limitations. These were known already to Arab linguists in the Middle Ages. Although the phenomenon has been discussed before (e.g. Aešcoly 1939; Cantineau 1960:199–202), the starting point for the current discussion is usually considered to be the work of Greenberg (1950). Greenberg showed that Arabic consonants divide into sections of homorganic consonants that tend not to co-occur within the same root, especially not at 'contact' positions within the root, like 1–2 or 2–3. According to McCarthy (1986), this tendency can be accounted for in terms of constraints against repeating homorganic consonants as an application of the OCP.

In nonconcatenative morphology, the OCP can be summarized as follows. In a given autosegmental tier (root), adjacent autosegments (i.e. consonants created at a close place of articulation) are prohibited. The major co-occurrence classes are:

i. Labials (*b, f, m*)
ii. Coronal obstruents (*t, d, ṯ, ḏ, ṭ, ḍ, s, z, ṣ, ḏ̣, š*)
iii. Velars (*k, g, q*)
iv. Gutturals (*x, ġ, ḥ, ʿ, h, ʾ*)
v. Coronal sonorants (*l, r, n*)

Consonants of these classes co-occur freely with consonants from other classes but tend not to co-occur with consonants from within their class, with two exceptions: velars cannot co-occur with the uvular approximants (*x, ġ*), but they can co-occur with other gutturals; among the coronal obstruents, fricatives and obstruents are more likely to co-occur than two fricatives or two occlusives. It should, however, be kept in mind that these constraints are much looser with nominal roots. Note that this rule is also violated in case of the reduplicated final consonant; in this theory, the 1–2–2 type of root is regarded as underlyingly 1–2 with spreading of the 2 to the empty C-slot in a triliteral pattern. Also the crossing of an underlyingly 1–2 root is prohibited (formations of the type 1–2–1). The domain (autosegmental tier) of this rule is the root, but across the morpheme boundaries, these constraints are not valid. Thus, while a root of the type *m-b-*3 is impossible, a formation consisting of prefix (e.g. *mu-*) and a root beginning with *b-* is quite common (*mu-BāRaK-un* 'blessed').

There are a number of exceptions to these rules (for a summary, see, e.g., Zaborski 1996). One could say that most of these violations consist of words with some emotional charge (e.g. *taṭā* 'to wrong a person, to tyrannize') or even onomatopoetic (*ġaqqa* 'to croak, crow').

It should be noted that the violations of the OCP principle are not very frequent in contemporary texts. For instance, while combinations like *ġ-h* in verbal roots do occur in dictionaries of Classical Arabic (three with this combination on the position 1–2 in de Biberstein-Kazimirski [1860], two of them verbal roots), they are very rare in contemporary Modern Standard Arabic. In a corpus of fifty million words, only one such instance of a verbal root has been found as an explanation of an archaic word (apart from foreign words such as *bīrminġhām* 'Birmingham'). This means that such violations may be expected especially in words with some additional (often emotional) semantic charge.

Other OCP violations are found in instances like the assimilation of the definite article (Watson 2002:220; the tier conflation takes place before the assimilation, which erases the phonological cue for morphemic distinction), various types of assimilation in the dialects of Arabic (cf. Watson 2002:222–224), etc. It should be noted that all of these OCP violations resolved by assimilations take place across the morpheme boundaries, i.e. across the boundaries of the autosegmental tiers, where normally the OCP rule should apply.

It should also be noted that the rule seems to be dynamic and to change in the course of time. The dynamicity of the rule can be illustrated also by the fact that other Semitic languages may have OCP rules of their own: the OCP rule on emphatic consonants in Akkadian (Geers 1945), for instance, does not apply in Arabic or Hebrew. However, there has as yet been no study of the dynamicity of this rule in Arabic, taking into account the historical development of the language.

The mental representation of the root (its existence in the minds of the speakers of Arabic) has been proven several times, both by experiments and empirical data. Prunet, Béland, and Idrissi (2000) show the mental representation of the root in Arabic/Semitic in the metathesis in aphasic speech which takes place within the root domain, while the pattern remains unchanged (e.g., *miNṬaQatun* 'area, region' > **miṬNaQatun*; *iḤ-ti-MāLun* 'probability' > **iḤ-ti-LāMun*; etc.).

There are a number of phonological processes that take place within the root. Among them, the most prominent are serialization of consonants and → metathesis. Serialization of consonants means that within the phonological correlations, it is possible for the consonants to change without changing the meaning of the root (e.g. *d-r-r/t-r-r* 'to spread, sprinkle'). According to Maizel' (1983), there are four such series (basically similar to the co-occurrence classes mentioned above; classes iii and iv would merge).

Metathesis within the root denotes the interchange of the root consonants, such as *bawš* – *wabaš* – *wišb* 'mob, rabble, crowd'. For details, see Maizel' (1983).

5. The original form of the root

The question of the original form of the root has been raised many times, mostly within a → Semitic and → Afro-Asiatic context. There are two main hypotheses: biliteralism (most frequently with a vocalic complement) and triliteralism (→ biradicalism; → stem). The main arguments of the biliteralist hypothesis are the alternations in the Semitic/Arabic root, the biliteral substantives, biliteral forms of weak verbs, the supposed biliteral character of Afro-Asiatic, and the 'root determinatives' in Indo-European.

The root in Arabic exhibits a great number of alternations of consonants at various positions (most frequently at the position R_3, but occasionally also at R_2 and rarely at R_1) while retaining the same or similar meaning. For instance, the roots *j-b-b, j-b-z, j-b-l, j-y-b, l-j-b*, etc. have a similar meaning 'to cut, divide into several pieces'. Such alternations can be observed also in other Semitic languages. These alternations are explained by the original biconsonantal character of the root in Proto-Semitic or by very common alternations of Arabic consonants, mostly based on phonological grounds (these cover especially alternations like *t/ṯ, t/d, t/t*, etc.; see the serialization of consonants as described by Maizel' 1983). An extreme application of the thesis of the original biconsonantal character of the root is found in Ehret (1995), who believes that the third consonant has a derivative meaning similar to the root determinatives in Indo-European (e.g. R_3 = *h*: amplificative; R_3 = *r*: diffusive; R_3 = *f*: iterative; etc.). The system of root extensions as presented by Ehret has a function similar to the verbal extension in the derived verbal Forms. Leaving aside the question of the reliability of his data, the fact that the alternations also take place at the position of R_2 and R_1 is not explained. Voigt (1988) is inclined to explain these alternations on the basis of phonetic/phonological considerations. For instance, for a QD basis (to which a third consonant can be added), alternations with Q and D (thus alternating the basis itself) are found. Thus, the consonant series at the position of the first radical could be *q/g/k/ġ/x*, and the series at the R_2 position could be *d/z/s/š/ṣ/ḍ/ṭ* (Voigt 1988:69).

In Bohas (1998), based on the analysis of the Arabic roots in de Biberstein-Kazimirski (1860), the lexicon (i.e. the 'root' system) is structured in the following way (→ lexicon: matrix and etymon model):

i. Matrix: Nonordered combination of features (labial, coronal, velar, uvular, pharyngeal, glottal) in a large semantic field; in other words, combination of two places of articulation; the OCP works at the level of features;
ii. Etymon: Nonordered combination of two consonants/phonemes emerging according to the matrix and representing the same semantic charge; e.g. ε{ğ, š} = *jašša*, *šajja* 'to break'; the OCP works at the level of phonemes;
iii. Roots: Combinations of the output of etymons with the skeleton and apophonic vowel; the OCP works as mentioned in Section 4.

In both Voigt and Bohas' analyses, processes of serialization and metathesis as described by Maizel' (1983) should be taken into account.

It should be noted that in these alternations, Arabic has the leading position among other Semitic languages. This may be due to the number of (consonantal) roots attested in various languages, whose number in Arabic is by far the largest among the Semitic languages, but also due to the approach that collects all the roots of the language without taking into consideration the historical data.

The biconsonantal words in Arabic (such as *yadun* 'hand', *ʾabun* 'father', *damun* 'blood'; for a detailed list, see Fleisch 1961:252–254) cover semantic fields usually labeled as members of the basic vocabulary (body parts, relatives, etc.) and thus interpreted as reflecting ancient stages of the language. Often, these words are assigned to the 1–2–2 type of root (e.g. *lubbun* 'heart') or to roots containing weak consonants (*w, y, ʾ*). This has led many authors to the conclusion that they reflect the original form of words in Proto-Semitic (CVC). Others (e.g. Voigt 1988:61–67) prefer to reconstruct the relevant protoform as containing one of the weak consonants, just as this is done for the explanation of the paradigmatic changes in the weak verbs, positing semivowels *w* and *y* at the deep level).

The supposedly biliteral character of Afro-Asiatic, used by some scholars to demonstrate the originally biliteral character of the root in Semitic, is challenged by many others. One could say that currently there are two principal schools in opposition to each other, and no consensus has been reached as yet.

The problem of the original form of the root in Proto-Semitic seems to be far from being solved. The only assertion that can be made is that when the Semitic languages separated from the Afro-Asiatic common stock, they possessed both biliteral and triliteral roots, and the triliteral type was probably prevailing.

Bibliographical references

Aešcoly, Aaron Zeev 1939. "Incompatibilités de phonèmes en hébreu et en araméen". *Groupe Linguistique des Etudes Camito-Sémitiques* 3.54–56.

Bawwāb, Marwān al-, Yaḥyā Mīr ʿAlam, Muḥammad Murayyātī, and Muḥammad Ḥassān aṭ-Ṭayyān. 1996. *ʾIḥṣāʾ al-ʾafʿāl al-ʿarabiyya fī l-muʿjam al-ḥāsūbī*. Beirut: Maktaba Lubnān Nāširūn.

Beesley, Kenneth R. 1999. "Arabic stem morphotactics via finite-state intersection". *Perspectives on Arabic Linguistics*, XII, ed. Elabbas Benmamoun, 85–100. Amsterdam and Philadelphia: J. Benjamins.

Biberstein-Kazimirski, Albert de. 1860. *Dictionnaire arabe-français*. 2 vols. Paris: Maisonneuve.

Bohas, Georges. 1998. *Matrices, étymons, racines: Eléments d'une théorie lexicologique du vocabulaire arabe.* Leuven: Peeters.

Cantineau, Jean 1960. *Etudes de linguistique arabe: Mémorial Jean Cantineau.* Paris: C. Klinksieck.

Ehret, Christopher. 1995. *Reconstructing Proto-Afroasiatic (Proto-Afrasian): Vowels, tone, consonants, and vocabulary.* Berkeley and Los Angeles: University of California Press.

Fleisch, Henri. 1961. *Traité de philologie arabe.* I. *Préliminaires, phonétique, morphologie nominale.* Beirut: Imprimerie catholique.

——. 1978. *Traité de philologie arabe.* II. *Pronoms, morphologie verbale, particules.* Beirut: Dar el-Machreq.

Geers, F.W. 1945. "The treatment of emphatics in Akkadian". *Journal of Near Eastern Studies* 45.65–67.

Greenberg, Joseph. 1950. "The patterning of the root morpheme in Semitic". *Word* 6.162–181.

Herdan, Gustav. 1962. *The calculus of linguistic observations.* The Hague: Mouton.

Kiraz, Georg Anton. 1999. "Computational tool for developing morphophonological models for Arabic". *Perspectives on Arabic Linguistics*, XII, ed. Elabbas Benmamoun, 101–110. Amsterdam and Philadelphia: J. Benjamins.

Maizel', S.S. 1983. *Puti razvitija kornevogo fonda semitskix jazykov* [The ways of development of the root inventory of the Semitic languages]. Moscow: Nauka.

McCarthy, John J. 1986. "OCP effects: Gemination and antigemination". *Linguistic Inquiry* 172.207–263.

Moujib, Ilham Dupont. 1998. "La formation des pluriels brisés issus des quinquilitères et des sextilitères en arabe". *Langues et Linguistique* (Fès – Meknès) 2.93–145.

Olmo Lete, G. del. 1998. "The monoconsonantal Semitic series". *Aula Orientalis* 16.37–75.

Petráček, Karel. 1960–1964. "Die innere Flexion in den semitischen Sprachen: Entstehnung und Entwicklung des Systems". *Archív Orientální* 28.547–606; 29.513–545; 30.361–408; 31.577–624; 32.185–222.

Prunet, Jean-François, Renée Béland, and Ali Idrissi. 2000. "The mental representation of Semitic words". *Linguistic Inquiry* 31.609–648.

Ratcliffe, Robert R. 1998. *The 'broken' plural problem in Arabic and comparative Semitic: Allomorphy and analogy in non-concatenative morphology.* Amsterdam and Philadelphia: J. Benjamins.

Voigt, Rainer M. 1988. *Die infirmen Verbaltypen des Arabischen und das Biradikalismus-Problem.* Stuttgart: F. Steiner.

Watson, Janet C.E. 2002. *The phonology and morphology of Arabic.* Oxford and New York: Oxford University Press.

Zaborski, Andrzej. 1996. "Some alleged exceptions to incompatibility rules in Arabic verbal roots". *Studies in Near Eastern languages and literatures: Memorial volume of Karel Petráček*, ed. Petr Zemánek, 631–658. Prague: Academy of Sciences of the Czech Republic, Oriental Institute.

Petr Zemánek (Charles University)

Ruqʿa

Ruqʿa lit. 'piece of paper' is a script which originated in the Ottoman chancery. It is known in that context as *riqʿa* (also *rikʿa, rika*), not to be confused with *riqāʿ* (the plural of *ruqʿa*), one of the 'proportioned' scripts (→ *ṯuluṯ*). This script is said to have developed from *dīwānī* (*divani*), the chancery script of the Ottoman Empire and represents an adaptation of the Persian *taʿlīq* (→ *nastaʿlīq*), most probably in the second half of the 12th/18th century.

It is said to have developed from *dīwānī* by simplifying its letter forms to such a point that there is little visible resemblance between these scripts. *Ruqʿa*, however, is much smaller, has straight and short ascenders, and is almost completely divested of the very pronounced, long, and left-sloping head-serif.

Ruqʿa became a standard script used for private correspondence and other purposes first among the Turks and later throughout the Arab world east of Tripolitania. It was also used for full texts, though rarely (Gacek and Yaycioğlu 1998:42, 62).

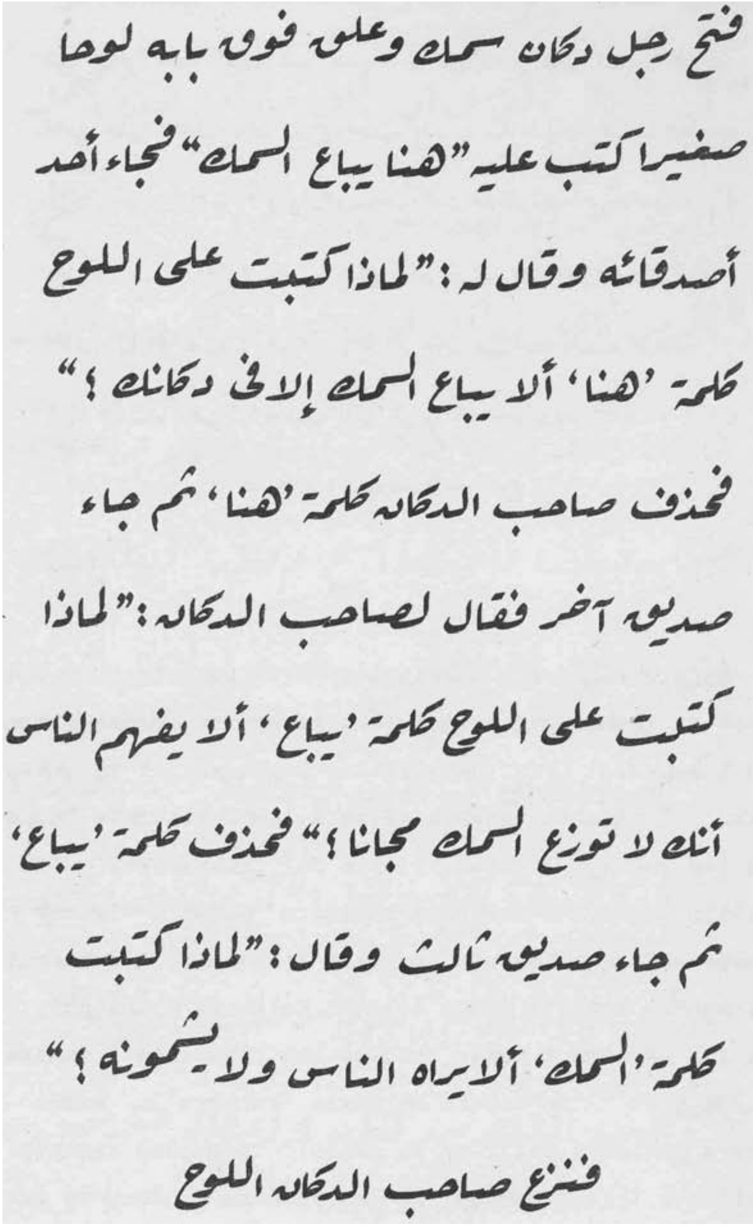
Mahmud Kemal Inal, the author of *Son hattlar*, a biographical dictionary of calligraphers, devotes an entire chapter to the masters of this script, who, as expected, were mainly high government officials (Schimmel 1984:23). *Ruqʿa* underwent a major reform in the hands of Mumtāz Bik (Beg), otherwise known as Mumtāz Afandī (Mumtaz Efendi; d. 1287/1871), a master calligrapher at the court of Sultan ʿAbd al-Majīd, and later Muḥammad ʿIzzat Afandī (Mehmet Izzet Efendi; d. 1320/1902); hence, it came to be known as *Mumtaz Efendi rikʿasi* or *Babi Ali rikʿasi*, and *Izzet Efendi rikʿasi* (Derman 1998:20).

In the Arab world, the main exponents of *ruqʿa* were ʿAbd ar-Rāziq ʿAwaḍ al-Miṣrī, Najīb Hawāwīnī, Muḥammad ʿAbd ar-Raḥmān, ʿAbd ar-Rāziq Sālim, ʾAḥmad al-Ḥusaynī, and Muḥammad al-Kurdī (ʿAfīfī 1980:154). An informal version of *ruqʿa* used for rapid writing is known in Turkish as *ruqʿa qirmasī* (*rikʿa kirmasi*) or 'broken' *ruqʿa* (Zakariya 1979:29).

Although there are a number of modern albums of calligraphy presenting the *ruqʿa* script, there is very little written on this script, its history, and salient characteristics (Muḥammad 1980). The only, and, one should add, very useful, manual of *ruqʿa* for Western students was made by T.F. Mitchell and published in 1953 (see Fig. 1). According to Derman, *ruqʿa* was always written with a reed pen whose nib was less than 1 mm wide (Derman 1998:20), and was always devoid of vocalization.

The spaces between letters and words in *ruqʿa* are equal. From the basic four letters (*ʾalif, bāʾ, ʿayn*, and *nūn*) the calligrapher can design all the letters of the alphabet. Thus, the letter *bāʾ* is transformed into the letter *kāf* by adding an *ʾalif* at the start. The letter *dāl* is drawn by omitting the last half of the letter *bā* (Al-Ali 1988:34).

Examining various specimens of *ruqʿa*, one notices the following: the two diacritical points (for *tāʾ, qāf*, and *yāʾ*) are represented by a short horizontal line, and the three points (in *ṯāʾ* and *šīn*) by an inverted tick. The *sīn* and *šīn* may be seriffed but have no 'teeth (denticles)'. No other



فتح رجل دكان سمك وعلق فوق بابه لوحا
صغيرا كتب عليه "هنا يباع السمك" فجاء أحد
أصدقائه وقال له: "لماذا كتبت على اللوح
كلمة 'هنا' ألا يباع السمك إلا في دكانك؟"
فحذف صاحب الدكان كلمة 'هنا' ثم جاء
صديق آخر فقال لصاحب الدكان: "لماذا
كتبت على اللوح كلمة 'يباع' ألا يفهم الناس
أنك لا توزع السمك مجانا؟" فحذف كلمة 'يباع'
ثم جاء صديق ثالث وقال: "لماذا كتبت
كلمة 'السمك' ألا يراه الناس ولا يشمونه؟"
فنزع صاحب الدكان اللوح

Figure 1. *Ruqʿa* specimen (Mitchell 1953:113)

letters (*ʾalif, lām, ṭāʾ*, etc) are seriffed, and the freestanding *ʾalif* is very short, while some letters in a word have a tendency to descend onto the baseline. Furthermore, except for the final forms of *mīm, jīm/ḥāʾ/xāʾ*, and *ʿayn/ġayn*, other letters such as *rāʾ, qāf, lām, nūn*, and *wāw*, with the traditional well-defined descenders, are written on the baseline.

A number of letters have very different forms depending on their position in a word (initial, medial, final). These include *šīn, ḍād, kāf, qāf, nūn*, and *hāʾ*. Furthermore, some letters in their final or isolated positions (e.g. *šīn, ḍād, qāf, nūn*) assimilate their diacritical points (one or two) into a final stroke ('penon'), attached to their tails. Finally, the numerals two and three in *ruqʿa* have specific forms: ٢ and ٣ (Mitchell 1953:107).

Although it was never used in printing, *ruqʿa* has survived to this day in the Arab world as the preferred script for correspondence and ordinary purposes of writing.

Bibliographical references

ʿAfīfī, Fawzī Sālim. 1980. *Našʾa wa-taṭawwur al-kitāba al-xaṭṭiyya wa-dawru-hā aṯ-ṯaqāfī wa-l-ijtimāʿī*. Kuwait: Wikālat al-Maṭbūʿāt.

Al-Ali, Salih. 1988. "Remarks on style". *Islamic calligraphy: Sacred and secular writings*, ed. David James a.o., 30–39. Geneva: Musée d'Art et d'Histoire.

Derman, M. Uğur. 1998. *Letters in gold: Ottoman calligraphy from the Sakip Sabanci collection, Istanbul*. New York: Metropolitan Museum of Art.

Gacek, Adam and Ali Yaycioğlu. 1998. "Ottoman Turkish manuscripts in the Islamic Studies Library and other libraries of McGill University". *Fontanus, from the collections of McGill University* 10.41–63.

Mitchell, T.F. 1953. *Writing Arabic: A practical introduction to* ruqʿah *script*. London: Oxford University Press.

Muḥammad, Hāšim al-Xaṭṭāṭ. 1980. *Qawāʿid al-xaṭṭ al-ʿarabī: Majmūʿa xaṭṭiyya li-ʾanwāʿ al-xuṭūṭ al-ʿarabiyya*. Baghdad: Maktabat an-Nahḍa; Beirut: Dār al-Qalam.

Schimmel, Annemarie. 1984. *Calligraphy and Islamic culture*. New York and London: New York University Press.

Zakariya, Mohamed U. 1979. *The calligraphy of Islam: Reflections on the state of the art*. Washington, D.C.: Center for Contemporary Arab Studies, Georgetown University.

Adam Gacek (McGill University)

Russian → Slavonic languages

S

Sabab

The word *sabab* literally means ‘a cord or ligature’, often ‘a tent rope’, signifying the tight structural bond between the fabric of the tent and the peg. In the Arab sciences, this concept has been exploited metaphorically in a number of ways, in prosody as the name for part of a metrical unit (another element being the *watid* lit. ‘tent peg’), in philosophy to denote a logical ‘cause’ or ‘reason’, and more generally (e.g. in law) for anything connected with or providing access to something else, such as relatives, dependents, or assets. In grammar it was first used by Sībawayhi (late 2nd/8th century) to refer to a semantic link between syntactic elements which produces a number of apparently irregular inflections. Thus, alongside the normal *zaydun laqītu ʾaxāhu* ‘Zayd [independent case as topic] I met his brother [comment sentence]’ = ‘I met Zayd’s brother’, we find *zaydan laqītu ʾaxāhu* ‘Zayd [dependent] I met his brother [explanatory sentence]’, where the dependent form of *zaydan* is accounted for by its being ‘semantically linked’ (*min sababihi*) to *ʾaxāhu* ‘his brother’, the true object of the verb (Sībawayhi, *Kitāb*, I, 32 Derenbourg/I, 43 Būlāq).

The semantic link is almost always indicated by an anaphoric pronoun, usually suffixed (such as *ʾaxāhu* ‘*his* brother’ in the example above), although occasionally reference may be through a ‘concealed’ pronoun (*ḍamīr mustatir*), e.g. *ʾanta fa-nḍur* “you, look [you]!”, where the 2nd person masculine singular agent pronoun is incorporated in the imperative verb *[u]nḍur* ‘look!’, or even by simple lexical repetition, as in the verse *lā ʾarā l-mawta yasbiqu l-mawta šayʾun* ‘I do not think death, anything will outrun death’, where the second instance of ‘death’ would, in prose, be replaced by the pronoun, scil. *yasbiquhu* ‘will outrun *it*’. That the pronoun is an essential feature of the construction is shown in the counterexample **mā zaydun munṭaliqan ʾabū ʿamrin* ‘Zayd [topic], ʾAbū ʿAmr’s father is not going away [comment sentence]’ (Sībawayhi, *Kitāb*, I, 24 Derenbourg/I, 31 Būlāq), which is disallowed even when it is known for a fact that ʾAbū ʿAmr is Zayd’s father (scil. *ʾabūhu*), contrast *mā zaydun munṭaliqan ʾabūhu* ‘Zayd, his father is not going away’, where *munṭaliqan* has dependent case as the negated predicate of ‘Zayd’ grammatically, but is logically the negated predicate of ‘*his* father’, with a marked semantic link.

The most common occurrence of this phenomenon is the construction later called the *naʿt sababī* ‘semantically linked adjective’. Compare the regular adjectival agreement in *marartu bi-rajulin ḥasanin* ‘I passed by a handsome man’, where *ḥasanin* ‘handsome’ agrees in gender, number, case, and definiteness with its head *rajulin* ‘man’ (masc. sg., obl., indef.), with the split agreement in *marartu bi-rajulin ḥasanatin ʾummuhu* lit. ‘I passed by a man beautiful his mother’, i.e. ‘I passed by a man with a beautiful mother’, where *ḥasanatin* agrees only in case and indefiniteness with its grammatical head *rajulin* (obl., indef.) and in gender and number with its logical head *ʾummuhu* (fem. sg.). This bidirectional agreement reflects the dual function of the adjective: syntactically it is an attribute of ‘man’, semantically it is a predicate of 'mother', and the connection (*sabab*) is manifested through the bound pronoun *-hu* in '*his* mother'.

An important subset of this structure is the so-called improper annexation, *ʾiḍāfa ġayr ḥaqīqiyya*, e.g. *marartu bi-r-rajuli l-ḥasani l-wajhi* lit. 'I passed by the man, the handsome *of* face', since it can be paraphrased in the 'semantically linked' form as _*bi-r-rajuli l-ḥasani wajhuhu_ 'by the man, the handsome *his* face', although this pattern is rare (→ *ʾiḍāfa*).

In an extension of the *sabab* construction, the 'semantic link' is somewhat weaker, and Sībawayhi terms it *iltibās* 'involvement'. Thus, in *marartu bi-rajulin muxāliṭihi dāʾun* lit. 'I passed by a man, a sickness infecting him' (Sībawayhi, *Kitāb*, I, 193 Derenbourg/I, 226 Būlāq; Mosel 1975:297), there is a sufficient 'involvement' between the head, *rajulin*, and the sickness, *dāʾun*, to permit the oblique case agreement in *muxāliṭi-* (definiteness agreement is neutralized in this situation). An even more complex relationship is identified by Sībawayhi as 'involved with something semantically linked' (*iltabasa bi-šayʾin min sababihi*), which we might regard as a transferred *sabab*, as in *marartu bi-rajulin muxāliṭin ʾabāhu dāʾun* lit. 'I passed by a man, a sickness affecting his father', where the semantic link (*sabab*) is now with '*his* father', *ʾabāhu*, itself the object of the participle *muxāliṭin*, which in turn is 'involved' through *dāʾun* 'sickness' with the head *rajulin*, and agrees with it in oblique case and indefiniteness.

Sībawayhi's analysis of the 'semantic link' and its effect on agreement is part of his overall preoccupation with constituent boundaries at the phrase level; these are determined by the interplay of contrasts between +/– agreement and +/– definiteness, marking inclusion in or exclusion from a constituent (see Carter 2002). Diachronic considerations as to the origins of the construction do not arise in the Arab sources, for which see Diem (1998), who shows, among other things, that attraction alone cannot account for the agreement features, nor can the structures be explained simply as derivatives of an underlying relative clause.

It has to be said that later grammarians seem to have had little appetite for the speculations of Sībawayhi, and his broad notion of *sabab* has almost disappeared from the literature, being virtually restricted to the *naʿt sababī* construction described above (already by the time of Ibn as-Sarrāj [d. 316/929], *Mūjaz* 62). There was, to be sure, a continued interest in the cohesive function of anaphoric pronouns, which led to a profusion of terms for something evidently taken for granted by Sībawayhi, since he does not use any of them, e.g. *rābiṭ(a)* 'binder, tie', → *ʿāʾid* 'referring element', *rājiʿ* 'returning element', *wuṣla* 'connector', *ʿulqa* 'bond, attachment'. But the emphasis now is on the mechanism of predication: in the absence of a verbal copula in Arabic, these joining elements were seen as the necessary link between subjects and predicates, a line of inquiry which no doubt resulted from the introduction of Greek logic into Arabic, raising issues which were scarcely perceptible when the *Kitāb* was being composed.

Bibliographical references

Primary sources

Ibn as-Sarrāj, *Mūjaz* = ʾAbū Bakr Muḥammad Ibn as-Sarrāj, *al-Mūjaz fī n-naḥw*, ed. Moustafa El-Chouémi [Muṣṭafā š-Šuʿaymī] and Bensalem Damerdji [Bin Sālim ad-Dāmirjī]. Beirut: A. Badran, 1965.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʾAmr ibn ʾUṯmān Sībawayhi, *al-Kitāb*. Ed. Hartwig Derenbourg. Paris: Imprimerie Nationale, 1881–1889. (Repr., Hildesheim and New York: Olms, 1970.) Ed. Būlāq, 1898–1900. (Repr., 2 vols. Baghdad: Muthanna [1965].) Ed. ʿAbd as-Salām Muḥammad Hārūn. 5 vols. Cairo: Dār al-Qalam, 1968–1977 and later eds.

Secondary sources

Carter, Michael G. 1985. "The term *sabab* in Arabic grammar". *Zeitschrift für arabische Linguistik* 15.53–66. [The bibliography lists sources additional to those mentioned in the present entry.]

——. 2002. "Patterns of reasoning: Sībawayhi's analysis of the *ḥāl*". *Proceedings of the 20th Congress of the Union of European Arabists and Islamicists. I. Linguistics, Literature, History*, ed. Kinga Devenyi, Tamás Iványi, and Avihai Shivtiel, 3–15. (= *The Arabist*, 24–25.) Budapest: Eötvös Loránd University.

Diem, Werner. 1998. fa-waylun li-l-qāsiyati qulūbuhum: *Studien zum arabischen adjektivischen Satz*. Wiesbaden: O. Harrassowitz.

Jahn, Gustav. 1895-1900. *Sībawaihi's Buch über die Grammatik, übersetzt und erklärt*. Berlin: Reuther und Reichard. (Repr., Hildesheim: Olms, 1965.)

Mosel, Ulrike. 1975. *Die syntaktische Terminologie bei Sībawaih*. Ph.D. diss., University of Munich.

Troupeau, Gérard. 1976. *Lexique-index du* Kitāb *de Sībawayhi*. Paris: Klinksieck.

ʿUḍayma, ʿAbd al-Xāliq. 1975. *Fahāris Kitāb Sībawayhi wa-dirāsa lahu*. Cairo: Maṭbaʿat as-Saʿāda.

Michael G. Carter (University of Sydney)

Ṣafaitic → Thamudic

Sajʿ

1. Introduction

Sajʿ is commonly known as rhymed prose. It is said to have → rhyme but no → meter, distinguishing it from poetry (*qarīḍ*), which features both rhyme and meter. *Sajʿ* is often associated with the text of the → *Qurʾān*, because large parts of the *Qurʾān* were composed in this type of rhymed prose. (In her study of the early Qurʾānic suras, Neuwirth (1981) expresses her doubt whether this rhyme can be considered *sajʿ*.)

The Arabic lexicographers usually derive the term *sajʿ* from the root *s-j-ʿ* in its sense of the 'cooing of doves', although a different etymology cannot be excluded. According to al-Fīrūzābādī's (d. 817/1414) *Tāj al-ʿarūs, s-j-ʿ* can also be associated with 'continuing an even, uniform course...following one order' (transl. Lane 1863–1893:IV, 1309–1310). An endorsement for this etymology would be that *sajʿ* as such is not restricted to the production of one fixed sound type, as the association with cooing would suggest. The common denominator of both etymologies would be repetition, rather than a certain kind of sound.

2. Form of *sajʿ*

Stewart (1990) develops a *poetica* for *sajʿ* on an empirical basis by analyzing the actual *sajʿ* structures in the *Qurʾān*. Alongside this analysis, he refers to the observations of medieval Arabic analysts like al-ʿAskarī (d. after 400/1010), Ibn al-ʾAṯīr (d. 637/1239), and al-Qalqašandī (d. 821/1418) for a theoretical framework. The result is a mixture of a prescriptive and descriptive canon of *sajʿ*. The rules of *sajʿ* as defined by Stewart can be summarized as follows:

i. The rhyming phrases are called *sajʿa, fāṣila, maqṭaʿ*, or *qarīna*;
ii. The rhyme of a *sajʿa* almost always ends in a consonant-*sukūn* cluster;
iii. *Luzūm* rhymes (i.e. rhymes based on more than one consonant) may occur;
iv. *Sajʿ* structures (or *sajʿ* units, i.e. a series of *sajʿa*s) may be introduced by a short line, not rhyming with the following *sajʿa*s and usually shorter than the *sajʿa*s themselves; Stewart on his own accord calls this line a *maṭlaʿ*, thus creating an ambiguous terminology: in *qarīḍ*, the first line of the poem is also called *maṭla*, usually a line that shows double rhyme; the technical term might better be reserved for this use;
v. The fundamental unit of a *sajʿa* is a word (*lafḏ̣a*), rather than a syllable or a *tafʿīla* (metrical foot or fixed combination of syllables);
vi. The number of words per *sajʿa* is normally roughly equal within each *sajʿ* unit, but varying quantities do occur; if the number of words is not equal, it tends to increase within the *sajʿ* unit, producing what Stewart calls 'pyramidic' constructions;
vii. *Sajʿ* has a qualitative accent, especially in the final word or *lafḏ̣a*; this would mean that verses like Q. 99/3 and 99/5 would have unexpected accents like this: ***māʾ*** *lahā* and *ʾaw**ḥāʾ** lahā* following the accents in parallel verse endings like *zilzālahā* and *ʾaṯqālahā*; this shows that a word like *lahā* is considered too short to be an independent *lafḏ̣a*, so it is combined with the preceding word to form one *lafḏ̣a*.

Stewart concludes that in form *sajʿ* is "a complex interplay of accentual meter, rhyme, and morphological pattern", producing "a type of composition distinct from both free prose (*naṯr mursal*) and syllabic verse (*naẓm*)".

The accentual nature of meter in *sajʿ* might be questioned, but this view receives support from an unexpected side. Although *sajʿ* is usually limited to prose texts, rhyming phrases within lines of poetry are not uncommon: Ibn ʾAbī l-ʾIṣbaʿ (d. 654/1256) in his *Taḥrīr at-taḥbīr fī ṣināʿat aš-šiʿr wa-n-naṯr wa-bayān ʾiʿjāz al-Qurʾān* lists a few figures of assonance or internal rhyme in *qarīḍ* poetry. The terminology he uses is quite complicated, and seems to depend on whether or not the internal rhyme coincides with the rhyme of the poem (*rawī*) and on the way the *sajʿa*s are distributed over the *tafʿīla*s.

The relevant categories that Ibn ʾAbī l-ʾIṣbaʿ distinguishes are *tasjīʿ* and *tarṣīʿ*. In *tasjīʿ* the *sajʿa*s share the rhyme with the *rawī* of the poem, the dominant end rhyme. Ibn ʾAbī l-ʾIṣbaʿ observes that in the two examples he cites, the meter of some of the *sajʿa*s is unsound (*baʿḍ*

ʾajzāʾihi ġayr muttazina zinatan ʿarūḍiyyatan). The two examples are from ʾAbū Tammam (d. 231/845):

> *tajallā bihi rušdī wa-ʾaṯrat bihi yadī // wa-fāḍa* [or: *ṭāba*] *bihi ṯamdī wa-ʾawrā bihi zandī*
>
> ‘Through him my guidance became clear [to me], through him my hand // became generous, through him my dried puddle flooded, through him my fire sticks gave sparks’ (meter *ṭawīl*, rhyme *-dī; Dīwān* I, 293)

and from Dīk al-Jinn (d. 236/850):

> *ḥurru l-ʾihābi wasīmahu barru al-ʾiyābi karīmahu maḥḍu n-niṣābi ṣamīmah*(*u*)
>
> ‘Free in its skin and beautiful, obedient and noble in returning, pure and unmixed of origin’ (meter *kāmil*, rhyme *-ābi…īmuhu; Dīwān* 154)

These examples are in fact metrically sound, but we do not know what versions Ibn ʾAbī l-ʾIṣbaʿ had in front of him. If he explicitly deems his versions metrically unsound, it would support Stewart’s opinion that the accentual rhythm dominates in *sajʿ*, rather than the quantitative *qarīḍ* type, even in cases when *sajʿ* is incorporated in poetry. Ibn ʾAbī l-ʾIṣbaʿ makes a similar observation in the chapter about *tarṣīʿ*, which resembles *tasjīʿ*: “The *tarṣīʿ* in the form of *sajʿ* consists of nonmetrical parts because *sajʿ* occurs in some of its parts” (*wa-l-musajjaʿ min at-tarṣīʿ ʾajzāʾ ġayr ʿarūḍiyya li-wuqūʿ as-sajʿ fī baʿḍ al-ʾajzāʾ*).

So in Ibn ʾAbī l-ʾIṣbaʿʾs view, these figures of speech share the unsoundness of meter. The main difference between *tasjīʿ* and *tarṣīʿ* seems to be that in *tarṣīʿ* the *sajʿa*s differ in rhyme from the end rhyme of the verse, thus producing the kind of independent, nonrhyming segments that Stewart calls *maṭlaʿ*, albeit on the wrong side of the *sajʿ* unit; these segments can be read as *maṭlaʿ*, though, if the reading of the *sajʿ* units starts at a different point in the line. The *sajʿa*s in the two examples by ʾAbū Tammām and Dīk al-Jinn coincide with the end of every second foot.

3. History of *sajʿ*

Beeston (1974) associates *sajʿ* with earlier forms of *parallellismus membrorum* in Old Egyptian, Ugaritic, Akkadian, and Hebrew. It is not clear whether or how this tradition survived the ages to reappear in Arabic literature, but it certainly might reflect a general Semitic literary form of expression. The oldest samples of *sajʿ* in Arabic are attributed to the *kuhhān* (pre-Islamic shamans), who used it for soothsaying and enchanting. Another example for its use is a series of statements by Jumʿa and Hind, daughters of al-Xass, who in pre-Islamic times were asked by al-Qulammas al-Kinānī about qualities and defects in camels, horses, goats, clouds, and men and women. This resulted in a number of monologues in rhymed prose in the *Balāġāt an-nisāʾ*, collected by ʾAḥmad ibn ʾAbī Ṭāhir (d. 280/893). Both of these examples are indications for the original functions of *sajʿ*: performative (enchanting) language, and mnemotechnic support (see below).

Apart from its functional purpose, *sajʿ* has an obvious artistic purpose as well: the use of adorned and formally structured language elevates utterances above everyday speech. In a genre like the dirge (*marṯiya*), which is closely related to the early performative lament (*niyāḥa*), one finds phrases that appear to have been originally conceived as *sajʿ* within the context of a *niyāḥa*, creating internal rhyme (*tarṣīʿ*) (Borg 1997:97). A *niyāḥa* structured in *sajʿ* and with the typical formulas of *sajʿ* is:

> *laqad kunta ṣaḥīḥa l-ʾadīm, manīʿa l-ḥarīm, ʿaḏīma s-silm, fāḍila l-ḥilm* [*luzūm*!], *wāriya z-zinād, rafīʿa l-ʿimād…*
>
> ‘You were healthy of skin, strong in defending private quarters, powerful in establishing peace, outstanding in kindness, caring for the wood fire, an elevated support…’

In *marṯiya*s, similar *sajʿ* structures can be found, e.g.

> *ḥammālu ʾalwiyatin šahhādu ʾanjiyatin qaṭṭāʿu ʾawdiyatin…*
>
> ‘A bearer of banners, a [reliable] witness to secrets, often crossing wadis…’ (Borg 1997:88)

and

> *yahdī r-raʿīl*(*a*) *ʾiḏā ḍāqa s-sabīl*(*u*) *bihim nahda t-talīl*(*i*) [*li-ṣaʿbi l-ʾamri rakkābā*]
>
> ‘He leads the fighting front when their way forward becomes narrow, [like] a high hill [always prepared to face difficulties]’ (Borg 1997:84–85)

In Arabic prose, *sajʿ* developed into a dominant means of expression in almost every field of *ʾadab* literature: book titles are often in *sajᶜ*, and as *sajᶜ* found its way into epistolography, letters, both official and unofficial, theoretical treatises, and other works that can be categorized as belletrie likewise are often in *sajᶜ*. *Sajᶜ* became the epitome of embellished language, creating an opportunity to exhibit lexical virtuosity as well as demonstrating the writer's broad knowledge of Arabic language and culture. Starting from the 11th century with the famous al-Ḥarīrī (d. 516/1122), *sajᶜ* came into use for a new genre in Arabic literature, the *maqāma*. The stylistic feature of parallel assonance in the *maqāma* developed quickly, and the genre became and remained popular until far into the 20th century, all over the Islamic world, from India to Spain.

Objections against *sajᶜ* as being artificial date back to the 14th-century scholar Ibn Xaldūn (d. 808/1406), but more recently the Orientalist Hamilton Gibb defended it as a stylistic means that contributed to the liveliness of the narrative. When in the 19th century the Arab world first discovered the Western genres of the novel and the short story, the *maqāma*—and with it the use of *sajᶜ*—proved to be the ideal medium to adapt Arabic narrative prose to these European models, because its main purpose (storytelling) could be used for European genres. In this way, the *maqāma* bridged the gap between Western content and Arabic form.

Storytelling in the Arabic tradition obviously knew pure prose forms, as in *The Arabian nights*, but probably this medium was considered too plain and too folkloristic to allow it to be used in prestigious literature, creating a gap for which the *maqāma* form was ideally fit. A typical feature of the *maqāma* was its comical intent, which could of course be supported by hilarious form features. A composer of *maqāmāt* like Bayram at-Tūnisī (1893–1961) used rhyme to underline comic situations, as in his *maqāma funūġrāfīya*, in which he forces a rhyme between the words *al-xayr* and *madrasat…al-furayr* 'École…des Frères', or *qawm* 'people' and *ṯawm* (i.e. *ṯūm*) 'garlic', probably to emphasize the class difference between an upstriving lower-class boy who becomes friends with an upper-class youth, who does not know Arabic.

Even in modern-day Cairo, mocking invectives in dialect can be heard in the cafés that are heartily exchanged between customers; these so-called *ʾāfiya*s (*qāfiya*) are short rhymed lines similar to *sajᶜ*.

4. FUNCTIONALITY OF *SAJᶜ*

Medieval Arab authors were reluctant to admit that the *Qurʾān* is in *sajᶜ*, although many do not hesitate to admit that large parts of the *Qurʾān* are composed in *sajᶜ*. The matter carries an ideological bias, especially with regard to the → *ʾiᶜjāz al-Qurʾān*: the conviction that the *Qurʾān* is God's word, unrivaled in its perfection. This explains the hesitation of the medieval authors: the occurrence of *sajᶜ* makes it hard to interpret the *Qurʾān* as God's word, because *sajᶜ* could obviously also be produced by man, or man-made *sajᶜ* might even surpass the quality of *sajᶜ* in the *Qurʾān*.

The use of *sajᶜ* in the *Qurʾān* needs to be seen in the context of its function in pre-Islamic times: early *sajᶜ* was used by the *kuhhān* for soothsaying and enchanting, a use of language which is highly performative; one might even say that it is language that almost equals action in order to change or grasp reality. A different purpose of *sajᶜ* may have been the case in enumerations like those mentioned in connection with Jumᶜa and Hind. Here, mnemotechnic support may have been a good reason for the use of *sajᶜ* because words in rhyme are probably more easily remembered and passed on than 'normal' sentences.

Sajᶜ in the *Qurʾān* can be associated with both of these traditions: it is found with a more complicated rhyme in short *sūra*s with a strong enchanting character, such as Q. 99 (*az-Zalzala* with a rhyme in *-ālahā*/*-ārahā*), whereas in longer *sūra*s repetitions of 'simple' rhymes are found, such as *-ūn* in, for instance, Q. 23 (*al-Muʾminūn*), Q. 27 (*an-Naml*), Q. 36 (*Yā Sīn*), Q. 44 (*ad-Duxān*), etc. This use of *sajᶜ* may well have contributed to the oral transmission of the Qurʾānic text. In terms of functionality, these two cases would coincide with the use of *sajᶜ* in pre-Islamic Arabia: performative in short *sūra*s and mnemotechnically supportive in longer *sūra*s. If *sajᶜ* is defined as a 'repetitive' structuring of language, it is not uncommon in other religious contexts, as, for instance, the litanies in the Catholic Church.

The use of *sajᶜ* in the *Qurʾān* strongly inspired its use in other contexts as an embellished and stylized linguistic practice. The question of whether or not it should effectively be regarded as poetry is a complicated one, because it involves Western and Arabic definitions of poetry. Stewart's position is a mitigated equilibrium: he qualifies *sajʿ* as a complex interplay of accentual meter, rhyme, and morphological patterns. But he also cites approvingly ʾAḥmad Šawqī's qualification: "*Sajʿ* is the second kind of poetry in the Arabic language" (*as-sajʿ šiʿr al-ʿarabiyyati aṯ-ṯānī*).

Bibliographical references

Primary sources

ʾAbū Tammām, *Dīwān* = ʾAbū Tammām Ḥabīb ibn ʾAws aṭ-Ṭāʾī, *Dīwān*. Ed. Dār al-Fikr al-ʿArabī, Beirut, 2001. [Together with the commentary by at-Tibrīzī.]

ʿAskarī, *Ṣināʿatayn* = ʾAbū Hilāl al-ʿAskarī, *Kitāb as-ṣināʿatayn*. Ed. ʿAlī Maḥmūd al-Bijāwī and Muḥammad ʾAbū l-Faḍl ʾIbrāhīm. Cairo, 1952.

Dīk al-Jinn, *Dīwān* = ʿAbd as-Salām ibn Raġbān Dīk al-Jinn al-Ḥimṣī, *Dīwān*. Ed. ʿAbdallāh Munīr Muhannā. Beirut, 1990.

Ibn Abī l-ʾIṣbaʿ, *Taḥrīr* = ʿAbd al-ʿAẓīm ibn ʿAbd al-Wāḥid Ibn ʾAbī l-ʾIṣbaʿ al-Miṣrī, *Taḥrīr at-taḥbīr fī ṣināʿat aš-šiʿr wa-n-naṯr wa-bayān ʾiʿjāz al-Qurʾān*. Ed. Ḥifnī Muḥammad Šaraf. Cairo, 1963. [Pp. 295–308.]

Ibn al-ʾAṯīr, *Maṯal* = Ḍiyāʾ ad-Dīn Ibn al-ʾAṯīr, *al-Maṯal as-sāʾir fī ʾadab al-kātib wa-š-šāʿir*. Ed. Dār Nahḍa Miṣr. Cairo, 1959–1962.

Qalqašandī, *Ṣubḥ* = ʾAḥmad ibn ʿAlī al-Qalqašandī, *Ṣubḥ al-ʾaʿšā fī ṣināʿat al-ʾinšāʾ*. Ed. al-Muʾassasa al-Miṣriyya al-ʿĀmma li-t-Taʾlīf wa-t-Tarjama wa-ṭ-Ṭibāʿa wa-n-Našr. Cairo, 1964.

Secondary sources

Beeston, Alfred F.L. 1974. "Parallelism in Arabic prose". *Journal of Arabic literature* 5.134–146.

Borg, Gert. 1997. *"Mit Poesie vertreibe ich den Kummer meines Herzens": Eine Studie zur altarabischen Trauerklage der Frau*. Leiden: Nederlands Historisch Archaeologisch Instituut te Istanbul.

Lane, Edward William. 1863–1893. *An Arabic-English lexicon derived from the best and most copious Eastern sources*. 8 vols. London and Edinburgh: Williams and Norgate.

Neuwirth, Angelika. 1981. *Studien zur Komposition der mekkanischen Suren*. Berlin and New York: Mouton de Gruyter.

Rowson, Everett K. 1998. "*Sajʿ*". *Encyclopedia of Arabic literature*, ed. Julie Scott Meisami and Paul Starkey, II, 677–678. London and New York: Routledge.

Šawqī, ʾAḥmad. 1970. *ʾAswāq aḏ-ḏahab*. Cairo.

Stewart, Devin J. 1990. "*Sajʿ* in the *Qurʾān*: Prosody and structure". *Journal of Arabic Literature* 21.101–139.

Gert Borg (University of Nijmegen)

Ṣanʿānī Arabic

1. General

Ṣanʿānī Arabic is the dialect of the original inhabitants of the Old City of Ṣanʿāʾ and its traditional suburbs, al-Bawniyah and al-Qāʿ (Qāʿ al-Yahūd). Ṣanʿānī belongs to the Eastern Muslim dialect type, and it is also spoken by Jews who emigrated from Ṣanʿāʾ to Israel after 1948. There are approximately one hundred thousand speakers in and around the Old City today, a figure which includes Ṣanʿānīs who left the Old City following the post–1991 Gulf War expansion of Ṣanʿāʾ. The number of Ṣanʿānīs who have emigrated to other Arab countries, Israel, mainland Europe, the United Kingdom, and North America and who still speak Ṣanʿānī Arabic is unknown.

Ṣanʿānīs in the Old City lead a traditional, sedentary lifestyle. Until the 1970s, life was particularly hard for women, who were required to draw water from wells, chop wood in the woods surrounding the Old City, and wash clothes at springs beyond the city walls. Life has eased with the introduction of electricity, piped water and gas, and, more recently, the importation of an increasing variety of foodstuffs and consumer goods.

Ṣanʿānī is probably the best known Yemeni dialect both within the country and without (→ Yemen). It is not, however, normally adopted by speakers from outside Ṣanʿāʾ, with the exception of women who marry into Ṣanʿānī families. The Old City retains a self-imposed isolation, actively discouraging the integration of outsiders, particularly males.

In addition to published historical documents such as *Ḥawliyyāt yamaniyya*, there are a number of collections of colloquial poetry, proverbs, and sayings which go back to the 19th century and earlier. Documental evidence from other Yemeni dialects, such as Rāziḥī (S. Weir p.c.), suggests that large bodies of handwritten documents relating to trade and legal disputes probably exist. Rossi's (1939) description of the dialect, based on fieldwork in the early 1930s, and Goitein (1934) provide points of comparison for studies considering recent development of the dialect. Research on the phonology, morphology, syntax, and semantics of the dialect has been carried out more recently by Naïm-Sanbar (e.g. 1994), Watson (1993, 1996, 2002), and Watson and Al-Amri (2000). No dictionary deals exclusively with Ṣanʿānī Arabic, although Yemeni glossaries by Behnstedt (1992, 1993, 1996) and dictionaries by Piamenta (1990–1991) and al-ʾIriyānī (1996) include some Ṣanʿānī words. Texts in Ṣanʿānī Arabic from Watson (1996), Watson and Al-Amri (2000), and Watson and Mutahhar (2002) can be heard on the Heidelberg Semitic Spracharchiv (http://semarch.uni-hd.de/).

Ṣanʿānī has been used by a number of public performers, particularly for humorous purposes, for instance by the comedians Muḥammad al-Maṭarī and ʿAbdullāh Šākir. The radio series *Musʿid wa-Musʿidih*, which attempts to improve social mores by satirizing undesirable aspects of Yemeni life, has been scripted and performed in Ṣanʿānī Arabic by ʿAbd ar-Raḥmān Muṭahhar since 1988. After Yemeni unification in 1990, the subsequently banned television series *Daḥbāš* satirized the accent and manners of the North.

2. LINGUISTIC DESCRIPTION

2.1 *Phonology*

2.1.1 Consonants

Ṣanʿānī has the consonant inventory in Table 1.

Ṣanʿānī also has a pharyngealized lateral, attested solely in the word *aḷḷāh* 'God' and derivatives. The Classical Arabic voiceless uvular stop, **q*, is not found in the dialect. Even religious and Standard Arabic words are pronounced with a voiced velar stop, /g/, as in: *algurʾān* 'the *Qurʾān*'.

In intervocalic and word-initial position, /ṭ/ and, to a lesser extent, /t/, are voiced. Emphasis is realized as pharyngealization with concomitant labialization. Labialization is one of a number of phonetic correlates of emphasis in several Arabic dialects. In Ṣanʿānī, however, labialization plays a more central role in the phonology of emphatics, and spreads within the word across several segments to target short high /i/ with a resulting realization of [u]. Examples include: *ṭawīl[u]h* 'long [sg. fem.]', *mustaṭīl[u]h*' lengthened [sg. fem.]'. Before /i/ and following /u/, labial and velar stops are often realized with labialization, as in: *ummʷī* 'my mother', *ḍubbʷī* 'fly', *šuggʷih* 'flat' (→ labiovelarization). In *rabbī* 'my Lord', /b/ is labialized, particularly in the vocative phrase *yā rabbʷī* 'oh my Lord!', forming a minimal contrast with *rabbī* 'bring [sg. masc.] up!'.

2.1.2 Vowels

Ṣanʿānī has three short vowels, /a, i, u/, and three long vowels, /ā, ī, ū/. The opposition between the short high vowels, /i/ and /u/, is maintained in most phonological contexts. Only in unstressed environments and in casual speech is the opposition not maintained. The long high vowels /ī/ and /ū/ have a closer articulation than that of their short counterparts, and /ā/ has a front articulation in all but emphatic environments.

2.1.3 Diphthongs

Ṣanʿānī has two diphthongs, /ay/ and /aw/, which are maintained in all phonological contexts. Phonetically, the /a/ is less open in /ay/ than in /aw/. The diphthongs are phonologically as long as long vowels or vC sequences.

Table 1. Consonant inventory

	labial	labiodental	interdental	dental-alveolar	palato-alveolar	velar	pharyngeal	laryngeal
plosive	b			t, d	j	k, g		ʾ
emph.				ṭ				
fricative		f	ṯ, ḏ	s, z	š	x, ġ	ḥ, ʿ	h
emph.			ḏ̣	ṣ				
nasal	m			n				
lateral				l				
tap				r				
glide					y	w		

2.1.4 Syllables

The syllable inventory includes one light, two heavy, and three superheavy syllable types (Table 2).

Table 2. Syllable inventory

light syllables	heavy syllables	superheavy syllables
Cv	CvC Cv̄	CvCC Cv̄C CvCCC/Cv̄CC

Light and heavy syllables can occur in any position; the superheavy syllable types CvCC and Cv̄C are always stem final; the superheavy syllable type Cv̄C may occur stem-finally when an *h*- or *n*-initial suffix is added (Watson 2002:69). The superheavy syllable type CvCC may occur stem-finally when the rightmost C is /t/ followed by an *h*-initial suffix.

kitābhā	‘her book’
baythum	‘their [masc.] house’
bābnā	‘our door’
absarthā	‘I/you [sg. masc.] saw her/it [sg. fem.]’
libisthā	‘I/you [sg. masc.] put it [sg. fem.] on’
xabazthā	‘I/you [sg. masc.] baked it [sg. fem.]’

The superheavy syllable type, CvCCC/Cv̄CC, is restricted to word-final position. This syllable type only ever surfaces in the specific derived environment of {2 sg. masc.}/{1 sg.} perfect verb + {negative}, essentially involving the concatenation of /t/ and /š/.

mā gambartš	‘I/you [sg. masc.] did not sit/stay’
mā libistš	‘I/you [sg. masc.] did not put on’
mā girītš	‘I/you [sg. masc.] did not read/learn’

Where the morphology otherwise potentially concatenates three consonants or a long vowel and two consonants and the suffix is not /h/ (or, in the case of a long vowel + two consonants, /n/) initial (as above), the consonant cluster CCC (v̄CC) is either broken up to CCvC (v̄CvC) by an epenthetic vowel [a]:

absart[a]nā	‘you [sg. masc.] saw us’
ḥubb[a]nā	‘our love’
bayt[a]kin	‘your [pl. fem.] house’

or a geminate consonant followed by the negative suffix /š/ is degeminated:

mā yiḥibb + š	*mā yiḥibš*	‘he does not like’
mā jarr + š	*mā jarš*	‘he did not take’

or, where the diphthong of the plural feminine suffix *-ayn* is followed by the negative suffix /š/, the diphthong is reduced to [a]:

mā + antayn + š	*mantanš*	‘not you [pl. fem.]’
mā + absartayn + š	*ma bsartanš*	‘you [pl. fem.] did not see’
mā + diriyayn + š	*mā diriyanš*	‘you [pl. fem.] did not know’

In Watson (2002:60), final /tš/ is analyzed as an affricate – the voiceless counterpart of /j/ – even though it is derived from two morphemes. Evidence for the analysis of /tš/ as a single consonant is adduced from the pronunciation of loanwords such as ‘jug’ as *tšāg* or *jāg* and ‘jelly’ as *tšīlī* or *jīlī*, and from the fact that pausal devoicing of /j/ results in [tʃ]. The sound [tʃ] is therefore not strange in Ṣanʿānī, providing a voiceless counterpart to the alveopalatal affricate /j/.

2.1.5 Stress

Ṣanʿānī is marked both by contextually fluctuating stress (Rossi 1939; Goitein 1934; Naïm-Sanbar 1994), and by a tendency to stress nonfinal syllables which end either in a long vowel or in the first half of a geminate (CvG). Stress fluctuates for expressive accent or emphasis, in prepause and postpause position, when the word is the second term in a genitive phrase, and after the definite article.

Where stress fluctuation fails to apply, or where words are pronounced in isolation, the following stress rules apply:

i. If the word contains Cv̄ or CvG syllable(s), stress the rightmost nonfinal syllable of this type, e.g. *ma'kātib* ‘offices’, *a'sāmī* ‘names’, *ba'sātīn* ‘gardens’, *'xārijīn* ‘going out [pl. masc.]’, *mit'axxirāt* ‘late [pl. fem.]’, *yi'ḥibbū* ‘they [masc.] love, like’, *'hākaḏahā* ‘like this’, *'sāfart* ‘I/you [sg. masc.] traveled’.

ii. Otherwise, stress a final Cv̄C/CvCC syllable, if there is one, e.g. *da'rast* 'I/you [sg. masc.] learned', *gam'bart* 'I/you [sg. masc.] sat', *diš'mān* 'rebel; poor [sg. masc.]', *ba'nāt* 'girls'.

iii. Otherwise, stress the rightmost nonfinal CvC syllable, if there is one, up to the antepenultimate, e.g. *'laflaf* 'he collected', *'maklaf* 'woman', *'madrasih* 'school', *'maklafih* 'his woman'.

iv. Otherwise, stress the leftmost Cv syllable, e.g. *'katab* 'he wrote', *'darasat* 'she learned', *'ragabatih* 'his neck', *mak'tabatī* 'my library'.

v. Stress final Cv̄ in disyllabic adjectives or di- or trisyllabic verbs in the imperfect, e.g. *yifʿa'lū* 'they [masc.] do', *ṭa'rī* 'fresh'.

2.2 *Phonotactics*

2.2.1 Assimilation

Ṣanʿānī is relatively conservative in terms of assimilation processes. Within a phonological word, a voiced plosive, affricate, or fricative is devoiced to the left of a voiceless consonant.

/wagt/	*wa*[k]*t*	'time'
/yudxul/	*yu*[t]*xul*	'he enters'
/nixbiz+hin/	*nixbi*[s]*hin*	'we bake them [fem.]'
/taḏkirih/	*ta*[θ]*kirih*	'ticket'

By contrast, a voiceless consonant only assimilates voice from a following voiced consonant under certain conditions. A voiceless stop is voiced before a voiced stop, but rarely before a voiced fricative; furthermore, voiceless coronal consonants are more likely to assimilate voice than voiceless velars. A voiceless fricative becomes voiced before a following fricative.

/yitgarṭaṭ/	*yi*[d]*garṭaṭ*	'he gets annoyed'
/xuṭbih/	*xu*[ᵭ]*buh*	'sermon'
/akbar/	*a*[g]*bar* ~ *a*[k]*bar*	'bigger; older'
/lafḏ̣/	*la*[v]*ḏ̣*	'expression; pronunciation'

The voiceless gutturals, /h/, /ʾ/, and /ḥ/, are not voiced before a voiced obstruent; however, they totally assimilate to a following voiced pharyngeal. Where a pharyngeal and a laryngeal are adjacent, total assimilation results in combining the voicing value of the rightmost guttural with the pharyngeal articulation of the stronger (pharyngeal) consonant.

/maʿhad/	*ma*[ħħ]*ad*	'institute'
/šibiḥ + hum/	*šibi*[ħħ]*um*	'he pulled them [masc.]'

The *-l* of the definite article assimilates to a following coronal plosive, sonorant, or fricative, namely /t, d, ṯ, ḏ, ḏ̣, s, š, ṣ, n, r, l/. As in Classical Arabic, *-l* does not assimilate to a following /j/.

2.2.2 Pausal phenomena

Ṣanʿānī exhibits → *ʾimāla* in pause, such that final /ā/ is realized as [eː]. Thus, /anā/ 'I' is realized prepausally as *an*[eː] and /absarnā/ 'we saw' as *absarn*[eː]. Where final /ā/ occurs in a word with a coronal emphatic, labialization characteristic of emphasis targets the resulting raised vowel to give /ō/. Thus, /ṣallaynā/ 'we prayed; washed' is realized as *ṣallayn*[oː].

The singular feminine ending *-ih*, which is realized with a lax mid vowel in pause, may well have originated from raising of /a/. Today, the morpheme is lexicalized as *-ih-* when a prepositional phrase with a pronoun suffix is suffixed to a singular feminine participle, the vowel of the morpheme being realized as a tense, high front [i], as in: /lābisih + lih/ = *lābisillih* 'dressing [sg. fem.] him' and /fāhimih + lih/ = *fāhimillih* 'understanding [sg. fem.] it [masc.]'.

A related phenomenon is diphthongization of long high vowels in pause, first noted by Jastrow (1984), and before a final consonant. Thus, /ṣābūn/ 'soap' is realized prepausally as *ṣābawn*, /migambirīn/ 'staying, sitting [pl. masc.]' as *migambirayn*, /gambarū/ 'they [masc.] stayed, sat' as *gambaraw*, /gambirī/ 'sit [sg. fem.]!' as *gambirey*. *ʾImāla* and prepausal diphthongization both involve a vowel shift from the periphery of the vocal tract to the central region, and can therefore be described as prepausal centralization.

A second pausal phenomenon discussed by Jastrow (1984) is glottalization of consonants. In the case of a sonorant (/n/, /l/, or /r/), glottalization often results in nonarticulation of the sonorant, as in: /samn/ = *samʾ* 'ghee', /fār/ = *fāʾ* 'mouse'. A final voiced pharyngeal is devoiced and postaspirated, rather than glottalized. Examples of pausal glottalization and aspiration in Ṣanʿānī can be found under ‘Pausalglottalisierung’ at http://semarch.uni-hd.de/.

2.2.3 Contextual phonological phenomena

Phonological phenomena within and across words which are not dependent on morphological information include the devoicing of geminate stops, intervocalic voicing of obstruents, and palatalization of stops before a palatal glide. Geminate stops are realized without voice. Thus, /ḥaggak/ ‘yours [sg. masc.]’ is realized as *ḥaggak* or *ḥakkak*, /ḏ̣ubbī/ ‘fly’ as [ð̴ub̥ːiː] or [ð̴upːiː], and /ḥajj/ ‘pilgrimage’ as [ħatːʃ]. Where a geminate stop results from an assimilation process, the resulting geminate is devoiced, as in /al-daymih/ ‘the kitchen’ = *addaymih* = [atːajmih].

All voiceless obstruents may be voiced intervocalically, particularly in casual speech; consonants most commonly affected by intervocalic voicing, however, are the alveolar stops /t, ṭ/.

Finally, weak → palatalization occurs in the dialect. In casual speech, a word-final stop may be palatalized before a word-initial palatal glide. Thus, /d/ in /walad/ ‘boy’ has a weak palatal off-glide before a word beginning in /y/. Strong palatalization (characteristic of some dialects spoken in the western mountain range of Yemen) and palatalization induced by a preceding high vowel /i/ do not occur in this dialect.

2.3 *Morphology*

2.3.1 Pronouns

2.3.1.1 Personal pronouns

Ṣanʿānī maintains the Classical Arabic gender distinction in the 3rd and 2nd person plural pronouns (Tables 3 and 4). Unlike some Yemeni dialects spoken in the southwestern mountain range, Ṣanʿānī makes no gender distinction in the 1st person singular.

Table 3. Independent pronouns (Ṣanʿāʾ)

	3rd pers.	2nd pers.	1st pers.
sg. masc.	*hū*	*ant*	*anā*
sg. fem.	*hī*	*antī*	
pl. masc.	*hum*	*antū*	*iḥnā*
pl. fem.	*hin*	*antayn*	

Table 4. Possessive pronouns

	3rd pers.	2nd pers.	1st pers.
sg. masc.	*-ih*/*-h*	*-ak*/*-k*	*-ī*
sg. fem.	*-hā*	*-iš*/*-š*	
pl. masc.	*-hum*	*-kum*	*-nā*
pl. fem.	*-hin*	*-kin*	

Object pronouns differ from possessive pronouns only insofar as the 1st person singular takes an initial *n*-, viz. *-nī*. The *n*- form of the 1st person singular pronoun is also suffixed to the preposition *fī* ‘in’: *fīnī* ‘in me’. Suffixes with initial /k/ or /n/ always, and suffixes with initial /h/ in careful speech, induce epenthesis ([u] before /kum/, otherwise [a]) when suffixed to a noun or verb ending in two consonants, as in /absart + kum/ = *absartukum* ‘I saw you [pl. masc.]’, /bint + nā/ = *bintanā* ‘our daughter’. Epenthesis occurs in careful speech on suffixation of a consonant-initial pronoun to a noun or verb ending in vCC or v̄C, as in /bayt + nā/ = *baytanā* ‘our house’, /bayt + kum/ = *baytukum* ‘your [pl. masc.] house’, /dīn + kum/ = *dīnukum* ‘your [pl. masc.] religion’.

2.3.1.2 Demonstrative pronouns

Demonstrative pronouns either occur independently or with a following or, less frequently, preceding noun. There are two sets of demonstrative pronouns: one which takes initial *hā*- (Tables 5 and 6), and one which does not. The demonstrative pronouns have a two-way number distinction, a two-way gender distinction in the singular, and a two-way distance distinction (near or far). Although there is no gender distinction in the plural demonstrative pronouns, a number of plural alternatives exist.

Table 5. Near demonstratives I

	singular	plural
masc.	*hāḏā*	*hāḏawlā* ~ *hawlā* *hāḏālā*
fem.	*hāḏī*	

In the *hā*-initial distal demonstratives, /ḏ/ may be emphasized in the singular masculine form.

Table 6. Distal demonstratives I

	singular	plural
masc.	*hāḏāk*	*hāḏawlāk*
fem.	*hāḏīk*	

The non-*hā* near and distal demonstratives are set out in Tables 7 and 8.

Table 7. Near demonstratives II

	singular	plural
masc.	*ḏayyā*	*hāḏawlāk*
fem.	*tayyih*	*ḏawlayyā* ~ *ḏawlāʾī*

Table 8. Distal demonstratives II

	singular	plural
masc.	*ḏayyāk*	*ḏawlayyak* ~ *ḏawlāk*
fem.	*tayyik*	*awlāk* ~ *awlāʾik*

2.3.1.3 Presentatives

Ṣanʿānī makes ample use of → presentative particles. The presentative particles include *inn, hā*, and *ḏā*. The latter two can function separately or combine with independent pronouns; the former takes a following noun or suffixed pronoun. The particle *inn* is most commonly introduced by *wa-*, as in *sāʿatayn ṯalāṯ w-innak fī tiʿizz* 'two or three hours and you're in Taʿizz'. The particle *hā(h)* in isolation usually has the imperative intention of 'take!'. The particle *ḏā* commonly combines with a preceding pronoun: *hūḏā* 'there it [masc.]/he is!', *hīḏā* 'there it [fem.]/she is!' A pronoun can additionally follow *ḏā*, as in *hūḏāhū* 'there it [masc.]/he is!'. The pronoun *hū* can function as a presentative in combination with following *ḏā* (+ pronoun), as in *hūḏanā šā-jī* 'here I am, I'm coming!', *hūḏaḥnā* 'here we are!', *hūḏantū* 'there you [pl. masc.] are!'. In the case of a following 3rd person plural pronoun, the pronoun is introduced with *la-*, as in *humḏāluhum* 'there they [masc.] are!', *hinḏālahin* 'there they [fem.] are!'. *Ḏā* with or without a following pronoun may also be suffixed to the particle *gad*, as in *gaḏḏā jīt* or *gaḏḏanā jīt* 'there, I've arrived!'. Distal presentative particles are derived by suffixation of /k/, as in *ḏakka* or *hūḏāk* 'there it [masc.]/he is [over there]!' and *ḏikkī* or *hūḏīk* 'there it [fem.]/she is [over there]!'. *Hūḏāk* is often realized with emphasis as *hūḏ̣āk*.

2.3.1.4 Relative pronoun

The relative pronoun is uninflected *allaḏī* or *allī*. As a tendency, speakers originating from the Old City favor *allaḏī*, while speakers from the suburbs of al-Bawniya and al-Qāʿ prefer *allī*. However, the same speaker may shift between the two forms within a single conversation. The relative pronoun usually relates to a definite noun, but for emphasis may relate to an indefinite noun, as in *gad maʿānā jāmiʿ fi l-ʿamal, ġurfi llī nṣallī fīhā* 'we have a mosque at work, a room that we pray in' (Watson and Al-ʿAmri 2000).

2.3.1.5 Interrogative pronouns

Interrogative pronouns occur clause-initially, although they may occur after the subject in topicalized questions. The majority of the interrogatives are reminiscent of those in Classical Arabic. They include *ayn* 'where?', *minayn* 'from where?', *ilayn* 'where to?' (*aynū* 'where is he?'), *man* 'who?' (*manū* 'who is it?'), *ayy* 'which?' (*ayyinū* 'which is it [masc.]?'), *kayf* 'how?' (*kayfū* 'how is he?'), *kayfant* or *kayfak* 'how are you [sg. masc.]?'), *lilmā* 'why?' (*lilmih* in pause), *mā* 'what?' (*māhū* 'what is it [masc.]?', *mih* in pause), *kam* 'how many?; how much?'. The pausal form of *mā* 'what?' functions as a tag question, sometimes preceded by *aw* 'or', as in *awmih* 'isn't it?'.

2.4 *Adverbs*

2.4.1 Temporal adverbs

The majority of temporal adverbs are similar to those of Classical Arabic. These include *alyawm* 'today', *ams* 'yesterday', *awwal ams* 'the day before yesterday', *alʿašī* 'this evening', *ḥīn* 'early', *awwal/min awwal* 'earlier', *ġudwuh* 'yesterday'. The word for now, *ḏalḥīn(ih)*, originates from *ḏā* + *alḥīn*. More innovative temporal adverbs include *ḏīksāʿ* 'at that time', from *ḏīk* + *assāʿah*, and *baḥīn* 'early', from *bi-* + *ḥīn*. 'Week' is most commonly *waʿd*, which is used adverbially in the expression *alwaʿd allawwal* 'last week'. The term *ḏirratayn* 'two specks' is used metaphorically in both the measure and temporal sense of 'a little'. More recently, *šwayyih* has come to function in the sense of 'in a little while'.

Expressions for the days of the week resemble those of Classical Arabic, with the exception of the terms for Tuesday and Wednesday:

yawm alaḥad ‘Sunday’, *yawm aliṯnayn* ‘Monday’, *yawm aṯṯalūṯ* ‘Tuesday’, *yawm arrabūʿ* ‘Wednesday’, *yawm alxamīs* ‘Thursday’, *yawm aljumʿah* ‘Friday’, *yawm assabt* ‘Saturday’.

Clock time is expressed adverbially as an annexion phrase by the older generation, e.g. *sāʿat xams* ‘5 o’clock’, but as a noun + predicate structure by the younger generation, e.g. *assāʿ(ah) xams*. In fast speech, the article may also be omitted.

2.4.2 Local adverbs

The main local adverbs are: *hānā* ‘here’, *hānāk* ‘there’, usually realized with emphasis, *hinīyih* ‘here’ and *hinayyik* ‘there’. According to Fischer (1969), the derivation of *hānā* is as follows: *hāhunā* > *hāhnā* > *hānā* ‘here’, and *hinīyih* ‘here’ was derived by suffixation of an emphatic element *-(ī)yih*. Other local adverbs include *yimān* ‘right’, *yisār* ‘left’, and several active participles, e.g. *ṭāliʿ* ‘up; upstairs’, *nāzil* ‘down; downstairs’, *dāxil* ‘inside’, *xārij* ‘outside’.

2.4.3 Manner adverbs

One set of manner adverbs are realized with or without the pronominal element *hā-*. These may take a following suffixal *-hā* for emphasis. They include *hākaḏā* ‘like this’, *hākaḏ̣āk* ‘like that’, *kaḏā* ‘like this’, *kaḏ̣āk* ‘like that’, *kaḏayyā* ‘like this’, *kaḏayyik* ‘like that’. Older speakers use *kaḏayyā, kaḏayyahā* and the more innovative forms *kaḏannanī, kaḏannahā* in the sense of ‘a little’, as in *iddālī ġarr kaḏayyahā* ‘he only gave me a little’. Other manner adverbs are derived from prepositional phrases and include *fīsāʿ* ‘quickly’ and *bisāʿat* + pronoun ‘immediately’.

2.4.4 Degree adverbs

The degree adverbs are *gawī* ‘very’, *xayrāt* ‘a lot’, less commonly *marrih* ‘very’. The adverb *gawī* can be repeated for emphasis. Other degree adverbs include *šwayyih* and *šwannanī* ‘a little’, *ġarr* and *bass* ‘only’, and words used in a metaphorical sense, such as *nār* ‘fire’, as in *ġālī nār* ‘really expensive’.

2.5 *Particles*

2.5.1 Genitive marker

The genitive marker *ḥagg* inflects for neither number nor gender, as in *arrajjāl ḥaggī* ‘my man/husband’, *albint ḥaggī* ‘my daughter’, *alkutub ḥaggī* ‘my books’. For emphasis, the genitive marker can be used to express familial relations. To emphasize possession, it may precede rather than follow the possessed noun. Whether the genitive marker is used or whether a possessed object forms the first part of an annexion phrase is often determined on stylistic or even rhythmic grounds. A phrase involving the genitive marker may be apposed to an annexion phrase, as in *alkaʿk ḥagg alʿīd…kaʿk alʿīd* ‘cake for the Eid…Eid cake’, and *yifʿalūlih samn fī gāʿatih…fi lgāʿah ḥakk aṣṣaḥn* ‘they put ghee on its base…in the base of the plate’ (Watson and Al-ʿAmri 2000).

2.5.2 Negative particles

The negative particles are *māšī* ‘no’, *miš* ‘not’, which usually negates the predicate; *mā* or *lā*, which are used in absolute negation; and the discontinuous *mā…-š*, which may also occur without final *-š* and negates verbs and prepositional phrases with pronominal suffixes, as in *mā maʿakš* ‘you [sg. masc.] do not have’ and *mā yiʿjibnī hāḏā* ‘I don’t like that’. The particle *māšī* functions additionally in the sense of ‘otherwise’, as in *māšī mā ʿād tilsāš alxubzih* ‘otherwise, the bread won’t stick’. The compound particle *wa-lā* expresses emphatic negation, as in *wa-lā šī* ‘nothing at all!’; *lā/mā…wa-lā* has the sense of ‘neither…nor’; and *mā…illā wa-* either conveys the sense of ‘by the time’, as in *mā yijī fawg almāyidih illa w-gadū šābiʿ* ‘by the time it reaches the table, he’s full’, or ‘only’, as in *mā yixazzin illā w-bih gāt sawā* ‘he’ll only chew if there is good qat’.

2.5.3 Existentials

The main existential particle is *bih* ‘there is’, negated as *mā bišš/biš*. In conditionals and questions, *šī* is more common, as in *lā šī xubz* ‘if there is any bread’, *šī gāt* ‘is there any qat?’ In questions, *šī* may be combined with *bih* and/or another prepositional phrase, as in *šī bih (maʿākum) gāt* ‘do you [pl. masc.] have any qat?’

2.5.4 Prepositions

The main prepositions in Ṣanʿānī are *min* ‘from’, *lā* ‘to’, *fī* ‘in’, *bi* ‘in; with’, *maʿ* ‘with’, *ʿind* ‘with; at the house of’, *fawg* ‘over’, *taḥt* ‘under’, *gafā* ‘behind’, *ṣalā* ‘toward’, *tijāh* ‘before; in front of’, *šigg* ‘next to’, *sāʿ* ‘like’, *gabl* ‘before’, and *baʿd* ‘after’. In contrast to many other dialects, possession is indicated through *maʿ* ‘with’ rather than *ʿind*, as in *maʿī sayyārih* ‘I have a car’, *kam maʿiš banāt* ‘how many daughters do you [sg. fem.] have?’.

2.5.5 Conjunctions, conjuncts, subordinators

The main conjunctions are *wa-* ‘and’, *aw/awlā/wallā* ‘or’, and *fa-* ‘then; so’. The latter is common in narratives, particularly when the narrator is deliberating. The conjunct *lākin* ‘but; however’ is pronounced as *laykin* by Old City speakers. Ṣanʿānī has a rich set of subordinators, of which probably the most common is *law-mā* ‘when; until’. Other subordinators include *lā* ‘if’, *ḥīn* ‘since’, *ʿalā sibb* ‘because’, *tijāh-mā* ‘before’, *sāʿ-mā* ‘like’. Subordinators and subordination are presented and discussed in detail in Watson (1993).

2.6 *Nouns and adjectives*

Nouns are either feminine or masculine. Feminine nouns without the feminine ending *-ih* include parts of the body which come in pairs; inherently female objects, as in *umm* ‘mother’, *bint* ‘girl; daughter’; and other terms, including *ḥarb* ‘war’, *šams* ‘sun’. Nouns and adjectives take the sound masculine or feminine plural, *-īn* and *-āt*, or one of a large number of broken plural patterns (Watson 2002:164ff.). The dual suffix, *-ayn*, is usually attested only for weight, measurement, or time, as in *yawmayn/yawmatayn* ‘two days’, *giršayn* ‘two riyals’, *sāʿatayn* ‘two hours’, and in the frozen phrase *wāliday* + pronoun ‘his, etc. parents’. In other cases, the appropriate form for ‘two’ precedes the noun in the plural, as in *ṯintayn banāt* ‘two girls’, *iṯnayn ʿiyyāl* ‘two boys’. Even paired parts of the body are referred to through the plural, with *iṯnayn/ṯintayn* added optionally for emphasis, as in *arjulī* ‘my legs’ and *arjulī aṯṯintayn* ‘both my legs’.

Diminutive patterns occur in place-names and personal tribal names throughout Yemen, such as *alzubayrī, alḥubayšī, aljunayd*, but the diminutive is not productive. The only adjective with a diminutive morphological pattern is *zuġayrī/zuġayyirī* ‘small’. In order to express smallness and affection, *-ī* can be suffixed to personal names, as in *ixlāṣī* ‘little Ikhlas’, *šahābī* ‘little Shahab’. Another function of *-ī* is to derive singulatives from collectives, as in *mawzī* ‘banana’ from *mawz, dūdī* ‘worm’ from *dūd*.

2.7 *Numerals 1–20*

The cardinal numerals for one and two are: *wāḥid, iṯnayn*, and from three to ten: *ṯalāṯih, arbaʿah, xamsih, sittih, sabʿah, ṯamāniyih, tisʿah, ʿašarih*. In numeral phrases, these forms combine with the plural of a masculine noun, and shortened forms with the plural of a feminine noun. The numerals from eleven to nineteen have two allomorphs, depending on whether they take a following noun or occur in isolation. Thus, *xamstʿašar bint* ‘fifteen girls’ compares with *xamstʿāš* ‘fifteen’, in which final *-ar* is not pronounced.

2.8 *Verbs*

2.8.1 Verbal Forms

The Form I verb takes one of three possible vocalisms, *faʿal, fiʿil, fuʿul*. The majority of *fiʿil* verbs are intransitive or grammatically transitive, denoting actions whose consequences are experienced by the agent, ‘agent moyen’. Exceptions include *šibiḥ* ‘to pull’, *libis* ‘to wear’. The *fuʿul* verbs have either a root emphatic or velar consonant, as in *kubur* ‘to grow up’, *gurub* ‘to approach’, *ruḏ̣ī* ‘to want’. The derived Forms of the triliteral verb in the dialect are II, III, IV, V, VI, VII, VIII, and X. The function of original Form IX is taken over by Form II, as in *ḥammar* ‘to be/become red’, *sawwad* ‘to be/become black’. Form II is the most productive verb form alongside the quadriliteral verb. Form IV is comparatively common in Ṣanʿānī; examples include *absar* ‘to see’, *ixlaṣ* ‘to take off’. Forms V and VI take initial *t-, ti-*, or *ta-*, as in *tilaṯṯam* ‘to wear a face veil’, *taḥākā* ‘to talk’. In the imperative, however, Forms V and VI take *it-*, as in *itḥākā* ‘talk [sg. masc.]!’ Form VII is relatively rare; examples include *ingaḍā* ‘to pass [of time]’. Form VIII is common.

Quadriliteral verbs are both common and productive in Ṣanʿānī. They are derived either from biliteral or triliteral verbs through reduplication, as in *dagdag* ‘to knock [continuously]’, *šamšam* ‘to sniff’, or infixation of a sonorant, as in *ḍalḥak* ‘to laugh [a lot]’ (cf. *ḍaḥak* ‘to laugh’), *tinagwal* ‘to move [from place to place]’ (cf. *nagal* ‘to move’); from nouns, as in *saynam* ‘to go to the cinema’ (cf. *sīnamā* ‘cinema’); or from adjectives, as in *timajnan* ‘to become mad’ (cf. *majnūn* ‘mad’), *tibayxal* ‘to be mean’ (cf. *baxīl* ‘mean’).

The internal passive, *u-i*, exists for a few verbs, including *kumil* 'to be finished', *wulid* 'to be born', *surig* 'to be stolen (from)', *summī* 'to be called', *wujid* 'to be found'. The passive voice may also be expressed by a Form VIII verb, as in *iḥtabas* 'to be imprisoned', but there is no evidence that this form is morphologically productive in Ṣanʿānī.

2.8.2 Inflection of perfect

In the perfect, sound verbs take the suffixes in Table 9.

Table 9. Perfect verb suffixes

	3rd pers.	2nd pers.	1st pers.
sg. masc.	-	*-t*	*-t*
sg. fem.	*-at*	*-tī*	*-t*
pl. masc.	*-ū*	*-tū*	*-nā*
pl. fem	*-ayn*	*-tayn*	*-nā*

Doubled verbs take an infix *-ay-* before a consonant-initial suffix, as in *šall-ay-t* 'I/you [sg. masc.] took'. Final weak verbs ending in *-ī* take the suffixes given above, e.g. *nisiy-at* 'she forgot', *ruḍiy-ayn* 'they [fem.] wanted'. In final weak verbs ending in *-ā*, however, vowel-initial suffixes take other forms which attach directly to the stem: the 3rd person singular feminine suffix is realized as *-it*, as in *awf-it* 'she finished', *mall-it* 'she filled'; the 3rd person plural masculine suffix is realized as *-aw*, as in *mall-aw* 'they [masc.] filled'; and the 3rd person plural feminine suffix is realized as *-ayn*, as in *mall-ayn* 'they [fem.] filled'.

2.8.3 Inflection of imperfect

In the imperfect, the sound verbs take the affixes in Table 10.

Table 10. Imperfect verb affixes

	3rd pers.	2nd pers.	1st pers.
sg. masc.	*yi-*	*ti-*	*a-*
sg. fem.	*ti-*	*ti-…ī*	
pl. masc.	*yi…ū*	*ti…ū*	*ni-*
pl. fem.	*yi…ayn*	*ti…ayn*	

The long-vowel imperfect suffixes *-ī*, *-ū* are realized in final-weak verbs as diphthongs *-ay*, *-aw*, as in *tišt-ay* 'you [sg. fem.] want', *titḥāk-aw* 'you [pl. masc.] speak', *yirḍ-aw* 'they [masc.] want'. In most quadriliteral verbs and the triliteral derived verbs of Forms II, III, IV, VII, VIII, and X, the final imperfect stem vowel is *-i-*, as in *yixazzin* 'he chews [qat]', *yijābir* 'he chats', *yibsir* 'he sees', *yinšaġil* 'he is occupied', *yištaġil* 'he works', *yistaʿmil* 'he uses'.

On suffixation of object pronouns or the negative suffix to feminine plural verbs, the ending *-ayn* is reduced to *-an* with gemination of *-n-* before an object pronoun, as in /absartayn + ih/ = *absartannih* 'you [pl. fem.] saw him', /yištayn + iš/ = *yištanniš* 'they [fem.] want you [sg. fem.]', /mā yibsirayn + š/ = *mā yibsiranš* 'they [fem.] don't see'.

2.8.3.1 Imperfect tense/aspect markers

Ṣanʿānī has two tense/aspect markers: one denoting continuous/habitual aspect and one future tense (Table 11). In both cases, the 1st person singular prefix differs from that for other persons, and has two allomorphs for the future prefix.

Table 11. Imperfect tense/aspect markers

person/gender	continuous/habitual	future
1st sg.	*bayn-*	*šā-/ʿad-*
other persons	*bi-*	*ʿā-*

On prefixation of *šā-*, *a-* of the 1st person singular imperfect prefix is not pronounced, e.g. *šā-sīr* 'I will go', *šā-bsir* 'I will see'. The prefix *ʿad-* differs from *šā-* by conveying that the speaker will do something at some time, while *šā-* indicates that the speaker will do something at a definite time.

Bibliographical references

Behnstedt, Peter. 1985. *Die nordjemenitischen Dialekte. I. Atlas*. Wiesbaden: L. Reichert.

——. 1992. *Die nordjemenitischen Dialekte*. II/1. *Glossar Alif – Dāl*. Wiesbaden: L. Reichert.

——. 1993. *Glossar der jemenitischen Dialektwörter in Eduard Glasers Tagebücher (II, III, VI, VII, VIII, X)*. Vienna: Verlag der Österreichischen Akademie der Wissenschaften.

——. 1996. *Die nordjemenitischen Dialekte*. II/2. *Glossar Ḏāl – ʿAyn*. Wiesbaden: L. Reichert.

——. 2006. *Die nordjemenitischen Dialekte*. II/3. *Glossar Fāʾ – Yāʾ*. Wiesbaden: L. Reichert.

Fischer, Wolfdietrich. 1969. *Die demonstrativen Bildungen der neuarabischen Dialekte*. The Hague: Mouton.

Goitein, Shlomo Dov. 1934. *Jemenica: Sprichwörter und Redensarten aus Zentral-Jemen*. Leiden: E.J. Brill.

——. 1970. *Jemenica: Sprichwörter und Redensarten aus Zentral-Jemen mit zahlreichen Sach- und Worterläutungen*. Leiden: E.J. Brill.

Ḥibšī, ʿAbdallāh Muḥammad al- (ed.). 1991. *Ḥawliyyyāt yamaniyya: al-Yaman fī l-qarn at-tāsiʿ ʿašar al-mīlādī*. Sanaa: Dār al-Ḥikma.

ʾIriyānī, Muṭahhar ʿAlī al-. 1996. *al-Muʿjam al-yamanī fī l-luġa wa-t-turāṯ*. Damascus.

Jastrow, Otto. 1984. "Zur Phonologie und Phonetik des Ṣanʿānischen". *Entwicklungsprozesse in der Arabischen Republik Jemen*, ed. Horst Kopp and Günther Schweizer, 289–304. Wiesbaden: L. Reichert.

Naïm-Sanbar, Samia. 1994. "Contribution à l'étude de l'accent yéménite: Le parler des femmes de l'ancienne génération". *Zeitschrift für Arabische Linguistik* 27.67–89.

Piamenta, Moshe. 1990–1991. *Dictionary of post-Classical Yemeni Arabic*. Leiden: E.J. Brill.

Rossi, Ettore. 1939. *L'arabo parlato a Ṣanʿāʾ*. Rome: Istituto per l'Oriente.

Watson, Janet C.E. 1993. *A syntax of Ṣanʿānī Arabic*. Wiesbaden: O. Harrassowitz.

——. 1996. Ṣbaḥtū! *A course in Ṣanʿānī Arabic*. Wiesbaden: O. Harrassowitz.

——. 2002. *The phonology and morphology of Arabic*. Oxford: Oxford University Press.

——. and ʿAbd al-Salām Al-ʿAmri. 2000. Waṣf Ṣanʿāʾ: *Texts in Ṣanʿānī Arabic*. Wiesbaden: O. Harrassowitz.

Watson, Janet C.E. and Abd al-Rahman Mutahhar. 2002. *Social issues in popular Yemeni culture*. Sanaa: al-Sabahi Press.

Janet Watson (University of Salford)

Sandhi

The term *sandhi* (from Sanskrit *saṃdhi* 'putting together') refers to phonological processes that apply when two morphemes (roots, stems, or affixes) are juxtaposed. This very general term has been used in a variety of senses. In its broadest sense, sandhi includes the selection of alternative forms (allomorphs) at the joining of two morphemes within a word (internal sandhi) as well as processes that apply between adjacent words in a phrase (external sandhi), and in this broad sense it encompasses a large part of the phonology of a language. Sandhi includes both automatic phonetic processes (including, for some linguists, allophonic patterns) and alternations that are morphologically or lexically conditioned. It would exclude only those patterns (morpheme structure conditions) that constrain the makeup of individual morphemes, such as the fact that Arabic triconsonantal roots frequently have identical second and third consonants (as in *r-d-d* in *radda* 'he returned', *rudūd* 'returns [noun, pl.]', but not identical first and second consonants (**r-r-d*); nearly all other phonology could be called sandhi. Because of the breadth of this term, many linguists use it to refer only to external sandhi, which is the phonology of the phrase, rather than of the word, and it is therefore also called postlexical phonology, meaning those phonological processes that apply only after a completely formed, whole word is inserted into a phrase.

What is significant about sandhi in all its senses is the fact that it depends on constituents and their boundaries, and therefore it marks or indicates those boundaries. Sandhi facilitates the hearer's identification of morphemes within a word or of the boundaries between phrases within a sentence. Thus, in Classical Arabic the alternation between the full and pausal forms of words helps mark phrase boundaries; for example, in *madīnatun kabīrah* 'a big city', the *-h* in the pausal form *kabīrah* (as opposed to the full form *kabīratun*, which occurs in other situations) marks it as being at the end of a phonological phrase, while the *-tun* ending of *madīnatun* marks it as being *not* phrase-final. There may or may not be audible silence after either of these words, but the appearance of a full or a pausal form shows where the phrase boundary is. Within a word, too, sandhi can help define boundaries and categories. Probably the most familiar sandhi process of Arabic applies to the definite prefix *al-* (*il* in many dialects), the *l* of which assimilates to a following coronal (i.e. dental, alveolar, or postalveolar) consonant, as in *al-ṭifl* 'the child', pronounced *aṭṭifl*. This assimilation does not apply between separate words (*bal ṭifl* 'but rather a child', not **baṭṭifl*), nor within a word (*ʾalṭaf* 'nicer', not **ʾaṭṭaf*). One could view the assimilation of the article as a mere phonological process (a helpful one, facilitating pronunciation, or an unfortunate one, obscuring the basic form of the prefix), but on the other hand, one could view it as helping to identify the article as opposed to other sequences. Linguists who use the word sandhi tend to take the latter view and see sandhi alternations as signs of constituent boundaries.

The vernacular Arabic of the Syrian dialect area provides an especially clear illustration of sandhi as a marker of a grammatical boundary. In these dialects, there is a difference between *msikna* 'we held' and *misikna* 'he held us'. Both words consist of a stem meaning 'to hold' and a suffix *-na* '1st person plural', with the stem taking different forms depending on whether the suffix marks a subject or an object: *msik* when followed by a consonant-initial subject suffix, *misik* before a consonant-initial object suffix or no suffix (and *misk* before a vowel-initial suffix). Within the word, there is a constituent consisting of the verb stem and a subject suffix, if any, while an object suffix lies outside that constituent, which can be marked with parentheses: *(msik-na)* 'we held', *(msik-t)-na* 'you [masc. sg.] held us' (i.e. '(held-you)-us'), versus *(misik)-na* 'he held us' (Brame 1973). Similarly, there are sandhi alternations in the stem of vernacular Arabic *(zur-na)* 'we visited', *(zur-t)-na* 'you [masc. sg.] visited us', versus *(zār)-na* 'he visited us', and in Classical Arabic *(raʾay-nā)* 'we saw', *(raʾay-ta)-nā* 'you [masc. sg.] saw us', versus *(raʾā)-nā* 'he saw us'. The sandhi alternations between *msik* and *misik,* between *zur* and *zār*, and between *raʾay* and *raʾā* inform the hearer whether the *-nā* is inside or outside the bracketed unit, marking a subject or an object.

A large part of any reference grammar of Arabic is occupied by the description of internal sandhi, that is, the different forms that morphemes take in different contexts. Therefore, the remainder of this entry focuses on external sandhi, processes that apply between words and at the attachment of the pronominal enclitics that mark the objects of verbs and prepositions and the possessors of nouns, especially some that are of particular interest.

Many sandhi processes serve to maintain the preferred structure of Arabic syllables. In Classical and Modern Standard Arabic, every syllable has an onset consisting of one consonant, a short or long vowel, and possibly a coda of one consonant, making three types, Cv, Cv̄, and CvC, plus the much more restricted Cv̄C. In addition, a word in its pausal form (at the end of a phrase) may end in two consonants (CvCC). There can never be three adjacent consonants (which the Arabic grammarians call *iltiqāʾ as-sākinayn* 'the meeting of two vowelless consonants'), nor two adjacent vowels. There are four phonological processes that apply automatically to alleviate violations of these syllable structure restrictions.

i. → Syncope: every word-initial vowel is deleted, except at the beginning of a phrase: *hāḏihi* 'this' + *al-madīna* 'the city': *hāḏihi lmadīna* 'this city', pronounced *hāḏihilmadīna*, with → resyllabification creating a syllable *-hil-* which begins in the first word and ends in the next; *baʿda* 'after' + *ibtidāʾ* 'a beginning' becomes *baʿda btidāʾ*, pronounced *baʿdabtidāʾ; fa-* 'so' + *uktub* 'write [imper.]!', pronounced *faktub*.
ii. A long vowel in a closed syllable is shortened (the same happens within a word as well): *fī* + *al-madīna* 'the city', pronounced *filmadīna*, with the long vowel of *fī* shortened.
iii. Where three consonants would come together, the process of → epenthesis applies, inserting the vowel *i* between the first and second of the three: *kānat al-madīna* becomes *kānatilmadīna* 'the city was'. Somewhat surprisingly, in a sequence like *fī al-ibtidāʾ* 'in the beginning', the first vowel shortens despite the presence of a vowel after the following consonant: *filibtidāʾ*, not **fīlibtidāʾ* (Wright 1896:21), and epenthesis applies, likewise surprisingly, in a sequence like *ʿan al-ibtidāʾ* 'about the beginning', pronounced *ʿanilibtidāʾ*, not **ʿanlibtidāʾ*.
iv. Another epenthesis process inserts a glottal stop ʾ when a vowel-initial word appears at the beginning of a phrase: *ʾalmadīna, ʾibtidāʾ*.

These four are default processes, in that they apply without any morphological or lexical restrictions. In certain morphological environments, however, vowels other than *i* are inserted. After the masculine plural suffixes *-hum, -kum*, and *-tum*, the epenthetic vowel is *u: katabtum* 'you [masc. pl.] wrote' + *ad-dars* 'the lesson': *katabtumuddars*. Between the preposition *min* 'from' and any word beginning with the vowel *a* (these are just the definite article *al-*, the forms of the relative pronoun *allaḏī*, and the word *aḷḷāh* 'God'), the epenthetic vowel is *a*: *min dars* 'from a lesson', *minibtidāʾ* 'from a beginning', but *minaddars* 'from the lesson', *minaḷḷāh* 'from God'. Furthermore, it is possible to consider the initial *i* of words like *ism* 'name', *ibtidāʾ*

'beginning', and *ifham* 'understand [imper.]!' to be epenthetic, and if so, then there is also epenthetic *u* in imperatives of verbs with the stem-vowel *u*, such as *uktub* 'write!', and in the perfect passive of verbs like *ustuxrija* 'it was extracted'; the rule is that the epenthetic vowel is *u* in verb forms with the stem vowel *u*, and *i* otherwise. Another morphologically conditioned alternation that preempts the default syllable-structure processes applies optionally to the 1st person singular enclitic *-(n)ī*, which may become *-(n)iya* if the following word begins with two consonants, so that *bayt-ī* 'my house' + *al-jadīd* 'new' can be pronounced *baytiyaljadīd*; if not for this, the vowel would shorten by the default rule (ii), producing *baytiljadīd*, which would obscure the enclitic entirely. (In Modern Standard Arabic, the ambiguity of *baytiljadīd* can be avoided by pronouncing 'my new house' with a tense [i], as opposed to the lax [ɪ] in [bajtɪlʤadiːd] 'the house [gen.] of the new one'.)

While in Standard Arabic epenthesis inserts a vowel between the first and second consonants of three in contact (CCC > C*i*CC), the modern vernacular dialects vary in the place of the epenthetic vowel. Watson (2007) has provided a detailed typology and analysis of this and related phenomena. Watson illustrates the main types with the word for 'I said to him'; the Standard Arabic phrase *qultu la-hu* has become a single word in modern vernacular Arabic, with a basic form reconstructable as *qultlahu*, containing three adjacent consonants: *ltl*. By the basic principles of Arabic syllable structure, which permit only one consonant in an onset and one in a coda, the *t* in this sequence cannot be part of either the preceding or the following syllable: *(qul)t(la)(hu)*. In Baghdadi Arabic, an epenthetic vowel is inserted before the unsyllabified *t: gilitla*, and hence this is labeled a vC dialect (vowel-consonant, for the sequence *it*). In Cairene Arabic, a Cv dialect, the epenthetic vowel is inserted after the *t: ˀultilu*. The Arabic of Ṣanʿāˀ, with *gultalih*, is in this respect like Cairene Arabic. In Moroccan Arabic, the unsyllabified consonant is tolerated, *qiltlu*, and hence this type of dialect is termed a C dialect. Several other properties tend to go along with each type. vC dialects like Baghdadi Arabic permit consonant clusters at the beginning of a phrase (*člāb* 'dogs', *b-báġdād* 'in Baghdad'), but do not allow phrase-final clusters (/čalb/ 'dog' is pronounced *čalib*). Cairene Arabic is the opposite, with unlimited final clusters (*baṭn* 'belly'), but no initial clusters. Many vC dialects also exhibit → metathesis of medial CC*i*C to C*i*CC (/jiktib-u/ → *jikitbu* 'they write'), and in some this goes along with an opaque stress pattern (*yíkitbu*, rather than *yikítbu*). Metathesis and opaque stress do not occur in Cairene Arabic, the prototypical Cv dialect. Cairene Arabic, for its part, shortens long vowels when they occur in closed syllables (/bāb-na/ *babna* 'our door', /kātib-a/ *katba* 'writing [fem. sg.]'), while vC dialects retain long vowels in closed syllables (*bābna, kātba*). Ṣanʿāni Arabic and a number of others form an intermediate type, in that they have Cv epenthesis like Cairene Arabic but otherwise have the characteristics typical of vC dialects like Baghdadi Arabic. Watson provides an elegant phonological analysis that ties together these apparently disparate properties. Standard Arabic is in most respects a Cv dialect like Cairene Arabic, with phrase-final consonant clusters (in pausal forms like *kalb*), no initial clusters (/btidāˀ/ *ˀibtidāˀ*), and shortening of long vowels in closed syllables (/ḥtāj-nā/ *ˀiḥtajnā* 'we needed'). However, as in vC dialects, there is metathesis (/yatmimu/ *yatimmu* 'it is accomplished'). There is no word-internal epenthesis, because the morphology does not produce sequences of three consonants.

A sandhi process that is nearly universal in modern vernacular Arabic is the apparent lengthening of a vowel before a suffix. This applies both to stem vowels (Palestinian Arabic *nisi* 'he forgot', *nsīna* 'we forgot', *nisī-na* 'he forgot us', *ma nisī-š* 'he did not forget') and to suffix vowels (*zāru* 'they visited', *zārū-na* 'they visited us', *ma zārū-š* 'they did not visit', *ma zārū-nā-š* 'they did not visit us'; *kunti* 'you [fem. sg.] were', *ma kuntī-š* 'you [fem. sg.] were not'). This raises the question of the directionality of the process, synchronically in modern Arabic: are these vowels underlyingly short and lengthened before a suffix, or are they underlyingly long and shortened in word-final position? McCarthy (2005) has argued that these vowels must be underlyingly long, both on formal theoretical grounds and from the fact that, while unstressed vowels are often shortened in Arabic and other languages, preaffixal lengthening is apparently unknown. Thus, McCarthy argues that /zārū/ is shortened to *zāru*, except when a suffix follows.

The historical origin of this alternation is not uniform. Some of the alternating vowels are long in Classical Arabic: *zārū* 'they visited', *zārū-nā* 'they visited us'. When these forms are pronounced with the patterns of stress and vowel shortening that are widespread in the modern vernaculars, they appear to alternate in a manner identical to the vernacular alternation: *zā́ru, zārū́na*. However, there is no other evidence that they were pronounced in that way in ancient times. Other vowels that alternate in the modern vernaculars were short in Classical Arabic both when final and when followed by a suffix: *zurti* 'you [fem. sg.] visited', *zurti-nā* 'you visited us'. (Although forms like *zurtī-nā*, with a long *ī*, are attested in Classical Arabic, they are not the norm; see Wright 1896:102). In only one situation does Classical Arabic show a regular alternation between short and long vowels that is similar to the modern alternation: in the past tense after the 2nd person masculine plural suffix *-tum*. Before a word beginning with two consonants, the epenthetic vowel after *-tum* (as mentioned above) is *u: samiʿtumu d-dars* 'you [masc. pl.] heard the lesson'. But before a pronominal suffix, *-tum* is extended by the addition of a long *ū: samiʿtumū-nā* 'you [masc. pl.] heard us'. Thus, there is a three-way alternation among *-tumu* (before a consonant cluster in the next word), *-tumū* (before a pronominal suffix), and *-tum* (otherwise). It should not be assumed that this particular suffix was the seed that brought forth the pervasive alternation in modern Arabic, but rather that both Classical Arabic and vernacular Arabic reflect an earlier stage in which some suffixes alternated in this manner, although we do not know which ones, nor when and where.

An extremely unusual and intriguing sandhi process occurs in Nigerian Arabic (Owens 1998:22–26). The past tense suffix *t* marking the 1st person and 2nd person masculine singular, as in *mašē-t* 'I went', *jib-t-uhum* 'I brought them', is invariably deleted whenever it would appear at the end of a word after a consonant; thus, /taxallas-t/ 'finish-I' is pronounced *taxallás*, in which the stress on the last syllable is a consequence of the underlying /t/ and itself indicates that this is a 1st or 2nd person singular form. There is one systematic exception: when such a verb is followed by a direct object marked with the definite prefix *al-*, the *t* is not deleted; Owens gives the example *xallas-t al-enwayessi* 'I finished the NYSC [youth corps]'. Phonologically identical sequences in which the *al-* is not the definite prefix, or the noun is not the direct object, do not preserve the *t*: *taxallás altahág* 'I finished [and] I joined', *katáb al-inta dawwart-a* 'I wrote what you [masc. sg.] wanted (it)', *inta gul al-muftá indak* 'you [masc. sg.] said the key was with you [masc. sg.]', nor does a direct object with *al-* that does not immediately follow the verb, as in *katáb bas al-maktūb* 'I wrote only the book' (Owens 1998:25). Owens argues that what preserves the *t* is that the definite prefix of a direct object is cliticized to the preceding verb. The enclitic article, like suffixes, prevents the elision.

Many other sandhi processes occur in Classical and vernacular Arabic, including consonant → assimilations, stress shifts, vowel changes, vowel shortening, devoicing or glottalization of final consonants and vowels, and the spreading of 'emphasis' (pharyngealization and backed articulation). Some of these are examined in detail in Watson (2002:226–267, 273–286).

BIBLIOGRAPHICAL REFERENCES

Brame, Michael K. 1973. "On stress assignment in two Arabic dialects". *A festschrift for Morris Halle*, ed. Stephen R. Anderson and Paul Kiparsky, 14–25. New York: Holt, Rinehart and Winston.

McCarthy, John J. 2005. "The length of stem-final vowels in Colloquial Arabic". *Perspectives on Arabic linguistics*, XVII–XVIII, ed. Mohammad T. Alhawary and Elabbas Benmamoun, 1–26. Amsterdam and Philadelphia: J. Benjamins.

Owens, Jonathan. 1998. *Neighborhood and ancestry: Variation in the spoken Arabic of Maiduguri, Nigeria*. Amsterdam and Philadelphia: J. Benjamins.

Watson, Janet C.E. 2002. *The phonology and morphology of Arabic*. Oxford: Oxford University Press.

——. 2007. "Syllabification patterns in Arabic dialects: Long segments and mora sharing". *Phonology* 24.335–356.

Wright, William. 1896. *A grammar of the Arabic language*, I. 3rd ed. Cambridge: Cambridge University Press. (Repr., 1971.)

ROBERT D. HOBERMAN (Stony Brook University)

Ṣarf

1. DEFINITION

Ṣarf, originally meaning 'shifting a thing from one state or condition to another' (Lane 1863–1893:II, 1680; Ibn Manẓūr, *Lisān* IV, 2434), is used in Arabic grammar as the technical term for morphology. It is linked with *taṣrīf*, which also has to do with change, and originally meant 'the turning of the winds from one state or condition, to another' (Lane 1863–1893:II, 1681; Ibn Manẓūr, *Lisān* IV, 2435). In relation to language, the science of *taṣrīf* is usually called *ʿilm aṣ-ṣarf*. Both indicate a change in the form of words, and both are used indiscriminately to designate the science of morphology.

Modern linguists sometimes distinguish between *ṣarf* as the study of the structure of words with their morphemes (cf. Hindāwī 1989:20ff.) and *taṣrīf* as the study of the → *ištiqāq* 'derivation' of words from a specific → *ʾaṣl* 'root' or base form and the different phonological rules that affect their form. In a still wider sense, *ṣarf* includes *taṣrīf* and is applied to the science of morphology in general (cf. Åkesson 2001:40–42). In the field of Arabic grammar (see Troupeau 1993), *ṣarf* is generally used rather than *naḥw* 'syntax'. While the latter denotes the syntactic position of the word as indicated by its ending (cf. Ibn Jinnī, *Munṣif* I, 4) – whether it is *muʿrab* 'declinable', depending on an operator, e.g. the agent/topic (voweled by a *ḍamma*), the object (voweled by a *fatḥa*), or the genitive (voweled by a *kasra*), or *mabnī* 'undeclinable' (for both terms, see Sībawayhi, *Kitāb* I, 1ff.; Ibn Jinnī, *Xaṣāʾiṣ* I, 35ff.; Bohas a.o. 1990:53–55) – there is one aspect of *ʾiʿrāb* 'declension' that refers to the morphological aspect of declension, namely the distinction between complete and incomplete declension of the word (cf. ʿUkbarī, *Masāʾil* 102–105; Carter 1981:37–38; Åkesson 2001:130). Indeed, one meaning of *muʿrab* 'declined' is synonymous with *munṣarif* 'fully declined', which is connected with *ṣarf* when it is applied to the inflected noun according to Sībawayhi's (d. 177/793) theory in the *Kitāb*. According to him, a noun is fully declinable when it can have the three case endings, nominative, accusative, and genitive, or when it can receive the *tanwīn* (cf. Versteegh 1995:173–174).

2. Development of *ṣarf*

Sībawayhi was one of the first grammarians to combine *ṣarf* and *naḥw*, while still distinguishing them from each other. He introduces the science of morphology with these words: "This is the section about what the Arabs formed of nouns, adjectives, and verbs that are sound and unsound...and what the grammarians name *taṣrīf* and *fʿl* [i.e. forms made to correspond to the derived measures of *fʿl*]" (*Kitāb* II, 343). The main issues studied in the second part of the *Kitāb* became the classical topics in the works on morphology.

Among those who developed the study of morphology as an independent discipline, the following grammarians may be mentioned: ʿAlī ibn al-Ḥasan al-ʾAḥmar (d. 194/809), al-Farrāʾ (d. 207/822), and al-ʾAxfaš al-ʾAwsaṭ (d. 215/830 or 225/839), who wrote treatises with the title *at-Taṣrīf*; ʾAbū ʿUmar al-Jarmī (d. 225/839), with the *Kitāb al-ʾabniya wa-t-taṣrīf*; ʾAbū ʿUṯmān al-Māzinī (d. 249/862), with the *Kitāb at-taṣrīf*; al-Mubarrad (d. 285/898), with his *at-Taṣrīf* or *at-Taṣārīf*; ar-Rummānī (d. 384/994), with the *Kitāb at-taṣrīf*; ʾAbū ʿAlī al-Fārisī (d. 377/987), with *at-Takmila*; Ibn Jinnī (d. 393/1002), with his *at-Taṣrīf al-Mulūkī (de Flexione), al-Munṣif*, and *Sirr aṣ-ṣināʿa*; al-Muʾaddib (active during the 4th/10th century), with the *Daqāʾiq at-taṣrīf*; al-Maydānī (d. 518/1124), with the *Nuzhat aḍ-ḍarf fī ʿilm aṣ-ṣarf*; Ibn Yaʿīš (d. 642/1245), with his *Šarḥ al-Mulūkī*; Ibn al-Ḥājib (d. 646/1249), with *aš-Šāfiya*; az-Zanjānī (d. 654/1256–1257), with his *Kitāb al-ʿIzzī*; Ibn ʿUṣfūr (d. 669/1270), with his *al-Mumtiʿ fī t-taṣrīf*; Ibn Mālik (d. 672/1273), with his *Lāmiyat al-ʾafʿāl*; Ibn Mālik's son, Badr ad-Dīn (d. 686/1286), with his commentary on the *Lāmiya* commentary; al-ʾAstarābāḏī (d. 686/1286), with his *Šarḥ Šāfiyat Ibn Ḥājib*; and ʾAḥmad ibn ʿAlī ibn Masʿūd (active at the end of the 7th/13th century or the beginning of the 8th/14th century), with his *Marāḥ al-ʾarwāḥ* (Åkesson 2001). A few grammarians, however, followed the model of *al-Kitāb*, e.g. az-Zamaxšarī (d. 538/1143–4), with his *al-Mufaṣṣal*, on which Wright (1985) is based, and Ibn Yaʿīš, with his *Šarḥ al-Mufaṣṣal*, on which Howell (1986) is based.

Ibn Jinnī's *Munṣif* in particular marks the integration of new methods of reasoning in the study of morphology (cf. Guillaume 1981:224). Between the 3rd/9th century and the 8th/14th century, which represents the end of grammatical productivity, the notion of *qiyās* 'analogy', nourished by the translations from the Greek (Versteegh 1977) and already in use in Sībawayhi's book (cf. Baalbaki 1979; Carter 1997:527), was developed and became well integrated in matters of morphophonology. In their study of *ṣarf* many grammarians strove to explain the reasons of many particularities (Åkesson 2001:27–35), such as the phonological changes, the avoidance of specific combinations, the addition or elision of a vowel (cf. Åkesson 1999) or a segment, the choice of a segment, vowel, or quiescence, and the transfer of a vowel or a segment in a word, by relying on fundamental rules and principles (for this term see Baalbaki 1988) in their reasoning. The discipline of *ṣarf* could include, for instance, the study of the selection of the vowels to indicate declension or undeclinability, e.g. the choice of the *ḍamma* for the indicative because of its resemblance to the noun or to the active participle (cf. Ibn Yaʿīš, *Šarḥ* VII, 6; Åkesson 2001:163–164), the *fatḥa* for the perfect because of its semiresemblance to the active participle (cf. Åkesson 2001:135–137), and the *sukūn* for the imperative and jussive because of its lack of resemblance to the noun (cf. Ibn Yaʿīš, *Šarḥ* VII, 4; Åkesson 2001:139).

3. PARTS OF *ṢARF*

There are two main parts of *taṣrīf* (cf. Ibn ʿUṣfūr, *Mumtiʿ* I, 31–32; Bohas a.o. 1990:76–80; Mokhlis 1997:24–25; Åkesson 2001:40–41). The first part is associated with the morphological *ištiqāq* 'derivation'. It deals with the derivation of words from roots and their shift from one form to another, according to the different patterns that assign them their specific meaning (cf. Ibn Jinnī, *Munṣif* I, 3–4). The root is mostly triliteral and is represented by the three consonants *f-ʿ-l* (the *f* standing for the first radical, the *ʿ* for the second, and the *l* for the third radical), which together with its derivatives are used as paradigms. There is a difference of opinion between the Basran and the Kufan grammarians concerning the root or source of the derivatives, as the Basrans consider this to be the → *maṣdar* 'verbal noun', e.g. *ḍarbun* 'a hitting', while the Kufans consider it to be the perfect of the verb, e.g. *ḍaraba* 'he hit' (cf. Ibn al-ʾAnbārī, *ʾInṣāf,* question 28; Versteegh 1995:72ff.).

The word can be regarded as consisting of a root morpheme, a specific voweling of its segments, and in many cases prefixes, infixes, and/or suffixes. These additional segments are ten in number and are combined in a hemistych attributed to al-Māzinī: *hawiytu s-simāna* 'I loved the plump women' (cf. Ibn Jinnī, *Munṣif* I, 98). For instance, the verbal form *ḍaraba* 'he hit', which has no additional segment attached to it, consists of the root morpheme *ḍ-r-b* 'notion of hitting' and the specific voweling of the segments *a-a-a*, the last *fatḥa* being the marker of undeclinability. This word corresponds to the measure *faʿala*, which indicates the 3rd person masculine singular of the perfect, represented by the pattern CvCvCv. From the same root *ḍ-r-b*, different variants of the fundamental meaning ('hitting') can be generated by vowel changes and/or the insertion of additional segments. For instance, one can derive the perfect *ḍarabtu* 'I hit', corresponding to the pattern *faʿaltu* with the *-tu* inserted as an agent pronominal suffix marking the 1st person singular; the imperfect *yaḍribu* 'he hits', corresponding to the measure *yafʿilu* with the *ya-* inserted as an imperfect prefix marking the 3rd person singular, the last *ḍamma* being the declensional marker; the active participle *ḍāribun* 'one who is hitting', corresponding to *fāʿilun* with the *ā* infix and the nunation *-un* indicating indefiniteness; and the passive participle *maḍrūbun* 'one who is hit', corresponding to *mafʿūlun* with the *ma-* prefix and the *ū* infix, etc.

Following Sībawayhi's lead, grammarians study in this part of *taṣrīf* the base forms and augmented forms of the nouns and verbs. The most common forms of the nouns are the verbal noun, the active participle, the deverbal adjective, the elative, the passive participle, and the nouns of time, place, and instrument.

Ibn Jinnī (*Taṣrīf*) includes in his introduction to the study of *taṣrīf* the various verbal forms. The inflection of the perfect, imperfect, and imperative of the base form are formed by vowel changes. Additional segments are inserted in all the forms (except the 3rd pers. masc. sg. perf.) and indicate person, number, and gender. Examples are the perfect forms with suffixes *ḍarab-tu* 'I hit', *ḍarab-nā* 'we hit', and *ḍarab-na* 'they hit [fem. pl.]', etc.; and the imperfect and imperative forms with prefixes and suffixes, e.g. *yaḍribūna* 'they hit [masc. pl.]', *yaḍribna* 'they hit [fem. pl.]', *iḍribū* 'hit! [masc. pl.]', etc. Additional changes are carried out to indicate active or passive voice, e.g. in the perfect *ḍaraba* 'he hit' and *ḍuriba* 'he was hit', in the imperfect *yaḍribu* ‘he hits’ and *yuḍrabu* ‘he is being hit’; and to indicate → mood, e.g. indicative *yaḍribu*, jussive *lam yaḍrib* ‘he did not hit’, subjunctive *ʾan yaḍriba* ‘that he shall hit’, → energicus *lā taḍribanna* ‘do not hit!’, or imperative energicus, e.g. *iḍribanna* ‘hit! [2nd pers. masc. sg.]’.

Nouns and verbs are divided into seven classes, namely: *ṣaḥīḥ* ‘strong verb’, which has three consonants as its radicals (e.g. *ḍaraba* ‘to hit’); *muḍāʿaf* ‘doubled verb’, whose second and third radicals are identical (e.g. *madada* > *madda* ‘to stretch’); *mahmūz* ‘hamzated verb’, which contains a *hamza* as one of its radicals (e.g. *ʾaxaḏa* ‘to take’, with the *hamza* as the first radical, *saʾala* ‘to ask’ with the *hamza* as the second radical, and *qaraʾa* ‘to read’ with the *hamza* as the third radical); *miṯāl* ‘verb with first radical *w* or *y* (e.g. *waʿada* ‘to promise’, *yasira* ‘to be easy’); *ʾajwaf* ‘verb with second radical *w* or *y* (e.g. *qawala* > *qāla* ‘to say’ and *bayaʿa* > *bāʿa* ‘to sell’); *nāqiṣ* ‘verb with third radical *w* or *y* (e.g. *ġazawa* > *ġazā* ‘to attack’, *ramaya* > *ramā* ‘to throw’); and *lafīf* ‘doubly weak verb’ (e.g. *waqaya* > *waqā* ‘to guard’, with first and third weak radical and *ṭawaya* > *ṭawā* ‘to fold’, with second and third weak radical).

The second part of *taṣrīf* deals with derivations from one underlying form to another by one or more phonological changes, without a semantic change, e.g. *qāla* from *qawala* (cf. Ibn ʿUṣfūr, *Mumtiʿ* I, 31–33). These phonological changes are due to many causes, such as the presence of two identical segments or two segments close to each other in place of articulation, or a *hamza* or one or more weak segments in a word. Ibn Jinnī (*Taṣrīf* 8) recognizes the following changes on the basis of Sībawayhi’s *Kitāb: ziyāda* ‘addition of a segment’, *badal* ‘substitution of one segment for another’, *ḥaḏf* ‘elision of a segment or a vowel’, *taġyīr bi-ḥaraka aw bi-sukūn* ‘a change with a vowel or with a quiescence’, and *ʾidġām* ‘assimilation of one segment to another’.

Two identical segments or two segments close to each other combined in one word or in two words can lead to → *ʾidġām* ‘assimilation’ (Sībawayhi, *Kitāb* II, 443ff.; Åkesson 2001:194ff.). An example is the assimilation of the *ḍād*s in the doubled verb *ʿaḍiḍa* > *ʿaḍḍa* ‘to bite’ and the *bā*s in two adjacent words *la-ḏahaba bi-samʿihim* > *la-ḏahabbi-samʿihim* ‘He would take away their faculty of hearing’ (Q. 2/19), read so by ʾAbū ʿAmr (cf. Zamaxšarī, *Mufaṣṣal* 195; Ibn Yaʿīš, *Šarḥ* X, 147; Åkesson 2001:210). The closeness of two different segments in point of articulation or manner can also lead to assimilation. In order to explain this phenomenon, Sībawayhi (*Kitāb* II, 452ff.) treats the subject of phonetics by introducing the 29 base segments, indicating for each its point of articulation and manner (cf. Åkesson 2001:222–225). In the *Sirr aṣ-ṣināʿa*, Ibn Jinnī studies these segments of words at the level of phonetics, phonology, and morphology. This kind of assimilation between two different segments is carried out in one word and in two successive words (cf. Åkesson 2001:194ff.). An example of its occurrence in one word is the assimilation involving the augmented *t* in some derived verbs, e.g. Form V *tazayyana* > *izzayyana* ‘to decorate itself’, Form VI *taṯāqala* > *iṯṯāqala* ‘to be borne down heavily’ with the insertion of a prosthetic hamza, and Form VIII perfect *iṯtaʾara* > *iṯṯaʾara* ~ *ittaʾara* ‘to get one’s revenge’ and imperfect *yabtasimu* > *yabbasimu* ‘to smile’. An example of its occurrence in two successive words is the assimilation of *t* to *ṯ* in *sakata ṯāmirun* > *sakaṯṯ āmirun* ‘a wealthy man fell silent’ and *t* to *š* in *ʾaṣabat šarban* > *ʾaṣabaššarban* ‘she obtained a drink’ (cf. Åkesson 2001:210). Sībawayhi’s study of phonetics is important for the study of *badal* or → *ʾibdāl* ‘substitution’ as well (cf. *Kitāb* II, 340–342; Ibn ʿUṣfūr, *Mumtiʿ* I, 319–415; Åkesson 2001:330ff.), as it explains why one segment can be substituted for another in a few contexts. The segments of substitution are combined in different mnemonic phrases, among them *istanjadahu yawma sāla zuṭṭa* ‘he asked him for help on the day some Zuṭṭ [sc. a people of Hindus] attacked’ (cf. Zamaxšarī, *Mufaṣṣal* 172; Ibn Yaʿīš, *Šarḥ* X, 7–8; Åkesson 2001:342–343). Examples of substitution are *waḥḥid waḥḥid* > *ʾaḥḥid ʾaḥḥid* ‘make the sign with one, one!’, *wišāḥun* > *ʾišāḥun* ‘baldric’, *ʾaraqtu* > *haraqtu* ‘I spilled’, etc.

The *hamza* behaves as an unstable segment in some contexts (cf. Åkesson 2001:240ff.). It can be made lighter by its change into *ā*, e.g. *raʾsun* > *rāsun* ‘head’, *hanaʾaki* > *hanāki* ‘it was pleasant for you’; *ū*, e.g. *luʾmun* > *lūmun* ‘blame’, *ʾuʾṯira* > *ʾūṯira* ‘he was preferred’; or *ī*, e.g. *biʾrun* > *bīrun* ‘well, spring’, *ʾiʾsir* > *ʾīsir* ‘capture! [2nd pers. masc. sg.]’. It can also be elided and have its vowel transferred to the vowel preceding it. This may occur within one word, e.g. *malʾakun* 'angel' > *malakun*, or in two successive words, e.g. *ʾAbuw ʾAyyūba* > *ʾAbuwa yyūba*.

The presence of a glide, *w* or *y*, in a word may lead to phonological changes called *ʾiʿlāl* 'changes due to unsoundness' (cf. Åkesson 2001:282ff.; → *ʿilla*). The most common changes are that the glide is made vowelless, e.g. *yaġzuwu* > *yaġzuw* > *yaġzū* 'he attacks', *yarmiyu* > *yarmiy* > *yarmī* 'he throws'; that it is changed into an *ā*, e.g. *bayaʿa* > *bāʿa* 'to sell' (cf. Åkesson 1996), *qawala* > *qāla* 'to say'; that its vowel is transferred to the segment preceding it before it is changed into an *ā* (*naql*), e.g. *yaxwafu* > *yaxawfu* > *yaxāfu* 'he is afraid', *yahyabu* > *yahaybu* > *yahābu* 'he is awed'; or that *w* is changed into *y*, e.g. *siwātun* > *siyātun* 'whips'.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Ibn al-ʾAnbārī, *ʾInṣāf* = ʾAbū l-Barakāt ʿAbd ar-Raḥmān ibn Muḥammad al-ʾAnbārī, *Kitāb al-ʾinṣāf fī masāʾil al-xilāf bayna n-naḥwiyyīna l-Baṣriyyīn wa-l-Kūfiyyīn*. Ed. Gotthold Weil. Leiden: E.J. Brill, 1913.

Ibn Jinnī, *Munṣif* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Munṣif fī šarḥ Taṣrīf al-Māzinī*. Ed. ʾIbrāhīm Muṣṭafā and ʿAbdallāh ʾAmīn. 3 vols. Cairo: Muṣṭafā al-Bābī al-Ḥalabī, 1373/1954–1379/1960.

——, *Sirr* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *Sirr ṣināʿat al-ʾiʿrāb*. Ed. Ḥasan Hindāwī. 2 vols. Damascus, 1405/1985.

——, *Taṣrīf* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *Kitāb at-taṣrīf: De Flexione Libellvs*. Ed. G. Hoberg. Lipsiae, 1885.

——, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. Ed. Muḥammad ʿAlī an-Najjār. 3 vols. Cairo, 1371/1952–1376/1956.

Ibn Manẓūr, *Lisān* = Jamāl ad-Dīn Ibn Manẓūr, *Lisān al-ʿArab*. 6 vols. Beirut, n.d.

Ibn ʿUṣfūr, *Mumtiʿ* = ʾAbū l-Ḥasan ʿAlī ibn Muʾmin Ibn ʿUṣfūr al-ʾIšbīlī, *al-Mumtiʿ fī t-taṣrīf*. Ed. Faxr ad-Dīn Qabāwa. Aleppo: Dār al-ʾĀfāq al-Jadīda, 1390/1970.

Ibn Yaʿīš, *Šarḥ* = Muwaffaq ad-Dīn ʾAbū l-Barāʾ Yaʿīš Ibn Yaʿīš, *Šarḥ al-Mufaṣṣal*. 2 vols. Beirut, n.d.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb: Le Livre de Sîbawaihi (Kitāb Sībawaihi), Traité de grammaire arabe*. Ed. Hartwig Derenbourg. 2 vols. Paris, 1881–1889. (Repr., 1970.)

ʿUkbarī, *Masāʾil* = ʿAbdallāh ibn al-Ḥusayn al-ʿUkbarī, *Masāʾil xilāfiyya fī n-naḥw*. Ed. Muḥammad Xayr al-Ḥalwānī. Aleppo, n.d.

Zamaxšarī, *Mufaṣṣal* = ʾAbū l-Qāsim Maḥmūd ibn ʿUmar az-Zamaxšarī, *al-Mufaṣṣal*. Ed. Jens Peter Broch. Christianiae: Libraria P.T. Mallingii, 1840.

SECONDARY SOURCES

Åkesson, Joyce. 1996. "Conversion of the *yāʾ* into an *alif* in Classical Arabic". *Zeitschrift für Arabische Linguistik* 31.27–33.

——. 1999. "Anomalous elision and addition of a vowel in Classical Arabic". *Zeitschrift für Arabische Linguistik* 36.21–31.

——. 2001. *Arabic morphology and phonology based on the* Marāḥ al-arwāḥ *by Aḥmad b. ʿAlī b. Masʿūd, presented with an introduction, Arabic edition, English translation and commentary*. Leiden: E.J. Brill.

Baalbaki, Ramzi. 1979. "Some aspects of harmony and hierarchy in Sībawaihi's grammatical analysis". *Zeitschrift für Arabische Linguistik* 2.7–22.

——. 1988. "A contribution to the study of technical terms in early Arabic grammar: The term *aṣl* in Sībawaihi's *Kitāb*", *A miscellany of Middle Eastern articles,* ed. A.K. Irvine a.o., 163–77. Essex: Longman.

Bohas, Georges, Jean-Patrick Guillaume, and Jamel Eddine Kouloughli. 1990. *The Arabic linguistic tradition*. London and New York: Routledge.

Carter, Michael G. 1981. *Arab linguistics: An introductory Classical text with translation and notes*. Amsterdam: J. Benjamins.

——. 1997. "Sībawayhi". *Encyclopaedia of Islam*, IX, 524–531. 2nd ed. Leiden: E.J. Brill.

Guillaume, Jean-Patrick. 1981. "Le statut des représentations sous-jacentes en morphonologie d'après Ibn Ǧinnī". *Arabica* 28.222–41.

Hindāwī, Ḥasan. 1989. *Manāhij aṣ-ṣarfiyyīn wa-maḏāhibuhum fī l-qarnayn aṯ-ṯāliṯ wa-r-rābiʿ min al-hijra*. Damascus.

Howell, Mortimer Sloper. 1986. *Grammar of the Classical Arabic language*. 4 parts in 7 vols. New Delhi: Gian Publishing House. (1st ed., 1883–1911.)

Lane, Edward William. 1863–1893. *Arabic-English lexicon*. 8 in 2 vols. London and Edinburgh: Williams and Norgate. (Repr., Beirut: Librairie du Liban, 1984.)

Mokhlis, Hassan. 1997. *Théorie du* taṣrīf *et traitement du lexique chez les grammairiens arabes*. Frankfurt a. Main: Lang.

Troupeau, Gérard. 1993. "*Naḥw*". *Encyclopaedia of Islam*, VII, 913. 2nd ed. Leiden: E.J. Brill.

Versteegh, Kees. 1977. *Greek elements in Arabic linguistic thinking*. Leiden: E.J. Brill.

——. 1995. *The explanation of linguistic causes: Az-Zaǧǧāǧī's theory of grammar*. Amsterdam: J. Benjamins.

——. 1997. "*Ṣarf*". *Encyclopaedia of Islam* IX, 53–54. 2nd ed. Leiden: E.J. Brill.

Wright, William. 1985. *A grammar of the Arabic language*. 3rd ed. rev. W. Robertson Smith and M.J. de Goeje. Repr. Cambridge.: Cambridge University Press.

JOYCE ÅKESSON (Lund)

Saudi Arabia

1. Introduction

Saudi Arabia is an area of considerable linguistic uniformity. With regard to languages that are native to the country, on the basis of the most recent data, only one language, Arabic, is spoken, although dialect diversity is considerable. In fact, in the southwest on the border of Yemen, unconfirmed reports have the dialect of Fayfa to be mutually unintelligible with local Arabic dialects and showing a substratum of the → Ḥimyaritic languages of Ancient South Arabia. Although modern descendants of these languages are spoken now only in pockets in the Yemen and Oman (→ South Arabian, Modern), they are thought to have been much more widespread in the pre- and early Islamic period, particularly in the west of the peninsula and probably covering more than the present territory of the state of Yemen; thus, they may have influenced the Arabic of the area considerably. To quote Rabin in his classic study of pre-Islamic dialects:

> The little we know of the Northern Yemenite dialects rather tends to suggest that there was a continuous chain of dialects from south to north without any dividing line between Yemen and Hijaz (Rabin 1951:25)

On this basis, it would seem that the Najdi dialect area has expanded southward into the area of Bīša and Najrān, probably since the beginning of the Islamic era, corresponding to the diminution of the political and commercial importance of the Yemen region.

By one theory, the Arabs and their language are thought to have originated in the north of the area around Dūmat al-Jandal, or the present city of al-Jawf, and spread southward into the peninsula, with their language gradually replacing or merging with the languages spoken there, which were possibly, but not necessarily, South Arabian.

With regard to actual minority languages presently spoken in Saudi Arabia, it is difficult to make accurate statements. As stated above, the only regional language is Arabic. However, the Ḥijāz in particular has long-standing immigrant populations from non-Arabic-speaking regions, many of whom came as pilgrims and stayed. These are referred to in general as Ḥijiz, literally people of the Ḥijāz, but denoting those of immigrant origin. Among them are, giving the local taxonomy: Bukharis (people from Bukhara) denoting Central Asians, Yawis (people from Java) denoting Indonesians and also Caucasians, and Africans, in particular from Nigeria. It is difficult to say how many of these populations retain a foreign native language beyond the second generation, although one does meet individuals of the younger generation who were born in the Ḥijāz and speak such languages as Chechen or Central Asian Turkic.

In addition, large numbers of temporary immigrants work in the country for long periods. These include Philipinos, Sri Lankans, and citizens of the state of Kerala (India), but most of these intend to return home eventually, and it is unlikely that their languages will retain a permanent footing. There is, however, a frequently heard type of 'pidgin Arabic' spoken both by these immigrants and their Saudi employers. This consists of sentences composed from Arabic and foreign nouns and undeclined verbal stems, often using the imperative form as the stem, in such utterances as *ana rūhi kalām bāba* 'I will go and speak to the master' and *ana rūhi jīb bānzēn* 'I will go and bring gas'. The phonology is simplified, with the Arabic pharyngeals and emphatics being replaced by simpler sounds, as in *rūhi* for *rūḥi* above. This pidgin is frequently encountered at gas stations and in small shops, especially in the smaller towns and villages, where customers might not be expected to speak English. It is also institutionalized in popular culture and occurs as the form of speech of foreign domestic servants in popular radio and television plays. English is of course widely used in large retail outlets, especially in Riyadh.

2. General dialect geography

The dialectology of Saudi Arabia forms a continuum with that of Yemen and Oman to the south, the Emirates and Qatar and Kuwait to the east, and Jordan and Iraq to the north. There are no discrete borders separating Saudi dialects from others. In the north, the border areas are desert, and the nomads form the link in dialect, while in the southwest, there is a continuation of the South Arabian cultural zone of the Yemen. To the southeast, the border zone is the Rubʿ al-Xālī or 'Empty Quarter', an area of only nomadic population. From what little we know of the dialects of that area, those of the south would seem to bear most resemblance to the dialects of Oman (→ Omani Arabic) and the Arab Emirates (→ Gulf Arabic), earlier known as the 'Trucial Oman'. Bertram Thomas describes the dialects of the area as follows:

> The sands (the Empty Quarter) know two dialects of Arabic, but here the division is latitudinal. There is a northern or Murri dialect and a southern or Rashid dialect, the latter also spoken by the ʿAwamir of the northeast (Thomas 1932:268)

And he cites an important feature of these dialects:

> In the Rashidi dialect of the southern sands as distinct from the northern dialect of the Murra, J is pronounced Y, thus Jaub = Yaub and Jiban = Yiban (Thomas 1938:164, n. 1)

The dialect of the Āl Murrah is referred to below; it is of the Najdi type, with certain southern features, and would seem to form the southernmost limit of Najdi dialect, at least when that tribe is residing in the Empty Quarter.

The population of Saudi Arabia includes settled and nomadic groups. Generally, the dialect of the nomads relates to the area from which they originate and shows links with the sedentary dialects of that region. With a degree of generalization, the following regions may be distinguished: (i) dialects of the Najd; (ii) dialects of the southwest; (iii) dialects of the eastern region; and (iv) dialects of the Ḥijāz. The southwest and the eastern region have old, established, settled populations, based on water supply, monsoon rains in the southwest, and oases in the east. The eastern region referred to means the area of al-Ḥasa, an oasis region of great antiquity, and the coastal region of Qaṭīf. Here, the dialects show specific local features, often highly innovating. Najd also has a core of settled population centered on oases, but with close links with nomadic populations in the northern part. The rural Ḥijāzi dialects seem to form a dialect continuum with the Yemen, but little is known of the northern Ḥijāz, while the urban Ḥijāz has been subject to immigration from outside the area and its dialects show Arabian and non-Arabian characteristics and a general leveling of the archaic features of the Najdi type, possibly attributable to population mixture (see Sieny 1978; Feghali 1991). Najd shows very archaic features, showing a considerable resemblance to Classical Arabic, in particular an extensive use of an *-in* ending, whose distribution is almost identical with the Classical Arabic → *tanwīn*, and extensive use of a passive marked by internal voweling of the same type as Classical Arabic. Another archaic feature is the occurrence of a preverbal particle *ǵid* or *ćid*, having a similar function to Classical Arabic *qad*, and the use of the negative with *b-*.

The interdentals **ṯ* and **ḏ* are also retained in Najd, rural Ḥijāz, and the southwest, but they have merged with **t* and **d* in Urban Ḥijāzi and occasionally in the eastern region. Also in the east in places, the voiceless interdental **ṯ* has merged with /f/, giving *fallāja* 'refrigerator', *falj* snow, ice', and *falāfa* 'three'. **ḍ* and **ḏ̣* have merged to /ḏ̣/ in Najd, rural Ḥijāz, and the southwest, but merged to /ḍ/ in Urban Ḥijāzi and the dialect of Qaṭīf in the east.

Feminine plural forms are retained in the morphology in some areas of Najd and the southwest, but not universally so. Urban Ḥijāzi does not retain them, nor does the eastern region. In certain cases, the feminine plural verbal suffixes *-an, -in*, etc. seem to be more regularly used for inanimate plurals and with animals, while actual human feminine plurals are referred to by the original masculine plural, which now indicates human plural.

The conservative features are retained in the dialects of the Bedouin, who can be thought to have originated in central Najd. These have been carried far into the Syrian Desert and at an earlier stage spread into the desert regions of Syria and Iraq, although many of these populations have now returned south to Saudi territory.

This archaic appearance is even more marked in the dialect of Najrān to the southwest of Najd, since here the fronting of /k/ and /q/ does not occur, nor, in most cases, does the → resyllabification of suffixed forms. Also here, the incidence of the *-in* suffix and the *gid* preverbal particle are particularly high. In addition, the southwest retains certain Classical Arabic-sounding lexical items, which do not occur so frequently in the North, in particular reflexes of *raʾā* 'to see', *ʾatā* 'to come', and *ḏahaba* 'to go', rather than the more commonly seen *šāf, jā*, and *rāḥ*, sometimes in co-occurrence with the latter. One possible interpretation of this concurrence of southern forms with Classical Arabic ones could be that the early grammarians, working in the north, in seeking to standardize the written form, selected the southern forms, since they did not occur in the spoken dialects to the north and therefore, in being strange to the local spoken form, had a more Classical Arabic ring. This, however, is only speculation.

A feature common to the whole area, and possibly to all spoken Arabic, is the almost complete disappearance of the glottal stop as a reflex of Classical Arabic → *hamza*. Although the form *saʾal* 'to ask' is common in the urban Ḥijāzi dialects and in formal speech in most areas, it has been replaced in much of the rest of the area in dialect speech by such forms as *nišad, sāl*, or *saʿal*, the latter with /ʿ/, recalling the *ʿanʿana* mentioned by the Old Arabic grammarians (Rabin 1951:194). This incidence of /ʿ/ can also be seen in the deictic *arʿih* 'see it!', traceable to a root *raʿa* 'to see' and occurring in the dialect of the Rwalah. Southern dialects also show reflexes of Classical Arabic *raʾā*, showing a stem **ra* without the *hamza*, as in *rēt* 'I saw', *rēna* 'we saw'; in most other instances, *hamza* is regularly omitted, as in *rās* 'head', *rūs* 'heads', *maddibah* 'reproach' cognate with *ʾadab* 'good manners'.

3. Local dialect features

Local dialect features can be examined in the context of the distribution of population centers, in the case of the settled population, and population movement, in the case of the nomads. The main population foci are: (i) the southwest on the borders of Yemen, with the Ḥijāz as an extension of it; (ii) central Najd and the Šammar highland region to the north of it; and (iii) the oasis and coastal region of the east. With that as background, one can also note a general historical movement northward and eastward within the Arabian Peninsula, so that the Āl Murrah and ʿIjmān tribes originating in Najrān have moved northeastward to the area of eastern Arabia, as far north as Kuwait. Tribes from western Najd, such as the Muṭayr and Ḥarb, have moved eastward to central Najd and the eastern region, and the ʿAnizah, who originate in the region of Xaybar, have moved to central Najd, with a later movement of some of them to the northern borders of Saudi Arabia. This historical movement is noted in local folk memory in such sayings as *al-yaman raḥim al-ʿArab wa-l-ʿIrāq qabruhum* 'Yemen is the womb of the Arabs and Iraq their grave', and *najd tuwallid wa lā tuġaḏḏi* 'Najd gives birth, but does not nourish'. The last refers in part to the memory of recurrent droughts, which caused individuals, groups, or whole tribes to leave the central area from time to time. Classical historical tradition refers to the breaking of the Maʾrib dam in Yemen as the cause of movement of many groups northward from Yemen. Other traditions remember the migration of the Banū Hilāl in the 9th century from Arabia to North Africa. In more recent times, the movement of the Šammar and Ḍafir to the Jazīrah in the 18th and 19th centuries, and of the ʿAnizah to Iraq and Syria at the same time, reflect this movement. These tribes retained the original Najdi form of their dialects, as shown by recent studies, possibly helped by the isolation of Bedouin life and the sentiment that the dialect was part of their tribal identity and to them represented a purer form of Arabic, than the dialects of the settled people they encountered there. The same sentiment was expressed by a non-Bedouin *sayyid* in Iraq in the form *sīd alḥači ʿind al-badu* 'the most noble of speech is heard among the Bedouin'. Johnstone (1964:85) also reports the impression evident in Kuwait that the dialect of the ʿIjmān tribe, relatively recent arrivals from the southwest, represented the 'best' Arabic. This impression was probably based on the absence of the affrication of /k/ and /q/ in that dialect, and possibly also on the other marked archaisms mentioned below for the southwest.

Occasionally, the same linguistic features are found in areas far apart. In some cases, these look like preservation of an older common form. In others, they may be independent developments based on universal linguistic tendencies (Ingham 1982). Possibly of the first type is a 2nd person singular masculine suffix *-ta*, as in *kitabta* 'you wrote', occurring in Sudayr in central Najd and also in the south in Maḥāyil in the Tihāmah and Bani Šihr in the ʿAṣīr. Also, the form *-iš* of the 2nd person singular feminine suffix occurs both in areas of the southwest and in the eastern region. This may be the result of a common historical link with the southern region, possibly seaborne in the case of the east, but not easily proven. Of the second type is a reflex /ž/ of *ǰ (→ *jīm*), occurring both in the dialect of the Ḥarb tribe in the region of Medina and also in the dialect of Ġāmid in the southwest, and a reflex /y/ of the same *ǰ in the dialect of Bani Ṯōr in al-Ḥasa and of a number of localities in the southwest. Bani Ṯōr has *wāyid* 'much', *rayyāl* 'man', *daray* 'stairs', while the southwest has examples such as *yabal* 'mountain', *yarād* 'locust', and *yāyiʿ* 'hungry'. Similarly, if one goes further afield, one notes that a form *-am* of the 3rd person plural masculine suffix of verbs, as in *rāḥam* 'they [masc.] went', occurs in dialects originating in the central Najd (cf. Cairene *-um*). With the present state of our knowledge, there seems no reason to suppose that these have a common origin.

Actual dialect features are examined below by reference to these original population foci. Those features that differ from Classical Arabic are pointed out, Classical Arabic being used as a general point of reference.

3.1 *Najd*

3.1.1 Central Najd

A feature which originates in central Arabia and has spread outward to the borders of Palestine and most of southern Iraq and the Gulf Coast is fronting of *k and *q in the environment of front vowels, *k going to /č/ and later /ć/ and *q going via /g/ to /j/ and later /ǵ/, respectively (→ *kaškaša/kaskasa*). This has produced a → phonological split of *k and *q, giving *kōn* 'battle' but *ćān* 'if', *galb* 'heart' but *ǵiddām* 'in front of'. Dialects on the periphery can show /č/ and /j/, such as the dialect of Hufūf toward the east, showing *čanʿad* 'mackerel', *yačwi* 'he brands', *rifīj* 'companion', and *ṭirīj* 'road'.

Other characteristics of high incidence are the raising of *a to /i/ or /u/ in short, open, nonfinal syllables, as in *kitab* 'he wrote', *hubaṭ* 'he went down'; resyllabication associated with guttural consonants producing, for instance, *ghawah* rather than *gahwah* 'coffee', and *nxalah* rather than *naxlah* 'palm' (→ *gahawah*-syndrome); and → resyllabication associated with a succession of short open syllables, giving forms such as *ktibat* 'she wrote', rather than *katabat*. These last features have been used as diagnostics to distinguish Najdi from non-Najdi dialects.

Central Najd, areas to the south, and some eastern dialects regularly show a negative with *b-*, as in *māni b-rāyiḥ* 'I am not going', *muhu b-zēn* 'it is not good'.

The dialect of Sudayr retains the form *-ta* for the 2nd person singular masculine in *simiʿta* 'you heard', and an ending *-am* for the 3rd person plural masculine occurs in some central Najdi dialects and also in the dialect of the émigré Rwalah in the Syrian desert.

3.1.2 Northern Najd

Northern Najdi, typified by the dialect of the Šammar tribe and the city of Ḥāyil, shows, in addition to the general characteristics shown above, innovations in verbal morphology and also marked → *ʾimāla* of the feminine singular and plural suffixes, resulting in change of final *-t-* to *-h-* or *-y-*, yielding *-eyh* for normal *-ah* or *–at* in the singular and *-āy* for more normal *-āt* in the plural. This yields such forms as *ṯlaṯ banāy ṭabbāxāy* 'three girls cooking' or 'girl cooks', *rāḥeyh* 'she went', *ghaweyh* 'coffee'. This feature is mentioned by the early Arabic grammarians as a feature of the Ṭayyiʾ tribe, who inhabited the area of Jabal Ṭayy, now known as Jabal Šammar (Rabin 1951:206). Verbal morphology shows *-an* for *-ni* as the 1st person singular object suffix, as in *ḥbisan* 'he imprisoned me'; *-ah* for the 3rd person singular feminine (*-h* postvocalically); *-uh* for the 3rd person singular masculine (*-w* postvocalically); and *-ham*, *-tam*, and *-kam* for the 3rd and 2nd person plural masculine, as opposed to *-hum, -tum*, and *-kum* elsewhere. It also differs from the central area in lacking the negative with *b-*.

3.2 *The Ḥijāz*

The rural dialects of Ḥijāz are not particularly well described and the description here depends mainly on Al-Ali (1981). Urban Ḥijāzi, meaning the dialects of Mecca and Jidda, are better known (Schreiber 1970; Ingham 1971; Al-Sasi 1972). Generally speaking, the two forms share basic phonological and, to a lesser extent, morphological features, while the urban form shows phonological leveling under the influence of dialects from outside the peninsula. *q is realized as /g/, *k as /k/ in both cities, and *ǰ as /j/. While the rural dialect preserves the interdentals in *hāḏa* 'this', *ṯāni* 'second', the urban dialect merges them with *d and *t to produce *hāda* and *tāni*. Old Arabic **ḏ̣* is realized as /ḏ̣/ among the older generation of rural speakers, but often as /ḍ/ with younger speakers, giving *waḍḍaḥ* or *waḏ̣ḏ̣aḥ* ‘to illustrate’. The urban dialect shows /ḍ/. What one might call ‘Arabian’ features occur in both, particularly the use of the particle *gid/gad*, here sometimes realized as *gīd*, the use of the possessive particle *ḥagg/ḥaggat*, and the verb *baġa/yabġi* ‘to want’. Examples include *mā gad/gīd ṭarētu* ‘I have never seen him’. Often, the speech of the older generation resembles Najdi in such lexical usages as *yiḏillūn* ‘they are afraid’, *ḥadart* ‘I went down [to the city]’, or *hubaṭna* ‘we went down’. Al-Ali’s (1981) study shows that the dialects of the Ḥijāz, at least in the area of Wādi Fāṭima, merge into the Najdi type as one goes east, affrication of **k* and **q* to /č/ and /j/ being shown in the dialect of Bukayriyyah, giving *čalb* ‘dog’ and *ṭirīj* ‘road’. This gradual eastward approximation to the Najdi type is also shown in detail in Al-Hazmi’s (1975) study for the dialect of the Ḥarb tribe near Medina.

3.3 *The southwest*

The dialects of the southwest can in some ways be considered the most archaic, since they show the archaic features of Najdi Arabic without the innovation of the fronting of **k* and **q*. They are not, however, particularly uniform. One remarkable feature of the southwest is the use of *-iš* for the 2nd person singular feminine object, and possessive affix in contrast to *-ik*, *-ič*, or *-ić* elsewhere (→ *kaškaša/kaskasa*). This feature is South Arabian in general and is found in Yemeni dialects and South Arabian languages and in fact also occurs in Semitic languages of Ethiopia. A further southern feature is the occasional incidence of /š/ for /s/ and /ž/ for /z/, in such forms as *šimiʿt* ‘I heard’ for *simiʿt*, *žyāra* for *zyāra* ‘visit’, and *ḥažm* for *ḥazm* ‘hill’. Prochazka (1988a) shows this for Sarāt ʿAbīdah, Rufayda, Tannūmah, and Najrān, and it is also recorded for the Āl Murrah. A parallel feature is the occurrence of a definite article *im-*, rather than the usual *al-* or *il-*. This is also common to South Arabian and parts of Yemen. The dialect of Najrān and associated Āl Murrah Bedouin have also developed a 2nd person singular suffix *-hant*, as in *gulhant* ‘you said’ and *šifhant* ‘did you see?’. They also show a marked usage of the preverbal particle *gid/gad*, which is associated with existential meaning, as in *wēn gadik ṯam* ‘where were you at that time?’, *hāk al-ḥīn gidni fissūg* ‘at that time I will already be in the market’, occurring with even more frequency than in the Najd.

The dialects of Bal Qarn, Bani Šihr, and Rijāl Alma in the ʿAṣīr show a suffix *-an* for the 3rd person singular feminine, rather than the usual *-at*, as in *atan* ‘she came’, *širiban* ‘she drunk’, *sāfaran* ‘she traveled’, and the dialect of Abha shows a negative particle *lis-*, reminiscent of Classical Arabic *laysa*, giving *lišu, lisu* ‘he is not’, *lisni* ‘I am not’.

Some areas show affrication of **k* and **q*, but not all. This may be thought to have spread into the area from Najd. Bani Bišr and Bal Qarn show /č/ and /j/, as in *čitif* ‘shoulder’, *čabb* ‘he spilled’, and Qahabah shows /č/, while Bīša shows /ć/, as in *rāćib* ‘riding’.

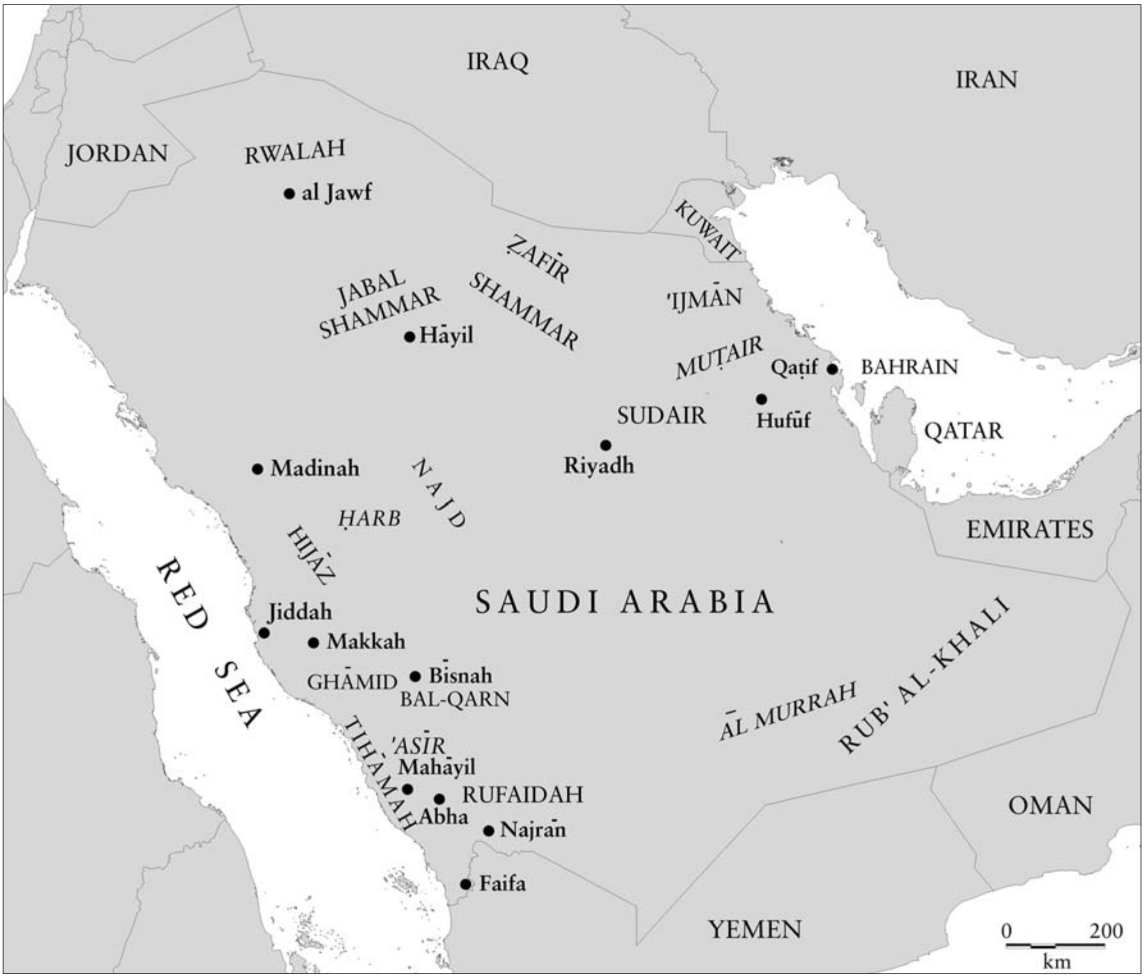
Bani Šihr and Maḥāyil also show the *-ta* suffix of the 2nd person singular masculine in verbs, as in *xarajta* ‘you went out’, *atēta* ‘you came’, and *šaribta* ‘you drunk’, which was also mentioned for Sudayr above.

Lastly, **k* occurs as a postpalatal fricative /x/ in many instances in the speech of Rijāl Alma in ʿAṣīr and Maḥāyil in the Tihāmah, giving *xān* ‘he was’ (for *kān*), *tanaxah* ‘a liquid measure’ (for *tanakah*), and *ġayrax* ‘other than you’ (for *ġayrak*).

3.4 *The eastern region*

What we know of the dialects of the eastern region is mainly due to the work of Prochazka (1988a). This reveals two types: one in al-Ḥasa shows some features of the South Mesopotamian type akin to the *gilit* dialects described by Blanc (1964), and one on the coast akin to the dialect of the Shiʿi Baḥārnah population of Bahrain. The first type, typified by the dialect of al-Thōr in the principality of ʿUmrān in the al-Ḥasa oasis has the Mesopotamian type of anaptyctic vowels separating final clusters, as in *šifit* ‘I saw’. It also shows verbal forms of the type *kitbat* ‘she wrote’ and *kitbaw* ‘they [masc.] wrote’, in contrast to the more usual *ktibat* type forms of Najd. However, it also shows features of the Emirates and Oman, namely the insertion of a suffix *-in* between an active participle and an object pronoun suffix, as in *māskinnah* ‘having caught [sg. masc.] him’ and *māskitinnah* ‘having caught [sg. fem.] him’, and the form *-iš* of the 2nd person singular feminine object suffix. Prochazka makes no mention of the affrication of **k* and **q* in this dialect, but neither does he give any counterexamples.

The dialect of Qaṭīf on the Gulf coast represents the Baḥārnah dialect, that of the older Shiʿi population of the eastern region, spoken also by that population in Bahrain. It is distinguished by having lost the interdentals, fusing them with /d/, /f/, and /ḍ/, giving *axad* 'he took', *fnēn* 'two', *falāfeh* 'three', and *ḍarab* 'he struck'. In contrast to the dialect of al-Ḥasa above, it does show affrication of **k*, giving *tḥačče* 'he talked' and *čēfeh* 'how is he'. Prochazka (1988a), however, makes no mention of the affrication of **q* > /j/, as one might have expected, although in fact the same is true for Muslim Baghdadi, which shows many instances of /č/, but few of /j/ as a reflex of **q*. It also shows the southern feature of *-iš* for the 2nd person singular feminine suffix, as in *ḍarabiš* 'he hit you [sg. fem.]'. In contrast also to the rest of the area, sequences of two short open syllables remain, as in *waladiš* 'your [sg. fem.] son', *ḍarabuk* 'he hit you [sg. masc.]', *katabat* 'she wrote', and *xašabateh* 'his piece of wood'. These contrast with elided forms in the Najdi and Mesopotamian type, such as *waladć* (Najdi), *waldič* (South Mesopotamian), *ktibat* (Najdi), *kitbat* (Mesopotamian), and seem to point to a greater antiquity of the dialects in this region.


IRAQ
IRAN
JORDAN
RWALAH
al Jawf
KUWAIT
ẒAFĪR
JABAL SHAMMAR
SHAMMAR
Hāyil
'IJMĀN
MUṬAIR
Qaṭif
BAHRAIN
SUDAIR
Hufūf
QATAR
Riyadh
Madinah
NAJD
ḤARB
HIJĀZ
EMIRATES
RED SEA
SAUDI ARABIA
Jiddah
Makkah
RUB' AL-KHALI
GHĀMID
Bīsnah
BAL-QARN
ĀL MURRAH
TIHĀMAH
'ASĪR
Mahāyil
RUFAIDAH
Abha
Najrān
OMAN
Faifa
YEMEN
0 200
km

Map 1. Saudi Arabia

4. DIGLOSSIA

As in the rest of the Arab world, a situation of diglossia exists, with Classical Arabic being opposed to the local spoken dialects. However, in Saudi Arabia, the spoken dialects of some groups often retain considerable prestige due to the prestige of the speakers. This is particularly true of the dialects of central Najd and the associated Bedouin dialects, due to the association of these populations with the ruling elite. Two factors seem to be responsible for this: the presence of a strong tradition of oral literature in the dialect, both prose and poetry (see Ingham 1986b; Kurpershoek 1994, 1995, 1999, 2002, 2005); and the archaic nature of the dialects, which means that they differ from Classical Arabic in less features than other more phonologically progressive dialects. The result of this is that the two forms are perceived locally to be forms of one language, *al-luġa al-ʿarabiyya*, differing mainly in regard to the affrication of **k* and **q*, and one has the impression that the perception of a very separate form called *fuṣḥā* is less marked among speakers.

In the 20th century, this oral literature has often been reduced to writing both by local scholars and foreign researchers and remains a very strong tradition (→ orality). Local poets such as the court poet Xalaf al-Haḏḏāl are called upon to produce poems for important occasions, and their poems are often recited on radio or television and recorded and circulated widely.

There is also a tradition of writing on subjects of local interest in Classical Arabic, but using local dialect vocabulary. As an example, one can cite Ibn Bišr's *ʿInwān al-majd fī tārīx Najd*, which uses such local vocabulary as *kōn* 'battle', *subul* 'expedition to the towns to buy provisions', *manāx* 'battle'.

For all of the above reasons, diglossia is possibly not felt to be as visible a phenomenon as in certain other areas of the Arab world, and in Saudi Arabia one hears the terms *fuṣḥā* and *dārija* less often than elsewhere.

BIBLIOGRAPHICAL REFERENCES

Abboud, Peter. 1979. "The verb in North Najdi Arabic". *Bulletin of the School of Oriental and African Studies* 42.467–499.

Al-Ali, Monira Dawod. 1981. *A critical and comparative study of contemporary Arabic fairy tales, collected from women mainly in the Wadi Fatima, near Mecca*. M.Phil. diss., University of Leeds.

Al-Azraqi, Munirah. 1998. *Aspects of the syntax of the dialect of Abha (South-west Saudi Arabia)*. Ph.D. diss., University of Durham.

Al-Hazmi, A. 1975. *A critical and comparative study of the spoken Arabic of the Ḥarb tribe in Saudi Arabia*. Ph.D. diss., University of Leeds.

Al-Sasi, Omar. 1972. *Sprichwörter und andere volkskundliche Texte aus Mekka*. Ph.D. diss., University of Münster.

Blanc, Haim. 1964. *Communal dialects in Baghdad*. Cambridge, Mass.: Center for Middle Eastern Studies, Harvard University.

Feghali, Habaa J. 1991. *Arabic Hijazi reader*. Wheaton, Ill.: Dunwoody Press.

Hess, Johann J. 1938. *Von den Beduinen des inneren Arabiens*. Zurich and Leipzig: M. Niehans.

Ibn Bišr, ʿUṯmān. N.d. *ʿInwān al-majd fī tārīx Najd*. Riyadh.

Ingham, Bruce. 1971. "Some characteristics of Meccan speech". *Bulletin of the School of Oriental and African Studies* 34.273–297.

——. 1982. *North East Arabian dialects*. London: Kegan Paul International.

——. 1986a. "Notes on the dialect of the Al Murrah of Eastern and Southern Arabia". *Bulletin of the School of Oriental and African Studies* 49.271–291.

——. 1986b. *Bedouin of northern Arabia: Traditions of the Al-Dhafir*. London and New York: Routledge and Kegan Paul.

——. 1994. *Najdi Arabic: Central Arabian*. Amsterdam and Philadelphia: J. Benjamins.

Johnstone, Thomas Muir. 1964. "Further studies on the Dōsiri dialect of Arabic as spoken in Kuwait". *Bulletin of the School of Oriental and African Studies* 27.77–113.

——. 1967. *Eastern Arabian dialect studies*. London: Oxford University Press.

Kurpershoek, Marcel. 1994. *Oral poetry and narratives from Central Arabia*. I. *The poetry of ad-Dindān, a Bedouin bard in Saudi Arabia*. Leiden: E.J. Brill.

——. 1995. *Oral poetry and narratives from Central Arabia*. II. *The story of a desert knight, the legend of Šlēwīḥ al-ʿAṭāwi and other ʿUtaybah heroes*. Leiden: E.J. Brill.

——. 1999. *Oral poetry and narratives from Central Arabia*. III. *Bedouin poets of the Dawāsīr tribe, between nomadism and settlement in southern Najd*. Leiden: E.J. Brill.

——. 2002. *Oral poetry and narratives from Central Arabia*. IV. *A Saudi tribal history: Honour and faith in the traditions of the Dawāsīr*. Leiden: E.J. Brill.

——. 2005. *Oral poetry and narratives from Central Arabia*. V. *Voices from the desert: Glossary, indices, and list of recordings*. Leiden: E.J. Brill.

Philby, H. St. John. 1933. *The Empty Quarter*. London.

Prochazka, Theodore. 1988a. *Saudi Arabian dialects*. London: Kegan Paul International.

——. 1988b. "The spoken Arabic of al-Thōr in al-Ḥasa". *Zeitschrift für Arabische Linguistik* 18.59–76.

——. 1990. "The spoken Arabic of al-Qaṭīf". *Zeitschrift für Arabische Linguistik* 21.63–70.

——. 1991. "Notes on the spoken Arabic of Tihāmat Beni Shihr". *Zeitschrift für Arabische Linguistik* 23.99–101.

Rabin, Chaim. 1951. *Ancient West-Arabian*. London: Taylor's Foreign Press.

Schreiber, Giselher. 1970. *Der arabische Dialekt von Mekka: Abriß der Grammatik mit Texten und Glossar*. Ph.D. diss., University of Münster.

Sieny, Mahmoud Esma'il [Maḥmūd ʾIsmāʿīl Ṣīnī]. 1978. *The syntax of Urban Hijazi Arabic (Saʿudi Arabia)*. London: Longman; Beirut: Librairie du Liban.

Sowayan, Saad A. 1992. *The Arabian oral historical narrative*. Wiesbaden: O. Harrassowitz.

Thomas, Bertram. 1932. *Arabia Felix*. London: Jonathan Cape.

Bruce Ingham
(School of Oriental and African Studies)

Ṣawt → Sound Symbolism

Script → Arabic Alphabet: Origin

Script and Art

The 28 (or 29, if we include the *lām ʾalif*) letters of the Arabic alphabet, as we know them today, developed from a primitive set of 17 basic letter shapes (graphemes), which included a number of same-looking forms (homographs; → Arabic alphabet: origin). These letter forms at the beginning of Islam had nothing in themselves that would indicate their future grand place in Islamic art. And yet, within perhaps several decades after the birth of Islam, the Arabic script began to take on qualities which later in the Abbasid period blossomed into beautiful handwriting used to adorn the pages of the *Qurʾān* and religious buildings. Although the script was still in some ways defective because it did not distinguish its various letter-forms properly through a system of diacritical points, and of course it had no vocalization, nevertheless, the shape of these letters and the way they were joined together to form lines of writing began to acquire progressively greater elegance.

This can be seen clearly when admiring the inscriptions containing the earliest known dated passages from the *Qurʾān*, on the Dome of the Rock in Jerusalem, constructed in 72/691–692 and credited to the Umayyad ruler ʿAbd al-Malik (→ epigraphy). What was the reason for this stunning development of beautiful writing, which the Arabs would later call *ḥusn al-xaṭṭ* (or *ḥusn al-kitāba*)?

The answer lies in the Muslim belief that God revealed Himself to Muḥammad, through the agency of the archangel Gabriel, by dictating to him a series of revelations which were later recorded and assembled in book form, the *Qurʾān*. To the believer, the recorded word of God naturally required a proper attitude of reverence and the desire to copy it in a manner and script worthy of the revelation. Hence, generations of Muslims, from that time on, paid special attention not just to the copying and embellishment of *Qurʾān*s but also to any form of writing as having a 'heavenly' connection.

At the end of the Umayyad period and the beginning of Abbasid rule (132/750), a very different picture emerges. With the change of the dynasty there was a change (albeit perhaps gradual) in the format of the book (from vertical to horizontal) and the introduction of large and heavy scripts. The reason for this radical change may have been either an apologetic dimension, or a desire to make a clear distinction between the text of the *Qurʾān* and other non-Qurʾānic texts, principally the *Ḥadīṯ*.

The numerous surviving *Qurʾān* fragments of the early Abbasid period with large and hieratic scripts testify to this endeavor. These impressive *Qurʾān*s were not made to be read but rather to be admired and cherished. The beauty of the script, enhanced by the oblong format of the codex and surrounded by multicolored geometrical and vegetal decoration, shone from the pages of these *Qurʾān*s and reflected the newly acquired confidence of an expanding religion (→ Kufic).

In this early period, we encounter a variety of old scripts used in the chancery and for the copying of books, the most important of which are the Early Abbasid scripts (the 'Kufic' family, to use the old, but misleading, appellation), the New Abbasid Style (NS), also referred to today as 'broken cursive', which had its heyday in the 4th/10th century, and the → *maġribī* family of scripts toward the end of this period.

The constant preoccupation with script and decoration, rather than with images of humans and animals, discouraged by theologians, meant that calligraphy became the most important form of art in Islamic civilization.

Indeed, calligraphy was so tightly bound with the religion of Islam that Annemarie Schimmel (1979:177–178) called it "an art which can be called without exaggeration the 'quintessential' art of Islam". Calligraphers saw this connection between calligraphy and faith reinforced by the fact that the numerical value of the full text of the profession of faith (*šahāda*), namely 691, was the same as that of the word *xaṭṭāṭ*, and that the numerical value of the second part of the *šahāda* (454) corresponded to the value of the letters in the word *al-kātib*, both terms used for calligraphers (Vernoit 1997:17).

Apart from the desire to copy the *Qur'ān* in elegant scripts, the promotion of formal writing and later developments in calligraphy were often instigated in the administrative apparatus of the state, principally the chancery. It was the secretaries (*kuttāb*) who were behind many innovations and even the creation of new styles and scripts. The early manuals on penmanship originated from the chancery or were written for the benefit of those engaged in drafting and copying of letters and documents. According to Arabic tradition, the major reform of writing initiated by the Abbasid vizier Ibn Muqla (d. 328/940) came from the state apparatus. A new calligraphic tradition, with a set of new scripts, quickly developed in the eastern part of the Muslim world around the figures of three outstanding masters: Ibn al-Bawwāb (d. 413/1022), Yāqūt al-Mustaʿṣimī (d. 698/1298), and Ḥamd Allāh al-ʿAmāsī (d. 926/1520).

Toward the end of the early Islamic period (beginning of the 4th/10th century), the calligraphic field began to be completely transformed. Arabic tradition links this period to Ibn Muqla and the introduction of 'proportioned' writing (*al-xaṭṭ al-mansūb, al-kitāba al-mansūba*), perhaps a new fashion under the impulse of the chancery, although there are no extant specimens of writing that can be confidently attributed to Ibn Muqla; in fact, even the authenticity of his treatise on calligraphy is questioned. The new scripts which emerged toward the end of the 4th/10th century are → *ṯuluṯ*, *tawqīʿ*, and *riqāʿ*, on the one hand, and → *muḥaqqaq*, → *nasx*, and *rayḥān*, on the other. Their standardization is attributed to Ibn al-Bawwāb.

The 4th/10th century was, indeed, a period of great and radical change in the way Qurʾānic and non-Qurʾānic manuscripts were copied. Ibn Muqla's reform of writing, whether it concerned the 'standardization' of the New Abbasid Style ('broken cursive', also likened in old sources to *muḥaqqaq*, otherwise known as *warrāqī* or *ʿirāqī* script) or not, was a reality, whether prompted by political or other considerations (Tabbaa 2001:25–52; Blair 2006:173–178).

Apart from this, it is evident from the surviving manuscripts of the late 4th/10th century and early 5th/11th century, and principally the Chester Beatty copy of the *Qur'ān* executed by Ibn al-Bawwāb in 391/1000–1001 (Rice 1955:19–22), that a completely new picture emerged during that era; in the later calligraphic tradition, this style came to be associated with the so-called 'proportioned' writing.

The Chester Beatty *Qur'ān* of Ibn al-Bawwāb exhibits a high quality of calligraphic performance using at least two new distinct scripts, one for the main text and the other for chapter headings. These and other 'proportioned' scripts are grouped in later literature into two main families: rectilinear (*muḥaqqaq*) and curvilinear (*ṯuluṯ*). All scripts here fall within the following categories: large (*jalīl, ṯaqīl*), medium, and small (*daqīq, xafīf, laṭīf*), and the format of the writing surface is connected to the size of the script, i.e. large format/large script, small format/small script.

The end of the early middle period, some two centuries after the introduction of the new 'proportioned' scripts (late 7th/13th century), witnessed the standardization of two other regional scripts, namely *taʿlīq* (Iran) and *bihārī* (India).

Although the calligraphic tradition traces its roots to Ibn Muqla and Ibn al-Bawwāb, a significant split occurred, perhaps as early as the 7th/13th century. This was the time when Yāqūt al-Mustaʿṣimī, the third major figure in calligraphy, came onto the scene. Whereas in the Mamluk tradition scripts were grouped into 'Five Pens' or 'Seven Pens' (*al-ʾaqlām al-ʾuṣūl*), with a clear distinction between the curvilinear (*ṯuluṯ, tawqīʿ, riqāʿ*) family and the rectilinear (*muḥaqqaq, maṣāḥif, nasx, rayḥān*) family, the Yāqūtī tradition of the 'Six Pens' (*al-ʾaqlām as-sitta*; later followed by the Persian and Ottoman calligraphers) paired these scripts in the following manner: *ṯuluṯ/nasx, muḥaqqaq/rayḥān*, and *tawqīʿ/riqāʿ*. The Six Pens tradition survived into the printing age and is being practiced by modern calligraphers.

A century later (8th/14th century) marks the emergence of another major regional script, → *nastaʿlīq*, the Persian script par excellence. And from the 10th/16th century onward (late Islamic period), we have the standardization of regional forms of *nasx* (particularly Ottoman Turkish and Persian) and the emergence of typically Ottoman scripts such as *dīwānī* and → *ruqʿa*.

This great preoccupation with beautiful writing naturally produced a wealth of literature on the subject in all three languages, Arabic, Persian, and Ottoman Turkish. This corpus of literature includes chapters, passages, and statements in various works, as well as complete compositions. In the Arabic language alone, some 33 works on penmanship have already been published in one form or another, and perhaps as many more are known to have survived in manuscript. The earliest extant compositions of interest were composed by such authors as ʿAbdallāh al-Baġdādī (fl. 3rd/9th century), Ibn Qutayba (fl. 3rd/9th century), Ibn Muqla (d. 328/940), aṣ-Ṣūlī (d. 335/946), ʾAḥmad an-Naḥḥās (d. 339/950), Ibn Durustawayh (d. 347/958), Ibn an-Nadīm (d. 380/990), Ibn al-Bawwāb (d. 413/1022), Ibn Xalaf an-Nayramānī (d. 414/1023), Muḥammad al-Ḥumaydī (d. 488/1095), ʾAbū Ḥayyān at-Tawḥīdī (d. after 400/1009), and Ibn Sīd al-Baṭalyawsī (d. 521/1127; Gacek 2004).

Later in the Mamluk period (648/1250–923/1517), a number of authors greatly augmented this corpus of literature by the most important works on calligraphy produced in the Arab world. Here, twelve authors stand out in particular: Ibn al-Waḥīd (d. 711/1311 or 1312), an-Nuwayrī (d. 733/1333), al-Kātib ad-Dimašqī (fl. 781/1379), Ibn al-Baṣīṣ (fl. 8th/14th century), aṣ-Ṣaydāwī (late 8th/14th century), al-Maqdisī (late 8th/14th century), az-Ziftāwī (806/1403 or 1404), al-ʾĀṯārī (d. 828/1429), al-Qalqašandī (d. 821/1418), Ibn aṣ-Ṣāʾiġ (d. 845/1441 or 1442), al-Hītī (d. 891/1486), and aṭ-Ṭayyibī (Ṭībī) (fl. 908/1502 or 1503; Gacek 2004).

Most of this literature still remains to be properly explored and analyzed, and although critical editions of many of these texts have yet to be established, even at this stage we can extract useful data from them for the study of various scripts, their relationship to one another, and their appellations. It is, indeed, thanks to the rich Mamluk literature that we have a better picture of the practice of calligraphy in that period (Déroche and Berthier 2000:213; Gacek 1987, 1989a).

All throughout the manuscript age, a great number of scripts developed. Already in the early Abbasid period, they were grouped into three categories: Qurʾānic scripts (*xuṭūṭ al-maṣāḥif*), scribal hands or bookhands for texts other than the *Qurʾān* (*xuṭūṭ al-warrāqīn*), and chancery (secretarial) hands (*xuṭūṭ al-kuttāb*). These major classifications remained, broadly speaking, valid throughout the manuscript period (Gacek 1989a).

Thus, for instance, the most often used scripts for the copying of the *Qurʾān* in the later middle period (roughly from the mid-7th/13th century to the 9th/15th century) were → *muḥaqqaq*, *rayḥān*, and *nasx*. As far as we know, *muḥaqqaq* was never used for the copying of non-Qurʾānic manuscripts. Among other scripts belonging to this category are New Abbasid Style (or 'broken cursive'), → *maġribī*, and *bihārī*.

A group of scripts known as *xuṭūṭ al-kuttāb* were traditionally associated with the chancery and the state administration. They include → *ṯuluṯ*, *tawqīʿ*, and *riqāʿ*, as well as *taʿlīq*, *dīwānī*, and → *ruqʿa*. The *ṯuluṯ* was also used for decoration on hard surfaces and for monumental inscriptions. The *xuṭūṭ al-warrāqīn* family included such scripts as → *nasx* (with its variants), → *maġribī*, and → *nastaʿlīq* (in the Persianate world).

The calligrapher occupied a very prominent position in Muslim society, although this did not necessarily translate into adequate remuneration for the service rendered. He was also often an illuminator (limner) or decorator. This fact demonstrates how closely these two arts, calligraphy and painted decoration, were connected. Although there is evidence of women involved in calligraphy, this was a predominantly male profession, so much so that ink was described as the perfume of men (Rosenthal 1961:18). The calligrapher's most important implement was the reed pen (calamus; *qalam, jazm, mizbar, mirqam*). The pen, ink, and other writing implements were often extolled in Islamic literature. Thus, for instance, Ibn an-Nadīm (d. 380/990) quotes Ibn ʾAbī Dāʾūd as saying, "The pen is the ambassador of the mind, its messenger, its furthest reaching tongue, and its best interpreter" (*Fihrist* 19; Rosenthal 1961).

Just like recipes for making ink, the cutting of the nib was a well-kept secret. Furthermore, just as there was a direct relationship between the format of the writing surface and the script, there was an interdependence between the width of the nib, the manner of cutting it, and the calligraphic style or script. In the eastern part of the Muslim world, the point of the nib was usually cut either straight (*al-qaṭṭ al-mustawī* or *al-mudawwar* or *al-murabbaʿ*) or at an angle, obliquely (*al-qaṭṭ al-muḥarraf*; Gacek 2001).

The effect of writing with an obliquely cut nib was to produce thinner strokes (*farakāt*) at angles, as well as thinner shafts (*muntaṣabāt*) in such letters as the *ʾalif* and the *lām*. It is said that Ibn al-Bawwāb wrote all scripts with a straight-cut nib, whereas Yāqūt al-Mustaʿṣimī used an obliquely cut nib for all scripts. Others varied the cut according to scripts. Thus, the scripts in the *muḥaqqaq* family were often executed with pens cut obliquely.

The oblique angle could differ greatly depending on the personal preference of a scribe or calligrapher. In the oblique cut, it was usually the right half-nib which was elevated, but, surprisingly, some scribes were in the habit of cutting the nib in such a way that the left half-nib was higher than the right half-nib. *Maġribī* scripts, however, were executed with a pen whose nib was pointed, and, therefore, its pen strokes were of the same thickness.

With time, primitive Arabic script developed a large variety of letter forms (allographs), some of them very characteristic of individual scripts. The rich terminology found in the literature on penmanship can be very helpful in the paleographical analysis. The early authors classified letter forms into various categories. Thus, for instance, the 5th/11th-century author Muḥammad al-Ḥumaydī (*Tashīl* 23–25) divides them into four distinct categories: *ʾaṭnāb* (those with ascenders), *ʾahdāb* (those with descenders), *nawājiḏ* (those with 'teeth'), and *maḥājir* (those with counters; see also Gacek 2001). The descenders are sometimes divided into *muʿarraqa* (the tail turned left) and *muʿaqqafa* (the tail turned right – as in the letters *jīm* and *ʿayn*). They can be flattened or tapered (*mabsūṭa*) or rounded (*muqawwara*, *muqawwasa*, *murattaba*). Those with counters can either be open (*maftūḥa, munawwara*) or closed (*maṭmūsa, muʿawwara, muʿammā*; Gacek 2001).

The most important letter was the *ʾalif*, often likened to the figure of a man. The shape or form of this letter is often the best clue as to the identity of the script. Throughout the manuscript age, the isolated *ʾalif* was written slanted to the right or left, very straight – almost vertical, bowed, seriffed or serif-less (sans serif; both head and foot), and the like. In the proportioned scripts, there was a direct relation, not only between the *ʾalif* and the other letters but also between its length and its thickness. Thus, if the *ʾalif* was seven rhombic dots in height, its thickness was 1/7, and if it had a head-serif, its length was also 1/7 of its height (Qalqašandī, *Ṣubḥ* III, 24, 47).

Another, perhaps the second most important, letter was the *lām ʾalif*. One of the most prominent ligatures in the Arabic script, the *lām ʾalif* was regarded as the 29th letter of the alphabet and was traditionally placed before *yāʾ* in the alphabet sequence. There were three main forms of this letter: *al-lām ʾalif al-muḥaqqaqa* – the *lām ʾalif* ligature characterized by an open (*maftūḥ*) loop (counter) at its base; *al-lām ʾalif al-muxaffafa* – the *lām ʾalif* ligature characterized by joining the foot of the *ʾalif* to the extremity of *lām* on the base line (ﻼ); and *al-lām ʾalif al-warrāqiyya* – the *lām ʾalif* ligature characterized by its triangular base (لا) and used in the Mamluk period exclusively in the rectilinear family of scripts: *muḥaqqaq*, *maṣāḥif*, *rayḥān*, and *nasx* (Gacek 2001:154).

The flexibility of the Arabic script and its preeminence in Islamic culture meant that it was used not just on the traditional soft writing surfaces such as papyrus, parchment, paper, and textiles (fabrics) but also, and from the earliest time onward, on hard surfaces, such as stucco, tile, and metal. Elegant inscriptions can be found on pillows, curtains, garments, belts and kerchiefs, flags and tents, as well as on golden and silver vessels, ivory, porcelain, ceramics, and helmets and swords, and even on large dry leaves (Schimmel 1970:10; 1984:25–26). Furthermore, superb *nastaʿlīq* calligraphy can be seen on coins and seals from Mughal India (Blair 2006:554–555).

Decorative arts also included decoupé work (*découpage*), invented in Herat in the 9th/15th century, which involved cutting out letters with scissors and placing them on colored paper (Schimmel 1979:10).

One of the earliest architectural inscriptions in 'Kufic' script is to be found on the Dome of the Rock. Calligraphers often exaggerated or stylized the features of the original script. Ornamental 'Kufic' became an important part of Islamic decorative arts as early as the beginning of the 2nd/8th century. The most important decorative styles of this script, which developed mostly, but not exclusively, on hard surfaces, are variously described as square, foliated, floriated, plaited, knotted, interlaced, and animated (Safadi 1978:11–12). The New Style scripts, dating from the 4th/10th century and later, also underwent this kind of stylized transformation. A good example here is the so-called Qarmatian *Qurʾān* or a copy of *Kitāb at-tiryāq* by pseudo-Galen, transcribed in 595/1199 (Blair 2006:198–200; Safadi 1978:12–13).

In addition to fancy and stylized scripts, Arabic decorative arts include zoomorphic calligraphy featuring both human and animal figures, which dates back to the 9th/15th century. The most popular animal to be formed by letters was the lion representing the Imam ʿAlī ibn ʾAbī Ṭālib, whose nickname was 'Lion of Allah'. In Turkey, the stork was popular, as it was considered a pious bird. Another one was the parrot, symbol of sweet talk and, because it was green, connected with paradise. The texts included the *basmala*, confession of faith, the names of the 'Seven Sleepers', etc. (Safadi 1978:31; Schimmel 1984:52, 1970:11–12; Blair 2006:449–451).

In contrast to monumental inscriptions, calligraphers sometimes employed micrography, or minute writing. The best-suited script for this was *ġubār* 'dust'. Originally used for pigeon post and secret messages, this script was variously referred to as *qalam al-janāḥ* or *qalam al-baṭāʾiq*, as well as *qalam al-ḥalba* (or *al-ḥilya*; Gacek 1989a:145). Mamluk authors were divided on its origin. Some authors, such as an-Nuwayrī, viewed it as a smaller (*xafīf*) version of *riqāʿ* script, just as *ḥawāšī* is a smaller version of *nasx*. Al-Qalqašandī (*Ṣubḥ* III, 48, 128), on the other hand, states that it was derived from both *riqāʿ* and *nasx* and that it is all curvilinear but without head-serifs (*tarwīs*). Al-ʾĀṯārī, on the other hand, regarded it simply as a smaller variety (*xafīf*) of *nasx* (Gacek 1989a:45). In later calligraphic circles, its derivation from *nasx* is almost a given. In Iran, a very small *nasx* is known as *nasx-i ġubār* (Schimmel 1984:25; Blair 2006:259–260).

Micrography was used for miniature *Qurʾān*s, often octagonal in shape, for amulets in the form of rolls/scrolls, and for a type of large calligraphy where single letters, as well as figures of human beings, animals, and flowers, were filled with minute writing, usually pious formulae, a technique known as *gulzār* 'garden of roses' (Safadi 1978:31; Blair 2006:451–454). Modern calligraphers have even attempted to write on such objects as a grain of rice or wheat or the shell of a hen's egg, and recently the entire text of the *Qurʾān* was fitted onto a page measuring no more than 55 × 45 cm (Safadi 1978:30).

After the standardization of the 'Six Pens', there was a corresponding increase in the production of calligraphic specimens (*qiṭʿa*), albums (*muraqqaʿ*), alphabetic exercises, practice sheets, calligraphic panels (*lawḥa*), and the like (Blair 2006:497–508). The Ottoman calligraphers were especially fond of calligraphic specimens, which they used for the granting of diplomas or licenses (*ʾijāzāt*). Indeed, the official diploma, although having its origins in the Mamluk period, became very popular with Turkish calligraphers, and numerous specimens of them have survived in various collections around the world (see, for example, Gacek 1989b).

Other forms of fancy scripts and techniques include the *ṭuġrā*, originally the signature of the ruler placed at the beginning of the document, but later, and especially in the Ottoman period, an elaborate device and a very sophisticated calligraphic motif. The word came to mean any kind of unusual joining of words into decorative shapes, representing trademarks or the signs of societies (Schimmel 1970:46).

Zulf-i ʿarūs 'the bride's lock of hair' is a decorative style, perhaps a hybrid of *rayḥān* and *nastaʿlīq*, which resembles tiny curls. *Muṯannā*, also known as *aynalī* or *maʿqūs*, is the art of mirror writing (Safadi 1978:30–31). First employed for seal cutting, this calligraphic technique is known to have been used as early as the 9th/15th century. It flourished especially in Turkey in the 13th/19th century. *Xaṭṭ-i nāxun* is fingernail writing, a technique of 'engraving' the text with the fingernail on the back of the paper. It was invented or made popular in the

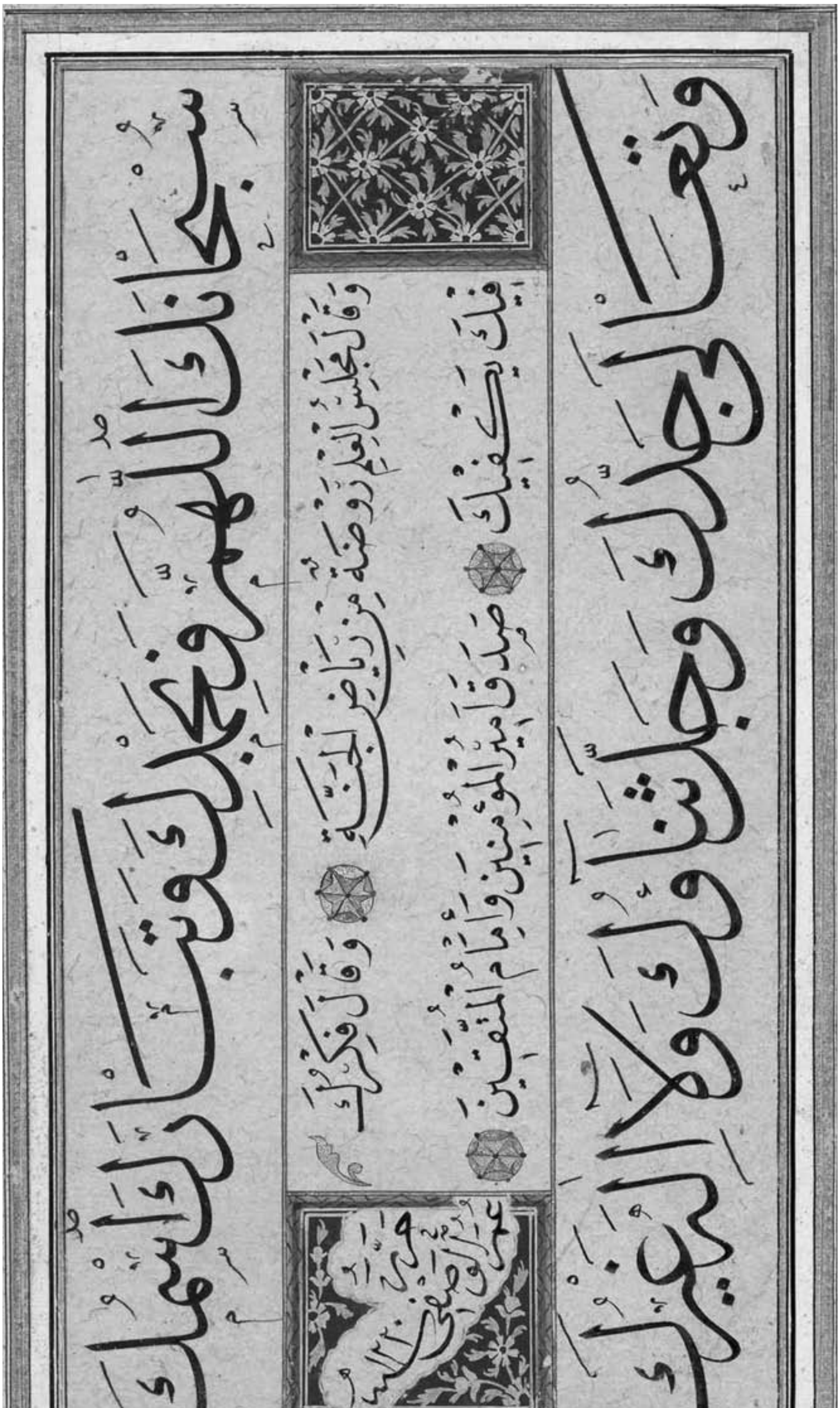


Figure 1. A leaf from an Ottoman album of calligraphy (*muraqqaʿ*) executed in *ṯulu*ṯ and *nasx* scripts by ʿUmar al-Waṣfī and dated 1220/1805.

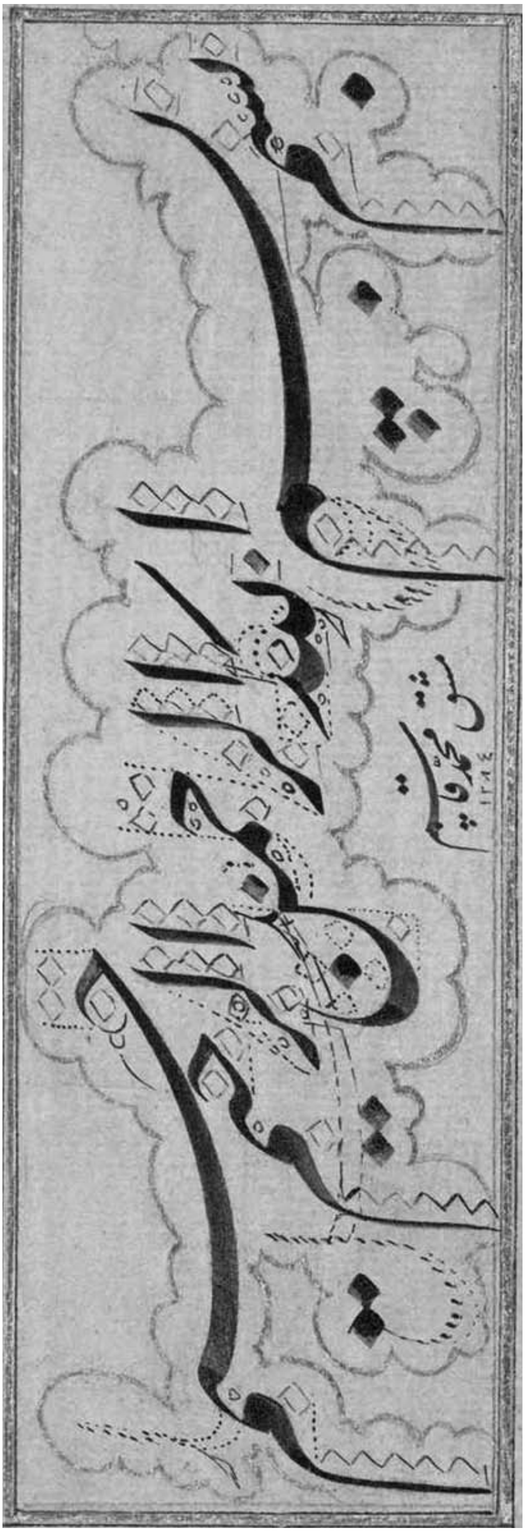
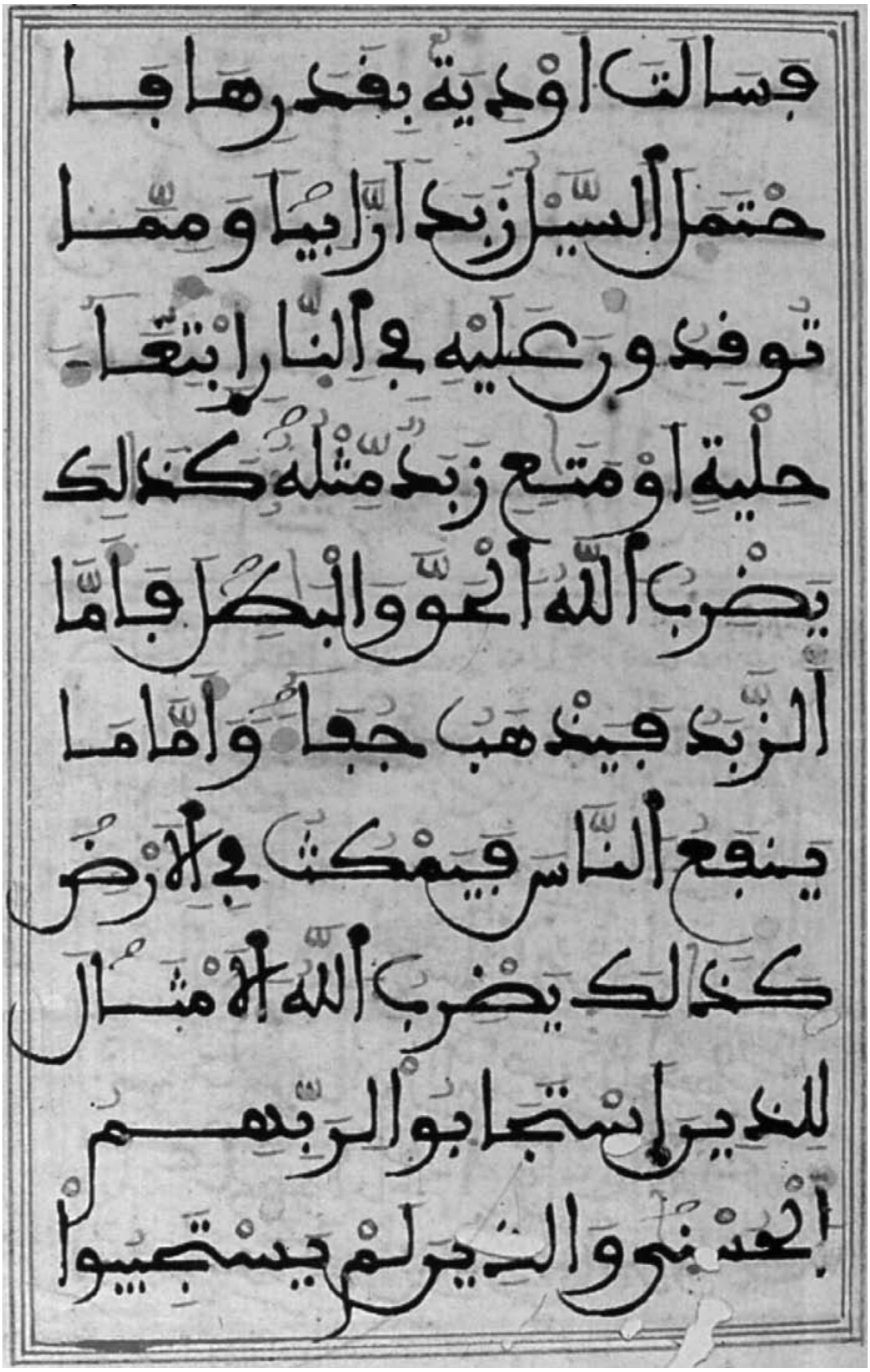


Figure 2. A calligraphic panel featuring the *basmalah* in *nastaʿlīq* script penned by Muḥammad Qāsim [Tabrīzī] in 1284/1867–8. (McGill library, RBC AC 158)



فسالت اودية بقدرها فا
حتمل السيل زبدا رابيا ومما
توقدون عليه في النار ابتغا
حلية او متع زبد مثله كذلك
يضرب الله الحق والبطل فاما
الزبد فيذهب جفا واما ما
ينفع الناس فيمكث في الارض
كذلك يضرب الله الامثال
للذين استجابوا لربهم
الحسنى والذين لم يستجيبوا

Figure 3. A page from a *Qur'ān* manuscript in *maġribī* script, dated 1144/1731. (McGill Library, RBD A21).

10th/16th century by Niẓām ad-Dīn Buxārī and is still practiced in Pakistan by a few artists (Schimmel 1970:11, 1984:32).

Arabic calligraphy survived into the printing age and was used extensively in such countries as Morocco, Egypt, Turkey, Iran, and India for printing by lithography. In fact, many books printed by lithography in the 13th/19th and the early 14th/20th century contain some of the best calligraphy of that period. Here, naturally, *nasx* predominates, but one also finds *maġribī*, *nastaʿlīq*, and *ṯuluṯ* family scripts used as display texts (Gacek 1996).

Today, Arabic calligraphy is practiced by a growing number of calligraphers not just in the Islamic world but also in the West. Many calligraphers follow the traditional methods of calligraphy, but many others also combine the traditional with the new, such as three-dimensional forms, calligraphic sculpture (*naḥt xaṭṭī*), and multimedia compositions. Among the practitioners of traditional calligraphy are such well-established calligraphers as Ḥabīb Allāh Faẓāʾilī (Iran), M.U. Derman (Turkey), Mohamed Zakariya (USA), Muṣṭafā Jaʿfar (originally from Iraq), ʿUṯmān Waqīʿ Allāh (Osman Waqialla; originally from Sudan), and Ahmad Moustafa (originally from Egypt), to mention just a few (Blair 2006:589–621).

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Ḥumaydī, *Tashīl* = ʾAbū ʿAbdallāh Muḥammad ibn Fattūḥ al-Ḥumaydī, *Tashīl as-sabīl ʾilā taʿallum at-tarsīl*. Facsimile ed. Frankfurt a. Main, 1985.

Ibn an-Nadīm, *Fihrist* = Muḥammad ibn ʾIsḥāq ibn an-Nadīm al-Warrāq al-Baġdādī, *Kitāb al-fihrist*. Ed. Yūsuf ʿAlī Ṭawīl. Beirut: Dār al-Kutub al-ʿIlmiyya, 1996.

Qalqašandī, *Ṣubḥ* = ʾAḥmad ibn ʿAlī al-Qalqašandī, *Ṣubḥ al-ʾaʿšā fī ṣināʿat al-ʾinšāʾ*. 14 vols. Cairo: al-Muʾassasa al-Miṣriyya al-ʿĀmma, 1963–1965. (First publ., aṭ-Ṭabʿa al-ʾAmīriyya', 1913–1918.)

SECONDARY SOURCES

Blair, Sheila. 2006. *Islamic calligraphy*. Edinburgh: Edinburgh University Press.

Déroche, François and Anne Berthier. 2000. *Manuel de codicologie des manuscrits en écriture arabe*. Paris: Bibliothèque Nationale de France. [English trans., *Islamic codicology: An introduction to the study of manuscripts in Arabic script*, trans. Deke Dusinberre and David Radzinowicz, ed. Muhammad Isa Waley. London: al-Furqān Islamic Heritage Foundation, 2006.]

Gacek, Adam. 1987. "Al-Nuwayrī's classification of Arabic scripts". *Manuscripts of the Middle East* 2.126–130.

——. 1989a. "Arabic scripts and their characteristics as seen through the eyes of Mamluk authors". *Manuscripts of the Middle East* 4.144–149.

——. 1989b. "Arabic calligraphy and the 'Herbal' of al-Ghāfiqī: A survey of Arabic manuscripts at McGill University". *Fontanus: From the collections of McGill University* 2.37–53.

——. 1996. *Arabic lithographed books in the Islamic Studies Library, McGill University: Descriptive catalogue*. Montreal: McGill University Libraries.

——. 2001. *The Arabic manuscript tradition: A glossary of technical terms and bibliography*. Leiden: E.J. Brill.

——. 2003. "The head-serif (*tarwīs*) and the typology of Arabic scripts: Preliminary observations". *Manuscripta Orientalia* 9:3.27–33.

——. 2004. "Scribes, amanuenses, and scholars: A bibliographic survey of published Arabic literature from the manuscript age on various aspects of penmanship, bookmaking, and the transmission of knowledge". *Manuscripta Orientalia* 10:2.3–29.

Rice, D.S. 1955. *The unique Ibn al-Bawwāb manuscript in the Chester Beatty Library, Dublin*. Dublin: Emery Walker.

Rosenthal, Franz. 1961. "Significant uses of Arabic writing". *Ars Orientalis* 4.15–23.

Safadi, Yasin Hamid. 1978. *Islamic calligraphy*. London: Thames and Hudson.

Schimmel, Annemarie. 1970. *Islamic calligraphy*. Leiden: E.J. Brill.

——. 1979. "Poetry and calligraphy: Thoughts about their interrelation in Persian culture". *Highlights of Persian art*, ed. Richard Ettinghausen and Ehsan Yarshater, 188–209. Boulder, Colo.: Westview Press.

——. 1984. *Calligraphy and Islamic culture*. New York and London: New York University Press.

Tabbaa, Yasser. 2001. *The transformation of Islamic art during the Sunni revival*. Seattle: University of Washington Press.

Vernoit, Stephen. 1997. *Occidentalism: Islamic art in the 19th century*. London and Oxford: The Nour Foundation [in association with Azimuth Editions and Oxford University Press] .

ADAM GACEK (McGill University)

Second Language Acquisition

1. INTRODUCTION

While the pedagogy of foreign language teaching (→ second language teaching) is concerned with the various approaches, methods, and techniques of how a second/foreign language is taught, the field of Second Language Acquisition (SLA) is concerned with how a second/foreign language (L2) is actually learned. Other terms are also used to refer to the same phenomenon, including L2 acquisition, second language development (L2 development), and

L2 learning. SLA focuses on the development of the learner's language, whether in a formal (i.e. a classroom) or a naturalistic setting. This language is known as interlanguage (IL), and as a natural language, it is subject to systematic development formally (i.e. in a classroom setting) or naturalistically, and is subject to systematic development either toward approximation of the target language (TL) or toward a fossilized, non-target-like state. The scope of SLA investigation includes learning factors, processes, and strategies. Although there is a consensus among researchers that no single factor can account for a complex human phenomenon such as second language acquisition, the exact role of any given factor is still subject to debate. These factors are varied. Some are considered internal (linguistic) to the acquisition process, others external (nonlinguistic). Internal factors include such phenomena as transfer, or the influence of the learner's first language, input and input frequency, and structural and processing complexity. External factors are more controversial and are harder to quantify; they include such factors as age, aptitude, attitude, anxiety, motivation, learning environment, social distance, and ethnicity.

Early SLA studies focused heavily on informing foreign language pedagogy. This emphasis shifted gradually, though not exclusively, toward more interdisciplinary research with implications, both theoretical and applied, that extend beyond language pedagogy. This was partly due to interaction with other disciplines, including theoretical linguistics, historical linguistics, psychology, psycholinguistics, neurolinguistics, language pathology, language testing, sociology, communication, etc. The subsequent discussion of Arabic SLA studies is organized chronologically following the development of the field.

2. Contrastive analysis and error analysis

Contrastive Analysis studies comprised the earliest formative period of SLA research (between the 1940s and 1960s), espousing a behaviorist view of language learning. Identifying learners' errors, predicted on the basis of differences between the native language (L1) and the target language (TL), was important to avoid bad 'habit formation'. Error Analysis studies started to appear with the advent of nativist and cognitive views of language learning as a process of 'rule formation'. Errors were no longer explained on the basis of differences between the native and target language systems, as was done during Contrastive Analysis days. Error Analysis studies were based on the output production of the learner, both written and verbal. A distinction was made between 'errors' and 'mistakes', the former being systematic and inevitable (competence-based) and the latter being random (performance-based; Corder 1967). A plethora of categories of errors was introduced, including interlingual errors (caused by L1 interference), intralingual errors (made regardless of L1), developmental errors (caused by the learner's constructed hypotheses of L2 systems according to extent of exposure to L2), overgeneralization errors, simplification errors, and induced errors (caused by instructional lapses or errors), among others. A major contribution of error analysis is the concept of the L2 learner's language or *interlanguage*, coined by Selinker (1972). It is usually defined as a continuum between L1 and L2, reflecting a systematic development of L2 by the learner, as well as learners' variations according to different learning experiences.

At least two studies on Arabic were conducted within Error Analysis framework. Al-Ani (1972–1973) analyzed 'major' errors in a 'limited' number of written composition assignments of adult (American) English speakers enrolled in an advanced Arabic course. The exact number of written samples was not reported. There was no attempt to analyze all written errors in the students' compositions and no attempt to offer any statistical measure. Errors identified were roughly categorized along three levels: orthographic and phonological (or spelling) errors, diction and dictionary usage, and grammatical errors. Al-Ani attributed most errors to L1 interference errors. These included errors such as the undersuppliance of the definite article with Arabic abstract nouns, as well as gender and number agreement mismatches. Other errors, such as the affixation of the definite article on proper names, were considered overgeneralization errors, and others were regarded as performance errors (e.g. affixing the article on the head noun in *ʾiḍāfa* constructions). Some were attributed as performance errors (such as the gender agreement error in *al-ḥarb* [fem.] *al-qāsī* [masc.] ==> *al-ḥarb* [fem.] *al-qāsiya* [fem.] 'the severe war'), rather than competence errors. Reflecting some of the limitations of Error Analysis, Al-Ani concluded that it was not always easy to categorize an error or identify its source.

Rammuny (1976) provides a more extensive study that investigated errors of adult native speakers of English learning Arabic as a second language. He analyzed data of an Arabic proficiency written test of 115 American university students with formal exposure to Arabic (between two and six years). Unlike Al-Ani's study, all errors were analyzed except those occurring fewer than five times. Four main categories of errors were identified, similar to those identified in Al-Ani's study: orthographic and phonological, lexical, structural (including Noun-Adjective agreement, Verb-Subject concord, preposition use, definiteness, plural forms, interrogatives, negation, conditionals, etc.), and stylistic errors. The categories were then grouped along four error causes, including inefficient 'teaching-learning strategies' (or induced errors in Error Analysis terms), 'interference' of L1, 'unfamiliarity' (or competence) and 'sociopsychological' issues (or performance). Accordingly, of the total (578) number of errors identified as structural errors, 406 tokens were classified as induced, 145 as L1 interference, 4 as competence, and 23 as performance errors. Of the total number of errors (1,520), 49 percent are induced errors, 27.9 percent L1 interference errors, 16.9 percent competence errors, and 6.2 percent performance errors. In both of the above studies, the number of tokens of correct rule application in the four categories is missing. Hence, only a partial account of the L2 learners' interlanguage systems is provided. However, this is a characteristic limitation of Error Analysis methodology in general.

3. PERFORMANCE/DEVELOPMENTAL STUDIES

The motivation for a new, more comprehensive approach aimed at investigating the overall performance of L2 learners (usually abbreviated as L2ers) soon arose. The approach of detecting consistent patterns in the productions of L2 learners, not just their errors, became popular in the 1970s and gave rise mainly to morpheme order, developmental sequences, and learning and communication strategies studies (see also Larsen-Freeman and Long 1991).

3.1 *Morpheme Order studies*

Morpheme Order acquisition studies attempted to investigate the claim that the acquisition process was guided by a universal internal mechanism, following Roger Brown's (1973) and other subsequent L1 studies which found a set of (14) grammatical morphemes acquired in the same order by children from different languages. The idea was particularly appealing to SLA researchers, since such findings would indicate the presence of a 'built-in syllabus', regardless of L1 backgrounds and exposure to L2 (Corder 1967). Morpheme Order studies analyzed children and adult SLA data for acquisition of a set of English grammatical morphemes, such as pronoun case, the definite and indefinite articles, singular copular verb, the progressive [-ing], regular plural, singular auxiliary, regular and irregular past tense, 3rd person present tense, and possessive [-'s]. The general findings indicate that language acquisition involves a process of creative construction rather than habit formation (Dulay and Burt 1973, 1974), that first and second language acquisition involve similar strategies and processes (also known as the L1 = L2 hypothesis; Bailey and Krashen 1974), and that there seems to be a 'natural order of acquisition' in SLA, irrespective of L1 background and L2 learning environment (e.g. Krashen 1981).

3.2 *Developmental sequence*

Interest in Morpheme Order subsided due to methodological limitations and as interests shifted toward identifying developmental sequences/stages and processes which, in mastering a given structure, an L2 learner goes through from the pre-target-like stage to that of target-like mastery (e.g. Wode a.o. 1978). Like previous studies, developmental investigations in SLA were triggered by L1 acquisition studies to examine whether or not the sequences in L1 and L2 acquisition were the same (→ language acquisition, first). Structures investigated include WH-questions, negation, relative clauses, and word order. For example, the acquisition of English negation was examined in L2 learners of different L1 backgrounds and was found to be acquired along predictable stages (similar to those found in L1): (i) use of the negation *no* in utterance-initial position; (ii) with *no/not* in internal preverbal position;

(iii) with *no* occurring with modals and auxiliaries as unanalyzed chunks; and (iv) with the negative particle used differentially from auxiliaries and models (e.g. Schumann 1979).

Al-Buainain (1986) investigated the acquisition of negation and interrogation by 53 adult native (British) English speakers learning Arabic as an L2. The participants were at five different proficiency levels ranging from one to five years, resulting in five groups. The cross-sectional data consisted of the participants' written performance on a translation task and a manipulation task. Due to limitations of the elicitation tasks, the study's main findings resulted in different acquisition substages within each of the two target structures. In negation, for example, three statistical measures produced three different acquisition orders. Two of these resulted in the following implicational orders: (i) *lā* ==> *lam* ==> *lan* ==> *laysa*, based on participants' performance on the translation task; and (ii) *lā* and *lam* ==> *lan* and *laysa*, based on the manipulation task. Al-Buainain also offered a brief qualitative analysis of negation based on translation data. For example, the acquisition of *lā* is explained to proceed along the following substages: (i) correct suppliance of *lā* in inappropriate position; (ii) correct placement of *lā* with improper order of the remainder of the sentence; (iii) redundant production of subject pronoun; and (iv) correct production and placement of *lā* and the verb following it. Al-Buainain's findings lend support to the general observation in developmental sequence studies that L2 acquisition generally progresses along intermediate stages, starting from highly deviant and simple interlanguage forms to more acceptable variants as the interlanguage system of the L2 learner approximates to the target language.

3.3 *Learning and communication strategies*

In the spirit of Developmental Studies and in an attempt to account for L2 variation along the interlanguage continuum, SLA researchers examined two different categories of strategies adopted by L2 learners: learning and communication strategies. Learning strategies are ways or methods that learners follow to improve their learning skills in 'perceiving', organizing, storing, manipulating, and recalling information. Communication strategies, on the other hand, are methods of 'achieving communication' and of 'encoding meaning' by L2 learners, given their limited knowledge of the target language (H. Brown 1980:83). These strategies include transfer from L1, generalizations, avoidance, circumlocution, message adjustment, prefabricated patterns, and code switching (e.g. Tarone 1980).

Fakhri (1984) examined the notion of communication strategies based on spontaneous data recorded over a month from a female adult English speaker learning Moroccan Arabic as an L2. The participant had lived in Morocco for three years. Fakhri identified the 'most frequent' communication strategies used by the participant as being similar to those reported in the literature. They included circumlocution, lexical borrowing from L1, elicitation of vocabulary (i.e. asking the native speaker interlocutor for words in the target language), expanded use of formulaic expression (i.e. using learned phrases or expressions in contexts not used by L1 speakers, as in *meskīna!* 'poor girl' ==> **meskīna bezzāf!* 'very poor girl!'), and morphosyntactic innovation (i.e. interlanguage-like forms). Fakhri suggested that communication strategies were probably further constrained by components of the specific narrative genre. For example, the participant resorted to lexical borrowing (73%) during the episodic part of the narrative more often than she did in the orientation part (33%), whereas she resorted to circumlocution (18%) and vocabulary elicitations (9%) in the episodic part far less than she did in the orientation part (39% and 28%, respectively). Fakhri argues that due to urgency in the episodic part to convey meanings/events to an 'intrigued' listener, the subject in the study made the narration more effective by resorting more to lexical borrowings than to other available strategies.

The strategies approach has been mainly criticized for vagueness of definitions and concepts and for lack of independently motivated explanation as to whether communicative strategies adopted by the L2 learner in response to the limited interlanguage system are different from adjustments in normal L1 use to maintain real-time language processing and communication (Oxford and Cohen 1992; Bialystok 1990).

4. The current scene

With increasing emphasis on providing a fuller account of L2 performance, SLA researchers started to focus on the role of input in SLA. This triggered the incorporation of discourse analysis and the acquisition of pragmatics into SLA research, previously focused on acquisition studies of phonology, semantics, morphology, and syntax. Competing models subscribing to different views of learning were also introduced to the field.

The discussion in the remaining part of this section focuses on two accounts of L2 development that seem to have inspired recent studies of SLA in Arabic: the cognitive-interactionist model and Universal Grammar-based proposals.

4.1 *Processability Theory: A cognitive-interactionist model of L2 development*

Pienemann's (1998) Processability Theory explains L2 grammatical development from a cognitive-interactionist perspective. The model assumes that L2 learners create their own language-specific processing prerequisites or procedures of the L2, claimed to emerge along a set implicational sequence. In morphology, for example, the emergence of processing prerequisite procedures is explained in terms of the production of three types of morphemes along five main stages. Below is an explanation of the first four stages:

Stage 1	Absence of any language-specific procedures, for example words and phrases being produced as unanalyzed chunks
Stage 2	Development of category procedures and 'lexical morphemes', for example the {-ed} tense marker in English
Stage 3	Development of 'phrasal procedures' where exchange of grammatical agreement within a constituent phrase occurs, such as Noun-Adjective agreement
Stage 4	Development of 'interphrasal morphemes' where exchange of grammatical agreement across two constituent phrases occurs, as, for example, in Subject-Verb agreement

The difference between Morpheme Order studies and Processability Theory is that while the former is concerned with identifying mere acquisition order of forms, the latter is concerned with identifying processing mechanisms that determine acquisition orders. For a review of Processability Theory, see Alhawary (1999, 2003).

Nielsen (1997) attempted to test the predictions made by Processability Theory on two adult Danish speakers learning Arabic as an L2 at the beginning level. The data were collected longitudinally over 15 months. Elicitation tasks comprised oral interviews, role plays, and presentations. The study examined acquisition of definiteness agreement within the noun phrase, demonstrative pronouns, and Noun-Adjective and Subject-Verb agreement for singular masculine and singular feminine. The most significant findings revealed that neither Noun nor Adjective nor Subject-Verb agreement emerged in one of the two participants' interlanguage systems. Both agreement forms emerged at the same time in the other participant's interlanguage system. Thus, Nielsen's study provides counterevidence to Processability Theory assumptions regarding processing prerequisites. According to this theory, Noun-Adjective agreement should have been acquired first, then Subject-Verb agreement.

Alhawary (1999, 2003) also investigated the predictions of the speech processing hierarchy claimed by Processability Theory. The study was conducted within a longitudinal setting that followed eight adult native (American) English speakers and one French speaker learning Arabic as an L2 (at the beginner level) over one school year. Data elicitation tasks included picture description, picture differences, picture sequencing, video story retelling, and informal interviews. The study investigated a number of structures. The most significant findings relate to the acquisition of Arabic Noun-Adjective and Subject-Verb agreement, where the former is predicted to be acquired before the latter (for results related to other structures, see Alhawary 1999). The findings showed that as many as six of the participants acquired Subject-Verb agreement before they did Noun-Adjective agreement, even though the participants were exposed to Noun-Adjective agreement before Subject-Verb agreement. In addition, Alhawary (2003) analyzed the same data of both structures, using two different acquisition criteria (an emergence criterion and a 90 percent correct in obligatory context measure), both of which converged on the same findings. Accordingly, the data strongly challenge the Processability Theory claims regarding the hypothesized speech processing hierarchy as untenable. Alhawary (1999, 2002, 2003) concludes that there seem to be other processing factors, such as L1 transfer, not accounted for by Processability Theory.

Mansouri (2000) is a third study that investigated Processability Theory. It is based on cross-sectional data consisting of four adult native (Australian) English speakers learning Arabic as an L2, two of whom were at the beginning level and two at the intermediate level. Two data samples were collected from each at 13-week intervals. Elicitation consisted of an oral interview and an oral transformation task (story retelling about an actor in the dual and plural). The findings produced mixed evidence. In particular, a number of morphological developmental patterns were found to be inconsistent with Processability Theory, such as the emergence of relative pronouns before pronoun clitics. Mansouri claimed, however, that a number of other structures supported Processability Theory. For example, he claimed that Noun-Adjective agreement had already emerged before the first data set had taken place, and Subject-Verb agreement emerged later in the second data set, when the learners produced 3rd person plural agreement (after having produced agreement marking for 1st person singular and 2nd person singular in addition to 3rd person singular masculine and feminine). However, relying on a more reliable emergence criterion such as that adopted by Nielsen (1997) and Alhawary (1999, 2003) and focusing on 3rd person singular and feminine (since one would not expect L2 learners to acquire almost the *entire* agreement paradigm at an early stage as Mansouri 2000 rather assumed), Mansouri's data would show instead that Subject-Verb agreement had already emerged in the interlanguage systems of the two beginning learners before the first data session took place. But, if both Noun-Adjective and Subject-Verb agreement structures had already emerged prior to the first data collection session, it would be difficult to determine if one structure emerged before the other, given the cross-sectional semi-longitudinal nature of the study. Hence Mansouri's (2000) findings neither support nor contradict Processability Theory with respect to the two structures (for a more detailed review, see Alhawary 2003).

4.2 *Nativist/UG-based accounts of L2 development*

Since the 1980s, the greatest emphasis in accounting for L2 development has been within the generative (nativist/rationalist), Universal Grammar (UG) view of language acquisition, particularly within the Principles and Parameters framework. Like earlier studies, investigation of L2 development was preceded by investigation of L1 (for a review, see Alhawary 2002). Among the central (and related) questions investigated in the SLA literature are Universal Grammar access/L1 transfer in L2 development, and the status of L2 competence.

4.2.1 Universal Grammar access/ L1 transfer in L2 development

The issue of Universal Grammar access in L2 development has been subject to extensive debate in SLA literature and has resulted in proposals belonging to two different camps. One camp of researchers argues for limited or no Universal Grammar access at all and instead appeal to universal cognitive principles accessed by the L2 learner. The other camp argues for access to Universal Grammar but claima that such access is indirect where the L2 learner has access to Universal Grammar through L1 as manifested in the L1 transfer phenomenon. Researchers of the latter camp advanced three main proposals that differ with respect to the relative amount of transfer they allow. The Minimal Tree hypothesis (e.g. Vainikki and Young-Scholten 1998) claims that only lexical categories transfer. The Weak Transfer hypothesis (e.g. Eubank 1996) claims that both lexical and functional categories transfer but the feature values associated with functional categories do not. The Full Transfer/Full Access hypothesis (e.g. Schwartz and Sprouse 1996) proposes that the entirety of L1 grammar (including lexical categories, functional categories, and abstract feature values associated with functional categories) is available to L2 learners from the early stages of L2 acquisition.

Alhawary (1999, 2002) argues that his longitudinal data, reported in Alhawary (1999), show that contrary to a cognitive-interactionist proposal such as Processability Theory, L2 learners need not wait until they are able to develop category procedures, then phrasal agreement procedures, then interphrasal agreement. Alternatively, if abstract knowledge of inflection (associated with functional categories) needed for acquiring Arabic Subject-Verb agreement is already available in English native speakers for transfer to L2 but absent in nominals (since English exhibits Subject-Verb agreement features but not Noun-Adjective agreement), the relatively earlier emergence of Subject-Verb agreement before Noun-Adjective agreement in most of the participants' interlanguage systems (in six of the nine participants) may be explained accordingly. In other words, because English exhibits the abstract feature of Subject-Verb agreement, the participants were able to access/transfer this abstract knowledge and figure out that Arabic involved Subject-Verb agreement sooner than they did Noun-Adjective agreement. Hence, of the three hypotheses advanced with respect to L1 transfer, the Full Transfer/Full Access proposal seems to provide the most adequate explanation of the data.

Alhawary's (2005) study provides further evidence for the issue of L1 transfer based on cross-sectional production data collected from 53 Arabic L2 learners, (American) English and French L1 speakers enrolled in Arabic classes in their home institutions in the United States and France. Participants of each L1 belonged to three groups according to length of exposure to formal Arabic instruction: first, second, and third year. The study investigated the (abstract) φ-feature gender exhibited in Arabic Subject-Verb and Noun-Adjective agreement. The clearest findings are those related to Noun-Adjective agreement. In terms of L1 transfer, one would predict the L1 French participants to transfer their knowledge of abstract gender feature agreement in Noun-Adjective construction and acquire the form sooner than their L1 English counterparts, since the latter's L1 (English) exhibits no abstract gender agreement feature in nominal agreement. The findings confirm this prediction. There was a statistically significant difference in the production of Noun-Adjective agreement between the French L1 and English L1 groups, with the French participants outperforming their English counterparts.

4.2.2 Status of L2 competence

A closely related issue to that of Universal Grammar access/L1 transfer is the exact status (or nature) of L2 competence as to whether or not ultimate attainment (or full competence) can be achieved in L2 on a par with L1. Three recent proposals have been explicitly posited. The Local Impairment hypothesis assumes that functional projections are attainable in L2, but that features associated with functional heads are permanently impaired, irrespective of L1 (Beck 1998). The Failed Functional Features hypothesis claims that the interlanguage system of the L2 learner, specifically the functional feature system, is constrained by what is available in L1, but Universal Grammar is partially unavailable in L2 beyond the critical period (e.g. Hawkins and Chan 1997). The Missing Surface Inflection hypothesis claims that the feature system is temporarily impaired at the morphophonological (surface) level due to complexity in mapping between surface forms and underlying abstract features (e.g. Bruhn de Garavito and White 2002). In essence, the first two hypotheses assume that ultimate (grammatical) attainment of L2, on a par with L1 acquisition, is not possible, due to permanent impairment in the L2 system beyond the critical period. The third hypothesis assumes that ultimate attainment is possible, but it is contingent upon the L2 learner's figuring out the mapping complexity between surface forms and underlying abstract forms.

In addition to addressing L1 transfer, Alhawary's (2005) study investigated the predictions made by these three proposals based on the cross-sectional production data collected from 53 Arabic L2 learners, (American) English and French L1 speakers, as reported above. The reported findings indicate that the L1 English participants had more difficulty with gender agreement than their L1 French counterparts. However, the findings also show that the L1 English advanced group performed (with a correct agreement score of at 80%) better than the L1 English beginning and intermediate groups (at 60% and 61%, respectively), suggesting that performance of the L1 English participants improved along length of exposure to L2. The findings also reveal that in Group 3 (of both L1 backgrounds), at least one L1 English participant had a 100-percent-correct score in singular feminine Noun-Adjective agreement and as many as three had a 100-percent-correct score in singular feminine Subject-Verb agreement (as opposed to three and two, respectively, in the French Group 3). Moreover, at least one of the L1 English participants (Group 3) had 100-percent-correct score in both forms (as did an L1 French counterpart). This observation suggests that even full (grammatical) attainment in SLA in both groups is possible. Accordingly, Alhawary (2005) concludes that the findings are in support of Bruhn de Garavito and White's (2000) Temporary Impairment hypothesis, assuming temporary rather than permanent impairment (cf. Beck 1998; Hawkins and Chan 1997).

However, Alhawary (2005) departs from Bruhn de Garavito and White (2000) in attributing the temporary impairment status to the L1 transfer factor (i.e. the presence or absence of a feature in L1) rather than only to the mapping complexity factor. Bruhn de Garavito and White (2000) justify their position of the form/function complexity factor based on the obscure nature of the Spanish L2 input which the participants (L1 English and L1 French) of their study received. Bruhn de Garavito and White argue that the generalization of the endings {-o} and {-a} as masculine and feminine gender markers is misleading (in Spanish) and insufficient to account for many other obscure cases. For example, there are a few feminine nouns with {-o} ending, numerous and common masculine nouns with {-a} ending, and many (invariant) adjectives lacking overt gender agreement. This makes Bruhn de Garavito and White go so far as to state that such gender markings amount to no more than word markers rather than gender markers. Alhawary (2005) argues that this was hardly the case with the input which the participants of his study received. The participants of his study received a clearer, more transparent input where the gender agreement is to a large extent regular (in Modern Standard Arabic). In the vast majority of cases, Noun-Adjective agreement for singular feminine is achieved by attaching the suffix {-a} in word-final (salient) position to both nouns and adjectives, resulting in a rhyming effect that can serve as a phonological clue to the Arabic L2 learner. For singular masculine, agreement is realized as zero {-Ø}, the stem form being the default masculine form. Thus, given the data and the Arabic L2 input, Alhawary (2005) argues that the presence or absence of a feature in L1 (i.e. L1 transfer), claimed to be irrelevant by Bruhn de Garavito and White (2000), is a factor in L2 development.

Thus, current SLA research attempts to readdress the perennial question in SLA with respect to the difference between ultimate attainment in L1 and L2 and the factors involved. In L2 acquisition, many adult L2 learners are observed to never fully attain native-like competence regardless of length of exposure to L2, but this is not true in L1 acquisition. A biological 'critical period' effect (between 6 and 15 years of age) is identified as a determining factor for achieving native competence status (Long 1991), while others claim that formal (conscious) learning never becomes (unconscious) acquisition (Krashen 1981). However, recent research suggests that even L1 attainment is not to be taken for granted. Studies in Specific Language Impairment (→ language impairment) have revealed that in fact at least as many as 10 percent of children in the United States and Britain exhibit language delay, and not all children eventually catch up on their L1 competence. More will be learned about SLA as inquiry into language acquisition phenomena in general expands.

Bibliographical references

Al-Ani, Salman H. 1972–1973. "Features of interference in the teaching of Arabic composition". *an-Nashra* 5–6.3–13.

Al-Buainain, Haifa. 1986. *Second language acquisition of Arabic: The development of negation and interrogation among learners in the UK*. Ph.D. diss., University of Edinburgh. (Publ., Qatar: Dar al-Thakafa, 1987.)

Alhawary, Mohammad T. 1999. *Testing processability and effectiveness of computer-assisted language instruction: A longitudinal study of Arabic as a second/foreign language*. Ph.D. diss., Georgetown University.

——. 2002. "The role of L1 in L2 acquisition of inflectional morphology". *Perspectives on Arabic linguistics,* XIII–XIV, ed. Dilworth Parkinson and Elabbas Benmamoun, 219–248. Amsterdam and Philadelphia: J. Benjamins.

——. 2003. "Processability theory: Counter-evidence from Arabic second language acquisition data". *al-ʿArabiyya* 36.107–166.

——. 2005. "L2 acquisition of Arabic morpho-syntactic features: Temporary or permanent impairment?" *Perspectives on Arabic linguistics*, XVII–XVIII, ed. Mohammad T. Alhawary and Elabbas Benmamoun, 273–312. Amsterdam and Philadelphia: J. Benjamins.

Bailey, Carolyn Madden and Steven Krashen. 1974. "Is there a natural sequence in adult second language learning?" *Language Learning* 24.235–243.

Beck, Maria-Luise. 1998. "L2 acquisition and obligatory head movement: English speaking learners of German and the local impairment hypothesis". *Studies in Second Language Acquisition* 20.311–348.

Bialystok, Ellen. 1990. *Communication strategies*. Oxford: Blackwell.

Brown, H. Douglas. 1980. *Principles of language teaching and learning*. Englewood Cliffs, N.J.: Prentice-Hall.

Brown, Roger. 1973. *A first language*. Cambridge, Mass.: Harvard University Press.

Bruhn de Garavito, Joyce and Lydia White. 2002. "The second language acquisition of Spanish DPs: The status of grammatical features". *The acquisition of Spanish morphosyntax*, ed. Ana Teresa Pérez-Leroux and Juana Muñoz Liceras, 153–178. Dordrecht: Kluwer Academic Publishers.

Corder, Stephen Pit. 1967. "The significance of learner's errors". *International Review of Applied Linguistics* 5.161–170.

Dulay, Heidi and Marina Burt. 1973. "Should we teach children syntax?" *Language Learning* 23.245–258.

——. 1974. "Natural sequences in child second language acquisition". *Language Learning* 24.37–53.

Eubank, Lynn. 1996. "Negation in early German-English interlanguage: More valueless features in the L2 initial state". *Second Language Research* 12:4.73–106.

Fakhri, Ahmed. 1984. "The use of communicative strategies in narrative discourse: A case study of Moroccan Arabic as a second language". *Language Learning* 34:3.15–37.

Hawkins, Roger and Yuet-Hung Chan. 1997. "The partial availability of Universal Grammar in second language acquisition: The 'failed functional features hypothesis'". *Second Language Research* 13:3.187–226.

Krashen, Stephen. 1981. *Second language acquisition and second language learning*. New York: Pergamon Press.

Larsen-Freeman, Diane and Michael H. Long. 1991. *An introduction to second language acquisition research*. New York: Longman.

Long, Michael H. 1991. "Maturational constraints on language development". *Studies in Second Language Acquisition* 12.251–285.

Mansouri, Fethi. 2000. *Grammatical markedness and information processing in the acquisition of Arabic L2*. Munich: Lincom Europa.

Nielsen, Helle Lykke. 1997. "On acquisition order of agreement procedures in Arabic learner language". *al-ʿArabiyya* 30.49–93.

Oxford, Rebecca and Andrew Cohen. 1992. "Language learning strategies: Crucial issues of concept and classification". *Language Learning* 3.1–35.

Pienemann, Manfred. 1998. *Language processing and second language development: Processability theory*. Amsterdam and Philadelphia: J. Benjamins.

Rammuny, Raji M. 1976. "Statistical study of errors made by American students in written Arabic". *al-ʿArabiyya* 9.75–94.

Schumann, John. 1979. "The acquisition of English negation by speakers of Spanish: A review of the literature". *The acquisition and use of Spanish and English as first and second languages*, ed. Roger W. Anderson, 3–32. Washington, D.C.: TESOL.

Schwartz, Bonnie D. and Rex A. Sprouse. 1996. "L2 cognitive states and the full transfer/full access model". *Second Language Research* 12.40–72.

Selinker, Larry. 1972. "Interlanguage". *International Review of Applied Linguistics* 10.209–231.

Tarone, Elaine. 1980. "Communication strategies, foreigner talk, and repair in interlanguage". *Language Learning* 20.417–431.

Vainikki, Anne and Martha Young-Scholten. 1998. "The initial state in the L2 acquisition of phrase structure". *The generative study of second language acquisition*, ed. Suzanne Flynn, Gita Martohardjiono, and Wayne O'Neil, 17–34. Mahwah, N.J.: Erlbaum.

Wode, Henning, Jens Bahns, Hannjorg Bedey, and Werner Frank. 1978. "Developmental sequence: An alternative approach to morpheme order." *Language Learning* 28.175–185.

MOHAMMAD T. ALHAWARY (University of Oklahoma)

Second Language Teaching

1. INTRODUCTION

This entry focuses on the didactics of teaching Arabic in a Western context, particularly at the university level, since this is where Arabic is most often taught in the West. In recent years, the teaching of Arabic has also spread to secondary schools and private language institutions, but they are still few in number and do not differentiate substantially, at least at this stage, from the teaching approach at universities.

The teaching of foreign languages is always set in a historic and social context, and this is the case for the teaching of Arabic in the West. Arabic has developed from being a scholarly language studied for religious and, at times, commercial reasons in the 16th and 17th centuries to serving as one of the main foreign and second languages in the 21st century. The teaching of Arabic in Europe and the United States today and the development of the instruction in form as well as content during the last decades are therefore best understood if we take as a starting point the role, status, and social relevance of the Arabic language in Western societies.

From a historical point of view, two phenomena have had a basic importance for the teaching of Arabic in the West from the 1970s onward. In the first place, the growing economic and political importance of the Arab countries on the international stage during the 1970s and the 1980s led to pressure on universities to deliver a more communicatively based teaching of Arabic – a development most clearly seen in the United States, although the same tendency, somewhat delayed, has been seen in Europe. In the second place, the economic and cultural globalization movement began to make itself felt in education in the 1990s, which has given Arabic a far more prominent position in Western societies than was the case earlier. Immigration, the abundance and availability of new media in Arabic, and the widely used interactive communication technology put new focus on choice of language varieties to consider in teaching, as well as on the content and the kinds of teaching materials to use. One might claim that the two tendencies, i.e. the change from a philological to a more communicative approach and the importance of globalization for the form and content of the instruction, in their different ways represent a paradigmatic change in the teaching of Arabic: first, a change of language skills to be considered in the teaching, second, the use of media and the content of the teaching.

The background for the paradigmatic changes are sketched here by summarizing the approach to the teaching of Arabic up to the 1970s, where reading skills and linguistic description were the main competencies to be taught. Then, the change from philology to communication and the subsequent problems concerning the choice of linguistic varieties and adequate teaching materials are discussed. And finally, the influence of new technologies and new media in the field of teaching Arabic from the 1990s onward is considered.

2. Reading Arabic text and the grammar-translation method

Arabic has been studied in Europe at least since the Middle Ages. At that time, Arabic was a most important tool in acquiring the scientific knowledge in fields such as mathematics, astrology, and medicine that was to be found in the many Arabic manuscripts which had ended up in Europe via Islamic Spain or the Crusades in the eastern part of the Mediterranean. However, as science moved out of the monasteries and into the newly established universities, and especially after the fall of Constantinople in 1453, scholarly interest began to focus more on Greek texts. By the end of the 15th century, Arabic had lost its importance as a language of science. During the following centuries, the study of Arabic focused mainly on theological aspects (→ Arabic studies in Europe). The growing interest in Biblical studies, in the wake of the Reformation in the 16th century, meant that a number of universities started to teach Hebrew, only to be followed later by languages such as Assyrian, Chaldaean, and Arabic (Versteegh 2001, 2006; McCarus 1992).

Biblical studies and theological interest in Islam and Arabic created a natural need for knowledge within the field of comparative linguistics and put focus on the study of texts which could give information on Arab and Islamic cultures – religious texts, poetry and literature, grammatical manuscripts, etc. During the 19th and the first half of the 20th century, interest grew due to colonization and the geostrategic interests of the Western countries in the Arab world, leading to the creation of a number of Western university institutes for comparative Semitic philology that focus mainly on reading skills and linguistic description. Even though many Arab countries achieved independence in the years after the Second World War, some of which grew in international importance because of the exploitation of oil, most universities kept their focus on these areas.

The common denominator for the reading of manuscripts, Bible studies, and the philological approach of comparative Semitic linguistics was the focus on reading skills, since gaining access to the contents could only be done through reading – Classical Arabic for religious texts and Modern Standard Arabic for more contemporary texts. Reading skills require a knowledge of vocabulary and grammar, and the most common way to teach this was what in applied linguistics has been termed the grammar-translation method. The core of this approach is translation, both when teaching the language and testing the learners' reading skills. The teacher would typically give a short introduction to the Arabic alphabet and then introduce rules and phenomena of the language – declension of nouns, the triliteral verb, the construct state, etc. Then followed translation exercises, in which the learner had to translate sentences into Arabic, often with the help of a bilingual wordlist. Next, students were tested in reading and translation into English, to see if they could recognize and render linguistic structures and meanings correctly (Byram 2004:635–636). And finally, when the basic linguistic phenomena had been taught, the teaching proceeded to the reading of relevant texts with focus on contents. This approach, which was the core of Arabic studies for centuries, is still used at many Western universities, not just for teaching Arabic but also for teaching other less commonly taught languages, as well as in many autodidactic Arabic courses, the only change being that grammar books have been replaced by teaching materials with a more modern vocabulary. Recent examples of this approach can be found in Haywood and Nahmad (1962), Thackston (1994), and Abu-Chacra (2001). It has often been argued that the grammar-translation method met most of the requirements of the Biblical and philological studies and as such was a suitable approach to the teaching of reading Arabic texts. But modern research in applied linguistics makes this doubtful, since the grammar-translation method focuses primarily on understanding of linguistic forms and needs to be supplemented by other competencies if the learner is to obtain good reading skills.

The academic focus on reading skills and linguistic description meant that people who needed to speak and understand spoken Arabic – merchants engaging in business agreements, Christian missionaries wishing to spread the message of the Bible, or persons with diplomatic problems to be solved – had to seek assistance elsewhere. Oral skills cannot just take their starting point in a spoken variety of the written language because of the important differences between spoken and written Arabic, and therefore oral skills had to be learned either through personal contacts with native speakers of Arabic or non-natives with a solid oral proficiency in the West, or by going to the Arab world. There are a few examples of scholars who learned spoken Arabic in order to help solve diplomatic or business-related problems, but normally this was done by people outside the universities. The acquisition of speaking and listening skills in Arabic was not an institutionalized part of university activities; and to the extent that a spoken variety of Arabic was used in scholarly work, it was primarily a linguistic object to be described, not a skill to be mastered.

3. From philology to communication

Terry Mitchell (1969) described the dilemma facing Arabic language instructors in Great Britain. On the one hand, universities considered the goal of Arabic studies to provide their students with a style of education that would enlarge their intellectual awareness, rather than provide them with vocational training in skills and techniques. Their approach primarily led to academic careers. On the other hand, there was a growing demand in society for people with skills in Arabic who could fill jobs in the Foreign Service, the British Council, the oil companies, and similar commercial organizations where more proficiency-oriented skills were demanded (Mitchell 1969:3–4).

Mitchell pointed to a problem which had been growing since the Second World War. With the onset of the Cold War, the American and British military realized that there was a need for people with skills in Arabic, who were able to do more than just read Arabic texts. In 1947, the American Army Language School started teaching Arabic – a school which later merged with the language school of the U.S. Navy to become the Defense Language Institute in Monterey, California – and in the same year the British government decided to establish the Middle East Centre for Arab Studies (MECAS) in Shemlan, Lebanon. The aim of both initiatives was to train Arabic-speaking persons who would be able to solve many of the practical military and intelligence-related tasks arising in the wake of the Cold War. The launching in the USSR of the Sputnik satellites in 1957 further promoted initiatives to change the conditions for the teaching of Arabic. The National Defense and Education Act (NDEA), passed by the U.S. Congress in 1958, provided means to further area-related specializations and applied language teaching, and in the field of Arabic there was now money to pay for teachers' salaries, grants, pedagogical workshops, and development of teaching materials. The NDEA emphasized speaking and listening skills and wanted foreign languages to become an integrated part of the newly established area study programs. In 1967, as a natural consequence of this applied approach, a consortium of American universities established the Center for Arabic Studies Abroad (CASA) at the American University in Cairo. Here, undergraduate and graduate students who had studied Arabic for a minimum of two years could take intensive Arabic classes during the summer and the academic year, if they were able to pass a fairly competitive entrance test. Today, almost forty years later, it is still a flourishing and academically well-established program which not only trains students but also offers further education in language skills and pedagogy to American teachers of Arabic.

The modernization of the teaching of Arabic in the United States in the 1960s and the 1970s was similar to the one taking place at the British universities. In the late 1940s, the Royal Commission, under the chairmanship of the Earl of Scarbrough, had recommended the establishment of area studies combining language teaching and civilizations with social and political sciences. In the 1960s, British students could take a university degree in modern Middle Eastern studies, thus combining Arabic (or Turkish or Persian) with political science, geography, or anthropology. This did not, however, lead to competence-based teaching of speaking and listening skills to the same degree as in the United States. First and foremost, the students were to learn to read newspapers, reports, and articles, whereas the spoken language had a very limited role, for instance in connection with fieldwork and short stays in Arabic-speaking countries (Mitchell 1969:6–7). If students were to acquire speaking and listening comprehension in Arabic, they had to go to governmental schools such as Shemlan (which was closed in 1967), or to one of the many language schools for foreigners that were established in the Arab countries during the 1970s. The University of Damascus established the Arabic Teaching Institute for Foreigners in 1973; the Institut Bourguiba des Langages Vivantes at the University of Tunis started offering courses in Arabic for foreigners in the mid-1970s; and the International Language Institute in Cairo was founded in 1977, to mention only a few. Later, a number of Arabic language schools followed, both private and state-supported, all over the Arab world, some of which only existed for a few years. The CASA approach and the many bilateral agreements among American and Arabic universities that also grew up during the 1980s did not really become popular in Europe. One reason for this may have been that communicative teaching of Arabic was not common in Europe at the time; another reason may have been the geographical closeness, which meant that an individually arranged stay at a summer school at in the Arab world was a feasible challenge for European students of Arabic.

In many ways, the development in the United Kingdom paved the way for what happened in the rest of Europe during the 1970s and the 1980s. Philological university institutes added Arabic or Middle Eastern area study programs, which included the teaching of Arabic with a focus on reading skills and with less importance given to oral skills. Even though the two types of study programs exist side by side today, there is no doubt that the establishment of area studies affected the philological studies. The proof of this is to be found in the classes of speaking and listening skills offered by many philological institutes in parallel with translation, grammar, and text analyses, as well as the discourse used by many philologically trained university teachers when describing their teaching of Arabic. They often argue that even a student who learns Arabic for a scholarly purpose will benefit from being taught by modern didactic methods and that more research is needed in the application of general didactic principles. In spite of their positive attitude to communicative teaching, also for students at traditional philological institutions, however, one can sometimes observe a certain reservation to the matter, when the same teachers suggest a differentiation between courses for scholarly and for practical purposes, so that university institutes can concentrate more on the former (see, e.g., Steppat 1978:36–38).

4. COMMUNICATIVE COMPETENCE AND THE PROFICIENCY MOVEMENT

There seem to be at least two different definitions of communicative language teaching in the field of Arabic. The one most widely used by teachers and students seems to equal 'communicative' with 'oral communication'; in that sense, one only needs to add oral skills and conversation exercises to the existing teaching approach to establish a communicative teaching of Arabic. For professionals in applied linguistics, however, communicative language teaching is a more comprehensive term, which focuses on processes as well as goals in classroom learning. It includes the four skills of speaking, listening, reading, and writing, and aims at integrating a wide range of subskills in linguistic, sociolinguistic, and discourse-related areas – all of this from the learner's perspective. Learning strategies and identification of the learner's communicative needs then provide a basis for the curriculum in the Arabic classroom (Byram 2004:124–129). In what follows, the second definition is used in connection with communicative language teaching.

The change from text-based Arabic studies focusing on reading skills and translation to communicatively oriented studies, which included the four skills from a learner's perspective, had both scholarly and pedagogical implications for the universities. Two questions in particular were pressing: what were the students supposed to master – i.e. which competences were they to practice, and how were those competences to be tested – and which language varieties were they to acquire in a foreign language with a considerable difference between the spoken and the written language.

The proficiency movement, which developed during the 1980s in the United States, provided an answer to the first question. Based on experiences gathered at the government language schools such as the Foreign Service Institute of the State Department, the Defense Language Institute in Monterey, and others, the proficiency movement defined a number of goals for how to introduce a more applied way of teaching languages in academia. In 1982, the American Council for Teaching Foreign Languages (ACTFL) published its first proficiency guidelines – a description of what learners were to master within the fields of speaking, listening, reading, and writing. The levels of language were defined on a scale of novice, intermediate, advanced, and superior, each category subdivided into low, mid, and high. The first draft was generic, i.e. comprising all languages, but the levels were later defined for the different languages, including Arabic, so that the various linguistic characteristics were considered. Also a tool especially designed to measure the learner's oral proficiency according to the ACTFL's guidelines – the oral proficiency interview (OPI) – was developed, and workshops were organized to train teachers in how to conduct these interviews.

The idea behind the proficiency guidelines and the oral proficiency interview was to create a sort of benchmark to measure language learner proficiency in the four skills and thus establish a rating scale that would allow comparison and description of progression for learners and institutions. As early as 1984, the first teachers of Arabic had courses in conducting the OPI (Allen 1992:237), and in 1989 the Arabic proficiency guidelines were published.

While the proficiency guidelines attempted to introduce a systematic description of competence goals for Arabic, another much more complicated question was pressing. Arabic is characterized by a considerable difference between the written language and the numerous spoken varieties in the Arab world. Therefore, if speaking and listening comprehension are to be included in the teaching of Arabic, it is necessary to determine how this sociolinguistic situation is to be reflected in teaching. The discussion of this question is further complicated by a widespread normative language attitude in the Arab world, which, taking as its starting point the prestige of the written language, has led to a situation where non-native speakers wishing to learn Arabic in the Arab world are often introduced to linguistic structures that many native Arab speakers do not command. It was only with the introduction of Ferguson's → diglossia model in 1959 that a proper description of the sociolinguistic reality was introduced, and even today there is a lack of more detailed descriptions of the way the spoken and written varieties interact with different users of language, for instance with the educated native speaker, whom the proficiency guidelines use as an idealized model. This linguistic situation has been one of the most commonly stated arguments against the introduction of communicative teaching of Arabic at Western universities: the pedagogical difficulty of teaching a foreign language with several linguistic varieties for which there are no well-documented descriptions of actual language use that could serve as a standard.

5. MODERN STANDARD ARABIC VS. COLLOQUIAL

The problem of language choice between Modern Standard Arabic and the colloquials in second language teaching was eagerly discussed in the pedagogical literature and at various conferences during the 1980s (see, for example, Agius 1990). Often, the discussions were focused on finding a general solution to the problem, which could benefit the teaching of Arabic as a whole, but later a clear understanding developed that the difference between written and spoken Arabic can only be handled sensibly in foreign language teaching if the starting point is taken in learner-related goals, and if the persons involved are prepared to make some necessary choices. Apart from considering learner-related goals, a minimum of three factors are to be considered in the organization of a communicative Arabic curriculum: (i) the sociolinguistic situation pertaining to the written language (hence Modern Standard Arabic) with respect to reading and writing and partly to comprehension of spoken language, and a spoken variety, a colloquial, which is bound to a geographic location, for speaking and listening comprehension, plus knowledge of how native speakers code-switch among the varieties in different speech situations; (ii) pedagogical considerations on learnability, so that the teaching is organized in such a way that it does not create unnecessary difficulties for the learner; and (iii) the language learner's motivation, which is the driving force in any language learning situation; hence, the learner must see the relevance of the way the teaching is being organized.

For learners who need to speak and understand Arabic to manage everyday situations, the teaching is pedagogically uncomplicated, since they only need to learn an appropriate, spoken variety. The same goes for learners whose goal is to learn how to read and write, because the teaching only needs to focus on Modern Standard Arabic or Classical Arabic, depending on the field of interest. In both cases, the goals of the learner and the linguistic variety chosen are connected, so the teaching can be organized in a pedagogic fashion to fit learner motivation. Problems arise, however, when the four skills are to be integrated in the acquisition process, because learners will have to deal with two varieties of Arabic, which overlap in certain fields but are distinctly different in others. In principle, there are three options to deal with this situation. The first option is to organize a curriculum in which the learners learn the two varieties, one after the other, i.e. first Modern Standard Arabic and then a spoken variety, or vice versa. The second option is to learn the two varieties simultaneously. The third option, focusing on Modern Standard Arabic to be used in all four skills, is problematic because it does not reflect the actual sociolinguistic situation in Arab countries. This third approach has, however, been introduced at a number of European universities on the pretext that it is communicative because it integrates oral proficiency.

A special case is the situation in the former Eastern European countries, where the primary goal of Arabic studies used to be the training of interpreters. Here, the emphasis was on proficiency in Modern Standard Arabic, especially the political and economical jargon. This approach is epitomized by the language course of Krahl and Reuschel (1974). Students from these countries sometimes achieved a remarkable level of proficiency in this rather stilted variety of the language.

There is no solid research to indicate which model is the more appropriate from a learnability point of view. Agius (1990:4) refers to an experiment with two groups of learners at the University of Arizona and at the Defense Language Institute at Monterey, one starting with dialect, the other with Modern Standard Arabic; the former group is reported to have shown much higher motivation and a more positive response, but no additional data are given. Therefore, communicative Arabic teaching has had to try out various models. During the 1980s many Western universities put pedagogical needs above the sociolinguistic reality and taught Modern Standard Arabic in all four skills without any use of a spoken variety. The students could then on their own initiative choose to combine this with a stay at a language school in an Arab country and thus learn speaking and listening in a relevant colloquial. An approach taking its starting point in Modern Standard Arabic and letting the students decide for themselves whether or not to further improve their skills in a spoken variety has the pedagogical advantage of not confusing the learner linguistically, since the teaching can concentrate on one language variety. But this approach comes with a price: it is often very demotivating for learners to realize that the language they have spent a lot of energy learning cannot be applied directly in spoken communication, making it very difficult indeed to understand what Arab speakers say to them.

In the 1990s, however, sociolinguistic reality started gaining a more prominent role in the university teaching of Arabic. Gradually, it became more common practice to use elements from the spoken language, especially in everyday situations in which students may find themselves when traveling in an Arab country. It became a characteristic feature of this approach that the use of the two varieties was tied rather strictly to specific situations. Colloquial features and vocabulary would typically be used in teaching everyday situations – greetings, introductions, shopping, etc. – while all other areas would be in Modern Standard Arabic, even in oral communication. The change from speaking exclusively in Modern Standard Arabic to communicating in a colloquial in specific situations pointed to the question of which colloquial to choose: depending on the historic, political, and economic ties to the Arab world of the individual countries, as a rule of thumb, Egyptian and Palestinian and, to a certain degree, various Gulf Arabic colloquials came to play an important role among Anglophone students of Arabic, whereas Syrian, Lebanese, and a mixture of North African colloquials were favored in the French teaching tradition.

Universities in the 1980s dealt primarily with Modern Standard Arabic and left it up to the learners to upgrade their skills in a colloquial; they later softened this approach so as to integrate some colloquial vocabulary into their curricula in the 1990s. Many government schools, on the other hand, particularly in the military, opted for the opposite: to make learners orally operative as quickly as possible; in those settings the teaching would start with a spoken variety and later proceed to Modern Standard Arabic. Unfortunately, the experience gained here is rarely publicly available – the same is true for the teaching materials used in these schools – and it is therefore difficult to evaluate the results of this approach. There is, however, information available about the language course at the University of Amsterdam, which follows this approach (Woidich 2007), and the teaching materials used in the course have been published (Woidich and Heinen-Nasr 1995, 1998).

Summing up the complex issue of which variety to teach in the communicative classroom including the four skills, two points must be made. First, it is important for learner motivation – and for the students' job possibilities after graduation – that the teaching take into consideration the sociolinguistic realities of the Arab world; hence, both the written language and a spoken variety should be taught. Secondly, as there is no solid research to clarify the issue from a learnability point of view, it is not possible to say anything decisive on the order in which the varieties should be taught – simultaneously or one after the other. In recent years, a few articles have reported on individual teacher and classroom experiences, and as valuable as these are to inspire new teaching practices, they do not make up for results based on systematically collected empirical data on learner behavior. So, while waiting for the results of research on the learnability issue, teachers must continue to try out various models, starting from learner-related goals and experience-based teaching, and perhaps seek inspiration from the growing research on second language acquisition and bilingualism.

6. Three generations of textbooks

The development in the communicative teaching of Arabic can be illustrated through three of the most commonly used textbooks at Western universities. The first and probably the best known is *Elementary Modern Standard Arabic*

(EMSA) published by Cambridge University Press in 1968 – a set of two textbooks that for generations of students of Arabic came to be known as 'The orange book' because of the color of the book cover. EMSA and its successor *Intermediate Modern Standard Arabic* were a result of the efforts which took place in the United States in the 1960s following passage of the National Defense and Education Act, described above. EMSA carefully follows the audiolingual method: each lesson comprises an introductory text in Arabic followed by a translation of the text into English. The learner is expected to listen to the text, read it aloud, and answer questions. Then follows a number of grammatical explanations accompanied by drills, i.e. sentences with the same grammatical structure, but with a varied vocabulary, which the learner is to read, repeat, or answer according to a set standard. The accompanying tapes allow the learners to listen to exercises and answer questions and drills, thereby practicing their oral proficiency. The variety used is Modern Standard Arabic, and the texts are compiled so that most of them deal with situations in which native speakers of Arabic would use this variety: short news items, texts on women, Bedouin, etc. By introducing listening and speaking skills in the classroom, using tapes for individual learners, and including a drill-based approach, EMSA illustrates the transition from the grammar-translation method toward the audiolingual approach. But pedagogical considerations still prevail over sociolinguistic reality, in the sense that everything is done in Modern Standard Arabic.

The second textbook is *al-Kitāb al-ʾasāsī* (1988), in three volumes, published by the Arab League's Organization for Education, Culture, and Science in Tunis. Both the content and the methodology are very different from EMSA. The plot of the first volume follows Yusif who arrives in Khartoum, books himself into a hotel, has dinner at a restaurant, goes shopping, goes to the post office, visits his physician, etc. – very similar to the plot of textbooks for other foreign languages, which aim to motivate the learner by focusing on phrases which can be used in everyday situations. But instead of using the variety which native speakers of Arabic would use in similar situations, i.e. a relevant colloquial, everything is done in Modern Standard Arabic. Each lesson comprises an Arabic text with vocabulary lists of new words, but without translation, only a few simple drawings to support understanding. This is followed by a number of drills, questions, and cloze tests – all in Arabic, even when it comes to instructions. There are no grammatical instructions, no translations, only a comprehensive vocabulary list at the end of the book with translations of words into English and French. *Al-Kitāb al-ʾasāsī* is an example of the direct method, which basically tries to imitate the way children learn their mother tongue. With its avoidance of translation, use of Arabic as the only medium of instruction, and its focus on everyday vocabulary, it is a method that depends strongly on the teacher. The form of the content and the strong focus on situations where oral skills are needed show the authors' intentions to make the material more pragmatic and useful to students of Arabic, but the sole use of Modern Standard Arabic in situations where native speakers would interact in colloquial strongly limits the learners' outcome of the first volume of the book.

With the first volume of *al-Kitab fī taʿallum al-ʿarabiyya*, published by Georgetown University Press in 1995, the teaching of Arabic acquired a genuine communicative textbook. Communicative teaching requires that the learner both gain knowledge and develop skills in linguistic, sociolinguistic, discourse, and strategic competences (Byram 2000), and *al-Kitāb* offers the necessary explanations to master these skills and practice them. In each lesson, the learner is introduced to new vocabulary through a listening text, followed by questions and tasks, and the text is later introduced in writing. Then come a number of grammatical explanations followed by exercises. The last part of the lesson consists of a number of task-based exercises in the four skills: games, role plays, pair and group activities, authentic reading and listening texts, etc., which engage the learner in meaningful communicative activities. The textbook uses Modern Standard Arabic, but in a number of places Egyptian colloquial is used, and the learners' attention is often drawn to the differences between Modern Standard Arabic and the colloquial on the one hand, and the differences between colloquials on the other, so as to make the learner aware of the sociolinguistic reality. The content is also different from the other two textbooks: the two main characters, Maha and Khalid, live with their families in New York and Cairo, respectively, thus offering the learner a perspective on Arabic language and culture as it is displayed in and outside the Arab world. Such an approach points to the globalized reality in which Arabic plays a part from the beginning of the 1990s.

7. TEACHING ARABIC IN A GLOBALIZED WORLD

Arabic became an important foreign language in many European and American universities from the mid-1990s. The number of students grew steadily, and although the September 11 attack undoubtedly had an impact on university students' interests in Middle Eastern studies in general, interest in learning Arabic as a second or foreign language began well before that. Recent figures from American universities show an increase of 92.5 percent between 1998 and 2002 – from 5,505 to 10,584 students (Welles 2004) – and a similar development has taken place in Europe. This development is most probably due to the ongoing cultural and economic globalization. Many Arabs have migrated to Europe and the United States, oil is still an important resource worldwide, and political factors such as the war in Iraq and the Palestinian/Israeli conflict add to the media's coverage of the region (Nielsen a.o. 2005).

The number of students has been growing as has their diversity. Whereas students who entered university programs in the 1970s and 1980s usually did not have any knowledge of the language and the culture they were to study, the 1990s saw an important influx of heritage learners, i.e. students of Arab background or Muslims, most of whom had some prior knowledge of Arabic, be it in the colloquial of their parents or in the form of reading skills for religious purposes. There were even students who had mastered Arabic rather fluently but wished to know more about the society and the culture of their parents. The percentage of this new group of students varies, of course, according to national and local conditions; some European universities report that they had up to 75 percent of new students with a Middle Eastern background in 2005 (Nielsen 2006). The inclusion of heritage learners in the university curricula is a pedagogical challenge for the teaching of Arabic at the universities, because they often have different qualifications linguistically and culturally, and because their interests and their expectations from the education might differ from that of other students.

The rising number of students coincided with the fact that Arabic has gained ground in everyday life in many Western cities. The number of Arab immigrants make Arabic an important minority language, spoken in many parts of the public sphere in Europe and the United States. Therefore, students might well hear the language in many everyday situations, be it in a local bus or neighborhood supermarket, or at work. Add to this the easy access to new media such as Arabic satellite channels, local Arabic radio and TV stations serving the Arab communities, and the many Arabic newspapers, reports, and advertisements freely available on the Internet. Never before have European and American university students of Arabic studying in the West had such good opportunities to compare the language they learn in class with actual language use.

The growing number of students and the frequent use of Arabic in the Western public sphere must necessarily affect the teaching of Arabic. At the educational level, the growing number of students has first and foremost caused a serious lack of qualified teachers in the communicative disciplines. It is therefore highly recommended to establish teacher training programs with a specialization in TAFL (Teaching Arabic as a Foreign Language). The situation also makes new demands on both form and content of the existing teaching. The fact that the spoken varieties play an increasingly larger role for the students, even when they study at a Western university, puts pressure on the balance between Modern Standard Arabic and colloquial. The solution, chosen by many universities during the 1980s and the 1990s, of teaching Modern Standard Arabic in all four skills, while students can optionally take colloquial classes at a summer school in the Arab world, is no longer appropriate, and the limited integration of a colloquial in a curriculum based on Modern Standard Arabic no longer seems adequate. The risk of the students losing their motivation for learning Arabic is simply too great, if they repeatedly experience an inability to understand the Arabs they meet in their daily lives, among them the heritage learners participating in their classes. Studying

Arabic in a globalized world requires that the curriculum consider the sociolinguistic realities and points to the need for developing materials integrating the two varieties, both for beginners and more advanced students.

Another important challenge for the teaching of Arabic in the coming years has to do with use of communication technology. Foreign language students of today have grown up with the Internet and the new media as an integrated part of their daily lives, and they therefore expect these tools to be part of their language learning. This situation makes new demands on the form and content of the teaching materials because books with a video or a CD-ROM attached are far from sufficient to attract students. Good on-line tools need to be developed, such as reliable dictionaries and grammars, large text corpora in which students can check vocabulary and grammar against actual language use, well-structured interactive exercises challenging the linguistic and cultural competences of the students, collaborative learning activities with native speakers in the Arab world, and Internet-based courses or course components to be used independently or as a component in a blended teaching approach – just as these are developed for the teaching of other foreign languages (Ditters 2006). Such tools are slowly finding their way to the Internet. A good example is *ʾAṣwāt ʿarabiyya*, made by Emory University (http://langqtss.library.emory.edu/arabic_listening/index.html), a collection of video clips to train the students' listening skills, and *The Visual Interactive Syntax Learning* (http://visl.sdu.dk/) from the University of Southern Denmark, where students can practice their grammar by doing sentence analyses, quizzes, and computer games (Nielsen and Carlsen 2003).

New media, interactive communication technology, and new sociolinguistic realities outline the agenda for the teaching of Arabic in the years to come. This will make new demands on the teaching profession, not just to develop teacher training for new teachers but also to offer further training to qualified teachers in the use of new technology, computer-assisted language learning, and e-learning within a pedagogical framework. Collaboration with teachers and researchers in other foreign languages will be highly advantageous, so that the Arabic teaching profession will not be isolated but will benefit from the inspiration and the pedagogical and linguistic progress which is a part of the teaching of other foreign languages, such as English, German, French, and Spanish.

Bibliographical references

Abboud, Peter F. 1974. *Intermediate Modern Standard Arabic*, I, II, III. Ann Arbor: University of Michigan.

—— and Ernest N. McCarus (eds.). 1968. *Elementary Modern Standard Arabic*, I, II. Cambridge, London, New York: Cambridge University Press.

Abu-Chacra, Faruk. 2001. *Arabic around the world: Grammar and exercises*. Helsinki and Lebanon: Ammantour-Press.

Agius, Dionisius A. (ed.). 1990. *Diglossic tension: Teaching Arabic for communication*. Leeds: Folia Scholastica.

Allen, Roger. 1992. "Teaching Arabic in the United States". *The Arabic language in America*, ed. Aleya Rouchdy, 222–250. Detroit: Wayne State University Press.

American Council on the Teaching of Foreign Languages. 1989. "ACTFL Arabic Proficiency Guidelines". *Foreign Languages Annals* 22:4.373–392.

Badawi, Said M. and Fathi A. Yunis. 1988. *al-Kitāb al-ʾasāsī fī taʿlīm al-luġa al-ʿarabiyya li-ġayr an-nāṭiqīna bihā*. Tunis: Arab League's Organization for Education, Culture, and Science.

Batal, Mahmoud al-, Kristin Brustadt, and Abbas al-Tonsi. 1995. *al-Kitāb fī taʿallum al-ʿarabiyya*, I. Washington, D.C.: Georgetown University Press.

Byram, Michael (ed.). 2000. *Routledge encyclopedia of language teaching and learning*. London and New York: Routledge. (Repr. 2004.)

Ditters, Everhard. 2006. "Technologies for Arabic language teaching and learning". Wahba a.o. (2006:239–252).

Ferguson, Charles A. 1959. "Diglossia". *Word* 15.325–340.

Haywood, John Alfred and Hayim Musa Nahmad. 1962. *A new Arabic grammar of the written language*. London: Lund Humphries.

Krahl, Günther and Wolfgang Reuschel. 1974. *Lehrbuch des modernen Arabisch*. 3 vols. Leipzig: VEB Enzyklopädie.

McCarus, Ernest. 1992. "History of Arabic study in the United States". *The Arabic language in America*, ed. Aleya Rouchdy, 207–221. Detroit: Wayne State University Press.

Mitchell, Terry F. 1969. "The teaching of Arabic in Great Britain". *Linguistic Reporter* 11:2.7–11.

Nielsen, Helle Lykke. 2006. "Kulturmøder i uddannelsesverdenen [Cross-cultural encounters in education]". *Socialexpress* 3.73–81. [In Danish.]

—— and Martin Carlsen. 2003. "Interactive Arabic grammar on the Internet: Problems and solutions". *Computer Assisted Language Learning Journal* 16:1, February, 95–112.

——, Gunna F. Hansen, and Maria K. Overgaard. 2005. "Al-Maqha: A low-tech Web café for foreign language learners". *Japanese Association of Language Teachers and Computer Assisted Language Learning Journal* 1:2, August, 75–86.

Steppat, Fritz. 1978. “The diffusion of the knowledge of Arabic language and literary civilization in the Federal Republic of Germany and West Berlin”. *Bulletin of the British Society for Middle Eastern Studies* 5.35–41.

Thackston, Wheeler. 1994. *An introduction to Koranic and Classical Arabic: An elementary grammar of the language*. Bethesda, Md.: Iranbooks.

Versteegh, Kees. 2001. “Arabic in Europe: From language of science to language of minority”. *Lingua e Stile* 36:2.335–346.

——. 2006. “History of Arabic language teaching”. Wahba a.o. (2006:3–12).

Welles, Elizabeth B. 2004. “Foreign language enrollments in United States institutions of higher education”. *ADFL Bulletin* 35:2–3.7–26.

Wahba, Kassem M., Zeinab A. Taha, and Liz England (eds.). 2006. *Handbook for Arabic language teaching professionals in the 21st century*. London and Mahwan, N.J.: Lawrence Erlbaum.

Woidich, Manfred. 2007. “From Colloquial to Standard Arabic: The Amsterdam approach to teaching Arabic as a foreign language”. Paper presented at *Foundations for a Pedagogy of Arabic Translation International Conference held at the University College Francisco Ferrer, Brussels, 19–20 October, 2007*.

—— and Rabha Heinen-Nasr. 1995. Kullu tamam: *Inleiding tot de Egyptische omgangstaal*. Amsterdam: Bulaaq. [English trans., Kullu tamam: *Introduction to Egyptian Colloquial Arabic*. Cairo and New York: American University in Cairo Press, 2004.]

—— and Rabha Heinen-Nasr. 1998. *al-ʿArabiyya al-Ḥadīṯa/Inleiding tot het Modern Standaard Arabisch*. Amsterdam: Bulaaq.

Helle Lykke Nielsen
(University of Southern Denmark)

Secret Languages

1. Definition of secret languages

Dissimulation, one of the current functions of language, is the process whereby communicants resort to various stratagems to conceal from outsiders the content of what are intended as private or restricted exchanges (→ slang; → jargon). Effected through the use of secret languages, this function is performed, basically, in one of two ways:

(i) The use of foreign or minority languages, or the use of metaphorical speech (usually slang, argot, etc.); this type, which consists in referring to things by different names, may be called *notional* or semantic secret languages.

(ii) The use of complex structural transpositions of fragments in the phonic and/or grammatical patterns of the language (in fact, very often, of the pig-, goat-, or dog-Latin kind of transpositions), which may be called *canonical* or *structural secret languages*.

Thus, these quasi-universal encrypting processes in secret languages predominantly involve playing around with the sounds – usually consonants or entire syllables – which are transposed from their ordinary positions in words, e.g. Moroccan Arabic *bala* ‘smart’ > *laba*. This encrypting may also involve the insertion of external fragments used as ‘noise’, i.e. disguise or distracting elements such as the nonsensical syllable *-bidj-* in the Anglo-American pig-Latin, e.g. *door* > *do**bidj**or*; or the syllable *-av-* in the French Javanais (no relationship with the Indonesian island), e.g. *porte* ‘door’ > *p**av**orte*. The degree of complexity of these secret languages depends to a great extent on the age of the users and on the functions to which the secret languages are put.

2. Names of secret languages in the vernaculars

Classical Arabic and vernacular secret languages, including Judaeo-Arabic varieties, are occasionally referred to by a generic name which may be that of the minority language used for concealment, as in the case of Syriac, Coptic, etc., in the Middle East; Berber in North Africa; or the Moroccan Jews’ *Ḥaketiyya* < Arabic *ḥ-k-y* ‘to tell, speak’. The latter hybrid language, for example, used to be spoken in the northern Moroccan cities of Tetuan, Tangier, and Larache and even further south in Fes, etc., and as far to the east as Oran in Algeria. The use of Ḥaketiyya, it is said, was restricted to the more ‘aristocratic’ old Jewish families of Andalusian descent. It was originally a creolized language, with heavy admixtures of Judaeo-Arabic and Medieval Spanish lexicon and an underlying Spanish morphology, e.g. *estoy vaḥleando y ženneando solo* ‘I’m stuck and angry all by myself’, where the Moroccan Arabic verbs *wḥel* ‘to get stuck’ and *žennen* ‘to get mad at someone or something’ are the Spanish participial inflexions of the verbs *vaḥlear* and *žennear* (see Marius 1976:21–22 on Moroccan

Judaeo-Arabic varieties, and 1976:76ff. for detailed bibliographies).

Used more specifically for concealment purposes, there existed languages such as the Judaeo-Moroccan Arabic slang called *Lašuniyya*, < Hebrew *lašon* (see Chetrit 1994:523; Levy 2001:194ff.); Levy also mentions a Judaeo-Algerian Arabic variety called *Išuruniyya* < *ʾišurun* 'Jewish', i.e. 'Jewish speech'. With appropriate admixtures of vernacular Arabic and Hebrew morphologies, along with predominantly 'slangish' Hebrew vocabulary, these Judaeo-Arabic varieties thus served as secret languages in urban and rural marketplaces for traveling Jewish vendors and salesmen of all kinds of trades, as well as conniving forms of speech, e.g. in Levy (2001:194) *ḥta y-**lex** o n-**debber lax*** [Hebrew items bolded] lit. 'until he goes away I'll talk to you', i.e. 'wait until he leaves and I'll explain things to you'. Being, therefore, morphophonic and lexicosemantic hybrids of Judaeo-Arabic and/or Spanish (with, to some extent, the ad hoc mixing of Biblical Hebrew and vernacular Arabic and/or Spanish in the case of Ḥaketiyya), these Judaeo-Arabic varieties were of the rather notional type of secret languages, spoken to beguile noninitiated Jews and gentiles alike, and they were functionally and sociolinguistically typical of minority languages in contact.

In Muslim Morocco, the term *ġoṣ*, *ġawṣ*, or *ġus* (a verbal noun from Arabic *ġ-w-ṣ* 'to dive') is used to refer to the structural types. Tunisian *gejmi* (a corrupted form of *ʿajami* 'Christian'?) seems to cover both notional and structural codes. Secret languages are also referred to as *luġet, hedret, klam* 'language, speech, parlance', etc., followed by the name of an ethnic or religious community or a brotherhood or a sect, for example *ġoṣ Hedawa*, a Sufi brotherhood of vagabond adepts of Santon Sidi Haddi, whose shrine lies in northern Morocco. Vowing to live in self-imposed degradation and ritual beggary, Heddawa men led precarious lives and often had to reckon with some amount of hostility from the bourgeois population.

The name of a secret language may also refer to a trade (the tanners of Marrakesh) or a region, tribe, city, or district (the Misfalah district in Mecca; see Bakalla 2002), in which secret languages are used by peers or associates for the communication of economic, trade, and handcraft transactions around the bazaars, or criminal or reprehensible activities in general, such as the Egyptian *luġat issīm* (no clear signification or etymology), which seems to be reserved for those slangs, i.e. notional secret languages. Of these, there are relatively substantial, albeit sketchy or at times rather imprecise, inventories and descriptions of the situations of use in the carrying out of precarious activities, such as those of street performers, musicians for bellydancing groups (the *ʿawālim*), who perform at family ceremonies, the activities of drug dealers, etc. (see, e.g., ʿĪsā 1988). Sometimes also called *luġat issīn* or *iṣṣād* (by reference to the letters?), the term ended up referring to all varieties of secret languages, structural as well as notional, so much so that the name has eventually been applied to usages ranging from the slang of Jewish or Christian goldsmiths in Cairo and Alexandria (and as far as Damascus; see Barbot 1974), to the varieties used by crooks and pickpockets, street urchins, pimps and prostitutes, etc. More currently, codes may be defined with reference to the distracting letters of the alphabet and to the syllables, such as *ġoṣ lkaf* in Morocco, or the nonsensical disguise words involved in *ġuṣ trisa* (see below, Sec. 4).

3. Names of secret languages in Classical Arabic

In Classical Arabic, the names for the encoding processes to render communication opaque (*muʿammā* 'hidden, obscure') cover too many phenomena to be of any use for typologies or classifications. In fact, the distinction between semantic and structural secret languages proposed above is all the more necessary as, in the Arabic literary tradition, secret languages have been regularly classified among the larger category of *luġz* 'enigma', pl. *ʾalġāz*, which includes more than seven types, ranging from real form-modification secret languages (called *tarjama* < *targum* lit. 'interpreting; translation', and *ṣafra* 'code', i.e. type (ii) above) to the rather symbolic types in the form of rebuses (*kināya*), puns (*tawriya*), charades (*tamṯīliyya taṣwīriyyah*), word riddles (*muḥājāt/ʾuḥjiya*), anagrams (*jinās taṣḥīfī*), etc., very often in verse and alliterative prose, some of the corpora dating back to the pre-Islamic period. The literary tradition (for example al-Ḥarīrī's *Maqāmāt*, or the *ʿAlf layla wa-layla*, etc.) abounds in settings in which this genre is indulged, covering such unlikely domains as jurisprudence, theology, Arabic grammar and lexis, arithmetic, the weather, or downright obscene riddles. Constituting thus a genre in its own right, and pertaining to language games more than to secret languages, the types above were engaged in, and didactically handed down through the generations, mostly for the pleasure of challenging the wit and breadth of the classical culture of protagonists (see Cherkaoui-Iqbal 1987 for an inventory with rich illustrations, and an exhaustive bibliography; Cherkaoui-Iqbal actually briefly reviews, in the text, many more titles than the one hundred appended references).

It is a fact that one finds it quite often difficult – even from a straightforward sociolinguistic point of view – to draw the line between true secret languages proper and many of the genres above, whose function is mainly one of gaming, with the primary aim of entertaining consenting interlocutors. However, it has been observed that (with the exception of the 'hardcore' secret languages used in situations in which the encrypting process is obviously carried out for purposes of concealment), there are occasions when the use of either of the notional or canonical types of secret languages serves exactly the same functions as those of language games, in terms not only of their ordinary communicative functions but also in terms of those feelings of solidarity, connivance, or group or peer identity for the expression of simple 'feelings of togetherness', etc. Were it in these capacities alone, these genres – whether slangs or language games of the classic kinds mentioned above – should deserve full-fledged, albeit separate treatments. However, leaving aside type (i) secret languages and language games for the obvious constraints (and due mostly to their rather conventional, unsystematic nature), in what follows the focus will be on the patterning of structural secret languages (type ii) only.

4. Types of structural secret languages

In Classical Arabic, the enciphering operation (from Arabic *ṣifr* 'zero' > *ṣafra* 'code') was used by the Abbasid and Mamluk élite for correspondence in diplomacy, secret services, war departments, sensitive business transactions, etc., with varying degrees of complexity and sophistication. The level of education of the users, therefore, is a (diglossic) divide in the repartition of the types of secret languages. On the one hand, there are common vernacular – and thus popular – secret languages, as opposed, on the other hand, to the more elitist types. The latter resort to literate encrypting, even when they may sometimes straddle the boundaries of literate and vernacular speech (e.g. the still-surviving secret languages of the *ṭolba*, the Qurʾānic 'scholars' in North Africa; see below).

As the following rapid and quite amazing illustration shows, there certainly had been phenomena in the pre-Islamic desert culture that indicated the Arabs' particular indulgence in structural encrypting. There are reports with full *ʾisnād* about the encyclopedic Arab scholar ʿAbdallāh Ibn al-ʿAbbās, who was born in the year 3 before the Hijra and died in 68 A.H. (quoted in Cherkaoui-Iqbal 1987:54ff.), who observed that the names of the lunar month's 28 nightly mansions (*manāzil*) coincided with the positions of the moon in the night sky in relation to 28 stars. The 28 names of the lunar mansions (see list in, among others, Cherkaoui-Iqbal 1987:54) could thus correlate with the 29 letters of the Arabic alphabet, which are actually 28 if one does not count the purely orthographical *lām-ʾalif*. In this respect, the name of each lunar mansion would correspond to one letter of the alphabet; the three letters in the pronoun *ʾana* 'I' <ʾ-n-ʾ>, for instance, would be encoded with the names of the 1st and 25th of the lunar mansions as *aš-šaraṭāni – saʿd al-ʾaxbiya – aš-šaraṭāni*). The first consonants in the lexical items in any nominal nomenclature or semantic field, such as the names of birds, wild animals, proper names, etc., could be and were used to encipher stretches of spoken or written text for concealment in the manner described above. However, the heavy communication cost of these systems and the heavy load they undoubtedly imposed on memory, not to mention the transmission effort, could not have made these encoding processes as largely widespread bases for enciphering protocols as one is made to believe.

However, from the 13th century onward, better-established classic types of written secret languages were developed, basically from advances in cryptology, which were themselves an outcome of the textual studies of the *Qur'ān*, particularly in the form of lexical word counts, statistical and frequency counts of the consonants, etc. These developments were spurred by the competition between scholars from Kufa, Basra, Baghdad, or Damascus to solve what seemed to be anagrams and hidden meanings within the text of the *Qur'ān*. Coupled with tremendous progress in the use of mathematics in various domains of learning, there ensued an unprecedented development of cryptography and its corollary, code-breaking techniques, to unravel the contents of enemy correspondence for use in the caliphal administration (see Qalqašandī [d. 821/1418], *Šubḥ* IX, 229–251; Bosworth 1963:17–33; Kahn 1967:80ff.).

This extraordinary offshoot of cryptography among the Arabs would eventually dwindle to be remembered at present only in rudimentary, exotic forms among *ṭolba*, traditionally educated scholars who make a living reciting the *Qur'ān* at family ceremonies, where the secrecy does not go beyond commenting on the environment or on strategies to stimulate the audience to generosity in their money offerings. The code consists of identifying the 29 letters/consonants of Arabic with numbers (from 1 to 10 for the first ten consonants; the following ten from 20 to 100; the remainder from 200 to 700 or 800). Messages are thus spelled or written down as sequences of digits (see Berjaoui 1994:512). More complex and certainly more imaginative processes in the vernaculars of Arabic, however, are the processes that involve the shuffling around and/or the insertion of structural elements of the pig-Latin types.

5. VERNACULAR TYPES OF SECRET LANGUAGES

The structural, vernacular secret languages observed in Arabic-speaking communities reveal a recurrence of virtually similar formal principles, many of which share features reported in other languages very distant from Arabic. The infinite phonic and morphological potential for encoding transformations for secret communication would make any attempt at an exhaustive inventory quite illusionary indeed. For the sake of economy, the following offers only a synthetic typology of the most current types of transformations and of their underlying principles, with relevant illustrations from only a few dialects, going from the most simple to the relatively more complex codes.

i. Talking backward in Moroccan Arabic
This is carried out with only the consonants displaced: $C_1VC_2V > C_2VC_1V$, e.g. *daṛi* 'my house' > *ṛaḍi*.

ii. Talking backward with syllable displacement
A quite simple and quasi-universal principle, this code involves the simple inversion of the syllables, thus $CV_1\ CV_2 > CV_2\ CV_1$, as in Moroccan Arabic *gaṛṛu* 'cigarette' > *ṛugga* (note the maintenance of gemination in the syllable in this example and of the feature of emphasis in (i)).

iii. Scheme with lexeme and morphological pattern transformation
In this type, the function (or grammatical) words are not affected, e.g. Moroccan Arabic *fitnaʿal* or C*itna*C*a*C, e.g. *ṛajel* 'a man' > *ritṇajal*. All kinds of other nonsensical or imaginary phonic and grammatical schemes (*'awzān*) can be found, e.g. Mauritanian Arabic *bil-fūʿil yūʿil*, i.e. *bi-l-*C*ū*C*il yū*C*i*C, thus, *ḥammad ḏahab* 'Hammad went' > *bilḥūmid yūmid biḏḏūhib yūhib*; Yemeni Arabic with the insertion of nonsensical *al-haytiyā'i* and transformation of lexemes into the scheme *al-faytiyā'il*, e.g. *yiktib* 'he writes' > *al-haytiyā'i al-kaytiyā'ib* (see Serjeant 1948).

iv. Distracting consonant or syllable insertion
The transformation occurs usually at syllable boundary. Examples are Mauritanian Arabic /g/, e.g. *'anta* 'you' > *'agantaga*; Abbadi Sheiks' insertion of *-ark-* at each syllable boundary, e.g. *kalib* 'dog' > *karkalarkib* (see Vycichl 1959).

v. Distracting syllable insertion replicating kernel vowel of original
Mekkan Misfalawiyyah, with kernel or original vowel (V) + distraction element *rb*, e.g. *Qur'ān* > *Qur'ārbān* (see Bakalla 2002).

vi. Phonematic structure inversion plus distracting syllabic elements
Moroccan Arabic distraction *k – U – an* insertion, e.g. *l-bent jmila* 'the girl [is] beautiful' ($C_1–C_2VC_3C_4\ \ C_1C_2VC_3V$) > *l-KentUbAN KmilaUjAN*, i.e. *l*$_1$*–ke*C_3C_4*u*C_2*an* *k*C_2VC_3V*u*C_1*an* (see Youssi 1977).

vii. Lexematic scheme transformation as in (iii) above, plus distracting, nonsensical words
The nonsensical distraction words alternatively take the canonical forms of verbal and nominal items, for example the Moroccan Arabic pattern *lemfaʿel* + alternating insertion of nonsensical (verbal) *terkus* and (nominal) *trisa*, e.g. *ssukkaṛ ġali* 'sugar [is] expensive' > *lemsakeṛ terkus lemġali trisa*.

The sociolinguistic distribution of secret language use across the categories of age, sex, socioeconomic and occupational stratification, etc. reveals that, notwithstanding the receding fate of these types due to various modern social dynamics, types (i), and (ii) above are used by children and (iii) and (vii) types by females, while most others, along with the great varieties of the ever-changing opaque slang varieties, are used among adolescent and adult males.

Bibliographical references

Primary sources

Marrākušī, *Ḍayl* = ʿAbū ʿAbdallāh Muḥammad ibn Muḥammad ibn ʿAbdallāh al-Marrākušī, *aḍ-Ḍayl wa-t-takmila*. Ed. Mohamed Bencherifa and ʾIḥsān ʿAbbās. Beirut: Dār aṯ-Ṯaqāfa, 1966. (2nd ed., Rabat: Academy of the Kingdom of Morocco, 1982.) [Pp. 253–256.]

Qalqašandī, *Ṣubḥ* = ʿAbū l-ʿAbbās ʿAḥmad ibn ʿAlī al-Qalqašandī, *Ṣubḥ al-ʾaʿšā fī ṣināʿat al-ʿašāʾ*. Repr. from the Amiri ed., Cairo: Ministry of National Culture and Orientation, 1918.

Secondary sources

Bakalla, Mohammad Hassan. 2002. "*Al-misfalawiyyah*: A subdialect of the Makkan dialect of Arabic". *Aspects of the dialects of Arabic*, ed. Abderrahim Youssi a.o., 412–421. Rabat: Amapatril.

Barbot, Michel. 1974. "Notes lexicographiques sur les orfèvres et bijoutiers de Damas". *Arabica* 8.72–83.

Berjaoui, Nacer. 1994. "Preliminary notes on a secret language of *ṯəlbās* in Morocco". Caubet and Vanhove (1994:511–518).

——. 1999. *A study of the* ġoṣ *in Morocco*. Ph.D. diss., El Jadida University.

Bosworth, Clifford Edmund. 1963. "The section on codes and their decipherment in Qalqashandi's *Subh al-aʿsha*". *Journal of Semitic Studies* 8.17–33.

——. 1976. *The medieval Islamic underworld*. 2 vols. Leiden: E.J. Brill.

Bright, William (ed.). 1992. *International encyclopedia of linguistics*, IV. New York and Oxford: Oxford University Press.

Caubet, Dominique and Martine Vanhove (eds). 1994. *Actes des Premières Journées Internationales de Dialectologie Arabe de Paris*. Paris: INALCO.

Cherkaoui-Iqbal, Ahmed [ʿAḥmad aš-Šarqāwī ʾIqbāl]. 1987. *Fī l-luġz wa-mā ʾilayhi*. Casablanca: Maṭbaʿat an-Najāḥ al-Jadīda.

Chetrit, Joseph. 1994. "Formes et structures du mixage linguistique dans les langues secrètes juives du Maroc". Caubet and Vanhove (1994:519–530).

Heath, Jeffrey. 1987. *Ablaut and ambiguity: Phonology of a Moroccan Arabic dialect*. New York: State University of New York Press. [Index s.v. 'Play language'.]

ʿĪsā, ʿAlī. 1988. *al-Luġāt as-sirriyya*. Cairo: Intissar Press.

Kahn, David. 1967. *The code breakers*. New York: Macmillan. (Abridged ed. 1973, Signet.)

Levy, Simon. 2001. *Essai d'histoire et de civilisation judéo-marocaines*. Rabat: Centre Tarik Ibn Ziyad.

Marius, Sala. 1976. *Le judéo-espagnol*. The Hague: Mouton.

Serjeant, Richard B. 1948. "'Cant' in a contemporary South Arabic dialect". *Transactions of the Philological Society* (1949) 121–126.

Vycichl, Werner. 1959. "A forgotten secret language of the ʿAbbadi sheiks and the slang of the Halab in Sudan". *Kush* 7.222–228.

Youssi, Abderrahim. 1977. "Les parlers secrets au Maroc". *La Linguistique* 13:1.135–143.

——. 1998. "Fieldwork 'fallouts' and the linguistic theory: /bi-/ and multilateral relations in the phonematic structure (with evidence from secret languages)". *Langues et Littératures* 16.199–224.

Abderrahim Youssi
(Mohamed V University)

Sectarian Varieties → Communal Dialects

Semantic Bleaching

The term 'semantic bleaching' refers to loss of lexical content or categorial status of a lexical item in the course of diachronic semantic change, typically resulting from a → grammaticalization process whereby certain lexical items develop into markers of grammatical relations, thus acquiring grammatical functions. This phenomenon is variously labeled 'desemanticization' (Lehmann 1995:127), 'semantic weakening' (Guimier 1985:157), 'abstraction' (Heine a.o. 1991:41–45; Heine 1991:155–157), 'semantic generalization', 'semantic reduction' (Bybee a.o. 1994:6), 'containment hypothesis' (Willett 1988:80), and 'semantic depletion' (Weinreich in Lehmann 1995:127). The common denominator to these diverse nomenclatures is that they all regard the former semantic content of a linguistic item that undergoes such change as the source against which subsequent change in semantic substance is to be measured and approximated (see Heine 1991:149). The label 'bleaching', the most frequently encountered term in literature, is likely to have derived from 19th-century scholars of grammaticalization, such as von der Gabelentz (1891), who used the German term *verbleichen* 'to bleach' to denote change occurring as part of grammaticalization (see Hopper and Traugott 2003:20).

1. Semantic bleaching and grammaticalization

Semantic bleaching is related intimately to the process of grammaticalization, in which a concrete lexical or less abstract form, in the process of being grammaticalized for use as a marker of grammatical relation, acquires an abstract or a more abstract meaning, respectively, and loses its concrete and referential meaning. The gradual nature of this process is frequently emphasized in studies of grammaticalization, and so too, semantic bleaching is assumed to be a continuous decline of the original lexical and referential meaning, leading to its ultimate bleaching at some stage of the grammaticalization process. In some cases, semantic bleaching (or depletion; Givón 1975) was used as a major criterion leading to diachronic change (grammaticization = grammaticalization). Divergent views have been expressed in the literature as to the existence of a parallel between semantic bleaching and phonological erosion. While one strand of research denies that bleaching consistently and uniformly correlates or entails phonological reduction (Haiman 1991:154), others (e.g. Lehmann 1995:127; Bybee a.o. 1994:19) suggest that semantic bleaching is synchronous with phonological erosion. This latter view envisages the loss in semantic content as taking place in advanced stages of grammaticalization, while conceding that bleaching is not limited to grammaticalization, and claiming that it occurs outside the grammaticalization domain as well (see Traugott and Konig in Traugott and Heine 1991b:II, 5). Hopper and Traugott (2003) corroborate the notion that semantic changes occur throughout the process of grammaticalization and that not every instance of semantic change involves reduction in semantic content. Nevertheless, they too reserve bleaching for later stages in the process (Hopper and Traugott 2003:76).

2. Scope of application

The term 'bleaching' and its variants are sometimes used to designate quite diverse phenomena in the field of historical semantics. The most common of these are loss or reduction of semantic components, typically the restrictive lexical specification of a word recruited for grammaticalization (Lehmann 1995:127). Heine and Reh (cited in Heine 1993:91) include any semantic development a particular linguistic item exhibits in the course of its grammaticalization as a case of bleaching or desemanticization, irrespective of the occurrence of loss (Heine 1993:90–91), whereas Hopper and Traugott (2003), while admitting that semantic changes occur throughout the grammaticalization process, reserve the term 'bleaching' for advanced stages of this process, as does Greenberg (1991). Moreover, Hopper and Traugott recognize other 'gains' the linguistic items acquire and regard these as compensation for loss of lexical content (see the discussion of the loss-and-gain model in Sec. 7 below).

3. Unidirectionality of semantic bleaching

It appears that semantic bleaching is motivated and not arbitrary; also, it is patterned in that the fuller lexical item, typically with concrete meaning, acquires a more abstract meaning for the same linguistic form. The theory of grammaticalization presupposes that irreversibility is inherent to bleaching and that bleaching is used so frequently as a label for such diachronic phenomena because of the emphasis placed on the permanent nature of semantic change. Gradual semantic loss, ultimately leading to bleaching, seems to offer a partial or negative view of what typically occurs in the process of grammaticalization, for it fails to recognize gains in the area of pragmatics (Hopper and Traugott 2003:94–98). When pragmatic 'enrichment' is taken into account in the process of semantic change resulting from grammaticalization, the increase in pragmatic meaning counterbalances (according to Rubba 1990, cited in Heine 1991), and in some cases even exceeds, the loss in semantic content for an emerging grammaticalized linguistic item.

4. Efficacy of bleaching

Bleaching of semantic content, within the framework of the bleaching model (Heine 1991), has enriched our understanding of semantic change, particularly by excluding randomness in type and direction of semantic change. The bleaching model clarifies that change is motivated and predictable. The applicability of such a model becomes particularly evident when patterns or degrees of semantic loss in crosslinguistic studies of areally as well as genetically unrelated languages are shown to be similar.

5. Generalization

Lexical meaning restricts the scope of usage for a given linguistic item, since in most prototypical cases it identifies a limited set of referents. With bleaching of the lexical portion of semantic material, the scope of usage of the linguistic item widens as it enters contexts and conceptual domains previously not permitted (e.g., the participle form *qāˁid* 'sitting' derived from the postural verb *qaˁada* 'to sit'). With the loss of its denotation of physical posture and with the retention of the generalized and more abstract durative aspect, this form shows much wider distribution as a progressive marker than its original lexical use would have allowed. Bleaching of physical semantic content also has further consequences. In rapid speech, for instance, *qaˁada* becomes morphologically defective (it loses the morphological distinctions for gender and number) and phonologically reduced: *sāmī qaˁ-yaḍrib/maryam qaˁ-taktib* 'Sammy is hitting/Miryam is writing' (Al-Najjar 1991).

Bleaching of semantic content may have far-reaching effects that go beyond the loss of the former lexical content in the affected form itself. Heine (1993) shows that simplification or bleaching of semantic content is responsible for the emergence of the past tense from the perfect, from which it is derived. In this vein, Holes (2004:217–218) claims that the recurrent association of completion of the action in Arabic verbs inflected for the perfect, and of pastness as its nonfocal sense, has given rise to the emergence of the past tense in Modern Standard Arabic, where this secondary development has become primary in certain contexts and usages.

Loss of restrictive lexical content through bleaching within the grammaticalization process of Arabic auxiliaries affords the selected lexical items for the process greater contextual frequency than was previously available. In Classical and Modern Standard Arabic, attested cases of verb auxiliaries within groupings – such as *kāna* and its sisters, the verbs of beginning, and those verbs that are part of adverbial expressions (e.g. *ṭālamā* 'how often', *qallamā* 'seldom', etc.) – shed their restrictive lexical content when undergoing bleaching of their former lexical content, thus enjoying greater textual frequency as a result of their new status as grammaticalized items.

In present-day Arabic dialects, a class of verbs of motion (e.g. *mša* 'to walk', *rāḥ* 'to leave in the evening time', *jā* 'to arrive') and postural verbs (e.g. *nāḍ* 'to rise, get up', *gām* 'to stand up', *gaˁad* 'to sit down') among others with emptied lexical content show bleaching of their original meaning. Originally, these verbs denoted various types of physical actions, but they were bleached to varying degrees when used as aspectual markers to indicate futurity as well as continuous, habitual, and progressive types of actions (see Brustad 2000:193 for a comprehensive list of bleached lexical verbs).

6. Alternative models to bleaching

The bleaching model has been criticized as overtly pessimistic in that it presupposes that lexical forms selected for grammaticalization gradually incur loss of their semantic content without any compensating gains. This criticism has induced some scholars (e.g. Sweetser 1988; Heine 1993) to question this notion, to delve deeper into the semantic and pragmatic aspects of items undergoing grammaticalization, and to conceive of alternative models of semantic and conceptual change. Two of these models are discussed here: the loss-and-gain model and the implicature model.

7. Loss-and-gain model

Draining lexical content from grammaticalized items has been acknowledged as a recurrent and perhaps essential ingredient for grammaticalization when it takes place. Yet, what has been overlooked in the grammaticalization process is that the same item accumulates gains when used in new domains that were not possible before the process was launched. To elucidate, when the verb of motion *rāḥa* 'to leave at nighttime' was emptied of its restrictive lexical content, it became possible for this verb to be used as an auxiliary, e.g. *rāḥat tudaxxin* 'she began to smoke' (Cantarino 1975:III, 259); in certain Arabic dialects (e.g. Syrian and Kuwaiti Arabic), this verb became the marker of futurity and prediction, as in Syrian Arabic *šū raḥ-tsammi l-mawlūd?* 'what are you going to name the baby?' (Brustad 2000:244), in spite of the past tense morphology of *raḥ*. These and other similar gains were first recognized by Sweetser (1988), and her hypothesis accounting for semantic shift in grammaticalized items was later labeled the 'loss-and-gain model' in Heine (1991).

8. Implicature model

The implicature model recognizes continuous conventionalization of inferences as a mechanism for diachronic semantic change, allowing the primary sense of a given linguistic form to become secondary, and vice versa. This model was developed by Dahl (1985), Willett (1988), and Heine (1993), under the name of 'implicature hypothesis' or 'implicature model'. This model overcomes the pitfalls of the bleaching model, building upon the loss-and-gain model, and is held by Heine (1993:95) to be the most suitable for accounting for meaning changes associated with grammaticalization. When compared with the two earlier models (bleaching and loss-and-gain), the implicature model presents the semantic changes that a given item undergoes, as a successive and uninterrupted series of emerging semantic modifications, with the loss of one focal sense being offset by gain. The schematization in Figure 1 (from Heine 1993) illustrates this change.

Figure 1. Comparison of the three models (after Heine 1993)

ab >	b	Bleaching model
ab >	bc	Loss-and-gain model
ab > bc > cd		Implicature model

The implicational model clarifies that most likely the semantic properties *cd* for the same linguistic form are neither directly related nor share in the sense of the *ab*. The two senses are related through the intermediate stage *bc*, where *b* is the part of the original sense that is maintained, and *c*, the part of the new sense that is subsequently introduced; for example, *kāna* 'to be' has its semantic origin in the notion of 'existence, creation', e.g. *ʾana ʾaʿrifuhu muḏ kāna* 'I have known him since he existed/was created' (Ibn Manẓūr, *Lisān* V, 3961; → *kāna wa-ʾaxawātuhā*). As such, it occurs in the past and without predicate (*kāna at-tāmma* 'complete *kāna*'). Several other senses of *kāna* have evolved historically that may not be related directly in a strict sense: (i) continuity, in the Qurʾānic example *kāna l-lāhu ġafūran raḥīman* 'Allah is much-forgiving, merciful' (also quoted by Ibn Manẓūr); (ii) copulative *kāna*, e.g. *kāna zaydun marīḍan* 'Zayd was ill'; (iii) auxiliary *kāna*, e.g. *kānat ʾummu ṣ-ṣabiyyi wa-ʾabūhu yajidāni laḏḏatan fī ʾan…* 'the boy's parents found pleasure in…' (Cantarino 1974:I, 71); and (iv) remoteness in propositional meaning, e.g. *law kuntu fī makānik* 'if I were in your place'. As an auxiliary, *kāna* has temporal meaning, namely pastness; in example (iv), pastness in temporal meaning was extended to include remoteness of the proposition and possible exclusion from occurring, counterfactuality, and nonattainment of the condition expressed in the clause.

Medieval Arabic grammarians, e.g. Ibn Barrī (d. 582/1187), recognized bleaching of lexical content from certain linguistic items that become markers of grammatical relations. In the domain of verbs, Ibn Barrī notes "semantic loss" (*salb ad-dalāla*) in *kāna* and its sisters (Ibn Manẓūr, *Lisān* V, 3962).

9. Challenges to semantic bleaching

As seen above in Sections 3–5, bleaching, while useful in describing the nature, direction, and pattern of diachronic semantic change, particularly for linguistic forms undergoing grammaticalization, offers at best a partial view when considered to be the sole factor responsible for the changes in the semantic composition of a given linguistic item (for critical views of the bleaching model, see Sweetser 1988; Traugott 1988; Heine 1993). Other types of change, such as semantic gain, are overlooked. Also, the purported bleaching of lexical content of a linguistic form having undergone grammaticalization can hardly account for the range of meanings the form in question confers on the construction in which it figures – as an illustration, the numerous meanings *mā* contributes to the syntactic constructions in which it appears make it more of a polysemous particle, despite its existence in a bleached form when in isolation. Likewise, it becomes even less useful when the umbilical cord between the source concept and its target, the grammaticalized form, is lost in the obliterating layers of history, and synchronic analysis fails to relate the two meanings (the former lexical and the resulting grammaticalized meanings) in any empirically reliable way. It is unclear from the literature whether loss of semantic content under the bleaching model should entail loss of all lexical content in all senses and subsenses of a given word alike, or merely the focal sense. Additionally, no consensus among scholars has as yet been reached on whether what is termed bleaching is simply a case of semantic shift rather than loss.

Bibliographical references

Primary source

Ibn Manẓūr, *Lisān* = ʾAbū l-Faḍl Jamāl ad-Dīn Mukarram ibn Mukarram Ibn Manẓūr, *Lisān al-ʿArab*. 20 vols. 3rd ed. Cairo: Dār al-Maʿārif.

Secondary sources

Al-Najjar, Balkees. 1991. "Grammaticalization of lexical markers in Kuwaiti Arabic". *Folia Linguistica* 25.665–675.

Brustad, Kirsten. 2000. *The syntax of spoken Arabic: A comprehensive study of Moroccan, Egyptian, Syrian, and Kuwaiti dialects*. Washington, D.C.: Georgetown University Press.

Bybee, Joan, Revere Perkins, and William Pagliuca. 1994. *The evolution of grammar, tense, aspect, and modality in the languages of the world*. Chicago and London: University of Chicago Press.

Cantarino, Vicente. 1975. *Syntax of modern Arabic prose*. III. *The compound sentence*. Bloomington and London: Indiana University Press.

Dahl, Östen. 1985. *Tense and aspect systems*. Oxford: Blackwell.

Givón, Talmy. 1975. "Serial verbs and syntactic change: Niger-Congo". *Word order and word order change*, ed. Charles Li, 47–112. Austin and London: University of Texas Press.

Greenberg, Joseph. 1991. "The last stages of grammatical elements: Contractive and expansive desemanticization". Traugott and Heine (1991b:I, 305–314).

Guimier, Claude. 1985. "On the origin of the suffix *-ly*". *Historical semantics, historical word formation*, ed. Jacek Fisiak, 155–170. Berlin: Mouton de Gruyter.

Haiman, John. 1991. "From v/2 to subject clitics: Northern Italian". Traugott and Heine (1991b: II, 135–157).

Heine, Bernd. 1991. "From cognition to grammar: Evidence from African languages". Traugott and Heine (1991b:149–187).

——. 1993. *Auxiliaries, cognitive forces and grammaticalization*. New York and Oxford: Oxford University Press.

Heine, Bernd, Ulrike Claudi, and Friederike Hünnemeyer. 1991. *Grammaticalization: A conceptual framework*. Chicago and London: University of Chicago Press.

Holes, Clive. 2004. *Modern Arabic: Structures, functions, and varieties*. Washington, D.C.: Georgetown University Press.

Hopper, Paul and Elizabeth Closs Traugott. 2003. *Grammaticalization*. 2nd ed. Cambridge: Cambridge University Press.

Lehmann, Christian. 1995. *Thoughts on grammaticalization*. Munich and New Castle: Lincom Europa.

Sweetser, Eva. 1988. "Grammaticalization and semantic bleaching". *Berkeley Linguistic Society, Proceedings of the Fourteenth Annual Meeting*, 389–405. Berkeley: University of California.

Traugott, Elizabeth Closs. 1988. "Pragmatic strengthening and grammaticalization". *Berkeley Linguistic Society, Proceedings of the Fourteenth Annual Meeting*, 406–416. Berkeley: University of California.

—— and Bernd Heine. 1991a. Introduction to Traugott and Heine (1991b:II, 1–14).

——. (eds.). 1991b. *Approaches to grammaticalization: Focus on theoretical and methodological issues*. 2 vols. Amsterdam and Philadelphia: J. Benjamins.

Willett, Thomas. 1988. "A cross-linguistic survey of evidentiality". *Studies in Language* 12:1.51–97.

Mohssen Esseesy
(The George Washington University)

Semantic Extension

Semantic extension refers to the process through which one or more semantic senses within the same or another conceptual domain are added to the core semantic sense or focus of a linguistic item. Defined as such, semantic extension has a diachronic dimension, in which a single linguistic form accumulates additional senses over time. These senses are sometimes demonstrably close and are thereby regarded within cognitive semantics as cases of polysemy; other cases in which no apparent set of semantic property links the senses together may be considered cases of homonymy. However, the original sense of a given word may not always maintain centrality through time. It, too, may become peripheral, and what was once a secondary sense may become a new dominant central sense with other extensions at its periphery.

Multiple senses attached to a single linguistic form are possibly widespread in other languages. In English, for example, almost 40 percent of entries in *Webster's seventh dictionary* show more than one meaning (Byrd 1987, cited in Ravin and Leacock 2000:1). A cursory review of Hans Wehr's root-based *Arabic-English dictionary* shows that the great majority of Arabic roots are polysemous, and the possibility for semantic extension of the vague core cluster of senses increases when the set of (tri)consonantal roots is expanded in derived forms. It is with great difficulty that word senses are distinguished and enumerated accordingly as discrete and separable senses for many words in Arabic. The set of (tri)consonantal roots, in which Arabic words are traditionally conceived, in itself contains several senses, many of which typically overlap. To illustrate, while a basic nominal *dam* 'blood' is nearly monosemous, another, *yad* 'hand', initially indicating a body part or an object, synchronically has numerous semantic extensions that cannot be enumerated precisely (see the example in Sec. 2 below). It is, therefore, difficult, if not impossible, to accurately establish the number of semantic extensions for a great majority of linguistic forms. Despite these challenges, medieval Arabic grammarians did pay attention to instances of polysemous words and their various extensions by cataloging putative cases encountered in Classical Arabic (see Ibn aš-Šajarī, *Mā ttafaqa lafḏuhu wa-xtalafa maʿnāhu*) with ample analyses.

1. Motivation for semantic extension

Semantic extension may be viewed as a natural consequence of using finite phonological forms to express potentially expandable and extendable word senses. It is further motivated on the grounds that human cognition facilitates the extension of meaning through one or more of the creative cognitive mechanisms, such as metaphor and metonymy, and socially motivated mechanisms, such as euphemism and conventionalization of meaning. The benefits of semantic extension of a single word lie in its contextual usage, which were previously unrealizable.

2. Mechanisms for semantic extension

Use of metaphor (Arabic → *istiʿāra*; → *majāz*) as a principle in expanding meaning was well known to medieval Arabic grammarians (e.g. in al-Jurjānī's *ʾAsrār*). Metaphors allow the conceptualization of one entity in one domain to be structured, understood, and expressed in terms of another domain through mapping. Normally, the more abstract is conceived of, understood, and expressed in terms of the more concrete (Lakoff and Johnson 1980). To demonstrate, *ʿayn* 'eye' as a body part functioning as the vision device is extended metaphorically in *ʿaynu l-māʾ* lit. 'eye of the water' 'water spring', whereby an aperture with flowing water in the ground is expressed in terms of the body part that has the capacity to shed tears (conceived of as water). Perhaps a contributing factor in metaphorical mapping across domains in this case is the vital importance of the eye's function to human life (also *bi-ʿaynihi* 'in person') and the quintessential role of water sources to survival in an arid climate. One observes from the cited example that the prototypical sense of *ʿayn* is somewhat more general when compared to the more specific one assumed to be its extension. Thus, sometimes the physical configuration or shape, function, and location of parts within a larger whole can be structured conceptually in one domain and mapped to another for increased expressivity and enhancement of understanding; consequently, a word's prototypical sense can expand.

Through metonymic relations (metonymy, called in Arabic → *kināya, majāz mursal*), the part stands for the whole, as in the *ḥadīṯ* (prophetic saying) *al-yadu l-ʿulyā xayrun min al-yadi s-suflā* lit. 'the upper hand is better than the lower hand' 'the giving hand is better than the seeking hand' (Ibn Manẓūr, *Lisān* VI, 4952), in which case *yad* stands for the whole person. The frequent association between the concept of giving or seeking with delivery through hands has given rise to highlighting of the hand in a conventional way to stand for the entire person. In this instance, *yad*, as a physical object, is extended to the domain of functions closely associated with it, so that it may stand for 'deliverance', 'handing over', 'receiving', as well as for the person who has the intention and capacity to perform such actions. Taylor (1995:124) awards metonymy a more privileged status than metaphor in semantic extension.

Under certain circumstances, words considered culturally → taboo or utilized in sarcasm may trigger the use of → euphemisms, which also extends meaning of certain lexicon. For example, *yā xabar ʾabyaḍ* lit. 'oh, white news!' 'what bad news!' extends the central sense of the word *ʾabyaḍ* 'white' belonging to the color domain to mean 'bad' in the context of news through euphemism – although in typical renditions, the color white in Arabic is often associated with good tidings.

3. Conventionalization of meaning

Metonymic principles are amenable to stereotypes and other social conventions, which are also subject to diachronic change. In Classical Arabic, for instance, *ṭawīlu l-yad* 'long of hands' was used to designate generosity and ability (Jurjānī, *ʾAsrār* 252). Through social convention, the same expression in Modern Arabic designates 'stealing' and is frequently used even to refer to physical intimidation. Conventionalization of meaning through social interaction facilitates the standardization of extended meaning, which in some cases bears very little or, as seen above, even converse meaning to that of the original.

4. Word meaning and interaction with context

Word meaning is neither discrete nor fixed; rather, it continuously evolves and is subject to constant extension that, in turn, serves as the basis for polysemy. Some researchers (Schütze 2000) have assigned context the primary role in appending word semantics and, therefore, have suggested that core word meaning is nonexistent. Others (e.g. Katz and Postal 1964; Katz 1972), mainly adherents to the classical approach to semantics, have suggested that word meanings be measured against a set of necessary and sufficient conditions to qualify as belonging to a certain word or category. Between these two poles, there is a range of views that grant both context and linguistic forms equal or near-equal influence on semantic senses and their extensions.

Some typically monosemic (i.e. having a single meaning) words extend their semantic sense through context as well. Examples include *kabīr* 'big, large' in *bināʾ kabīr* lit. 'big' 'large building' and *masʾūl kabīr* lit. 'big' 'high-level official', particularly where the adjective does not modify the physical size of the official but rather his/her ranking within an organization or place of employment and the like. Another illustration is *raʾs* 'head', whose prototypical sense denotes a body part, an object. It has numerous extended senses, some of which are arguably less prototypical, as in the case of *raʾsu s-sana* lit. 'the head of the year' 'New Year's Day', which pertains to the domain of time when seen from the perspective of its central denotation.

It is unnecessary that one sense at a time be detectable in context only, since there are cases in which more than one sense is present: *ʾakmalnā l-bināʾ* 'we completed the building' refers to both the completion of the building action and its result (see Pustejovsky 1995). Context plays a crucial role in determining and restricting the appropriate referent denoted by a polysemic word; for instance, *bint* 'girl; daughter' conveys an independent entity, 'girl', and an entity within a (family) relationship, 'daughter'. While *walad wa-bint* means 'boy and girl', *ibn wa-bint* refers to 'son and daughter'. Moreover, *bintī* communicates typically, at least in Modern Standard Arabic, 'my daughter' and 'my girl' (see Goddard 2000:133 for a similar discussion of French *fille*).

5. Semantic generalization

It is not always required for a word to retain its former semantic sense before undergoing semantic extension. Under certain circumstances, such as in the case of → grammaticalization, words can undergo the loss of some of their original and somewhat restrictive lexical semantic content and, over time, broaden their semantic sense, thus becoming suitable for use in a wider range of contexts than were available previously. In Classical Arabic, for instance, *rāḥa* 'he went, left' initially denoted only going or leaving during the nighttime. The same verb was later used to designate leaving at any time (see Ibn Manẓūr, *Lisān* III, 1769); in the modern dialects, *rāḥa*, after losing the sense of physical motion, is used in a modified phonological form as an auxiliary – an extension in semantic and contextual functions that signifies futurity, e.g. *šu raḥ tsammi l-mawlūd* 'what are you going to name the baby?' (Brustad 2000:244). Change of grammatical function and categorization may therefore result in extending or generalizing the original meaning to enable the linguistic form to perform a newly acquired function.

Semantic extension also may occur as a result of conceptual change underlying the linguistic form. For example, *ṣāḥib* 'companion, comrade' referred originally to a human companion, as in *ṣāḥibī* 'my companion [i.e. my friend']. When this word conceptually shifted to enter the domain of possession and ownership, it came to include the relation with nonhuman entities, e.g. *ṣāḥib bayt* 'landlord', *ṣāḥib maraḍ* 'chronically ill', *ṣāḥib fikra* 'inventor of an idea'.

6. Semantic specialization

As the converse of semantic broadening, word senses may become more restricted in their referential denotation. *Laḥm* was used in Classical Arabic to designate any type of meat, including flesh (edible or not), and even the core of fruit. In present-day Arabic, the same word, while still used to designate flesh and still within the domain of edible meats, conveys (red) meats almost exclusively, while other types of meats are referenced often by the name of their animal source (e.g. *dajāj* 'chicken').

7. Enantiosemantics

Semantic extension may lead to encroachment in the domain of converse meaning. A particular case of polysemy in Arabic – and perhaps in other languages as well – is the use of one sense or its extension to mean one thing and its converse at the same time. The term → *ḍidd* (pl. *ʾaḍdād*), i.e. enantiosemic words, is used to refer to such words (see Ibn al-ʾAnbārī, *ʾAḍdād*; al-Ḥalabī, *Kitāb al-ʾaḍdād*; for a modern treatment of this category of words, see Justice 1987:195–215). Examples of such semantic polarities include *fawqa* 'above (over); below (under)', *rāʿa* 'to frighten; to please', *muḥtall* 'occupied; occupier'.

Such semantic peculiarities have been defended by medieval Arabic grammarians against charges of ambivalence, lack of wisdom, and the dubious nature of the Arabic language – all of these claims leveled by non-Arabs in the Islamic Empire. The grammarians dismissed the charge of apparent contradiction and regarded them as normal cases of polysemy. Arabic grammarians thus resorted to the exclusive role of context in selecting single unequivocal interpretation (disambiguation), thereby eliminating inherent vagueness in Arabic lexis. Other explanations included the assertion that each of the enantiosemic words was first monosemic, and due to subsequent semantic expansion, acquired the other (i.e. contradictory) meaning. Another explanation centered on mergers between Arabic dialects (*tadāxul al-luġāt*), where in the premerger stage one dialect used a word in one meaning, whereas another dialect used the same word in the opposite meaning. In the postmerger stage, one of the two senses won wider recognition and became the focal sense.

In modern times, this phenomenon involving the 'transfer' between two contradictory word senses has been explained by appealing to general human cognitive traits, by euphemisms, by cataloging differences in interdialectal lexical semantics, and by pragmatic factors turning the originally neutral sense to one of the two senses (see Justice 1987, Chap. 7 for further analysis). Most of these enantiosemic words have undergone diachronic change through narrowing of their semantic senses and hence appear in Modern Standard Arabic in one of their two meanings only, while the converse meaning has become obsolete.

8. Extension as a mechanism for lexical expansion

Loan translations (*calques*) can also be responsible for the semantic extension of native words; French *cadre*, for instance, was substituted in the Moroccan dialect by the nonidentical native Arabic *ʾiṭār* 'frame, framework'. In this usage, the Moroccan dialect extends the original meaning of *ʾiṭār* to include among its various meanings 'the higher staff of an agency, firm', which was previously unavailable (Ali 1987:118). The word *nawā* 'pit' originally designated pits found in dates and the like; semantic extension occurred in this native word under foreign influence, in particular to accommodate modern scientific concepts. The word *nawā* is extended semantically to ‘nuclear’, as in *ḥarb nawawiyya* ‘nuclear war’. The word *mirwaḥa* ‘fan’ was extended semantically to accommodate ‘propeller’ in scientific contexts; and *jammada* lit. ‘to make solid’ came to be used for ‘to freeze’ (see Holes 2004, Chap. 8; Versteegh 1997:177–183). Along with foreign influence and the increasing need for scientific terms in Modern Arabic, certain religious terms in Classical Arabic (e.g. *yawmu l-qiyāma* ‘the Day of Resurrection’, a loan translation of the Greek *anástasis* via Syriac) show semantic expansion beyond the intended original meaning of native Arabic words (see Versteegh 1997:60–61).

In addition to the direct influx of foreign terms into Arabic, numerous previously nontechnical native lexical items in Arabic in various fields (e.g. theology, medicine, and science) have acquired technical meanings through semantic extension: *ʾaslama* ‘to surrender oneself’ > ‘to submit to God; to convert to Islam’ (Versteegh 1997:61); *šarba* ‘a sip’ > ‘laxative’ (Abboud and McCarus 1992:28); and *mirwaḥa* (see above).

Certain patterns in word derivation in Modern Standard Arabic have been used frequently and regularly in coining new terminology. The pattern *fuʿāl*, used in generating lexical items that did not belong to a special field or to specialized terminology (e.g. *furāq* ‘separation’, *ḥuṭām* ‘wreckage’, *suʾāl* ‘question’), is often used, both in Classical and Modern Arabic to coin names of diseases and medical conditions (e.g. *zukām* ‘common cold’, *juḏām* ‘leprosy’, *suʿāl* ‘coughing, *buhāq* ‘vitiligo alba’, *ṣudāʿ* ‘headache’). When put in this pattern, semantically diverse word roots generate certain medical terms systematically.

9. DIACHRONIC SEMANTIC ACCRETION

Whether words are originally monosemic or not in the early stages of their introduction to language, they typically accrue (or sometimes lose) additional related senses over time. Cognitive linguists argue, therefore, in favor of the existence of semantic structures in the form of family resemblance. There is a network of semantic relations among different senses within which a prototype emerges that captures in a holistic way the prototypical meaning of the word and its semantic relations with its submeanings. In stark opposition to classic accounts of semantics, which stipulate as a rigid condition the sharing of a set of all necessary and sufficient criteria among all members of a semantic category, cognitive semantics relaxes such requirement and instead suggests that relations among different meanings or a word or category are chain-like, whereby each two senses share one or more criteria, but a single set of definitional properties is not required to be entirely present in all word senses. Cognitive semanticists argue instead for graded membership and ‘fuzzy’ boundaries among word senses, including the prototypical or central sense and its extensions.

10. POLYSEMY AND HOMONYMY

Linguists have traditionally struggled in their efforts to distinguish polysemy from homonymy through straightforward analysis. Although on intuitive grounds polysemy relates two or more distinct but related senses to a linguistic form (e.g. a word) and homonymy two or more unrelated meanings to a single linguistic form, the demarcation between the two sometimes cannot be drawn with ease. In cases where the historical process of wear and tear disconnects the word from its complete history and empirical data is not readily available, synchronic analysis of meaning for a given linguistic item becomes all the more difficult. Consequently, whether two or more linguistic items are polysemic or homonymous can be decided on an ad hoc basis. There is a fundamental difference between the two, however. Whereas polysemy usually involves chaining of related meanings in terms of an intermediate diachronic stage in which two meanings (perhaps old and new), denoted by a single linguistic form, overlap (as in the case of grammaticalization), homonymy does not involve sharing of meaning among phonologically identical forms. Nonetheless, if relatedness in meaning in the course of history is obscured, one cannot determine with certainty whether homonymy or polysemy is the cause.

A clear case of homonymy may occur as a result of sound change. In the Egyptian Cairene dialect, *ʾalam* ‘pain’ has become homonymous with *ʾalam* ‘pen, pencil’ as a result of the phonetic deviation of that dialect from Modern Standard Arabic, in which the word in question is pronounced *qalam*. Yet, in some cases homonymous words – i.e. synchronically showing unrelated meanings among the senses – may have been polysemic words whose different meanings have separated over time. One cannot with certainty decide whether *mašrūʿ* 'legal' and at the same time 'project', or *naqd* 'critique/criticism' and 'monetary', are instances of polysemy or homonymy, because each pair may conceptually co-occur in a single context (e.g. *mašrūʿ binā' al-jisr ġayr mašrūʿ* 'the building project of the bridge is illegal', or the ambiguous *an-naqdu bi-lā naqd* 'money without criticism/criticism without [payment of] money'. At this juncture in the research, it is challenging to draw a dividing line on the continuum between polysemy and homonymy. In many cases, these efforts are at best arbitrary.

11. Issues in the field of semantic extension

Despite progress in identifying mechanisms that facilitate semantic extension for a given lexical item, there remain some vexing problems in the field, among them the problem that polysemy results from packing a linguistic form with additional senses. Questions thus remain: (i) When do senses cease to be added to a given linguistic form, instead starting a new form or, put differently, what factors (linguistics, cognitive, etc.) determine whether a new sense should be paired with a new linguistic form instead of being appended to an already existing word (Fellbaum 2000); (ii) how much sense extension can a concept accommodate; and (iii) at what point are senses too numerous to be subsumed under a single linguistic form Thus far, these and other questions pertaining to the nature of the role of immediate, topical, and broader contexts in amplifying and identifying word senses and their extensions remain subject to inquiry.

The notion of extension is found within the field of linguistic semantics, where various theories still endeavor to offer insight into the complex domain of word senses. A common inquiry into the nature of word meaning poses the question: Do lexical items have inherent semantic content independently of context, or does context provide meaning for otherwise semantically underspecified lexicon? Scholars continuously endeavor to put forth answers to these vexing questions, but the issue has yet to be settled.

Bibliographical references

Primary sources

Ibn Manẓūr, *Lisān* = 'Abū l-Faḍl Jamāl ad-Dīn Mukarram ibn Mukarram Ibn Manẓūr, *Lisān al-ʿArab*. 20 vols. Cairo: Dār al-Maʿārif, n.d.

Ibn aš-Šajarī, *Mā ttafaqa* = 'Abū s-Saʿādāt Hibatallāh ibn ʿAlī Ibn aš-Šajarī al-ʿAlawi al-Ḥasanī, *Mā ttafaqa lafḏuhu wa-xtalafa maʿnāhu*. Ed. Attia Rizk. Stuttgart: F. Steiner, 1992.

Jurjānī, *'Asrār* = 'Abū Bakr ʿAbd al-Qāhir ibn ʿAbd ar-Raḥmān al-Jurjānī, *'Asrār al-balāġa fī ʿilm al-bayān*. Beirut: Dār al-Kutub al-ʿIlmiyya, 2001.

Secondary sources

Abboud, Peter and Ernest McCarus. 1992. *Elementary Modern Standard Arabic*, II. Cambridge: Cambridge University Press.

Ali, Abdul Sahib Mehdi. 1987. *A linguistic study of the development of scientific vocabulary in Standard Arabic*. New York and London: Kegan Paul International.

Brustad, Kristen. 2000. *The syntax of spoken Arabic: A comprehensive study of Moroccan, Egyptian, Syrian, and Kuwaiti dialects*. Washington, D.C.: Georgetown University Press.

Fellbaum, Christiane. 2000. "Autotroponymy". Ravin and Leacock (2000:52–67).

Goddard, Cliff. 2000. "Polysemy: A problem of definition". Ravin and Leacock (2000:129–151).

Holes, Clive. 2004. *Modern Arabic: Structures, functions, and varieties*. Washington, D.C.: Georgetown University Press.

Justice, David. 1987. *The semantics of form in Arabic in the mirror of European languages*. Amsterdam and Philadelphia: J. Benjamins.

Katz, Jerrold. 1972. *Semantic theory*. New York: Harper and Row.

—— and Postal, Paul M. 1964. *An integrated theory of linguistic descriptions*. Cambridge, Mass.: MIT Press.

Lakoff, George and Mark Johnson. 1980. *Metaphors we live by*. Chicago and London: University of Chicago Press.

Pustejovsky, James. 1995. *The generative lexicon*. Cambridge, Mass., and London: MIT Press.

Ravin, Yael and Claudia Leacock. 2000. *Polysemy: Theoretical and computational approaches*. Oxford: Oxford University Press.

Schütze, Hinrich. 2000. "Disambiguation and connectionism". Ravin and Leacock (2000:205–219).

Versteegh, Kees. 1997. *The Arabic language*. New York: Columbia University Press.

Mohssen Esseesy
(The George Washington University)

Semitic Languages

1. Arabic as an archaic Semitic language

Historically, the core region of the Semitic peoples during the 3rd and 2nd millennia B.C.E. lay in the Fertile Crescent (Palestine – Syria – Mesopotamia). Therefore, their assumed shared original homeland cannot have been situated very far from there. Applying a genetically based distribution model of the individual Semitic peoples, it may be assumed that they emerged from the Syrian desert/steppe and infiltrated the fertile agrarian lands to the east, west, and north of this hypothesized homeland. This process begins at around 3000 B.C.E. with the migration of the – later so to be named – Akkadians into Mesopotamia, and continues with the spread into cultivated lands by Amorites, Aramaeans, Hebrews, and Old South Arabians. Those tribes that remained in the Syrian steppe, and whose language had most likely already split into several dialects, are called Arabs (→ *ʿArab*). They were subjected much less than the others to the influences of the civilizations around them and thus were able to preserve archaic linguistic features for much longer. They were the last Semitic people to develop a writing system for their language. Although 'Arabs' and their rulers (e.g. Gindibu < Arabic *jundub* 'locust') are mentioned as early as the 9th century B.C.E., occasional texts written in Arabic are not found before the beginning of the Common Era. Only with the expansion of Islam did Arabic become an important fully fledged literary Semitic language.

Arabic has preserved many Proto-Semitic linguistic features, yet it would be wrong to assume that this is invariably the case; rather, any given feature has to be examined to decide whether retention or innovation applies.

2. The genetic classification of Arabic

The fragmentation and formation of the individual Semitic languages and peoples took place within the framework of the historical processes briefly sketched above. An exact match between the historical events, as we comprehend them, and the linguistic developments has so far not been achieved.

Arabic belongs to that large Semitic group of languages that is left after the separation from East Semitic (Akkadian) and → South Semitic (i.e. Modern South Arabian and Ethio-Semitic). This linguistic group can be termed Central Semitic (in the wider sense) or, better, Middle Semitic. After the exodus of the (later to be called) Old South Arabians (= separation from Southwest Semitic), what remains is → Northwest Semitic, which forks into an older and a younger branch. The younger branch has been called 'Jungsemitisch' (i.e. Neo-Semitic) by Rössler and 'Central Semitic' by Hetzron. Among the Central Semitic languages, Arabic belongs more closely to Canaanite than to Aramaic.

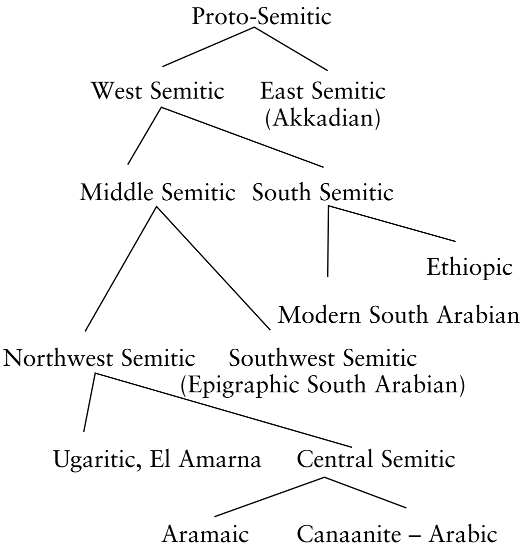
It may be that the final word on the classification of the Semitic languages has not yet been spoken. For almost any grouping, contradictory evidence can be quoted. For Arabic, for instance, the following points of agreement with Epigraphic South Arabian and Ethiopic used to be highlighted in support of its inclusion into the South Semitic group:

i. sound change *p* > *f*
ii. morphological correspondence of internal plurals that do not exist otherwise in such prominent form
iii. morphological correspondence of derived verbal Form III (Arabic *fāʿala* = Ethiopic *gābärä*) and Form VI (Arabic *tafāʿala* = Ethiopic *tägābärä*), which only exist in a very restricted sense outside this group; in Hebrew, for instance, in verbs II*w*, the *pôlel* represents a substitute for the missing **pôʿel* (= Arabic Form III *fāʿala*); *råm* (*r-u̯-m*) 'to be high, to rise' forms a stem *rômẹm* 'he rose, made high' (as if formed from *r-m-m*).

Even though Hetzron has put forward valid arguments against the relevance of these criteria, it is difficult to view such strong morphological agreements as nothing more than areal features. The assumption of a kind of *Sprachbund*, in which in the course of time morphological approximation has taken place, may not be unreasonable. However, the historical conditions for such a symbiosis of early Arabic with early Ethiopic to take shape must be shown clearly to have existed, which may prove problematic.

There is no generally accepted tree of Semitic subgroupings and their designations (see Fig. 1).

Figure 1. Tree of Semitic subgroupings and their designations.



3. PHONOLOGY

3.1 *Consonants*

For the reconstruction of Proto-Semitic phonology, Arabic plays a key role as the Semitic language that has preserved the largest number of reflexes of the Proto-Semitic consonants, i.e. 28, second only to Ancient South Arabian, which distinguishes 29 consonant symbols in its script. Ancient South Arabian is only known from inscriptions, and although these are numerous, the language has not yet been fully researched lexicographically; therefore, it cannot compete with the lexicographical progress made in Arabic. But compared to its great etymological importance for positing the Proto-Semitic phonemic units, the value of Arabic for fleshing out the phonetic character of these units is rather small.

Using the current notation and its feature implications, we have the array of Arabic consonants shown in Table 1 (except for the liquids *l*, *r*, the nasals *m*, *n*, and the semivowels *w*, *y*); note also the fricative realization (as with *f*) and the affrication (as with *j*):

Table 1. Consonants in Arabic

	[– voice] [– ‘emph’]	[+ ‘emph’] [– voice]	[+ ‘emph’] [+ voice]	[+ voice] [– ‘emph’]
labial	*f*	–	–	*b*
dental	*t*	*ṭ*	*ḍ*	*d*
interdental	*ṯ* [θ]	–	–	*ḏ* [ð]
alveolar	*s*	*ṣ*	*ḏ̣* [ð̣]	*z*
postalveolar	*š*	–	–	*j* [dʒ]
velar	*k*, *x* [x]	–	–	*ġ* [ɣ]
uvular	*q*	–	–	–
pharyngeal	*ḥ* [ħ]	–	–	ʿ [ʕ]
glottal	*h*	ʾ [ʔ]	–	–

This does not result in a structured system, contrary to other Semitic languages like Classical Ethiopic.

Emphasis has been chosen here as the distinctive phonemic feature, which is phonetically realized in each case by velarization or pharyngealization, respectively.

Restricting the attempt to the dental-sibilant series, which shows four columns, the core is formed by these rows of four: *t* - *ṭ* - *ḍ* - *d* and *s* - *ṣ* - *ẓ(ḏ̣)* - *z*. The emphatic dental *ḍ* (ض; → *ḍād*), which is voiced today, must previously have been lateral, as is shown by loans taken from Arabic (e.g. Spanish *alcalde* < *al-qāḍī*). The lateral feature is also confirmed by the Classical Arabic grammarians (8th c.). For Proto-Semitic, Arabic *ḍ* is to be postulated as voiceless **ᵗśˁ* [tɬˀ] (cf. the Assyrian rendering *Ruldāʾu* and the Greek rendering λτ (for τλ) in the name of the Arabian god Οροταλτ (from *(ο)ροτλα-) = Ruḍāʾ/y (Knauf 1989:85). This results in an original dental triad: *t* - *ṭ* - *d*.

The second row is based on the inexact representation of ظ by *ẓ*, which, according to the more educated pronunciation of Modern Standard Arabic, is better rendered as *ḏ̣*. The usual transcription of this sound represents a dialectally tinged pronunciation. If we transcribe with *ḏ̣*, the result is an interdental triad: *ṯ* - *ḏ* - *ḏ̣* [θ - ð - ð̣]. Correspondences with other Semitic languages clearly show that the emphatic interdental (*ḏ̣*) of Arabic must have been voiceless previously (**tθ̣* [tθˀ]; see below). The following original triads can still be gleaned from Arabic:

dental	*t* - *ṭ* - *d*
interdental	*ṯ* - *ḏ̣* - *ḏ*
alveolar	*s* - *ṣ* - *z*
lateral	*š* - *ḍ* - -

The view that *š* was a lateral is based on older renderings like Greek βάλσαμον < Semitic (Arabic) *bašām*, Hebrew *bośẹm* (see Steiner 1977:123).

In this way we have arrived at the Proto-Semitic triads, but the phonetic realization does not yet reflect the Proto-Semitic situation because Arabic deviates in important points from the Proto-Semitic sounds:

i. The emphatic column was originally voiceless. This is based on the observation that the other Semitic languages only show voiceless emphatic sounds.
ii. All sounds that are fricatives today must originally have been affricates. This is especially true of the alveolar row of sibilants: *s* - *ṣ* - *z*; even today, ^{t}*ṣådẹ*^{y} is realized as an affricate in Hebrew (see Steiner 1982). But even the interdental and lateral sounds must originally have shown affrication, cf. Proto-Semitic **ʾar*^{t}*ś’-* > Arabic *ʾarḍ-*, Hebrew *ʾęręṣ-* ‘earth’, Proto-Semitic ***^{t}*θ’ill-* ‘shadow’ > Arabic *ḏ̣ill-*, Hebrew *ṣẹl*. The sound changes active in these cases presuppose an affricative realization.
iii. Emphasis must originally have been realized as glottalization, which differs from Arabic. For a long time, modern Arabic velarization was regarded as original, but by now glottalization – as encountered in the Ethio-Semitic languages and partially still in Modern South Arabian – is probably to be regarded as a more likely candidate.
iv. Arabic *š* poses a special problem. This sound corresponds with *ś* (*s*^{2}) in other Semitic languages. It is in opposition with the Arabic sound *s*, the result of a merger of Proto-Semitic **s*^{1} (the unmarked sibilant, which is often transcribed *š*) and deaffricated **s*^{3} (the affricated sibilant).
v. Arabic *ǧ* [dʒ] goes back to **g*, and *q* was articulated further forward in the larynx ([k’]).

According to the traditional view, the merger of **s*^{1} [s] (‘*š*’) and **s*^{3} [ts] represents the only merger of Proto-Semitic phonemes in Arabic. In this opinion, the phonemes of Ancient South Arabian as represented by their own written graphic symbols are identical with the phonemes of Proto-Semitic. However, there are indications of further Proto-Semitic phonemic units which seem to fill some gaps in the system: a voiced representative of the lateral row (*^{d}*lʒ*) has been postulated (see Voigt 1992), and an emphatic representative of the velar row (**x̣*; see Huehnergard 2003).

In spite of the considerable differences between the Proto-Semitic and Arabic sounds, the sound change that led to Arabic *s* (< **s*^{1} and **s*^{3}) may throw some light on the original pronunciation of sibilants. The old idea that original ‘*š*’ (i.e. *s*^{1}) changed to *s* in Arabic is still upheld by many, but it is phonetically less probable. It is much more in congruence with sound changes in Semitic to assume it developed from **s*^{3} [ts], which was originally affricated, and a general type sibilant **s*^{1} [s], which is outside the (occlusive and affricated) consonant block (cf. *s* in Greek, which is outside the truly occlusive consonants). The genesis of Arabic *s* lies in the deaffrication of *s*^{3} [ts] > [s], which led to the merger with *s*^{1} [s].

In Proto-Semitic, the units **s*^{1}, **h*, and *ʾ stand outside the (occlusive/affricative) consonantal system, and it is possibly not a coincidence that these three consonants in particular are attested as causatives in the Semitic languages. They are connected through the sound developments *s*^{1} > *h* > ʾ or Ø (cf. the same sound changes that led from Latin *sex* to Greek ἕξ [heks], later on [eks] ‘six’).

These considerations concerning the reconstruction of Proto-Semitic lead to the original system of Table 2, of which only a few phonemes have preserved their original phonetic realization in Arabic. In the four rows of the dental-sibilant series, it is only *t, d* (leaving apart the question of aspiration), and *s* to some extent.

The following sound changes have taken place from Proto-Semitic to Classical Arabic:

i. In the P part: fricativization of the stop: **p* > *f*
ii. In the T part:
    a. deaffrication in the interdental, alveolar, and lateral row: e.g. *^{t}*θ* > *θ* (*ṯ*), *^{d}*z* > *z*, *^{t}*s* (*s*^{3}) > *s*
    b. realization of emphasis (*glottalization) as velarization/pharyngealization: e.g. **t’* > *ṭ* [tˠ/tˤ]

Table 2. Consonant system of Proto-Semitic

			[– voice] [– glott]	[+ glott] [– voice]	[+ voice] [– glott]
P	labial		*p ف	–	*_b_ ب
T	dental		*_t_ ت	*_t’_ ط	*_d_ د
	interdental		*^{t}_θ_ ث	*^{t}_θ’_ ظ	*^{d}ð ذ
	alveolar	*_s_^{1} (_š_) س	*^{t}_s_ (_s_^{3}) س	*^{t}_s’_ ص	*^{d}_z_ ز
	lateral		*^{t}_ś_ (_s_^{2}) [tɬ] ش	*^{t}_ś_ [tɬʔ] ض	(*^{d}_lʒ_)
K	velar		*_k_ ك	*_k’_ ق	*_g_ ج
	uvular		*_x_ خ	(*_x̣_)	*_ġ_ غ
	pharyngeal		*_ḥ_ ح	–	*ʿ ع
	glottal		*_h_ ه	* ʾ ء	–

c. voicing of the interdental and the lateral emphatics: *^{t}_θ’_ >…> _ḏ̣_, *^{t}_ś’_ [tɬʔ] >…> _ḍ_
d. delateralization of the lateral emphatics: *^{t}_ś’_ [tɬʔ] >…> _ḍ_

iii. In the K part: uvularization of the velar *_k’_ (> _q_), palatalization of *_g_ > _j_

Rule ordering has to be observed in the T part. Deglottalization/velarization is likely to come first, followed by voicing, with delateralization at the end of this historical process. The following derivation may therefore be posited for the velarized _ḍ_ of present-day Arabic: (the voiceless affricate glottalized lateral) *^{t}_ś’_ [tɬˀ] > (deglottalization/velarization) *^{t}_ṣ́_ [tɬˠ] > (voicing) *^{d}_ẓ́_ [dlʒˠ] > (delateralization/desibilation) *_ḍ_ [dˠ]. For the voiced glottalized interdental _ḏ̣_ we assume a parallel derivation: (the voiceless affricate glottalized interdental) *^{t}_θ’_ > (deglottalization/velarization) *^{t}_θ̣_ [tθˠ] > (voicing) *^{d}ð̵ [dðˠ] > (deaffrication) ð̵ [ðˠ] (‘_ẓ_’).

Arabic is needed for positing the Proto-Semitic phonemes because – as already mentioned – all other Semitic languages except Epigraphic South Arabian have experienced more phoneme mergers. Consider Hebrew, where _ṣ_ is the result of a merger of Proto-Semitic *^{t}_s’_, *^{t}_ś’_, and *^{t}_θ’_. Only a comparison with Arabic roots, where these sounds have one-to-one reflexes – though phonetically strongly modified – _ṣ_, _ḍ_, and _ḏ̣_ (‘_ẓ_’), allows the safe establishment of the Proto-Semitic character of a Hebrew root with _ṣ_.

3.2 _Vowels_

The Proto-Semitic vocalic system is best reflected in Arabic: three short vowels and three long vowels (Table 3).

Table 3. Proto-Semitic vowels

_i_		_u_	_ī_		_ū_
	_a_			_ā_	

Concerning long vowels, morphological length (e.g. the adjective _faʿīl_ or the verbal Form III _fāʿala_) must be distinguished from contraction length (e.g. with _biiḍ_ < *_buiḍ_ ‘white [pl.]’). The two diphthongs (_ay_ and _aw_) have survived unchanged.

The two semivowels (_w_ and _y_) are only distinguished from their homologous vowels in relation to their positioning in the syllable. Syllable-initially and syllable-finally only _y_ and _w_ occur, but in the core of the syllable only _i_ and _u_. Other vowel sequences are assimilated like *_ui_ > _ii_; see the more complex development *_sauiid_ (*_sawīd_) > *_saiiid_ (= *_sayīd_) > _saiiid_ (i.e. _sayyid_) ‘master’. Word-initial _iu_- (or _yu_-) remains unchanged.

4. ROOT SYSTEM

All Semitic languages have a verbal form and a nominal form (except for functional words and particles), characteristically consisting of a triradical → root and a vocalic pattern which may also require the addition of further consonants. The meaning of a verbal or a nominal form can go beyond the meaning of root and pattern (lexicalization of meaning). Thus, *muʿjam* 'dictionary' belongs to the root *ʿ-j-m* 'to examine, vocalize', and its pattern is *mu12a3* (passive participle of Form IV = *nomen instrumenti*), with 1-2-3 indicating the three root radicals in the word.

Any phoneme except the vowel *a* can be the radical of a root. From the strong consonants *m, r* and the weak radicals *u, i*, and ʾ (glottal stop), the following roots can be formed in Modern Standard Arabic: *m-r-r* ,to pass; to be bitter', *m-u̯-r* 'to move', *m-i̯-r* 'to provide', *m-r-m-r* 'to be bitter', *r-m-m* 'to repair; to decay', *r-u̯-m* 'to desire', *r-i̯-m* 'to move', *r-m-i̯* 'to throw', *ʾ-m-r* 'to order', *ʾ-r-m* 'to bite', etc. These examples show that the weak elements *u, i*, and ʾ cannot simply be regarded as additional elements. As can be seen with *marmara* = *marra* 'to be bitter', verb types can shift from one type to another. Some root types are more common than others; e.g., type 1-2-2 (geminated verbs) is common, type 1-2-1 (e.g. shortened from type *1-2-1-2) is rare, while type 1-1-2 is even rarer.

This root and pattern system has been preserved particularly clearly in Arabic, whose vowels have undergone only very limited sound changes and where there is a general absence of prop vowels (but e.g. *ʾibil* < **ʾibl* 'camels'), vowel elision (but *katf* < *katif* 'shoulder'), and vowel assimilation (but *qiddīs* < **qaddīs* 'holy'). Roots with three and, less frequently, four radicals are almost always clearly recognizable. Only when the vowels *u* or *i* appear as root radicals do vowel contractions occur, e.g. **ramaya* (from the root *r-m-i̯*) > Arabic *ramā* 'he threw', Hebrew *råmå*^{h}, but cf. Classical Ethiopic *rämäyä*, where **aya* appears (→ weak verbs).

In the case of so-called primary nouns, it is difficult to attribute concrete meanings to their underlying roots. The reason may be that some of these often well-documented nouns belong to the oldest stratum of the Semitic lexicon, and this very remoteness may prevent their subsumption under an attested root with a discernible lexical meaning. In other conceivable cases, their roots may by chance not have survived.

Some of these nouns have only two radicals, e.g. *dam* 'blood', *yad* 'arm, hand'; but in these examples it is possible to assign them to the roots *ʾ-d-m* 'to be red' (Arabic *ʾadima*) and *u̯-d-i̯* 'to throw' (Hebrew *yådå*^{h}, Ethiopic *wädäyä*), respectively (→ biradicals). Other so-called biradical nouns are really triradical – observe the weak third radical in the words 'father', 'brother', and 'father-in-law', e.g. Arabic *ʾab(u/i/an), ʾabū/ī/ā*-[suffix] 'father', with similar patterns in many other Semitic languages. The underlying root is *ʾ-b-u̯*, cf. the root *ʾ-b-i̯* 'to refuse', which in Hebrew has the meaning 'to want to do'. This could point to a connection, but it is neither of a synchronic nature nor is it felt to be morphemically connected by native speakers.

5. Nominal morphology

5.1 *Inflection*

There are only two Semitic languages that have fully preserved the Proto-Semitic noun declension: Akkadian (mainly Old Akkadian) and Arabic. In the later periods of Akkadian (Babylonian and Assyrian), as well as in the dialects of Arabic, nominal endings show gradual reduction/elimination (see Table 4).

Observe the almost total identity of the paradigms – except for the external plural of nouns and adjectives. These are the particular differences:

i. Akkadian word-final *-m* corresponds to Arabic *-n*, showing a sound-change *m*/___ #> *n* (cf. Lat. *lupum* with Greek λύκον 'wolf [acc.]').
ii. The *-n* of external masculine plurals is derived from *-m*, if the word-final *-a* of Arabic is seen as a later prop-vowel; Arabic *-ūn-a/-īn-a* is therefore derived from **-ūm/*-īm* (or **-ūt-um/*-ūt-im*).
iii. In the Akkadian forms with *-ū/-ī* in the nominative masculine plural, the word-final nasal must have been dropped. The fact that here etymologically cognate examples cannot be found is easily explainable since external plurals, which are very common in Akkadian, are restricted only to adjectives (and participles) in Arabic. In Arabic, nouns have internal pluralization, while they have

Table 4. Declension in Akkadian and Arabic

		Akkadian	Arabic	
sg. masc.	nom.	*bīt-um*	*bayt-un*	‘house’
	gen.	*bīt-im*	*bayt-in*	
	acc.	*bīt-am*	*bayt-an*	
sg. fem.	nom.	*kalb-at-um*	*kalb-at-un*	‘bitch’
	gen.	*kalb-at-im*	*kalb-at-in*	
	acc.	*kalb-at-am*	*kalb-at-an*	
du.	nom.	*išd-ān*	*ist-ān-i*	‘fundament; [Arabic] posterior’
	gen./acc.	*išd-ī/ēn*	*ist-ayn-i*	
pl. masc. subst.	nom.	*šarr-ū*, *šarr-ān-u*,	*mālik-ūn-a*	‘king; [Arabic] ruling’ [Akkadian and Arabic etyma are not cognate]
	gen./acc.	*šarr-ī*, *šarr-ān-i*	*mālik-īn-a*	
pl. masc. adj.	nom.	*ṣeḫr-ūt-um*	*ṣaġīr-ūna*	‘small’
	gen./acc.	*ṣeḫr-ūt-im*	*ṣaġīr-īna*	
pl. fem.	nom.	*ṭāb-āt-um*	*ṭayyib-āt-un*	‘nice’
	gen./acc.	*ṭāb-āt-im*	*ṭayyib-āt-in*	

external pluralization in Akkadian. In this case, Arabic probably reflects the original situation.

iv. The Akkadian plural in *ān-u/-ān-i* (< *-*ān-ū*/*-*n-ī*) has no equivalent in Arabic, but it has in Old Ethiopic (masc. pl. -*ān*).

The astonishing agreement between Akkadian and Arabic shows the archaic character of Arabic in this respect and allows for easy Proto-Semitic reconstructions.

5.2 *Plural of noun*

The morphological diversity of broken (internal) plurals in Arabic (→ number) is only found elsewhere on a comparable level in Old South Arabian, Modern South Arabian, and Ethiopic. Possibly, some plural patterns may have arisen in individual languages, but many patterns are shared by languages and language groups, e.g. Arabic *kanaf*, pl. *ʾaknāf* = Classical Ethiopic *kənf*, pl. *ʾaknāf, kənäf* ‘wing’; Arabic *ʾalf*, pl. *ʾālāf* (< *ʾaʾlāf*), *ʾulūf*, Sabaic *ʾlf* (*ʾalf*), pl. *ʾʾlf* (*ʾaʾlāf*) ‘thousand’, Classical Ethiopic *ʾəlf*, pl. *ʾaʾlāf* (*ʾaʾəlāf*) ‘thousand’. Plural formation is in many cases quite independent from the singular form. The formation 1V*wa*22V3 is attested in Modern South Arabian and Tigre, cf. Mehri *ġōbər*, pl. *ġəwabbər* ‘pregnant camel’ with Tigre *derho*, pl. *dawarrəh* ‘chicken’.

All the other Semitic languages only show relics of internal plural formation, like Syriac *qrîṯå*, pl. *quʷryå* ‘village’ (cf. Arabic *qarya*, pl. *quran/al-qurā*), Hebrew *zåḵår*, pl. *zəḵūr* ‘male’ (cf. Arabic *ḏakar*, pl. *ḏukūr*; Sabaic *ḏkr*, pl. *ʾḏkr, ḏkwr*).

It is to be assumed that Proto-Semitic had internal plural formation, which experienced expansion in the (former?) ‘South Semitic’ languages, while it became vestigial in the remaining Semitic languages (see Brockelmann 1908–1913:I, 426ff.).

5.3 *Noun formations*

In broad outline, Proto-Semitic noun formation has been well preserved in Arabic. Many nouns of the basic vocabulary have the form *fVʿl* (with V = *a, i, u*), like *bayt* ‘house’, *jins* (< Greek γένος) ‘kind, genus’, *xubz* ‘bread’. In nominal prefixes (almost) exclusively *m*-, *t*-, and ʾ- occur. The prefix *mu*- only occurs in participles (except in the basic stem); *mi*- denotes preferably *nomina instrumenti*, e.g. *miftāḥ* ‘tool with which to open [*f-t-ḥ* ‘to open’], key’; *ma*- serves to denote *nomina loci*, e.g. *maktab* ‘place where to write (*k-t-b* ‘to write’], school’. Also the number of afformatives is very limited: -*ā* (written with y), -*āʾ*, -*ān*, and -*ūt* (adopted from Aramaic), e.g. *sakrān*, fem. *sakrā* ‘drunk’ (*s-k-r* ‘to inebriate oneself’).

Table 5. Nominal and verbal endings

	noun (st. indet.)	noun (st. constr.)	perfect	imperfect/ present	aorist/jussive (apocopate)
sg. masc.	*(-un)*	*-u*	*(-a)*	*-u*	*(-ø)*
du. masc.	*-ā-ni*	*-ā*	*-ā*	*-ā-ni*	*-ā*
pl. masc.	*-ū-na*	*-ū*	*-ū*	*-ū-na*	*-ū*

All languages show some cases where reconstructing is made difficult by idiosyncratic developments and analogical formations (cf. Barth 1894; Fox 2003). Thus, in the case of Arabic *katif/kitf*, Hebrew *kåṯẹp̄* (< **katip*), Syriac *kaṯpå* 'shoulder', it cannot be determined whether the collateral form *kitf* is Proto-Semitic or a later development of Arabic.

6. NOMINAL AND VERBAL INFLECTION

A comparison of the nominal and verbal endings shows extensive parallel formation (see Table 5). With nouns, forms with a nasal are used in the indeterminate state (*bayt-un* 'house'); in the construct state, i.e. before genitive, this nasal disappears (*bayt-u l-maliki* 'house of the king'). To a great degree these nominal endings are identical to the verbal endings (in the 3rd pers. masc.) of the perfect, the imperfect/present (with nasal and *-u* ending), and the aorist/jussive (without these endings, hence the name 'apocopate').

The consonantal elements in the prefix conjugation are in the 3rd person masculine *y-*, feminine *t-*, and in the 2nd person *t-*. The 1st person singular has *ʾ-*, the 1st person pl. *n-*. These elements are ubiquitous in Semitic (with the partial exception of Neo-Aramaic) and in addition also in Semito-Hamitic (Afro-Asiatic).

7. VERBAL MORPHOLOGY

7.1 *Verbal stems*

The verbal stems or themes of Arabic show a fundamental reorganization of the Proto-Semitic situation by adding to the four active basic Forms (Ia–IVa) the respective passive basic Forms (Ip–IVp), as well as to the active reflexive Forms (VIIIa, Va, VIa, Xa) the passive reflexive Forms (VIIIp, Vp, VIp, Xp). The N-Form stands isolated just like other rare Forms, like IX *(i)fʿall-a*, which is related to the elative *ʾafʿal*^{u}, e.g. *(i)ḥmarr-a* 'to become red' – *ʾaḥmar(u/a)* 'red'. This produces the three-dimensional system of Table 6 (with all forms shown in the perfect, and final *-a* indicating the 3rd pers. masc. sg.).

Table 6. Verbal Forms

	active basic Forms		active reflexive Forms
Ia	*faʿal-a*; *faʿil-a*, *faʿul-a*	VIIIa	*(i)ftaʿal-a*
IIa	*faʿʿal-a*	Va	*tafaʿʿal-a*
IIIa	*fāʿal-a*	VIa	*tafāʿal-a*
IVa	*ʾafʿal-a*	Xa	*(i)stafʿal-a*
	passive basic Forms		passive reflexive Forms
Ip	*fuʿil-a; ø, ø*	VIIIp	*(u)ftuʿil-a*
IIp	*fuʿʿil-a*	Vp	*tufuʿʿil-a*
IIIp	*fūʿil-a*	VIp	*tufūʿil-a*
IVp	*ʾufʿil-a*	Xp	*(u)stufʿil-a*

In the active basic stem, the most common pattern by far is *faʿal-a* for verbs of action; much rarer are the patterns *faʿil-a* and *faʿul-a* for stative verbs, as is true for all West Semitic languages. Contrast this with the related stative (permansive) of Akkadian, which in most cases has *i*-vocalization (*paris*). This suggests that the *a*-vowel patterning for active verbs is a West Semitic innovation.

Another innovation of Arabic is the generalized passive formation, although Hebrew has some internal passive forms as well. In the perfect passive, the Arabic vowel pattern *(a-)a-a* is replaced by *(u-)u-i*.

This uniform vocalization of the active verbal stems with generalized *a* also represents an innovation of Arabic (and Ethiopic). All the other Semitic languages show diverse vocalizations. Consider for comparison's sake the Hebrew and Syriac forms in Table 7:

Table 7. Vocalization of verbs in Arabic, Hebrew, and Syriac

Arabic	Hebrew	Syriac
IIa *faʿʿal-a*	pi. *qiṭṭẹ/al, qiṭṭal-tī*	pa. *katteḇ*
IIIa *fāʿal-a*	*(pôlẹl)*	–
IVa *ʾafʿal-a*	hif. *hiqṭîl*	af. *ʾaḵteḇ*

The *i*-vocalization in the intensive and causative stems of Hebrew (**qiṭṭil* and **hiqṭil*) is conspicuous. Here, just as in Arabic, Hebrew has assimilated the vowels, where an original **paʿʿil-* and **hapʿil-* became **piʿʿil-* and **hipʿil-* and in Arabic *faʿʿal-* and *ʾafʿal-*. The original vocalization **paʿʿil-* and **ʾap̄ʿil-* is preserved in Aramaic.

In Hebrew the *pôlẹl* of verbs II*w* constitutes the substitute for the missing **pōʿel* (= Arabic Form III *fāʿala*); *råm* (*r-u̯-m*) 'to be high, to rise'; the stem *rômẹm* 'he erected, elevated' is formed (as if derived from *r-m-m*).

The imperfect shows better correspondence in the individual languages (see Table 8).

If one ignores *u* as the mark of the imperfect of these stems, then the original core morphemes, i.e. that of the imperfect and the perfect, are not distinguished. This is, for instance, the situation preserved in Syriac. By adopting a more generalized *a*-vowel pattern, Arabic has now achieved a clearer distinction between perfect and imperfect forms.

7.2 *Personal elements in the verb conjugations*

As with all Semitic languages, with the exception of a part of Neo-Aramaic, it is convenient to distinguish between a suffix conjugation, called 'perfect' in West Semitic, and one or more prefix conjugations. Almost always the personal elements correspond well in essence.

The dental elements in the perfective endings (e.g. *faʿal-tu* 'I made', *faʿal-ta* 'you [masc. sg.] made'), which correspond with Ethiopic (*-**ku***, *-**kä***) and a velar element in Modern South Arabian, have come about by paradigmatic pressure from *k* in the 1st person singular and *t* in the other forms. The paradigm is preserved in Akkadian (1st pers. sg. *pars-**āku***, 2nd pers. masc. *pars-**āta***). Arabic no longer shows any traces of this ***ā*** (but more can still be found in Hebrew, e.g. *bînôtå*).

The essential difference in the prefix conjugations concerns the endings of the 2nd and 3rd person feminine plural (see Table 9).

Table 8. Imperfect in Arabic, Hebrew, and Syriac

Proto-Semitic	Arabic	Hebrew	Syriac
**yupaʿʿil-*	II^a *yufaʿʿil-*	pi. *yəqiṭṭẹl*	pa. *nḵatteḇ*
**yupāʿil-*	III^a *yufāʿil-*	*(yəp̄ôlẹl)*	–
**yus^{1}apʿil-*	IV^a *yufʿil-*	hif. *yaqṭīl*	af. *naḵteḇ*
(> **yuhapʿil-*)	(<**yuØapʿil-*)	(<**yuØaqṭil-*)	

Table 9. Endings of the 2nd/3rd person feminine plural in the imperfect

		Arabic	Hebrew	Classical Ethiopic	Syriac
pl.	3 masc.	*yaktubū*	*yiḵtəḇū*	*yəṣḥafū*	*neḵtḇûn*
	fem.	*yaktub**na***	*tiḵtōḇ**nå***^{h}	*yəṣḥaf**ā***	*neḵtḇ**ān***
	2 masc.	*taktubū*	*tiḵtəḇū*	*təṣḥafū*	*teḵtḇûn*
	fem.	*taktub**na***	*yiḵtōḇ**nå***^{h}	*təṣḥaf**ā***	*teḵtḇ**ān***
	1	*naktubu*	*niḵtōḇ*	*nəṣḥaf*	*neḵtôḇ*
		(imperfect)	(imperfect)	(jussive)	(imperfect)

Here the Arabic and Hebrew forms with masculine *-ū*, feminine *-na/nā* are isolated; most other Semitic languages (like Old Ethiopic and Syriac) have masculine *-ū(n)*, feminine *-ā(n)*. Since Akkadian has the endings masculine *-ū*, feminine *-ā*, these are deemed the best matches for Proto-Semitic reconstruction. The nasal element in the 2nd and 3rd person feminine of Arabic and Hebrew can be explained thus: the endings of feminine plural were originally *-āt* and feminine dual *-ān*. Because of the elision of word-final *-t* and *-n*, plural and dual could no longer be distinguished, which led to the adoption of the feminine ending *-na* of the independent personal pronoun. According to Hetzron (1976), this is the decisive isogloss for classifying both languages in one grouping, Arabo-Canaanite.

7.3 *Tense/aspect forms*

The suffix conjugation expresses in nearly all West Semitic languages a past in active verbs (vocalized with internal *a*) and stative meaning with stative verbs (with internal *i* or *u*).

In contrast, the prefix conjugations have shown greater diversity since Proto-Semitic times. Of the three prefix conjugations of Proto-Semitic and Akkadian (aorist *ipr*V*s*, present *iparr*V*s* and the (East Semitic) perfect *iptar*V*s*), only *ipr*V*s* has remained in Arabic in different verbal categories:

i. Imperfect: 3rd pers. masc. sg. *yaktub-u*, pl. *yaktub-ūna*: This form replaces in its function as a present (progressive form) the Old Semitic present (*iparr*V*s*) and continues the Akkadian subordinative (subjunctive), i.e. the form of the aorist used for instance in relative clauses (*modus relativus*, e.g. *ša iprus-u* 'who sent'). This form corresponds with Hebrew *yåqûm* (< **yaqūmu*) 'he rises'.
ii. The old aorist (= Akkadian preterite) 3rd pers. masc. sg. *yaktub*, pl. *yaktub-ū* has in Arabic only been preserved in conjunction with the particles *lam* 'not', *lammā* 'not yet', e.g. *lam yaqum* 'he did not rise'. This form corresponds with the Hebrew short form *way-yåqåm* (< **wa-yaqūm*) 'and he rose'.
iii. Jussive 3rd pers. masc. sg. *yaktub*, pl. *yaktub-ū* is preserved in connection with the conjunction *li-* (*li-yaqum* 'he is to rise') and the negative *lā* (*lā taqum* 'you [masc. sg.] are not to rise'). This form corresponds with the Hebrew jussive *yåqom* (< **yaqum*) 'he may rise'. The twofold use of the short form as a narrative (past) as well as jussive is also attested in Akkadian and can be assumed to be Proto-Semitic. Both forms were distinguished by the accent: narrative **yáktub* (Hebrew *way-'yåqåm*) as opposed to jussive **yaktúb* (Hebrew *yå'qom*).
iv. Conjunctive 3rd pers. masc. sg. *yaktub-a*, pl. *yaktub-ū* only occurs after conjunctions, e.g. *ʾan yaqūma* 'that he rise'. This form continues the ventive of Akkadian (and Proto-Semitic) in *-am*, pl. *-nim*.
v. → Energicus 3rd pers. masc. sg. *yaktub-an(na)*, pl. *yaktub-un(na)* (< **yaktub-ūn(na)*) is presumably derived from the ventive. It corresponds with the Hebrew cohortative in *-ā: ʾeqṭəlåh* (pausa *ʾeqṭolåh*) 'I want to kill' and the forms with reinforced object suffix: *yəqṭəl-ęnnū*, *yəqṭəl-ęn-hū* (< **yiqṭol-an-hū*) 'he kills him'.

This system corresponds in large part with that of Hebrew, so that a particularly close relationship between these two languages must be postulated. The special difference with Hebrew lies in this construction, where with the perfect (as with the aorist) a distinction is made between narrative and jussive, and these forms (as with the aorist) are marked by differently stressed syllables, cf. *qå'ṭáltå* 'you [masc. sg.] killed' vs. *wə-qåṭal'tå* 'you [masc. sg.] will kill'. This trait seems to be a feature of Proto-Semitic and not a later development in Hebrew (in analogy to the opposition narrative **yáktub* vs. jussive **yaktúb*).

8. LEXICOGRAPHY AND SEMANTICS

Of all Semitic languages, Arabic has the greatest volume of rich literature. The lexicographical work of the indigenous grammarians exceeds that of the other Semitic languages by far. This gives special weight to Arabic in etymological comparisons, but it can also sometimes prove to be a disadvantage because these comparisons may become biased toward Arabic. The enormous lexicon of Arabic with many roots similar in shape and meaning tends to induce some scholars to favor the reduction of triradical to biradical roots.

The disadvantage of the copious indigenous dictionaries is that too many meanings are offered and mere ad hoc rhyme words are often included. Therefore, it is advisable to proceed with caution in any comparison with Arabic.

Bibliographical references

Barth, Jacob. 1984. *Die Nominalbildung in den semitischen Sprachen*. 2nd ed. Leipzig. (Repr., Hildesheim: G. Olms, 1967.)

Bauer, Hans and Pontus Leander. 1922. *Historische Grammatik der hebräischen Sprache des Alten Testaments*. Halle/Saale. (Repr., Hildesheim: G. Olms, 1965.)

Brockelmann, Carl. 1908–1913. *Grundriss der vergleichenden Grammatik der semitischen Sprachen*. 2 vols. Berlin. (Repr., Hildesheim: G. Olms, 1966.)

Dolgopolsky, Aron. 1999. *From Proto-Semitic to Hebrew: Phonology*. Milan: Centro di Studi Camito-Semitici.

Fleisch, Henri. 1961, 1979. *Traité de philologie arabe*, I, II. Beirut: Imprimerie Catolique.

Fox, Joshua. 2003. *Semitic noun patterns*. Winona Lake, Ind.: Eisenbrauns.

Hetzron, Robert. 1969. “The evidence for perfect **y'aqtul* and jussive **yaqt'ul* in Proto-Semitic”. *Journal of Semitic Studies* 14.1–21.

——. 1976. “Two principles of genetic reconstruction”. *Lingua* 38.89–102.

Huehnergard, John. 1992. “Historical phonology and the Hebrew Piel”. *Linguistics and Biblical Hebrew*, 209–229. Winona Lake, Ind.: Eisenbrauns.

——. 2003. “Akkadian *ḫ* and West Semitic **ḥ*”. *Orientalia – Trudy Instituta Vostočnych Kul'tur* 3.102–119.

Knauf, Erst Axel. 1989. *Ismael: Untersuchungen zur Geschichte Palästinas und Nordarabiens im 1. Jt. v.Chr.* 2nd ed. Wiesbaden: O. Harrassowitz.

Steiner, Richard C. 1977. *The case for fricative-laterals in Proto-Semitic*. New Haven, Conn.: American Oriental Society.

——. 1982. *Affricated* ṣade *in the Semitic languages*. New York: American Academy for Jewish Research.

Voigt, Rainer. 1987. “The classification of Central Semitic”. *Journal of Semitic Studies* 32.1–21.

——. 1992. “Die Lateralreihe */ś ṣ ź/* im Semitischen”. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 142.37–52.

Rainer Voigt (Free University of Berlin)

Senegal

1. Arabic in Senegal

Senegal is a predominantly Muslim country of sub-Saharan Africa, situated on the west coast of Africa. It is bordered by → Mauritania to the north, → Mali to the east, and Guinea and Guinea Bissau to the south. The Arabic language was introduced into Senegal as early as the 11th century C.E. with the southward spread of Islam, brought by Berber merchants and militants from Morocco and Mauritania. Currently, Arabic is used as a religious language and as a medium of education in religious circles; it is taught in the public school system and is used among certain segments of the Senegalese population as a written language. There are also small but significant minorities of native Arabic-speaking populations of Mauritanian, Syrian, and Lebanese origin living in Senegal. Because of their proximity to Senegal, Mauritanians have interacted in various ways with Senegalese populations for centuries. Syrians and Lebanese arrived under French colonial domination and typically served as middleman merchants during the colonial period. Many of the Syrian and Lebanese families living in Senegal today have been there for at least three generations, and many continue to run businesses.

In both its spoken and written forms, Arabic has influenced indigenous Senegalese languages. These languages have incorporated a sizable number of terms and expressions from Arabic into their lexicons, and the Arabic script has also been adapted as a writing system for local languages, primarily Pulaar (Fula; → Fulfulde), → Wolof, and Mandinka.

2. Islam in Senegal

According to reports dating from 1068 C.E. by the Andalusian writer al-Bakrī, by the 11th century Islam appears to have been well established in the Pulaar-speaking royal courts in the kingdom of Tekrur, located in the lower Senegal River valley in an area now known as the Fuuta Toro. By the late 18th century, Muslim clerics of the Fuuta Toro had established themselves as a clerical class known as the Torodɓe. The most famous member of this group is undoubtedly al-Ḥājj ʿUmar Tal, a Tijani Jihadist born in Podor at the end of the 18th century who led jihads against various neighboring non-Muslim regimes, so that by the late 19th century Islam was firmly implanted in this part of West Africa. French colonization of Senegal and the resultant dissolution of the precolonial kingdoms served as a further impetus for mass conversions to Islam at the end of the 19th and beginning of the 20th centuries, especially among the Wolof in the interior of the country. It is at this troubled point in Senegalese history that many of the most prominent religious figures emerged. Primary among these are the marabouts al-Ḥājj Malik Sy (1855–1923), founder of the Tivaouane *zāwiya* of the Tijaniyya in Senegal, and Amadu Bamba Mbacke (1855–1927), founder of the indigenous Murid Sufi order centered on the holy city of Touba. Both of these men were savants and literary figures who produced important bodies of writing in Arabic. With the last substantial wave of conversions to Islam taking place in the mid-20th century, current census figures show that of Senegal's ten million inhabitants, today approximately 94 percent are Muslim.

Islam in Senegal is practiced almost exclusively within the Saharan Sufi model that emphasizes the role of the shaykh or spiritual leader, known as a *marabout* in West Africa (Robinson 2004). The word is derived from *al-murābiṭūn* 'those of the fortress', referring to the Almoravids, the militant Saharan Berbers whose Islamic Empire (1042–1148 C.E.) encompassed Morocco, Mauritania, and parts of Spain and western Algeria. The predominant Sufi orders in Senegal today are the Tijaniyya, the Qadiriyya, and two indigenous orders, the Layene order founded by Seydina Limaamu Laye, a Mahdist figure who died in 1326/1909, and the influential Muridiyya, founded by Amadu Bamba Mbacke (d. 1345/1927). A mendicant subsect of the Muridiyya emerged around the figure of Bamba's most ardent follower, Shaykh Ibra Fall. His followers are known as the Baye Fall and substitute hard physical work for prayer and fasting.

The influence of Islam in Senegal spreads beyond the domain of religion into the practice of daily life and popular culture. Popular art forms such as the renowned Senegalese reverse glass paintings frequently depict religious scenes, especially miracles from the life of Shaykh Amadu Bamba, and hagiographic portraits of the most important marabouts. Likewise, shopkeepers and restaurant owners commission such portraits to decorate their businesses and attract customers, and portraits of marabouts often turn up on city walls. Forms of transportation such as buses and small boats are also decorated with religious motifs, especially short prayers and blessings in Arabic that serve as protection against misfortune and accidents. The role of the marabout as a spiritual guide is widely celebrated in Senegalese popular music of all styles, ranging from Afro-Cuban salsa to rap, representing an adaptation of the praise-singing tradition of the Sahelian griots (Mc Laughlin 1997). Members of certain Sufi orders, especially youth, are often readily identifiable by their dress, and certain aspects of specifically Sufi garb have carried over into mainstream fashion. Most conspicuously, the patchwork clothing and dreadlocks of the mendicant Murid subsect, the Baye Fall, have been adopted as a popular look among young people. As these cultural practices illustrate, Islam is an essential component of identity for the vast majority of Senegal's Muslim population.

3. ARABIC AS A RELIGIOUS LANGUAGE

As the language of Islam and the *Qurʾān*, Arabic enjoys high status as a religious language throughout the country. So closely is the language associated with religion that Wolof speakers often refer to Arabic as *alxuraan*, the Wolof word for the *Qurʾān*. Widespread attendance at Qurʾānic school or *daara* (probably < Arabic *dār* 'house'), as it is known in Senegal, is the norm for both boys and girls, and most people are introduced to the Arabic language and writing system in this way. For the majority of the population, Qurʾānic school is their only formal training in Arabic, and even then, the focus is on rote memorization and recitation rather than on learning the Arabic language. Thus, although the vast majority have been exposed to Arabic, it is only a small minority who actually master the language. There is, however, a strong tradition of Arabic transmission among certain religious communities both in the Pulaar-speaking Fuuta Toro and the Wolof heartland.

Given the history of Senegal and the nature of the historical moments that favored conversion to Islam, certain pre-Islamic practices have been incorporated into the practice of Islam among the general population. Chief among these is the conflation of the role of the marabout as Islamic savant and teacher, healer, and even diviner. This has in turn had an important influence on the use of spoken and written

Arabic and Qurʾānic verses as instruments of healing and protection. In many cases, the newer Islamic practices are fused with practices that come from pre-Islamic knowledge and belief systems. Foremost among these is the use of leather amulets, known as *téere* in Wolof, that contain written verses of the *Qurʾān* along with substances such as powders, leaves, or animal parts believed to have healing or protective properties. Prayers and Qurʾānic verses may also be recited during the fabrication of an amulet in order to enhance its efficacy. These amulets are worn on the body under clothing, or hung inside or outside a house to protect the inhabitants. Amulets may also contain writings based on esoteric Sufi knowledge of the mysteries associated with the 28 letters of the Arabic alphabet and their numerical values, arranged in appropriate geometrical patterns. Such drawings or mystical figures, known as *xaatim* in Wolof (< Arabic *xātim* 'seal'; the *xaatim* is the marabout's seal that he puts on a drawing to validate it), are widely used in the Sahel and the Sahara. *Xaatim* can also be inscribed on special shirts that resemble hunters' tunics, providing the wearer with protection or blessings (Roberts and Roberts 2003:174–177).

Another common practice involving the written word in Arabic is that of the preparation and consumption of *saafara*, or holy water. Appropriate Qurʾānic verses or other prayers are written in black ink on a surface such as a wooden board, and the words are then rinsed off with water. The water containing these words is saved, and, according to maraboutic instructions, the afflicted person may use the *saafara* to bathe in or drink in order to be healed. Similar practices involving holy water are prevalent throughout the Muslim parts of West Africa. Other examples of the use of written and spoken Arabic come from the domain of child rearing. Amulets are generally fastened around children's bodies at birth to protect them, but there are other more specific practices, including those related to weaning and teething. A widespread weaning practice involves writing an appropriate verse of the *Qurʾān* on a piece of bread that is given to a nursing child to eat. After the child consumes the bread, he or she is weaned. Teething necklaces are made by stringing small beads together with knots between them. At the tying of each knot a specific prayer is said, imparting the necklace with properties to assuage the child's teething pain. These are just some of the many ways in which the spiritual properties of spoken and written Qurʾānic verses and prayers have been incorporated into practices of healing and protection in Senegal.

Religious practice in Senegal involves numerous Sufi ceremonies throughout the Muslim calendar year. These include the *Grand Màggal* of the Murid order, which commemorates the return of Shaykh Amadu Bamba from exile in Gabon, the annual *Gàmmu* of the Tijaniyya, and other ceremonies dedicated to the glory of the Prophet and various Sufi saints. During these ceremonies there is extensive preaching by marabouts and other distinguished figures, and while most of it is done in local languages, it is almost always punctuated by long passages in Arabic, including recitations of parts of the *Qurʾān*. Groups of singers also participate in these ceremonies, performing musical versions of the religious writings in Arabic of renowned Senegalese poets, which are usually hagiographies or praise poems to the Prophet.

The use of Arabic writing for decorative purposes is widely appreciated in Senegal and has often been adapted to local aesthetics. Iconic portraits of local marabouts can be rendered in calligraphy, and the *xaatim*, or magic squares, are often incorporated into men's jewelry, and especially rings, as part of aesthetic expression. Some contemporary artists have also explored the theme of calligraphy in their work, producing series of paintings and intricate paper cutouts of the letters of the Arabic alphabet.

As a religious language, Arabic is used regularly in Senegal in prayers and religious ceremonies, and in healing practices. While almost all Muslims know how to recite the basic prayers and sections of the *Qurʾān* in Arabic, it is only a small religious elite who actually master the language. Mastery of Arabic is greatly admired, and those who have accomplished such a task can use it to their advantage in a career in the religious and spiritual domain.

4. Sufi mystics and their writings

Senegal has produced a sizable body of literature in Arabic, most of which is the work of Sufi mystics, Jihadists, and poets who have written religious poetry and hagiographies of the Prophet. Although many scholars suspect that there may well have been a body of literature dating from earlier periods but that has now been lost, the earliest known texts currently date from the middle of the 19th century (Gérard 1981). What is clear, however, is that this tradition reflects fluency in Arabic and extensive knowledge of Arabic poetic traditions among a substantial segment of the Senegalese religious elite (Abdullah 2004). The appearance of this literature coincides with the founding and development of several *zāwiyas*, or centers of learning, throughout the country, and the intellectual and spiritual environment that the *zāwiyas* offered encouraged the production of scholarly and religious writing in the Arabic language (Samb 1972). The earliest Arabic writings are the work of the Tijani Jihadist from Fuuta Toro, al-Ḥājj ʿUmar Tal, who was also a scholar and poet. His most influential work, which has subsequently become one of the most important books for the Tijani order, is entitled *Rimāḥ ḥizb ar-raḥīm ʿalā nuḥūr ḥizb ar-rajīm* 'The lances of God's party against the throats of the Satanic party'. This work was completed in 1845, seven years before he launched his West African jihad, and describes customs and practices of the Tijaniyya, as well as al-Ḥājj ʿUmar Tal's own account of his pilgrimage to Mecca and his sojourn in Sokoto in Hausaland on his way back to Fuuta Toro.

Two other Senegalese literary luminaries who wrote in Arabic were al-Ḥājj Malik Sy, founder of the Tivaouane *zāwiya* of the Tijaniyya, and Shaykh Amadu Bamba Mbacke, founder of the Muridiyya. Al-Ḥājj Malik's work comprises both prose and poetry and encompasses subjects as diverse as linguistics, history, theology, and the correct practice of Islam. His religious poetry in Arabic, including his biography of the Prophet entitled *Xilāṣ aḏ-ḏahab* 'Pure gold', is often sung at Tijani religious ceremonies today. The Tivaouane *zāwiya* was renowned as a center of learning and writing and produced many other fine poets, some of whom were sons of al-Ḥājj Malik, who wrote in Arabic. Shaykh Amadu Bamba Mbacke, founder of the indigenous Sufi order of the Muridiyya, and of the Murid *zāwiya* in Touba in the late 1880s, was also a prolific writer, and his Arabic language poetry, especially his *xasaayids* (> Arabic *qaṣāʾid*), are readily available for sale in any Senegalese market to this day. Repeated recitation of Bamba's *xasaayids*, which speak, among other things, of the disciple's longing for God, is known to put disciples into a spiritual trance, a phenomenon known as *daanu leer* in Wolof.

A second influential Tijani *zāwiya* is that of Kaolack, which has been dominated by the maraboutic Niasse family since the beginning of the 20th century. Al-Ḥājj Ibrahima Niasse (1902–1975) and his brother Mohammadou Niasse (1881–1957) wrote collections of poetry in Arabic, some of which were published in Kano in northern Nigeria, and Ibrahima's daughter, Roqaya Niasse, published a pamphlet in Arabic entitled *Tanbīh al-bint al-muslima fī d-dīn wa-d-dunyā* in Dakar in 1954 concerning the comportment and duties of Muslim women, in which she shows herself to be a strong supporter of girls' and women's education (Samb 1972:236–241).

In addition to these major religious figures and members of their families who wrote important works in Arabic, there were many other writers who contributed to this body of literature with works both religious and secular. In most cases, the principles of Classical Arabic verse and rhyme were maintained, and eventually even transferred over to poetic works written in indigenous languages like Pulaar and Wolof. Most Senegalese Muslims have been exposed to these Arabic writings in their written, but more often oral, forms, at religious ceremonies, and many have memorized verses or passages from them. The tradition of Arabic literary creation in Senegal is thus considered to be a valued aspect of the country's cultural legacy and merits more scholarly attention than it has thus far received.

5. The influence of Arabic on Senegalese languages

The major languages of Senegal belong to the Atlantic and Mande branches of the vast Niger-Congo family. Because of the long history of contact with Arabic as a religious language, many Senegalese languages have borrowed lexical items from Arabic in certain semantic domains (Mouradian 1940). As the lingua franca of Senegal, → Wolof, an Atlantic language, will be used to illustrate some of these loans. In the domain of religion, Wolof unsurprisingly has numerous words of Arabic origin, including *barke* 'blessing, grace' (< *baraka*), *jumaa* 'mosque' (< *jumʿa* 'Friday'), *malaaka* 'angel' (< *malʾak*), *àjjana* 'heaven' (< *al-janna* 'paradise'), *àllaaji* 'male pilgrim to Mecca' (< *al-ḥājj*), *sajjaada* 'prayer rug' (< *sajjāda*), *ilimaan* 'imam' (< *al-ʾimām*), *seytaane* 'devil' (< *šayṭān*), *xalwa* 'spiritual retreat' (< *xalwa*). Many time expressions are similarly borrowed from Arabic, including *saa* 'moment' (< *sāʿa* 'hour'), *waxtu* 'hour' (< *waqt* 'time'), *fajar* 'dawn' (< *fajr*), *suba* 'morning, tomorrow' (< *ṣubḥ)*, and the days of the week from Monday to Friday: *altine, talaata, àllarba, alxames, àjjuma*. Numbers have not been borrowed, although the Arabic word for 'first' has been borrowed as *lëwël*, to mean specifically the first of three glasses of tea that are drunk in a sequence. There is an extensive system of greetings in Wolof, and although most of the formulaic sequences do not show the influence of Arabic, the greeting is generally bounded by expressions that come from Arabic at the beginning, with *salaam aleekum* 'peace be with you' (*as-salāmu ʿalaykum*), and at the end, with *alxamdulilay* 'by the grace of God' (< *al-ḥamdu lī-llāhi*). Other common expressions that come from Arabic include *astafurlaa* 'may God forgive me' (*ʾastaġfiru llāh*), which is used when correcting oneself, and the word *naam* 'yes' (< *naʿam*), which is used to respond when one's name is being called, although the general Wolof word for yes is *waaw*. The *naamu-naamu*, among the Mande-speaking population, is the person who lends veracity to the griot's performance of a narrative or epic by repeating formulaic praises such as 'Yes, what he says is true'.

Arabic given names, too, are very frequent among the Muslim population in Senegal. Additional Arabic-inspired names come from titles such as Maalik 'king', Seex 'shaykh', and As 'al-Ḥājj'. There are also some Senegalese place names that come from Arabic, such as the popular Medina, as well as Fas (Fes), and Lam Lam (two of the letters that form the word for God). There are many other loans from Arabic in Wolof and other Senegalese languages, but these examples give a sense of the distribution of such loans in the languages.

6. Ajami writing

Some of Senegal's main languages, especially Wolof, Pulaar, and Mandinka, have a well-developed and fairly standardized writing system in the Arabic script, usually called *ajami*. These scripts were once widely used for keeping records and writing letters, often through the services of a scribe, and although they are still used extensively in certain circles, increased rates of literacy in the Roman alphabet, especially in urban areas, mean that reliance on the Arabic script is only one of two choices for writing in indigenous languages, and many younger people prefer to use the Roman script, which is more familiar to them. Wolof written in the Arabic script is called *wolofal*, a word composed of the stem Wolof plus a causative suffix to give the meaning 'to make Wolof.' *Wolofal* is widely used in record keeping and correspondence in the rural areas of the Wolof heartland in northern central Senegal, and by members of the Murid Sufi order, which has its origins in that part of the country (Dème 1996). Signs and public messages in rural areas of Senegal, and to a lesser extent in urban areas, are often in *wolofal* (see Figs. 1 and 2). The Arabic alphabet has been adapted, primarily through a series of diacritics, to accommodate Wolof phonemes that do not exist in Arabic, such as the prenasalized stops [^{m}b], [^{n}d], [^{n}j], [^{n}g], and [^{n}q], and various other nasal sounds. There are also many publications in *wolofal*, including books of poetry and religious instruction (see Fig. 3). Similar adaptations of the Arabic alphabet have been made for Pulaar phonemes, such as the implosive stops [ɓ], [ɗ], and [ʄ]. Ajami script is generally preferred for writing of a religious nature, and even today, religious poetry written in Pulaar in the Fuuta Toro region of northern Senegal is written in ajami.

7. Arabic in education

In the Muslim tradition, in Senegal as elsewhere in the Muslim world, religious education is centered on the study of the *Qurʾān*. Senegalese Qurʾānic schools or *daara*, as they are known locally, are run by marabouts who undertake the spiritual and religious education of children. In many cases, children attend the *daara* on a regular basis until the age of seven, when they start state-run primary school. Many continue to go to the *daara* in the afternoon or evening, when they return from regular school, or on weekends. In other cases, children may be sent to live in a marabout's household with other boarders, sometimes for a period of several years, and may not attend the state-run primary school at all during this period

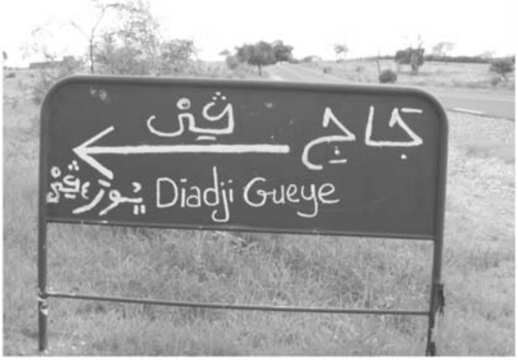


Figure 1. A roadside sign in rural Senegal in *wolofal* and Roman script, indicating the way to the house of a certain Diadji Gueye.

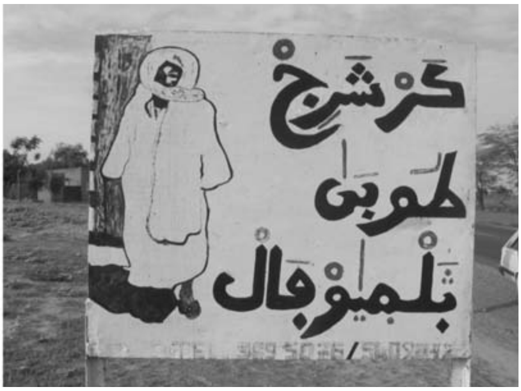


Figure 2. A roadway sign in written in *wolofal* with a portrait of Shaykh Amadu Bamba, marking the location of a Murid marabout's house.

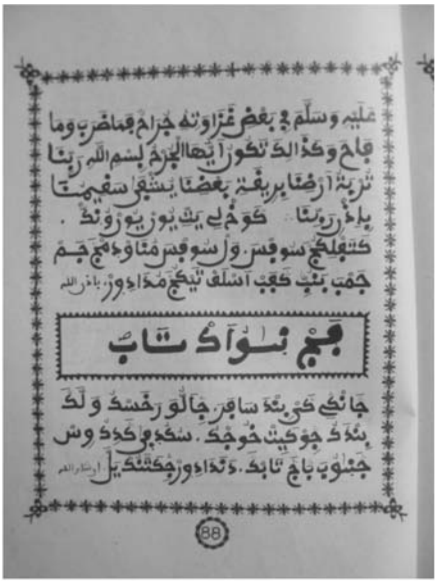


Figure 3. Page from a Muslim's guidebook in *wolofal* showing how to treat a boil with *saafara.*

of their lives. In such cases, pupils may also be required to perform other duties such as engaging in agricultural labor for the marabout in rural areas or sometimes even begging for alms in urban areas, although this latter type of activity is discouraged by the government, and many nongovernmental organizations have launched campaigns to help get such children off the streets. Cheikh Hamidou Kane's classic novel *L'aventure ambiguë* 'Ambiguous adventure', published in 1961, gives a vivid fictional account of a Qurʾānic school education in Torodɓe society in northern Senegal at the middle of the 20th century.

Depending on the size of the school, the marabout may be responsible for all instruction, but in most cases there are other qualified teachers, or even promising senior students, who teach the recitation and writing of the *Qurʾān* to young pupils. Instruction consists primarily of rote memorization and in some cases reading of the *Qurʾān* and *ḥadīṯs* in Arabic. In schools where reading is taught, pupils write the day's lesson, usually a verse or two from the holy book, on a wooden board known as an *àlluwa* in Wolof, and recite it until they have it memorized. In certain Murid *daaras* pupils may also learn to recite some of Shaykh Ahmadu Bamba's poetry, and especially the *xasaayids*, in Arabic, or poems in Wolof about his life and the miracles he performed. The primary goal of early religious education is to learn to recite the entire *Qurʾān*, although few manage to achieve this goal. The accomplishment of those promising pupils who manage to learn the entire *Qurʾān* is celebrated in a night of performance attended by their families and friends in addition to their fellow pupils and religious leaders.

Students who continue their education with a marabout may also add other subjects, such as the theory of Qurʾānic recitation, *šarīʿa*, and Arabic language and grammar, to their educational program. While much of this kind of teaching in Arabic is part of Senegal's informal educational system, there are nonetheless certain private schools that provide a general primary and secondary education in Arabic as well. Promising students from such schools may then go on to study at universities in the Arabic-speaking world, typically in Morocco, Egypt, or Sudan.

The medium of instruction in the Senegalese public school system is French, the official language. Arabic, however, is offered as an elective subject at all levels of primary and secondary school and is included in the list of foreign languages for the *baccalauréat* exam, but it is less popular than other languages like English because of its association with religion rather than opportunity. The study of Arabic in the public school system is completely secular in nature, focusing on Arabic grammar and reading, writing, and spoken Modern Standard Arabic, thus the curriculum for learning Arabic as a foreign language is similar to that for other foreign languages such as English or Spanish. Arabic teachers in the primary and secondary school systems are generally Senegalese and come from a wide variety of educational backgrounds, including those who have studied at local Arabic schools as well as many who have studied abroad in Arabic-speaking countries. There are many private schools at the primary and secondary levels that incorporate Arabic into their curriculum. In some private schools, all classes are taught in Arabic, but much more prevalent are the so-called Franco-Arabic schools where students receive a bilingual education that has both secular and religious aspects to it. These hybrid schools have increased significantly in popularity during the last decade or two, and a substantial number of them receive external funding from Arab countries or other predominantly Muslim countries, such as Turkey.

It is possible to pursue a postsecondary education in Senegal focusing on Arabic, but in order to be admitted to state universities in Senegal, students must have passed the *baccalauréat* exam, which is given in French, thus those who have studied exclusively in Arabic-speaking schools rarely pursue this option. Instead, they may win scholarships to go abroad to pursue university degrees in Arabic-speaking countries. At Senegal's main university, Université Cheikh Anta Diop in Dakar, the Department of Arabic offers degrees up to the level of the doctorate in Arabic language and literature.

8. Conclusion

Arabic in its spoken and written form is a vital part of Senegal's linguistic environment, reflecting the country's Muslim heritage. Arabic has been adapted in many ways and has influenced local languages in their spoken and written forms as well. Many cultural practices, especially in the religious domain, have incorporated the use of Arabic as an efficacious tool in protecting and healing individuals. The Arabic language holds a position of great prestige in Senegal, a situation that is unlikely to change given the immense respect for Islam that characterizes the society as a whole.

Bibliographical references

Abdullah, Abdul-Samad. 2004. "Arabic poetry in West Africa: An assessment of the panegyric and elegy genres in Arabic poetry of the 19th and 20th centuries in Senegal and Nigeria". *Journal of Arabic Literature* 35.368–390.

Dème, Abdoulaye. 1996. "Écriture dans une société à tradition orale: Le cas du wolof au Sénégal: le 'wolofal'". *Les politiques linguistiques, mythes et réalités*, ed. Caroline Juillard and Louis-Jean Calvet, 149–153. Beirut: Fiches du Monde Arabe.

Gérard, Albert S. 1981. *African language literatures: An introduction to the literary history of sub-Saharan Africa.* Washington, D.C.: Three Continents Press.

Kane, Cheikh Hamidou. 1961. *L'aventure ambiguë.* Paris: Julliard.

Mc Laughlin, Fiona. 1997. "Islam and popular music in Senegal: The emergence of a 'new tradition'". *Africa* 67.560–581.

Mouradian, Jacques. 1940. "Notes sur quelques emprunts de la langue wolof à l'arabe". *Bulletin de l'Institut Français d'Afrique Noire* (Dakar) 2.269–294.

Roberts, Allen F. and Mary Nooter Roberts. 2003. *A saint in the city: Sufi arts of urban Senegal.* Los Angeles: UCLA Fowler Museum of Cultural History.

Robinson, David. 2004. *Muslim societies in African history.* Cambridge: Cambridge University Press.

Samb, Amar. 1972. *Essai sur la contribution du Sénégal à la littérature d'expression arabe.* Dakar: Institut Fondamental d'Afrique Noire.

Fiona McLaughlin (Gainesville)

Sentence

1. The simple sentence

This entry examines salient syntactic differences in simple and complex sentences between Standard Arabic and Moroccan Arabic (see also → verbal clause; → nominal clause). In both varieties, the → word orders Verb-Subject-Object (VSO), as in (1a) and (2a), and Subject-Verb-Object (SVO), as in (1b) and (2b), are used alternately, although the latter is the unmarked surface order.

Standard Arabic
(1a) *kataba* *ḥātim-un* *risālat-an*
wrote.3ms Hatim-Nom letter-Acc
'Hatim wrote a letter'
(1b) *ḥātim-un* *kataba* *risālat-an*
Hatim-Nom wrote.3ms letter-Acc
'Hatim wrote a letter'

Moroccan Arabic
(2a) *kteb* *ḥatim* *bra*
wrote.3ms Hatim letter
'Hatim wrote a letter'
(2b) *ḥatim* *kteb* *bra*
Hatim wrote.3ms letter
'Hatim wrote a letter'

In both varieties, SVO is the most commonly used word order. In Standard Arabic, but not Moroccan Arabic, the morphological endings on the arguments of the verb determine the functions of the noun phrases. In examples (1a-b), the subject has nominative case *-un*, and the object has the accusative → *tanwīn -an*. These case endings are not influenced by the word order change. In Moroccan Arabic, these grammatical functions are not indicated by case marking but by the syntactic position of the relevant phrase in the sentence. However, there are instances in Standard Arabic where the subject and the object noun phrases have no overt case endings, as in (3a–b).

(3a) *ḍaraba* *ʿīsa* *mūsā*
hit.3ms Issa Moussa
'Issa hit Moussa'
(3b) *ḍaraba* *mūsā* *ʿīsā*
hit.3ms Moussa Issa
'Moussa hit Issa'

In (3a), the subject noun phrase *ʿīsā* and the object noun phrase *mūsā* are determined by their syntactic position. In (3b), the meaning is different because the first noun phrase after the verb is the subject and the second one is the object.

Standard Arabic and Moroccan Arabic may also exhibit Verb-Subject-Object (VOS), as in (4a), and Object-Verb-Subject (OVS) word order, which are less common and morphologically marked, as in (4b).

Standard Arabic
(4a) *kataba* *r-risālat-a*
wrote.3ms the-letter-Acc
ḥātim-un (OVS)
Hatim-Nom
'Hatim wrote the letter'
(4b) *ar-risālat-a* *kataba*
the-letter-Acc wrote.3ms
ḥātim-un (OVS)
Hatim-Nom
'Hatim wrote the letter'
(4c) *ar-risālat-u* *kataba-hā*
the letter-Acc wrote.3ms-cl3fs
ḥātim-un
Hatim-Nom
'The letter, Hatim wrote it'

In Standard Arabic, the VOS order requires the object to be definite. In Moroccan Arabic, the definite article is not compulsory, but a pause is necessary after the object. In OVS sentences in Standard Arabic, a resumptive pronoun is not required, as in (4b). If the object is in topic position, it receives nominative case; in this case, a resumptive pronoun is necessary, as in (4c). In Moroccan Arabic, a resumptive pronoun is required, as in (5).

Moroccan Arabic
(5) *l-bra* *kteb-ha* *ḥātim*
the-letter wrote.3ms-cl3fs Hatim
'Hatim wrote the letter'

The resumptive pronoun here has the same reference as the object noun phrase. The subject and/or the object may be omitted if it has already been introduced in the discourse; the result is an inflected verb which by itself constitutes a meaningful and grammatical sentence.

Standard Arabic
(6a) *kataba-hā*
wrote.3ms-cl3fs
'He wrote it'

Moroccan Arabic
(6b) *kteb-ha*
wrote.3ms-cl3fs
'He wrote it'

The possibility of using subject and object clitics allows the omission of both the subject and the object lexical noun phrases.

Clitics are related to agreement in both Standard Arabic and Moroccan Arabic. One significant difference between the two varieties is the absence in Standard Arabic of agreement in number between the verb and its subject, as in (7a). Subject-verb agreement is required in subject-initial sentences, as in (7b).

Standard Arabic
(7a) *xaraja* *l-ʾawlād-u*
went.out.3ms the-children-Nom
(7b) *al-ʾawlād-u* *xaraj-ū*
the-children-Nom went.out.3mp
'The children went out'

Moroccan Arabic
(8a) *xerj-u* *l-ulad*
went.out.3mp the-children
(8b) *l-ulad* *xerj-u*
the-children went.out.3mp
'The children went out'

An analysis of the sentences in (7) and (8) reveals that in Standard Arabic no agreement in number is required between the subject and its verb when the latter is initial (7a). We assume that in (7b) *al-ʾawlādu* is in topic position, hence the apparent agreement on the verb marked by the morpheme *-ū*. In the case of Moroccan Arabic, verb-subject agreement is a morphosyntactic requirement because the agreement has to occur between the verb and the subject, irrespective of the word order. This implies that, although Standard Arabic and Moroccan Arabic permit the subject to be either initial or to follow the verb, agreement phenomena are dissimilar in both languages.

2. The complex sentence

A number of syntactic phenomena are associated with the complex sentence in Standard Arabic and Moroccan Arabic. In view of space limitations, only a few areas are discussed here; most of the main types of subordinate clause (relative clauses, complement clauses, and coordination) are also dealt with briefly (→ relative clause; → subordination; → sentence coordination).

2.1 *Relative clauses*

A relative clause consists of an embedded sentence used as modifier of a noun phrase. In Standard Arabic, → relative pronouns inflect for person, number, and gender (and the dual relative pronoun inflects for case as well): *allaḏī* (masc. sg.), *allatī* (fem. sg.); *allaḏāni/allaḏayni* (masc. du. nom./obl.), *allatāni/allatayni* (fem. du. nom./obl.), *allaḏīna* (masc. pl.), *allāti* ~ *allawāti* (fem. pl.), but in Moroccan Arabic, they do not: *lli* 'that, who', *aš* 'what', *mn* 'whom'. Both Standard Arabic *allaḏī* and Moroccan Arabic *lli* relative pronouns do not inflect for animacy, i.e., the same relative pronoun is used, irrespective of whether it refers to animate or inanimate heads, as in (9)–(10).

Standard Arabic
(9a) *ar-rajul-u* *llaḏī* *iltaqay-ta*
the-man-Nom Rel met-2ms
bi-hi
with-him
'the man whom you met'
(9b) *al-faras-u* *llaḏī* *imtaṭā-hu*
the-horse-Nom Rel rode.3ms-cl3ms
ʾab-ī
father-my
'the horse that my father rode'

Moroccan Arabic
(10a) *l-mra* *lli* *tlaqi-t*
the-woman Rel met-1s
'the woman whom I met'
(10b) *l-kelb* *lli* *mat*
the-dog Rel died.3ms
'the dog that died'

Moroccan Arabic *aš* 'what' and *mn* 'whom' are used solely to identify inanimate and animate antecedents respectively. So far as their functions are concerned, Table 1 shows that in Standard Arabic the relative pronoun *allaḏī* and its morphological variants, like Moroccan Arabic *lli*, can function as subject, direct object, and indirect object as well as genitive, as in (11a–f); Moroccan Arabic bound morphemes, however, can function only as indirect objects or prepositional objects since, as bound morphemes, they cannot stand by themselves, as in (12a–h).

Table 1. Functions of relative pronouns in Standard and Moroccan Arabic

relative pronouns	subject	direct object	indirect object	genitive
allaḏī	+	+	+	+
lli	+	+	+	+
aš	–	–	+	–
mn	–	–	+	–

Standard Arabic
(11a) *al-walad-u* *llaḏī* *jāʾa*
the-boy-Nom Rel came.3ms
'the boy who came'
(11b) *al-walad-u* *llaḏī* *raʾay-ta-hu*
the-boy-Nom Rel saw-2ms-cl3ms
'the boy whom you saw'
(11c) *al-walad-u* *llaḏī* *ʾaʿṭay-ta-hu*
the-boy-Nom Rel gave-2ms-cl3ms
l-kitāb-a
the-book-Acc
'the boy to whom you gave the book'
(11d) **al-walad-u* *llaḏī* *ʾaʿṭay-ta*
the-boy-Nom Rel gave-2ms
l-kitāb-a
the-book-Acc
(11e) *al-walad-u* *llaḏī* *sāfara*
the-boy-Nom Rel traveled.3ms
ʾabū-hu
father-his
'the boy whose father traveled'
(11f) **al-walad-u* *llaḏī* *sāfara*
the-boy-Nom Rel traveled.3ms
ʾabū
father

Moroccan Arabic
(12a) *l-mra* *lli* *matat* *sġira*
the-woman Rel died.3fs young.fs
'the woman who died young'
(12b) *l-mra* *lli* *šef-ti-(ha)*
the-woman Rel saw-2ms-cl3fs
'the woman whom you saw'
(12c) *l-mra* *lli* *tkllem-ti* *mʿa-ha*
the-woman Rel talked.2ms with-her
'the woman whom you talked to'
(12d) **l-mra* *lli* *tkllem-ti* *mʿa*
the-woman Rel talked.2ms with-her
(12e) *l-mra* *lli* *rajl-ha* *ḥmaq*
the-woman Rel husband-her fool
'the woman whose husband is a fool'
(12f) **l-mra* *lli* *rajl* *ḥmaq*
the-woman Rel husband fool
(12g) *l-mra* *mʿa* *mn* *šef-t-k*
the-woman with Rel saw-1s-cl2ms
'the woman with whom I saw you'
(12h) *l-kursi* *f-aš* *kun-ti* *gals*
the-chair in-Rel were-2ms sitting
'the chair on which you were sitting'

Note that *-mn* is used to refer back to the animate head noun phrase *l-mra* while *-aš* refers back to the inanimate head noun phrase *l-kursi*.

Standard Arabic and Moroccan Arabic also exhibit important similarities: anaphoric pronouns are optional when *allaḏī* and *lli* relativize direct objects (as in 11b and 12b), and obligatory when they relativize indirect objects (as in 11c and 12c); they are obligatorily deleted when they relativize subjects, as (11a) and (12a) show.

2.2 *Complement clauses*

Complement clauses in both Standard Arabic and Moroccan Arabic are introduced by complementizers (or clause-introducing particles). In Standard Arabic, there exist three main complementizers: *ʾanna* 'that', *ʾan* 'for', and *hal/ʾa* 'whether', the latter being interchangeable as they never vary meaningfully. In Moroccan Arabic, three complementizers are used: *blli* 'that', *baš* 'for', and *waš* 'whether'. Examples are given in (13)–(14).

Standard Arabic
(13a) *ʿalim-tu* *ʾanna* *l-walad-a*
knew-1s that the-boy-Acc
najaḥa
succeeded.3ms
'I knew that the boy had succeeded'
(13b) **ʿalim-tu* *ʾanna* *najaḥa*
knew-1s that succeeded.3ms
l-walad-a
the-boy-Acc
(13c) *ʾu-rīdu* *ʾan* *ta-štariy-a*
1s-want for you-buy-Subj
s-sayyārat-a
the-car-Acc
'I would like for you to buy the car'
(13d) **ʾu-rīdu* *ʾan* *as-sayyārata*
1s-want that the-car-Acc
ta-štariy-a
2ms-buy-Subj
(13e) *lā* *ʾa-drī* *hal/ʾa*
not 1s-know whether

tawaṣṣal-ta *bi-risālat-ī*
received-2ms with-letter-my
'I don't know whether you received my letter'

Moroccan Arabic

(14a) *ḍann* *aḥmed* *blli*
thought.3ms Ahmed that
weld-u *mriḍ*
son-his ill
'Ahmed thought that his son was ill'

(14b) *ḍann* *aḥmed* *blli*
thought.3ms Ahmed that
mriḍ *weld-u*
ill son-his
'Ahmed thought that his son was ill'

(14c) *bġi-t-u* *baš* *i-mši*
want-1s-cl3ms for 3ms-go
'I want him to go'

(14d) *ma-smaʿ-t-š* *waš* *aḥmed*
Neg-heard-1s-Neg whether Ahmed
ja
came.3ms
'I didn't hear whether Ahmed came'

(14e) *ma-smaʿ-t-š* *waš* *ja*
Neg-heard-1s-Neg whether came.3ms
aḥmed
Ahmed
'I didn't hear whether Ahmed came'

These examples show that Standard Arabic complement clauses are more syntactically constrained than Moroccan Arabic ones. First, the word order after *ʾanna* must be SV, but never VS (see 13b). Second, with *ʾan*, it is the opposite: only VS order is possible in the embedded clause. Third, by contrast, in Moroccan Arabic the word order is free in the sense that either SV or VS order is possible, regardless of whether the complement clause is introduced by *blli, baš*, or *waš*.

In addition, *ʾanna* in Standard Arabic assigns accusative case to the following noun phrase (as in 13a), whereas *ʾan* requires the subjunctive verb form in the embedded clause (as in 13c), and *hal/ʾa* is used to express the indicative mood.

In Moroccan Arabic, the complementizer *blli* introduces declarative complement clauses involving a perfective aspect, while *baš* is used to introduce clauses in the imperfective aspect, and *waš* serves to introduce indirect questions.

From a semantic point of view, it is plausible to state that first, not all verbs in Standard Arabic and Moroccan Arabic take complementation, and second, those that do take complementation behave differently and determine the type of complement clause they allow. In both languages, the verbs that can take complementation fall into three categories. The first category includes verbs followed by a perfective declarative complement; the second consists of verbs which take imperfective complement clauses, and the third category consists of verbs which allow interrogative complement clauses. Interestingly, almost all those verbs which do not take complement clauses generally share the semantic property of not allowing abstract object noun phrases, as in (15)–(16).

Standard Arabic

(15) **laʿiba* *l-walad-u* *ʾanna*
played.3ms the-boy-Nom that
s-samāʾ-a *zarqāʾu*
the-sky-Acc blue
'*The boy played that the sky was blue'

Moroccan Arabic

(16) **attat* *aḥmed* *blli*
furnished.3ms Ahmed that
xeṣṣu *i-rbaḥ*
needs 3ms-win
'*Ahmed furnished that he must win'

By implication, complement-taking verbs are verbs which allow abstract objects (cf. Ennaji 1985).

2.3 *Coordination*

Coordination is the process whereby constituents are linked on the same structural level, so that the elements which are coordinated represent the elements of a larger constituent (see Dik 1968; Hudson 1976). For instance, each of the compound structures in (17)–(18) is a sentence consisting of the compounding of two sentences (→ sentence coordination).

Standard Arabic

(17) *kataba* *mūsā* *risālat-an*
wrote.3ms Musa letter-Acc
wa-kataba *ʾaḥmad-u* *kitāb-an*
and-wrote.3ms Ahmed-Nom book-Acc
'Musa wrote a letter and Ahmed wrote a book'

Moroccan Arabic

(18) *šra* *ʾaḥmed* *dar*
bought.3ms Ahmed house
u-šra-t *faṭima* *tomobil*
and-bought-3fs Fatima car
'Ahmed bought a house and Fatima bought a car'

The two clauses involved in (17) and (18) above are independent in the sense that individually each one may constitute a sentence on its own. In both Standard Arabic and Moroccan Arabic, the → connectives *wa-* and *u-* 'and', respectively, usually occur overtly, unlike in Berber, where only a juxtaposition of sentences is sufficient to indicate coordinate structures (cf. Ennaji 1985; Sadiqi 1986).

Both *wa-* and *u-* 'and' may link an arbitrarily large number of clauses. Unlike English, where *and* may not appear in all the conjoined clauses but only in the last one, in Standard Arabic the connector *wa-* is used to link every conjoined clause, whereas in Moroccan Arabic *u-* may appear either to link every conjoined clause including the last one, or only to conjoin the last two clauses, as in (19)–(20).

Standard Arabic

(19) *kul* *wa-šrab*
eat.Imper. and-drink.Imper
wa-kun *fariḥ-an*
and-be.Imper merry-Acc
'Eat, drink, and be merry!'

Moroccan Arabic

(20) *faq* *aḥmed* *f* *ṣ-ṣbaḥ*
woke.3ms Ahmed in the-morning
(u-)ġsel *wajh-u*
and-washed.3ms face-his
(u-)fṭar *u-mša*
and-had.breakfast.3ms and-went.3ms
l *xdma*
to work
'Ahmed woke up, washed his face, had his breakfast, and went to work'

It is possible to conjoin sentences which are naturally associated in some way, and to conjoin sets of sentences with others. The connectives *wa-* and *u-* are semantically the most neutral coordinators, and their meaning tends to be dependent on factors like sequence of actions and events. The usual notion which they express is that of additionality (adjunction). Likewise, they are commonly used to mark a certain continuity of the narrative, usually in the middle of discourse, never at the beginning of a conversation. This is why they are very difficult to translate out of context.

Coordination takes place not only when two sentences are conjoined but also when there are disjoined sentences. The coordinators *ʾaw* in Standard Arabic and *wlla* in Moroccan Arabic are generally used as devices of disjunction. Unlike in Standard Arabic where the connector *ʾaw* must occur in every coordinated clause, in Moroccan Arabic the connector *wlla* may be omitted from all clauses except the last one; otherwise, two simple sentences are derived instead of one disjunction, as in (21)–(22).

Standard Arabic

(21) *rubbamā* *bāʿa* *zayd-un*
perhaps sold.3ms Zayd-Nom
sāʿat-a-hu *ʾaw* *bāʿa*
watch-Acc-his or sold.3ms
kitāb-a-hu *ʾaw* *bāʿa*
book-Acc-his or sold.3ms
ḥaqībat-a-hu
bag-Acc-his
'Perhaps Zayd sold his watch, or sold his book, or sold his bag'

Moroccan Arabic

(22) *yemkn* *li-k* *t-bqa* *hna*
possible to-you 2ms-stay here
t-mši *l* *mktaba*
2ms-go to library
wlla *t-mši* *l* *d-dar*
or 2ms-go to the-house
'You can stay here, go to the library, or go home'

In the Moroccan Arabic example in (22), the coordinator is deleted in the first clause, but because connectors occur before the final conjunct, we can easily understand that the sentences involved are disjoined since they carry alternative propositions and do not denote additionality (adjunction).

The third type of connective is *lākin* 'but' in Standard Arabic and *walakin* 'but' in Moroccan Arabic. The use of these two connectives depends on many factors, such as the semantic relationship between the coordinated sentences. The notion of contrast or semantic opposition is basically what allows the use of the connectives *lākin* and *walākin*. Examples are given in (23)–(24).

Standard Arabic
(23) *sayyārat-ī* *jadīdat-un* *lākin* *lā*
car-my new-Nom but not
ʾu-ḥibbu *lawn-a-hā*
1s-like color-Acc-its
'My car is new, but I don't like its color'

Moroccan Arabic
(24) *qra* *ḥasan* *l-ʿam* *kullu*
read.3ms Hasan the-year all
walakin *ma-njaḥ-š*
but Neg-succeeded.3ms-Neg
'Hasan studied the whole year, but he didn't succeed'

In both examples, a contrast is obviously conveyed by the two conjuncts of every coordinate sentence. The connectors *lākin* and *walakin* can only connect two sentences, unlike other types of coordinators. The reasons for this appear mainly to relate to the fact that *lākin* and *walakin* involve the opposition of two states of affairs, while the conjunctions *wa-* and *u-* 'and' and the disjunctions *ʾaw* and *wlla* 'or' can conjoin or disjoin any number of sentences.

3. CONCLUSION

Thus, Standard Arabic and Moroccan Arabic make use of different relative pronouns and complementizers for different functions; however, Moroccan Arabic has fewer relative forms.

Complement clauses occur only in postverbal position in both Arabic varieties, due to the fact that their basic word order is VSO. In Standard Arabic, complement clauses are coordinated in accordance with coordinate sentence requirements, explained above. In Moroccan Arabic, however, they require the use of an overt connector to link two or more clauses, specifically when the second clause has VSO order. When more than two clauses are coordinated, the connectives may be deleted in Moroccan Arabic from all conjuncts except the last one, but in Standard Arabic they remain compulsory.

BIBLIOGRAPHICAL REFERENCES

Dik, Simon C. 1968. *Coordination: Its implications for the theory of general linguistics.* Amsterdam: North-Holland.

Ennaji, Moha. 1985. *Contrastive syntax.* Wurzburg: Königshausen und Neumann.

——. 2002. "On the main differences between Standard Arabic and Moroccan Arabic complex sentences". *Aspects of the dialects of Arabic today*, ed. Abderrahim Youssi, Fouzia Benjelloun, Mohamed Dahbi, and Zakia Iraqi-Sinaceur, 347–353. Rabat: Amapatril.

Fassi Fehri, Abdelkader. 1976. "Relatives et adjectifs en arabe: Le problème de la détermination". *Lingua* 38.125–152.

——. 1978. "Relatives restrictives en arabe: Déplacement de WH ou éffacement par dessus une variable?". Ms., Université de Paris-7, Vincennes.

Hudson, Richard A. 1976. "Conjunction reduction, gapping, and right-node raising". *Language* 52.535–561.

Sadiqi, Fatima. 1986. *Studies in Berber syntax: The complex sentence.* Würzburg: Königshausen und Neumann.

——. 2002. "The major syntactic differences between Standard Arabic and Moroccan Arabic simple sentences". *Aspects of the dialects of Arabic today*, ed. Abderrahim Youssi, Fouzia Benjelloun, Mohamed Dahbi, and Zakia Iraqi-Sinaceur, 394–401. Rabat: Amapatril.

Woolford, Ellen 1978. "Relatives and other base-generated WH constructions". *Papers from the Fourteenth Regional Meeting of the Chicago Linguistic Society*, 482–489. Chicago: University of Chicago.

Youssi, Abderrahim. 1992. *Grammaire et lexique de l'arabe marocain moderne.* Casablanca: Wallada.

MOHA ENNAJI (Rutgers University)

Sentence Coordination

Sentence coordination refers to the process of creating compound sentences through combining two or more independent clauses. Coordination may be 'syndetic', which means that the conjuncts are combined using a conjunction such as *wa-* 'and' or *fa-* 'and so', or coordination may be 'asyndetic', in which case the sentences are combined without the use of any conjunctions (Cantarino 1974–1975:III, 7).

Syndetic coordination is more common than asyndetic coordination, and the conjunctions that conjoin sentences are also used for many other purposes, including conjoining phrases (→ conjunctions) and attaching subordinate structures (→ subordination). Subordination is distinct from coordination: in subordinate clauses, one clause modifies the other, while in sentence coordination, the sentence is made of two equal parts. The distinction is useful but can at times be extremely difficult to define (Ryding 2005). Many of the same conjunctions (e.g. *wa-* and *fa-*) are also used to begin utterances as discourse markers (→ connectives), thus connecting larger units of text.

1. HISTORY

Modern Standard Arabic prose differs from older written prose by the increased use of subordination in Modern Standard Arabic and by the relatively reduced use of a few particles in simple coordination, toward more complex sentences (Holes 1995). In Classical Arabic, a small set of conjunctions connected all sentences and served as 'punctuation', but with the introduction of Western → punctuation, Modern Standard Arabic has developed a system that uses both conjunctions and punctuation to connect ideas (Badawi a.o. 2004), with the end result being a sparser use of *wa-* and *fa-* in Modern Standard Arabic than in Classical Arabic (Kammensjö 2004).

The meaning of asyndetic coordination has also changed from Classical Arabic to Modern Standard. While it was used in Classical Arabic to stress immediacy or vividness, it is now used in less marked contexts, such as (1), in which the first clause gives the context for the second.

(1)	*naḥnu*	*l-bašaru*	*ʾaġbiyāʾu*
	we	the-humans	stupid.p
	na-ʿtaqidu	*kulla*	*šayʾin*
	1p-believe	all	thing

'We humans are stupid we believe everything' (Badawi a.o. 2004:540)

The ordering of the coordinated clauses has also become more flexible in Modern Standard Arabic. In Classical Arabic, clauses that described the circumstances of the agent or patient at the time of the event were strictly required to occur after the clause describing the event, but that strict word order requirement is now relaxed, due in part to influences from the dialects, which blur the distinctions between coordinated and subordinated conjunction (Rosenhouse 1978), and also due in part to language contact with English and French (Holes 1995).

2. TYPES OF COORDINATION

The two most frequent conjunctions used to coordinate sentences are clitics (→ enclisis), *wa-* and *fa-*. They are lexically unstressed and, in written language, are frequently attached to the word that follows them. Other conjunctions that coordinate sentences are lexical words, and they include *ṯumma* 'then', *lākin(na)* 'but', *bal* 'but rather', *ʾaw* 'or [inclusive]', *ʾam* 'or [exclusive]', and *ʾimmā* 'either'. Conjunctions can also be attached to other conjunctions or particles to form complex clitics such as *wa-lākin* 'but', and *wa-lā, wa-laysa* 'and not'.

Syndetic coordination captures a wide range of semantic relationships. The coordinated sentences may be additive, adversative, or disjunctive, one sentence may provide background information for the other, the sentences may depict cause-and-effect relationships, they may outline the order of events, or the second sentence may explain the first. The most common particles, *wa-* and *fa-*, are used to connect sentences within a wide range of semantic relationships, and the meaning is generally inferable from context rather than from the choice of particles (Cantarino 1974–1975:III, 11–34).

In contrast to the wide range of uses for syndetic coordination, asyndetic coordination is possible only for a restricted set of semantic relationships. For example, asyndetic coordination is possible only when there is no temporal sequence between the two clauses, and it is becoming more common when describing people and identities. When two actions occur simultaneously, either syndetic or asyndetic linkage is possible (Beeston 1973; Holes 1995).

In terms of punctuation, sentences can be coordinated in a variety of ways. As shown in (1), the two independent clauses may be asyndetically coordinated, with no intervening punctuation. They may also be syndetically coordinated with no punctuation (2). If punctuation is used, the clauses may be separated by commas, by dots '...' (3), or by a combination of dots and other punctuation marks, e.g. '...?' (Badawi a.o. 2004).

(2) *inṣaraf-a r-rajul wa-rakib-a*
left-3ms the-man and-mounted-3ms
ṣadīqu-hu d-darrāja
friend-his the-bicycle
'The man left, and his friend mounted the bicycle' (Modern Standard Arabic; Holes 1995:269)

(3) *laʿiba-tā maʿan… mariḥa-tā*
played-3fd together had.fun-3fd
maʿan… ḏākara-tā maʿan…
together studied-3fd together
'They played together, had fun together, studied together…' (Modern Standard Arabic; Badawi a.o. 2004:540)

The coordinated sentences may be nominal, verbal, or imperatives, but they follow the rule of the Law of Coordination of Likes (Pullum and Zwicky 1986), meaning that nominal sentences are coordinated with other nominal sentences (4), verbal sentences are coordinated with other verbal sentences (5), and imperatives are coordinated with imperatives (6).

(4) *aḍ-ḍawʾu nāṣiʿ-un*
the-light clear-Nom
wa-l-jawwu ḥārr-un
and-the-weather hot-Nom
'The light is clear, and the weather is hot' (Modern Standard Arabic; Badawi a.o. 2004:542)

(5) *šakar-tu-hu wa-nṣarafa*
thanked-1s-him and-left.3ms
'I thanked him, and he left' (Modern Standard Arabic; Cantarino 1974–1975: III, 14)

(6) *iġsilī-hi wa-lā tasʾalī*
wash.2fs.Imper-it and-Neg ask.2fs.Juss
'[Just] wash it and don't ask' (Modern Standard Arabic; Badawi a.o. 2004:543)

3. Coordination and ellipsis

The linguistic difficulty in analyzing coordinated sentences comes from the possibility of ellipsis, which allows certain elements to be elided in the coordinated clause (→ ellipsis) if they are redundant or recoverable from context (Merchant 2001). The most frequent elements to undergo elision are prepositions and complementizers. But even though these elements are not pronounced, their inflection is still preserved in the coordinated elements, as in (7), where both nouns have accusative marking because of the complementizer *ʾanna*, which is elided before *karāmata-hu* 'his honor'.

(7) *bi-ʾanna ḥuqūq-a-hu muṣānatun*
with-Compl rights-Acc-his protected
wa-karāmat-a-hu maḥfūḏatun
and-honor-Acc-his preserved
'…that his rights are protected and [that] his honor is preserved' (Modern Standard Arabic; Badawi a.o. 2004:546)

The important question is how to determine what was coordinated and what was elided. In sentence (8), the surface structure indicates coordination of prepositional phrases, *fī ṣ-ṣayf* 'in the summer' and *fī l-xarīf* 'in the fall'. An alternate analysis, as proposed by Cantarino, is that this sentence is an example of clausal coordination, with elision of all redundant material, leaving only the prepositional phrase to be pronounced in the second clause.

(8) *fī ḍayʿati-nā lā yuʿayyidūna*
in village-our Neg celebrate.feast.3mp
fī ṣ-ṣayf wa-lā fī
in the-summer and-Neg in
l-xarīf
the-autumn
'In our village, people do not celebrate a feast either in summer or in autumn' (Modern Standard Arabic; Cantarino 1974–1975:III, 57)

The possibility of ellipsis becomes important in analyzing first conjunct agreement, one of the best-known patterns of subject/verb agreement in Arabic (→ agreement; → *tanāzuʿ*). First conjunct agreement arises with postverbal conjoined noun phrases, such as in (9), and has been attested in Standard Arabic and in several dialects, including Moroccan Arabic and Lebanese Arabic (Aoun a.o. 1994; Harbert and Bahloul 2002).

(9a) *xaraj-at al-bintu wa-l-waladu*
left-3fs the-girl.f and-the-boy.m
'The girl and the boy left'

(9b) *xaraj-a* *l-waladu* *wa-l-bintu*
left-3ms the-boy.m and-the-girl.f
'The boy and the girl left'

(9c) *xaraj-ti* *ʾanti* *wa-huwa*
left-2fs you.2fs and-he.3ms
'You and he left' (Modern Standard Arabic; Harbert and Bahloul 2002:51)

Sentences (9a–c) show that the verb *xaraja* agrees with the closest noun, the first conjunct, in person and gender. (Number agreement only occurs with pronominal postverbal subjects in Modern Standard Arabic.) If the verb had agreed with both conjuncts, masculine singular agreement would be expected in both (9a) and (9c). Instead, verbs that agree with postverbal conjoined subjects may enter into an agreement relationship with just the first conjunct, rather than with the resolved feature values from the whole conjunction.

One explanation (Aoun a.o. 1994) is based on the possibility of ellipsis in coordinated sentences. This approach argues that first conjunct agreement is due to clausal coordination with gapping, as in (10), which shows sentence (9c) as it would be if the verb in the second conjunct had not undergone elision.

(10) *xaraj-ti* *ʾanti* *wa-*[*xaraj-a*] *huwa*
left-2fs you.2fs and-[left-3ms] he.3ms
'You left, and he left' (Modern Standard Arabic; Harbert and Bahloul 2002:51)

One particularly convincing piece of evidence against the gapping analysis of first conjunct agreement (Aoun a.o. 1994) comes from the presence of 'mixed agreement', which has recently been demonstrated in Lebanese Arabic through sentence elicitation tasks and confirmed via grammaticality judgments (Lorimor 2007). The pattern of mixed agreement arises when a sentence contains two agreement targets that express conflicting feature values, such as the verb and adjective in (11).

(11) *kān-it* *el-baṭṭa*
was-3fs the-duck.f
w-el-wazzi *xuḍer*
and-the-swan.m green.p
'The duck and the swan were green' (Lebanese Arabic)

The sentences were elicited by showing participants pictures of objects that briefly changed in color, from black to either red, blue, or green. Participants were instructed to ask questions or to make statements using a particular word order, as modeled by the experimenter. In (11), the verb *kānit* agrees with the first conjunct, while the adjective *xuḍer* shows plural agreement, which is only possible if it agrees with the resolved features of the conjoined noun phrase. The gapping explanation would predict the opposite pattern: a singular adjective in sentence-final position, since each clause would only contain one noun with which to agree. Mixed agreement is therefore solid evidence that first conjunct agreement cannot be attributed to gapping. Similar analyses of mixed agreement in other languages, including Finnish, Greek, Slovenian, and Welsh, have also rejected the notion that first conjunct agreement is due to sentence coordination with ellipsis (see Badecker 2007 and Corbett 2006 for further discussion).

Several alternate explanations for first conjunct agreement have also been proposed, the most promising of which link first conjunct agreement to other patterns of partial agreement that occur with only postverbal subjects (Harbert and Bahloul 2002; Lorimor 2007) by relying on a distinction between specifier/head agreement and agreement under government (Benmamoun 1992).

4. Sentence coordination in the dialects

Sentence coordination in the dialects is largely the same as in Modern Standard Arabic, as *wa-*, *fa-*, and *ʾaw* are attested across nearly all registers and dialectal variants of Arabic. Register does play a role, however, in the prevalence of coordination and subordination, as informal and uneducated speech contain more instances of coordination, while written Modern Standard Arabic contains more instances of subordination (Holes 1995). Another source of variation between Modern Standard Arabic and the dialects is a general relaxation of the rules that restrict the ordering of coordinated sentences. For example, in Cairene, Damascene, Baghdadi, and Palestinian Arabic, there is no longer a restriction requiring that a circumstantial clause follow the action it is describing (Rosenhouse 1978).

BIBLIOGRAPHICAL REFERENCES

Aoun, Joseph, Elabbas Benmamoun, and Dominique Sportiche. 1994. "Agreement, conjunction and word order in some varieties of Arabic". *Linguistic Inquiry* 25.195–220.

Badawi, Elsaid, Michael G. Carter, and Adrian Gully. 2004. *Modern Written Arabic: A comprehensive grammar*. London: Routledge.

Badecker, William. 2007. "A feature principle for partial agreement". *Lingua* 117.1541–1565.

Beeston, Alfred F.L. 1973. "Coordination in Literary Arabic". *Mélanges de l'Université Saint-Joseph* 48.167–174.

Benmamoun, Elabbas. 1992. "Structural conditions on agreement". *Proceedings of the North East Linguistic Society* 22.17–32.

Cantarino, Vicente. 1974–1975. *Syntax of Modern Arabic prose*. 3 vols. Bloomington: Indiana University Press.

Corbett, Greville G. 2006. *Agreement*. Cambridge: Cambridge University Press.

Harbert, Wayne and Maher Bahloul. 2002. "Postverbal subjects and the theory of agreement". *Themes in Arabic and Hebrew syntax*, ed. Jamal Ouhalla and Ur Shlonsky, 45–70. Dordrecht: Kluwer Academic Publishers.

Holes, Clive. 1995. *Modern Arabic: Structures, functions, and varieties*. London and New York: Longman.

Kammensjö, Heléne. 2004. *Discourse connectives in Arabic lecturing monologue*. Ph.D. diss., Göteborg University.

Lorimor, Heidi. 2007. *Conjunctions and grammatical agreement*. Ph.D. diss., University of Illinois at Urbana-Champaign.

Merchant, Jason. 2001. *The syntax of silence: Sluicing, islands, and the theory of ellipsis*. Oxford: Oxford University Press.

Pullum, Geoffrey K. and Arnold M. Zwicky. 1986. "Phonological resolution of syntactic feature conflict". *Language* 62.751–773.

Rosenhouse, Judith. 1978. "Circumstantial clauses in some Arabic dialects". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 128.226–237.

Ryding, Karin. 2005. *A reference grammar of Modern Standard Arabic*. Cambridge: Cambridge University Press.

HEIDI LORIMOR (Bucknell University)

Serial Verbs

The term 'serial verbs' is used in the literature to indicate a verbal syntagm consisting of two (or more) finite verbs without a formal coordinating marker but with the same argument structure, one of which is semantically demoted, often grammaticalized, and lexically restricted (Sebba 1987:39). Constructions with serial verbs are familiar from a wide group of languages, ranging from Mandarin Chinese to West African languages like Yoruba. Examples are given in (1) – (3).

(1) (Mandarin Chinese)
ta lai shang ban
he come go.up shift
'He comes to work' (Dai 1990:327)

(2) (Sranan)
a waka go na wowoyo
he walk go Loc market
'He walks to the market' (Sebba 1987:46)

(3) (Sranan)
kofi tjari den fisi kon gi mi
Kofi bring Det fish come give me
'Kofi brought me the fish' (Holm 1988:184)

In (1) and (2), the second verb (*shang, go*) is desemanticized and serves only to indicate the locational direction of the event; in (3), the second verb (*kon*) is directional and the third verb (*gi*) indicates a dative.

Both in Classical Arabic and in modern Arabic dialects, strings of finite verbs without coordinating markers occur, but whether these are to be classified as serial constructions is a controversial issue. In the dialects, asyndetic constructions are common after modal expressions, but in these constructions the asyndetically connected verb is part of a subordinate clause, as in (4) and (5).

(4) (Egyptian Arabic)
lāzim tu-dxul
necessary 2ms-enter
'You must enter'

(5) (Syrian Arabic)
bədd-i ʾə-ktob
wish-my 1s-write
'I want to write'

In these examples, even if the modal expression is counted as a verb, the second verb is subordinated and does not need to have the same tense as the main clause. According to most definitions of serial verbs, this cannot count as an instance of serialization because there is no desemanticization and both clauses can be negated separately.

Yet, same-tense finite verbs occur in asyndetic constructions in Arabic, too. Potentially, at least, these might be counted as an instance of serialization (Hussein 1990). The most extensive analysis so far of asyndetic same-tense finite verbs is Woidich (2002). He concludes that in Egyptian Arabic such constructions have nothing in common with the serial verbs in other languages, hence he prefers to call them instead 'verbal phrases with asyndetic perfect' ("Verbalphrasen mit asyndetischem Perfekt"). Woidich distinguishes the following types in Egyptian Arabic (2002:125–127).

(i) Verbs expressing the → Aktionsart ('to do again', 'to do quickly', etc.), as in (6)

(6) (Egyptian Arabic)
rigiʿ hirib tāni
3ms.returned 3ms.fled second.time
'He fled a second time' (Woidich 2002:128)

Such combinations are well known from Classical Arabic, but there they are always realized with a subordinated verb in the imperfect tense, as in (7).

(7) (Classical Arabic)
ʾin ʿud-tu ʾa-smaʿu
if returned–1s 1s-hear
'if I hear again' (Reckendorf 1921:296)

(ii) Motion or posture verbs (e.g. *ʾām* 'to get up', *rāḥ* 'to go away'), as in (8)

(8) Egyptian Arabic (Baris, Kharga)
lamma tammi-t sana w
when completed–3fs year and
nuṣṣ arāḥi-t ḥibili-t
half went–3fs became.pregnant–3fs
'When she had spent (there) a year and a half, she became pregnant' (Woidich (2002:152)

In this group, the verb whose grammaticalization has progressed the most is *qām/ʾām*; it may even develop into a noninflected particle. The function of this particle is, according to Woidich (2002:141), to indicate that an activity or event takes place that is connected with something that has been mentioned before (see also Woidich 1995:265–266, where the loss of inflection is connected with the substandard language of the lower classes), as in (9).

(9) (Egyptian Arabic)
ʾām inta ʿamal-t^{i} ʾē
then you did–2ms what
'Then, what did you do?' (Woidich 2002:148)

In other dialects, too, verbs of motion and posture are frequently used in similar constructions; in some of them, the use of *qām/ʾām* as a noninflected particle has been documented as well. In his grammar of Syrian Arabic, Grotzfeld (1965:89–90) states that this use of *ʾām* signals the occurrence of a new, nondurative event in the past ("ein neu eintretendes, nicht andauerndes Geschehen in der Vergangenheit"), as in (10).

(10) Syrian Arabic
ʾām ʾəžət-na sayyāra
then came.to3fs-us car
'Then, a car came toward us' (Grotzfeld 1965:89)

According to Firanescu (2003), this use of *qām/ʾām* as an "inchoatif événementiel" contrasts with its use as an auxiliary in which it preserves its full verbal value; accordingly, the loss of agreement with the subject and development into a noninflected particle is connected with the former, rather than the latter.

Such lexical combinations already occur in Middle Arabic texts, as in (11) and (12).

(11) (Middle Arabic)
qām-ū taqātal-ū
got.up–3mp fought.eachother–3mp
'They began to fight with each other' (Brockelmann 1965:181, n. b)

(12) (Middle Arabic)
qum uxṭub-hā
get.up.Imper court.Imper-her
'Start courting her!' (Brockelmann 1965:181, n. b)

(iii) Lexicalized phrases like *daxal nām* 'to go to sleep', *baʿat gāb* 'to fetch', as in (13)

(13) (Egyptian Arabic)
baˁat *gāb*
3ms.sent 3ms.brought
ˁilbit-ēn *sagāyir*
packet-dual cigarettes
'He sent for two packets of cigarettes' (Woidich 2002:177)

The reason why Woidich does not analyze any of these constructions as serial verbs is that only the first verb in the chain is lexically restricted, i.e., only a few verbs may be used as first verb, whereas the choice for the second verb is more or less free. He also points out that the subject of both verbs may come between the first and the second verb, and that negations may have scope over the second verb alone. In his view, this demonstrates that the two verbs do not constitute a semantic unit and that they continue to indicate two events. The first point is demonstrated by (14).

(14) (Egyptian Arabic)
ʾām *is-sīx* *gih*
got.up3ms the-skewer came.3sm
fi *ˁēn-u* *ʾalaˁha*
in eye-his took.out3sm-it
'Then the skewer hit him in the eye and took it out' (Woidich 2002:143)

Here, the subject *issīx* comes between the two verbs, which demonstrates that the first verb has its own arguments and complements. A possible distinction that might be helpful here is that between core, nuclear, and clausal juncture (Foley and Olson 1985). Nuclear juncture takes place at the level of the bare verb, core juncture at the level of the verb with its tense/aspect markers, and clausal juncture at the level of the verb with its markers and arguments. Asyndetic constructions in Arabic normally operate at the core level, but those constructions in which the subject intervenes between the two verbs might be analyzed as instances of clausal juncture. With respect to negation, the situation is not entirely clear. Hussein (1990:346–347) rejects independent negation of the second verb, at least in strings containing an imperative.

The most serious objection against a serializing analysis is that in most 'classic' serial constructions, it is the second verb that is lexically restricted, as in (1) – (3). Yet, some researchers (e.g. Payne 1997:311–312) also accept verbs in second position as serial verbs, in particular when a verb of motion is involved. Payne (1997:307) calls such constructions marginal cases of serialization, and quotes from English the example in (15).

(15) *run go get me a newspaper*

In his approach – which is not followed by everyone (for a discussion, see Pullum 1990) – such constructions may be called 'serial verbs', although they are less prototypical than the examples quoted above in (1) – (3). In Arabic, it is almost exclusively the first verb that is lexically restricted, as in the three types mentioned by Woidich.

According to Fischer (2002), the asyndetic construction in Arabic developed from two coordinated verbs through deletion of the coordinating marker, e.g. *rigiˁ katab* 'he wrote again' (lit. 'he returned he wrote') developing from *rigiˁ wi katab*. Woidich (2002:171–176) assumes, however, that the use of asyndetic finite verbs in the same tense developed as the result of grammaticalization from subordinate constructions: along with *xallētu yimḍi* 'I made him sign', *xallētu maḍa* came to be used when the event of signing had actually taken place. This provided the model for other cases of asyndetic verbs, *rigiˁ yiktib* being replaced by *rigiˁ katab* when the event was completed. Subsequently, the first verb was subjected to → grammaticalization and → semantic bleaching, losing its own lexical content and sometimes being demoted to a noninflected particle.

Apart from the three types with premodifying verbs, Woidich also mentions two constructions in Egyptian Arabic in which the second verb is lexically restricted (called by him 'pseudocomplementation', after Seuren 1991). The first is the use of *rāḥ* after a number of verbs of motion as a directional marker, as in (16)

(16) (Egyptian Arabic)
miši *rāḥ* *fēn*
3ms.walked 3ms.went.away where
'Where did he go?' (Woidich 2002: 181)

In Egyptian Arabic, the use of *rāḥ* (and other verbs, like *waddā* in standard Egyptian and *jāb* 'to bring', *žā* 'to come' in rural dialects) is obligatory, because the argument structure of verbs like *miši* does not allow for the indication of a goal. Such verbs are also found in other dialects, for instance in Chad Arabic (Roth 1979:58) and in → Uzbekistan Arabic, as in (17),

(17) (Uzbekistan Arabic)

ʿal	*čūl*	*ṭalaʿ*
to	steppe	3ms.went.up
ġadāk		
3ms.went.away		

‘He went into the steppe’ (Fischer 1961:258)

where *ġadāk* serves only to indicate motion away from the speaker (for motion toward the speaker, *jāk* is used). Unlike Egyptian Arabic, however, there is no syntactic motivation in Uzbekistan Arabic for the use of a motion verb.

The second construction is the use of *qāla* (*gāl, ʾāl*) as a complementizer, in order to provide a syntactic slot for an object, as in (18)

(18) (Egyptian Arabic)

huwwa	*raddi*	*ʿalē-k*
he	answered	to- you
ʾal-l-ak		*ʾē*
3ms.said-to-you		what

‘What did he answer you?’ (Woidich 2002:183)

Although Woidich does not accept these last two types as serial constructions, either, it is difficult to escape the impression that they are indeed related to constructions in other languages that are usually analyzed as serializations (Versteegh 2006). With respect to the movement verbs, the Sranan example in (2) has the same obligatory use of *go* (without it, the sentence would be interpreted in a locative sense: ‘he was walking around at the market’).

With respect to *qāla* as a complementizer in Egyptian Arabic, Woidich (2007:692–696) remarks that unlike serial verbs in creole languages, it may be replaced with a coordinating construction with *wi-* ‘and’, and it is usually marked for tense and aspect. Besides, the verb *ʾāl* may have its own overt nominal argument. In his view, the verb *ʾāl* in this construction should therefore be analyzed differently from *se* (probably < English *say*) in many French and English creoles as a complementizer after verbs of saying and even after verbs of cognition (see Holm 1988:185–188), as in (19).

(19) (Krio Creole English)

a	*no*	*se*	*yu*	*bizi*
I	know	Comp	you	busy

‘I know that you are busy’ (Holm 1988:185)

In the function of complementizer, *qāla* occurs frequently in pidginized and creolized forms of Arabic, as in (20) from → Ki-Nubi (cf. also Miller 2001, on the grammaticalization of *gale* in → Juba Arabic).

(20) (Ki-Nubi)

a'jol	*'de*	*'kelem*	*'gal*	*'ai*
person	DET	spoke	said	yes

‘The person said (that) yes’ (Wellens 2005:282)

Grammaticalization has progressed further here than in Egyptian Arabic because the complementizer is used not only after verbs of ‘saying’ but also after verbs of cognition, such as ‘to think’.

The crucial issue is not whether any of the constructions mentioned here actually involve serialization, but the fact that in the emergence of the modern dialects a structurally new type of construction has developed, in which full verbs have become grammaticalized and semantically demoted. In an inflecting language like Arabic, it is of course harder for verbs to lose inflection than in other languages with less morphology. As a result, asyndetic constructions of finite verbs in Arabic always take place at the level of the verbal core or at the clausal level. They could, therefore, be analyzed as less prototypical examples of serial verbs than in other languages (Versteegh 2003–2005).

Historically, the most important issue is why these processes of loss of inflection and semantic status take place, and whether they are the result of a single process that took place at a certain moment in time and left its traces in later varieties of the language, or of a synchronic ongoing process. Woidich (2002:172) believes that the process of grammaticalization is a gradual one, resulting in a linguistic layering of increasingly atrophied verbal elements (his term is “Schichtung”, what Firanescu [2003:490] calls a “principe de stratification”). In this view, asyndetic constructions are regarded as the normal outcome of an ongoing process rather than a sudden transition in a process of language shift.

Bibliographical references

Brockelmann, Carl. 1965. *Arabische Grammatik, Paradigmen, Literatur, Übungsstücke und Glossar.* 16th ed. Ed. Manfred Fleischhammer. Leipzig: VEB Verlag Enzyklopädie.

Dai, John Xiang-ling. 1990. "Syntactic constructions in serial verb expressions in Chinese". Joseph and Zwicky (1990:316–339).

Firanescu, Daniel Rodica. 2003. "Le modalisateur aspectuel-temporel *qām* dans le parler syrien". *Proceedings of the 5th Conference of the Association Internationale de Dialectologie Arabe, Cádiz, September 2002*, ed. Ignacio Ferrando and Juan José Sánchez Sandoval, 481–492. Cadiz: University of Cadiz.

Fischer, Wolfdietrich. 1961. "Die Sprache der arabischen Sprachinsel in Uzbekistan". *Der Islam* 36.232–263.

——. 2002. "Unterordnende und nebenordnende Verbalkomposita in den neuarabischen Dialekten und im Schriftarabischen". *Sprich doch mit deinen Knechten aramäisch, wir verstehen es!*, ed. Werner Arnold and Hartmut Bobzin, 147–164. Wiesbaden: O. Harrassowitz.

Foley, William A. and Mike Olson. 1985. "Clausehood and verb serialization". *Grammar inside and outside the clause: Some approaches to theory from the field*, ed. Joanna Nichols and Anthony C. Woodbury, 17–60. Cambridge: Cambridge University Press.

Grotzfeld, Heinz. 1965. *Syrisch-Arabische Grammatik.* Wiesbaden: O. Harrassowitz.

Holm, John. 1988. *Pidgins and creoles.* I. *Theory and structure.* Cambridge: Cambridge University Press.

Hussein, Lutfi. 1990. "Serial verbs in Colloquial Arabic". Joseph and Zwicky (1990:340–354).

Joseph, Brian D. and Arnold M. Zwicky (eds.). 1990. *When verbs collide: Papers from the 1990 Ohio State Mini-Conference on Serial Verbs.* Columbus: Ohio State University, Department of Linguistics.

Miller, Catherine. 2001. "Grammaticalisation du verbe *gale* 'dire' et subordination en Juba-arabic". *Leçons d'Afrique: Filiations, ruptures et reconstitution de langues: Un hommage à Gabriel Manessy*, ed. Robert Nicolai, 455–482. Louvain and Paris: Peeters.

Payne, Thomas E. 1997. *Describing morpho-syntax: A guide for field linguistics.* Cambridge: Cambridge University Press.

Pullum, Geoffrey. 1990. "Constraints on intransitive quasi-serial verb constructions in modern colloquial English". Joseph and Zwicky (1990:218–239).

Reckendorf, Hermann. 1921. *Arabische Syntax.* Heidelberg: C. Winter.

Roth, Arlette. 1979. *Esquisse grammaticale du parler arabe d'Abbéché (Tchad).* Paris: Geuthner.

Sebba, Mark. 1987. *The syntax of serial verbs.* Amsterdam and Philadelphia: J. Benjamins.

Seuren, Pieter. 1990. "Serial verb constructions". Joseph and Zwicky (1990:14–33).

——. 1991. "The definition of serial verbs". *Development and structure of creole languages: Essays in honor of Derek Bickerton*, ed. Francis Byrne and Thom Huebner, 193–205. Amsterdam and Philadelphia: J. Benjamins.

Versteegh, Kees. 2003–2005. "Some remarks on verbal serialization in Arabic dialects". *Cahiers de Linguistique d'INALCO* 5.49–69.

Wellens, Ineke. 2005. *The Nubi language of Uganda: An Arabic creole in Africa.* Leiden: E.J. Brill.

Woidich, Manfred. 1995. "Some cases of grammaticalization in Egyptian Arabic". *Proceedings of the 2nd International Conference of l'Association Internationale pour la Dialectologie Arabe held at Trinity Hall in the University of Cambridge, 10–14 September 1995*, ed. Joe Cremona, Clive Holes, and Geoffrey Khan, 259–268. Cambridge: University of Cambridge.

——. 2002. "Verbalphrasen mit asyndetischem Perfekt im Ägyptisch-Arabischen". *Estudios de Dialectología Norteafricana y Andalusí* 6.121–192.

——. 2007. "Everything you always wanted to know about *fāl, yifūl* in Egyptian Arabic". *Approaches to Arabic linguistics presented to Kees Versteegh on the occasion of his sixtieth birthday*, ed. Everhard Ditters and Harald Motzki, 675–700. Leiden: E.J. Brill.

Kees Versteegh (University of Nijmegen)

Shift → Language Shift: Amazigh

Šiʿr

1. Introduction

Studying the language of Arabic poetry (*šiʿr*) poses a problem that may not be common in the study of language in poetry in other cultures. The problem arises from the fact that the Arabic language has been bequeathed to us largely in the form in which it was used in that poetry. In other words, the language of poetry, supplemented by the properties of the language of the Qurʾānic text, generated our concept of a *norm* in the use of Arabic. Thus, if we try to examine the language of poetry in terms of some modern notions of poeticality, we find ourselves in a vicious circle. Popular notions such as 'ungrammaticalness', 'deviation', a 'special poetic language', etc. break down as tools of analysis if we try to see how Arabic poetry can be said to be poetry on the grounds that the language in it 'deviates' from the norm that is its very language, or that it is ungrammatical in relation to a grammatically constructed language that is the language of itself. What we are left with, if we use such notions as those glorified by critics like Roman Jakobson (1987) and Michael Riffaterre (1978), from two distinct ends, will be some odd, individual, localized instances of deviation and ungrammaticalness, which are too few to define poeticality in terms of their properties. Most such instances were studied and classified by ancient Arabic scholars – who studied language in *manḏūm* ‘poetry’ and *manṯūr* ‘prose’ in the same vein – in a level-headed fashion under the heading of *aḍ-ḍarūra aš-šiʿriyya* (usage necessitated by poetic needs, perhaps a more accurate concept and term than the English → ‘poetic license’). The largest collection of lines of poetry in which one or another form of ‘ungrammaticalness’ or ‘deviation’ can be said to occur is perhaps the one made by ʾAbū ʿAlī al-Fārisī (d. 377/987) in his *Kitāb aš-šiʿr* ‘Book of poetry’. According to its editor, this book contains about 815 lines which al-Fārisī characterized as being *muškil al-ʾiʿrāb* ‘with problematic declension’. In relation to the body of Arabic poetry produced up to the time of al-Fārisī, this number is insignificant. Furthermore, being *muškil al-ʾiʿrāb* does not necessarily mean that the line in view is ungrammatical. It often means that it has a certain degree of ambiguity akin to one or other of the cases described by William Empson in *Seven types of ambiguity*. Al-Fārisī endeavors to show how the *ʾiʿrāb* of a line he is considering can be formulated in a fashion that will make the grammaticalness explicit in accordance with familiar norms, and rules, in the Arabic language. His effort is almost always directed toward removing ambiguity on the grammatical level in relation to the semantic level. However, many of the cases he considers show that the ungrammaticalness can only be removed by assuming an act of elision (sometimes of a drastic nature, such as deleting a verb), or the presence of an extraneous element of language in a line of poetry. To the modern reader, such cases may appear to represent an excessive degree of deviating from the norms. No modern poet will today have the courage to use language as encountered in some of the lines considered by al-Fārisī. It is not wholly incorrect to say that, particularly for pre-Islamic poets, almost anything seems to have been doable with language on the level of grammar and syntax, and, sometimes, even on that of the phonetic composition of signs. And that was not always dictated by a *ḍarūra šiʿriyya*. The poetic *ḍarūra* ‘necessity’ appears in some cases to have given the poet the freedom to shape the dough of language in his own image without any sense of inhibition. Yet, the cases of real *ḍarūra* encountered in al-Fārisī’s book are very limited and do not represent major acts of deviation. The dictating *ḍarūra* is often imposed by the demands of rhyming and rhythm, but not always so. Nevertheless, it is fortunate that the number of excessive cases of deviation we encounter in records like al-Fārisī’s and Ibn ʿUṣfūr’s (d. 670/1271) is very small, that they do not occur too many times in actual poems but often appear as isolated single lines, and that Arab grammarians did not take them as a basis for the construction of the grammar of Arabic; we would have ended up with a language more like English. In any case, a cursory comparative glance enables one to say that grammar and syntax, as well as the structure of individual words in Arabic, as manifested in poetry, have changed very little over at least sixteen hundred years, in comparison with those of English between Shakespeare and the present moment, a period covering about four hundred years.

However, there are many studies, the best of which is by Ibn Jinnī (d. 392/1002), which acknowledge that in some lines of poetry there is a discordance between the semantic constituent and the grammatical one, which leads to lack of expressiveness. The poetry of al-Mutanabbī (d. 354/965), a contemporary of Ibn Jinnī, was particularly investigated by his detractors for such discordance.

Two further observations need to be made: first, a very significant proportion of the lines recorded by al-Fārisī come from the poetry of a handful of poets who were still close to the desert, for instance Ḏū r-Rumma (d. 117/735), al-Farazdaq (d. 110/728), Ru’ba ibn al-ʿAjjāj (d. 145/762), al-’Aswad ibn Yaʿfur (6th century C.E.), and ʾAbū Ḏu’ayb al-Huḏalī (6th century C.E.), and who were known for the ‘difficulty’ of their language. The major poets of the *Muʿallaqāt* and the poets who are most prominent in the main anthologies of Arabic poetry represent a small minority in al-Fārisī’s collection. The pre-Islamic poet best known for revising his poetry, Zuhayr ibn ʾAbī Sulma (6th century), hardly figures in this collection (only four lines of his appear). Second, the vast majority of the lines quoted belong to the pre-Islamic and early Islamic eras; an insignificant number of lines are by poets of the Abbasid era, the age of writing. It is very likely that oral narration is responsible for many, if not most, of the more excessive cases of deviation within this collection. Further research, however, is needed to verify both of the points made here.

Orality was also responsible for some other features in poetry, at least up to the end of the 2nd century A.H. Oral composition generated a relatively significant ratio of formulaic expressions, varying from a short phrase (*li-man ṭalalun*) to a whole line (as in two lines by the 6th century C.E. poets Imruʾ al-Qays and Ṭarafa, for instance). Yet, the number and nature of formulas (added to the nature of Arabic poetry itself) are not sufficient to legitimize an approach to Arabic poetry based on theories of → orality developed by scholars like Parry and Lord (→ poetic koine). Oral narration, from another perspective, is perhaps what gives many words their semantic density, complexity, and at times oppositional and contradictory 'senses' or meanings. This at times reaches a point where a given line becomes difficult to explain or interpret in terms relevant to the context in which it appears or to the overall context of the culture itself. New approaches to poetry need to be adopted in order to resolve such possibilities of contradiction or discordance. The early commentators on pre-Islamic poetry in particular find themselves at times at a loss when handling such lines. Their interpretations of single words as well as whole lines sometimes sound artificially constructed and often diminish the scope of richness of the poetic material.

In addition to oral composition and oral transmission, it is likely that many of the 'deviations' we encounter in pre-Islamic and early Islamic verse are due to dialectal variations. This, however, should not be overemphasized. The corpus of Arabic poetry as we have inherited it does not, perhaps surprisingly, reveal the existence of a widespread dialectal material or the existence of fully fledged dialects in general in this age, although evidence of dialectal usages exists outside poetry (→ pre-Islamic Arabic). It is feasible that the Arabs used for poetry in the past, just as they do today, a unifying literary language that did not partake much of the spoken dialects of the area (→ poetic koine). The language of the Qurʾānic text shares this property with the language of poetry. There are claims that this unifying language represented the 'dialect' – if it can be legitimately called that – of the tribe of Qurayš.

2. Deviations from the grammatical norms

In a line of poetry attributed to the famous pre-Islamic poet Imruʾ al-Qays, a verb in the imperfect tense that is expected to be vocalized with *u* (*ḍamma*) at the end, as it is in the nominative case, is not; instead of *al-yawma ʾašrabu* 'today, I drink', the line has *al-yawma ʾašrab*, ending with a *sukūn*, a vocalization normally found only on the endings of verbs in the jussive case (*jazm*). If the line is taken on its own, as a complete and independent statement, the verb can be read *al-yawma išrab*, in the command form, which will render it fully 'grammatical'. In context, however, it is clear that the speaking voice, I, is the agent of the verb, thus, it is to be in the nominative case: *ʾašrabu*, as it has been recorded by narrators.

On what basis the poet found it legitimate to turn *ʾašrabu* into *ʾašrab*, we do not know. It could be the sense of metrical harmony, i.e. rhythm, which imposed the reading *ʾašrab*, grammaticalness being sacrificed for the sake of metrical harmony. It is worth noting that this happened a long time before al-Xalil ibn ʾAḥmad (d. 170/786 or 175/791) established the science of meter (*ʿarūḍ*) and in the absence, as far as we now know, of any theoretical awareness of demands of rhythm or prosodic 'correctness'.

This instance of what we now call 'ungrammaticalness' is rather rare in Arabic poetry throughout its history, with the exception of the pre-Islamic age. Verbs are much less exposed to the whims, conscious or unconscious, of the poetic Muse. Nouns are normally the field within which much deviation from the norms of implicit grammar occurs. This is true particularly of some proper names. No poet will use the noun, as a nonproper name, in a truncated form; this is because, naturally, a noun is the sign that helps the recipient to identify the signified, object or entity, of which it is a sign. Thus, for instance, *fara* would not be used in a poem about a battlefield, because no one would understand that what is meant is *faras* 'a horse'. However, as proper nouns tend to be less context-dependent and are not usually generative of larger units of meaning, and because they stand out as injected or inserted elements in a text, they can be more easily truncated. *Fāṭimah* is thus used as *Fāṭim* or *Fāṭima* or *Fāṭimu*, but only in a context where it is clear that a woman is the center of interest and that the word cannot be referring to a woman who has weaned a baby off her breast: *faṭama-faṭamat-fāṭimah*.

3. THE LEXICON OF POETRY

As we look at the sets of vocabulary prominent in Arabic poetry, we realize with ease that the phonetic composition of individual words evolved greatly from the first recorded samples to the present day. The pre-Islamic era undoubtedly has many more (in fact, a large body) of what to the modern ear at least may sound unmellifluous, lacking in lucidity, fluency, and *faṣāḥa*. And poetry appears at that point in history – perhaps surprisingly – to accommodate, without any difficulty, various strands of words, from items designating everyday life in its very practical aspects, to others relating to sexuality, the body, and bodily functions, to others still which belong to the semantic space of morality, spirituality, mythology, and divinity. (On this specific level, Modernist poetry appears by comparison to be much less adventurous or accommodating). But within that pre-Islamic space, the vocabulary differs significantly from one area to another, one social setting or group or even class to another, and from the desert (e.g. the Najd) to urban centers (Mecca and Medina, al-Ḥīra) and naturally fertile regions (such as al-Yaman). The most obvious contrast can be seen in the *Muʿallaqāt*, for instance, or the poetry of al-ʾAʿšā (d. after 629 C.E.) or ʿAdiyy ibn Zayd al-ʿIbādī (2nd half 6th century C.E.), in particular, on the one hand, and, on the other, the poetry of the satanic, particularly Taʾabbaṭa Šarran.

Traces of the 'heavy' qualities of pre-Islamic vocabulary survive well into Islam, especially in the poetry of figures like Ḏū r-Rumma and al-Farazdaq and the composers of → *rajaz*. But an emerging religious vocabulary, much more chiseled phonetically, begins to make its appearance. In addition, there soon develops to a remarkable degree the vocabulary of the semantic space of love, *ġazal*, and the body of the female, in *ʿuḏrī* and sensual poetry. Together with the new arts, singing and music, this has all caused something of a semantic explosion and a transformation of the phonetic and morphological structures of the vocabulary of poetic language. This is exemplified at its finest in the poetry of ʿUmar ibn ʾAbī Rabīʿa, Majnūn Layla, Jamīl ibn Muʿammar, and al-ʾAhwas, the first of whom, in particular, sounds as contemporary and refined today as he did in the 1st century A.H. The culmination of this process was reached in the poetry of the most refined, elegant, distilled, and musically crafted languages of poetry up to the end of the 2nd century A.H., that of the Caliph (no less!) al-Walīd ibn Yazīd (d. 126/744), the real founder of the poetics of wine, love, song, and sexuality in Arabic. We no longer find Imruʾ al-Qays's *mustašziratun* and *mutaʿaṯkili*, for instance, in al-Walīd's vocabulary, nor do we find al-Farazdaq's cumbersome syntax and labored diction. And this evolving space, with the more chiseled and refined vocabulary instituted in it, was to become the most solid foundation for the language of poetry in Arab history as it was further enhanced by poets like al-Buḥturī (d. 284/897), ʾAbū Tammām (d. 231/845), al-Mutanabbī (d. 354/965), and many others. It still is today, having been enriched throughout the centuries by the languages of logic, philosophy, *ʿilm al-kalām*, science, the sciences of language, schools of *fiqh*, Sufism, nature, military conquests, newly adopted traditions of fashion, dress, cooking, dance, singing, musical composition, administration, military organization, political, social and religious conflicts, and much else in city life and in new geographies, etc. etc. And throughout, it has also been enriched by the language of the *Qurʾān*, in more aspects than one.

This process of evolution, refinement, and selective usage of individual words was possibly responsible for the emergence of a critical tendency to attribute qualities of beauty and *faṣāḥa* to individual words and assign virtue to the → *lafḏ* 'words' (whatever this might mean) in poetic composition at the expense of the → *maʿnā* 'meaning'. A whole body of critical work was then devoted to the study of *faṣāḥa* in vocabulary and to the dichotomy of *maʿnā* and *mabnā* 'structure, composition' (or, in fact, *lafḏ*). Al-Jāḥiẓ (d. 255/869) has been wrongly said to have initiated the doctrine of favoring *lafḏ*. Ibn Qutayba (d. 276/889) judged poetry in complicated terms based on the quality of its *maʿnā* and *lafḏ* and the proportional degree of excellence of each of these two items in relation to the other in the same poem or lines of poetry. The dichotomy became very important for the exploration of the issue of the → *ʾiʿjāz al-Qurʾān* 'inimitability of the *Qurʾān*; miraculous nature of the *Qurʾān*', about which scholars were greatly divided. The most significant voice against the view that individual words in themselves may be (or may not be) beautiful or more suitable for use in poetry – or in fine writing outside poetry – was that of ʿAbd al-Qāhir al-Jurjānī (d. 471/1078), who rejected this view altogether and argued that any word whatsoever can be beautiful if used in a context in which it functions in harmony with the other constituent elements of that context. His revolutionary views precede ideas expressed in modern critical theory, both in the Arab world and outside it – in Western cultures – by ten whole centuries.

History was on al-Jurjānī's side, as the actual poetic tradition in the centuries that followed showed a huge expansion in the semantic and social space from which the vocabulary of poetry derived. Ironically, this is true especially of the period of history most critics wrongly label *ʿuṣūr al-inḥiṭāṭ* 'ages of decadence'. In these centuries, the language of poetry changes and expands and is enriched by daily life in its ordinary aspects, by tools, experiences, events, activities such as cooking, sport, gambling, sex, wandering, and all sorts of experiences from social deprivation to rebellion against poverty and exploitation by governments, states, or officials or simply by tricksters and conmen. Needless to say, the language of the more established realms of poetic inspiration continued to flourish side by side with these inventive new features of poetic language.

It should be noted, however, that while poets, the creative agents in the poetic space, were transforming poetry and the Arabic language as a whole in their own times as from the 2nd century A.H. and throughout the Abbasid age, scholars of the 3rd and 4th centuries A.H. recorded and studied the vocabulary of the earlier periods, especially the pre-Islamic one, with real zeal but paid much less attention to the language of their own times. Thus, a historical perspective on the language of poetry has never developed in Arabic studies, despite a couple of philologically oriented works on the emergence of a new set of vocabulary inspired mostly by religion. And when scholars did study the vocabulary, they studied it mostly in terms of *ġarīb* and *waḥšī* lit. 'strange and wild', but meaning in reality 'unfamiliar to them', 'difficult to assimilate semantically (or, at times, morphologically), with immediacy and without hard labor'. The only area in which scholars made valuable contribution was the literal and nonliteral/figurative (*majāzī*) uses of language (→ *majāz*). This branch of knowledge developed to an impressive degree right up to the 8th century A.H., and it did so within a reasonably clear sense of history, as it involved the debates and disputes over the ancient and the modern (*al-qadīm wa-l-muḥdaṯ; al-qudamāʾ wa-l-muḥdaṯīn*). This aspect of their work deserves attention in a more specialized study than the present entry.

4. THE MYSTICAL DIMENSION OF POETRY

In the Judaeo-Christian tradition, in the beginning was The Word. Language is at the initiation and commencement of existence. In the Arabic tradition, we have no such initial status for language, but a mystical dimension of the Word as a pure sign, a signifier without an easily determinable signified or even significance, makes an early appearance in pre-Islamic poetry. Perhaps if more of the rhyming chants of the *kuhhān* 'soothsayers; priests' before Islam had been preserved (→ *sajʿ*), we would have been able to explore this dimension of language in a more informed manner. Yet, even in the absence of such a corpus, there are linguistic features in poetry that radiate this indeterminable dimension, this mystical halo that surrounds language. Despite traditional interpretations which tried to give logical, rational explanations to the three or four phenomena to be mentioned here, one may without hesitation think of these phenomena as pure signifiers connected with a deeply rooted ritualistic, mystical dimension of the Word. Some pre-Islamic poems begin with the sound *ʾalā*, which has no contextually definable function except on the rhythmic level: it is part of the metrical composition of the line of poetry in which it occurs. The phrases *ʾalā, lahi* (part of *tallāhi*), perhaps meaning 'by God', *lā ʾabā laka* lit. 'you have no father', *laʿamrī* 'upon my life!', *laʿamruka* 'upon your life!', *yā ṣāḥi* 'o, companion!', *xalīlayya* 'my two friends!', *yā ṣāḥibayya* 'o, my two companions!' belong in this category of indeterminability. The → *Qurʾān*, while negating all connections with Arabic poetry, uses the same way of opening many of its suras, for instance, *ʾALM*, in sura 29; the only difference is that these were uttered as single phonemes rather than as a morpheme. And they are not, as far as is known now, part of a metrical unit that determines the rhythmic structure of the following sentence. But further research is needed to test this inherited wisdom.

It was not long before the Sufis explored the mystical dimension of language and focused on individual letters in their visual as well as phonetic features. Ibn ʿArabī (d. 638/1240) initiated a whole philosophy of mystical contemplation of language from this perspective. But before him, al-Ḥallāj (d. 309/922), as-Suhrawardī (d. 587/1191), an-Niffarī (d. 354/965?), and others had made significant contributions to this perception of language. Much originated in the Qurʾānic statement (Q. 2/31) that God taught Adam *al-ʾasmāʾa kullahā* 'all the names' (→ *waḍʿ al-luġa*); thus, words possessed this existentially mystical, powerful function of being the key to knowledge as *ʾasmāʾ* 'names, signs', not, significantly, as *ʾafʿāl* 'actions, deeds'.

The emergence of a distinct Arab art, called in some quarters, rather strangely, Islamic, in which the word constituted the *art object* in itself and was used as a Qurʾānic element and as a poetic element, contributed greatly to this mystical treatment of language. It also generated the interest in the purely visual aspect of words in poetry. Although the classical tradition itself did not develop this greatly, a whole art (that has also developed in Western poetry recently and has been called 'concrete poetry') grew and became hugely popular. The very zenith of this is the work of al-Jilyānī al-ʾAndalusī in the 6th century A.H., especially in his *Dīwān at-tadbīj*.

The Modernists turned this into a major aspect of their fresh handling of language, in poetry and art. Arabic as a medium of the plastic arts became the plastic art itself. The letter or letters of the alphabet are the medium and the artistic object at one and the same time. Some artists used this out of a Sufi vision, others purely as an esthetically motivated exploration of visually appealing forms. Diya al-ʿAzzawi (b. 1939) is one such artist, and others are Kamal Bullata (b. 1942) and Shadia Alem (b. 1960).

In poetry, the alphabet constitutes a central theme in the poetry of Adonis (b. 1930) at various points in his development into a great poet. Of the younger generation, Adeeb Kamal al-Din has distinguished himself mostly as a poet of *al-ḥarf* 'the letter'. His latest book is called *Šajarat al-ḥurūf* 'The tree of letters' (2007). Just as in art, the letters, the word, are at times used by some poets as a sign with a mystical significance and at times purely as form. In the very formation of the letter, the word, the phrase, then in their distribution on the page, totally without reference to any immediate semantic function, Modernist poetry asserted the multiple and vital role of language in the formation of the poetic text.

This, however, is countered by an extremely interesting treatment of the language of poetry in the work of a number of poets who explore what might well be called 'the function of absence' in the poetic text. This in reality is an aspect of the interplay between language and the absence of language, and it has so far manifested itself in two different forms. In one form, a poetic text using language in one part suddenly shifts to blank spaces (devoid of words); in the other form, the body of words with their material shapes and sounds is countered by passages using dots over a large space exploring the expressive power of silence in an environment, or within a structure, of sound and noise.

Finally, a more complex feature of language in poetry has won a great degree of popularity: the mixing of levels of language (including mixing dialects with *fuṣḥā*, popular song with formal Arabic), mixing of languages (Arabic, English, French), mixing of genres in what is called 'open texts', and of language as a semantically oriented body of signs with the language of pure sound or pure form: music and painting.

5. Poetry and modernity

It is reported that Xalaf al-ʾAḥmar (d. ca. 180/796), the master narrator of Arabic poetry in the Abbasid age, heard a poem by Baššār ibn

Burd (d. 167/784), the early Abbasid poet, in which the following line occurs:

> *bakkirā ṣāḥibayya qabla l-hajīr ʾinna ḏāka n-najāḥa fī t-tabkīr*
>
> ‘Start early, my two companions, before the heat of the noon: indeed that success is in the early start’

Xalaf apparently said to the poet: If only you had said *bakkirā fa-n-najāḥu fī t-tabkīr* ‘Start early, for success is in the early start’, to which Baššār responded that he composed the line in the manner of the *ʾAʿrāb* and the Bedouin, as this is the way they speak; adding that if he had said *bakkirā fa-n-najāḥu*…, it would have been “the way of the *muwalladīn*, the moderns, and not in keeping with the *maʿnā* of the poem”.

We are not very likely ever to know for sure what it is that Baššār identified – or felt – to be the qualities that give the second alternative reading the character of ancientness and convention and what it is that gives his rendering the stamp of modernity. We can venture and conjecture and guess. It is perhaps the repetition of the verb *bakkirā* and the more tightly logical cause/effect relationship established between *tabkīr* and *najāḥ*. But it may be the way the alternative choice breaks the link between the first and the second *šaṭr* by starting the second with *ʾinna*, as though a new sentence – semantic unit – is initiated, or the insertion of *ḏāka*, which is a demonstrative pronoun, to refer to *an-najāḥ* when *an-najāḥ* has not been previously mentioned. Or it may be other things.

The significance does not lie here. It lies in an implicit aspect of enormous importance: here is a poet who is marking a whole history of a culture and establishing lines of rupture and qualitative divisions in it merely on the basis of the way language is used in a single line of poetry. A stunning achievement. It took European culture ten centuries after that moment in history to assign to language this ontologically staggering role. And it has not yet even gotten to the point reached by Baššār.

6. THE POETIC REVOLUTION OF THE UMAYYAD PERIOD

A poetic revolution began early in the Umayyad period with the art of the Caliph al-Walīd ibn Yazīd. Al-Walīd urbanized the language of poetry fully, not only on the level of vocabulary, which derived almost exclusively from an urban, affluent setting, but also on the level of sensibility, rhythm, and imagery. His work was echoed by another Qurašī and Umayyad poet, ʾAbū l-Hindī, the poet of love, homosexuality, wine, and enchanting rhythm and imagery, from whom ʾAbū Nuwās (d. 199/813) was to derive a great deal of his art. But in both cases, language generally remained rather tame in its exploration of connections and relationships between aspects of experience, the world, and art. The message remained of supreme dominance.

It was with the burgeoning of *al-badīʿ*, with poets like Baššār ibn Burd, Muslim ibn al-Walīd, ʾAbū Tammām, and others, that a genuine poetic revolution took shape. *Al-badīʿ* was not simply the beautiful, the invented, etc. but a genuine shift in paradigm in Arabic culture as a whole, a shift from the message to the code: the code suddenly began to be the focus of attention in poetry as well as in prose and fashion, food, architecture, music, and many other aspects of human existence. And language as such became the center of gravity, at times taking priority over experience or thought processes. Within this shift, the most important aspect of the evolution and transformation had less to do with grammar and syntax than with assigning language a totally new role: the role of embodying a new urban, cultured, refined sensibility. And it was ʾAbū Tammām who embodied this shift throughout his poetry but especially in images and lines like this:

> *raqīqu ḥawāšī l-ḥilmi law ʾanna ḥilmahu // bi-kaffayka mā rayta fī ʾannahu burdu*
>
> ‘Thin and soft are the trails of his forbearance; if you hold his forbearance // in your palms, you will not doubt that it is woven of silk’

and this:

> *maṭarun yaḏūbu ṣ-ṣaḥwu minhu wa-baʿdahu // ṣaḥwun yakādu mina n-naḍārati yumṭiru*
>
> ‘Rain from which clarity dissolves, followed by // clarity that is so fresh it almost rains’

The Arabs used to describe *ḥilm* ‘forbearance’ in terms of heaviness, solidity, bulkiness, immobility, and ʾAbū Tammām knew that well; after all, he is the poet who compiled *al-Ḥamāsa*, one of the finest anthologies of

Arabic poetry ever compiled. But he advanced a new perception of *ḥilm; ḥilm* to him clearly embodied not heaviness and solidity but tenderness, silkiness of manner, as an urban, cultured, complex environment demanded its role to be. Thus, his *mamdūḥ* is *raqīqu ḥawāšī l-ḥilm*; not only is *ḥilm* a concrete, material matter, not only does *ḥilm* have *ḥawāšī*, but it is also thin and soft, so soft that if you hold it in your fingers you would be certain it is woven of silk. The whole culture has changed, and this change was embodied by a brilliant poet in just one metaphor.

In the second line, this new sensibility is evident in the abolition of the binary opposition *maṭar/ṣaḥw* 'rain/absence of rain'; no longer is the relation between these two felt or seen or conceived as an opposition. Much in pre-Islamic culture was deeply rooted in a sense of the oppositional nature of human existence and experience, and there was much in Islam to reinforce and deepen this type of sensibility; but to ʾAbū Tammām, binaries are now inherently connected; they are inwardly linked by an existential bond. Beauty, clarity, aesthetic values do not arise out of one phenomenon negating its opposite, but from the very innate procreational relationship that interlaces the two to each other. Each generates the other and is in turn generated by it, in a cyclical motion that represents a dialectical process rather than a process of one-way negation or affirmation.

Language now was embodying culture in a new, transformed shape, in the process of its evolution.

7. The phonetic structure of vocabulary in Arabic

Arabic poetry, perhaps more than any other force in the culture, has preserved the lucidity and purity of the language, first on the level of the structure of vocabulary, second on the level of its syntax and grammar.

The first is due to a large extent to an extremely important feature of the structure of vocabulary and the rhythmic structure in poetry (→ meter). Exhaustive statistical analysis has revealed that the structure of the basic feet that al-Xalīl identified as forming the basis for rhythmic formations in Arabic poetry is also that which forms the basis for the structure of vocabulary in the language, and that the rules which apply to the former also apply to the latter, with only minor exceptions and modifications. What this rather vague formulation means is that the basic feet *fāʿilun, faʿūlun, mustafʿilun, mafāʿīlun, fāʿilātun, mutafāʿilun, mufāʿalatun* comprise all the possible sequences of *mutaḥarrik/sākin* (or syllabic elements) which occur in Arabic poetry and in the morphology of vocabulary in the Arabic language itself. Significantly, the eighth foot in *ʿarūḍ, mafʿūlātu*, is superfluous in this respect and has no independent presence in this form in any word in Arabic. It is striking that some prosodists after al-Xalīl did not include it as one of the *tafʿīlāt*, and in the system proposed here for the rhythmic structure of Arabic, it has no existence.

The seven feet can be organized into three groups:

i. The *fāʿilun* group: *fāʿilun, mustafʿilun, fāʿilātun*
ii. The *faʿūlun* group: *faʿūlun, mafāʿīlun*
iii. The *mutafa* group: *mutafāʿilun, mufāʿalatun*

Al-Xalīl, for important reasons, among which was the desire to keep his system a binary one, considered the *mutafa* group to be a combination of *muta* + *fa* (- -) and (- o).

Every individual Arabic word is structured on the model of one of these feet with one single modification: just as in poetry some of these feet can end with an additional (– o) or (*fa*) as it is called here, some words in Arabic can be structured or are structured in some cases in a manner that corresponds to one *tafʿīla* with an additional (*fa*). Thus, in poetry *mustafʿilun* can be: *mustafʿilun* + *fa* = *mustafʿilatun*, and *mutafāʿilun* can become *mutafāʿilatun.* If *fāʿilun* is extended, it becomes *fāʿilatun*, which is nothing new. If *faʿūlun* is extended, it becomes *mafāʿīlun*, which again is nothing new. The *mutafa* group can be extended only into *mutaffa* = *fāʿilātun* or (- - - - o).

Al-Xalīl brilliantly discovered also that Arabic poetry rarely accepts the *mutaffa* sequence and never goes beyond it to form, for instance, *mutaffa* = (- - - - - o). Similarly, the *mustafʿilun* group, while accepting an extension at the end of a (–o), cannot easily accept an extension at the beginning that will turn it into (-o-o-o- - o).

Quite naturally, but amazingly too, when the Rubāʿī metrical form came into being in the later part of the Abbasid age, the prosodists could not describe it properly as it really is and resorted to peculiar metrical structures to describe it. The reason is precisely that the Rubāʿī begins with the unit (-o-o-o- -o) which, as mentioned above, Arabic poetry could not accommodate, up to al-Xalīl's time at least.

An examination of morphology and the numerous forms that *ʿilm aṣ-ṣarf* designates as the forms of words in Arabic reveals this amazing fact: the countless forms of → *ṣarf* are all comprehensively contained in the abstracted phonetic/morphological composition of the feet described above. For instance, *istaqbala*, *istaqalla, intaḍama, taʿabbada, tašāġala, taʿarrā, marasa, ixdawdara* have the following abstracted phonetic structures, which correspond with the *tafʿīlāt*:

-o-o- -o	*mustafʿilun*
-o- -o-o	*fāʿilātun*
- o- - -o	*muftaʿilun*
- - o - - o	*faʿūlun*
- - o - -o	
- - o - o	
- o - - o	*fāʿilun*
- o -o - - o	*mustafʿilun*

These are, in fact, the *muftaʿilun, mafāʿīlun, faʿūlun, fāʿilun, mustafʿilun*. Despite the fact that some of them are different as actual, derived *ʾawzān* of *ṣarf*, they are identical on the abstract phonetic level as units of sound, i.e. as metrical units. The entire body of the Arabic language is covered by the few units that we abstract as *tafʿīlāt aš-šiʿr*. This amazing quality of Arabic vocabulary is an instinctive key that allows us often enough to say reasonably quickly that a certain word is or is not Arabic. *Dīmūqrāṭiyya* cannot be Arabic, not because of anything to do with democracy as a concept and practice, but because the word itself does not fall within the scope of structuring the *tafʿīlāt* of Arabic poetry. *Dīmūqrāṭiyya* is this: (- o - oo- o- o - o), and Arabic does not accept any sequence that has five *sabab xafīf* or *fa* without a break, especially if two *sākin*s occur in it next to each other. The maximum is four, as in *mafulatun*. *Šarikatun*, on the other hand, is an Arabic word because it has no more than four *mutaḥarriks* in a sequence (- - - - o), and this is the maximum that is embodied in the rare *tafʿīla, fāʿilātun*.

Arabic poetry thus preserved the purity of structure and vocabulary of the language, the musicality, the rhythmic richness, and the lucidity, while prose, especially in the modern period, in the form of both journalistic and colloquial Arabic, corrupted these qualities of the language. The reason is that Arabic poetry throughout its history has had a remarkable sensitivity to the structure of individual words and has never – with very, very few exceptions – allowed imported words to be prominent. Even at the height of the ages of science, philosophy, religious debate, etc. in the Abbasid period and in the 20th century, poetry avoided and indeed resisted the use of such words (notwithstanding the fact that in very few intentionally composed lines, strange Persian words and phrases were used in poetry in the Abbasid period; none of them has, however, slipped into the main body of the language or poetry). Prose and spoken Arabic accepted them more readily and as a result have done much to blemish the eloquence inherent in the language.

8. ‘Ungrammaticalness’ in poetry

Under the heading of *ḍarūrat aš-šiʿr*, much was listed which can be examined from a different perspective, that of function within the overall structure of a poem, or at least a localized context. One of the best examples that may be referred to here does not involve a purely linguistic feature, but one related to thought processes. In a brilliant study of the → *istiʿāra ġayr al-mufīda* lit. ‘metaphor that is not beneficial’, ʿAbd al-Qāhir al-Jurjānī looked at the use of the word *mašāfiruh* to describe a human being, when it is in actual dictionary usage a description of the lips of a camel. This is a proper form of ‘ungrammaticalness’, but on the semantic level. He, however, rejected the view that it is ‘unuseful’ or devoid of significance, arguing that the poem was giving a negative image of a black man who was thick lipped. The poet used *ġalīḏun mašāfiruh* in order to connect the image of the man with that of a camel, thus presenting him as aesthetically more unpleasing than if he had said ‘thick of lips’.

Two other instances of ungrammaticalness to be considered are the *ʾiqwāʾ* in rhyming in two examples, one pre-Islamic and one Abbasid, and the use of a singular verb to qualify a dual subject.

The first case occurs in a poem by ʿUrwa ibn al-Ward, the leading poet of the *ṣaʿālīk* of pre-Islamic Arabia. The rhyme in the poem is based on the phoneme /r/ in the genitive case: *muʿṣiri, muxtiri, mutanaḏḏari*, etc. Suddenly, the poem shifts to *mušahharu*. *ʾIqwāʾ* thus is rhythmic ungrammaticalness, or phonetic ungrammaticalness; it is the violation of a rule on the phonetic level. It is the opposite of a *ḍarūra šiʿriyya* that is dictated by requirements of meter or rhyme, wherein a rule of grammar is violated in order to preserve the validity of meter or rhyme.

Superficially, this is a case of poetic license or *ḍarūra šiʿriyya* dictated by the rules of grammar at the expense of the requirements of phonetics (rhyme). Yet, ʿUrwa's violation could be regarded as aesthetically thrilling. The poem is by a leading *ṣuʿlūk*, a rejectionist, a violator of rules of the tribe and its values. In that violation of the tribal, communal requirements of the singularity of phonetic features of the rhyme of a poem, he embodies, within the linguistic structure and in his treatment of it, the very rejectionist attitude he practices on the level of actual behavior vis-à-vis the tribal, communal social order.

The second case occurs in the poetry of ʾAbū l-Hindī; once more it involves *ʾiqwāʾ*, violating a rule of phonetic grammar, and once more may be considered aesthetically thrilling. The poem depicts the experience of wine drinking, the pleasure of drinking in the morning and continuing throughout the day. Rhymewise, it is based on the phoneme /b/ in the genitive case (with a *kasra*). In the last line, an image of the wine is given that depicts its color as red, then closes with an image of a chameleon jumping around unable to bear the burning heat of the noon sun. An animal turning around unsettled, tossing itself over, restless, shuffling, etc. The word used is a verb meaning turning over, *yataqallabu*. At that point the phonetic feature dominating the rhyming of the poem is literally *turned over*: from *bi* to *bu*, and from nouns to a verb in the imperfect.

The third case comes from ʾAbū Tammām.

> *ḥattā ġadat wahdātuhā wa-najāduhā // fiʾatayni fī xalʿi r-rabīʿi tubaxtiru*
>
> 'Until its low lands and high lands became // two groups which in the garments of spring saunter'

This is a straightforward instance of *ḍarūra šiʿriyya*: a grammatical rule is violated to meet the demands of rhyme. Aesthetically, however, it is a supreme act of creativity. The entire poem is about the abolition of oppositions: seeing nature as a unified body in which winter and spring are not oppositional but interactive. The poet thus draws the entities gradually closer in each line of the poem until he gets to this point where he relates the two groups which perform the same act of sauntering. By not selecting the dual form, he has in fact reached the ultimate point of possibility in unifying the binary systems. They saunter as one body in unified colors wherein, in the following line, he drops any conjunction, saying, *muṣfarratan muḥmarratan*, rather than *muṣfarratan wa-muḥmarratan*, then saying *fa-ka-ʾannahā*.

9. Syntactic violations in poetry

Although cases of genuine ungrammaticalness in actual poems that form the main body of pre-Islamic poetry are generally of a restricted and confined nature, there are instances where the poetic structure is violated in a substantial fashion on the syntactic level. Of these, one of the most significant is that of Ṭarafa ibn al-ʿAbd, who, in a line of his *Muʿallaqa*, splits the adjective from the noun it qualifies by a verb:

> *wa-karrī ʾiḏā nādā l-muḍāfu muḥannaban // ka-sīdi l-ġaḍā nabbahtahu l-mutawarridi*
>
> 'And my rushing into attack when a guest calls for help // like a wolf out of the thickets you have provoked, darkish red'

The 'natural' linguistic string and word order here will be *ka-sīdi l-ġaḍā al-mutawwaridi wa-qad nabbahtahu*. The violation, however, enhances the power of the verb *nabbahtahu* and the image of the provoked, fearsome lion darting out of his bushes. But cases of syntactic violation are normally much less acute. They often take the form of a certain mode of formulation which allows the possibility of relating one linguistic element, e.g. a single word, to more than one agent that precedes it. This generates a certain degree of ambiguity, which is often enriching on the semantic level. An example in fact occurs in Ṭarafa's line itself in the word *muhannaban*. Initially, *muhannaban* can be a *ḥāl* of both agents, the 'I' of the speaking voice or the 'he' of the *muḍāf*; on further contemplation, though, the second possibility can be dropped. Al-Jurjānī lavishes great praise on some such expressions, as they reward the mind, after guiding it into different paths, with great pleasure when it eventually attains the more convincing and coherent meaning. Yet, if this process of violation is pushed to extreme degrees, the language structure can obstruct the process of understanding and at times render it impossible. Such cases are viewed negatively and are studied in some texts as instances of *muʿāzalat at-tarākīb* and considered to lack *faṣāḥa* and the power of *bayān* 'revelation, revealing'. Al-Jurjānī, among others, devoted a great deal of energy to this issue. So did ʿAbd al-ʿAzīz (al-Qāḍī) al-Jurjānī (d. 366/976), and al-ʾĀmidī (d. 631/1233), who devoted whole sections of his book *al-Muwāzana* to comparing the ways in which al-Buḥturī and ʾAbū Tammām behaved in their use of language.

There were times, in the history of poetry, when violating the rules of grammar was an intentional act of an explicit nature. Most famous of such cases was the action taken by al-Farazdaq at a time of a burgeoning conflict between poets and grammarians, whose authority was just beginning to be felt and asserted. It is reported that the grammarian ʿAbdallāh al-Ḥaḍramī objected to a line of poetry by al-Farazdaq. Incensed by the audacity of the grammarian to correct him, being the great poet and pure Arab who grew up in the desert where Arabic came into being, al-Farazdaq vowed to satirize the grammarian in a line "that will remain forever on the tongues of grammarians". And he honored his oath, because the line and the story behind it are still quoted today.

> *fa-law ʾanna ʿabdallāhi mawlā hajawtuhu // wa-lākinna ʿabdallāhi mawlā mawāliyā*
>
> 'A slave of a slave you are; you are not even worthy of being satirized'

The second *mawāliya*, of course, should have been, to be grammatically correct, *mawālin*.

Intention or the lack of it, however, does not affect the aesthetic and expressive value of violations of grammar, syntax, or other rules. And it is almost impossible in most cases to determine whether a violation is or is not intended by the poet. One such ambiguous act, but one that is among the most brilliant violations of language in the Arabic poetic tradition, is the act performed by the great, playful poet ʾAbū Nuwās (d. 199/813) in one of his finest wine poems, which opens with:

> *daʿ ʿanka lawmī fa-ʾinna l-lawma ʾiġrāʾu // wa-dāwinī bi-llatī kānat hiya d-dāʾu*
>
> 'Drop your chiding of me, your chiding is temptation // and cure me with that which was itself the ailment'

The very opening line of this poem, the one just quoted, creates an ambiguous situation within which a violation of grammar is one possibility. The phrase *kānat hiya d-dāʾu* should really be *kānat hiya d-dāʾa*, the last word being the *xabar* 'predicate' of *kāna*. But there it is, used 'incorrectly' by a man with enormous knowledge of Arabic. Grammarians had to look for some 'rational' explanation that would render the sentence grammatically correct. Perhaps what they came up with is not convincing. But the magical act of violation, which is evident already in the first line (violating the moral value system and, at the same moment, violating the rules of the grammar of the language of that moral system) is still to come. ʾAbū Nuwās goes on to articulate his scorn for the traditions of Bedouin life, wherein the Arabs and their culture originated, and his preference for urban, affluent life in the city. He thinks of this in terms of an opposition between desert, tents, camels, sheep – the constituent elements of desert life – on the one hand, and on the other hand, pearls – the very symbol of affluence and civilized existence in which wine is produced, served, and enjoyed. Thus, he gives wine a proper name, without even hinting that he does that. He calls her (it?) *Durra* 'Pearl'. He then constructs this brilliant line of poetry:

> *ḥāšā li-durratin ʾan tubnā l-xiyāmu lahā // wa-ʾan tarūḥa ʿalayhā l-ʾiblu wa-š-šāʾu*
>
> 'God forbid that Durra should have tents built for her // or that it/she be passed by camels and sheep'

treating the word *durra* as a proper noun, a name of a real, human female. By doing so, the name becomes *mamnūʿ min aṣ-ṣarf*, i.e., grammatically it cannot show an *i* (*kasra*) at the end in genitive (*jarr*), when governed by *li-*, but has to show an *a* (*fatḥa*) instead.

But all of this is conjecture. The poem does not tell us any of this. So, how can we justify this conjecture? Because, one might argue, if *durra* is treated as a normal word, i.e. not as a proper name, the correct reading should be *ḥāšā li-durratin ʾan tubnā l-xiyāmu lahā*, and metrically, this is wrong, and a poet like ʾAbū Nuwās would not produce a line so metrically 'ill'!

The alternative description of the line would be to say simply that we have here a violation of a grammatical rule, namely that *durra* did not behave correctly and has to be read with a *fatḥa* rather than the *tanwīn* of *kasr* to maintain the metrical correctness of the line, and with no justification of this except to say that it is a *ḍarūra šiʿriyya* 'poetic necessity'. That would be sheer laziness and displays a lack of desire to explore subtleties of poetic language that lie beyond the visible and immediately perceptible.

10. Poetic imagery

The most exciting aspect of language in Arabic poetry is perhaps the language of poetic imagery in all its forms and types. Aristotle considered the ability to use metaphor well a sign of genius; accepting his criterion will place a huge number of poets throughout the history of Arabic poetry, from those who composed only a few lines to those who composed volumes of poetry comprising hundreds of pages, into the category of genius without any need for much debate. A certain bias, or in fact a misconception, in much Western critical theory holds the view that simile is the simpler form of constructing images, which relates to a primitive, undeveloped stage of human culture and which therefore dominates early literatures in the world including Greek poetry. According to this view, metaphor, the more sophisticated device, distinguishes later, complex stages of human culture and is therefore less common in early literatures of the world. Nothing can invalidate such bias, or correct such a misconception, more effectively than to look closely at simile and metaphor in pre-Islamic poetry and throughout the centuries of poetic activity in the Arab world. As early as Imruʾ al-Qays, we find stunningly complex metaphors rising here and there in texts of genuine intellectual complexity; moreover, the predominance of simile itself is questionable. Though frequent in a number of the *Muʿallaqat*, in many other poems simile is by no means the dominant form of creating imagery. Metaphor plays as vital a role as simile in shaping the vision and creative process of much early poetry. The *Qurʾān*, being a text deeply rooted within this cultural climate, abounds in metaphors of real complexity just as does poetry in the age of the Umayyads. By the 2nd century A.H., the use of complex metaphors and wildly inventive images was becoming a feature of a new sensibility. *Al-badīʿ* was a product of this intellectually and imaginatively, i.e. linguistically, complex climate of thought and imaginative conceptualization (→ *istiʿāra*; → *majāz*).

It took a couple of centuries of contemplating poetry and designating certain of its features with specific terms before an inspired critic was able to formulate a theoretical framework for the identification and analysis of *al-Badīʿ* as a comprehensive structure of poetic and linguistic devices. Ibn al-Muʿtazz (d. 295/908) identified *al-badīʿ* as consisting of five elements; the overall landscape of *al-badīʿ* that these elements create is one in which the emphasis in poetic language shifts from the message to the code, to the linguistic, phonetic, and semantic relationships between signs within a line or a text. None of the elements, Ibn al-Muʿtazz correctly argued, was novel; they all existed in pre-Islamic poetry, in various sayings, as well as in the *Qurʾān*, the *Ḥadīṯ*, normal speech, and poetry before the Abbasid period. One exception was what he called *al-maḏhab al-kalāmī*, which he said did not exist in the *Qurʾān* (see Abu Deeb 1994 for an explanation of why he made this exception).

11. The Abbasid Modernists

This first revolution within the structure of poetic language was not restricted to poetry; it became in no time a feature of a whole cultural epoch. Language assumed a new status as a target in itself, in most of its aspects. And the *maqāma* as a genre embodied this new status in a most powerful fashion. So did poetry. This continued right up to the modern era. Thus, when Modernists talked about the centrality of language in poetry, about the crucial role of imagery, from metaphor to symbol, they were not breaking virgin territory at all, nor were they copying Western literary production. They were digging in familiar earth, plowing familiar fields but necessarily with a different emphasis on function and relation to the overall body of language in a poetic text. One thing they have definitely done, however, is to extend the imaginative space within which various phenomena in the world can be conceived as being related. And in this they were not inventing *ex nihilo*, nor imitating Western poetry: they were repositioning the Arabic poetic tradition and subverting the relationship between center and margins in it. They placed the Sufi imagination at the center and moved the accepted, celebrated body of thought to the margin. And the Sufis were the very visionaries who had invented the language of irrational, illogical interlacing and the notion of the unified nature of the universe and abolished the barriers between man and God, as well as between the constituents of the world they had created. An-Niffarī became Adonis's model, not only intellectually but even as a structure of linguistic formulations. His prose was declared to be supreme poetry and has come to play a hugely inspirational role in the rise of the prose poem and the transformation of the language and structure of Arabic poetry. A similar role was played by Ibn ʿArabī, especially in his mystical treatment of language and his conception of the imagination. Al-Ḥallāj, though more as a sacrificial hero than a maker of discourse, received no less attention.

The Modernists, however, shifted the emphasis from the phonetic aspect of the language of poetry to its conceptual, basically semantic, aspect. The least frequent element in Modernist poetry of the five elements identified by Ibn al-Muʿtazz, is *jinās* → 'paronomasia', and the most frequent is the complex, far-fetched metaphor, followed by *ʾiṭbāq* 'antithesis'. The intellectual element lying at the heart of the process Ibn al-Muʿtazz called *al-maḏhab al-kalāmī* forms a natural constituent of Modernist poetry and has become far more decisive from the 1990s onward, especially in the *poème en prose*. The use of this phenomenon could be called the poetry of *fatla* 'twist', which tends to form the last line or two of a poem functioning more like a 'capping' of the text in a manner similar to what Shakespeare is thought to have done in many of his sonnets.

12. The contemporary Modernist movement

While the Abbasid Modernists did their work without much theoretical debate or theorizing about language in poetry, the Modernists of the contemporary era have done a great deal of talking about language in poetry and language in culture. The slogan propagated by many pioneers in this period, *tafjīr al-luġa* 'exploding language', has gained great currency and has had an immense impact on the younger generation of poets. Part of the very ideology of the Modernists has been that language is the great bearer and carrier of traditions in Arab life. Therefore, the most important task in challenging and changing these traditions has to be the rejection of the language of those traditions and the invention of a new language for a new age. The nucleus of this notion emerged in a famous essay by Jubran Khalil Jubran (d. 1913), back in the early part of the 20th century, *lakum luġatukum wa-lī luġatī* 'you have your language, and I have mine'. In another essay, Jubran connected the future of Arab life, language, and culture to the future of the language itself and to the fate of the creative impulse in Arab life. With time, language has become a central ideological as well as aesthetic issue within the actual written poetic text, not only within critical discourse. According to Abu Deeb (1994), the so-called Arab revolutionary movements in the 20th century failed because they never had the courage to face and challenge two major constituents of life and culture: the Arabic language and religion, mainly Islam. He also shows that language is a fundamental formative element of what he calls the *ḥadaṯī* (roughly, 'modernist') consciousness, becoming not solely a *medium* of expression but an active agent in the poetic process itself within the individual text. Phrases like "My language writes me, my blood reads me" etc. are very common, and the idea of culture being a text or the body being a text has become very popular. The terminology and linguistic idiom of writing, reading, textuality, deciphering, codes, signifiers, signs, signifieds, significance, etc. are very much part of the jargon and surface of poetic texts in the modern age.

Yet, ironically perhaps, the Modernist poetic tradition is almost free of ungrammaticalness on the level of single grammatical elements. *Ḍarūrat aš-šiʿr* seems to be something of the past, and poetry now is much more pure and conforming to grammatical rules than ever before. This is true of the main corpus of poetry, but it has been a major complaint in modern critical discourse that the poetry of the younger generations throughout the Arab world displays a shocking lack of knowledge of Arabic grammar. Many linguistic occurrences in today's poetry are simply considered 'mistakes', and are not studied even as poetic necessities. The other side of this coin is that poetry is much more adventurous and violative on every other level from prosody, rhyme, and metaphorical language to the overall structuring of a poetic text. Freedom appears to be a complex process and to drive people in very different, often contradictory, directions. There is almost an inversion of the classical tradition: ungrammaticalness on the level of grammar vs. ungrammaticalness on the level of imagery and semantic coherence. The modern (with few exceptions) is much more adventurous with semantics and imagery, and less so with grammar; the ancients and old Modernists were more adventurous with grammar and less so with semantics and imagery.

Within the Modernist movement, a great range of manners of handling the established patterns of Arabic syntax have emerged. In general terms, there is a far greater degree of freedom from the dictates of syntactic rules, but at the same time there is a much more conservative spirit that abides by many of the established norms. Phenomena such as starting a sentence with a *ḥāl*, or *tamyīz*, or another one of the *mafāʿīl*, have become common, the qualifying verb coming much further along the linguistic string. At times, the gap between the surface structure and the deep structure of a sentence is untraditionally wide. And a number of texts by Adonis and other poets explore possibilities of structuring never attempted before. This is especially evident in long texts of prose poetry or poetry using classical meters in modified forms. In "This is my Name", for instance, in a string of words 1 2 3 4 5 6 7 8 9 10, numbers 3 and 4 can be read as a part of 1 2 3 4 5 or as a part of 3 4 5 6 7, as can 8 and 9. This ambiguity is intensified in some texts which do away with most punctuation marks and allow the text to flow almost as texts do in manuscripts of Arabic prose of the classical age. It all appears to be, in addition to having aesthetic and artistic aims, a manifestation of the desire to tear apart the language of the tradition and invent a new language for a modern sensibility. One interesting aspect of this same process seems to be the growing tendency to use words that possess strong audible or visual properties and distribute them on the page in a manner that reveals these salient properties, often imbuing the poem with a strong onomatopoetic character or a touch of the qualities of sculpture or visual representation of space. Combined with other attributes of poetic language in the modern period (such as tone, suggestiveness, the shift in many texts from the voice of the first person to the second, the language of symbolism and the use of words and combination of words in a totally fresh, unconventional fashion), such tendencies generate the feeling that the continuity of the Arabic poetic tradition has been disrupted and a rupture has occurred that almost completely severs the umbilical cord between the past fourteen hundred years and the last fifty years. Needless to say, this is not true across the board, for the language of Arabic poetry in many circles and across vast spaces has not dramatically changed from the way it has been since the Abbasid age.

13. The coherence of the image

One could perhaps formulate the law, or implicit rule, in accordance with which the nature of poetic imagery has evolved, changed, been transformed and revolutionized in terms of a *gap* or *distance of tension*, as Abu Deeb (1979) calls it. The notion of 'gap' can be expressed in different terms. In a statement relating A to B, whether through similarity or contradiction or any other relations, a gap between A and B exists. In order for the statement to have validity, the coherence of the statement must be established. From "a night like the waves of the sea", by Imruʾ al-Qays, to the "black snow" by Rimbaud, or "green perfumes like the flesh of young children" by Baudelaire, coherence can be established with reference to a wide variety of things. There are times, however, when coherence eludes us. Future generations may or my not be able to establish (or discover) it. Alternatively, the statement can remain incoherent for an indefinable time or even for ever. At times, the incoherence of the statement can itself be the functional element in the text, incoherence thus being a source of coherence on a higher level of existence.

In these terms, it can be said that the coherence of the image in ancient Arabic poetry was generally evident or possible to establish with reference to the immediate linguistic and imaginative context that can be one line, two, or, in some extended similes, ten or more. The historical curve of Arabic poetry shows that as from the Abbasid period, the coherence of the image cannot be established only with reference to the immediate context but needs to be placed in a wider context within the text itself or, at times, outside it. In modern poetry, the latter situation has almost become the norm, rather than the exception: a coherent image cannot be established except with reference to the widest possible context of the text, the entire oeuvre of the poet, or even the entire cultural context and beyond. Sometimes, even that does not help us to see/feel/recognize a coherent image or statement.

From the pre-Islamic era to the appearance of *al-badīʿ* as a total structure, the gap between the two constituents of the image was generally narrow and conceivable or perceptible in logical, real, or intellectual terms. *Al-badīʿ* witnessed the first revolution that widened the gap and intensified the tension at times to a breaking point. Al-ʾĀmidī, who in the 4th century A.H. formulated the traditional criteria for this aspect of Arabic poetry, especially for metaphorical language, screamed in despair at some of ʾAbū Tammām's metaphors and, on one occasion at least, accused him of madness. ʿAbd al-Qāhir al-Jurjānī glorified images that were based on discovering similarity between widely dissimilar objects or entities, considering the most beautiful metaphors to be those that represent *šiddat iʾtilāf fī šiddat ixtilāf* 'intense connectedness in intense difference'. The modern age witnessed a return to classical models in the early stage, in the work of poets like Ahmad Shawqi (1868–1932) and Hafiz Ibrahim (1872–1932), before the semi-Romantic impulse began to widen the gap again. The Modernist movement has produced a metaphorical language of immense complexity in which the gap, the distance of tension, approaches a surrealist and at times postsurrealist level. Adonis, ʾUnsi al-Hajj (b. 1937), Muhammad al-Maghut (1934–2006), Qasim Haddad (b. 1948), Mahmud Darwish (1941–2008), Khalil Hawi (1919–1982), Shawqi Abi Shaqra (b. 1935), Kamal Abu-Deeb, Salim Barakat (b. 1951), Walid Khazendar (b. 1950), and many others have all shown special fondness for this type of imaginative construction. The younger generation writing today, in its finest examples, partakes of this new trend and pushes the gap to a point beyond breaking, a point at which similarity, links, connections, bonds are simply impossible to discern, detect, or even perceive or visualize. It is a point wherein the *absence of signification* replaces the older model in which signification, of various degrees of hiddenness and clarity, was still attainable at one point or another of the process of reception. However, within the corpus produced by the younger generation, something of a return to simpler, almost classical narrowness of the gap between the constituents of the image is also observable.

Finally, it is worth noting that within this Modernist corpus, two distinct directions have been taken in relation to the nature of poetic language and its relation to the language of daily life. One current of thought aspires to approximate the language of poetry, from its vocabulary to its syntax and rhythm, to the language used in daily life or to language written for general purposes in the Arab world; another has sought to purify language and distil it to a point beyond the mendacity of current usage, familiarity, and prosaic qualities. Each represents not only an artistic and aesthetic choice but fundamentally a political and ideological one as well. The conflict between these two powerful trends continues today, with the balance so far tilting in favor of the second of them. But the first has been gradually gaining ground especially in the poetry of the very young generation and particularly in prose poetry written by young women. And as this seems to be the age of women, the map of the future of language in Arabic poetry appears to have been charted already.

14. CONCLUSION

Throughout its history, the language of Arabic poetry has been intense, often dramatic and highly sensuous, as evident in particular in the humanization and animation of the natural, nonhuman world as well as abstract concepts and notions. Concreteness of vocabulary was prominent in pre-Islamic poetry, notwithstanding the presence of a strong element of abstract, nonconcrete features in the poetry of contemplation and *ḥikam* 'aphorisms' that we encounter in many major poems. Gradually, this abstracted, intellectual aspect began to flourish and, at times, came to be the hallmark of the work of certain poets (e.g. al-Mutanabbī and ʾAbū l-ʿAlāʾ al- Maʿarrī). In one variety of it, it became far too abstract and cerebral, lacking both concreteness and a feel of real, individual experience, as well as emotional intensity, especially in poems written to illustrate a certain view of the world or a religious doctrine. Sufi poetry at its worst was of this type, but in many instances it was poetry of experience of the highest order. In modern poetry, this feature of cerebrality has faded away, and a different manifestation of intellectualism has flourished in its place: highly contemplative, intellectually sophisticated, confessional poetry, often mixed with a Sufi touch and vocabulary, as in many poems by Yusuf al-Khal (1917–1987), Salah ʿAbd al-Sabur (1931–1981), Adonis, Tawfiq Sayigh (1923–1971), and Samir al-Sayigh.

Furthermore, the language of Arabic poetry has always been a language of strong patterning, in various ways of creating patterns, such as parallelism, repetition, and symmetry. But in one of its forms, patterning is a most distinctive feature in the very prosodic and rhyming requirements of poetry in its most widely accepted definition.

In the latter form, this reached its climax in the poetry of ʾAbū l-ʿAlāʾ al-Maʿarrī (d. 449/1057), especially in his *Luzūmiyyāt: Luzūm mā lā yalzam* 'The imperativeness of the nonimperative', which is a virtuoso display of the incredible linguistic flair al-Maʿarrī had. In modern poetry in general, there has been a tendency to avoid most aspects of the poetic 'craft' which reveal a conscious sense of craftsmanship, of *making* poems, rather than just *writing* them or allowing them to *flow*, just as there has been a strong desire to be freer of most preconditions and requirements of the Art. This has climaxed in present-day prose poetry in much of whose instances we have the nearest thing we get to what is actually *prosaic prose*, which has very little to do with the language of poetry either in its classical or most modern definitions.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Fārisī, *Šiʿr* = ʾAbū ʿAlī al-Ḥasan ibn ʾAḥmad al-Fārisī, *Kitāb aš-šiʿr*. Ed. Maḥmūd aṭ-Ṭanāḥī. Cairo: Maktabat al-Xānjī, 1988.
Ibn ʿUṣfūr, *Ḍarāʾir* = ʾAbū l-Ḥasan ʿAlī ibn Muʾmin Ibn ʿUṣfūr al-ʾIšbīlī, *Ḍarāʾir aš-šiʿr*. Beirut: Dār al-ʾAndalus, 1980.
Jurjānī, *ʾAsrār* = ʾAbū Bakr ʿAbd al-Qāhir ibn ʿAbd ar-Raḥmān al-Jurjānī, *ʾAsrār al-balāġa*. Ed. Helmut Ritter. Istanbul: Government Press, 1954.
——, *Dalāʾil* = ʾAbū Bakr ʿAbd al-Qāhir ibn ʿAbd ar-Raḥmān al-Jurjānī, *Dalaʾil al-ʾiʿjāz*. Ed. Fāyiz ad-Dayah and M.R. ad-Dayah. Damascus: Maktabat Saʿd ad-Din, 1987.

SECONDARY SOURCES

ʿAbd al-Muṭṭalib, Muḥammad. 2007. "Rakāʾiz at-tajdīd fī šiʿriyyat al-ḥadāṯa". *Ġayman* (Ṣanʿāʾ) 1.
Abu Deeb, Kamal. 1979. *Al-Jurjani's theory of poetic imagery*. Warminster: Aris and Phillips.
——. 1979. *Jadaliyyat al-xafāʾ wa-t-tajallī*. Beirut: Dār al-ʿIlm li-l-Malāyīn.
——. 1986. *ar-Ruʾā al-muqannaʿa*. Cairo: al-Hayʾa al-Miṣriyya al-ʿĀmma li-l-Kitāb.
——. 1987. *Fī š-šiʿriyya*. Beirut: Muʾassasat al-ʾAbḥāṯ al-ʿArabiyya.
——. 1994. "al-Luġa mukawwinan min mukawwināt al-waʿy al-ḥadaṯī". *Nizwa* (Muscat) 1.
Ḍayf, Šawqī. 1960. *al-Fann wa-maḏāhibuhu fī š-šiʿr al-ʿarabī*. 4th ed. Cairo: Dār al-Maʿārif.
Jakobson, Roman. 1987. *Language in literature*, ed. Krystyna Pomorska and Stephen Rudy. Cambridge, Mass.: Harvard University Press.
Kristeva, Julia. 1984. *Revolution in poetic language*. Trans. Margaret Waller. New York: Columbia University Press.
Mukařovský, Jan. 1976. *On poetic language*. Trans. John Burbank and Peter Steiner. Lisse: Peter de Ridder Press.
Riffaterre, Michael. 1978. *Semiotics of poetry*. Bloomington: Indiana University Press.
Saʿīd, Xālida. 1979. *Ḥarakiyyat al-ʾibdāʿ*. Beirut: Dār al-ʿAwda.
Sāmarrāʾī, ʾIbrāhīm as-. N.d. *Luġat aš-šiʿr bayna jīlayn*. Beirut: Dār aṯ-Ṯaqāfa.
——. 1984? *Fī luġat aš-šiʿr*. Amman: Dār al-Fikr.
Warrāqī, as-Saʿīd al-. 1979. *Luġat aš-šiʿr al-ʿarabī al-ḥadīṯ*. Cairo: al-Hayʾa al-Miṣriyya al-ʿĀmma li-l-Kitāb.
Xayr Beik, Kamāl. 1982. *Ḥarakat al-ḥadāṯa fī š-šiʿr al-ʿarabī al-muʿāṣir*. Beirut: Dār an-Nahār.

KAMAL ABU DEEB
(School of Oriental and African Studies)

Sicily

The geographical location of Sicily in the central Mediterranean, between the mainland areas now referred to as Italy and Tunisia, largely determined the complex linguistic history of the island in the ancient and Medieval periods. Thus, for most of the Classical and early Medieval periods (ca. 835 B.C.E.–535 C.E.), dialects of Greek and Latin came to predominate over the miscellany of tongues used by a range of indigenous and immigrant peoples. In addition, Neo-Punic was also attested but to a much lesser extent, and only in the west of the island. Nonetheless, the Romano-African author Apuleius could still speak of the *Siculi trilingues* in the mid-2nd century C.E., a reference to multilingualism that was picked up again a millennium later. However, on the eve of the Arab-Muslim invasion from Aghlabid Ifrīqiya that precipitated the disintegration of Byzantine rule (535–827 C.E.), forms of Sicilian Greek are thought to have been the island's main language, although the extent to which Latinate dialects had continued to coexist remains a matter of debate.

The subsequent Muslim conquest and settlement not only introduced Islam as the main religion but with it Arabic as the island's prestigious new lingua franca. These factors strongly determined the direction of acculturation for both the indigenous and immigrant population as a colony of Ifrīqiya under Arab-Islamic rule (827–ca. 1072 C.E.). During this period, the Byzantine Greek church in Sicily came close to total collapse, yet there remained strong concentrations of Christian influence in the mixed communities of the island's north-eastern corner, and to some degree this is borne out by the mottled distribution of Arabic and Greek toponymy. Although there is little reason to doubt the Arabic sources' claim that most people converted to Islam, the remaining Christian communities, whose strong religious identity was bound up with the Greek language of their liturgy, intermittently provided stiff resistance throughout the two centuries of Muslim rule. Consequently, by the end of the Islamic period, many of these Sicilian Christians were likely to have succumbed to varying degrees of Arabic-Greek bilingualism, while the majority of the island's population was Arabic speaking and Muslim. The main urban environments, which were probably subject to substantial repopulation, quickly appear to have assumed an Arab-Islamic character, with political and cultural life being heavily concentrated in the island's largest town, Palermo. In contrast, sketchy evidence suggests that inland rural areas underwent much slower processes of social, religious, and linguistic assimilation. A small number of short-lived colonies were established on the south Italian mainland and possibly in Sardinia, too, while the peripheral southerly islands of Pantelleria and → Malta were repopulated with Arabic speakers, whose dialects would persist beyond those of the Sicilians. Whether these islands were populated from Sicily or North Africa or both is unknown, and many of the arguments designed to describe the relationship between Sicilian Arabic and Medieval → Maltese have proved difficult to establish with certainty.

If Berber dialects managed to survive at all in the crowded and competitive language situation of fiercely anti-Berber Sicily, their impact has left barely the faintest trace in Sicilian Arabic, and no convincing examples can be found in later Romance-based Sicilian dialects. A small amount of Sicilian toponymy reflects Berber tribal names.

The Norman period (ca. 1061–ca. 1194) witnessed the chaotic end to Muslim dynastic rule, the introduction of the Latin church, and a wide range of colonists from the European mainland. A new ruling elite began to emerge which increasingly included 'Latin' Christians who were not native to the island, as well as some Muslims, converts, and multilingual Christian administrators. In addition, overwhelming numbers of settlers were attracted from the Italian mainland, particularly from the northern regions covered by modern Piedmont, Liguria, and Lombardy. Thus, to a large extent, the introduction of Arabic-speaking elites and colonists along with the socioreligious and linguistic assimilation toward Arab-Islamic norms that had been brought by the Muslim conquests were reversed in the Norman period.

Arabic was said by independent Medieval Arabic sources (Ibn Saʿīd al-Maġribī and Ibn Jubayr) to be known by the Norman kings and was prominent as the principal working language of the royal palaces and fiscal administration. Nonetheless, in the light of a deteriorating politico-religious situation around the island from the 1160s, the association of Arabic with Islam was viewed negatively by the recent settlers and insurgent landholding classes whose miscellaneous Italo- and Gallo-Romance dialects unambiguously identified them as adherents of the Latin church. As such, the prestige of Arabic as one of the royally adopted languages of inscriptions, coinage, and chancery documents cannot always be reconciled with its decline outside the palaces. Similarly, Frederick II's (d. 1250) harsh repression of the Sicilian Muslims did nothing to reduce his eclectic use of Oriental imports that consciously echoed his Norman predecessors in Palermo.

The survival of a few Arabic terms in later Sicilian dialects (e.g. *cajitu* 'local leader' < *qā'id; taibbu* 'a fine wine' < *ṭayyib* 'good'; *defetari* 'record books' < *dafātir; saia* 'water-irrigation channel' < *sāqiya*) suggests a degree of transitory Romance-Arabic bilingualism in some quarters. However, given the relatively brief (ca. 1100–ca. 1250), antagonistic, and privileged presence of Romance-speaking Christian settlers in areas where the low-prestige language of the Muslims was also used, forms of Latin-Arabic bilingualism were short-lived by comparison with the much longer history of Greek-Arabic social intermingling, religious conversion, acculturation, and bilingualism on the island (ca. 850–ca. 1250). When Muslims were assimilated into Christian communities during the 12th and 13th centuries, the evidence points to absorption principally by their old bilingual Greek neighbors, rather than by the immigrant, nonindigenous 'Latin' communities. For their part, under renewed Christian rule of the Norman period, the bilingual Arabic-Greek Christians are thought to have increasingly resorted back to Greek dialects, which were becoming ever more Italo-Greek in nature. However, many of the finer details of this period's complex socioreligious history that underpin considerations about the wider language situation cannot be established with certainty, and were evidently subject to many local variations.

It is clear that even prior to the year 1100 many Muslims who could afford to had quickly abandoned the island for the safety of North Africa, al-Andalus, or Egypt. In doing so, the island's intellectual output in Arabic collapsed, with al-ʾIdrīsī and Ibn Qalāqis the most notable (but nonnative) authors of the Norman period. The island's remaining Muslims became ever more concentrated in the regions toward the southwest of the island. A long series of Muslim revolts began in 1189 and ended with large-scale deportations to the colony of Lucera on the Italian mainland during the 1220s and 1240s under Frederick II. The rapid decline of Arabic on the island from the end of the 12th century was thus accelerated even further during the 13th. Arabic continued to be used in the large Lucera colony until its dissolution in 1300, while forms of Judeo-Arabic persisted on the island until the expulsion of the Jews at the end of the 15th century (Rocco 1995). Apart from the Jews, it is doubtful whether the remaining population of Arabic speakers, which now consisted of increasingly exiguous numbers of bilingual Christians and converts from Islam, could have reproduced themselves for very long as Arabic-speaking communities on Sicily much beyond ca. 1250.

Along with al-Andalus, the multilingual environment of Sicily provided a medium of transmission for a relatively small amount of Arabic vocabulary into various modern European languages. In almost all cases, these consist of nouns, and most are derived from the fields of commerce, technology, and material culture. Although later Medieval and modern Sicilian dialects are distinguished from other Italian dialects by the presence of Arabic interferences and loanwords, those elements are both slight and superficial, the language having been effectively obliterated by events of the 13th century. Most Arabic loanwords in Italo-Romance dialects of the later Medieval period are nouns that relate to the fields of daily life activities, commerce, flora, fauna, farming, fishing, and physical geography (Pellegrini 1972; Caracausi 1983). There are relatively very few adjectives or adverbial expressions, and it might be noted that lexical items in Italian dialects that derive from Arabic are significantly increased by inclusion of the exceptional dialect of the remote island of Pantelleria. The likelihood that some Arabic interferences and/or loanwords might have been introduced by migrants from the Spanish peninsula while Sicily was under Spanish rule for 400 years until 1713 compounds the problem of describing the relationship between Andalusi and Sicilian Arabic. Indeed, the same problem of interference from the Iberian Peninsula cannot be excluded from the study of Sicilian surnames derived from Arabic. Nor can one exclude the largely undocumented transmission of Arabic terms into Italianate dialects from the seventy thousand Italian, mainly Sicilian, workers who were living and working in Tunisia by the 1880s (where there was also a strong Maltese presence), or from soldiers serving in North Africa during the Second World War. Toward the end of the 20th century, a few thousand migrant workers from Tunisia have come to live and work mainly around the southern Sicilian ports, but they have had a negligible impact on the island's main dialects to date.

In spite of its obvious interest, the study of Medieval Sicilian Arabic is still in its infancy. Scholarly pioneering efforts have now recorded and classified nonspeculative examples of Arabic elements in later Sicilian dialects, and these have been accompanied by works seeking to outline the phonetic features of Sicilian Arabic and highlight resemblances to Maltese and Andalusi Arabic. Attention is now refocusing on the complex underlying problems of methodology and the need to establish reliable readings from the source material, as well as a reexamination of the wider language situation and the particular contexts in which the linguistic evidence occurs. As such, the perception of the limits and possibilities surrounding these issues is likely to undergo continuing revision.

Of written material containing some element of Arabic, excluding those merely appended with Arabic signatures or witness lists, there are 33 extant royal *dīwānī* and 22 private documents (Cusa; Johns 2002:301–325). These date from between 1095 and 1242, with the majority issued between 1133 and 1183. A few exist in fragmentary form, but almost all are in legible condition, being written mainly on durable and high-quality parchment. They consist of endowment charters, privileges, donations, decrees, writs, a draft loan agreement, deeds of sale and purchase, various letters patent (one of which is in Judaeo-Arabic), inquest proceedings, and sometimes long descriptions of boundaries (*jarāʾid al-ḥudūd*) and lists of men (*jarāʾid ar-rijāl)* who lived on crown lands. Most of the *jarāʾid* were bilingual (Arabic-Greek), although one significant and extensive boundary definition was composed in Arabic and Latin. In many cases of bilingual documents, the Greek or Latin had been translated or transcribed from the Arabic. The vast majority are available for consultation in state, regional, and church archives in Sicily, although a handful of important Arabic documents are located in the Archivo de Casa Ducal de Medinaceli in Toledo, where their availability is restricted. Almost all the Arabic charter material currently located in Sicily was published between 1868 and 1882 (Cusa, *Diplomi*). This edition, which was reprinted in 1982 without additions or corrections, contains neither proper indices nor translations. Moreover, it is well known to be riddled with errors that continue to undermine the reliability of attempts at investigating Sicilian Arabic from a detailed linguistic perspective. International projects are now underway to produce modern critical editions of all the material.

The *Biblioteca arabo-sicula* (*BAS*) contains extracts of most Medieval authors who have written about Sicily in Arabic. Very few of these authors were native to the island, and their contribution is to our understanding of Sicily's history, geography, and poetry rather than its language. There is no extant Sicilian Arabic poetry in colloquial form equivalent to the Andalusi *zajal*. A collection of the Arabic inscriptions of Sicily was originally recorded by Michele Amari (*Le epigrafi arabiche*).

A single source of *laḥn al-ʿāmma* 'mistakes of the common people' literature survives for Sicily, written by Ibn Makkī, who emigrated from the island in the second half of the 11th century. While his comments on the speech errors of the *ʿāmma* and *xāṣṣa* are ultimately inconclusive, he makes some intriguing observations relating to morphology, hypercorrections, and gender switching (Agiùs 1996:123–157). However, the force of these observations remains open to interpretation.

Many deviations from Classical Arabic norms that are found in Sicilian Arabic are quite usual for the loose scribal conventions found in Medieval Arabic administrative texts. Not uncommon examples of Sicilian Middle Arabic include inconsistent use of the relative adjective, the dual form, avoidance of double *ʾiḍāfa*, and a tendency toward analytic possessive constructions (Agiùs 1996:401–403). A very commonly attested characteristic of Sicilian Arabic was the use of noun duplication to indicate extent, e.g. *aṭ-ṭarīq aṭ-ṭarīq* 'right along the road'. The origin of this construction may be Greek, and finds parallels in contemporary Sicilian Latin as well as modern Sicilian dialects.

Given the large-scale immigration into Sicily during the Arab-Islamic period, one might reasonably expect that Sicilian Arabic would be related to whatever Arabic dialects were spoken in Aghlabid and Fatimid Ifrīqiyā, particularly in the coastal towns from where most settlers seem to have originated. However, linguistic evidence to corroborate this is minimal, conspicuous examples being *k.nīsya* for *kanīsa, zawj* 'two', and *m.tāʾ* 'belonging to' (Sgroi 1986). Several variants of the Maghribi form *bi-z-zāf* 'in excess, much' are attested in Sicilian dialect but significantly not before the 16th century. It should also be pointed out that neither Medieval documents nor later dialects offer any evidence of Maghribi aspectual markers or 1st person verb forms in Sicilian Arabic. Occasionally, there is found some → *maġribī* pointing in both royal and private documents, although this may merely indicate the provenance of a particular scribe. Moreover, since the fiscal administration (Arabic *dīwān*) of the Normans came to be based on the offices of Fatimid Cairo, it is not inconceivable that some of the scribes who worked in Sicily were from Egypt. One might note, for example, in a boundary description of crown lands, the use of *baḥrī* and *qiblī* to indicate 'north' and 'south'. Not only is this usage particularly associated with Egypt, but one also wonders whether *baḥrī* could ever have been in vernacular use on an island where every direction is necessarily 'seaward' (De Simone 1986:483–484). The possibility that non-Sicilian scribes were employed in the Norman Sicilian *dīwān* thus poses a threat to wider linguistic and diplomatic comparisons.

The exceptional importance of the Arabic material in Sicily lies in the fact that many documents are bilingual, and Arabic-Greek documents in particular offer the most reliable opportunity to reconstruct aspects of Sicilian Arabic phonology (De Simone 1979). For example, the letters *ʾalif* and *fatḥa* were consistently rendered in Greek with *epsilon* (rather than *alpha*), suggesting the strong presence of *ʾimāla* (fronting of *a*) in Sicilian Arabic. In addition, the Greek transcriptions of Arabic vowels suggest the inhibition of *ʾimāla* in the environment of the emphatic consonants. Evidence for *tafxīm* (velarization) is also common, and there is some suggestion that nasalization may have been a common feature of the island's main 12th-century dialects. For example, we find Sicilian Arabic *injāṣa* 'a pear orchard' for Classical Arabic *ʾijjāṣa*, and *ḥajjām* > χαγγέμης 'a barber', from which the modern Sicilian surname 'Cangemi' is derived. Similarly, although the Arabic letter *ḏ̣āʾ* is sometimes found represented by a Greek *delta*, rather than a *zeta*, this type of observation begs questions about the phonology of Sicilian Greek, orthographic consistency, and the route of transmission of Sicilian Arabic elements. Evidence that might have provided a means to reconstruct Sicilian Arabic stress patterns by noting where the corresponding accent was marked in Greek transliterations has proved inconsistent to date.

Sicilian Arabic is distinguished by a small number of loanwords and interferences from South Italian Greek dialects (Caracausi 1990) and Gallo-Romance (Várvaro 1981:196–204). According to some researchers, hybrid forms with Arabic and Romance elements attested in later Romance-based Sicilian dialect are evidence that Sicilian Arabic underwent varying degrees of pidginization and creolization of a type which parallel linguistic developments in Maltese (Agiùs 1996). It should be noted, however, that the vast majority of examples of this hybrid type are attested in later Romance-based Sicilian dialects or Sicilian Greek, and there are doubts about the reliability and validity of arguments that seek to infer the nature of Sicilian Arabic anachronistically from non-Arabic dialects of later periods. Recent works have raised concerns over the linguistic status of many of these hybrid forms since they are often attested in translations and transumpts written by scribes with a strong tendency toward code-switching and whose loose concepts of how to write the words of one language in another often included a capricious blend of translation and transliteration (Metcalfe 2001). This has undoubtedly cast a veil over the evidence and, as such, the wider implications of Romance and/or Greek elements attested in Sicilian Arabic itself, e.g. *al-kh.nzārī* 'pig farmer' (< *xinzīr* + Romance *-ari(us)*), remain as intriguing as they are uncertain.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

BAS = *Biblioteca arabo-sicula, ossia raccolta di testi arabici*. 2 vols. 2nd ed. Ed. Michele Amari and Umberto Rizzitano. Palermo: Edizione nazionale delle opere di Michele Amari, 1988.

Cusa, *Diplomi* = *I diplomi greci ed arabi di Sicilia, pubblicati nel testo originale, tradotti ed illustrati*. 2 vols. Palermo, 1868–1882. (Repr., Cologne and Vienna, 1982.)

Ibn Makkī, *Taṯqīf al-lisān* = Ibn Makkī, *Taṯqīf al-lisān wa-talqīḥ al-janān*. Ed. ʿAbd al-ʿAzīz Maṭar. Cairo, 1969.

Le epigrafi arabiche = *Le epigrafi arabiche di Sicilia, trascritte, tradotte, e illustrate. Parte prima: iscrizioni edili*. Palermo, 1878. (Repr., but with far fewer plates, 1971.)

SECONDARY SOURCES

Agiùs, Dionisius A. 1996. *Siculo Arabic*. London and New York: Kegan Paul International.

Caracausi, Girolamo. 1983. *Arabismi medievali di Sicilia*. Palermo: Centro di studi filologici e linguistici siciliani.

——. 1990. *Lessico greco della Sicilia e dell'Italia meridionale (secoli X–XIV)*. Palermo: Centro di studi filologici e linguistici siciliani.

De Simone, Adalgisa. 1979. *Spoglio antroponomico della giaride* (ǧarā'id) *arabo-greche nei* Diplomi *editi da Salvatore Cusa, parte I*. Roma.

——. 1986. "Su alcune corrispondenze lessicali in diplomi arabo-latini della Sicilia medievale". *Gli interscambi culturali e socio-economici fra l'Africa settentrionale e l'Europa mediterranea*, ed. Luigi Serra, 469–484. Naples.

Johns, Jeremy. 2002. *Arabic administration in Norman Sicily: The Royal* Dīwān. Cambridge: Cambridge University Press.

Metcalfe, Alex. 2001. "*De Saracenico in Latinum transferri*: Causes and effects of translation in the fiscal administration of Norman Sicily". *Al-Masāq: Islam and the Medieval Mediterranean* 13.43–86.

——. 2003. *Muslims and Christians in Norman Sicily: Arabic speakers and the end of Islam*. London and New York: RoutledgeCurzon.

Pellegrini, Giovan Battista. 1972. *Gli arabismi nelle lingue neolatine, con speciale riguardo all'Italia*. 2 vols. Brescia: Paideia editore.

Rocco, Benedetto. 1995. "Le tre lingue usate dagli ebrei in Sicilia dal sec. XII al sec. XV". *Italia Judaica* 5.355–369.

Sgroi, Salvatore. 1986. *Interferenze fonologici, morfosintattiche e lessicali fra l'arabo e il siciliano*. Palermo: Centro di Studi Filologici e Linguistici Siciliani.

Várvaro, Alberto. 1981. *Lingua e storia in Sicilia (dalle guerre puniche alla conquista normanna)*, I. Palermo: Sellerio editore.

ALEX METCALFE (Lancaster University)

Ṣifa

The term *ṣifa* lit. 'feature, attribute, property', from the root *w-ṣ-f* 'to describe', belongs to the earliest stock of Arabic grammatical terminology. In later grammar, its meaning became more or less fixed for a category of words corresponding to the adjective and the attribute in the Greco-Latin tradition, but originally it was used for a variety of meanings.

The term *ṣifa* is one of a functional pair *ṣifa/mawṣūf*, which is analogous in meaning to the terminological pairs *musnad/musnad ʾilayhi* (→ *ʾisnād*) and *muxbar bihi/muxbar ʿanhu* (→ *xabar*), as al-Fārābī (d. 339/950; *ʾAlfāḏ* 57) explains. Within the Aristotelian tradition in Arabic philosophy, the verb as the predicate par excellence was defined as a *ṣifa ġayr mawṣūf*, a predicate of which nothing can be predicated (e.g. Zajjājī, *ʾĪḍāḥ* 53; cf. Fārābī, *ʾAlfāḏ* 57; Xwārizmī, *Mafātīḥ* 142.11ff.; Ġazzālī, *Miḥakk* 23.28). According to al-Ġazzālī (d. 505/1111), *ṣifa/mawṣūf* was the preferred terminology for 'subject/predicate' in theological treatises. The concept of the verb's function as that of a predicate is not very compatible with the mainstream of Arabic grammar, however, and in the grammatical tradition, *ṣifa* is predominantly used for an attribute rather than a predicate.

In Sībawayhi's *Kitāb*, the term *ṣifa* is by far the most frequent term for adjectives or attributes (450 times; Troupeau 1976:215, who distinguishes between 296 times 'qualité' and 154 times 'qualification'); it is a synonym of the less frequently used *naʿt* (59 times; Troupeau 1976:203). The latter was known in the later tradition as a typically Kufan term (although it remained in use as a synonym for *ṣifa*; Carter 1981:239, n. 11.0), and it is indeed the preferred term for 'attribute' in al-Farrāʾ's *Maʿānī l-Qurʾān*, where *ṣifa* is used infrequently in this sense (Kinberg 1996:914–918). The use of *naʿt* by the Kufan grammarians tallies with its use in the sense of 'adjective' in the early exegetical writings (Versteegh 1993:108), which in other terms as well exhibit a decidedly Kufan slant. In these writings, *ṣifa* is not unknown, but it usually has a nontechnical meaning, for instance when it is used in referring to God's attributes (*ṣifāt Allāh*; Versteegh 1993:118).

The term *ṣifa* is used by al-Farrāʾ and other Kufan grammarians frequently, but in a different sense, that of 'prepositional phrase' or 'locative phrase', which was known in the mainstream tradition as *ḍarf* (Owens 1990:144–146; Kinberg 1996:909–914). According to Talmon (1995), this Kufan usage is a clue to its original meaning, which he claims goes back to Greek logic: the use of *ṣifa* for the prepositional phrase derived from the predicative function of these phrases. In some of the translations of Aristotle and in the *Rasāʾil ʾIxwān aṣ-ṣafāʾ*, the term is used, indeed, as equivalent of the Greek *katègoroúmenon* 'predicate' (Zimmermann 1972:530–531). In Ibn al-Muqaffaʿ's (d. 142/759) *Kitāb al-manṭiq*, the term is used both for the Aristotelian category of 'quality' (*poíon*) and for all categories together, with the exclusion of the substance (*ousía*). The latter meaning is the one that developed into that of adverbials of time and place or prepositional phrases in general, as it was used in Kufan grammar. The term *ṣifa* could even be used for the pre-positions themselves, for instance in Xalaf al-ʾAḥmar's (d. 180/796) *Muqaddima fī n-naḥw* (on this treatise and its alleged author, see Talmon 1990; Owens 1990:200–202), where *ḥurūf aṣ-ṣifa* is used for those particles that put the following noun in the genitive (*Muqaddima* 3, 43).

The replacement of the original term for 'adjective' or 'attribute', *naʿt*, with *ṣifa*, and the exclusion of prepositional phrases from the definition of *ṣifa* seem to have been innovations in Sībawayhi's *Kitāb*. In the *Kitāb al-ʿayn*, the old meaning is still visible (II, 43, 52, 246; Talmon 1995), but in the *Kitāb Sībawayhi*, although the old term *naʿt* still occurs occasionally, the preferred term has become *ṣifa*, and in this point, as in many other cases, the *Kitāb Sībawayhi* therefore marks a transition in grammatical theory and terminology. Mosel (1975) analyzes Sībawayhi's use of *ṣifa* extensively and shows that he uses it for two different functions: to denote a substantial category (adjective), and to denote a functional category (attributive). In its first function, it is used for words that fall under the following headings (Sībawayhi, *Kitāb* II, 219.9–2226.21; Mosel 1975:138ff.): (i) illnesses (e.g. *marīḍ* 'ill'); (ii) the pattern *faʿlān* (e.g. *ʿaṭšān* 'thirsty'); (iii) the pattern *ʾafʿalu* (e.g. *ʾaʿwaru* 'one-eyed'); and (iv) (moral or esthetic) properties of things (e.g. *jabān* 'cowardly').

According to Diem (1970–1971), the Arabic grammarians did not distinguish 'adjectives' as a separate category, apart from the nouns. But Mosel (1975:143) interprets Sībawayhi's classification differently. In one passage (*Kitāb* II, 5.11), Sībawayhi states that the Bedouin use some *ṣifāt* so often instead of a noun that they even use them without a noun, for instance the word *ʾaswadu* in the sense of 'black [snake]'. She shows (Mosel 1975:141) that Sībawayhi very often calls them *ṣifāt* rather than *ʾasmāʾ*, and even though he treats them syntactically as nouns, he distinguishes them from the nouns semantically.

The difference between the *ṣifāt* and the nouns is apparent in the fact that there are some constructions in which a *ṣifa* cannot be used. For instance, in the sentence *sīra ʿalayhi ṭawīlan* 'people went along it for a long time [lit. 'there was going along it for a long time']', the adjective cannot be raised to the nominative position without an additional substantive (thus *sīra ʿalayhi laylun ṭawīlun*, but **sīra ʿalayhi ṭawīlun*; Mosel 1975:149). This difference argues for a special status for the adjectives, separating them from the other nouns (Sībawayhi, *Kitāb* I, 99.13ff.; Mosel 1975:145ff.). Sībawayhi calls them *ṣifa mušabbaha* 'attribute which is being made similar to [the participle *ism al-fāʿil*]', because they share with the participles the possibility of being used in a construct state. A sentence like *huwa ḥasanu l-wajhi* 'he is beautiful of face' is structurally similar to *huwa ḍāribu zaydin* 'he is hitting Zayd'. Both phrases can be used with an article: *(huwa) aḍ-ḍāribu r-rajuli* '(he is) the one hitting the man', and *(huwa) al-ḥasanu l-wajhi* '(he is) the one with the beautiful face'. They are not completely parallel, however, because the participle serves as a predicate for the topic of the sentence, whereas in the sentence (*huwa*) *al-ḥasanu l-wajhi* the word *ḥasan* 'beautiful' serves as the underlying predicate for the word *wajh* 'face' (Mosel 1975:146).

In later grammar, adjectives remained within the class of the nouns, but from time to time, grammarians refer to their special character, either from a semantic or a morphological point of view. Ibn Kaysān (d. 299/912 or 320/932) states that the plural pattern *faʿalāt* is used with a singular *faʿla*, but only with nouns, not with adjectives (*mā kāna min al-ʾasmāʾ dūna ṣ-ṣifāt*, quoted by Ibn al-ʾAnbārī, *ʾInṣāf* 19.7).

Az-Zajjājī (*ʾĪḍāḥ* 89) uses the expression *al-ism ʾaw aṣ-ṣifa ʾaw al-laqab* ‘the noun, or the attribute, or the name’, as if these are three subclasses of the category ‘noun’.

The functional sense in which *ṣifa* is used in the *Kitāb* may be defined as that of a qualifier, qualifying a preceding noun (Mosel 1975:287–336). Its main function is to serve as an ‘ornament’ (*taḥliya*), e.g. *aṭ-ṭawīlu* ‘the tall one’, or to denote ‘kinship’ (*qarāba*), e.g. *ʾaxūka* ‘your brother’, or to emphasize the preceding noun, just like the *ʾasmāʾ mubhama*, e.g. *marartu bihim kullihim* ‘I passed all of them’ and *hāḏā* ‘this one’ (Sībawayhi, *Kitāb* I, 223.6ff.; Mosel 1975:292).

The relationship between *ṣifa* and *mawṣūf* is similar to that between the → *ṣila* and the *mawṣūl* in one respect: the *ṣifa* constitutes a ‘completion’ (*tatimma*) of the noun to which it belongs, just like the *mawṣul* completes the *ṣila*. Unlike the *mawṣul*, however, it is syntactically dispensable (Mosel 1975:292–293). After a *ṣila*, for instance the relative pronoun *allaḏī*, there must be a *mawṣūl* in the form of a relative phrase, whereas the *mawṣūf* may or may not be completed by a *ṣifa*, even when this *ṣifa* is necessary in a pragmatic or semantic sense.

The attributive adjective agrees with the word it qualifies in case. Sībawayhi (*Kitāb* I, 210.1ff.; Mosel 1975:293ff.) simply states that it receives the same case because it is identical with the word to which it is attributed. In later syntax, the attribute was regarded as one of the *tawābiʿ*, i.e. words whose case assignment is caused by their ‘following’ another word (*tabʿiyya*; Carter 1981:238–274). Owens (1988: 154–156) calls this category ‘modifiers’, which include the attributive adjective, the permutative (→ apposition), the emphasizer (*taʾkīd*), the explicative (*ʿaṭf bayān*), and the conjunct (*ʿaṭf nasq*). This category created a special problem in the dependency framework of Arabic grammar, since their case assignment could not be explained by governance (→ *ʿamal*).

The general test of whether a word can be an attribute is when it can function as the noun’s predicate (*mabnī ʿalā l-mubtadaʾ*; Sībawayhi, *Kitāb* I, 275.24; Mosel 1975:294ff.): if it is possible to say *zaydun (huwa) ṭ-ṭawīlu* ‘Zayd is the tall one’, *aṭ-ṭawīlu* may also function as its attribute (*zaydun aṭ-ṭawīlu* ‘Zayd, the tall one’). The following categories of words can serve as *ṣifa* (for a complete list and extensive commentary, see Mosel 1975:295ff.):

i. active and passive participles, e.g. *zaydun aḍ-ḍāribu* ‘Zayd, the one who hits’
ii. *ṣifa mušabbaha*, e.g. *zaydun aṭ-ṭawīlu* ‘Zayd, the tall one’
iii. other nouns which do not belong to i or ii
iv. verbs, e.g. *rajulun ḍarabtuhu* ‘a man I hit’
v. prepositional phrases, e.g. *kitābun laka* ‘a book belonging to you’
vi. kinship terms, e.g. *zaydun ʾaxūka* ‘Zayd, your brother’
vii. *ism mubham*, e.g. *zaydun hāḏā* ‘this Zayd’ (in *hāḏā r-rajulu* ‘this man’, it is the noun that serves as attribute to the demonstrative)
viii. *ʿalāmat al-muḍmar*, e.g. *kulluhum* ‘all’ or *nafs* ‘self’
ix. numerals, e.g. *miʾatun* ‘one hundred’
x. *ʾillā*, e.g. *ar-rijālu ʾillā zaydun* ‘the men, except Zayd’ (→ *istiṯnāʾ*)

The function of the *ṣifa* as a constituent is clear in constructions like *marartu bi-rajulin ḥasanin ʾabūhu* ‘I passed a man whose father was good’. In this sentence, *ḥasan* is an attribute of *ʾabūhu*, but at the same time it is *ṣifa* of *rajulin*, as evidenced by the fact that it has the same case ending (Mosel 1975:296ff.). In later grammar, this construction was called *naʿt sababī*, a semantically linked attribute (Carter 1981:245ff.; → *sabab*).

The term *waṣf* is used synonymously with *ṣifa*, but may also be used in a more general, descriptive sense, for instance for the circumstantial clause (→ *ḥāl*) in *sīra ʿalayhi šadīdan* lit. ‘there was traveling along it in an intensive manner’: here, *šadīdan* is said to describe the mode of the traveling (Sībawayhi, *Kitāb* I, 116.17; Mosel 1975:288).

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Fārābī, *ʾAlfāḏ̣* = ʾAbū Naṣr Muḥammad ibn Muḥammad al-Fārābī, *Kitāb al-ʾalfāḏ̣ al-mustaʿmala fī l-manṭiq*. Ed. Muḥsin Mahdī. Beirut: Dār al-Mašriq, 1968.

Farrāʾ, *Maʿānī* = ʾAbū Zakariyyāʾ Yaḥyā ibn Ziyād al-Farrāʾ, *Maʿānī l-Qurʾān*. Ed. Muḥammad ʿAlī an-Najjār. 3 vols. Cairo: ad-Dār al-Miṣriyya, 1966–1972.

Ġazzālī, *Miḥakk* = ʾAbū Ḥāmid Muḥammad ibn Muḥammad al-Ġazzālī, *Kitāb miḥakk an-naḏ̣ar fī l-manṭiq*. Ed. Muḥammad Badr ad-Dīn ʾAbū Firās an-Naʿsānī al-Ḥalabī and Muṣṭafā al-Qabānī ad-Dimašqī. Cairo: al-Maṭbaʿa al-ʾAdabiyya, n.d.

Ibn al-ʾAnbārī, *ʾInṣāf* = ʾAbū l-Barakāt ʿAbd ar-Raḥmān ibn Muḥammad al-ʾAnbārī, *Kitāb al-ʾinṣāf fī masāʾil al-xilāf bayna n-naḥwiyyīna l-Baṣriyyīna wa-l-Kūfiyyīn*. Ed. Gotthold Weil. Leiden: E.J. Brill, 1913.

Ibn al-Muqaffaʿ, *Manṭiq* = ʾAbū Muḥammad (ʾAbū ʿAmr) ʿAbdallāh Ibn al-Muqaffaʿ, *al-Manṭiq*. Ed. Muḥammad Taqī Danešpažuh. Teheran: Imperial Iranian Academy of Philosophy, 1398/1978.

Ibn as-Sarrāj, *ʾUṣūl* = ʾAbū Bakr ibn as-Sarī Ibn as-Sarrāj, *Kitāb al-ʾuṣūl fī n-naḥw*. Ed. ʿAbd al-Ḥusayn al-Fatlī. 3 vols. Beirut: Muʾassasat ar-Risāla, 1985.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. 2 vols. Ed. Bulaq, 1316 A.H. (Repr., Baghdad: Al-Muthanna Library, n.d.)

Xalaf, *Muqaddima* = Xalaf al-ʾAḥmar, *Muqaddima fī n-naḥw*. Ed. ʿIzz ad-Dīn at-Tanūxī. Damascus, 1961.

Xwārizmī, *Mafātīḥ* = ʾAbū ʿAbdallāh Muḥammad ibn ʾAḥmad al-Xwārizmī, *Kitāb mafātīḥ al-ʿulūm*. Ed. Gerlof van Vloten. Leiden: E.J. Brill, 1895.

Zajjājī, *ʾĪḍāḥ* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān ibn ʾIsḥāq az-Zajjājī, *al-ʾĪḍāḥ fī ʿilal an-naḥw*. Ed. Māzin al-Mubārak. Cairo: Dār al-ʿUrūba, 1959.

SECONDARY SOURCES

Carter, Michael G. 1981. *Arab linguistics: An introductory classical text with translation and notes*. Amsterdam: J. Benjamins.

Diem, Werner. 1970–1971. “Nomen, Substantiv und Adjektiv bei den arabischen Grammatikern”. *Oriens* 23–24.312–332.

Kinberg, Naphtali. 1996. *A lexicon of al-Farrāʾ's terminology in his* Qurʾān *commentary with full definitions, English summaries, and extensive citations*. Leiden: E.J. Brill.

Mosel, Ulrike. 1975. *Die syntaktische Terminologie bei Sībawaih*. Ph.D. diss., University of Munich.

Owens, Jonathan. 1988. *The foundations of grammar: An introduction to medieval Arabic grammatical theory*. Amsterdam and Philadelphia: J. Benjamins.

——. 1990. *Early Arabic grammatical theory: Heterogeneity and standardization*. Amsterdam and Philadelphia: J. Benjamins.

Talmon, Rafael. 1990. “Kitāb *Muqaddima fī n-naḥw* al-mansūb ʾilā Xalaf al-ʾAḥmar: Dirāsa wa-muṣṭalaḥāt”. *al-Karmil* 11.129–199.

——. 1995. “*Ṣifa*”. *Encyclopaedia of Islam* IX, 551. 2nd ed. Leiden: E.J. Brill.

Troupeau, Gérard. 1976. *Lexique-index du* Kitāb *de Sībawayhi*. Paris: Klincksieck.

Versteegh, Kees. 1993. *Arabic grammar and Qurʾānic exegesis in early Islam*. Leiden: E.J. Brill.

Zimmermann, F.W. 1972. “Some observations on al-Fārābī and logical tradition”. *Islamic philosophy and the Classical tradition: Essays presented by his friends and pupils to Richard Walzer on his seventieth birthday*, ed. Samuel M. Stern, Albert Hourani, and Vivian Brown, 517–546. Oxford: Cassirer.

KEES VERSTEEGH (University of Nijmegen)

Sign Languages

Sign languages are visual-gestural languages that are produced by hand movements, facial expressions, and head/body postures and are perceived by the eyes. They are natural human languages that have arisen wherever Deaf people have come together in communities to meet their communicative needs through signing. Sign languages have complex structures at all levels of linguistic organization, with lexicon and grammar being independent of and different from the spoken languages used in the same region. In the Arab world, several sign languages and sign language dialects are in use in Deaf communities. They are distinct from the surrounding varieties of spoken Arabic, and their linguistic boundaries do not coincide with the political borders of the region.

1. DEAF COMMUNITIES AND SIGN LANGUAGES IN THE ARAB WORLD

Very little is known about the signed languages used in the Arab world and particularly in North Africa. Even the number of deaf people living in Arab countries is unclear, as widely differing statistics are used. Some sources speak of two million hearing-impaired Arabs in Egypt alone (Brelje 1999:72–73), but this figure probably includes everyone with a hearing loss, including those who have lost their hearing as a result of old age. The same source gives a more realistic number of around ten thousand deaf people (0.27 percent of the population) in Lebanon (Brelje 1999:224). The *Gallaudet encyclopedia* states that in Israel “the overall incidence of deafness...in the population up to 18 years of age is about 1.2 per 1000”, but among minorities, like the Druze, the Bedouin, and the general Arab population, the incidence of deafness is higher. This is probably due to the high incidence of consanguineous marriages among these groups (Van Cleve 1987: I, 102). The incidence of deafness is influenced by the standard of health care, since good health care reduces prevalent medical causes of deafness, like rubella and otitis media. Since Lebanon and Israel have better health care than many other Arab countries, we may assume that countries like Egypt have an even higher incidence of deafness. An estimate of 0.4 percent of the population of the Arab world as a whole would mean that there are between one and two million deaf people in this part of the world. In comparison, the incidence of deafness in Western countries is around 0.05–0.1 percent of the population.

Contrary to a common misunderstanding among hearing people, there are many signed languages across the world rather than one universal sign language, and there are well-developed indigenous sign languages in the Arab world. Although detailed survey work into the different dialects and possibly the different languages still needs to be done, it would seem that there is one common sign language (albeit with dialectal differences) in Lebanon, Palestine, Syria, and Jordan. This could be called Levantine Arabic Sign Language. Iraqi Sign Language is closely related to this language, but it may be too distinct to call it a dialect. The sign languages of the Gulf seem to be influenced by Egyptian Sign Language because many teachers of the deaf there are Egyptian, but their roots are closer to the sign language of the Levant area. Around 50 percent of the vocabulary of Yemeni Sign Language is similar to that of the Levant even though it is clearly influenced by Egyptian Sign Language. Egyptian Sign Language itself is quite distinct from the sign languages in the Levant, though still related. Very little is known about any of the other sign languages of North Africa.

To what extent these varieties are different branches of the same language family or just different dialects of the same language still needs to be investigated. An initial approximation to the relationship between sign language varieties in the Arab world can be gleaned from a comparison of basic vocabulary items across sign language varieties of the region, as is shown in Table 1. The 185 vocabulary items are drawn from a list of frequently occurring signs; signs are omitted that are clearly iconic in order to avoid an overrepresentation of similar/same signs that might be due to shared iconicity rather than common origin or contact between the sign varieties (see sec. 4 about iconicity). Table 1 confirms the above characterization of Arab sign language varieties, that is, varieties of the Levant are more similar to each other than to other varieties, with Iraq, Egypt, and Yemen representing distinct groupings.

As a control, Brazilian Sign Language is included in the table as representing a language clearly unrelated to any of the Arab sign varieties. There are some indications, including vocabulary similarities, of possible historical contact between Turkish Sign Language and Arab sign language varieties, and accordingly, the table does show a higher percentage of overlap for Turkish Sign Language and Arab sign varieties in comparison with the Brazilian control items. However, on structural grounds, Turkish Sign Language is clearly a separate language, and so is Israeli Sign Language. In spite of considerable variation in vocabulary across these urban sign languages in the Arab world, the sign varieties appear to be very similar grammatically, and this is a significant factor that requires further study.

Another case apart is the existence of village-based sign languages, as recently documented in an Arab Bedouin tribe in Israel (Sandler a.o. 2005). It seems that individual villages with a high incidence of hereditary deafness give rise to localized sign languages that are in principle independent of the urban sign language varieties in the region. There are probably several such ‘Deaf villages’ in the Arab world, but the linguistic status and structural characteristics of these sign languages are mostly undocumented to date.

2. The sociolinguistic situation

Not all Arabs who are born with a significant hearing loss automatically learn a signed language. Parents of deaf children do not generally learn the local sign language, and deaf children tend to live their first years in a communicative vacuum. Their communication may well consist only of a few ‘homemade signs’ understood by their immediate family but not forming a true language. Learning the spoken language by means of lipreading and speech therapy, if viable at all, takes up a lot of time, money, and effort and requires very intense training, which is not usually available in the Arab world.

With one or two exceptions, there was no education for the deaf in the Arab world until the second half of the 20th century. In Saudi Arabia, for example, the first two specialized institutes for the education of deaf children were opened in 1964. In 1989, more than two

Table 1. Vocabulary comparison of sign languages

185 words (non-iconic)	1	2	3	4	5	6	7	8	9	10	11	12	13	14
	Jordan-Salt1	Jordan-Salt2	Jordan-Salt3	Egypt-alMinya	Egypt-Cairo	Yemen-Aden	Yemen-Sana'a	Yemen-Hadramawt	Syria-Aleppo	Iraq-Baghdad1	Iraq-Baghdad2	Jordan-Amman	Turkey	Brazil
Jordan-Salt1	x													
Jordan-Salt2	94%	x												
Jordan-Salt3	95%	93%	x											
Egypt-alMinya	37%	39%	39%	x										
Egypt-Cairo	36%	37%	38%	61%	x									
Yemen-Aden	43%	45%	44%	36%	42%	x								
Yemen-Sana'a	42%	44%	43%	43%	49%	64%	x							
Yemen-Hadramawt	42%	42%	43%	47%	45%	67%	66%	x						
Syria-Aleppo	61%	60%	61%	35%	43%	38%	41%	40%	x					
Iraq-Baghdad1	51%	53%	53%	32%	34%	37%	41%	40%	51%	x				
Iraq-Baghdad2	53%	53%	53%	33%	38%	36%	39%	37%	53%	62%	x			
Jordan-Amman	74%	73%	74%	36%	37%	42%	43%	46%	60%	53%	52%	x		
Turkey	25%	25%	24%	19%	18%	20%	16%	18%	24%	21%	25%	21%	x	
Brazil	13%	13%	13%	5%	7%	10%	9%	10%	11%	9%	10%	15%	11%	x

thousand students were enrolled in 14 institutes for the deaf (Al-Muslat 1994:277). With a population of more than 25 million and an estimated number of more than one hundred thousand deaf people, it may be clear that this number represents only a small, privileged group of deaf children in Saudi Arabia. Similar situations exist in other Arab countries.

Education for deaf children in government day schools is usually oral, sometimes with some supporting signs, and focuses on lipreading and voicing words, reading, writing, and arithmetic. Without good methods and with very little communication between children and teachers, even these skills, however, are not learned well, and therefore most of the relatively small number of deaf people who attended school are functionally illiterate. An additional problem is the difference between written Standard Arabic and the local spoken dialects. Those deaf people who can write tend to write the Arabic words that are used in the local dialect but not normally written. In most countries, there is no secondary education for the deaf, and vocational training for deaf people is almost exclusively geared toward manual skills, such as carpentry or car mechanics for men and needlework for women. Advanced education, at college or university level, has only just become available for some deaf students in Jordan, but is impossible in most other Arab countries, partly due to the lack of trained sign language interpreters.

Because deaf education in the Arab world is mostly oral and signed languages are viewed by many people, including some deaf people, as substandard, some deaf people refuse to use the signed language, because they regard it

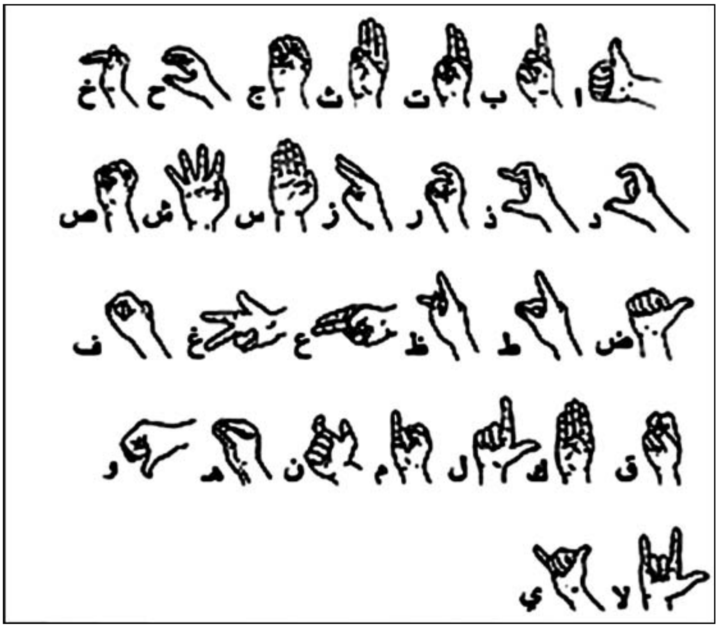


Figure 1. The Arabic finger-spelling alphabet as used in Jordan (from right to left).

as inferior compared to the spoken language. Even Deaf people who do use sign language do not generally realize it is a real language with its own grammar, and they may look upon a hearing person's sign language as more 'standard' than their own. Deaf people sometimes distinguish between 'hearing signing' and 'deaf signing' and view the latter as their own slang. When signing to a hearing person, they may modify their own sign language to become more 'hearing', without realizing that this makes it less well formed or grammatical. That this also affects the way they view language in general is clear when they distinguish 'hearing Arabic' (which is grammatically well-formed Arabic) from 'deaf Arabic' (usually a word-by-word translation of their sign language).

In some Arab countries a finger-spelling alphabet is used to represent Arabic letters (see Fig. 1). The alphabet below is used in several Middle Eastern countries with some variations and is mostly used to spell names and unfamiliar words. In some other countries, an alphabet that is based on the sounds of the Arabic language is used by Deaf people. The finger-spelling alphabet is not an integral part of any signed language but provides a visual way of representing an oral language without the need of paper. It can therefore only be used by the more educated Deaf Arabs.

Because Deaf people who have learned the local sign language can communicate freely with each other and will always have problems understanding those who do not know sign language, they tend to stick together and form a close-knit community of their own. In the Middle East, as in other parts of the world, there are many Deaf clubs, where Deaf people mix and talk together. Many Deaf people marry other Deaf and have Deaf friends. Thus, the Deaf form a subculture, with their own language, their own humor, their own problems, and their own values. Because of the high number of Deaf people in the Middle East, this community is quite strong. Deaf culture is very visual and focuses on physical characteristics (most sign names of Deaf people are based on physical characteristics, like a scar or a certain haircut). Although in most cases hearing people who have learned the local sign language well are welcomed with open arms, in some cases they may be viewed as intruders who want to take advantage of the Deaf.

The indigenous Arab sign languages are not officially recognized as minority languages in any Arab country. There is no interpreter certification and there are no government-sponsored interpreter services for the deaf. This means that for visits to a physician or lawyer, or in any meetings with hearing people, Deaf people usually have to rely on friends or relatives to interpret for them. In some cases Deaf clubs have interpreters that are available to their members. Lack of training in sign language skills and proper interpreter ethics are a big problem. In some Arab countries, like Lebanon, Jordan, Yemen, and Egypt, sign language interpretation is provided for news bulletins and sometimes for the televised mosque services on Friday. In Jordan, some public universities now have sign language interpreting for their deaf students.

Recently, there has been an attempt at creating a unified Arabic Sign Language. The idea behind this project was that a standard language was needed by deaf Arabs in the same way that Modern Standard Arabic functions as a standard language among all hearing Arabs. However, the project was not informed by linguistic considerations and documented facts about the sign languages in the region. As a result, the 'unified Arabic Sign Language' is merely a list of signs compiled from different Arab sign languages in an artificial and communicatively unacceptable way. However, the signs from the unified Arabic Sign Language dictionary are used on pan-Arab television channels, like Al-Jazeera.

Very little has been published on the signed languages in the Arab world. For most countries, all that is available is a 'dictionary' that gives a picture of a sign and an Arabic (and in some cases also English) gloss of the sign. Often there are no more than about one thousand signs in these word lists. An introductory grammar has only recently been published for Jordanian Sign Language (Hendriks 2004).

3. General comments on linguistic structure

Sign languages in the Arab world are not Semitic languages, and their grammatical structures do not have much in common with the spoken languages of the region. Most sign language varieties in the Arab world are severely underdocumented. This is true in particular of village-based sign languages but also, to a lesser extent, of urban varieties.

Urban sign languages in the Arab world share a number of structural characteristics with sign languages in other parts of the world, but differ from the latter in other respects. The description of linguistic structures in the following sections is based on Levantine Arabic Sign Language (LASL), in particular the dialects used in Jordan and in Lebanon. It is not known to what extent other sign language varieties in the Arab world are similar to or different from Levantine Arabic Sign Language, since the grammar of other varieties is undocumented.

Although Levantine Arabic Sign Language is not genetically related to spoken Arabic, influence from conventional hand and head gestures used in the region can be seen in a limited number of signs, for example the signs for *šū* 'what?' and *ʕamal maʕrūf* 'please' (see Figs. 2 and 3), as well as a backward head tilt used for negation (see Fig. 22).

4. Lexicon

All sign languages have an extensive vocabulary with thousands of words (signs), as well as specialized strategies for creating new words. One important characteristic of signs is that they are much more often iconic than words in spoken languages, that is, the form of signs often resembles their meaning in one way or another. This is comparable to onomatopoeia and → sound symbolism in spoken languages, e.g. in English words such as *cuckoo, splash, roar*, etc., but since there is so much more potential for iconicity in the visual modality, a nonarbitrary relationship between form and meaning

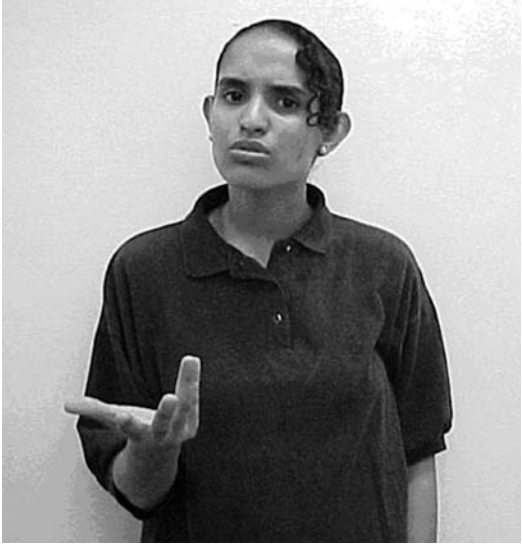


Figure 2. WHAT.

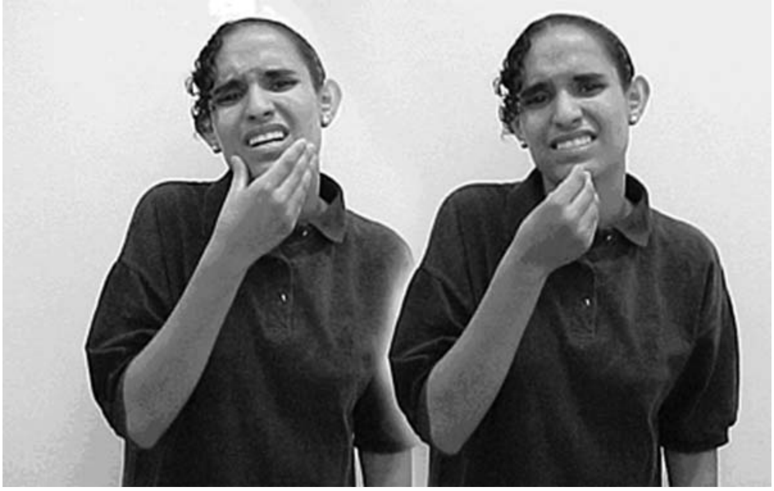


Figure 3. PLEASE.

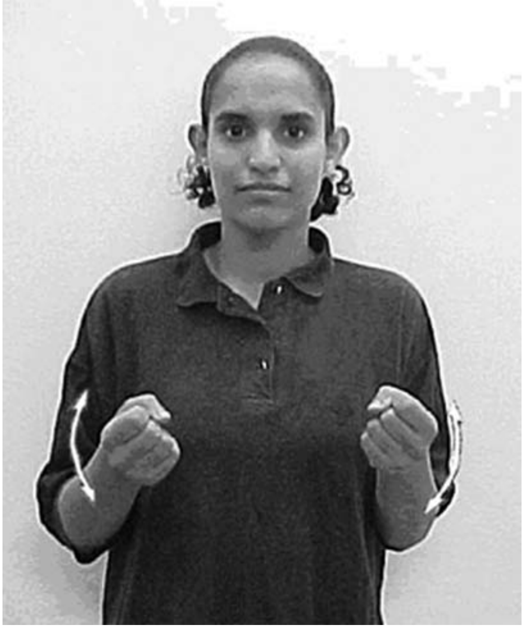


Figure 4. CAR.

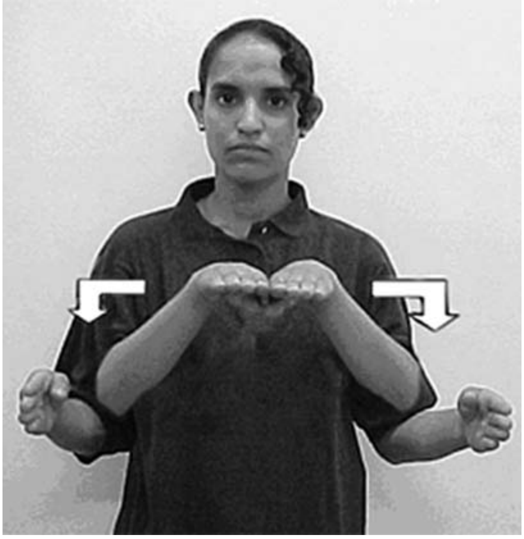


Figure 5. TABLE.

is much more common in sign languages (see Figs. 4 and 5). However, even fully or partially iconic signs are completely conventional, and it is this conventionality within a language community that is crucial for the status of a linguistic signal. In Levantine Arabic Sign Language, as used in Jordan, it has been found that only about 15–20 percent of signs are transparently iconic, that is, directly recognizable by nonsigners, and about 50 percent of signs are partially iconic, so that nonsigners can recognize the iconic connection after they have been told the meaning of the sign (Hendriks 2004:25).

The lexicon of sign language varieties in the Arab world is structured differently from and independently of the coexisting spoken languages. This can be seen clearly in a number of basic lexical domains, such as pronouns, numbers, color, and kinship terms. Table 2 summarizes some of the properties of lexicon organization in Levantine Arabic Sign Language, as compared to the spoken Arabic of the region.

Table 2. Comparison between spoken Arabic and Levantine Arabic Sign Language

Domain	Spoken Arabic	Levantine Arabic Sign Language
pronouns	paradigms of free personal pronouns and relative pronouns, suffixing for possessive pronouns number distinctions: singular, dual, plural gender distinctions: masculine and feminine	paradigms of free personal and emphatic/reflexive/possessive pronouns, no relative pronouns, no suffixing number distinctions: singular, dual, trial, quadruple, quintuple, plural gender distinctions: none
numbers	multiples of 10 are morphologically derived by adding a suffix (sequential morphology) special dual forms with the dual suffix for 20, 200, 2000	most multiples of 10 are morphologically derived by adding a side-to-side movement (simultaneous morphology) no special dual forms
color	most color words have the same morphological template	no morphological relationship between color words
kinship	no gender-neutral kinship terms, but several pairs of a basic masculine term and a derived feminine term with a suffix	gender-neutral terms for most kinship relationships, compounded with a sign for the gender (e.g. FEMALE SIBLING "sister")

5. SEQUENTIAL AND SIMULTANEOUS MORPHOLOGY

All sign languages that have been documented so far display a preference for a particular type of morphological organization that is significantly different from spoken languages. In spoken languages, the predominant type of morphology is sequential in nature, including compounding, cliticization, and, most commonly, affixation (prefixes and suffixes). Cases where a morpheme is realized simultaneously with its stem, such as changing a tone to indicate past tense, or changing vowel quality in ablaut, are comparatively rare. Sign languages show exactly the opposite pattern. Thus, in the Jordanian variety of Levantine Arabic Sign Language, there is little evidence of sequential morphology other than a negative affix (see Fig. 6) and a limited amount of compounding. As in other sign languages, various patterns of simultaneous morphology predominate which may involve both the hands and the nonmanual articulators.

The signs in Figures 7 and 8 are instances of numeral incorporation, a sign type that is common in most sign languages. The sign consists of a handshape indicating a number, and the rest of the sign indicating a unit of quantification such as time concepts (year, week, minute) or monetary units. Both elements are produced at the same time, forming a single complex sign.

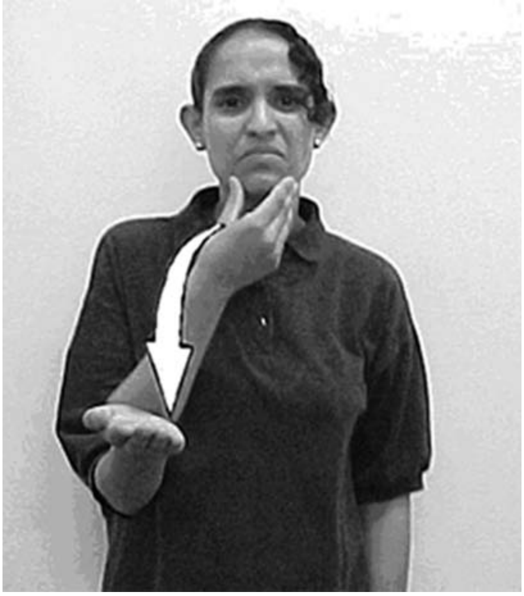


Figure 6. Negative affix in Levantine Arabic Sign Language.

Another important process is found in the domain of aspect marking. Like most sign languages, Levantine Arabic Sign Language has

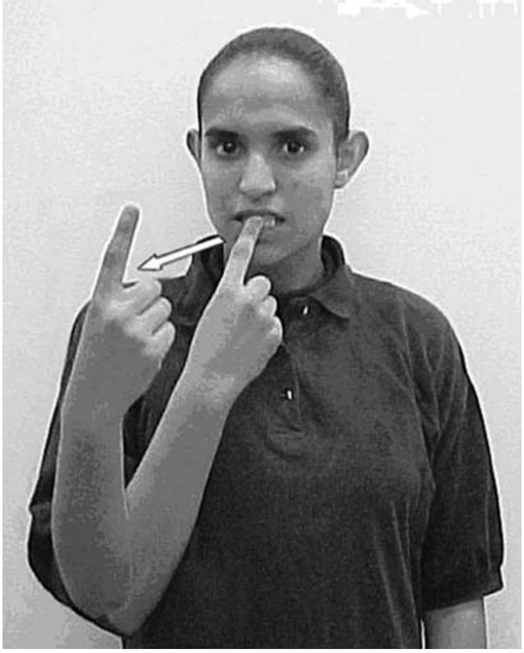


Figure 7. ONE-YEAR.

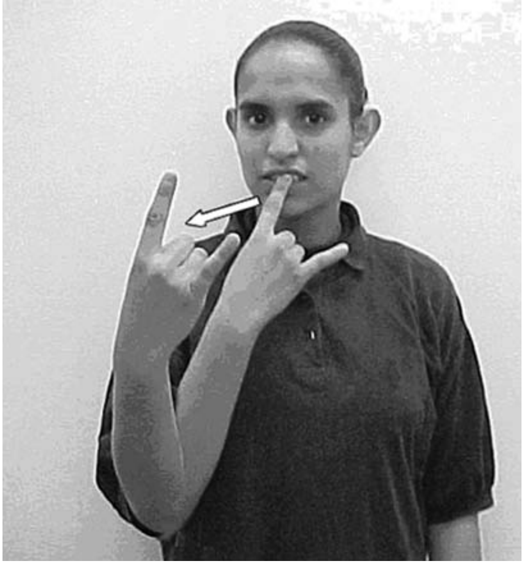


Figure 8. TEN-YEAR.

no grammatical category of tense. Time is indicated by individual lexemes at the beginning of a discourse paragraph, and a spatial metaphor ('time line') is used in this subsystem. The time line is an imaginary line running through the signer's body from back to front, where the past is located behind the signer and the future is located in front (see Fig. 9). Aspect marking, on the other hand, involves morphologically complex forms. A basic sign can occur with a number of different movement patterns to indicate repeated action, spatially distributed action, reciprocal action, and the like (see Figs. 10 and 11). Although this process is in some ways akin to the templatic morphology commonly found in Semitic languages, the expression of tense and aspect in Levantine Arabic Sign Language is in itself not at all similar to any variety of Arabic.

Movement patterns are also important for differentiating between related pairs of signs in Levantine Arabic Sign Language, where the first sign has a verbal and the second sign a nominal reading. In the various patterns, the nominal signs are usually characterized by restrained movement, sometimes with repetition of movement. Semantically, the most common pattern indicates an object on the one hand and an action involving that object on the other hand, e.g. 'light' – 'to turn on light', 'boat' – 'to go by boat', 'medicine' – 'to take medicine', etc.

Several processes of simultaneous morphology crucially depend on the use of the sign space for grammatical purposes (see Sec. 6). Moreover, the grammatical use of facial expressions (see Sec. 7) also occurs simultaneously with manual signs.

6. THE USE OF SPACE IN SIGN LANGUAGE GRAMMAR

All sign languages documented so far use the sign space, that is, the space around the signer's torso, for grammatical purposes. The most important aspects of the grammatical use of space include the localization of discourse referents, the spatial marking of subject and object, and the use of complex classificatory constructions.

In a signed text, persons, places, and objects are often associated with a particular location in the sign space. This process is known as localization and can be achieved in various ways, for example by signing a proper noun in combination with the index finger pointing at a certain location in the sign space. The signer can then refer back to the noun by directing a pronoun or a spatial agreement verb toward this location. In this way, signers construct complex spatial layouts in front of themselves that are consistently used within a discourse to

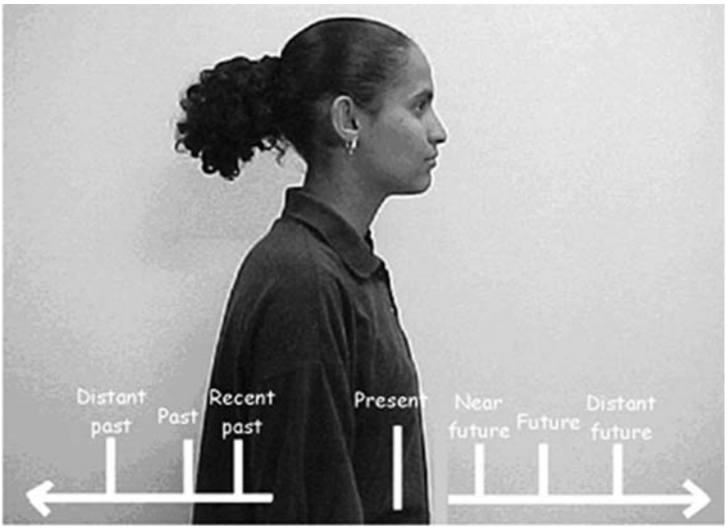


Figure 9. Time line.

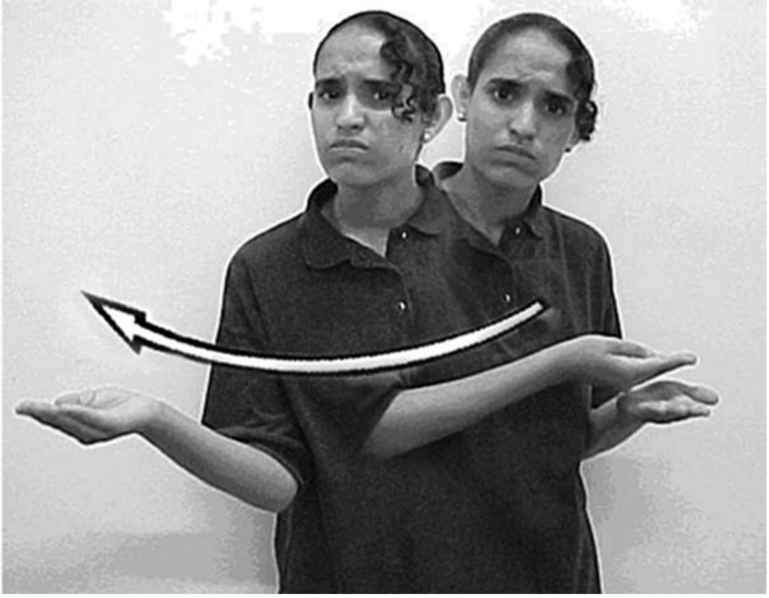


Figure 10. GIVE-ALL.

keep track of the different discourse referents. Figure 12 shows the signing space with the signer (no. 1) and the addressee (no. 2). The signer is pointing at a location in the sign space representing an absent third person (no. 3).

One of the most important uses of the sign space is the expression of subject-object relationships in agreement verbs. These are morphologically complex verbs that change movement direction to show who is doing what to whom. These signs usually begin at the subject or 'source' location and move toward the object or 'goal' location (see Figs. 13 and 14). Levantine Arabic Sign Language has at least fifty transitive verbs that use movement for indicating subject-object relationships. This grammatical mechanism closely interacts with the more general principle of localization, since it depends on associating discourse referents with locations in the sign space. Spatial agreement verbs are in some respects similar to multiple person marking on verbs where bound

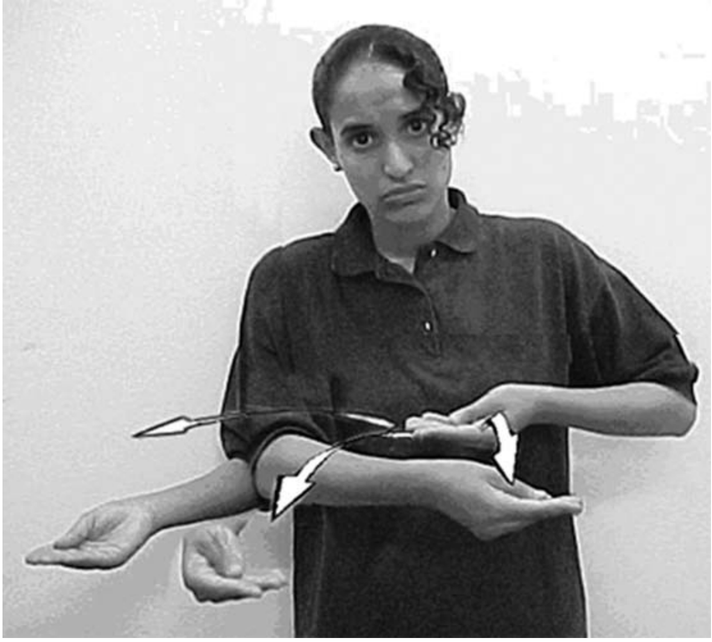


Figure 11. GIVE-EACH.

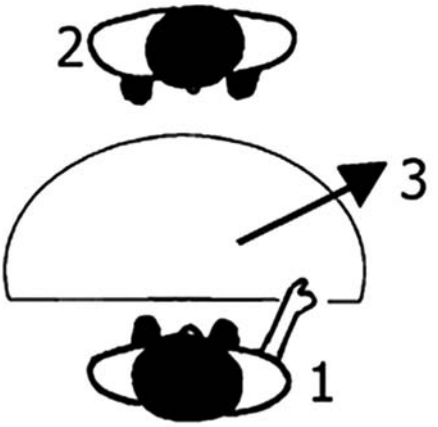


Figure 12. Localizing an absent third person in the sign space.

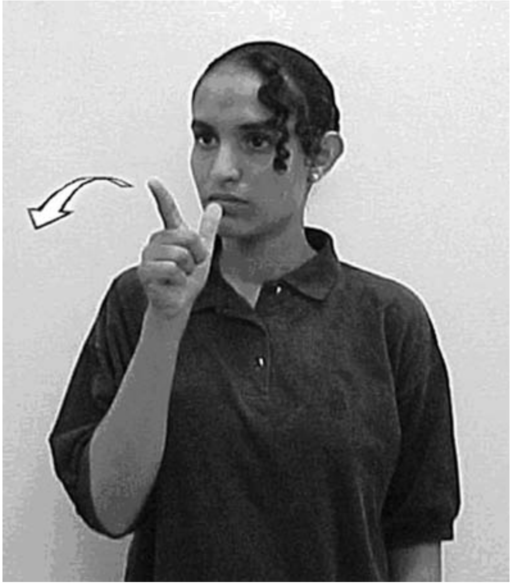


Figure 13. TELL-YOU.

pronouns represent subject and object (cf. Arabic *ya-s'alu-nī* 'he-ask-me'), except that the 'pronouns' in spatial agreement verbs consist of locations in sign space rather than spoken syllables.

The most complex uses of sign space are associated with so-called classifier constructions (for classifiers in a Jordanian sign language dialect, see van Dijken 2004). In these constructions, a salient feature of a class of referents is mapped onto the shape of the hand, while movement and location of the hand represent the real-world movement and spatial arrangement of the referent. The handshape can represent the spatial movement and position of entities such as vehicles, persons, or animals (see Figs. 15 and 16), or can be related to the geometrical properties of objects. In the latter case, objects are represented either in terms of their inherent size and shape properties or in terms of the way

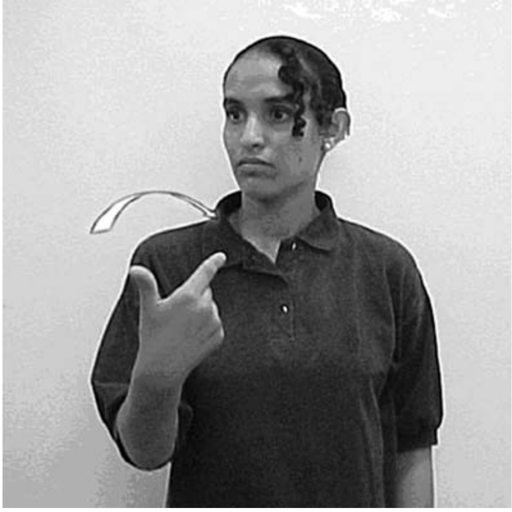


Figure 14. TELL-ME.

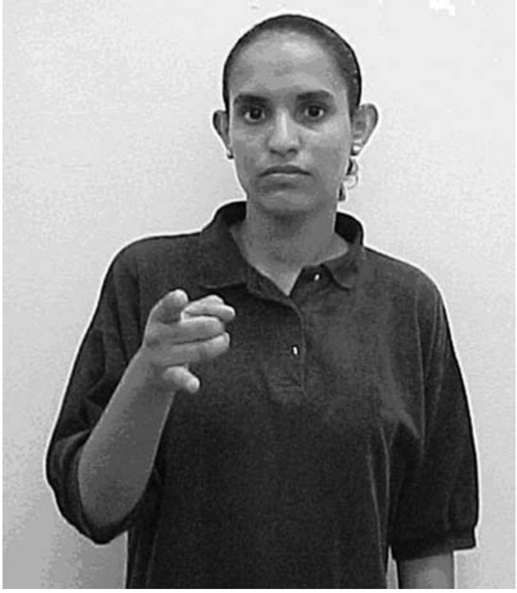


Figure 15. Classifier for VEHICLE.

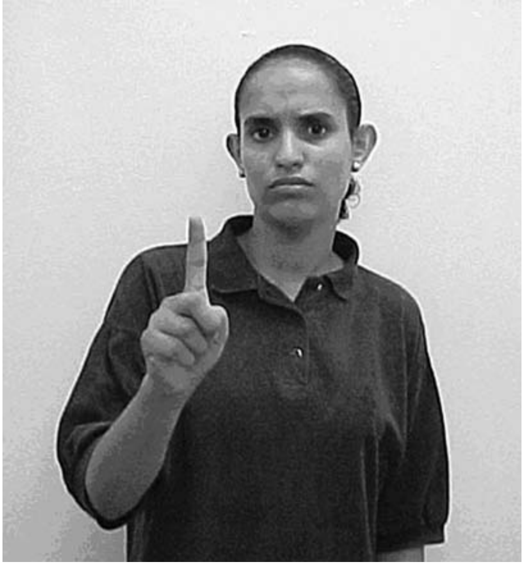


Figure 16. Classifier for PERSON.

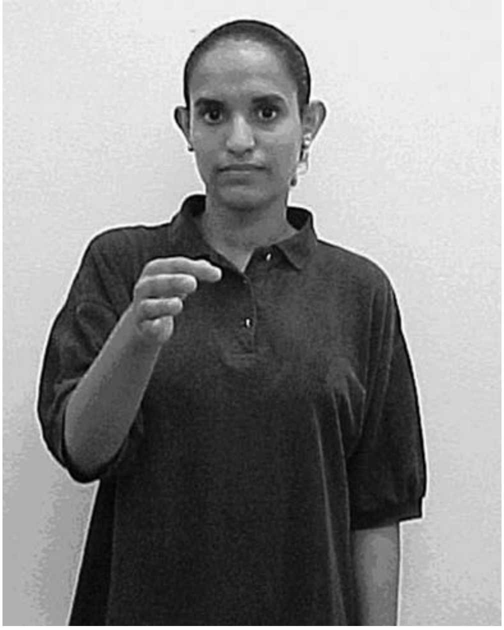


Figure 17. GIVE-FLOWER.

they are handled by a person (see Figs. 17 and 18). The possibility of mapping a real-world situation onto the hands in conventionalized ways allows for a large number of very productive combinations, especially when two-handed combinations are used.

7. Nonmanual aspects of sign language grammar

Sign languages do not only use the hands to code linguistic information. Nonmanual aspects of signing also contribute significantly to sign language grammar, with head movements and facial expressions as the most important features. Nonmanual behaviors in sign languages are functionally equivalent to the use of intonation in spoken languages. They are suprasegmental and can spread over one or more manual signs, with this co-occurrence indicating the scope of the nonmanual signal. Two major uses of nonmanual behaviors include clause typing and adverbial functions.

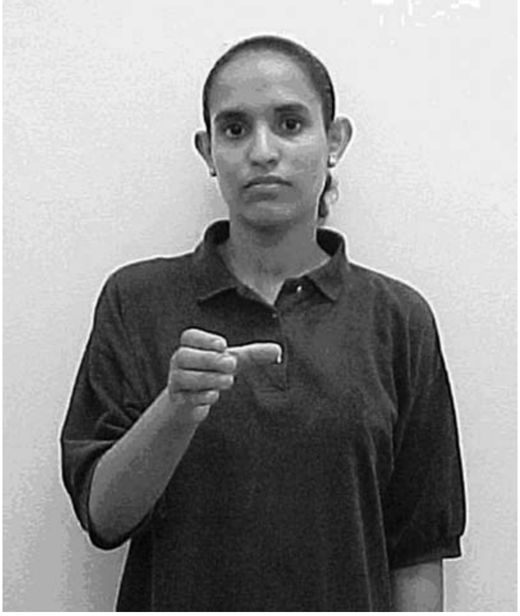


Figure 18. GIVE-BUNCH-OF-FLOWERS.

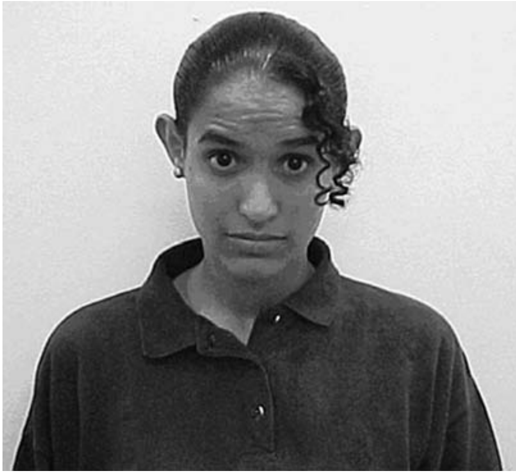


Figure 19. Nonmanual marking of yes-no questions in Levantine Arabic Sign Language.

As in other sign languages, various clause types are marked by a particular nonmanual configuration in Levantine Arabic Sign Language. This includes various types of questions, negation, imperatives, and conditional clauses, as well as topicalization.

As no question particles exist in Levantine Arabic Sign Language, yes/no questions are typically marked nonmanually only, with raised eyebrows, wide open eyes, and the head tilted forward (see Fig. 19). The facial expression used in WH-questions is more variable and partly depends on the context of the utterance. The most basic question word is the sign WHAT (Fig. 2). Different dialects of Levantine Arabic Sign Language differ with respect to the paradigm of additional question words. The Jordanian variety in Salt has separate signs for 'where?', 'who?', 'when?', and 'why?', the latter being the same sign as the sign for 'reason' (Fig. 20). Other dialects have fewer question words, subsuming them under a general interrogative.

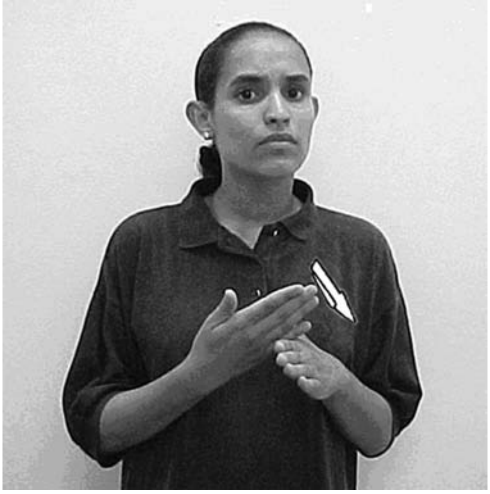


Figure 20. REASON/WHY.

The interplay of negative signs and nonmanual negative behaviors is complex in Levantine Arabic Sign Language. The most common manual negator consists of a repeated side-to-side movement with an extended index finger (Fig. 21), which is usually accompanied by the most common nonmanual negator, the side-to-side headshake. The complete paradigm of all negative signs again differs somewhat across dialects. Unlike other sign languages, Levantine Arabic Sign Language normally requires a manual negative sign to be present in the clause, that is, a headshake by itself cannot usually negate a clause (Hendriks, forthcoming). In addition to the headshake, a backward head tilt with or without a tongue click, formally identical to the negative gesture used by hearing people, occurs in Levantine Arabic Sign Language to mark negation (Fig. 22). The distribution of this head movement is more restricted than for the headshake. In the Jordanian variety, it mostly occurs by itself as a short negative response. However, in the Lebanese variety, the backward head tilt

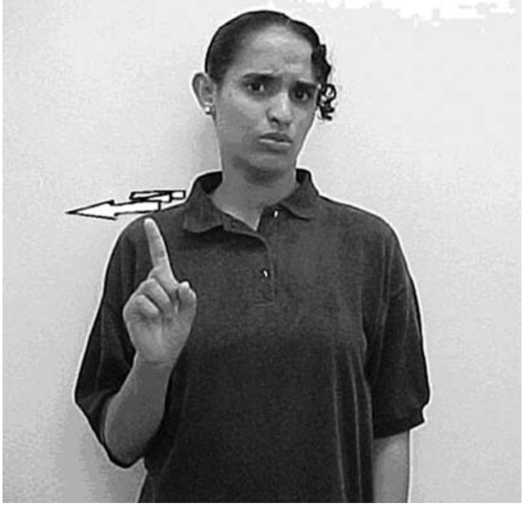


Figure 21. Manual negator in Levantine Arabic Sign Language.

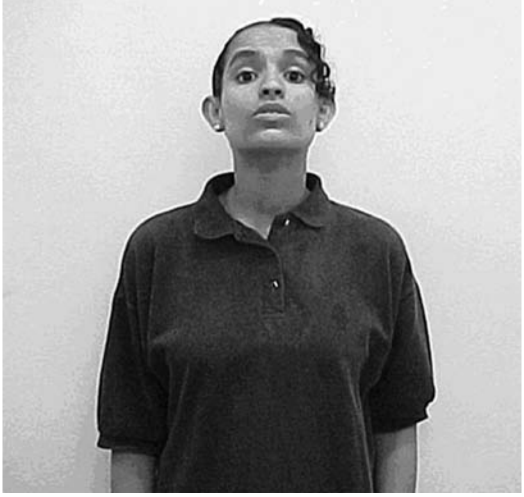


Figure 22. Backward head tilt for negation.

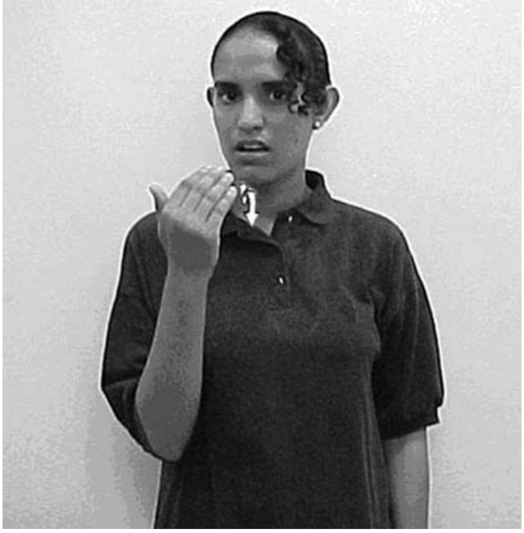


Figure 23. THERE-IS-NO.

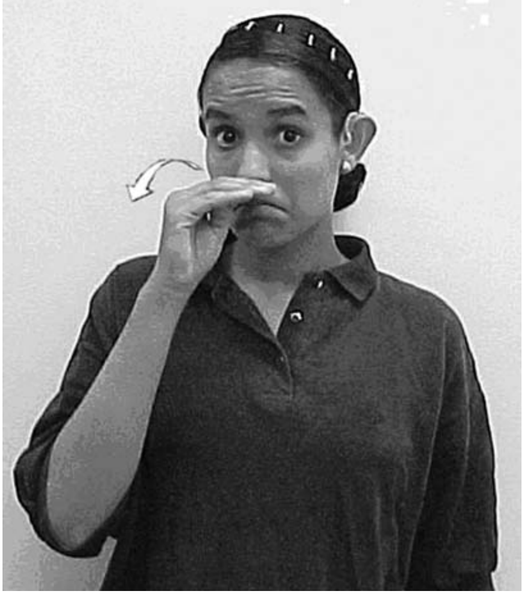


Figure 24. NOT-KNOW.

occurs regularly with a number of signs, including the negative existential THERE-IS-NO and the signs for NOT-KNOW and NOT-LIKE (Zeshan 2004). The first two are suppletive forms in Levantine Arabic Sign Language, that is, they differ from their positive counterparts in irregular ways (see Figs. 23 and 24).

Bibliographical references

Al-Muslat, Zaid Abdullah. 1994. “The history of deaf education in the Kingdom of Saudi Arabia”. *The Deaf Way: Perspectives from the International Conference on Deaf Culture*, ed. Carol J. Erting a.o., 275–282. Washington, D.C.: Gallaudet University Press.

Brelje, Herman W. (ed.). 1999. *Global perspectives on the education of the deaf in selected countries*. Hillsboro, Ore.: Butte Publications.

Dijken, Lianne van. 2004. *Creating perspective: Hand classifiers and referent projections in* lughat il-ishaara il-urdunia. M.A. thesis, University of Amsterdam.

Hendriks, Bernadet. 2004. *An introduction to the grammar of Jordanian Sign Language*. Salt: Balqa’ University.

——. Forthcoming. “Negation in Jordanian Sign Language: A cross-linguistic perspective”. *Sign languages: A cross-linguistic perspective*, ed. Pamela

Perniss, Roland Pfau, and Markus Steinbach. Berlin: Mouton de Gruyter.
Sandler, Wendy a.o. 2005. "The emergence of grammar: Systematic structure in a new language". *Proceedings of the National Academy of Sciences of the United States of America* 102.2661–2665.
Van Cleve, J.V. 1987. *Gallaudet encyclopedia of Deaf people and deafness*. 3 vols. New York: McGraw-Hill.
Zeshan, Ulrike. 2004. "Hand, head and face: Negative constructions in sign languages". *Linguistic Typology* 8:1.1–58.

Bernadet Hendriks
(Holy Land Institute for the Deaf)
Ulrike Zeshan
(University of Central Lancashire)

Ṣila

Derivatives of the root *w-ṣ-l* are used in Arabic grammatical theory to express the general idea of 'connecting' two linguistic units. Two terms derived from this root, *waṣl* and *ṣila*, are used as technical terms, along with *mawṣūl* as the correlate of *ṣila*, and the verbs *waṣala, ʾawṣala*, and *ittaṣala*.

The use of *ṣila* as a technical term goes back to the earliest Qurʾānic commentaries. Here, the term is used for redundant elements whose only function is to come between two linguistic units. Examples are found in Muḥammad al-Kalbī's (d. 146/763) commentary, for instance when he uses the term *ṣila fī l-kalām* for the preposition *min* in Q. 24/30 *yaġuḍḍū min ʾabṣārihim* 'that they cast down their eyes' (*Tafsīr* 146b4), or for the *kāf* in Q. 28/82 *wayka'annahu* 'woe!' (*Tafsir* 163a14; Versteegh 1993:118–119). The term is particularly frequent in Muqātil ibn Sulaymān's (d. 150/767) *Tafsīr* (Versteegh 1993:141–146).

According to Ibn Yaʿīš (*Šarḥ* VIII, 128–139), the term *ṣila* in this sense belonged to the terminology of the Kufan grammarians. It is indeed used frequently in al-Farrāʾ's *Maʿānī* (e.g. III, 187.11–15). Kinberg (1996:923–930) distinguishes between six different uses of the term in the *Maʿānī*: (i) redundant element; (ii) modifier of an indefinite noun; (iii) modifier after a relative pronoun; (iv) complement of a verb (e.g. a prepositional phrase); (v) attached element (e.g. *-mā* in *kullamā, niʿma mā*); and (vi) prolongation of a final vowel in pause or rhyme. The original exegetical meaning of 'redundant element' recurs in (i) and (v). The use of *ṣila* in the sense of 'relative clause as modifier', which is present in (ii) and (iii), is an innovation that al-Farrāʾ shares with the Basran grammarians (see below). The meaning of 'complement of the verb' in (v) is found for instance in Ibn as-Sarrāj (Taha 1993). The phonetic meaning under (vi) is something new and not found in later grammarians; perhaps, this meaning is related to that of 'redundant element', i.e. an element that serves merely to bridge the gap between two other elements, without having any other function.

According to Talmon (1995; 2003:222–232), the origin of the term *ṣila* lies in the translations of Greek logical writings. These preserved the Aristotelian tripartition of the parts of speech into two meaningful classes, nouns and verbs, and one meaningless class, the *sundesmós*, whose function was to conjoin the other parts. This is also the case in Jābir ibn Ḥayyān's terminology, where *ṣila* indicates the preposition which determines the relations between the noun and the verb (Kraus 1942:250). If Talmon is right about the Greek provenance of the term, the use of *ṣila* for redundant elements without any semantic content must have preceded its use in syntax in the sense of connective elements with their own meaning and function. Note that al-Fārābī (d. 337/949; *ʾAlfāḏ* 44) uses a related term, *wāṣila* (probably translated from the Greek term *sundesmós*), for a mixed group of words, consisting of the definite article, the vocative particle *yā*, and the quantifiers *kull* and *baʿḍ*.

In the Basran tradition, *ṣila* was not used for redundant elements, for which other terms existed, such as *zāʾid, laġw* (e.g. Sībawayhi, *Kitāb* I, 433.12 *mā* in *mahmā*; I, 405.5 *ḏā* in *mā-ḏā*), or *ḥašw* (as well as some other terms; see the exhaustive list in Talmon 2003:222ff.). The only Basran grammarian using *ṣila* in the sense of 'redundant element' is al-ʾAxfaš (d. 215/830 or 221/835), in his *Maʿānī l-Qurʾān* (e.g. *Maʿānī* 347), which is dedicated to Qurʾānic exegesis, just like al-Farrāʾ's. The other Basran grammarians did retain the term *ṣila*, but for a different function, to denote modifiers connected to the main clause by a connecting element called *mawṣūl*. In Sībawayhi's *Kitāb*, *ṣila* is frequently used in this sense (Troupeau 1976:217: 'adjunction', 70 times, sometimes in combination with the verb *waṣala* 'to connect', which occurs 25 times). The related term *waṣl* is predominantly used to indicate a morphophonological junction (the opposite of *waqf* 'pause'), as in the frequently used term *ʾalif al-waṣl* (45 times), although it is sometimes used in the sense of a syntactic connection as well. According to Mosel, the modifiers that are called *ṣila* in the *Kitāb* include relative clauses (1975:155–157), clauses introduced by *ʾan* (1975:192–193), and clauses introduced by *ʾanna* (1975:199–204).

The meaning of 'dependent clause as modifier' was to become the usual meaning of *ṣila* in later Arabic grammar, in which Basran terminology predominated. The earlier use of *ṣila* in the sense of 'redundant element' disappeared more or less from grammatical terminology, although in the exegetical literature it is still found from time to time, for instance when the word *mā* in its function as emphasizer rather than relative pronoun is referred to as a *ṣila* or *ḥašw* (Taha 1993:236). In later grammatical terminology, traces of this use of the term are also visible in a category of particles called *ḥurūf aṣ-ṣila*. According to Ibn Yaʿīš' definition (*Šarḥ* VIII, 128–139), this category includes the words *ʾin, ʾan, mā, lā, min, bi-*, all of which can be used without producing any additional meaning (*duxūluhu ka-xurūjihi min ġayr ʾiḥdāṯ maʿnā*).

Ibn as-Sarrāj (d. 316/928) uses the term *ṣila* more or less in the same sense as Sībawayhi: it is the name for constructions with the particle *mā*, with the relative pronoun, with the particle *ʾayy*, with *ʾan*, and with oaths (Taha 1993:237). In all of these structures, *ṣila* refers to an item that is connected with a noun in order to form a complete syntactic (and semantic) unit; the noun with which it is connected (i.e. the relative pronoun, the article in a participial construction, or the particle *ʾayy*) is called its *mawṣūl*. That the *ṣila* and its *mawṣūl* form one unit is shown among other things by the fact that neither the *ṣila* itself nor any part of it can be preposed (Ibn as-Sarrāj, *ʾUṣūl* II, 222–223). In the phrase *allaḏī ḍaraba zaydan ʿamrun* 'the one who is hitting Zayd is Amr', for instance, one cannot front *zaydan*, saying **zaydan allaḏī ḍaraba ʿamrun* (*ʾUṣūl* II, 223.11–12). Likewise, in the Qurʾānic verse *wa-kānū fīhi min az-zāhidīna*, usually interpreted as 'they did not put a great value on him' (Q. 12/20), Ibn as-Sarrāj does not analyze *fīhi* as belonging syntactically to *zāhidīna* (*fa-lā yajūzu ʾan tajʿala fīhi fī ṣ-ṣila*). In both sentences, the fact that a word is part of a *ṣila* prevents it from being moved before its *mawṣūl* (the *al-* in *az-zāhidīna*). In coordination, the entire phrase consisting of the *mawṣūl* and its *ṣila* must be coordinated, e.g. *ḍarabtu llaḏī fī d-dār wa-zaydan* 'I hit the one in the house and Zayd', while it is impossible to say **ḍarabtu allaḏī wa-zaydan* (Ibn as-Sarrāj, *ʾUṣūl* II, 69; Owens 1988:78–79).

According to Taha (1993:240–241), the term *ṣila* in the expression *min ṣilatihi* refers to a semantic connection, similar to the one expressed by Sībawayhi with the term → *sabab*. For instance, in the sentence

ʾaʿjaba *rukūb-u*
amazed.3ms riding-Nom
d-dābbat-a *zayd-un* *ʿamr-an*
the-riding.animal-Acc Zayd-nom ʿAmr-Acc
'The fact that Zayd rode the animal amazed ʿAmr' (Ibn as-Sarrāj, *ʾUṣūl* I, 138)

the word *dābba* is said to be connected (*min ṣilatihi*) to *rukūb*, hence the impossibility to prepose it; likewise, the word *zayd* cannot come before *rukūb* (Taha 1993:240). On the other hand, one could also analyze this in the sense of 'belonging to the syntagm of *rukūb*'. The expression *min sababihi* is also used by Ibn as-Sarrāj, but only in reference to nonverbal constructions (Taha 1993:241).

The terminology of the root *w-ṣ-l* 'to connect' is further used by Ibn as-Sarrāj in connection with his theory of transitivity (→ *taʿaddin*; Taha 1995): in his classification of verbs, he calls those transitive verbs *wāṣil* whose action actually goes from the agent to the object (*ʾUṣūl* I, 73), e.g. *ḍaraba* 'to hit', as against other transitive verbs such as *ḏanna* 'to regard as'. This usage is extended by Ibn as-Sarrāj to the link between the preposition and the noun governed by it (*ʾUṣūl* I, 408; Taha 1993:242), and, accordingly, it is also applied by him to cases like *marartu bi-zaydin* 'I passed Zayd', in which a verb is connected with its (semantic) object by means of a preposition. Ibn as-Sarrāj claims that in such cases the action of the verb passes to the noun (*ʾUṣūl* II, 13.14–16; Taha 1993:240), "because *bi-zaydin* is a patiens and that what is reached by the verb with the help of a preposition is equivalent in meaning to that which is reached by the verb itself, because the meaning of the expression *bi-zaydin* is *ʾataytu zaydan* 'I came to Zayd'" (*li-ʾanna bi-zaydin*

mafʿūl wa-l-wāṣil ʾilayhi al-fiʿl bi-ḥarf fī l-maʿnā ka-llaḏī yaṣilu ʾilayhi l-fiʿl bi-ḏātihi li-ʾanna qawlaka marartu bi-zaydin maʿnāhu ʾataytu zaydan).

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

ʾAxfaš, *Maʿānī* = ʾAbū l-Ḥasan Saʿīd ibn Masʿada al-ʾAxfaš al-ʾAwsaṭ, *Maʿānī l-Qurʾān*. Ed. Fāʾiz Fāris. 2 vols. Kuwait, 1979. (2nd. ed., Beirut: Muʾassasat ad-Dirāsa, 1974.)

Fārābī, *ʾAlfāḏ* = ʾAbū Naṣr Muḥammad ibn Muḥammad al-Fārābī, *Kitāb al-ʾalfāḏ al-mustaʿmala fī l-manṭiq*. Ed. Muḥsin Mahdī. Beirut: Imprimerie Catholique.

Farrāʾ, *Maʿānī* = ʾAbū Zakariyyāʾ Yaḥyā ibn Ziyād al-Farrāʾ, *Maʿānī l-Qurʾān*. Ed. Muḥammad ʿAlī an-Najjār. 3 vols. Cairo: ad-Dār al-Miṣriyya, 1966–1972.

Ibn as-Sarrāj, *ʾUṣūl* = ʾAbū Bakr ibn as-Sarī Ibn as-Sarrāj, *Kitāb al-ʾuṣūl fī n-naḥw*. Ed. ʿAbd al-Ḥusayn al-Fatlī. 3 vols. Beirut: Muʾassasat ar-Risāla, 1985.

Ibn Yaʿīš, *Sarḥ* = Muwaffaq ad-Dīn Yaʿīš Ibn Yaʿīš, *Šarḥ al-Mufaṣṣal*. 10 vols. Beirut: ʿĀlam al-Kutub; Cairo: Maktabat al-Mutanabbī, n.d.

Muḥammad al-Kalbī, *Tafsīr* = ʾAbū n-Naḍr Muḥammad ibn as-Sāʾib al-Kalbī, *at-Tafsīr [Tafsīr Ibn ʿAbbās]*. Ms. Chester Beatty, 4224.

Muqātil, *Tafsīr* = ʾAbū l-Ḥasan Muqātil ibn Sulaymān al-Balxī, *at-Tafsīr*. Ed. ʿAbdallāh Maḥmūd Šiḥāta. 4 vols. Cairo: al-Hayʾa al-Miṣriyya al-ʿĀmma li-l-Kitāb, 1980–1987.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. 2 vols. Ed. Bulaq, 1316 A.H. (Repr., Baghdad: Al-Muthanna Library, n.d.)

SECONDARY SOURCES

Kinberg, Naphtali. 1996. *A lexicon of al-Farrāʾ's terminology in his* Qurʾān *commentary with full definitions, English summaries, and extensive citations*. Leiden: E.J. Brill.

Kraus, Paul. 1942. *Jābir ibn Ḥayyān, contribution à l'histoire des idées scientifiques dans l'Islam: Jábir et la science grecque*. Cairo. (Repr., Paris: Les Belles Lettres, 1986.)

Mosel, Ulrike. 1975. *Die syntaktische Terminologie bei Sībawaih*. Ph.D. diss., University of Munich.

Owens, Jonathan. 1988. *The foundations of grammar: An introduction to medieval Arabic grammatical theory*. Amsterdam and Philadelphia: J. Benjamins.

Taha, Zeinab A. 1993. "The term *ṣila* in early Arab grammatical theory: The case of Ibn as-Sarrāǧ". *Proceedings of the Colloquium on Arabic Lexicology and Lexicography*, ed. Kinga Dévényi, Tamás Iványi, and Avihai Shivtiel, 233–244. Budapest: Csoma de Kőrös Society and Eötvös Loránd University.

——. 1995. *Issues of syntax and semantics: A comparative study of Sibawayhi, al-Mubarrad, and Ibn as-Sarraaj*. Ph.D. diss., Georgetown University.

Talmon, Rafael. 1995. "*Ṣila*". *Encyclopaedia of Islam* IX, 603. 2nd ed. Leiden: E.J. Brill.

——. 2003. *Eighth-century Iraqi grammar: A critical exploration of pre-Ḫalīlian Arabic linguistics*. Winona Lake, Ind.: Eisenbrauns.

Troupeau, Gérard. 1976. *Lexique-index du* Kitāb *de Sībawayhi*. Paris: Klincksieck.

Versteegh, Kees. 1993. *Arabic grammar and Qurʾānic exegesis in early Islam*. Leiden: E.J. Brill.

——. 1995. *The explanation of linguistic causes: Az-Zaǧǧāǧī's theory of grammar*. Amsterdam and Philadelphia: J. Benjamins.

KEES VERSTEEGH (University of Nijmegen)

Sinai Arabic

1. GENERAL

1.1 *Geographical*

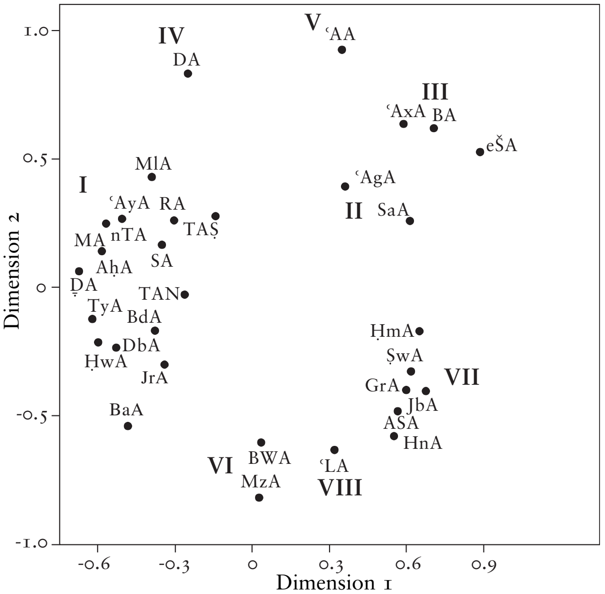
Wedged between North Africa and southwestern Asia lies the Sinai Peninsula. Through the ages, the area has served as a land bridge between the two continents, and across it Islam and the Arabic language were spread to Africa by Arabs from the Arabian Peninsula.

The most populous region of the Sinai Desert is its northern littoral. The central plain of at-Tīh and the mountainous region of the south (aṭ-Ṭūr) are only thinly populated. The majority of the population in the south live near the Gulf of Suez and the Gulf of Aqaba, or in towns on the coast, near St. Catherine's Monastery and in Wādiy Fēṛān, where the main road leads through central southern Sinai from St. Catherine's to the Gulf of Suez.

1.2 *Native speakers*

The total number of inhabitants of Sinai is well over 350,000 today (more than 300,000 live in North Sinai – with about half of these in al-ʿArīš – and more than 60,000 in South Sinai). Of this total, some 80,000 can claim Bedouin descent. Among Sinai's non-Bedouin population are the 'original' inhabitants of al-ʿArīš, while the majority of non-Bedouin newcomers are from the Nile Valley or Delta, having arrived via large-scale government settlement programs. A minority are Palestinians.

In older times, Bedouin in the northeast worked as farmhands during harvest times in Palestine, while those in the northwest used to frequent the eastern Nile Delta. These tribes almost all became fully sedentary, after completion of the Suez Canal in 1869 and creation of



Map 1. Multi-Dimensional Scaling Plot (by Geer Hoppenbrouwers), based on 87 criteria producing differences inside Sinai. Clusters are typologically relatively similar dialects in Sinai.

the state of Israel in 1948 made their seminomadic lifestyle impossible. Seasonal trekking, however, still occurs on a very limited scale, mainly in the center and south of Sinai.

Over the past decades, Bedouin have started settling near schools to secure an education for their children (the first schools in Sinai were built in the 1950s). Also, better health care is often cited as a reason to settle in more populated areas.

For details on the arrival of tribes in Sinai, see Bailey (1985).

1.3 *Tribes and dialect groups in Sinai*

Eight typological groups can be identified in Sinai, as illustrated by the Multi-Dimensional Scaling Plot (→ dialect geography; see Map 1). It shows relative typological distances between dialects. Every black dot represents a tribal dialect. Abbreviations used for the dialects of the tribes follow in brackets in the main text below.

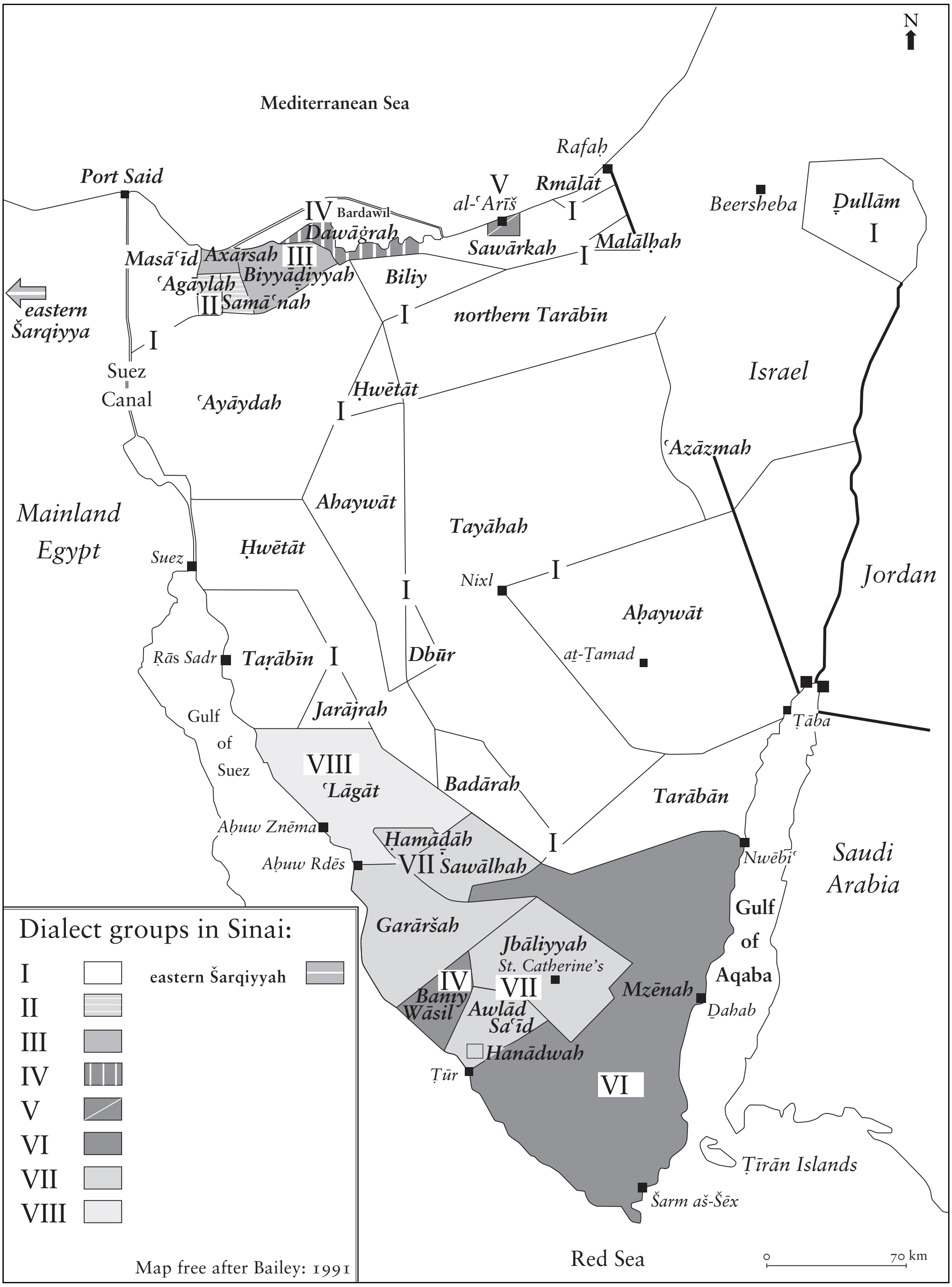
The geographical distribution of the tribes in Sinai is approximately as indicated in Map 2 (dialect groups are in roman numerals).

The eight typological groups are the following:

> Group I: Ḍullām in the Negev (Blanc 1970; abbreviation ḌA), northern Taṛābīn (nTA), Rmēlāt (RA), Sawārkah (SA), Biliy (BaA), the Masāʿīd (MA), the ʿAyāydah (ʿAyA). Group I tribes living more toward the center and south of Sinai are: the Malālḥah (MlA), Aḥaywāt (AḥA; Stewart 1987, 1988, 1990), Taṛābīn of Nwēbiʿ (TAN) and Rās Ṣadr (TAṢ), Jarājrah (JrA), Ḥwēṭāt (ḤwA), Tayāha (TyA), Badāṛah (BdA), and Dbūr (DbA).
>
> Group II: The dialect type of the ʿAgāylah (ʿAgA) and Samāʿnah (SaA).
>
> Group III: Typologically nearest to the sedentary dialect of the eastern Šarqiyya (eŠA; as described in Abul Fadl 1961; Woidich 1979, 1980; Behnstedt and Woidich 1987), and spoken by the Axārsah (AxA) and Biyyāḏiyyah (BA).

The dialects of Groups I–III form a continuum, or area of transition, between the Negev (I) type of Bedouin dialects and eŠA.

> Group IV: The dialect of the Dawāġrah (DwA), who are treated as pariahs by their neighbors. Their dialect has conserved many of its original (presumably central or southern) Najdi features, and it is not Northwest Arabian.


N
Mediterranean Sea
Port Said
Rafaḥ
V
Rmālāt
al-ʿArīš
IV Bardawil
Dawāġrah
I
Beersheba
Ḏullām
I
Masāʿīd
Axārsah
III
Sawārkah
Malālḥah
I
ʿAgāylah
Biyyāḏiyyah
Biliy
eastern Šarqiyya
II Samāʿnah
I
northern Tarābīn
I
Suez Canal
ʿAyāydah
Ḥwēṭāt
I
Israel
ʿAzāzmah
Mainland Egypt
Ahaywāt
Tayāhah
Suez
Ḥwēṭāt
I
Nixl
Jordan
I
Aḥaywāt
Rās Sadr
Tarābīn
I
Dbūr
at-Tamad
Jarājrah
Ṭāba
Gulf of Suez
VIII
ʿLāgāt
Badārah
Tarābān
Aḅuw Znēma
Ḥamādāh
I
Nwēbiʿ
Saudi Arabia
Aḅuw Rdēs
VII Sawālḥah
Gulf of Aqaba
Garāršah
Jbāliyyah
St. Catherine's
IV
VII
Mzēnah
Bany Wāsil
Awlād Saʿīd
Ḏahab
Hanādwah
Ṭūr
VI
Ṭīrān Islands
Šarm aš-Šēx
Dialect groups in Sinai:
I
eastern Šarqiyyah
II
III
IV
V
VI
VII
VIII
Map free after Bailey: 1991
Red Sea
0
70 km

Map 2. Approximate geographical distribution of the tribes in Sinai. Shades of gray indicate clusters of typologically relatively similar dialects (cf. Map 1).

Group V: The dialect of al-ʿArīš spoken by its original inhabitants (the Fawaxriyyah). It is not Northwest Arabian but rather a type of mixture of the southern Transjordanian type and the urban type (ʿAA; cf. Palva 1984:371; de Jong 2000:588).

Group VI: The dialects of Mzēnah (MzA) and Baniy Wāṣil (BWA). BWA was presumably originally much like the Group I type, but it now clearly 'leans' toward MzA in a dialect-typological sense (cf. de Jong 2004b).

Group VII: The dialects of Garāršah (GrA), Awlād Saʿīd (ASA), Ṣawālḥah (ṢwA), Jbāliyyah (JbA; cf. Nishio 1992), and the Ḥamāḏ̣ah (ḤmA). The Hanādwah (HnA) are an example of several non-Bedouin families living in Wādiy aṭ-Ṭūr.

Group VIII: The ʿLēgāt (ʿLg). These and notably ḤmA of Group VII have a number of remarkable features in common with Group II.

1.4 *Position of the dialects and linguistic type*

The dialects of the Bedouin tribes of Groups I–III and VI–VIII form a western extension into Sinai of the → Northwest Arabian (NWA) dialects, as proposed by Palva (1991; see below Sec. 1.5).

Members of all Bedouin tribes in Sinai are found in al-ʿArīš, and the dialect of this town also serves as a regional vehicle of communication in the north, although farther to the west, influences of the Cairo (or more or less Central Delta) variety are stronger.

On occasions like court sessions, or when reciting poetry, Bedouin often (attempt to) speak in a dialect type very similar to that of Group I, which is considered the appropriate (and prestigious) 'Bedouin' register for such social settings.

From a dialect-typological perspective, as one moves along the north coast from Group I in the east, via Groups II and III toward eŠA in the west, 'Bedouin' dialect characteristics (of the → Negev and southern Jordan) give way to more typically sedentary features heard in the eastern Nile Delta. The treatment of interdentals in northern Sinai illustrates this transition (see Fig. 1).

1.5 *Dialect notes in historical sources*

Apart from some notes by travelers or historians (e.g. in Šuqayr 1916:341–343), not much is known of the history of the dialects of Sinai.

1.6 *Dialects in surrounding areas*

For a selection of publications on dialects in (or adjacent to) the area, see bibliographical references below. Research into the dialects of the middle and south of Sinai was completed only recently (de Jong 2007).

Most tribes claim the Ḥijāz as their land of origin. This may well be true, but hardly anything is known of the route by which these tribes came to Sinai, or of the dialects spoken there.

A plausible hypothesis (Palva 1991) is that a Northwest Arabian group of dialects in southern Jordan (among these the Ḥwēṭāt and Bani ʿAṭīye; see Palva 1984–1986), the Negev, and Sinai links up with dialects in the southern Ḥijāz and the Tihāmah through as-yet-unknown dialects in the northern Ḥijāz. To test this hypothesis, research into dialects spoken on the Red Sea coast of Saudi Arabia remains a great desideratum.

Figure 1. Treatment of interdentals in northern Sinai

WEST				EAST
eŠA	III		II*	I*
	AxA	BA		
t, d, ḏ	*t ~ ṯ, ḏ ~ d, ḍ̱*	*t, d, ḏ*	*ṯ, ḏ, ḍ̱*	*ṯ, ḏ, ḍ̱*

* Also in Groups IV and VI–VIII, all three interdentals are present. In Group V, these have all been replaced by stops /t/, /d/, and /ḍ/.

2. Linguistic description

2.1 *General*

2.1.1 Northwest Arabian type vs. Najdi type

Prominent Northwest Arabian features of the dialects in Sinai, distinguishing these from → Najdi-type dialects, are the following (cf. Palva 1991:154–156):

- absence of *tanwīn* and its residues
- absence of affricated reflexes of **q* and **k*
- absence of final /n/ in the imperfect 2nd person singular feminine, and 2nd and 3rd person plural masculine (except in Group IV)
- absence of resyllabification of sequences (v = v or v̄) C*a*C*a*Cv > CCvCv (but present in Group IV)

2.1.2 Other general characteristics of Sinai dialects

Some major characteristics shared by all dialects in Sinai (unless indicated otherwise) are the following (cf. Palva 1991):

- presence of interdentals (but for BA of Group III and Group V, see Sec. 1.3)
- /j/ for **j* (allophones vary from *ǧ* to *ž* (IPA$[d^z]$ via [dʒ] to [ʒ] respectively))
- /g/ for **q*
- an active → *gahawa(h)*-syndrome (not Groups III and V), resyllabifying a sequence (C)*a*XC(v) > (C)*a*X*a*C(v)
- gender distinction in the 2nd and 3rd person plural in the nominal and verbal systems (not Group V)
- productivity of Form IV (not HnA, SaA of Group II, Groups III and V)
- occurrence of typical Bedouin lexical items (not Group V)
- occurrence of stressed 1st person singular pronominals *-ī* and *-nī* (not Groups IV and V)
- partial lack of phonemic distinction between /i/ and /u/
- reduction of geminates: $vC_aC_aC_bv > vC_aC_bv$
- 2nd person plural masculine pronominal suffix *-kuw* (but Groups VII, VIII, and SaA of Group II often also have *-kuṃ*)
- use of the *b*-imperfect (not Group IV)

Notwithstanding such similarities, there are differences between Sinai dialects, yielding the eight typological clusters (groups) mentioned above (see Map 1).

2.2 *Phonology*

2.2.1 Consonants

The inventory of consonants varies slightly between dialects (Table 1). The greatest differences are:

- an additional phoneme /ḳ/ in Group II and Groups VI–VIII
- only one interdental /ḏ̣/ in BA of Group III, and none in Group V

Table 1. Inventory of consonants

	bilabial		labio-dental		alveolar		inter-dental		postal-veolar		palatal		velar		uvular		phar-yngeal		laryngeal	
	vl	vd	vl	vd	vl	vd	vl	vd	vl	vd	vl	vd	vl	vd	vl	vd	vl	vd	vl	vd
plosive		b			t	d							k	g	(q)				(ʾ)	
emph.					ṭ	*ḍ*							*ḳ*							
nasal		m				n														
fricative			f		s	z	*ṯ*	*ḏ*	š	(*ž*)			x	ġ			ḥ	ʿ	h	
emph.					ṣ	(ẓ)		*ḏ̣*												
affricate												j								
trill						r														
emph.						(ṛ)														
lateral						l														
emph.						ḷ														
glides		w				y						y								

vd = voiced, vl = voiceless, emph. = emphatic
Note: Italics indicate differences in phoneme inventories of the dialects:
ḍ only in Group V. For interdentals, see Sec. 1.3. For *ḳ*, see Sec. 2.4.1. Marginal or secondary phonemes are in brackets.

Exceptions to regular reflexes occur in the case of interdentals in forms loaned from ‘Egyptian’ or Standard Arabic, such as (in Groups VI–VIII) _tillājih_ ‘refrigerator’, _talj_ ‘snow’, and (also elsewhere) _ḥādsih_ ‘accident’, _masalan_ ‘for instance’, _bizr_ ‘seed’, _maẓbūṭ_ ‘correct’, _ẓabbaṭ/yẓabbiṭ_ ‘to do properly’. *ʾ may have a reflex _w_, as in _ánwikal_ ‘it was eaten’, sometimes ʿ, as in _saʿal_ ‘he asked’ or _irʿ_ ~ _arʿ_ ‘behold!’ (*_r-ʾ-y_). Usually, *ʾ has simply disappeared, leaving behind a complementary lengthened vowel (e.g. *_raʾs_ > _ṛās_), a _y_ instead (e.g. *_ʿāʾilah_ > _ʿāylah_ ‘family’), or no trace at all, e.g. *_ruʾūs_ > _ṛūs_ ‘heads’.

2.2.2 Vowels

All dialects have three short vowel phonemes: /a/, /i/, and /u/. Although phonemic opposition between /i/ and /u/ is limited, the minimal pair _xiḏ̣r_ [proper name] - _xuḏ̣r_ ‘green [pl.]’ is found in all dialects. Generally, however, _u_ appears in velarized and/or labial environments and _i_ elsewhere.

Older /i/ and /u/ have been elided from open initial syllables, resulting in initial CCv sequences (see Sec. 2.2.4), e.g. _lsān_ ‘tongue’ (*_lisān_) and _ṛkab_ ‘knees’ (*_rukab_). The ‘vanished’ vowel /u/ (Blanc 1970:18) often left velarization behind (as in the second example).

Underlying /a/ in open syllables often surfaces as /i/ or /u/. The resulting unstressed high vowel is generally not elided in Bedouin dialects and is still underlyingly /a/, e.g. _yínfitiḥ_ (_yinfatiḥ_) ‘it is opened’, _simiʿt_ (_samiʿt_) ‘I heard’, _kibīr_ ‘big’ (_kabīr_), _kitabt_ (_katabt_) ‘I wrote’, _dikātrih_ (_dakātrah_) ‘doctors’. However, in Group V, forms like _yíniftiḥ, smiʿt_ (the latter also in BA), and _ikbīr_ are found.

In dialects of Groups I, IV, and VI–VIII, /a/ reappears in closed syllables in derived forms, e.g. _yinfatḥuw_ ‘they are opened’, _tištaġḷiy_ ‘she works’. In some dialects of Group I (not all), /a/ also reappears in closed syllables in verbal Form I, e.g. _šarbit_ ‘she drank’ (e.g. Ḏ̣A, ḤwA, TyA), but in Group VI it surfaces as _simʿat_ ‘she heard’ and _kubrin_ ‘they [fem.] grew older’. In Group II, Form _n_-I also has underlying /a/, but it does not reappear as /a/ in closed syllables, e.g. _yinfítḥin_ ‘they [fem.] are opened’.

All dialects have the phonemes /ī/, /ē/, /ā/, /ō/, and /ū/. In neutral environments, *_ay_ and *_aw_ generally have /ē/ and /ō/ reflexes in all dialects, but Groups I, IV, and VI also have phonetically conditioned diphthongs _ay_ and _aw_ (see Sec. 2.2.3).

Other dialects have monophthongs irrespective of phonetic environment: _ʿēn, ḏ̣ēf, ḥōl, ṣōm_. In Groups VII and VIII, these are realized low (slightly below IPA [ɛː] and [ɔː], respectively), when preceded by X (i.e. any back spirant) or a velarized consonant.

In Group I (mainly central and northeastern dialects), long vowels /ē/ and /ī/ may phonetically overlap, e.g. _bīt_ ‘house’, _sīf_ ‘sword’.

2.2.2.1 Final _-ā(ʾ)_

Final long vowel _-ā_ has been raised in neutral environments to _-iy_ in Groups I and IV, e.g. _štiy_ ‘winter’, _sōdiy_ ~ _sōdíy_ ‘black [sg. fem.]’, and (only Group IV) _jiy_ ‘he came’. When it is not raised, *_-ā_ has often remained long in Groups IV and VI–VIII. Final long (unraised) _-ā_ is often accompanied by a glottal catch, usually in pause, e.g. _ḥamṛāʾ#_ ‘red [sg. fem.]’. In Group I, this unraised *_-ā_(ʾ) is usually short, e.g. _bēḏ̣áʾ#_ ‘white [sg. fem.]’.

In BaA (of Group I) and Groups VI–VIII, final _-ā(ʾ)_ has often been raised to _-iʾ_ ([ɪʔ] (BaA) _hniʾ_ ‘here’, _štiʾ_ ‘winter’ (JbA _íštiʾ_), and (Group VI) _ġadíʾ_ ‘dinner’, but in the singular feminine adjectival pattern *C_a_CC_āʾ_: _sōdíy_ ‘black’ and _šōlíy_ ‘left-handed’. In Groups II, III, and V, the reflexes of *_-ā(ʾ)_ have become more like current realizations of the feminine morpheme, e.g. _sōda_ or _sōdih_.

2.2.3 Diphthongs

In Groups I, IV, and VI, diphthongal *_ay_ and *_aw_ have remained in positions directly preceded by back spirants or velarized consonants, e.g. _ʿayn_ ‘eye’ and _ṭayr_ ‘birds’, _ṣayf_ ‘summer’, and _ḥawl_ ‘year’ and _ṣawm_ ‘fasting’.

2.2.4 Syllables

Possible syllables in Sinai are (examples underlined): Cv (_ka-tab_), Cv̄ (_ʿā-rif_), CvC (_ka-tab_), Cv̄C (_ʿār-fih_), CvCC (_bint_), Cv̄CC (_ḥāṭṭ_). These may also appear as such on the surface. Underlyingly, the following syllable types are also possible: CCv(C) (_ṛkab_) and CCv̄(C) (_lsān_). Surface forms (anaptyctics are indicated as _ə_) are, however, bisyllabics: ʾəC-Cv(C) and ʾəC-Cv̄(C). In Group V, underlying CC-initial syllables are not current: contrast for instance (Bedouin) _iṣṣgūr_ or _aṣṣgūr_ with _liṣgūr_ ‘the falcons’ in Group V.

2.2.5 Elision, consonant clusters, and anaptyxis

2.2.5.1 Elision

Dialects of Sinai (except Group IV*) are '*différentiel*'; (only underlying) high vowels /i/ and /u/, but not /a/, are elided in unstressed open nonfinal syllables. Group IV has additional rules for resyllabication (cf. Sec. 2.3.6), the feminine morpheme in construct state (cf. Sec. 2.6.1), and the pronominal suffix -*k* (cf. n. 5 in Sec. 2.4.1). The general rule for the dialects of Sinai is:

I > Ø / (v)$C_a(C_b)_C_c$v
I = /i/ or /u/; C = any consonant; v = any vowel

Resulting clusters may then become eligible for the anaptyxis rule (see Sec. 2.2.5.2). If, however, C_aC_b is a geminate, /i, u/ are not elided, but the geminate may be reduced, e.g. *njaddid* + *uh* > *njádiduh* 'we renew it'.

2.2.5.2 Anaptyxis

The anaptyxis rule is:

Ø > ᵓ / (C)$C_a\_CC_a$
C_a = C or #

Clusters #CC are resolved, e.g. *#ḥmāṛ* > *#(ᵓ)əḥmāṛ* 'donkey'. Clusters CCC or CC# may be left intact, depending on the degrees of sonority of the consonants involved (e.g. *bint#* 'girl', but *ʿabəd#* 'slave'). Clusters CCCC are resolved, e.g. *rikibt* + *ḥmāṛ* > *rikibt əḥmāṛ* 'I rode a donkey'. Geminates in a cluster $C_aC_aC_b$ may be phonetically reduced, e.g. *nṭabbiʿ* + *-ih* > *nṭabbʿih*, in IPA [ʔən'ɫɑbʕeh]) 'we train him'.

The phonetic quality of the anaptyctic vowel is colored by its surroundings: toward IPA [ʊ] in a labial and/or velarized environment (anaptyctic vowel underlined), e.g. *yúḍurbuw* 'they hit', otherwise toward IPA [ɪ], e.g. *yíkitbin* 'they [fem.] write'.

2.2.6 Stress

As a general rule, the vowel of the first heavy sequence (of the morphological base) from the right is stressed for Sinai. A heavy sequence is defined as vCC or v̄(C). All dialects in Sinai are of the *mádrasah* stress type, except BA of Group III, where stress is *madrásah*.

The domain of stress, however, varies, and in the absence of heavy sequences, different dialects have different solutions.

Stress is placed in the following manner (by group) in the following words: (a) 'camel'; (b) 'she wrote'; (c) 'the boy'; (d) 'it was opened'; (e) 'it is opened'; (f) 'she wrote it [masc.]'; (g) 'your [sg. masc.] neck'; (h) 'he stretches' (roman numerals refer to dialect groups).

> Group I: (a) *jimál* ~ *jámal*; (b) *kátabat* ~ *kitábat*; (c) *álwalad*; (d) *ánfitaḥ*; (e) *yínfitiḥ*; (f) *kitábatih* ~ *kátabatih*; (g) *ṛagábatak* ~ *ṛágabatak*; (h) *ymidd*.
>
> Groups II, VII, and VIII: (a) *jámal*; (b) *kátabat*; (c) *ilwálad* (II, VII), *álwalad* (VIII); (d) *infátaḥ*; (e) *yínfitiḥ*; (f) *kátabatu*(*h*); (g) *ṛagabátḳ* (cf. n. 2 in Sec. 2.4.1), (h) *ymidd*.
>
> Group III: (a) *jámal*; (b) *kátabat*; (c) *ilwálad*; (d) *infátaḥ*; (e) *yinfátiḥ*; (f) *katabátu* (~ *kátabatu* in AxA); (g) *ṛagabátak*; (h) *ymidd*.
>
> Group IV: (a) *jámal*; (b) *ktíbat*; (c) *álwalad*; (d) *ánfitaḥ*; (e) *yínfitiḥ*; (f) *ktíbtah*; (g) *ṛgúḅatk* (cf. n. 6 in Sec. 2.4.1); (h) *yímidd*.
>
> Group V: (a) *jámal*; (b) *kátabat*; (c) *ilwálad*; (d) *infátaḥ*; (e) *yíniftiḥ*; (f) *kátabatu*; (g) *ṛágabatak*; (h) *iymidd*.
>
> Group VI: (a) *jimál* ~ *jámal* (BWA only *jimál*); (b) *kátabat*; (c) *álwalad*; (d) *ánfitaḥ*; (e) *yínfitiḥ*; (f) *kátabatuh* (~ MzA *kátbituh*); (g) *ṛagabát*^{u}*ḳ* (cf. n. 2 in Sec. 2.4.1), (h) *ymidd*.

Rule ordering in all dialects is (i) elision, (ii) stress, (iii) anaptyxis.

2.3 *Phonotactics*

2.3.1 Raising of the feminine morpheme -*ah*

The feminine morpheme -*ah* tends to be raised in nonvelarized environments. The degree of raising differs somewhat in the various groups. Generally, raising may reach IPA [ɪ^{h}] in Groups I, II, and IV and VI–VIII. In Group III, raising is slightly lower and depends on pausal position. In Group V, such raising – only in pause – reaches up to a value between IPA [ɛ] and [e].

2.3.2 Assimilation

Apart from regular assimilations of ‘sun letters’, /l/ of the article may assimilate to /j/ (or /ž/) in some dialects.

Assimilations that occur may be (a) partial or (b) total and (x) regressive, (y) progressive, or (z) reciprocal. Instances are: (bx) *ridt* > *ritt* ‘I wanted’, *bitzaġriṭ* > *bizzaġriṭ* ‘she ululates’; (ax) *bnalga* > *mnalga* ‘we find’, *nbāt* > *mbāt* ‘we spend the night’; (by) *šēxhuṃ* > *šēxxuṃ* ‘their sheikh’, *yšūfhiy* > *yšūffiy* ‘he sees her’; (bz) *miʿhiy* > *miḥḥiy* ‘with her’. Total assimilation of initial /h/- of pronominal suffixes to preceding voiceless consonants as in (by) occurs mainly in Groups I and VI–VIII.

2.3.3 The *gahawah*-syndrome

The → *gahawah*-syndrome is active in Groups I, II, IV, and VI–VIII. In Groups III and V, it is not regular. Thus in all dialects one finds *baḥaṛ* ‘sea’ (even in Group V, where *gahawah* forms are rare) and *šahaṛ* ‘month’ and *laḥam* ‘meat’ (and suffixed, e.g. *šaháṛha* ‘her month’), but no *gahawah* vowels appear in, for example, (Group V) *naxl* ‘palm trees’, *aḥmaṛ* ‘red’, and (Group III) *gahwah* ‘coffee’, *naʿjah* ‘ewe’. Suffixing such forms then results in, for example, *gáhiwtu* ‘his coffee’.

The *gahawah*-syndrome affects Form I verbs as well, e.g. *taḥariṯ* ‘you plow’. When such verb forms (in dialects that stress C*á*C*a*CvC) are suffixed, stress is *táḥarṯih* ‘you plow it’. Dialects that stress *taḥáriṯ* will also stress *taḥárṯih*.

When nominals are suffixed, Group I and ʿAgA have forms like *gáhawtih* ‘his coffee’, but SaA and Groups VI–VIII have *láḥamatu(h)* ‘his piece of meat’ (in MzA ~ *láḥmituh*), although verb forms in Group VI are like those in Groups I and II, e.g. *yáḥarṯin* ‘they [fem.] plow’. In Group IV, all sequences C*a*C*a*Cv are subject to the Najdi rule of → resyllabication (see Sec. 2.3.6).

Derived forms and quadriradicals are not affected by the *gahawah*-syndrome, e.g.: *aʿṭat* ‘she gave’, *gahwāk* ‘he served you coffee’, *istaġṛabt* ‘I wondered’.

2.3.4 The *bukaṛa*-syndrome

The → *bukaṛa*-syndrome often creates intrusive vowels, e.g. (underlined), *nidris* > *nidi̲ris* ‘we thresh’, *bukṛah* > *buka̲ṛah* ‘tomorrow’.

2.3.5 Metathesis and the mutual influence of sibilants

→ Metathesis is especially frequent in Groups VI and VII when sibilants are involved, e.g. *šizih* ‘[the game of] *sījah*’, *šāz* ‘baking sheet’ (more commonly *ṣāj*), *šizn* ‘prison’ (**sijn*), *mšazzil* ‘recorder’ (**msajjil*), *našiz*(#) ‘weaving’ (**nasj*), but usually *šajaṛah* ‘tree’ (often *sajaṛah* elsewhere) and in many dialects *šamš* ‘sun’ (**šams*).

2.3.6 Najdi resyllabication

A conspicuous characteristic of Group IV is the Najdi rule for → resyllabication:

C*a*C_a*a*C_bv > CCICv
C*a*C*a*Lv > CC*a*Lv
(after *gahawah*-syndrome) C*a*X*a*Cv > CX*a*Cv

C = any consonant; C_a = any consonant but X; C_b = any consonant but /l/ or /r/; I = /i/ or /u/; L = /l/ or /r/; X = /x, ġ, ḥ, ʿ/ or /h/

Examples: *samakah* > *smikah* ‘a fish’, *ḥaṭabah* > *ḥṭubah* ‘a piece of firewood’, *baṣalah* > *bṣalah* ‘an onion’, *šajaṛah* > *šjaṛah* ‘a tree’, *gahawah* > *ghawah* ‘coffee’.

2.3.7 Morphophonology: The feminine morpheme in construct state

In Group I, ʿAgA of II, III, and V, the feminine morpheme -*ah* > *at* when (historical) *a*C directly precedes. In SaA and Groups VI–VIII, this is also the case when /a/ is a *gahawah*-vowel (e.g. *gahawatī* ‘my coffee’, MzA *gahawathin* ‘their [fem.] coffee’). In Group I and ʿAgA, the feminine morpheme following (*gahawah*) X*a*C > -*it* when suffixed with vowel-initial suffixes (/i/ is then elided), e.g. *náxaḷtu* ‘his palm tree’. When sequences other than *a*C directly precede, the feminine morpheme > -*it* (or -*t* when v precedes), e.g. *ʿilbitha* ‘her pack’, *ʿílibtu* ‘his pack’.

In Group IV, the feminine morpheme > -*at* when C precedes, e.g. *ʿilbatha* ‘her packet’, *ṛguḅátyah* ‘my neck’, *gháwatk* ‘your coffee’. The *a* of -*at* (and also of the 3rd sg. fem. ending -*at*; cf. Sec. 2.2.6, for Group IV example (f)) is dropped in open syllables, e.g. *nāgtah* ‘his she-camel’.

2.4 *Morphology*

In nominal morphology, short /a/ is often raised in the pattern C_1*a*C_2*ī*C_3 > C_1*i*C_2*ī*C_3: in most dialects of Group I: C_1*i*C_2*ī*C_3, e.g. *kibīr, ʿirīs*.

In Groups II, IV, VI–VIII, and BaA, TAṢ, TAN, JrA, TyA, and BdA of Group I, raising is optional, e.g. *kibīr* ~ *kabīr, ʿirīs* ~ *ʿarīs*.

In Group III, raising is phonetically conditioned, e.g. *kibīr*, but no raising after X (i.e. any of the back spirants), e.g. *ʿarīs*. In some villages (e.g. Gṭayyiʿ and Najīlah) of Group III, there is no raising, e.g. *kabīr, ʿarīs*.

In Group V, a pattern *i*C_1C_2*ī*C_3 is used (after raising of *a* and its subsequent elision: morphological restructuring), but C_1*a*C_2*ī*C_3 is also in use, mainly when C_1 = X, e.g. *ikbīr, ʿarīs*.

2.4.1 Pronominals

Pronominals in Sinai are given in Table 2 (forms in parentheses occur less regularly).

The pronominal suffixes are given in Table 3.

2.4.2 Demonstratives

The singular masculine form for near deixis may be velarized in Group I: *hāḏ̣a* ~ *hāḏa*. In RA and SA (of Group I) and in Group II, *hāḏa* is current, *hāda* in Groups III and V (~ *da* ~ *dih* in Groups II and III). Groups VI–VIII have *ḏa* ~ *ḏiʾ* (~ *hāḏa*). In Group IV, final *-ā* (of **hāḏā*) of the singular masculine is raised and stressed: *hāḏíy*, contrasting only in stress with the singular feminine *hāḏiy*. For ḤwA and JrA, postpositioned *ha* is typical for the singular masculine, e.g. *álwalad ha*(ʾ) 'this boy'.

Singular feminine in Group I: *hēḏiy* (~ *hāḏiy* in RA, SA, BaA, and ḤwA) and in Groups II, III *hāḏiy* (respectively *hādiy*) (~ *diy* in both) and V. Groups VII–VIII have *hāḏiy* ~ *ḏiy* (~ *ḏiyyih* in ASA and GrA; in GrA, *hāḏiy* is rare) and MzA *hāḏiy*(*yih*).

Plural common forms in Groups VI–VIII are unvelarized (*hā-*)*ḏill*(*-ih*) ~ *ḏillēl*(*-ih*) (~ K-form (*hā-*)*ḏōl* or *dōl*). Group IV distinguishes gender in the plural: (masc.) *hāḏ̣ōḷ*(*-ḷah*) ~ *hāḏ̣oḷḷāw* and (fem.) *hāḏ̣oḷḷayn*. Elsewhere, only velarized plural forms occur: *hōḏ̣aḷ*(*-ḷah*)/*hawḏ̣aḷ*(*-ḷah*), *hāḏ̣aḷ*(*-ḷah*) and/or *hāḏ̣ōḷ*(*-ḷah*).

For far deixis, plural forms with suffixed *-k* (+ *-ah* ~ *-ih*) are used, e.g. (pl. comm.) forms like *hōḏ̣aḷḷāḳ*(*-ah*), *hāḏ̣oḷḷāḳ*(*-ah*) or (*ha-*)*ḏ̣aḷḷāḳ*

Table 2. Pronominals in Sinai Arabic

Group:		I	II	III	IV
	sg.				
3rd	masc.	*hū*	*hū (~ huwwa)*	*hū (~ huwwa)*	*íhwa ~ úhwa*
	fem.	*hī*	*hī (~ hiyya)*	*hī (~ hiyya)*	*íhya*
2nd	masc.	*int(a) (~intih)*	*inta (~int)*	*inta*	*int(a)*
	fem.	*intiy*	*intiy*	*inti*	*intiy*
1st	comm.	*aná(ʾ) ~ ána*	*áni (~ ána, anī)*	*áni (~ ána)*	*ána*
	pl.				
3rd	masc.	*huṃ(ṃa)*	*huṃ(ṃa)*	*huṃ(ṃa)*	*huṃṃa*
	fem.	*hin(na)*	*hinna*	*hin(na)*	*hinna*
2nd	masc.	*intuw*	*intuw (~ intum)*	*intu*	*intuw*
	fem.	*intin*	*intin*	*intin*	*intin*
1st	comm.	*iḥna**	*iḥna*	*iḥna**	*iḥna*

* In ḤwA, Ḏ̣A, and BA, *aḥna*

Group:		V	VI	VII–VIII	
	sg.				
3rd	masc.	*huwwa*	*hū*	*hū ~ huwwa*	
	fem.	*hiyya*	*hī*	*hī ~ hiyya*	
2nd	masc.	*inta*	*int(ih)*	*inti(h)*	
	fem.	*inti*	*intiy*	*intiy*	
1st	comm.	*áni*	*ána*	*ána*	
	pl.				
3rd	masc.	*huṃṃa* (comm.)		*huwwa*	*huṃ(ṃa)*
	fem.	-	*hinna*	*hin(na)*	
2nd	masc.	*intu* (comm.)	*intuw*	*intuw ~ intum*	
	fem.	-	*intin*	*intin*	
1st	comm.	*iḥna*	*iḥna*	*iḥna*	

Table 3. Pronominal suffixes in Sinai Arabic

Group:		I	II	III
	sg.			
3rd	masc.	C-*ah* ~ C-*ih*, v-*h*	C-*u(h)*, v-*h*	C-*u*, v-*h*
	fem.	*-hiy or -ha (~ -hiʾ BaA)*	*-ha*	*-ha*
2nd	masc.	C-*ak,* v-*k*	CC-*uḳ*, v(C)-^{u}*ḳ*^{*2}	C-*ak*, v-*k*
	fem.	*-kiy (*CC*-ik BaA)*	CC-*ik*, v(C)-*k*	*-ki*
1st	comm. (gen.)	C-*ī*, v-*y (~* C*-yah in BaA)*	C-*ī*, v-*y*	C-*ī*, v-*y*
	comm. (obj.)	*-nī ~ -ni*	*-nī ~ -ni*	*-nī ~ -ni*
	pl.			
3rd	masc.	*-huṃ*^{*4}	*-huṃ*	*-huṃ*
	fem.	*-hin*	*-hin*	*-hin*
2nd	masc.	*-ḳuw*	*-ḳuṃ ~ -ḳuw*	*-ku*
	fem.	*-kin*	*-kin*	*-kin*
1st	comm.	*-na (~ -niʾ in BaA)*	*-na*	*-na*

Group:		IV	V	VI–VIII
	sg.			
3rd	masc.	M-*ah*^{*1}, C-*ih*	C-*u*, v-*h*	C*u(h)*, v-*h*
	fem.	*-ha*	*-ha*	*-ha ~ -hiʾ*
2nd	masc.	*-k*^{*6}	C-*ak*, v-*k*	CC-*uḳ*, v(C)-^{u}*ḳ*^{*2} *(~* C*-uḳ)*
	fem.	-(C)C-*ik*, v-*kiy*	C-*ik*, v-*ki*	CC-*ik*, v(C)-*k*^{*3}
1st	comm. (gen.)	*-ya(h)~ -yih*	C-*i*, v-*ya*	C-*ī* ~ C-*i*, v-*y*
	comm. (obj.)	*-nya(h) ~ -nyi(h)*	*-ni*	*-nī ~ -ni*
	pl.			
3rd	masc.	*-huṃ ~ -huw*	*-hum* (comm.)	*-huṃ*^{*4}
	fem.	*-hin*	-	*-hin*
2nd	masc.	*-kuw*	*-ku* (comm.)	*-ḳuw~ -ḳuṃ*^{*5}
	fem.	*-kin*	-	*-kin*
1st	comm.	*-na*	*-na*	*-na ~ –niʾ*

n. 1 M = velarized C

n. 2 Superscript ^{u} indicates strong velarization accompanied by lip rounding (not a vowel in terms of syllabication; stress is, for example, *ḍarabát*^{u}*ḳ* 'she hit you [sg. masc.]'). In Group II, these forms co-occur with C-*ak* and v-*k* (used by younger generations).

n. 3 In BWA ~ *-kiy*.

n. 4 *-huw* in Group VI. *-huw ~ -huṃ* in JrA, BaA, and TAN.

n. 5 Only *-ḳuw* in Group VI.

n. 6 Suffixing *-k* will not cause stress to shift (contrast n. 2), nor is a cluster C*k#* resolved, e.g. *jámalk#* 'your camel'. A cluster CC*k#* is resolved, e.g. *šúfitk#* 'I saw you [sg. masc.]'.

(*-ah*) in Groups I, II, and IV. Group VI has (*hā-*)*ḏ̣áḷḷaḳ*(*-ah*), Groups VII–VIII *ḏ̣aḷḷāḳ*(*-ah*) (~ *hāḏōḳah* in ḤmA). In Group III, (*hā-*)*daḷḷāk* is used for the plural masculine, and (*hā-*) *daḷḷāka* for the plural feminine. In Group V, plural common is *hadlāk*(*-a*) and the form *hadkúm* (*-ma*).

The singular forms for far deixis are usually suffixed with *-k*(*ah*) or *-k*(*-ih*), e.g. in Group IV, singular masculine *hāḏāk* (with long *ā*) and singular feminine *-hāḏīk*(*-ah*). In Group I, dialects with *hāḏa* and/or *hāḏa* have *hāḏāk*(*-ah*) and/or *hāḏ̣āk*(*-ah*). Those with *hāḏiy* and/or *hēḏiy* have *hāḏīk*(*-ih*) and/or *hēḏīk*(*-ih*). Group II has *hāḏāk* and *hāḏīk*(*-ih*), Group III has *hādāk* and *hādīk*(*-a*), and Group V has *hadāk*(*-a*) and *hadīk*(*-a*). In Groups VI–VIII, forms are usually without initial *hā-: ḏāḳ*(*-ah*) and *ḏīk*(*-ah*).

2.4.3 Presentatives

Presentatives are *hā-* or (heard in Group II) *hē-*, e.g. *hēhū jiʾ* 'here he has come', *hāhī jat* 'here she has come'. Suffixed with *-y* (*hāy/hay*), it is used for a nearer deixis – 'here I/we have' – and with suffixes *-k/-kiy* (*hāk/hākiy* 'here you [masc./fem.] have') for farther deixis.

2.4.4 Relative pronoun and article

The relative pronoun is *alliy* for those dialects that have *al-* as the article, and *illiy* for those that have *il-*. Groups II and VI, which use both articles, have *illiy*. In Group I, the article is predominantly *al-*, and in Group VI it usually harmonizes with the vowel color of the following noun (similarly in Group IV), e.g. *ánngaṛ* 'the potholes', *ílihṣiy* 'the rocks'. Groups III, V, and VII use *il-* and *illiy*.

2.4.5 Interrogatives

Among interrogatives used in Sinai are: 'what?' *ēš* (~ *īš* in Group I) and *ēh* (Groups I–III). *Ēh* is common in Group IV (and predominant in Group III), and *ēš* is current in Group V. *Ēš* is predominant in Group VI and *ēh* in VII–VIII. The distribution of 'why?' *lēh* (~ *līh*) and *lēš* (~ *līš*) is parallel to that of *ēh* and *ēš*. In ḎA and ḤwA, both *ēš* ~ *wišš* occur as variants.

Other interrogatives are:

'where?' – *fēn* (Group III) and *wēn* (elsewhere)
'how?' – *izzāy* (Group III) and *kēf* ~ *kīf*, sometimes *kayf* (elsewhere)
'who?' – *mīn* (SaA, Groups III, V–VIII); *min* in Group I (~ *man* in BaA) ʿAgA and *man* in Group IV
'how much/many?' – *kam* (Groups I, III, IV, VI–VIII), *ákam* (Group V), *akám* (ʿAgA), *kām* (SaA, Group III)
'when?' – *wagtēh* (Groups I, III, IV (~ *míta*), VI), *wagtēš* (Groups I, II, V), *matá* or *matā* (Group I, BWA, but ḤwA *matān*), MzA *mitēh*, ASA *imtēh*, GrA and JbA *mitēn*
'which?' – *yāt* in Groups I, IV, VI; *iyyāt* in Groups II, III (~ *anhu/anhi*), VII, VIII

ʿasa + pronominal suffix expresses the hope that the addressee is well (or it is used to inquire after his or her well-being), as in *ʿasāk ṭayyib* 'are you [sg. masc.] well?' (Group I). To inquire 'what is the matter?' (mainly) in dialects of Groups I and VI–VIII, *ʿalām* + pronominal suffix is heard, as in *ʿalām*^{u}*ḳ ya-ḫuw zēd* 'what is the matter, oh Abu Zayd?' (GrA).

2.4.6 Adverbs

Temporal: (*h*)*alḥīn*/*ilḥīn* is used throughout Sinai for 'now'. Variations include: *halḥīnit* (MzA), (*h*)*alḥīniy* (BaA, SaA, Group IV), and *ilḥīnih* (GrA, JbA). The K-form *dilwagtiy* is heard predominantly in the northwest.

Local: *hniy* is 'here' in Groups I (~ *hniʾ* in BaA, *hāna* in MlA) and IV, and usually *nihāʾ* ~ *nihāniy* in Groups VI–VIII. In Group II *íhnih* (sporadically *hāna*) is used. In Group III, *hāna* is current (~ *ihna* in BA). The current form in Group V is *hína* (also as K-form elsewhere). Variations are: *hniyāt, hniyāniy, hniyāntiy, hniyyih, hniyyan*, and others. In some dialects (notably of Group I) also *fi hāḏ̣a*/*fi hāḏa* for 'here' is used. 'There' is *hnuh* in Groups I, IV, MzA (~ *hnūtiy*), ASA and ḤmA (~ *hnōtiy*). In Groups II, III, and V, GrA and JbA (~ *hnuh*) *hnāk* is current (also as K-form elsewhere). For 'there [far away]', *ġād* or *ġādiy* is used. Some dialects have *faḏ̣āḳ* (*fi* + *ḏāḳ*) for '[over] there'.

2.4.7 Analytic genitive

The analytic genitive is formed with *šuġl* in Groups I, II, and IV (~ K-form *btāʿ*). In GrA, JbA, and Group III, *šuġl* ~ *btāʿ* (the latter is preferred in BA), in Group V (*b*)*tāʿ*, and in Group VI *šuġl* and *ḥagg*.

2.4.8 Negation

Negation of verb forms is generally done with preceding *mā* in Group I and MzA (the compound negation is much less regular). Other dialects use *mā* + verb form + *š*(*i*), while the single negation usually expresses more emphatic negation (often in combination with *xāliṣ* 'at all').

'There is none' is *mā fīh* (in Group I, ~ sometimes also *ma fīš*) or *mā fīš* (in Groups II-IV, VI; often *ma fišši* in Group V). In Group I, *māš* is also regular.

2.5 *Verbs*

All dialects are of the *aktib/niktib* type. Groups III and V have vowel harmony in the imperfect prefixes of the *i-* and *u* types (e.g. *yiktib* and *yugʿud*), but not in the *a-* type (e.g. *yišṛab*). Other dialects also show vowel harmony in the imperfect prefixes of the *a-* type: *yiktib, yugʿud*, and *yašṛab*. In Groups VI–VIII, this vowel harmony is applied throughout the paradigm, including the 1st person singular of *i-* and *u*-type imperfects: *iktib-niktib, uḏrub-nuḏrub* (and *ašṛab-našṛab*). Elsewhere in Sinai, initial *a-* in the 1st person singular of imperfects is current in all three vowel types: *aktib, agʿud*, and *ašṛab*.

In Group IV, imperfect endings are -v*n* in the 2nd and 3rd person plural masculine and the 2nd person singular feminine, e.g. *tikitbūn, yikitbūn, tikitbīn* (respectively); elsewhere, forms are *tíkitbuw, yíkitbuw*, and *tíkitbiy* (in SaA and ḤmA ~ *-um* in the pl. masc.), and perfect and imperfect plural feminine forms are with -v*n* (i.e. *-in* or *-an*), e.g. *yíkitb*v*n, yášṛab*v*n* and *širb*v*n, kátab*v*n*. In Group I, v harmonizes with the base vowel (see below). In Group IV, the endings are vowelless *-n* (e.g. *yašṛábn, yḥalíbn, katábn*).

In Group I, vowel-initial endings of the *a*-type perfect and *a*-type imperfect harmonize as well, e.g. *tasmaʿay* 'you [sg. fem.] hear' and *t/yasmaʿaw* 'you/they [pl. masc.] hear' and *t/yasmaʿan* 'you/they [pl. fem.] hear'. In the perfect, e.g. *katabaw* 'they [masc.] wrote', *kataban* 'they [fem.] wrote', and *katabat* 'she wrote'. In Groups II, III, IV and VI, VII, and VIII, endings are (imperfect) *-iy, -uw, -in*, respectively, and (perfect) *-uw, -in* and also *-at* (SaA also has pl. masc. *-um*). For the 3rd person singular feminine of the *a*-type perfect, all have *-at* (but e.g. *širbit, kubrit*). Groups IV and VI have *-at* in all vowel types of the perfect (e.g. *širbat, kubrat*).

In *gahawah*-forms of verbs (i.e. when $C_2 = X$; cf. Sec. 2.3.3), older /a/ of the prefix has been preserved, e.g. *yaxabuṭ* 'he knocks', *yaḥalib* 'he milks'. In Group IV, such *gahawah*-forms are resyllabified (cf. Sec. 2.3.6): *yxábuṭ* and *yḥálib*. In Groups III and V, forms are *yuxbuṭ, yiḥlib*, etc. (For treatment of underlying /a/ in verb forms, cf. Sec. 2.2.2.)

2.5.1 Vowel types of Form I verbs

Vowel types of Form I verbs for the perfect are: (*a*-type perfect) *katab* 'to write', (*i*-type perfect) *simiʿ* 'to hear', and (*u*-type perfect, quite rare) *kubur* 'to grow old'. For the imperfect (*a*-type) *yasmaʿ* or *yismaʿ*, (*i*-type) *yiktib*, and (*u*-type) *yuḍrub* or *yukbur*.

2.5.2 Derived verbal Forms

Dialects in Sinai have an active Form IV, e.g. *aʿṭa/yiʿṭiy* 'to give' (except in Groups III, V, SaA (uncertain), and HnA).

Form II patterns are: $C_1aC_2C_2aC_3/yC_1aC_2C_2iC_3$.

Form *ta*-II patterns are: $taC_1aC_2C_2aC_3/ytaC_1aC_2C_2aC_3$. In northern dialects of Groups I and in IV, these patterns co-occur with $(i)tC_1aC_2C_2aC_3/yitC_1aC_2C_2aC_3$ (especially the latter is current). In southern Group I dialects, the $ytaC_1aC_2CaC_3$ pattern is much more regular.

Form III patterns are: $C_1\bar{a}C_2aC_3/yC_1\bar{a}C_2iC_3$.

Form *ta*-III patterns are: $taC_1\bar{a}C_2aC_3/ytaC_1\bar{a}C_2aC_3$. Here, too, the $yitC_1\bar{a}C_2aC_3$ pattern is used almost to the exclusion of the $ytaC_1\bar{a}C_2aC_3$ pattern in northern Group I and in Group IV, while $itC_1\bar{a}C_2aC_3$ co-occurs with $taC_1\bar{a}C_2aC_3$. In southern Group I dialects, $ytaC_1\bar{a}C_2aC_3$ is much more regular.

In Groups VI–VIII, *ta-* prefixes are current in Forms *t*-II and *t*-III in perfect and imperfect; elsewhere, (*i*)*t-* prefixes are used.

In dialects with *ta*-prefix in the imperfect of Forms *t*-II and *t*-III, an initial *tt-* cluster is reduced, e.g. (*ttaġadda* >) *taġadda* 'she has lunch'.

As the basic passive for Form I, Form *n*-I is used. As for vowel distribution and stress, Form I-*t* runs parallel to Form *n*-I (see examples d and e in Sec. 2.2.6).

Form *ista/asta*-I: alternating vowels in Groups I, IV–VIII $ast\acute{a}C_1C_2aC_3$ (Groups V–VIII $ist\acute{a}C_1C_2aC_3$)/$yist\acute{a}C_1C_2iC_3$; fixed *a* in Groups II, III $ist\acute{a}C_1C_2aC_3/yist\acute{a}C_1C_2aC_3$.

Patterns of quadriliteral verbs run parallel to Form II: $C_1aC_2C_3aC_4/yC_1aC_2C_3iC_4$. With the *ta-* or (*i*)*t-* prefix, the parallel is with Form *t*-II): $taC_1aC_2C_3aC_4/ytaC_1aC_2C_3aC_4$ or $(i)tC_1aC_2C_3aC_4/yitC_1aC_2C_3aC_4$.

2.5.3 Geminated verbs

The imperfect vowel is usually /u/ in velarized environments and /i/ in neutral environments, e.g. *yḥuṭṭ* 'he places', *yjuṛṛ* 'he pulls', but *ymidd* 'he stretches', *yliff* 'he wraps, turns'. In Group IV (where stress is *yúḥuṭṭ, yímidd*, etc.), the plural feminine endings for medial geminate verbs are (perf.) *-ann* and (imperf.) *-inn*, e.g. *šaddánn* 'they [fem.] pulled' and *yxuššínn* 'they [fem.] enter'.

In Groups II and VI–VIII and some dialects of Group I, the perfect base vowel /a/ is raised > /i, u/ when preceding stressed /ē/, e.g. *middēt* 'I stretched', *suwwēt* 'I did'.

2.5.4 Verbs Iʾ

The perfect of Iʾ verbs 'to eat' and 'to take': Groups IV–VI, ḌA, nTA, AḥA, BaA, MA, ḤwA, and TyA have *akal* and *axaḏ*. Elsewhere forms without initial *a-* are current.

The imperfect vowel and degree of velarization varies: *yākil, yāxiḏ* or *yākul, yāxuḏ* elsewhere. The active participle is *wākil* and *wāxid* in Groups III and V (~ *ʾākil, ʾāxid* in Group V), *mākil, māxiḏ* elsewhere.

Imperatives: *úxuḏ* or *xuḏ*, *uxḏiy, uxḏuw, uxḏin* (in AḥA, and Groups II and VII), *xuḏ, xḏiy, xḏuw, xḏin* (in BaA, SA, ḎA, AḥA, TAN, TAṢ, ḤwA, Groups VI and VIII), and *xuḏ, xuḏiy*, etc. elsewhere. Some dialects (e.g. AḥA, JbA) have the various forms side by side.

2.5.5 Verbs I*w*

I*w* verbs may have *w* incorporated in the imperfect, as in *yōgaf* 'he stands', *yōjiʿ* 'he hurts' (*yawgaf, yawjiʿ* occur predominantly in Groups I and IV). Such forms, all with (originally) *a* in the preformatives and subsequently monophthongized to *ō*, are current in all groups, except in Group V (there *yiwrid, yiwṣal*).

In some dialects, such as AḥA and Group IV, forms without incorporated *wāw* occur as well, e.g. *yiríd* 'he waters', *yišíl* 'he arrives'.

2.5.6 Verbs II*w/y*

In Groups VI–VIII, the long base vowel is often shortened in the 2nd person singular masculine of the imperfect, e.g. *túgul* 'you say', *tíġib* 'you are absent',

The singular masculine imperative is *šil* 'carry!'. In some dialects, a vowel precedes, e.g. *úgum* 'get up!' and *íšil* 'carry!'.

Group IV has plural feminine forms with vowelless *-n*, triggering the use of short base vowels, e.g. *guln* 'they [fem.] said' and *yguln* 'they [fem.] say'.

2.5.7 Verbs III*w/y*

Apocopated imperatives for the singular masculine occur in Groups I, IV, and VI–VIII, e.g. *irm* 'throw!', *saww* 'do, make!' and *iʿḍ* 'give!'.

In Groups VI–VIII, 2nd person singular masculine imperfects are apocopated as well (both *i*- and *a*-types in all forms), e.g. *timš* 'you go', *tans* 'you forget', *tsaww* 'you make/do', *tlāg* 'you find', *tiʿṭ* 'you give', *taġadd* 'you have lunch'.

III*y a*-type verbs tend to have final *-a* that is raised to *-iʾ* in Group VI, e.g. *mišíʾ* 'he went'.

In Group IV, the plural feminine ending *-n* is doubled: *nasánn* 'they [fem.] forgot' and *yimšínn* 'they [fem.] go'.

2.5.8 Verb 'to come'

Perfect

Group:		I		II		III	
		sg.	pl.	sg.	pl.	sg.	pl.
3rd	masc.	*ja(ʾ)*	*jaw*	*jih*	*jum*	*īja*	*īju*
	fem.	*jat*	*jan*	*jāt*	*jin*	*ījat*	*ījin*
2nd	masc.	*jīt*	*jītuw*	*jīt*	*jītum*	*jēt*	*jētu*
	fem.	*jītiy*	*jītin*	*jītiy*	*jītin*	*jēti*	*jētin*
1st	comm.	*jīt*	*jīna*	*jīt*	*jīna*	*jēt*	*jēna*

Group:		IV		VI–VIII		V	
		sg.	pl.	sg.	pl.	sg.	pl.
3rd	masc.	*jiy*	*jāw (uw)*	*jiʾ*	*juw***	*īja*	*īju* (comm.)
	fem.	*jāt*	*jinn*	*jāt*	*jin*	*ījat*	
2nd	masc.	*jīt*	*jītuw*	*jīt*	*jītuw***	*ijēt*	*ijētu* (comm.)
	fem.	*jītiy*	*jītin*	*jītiy*	*jītin*	*ijēti*	
1st	comm.	*jīt*	*jīna*	*jīt*	*jīna*	*ijēt*	*ijēna*

** In ḤmA and ʿLA ~ *jum* and *jítum*.

Imperfect

Group:		I		II and III*		IV	
		sg.	pl.	sg.	pl.	sg.	pl.
3rd	masc.	*yjiy*	*yjuw*	*yījiy*	*yījuw***	*yíjiy*	*yjūn*
	fem.	*tjiy*	*yjin*	*tījiy*	*yījin*	*tíjiy*	*yjinn*
2nd	masc.	*tjiy*	*tjuw*	*tījiy*	*tījuw***	*tíjiy*	*tjūn*
	fem.	*tjiy*	*tjin*	*tījiy*	*tījin*	*tjīn*	*tjinn*
1st	comm.	*ajíy*	*njiy*	*ājiy*	*nījiy*	*ájiy*	*níjiy*

* Endings *–uw* and *-iy* in Group III are less diphthongal.
** In SaA, forms are *yījum* and *tījum*.

Group:		VII, VIII		VI*		V	
		sg.	pl.	sg.	pl.	sg.	pl.
3rd	masc.	*yījiy*	*yījuw*	*yijíy*	*yijúw*	*yīji*	*yīju* (comm.)
	fem.	*tījiy*	*yījin*	*tijíy*	*yijín*	*tīji*	
2nd	masc.	*tij ~ tīj(iy)*	*tījuw*	*tij*	*tijúw*	*tīji*	*tīju* (comm.)
	fem.	*tījiy*	*tījin*	*tijíy*	*tijín*	*tīji*	
1st	comm.	*ījiy*	*nījiy*	*ijíy*	*nijíy*	*āji*	*nīji*

* For *j* ~ *ī*, see Sec. 2.1.2.

2.6 *Innovations in the verbal system*

Throughout the area (except in Group IV), the *b*-imperfect – one of the hallmarks of sedentary dialects – expresses the habitual present tense. In many dialects (mostly in the north), initial *y*- of the imperfect may be dropped (mainly in *i*- and *u*-types) when *b*- precedes, e.g. *búguʿduw* 'they sit' and *biktib* 'he writes'.

2.7 *Future and volition*

To express volition (with an added sense of 'futurity'), *widd* (Groups I and IV) or *bidd* (elsewhere) is used, e.g. *widdna ngōṭir* 'we want to [shall] go'. Alternatively, *ha-* (mainly *ḥa-* in Group III) prefixed to the imperfect (or less frequently with *ṛāḥ* ~ *ṛaḥ* preceding < *rāyiḥ*) may be used for the future, e.g. *ṛāḥ aḥkílak zayy mā ṣār* 'I'll tell you like it happened'. Another way to express future is to use the simple present tense.

2.8 *Some additional remarks*

2.8.1 Agreement

Limited or countable numbers of objects, and even male persons – whatever the grammatical gender of the singular noun – tend to be referred to in the plural feminine, e.g. *w ixwān xamsithin iwlād ḥamdān* 'and the brothers – the five of them – are the sons of Ḥamdān' (AḥA).

2.8.2 Enclitic suffixing

Suffixed prepositions *l* and (much rarer) *b* may in turn be enclitically suffixed to verb forms, e.g. (preposition *l*) *btugˁúd-ilhiy* 'she stays [for herself]' (RA of Group I), *tašṛáb-ilk* 'you drink [for yourself]' (Group IV); (preposition *b*) *byifríg-ibh-ássalab* 'he separates the rope with it [sg. fem.]' (BaA).

2.8.3 Grammaticalizations

Unconjugated *gām* is often used as a particle expressing a notion of ingressiveness/suddenness of the action expressed in the verb that follows, *gām gult* 'I then said'.

Xāf or *xāfaḷḷah* expresses doubt and can be translated with 'maybe, perhaps' and seems to be used mainly for undesirable possibilities, e.g. *xāfaḷḷah minta ṛajil* 'maybe you're not a (real) man'.

For desired possible developments, *kīd* is used, e.g. *kīd síˁirhin yaṭlaˁ* 'maybe their price will go up' (said by someone planning to sell).

Min xawf/xōf (*ma/la*) lit. 'for fear that' has developed into a conjunction expressing 'lest', e.g. *binˁalligha min xawf la tawgaˁ* 'we tie it up, lest it should fall'.

BIBLIOGRAPHICAL REFERENCES

Abul Fadl, Fahmi. 1961. *Volkstümliche Texte in arabischen Bauerndialekten der ägyptischen Provinz Šarqiyya mit dialektgeographischen Untersuchungen zur Lautlehre*. Ph.D. diss., University of Münster.

Bailey, Clinton. 1985. "Dating the arrival of the Bedouin tribes in Sinai and the Negev". *Journal of the Economic and Social History of the Orient* 28.20–49.

——. 1991. *Bedouin poetry from Sinai and the Negev*. With a foreword by Wilfred Thesiger. Oxford: Clarendon Press.

Bani Yasin, Raslan and Jonathan Owens. 1984. "The Bduul dialect of Jordan". *Anthropological Linguistics* 26.202–232.

Behnstedt, Peter and Manfred Woidich. 1987. *Die ägyptisch-arabischen Dialekte*. III. *Texte*. 1. *Delta-Dialekte*. Wiesbaden: L. Reichert.

Blanc, Haim. 1970. "The Arabic dialect of the Negev Bedouins". *Proceedings of the Israel Academy of Sciences and Humanities* 4.112–150.

De Jong, Rudolf. 2000. *A grammar of the Bedouin dialects of the northern Sinai littoral: Bridging the linguistic gap between the Eastern and Western Arab world*. Leiden: E. J. Brill.

——. 2004a. "Characteristics of Southern Sinai Bedouin dialects: Preliminary observations". *Approaches to Arabic dialects: A collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday*, ed. Martine Haak, Rudolf de Jong, and Kees Versteegh, 151–175. Leiden: E.J. Brill.

——. 2004b. "The dialect of the Mzēnih of South Sinai". *Estudios de Dialecología Norteafricana y Andalusí* 8.59–79.

——. 2007. "Notes on the dialects of the ˁLēgāt and Ḥamāḏ̣ah of southern Sinai". *Approaches to Arabic linguistics presented to Kees Versteegh on the occasion of his sixtieth birthday*, ed. Everhard Ditters and Harald Motzki, 565–593. Leiden: E.J. Brill.

Murray, George William. 1935. *Sons of Ishmael: A study of the Egyptian Bedouin*. London: Routledge.

Nishio, Tetsuo. 1992. *A basic vocabulary of the Bedouin Arabic dialect of the Jbāli tribe*. (= *Studia Sinaitica*, I.) Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.

Oppenheim, Max Freiherr von. 1943. *Die Beduinen*. II. *Die Beduinenstämme in Palästina, Transjordanien, Sinai, Ḥedjāz*. Leipzig: O. Harrassowitz.

Palva, Heikki. 1977. "The descriptive imperative of narrative style in spoken Arabic". *Folia Orientalia* 18.5–26.

——. 1984. "A general classification for the Arabic dialects spoken in Palestine and Transjordan". *Studia Orientalia* 55.359–376.

——. 1984–86. "Characteristics of the Arabic dialect of the Ḥwēṭāt tribe". *Orientalia Suecana* 33–35.295–312.

——. 1991. "Is there a North West Arabian dialect group?". *Festgabe für Hans-Rudolf Singer, zum 65. Geburtstag am 6. April 1990. überreicht von seinen Freunden und Kollegen*, I, ed. Martin Forstner, 151–166. Frankfurt: P. Lang.

Procházka, Theodore Jr. 1988. *Saudi Arabian dialects*. London: Kegan Paul International.

Rosenhouse, Judith. 1984. *The Bedouin Arabic dialects: General problems and a close analysis of north Israel Bedouin dialects*. Wiesbaden: O. Harrassowitz.

Salonen, Erkki. 1979. *Zum arabischen Dialekt von Gaza*, I. *Studia Orientalia* (Helsinki) 51:10.
——. 1980. *Zum arabischen Dialekt von Gaza,* II. *Annales Academiae Scientiarum Fennicae, SARJA – Ser. B NIDE*, CCXIII. Helsinki: Suomalainen Tiedeakatemia.
Stewart, Frank Henderson. 1987. "A Bedouin narrative from Central Sinai". *Zeitschrift für Arabische Linguistik* 16.44–92.
——. 1988. *Texts in Sinai Bedouin law*. I. *The texts in English translation*. Wiesbaden: O. Harrassowitz.
——. 1990. *Texts in Sinai Bedouin law*. II. *The texts in Arabic*. Wiesbaden: O. Harrassowitz.
Šuqayr, Naʿīm. 1916. *Tārīx Sīnā al-qadīm wa l-ḥadīṯ wa juġrāfīyatuhā*. (Repr., Beirut: Dār al-Jīl, 1991.)
Woidich, Manfred. 1979. "Zum Dialekt von il-ʿAwāmṛa in der östlichen Šarqiyya (Ägypten)". *Zeitschrift für Arabische Linguistik* 2.79–99.
——. 1980. "Zum Dialekt von il-ʿAwāmṛa in der östlichen Šarqiyya (Ägypten). II. Texte und Glossar". *Zeitschrift für Arabische Linguistik* 4.31–60.
Yrttiaho, Kaarlo. 1988. "Texts from Arabia Petraea in the dialect of the semi-nomadic an-Nʿēmāt tribe of the Shara Mountains (Jordan)." *Studia Orientalia* 64.145–168.

RUDOLF DE JONG (University of Amsterdam)

Singular → Number

Slang

'Slang' is an English term which is sometimes used erroneously to refer to what in Arabic is termed *al-ʿāmmiyya* or *ad-dārija* or *al-lahja*, i.e. linguistic entities which are more appropriately referred to in English as colloquials or dialects. A more precise correspondence to the English term 'slang' might be the Arabic term *sīm* or the phrase *raṭāna ʿāmmiyya*. The former term is more closely akin to a very limited and specific form of slang which in English is termed 'argot' or 'cant' (also 'lingo'), while the latter phrase captures that aspect of slang which is temporary and fast changing and more closely fits the English term 'slang'.

Due to the differences in the linguistic situation between English and Arabic, it is necessary to define each of the above terms in order to clarify their exact meanings within the context of each language. First, all of the terms refer to forms of the spoken language, and all denote a deviation in some way from a standard. For English, it is a spoken standard, based on the country under consideration (e.g., standard American English differs from standard British English or 'Received Pronunciation', but one would not be considered 'slang' with regard to the other).

In an English language context, the term 'colloquial speech' refers to speech that differs from a standard based on *level of informality*. Colloquial speech is less formal than 'standard' speech, based on the presence of certain patterns of speech which are considered inappropriate in formal contexts but are nonetheless in wide use and understood by all. This would include, for example, the use of contractions, or lexical items not ordinarily found in formal contexts, including certain slang or dialectal terms which have come into general use. The term 'colloquial' is originally related to the notion of 'conversational', since it is the form of speech used in ordinary conversations and does not have the connotation of a stigmatized uneducated or unrefined style of speech, as it may have when applied to Arabic contexts. The term 'dialect', on the other hand, is speech that deviates from a standard based on *locale*. In defining a dialect, one must also recognize the definitive role that politics play in deciding which variety of a language is recognized as the standard and which is the dialect. In most cases, the language of the capital city comes to be recognized as the standard form of both speech and writing, while the other varieties become perforce 'dialects'.

Slang, in opposition to both of the above terms, deviates from a standard (as well as from colloquial speech and dialects) based on its *novelty*. That is, slang is speech which employs either newly minted words or ordinary words with newly developed meanings to impart a vividness to one's speech, as well as to set the speaker apart as a member of a certain 'in' group. The evocation of group solidarity is an important element of slang usage, and although group solidarity may play a role in an individual's usage of a dialect or more colloquial speech, the use of slang terms tends to evoke a much more limited and socially circumscribed grouping, often of groups that fall outside the power structure, e.g. youth (→ youth speech), racial or ethnic minorities, musicians, etc. Related to the term 'slang' are the terms 'cant' or 'argot', which both refer to the specialized speech of certain social groups, some stigmatized (such as prison inmates, homosexuals, prostitutes, and drug users), to which the term 'cant' is usually applied, and others professional in nature, to which the term ‘argot’ is applied (although this distinction is often not made). What distinguishes cant/argot from slang is its relative persistence as opposed to the impermanence of slang, as well as the heightened importance of secrecy in the use of cant/argot (see also → jargon; → secret languages).

In the Arabic situation, due to → diglossia, defining a spoken standard is much more difficult. The standard written language, referred to as *al-fuṣḥā*, is used in certain spoken contexts in the media as well as in religious functions, and although highly valued and respected, it does not function as a spoken standard in the same way that standard American English does. The English term ‘colloquial’ is applied to commonly spoken forms of Arabic, not because these deviate from a standard but more due to the connotation of ‘conversational’ speech. They may also be described as informal speech, but simply because their informality follows from the fact that they are used in everyday conversation, the determination of what is a formal and informal context may vary widely. The term ‘dialect’ may as well be misunderstood when applied to the widely varying forms of spoken Arabic. In an English context, a ‘dialect’ is defined as such because it deviates from the standard *spoken* form, based on locale. In an Arabic context, a dialect is a much different matter, due to the nature of the standard: *al-fuṣḥā* is a supranational standard primarily for written functions, with its spoken functions being very limited and circumscribed (but extremely important nonetheless). In lieu of a supranational spoken standard, certain ‘dialects’ function to all intents and purposes as local spoken standard languages, the dialect usually associated with the capital or largest city. Although these are unofficial standards (in the sense that they are not recognized politically, and culturally do not have institutions associated with their development and upkeep), they nevertheless do exert a strong influence over the linguistic practices of individuals living within those states. When applying the word ‘slang’ to an Arabic context, one is therefore describing forms of speech which deviate from *this* kind of spoken standard based on its novelty and ephemeral nature, and not on deviance from the written standard. Dialects as well may be divided between those which function as national standards (Cairene, Damascene, Baghdadi, etc.) and those which do not (Ṣaʕīdi, Ḥomsi, Muṣili, etc.).

Because of the nature of the linguistic situation in Arabic, therefore, finding the best possible Arabic equivalent for the term ‘slang’ in Arabic is a problem. Traditionally, Arabic opposes to *al-fuṣḥā* ‘the most eloquent’ form of the language various terms in common use, none of which are equivalent to ‘slang’: *al-ʕāmmiyya* or the language of *ʕāmmat an-nās* ‘common people’, and *ad-dārija*, the language which is ‘in circulation’, both of which may be equivalent to ‘colloquial’ Arabic. The term *lahja* is the best translation for ‘dialect,’ although in earlier stages of Arabic culture the term → *luġa* (today the common equivalent to ‘language’) was also used in the sense of ‘dialect’. The term *sīm* is used in Egyptian Arabic primarily as a translation equivalent of ‘argot/cant’, but in some instances may connote a kind of slang as well. One dictionary consulted provided *raṭāna ʕāmmiyya* as a translation of ‘slang’, and while that does seem like the most appropriate, it does not seem to be widely understood or commonly used in exactly that sense.

Rather than trying to find or coin a term in Arabic which unequivocally connotes the same thing as the English term ‘slang’, it may be better to consider examples of such language. Allam (2000) examines ‘youngsters’ new vocabulary’ (*kalām aš-šabāb*) in Cairo – or, in other words, teenage slang (→ youth speech). She takes twenty words, almost all of which express an intensification of a good or bad evaluation of something or someone, classifies them based upon the degree of standardization, and then examines their use among different segments of the population.

Allam first classifies the expressions into three groups: (i) those with both a standard (dictionary) reflex and a slang reflex which have nearly the same meaning; (ii) those with both a standard (dictionary) reflex and a slang reflex which have different meanings; (iii) those without a standard (dictionary) reflex but with a slang reflex. Each of these subcategories shows the degree of acceptance into the standard dialect of these words. The first group, by being included in a dictionary (a good test for ‘standard-ness’, especially in what is primarily an oral language), as well as sharing the same

Table 1. Cairene teenage slang terms (Allam 2000)

	Dictionary meaning		Slang meaning	Understand	Use
I: *inšakaḥ*	‘to fill with pleasure’		same meaning	92	36
falsaʿ	‘skedaddle, scram’		same meaning	76	24
ʾišṭah	*ṣabāḥ il-ʾišṭa* ‘good morning’		same meaning: ‘okay, good, alright’	44	0
iṭṭaḥan	‘to be ground down; to become exhausted’		same meaning	32	0
xaryaṭ	‘exhausted’		same meaning	0	0
II: *haṭal*	‘to be slow, weak, undisciplined’	>	‘silly, stupid’	64	0
kabbar	‘to say *allāhu ʾakbar*’	>	‘to ignore, not to pay attention’	60	40
bīʾah	‘environment’: *min bīʾa waṭya*	>	‘low class, low standard’	52	12
lasaʿ	‘to scorch, sting, steal’	>	‘to go bonkers, to flip’	52	12
ʾustāz	[respectful title]	>	‘excellent, great, of high quality [adj.]’	48	0
riwiš	‘to distract, disturb, be unnerving’	>	‘cute, cool, beautiful, stylish’	44	12
xayyiš	‘to cover with sackcloth’	>	‘to fail, mess up’	24	0
ʾafaš	‘to seize, catch in the act’	>	‘to get angry at someone because of something they’ve done’	20	0
itbaʿat	‘to be sent’	>	‘to be fooled, tricked, deceived’	16	0
sabbit	‘to fix, provide evidence’	>	‘to put someone in his place by convincing argument’	8	0
sayyaḥ	‘to fill to overflowing’	>	‘to spread around a secret’	4	0
III: *mifayyaṣ*	–		‘exhausted, unable to perform’	72	16
ṭaḥn	–		‘extremely good, very positive’	48	8
ʾantax	–		‘to relax, sit doing nothing’	20	0
hartil	–		‘to go crazy, talk nonsense’	16	0

meaning as the current slang term, indicates that they have become, if not accepted as standard, at least more recognized and perhaps more used as such than the other items, since they have likely been in use for some time. The second group of words indicates a departure from the standard: while the standard meaning of each item is easily relatable to its slang meaning, there is enough divergence to create a miscommunication for those not ‘in the know’. The lack of recognition on the part of the lexicographers of the third group of terms indicates its marginality and perhaps its novelty as well.

Allam provides another sort of grouping based on the understanding of these words by a group of older individuals. As opposed to the near unanimous understanding of these words in their slang meaning by younger individuals, for older individuals there is a much greater disparity in understanding some of them than others. Interestingly, this does not line up with the degree of standardization indicated by the first category. For example, while the category (i) words *inšakaḥ* and *falsaʿ* are the most recognized among the older segment of the population, a category (iii) word, *mifayyaṣ*, is almost as widely recognized. Furthermore, the slang meanings of two other category (i) words, *ʾišṭah* and *iṭṭaḥan*, are not recognized by more than half of the older population, while the slang connotation of another category (i) word, *xaryaṭ*, is not understood at all. Category (ii) words (those with a new sense for old) show a relatively high degree of understanding among the older group, with about half being understood by about half or more of the older individuals, while two of the four category (iii) words are also understood by a significant number of older individuals. This second grouping (based on understanding among different age groups) indicates that there is a significant difference between the younger group, for whom these words represent an aspect of their age group identity, and the older group, for whom they do not. Understanding of some of the slang terms among this older group may be due to various historical and sociocultural factors: the understood slang items may be a generational leftover from a previous generation. Adults may understand the slang simply because they have to interact with their children of this age or they may be exposed to it through films and television. In any case, age-based understanding differences indicate a different kind of standardization from that found in dictionaries, namely one that is fluid and ever changing, based on personal interactions as well as media exposure.

The third grouping provided by Allam is based on differences in use between the younger and older generations, and it is here that the true nature of these terms as teenage slang is made clear. While almost all of the terms are used by a large majority of teenagers (the exception being *mixaryaṭ*), 12 out of the 20 terms are not used at all by the older generation, while the rest are used by only a very few of the older individuals questioned. The most commonly understood slang term among the older group (*inšakaḥ*, understood by 92 percent of those questioned) is only used by one-third of them. The nonuse of these teen slang terms, even those ostensibly understood by adults, gives perhaps the clearest indication of what slang is about, namely to bolster a sense of group identity among a marginal social group. Use of these terms among adults is inhibited either due to lack of understanding of the terms by their association with a youth subculture or by their being associated with immoral or sexual connotations. An indication of the importance of these kinds of connotations may be seen in the one term which was used by the largest percentage of the adults questioned, namely *kabbar* (used by 40 percent of the adults). Perhaps because of its probable relation with the Muslim invocation *allāhu ʾakbar*, it is much more readily accepted among adults than are the other terms understood by adults, which may have (in their mind) a hint of immorality or a sexual connotation.

In addition, it is interesting to note that category (iii) may not be as unrepresentative in the Arabic lexicon as Allam imagines. In other dictionaries there exist root entries similar to these particular lexical items. For example, *ʾantax* may be seen as derived Form VIII from the root *n-x-x* (with the meaning 'to kneel [camel]; to become tired'), making it originally *intaxx*, which has become reduced in the crucible of slang to *ʾantax*). The lexical item *mifayyiṣ* as well may be a reflex of an ancient Arabic root, which has the connotation of 'to cease; to tire of; to finish'. Under the root *f-y-ṣ* in Lane's *Lexicon* and in the *Lisān al-ʿArab* are listed the following lexemes: *fāṣa* (*wallāhi mā fiṣtu* 'by God I did not cease'; *mā fiṣtu ʾafʿalu* 'I have not tired of doing it'), *ʾafāṣa* (*ʾafāṣa bi-bawlihi* 'he ejected all his urine'), and *mafyaṣ* (*mā ʿanhu mafyaṣ* 'there is no place to quit it'). The word *hartil* has a much less direct but much more interesting (possible) etymology. The first thing to note about this root is that it shares its first two consonants with a very large number of lexemes whose meaning involves detailing the negative qualities of a person or situation, often involving senility, weakness, and disease. Secondly, the word *harkal* 'to become old, decrepit' (listed in Hava, Lane, and the *Lisān al-ʿArab*) may be seen as a possible historical source for the present *hartil* 'to go crazy, talk nonsense', due to the similarity in meaning, as well as in form (note that the phonemes /k/ and /t/ are acoustically very similar and are related even in Arabic verbal morphology as markers of the 2nd person). Rather than referring the origins of *hartil* and similar words back to the biradical hypothesis (which is the most common response), it seems preferable to see them as the result of two phenomena, phonaesthetics and the presence of cant or argot. While the notion of phonaesthemes cannot be explained here in detail (→ sound symbolism), suffice it to say that it has to do with the grouping of lexemes in Arabic based on shared root consonants (usually the first two) as well as a shared semantic field, something which has been remarked upon for centuries. One of the first and best-known analyses of this was given by Ibn Jinnī (d. 392/1002) and his theory of *al-ištiqāq al-kabīr*, in which he tried to extend a semantic value not just to clusters of two root consonants but to individual consonants themselves (*Xaṣāʾiṣ* II, 133–139). More recently, Arabists and Semiticists have explained this phenomenon as being due to the fact that Semitic roots were originally biradical, with the third consonant added on to provide a nuance of meaning (→ biradicalism). An alternate way to understand this phenomenon is to say that these groups of lexemes have developed over the centuries as a normal part of language evolution and change, not necessarily deriving from the same root but rather converging on a group of sounds (phonaesthemes), which have become associated in a speech community's practice with a particular semantic value (e.g., in English the /sl-/ complex is found in a great many words which denote a wet, slippery texture). While this proposal is still somewhat tentative and in need of further research, it is quite possible that such processes have been involved in the development of slang and may have played an important role in the development of Arabic lexicon, which is a major point of Allam's study.

The second phenomenon of relevance to the discussion of these slang terms and how they came to be is cant or argot, which is the specialized vocabulary of specific social groups. While it is difficult to say, based on the available evidence, how many of the above slang terms derive from more established and underground argots, such a possibility exists and should be considered before being discarded. However, even if such a source exists, one should not overlook the great differences between slang and argot. Slang is ephemeral, argot more long lived, and slang usually involves native words, either semantically reconceived or conceived anew, whereas argot quite often is based on foreign language sources. Both tend to be confined to 'in' groups, but of quite different types. Particular types of slang may be shared by vaguely defined groups in a society, with a great deal of variability in the size and extent of the group at any one time, while argot tends to be confined to clearly identifiable groups of individuals who use it among themselves both to bolster their group identity and also to hide aspects of their interaction from outsiders with whom they deal (→ jargon).

In contrast to the lack of research on contemporary slang in Arabic, there have been a number of studies of argot in Arabic, especially in the medieval period. The best of them is undoubtedly Bosworth (1976), who in the final chapter examines the jargon of the medieval genre of *qaṣīda sāsāniyya*, poems describing the underworld of thieves and robbers in the urban centers of the Islamic world. He categorizes the jargon words based on their origin and how they arrived at their form or meaning as presented in the texts, including derivations based on phonetic, morphological, or semantic changes as well as those derived from loanwords, summarized in Table 2 and Table 3. He, too, notes the lack of a widely accepted Arabic term for 'jargon', although the term *munākāt Banī Sāsān* (= *munāġāt*) functions essentially in the same way that 'jargon' does, but, he notes, it was not used widely in medieval times nor is it used in such a way in modern times.

For the modern period, Rowson (1982) deals with cant and argot (Arabic: *sīm* or *sīn*) in Cairene Arabic. Rowson (p.c.) distinguishes 'cant' (the special language of a marginal group) from 'argot' (the secret language of a professional group), but in Rowson (1982), the terms are used interchangeably. He begins by noting the long life of some of these argot words, having found reflexes of them in works (some detailed in Bosworth 1976) from the 10th century (the *Qaṣīda sāsāniyya* of ʾAbū Dulaf al-Xazrajī), the 13th century (Ibn Daniyāl's shadow plays), the 14th century (the *Qaṣīda sāsāniyya* of Ṣafī d-Dīn al-Ḥillī), and the 16th century (aš-Širbini's *Hazz al-quḥūf*). He then delimits several major argots in Cairo (ca. 1982). The first two are more basic, while the next two are derived from them, with the rest being either too difficult to record or on their way to extinction (summarized in Table 4 with examples in Table 5).

The first of these argots, *sīm is-sāġa* (*sīm it-tuggār, sīm xān il-xalīli*) is Hebrew-based, as evidenced in the use of Hebrew numbers, when they are not too close to Arabic. The second type, *sīm il-fannānīn* (*sīm il-ʿawālim, sīm il-mumassilīn, bitūʿ il-malāhi*) is based on the → Gypsy language (Rom, Romany, or Rotwälsch), with a great number of Italian words apparently derived from Italian *commedia dell'arte* troupes that performed in the Egyptian countryside in the 19th century. The third type, *sīm il-xartiyya* (for which no examples are given), is derived

Table 2. Jargon terms from ʾAbū Dulaf's *Qaṣīda sāsāniyya* (Bosworth 1972)

	English translation	Origin
Phonetic change		
muqanwin	'who claims to be Muslim convert from *ʾahl al-kitāb*'	*qawnana* 'to punish'
faššaša	'to fart in a mosque to annoy worshippers'	*fašša* 'to fart softly'
bašraka	'to dress like ascetic holy man'	*b-š-r* 'to put close to body'
Semantic change		
ḏarraʿa	'to beg for food from *harīsa* maker'	'to stretch forth arms'
kabbasa	'to extort money from someone who cashed check'	*kabasa* 'to seize by force'
šaṭṭaba	'to mutilate oneself and allege bandits attacked'	'to scar; to let blood'
fakkaka	'to escape from bonds as feat of skill'	'to separate two things fastened'
xušni	'nonbeggar, outside beggar's circle'	*xašin/xušn* 'rough, coarse'
kudda	'woman who begs with husband in mosque'	*kadda* 'to labor, exert oneself'
Semantic transfer		
matr	'coitus'	*matara* 'to stretch a rope taut'
ʾabū šukr	'salt'	'father of gratitude'
ʿallāfa	'female breadwinner because of begging'	*ʿallāf* 'fodder merchant'
labūsa	'vulva [pubic hair]'	'what covers naked body'
bāz	'skillful beggar woman'	'hawk'
ṣaqr	'skillful beggar woman'	'falcon'
ġāliyāt al-juhr	'feces'	'precious unguent of anus'
qaṣr	'stokehold of a bath'	'palace'
iṣṭabl	'mosque'	'stable'
šaġāṯa	'mosque'	*šaʿiṯa* 'to be disordered, ruinous'
ʾabū mūsā	'nonbeggar'	[proper name]
Mālik	'stoker of the bath'	'angel Malik, guardian of hell'
Loanwords		
qaynūn	'place where beggars meet to share profits'	Greek *κοινωνία* 'community'
šawlasa	'to beg in ascetic guise'	Greek *σάλος* 'idiot, fool'
qannāʾ	'pretended convert who recites non-Muslim scriptures'	Hebrew *qannāʾ* 'zealot'
hāḏūr	'circle of fortune tellers and tricksters'	Aramaic *hādūr*
kāġ/kāġa	'pretended lunatic'	Persian *kāġa* 'mad'
buštadāriyyūn	'porters who transport beggars' booty'	Persian *pušti-darī* 'he carries on back'
xušbūyi	'drugged stew'	Persian *xūš-būi* 'aromatic'
kurs	'hunger'	Persian *gurs*

primarily from the first type (*sīm il-sāġa*), with an admixture from the other types as well. The fourth type (*sīm il-kawanīn*), on the other hand, is derived from the second type (*sīm il-fannānīn*), with much more development and a good number of French and English loanwords. Meager information is given about the fifth type, *sīm il-ḥarāmiyya* (*sīm in-naššālīn, sīm is-sakka*), which is said to be due to the difficulty in finding informants, and very little information is given about the sixth type as well (*sīm il-miʿarriṣīn*), which Rowson speculates might be slowly dying out due to the changing nature of prostitution in the latter half of the 20th century.

Rowson's article provides only a tantalizing glimpse at the argots of these groups, and it is hoped that more information will become available in the future. More research must be done in this area to delineate and clarify not simply the inventory of argots but their social and individual function and use as well. Rowson makes a few intriguing observations touching on this point, such as the fact that the *sīm al-xawanīn* seems to be used more among members of its in group even when outsiders are not around, which is what one finds with slang rather than argots. Examining both modern slang and argots in Arabic in light of more recent work in sociolinguistics and anthropological linguistics

Table 3. Jargon terms from Ṣafī d-Dīn's *Qaṣīda sāsāniyya* (Bosworth 1972)

	English translation	Origin
Phonetic change		
q.ṯyān	'beggar chiefs'	*qitnān* 'groups'
nāṭūr	'swindler's accomplice in audience'	Aramaic 'watcher'
al-hāt	'the tongues'	*al-lahāt* 'uvula'
Morphological change		
ṣūfān	'Sufi shaykhs'	Arabic/Persian (?) plural ending /-ān/
qalī	'storyteller'	Persian (?) /-ī/
talaġġa	'to speak'	*laġā* 'to talk, babble'
mayyama	'to convert to the faith of Muḥammad'	/m/ < Muḥammad
tamyīm	'claim of ability to find buried treasure'	/m/ < *muṭālib* 'treasure seeker'
Semantic change		
nas(a)b	'house of God' > 'Kaʿba'	'house'
kassāb	'astrologer'	'person who earns living'
Semantic transfer		
šaḥḥama	'to gain someone's favor'	'to fatten up'
ṣahl	'horse'	*ṣahala* 'to whinny, neigh'
muzaffat	'destitute, penniless'	'smeared with pitch'
bāxīs	'one [number]'	*baxasa* 'to diminish, deprive'
ḏābūl	'whore'	*taḏabbala* 'to walk with mincing step'
ḏābil/ḏablān	'whore'	'having languid eyes, coquettish'
qašm	'(immature) boy'	'unripe but edible dates'
šamūl	'bread'	'all-embracing, staff of life'
qarūda	'rat'	'the gnawing one'
laqy	'copulation'	'act of meeting someone'
taʿdīl	'copulation'	'making straight, level'
malāqim	'mouths'	'place where one takes in mouthfuls'
nayafa	'hunger'	*nayyif, nayf* 'excess'
zīh	'hashish'	uncertain
marqān	'dinar'	uncertain
qārūb	'beggar chief'	uncertain
kazākī	'governors, rulers'	uncertain
bazwān	'Persian; Kurd'	uncertain
karajīm	'the jinn'	uncertain
ḥamdānī	'pimp'	uncertain
dabašrī	'ox'	uncertain
Loanwords		
hankama	'to gather round'	Persian *hangāma* 'assembly'
buštadār	'slaveboy'	Persian *pušti-darī* 'carries on back'
q.nta	'town'	Persian *kant* 'town'
k.b.š.t.r	'camel'	Persian *uštur* 'camel'
raxtānī	'itinerant drug peddlar'	Persian *raxt* 'goods, equipment'
barġāšāt	'ears'	Persian *bar gūš āmadan* 'hear'
buštakānī xurda	'peddlar of trashy goods'	Persian: *pušti-darī xurda* 'carries on back small object'
jarraxa	'to dance'	Persian *čarx* 'circle of dancers'
xandaja	'to laugh'	Persian *xandagī* 'laughter'
mard	'dirham'	Persian *mard* 'man'
tank	'copper coin'	Armenian *tank* 'one quarter of dirham'
šann	'two [number]'	Hebrew *šᵊnayim* 'two'
dast	'ten'	Indo-European *dekṃ* 'ten'
manj	'divination tool of astrologers'	Greek *μάγγανον* 'machine'
qalaftūriyya	'talismanic formula'	Greek *φυλακτήρια* 'writing of phylacteries'
mufawwil	'one who draws omens'	Middle Arabic *mufaʾʾil*

Table 4. Cairene Argots (Rowson 1982)

1. *sīm is-sāġa* (*sīm it-tuggār*, *sīm xān il-xalīli*)	the argot of the goldsmiths, silversmiths, and merchants in the Khan il-Khalili neighborhood
2. *sīm il-fannānīn* (*sīm il-bawālim, sīm il-mumassilīn*, *bitūʿ il-malāhi*)	the argot of the musicians and entertainers who perform at moulids and weddings
3. *sīm il-xartiyya*	the argot of the young men who attach themselves to tourists and try to profit from them in some way
4. *sīm il-kawanīn*	the argot of homosexuals
5. *sīm il-ḥarāmiyya* (*sīm in-naššālīn*, *sīm is-sakka*)	the argot of criminals, thieves, pickpockets
6. *sīm il-miʿarriṣīn*	the argot of pimps and prostitutes

Table 5. Terms of Cairene argots (Rowson 1982)

	English translation	origin
sīm is-sāġa		
maʿūṯ	‘money’	uncertain
yāfit	‘good, rich; in buying mood’	Hebrew
ḥamišša	‘five [number]’	Hebrew
šimunya	‘eight [and most other numbers]’	Hebrew
kwatrīn	‘forty’	French
hāt il-gaft	‘you are being indiscreet, silent’	*jaffata* ‘to hide’
hāt iššāy min ʿand abu yaʿʾūb	‘don’t bring tea’	Arabic
ʾarnab	‘one million pounds’	Arabic
habastaʾān	‘hashish’	uncertain
šall	[reflexive]	Hebrew
iddafš illi fi šallak ašfūr	‘the guy next to you is bad’	Hebrew
sīm il-fannānīn		
bišbāš	‘mustache’	Gypsy
ʾamr	[reflexive]	Gypsy
šallaftu barġašāti ʾamri	‘I’ve hurt my own ears’	Gypsy
fi ʾamr	*ʿandi* ‘with me’ (Kahle 1927)	Gypsy
il-barġal illi f-amrak šalaf	‘the guy next to you is bad’	Gypsy
kāra	‘buttocks, anus’	(Ibn Dāniyāl)
furti	‘hurry up!’	Italian
salūti	‘leave taking’	Italian
ʾamarūz	‘protagonist’	Italian
ʾamriz	‘to play the lead’	Italian
sīm il-kawanīn		
kawanīn (vs. standard *xawal*)	‘homosexual(s) [noun and adj.]’	Arabic (?)
da kawanīn	‘he is a homosexual’	
biyiʿmilu kawanīn	‘they perform homosexual acts’	
kawnana	‘homosexuality’	
ikkawnin	‘to participate in homosexual act’	
barġal	‘man [active partner in act]’	*sīm al-fannānīn*
kudyāna	‘woman [passive partner in act]’	*sīm al-fannānīn*
dublifās/dabalfās	‘both active and passive partner in act’	French/English
dinyāra	‘woman [real]’	
kēl	[intensifier]	French *quel*
kēl ʿala	‘(I) was so upset at…’	French
nās ʾamēh	‘nonhomosexuals’	French
tīb	‘type’	English
huwwa tībi	‘he’s my type’	English
slō	‘romantic’	English
biznis/bazānis	‘picking up tourists for sex’	English
sīm il-ḥarāmiyya		
baġbaġān	‘transistor radio’	
tirmisa	‘watch’	
bazargi	‘policeman’	
gaft /ʾamēh /ništa /naḥw	‘cheese it!’	
ʾammin	‘to pat a pocket’	
fardit šuġl	‘gold-plated brass necklace’	

(especially as indicative of alternative regimes of authority) is greatly needed and will shed much-needed light on the entire spectrum of Arabic linguistic practice, both high and low.

Bibliographical references

Primary sources

Ibn Jinnī, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. Ed. Muḥammad ʿAlī an-Najjār. 3 vols. Cairo, 1952–1956. (Repr., Beirut: Dār al-Hudā, n.d.)

Širbīnī, *Hazz* = Yūsuf aš-Širbīnī, *Hazz al-quḥūf fī šarḥ ʾAbī Šādūf*. Cairo, 1963. [Critical ed. Humphrey Davies. Louvain: Peeters, 2005.]

Secondary sources

Allam, Jehan. 2000. "Youngsters and new vocabulary". *Language Series of Papers in Linguistics* (Cairo Linguists Group, Arab Research Center) 2.112–135.

Barbot, Michel. 1974. "Notes lexicographiques sur les orfèvres et bijoutiers de Damas". *Arabica* 21.72–83.

Ben Cheneb, Rachid. 1942. "L'argot des arabes d'Alger". *Revue Africaine* 86.72–101.

Ben Cheneb, Saadeddine. 1941. "L'argot a l'époque classique". *Bulletin des Etudes Arabes* (Algiers) 1:4.109–110.

Bosworth, C. Edmund. 1976. *The medieval Islamic underworld*. Leiden: E.J. Brill.

Cohen, Marcel. 1919. "Note sur l'argot". *Bulletin du Société de Linguistique* 21.132–147. (Repr., *Cinquante années de recherches linguistiques, ethnographiques, sociologiques, critiques, et pédagogiques*, 133–139. Paris, 1955.)

Kahle, Paul. 1926. "Eine Zunftsprache der ägyptischen Schattenspieler". *Islamica* 2.313–322.

Littmann, Enno. 1920. *Zigeuner-Arabisch*. Bonn: K. Schroeder.

Parkinson, Dilworth. 1988. "Egyptian Arabic abuse". *Maledicta* 10.145–161.

Rowson, Everett. 1982. "Cant and argot in Cairo colloquial Arabic". *American Research Center in Egypt Newsletter* 122.13–24.

Sergeant, R.B. 1948. "'Cant' in a contemporary South-Arabic dialect". *Transactions of the Philological Society* 121–126.

JOHN EISELE (The College of William and Mary)

Slavonic Languages

1. Introduction

The number of words of Arabic origin (or words borrowed from other sources via Arabic) in the Slavonic languages differs from one language to another. There are, for instance, more than fifty Arabic loanwords in Czech (Machek 1968; Rejzek 2001), approximately seventy in Russian (Vasmer 1950–1958), almost three hundred in Macedonian (Jašar-Nasteva 2001), and more than four hundred in Bulgarian (BER 1962ff.) and Serbo-Croatian (Skok 1971–1974). Although direct contacts between Slavs and Arabs are documented as early as the 7th century C.E. (Theophanes, *Chronographia*), most borrowings were mediated. There were two dominant routes of mediation:

i. Western: Arabic > (Spanish) > French or Italian > (German or Dutch) > (especially) West and East Slavonic;
ii. Eastern: Arabic > (Persian) > Turkish (exceptionally Tatar) > (Middle Greek) > (especially) South and East Slavonic.

In the case of Bulgarian and Serbo-Croatian, the Arabic loanwords were mostly transmitted through the mediation of Ottoman Turkish during the period when the speakers of these languages were subjects of the Ottoman Empire. In some cases, the Arabic words also penetrated into Italian or French through Turkish. There are also cases in which the Arabic loans into Serbo-Croatian were mediated by Italian or borrowed directly, especially in Bosnia.

Almost all loanwords concern nouns, but at least one grammatical item was borrowed: *ʾammā* 'but; but as to' > Turkish *ama* > Serbo-Croatian *àma* 'but; just', Bulgarian *amà* 'id.' (Sk I, 32; BER 10).

2. The western route

To the western route of borrowing belong the Arabic loanwords that were borrowed into European languages, usually through the languages of the Iberian Peninsula, or through direct trade with Italy (→ Italian; → Ibero-Romance; → English).

One category of loanwords is that of words connected with Islam. Common European terms of Arabic origin connected with Islam have been omitted here (e.g. *Allāh, Islām, Qurʾān, šarīʿa(t)*, etc.). Other loanwords are:

> *badāwiyy* ~ *bidāwiyy* 'inhabitant of the desert' (from *badw* 'desert') > colloquial Arabic *bedāwī* > Old French *beduin*, French *bédouin* > German *Beduine* > Czech, Russian *beduín* (V I, 143).
>
> *ḥarām* '[that which is] prohibited', hence 'women's apartments' (from *ḥarama* 'to prohibit') > Turkish

harem ‘Frauenräume, Privaträume’ > French *harem* > German *Harem* > Czech *harém*; Turkish > Serbo-Croatian *hārem*; French > Russian *garém* (R 195; V I, 393, IV, 265; KS 356; H 210; St 39).

kāfir, pl. *kifār* ‘infidel; renegade’ > Persian *gäbr* > Medieval Turkish *giaur* > Czech *d’aur*; Russian *gjaúr*; Serbo-Croatian *đàurin* (V I, 480; St 33).

mamlūk ‘possessed; slave, servant’ > Italian *mammalucco* > German *Mameluck* > Old Czech *mamalík*, Czech *mameluk* ‘a member of the Egyptian sultan’s guard’ (M 349; KS 536).

masjid ‘mosque’ > Early Modern High German *Meschit* or Early Italian *meschita* > Czech *mešita* (R 374); (Russian *mečet’* probably through Ottoman Turkish *mescit*).

nawwāb ‘deputy’ > Urdu *nawāb* > Portuguese *nababo* > English *nabob* ‘Muslim official acting as deputy governor in the Mogul Empire; rich person’ > Czech *nabob* (R 400; H 308).

Some of the Arabic loanwords are related to military organization, weapons, etc.

ʾamīru l-baḥri ‘admiral’ lit. ‘commander of the sea’ > Old French *amiral* > French *admiral* (contaminated with Latin *admirārī* ‘to admire’) > German *Admiral* > Czech *admirál*; Russian *admiral* (R 45; V I, 62; H 6).

al-qubba(t) ‘cupola, vault; alcove; tower; cathedral, dome; tent, tabernacle; sunshade; palanquin’ > Spanish *alcoba* > French *alcôve* > German *Alkove* ~ *Alkoven* > Slovak *alkovňa*, Czech *alkovna* ‘room without windows’ (M 35; KS 27); without the article: *qubba(t)* > Turkish *kubbe* > Serbo-Croatian *kùbe* ‘vault’ (St 48).

fulūka(t) ~ *fulūqa(t)* ‘ship, barque, felucca’ (< Greek *ephólkion*) > Spanish *faluca* > French *felouque* > Russian *feljúga*, South Russian *felúka* ‘small Turkish boat’ (V IV, 189).

manāra(t) ‘lighthouse; minaret of a mosque’ (from *nār* ‘fire’) > Turkish *minâre* > French *minaret* > German *Minarett* > Czech, Polish *minaret* ‘tower of a mosque’ (R 379); Turkish > Serbo-Croatian *minare* ‘id.’ (Miklosich 1888, 85).

taʿrīf ‘instruction; description, definition; tariff’ (from *ʿarrafa* ‘to notify, make known’) > Turkish *tarife* > Italian *tariffa* > French *tarif* > German *Tarif* > Czech *tarif* (R 651; KS 816).

šayx ‘old man; head of a family or tribe’ > French *cheik* > English *sheik* > Czech *šejk* (R 626).

qālib ‘mold for casting metal’ from *qalaba* ‘to turn, convert’ (or < Greek *kālopódion* ‘Schusterleisten’) > Italian *calibro* > French *calibre* > German *Kaliber* > Czech *kalibr* ‘caliber’ (R 258; KS 419; H 59).

A large number of loanwords are connected with science, especially pertaining to the calendar, astronomy, geography, and medicine.

al-jabr ‘predestination; algebra’, *jabr* ‘force; reunion of what has been separated; reduction of fractures’ (from *jabara* ‘to set a broken bone; to contract; to help’) > Medieval Latin *algebra* > German *Algebra* > Czech *algebra* (R 50; KS 26).

al-kīmiyāʾ ‘philosopher’s stone; elixir’ (perhaps < Greek *chēmeía* ‘art of transmuting metals’) > Spanish *alquimia* > Medieval Latin *alchimia* > German *alchimie* > Czech *alchymie* (KS 25; H 10).

al-manāx ‘calendar’ > Spanish *almanaque* > French *almanac* > Middle Low German *almanak* > German *Almanach* ‘yearbook’ > Czech *almanach* (R 50; KS 29; ML 5281).

as-sumūt ~ *as-simūt* ‘azimuth’ (pl. of *samt* ‘road’) > French *azimut* > Czech *azimut* ‘id.’ (R 64; H 30).

al-Xwārizmī [the name of a 9th century C.E. Arabic mathematician from Khwarezm] > Medieval Latin *algorismus* > Middle High German *algorismus* (contaminated with Greek *arithmós* ‘number’) > German *Algorithmus* > Czech *algoritmus* (R 50; KS 26).

naḏīr as-samt ‘opposite, in front of; parallel; nadir’ > French *nadir* ‘point opposite to the zenith’ > Czech *nadir* ‘id.’ (R 401; H 308).

ramaḍān ‘the [ninth] month of fasting’ > Medieval Turkish *ramadan* > Czech *ramadán*; Serbo-Croatian *ramàdān*; Arabic > Turkish *ramazan* > Russian *ramazán*, Serbo-Croatian *ramàzān* (V III, 440; St 58).

samt ar-raʾs ‘path over the head’ > Medieval Latin, Old French *cenit* ‘zenith’, i.e. ‘point of the sky directly overhead’, Italian *zenit* > German *Zenit* > Czech *zenit* (R 738; H 551; KS 907).

ṣifr, pl. *ṣifrāt* ‘zero’, from *ṣifr* ‘empty’ (the idea of ‘zero’ was borrowed from India, cf. Sanskrit *śúnya-* ‘empty; zero’) > Late Latin *cifra* ‘zero; figure’ > Late Middle High German *zif(f)er* ‘zero’, German *Ziffer* ‘figure’ > Slovak, Czech *cifra*, Upper Sorbian, Polish *cyfra*; Russian *cifra* ‘figure’ (M 85; V IV, 303); independently Arabic *ṣifr* > early Italian *zefiro* > *zero* > French, English *zero*, and Arabic *ṣifr* > Medieval Latin *cifra* > Old French *cifre* ‘zero; figure’ > French *chiffre* ‘figure, number; secret code’ > German *chiffre* ‘secret code’ > Czech *šifra* ‘id.’; Russian *šifr* (KS 910; 153; V IV, 444).

Many loanwords are names of minerals, plants, fruits, or animals.

al-birqūq ‘apricot, yellow plum’ > Spanish *albercoque* ~ *albaricoque* > French *abricot* > Dutch *abrikoos* > Russian *abrikos* ‘apricot’ (V I, 57).

al-kuḥl ‘antimony in powder form to paint the eyelids black; collyrium’ > Spanish *alcohol* ‘fine metallic powder’ > English *alcohol* ‘fine metallic powder, especially as produced by sublimation (16th c.); distilled or rectified spirit of wine’ (from 18th c.), German *Alkohol* ‘reiner Weingeist’ > Czech *alkohol*; Russian *alkogol’* etc. (V I, 71; KS 27).

al-qily ‘Laugensalz’ > colloquial Arabic *al-qalī* > Spanish *álcali* > French *alcali* > German *Alkali* > Czech *alkaloid*, adj. *alkalický* (R 50; KS 26–27).

ʾaṣfarān ‘saffron’ (from *ʾaṣfar* ‘yellow’) > Latin *safranum* > Bulgarian, Serbo-Croatian *čafran*, Slovenian *žefran*; Slovak *šafran*, Old Czech *šafrán ~ šefrán*, Czech *šafrán*, dialectal (South) *šefrán*, Polish *szafran ~ szefran*; Russian *šafran*; the forms with *-e-* indicate Turkish mediation (M 600). But the source could also be sought in Arabic *zaʿfarān* ‘saffron’, with the following chain of borrowings: Spanish *azafrán* > Old French *safran* > Middle High German *saf(f)rān* ‘id.’ (KS 699; ML 9588).

bab(b)aġāʾ ~ babaġāl ‘parrot; nightingale’ > Spanish *papagayo* > Old French *papegai* > Middle Low German *papagoie*, Middle High German *papagey* > Bulgarian *papagál* (cf. Italian *pappagallo*), Serbo-Croatian, Slovenian *papiga*; Slovak *papagáj* (maybe via Hungarian *papagáj*), Old Czech *papúch*, Czech *papoušek*, Upper Sorbian *papaguj*, Polish *papuga*; Ukrainian *papúha*, Russian *popugáj* ‘parrot’ (M 433; V III, 328; KS 610).

bādinjān ‘love apple (Melongena)’ > Turkish *patlıdžan* > Russian *baklažán* > Czech *baklažán* ‘eggplant’, besides dialectal Russian (Astrakhan) *badaržán* (M 43; V I, 110).

būraq ‘borax, saltpeter’ > Persian *bûräh* > Russian *burá*; Arabic > Medieval Latin *borax* > German *Borax* > Czech *borax* (V I, 242; KS 126).

lāzward ‘lapis lazuli’, *lāzwardiyy* ‘azure blue’ > Medieval Latin *lazurium ~ lasurium* ‘lapis lazuli’, besides *azurum* (with the lost *l-*, identified with the Romance definite article) > French *azur* > Czech *azur* ‘azure blue’ (R 64; KS 504).

tūtiyāʾ ‘zinc’ > Qumanic *tutiya* > Italian *tuzia* > Russian *túcija* (V IV, 128; Räs 502).

ḥašīš ‘grass, hay, herb; intoxicating extract of hemp’ > English *hashish* (from 16th c.), German *Haschisch* (from 19th c.) > Czech *hašiš* ‘a drug from hemp’ (R 196; KS 358; H 211).

jarnayṭ ‘civet cat’ > Spanish *ginetta* > French *genette* > Dutch *genetkat*, German *Genettkatze* > Russian *jenót* ‘Procyon lotora’ (V II, 20; ML 3943b).

kāfūr ‘blossom or spathe of the palm tree’ (< Prakrit *kaṃpura-* < Sanskrit *karpūra-*) > Spanish *canfora ~ alcanfor* > French *camphre* > Old Italian *cafura* > Italian *canfora* > Russian *kamforá* (V II, 176; ML 4656).

lakk ‘a plant for dyeing goatskins; mixture; lacquer’ > Italian *lacca* > German *Lack* > Czech *lak* ‘paint, lacquer’; originally from Middle Indian (Pali *lākkhā-* < Sanskrit *lākṣā-*), perhaps through Persian *lāk* (M 318; KS 498).

Arabic/Persian *nāranj* ‘Seville orange, bitter orange’ > Old Russian *naranža* (V III, 43); cf. also Italian *arancio* > Provençal *arange* > French *orange* > Czech *oranžový* ‘orange color’, and further Italian *pomarancia* ‘orange’ (*pomo* ‘apple’) > Viennese German *Pomerantsche* > Czech *pomeranč* (R 431, 486).

rībās ‘a kind of sorrel’ (< Persian *rībās* ‘a kind of rhubarb’) > Italian *ribes* > German *Ribis(el)* ‘Johannisbeere’ > Slovak *ribezl’a*, Old Czech *rybés*, Czech *rybíz* ‘currants’ (M 526; KS 685).

suwwād ‘name of the plant from whose ash soda was produced’ > Spanish *soda* > German *Soda* ‘Natriumsalz der Kohlensäure’ > Slovak *sóda*, Czech *soda* ‘id.’ (M 565; KS 768).

sarxūn ‘estragon (Artemisia dracunculus)’ > Medieval Latin *tarcon* > Old French *targon* > French *estragon* > Russian *ėstragón*; Arabic > Turkish *targun* > Russian *turgún*, Ukrainian *turhún* ‘Artemisia dracunculus’ (V IV, 522; ML 8581).

ṭalq ‘talc’ > Spanish *talco*, French *talc* > German *Talk* > Slovak, Czech *talk*, Polish *talek*; Russian *tal’k* (M 635; KS 814).

ʿuṭārid ‘quicksilver’ > Turkish *utaryd* > Slovak *ortut’*, Old Czech *rtut*, Czech *rtut’*, dialectal (Sušice) *trut’*; Polish *rtęć*; Ukrainian, Russian *rtút’* (M 522; otherwise V III, 509–510).

zadwār ~ zidwār ‘Zitterwurzel’ (< Persian *zädwār*) > Medieval Latin *cedoarium* > Italian *zettovario* > Old High German *zit(a)war*, Middle High German *zitwar*, German *Zitwer* ‘aromatisch duftendes Kraut’ > Bulgarian *citvar*; Slovak, Czech *cicvár*, Czech also *citvar*, Lower Sorbian *cytwar*, Polish *cytwor, cytwar* (M 85; KS 913; ML 9617a).

zaytūn ‘olive’ > Turkish *zejtin* > Serbo-Croatian *zéjtin* ‘id.’; Crimean-Tatar *zäitin* > Russian *zétin, zitín* ‘id.’ (St 72; V II, 96).

zabād ‘civet’ > Medieval Latin *zibethum*, Italian *zibetto* > Serbo-Croatian *cibet* ‘id.’ (Sk I, 259; H 78).

zur(r)āfa(t), pl. *zurāfa* ‘giraffe’ > Italian *giraffa*, French *girafe* > German *Giraffe* > Slovak, Czech *žirafa* ‘id.’ (M 728; KS 325; ML 9632b).

A number of loanwords are connected with food, drinks, and spices.

ʿaraq ‘milk, juice, sap, date honey; distilled liquor, brandy’ > French *arac* > Czech *arak*; Arabic > Turkish *raky* ‘brandy’ > Serbo-Croatian *ràkija* ‘brandy’ > Czech *rakije*; Russian *raká* (M 38; R 57; V III, 438; St 57–58).

ʾisfināj ~ ʾisfānāx ‘spinach’ (< Persian *ispanāj ~ ispānāx*) > Spanish *espinaca* > Middle German *spināt*, German *Spinat* > Czech *špenát* ‘id.’ (R 638; KS 779).

kabar, pl. *kibār* ‘capers [spice]’ > Greek *kapparis* ‘Capparis spinosa’ > Medieval Latin *capparis* > German *Kaper* > Old Czech *kapar* ‘a kind of spice’ (M 239; KS 424).

kubāba(t) ‘cubeb’ > Spanish *cubeba* > French *cubèbe*, Italian *cubebe* > German *Kubebe* > Russian *kubeba* ‘a kind of brandy’ (V II, 396; ML 4788c).

qahwa(t) ‘coffee’, pl. *qahāwī* ‘coffeehouse’ or ‘coffee room’ > Turkish *kahve* > Hungarian *kávé* or Serbo-Croatian *káhva ~ kava* > Czech *káva*, while forms such as Russian *kófe* were mediated by Armenian, through Italian *caffé*, French *café* > German *Kaffe*, Dutch *koffie*, English *coffee* (M 246; KS 416; St 47; Räs 221).

rubb 'jam, syrup' > Old Italian *robbo* > French *rob* > German *Robb* > Polish *rob* 'apothecary's jam' (M 516; ML 7401b).

samīd 'white bread; finest flour' > Turkish *simit* 'roll, bread crumbs' > Bulgarian *simìd* 'id.'; + suffix of the *nomen agentis -či* ⇒ *simitči* > Bulgarian *simì(t)če*, Serbo-Croatian *simìčija* 'baker who bakes rolls' (St 60–61; BER VI, 653).

šarāb 'drink' (from *šariba* 'to drink') > Medieval Latin *sirup(p)us*, *sirop(p)us* > Middle German *sirup(e)*, *sirop(e)*, German *Sirup* > Czech *sirup* (R 573; KS 765).

sukkar 'sugar' > Greek *sákkharon* > Medieval Latin *saccharon* > Italian *zuchero* > Middle High German *zucker* > Serbo-Croatian *cukar*, Slovenian *cuker*; Slovak *cukor*, Old Czech *cukar* ~ *cuker*, Czech *cukr*, Upper Sorbian *cukor* ~ *cokor*, Lower Sorbian *cukor*, Polish *cukier* > Ukrainian *cukor* 'sugar'; Arabic < Persian *šäkär* < Sanskrit *śarkarā* 'Sandzucker' lit. 'Kies, Griess' (M 90; KS 916).

Three loanwords denote vessels or dishes.

al-ʾanbīq ~ *al-ʾanbīk* 'alembic' (< Greek *ámbiks, ámbikos* 'cup, cap of still') > Spanish *alambique*, French *alambic* ~ *alembic* > Bulgarian *alambik*; Czech *alambik, alembik*; Russian *alembik*; Arabic > Italian *lambicco* > Serbo-Croatian *lambik* (M 35; ML 442).

ġarrāfa(t) 'carafe' > Spanish *garrafa* > Italian *caraffa* > French *caraffe* > German *Karaffe* > Bulgarian *karafa*; Slovak, Czech, Polish *karafa* 'id.' (M 241; KS 426; H 62; ML 3690b).

ṭās 'cup, saucer; plate, tray' (< Persian *täšt* 'Becken, Schale') > Italian *tazza* > German dialectal *Täzze* > Slovak *táca*, Czech *tác* 'tray', Polish *tac(k)a*; Arabic > French *tasse* > German *Tasse* 'cup' (M 633; KS 816).

Another group of loanwords contains terms for cloth, dress, jewelry, and cosmetics.

ʾaṭlas 'satin' lit. 'fine' > French *atlas* > German *Atlas* > Czech *atlas* 'satin' (M 39; KS 59).

barrakān 'cloth from camel hair' > Spanish *barragán* > Portuguese *barracan*, French *baracan* > Middle High German *barchant*, *barkān*, etc. 'auf einer Seite aufgerauhter Baumwollflanell' > Slovenian *barchan(t)*; Slovak *barchan*, Czech *barchan*, *parchan*, Polish *parchan*; French or Italian > Polish, Ukrainian, Russian *barakan*; Dutch *barkaan* > Russian *barkán* (M 47; V I, 123; KS 80; ML 941).

burnus 'a kind of hood' (< Greek *bírros* 'Art Überwurf') > Turkish *burnus* > Russian *burnus* > Slovak, Czech, Polish dialectal *burnus*, or through French *bournous* > German *burnus* (M 77; V I, 247; KS 146; ML 1223).

jubba(t), pl. *jibāb* ~ *jabab* 'vest or jacket with wide sleeves' > French *jupe* or Italian *giuppa* 'Jacke, Wams' > Middle High German *jop(p)e* ~ *juppe* > Slovenian *jopa*; Slovak, Czech *jupka*, dialectal (South) *jupa*, Upper Sorbian *jupa*, Lower Sorbian *jopa*, Polish *jupa*, Ukrainian *jupka*; the same origin is proposed for Czech dialectal (Morava) *župa* 'fur coat', and through Middle High German *schûbe, schoube* the word *šuba* 'fur coat' is attested in Bulgarian, Serbo-Croatian, Slovenian, Slovak, Czech, Upper and Lower Sorbian, Polish (*szuba*), Russian, Belorussian, Ukrainian (M 232, 629, 731; V IV, 482, 525; KS 412; ML 3951).

lubān jāwī 'frankincense of Java' > Catalan *lo benjuí* 'Weihrauch dienende Harz aus der Levante' (*lu-* was identified with the definite article) > French *benzoin* > German *Benzin* > Czech *benzin* 'benzoin'; Russian *benzin* etc. (R 76; V I, 151; KS 97; H 39).

muxayyar 'Stoff aus Ziegenhaar' > French *mohair* > Dutch or Early Modern High German *Machaier* > Old Czech *mochejr* ~ *muchejr*, Polish *mochajer* ~ *muchaj(e)r*; Russian *muxojar*; French > English *mohair* > German *Mohair* > Czech, Slovak *mohér* (M 372; V III, 19; KS 565).

qazz 'raw silk, floss silk; silk' (< Persian *käz*) > Spanish *gase* > French *gaze* > Dutch *gaas* > German *Gaze* > Russian *gaz* 'a kind of cloth' (V I, 382; KS 302; ML 3710).

quṭ(u)n 'cotton' > Italian *cotone* , French *coton* > Dutch *katoen* > German *Kattun*, dialectal (Silesia) *Kartun* > Slovak *kartún*, Czech *kartoun*, Upper Sorbian *kortun*, Lower Sorbian, Polish *kartun* (M 243; KS 243); Arabic > Turkish *kutnu*, *kutni* > Russian *kutnjá* 'Asiatic semisilk cloth' (V II, 435).

sammūr 'Siberian weasel; sable fur' > Spanish *zamarra* > Middle French *chamarre* > Dutch *samaar* 'long dress' > Russian *samára* 'id.'; Arabic > Italian *cimarra* 'long coat' > Polish *czamara* > Russian *čemára* (V III, 552; IV 331; ML 7563a).

There is one loanword denoting a musical instrument:

al-ʿūd 'lute, cittern' (lit. 'wood') > Portuguese *alaúde*, Spanish *laúd* > Old French *leüt* > Late Middle High German *lūte* > Slovak *lutna*, Old Czech *lútna*, Czech *loutna*, Upper Sorbian *lutna*, Polish *lutnia* 'lute' (M 342; ML 388).

Other loanwords include the following:

ʿawār 'defect, fault' > Italian *avaria* > French *avarie* > Dutch *averij*, Low German *Haverye* > Czech *havárie*; Italian or French > Russian *avárija* (V I, 58; KS 361).

masraḥ 'mattress' > Italian *materasso* > Old French *materas* > Early New High German *mat(e)raz* 'Matratze' > Bulgarian, Serbo-Croatian, Slovenian *matrac*, Serbo-Croatian also *matarac*; Slovak *matrac*, Czech *matrace*, Old Polish *matrac*, Polish *materac*; Russian *matrac*; Italian or Dutch *matras*, Middle Low German *matrasse* > Old Czech *matras* ‘mattress’; Russian *matras* (BER III, 688; M 355; V II, 582–583; KS 545).

masxara(t) ‘jest, joke, mockery; mummery, masquerade; mask, masked person’ > Turkish *masχara* > Italian *maschera* > German dialectal (Bavaria) *maškərə* > Slovenian *maškara*; Czech *maškara*, dialectal (Domažlice) *maškera*, Polish *maszkara* > Ukrainian *maškará*, Russian *maškara*; further Italian > French *masque* > German *Maske* > Czech *maska*; Serbo-Croatian *maskara* was borrowed through Turkish (M 353; V II, 586; KS 543; Räs 329).

mūmiyā ‘bitumen; mummy’ (< Persian *mūm* ‘wax’) > Italian *mummia* > German *Mumie* > Old Czech *mumie* ‘pitch mixed with glue; embalmed body’, Czech *mumie*; Russian *mumija*; Bulgarian *mùmija* ‘mummy’ ‘embalmed body’ (M 383; BER IV, 321; KS 574).

qalfasa ‘abdichten’ > Turkish *kalafat* and Medieval Greek *kalafatî* > Spanish *calatafear*, Italian *calafatare* > French *calfater* > Dutch *(op)kalefateren* > German *kalfatern* > Russian *konopátit’* (V II, 311; ML 4663).

rāḥa(t) ‘palm of the hand’ > Italian *racchetta* ‘Ballnetz, Federball, Rakete’, French *raquette* ‘palm of the hand’ > English *racket* > German *Racket* > Slovak, Czech *raketa* ‘racket’ (M 507; KS 662–663; ML 7013).

raʾs ‘head; origin’ > Spanish *raza*, Italian *razza*, French *race* > German *Rasse* > Czech *rasa* ‘race’ (M 509; KS 668).

ṭalismān, du. from *silasm* ‘talisman, amulet’ (< Greek *télesma* ‘bestätigendes Abbild’) > Spanish *talismán* > Italian *talismano*, French *talisman* > German *Talisman* > Czech *talisman* (R 650; KS 814).

ṭarḥa(t) ‘wrapper, overall’ (from *ṭaraḥa* ‘to remove’) > Italian *tara* > German *Tara* ‘(Gewicht der) Verpackung einer Ware’ > Russian *tára* (V IV, 20; KS 815).

3. The eastern route: Old Arabic loanwords in South and East Slavonic languages

A number of loanwords were borrowed at an early period from Arabic through Turkish or a Turkic language into Old Church Slavonic or Old Russian.

ʿanqāʾ ‘griffin, phoenix’ > Serbian Church Slavonic *inogъ* ~ *inegъ*, Old Czech *noh* ‘griffin’, Old Russian *nogъ* ‘hawk’ (M 401; V II, 134).

busra(t), pl. *busur* ‘glass bead’ (Ibn Faḍlān) > Old Bulgarian (cf. Chuvash *pĕšĕr* ‘pearl’) or Turkish *büsre* > Old Church Slavonic *bisъrъ*, *bisъrъ*, *biserъ*, Bulgarian, Macedonian, Serbo-Croatian *biser*; Church Slavonic > Old Czech *bíser*; Old Russian, Russian *bíser*, Ukrainian *býser* ‘pearl’ (M 54; Sk I, 156–157; V I, 168; Peňáz, ESJS 1, 63).

naqd ‘payment in cash; readiness’ > Crimean-Tatar *naχd*, Turkish *nakt* > Old Russian *nogáta* ‘one-half of grivna’ (V III, 79).

Šām, Šaʾam ‘Syria; Damascus’ > Turkish *Šam* ‘Syria’ > Old Russian *šamskij* ‘eastern’ (V IV, 403).

raḥmān ‘merciful, compassionate’ > Turkish *rahman* ‘id.’ > Russian *raxmánnyj* ‘lazy; quiet, silent, peaceful; peculiar’ (V III, 449–450).

rāḥat ḥulqūm ‘rest, repose, ease’ + ‘throat, gullet’ > Turkish *rahat lokum* > Russian *raxat-lukúm* ‘Oriental sweet from rice flour, sugar, fruit juice, and milk’ (V III, 449).

ṣandūq ‘box, chest, trunk’ (< Greek *sundokheíon*) > Turkish, Crimean-Tatar *sandyk*, Kypchak *sunduq*, *synduq*, Chuvash *sundəχ* ‘box’ > Russian, Ukrainian *sundúk* (V III, 803).

fūṭa ‘apron; napkin; handkerchief; purse’ > Turkish *futa*, *fota* ‘apron, striped cloth of Indian origin’ > Old Russian *fota*, Russian *fatá*, Ukrainian *fóta* ‘woman’s belt’ (V IV, 187).

faras ‘horse’ > Medieval Greek *fárēs* ‘Arabic horse’ > Old Serbian *farisъ*, Serbian *fariz*; Old Russian *farisъ*, Russian archaic *farъ* ‘horse of a good race’ (Sk I, 507; V IV, 187).

ḥinnā ‘henna (Lawsonia inermis)’ > Turkish *kyna* > Russian *xna* ‘henna’ (V IV, 251).

kīs ‘bag, purse’ > Turkish, Crimean-Tatar *käsä*, Tatar *kisä* > Polish *kiesa*; Ukrainian *kysá*, Russian *kisá* ‘sack’ (V II, 239).

maxzūla(t) ‘litter, rubbish’ > Turkish? > Russian *mazút* > Czech *mazut* (V II, 558; R 369).

muslim ‘believer’ > Persian pl. *muslimån* > Turkish *müslümän*, Tatar, Kazakh *musulman*, Kirgiz, Balqar *busurman* > Old Russian *busurmán* ‘pagan, Muslim’ (V I, 133).

nāʾib ‘substitute’ > Turkish *naib* > Russian *naíb* (V III, 39).

naqqāra(t) ‘small kettledrum, cymbal’ > Turkish *nağara* > Russian *nagará* ‘tympanum’ (V III, 36).

nawba(t) ‘guard; music, orchestra’, pl. *nawbāt* ‘drums, which are beaten from time to time at a great man’s residence’ > Turkish *näübät*, *nevbet*, *nöbet* ‘Reihenfolge, Wache’ > Russian *nabát* (V III, 34).

qalʿa(t) ‘fortress, castle’ > Turkish *kala* ‘id.’, Crimean-Tatar *kalä* > Russian *kalančá* (V II, 165).

qaydāni, du. of *qayd* ‘fetter, chain, bond; strap’ > Turkish *kajd* > Polish pl. *kajdany*; Ukrainian *kajdány*, Russian *kajdály*, dialectal *kajdány* ‘id.’ (V II, 161, 178).

ribāṭāt ‘solidly built building, blockhouse’ > Medieval Greek *rapátion̦* > Old Russian *ropat*, Russian, Ukrainian *rópat’* ‘pagan temple’ (V III, 502).

sabaniyyat ‘cloth from Saban near Baghdad’ > Middle Greek *sábanon* > Church Slavonic, Old Russian *savanъ*, Russian *sávan* (V III, 542).

ṣabr ‘myrrh’ > Turkish, Kypchak *sabur*, *sabyr* > Russian *sabúr* (V III, 542).

ṣandal ‘sandal(-wood)’ (< Persian *čandal* < Sanskrit *candana-*) > Turkish *sandal* > Old Russian *sandalъ*, Russian *sandál*; while Czech *santál* is borrowed from Medieval Greek *sándalon* (V III, 556; KS 704).

šayṭān ‘Satan, devil’ > Tatar *šaitan*, Turkish *šäitan* ‘devil’ > Russian, Ukrainian *šajtán* ‘id.’ (M 538; V IV, 395).

sayyid ‘master, lord, prince’ > Turkish *säyid* > Russian archaic *sejit* ‘noble dignitary at the sultan’s court’ (V III, 591; Räs 408).

ṣūf ‘wool’ > Turkish > Russian *zuf’* ‘haircloth’ (V II, 109).

sulṭān ‘absolute power, dominion; violence; absolute ruler, emperor, sultan’ > Turkish *sultan* > Old Russian *sultanъ*, Russian *sultan*, besides Old Russian *soltanъ*, *saltanъ* (V III, 551, 801).

ṭāqiya ‘undercap, fillet’ > Turkish, Tatar *takja*, Chuvash *toxja* ‘cap’ > Old Russian *taf’já* (V IV, 29).

xardal ‘mustard’ > Turkish *hardal* > Russian *gardál* ‘id.’ (V I, 393).

xilʿa(t) ‘robe of honor’ > Turkish *χilat* > Russian, Ukrainian *xalát* ‘long coat’ (V IV, 217).

Persian-Arabic *zum(ur)rud* ‘emerald’ (< Greek *smáragdos* ‘id.’) > Turkish *zümrüt* > Old Russian *izumrutъ* > Russian *izumrúd* ‘id.’, also ‘green’ (V II, 123).

4. Arabic words in Serbo-Croatian and Bulgarian, through the Ottoman Empire

Arabic loanwords in Serbo-Croatian and Bulgarian often date from the period of the Ottoman Empire; they were borrowed through the mediation of Ottoman Turkish. Many of them are administrative terms.

ʾamān ‘security, protection; faith’ > Turkish *aman* > Serbo-Croatian *àmān* ‘mercy’, Bulgarian *amàn* ‘id.’ (St 14; BER 10).

ʾamīn ‘faithful, trustworthy; minister’ > Turkish *emin* ‘district official; chief’ > Serbo-Croatian *èmīn* ‘id.’ (St 35).

ʾamīr ‘leader, prince’ > Turkish *ämir* > French *émir* > Czech *emír*, Russian *ėmír*; Arabic > Old Russian *amirъ*; Turkish > Serbo-Croatian *àmīr* ‘title of the Turkish sultans’, Bulgarian *emìr* (V IV, 518; St 15; BER 496).

ʾaʿyān pl. ‘great men, grandees’ > Turkish *âyan* ‘aristocrats’ > Bulgarian *ajànin*, Serbo-Croatian *àjān* ‘chief of an administrative unit in Turkey’ (St 12; BER 22).

ḥājj, pl. *ḥajīj* ‘pilgrim to Mecca’ > Turkish *hağy* > Serbo-Croatian *hàdžija* ‘pilgrim to Mecca or to Jerusalem’ (St 37).

ʾimām ‘leader, moderator; title of the first caliphs’ > Turkish *imam* > Serbo-Croatian *ìmām* ‘Muslim cleric, scholar’, Bulgarian *imàm* (St 42; BER II, 70).

jāmiʿ ‘great mosque’ > Turkish *ğami* > Bulgarian *džamìja*, Serbo-Croatian *džámija* ‘id.’ (St 32; BER I, 354).

jinn ‘demon, genie, fairy’ > Turkish *cin* > Czech *džin*; Serbo-Croatian *džìn*, Bulgarian *džin* ‘id.’ (Sk I, 474; BER I, 364).

kitāb ‘book’ > Turkish *kitab* ~ *kitap* ‘book; script’ > Serbo-Croatian *ćìtāp* ‘book; the *Qurʾān*’, Bulgarian *kitàb* ~ *kitàp* ‘book, document’ (St 29; BER II, 394).

mubāšir ‘inspector’ > Turkish *mübašir* > Serbo-Croatian *mubàšīr* ‘id.’ (St 51).

mudīr ‘prefect of a district, governor, director’ > Turkish *müdir* > Serbo-Croatian *mùdīr*, Bulgarian *mudìr* ‘Turkish governor, counselor’ (St 51; BER IV, 301).

musallim ‘counselor, deputy’ > Turkish *müsellim* > Serbo-Croatian *musèlim* ‘magistrate; deputy of a vizier’ (St 53).

nāḥiya(t) ‘district, canton’ > Turkish *nahije* > Serbo-Croatian (archaic) *náhija*, Bulgarian *naxìja* ‘id.’ (St 53; BER IV, 568).

wakīl ‘agent, deputy, governor’ > Turkish *vekil* ‘minister, deputy’ > Serbo-Croatian *vèćīl* ‘deputy of the sultan in Turkey’, Bulgarian *vekìl* (St 69; B 130).

wālī ‘regent, governor, prefect of a province’ > Turkish *vali* > Serbo-Croatian *válija* ‘governor in Turkey’, Bulgarian *valìja* ‘id.’ (St 69; BER I, 114).

wilāyat ‘empire, country, province, district’ > Turkish *vilâyet* > Serbian *vilájet* ‘district, province’, Bulgarian *vilaèt* ‘id.’ (St 70; BER I, 148).

wazīr ‘vizier, minister’ (from *wazara* ‘to carry weight’) > Turkish *vezir* > Czech *vezír*; Serbo-Croatian *vèzīr*, Bulgarian *vezìr* ‘id.’ (R 709; St 69; BER I, 129).

zaʿīm ‘master, lord, leader; owner of a large fief’ > Turkish *zaim* > Serbo-Croatian *záim*, Bulgarian *zaìm* ‘owner of a big fief’ (St 71; BER I, 589).

Other loanwords are connected with the army.

ʿalam ‘sign, mark; flag’ > Turkish *alem* > Serbo-Croatian *àlem* ‘id.’ (Sk I, 27–28).

ʿaskar ‘soldier, army, troop’ (from Latin *exercitus*) > Turkish *asker* > Bulgarian *askèr*, Serbo-Croatian *àsker* ‘Turkish mercenary; Turkish army’ (St 17; BER I, 18).

badan ‘body, trunk, member; rampart, wall’ > Turkish *beden* > Serbo-Croatian *bèden* ‘wall; trench’, Bulgarian *bedèn* ‘id.’ (St 21; BER I, 39).

ḍābiṭ ‘officer; policeman’ > Turkish *zabit* > Bulgarian *zabìt(in)*, Serbo-Croatian *zàbit* ‘Turkish officer’ (St 70; BER I, 569).

maḥalla(t) ‘quarter of a town; street; station’ > Turkish *mahalle* > Serbo-Croatian *màhala*, Bulgarian *maxalà* ‘id.’ (St 50; BER III, 690).

qaṣaba(t) ‘capital, town, citadel’ > Turkish *kasaba* > Serbo-Croatian *kasàba*, Bulgarian *kasabà* ‘village’ (St 48; BER 259–260).

nafar ‘common soldier’ > Turkish *nefer* > Serbo-Croatian *nèfer*, Bulgarian *nefèr* ‘id.’ (BER IV, 629–630).

niḏ̣ām ‘good order, arrangement, organization’ > Turkish *nizam* ‘id.; regular Turkish army’ > Serbo-Croatian *nìzam* ‘regular Turkish army’, Bulgarian *nizàm* ‘army’ (St 54; BER IV, 640).

silāḥ ‘arms’ + Persian suffix *-dār* ⇒ Persian *silāḥdār* ‘bearer of arms’ > Turkish *silihtar* > Serbo-Croatian *silìhtār* ‘id.’ (St 60).

taʿyīn ‘daily salary, ration’ > Turkish *tayın* > Serbo-Croatian *táin* ‘daily pay of a mercenary’ (St 63).

ṭināb ‘tent rope; rope ladder’ > Turkish *tenef* ‘thin rope’ > Serbo-Croatian *tènef* ‘rope’ (St 65).

Other loanwords stem from the judicial and legal system.

ʿādat ‘custom, habit, manner’ > Turkish > Bulgarian *adèt*, Serbo-Croatian *(h)ádet* ‘id.’ (St 11; BER 4).

daʿwā ‘process, lawsuit’ > Turkish *dava* > Bulgarian *davà*, *davìja*, Serbo-Croatian *dávija* ‘id.’ (St 30; Sk I, 384; BER I, 312).

ḍabṭ ‘keeping in order; government; regulation; control; confiscation’ > Middle Turkish *zapt* (modern Turkish *zabıt*) > Bulgarian *zapt*, Serbo-Croatian *zàpt* ‘discipline, subordination’ (St 72; BER I, 605).

ḏ̣ulm ‘injustice, oppression, tyranny’ > Turkish *zulüm* > Serbo-Croatian *zȕlum*, Bulgarian *zulùm* ‘violence, injustice, brutality’ (St 73; BER 657).

ḥabs ‘prison’ > Turkish *haps* > Serbo-Croatian *(h)àps* ‘id.’ (St 39).

ḥukm ‘judgment; jurisdiction’ > Turkish *hüküm* > Serbo-Croatian *hùćum* ‘id.’ (St 40).

ḥalāl ‘lawful, legitimate’ > Turkish *halal* > Serbo-Croatian *(h)àlāl* ‘blessing, forgiveness’ (St 38).

ʾiḏn ‘permission’ > Turkish *izin* > Serbo-Croatian *ìzun* ~ *ìzam*, Bulgarian *izìn* ‘id.’ (St 42; BER II, 31).

iftirāʾ ‘lie, insult, offense’ > Turkish *iftira* > Serbo-Croatian *iftìra*, Bulgarian *iftirà* ‘id.’ (St 41; BER II, 93).

ʾimdād ‘help, succor [in money, troops]; subsidies’ > Turkish *imdat* > Serbo-Croatian *ìndāt*, Bulgarian *imdàt* ‘id.’ (St 42; BER II, 70).

ʾiṭāʿat ‘obedience, submission’ > Turkish *itaat etmek* ‘to make obedient’ > Serbo-Croatian *itájet čìniti* ‘id.’ (St 42).

kafīl ‘who stands bail or security’ > Turkish *kefil* > Serbo-Croatian *ćèfīl*, Bulgarian *kefìl* ‘witness’ (St 28; BER II, 347).

maḥkama(t) ‘court of justice, tribune’ > Turkish *mahkeme* > Serbo-Croatian *mehćèma*, Bulgarian *mexkemè* ‘id.’ (St 51; BER III, 773).

maxārij pl. ‘expenses, costs’ > Turkish > Old Russian *mogoryč* ‘payment, wages’, Russian *mogoryč* ~ *magaryč* ‘reception; drinking after a bargain’, Belorussian *maharyč*, Ukranian *mohoryč* > Polish *mohorycz* ‘id.’; maybe also Old Czech *mochodrž* ~ *chomodrž* ‘a kind of a play’ lit. probably ‘payoff’ (M 202; V II, 635).

mufattiš ‘seeking; examining judge; censor, syndic’ > Turkish *müfettiš* > Serbo-Croatian *mufètiš* ‘inspector’ (St 52).

murāsala(t) ‘correspondence’ > Turkish *mürasele* > Serbo-Croatian *murásela* ‘law permission’ (St 52).

qāḍī ‘one who decides definitely, judge’ > Turkish *kady* > Serbo-Croatian *kàdija*, Bulgarian *kadìja* ‘clerical judge’ (St 45; BER II, 131).

sijillāt pl. ‘roll of parchment to write upon, volume; public document, diploma; edict; protocol’ (< Latin *sigilla* ‘seal’) > Turkish *siğilat* > Serbo-Croatian *sìdžilat* ‘court protocol’ (St 60).

šāhid ‘witness’ > Turkish *šahit* > Serbo-Croatian *šàit* ‘id.’ (St 62).

taḏkira(t) ‘remembrance; memorandum; passport’ > Turkish *tezkere* > Serbo-Croatian *teskera* ‘document; written permission’ (St 66).

taftīš ‘inspection, exam’ > Turkish *teftiš* > Serbo-Croatian *tèftiš* ‘Untersuchungskommission’ (St 64).

tamassuk ‘seizing, attachment’, pl. ‘written obligation’ > Turkish *temessük* > Serbo-Croatian *temesuć* ‘certificate’ (St 64).

ṯawāb ‘reward, recompense; requital’ > Turkish *sevap* > Serbo-Croatian *sèvāp/b* ‘charitable gift; good turn’, Bulgarian *sevàp* ‘id.’ (St 60; BER VI, 572).

ʿulūfa(t) ‘fed in the stable; pay, wages, salary’ > Turkish *ulûfa* > Serbo-Croatian *ulàpa* ‘pay for the Turkish mercenaries in Kraina’ (St 68).

ʿušr ‘tenth part, tithe; piece’ > Turkish *öşür* > Serbo-Croatian *ȕšur* ‘in-kind pay for a miller’ (St 69).

xanjar ‘large dagger’ > Turkish *χanžar* > Serbo-Croatian *hàndžār*, *hanžar*, *hančar* ‘id.’; Turkish > Kumyk *χynžal* > Russian *kindžal* ‘curved dagger’ (Sk I, 656; Räs 155).

zahr ‘die [play]’ > Turkish *zar* > Serbo-Croatian *zare* ‘id.’ (St 72).

Loanwords connected with the calendar include the following.

jumādā l-ʾāxir ‘sixth month of the lunar calendar’ > Turkish *ğemaziel-ahir* > Serbo-Croatian *džemáziful-áhir* ‘December’ (St 32–33).

šawwāl ‘the tenth month of the Arabic lunar calendar’ > Turkish *ševval* > Serbo-Croatian *šèvāl* ‘id.’ (St 62).

zamān ‘time, epoch, season, moment’ > Medieval Turkish *zaman*, Turkish *zeman* > Serbo-Croatian *zàmān*, *zèmān*, Bulgarian *zamàn* ‘time’ (St 70; BER I, 596).

A number of loanwords are connected with trades, professions, and crafts.

ʿanbar ‘ambergris’ > New Greek *ámbra* > Serbo-Croatian *àmber*; Bulgarian *ambér* (Sk I, 34).

ʾaṣnāf ‘forms, kinds’ > Turkish *esnaf* ‘craftsman’ > Serbo-Croatian *èsnāf* ‘craft’ (St 35).

ballūr ~ *billawr* ‘crystal, beryl’ (from Medieval Greek *bérulos*) > Turkish *bill’ur* > Serbo-Croatian *bìljūr*, Bulgarian *biljùr* ‘id.’ (St 22; BER I, 48).

bāmiya(t) ‘Hibiscus esculentus’ > Turkish *bamye* > Bulgarian *bàmja*, Serbo-Croatian *bàmnja*; Russian *bámija*, Ukrainian *bamija* ‘id.’ (Sk I, 104; BER 30).

baqqāl ‘greengrocer’ > Turkish *bakkal* > Bulgarian *bakàlin* ‘id.’; Turkish > Russian *bakalejščik* : *bakaléja* ‘dried fruits’ (BER I, 27; V I, 109).

daxīra(t) ‘stores, provisions’ > Turkish *zahire* > Bulgarian *zajrè*, Serbo-Croatian *zahíra* ‘id.’ (St 70; BER I, 590).

ḥalwa ‘sweets’ > Turkish *halva* > Bulgarian *xalvà*, Serbo-Croatian *(h)àlva* ‘sweet food consisting of nuts and wheat flour’ (Sk I, 652).

ḥirfa(t) ‘trade, profession; guild, corporation; industry’ > Turkish *hirfet* > Serbo-Croatian *rúfet* ‘guild; craft’ (St 58).

muštari ‘buyer’ > Turkish *müšteri* > Serbo-Croatian *muštèrija*, Bulgarian *mùščerija*, *mjuščerìja* ‘client; clientele’ (St 53; BER IV, 442).

nušādir ‘sal ammoniac’ > Turkish *nyšadyr* > Bulgarian *nišadăr*, Serbo-Croatian *nišàdor*; Arabic > Tatar *nyšatyr* > Ukrainian *našatýr*, Russian *našatýr’* (V III, 51).

ṣanʿa(t) ‘work; craft; deed’ > Turkish *zanaat* (modern Turkish *sanat*) > Serbo-Croatian *zànāt* ‘trade, craft’, Bulgarian *zanajàt* ‘id.’ (St 71; BER I, 599).

tarzī ‘tailor’ > Turkish *terzi* > Serbo-Croatian *tèrzija* ‘id.’ (St 65–66).

Other loanwords are connected with material culture.

ḥammām ‘warm bath’ > Turkish *hamam* ‘Turkish bath’ > Serbo-Croatian *(h)àmām* ‘id.’ (St 38).

ʾibrīq ‘water jug’ > Turkish *ibrik* > Serbo-Croatian *ìbrik*, Bulgarian *ibrìk* ‘id.’ (St 41; BER II, 2).

mašʿala(t) ~ *mašʿal* ~ *mišʿal* ‘light, lamp, torch’ > Turkish *meşal* > Russian *mášal* ‘torch’; Serbo-Croatian *mašàla*, Bulgarian *mašàl(a)* ‘id.’ (V II, 586; St 50; BER III, 699).

ṣaḥn ‘plate; large cup’ > Turkish *sahan* > Serbo-Croatian *sàhan*, Bulgarian *saxàn* ‘copper plate or cup’ (St 58; BER VI, 516).

A number of loanwords denote cloth or items of clothing.

ʿabāʾ ‘coarse cloth; wrapper, cloak’ > Turkish *aba* > Serbo-Croatian *abà*; Slovak pl. *háby*, Czech *háb* ‘dress’, Polish *haba* ‘coarse cloth’; Ukrainian *haba*, Belorussian, Russian *aba* ‘id.’ (M 153; St 11).

bazz ‘fine linen, silk, clothing’ ⇒ Persian *bezistān* ‘place of business’ > Turkish *bezesten* > Serbo-Croatian *bezìsten*, Bulgarian *bezistèn* ‘business hall’ (St 22; BER I, 40).

farāja(t) or *farajiyya(t)* ‘upper garment, fur coat’ (< Medieval Greek *phoresiá* ‘Kleid, Strassenmantel’) > Medieval Turkish *fereǧe* (modern Turkish *ferace*) > Serbo-Croatian *fëredža* ‘long and wide coat for ladies or Muslim clerics’ (St 36).

jawhar ‘jewel, pearl’ > Turkish *ǧevher* ‘jewel’, hence the verb *ǧevherlemek* ‘to decorate with jewels’ > Serbo-Croatian unattested verb **ževrleisati*, hence adjective *ževrleisan* ‘decorative, ornamental’ (St 73).

qaṭīfa(t) ‘velvet’ > Turkish *kadife* > Serbo-Croatian *kadífa*, Bulgarian *kadifè* ‘id.’ (St 45; BER II, 131).

qumāš ‘material for clothes, stuff; linen’ > Turkish *kumaš* > Serbo-Croatian *kùmaš* ‘atlas; silk band’, Bulgarian *kumàš* ‘a kind of cloth’; Russian *kumáč* > Slovak, Czech *kumač* ‘id.’ (St 49; BER III, 128; M 306–307).

Other cultural or abstract terms include the following.

ʿalāqa(t) ‘relations, interest’ and *ʿilāqa(t)* ‘straps, ropes’ > Turkish *alâka* > Serbo-Croatian *iláka* ‘bond; interest’ (St 41).

baraka(t) ‘blessing, happiness, abundance; fertility’ > Turkish *bereket* > Bulgarian *berekèt*, Serbo-Croatian *berìćet* ‘id.’ (St 22; BER I, 44).

ḍarar ‘damage, loss, injury; difficulty; distress, need’ > Turkish *zarar* > Bulgarian *zarar*, Serbo-Croatian *zàrār* ‘id.’ (St 72; BER I, 607).

ḥuḍūr ‘presence, appearance, attendance; quietude’ > Turkish *huzur* > Serbo-Croatian *(h)ùzūr* ‘id.’ (Sk I, 695–696).

ḥurma(t) ‘anything forbidden, sacred’ > Turkish *hürmet* > Serbo-Croatian *ùrmet* ‘honor’ (St 68).

ʿilāǧ ‘medical treatment, cure; medicine’ > Turkish *ilâç* > Serbo-Croatian *ílač*, Bulgarian *il(j)àč* ‘id.’ (BER II, 65).

ʿirḍ ‘good reputation, honor’ > Turkish *hırsiz* > Serbo-Croatian *(h)Ýz* ‘id.’ (Sk I, 655).

murakkab ‘compound, mixed; mixture’ > Turkish *mürekkep* > Serbo-Croatian *murèćep*, Bulgarian *murakèb* ‘id.’ (St 52; BER IV, 331).

naḏar ‘look, outlook, view; aspect; doubt; the evil eye’ > Turkish *nezer* > Serbo-Croatian *nèzer* ‘the evil eye; magic’ (St 54).

nūr ‘light’ > Turkish *nur* > Serbo-Croatian *nûr*, *nùhur* ‘ray of light, shining from the holy grave’, Bulgarian *nur* ‘fire, light’ (St 55; BER IV, 704).

qisma(t) ‘share, portion, lot; destiny, fate’ > Turkish *kismet* > Serbo-Croatian *kìsmet* ‘destiny, fate’ (St 48).

tafarruj ‘viewing, observation’ > Turkish *teferrüč* > Serbo-Croatian *tefèrič* ‘picnic, place of trip’ (St 63).

tartīb ‘method, plan, system’ > Turkish *tertip* > Serbo-Croatian *tèrtip* ‘id.’ (St 65).

xāṭir ‘mind, soul, consciousness’ > Turkish *hatyr* ‘memory’ > Serbo-Croatian *hátar* ‘mind, memory; love, attachment, affection; regard’ (St 39).

One animal’s name was borrowed through Turkish:

fīl ‘elephant’ > Turkish *fil* > Serbo-Croatian *fîl* ~ *fîlj*; Bulgarian *fil* ‘id.’ (Sk I, 510).

Bibliographical references

BER = *Bălgarski etimologičen rečnik*, ed. Vladimir Georgiev. 6 vols. Sofia.
ESJS = *Etymologický slovník jazyka staroslověnského*, vols. 1–13, ed. Eva Havlová and Adolf Erhart. Prague: Academia, 1989ff.
H = Hoad, Terry F. (ed.). 1986. *The concise Oxford dictionary of English etymology*. Oxford: Oxford University Press.
Jašar-Nasteva, Olivera. 2001. *Turskite leksički elementi vo makedonskiot jazik*. Skopje: Makedonska Akademija na Naukite i Umetnostite.
Korsch, Theodor 1885–1886. Review of Miklosich (1884). *Archiv für Slavische Philologie* 8.637–651; 9.487–520, 653–682.
KS = Kluge, Friedrich and Elmar Seebold. 1999. *Etymologisches Wörterbuch der deutschen Sprache*. 23rd ed. Berlin and New York: de Gruyter.
M = Machek, Václav. 1968. *Etymologický slovník jazyka českého*. Prague: Československá Akademie Věd.
Miklosich, Franz. 1884. *Die türkischen Elemente in den südost- und osteuropäischen Sprachen (griechisch, albanisch, rumunisch, bulgarisch, serbisch, kleinrussisch, grossrussisch, polnisch)*. Vienna: k.k. Akademie der Wissenschaften in Wien.
——. 1888. *Die türkischen Elemente in den südost- und osteuropäischen Sprachen (griechisch, albanisch, rumunisch, bulgarisch, serbisch, kleinrussisch, grossrussisch, polnisch): Nachtrag*. Vienna: k.k. Akademie der Wissenschaften in Wien.
ML = Meyer-Lübke, Wilhelm. 1935. *Romanisches etymologisches Wörterbuch*. Heidelberg: C. Winter.
R = Rejzek, Jiří. 2001. *Český etymologický slovník*. Prague: Leda.
Räs = Räsänen, Martti. 1969. *Versuch eines etymologischen Wörterbuchs der Türksprachen*. Helsinki: Suomalais-ugrilainen Seura.
Sk = Skok, Petar. 1971–1974. *Etimologijski rječnik hrvatskoga ili srpskoga jezika*. Zagreb: Jugoslavenska Akademija Znanosti i Umjetnosti.
St = Stachowski, Stanisław. 1967. *Studia nad chronologiją turcyzmów w języku serbsko-chorwackim*. Cracow.
T = Turek, Wacław P. 2001. *Słownik zapożyczeń pochodzenia arabskiego w polszczyźnie*. Cracow: Universitas Kraków.
V = Vasmer, Max. 1950–1958. *Russisches etymologisches Wörterbuch*. Heidelberg: C. Winter. [Quoted after the Russian translation, published in Moscow, 1986–1988.]

Václav Blažek (Masaryk University of Brno)

Slips of the Tongue

Slips of the tongue (also known as → speech errors) are unintentional deviations from the speaker’s intended production of a string of linguistic units. Slips have been of interest to Arabic grammarians as far back as the 8th century (al-Kisāʾī’s [d. 189/805] *Mā talḥanu fīhi l-ʿawāmm* ‘Errors of the populace’), insofar as they believed that slips provide clues as to how language changes. But it was not until the publication of Meringer and Mayer’s corpus (1895) that slips of the tongue began to receive increased attention. At the end of the 19th century, the term ‘spoonerism’ was coined after the Reverend William Spooner, warden of New College, Oxford, who was notorious for speech errors such as *the queer old dean*, when referring to the *dear old queen*. Sigmund Freud (1904) suggested that slips of the tongue resulted from repressed emotions and desires. More recently, however, it has been shown that slips of tongue reveal much more about our knowledge of the structure of language, as well as how we use this knowledge in our linguistic interactions, than about language change or repressed emotions or desires.

Slips of the tongue are of particular interest to linguists because of the widely accepted assumption that the rules of grammar guide the language-processing mechanisms and that evidence from language acquisition and ‘use’ (comprehension and production) have a bearing on the theory of grammar. Errors potentially shed light on the nature of the language system, since they mark points at which the system breaks down. In principle, errors carry information on where in the system the breakdown has occurred and how (i.e. the possible mechanisms involved). There is little evidence from naturally occurring slips, however, that suggests that their roots lie in repressed thoughts or fears. Language use is a cognitively complex process that requires many processes which can break down at any point.

1. Distribution

Slips of the tongue in Arabic exhibit the same distributional characteristics as slips in other languages (for Arabic, see Abd El-Jawad and Abu-Salim 1987; Safi-Stagni 1990, 1994; Berg and Abd El-Jawad 1996; for English, Fromkin 1988; Garrett 1988; Shattuck-Hufnagel 1991; for German, Meringer and Mayer 1895; for Mandarin Chinese, Jin 1995; for Thai, Gandour 1977). The highest percentage of errors is reported to occur between sounds (phonemes), or what are known as segmental errors, followed by errors between words, also known as lexical errors. Speech errors involving phonetic segments or words can either be anticipated, where a later segment replaces an earlier target segment (*ismu hayṯam* > *iṯmu hayṯam* 'his name is Haytham'), or perseverated, where an earlier segment replaces a later one (*mukāfaḥat al-muxaddarāt* > *mukāfaḥat al-muḥaddarāt* 'war on drugs'). Errors can result from exchanges or substitutions; in the former, two sounds or two words exchange places (*fāks* ***maysa*** > ***māks*** *faysa* 'Maisa's fax'; *kul-i l-ʾ**akil** u ʾašrab-i l-ʿ**aṣīr*** > *kul-i l-ʿ**aṣīr** u ʾašrab-i l-ʾ**akil*** 'eat the *food* and drink the *juice*'), while in the latter a segment or a word from an unknown source replaces the target segment or word (*sūrat tabāra**k*** > *sūrat tabāra**q*** 'sura *Tabāraka*' and *al-mōya* ***ḍaʿīfa*** > *al-mōya* ***naḥīfa*** 'the water is *weak* > *thin*'). Other errors are the result of either blending, where two competing words blend into one (***lamm**ēt* 'gathered' + *jamma**ʿat*** 'collected' > *lammaʿat*), or addition or deletion (*ʿulba* > *ʿuba* 'box'). The examples cited show the intended utterance on the left and the actual utterance on the right, with the units involved in the error in bold.

Speech errors affect not only phonetic segments or words but also involve phonetic or distinctive features, morphemes, phrases, and clauses. Phonetic features can be anticipated/perseverated (*bixēr u ʿāfya* > *bixēr u ḥāfya*). Morphemes can be shifted or stranded (*bārik ʿalā rasūli-**nā*** > *bārik-**nā** ʿalā rasūl*... 'bless our Prophet...'). Two intended phrases can blend into one (*al-ʿuṣūr al-wusṭā* + *al-qurūn al-wusṭā* > *al-qurūn al-ʿuṣrā* 'the Middle Ages'). Similarly, two clauses or sentences can fall victim to blending, with the resultant sentence showing units from both targets (*ʾaṭlaʿ-i ṣ-ṣuṭūḥ u jīb-i l-ġasīl* 'go up to the roof and get the laundry' + *jīb-i l-ġasīl min aṣ-ṣuṭūḥ u ʾanzul-i* 'bring the laundry from the roof and come down' > *ʾaṭlaʿ-i jīb-i ṣ-ṣuṭūḥ u ʾanzul-i* 'go up, get the roof, and come down').

2. Constraints

In natural speech, slips of the tongue are considered rare (one in every 1,000 words). Their occurrence, however, seems to be governed by universal as well as language-specific constraints. Lexical errors, for example, respect syntactic category constraints whereby nouns exchange with nouns (*ʾawarrīk an-**nujūm** fi ʿizz aḍ-**ḍuhur*** > *ʾawarrīk aḍ-**ḍuhur** fi ʿizz an-**nujūm*** 'I'll show you the *stars* in the height of *noon*') and verbs replace verbs (*ʾarūḥ* ***ʾatwaḍḍa*** > *ʾarūḥ* ***ʾastaḥamma*** 'I'm going to make an *ablution* > take a *shower*'), prepositions for prepositions (***gabil*** *ʿāmēn* > ***baʿad*** *ʿāmēn* '*before* > *after* two years'), and so on. Although segmental errors can occur between open and closed class items (*rāḍya ʿan šuġl-i* > *rāḍʿa ʿan šuġl-i ya daktōra* 'are you pleased with my work, doctor?'), lexical errors respect word class membership where errors do not occur between open and closed class words. Both lexical and segmental errors are subject to a position similarity constraint where an error and its source occupy an identical word or syllable position. For example, clitic pronoun exchanges respect word positions (*tilifōn-**ak** ʿind-**aha*** > *tilifōn-**aha** ʿind-**ak*** '*your* [masc. sg.] telephone number is with *her*') and segments occupying syllable onset/coda position interact with similar segments (***b**ūsi u **n**ūra* > ***n**ūsi u **b**ūra* 'Busi and Nura' and *muš bas mawḍūʿ-ik* > *mus baš mawḍūʿ-ik* 'not just your topic'). This seems to hold even if the segments are within one word (*samīḥa* > *saḥīma* 'Samiha'). Arabic segmental slips, however, have been shown to be insensitive to this positional similarity constraint (Safi-Stagni 1994; Berg and Abd El-Jawad 1996). Examples such as *ʿaṭšāna* > *ʿašṭāna* ‘thirsty’ and *ad-dā**nūb*** > *ad-dā**būn*** ‘the Danube’ show an interaction between onset and coda across syllable boundary as well as within the same syllable, respectively. Because of the nonconcatenative nature of Arabic morphology (McCarthy 1981; McCarthy and Prince 1990), consonantal segments are represented on separate tiers from their vocalic templates and are unanchored in the underlying structure of the word. The consonantal segments can therefore move independently of the vocalic templates and any affixes, or can be freely attached to onsets or codas depending on the template (e.g. *daraj* ‘stairs’ + C*a*C*ā*C*i*C from *salālim* ‘stairs’ > *darārij; yasīr-u bi-**x**uṭ-an **ḥ**aṯīṯa* > *yasīr-u bi-**ḥ**uṯ-an **x**aṭīṭa* ‘walks with rapid strides’).

Although initial segments are more error-prone than others, errors are also more common when the rest of the words are phonologically similar, e.g. *sitti rigayya **garība*** > *sitti rigayya **ragība*** ‘Grandma Rugayya is nearby’. Errors tend to result in existing words or stems, what has become known as the lexical bias, e.g. *ʾaji **maʿā-kum*** > *ʾaji **ʿamā-kum*** ‘I come *with you* > *your blindness*’, ***šāʿir** az-zajal* > ***šāriʿ** az-zajal* ‘zajal *poet* > *street*’, *kul-u ya **ḥarīm*** > *kul-u ya **ḥamīr*** ‘eat, oh *ladies* > *donkeys*’. When segmental slips, however, result in novel forms, these forms do not violate the phonotactics of the language. Mislocated segments accommodate to their new environment. For example, in *kīlu xamsa* > *kīlu xaṣma* ‘five kilo’, the /s/ of *xamsa* picks up [+back] due to its proximity to /x/, while in *ʿašān as-**sana** aj-jadīda* > *ʿašān al-**ʿīd** aj-jadīd* ‘because of the new *year* > *Eid*’ both the definite marker *al-* and the adjective *jadīd* ‘new’ show accommodation to the new error rather than to the target.

3. Slips and the mental lexicon

Speech errors not only provide evidence for the units involved in speech production (phonetic segments, morphemes, words, phrases, etc.), they also shed light on the nature of the representation and organization of the mental lexicon. Consonantal roots and word pattern movements and substitutions, such as ***šayyad*** *tāj maḥal* > ***tayyaj*** *tāj maḥal* ‘he erected the Taj Mahal’ and *ḥugūga-hum mahḍūma* > ***huḍūma**-hum **maḥgūga*** ‘their rights are infringed’, as well as *lā tjība-ha **mṭafaḥa*** > *la tjība-ha **maṭfūḥa*** ‘don’t bring it full!’, show that they are represented on separate tiers in the lexicon and are assembled during the course of production. Lexical entries include syntactic category information, since targets and errors in lexical exchanges (e.g. *šūf-i ʾ**axū**-ki **rās**-u ʾatxabaṭ* > *šūf-i **rās**-ik ʾ**ax**-u ʾatxabaṭ* ‘see your *brother*, his *head* is hit’) and lexical substitutions (e.g. *ma-hum **katīr*** > *ma-hum **galīl*** ‘they are not *much* > *little*’) share the same grammatical category and respect word class membership.

Additionally, lexical entries seem to be organized into neighborhoods of semantically and phonologically related items, since substitutions are either semantically triggered (e.g. ***ṣūm**-i talāta ʾayyām* > *ʾ**atṣaḍag**-i talāta ʾayyām* ‘*fast* > *donate* three days’), or phonologically triggered (e.g. *yōm al-**giyāma*** > *yōm al-**gumāma*** ‘*Judgment* > *garbage* Day’), or both (e.g. *ʾ**idn**-i marra taglāna* > ***dign**-i marra taglāna* ‘my *ear* > my *chin* is very heavy’). Shifts and stranding errors, such as *nafs **ar**-ragum* > ***an**-nafs ragum* ‘the same number’, and *šīl-i **jazmat**-ik u **bōt** ʾaxū-ki* > *šīl-i **bōt**-ik u **jazmat** ʾaxū-ki* ‘pick up your *shoes* and your brother’s *boots*’, respectively, show that complex words are stored decomposed where inflectional and grammatical markers (possibly including tense, e.g. *ʾa**dug** al-bāb ma-ʾaḥad yiʾa**fuk**-al-i* > *ʾa**fuk** al-bāb ma-ʾaḥad yi**dug**-al-i* ‘I *knock* on the door and nobody *opens*’), as well as clitic pronouns, are stored separately from stems in the lexicon.

4. Slips and models of language production

Although physical evidence in the form of spectrograms shows that natural speech is continuous, slips of the tongue provide evidence that phonetic segments, morphemes, words, phrases, and clauses are psychologically discrete units that form the building blocks of the computations involved in language production. Several models have been proposed to account for speech production using evidence from speech errors (Bierwisch 1981; Butterworth 1981; Dell 1986; Fromkin 1971, 1973; Garrett 1980, 1988; Levelt and Cutler 1983; Levelt 1989).

The two most prominent models in the 1970s and 1980s were those of Fromkin (1973) and Garrett (1988). Both models assume several levels of autonomous representations and sets of serial ordered processes that apply at each level. Lexical substitution errors, such as *taʿāl-i **kul**-i hina* > *taʿāl-i **nām**-i hina* ‘come *eat* > *sleep* here’ suggest a procedure of retrieval of stored linguistic information. On the other hand, exchange errors such as *ṣuk **xušm**-ak u ʾaftaḥ **fam**-ak* > *ṣuk **fam**-ak u ʾaftaḥ **xušm**-ak* ‘shut your *nose* and open your *mouth*’ suggest a phrasal ‘integration’ procedure. Semantically driven retrieval errors (e.g. ***šarrib**-ih mōya* > ***ʾakkil**-ih mōya* ‘give him water to *drink* > *eat*’), and the fact that exchanges occur between words of the same grammatical class with no phonological similarity between the target or source (e.g. *ḥuṭṭ-i l-**kafatīra** fōg al-**butagaz*** > *ḥuṭṭ-i l-**butagaz** fōg al-**kafatīra*** ‘put the *kettle* on the *stove*’) both argue for a level of representation where a syntactic frame is constructed and a semantically driven lexical insertion process retrieves lexical items (content words) from the lexicon and assigns them to their structural slots. Inflectional and derivational morphemes (and, in Arabic, clitic pronouns, as well as the definite marker *al-*) are rarely involved in lexical exchanges and are usually stranded (e.g. ***mōya** fī l-**barmīl*** > ***barmīl** fī l-**mōya*** ‘*water* in the *barrel*’ and *hāda **duxān sayyāra**-t-na* > *hāda **sayyār duxāna**-t-na* ‘that’s the *smoke* of our *car*’). They are, therefore, assigned to the syntactic frame independent of lexical items. On the other hand, phonologically driven retrieval errors (substitutions such as *fēn al-ʿ**alāga*** > *fēn al-**milʿaga*** ‘where is the *hanger* > *spoon*’) and phoneme and morpheme exchanges (such as ***k**uḥl-i **s**āyil* > ***s**uḥl-i **k**āyil* ‘my eyeliner is running’, and *murā**sil** min al-mar**kaz*** > *murā**kiz** min al-mar**sal*** ‘a courier from the center’), where there is often a strong phonological similarity among phonemes/words plus same syllable structure, both argue for a level where a phonologically driven lexical insertion process retrieves the phonological form and assigns it to the surface slots. At this level, phonological forms of function words (articles, prepositions, and affixes) and prosody (stress and intonation) are assigned to the syntactic frame. Lexical blends such as *fōg al-kursi* ‘chair’ + *al-kanaba* ‘sofa’ > *fōg al-karasa* occur at this level, since they are the result of multiple words being activated and are competing for the same phrasal slot at the same time. Finally, accommodation errors such as *as-sawwā**g** bi-y**ṭ**awwi**ṭ*** > *a**ṭ**-**ṭ**awwā**ṭ** bi-ysawwi**g*** ‘the driver is beeping his horn’ and *ʾasʾal **aṭ-ṭabbāx aṭ-ṭulyāni** tikaffi* > *ʾasʾal **aṭ-ṭulyāni aṭ-ṭabbāx** yikaffi* ‘ask the *chef* if the *sheep* are enough’ suggest the last stage of the production process, where the phonetic forms are specified and adjusted to the final phonetic rules of the language. Shattuck-Hufnagel and colleagues (1979, 1983, 1987, 1991) proposed a serial ordering mechanism to account for the behavior of segmental errors.

Fromkin’s and Garrett’s autonomous stage and serial approach to language production, however, did not account for some other properties of speech errors, such as mixed substitution errors, or for the lexical bias (i.e., errors more often than not result in real words), observed in naturally occurring as well as experimentally induced speech errors. In mixed substitution errors (e.g. *ʾaḥubb ʾ**atfašxar** guddām an-nās* > *ʾaḥubb ʾ**atfaršax** guddām an-nās*), it is not clear whether the error is the result of a lexical substitution or a segmental exchange. Motley and Baars (1976) found that in laboratory-induced phonological errors there was a higher likelihood of the errors being real words than neologisms (e.g. *darn bore* > *barn door* rather than *bart doard*). These observations have been used by both Dell (1986) and Levelt (1989) to formulate a bidirectional interactive (connectionist) model of representation involving top-down and bottom-up parallel distributed processing with spreading activation between nodes organized into networks. Connections between the nodes are based on semantic and phonologic similarities. Frequency plays a role in determining the strength of connections between different nodes. Contamination slips (e.g. *ḥuṭṭ-i r-ruzz **fi t-tallāja*** > *ḥuṭṭ-i r-ruzz **taḥt al-maxada*** ‘put the rice *in the fridge* > *under the pillow*’) where the substitution of the lexical item is neither semantically nor phonologically triggered (the subject was tucking the bedsheets under the pillow and giving instructions regarding the rice) have been used to argue that the language production mechanism is not informationally encapsulated, as has been suggested by Garrett (Safi-Stagni 1990).

5. SLIPS AND OTHER PERFORMANCE BREAKDOWNS

Apparently, the same processes that govern slips of the tongue are at work in other linguistic and nonlinguistic phenomena (see Garnes and Bond

1980 on slips of the ear; Ellis 1979 and Hotopf 1983 on slips of the hand; Ruch 1972 on slips of the eye; and Norman 1981 on action slips). Similar reading, writing, and action errors and slips of the ear have been attested in Arabic but have not been systematically collected or analyzed yet. Additionally, Buckingham (1979, 1980), Talo (1980), and Garrett (1982) have correlated slips of the tongue with similar errors in the performance of aphasic patients. Safi-Stagni (1991) reported similar processes governing aphasic errors with slips of the tongue in Arabic. More recently Prunet a.o. (2000) correlated aphasic errors with speech errors in Semitic languages (both Hebrew and Arabic). Poulisse (1999) looked at speech errors in first and second language production and found that both lemmas (a representation proposed by Levelt that contains the semantic and syntactic properties of the lexical item) and lexemes (the representation that holds the phonological information of the lexical item, also proposed by Levelt) are activated simultaneously and that phonologically related word forms from different languages may spread activation to each other.

Bibliographical references

Primary sources

Kisāʾī, *Mā talḥanu* = ʾAbū l-Ḥasan ʿAlī ibn Hamza al-Kisāʾī, *Kitāb mā talḥanu fīhi l-ʿawāmm*. Ed. Carl Brockelmann, *Zeitschrift für Assyriologie* 13 (1898) 29–46. [English trans. Mohammad S. Anwar, "Remarks on a collection of speech errors", *International Journal of Psycholinguistics* 6 (1979) 59–72.]

Secondary sources

Abd El-Jawad, Hassan and Issam Abu-Salim. 1987. "Slips of the tongue in Arabic and their theoretical implications". *Language Sciences* 9.145–171.

Baars, Bernard. 1980. "On eliciting predictable speech errors in the laboratory". Fromkin (1980:47–65).

Berg, Thomas and Hassan Abd El-Jawad. 1996. "The unfolding of suprasegmental representations: A cross-linguistic perspective". *Journal of Linguistics* 32.291–324.

Bierwisch, Manfred. 1981. "Linguistics and language errors". *Linguistics* 19.583–626.

Buckingham, Hugh. 1979. "Linguistic aspects of lexical retrieval disturbances in posterior fluent aphasias". *Studies in Neurolinguistics*, IV, ed. Harry Whitaker and H.A. Whitaker, 269–291. New York: Academic Press.

——. 1980. "On correlating aphasic errors with slips-of-the-tongue". *Applied Psycholinguistics* 1.199–220.

Butterworth, Brian. 1981. "Speech errors: Old data in search of new theories". *Linguistics* 19.625–662.

Dell, Gary. 1986. "A spreading-activation theory of retrieval in sentence production". *Psychological Review* 93.283–321.

Ellis, Andrew. 1979. "Slips of the pen". *Visible Language* 13.265–282.

Freud, Sigmund. 1904. *Zur Psychopathologie des Alltagslebens: Über Vergessen, Versprechen, Vergreifen, Aberglaube und Irrtum*. Berlin: S. Karger. (1st ed., *Monatsschrift für Psychiatrie und Neurologie* 10.1–32.) (Repr., Frankfurt/Main: Fischer, 1954.) [English trans., *The psychopathology of everyday life*. Hammondsworth: Penguin, 1975.]

Fromkin, Victoria. 1971. "The non-anomalous nature of anomalous utterances". *Language* 47.27–52.

——. 1973. *Speech errors as linguistic evidence*. The Hague: Mouton.

——. (ed.). 1980. *Errors in linguistic performance: Slips of the tongue, ear, and hand*. New York: Academic Press.

——. 1988. "The grammatical aspects of speech errors". *Linguistics: The Cambridge survey*, II, ed. Frederick Newmeyer, 117–138. Cambridge: Cambridge University Press.

Gandour, Jack. 1977. "Counterfeit tones in the speech of southern Thai bidialectals". *Lingua* 41.125–143.

Garnes, Sara and Zinny Bond. 1980. "A slip of the ear: A snip of the ear? A slip of the year?". Fromkin (1980:231–239).

Garrett, Merrill. 1980. "Levels of processing in sentence production". *Language Production*, I, ed. Brian Butterworth, 177–220. London: Academic Press.

——. 1982. "Production of speech: Observations from normal and pathological use". *Normality and pathology in cognitive functions*, ed. Andrew Ellis, 19–76. New York: Academic Press.

——. 1988. "Processes in language production". *Linguistics: The Cambridge survey*, III, ed. Frederick Newmeyer, 69–96. Cambridge: Cambridge University Press.

Hotopf, W.H.N. 1983. "Lexical slips of the pen and tongue: What do they tell us about language production". *Language production*, II, ed. Brian Butterworth, 147–199. London: Academic Press.

Jin, Zhou. 1995. *A study of speech errors in Mandarin Chinese*. Ph.D. diss., University of Oxford.

Levelt, Willem. 1989. *Speaking: From intention to articulation*. Cambridge, Mass.: MIT Press.

——. and Ann Cutler. 1983. "Prosodic marking in speech repair". *Journal of Semantics* 2.205–217.

McCarthy, John. 1981. "A prosodic theory of non-concatenative morphology". *Linguistic Inquiry* 12:3.373–416.

——. and Alan Prince. 1990. "Prosodic morphology and templatic morphology". *Perspectives on Arabic linguistics*, II, ed. Mushira Eid and John McCarthy, 1–54. Amsterdam and Philadelphia: J. Benjamins.

Meringer, Rudolf and Karl Mayer. 1895. *Versprechen und Verlesen: Eine psychologisch-linguistische Studie*. Stuttgart: Goschen'sche Verlagsbuchhandlung. [Reissued, with introductory essay by Ann Cutler and David Fay. Amsterdam: J. Benjamins, 1978.]

Motley, Michael and Bernard Baars. 1976. "Semantic bias effects on the outcomes of verbal slips". *Cognition* 4.177–187.

Norman, Donald. 1981. “Categorization of action slips”. *Psychological Review* 88.1–15.

Poulisse, Nanda. 1999. *Slips of the tongue: Speech errors in first and second language production*. Amsterdam and Philadelphia: J. Benjamins.

Prunet, Jean-Francois, Renee Beland, and Ali Idrissi. 2000. “The mental representation of Semitic words”. *Linguistic Inquiry* 31:4.609–648.

Ruch, R. 1972. “An account of the Freudian slip in reading and writing”. *Journal of Technical Writing and Communication* 2.343–352.

Safi-Stagni, Sabah. 1990. “Slips of the tongue in Arabic”. *Perspectives on Arabic linguistics* I, ed. Mushira Eid, 271–290. Amsterdam and Philadelphia: J. Benjamins.

——. 1991. “Agrammatism in Arabic”. *Perspectives on Arabic linguistics*, III, ed. Bernard Comrie and Mushira Eid, 251–270. Amsterdam and Philadelphia: J. Benjamins.

——. 1994. “Arabic segmental errors and segmental phonology”. *Perspectives on Arabic linguistics*, VI, ed. Mushira Eid, Vicente Cantarino, and Keith Walters, 169–184. Amsterdam and Philadelphia: J. Benjamins.

Shattuck-Hufnagel, Stefanie. 1979. “Speech errors as evidence for a serial-ordering mechanism in sentence production”. *Sentence processing: Psycholinguistic studies presented to Merrill Garrett*, ed. William E. Cooper and Edward C. Walker, 295–342. Hillsdale, N.J.: Erlbaum.

——. 1983. “Sublexical units and suprasegmental structure in speech production planning”. *The production of speech*, ed. Peter F. MacNeilage, 109–136. New York: Springer Verlag.

——. 1987. “The role of word-onset consonants in speech production planning: New evidence from speech error patterns”. *Motor sensory processes*, ed. Eric Keller and Myrna Gopnik, 17–51. Hillsdale, N.J.: Erlbaum.

——. 1991. “The role of word structure in segmental serial ordering”. *Lexical access in speech production*, ed. Willem Levelt, 213–260. Cambridge, Mass.: Blackwell.

——, and Dennis Klatt. 1980. “How single phoneme error data rule out two models of error generation”. Fromkin (1980:35–46).

Talo, Ewa S. 1980. “Slips of the tongue in normal and pathological speech”. Fromkin (1980:81–86).

Sabah M.Z. Safi (King Abdulaziz University)

Somali

Arabic and the Somali language are related, though distantly, because Arabic belongs to the Semitic branch and Somali to the Cushitic branch of the → Afro-Asiatic/Hamito-Semitic family. This does not mean that the Somalis are Arabs, and as a matter of fact the membership of → Somalia in the League of Arab States is based on a political decision rather than on linguistic relationship. The contacts between Arabic and Somali are as old as the cultural contacts that culminated in the Islamization of the Somali people centuries ago. There is no precise dating, but more than one thousand years of contact is a safe guess. It is well known that Arabic was not only the language of religion but also the second language of culture in the 19th and 20th centuries. When Somalia regained independence in 1960, Arabic was recognized as one of the official languages, together with Italian and English. Some people advocated the use of Arabic script for Somali, but the Latin script was introduced officially on October 21, 1972. Note that in the current Somali writing system *ʿayn* is spelled as *c*, as in the Arabic name ʿAlī, which is spelled Cali; the retroflex /ɖ/ is spelled as *dh*; uvular /x/ is spelled *kh*; pharyngeal /ḥ/ is spelled as *x*; and the glottal stop is spelled as an apostrophe; long vowels are doubled in writing.

A large number of Somalis have only limited, sometimes very basic, knowledge of Arabic, while many others are quite fluent in Literary Arabic and/or a variety of spoken dialects. The influence of spoken varieties of Arabic, mainly the coastal dialects of Yemen and Oman, and of Literary Arabic (represented by the *Qurʾān* and some theological and legal writings known in Somalia) has continued over a long period of time in Somali. Due, however, to the considerable grammatical differences between the two languages (in spite of some evident inherited features they have in common, such as the few Somali prefix-conjugated verbs, which are inflected in much the same way as in Arabic, e.g. *ya-qaan* ‘he knows’, *ta-qaan* ‘you know’, *na-qaan* ‘we know’, etc.), this influence has been limited on the morphological, syntactic, and stylistic levels. On the other hand, the phonological systems of Somali and Arabic show several similarities, and because of the presence in Somali of such consonants as /ʿ/, /h/, /ḥ/, /x/, /q/, /ʾ/, as well as the glide /w/ and long vowels, lexical influence was facilitated phonetically.

Lexical borrowing has been quite considerable, and this has resulted in a large number of Arabic loanwords in Somali, which is one of the languages with a relatively high number of Arabic loanwords. Soravia (1994) found 1,436 Arabic loanwords in the best Somali dictionary available (Agostini a.o. 1985), which contains approximately thirty thousand lemmas.

Many of the Arabic loanwords, the majority of which are nouns, are frequently used. It must be emphasized that some Arabic loanwords may have been more or less ephemeral because some of them, found by Zaborski (1967:122) in older sources, are absent in Agostini a.o. (1985). The process of lexical borrowing from Arabic continues to be strong, but its extent has not yet been investigated.

It seems that the bulk of the loanwords were taken from Literary Arabic, although with some phonetic and phonological features of different Arabic dialects. The intermediaries were in most cases Somali men, learned in Islam, so that the majority of loanwords do not betray a pure dialect character. As far as the semantics of the loans is concerned, there are both words borrowed out of practical necessity and prestige loans; sometimes, the reason for borrowing is rather obscure.

Arabic emphatic /ḍ/ usually corresponds with Somali /d/, but sometimes with Somali /l/, e.g. *ḍaʿīf* 'weak, ill' > Somali *laciif/daciif; rāḍī* 'pleased' > Somali *raalli; ḍamān* 'warranty' > Somali *dammaan/lammaan; qāḍī* 'judge' > Somali *qaalli/qaaddi*. Very seldom, Arabic /ḍ/ corresponds with Somali retroflex /dh/, e.g. *waḍaf* 'sling, catapult' > *wadhaf* (Abraham 1962:246; *waraf* in Agostini a.o. 1985). These different renderings of Arabic /ḍ/ indicate that the words concerned derive from different Arabic dialects, with different pronunciation of this consonant (→ *ḍād*), some regional types of Arabic realizing this phoneme with a lateral feature. Idiosyncratic is *ḍarūra* 'necessity' > *naruuro* 'necessity', due to dissimilation. Arabic emphatic /ṭ/ usually corresponds with Somali /d/, e.g. *xuṭba* 'sermon' > *khudbad* 'speech, sermon', but there are also examples in which it is rendered by Somali /t/, e.g. *ʿaṭr* 'perfume' > *catar*. Arabic emphatic /ḏ̣/ is rendered as /d/, e.g. *ḏ̣ālim* 'unjust, evil' > *daalin* (with *-m* > *-n*) 'dishonest', and Arabic emphatic /ṣ/ is rendered as Somali /s/, e.g. *ṣarāḥa* 'sincerity' > *saraaxad*.

Original Arabic interdental /ṯ/ is rendered either by Somali /s/ or /t/; there are, for instance, variants *salaasa* and *talaado* 'Tuesday' (< Arabic *ṯalāṯa* 'three', with the second /ṯ/ rendered by Somali /d/!); the merging of original /ṯ/ with either /s/ or /t/ is already found in several Arabic dialects. Arabic interdental /ḏ/ is rendered in Somali by /d/, e.g. *ḏubāla* 'wick' > *dubaalad*.

Interestingly, although Arabic /s/ usually corresponds with Somali /s/, there are some cases in which it corresponds with Somali /sh/, and Arabic /š/ is realized usually as /sh/, but there are some of examples of a correspondence with Somali /s/, e.g. *kīs* 'bag, purse' > Somali *kiish/kiis*. Arabic /z/ is generally rendered by Somali /s/, e.g. *wazīr* 'minister' > *wasiir*.

Uvular /x/ occurs in Somali almost exclusively in Arabic loanwords (see Agostini a.o. 1985:363–367). This phoneme may have been borrowed from Arabic; it sometimes interchanges with Somali /q/ and /k/, e.g. *maxzin* 'store, magazine' > *maqsin; ʾaxḍar* 'green' > *akhdar/akhtar/aktar; maxlūq* 'creature' > *makluuq*. Somali /q/ also renders both Arabic /q/ and /ġ/, e.g. *ġaniyy* 'rich' > *qani*, but there are some cases where, following Bedouin dialects of Arabic, Arabic /q/ is rendered in Somali as /g/, e.g. *milʿaqa* 'spoon' > *maclagad/macalgad*; sometimes, it is rendered by /k/, e.g. *qism* 'subdivision' > *kasmo*, and /kh/, e.g. *ṣandūq* 'box, case' > *santuukh; baqqāl* 'greengrocer' > *bakhaar* 'shop', while /ġ/ is sometimes rendered by /kh/, e.g. *maġrib* 'sunset' > *makhrib; ṣiyāġa* 'goldsmithing' > *siyaakhad* 'a piece of jewelry'. Arabic /k/ is rendered as /k/ in initial and medial position but as /g/ in final position and sometimes medially, e.g. *kiḏb* 'lie' > *kidib; šukr* 'thanks' > *shugri; šarīk* 'companion' > *shariig*; there are also cases of Arabic /k/ > /q/ or even /kh/, e.g. *sakrān* 'drunk' > *saqraan/sakhraan/sarqaan*.

Arabic /j/ is usually rendered in Somali as /j/, but in a total of nine loanwords it corresponds with /g/ (Callegari 1987–1988:458–459), possibly indicating Cairene Arabic origin. This is, however, not quite certain since the loans in question are semantically marginal, and it is difficult to say why they would have been borrowed precisely from Cairene Arabic. Sometimes, doublets may indicate origin from different dialects and different chronology, e.g. *jayš* 'army' > *gaas* 'a division of soldiers' and *jeysh* 'army'. There are very few cases of Arabic /j/ corresponding with /y/, e.g. *jār* 'neighbor' > *yaar; dajāj* 'poultry' > *diyaaj/diyaad*, which shows its origin from some Yemeni, Omani, or Gulf Arabic dialect, while there is also *digaag*. Very rarely, Arabic /j/ is rendered as Somali /sh/, e.g. *mujarrab* 'tentative' > *sharrib; xarāj* 'land tax' > *kharash* 'expenditure', but this is rather an internal Somali change (viz. devoicing), conditioned by *-r(r)-*, which causes dissimilation. There is also *ḥajj* 'pilgrimage' > *xaj/xad* and *sirāj* 'lamp, lantern' > *siraaj/siraad*.

Sometimes Arabic /l/ > /r/, e.g. *wālid* 'father' > *waarid*. There are also some cases of /m/ > /b/, e.g. *zamān* 'time' > *saman/samaan/saben/sabaan*.

Consonant clusters are realized in Somali with an anaptyctic vowel, and in monosyllabic words final gemination is lost. Sometimes, there is metathesis, e.g. *ʾibra(t)* 'needle' > *irbad; qalʿa(t)* 'fortress' > *qalcad/qalco* or *calqad; jins* 'sort, species, genus' > *jinsi/sinji* 'gender, race'; *laʿna(t)* 'curse' > *lacanad/lacnad/nacallad/nacdal; bunduq(iyya)* 'shotgun, rifle' > *dumbuq* as well as *bunduq/buntuq; ʿafrīt/ʿifrīt* 'female evil spirit' > *cifriid/cirfiid* 'evil spirit'.

As far as morphological interference is concerned, Arabic feminine singular *-a(t)* is usually preserved in Somali, e.g. *muʿallim/muʿallima* 'male/female teacher' > *macallin/macallim-ad*, although in some cases it is only represented by the *-o* allomorph, e.g. *barak-o* 'blessing', but *baraka-ad-ii* 'the blessing'. Gender distinction in nominals is preserved in the singular, although there are some exceptions, e.g. *sana(t)* 'year' > *sanad/sannad*, which can be either feminine or masculine (Agostini a.o. 1985:535, *pace* Soravia 1994). According to the rules of Somali grammar, though, the gender changes in the plural, a process known as 'gender polarization', e.g. *dār* 'house' > *daar-ta* 'the house [made of stone; fem.]', *daar-o-ha* 'houses' [masc.]; *al-bāb* 'the door' > *albaab-ka* 'the door [masc.]', *albaabb-o-ta* 'the doors [fem.]'; *muʾallif* 'writer, author' > *allife* 'the writer, author [masc.]', *allifaad/alifid* 'the female writer, author' [fem. sg.], *allif-a-yaal* 'male authors [fem.]', but the plural *dukaan-la-yaal* (< Arabic *dukkān* 'shop') is either masculine or feminine. Some Arabic internal plurals have been preserved, sometimes with an additional Somali plural ending, e.g. *tājir/tujjār* 'merchant' > *taajir* 'rich man, merchant', plural *tujaariin* or *taajirro*. Sometimes, Arabic internal plurals have been reinterpreted as singular forms, e.g. *yawm/ʾayyām* 'day' > *ayaan* [sg.]; *masjid/masājid* 'mosque' > *masaajid* [sg.]. Monosyllabic nouns follow the rules of Somali grammar, i.e., they make their plural forms with partial reduplication, e.g. *ṣawt* 'voice' > *sawd*, pl. *sawdad; fāʾs* 'axe' > *faash*, pl. *faashash*. Sometimes, nouns have been borrowed together with the Arabic definite article, e.g. *albaab-jooge* 'doorman' (Arabic *al-bāb* 'the door'); *alleyl* 'night' (Arabic *al-layl* 'the night'), *adduun/addunyo/duunyo* 'world' (Arabic *ad-dunyā* 'the world'). There are also Somali nouns derived with Somali suffixes from Arabic nominals, e.g. *ʿāqil* 'intelligent' > *caaqil: caaqil-ni-mo* 'intelligence'; *ʾamīr* 'emir' > *ammir: ammir-ni-mo* 'emirate'; *dukkān* 'shop' > *dukaan: dukaan-le* 'shop owner'.

With verbs, Somali derivational suffixes can be used, e.g. *bāraka* 'to bless' > *barakee* 'to bless', *barak-so* (with a 'causative' suffix of Afro-Asiatic origin, cf. *-s-* in Arabic *i-s-tafʿala*) 'to give charity in the hope of gaining God's favor'. There are some examples of Arabic participles and verbal nouns (*maṣdar*s) used as verbs, e.g. *mamnūʿ* 'forbidden' > *mamnuuc* 'to forbid'; *muquur* 'to dive' (< Arabic *ġāra* 'to sink, go down, penetrate into'?); *musāfir* 'traveler' > *masaafiri* 'to expel'; *suʾāl* 'question' > *suʾaal* 'to ask'; *ṭālib* 'student; seeking' > *daalib* 'male student' (*daalib-id* 'female student') and 'to seek'; *ḥayāt* 'life' > *xayaad* 'to live'; *imtiḥān* 'examination' > *imtixaan* 'to examine'; *istiʿmār* 'colonization' > *isticmaar-so* 'to administer a country as a colony'; *ʾislām* 'surrendering to God; Islam' > *islaan* 'to become a Muslim'; *ʿāfiya* 'health' > *caafimaad* 'to recover health', *caafimaad-san* 'to be in good health'; *al-ḥamdu lillāhi* 'praise the Lord' > *alxamdulillay-so* 'to thank God'. Rather idiosyncratic is *akhri* 'to read', which may go back to the Arabic imperative (Callegari 1987–1988:448) or may be a reinterpretation of an Arabic causative *ʾaqraʾa* (cf. Somali *aqbal* 'to accept').

Some Arabic idioms and genitive constructions functioning as compound words have been borrowed, e.g. *bayt al-māʾ* 'w.c.' > *beytelmay; maʿa s-salāma* 'farewell!' > *macasalaamo* 'farewell', also used as a verb *macasalaamee* 'to say goodbye'.

Bibliographical references

Primary sources

Abraham, R.C. 1962. *Somali-English dictionary*. London: University of London Press.

Agostini, Francesco, Annarita' Puglielli, and Ciise Moxamed Siyaad. 1985. *Dizionario somalo-italiano*. Rome: Gangemi Editore.

Secondary sources

Callegari, Laura. 1987–1988. *I prestiti arabi nella lingua somala*. 2 vols. Tesi di Laurea, Università degli Studi di Torino.

Cardona, George R. 1988. "Somalia and the Indian Ocean: Cultural and linguistic contacts". *Proceedings of the Third International Congress of Somali Studies*, ed. Annarita' Puglielli, 22–35. Rome: Università "La Sapienza".

Saeed, O. Gulaid. 1973. *ʾAlfāḏ̣ ʿarabiyya fī l-luġa aṣ-ṣūmāliyya*. Aden: Maṭābiʿ aṯ-Ṯawrī.
Soravia, Giulio. 1994. “Gli imprestiti arabi in somalo”. *Sem – Cam – Iafet: Atti della 7a Giornata di Studi Camito-Semitici e Indoeuropei*, ed. Vermondo Brugnatelli, 199–224. Milan: Centro Studi Camito-Semitici.
Xuseen, M. Cabdi. 1984. “Note su alcuni prestiti nel somalo”. *Incontri Linguistici* 9.197–199.
Zaborski, Andrzej. 1967. “Arabic loan-words in Somali: Preliminary survey”. *Folia Orientalia* 8.125–175.

Andrzej Zaborski (University of Cracow)

Somalia

1. Land and people

On February 14, 1974, Somalia joined the League of Arab States and made Arabic an official language of the country. In doing so, Somalia became the first Arab League country to use a second official language along with Arabic. Somali, which had become the official language in 1973, was based on the spoken form, Af-Mahaa, of the northern and central regions of the Somali peninsula. A Roman-based script had been introduced by the state in 1972. After the collapse of the Somali state in 1991, speakers of other Somali languages introduced at least two additional Somali languages, again mostly in Roman scripts: Af-Maay (Mukhtar 2003:142) and Af-Jiddu (Ibro 1998), spoken in the interriverine and coastal districts of southern Somalia (Lamberti 1986). Today, Somalia is the only country in sub-Saharan Africa with no functioning central government, thus, no reliable data on literacy rates exist for any of the languages spoken there.

Somalis occupy the Horn of Africa, and most of them claim to be of Arab stock, particularly of the Quraysh clan, the household of the Prophet Muḥammad. All Somalis, 10–12 million, are Muslims, and they live in the former Somali Republic, the Republic of Djibouti, the Ethiopian Zone Five, the Northeastern Province of Kenya, and the diaspora in Canada, the United States, Europe, Australia, and the Middle East.

Historically, Somalia was known to Ancient Egyptians as part of the land of *mayddi* ‘myrrh’ or *Punt* ‘God’s Land’. Greeks and Romans traded in that part of the world, but Arabs, Persians, and Southeast Asians migrated and settled on the coast, influencing language and culture. Burton (1987:72) reports that Somali may be traced etymologically to the Arabic word *samala* ‘to thrust out’ because an Arab, presumably the progenitor of the Somali race, had thrust out his brother’s eye and run away to Somalia. Another theory suggests that the word is derived from the Arabic *ḏū māl* ‘the wealthy one’, referring to the son of a wealthy Arab who migrated with his assets to northeast Africa and fathered the Somali people (Drake-Brockman 1912:15). Somalia’s contacts with the Arabs date back to Pharaonic times (Neville 1894), when as early as the 3rd century B.C.E., South Arabians established trade links between the Horn and the western Indian Ocean world (Hourani 1951:20–21).

2. Status of Arabic in Islamic Somalia

Contacts between Arabs and Somalians became more prominent with the rise of Islam in Arabia in the 7th century C.E. The Meccan persecution, beginning in 615, of the Companions of the Prophet caused the wave of refugees to the Horn of Africa known as *al-Hijra ʾilā l-Ḥabaša* ‘the migration to Abyssinia’ (Ibn Hišām, *Sīra* 266). Successive migrations followed upon the *ridda* ‘apostasy’ and *fitna* ‘civil disobedience’ wars during the Orthodox Caliphate (632–661) and the Umayyad Dynasty (661–750). The interaction between the migrants and the indigenous population led to the emergence of Islamic centers on the Somali coast. The discovery of inscriptions on the tombstones in Mogadishu of Fāṭima bint ʿAbd aṣ-Ṣamad, who died on 22 Jumādā l-ʾŪlā 101/719, and Ḥāja bint Miqdām, who died on 5 Ḏū l-Ḥijja 138/755, indicates the presence of Islam in its first century (Cerulli 1957). The Somali lexicon was expanded by Arabic vocabulary derived from Islamic religion and civilization, especially in areas of theology, trade, politics, geography and seafaring, social relations, poetry, folk stories, and means of teaching (Zaborski 1967).

Arabic influence was at its greatest in the 9th century in towns on the Banadir coast, the Awdal-Harar corridor, and in the hinterland. Each city had a center famous for specialized teaching and learning. Mogadishu and Marka, for example, were strong in *tafsīr* and *ḥadīṯ*, and Bardera and Sarmaan for *ḥifḏ* 'memorization' of the *Qurʾān*. Students who excelled in the writing and reading of Arabic at an early age then pursued further learning at centers in Qulunqūl and Harar and elsewhere (Mukhtar 1995:10). Many graduates of the above centers went on to Mecca, Sanʾa, Kairouan, Damascus, and Cairo and returned home to teach and proselytize in the Horn and East Africa (Sālim 1964–1965:I, 353).

By the 13th century, a well-educated elite in Somalia spoke and wrote Arabic. During his sojourn in Mogadishu, Ibn Baṭṭūṭa met the sultan and commented, "His name is ʾAbū Bakr ibn Šayx ʿUmar, he is of Berber origin and speaks in the Maqdishi (local) language, though he knows Arabic" (*Riḥla* 169). In this period, the Somali instructors invented a new method of teaching Arabic known as *Laqbo* 'translation', so that students became bilingual and were able to master complex subjects through their mother tongue (Mukhtar 2003:136–137).

From the 15th to the 17th centuries, Arab-Islamic administrations such as the Faxruddīn and Muẓaffar dynasties flourished in the city-states of coastal Banadir in the south, while the Maxzūmī and Asmaʾa dynasties ruled in Awdal and Harar in the north. During these administrations, Arabic became the language of education, communication, and trade. The rulers and the administrative staff used Arabic titles, e.g. *sulṭān, šayx, ʾamīr, qāḍī, kātib, faqīh*, and *wazīr*. The *fatwas* were issued in Arabic, and appeals were typically addressed in writing to the sultan, who, after reviewing the cases, wrote his reply on the back of the appeal paper (Ibn Baṭṭūṭa, *Riḥla* 171).

3. ARABIC DURING AND AFTER THE COLONIAL PERIOD

During colonial times, Arabic remained the chief language of public communication, governance, and diplomacy. All colonial agreements and contracts between Somalis and foreigners had an Arabic text beside the colonial language (Sālim 1964–1965:II, 547–570). Somali sultans, chiefs, and leaders either signed or put their seal on the Arabic version (Mukhtar 1973). Arabic was also the language of correspondence between the leaders of the Somali militant movements fighting colonial occupation and the colonial officials, as well as between the Somali sheikhs of different *ṭarīqa* orders. Family or social correspondences, trees of genealogies, birth and death certificates, and many other significant documents were all recorded in Arabic (Cerulli 1957). The official colonial daily, *Corriere della Somalia*, in Italian Somalia, and the *Somaliland News*, in British Somaliland, always included Arabic supplements.

The constitutions, by-laws, and records of Somali political parties were kept in Arabic; they submitted their petitions to the government or to the United Nations in Arabic, and even their logos and names were associated with Arab political organizations and Muslim movements. The *Rābiṭa Waḥdat aš-Šabab aṣ-Ṣūmālī* 'Somali Youth League (SYL)', which emerged in the early 1940s, developed along the same lines as the emergence of the Young Arabs' *al-Fatā*, or the Young Turks during the 1920s. The sword, crescent, star, and the name used in the *Ḥizbiyya Dastūr Mustaqill aṣ-Ṣūmāl* (HDMS) 'Somali Independent Constitutional Party' flag are also typical emblems of Arab and Muslim movements. Moreover, the party anthems were all sung in Arabic: *Yā ʾayyuhā ḥizbiyyatun* of HDMS and *Yā ʾayyuhā ṣ-Ṣūmāliyyūn* of SYC (Mukhtar 1987:152). In 1950, both the SYL and HDMS asked the Trusteeship administration to recognize Arabic as the official language of the country (Somaliya 1969:24–25). And in 1951, the Territorial Council, in its deliberations on the future of independent Somalia, unanimously agreed that Arabic should be the national language (Somaliya 1969:47–50).

4. ARABIC AND EDUCATION

Arabic has always been the medium of instruction. The *duksi* Qurʾānic schools played a significant role in diminishing the level of illiteracy. Starting in the late 18th century, *Jamāʿa* settlements and roving *ḥīr* and *ṣūfīs* in the countryside vigorously promoted Arabic (Hersi 1977:266–70). During colonial times, Arabic was the second language of education. In fact, the Italian colonial administration created teacher training for the teaching of Arabic and religious subjects in all schools and levels.

The newsletters and bulletins of the Somali political parties were in Arabic: *al-Waḥda* of the SYL, *aš-Šaʿb* of the SDU, and *al-Qarn al-ʾIfrīqī* of the SNL (Mukhtar 1987:152). Many nonpolitical newspapers and magazines were in Arabic, for example *aṭ-Ṭalīʿa* and *al-Ḥaqīqa*. In the 1950s, Arab countries, particularly Egypt, supported Somalia with teachers for all grades and established a modern school system offering all subjects in Arabic and granting diplomas equivalent to Arab schools. In addition, scholarships were granted to the brightest high school graduates for further study in Cairo, Damascus, Baghdad, and elsewhere.

Even in independent Somalia, the national dailies, which were still printed in Italian and English, also published an Arabic supplement, *Barīd aṣ-Ṣūmāl*, until Af-Mahaa was adopted as the national language in 1972; even then the new *Xidigta October* 'October Star' continued to publish an Arabic supplement, *Najma ʾUktūbar*. Radio and television programs were broadcast in both Somali and Arabic formats until the collapse of the state in 1991. The use of Arabic remained pervasive even after the adoption of the Somali script in 1972 and after Somalia joined the Arab League in 1974.

5. SOMALI WRITERS IN ARABIC

The educated elite has continued to use Arabic in Islamic studies such as theology, philosophy, literature, *ḥadīṯ*, and history. ʿUṯmān ibn ʿAlī az-Zaylaʿī (n.d.) wrote one of the most authoritative books in the *Ḥanafī maḏhab*. His student and co-national ʿAbdallāh ibn Yūsuf az-Zaylaʿī (1938) became one of the best editors in *ḥadīṯ*.

In literature, ʿAbd ar-Raḥmān az-Zaylaʿī (1972) left behind a collection of religious poems. He also published a book on Arabic grammar and morphology (az-Zaylaʿī 1938). Sheikh Ṣūfī (n.d.) published a collection of mystic literature. Sheikh ʿUways al-Barāwī is considered the most prolific poet of his time, composing *qaṣāʾid*, poems in praise of Allah, of the Prophet, and of Sheikh ʿAbd al-Qādir al-Jilānī (ʿIlī 1954a).

Early-20th-century historians dealt with the ancient history of Somalia, its major kingdoms, and relations with its neighbors (ʿAydarūs 1955). Scholarly publications on the history of Islam and the emergence of sultanates in the coastal regions and the hinterlands are now being published (Rirash 1974). Colonial and postcolonial history and issues dealing with the history of anticolonial movements and leaders of political parties fighting for independence are covered by writers of the late 20th century (Mukhtar 1982). The turning point in the spread of Islam and the history of Sufi brotherhoods is documented by hagiographers (Ḥājjī Yūsuf 1912). The life stories of Sufi saints (ʿIlī 1954b) and the history of the rivalry between Sufi orders are well documented in Arabic. Out of five books, known as *Majmūʿāt al-Qulunqūlī*, al-Quṭbī (n.d.) dedicated two books, *Taḥḏīrāt bāliġa* and *Naṣr al-muʾminīn*, to the major doctrinal differences between Qādiriyya and Ṣāliḥiyya, which virtually justify the war against the Ṣāliḥiyya followers. There is a significant publication in Arabic on the history of the dervish movement and a biography of its leader (Aw Jamaʾ 1965a). Aw Jamaʾ (1965b) published another book on the founders of modern Somali political parties in the fight for independence.

Somali intellectuals in the late 19th and early 20th centuries unsuccessfully urged the adoption of the Arabic script for the Somali languages. Sheikh ʿUways al-Barāwī used Arabic script for his Af-Maay *qaṣāʾid* (Moreno 1955). Makāhīl (1926) also used Arabic for the northern Issaq dialect (Lewis 1958). In the 1960s, Mahamūd (1963) advocated the possibility of writing Somali using an Arabic script. In the 1970s, a similar study was produced by Gūled (1973), who discussed the roots of the Somali language and its relation with Arabic.

6. CONCLUSION

Arabic has been the language of trade, education, and religion throughout most of Somalia's recorded history. Because no Somali language was written until 1972, Arabic remained the major vehicle for record keeping and correspondence. From 1974, when Somalia joined the Arab League, Arabic became an official second language. The government launched a campaign in 1980 to promote its instruction nationwide, as it had done during the rural literacy campaign for Somali in 1974. Somalia developed for the first time two equal tracks for education, Somali and Arabic from elementary school to university. It also instituted an Arabic program at the Lafoole Teacher-Training College. After the collapse of the state, the country was left without formal education. However, efforts to reconstruct the country have seen a mushrooming of privately owned schools whose media of instruction are usually Arabic and English. Thus, Arabic has continued to flourish in all of the self-declared regional states.

Bibliographical references

Primary sources

Ibn Baṭṭūṭa, *Riḥla* = ʾAbū ʿAbdallāh Muḥammad Ibn Baṭṭūṭa aṭ-Ṭānjī, *Riḥlat Ibn Baṭṭūṭa*. Beirut: Dār al-Kitāb al-Lubnānī, 1960.

Ibn Hišām, *Sīra* = ʾAbū Muḥammad ʿAbd al-Malik Ibn Hišām, *as-Sīra an-Nabawiyya*, IV. Ed. Jamāl Ṯābit a.o. Cairo: Dār al-Ḥadīṯ, 1996.

Secondary sources

Aw Jamaʾ, ʿUmar ʿĪsā. 1965a. *Tārīx aṣ-Ṣūmāl fī l-ʿuṣūr al-wusṭā wa-l-ḥadīṯa*. Cairo: Maṭbaʿat al-ʾImām.

——. 1965b. *Tārīx zuʿamāʾ al-ḥaraka as-siyāsiyya fī ṣ-Ṣūmāl*. Mogadishu: Government Printing Agency.

ʿAydarūs, Šarīf ʿAydarūs ibn Šarī ʿAlī. 1955. *Buġyat al-ʾamal fī tārīx aṣ-Ṣūmāl li-baʿḍ mulūkihā wa-sukkānihā wa-ʿumrānihā wa-d-dīn allaḏī yaʿbudūnahu qabl al-ʾIslam bi-ṯamāniya qurūn ʾilā l-ʾān*. Mogadishu: Stamperia A.F.I.S.

Burton, Richard F. 1987. *First footsteps in East Africa or An exploration of Harar*, II. New York: Dover Publication.

Cerulli, Enrico. 1957, 1959, 1964. *Somalia: Scritti vari editi ed inediti* [Somali: Edited and unedited manuscripts], III. Rome: Istituto Poligrafico dello Stato.

Drake-Brockman, Ralph E. 1912. *British Somaliland*. London: Hurst and Blacket.

Gūled, Saʿīd ʿUṯmān. 1973. *Alfāḍ ʿarabiyya fī l-luġa aṣ-ṣūmaliyya: Baḥṯ maydanī*. Aden: Maṭābiʿ aṯ-Ṯawrī.

Ḥājjī Yūsuf ʿAbd ar-Raḥmān. 1912. *Kitāb rabīʿ al-qulūb fi ḏikr manāqib wa-faḍāʾil Sayyidinā aš-Šayx Nūr Ḥusayn*. Cairo: Maṭbaʿat al-Futūḥ al-ʾAdabiyya. [Contains another book, more on deeds.]

Hersi, Ali Abdirahman. 1977. *The Arab factor in Somali history: The origins and the development of Arab enterprise and cultural influence in the Somali Peninsula*. Ph.D. diss., University of California.

Hourani, George Fadlo. 1951. *Arab seafaring in the Indian Ocean in ancient and early Medieval times*. Princeton, N.J.: Princeton University Press.

Ibro, Salim Alio. 1998. *English Jiddu Somali mini-dictionary*. Victoria: La Trobe University.

ʿIlī, Šayx ʿAbdī. 1954a. *Jalāʾ al-ʿaynayn fī manāqib aš-šayxayn, aš-Šayx al-Wālī Ḥājjī ʿUways al-Qādirī wa-š-Šayx al-Kāmil ʿAbd ar-Raḥmān az-Zaylaʿi*. Cairo: Maṭbaʿat al-Mašhad al-Ḥusaynī.

——. 1954b. *Rāḥat al-qalb al-mutawalliʿ fī manāqib aš-Šayx ʿAbd ar-Raḥmān az-Zaylaʿi*. Cairo: Maṭbaʿat al-Mašhad al-Ḥusaynī.

Lamberti, Marcello. 1986. *Map of Somali dialects in the Somali Democratic Republic*. Hamburg: H. Buske.

Lewis, Ioan Myrddin. 1958. “The Gadabursi Somali script”. *Bulletin of the School of Oriental and African Studies* 21.134–156.

Mahamūd, ʾIbrāhīm Ḥāšī. 1963. *aṣ-Ṣūmāliyya bi-luġat al-Qurʾān: Muḥāwala waṭaniyya li-kitāba luġat al-ʾumm*. Cairo: Dār aṭ-Ṭibāʿa al-Ḥadīṯa.

Makāhīl, Šayx Muḥammad ʿAbdī. 1345/1926. *ʾInšāʾ al-mukātabāt al-ʿaṣriyya fī l-luġa aṣ-ṣūmāliyya*. Bombay.

Moreno, Martino M. 1955. *Il somalo della Somalia: Grammatica e testi del Benadir, Darod e Dighil*. Rome: Istituto Poligrafico dello Stato.

Mukhtar, Mohamed Haji. 1973. *Tārīx al-istiʾmār al-ʾīṭālī fī ṣ-Ṣūmāl ḥattā 1908*. M.A. thesis, Al-Azhar University, Cairo.

——. 1982. *aṣ-Ṣūmāl al-ʾĪṭālī fī fatrat al-wiṣāya ḥattā l-ʾistiqlāl, 1950–1960*. Ph.D. diss., Al-Azhar University, Cairo.

——. 1987. “Arabic sources on Somalia.” *History in Africa: A Journal of Method* 14.141–172.

——. 1995. “Islam in Somali history: Fact and fiction”. *The invention of Somalia*, ed. Ali Jimale Ahmed, 1–27. Lawrenceville, N.J.: Red Sea Press.

——. 2003. *Historical dictionary of Somalia*. 2nd ed. Lanham, Md.: Scarecrow Press.

Neville, Edward. 1894. *The Temple of Deir al-Bahri*. London: Kegan Paul.

Quṭbī, Šayx ʿAbdallāh ibn Muʿallim Yūsuf al-. N.d. *al-Majmūʿa al-mubāraka*. Cairo: Maṭbaʿat al-Mašhad al-Ḥusaynī. [Known also as *Majmūʿāt al-Qulunqūlī*; contains five books.]

Rīrāš, Šayx ʾAḥmad ʿAbdallāh. 1974. *Kašf as-sudūl ʿan tārīx aṣ-Ṣūmāl wa-mamālikihim as-sabʿa*. Mogadishu: State Printing Agency.

Sālim, Ḥamdī as-Sayyid. 1964–1965. *aṣ-Ṣūmāl qadīman wa-ḥadīṯan*. 2 vols. Cairo: ad-Dār al-Qawmiyya li-ṭ-Ṭibāʿa wa-n-Našr.

Somaliya. 1969. *Antologia storico culturale: Il problema della lingua somala* [Cultural and historical anthology: The problem of the Somali language], I. Mogadishu: Ministry of Education, Government Publishing Agency.

Ṣūfī, Šayx Muḥammad. N.d. *Dalīl al-ʿibād ʾilā sabīl ar-rašād*. Cairo: Maṭbaʿat al-Kaylānī aṣ-Ṣaġīr.

Zaborski, Andrzej. 1967. “Arabic loan words in Somali: Preliminary survey”. *Folia Orientalia* 8.125–175.

Zaylaʿī, Šayx ʿAbd ar-Raḥmān az-. 1938. *Fatḥ al-laṭīf šarḥ Ḥadīqat at-taṣrīf*. Cairo: Maṭbaʿa Muṣṭafā al-Bābī al-Ḥalabī.

——. 1972. *Majmūʿa muštamila*. Cairo: Maktabat al-Jundī.

Zaylaʿī, Šayx ʿAbdallāh ibn Yūsuf az-. 1938. *Naṣb ar-rāya fī taxrīj ʾaḥādīṯ al-hidāya*. 4 vols. Cairo: Dār al-Maʾmūn.

Zaylaʿī, Šayx ʿUṯmān ibn ʿAlī az-. N.d. *Tabyīn al-ḥaqāʾiq šarḥ Kanz ad-daqāʾiq*. 6 vols. Beirut: Dār al-Maʿrifa.

MOHAMED HAJI MUKHTAR
(Savannah State University)

Songhay

1. The Songhay Empire and Islam

Songhay is spoken by about 700,000 people in Mali, Niger, Upper Volta, Dahomey, and Nigeria and is the westernmost subgroup of Nilo-Saharan. Its dialects or closely related languages are Dendi (in Dahomey) and Dyerma (Zarma, especially in Niger), and the intercomprehension does not seem total. Songhay was considered an isolated linguistic group until Greenberg classified it as one of the six branches of the Nilo-Saharan family (Tersis 1972:17–18). It is a language relatively little described but on which we have information, some of it rather old. It is spoken by the descendants of those who founded, between the 7th and 17th centuries, what would become, thanks to Sonni Ali Beer (r. 1464–1492), the Songhay Empire. At the beginning of the 11th century, in the year 1010, the king, Dia Kossoi, became Muslim (Davidson 1966:67), and, according to al-Muhallabī (d. 380/990), "Gao was not just a town but also the centre of a small Islamic kingdom of the Niger Bend in his day. It therefore seems likely that the Zā dynasty, which perhaps replaced earlier local Muslim rulers in Gao, converted to Islam before 399–400/1009" (Hiskett 1984:32–33).

Tārīx as-Sūdān, which lists the first fourteen sovereigns of the Zā dynasty, states in this regard: "None among them believed in God nor in the Prophet. They died without embracing the Muslim faith". Then, the author goes on: "Zā Kosoï, who came after, accepted Islam. They nicknamed him *muslim-dam*, which means full Muslim. This event happened during the year 400 after Hegira [i.e. toward 1009]". Almost immediately after his conversion, Zā Kosoï moved the capital of Songhay from Kukiya to Gao, and al-Bakrī, writing in 1068, says: "The ruler of Gao professed Islam and when he was enthroned he was presented with a copy of the *Qurʾān*, a sword and a shield, all sent from the Caliph (leader) of the Muslim world who at that time resided in Baghdad" (Clarke 1982:47). But even if the king was a Muslim, as were all his dignitaries, "The ceremonials at the royal court continued to be based on traditional customs and beliefs, and the majority of the people in Gao were non-Muslims" (Clarke 1982:47).

The phonetic treatment of Arabic loans in Songhay is based on words collected in Ducroz and Charles (1978), but when the other dialects give different data for Arabic loans in Songhay, material from different sources is given.

2. Phonology

Neither northern nomadic Songhay nor eastern and western Songhay, including Korandje, seem to have a tonal system (Nicolaï 1981:25). Songhay is a language with a tendency to phonetic erosion of borrowed words. This phenomenon does not seem to correspond to the position of the syllable or to the consonant or vowel eroded in the word: *ṣadaqa* > *saraa* (in the dialects of Koyra Chiini and Djenné Chiini) 'alms'.

In Songhay, gemination is important: "Nine percent of dissyllabic words have a gemination in the intervocalic position" (Tersis 1972:54), and there is a tendency to preserve within loanwords the original Arabic geminates: *dābba* > *dábbè* (in the dialects of Koyra Chiini, Djenné Chiini, and Koroboro Senni, *addabba*) 'animal'; *ḥijj* 'pilgrimage to Mecca' > *hízzà* 'to perform the pilgrimage'; *ḥadd* > *híddí* 'border'; *šakk* 'doubt' > *síkkà* 'to doubt'; *in šāʾa llāh* > *ìnsállà* 'God willing'; *janna* > *àlzánnà* (in the dialects of Koyra Chiini and Koroboro Senni, *aljanna*; Dendi *àRzánnà*) 'paradise'. There are a few exceptions, which are probably due to colloquial use: *muʿallim* 'teacher' (? via Fulfulde *mallum*) > *maale* 'patron', or to the interference of the Berber intermediary *wa-aḥažžam: ḥajjām* 'cupper' > *wànzâm* 'barber' (cf. Fulfulde *wanjamījo*, Hausa *wànzāmì̄*, Kanuri *wanzâm*).

Individual Arabic phonemes are realized in Songhay as follows (for more details, see Baldi 1994):

/ʾ/ > Ø: *ʾiblīs* > *ìblísì* 'devil, Satan'; *al-qurʾān* > *àlkùráàn* 'the *Qurʾān*'; *al-ʾarbiʾāʾ* > *àlárbá* 'Wednesday'; but /ʾ/ > /w/ only in one case: *balāʾ* 'misfortune, plague' > *bàláw* 'epidemic';

/b/ > /b/: *al-baraka* > *àlbárkà* 'blessing'; *bāliġ* 'mature; of age, legally major' > *bàlíjì* 'adult'; *muṣība* > *màsíibà* 'misfortune'; *al-ʿarab* > *làaráabù* 'Arab'; but /b/ > /p/ only in one case: *as-sabt* > *àsíptì* 'Saturday';

/t/ > /t/: *tawʾam* > *táwéy* ‘twin’; *fatīl* > *fìtíllà* ‘lamp’; *al-waqt* > *wáatì* ‘time’; but in one case /t/ > /l/: *ḥattā* > *hál* ‘until, up to’, probably via Hausa (*har*) and Tamasheq (*ar*); in fact, for Dendi we have *háR*;

/ṯ/ > /t/: *ṯawāb* ‘recompense, reward’ > *tìáabù* ‘recompense, pardon of God’; *al-iṯnayn* > *àtínnì* ‘Monday’;

/j/ > /z/: *jahannam* > *zàhánnàm* ‘hell’; *daraja* ‘rank’ > *dárzà* ‘importance, glory’; *ḥijj* ‘pilgrimage’ > *hízzà* ‘to perform the pilgrimage’; but /j/ > /j/ in some cases: *jamāʿa* ‘community’ > *jámâ* ‘crowd, people’; *al-jāhil* > *àjáhílì* ‘ignorant’; in one case /j/ > /d/, probably because of regressive assimilation: *sujūd* ‘prostration, adoration’ > *sùdúudù* ‘to prostrate oneself, to worship’; in a doubtful loan we have *jawwada* II ‘to recite [the *Qurʾān*]’ > *céw* (and in Dendi *tyóó*) ‘to read’. In the dialect of Koroboro Senni, we also have /j/ > /ž/: *jumʿa* > *alzuma* ~ *alžumaa* ‘Friday’;

/ḥ/ > /h/: *al-ḥurma* ‘reverence, deference; that which is holy, sacred’ > *àlhórmà* ‘grace, favor’; *al-ʾaḥad* > *àlàhádì* ‘Sunday’; *lawḥ* > *wàlàhā́* ‘slate, board’; but /ḥ/ > Ø in final position, almost always, and sometimes in initial position because of assimilation with the Arabic article: *al-qamḥ* > *àlkámà* ‘wheat’; *al-ḥisāb* ‘arithmetic’ > *làsáabù* ‘to reflect, point out’;

/x/ > /h/: *xaṭṭ* ‘writing’ > *hàntúm* ‘to write’; *al-ʾaxbār*, pl. of *al-xabar* > *àlhàbáarù* ‘news’; /x/ > Ø because of assimilation with the Arabic article: *al-xayma* ‘tent’ > *léemà* ‘umbrella’; but sometimes /x/ > /k/: *baxīl* > *bàkíilù* ‘avaricious’; *al-xamīs* > *àlkàmíisá* ‘Thursday’;

/d/ > /d/: *dalīl* ‘sign’ > *dàlíilì* ‘reason’; *ʾabadan* ‘forever; ever, never’ > *àbádá* ‘never’; *al-ʾaḥad* ‘Sunday’ > *àlàhádì*; in some cases there is assimilation to the preceding consonant: *walad* ‘child, son, boy’ > *wáddè* ‘companion of the same age’; /d/ > /r/: *ṣadaqa* ‘alms; almsgiving’ > *sárgà* ‘to make a sacrifice, an offering’; in a doubtful case, it disappears: *jawwada* II ‘to recite [the *Qurʾān*]’ (?) > *céw* ‘to read’, and it also disappears in the dialects of Koyra Chiini, Djenné Chiini, and Koroboro Senni: *ṣadaqa* ‘alms; almsgiving’ > *saraa* ‘alms; to give (something) as alms’;

/ḏ/ > /z/ in the initial position of a word: *ḏanb* ‘sin, crime’ > *zàmbà* ‘to trick’; in other cases /ḏ/ > /d/: *al-ʾaḏān* ‘call to prayer’ > *àlàadân* ‘muezzin’. In the dialects of Koyra Chiini and Djenné Chiini, /ḏ/ > /j/ and/or /ž/: *al-ʿaḏāb* ‘pain, torment, torture; punishment’ > *laajaaba/laažaaba* ‘suffering; to make suffer’;

/r/ > /r/: *ribḥ* ‘interest [on money]’ > *ìríibì* ‘benefit’; *al-ʾarbiʾāʾ* > *àlárbá* ‘Wednesday’; *al-ʿaṣr* > *àláasárú* ‘afternoon; afternoon prayer’; sometimes in Kaado (and more often in the dialects of Koyra Chiini and Djenné Chiini), it is geminated with the following consonant: *wird* ‘private worship’ > *wíddì* ‘to recite the Koran’; in two cases, /r/ > Ø: *rizq* > *àlzàkà* ‘wealth’; *rikāb* > *àlcébù* ‘stirrup’; once in Kaado, /r/ > /l/: *barīm* ‘rope; string, cord, twine’ > *bìlím* ‘to roll in the earth’, and quite regularly in the dialects of Koyra Chiini and Koroboro Senni: *ar-riyāl* > *allaara/alliyaara* ‘rial; a silver coin’; but in the Songhay of Koyra Chiini, we have /r/ > /d/: *bārūd* ‘saltpeter; gunpowder’ > *albaadũ* ‘(gun)powder’;

/z/ > /z/: *zamān* > *zàbànì* ‘time’; *ar-rizq* > *àlzàkà* ‘wealth’; but in the dialects of Koyra Chiini and Djenné Chiini, we have /z/ > /j/: *zamān* ‘time’ > *jaman* ‘era, season’;

/s/ > /s/: *as-sabt* > *àsíptì* ‘Saturday’; *ṭāsa* > *táasà* ‘metal bowl’; *ʾiblīs* > *ìblísì* ‘devil, Satan’. In the dialect of Koroboro Senni, we have /s/ > /š/: *ʾislām* ‘the religion of Islam’ > *alšilaama* ‘Muslim’;

/š/ > /s/: *šahāda* > *sàhádù* ‘creed formula’; *ʿušr* > *àláasárú* ‘one-tenth; tithe’, as well as in the dialects of Koyra Chiini and Djenné Chiini: *šarṭ* ‘condition; stipulation [of a contract]’ > *satti/setti/serti/sarti* ‘deadline; fixed date or schedule’. In the dialects of Koyra Chiini and Djenné Chiini, there are a few cases where /š/ > /š/: *šarīʿa* > *aššaraa* ‘Muslim law’; *šayṭān* > *aššeytaan* ‘Satan, devil’;

/ṣ/ > /s/: *ṣubḥ* ‘dawn; daybreak; morning’ > *súbà* ‘tomorrow’; *muṣība* ‘misfortune’ > *màsíibà*; *maqāṣṣ* [pl.] ‘scissors’ > *mágásà*; but in the dialects of Koyra Chiini and Koroboro Senni, we have /ṣ/ > /š/: *ʾaṣl* > *aššil/aššel* > ‘origin’. In Kaado, once /ṣ/ > /z/: *ṣirf* ‘pure, unadulterated’ > *nzórfù* ‘silver’, with a Tamasheq intermediary *âżref*;

/ḍ/ > /l/: *al-qāḍī* > *àlkáali* ‘judge’; *al-waḍāʾa* > *àlwàláà* ‘ceremonial purity’; but /ḍ/ > /d/ in *aḍ-ḍuḥā* ‘forenoon’ > *àddùhá* ‘morning, at about 9 a.m.’;

/ṭ/ > /t/: *ṭālib* ‘student, scholar’ > *táalíbì* ‘disciple of a marabout’; *saṭl* ‘bucket, pail [of wood or metal]’ > *sàtállà* ‘kettle [for ritual ablutions]’; but /ṭ/ > /d/ in *qarṭas* ‘paper’ > *kàrdáasì* ‘paper not written’; and in one case, /ṭ/ > /c/: *fiṭr* ‘fast breaking’ > *wícíró* ‘late afternoon greetings’;

/ḏ̣/ > /z/: *waʿḏ̣* ‘admonition; sermon’ > *wáazù* ‘to pray; to indoctrinate’, but in the dialects of Koyra Chiini and Djenné Chiini it becomes *waaju* ‘to pray; to give religious advice’;

/ʿ/ > Ø: *al-ʿāda* > *àláadà* ‘habit, custom’; *jumʿa* > *àlzúmá* ‘Friday’; *midfaʿ* > *málfà* ‘gun, cannon’;

/ġ/ > /g/ or Ø: *al-ġarīb* ‘strange’ > *àlgàríbù* ‘to beg’; *maġrib* ‘place or time of sunset; prayer at sunset’ > *àlmáaríŋ* ‘evening’; in one case, /ġ/ > /j/: *bāliġ* ‘mature; of age, legally major’ > *bàlíjì* ‘adult’, but in the dialect of Koyra Chiini, it becomes *baaliči/baaliki*;

/f/ > /f/: *al-fitna* > *àlfítínà* ‘discord, dissension’; *ʿafā* > *yàafà* ‘to forgive’; *šarīf* ‘sherif, title of the descendants of Muḥammad’ > *sìfírì* ‘rainbow; sherif’;

/q/ > /k/: *qalam* ‘pen’ > *kàlámì* ‘pen, pencil’; *al-ʿaql* > *làkkál* ‘intelligence’; *al-munāfiq* > *àlmùnàafíkì/mùnàafíkì* ‘hypocrite’; but also /q/ > /g/: *al-qadar* > *algadar* ‘fate, destiny’; *ṣadaqa* ‘alms, charitable gift; legally prescribed alms tax’ > *sárgà* ‘to make a sacrifice, an offering’; there are a couple of cases in Kaado (but very frequently in the dialects of Koyra Chiini, Djenné Chiini, and Koroboro Senni) where

/q/ > Ø: *al-waqt* > *wáatì* ‘time’; *šafaq* ‘evening glow, twilight, dusk’ > *sáafó* ‘evening (religious word)’. In the dialect of Koyra Chiini, we have *qibla* ‘kiblah, direction to which Muslims turn in praying [toward the Kaaba]’ > *alčibla* ‘east’;

/k/ > /k/: *kull* > *kúlú* ‘totality; all’; *wakīl* > *wàkíilù* ‘representative’; *šakk* ‘doubt’ > *síkkà* ‘to doubt’; in a couple of cases, /k/ > /c/: *kāfir* ‘infidel’ > *céférì* ‘to be infidel’; *rikāb* > *àlcébù* ‘stirrup’;

/l/ > /l/: *luqma* > *lóomò* ‘bite; mouthful’; *dawla* ‘power’ > *dóolè* ‘to force’ (cf. Fulfulde *dōle*, Hausa *dōlè*, and Swahili *dola*); *baxīl* > *bàkíilù* ‘avaricious’; in one case, there is assimilation with the following consonant: *walad* ‘child; son; boy’ > *wáddè* ‘companion of the same age’; and in another one /l/ > /j/: *šaġala* > *sàajì* ‘to occupy, to keep someone busy’; on the contrary, in a couple of cases the gemination disappears: *taʿallala* V ‘to distract oneself; to use as an expedient’ > *taali* ‘to wrong someone’; *kull* > *kúlú* ‘totality; all’; but in Dendi we have /l/ > /r/: *ʾaṣl* > *àsàRíí* > ‘origin’. In all dialects, in final position, in a few cases /l/ > /n/: *māl* ‘property; wealth; money; goods’ > *àlmén* ‘domestic animal’. One case of gemination with the following consonant: *walad* ‘descendant, offspring; child, son, boy’ > *wáddè* ‘same-age companion’;

/m/ > /m/: *malāʾika*, pl. of *malʾak* > *màléykò* ‘angel’; *ʾammā* > *àmá* ‘but’; *jahannam* > *zàhánnàm* ‘hell’; in one case /m/ > /b/: *zamān* > *zàbànì* ‘time’;

/n/ > /n/: *nafaʿa* > *nàfà* ‘to be useful’; *al-janna* > *àlzánnà* ‘paradise’; *dīn* > *àdíinà* ‘religion’; but /n/ > /m/ before /b/: *ḏanb* ‘sin, crime’ > *zàmbà* ‘to cheat’; sometimes /n/ > Ø, particularly in final position: *ʾabadan* > *àbádá* ‘never’;

/h/ > /h/: *hamm* ‘worry’ > *àlhém* ‘unhappy’; *šahāda* > *sàhádù* ‘creed formula’; in final position /h/ > Ø: *wa-llāh* ‘by God!’ > *wálá* ‘or this [at end of an utterance]’ and *wàl* ‘or that [disjunctive morpheme]’;

/w/ > /w/: *waqt* > *wáatì* ‘time’; *dawā* ‘inkwell’ > *dá* /*dáwà* ‘ink’;

/y/ > /y/, but very often it disappears: *bayān* ‘explanation’ > *béy* ‘to know’; *dunyā* > *àdùñà* ‘world’; *walīy* > *wèĺì* ‘holy man’; *al-yatīm* > *àlàatîm* ‘orphan’.

3. MORPHOLOGY AND SYNTAX

Arabic article. In Songhay, the Arabic definite article *al-* is not retained as frequently as in Hausa. If it appears in loans, it is very rare in Kaado compared to the dialects of Koyra Chiini, Djenné Chiini, and Koroboro Senni: in fact, we have: *baxīl* > *bàkíilù* ‘avaricious’ besides the variants *albačir* (Koyra Chiini) and *albahiiri* (Koroboro Senni) or *zināʾ* ‘adultery; fornication’ > *zìnà* ‘to commit adultery’ besides the variants *azzinaa* ‘(act of) adultery’ (Koyra Chiini) and *aljinaa* ‘adulterer’ (Djenné Chiini). On the other hand, in Kaado we notice the agglutination of the Arabic article, modified according to what happens in Arabic Maghrebi dialects: *ḥisāb* ‘arithmetic, reckoning; calculation’ > *làsáabù* ‘to reflect, observe’. This phenomenon is also found in Dendi: *ʾimām* > *léémàm* ‘imam’.

Nouns. Nominal loanwords mostly derive from the singular form; loans coming from plural forms are rare: *ʾaxbār*, pl. of *xabar* > *àlhàbáarù* ‘news’; *naṣārā*, pl. of *naṣrānī* > *ànsáará* ‘Christian’. As in Dagbani and Hausa, there is one case of a double loan coming from a singular and plural form: *ḏanb* ‘sin, crime’ > *zàmbà* ‘to cheat’ (cf. Dagbani *samba* ‘slanderer’, Hausa *zàmba* ‘fraud’) and *ḏunūb* ‘sin [pl.]’ > *zùnúbù* ‘to sin’ (cf. Dagbani *zunubi* ‘sin’, Hausa *zùnubìi/zùnufìi* ‘sin’).

Verbs. Loans in Songhay are almost all nominal; verbal loans are very few, less than 10 percent (Baldi 2004: Table of loanwords); most of these derive from the Arabic imperfect: *yaʿfā* (< *ʿafā* ‘to forgive’) > *yàafà* ‘forgiveness; to forgive’. This is found not only in Kaado but also in other dialects: *yaʾmuru* (< *ʾamara* ‘to order, command’) > *yaamar* ‘to predominate, order’ (Koyra Chiini). There is a tendency in the Kaado dialect to borrow nominals with the Arabic article as verbs: *ḥuzn* ‘sadness’ > *àlhúuzù* ‘to try to frighten someone’; *ḥanna* ‘to pity, have mercy’ > *àlhánnà* ‘to ask forgiveness’; *ġarīb* ‘strange’ > *algaribu* ‘to beg’; *faʾl* ‘good omen’ > *àlfáalà* ‘to make a wish come true’.

Semantic analysis. Arabic loanwords are found in all domains of Songhay cultural vocabulary: jurisprudence (*àlkáalì* < *al-qāḍī* ‘judge’), trade (*táazírì* ‘rich man’ < *tājir* ‘merchant’), religion (*hàrdám̀* ‘to be unlawful, forbidden by religion’ < *ḥarām* ‘forbidden, unlawful’), science and culture (*làkkál̀* < *ʿaql* ‘intelligence’; *táalíbì* ‘Qurʾānic student’ < *ṭālib* ‘student, scholar’), names of some everyday objects (*sàafún* < *ṣābūn* ‘soap’).

Songhay as a medium of spreading Arabic loans. Songhay not only received Arabic loans but it was also a donor language for Arabic loans in many languages in West Africa: Arabic *faqīh* ‘legal scholar, expert of *fiqh*’ > Songhay *àlfá* ‘(Muslim) priest’ > Busa *arfā̀ná*; Dagbani *álefa*; Ewe *alufa*; Fulfulde *alfaa*; Kanuri *alfakí*; Mandinka *àlfaa*; Nupe *àfâ/àlùfâ*; Serer *alfa*; Yoruba *àlùfáà*.

Bibliographical references

ʿAbd ar-Raḥmān ibn ʿAbdallāh ibn ʿImrān ibn ʿĀmir as-Saʿdī, *Tārīx as-Sūdān*. Transl. from Arabic by Octave Houdas, *Tarikh-es-Soudan*. (= *Publications de l'École des Langues Orientales Vivantes. Documents arabes relatifs à l'histoire du Soudan.*) Paris, 1981.

Baldi, Sergio. 1994. "Le traitement phonétique des emprunts arabes dans le songhay (parler kaado du Gorouol)". *Deuxième Table Ronde Internationale du Réseau Diffusion Lexicale en Zone Sahelo-Saharienne (Prague 23–28 Août 1993)*, II, ed. Petr Zima, Stanislava Boušková, and J. Urbanová, 12–20. Prague: Center for Theoretical Study.

——. 1995. Review of: Zima (1994). *Annali dell'Istituto Universitario Orientale* 55:3.373–374.

——. 2004. "L'influence de la langue arabe en Afrique". *Mémorial Vycichl*, ed. Kamal Naït-Zerrad and Gabor Takács. Frankfurt a. Main: Johann Wolfgang Goethe-Universität.

Clarke, Peter B. 1982. *West Africa and Islam: A study of religious development from the 8th to the 20th century*. London: E. Arnold.

Davidson, Basil. 1966. *A history of West Africa to the nineteenth century*. New York: Anchor Books.

Ducroz, Jean-Marie and Marie-Claire Charles. 1978. *Lexique soŋey (songay)-français, parler kaado du Gorouol*. Paris: L'Harmattan.

Heath, Jeffrey. 1998. *Dictionnaire songhay-anglais-français*. I. *Koyra chiini*; II. *Djenné chiini*; III. *Koroboro senni*. Paris: L'Harmattan.

Hiskett, Mervyn. 1984. *The development of Islam in West Africa*. London and New York: Longman.

Nicolaï, Robert. 1981. *Les dialectes du songhay: Contribution à l'étude des changements linguistiques*. Paris: SELAF.

Tersis, Nicole. 1972. *Le zarma (République du Niger): Étude du parler djerma de Dosso*. Paris: SELAF.

Zima, Petr. 1994. *Lexique dendi (Songhay): (Djougou, Bénin), avec un index français-dendi*. Cologne: R. Köppe.

Sergio Baldi
(University of Naples)

Soqoṭri → South Arabian, Modern

Sorani → Kurdish

Sound Symbolism

1. Definition of sound symbolism

In the introduction to their standard work on sound symbolism, Hinton a.o. (1994b) distinguish four types of sound symbolism, which they define as "the direct linkage between sound and meaning" (1994b:1):

(i) Corporeal sound symbolism: "the use of certain sounds or intonation patterns to express the internal state of the speaker, emotional or physical" (Hinton a.o. 1994b:2);
(ii) Imitative sound symbolism: "the use of onomatopoeic words and phrases representing environmental sounds" (Hinton a.o. 1994b:3), e.g. sounds of animals, often conventionalized when represented in writing; this is what Ullmann (1962:83) calls 'primary onomatopoeia';
(iii) Synesthetic sound symbolism: "the acoustic symbolization of non-acoustic phenomena" (Hinton a.o. 1994b:4); this is what Ullmann (1962:83) calls 'secondary onomatopoeia'; and
(iv) Conventional sound symbolism: "the analogical association of certain phonemes and clusters with certain meanings" (Hinton a.o. 1994b:5).

2. Sound symbolism in the Arabic linguistic tradition

For most Arabic grammarians, phenomena of sound symbolism held no great interest, and the grammarians generally regarded them as linguistically irrelevant. Sound symbols of the first three categories were subsumed under the category of → *ism al-fiʿl*, a term that comes closest to the modern concept of → 'interjection' (Sībawayhi, *Kitāb* I, 122–128; Mubarrad, *Muqtaḍab* III, 202; Levin 1991). Az-Zamaxšarī (*Mufaṣṣal* 61–67) mentions many of these in a list, distinguishing between those *ʾasmāʾ al-fiʿl* which express an order and those which express a predication. The category of those interjections expressing an order contains words like *ruwayda* 'gently!', *hālumma* 'get up!, onward!', *balha* 'let alone!', *ṣah* 'hush!', *ḥaḏāri* 'careful!', etc. The category of interjections expressing a predication is probably identical with the one az-Zamaxšarī designates as *ʾaṣwāt* 'sounds' (*Mufaṣṣal* 66.2), e.g. *way* as an expression of surprise or regret, and *miḍḍi* as an expression of refusal by smacking one's lips.

The inclusion of such sounds in az-Zamaxšarī's list often seems to have been inspired by the fact that they occur in the *Qurʾān* or in poetry. *Way*, for instance, occurs in the verse

> *waykaʾannahu lā yufliḥu l-kāfirūna*
> ‘Woe, the unbelievers are not successful’
> (Q. 28/82)

and *miḍḍi* is used in a poetic line

> *saʾaltuhā l-waṣla fa-qālat miḍḍi*
> ‘I asked her for a reunion and she said “pfui”’

To the same category in az-Zamaxšarī’s list belong various sounds used to spur animals on, e.g. *kixxā* or *halā* to a horse, *ʿadas* to a mule, and *hayda* (or *haydi* or *hād*) to a camel. Az-Zamaxšarī calls some of these nouns an ‘imitation’ (*ḥikāya*) of animal sounds (including calls to animals) and natural sounds. The word *ṭīxi*, for instance, is an imitation of the sound of someone laughing, *māʾi* an imitation of the bleating (*buġām*) of a goat, and *ṭāqi* an imitation of the sound of hitting.

Grammatically, the Arabic grammarians analyzed all *ʾasmāʾ al-fiʿl* as nouns, apparently because they could be used with an indefinite ending. The sound of a crow, for instance, is represented as *ġāqi*, but it may be used in a sentence like *qāla l-ġurābu ġāqin* ‘the crow uttered one “caw”’, where it is supplied with the → *tanwīn* ending to indicate indefiniteness (Zajjājī, *ʾĪḍāḥ* 99). A modern Arab linguist, Tammām Ḥassān (1973), proposed the name *xawālif* (sg. *xālifa*) for a new part of speech, comprising all those words with an affective function that are used to express emotional feelings, including the *ʾasmāʾ al-ʾafʿāl* of the classical Arabic grammarians and such expressions as *niʿma*/*biʾsa* ‘how good/bad is…’. He distinguished between assertive statements in a sentence like *ʾamdaḥu zaydan* ‘I praise Zayd’ and → exclamations, such as *niʿma zaydun rajulan* ‘what a wonderful man Zayd is!’, calling the latter *xawālif* (Firanescu 2003:108–111). In this, he may have followed an otherwise unknown grammarian, Ibn Ṣābir, whose views are reported by as-Suyūṭī (*ʾAšbāh* III, 2.8–9), and who proposed to call the → *ism al-fiʿl* by this name.

The task of making a complete inventory of linguistic representations of sounds fell to the lexicographers, who were particularly interested in onomatopoetic words of the second category, imitations of natural and animal sounds. Ibn Sīdah (d. 458/1968) painstakingly lists the sounds of the animals in his *Muxaṣṣaṣ*, but always from a lexicographical point of view. He says, for instance, that the sound of a horse is called *ṣahīl* (*Muxaṣṣaṣ* VI, 157), that of a cow *xuwār* (*Muxaṣṣaṣ* VIII, 41), that of a donkey *nahīq* (*Muxaṣṣaṣ* VIII, 49), that of a lion *zaʾīr* or *nahīt* (*Muxaṣṣaṣ* VIII, 64), and that of a hyena *ḥafḥafa* (*Muxaṣṣaṣ* VIII, 72). In all such cases, the noun is the name of the sound, not the sound itself, although obviously in some cases, as in that of the hyena, it does imitate the sound.

When speaking about the sounds of birds (*Muxaṣṣaṣ* VIII, 133–136), Ibn Sīdah adds several times that the noun is an imitation (*ḥikāya*) of the sound the bird makes, for instance when he says that *ʿaqʿaqa* is an imitation of the sound of the *ʿaqʿaq* ‘magpie’ (*Muxaṣṣaṣ* VIII, 136.6). Interestingly, *hadhada* is said to be the name of the sound of the pigeon, and Ibn Sīdah states that the name of the hoopoe, *hudhud*, is derived from this, although the verb denoting the sound of the hoopoe is *nabaḥa* ‘to bark’ (*Muxaṣṣaṣ* VIII, 136.3); this verb probably refers to the screeching sound the hoopoe makes when it flies up, rather than to the better known sound to which it owes its name (on the hoopoe in Islam, see Venzlaff 1994). Such verbs, nouns, and adjectives were collected in the same way as the names for an animal’s young and the correct noun to be used for a flock or a herd of animals. The lexicographers were mainly interested in knowing what the correct Arabic word was, rather than the sound itself. Thus, they limited themselves to listing the lexical items referring to the sounds of the animals and did not attempt to imitate the sounds as such.

In his *Fiqh al-luġa*, the lexicographer aṯ-Ṯaʿālibī (d. 429/1038) goes one step further, because he distinguishes between actual sound imitations and the name of the sound (cf. Shivtiel 2000). In the section about *ʾaṣwāt wa-ḥikāyāt* (*Fiqh* 202–216), aṯ-Ṯaʿālibī states that the Arabs call the sound of the crow *ġāqi ġāqi* and the sound of hitting *ṭāqi ṭāqi*. He then adds that the *ḥikāya* of the latter sound is *ṭaqṭaqa*. A special case of *ḥikāya* is that of the sound of a horse’s hooves on the ground, *ḥabaṭiqṭiq*, which derives from a poem by an anonymous poet (*Fiqh* 216.2; cf. Xalīl, *ʿAyn* III, 339):

> *jarat al-xaylu fa-qālat ḥabaṭiqṭiq*
> ‘The horses walked and went clippety-clop’

Aṯ-Ṯaʿālibī mentions yet other cases of *ḥikāya*, such as *ġiq ġiq* (the sound of a kettle boiling) and *šiba šiba* (the sound of a camel drinking water). In general, his terminological use of the terms seems to be as follows. When he calls the sound of a donkey *nahīq*, what he means is that the substantive *nahīq* signifies 'the braying of a donkey'. But since *ṭāqi ṭāqi* represents the sound (of hitting) itself, *ṭaqṭaqa* is called its *ḥikāya*. Likewise, *qah qah* is the sound of someone laughing; therefore, its *ḥikāya* is called *qahqaha* 'to roar with laughter, to guffaw'. Words like *faxīx, naxīx, ġaṭīṭ*, and *xajīf*, which indicate increasingly noisy sounds of snoring made by someone sleeping, are called sounds (*ʾaṣwāt*) rather than *ḥikāyāt*. This looks like a terminological difference between straightforward onomatopoeia, on the one hand, and names of sounds, such as *ṣafīr* 'chirping', *nahīq* 'braying [of a donkey], *ṯuġāʾ* 'roaring [of a lion]', *muʿāʾ* 'mewing [of a cat]', *ṣaʿīy* 'trumpeting [of an elephant]', etc., on the other. Only rarely does aṯ-Ṯaʿālibī call such a sound a *ḥikāya*, e.g. *qahqāʿ*, which is the *ḥikāya* of the sound of a bear (*Fiqh* 211), and *baqbaqa*, which is the *ḥikāya* of the bubbling sound of a jug or jar submerged in water (*Fiqh* 212.15).

It should be added that in grammatical literature, the term *ḥikāya* had been in use for a long time for a mixed category of elements, classified according to both syntactic and morphological elements. Elements assigned by Sībawayhi to this category on the basis of syntactic criteria seem to include direct quotations, for instance when asking *man zaydin* 'Zayd who?' of someone who says *marartu bi-zaydin* 'I passed Zayd' (for a thorough analysis of this use of *ḥikāya*, see Baalbaki 2007:13–21). In this sense, *ḥikāya* is not connected with the concept of onomatopoeia. The morphological classes that are assigned to the category of *ḥikāya* do seem to have such a connection; they include the opening letters in some of the Qurʾānic suras, the names of the letters of the alphabet, and onomatopoetic words, such as *ġāqi* for the sound of the crow (Talmon 2003:220–222).

3. Sound symbolism in the modern dialects

With respect to onomatopoeia and sound symbolism in the contemporary Arabic dialects, there are relatively few references in the literature. Only one modern dialect, Nigerian Shuwa Arabic, seems to have developed an entirely new class of ideophones, probably due to contact with neighboring languages, such as → Kanuri (Owens 2004). This class of ideophones has become grammaticalized to an extent not found in other Arabic dialects. Ideophones are used to indicate sounds, e.g. *al-bagar akalan al-qalla rūs* 'the cows gnashed up the grain', *al-iyāl liʾib fi alme cabaq* 'the children splashed in [the] water', where *rūs* and *cabaq* are ideophones connected with the sound of gnashing and splashing, respectively (Owens 2004:211). There are also ideophones for qualities, for instance *co* to indicate hotness in *al-alme hāmi bilhēn co* 'the water is very hot' (Owens 2004:210).

In the other Arabic dialects, sound symbols have not become grammaticalized to such an extent, and, grammatically speaking, they play a marginal role. Nonetheless, as in other languages, the lexicon of Arabic dialects contains a large number of sound symbols in the categories mentioned above under (i) and (ii) (corporeal and environmental sounds). An unexpected data source for such symbols are Arabic comics, usually translations of Western comics, such as Asterix (*ʾAstirīks*) or Tintin (*Tān Tān*). The language of most of these comics is strictly Standard Arabic, but it seems safe to assume that the representations of sounds derive from spoken speech. Comics abound with representations of corporeal and environmental sounds, and although these may occasionally be the result of the translation process, most of them do seem to be peculiar to Arabic, rather than imitations of their representation in the French original (see Table 1). Some sounds are used in swearing or insulting other people, e.g. *ʾaxx* or *ʾaḫḫ* (→ taboo).

Not much is known about imitations of the sounds of animals in Arabic dialects, because these are not usually listed in dictionaries and wordlists. It seems to be the case that the verbs referring to these sounds are often reduplicated, e.g. in Egyptian Arabic *nawnaw* 'to mew', *hawhaw* 'to bark' (cf. Badawi and Hinds 1986:917 *lamma thawhaw* '[not even] if you should bark [i.e. never]'), *ṣawṣaw* 'to chirp', *kākkāk* 'to quack [duck]' (but according to Badawi and Hinds 1986:728, *kāk* is the sound of the crow, as in the proverb *zayyi gamʿiyyit ilġirbān ʾawwilha kāk wi-ʾāxirha kāk* 'like a flock of crows, the first [sound] is "caw", and the last one is "caw"'). Imitations of animals

Table 1. Representation of corporeal and environmental sounds in comics

Transcribed sound	Context
būf	sound of engine failing
ṭāx	sound of hitting or gunshot
ṭrāx	sound of crashing into something; intensive form of *ṭāx*
tāk	sound of knocking
hrrr	sound to chase an animal
tṣaw tṣaw	sound of the creaking of a rusty machine
trrrn	sound of the ringing of a doorbell, an alarm, or a telephone
krāš	sound of glass breaking
klang	sound of bell of fire engine
krāk	sound of something breaking
būm	sound of heavy object falling down
tūt	sound of traction engine or the horn of a ship or a car
tš	sound of water splashing
hiʾ	sound of hiccup
yāh	sound of surprise
ʾāy	cry of pain
ʾāh	sound of fear or pain
xāxx	sound of snoring
kr kr kr	sound of gulping down a drink
ʾāx	sound of disgust
ʿaṭsa	sound of sneezing
fš	hissing sound
zībīʾ	sound of someone slipping or of screeching tires of a car skidding
ṭūf	sound of disgust, ugh!, yuck!

seem to have the same form, as in a well-known Egyptian children's song:

> *ḏahaba l-laylu ṭalaʿa l-fagru*
>   *wi-l-ʿaṣfūr ṣawṣaw ṣawṣaw*
> *šāf il-ʾuṭṭa ʾāllaha biss biss*
>   *ʾālitlu nawnaw nawnaw*
>
> 'The night departed, the dawn came, the bird chirped, it saw the cat and said to her "puss puss", and she said to him "meow" '

Since imitations of animal sounds are conventional, they may be expressed differently in languages. The sound of both a rooster and a cuckoo, for instance, is represented in Arabic as *kū kū* (cf. Goscinny and Uderzo, *Minjal* 22, and Hergé, *Sirr* 42, respectively), but in English the sound of the rooster is represented as *cock-a-doodle-doo*, in French as *cocorico*, and in Dutch as *kukeleku*, whereas in all three languages the cuckoo calls out its own name, *cuckoo, coucou*, and *koekoek*, respectively (for the name of the cuckoo in Arabic, ʿInānī [1993:184, n. 1] proposes *kūkū* instead of the traditional name *waqwāq*). Therefore, the imitations are bound to be different across Arabic dialects, and so are the names of birds (for bird names in Syro-Palestinian dialects, see Dalman 1913). They appear, however, to have in common that they do more justice to the actual sound of the animal than the Standard Arabic ones. Thus, for instance, the verb used for 'to mew' in Classical Arabic is *māʾa* or *māʿa* (Ibn Sīdah, *Muxaṣṣaṣ* VIII, 85), which seems to be less mimetic than Egyptian Arabic *nawnaw* (note that both in Classical Arabic and in Modern Standard Arabic *māʾ māʾ* is the sound made by a goat; cf. Zamaxšarī, *Mufaṣṣal* 66.21 and Hergé, *Jazīra* 9).

Not surprisingly, many birds have onomatopoetic names in Arabic, often in the form of reduplicated words of the pattern CvCCvC or CvCCv̄C (cf. ʿInānī 1993:28), e.g. *waqwāq* 'cuckoo', *laqlaq* 'stork', *hudhud* 'hoopoe', *bulbul* 'bulbul', *yuʾyuʾ* 'merlin', *ʿaqʿaq* 'magpie', *qaṭqāṭ* 'plover; lapwing', *wirwār* 'bee-eater', *ʿulʿul* '(short-toed) lark' (with a diminutive *ʿulayʿila* 'lesser short-toed lark'), *siksaka* 'chiff-chaff', *zarzūr* 'starling' (in Syro-Palestinian this means 'cricket'; Procházka 2004:159), *basbūsa* 'serin' (Procházka 2004:156). Note that not all birds with onomatopoetic names have this pattern, e.g. *qaṭā* 'sandgrouse', *sunūnū* 'swallow', *ṭīṭawī* 'redshank', *baqwīqa* 'godwit', and *šaraqraq* 'green woodpecker'; according to some authors, *ʿandalīb* 'nightingale' is onomatopoetic as well. Conversely, not all animal names with the pattern CvCCvC are names of birds, e.g. *judjud* 'cricket', *waṭwaṭ* 'bat', *ṣurṣur* (*ṣarṣūr*) 'cockroach' (Nöldeke 1904). Procházka (2004:159) points out that generally speaking only the reduplicated quadriliteral names are onomatopoetic in nature, whereas other animal names of the pattern *faʿlūl* or *faʿʿūl* are usually related to triradical roots, and not onomatopoetic.

In child language and in baby talk, animal names are often formed from interjections. Woidich (2003:576–577) adduces some examples of this from the dialect of Bašandi in the Daxla oasis in Egypt. In the baby talk register of this dialect, the name for the donkey, *hošša*, and the name of the chicken, *hišša*, are derived with a derivational suffix *-a* from the sound made to halt a donkey, *hošš*, and the one made to chase a chicken away, *hišš*. Other names derive from imitations of the sound of the animal, e.g. *ilʿawa* 'dog' (cf. *yiʿawwēy* 'to bark') and *ilmāʾ* 'goat' (cf. *ymiʾmēʾ* 'to bleat', and *māʾ māʾ* in comics; see above). That this is not limited to baby talk is shown by Syro-Palestinian Arabic *barbūr* 'lamb, sheep', if it is indeed connected with *birbir*, a sound to lure sheep (Procházka 2004:150–151).

Both imitative and synesthetic sound symbolism are often expressed in dialectal Arabic by quadriliteral reduplicated verbs; these represent what Holes (2004:100) calls "mimetic of extended or repeated sounds, movements and actions which occur in the physical world" (→ reduplication). He adds that there is "very often a direct relationship in this verb type between physical phonetic form and semantic function", at least when they are not obvious denominative derivations. Many of the examples quoted by Holes from Eastern Arabian dialects are reduplicated verbs that indicate intensity or repetition of the action, e.g. *dagg* 'to knock' vs. *dagdag* 'to knock repeatedly'; some of them, however, seem to be used to imitate sounds without emphasis, e.g. *xaḍxaḍ* 'to rattle', *xašxaš* 'to jangle', *wašwaš* 'to whisper', *tnahnah* 'to sob' (Rieschild 2006:5). The reduplicated verbs may be extended by a liquid, e.g. *barbag* 'to blow, make bubbles in water', *xarxaš* 'to jingle, rustle'; or by a nasal, e.g. *bambaʿ* 'to bleat [sheep], stammer [with fright, people]'.

Quadriliteral reduplicated verbs are not new in the Arabic lexicon. Classical Arabic lexicographers already drew attention to such verbs. According to aṯ-Ṯaʿālibī (*Fiqh* 209), they are used to indicate the intensity of a sound: the sound of a camel, for instance, is represented by the verb *kaššā*, but if the camel is very loud, the sound is called *kaškaša* (or *qašqaša*). In his analysis of aṯ-Ṯaʿālibī's text, Shivtiel (2000:141) provides numerous examples of such verbs, both derived (e.g. *kaḥḥa/kaḥkaḥa* 'to cough', *xarra/xarxara* 'to snore') and underived (e.g. *wašwaša* 'to whisper', *ṭamṭama* 'to stutter'). Of special interest are those cases where the simple reduplicated verb does not seem to be onomatopoetic (e.g. *hamma* 'to distress', *šaqqa* 'to split'), whereas the derived quadriliteral does (e.g. *hamhama* 'to hem, mutter, growl', *šaqšaqa* 'to twitter, chirp').

In his analysis of reduplicated verbs in Arabic, Procházka (1993) shows that about three-quarters of these verbs belong to only two categories: they denote either intensive or rhythmic motion, e.g. *qasqasa* 'to hurry', *tazaʿzaʿa* 'to wobble', or they refer to acoustic phenomena, e.g. *taʿtaʿa* 'to stammer', *faʿfaʿa* 'to stutter', *xarxara* 'to snore'. In his view, this demonstrates their onomatopoetic origin, because these two semantic fields are precisely those which are the main function of onomatopoeia. Accordingly, when there are word pairs like *takka* and *taktaka*, both meaning 'to tick', Procházka (1993:102) believes that the shorter form was derived from the longer form, rather than the other way around.

In some dialects, special procedures exist to derive and develop onomatopoetic forms. Prunet a.o. (2000:630) point out that onomatopoetic forms are often created by the same procedures as those used in language games and in slips of the tongue, for instance → metathesis or the insertion of sonorants, quoting several examples from Lmnabha Arabic (Elmedlaoui 1995:57), which have now become current in Moroccan Arabic in general, e.g. *ttaq*, reinforced form *ttraq*, superreinforced form *trtllaq*, with the derived verbs *tqtəq* and *trtəq; bbaʿʿ*, reinforced form *bbraʿʿ*, super-reinforced form *brbllaʿʿ*, derived verbs *bʿbəʿ* and *brbəʿ*, indicating various degrees of intensity in producing certain sounds, voices, etc.

4. CONVENTIONAL SOUND SYMBOLISM

Sound symbolism in the synesthetic and the conventional senses (categories iii and iv) was not a frequently studied topic in the Arabic grammatical tradition. There was a certain connection with theories about the origin of speech (→ *waḍʿ al-luġa*). Since the discussion of the origin of speech never gained much popularity, the topic in general was usually dealt with only for the sake of curiosity. One of the authors mentioned in this respect is ʿAbbād ibn Sulaymān (d. 844/230), a Muʿtazilite, who maintained that there was a natural relationship (*munāsaba ṭabīʿiyya*) between sound and meaning, but no details about exactly how he viewed this relationship are given (Weiss 1974:35). It is unknown whether ʿAbbād's theory was in any way related to the speculations about the nature of language of the alchemist and philosopher Jābir ibn Ḥayyān (2nd/8th century), who maintained that the relationship between sounds and their referents was a necessary one: he believed that the nature of the objects in the world is expressed naturally by the words referring to them (Kraus 1942:239–262). By analyzing the words, the philosopher can go back to the original word (*ʾaṣl*), which expresses the true nature of the object.

One of the few grammarians who professed an interest in (conventional) sound symbolism is Ibn Jinnī (d. 392/1002). His *Xaṣāʾiṣ* contains three chapters about this topic (cf. Mehiri 1973:252–259): "Chapter on sounds that follow the meaning" (*Bāb fī taṣāqub al-ʾalfāḏ li-taṣāqub al-maʿānī; Xaṣāʾiṣ* II, 145); "Chapter on the strength of the sound as corresponding to the strength of the meaning" (*Bāb fī quwwat al-lafḏ li-quwwat al-maʿnā; Xaṣāʾiṣ* III, 264); and "Chapter on sounds that imitate the meaning" (*Bāb fī ʾimsās al-ʾalfāḏ ʾašbāh al-maʿānī; Xaṣāʾiṣ* II, 152). According to Ibn Jinnī, Arabic often uses the sounds of the word to represent the event the word stands for, for instance in the case of the two verbs *xaḍama* and *qaḍama*, which both mean 'to chew something', but the former is used for eating fresh vegetables, while the latter is used for crunching something hard and crispy (*Xaṣāʾiṣ* II, 157.13–158.2), or in the case of the words *ʾasaf* 'despair' and *ʿasaf* 'violence', where the sound ʿ in the latter symbolizes the force which is expressed by the action the word stands for (*Xaṣāʾiṣ* II, 146.15–17). The sound symbolism is contained here in the sounds *q/x* and *ʾ/ʿ*, respectively. In morphology, a similar principle of iconicity is at work when the reduplication of the second radical in Form II of the verb signals the repetition of the action expressed by the verb (*Xaṣāʾiṣ* II, 155.3ff.).

Ibn Jinnī's ideas about sound symbolism were combined with his ideas about etymology (*ištiqāq kabīr*), in a grand theory about the nature of the Arabic lexicon, according to which all permutations of radicals express the same semantic primitive, and each radical in these combinations contributes to a part of its semantic load. The semantic symbolism of the sounds may also be expressed by combinations of sounds, for instance when Ibn Jinnī states that the combination of the sound *f* with any of *d, t, ṯ, r, l, n* symbolizes a weakness or a lack of something (*Xaṣāʾiṣ* II, 166.3), or when he says that the combination *s-l-m* in any permutation expresses the notion of 'gentleness'. Later grammarians sometimes quote Ibn Jinnī, but without much enthusiasm. It seems that in general most scholars were convinced of the special characteristics of certain 'letters', but apparently, they did not wish to go as far as Ibn Jinnī did in his etymological ideas.

Because of its role in literature, one might have expected that rhetoricians would be particularly interested in sound symbolism of this category. However, even if they remarked on such phenomena, they did not deal with them systematically. In his *Miftāḥ al-ʿulūm*, as-Sakkākī (d. 626/1229) limits himself to a brief remark about the special character of individual consonants, giving examples such as the difference between *faṣama* 'to split, crack' and *qaṣama* 'to break, shatter', which he explains with the character of its first consonants: in the former, the cracking is slight and hardly visible, whereas in the latter it is clearly visible; likewise, in *ṯulm* 'nick, notch' the *m* indicates the fact that it is just a slight crack in a wall, whereas in *ṯulb* 'slander, defamation', the *b* indicates a major mistake pointed out by someone in criticizing someone else. Along the same lines, the *f* in *zafīr* represents the sound of the donkey, and the glottal stop in *zaʾīr* represents the sound of the lion (*Miftāḥ* 357).

Sound effects and the use of interjections were especially popular in *rajaz* poetry, in which all kinds of phonetic imagery are used to enhance the effect of the poem. Manfred Ullmann (1966:209–213) gives many examples of the use of interjections to achieve this effect. Actual interjections are used, for instance in the following lines from a *rajaz* poem by Dalam ʾAbū Zuʿayb al-ʿAbšamī (M. Ullmann 1966:211).

> *ʾanūʾu li-l-qiyāmi ʾāhan ʾāyah*
> *ʾamšī ruwaydan tāha tāha tāyah*
> 'I have trouble getting up, ow and ouch,
> and I walk slowly, thump, thump, thump'

In other poems, interjections are used in nominalized form, for instance in the following verse by al-ʿAjjāj (M. Ullmann 1966:212), in which the sound *kx*, expressing disgust, is used as a substantive:

> *wa-ṣāra waṣlu l-ġāniyāti kixxan*
> ‘and [when] meetings with beautiful women [end up in their saying] tsk’

Onomatopoetic effects are also achieved by the juxtaposition of words, so for instance in a poem by an anonymous poet describing a garrulous woman (M. Ullmann 1966:208):

> *ḥattā taniqqī ka-naqīqi d-dīki*
> ‘until you cackle like a rooster’s crowing’

The verbs *naqqa* and *naqnaqa* usually indicate the cackling of a hen after laying an egg, as well as sounds made by other birds, but here, *naqqa* is used for the crowing of a rooster.

There are a few studies of the use of sound symbolism in modern literature, for instance by Shunnaq and Al-Thebyan (2003), who studied the English translation of one of Edwār al-Xarrāt’s novels, *Turābuhā zaʿfarān*. They emphasize the role of what they call ‘phonological word-strings’ in these novels, which pose a special challenge to the translator. In such strings, the semantic relations between the words are strengthened by the sounds, for instance in the following fragment, which Shunnaq and Al-Thebyan (2003) compare with the existing translations.

> *lam tabqa fīhā ʾillā ʾiṯāra ḥayāt tuġamġimu bihā wa-tuzaqziqu wa-tuwaḥwiḥu bi-luġa lā naʿrifuhā*
> ‘[faces] in which only a shadow of life remains, with which they mumble, cheep, and whine in a language we do not recognize’

At an even higher text level, Fatani (2005:177) claims that “the articulatory structure of key content words [in sura 113 *al-Falaq*] directly enacts the kinetic movements of the ‘splitting’ process referred to in the text”. In his view, the patterning of fricatives and plosives, in particular in the end rhyme in this sura, correlates with the semantic process of ‘splitting’, which he regards as the key issue of the sura.

Modern studies about conventional sound symbolism often deal with this phenomenon in the context of general Semitic linguistics. They usually take as their point of departure that Semitic languages have semantically associated series of roots sharing the first one or two radicals. These series have played an important role in the debate about → biradicalism because they were analyzed as originally biradical roots with an affix. The most famous example of these is the series ‘to cut’, whose members share the radical *q* (e.g. *qaṣṣa* ‘to cut, clip, shear off’, *qasama* ‘to divide, cut’, *qaṭaʿa* ‘to cut’, *qatala* ‘to kill’, *qaṣaba* ‘to cut up, carve up’, *qaraḍa* ‘to cut, sever’, etc.; Fox 1982:56); this series is connected with other series in which the shared radical is *k, j, ḥ*, or *x* (e.g. *jazara* ‘to slaughter’, *jadaʿa* ‘to cut off, amputate’, *xazza* ‘to pierce’, *xazala* ‘to cut off, sever’, etc.; Fox 1982:57). One explanation of this phenomenon is root-fusion, whereby two separate roots were blended into one. Fox (1982) proposes, instead, that the series originated through sound-symbolic contamination through a similar mechanism as in English *slide, slip, slither, slime, slush*, etc. (cf. Samuels 1972:46–47). According to Fox, the combinations *q-ṣ*, *j-z*, etc. have a similar origin and were never independent roots. A similar proposal, based on the function of initial consonants, has been made more recently by bar-Lev (2005–2006).

Bohas makes a proposal that reaches even further. In a series of studies (Bohas 1997, 2000; Bohas and Dat 2007), he developed a theory about the relationship between the phonetic structure of the Arabic lexicon and its semantic referents (→ lexicon: matrix and etymon theory). In Bohas’ view, both the arbitrary relationship between sound and meaning and the linear nature of sound combinations should be called into question. The matrices he reconstructs for the Arabic lexicon consist of formal phonetic invariants defined by phonetic features, rather than phonemes, and they stand for a semantic core that persists in all resulting roots. According to Bohas, there is a mimophonic relationship between the matrices and the world, i.e., the sounds express in some way the exterior world, so that for instance the presence of pharyngeals in the matrix both articulatorily and semantically express the notion of ‘constriction’. Presumably, such proposals fall under category (iv) mentioned above, but they remain hypotheses that have not yet found general acceptance.

Gordon and Heath (1998) give a biological dimension to the study of sound symbolism; they claim that at least some of the sound-symbolical correlations existing in languages are related to sex differences. The contrast between /i/ vowels and /o, u/ vowels, for instance, seems to be associated by many speakers with ‘smallness’ and ‘bigness’. According to Gordon and Heath (1998:436–438), some of the data reported in gender-related variationist studies in Arabic may be explained by the contrast between female and male speech, for instance the more pronounced tendency in male speech toward pharyngealization of vowels adjacent to emphatic consonants. In their view, this finding and the concomitant evaluation of lower pharyngealization as 'effeminate' (Kahn 1975:41) are consistent with the idea that women are more attracted to high front unrounded vowels, while men typically prefer back vowels (Gordon and Heath 1998:423).

Bibliographical references

Primary sources

Goscinny and Uderzo, *Minjal* = René Goscinny and Albert Uderzo, *ʾAstirīks wa-l-minjal aḏ-ḏahabī*. Cairo: Dār al-Maʿārif, 1980. (Trans. from the French *La serpe d'or*. Paris: Dargaud, 1962.)

Hergé, *Jazīra* = Hergé [= George Rémi], *Tān Tān fī l-jazīra as-sawdāʾ*. Cairo: Dār al-Maʿārif, 1989. (Trans. from the French *L'Île Noir*. Tournai: Casterman, 1945, 1979.)

——, *Sirr* = Hergé [= George Rémi], *Sirr al-xartīt*. Cairo: Dār al-Maʿārif, 1989. (Trans. from the French *Le secret de la Licorne*. Tournai: Casterman, 1945, 1979.)

Ibn Jinnī, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. Ed. Muḥammad ʿAlī an-Najjār. Cairo, 1952–1956. (Repr., Beirut: Dār al-Hudā, n.d.)

Ibn Sīdah, *Muxaṣṣaṣ* = ʾAbū l-Ḥasan ʿAlī ibn ʾIsmāʿīl Ibn Sīdah al-ʾAndalusī, *al-Muxaṣṣaṣ*. 5 vols. Beirut: Dār al-ʾĀfāq al-Jadīda, n.d.

Mubarrad, *Muqtaḍab* = ʾAbū l-ʿAbbās Muḥammad ibn Yazīd al-Mubarrad, *al-Muqtaḍab*. Ed. Muḥammad ʿAbd al-Xāliq ʿUḍayma. 4 vols. Cairo: Dār at-Taḥrīr, 1965–1968.

Sakkākī, *Miftāḥ* = Sirāj ad-Dīn ʾAbū Yaʿqūb Yūsuf ibn ʾAbī Bakr as-Sakkākī, *Miftāḥ al-ʿulūm*. Ed. Nuʿaym Zarzūr. Beirut: Dār al-Kutub al-ʿIlmiyya, 1983.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. 2 vols. Ed. Bulaq, 1316 A.H. (Repr., Baghdad: Al-Muthanna Library, n.d.)

Suyūṭī, *ʾAšbāh* = Jalāl ad-Dīn ʾAbū l-Faḍl ʿAbd ar-Raḥmān ibn ʾAbī Bakr as-Suyūṭī, *al-ʾAšbāh wa-n-naḏāʾir fī n-naḥw*. 2nd ed. 4 vols. Hyderabad, 1359–1361 A.H.

Ṯaʿālibī, *Fiqh* = ʾAbū Manṣūr ʿAbd al-Malik ibn Muḥammad ibn ʾIsmāʿīl aṯ-Ṯaʿālibī, *Fiqh al-luġa wa-sirr al-ʿarabiyya*. Beirut: Dār al-Kutub al-ʿIlmiyya, n.d.

Xalīl, *ʿAyn* = ʾAbū ʿAbd ar-Raḥmān al-Xalīl ibn ʾAḥmad al-Farāhīdī, *Kitāb al-ʿayn*. Ed. Mahdī al-Maxzūmī and ʾIbrāhīm as-Samarrāʾī. 8 vols. Beirut: Muʾassasat al-ʾAʿlamī li-l-Maṭbūʿāt, 1988.

Zajjājī, *ʾĪḍāḥ* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān ibn ʾIsḥāq az-Zajjājī, *al-ʾĪḍāḥ fī ʿilal an-naḥw*. Ed. Māzin al-Mubārak. Cairo: Dār al-ʿUrūba, 1959.

Zamaxšarī, *Mufaṣṣal* = ʾAbū l-Qāsim Maḥmūd ibn ʿUmar az-Zamaxšarī, *Kitāb al-mufaṣṣal fī n-naḥw*. Ed. Jens Peter Broch. Christiania: Libraria P.T. Mallingii, 1870.

Secondary sources

Baalbaki, Ramzi. 2007. "Inside the speaker's mind: Speaker's awareness as arbiter of usage in Arab grammatical theory". *Approaches to Arabic linguistics presented to Kees Versteegh on the occasion of his sixtieth birthday*, ed. Everhard Ditters and Harald Motzki, 3–23. Leiden: E.J. Brill.

Badawi, El-Said and Martin Hinds. 1986. *A dictionary of Egyptian Arabic: Arabic-English*. Beirut: Librairie du Liban.

bar-Lev, Zev. 2005–2006. "Arabic key-consonants". *Journal of Arabic and Islamic Studies* 6.24–63.

Bohas, Georges. 1997. *Matrices, étymons, racines: Eléments d'une théorie lexicologique du vocabulaire arabe*. Paris and Louvain: Peeters.

——. 2000. *Matrices et étymons: Développements de la théorie*. Lausanne: Editions du Zèbre.

—— and Mihai Dat. 2007. *Une théorie de l'organisation du lexique des langues sémitiques: Matrices et étymons*. Paris: ENS Editions.

Dalman, Gustaf. 1913. "Arabische Vogelnamen von Palästina und Syrien". *Zeitschrift des Deutschen Palästina-Vereins* 36.165–179.

Elmedlaoui, Mohamed. 1995. "Géométrie des restrictions de cooccurrence de traits en sémitique et berbère: Synchronie et diachronie". *Canadian Journal of Linguistics* 40.39–76.

Fatani, Afnan H. 2005. "The iconic-cognitive role of fricatives and plosives: A phono-semantic analysis of a Classical Arabic prayer *Al-falaq*". *Outside-in – Inside-out: Iconicity in language and literature*, IV, ed. Costantino Maeder, Olga Fischer, and William J. Herlovsky, 173–193. Amsterdam and Philadelphia: J. Benjamins.

Firanescu, Daniela Rodica. 2003. *Exclamation in Modern Literary Arabic: A pragmatic perspective*. Bucharest: Editura Universităţii din Bucureşti.

Fox, Andrew. 1982. "Sound-symbolism in Semitic". *Cahiers de l'Institut de Linguistique de Louvain* 8.55–61.

Gordon, Matthew and Jeffrey Heath. 1998. "Sex, sound symbolism, and sociolinguistics". *Current Anthropology* 39.421–449.

Haak, Martine, Rudolf de Jong, and Kees Versteegh (eds.). 2004. *Approaches to Arabic dialects: A collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday*. Leiden: E.J. Brill.

Ḥassān, Tammām. 1973. *al-Luġa al-ʿarabiyya: Maʿnāhā wa-mabnāhā*. Cairo.

Hinton, Leanne, Johanna Nichols, and John J. Ohala (eds.). 1994a. *Sound symbolism*. Cambridge: Cambridge University Press.

——, Johanna Nichols, and John J. Ohala. 1994b. "Introduction: Sound symbolic processes". Hinton a.o. (1994a:1–12).

Holes, Clive. 2004. "Quadriliteral verbs in the Arabic dialects of Eastern Arabia". Haak a.o. (2004:97–116).

ʿInānī, Muḥammad Muḥammad. 1993. *Ṭuyūr Miṣr*. Cairo: al-Hayʾa al-Miṣriyya al-ʿĀmma li-l-Kitāb.

Kahn, Margaret. 1975. "Arabic emphatics: The evidence for cultural determinants of phonetic sex-typing". *Phonetica* 31.38–50.

Kraus, Paul. 1942. *Jābir ibn Ḥayyān, contribution à l'histoire des idées scientifiques dans l'Islam: Jâbir et la science grecque*. Cairo. (Repr., Paris: Les Belles Lettres, 1986.)

Levin, Aryeh. 1991. "The category of *ʾasmāʾ al-fiʿl* in Arabic grammar". *Proceedings of the Colloquium on Arabic Grammar*, ed. Kinga Dévényi and Tamás Iványi, 247–256. Budapest: Eötvös Loránd University and Csoma de Kőrös Society.

Mehiri, Abdelkader. 1973. *Les théories grammaticales d'Ibn Jinnī*. Tunis: Université de Tunis.

Nöldeke, Theodor. 1904. "Tiernamen mit Reduplikation". Theodor Nöldeke, *Beiträge zur semitischen Sprachwissenschaft*, 107–123. Strasbourg: K. Trübner.

Owens, Jonathan. 2004. "Remarks on ideophones in Nigerian Arabic". Haak a.o. (2004:207–220).

Procházka, Stephan. 1993. "Some remarks on the semantic function of the reduplicated quadriliteral verb (structure *faʿfaʿa*)". *Proceedings of the Colloquium on Arabic Lexicology and Lexicography, Budapest, 1–7 September 1993*, ed. Kinga Dévényi, Tamás Iványi, and Avihai Shivtiel, I, 97–103. Budapest: Eötvös Loránd University and Csoma de Kőrös Society.

——. 2004. "Lämmer, Geckos, junge Hähne: Tiernamen der Schemata *faʿʿūl* und *faʿlūl* in den syro-palästinensischen Dialekte". *Estudios de Dialectología Norteafricana y Andalusí* 8.149–162.

Prunet, Jean-François, Renée Béland, and Ali Idrissi. 2000. "The mental representation of Semitic words". *Linguistic Inquiry* 31.609–648.

Rieschild, Verna Robertson. 2006. "Emphatic repetition in spoken Arabic". *Proceedings of the 2004 Conference of the Australian Linguistics Society*. Available at: http://hdl.handle.net/2123/274.

Samuels, M.L. 1972. *Linguistic evolution*. Cambridge: Cambridge University Press.

Shivtiel, Avihai. 2000. "Onomatopoeia in Arabic". *Hunter of the East: Arabic and Semitic studies in honour of Edmund Clifford Bosworth*, ed. Ian R. Netton, 138–144. Leiden: E.J. Brill.

Shunnaq, Abdullah and Qusai Al-Thebyan. 2003. "The translation of onomatopoeia in Edwar Al-Kharrat's *Turābuhā zaʿfarān* into English". *Interface* (Brussels) 17:2.127–144. (Also available at: www.arabswata.org/site/researches/87.html.)

Talmon, Rafael. 2003. *Eighth-century Iraqi grammar: A critical exploration of pre-H̲alīlian Arabic linguistics*. Winona Lake, Ind.: Eisenbrauns.

Ullmann, Manfred. 1966. *Untersuchungen zur* Rağaz*poesie: Ein Beitrag zur arabischen Sprach- und Literaturwissenschaft*. Wiesbaden: O. Harrassowitz.

Ullmann, Stephen. 1962. *Semantics: An introduction to the science of meaning*. Oxford: Blackwell.

Venzlaff, Helga. 1994. Al-Hudhud*: Eine Untersuchung zur kulturgeschichtlichen Bedeutung des Wiedehopfs im Islam*. Frankfurt a. Main: P. Lang.

Weiss, Bernard. 1974. "Medieval Muslim discussions of the origin of language". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 124.33–41.

Woidich, Manfred. 2003. "Baby-talk from ilBašandi, Dakhla oasis, Egypt". *AIDA 5th Conference Proceedings, Cádiz, September 2003*, ed. Ignacio Ferrando and Juan José Sánchez Sandoval, 571–581. Cadiz: Servicio de Publicaciones Universidad de Cádiz.

Kees Versteegh (University of Nijmegen)

South Africa

At the end of the 15th century, European seafaring nations established contact with southern Africa. The first to arrive were the Portuguese, but in the 17th century the Dutch established a monopoly on trade in this region when Jan van Riebeeck founded the Cape Colony in 1652. The first Muslims arrived almost simultaneously with the Dutch, because the Dutch East Indies Company (V.O.C.) started to use the colony as an exile place for slaves and political prisoners from the Dutch East Indies and India. These slaves constituted "the embryo of the Cape Muslim community" (Davids 1980:xv). In the literature, they are called Cape Malays, Malay Muslims, or Kaapmaleiers, which is not quite accurate because most of them did not come from Malaysia, but from India and the Indonesian archipelago. A second group was formed by the so-called Mardyckers (< Malay *merdeka* 'freedom'), free servants from the East Indies who came to the Cape voluntarily. Table 1 shows the main countries of origin of the slaves that were brought to the Cape by the Dutch.

Table 1. Countries of origin of slaves in the Cape Colony between 1652 and 1818 (Da Costa and Davids 1994:2)

Country of origin	Percentage
Africa	26.65%
Ceylon	3.1%
India	36.4%
East India	31.47%
Mauritius	0.18%
Malaysia	0.49%
other	0.4%
unidentified	1.31%

Most Muslims arriving from the East Indies used Malay as their lingua franca, which from the 14th century onward had been the language in which Islam was spread in Southeast Asia.

Malay as used by Muslims was written in Arabic script and had a long tradition of incorporating Arabic loanwords (→ Indonesian/Malay). The first generations of Muslims who arrived as slaves in the Cape Colony were forced to adopt the local form of Dutch, known as Afrikaans, as their new lingua franca. This was not the standard variety of Dutch, but rather the form that had developed in the Cape Colony, probably as the result of a process of creolization by the indigenous Khoisan population, although this remains a controversial issue (Valkhoff 1972; Coetzee 2001). The newcomers from the East Indies must have played an important role in the creolization process, because the next generations of Muslims began to speak Afrikaans as their native language (Davids 1994a, b, 1996). It is fairly certain that the newcoming Muslims were the first to write Afrikaans down, using the Arabic script, just as they had done for Malay in Southeast Asia.

Documents in Afrikaans written with Arabic script were probably circulating already at the beginning of the 18th century (for a general survey, see Haron 2001). One of the oldest official publications in Arabic Afrikaans is the *Bayān ad-dīn* by the Ottoman scholar Abu Bakr Effendi, who taught in the Cape Colony; his work was published in Istanbul in 1877 (Selms 1951; Kähler 1971:71–79). His book may have been preceded by the publication of ʾAḥmad al-ʾĪšmūnī's *Kitāb al-qawl al-matīn fī bayān ʾumūr ad-dīn* (Cape Town, 1856; Selms 1953). Most of the documents are of a religious nature, but some texts belong to other genres, such as grammar. Apart from published books, of which Kähler (1971:70–171) lists 51 items, there are some examples of personal letters written in Afrikaans in Arabic script, preserved in the Davids Collection in the National Library in Cape Town. The Afrikaans literature in Arabic script flourished in the period between 1890 and 1918, but publication did not stop after that time. A grammar book, *Nayl al-ʿArab fī luġat al-ʿArab*, by *šayx* ʾIsmāʿīl ibn Muḥammad Ḥanīf al-ʾAzharī, was published as late as 1948 in Cape Town (Ebrahim 2004; Kähler [1971:86, 104] lists other grammatical writings by this author; see Fig. 1). The *madrasa*s in Cape Town used books in Arabic script in their teaching, hence the popularity of these texts. Arabic script continued to be employed in private correspondence among Muslims, but this does not seem to be the case anymore.

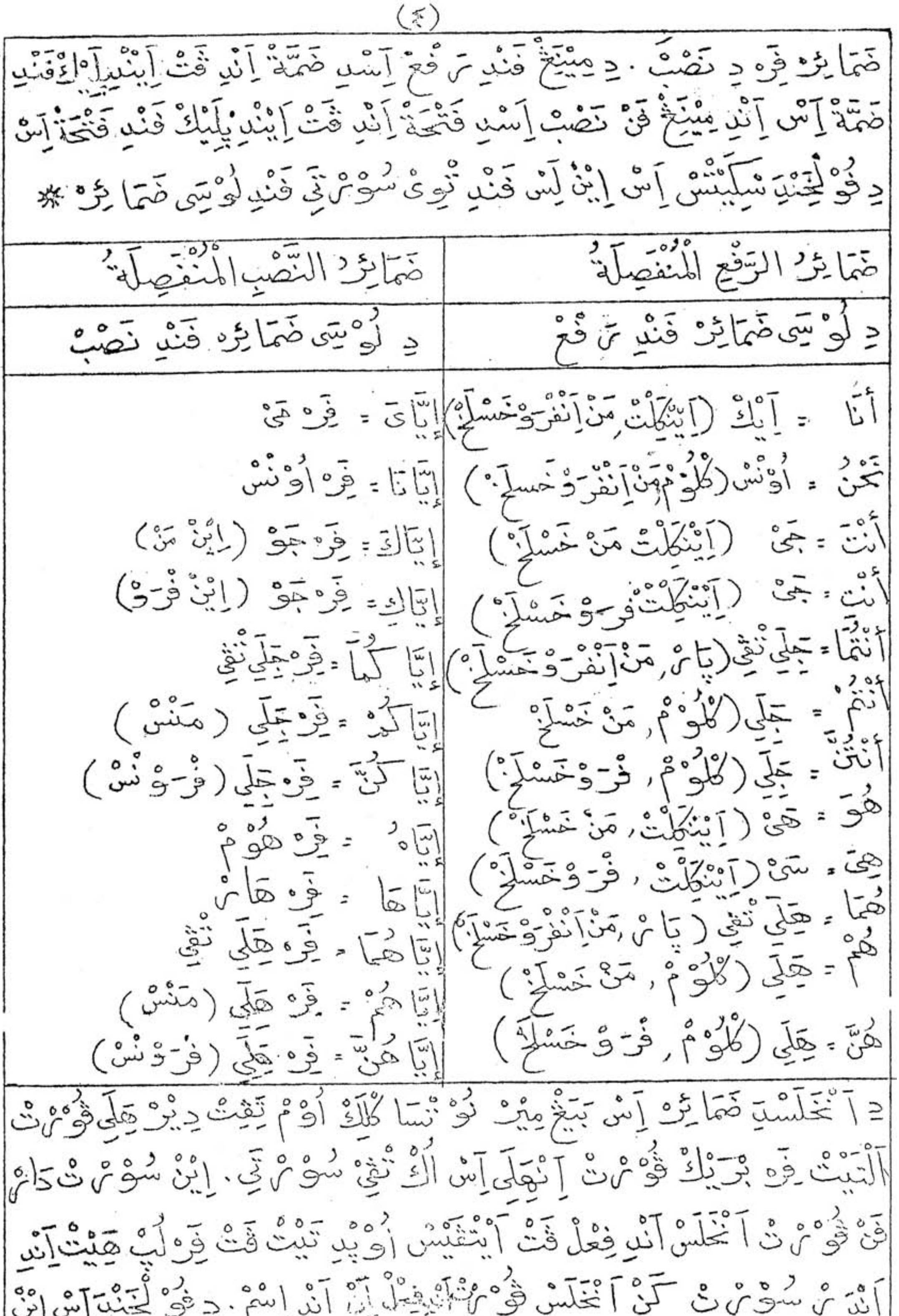
Most of the additional consonants for Afrikaans in Arabic script were borrowed from the orthography of Turkish (→ Arabic alphabet for other languages), rather than from Malay orthography. Afrikaans /p/, for instance, is represented by پ rather than ڤ, which is used in these texts for Afrikaans /w/. The character ف is used for both Afrikaans /f/ and /v/. For /č/, both in English and in Malay words (and in Afrikaans words like *bitjie* 'a bit, a little'), the character چ is used; for /ŋ/, either a combination of ن and ك, or the character ڠ is used.

The representation of the Afrikaans vowels was more complicated, and in this respect the writers of Afrikaans in Arabic script showed their ingenuity by devising entirely new combinations to represent the Afrikaans vowels, in particular by combining two vowel signs on one letter, or by distinguishing between lengthening letters with and without *sukūn*. Table 2 presents some of the combinations that are used in one of the transcription systems, that of the *Nayl al-ʿArab*; other texts use similar combinations, but there is a certain measure of variation in the texts.

Table 2. Representation of Afrikaans vowels in Arabic script (*Nayl al-ʿArab*)

Afrikaans vowel	Arabic script	Example	Gloss
/a/	ـَ	مَنْ	*man* 'man'
/a:/	ـَا	نَام	*naam* 'name'
/ɛ/	ـَِيْ	مَِيْسْ	*mes* 'knife'
/ɛ:/	?	?	?
/e:/	ـِيْ	تْوِىْ	*twee* 'two'
/i:/	ـِيْ	رَفِيْرْ	*rivier* 'river'
/o/	ـُوْ	اُوْنْسْ	*ons* 'us'
/o:/	ـُوْ	ڤُوْرْتْ	*woord* 'word'
/ə/	ـَِ	تَافَِلْ	*tafel* 'table'
/y:/	ـِيْ	اِيْرَِ	*ure* 'hours'
/eu/	ـِيْ	دِيْرْ	*deur* 'door'
/ei/ (= *ei*, *y*)	ـَيْ	سْكْرَيْفْ	*skryf* 'write'
/ou/	ـَو	نَوْ	*nou* 'now'
/ui/	ـَيْ	بَيْكْ	*buik* 'belly'

Figure 1. A page from *Nailoe ʾarabic fee loeghatiel ʾarabie* (Cape Town, ca. 1948)



(٣)

ضَمَائِرْ فِرْ دِ نَصْبْ . دِ مِيْنِيْغْ فَنْدِ رَفْعْ اَسْ ضَمَّةْ اَنْدِ ڤَتْ اِيْنِيْلْ آهْ اِءْ فَنْدِ
ضَمَّةْ اَسْ اَنْدِ مِيْنِيْغْ فَنْ نَصْبْ اِسْ فَتْحَةْ اَنْدِ ڤَتْ اِيْنْدِيْلِيْكْ فَنْدِ فَتْحَةْ اَسْ
دِ فُوْلْخَنْدِ نَسْكِيْبِيْشْ اَسْ اِيْنْ لَسْ فَنْدِ تْوِيْ سُوْرْتِيْ فَنْدِ لُوْسِيْ ضَمَائِرْ ※

ضَمَائِرُ الرَّفْعِ الْمُنْفَصِلَةُ	ضَمَائِرُ النَّصْبِ الْمُنْفَصِلَةُ
دِ لُوْسِيْ ضَمَائِرْ فَنْدِ رَفْعْ	دِ لُوْسِيْ ضَمَائِرْ فَنْدِ نَصْبْ
أَنَا = اَيْكْ (اَيْنْكَلْتْ، مَنْ اَنْفِرْ وْ خَسْلَخْ)	إِيَّايَ = فِرْ مِيْ
نَحْنُ = اُوْنْسْ (كُلُوْمْ مَنْ اَنْفِرْ وْ خَسْلَخْ)	إِيَّانَا = فِرْ اُوْنْسْ
أَنْتَ = جِيْ (اَيْنْكَلْتْ مَنْ خَسْلَخْ)	إِيَّاكَ = فِرْ جُوْ (اِيْنْ مَنْ)
أَنْتِ = جِيْ (اَيْنْكَلْتْ فْرَوْ خَسْلَخْ)	إِيَّاكِ = فِرْ جُوْ (اِيْنْ فْرَوْ)
أَنْتُمَا = جُلِيْ تْوِيْ (پَارْ، مَنْ اَنْفِرْ وْ خَسْلَخْ)	إِيَّاكُمَا = فِرْ جُلِيْ تْوِيْ
أَنْتُمْ = جُلِيْ (كُلُوْمْ، مَنْ خَسْلَخْ)	إِيَّاكُمْ = فِرْ جُلِيْ (مَنْسْ)
أَنْتُنَّ = جُلِيْ (كُلُوْمْ، فْرَوْ خَسْلَخْ)	إِيَّاكُنَّ = فِرْ جُلِيْ (فْرَوْنْسْ)
هُوَ = هِيْ (اَيْنْكَلْتْ، مَنْ خَسْلَخْ)	إِيَّاهُ = فِرْ هُوْمْ
هِيَ = سِيْ (اَيْنْكَلْتْ، فْرَوْ خَسْلَخْ)	إِيَّاهَا = فِرْ هَارْ
هُمَا = هُلِيْ تْوِيْ (پَارْ، مَنْ اَنْفِرْ وْ خَسْلَخْ)	إِيَّاهُمَا = فِرْ هُلِيْ تْوِيْ
هُمْ = هُلِيْ (كُلُوْمْ، مَنْ خَسْلَخْ)	إِيَّاهُمْ = فِرْ هُلِيْ (مَنْسْ)
هُنَّ = هُلِيْ (كُلُوْمْ، فْرَوْ خَسْلَخْ)	إِيَّاهُنَّ = فِرْ هُلِيْ (فْرَوْنْسْ)

دَ اَخْتَلَسْدِ ضَمَائِرْ اَسْ بَيْغْ مِيْرْ نُوْ نَسَا كُلَكْ اُوْمْ تِيْفَتْ دِيْرْ هُلِيْ ڤُوْرْتْ
اَلْتِيْتْ فِرْ بْرِيْكْ ڤُوْرْتْ اِنْهِلِيْ اَسْ اَكْ تْوِيْ سُوْرْتِيْ. اِيْنْ سُوْرْتْ دَارْ
فَنْ ڤُوْرْتْ اَخْتَلَسْ اَنْدِ فِعْلْ ڤَتْ اَيْتْغَيْسْ اُوْرِيْدْ تَيْتْ ڤَتْ فِرْ لُوْپْ هَيْتْ اَنْدِ
اَنْدِرْ سُوْرْتْ كَنْ اَخْتَلَسْ ڤُوْرْتْ اِيْفِعْلْ اِيْ اَنْدِ اِسْمْ . دِ فُوْلْخَنْدِ اَسْ اِيْنْ

Most of the Arabic words in the Afrikaans texts belong to the realm of religion (or, in the case of the grammar book mentioned above, they are grammatical technical terms). The most striking aspect of these Arabic words is that they are fully integrated in the language. Arabic words are used with the Afrikaans article *die* and regularly receive Afrikaans plurals in *-s* or *-e*, for instance *kitābs* 'books', *ʿālims* 'scholars', *rasūls* 'envoys', but also *kitāpe* and *jine* 'jinns' (Hoedemaekers 2006:63). Arabic words are also used in compounds, e.g. *qiyāma dāg* 'Day of Judgment' (< Arabic *qiyāma* + Afrikaans *dag* 'day'), or in the curious example *naḥw kitāpe* 'grammar books' (Hoedemaekers 2006:75).

In the case of verbs, there is a tendency to use compounds with a dummy verb *māk* 'to make, to do' (→ Persian), e.g. *māk ʿibāda* 'to serve', *māk sujūd* 'to kneel in prayer', *ṣalā māk* 'to pray' (Hoedemaekers 2006:64). These seem to be independent innovations because in Malay, Arabic verbs are borrowed directly. There are some examples, however, of Arabic nouns being used as verbs without the help of a dummy verb, e.g. *om teṣalāt* 'in order to pray' (< Arabic *ṣalāt* 'prayer'). Interestingly, these are paralleled by the use of Malay verbs in Afrikaans used in the same way, e.g. *gebacha* 'read [past participle]' and *tebacha* 'to read' (< Malay *(mem)baca* 'to read'; Kähler 1971:47–64; Hoedemaekers 2006:80). Later texts also contain a number of English loanwords (Hoedemaekers 2006:66), as in (1), from the *Nayl al-ʿArab* (1.9–10).

(1) *det esnodag omte witdie*
it is-necessary to know.the
defrent patrone fan eilke ḥarf
different patterns of each letter
'It is necessary to know the different patterns of each letter'

The presumably written transmission of the Arabic loanwords, often in texts that were translated directly from Arabic and printed together with the Arabic original, has led to a high degree of syntactic interference. This is quite similar to the use of Arabic loanwords in Malay (→ Indonesian/Malay); beyond the domain of Arabic, it may be compared to the syntactic and stylistic interference of Latin in all European languages on the basis of the translations from Latin, in particular the Bible. In some cases, this interference looks like typical translation errors, for instance when the Afrikaans copula is left out in phrases like (2) and (3),

(2) *die galūk op-hele tiwie*
the happiness on-them two
'Peace be upon both of them' (Arabic *ʿalayhimā s-salām*; Abu Bakr Ibn al-Fāḍil, *Tanbīhāt* 28.10)

(3) *wie māk ḥarām fit fer jahanam*
who makes forbidden fit for hell
'Who does forbidden things is destined for hell' (from a religious document printed in 1911; Hoedemaekers 2006:81)

or when the referential pronoun in the relative clause is faithfully rendered in the Afrikaans text, as in (4):

(4) *die weig fan die gīnege wat u die*
the way of the ones who you the
hidāya op hele fergent het
guidance on them provided have
'The way of those you have guided' (Arabic *ṣirāṭ allaḏīna ʾanʿamta ʿalayhim*; (Abu Bakr Ibn al-Fāḍil, *Tanbīhāt* 43; Hoedemaekers 2006:53)

Likewise, the verb *seig* (Afrikaans *sê*) 'to say' is construed with a preposition *fer* (Afrikaans *vir* 'for') in a phrase like (5), which also illustrates the word order VSO that mirrors the Arabic word order:

(5) *seig die nabī fer jibrīl*
say the prophet for Jibrīl
'The Prophet said to Jibrīl' (Arabic *qāla n-nabī li-jibrīl*; Abu Bakr Ibn al-Fāḍil, *Tanbīhāt* 44; Hoedemaekers 2006:53)

The vocative in these texts is sometimes followed by the article, just like Arabic *yā ʾayyuhā*: *oodie broofit* 'oh, Prophet!' (Afrikaans *oh profeet*). A semantic misnomer occurs when the Arabic word *ʿayn* 'eye; source' is translated with *oog water* 'eye water'.

In other cases, the pervasive influence of Arabic may have led to the development of new syntagms even in Afrikaans spoken by Muslims, but this influence has not yet been researched systematically. A possible example is the use of *op hole es* 'they must' lit. 'on them is' as a calque of the Arabic syntagm *kāna ʿalayhim* (Kähler 1971:67). The syntactic interference is matched with semantic interference in the form of calques. One of the most frequent examples is that of the Afrikaans word *sklaaf* 'slave', which is used for 'servant of God' on the basis of Arabic *ʿabd*.

Not much is known about the present-day use of Arabic loanwords in spoken Afrikaans by Muslims (on this variety of Afrikaans, see Kotzé 1983). Muslim Cape Afrikaans remains distinct from other varieties of Afrikaans, not only by the presence of Malay and Arabic loanwords but also by syntactic peculiarities (Van Rensburg 1989). Moosa (2004) states that most *ʿulamāʾ* have switched to Roman script when writing Afrikaans, but according to Tayob (p.c.), Muslim preachers still tend to use many Arabic words in their sermons (on the present-day situation of the Muslim community, see Tayob 1999), and there are some indications that even in ordinary speech Arabic loanwords continue to be used.

Bibliographical references

Primary sources

Abu Bakr Effendi, *Bayān ad-Dīn*, ed. Adrianus van Selms, *Abu Bakr se* 'Uiteensetting van die godsdiens'. (= *Verhandelingen der Koninklijke Nederlandse Akademie van Wetenschappen, Afdeling Letterkunde*, NS CX.) Amsterdam: Koninklijke Nederlandse Akademie van Wetenschappen, 1979. [English trans., *The religious duties of Islam as taught and explained by Abu Bakr Effendi: A translation from the original Arabic and Afrikaans*, trans. Mia Brandel-Syrier. Leiden: E.J. Brill, 1960.]

Abu Bakr Ibn al-Fāḍil, *Tanbīhāt* = Abu Bakr Ibn al-Fāḍil, *at-Tanbīhāt al-lāʾiqa bi-ʾaḥwāl al-muslimīn*. [Cairo], 1330/1911. [Kähler 1971:154.]

Nayl al-ʿArab = Ganief, Ismail. *Nailoel ʾarabie fee loeqhatiel ʾarabie: Handboek van Arabies en Arabiese grammatika*. Cape Town, ca. 1948.

Secondary sources

Coetzee, Anna. 2001. "Nederlands se Afrikaanse kreool-kinders" [The African creole children of Dutch]. Available at www.ned.univie.ac.at/CMS/Afrikaans/Grammatika/Bydraes/Anna_Coetzee_Nederlands_se_Afrikaanse_Kreool_kinders/.

Da Costa, Yusuf and Achmat Davids. 1994. *Pages from Cape Muslim history*. Pietermaritzburg: Shuter and Shooter.

Davids, Achmat. 1980. *The mosques of Bo Kaap: A social history of Islam at the Cape*. Cape Town: Cape and Transvaal Printers.

——. 1990. "Words the Cape slaves made: A socio-historical linguistic study". *South African Journal of Linguistics* 8:1.1–24.

——. 1991. "Abubakr Effendi: His creation of the Afrikaans letter *e* in Arabic script". *South African Journal of Linguistics* 9:1.1–17.

——. 1992. *Die Afrikaans van die Kaapse Moslems, 1815–1915: 'n Sociolinguistiese studie* [The Afrikaans of the Cape Muslims: A sociolinguistic study]. M.A. thesis, University of Natal.

——. 1993. "The early Afrikaans publications and manuscripts in the Arabic script". *Festschrift in honour of F.R. Bradlow*, ed. P. Wiestra and B. Warner, 67–82. Cape Town: Friends of the South African Library.

——. 1994a. "Afrikaans die produk van akkulturasie" [Afrikaans, the product of acculturation]. *Nuwe perspektiewe op die geskiedenis van Afrikaans: Opgedra aan Edith H. Raidt*, ed. Anna Coetzee and Gerrit Olivier, 110–119. Halfweghuis: Southern Boekuitgewers.

——. 1994b. "The contribution of the slaves to the genesis of Afrikaans". *Taal en identiteit*, ed. Vernon February, 39–53. Cape Town Bpk.: Tafelberg Uitgewers.

——. 1996. "Laying the lie of the 'Boer' language: An alternative view of the genesis of Afrikaans". *Matatu: Zeitschrift für Afrikanische Kultur und Gesellschaft* 15–16.13–58.

Ebrahim, Mogamat Hoosain. 2004. *Shaykh Ismail Hanif Edwards: His life and works*. Paarl: Paarl Print.

Haron, Muhammad. 1997. *Muslims in South Africa: An annotated bibliography*. Cape Town: South African Library.

——. 2001. "The making, preservation and study of South-African *ajami* mss and texts". *Sudanic Africa Volume* 12.1–14. [Also available at http://hf.uib.no/smi/sa/12/12Ajami.pdf.]

Hoedemaekers, Iris. 2006. *'Die kitāb wat sal prāt op die rules fan die naḥw...': Het Arabisch-Afrikaans*. M.A. thesis, University of Nijmegen.

Jeppie, Shamil. 2006. *Language, identity, modernity: The Arabic Study Circle of Durban*. Cape Town: Human Sciences Research Council Press.

Kähler, Hans. 1971. *Studien über die Kultur, die Sprache und die arabisch-afrikaanse Literatur der Kap-Maleien*. Berlin: Reimer.

Kotzé, Ernst F. 1983. *Variasiepatrone in Maleier-Afrikaans* [Patterns of variation in Malay Afrikaans]. Ph.D. diss., University of Witwatersrand.

Moosa, Ebrahim. 2004. "Afrikaans in Arabic script". *Encyclopaedia of Islam*, XII Supplement, 754. 2nd ed. Leiden: E.J. Brill.

Selms, Adrianus van. 1951. *Een tweetalige (Arabiese en Afrikaanse) Kategismus* [A bilingual Arabic and Afrikaans catechism]. Amsterdam: Noord-Hollandsche Uitgevers Maatschappij. [On Ahmed Behardien's *Kitāb at-tawḥīd*.]

——. 1953. "Die oudste boek in Afrikaans: Isjmoeni se '*Betroubare woord*'" [The oldest book in Afrikaans: Isjmoeni's *Reliable word*]. *Hertzog Annale van die Suid-Afrikaanse Akademie vir Wetenskap en Kuns* 2.61–103.

Tayob, Abdelkader. 1999. *Islam in South Africa: Mosques, imams, and sermons*. Gainesville: University Press of Florida.

Valkhoff, Marius F. 1972. *New light on Afrikaans and 'Malayo-Portuguese'*. Louvain: Peeters.

Van Rensburg, M.C.J. 1989. "Soorte van Afrikaans [Varieties of Afrikaans]". *Inleiding tot die Afrikaanse taalkunde*, ed. Theunis J.R. Botha, 436–467. Pretoria: Academica.

Iris Hoedemaekers and Kees Versteegh
(University of Nijmegen)

South Arabian Loanwords

Pre-Islamic Arabia was far from immune from cultural and linguistic influences exerted by two prestigious cultural centers, namely the Fertile Crescent in the north and the Yemen in the south (Beeston 1981:180). The material civilization of the Yemenite kingdoms of Sabaʾ, Maʿīn, Qaṭabān, and Ḥaḍramawt, renowned for their achievements in agriculture, administration, architecture, and commerce, undoubtedly attracted the Arabs' attention. Robin (1991–1993:72, 77) reports that, as early as the 6th century B.C.E., Arab nomads are mentioned in a Minean inscription, and that by the 2nd century B.C.E. a number of Arabs had settled in the Yemenite heartland. Moreover, for centuries before the beginning of the Common Era, the incense trade had brought South Arabian caravaneers traveling across Arabia in direct contact with the Arabs (Rabin 1984:126). Conflict between the Sabeans and the Arabs reached its climax in the mid-2nd century C.E., but shortly after, during the 3rd century C.E., the fighting skills of Arab Bedouin were sought by various kingdoms and their allies. Such recourse to Arab auxiliaries greatly enhanced the Arab penetration in the Yemen and integration into South Arabian society. This is confirmed by a number of hybrid inscriptions composed in Sabean mixed with elements of Arabic. After the expulsion of the Abyssinians, toward the end of the 3rd century, from western Yemen, the Himyarites and their Kinda (Arab) auxiliaries ruled supreme. Beginning in 525 C.E., vast areas of South Arabia again became an Abyssinian protectorate, for about fifty years. By this time, the South Arabian kingdoms were in an advanced state of decline (Rippin 1990:155–156), and a number of tribes migrated toward the center of the Arabian Peninsula. By the end of the 6th century, most of eastern Yemen was in the hands of Arab nomads, and during the first years of the Islamic era, the Hamdān confederation represented what remained of the ancient South Arabian tribes.

The adoption by the Arabs of features of South Arabian cultural models inevitably resulted in borrowing the terminology related to these models. Ancient (or Epigraphic) South Arabian, attested as early as the beginning of the 1st millennium B.C.E., comprised the following closely related Semitic languages: Sabean, Minean, Qatabanian, and Hadramitic (Kogan a.o. 1997:220–221). Given that South Arabian and Arabic are genetically related, sharing many inherited linguistic features, it is an arduous task to extricate loanwords from the cognate material. The criteria outlined by Militarev and Kogan (2000:l–lvii) can surely help one to avoid pitfalls. Medieval Muslim philologers and exegetes engaged in unending polemics concerning the presence of foreign elements in the *Qurʾān* in particular, and in Arabic in general (Kopf 1956:40–45; az-Zubaydī 1994). Schall (1982:144–146) reports that some loanwords which had penetrated the Arabic lexicon during pre-Islamic times did so only transitorily and were soon forgotten, causing much confusion in medieval Muslim philological circles. Baalbaki (1983:124) acknowledges these scholars' contribution to the comparative study of Arabic. Nevertheless, they lacked knowledge about South Arabian, to such an extent that even authors like al-Hamdānī (d. 334/946) and Ibn an-Nadīm (d. 380/990), of South Arabian extraction themselves, had only an imperfect notion of the subject. Rippin (1990:160–161) states that the latter could not distinguish between → Ḥimyaritic and South Arabian, and both are considered unreliable by modern standards of scholarship.

The earliest exegetes of the 1st century A.H. acknowledged the existence of loans in Qurʾānic Arabic (Versteegh 1993:88–89). Pious Muslim scholars such as aš-Šāfiʿī (d. 205/820) would not subscribe to anything foreign in the *Qurʾān*, and suspected that the Šuʿūbiyya movement was behind such efforts to trace as many foreign elements in Arabic as possible (Kopf 1976:257–258). Others, like aṭ-Ṭabarī (d. 311/923) and as-Suyūṭī (d. 911/1505), sought to strike a middle course in this debate.

In Jeffery's (1938) monograph on approximately 322 loanwords in the *Qurʾān*, loanwords originating in Ethiopic and South Arabian constitute about 13 percent of his list, with the Ethiopic elements outnumbering the South Arabian ones (Zammit 2002:57–60). Indeed, the latter are not very numerous, and it cannot be otherwise, given the limited time frame during which they were introduced into Arabic, the fragmentary nature of South Arabian linguistic evidence (Ullendorff 1956:198), its restricted subject matter, and the possible contamination from Ḥimyaritic. South Arabian loanwords reached Arabic via a number of channels of diffusion: (i) directly from Yemen: Arabic *baʿīr* 'camel', *baġl* 'mule' (South Arabian *bġl*; Ethiopic *baql*), *baliya* 'to be consumed [a corpse]', *jazama* 'to swear an oath', *ḥiṣn* 'fortress', *Sabaʾ* [name of a country], *safīna* 'ship', *ʿarim* name of Sabaʾ's inundation, *fulk* 'ship', *yaġūṯ* [name of a particular idol], *yahūd* 'Jews'; (ii) loans attested in South Arabian and Ethiopic, which could have reached Arabic directly from South Arabian: Arabic *burhān* 'evident proof', *tubbaʿ* [title of Himyarite kings], *ḥizb* 'a party, sect', *xayma* 'a tent', *širk* 'associating anyone with God', *ṣuḥuf* 'pages of writing', *muṣḥaf* 'codex of the *Qurʾān*', *ṣarḥ* 'tower', *waṯn* 'idol', *fatḥ* 'judgment, decision', *Miṣr* 'Egypt', *tārīx* 'date' (South Arabian *wrx* 'month; date'), *wariq* 'silver; silver coins'; (iii) probable South Arabian loans: *ḥamida* 'to praise', *ʾasāṭīr* 'fables, tales'; (iv) South Arabian loans penetrating Arabic via Ḥimyaritic: Arabic *jafn* 'vine', *waṯaba* 'to sit'; (v) loanwords entering Arabic from other languages via South Arabian: a. from Aramaic sources (Ryckmans 1975:461–462): Arabic *bīʿa* 'church' (< Syriac); *ar-raḥmān* 'the Merciful', *rūḥ al-qudus* 'Holy Spirit', *masjid* 'mosque' (< 'sanctuary'), *ṣalawāt* 'prayers; places of worship', *qurbān* 'sacrifice, offering', *qissīs* 'priest', *kaffara* 'to expiate', and, very likely, also *saṭara* 'to write' and *ṣūra* 'form, picture'; b. from the pre-Semitic substratum in Akkadian and other Semitic languages (Schall 1982:146): *wayn* 'grapes [black or white]; wine'; and (vi) loans, such as *baʿal* 'Baʿal', whose origins are equally contended by the Aramaean North and South Arabian.

South Arabian, attested exclusively in consonantal inscriptions, and Arabic phonologies have much in common in their consonantal segments, and therefore South Arabian loans underwent few phonological adaptations. Furthermore, these loanwords needed very little, if any, morphological adaptation for full assimilation with Arabic patterns. Given that most South Arabian loans are nouns, it is difficult to make any comments of a syntactic nature. Semantically, the religious terms reflecting the transition from paganism to Judaism and Christianity (second half of the 4th century C.E. onward) are very conspicuous. Next in importance are loans reflecting material culture.

South Arabian influence is also felt in ancient Arabic onomastics, but the advent of Islam, signaling a break with the pagan past, left little scope for the propagation of the old names, including theophoric ones ending with *-īl*, like *šuraḥbīl*, *šarāḥīl*, which were perhaps borrowed from South Arabia (Corriente 1976:97–98).

The decline of the South Arabian languages and the subsequent emergence of Arabic inevitably led to the assimilation by Yemenite Arabic dialects of South Arabian elements (Beeston 1981:180). Future research will certainly identify a greater number of South Arabian loanwords in Arabic which were previously derived from other languages.

BIBLIOGRAPHICAL REFERENCES

Baalbaki, Ramzi. 1983. "Early Arab lexicographers and the use of Semitic languages". *Berytus Archaeological Studies* 31.117–127.

Beeston, Alfred F.L. 1981. "Languages of Pre-Islamic Arabia". *Arabica* 28.178–186.

Corriente, Federico. 1976. "From Old Arabic to Classical Arabic through the pre-Islamic Koine: Some notes on the native grammarians' sources, attitudes and goals". *Journal of Semitic Studies* 21.62–98.

Jeffery, Arthur. 1938. *The foreign vocabulary of the* Qur'ān. Baroda: Oriental Institute.

Kogan, Leonid E. and Andrey V. Korotayev. 1997. "Sayhadic (Epigraphic South Arabian)". *The Semitic languages*, ed. Robert Hetzron, 220–241. London: Routledge.

Kopf, Lothar. 1956. "Religious influences on Medieval Arabic philology". *Studia Islamica* 5.33–59.

——. 1976. "The treatment of foreign words in Mediaeval Arabic lexicology". *Studies in Arabic and Hebrew lexicography*, ed. Moshe H. Goshen-Gottstein, 247–261. Jerusalem: Magnes Press.

Militarev, Alexander and Leonid Kogan. 2000. *Semitic etymological dictionary*. I. *Anatomy of man and animals*. Münster: Ugarit-Verlag.

Rabin, Chaim. 1984. "On the probability of South-Arabian influence on the Arabic vocabulary". *Jerusalem Studies in Arabic and Islam* 4.125–134.

Rippin, Andrew. 1990. "Epigraphical South Arabian and Qur'ānic exegesis". *Jerusalem Studies in Arabic and Islam* 13.155–174.

Robin, Christian. 1991–1993. "La pénétration des Arabes nomades au Yémen". *L'Arabie antique de Karib'īl à Mahomet: Nouvelles données sur l'histoire des Arabes grace aux inscriptions. Revue du Monde Musulman et de la Méditerranée* 61.71–88.

Ryckmans, Jacques. 1975. "Les inscriptions sud-arabes anciennes et les études arabes". *Annali dell'Istituto Universitario Orientale di Napoli* 25.443–463.

Schall, Anton. 1982. “Der arabische Wortschatz”. *Grundriß der arabischen Philologie. I. Sprachwissenschaft*, ed. Wolfdietrich Fischer, 142–153. Wiesbaden: L. Reichert.

Ullendorff, Edward. 1956. “The contribution of South Semitics to Hebrew lexicography”. *Vetus Testamentum* 6.190–198.

Versteegh, Kees. 1993. *Arabic grammar and Qurʾānic exegesis in early Islam*. Leiden: E.J. Brill.

Zammit, Martin R. 2002. *A comparative lexical study of Qurʾānic Arabic*. Leiden: E.J. Brill.

Zubaydī, ʿAbd al-Karīm az-. 1994. “al-ʾAlfāḍ al-ʾaʿjamiyya fī l-*Qurʾān* al-Karīm: Dirāsa tārīxiyya muʿjamiyya”. *Proceedings of the Colloquium on Arabic Lexicology and Lexicography, Budapest, September 1–7, 1993*, ed. Kinga Dévényi, Tamás Iványi, and Avihai Shivtiel, 103–123. Budapest: Eötvös Loránd University.

Martin R. Zammit (University of Malta)

South Arabian, Modern

The Modern South Arabian languages (Johnstone 1977, 1981, 1987; Leslau 1938, Lonnet 1994), which are still spoken, constitute the substratum for Arabic in the southern part of the Peninsula, as they did for Ancient South Arabian. Their relationship with Arabic has not always been the same, depending on historical circumstances, and in particular it is noteworthy that mere occasional contacts with rival Arabic tribes did not produce the same linguistic result as the inexorable increase in the cultural weight of the language of Islam.

The Modern South Arabian language family was shaped by two historical events: the west-east split – doubtless a response to the settling of the ancient South Arabian kingdoms – and the migration of a group belonging to the eastern branch toward Soqotra Island. Western Modern South Arabian, i.e. Mehri (*mähri*), is deeply marked by continuous contact, over more than three thousand years, with the languages of the ruling communities, first Ancient South Arabian, then Arabic. Eastern Modern South Arabian, continental Jibbali (*jibbāli*) and insular Soqotri (*suquṭri*), developed well away from these influences, in its remote geographical location. This picture became more complicated a few (six?) centuries ago, as Mehri tribes coming from the west seized Dhofar, the eastern zone of the continental Modern South Arabian-speaking area, and adopted the language of the subdued populations (Jibbali), except for two minorities that acquired the language of their conquerors: the groups which now speak Harsusi (*ḥarsūsi*) and Bathari (*baṭḥari*) – both languages being therefore dialects of Mehri. From this outline, one can draw the following broad conclusions: the westernmost Mehri dialects in Yemen have been rather deeply altered by the contact with Arabic, while Dhofari dialects have been much less so. Among the ruling (Mehri) tribes, the Ǝḥkloʾ (Jibbali name) or Qarā (Mehri name), Jibbali is marked by Arabic to a limited extent only, and even less among the nontribal dependent groups, the Śḥäroʾ or Gəbloʾ. The most remote dialects of Soqotri, in the high mountains, almost completely escaped the influence of Arabic, which was not possible for coastal dialects. Moreover, it is probable that Arab tribal factions from the Yemenite highlands migrated to Soqotra, where they acquired the Soqotri language but without losing Arabic.

The Modern South Arabian-speaking communities may individually have close relationships with Arabic-speaking communities: such is the case of the Ḥarāsīs, whose language contains quite an important Arabic element and presents symptoms of extinction. It is also the case of the inhabitants of ʿAbd-el-Kūrī, a small island close to Soqotra, who are in regular and exclusive contact with certain (Arabic-speaking) villages of Hadramawt, to such an extent that their dialect has evolved in a very peculiar way under the influence of Arabic. It remains to say a word about Baṭḥari, which is becoming extinct under the pressure of Jibbali as much as under that of Arabic, and about Hobyot (*həwbyōt*), a very well preserved language within a small area on the border of Dhofar and Mahra, deeply marked by Mehri and Jibbali and hardly by Arabic.

In modern times, a new wave of borrowings from Arabic has its origin in modernity itself. The states of Yemen and Oman are using Arabic in an effort to strengthen their authority, settle the nomads, and support centralized – therefore Arabicized – economic development. These factors do not favor the survival of local cultural identities; furthermore, the modern means of language circulation (e.g. radio, television, and cassette tapes) are almost entirely dedicated to Arabic.

Multilingualism is standard among Modern South Arabian speakers, who nearly all speak

Arabic and, except in Soqotra, one or two other Modern South Arabian languages as well.

Arabic borrowings in Modern South Arabian may be from Classical Arabic – when they are connected with Islamic culture – but also, and much more abundantly, from Yemenite and Omani colloquial Arabic. The most recent loans may be attributed to international Standard Arabic. As for the oldest ones, they are to be situated in the era preceding the arrival of the Sabeans, when the Peninsula was shared between North Arabian and South Arabian dialects.

Since the end of the ancient South Arabian kingdoms, Modern South Arabian has been in contact only with Arabic on dry land (leaving aside the British colonial episode, which hardly affected it directly). However, the nautical traffic of men and their material and cultural goods – and among them languages and vocabulary – has always been very heavy in the northeast of the Indian Ocean, and Arabic as well as Modern South Arabian played its part on that scene, giving and receiving through the centuries. It follows that it is not always easy to determine the source and the circuitous route of a given loanword.

The kinship proximity of Modern South Arabian with Arabic is an important characteristic; it appears in their extensive common lexical/radical stock, their analogous morphology, and their partially identical phonology, which highly facilitates the integration of loanwords. The languages can easily be graded according to this 'similarity' to Arabic: Harsusi, Mehri, Bathari, Hobyot, Jibbali, and Soqotri (in decreasing order of similarity). The morphological and phonetic differences, illustrated below, allow us to identify the most recent borrowings. No more than a sketch is presented here, as it disregards the very rich dialectal diversity in Arabia. Nor does it take into account the crucial sociolinguistic facts labeled 'code-switching': the borrowings considered here are those present in (as much as possible) monolingual speech.

Any part of speech may be borrowed. For instance, in the pronominal system, Modern South Arabian has a characteristic *-i* in the 1st person singular suffix in all contexts, but there are, especially in Harsusi, occurrences of *-ni* after verbs and particles. This is obviously due to the pressure of Arabic.

The simplest case is that of lexical borrowing. The noun for 'cow' gives a good example of the various situations. The Arabic root *b-q-r* and the Modern South Arabian root *l-ʾ-y/l-h-y* are in competition: *bəḳərēt*, pl. *bəḳār*, a fully integrated borrowing (it even has a diminutive, *bəḳərēnōt*), is found in Harsusi and Western Yemenite Mehri; the Modern South Arabian root appears only with the collective *lháytən* in some Mehri dialects, but in all numbers in the Mehri of Jādib (Yemen) – sg. *leh*, pl. and coll. *lháyten* – and in all the other Modern South Arabian languages, which lack the Arabic root. The root *l-ʾ-y* being present in the north of the Yemen, away from the Modern South Arabian area, the hypothesis of very ancient borrowing (in either direction) must not be discarded. It is not unusual for two words of different origin to compete in expressing one meaning. In the following pairs of examples, in Hobyot, the first item is Modern South Arabian and the second is an Arabic loanword: 'heart' *ʾəlbēb* and *ḳalb*; 'sky' *hétəm* and *sēmɛ́ʾ*; 'to spit' *fosġ* and *tfōl*; 'dress' *xəllōḳ* and *ksuwɛ́t*. The loanword may take on a meaning somewhat diverging from the original: thus in Qishn Mehri, in *ḥarmēt*, pl. *ʾāgzōn* 'woman', one recognizes Arabic loanwords, but their meaning is that of Arabic *marʾa* 'woman' and *nisāʾ* 'women'; they have lost the specific value of Arabic *ḥurma* 'spouse' or *ʿajūz* 'elderly woman'. The loanword's root may be present in Modern South Arabian with a different meaning; thus, Jibbali *ż́ɛrbɛ́t* 'hit' is an Arabic loanword, while the Modern South Arabian root *ż́-r-b* means 'to suffer pain' and not 'to hit' (roots *k-r-ś, r-ġ-ṯ, s-b-ṭ, s-ḳ-ḥ*, etc.).

Borrowings are often visible at the phonetic level. In all the Modern South Arabian languages, *ż́* and *ṯ̣* are clearly distinct (the latter being often realized as *ḏ̣*, especially in Harsusi): the root of Mehri *ż́əfərēt*, Harsusi *ż́əfərōt*, Jibbali *ż́əfrét*, Soqotri *ż́afreh* 'plait of hair' differs from that of Mehri *ṯ̣ayfēr* etc. (Soqotri *ṯ* instead of *ṭ*) 'fingernail'. Such is not the case of the neighboring vernacular Arabic, which merges the Arabic reflexes of *ż́* and *ṯ̣*, i.e. *ḍād* and *ḏ̣āʾ*, respectively, into one phoneme *ḏ̣*; therefore, conclusions can be drawn concerning the borrowings: Harsusi, Mehri, Jibbali, Hobyot *ṯ̣arb*, Soqotri *ṭarb* 'piece of wood' is Modern South Arabian (root *ṯ̣-r-b*), while Harsusi *ṯ̣arb* 'kind [n.]', Harsusi Mehri *ṯ̣ərbēt*, Jibbali *ṯ̣arbɛ́t* 'gust of wind', Jibbali also 'gunshot', are Arabic loanwords (root *ḍ-r-b*). Similarly, Mehri *ṯ̣ərūr* ‘harm’ is borrowed (Arabic *ḍ-r-r*), but Harsusi *ẓ́ərr*, Jibbali *ẓ́err*, Mehri *ẓ́ər* ‘to harm’ are Modern South Arabian. An interesting case is that of the Semitic root **ṯ̣-h-r*: to it pertain Harsusi *ṯ̣əhōr*, *ṯ̣əhār*, *ẓ́əhōr*, *ləhōr*, Mehri *ẓ́əhēr*, *ẓ́əhār*, Jibbali *ẓ́ɛhɛ́r* ‘to appear’, and many derived forms with *ẓ́-h-r*, as well as Mehri, Harsusi *ṯ̣ahr*, Jibbali *ṯ̣ɔ́hur* ‘noon’, Harsusi *ṯ̣ahr*, *ẓ́ahr*, Mehri *ẓ́āhər*, *ẓ́áhər*, Jibbali *ẓ́ɛ́hər* ‘back’. Doubtless, these words are the more or less integrated results of a series of borrowings from Arabic, spaced in time. As for Jibbali *ṯ̣ɛhɛ́r* ‘to be finished’, Soqotri *ṭáhar* ‘to go’, they are genuine Modern South Arabian.

Irregular correspondence of a Semitic sibilant is a clue: e. g., Mehri *sənēt* ‘year’ does not fit, as **hənēt* is expected. Actually, some dialects have *sənéh*, where the Arabic feminine is still visible.

Harsusi shows a *g/y* variation that reveals borrowing from (Omani) Arabic words in *y* (< *j*). Nonborrowed words have no *y* variant. Examples: *fagr, fayr* ‘dawn’, *negáḥ, neyáḥ* ‘to succeed’, *wāyeb* ‘necessary’; compare *yann* ‘madness’ and *génni* ‘jinni’, both loanwords, the latter from more standard Arabic (the Modern South Arabian word for ‘jinni’ is *ḳēʿi*). Dhofari Mehri shows, more rarely, a corresponding *g/j* variation; compare *jamʿ* ‘fist’ and *gamʿ* ‘communal prayer’, borrowed from a local dialect and – with better integration – from Standard Arabic, respectively.

Likewise for *ḳ/g(j)*: Harsusi borrows words with *g* and even *j* (< *q*): thus, *gəbāyli* ‘tribesman’ (< Arabic) against *ḳəbīyəl* ‘tribes’, *tagrīb* ‘approximately’ against *ḳarb* ‘proximity’, Harsusi *ghâweh* against Mehri *ḳahwēt* ‘coffee’, Harsusi *gəṣébəh* ‘reed’ against *ḳəṣəbə́t* ‘crowbar’, Mehri both *ḳəṣəbēt*. Nonborrowed words have no *g(j)* variant, but better integrated loanwords have *ḳ*: *məḳlūʾ* ‘detestable’ (root *ḳ-l-ʾ*).

Some Modern South Arabian dialects show a tendency to realize the emphatic consonants pharyngealized-uvularized instead of the original ejective articulation. This phenomenon first spreads in the fricatives: *ṯ’* > *ḏ̣* (IPA: θ’ > ðˤ) *ś’* > *ẓ́* (IPA: ɬ’ > ɮˤ), *š’* > *ẓ̌* (IPA: ʃ’ > ʒˤ), *s’* > *ẓ*, *ṣ* (IPA: s’ > sˤ) – notice the voicing: its first stage is a creaky voice (laryngealization) –, and then in the plosives *t’* > *ṭ* (IPA: t’ > tˤ), *k’* > *ḳ*, *Ģ* (IPA: k’ > kˤ, ɢˤ). This change, due to the pressure of Arabic, is a recurrence of what Arabic itself underwent in its early history. The reverse influence also occurs: in some border areas one notices Arabic dialects whose emphatics (*ṭ*, *ṣ*, *ḍ*, *ḏ̣*, *q*) are realized as ejectives (IPA: t’, s’, ɬ’, θ’, k’).

Interference may arise with a nearly identical Arabic word, the Arabic stress imposing itself on the Modern South Arabian word, for example: ‘she began’ Soqotri *bə́d'ɔh*, (local) Arabic *bədá'at*, Soqotri with an Arabic stress *bədɔ́'ɔh*.

It is sometimes stated that Soqotri words with *x* or *ġ* have to be Arabic loanwords. This is an old mistake due to the fact that the first-studied dialects are still taken for the language as a whole. Anyone visiting Socotra can hear *ġayg* ‘man’, *xa* ‘mouth’, which are not borrowings.

In many cases, the borrowing is revealed by a morphological clue, such as a pattern unknown to Modern South Arabian. For example, the pattern of the past participle of simple verbs is *m*vCC*ū*C in Arabic but *m*vCC*ī*C in Modern South Arabian. In Soqotri *máxlɔḳ* ‘human being’, one easily identifies Arabic *maxlūq* ‘human being < created’, as opposed to Harsusi *mxəlīḳ* ‘human being’, Jibbali *məxléḳ*, Mehri *məxlīḳ* ‘poor fellow’; and similarly for Mehri *məbrūk*, Harsusi *əmbərūk* vs. Jibbali *məbrék* ‘blessed’.

Sometimes a doublet appears: Mehri *bənnāy* (< Arabic) and *mənnōy* ‘builder’, Harsusi *xálləṣ* (< Arabic) and *xāləṣ* ‘to finish’; most probably, the two words not designated as loanwords were actually borrowed long ago and fully integrated.

Quadriliteral nouns *CvCCvC yield Modern South Arabian *CvCC*ē*C. Therefore, borrowing is obvious in such nouns as Harsusi *mébrad* ‘file’ (vs. Mehri *məbrēd*), Mehri and Jibbali *mə́xləṣ* ‘sincere’ (< Classical Arabic *muxliṣ*).

Borrowing a verb can produce a supernumerary derived theme: Harsusi *gátma*, *egtōma*, *əgtəmáʾ* > ‘to gather’: the two regular verbs with infixed *-t-* and an Arabic loanword, respectively; Harsusi *aḳāṣər* ‘to fall short’, *ḳáṣṣər*, *gáṣṣər* ‘to pass [time]’: Modern South Arabian and borrowed Arabic D-forms, respectively. The Harsusi participle *mətḥáffi* ‘barefoot’ must be a loanword (Omani Arabic *mitḥáffi*), since Modern South Arabian verbal derivation admits no *t-* prefixation (it is always infixed), nor gemination.

Many Arabic borrowed nouns stand out by their conspicuous Arabic feminine ending in -v*(h)* instead of Modern South Arabian (except Soqotri) *-(*v*)t*. Examples: Jibbali *ṯórəh* ‘revolution’, *sáʿáh* ‘hour; watch’, but some loanwords have been adapted morphologically: Jibbali *rišt* ‘trigger’ < Omani Arabic *rīšah*.

Since Modern South Arabian and Arabic morphology and phonetics are so close, and their lexical material shares so many fields and items pertaining to (South) Arabia and its ancestral Bedouin culture – not to mention the fact that the Mahra were among the first communities to embrace the Prophet’s religion – and with such an extent of bilingualism, it is not surprising that no lexical domain stands out in Arabic borrowings, apart from that of modernity (e.g. the lexicon of administration, politics, and technology).

Traditional measuring, valuation, and trade are still performed with the inherited numbers and units, but activities oriented toward the outside world have introduced enough Arabic to eliminate, for instance, most Modern South Arabian numbers above ten, and to produce mixed and disrupted systems of enumeration: Jibbali *stīn* ‘sixty’ (Arabic) against *šɛt*, fem. *štət* ‘six’; Mehri *sōdəs* (Arabic *sādis*) against its fem. *šədṯīt* ‘sixth’; Harsusi *ṯəlaṭṭáʿəš* ‘thirteen’ (Arabic), against *śəlayś*, fem. *śāfayt* (< *śaʾṯayt*) ‘three’; *śēləṯ* or *śēləś* ‘third’. However, certain Soqotri dialects keep using their traditional decades, such as *ʰyaʿt ʿəśárhən* ‘sixty [six tens]’ (*ʰy*: murmured palatal approximant).

The Arabic element shows remarkably in the large number of borrowed particles: Mehri and Harsusi *yəkūn* ‘maybe’, against Jibbali *(ḥa-)yékən* ‘maybe; about’, Mehri *yə́mkən* ‘perhaps’; Mehri and Harsusi *yā* ‘oh…[voc.]’ (against Modern South Arabian *ʾā*); Harsusi *yā-rēt*, Mehri *yā-rayt*, Jibbali *ya-rét* ‘would that…!’; Jibbali *bdan*, Mehri *ʾábdan* ‘never, not at all’, against Modern South Arabian, Mehri and Harsusi *bəháwʾ*, Jibbali *bhóʾ*, borrowed by Omani Arabic: *bhawʾ*; Soqotri *mseb* ‘because’ < Arabic *min sabab*; Mehri and Jibbali *ʾəkīd* ‘certainly!’; Mehri and Jibbali *ʾáywa(h)*, *ʾēwa* ‘yes’, against Modern South Arabian, Mehri *yēya*, *ʿahā*, Jibbali *yɛ́yɛ́*, *‘ahā̃n*; Mehri *hɛ̄śən* ‘what?’ < Arabic *ʾāšən* < *ʾayy šayʾ-in* ‘which thing?’. A more ‘threatening’ borrowing is that of Arabic *li-ʾanna* ‘because’: Mehri and Jibbali *ən-*, *lən-*, Harsusi *yənn-* + Arabic suffixed pronoun.

Expected borrowings that did *not* occur are culturally significant. ‘Allah’ is God’s name in Soqotri but not in Mehri, which, in Islamic everyday phraseology – and even in proper reference to God – uses other words: *(a)bā̆li* (‘my lord’, root *b-ʿ-l*), *rəḥmōn* (a pre-Islamic divine name; ‘the Merciful’), *ʾāl* (either Semitic **ʾal* ‘god’ or borrowed from Jibbali: **a-baʿl > āʿl* > *ʾāl*), hence Mehri *Abdāl* = *ʿAbdallah*. Several exclamations based on *Allāh* were borrowed, their semantic motivation being lost: Jibbali (‘God’ is *ʾōź*, *ʾóʿoź*, *ā̆rḥəmún*) *ʾəllɛ́!* ‘well then!’, *bəllɛ́!?* ‘really!?’, *wəllɔ́!* ‘true!’, Mehri *ʾállah!* ‘indeed!’.

Bibliographical references

Johnstone, Thomas M. 1977. *Ḥarsūsi lexicon and English-Ḥarsūsi word-list*. London: Oxford University Press.

——. 1981. *Jibbāli lexicon*. London: Oxford University Press.

——. 1987. *Mehri lexicon and English-Mehri word-list*. London: School of Oriental and African Studies. [Bound with *Index of the English definitions in the* Jibbāli lexicon, comp. G. Rex Smith.]

Leslau, Wolf. 1938. *Lexique soqoṭri (sudarabique moderne) avec comparaisons et explications étymologiques*. Paris: C. Klincksieck.

Lonnet, Antoine. 1994. “Quelques résultats en linguistique sudarabique moderne”. *Quaderni di Studi Arabi* 11.37–82.

——. 2006. “Les langues sudarabiques modernes (sémitique)”. *Faits de Langue* 27.27–43.

Antoine Lonnet (CNRS)

South Semitic Languages

1. Extralinguistic situation

In this discussion of the relationship between Arabic and the South Semitic languages, Arabic is represented by its two forms: (i) the literary written form, Classical Arabic, and its somewhat modified modern variant, Modern Standard Arabic; and (ii) the spoken form, called Neo-Arabic, represented by a large number of local and social dialects and ‘intermediate’ interregional forms of spoken koine (usually called → Middle Arabic). The modern Arabic dialects, which are spread over vast territories in the Middle East and North and Northeast Africa, are not written. In their phonetic/phonological and morphosyntactic features, the dialects differ from Classical and Modern Standard

Arabic, as well as among themselves. Some of these dialectal differences may be inherited from the ancient Arabic dialects. Some of the isoglosses between Arabic and the South Semitic languages may be common only with some particular group of dialects but not with Classical Arabic, or vice versa.

The South Semitic languages are also heterogeneous from the extralinguistic point of view: some of them are represented by ancient languages, preserved in written texts, others are represented by modern languages with their own written/literary form, and, finally, some of the modern South Semitic languages are never written.

In South Arabia, several ancient languages (Old South Arabian or Epigraphic South Arabian, also called Sayhadic) existed, including Sabaean, Qatabanian, Hadramitic, and Minaean (also called Madhabic). The period covered by the epigraphic texts in these languages is between the beginning of the 1st millennium B.C.E. and the middle of the 6th century C.E.

Along with the epigraphic languages in Yemen, there existed in the first centuries C.E. an unwritten language, called by early medieval Arabic sources from Yemen → Ḥimyaritic. This language remained in use in some regions of Yemen until the 10th–13th centuries C.E. According to Arab Yemenite authors, the language had some features distinguishing it from the older epigraphic languages.

At present in the territory of South Yemen and the Sultanate of Oman, and on the island of Soqoṭra and the neighboring islets of ʿAbd al-Kūri and Samḥa, a group of unwritten languages is found, the so-called Modern South Arabian languages (→ South Arabian, Modern). This group consists of the following languages: Mehri, Ḥarsūsi, Baṭḥari, Hobyot, Jibbāli (Šxauri), and Soqoṭri.

The Ethiopian Semitic (Ethiopic Semitic/Ethio-Semitic) languages, too, are represented both by ancient and modern languages. The Ethio-Semitic languages were and are used in the territories of modern Ethiopia and Eritrea. They are divided into two branches: North Ethiopic, including Geʿez, Tigrinya, and Tigré; and South Ethiopic, including Amharic and Argobba, Harari, East Gurage (Gafat, Soddo, Goggot, Muher), West Gurage (Mäskan), Central Gurage (Ezha, Chaha, Gura), and Peripheral (Gyeto, Ennemore, Endegen).

The oldest written language from the Ethio-Semitic group is 'Old Ethiopian', or Geʿez. The first written texts in this language date from the mid-4th century C.E. As a spoken language, Geʿez existed until approximately the 10th century C.E. At present, the official language of Ethiopia is Amharic, a language with both written/literary and spoken forms. The first written texts in Amharic date from the 14th century C.E. The official language of Eritrea is Tigrinya (Tigray/Tigrai). Its written form dates from the beginning of the 19th century. Another modern language of Eritrea, also in both written and spoken form, is Tigré. The other modern Ethio-Semitic languages – Argobba, Harari, the Silte group (Silte, Inneqor, Wolane, Zway), Gafat, the Gurage group, and others – exist as spoken languages only.

The question of the relationship between Arabic and the South Semitic languages is connected with the difficult problem of positioning Arabic within the genetic subgrouping of the → Semitic languages. Arabic in general is characterized both by features uniting it with the → Northwest Semitic languages and by features it shares with the South Semitic languages. As a result, in various genetic and typological classifications, Arabic occupies a middle (perhaps even 'intermediate') position between the North and Northwest Semitic languages on the one hand, and the languages of differing genetic proximity, which are usually, mostly on geographic grounds, classified as South Semitic, on the other (for the history of this question, see Faber 1997).

At the same time, Classical Arabic and modern Arabic dialects represent different stages of historical development and different effects of substratal and adstratal influence (→ substrate). Their genetic proximity may be overridden by various structural, typological, and areal features. The genetic proximity between the South Semitic languages may similarly be overridden by their structural, typological, and areal features.

Accordingly, some phonological and morphosyntactic characteristics may be common to Arabic in general or the South Semitic languages in general. Other characteristics are only present in a particular subset of Arabic dialects and/or a particular subset of South Semitic languages, forming features that are partially shared (areal isoglosses). Thus, when establishing relations between Arabic and the South Semitic languages, a distinction should be made between their shared features and areal isoglosses.

The question of shared and distinguishing features of Arabic and the South Semitic languages are addressed here, without aiming at a classification in the sense of determining the place of Arabic within the Semitic family of languages.

2. PHONETICS AND PHONOLOGY

The phonological systems of the languages under consideration differ in degree of preservation and in degree of reduction and innovation relative to the reconstructed phonological system of Proto-Semitic. The richest consonant systems are present in Modern South Arabian languages, with the exception of Soqoṭri, in Epigraphic South Arabian (Kogan and Korotayev 1997:222–223), and in Classical Arabic. The consonant systems of modern Arabic dialects, Soqoṭri, and Ethio-Semitic languages are characterized by both shared and particular tendencies of phonetic/phonological inventory change.

The following tendencies can be classified as shared: the reduction of interdental fricatives: *ṯ* > *s*/*t*; *ḏ* > *z*/*d*; *ṯ* > *ṣ*/*ḍ*/*ḏ*, also *ġ* > ʿ, *x*; and the reduction of the laryngeal plosive ʾ > Ø/*w*/*y*. These tendencies and their results are present, however, in the phonetic/phonological systems of most known Semitic languages in general. Among the Arabic dialects, the interdentals and *ġ* are preserved in most dialects of Yemen and Iraq and also in East Arabian dialects (Johnstone 1967: II(A); Jastrow 1982:129; Kaye and Rosenhouse 1997:275). In some dialects of Yemen, preservation of the lateral articulation of the emphatic *ḍ* has been observed (Landberg 1901:637; Rhodokanakis 1911:82; Fischer 1982:85; → *ḍād*).

The phonetic-phonological system of each South Semitic language has its own particular features. Thus, the consonant systems of Geʿez, Tigrinya, Amharic, and some of the other Ethio-Semitic languages of the South Ethiopian branch are characterized by the presence of labiovelar phonemes or labialized allophones (Gragg 1997:244; Kogan 1997:424; Hudson 1997:458; Hetzron 1997:536–538), as well as a number of particular features.

Comparing the vowel systems of Arabic and the South Semitic languages, especially those without an established written tradition, is a difficult task. In spoken forms, the quantitative and qualitative characterization of vowels and their phonetic/phonological status cannot always be established, since these characteristics are often determined by the syllable structure of the word, position of stress, and consonantal context. In general, modern Arabic dialects and unwritten South Semitic languages share a tendency to enhance the vowel inventory, to weaken or reduce the original short vowels in phonological function, and to create new long or short vowels as a result of contraction of original combinations -V+ʾ-, -V+*w*-, -V+*y*-, etc.

Despite the great variation among the phonetic/phonological systems of Classical Arabic, modern Arabic dialects, and South Semitic languages, their sound systems share one feature, and a number of areal isoglosses may be drawn. The common shared feature of these systems is the presence of /f/, corresponding to /p/ in Northeast and Northwest Semitic languages. No historical or prosodic explanation has as yet been offered for this feature of Arabic and all South Semitic languages. The ancient Epigraphic South Arabian languages are also included in the *f*-languages group. Even though the phonetic realization of their consonants is unknown, the position of these languages relative to the neighboring Arabic and South Arabian languages and the close genetic ties to the Ethio-Semitic language Geʿez make it possible to interpret the corresponding grapheme as /f/. Its phonetic characterization is confirmed by data from early Arabic Yemenite authors about Ḥimyaritic, where /f/ is also found. The regular character of Arabic and South Semitic correspondences /f/:/p/ is illustrated by the following examples:

> ‘mouth’: Akkadian *pū*; Hebrew *pä*; Mandaic *pum* vs. Arabic (Classical and dialects) *fū(h), fam, fum*; Geʿez *ʾaf*; Tigré, Tigrinya *ʾaf*; Amharic, Argobba, Harari, Gurage *af* (Militarev and Kogan 2000: 195–197).
>
> ‘soul, breath’: Akkadian *napišt-u*; Ugaritic *npš*; Hebrew *näpäš*; Syro-Aramaic *napš-a* vs. Arabic (Classical and dialects) *nafs-*; Epigraphic South Arabian *nfs*^{1}; Geʿez *nafs*; Tigré *näfəs*; Tigrinya *näfsi*; Amharic *näfs* (other Ethio-Semitic languages also have *näfs*); Mehri, Ḥarsūsi *nəfsēt* ~ *nefesét*; Jibbali *nəfsét*; Soqoṭri *nófos* (pl.) (Militarev and Kogan 2000:307–308).

Along with this single phonetic feature common to Arabic and all South Semitic languages, a number of areal isoglosses are found in the phonetic systems of these languages, uniting territorially close languages and dialects:

i. Palatalization of *k, g: k* > affricate *č* or fricative *š; g* > affricate *j* or fricative *ž*, up to palatal *y*. The palatalization *g* > *j* was already firmly present in Classical Arabic by the 10th century. In modern Arabic dialects of the Arabian Peninsula, the palatalization process is stronger to the east and southeast, cf. Classical Arabic *dajāja*, Syro-Palestinian *(da)jāja*, Kuwaiti *dayāya* 'hen'. Palatalization *k* > *č* in modern Arabic dialects: Iraqi, Kuwaiti *čebīr* < *kabīr* 'big'. An analogous process partly affects the *q* phoneme before fronted vowels: Iraqi *rafīj* < *rafīq* 'companion, friend' (Johnstone 1967:5; Kaye and Rosenhouse 1997:271–275). The isogloss for palatalization of stops can be traced to the north of Mesopotamia and covers some of the modern Aramaic dialects (Tsereteli 1991:238; Jastrow 1997:350). The palatalization isogloss is being 'skipped' by the Yemenite dialects, but it is found again in the Modern South Arabian languages (Simeone-Senelle 1997:384), e.g. Classical Arabic *kabid* ~ *kibd*, Iraqi dialect *čibde*, Mehri *šəbdīt*, Jibbāli *šubdet*, Soqoṭri *šíbdeh* 'liver'; Classical Arabic *kariš/kirš*, East Arabian dialects *čerš*, Yemenite dialects *käriš*, Mehri *kīrəš*, Jibbāli *širŝ*, Soqoṭri *šéreš*, but in Geʿez and other Ethio-Semitic languages again *karš/kärs* 'belly, stomach' (Johnstone 1965:215–216, 234–241; Simeone-Senelle and Lonnet 1991:1461, 1474; Militarev and Kogan 2000:136–137). In Modern South Arabian, irregular cases of palatalization *q* > *č/š* are also attested: Classical Arabic *qidr*, Iraqi dialect, Bedouin dialects *jider* (but Yemenite dialect *qidr*), Mehri *qadər*, Soqoṭri *žádhər* ~ *šádher* 'pot' (Johnstone 1987:224); Classical Arabic *qaryat*, East Arabian dialects *jerye*, Yemenite dialects *qarya*, Šxauri/Jibbali ṣ*erét* ~ *čiret*, but Soqoṭri *qéryeh* 'town, settlement' (Müller 1907:146; Leslau 1936: 385). Similar effects can be observed with reflexes of Semitic **b-k-y* 'to weep': Classical Arabic *ya-bkī* (imperf. 3rd pers. masc. sg.), Geʿez *yə-bki* (jussive 3rd pers. masc. sg.), Mehri *yə-bkē* (subj. 3rd pers. masc. sg.), but Soqoṭri *i-béśe* (imperf. 3rd pers. masc. sg.), Iraqi Arabic *yi-bčī* (imperf. 3rd pers. masc. sg.).

ii. Only Modern South Arabian languages and the Arabic dialects of South Arabia in contact with them are characterized by the lateral articulation of the ancient Proto-Semitic emphatic lateral **ṣ̂* (Simeone-Semelle 1997:382). In Classical Arabic, lateral pronunciation of the *ḍ* phoneme was attested by early Arabic philologists in the 8th century (Sībawayhi, *Kitāb* II, 404–406, 429; → *ḍād*). But already in the 10th–12th centuries, Arabic Yemenite authors considered the alternation *l/š/ḍ* a feature of the language of Yemen inhabitants, in contrast to Classical Arabic: *šaḥm/laḥm* 'prey; gain' (Hamdāni, *ʾIklīl* X, 26); *nāḍa-hu/nāša-hu* 'he removed it' (Našwān, *Šams al-ʿulūm* 107). In modern dialects of the southern regions of Yemen, the lateral character of the *ḍ* phoneme is attested: *ḥawl̩/ḥawḍ* 'cistern, reservoir'; *ġaral̩* < *ġaraḍ* 'target' (Landberg 1901:637; Rhodokanakis 1911:82).

iii. Free alternations *f/ṯ* and *ṯ/f* are attested in the whole South Arabian area, for Arabic dialects as well as for some of the Modern South Arabian languages: *fum* ~ *ṯum* 'mouth'; Ḥarsūsi *ṯawr*, but Mehri, Jibbāli *fōr* ~ *fʿór* 'ox, bull' (Landberg 1901:538; Rhodokanakis 1911:83; Behnstedt 1985: 185; Johnstone 1987:87).

iv. In Arabic in general, as well as in the South Semitic languages, preservation of *w* as the first radical in the initial position is attested, unlike the Northwest Semitic languages, where the shift *w* > *y* occurs at the beginning of a word, e.g. Arabic *wrd*, Sabean *wrd*, Mehri *wrd*; Geʿez, Tigré, Tigrinya, Amharic, Harari *wrd* vs. Hebrew, Aramaic, Syriac, Phoenician, Ugaritic *yrd* 'to go down [to the water]' (Leslau 1987:618); Arabic *wark* ~ *wirk*; Sabean *wrk*; Mehri *wərkit*; Ḥarsūsi *wərkēt* vs. Hebrew *yārēk*; Ugaritic *yrk* 'hip, thigh, side' (Militarev and Kogan 2000:258–259).

3. MORPHOLOGY

3.1 *Pronominal system*

Independent and bound personal pronouns, in Arabic as well as in South Semitic languages, represent different degrees of completeness and reduction. The richest personal pronoun systems are attested in Classical Arabic and Modern South Arabian languages. These systems include not only singular and plural forms, with gender distinction for 2nd and 3rd person plural, but also dual forms. Soqoṭri, where in the 2nd person masculine and feminine plural have merged, is an exception (Simeone-Senelle 1997:387–388). The pronominal system of all Arabic dialects and the Ethio-Semitic languages lacks dual forms. The pronominal system of Epigraphic South Arabian languages is known partially; the 3rd person pronouns have singular, dual, and plural forms.

The next degree of reduction is represented by systems where 2nd and 3rd person plural lack gender distinction. According to this feature, the languages and dialects form the following groups: those where gender in 2nd and 3rd person plural is preserved (giving a total of ten forms): Yemenite Arabic dialects and some Eastern Arabic Bedouin dialects, with the dialects of western Yemen having two distinct forms for the independent 1st person pronoun: *ˀana* (1st pers. masc. sg.), *ˀani* (1st pers. fem. sg.) (Behnstedt 1985:71); among Ethio-Semitic languages, the full system is present in Geʿez, Tigré, and Tigrinya of the northern branch and Soddo, Muher, Ezha, and Inor of the southern branch.

The gender distinction in 2nd and 3rd person plural is lacking (giving a total of eight forms) in Arabic sedentary dialects of Iraq (Baghdad), Syria (Damascus), Egypt (Cairo), the so-called Maghrebinian dialects of North Africa, Maltese (Zavadovskij 1962:101–102), and Ethio-Semitic of the southern branch (Amharic, Argobba, Harari, East Gurage, Gafat).

Regardless of the material expression of these pronouns, the 'full personal pronoun system' isogloss unites Classical Arabic and the Modern South Arabian languages. The isogloss of 'incomplete system with ten members' covers part of the Arabic dialects and part of the Ethio-Semitic languages, while the isogloss of 'incomplete system with eight members' also covers part of the Arabic dialects and part of the Ethio-Semitic languages of the southern branch.

With respect to the material expression of the personal pronouns, Classical Arabic, Arabic dialects, and the South Semitic languages represent various degrees of proximity to the Common Semitic system and a wide range of innovations, especially widespread in the Ethio-Semitic languages.

There is one major isogloss in this complicated picture of pronoun forms, the suffixed 2nd person feminine personal pronoun *-ki/ik: -(i)č/(i)š*. This isogloss divides Classical Arabic and some Arabic dialects. In the same way, it divides the northern and southern branches of the Ethio-Semitic languages. The second variant -(V)*š* includes all the Modern South Arabian languages.

The isogloss *-ki/-ik/-k* includes Classical Arabic, Maghrebinian, Egyptian, Syro-Palestinian dialects, and part of the dialects of Iraq and the Arabian Peninsula, on one side, and the languages of the northern Ethio-Semitic branch (Geʿez, Tigrinya, Tigré), on the other, e.g. Classical Arabic *baytu-ki*, Syrian dialect *bēt-ik* 'your [fem. sg.] house)', Geʿez *məgbāˀe-ki* 'your [fem. sg.] refuge', Tigrinya *ʿof-ki* 'your [fem. sg.] bird'.

The *-(i)č/-(i)š/-c* isogloss includes some dialects of Iraq, Eastern Arabian dialects, Gulf Arabian dialects, and the Arabic dialects of Dhofar, Oman, and East Yemen, e.g. *ˀaxt-ič* 'your [fem. sg.] sister' (Johnstone 1967:I, 8; V); cf. also West Yemen *ˀabū-k/ˀabī-k* 'your [masc. sg.]/your [fem. sg.] father' (Behnstedt 1985:82–85) and East Yemen *ˀabū-š/ˀabū-č* 'your [fem. sg.] father'; Dhofar *gism-iš* 'your [fem. sg.] part'; *ṯmī-š* 'your [fem. sg.] mouth' (Rhodokanakis 1911:106; Diem 1973:36).

The forms -(V)*s*/-(V)*š*/-(V)*č* for the bound 2nd person feminine singular pronoun were already noted by the early Arabic philologists as a characteristic of the tribes Muḍar, Rabīʿa, and Tamīm (Brockelmann 1908:I, 309–310), and of the Yemenite dialects (→ *kaškaša/kaskasa*). The Arabic grammarians called these features *kaskasa Hawāzin, Bakr; kaškaša Tamīm, Rabīʿa, Muḍar; šanšanat al-Yaman* (Sallūm 1976:54–60, 83, 256; Retsö 2000:111–118).

The isogloss of the -(V)*š* form includes all the Modern South Arabian languages (Simeone-Senelle 1997:388): Mehri *abət-š* 'your [fem. sg.] house', *ġəwqay-š* 'he looked at you, saw you [fem. sg.]' (Johnstone 1987:xv–xviii), Ḥarsūsi *tə-š* 'for you [fem. sg.]' (Simeone-Senelle 1993:257), Soqoṭri *nhof-š* 'yourself [fem. sg.]', *teˀbar-š* 'they will come to you [fem. sg.]' (Naumkin-Porxomovskij 1981:22, 108–110).

This isogloss is also shared by the Ethio-Semitic languages of the southern branch:

Amharic *-(ə)š: bet-əš* 'your [fem. sg.] house' (Hudson 1997:462), Harari, East Gurage, Gafat, Soddo, Inor *-(a)š* (Wagner 1997:489–491; Gutt 1997:511; Hetzron 1997:540).

Other systems of deictic morphemes in the languages under consideration are either derived from the Common Semitic vocabulary or represent independent developments in particular languages.

3.2 *Verbal system*

Personal conjugated forms of the verb are closely connected to the corresponding personal pronoun systems in all of the languages and dialects, and may be defined as full, less full, and reduced according to the same principle. A full conjugation paradigm, including dual forms, is a feature of Classical Arabic and Modern South Arabian languages, with the corresponding exception of a more reduced system in Soqoṭri (Simeone-Senelle 1997:402). A less full conjugation paradigm, without dual forms, is found in Eastern Arabian and Yemenite dialects, and in the northern branch of the Ethio-Semitic languages (Geʿez, Tigrinya, Tigré), and also in several languages in the southern branch of the Ethio-Semitic languages (Soddo, Muher, Ezha, Inor; Kaye and Rosenhouse 1997:292–293; Gragg 1997:252; Kogan 1997:438; Raz 1997:452; Hetzron 1997:545). A reduced conjugation paradigm, without gender forms in the 2nd and 3rd person plural, is found in most modern Arabic dialects and languages of the southern branch of the Ethio-Semitic languages (Amharic, Argobba, Harari, East Gurage, Gafat; Kaye and Rosenhouse 1997:292; Hudson 1997:470–481; Wagner 1997:497; Gutt 1997:522–523; Hetzron 1997:545).

The opposition of two verbal stems, the suffixal conjugation stem (perfect) and the prefixal conjugation stem (imperfect), is found in all variants of Arabic and in all South Semitic languages. However, modern languages and dialects have created new forms and categories, with new ways of expression. Newer systems in modern languages demonstrate various evolutionary paths.

One of the most notable features distinguishing all the South Semitic languages from the other Semitic languages, including Classical Arabic and Arabic dialects, is the so-called *k*-perfect. In contrast to the Arabic forms for 1st person singular, 2nd person masculine and feminine singular *-tu/ta/ti* and plural *-tum/tunna*, the corresponding perfect forms of the South Semitic languages have as endings singular -*ku/ka/ki*, plural -*k*(V)*m*/*k*(V)*n*, or their phonetic reflexes.

Several groups of Yemenite dialects, however, form an exception to the general Arabic continuum of *t*-perfect. The Yemenite dialects of the southwestern mountain range and highland have the following perfect paradigm: 1st person singular *nazalk*, 2nd person masculine singular *nazalk*, 2nd person feminine singular *nazalki* (Reim, Banī Matar); 2nd person masculine plural *nazalkum* (Reim), *nazalkō* (Banī Matar), 2nd person feminine plural *nazalkun* (Reim), *nazalkayn* (Banī Matar) 'to descend, go down; to depart' (Karimov 1973:45). Some local dialects have a special *k*-perfect paradigm: 1st person singular *sirk(u)*; 2nd person masculine singular *sirk*, 2nd person feminine singular *sirč*, 2nd person masculine plural *sirku*, 2nd person feminine plural *sirkun* 'to go' (Prochazka 1974:439–441); or forms of 2nd person feminine singular with another phonological variant *katabš/katabč* 'you wrote' (Behnstedt 1985:116–118).

New data from Epigraphic South Arabian languages show that these languages, too, had *k*-perfect forms (Kogan and Korotajev 1997: 234). Ḥimyaritic also belongs to the *k*-perfect languages, according to the data from Arabic Yemenite authors (Belova 1996:60–62).

The South Semitic isogloss of *k*-perfect thus includes some dialects of Yemen. The Yemenite paradigms with 2nd person feminine singular *-(i)k(i)* or 2nd person feminine singular *-(i)č/š* show that some of them are not only part of the *k*-perfect isogloss but also of the sub-isogloss with 2nd person feminine singular *-č/š*, which includes only part of the South Semitic languages. Compare this form in Yemenite *sir-č* (Prochazka 1974:440–441); Mehri *səyə́r-š* (Johnstone 1987:xxx, 355) 'to go'; Amharic *näggär-š* 'to tell'; Argobba *säddäb-č(i)* (Hudson 1997:470) 'to insult'; Harari *sabar-ši* (Wagner 1997:497) 'to break'; East Gurage *masak-š* (Gutt 1997:521) 'to guide'; the Gafat form has the same inflection *-š/h*^{y} (Hetzron 1997:545).

Thus, in a small group of Yemenite Arabic dialects lie two isoglosses: the *k*-perfect isogloss for all the conjugated verb forms, characteristic of Epigraphic South Arabian languages, Ḥimyaritic, and the languages of the northern branch of the Ethio-Semitic languages (Geʿez, Tigrinya, Tigré); and the sub-isogloss for *č/š*-perfect for 2nd person feminine singular, including some subdialects of Yemen, all the Modern South Arabian languages, and the southern branch of the Ethio-Semitic languages.

On the whole, the verbal systems of Classical Arabic, Arabic dialects, and South Semitic languages have the Common Semitic paradigms of prefixal and suffixal conjugation. However, they differ in many respects with regard to the development of more differentiated categories of verbal tenses and moods and in their systems of deverbal nouns (participles, action nouns) and their inclusion in the verbal paradigm. It is difficult to find in the languages involved innovations constituting isoglosses common to Classical Arabic or Arabic dialects and to Ethio-Semitic languages. Each language develops its own innovations. Thus, it is characteristic of Modern Standard Arabic and modern Arabic dialects to develop analytical forms to express tense with auxiliary verbs and particles. Modern Standard Arabic undergoes a process of → participle verbalization and its inclusion into the paradigm of personal verb forms. New categories on the base of deverbal noun stems have developed in Ethio-Semitic languages, the so-called converb or gerund or gerundive, and the conjunctive. The corresponding processes of development of a particle system, development of auxiliary verbs, and inclusion of participles in the verb paradigm are also found in Modern South Arabian languages. However, each language uses its own means of expression for the new forms and categories, such as verbonominal forms, participles, and preverbs (Simeone-Senelle 1997:408–409).

The system of active/passive voices is present in all the languages under consideration, but the languages differ in the way in which the voices are expressed. Passive forms effected by internal vowel change, the so-called internal passive, are characteristic of Classical Arabic and Modern South Arabian languages. Compare Classical Arabic *katab-/ya-ktub-* 'to write', passive *kutib-/yu-ktab* 'to be written' with Mehri *rəkūz/yərūkəz* 'to straighten something, to stick something upright', passive *rəkēz/yərkōz; ṯəbur/yə-ṯūbər* 'to break [someone's leg]', passive *ṯībər/yəṯbōr* 'to be broken' (Johnstone 1987: xix, xxi–xxii, 324, 414–415; Simeone-Senelle 1997:397–398, 407).

Modern Arabic dialects represent different stages of preservation or loss of the internal passive and its replacement with external means. The internal passive forms are most frequently observed in the dialects of Yemen: *ḥad qutil hina?* 'was somebody killed here?' (Landberg 1913:67, 220), Qaṭar dialect *ʾuxiḏū* 'they were taken [prisoner]', *qaṣr yusammā* 'the palace named…' (Johnstone 1967:142). In the other dialects in the east and the west of the Arabic area, only a few cases of internal passive usage are preserved (Kaye and Rosenhouse 1997:294, 297–298). All modern Arabic dialects tend to express the passive by external means via derived verbal stems, so-called reflexives with preformatives *ta-/n-/-t-/-st-*.

A similar situation is found in Ethio-Semitic languages, where the passive is expressed by prefixes or reflexive stems *t*(V)-, (V)*st*-, cf. Geʿez *kadana* 'to cover', *ta-kadna/ta-kadana* 'to be covered'; *maḥara/məḥra* 'to show mercy, have mercy', *ʾasta-mḥara* 'to be merciful, be compassionate' (Dillmann 1907:141,151–159). In Tigrinya, Tigré, Amharic, Harari, and other southern Ethio-Semitic languages, passive forms are also formed on the basis of the prefix *t*(V)-, common for passive and reflexive (Kogan 1997: 436; Raz 1997:453; Hudson 1997:479; Wagner 1997:497–498; Gutt 1997:518, 524; Hetzron 1997:544).

In the Arabic dialects, one finds Moroccan *ləssəq* 'to glue, attach', *t-ləssəq* 'to be glued, attached'; *trək* 'to leave', *n-trək* 'to be left'; an innovated form *ttə-ḥrəq* 'to be burnt' (Kjamilev 1968:54–57); Algerian *n-bāʿ* ~ *t-bāʿ* 'to be sold' (< *b-y-ʿ* 'to sell'), *nə-ḍrəb* 'to be beaten' (< *ḍ-r-b* 'to beat') (Mishkurov 1982:58–59; cf. also Zavadovskij 1962:83–85; Rhodokanakis 1911:172–173).

Since the system of derived verbal forms is a Common Semitic feature, Arabic in general and all the South Semitic languages have this derivational verbal category. However, each language group has its own characteristics, which coincide only partially.

The derived stems in Classical Arabic and modern Arabic dialects are characterized by vocalic modification, by gemination of the second radical, or by prefixation and infixation. In Classical Arabic, the regular paradigm is represented: basic stem and derived Forms II-X of the verbal or nominal consonantal roots. Certain lexical or grammatical meanings are associated with each of the derived forms, changing the base meaning of the verbal or nominal root.

Modern Arabic dialects in general are characterized by reduction of the regular system of derived forms: usually the causative Form (IV or *ʾa*-stem) is superseded by the geminated Form (II or D-stem); regular formation of Forms III and VI with a long first vowel is reduced; derived Forms V, VII, VIII often function as passive/reflexive forms, as mentioned above. Reduction of short vowels both in base and derived forms shifts the expression of lexico-grammatical meanings to external affixation and root consonant gemination.

Prefixal patterns of derived forms are represented in Classical Arabic and in modern Arabic dialects, as well as in all South Semitic languages. The distribution of other patterns, however, exhibits more variation than common features.

The prefixal *n*-form is found in Classical Arabic and modern Arabic dialects (Form VII) and in Modern South Arabian languages with the following meanings: intransitive, middle, reflexive, reciprocal (Simeone-Senelle 1997: 401). This model of derivation is absent from the other South Semitic languages.

The geminated verbal form (Arabic Form II/D-stem) with the meanings of intensity, causative, factitivity, etc., as a member of the system of derived verbal forms, is present only in Classical Arabic and modern Arabic dialects. In Epigraphic South Arabian languages, this form may only be hypothesized (Beeston 1984:12–13); geminated forms are not found in Modern South Arabian languages (Simeone-Senelle 1997:397–398). In Ethio-Semitic languages, geminated forms are not part of the system of derived forms but represent one of the base lexical models.

The class of verbs with stem vowel *ā* after the first radical is traditionally considered an important diagnostic feature uniting Arabic with South Semitic languages (Brockelmann 1908:I, 511–513; Hecker 1982:14; Faber 1997:12). The presence of such a stem in Epigraphic South Arabian also remains hypothetical (Beeston 1984:12–13). As a separate derivational stem with the meaning of factitive or causative, this pattern is found in Modern South Arabian: Mehri *rēkəb* 'to ride' – *ā*-stem *rōkəb* 'to put [a pot] on the fire' (Simeone-Senelle 1997:398–399). In Ethio-Semitic languages, *ā*-stems, like geminated ones, do not participate in the system of derived forms but rather exist as one of the base lexical models (Gragg 1997:252–254; Kogan 1997:436–437; Raz 1997:452–453; Hudson 1997:468, 477–479; Wagner 1997:494–495; Gutt 1997:516–517; Hetzron 1997:544).

The derivational model of the *ā*-pattern type is indeed characteristic of all known South Semitic languages and unites them with Classical Arabic and modern Arabic dialects, but this model has different grammatical and lexical functions in the languages of the Arabian Peninsula and in Ethio-Semitic languages. The same can be said about the D-stem (Form II) derivational model in Arabic and Ethio-Semitic languages in general, which is also absent from the system of derived forms in Modern South Arabian languages.

One unique pattern among the derived forms in Arabic (in general) and the South Semitic languages is formed by stem vowel alternation accompanied by gemination of the third radical consonant. Synchronically, the finite forms of this model actually exhibit geminated third radical. However, as verbal noun or imperative, such forms demonstrate reduplication of the stem. These are the so-called Form IX and its variant Form XI, which are related to the adjectives of color and bodily characteristics of the pattern *ʾafʿal* (masc.)/*faʿlāʾ* (fem.), e.g. Arabic *ʾaḥmar* 'red' – *iḥmarr-a*/*ya-ḥmarr-u* (Form IX)/*iḥmārra*/*ya-ḥmārr-u* (Form XI) 'to redden, become red' (imperative *iḥmarir*, verbal noun *iḥmirār-*). Only in Modern South Arabian is a separate model of verbal stem specialized for color adjectives found, formed by reduplication of the third radical consonant: Mehri *ʾōfər* 'red' – *ʾāfērōr*/*yāfērōr* 'to go red' (Johnstone 1987:14); Soqoṭri *aʾférir* 'to go red' < *ʾáfer* 'red' (Leslau 1936:320). Cf. also Mehri *əwbēnūn* 'to go white' < *awbōn* (< **l-b-n*) 'white' (Johnstone 1987:251) and Soqoṭri *yalbínen* (imperfect) 'to become white' < *libehon* 'white' (Leslau 1936:228). Unlike Classical Arabic and Arabic dialects, in Modern South Arabian languages this verbal stem is associated only with adjectives of color, not with adjectives of bodily features. Compare Mehri *ḥəwīrūr*/*yənḥīrūr* 'to turn black' < *ḥōwər* 'black' (Johnstone 1987:125), *həźīrūr* 'to turn green, pale' < *həźáwr* 'green, yellow' (Johnstone 1987:163), but *ʾáywer*/*yāwōr* 'to be, to go blind' < *ʾāwēr* 'blind' (Johnstone 1987:37), where Arabic has *iʿwarr-a/yaʿwarr-u* 'to be, become one-eyed' < *ʾa-ʿwar-u* 'one-eyed'.

Only Classical Arabic, Modern South Arabian languages, and Geʿez preserve two different patterns for the basic stem according to the transitivity or intransitivity of its meaning. In Classical Arabic this distinction is not always regular, but it can still be traced. Compare active verbs: Classical Arabic *ṯabara/ya-ṯbur-u* 'to push off, disturb, hinder', Geʿez *sabar-a/ yə-sbər* (jussive 3rd pers. masc. sg.) 'to break' (Leslau 1987:485), Mehri *ṯəbūr/yə-ṯūbər* 'to break' (Johnstone 1987:414); middle verbs: Classical Arabic *labis-a/ya-lbas-u*, Geʿez *labs-a/ yə-lbäs* (jussive 3rd pers. masc. sg.) 'to dress oneself', Mehri *lībəs/yə-wbōs* (with *w* < **l*) 'to wear, put on' (Johnstone 1987:251).

In most Modern Arabic dialects, the distinction between the two patterns of the base verbal stem has disappeared as a result of short vowel reduction and disappearance of short vowels in open unstressed syllables.

3.3 *Nominal system*

The nominal morphology of Arabic and the South Semitic languages in general is characterized by grammatical categories of Common Semitic origin. However, the categories of case, gender, and number in the languages under consideration represent various stages of change and reduction. The category of definiteness/indefiniteness is not represented in all languages, but in each language where it is present the systems of definiteness markers (articles) are the result of relatively recent independent development in this particular language.

The category of case is represented by a full paradigm only in Classical Arabic; existence of a case system is also assumed in Epigraphic South Arabian languages, with some evidence of reduction (Beeston 1984:32). Geʿez retains only one case ending, formed by suffixation of *-ä* which is used both in accusative and possessive constructions, and may be called polyfunctional (Gragg 1997:248–249; Dillmann 1907:320–323).

In modern Arabic dialects and other South Semitic languages, relations between words are expressed analytically by means of relational words, prepositions, and word order.

The category of gender is represented both in Arabic (in general) and in South Semitic languages, but the means of expression for this category differ considerably. In Classical Arabic and Arabic dialects, the feminine gender is mostly expressed by suffixes: masc. Ø vs. fem. *-(a)t* ~ *(a)h*. The same suffixal means of expression is characteristic of Epigraphic South Arabian languages, Modern South Arabian languages, and Geʿez, in which names of human beings and animals with lexical expression of gender form an exception. However, in Arabic (both Classical and dialects), Soqoṭri, Geʿez, and Tigrinya, a class of adjectives and participles is found in which gender forms are marked by internal vocalic change or by employing different patterns. In Arabic, these are the adjectives of color and bodily characteristics: *ʾaṣfar-/ṣafrāʾ-* (pl. *ṣufr-*) 'yellow'; in Soqoṭri *sígded/sígdíd* 'avaricious', *ʿáṯmham/ʿaṯmhim* 'generous' (Leslau 1936:281, 305; Simeone-Senelle 1997:390); in Geʿez *ṭäbib/ṭäbbab* 'wise' (Leslau 1987:585), *ḥäddis/ḥäddas* 'new' (Gragg 1997:249–250); Tigrinya *ṣällim/ṣällam* 'black' (Kogan 1997:434).

In most modern Ethio-Semitic languages, grammatical gender is determined by agreement categories of the verb, pronoun, and definite article. In Tigré, suffixal gender markers are found for participles: sg. Ø*:* (-V)*t*, pl. *-am* (masc.)*/-āt* (fem.) (Raz 1997:448–450).

The systems of number forms also differ in the languages under consideration both in degree of paradigm completeness and means of expression.

The number paradigm including dual is represented in Classical Arabic, Epigraphic South Arabian languages, and Modern South Arabian languages. Dual forms are expressed externally, e.g. Classical Arabic *yawm-āni/yawm-ayni*; Qaṭabanian *ywm-myw* 'two days' (Kogan and Korotajev 1997:228); Mehri *ġaj-i śeruh*, Šxauri (Jibbāli) *ġaíg-i ṭro*, Soqoṭri *tra ʿaig-i* 'two men'; Mehri *junít-i ṯirît*, Soqoṭri *tri gunīt-i* 'two bags' (Müller 1907:9, 38). Dual forms are also accompanied by the numeral 'two'.

In modern Arabic dialects, dual forms are partially preserved in a specific class of nouns, mostly to denote paired body parts or time periods: Iraqi *rijl-ēn*, Egyptian *rigl-ēn* 'two legs', Iraqi *dagigt-ēn* 'two minutes', *sent-ēn* 'two years' (→ pseudodual).

Ethio-Semitic languages have a paradigm of two → number forms, singular and plural. The languages under consideration are characterized by two kinds of plural formation: the so-called external (or 'sound') plural, marked by suffixes, and the so-called internal ('broken') plural, marked by internal vocalic change or by a combination of external and internal formants.

Traditionally, the broken plural is considered one of the most important features uniting Arabic with the South Semitic languages. This thesis, however, needs qualification. In the southern branch of Ethio-Semitic languages, all the languages (Amharic and others) have only one system of plurals, that of external plural forms. Thus, not all South Semitic languages participate in the isogloss of the broken plural.

In all languages concerned, the external plural is represented by two endings, which descend from Common Semitic and do not constitute any distinguishing feature of the South Semitic languages. The most common and widely distributed ending, in Arabic as well as in South Semitic languages, is *-(ā)t* and its corresponding variants: in Epigraphic South Semitic languages -(V)*t*; in Geʿez, Tigrinya, Tigré *-(ā)t/-*(*o*)*tat*; in Modern South Arabian languages -(V)*t*/-(V)*tə*(*n*); in the southern branch of the Ethio-Semitic languages (Amharic and others) *-oččl-(a)čl-ča*. More restricted in its functions is the external plural marker for masculine nouns, adjectives, and participles. In Classical Arabic, this marker has the form *-ūn/-īn*; modern Arabic dialects have *-īn*; Epigraphic South Arabian languages -(V)*n*; Soqoṭri *-īn/-ihon*; and Geʿez and Tigrinya *-an*, Tigré *-ām*.

Broken plurals are a productive way of plural formation in Classical Arabic and modern Arabic dialects, in Epigraphic South Arabian languages, in Modern South Arabian languages, and in Ethio-Semitic languages of the northern branch (Geʿez, Tigrinya, and Tigré; Diakonoff 1988:63–66; Ratcliffe 1998:117–149).

Classical Arabic has the largest number of broken plural patterns (more than thirty). Many singular nouns can have several patterns of plural formation. Modern Arabic dialects in general have a more limited number of patterns, although some new patterns are present in certain languages. Most singular nouns have one corresponding plural form.

In South Semitic languages, the number of patterns usually does not exceed ten. Compare Classical Arabic *faras/furūs ~ ʾafrās ~ fursān*, modern Arabic dialects (V)*frās ~ fursān*, Sabean *frs/ʾfrs*, Geʿez *faras/ʾafrās*, Tigrinya *färäs/ʾafrās* (~ *ʾafrus*) 'horse, mare'.

The multitude of plural patterns with respect to one singular pattern in Classical Arabic is partially conditioned by pattern variation in ancient dialects, later collected and formalized in the grammar of Classical Arabic. In South Semitic languages, such cases are not so widespread. Compare Classical Arabic *baḥr/buḥūr- ~ biḥār- ~ ʾabḥur-*, modern Arabic dialects: pl. *buḥūr ~ b(i)ḥār*, Sabean *bḥr/ʾbḥr*, Geʿez *bāḥr/ʾabḥərt ~ bāḥrāt*, Ḥarsūsi pl. *beḥéwr* 'sea; river'.

As a general rule, in Arabic as a whole as well as in those South Semitic languages which use the broken plural, there is a certain correlation between particular patterns and the consonantal-vocalic structure of the singular nominal pattern: bi-/triconsonantal structure or quadriconsonantal structure (Ratcliffe 1998:166–203), e.g. Classical Arabic *kalb/kilāb ~ ʾaklub*, all modern Arabic dialects pl. *k*(V)*lāb*, Modern South Arabian languages pl. *kel(ōb) ~ ha-kwebet*, Geʿez pl. *aklāb ~ ʾakləbt*, also *kalab-āt*, Tigrinya pl. *akaləbt*, Tigré pl. *ʾaklub* 'dog, wolf'. For quadriconsonantal stems (derived and nonderived): Arabic (all) *maṣnaʿ/maṣāniʿ* 'workshop; building; palace'; Epigraphic South Arabian languages *mḥfd/mḥfdt* 'tower'; Modern South Arabian languages *mənxal/mənōxəl* 'sieve'; Geʿez *maṣḥaf/maṣāḥeft* 'book'; Tigrinya *känfär/känafər ~ känäffər* 'lip'. This feature is found with corresponding variations in all the languages under consideration.

In spite of the large number of broken plural patterns in Classical Arabic and Modern Arabic dialects, a number of isoglosses may be established on the basis of the most widespread patterns. These isoglosses connect Arabic patterns with the corresponding major patterns of various South Semitic languages (Ratcliffe 1998:204–213). The most widespread of these patterns is *ʾa-qtāl*, which mostly represents an isogloss that covers Arabic (in general), Epigraphic South Arabian (where this pattern is the most frequent one), and North Ethio-Semitic languages, e.g. Arabic Syrian and Yemenite dialects *ṯōr/ʾaṯwār*, Sabean *ṯwr ~ ṯr/ʾṯwr*, Geʿez *sōr/ʾäswār*, Tigrinya *sor/aswar*, Ḥarsūsi *ṯawr/heṯwéret* (a rare pattern in Modern South Arabian) 'bull'; Arabic *sinn/ʾasnān- ~ ʾasunn*, Geʿez *sənn/ʾasnān ~ sənān*, Tigrinya *sənni/ašnan* ‘tooth’.

The *qutūl* pattern forms a narrower isogloss, including Arabic (in general), assumed forms in Epigraphic South Arabian languages, and Modern South Arabian, e.g. Arabic *qarn/qurūn*, Mehri *qōn/qərūn*, Ḥarsūsi *qōn/qərōn*, Jibbali *qun/qérun*, Soqoṭri *qan/qirihon* ‘horn’ (Militarev and Kogan 2000:151–152). It may be assumed that some patterns found in Ḥarsūsi correspond etymologically to the Arabic pattern *qatūl* or *qitāl*; Arabic *nimr ~ namir/numūr ~ nimār*, Modern Arabic dialects pl. *n*(V)*mūr*[1] ~ *n*(V)*mār*, Ḥarsūsi *nemr/nemáwr* ‘tiger, panther’; Arabic (Classical and dialects) *ḏ̣ahr ~ ẓahr ~ ḍahr/ḏ̣uhūr/ẓ*(V)*hūr/ḍ*(V)*hūr*, Ḥarsūsi *ḏ̣ahr/ḏ̣eháwwer* ‘back’.

In Arabic Yemenite dialects, interesting plural patterns are found for the substantives of the structure *qatīl/qitāl: ṭarīg/ṭurwag ~ ṭirwag* ‘road, way’; *bilād/bilwid ~ belawwid* ‘country, land’ (Diem 1979:29–30). It is possible that these patterns go back to a local substrate whose traces are also found in Epigraphic South Arabian languages, cf. Sabean pl. *ʿrwn* ‘wooded land’ (Beeston a.o. 1982:19). The pattern might be connected with Classical Arabic *ʿarīn/ʿurun* ‘wood, forest’; cf. also Sabean pl. *blwd* ‘settlement’ (Beeston a.o. 1982:28).

A number of peripheral forms serve as evidence for the existence of certain formants (elements) or tendencies common to the Arabic periphery and some South Semitic languages. One such pattern is a special paradigm of masculine and feminine plural forms for the adjectives of color and bodily characteristics in the Arabic dialects of South Arabia: masc. pl. *ʿujm-ān*, fem. pl. *ʿujm* ‘dumb, mute’; masc. pl. *ḥumr-ān*, fem. pl. *ḥumr* ‘red’; etc. (Landberg 1901:23–24). The suffix *-ān* is relatively widespread in Classical Arabic and other Arabic dialects; it corresponds to the external suffix of the masculine plural of adjectives and participles in Geʿez.

Also worth mentioning are some patterns for plural formation in the Arabic dialects and some Ethio-Semitic languages that are historically in contact with the languages of East and Central Africa.

In the Arabic dialects of Upper Egypt, Sudan, Nigeria, and the region of Lake Chad, plural forms are found with gemination of the third radical, e.g. South Egyptian *bnitta* (sg. *bint* ‘girl’), *ibṛakk* (sg. *birka* ‘pond’); Sudanese Arabic *usudda* (sg. *asad* ‘lion’); Nigerian Arabic *dugunne* (sg. *digin* ‘beard; chin’). Some languages of the southern branch of Ethio-Semitic along with the regular pattern of plural formation use reduplication of the last consonant accompanied by a vowel, e.g. Amharic *wəndəmam-ač* (< sg. *wəndəm* ‘brother’), East Gurage *alagāgo* (< sg. *alaga* ‘stranger’), Soddo *gurzazä* (< sg. *gurz* ‘old man’).

Thus, a peripheral local isogloss of gemination/reduplication encompasses Arabic dialects, which are a part of the large broken plural isogloss, and part of the Ethio-Semitic languages, which as a whole belong to the sound/external plural isogloss.

Comparing the paradigms of external and internal plural forms in general in the languages under consideration, it should be noted that every language has a similar distribution scheme for the forms of the first and second types. The scheme is asymmetric for external and internal forms, as well as for suffixal forms within the external paradigm. The paradigm of suffixal forms has the most restricted distribution, including masculine, mostly derived, nouns denoting persons, and also adjectives and participles denoting persons or agreeing with masculine person denoting nouns: Arabic *-ūn/-īn*, Soqoṭri -(V)*n*, Ethio-Semitic *-ān*. The external plural *-āt/*-(V)*t* paradigm includes not only feminine but also masculine nouns, both animate and inanimate. As the Ethio-Semitic data show, the function of the plural marker *-āt* has a tendency to expand and take over the functions of other markers.

Broken plural paradigms overlap with external plural paradigms, i.e., they include names of different lexical-grammatical categories, among them those that can form plural forms using suffixes, e.g. in Arabic: *nabiyy-īn* (Q. 2/61) and, in the same sura, *ʾanbiyāʾ* (Q. 2/91) ‘prophets’. Comparative studies of the external/internal plural in the Semitic languages show that in the languages where both ways of plural formation coexist, “the internal plural is either the obligatory or at least the only productive plural for underived, unmarked nouns of three or fewer consonants (stem shapes CvC, CvCC, CvCvC), while the external plural is generally obligatory for productively derived nouns such as participles and verbal nouns” (Ratcliffe 1998:219–242). In other words, the choice between external and internal plural is determined not by the noun's lexical class but rather by the consonant and syllable structure of its stem. Still, the person category retains the preference for the suffixal form.

It should be noted, however, that in some peripheral Arabic dialects, which coexist with languages of a different grammatical type, there is a tendency for the external plural paradigm to somewhat expand its usage. For example, in the Arabic dialect of Bukhara: *wazīr/wazīr-īn* 'minister, vizier'; *axū/uxaw ~ uxw-īn* 'brother'; *qamar/qamar-āt* 'month; moon', or: *ʿašrāt (ə)ḥmār-āt* 'ten donkeys'; the Arabic dialect of Balkh (Afghanistan): pl. *kitāb-āt-ak* 'your books'.

New loanwords in Modern Standard Arabic and in modern Arabic dialects, when their syllable structure does not fit into the patterns of the broken plural, also form their plural forms by using the suffix *-āt*. Arabic *duktūr/dakātir(a)* 'physician' fits in an existing pattern of broken plural, but other loanwords use a more universal means of plural formation, e.g. *lūrd/lurd-āt* 'lord', *brins/brins-āt* 'prince'. The choice of internal or external plural for loanwords designating inanimate objects is also determined by their syllable structure, e.g. *tāksī/taksi-h-āt ~ taksi-y-āt ~ takāsī* 'taxi', but for a complex stem only an external plural is possible, e.g. *kumbyūtar/kumbyūter-āt* 'computer'.

The asymmetry in the distribution of external and internal plural forms is present in all the languages where both external and internal plural formation is possible; when the 'external' plural formation is chosen, there is a tendency to prefer the Common Semitic formant *-āt*, which loses its connection to grammatical gender and becomes a universal plural marker. At the same time, it should be stressed that, in Modern Standard Arabic and modern Arabic dialects, the internal plural remains productive and is used for all nominal parts of speech with appropriate syllable structure.

Derived names (nouns, adjectives, and participles) are formed in Arabic as well as in South Semitic languages by means of Common Semitic morphemes and internal derivation models. Each language also has its own innovations. It is hard to distinguish any common isoglosses in the nominal derivation system which are only applicable to the languages under consideration or groups within them.

With all the great variation in forms and means of nominal derivation, it is nevertheless worth noting that derivation of adjectives and verbal nouns has a certain parallelism in Arabic (in general) and Soqoṭri. It should be stressed that each of these languages uses its own grammatical means, but these means distinguish in both languages a separate lexical-grammatical category: adjectives of color and bodily feature, e.g. in Arabic: masc. sg. *ʾarmad*/fem. sg. *ramdā* 'ashy pale', Soqoṭri masc. sg. *ráḥmam*/fem. sg. *ráḥmim* 'id.'; Arabic masc. sg. *ʾadrad*/fem. sg. *dardā* 'having rotten teeth'; Soqoṭri masc. sg. *gimses*/fem. sg. *gimsis* 'having crooked teeth'. For such adjectives Arabic uses a pattern of internal word derivation; Soqoṭri uses reduplication of the third consonant of the root.

The systems of numerals, deictics, adverbs, and particles (prepositions, conjunctions, and modality markers) are also represented in the languages under consideration partly by Common Semitic stems, partly by innovations particular to each language. It is difficult to distinguish any common or local isoglosses among the innovations. One can only speak of a common tendency characteristic of all languages that lose synthetic means of expressing syntactic relations between words. Modern Arabic dialects as well as all modern South Semitic languages exhibit an increase in the functional role of pronouns, particles, auxiliary verbs, and word order.

The system of definiteness/indefiniteness marking is most fully represented in Classical Arabic and in Epigraphic South Arabian languages. In these languages, three states of a noun are distinguished, each of them having its own marker: indefinite (or absolute) state, determinate (or definite) state, and construct state. The construct state in Classical Arabic and Epigraphic languages has a Common Semitic way of expression: positional – head+modifier, in which the head occurs without the postpositional state marker, while in Classical Arabic, the modifier is put in the Genitive case.

Modern Arabic dialects in general also retain the system of the three states, but the absence of case endings makes the role of the word order more important in the dialects. Apart from this, in most dialects new analytical ways of expressing genitive relations have developed by means of special genitive markers, e.g. Iraqi Arabic *sayyārat il-ḥukūme* or *sayyāra mālt il-ḥukūme* 'the government car'; Syrian Arabic *il-ʾalam tabaʿ il-walad* 'the boy's pencil' (→ analytical genitive).

While retaining a common system of three states, Arabic (Classical) and Epigraphic South Arabian languages use different markers of definiteness/indefiniteness: Classical Arabic and modern Arabic dialects are characterized by a system in which the article (V)*l-* is the prepositional marker of definiteness, while the marker of indefiniteness is in Classical Arabic the postpositional *-n*, and in the dialects Ø. In Epigraphic South Arabian languages, both markers (definiteness – *-n*, indefiniteness – *-m*) are postpositional.

In Modern South Arabian languages (Mehri, Ḥarsūsi, Jibbāli), traces of a definite article in preposition to a name are found.

Modern Ethio-Semitic languages develop their own systems of definiteness markers. In the southern branch of Ethio-Semitic (Amharic, Harari, East Gurage, and others), the new system of definiteness markers is developed on the basis of deictic elements. The definiteness markers are postpositional; they carry an additional function of distinguishing grammatical gender.

Syntactic systems in the languages under consideration are determined both by their grammatical structure and by possible substrate influences. In general we can only mention the most basic isoglosses for word order in a verb-centric statement: in Classical Arabic and Epigraphic South Arabian languages, the basic word order is VSO. In Geʿez, possible variations are VSO, SVO, SOV. Modern Ethio-Semitic languages have the basic SOV word order.

Agreement for adjectives, pronouns, and verbs in gender and number is preserved in Classical Arabic, Epigraphic languages, and Modern South Arabian languages and in modern Arabic dialects. In modern Ethio-Semitic languages, especially in their southern branch, agreement is limited to verbs and pronouns.

4. VOCABULARY/LEXICON

The lexicon of Arabic and the South Semitic languages presents a complicated picture. The basic vocabulary (the basic root inventory) in all these languages goes back to the Common Semitic vocabulary.

Naturally, some lexical areal isoglosses unite Classical Arabic and Arabic dialects with the South Semitic languages of Arabia and Ethiopia. In most cases, such isoglosses are evidence of the common cultural history of the speakers of these languages.

Part of the areal Arabic/South Semitic common vocabulary consists of terms and names related to the most ancient traditional economic activities of the South Arabian population. Agricultural terminology belongs to this group of lexical isoglosses: names of agricultural activities, seasons, types of irrigation systems, relief features, and weather, e.g. Classical Arabic *jirbat*, Yemenite *girba* 'plowed or sown field', Sabean *grb* 'to lay out fields in terraces', Harari *gäräb* 'section', *gäräban gäräb* 'system of division of farmland'; or Yemenite, Ḥimyaritic *ṣirāb*/*ṣarāb* 'harvest, harvest season', *ṣrb* 'to harvest a crop by cutting', Sabean *ṣrb* 'harvest, harvest season', Mehri *ṣáyrəb* 'late autumn', Geʿez *ṣrb* 'to cut, to hew trees', Gafat *ṣrb* 'to prune trees'.

Another areal isogloss is formed by terms for building and construction, known since the earliest period of South Arabian civilization, e.g. Classical Arabic < Yemenite *ḥisn* ~ *huṣn* 'protected place, fortress', Sabean *ḥṣn* 'to take under protection', Jibbāli *oḥóṣun* 'to build, to fortify', *ḥeṣn* 'castle', Soqoṭri *ḥoṣon*, Geʿez *ḥəṣn* 'fortress, castle'; Classical Arabic *hajar* (< Southern Arabia) 'town', Ḥimyaritic, Sabean *hgr*/*hgrt* 'town, populated locality', Geʿez, Tigrinya *hägär*, Amharic *agär*, Gafat *agärä* 'populated locality, land, district; town' (Belova 1988:30–31).

The verb *hll* 'to be, exist' unites Yemenite dialects and Ḥimyaritic with the Ethio-Semitic languages. Modern Yemenite dialects preserve a relic of this verb as a fixed form *halla*. This verb is used in a restricted set of contexts of 'presence/absence', e.g. *halla u mahalluš məḥammad?* 'is Muḥammad here or not?', or in a Ḥimyaritic utterance cited by al-Hamdānī: *dw hl qylʾ* = *laysa bi-malikin* 'is not a king' (*ʾIklīl* II, 352–353). In Ethio-Semitic languages, the corresponding words are: Geʿez *halläwä*/*hallo*, Tigrinya *ʾallo*, Tigré *halla*, Amharic *allä*, Argobba *halla*, Harari *ḥal*, Gurage *ala* 'to be, to exist, to be available' (Belova 1996:69–70).

The use of lexical data to establish genetic relationships between Arabic and other Semitic languages developed only recently. At present, a glottochronological method is developed for genetic classification of Semitic languages. The method is based on Swadesh's lexicostatistics, which uses the so-called 100-word list of basic lexicon. This method was modified by Sergei Starostin and applied by Alexander Militarev to establish an absolute chronology of divergence for Semitic languages. According to Militarev's calculation of the degree of preservation of the basic vocabulary in the Semitic languages known to us, the genetic relationships among the languages involved here are as follows. The first division of Proto-Semitic into two large branches of North Semitic and South Semitic languages took place in the second half of the 5th millennium B.C.E. South Semitic is the ancestor of the Modern South Arabian languages, which diverge into separate languages known to us in the period between the middle of the 1st millennium B.C.E. and the middle of the 1st millennium C.E. The ancestors of Ethio-Semitic and Arabic diverge much later (beginning of the 3rd millennium B.C.E.), splitting off the common Northwest Semitic branch. In its turn, the division of Proto-Ethio-Semitic, giving North Ethiopian and South Ethiopian branches, occurs in the first half of the 1st millennium B.C.E. Epigraphic South Arabian is not present in this scheme, since it does not have a whole list of basic words, necessary for the method to apply (Militarev and Kogan 2000:xxxix–xlii).

The scheme of genetic relationships between Arabic and the South Semitic languages built on the basis of vocabulary differs considerably from the classification of Semitic languages based on shared morphological innovations. In this second classification, Arabic remains in the central position within West Semitic/Central Semitic, but forms an opposition to the Northwest Semitic (Ugaritic, Canaanite, Aramaic). The classification of South Semitic includes the Eastern group (Modern South Arabian) and the Western Group (Epigraphic South Arabian and Ethiopian Semitic; Faber 1997:5–13).

The glottochronological classification establishes temporal parameters of the genetic relationships among languages, while the morphological classification and the system of isoglosses reflects to a greater degree their geographical parameters and their historical contacts.

The problem of genetic proximity of Semitic languages is connected in a natural way to the problem of their geographical origin and the problem of the movements of their speakers in time and space.

Considering the fact that the ancestors of the Ethio-Semitic languages have moved to East Africa from the territory of South Arabia, South Arabia may be assumed to have been the center and origin of all common South Semitic isoglosses. Spreading north along the Arabian Peninsula, the common isoglosses also encompass part of the Arabic dialects. From the historical point of view, Arabic should not be viewed as homogeneous; not only does Classical Arabic retain archaic features of phonetics and grammar, but so-called Neo-Arabic is also represented by dialects with varying degrees of innovation and archaic features. It should be enough to mention the causative-factitive stem in the peripheral Ḥassāniyya dialect of Mauritania with the prefix *sa-*: *sa-kḥal/i-sa-kḥal* (participle *mu-sa-kḥal*) 'to make black' (< *akḥal* 'black'), *sa-gbal* 'to go south' (cf. Arabic *ʾa-šraq-a* 'to go east'), etc. The prefix is found with this function in a number of Epigraphic South Arabian languages (Minean, Qatabanian): *s-ʿrb/y-s-ʿrb* 'to offer'. The same prefix remains as a relic in some verbal stems of Arabic: *sajar-a* 'he filled the jar' (cf. *jarrat-* 'jar, vessel'); *sadam-a* 'he plugged, corked up' (cf. *d-m-m* 'to plug, to cork up'). The features of the dialects of central and western Yemen are so peculiar that it is hard to decide whether they reflect South Arabian substrate (Diem 1979:12–80) or are remains of ancient South Arabian languages (such as Ḥimyaritic), on which a later Arabic superstratum has been imposed.

Another approach to lexical data for the purpose of determining the place of Arabic among the Semitic languages has been attempted in a comparative lexical study of Qurʾānic Arabic. A comparison between the lexical corpus of the *Qurʾān* and the vocabulary of known Old Semitic languages (Akkadian, Ugaritic, Phoenician, Ancient Hebrew, Aramaic, Syrian) and South Semitic languages (Geʿez and Epigraphic South Arabian) has established that Qurʾānic Arabic shares cognates with Northwest Semitic (40.44%) and with South Semitic (30.26%) (Zammit 2002:526–563, 586–587). Note, however, that lexical material of only two South Semitic languages was taken into account and that the known vocabulary of Epigraphic South Arabian languages is limited. Further studies in the area of comparative Semitic lexicology on the basis of established written texts could also contribute to the solution of the problem of relations between Arabic and other Semitic languages.

5. CONCLUSION

The following common isoglosses have been established: the isogloss of '*f*-languages', which unites Arabic in general and all the South Semitic languages, in contrast to all the other Semitic languages; the isogloss of the first root consonant *w*-, which corresponds to the first root consonant *y*- in the other Semitic languages; and the isogloss of the derived verbal stem with the long vowel *ā* after the first consonant of the stem, which is unique to the languages being considered here.

The other isoglosses are not shared by all the languages involved. The most 'prominent' isogloss of the broken plural does not encompass the Ethio-Semitic languages of the South branch. The isogloss of *k*-perfect, present in all the South Semitic languages, encompasses only a limited group of Yemenite dialects. The isogloss of the suffixed pronoun of the 2nd person singular feminine -(V)*č*/-(V)*š* unites the Arabic dialects of East Arabia and eastern Yemen with the Modern South Arabian languages and with the southern branch of Ethio-Semitic languages. A variant of this isogloss is the conjugated form of the 2nd person feminine singular -(V)*č*/-(V)*š* in the languages with *k*-perfect, but this isogloss does not spread over Epigraphic South Arabian languages and North Ethio-Semitic.

The distribution of the other local phonetic and morphological isoglosses discussed earlier in the text seems to be even narrower.

Most local isoglosses intersect but do not coincide completely. With respect to the Arabic language continuum, it is worth mentioning that a cluster of general basic local isoglosses is concentrated only in a limited group of Yemenite dialects. This cluster unites them with Modern South Arabian languages and with the northern branch of the Ethio-Semitic languages. Classical Arabic has more features distinguishing it from the South Semitic languages. This is why Zaborski says, contrasting this to a vast and widely divergent dialect continuum, "The Arabic dialect group has an intermediate position between North West Semitic and South Arabian" (1991:365). Indeed, irrespective of one's classificatory principles, Arabic in general occupies a central position among the other Semitic languages.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Hamdānī, *ʾIklīl*, II = ʾAbū Muḥammad al-Ḥasan ibn ʾAḥmad ibn Yaʿqūb al-Hamdānī, *Kitāb al-ʾiklīl*, II. Ed. Muḥammad ibn ʿAlī al-ʾAkwaʿ. Baghdad, 1980.
——, *ʾIklīl*, VIII = ʾAbū Muḥammad al-Ḥasan ibn ʾAḥmad ibn Yaʿqūb al-Hamdānī, *Kitāb al-ʾiklīl*, VIII. Ed. Anastās al-Karmālī. Baghdad, 1931.
——, *ʾIklīl*, X = ʾAbū Muḥammad al-Ḥasan ibn ʾAḥmad ibn Yaʿqūb al-Hamdānī, *Kitāb al-ʾiklīl*, X. Ed. Muḥyī d-Dīn al-Xāṭib. Cairo, 1368 A.H.
Našwān, *Šams al-ʿulūm* = Našwān ibn Saʿīd al-Ḥimyarī, *Muntaxabāt fī ʾaxbār al-Yaman min Kitāb šams ʿulūm al-ʿArab wa-dawāʾ kalām al-ʿArab min kulūm*. Ed. ʿAẓīm ad-Dīn ʾAḥmad. 2nd ed. Damascus, 1981.
Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*, II. Ed. Būlāq, 1316 A.H.

SECONDARY SOURCES

Beeston, Alfred F.L. 1984. *Sabaic grammar*. Manchester: Manchester University Press.
——, M.A. Ghul, W.W. Müller, and J. Ryckmans. 1982. *Sabaic dictionary (English-French-Arabic)*. Louvain-la-Neuve: Peeters; Beirut: Librairie de Liban.
Behnstedt, Peter. 1985. *Die nordjemenitischen Dialekte. I. Atlas*. Wiesbaden: L. Reichert.
Belova, Anna G. 1988. "The position of Himyaritic within the South Semitic group of Semitic languages (Yemenite-Ethiopian isoglosses)". *Proceedings of the XIXth International Congress of Ethiopian Studies, Moscow, 26–29 August 1986*, 28–24. Moscow: Nauka.
——. 1996. *Himyaritskij jazyk (arealnyje issledovanija k istorii arabskogo jazyka)* [The Ḥimyaritic language (areal studies on the history of Arabic)]. Moscow: Nauka.
Brockelmann, Carl. 1908–1913. *Grundriss der vergleichenden Grammatik der semitischen Sprachen*. 2 vols. Berlin. (Repr., Hildesheim: G. Olms, 1966.)
Diakonoff, Igor M. 1988. *Afrasian languages*. Moscow: Nauka.
Diem, Werner. 1973. *Skizzen jemenitischer Dialekte*. Beirut: Orient Institut der Deutschen Morgenländischen Gesellschaft and Steiner Verlag.
——. 1979. "Studien zur Frage des Substrats im Arabischen". *Der Islam* 56.12–80.
Dillmann, August. 1907. *Ethiopic grammar*. 2nd ed. London: Williams and Norgate. (English trans. of German *Grammatik der äthiopischen Sprache*, 2nd rev. ed. Carl Bezold, Leipzig, 1899.)
Faber, Alice. 1997. "Genetic subgrouping of the Semitic languages". Hetzron (1997:3–15).
Fischer, Wolfdietrich. 1982. "Frühe Zeugnisse des Neuarabischen". *Grundriss der arabischen Philologie. I. Sprachwissenschaft*, ed. Wolfdietrich Fischer, 83–95. Wiesbaden: L. Reichert.

Gragg, Gene. 1997. "Geʿez (Ethiopic)". Hetzron (1997:242–260).

Gutt, Ernst-August. 1997. "The Silte group (East Gurage)". Hetzron (1997:509–534).

Hecker, Karl. 1982. "Das Arabische im Rahmen der semitischen Sprachen". *Grundriss der arabischen Philologie*. I. *Sprachwissenschaft*, ed. Wolfdietrich Fischer, 6–16. Wiesbaden: L. Reichert.

Hetzron, Robert. 1997. "Outer South Ethiopic". Hetzron (1997:535–549).

——. (ed.). 1997. *The Semitic languages*. London: Routledge.

Hudson, Grover. 1997. "Amharic and Argobba". Hetzron (1997:457–485).

Jastrow, Otto. 1982. "Die Struktur des Neuarabischen". *Grundriss der arabischen Philologie*, I, ed. Wolfdietrich Fischer, 128–140. Wiesbaden: L. Reichert.

——. 1997. "The Neo-Aramaic languages". Hetzron (1997:334–377).

Johnstone, T.M. 1965. "The application of *kaf* and *gaf* in the Arabic dialects of the Arabian Peninsula". *Journal of Semitic Studies* 7.215–242.

——. 1967. *Eastern Arabian dialect studies*. London: Oxford University Press.

——. 1987. *Mehri lexicon and English-Mehri word-list, with index of the English definitions in the Jibbali lexicon*, comp. G. Rex Smith. London: School of Oriental and African Studies.

Karimov, Akbar U. 1973. *Jemenskij dialekt arabskogo jazyka* [Yemenite dialect of Arabic]. Ph.D. diss., Institut Vostokovedenija, Moscow.

Kaye, Alan S. 1986. *Nigerian Arabic-English dictionary*, II, ed. Robert Hetzron. Malibu: Undena.

—— and Judith Rosenhouse. 1997. "Arabic dialects and Maltese". Hetzron (1997:263–311).

Kjamilev, Said H. 1968. *Marokkanskij dialekt arabskogo jazyka* [The Moroccan dialect of Arabic]. Moscow: Nauka.

Kogan, Leonid. 1997. "Tigrinya". Hetzron (1997: 424–445).

—— and Andrej Korotajev. 1997. "Sayhadic (Epigraphic South Arabian)". Hetzron (1997:220–241).

Landberg, Carl. 1901. *Études sur les dialectes de l'Arabie méridionale*. I. *Ḥaḍramoût*. Leiden: E.J. Brill.

——. 1913. *Études sur les dialectes de l'Arabie méridionale*. II. *Daṯīna*. Leiden: E.J. Brill.

Leslau, Wolf. 1936. *Lexique soqoṭri*. Paris.

——. 1987. *Comparative dictionary of Geʿez*. Wiesbaden: O. Harrassowitz.

Militarev, Alexander and Leonid Kogan. 2000. *Semitic etymological dictionary*. I. *Anatomy of man and animals*. Münster: Ugarit-Verlag.

Mishkurov, Edward N. 1982. *Alžirskiy dialekt arabskogo jazyka* [The Algerian dialect of Arabic]. Moscow: Nauka.

Müller, David H. 1907. *Die Mehri- und Soqoṭri-Sprache*. (= *Südarabische Expedition*,VII/3.) Vienna: A. Hölder.

Naumkin, Vitalij and Viktor Porxomovskij. 1981. *Očerki po etnolingvistike Sokotry* [Ethnolinguistic studies of Soqoṭra]. Moscow: Nauka.

Prochazka, Theodore. 1974. "Perfect-*k* in Taʿizz dialect". *Bulletin of the School of Oriental and African Studies* 37.439–441.

Ratcliffe, Robert R. 1998. *The 'broken' plural problem in Arabic and comparative Semitic*. Amsterdam and Philadelphia: J. Benjamins.

Raz, Shlomo. 1997. "Tigré". Hetzron (1997:446–456).

Retsö, Jan. 2000. "*Kaškaša*, *t*-passives and the ancient dialects in Arabia". *Oriente Moderno* N.S. 19.111–118.

Rhodokanakis, Nikolaus. 1911. *Der vulgärarabische Dialekt im Ḏ̣ofār (Ẓfār)*. (= *Südarabische Expedition*, X/2.) Vienna: A. Hölder.

Sallūm, Dāʾūd. 1976. *Dirāsat al-lahjāt al-qadīma*. Lahōr-al-Kuwait.

Simeone-Senelle, Marie-Claude. 1993. "L'expression du 'futur' dans les langues sudarabiques modernes". *Matériaux Arabes et Sudarabiques*, N.S. 5.249–278.

——. 1997. "The Modern South Arabian languages". Hetzron (1997:378–423).

—— and Antoine Lonnet. 1991. "Lexique soqotri: Les noms des parties du corps". *Semitic studies in honor of Wolf Leslau*, ed. Alan S. Kaye, II, 1443–1487. Wiesbaden: O. Harrassowitz.

Tsereteli, Konstantin G. 1991. "Novoaramejskij jazyk" [The Neo-Aramaic language]. *Jazyki Azii i Afriki*, IV/1, 233–249. Moscow: Nauka.

Wagner, Ewald. 1997. "Harari". Hetzron (1997:486–508).

Zaborski, Andrzej. 1991. "The position of Arabic within the Semitic dialect continuum". *Proceedings of the Colloquium on Arabic Grammar*, ed. Kinga Dévényi and Tamás Iványi, 365–375. Budapest: Eötvös Loránd University.

Zammit, Martin R. 2002. *A comparative lexical study of Qurʾānic Arabic*. Leiden: E.J. Brill.

Zavadovskij, Jurij N. 1962. *Arabskije dialekty Maġriba* [The Arabic dialects of the Maghreb]. Moscow: Izdatel'stvo Vostočnoj Literatury.

ANNA GR. BELOVA (Moscow)

Specificity

1. SPECIFICITY IN LINGUISTIC THEORY

Specificity has been used by linguists in several theoretical traditions to describe a property of noun phrase interpretation. It is most often used in describing the interpretation of indefinite noun phrases, which is the focus of this entry.

The term has been used inconsistently and with little precision to describe a number of related but logically distinct notions. An extensive literature in the generative-compositional semantic tradition is devoted to unpacking these different notions, using a variety of formal devices and debating whether specificity is a matter of truth-conditional semantic or pragmatic inference (for references and useful overview, see Farkas 2002; von Heusinger 2002; Schwarzschild 2002).

Farkas (2002) identifies three separate kinds of specificity discussed in the literature: partitive specificity, scopal specificity, and epistemic specificity.

Partitive specificity describes a noun phrase interpreted as a subset or part of a set referent presupposed in a discourse (Enç 1991). For example, the sentence *ʾaḥbabtu kulla tilka l-kutubi* ‘I liked all of those books’ refers to a particular set of books. The continuation *wa-lākin ištaraytu iṯnayni (min-hā) faqaṭ* ‘but I bought only two (of them)’ refers back to this set (either implicitly or with *min-hā* ‘of them’) and introduces two of these books that the speaker bought. Accordingly, *iṯnayni (min-hā)* is partitively specific.

A scopally specific noun phrase does not have its interpretation restricted by any other superordinate scope-taking operator, and it is often referred to as taking ‘widest’ scope or as ‘scopeless’ (Fodor and Sag 1982). For example, *ʾaḏunnu ʾanna ʾuxtī turīdu ʾan tatazawwaja bi-jundiyyin* ‘I think my sister wants to marry a soldier’ can be understood either as a description of the speaker’s beliefs about the kind of person his sister wants to marry, without commitment either to the truth of his belief or to the existence of a soldier she likes, or as an assertion that he has a particular soldier in mind whom he believes that his sister wants to marry. On this second reading, the interpretation of *jundī* is restricted by neither *ʾaḏunnu* nor *turīdu*, and so takes wide scope with regard to both.

An epistemically specific noun phrase expresses a speaker’s commitment to the existence of an appropriate referent (Lumsden 1988; Lambrecht 1994; Farkas 2002; von Heusinger 2002; Schwarzschild 2002; see also Givón’s 1979 term ‘referentiality’). In English, epistemic specificity is expressed by the adjectives *certain* or *particular* or, in more colloquial registers, the ‘indefinite’ use of *this* (Prince 1981). The Levantine dialects of Arabic have an analogous indefinite use of the demonstrative prefix *hal-* ‘this’ (cf. Blau 1960:20), found frequently in storytelling, jokes, and so on (see below for examples).

Two more kinds of specificity can be adduced from the literature: descriptive specificity (corresponding to some definitions of Arabic *taxṣīṣ*; see Brustad 2000:21–31); and quantificational specificity (corresponding to Arabic *tafrīd* ‘individuation’).

A descriptively specific indefinite noun phrase has some degree of restrictive modification: a more richly modified noun is more restricted in its potential reference than a less modified noun (Wright 1898:II, 198, 260–263; Mohammad 2000:111). For example, *rajulun ṭawīlun wa-jamīlu l-wajhi* ‘a tall man with a handsome face’ is more specific than *rajulun ṭawīlun* ‘a tall man’ because it describes a more narrowly defined subset of the set of men.

Quantificational specificity is the degree to which a noun phrase describes a discrete quantity rather than an undifferentiated mass. A singular noun phrase such as *ʿarabī* ‘an Arab’ is more quantificationally specific than the paucal plural *ʿarabiyyūna* ‘some Arabs’, which, in turn, has more quantificational specificity than the mass term *ʿarab* ‘Arabs’.

The different notions of specificity frequently overlap but do not entail each other. For example, *sayyāratan lawnuhā ʾaḥmaru zahīyun wa-maftūḥatan min al-fawq* ‘a bright red convertible’ in (1a) (Doron and Haycock 1999) is more heavily modified than *sayyāratan raʾaytu-hā ʾamsi* in (1b) and is therefore more descriptively specific.

(1a) *ʾurīdu* *ʾan* *ʾaštariya*
want.1s Compl buy.1s
sayyārat-an *lawn-u-hā*
car.fs-Acc color.ms-Nom-cl3fs
ʾaḥmar-u *zahīy-un*
red.ms-Nom bright.ms-Nom
wa-maftūḥat-an *min* *al-fawq*
and-open.fs-Acc from the-above
‘I want to buy a bright red convertible’

(1b) *ʾurīdu* *ʾan* *ʾaštariya*
want.1s Compl buy.1s
sayyārat-an *raʾaytu-hā* *ʾamsi*
car.fs-Acc saw.1s-cl3fs yesterday
‘I want to buy a car that I saw yesterday’

However, (1a) can be used by a speaker who has no actual car in mind and who is only describing his or her ideal car, in which case it has low epistemic and scopal specificity. In contrast, *sayyāratan raʾaytu-hā ʾamsi* ‘a car I saw yesterday' in (1b) has high epistemic specificity because the speaker saying it asserts the existence of a particular car.

2. Specificity in nominal clauses

It is often claimed that only a definite noun phrase can be used as the *mubtadaʾ* (→ *ibtidāʾ*) in a → nominal clause and that indefinites cannot be, unless they have generic reference, are pronounced with emphatic intonation, are modified, or are conjoined with a modified noun (Wright 1898:II, 260–264; Cantarino 1975:31; Bakir 1980:62–63; Brustad 2000:21–31, 332; Mohammad 2000:111–141; Holes 2004:252).

However, in both Standard Arabic and the dialects, noun phrases meeting none of these criteria undergo *ibtidāʾ*, particularly at the beginning of jokes or narratives, as in (3) and (4).

(3) Standard Arabic (Khan 1988:33)

ʾasad-un	marrat-an
lion.ms-Nom	time.fs-Acc
ištadda	ʿalay-hi
weighed.3ms	upon-cl3ms
ḥarr-u	š-šams
heat.ms-Nom	the-sun

'a [certain] lion, once the heat of the sun weighed upon him'

(4) Bir Zeit Palestinian Arabic (Schmidt and Kahle 1930:§4)

ḥaṭṭāb	*b-irūḥ*	*čill*
woodcutter.ms	Ind-go.3ms	every
yōm	*ʿa-l-ḥīš*	
day	to-the-forest	
b-ijīb-l-e	*ḥimil*	*ḥaṭab*
Ind-bring.3ms-tocl3ms	load	wood

'A [certain] woodcutter, he would go every day to the forest and get himself a load of wood'

It is not clear that any one kind of specificity is uniquely associated with *ibtidāʾ*. For example, a *mubtadaʾ* can have generic reference and hence low quantificational specificity, while the *mubtadaʾ* in each of (3) and (4) is unmodified and hence descriptively nonspecific, although it is specific epistemically.

Rather, following Khan (1988), *ibtidāʾ* may be a strategy for marking topichood (Grimes 1975; Li and Thompson 1976, among others; see also Belyayeva 1996). Different kinds of specificity may be used to reinforce topic-hood, but none of them is necessary for doing so.

3. Specificity and noun modification

Many dialects allow modification of indefinites with 'definite' relative clauses to increase their specificity, an apparent exception to a well-known rule that indefinite nouns are modified only by 'indefinite' modifiers (see Brustad 2000:91–99). Examples have been noted in Egyptian (Mitchell 1956), Palestinian (Schmidt and Kahle 1918, 1930; Blau 1960; Belyayeva 1996), Moroccan (Harrell 1962), and Syrian Arabic (Cowell 1964), and in various Gulf Arabic dialects (Holes 1990).

For example, *walad* 'boy' in (5) is modified by the definite relative clause *illi ibn ʿarūs katal abū* 'whose father Ibn Arus had killed' and introduces a new character into a narrative. Accordingly, the use of the definite relative clause increases the epistemic specificity of *walad*:

(5)

bāki		*hāna*	*walad*
be.Part.Act.ms		here	child
illi	*ibn*	*ʿarūs*	*katal*
Rel	son	Arus	kill.Perf.3ms
abū			
father-cl3ms			

'There was a child here whose father Ibn Arus had killed' (Schmidt and Kahle 1930:§34)

In Syrian Arabic, the verbal mood of a relative clause can reinforce the degree of epistemic specificity with which a noun phrase is interpreted. Cowell (1964:356–357), for instance, provides the pair of sentences in (6), which contrast in terms of the epistemic specificity with which *wāḥdi* '[female] person' is interpreted.

(6a)

ʿandi	*wāḥdi*	*btaʿref*	*ʾənglīzi*
at-cl1s	one.fs	Ind-know.3fs	English

'I have someone who speaks English'

(6b)

bədd-i	*wāḥdi*	*taʿref*
want–1s	one.fs	know.3fs
təḥkī-l-ha	*šwayyet*	*ʾənglīzi*
speak.3fs-to-cl3fs	little	English

'I want someone who can speak a little English'

In (6a), the relative clause *btaʿref ʾənglīzi* ‘who can speak English’ contains a verb marked in the indicative mood. Based on the gloss, *wāḥdi* is epistemically specific because the sentence expresses the speaker’s knowledge of a particular person. In contrast, the subjunctive relative clause *taʿref təḥkī-l-ha šwayyet ʾənglīzi* ‘who can speak a little English’ in (6b) reinforces that the speaker does not have a particular person in mind, and hence *wāḥdi* is epistemically nonspecific.

4. Specificity and functional morphology

In several dialects, such as the → Najdi Bedouin dialects of Saudi Arabia (Abboud 1964; Ingham 1994; → noun phrase), vestigial → *tanwīn* is used to mark indefinite nouns as specific. For example, Ingham (1994:47) contrasts the bare stem *bēt* ‘house’, with *bēt-in* ‘a [particular] house’ marked with the *-in* suffix and glossed to indicate higher epistemic specificity.

In several dialects, indefinite articles (→ article, indefinite) are used to express various kinds of specificity (Brustad 2000:26–27), including cognates of *šī* ‘some’ in Syrian (7a) and Moroccan (7b), and *fadd* (7c) in various varieties of Iraqi Arabic (Erwin 1969:355–358; Jastrow 1990):

(7a) *štarēt ᵊs-sayyāra mn*
bought.2s the-car from
ᵊl-wakīl wəlla mən šī šaxᵊṣ?
the-dealer or from some person
‘Did you buy the car from the dealer or from some person?’ (Cowell 1964:470)

(7b) *w-āna ʿandi ši nās ḍīfān*
and-I at-cl1s some people guests
‘I had some people as guests’ (Brustad 2000:27)

(7c) *ez-zawāǰ mālna, ida wēḥəd*
the-marriage Poss-cl1s if one
kār-rād fad bənt, nəḥne
want.3ms some girl we
ʿədd-na mā kān ak-u
custom-cl1p not was that-cl3ms
yəmši maʿa, yəǰi,
walk.3ms with-cl3fs come.3ms
yəmši, la
go.3ms no
‘Our marriage, if someone wanted some girl, our custom was not that he would go out with her, [and that he would] come [and] go, no’ (Jastrow 1990:166–167)

As noted above, in Levantine Arabic the demonstrative *hal-* ‘this’ (Blau 1960:20) has an ‘indefinite’ use paralleling the ‘indefinite’ use of *this* in vernacular English (Prince 1981). In (8), a fragment of naturally occurring discourse from Levantine Arabic, *hal-wāḥad* ‘this one’ is indefinite because it is used to introduce a novel discourse referent.

(8) Q: *bi-šū bi-tfakkir halla’?*
in-what Ind-think.2fs now
‘What are you thinking about now?’

A: *b-astanna bi-hal-wāḥad u-miš*
Ind-wait.1s in-this-one.ms and-not
rāḍi ybayyin
ready.ms appear.3ms
‘I’m waiting for this person and [he] has not agreed to show up’
(www.mahjoob.com/ar/forums/printthread.php?t=154453&page=40&pp=25)

5. Specificity and agreement marking

In several dialects, existential or → locative sentences (Arabic *jumal ḏarfiyya* ‘locative sentences’; → *mafʿūl fīhi*) alternate between ‘full’ agreement marking on the verb and ‘reduced’ or ‘neutralized’ 3rd person masculine singular agreement (Cowell 1964; Belnap 1991; Hoyt 2000, 2002).

In some rural dialects of Palestinian Arabic, speakers have subtle preferences for full or neutralized agreement, depending on the form of the postverbal subject or ‘pivot’ (see Halila 1992:353 and Hoyt 2002:125–126 for similar phenomena in Tunisian Arabic). With richly modified pivots, there is a slight preference (indicated with ‘!!’) for full agreement:

(9) *bāki /!!bākye hanāk*
be.Part.Act.ms/be.Part.Act.fs there
ḥayye bidd-ha tōčil frāx
snake.fs want-cl3fs eat.Imperf.3fs chicks
iṭ-ṭēr
the-bird
‘There was a snake there [that] wanted to eat the bird’s eggs’ (Hoyt 2002)

If the pivot is premodified by a numeral, a slight preference for neutralized agreement emerges:

(10) *bākyīn/!!bāki* *fī* *dār*
be.Part.Act.mp/be.Part.Act.ms in house
abū-ha *sabiʿ* *ṯaman* *ḥarrāṯīn*
father-cl3fs seven eight plowmen
'In her father's household were seven or eight plowmen' (Hoyt 2002)

A pivot prefixed with indefinite *hal-* is also associated with a slight preference for neutralized agreement, as in (11).

(11) *čānat /!!čān* *tiḥt*
be.Perf.3ms/be.Perf.3fs under
sēr-e *haṭ-ṭabanje* *mnazzale*
belt-cl3ms this-pistol.fs inlaid.fs
bi-l-fiḏ̣ḏ̣e
with-the-silver
'Under his belt was a pistol inlaid with silver' (Hoyt 2002:115)

The preference for full agreement may not correspond directly to any one notion of specificity, but rather to topicality in Khan's (1988) sense. However, indefinite *hal-* and rich restrictive modification increase the epistemic and descriptive specificity of *ṭabanje* 'pistol', signaling to a listener that it is a topic. The preference for neutralized agreement with a prenominal number may decrease a pivot's potential topicality by reducing its quantificational specificity, as it focuses on the quantity of an unindividuated group rather than on the properties of the individual or individuals that make up the group. This may correlate with the preference for neutralized agreement.

A particularly interesting contrast is between the preference for the full agreement in (9) and neutralized agreement in (11). The noun phrase *haṭ-ṭabanje mnazzale bi-l-fiḏ̣ḏ̣e* is descriptively, epistemically, and quantificationally specific and introduces a discourse topic. Accordingly, we would expect to find a preference for full agreement. It may be that the epistemic specificity contributed by *hal-* is sufficient to reinforce the topicality of the noun phrase, while neutralized agreement distinguishes the indefinite use of *hal-* from its definite-deictic use.

BIBLIOGRAPHICAL REFERENCES

Abboud, Peter F. 1964. *The syntax of Najdi Arabic.* Ph.D. diss., University of Michigan, Ann Arbor.

Bakir, Murtada. 1980. *Aspects of clause structure in Arabic.* Bloomington, Ind.: Indiana University Lin-guistics Club.

Belnap, Kirk. 1991. "The meaning of deflected/strict agreement in Cairene Arabic". *Perspectives on Arabic linguistics*, V, ed. Mushira Eid and Clive Holes, 97–118. Amsterdam and Philadelphia: J. Benjamins.

Belyayeva, Dina. 1996. "Definiteness realization and function in Palestinian Arabic". *Perspectives on Arabic linguistics*, X, ed. Mushira Eid and Robert R. Ratcliffe, 47–67. Amsterdam and Philadelphia: J. Benjamins.

Blau, Joshua. 1960. *Syntax des palästinensischen Bauerndialektes Bir Zeits.* Walldorf, Hessen: Verlag für Orientkunde.

Brustad, Kirsten. 2000. *The syntax of spoken Arabic: A comprehensive study of Moroccan, Egyptian, Syrian, and Kuwaiti dialects.* Washington, D.C.: Georgetown University Press.

Cantarino, Vicente. 1975. *The syntax of modern Arabic prose.* 3 vols. Bloomington, Ind.: Indiana University Press.

Chafe, Wallace L. 1976. "Givenness, contrastiveness, definiteness, subjects, topics, and point of view". Li (1976:25–56).

Cowell, Mark. 1964. *A reference grammar of Syrian Arabic.* Washington, D.C.: Georgetown University Press.

Doron, Edit and Caroline Heycock. 1999. "Filling and licensing multiple specifiers". *Specifiers: Minimalist approaches*, ed. David Adger, Susan Pintzuk, Bernadette Plunkett, and George Tsoulas, 69–89. Oxford: Oxford University Press.

Enç, Murvet. 1991. "The semantics of specificity". *Linguistic Inquiry* 22.1–25.

Erwin, Wallace. 1969. *A basic course in Iraqi Arabic.* Washington, D.C.: Georgetown University Press.

Farkas, Donka. 2002. "Specificity distinctions". *Journal of Semantics* 19.213–243.

Fodor, Janet Dean and Ivan A. Sag. 1982. "Referential and quantificational indefinites". *Linguistics and Philosophy* 5.355–398.

Givón, Talmy. 1979. *Discourse and grammar.* New York: Academic Press.

Grimes, Joseph. 1975. *The thread of discourse.* Berlin and New York: de Gruyter.

Halila, Hafedh. 1992. *Subject specificity effects in Tunisian Arabic.* Ph.D. diss., University of Southern California.

Harrell, Richard. 1962. *A short reference grammar of Moroccan Arabic.* Washington, D.C.: Georgetown University Press.

Heusinger, Klaus von. 2002. "Specificity and definiteness in sentence and discourse structure". *Journal of Semantics* 19.245–274.

Holes, Clive. 1990. *Gulf Arabic.* London: Routledge.

——. 2004. *Modern Arabic: Structures, functions, and variables.* 2nd ed. New York: Longman.

Hoyt, Frederick. 2000. *Word order, agreement, and specificity effects in rural Palestinian Arabic existential constructions.* M.A. thesis, Cornell University.

——. 2002. "Reduced agreement as a specificity effect in rural Palestinian Arabic". *Perspectives on Arabic linguistics,* XIII–XIV, ed. Elabbas Benmamoun, 111–141. Amsterdam and Philadelphia: J. Benjamins.

Ingham, Bruce. 1994. *Najdi Arabic: Central Arabian.* Amsterdam and Philadelphia: J. Benjamins.
Jastrow, Otto. 1990. *Der arabische Dialekt der Juden von ʿAqra und Arbīl.* Wiesbaden: O. Harrassowitz.
Khan, Geoffrey. 1988. *Studies in Semitic syntax.* Oxford: Oxford University Press.
Lambrecht, Knud. 1994. *Information structure and sentence form.* Cambridge: Cambridge University Press.
Li, Charles N. (ed.). 1976. *Subject and topic.* New York: Academic Press.
—— and Sandra A. Thompson. 1976. "Subject and topic: A new typology of language". Li (1976:457–489).
Lumsden, Michael. 1988. *Existential sentences: Their structure and meaning.* London: Croom Helm.
Mitchell, T.F. 1956. *An introduction to Egyptian Colloquial Arabic.* London: Oxford University Press.
Mohammad, Mohammad A. 2000. *Word order, agreement, and pronominalization in Standard and Palestinian Arabic.* Amsterdam and Philadelphia: J. Benjamins.
Prince, Ellen F. 1981. "Inferencing of indefinite-*this* NPs." *Elements of discourse understanding*, ed. Aravind Joshi, Bonnie L. Webber, and Ivan A. Sag, 231–250. New York: Cambridge University Press.
Schmidt, Hans and Paul Kahle. 1918. *Volkserzählungen aus Palästina*, I. Göttingen: Vandenhoeck und Ruprecht.
—— and Paul Kahle. 1930. *Volkserzählungen aus Palästina*, II. Göttingen: Vandenhoeck und Ruprecht.
Schwarzschild, Roger. 2002. "Singleton indefinites". *Journal of Semantics* 19.289–314.
Wright, William. 1898. *A grammar of the Arabic language.* 2 vols. London: Cambridge University Press.

Frederick M. Hoyt
(University of Texas at Austin)

Specifier → X-Bar Syntax

Speech Accommodation

What kind of language do speakers of Arabic from geographically remote parts of the Arab world use when they meet? Which language do Moroccans use, for instance, when communicating with Egyptians, and why? These are commonly asked questions posed by non-Arabic speakers upon realizing how linguistically diverse the varieties of Arabic (or 'dialects') are. With the tremendous amount of linguistic variation not only across borders in the Arab world but also within the boundaries of the same country, questions similar to the following also arise: What would a young urban Tunisian female (typically a code-switcher in Arabic and French, with a high school education as a minimum) use to ask directions from an older Tunisian male construction worker (typically a speaker of rural Tunisian Arabic with minimal education and possibly no bilingual education)? Which element in the context of the interaction would most determine her linguistic choices here? Will the interaction succeed if the speaker approaches her interlocutor in her gender-, class-, and urban-marked Arabic-French code-switching? Or will she perhaps make some other choices, drawing on her linguistic repertoire and assessment of the situation? What will motivate her linguistic choices?

Communication Accommodation Theory (originally known as Speech Accommodation Theory) offers a particularly useful and comprehensive framework for understanding some of the social-psychological factors motivating speakers to make particular linguistic choices in particular interactions. Communication Accommodation Theory (CAT) argues that speakers modify their language in relation to the language of their addresses and their perception of their own and the latter's social identity. It is particularly relevant to comprehending communication in Arabic, a language rich in geographically, economically, and socially based varieties (or 'dialects'). An increase in opportunities for linguistic contact among Arabic varieties, both real and virtual through television and the Internet, has been witnessed especially during the last few decades. Opportunities for linguistic accommodation consequently became more pressing for Arabic speakers, possibly fueling some of the hybrid linguistic forms and the linguistic shifts observed in progress today. Applying Communication Accommodation Theory to the study of languages and varieties in contact might allow prediction of the direction that variation and change as well as language attrition and death (→ language loss) might take in a given community.

The term 'Speech Accommodation Theory' (SAT) was first introduced by Giles (1973) in a study in which he noticed that interviewees adjusted their accents in order to sound more like their interviewers. Speech Accommodation Theory thus emerged to account for the social-psychological motives and effects underlying speakers' modification of their language in relation to the language of their interlocutors. Beebe and Giles point out that Speech Accommodation Theory was "devised to explain some of the motivations underlying certain shifts in people's speech styles during social encounters, and some of the social consequences arising from them. More specifically, it originated in order to elucidate the cognitive and affective processes underlying speech convergence and divergence" (1984:7). The designation 'Speech Accommodation Theory' was later changed and expanded to 'Communication Accommodation Theory' (CAT; Giles a.o. 1987), when it became clear that accommodation affected forms of communication other than speech, such as utterance length, pauses, or facial expressions. It also became clear that accommodation was a prevalent and integral feature of communication. "Each one of us will have experienced 'accommodating' verbally and non-verbally to others, in the general sense of adjusting our communication actions relative to those of our conversation partners, and been aware of others accommodating (or failing to accommodate) to us" (Giles and Coupland 1991:60). Accommodation may manifest itself in a multitude of forms; speaking louder to foreigners or to the blind, code-switching into another language or variety, and 'baby talking' to older people represent some of the illustrations of this phenomenon (Giles a.o. 1987). Although this phenomenon was noted prior to Giles (1973), it is with the development of Communication Accommodation Theory that a more complex framework emerged to account for some of the intricacies in communication. Before surveying the applications of accommodation to the study of Arabic, a few relevant constituents of this theory are delineated below.

According to Communication Accommodation Theory, accommodation can be convergent or divergent. Convergent accommodation is defined as "a strategy whereby individuals adapt to each other's communicative behaviors in terms of a wide range of linguistic-prosodic-nonverbal features including speech rate, pausal phenomena and utterance length, phonological variants, smiling, gaze, and so on" (Giles a.o. 1991:7). Because it reduces linguistic dissimilarities between individuals, this behavior "can lead persons to attribute to the converger the traits of friendliness, warmth, and so on" (Giles a.o. 1987:15). More specifically, Giles a.o. explain that "during interaction individuals are motivated to adjust (or accommodate) their speech styles as a strategy for gaining one or more of the following goals: evoking listeners' social approval, attaining communicational efficiency between interactants, and maintaining positive social identities. In addition, it is the individual's *perception* of the other's speech that will determine his or her evaluative and communicative responses" (1987:14–15).

Furthermore, convergence has been considered a "reflection (often unconscious) of a speaker's or group's need for social integration or identification with another" (Giles a.o. 1987:16), except when the convergence is meant to mock the speech style of that interlocutor (Giles a.o. 1987:17).

Power is also an integral variable in determining the degree of convergence exhibited by an individual or a group. Hence, a subordinate individual or group will converge more toward the language of the dominant party than vice versa. Wolfram's (1973) work cited in Giles a.o. (1987:22), for example, reports that because blacks in New York City have more power and prestige than Puerto Ricans, the latter adopt the dialect of the former more than vice versa. However, when friction exists at the intergroup level, i.e., when, according to Communication Accommodation Theory, a situation is described as high in intergroup and low in interindividual terms, speakers may highlight their identity by accentuating their 'ingroup' (us) language patterns to distinguish themselves from the 'outgroup' (them).

Divergent accommodation and maintenance, on the other hand, involve accentuating differences between speakers and their interlocutors' speech styles or speakers maintaining their own speech styles respectively. Both strategies can be evaluated negatively by recipients if they are perceived as dissociative, but favorably if the encounter is interpreted in intergroup terms (Beebe and Giles 1984:8–9). For instance, Bourhis and Giles (1977) show how a group of Welsh students assert their Welsh identity in the presence of an Englishman who questions the vitality of the Welsh language, by broadening their Welsh accent while speaking English, by code-switching, or by shifting completely to Welsh. By shifting to Welsh (even if it is merely by conjugating a Welsh verb as one student does), this group announces its divergence from the Englishman and affirms its Welsh ingroup membership.

Both convergence and divergence can be upward or downward. Hence, one can adjust one's style to a variety that is higher or lower than one's own in terms of status.

Another relevant manifestation of accommodation is stereotypical convergence: the process by which speakers will converge toward the speech style they believe their interlocutors to possess. Bell (1982), for instance, shows how broadcasters in New Zealand phonologically adjust their reading of the same news to the different types of audiences they believe are listening. Also, in their research on speakers of different status, Thakerar a.o. (1982:236) found that a lower-status speaker and a higher-status speaker both adopted the speech stereotype associated with their partner. Their conclusion was that although they diverged linguistically from each other, these speakers might not only have converged psychologically toward their partners, they might also "have been attempting to converge linguistically to what they *believed* the speech of the other to be" (Thakerar a.o. 1982:235). This is confirmed by Giles a.o.: "In face-to-face interactions where individuals act as group members, it has been found that people communicatively accommodate to group stereotypes" (1987:63). A well-intentioned manifestation of this type of accommodation that may be perceived as discriminatory is the language able-bodied people use with the disabled or the visually impaired (as reported by Markova in 1990 and Klemz in 1977, respectively, mentioned in Giles and Coupland 1991).

Despite its apparently harmless manifestations, accommodation has been found to reflect existing power relations and sociopolitical relations in interpersonal and intergroup encounters across age, gender, class, and ethnic lines (see, for instance, Coupland a.o. 1991; Genesee and Bourhis 1982; Giles a.o. 1991; Giles and Coupland 1991; Gudykunst 1988). The few studies to date that investigate accommodation among speakers of Arabic confirm that Arabic is no exception.

Because of the vast geographical expanse occupied by the Arab world and the historical and political changes that it has witnessed, it is not surprising that Arabic possesses such a great number of spoken varieties or dialects in addition to the different forms of Modern Standard Arabic and Classical Arabic. In addition to the four generally recognized major regional varieties of Arabic (Egyptian, Levantine, Gulf/Iraqi Arabic, and Arabic of the Maghreb), numerous distinct varieties can be isolated within each country, often along the lines of gender, religion, education, geographical location, ethnicity, class, and age (→ variation). Conversely, because migration from the Arabian Peninsula to what is now the Arab world occurred in waves, certain varieties in distant locations (such as the so-called → Bedouin Arabic variety) are occasionally more mutually intelligible across borders than they are with geographically closer varieties. Urban/rural varieties within the same country are a good example of this linguistic distance within the same geopolitical borders. Since linguistic prestige and power tend to be ascribed to urban centers, each Arabic national variety would thus possess its own 'standard' variety (Holes 1995b), most likely located in its capital city. Even though Arabic dialect classification criteria seem rather complex, speakers of this large number of varieties of spoken Arabic tend to have a clear sense of the features that distinguish their varieties from those of others. The distinction could be as simple as a single phoneme or some lexical items that set them apart from others. As they become cognizant of the sociolinguistic attitudes associated with their varieties and those of others, different dialect speakers acquire an understanding of the kind of accommodation required to achieve the desired purposes in encounters with speakers of other varieties. This is equally true of the experience of Arabic speakers from adjacent villages and speakers of remote urban dialects from major cities of the Arab world.

Numerous causes of contact among Arabic varieties can be identified. The focus here is on two of the most prominent ones. Population mobility, mainly triggered by urbanization, economic migration within the same country or across borders, and by the different conflicts in the region, is the first major factor in the intensifying contact that obtains between the different spoken varieties of Arabic. The second important element of contact, albeit of a more virtual nature, is the growing exposure to other Arabic varieties that has ensued from the spread of Arab satellite television channels starting in the mid-1990s (→ media) and the growing use of the Internet in more recent years.

Because of the differences that exist between the spoken varieties of Arabic, issues of mutual comprehensibility routinely arise both within the same country and across borders. The linguistic situation in the Arab world appears more multifarious when one takes into account the fact that the linguistic repertoire of many speakers of Arabic involves diglossic, bilingual, multilingual, and code-switching proficiency (→ diglossia; → multilingualism; → code-switching). In addition to speaking their own dialect or dialects, many acquire through education Modern/Classical Arabic and Educated Spoken Arabic and will possibly speak other indigenous and/or colonial languages, and will adequately code-switch in these languages and varieties. Comprehensibility across varieties attracts a lot of attention, however. The differences between some of the local varieties are so stark that it is sometimes argued that it would be more accurate to speak of 'languages' of Arabic rather than 'dialects' of the same language. The wider held belief among many Arabs, however, is deeply entrenched in Arabic language ideology. This belief contends that Arabs speak the same language with 'slight variations'. Regardless of where they stand ideologically on this matter, Arabic speakers are increasingly faced with situations that call for communication with speakers of other, potentially incomprehensible dialects. Various forms of communication accommodation are thus observed as Arabic speakers make linguistic choices in different contact situations.

Research that specifically uses Communication Accommodation Theory in analyzing the language choices used in contact situations in Arabic remains scarce to date (Abu-Melhim 1991, 1992; Walters 1991; Lawson-Sako and Sachdev 1996; S'hiri 2002; Suleiman 2004). Research that touches on the subject from other perspectives can be noted, but remains limited as well.

Two main kinds of contact situation are documented in the literature. The first kind obtains between different groups from the same community or country. The second can be identified between speakers of different national varieties in contact. The contact situation might be one of the interlocutors' countries, or a location outside the Arab world altogether. Examples of intracountry contact situations calling for accommodation could be Jordanian women interacting today with their male compatriots, or it could be Muslim, Christian, and Jewish Iraqi speakers going about their daily business (in the 1940s). Instances of communication across national borders are Maghrebi Arabic speakers in contact with speakers of the Arabic of the Levant. A more complex example would be speakers of two varieties that belong to the same group (e.g. Levantine Arabic), but whose distinction has national identity repercussions (e.g. Jordanian and Palestinian Arabic speakers living in Jordan).

Among the studies documenting occurrences of accommodation within the same country, Blanc's (1960) is perhaps the first, even though he does not use this terminology. Blanc refers to 'variations of style' when discussing 'modifications' commonly introduced within Arabic dialects. He identifies two processes by which these modifications are implemented: one is → 'leveling', usually in crossdialectal situations, and 'Classicization', usually present in 'educated' speech and when more formal speech is deemed necessary (→ classicism). Blanc (1960) classifies these 'style varieties' into five categories that tend to intermix constantly in usage: plain colloquial, koineized colloquial, semiliterary or elevated colloquial, modified Classical, and standard Classical. Although he refers to the interlocutor as a potential reason for selecting a particular style variety at a certain point in the course of communication, he does not discuss it in greater detail from this perspective.

Later works by Walters (1991), Lawson-Sako and Sachdev (1996, 2000), and Holes (1986, 1995a) explicitly analyze style variation as a function of addressee accommodation. In all of these studies, instances of convergent and divergent accommodation consistently signal change and the emergence of new varieties in response to social and political constraints. The first two studies focus on Tunisia, while Holes investigates Bahrain, Jordan, and Iraq (without using the terminology of Communication Accommodation Theory). Lawson-Sako and Sachdev (1996) concentrate on linguistic convergence and divergence among Tunisians and foreigners in the coastal town of Sousse, a tourist destination accustomed to regular linguistic contact the year around between the locals and visitors (mainly European). The study concludes that while the Tunisians in the study generally converged linguistically toward their interlocutors, their behavior varied according to the ethnicity, gender, and language of the researcher. The Tunisian subjects were asked directions to the post office in Tunisian Arabic and French. The language choices made by the Tunisian subjects in responding to the researchers (Tunisian Arabic, French, code-switching) were not determined by linguistic competence. The choices were "identity choices, varying systematically as a function of the language, ethnicity and gender of the researcher" (Lawson-Sako and Sachdev 1996:76). The study found, for instance, that male researchers, regardless of whether they used French or Tunisian Arabic, encountered a high level of initial convergence by Tunisian subjects. Tunisian female researchers speaking French, however, received the lowest levels of convergence and the highest levels of divergence out of all researchers using French (Arab male, European male, and European female). There was strong divergence in response to the African researcher's requests, regardless of whether he used French or Tunisian Arabic. The study further supports Giles a.o.'s claim that individuals accommodate to group stereotypes when in situations that dictate that they act as group members.

Lawson-Sako and Sachdev (2000) investigate attitudes toward Tunisian Arabic/French code-switching among college students and the Tunisian population at large. They note that convergence to requests in Tunisian Arabic was higher than to requests in French, and divergence was higher to requests in French than it was to requests in Tunisian Arabic. Again, there was higher convergence to Tunisian Arabic-speaking males, whether European or Arab, while female Tunisians received a higher amount of code-switching responses. The latter finding has been explained as an instance of stereotypical convergence, because women are perceived to code-switch more than men. The findings also demonstrated that the negative attitude toward code-switching elicited during the matched-guise technique part of the study was not reflected in the students' self-reports and the general population's actual behavior on the streets. Code-switching was instead used mostly with in-group members, i.e. friends, family, and other Tunisians and was used less with non-Arabs and outside people such as teachers.

These findings have been confirmed by Walters (1991). His study suggests that code-switching appears to be a variety used mainly in intraethnic situations, especially in situations involving friends. In other words, it is used with the in-group (other Tunisians) to foster a feeling of solidarity through the use of Tunisian Arabic, combined with a component of 'socioeducational competence' through the use of French. Walters further claims that code-switching can be considered a distinct language variety in Tunisia, used in private settings and social activities. Among friends, Walters found that his subjects (who were college students of English) exploited their linguistic repertoire in full. They used code-switching speech as their main language, combined with some Tunisian Arabic as well as a little Modern Standard Arabic, French, and English.

Moving to the east of the Arab world, Holes (1995a) focuses on the role that literacy and population mobility play in reshaping varieties and the relationships between them in three countries in the Arab Middle East: Bahrain, Jordan, and Iraq. Holes does not use Communication Accommodation Theory to analyze his data, but refers with the terms 'assimilation', 'shift' and 'convergence' to convergent accommodation. In Bahrain, Holes (1986, 1995a) documents the emergence of a new 'standard' dialect as a result of the move away from sectarian separation between ʿArab and Baḥārna with the creation of new job opportunities and the mixing of schools beginning in the 1970s. Initially, the Baḥārna community marker /ṯ/ was realized as /f/ and the ʿArab's community marker /j/ was realized as /y/. During the process of contact and integration between these two communities, literate Baḥārna speakers completely abandoned stereotypical variants. The /q/ that used to be realized as /ḳ/ is now realized as /g/, which is not standard, but rather part of ʿArab speech. The Baḥārna /f/ and /d/ have been replaced completely by /ṯ/ and /ḏ/, which have double force because they are part of both Modern Standard Arabic and the ʿArab dialect. The variant /č/ has disappeared except for an occasional realization of /k/ when it is shared by the ʿArab. Interestingly, /j/ has been replaced by /y/, even though /y/ is not the Modern Standard Arabic form but rather part of the ʿArab dialect. Conversely, stereotypical ʿArab variants that are non-Modern Standard Arabic remained strong (e.g. /q/ realized as /g/ and /j/ realized as /y/), suggesting that "the convergence of educated A and B speakers was invariably asymmetrical, and strongly in the direction of a modified form of the A dialect" (Holes 1995a:276). As the correlation between community and dialect gradually broke down, especially among the young literate population, a "modified form of the A [ʿArab] dialect is now widely heard in Manama (and indeed in the rest of Bahrain), and seems to have acquired the role, in public, particularly in intercommunal speech contexts, of a neutral 'standard' for both A and B [Baḥārna] speakers" (Holes 1995a:275).

Holes then notes how an Amman urban dialect is beginning to form, in parallel to three other dialects and Modern Standard Arabic (→ Jordanian Arabic: Amman). The new dialect primarily relies on elements from the *madanī* (urban) dialect (brought to Jordan by Palestinian immigrants) as well as elements from the Jordanian Bedouin dialect, with little value given to their *fallāḥī* (rural) counterpart. This new dialect is motivated by local, political, and social reasons. In a more recent analysis, Suleiman (2004) points out that the Bedouin (Jordanian) variety (characterized here with the /g/ variant) is given institutional support in the arts and electronic media. Male speech shifted in favor of the indigenous Bedouin /g/ variant after Black September of 1970, while prestigious female speech remained associated with the use of the variant /ʾ/, also typical of the urban dialect. Suleiman (2004:133) suggests, furthermore, that the "linguistic feminization of the Palestinians through [ʔ] and the linguistic masculinization of the Jordanians through [g] are in fact no more than metaphors for this asymmetry in power distribution in the country". He explains that as "the dialect of the in-group (and political dominant group), the Bedouin variety acts as the target of speech convergence for the Palestinians. This reflects the difference in power allocation between Palestinians and Jordanians at various levels" (Suleiman 2004:130).

When investigating the language situation in Baghdad, Holes identifies a new 'standard' dialect based on the rise to prominence of variants from the hitherto secondary Muslim Baghdadi dialect, as it distanced itself from its Bedouin roots and acquired sedentary variants (→ Baghdadi Arabic: Muslim). The other dialects that coexisted with the Muslim Baghdadi were the Christian Baghdadi (→ Baghdadi Arabic: Christian) and the Jewish dialect, which has now disappeared from Baghdad but can still be found among Baghdadi Jews living in Israel and elsewhere. Holes refers to Muslim Baghdadi as a *gilit*-dialect (*gilit* meaning 'I said') and to Christian Baghdadi and Jewish Baghdadi as a *qeltu*-dialect, indicating their stereotypical marked use of the variants /g/ and /q/, respectively. The *qeltu*/*gilit* distinction was first noted and coined by Blanc (1964) in his study of Baghdad's → 'communal dialects', i.e. the dialects spoken by different religious communities (Muslims, Christians, and Jews). According to Blanc (1964), the *qeltu* variety was most likely shared by all three religious groups in urban centers since the times of the Abbasids until a later wave of urbanization brought Bedouin (who are *gilit* speakers) to Baghdad. Eventually, all Baghdadi Muslims became *gilit* speakers, while Christian Baghdadi and Jewish Baghdadi continued to be *qeltu* speakers. Most Jewish Baghdadi speakers left Baghdad in 1950–1951 to live mainly in Israel.

Starting in the 1960s, Muslim Baghdadi was used by Christians and Jews in intercommunal situations because it had become the local lingua franca thanks to the postindependence rise in power of the Sunnis. "In relative terms, both the size and the political influence of the non-Muslim population in Baghdad shrank, and provided the incentive for a dialectal accommodation to M[uslim] B[aghdadi] which had previously been unnecessary" (Holes 1995a:285). During the 1990s, Muslim Baghdadi began to be used sometimes even within the Christian community (i.e. outside intercommunal situations). As in the case of Bahrain and Amman's new 'standard' dialects, the new 'standard' dialect of Baghdad originated in the rise in political importance of the community and the community's size but gradually lost its associations with any individual community.

In examining accommodation across the four major dialect groups in Arabic, it is noticeable that very little research has been conducted to date (Abu-Melhim 1991, 1992; S'hiri 2002). The most commonly held assumptions about communication across the major Arabic dialect groups seem to emanate from linguistic attitudes imbued in language ideology. A widely held belief among Arabic speakers is that → Educated Spoken Arabic (El-Hassan 1978; Meiseles 1980; Mitchell 1986), generally understood as a codified mixture of spoken and Modern Standard Arabic used by educated speakers of Arabic, is the 'third variety' that educated Arabic speakers resort to in contact situations. This assumption about the behavior of educated Arabic speakers is, however, challenged in the research delineated below. Besides, Modern Standard Arabic remains the prerogative of the educated few because the majority of the Arabic-speaking population are still illiterate with little or no competence in Modern Standard Arabic (Holes 1995b:3–5). A large portion of the population mobility across Arab borders described above is actually undertaken by unskilled laborers who have little or no education and therefore no proficiency in Educated Spoken Arabic.

Abu-Melhim investigates diglossic and bilingual code-switching in relation to linguistic accommodation between Jordanians and Egyptians (1991), and a wider number of Arabic speakers, from the Middle East (Egyptian, Lebanese, Jordanian, and Saudi) and from North Africa (Moroccan; 1992). His findings suggest that speakers of different national varieties of Arabic use a number of linguistic and paralinguistic strategies (e.g. repetition, paraphrasing, intonation, voice tone and quality, vocal stress) to communicate successfully with their interlocutors. The most common linguistic strategy, however, is → code-switching. Rather than switching entirely to Modern Standard Arabic, as posited by previous research, interlocutors mainly retained their colloquial varieties and code-switched into other Arabic varieties and into English, as well as Educated Spoken Arabic. According to Abu-Melhim's (1992) study, neither Modern Standard Arabic nor Cairene Arabic (the widest known among all spoken varieties) was the medium of choice in these contact situations. Holes (1995b:5) later asserts that "in normal face-to-face conversation, as opposed to writing…a blanket switch dialect to 'pure' MSA is rare indeed, even if it were within the ability of most Arabic speakers, and is a strategy which is resorted to only when all else fails". Holes further reports that crossdialectally, Arabic speakers use 'hybridized forms'. Abu-Melhim (1992) shows that when Moroccans and Iraqis converse, being the furthest apart linguistically and geographically, they use substantially more Educated Spoken Arabic and bilingual code-switching into English than they do with other interlocutors. The different varieties of Arabic from the East of the Arab world that Abu-Melhim studied seemed to be overall mutually comprehensible. Difficulties of comprehension were only reported with Moroccan Arabic, the only North African variety in the study. Overall, Classical Arabic was used very infrequently and merely in quotations in these data.

S'hiri's study (2002), which explores linguistic accommodation among speakers of Tunisian Arabic and speakers of Arabic from the Middle East (i.e. speakers of non-Maghrebi dialects such as Egyptian, Levantine, and Gulf Arabic), confirms Abu-Melhim's conclusion that, crossdialectally, Arabic speakers code-switch from their original national varieties to Modern Standard Arabic, to other prestige varieties, and to foreign languages. The Tunisian Arabic speakers in this study were journalists working at Arabic radio and television stations in London. The focus is on their interactions with their colleagues from the East of the Arab world, mainly from Egypt, Lebanon, and Palestine. The study found that Tunisians tend to converge unilaterally to their Eastern colleagues by avoiding colloquialisms and bilingual code-switching into French (a common practice among educated Tunisians, as noted in the studies presented above). Tunisians also adopt morphological and lexical elements from their colleagues' varieties of Arabic, even when successful communication is not at stake, in addition to code-switching into Modern Standard Arabic and English. The first explanation S'hiri proposes for this 'unnecessary' convergence, i.e. convergence that occurs where there is no risk of miscommunication, pertains to the imbalance in prior exposure to the Tunisian variety among Eastern Arabic speakers. While Eastern Arabs are barely acquainted with Tunisian Arabic, Tunisians are well versed, albeit passively, in Egyptian and Lebanese Arabic because of the long-term exposure to the latter's (especially Egyptian) films, songs, and soap operas. Second, S'hiri attributes 'unnecessary' convergent accommodation to Tunisians' experience and pride in multilingualism and openness to the other. Finally, she argues that this convergence bespeaks a linguistic insecurity on the part of Tunisians in an environment that questions their native speaker status and puts them on the margins of Arab identity. The Tunisian subjects interviewed in the study report that most of their Eastern colleagues believe their Tunisian variety to be 'corrupted' with Berber and French. This linguistic attitude indirectly puts in question these Tunisian journalists' ability to perform adequately at their job. Convergence to Eastern Arabic, even in a stereotypical manner, is their unconscious way to minimize differences with the Arabs of the 'center' and thereby claim proficiency and membership in the Arabic-speaking Nation (which meets the need of the 'in-group' to integrate into the 'out-group').

Unintelligibility is not the main cause for asymmetrical linguistic convergence of Tunisians to Eastern Arabic speakers. Instead, self-image, mutual perceptions, and sociopolitical and cultural relations seem to play a major role in determining the characteristics of linguistic interactions between these two groups. This convergent accommodation toward the Arabic of the East is yielding hybridized forms that are increasingly predictable and codified and may therefore signal an imminent shift in the Arabic of Tunisians.

In conclusion, it is worth noting that despite the insights offered by the above-referenced and other studies, the study of linguistic accommodation in Arabic remains a nascent field that is in need of greater attention.

Bibliographical references

Abu-Melhim, Abdel-Rahman Husni. 1991. "Code-switching and linguistic accommodation in Arabic". *Perspectives on Arabic linguistics*, ed. Bernard Comrie and Mushira Eid, 231–250. Amsterdam and Philadelphia: J. Benjamins.

——. 1992. *Communication across Arabic dialects: Code-switching and linguistic accommodation in informal conversational interactions*. Ph.D. diss., Texas A & M University.

Beebe, Leslie M. and Howard Giles. 1984. "Speech Accommodation Theories: A discussion in terms of second-language acquisition". *International Journal of Social Psychology* 46.5–32.

Bell, Allan. 1982. "Radio: The style of news language". *Journal of Communication* 32.150–164.

Blanc, Haim. 1960. "Style variation in spoken Arabic: A sample of interdialectal educated conversation." *Contributions to Arabic linguistics*, ed. Charles A. Ferguson, 79–161. Cambridge, Mass.: Center for Middle Eastern Studies, Harvard University.

——. 1964. *Communal dialects in Baghdad*. Cambridge, Mass.: Center for Middle Eastern Studies, Harvard University.

Bourhis, Richard Y. and Howard Giles. 1977. "The language of intergroup distinctiveness". *Language, ethnicity, and intergroup relations*, ed. Howard Giles, 119–135. London: Academic Press.

Coupland, Nikolas, Justine Coupland, and Howard Giles. 1991. *Language, society and the elderly: Discourse, identity and ageing*. Oxford and Cambridge, Mass.: Blackwell.

El-Hassan, S.A. 1978. "Educated Spoken Arabic in Egypt and the Levant: A critical review of diglossia and related concepts". *Archivum Linguisticum* 9:1.32–57.

Fraser, Colin and Klaus R. Scherer (eds). 1982. *Advances in the social psychology of language*. Cambridge: Cambridge University Press.

Genesee, Fred and Richard Y. Bourhis. 1982. "The social psychological significance of code switching in cross-cultural communication". *Journal of Language and Social Psychology* 1.1–28.

Giles, Howard. 1973. "Accent mobility: A model and some data". *Anthropological Linguistics* 15.87–105.

——, Anthony Mulac, James Bradac, and Patricia Johnson. 1987. "Speech Accommodation Theory: The first decade and beyond". *Communication Yearbook*, X, ed. Margaret McLaughlin, 13–48. Newbury Park, Cal.: Sage Publishers.

——. and Nikolas Coupland. 1991. *Language: Contexts and consequences*. Pacific Grove, Cal.: Brooks/Cole Publishing Company.

——, Justine Coupland, and Nikolas Coupland. 1991. *Contexts of accommodation: Developments in applied sociolinguistics*. Cambridge: Cambridge University Press.

Gudykunst, William B. (ed.). 1988. *Language and ethnic identity*. Clevedon and Philadelphia: Multilingual Matters.

Holes, Clive. 1986. "The social motivation for phonological convergence in three Arabic dialects". *International Journal of the Sociology of Language* 61.33–51.

——. 1995a. "Community, dialect and urbanization in the Arabic-speaking Middle East". *Bulletin of the School of Oriental and African Studies* 58:2.270–287.

——. 1995b. *Modern Arabic: Structures, functions, and varieties*. London: Longman.

Lawson-Sako, Sarah and Itesh Sachdev. 1996. "Ethnolinguistic communication in Tunisian streets: Convergence and divergence". *Language and identity in the Middle East and North Africa*, ed. Yasir Suleiman, 61–79. Richmond: Curzon Press.

——. and Itesh Sachdev. 2000. "Codeswitching in Tunisia: Attitudinal and behavioural dimensions". *Journal of Pragmatics* 32.1343–1361.

Meiseles, Gustav. 1980. "Educated Spoken Arabic and the Arabic language continuum". *Archivum Linguisticum* 11:2.118–148.

Mitchell, Terence F. 1986. "What is Educated Spoken Arabic?". *International Journal of the Sociology of Language* 61.7–32.

S'hiri, Sonia. 2002. "Speak Arabic please! Tunisian Arabic speakers' linguistic accommodation to Middle Easterners". *Language contact and language conflict in Arabic*, ed. Aleya Rouchdy, 149–174. London: Routledge Curzon.

Suleiman, Yasir. 2004. *A war of words: Language and conflict in the Middle East*. Cambridge: Cambridge University Press.

Thakerar, Jitendra N., Howard Giles, and Jenny Cheshire. 1982. "Psychological and linguistic parameters of Speech Accommodation Theory". *Advances in the social psychology of language*, ed. Colin Fraser and Klaus R. Scherer, 205–255. Cambridge: Cambridge University Press.

Walters, Keith. 1991. "Women, men, and linguistic variation in the Arab world". *Perspectives on Arabic linguistics*, III, ed. Bernard Comrie and Mushira Eid, 199–229. Amsterdam and Philadelphia: J. Benjamins.

Sonia Shiri (University of California at Berkeley)

Speech Acts

1. Speech act theory and modern pragmatics

Verbal communication as a form of human action has preoccupied European linguists, more conspicuously since the 19th century (Wilhelm von Humboldt, Karl Bühler, Roman Jakobson, and others), but the first elaborate theory of speech acts appeared only in the second half of the 20th century with the work of the English philosopher J.L. Austin, in a series of lectures that were included in the well-known work *How to do things with words* (1962). Reflecting on 'enunciation' as an essential process in verbal communication, Austin drew a fundamental distinction between 'constative' statements, which describe an action, an event, or a fact (e.g. *John came*), and → 'performative' ones, which not only describe an action but whose enunciation means the very accomplishment of the respective action (*I'm asking/commanding you to come*). The contents of constative statements can be subjected to a true/false test, while a performative phrase is neither true nor false: it is a phrase through which speakers do not just say but actually perform something. This category of the 'performative' drew the attention of other linguists who, without naming it as such (Benveniste 1966), or even calling it as such (Récanati 1981), brought important contributions to the study of this pragmatic concept.

Austin distinguishes between 'explicit performatives' and 'nonexplicit performatives', while acknowledging that the same phrase can be uttered performatively or constatively. In his view, it is more accurate to speak of simultaneous speech 'acts', which the speaker performs by uttering a sentence. He identifies three types of such acts: locutionary, illocutionary, and perlocutionary. A locutionary act consists in the articulation and combination of sounds so as to form meaningful words in the language, as well as grammatically well-formed syntactic combinations of the words. In an illocutionary act, the fact of uttering something has a certain action value in transforming the relations between the interlocutors, as in acts of promising or commanding where a promise or command is made by saying *I promise you that* or *I command you to*, both considered to be conventional illocutionary acts. A perlocutionary act aims at producing some effect on the interlocutor, which may or may not be perceived by the interlocutor, depending on a series of contextual factors and adequate emission/reception conditions of enunciation, referred to as 'felicity conditions'.

A few years after Austin's elaboration of the theory of speech acts, the American philosopher J.R. Searle dealt especially with the definition and analysis of the structure of illocutionary acts (1969) – 'indirect speech acts' in the terminology of Searle (1979) – and with their taxonomy (1975). Searle also elaborated related notions such as 'illocutionary force', 'conditions of success of illocutionary acts', 'sincerity conditions', and 'illocutionary commitments' (Searle and Vanderveken 1985). For him, "the minimal units of human communication are speech acts of a type called illocutionary acts", which he describes as follows (1985:1):

> Whenever a speaker utters a sentence in an appropriate context with certain intentions, he performs one or more illocutionary acts. In general an illocutionary act consists of an illocutionary force F and a propositional content P. For example, the two utterances 'You will leave the room' and 'Leave the room!' have the same propositional content, namely that you will leave the room; but characteristically the first of these has the illocutionary force of a prediction and the second has the illocutionary force of an order.

Illocutionary acts are, therefore, made in an appropriate context and with certain intentions, two points that are relevant for the Arabic grammarians' views on the subject.

Searle (1975; Searle and Vanderveken 1985) establishes five types of speech acts corresponding to the five types of 'illocutionary points' (1985:37–38):

i. assertives (to say how things are)
ii. commissives (to commit the speaker to doing something)
iii. directives (to try to get other people to do things)
iv. declaratives (to change the world by saying so)
v. expressives (to express feelings and attitudes)

These illocutionary forces are semantically named by illocutionary verbs, later called 'performative verbs', specifically "those verbs that imply an illocutionary point as part of their meaning" (Searle 1985:180). Searle discusses more than one hundred English illocutionary verbs (1985:182–216) corresponding to each of the five types of illocutionary acts. For assertives he includes such verbs as *to assert*, *to affirm*, *to rebut*, *to inform*, *to insist*, *to accuse*, *to blame*, *to praise*, *to complain*, *to lament*; for commissives, he includes *to commit*, *to promise*, *to swear*, *to accept*, *to consent*, *to refuse*, *to offer*, *to bid*, *to assure*; for directives, he includes *to direct*, *to ask*, *to request*, *to urge*, *to tell*, *to demand*, *to command*, *to order*, *to forbid*, *to permit*, *to suggest*, *to insist*, *to beg*, *to supplicate*, *to beseech*, *to implore*, *to pray*; for declaratives, he includes *to declare*, *to approve*, *to affirm*, *to confirm*, *to disapprove*, *to renounce*, *to disclaim*, *to repudiate*, *to bless*, *to curse*, *to name*, *to call*; for expressives, he includes *to apologize*, *to thank*, *to congratulate*, *to condole*, *to lament*, *to protest*, *to deplore*, *to compliment*, *to praise*, *to greet*.

2. Speech act theory in the Arabic grammatical tradition

Many linguists working within Arabic studies, especially since the 1980s, have focused on pragmatic concepts in traditional Arabic linguistic thought, particularly issues of 'meanings of speech' and 'speech acts' (see Frank 1981; Moutaouakil 1982; Bohas a.o. 1990; Larcher 1990, 1991, 1992, 1993; Versteegh 1997, 2004).

Arabic linguistic thought is based on the field of Qurʾānic exegesis. The early commentators aimed at explaining ambiguities in this text and were, therefore, not interested in the formal analysis of utterances. Their primary aim was "to find out God's intention and the categories they set up belong, therefore, to the realm of intentions of the speaker" (Versteegh 2004). Long before the great rhetoricians of the 7th and 8th centuries A.H., and following the traditions instituted by the exegetes, the classical grammarians before the 6th century A.H., which marked the beginning of the 'postclassical' period in Arabic linguistic thought, were interested in the forms of transmitting the communicative intention through different kinds of sentences. They studied the relation between the linguistic form or 'expression' and the intentional meaning transmitted through that expression. In the course of a few centuries, they elaborated a complex theory of meaning (→ *maʿnā*), dealing with a large number of issues related to speech acts.

Despite the terminological ambiguity of the polysemantic term *maʿnā*, which Sībawayhi uses in his analysis of speech, some trends in his approach are clear. In many passages of the *Kitāb*, he explains the grammatical form and syntax of different phrases on the basis of communicative intention. In other places (*Kitāb* I, 162ff.), he explains the syntactic structure of different expressions on the basis of the 'presence', in the deep structure, of a verb in the 1st person imperfect (the same method used by modern theoreticians, such as Austin, in their analysis of performative verbs). Dealing with statements containing a vocative (*Kitāb* I, 318ff.), Sībawayhi refers to a series of meanings which may be expressed by these, such as mourning, complaining (*nudba*), calling for help (*istiġāṯa*), threatening (*waʿīd*), menacing (*tahaddud*), extolling (*taʿḏīm*), boasting (*iftixār*), etc. Sībawayhi's discourse refers to indirect speech acts that may be realized by uttering statements in the vocative but which convey various meanings (see Buburuzan [Firanescu] 1993:421–437).

Early grammarians such as al-ʾAxfaš al-ʾAwsaṭ (d. 215/830 or 221/835) and Quṭrub (d. 206/821), set up a list of 'categories of speech', actually sentential types (*ʾaqsām al-kalām*), of which there are four (proposition, question, request, vocative), six (proposition, question, command, prohibition, vocative and wish), or even more, up till ten (see Versteegh 2004). Ibn Fāris (d. 395/1004), in his *Ṣāḥibī* (*Bāb maʿānī l-kalām* 'Chapter about the meanings of speech'), seems to have been the first to deal exhaustively with speech acts in a precise terminology. He distinguishes between a type of sentence that is characterized by the concordance between 'formal aspect/explicit meaning' and 'content/implicit meaning' and serves in performing direct speech acts; and a type that is not characterized by this concordance and serves to perform various indirect or illocutionary speech acts. Among the illocutionary speech acts, some are conventionalized, while others are nonconventional. A great variety of nonconventional illocutionary acts can be accomplished by uttering interrogative sentences (which express exclamation, overrating, reproach, affliction, blame, proposal, challenge, refutation, etc.) or imperative sentences (which institute or establish a reality, charge or commission, admiration, wish, engagement, etc). Besides a large number of examples of illocutionary acts, Ibn Fāris offers suggestions that could be of interest for the analysis of speech acts in the framework of modern pragmatics (for more details, see Buburuzan [Firanescu] 1995:103–114).

3. SPEECH ACT THEORY AND RHETORIC

Speech acts are studied from different perspectives in all disciplines dealing with language, or, as Moutaouakil (1982:162) notes:

> The phenomenon of the speech acts is treated in all disciplines (...) from perspectives which differ according to the preponderant interest of each discipline: the fundamentalists [i.e. the *'uṣūliyyūn*] are especially interested in the act of the 'order' (positive or negative) and in its various manners of expression; the grammarians center their analyses on the formal aspect of the acts, respectively 'affirmation', 'negation', 'interrogation', 'order', etc...conceived of as being essentially syntactic categories.

Rhetoric, however, is the field in which the largest number of notions linked with the theory of speech acts is found. In an important work entitled *'Asās al-balāġa* 'The foundations of rhetoric', the Muʿtazilite scholar az-Zamaxšarī (d. 539/1143) develops a theory of the metaphoric meaning (→ *majāz*) with regard to the proper or 'literal' meaning of sentences. With as-Sakkākī (d. 626/1229) and his encyclopedic work *Miftāḥ al-ʿulūm* 'The key of the sciences', the theory of speech acts reaches a new level. In the second chapter of this work, the author distinguishes within the field of rhetoric a separate discipline, the 'science of meanings' (*ʿilm al-maʿānī*), "a study which derived from the earlier work of linguists such as Sarrâj, Ibn Fâris, Thaʿâlibî, and especially Jurjânî, among others" (Owens 1988:243). In this discipline "would be treated all the questions relating to grammatical semantics and pragmatics" (Bohas a.o. 1990:118–119).

According to Versteegh (1997:263), "the meanings studied by *ʿilm al-maʿānī* are identical with the *maʿānī l-kalām*". He compares these to the notion of 'sentential types' or 'moods' and notes that "as-Sakkākī's interest concentrates on the study of sentential meanings in the sense of the structure of the pre-verbal message in which the speaker indicates his 'mood' towards his message" (Versteegh 1997:263). A similar point of view is expressed by Owens (1988:243), who emphasizes that "for Sakkâkî, *ʿilm al maʿânî* dealt primarily with the correlation between word order variation on the one hand and on the other the different sentential meanings associated with this variation, and the pragmatic implications of the different choices".

As-Sakkākī (for a detailed treatment of his linguistic thinking, see Simon 1993) deals with the opposition (based on logical principles) between two terms describing the two categories of speech: → *xabar* 'proposition; assertion; information' and *ṭalab* 'request; nonassertion'. The term *ṭalab* has an extended meaning, and this makes it possible to group together five forms of request or 'primary values' under two types: wish forms the first type of *ṭalab*, while question, order, interdiction, and call represent the second type.

Some notions relating to speech acts appear in the discourse of the rhetoricians when they discuss the exclamatory sentence, because this kind of statement is not entirely of the type 'proposition' nor that of 'request'. In order to explain this kind of sentence – and, generally, the sentence of the type *ṭalab* – as-Sakkākī deals with a concept named by him *tawallud* 'semantic engendering' (for more details, cf. Firanescu 2003:87–92). As-Sakkākī shows that "when these types of statements [the five primary types of *ṭalab*] stop being used with their initial meaning, we say that they generate new significations appropriate to the context". These 'new significations appropriate to the context' are called 'secondary meanings'. Many specialists in Arabic linguistics have recognized in the dichotomy *xabar/ṭalab* the opposition used by modern pragmaticians, in particular

Austin, between constative and performative speech acts, and in the mechanism of semantic 'engendering' they recognize what Searle (1975) calls 'an illocutionary derivation' (cf. Larcher 1998:107).

The long passage about this mechanism in as-Sakkākī obviously refers to indirect speech acts, which can be realized by uttering statements of the type *ṭalab* in a certain context (*maqām*). Two important terms in this connection are *maqām* 'context; contextual situation' and *qarīnat al-ḥal* (pl. *qarāʾin al-ḥāl*), used by the author in the sense of what is known in pragmatics and the theory of speech acts as 'conversational implicatures'. The *qarīnat al-ḥāl* is a concept that in as-Sakkākī seems to have passed beyond the embryonic stage (cf. Firanescu 2003:88); for the interpretation of the 'conversational implicatures' in as-Sakkākī's theory of semantic engendering, see Moutaouakil (1986:96ff., 1990:233–235). Semantic engendering in as-Sakkākī has two stages, which may be compared with the locutionary and illocutionary levels in modern pragmatics.

The following example from as-Sakkākī (*Miftāḥ* 305–306) offers a model of analysis in 'pragmatic' terms, which may be compared to the modern pragmatic theory of speech acts:

> If you say to a slave who insulted his master and to whom you have given his due punishment or whom you have threatened properly: 'Insult your master!', it is not possible that your intention be that of ordering the slave to insult his master, the situation being the above mentioned. By the help of the conversational implicature, [your command] takes the way/takes the direction of the sentence 'Mind whom you insult!' and, thus, the threat is generated.

At the locutionary level, in the 'contextual situation' (the sentence is uttered by the master and addressed to the slave who insulted him) or 'enunciation circumstances' (*maqām*), an 'obstruction' (*imtināʿ*) occurs in the transmission of the literal meaning of the sentence (or in the realization of the directive act performed by uttering the imperative sentence). So, at the illocutionary level, through the conversational implicature (*bi-maʿūna qarīnat al-ḥāl*), the result is a release of transmission that 'makes its way/takes the direction' (*yatawajjahu ʾilā naḥw*) toward the intentional meaning of the speaker (*al-murād*), and thus the act of threatening (expressive-directive) is engendered indirectly (*wa-tawallada min-hu t-tahdīd*).

Of the five subcategories of the 'request'-type sentences (*ṭalab*) enumerated by as-Sakkākī – wish, question, order, interdiction, and call (vocative) – the first one, when marked by the particle *layta*, serves in performing a direct speech act of wishing or desiring, in which the speaker does not claim that the achievement of the wish is possible. In Searle's terms, this is an 'expressive act'. But in the same sentential category of 'request', as-Sakkākī also introduces sentences preceded by the "particles that are useful to provoke the regret of the interlocutor and to urge or incite him" *(ḥurūf at-tandīm wa-t-taḥḍīḍ)*, i.e. *hallā, ʾalā, law lā, law mā*. He says that any statement containing one of these particles conveys the same meaning as a statement containing the particle *layta*, and, moreover, it conveys another meaning, that of stirring regret (*tandīm*). In pragmatic terms, it is understood that these marked sentences, of the formal type 'wish', are specialized in performing a conventionalized complex directive-expressive illocutionary speech act (cf. the taxonomy of illocutionary acts in Searle and Vanderveken 1985). The other four subtypes of 'request' acts can serve to perform some nonconventional illocutionary acts in the following way: the interrogative statement performs an unfulfilled desire, disapproval and opposition, disapproval and reprimand, threat and opposition; the imperative statement performs helplessness and defiance (orientated toward a person), or threat; the prohibitive statement performs threat; the vocative or 'call' statement performs temptation or stirring up.

The complexity of the nonconventional hybrid acts is pointed out by the author: they have a directive component, given by the general category in which they are enlisted, that of the performative 'to request', but they also have various components that are grafted onto the main component. This type of hybrid illocution, with several possible components, was discussed by Searle as well, and represents a topic that is studied in the framework of the modern theory of speech acts.

4. THE NOTION OF *ʾINŠĀʾ*

In Arabic linguistic thought, the term *ʾinšāʾ* 'performative' is surely the one that most clearly reflects Arabic linguistic thought on the concepts of 'performative' and 'speech acts'. This term has been treated exhaustively by Larcher

(1990, 1991, 1993, 1994, 1998). According to Larcher (1993:259), "this is not only a pluridisciplinary category, but transdisciplinary as well. It is found in grammar (*naḥw*), logic (*manṭiq*), rhetoric (*balāġa*), jurisprudence (*fiqh*), foundations of jurisprudence (*ʾuṣūl al-fiqh*), theology (*kalām*)". It seems that *ʾinšāʾ* 'performative', in Austin's (1962) sense, was used first by the jurists (*fuqahāʾ*), who contrasted it with *ʾixbār* 'information, assertion'. But as a term designating one of the two categories of speech (*kalām*) – the other being *xabar* 'assertion' – *ʾinšāʾ* was not used before the 7th or 8th century A.H. (Larcher 1993:260). The shift from the opposition *ʾixbār/ʾinšāʾ* to the opposition *xabar/ʾinšāʾ* is due to the legal scholars (*ʾuṣūliyyūn*) who, for practical reasons, analyzed mainly the utterances of the *ṭalab* 'request' type, concentrating on the 'imperative' (*ʾamr*) and 'prohibitive' (*nahy*) subtypes. The same legal scholars are also responsible for the incorporation of *ṭalab* into the more extended category of *ʾinšāʾ* (cf. Larcher 1993:260).

Most specialists in Arabic linguistics agree that, generally speaking, the term *ʾinšāʾ* in Arabic rhetoric refers explicitly to 'performative' utterances in opposition to 'constative' utterances (*xabar*) in Austin's (1962) terms. This opposition was interpreted also in terms of the opposition 'referential enunciation/the objective mood of speech' vs. 'nonreferential enunciation/the subjective mood of speech' (cf. Larcher 1991:257, 261).

Rhetoricians of the 8th century A.H. took over the opposition *xabar/ʾinšāʾ*, which they discuss, according to Bohas a.o. (1990:121), "in reference with some pragmatic notions as 'the requirements of the situation' (*muqtaḍā al-ḥāl*), or 'the situations of communication' (*maqāmāt al-kalām*)". The term 'utterance' (*kalām*), in the pragmatic acceptation, replaces in the rhetoricians' vocabulary the term 'statement' or 'sentence', with which the grammarians had operated. The process of communication is presented as the interaction between speaker (*mutakallim*) and addressee (*muxāṭab*).

The great rhetorician al-Qazwīnī (d. 739/1338) deals with the types of utterances and the opposition *xabar/ʾinšāʾ* systematically and concisely, detailing the values of the performative utterances and the typology of the illocutionary acts. In his *ʾĪḍāḥ*, he classifies performative utterances in two types: 'request' (*ṭalab*) and 'nonrequest' (*ġayr ṭalab*). In the case of the utterances of the type *xabar*, there is a "concordance between the expression and the requirements of the situation" (*muṭābaqat al-lafḏ li-muqtaḍā l-ḥāl*), but even in their case, in a certain discursive context (*maqām*), under certain conditions of uttering, which determine certain conversational implicatures (*qarāʾin*), these are not real assertions but rather illocutions with a marked expressive component: praise, insult/affront, humiliation, etc. In the case of the performative utterances of the type 'request', the condition is that the thing requested has not happened or has not yet been achieved at the moment of the request. The subtypes of the performative utterance 'request' (*ʾinšāʾ ṭalabī*) are: wish, interrogation, command, interdiction (Qazwīnī, *ʾĪḍāḥ* 98–108). Each of these acts is accomplished by uttering a sentence containing in its structure a specialized, established (*mawḍūʿ*; → *waḍʿ al-luġa*) particle, whose function is demonstrated by its use (*istiʿmāl*). In pragmatic terms, one could say that the author refers to conventionalized illocutions. But, in certain conditions of uttering, these utterances can be used to realize nonconventional illocutions. This is the case, for example, of the interrogative utterance marked by particles specialized in performing the interrogation, subtype of the request. But the interrogative particles "are usually used to express other meanings beside the interrogative one, meanings which correspond to a specific discursive context" (Qazwīnī, *ʾĪḍāḥ* 103). By uttering a marked interrogative sentence (at the locutionary level), various illocutions can be achieved, named by performative verbs, such as *to hurry*, *to press somebody*, *to exclaim*, *to warn* or *incite*, *to threaten*, *to order*, *to confirm*, *to deny*, *to contest*, *to reproach*, *to reprimand*, *to ironize*, *to permit*, *to deride*, etc. In some cases, al-Qazwīnī even indicates the perlocutionary level, more precisely the effect aimed at by speakers through the complex illocution they are performing, whose aim is to express the communicative intention and to achieve the effect. An utterance like *Do you forget that this man treated you well in the past?* is analyzed by the author (*ʾĪḍāḥ* 104) as an interrogation (locutionary) "serving to disapprove in the sense of reproach (illocutionary)", but "the purpose is to call the hearer's attention to the need to review his behavior, to be ashamed, not to do what he wanted to" (perlocutionary).

5. The communicative intention

Some rhetoricians use the term *ġaraḍ* (which seems to be equivalent to 'communicative intention', 'purpose', and 'expected effect') as a close parallel of the 'perlocutionary level' or 'perlocutionary effect'. This is the case, for instance, of the rhetorician Ḥāzim al-Qarṭājannī (d. 684/1285), who was at the same time a grammarian, a poet, and a critic. In his *Minhāj al-bulaġāʾ*, he deals with the expressive speech acts realized in poetical discourse. For him, the 'intentional meaning' (*qaṣd*) is subordinated to the scope, indicated by the term *ġaraḍ* (pl. *ʾaġrāḍ*), which designates both 'internal psychical act', achieved in the soul of the poet, and the 'perlocutionary effect on the receiver', which is supposed to be similar. These acts and effects (termed *ʾaġrāḍ ʾuwwal* 'primary acts/effects') are three: satisfaction, dissatisfaction, and intermediate [between satisfaction and dissatisfaction] states of minds or 'moods'. The subcategories of these primary acts, called *ʾanwāʿ* 'kinds, types', are in fact the principal illocutions (having a positive or negative connotation, following Aristotle's idea regarding the two poles of affectivity) that are realized in poetry: hope, aspiration, acceptance (contentment), and the wish to follow or learn something, on the positive side, and fear, rejection, fury, and amazement, on the negative side. Thus, the poetical genres (panegyric, erotic preamble, elegy, description, comparison, satire) represent the 'lexical vehicles' conveying the sense of the illocutions. Al-Qarṭājannī analyzes the relation between the poetical expression and the moods of poetical discourse, a relation that he views as interactive, and offers a detailed analysis of a variety of expressive illocutionary acts, proper to poetical works, such as begging, warning, demanding forgiveness, tempting, frightening, blaming, challenging (for more details, see Firanescu 2003:100–101, 124). It is obvious that these interpretations and suggestions in al-Qarṭājannī's work are of interest to modern research on literary pragmatics and especially speech acts in literature.

6. Later elaborations

In the works of later (post-13th century) Islamic theologians (*mutakallimūn*), speech act theory becomes more nuanced because of the theologians' constant effort to explain 'the indirect mood of speech' (*xurūj al-kalām ʿan ḥaqīqati-hi*) in the Qurʾānic text. Az-Zarkašī (*Burhān* II, 45), for instance, when discussing the interrogative sentence in the Qurʾānic text – which for theologians is a pseudo-interrogation because of God's 'prescience' (*maʿrifa qadīma*) – analyzes this type of sentence, finding a great number of indirect speech acts that are realized by uttering an interrogative sentence in a constative sense (*istifhām bi-maʿnā l-xabar*), e.g. reprimand/admonition, reproach/verbal punishment, rough reprimand, glorification, fright, repentance; or an interrogative sentence in a performative sense (*istifhām bi-maʿnā l-ʾinšāʾ*), e.g. interdiction, warning, stirring, hope, begging (of God), proposal and challenge, impatience, irony and mockery, humiliation, wonder, reprimand (cf. Firanescu 2003:153–157, 2004).

7. Conclusion

The framework of the speech acts is a part of the pragmatic perspective that was developed in Arabic traditional linguistic thinking. It is worth taking this framework into consideration because of the exceptional degree of subtlety and the many important suggestions regarding some aspects of speech behavior that are yet to be dealt with in modern studies.

Bibliographical references

Primary sources

Ibn Fāris, *Ṣāḥibī* = ʾAbū l-Ḥusayn ʾAḥmad Ibn Fāris, *aṣ-Ṣāḥibī fī fiqh al-luġa al-ʿarabiyya wa-sunan al-ʿArab fī kalāmi-hā*. Ed. ʿUmar Fārūq aṭ-Ṭabbāʿ. Beirut: Maktabat al-Maʿārif, n.d.
Qarṭājannī, *Minhāj* = Ḥāzim al-Qarṭājannī, *Minhāj al-bulaġāʾ wa-sirāj al-ʾudabāʾ*. Ed. Muḥammad al-Ḥabīb Ibn al-Xawāga. Tunis, n.d.
Qazwīnī, *ʾĪḍāḥ* = Jalāl ad-Dīn Muḥammad ibn ʿAbd ar-Raḥmān al-Qazwīnī (al-Xaṭīb), *al-ʾĪḍāḥ li-Muxtaṣar Talxīṣ al-miftāḥ*. Ed. *Šurūḥ at-Talxīṣ*. 4 vols. Cairo: Maṭbaʿa ʿĪsā al-Bābī al-Ḥalabī, 1937.
——, *Talxīṣ* = Jalāl ad-Dīn Muḥammad ibn ʿAbd ar-Raḥmān al-Qazwīnī (al-Xaṭīb), *at-Talxīṣ fī ʿulūm al-balāġa*. Ed. ʿAbd ar-Raḥmān al-Barqūqī. Beirut: Dār al-Kitāb al-ʿArabī, 1904.
Sakkākī, *Miftāḥ* = ʾAbū Yaʿqūb Yūsuf ibn ʾAbī Bakr Muḥammad ibn ʿAlī as-Sakkākī, *Miftāḥ al-ʿulūm*. Ed. Naʿīm Zarzūr. Beirut: Dār al-Kutub al-ʿIlmiyya, n.d.
Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. 5 vols. Ed. ʿAbd as-Salām Hārūn. Cairo: Dār al-Qalam etc., 1966–1977.

Zamaxšarī, *ʾAsās* = ʾAbū l-Qāsim Maḥmūd ibn ʿUmar az-Zamaxšarī, *ʾAsās al-balāġa*. Ed. ʿAbd ar-Raḥīm Maḥmūd. Beirut: Dār al-Maʿrifa, 1979.

Zarkašī, *Burhān* = Badr ad-Dīn Muḥammad ibn ʿAbdallāh ibn Bahādir az-Zarkašī, *al-Burhān fī ʿulūm al-Qurʾān*. Ed. Muḥammad ʾAbū l-Faḍl ʾIbrāhīm. 4 vols. Cairo: Maktaba Dār at-Turāṯ, n.d.

Secondary sources

Austin, John Langshaw. 1962. *How to do things with words*. London: Oxford University Press.

Benveniste, Émile. 1966. *Problèmes de linguistique générale*. Paris: Gallimard.

Bohas, Georges, Jean-Patrick Guillaume, and Djamel Eddine Kouloughli. 1990. *The Arabic linguistic tradition*. London: Routledge.

Buburuzan [Firanescu], Daniela Rodica. 1993. "Exclamation et actes de langage chez Sībawayhi". *Revue Roumaine de Linguistique* 38:5.421–437.

——. 1995. "Significations des énoncés et actes de langage chez Ibn Fāris". *Proceedings of the Colloquium on Arabic Linguistics, Bucharest, 29 August to 2 September 1994*, ed. Nadia Anghelescu and Andrei Avram, I, 103–114. Bucharest: University of Bucharest, Center of Arab Studies.

Firanescu, Daniela Rodica [see also Buburuzan]. 2003. *Exclamation in Modern Literary Arabic: A pragmatic perspective*. Bucharest: Editura Universităţii din Bucureşti.

——. 2004. "*Maʿānī l-kalām* chez al-Zarkašī: Interrogation et performance". *Romano-Arabica* N.S. 3.93–117.

Frank, Richard M. 1981. "Meanings are spoken of in many ways: The earlier Arab grammarians". *Le Museon* 94:3–4.259–314.

Larcher, Pierre. 1990. "Éléments pragmatiques dans la théorie grammaticale arabe postclassique". *Studies in the history of Arabic grammar*, II, ed. Michael G. Carter and Kees Versteegh, 193–214. Amsterdam and Philadelphia: J. Benjamins.

——. 1991. "Quand, en arabe, on parlait de l'arabe… II. Essai sur la catégorie de *ʾinšāʾ* (vs. *ḫabar*)". *Arabica* 38.246–273.

——. 1992. "Quand, en arabe, on parlait de l'arabe… III. Grammaire, logique, rhétorique dans l'islam postclassique". *Arabica* 39.358–384.

——. 1993. "Les arabisants et la catégorie de *ʾinšāʾ*: Histoire d'une 'occultation'". *Historiographia Linguistica* 20.259–282.

——. 1994. "Les *ʾAmālī* d'Ibn al-Ḥāǧib ou les *Annales* d'un grammairien". *Arabica* 41.273–280.

——. 1998. "Une pragmatique avant la pragmatique: 'Médiévale', 'arabe' et 'islamique'". *Histoire – Épistémologie – Langage* 20.101–116.

——. 2003. *Le système verbal de l'arabe classique*. Aix-en-Provence: Publications de l'Université de Provence.

Moutaouakil, Ahmed. 1982. *Réflexions sur la théorie de la signification dans la pensée linguistique arabe*. Rabat: Publications de la Faculté des Lettres et des Sciences Humaines de Rabat.

——. 1986. *Dirāsāt fī naḥw al-luġa al-ʿarabiyya al-waḍīfī* . Casablanca: Dār aṯ-Ṯaqāfa.

——. 1990. "La notion d'actes de langage dans la pensée linguistique arabe ancienne". *Studies in the history of Arabic grammar*, II, ed. Michael G. Carter and Kees Versteegh, 229–237. Amsterdam and Philadelphia: J. Benjamins.

Owens, Jonathan. 1988. *The foundations of grammar: An introduction to medieval Arabic grammatical theory*. Amsterdam and Philadelphia: J. Benjamins.

Récanati, François. 1981. *Les énoncés performatifs: Contribution à la pragmatique*. Paris: Minuit.

Searle, John R. 1969. *Speech acts*. Cambridge: Cambridge University Press.

——. 1975. "A taxonomy of illocutionary acts". *Language, mind and knowledge*, ed. Keith Gunderson, 344–369. (= *Minnesota Studies in the Philosophy of Science*, VII.). Minneapolis: University of Minnesota Press. [Repr., Searle (1979:1–29).]

——. 1979. *Expression and meaning*. Cambridge: Cambridge University Press.

—— and Daniel Vanderveken. 1985. *Foundations of illocutionary logic*. Cambridge: Cambridge University Press.

Simon, Udo Gerald. 1993. *Mittelalterliche arabische Sprachbetrachtung zwischen Grammatik and Rhetorik*: ʿIlm al-maʿānī *bei as-Sakkākī*. Heidelberg: Heidelberger Orientverlag.

Versteegh, Kees. 1997. "The Arabic tradition". Wout van Bekkum, Jan Houben, Ineke Sluiter, and Kees Versteegh, *The emergence of semantics in four linguistic traditions: Hebrew, Sanskrit, Greek, Arabic*, 227–284. Amsterdam and Philadelphia: J. Benjamins.

——. 2004. "Meanings of speech: The category of sentential mood in Arabic grammar". *Le voyage et la langue: Mélanges en l'honneur d'Anouar Louca et d'André Roman*, ed. Joseph Dichy and Hassan Hamzé, 269–287. Damascus: Institut Français du Proche-Orient.

Daniela Rodica Firanescu
(Dalhousie University)

Speech Errors

A speech error is an unintended deviation from the intended utterance. Speech errors can involve units from articulatory gestures of individual phonemes to entire phrases (→ slips of the tongue).

1. General introduction to speech errors

The easy and automatic production of speech makes the cognitive processes that are involved in speech production difficult to detect. The study of mistakes in speech production provides a window through which to view the mental organization of language and the processes that are used to turn language into speech (Dell 1995). The vast majority of speech error research has been conducted on English, and the received wisdom on speech error patterns is based on English data. Studies of speech errors in Arabic have highlighted similarities and differences with speech errors produced in other languages.

The most commonly occurring errors in English speech involve phonemes, words, and morphemes. These errors provide evidence that these structures are psychologically real and are used by speakers in the production of speech. Errors in speech production frequently involve the interaction between two units in the utterance plan, and such errors are also constrained by proximity and similarity (Fromkin 1971). Interactions are more common when the interacting units are similar and when they are located near one another in the utterance. For example, segmental errors frequently involve the confusion of onset segments between nearby words (e.g. [sɛɾɪŋ] for 'getting' in 'getting such bad luck'). In this case, the error is between phonemes in adjacent words, and the phonemes in error are in similar structural positions in the intended utterance. The units that interact in speech errors are usually units of the same type (e.g. noun, syllable onset).

Speech error data can be collected either opportunistically, as heard in normal conversation (e.g. Fromkin 1971; Dell and Reich 1980; Stemberger 1985), or experimentally through procedures designed to elicit speech errors (e.g. Baars, Motley, and MacKay 1975; Shattuck-Hufnagel 1992). While naturally occurring speech errors might better reflect the normal production of speech, the lack of control over the types of errors produced or the means by which the data are collected complicates a systematic analysis. Nonetheless, a number of researchers have collected corpora of naturally occurring speech errors in an attempt to discover the error processes commonly present in normal conversation. Experimental studies of speech errors involve testing particular hypotheses about the speech production process, based on systematic manipulation of the stimuli to be produced. Experimentally elicited speech errors generally focus on creating phonological segment errors and can be recorded for instrumental analysis (e.g. Frisch and Wright 2002). Recently, researchers have begun to use articulatory instrumentation to directly examine the activity of the articulators during speech error production, and they have found evidence that individual articulatory gesture errors occur (see, for example, Pouplier and Goldstein 2005).

2. Historical perspective on the study of speech errors in Arabic

The beginning of the systematic linguistic study of speech errors is usually attributed to Meringer and Mayer (1895/1978), who collected the first modern corpus of speech errors in German and analyzed their linguistic patterns. Using a similar corpus, Fromkin's (1971) seminal work applied speech error data to theoretical topics in modern generative linguistics. She used speech errors to argue for the reality of the abstract, underlying representations of generative grammar as part of the speech production process. However, Sami Anwar (1981:249) argues that the roots of speech error analysis in Arabic began "more than eleven centuries ago and that the study of speech errors helped greatly in the development of Arabic linguistic theory", which significantly predates these works.

In an overview of primary sources, Sami Anwar (1981) reports manuscripts dated between 858 and 1362 C.E. containing collections of errors and error analyses. These manuscripts cover a variety of topics, including phonological errors such as sound substitutions and metatheses; grammatical errors in inflection and derivation; semantic errors such as metonymy and antonymy; and errors in speech perception, writing, and reading. Some of the same generalizations presented by Fromkin (1971), concerning the proximity of interacting units and similarity between interacting units, were reached in these historical works. Sami Anwar (1981:253) states that the "development of Arabic studies of phonetics, grammar, lexicography and dialectology, as well as the writing system, owes a great deal to the interest of Arab linguists in speech errors". Given their motivation to explain speech errors, he claims, these early researchers were led to develop systematic structural descriptions of Arabic.

Some of these early studies are probably better classified as error analyses like those found today in applied linguistics research. Nonetheless, many of the studies appear to have had methods and goals comparable to current research on naturally occurring errors. In some other cases, the authors were motivated by a desire to eliminate errors in reading the *Qurʾān* based on ambiguities in the writing system that led to noncanonical interpretations. Another common type of study examined errors from dialectal and cultural perspectives (→ *laḥn*). Both of these types of study fall outside the current realm of speech error analysis, although once again there are some links. The early Arabic linguists were interested in potential relationships between speech errors and historical sound changes that might lead to dialectal and cultural differences. This idea is not unlike some modern proposals linking errors to historical sound change (e.g. Ohala 1981).

3. Modern studies of speech errors in Arabic

There have been very few modern studies of speech errors in languages other than English. There are four published papers on speech errors in Arabic in the literature, based on two corpora of errors. The major conclusions are summarized here.

Safi-Stagni (1990, 1994) collected a small corpus of approximately one hundred naturally occurring speech errors in Arabic. The corpus consisted of errors heard in social conversations, which were recorded by phonetic transcription. Safi-Stagni provides an overview of the errors observed and finds for the most part that the same sorts of error are observed in Arabic as have been found in other languages. In particular, phonological errors involving segments were common, occurring 45 times. These errors involved exchanges, anticipations, perseverations, and substitutions, all of which are found in English, German, and French speech errors. For example, the production of *sūfi sihām* for *šūfi sihām* ‘look at Sihām!’ involves an anticipation of the /s/ of the second word onto the first. In this particular case, the interaction between /s/ and /š/ is also typical of errors in English (Stemberger 1991). Safi-Stagni also observed 42 word-level errors that are found in other languages, including semantic substitutions, word exchanges, and word blends. For example, the production of *fī ṣāla fi ttilifōn* for *fī tilifōn fi ṣṣāla* ‘there is a telephone in the hall’ is a case of word exchange. In this case, there is also accommodation of the exchanged lexical items to their new morphological context, which is typical of errors in other languages as well. These segmental and word-level errors were the majority of the errors observed (87 out of 100), which is also typical of errors in other languages. The remaining errors included a few grammatical errors and some other mysterious productions that are not atypical of other error corpora.

Abd-El-Jawad and Abu-Salim (1987) examined a larger corpus of more than nine hundred naturally occurring errors. They observed many of the same errors as Safi-Stagni (1990), involving word and phoneme interactions, substitutions, and blends. They also found that word exchanges involved words of the same syntactic category, and that a semantically similar word frequently appeared in substitutions. Given the larger size of their corpus, they also observed errors not observed by Safi-Stagni. One such case is morpheme errors involving the consonantal roots. They observed exchange errors where the consonants were exchanged between words, with the vowel pattern left in place, for example *šaqar iḥsēn* for *ḥasan išqēr* ‘Hassan Shuqair’. This type of error was observed 57 times (6% of the corpus) and so cannot be dismissed as a bizarre error that only appears to involve the root consonant morpheme. Obviously, errors of this type have not been observed in other languages with a very different morphology than Arabic. In these cases, the exchange of roots also resulted in other prefixes, suffixes, or infixes being left in their original location, which is typical of errors in other languages. For example, the *-i* suffix remains in place in *mart ʾabūy-i* for *ʾabū marat-i* ‘father of my wife’. This example also shows another pattern that is typical of other languages, namely phonological accommodation of the erroneous roots to their new environment. In this example, the accommodation is specific to the weak noun morphology of Arabic as the intended /u/ in the utterance is followed by /y/ in the error to fit the new environment. Abd-El-Jawad and Abu-Salim took these errors to be strong evidence for the analysis of Arabic roots as a nonconcatenative morphological system with separate morphemes for the consonant root combined with grammatical morphemes containing vowels and consonant position slots (→ root).

Abd-El-Jawad and Abu-Salim (1987) also observed many cases of segment interaction errors within words and make some comments on the effects of word and syllable structure on errors, based on a comparison with statements by Fromkin (1971) and Boomer and Laver (1973) for patterns in other languages. Berg and Abd-El-Jawad (1996) conducted a more systematic comparison between a corpus of German errors and the error corpus of Abd-El-Jawad and Abu-Salim (1987), and so the discussion of the topic in this entry uses the later, more comprehensive analysis.

Berg and Abd-El-Jawad (1996) focus on the role of suprasegmental representation in speech production and hence error production. They examine the effects of word level structure and syllable structure on segmental errors in Arabic and German, and also in English in cases where Stemberger's (1985) corpus provided comparable data. Berg and Abd-El-Jawad (1996) first note a strong difference in the effect of word structure on errors. In English and German, errors between consonant phonemes within words are relatively uncommon, making up only about 10 percent of the consonant errors. In Arabic, within-word errors were found in about 80 percent of the consonant error cases. Berg and Abd-El-Jawad also examined whether errors typically occurred between consonants in parallel syllable structure positions (e.g. onsets with onsets). Almost all English and German errors that were between-words preserved syllable position (more than 95%), although errors involving changes in syllable position were found more commonly in within-word than between-word errors (80% preserved position). For Arabic, within-word errors frequently involved different syllable positions (38% preserved position), but between-words errors in Arabic preserved position more often, though not as often as in English and German (81% vs. more than 95%). Finally, it is more commonly the case in English and German that errors are made in syllable-onset position than in other positions, although this conclusion is primarily based on errors in word-onset position in between-word errors. Berg and Abd-El-Jawad (1996) demonstrate that this generalization holds for within-word errors in German. Errors in German involving onsets are more common than errors involving codas. In Arabic, by contrast, there seems to be no effect of position within the syllable. The likelihood that an error would involve an onset rather than a coda appeared to be no different from the overall likelihood of a consonant in a word being an onset.

Berg and Abd-El-Jawad (1996) interpreted the differences in segmental speech error frequency between Arabic and German as processing evidence for the root-and-pattern morphological system of Arabic. In particular, they claim that the lexical representation of Arabic consonant roots is very different from the representation of words in German or English. In German or English, the prosodic position of a consonant rarely changes in different morphological contexts. In Arabic, however, root consonants can appear in onset or coda position depending on the form of the vowel template for the particular verb form to be used. Given the difference in regularity of association between segments and syllable positions, they claim that associations between segments and syllable positions are stored in the lexicon in German and English, but not in Arabic. They claim consonantal roots in Arabic are stored without syllable position information and so can appear in errors in any syllable position with equal frequency. In addition, the lack of association of consonant roots to specific positions as part of the lexical representation also accounts for the high frequency of within-word errors. Consonants with less attachment to structural positions are assumed to be more vulnerable to speech errors where consonants appear in an unintended position. This is less likely in a language where consonants generally appear in the same place in a word each time the word is used.

4. CONCLUSION

The study of speech errors in Arabic fulfills the same role as the study of speech errors in other languages, providing evidence for the nature of language structure and processing. While overall factors in speech error production such as similarity and proximity are also found in Arabic, the pattern of speech errors is in some ways different from the pattern observed in other languages. Arabic has a root-and-pattern morphological system that is not found in the other languages that have been studied, and this system leads to a set of frequent errors in Arabic that are not found in other languages.

BIBLIOGRAPHICAL REFERENCES

Abd-El-Jawad, Hassan and Issam Abu-Salim. 1987. “Slips of the tongue in Arabic and their theoretical implications”. *Language Sciences* 9.145–171.

Anwar, Mohamed Sami. 1981. “The legitimate fathers of speech errors”. *Historiographia Linguistica* 8.249–265.

Baars, Bernard, Michael Motley, and Donald MacKay. 1975. “Output editing for lexical status from artificially elicited slips of the tongue”. *Journal of Verbal Learning and Verbal Behavior* 14.382–391.

Berg, Thomas and Hassan Abd El-Jawad. 1996. “The unfolding of suprasegmental representations: A cross-linguistic perspective”. *Journal of Linguistics* 32.291–324.

Boomer, Donald and John Laver. 1973. “Slips of the tongue”. *Speech errors as linguistic evidence*, ed. Victoria Fromkin, 120–131. The Hague: Mouton.

Dell, Gary S. 1995. “Speaking and misspeaking”. *An invitation to cognitive science*, ed. Lila Gleitman and Mark Liberman, 183–208. Cambridge, Mass.: MIT Press.

—— and Peter Reich. 1980. “Toward a unified theory of slips of the tongue”. *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand*, ed. Victoria Fromkin, 273–286. New York: Academic Press.

Frisch, Stefan A. and Richard A. Wright. 2002. “The phonetics of phonological speech errors: An acoustic analysis of slips of the tongue”. *Journal of Phonetics* 30.139–162.

Fromkin, Victoria. 1971. “The non-anomalous nature of anomalous utterances”. *Language* 47.27–52.

Meringer, Rudolf and Carl Mayer. 1978. *Versprechen und Verlesen: Eine psychologisch-linguistische Studie*. Amsterdam and Philadelphia: J. Benjamins. (First publ., Stuttgart: Göschen’sche Verlagsbuchhandlung, 1895.)

Ohala, John. 1981. “The listener as a source of sound change”. *Proceedings of the Chicago Linguistics Society: Papers from the Parasession on Language and Behavior*, 178–203. Chicago: Chicago Linguistics Society.

Pouplier, Marianne and Louis Goldstein. 2005. “Asymmetries in the perception of speech production errors”. *Journal of Phonetics* 33.47–75.

Safi-Stagni, Sabah. 1990. “Slips of the tongue in Arabic”. *Perspectives on Arabic linguistics*, I, ed. Mushira Eid, 271–290. Amsterdam and Philadelphia: J. Benjamins.

——. 1994. “Arabic segmental errors and segmental phonology”. *Perspectives on Arabic linguistics*, VI, ed. Mushira Eid, Vicente Cantarino, and Keith Walters, 169–194. Amsterdam and Philadelphia: J. Benjamins.

Shattuck-Hufnagel, Stefanie. 1992. “The role of word structure in segmental serial ordering”. *Cognition* 42.213–259.

Stemberger, Joseph P. 1985. *The lexicon in a model of language production*. New York: Garland.

——. 1991. “Apparent anti-frequency effects in language production: The addition bias and phonological underspecification”. *Journal of Memory and Language* 30.161–185.

STEFAN A. FRISCH (University of South Florida)

Speech Processing → Automatic Language Processing

Split → Phonological Split

Stem

Two broad perspectives underlie the research concerned with the question of what should be taken to be the set of basic and descriptively adequate notions needed for Arabic morphology. In a time-honored view, called here the root-based approach, words are formed from (consonantal) → roots, sequences of consonants identifying a common invariant among various related word forms. In the other, less widely explored view, the stem-based approach, words are formed from stems, i.e. forms that may consist of vowels as well as prosodic features such as vocalic or consonantal length. In this view, consonantal roots are considered to be a by-product or an emergent property of the organizational principles in the linguistic grammar.

This entry presents the stem-based view. Since linguistic morphology is primarily concerned with systems of relations between words, it should first be noted that, as in other languages with rich inflectional morphology, Arabic organizes words in paradigms. These can be described as sets of words built from combinations of stems with inflectional markers, the latter designating various morphosyntactic categories. As an example, consider a fragment of the Arabic verbal paradigm in Table 1. The Arabic verb is described as having two sets of forms or ‘Tense/Aspect’ categories, known as the imperfect and the perfect. The words in Table 1 illustrate the indicative mood of the imperfect aspect of the lexeme ‘to write’. Fully inflected words are formed by placing /ktub/ in the context of the appropriate prefix—suffix pair. These pairs consist of {ya—u, ta—u, ta—u, ta—īna, ʾa—u, ya—ūna, ya—na, ta—ūna, ta—na, na—u, ya—āni, ta—āni, ta—āni, ta—āni}, the exponents of the morphosyntactic categories of Mood (Indicative), Person (1st, 2nd, 3rd), Number (singular, plural, dual), and Gender (masculine, feminine). In this article, the set of phonological forms created from the exponents of the morphosyntactic categories of a particular paradigm is referred to as the inflectional context of that paradigm, and the form /ktub/ is referred to as the verbal stem. To avoid ambiguity, the term 'stem' refers to that phonological form of a lexeme to which an affix is attached. This sense of 'stem' is essentially the same as that assumed in modern lexeme-based theories of morphology such as those of Matthews (1972), Aronoff (1992), and Anderson (1992), and consequently it is not specific to Arabic or Arabic-like morphologies.

Table 1. Imperfect, indicative of *kataba* 'to write'

		singular	plural	dual
3	masc.	*ya-ktub-u*	*ya-ktub-ūna*	*ya-ktub-āni*
	fem.	*ta-ktub-u*	*ya-ktub-na*	*ta-ktub-āni*
2	masc.	*ta-ktub-u*	*ta-ktub-ūna*	*ta-ktub-āni*
	fem.	*ta-ktub-īna*	*ta-ktub-na*	*ta-ktub-āni*
1		*ʾa-ktub-u*	*na-ktub-u*	

The vowel of the verbal stem must be lexically specified: [ja-ktub-u] 'he writes', [ja-lbas-u] 'he dresses', [ja-ḍrib-u] 'he hits'. There are also a few minimal pairs of stems distinguished solely on the basis of this vowel: [ja-smar-u] 'he is brown' vs. [ja-smur-u] 'he spends the night conversing', and [ja-ħzun-u] 'he saddens' vs. [ja-ħzan-u] 'he is sad'. Given these facts, some researchers have proposed that the verbal morphology of triliterals builds on the stem /ktub/ (Schramm 1962, 1991:1403; Kuryłowicz 1972:34, 43; McOmber 1995:179; Ratcliffe 1998:33; Benmamoun 1999:176, among others). For example, Kuryłowicz (1972:43) writes,

> The fundamental form of the Sem. conjugation, the so-called 'imperfect(ive)' yaqtul(u), shows a characteristic vowel after (R2) which is unpredictable, i.e. independent of any grammatical rule, hence basic. Therefore the verbal root is not a consonantal skeleton (q-t-l), but contains an essential vocalic component (u of qtul).

Put in present terms, Kuryłowicz's view consists of the claim that the verbal morphology is stem-based. This may be a plausible hypothesis, but it is not the generally accepted view in Arabic linguistics. For verbs, traditionally, morphology is assumed to operate on the consonantal root (see Cantineau 1950; Fleisch 1956; McCarthy 1979; Yip 1988; Hoberman 1988; Goldenberg 1994, among others; and Goldenberg 1994 and Hoberman 1995 for two recent reviews). As Schramm (1991:1402) writes,

> The conventional statement of Semitic morphological typology for the last thousand years or so has always reflected the view that all verbs and most nouns are to be derived by a process of interdigitating discontinuous consonantal root morphemes, expressing lexical content, and vocalic pattern morphemes which express grammatical content.

For nouns, in contrast, the stem-based hypothesis has made significant contributions to the understanding of the lawful relationships between noun forms. Some important studies on Arabic singular/plural morphology (→ number), in particular, have established that surface properties of the noun stem such as vocalic and consonantal length condition in crucial ways the form of the corresponding plural form (see Hammond 1988; McCarthy and Prince 1990, and references therein).

In recent work, however, the stem-based view for Arabic verbal morphology has witnessed more systematic development, for instance by McCarthy (1993), McOmber (1995), Ratcliffe (1998, Chap. 2), Benmamoun (1999), and Gafos (2003) for Classical Arabic. For notable examples of the stem-based view for modern Arabic dialects, see Cowell's (1962) grammar of Syrian Arabic and Heath's (1987) monograph on Moroccan Arabic. In what follows, some of the virtues of the stem-based approach to Classical Arabic verbal morphology are sketched, starting with the set of facts related to doubled verbs, also known as biconsonantal or geminated verbs, whose explanation has consistently relied on the root-based view. In the perfect, geminated verbs show two allomorphs, [madd] and [madad], as shown in Table 2 for the lexeme 'to stretch'. Henceforth, [madd] will be called the GEMINATE allomorph and [madad] the STRONG allomorph – strong due to its resembling the nonalternating, so-called strong verbs like [katab] 'to write'.

The distribution of the allomorphs is also shown on the left in Table 2. The geminate

Table 2. Perfect of *madda* 'to stretch'

		singular	plural	dual	allomorph distribution
3	masc.	*madd-a*	*madd-ū*	*madd-ā*	
	fem.	*madd-at*	*madad-na*	*madd-atā*	Geminate [madd] /__V
2	masc.	*madad-ta*	*madad-tum*	*madad-tumā*	Strong [madad] /__C
	fem.	*madad-ti*	*madad-tunna*	*madad-tumā*	
1		*madad-tu*	*madad-nā*		

Table 3. Imperfect, indicative of *madda* 'to stretch'

		singular	plural	dual
3	masc.	*ya-mudd-u*	*ya-mudd-ūna*	*ya-mudd-āni*
	fem.	*ta-mudd-u*	*ya-mdud-na*	*ta-mudd-āni*
2	masc.	*ta-mudd-u*	*ta-mudd-ūna*	*ta-mudd-āni*
	fem.	*ta-mudd-īna*	*ta-mdud-na*	*ta-mudd-āni*
1		*ʾa-mudd-u*	*na-mudd-u*	

allomorph occurs before vowels, and the strong allomorph before consonants. Which one of these two allomorphs underlies the alternation? One answer to this question is suggested by the following observation. The shape of [madad] is the shape of the nonalternating triliteral verbs like [katab-a] 'he wrote', [katab-tu] 'I wrote'. Conventionally, triliterals are assumed to be the 'canonical' verbs in Arabic, and by extrapolation rather than logical necessity, their shape is assumed to be the canonical shape for verbs. This is a widespread assumption (see Wright 1896:68–71; Cantineau 1946:133; Brame 1970:119; McCarthy 1979:265–267). Specifically, this assumption implies an analysis that consists in the following steps. The root /md/ first assumes the shape of a CvCvC sequence. Because the root consists of only two consonants, its final consonant /d/ extends to occupy two positions, hence /madad/. This intermediate form is then converted to [madd] before a vowel-initial suffix, via a process of syncope, as in /madad+V/ #⇒# [maddV], and in some cases via a process of metathesis as in (the imperfect) /ya+mdud+V/ #⇒# [jamuddV]. However, as McCarthy (1986:247–248) observes, this analysis treats the alternation as 'morpholexical' in character. There does not seem to be any reason why /madad/ should change to [madd] or why /ya-mdud-/ should change to [ja-mudd-]. If this alternation were phonological, it would falsely predict that /katab-/ should change to [katb-] and /ya-ktub/ should change to [ja-kutb-] before a vowel. It can thus be seen that the alternation, as formulated in the syncope/metathesis rule, is arbitrary in the sense that there is no phonological motivation for the particular form that this alternation takes.

However, there is an alternative (Gafos 2003): the underlying stem is /madd/, and [madad] is a surface variant of /madd/. Surprisingly, this alternative has not been pursued. It is standard methodology in generative grammar that, given an alternation like [madd] ~ [madad], we consider at least the two hypotheses outlined above, and contemplate their consequences for the rest of the grammar. If /madd/ is the basic verbal stem, then suffixation with a vowel-initial suffix gives [madd-a], an attested form. Suffixation with a consonant-initial suffix, however, results in an illicit triconsonantal sequence, */madd-tu/. As in many other languages with geminates, Arabic bans geminates from syllable codas (for syllabification in Arabic, see Angoujard 1988; Broselow 1992; Itô 1986; Farwaneh 1995). The illicit consonant sequence is therefore split to satisfy syllabification, [madad-tu]. There is no need for morphological stipulation or intermediate, unmotivated steps. The alternation is driven by pure phonotactic canons or constraints that govern the admissible sequencing of phonemes in the language.

The geminated verb allomorphy is also found in the imperfect. As Table 3 shows, the conditioning of the two allomorphs is the same as that in the perfect. The geminate allomorph occurs before vowels, the strong elsewhere: [ja-mudd-u], [ja-mdud-na]. The forms in Table 3 illustrate the indicative mood of the imperfect aspect of the lexeme 'to stretch'. The other verbal moods built on the imperfect stem (subjunctive, jussive, imperative, and the rare → energicus) are in all relevant respects similar to the indicative. That is, prefixes are vowel-final and suffixes are vowel-initial, consonant-initial, or null. Moreover, the geminated verb alternation in these moods is identical to that found in the indicative (e.g. jussive 3rd pers. masc. sg. [ja-mdud], 3rd pers. fem. sg. [ta-mudd-a]).

If we assume that /mudd/ is the stem, the alternation follows the same pattern as in the perfect: in combination with a V-initial suffix, the stem surfaces as in [ya-mudd-u], but with a C-initial or null suffix, a geminate coda would result, *[ya-mudd(-na)]. The ban on coda geminates enforces alternation to [ya-mdud(-na)].

Up to now, discussion of verbal allomorphy has been confined to Form I of the Arabic verb. If, as argued, the allomorphy is due to phonological principles rather than morphologically conditioned idiosyncrasies of certain forms, then it is predicted that the alternation will be found whenever its phonological conditions are met. This prediction is confirmed. The alternation is also met in verbs of Form IX, XI, and QIV under conditions identical to Form I. Traditionally, Form IX is identified with the pattern *ktabab* (Wright 1896:43). As far as known, all subsequent work in the generative tradition has assumed that *ktabab* is the canonical Form IX of verbs. However, stems in Form IX surface as *ktabab* only before consonant-initial suffixes, for reasons familiar by now. Representative examples of the alternation are given under the verbal part of Table 4.

Form IX verbs like [i-ħmarr-a] 'he blushed' are related to adjectives of color and bodily defects, here [ʔaħmar-u] 'red' and its corresponding nominal form [ħumr-un] 'red.plural' (for an illustration of exactly this morphology in a modern dialect, consider Form IX verbs from Syrian, e.g. [ʔaħmar] 'red' ~ [ħmarr] 'to blush', [ʔas̴far] 'yellow' ~ [s̴farr] 'to turn pale', and so on; Cowell 1962:101, 250). In the verbal form [i-ħmarr-a], putting aside the transparently epenthetic [i-] and the suffix [-a], the final consonant is the long version of its corresponding segment in the noun or the adjective. There are a few different ways to state the morphological link between the verb and its derivationally related forms. One such way is to derive the verbal stem by adding a suffixal mora (μ) to the simpler stem /ḥmVr/, underlying the noun or the adjective: /ḥmVr-/$^{\text{Stem}}$ /+μ #⇒# /ḥmarr-/$^{\text{Verb-stem}}$. What is important for current purposes is that once the verbal stem is placed in its paradigm, it is clear that what is involved in the allomorphy [i-ħmarr-a] ~ [i-ħmarar-tu] is the by now familiar phonologically determined alternation.

The same alternation applies to quadriliteral verbs in Form QIV, [i-ʃmaʕall-a] 'he hastened', [i-ʃmaʕlal-tu] 'I hastened' and [i-t̴maʔann-a] 'he was tranquil', [i-t̴maʔnan-tu] 'I was tranquil'. These examples are all perfect forms. In the imperfect, the same alternation is found, e.g. [ja-t̴maʔinn-u] and [ja-t̴maʔnin-na] 'he/they [fem.] are tranquil' (Schramm 1962:362). The conditions for the alternation and the form that this alternation takes are identical throughout.

Another systematic property of the verbal system is that all verbal moods are based on the imperfect form CCvC, Kuryłowicz's 'fundamental form'. The indicative is shown in Table 1. The subjunctive differs from the indicative in superficial ways that do not affect the ensuing discussion. The jussive and the imperative are presented in Table 5 and Table 6, respectively. In the rare variant of the jussive, the → energicus (Schramm 1962:364), the affixes are in all relevant respects similar to the other moods, that is, all prefixes are vowel-initial and suffixes are vowel- or consonant-initial.

Table 4. Alternation in Form IX (Perfect)

Adjectives		Verbal alternation in Form IX (perfect)		
ʾa-ḥmar-u	'red'	*i-ḥmarr-a*	*i-ḥmarar-tu*	'he/I blushed'
ʾa-ṣfar-u	'yellow'	*i-ṣfarr-a*	*i-ṣfarar-tu*	'he/I became yellow'
ʾa-qbal-u	'cross-eyed'	*i-qball-a*	*i-qbalal-tu*	'he/I became cross-eyed'
cf. Verbs		Form I (perfect)		
/radd/	'to return'	*radd-a*	*radad-tu*	'he/I returned'

Table 5. Jussive of *kataba* 'to write'

		singular	plural	dual
3	masc.	*ya-ktub*	*ya-ktub-ū*	*ya-ktub-ā*
	fem.	*ta-ktub*	*ya-ktub-na*	*ta-ktub-ā*
2	masc.	*ta-ktub*	*ta-ktub-ū*	*ta-ktub-ā*
	fem.	*ta-ktub-ii*	*ta-ktub-na*	*ta-ktub-ā*
1		*ʾa-ktub*	*na-ktub*	

Table 6. Imperative of *kataba* 'to write'

		singular	plural	dual
2	masc.	*u-ktub*	*u-ktub-ū*	*u-ktub-ā*
2	fem.	*u-ktub-ī*	*u-ktub-na*	*u-ktub-ā*

In the remaining discussion, an important question is addressed that has not been addressed so far by the proponents of stem-based morphology. Observe that there is no contrast between [ja-CCvC-u] and [ja-CvCC-u]. Both [ja-CCvC-u] and [ja-CvCC-u] are phonologically well formed, but only the former is attested. Whence the [CCvC] invariance of the fundamental form?

This question turns out to have a simple answer when we take into account some independent properties of the language. First, Arabic does not allow complex syllable onsets or codas. Second, whereas all prefixes in Tables 1, 5, and 6 end in vowels, some suffixes begin with a consonant or are null. Thus, a /CvCC/ stem would raise a phonotactic problem before a consonant-initial or null suffix, since *[Cv-CvCC-Cv] is banned. A /CCvC/ stem presents no phonotactic problem because prefixes end in vowels. A [Cv-CCvC-Cv] is permissible because the first stem consonant can be parsed as a coda. Hence, we can begin to see how the inflectional context coupled with phonotactics requires that the CC cluster be at the left edge of the stem.

It is instructive to contrast this approach to a well-known alternative. To account for the lack of *[ja-CvCC-u] or, equivalently, the lack of contrast between [ja-CCvC-u] and *[ja-CvCC-u], the lexicon is restricted to include only /CCvC/. The lexicon is thus preconfigured so that [ja-CvCC-u] surface forms cannot arise, and this is done by imposing a restriction on the set of admissible grammar inputs or a 'morpheme structure constraint' (Chomsky and Halle 1968), as is commonly known. Conceptually, this approach is quite different from the one argued for presently, which seeks to derive the observed pattern as the lawful consequence of systemic factors, here, the inflectional context and phonotactics. Saying that the pattern is derived means that there is no unique locus in the grammar or the lexicon where the ban against [CvCC] forms or /CvCC/ stems is stated. Rather, it is the interaction of a few independent factors that effectively bans these forms (see Kisseberth 1970 on phonotactic 'conspiracies' and apparent constraints on inputs).

A more important reason in support of the proposed model derives from its predictive power. Observe that morpheme structure constraints do not make any predictions beyond their highly specific assertions, e.g., there is no /CvCC/ verbal stem in the Arabic lexicon. The model promoted here instead employs general principles in a theory of grammar, and consequently makes predictions beyond specific data. The stem-in-paradigm approach predicts that in a different paradigm with vowel-initial suffixes /CvCC/, stems would be possible. The example needed to test this prediction is provided by the morphology of the noun. As shown in Table 7, the inflectional context for nouns consists of vowel-initial suffixes ([*stem*-un] in the indefinite, [ʔal-*stem*-u] in the definite). It is thus expected that the /CvCC/ stem banned in the verb should now be possible in the noun. This is indeed the case, as shown by a few representative forms from the well-populated class of triliteral nouns, [nafs-un] 'soul', [baħr-un] 'sea', [qufl-un] 'lock', [burd-un] 'robe', and so on.

Table 7. Noun endings

	masc. sg.	masc. pl.	fem. sg.	fem. pl.
nominative	- *un*	- *ūna*	- *atun*	- *ātun*
genitive	- *in*	- *īna*	- *atin*	- *ātin*
accusative	- *an*	- *īna*	- *atan*	- *ātin*

To review, two related specific ideas are promoted here. The first is that we can make sense of the alternation between /madd/ and /madad/ if we assume that the verbal stem is /madd/. This is the basic form of the stem on which inflectional affixes are attached. The other surface form of a doubled verb, /madad/, results by splitting the geminate of /madd/ when that form is combined with a consonant-initial or null suffix. This happens because Arabic does not permit geminate codas as in *[madd-tu] or [madd]. The geminated verb alternation is not arbitrary, in the sense of being morphologically conditioned, and there is no need for rules that sometimes result in metathesis and sometimes deletion of vowels. The second related idea is that the inflectional structure of the paradigm coupled with phonotactics provides a powerful source of constraints on the theoretically possible diversity of stem forms within that paradigm. The paradigm molds stems to fit the inflectional context of their realizations. This allows us to explain why certain stem shapes are found while others are not attested.

In sum, this entry presents the hypothesis that Arabic morphology, the system of lawful relations between words, must have access to more richly specified underlying representations than is allowed by consonantal roots. Specifically, core areas of the verbal morphology require reference to stems specified for properties such as vocalism and consonantal length, e.g. /ṣubb/ 'to pour [liquid]', /ḥabb/ 'to love passionately', /ṣal/ 'to arrive', which are not admissible as part of consonantal roots. It is at the stem level where generalizations about the morphology and phonology of the Arabic verb can be observed and stated in the form of a testable theory.

Bibliographical references

Anderson, Stephen. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press.

Angoujard, Jean-Pierre. 1988. "Gémination et redoublement". *Langues Orientales Anciennes Philologie et Linguistique* 1.1–15.

Aronoff, Mark. 1992. "Stems in Latin verbal morphology". *Morphology now*, ed. Mark Aronoff, 5–32. Albany: State University of New York Press.

Benmamoun, Elabbas. 1999. "Arabic morphology: The central role of the imperfective". *Lingua* 108.175–201.

Brame, Michael. 1970. *Arabic phonology: Implications for phonological theory and historical Semitic*. Ph.D. diss., Massachusetts Institute of Technology.

Broselow, Ellen. 1992. "Parametric variation in Arabic dialect phonology". *Perspectives on Arabic Linguistics* IV, ed. Ellen Broselow, Mushira Eid, and John McCarthy, 7–45. Amsterdam and Philadelphia: J. Benjamins.

Cantineau, Jean. 1946. "Esquisse d'une phonologie de l'arabe classique". *Bulletin de la Société Linguistique de Paris* 43.93–140.

——. 1950. "Racines et schèmes". *Mélanges William Marçais*, 119–124. Paris: Maisonneuve.

Chomsky, Noam and Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.

Cowell, W. Mark. 1962. *A reference grammar of Syrian Arabic based on the dialect of Damascus*. Washington, D.C.: Georgetown University Press.

Eid, Mushira (ed.) 1995. *Perspectives on Arabic Linguistics*, VII. Amsterdam and Philadelphia: J. Benjamins.

Farwaneh, Samira. 1995. *Directionality effects in Arabic dialect syllable structure*. Ph.D. diss., University of Utah.

Fleisch, Henri. 1956. *L'arabe classique: Esquisse d'une structure linguistique*. Beirut: Imprimerie Catholique.

Gafos, Adamantios. 2003. "Greenberg's Asymmetry in Arabic: A consequence of stems in paradigms". *Language* 79.317–355.

Goldenberg, Gideon. 1994. "Principles of Semitic word-structure". *Semitic and Cushitic studies*, ed. Gideon Goldenberg and Shlomo Raz, 29–64. Wiesbaden: O. Harrassowitz.

Goldsmith, John. 1995. *Handbook of phonological theory*. Oxford: Blackwell.

Hammond, Michael. 1988. "Templatic transfer in Arabic broken plurals". *Natural Language and Linguistic Theory* 6.247–270.

Heath, Jeffrey. 1987. *Ablaut and ambiguity: Phonology of a Moroccan Arabic dialect*. New York: State University of New York Press.

Hoberman, Robert. 1988. "Local and long-distance spreading in Semitic morphology". *Natural Language and Linguistic Theory* 6.541–549.

——. 1995. "Current issues in Semitic morphology". Goldsmith (1995:839–847).

Itô, Junko. 1986. *Syllable theory in prosodic phonology*. Ph.D. diss., University of Massachusetts, Amherst. (Publ., New York: Garland Publishers, 1989.)

Kaye, Alan S. (ed.). 1991. *Semitic studies in honor of Wolf Leslau on the occasion of his eighty-fifth birthday*. Wiesbaden: O. Harrassowitz.

Kisseberth, Charles. 1970. "On the functional unity of phonological rules". *Linguistic Inquiry* 1.291–306.

Kuryłowicz, Jerzy. 1972. *Studies in Semitic grammar and metrics*. London: Curzon Press.

Matthews, P.H. 1972. *Inflectional morphology*. Cambridge: Cambridge University Press.

McCarthy, John J. 1979. *Formal problems in Semitic phonology and morphology*. Ph.D. diss., Massachusetts Institute of Technology. (Publ., New York: Garland Publishers, 1985.)

——. 1986. "OCP effects: Gemination and antigemination". *Linguistic Inquiry* 17.207–263.

——. 1993. "Template form in prosodic morphology". *Papers from the Third Annual Formal Linguistics Society of Midamerica Conference*, ed. L. Smith Stvan, 187–218. Bloomington, Ind.: IULC Publications.

—— and Alan Prince. 1990. "Foot and word in prosodic morphology: The Arabic broken plurals". *Natural Language and Linguistic Theory* 8.209–282.

McOmber, Michael L. 1995. "Morpheme edges and Arabic infixation". Eid (1995:173–189).

Ratcliffe, Robert R. 1998. *The 'broken' plural problem in Arabic and comparative Semitic: Allomorphy and analogy in Arabic and comparative Semitic*. Amsterdam and Philadelphia: J. Benjamins.
Schramm, Gene M. 1962. "An outline of classical Arabic verb structure". *Language* 38.360–375.
——. 1991. "Semitic morpheme structure typology". Kaye (1991:1402–8).
Wright, William. 1896. *A grammar of the Arabic language*. 3rd ed. Cambridge: Cambridge University Press.
Yip, Moira. 1988. "The Obligatory Contour Principle and phonological rules: A loss of identity". *Linguistic Inquiry* 19.65–100.

ADAMANTIOS I. GAFOS (New York University)

Stress

1. INTRODUCTION

All modern Arabic dialects have word stress: one syllable in a word is prominent as compared to other syllables. The position of the strong stress is usually limited to the last three syllables of the word, and depends on the overall pattern of short and long syllables. Compare Cairene colloquial *kiˈtaab* 'book', *kaˈtabt* 'I wrote', *kaˈtabna* 'we wrote', *ˈkatabu* 'they wrote', *kataˈbitu* 'she wrote it'. Word stress is not distinctive in Arabic, that is, it does not serve to distinguish meanings, although the morphological structure of words often affects stress, compare Cairene *raˈmitu* /ram-it u/ 'she threw it' vs. *ˈkatabu* /katab-u/ 'they wrote'.

Arabic stress patterns have properties which are crosslinguistically common in stress languages: each word has at least one prominent syllable (the *cumulative* property), which is located near the beginning or the end of words (the *demarcative* property). Stress placement is sensitive to long and short syllables ('quantity-sensitivity'), while longer words show alternating patterns of strong and weak syllables, at least in some dialects. The phonetic realization of stress in Arabic is also crosslinguistically common: stressed syllables have overall higher pitch levels, longer duration, and greater loudness than unstressed syllables (Al-Ani 1992a, b).

From a linguistic point of view, the main interest of Arabic stress resides in the patterns of variation among the dialects, which show a remarkable unity in diversity. Dialects share a number of basic patterns, such as stressing superheavy syllables (CvCC or CvvC) in ultimate position (*kiˈtaab* 'book', *kaˈtabt* 'I wrote'), and stressing heavy syllables (CvC or Cvv) in penultimate position (*kaˈtabna* 'we wrote', *kiˈtaabi* 'my book'). However, dialects differ in their stressing of certain word types. For example, words with a heavy antepenult followed by two light syllables are stressed on the penult (*madˈrasa* 'school', Cairo) or antepenult (*ˈmadrase* 'school', Damascus). Another source of crossdialectal variation resides in words of the type CvCvC, which have initial stress (*ˈkatab* 'he wrote', Cairo) or final stress (*kiˈtab* 'he wrote', Bedouin dialect of the Cyrenaican Jebel).

Stress differences among Arabic dialects have been thoroughly studied in recent decades because of their relevance for metrical theory (Kenstowicz 1983; Halle and Vergnaud 1987; Hayes 1995). Crossdialectal differences are analyzed in terms of a small set of options in metrical theory, known as 'parameters'. Parametric differences involve the type of metrical foot (feet are rhythm units whose initial or final syllable is strong), the direction of metrification (starting at the word beginning or the end), and the stressability of the final syllable. Yet another reason for which phonologists have studied Arabic dialects resides in the complex interactions of stress assignment, syllabification, and processes of → syncope and → epenthesis, often resulting in opaque stress patterns (e.g. Brame 1974; Broselow 1982; Kenstowicz 1983; Al-Mozainy a.o. 1985; Kiparsky 2002).

This entry examines the stress patterns of the major modern dialects in some detail, starting with Egyptian Arabic, followed by the Eastern and Western dialects.

2. EGYPTIAN ARABIC

In Egyptian Arabic, specifically the colloquial spoken around Cairo, stress placement is governed by the following set of generalizations (Mitchell 1956:110–111, 1975:81; Harrell 1957):

(1) a. A final superheavy syllable (CvvC, CvCC) or heavy (Cvv) syllable is stressed.

maf.ˈhuum	'understood'
ḍa.ˈrabt	'I/you hit'
mas.ˈkaa	'holding [fem. sg.]'

b. Otherwise, a prefinal heavy syllable (CvC, CvC) is stressed.

fih.'muu.ha	‘they understood her’
mu.'ʿal.lim	‘teacher’

c. Stress falls on the antepenult in words of the types LLL and HLLL.

'ḍa.ra.bit	‘she hit’
'ka.ta.bu	‘they wrote/he wrote it’
ʾin.'ka.sa.rit	‘it was broken’
mux.'ta.li.fa	‘different [fem. sg.]’

d. Stress falls on the penult in words of the types LL, LLLL, HLL, LHLL.

'ka.tab	‘he wrote’
ka.ta.'bi.tu	‘she wrote it [masc.]’
mad.'ra.sa	‘school’
mu.dar.'ri.sit	‘teacher

A light syllable is defined as an open syllable that contains a short vowel (Cv). A heavy syllable is one that contains a long vowel (Cvv) or a short vowel and a consonant (CvC). In this entry a standard mora-based representation is adopted to represent syllable weight (McCarthy and Prince 1991; Hayes 1995).

```
(2)  a.   σ       b.    σ        c.    σ
          |            / \            / \
          μ            μ  μ           μ  μ
          |            |  |           \ /
         t a          t a b           t a ː
```

Forms such as *'ḍarab* show that CvC syllables count as light in word-final position, while forms such as *ʿa'malti* show that CvC counts as heavy in nonfinal position. In Cairene, as in most other dialects, the word-final consonant does not contribute to syllable weight. That is, final CvC syllables count as light, while superheavy syllables, which contain a long vowel plus a consonant (CvvC), or a short vowel plus two consonants (CvCC), are heavy. Invisibility (‘extrametricality’) of the final consonant is represented by angled brackets <…>.

```
(3)  a.   σ        b.    σ         c.    σ
          |             / \             / \
          μ             μ  μ            μ  μ
          |             |  |            \ /
         t a <b>       t a b <t>       t a : <b>
```

Detailed analyses of syllable structure in Arabic dialects, in relation with syncope and epenthesis, are offered by Broselow (1980, 1982, 1992), Kenstowicz (1986), and Kiparsky (2002).

As McCarthy (1979b) points out, the Cairene pattern raises interest because of its stressing of words ending in two light syllables. The penult or antepenult is stressed, depending on the length of the preceding string of light syllables. In particular, the penultimate stress of words ending in a sequence heavy-light-light presents a paradox. Crosslinguistically, heavy syllables tend to attract stress (Prince 1983). Cairene, however, stresses a light penult rather than the heavy antepenult in words such as *mad'rasa*, while it stresses a light antepenult in words composed of three light syllables, such as *'katabu*.

The metrical analysis of the Cairene colloquial pattern by Hayes (1995) incorporates the idea that a heavy syllable is quantitatively and metrically equivalent to two light syllables (Allen 1973; Halle and Vergnaud 1987; McCarthy 1979a, b; Prince 1983).

```
(4)  a. Foot assignment: parse the word from
        left to right into moraic trochees.
        Moraic trochee:
        (* .)               (*)
        L L                 H
        two light     or    a single heavy
        syllables           syllable

     b. Word layer construction: group feet
        into a right-headed word constituent.
```

In the examples of metrical structures below, feet are represented by pairs of parentheses. Their head (strong element) is indicated by an asterisk ‘*’, and their nonhead (weak element) by a dot ‘.’. The metrical layer above the feet indicates the placement of the main stress.

```
(5)  a. ( .    *)          b. ( .    *)
        (*)  (*)              (*)  (*) .
        H    H                H    H L
        maf.'huu<m>           fih.'muu.ha

     c. (*  )        d. (*   )         e. (.      *  )
        (* .)           (* .) .           (* .) (* .)
        L L             L L  L            L L   L L
        'ḍa.ra<b>       'ḍa.ra.bi<t>      ḍa.ra.'bi.tu
```

```
f. (.    *  )       g. ( .    *  )
   (*) (*  .)          (*) (*  .) .
    H   L  L            H   L  L L
   mad.'ra.sa          ʾin.'ka.sa.ri<t>
```

Note that a foot cannot consist entirely of a single light syllable, hence the lack of final stress in (5b, d, f). This ban on 'degenerate' feet captures the fact that the minimal word in Arabic is a heavy syllable (McCarthy and Prince 1990). An alternative analysis (Halle and Vergnaud 1987; Halle 1991) allows degenerate feet, and marks the final mora as extrametrical before foot assignment.

There are two well-known classes of exceptions to these stress rules (Mitchell 1956:111, 1975:81–82). Both involve words consisting of three light syllables, whose penult is exceptionally stressed:

(6) a. Verbs in which a vowel-initial suffix is attached to 3rd person feminine singular perfect of weak verbs
ra.'mi.t+u 'she threw it' (cf. 'kata.b+u 'he wrote it')
ša.'fi.t+ak 'she saw you'

b. Plurals containing two identical high vowels (i-i or u-u) in their first two syllables
du.'bu.ʿa 'hyenas' (cf. 'ku.tu.b+u 'his books')
ġi.'ri.ba 'crows'

These exceptional patterns can be specified for undergoing a morphologically triggered reversal of the direction of foot assignment, which becomes right-to-left in these cases (Watson 2002:97–98).

Mitchell (1975) discovered important evidence for the colloquial Cairene pattern by studying the way in which teachers of Classical Arabic (*ʿulamāʾ*) place stress when reciting the *Qurʾān*. Patterns reported by Mitchell cover colloquial and classical pronunciations. The generalizations in (7) represent interpretations by generative analysts (Langendoen 1968; Halle and Vergnaud 1987; McCarthy 1979a, b).

(7) a. Stress a superheavy ultima.

L'S	ḍa.'rabt	'I hit'
LL'S	sa.ka.'kiin	'thieves'

b. Otherwise, stress a heavy penult.

L'HL	ha.'ḏaa.ni	'these [fem. du.]'
L'HL	ʿa.'mal.ti	'you [fem. sg.] did'

c. Otherwise, stress the penult or antepenult, whichever is separated by an even number of syllables from the rightmost nonfinal heavy syllable or, if there is no heavy syllable, from the left boundary of the word.

'LLL	'ka.ta.ba	'he wrote'
LL'LL	ka.ta.'bi.tu	'they wrote'
LL'LLL	ša.ja.'ra.tu.hu	'his tree'
LLLL'LL	ša.ja.ra.tu.' hu.maa	'their [du.] tree'
H'LL	mar.'ta.ba	'mattress'
H'LLL	mar.'ta.ba.tu	'his mattress'
HLL'LL	ʾad.wi.ya.'tu.hu	'his drugs'
HLL'LLL	ʾad.wi.ya. 'tu.hu.maa	'their [du.] drugs'
LH'LL	mu.ʿal.'li.muun	'teachers'
LH'LLL	mu.qaa.'ti.la.tun	'fighter'

These forms are predictable by the set of metrical rules in (4) with the addition of a rule that marks the last mora of a final syllable as extrametrical (8b, d):

```
(8) a. (.      *   )      b. (.     .        *   )
       (*  .) (*  .)  .      (*  .) (*  .) (*  .)
        L  L  L  L  L         L  L  L  L  L  L
       ša.ja.'ra.tu.hu       ša.ja.ra.tu.'hu.ma<a>

    c. ( .   .       *   )  d. (.    .      *   )
       (*) (*  .) (*  .)       (*) (*  .) (*  .)  .
        H  L  L  L  L           H  L  L  L  L  L
       ʾad.wi.ya.'tu.hu        ʾad.wi.ya.'tu.hu.ma<a>
```

This analysis predicts secondary stresses on every foot head which is not main-stressed. Evidence for secondary stress is rather unclear. Mitchell (1975) does not mention it, whereas Weldon (1980) reports a secondary stress pattern deviating from the analysis above.

3. Eastern Arabic

The Palestinian pattern minimally differs from the Cairene pattern in stressing the antepenult, not the penult, in words of the type HLL (e.g. *'martaba*). This creates uniform antepenultimate stress in words ending in two light syllables (Brame 1973:20, 1974:41; Johnson 1979:154; Abu-Salim 1980:1; Kenstowicz and Abdul-Karim 1980; Kenstowicz 1983:207; Younes 1995:160).

(9) a. Stress a superheavy ultima.

L'S	ja.'waab	‘answer’	(Younes 1995: 160)
HL'S	ʿaa.la.'meen	‘two worlds’	(Younes 1995: 163)

b. Otherwise, stress a heavy penult.

L'HL	ka.'tab.na	‘we wrote’	(Abu-Salim 1980:1)
H'HL	mos.'taš.fa	‘hospital’	(Johnson 1979:154)

c. Otherwise, stress the antepenult.

LLL	'ḍa.ra.bu	‘they hit’	(Kenstowicz 1983:207)
LL'LLL	ša.ja.'ra.tu.hu	‘his tree’	
'HLL	'baa.ra.ku	‘they blessed’	(Kenstowicz 1983:207)
'HLL	'ʿal.la.mat	‘she taught’	(Kenstowicz 1983:207)
H'LLL	ḥaa.'ra.ba.to	‘she fought him’	(Younes 1995:163)
H'LLL	ʿal.'la.ma.tu	‘she taught him’	(Kenstowicz 1983:207)
L'HLL	mo.'naa.fa.se	‘competition’	(Johnson 1979:154)

This ‘Eastern Arabic stress rule’ is identical to the Latin stress rule, except for clause (9a), which Latin lacks. The antepenult is reached by marking the final syllable extrametrical, and constructing a quantity-sensitive trochee at the right edge (Kenstowicz 1983).

(10) a. Mark the final syllable as extrametrical (except when it is superheavy).

b. Foot assignment: assign a single quantity-sensitive trochee at the end of the word.

QS trochee: (* .) or (* .) or (*)
L L H L H

c. Word layer construction: group feet into a right-headed word constituent.

This analysis produces the following example metrifications:

(11) a. (*)
. (*)
L H
ja.'waa<b>

b. (*)
. (*)
L H
ka.'tab.<na>

c. (*)
(* .)
H L L
ʿal.'la.ma.<tu>

d. (*)
(* .)
H L L
'baa.ra.<ku>

This analysis features the uneven trochee (HL) in (11d), which is controversial in metrical theory. In Hayes' (1995) framework, trochees are strictly bimoraic (LL) or (H). Hayes (1995:128) presents a reanalysis of Palestinian stress, based on left-to-right metrification by strictly bimoraic trochees. This analysis is almost identical to the analysis of Cairene (4), with the single difference that a rule is added which marks the final foot as extrametrical in absolute final position.

(12) a. Foot assignment: parse the word from left to right into moraic trochees.

b. Mark a foot as extrametrical at the right edge of the word.

c. Word layer construction: group feet into a right-headed word constituent.

(13) a. (*)
(*)<(* .)>
H L L
‘baa.ra.ku

b. (*)
(* .)<(* .)>
L L L L
‘ša.ja.ra.tun

c. (*)
(*) (* .) .
H L L L
ʿal.'la.ma.tu

d. (. *)
(* .) (* .) .
L L L L L
ša.ja.'ra.tu.hu

To support the psychological reality of the rightward counting pattern, Hayes cites forms from Classical Arabic as produced by Palestinian speakers (Kenstowicz 1981a): *'šajaratun* 'a tree' and *šaja'ratuhu* 'his tree'. These forms motivate an addition to the generalizations in (9):

(14) Stress the pre-antepenult in words composed of four light syllables.

a.

'LLLL	'ša.ja.ra.tun	'a tree'	(Classical; Kenstowicz 1981)

b.

'LLLL	'ḍa.ra.ba.tu	'she hit him'	(Kenstowicz 1981:207)	(>'ḍar.ba.tu)
'LLLL	'ba.ka.ri.to	'his cow'	(Kenstowicz and Abdul-Karim 1980)	(>'ba.kar.to)

Words consisting of four light syllables do not occur in Palestinian colloquial, due to processes of syncope and epenthesis, familiar from Levantine dialects. These processes render stress opaque in (14b). For example, /bakar-it o/ 'his cow' is syncopated into ['bakarto], with antepenultimate stress across a closed penult, an apparent violation of quantity-sensitivity (9b). The opaque stress pattern of such forms thus presents additional evidence for the analysis in (12), on the assumption that stress is assigned prior to syncope and epenthesis (Brame 1974; Kiparsky 2002).

Damascene colloquial (Cowell 1964:180; McCarthy 1979b:459, 1980:79; Halle and Kenstowicz 1991:485) is apparently indistinguishable from Palestinian.

(15) a. Stress a superheavy ultima.

H'S	da.'rast	'I/you [masc. sg.] studied'
H'S	zaa.'ruuk	'they visited you [masc. sg.]'

b. Otherwise, stress a heavy penult.

L'HL	ka.'tab.ti	'you [fem. sg.] wrote'
L'HL	ma.'daa.res	'schools'

c. Otherwise, stress the antepenult.

'LLL	'da.ra.su	'they studied'
'HLL	'mad.ra.se	'school'
H'LLL	mut.'ta.ḥi.de	'she united' (Literary Arabic)

There are some subtle differences between Damascene and Palestinian in the enclitic stress system which are partly reviewed below.

The Lebanese dialect (Haddad 1984:19–21) also displays the 'Eastern Arabic stress rule'. Haddad presents two sources of evidence for a three-syllable window. First, Lebanese speakers stress the antepenult of classical words which contain long sequences of light syllables, such as *ḍa'rabana* 'he hit us'. Second, speakers mispronounce English words such as *ne'cessary* and *par'ticiple* with antepenultimate stress. Here, the final syllable is extrametrical, while a trochee places stress on the antepenult. HLL words can be analyzed either by an uneven trochee (Kenstowicz 1983) or by an even trochee (Hayes 1995).

(16)

a. uneven trochee	b. even trochee
(*)	(*)
(* .)	(*) .
H L	H L
'mad.ra.<se>	'mad.ra.<se>

In (16b) no monosyllabic foot is built on the penultimate syllable due to the ban on degenerate feet (Hayes 1995).

In sum, the stress patterns of the Eastern dialects are to some extent ambiguous between a right-to-left analysis with syllable extrametricality (10), where the foot is either an uneven trochee (16a) or an even trochee (16b), and a left-to-right analysis with moraic trochees and foot extrametricality (12). The ambiguity can be resolved using words with long sequences of light syllables, which the dialects strongly avoid by various constraints against open syllables, and which are only available from the pronunciation of words from Classical Arabic, or stress in loanwords.

Another way to resolve the ambiguities between uneven and even trochees is based on studying patterns of enclitic stress, specifically the accentuation of 3rd person feminine singular perfect verbs when followed by vowel-initial pronominal object suffixes. The examples in (17) are from Damascene (McCarthy 1980:84–85), but analogous cases have been reported for the dialects of Beirut (Abdul-Karim 1979) and Bani-Hassan Bedouin (Irshied and Kenstowicz 1984).

(17)

(*)	(*)	(. *)
(* .)	(* .)	(* .) (*)
H L	H L	H L L
'ʿal.la.<met> >	'ʿal.la.met # o >	ʿal.la.'mə<to>
‘she taught’	‘she taught him’	

The pronominal object form has penultimate stress, whereas antepenultimate stress is expected if the form were nonenclitic (see *mut'taḥide* (15c)). McCarthy (1980) and Halle and Kenstowicz (1991) attribute penultimate stress in the pronominal object form to the foot over the first two syllables of the base ('ʿal.la)<met>, a disyllabic trochee of the shape heavy-light (HL). This foot is respected by the construction of a foot on the penultimate syllable in the pronominal object form. If the foot on ‘she taught’ were a bimoraic trochee instead (compare Palestinian, structure (13a)), the incorrect prediction would be made of antepenultimate stress since the penult would be free to form a bimoraic trochee with the antepenult, as in (ʿal)('la.mə)<to>.

4. Classical Arabic

The stress pattern of Classical Arabic is reconstructed, due to a lack of native speakers, while the orthoepic tradition provides no explicit guidelines for accentuation. Current accentuations in use for reciting *Qurʾān* verses are influenced by native stress patterns of modern dialects. Methods of reconstructing Classical Arabic accentuation therefore include diachronic comparison with other Semitic languages (Brockelmann 1982), and identification of constant patterns in crossdialectal analysis (Janssens 1972). According to some authors (Wright 1859; Brockelmann 1982), the stress fell on the rightmost heavy (Cvv or CvC) syllable, and otherwise on the initial syllable. McCarthy (1979) states this rule as follows:

(18) a. Stress a superheavy ultima (limited to pausal forms, before a major syntactic break).

L'S	ya.'quul	‘he says’
L'S	ḍa.'rabt	‘I/you hit’

b. Otherwise, stress the rightmost nonfinal heavy syllable.

LH'HL	ma.naa.'dii.lu	‘kerchiefs’
LHL'HL	mu.dar.ri.'suu.na	‘teachers’
L'HLL	yu.'šaa.ri.ku	‘he participates’
H'HLL	kas.'sar.tu.hu	‘I smashed it’
'HLLH	'mam.la.ka.tun	‘kingdom’
'HLLLL	'mas.ʾa.la.tu.ha	‘her problem’

c. Otherwise, stress the first syllable.

'LLH	'ka.ta.buu	‘they wrote’
'LLLH	'ka.ta.ba.taa	‘they wrote [fem. du.]’
'LLLLL	'ma.li.ka.tu.hu	‘his queen’
'LLLLLH	'qa.ṣa.ba.tu.hu.maa	‘their [du.] flute’

McCarthy notes that this pattern occurs in modern dialects such as Egyptian Saʿīdī (Khalafallah 1969) and Yemen Plateau (Diem 1973). This pattern can be analyzed as follows:

(19) a. Mark the final syllable (CvC or Cvv) as extrametrical.

b. Foot assignment: assign a single quantity-sensitive *unbounded* trochee at the end of the word.
QS unbounded trochee:
(* … .) or (* … .) or (*)
L … L H … L H

c. Word layer construction: group feet into a right-headed word constituent.

Under this analysis, Classical Arabic is nearly identical to the Palestinian/Damascene pattern

(10b), from which it differs only by lacking the requirement that feet are maximally two syllables long.

```
(20) a. (*       )        b.         (*)
        (*  .  .)            .  .  .  (*)
        H  L  L              L  H  L  H
        'mam.la.ka.<tun>     mu.dar.ri.'suu.<na>

     c. (*          )
        (*  .  .  .)
        L  L  L  L
        'ma.li.ka.tu.<hu>
```

A different pattern for the classical language was deduced from the modern dialects by Abboud a.o. (1968), Abdo (1969), Brame (1970), Angoujard (1990), Versteegh (1997), and Gordon (2000). Here stress is restricted to fall on one of the last three syllables of the word.

(21) a. Stress a superheavy ultima (see examples above).
b. Otherwise, stress a heavy penult (see examples above).
c. Otherwise, stress the antepenult (only forms deviating from above).

L'LLH	ka.'ta.ba.taa	'they wrote [fem. du.]'
H'LLH	mam.'la.ka.tun	'kingdom'
LL'LLL	ma.li.'ka.tu.hu	'his queen'
HL'LLL	mas.ʾa.'la.tu.ha	'her problem'
LLL'LLH	qa.ṣa.ba.'tu.hu.maa	'their [du.] flute'

This pattern is identical to Palestinian, Damascene, and Lebanese. See (10) for a metrical analysis.

5. Western Arabic

Finally, we briefly turn to some of the Western dialects. The Bedouin dialect spoken in the Negev (Blanc 1970:120–122; Kenstowicz 1983; Hayes 1995) displays an iambic stress pattern. The generalizations below were taken from Hayes (1995:226):

(22) a. Stress a superheavy ultima.

LLL'S	ġa.na.'maat	'[several individual] sheep'

b. Otherwise, stress a heavy penult.

L'HL	ġa.'nam.na	'our sheep'

c. Otherwise, stress a heavy antepenult.

'HLH	'al.ġa.nam	'the sheep'

d. Otherwise, stress the ultima of a two-syllable word beginning with a light syllable.

L'L	ki.'tab	'he wrote'
L'H	ġa.'nam	'sheep'

e. Otherwise, stress the penult or antepenult, whichever is separated by an odd number of light syllables from the nearest preceding heavy syllable, or in the absence of such a syllable, from the beginning of the word.

L'LH	za.'la.mah	'man'
L'LLH	za.'la.ma.tak	'your man'
HL'LH	al.ġa.'na.mah	'the sheep [sg.]'
HL'LH	ḥaa.ṛa.'ba.tih	'she fought him'

The analysis (Hayes 1995:227) is the exact iambic counterpart of the Palestinian pattern:

(23) a. Foot assignment: parse the word from left to right into iambs.

```
Iamb: (.  *)  or  (.  *)  or  (*)
       L  L        L  H        H
```

b. Mark a foot as extrametrical at the right edge of the word.

c. Word layer construction: group feet into a right-headed word constituent.

```
(24) a. (      *)         b. (*)
        (.     *)            (*) (.   *)
        L      H              H   L   H
        ġa.'nam              'al.ġa.nam

     c. (      *)         d. (.          *)
        (.  *) (.   *)       (*) (.  *) (*)
        L   L   L   H         H   L   L   H
        za.'la.ma.tak        al.ġa.'na.mah
```

The Bedouin dialect of Cyrenaican Jebel, spoken in eastern Libya (Owens 1984:32–35; Mitchell 1975:83–92), has rather complex interactions between stress assignment and processes of syncope and epenthesis, affecting syllabification (→ resyllabification). In the examples below, effects of syncope and epenthesis have been ignored.

(25) a. Stress a superheavy ultima.

L'S	ḥa.'šiiš	'grass'	
LH'S	fi.naa.'jiil	'cups'	

b. Otherwise, stress the ultima of a two-syllable word beginning with a light syllable.

L'H	si.'maa	'shy'	
L'H	ka.'tab	'he wrote'	(> ki'tab)

c. Otherwise, stress a heavy penult.

L'HH	ka.'tab.tan	'you [fem. pl.] wrote'	(> ki'tabtan)
HL'HL	maˁ.ra.'kɨt.ha	'her quarrel'	(> maˁrɨ'kɨtta)

d. Otherwise, stress a heavy antepenult.

'HLH	'mak.ta.bih	'his office'	(> 'maktibih)
L'HLH	ta.'raa.fa.gan	'they [fem.] accompanied'	(> ti'raafagan)

e. Otherwise, stress the penult.

L'LH	ka.'ta.bat	'she wrote'	(> ik'tibat)
HL'LH	in.ga.'ta.lat	'she was killed'	(> inig'tilat)

Hayes (1995:228–239) analyzes the stress pattern by an iambic (weak-strong) foot, as in the Negev Bedouin dialect:

```
(26) a. (    *)          b. (   *)
        (.   *)             (.  *) (.  *)
        L    H              L   H  L   H
        ka.'tab             ta.'raa.fa.gan

     c. (*)              d. (.        *)
        (*) (.  *)          (*) (.    *)  .
        H   L   H           H   L     H   L
        'mak.ti.bih         maˁ.ra.'kɨt.ha
```

The most interesting metrical property of the Cyrenaican Bedouin dialect is its interaction between stress assignment and → syncope. Hayes (1995) argues that syncope of a stressed syllable (the head of an iambic foot) causes a retraction of stress to the unstressed syllable of the foot, shown in (27):

(27)

	After footing		After syncope
a.	(ki.'ti).(bih)	>	(ki.t).(bih)
b.	(faa).(ki.'hi).(tih)	>	(faa).(ki.h).(tih)

After stress shift

>	('kit).(bih)	'his books'
>	(faa).('kih).(tih)	'his fruit'

Similar stress shifts under deletion of the stressed syllable have been reported for Hijazi Bedouin (Al-Mozainy a.o. 1985) and Bani-Hassan Bedouin (Kenstowicz 1983; Irshied and Kenstowicz 1984).

Finally, we turn to the stress pattern of the Maghreb dialects, specifically Moroccan (Keegan 1986; Harrell 1962; Boudlal 2001). The following examples, all isolation forms, are taken from Boudlal (2001:122), who does not provides glosses.

(28) a. Stress a heavy ultima.

H'H	law.'yin
L'H	li.'mun
LL'H	mər.məd.'nak
HLL'H	ban.ya.ha.'lih

b. Otherwise, stress the penult.

'HL	'bab.ha
'LL	'məl.məl
L'LL	li.'mu.na
LLL'LL	di.ri.ha.'li.ha

The weight distinction is between heavy syllables (CvC) and light syllables (Cv and CəC). Heavy syllables in final position are heads of iambic feet. The penult is reached by marking final syllables extrametrical if they are light.

(29) a. Mark a final light syllable (Cv and CəC) as extrametrical.

b. Foot assignment: assign a single iamb at the end of the word.

c. Word layer construction: group feet into a right-headed word constituent.

```
(30) a.      (   *)        b. (*)
             (.  *)           (*)
          H  L  L H           H
          ban.ya.ha.'lih      'bab.<ha>

     c.      (   *)
             (.  *)
          L  L  L  L
          di.ri.ha.'li.<ha>
```

Moroccan Arabic thus shares the iambic foot with Negev and Cyrenaican Bedouin dialects, but resembles Damascene and Lebanese Arabic in its directionality: a single foot is constructed at the right edge, vis-à-vis extrametricality.

6. Conclusion

In sum, word stress patterns of Arabic dialects differ along a number of dimensions, which can be captured by metrical frameworks. The major parametric differences between the dialects involve (i) foot type (trochee or iamb), (ii) direction of metrification (left-to-right or right-to-left), and (iii) different types of extrametricality (consonant, mora, syllable, or foot).

Bibliographical references

Abboud, Peter F., Najm Bezirgan, Wallace Erwin, Munah Khouri, Ernest McCarus, and Raji Rammuny. 1968. *Introduction to Modern Standard Arabic pronunciation and writing*. Ann Arbor, Mich.: Interuniversity Committee on Near Eastern Languages.

Abdo, Daud A. 1969. *On stress and Arabic phonology*. Beirut: Khayats.

Abdul-Karim, Kamal Wadih. 1979. *Aspects of the phonology of Lebanese Arabic*. Ph.D. diss., University of Illinois at Urbana.

Abu-Salim, Issam M. 1980. "Epenthesis and geminate consonants in Palestinian Arabic". *Studies in the Linguistic Sciences* 10.1–11.

——. 1987. "Vowel harmony in Palestinian Arabic: A metrical perspective". *Journal of Linguistics* 23.1–24.

Al-Ani, Salman H. 1992a. "Lexical stress variation in Arabic: An acoustic spectrographic analysis". *The Arabist: Budapest Studies in Arabic* 3–4.9–27.

——. 1992b. "Stress variation of the construct phrase in Arabic: A spectrographic analysis". *Anthropological Linguistics* 34.256–276.

Allen, W. Sidney. 1973. *Accent and rhythm*. Cambridge: Cambridge University Press.

Al-Mozainy, Hamza Qublan. 1981. *Vowel alternations in a Bedouin Hijazi Arabic dialect: Abstractness and stress*. Ph.D. diss., University of Texas at Austin.

——, Robert Bley-Vroman, and John J. McCarthy. 1985. "Stress shift and metrical structure". *Linguistic Inquiry* 16.135–144.

Angoujard, Jean-Pierre. 1990. *Metrical structure of Arabic*. Dordrecht: Foris.

Birkeland, Harris. 1954. *Stress patterns in Arabic*. Oslo: Dybwad.

Blanc, Haim. 1970. "The Arabic dialect of the Negev Bedouins". *Proceedings of the Israel Academy of Science and Humanities* 4.112–150.

Boudlal, Abdelaziz. 2001. *Constraint interaction in the phonology and morphology of Casablanca Moroccan Arabic*. Ph.D. diss., Mohammed V University, Rabat.

Brame, Michael K. 1970. *Arabic phonology: Implications for phonological theory and historical Semitic*. Ph.D. diss., Massachusetts Institute of Technology.

——. 1973. "On stress assignment in two Arabic dialects". *A Festschrift for Morris Halle*, ed. Stephen R. Anderson and Paul Kiparsky, 14–25. New York: Holt, Rinehart and Winston.

——. 1974. "The cycle in phonology: Stress in Palestinian, Maltese and Spanish". *Linguistic Inquiry* 5.39–60.

Brockelmann, Carl. 1982. *Grundriss der vergleichenden Grammatik der semitischen Sprachen*. Repr. Hildesheim: G. Olms. (First publ., Berlin: Reuther und Reichard, 1908.)

Broselow, Ellen. 1980. "Syllable structure in two Arabic dialects". *Studies in the Linguistic Sciences* 10.13–24.

——. 1982. "On predicting the interaction of stress and epenthesis". *Glossa* 16.115–132.

——. 1992. "Parametric variation in Arabic dialect phonology". *Perspectives on Arabic linguistics*, IV, ed. Mushira Eid and John McCarthy, 1–42. Amsterdam and Philadelphia: J. Benjamins.

Cowell, Mark W. 1964. *A reference grammar of Syrian Arabic*. Washington, D.C.: Georgetown University Press.

Diem, Werner. 1973. *Skizzen jemenitischer Dialekte*. Wiesbaden and Beirut: F. Steiner.

Fischer, Wolfdietrich and Otto Jastrow. 1980. *Handbuch arabischer Dialekte*. Wiesbaden: O. Harrassowitz.

Gordon, Matthew. 2000. "Re-examining default-to-opposite stress". *Proceedings of the Berkeley Linguistics Society* 26.101–112.

Guindy, A.-K. 1988. "On the stress in the 'madrasa' word structure in Cairene colloquial Arabic". *Zeitschrift für Arabische Linguistik* 18.33–58.

Haddad, Ghassan F. 1984. *Problems and issues in the phonology of Lebanese Arabic*. Ph.D. diss., University of Illinois at Urbana-Champaign.

Halle, Morris. 1991. "Respecting metrical structure." *Natural Language and Linguistic Theory* 8.149–176.

—— and Michael Kenstowicz. 1991. "The free element condition and cyclic versus noncyclic stress". *Linguistic Inquiry* 22.457–501.

—— and Jean-Roger Vergnaud. 1987. *An essay on stress*. Cambridge, Mass.: MIT Press.

Harms, Robert T. 1981. "A backwards approach to Cairo Arabic stress". *Linguistic Analysis* 7.429–50.

Harrell, Richard. 1957. *The phonology of Colloquial Egyptian Arabic*. New York: American Council of Learned Societies.

——. 1960. “A linguistic analysis of Egyptian Radio Arabic. *Contributions to Arabic linguistics*, ed. Charles A. Ferguson, 3–77. Cambridge, Mass.: Harvard University.

——. 1962. *A short reference grammar of Moroccan Arabic*. Washington, D.C.: Georgetown University Press.

Hayes, Bruce. 1981. *A metrical theory of stress rules*. Ph.D. diss., Massachusetts Institute of Technology.

——. 1995. *Metrical stress theory: Principles and case studies*. Chicago: University of Chicago Press.

Irshied, Omar and Michael Kenstowicz. 1984. “Some phonological rules of Bani-Hassan Arabic: A Bedouin dialect”. *Studies in the Linguistic Sciences* 14.109–147.

Janssens, Gerard. 1972. *Stress in Arabic and word structure in the modern Arabic dialects*. Leuven: Peeters.

Johnson, C. Douglas. 1979. “Opaque stress in Palestinian”. *Lingua* 49.153–168.

Kager, René. 1995. “Surface opacity of metrical structure in Optimality Theory”. *The derivational residue in phonological Optimality Theory*, ed. Ben Hermans and Marc van Oostendorp, 207–245. Amsterdam and Philadelphia: J. Benjamins.

——. 1999. *Optimality Theory*. Cambridge: Cambridge University Press.

Keegan, John M. 1986. *The phonology and morphology of Moroccan Arabic*. Ph.D. diss., New York University.

Ken‘aan, S.A.R. 1989. “Word stress in English and Arabic”. *Indian Journal of Applied Linguistics* 15.1–22.

Kenstowicz, Michael. 1980. “Notes on Cairene Arabic syncope”. *Studies in the Linguistic Sciences* 10.39–53.

——. 1981a. “A note on cyclic stress in Levantine Arabic”. *Theoretical issues in the grammar of Semitic languages*, ed. Hagit Borer and Joseph Aoun, 21–31. (= *MIT Working Papers in Linguistics*, 3.) Cambridge, Mass.: Department of Linguistics and Philosophy, Massachusetts Institute of Technology.

——. 1981b. “The metrical structure of Arabic accent”. Paper delivered at the UCLA-USC Conference on Nonlinear Phonology, Lake Arrowhead, California.

——. 1983. “Parametric variation and accent in the Arabic dialects”. *Proceedings of the Chicago Linguistic Society* 19.205–213.

——. 1986. “Notes on syllable structure in three Arabic dialects”. *Revue Québécoise de Linguistique* 16.101–128.

——. 1994. *Phonology in generative grammar*. Oxford: Blackwell.

—— and Kamal Abdul-Karim. 1980. “Cyclic stress in Levantine Arabic”. *Studies in the Linguistic Sciences* 10.55–76.

Khalafallah, Abdelghany A. 1969. *A descriptive grammar of Saʕi:di Egyptian Colloquial Arabic*. The Hague: Mouton.

Kiparsky, Paul. 2000. “Opacity and cyclicity”. *Linguistic Review* 17.351–367.

——. 2002. “Syllables and moras in Arabic”. *The syllable in Optimality Theory*, ed. Caroline Féry and Ruben van de Vijver, 147–182. Cambridge: Cambridge University Press.

Langendoen, D. Terence. 1968. *The London School of Linguistics: A study of the linguistic theories of B. Malinowski and J.R. Firth*. Cambridge, Mass.: MIT Press.

McCarthy, John J. 1979a. *Formal problems in Semitic phonology and morphology*. Ph.D. diss., Massachusetts Institute of Technology.

——. 1979b. “On stress and syllabification”. *Linguistic Inquiry* 10.443–65.

——. 1980. “A note on the accentuation of Damascene Arabic”. *Studies in the Linguistic Sciences* 10.77–98.

—— and Alan S. Prince. 1990. “Foot and word in prosodic morphology: The Arabic broken plural”. *Natural Language and Linguistic Theory* 8.209–284.

Mitchell, Terence F. 1956. *An introduction to Egyptian Colloquial Arabic*. London: Oxford University Press.

——. 1960. “Prominence and syllabification in Arabic”. *Bulletin of the School of Oriental and African Studies* 23.369–389.

——. 1975. *Principles of Firthian linguistics*. London: Longmans.

——. 1993. *Pronouncing Arabic*, II. Oxford: Clarendon Press.

Owens, Jonathan. 1980. “The syllable as prosody: A reanalysis of syllabification in Eastern Libyan Arabic”. *Bulletin of the School of Oriental and African Studies* 43.277–287.

——. 1984. *A short reference grammar of Eastern Libyan Arabic*. Wiesbaden: O. Harrassowitz.

Prince, Alan S. 1983. “Relating to the grid”. *Linguistic Inquiry* 14.19–100.

Selkirk, Elisabeth O. 1981. “Epenthesis and degenerate syllables in Cairene Arabic”. *Theoretical issues in the grammar of Semitic languages*, ed. Hagit Borer and Joseph Aoun, 209–232. (= *MIT Working Papers in Linguistics*, 3.) Cambridge, Mass.: Department of Linguistics and Philosophy, Massachusetts Institute of Technology.

Stoetzer, Willem. 1981. “Stress in Koranic Arabic: Its relation to *ʾiʕrāb*”. *Proceedings of the Ninth Congress of the Union Européenne des Arabisants et Islamisants, Amsterdam, 1st to 7th September 1978*, ed. Rudolph Peters, 261–269. Leiden: E.J. Brill.

Versteegh, Kees. 1997. *The Arabic language*. New York: Columbia University Press.

Watson, Janet C.E. 2002. *The phonology and morphology of Arabic*. Oxford: Oxford University Press.

Welden, Ann. 1980. “Stress in Cairo Arabic”. *Studies in the Linguistic Sciences* 10.99–120.

Wright, William. 1859. *A grammar of the Arabic language*. Cambridge: Cambridge University Press. (Repr., 1967.)

Younes, Munther A. 1995. “On vowel shortening in Palestinian Arabic”. *Perspectives on Arabic linguistics*, VII, ed. Mushira Eid, 157–171. Amsterdam and Philadelphia: J. Benjamins.

René Kager (University of Utrecht)

Subject

The subject of a sentence is a constituent that performs a certain grammatical function in relation to the → predicate or verb of the clause. It signifies the topic around which the rest of the clause revolves, or represents the → agent (doer) of the verb. Typically, the subject is a → noun phrase. It can also be an adjective, a pronoun, or a clause. In Classical and Modern Standard Arabic, the subject is explicitly marked for Case; it mainly receives nominative case but, under certain conditions, can also be marked for accusative. In the spoken varieties, the subject is not morphologically marked for Case.

1. Position in the sentence

There are two types of sentences in Arabic, → verbal and → nominal, exhibiting two different word orders. The position of the subject in Arabic depends on the sentence type in which it appears. It may occupy either a sentence-initial position or appear following the verb or predicate. Arabic grammarians, as a result, give different names to the subject, depending on its position in the sentence. When it appears following the verb, the subject is called → *fāʿil* lit. 'doer [of the verb]' or 'agent', as the verb cannot stand alone without it. In subject/predicate sentences, where no verb may be used, the subject is called *mubtadaʾ* or *musnad* 'inchoative' (→ *ʾisnād*; see next section for details).

In verbal sentences, the subject (or agent) occurs following the verb, as in (1). When the sentence is nominal, the subject occurs before either the verb, as in (2), or the predicate, as in (3).

(1)	*ʾakala*	*l-walad-u*
	ate.3ms	the-boy-Nom

'The boy ate'

(2)	*al-walad-u*	*ʾakala*
	the-boy-Nom	ate.3ms

(3)	*al-ġināʾ-u*	*jamīl*
	the-singing-Nom	lovely

'The singing is lovely'

In nominal, subject/predicate sentences (also called 'equational', e.g. in Badawi a.o. 2004), the subject normally occurs first, as in (3). However, when the subject is indefinite and the predicate of the sentence is a prepositional phrase, the subject is delayed or postposed, i.e. occurs after the predicate, while retaining its nominative case, as in (4).

(4)	*fī*	*l-maktabati*	*ṭifl-un*
	in	the-library	child.ms-Nom

'A child is in the library'/
'There is a child in the library'

The subject of a → passive sentence can occur either pre- or postverbally, e.g. *fuqida l-kitābu/al-kitābu fuqida* 'the book was lost'.

In spoken (Egyptian) Arabic, the subject precedes the verb. When the verb precedes the subject, it receives contrastive stress and is followed by a pause, indicating a topicalized structure (see Ouhalla 1997; → topicalization).

2. Terms used to describe the subject

Traditional Arabic grammarians use different names to refer to the subject, depending on whether it appears in a verbal or nominal sentence. Thus, in Classical Arabic grammar, the subject of a verbal sentence is called *fāʿil* 'agent', a word derived from the word *fiʿl* 'verb', indicating the close relationship between both (Sībawayhi, *Kitāb* I, 13). The subject of the nominal, subject-predicate sentence is called *mubtadaʾ* 'that with which is begun', i.e. the inchoative (→ *ibtidāʾ*). The rest of the sentence is called → *xabar* 'predicate'. Sībawayhi (*Kitāb* I, 7) uses both *mubtadaʾ* and *musnad* 'that which leans' to refer to the subject of a nominal sentence; he calls the predicate *musnad ʾilayhi* 'that upon which the subject leans' (→ *ʾisnād*). The subject of nominal, subject/verb sentences (SV) is sometimes called a 'topic', and the sentence structure, a → topic/comment sentence (Badawi a.o. 2004:326). When the predicate is fronted so that it precedes the subject, the subject is called *mubtadaʾ muʾaxxar* 'delayed [postposed] subject'.

In Arabic grammar, the subject of a passive sentence, which is semantically the object (patient) of the verb, is called *nāʾib al-fāʿil* 'the substitute of the subject/agent', a name that describes the new role of the object in the passive sentence.

3. CONCORD/AGREEMENT WITH THE VERB

Arabic has two patterns of subject-verb → agreement, depending on the → word order used. When the word order is verb-subject (VS), agreement is partial. The verb agrees with only two features of the subject: person and gender. Number agreement is not marked in this order – default agreement with all subjects being singular, as in (5) and (6).

(5) *jāʾa* *l-ʾawlād-u*
came.3ms the-boys.mp-Nom
'The boys came'

(6) *jāʾat* *al-banāt-u*
came.3fs the-girls.fp-Nom
'The girls came'

Agreement in SV sentences, by contrast, must be complete in person, gender, and number, as in (7)–(10).

(7) *al-walad-āni* *jāʾā*
the-boys-mdu.Nom came.3mdu
'The two boys came'

(8) *al-bin-tāni* *jāʾatā*
the-girls-fdu.Nom came.3fdu
'The two girls came'

(9) *al-ʾawlād-u* *jāʾū*
the-boys.mp-Nom came.3mp
'The boys came'

(10) *al-banāt-u* *jiʾna*
the-girls.fp-Nom came.3fp

Changing the agreement patterns of either order results in ungrammatical sentences when the subject is masculine plural or dual. as in (11) and (12); feminine subjects tolerate singular agreement in the SV order, as in (13).

(11) **al-ʾawlād-u* *jāʾa*
the-boys.mp-Nom came.3ms

(12) **jāʾā* *l-walad-āni*
came.3mdu the-boys-mdu.Nom

(13) *al-banāt-u* *jāʾat*
the-girls.fp-Nom came.3fs

A fully agreeing verb in what seems to be a VS order is possible only when the verb receives contrastive stress and is followed by a pause – indicating a topicalized structure, as in (14).

(14) *jāʾū* *l-ʾawlād-u*
came.3mp the-boys.mp-Nom

When the subject is a feminine singular substantive noun, gender agreement is required in VS only when the verb immediately precedes the subject, as in (15). However, if the feminine subject is separated from its verb by the object, the verb may carry masculine singular agreement, as in (16).

(15) *ḥaḍarat* *imraʾat-un*
came.3fs woman.fs-Nom
'A woman came'

(16) *ḥaḍara* *l-qāḍiy-a*
came.3ms the-judge.ms-Acc
mraʾat-un
woman.fs-Nom
'A woman came before the judge'
(Wright 1967:II, 289)

The above rule applies to nouns that are explicitly marked for feminine, i.e. when → gender signifies the sex of a noun (female) or when gender is grammatical (feminine by form, e.g. *aš-šajara* 'the tree'). Some nouns are feminine by convention or 'usage' (Wright 1967:I, 179–180), i.e. not explicitly marked for feminine gender; yet, when one of these nouns is the subject of a sentence, feminine agreement is preferred.

(17) *ġarabat* *aš-šams-u*
set.3fs the-sun.fs-Nom
'The sun set'

Agreement with irregular (broken) plural subjects is feminine, particularly with nonhuman nouns, as in (18).

(18) *ṯumma* *qasat*
then hardened.3fs
qulūb-u-kum
hearts-Nom-your
'Then, your hearts became cruel'
(Q. 2/74; Wright 1967:II, 290)

Verbal agreement with a collective noun subject can be realized as either feminine or masculine; feminine agreement is preferred, though, as in (19).

(19) *mātat* *al-ġanam-u*
died.3fs the-sheep-Nom
'The sheep died'

In spoken varieties of Arabic, the subject/verb agreement system is reduced. The verb agrees with a singular subject in all features of person, gender, and number. Dual agreement is no longer expressed on the verb; instead, plural agreement is used for both dual and feminine plural subjects, as in (20).

(20) *il-bint-ēn* *bi-yzakru*
the-girls-du Progr-study.3p
'The two girls are studying'

Interestingly, when the subject is plural, two patterns of agreement can be expressed on the verb. The first is plural agreement, and the second is feminine singular agreement (called 'deflected' agreement). The last pattern is remarkable as it is the emerging pattern of agreement with masculine plural subjects, particularly among the younger generation, as in (21) and (22).

(21) *il-banāt*
the-girls
bi-yzakru/bi-tzākir
Progr-study.3p/Progr-study.3fs
'The girls are studying'

(22) *iṭ-ṭalaba*
the-students.mp
bi-yzakru/bi-tzākir
Progr-study.3p/Progr-study.3fs
'The students are studying'

For more details on the realization and rules of use of deflected agreement in spoken varieties of Arabic, see Belnap (1993).

4. PROPERTIES OF THE SUBJECT: TYPES AND DEFINITENESS

Any noun – a proper name, a common noun, a pronoun (23), a demonstrative pronoun (24), or a whole clause (25) – can function as the subject of a sentence.

(23) *huwa* *mujtahid*
he hardworking.ms
'He is hardworking'

(24) *hāḏā* *raʾiʿ-un*
this great.ms-Nom
'This is great/wonderful'

(25) *ʾan* *taṣūmū* *xayr-un* *la-kum*
Compl. fast.2mp better-Nom to-you
'For you to fast is for your good'
(Q. 2/184)

The subject of a subject-predicate sentence is overwhelmingly definite, and the predicate agrees with it in gender and number. The subject of the verbal sentence can be either definite or indefinite. In Classical and Modern Standard Arabic, an indefinite noun can occur as the subject of a nominal sentence. In such cases, the word *hunāka* 'there' is usually used to introduce the sentence in Modern Standard Arabic (Badawi a.o. 2004:316–317), as in (26).

(26) *hunāka* *rajul-un*
there man.ms-Nom
yantaḏ̣iru-ka
Imperf.wait.3ms-cl2ms
'There is a man waiting for you'

The equivalent of *hunāka* in spoken Arabic is the existential particle *fī* (→ locatives).

5. SUBJECTLESS SENTENCES (PRO-DROP)

Arabic allows the optional omission of the subject pronoun of a sentence, a phenomenon sometimes referred to as → 'pro-drop' and found in languages with rich subject/verb agreement systems. The use of this grammatical phenomenon depends on semantic and pragmatic factors, hinging on the familiarity of the reader/interlocutor with the topic and context of the text/conversation. It follows that, given the two possible positions the subject can occupy in Arabic and the matching agreement patterns realized on the verb, pro-drop is possible only in the SV order, where agreement features on the verb represent the complete set of the missing subject pronoun: person, gender, and number. In Arabic grammar, the pronominal features expressed on the verb in subjectless sentences are treated as 'an implied pronoun' (Wright 1967:II, 250) or 'a pronominalized agent' (Badawi a.o. 2004:352). An example is given in (27) and (28).

(27) *hum intahaw min al-mašrūʿ*
they finished.3mp from the-project
'They have finished the project'

(28) *intahaw min al-mašrūʿ*
finished.3mp from the-project
'They have finished the project'

In general, pro-drop applies freely, subject to recoverability of deletion, i.e. the interlocutor's ability to reach the appropriate interpretation of the missing pronoun. Pro-drop, therefore, cannot apply in subject-predicate sentences, since as a result of the missing verb in these sentences, the remnant (i.e. the predicate) would be a truncated constituent failing to be construed as a complete proposition.

When the subject is singular, as in (29)–(32), the subjectless sentence may seem ambiguous between VS and SV word orders, since singular subjects do not require an accompanying difference in the agreement patterns in each order.

(29) *ʾakala ṭ-ṭifl-u*
ate.3ms the-child.ms-Nom

(30) *aṭ-ṭifl-u ʾakala*
the-child.ms-Nom ate.3ms

(31) *ʾakalat aṭ-ṭiflat-u*
ate.3fs the-child.fs-Nom

(32) *aṭ-ṭiflat-u ʾakalat*
the-child.fs-Nom ate.3fs

In each of the above sentences, the verb can stand alone as a sentence: *ʾakala* 'he ate' and *ʾakalat* 'she ate'. Yet, it is generally accepted that since subjects with features other than singular require richer agreement on the verb in SV, this is the word order that allows the subject to be absent.

In the spoken varieties of Arabic, pro-drop is widely used in conversation. Reduced (or deflected) agreement with masculine and feminine plural subjects, however, does not allow pro-drop, due to impossibility of appropriate interpretation. Thus, a sentence consisting only of the verb *bi-tzākir* [Progr-study.3fs] in examples (21) and (22) can never be interpreted to mean 'they [fem./masc. pl.] are studying'. Standing alone without a subject, it can only mean 'she is studying'.

One of the uses of pro-drop in the spoken variety, with plural agreement on the verb, is to indicate an impersonal subject, as in (33). Pro-drop in this case is used to express a proposition in which the agent is unimportant or one that the speaker chooses to suppress.

(33) *iktašafu ġāz ṭabīʿi fi l-baḥr*
discovered.3p gas natural in the-sea
'They discovered natural gas in the sea'

This use replaces that of the impersonal → passive in the Classical/Standard variety.

6. Case marking of the subject

In both Classical Arabic and Modern Standard Arabic, the subject normally receives nominative case, whether the word order is VS or SV. In Arabic grammar, the realization of nominative case varies, depending on the nature of the noun receiving Case, i.e. depending on its number (singular, dual, or plural and type of plural) and gender (masculine or feminine). Generally, nominative case is termed *rafʿ* lit. 'raising', and realized as the back, high, rounded vowel /u/ (*ḍamma*), which is suffixed to the noun. The following are the main patterns of nominative case marking: (i) Masculine and feminine singular nouns, broken plurals, and regular (or sound) feminine plurals are marked for nominative by the addition of the suffix /u/ as illustrated by the examples of Case realization on these noun types, respectively: *al-kitāb-**u*** 'the book', *al-bint-**u*** 'the girl', *al-kutub-**u*** 'the books', and *al-banāt-**u*** 'the girls'; (ii) the suffixal morpheme /-ūna/ is used to mark nominative case on sound (regular) masculine plurals, e.g. *al-muhandis**ūna*** 'the engineers'; and (iii) dual forms are marked for nominative by the suffix /āni/, e.g. *kitāb**āni*** 'two books'.

Nominative case marking on indefinite nouns, other than sound plurals, requires attaching the consonant /n/ to the /u/ suffix. This is called → *tanwīn* (→ nunation), whose nature and the motivation for its use have been argued to be phonological rather than morphological (Gaballa 1986), thus *ar-rajul**u**/rajul**un*** (masc. sg.);

al-banātu/banātun (fem. pl.); *ar-rijālu/rijālun* (masc. pl.). For more details on plural formation and syllable structure in broken and sound plurals in Arabic, see Idrissi (1996).

Nominative case is not the only Case the subject can receive in Arabic. In subject-predicate sentences, the subject of a main (or kernel) sentence can be preceded by one of a set of particles that affect the case the subject receives. The most frequently used of these particles is *ʾinna*, which is why the particles have been named after it; → *ʾinna wa-ʾaxawātuhā*. There is no consensus in the literature on how to translate *ʾinna*. Some, for example Cowan (1958), argue that there is no English translation equivalent for this particle. However, Cowan points out that in an *ʾinna* sentence, when the emphatic particle *la-* is prefixed to the predicate, *ʾinna* may be translated as 'indeed'. Badawi a.o. (2004:320) also translate it as 'indeed' and call it a sentence modifier with an emphatic or focus function, as in (34).

(34) *ʾinna* *l-ʾinsān-a*
ʾinna the-human.being-Acc
la-fī *xusr*
Emph-in loss
'Man is indeed in loss' (Q. 103/2; Cowan 1958:65)

Wright (1967:II, 78) treats *ʾinna* as an adverb meaning 'truly' or 'certainly'. Subjects occurring after *ʾinna* receive accusative case, as in (35).

(35) *ʾinna* *l-mudarrisat-**a***
ʾinna the-teacher.fs-Acc
muxliṣat-un
devoted.fs-Nom
'The teacher is devoted'

The subject is marked for accusative case after the *ʾinna*-type particles, even when indefinite, i.e. when separated from the particle by a PP predicate.

(36) *ʾinna* *fī* *l-maktabat-i* *walad-**an***
ʾinna in the-library-Gen boy.ms-Acc
'There is a boy in the library'

Note that in terms of their meaning, sentences with *ʾinna* followed by a delayed subject, as in (36), are equivalent to English existential sentences, as the translation indicates (see Haywood and Nahmad 1965:396).

Definite subjects do not usually allow material to intervene between them and the particle *ʾinna*, which is why (37) is ungrammatical.

(37) **ʾinna fī* *l-maktabat-i*
ʾinna in the-library-Gen
*ṭ-ṭifl-**a***
the-child.ms-Acc
'There is the child in the library'

When the indefinite subject is preceded by the negator *lā,* it is marked for accusative case without nunation, as in (38).

(38) *lā* *ṭifl-**a*** *fī* *l-maktaba*
no child.ms-Acc in the-library
'There is no child in the library'

Finally, if the subject of an *'inna* sentence is a pronoun, it attaches/cliticizes to this particle, without being overtly marked for Case, as pronouns have frozen forms whether in their free or bound forms.

(39) *ʾinna-**hā*** *nāʾima*
ʾinna-cl3fs sleeping.fs/asleep
'She is sleeping/asleep'

A personal pronoun cannot replace the clitic pronoun in (39), rendering (40) ungrammatical.

(40) **ʾinna* ***hiya*** *nāʾima*
ʾinna she sleeping.fs/asleep

In subordinate clauses, the complementizer *ʾanna*, being a sister of *ʾinna*, performs the same function as *ʾinna* in (43) and (44) in terms of subject Case marking.

In spoken Arabic, Case is not marked morphologically on the subject. Regular masculine plurals have the accusative form, as a frozen form that does not inflect for Case, e.g. *il-mudarrisīn* 'the teachers'. Further, there is no equivalent to *ʾinna* in subject/predicate kernel sentences in the spoken variety.

7. THE CATEGORIAL STATUS OF THE SUBJECT

The subject in Arabic is normally a noun (common or proper) or a pronoun. When common, it can be modified by one or more adjectives or a → relative clause. In Arabic, however, it is common for adjectives to function as nouns and, hence, to occupy the subject position, as in (41) and (42).

(41) *al-marīḍ-u* *šufiya*
the-sick.ms-Nom recovered.3ms.Pass
lit. 'The sick person was recovered'

(42) *al-murhaqat-u* *nāmat*
the-tired.fs-Nom slept.3fs
lit. 'The exhausted [female] person slept'

Subordinate noun clauses can function as the subject of a sentence. They can occur as complements of some verbs such as *balaġa-nī* 'I heard [lit. 'it reached me]'. Traditional Arabic grammar assumes that these noun clauses function as the subject of such verbs, as in (43).

(43) *balaġa-nī* *ʾanna*
reached.3mp-me that
ṯ-ṭalabat-a *lā* *yuḏākirūna*
the-students-Acc Neg study.3mp
'I heard that the students do not study'

Similarly, the subject of some passive verbs can be a noun clause, as in (44).

(44) *ʾušīʿa* *ʾanna*
Pass.rumored.3mp that
l-waḍʿ-a *mutadahwir*
the-situation-Acc deteriorating/bad
'It is rumored that the situation has deteriorated'

In both (43) and (44), the clausal subjects must follow rather than precede the verb, i.e., the word order must strictly be VS.

The subject of a predicate can be a subordinate clause. Example (45) from the *Qurʾān* 2/184 is cited in Wright (1967:II, 252).

(45) *ʾan* *taṣūmū* *xayrun* *la-kum*
that fast.2mp better for-you.2mp
'For you to fast is for your good'

The clausal subject in (45) begins with the particle *ʾan,* which must be followed by an imperfect verb. This particle is used before verbal clauses. A nominal SV sentence can be used as a clausal subject, as in (46).

(46) *yuqālu* *ʾanna* *l-kitāb-a*
Pass.said.3mp that the-book-Acc
sa-yunšaru
Fut-Pass.publish.3mp
'It is said that the book will be published'

With respect to Case, Arabic grammarians assume that clausal subjects receive nominative case by virtue of their function. In more recent treatments of these clauses in Arabic, sentences such as (44) and (46) have been argued to contain an impersonal, expletive subject equivalent to English *it*, the subject being the complement of the verb (e.g., among many others, Mohammad 1990).

While there is no equivalent for *ʾinna* in spoken Arabic, the equivalent of *ʾanna* is *ʾinn*, a widely used noun clause complementizer.

Bibliographical references

Primary source
Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*, I. Bulaq, 1898.

Secondary sources
Badawi, Elsaid, Michael G. Carter, and Adrian Gully. 2004. *Modern Written Arabic: A comprehensive grammar.* London and New York: Routledge.
Belnap, R. Kirk. 1993. "The meaning of deflected/strict agreement variation in Cairene Arabic". *Perspectives on Arabic linguistics*, V, ed. Mushira Eid and Clive Holes, 119–147. Amsterdam and Philadelphia: J. Benjamins.
Cowan, David. 1958. *An introduction to Modern Literary Arabic.* Cambridge: Cambridge University Press.
Eid, Mushira and Robert R. Ratcliffe (eds.). 1997. *Perspectives on Arabic linguistics*, X. Amsterdam and Philadelphia: J. Benjamins.
Gaballa, Hassan. 1986. *Case-marking and Classical Arabic: A semantic and historiographic study.* Ph.D. diss., University of Delaware.
Haywood, John A. and Hayim Musa Nahmad. 1965. *A new Arabic grammar of the written language.* London: Lund Humphries.
Idrissi, Ali. 1996. "Plural formation in Arabic". Eid and Ratcliffe (1997:123–146).
Mohammad, Mohammad. 1990. *The sentential structure of Arabic.* Ph.D. diss., University of Southern California, Los Angeles.
——. 2000. *Word order, agreement and pronominalization in Standard and Palestinian Arabic.* Amsterdam and Philadelphia: J. Benjamins.
Ouhalla, Jamal. 1997. "Remarks on focus in Standard Arabic". Eid and Ratcliffe (1997:9–46).
Wright, William. 1967. *A grammar of the Arabic language.* 3rd ed. Cambridge: Cambridge University Press.

Amira Agameya (Cairo University and The American University in Cairo)

Subjunctive → Mood

Subordination

Subordination refers to a syntactic dependency between clauses in a multiclausal sentence, in which the subordinate clause must be annexed to an independent clause but not conversely. As such, subordination in Arabic is a large topic, various aspects of which are dealt with in other entries (→ circumstantial clauses; → *ḥāl* clauses; → *ʾinna wa-ʾaxawātuhā*; → relative clauses). This entry deals with the form of subordination known as complementation, whereby a subordinate clause bears a nominal function signaled by a particle introducing it, known as a complementizer.

As nominal constituents, complement clauses bear the full range of grammatical functions assigned to noun phrases (Abdul-Ghany 1981: 20–21; Badawi a.o. 2004:604–608): subject (with inversion of complement clause), as in (1); predicate (in an equational sentence), as in (2); direct object (of a transitive verb), as in (3); and object of a preposition (that is required by the matrix verb), as in (4).

(1) *ṣaḥīḥun ʾanna l-ʾujrat-a munxafiḍatun*
true that the-rate-Acc low
'It is true that the rate was low'

(2) *al-muškilatu hiya ʾanna-ka*
the-problem she that-you.ms
lā taʾtī mubakkiran
not 2fs.come early
'The problem is that you don't come early'

(3) *raʾay-tu ʾanna-hā tuḥaddiṯu*
saw-1s that-she 3fs.speak
šaxṣan mā
person some
'I saw that she was talking to somebody'

(4) *ʾaṯiqu bi ʾanna ʾallāh-a maʿ-ī*
1s.trust by that God-Acc with-me
'I trust that God is with me'

Traditional Arabic grammar, and much generative research, recognizes three major complementizers in Classical and Modern Standard Arabic. (Due to its marginal and anachronistic character [Cantarino 1975:II, 233–234], a fourth, *ʾin*, is not discussed here.) The first is *ʾanna* 'that', which introduces clauses expressing propositions that denote factual assertions (Abdul-Ghany 1981:15; Ryding 2005:425), as in (5).

(5) *ḍann-a zayd-un*
supposed-3ms Zayd-Nom that
ʾanna ʿamr-an sāfar-a
Amr-Acc traveled-3ms
'Zayd supposed that Amr (had) traveled'

(5) illustrates the salient properties of an *ʾanna* clause: (i) it is (if not a subject) the complement of a lexical head, here the verb *ḍanna* 'to suppose'; (ii) the first element is not a verb; (iii) the verb is in the indicative mood; and (iv) the noun phrase to the left of the verb is in the accusative case, with the accusative suffix *–a* when lexical, as here and in (1) and (4) above, otherwise as an accusative clitic, as in (2) and (3).

A variant of *ʾanna* is *ʾinna*, which occurs only after *qāla* 'to say'. Shlonsky (2000:336–337) proposes that the two complementizers differ in terms of the illocutionary feature Force: the asseverative complementizer *ʾinna* is [+Force], while *ʾanna* is [-Force]. Inasmuch as verbs that select *ʾanna* share asseverative force with *qāla*, e.g. *zaʿama* and *iddaʿā* 'to claim', this featural distinction is little more than a diacritic.

The other two complementizers have the form *ʾan*, a truncated form of *ʾanna*. The first of these introduces clauses expressing propositions that do not denote factual assertions, as in (6).

(6) *ʾarād-a zayd-un ʾan*
desired-3ms Zayd-Nom that
yu-sāfir-a ʿamr-un
3ms-travel-Subj Amr-Nom
'Zayd desired that Amr travel'

Substituting *ʾan* for *ʾanna* triggers three other differences from (5). First, the initial element in the subordinate clause is the verb *yusāfira* 'that he travel'. Second, the mood of the verb is not indicative because the clause denotes a future contingency rather than an accomplished fact (Cantarino 1975:III, 106; Holes 2004:279). Since it denotes an as-yet-unrealized act, this form of the verb (*al-muḍāriʿ al-manṣūb*) is often designated as subjunctive in Western grammars of Arabic (Badawi a.o. 2004:589). Finally, the postverbal subject *ʿamrun* has nominative rather than accusative case. This *ʾan* is called by the Arabic grammarians *ʾan al-maṣdariyya* 'the *ʾan* of the *maṣdar*' because its clause paraphrases with the corresponding verbal noun or → *maṣdar* when the subject of the matrix and complement clauses are coreferential (Wright 1896:II, 26; Badawi a.o. 2004:588), as in (7) and (8).

(7) *ʾarād-a zayd-un ʾan*
desired-3ms Zayd-Nom that
yu-sāfir-a
3ms-travel-Subj
'Zayd$_1$ desired that he$_1$ travel'

(8) *ʾarād-a zayd-un as-safar-a*
desired-3ms Zayd-Nom the-travel-Acc
'Zayd wanted to travel'

This interchangeability between an *ʾan* clause and a corresponding *maṣdar* is evidently why the *ʾan* clause in (6) is sometimes referred to as nonfinite (Bakir 2006:70–71), even though the subjunctive occurs in simple sentences when negated by *lan* (see below).

The other variant of *ʾan* is called *ʾan al-muxaffafa* 'the light[en]ed *ʾan*'. Like *ʾanna*, lighted *ʾan* takes a verb in the indicative mood and a subject in nominative case; unlike the other two complementizers, its word order is either SV or VS (Abdul-Ghany 1981:9), as in (9) and (10).

(9) *ḏ̣ann-a zayd-un*
supposed-3ms Zayd-Nom that
ʾan ʿamr-un sāfar-a
Amr-Nom traveled-3ms

(10) *ḏ̣ann-a zayd-un*
supposed-3ms Zayd-Nom that
ʾan sāfar-a ʿamr-un
traveled-3ms Amr-Nom
'Zayd supposed that Amr (had) traveled'

Implicit in the preceding examples is the idea that the choice of complementizer depends on the lexical semantics of the predicate that selects the complement clause. Because *ʾanna* introduces factual statements, it is selected by asseverative verbs, such as *zaʿama* 'to claim', or epistemics, such as *ḏakara* 'to remember'. The same stricture applies to *ʾan al-muxaffafa*, as (9)–(10) show. *ʾAn al-maṣdariyya*, which introduces nonfactual clauses, is selected by predicates expressing a wider range of modal meanings: volition (*ʾarāda* 'to want'), possibility (*ʾamkana* 'to be possible'), obligation (*wajaba* 'to have to, to be necessary'), command (*ʾamara* 'to order'), or fear (*xāfa* 'to fear') (Abdul-Ghany 1981:11–13; Badawi a.o. 2004:590–593). In all these cases, selection of the complementizer is obligatory. A few verbs select clauses with or without the complementizer; for example *ḏ̣anna* in (5), and verbs of appropinquation (Abdul-Ghany 1981:19; Ryding 2005:452) such as *ʾawšaka* or *kāda* 'almost; to be on the verge of', as in (11)–(13).

(11) *ḏ̣ann-a zayd-un ʿamr-an*
supposed-3ms Zayd-Nom Amr-Acc
sāfar-a
traveled-3ms
'Zayd supposed Amr (had) traveled'

(12) *ʾawšak-a ʾan*
was on the verge-3ms that
na-squṭ-a
1p-fall-Subj
'We almost fell [i.e. 'we were on the verge of falling']'

(13) *ʾiḏā ʾa-xraj-a yad-a-hu lam*
when stretched-3ms hand-Acc-his not
ya-kad ya-rā-hu
3ms-was.about.to 3ms-see-him
'When he stretched out his hand, he almost couldn't see it'

Recent research has offered more fine-grained analyses of complementizers and their selection, both semantically and syntactically. Awad (1995) shows that while in some cases the presence or absence of the complementizer *ʾinna* in Palestinian Arabic has no semantic effect, as in (14), in others its presence signals the lesser evidentiality of the complement clause and, consequently, a lower degree of commitment on the part of the speaker to its truth, as in (15)–(16).

(14) *kāl-at (ʾinna) iṭṭaqs*
said-3fs that the-weather
tḥassan
improved.3ms
'She said (that) the weather improved'

(15) *smiʿ-na* *l-wlād* *b-ilʿabu*
heard-1p the-children Progr-play-3pl
maʿ *il-xurfān*
with the-sheep
'We heard the children playing with the sheep'

(16) *smiʿ-na* *ʾinna* *l-wlād*
heard-1p that the-children
b-ilʿabu *maʿ* *il-xurfān*
Progr-play-3p with the-sheep
'We heard that the children are playing with the sheep'

Such cases apart, *ʾinna* is normally optional with verbs in the (perfect) indicative due to its semantic vacuity, as in (14). With the subjunctive, the complementizer must be omitted when the matrix predicate is a manipulative verb like *xalla* 'to make', as in (17).

(17) *layla* *xall-at* *xālid* *y-rūḥ*
Laila made-3fs Khalid 3ms-go
'Laila made Khalid go'

When the matrix predicate is an emotive verb like *ḥabb* 'to love, like', *ʾinna* is disallowed if the matrix and complement subjects co-refer (18–19), and allowed if they do not (20) (Awad 1995:4–5).

(18) *ḥabb-ēt* *a-rūḥ*
liked-1s 1s-go

(19) **ḥabb-ēt* *ʾinn-i* *a-rūḥ*
liked-1s that-1s 1s-go
'I liked to go'

(20) *ḥabb-ēt* *ʾinn-him* *y-rūḥ-u*
liked-1s that-them 3m-go-3p
'I liked them to go'

In her study of the lexical semantics of the predicates that select complement clauses, Persson (1999) argues for a correlation between the semantic category of a matrix verb and the degree to which a complement clause is integrated syntactically into the matrix clause. If the matrix verb selects as its subject an agent that exerts strong control over the agent in the complement clause, the complement subject is raised to become the direct object of the verb, as with *-hā* in (21). If the influence of the higher agent on the lower is less direct, the complement subject remains in the complement clause as in (22) (Persson 1999:193–194).

(21) *ʾajbar-a-hā* *ʿalā* *ʾan*
forced-3sm-her on that
ta-skut-a
3fs-be.silent-Subj
'He forced her to be quiet'

(22) *ʾaṣarr-a* *al-qāḍī* *ʿalā*
insisted-3ms the-judg-Nom on
ʾan *ya-jlis-a* *farġalī*
that 3ms-sit-Subj Farghali
'The judge insisted that Farghali sit'

On the syntactic front, Shlonsky (2000) applies the split-complementizer hypothesis of Rizzi (1997) to *ʾanna* clauses in Modern Standard Arabic. He proposes that the operator layer of the clause, designated as Complementizer Phrase (Shlonsky 1997:2–6; Ouhalla and Shlonsky 2002b:1–4), be decomposed into the series of functional projections, as in (23) (Shlonsky 2000:341).

(23) ForceP[hrase] > TopicP > AgrP > ℵP > TopicP* > FocusP…

ʾAnna originates as the head ℵ° of ℵP, with its associated accusative noun phrase in [Spec, ℵP]. Because such noun phrases are sometimes topics, they raise to [Spec, TopicP]. In order for the complementizer to precede the topic noun phrase, *ʾanna* must adjoin to the head of a higher projection, AgrP. However, when the accusative subject is a clitic, it serves as an agreement affix to *ʾanna* and must therefore adjoin to Force°. Thus, *ʾanna* is not licensed by the thematic layer of the clause, which is instantiated by AgrP, but by the operator layer(s), including ForceP and TopicP. One piece of evidence on which Shlonsky bases this conclusion is the well-known fact that the agreement clitic cannot be phonetically null, i.e. *pro* (Shlonsky 2000:338–339; Mohammad 1990:99; 1999:93), as shown by (24).

(24) **zaʿam-a* *zayd-un* *ʾanna*-pro
claimed-3ms Zayd-nom that-*pro*
sāfar-a
traveled-3ms
'Zayd claimed that [he] traveled'

If the clitic pronoun were governed by Agreement or Tense, then the null subject *pro* should be licensed, contrary to fact. That it is not licensed suggests that the clitic is instead licensed by *ʾanna*, within the operator layer of the clause. LeTourneau (1993) accounts for this fact by appeal to the Minimality Condition: the minimal projection containing the clitic is the intermediate projection C' rather than T' or Agr', thereby requiring that the clitic, which originates in subject position in order to be assigned its thematic role, incorporates into C as an affix.

Recently, the status of *ʾan al-maṣdariyya* as a complementizer has been called into question. Following the traditional classification, Aoun (1985:57) proposed that the complementary distribution of *ʾanna* and *ʾan* as nominal and verbal complementizers, respectively, could be captured by treating them as realizations of the feature [mood] on C: [+mood] would realize C as *ʾan*, [-mood] as *ʾanna*. Going a step further, Ouhalla and Shlonsky (2002b:18) remark that it is plausible to regard *ʾan* as the head of a Mood projection rather than as a complementizer. Ayed (2003) argues in detail for this analysis. He observes that *ʾan* distributes like the negative particle *lan*, as against *ʾanna*, in three respects: (i) it selects subjunctive mood; (ii) its complement is verb-initial; and (iii), as a special case of (ii), it cannot take a small clause complement. The first two properties were illustrated in (6), repeated here as (25); the third, and the corresponding properties of *lan,* follow in (26)–(29) (Ayed 2003:33, 35–36).

(25) *ʾarād-a zayd-un ʾan*
desired-3ms Zayd-Nom that
yu-sāfir-a ˁamr-un
3ms-travel-Subj Amr-Nom
'Zayd desired that Amr travel'

(26) **ufaḍḍilu ʾan al-walad-u*
prefer-1s that the-boy-Nom
saˁīd-un
happy-Nom
'I prefer that the boy [be] happy'

(27) *lan ya-ˁrif-a al-walad-u*
not.Fut 3ms-know-Subj the-boy-Nom
al-ḥaqīqat-a
the-truth-Acc
'The boy will not know the truth'

(28) **lan al-walad-u*$_1$ *ya-ˁrif-a*
not.Fut the-boy-Nom 3ms-know-Subj
pro$_1$ *al-ḥaqīqat-a*
the-truth-Acc
'The boy$_1$, he$_1$ will not know the truth'

(29) **lan al-walad-u saˁīd-un*
not.Fut the-boy-Nom happy-Nom
'The boy will not [be] happy'

Moreover, *lan* can co-occur with *ʾanna* (30) but not with *ʾan* (31) (Ayed 2003:36–37):

(30) *ʾa-rjū ʾanna al-walad-a*
1s-hope that the-boy-Acc
lan ya- ˁrif-a al-ḥaqīqat-a
not.fut 3ms-know-Subj the-truth-Acc

(31) **ʾa-rjū ʾan lan ya-ˁrif-a*
1s-hope that not.fut 3ms-know-Subj
al-walad-u al-ḥaqīqat-a
the-boy-Nom the-truth-Acc
'I hope that the boy will not know the truth'

From these facts, Ayed deduces that *ʾan* is not a complementizer at all. If this is correct, then the nominal vs. verbal contrast between complementizers has to be stated between *ʾanna* and *ʾan muxaffafa* rather than *ʾan maṣdariyya.* If the latter is reclassified as a formative realizing mood, then the two remaining complementizers differ not in mood selection but only in freedom of word order: *ʾanna* clauses are uniformly SVO, while *ʾan* clauses are either SVO or VSO.

The phenomenon of word order variation in complement clauses raises the question of their underlying word order. Though rarely the object of direct investigation, this has been studied in the context of three other lines of inquiry: (i) extraction out of complement clauses; (ii) the expletive hypothesis; and (iii) the status of the accusative noun phrase following *ʾanna* as a subject or a topic.

Majdi (1990) and Mohammad (1990, 1999) base their conclusions about unmarked word order on the possibilities for extraction of interrogative operators such as *man* 'who' or *māḏā* 'what' from complement clauses. (Mohammad and Majdi both take *ʾan* to be a complementizer and the preverbal noun phrase in *ʾanna* clauses to be a subject.) With *ʾan maṣdariyya*, subjects and objects behave symmetrically; either one may be extracted (Majdi 1990:144–145), as in (32) and (33).

(32) *māḏā*$_1$ *tu-rīdu ʾan ya-ktub-a*
what 2ms-want that 3ms-write-Subj
t$_1$ *salīm-un*
Salim-Nom
'What do you desire that Salim write?'

(33) *man tu-rīdu ʾan ya-ktub-a*
who 2ms-want that 3ms-write-Subj
at-taqrīr-a
the-report-Acc
'Who do you want to write the report?'

With *ʾanna* clauses, objects may be extracted whether the complementizer is present (overt) or absent (null) (Majdi 1990:146–147), as in (34) and (35).

(34) *man*$_1$ *ta-ḏ̣unnu salīm-an*
who 2ms-suppose Salim-Acc
raʾā-hu *t*$_1$
saw.3ms-him

(35) *man*$_1$ *ta-ḏ̣unnu ʾanna salīm-an*
who 2ms-suppose that Salim-Acc
raʾā-hu *t*$_1$
saw.3ms-him
'Who(m) do you suppose that Salim saw?'

However, a subject operator can be extracted from a clause with a declarative complementizer only if C is null (Majdi 1990:147), as in (36).

(36) *man ta-ḏ̣unnu raʾā*
who 2ms-suppose saw.3ms
salīm-an pro
Salim-Acc
'Who do you suppose saw Salim?'

(37) **man*$_1$ *ta-ḏ̣unnu ʾanna t*$_1$
who 2ms-suppose that
raʾā salīm-an pro
saw.3ms Salīm-Acc
'Who do you suppose that saw Salim?'

In this case, a repair strategy is available: lexicalizing the trace as a clitic pronoun in (38) renders (37) grammatical (Majdi 1990:148).

(38) *man*$_1$ *ta-ḏ̣unnu ʾanna-hu*$_1$ *raʾā*
who 2ms-suppose that-him saw.3ms
salīm-an pro
Salim-Acc
'Who do you suppose that [he] saw Salim?'

Both Majdi and Mohammad assume that a verb and its direct object form a verb phrase constituent in Standard Arabic, but they posit different basic word orders. Majdi takes the underlying order to be [VO]S, as in *ʾan* clauses, while Mohammad takes it to be S[VO], as in *ʾanna* clauses. Majdi's assumption of VOS order is the reason there is a clause-final *pro* subject in (36)–(38). To derive SVO order in *ʾanna* clauses, Majdi must assume that *salīmun* adjoins to the verb *raʾā* in (34)–(35) in order to be governed by it. Mohammad's analysis avoids this artifice (and its necessity under government, eliminated as a superfluity from minimalist theories), in favor of an independently required raising of the verb to T[ense] to check its tense feature and, in the process, derive VSO word order. (Mohammad relies in this derivation on the VP Internal Subject Hypothesis, under which the subject originates in the specifier of VP, another well-motivated proposal.) In the light of extraction phenomena, SVO seems the better choice for underlying word order.

SVO order is also implicit in the Expletive Hypothesis (Mohammad 1990, 1999), proposed to account for impoverished number agreement in VSO sentences, in which the verb is singular even when the postverbal subject is plural. The Expletive Hypothesis postulates that in VSO sentences, the verb is invariably singular because it agrees with a preceding null expletive pronoun similar to *it* in *it is odd that she left*. Mohammad justifies postulating such a null expletive based on agreement patterns with verbs in Standard Arabic that select clausal complements. The subject of the complement clause may appear to the left of the matrix predicate; when it does, the sentence (superficially) exhibits SVO order. SVO order exhibits full agreement; if the preverbal noun phrase were the subject of that verb, the latter should agree with the noun phrase in all features. However, such left-peripheral noun phrases exhibit only partial agreement with the verb, for person and (masculine) gender but not number. This pattern holds for 'raising' verbs like _badā_ 'to seem' in (39)–(40), modal verbs like _wajaba_ 'to be necessary' in (41)–(42), '_tough_ movement' predicates like _sahlun_ 'easy' in (43)–(45), and impersonal passives like _xašā_ 'to (be) fear(ed)' in (46)–(48) (Mohammad 1990:98–109; 1999:95–106).

(39) *_al-ʾawlād-u   yabd-ūna   ʾanna-hum_
the-boys-Nom   seem-3mp   that-them
sāfar-ū
traveled-3mp
*'The boys seem that they traveled'

(40) _al-ʾawlād-u_   pro   _yabd-ū_
the-boys-Nom   seem-3ms
ʾanna-hum   sāfar-ū
that-them   traveled-3mp
'The boys it seems that they traveled'

(41) _ʾayy-u   l-ʾawlād-i_   pro   _yajib-u_
which-Nom   the-boys-Gen   must-3ms
ʾan   yu-sāfir-ū
that   3m-travel-Subj.p
'Which boys must travel?'

(42) *_ʾayy-u   l-ʾawlād-i   yajib-ūna_
which-Nom   the-boys-Gen   must-3mp
ʾan   yu-sāfir-ū
that   3m-travel-Subj.p
*'Which boys are necessary that they travel?'

(43) pro   _sahl-un   ʾan_
easy.ms-Nom   that
ta-njaḥ-a   al-bint-u
3fs-succeed-Subj   the-girl-Nom
'It is easy for the girl to succeed'

(44) _al-bint-u_   pro   _sahl-un   ʾan_
the-girl-Nom   easy.ms-Nom   that
ta-njaḥ-a
3fs-succeed-Subj
'The girl, it is easy for her to succeed'

(45) *_al-bint-u_   pro   _sahl-at-un   ʾan_
the-girl-Nom   easy-fs-Nom   that
ta-njaḥ-a
3fs-succeed-Subj
'The girl is easy to succeed'

(46) pro   _yuxšā   ʾan   ta-hrub-a_
fear.Pass.3ms   that   3fs-escape-Subj
as-sajīn-at-u
the-prisoner-fs-Nom
'It is feared that the prisoner [fem.] (may) escape'

(47) *_as-sajīn-at-u   tu-xšā   ʾan_
the-prisoner-fs-Nom   3fs-fear.Pass   that
pro   _ta-hrub-a_
3fs-escape-Subj
'The prisoner [fem.] is feared that she (may) escape'

(48) *_ʾayy-u   sajīn-at-in_
which-Nom   prisoner-fs-Gen
tu-xšā   ʾan   pro   _ta-hrub-a_
3fs-fear.Pass   that   3fs-escape-Subj
*'Which prisoner [fem.] is feared that she (may) escape?'

The impoverished agreement on the verb in all these case is the result of the (full) agreement with the null expletive subject in SVO order. The absence of number agreement with a preverbal noun phrase in these constructions is one major cluster of evidence for the conclusion that raising of noun phrases to matrix subject position from subordinate clauses (as in English _the child seems likely to leave ~ the children seem likely to leave_, with full [number] agreement) does not exist in Arabic (Farghal 1993:103; Mohammad 1999:97). The reason, Mohammad speculates, is that Arabic lacks infinitives which, because of their defective inflection, require their overt subjects to raise to a finite clause to check their Case (or perhaps Extended Projection Principle [EPP]) feature. Hence, as the glosses indicate, the left-peripheral noun phrases in the grammatical sentences are construed as topics. (In keeping with its limited purpose, the preceding discussion ignores criticisms of the Expletive Hypothesis offered by Bahloul and Harbert 1993; Fassi Fehri 1993; Aoun a.o. 1994; and Benmamoun 2000. Replies to the first three of these works can be found in Mohammad 1999, Chap. 4.)

As the preceding discussion anticipates, the question of basic word order is also related to whether a preverbal noun phrase in a complement clause is a subject or a topic. That question arises because root SV(O) sentences, such as (49), are sometimes ambiguous between a subject and a topic reading (Fassi Fehri 1993:27).

(49) *al-banāt-u ji*ʾ*-na*
the-girls-Nom came-3fp
'The girls came'/'The girls, they came'

(49) contrasts with unambiguous cases in which the noun phrase is shown to be a topic by the occurrence of a resumptive pronoun, as in (50), or in which it is a subject because it is indefinite, as in (51), a feature incompatible with topicality (Fassi Fehri 1993:28):

(50) *al-*ʾ*awlād-u*$_1$ *ḍarab-tu-hum*$_1$
the-boys-Nom beat.Perf-1s-them
'The boys, I beat them'

(51) *baqar-at-un takallam-at*
cow-fs-Nom spoke-3fs
'A cow spoke'

With regard to the subject or topic status of initial noun phrases in ʾ*anna* clauses, Majdi (1990:144) is agnostic; Shlonsky (2000) and Abdul-Ghany (1981) argue that it must be a topic. Shlonsky (2000:339) denies that the NP can be a subject, because he argues, as noted above, that it is licensed by the operator layer of the clause, which includes TopicP, rather than the thematic layer, instantiated by AgrP and T[ense]P, which licenses subjects. Abdul-Ghany (1981:133) observes that the pattern of a left-peripheral noun phrase co-occurring with a coindexed resumptive pronoun – which is diagnostic for topics cross-linguistically (Rizzi 1997) – occurs in complement as well as root clauses, as in (52).

(52) *ḏ̣anna zayd-un* ʾ*anna* ʿ*amr-an*$_1$
supposed Zayd-Nom that Amr-Acc
qābal-a-hu$_1$ *xālid-un*
met-3ms-him Khalid
'Zayd supposed that Amr, Khalid met him'

Abdul-Ghany then assimilates clauses without a resumptive pronoun like (53) to the topic/comment structure of (52) by assuming that the *pro* subject of the verb is a null resumptive.

(53) *ḏ̣anna zayd-un* ʾ*anna* ʿ*amr-an*
supposed Zayd-Nom that Amr-Acc
sāfar-a pro
traveled-3ms
'Zayd supposed that Amr, he traveled'

The *pro* subject is postverbal on analogy with *xālidun* in (52). On this analysis, the basic word order in a complement clause is VS(O). The argument is not decisive in view of the (alleged) ambiguity of (49).

Possibly more telling is the fact that the left-peripheral noun phrase in declarative complement clause is normally definite – in fact, invariably so, according to Abdul-Ghany. As illustrated by (50), topics in Arabic are definite, and they cannot be indefinite, as shown by (54).

(54) ?**baqar-at-un*$_1$ *ḏabaḥ-tu-hā*$_1$
cow-fs-Nom slaughtered-1s-her
'A cow, I slaughtered it'

Conversely, indefinite noun phrases (hence, subjects) cannot antecede resumptive pronouns (Fassi Fehri 1993:29), as shown by (55).

(55) ?**kull-u rajul-in*$_1$ ʾ*a-ḥtarimu-hu*$_1$
every-Nom man-Gen 1s-respect-him
'Every man, I respect him'

On the basis of facts like these, Abdul-Ghany (1981:134) finds that an indefinite noun phrase cannot be left peripheral in a complement clause, as shown by (56).

(56) **ḏ̣anna zayd-un*
supposed-3ms Zayd-Nom that
ʾ*anna rajul-an sāfar-a*
man-Acc traveled-3ms
'Zayd supposed that a man traveled'

However, Mohammad (1999:9–10, 22) reports that the prohibition on indefinite subjects in initial position in matrix clauses, as in (57), is relaxed in complement clauses, as in (58).

(57) **rajul-un qābal-a al-walad-a*
man-Nom met-3ms the-boy-Acc
'A man met the boy'

(58) *ḥasib-a zayd-un* ʾ*anna*
thought-3sm Zayd-Nom that
ʾ*af*ʿ*-an ladaġ-at* ʿ*amr-an*
snake-Acc bit-3fs Amr-Acc
'Zayd thought that a snake bit Amr'

Thus, Mohammad agrees with Abdul-Ghany against Fassi Fehri that root clauses cannot host initial indefinites, but disagrees with him about the possibility of initial indefinites in complement clauses. (Fassi Fehri does not discuss examples like (56).) In view of this data dispute, the topic or subject status of noun phrases in *ʾanna* clauses remains unresolved.

BIBLIOGRAPHICAL REFERENCES

Abdul-Ghany, Muhammad Kamel El-Deen. 1981. *Government-binding in Classical Arabic.* Ph.D. diss., University of Texas at Austin.

Aoun, Joseph. 1985. *A grammar of anaphora.* Cambridge, Mass.: MIT Press.

——, Elabbas Benmamoun, and Dominique Sportiche. 1994. "Agreement, conjunction, and word order in some varieties of Arabic". *Linguistic Inquiry* 25.195–220.

Awad, Maher. 1995. "Commitment and evidence in Arabic complementation". Paper presented at the 26th Annual Conference on African Linguistics, Los Angeles, March 24–26, 1995. ERIC: ED 391–364.

Ayed, Hela Ben. 2003. *Mood and functional projections.* Ph.D. diss., McGill University.

Badawi, El-Said, Michael G. Carter, and Adrian Gully. 2004. *Modern Written Arabic: A comprehensive grammar.* New York: Routledge.

Bahloul, Maher and Wayne Harbert. 1993. "Agreement asymmetries in Arabic". *The Eleventh West Coast Conference on Formal Linguistics,* ed. Jonathan Mead, 15–31. Stanford, Cal.: Center for the Study of Language and Information.

Bakir, Murtada J. 2006. "Investigating control in English and Arabic". *International Journal of Arabic-English Studies* 7.65–80.

Benmamoun, Elabbas. 2000. *The feature structure of functional categories: A comparative study of Arabic dialects.* New York: Oxford University Press.

Cantarino, Vicente. 1975. *The syntax of modern Arabic prose,* II, III. Bloomington, Ind.: Indiana University Press.

Eid, Mushira (ed.). 1990. *Perspectives on Arabic linguistics*, I. Amsterdam and Philadelphia: J. Benjamins.

Farghal, Mohammad. 1993. "The theory of PRO and Arabic empty categories". *Papers and Studies in Contrastive Linguistics* 28.99–108.

Fassi Fehri, Abdelkader. 1993. *Issues in the structure of Arabic clauses and words.* Dordrecht: Kluwer Academic Publishers.

Holes, Clive. 2004. *Modern Arabic: Structures, functions, and varieties.* 2nd rev. ed., Washington, D.C.: Georgetown University Press.

LeTourneau, Mark S. 1993. "Case-marking and binding of subject clitics in Arabic complement clauses". *Perspectives on Arabic linguistics*, V, ed. Mushira Eid and Clive Holes, 261–290. Amsterdam and Philadelphia: J. Benjamins.

Majdi, Bassam. 1990. "Word order and proper government in Classical Arabic". Eid (1990:127–153).

Mohammad, Mohammad. 1990. "The problem of subject-verb agreement in Arabic: Towards a solution". Eid (1990:95–125).

——. 1999. *Word order, agreement and pronominalization in Standard and Palestinian Arabic.* Amsterdam and Philadelphia: J. Benjamins.

Ouhalla, Jamal and Ur Shlonsky (eds.). 2002a. *Themes in Arabic and Hebrew syntax.* Dordrecht: Kluwer Academic Publishers.

——. 2002b. "Introduction". Ouhalla and Shlonsky (2002a:1–43).

Persson, Maria. 1999. "Semantic considerations in the syntactic structure of complement clauses in Modern Literary Arabic". *Arabic grammar and linguistics*, ed. Yasir Suleiman, 182–202. Richmond (Surrey): Curzon.

Rizzi, Luigi. 1997. "The fine structure of the left periphery". *Elements of grammar,* ed. Lilian Hageman, 282–337. Dordrecht: Kluwer Academic Publishers.

Ryding, Karen C. 2005. *A reference grammar of Modern Standard Arabic.* Cambridge: Cambridge University Press.

Shlonsky, Ur. 1997. *Clause structure and word order in Hebrew and Arabic: An essay in comparative Semitic syntax.* Oxford and New York: Oxford University Press.

——. 2000. "Remarks on the complementizer layer of standard Arabic". *Research on Afroasiatic grammar: Papers from the Third Conference on Afroasiatic Languages*, ed. Jacqueline Lecarme, Jean Lowenstamm, and Ur Shlonsky, 325–343. Amsterdam and Philadelphia: J. Benjamins.

Wright, William. 1896. *A grammar of the Arabic language.* 2 vols. Cambridge: Cambridge University Press. (Repr., Beirut: Librairie du Liban, 1974.)

MARK S. LETOURNEAU (Weber State University)

Subsaharan Africa → West Sudanic Arabic

Substrate

1. INTRODUCTION

The term 'substrate' stems from the Latin *substratum*, which means 'underlying layer'. The two aspects of 'underlying' and 'being layered' are both essential to the various connotations of substrate in linguistics as well as in other sciences.

A substrate language is defined as a language that underlies another language. Substrate effects are the traces that a substrate language has left on another language. This underlying position of the substrate language must primarily be understood in a linguistic sense: in a language system there is an exogenous hidden layer that influences and molds the surface form of the language. The substrate effects belong to the fundamental structure of the language and do not involve only superficial phenomena, such as peripheral items of the lexicon. This underlying structured layer or stratum results from language contact with speakers of the substrate language. In most cases, the substrate effects are imposed by speakers who have abandoned their first language and switched to the new language. It was in this context of language change and language contact that the term 'substrate' emerged in the second half of the 19th century. Later, these terms were used in all kinds of historical linguistics, and in the latter half of the 20th century they became common terms in creole studies as well (cf. Goodman 1993).

2. SUBSTRATE AS SOCIALLY AND GEOGRAPHICALLY UNDERLYING

Although it is a linguistic concept, social and geographical meanings are also attached to the prefix *sub-* in substrate. Hock (1986:411) discusses three types of language contact situations, defined by lower, equal, or higher social status of one language compared with the other. In his terminology, only the language that is underlying in terms of social position is called 'substrate'; he calls the other two cases 'adstrate' and 'superstrate'. Native American Indian languages would then be substrate languages in the American colonization era, while French and Scandinavian as introduced by the Norman and Scandinavian conquerors in Britain would be superstrates and adstrates. Since social domination usually coincides with linguistic domination, a substrate language in the linguistic sense has indeed often become a *sub*strate because of its socially underlying position.

In Romance historical linguistics, Delattre (1970:480) defines the difference between substrate and superstrate in terms of (geographically) *local* vs. *strange*: "Si les habitudes articulatoires locales se maintiennent malgré l'adoption d'un vocabulaire étranger, on parle de substrat". Delattre then discusses which local language, Celtic or Germanic, influenced the 'strange' language of the Roman invaders. In this case, the geographical status of being foreign runs parallel with both a linguistic meaning of 'underlying' in the French language and a social meaning in the sense that the Roman invaders were socially dominant.

Although often the three senses of underlying run parallel, this is not necessarily the case. Speakers of creole language have indeed been repressed socially by speakers of a dominant language. However, with respect to the creoles that emerged on the plantations in the Western hemisphere, it makes no sense to claim that the language of the dominant group or that of the repressed group would be local or foreign, because both languages were introduced from other continents.

The social aspect of the underlying position seems to be related more closely with linguistic underlying than with the geographical aspect, but the two aspects cannot be taken as equal at the risk of circularity. In many instances, more or less apparent traces of substrate languages can be pointed at, but if their social history is unknown, it is not justified to infer social dominance for the nonsubstrate language. In instances where language contact took place long ago, the term 'substrate' is used on the basis of linguistic data alone. Therefore, only in combination with knowledge about social history is it proper to speak of substrates vs. superstrates, but when the social history is not known, 'substrate' is used here as a general term for a language underlying another language in a linguistic sense.

Nevertheless, the linguistic characterization of substrates implies that substrates are often more associated with interference from (nondominant) speakers giving up their own language than with (dominant) speakers borrowing from another language. Because of this confusing status of substrate between social and linguistic definitions, Thomason and Kaufman (1988:118) prefer to use the term 'interference through shift' for situations where there is no historical evidence about social positions. However, while traces of substrate languages may typically be a result of 'interference through shift', in many instances additional mechanisms of language transfer, like borrowing, → convergence, and → code-switching, have played a role as well, and the term 'interference through shift' seems too narrow to comprise all language change mechanisms that result in what are called substrate traces.

3. FOUR ASPECTS OF SUBSTRATES

The definition of a substrate as a linguistic layer that underlies another language entails four related aspects.

First, the term 'substrate' is used exclusively on the level of language as a *system*, a code or a structure. In the actual usage or performance, for instance in cases where immigrants speak a second language variety or interlanguage – influenced by their first language – we do not speak of a substrate. We only speak of a substrate when there is *structural* influence of another (substrate) language throughout a language or dialect. Although the origin of such substrate effects may very well originate from unstable and variable interlanguage phenomena, these are only regarded as substrate influence when the features have stabilized and become permanent. When exactly this is the case remains disputed.

The second aspect is that stabilization is more likely to occur when contact between speakers of the substrate language and the receiving language has ended, because then the extent of the influence and prestige of the substrate language can no longer change substantially. Therefore, it is unusual to speak of a substrate when the original language is still present. For instance, English as spoken by speakers of Indian languages for outgroup communication in India is clearly influenced on a systemic level in its rhythmic patterns and phonemic qualities by Indian languages. Nevertheless, the term 'substrate' is seldom used in such cases, precisely because the original language is still there. The term 'substrate' here has connotations related to its use in geology, where it denotes a layer lying below the present surface layer and still affecting more recent layers, although it is no longer visible. When a 'substrate' language is still spoken, some authors explicitly do not consider it to be a substrate. Diem (1979:54), for instance, rules out substrate Berber influence in Arabic in cases where speakers still speak Berber. Actual bilingualism is an indicator for Diem that we cannot yet speak of a substrate: "…[es ist nicht klar], ob die betreffende Erscheinung als Substraterscheinung, d.h. in rein arabischsprachigem Gebiet, oder als lebende berberische Interferenz in der Situation berberisch-arabischer Zweisprachigkeit einzuordnen ist". For the same reason, Versteegh (1997:104) considers Berber to be an adstrate. This means he uses the term 'adstrate' in a different sense than Hock (1986:411), because for him it has no connotations with the social position of the languages.

Third, this absence of the original substrate language also means that a substrate has a *hidden* underlying character; a substrate language is only known by its effects in another language, although it may be spoken by speakers of the substrate language in another place. This invisibility, together with its systemic character, implies that it is primarily not the lexicon but the less salient, more structural aspects of a language, like phonology or syntax, that determine whether we speak of a substrate.

Finally, what lies below must have been formed earlier – at least this is a rule of thumb in geological and archeological research of earthly (sub)strata. In linguistics, substrates also have this connotation of being older than the new languages that overrode the original autochthonous language. In cases where a Basque or Iberian substrate is suggested for Spanish, or a Dravidian substrate for Indian languages, this may be correct, both in a linguistic and an ethnohistorical sense of antiquity. However, this sense of being older must often be understood only relatively. Consider the case of creole languages, which are mergers of a language that provides lexical forms (in creole studies called the superstrate or lexifier) and a language (or languages) that to some extent provide phonological patterns or syntactic form (called the substrate). These latter languages are in no absolute sense, either genealogically or geographically, older than the superstrate. It may well be the case that in the Arabic creole, → (Ki-)Nubi, nowadays spoken in Uganda and Kenya, the African substrate languages have shorter histories than the language that provided the lexical material, Arabic. 'Younger' or 'older' in this context has to be understood only from the perspective of the speech communities involved. In an earlier period, these speech communities may have spoken African languages, while later the speakers switched toward a form of Arabic that finally resulted in a new language.

4. HOW A LANGUAGE BECOMES A SUBSTRATE

Although substrate languages are defined on a systemic level, the initial processes through which one language becomes the substrate of another take place on the individual level. Conceptually, there are three different ways that this can happen, although in practice they overlap and are not easily distinguishable. In all three cases there must have been at least some contact between speakers of language A and language B.

The first and most typical situation is when speakers of language A learn language B and retain traces of language A in language B. In this scenario, substrate effects are systemic consequences of massive second language learning. It is essential that the errors or, more neutrally, the modifications that the learners impose on the language spread from the individual to the systemic level, stabilize, and become part of the language as a system. For instance, in several Egyptian Arabic varieties, /k/ has been palatalized to /kʸ/ before short front vowels. This may be attributed to substrate influence from Coptic. Under this substrate effect, one must assume a stage where native Coptic speakers acquired a form of Arabic onto which they imposed their own phonetic assimilation rules from Coptic with respect to /k/. These individual modifications must have spread and become stabilized in varieties and during generations when the original Coptic influence was no longer there (→ Coptic loanwords).

In such cases, substrate influences are a result of interference or transfer from a source language to the target language of second language learning. Without further modification (like the one posited by Thomason and Kaufman, who add "through language shift" [1988:118]), terms like 'transfer' and 'interference' are ambiguous, because they do not make explicit *who* is the agent of the transfer or interference. Rejecting these terms, Van Coetsem (1988) proposes a model where the agent of language contact phenomena is accounted for, using the apt term 'imposition' when the transfer is effected by speakers of the source language. To distinguish such individual instances of transfer from systemic enduring effects, Van Coetsem (1988:78) distinguishes between 'inclusion' (into the individual message) and 'integration' (into the code or system).

A phenomenon is integrated in a language when it is part of the structure or code. However, it is not always clear whether something belongs to the system or only to individual usage. For instance, in modern urban dialects in countries like the Netherlands, Belgium, and France, interference from immigrants from the Maghreb results in a peculiar modification of the voicedness and place of articulation of sibilants. This new kind of Dutch, influenced by Berber and Moroccan Arabic, is spoken by at least two generations, and nowadays it is also in use among some Dutch without Berber or Moroccan background. This means that these features are spreading (cf. Hinskens a.o., forthcoming). These are indications that in the future we may speak of an Arabic or Berber substrate in some Western European urban dialects, quite apart from the question of whether this may count as substrate or adstrate, because this depends on whether or not we take the geographical detachment of the substrate speakers into account. However, it is as yet unknown how long such features will remain in the urban dialects. It may turn out that after one or several generations they are leveled out, depending on future social and political developments. A corresponding case is found in Medieval Arabic. Diem (1979:57) conjectures that in the first centuries after the Hijra there may have been more divergence due to substrate influence, but that this was leveled out through later population movements, of which some examples are found in modern Arabic as well.

When considering interference from second language learning as one of the most important factors of becoming a substrate, it should not be forgotten that interference through second language learning consists not only of errors and new formations due to the first language of the speakers, but also of general tendencies and universals of acquisition, which are not due to the first language in question. In creole formation, such changes as reduction and simplification are particularly prevalent. (Ki-)Nubi for instance, has a Cv syllable pattern. This may be due to substrate influence because the substrate languages in the region also tend to open syllables (Wellens 2005:301ff.). However, a tendency to use Cv patterns may also be a universal aspect of second language acquisition and creolization, since most creoles tend to open syllables. Similarly, the replacement of interdentals by dentals in Arabic dialects spoken in Syria and Egypt may be an effect of the substrate languages, Syriac and Coptic. However, the same change also took place in other varieties of Arabic, and it is common in language change in general. Therefore, in such cases a substrate may have helped to reinforce a change that was on its way in any event (cf. Versteegh 1997:105).

5. Borrowing, code-switching, and language areas

Although substrate influence is most often considered to be a result of interference through language shift, it is not so by definition. Substrate influence may also involve transfer activity of a speaker of the recipient language. In such a case, the term 'borrowing' is used. When borrowing only takes place on the lexical level, substrate is an unusual term, especially when the language from which borrowing takes place has (had) a socially dominant position, as, for instance, in the case of French and English words used in modern Arabic. Lexical borrowings are too visible, too much on the surface, to become a substrate phenomenon. If the association of substrate with a socially underlying group is to be retained, borrowing is not typical for a substrate, either. Nevertheless, the term substrate is used for languages in regions where toponyms stem from another language. The names of some American states, like Illinois, Michigan, and Wisconsin, for instance, stem from Amerindian languages. Yet, they may very well have been the result of borrowing by English speakers from native Amerindian speakers. In this case, only the geographical primacy of the Amerindian speakers has led to the use of the term 'substrate'.

When in more intense language contact situations more structural phenomena are borrowed, the term 'substrate' is also used (cf. Leslau 1945). The domain from which a speaker borrows a phenomenon from another language may first of all be the speech of a second language learner. In that case, borrowing is the result of imitating an imposition effect of a second language learner, and borrowing and imposition are thereby part of the same process of integration of new phenomena into a language. In such a case, the speaker of the recipient language need not be bilingual, receiving the borrowings through the intermediary of the second language learner. Secondly, borrowing may take place directly from another language. The underlying mechanism is partial or full bilingualism of the speaker of the recipient language. In such more intense language contact situations, it is hard to distinguish what exactly is due to borrowing and what to imposition. In the Algerian dialect of Djidjelli, for instance, Marçais (1956) found Berber lexical items, but also Berber gender assignment, Berber prefixes, and Berber syntax (→ Berber loanwords). In this case, it is practically impossible to find out which Berber phenomena were borrowed and which were imposed. A similar case is found in Ethiopic languages, where a Cushitic substrate has influenced all domains of the originally Semitic languages (cf. Leslau 1945).

Substrate influence may be scaled with regard to its intensity. The intensity depends on the relative number of speakers of the substrate language compared to those of the recipient language, their prestige or influence, the time scale, and the mode of acquisition of the recipient language by the speakers of the substrate language. As Thomason and Kaufman (1988:120) show, however, these factors do not yield a predictive model, which is why they rate the intensity of substratal influence purely along linguistic lines.

Within the context of bilingualism, other kinds of situations may be found where speakers accommodate to each other's patterns of speech. This may result in mutual structural convergence. A classic example of this is the Balkan language area. In the southern Balkans, several languages – Albanian, Greek, Romanian Bulgarian, Macedonian, and Serbo-Croatian – are spoken that are only indirectly related, but they share more properties than would be expected from their linguistic genealogy. These concern, among other properties, absence of nasality and length in vowels, postposed enclitic definite articles, merger of dative and genitive case, and the loss of infinitival structures in favor of finite complements. Earlier explanations have been given in terms of superstrate from Greek and in terms of unknown substrates. More recently, it has been assumed that the languages involved function as each other's substrates, although for many traits it is no longer clear what the substrate language was (Joseph 1983). Such areal influences disturb the traditional, 'clean' model of a genealogical tree. It has been argued for the Semitic languages as well that many traits must be attributed to mutual influence rather than common origin (cf. Garbini 1984). As in the Balkan case, common features in a specific language area are not sufficient by themselves to claim influence from an unknown substrate.

6. SUBSTRATE AS A FACTOR IN LANGUAGE CHANGE

Explanations of language change based on the hypothesis of substratal influence meet with several problems. First of all, if there are changes in one language or in one dialect that do not occur in other varieties, these may be ascribed to some unknown substrate language. However, when there is no linguistic evidence of languages being originally spoken in the area, such explanations are vacuous. Hock (1986:484) even argues that invoking a mysterious influence from an unknown substrate which cannot be further examined may even block further research into mechanisms of language change.

The substrate factor as an explanation for change becomes stronger when there are several languages or dialects that have changed in a common direction, in an area where one known or unknown language (family) was spoken previously. Uralic as an underlying factor of both Slavic and Baltic languages is a case in point. The Baltic languages Latvian and Lithuanian share common changes in stress pattern with Russian dialects that are spoken in the area where in the past Uralic speakers may have lived, and the word-initial stress pattern can be related to Uralic languages spoken in the area (Thomason and Kaufman 1988:238ff.).

Another problem with the substrate factor is that the more common a change is, the more difficult it is to claim that the change is caused solely by substratal influence. Such changes may also be due to chance and need no extra substrate factor as explanation. Now, since substratal influence takes place on the structural level, where in principle less variation is possible than in the lexicon, it is harder to prove than change through borrowing and is more easily confused with accidental change. For instance, the change of interdentals to dentals in Egyptian Arabic, referred to above, or changes in word order from SOV to SVO, are less likely to be due only to substratal influence.

In traditional historical linguistics, the tendency has, therefore, been to exclude all external explanations of language change unless the external explanation is completely clear-cut (cf. especially the discussion on Dravidian influence in Indian languages in Emeneau 1956). In instances of SOV/SVO word order changes, it is, therefore, less usual to invoke substratal influence. Yet, even though it cannot be proved that the substrate alone is responsible for a particular change, in many cases it seems likely that the structure of the substrate language at least has helped to change the language in a particular direction. For SVO/SOV changes in particular, it has been shown that crosslinguistically there is a quite strong areal influence on word order (Nichols 1992). When many Austronesian languages have SVO order, while most Austronesian languages that were in contact with Papuan languages have SOV order, we may not be able to prove substrate causation in each particular instance, but substrate, or at least adstrate, influence is likely, to say the least. Substratal influence may also be invoked in such a case as the sentence position of interrogative pronouns in Egyptian Arabic (→ Coptic loanwords). While there were two alternative interrogative positions in older Arabic, substratal influence may well have pushed the language in the direction of keeping to keep interrogatives *in situ* on their structural position in the sentence instead of fronting them (Versteegh 1997:106).

Diem (1979) explicitly rules out substrate influence when a certain language change also occurs in varieties that cannot be linked to the substrate language. When considering Aramaic substrate influence in Syrian-Lebanese-Palestinian Arabic varieties, he excludes changes like the voiceless realization of /q/ and elision of /u/ and /i/ because these changes also occur in several other Arabic varieties. Therefore, although substratal influence cannot be proven, it is again not unlikely that, since these changes correspond to changes in Aramaic, influence from Aramaic has at least not hindered these developments. In the Mosul dialect of Iraqi Arabic, there is a final conjunction *di-* that corresponds to an Aramaic particle. This may seem a good candidate for Aramaic substratal influence since it is not found elsewhere. However, since related Iraqi varieties have similar particles, such as *tə-* or *da-*, Diem (1979:42) doubts that this is a case of substratal influence. Again, we may be dealing here with an independent internal change which has, however, received its shape due to Aramaic substratal influence.

Historical linguists tend to be rather skeptical with regard to substrate explanations (cf. Hock 1986:483; Diem 1979). This skepticism is a reaction to overly bold claims by earlier historical linguists, but it may also be a proper characteristic of the discipline of historical linguistics and its methodology. Traditional historical linguistics seeks to construct genealogical trees and protolanguages. Explanations in terms of internal laws are therefore more attractive, since with their help a genealogical tree can be built. External explanations, however, are considered to be more like disturbances of the ideal tree structure. In modern creole studies and language contact studies, the opposite is true (cf. Thomason and Kaufman 1988). Although historical explanations of creole structures have also been given in terms of the superstrate or lexifier language, nowadays there is a strong research program in creole studies that attempts to assess to what extent, for instance, West African languages have influenced creoles that have as their lexical bases European languages. The aim of this program is not to (re)build family trees, but to assess what sources are used in particular creole formation and to gain insight into language contact in general. In their turn, however, creolists working from a generative perspective tend to be more skeptical of supposed substrate influence, since they hope to find structures from an innate bioprogram preferably not contaminated by inherited structures from substrates.

Nevertheless, irrespective of whether one takes a ‘substratomaniac’ or a ‘substratophobic’ position, the case for substrate influence becomes much stronger when independent knowledge about the substrate language(s) and their social history is available. In fact, Diem (1979:16) claims that explanations in terms of substrate may only be invoked if the original language is known. A case from Arabic are the varieties spoken in the western Yemenite mountains. Southeast Semitic languages spoken in Zufar and Soqotra, as well as Ethiopic languages, have been named as substrates in order to explain a range of striking differences from other Arabic varieties (cf. Diem 1979:18ff.). The explanation in this case does fit in with what we know about the history of the area.

The specificity of how precise our knowledge of the original language must be varies. In the Arabic creole (Ki-)Nubi, /l/, /r/, /d/, and /n/ have irregularly alternated, and in some instances they are allophonic, something which does not occur in any other Arabic variety. This may safely be attributed to influence from African languages spoken in the area, which commonly have these apicals in complementary distribution, even though the exact substrate language may be unknown (Wellens 2005:295ff.).

On the other hand, even knowledge about the substrate language does not warrant that substratal influence can be attested beyond any doubt. It still remains to be seen whether the knowledge about a language corresponds exactly to the stage of the language at the moment of language contact. For instance, Diem rejects an explanation in terms of Aramaic substrate in the Palestinian consonantal change of /q/ > /k/ and /k/ > /c/. Although in New Aramaic dialects there are varieties in which /q/ is lacking and in which there is palatalization, the diversity in this respect among the Aramaic dialects leads Diem to reject substratal influence here. In yet other cases, data from social history are available but they do not correspond with what is known about the language. Changes like the consonant lenition in West Germanic and Romance varieties could be attributed to a Celtic substrate. However, since these lenitions took place hundreds of years after the last Celts stopped speaking Celtic, this explanation loses in force. Yet, although obviously *direct* substratal influence in these cases is out of the question, *indirect* substrate influence cannot be excluded either. Influence can be indirect in the sense of historical delay, alluded to above. If substrate factors were involved, for instance in the stress shift in early Germanic, their effects would indeed range over hundreds of years, because even today deflection is continuing in several Germanic varieties that might be attributed to this early language contact (cf. Salmons 1992). Thus, the case of Celtic substrate may not be unfounded after all.

Another form of *indirect* substratal influence does not concern the transfer of properties of the substrate language themselves to the surface language but rather the introduction of structural changes by the mere fact of language shift and second language learning. In Ecuadorian Quechua, the loss of object agreement and complex affixes can be attributed to substrate Barbacoan languages that also lack such structures. Although one might hesitate to call this substratal influence, the makeup of modern Quechua in Ecuador seems to be shaped under the influence of substrate speakers. In this case, no specific phonological or syntactic patterns could be attributed to any substrate, yet, specific influence is observed when Ecuadorian Quechua is compared with Bolivian Quechua. In Bolivia, the substrate language of Aymara, which is typologically much closer to Quechua, has probably been involved in the retention of much of Quechua verbal morphological structure, while the modifications of verbal plural agreement affixes do indeed run parallel to plural formation in Aymara (see Kusters 2003:276ff.). Thus, in the Bolivian case one might paradoxically say that substrate influence has been such that it has resulted in relatively little change, while Ecuadorian Quechua may have been deflected under indirect influence of particular substrates.

The issue of whether changes can be attributed directly to a substrate or indirectly to processes of second language acquisition, leveling, and simplification is also at stake in Arabic linguistics. In some cases, new, more analytic and transparent constructions like the analytical possessive or the new demonstrative pronoun in modern Arabic varieties have various shapes, which suggests that combinations of local substratal influence, dialect spread, and universals of leveling and simplification are all involved. The exact proportion of each of these factors has been under debate for the last three decades (Diem 1979; Versteegh 1984, 1997:107ff.; Kusters 2003:150ff., and references cited there). Since some of these changes have even occurred in varieties that had hardly any contact with nonnative speakers, other factors, such as ‘ immanent’ processes of change, must be invoked as well (Diem 1979:63).

7. Substrates and theoretical linguistics

From a historical perspective, substrates become interesting when one variety is compared with related varieties that lack a specific change. In theoretical linguistics, substrates as well as adstrates and superstrates have a quite distinct role. While in historical linguistics a feature is the more typically substratal the more it permeates the whole language, in theoretical linguistics the focus lies on variation and separation of two or more strata within one language. For instance, English is assumed to have a theoretically interesting (ad)strate of Romance/French because many rules like stress shift and affixation only apply to a limited set of the lexicon that may be marked as +Romance. The task of theoretical linguistics is, on the one hand, to construct a plausible grammar in which rules are as extensive as possible and do not differ for (groups of) lexical items on an ad hoc basis. On the other hand, when there is a separate layer to which different rules apply, this must also be acknowledged. In syntax one is hesitant to build variation into the grammar in order to account for substrate effects. Instead, it is preferable to speak of two language systems. In phonology, more attempts are made to reconcile different components within one grammar (cf. Ito and Mester 1995). Thus, in theoretical linguistics the focus is on the not fully integrated substrate, while in comparative and historical linguistics the incomplete integration in the genealogical tree is at stake.

In Arabic linguistics, many of the issues mentioned above come together. In the heterogeneous speech communities of the larger Arab cities, such as Cairo, dialectal and standard registers meet. Here we find dialect contact, language contact, the emergence of new features, and the maintenance of older features, which stem from older substrate languages such as Coptic, as well as from substrate dialects, such as Bedouin dialects. Questions about what counts as underlying in a social, geographical, or linguistic sense also return here.

Bibliographical references

Delattre, Pierre. 1970. “La théorie celtique et les substrats”. *Romance Philology* 23.480–491.
Diem, Werner. 1978. “Divergenz und Konvergenz im Arabischen”. *Arabica* 25.128–147.
——. 1979. “Studien zur Frage des Substrats im Arabischen”. *Der Islam* 56.12–80.
Emeneau, Murray B. 1956. “India as a linguistic area”. *Language* 32.3–16. (Repr., *Language and linguistic area: Essays by Murray B. Emeneau*, ed. Anwar S. Dil, 38–65. Stanford, Cal.: Stanford University Press, 1980.)
Garbini, Giovanni. 1984. *Le lingue semitiche: Studi di storia linguistica*. 2nd ed. Naples: Istituto Universitario Orientale. (1st ed., 1972.)
Goodman, Morris. 1993. “African substratum: Some cautionary words”. *Africanisms in Afro-American*

language varieties, ed. Salikoko Mufwene, 64–73. Athens: University of Georgia Press.
Hinskens, Frans, Esther van Krieken, Wouter Kusters, and Pieter Muysken. Forthcoming. *Studying ethnolects: Traditions, issues, and perspectives.*
Hock, Hans Henrich. 1986. *Principles of historical linguistics.* Berlin: Mouton de Gruyter.
Ito, Junko and Armin Mester. 1995. "The core-periphery structure of the lexicon and constraints on reranking". *University of Massachusetts Occasional Papers in Linguistics.* XVIII. *Papers in optimality theory*, ed. Jill Beckman, Laura Walsh Dickey, and Suzanne Urbanczyk, 181–210. Cambridge, Mass.: University of Massachusetts.
Joseph, Brian. 1983. *The Balkan infinitive.* Cambridge: Cambridge University Press.
Kusters, Wouter. 2003. *Linguistic complexity: The influence of social change on verbal inflection.* Utrecht: LOT Dissertation Series.
Leslau, Wolf. 1945. "The influence of Cushitic on the Semitic languages of Ethiopia: A problem of substratum". *Word* 1.59–82.
Marçais, Philippe. 1956. *Le parler arabe de Djidjelli.* Paris: Librairie d'Amérique et d'Orient Adrien-Maisonneuve.
Salmons, Joe. 1992. *Accentual change and language contact: Comparative survey and a case study of early northern Europe.* Stanford, Cal.: Stanford University Press.
Thomason, Sarah G. and Terence Kaufman. 1988. *Language contact, creolization, and genetic linguistics.* Berkeley: University of California Press.
Van Coetsem, Frans. 1988. *Loan phonology and the two transfer types in language contact.* Dordrecht: Foris Publications.
Versteegh, Kees. 1984. *Pidginization and creolization: The case of Arabic.* Amsterdam: J. Benjamins.
——. 1997. *The Arabic language.* Edinburgh: Edinburgh University Press.
Wang, William. "Sound change: Actuation and implementation". *Language* 51.255–281.
Wellens, Ineke. 2005. *The Nubi language of Uganda: An Arabic creole in Africa.* Leiden: E.J. Brill.

Wouter Kusters (Meertens Institute)

Sudan

1. Arabic within the linguistic map of the Sudan

Sudan is the largest country in Africa, with an area of almost 1 million square miles inhabited by about 33 million people. Like many of the tropical African countries, it is characterized by linguistic density and diversity, with more than one hundred languages (113, according to the 1956 census). These languages belong to three out of the four language families into which all the languages of Africa are classified (Greenberg 1966). The 1956 census showed that 51.4 percent of the people spoke Arabic as their mother tongue and the remaining 112 languages were spoken as native language by 48.6 percent of the total population. All recent language surveys agree that Arabic is spoken by about 80 percent as a first, second, or third language. As such, it serves as a lingua franca and facilitates communication among groups speaking more than one hundred mutually unintelligible languages spreading all over the country, including the southern region. Mother-tongue Arabic speakers make up the most economically affluent, socially prestigious, and culturally dominant ethnic group in the country, and Arabic derives its prestige from their status. It is the official language of the State according to the Constitution, and the dominant language in all other official and semiofficial domains, including, for example, mass media, politics, administration, and the army. After the Arabicization of higher education in 1991, Arabic became the official medium of instruction at all educational levels. Other languages with a large number of speakers include Dinka, Nuer, Zande, and Bari in southern Sudan; Beja in eastern Sudan; Fur and Masalit in Darfur in western Sudan; Koalib and Nyimang in the Nuba Mountains in Kordofan and Fellata or Fulani in different areas in western, central, and eastern Sudan. Although the → Nubian languages of northern Sudan, i.e. Nobiin and Dongolese, do not belong to the most widely spoken languages, their demographic minority position is counterbalanced by their historical importance and their considerable contribution to the development of Sudanese Colloquial Arabic (see below, Sec. 3).

2. Sudanese Arabic dialects and varieties

Arabic in the Sudan exists in the form of one main central dialect and a number of other regional dialects, each with distinctive linguistic features. There are three other distinctive Arabic varieties spoken by people whose mother tongue is a language other than Arabic. Qāsim (1989) attributes the multiplicity of regional dialects in the Sudan to environmental and linguistic factors. Regarding the first factor, the early immigrant Arab tribesmen of the 7th–16th centuries, in their search for pasture, settled far apart from one another, and eventually adapted themselves to the indigenous languages and environmental conditions of their new habitat. Linguistically, these immigrants originated from a number of well-known tribes that existed in Arabia at that time. Each group, due to isolation, maintained its dialectal features; the following features are only a few examples of the variants that are found (Qāsim 1989):

i. Realization of /ḥ/ as /h/ in the Hamar dialect of Kordofan, e.g. *ḥārr* > *hār* 'hot'; this is a feature of the Bani Saʿd Ibn Zayd Manāt dialect in Arabia.
ii. Deletion of the final consonants of a word (known as *tarxīm*), e.g. *alkalām* > *alkala* 'speech'; this characteristic of the Rubatab dialect of northern Sudan was also a feature of the Ṭayyiʾ dialect.
iii. Realization of the word *ʾaʿṭā* as *anṭa* 'to give' by some tribes of Kordofan was recorded in the dialects of Saʿd Ibn Bakr, Huḏayl, ʾAzd, and Qays.
iv. Rendering of the definite article *al-* as *am-* in some words all over the Sudan, e.g. *albāriḥa* > *umbāriḥ* 'yesterday', which is a feature of the Yemeni dialects.

The inhabitants of western Sudan, in particular Darfur, belong to two major distinct linguistic stocks: speakers of Nilo-Saharan languages (Fur, Masalit, Maba, Zaghawa, etc.), and the Arabic speakers who immigrated to this region mainly from across the Sahara through Fezzan (Libya). The latter speak the → Ḥassāniyya Arabic dialect, similar to the one spoken by the Shuwa Arabs in Chad and Nigeria. It is this dialect which constitutes the substratum for the → West Sudanic Arabic variety developed and used by the non-Arab communities of the region. It is the same variety that is spoken in → Chad. Its main outstanding feature is the inability of its speakers – being nonnative speakers of Arabic – to properly pronounce the Arabic emphatic sounds /ṭ, ḍ, ṣ/, and /ḏ̣/ (pronounced in Sudan as /ẓ/), the velar fricative /ġ/, and the pharyngeal sounds /ḥ/ and /ʿ/. The emphatic sounds are replaced by their simple correspondents: *ṭibb* > *tibb* 'medicine', *ḍābiṭ* > *dābit* 'officer', etc. The sound /ġ/ is realized as /x/: *ġāba* > *xāba* 'forest'. The sounds /ḥ/ and /ʿ/ are replaced by /h/ and /ʾ/, respectively, and they may disappear in final position: *ḥarb* > *harib* 'war', *ṣaḥīḥ* > *sahi* 'correct'; *ʿali* > *ʾali* [proper name], *nawʿ* > *nō* 'variety'.

The Eastern Sudanese Arabic variety is the one spoken by the Beja in the area between Kassala northward through Port Sudan, along the Red Sea up to the Egyptian border. In addition to the phonological features described for the West Sudanic dialect, the Eastern Sudanese dialect is also distinguished by its deviating word order, having SOV as a basic pattern of word order (instead of the SVO of the other dialects): *alwalad jāb almōya* 'the boy brought the water' > *alwalad almōya jāb* 'the boy the water brought'. For a descriptive grammar of the dialect of the Shukriyya tribe, see Reichmuth (1983).

The Southern Sudanese variety of Arabic is, in fact, a kind of creole Arabic that developed during the Turkish-Egyptian rule (1821–1882). It is also known as → Juba Arabic, although it is spoken with a number of variations in different parts of the southern region. It has assimilated and adapted a large number of vocabulary items and concepts from the surrounding local languages. However, its most salient feature is the drastic reduction of its grammar: no gender distinction, minimal and peculiar application of number (e.g. *ita* 'you [sg. masc. and sg. fem.]', *itakum* 'you [pl. masc. and pl. fem.]'), and reduction in tenses (e.g. *ita akal* 'you ate', *ita ge akal* 'you are eating').

3. Sudanese Colloquial Arabic (Sudan Colloquial Arabic)

The standard version of spoken Arabic from which these regional dialects vary is known in the literature as 'Sudanese Colloquial Arabic', also known as → 'Khartoum Arabic' or 'Omdurman Arabic'. It acts as the central or model dialect by which the other varieties are measured and which speakers of the other dialects and varieties strive to approximate when trying to speak more elegantly. It is spoken over a vast area extending from northern Sudan (below Nubia), along the Nile, through Greater Khartoum, and then down between the White and Blue Niles into the Gezira and further down to the edges of the Southern Blue Nile regions. Its center of gravity in central Sudan is the most developed part of the country, both socially and economically. However, it is also the form of Arabic spoken in all the large towns in the Sudan where the riverine Arabs are present in significant numbers.

Qāsim (1975) compares the conditions which led to the supremacy of this dialect with those which led to the supremacy of the Qurayš dialect of Mecca over the other dialects of Arabia. The Nile, which earlier provided a suitable settlement for the sedentary Arabs, had the same function as the Kaʿba for the Bedouin, thus rendering the area from Barbar south to the remote areas of the Gezira a melting pot of Sudanese cultural, linguistic, and economic heritage. Qāsim concludes: "The language which sprang up (in this area) is to some extent representative of all Arab tribes in the Sudan and reflects their continuous interaction with the indigenous groups" (1975:100).

Sudanese Colloquial Arabic is indebted to two local linguistic sources, Nile → Nubian and Beja, for the greater part of its non-Arabic vocabulary. Nile Nubian provided it with a large number of words pertaining to various semantic fields, including farming, fauna, and handicrafts. Most of the Nile Nubian words found in Sudanese Colloquial Arabic can be morphologically identified by the Nubian accusative/dative case marker, -(v̄)*g*/-(v̄)*k*. Examples of these are: *mārēg* 'sorghum', *kōrēg* 'shovel', *safarōg* 'curved stick used for throwing', *kadīs* 'cat', *tagarōga* ~ *tabarōga* 'round mat of palm fronds', *kabdilo* 'granary-shaped deep and broad basket made of palm fronds [used for keeping dry food stuff]', etc. Beja influence, on the other hand, is encountered in many words relating to marriage customs and traditions, such as *jirtik* 'piece of red silk worn on the wrist of bride and groom at their wedding', *suksuk* 'small beads [for decorating the bride]', *šabbāl* 'the act whereby a woman who is dancing allows her hair to touch the face of a man who approaches her to express his admiration for her dance', etc. Other fields include names of some animals and birds, such as *baʿašōm* 'fox' and *baʿanēb* 'a kind of eagle'. Moreover, all words morphologically marked by the suffix -v̄*b* originate from Beja, such as *ʿangarēb* 'wooden bed', *ʿankōlīb* 'a kind of sugarcane', *wēkāb* 'an ingredient of sauce'. This suffix is also used for designating ethnic affiliation, as in *ʿAbdallāb* 'members of Abdalla's clan or his descendants', *ʿUmarāb, ʿAliyāb*, etc.

In the course of its development, Sudanese Colloquial Arabic also acquired lexical items from other (foreign) languages, the most important of these being:

From Turkish: Mostly words in the semantic fields of professions, administration, and the military, often morphologically identified through the affixes *-ji* denoting profession, *-xāna* 'place', and *bāš-* 'senior'. Examples: *busṭaji* 'postman', *tamarji* '[male] nurse', *ʾadabxāna* 'WC', *ʾijzaxāna* 'pharmacy', *bāškātib* 'senior clerk', *ṭābiya* 'military headquarters', *gišlāg* 'military residential compound', *šāwīš* 'sergeant, corporal'. Many of these words are gradually being replaced by their Standard Arabic forms.

From Persian: Mostly names of flowers (often used as proper names for females), food, carpets, work and musical instruments, and other miscellaneous items. Examples: *yāsmīn* 'jasmine', *xušāf* 'dessert of mixed fruits and juice', *ʾiklīm* 'a kind of precious Persian carpet', *šakūš* 'hammer', *ṭambūr* 'tambourine', *yāfṭa* 'signboard', *nīšān* 'decoration [for honoring a person]'.

From English: Mostly words in the field of 'modernism'. Examples: *rādiyo* 'radio', *warša* 'workshop', *barlamān* 'parliament', *lōri* 'lorry'.

From French: Only a few words, through Egyptian Arabic. Examples: *tilifziyōn* 'television', *jarasōn* 'waiter', *(al-)madām* 'wife'.

For the last thirty years, socioeconomic and sociocultural booms have had a great impact on many aspects of life in central Sudan, including language. These are the booms of *iġtirāb*, i.e. emigration to the oil countries in the Gulf for work, and the 'mania' of the Egyptian television film series. The *muġtaribīn* 'emigrants' brought back Arabian words and concepts such as *sayyāra* (instead of *ʿarabiyya*) 'car', *dawām* 'working hours of the day', *jannab* 'to put aside, park [a car]' and *ʿamm* 'uncle [in the sense of boss]'. Moreover, through these emigrants, Sudanese Colloquial Arabic words relating to travel abroad, which were hitherto very infrequently used, have now been promoted to the status of active vocabulary. Examples: *taʾšira* or *fīza* 'visa', *okke* 'OK [in flights booking]', *šēk siyāḥi* 'traveler's check'.

The influence of Egyptian television serials, on the other hand, is reflected by the newly spreading trend of using the 3rd person possessive pronoun with kinship terms as terms of address for the 1st person, e.g. *ʿammu* 'his paternal uncle' and *xālu* 'his maternal uncle', when addressing one's own uncles. Another new fashion is the naming of children, especially females, after Egyptian film stars (Shirin, Nasrin, etc.).

At present, Sudanese Colloquial Arabic is in the process of borrowing and integrating new words from computer language. These are mostly verbs adapted from English, such as *farmat* ‘to format’, *sakkan* ‘to scan’ (particularly when referring to scanning of pictures and photographs by journalists), and *sayyaf* ‘to save’. Of course, such words still have a very restricted degree of frequency, but with the progress in computer literacy they are expected to increase in number and become more widely circulated.

3.1 *Phonology*

Four out of the 28 Old Arabic consonants do not exist in Sudanese Colloquial Arabic: /q/, /ṯ/, /ḏ/, and /ḏ̣/. They are replaced by the following phonemes:

> **q* > /g/ or /k/ and (in a few cases) /ġ/: *qāla* > *gāl* ‘to say’, *baqar* > *bagar* ‘cattle’; *qatala* > *katal* ‘to kill’, *waqt* > *wakit* ‘time’; *qānūn* > *ġānūn* ‘law’, *laqab* > *laġab* ‘nickname’
>
> **ṯ* > /t/ or /s/: *ṯalāṯa* > *talāta* ‘three’, *juṯṯa* > *jitta* ‘corpse’; *ṯabbata* > *sabbat* ‘to fix’, *wāṯiq* > *wāsiġ* ‘sure, confident’
>
> **ḏ* > /d/, /ḍ/, or /z/: *jabaḏa* > *jabad* ‘to pull’, *ḏawwaba* > *dawwab* ‘to melt [trans.]’, *ḏanab* > *ḍanab* ‘tail’, *ʾuḏn* > *ʾaḏ̣ān* ‘ear’; *ḏakara* > *zakar* ‘to mention’, *laḏīḏ* > *lazīz* ‘delicious’
>
> **ḏ̣* > /ḍ/ or /ẓ/ (emphatic alveolar voiced fricative): *ḏ̣ahr* > *ḍahar* ‘back’, *naḏ̣īf* > *naḍīf* ‘clean’; *ḏ̣āhir* > *ẓāhir* ‘clear’, *waḏ̣īfa* > *waẓīfa* ‘job’

Apart from the above changes, the following are represented by at least one example:

> **ʾ* > /w/: *ʾayna* > *wēn* ‘where?’; **b* > /m/: *barġūt* > *margōt* ‘flea’; **j* > /d/: *jayš* > *dēš* ‘army’; **ḥ* > /x/: *ḥaṭṭa* > *xatta* ‘to put’; **d* > /t/: *zaġrada* > *zaġrat* ‘to ululate’; **r* > /l/: *ḍarfa* > *ḍalfa* ‘movable part of the door, window, or cupboard’; **ṣ* > /s/: *ṣurra* > *surra* ‘navel’; **s* > /z/: *saʿaf* > *zaʿaf* ‘palm fronds’; **l* > /n/: *jibrīl* > *jibrīn* [male proper name]; **m* > /n/: *fāṭima* > *fāṭna* [female proper name]; **y* > /j/: *yarbūʿ* > *jarbūʿ* ‘mean person’ (cf. also Qāsim 1989).

The Sudanese Colloquial Arabic consonant inventory includes three consonants that do not exist in Old Arabic:

> /ẓ/, realization of Modern Standard Arabic /ḏ̣/ (see the examples *ẓāhir, waẓīfa*, above)
>
> /č/ [tʃ] (alveo-palatal voiceless affricate), as in *čat* ‘all’ (only one example)
>
> /ny/ [ɲ] (nasal palatal): *nyamak* ‘sauce of rice mixed with lentil’ (only one example)

The Sudanese Colloquial Arabic vowel system includes two vowels that do not exist in Old Arabic: /e/ and /o/ (short and long).

> /e/ may be related to Old Arabic /i/, as in *be-llēl* < *bi-llayl* ‘in the night’. Its long form /ē/ always corresponds to the diphthong /ay/: *bēt* < *bayt* ‘house’, *yārēt* < *yā layta* ‘hopefully’.
>
> /o/ may be related to Old Arabic /ū/: *gālo* < *qālū* ‘they said’. Its long form /ō/ always corresponds to the diphthong /aw/: *tōr* < *ṯawr* ‘bull, ox’, *mōt* < *mawt* ‘death’.

The syllabic system of Old Arabic is based on six types of syllables: Cv, Cv̄, CvC, Cv̄C, CvCC, and Cv̄CC (usually a gemination and only in pausa). Sudanese Colloquial Arabic uses the first four types only, illustrated by the sentence *ʾad.dē.tu ʿa.ṣīr* ‘I gave him juice’. When Modern Standard Arabic words are pronounced colloquially, the geminated consonant of the syllable CvCC is realized as a simple (nongeminated) consonant, e.g. *ḥādd* > *ḥād* ‘sharp’. The CvCC (nongeminated) syllable type is also avoided in colloquial pronunciation through the insertion of epenthetic vowels /a/, /i/, and /u/) between C_2 and C_3, leading eventually to the split of the syllable into two: CvCC > Cv.CvC. Insertion of a specific vowel is roughly conditioned by the grammatical category of the word in question and the type of C_2 consonant: /a/, when the word is a noun and C_2 is a back fricative consonant, e.g. *naxl* > *naxal* ‘palm trees’, *laḥm* > *laḥam* ‘meat’; /i/, when (i) the word is a noun and C_2 is not a back fricative consonant, as in *waqt* > *wakit* ‘time’, *samn* > *samin* ‘ghee, butter’; and (ii) the word is a verbal noun, as in *jald* > *jalid* ‘slashing with a whip’, *ʾaxd* > *ʾaxid* ‘taking’; /u/, when the word is a noun, C_2 is not a back consonant, and C_3 is /m/, /r/, or /l/, as in *xašm* > *xašum* ‘mouth’, *faqr* > *fagur* ‘poverty’, and *raṭl* > *raṭul* ‘pound’ (cf. also Badawi 1962).

Stress plays a more important role in Sudanese Colloquial Arabic than in Modern Standard Arabic. Here, two types of stress can be distinguished: lexical and grammatical. The position of lexical stress in Sudanese Colloquial Arabic words is as follows:

i. Monosyllabic words are usually stressed when pronounced in isolation, but in strings of utterances, the presence of stress depends on the syntactic environments in which they occur.
ii. In disyllabic words, the stress falls on the ultimate syllable, if it is of the type Cv̄C or CvC going back to CvCC-gemination.

Otherwise, it falls on the other (first) syllable: *mu.dī́r* ‘director’; *mu.hím* ‘important’, but *ġá.bi* ‘idiot’; *ʿā́r.fak* ‘I know you’.

iii. In multisyllabic words, the stress falls on:
   (a) the ultimate syllable, if it is of the type Cv̄C or CvC going back to CvCC-gemination, as in *ma.ṣā.rī́f* ‘pocket money’; *mus.ta.bíd* ‘tyrannical’;
   (b) the penultimate, if (a) does not apply and the penultimate is of the type Cv̄ or CvC, as in *dag.gṓ.na* ‘they beat us [past]’; *šā.kál.tak* ‘I quarreled with you’; or
   (c) the preceding syllable, if (a) and (b) do not apply, as in *muz.dá.li.fa* [one of the places Muslim pilgrims stop at during the pilgrimage to Mecca].

Further preceding syllables never carry stress.

Grammatical stress, on the other hand, operates in Sudanese Colloquial Arabic as follows. When the 1st person object and possessive pronouns are suffixed to nouns, they carry the stress irrespective of the syllabic structure of the noun, as in *wa.ṭa.ní* ‘my nation’ (contrasting with *wá.ṭa.ni* ‘national’), *ma.xad.dā.tí* ‘my pillows’, *ad.dā.ní* ‘he gave me’, *is.taġ.fal.ní* ‘he fooled me’. When the object pronoun of the 3rd person singular masculine *-hu* is attached to a 3rd person plural masculine verb, the pronoun is usually deleted, the plural morpheme *-ū* is shortened and lowered to *-o*, and the stress is shifted to the final syllable irrespective of the syllabic structure of the verb, as in *ʾa.ka.ló* (< *ʾakalō-hu* < *ʾakalū-hu*) ‘they ate it’ (contrasting with *ʾá.ka.lo* ‘they ate’), *ʾit.ʿal.la.mó* ‘they learned it’. This stress applies to all tenses and moods: *biyākló* ‘they will eat it’, *ʾukló* ‘eat [pl.] it!’. In Old Arabic, the feminine marker *-āʾ* of adjectives whose masculine form is of the pattern *ʾa*CC*a*C is shortened to *-a*, with the stress still maintained on the final syllable (against the lexical stress rules), as in *bē.ḍá* ‘white’, *ʿō.rá* ‘one-eyed woman’ (cf. also Badrī 1974).

Metathesis occurs in *nijiḍ* < *naḍija* ‘to ripen, become cooked’, *naʿal* < *laʿan* ‘to curse’, *maʿlaga* < *milʿaqa* ‘spoon’, *ṣāgaʿa* < *ṣāʿiqa* ‘thunder’. A consonant is added in *nagraš* < *naqaša* ‘to decorate’, *ʿangara* < *ʿunq* ‘(thick) neck’, *laxbaṭ* < *xalaṭa* ‘mix’. Consonants are deleted in *wad* < *walad* ‘boy, son’, *nuṣ* < *niṣf* ‘half’, *ʿab* < *ʿabd* ‘slave’. Assimilation occurs in *itta* < *ʾanta* ‘you [sg. masc.]’, *gutta* < *qulta* ‘you said’.

New words are coined out of whole phrases or sentences, as in *maʿleš* < *mā ʿalayhi šayʾ* [an expression for excuse], *ḥabābak* < *marḥaban bika* ‘you are welcome’, *minu* < *man huwa* ‘who [sg. masc.]?’, *mini* < *man hiya* ‘who [sg. fem.]?’, *šinu* < *ʾayyu šayʾin huwa* ‘what?’ (→ interrogative pronoun), *hassaʿ* < *ha-ssāʿa* (< *hāḏihi s-sāʿa*) ‘now’.

Blending takes place, mostly with the Beja prefix morpheme of relation *-āb*, as in *gurbāb* < Arabic *qurb* ‘loin’ + Beja *-āb* ‘sheet of women’s clothing tied around the loin’.

3.2 *Grammatical features*

The basic word order in a simple sentence is SVO, with complete disappearance of case marking: *alwalad katab alʾimtiḥān* ‘the boy wrote the examination’. In verbs conjugated in the past tense, no distinction is made between the 1st person and the 2nd person singular masculine subject pronoun, which both end in *-a*, e.g. *ʾakalta* ‘I/you [sg. masc.] ate’; weak verbs are conjugated without the *-a*: *ṣallēt* ‘I/you [sg. masc.] prayed’. In general, there is no specific marker for the future tense; both the future and the present tense are expressed through the morpheme *bi-*, and the targeted tense is understood either from the accompanying adverb or merely from the context, e.g. *bištaġil fi ljāmiʿa* ‘I work in the university’, *bukra bištaġil fi ljāmiʿa* ‘tomorrow I will work in the university’. However, a few groups of speakers use the morphemes *ḥa-* and *rāḥ* for expression of the future: *ḥa- ~ rāḥ yištaġil fi ljāmiʿa* ‘he will work in the university’. The habitual and progressive aspects are built through the grammaticalized participle *gāʿid* ‘sitting’: *gāʿid yimši lmadrasa* ‘he goes to school [regularly]’, *gāʿid yākul* ‘he is eating’, *gāʿid yijri* ‘he is running’. With the remaining tenses and moods, Sudanese Colloquial Arabic does not differ much from Modern Standard Arabic.

Gender is generally marked for both nouns and verbs: *almalik māt* ‘the king died’, *almalka mātat* ‘the queen died’. However, the dual is restricted to nouns; it is marked with the ending *-ēn: dēl waladēn* ‘these are two boys’. Nouns in dual number concord with plural verbs: *alwaladēn ʾakalu* ‘the two boys have eaten’. The sound plural is formed with *-īn*: *šufta lmudarrisīn* ‘I saw the teachers’ and *almudarrisīn jo* ‘the teachers have come’.

3.3 *Semantic features*

A considerable number of Classical Arabic words still survive in the Bedouin dialects of the Butana region of central Sudan and Kordofan (western Sudan), especially words pertaining to the realm of camel rearing and natural phenomena (e.g. weather changes, different kinds of clouds, etc.). Sudanese Colloquial Arabic, on the other hand, maintains only a small number of such archaic words, some of which may even be derived from lexical sources that had become obsolete. Examples of these are *sagaṭ* ‘cold’, derived from *saqīṭ* ‘ice, snow’, and the verb *itḥanfaš* ‘to become furious and ready to fight’, derived from the noun *ḥinfiš* ‘a kind of snake, whose veins swell when it is enraged’.

There is another group of Classical Arabic words whose original meaning has undergone various degrees of change in Sudanese Colloquial Arabic. They may have undergone semantic extension, as in the word *aštar*, which originally meant ‘a person with one eye lower than the other’ and ‘a person who does not sing or clap in harmony with the group’. The former meaning is maintained as it is in Sudanese Colloquial Arabic, whereas the latter has been extended to mean ‘someone lacking in harmony’ in general. Another example is *jadaʿ* ‘to stop feeding a beast [i.e. by throwing it away]’; this has received the general meaning ‘to throw’. A final example is *fanjari* ‘an expert in camel riding’, which has taken the meaning of ‘an elegant person’. Semantic restriction is found in the word *šallag*, originally ‘to cut or pierce vertically’, but in Sudanese Colloquial Arabic specifically ‘to carry out eye surgery’ (hence *šallāg* ‘eye surgeon’). Another example is *dabīb* < *dābb* ‘creeping animal’, which is restricted to snakes in Sudanese Colloquial Arabic.

Additionally, Sudanese Colloquial Arabic includes a substantial number of idiomatic expressions whose origin can hardly ever be traced to Classical or Modern Standard Arabic. Examples of these are *atturāba fi xašmak* lit. ‘ground is in your mouth’ and *xum wu ṣur* ‘gather and make a bundle’, which both mean ‘you are just wasting your time’, said in anticipation of disappointment. There also exists an entire group of idioms constructed with the names of bodily parts for the description of positive or negative qualities. Examples of these are *baṭnu ġarīga* lit. ‘his stomach is deep’, i.e. ‘he hardly ever reveals information on his personal affairs to others’; *galbu ʾabyaḍ* lit. ‘his heart is white’, i.e. ‘he is kindhearted’; *galbu ʾaswad* lit. ‘his heart is black’, i.e. ‘he is cruel’; *lisānu zifir* lit. ‘his tongue is rancid’, said of someone who speaks obscenities; *ʿēnu ḥārra* lit. ‘his eye is hot’, i.e. ‘he may cast an evil eye’; *ʾīdu xafīfa* lit. ‘his hand is light’, i.e. ‘he is a thief’, but *ʾīdu lāḥga* i.e. ‘his hand is long enough [to attain the target]’, which means that he is influential.

4. GENDER LANGUAGE AND SOCIAL JARGONS

Aside from the regional Arabic dialects, Arabic varieties, and the Sudanese colloquial dialect, there are a number of linguistic styles related to gender, as well as social and occupational → jargons. The best examples among these are women’s language and the university students’ jargon.

Arabic in the Sudan includes a sizable corpus of vocabulary used by women (→ language and gender). There is also a distinguished speech art or style marked as exclusively ‘women’s language’, which falls outside the realm of the speech or usage of men. Examples of such vocabulary are *sajami* ‘my ashes’, said in reaction to bad luck; *gādir allah* ‘God is omnipotent’, said as an expression of wonder; *wōb ʿaley* ‘lamentation on me’, said to express sorrow; *bari* (< *barīʾa* ‘innocent’), said to express denial or negation when something is unpleasant; etc.

Women’s language in Sudan is also characterized by the frequent use of stylistic embellishments such as alliteration, simile, and proverbs, e.g. *albiḍman arrijāl yiḍman mōya fi lġarbāl* ‘having confidence in men is like trusting a sieve to hold water’. Other remarkable features of this language relate to the special use of colors, such as *lēmūni* (the color of the lemon) ‘dark green’ and *lōn zeynab* ‘Zeynab’s color’, i.e. ‘light green’. In addition to these is the women’s peculiar use of the adjective *ḥilu* ‘sweet’ (in the sense of ‘beautiful’), which may be used in describing, for example, a dress: *fustān ḥilu* ‘a sweet [i.e. beautiful] dress’, *ʿindā ḥalāt ʿigid* ‘what a sweet necklace she has’.

The university students’ jargon predominates mostly in the well-established universities in

Greater Khartoum, such as the University of Khartoum and the Ahlia and Ahfad Universities in Omdurman. A large number of students from wealthy families in high social classes enroll in these universities. The vocabulary that characterizes their jargon derives mainly from the nature of the university milieu that is mostly adapted from English. Examples: *barlōm/barālma* ‘preliminary year student [i.e. a student who is easy to fool, being alien to the university milieu]’, *lekšar/lakāšir* ‘lecture’, *gāb/gābāt* ‘gaps in examinations between one paper and another’, *murabbit* ‘repeater’, *mudabris* ‘depressed’, *mutanšin* ‘under tension’, *kēs* ‘a girl student friend’, *mukayyis* ‘in love with a girl student’, *xāliṭ* ‘mixing [i.e. accompanying a female colleague]’, etc.

BIBLIOGRAPHICAL REFERENCES

Abu-Manga, Al-Amin. 1999. *Hausa in the Sudan: Process of adaptation to Arabic*. Cologne: R. Köppe.

Badawī, as-Saʿīd Muḥammad. 1962. “Dirāsa muqārina fī baʿḍ ʾāṯār al-xilāf al-maqṭaʿī bayna l-fuṣḥā wa-ʿāmmiyya ʾUmdurmān”. *Majallat ad-Dirāsāt as-Sūdāniyya* 1.1–29.

Badrī, Kamāl. 1974. “al-Qīma al-xilāfiyya li-n-nabr fī ʿāmmiyya ʾUmdurmān”. *Majallat ad-Dirāsāt as-Sūdāniyya* 4.127–142.

Greenberg, Joseph H. 1966. *The languages of Africa*. 2nd rev. ed. The Hague and Paris: Mouton.

Ḥasan, Yūsuf Faḍl. 1973. *The Arabs and the Sudan*. Khartoum: Khartoum University Press.

Hillelson, S. 1935. *Sudan Arabic texts*. Cambridge: Cambridge University Press.

Miller, Catherine. 1979–1984. “Aperçu du système verbal en Juba-Arabic”. *Comptes Rendus du Groupe Linguistique des Etudes Camito-Sémitiques* 24–28:2.295–315.

——. 1983. “Le Juba-Arabic, une lingua-franca du Sudan méridional: Remarques sur le fonctionnement du verbe”. *Cahiers du MAS-GELLAS* 1.105–118.

Qāsim, ʿAwn aš-Šarīf. 1972. *Qāmūs al-lahja al-ʿāmmiyya fī s-Sūdān*. Khartoum: Šuʿba ʾAbḥāṯ as-Sūdān bi-Jāmiʿat al-Xurṭūm.

——. 1975. “Sudanese Colloquial Arabic in social and historical perspective”. *Directions in Sudanese linguistics and folklore*, ed. Sayyid H. Ḥurreiz and Herman Bell, 94–101. Khartoum: Khartoum University Press.

——. 1989. *al-ʿArabiyya wa-l-ʾIslām fī s-Sūdān*. Beirut: Dār al-Jīl.

Reichmuth, Stefan. 1983. *Der arabische Dialekt der Šukriyya im Ostsudan*. Hildesheim: G. Olms.

Smith, Ian and Morris T. Ama. 2005. *Juba Arabic – English dictionary. Kamuus ta arabi juba wa ingiliizi*. Kampala: Fountain Publishers.

Watson, Richard L. 1989, “An introduction to Juba Arabic”. *Occasional Papers in the Study of Sudanese Languages* 6.95–117.

AL-AMIN ABU-MANGA (University of Khartoum)

Superlative → Elative

Swahili

1. SWAHILI AND ARABIC

Swahili is a Bantu language, more specifically a member of the Sabaki subgroup of North East Coast Bantu. It has been suggested that the ancestor of the modern dialects was spoken in an area along the East African coast, somewhere between the Webi Shebelle River in what is now Somalia and the Tana River in Kenya (Nurse and Spear 1985:46; Nurse and Hinnebusch 1993:490–496). Between 1100 and 1500 C.E., the Swahili dominated trade between the African interior and the Indian Ocean, a hegemony that was interrupted in the 16th century by invasions from the south by the Portuguese and from the north by Orma and Somali raids. In the course of the 18th and 19th centuries, under the rule of Omani Arab sultans who ousted the Portuguese and established a court in Zanzibar, the Swahili regained their position as mediators of trade between the mainland and the Indian Ocean, establishing caravan routes and trading outposts between Zanzibar and places as far west as Uganda, Burundi, and Zaire (Contini-Morava 1997:841).

Swahili was first written in Arabic script, in which there are manuscripts dating back to the early 18th century. The earliest extant Swahili manuscript poem known to scholars is *Utendi Wa Tambuka*, an epic or heroic poem written in Pate for Fumo (Sultan) Laiti Nabhani and dated 1728 C.E. (in the Library of the Seminar für afrikanische Sprachen, Hamburg, no. 3554 H. 119). A Roman alphabet which is now standard was introduced during the colonial period, although some Swahili-speaking Muslims continue to use Arabic script, especially in private correspondence. The cultural importance of Islam is reflected in the large number of loanwords from Arabic. Indeed, this misled some early scholars to describe Swahili as a ‘mixed language’, a view that persists today among many East Africans (Contini-Morava 1997:842), probably also because its name is derived from an Arabic word *sawāḥil*, plural of *sāḥil* ‘coast’ with the suffix *-ī*, i.e. ‘the coastal language’. In the middle of the 20th century, Tucker (1946–1947:854, n. 3) observed: "The often heard view that Swahili is nothing but an Arabic patois is due to lack of knowledge of what characterizes a Bantu language, and has already been sufficiently dealt with by previous writers to render unnecessary a discussion of the subject here". The numerous Arabic loanwords in Swahili are clustered in several fields of cultural vocabulary, relating to jurisprudence (*sheria/sharia* < *šarīˁa* 'Muslim law'), trade (*tajiri* < *tājir* 'merchant'), religion (*hutuba/hotuba* < *xuṭba* 'Muslim Friday sermon'), nonindigenous flora (*zeituni* < *zaytūn* 'olive'), maritime affairs (*merikebu* < *markab* 'ship'), science and culture (*elimu* 'knowledge, education' < *ˁilm* 'knowledge'; *lafudhi/lafidhi* < *lafḏ̣* 'pronunciation'), and names of some everyday objects (*sabuni* < *ṣābūn* 'soap'; *subili* < *ṣabir, ṣabr* 'aloe'). This has led to statements that up to 50 percent of the Swahili lexicon is of Arabic origin. But the level of frequency of Arabic loans in basic vocabulary is much lower (Nurse and Spear 1985:15). In fact, Bertoncini (1971:150) gives a percentage of 29.48 percent in magazines, 30.56 percent in newspapers, and 34.28 percent in *Swahili* (the journal of the Swahili Institute in Dar es Salaam). Since the author marks as Arabic words from any 'Oriental' language, the real presence of Arabic words is probably lower. In another work, Bertoncini (1973:302–303) gives for all words from Oriental languages a range varying between 31.84 percent in journalistic texts, 36.03 percent in miscellaneous contemporary texts, and 62.42 percent in ancient texts.

The grammatical structure and the core vocabulary of Swahili are unambiguously Bantu, and the majority of Arabic loanwords entered the language relatively recently, most dating back only as far as the period of Omani Arab domination in the 18th and 19th centuries (Nurse and Hinnebusch 1993:315). Although the grammatical structure of the language has been unaffected by its contact with Arabic, the phonological system has absorbed some Arabic sounds along with the borrowed vocabulary (Contini-Morava 1997:842).

Standard Swahili is based on the dialect of Zanzibar City, part of the Southern group (Nurse and Spear 1985:61–62; see also Batibo [1989] for discussion of differences between this variety and Standard Swahili, especially that of Dar es Salaam, the capital of Tanzania).

2. PHONOLOGY

2.1 *Phonemes*

Arabic has introduced into Swahili some new phonemes, realized as follows by the average educated speaker:

/ṯ/ [θ]	(unvoiced dental fricative): *thabiti* (< *ṯābit*) 'firm; brave'; *thelathini* (< *ṯalāṯīna*) 'thirty'; *theluji* (< *ṯulūj*, pl. of *ṯalj*) 'snow'; *-thubutu* 'to have courage' (< *ṯubūt* 'sureness')
/x/ [x]	(unvoiced velar fricative): *habari* (< *xabar*) 'news'; *hofu/hawafu* (< *xawf*) 'fear'; *husuma* (< *xuṣūma*) 'dispute'
/ḏ/ [ð]	(voiced dental fricative): *dhahabu* (< *ḏahab*) 'gold'; *dhikiri* (< *ḏikr*) 'mention of God's name'); *dhiraa* (< *ḏirāˁ*) 'cubit'
/ġ/ [ɣ]	(voiced velar fricative): *ghali* 'scarce' (< *ġālī* 'expensive, valuable'); *ghera* (< *ġayra*) 'jealousy'; *ghofira* (< *ġafr*) 'forgiveness'

There is some variation among Swahili speakers in the pronunciation of these loanwords. The borrowed phonemes are most likely to occur in the speech of Muslim native speakers from the coast, who have had some exposure to Arabic, and for whom pronunciation of these sounds as close as possible to the Arabic model is a matter of prestige. In the speech of non-Muslims and nonnative speakers, the phonemes /ṯ x ḏ ġ/ are generally replaced with /s h z ɠ/ respectively. It should also be pointed out that loanwords have reinforced the functional load of /h/, /r/, /ɓ/, /ɗ/, and /š/, which originally had a much more restricted distribution than they do now (Nurse and Hinnebusch 1993:312).

In highly formal speech, such as a recitation in a mosque, an even more Arabized pronunciation of Arabic loanwords may be encountered, including pharyngealized-velarized ('emphatic') pronunciation of /t/, /s/, and /ḏ̣/ with the appropriate allophones of the following vowel, velarized [ɫ], [q] for /k/, interdental [t̤] for /t/, use of pharyngeal fricatives [ħ] and [ʕ] (Polomé 1967:45–46; Tucker 1946:861–867), and geminated consonants (Nurse and Hinnebusch 1993:567; Tucker and Ashton 1942:99). These do not normally occur in casual speech (Contini-Morava 1997:849). In fact, we have: *dubu* (< *dubb*) 'bear'; *hata* (< *ḥattā*) 'until, up'; *Rok* (< *ruxx*) 'the gigantic bird of Eastern tales'; *budi* (< *budd*) 'way out, escape'; *mwadhini* (< *muʾaḏḏin*) 'muezzin'; *hari* (< *ḥarr*) 'heat'; *hasa* 'expressly' (< *xāṣṣ* 'special'); *hati* (< *xaṭṭ*) 'writing'; *bazazi* 'trader' (< *bazzāz* 'draper, cloth merchant'); *hadhi* 'comfortable circumstances' (< *ḥaḏ̣ḏ̣* 'good luck'); *dafi* (< *daff*) 'tambourine'; *haki* (< *ḥaqq*) 'right'; *mdeki* (< *midakk*) 'ramrod'; *ila* (< *ʾillā*) 'except'; *umati* 'crowd' (< *ʾumma* 'nation, people'); *ina* (< *ʾinna*) 'truly'; *bawabu* (< *bawwāb*) 'doorman'; *ubaini* 'clearness' (< *bayyin* 'clear'). In a few cases, however, double consonants may still be heard: *Allah, hatta, henna, Sunni, umma*.

In Arabized speech styles, [ai], [ei], and [au] in Arabic loanwords such as *sháuri* 'intention' may be pronounced as diphthongs, but there is a tendency either to give syllabic value to each part of the diphthongs, resulting in a disyllabic pronunciation ([ʃauri]), or to coalesce the diphthong into a monophthong, e.g. [ʃaix] ~ [ʃeix] ~ [ʃeːx] 'chief' < Arabic *šayx* (Tucker 1946:870; Polomé 1967:48). Kaye (as cited by Contini-Morava 1997:850) points out that Omani Arabic, the source for most Arabic loanwords in Swahili, was a colloquial dialect in which the pronunciation of many of these sounds as diphthongs in formal speech is due to influence from the Classical language – the holy language of the *Qurʾān* – rather than familiarity with spoken Arabic (Contini-Morava 1997:850).

The presence of the Omani Arabic dialect is attested also by some vowels not existing in Classical Arabic: *dola/daulati* 'government, authority' (< *dawla* 'dynasty; state'); *robo* (< *rubʿ*) 'quarter'; *soko* (< *sūq*) 'market'.

The individual Arabic phonemes are rendered in Swahili as follows:

/ʾ/ > Ø: *ʾujra* > *ujira* 'hire, wages'; *ruʾyā* > *ruya* 'vision, dream'; *sawāʾ* > *sawa* 'equality'. In native and Standard Swahili there is only one case of initial *ʾa* in Arabic loans becoming *ha*-: *az-zayt* 'oil [edible, fuel, motor oil, etc.]' > *halzeti* 'olive oil'. In syllable-final position, the vowel before the /ʾ/ is geminated: *maʾrab* > *maarubu* 'purpose, intention'; *maʾkal* > *maakuli* 'food'; *maʾmūn* > *maamuna/mahamuna* 'reliable'; *juzʾ* > *juzuu* 'part (especially of the *Qurʾān*)'.

/b/ > /b/: *bizr* > *bizari* 'spice'; *jabal* > *jabali* 'mountain'; *sabab* > *sababu* 'reason, cause'.

/t/ > /t/: *tāj* > *taji* 'crown'; *xātima* > *hatima* 'end, conclusion'; *mayyit* > *maiti* 'dead, deceased'.

/ṯ/ > /th/: *ṯaman* > *thamani* 'price, value'; *maṯalan* > *mathalan* 'for example'; *ḥadīṯ* > *hadithi* 'hadith, traditions about the Prophet'. This sound reached Swahili through Arabic, which is why many Africans find it difficult to pronounce it and transform it: /ṯ/ > /s/: *ṯumn* > *sumni* 'one-eighth'; *ṯūm* > *saumu* 'garlic'.

/j/ > /j/: *jāh* > *jaha* 'honor, glory'; *najis* > *najisi* 'impure, unclean'; *ḥajj* > *haji* 'pilgrimage to Mecca'. There is one case of /j/ > /k/, because of the influence of Egyptian Arabic, in which /j/ is realized as *g: masjid* > *msikiti* 'mosque'.

/ḥ/ > /h/: *ḥukm* > *hukumu* 'judgment'; *ʾiḥrām* > *ihramu* 'garments of the Mecca pilgrim'; *lawḥ* 'slate, board' > *laha* 'a sheet of paper'.

/x/ > /h/: *xabar* > *habari* 'news'; *maxlūq* > *mahluki/mahluku* 'human being'; *barzax* > *barazahi* 'interval [from death to resurrection]'; once /x/ > Ø : *muxtaṣar* > *muhtasari/mutasari* 'summary, abstract'; in one case, /x/ > /k/: *maxṣīy* > *maksai/mahsai* 'castrated'.

/d/ > /d/: *dars* 'study; chapter [of a textbook]' > *darasa* 'class for reading or study'; *maʿdin* > *madini* 'metal'; *raʿd* > *radi* 'thunder'.

/ḏ/ > /dh/: *ḏikr* > *dhikiri* 'mention of God's name'; *ʾaḏā* > *adha* 'trouble'; *nāfiḏ* 'piercing; effective' > *-nafidhi* 'to save, help'; /ḏ/ > /l/ only in a very few cases: *bāḏinjān/bayḏinjān* > *bilingani* 'eggplant'; in one case, /ḏ/ > /th/: *juḏām* > *jethamu/jedhamu* 'leprosy'.

/r/ > /r/: *rizq* 'livelihood, subsistence; blessing [of God]' > *riziki* 'means of life'; *marham* > *marhamu/marahamu* 'ointment'; *ziyāra* > *ziara* 'visit'; but once: /r/ > /l/: *rāwaġa* III 'to cheat' > *ragai/laghai* 'a cheating person'.

/z/ > /z/: *zakāh* > *zaka/zakati* 'alms tax'; *ḥuzn* > *huzuni* 'grief, sorrow'; *ʿajūz* > *ajuza* 'old woman'.

/s/ > /s/: *samāwāt* (pl. of *samāʾ*) > *samawati* 'heaven, sky'; *nasab* > *nasaba* 'lineage'; *waswās* > *wasiwasi* 'doubt'.

/š/ > /sh/: *šamāl* > *shemali* 'north; north wind'; *rušwa* > *rushwa* 'bribe'; *jayš* > *jeshi* 'army, troop'.

/ṣ/ > /s/: *ṣadaqa* > *sadaka* 'alms; charity'; *xuṣūma* > *husuma* 'quarrel'; *naqṣ* > *nakisi* 'blemish'.

/ḍ/ > /dh/: *ḍāmin* > *dhamini* 'surety, guarantor'; *qāḍī* > *kadhi* 'judge'; *farḍ* > *faradhi/faridhi* 'religious duty'.

/ṭ/ > /t/: *ṭibb* > *tiba* ‘medicine’; *aṭlas* > *atlasi* ‘satin’; *ṣirāṭ* > *sirati* ‘way, path [over hell from which sinners fall]’.

/ḏ̣/ > /dh/: *ḏ̣ulm* > *dhulumu* ‘injustice’; *manḏ̣ar* > *mandhari* ‘appearance, aspect’; *waʿḏ̣* > *waadhi* ‘sermon’.

/ᶜ/ > Ø: *ʿidād* > *idadi* ‘number’; *sāʿa* > *saa* ‘hour’; *rubʿ* > *robo* ‘one-quarter’. In native and Standard Swahili, there is only one case of initial *ʿa* in Arabic loans becoming *ha-: ʿarūs* ‘bridegroom’ > *arusi/harusi* ‘nuptials, wedding’. In final position of a syllable and followed by a consonant, this phoneme gives in Swahili, very often (but not always, e.g. *raʿd* > *radi* ‘thunder’) a vowel, which is identical with the preceding vowel: *baʿda* > *baada* ‘after’; *baʿḍ* > *baadhi* ‘portion’; *daʿwā* > *daawa* ‘legal claim’. But /ᶜ/ disappears when followed by a vowel: *daʿib* ‘joking, jolly’ > *daba* ‘fool, simpleton’; *duʿāʾ* > *dua* ‘prayer’; *dufʿa* > *defa* ‘time’.

/ġ/ > /gh/: *ġaraḍ* > *gharadhi* ‘aim, object’; *maġrib* > *magharibi/mangharibi* ‘prayer at sunset’; *ʾaġlab* ‘prevalent’ > *aghalabu/aghlabu* ‘usually’.

/f/ > /f/: *fāsiq* > *fasiki* ‘profligate’; *kāfir* > *kafiri* ‘infidel’; *ḥarf* > *herufi* ‘letter’.

/q/ > /k/: *qabr* > *kaburi* ‘grave’; *nuqṭa* > *nukta* ‘point’; *ḥaqq* ‘truth’ > *haki* ‘justice’.

/k/ > /k/: *kalima* > *kalima* ‘word’; *baraka* > *baraka* ‘blessing’; *šakk* > *shaka* ‘doubt’.

/l/ > /l/: *lawn* > *launi* ‘color’; *jumla* > *jumla* ‘total’; *raṭl* > *ratli/ratili* ‘a weight’.

/m/ > /m/: *miʾa* > *mia* ‘hundred’; *ʾamr* > *amri* ‘command, order’; *qalam* > *kalamu* ‘pen’.

/n/ > /n/: *niyya* > *nia* ‘intention’; *janāba* > *janaba* ‘major (religious) impurity’; *qarn* > *karini/karne/karni* ‘century’; /n/ > /m/ before *b: minbar* > *mimbari* ‘minbar, pulpit’.

/h/ > /h/: *haram* > *haram* ‘the Pyramids’; *šahāda* > *shahada* ‘creed formula’; *wajh* > *wajihi* ‘appearance’.

/w/ > /w/: *wājib* > *wajibu* ‘obligation’; *jawāb* > *jawabu* ‘answer, reply’; *naḥw* > *nahau* ‘explanation’.

/y/ > /y/: *yābis* > *yabisi/yabis* ‘dry, arid’; *qiyāma* > *kiyama* ‘resurrection’; *raʾy* > *rai* ‘opinion’.

2.2 *Syllable structure*

As a rule, Swahili words end in a vowel. Borrowed words ending in a consonant acquire additional vowels, whose nature is determined by the nature of the final consonant: thus, after labials, *u* or *o* is added; after *t, n, l/r, i*, or *e* is added (Myachina 1981:12): *adabu* ‘good manners’ (< *ʾadab*); *wakati* ‘time’ (< *waqt*); *imani* ‘faith, belief’ (< *ʾīmān*); *jahili* ‘ignorant’ (< *jāhil*); *bizari* ‘spice’ (< *bizr*).

2.3 *Borrowed consonant clusters*

If Arabic loanwords contain consonant clusters (like *st, lt, lf*, or *ks, kr, kt*) outside the Bantu phonetic pattern, Swahili tends to insert an extra vowel between the two consonants, its character being determined by the same constraints governing final vowels: *u* is inserted after labial consonants, otherwise *i: bikira* ‘virgin’ (< *bikr*); *fikira/fikara* ‘thought’ (< *fikra*); *hitilafu* ‘difference’ (< *ixtilāf*).

Sometimes a vowel is inserted that matches the vowel in the preceding or following syllable: *bahari* (< *baḥr*) ‘sea’; *huzuni* ‘grief, calamity’ (< *ḥuzn* ‘sadness’); *ibilisi* (< *ʾiblīs*) ‘devil, Satan’.

2.4 *Stress*

The general rule is that primary stress is on the penultimate syllable (which may be a syllabic nasal), in polysyllabic words. Some polysyllabic loanwords are exceptional in being stressed on the antepenultimate: *núsura* ‘almost’ (< *nazr* ‘little’). Some show variable stress placement: *lázima/lazíma* ‘necessity’ (< *lazima* ‘to be necessary’). Vitale (1982:327) suggests differentiating between ‘historical loanwords’ and ‘phonological loanwords’; the latter either are not assimilated (like [áfrika]) or are variably assimilated (like [lázima] ~ [lazíma]), and can be marked as such in the lexicon.

3. Morphological adaptation

3.1 *Arabic article*

The Arabic article is almost never agglutinated in loanwords, differently from what happens in other languages (e.g. → Hausa). The examples are very few: *alasir* (< *al-ʿaṣr*) ‘afternoon’; *alfajiri* (< *al-fajr*) ‘dawn’; *alhaji* (< *al-ḥājjī*) ‘pilgrim’; *Alhamdulillahi!* (< *al-ḥamdu lillāh*) ‘praise be to God’; *Alhamisi* (< *al-xamīs*) ‘Thursday’.

3.2 *Adverbs*

As distinct grammatical or lexical items, adverbs hardly exist in Swahili. Most of them are derived from nouns, verbs, or pronouns. The nonderived adverbs are very few in number, most of them borrowings from Arabic: *abadan* (< *ʾabadan*) ‘never’; *afadhali* ‘rather, better’ (< *ʾafḍal* ‘better’); *aghalabu/aghlabu* (< *ʾaġlab*) ‘usually’; *baada* (< *baʿda*) ‘after’; *bado* ‘not yet’ (< *baʿdu* ‘then; still, yet’); *dahari* ‘always’ (< *dahr* ‘time; age’); *daima* ‘perpetually’ (< *dāʾim* ‘lasting; perpetual’); *dike/tike* ‘exactly’ (< *bi-diqqa*); *fauka/foko* ‘more’ (< *fawqa*) ‘above’; *ghafula* ‘suddenly’ (< *ġafla* ‘negligence’); *hadhara* ‘before’ (< *ḥaḍra* ‘presence’); *halafu* ‘after a bit’ (< *xalfu* ‘back’); *hasa* ‘specially’ (< *xāṣṣ* ‘special’); *hobelahobela* ‘anyhow’ (< *xabal* ‘confusion’); *hususa* ‘expressly’ (< *xuṣūṣan* ‘especially’); *kadha wa kadha* ‘thus and thus’ (< *ka-ḏā wa-ka-ḏā* ‘so and so’); *kadhalika* ‘in like manner’ (< *ka-ḏālika* ‘so, like so’); *labda* ‘possibly’ (< *lā budda* ‘definitely’); *nusura* ‘almost’ (< *nazr* ‘little’); *salimini* ‘safely’ (< *salām* ‘safety’); *sana* ‘very much’ (< *ṯanāʾ* ‘praise’?); *sawia* ‘then’ (< *sawiyyan* ‘equally’); *tasihili* ‘quickly’ (< *tashīl* ‘facilitation’); *wahedu* ‘alone’ (< *wāḥid* ‘one; sole’); *zamani* ‘formerly’ (< *zamān* ‘time’).

3.3 *Class system*

Swahili inflection is characterized by the Bantu class-prefix system. Many Arabic loanwords were included in a specific Swahili class because they fitted its semantic function and not according to their initials, which by chance could be similar to Swahili prefixes. In fact, we have in class 6: *mahari* ‘dowry’ (< *mahr*). Other Arabic loans with initial {ma} have been interpreted as forms with the zero allomorph of the {n} prefix of classes 9 and 10 (Polomé 1967:187): *maharazi* ‘shoemaker’s awl’ (< *maxāriz*, pl. of *mixraz* ‘awl’); *marijani* ‘red coral’ (< *marjān*); *mansuli* ‘woolen material’ (< *musūḥ*, pl. of *misḥ*); *majuni* ‘intoxicating sweetmeat containing Indian hemp’ (< *maʿjūn* ‘paste, cream’); *magharibi* ‘sunset, west’ (< *maġrib*); *mashariki* ‘east’ (< *mašriq*); *maskini* ‘poor’ (< *miskīn*); *majununi* ‘buffoon’ (< *majnūn*). Classes 9 and 10 contain many nouns of foreign origin, mainly Arabic, and being loanwords, these do not follow the rules of phonetic change in Swahili. Many such words have no prefix at all: *barua* ‘letter’ (< *barwa* ‘waste, scrap’); *dawa* ‘medicament’ (< *dawāʾ*); *jinsi* ‘kind, sort’ (< *jins*); *daraja* ‘bridge; rank’ (< *daraja* ‘rank’); *kofia* ‘fez, hat’ (< *kāfiyya* ‘kaffiyeh’); *safari* ‘journey’ (< *safar*); *saa* ‘hour’ (< *sāʿa*); *sahani* ‘plate’ (< *ṣaḥn*); *sabuni* ‘soap’ (< *ṣābūn*). Sometimes, Arabic loans were included in a noun class because of their initial consonants, which coincidentally fitted the Swahili system (class 7), e.g. *kitabu* ‘book’, pl. *vitabu* (< *kitāb*).

3.4 *Conjunctions*

In Swahili there are no original Bantu words functioning as conjunctions except *na*, which is composed of *-a* of relationship and *n-* of association. There are, however, various ways of joining words and sentences (Ashton 1947:197). Some are borrowings from Arabic: *ama…ama* ‘either…or’ (< *a-mā* ‘or?’); *au* ‘or’ (< *ʾaw*); *bali* ‘but’ (< *bal*); *ila* ‘except’ (< *ʾillā*); *ili* ‘in order that’ (< *li-* ‘in order that’); *kama* ‘if, whether’ (< *ka-mā* ‘as, equally, likewise’); *kusudi* ‘with the object of’ (< *qaṣada* ‘to intend’); *lakini* ‘but, nevertheless’ (< *lākin* ‘however, yet, but’); *wala…* ‘neither…nor’ (< *walla* ‘or’).

3.5 *Numerals*

Of the first ten numerals, three are of Arabic origin: *sita* ‘six’, *saba* ‘seven’, *tisa* ‘nine’. The numerals 11 to 19 coexist with the Bantu terms, but the numerals from 20 to 90, as well as the word for ‘one hundred’ (*mia*) and ‘one thousand’ (*elfu*) are all of Arabic origin. In such cases there is no concord.

3.6 *Prepositions*

In Swahili there are no original Bantu words functioning as prepositions (Ashton 1947:195), but some Arabic loanwords function as prepositions: *bila* ‘without’ (< *bi-lā*); *hata* ‘until, up to’ (< *ḥattā*).

3.7 *Verbs and verbal nouns*

Verbs and verbal nouns of Arabic origin generally show a difference in vowel quality because they were adopted directly from the corresponding Arabic forms: *abudu* ‘to worship’ (< *ʿabada*), *ibada* ‘worship’ (< *ʿibāda*); *hasibu* ‘to count’ (< *ḥasiba*), *hesabu* ‘arithmetic’ (< *ḥisāb*); *amini* ‘to believe’ (< *ʾāmana*), *imani* ‘faith’ (< *ʾīmān*). On the other hand, nouns may also be derived from Arabic verbs in a Bantu manner: *safiri* ‘to travel’ (< *sāfara*) > *msafiri* ‘a traveler’. This accounts for synonyms like *hasidi/husudu* ‘to envy’ (< *ḥasada*) > *uhasidi/husuda* ‘envy’.

3.8 *Uninflected loanwords*

These are mostly loanwords from Arabic. The following few examples, with nouns from different classes, should be sufficient to show the absence of concord for loans in adjectival position (Ashton 1947:49–50): *mtu hodari* 'a clever man', *mti hodari* 'strong, hard wood'; *watu tele* 'many people', *maji tele* 'plenty of water'; *chumba safi* 'a clean room', *maneno safi* 'a straightforward statement'; *mwezi kamili* 'a full month', *maneno haba* 'a few words'. Many of these words have abstract nouns corresponding to them: *u-hodari* 'courage; capability'; *u-safi* 'cleanliness; purity; honesty'; *u-kamili* 'completeness'; (*u-*) *haba* 'scarcity, rarity' (all in class 11); *tele* 'abundance, plenty' (in class 9).

4. SEMANTIC ANALYSIS

Arabic loanwords were introduced in all domains of the Swahili cultural lexicon (for more details, see Baldi 1988:10–53). Some examples are:

i. Nature: *hewa* (< *hawāʾ*) 'air'; *nuru* (< *nūr*) 'light'; *Thurea* (< *ṯurayyā*) 'Pleiades'; *ardhi* (< *ʾarḍ*) 'earth'; *zebaki* (< *ziʾbaq*) 'quicksilver'; *zafarani* (< *zaʿfarān*) 'saffron'; *hudhud* (< *hudhud*) 'hoopoe'.
ii. Man as a physical being: *jamala* (< *jamāl*) 'beauty'; *raha* (< *rāḥa*) 'rest, repose'; *barasi* (< *baraṣ*) 'leprosy'; *haraka* 'haste' (< *ḥaraka* 'movement'); *sahani* (< *ṣaḥn*) 'dish, plate'; *juba* (< *jubba*) 'jubbah'; *hema* (< *xayma*) 'tent'.
iii. Man as a spiritual being: *bayana* (< *bayān*) 'explanation'; *rehema* (< *raḥma*) 'mercy'; *muhali* (< *muḥāl*) 'impossible, absurd'; *duni* (< *dūn*) 'low, inferior'; *maana* (< *maʿnā*) 'meaning, sense'; *jarida* 'journal' (< *jarīda* 'newspaper'); *hekaya* (< *ḥikāya*) 'story, tale'; *Jahim* (< *jaḥim*) 'sixth of the Muslim hells'.
iv. Man as a social being: *ajali* (< *ʾajal*) 'deadline'; *dhuria* (< *ḏurriyya*) 'descendant'; *taa* (< *ṭāʿa*) 'obedience'.
v. Social organization and politics: *asili* (< *ʾaṣl*) 'origin'; *tuhuma* (< *tuhmā*) 'suspicion'; *alamu* (< *ʿalam*) 'flag'; *daftari* (< *daftar*) 'register'; *ala* (< *ʾāla*) 'tool; *badala* (< *badal*) 'substitute'; *himila* (< *ḥiml*) 'load'.
vi. Natural laws: *jinsi/jinsi* (< *jins*) 'kind'; *sudusu* (< *suds, sudus*) 'one sixth'; *mahali* (< *maḥall*) 'space'; *saa* (< *sāʿa*) 'hour'.
vii. Interjections and conjunctions: *Bismillahi* (< *bi-smi llāhi*) 'in the name of God'; *mathalan* (< *maṯalan*) 'for example'; *au* 'or' (< *ʾaw*).

5. SWAHILI AS A MEDIUM OF SPREADING ARABIC LOANS

Swahili not only received Arabic loans but it was also a donor language. Many languages in the region, both Bantu and non-Bantu, received Arabic loans through Swahili:

Arabic *sāʿa* 'while; hour; timepiece' > Swahili *saa* 'time; watch' > liNgala *sâ (sáa)* 'watch' > Sango *sâà* 'watch'.
Arabic *māl* 'money' > Swahili *mali* 'wealth' > Ila *madi* 'money'; Shona *mari* 'money (cash)'; isiXhosa *imali* 'money'.
Arabic *qahwa* 'coffee' > Swahili *kahawa* 'coffee' > liNgala *káwa* 'coffee' > Sango *káwà* 'coffee'.
Arabic *qarṭas* 'paper' > Swahili *karatasi* 'paper, a piece of paper' > Acholi *kàrtacì* 'sheet of paper'.
Arabic *kāfiyya* 'kaffiyeh' > Swahili *kofia* 'fez' > Acholi *kòfíà* 'tarboush'.

The fact that Swahili was one of the first languages in Africa to be appointed as a national language, in Tanzania (1967) and Kenya, has increased its role, even outside the national boundaries, so that in the near future it will no doubt continue to spread Arabic loanwords.

BIBLIOGRAPHICAL REFERENCES

Ashton, E.O. 1947. *Swahili grammar (including intonation)*. London: Longman.
Baldi, Sergio. 1988. *A first ethnolinguistic comparison of Arabic loanwords common to Hausa and Swahili.* (= Suppl. 57 to *Annali dell' Istituto Universitario Orientale di Napoli* 48:4). Naples.
Batibo, Herman M. 1989. "Evolution and dialectisation of Standard Kiswahili". *Le Swahili et ses limites: Ambiguïté des notions reçues. Table ronde internationale du CNRS (Sèvres, 20–22 avril 1983)*, ed. Marie-Françoise Rombi, 149–156. Paris: Éditions Recherches sur les Civilisations.
Bertoncini, Elena. 1971. "Ricerche sul lessico di alcuni racconti swahili". *Annali dell' Istituto Universitario Orientale di Napolo* 31:1.141–154.
——. 1973. "A tentative frequency list of Swahili words", *Annali del Istituto Orientale di Napoli* 43:3.298–363.

Contini-Morava, Ellen. 1997. “Swahili phonology”. *Phonologies of Asia and Africa*, II, ed. Alan S. Kaye, 841–860. Winona Lake, Ind.: Eisenbrauns.

Johnson, Frederick. 1939. *A standard Swahili-English dictionary*. London and Oxford: Oxford University Press.

Myachina, Ekaterina Nikolaevna. 1981. *The Swahili language: A descriptive grammar*. London: Routledge and Kegan Paul.

Nurse, Derek and Thomas Hinnebusch. 1993. *Swahili and Sabaki: A linguistic history*. Berkeley and Los Angeles: University of California Press.

—— and Thomas Spear. 1985. *The Swahili: Reconstructing the history and language of an African society, 800–1500*. Philadelphia: University of Pennsylvania Press.

Polomé, Edgar C. 1967. *Swahili language handbook*. Washington, D.C.: Center for Applied Linguistics.

Tucker, Archibald. 1946–1947. “Foreign sounds in Swahili”. *Bulletin of the School of Oriental and African Studies* 11:4.854–871, 12:1.214–232.

—— and E.O. Ashton. 1942. “Swahili phonetics”. *African Studies* 1.77–103, 161–182.

Vitale, Anthony. 1982. “Problems of stress placement in Swahili”. *Studies in African Linguistics* 13:3.325–330.

Sergio Baldi
(University of Naples)

Swearing → Insults

Syllable Structure

1. What is a syllable?

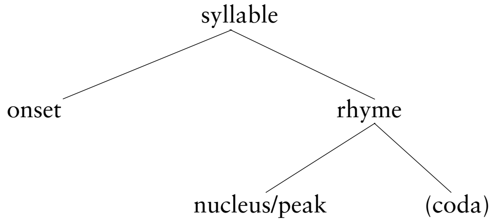
The syllable is a fundamental unit of speech in any language studied both on the phonetic and phonological level of analysis. Phonetically, syllables “are usually described as consisting of a centre which has little or no obstruction to airflow and which sounds comparatively loud; before and after that centre...there will be greater obstruction to airflow and/or less loud sound” (Roach 1991:67). Phonologically, Laver (1994:114) defines the syllable as “a complex unit made up of nuclear and marginal elements”. The nuclear elements are the *vowels* or syllabic segments, while the marginal elements are the *consonants* or nonsyllabic segments.

The syllable has a psychological reality as a unit that speakers of a language can identify; they can count the number of syllables in a word and can often tell where one syllable ends and the next begins (Cox a.o. 2004). The division of a particular word into syllables may vary from one individual to another, but it always remains easy and possible. In Arabic, for instance, the words /qaalat/ ‘she said’ and /banat/ ‘she built’ are divided into their component syllables as /qaa-lat/ and /ba-nat/, respectively. What the listeners hear are ‘peaks of sonority’ or peaks of relative loudness that represent the vocalic segments that occur in that sequence of sounds. The presence of vowels or of a sound having a high degree of sonority is an obligatory element in a syllable. Thus, because in Arabic the vowel, whether short or long, occurs only as the nucleus, and all the consonants, including the sonorants /y/ and /w/ occur only as the marginal elements, there is a clear-cut division of vowels and consonants (Al-Ani and May 1978). This facilitates the process of segmenting correctly and easily almost any utterance in Arabic.

2. Syllable structure

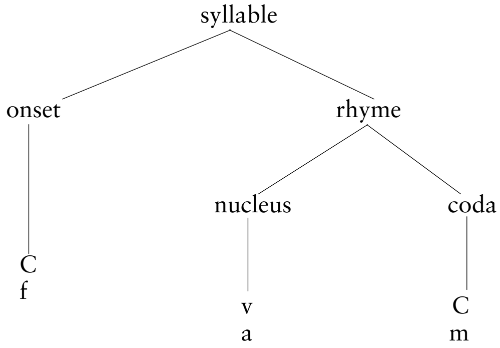
It is well established that the syllable has constituent or hierarchical rather than linear structure. In Arabic, the syllable structure can be divided into two constituents: the obligatory *onset*, which consists of a single consonant that precedes the nuclear element (the vowel), and the *rhyme*, which contains the nuclear element (or the *nucleus*) – also known as *peak* – as well as any optional marginal elements (consonants) (or the *coda*) that might follow it. The nucleus is always the most prominent or sonorant element of the Arabic syllable. It must be composed of any of the three short vowels or their long counterparts. The coda includes all consonants that follow the nucleus in a syllable. The coda may consist of zero, one, or two consonants. The syllable structure can thus be represented in the tree diagram in Figure 1:

Figure 1. Syllable structure



A word such as /fam/ ‘mouth’ has [f] as onset, [a] as nucleus, and [m] as coda. This monosyllabic word can be represented as CvC and has the tree diagram shown in Figure 2:

Figure 2. Syllable structure of /fam/



3. Syllable types

Unlike other languages such as English and German that allow complex consonantal clusters in both the onset and the coda of syllables and hence have a large number of syllable types (Kiraz and Möbius 1998), Arabic exhibits a restricted inventory of syllable types. According to most researchers (e.g. Al-Ani 1970), there are five syllable types in Arabic. Others (e.g. Al-Ani and May 1978), however, suggest six types:

(1) Cv	/li/	prepositional prefix ‘for, to’
(2) Cvv	/maa/	‘what’
(3) CvC	/sin/	‘tooth’
(4) CvvC	/baab/	‘door’
(5) CvCC	/nahr/	‘river’
(6) CvvCC	/maarr/	‘passerby’

Types (1), (2), and (3) are the most basic ones; they occur frequently and freely without restrictions in initial, medial, or final positions in an utterance in the Arabic language. The Cv type can occur as a word by itself, i.e. as a monosyllabic word like the imperative verbs /fi/ ‘keep your promise!’ and /qi/ ‘protect!’ or as the prepositional prefix /bi/ ‘in’, although it occurs more frequently in different positions within words, as in the following examples, /na-mat/ ‘it grew’, /mak-ta-bii/ ‘my office’, and /ka-tab-tu/ ‘I wrote’. The Cvv type can occur in some monosyllabic words, like the negative particle /laa/ ‘not, no’, and in different positions, as in /laa taktub/ ‘do not write!’, /saa-ʾiq/ ‘driver’, /si-baa-ḥa/ ‘swimming’, and /ka-tab-tu-maa/ ‘you [du.] wrote’. However, this pattern changes to CvC when it is immediately followed by any syllable beginning with *hamzat al-waṣl*, the assimilatory glottal stop, e.g. /fii ʾal-kitaab/ ‘in the book’. In this case, it becomes /fil-kitaab/, because the long vowel of the first syllable is shortened, and the following glottal stop and its accompanying vowel assimilate to zero (Al-Ani and May 1978).

Types (4) and (5) are restricted to the final position of words and utterances in pause form (Al-Ani and May 1978; Broselow 1992). Thus, when the inflectional ending /-un/ is added to /naas/ ‘people’ or /faʾs/ ‘axe’, their syllabic configurations change into those of Cvv /naa-sun/ and CvC /faʾ-sun/, respectively.

Types (1) and (2) are called *open syllables*, because they do not have a coda but end in the nucleus, as in the prepositional prefix /li/ ‘for, to’ and /maa/ ‘what’, respectively. Types (3), (4), (5), and (6) that have a coda and therefore end in a consonant are called *closed syllables*, as in /man/ ‘who’, /fiil/ ‘elephant’, /waqt/ ‘time’, and /daar/ ‘house’, respectively. An open syllable of the type Cv mostly occurs in polysyllabic words like /qa-ra-ʾa/ ‘he read’. Moreover, a syllable that is open and ends in a short vowel is called a *light syllable*. If the syllable is open but the vowel in its nucleus is long, i.e. Cvv, it is called a *heavy syllable*.

There is yet another type that occurs in Arabic, although its distribution, as pointed out earlier, is restricted to final position; this syllable, called *superheavy* by McCarthy (in Broselow 1992), contains either a long vowel followed by a consonant CvvC or a short vowel followed by two consonants CvCC.

Type (6), according to Al-Ani and May (1978), occurs only as the final syllable of utterances or of words in pause form. In this respect, the two identical consonants in /maarr/ are called a geminate consonant, where prolongation of the continuants or a longer closure of the stops is involved (Al-Ani 1970). However, when the inflectional ending /-un/ is added to /maarr/, its syllabic configuration changes and becomes of the type CvvC: /maar-run/.

4. Phonotactic constraints

It is clear from the above illustration that Arabic does not have a flexible syllable structure in comparison with other languages like English or German, where the nucleus can be preceded or followed by several consonants. Nevertheless, there are some constraints on which phoneme sequences are permissible in Arabic syllables. Such constraints are called *phonotactic constraints*, and these constraints determine all possible sound sequences within a particular language. They are language specific as they vary from one language to another. This means that two or more languages with similar phoneme inventories like English and German may have different rules governing the distribution of phonemes in morphemes, words, or syllables (Lass 1984).

In this respect, we find that Arabic does not allow empty onsets or words to start with more than one consonant in the onset position. This is connected with the fact that the syllable in Arabic is of the type (Cv-), and never of the (*vC-) or (*CC) structures. Conversely, the coda, as mentioned earlier, can be empty or have one or two consonants. There are, however, certain restrictions on the sequence of the two consonants that may appear in the coda position. For example, Arabic disallows the sequence or combination of /ḥ/ and /ġ/ because /baḥġ/ is not a possible word. Some more combinatory constraints in Arabic are /*bm/, /*fb/, /*rl/, /*kq/, /*xġ/, etc.

Bibliographical references

Al-Ani, Salman. 1970. *Arabic phonology: An acoustical and physiological investigation*. The Hague: Mouton.

—— and Darlene R. May. 1978. "The phonological structure of the syllable in Arabic". *Readings in Arabic linguistics*, ed. Salman Al-Ani, 113–126. Bloomington: Indiana University Linguistics Club.

Broselow, Ellen. 1992. "Parametric variation in Arabic dialect phonology". *Perspectives on Arabic linguistics*, IV, ed. Ellen Broselow, Mushira Eid, and John McCarthy, 7–45. Amsterdam and Philadelphia: J. Benjamins.

Cox, Felicity, Jonathan Harrington, and Robert Mannell. 2004. "The syllable and the foot". www.ling.mq.edu.au/units/ling 210–901/phonology/syllable/. (Accessed 30 March 2004.)

Kiraz, George A. and Bernd Möbius. 1998. "Multilingual syllabification using weighted finite-state transducers". www.ims.uni-stuttgart.de/~moebius/papers/synt98_R07.pdf. (Accessed 20 April 2004.)

Lass, Roger. 1984. *Phonology: An introduction to basic concepts*. Cambridge: Cambridge University Press.

Laver, John. 1994. *Principles of phonetics*. Cambridge: Cambridge University Press.

Roach, Peter. 1991. *English phonetics and phonology: A practical course*. 2nd ed. Cambridge: Cambridge University Press.

Maher Jesry (King Saud University)

Syncope

Syncope involves the loss of a vowel. In Arabic dialects, syncope is typically driven by metrical constraints. The most common targets of syncope are high vowels and central vowels, which are relatively low in sonority, but there are dialects in which even the highly sonorous [a] syncopates. Two fairly representative dialects are Bedouin Ḥijāzī Arabic and Cairene Arabic.

1. Bedouin Ḥijāzī Arabic

Bedouin Ḥijāzī Arabic is remarkable among Arabic dialects in that the most sonorous vowel, [a], is the target of syncope (→ *gahawa*-syndrome). Examples in (1) are from Al-Mozainy a.o. (1985:136). Syncopated vowels are underlined.

The result of syncope in each case is what might be called a degenerate syllable, i.e. a syllable in which the nucleus dominates no segment. Empty nuclei are represented here as [·], e.g. [.n·.xá.lah.].

Following McCarthy (2003), it may be assumed that Bedouin Ḥijāzī has iambic feet, except when this would result in final stress on a nonsuperheavy syllable. Iambic languages show a strong preference for uneven feet, i.e. those in which the strong syllable is heavier than the weak syllable (Prince 1990; Hayes 1995). The greater the difference in duration and/or sonority, the better the iamb; call this iambic harmony. The drive to improve iambic harmony outweighs syllable structure constraints, and thus syncope occurs. What accounts for the fact that [a], and not a less sonorous vowel, is targeted, is that more sonorous vowels are dispreferred as the heads of weak syllables, as in (2).

In (2a), the grammatical form has a high iambic harmony rating, while the ungrammatical form without syncope has a low rating because the strong and weak syllables are equally heavy and sonorous. In (2b), however, the strong syllable is already heavier than the weak, and thus there is no need for syncope to apply.

2. CAIRENE ARABIC

Cairene Arabic is a prototypical moraic trochee language, in which all feet are composed of either a single heavy syllable (H) or two light syllables (LL). According to the predominant analysis (McCarthy 1979; Hayes 1995), the Cairene stress pattern involves exhaustive construction of moraic trochees from the left word edge, as in (3). So, while a form may have only one stress, it may nonetheless have multiple feet. The rightmost foot gets primary stress, while all other feet bear no stress.

Foot construction interacts with a process of high vowel syncope. It involves deletion of short /i/ and /u/ in open unstressed syllables, and subsequent → resyllabification (Broselow 1976; Welden 1977; Kenstowicz 1980). This occurs at both the word level (4a) and the phrasal level (4b).

Syncope can apply not only to vowels that would normally have been unstressed but also to vowels that would have been stressed by the normal stress assignment algorithm. In this way, stress shifts onto the preceding vowel, which becomes the head of a heavy syllable, as in (5).

A heavy syllable makes a better foot-head than does a light syllable, and a much better foot-head than a light syllable with a low-sonority nucleus. Therefore, the ostensible reason for syncope in these cases is improvement of metrical structure.

But not all open syllables are subject to high vowel syncope. Feet constructed in the word-level phonology generally resist syncope at the phrasal level, as demonstrated by Kenstowicz (1980) in (6).

(1)	a.	*gʸálaˁ*	‘castles’	/gʸa̱laˁah/	> *gʸláˁah*	‘a castle’
	b.	*sáḥab*	‘he pulled’	/sa̱ḥabat/	> *sḥábat*	‘she pulled’
	c.	*saḥábna*	‘we pulled’	/sa̱ḥabaw/	> *sḥábaw*	‘they [masc.] pulled’
	d.	*náxal*	‘palm trees’	/na̱xalah/	> *nxálah*	‘a palm tree’

(2)	a.	/naxalah/	> .(n·.xá).lah.,	*.(na.xá).lah.	‘a palm tree’
	b.	/ḍarabtukum/	> .(ḍa.ráb).tu.kum.,	*.(ḍ·.ráb).tu.kum.	‘I hit you [pl.]’

(3)	*xára*	(LL)	‘shit’
	ˁárabi	(LL)L	‘Arabic/Arab’
	ríḥla	(H)L	‘trip’
	madrása	(H)(LL)	‘school’
	mudarrísa	L(H)(LL)	‘teacher [fem.]’

(4)	a.	/fihi̱m-it/	.fíh.mit.	‘she understood’
		/bi-ni̱-fuut/	.bin.fúut.	‘we are passing’
	b.	/huwwa mu̱darris/	.hùw.wam.dár.ris.	‘he is a teacher’
		/gineeni̱t irraagil/	.gi.nìn.tir.ráa.gil.	‘the man’s garden’

(5)	/kanak-i̱t-i/	.ka.(nák).ti.	*.(ka.na).(kí.ti).	‘my coffeepot’
	/kanab-i̱t-u/	.ka.(náb).tu.	*.(ka.na).(bí.tu).	‘his couch’
	/faahi̱m-a/	.(fáh).ma.	*.(faa).(hí.ma).	‘understanding [fem. act. part.]’

(6)	a.	*hùma ʾrádtu*	**hùma ʾirádtu*	‘they are his monkeys’
	b.	*hùma ʾírada*	**hùma ʾráda*	‘they are monkeys’
	c.	*hùma ʾiradítna*	**hùma ʾradítna*	‘they are our monkeys’

In (6a), syncope applies at the phrasal level, because the initial syllable is not footed at the word level: [.i.(rád).tu.]. However, the initial syllable is footed in (6b) and (6c): [.(í.ra).da.] and [.(i.ra).(dít).na.]; hence, they resist syncope at the phrasal level.

The same difference in behavior between footed and unfooted high vowels applies to word-final syllables. Phrasal syncope does not affect footed final syllables, as in (7).

(7) *širib ilʾáhwa* **širb ilʾáhwa*
'he drank the coffee'

Phrasal resyllabification of [b] leaves the syllable [.ri.], which is open and in the weak position of a foot, yet it does not undergo syncope because it has already been footed in the word-level phonology.

Bibliographical references

Al-Mozainy, Hamza Qublan, Robert Bley-Vroman, and John McCarthy. 1985. "Stress shift and metrical structure". *Linguistic Inquiry* 16.135–144.
Broselow, Ellen. 1976. *The phonology of Egyptian Arabic*. Ph.D. diss., University of Massachusetts, Amherst.
Hayes, Bruce. 1995. *Metrical stress theory: Principles and case studies*. Chicago: University of Chicago Press.
Kenstowicz, Michael. 1980. "Notes on Cairene Arabic syncope". *Studies in the Linguistic Sciences* 10.39–53.
McCarthy, John. 1979. "On stress and syllabification". *Linguistic Inquiry* 10.443–465.
——. 2003. "Sympathy, cumulativity, and the Duke-of-York gambit". *The syllable in optimality theory*, ed. Caroline Féry and Ruben van de Vijver, 23–76. Cambridge: Cambridge University Press.
Prince, Alan. 1990. "Quantitative consequences of rhythmic organization". *Parasession on the Syllable in Phonetics and Phonology*, ed. Karen Deaton, Manuela Noske, and Michael Ziolowski, 355–398. Chicago: Chicago Linguistics Society.
Welden, Ann. 1977. *Prosodic aspects of Cairo Arabic phonology*. Ph.D. diss., University of Texas, Austin.

David Teeple
(University of California, Santa Cruz)

Synonym → Mutarādif

Syntax

1. Introduction

Syntax is the study of phrasal and sentential patterns of natural language. It is the engine that combines the sound/gesture and meaning components of language. Syntax deals primarily with how words combine to form phrases and sentences, and the dependencies that obtain between the constituents of the phrase or sentence. Such dependencies include agreement, Case, anaphoric relations, filler gap/pronominal relations, and thematic relations, among many others. Another area that has occupied a prominent position within syntactic debates concerns word order alternations, particularly as they correlate with specific semantic and discourse interpretations, such as question (→ WH-movement), relatives (→ relative clause), → topicalization, and → focus constructions, to mention just the most prominent ones.

Within the generative paradigm and particularly the Principles and Parameters framework (Chomsky 1981, 1995, 2001), the above specific syntactic issues have been approached with two main goals in mind. The first goal is to explore syntactic principles and properties that have crosslinguistic manifestation and validity and that may help determine the nature of the syntactic dimension of the, possibly innate, human linguistic faculty, i.e. Universal Grammar. The second goal, intimately related to the first, is to provide in-depth descriptions, analyses, and comparisons of different languages, so that one can determine the extent and limit of language variation and how that variation relates to the basic Universal Grammar core.

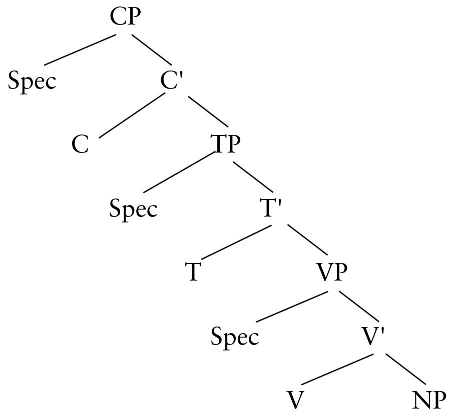
Generative approaches of Arabic, both formal (Classical and Modern Standard) and colloquial, have dealt with a number of topics that relate to the issues mentioned above. The most prominent issues that have received a great deal of attention in the last four years include lexical and functional categories, clause structure and Verb Subject (VS) order (→ word order), → agreement, particularly the correlation between the richness of agreement and the position of the subject relative to the verb, verbless sentences and copular constructions (→ copula), → negation, questions, relatives, the status of the 'subject' in the Subject Verb (SV) order, Case, Tense/Aspect, and the structure of the noun phrase, with the so-called → Construct State getting more scrutiny.

One key assumption of the Principles and Parameters framework is that the syntactic configuration and the structural relations between the members of the configuration play a critical role in accounting for syntactic generalizations such as those governing agreement, Case, licensing of polarity items, binding relations, and filler/gap dependencies. Every syntactic constituent must have a head, which may be combined with a complement or two, as in transitive or ditransitive verbs, a specifier, such as the subject, and an adjunct, such as an adverb, a phrase, or a clause that functions as modifier.

Clauses usually consist of functional and lexical projections. Lexical projections are headed by lexical categories, such as nouns and verbs, while functional categories are headed by grammatical categories, such as complementizers and tense.

The dependencies between functional and lexical categories are argued to be key to accounting for the distribution of the latter. For example, the dependency between the verb and tense may explain the displacement of the verb from its base position. Similarly, tense seems to play a role in determining the nominative Case of the subject, argued to be generated in the specifier of VP, where it receives its thematic role. Relative pronouns and question operators target the CP domain and bind a trace or a pronoun in a position within the same clause or a lower clause. Some dependency relations are the result of movement, as in questions with the gap strategy, while others involve relations between filled position (for example an anaphor and its antecedent). The relations between elements occupying different positions within a syntactic configuration such as (1) are subject to a limited number of principles and restrictions. For example, the verbal head cannot cross over the tense head to move to C, otherwise, a violation of minimality would result. The minimality restriction essentially entails that some categories (Heads, NPs, and WH-phrases) may block dependencies between categories of the same type. For example, a head that has moved to a higher position in the configuration cannot establish a dependency with its original position across an intervening head.

(1)



There is an ongoing debate about the features on the functional categories in (1) that drive syntactic operations such as movement, agreement, and Case assignment. Likewise, it is still a matter of debate how many functional categories there are and how they are ordered relative to each other. For a basic introduction to the Principles and Parameters framework, the reader is referred to Haegeman (1991), Adger (2003), and Carnie (2007).

The discussion below is focused on a number of constructions and aspects of Arabic phrasal and sentential syntax that have been subjected to analyses within the Principles and Parameter framework.

2. Sentence structure of Arabic

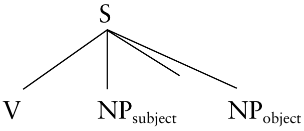
Standard Arabic is considered a verb-initial language on a par with Irish and other Celtic languages. The basic order is claimed to be VSO, as illustrated in (2):

(2a) *kassara* *l-walad-u*
broke.3ms the-boy-Nom
l-ka'š-a
the-glass-Acc
'The boy broke the glass'
(2b) *kāna* *l-walad-u* *fī*
was.3ms the-boy-Nom in
l-bayt-i
the-house-Gen
'The boy was in the house'

From the linear order in (2), it is clear that the verb and the object do not constitute a VP constituent that excludes the subject. The subject intervenes between the verb and the object, which indicates that the three elements, the verb, the subject, and the object, do not form a VP. In this respect, Arabic is different from English, where the subject precedes the verb and the verb clearly combines with the object to form a VP constituent. The question then is whether there is a VP constituent in Arabic, and whether there is a fundamental difference across languages in terms of VP constituency whereby the verb and the object form a constituent headed by the verb to capture the semantic and Case dependencies between them.

Earlier accounts of Arabic provided a flat structure for VSO sentences, as in (3).

(3)



The assumption behind this analysis is that the VP constituent, or the constituent that combines the predicate and its complement, is not universal, i.e., languages may vary in whether or not they display a VP constituent.

Mohammad (1999a) points out that structures such as (3) imply that the subject and object are in a symmetric relation since they are immediately dominated by the same mother node (S). Binding facts, however, show that the subject is syntactically more prominent than the object, with the subject always functioning as the antecedent, or binder, of the object, as illustrated in (4).

(4a) *lāma* *ṭ-ṭālib-u*
blamed.3ms the-student-Nom
nafs-a-hu
self-Acc-his
'The student blamed himself'
(4b) **lāmat* *nafs-u-hu*
blamed.3fs self-Nom-his
ṭ-ṭālib-a
the-student-Acc

(4a) is grammatical with subject binding the reflexive object; (4b) is ungrammatical because the reflexive, which needs to be bound, functions as subject and its putative binder as object (→ binding). According to the Principles and Parameters framework, the binder must occupy a configurationally more prominent position in the sentence than the element it binds. This is clearly not the case in (3), where configurationally both the subject and the object are on a par (i.e., neither c-commands the other).

Binding is not the only problem with the structure in (3). Mohammad also points out that, as in other languages, the verb and the object in Arabic can combine to give an idiomatic meaning, as in (5).

(5) *ʾeḥmad* *ḍayyaʿ* *ʿagl-u*
Ahmed lost.3ms mind-his
(Palestinian Arabic)
'Ahmad went crazy'

Such idioms are taken to argue for a VP constituent grouping the object and the verb, an option not available under the flat representation in (3).

There are additional arguments against the structure in (3). In the context of an auxiliary verb such as *kāna* 'to be', the verb and the object combine to form a constituent VP with the subject preceding the Verb-Object sequence. The basic order in an Arabic sentence with two such verbs is illustrated by (6).

(6) *kāna* *l-walad-u* *yaqraʾu*
was.3ms the-boy-Nom read.3ms
l-kitāb-a
the-book-Acc
'The boy was reading the book'

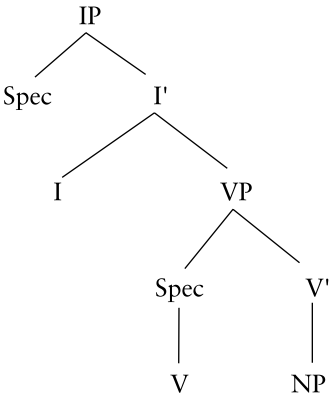
The first verb, the auxiliary (*kāna*), is followed by the subject. The second verb, the main thematic predicate (*yaqraʾu*), follows the subject and is followed by the object. → Sentence coordination provides strong evidence that in sentences such as (6), the main verb and the object form a VP constituent. The two elements can be conjoined in a constituent that excludes the subject, as (7) from Moroccan Arabic clearly illustrates:

(7) *kan* *l-wəld* *ta-y-qra*
was.3ms the-boy Asp-3ms-read
lə-ktab *w* *ya-kul* *lə-fṭur*
the-book and 3ms-eat the-breakfast
'The boy was reading the book and eating breakfast'

A similar situation exists in other Arabic dialects and in Standard Arabic.

These and other facts led Mohammad (1989) and Fassi Fehri (1993) to argue for a different representation of the Arabic structure. They proposed a representation like (8), already available in the literature for other languages. The representation in (8) provides two positions for the subject, a lower position within the lexical layer of the sentence headed by verb (the VP) and a second position within the functional category that carries tense (and possibly agreement).

(8)



Although the labels may differ and arguments may be made for additional projections in (8), there is a consensus within Principles and Parameters approaches to Arabic that it is basically correct. There is also widespread agreement within the theory that universally the subject starts within the VP (possibly in the Specifier position). Languages then have a number of options. In English, the subject raises from the Spec of VP to the Spec of IP, while the verb remains in the VP. This yields the basic SVO order that characterizes English. In a language such as French, the subject is generated in the VP and is raised to the Spec of IP. The verb also raises to adjoin to the head of IP (I), yielding the SVO order of French (Pollock 1989). With regard to Arabic, most analyses agree that the verb moves at least to I, while the subject remains lower within the VP, or at least a projection lower than the projection hosting the verb (Aoun, Benmamoun, and Sportiche 1994). This is the analysis of Mohammad (1989, 1999a) and Fassi Fehri (1993). Movement of the verb beyond the subject results in the VSO order that distinguishes Arabic from English. If we assume that the auxiliary verb is in I in (8), or starts in a lower functional position and moves to I, there is no need for the main verb to move to I, which in turns implies that the main verb would follow the subject, which is exactly what we find in (6) and (7).

In an analysis of Arabic based on the presentation in (8), combined with movement of the verb, the universality of the VP constituent can be maintained. The difference between English and Arabic is then attributed to options in verb movement and subject movement rather than to basic clause structure. In other words, the difference is attributed to choices available in the application of a limited set of universal principles. Furthermore, the structure in (8) allows for a principled treatment of idioms and coordinations involving verb and object and of other syntactic relations, such as binding relations, which require the subject to be more configurationally prominent than the object. In (8), the subject is in a higher position than the object.

The representation in (8) has two subject positions. In English, they are filled by the same NP subject, which moves from Spec of VP to Spec of IP. In principle, one would expect to find languages where the two positions are filled by different elements. Mohammad (1999a) argues that this is exactly the situation with the so-called expletive subjects found in sentences such as (9).

(9) *qultu* *ʾinna-hu* *waṣala*
said.1s that-it arrived
l-ʾawlād-u
the-children-Nom
'I said that the boys arrived'

In (9), the clitic pronoun on the complementizer presumably originates in the Spec of IP, while the thematic subject remains in the Spec of VP. The verb is located between two NPs, as predicted by the representation in (8), which allows for two subject positions. The expletive subject can occur in Spec of IP because the position is not thematic. On the other hand, the Spec of VP can only be occupied by an element that can bear a semantic role. Thus, the expletive subject will always be in a higher position than the thematic subject, which Mohammad shows to be the case in Arabic.

The preverbal position can be occupied by the NP that functions as subject. It is debatable whether that position should be treated on a par with preverbal subject positions in English and French, for example as in (10).

(10) *al-walad-u* *kassara*
the-boy-Nom broke.3ms
l-kaʾs-a
the-glass-Acc
'The boy broke the glass'

If the preverbal NP is treated as being in a subject position (A-position), it should be considered a regular subject. But if the subject in Arabic is restricted to the postverbal position, the preverbal position cannot be considered a subject position. It may, however, be a discourse position associated with focus and topicalization constructions (A'-positions). Both analyses have been advanced for Arabic, although the idea that the preverbal NP in sentences such as (10) is not an argument position has had a long history within generative approaches to Arabic and approaches inspired by insights from the Arabic linguistic tradition, the latter having dealt with this topic more than ten centuries ago. For proponents of the theory that the subject is not restricted to the postverbal position, the analysis of (10) involves movement of the NP *al-walad-u* from the lower position within the VP to the preverbal position, leaving a trace behind. In this respect, the SVO sentences receive an analysis similar to SVO sentences in English and French. For proponents of the theory that the subject is restricted to the postverbal position, the NP *al-walad-u* is in a topic or left-dislocated position, while the real subject is a null pronominal (→ pro drop) or an incorporated pronoun. The debate continues in view of evidence pulling in both directions (Ayoub 1981; Benmamoun 1992; Fassi Fehri 1993; van Gelderen 1996; Mohammad 1999a; Soltan 2007; Doron and Heycock 1999).

3. Verbless sentences

Another topic that has figured prominently in analyses of the syntax of Arabic concerns the so-called verbless sentences. These are well-known constructions found in Arabic and other languages, such as Hebrew and Russian, where no verbal copula is present in sentences with present tense interpretation, as in (11).

(11a) *al-walad-u* *fī* *l-bayt-i*
the-boy-Nom in the-house-Gen
'The boy is in the house'
(11b) *al-walad-u* *nāʾim-un*
the-boy-Nom sleeping-Nom
'The boy is sleeping'
(11c) *al-bayt-u* *kabīr-un*
the-house-Nom big-Nom
'The house is big'
(11d) *zayd-un* *muᶜallim-un*
Zayd-Nom teacher-Nom
'Zayd is a teacher'

Three competing analyses have been proposed for Arabic verbless constructions (Bakir 1980; Ayoub 1981; Jelinek 1981; Eid 1983, 1991; Mouchaweh 1986; Heggie 1988; Bahloul 1994; Shlonsky 1997; Benmamoun 2000, 2008). The first approach argues that they are small clauses on a par with the embedded small clauses found in English constructions such as *I left the door open* (Mouchaweh 1986). The main idea behind the small clause approach is that verbless constructions lack functional categories such as tense and they lack a VP projection. The second approach, which predates the first one, argues that the so-called verbless sentences are actually full clauses with a null or deleted copular verb that heads a VP (Bakir 1980; Fassi Fehri 1993). In this respect, the sentences in (11) are not syntactically different from their counterparts that contain the verb *kāna* in the past (12a) and future tenses (12b).

(12a) *kāna* *l-bayt-u* *kabīr-an*
was.3ms the-house-Nom big-Acc
'The house was big'
(12b) *sa-yakūnu* *l-bayt-u* *kabīr-an*
Fut-be.3ms the-house-Nom big-Acc
'The house will be big'

Under the second (full clause) approach, copular constructions in all tenses have identical syntactic representations, exactly like their English counterparts.

The third approach takes a middle position between the two (Jelinek 1981; Benmamoun 2000). It agrees that verbless sentences are not full clauses, because they do not contain a null copula and a null VP, but it maintains that they have a functional category Tense that contains present tense features. The evidence against a null copular verb comes from the fact that in sentences with a copular verb, like (12), the predicate is in accusative Case, but in verbless sentences it gets nominative Case, probably as default Case in this context. This is taken to suggest that a copular verb is absent. The evidence for a Tense projection in verbless sentences comes from expletive subjects: they can occur in these constructions, as in (13), and the subject gets nominative Case, which is assumed to be assigned/checked by Tense within the Principles and Parameters framework.

(13)	*hunāka*	*walad-un*	*fī*	*l-bayt-i*
	there	boy-Nom	in	the-house-Gen
	'There is a boy in the room'			

Expletives are not thematic elements and, therefore, cannot occur in the Specifier of a lexical, thematic role assigning, category. The standard assumption within the Principles and Parameters framework is that expletives are inserted to satisfy properties of functional categories such as Tense. But the nature of the relevant properties is subject to debate. One of the most discussed alternatives argues that sentences need subjects, and subjects must occur in the specifier of the tense projection; another argues that functional projections such as tense require that an overt element occupies their specifier position (Chomsky 2001). The overall generalization, however, is that verbless sentences contain more structure than small clauses but are not as big as sentences with verbal copulas. For detailed discussion of the various proposals and alternatives, see Benmamoun (2000, 2008) and references cited therein.

The topic of verbless sentences in Arabic brings into sharp focus the issue of language universals and language variation. The full clause analysis proposals are consistent with the theory that clause structure is universal, which in turns entails positing highly abstract representations to make languages adhere to the universal schema. The small clause analysis and the third (middle) alternative assume that languages may differ with regard to their clause structures depending on their own specific properties. If a language does not need a verbal copula in the present tense, it will not display a VP in such a construction. On the other hand, in constructions such as (13), where a verbal copula is needed, a VP constituent will be projected. Fortunately, the issue is empirical in the sense that evidence can be found from within the language to argue for a particular approach, with obvious consequences for our view on universal aspects of language and the nature of language variation. Research on verbless sentences in Arabic demonstrates clearly how systematic comparative syntactic studies exploring the workings of different constructions in unrelated language, as advocated and promoted by the Principles and Parameters framework, can shed light on important theoretical debates.

4. Subject/Verb agreement

Subject/Verb agreement in Arabic has also received a lot of attention within the Principles and Parameters framework (Ayoub 1981; Mohammad 1989; Benmamoun 1992; Fassi Fehri 1988, 2003; van Gelderen 1996; Bahloul and Harbert 1993; Benmamoun 2000; Benmamoun and Lorimor 2006; Soltan 2007). This is not surprising, because of the well-known agreement asymmetry in Standard Arabic, illustrated in (14).

(14a)	*daxala*	*l-muʿallim-ūna*
	entered .3ms	the-teachers-Nom.mp
	'The teachers came in'	
(14b)	*al-muʿallim-ūna*	*daxal-ū*
	the-teachers-Nom.mp	entered.3mp
	'The teachers came in'	

One popular account for the agreement asymmetry in (14) assumes that in (14a) there is a genuine gender and person agreement relation between the verb and the subject. In (14b), however, full agreement on the verb is a realization of the pronominal subject that has been incorporated into the verb and that is resumed by the left-dislocated or topicalized preverbal 'subject'. Another alternative account assumes that there are agreement relations between the subject and the verb in both sentences in (14), but that number fails to be morphologically realized in (14a), either because number agreement takes place at a later abstract point in the derivation and therefore does not feed the morphological component, or because the verb and the postverbal subject merge into a prosodic unit, which obviates the need to spell out number through a number affix since the merged subject can fulfill that role.

Both analyses, however, take it as given that the syntactic configuration is critical in accounting for the agreement asymmetry. Partial agreement takes place only when the verb c-commands (and is higher than) the subject; full agreement takes place when the subject is in a Spec-head relation with the verb. It is debatable within the Principles and Parameters framework whether both c-command (so-called Agree) and Spec-head configurations are needed to account for agreement in general and agreement asymmetries in particular. Again, the facts in Arabic provide testing grounds for the debate.

Another agreement asymmetry that has sparked debate within the Principles and Parameters framework concerns agreement with conjoined subjects (Aoun, Benmamoun, and Sportiche 1994, 1999; Munn 1999; Harbert and Bahloul 2003; Lorimor 2007; Soltan 2007). The examples in (15) from Moroccan Arabic illustrate the phenomenon.

(15a) *ja* *omar* *w* *karim*
came.3ms Omar and Karim
'Omar and Karim came'
(15b) *jaw* *omar* *w* *karim*
came.3p Omar and Karim
'Omar and Karim came'
(15c) *omar* *w* *karim* *jaw*
Omar and Karim came.3p
'Omar and Karim came'
(15d) **omar* *w* *karim* *ja*
Omar and Karim came.3ms

The verb in Moroccan Arabic can agree with the first conjunct to its right (15a) or with both conjuncts to its right (15b). However, it must agree with both conjuncts to its left (15c) and cannot agree with only one of the conjuncts on its left. Although this is another agreement asymmetry sensitive to word order, it is different from the agreement asymmetry in Standard Arabic discussed earlier. Unlike Standard Arabic, in Moroccan and other Arabic dialects, the verb fully agrees with the subject regardless of its linear position in relation to the subject, as illustrated in (16).

(16a) *wəqfu* *lə-wlad*
stood.3p the-children
'The children stood up'
(16b) *lə-wlad* *wəqfu*
the-children stood.3p
'The children stood up'

To deal with agreement asymmetry in the context of coordination, Aoun, Benmamoun, and Sportiche (1994) advanced a clausal/gapping analysis of coordination. The idea is that each conjunct is the subject of its own clause. The second conjunct is the subject of a clause whose verb is gapped and is recovered under identity with the first verb. The main argument given for the clausal/gapping analysis comes from number-sensitive items in Arabic, which require a plural subject. The prediction is that such elements should not be compatible with close single conjunct agreement. This prediction seems to be correct, as illustrated in (17) from Moroccan Arabic.

(17a) **gləs* *omar* *w* *karim*
sat.3ms Omar and Karim
ḥda *bəˁḍhum*
near each.other
(17b) *gəlsu* *omar* *w* *karim*
sat.3p Omar and Karim
ḥda *bəˁḍhum*
near each.other
'Omar and Karim sat near each other'

The → reciprocal in (17) requires a plural antecedent, which is not available in a clausal analysis of (17a). Other number-sensitive items and plural predicates provide further support for the analysis.

Munn (1999) and Soltan (2007) provide alternative analyses that do not posit clausal coordination or gapping. For them, close conjunct agreement obtains because only the first conjunct is accessible to the verb under specific assumptions about the structure and workings of coordination. Lorimor (2007) provides the first experimental study of close conjunct agreement in Lebanese Arabic, which proves the prevalence of close conjunct agreement in the VS order. The debate continues and includes data from a variety of languages that display the same phenomenon. Close conjunct agreement brings into the picture syntax, linear order, and semantics, and it is yet to be determined what role, if any, each one plays.

5. The syntax of heads: Tense, verbs, and negation

The work of Pollock (1989), among many others, has demonstrated that various heads of projections interact with each other, which may explain some aspects of the syntax of natural language. The heads that have received close scrutiny, particularly in Arabic, are tense, negation, and verbs. The most prominent question in relation to tense is whether all tenses, including the present tense, correspond to a syntactic projection and whether they all attract the verb. In the case of negation, the main issue is whether the sentential negative heads its own syntactic projection and whether this projection is located between Tense and Verb or in a position higher than both. Finally, the question with respect to the verb is whether it undergoes movement all the way to the highest functional projection, or whether the movement is more restricted.

Starting with tense, most syntactic studies of Arabic within the Principles and Parameters framework have argued that past, present, and future project their own tense projections. The labels may vary, but the main idea is that the time reference of the sentence in Arabic corresponds to a functional projection. It is controversial, however, whether the morphology and vocalic melodies that the verb displays in the present and past tense are realizations of those tenses. Benmamoun (2000) provides arguments that both the past and present tense are not realized overtly on the verb in Arabic. This is partly due to the fact that the vocalic melodies are not uniform for all verbs, which would be surprising for a temporal morpheme. Moreover, some dialects, such as Moroccan Arabic, do not have any vocalic melodies, and yet the verbs have the same distributions (perfective in the past tense and imperfective in the present tense). The agreement morphology on the verb also does not seem to carry tense information because the same morphology is found on the negative *laysa*, which occurs in the present tense but whose agreement morphology is the same as the one carried by the past tense verb. In short, both the present tense and past tense seem to be abstract and null projections in Arabic. Their syntax, though, may be different, as their interaction with negation and the verb indicates.

The topic of the syntax of sentential negation in Arabic, both Standard and Colloquial, has garnered extensive attention within the Principles and Parameters framework (Moutaouakil 1987; Benmamoun 1992, 2000; Halila 1992; Fassi Fehri 1993; Bahloul 1994; Ouhalla 1993; Al-Tamari 2001).

Standard Arabic negatives are intriguing in that they co-vary with tense. Thus, *lā* occurs in present tense sentence, *lam* in past tense sentences, and *lan* in future tense sentences (18).

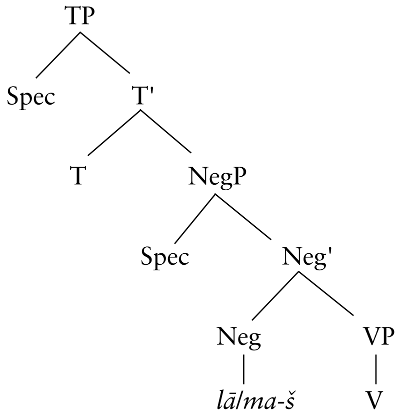
(18a) *aṭ-ṭullāb-u* ***lā***
the-students-Nom Neg
ya-drus-ū-n
3m-study-mp-Ind
'The students do not study'
(18b) *aṭ-ṭullāb-u* ***lam***
the-students-Nom Neg.past
ya-drus-ū
3m-study-mp
'The students did not study'
(18c) *aṭ-ṭullāb-u* ***lan***
the-students-Nom Neg.fut
ya-drus-ū
3m-study-mp
'The students will not study'

In this respect, Standard Arabic differs not only from many other languages where the verb carries tense information regardless of the polarity of the sentence, but also from the colloquial Arabic dialects where the same negative is used in affirmative and negative sentences and in all the tenses, as shown in (19) from Moroccan Arabic.

(19a) *ma-ta-yqra-š*
Neg-Asp-read.3ms-Neg
'He is not reading'
(19b) *ma-qra-š*
Neg-read.past.3ms-Neg
'He didn't read'
(19c) *ma-ġadi-yqra-š*
Neg-Fut-read.3ms-Neg
'He will not read'

A uniform analysis of both Moroccan and Standard Arabic is provided in Benmamoun (1992, 2000), where it is argued that the negative projection is located between the tense projection, TP, and the verbal projection, as shown in (20).

(20)



In Moroccan Arabic, the verb merges with negation and then with tense. In Standard Arabic, by contrast, the negative *lā* is a possible host of tense and therefore obviates the need for the verb to move to the latter. The principle within the Principles and Parameter framework at work is 'Minimality', which, despite its many reincarnations, essentially requires that the closest head that can move to tense should do so. In Standard Arabic, *lā* can move to tense, which can be spelled out as *lam* or *lan*, depending on the value of the tense head, as past or present.

Ouhalla (1993) proposes a similar representation for the Standard Arabic negative *lā* but suggests that the negative *mā* is located in a projection higher than tense; in other contexts where it negates and focuses the element it attaches to, it is generated on its host and the whole complex is moved to the focus projection located above TP. The projection that hosts *mā* is associated with the contrastive focus interpretation; hence, it is at a higher location, which in turn explains its lack of interaction with verb movement.

Returning to verb movement to tense, the distribution of sentential negation provides some clues as to whether the verb needs to move to tense in both present and past tenses. In Egyptian Arabic, it is possible for the present tense verb not to merge with negation, while the past tense verb must always do so, as the examples (21) and (22) from Egyptian Arabic show (Jelinek 1981).

(21a) *ma-bi-yiktib-š*
Neg-Asp-writes.3ms-Neg
'He doesn't write'
(21b) *mi-š* *bi-yiktib*
Neg-Neg Asp-writes.3ms
'He isn't writing'

(22) *ᶜumar* *ma-katab-š* *ig-gawāb*
Omar Neg-wrote.3ms-Neg the-letter
'Omar didn't write the letter'

The syntactic account for contrast in the syntax of verbs in present and past tense sentences is that the former do not need to force verb raising to tense, while the latter do. In order for the verb to raise to tense in the past tense, it must merge with negation to avoid a minimality violation. The contrast between the two sentences is also reflected in the syntax of what Ferguson (1983) refers to as God wishes (Benmamoun 2000). There is a tendency in these expressions for past tense sentences to have the verb precede the subject and to have the opposite order in present tense sentences, as in (23).

(23a) *raḥm-u* *llah*
blessed.3ms-him God
'May God bless him!'
(23b) *llah* *y-rəḥm-u*
God 3ms-bless-him
'May God bless him!'

Again, the contrast would follow from the necessity of verb movement in the context of the past tense and its optionality in the context of the present tense.

6. CONCLUSION

The essential assumption of all the accounts of the syntax of Arabic described above is that the configurational properties of the sentence as articulated by the theory of Principles and Parameters, coupled with the rules and constraints of the theory, are argued to be key to understanding the workings of the Arabic sentence, particularly its clause structure, the placement of its verb and subject, the interaction between the verb and the subject, and the asymmetries that arise due to the specificities of word order patterns in Arabic. The picture that emerges is that the syntax of Arabic is not as radically different from the syntax of other languages. Arabic deploys the same principles of combining words into phrases and phrases into sentences. The same universal principles govern the distribution of heads and phrases in the sentence. Any variation has to do with the specific properties of grammatical and lexical categories which may have far-reaching consequences. For example, the property of the present tense that prevents it from requiring a verb explains the VSO order, the ordering pattern in idiomatic constructions, and the lack of a copula in the present tense, among other properties.

The same line of analysis has been applied to a number of constructions, such as null pronouns (Fassi Fehri 1998, 2003; Kenstowicz 1989), question formation (Fassi Fehri 1982; Wahba 1984, 1991; Aoun and Choueiri 1999; Shlonsky 2002), relatives (Ouhalla 1996; Choueiri 2002; Aoun and Li 2003), focus (Ouhalla 1994), construct state (Mohammad 1988, 1999b; Benmamoun 2000), verbal nouns (Fassi Fehri 2003; Hazout 1990), quantifier floating (Benmamoun 1999), negative polarity licensing (Benmamoun 1996, 1997), clitic left-dislocation, bound pronouns, resumption, and topicalization (Aoun and Benmamoun 1998; Doron and Heycock 1999; Aoun and Choueiri 2000; Aoun, Choueiri, and Hornstein 2001), indefinite subject and agreement (Hoyt 2002), and pronominal copula (Jelinek 1981; Eid 1983). These approaches have helped elucidate important aspects of Arabic syntax in particular and the syntax of natural language in general. There is a great deal that remains to be done, including further development of generative approaches to the diachronic syntax of Arabic, computational linguistic study of the formal and spoken varieties, and psycholinguistic and neurolinguistic research on the language. These areas have not received as much attention within the Arabic linguistic community as the areas discussed here, although there are promising attempts in these directions.

Bibliographical references

Adger, David. 2003. *Core syntax: A Minimalist approach*. Oxford: Oxford University Press.

Al-Tamari, Ahmed. 2001. *Sentential negation in English and Arabic: A Minimalist approach*. Ph.D. diss., University of Kansas.

Aoun, Joseph and Elabbas Benmamoun. 1998. "Minimality, reconstruction, and PF movement". *Linguistic Inquiry* 29.569–592.

——, Elabbas Benmamoun, and Dominique Sportiche. 1994. "Agreement and conjunction in some varieties of Arabic". *Linguistic Inquiry* 25.195–220.

——, Elabbas Benmamoun, and Dominique Sportiche. 1999. "Further remarks on first conjunct agreement". *Linguistic Inquiry* 30.669–681.

Aoun, Joseph and Lina Choueiri. 1996. "Resumption and last resort". Ms., University of Southern California.

—— and Lina Choueiri. 1999. "Modes of interrogation". *Perspectives on Arabic linguistics*, XII, ed. Elabbas Benmamoun, 7–26. Amsterdam and Philadelphia: J. Benjamins.

—— and Lina Choueiri. 2000. "Epithets". *Natural Language and Linguistic Theory* 18.1–39.

——, Lina Choueiri, and Norbert Hornstein. 2001. "Resumption, movement, and derivational economy". *Linguistic Inquiry* 32.371–403.

Aoun, Joseph and Yen-Hui Audrey Li. 2003. *Essays on the representational and derivational nature of grammar: The diversity of* wh-*constructions*. Cambridge, Mass.: MIT Press.

Ayoub, Georgine. 1981. *Structure de la phrase en arabe standard*. Ph.D. diss., Université de Paris-7.

Bahloul, Maher. 1994. *The syntax and semantics of taxis, aspect, tense and modality in Standard Arabic*. Ph.D. diss., Cornell University.

—— and Wayne Harbert. 1993. "Agreement asymmetries in Arabic". *Proceedings of WCCFL 11*, ed. Jonathan Mead, 15–31. Stanford, Cal.: Center for the Study of Language and Information.

Bakir, Murtadha. 1980. *Aspects of clause structure in Arabic*. Ph.D. diss., Indiana University.

Benmamoun, Elabbas. 1992. *Inflectional and functional morphology: Problems of projection, representation and derivation*. Ph.D. diss., University of Southern California, Los Angeles.

——. 1996. "Negative polarity and presupposition in Moroccan Arabic". *Perspectives on Arabic linguistics*, VIII, ed. Mushira Eid, 47–66. Amsterdam and Philadelphia: J. Benjamins.

——. 1997. "Licensing of negative polarity in Moroccan Arabic". *Natural Language and Linguistic Theory* 15.263–287.

——. 1999. "The syntax of quantifiers and quantifier floating". *Linguistic Inquiry* 30.621–641.

——. 2000. *The feature structure of functional categories: A comparative study of Arabic dialects*. Oxford: Oxford University Press.

——. 2006. "Licensing configurations: The puzzle of head negative polarity items". *Linguistic Inquiry* 37.141–149.

——. 2008. "Clause structure and the syntax of verbless sentences". *Foundational issues in linguistic theory: Essays in honor of Roger Vergnaud*, ed. Robert Freidin, Carlos P. Otero, and Maria Luisa Zubizarreta, 105–131. Cambridge, Mass.: MIT Press.

—— and Heidi Lorimor. 2006. "Featureless expressions: When morphophonological markers are absent". *Linguistic Inquiry* 37.1–23.

Carnie, Andrew. 2007. *Syntax: A Generative introduction*. 2nd ed. Oxford: Blackwell.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

——. 1995. *The Minimalist program*. Cambridge, Mass.: MIT Press.

——. 2000. “Minimalist inquiries: The framework”. *Step by step: Essays on Minimalist syntax in honor of Howard Lasnik,* ed. Roger Martin, David Michaels, and Juan Uriagereka, 89–155. Cambridge, Mass.: MIT Press.

——. 2001. “Derivation by phase”. *Ken Hale: A life in language*, ed. Michael Kenstowicz, 1–52. Cambridge, Mass.: MIT Press.

Choueiri, Lina. 2002. *Issues in the syntax of resumption: Restrictive relatives in Lebanese Arabic*. Ph.D. diss., University of Southern California.

Déchaine, Rose-Marie. 1993. *Predicates across categories*. Ph.D. diss., University of Massachusetts at Amherst.

Doron, Edith. 1996. “The predicate in Arabic”. *Studies in Afroasiatic grammar*, ed. Jacqueline Lecarme, Jean Lowenstamm, and Ur Shlonksy, 77–87. Leiden: Holland Academic Graphics.

—— and Caroline Heycock. 1999. “Filling and licensing multiple specifiers”. *Specifiers*, ed. David Adger, Susan Pintzuk, Bernadette Plunkett, and George Tsoulas, 69–89. Oxford: Oxford University Press.

Eid, Mushira. 1983. “The copula function of pronouns”. *Lingua* 59.197–207.

——. 1991. “Verbless sentences in Arabic and Hebrew”. *Perspectives on Arabic linguistics*, III, ed. Bernard Comrie and Mushira Eid, 31–61. Amsterdam and Philadelphia: J. Benjamins.

Eisele, John. 1988. *The syntax and semantics of tense, aspect, and time reference in Cairene Arabic*. Ph.D. diss., University of Chicago.

Fassi Fehri, Abdelkader. 1982. *Linguistique arabe: Forme et interprétation*. Rabat: Publications of the Faculty of Letters.

——. 1988. “Agreement in Arabic, binding and coherence”. *Agreement in natural language: Approaches, theories, description*, ed. Michael Barlow and Charles Ferguson, 107–158. Stanford, Cal.: Center for the Study of Language and Information.

——. 1993. *Issues in the structure of Arabic clauses and words*. Dordrecht: Kluwer Academic Publishers.

——. 1998. “Layers in the distribution of Arabic adverbs and adjectives and their licensing”. *Perspectives on Arabic linguistics*, XI, ed. Elabbas Benmamoun, Mushira Eid, and Niloofar Haeri, 1–45. Amsterdam and Philadelphia: J. Benjamins.

Fassi Fehri, Abdelkader. 2003. “Verbal plurality, transitivity, and causativity”. *Research in Afroasiatic grammar,* II, ed. Jacqueline Lecarme, 151–185. Amsterdam and Philadelphia: J. Benjamins.

Ferguson, Charles A. 1983. “God wishes in Syrian Arabic”. *Mediterranean Language Review* 1.65–83.

Gelderen, Elly van. 1996. “Parametrising agreement features in Arabic, Bantu languages, and varieties of English”. *Linguistics* 34.753–767.

Haegeman, Liliane. 1991. *Introduction to Government and Binding Theory*. Oxford: Blackwell.

Halila, Hafedh. 1992. *Subject specificity effects in Tunisian Arabic*. Ph.D. diss., University of Southern California.

Harbert, Wayne and Maher Bahloul. 2002. “Postverbal subjects in Arabic and the theory of agreement”. *Themes in Arabic and Hebrew syntax*, ed. Ur Shlonsky and Jamal Ouhalla, 45–70. Dordrecht: Kluwer Academic Publishers.

Hazout, Ilan. 1990. *Verbal nouns: Theta-theoretic studies in Hebrew and Arabic*. Ph.D. diss., University of Massachusetts at Amherst.

Heggie, Lorie. 1988. *The syntax of copular structures*. Ph.D. diss., University of Southern California.

Hoyt, Frederick. 2002. “Impersonal agreement as a specificity effect in rural Palestinian Arabic”. *Perspectives on Arabic linguistics*, XIII–XIV, ed. Dilworth Parkinson and Elabbas Benmamoun, 111–141. Amsterdam and Philadelphia: J. Benjamins.

Jelinek, Eloise. 1981. *On defining categories: Aux and predicate in Egyptian Colloquial Arabic*. Ph.D. diss., University of Arizona.

Kenstowicz, Michael. 1989. “The null subject parameter in modern Arabic dialects”. *The null subject parameter,* ed. Osvaldo Jaeggli and Ken Safir, 263–275. Dordrecht: Kluwer Academic Publishers.

Lorimor, Heidi. 2007. *Conjunctions and grammatical agreement*. Ph.D. diss., University of Illinois at Urbana-Champaign.

Mohammad, Mohammad. 1988. “On the parallelism between IP and DP”. *Proceedings of WCCFL VII,* ed. Hagit Borer, 241–254. Stanford, Cal.: Center for the Study of Language and Information.

——. 1989. *The sentence structure of Arabic*. Ph.D. diss., University of Southern California.

——. 1999a. *Word order, agreement and pronominalization in Standard and Palestinian Arabic*. Amsterdam and Philadelphia: J. Benjamins.

——. 1999b. “Checking and licensing inside DP in Palestinian Arabic”. *Perspectives on Arabic linguistics*, XII, ed. Elabbas Benmamoun, 27–44. Amsterdam and Philadelphia: J. Benjamins.

Mouchaweh, Lina. 1986. *De la syntaxe des petites propositions*. Ph.D. diss., Université de Paris-7.

Moutaouakil, Ahmed. 1987. *Min qaḍāyā r-rābiṭ fī l-luġa al-ᶜarabiyya*. Casablanca: ʿOqāḏ.

Munn, Alan. 1999. “First conjunct agreement against a clausal analysis”. *Linguistic Inquiry* 30.643–668.

Ouhalla, Jamal. 1993. “Negation, focus and tense: The Arabic *mā* and *lā*”. *Rivista di Linguistica* 5.275–300.

——. 1994. “Focus in standard Arabic”. *Linguistics in Potsdam* 1.65–92.

——. 1996. “Remarks on the binding properties of *wh*-pronouns”. *Linguistic Inquiry* 27.676–707.

——. 1997. “Remarks on focus in Standard Arabic”. *Perspectives on Arabic linguistics*, X, ed. Mushira Eid and Robert Ratcliffe, 9–45. Amsterdam and Philadelphia: J. Benjamins.

Pollock, Jean-Yves. 1989. “Verb movement, UG and the structure of IP”. *Linguistic Inquiry* 20.365–424.

Shlonsky, Ur. 1997. *Clause structure and word order in Hebrew and Arabic: An essay in comparative Semitic syntax*. Oxford: Oxford University Press.

——. 2002. “Constituent questions in Palestinian Arabic”. *Themes in Arabic and Hebrew syntax*, ed. Jamal Ouhalla and Ur Shlonsky, 137–160. Dordrecht: Kluwer Academic Publishers.

Soltan, Usama. 2007. *On formal feature licensing in Minimalism: Aspects of Standard Arabic morphosyntax*. Ph.D. diss., University of Maryland.

Wahba, Wafaa. 1984. Wh-*constructions in Egyptian Arabic*. Ph.D. diss., University of Illinois at Urbana-Champaign.

——. 1991. “LF-movement in Iraqi Arabic”. *Logical structure and linguistic structure,* ed. C.-T. James Huang and Robert May, 253–276. Dordrecht: Kluwer Academic Publishers.

Elabbas Benmamoun
(University of Illinois at Urbana-Champaign)

Syria

1. Arabic and minority languages

In addition to Arabic, the following languages (in order of number of speakers) are spoken in Syria (see Map 1). → Kurdish is spoken on the northern border with Turkey, the main areas being to the northwest of Aleppo and in the northeast in the Qāmišli area. Approximately one hundred thousand Kurds from Turkey live in northeastern Syria. Most linguistic maps used in connection with the Kurdish problem are not exact and ascribe too large an areal distribution to Kurdish (cf. the map in Behnstedt 1992a for northeastern Syria). Turkish dialects are spoken north of Aleppo on the Syrian-Turkish border and on the northern coast, also near the Turkish-Syrian border. There are Turkish language islands in the Qalamūn area and the Ḥomṣ area. Four dialects of → Neo-Aramaic are spoken in Syria: Neo-West Aramaic in three villages near Damascus, the most famous one being Maʿlūla. This is the only autochthonous Aramaic spoken in Syria. Ṭuroyo speakers from the Ṭūr ʿAbdīn mountains (Turkey) have settled in several places in the province of al-Ḥasaka, e.g. in Qāmišli, ilQaḥṭānīye, and several small villages. A relatively large linguistic island is formed by the Assyrians along the Khabur River, their dialects being spoken in twenty localities. The speakers originate from Iraq and ultimately from the Hatari province in Turkey. Other Aramaic speakers, so-called Kildān (Chaldeans), have also been transplanted from Iraq and are found in two villages in the far northeast of Syria. From the Caucasus, two languages have been brought to Syria, namely Circassian, spoken in some villages south of Aleppo, in the Ḥomṣ area, and on the Golan (Ġōlān) Heights. Many of the Circassian speakers have left the Golan and are living in Damascus. Chechenian is spoken in two villages on the Khabur River, but it seems to be on the verge of dying out, especially among male speakers. Armenian is spoken mainly in the large towns, such as Aleppo and Damascus, and in one small town exclusively, in Kasab on the coast near the Turkish border. Finally, Greek is spoken in the village of ilḤamīdīye on the coast near the Syro-Lebanese border. Its inhabitants originate from Crete and came there via Tripoli (Lebanon). Being Muslims, they preferred to emigrate when the Ottoman rule came to an end in Crete.

None of these minority languages has any official status. A demand of the inhabitants of ilḤamīdīye to be allowed to teach Greek in their schools was rejected with the argument that they are Muslims. An exception is Armenian: the Armenians have their own schools where Armenian is taught and used in addition to Arabic. They are also the only minority using their own script, e.g. on shop signs. The Jacobites (Arabic speakers), Nestorians, and Ṭuroyos (Jacobites) are allowed to use Aramaic (Syriac) in church, and Aramaic is also taught privately. As for the Arabic-speaking Jacobites, Syriac is allowed to be taught in church institutions; in Aleppo, for example, it is taught in a modernized version comparable to Modern Standard Arabic (cf. *radōytō* ‘car’).

The interaction between these languages is manifold. Syrian Arabic dialects borrowed many loanwords from → Turkish during Ottoman rule. Conversely, the vocabulary, for instance of Northwest Aramaic, is highly Arabicized (→ Aramaic/Syriac). The same is true to a lesser extent for → Kurdish, Turkish, and the other minority languages. Eastern Arabic dialects have integrated some Kurdish words.

2. The history of Arabic in Syria

Arabic was spoken in Syria long before the Islamic conquest. Suffice it to mention the Itureans (2nd c. B.C.) in the Lebanon and Anti-Lebanon, the Roman emperors of Syrian origin Philippus Arabs and Heliogabalus, the Nabataeans in the southern Ḥōrān, and of course the Arabs in Palmyra, who are attested there at least seven centuries before the Islamic conquest (further details in Cantineau 1934; Oppenheim 1939; Cantineau 1946; Altheim-Stiehl 1964–1969; Shahid 1989; Retsö 2006). Perhaps the Arabic spoken in localities of the Syrian steppe or on its fringe (Palmyra, Soukhne, ilQaritēn, etc.) goes back in part to this pre-Islamic Arabic, showing quite archaic features, such as the scheme *katīr* instead of *ktīr*, and distinction of masculine and feminine in the plural in some of these dialects. Yet, some C and D dialects also show remarkable archaisms (see below).

The substratal influence of Aramaic in Syrian Arabic proper (see below) is important and is present in phonology, morphology, lexicon, and syntax (→ substrate). Examples from phonology are the distribution of the diphthongs *ay, aw* and the monophthongs *ē*, *ō*, following the Aramaic model in dialects of the coast and some of the Qalamūn (*bēt/baytu* 'house/his house', *sōṭ/sawṭu* 'voice/his voice'), or the distribution of *ō* and *a* in the Qalamūn dialects and the dialect of Mḥardi (Northwest Aramaic: *fallōḥa/fallaḥō* 'peasant/peasants', Arabic: *fallōḥ/fillaḥīn*; Arnold and Behnstedt 1993:73ff.; for Mḥardi, see Behnstedt 1992b), and the elision of *a* in open unstressed syllables (Diem 1997:47). The personal pronouns of the 3rd person plural *hinne, hinnen* common, *hinnon* masculine – *hinnin* feminine, etc. undoubtedly have been influenced by the Aramaic ones (cf. Behnstedt 1991; Arnold and Behnstedt 1993:75ff.). As for proper loans (which are a phenomenon of 'borrowing', not of substratal influence *strictu sensu*, i.e. 'imperfect learning'), there are many lexical items, mainly in agriculture, local fauna and flora, and local culture (jars, baskets, etc.), like *ṣumd* 'plow', *buṛk* 'plow beam', *ʾaṭrīb* 'peg on the yoke', *kasaḥ* 'to trim the vines', *kāṛa* 'the cushion with which bread is flapped at the inner wall of the oven', *ʾarṭal* 'basket', *ḥardōn* 'lizard', *ṣāʿūr* 'billy goat', *širš* 'root', *nāṭūr* 'field-guard' (with the Aramaic nominal scheme), but also very common vocabulary like *šōb* 'heat' (for details, see Feghali 1920–1922; Arnold and Behnstedt 1993:80ff.; Aramaic etymologies are also indicated by Barthélemy 1935–1969). A syntactic Aramaism might be found in constructions of the type *ʾəltillo la-əbno* 'I told his son', *šəft(u) l-xayyak* 'I saw your brother', *ʿammo la-flān* 'the brother of somebody' (cf. Diem 1979:47ff.). Another example is that of *šī* to indicate indefiniteness, which according to Diem (1997:49) reflects Syriac *meddem*, Neo West Aramaic *mette*, e.g. *ši ʿelle* = Neo West Aramaic *mett ʿellṯa* 'some disease'.

Some loanwords in Syrian Arabic proper might even be derived from a pre-Aramaic Semitic substrate, e.g. *ṣifdaʿa* 'frog', *ṣumd* 'plow', with *ṣ* for **ḍ* in coastal dialects.

3. DIALECTOLOGY

There is a difference between 'Syrian Arabic' and 'Arabic dialects spoken in Syria'. Due to migration, Bedouin dialects from what is now Saudi Arabia and Iraq and sedentary dialects from Iraq and Anatolia are also spoken in Syria. As a first step, Arabic in Syria may be divided into two main groups: sedentary dialects and Bedouin dialects.

3.1 *Bedouin dialects*

Bedouin Arabic covers the largest part of the country. The Bedouin dialects themselves can be divided into two subgroups, the so-called Šāwi dialects, which are the dialects of (former) sheep breeders, and the North Arabian dialects of the (former) camel breeders (see Cantineau 1936, 1937; Johnstone 1967:2ff with general description; Ingham 1982; Behnstedt 2000). The North Arabian dialects in Syria are recent and have been brought into the country with the intrusion of the ʿAnaza tribes in the 18th century, later with the Šammar coming through Iraq and settling mainly in northeastern Syria. The area of the North Arabian dialects has been somewhat reduced because many of the Rwaḷa tribe, attracted by the wealth of Saudi Arabia, have returned to their old homeland. Most of the former Bedouin have settled, and only approximately fifty thousand true nomads are left. Transhumance, however, still is very common. One main difference in phonology between these two types are the reflexes of **k* and **q* in front environment: Šāwi dialects *č* and *j*, e.g. *čalb* 'dog', *jiddām* 'in front of', North Arabian *c* and d^z: *calb*, d^z*iddām*, a distinction which seems to be on the verge of disappearing, *c* and d^z being replaced by the more common *č* and *j*. Both types have undergone a certain 'Syrianization', and some features are in regression, e.g. the plural feminine ending *-āh* being replaced by *-āt* in North Arabian, and *čēf* 'how?', the main form quoted by Cantineau (1937:208), for which nowadays *šlōn* is used (more details in Behnstedt 1994b:424).

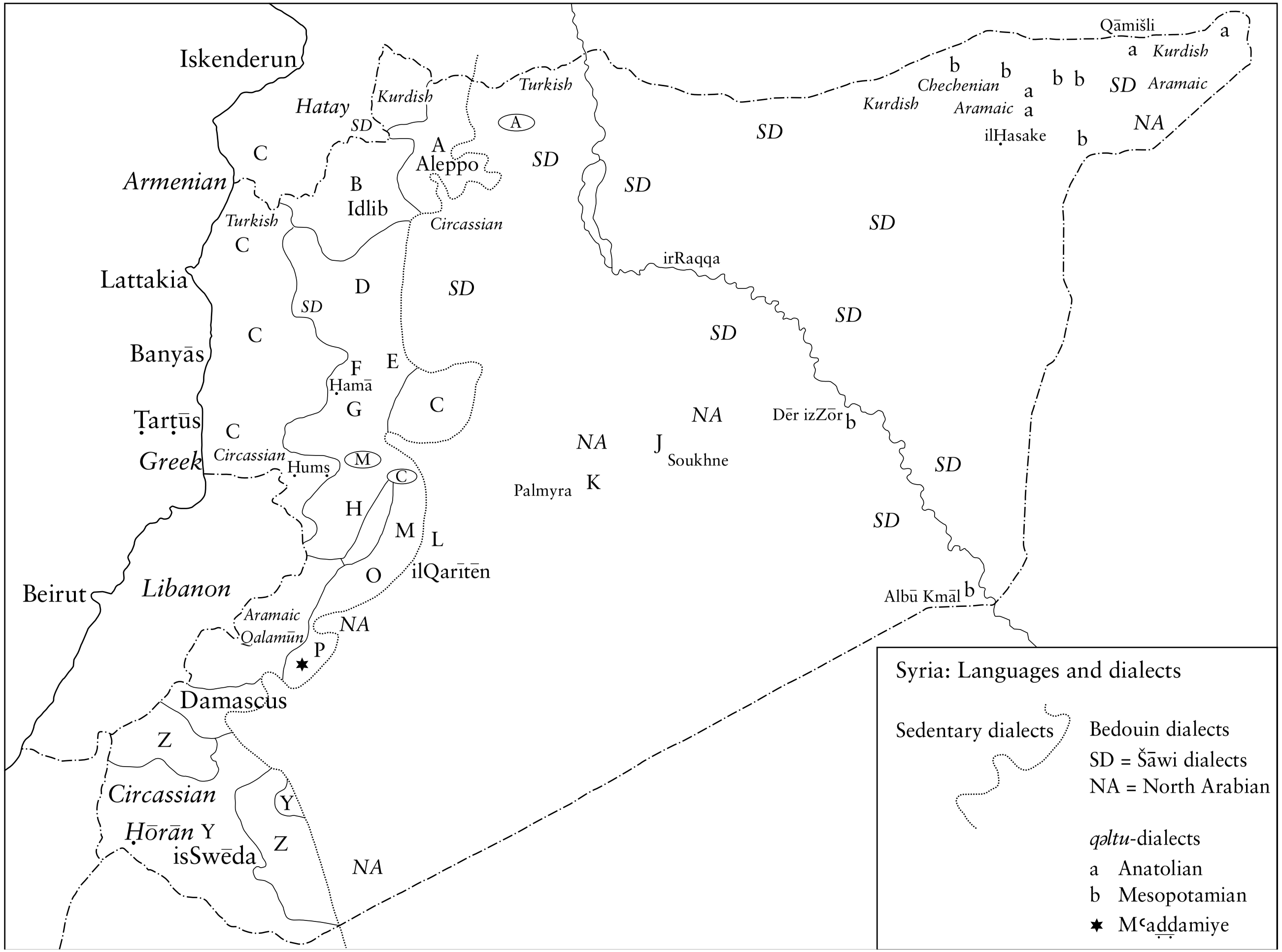
3.2 *Sedentary dialects*

Syrian Arabic as a sedentary dialect type is spoken in the western part of the country. It should not be confused with Damascus Arabic. The merging of *i* and *u* except in final closed syllable into *ə* in the dialect of Damascus does not apply to most Syrian dialects (cf. Behnstedt 1997b:22–25). As a dialect type, Syrian Arabic is spoken not only in Syria but also in Lebanon, → Antiochia, and → Cilicia and is closely related to Urban → Palestinian. Here, only the Arabic of Syria proper is dealt with, with the exception of → Damascus Arabic.

The question of what is typical for Syrian Arabic is difficult to answer. Within the frame of Near Eastern sedentary dialects, perhaps the intonation combined with certain pausal features, some phonetic and morphophonemic phenomena (treatment of diphthongs, → *ʾimāla* in many dialects, treatment of **a*), certain morphological forms like the personal pronouns and pronominal suffixes or imperative forms like *šṛāb* 'drink!', *lbēs* 'dress!', and some lexical items (*šōb* 'heat', *hōn* 'here', *hēk* 'so') are characteristic. But actually, there is a rich variety of dialectal features, and the statement in Fischer and Jastrow (1980:27), according to which Arabic in Syria was "one of the best investigated dialect groups and that hardly any linguistic detail had not been registered somehow", was precipitate. It is impossible in this entry to give a fair idea of the richness of linguistic features of Syrian Arabic (almost a hundred forms for 'those' have been attested), and vocabulary has to be neglected (see above for Aramaic loans). The following sketch follows the classification given in Behnstedt (1997a, Maps 499–518).

The shibboleth form of Northern dialects (Aleppo and surroundings = A) is the 1st person singular imperfect with *a-* (and not *i-* or zero as in most other Syrian sedentary dialects): *ašṛab* 'I drink', *ašūf* 'I see'. *ʾImāla* of the type *sēfaṛ/ysēfer* 'to travel' is widespread. Both *q* and *ʾ* are found. In another Northern group (B) with *išṛab*, diphthongs in every position are widespread, the *ʾimāla* is partly of the type *sāfaṛ/ysēfer*, and there are first traces of *a-* elision (*ktab+t > katabt*, but *katab+it > katabit* in some localities). These two dialect groups might be described as showing no extreme evolutions. One of their peculiarities is a phoneme *č* (from Turkish).

The dialect group C (coast and coastal mountains, with linguistic islands outside the area due to migration) covers the largest area of the west, continuing into Lebanon, Antiochia, and Cilicia in Turkey. Here we are confronted with a maximum of deviations compared to the Northern dialects. The reflexes of the Old Arabic diphthongs are *ē* – *ay*, *ō* – *aw* according to the position 'close' – 'open' (*bēt/bayti* 'house/my house', *ṣōt/ṣawti* 'voice/my voice'), but another distribution is attested as well, namely monophthongs in unsuffixed forms and diphthongs in suffixed ones (*baytna* 'our house', *ṣawtna* 'our voice'). Also, *ā* is found in many lexemes for both **ay* and **aw* (*sāf, yām*), and analogical formations like **jāj* > *jēj/jayji* 'chickens/chicken', and accordingly *tmayni* 'eight', **fāṛ* > *fōṛ/fawṛa* 'mice/mouse', occur. In the northern half, the phonemes **i* and **u* seem to have merged into *i* (as in Antiochian and Cilician Arabic), while in the south, the opposition is maintained. The *ʾimāla* in the north is conditioned, in the south unconditioned: *hāda* – *hēda* 'this'. An exception are forms of the type *nayyim* 'sleeping' (for which one has rather to suggest analogy to adjectives of the type *jayyid* than omission of the *ʾimāla* due to *ʾ* as in Classical Arabic *nāʾim*). In some dialects, a split of **ā* into *ō* and *ē* is found, with rather complex rules, e.g. Mḥardi: *ṛōs/ṛāsēn* 'head/two heads', *qōl/qālit* 'he/she said'. The treatment of unstressed *a* is certainly the most striking feature of this dialect group. Whenever possible, it is elided or raised to *i* and *u* (diachronically and synchronically): *katab+t > ktabt, katab+it > katbit, sallam+it > sallmit, sallam+t > sillamt, ḥaṭṭ+ayt > ḥiṭṭayt, trawwaq+t > truwwaqt,* **madrasa > madrsa > mádərsa, madírsi,* **fallāḥ > fillāḥ*. Another characteristic development (also attested in other Syrian dialect groups) is the split of *a* into *i* (*e*) and *o*: *aswad > aswid* (*aswed*), *axḍar > axḍor*; cf., however, *balid* (*baled*) but *baladi*. Fronting of *-a-*, *-a* leads to complex systems of short vowels: **inta > inti* 'you [sg. masc.]', **inti > inte* 'you [sg. fem.]', **hawdat > hawdet* 'these [masc.]' vs. *hawdit* 'these [fem.]'. This also explains forms like *ʿindik* 'you [sg. fem.] have' vs. *ḍarbet* 'she has beaten' < **ḍarbat*. Word accent in part of the dialects is determined not by


Syria: Languages and dialects
Sedentary dialects
Bedouin dialects
SD = Šāwi dialects
NA = North Arabian
qəltu-dialects
a Anatolian
b Mesopotamian
★ Mʿaḏḏamiye
Iskenderun
Hatay
Kurdish
Turkish
Armenian
Lattakia
Banyās
Ṭarṭūs
Greek
Beirut
Libanon
Aramaic
Qalamūn
Damascus
Circassian
Ḥōrān
isSwēda
Aleppo
Idlib
Ḥamā
Ḥumṣ
ilQarīten
Palmyra
Soukhne
irRaqqa
Dēr izZōr
Albū Kmāl
ilḤasake
Qāmišli
Chechenian
SD
NA
A
B
C
D
E
F
G
H
J
K
L
M
O
P
Y
Z
a
b

Map 1. Syria: Languages and dialects.

syllable structure but by grammatical category, e.g. *tinzíli* ‘you [sg. fem.] go down’, *tinzílu* ‘you [pl.] go down’ vs. *mádərsa* ~ *madírsi*. Pausal phenomena, especially lowering of *-i* > *-e*, *-a* are strongly marked. Such pausal forms are also used partly in context.

Among the striking morphological forms are feminine forms in the plural with pronouns, demonstratives (rare: *hawdet/hawdit* ‘these [masc./fem.]’, *hawket/hawkit* ‘those [masc./fem.]’), and verbs in the center of the C dialects: *(h)intu ktabtu/(h)intni ktabtni/hinnun katbu/hinnin ktabni* ‘you [masc. sg.]/you [sg. fem.]/they [masc.]/they [fem.] wrote’, etc. Other deviating formations are *nzēl* ‘I go down’, *šṛāb* ‘I drink’, and even *kēl* ‘I eat’, identical with the imperative. Remarkable are the pronominal singular forms of 1st and 2nd persons with *h*–: 1st person singular *hana*, 2nd person singular masculine *hint*, 2nd person singular feminine *hinte*, and rarities like *liḥna* or *riḥna* in some local dialects. Negation is of the type *a-katab* (perf.), *a-mā-yiktub* (imperf. with verb modifier present *mā*).

The Central dialects (D – H) present a certain degree of leveling without such striking features (monophthongs, few *a*-elisions, predominantly no *ʾimāla*), but the north (D) still has feminine forms in the plural. In the center, one of the most curious Syrian dialects is found, the dialect of Ṣōran and Ṭayybit ilImām, with a C background but with extreme developments within the pronominal and verbal system (most probably due to contact with Bedouin Arabic), internal evolution caused by lowering of vowels, *ʾimāla*, and analogical formation. The verbal endings *-am, -aw* (Bedouin) and *-o* (sedentary) have merged into *-a: gāḷam* x *qālo* > *qāla* ‘they said’. Consequently, *yqūlu* > *yqūla* (imperf.), *intu* > *inta* ‘you [pl.]’; *inti* ‘you [sg. fem.]’ > *inte, inta* ‘you [sg. masc.]’ > *inti*. Cf. also *hinhan* ‘they [masc.]’ – *hinhin* ‘they [fem.]’, or *ake* ‘he ate’ (following *nise*) – *yāka* (following *yinsa*), but *yākúla* ‘they eat’.

The dialects spoken in the Syrian steppe from Soukhne to the fringes of the Qalamūn (J – Q) do not form a homogeneous group but rather consist in a patchwork forming a continuum with some striking common features found nowhere else. As for the dialects of Soukhne and Palmyra, see mainly Behnstedt (1994a) and Cantineau (1934). One striking feature is the conservation of the nominal scheme of the type *katīr* ‘much’ and consequently *faṭūr* for **fuṭūr* ‘breakfast’. To give one example of the gradual transitions within this group: Soukhne, Arak, and Palmyra share the *i*-type of verbs I (strong and final y): *ʾílbis, ʾínsi*; ilQarītēn, Huwwērīn, and Mhīn Fruqlus still have *ʾínsi*, but *lbīs*; the dialects farther to the west of L-iḤmēra, Dēr ʿAṭīye, and inNabk have *lbīs* (*lbēs*), *nsī* (*nsē*). In this ‘group’, one of the most deviating dialects is that of inNabk, with extreme evolutions within the vowel system: monophthongization of **ay* and **aw* > *ā: banōt* ‘girls’ – *banāt* ‘I have built’; *fōʾ* ‘he got up’ – *fāʾ* ‘on, upon, above’; almost regular shift *a* > *i, u* in C*a*CC and similar structures: *kalb* > *kilb, xadd* > *xudd, šadd* > *šudd, kamašt* > *kamušt, nazzalt* > *nazzult, sēfoṛ* but *sēfuṛt*, lengthening of *a* in -C*a*C: *walad* > *walēd, maṭar* > *maṭōr, sakan* > *sakōn*. Another striking dialect is that of Mʿaḏḏamīye, a so-called *qəltu* (*ʾəltu*) dialect (→ see Iraq), spoken approximately 40 kms northeast of Damascus, which presents other features of the *qəltu* dialects, like *ṣafṛā* ‘yellow [fem.]’, lexical elements hinting at an eastern origin, but also features from the area, such as an unconditioned *ʾimāla* of *ā*. Here, too, strange developments are found, such as **fawʾ* > *fōʾ* ‘above’ vs. **banayt* > *banāt* ‘I have built’, distinguished from *banēt* ‘girls’.

The dialects of the Qalamūn area proper are a similar patchwork; dialects vary from village to village. In some places *č* is the reflex of *j*. Interdentals are maintained in the major part of it. The conservation of diphthongs, splitting of *ā* into *ē* and *ō*, an *ʾimāla* of the type *lsān* > *lsīn*, and many pausal phenomena relate some of its dialects to Lebanon and the C dialects. But *a*-elision is not very marked, and in some localities even *i* is not elided in forms like *libisit, wisixa*, which contradicts entirely the existing image of Syrian Arabic. Shortening of unstressed long vowels is characteristic: **sakākīn* > *sakakīn* ‘knives’, *fallōḥ/fillaḥīn* ‘peasant/peasants’, or *fillōḥ/filliḥīn*, as in Northwest Aramaic. Characteristic pausal features are diphthongization (*kbīr* > *kbeyr*#, *yṛāḥ*#> *yṛawḥ* , *tīn* > *tī*^{a}*n*#), lowering (*jbīn* > *jbēn*#, *ʿēli* > *ʿēla*#). In the dialect of Drayj, the perfect of the *a*-type of the verbs III*y* has been analogically restructured: *ṣʾayt* (1st pers. sg.) – *aṣʾa* (3rd pers. sg. masc.), *rmayt* – *arma*, etc. As for negation, the type *mā- -š* is already attested along with the simple negation.

The dialects in the vicinity of Damascus (W) show less-striking features; they have plosives and not interdentals like the Qalamūn dialects, and monophthongs, *ž* for **g*, and thus, they are closer to Damascus itself (cf. also *hadunke* 'those', similar to the Damascus form). But the opposition of *i* and *u* in other environments than final closed syllable is maintained till the suburbs of Damascus.

Going farther south, the Z dialects on Mount Hermon, in the north of the Ḥōrān, and in the Jabal idDrūz are dialects close to Southern Lebanese (unconditioned *ʾimāla, katabat*). The Druze have immigrated to this area from Lebanon, from the 17th century onward. One of the common features of all these Syrian sedentary-type dialects of the extreme south is the presence of interdentals. There are regional differences, such as for *jīm* (*j* ~ *ž*) and the 1st person singular imperfect (*ašṛab, išṛab*).

3.3 *Mixed dialects*

All dialects in the towns and villages of the Syrian steppe or dialects on the fringe of it have a strong admixture of Bedouin Arabic (e.g. Palmyra, Ṭayybit illmām, dialects in the surroundings of Aleppo), the highest degree certainly being found in the dialect of Soukhne. Examples are, for Palmyra and Soukhne, the emphatic *ḷ* in forms like *naxḷa* 'palm tree', *qaḷb*, *kaḷb* 'heart', and *č*, rare in the dialect of Palmyra but common in the dialect of Soukhne. There are many loanwords with *g* in the dialects of Palmyra and Soukhne. The Afāḏ̣le, sedentarized Bedouin east of Aleppo, for instance, have more or less given up their original dialect but have maintained interdentals; on the other hand, they have not 'adopted' the *ʾimāla* characteristic of the sedentary-type dialects of the area. A Bedouin feature in this 'group' is *m-* in *mākil, māxiḏ* instead of *ākil, āxiḏ* (also to be found in the Ḥōrān dialects). The pronoun *aham* 'they' in Soukhne is a contamination of *ahu* (as in Palmyra) and *ham* (found in surrounding dialects). As for vocabulary, these dialects use Bedouin *bī* 'there is' instead of *fī, hnāk* 'there' instead of *hōnīk, hnīk* or similar forms, and sheep-rearing vocabulary like *gaḥam* 'five-year-old wether' or *wardi* 'fat young lamb'.

3.4 *Ḥōrān dialects*

The Ḥōrān dialects (Y) are closely related to Jordanian and Palestinian rural dialects, and according to Cantineau (1946:418), they are not Syrian Arabic proper. They have clear Bedouin features; they differ, however, in syllable structure from North Arabian and Šāwi dialects (*ḏ̣arbat – ḏ̣arabu – ḏ̣arabin* vs. *ḏ̣rubat – ḏ̣rubam/ḏ̣rubaw – ḏ̣ruban*) and most probably present an older dialectal layer. As for Cantineau's classification (adopted by Behnstedt 1997a), it has to be partially revised. The dialects of the Ḥōrān mountains, considered by Cantineau to be a heterogeneous group, rather have to be divided according to dialectometrical measurings into two subgroups, a Northern one, almost identical with Ḥōrāni proper, and a Southern one, which shows more traits of a transitional area.

3.5 *Anatolian dialects*

Due to emigration, many speakers of → Anatolian *qəltu* dialects have settled in the first quarter of the 20th century in the northeast of Syria in Dirbasīye, Rās ilʿAyn, ʿĀmūda, ilḤasake, and Qāmišli, and surrounding villages (see Behnstedt 1992a; Isaakson 2000). The Anatolian dialect of Āzəx, meanwhile extinct in Turkey, has survived in əlMālkīye in northeastern Syria.

3.6 *Mesopotamian dialects*

There is only one certainly autochthonous Syro-Mesopotamian *qəltu* dialect, namely that of alXatuniyye, first sketched by Behnstedt (1992a), and described with more details about its areal distribution in a monograph by Talay (1999), who indicates 27 localities where this dialect is spoken. As for Dēr izZōr (cf. Jastrow 1978:26), doubts of the Syrian origin of its dialect are allowed, since Dēr izZōr was uninhabited during a part of the Middle Ages; the modern town was founded by the Ottomans in 1867, when they established new garrisons on the Euphrates, which they settled with people from Iraq. The inhabitants of Albū Kmāl, who also speak a *qəltu* dialect (to be more precise a *gəltu* dialect), hail from ʿĀna (90%) and Rāwa (10%) in Iraq. As for the *ʾəltu* dialect of Mʿaḏ̣ḏ̣amīye, its origin from

Mesopotamia could recently be established (Behnstedt 2000:324). All Mesopotamian dialects proper are heavily mixed with Bedouin Arabic.

4. DIGLOSSIA

For ideological reasons, the use of dialect is not welcome in official life. Teachers are obliged to speak only Modern Standard Arabic with their pupils (which they actually do only partly). In the media, dialect is used to a lesser extent than in Egypt. A form like *w alʾān binaltaqi bi* 'and now we meet', which may be heard on Egyptian television, is unthinkable on Syrian television. In television documentaries, too, strictly Modern Standard Arabic is spoken; a mixture of dialect and Modern Standard Arabic, as is quite common in Egyptian documentary films, is not usual. Dialect in the electronic media is more or less restricted to movies, *musalsalāt*, theater plays (often in Damascus dialect), or satirical sketches. Besides Damascus Arabic, Bedouin Arabic, or other dialects (for instance, in one *musalsal* the dialect of Dēr izZōr) may be used in the *musalsalāt*. In oral poetry and songs, dialect is widely used, and Bedouin dialects enjoy a certain prestige. The quite-famous poem of al-Ḥamda was composed by a poet from ilQaritēn (where a sedentary dialect is spoken) in a Bedouin dialect.

Language policy fostering Modern Standard Arabic has led to a certain Arabicization of the dialect vocabulary. Many lexical items of foreign origin found in Barthélemy (1935–1969) have meanwhile been replaced by Arabic terms, e.g. *tumbīl* 'car' by *sayyāṛa; ṭarnawīz* 'screwdriver' by *mafakk barāġi; abukāto* 'lawyer', the only form found in Barthélemy (1935–1969), by *muḥāmi* (more details in Behnstedt 1996). For the regression of Turkish loans. see Barbot (1961).

5. EMERGENCE OF A NATIONAL VARIETY OF THE COLLOQUIAL

There is no national standard colloquial comparable to the Cairo dialect in Egypt. The dialect of Damascus does not play this eminent role. There is what one may call a 'common Syrian Arabic', which, however, allows regional variants. There are rather local koines like the Damascus dialect, the dialects of Aleppo and Ḥama, or the Mardīn dialect of Qāmišli in northeastern Syria. Or to put it otherwise: acceptable is what is not too deviant. Thus the pronunciations of *jīm* as *ž* and *j* are both equally acceptable, while others, like *č* (Qalamūn, Palmyra) or *c* (Soukhne), are not; according to the sociolinguistic context, these are replaced by locally acceptable variants, i.e. *ž* in the Damascus area and *j* in Palmyra or Soukhne. Speakers of Aleppo Arabic in a contact situation might replace *sēfaṛ* 'he traveled' by *sāfaṛ*, but will stick to their *j*. *ʾImāla* of the type *ṯēni* (*tēni*) or *ṯīni* (*tīni*) 'second' is indeed often avoided. The same is true for *ō* and *ē* < **ā* in coastal and Qalamūn dialects, pronunciations that are replaced by *ā*. The pronunciation of → *qāf* as a glottal stop is prestigious and is gaining ground in the surroundings of bigger towns like Damascus or Aleppo. The *q*- speakers are often mocked as *byiḥku b-luġat il-qaqaqa*. But *q* is prestigious among speakers of Bedouin dialects. The Šāwi dialect speakers, however, having *q* < **ġ* (**ġanam* > *qanam* 'sheep') often produce hypercorrections like *ġāl* 'he said', *astaqfiru ḷḷāh* 'I ask God's forgiveness!'. For other interferences from sedentary dialects in Bedouin dialects, see the texts in Behnstedt (2000:460ff.).

6. ATTITUDES TOWARD FOREIGN LANGUAGES

The sole language of instruction is Arabic, including instruction in universities. Only one foreign language is taught in the secondary school; 60 percent of the pupils have to choose English, 40 percent French (if available). English is used in a second television channel and in the press. Middle- and upper-class Arab Christians in the big towns like Aleppo and Damascus are still very much orientated toward French and use it as a secondary cultural language. Code-switching is common, and if they speak Arabic, it is often highly mixed with French, as in *bonjour!*, to which the answer is: *bonjourāt!*; or *ʾiltillo l'emballage maw kwayyes, alors riḥət ʿand il-concurrent w laʾət ġēr situation* 'I told him the packing material wasn't good, so I went to the competitor and found another situation'.

Bibliographical references

Altheim, Franz and Ruth Stiehl. 1964–1969. *Die Araber in der alten Welt.* 4 vols. Berlin: W. de Gruyter.

Arnold, Werner and Peter Behnstedt. 1993. *Arabisch-aramäische Sprachbeziehungen im Qalamūn (Syrien).* Wiesbaden: O. Harrassowitz.

Barbot, Michel. 1961. "Emprunts et phonologie dans les dialectes citadins syro-libanais". *Arabica* 8.174–188.

Barthélemy, A. 1935–1969. *Dictionnaire arabe-français dialectes de Syrie: Alep, Damas, Liban, Jérusalem.* Paris: P. Geuthner.

Behnstedt, Peter 1991. "Noch einmal zum Problem der Personalpronomina *hənne* (3.Pl.), *-kon* (2.Pl.) und *-hon* (3.Pl.) in den syrisch-libanesischen Dialekten". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 141.235–252.

——. 1992a. "*qəltu*-Dialekte in Syrien". *Zeitschrift für Arabische Linguistik* 24.35–59.

——. 1992b. "Zur Spaltung von *ā in ō/ē im Qalamūn (Antilibanon)". *Welt des Orients* 23.77–100.

——. 1994a. *Der arabische Dialekt von Soukhne (Syrien).* II. *Phonologie, Morphologie, Syntax.* III. *Glossar.* Wiesbaden: O. Harrassowitz.

——. 1994b. "Dialektkontakt in Syrien". *Actes des Premières Journées de Dialectologie Arabe de Paris*, ed. Dominique Caubet and Martine Vanhove, 417–430. Paris: Publications Langues'O.

——. 1996. "Romanisches Lehngut im Syrisch-Arabischen". *Romania Arabica: Festschrift für Reinhold Kontzi zum 70. Geburtstag*, ed. Jens Lüdtke, 63–72. Tübingen: G. Narr.

——. 1997a. *Sprachatlas von Syrien.* Wiesbaden: O. Harrassowitz.

——. 1997b. *Sprachatlas von Syrien. Beiheft.* Wiesbaden: O. Harrassowitz.

——. 2000. *Sprachatlas von Syrien.* II. *Volkstümliche Texte.* Wiesbaden: O. Harrassowitz.

Cantineau, Jean. 1934. *Le dialecte arabe de Palmyre.* I. *Grammaire.* II. *Vocabulaire et textes.* Beirut: Mémoires de l'Institut Français de Damas.

——. 1936. "Etudes sur quelques parlers de nomades arabes d'Orient". *Annales de l'Institut d'Études Orientales* (Alger) 2.1–118.

——. 1937. "Etudes sur quelques parlers de nomades arabes d'Orient" (2nd article). *Annales de l'Institut d'Études Orientales* (Alger) 3.119–237.

——. 1940. *Les parlers arabes du Ḥōrân: Atlas.* Paris: C. Klincksiek.

——. 1946. *Les parlers arabes du Ḥōrân.* Paris: C. Klincksiek.

Diem, Werner. 1979. "Studien zur Frage des Substrats im Arabischen". *Der Islam* 56.12–80.

Feghali, Michel T. 1920–1922. "Etudes sur les emprunts syriaques dans les parlers arabes du Liban". *Mémoires de la Société de Linguistique de Paris* 21.210–248, 21.257–282; 22.13–42.

Fischer, Wolfdietrich and Otto Jastrow. 1980. *Handbuch der arabischen Dialekte.* Wiesbaden: O. Harrassowitz.

Ingham, Bruce. 1982. *North East Arabian dialects.* London and Boston: Kegan Paul International.

Isaakson, Bo. 2000. "Report from a research project and a brief account of the dialect of ʿĀmūda". *Aspects of the dialects of Arabic today: Proceedings of the 4th Conference of the International Arabic Dialectology Association (AIDA), Marrakesh, Apr. 1–4*, ed. Abderrahim Youssi, Fouzia Benjelloun, Mohamed Dahbi, and Zakia Iraqui-Sinaceur, 201–210. Rabat: AMAPATRIL.

Jastrow, Otto. 1978. *Die mesopotamisch-arabischen* qəltu-*Dialekte.* I. *Phonologie und Morphologie.* Wiesbaden: O. Harrassowitz.

Johnstone, Thomas Muir 1967. *Eastern Arabian Dialect Studies.* London: Oxford University Press.

Oppenheim, Max Freiherr von. 1939. *Die Beduinen.* I. *Die Beduinenstämme in Mesopotamien und Syrien.* Leipzig: O. Harrassowitz. (Repr., Hildesheim: G. Olms, 1983.)

Retsö, Jan. 2006. *The Arabs in Antiquity: Their history from the Assyrians to the Umayyads.* London and New York: RoutledgeCurzon.

Shahid, Irfan. 1989. *Byzantium and the Arabs in the fifth century.* Washington, D.C.: Harvard University Press.

Talay, Shabo. 1999. *Der arabische Dialekt der Khawētna.* I. *Grammatik.* Wiesbaden: O. Harrassowitz.

Peter Behnstedt (Chipiona)

Syriac → Aramaic/Syriac Loanwords

T

Taʿaddin

1. Definition

The Arabic verb *taʿaddā* and its derivatives *mutaʿaddin* and *taʿdiya* express the lexical concept of something going beyond something else. These terms are used by Arabic grammarians to refer to the concept of transitivity (for a general treatment of transitivity in Arabic grammar, see Owens 1988:167–172; Taha 1995). In the linguistic sense, the verb is said to go beyond its agent (→ *fāʿil*) to an object (→ *mafʿūl*). Verbs whose action goes beyond their agents to their direct objects are called transitive (*mutaʿaddin*); verbs whose action does not go beyond the agent to a direct object are intransitive (*ġayr mutaʿaddin* or *lāzim*). According to Sībawayhi (*Kitāb* I, 41), intransitive verbs are equivalent to transitive verbs in that both have the ability to go beyond their agent to other nominal complements in the accusative:

> Know that the verb which does not go beyond its agent [to a direct object] goes beyond it to the event noun, which is derived from it,...and it goes to the [adverbial of] time... and it goes to [the adverbial of] place (*iʿlam ʾanna l-fiʿl allaḏī lā yataʿaddā l-fāʿil yataʿaddā ʾilā ism al-ḥadaṯān allaḏī ʾuxiḏa minhu...wa-yataʿaddā ʾilā z-zamān...wa-yataʿaddā ʾilā l-makān*)

Transitive verbs can govern an accusative object directly without the help of a preposition, while intransitive verbs are unable to govern an object by itself, but only through a preposition. In the latter case, the noun is the object in meaning but not in form, because in the surface structure it is in the genitive case.

> Its form is the genitive, but its syntactic position is accusative because it is the object; therefore, coordination with this word can take place in two ways, with a genitive and an accusative, as in the expression *marartu bi-zaydin wa-ʿamran* or *ʿamrin* 'I passed Zayd and Amr' (*lafḏ̣uhu majrūr wa-mawḍiʿuhu naṣb li-ʾannahu mafʿūl, wa-li-ḏālika yajūzu fīhi l-ʿaṭf ʿalayhi wajhāni, al-jarr wa-n-naṣb naḥwa qawlika marartu bi-zaydin wa-ʿamran ʾaw ʿamrin* (Ibn Yaʿīš, *Šarḥ al-Mufaṣṣal* VII, 65)

Transitivity is a syntactic function that is not restricted to verbs: verb-like elements, such as the verbal nouns, active and passive participles, interjections, and adjectives that have the same status as active participles (→ *ṣifa mušabbaha*), are also called (in)transitive. They can all have accusative complements, if they are transitive.

2. Characteristics of (in)transitive verbs

Transitivity, according to Sībawayhi, involves a relationship between the verb and its complements. What distinguishes a verb like *ḏ̣anna* 'to believe', e.g. *ḏ̣anantu zaydan ʿamran* 'I believed Zayd to be Amr', from a verb like *ḍaraba* 'to hit', e.g. *ḍaraba zaydun ʿamran* 'Zayd hit Amr', is that if the former is not operational (*mulġā*; → *ʾilġāʾ*), the relationship between the agent and direct object reverts to the original relationship between a topic and a predicate, whereas, when the verb *ḍaraba* is omitted from the sentence, there is no relationship between the agent and the direct object to keep the sentence meaningful. The relationship is only kept intact with the inclusion of the verb in the construction (Sībawayhi, *Kitāb* I, 118–127).

Sībawayhi also introduces the concept of a verb being transitive by expansion (*ʿalā s-saʿa*) when he discusses the verb *daxala* 'to enter'. According to him, *daxala* is an intransitive verb that can only be transitive by expansion (Versteegh 1990). Normally, this verb is used with the preposition *fī* 'in', and the omission of this preposition expands the domain of the verb's government so that it now reaches the nominal complement, which becomes the object of the verb.

Hassan (1982) reviews the two linguistic tests that were commonly used by the Arabic grammarians to decide whether or not a verb is transitive. The first test is the suffixing of a pronoun to the verb referring to a noun used before the verb. If the meaning is complete, then the verb is transitive. If the result is not a complete meaning unless there is a preposition, then the verb is intransitive. He gives the following two examples to illustrate this rule: *aṣ-ṣuḥuf ʾaxaḏtuhā* 'the papers, I took them' and **al-ġurfatu qaʿadtuhā* 'the room, I sat it down'. The former sentence is meaningful, while the latter is meaningless unless it is rephrased as *al-ġurfatu qaʿadtu fīhā* 'the room, I sat down in it', i.e. using a preposition to connect the verb with the object. Therefore, the first verb, *ʾaxaḏa*, is transitive, while the second, *qaʿada*, is intransitive unless a preposition is added, which makes it transitive to a prepositional object.

The second test is to use the passive participle of the verb in question, as in *aṣ- ṣuḥuf maʾxūḏa* 'the papers are taken', which confirms that *ʾaxaḏa* is transitive, while **al-ġurfatu maqʿūda* is meaningless, unless one says *al-ġurfatu maqʿūdun fīhā* 'the room is sat in'.

The following are the most common classes of verbs cited by grammarians in describing intransitive verbs:

i. verbs referring to a permanent or recurrent characteristic in a person, such as *ḏarufa* 'to become nice', *nahima* 'to become greedy';
ii. verbs referring to a temporary state, such as *faziʿa* 'to be scared', *ḥazina* 'to be sad';
iii. verbs referring to a color or a defect, such as *ḥamira* 'to turn red', *ʿamiya* 'to become blind';
iv. verbs with one of the patterns *ifʿanlala, ifʿalla, faʿila, istafʿala, tafaʿʿala, iftaʿala*, and quadriliteral patterns.

Early grammarians such as Sībawayhi and al-Mubarrad (d. 295/898) engaged in an extensive discussion of the different verbal patterns and the meanings conveyed by each of these patterns. They established the morphological relationship between pairs of verbs, one of which was transitive while the other was not. Thus, for instance, they pointed out the analogy between the pair *kasara* 'to break [trans.]' and *inkasara* 'to break [intrans.]', on the one hand, and *ʾadxala* 'to cause someone to enter' and *daxala* 'to enter', on the other.

Within the framework of early grammatical theory, Hassan (1982:II, 150–185) summarizes the various methods mentioned by grammarians to convert verbs from intransitive to transitive. These are:

i. change from Form I of the verb to Forms II, III, IV, or X;
ii. use of a preposition to make the verb transitive;
iii. use of an intransitive verb without its preposition, so that the verb governs an accusative complement by expansion (*ʿalā s-saʿa*);
iv. change within Form I from the *faʿula* pattern to *faʿala*, as in *karuma* 'to be honored' and *karama* 'to honor' (Hassan 1982:II, 150–185); and
v. change from transitive to intransitive verb by giving them the *muṭāwaʿa* pattern of the Forms VII and VIII (→ middle), e.g. *jamaʿtuhu fa-jtamaʿa* 'I assembled it, so it became assembled (Hassan 1982:II, 233).

3. The role of the agent/doer and the patient in transitivity

The roles of the agent/doer (→ *fāʿil*) and the patient (→ *mafʿūl*) are crucial for the concept of transitivity, because there cannot be an action without a doer, and there cannot be a transitive action without a patient. Agents are the second indispensable component of a verbal sentence and together with the verb, they make the proposition syntactically complete. In discussing the verb's governance, early grammarians varied in their interpretation of the role of the agent. This is illustrated by their classification of verbs and the role they assigned to the agent in carrying out actions.

In discussing the meaning denoted by the verb, Sībawayhi employed a pattern-based approach in illustrating how (in)transitivity is connected with the pattern of verb used. Al-Mubarrad, too, employed morphological analysis of verbal forms and patterns to explain transitivity, but in addition he was the first grammarian to focus on the semantics of the verbs, which he classifies into real and non-real. Non-real verbs are those items which are similar to verbs in one or more aspects, though not in all, such as *kāna* and its sisters (→ *kāna wa-ʾaxawātuhā*), the particle *mā*, and the verbs of → exclamation. Real verbs can be either transitive or intransitive. For real transitive verbs, al-Mubarrad distinguishes between those which reach an object and affect it, such as the verb *ḍaraba* 'to hit', and those which neither reach the object nor have any effect on it, such as *ḏakara* 'to mention'. The concept of 'effect' is also implied in al-Mubarrad's argument of the intransitivity of verbs such as *šarufa* 'to become honorable', *ḍarufa* 'to become nice' (*Muqtaḍab* III, 188):

> Any verb in the pattern *faʿula* is not transitive, because it expresses the change in the status of the agent from one state to another. Therefore, there is no meaning of transitivity. This is your saying: 'Zayd became generous' and 'Abdallah became honored'. The interpretation of this is: 'he was not generous, and he became generous', and 'he was not honorable, and [now] he is honorable'. This is one kind of verb (*wa-kull mā kāna fiʿluhu ʿalā faʿula fa-ġayr mutaʿaddin, li-ʾannahu li-ntiqāl al-fāʿil ʾilā ḥāl ʿan ḥāl fa-lā maʿnā li-t-taʿaddī, wa-ḏālika qawluka: karuma zayd, wa-ḍarufa ʿabdullāh, wa-t-taqdīr: mā kāna karīman wa-laqad karuma, wa-mā kāna šarīfan wa-laqad šarufa fa-hāḏā naḥw min al-fiʿl*)

Like Sībawayhi, al-Mubarrad holds that if a genitivizer, be it a preposition or an oath particle, is omitted, the verb will 'reach' the object and cause it to have the accusative case. Al-Mubarrad uses the verb *waṣala* to refer to this process. In his explanation, he makes a distinction between the direct object, as acted upon by the verb, and other accusative complements, such as the adverbs of time and place, which refer to the time and place in which the action is performed. He adds that every verb/action happens at a time and in a place, whether it is transitive or intransitive (al-Mubarrad, *Muqtaḍab* IV, 299).

Al-Mubarrad's classification of the verbs into real and non-real reflects the way he regards the role of the doer (*fāʿil*) in terms of 'agency'. The distinction between the two terms *mutaʿaddin* and *wāṣil* in the *Muqtaḍab* shows that he made a distinction between the syntactic effect of the verb governing the agent and the accusative noun complements vs. the semantic connection between the action denoted by the verb on one hand and its doer and patient on the other.

The classification of verbs by Ibn as-Sarrāj (d. 316/928) in his *Kitāb ʾuṣūl an-naḥw* further elaborates the theory of transitivity (Taha 1995). He obviously takes semantic as well as syntactic considerations into account when he draws a distinction between the different types of verbs. Verbs are real and non-real, but then, the real transitive verbs are further classified into two categories: those which have an effect, and those which do not. The effect is explained as a physical one, inflicted on the patient of the verb by the doer. Thus, he considers verbs such as *saqaṭa* 'to fall' and *māta* 'to die' as non-real verbs, since the agent is not a doer but rather a recipient of the action. Here, agency and the role of the agent as a doer have become a determining factor in the classification of verbs.

With respect to the effect of the verb, Ibn as-Sarrāj introduces the term *mulāqin* 'encountering' to refer to transitive verbs and differentiate them from intransitive (*ġayr mulāqin*) verbs. The action of transitive verbs is described by him as follows (Ibn as-Sarrāj, *ʾUṣūl* I, 169):

> These [verbs] are of two kinds: one of them contacts something and affects it, [and another does not contact anything or affect it]. The verb that denotes an encounter is called transitive, and the verb that does not denote an encounter [is called] intransitive (*wa-lammā kānat hāḏihi takūnu ʿalā ḍarbayni: ḍarb fīhā yulāqī šayʾan wa-yuʾaṯṯiru fīhi, fa-summiya l-fiʿl al-mulāqī mutaʿadiyyan wa-mā lam yulāqi ġayr mutaʿaddin*)

Transitive verbs are subdivided into three subgroups: verbs of the senses; verbs denoting bodily movement in which contact is made with an entity other than itself; and verbs denoting reciprocal actions. He explains (*ʾUṣūl* I, 170):

> As for the transitive verb, [it is] every movement of the body that makes contact with something else, as well as what resembles this from the verbs of the psyche; the verbs of the five senses are all transitive and encountering, such as 'I looked', 'I smelled', 'I heard', 'I tasted', and 'I touched'. All the verbs that have similar meanings are also transitive.

> Moreover, with every movement of the body when it makes contact with something, the corresponding verb is considered transitive, such as ‘I reached Zayd’, ‘I stepped [in] your country, or your house’ (*wa-ʾammā l-fiʿl allaḏī yataʿaddā, fa-kull ḥaraka li-l-jism kānat mulāqiyatan li-ġayrihā wa-mā ʾašbaha ḏālika min ʾafʿāl an-nafs wa-ʾafʿāl al-ḥawāss min al-xams kulluhā mutaʿaddiya mulāqiya naḥwa naḏ̣artu wa-šammamtu wa-samiʿtu wa-ḏuqtu wa-lamastu wa-jamīʿ mā kāna fī maʿānīhinna fa-huwa mutaʿaddin wa-kaḏālika ḥarakat al-jism ʾiḏā lāqat šayʾan kāna l-fiʿl min ḏālika mutaʿaddiyan naḥwa ʾataytu zaydan wa-waṭiʾtu baladaka wa-dāraka*)

With respect to verbs with → reciprocal actions, Ibn as-Sarrāj says (*ʾUṣūl* I, 170): “Their meaning [is]: ‘I did as he did’. You made the two actions reciprocal. This reciprocity is only perceived through the [act of] encountering…” (*fa-ʾinnamā maʿnāhu faʿaltu kamā yafʿalu wa-sāwayta bayna al-fiʿlayni, wa-l-musāwā ʾinnamā tuʿlamu bi-t-talāqī*…).

Thus, Ibn as-Sarrāj’s concept of ‘encountering’ reflects the direction and extent of the action of the verb. When the verb’s action exceeds the doer externally to encounter and come into contact with a patient, the verb is considered transitive. On the other hand, when the action does not go beyond the doer to encounter a patient, the action is contained within the doer. The concept of ‘encountering’ was not taken over in later grammarians’ descriptions of the effect and nature of transitive verbs.

The verb and the agent constitute a complete structural unit (Mubarrad, *Muqtaḍab* I, 146), which can stand independently of any other units (*yaḥsunu s-sukūt ʿalayhā*) and has a communicative function (*fāʾida*). But when a transitive verb is used, one realizes that the action was not only carried out by someone, it also reached another entity. This entity is its direct object. This is why the transitive verb is defined as that whose action goes beyond the agent to reach a direct object. Al-Mubarrad distinguishes, for instance, between the verbs *ḍaraba* ‘to hit’ and *qāma* ‘to stand up’: in the former, an action goes from the agent to the patient, while in the latter, the action involves only the agent (*Muqtaḍab* IV, 335). All other accusative complements occur with both transitive and intransitive verbs and hence cannot be regarded as an essential part of the definition of either transitive or intransitive verbs.

The direct object/patient is essential for the transitive verbs, as far as Ibn as-Sarrāj is concerned. He says (*ʾUṣūl* I, 171): “These transitive verbs are not complete, nor do they exist except with the existence of the [direct] object” (*wa-lā tatimmu hāḏihi l-ʾafʿāl al-mutaʿaddiyya, wa-lā tūjadu ʾillā bi-wujūd al-mafʿūl*). Elsewhere (*ʾUṣūl* I, 412), he states that “the direct object of the verb has a share in the verb, as is the case for the agent” (*li-l-mafʿūl ḥiṣṣa min al-fiʿl kamā li-l-fāʿil*); the object is thus a structural *faḍla* that may or may not be present in the linear surface structure, but it is never a semantic *faḍla*.

The number of objects a transitive verb can take depends on the meaning and type of the verb. Generally speaking, transitive verbs take one direct object, but there are verbs that take two or three objects. With verbs of giving, for instance *ʾaʿṭaytu zaydan dirhaman* ‘I gave Zayd a dirham’, some early grammarians (followed in this by some contemporary grammarians) explain that the second object is governed by a suppressed verb. In the example given here, the second object *dirhaman* is then explained as the object of an underlying verb *ʾaxaḏa* ‘he took’, which causes the accusative in *dirhaman*. Verbs of cognition, such as *ḏ̣anna* ‘to believe’, take two objects that were originally the topic and the predicate of a nominal sentence. The role of the patient as the recipient of the action differs, according to how grammarians classify the verbs (Hassan 1982:II, 15 175).

4. On the terminology of transitivity: *mutaʿaddin* vs. *wāṣil*

The term *mutaʿaddin* fails to account for all verbs that have direct objects, because those verbs vary in the way they link the meanings denoted by them to the doer and the patient. Early grammarians therefore introduced the term *wāṣil* to refer to specific semantic relations within the larger category of transitive verbs. *Mutaʿaddin* remained, however, the general structural term for the syntactic process whereby the verb’s governance applied to more than just its agent.

The term *waṣala* and related terms were first used by Sībawayhi, and more frequently by later grammarians, to refer to the verb’s effect reaching the object beyond the agent. In many passages in the *Kitāb* (according to Troupeau 1976:217, the verb *ʾawṣala* is used eleven times, e.g. *Kitāb* I, 155), Sībawayhi refers to the process of transitivity by saying "you made the verb reach the direct object" (*ʾawṣalta l-fiʿl li-l-mafʿūl*). Other verbs are said by him to reach their object by the use of a preposition, the preposition being regarded as part of the verb (e.g. *Kitāb* I, 157). The verb *ʾawṣala*, however, was for Sībawayhi a term to refer to the verb's governance over its nominal complements in general, not specifically the direct object. Thus, it denotes transitivity in its wide sense.

The concept of *taʿaddin* is discussed in several places in the *Kitāb*, often in connection with the meaning of morphological patterns. Categories of real vs. non-real verbs are not well developed by Sībawayhi, but in later grammatical theory these categories were introduced to explain the semantic relations between the verb and the agent in (in)transitive constructions. Sībawayhi uses the term *wāṣil* 'reaching' in order to refer to the act of the verb in affecting its object. The term *wāṣil* refers to the relationship between a verb and its accusative or genitive complement (the latter in the case of verbs with prepositions). It is not used by him, however, to refer to verbs that introduce topic and predicate, such as *kāna* and *ḍanna*. Although Sībawayhi considers *kāna* to be a transitive verb, he never refers to it with the term *wāṣil*. In later grammarians, the term *wāṣil* only refers to a semantic process consisting in the verb's (or rather, the action denoted by the verb) 'reaching' first the agent, and then the patient, irrespective of whether the patient is a direct object in the accusative case or a genitive object of a preposition.

A few times, Sībawayhi uses the term *nafaḏa*/*ʾanfaḏa* 'to pass through/to make pass through' (e.g. *Kitāb* I, 204.12; see Troupeau 1976:204), in the case of verbs like *imtalaʾa* 'to be filled'. Such verbs have an object (*imtalaʾtu māʾan* 'I was filled with water'), but according to Sībawayhi, their force (*quwwa*) does not reach the object because their transitivity is less strong than that of 'real' transitive verbs (Baalbaki 2008:124).

Mutaʿaddin, on the other hand, continued to refer to the syntactic effect of transitivity. The term is employed by Sībawayhi to refer, for instance, to the oath verb (*fiʿl al-qasam*), which has a cognate accusative, and to the locative of time and place (Mosel 1975:65–70). The term *mutaʿaddin* in this sense refers to the verb's syntactic effect, passing over and beyond the agent to an accusative complement (see Levin 1979). This accusative complement may be a cognate accusative or a locative of time or place. The term *mutaʿaddin* is also used to refer to the operation of verb-like elements such as interjections (→ *ism al-fiʿl*). These interjections perform the role of imperative verbs and 'go beyond' the person commanded (*manhiyy*) to an action or thing that is prohibited or banned (*manhiyy ʿanhu*). Examples of usages that 'go beyond' the person commanded are *ruwayda zaydan* 'take it easy, Zayd!' and *hālumma zaydan* 'hurry up, Zayd!'. Examples of usages that do not go beyond the *manhiyy* include simple commands such as *mah* 'quit!' and *ṣah* 'shut up!' (*Kitāb* I, 241–242; Mubarrad, *Muqtaḍab* III, 202–205).

Like Sībawayhi, al-Mubarrad emphasizes the meaning of the morphological patterns of the verbs as an indication of verbal (in)transitivity (*Muqtaḍab* II, 102, 109, 125–126). He speaks of relationships such as compliance in the *ʾafʿāl al-muṭāwaʿa*, which are important in determining the transitivity of a certain verb. In this connection, he discusses pairs of verbs such as *kasara fa-nkasara* 'he broke [it] and it broke'. In addition, he specifies specific patterns of form, in which one verb is transitive while others are intransitive.

As discussed above, al-Mubarrad introduced a division into real vs. non-real verbs. Within this new division, the term *wāṣil* came to be used for the semantic bond between the (action denoted by the) verb and its agent and patient. The term *mutaʿaddin* continued to refer to the syntactic governance of verbs. The term *taʿaddā* continued to be used by al-Mubarrad to cover the accusative case in nouns other than the direct object, such as the cognate accusative (*mafʿūl muṭlaq*), adverbials of time and place (→ *mafʿūl fīhi*), and circumstantials (→ *ḥāl*), while *wāṣil* referred only to the direct object. This new semantic terminology was elaborated in the work of al-Mubarrad's successor Ibn as-Sarrāj.

Not all transitive verbs with an accusative object are to be considered *wāṣila* verbs, as far as al-Mubarrad is concerned. Verbs such as *ḍanna* and *ʿalima* are not *wāṣila,* while *ḍaraba* is. He explains that in the former one does not see an effect reaching the direct object, while in the latter the effect is clear (*Muqtaḍab* IV, 403). The two verbs *ḏakara* 'to mention' and *šatama* 'to vilify' are used by al-Mubarrad as examples of transitive but not *wāṣila* verbs. This classification represents the apparent confusion and overlapping which characterizes the classification of the verbs in *al-Muqtaḍab* in general. The problem of classifying the two verbs as transitive but not 'reaching' lies in the fact that they, like *ḍaraba* 'to hit' and *qatala* 'to kill', which are both *wāṣila*, have direct objects. The direct object, which constitutes a syntactic reflection of the patient, is involved in the same action as the doer. The difference is that in those *wāṣila* verbs which have an effect on their patient, such as hitting and killing, one actually sees the physical effect on the patient. By way of contrast, in the case of the verbs *ḏakara* and *šatama*, there is no visible physical effect on the patient.

The semantic analysis of verbal acts is almost entirely absent in Sībawayhi's *Kitāb*. For Sībawayhi, both *ḏakara* and *ḍaraba* are transitive verbs whose syntactic effect applies to their direct objects; therefore, they are of the same type, as far as transitivity is concerned. In al-Mubarrad's work, the meaning denoted by the verb begins to play an important role in the classification. But he equates the concept of *wāṣil* with the concept of physical effect and fails to distinguish between *wāṣila* verbs with and without physical effect on the patients. This subdivision was introduced by later grammarians, such as Ibn as-Sarrāj, who classified verbs like *ḏakara* as *wāṣil*, but not *muʾaṯṯir*.

Ibn as-Sarrāj continues to use the term *mutaᶜaddin* in the same way as his predecessors, to refer to both transitive and intransitive verbs with respect to their nominal complements, while consistently using *wāṣil* to refer to verbs with direct objects and to the occurrence of other nominal compliments with verbs. Although he sometimes refers to transitive verbs as *mutaᶜaddin*, strictly referring to the syntactic role they play with the nominal complements, he confines the use of *wāṣil* to those verbs which have direct objects. The category of real verbs in Ibn as-Sarrāj consists of verbs whose agents cause an action to take place. Thus, the verbs *māta* 'to die' and *saqaṭa* 'to fall' are non-real verbs, since the agents are not actually doers of the action denoted by the dying and falling.

5. CONCLUSION

Transitivity as a syntactic process was important for the Arabic grammarians, since it allowed them to discuss case inflections and predication in the verbal sentence. The general term used by Arabic grammarians to refer to transitivity was the term *taᶜaddin*. However, after Sībawayhi, early grammarians started to look into the semantico-syntactic relations that transitive verbs had with their nominal complements. Hence, terms such as *waṣala* and *wāṣil* started to be used to refer to the process of linking the verb's denotations to other semantic entities, such as the doer, the experiencer, the beneficiary, the recipient, and the patient. The most striking difference between the 10th-century grammarian Ibn as-Sarrāj and Sībawayhi is that the morphological aspect of determining transitivity is almost absent in Ibn as-Sarrāj's treatment of the issue. His analysis was mostly, if not entirely at times, based on the meaning denoted by the verb itself.

The shift from noun-centrism to verb-centrism, from Sībawayhi's first reference to transitivity as *fiᶜl fāᶜil* 'agent's action' and *fiᶜl mafᶜūl* 'patient's action' to the classification of types of actions and effects according to the verb's denotation, inaugurated a different approach to the description of grammatical processes. It seems as though the description of Arabic grammar has gone full circle, ending with contemporary grammarians attempting to explain rules of grammar with minimal reliance on semantic categories. The term *mutaᶜaddin* has remained the technical term for the notion 'transitive', whether the verb is transitive to one, two, or three objects, and whether it is a verb that introduces a topic and predicate, a verb of giving, or a regular verb with agent and direct object. Whenever the term *waṣala* is used, it is used concurrently with *mutaᶜaddin* to describe the act of going beyond the agent to the direct object. The term *wāṣil* has disappeared from technical grammatical terminology.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Ibn as-Sarrāj, *ʾUṣūl* = ʾAbū Bakr ibn as-Sarī Ibn as-Sarrāj, *Kitāb al-ʾuṣūl fī n-naḥw*. Ed. ᶜAbd al-Ḥusayn al-Fatlī. 3 vols. Beirut: Muʾassasat ar-Risāla, 1985.

Ibn Yaʿīš, *Šarḥ* = Muwaffaq ad-Dīn Yaʿīš Ibn Yaʿīš, *Šarḥ al-Mufaṣṣal*. 10 vols. in 5. Cairo: ʾIdārat aṭ-Ṭibāʿa al-Munīriyya, n.d.

Mubarrad, *Muqtaḍab* = ʾAbū l-ʿAbbās Muḥammad ibn Yazīd al-Mubarrad, *al-Muqtaḍab*. Ed. Muḥammad ʿAbd al-Xāliq ʿUḍayma. 4 vols. Cairo: Dār at-Taḥrīr, 1965–1968.

Sībawayhi, *Kitāb* = Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. Ed. ʿAbd as-Salām Muḥammad Hārūn. 4 vols. Cairo: al-Hayʾa al-Miṣriyya li-l-Kitāb, 1977.

SECONDARY SOURCES

Baalbaki, Ramzi. 2008. *The legacy of the* Kitāb: *Sībawayhi's analytical methods within the context of the Arabic grammatical theory*. Leiden: E.J. Brill.

Hassan, ʿAbbās. 1982. *an-Naḥw al-wāfī*. 2 vols. Cairo: Dār al-Maʿārif.

Levin, Aryeh. 1979. "The meaning of *taʿaddā l-fiʿl ilā* in Sībawayhi's *Kitāb*". *Studia orientalia D.H. Baneth dedicata*, 193–210. Jerusalem: Magnes Press.

Mosel, Ulrike. 1975. *Die syntaktische Terminologie bei Sībawaih*. Ph.D. diss., University of Munich.

Owens, Jonathan. 1988. *The foundations of grammar: An introduction to medieval Arabic grammatical theory*. Amsterdam and Philadelphia: J. Benjamins.

Taha, Zeinab A. 1995. *Issues of syntax and semantics: A comparative study of Sibawayhi, al-Mubarrad, and Ibn as-Sarraaj*. Ph.D. diss., Georgetown University.

Troupeau, Gérard. 1976. *Lexique-index du* Kitāb *de Sībawayhi*. Paris: Klincksieck.

Versteegh, Kees. 1990. "Freedom of the speaker? *Ittisāʿ* and related concepts in Arabic grammar". *Studies in the history of Arabic grammar*, II, ed. Michael G. Carter and Kees Versteegh, 281–293. Amsterdam: J. Benjamins.

ZEINAB TAHA
(The American University in Cairo)

Taboo

1. INTRODUCTION

Originally from the Tongan language, which belongs to the Polynesian group of languages, the word 'taboo' was first introduced into English in 1777 by the naval explorer Captain James Cook (1728–1779), and later entered other European languages. The concept refers to anything sacred and inviolable and later, by extension, to anything that is in contradiction to moral standards or good manners and is, therefore, forbidden or prohibited. 'Taboo language' refers to words and expressions that may cause embarrassment or offense and should be 'mitigated' or avoided altogether. Thus, names, words, and concepts that are sacred, or that denote profanity or blasphemy, may upset a religious person, in the same way that obscenity, swearing, and coarse and vulgar language may hurt a sensitive or well-bred person.

Taboo in Arabic is mainly used in reference to certain domains, such as religion, politics, honor, females, sex, body organs and bodily functions, unclean places, animals, unclean objects, the evil eye, illnesses, and death, but it may also refer to swear words, expletives, and vulgarisms at large (→ dysphemisms). Hence, social norms require the use of additions, explanations, circumlocutions, or → euphemisms in the case of the first type of taboo, and their total avoidance in the case of the second type, since their use is *ʿayb* 'disgrace'.

2. TABOOS IN VARIOUS DOMAINS

i. Religion
When referring to the Prophet Muḥammad by name or by his common epithets *ar-rasūl/rasūl allāh/an-nabī* 'the Messenger/the Messenger of God/the Prophet', it is customary to add the formula *ṣallā llāhu ʿalayhi wa-sallam* 'God bless him and grant him salvation'. Using the Prophet's name alone may upset an orthodox Muslim and will be regarded as bad manners. In the same way, naming any other book *al-Qurʾān* will be regarded as blasphemy, although *al-Kitāb* 'The Book' may refer to the *Qurʾān* as well as to the famous grammar book written by Sībawayhi (d. ca. 796 C.E.).

ii. Politics
Although in a democratic society the right to criticize the government is taken for granted, voicing reservations about a ruler or a regime or its policies may still be too dangerous for individuals in some countries around the Middle East, and therefore circumlocutions are used instead. *Tierfabeln* were occasionally used as indirect reference to the relationship between ruler and subjects, as in the famous work by ʿAbdallāh Ibn al-Muqaffaʿ (d. ca. 760 C.E.) entitled *Kalīla wa-Dimna* (based on the Sanskrit *Fables of Bidpai*), which is assumed to have used animals to represent human beings, especially in reference to despotic rulers.

iii. Honor

Arabic terms for this concept include *ʿirḍ, šaraf, karāma, waqār, iḥtirām* (the last word shares the root with the words *ḥaram/ḥarām* ‘forbidden; holy; taboo’, *ḥurma* ‘inviolable; woman’, *ḥarīm* ‘holy place; harem; women’, *ʾiḥrām* ‘state of ritual consecration of the pilgrim to Mecca’). Any conduct that may be interpreted as an insult to the individual or to one’s honor or the honor of the family or the tribe is regarded as taboo.

iv. Females

In some Muslim societies where a strict code of conduct concerning females exists, nonmembers of the family often use various words and expressions when referring to females, rather than their names or the ordinary words for ‘woman, wife’ (Arabic *imraʾa, zawja*). Thus, it is customary to use the general words *al-ʾahl* ‘the family’ or *ʾumm al-ʾawlād* ‘the mother of the children’, when referring to someone else’s wife, while the words *al-wālida* ‘the one who has given birth [i.e. the mother]’ or *al-karīma* ‘the respected [i.e. the daughter]’ are used in reference to other females in the family. Also, the word *al-ʾuxt/ʾuxtī* lit. ‘the sister/my sister’ may be used as a respectful reference to a woman. The common word for a woman in literary Arabic (*imraʾa*, colloquial Arabic *mara*) may have a negative connotation in spoken Egyptian Arabic and should therefore be avoided, and replaced by the word *zōga* (→ politeness).

v. Sex

Direct reference to sex, sexual intercourse, or sexual practices regarded as perversion (e.g. homosexuality or pedophilia) is a taboo, although equivalent words and expressions concerning sexual activities are found in Arabic, often with scores of synonyms. Instead, euphemisms and circumlocutions are used, e.g. *duxūl* ‘first coitus in marriage [lit. ‘entering’]’ (hence *laylat ad-duxla* ‘wedding night’); *jimāʿ* ‘intercourse’ (from the root *j-m-ʿ* ‘to gather’). The following words are considered taboo: *šarmūṭa, qaḥba, zāniya, mūmis, baġīy, ʿāhira, bint al-hawā, bint aš-šāriʿ* ‘prostitute’; *qawwād, ʿarṣ, dayyūṯ* ‘pimp’; *ibn ḥarām, bundūq, naġl* ‘illegitimate child, bastard’; *māxūr, bayt ad-diʾāra, karaxāna* ‘brothel’; *šabaq* ‘lust, sexual desire’; *nayk* ‘sexual intercourse’; *dalama* (colloquial Egyptian) ‘clitoris; orgasm’; *musāḥaqa* ‘lesbianism’; *liwāṭ(a)* ‘homosexuality’. The reason for the disappearance of the Classical Arabic word *miḏāʾ* ‘orgy’ from present-day Arabic may be that it is regarded as a taboo word.

vi. Body organs and bodily functions

Certain members of the body and bodily functions and discharges are considered taboo and are often expressed by circumlocutions or replaced by euphemisms. Thus, *al-ʾaʿḍāʾ at-tanāsuliyya* ‘sexual organs, genitals [lit. ‘reproduction members’]’ replaces such words as *farj* ‘vulva’; *ʾayr, ḏakar, qaḍīb, ʾihlīl* ‘penis’. Other words, usually in the dialect, which are regarded as taboo are *zubb*, *zubr* ‘penis’; *kuss* (which is also used in several strong swear words) ‘vulva’; *ṭīz* ‘buttocks, bottom’. The verbs *ḍaraṭa* ‘to break wind noisily’ and *fasā* ‘to break wind noiselessly’ may be replaced by *ʾaxraja rīḥ karih(a)* lit. ‘to let off a bad smell’. Other taboo words include *xara* ‘shit’, which may also be used as a derogatory reference to a human being, or to bad conditions, situations, or objects of bad quality; *šaxx* or *šaxāx* ‘urine, piss’, which is less common then *bawl*; and the two onomatopoeic words *ʾuff* ‘dirt in the ears, earwax’ and *tuff* ‘dirt under the nails’. *Istimnāʿ* ‘masturbation, onanism’ is often replaced by the euphemism *al-ʿāda as-sirriyya* lit. ‘the secret practice’ or *al-maʿrūfa* lit. ‘the known one’.

vii. Uncleanliness

Places known to be unclean or nonhygienic are also regarded as taboos, and, therefore, they are usually replaced by euphemisms. For *šišma* ‘toilet, loo’ (< Turkish *çişme*), for instance, the euphemisms *ḥammām* ‘bathroom’, *mirḥāḍ* ‘washing place, lavatory’, *bayt al-ʾadab* lit. ‘the house of good manners’, *bayt ar-rāḥa* lit. ‘the house of rest’ (cf. American *restroom*), *bayt al-xalāʾ* lit. ‘outdoor house’, *bayt al-māʾ* lit. ‘water house’, *mustarāḥ* ‘resting place’, *ʾadabxāna* lit. ‘the room of good manners’ (a Persian/Arabic compound), and, rarely, *kanīf* and a few more are used. The word *bālūʿa* ‘sewer, drain’ may be used derogatorily for someone’s mouth (= ‘gob’; see also viii, ix below).

viii. Animals

Several animals are considered taboo because they are unclean (*nijs*) or stupid (*ʾaḥmaq*) or bring bad luck (*naḥs*). These include the pig (*xinzīr*); the dog (*kalb*); the donkey (*ḥimār*); the hyena (*ḍabʿ*); the raven, the crow (*ġurāb*); and the owl (*būma*). In the case of a pig, the derogatory epithet *al-xabīṯ* 'the repulsive' is preferred, while in reference to a dog or a donkey the parenthetical remarks *baʿīd ʿannak* lit. 'far [be it] from you' or *ʾajallak* 'you are too respected [to be associated with such an animal]' are usually added when addressing someone.

ix. Objects

Shoes, in Arabic *ḥiḏāʾ, naʿl, kundara*, and in particular *jazma* and *ṣurmāya*, are regarded as unclean objects because they touch the ground and therefore are considered taboo. Hence, reference to them requires the expressions of reservation *baʿīd ʿannak* or *ʾajallak* to be added to the sentence.

x. Evil eye

The belief in, and therefore the fear of, the evil eye (*al-ʿayn* or *al-ʿayn aš-širrīra*) is characteristic of Arab culture. Being a taboo, the evil eye, like a spell, may affect children and adults, property, and, as a matter of fact, anything which may cause jealousy or envy of another person. There are therefore scores of words and expressions used against the evil eye and envious people or, alternatively, used to 'cure' those who were smitten by it. Hence, the popular formulas *mā šāʾa llāh* (*māšallāh*) 'God willing' or (*i*)*sm allāh ʿalēk/h* 'God's name on you/him', etc. or *xamsa* lit. 'five', i.e. the open hand, which will stop the evil eye from harming, are normally added when expressing surprise and amazement at something, or when referring to an achievement by a member of the family or a close friend, or in reference to a large quantity, size, number, amount, and the like. Examples of their use are *kam walad ʿandak, ma šāʾa llāh?* 'how many children do you have, God willing?', or *mabrūk ʿalā l-bēt al-jadīd, mā šāʾa llāh* 'congratulations on your new house, God willing'. This expression may also be used in reaction to a statement made by another speaker; for instance, when someone asks *kam walad ʿandak?* 'how many children do you have?' and the other answers *talāta* 'three', the appropriate reaction is *mā šāʾa llāh!* 'God willing!'.

xi. Illnesses

Reservations regarding illnesses, diseases, and disablement, on the one hand, and optimism and hopes for a speedy recovery, on the other, have given birth to numerous circumlocutions as well as scores of expressions of well-wishing. Venereal diseases are sometimes referred to as *al-ʾamrāḍ as-sirriyya* lit. 'the secret illnesses', while the word *šarr* 'evil' is often used in reference to all kind of illnesses, e.g. *rāḥ iššarr* 'may the illness be gone!'.

xii. Death

Although 'passing away' is believed to be equal to 'moving to a better place', direct reference to the concept is a taboo. Hence, death and words associated with it are often replaced by circumlocutions or euphemisms. The verb *māta* 'he died', in reference to a human being, is often replaced by the verb *tuwuffiya*, which is the passive form of the verb *tawaffā*, referring to God as the One who has taken the deceased. The word *mayyit* 'dead person' is hardly ever used and is often replaced by other words or expressions and idioms, such as *al-marḥūm* lit. 'the one who has received mercy', *ar-rāḥil* lit. 'the one who has traveled', *al-maġfūr lahu* lit. 'the one who has been pardoned', *sākin al-jinān* lit. 'the dweller of Paradise'; *labba nidāʾ rabbihi* lit. 'he responded to the call of his Master', *qaḍā naḥbahu* lit. 'he completed his term', or in Classical Arabic *qubiḍa* lit. 'he was taken away'.

3. Swear words, expletives, and vulgarisms

The second type of taboo may be divided into two main categories:

i. Grossness

Gross language denotes the use of rude style, often involving intonation or raising the voice and gestures, when communicating with people, showing anger, contempt, or superiority. This behavior runs counter to

→ politeness. It is regarded as impolite to order someone to do something by using the imperative, especially in a tone of command, e.g. *ijlis, uqʿud* 'sit down!' instead of *ijlis, min faḍlak* 'sit down, please!' or *istrīḥ* 'have a rest [i.e. sit down]!'; or to ask someone to leave the place, e.g. in Classical Arabic *uxruj, inṣarif* 'get out!, go away!', or in spoken Arabic *imši, iṭlaʿ, inqiliʿ* lit. 'be uprooted!', *txayyab* lit. 'disappear!', *yaḷḷāh rūḥ min hōn* 'go away!', *farjinā ẓahrak* lit. 'show us your back!', *farjinā ʿurḍ ktāfak* lit. 'show us the width of your shoulders!', i.e. 'shove off!'. It is also impolite to order someone to keep quiet, in Classical Arabic with *ṣah*, *uskut* and in dialect with *ixras, sakkir tummak* lit. 'shut your mouth'/*būzak* lit. '[shut] your muzzle, snout', *sudd ḥalqak* lit. 'block up your throat!'/*nīʿak* 'shut up!'. Drawing comparisons or using similes or metaphors can be very impolite, e.g. calling someone *šayṭān* 'devil', or saying things like *mā tibqāš kurdi* 'don't be a fool [lit. 'don't be a Kurd!']', *ḥissak ḥiss ilmotor* 'your voice sounds like an engine', *byākol mitl/zayy ilḥayawān* 'he eats like a beast'. Note, however, that comparisons with animals are drawn already in the *Qurʾān*, for instance when the Jews are compared with donkeys (Q. 62/6) or monkeys (Q. 2/65, 7/166), and when both Jews and Christians are said to have turned into monkeys and pigs (Q. 5/59)]. Some → proverbs are impolite when applied to a person, for instance *kalb iššēx šēx ilkilāb* 'the sheikh's dog is the dogs' sheikh', i.e. this person is a parvenu, or *min qillat ilxēl šaddu ilkilāb* 'because of the shortage of horses, they harnessed the dogs', which is used as a criticism when an unsuitable person has been appointed to a post, or *iṭṭabl ilʾajwaf ṣōto ʿāli* 'the empty drum has a loud sound', which means that someone is a braggart.

ii. Swearing and cursing

There are a vast number of swear words, 'four-letter words', 'dirty words', and curses in Arabic (*sabb, šatm, qaḏʿ, muhātara, duʿāʾ*); they are usually used verbally, and, therefore, most of them are dialectal (→ insults). Swear words are used far less frequently in Classical Arabic than in colloquial Arabic. However, as one of the uses of the perfect tense (and sometimes the imperfect as well) is to express good wishes or curses, the number of possibilities is in fact unlimited, e.g. *qatalahu llāh* 'may God kill him!', *laʿanaka llāh/laʿnat allāh ʿalayka* 'may God curse you!', *qabbaḥaka llāh* 'may God make you look ugly!', *xasafa llāh bihi l-ʾarḍ* 'may God make him sink into the ground!'. The jussive, especially with *fa-* + *la-* as prefixes, is also used in swearing, e.g. *fa-l-yaḏhab fī dāhiya* 'let him go to hell!'. Another popular pattern is an elliptical sentence from which the verb and subject/agent have been omitted, e.g. *tabban lahu* 'may evil befall him!, may he perish!', *jaḏʿan laka* 'may your nose be cut off!', *buʿdan/suḥqan lahu* 'to hell with him!', *tuffan laka* 'fie on you!', *al-wajaʿ bi-kabidika* 'may pain strike your liver!'.

Sociologically speaking, males of all layers of society, from children and teenagers to adults, are apt to use swear words more often than females, while females who belong to the middle and upper class will avoid using them, at least in public. So far as use of expressions by adults but not by children and vice versa is concerned, one may assume that certain expressions involving concepts such as 'honor' or 'religion' will only be used among adults but not by small children, though this is an area that would require more research. Moreover, and as a general remark, as the live material on the subject of taboo is so scarce, the information gathered comes mainly from dictionaries (e.g. Hinds and Badawi 1986) and from speakers of Arabic in various Arab countries. In the following categories of taboo words, most of the examples are taken from Egyptian and Levantine Arabic.

i. God is called upon to afflict the cursed person with calamity, e.g. *yimḥi ismak* 'may [God] strike off your name!', *allāh yāxdak* 'may God take you!', *allāh yiḥraqak* 'may God burn you!'.

ii. Harm is sought to be caused to the cursed person, e.g. *trāb yixabbik* 'may the soil cover you!', *dāhiya tāxdak/jahannam yāxdak* 'go to hell [lit. 'may hell take you']!', *ʿifrīt yāxdak* 'may the demon/imp take you!', *ḥabbak ilbalāʾ* lit. 'may you be loved by disaster!'.

iii. Disasters will happen without referring directly to God as their cause, e.g. *yixrab* *bētak* ‘may your house be ruined!’, *ḍarba fī qalbak* ‘may your heart be smitten!’, *yuqṭaʿ ʿumrak* ‘may your life be shortened!’.

iv. A large number of swear words begin with the exclamatory particle *yā,* denoting a direct approach to the cursed person, e.g. *yā malʿūn* ‘you cursed one!’, *ya-bn iššarmūṭa* ‘you son of a whore!’, *ya-bn kalb* ‘you son of a dog’.

v. Many imprecations and swear words involve taboos of the first type:

  a. Curses involving religion, e.g. *yilʿan dīnak* ‘may your religion be cursed!’, *yilʿan dīn ʾabūk wi-ʾabū ʾabūk* ‘may your father’s and grandfather’s religion be cursed!’, *yiḥreq dīnak* ‘may your religion burn!’, *yilʿan rabbak* ‘may your god be cursed!’, *ʾuzraʾīl* (in some areas: *ʾuzrayīn*) *yāxdo* ‘may the Angel of Death take him!’, *rūḥ liljahannam* ‘go to hell!’.

  b. Politics. Recorded political swear words are relatively few and are usually coined according to the imagination of the speaker or the writer. An example is the derogatory name given in the 1960s by the media of some hostile Arab countries to the late King Hussein of Jordan: *al-malik at-tranzistor* lit. ‘the transistor king’ or *al-mulayk* (a diminutive form of *malik*) ‘the small king’.

  c. Family, e.g. *yilʿan/yinʿal ʾabūk* ‘may your father be cursed!’, *yilʿan/yinʿal ʾabūk wi-ʾabū ʾabūk* ‘may your father and his father be cursed!’, *yilʿan ʾabu ddāye illi jābatak* ‘may the father of the midwife who delivered you be damned!’, *dāhiya fī ʾummak* ‘may misfortune overtake your mother!’.

  d. Honor, e.g. *ibn ḥarām* ‘bastard’, *ibn iššarmūṭa/walad ilqaḥba* lit. ‘son of a whore’.

  e. Females. Except for the various words for ‘ prostitute’ mentioned above under ‘sex’, which are taboo, some of the swear words used for males may apply to females as well, e.g. *bint ilḥarām* ‘bitch’, *bint iššarmūṭa* ‘the daughter of a whore’, *xabīṭa* ‘wicked’, *šaršūḥa* ‘foul-mouthed; low-class’, *ḥayzabūn* ‘old hag’.

  f. Sex, e.g. *manyūk* ‘one on whom sex has been performed’; *zubbi fī ṭīzak* ‘my penis is inside your rectum’; *ya-bn* (*ilmaṛa*) *ilmitnāka/ilqaḥba* ‘you son of a whore!’; *ʿarṣ/qawwād* (often pronounced *gawwād*) ‘a pimp’ (in certain contexts the words may mean ‘shrewd, smart, clever’); *kuss ʾummak /ʾuxtak/mart ʾabūk* ‘the vulva of your mother/sister/your father’s wife’, probably to be interpreted as a threat to have sex with the person named. By a peculiar case of → semantic bleaching, the word *kuss* is sometimes combined with other substantives involving the honor or even the religion of the person who is insulted, e.g. *kuss ʿarḍak* ‘your honor is worthless [lit. ‘(I penetrate) the vulva of your honor’]’.

  g. Body parts and bodily functions, e.g. *šxāx fī wuššak* ‘urine on your face’, *ilḥas ṭīzi/zubbi* ‘lick my rectum/penis’, *muxxak taʿbān* ‘you are an idiot [lit. ‘your brain is tired’]’.

  h. Unclean, e.g. *wisix*, *muximm* ‘dirty’, *laʾīm* ‘filthy’, *xabīṯ* ‘disgusting’; *lisāno zifr* ‘he is foul-mouthed’.

  i. Animals, e.g. *ḥumār* ~ *ḥmār/jaḥš* ‘donkey/young donkey’, i.e. ‘idiot, bloody fool’; *kalb/ibn kalb* ‘son of a bitch’; *tēs* ‘stupid, idiot [lit. ‘billy goat’]’; *waḥš* ‘a wild animal’; *dubb* ‘clumsy [lit. ‘bear’]’; *ḥayya* lit. ‘snake’; *ʿaqrab* ‘harmful person, snake in the grass [lit. ‘scorpion’]’.

  j. Objects, e.g. *barmīl* ‘fat and dumb [lit. ‘barrel’]’, *ġardal* ‘dolt, sucker [lit. ‘bucket’]’, *ʾinta taḥt ṣurmāyti* ‘you are nothing to me [lit. ‘you are under my shoe’]’, *rāgil zift iṭṭīn* ‘bastard [lit. ‘a man of the worst mud’]’.

  k. Evil eye, e.g. *tiṭlaʿ ʿēnak/ʿyūnak* ‘may your eye/s be gouged’, *ḥaswāya fī ʿēn ilḥassāda/ʿēn ilḥasūd fīha ʿūd* ‘may the evil eye be struck blind’.

  l. Illness, e.g. *majnūn* ‘mad’, *xarfān* ‘senile’, *ʿamā fī qalbak/ʿēnek* ‘may you be struck with blindness in your heart/eyes!’, *mā fīš fī wuššak dam* ‘you have no shame [lit. ‘you don’t have blood in your face’]’, *yinhirū maṣārīnak* ‘may your bowels disintegrate’, *mišyet ʿalēk baṭnak* ‘may you suffer diarrhea’, *ʿamā* (as an interjection) ‘hell [lit. ‘blindness’]!’. Bodily defects and disablement may also be used in a derogatory manner, e.g. *yā ʾaṭraš* ‘you, deaf one!’, *yā ʾaʿwar* ‘you, one-eyed one!’, *yā ʾaʿraj* ‘you, lame one!’.

m. Death, e.g. *ruḥ fī* (*sittīn/ʾalf*) *dāhiya* ‘go to (sixty/thousand) calamities!’, *trāb yixabbik* ‘may you be buried [lit. ‘may sand cover you’]!’, *tiṭlaʿ rōḥak* ‘may your soul leave [your body]’.

Nouns and attributes denoting negative characteristics may also be considered taboo if used in addressing human beings, e.g. *ḥarāmi*, *sarrāq* ‘thief’; *kaḏḏāb/ġaššāš* ‘liar’; *naṣṣāb/makkār* ‘crook’; *laʾīm* ‘ignoble’; *ʾazʿar/ʿakrūt* ‘scoundrel’; *ʾahbal, ʾablah, ʾawrah, ġabīy, ʿābiṯ, ʾaḥmaq, balīd, miġaffal* ‘stupid, idiot’; *saxīf* ‘foolish [something said or done]’; *qalīl al-ʾadab* ‘impolite’; *qalīl al-ḥayāʾ* ‘shameless’; *danīy al-ʾaxlāq* ‘immoral’; *jāhil* ‘ignorant’; *fallāḥ* ‘peasant, uneducated’; *dallūʿa/midallaʿ* ‘spoiled child; [teacher’s/boss’s] pet’.

Finally, some onomatopoeic → interjections are also used as swear words, e.g. *ṭuzz fīk* ‘the hell with you’; *ʾixṣ ʿalēk/itfū ʿalēk wi-ʿalā šaklak* ‘you are disgusting; shame on you’; *ʾaḥḥa yā ʿumar* ‘fuck that!’ (where *ʾaḥḥ* is an exclamation expressing extreme disapproval but also an exclamation of women’s pleasure during sexual intercourse); *ʾaxx/ʾixx* ‘ugh, shame, disgusting!’ (→ sound symbolism).

Bibliographical references

Chebel, Malek. 2006. *Le kama-sutra arabe*. Paris: Pauvert.

Goldenson, Robert and Kenneth Anderson. 1987. *Dictionary of sex*. London: Wordsworth-Bloomsbury.

Graves, Robert. 1929. *The future of swearing and improper language*. London: Kegan Paul, Trench.

Hinds, Martin and Elsaid Badawi. 1986. *A dictionary of Egyptian Arabic*. Beirut: Librairie du Liban.

Holder, Robert W. 1995. *A dictionary of euphemisms*. Oxford.

Hughes, Geoffrey. 1998. *Swearing: A social history of foul language, oaths and profanity in English*. London: Penguin Books.

McDonald, James. 1988. *A dictionary of obscenity, taboo and euphemism*. London: Sphere.

Montagu, Ashley. 1967. *The anatomy of swearing*. London: Rapp and Witing.

Morton, James. 1989. *Low speak*. London: Angus and Robertson.

——. 2005. *Dirty words*. London: Atlantic Books.

Pearce, Dan. 2003. *Talking dirty*. London: Cassell.

Piamenta, Moshe. 1983. *The Muslim conception of God and human welfare*. Leiden: E.J. Brill.

Wehr, Hans. 1979. *A dictionary of Modern Written Arabic*. Wiesbaden: O. Harrassowitz.

Avihai Shivtiel (University of Leeds)

Tafxīm

The term *tafxīm* is derived from the Arabic triliteral root *f-x-m*, generally signifying ‘thickening, magnifying, enlargement, emphasizing’. This notion was applied to the Arabic sound system to differentiate between certain sound groups. One class includes the four *muṭbaqa* consonants (→ *ʾiṭbāq*): /ḍ/ ض, /ṣ/ ص, /ṭ/ ط, and /ḏ̣/ ظ. In addition, three more consonants, /q/ ق, /ġ/ غ, and /x/ خ, are grouped with these four consonants to form a larger class, known as *mustaʿliya* ‘elevated, raised [consonants]’. They are called thus because of “the raising of the back of the tongue towards the soft palate when they are pronounced” (Ibn Jinnī, *Sirr* I, 62). These seven consonants share this particular characteristic, and, additionally, they all disallow the occurrence of → *ʾimāla*. In contrast, all the other consonants are termed *mustafila* or *munxafiḍa* ‘depressed, lowered’, because the back of the tongue is not raised when they are articulated. Ibn Jinnī also makes a distinction between the *muṭbaqa* consonants and the other three by saying: “The former four are *mustaʿliya* besides being velarized, whereas the latter three are *mustaʿliya* only, but not velarized” (*Sirr* I, 62). This clearly indicates that early phoneticians recognized that velarization is the result of a secondary articulation, while the raising of the back of the tongue in pronouncing /q/, /ġ/, and /x/ is a primary articulation, by nature (cf. Bakalla 1982:144). Both /ṭ/ and /q/ are grouped with /b/, /j/, and /d/, constituting the *qalqala* class (Sībawayhi, *Kitāb* IV, 174). Commenting on this feature, Al-Nassir states: “It is clear that Sibawayh is dealing with the release phase of producing plosive consonants.... The ‘small sound’ produced when pausing on the plosives concerned is termed a *ṣuwayt* (diminutive form of *ṣawt*). This small sound is, as attested, a brief vowel sound, similar to the English Shewa” (1993:52).

The term *mufaxxama* or *tafxīm* is used more frequently than *mustaʿliya* in treatises on Qurʾānic recitation (→ *tajwīd*). The contrast between *mustaʿliya* or *mufaxxama* and *mustafila* or *munxafiḍa* is rendered by many modern phoneticians as emphatic vs. nonemphatic speech sounds. In their analysis, *tafxīm* or emphasis may involve more processes than

Figure 1. Features of emphatic vs. nonemphatic consonants

FEATURES	q	x	ġ	ṣ	ḍ	ṭ	ḏ̣	ʾ	h	ʿ	ḥ	k	š	y	j	s	z	d	t	l	r	n	ḏ	ṯ	f	b	w	m
consonantal	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
vocalic	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
voiced	(+)	-	+	-	+	(+)	+	o	-	+	-	-	-	+	+	-	+	+	-	+	+	+	+	-	-	+	+	+
voiceless	-	+	-	+	-	-	-	o	-	-	+	+	+	-	-	+	-	-	+	-	-	-	-	+	+	-	-	-
muṭbaq	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ġayr muṭbaq	+	+	+	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
mustaʿlī	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ġayr mustaʿlī	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

KEY to symbols: + = presence of a feature - = absence of a feature o = neutral

just velarization. It may include pharyngealization together with some degree of labialization and/or glottalization. In his monumental paper on the Arabic *mufaxxama* phonemes, Jakobson (1962:511ff.) states:

> The characteristic articulatory feature of all the 'emphatic' phonemes is the contraction of the upper pharynx. Native informants usually point to their throat to elucidate the 'emphatic' articulation...usually the production of pharyngealized buccal phonemes is accompanied by a velarization....Furthermore, there is a tendency to emit the pharyngealized phonemes with a protrusion and slight rounding of the lips....Whatever orifice is contracted, there appears a concomitant velarization.

The matrix in Figure 1 displays the binary features of the emphatic vs. the nonemphatic consonants as conceived and reported by the early Arab and Muslim phoneticians and orthoepists.

Acoustically, the emphatic consonants display a markedly lower pitch. According to Jakobson (1962:512),

> The lowered pitch is a striking perceptual mark of pharyngealized phonemes which is synaesthetically expressed by native grammarians as 'dark, fat, thick, heavy....Spectrograms confirm that the pharyngealized consonant displays energy in a lower frequency region and affects the second formant of the following vowel in a downward direction.

In addition to the basic Arabic speech sounds or phonemes (*ḥurūf ʾuṣūl*; → *ḥarf*), the early phoneticians recognize sets of variants which Sībawayhi (*Kitāb* IV, 431ff.) calls *ḥurūf furūʿ*. They are of either allophonic, dialectal, stylistic, or free significance. The variants that concern us here are:

1. *ʾalif tafxīm* ([ɑː] and [ɑ]), e.g. [qaːma] > [qɑːma] (Ibn Jinnī, *Sirr* I, 50)
2. *ʾalif ʾimāla* ([eː] and [e]), e.g. [xaːtim] > [xeːtim] (Ibn Jinnī, *Sirr* I, 50)
3. /ṣ/ which sounds like [z]
4. the weak /ḍ/
5. /ṣ/ which is pronounced like [s]
6. /ṭ/ which sounds like [t]
7. /ḏ̣/ which resembles /ṯ/

Of interest to our discussion here is the short [ɑ] and long vowel [ɑː]. Both are open back and low vowels which usually occur before or after the emphatic consonants in most of the modern varieties of Arabic. This brings us to the question of the role the emphatic consonants can play in connected speech from a syntagmatic standpoint. It is apparent from the scattered statements regarding emphasis that it is seen as a prosodic feature in that it can influence neighboring vowels and consonants whether they are in relatively far or near proximity. A general rule for the phonetic emphasis in Arabic can be formulated as follows (where v = vowel, C = consonant, / = in the context of, ____ = after or before an emphatic (*emph*) consonant, and the parentheses ( ) indicating optionality of occurrence):

(C)	v	(C)	(C)	v
- *emph*	+ low - back	- *emph*	+ *emph*	- low

(C)	C
+ *emph* / ——	*mustaʿlī* (——
+ back	or *emph*

The best treatment of the two consonants /l/ and /r/ is found in the treatises on the phonetics of the *Qurʾān*, i.e. the science of *tajwīd* (cf. Nasr 1992:159ff.). Each of them comes in

Table 1. Frequency of Arabic phonemes

sequential order	phoneme	frequency of occurrence	percentage
1	/a/	7,707	16.74%
2	/i/	4,894	10.63%
3	/l/	3,482	7.50%
11	/r/	1,384	3.00%
21	/q/	551	1.20%
25	/ṭ/	366	0.80%
26	/ṣ/	339	0.74%
29	/x/	236	0.51%
30	/ġ/	234	0.51%
32	/ḍ/	205	0.45%
33	/ḏ̣/	167	0.36%

a pair of emphatic (*mufaxxam*) and nonemphatic (*muraqqaq*) consonants. In his paper on the emphatic /ḷ/, Ferguson (1956:446ff.) suggests that "the emphatic ḷ must be regarded as an independent phoneme in Classical Arabic and in most if not all the modern dialects". The early Arab and Muslim phoneticians were aware of this fact, but they chose a less complicated and more general approach, because only a very small set of lexical items is involved. Al-Nassir (1993:48–49) outlines the general rule for /ḷ/ as follows:

> The most common form of the Lām is the clear [l]. In certain phonetic contexts it appears as a dark [l,] as in the neighborhood of velarized or back consonants in some dialects.…In the proper noun /ʾaḷḷāh/ (God), if the Lām is preceded by the palatal short vowel /i/ it is produced as clear [l] /lillāhi/ (for God); when the preceding short vowel is the velar /u/ or the pharyngeal /a/ the Lām is produced as dark [ḷ], /ismuḷḷāhi/ (name of God) and /waḷḷāhi/ (by God).

As for the trilled /r/, it appears in Arabic as emphatic and nonemphatic. Al-Nassir (1993:49) states:

> In certain phonetic environments the Rāʾ develops a degree of *Tafkhīm*. This development, somewhat similar to 'darkening' the Lām, takes place when the Rāʾ occurs in the neighbourhood of the 'high' consonants *Ḥurūf al-ʾIstiʿlāʾ*,…as well as the two short vowels Fatḥah /a/ and Ḍammah /u/. Elsewhere it is a clear Rāʾ.

As for their frequency of occurrence, the *ḥurūf istiʿlāʾ* or emphatics show a very low phonemic load in the language. A statistical study on Arabic phonemes, carried out by al-Xūli (1984:52–53) on 46,029 phonemes in a random sample, gave the results in Table 1.

To conclude, *tafxīm* is treated by early Arab and Muslim grammarians, phoneticians, and orthoepists (cf. Ben Cheneb 1934:601), both in terms of place and manner of articulation, and within a simplified binary framework. Their method for description is both informative, with a high degree of accuracy, and interesting to any modern linguist or phonetician alike.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Ibn Jinnī, *Sirr* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *Sirr ṣināʿat al-ʾiʿrāb*, I. Ed. Ḥasan Hindāwī. Damascus: Dār al-Qalam, 1985.

Ibn Sīnā, *Risāla* = ʾAbū ʿAlī Ḥusayn ibn ʿAbdallāh Ibn Sīnā, *Risāla fī ʾasbāb ḥudūṯ al-ḥurūf*. Ed. Khalil I. Semaan, *Arabic phonetics*. Lahore: Muhammad Ashraf, 1963.

Ibn aṭ-Ṭaḥḥān, *Maxārij* = ʾAbū l-ʾAṣbaġ ʿAbd al-ʿAzīz Ibn aṭ-Ṭaḥḥān, *Maxārij al-ḥurūf wa-ṣifātuhā*. Ed. Muḥammad Yaʿqūb Turkustānī. 2nd ed. Private publication by the editor, 1991.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān ibn Qanbar Sībawayhi, *al-Kitāb*. Ed. ʿAbd as-Salām Muḥammad Hārūn, IV. Cairo: al-Hayʾa al-Miṣriyya al-ʿĀmma li-l-Kitāb, 1975. (Repr., Beirut: Dār al-Jīl.)

SECONDARY SOURCES

Al-Nassir, A.A. [ʿAbd al-Munʿim ʿAbd al-ʾAmīr an-Nāṣir]. 1993. *Sibawayh the phonologist: A critical study of the phonetic and phonological theory of Sibawayh as presented in his treatise* Al-Kitab. London: Kegan Paul International.

Bakalla, Muhammad Hasan. 1982. *Ibn Jinnī: An early Arab Muslim phonetician*. London and Taipei: European Language Publication.

Ben Cheneb, Mohammed. 1927. "*Tadjwīd*". *Encyclopaedia of Islam*, VII, 601. Leiden: E.J. Brill.

Dūsirī, ʾIbrāhīm ad-. 2004. *Muʿjam al-muṣṭalaḥāt fī ʿilmay at-tajwīd wa-l-qirāʾāt*. Riyadh: Al-Imam University.

Elgendy, Ahmed M. 2001. *Aspects of pharyngeal coarticulation*. Ph.D. diss., University of Amsterdam.

Ferguson, Charles A. 1956. “The emphatic *ḷ* in Arabic”. *Language* 32.446–452.
Jakobson, Roman. 1962. “*Mufaxxama*: The ‘emphatic’ phonemes in Arabic”. Roman Jakobson, *Selected writings*. I. *Phonological studies*, 510–522. The Hague: Mouton.
Naṣr, ʿAṭiyya Qābil. 1992. *Ġāyat al-murīd fī ʿilm at-tajwīd*. Cairo and Riyadh: Private publication by the author.
Xūlī, Muḥammad ʿAlī al-. 1984. “at-Taḥlīl al-ʾiḥṣāʾī li-ʾaṣwāt al-luġa al-ʿarabiyya”. Mecca: *Majalla Maʿhad al-Luġa al-ʿArabiyya* (ʾUmm al-Qurā University) 2.49–92.

Muhammad Hasan Bakalla
(Emeritus Professor; King Saud University)

Tajik

Tajik Persian (*zabån-e fårsi-e tåjiki*; as transliterated from Cyrillic, *zaboni forsi-i tojikī*) is the New Persian dialect of Central Asia, a descendant together with Persian of Iran (*fårsi*) from the spoken Middle Persian of the Sasanian Empire at the time of the Arab Muslim conquest of 640–712 C.E. The Arabic alphabet and vocabulary were instrumental in the rise of this language, which, in the form of Classical Persian, furnished a common idiom for writers of Iran, India, and Central Asia over many centuries. This survey accordingly focuses on the differences of distribution and form of Arabic loanwords in Central Asian Persian in comparison with other varieties, notably Standard → Persian of modern Iran.

Until the early 20th century, there was little difference in the style and vocabulary of Persian as written in Iran, Central Asia, or India. After the Bolshevik revolution in Bukhara (1920) and the creation of the Tajik SSR, a literary language called Tajik, based on vernacular Persian of Central Asia and written in a Latin alphabet (from 1928), then a Cyrillic alphabet (from 1939), was fostered as the language of the Soviet nationality of the Tajiks. (The term *Tåjik*, derived from the Middle Persian *Tāzīk* ‘Arab’, was an ethnonym distinguishing Persian speakers from Turks – specifically, in modern times, from Uzbeks.) As in the case of modern Turkish, the privileging of the vernacular and the break with the Arabic script provide for the first time a realistic indication of the status of Arabic loanwords in the language. Few statistics are available for Arabic loans in Tajik; the proportions are essentially similar to those in Standard Persian.

Arabic loanwords were originally assimilated into Persian of Central Asia in the same ways as into Persian of Iran. Over the centuries the two vowel systems diverged (see Table 1; → Persian), and a few consonants changed: /q/ and /ġ/ collapsed in Standard Persian but have remained distinct in Tajik. In some dialects (notably that of the Jews of Bukhara and Samarkand), the uvular pronunciation of *ʿayn* and postpalatal *ḥāʾ* is preserved. The mid back vowel /å/ of Tajik is rounded, and the contrast between long and short vowels has been neutralized in many dialects and in the standard orthography (contrast Standard Persian; see Perry 2004, § 1.1–1.2).

Table 1. Correspondence of Arabic and Tajik vowels

	long vowels			short vowels			diphthongs	
Arabic	ā	ī	ū	a	i	u	ay	aw
Tajik	å	ī, e	ū, ů	a	i, e	u, ů	ay	av, aw

Tajik has preserved the so-called *majhūl* vowels of early New Persian, originally *ē* and *ō*, distinct from *ī* and *ū*, although both are written with *wāw*; these appear in Table 1 as modern *e* and *ů* (a central vowel between *u* and *ü*, shared with Uzbek). Both these vowels are found in Arabic loanwords as allophonic variants of /i/ and /u/ respectively, mainly before /h/ and /ʾ/ (< Arabic ʿ): *ů:da* ‘responsibility’ (Arabic *ʿuhda*), *neᵉmat* ‘affluence’ (Arabic *niʿma*); /ů/ is also found in some Arabic loans of the pattern *mafʿūl*, instead of the more usual /u/: e.g. *ma:ruf* ‘familiar’, but *ma:ruza* ‘presentation, lecture’ (*maʿrūḍa*), and /e/ occurs initially in some loans of the pattern *ʾifʿāl*, e.g. *imån* ‘belief’ (*ʾīmān*), but *ejåd* ‘creation, production’ (*ʾījād*).

Some loanwords that are well established in (Standard) Persian have only Persian or Uzbek equivalents in everyday modern Tajik, e.g. *ġun dåštan* ‘to gather’ (Persian *jam kardan* < Arabic *jamʿ*); *pešvåz giriftan* ‘to meet, welcome’ (Persian *esteqbål kardan*); *gusel kardan* ‘to see off/out’ (Persian *mošāyeʾat kardan*). This was perhaps the result of a lower level of literacy, and competition with Persian and Uzbek vocabulary in the spoken language.

The lexicalization of Arabic broken plurals as singulars (a vernacular feature) appears more commonly in Tajik than in Persian, e.g. *a:zå-yi institut* 'a member of the institute' (Arabic *ʾaʿḍāʾ*, pl. of *ʿuḍw*), *yak maråtiba* 'once' (a blend of singular and plural; cf. Arabic *martaba* 'degree', pl. *marātib*, as borrowed in Persian). Variations on the Arabic(ate) feminine plural (*-åt, -jåt, -våt*) are more frequent, denoting collectives: *hayvånåt* 'fauna, livestock' (Arabic *ḥayawān*), *mevajåt* 'fruit', *sabzavåt* 'vegetables, greens' (cf. Persian *sabzijåt*; the last two are Arabicate plurals of Persian words). Other feminine plurals have developed a singular meaning: *kåinåt* 'cosmos' (a mass noun, < Arabic *kāʾināt* 'entity [pl.]'), *taškilåt* 'organization' (< Arabic *taškīlāt* [pl.]) , *hašaråt* 'insect' (< Arabic *ḥašarāt* [pl.]) with regular plurals *taškilåt-hå* and *hašaråt-hå*). Like early everyday Arabic loans for which there is no ready Tajik substitute (e.g. *kitåb* 'book', *havå* 'air, weather'), these later scholarly borrowings have survived the general Soviet condemnation of archaisms because they fill a useful niche.

Arabic loans in the feminine ending, being differentially assimilated (Arabic loanwords in → Persian), are an index of comparative vernacularization of borrowings as between Tajik and (Standard) Persian. Tajik reflexes in *-a* (more numerous and more vernacular, usually shared with Uzbek), contrast with the more literary Persian reflexes in *-at*: e.g. *jamåa* 'community; (madrasa) class, village, village soviet' (Arabic *jamāʿa*); *rioya kardan* 'to respect, maintain' (*riʿāya*); *ḥikåya* 'tale'; *himåya* 'protection, patronage'; *tarbiya* 'training, schooling'; *rivåya* 'narrative; fatwa'. The same forms sometimes took different semantic paths: *muråjiat* 'recourse, appeal' (Persian *moråjeat kardan* 'to return'; the doublet *moråjeʾe* is 'reference, recourse'; < Arabic *murājaʿa*). Tajik doublets in *-a* (where Persian has only the *-at* reflex) include *ibåra* 'idiom, term' (in addition to *ibårat az* 'consisting of', the widespread calque on Arabic *ʿibāratun min*); *ġåya* 'aim, goal'; and *kifåya* 'enough' (Perry 1984).

From the 16th century, cultural contacts between Central Asia and Iran diminished, as did the quality of education in Bukhara. Modern (pre-Soviet) neologisms, which often began as Arabicate coinages by the Turcophone intelligentsia of separate centers (Istanbul, Kazan, Baku), took different forms in Central Asia and Iran: *madaniyat* 'civilization' (also in Afghanistan; Persian *tamaddon*); *ittifåq* '(trade, professional) union' (Persian *ettehåd*); *ziåi* '(liberal) intellectual' (< Arabic *ḍiyāʾ* 'enlightenment'; in Persian *rowšan-fekr*, lit. '(of) bright thought', a Persian-Arabic hybrid; see Lazard 1956:178–182; Perry 2004, § 5.23–5.24).

Much Arabic vocabulary, as well as native Persian material, was displaced by Russian borrowings during the Soviet period. Since the late 1980s (and officially since independence in 1992), Russian vocabulary is being replaced by recourse to Persian words and morphs (both Tajik and Persian of Iran), and Arabic script is again being taught. However, Arabic is no longer a living lexical fount for Central Asian languages.

Bibliographical references

Lazard, Gilbert. 1956. "Caractères distinctifs de la langue tadjik". *Bulletin de la Société de Linguistique de Paris* 52:1.117–186.

Perry, John R. 1984. "Differential assimilation of some Arabic loanwords in Tajik and Uzbek". *Papers from the Third Conference on the Non-Slavic Languages of the USSR*, ed. Howard I. Aronson. *Folia Slavica* 7:1–2.268–282.

——. 2004. *Tajik Persian reference grammar*. Leiden: E.J. Brill.

John R. Perry (University of Chicago)

Tajnīs → Jinās

Tajwīd

1. Definition and material of *tajwīd*

Tajwīd is the system of rules regulating the correct and clear rendering of the *Qurʾān*, preserving the nature of a religious revelation whose meaning is expressed as much by its sound as by its content and expression. Muslims believe that *tajwīd* is the codification of the sound of the revelation as it was revealed to the Prophet Muḥammad and as he subsequently rehearsed it with the Angel Gabriel. Therefore, in Islam, there can be no history of the development of *tajwīd* except in terms of its scholarly codification, and it is generally believed that the rules were codified and written down in the 8th century C.E. The Qurʾānic verse "Recite the

Qurʾān with *tartīl*" (Q. 7/4) is interpreted to mean 'recite the *Qurʾān* according to the rules of *tajwīd*'. Correct recitation is therefore authorized by divine command and is the acknowledged duty of every Muslim. The rules not only govern the parameters of sound production (timbre and duration and articulation of syllables), but they also act to preserve the clarity of the meaning.

It is *tajwīd* that accounts for the characteristic and unique sound of the recited *Qurʾān*, a sound immediately recognizable in all its contexts throughout the Islamic world: learning to recite the *Qurʾān* means learning to reproduce this sound. The correct recitation is transmitted orally; the student listens to the teacher's recitation of a phrase or verse of the Qurʾānic text. The student imitates the sounds produced by the teacher, who then corrects the student's articulation. An increasing number of students also learn the correct recitation from cassette tapes and CDs; the same practice of listening, then imitating, applies. Learning the rules that produce the correct sound enables the student to predict the sound of a particular syllable.

The oral transmission process is crucial to both native and nonnative speakers of Arabic as many of the rules for pronunciation are not part of the sound system of literary and spoken Arabics but rather are applied uniquely to the Qurʾānic text. By themselves, written explanations and descriptions are not sufficient guides to reproducing the required sound. For example, *ʾixfāʾ*, a principle governing pronunciation of syllable-final /n/ or /m/ before certain consonants, is variously defined as "the pronunciation of a letter between full pronunciation and assimilation, free of doubling and with nasality of the letter" (Muḥayṣin 1970:809), "a state between full pronunciation and full assimilation" (al-Qāriʾ 1948:44), and "the sound between nasalized [n] and the ordinary sharp sound of [n]" (Tufayl 1974:85). Similarly, *qalqala*, the sound of certain syllable-final consonants which must be pronounced with following epenthetic schwa is traditionally described as vibrating the place of articulation so that a strong tone is heard (Naṣr 1930:62; al-Qāriʾ 1948:28) or as "consonnes bruyantes" (Boubakeur 1968:396). It is clear, therefore, that the role of the many manuals setting out the rules of *tajwīd* is to supplement the process of oral transmission, functioning only as an aid to help the student identify and remember what he or she has heard and taking for granted the student's access to the sound.

The material of *tajwīd* is largely phonetic, accounting for the unique sound. The rhythm of the recited text is governed by the rules of *madd* or syllabic duration. The rules for pause and beginning (*al-waqf wa-l-ibtidāʾ*) differ in nature from the other rules of *tajwīd* because they act to preserve the meaning rather than the sound; they mark the places in the text where it is appropriate or inappropriate for the reciter to interrupt the flow of meaning. Also included are prescriptions for pronunciation of the opening and closing formulae (the *istiʿāḏa, basmala*, and *ṣadaqa*), guidelines for the appropriate uses of the different *qirāʾāt* or reading systems, and the appropriateness of the *ṭuruq*, which are a matter of tempi and their associations with the reading systems (for a comprehensive discussion of the rules as they relate to the sound of recitation, see Nelson 2001:2–31).

2. Articulation of consonants and syllables

Most manuals and teachers of *tajwīd* begin with the physical points of articulation (*maxārij al-ḥurūf*) of the consonants of the Arabic alphabet and their phonetic realization. The consonants of the alphabet are classified according to their position in the vocal tract, starting with the larynx (*ʾaqṣā l-ḥalq*) and proceeding forward to the lips (*šifatāni*). This aspect of *tajwīd* spills over into non-Qurʾānic arenas. Singers and speakers in the Arab world are admired for their clear and correct pronunciation of Arabic, and this is usually attributed to their mastery of (at least this aspect of) the rules of *tajwīd*. The subject holds its religious authority, and aspiring singers and actors still learn their elocution from religious teachers.

The classification of phonemes as to manner of articulation (*ṣifāt al-ḥurūf*) corresponds to standard phonemic classifications, such as qualities of timbre (nasal), voiced/unvoiced, and retracted/unretracted, as well as by the categories of liquids, spirants, and dentals. However, the classification of 'strong' or 'weak', important for the rules of assimilation, does not wholly correspond to non-Arabic phonetic categories; for instance, when two consonants, sharing a point of articulation but differing in manner of articulation, are in sequence, the first assimilates to the second if it is 'weaker', as in [qālat ṭāʾif] becoming [qālaṭ ṭāʾif].

3. CONDITIONED ALLOPHONES

The characteristic sound of Qurʾānic recitation is largely determined by the principles governing nasality (*ġunna*), assimilation (→ *ʾidġām*, *ʾixfāʾ*, and *ʾiqlāb*), retraction or pharyngealization (→ *tafxīm*), and syllabic duration (*madd*). In addition, some of the reading systems variously dictate the principles of the softened glottal stop (*al-hamza al-musahhala*) within a word (e.g. [mūminūn] vs. [muʾminūn]), the fronted and raised /ā/ or → *ʾimāla* (e.g. [banāyha] vs. [banāha]), both of which are characteristic of the *qirāʾa Warš*; and the /ṣ/ tending toward /ž/ or /s/. The barely audible, if not often inaudible, articulation of final vowels, called *rawm* or *ʾišmām*, and the shift of accusative nunation from /an/ to /ā/ are determined by pause position but are also variously applied according to the particular reading system.

Nasality (*ġunna*) regulates which phonemes and syllables are to be articulated through the nasal cavity. Although phoneticians would classify certain phonemes/consonants as naturally nasal in timbre, such as geminate /m/ and /n/ (as in *janna* and *lamma*), the rules of *tajwīd* emphasize this quality and extend it to other phonemes. The effect is an intensified and conscious nasality which also affects the rhythm by prolonging the duration of the particular syllable (→ nasalization). The phonemes affected by *ġunna* are

a. geminate /m/, /n/, /w/, /y/, by virtue of full assimilation
b. /m/ and /n/, by virtue of partial assimilation (*ʾixfāʾ* and *ʾiqlāb*)

Assimilation is divided into full assimilation (*ʾidġām*), partial assimilation with change (*ʾiqlāb*), and partial assimilation without change (*ʾixfāʾ*).

Full assimilation governs the conjunction of syllable-final /n/ with the consonants /r/ and /l/ and /m/, /n/, /w/, /y/, the latter articulated with *ġunna*. The syllable-final /n/ assimilates to the following consonant, yielding a geminate, as in Q. 24/35: [šajaratim̃ m̃ubārakatĩ(ñ) zaytūnatil lā šarqīyyatiw̃ w̃a-lā ġarbīyya] instead of [šajaratin mubārakatin zaytūnatil lā šarqīyyatin wa-lā ġarbīyya].

Partial assimilation with change concerns syllable-final /n/ before /b/. The preceding vowel is nasalized, while the lips shape the /m/ sound but do not close until the /b/ is articulated, as in [mĩ(m̃) baʿd] instead of [min baʿd].

Partial assimilation (*ʾixfāʾ*) of /n/ takes place before consonants /t/, /ṭ/, /j/, /d/, /ḏ/, /z/, /s/, /š/, /ṣ/, /ḍ/, /ṭ/, /ḏ̣/, /f/, /k/, and /q/ and syllable-final /m/ before /b/. The lips do not quite close for the /m/, and the tongue does not quite touch the alveolar ridge for the /n/, the vocal cavity holding the shape of the preceding vowel and the sound articulated through the nasal cavity. Conjunction may occur within a word or between two words, as in [kũ(ñ)tũ(m̃) bihi], instead of [kuntum bihi].

The phenomenon of *qalqala* affects syllable-final consonants /q/, /ṭ/, /b/, /j/, and /d/ with following epenthetic schwa /ə/. The effects of *qalqala* on stress and rhythmic patterns contribute to the unique sound of the Qurʾānic text. Compare the same line (Q. 96/19) with and without *qalqala*: [la̯-q̱a̱d xa̯lá̱qṉa l-ʾi̱nsā̱́na̯ f̱ī̱] with [la̯-q̯adə̯xa̯lá̯qə̯na̱ l-̱ʾ̱ĩ̱(̱ṉ̃)sā̱́na̯ fī̱] and [w-á̱sju̱d w-a̱qṯ̯áṟi̱b] with [w-á̱sju̯də̯w-á̯qə̯ta̯ri̯ bə̯].

Tafxīm is retraction/pharyngealization. In addition to the consonants that in most varieties of Arabic govern the backing of short and long vowel /a/ (/q/, /ṣ/, /ḍ/, /ṭ/, and /ḏ̣/, or the *ḥurūf al-istiʿlāʾ*), Qurʾānic Arabic (and some dialects) add /x/ and /ġ/. Qurʾānic pronunciation is therefore [xạ̄lidūna] vs. [xālidūna] and [ġạyr] vs. [ġayr].

Tafxīm also governs the pharyngealization of /l/ following /a/ and /u/, as in [w-ʾaḷḷạ̄h] vs. [billāh], and of the pharyngealized vs. the trilled /r/, as in [ṛạbb] vs. [karīm].

Another feature in which Qurʾānic pronunciation differs from other varieties of Arabic is that the phenomenon of *tafxīm* is only immediately progressive, not regressive, e.g. [laqạb] not [lạqạb], and [ʾaṛḍ], not [ʾạṛḍ], and does not extend to the whole word but affects only an immediately following /a/ or /ā/, as in [xạlaqạ] vs. [xạlạqạ], and [ṣiyām] not [ṣiyạ̄m].

Since *tafxīm* is restricted in context, the sound of Qurʾānic recitation is characterized by more frequent alternation between the pharyngeal and the nonpharyngeal sounds than is found in other varieties of Arabic, as in [ṣạdaqạ] vs. [ṣạdạqạ].

4. Lengthening (*madd*)

The rules of *madd* or duration codify the rhythms of the Qurʾānic language. Together with *ġunna*, which tends to prolong the given duration of the syllable, *madd* accounts for the rhythm of a text that is distinct from Arabic prosody (with the exception of a few lines) and from the rhythmic system of Arabic music. Therefore, although recitation may share features with the melodic system of Arabic music, the rhythm of the Qurʾānic text is considered to be divinely given and supports arguments for the uniqueness of the text *vis-à-vis* music and poetry.

Arabic prosody classifies the syllable into short and long durations (˘ and ¯), one long being approximately equivalent to two short. The durations of syllables in Qurʾānic recitation range from one to six beats (*ḥarakāt*), and the rules of *tajwīd* divide the syllables into categories of fixed and variable duration. The minimum duration of one beat is fixed and applied to the short, open syllable (Cv, as in [mu̮bāra̮ka̮tin]), while the greatest fixed duration of six beats (but often longer) applies to the long vowel with following geminate consonant (e.g. [d̮̲ā̲l̲l̲ī̲n̲] and [L̲ā̲m̲ Mī̲m]). The pausal duration is the final syllable before a breath or the pause taken at the end of a verse (characteristically CvC or CvC), and the reciter may choose the shorter or longer duration (from two to six beats).

While the rules governing duration are detailed and comprehensive, covering every type of syllable in every conceivable context, at the same time variation is built into the system: some options are regulated by the reading systems (e.g., ʾv̄ or *badal* takes two beats in all of the reading systems except that of Warš, where it may extend to four or six beats), and otherwise the reciter is free to choose his durations and obligated only to be consistent within a single recitation. For example, if the reciter begins reciting with a pausal duration of six beats, he may not shorten it within the same recitation. Finally, although the underlying pulse is set by the shortest duration, the 2:1 ratio of the long and short syllables is tempered by *ġunna*. Compare the previous examples: [ša̮ja̮ra̮tim̃ m̲̃u̮bā̲ra̮ka̮tī̲(̲ñ̲) za̲ytū̲na̮ti̲l̲ lā̲ ša̲rqī̲yya̮ ti̲w̲̃ w̲̃a̮-lā̲ ġa̲rbī̲yya̮] vs. [ša̮ja̮ra̮ti̲n mu̮bā̲ra̮ka̮ti̲n̲ za̲ytū̲na̮ti̲l̲ lā̲ ša̲rqī̲yya̮ti̲n̲ wa̮-lā̲ ġa̲rbī̲yya̮] and [kū̲(̲ñ̲)tū̲(̲m̲̃) bi̮hi̮] vs. [ku̲ntu̲m bi̮hi̮].

The choice of durations is further influenced by the tempo and style of the recitation: whereas the aim of most Muslims is simply to recite correctly, the skilled professional reciter manipulates the rules for maximum aesthetic effect, the intent of the ideal recitation being to engage fully the emotional participation of the listeners. The public and melodic style of recitation is therefore characterized by the longer durations as they present greater opportunity for melodic development.

The durational options within a particular reading system, as well as the phonetic characteristics that differentiate the reading systems, allow for significant variation in the overall sound of recitation. At the same time, the sound of Qurʾānic recitation is unique: the rules of *tajwīd* both encode and make perceptible the singular nature of the text and the significance of its sound in the Muslim community.

Bibliographical references

Boubakeur, Si Hamza. 1968. "Psalmodie coranique". *Encyclopédie de musiques sacrées*, ed. Jacques Porte, I, 388–403. Paris: Editions Labergerie.

Muḥaysin, Muḥammad Sālim. 1390/1970. *Muršid al-murīd ʾilā ʿilm at-tajwīd*. Cairo: Maktabat al-Kulliyyat al-ʾAzhariyyah.

Naṣr, Muḥammad Makkī. 1349/1930. *Nihāyat al-qawl al-mufīd fī ʿilm at-tajwīd*. Cairo: Maṭbaʿa Muṣṭafā al-Bābī al-Ḥalabī wa-ʾAwlādih.

Nelson, Kristina. 2001. *The art of reciting the Qurʾan*. Cairo and New York: American University in Cairo Press.

Qāriʾ, ʿAlī ibn Sulṭān Muḥammad al-. 1948. *al-Minaḥ al fikriyya: Šarḥ al-Muqaddima al-Jazariyya*. Cairo: Maktabat wa-Maṭbaʿa Muṣṭafā al-Bābī al-Ḥalabī wa-ʾAwlādih.

Tufayl, S. Muhammad. 1974. *The Qurʾān reader*. Trinidad: Ladies Islamic Association.

Kristina Nelson (Cairo)

Taʿlīq

The term *taʿlīq* (or → *ʾilġāʾ*) 'hanging', i.e. 'abrogation, interruption, cancellation', indicates the interruption of grammatical effect when a word is not used in a governable position although it is a governable word, as shown by the fact that it exhibits the normal case or mood distinction (Owens 1988:50–51). Cancellation of government is mostly connected with judgment verbs (*ʾafʿāl al-qalb*) such as *ḥasiba* 'to consider, to reckon'; *xāla* 'to believe, to imagine', *darā* and *ʿalima* 'to know'; *raʾā* 'to see, to consider', *ḏanna* 'to think, to believe'; *ʿadda* 'to consider; to regard'; *zaʿama* 'to assert, to claim'; *wajada* and *ʾalfā* 'to find'; etc.

Taʿlīq occurs in parenthesis rather than in initial position (Kinberg 1996:753), in a verbal expression in a sentence with the purpose of creating an obstacle. After *taʿlīq*, the sentence is in the syntactic position of *naṣb*, blocking a direct object, for example: *ḥasibtu la-ʿabdullāhi qādimun* 'I considered that, necessarily, ʿAbdallāh would come'. Here, *ḥasibtu* is the predicate of the agent of the verbal clause. The particle *la-* puts the following word in the position of the topic, thereby causing the action of the governing word *ḥasibtu* to be blocked. *Qādimun* is the second object of the judgment verb *ḥasibtu* and the predicate of the word *ʿabdullāh*. The nominal clause *ʿabdullāh qādimun* consists of the topic (*mubtadaʾ*) *ʿabdullāh* and the predicate *qādimun*. It occupies the position of the object (*naṣb*) and as such is dependent on the verb *ḥasiba*, but the government of this verb is blocked. It is significant that this government may become unblocked in a conjunctional apposition, for example: *ḏanantu la-zaydun munṭaliqun wa-ʿamran* 'I thought that Zayd was leaving, and ʿAmr [as well]' (Girgas 1873:117).

All verbs of judgment are characterized by *taʿlīq*, except for the verbs *hab* 'suppose that, assuming that' and *taʿallam* 'know!', because they are considered to be nonderivatives inasmuch as they are used only in this form (Junaydī 1981:320).

The cancellation of the government of a judgment verb takes place under the following conditions:

i. When a verb is located between two nouns (objects), for example: *zaydun ḥasibtu ʿāqilun* 'Zayd, I believed, is clever' (Girgas 1873:117; Grande 1963:442);
ii. When a verb is located after two nouns, for example: *zaydun jāhilun ḏanantu* 'Zayd is a fool, I thought'; *aṭ-ṭarīqu qaṣīrun ʾaḏunnu* 'the road is short, I think' (Girgas 1873:117; Grande 1963:442);
iii. When a verb is followed by a negation particle such as *mā*, *ʾin*, *lā*, for example: *ḏanantu mā ʿabdullāh qādimun* 'I thought that ʿAbdullāh would not come'; *ḥasibtu mā raʾyu-ka nāfiʿun* 'I considered that your opinion was not useful'; *ʿalimtu ʾin al-ḥayātu dāʾimatun* 'I have learned (found out) that life is not eternal'; *xaltu lā ṣāḥibu sūʾin maḥbūbun* 'I believed that the malicious man was disliked'; *ḥasibtu lā l-faqru dāʾimun wa-lā l-ġināʾ* 'I considered that neither poverty nor riches are eternal' (Girgas 1873:117; Grande 1963:442; Junaydī 1981:320);
iv. When a verb is followed by a proper name with the confirmative particle *la-* 'necessarily', 'obligatory', which causes the following word to be in the position of topic, for example: *ḏanantu la-zaydun jāhilun* 'I believed that Zayd was really ignorant'; *ʿalimtu la l-qiṭāru qādimun* 'I have found out that the train will certainly arrive'; *ʿalimtu ʾanna ḥaqāʾiba-ka la-maḥmūlatun fī l-qiṭāri* 'I have found out that your suitcases will really go by train' (Grande 1963:442; Junaydī 1981:320);
v. When a verb is followed by the particle of oath *la-* (*lām al-qasam*), for example: *ʾarā t-taqṣīru fī l-ʿamali huwa ʾisāʾatun la-l-waṭanu* 'I regard remission in work as evil, I swear by the fatherland' (Junaydī 1981:320);
vi. When a verb is followed by a question:
    (a) with a particle, for example: … *wa-ʾin ʾadrī: ʾa-qarībun ʾam baʿīdun mā tūʿadūna* (Q. 21/109) '…even though I know not, whether near or far is that you are promised' (Arberry 1996:II, 26);
    (b) with a noun, for example: … *li-naʿlama ʾayyu -l-ḥizbayni ʾaḥṣā li-mā labiṯū ʾamadan*? (Q. 18/12) '…that We might know, which of the two parties would better calculate the while they had tarried' (Arberry 1996:I, 317); … *la-taʿallamunna: ʾayyu-nā ʾašaddu ʿaḏāban*? (Q. 20/71) '…you shall know of a certainty which of us is more terrible in chastisement' (Arberry 1996:I, 343).

An interrogative noun can be construed with a judgment verb interrupting government, if it functions as:

i. Topic of a nominal clause, as in the above-mentioned Qurʾānic passages;
ii. Predicate of the topic of a nominal clause, for example: *ʿalimtu matā r-raḥīlu*? ‘I have found out when the departure [is going to be]?’. Here, the adverb of time *matā* ‘when?’ acts as an interrogative noun. It functions as a word that makes the topic of the nominal clause (*raḥīlu*) to be used in *rafʿ*;
iii. The second component of *ʾiḍāfa* of a topic, for example: *ʿalimtu mawqifu ʾayyihim ʾaqwā* ‘I have found out the position of whoever of them is stronger’.

The interruption of government is carried out by means of words that are principal (basic) in a clause, such as the interrogative pronoun *kam* ‘how much?’ as a predicate (*kam al-xabariyya*), for example: *daraytu kam šajaratin ġarasta-hā* ‘I have found out how many trees you have planted’, or particles like *ʾinna* (except for *ʾanna*, as it cannot be principal), for example: *ʿalimtu ʾinna-ka la-munṣifun* ‘I found out, you really are a fair man’. After the verb *darā*, the particle *laʿalla* acts as an instrument canceling its government, for example: *ʾadrī laʿallaka tadrī mā qīmatu ṣ-ṣiḥḥati* ‘I know that you probably know what is the value of health’.

The elimination of government occurs at the expense of conditional particles (in *jazm* and without *jazm*), for example: *lā ʾaʿlamu ʾin kāna ṭ-ṭaqsu mulāʾiman li-l-ʿamali* ‘I do not know whether the weather will be suitable for the work’; *ʾaḏ̣unnu law tafāhama l-ḥākimu wa-l-maḥkūmu la-ḥaqqaqā l-fawza li-l-waṭani* ‘I think, if the judge and the defendant had understood each other, they would have ensured a victory for the fatherland’.

In addition to judgment verbs, other transitive verbs with a question can sometimes eliminate government of an object, for example ... *fa-l-yanḏ̣ur ʾayyu-hā ʾazkā ṭaʿāman* (Q. 18/19) ‘... let him look for which of them has the purest food’ (Arberry 1996:I, 318).

BIBLIOGRAPHICAL REFERENCES
Arberry, Arthur J. 1996. *The Koran interpreted*. New York: Touchstone.
Girgas, Viktor. 1873. *Očerk grammatičeskoj sistemy arabov* [Survey of the grammatical system of the Arabs]. St. Petersburg: Tipografija Imperatorskoj Akademii nauk.
Grande, Bentsion M. 1963. *Kurs arabskoj grammatiki v sravnitel’no-istoričeskom osveščenii* [Course of Arabic grammar in comparative historical perspective]. Moscow: Izdatel’stvo Vostočnoj Literatury.
Junaydī, Muḥammad. 1981. *Aš-Šāmil: Muʿjam fī ʿulūm al-luġa al-ʿarabiyya wa-muṣṭalaḥāti-hā*. Beirut: Dār al-ʿAwda.
Kinberg, Naphtali. 1996. *A lexicon of al-Farraʾ’s terminology in his* Qurʾān *Commentary*. Leiden: E.J. Brill.
Owens, Jonathan. 1988. *The foundations of grammar: An introduction to medieval Arabic grammatical theory*. Amsterdam and Philadelphia: J. Benjamins.

VALERIY RYBALKIN
(National Academy of Sciences of Ukraine)

Taltala

The ancient Arabic grammarians made a distinction between two types of Arabic according to the nature of the vowel of the imperfect preformative: some Arab tribes used *ya-*, while others used *yi-*. The name which designates the usage of *yi-* is *taltala*. According to Rabin (1951:60), who carefully collected all available data mentioned by the Arabic grammarians, a group of tribes who lived in a large area from Egypt to Iraq, the Quḍāʿa, had the full *taltala*. Some other tribes had a partial *taltala* only, i.e., they used *yi-*, *ti-*, *ʾi-*, *ni-* for the different persons of imperfect and imperative, but for the 3rd person masculine singular, they had *ya-*: Qays, Tamīm, ʾAsad Rabīʿa, and *ʿāmmat al-ʿArab* without further definition (the situation is not clear as far as Ṭayyiʾ are concerned). On the other hand, all tribes who lived in al-Ḥijāz, the ʾAzd Sarāt, a part of the Hawāzin, and a part of the Huḏayl, in addition to the tribes from Yemen, had the imperfect preformative in *ya-* (Rabin 1951:61, 158; the sources are Sībawayhi, Ibn Hišām, al-ʾAstarābāḏī).

Sībawayhi (*Kitāb* II, 256–257) contrasts two basic forms in the *taltala* dialects: an intransitive *faʿila/yifʿalu* in contrast with a transitive *faʿala/yafʿalu*, *yafʿulu*. He considers the non-*taltala* dialects of al-Ḥijāz as having the regular form *faʿila/yafʿalu*, arguing that *faʿula/yafʿulu* is pan-Arabic and not *yufʿulu*. Al-Bayḍāwī mentions a form *ʿabada/yiʿbudu* in the Tamīm tribe (Grand’Henry 1990:40–46).

Barth's (1894) principle points to the fact that intransitive *faʿila/yifʿalu* and transitive *faʿala/yafʿulu* are forms which should be considered to date back to the Proto-Semitic language. It seems that two basic processes of extension through analogy should be taken into consideration: on the one hand, the Proto-Semitic and Arabic *taltala* pattern finally became widespread throughout the modern Arabic dialect area. On the other hand, the non-*taltala* Ḥijāzī and Yemeni pattern seems to have been carried over a long distance, since there is nothing but *ya-* as imperfect preformative in al-Andalus old dialects (Valencia, Granada; Grand'Henry 1990:45–46).

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCE

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. 2 vols. Bulaq, 1316 A.H. (Repr., Baghdad: Al-Muthanna Library, n.d.)

SECONDARY SOURCES

Barth, Jakob. 1894. "Zur vergleichenden semitischen Grammatik: Zu den Vocalen der Imperfect-Präfixe." *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 48.4–6.

Grand'Henry, Jacques. 1990. "La *taltala* des grammairiens arabes et le parler d'al-Andalus". *Zeitschrift für Arabische Linguistik* 21.40–46.

Rabin, Chaim. 1951. *Ancient West Arabian*. London: Taylor's Press.

JACQUES GRAND'HENRY
(Catholic University of Leuven)

Tamakkun

The lexical meaning of the verb *tamakkana* is 'to be powerful, to be able to do something'. In grammatical terminology, *tamakkun* is used for a general grammatical and semantic category indicating the ability of words to be inflected and to perform various grammatical functions (for a survey of previous reconstructions of this term's function, see Danecki 1993; Talmon 2003:287, n. 2). A part of speech possessing *tamakkun* is called *mutamakkin*; it should have regular inflection and regular form. That which is irregular is classified as *ġayr mutamakkin*. The more functions a word can potentially perform, the more *mutamakkin* it is. For instance, a fully inflected *mutamakkin* noun may potentially have all inflectional suffixes – it can be declined, but it can also become determined or indetermined, or it can change from singular to plural. An uninflected noun is not capable of taking these suffixes, hence, it is *ġayr mutamakkin*. Thus, the *mutamakkin* is the unmarked element of the opposition, while *ġayr mutamakkin* is the marked one (cf. Owens 1988:202–203).

The category of *tamakkun* stands in direct relationship to the idea of *xiffa/ṯiqal* 'lightness/heaviness'. A word that is *mutamakkin* is at the same time *xafīf* 'light', not (yet) burdened with grammatical markers and functions but potentially ready to perform these functions. It is open to receive certain forms, such as case endings, provided they have formal markers. If the case endings are not formally marked on the surface level, the words remain *ġayr mutamakkin*. Yet, at the same time, they are now *ṯaqīl*, because they perform more than one syntactic function: the case endings have merged, as for instance in the word *ḏikrā* 'remembrance' (the underlying forms of the ending *-ā* are the case endings: *-u, -i*, and *-a*), hence, it is *ṯaqīl* (for the meaning of *xiffa/ṯiqal* in Arabic phonetics, see Bohas 1981, 1984).

The category of *tamakkun* was used and discussed by Sībawayhi in his *Kitāb*, where it occurs frequently (according to Troupeau 1976:196, both *tamakkun* and *mutamakkin* occur 35 times, and *ġayr mutamakkin* another 21 times). It may have been current in earlier grammatical theory, although Talmon (2003:287) states that he was unable to trace its origins in pre-Sībawayhian grammatical doctrine. Al-Xalīl applied this category solely to inflection (*taṣrīf*; Talmon 1997:164–165). In Kufan grammar, the term does not seem to have been current. In al-Farrāʾ's *Maʿānī* (I, 165.12), it is used only once, in a different context (Kinberg 1996:778): al-Farrāʾ states that the difference between *mā laka ʾan taqūma* 'you're not allowed to get up' and *mā laka l-qiyāmu* resides in the fact that *qiyām* is a full noun (*ism ṣaḥīḥ*) whereas the complementizer *ʾan* is 'a particle that has no standing in the nominals' (*ḥarf laysa bi-mutamakkin fī l-ʾasmāʾ*).

Sībawayhi uses *tamakkun* as a general category applied to all parts of speech – nouns, verbs, and the *ḥurūf* (prepositions, particles, conjunctions) – as well as to phonemes, and even to the semantic idea of nearness and remoteness (Mosel 1975:75–80). The word *qurb* 'vicinity' is classified by him as *mutamakkin*, while *buʿd* 'remoteness' is classified as *ġayr mutamakkin*, since it cannot be used in phrases of the type *qurbaka* 'near you' (Sībawayhi, *Kitāb* I, 284. 13–14). Parts of speech are categorized according to the degree of *tamakkun*. Nouns are more *mutamakkin* than verbs, while *ḥurūf* are classified as non-*mutamakkin*. Within nouns, *tamakkun* is gradual: fully inflected nouns are *mutamakkin*, while others are either less *mutamakkin* or *ġayr mutamakkin*.

Grammatical categories such as case, state (determination/indetermination), number, and gender are described by *tamakkun*. An indetermined word is more *mutamakkin* than a determined one, since it is liable to come determined. The same applies to number and gender: singular can become plural, hence, it is more *mutamakkin*; similarly, masculine can become feminine (as is the case with adjectives). Fully inflected nouns and verbs are *mutamakkin*, while defective verbs like *laysa* are *ġayr mutamakkin*.

Sībawayhi also applies the category of *tamakkun* to phonetic processes. The eliding *hamza* (*hamzat al-waṣl*) has greater potential power to change (is more *mutamakkin*) than the dividing *hamza* (*hamzat al-qaṭʿ*). Likewise, the semivowel *y* is more *mutamakkin* than *w* since *w* is liable to change into *y*, as in **maʿduww* becoming *maʿdiyy* (*Kitāb* II, 260.7–10).

Since the category was applied predominantly to nominal inflection, the later grammatical tradition limited the term *tamakkun* to nominal inflection. This explains why in modern grammars it is generally understood as the 'inflectability' of nominal parts of speech. Along with *tamakkun*, the term *ʾamkaniyya* is also used in modern grammars (as, for instance, by Ḥasan 1974:I, 33).

The category of inflectional *tamakkun* was introduced by later grammarians, who distinguished degrees of *tamakkun: mutamakkin ʾamkan, mutamakkin ġayr amkan*, and *ġayr mutamakkin* (Ḥasan 1974:I, 37). To the first category, *mutamakkin ʾamkan*, belong fully inflected nominal parts of speech, and to the second category, *mutamakkin ġayr ʾamkan*, partially inflected nominals such as diptota (→ diptosis), also called *ġayr munṣarif*, which have lesser *tamakkun* and are just *mutamakkina*. The third group, *ġayr mutamakkina*, contains the so-called *mabniya* words, such as the proper name *Xālawayhi*, or exclamations, such as *ṣah* 'hush!' or *ʾīhi* 'come on!'. Thus, in modern grammatical thought, the category of *tamakkun* becomes synonymous with the categories of → *ʾiʿrāb* and → *bināʾ*. A word that is *muʿrab munṣarif* or, in later terminology, *maṣrūf* is *mutamakkin ʾamkan*; a word that is *muʿrab ġayr munṣarif* is *mutamakkin*, and a *mabnī* word is *ġayr* (this is stated explicity by Ḥasan 1974:I, 41, 75). A parallel new terminology was introduced, and partially declined or undeclined words are called *mamnūʿ min aṣ-ṣarf* 'exempt from declension' (→ *ṣarf*).

The relationship between the terminologies of weight (lightness/heaviness), power (*tamakkun*), and self-determination (*taṣarruf*) is highly significant. In folk taxonomies in the Arab world, the adjectives *xafīf* and *ṯaqīl* are applied to the difference between men and women: women are regarded as light, which is why they have no power but are able to move around. Men, on the other hand, are seen as heavy, which explains why they are powerful and move slowly and in a dignified fashion (Jansen 1987:183). The same dichotomy applies to words in the view of the Arabic grammarians: nouns are light, cannot govern (→ *ʿamal*), and are liable to change. Verbs, on the other hand, are heavy, entitled to government, and not liable to change.

Bibliographical references

Primary sources

Farrāʾ, *Maʿānī* = ʾAbū Zakariyyāʾ Yaḥyā ibn Ziyād al-Farrāʾ, *Maʿānī l-Qurʾān*. Ed. Muḥammad ʿAlī an-Najjār. 3 vols. Cairo: ad-Dār al-Miṣriyya, 1966–1972.

Ibn as-Sarrāj, *ʾUṣūl* = ʾAbū Bakr ibn as-Sarī Ibn as-Sarrāj, *Kitāb al-ʾuṣūl fī n-naḥw*. Ed. ʿAbd al-Ḥusayn al-Fatlī. 3rd ed. 3 vols. Beirut: Dār ar-Risāla, 1996.

Mubarrad, *Muqtaḍab* = ʾAbū l-ʿAbbās Muḥammad ibn Yazīd al-Mubarrad, *Kitāb al-muqtaḍab*. Ed. ʿAbd al-Xāliq ʿUḍayma. 2nd ed. 4 vols. Cairo, 1979.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. Būlāq, 1318 A.H.

Secondary sources

ʿAfīfī, ʾAḥmad. 1996. *Ḏ̣āhirat at-taxfīf fī n-naḥw al-ʿarabī*. Cairo.

Bohas, Georges. 1981. "Quelques aspects de l'argumentation et de l'explication chez les grammairiens arabes". *Arabica* 28.204–221.

——. 1984. "L'explication en phonologie arabe". *Studies in the history of Arabic grammar*, ed. Hartmut Bobzin and Kees Versteegh, 45–52. (= *Zeitschrift für Arabische Linguistik* 15.) Wiesbaden: O. Harrassowitz.

Chairet, Mohamed. “*Ḫiffa, ṯiqal* et *tamakkun*: Régime d’incidence et classes de mots”. *Langues et Littératures du Monde Arabe* 1.213–226.
Danecki, Janusz. 1993. “The notion of *tamakkun* in Sībawayhi’s grammar”. *Rocznik Orientalistyczny* 48.121–130.
Diem, Werner. 1974. “Nomen, Substantiv und Adjektiv bei den arabischen Grammatikern”. *Oriens* 23.312–332.
Ḥasan, ʿAbbās. 1974. *an-Naḥw al-wāfī*. 3rd ed. 4 vols. Cairo: Dār al-Maʿārif.
Jansen, Willy. 1987. *Women without men: Gender and marginality in an Algerian town*. Leiden: E.J. Brill.
Kinberg, Naphtali. 1996. *A lexicon of al-Farrāʾ’s terminology in his* Qurʾān *commentary with full definitions, English summaries, and extensive citations*. Leiden: E.J. Brill.
Mosel, Ulrike. 1975. *Die syntaktische Terminologie bei Sībawaih*. 2 vols. Ph.D. diss., University of Munich.
Owens, Jonathan. 1988. *The foundations of grammar: An introduction to medieval Arabic grammatical theory*. Amsterdam and Philadelphia: J. Benjamins.
Talmon, Rafael. 1997. *Arabic grammar in its formative age:* Kitāb al-ʿAyn *and its attribution to Ḫalīl b. Aḥmad*. Leiden: E.J. Brill.
——. 2003. *Eighth-century Iraqi grammar: A critical exploration of pre-Ḫalīlian Arabic linguistics*. Winona Lake, Ind.: Eisenbrauns.
Troupeau, Gérard. 1976. *Lexique-index du* Kitāb *de Sībawayhi*. Paris: Klincksieck.

Janusz Danecki (University of Warsaw)

Tamil

At first glance, Tamil, a Dravidian language spoken by more than seventy million people in India, Sri Lanka, Malaysia, and Singapore, appears to be largely devoid of loanwords from Arabic, especially if compared to other South Asian languages. Even with reference to Muslim religious practice, Tamil words often replace Arabic ones, such as Tamil *toḻukai* instead of Arabic *ṣalāt*. Nevertheless, Arabic words are used by Tamil-speaking Muslims in both everyday conversation and Islamic literary texts dating back to the 16th century, and even among non-Muslims, Arabic terms can be found in administrative and nautical vocabulary.

Tamil came into contact with Arabic through two different routes roughly corresponding to the routes by which Islam was transmitted to South India, although the spread of Arabic vocabulary at times occurred independently of processes of Islamization. One route involved the Arab and Persian merchants who frequented the port cities of Southeast India and Ceylon since at least the beginning of the 2nd millennium C.E., thereby giving rise to many Muslim communities along the coasts of these regions. Arabic was transmitted to Tamil-speaking areas through a second route, the Persianate culture of states in central and northern India, which repeatedly conquered the region from the 14th century onward (cf. generally Fanselow 1989; Nainar 1942).

Little research has been undertaken on the topic. Most contributions are descriptive, with little phonetical, morphological, or semantic analysis (Nainar 1941; Mukamatu 1996:125–154; Uwise 1976:355–405, 1983; Vaidyanathan 1958; cf. also Bausani 1971). The focus of most studies is on Arabic loanwords in Islamic Tamil literature, with examples from non-Islamic literature and contemporary spoken Tamil mentioned only occasionally. Studies of contemporary spoken Muslim dialects are confined to a few, generally not easily accessible, dissertations (cf. nos. 2 and 1349 in Agesthialingom and Sakthivel 1973). Examples from the Muslim dialect of the Kanniyakumari district given below are therefore taken from a well-known novel (Tōppil Muhammatu 2004).

1. Script and phonological assimilation

Given the dearth of research on Arabic loanwords in spoken varieties of Tamil, most of our information regarding the assimilation of Arabic sounds to Tamil phonology has to depend on written sources. These employ both the Tamil and a modified version of the Arabic script which shows similarities to versions in use in Southeast Asia (cf. Shuʾayb 1993:95–99; Tschacher 2001:6–18; Vinson 1895). Obviously, while the use of the Arabic script often facilitates reading, it usually obscures sound changes. Yet, use of the Tamil script does not necessarily facilitate the identification of such changes, either. The distinction between voiceless and voiced consonants and between plosives and fricatives is not phonemic in Tamil nor are there separate letters to write these phones (cf. Andronov 2003:27–31). For example, the phoneme /c/ is pronounced [c], [ɟ], [s], or [ç] in different contexts, but it is written with just one letter, transliterated as *c*. This often makes it difficult to glean the actual pronunciation of Arabic words from the orthographic representation. Furthermore, there is substantial orthographic variation. While earlier spellings often used epenthetic letters to indicate actual pronunciation, there has been a tendency to drop such letters in recent times and to come closer to a transliteration of Arabic words with the help of so-called *grantha* letters originally used solely to write Sanskrit (cf. Das's [1981:346] distinction of 'scientific' and 'vulgar' systems of transliterating Arabic in Tamil).

There is no space to discuss all phonological assimilations and orthographic conventions here; for a short overview of the latter, see Das (1981). It will suffice to indicate some of the more common and interesting ones. Vowels remain largely unchanged; final *-ī*/*-ū* are usually shortened. Initial *wa-*, *ya-* commonly change to *o-*, *e-*, e.g. *oli* 'saint; guardian' (< *walī*), *ekūti* 'Jew' (< *yahūdī*). /ʾ/ tends to disappear in syllable-final position but lengthens a preceding vowel, e.g. *mūmiṉ* 'believer' (< *muʾmin*). /ʾ/, /ʿ/, and /ḥ/ tend to become /y/ in intervocalic position preceding *i*/*ī*, e.g. *kāyip* 'hidden' (< *ġāʾib*), *cayītu* 'fortunate' (a personal name, < *saʿīd*), *cāyapu* 'companion' (also an honorific title; < *ṣāḥib*). In words which are in common spoken use, more radical changes occur, e.g. *mōtiṉ* 'muezzin' (< *muʾaḏḏin*) or the personal name *Meytīṉ* (< *Muḥyī d-Dīn*). Intervocalic voiceless stops and fricatives are occasionally voiced, e.g. *āġir* 'last' (< *āxir*; cf. Vinson 1895:154, 159).

Most interesting are the reflexes of Arabic /ḍ/ and /ḏ̣/, as these often allow one to distinguish loans borrowed through Indian Ocean networks from those borrowed from northern India. The most widespread reflex of both phonemes, common in earlier Islamic Tamil literature but also widespread in spoken Tamil among Muslims in both India and Sri Lanka, is a lateral, either dental /l/ or retroflex /ḷ/, e.g. *paṟulu*/*paṟuḷu* 'duty' (< *farḍ*). A lateral reflex of /ḍ/ is found also in several Southeast Asian and West African languages (→ *ḍād*), and in the Tamil context obviously reflects borrowing through Indian Ocean networks. More surprising is that /ḍ/ and /ḏ̣/ have identical reflexes, as most languages exhibiting a lateral reflex of /ḍ/ treat the two phonemes differently. Yet, it is possible that there were originally different reflexes of /ḍ/ and /ḏ̣/, one represented by /l/ or /ḷ/ and one by the retroflex approximant /ḻ/ [ɻ]. From the 13th century onward, the phoneme /ḻ/ disappeared from spoken Tamil, merging most commonly with /ḷ/, /l/, or /y/ (cf. Andronov 2003:39, 86), thus obscuring the distinction between the reflexes of /ḍ/ and /ḏ̣/. Occasional use of /ḻ/ for both /ḍ/ and /ḏ̣/, e.g. in the divine name *Kapīḻu* 'Guardian' (< *(al-)ḥafīḏ̣*), and in the Kanniyakumari dialect word *hāḻir* 'presence' (< *ḥāḍir*), may corroborate this scenario. Whatever the case, it is probable that Tamil is the source of those Arabic loanwords in Malay that exhibit /l/ for /ḏ̣/ (→ Indonesian/Malay).

In contrast to these lateral reflexes, words borrowed through Persian or Urdu tend to be written with the *grantha* letter *j*, e.g. *kāji* 'judge' (< *qāḍī*). This reflex is especially common in administrative terms, most of which derive from Indo-Persian vocabulary. In rare cases, one also encounters /t/ for /ḍ/, obviously influenced by contemporary Arabic pronunciation.

2. MORPHOLOGICAL ASSIMILATION

The vast majority of Arabic loanwords in Tamil are nouns. Rather than borrowing verbs, nouns are combined with a Tamil verb to produce a new verb phrase, e.g. in Kanniyakumari dialect *hāḻir ā-* 'to become present, appear'. Words are usually borrowed in the pausal form without article. The common word-final *-u* (rarely *-i*) is not a reflex of Arabic case endings but rather a paragogic vowel added to avoid phonotactically restricted final consonants, e.g. *napucu* 'desire, lust; soul' (< *nafs* 'self, soul'), *oki* 'revelation' (< *waḥy*). This phenomenon shows great similarities with comparable forms in Malay and other Southeast Asian languages (cf. Versteegh 2003; → Acehnese; → Indonesian/Malay). That there is no obvious explanation for this phenomenon in Southeast Asian languages, while it is clearly phonotactically motivated in Tamil, makes it likely that Tamil was the source of the respective Arabic loanwords in Malay, as was already suggested by Bausani (1971:475, 477, n. 11) with regard to the same phenomenon in Persian loans. Tamil has both *-ā* and *-attu* as reflexes of the Arabic feminine ending *-a(t)*. Which of the two reflexes is used seems to be tied to the particular loanword, e.g. *kalimā* 'the profession of faith' (< *kalima* 'word') vs. *cūrattu* 'chapter of the *Qurʾān*' (< *sūra*). There are few semantic doublets of the kind seen in Persian, and those that exist seem to have been borrowed from either → Persian or → Urdu.

While Arabic plurals, personal suffixes, or whole phrases are occasionally encountered, these borrowings are either of a scholarly nature or lexicalized, with the plurals often being treated as singulars to which the Tamil plural suffix *-(k)kaḷ* is added, e.g. *ulamākkaḷ* 'religious scholars' (< *ʿulamāʾ*). Tamil suffixes are freely added to Arabic loanwords, e.g. *cūmaṉ* 'one who causes misfortune, the Devil' (< *šūm* 'misfortune' + 3rd pers. masc. sg. suffix *-aṉ*), *akatāy* 'as one' (< *aḥad* 'one' + suffixed verbal participle *-āy* 'having become'), Kanniyakumari dialect *kāpiricci* 'infidel woman' (< *kāfir* 'infidel' + colloquial 3rd pers. fem. sg. suffix *-(i)cci*). Borrowed nouns in common use which end in *-m* change this to *-ttu* in the oblique case just like native nouns, e.g. *iculām* 'Islam' > *iculāttu* 'Islam's'.

3. SEMANTIC DOMAINS

Most Arabic loanwords in Tamil belong to just a few semantic domains. Probably the single most important one is the domain of Islamic practice and thought, as many of the examples given above attest to. There is also a common tendency for Arabic loanwords to be more religiously circumscribed in meaning in Tamil than they are in Arabic, e.g. *kitāb* 'book' > *kittāpu* 'Islamic religious book; book using Arabic characters' (cf. Thurston 1909:IV, 205). Even though many common institutions and practices of Islam are referred to by Tamil terms in everyday speech by Muslims and non-Muslims alike, often Arabic loans exist side by side with these Tamil words, e.g. Tamil *paḷḷivācal* 'mosque' and its Arabic-derived equivalent *macūti* (< *masjid*).

Another important semantic domain in which Arabic loanwords are not uncommon is administration and bureaucracy. The largest number of Arabic loans in common use by non-Muslim Tamils may belong to this category. Most of this vocabulary is borrowed either directly from Indo-Persian or through Urdu or English, and thus shows reflexes of Persian or Urdu pronunciation, e.g. *jillā* 'district' (< *ḍilʿa* 'side'). While terms for specific administrative divisions, such as *tālu(k)kā* 'subdistrict' (< *taʿalluqa* 'connection'), are generally known only in those countries where the term is used in administration, other loans in this domain are more widespread, e.g. *okkīl*, *vakkīl* 'advocate' (< *wakīl*); *pākki* 'rest, remainder' (< *(al-)bāqī*).

A third domain is constituted by what Bausani (1971:477, n. 2) has called the "sailors' international vocabulary of the Indian Ocean coasts", which includes loans from Persian and Malay in addition to Arabic. Among the Arabic loans are *mālimi*, *mālumi* 'captain, sailor' (< *muʿallim* 'teacher, master') and *campōkku* 'boat' (< *sunbūq*). Interestingly, many of these terms show signs of being borrowed through Persian, e.g. *cukkāṉ* 'rudder, helm' (< *sukkān*), *cukkāṉkiri* 'helmsman' (< Persian *sukkān-gīr*).

Not all Arabic loanwords in Tamil fit into such neat semantic categories. Especially in literary texts, words often seem to be borrowed ad hoc. Colloquial Tamil, especially as spoken by Muslims, contains many Arabic loanwords beyond the confines of religious, administrative, and nautical vocabulary, such as *capar* 'journey' (< *safar*) or *mauttu* 'death' (< *mawt*). This also includes kinship terminology, such as the Shafiʿi Muslim kinship term *ummā* 'mother', a curious mixture of the respective Arabic and Tamil equivalents *ʾumm* and *ammā*. Like many other aspects of Arabic loans in Tamil, the semantic domains covered by Arabic vocabulary require further research (for Arabic loanwords in other Dravidian languages, → Malayalam; → Telugu).

BIBLIOGRAPHICAL REFERENCES

Agesthialingom, S. and S. Sakthivel. 1973. *A bibliography of Dravidian linguistics*. Annamalainagar: Annamalai University.

Andronov, Michail S. 2003. *A comparative grammar of the Dravidian languages*. Wiesbaden: O. Harrassowitz.

Bausani, Alessandro. 1971. "Notes on Persian words in Tamil". *Proceedings of the Second International Conference Seminar of Tamil Studies, Madras – India January 1968*, I, ed. Ronald Eaton Asher, 460–478. Madras: International Association of Tamil Research.

Das, Rahul Peter. 1981. "Zur Transliteration und Transkription des Arabischen im Tamil". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 131.343–348. [Cf. also errata in *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 132.*4*.]

Fanselow, Frank S. 1989. "Muslim society in Tamil Nadu (India): An historical perspective". *Journal Institute of Muslim Minority Affairs* 10:1.264–289.

Mukamatu Ārip Miyāṉ. 1996. *Cīṟāt tamiḻ* [Tamil of the *Cīṟāppurāṇam*]. Pondicherry: Tamiḻnilaip Patippakam.

Nainar, S. Muhammad Husayn. 1941. “Arabic and Persian words in the Tamil language”. *Proceedings and Transactions of the Tenth All-India Oriental Conference Tirupati March, 1940*, 719–723.

——. 1942. *Arab geographers' knowledge of southern India*. Madras: University of Madras.

Shu'ayb, Tayka. 1993. *Arabic, Arwi and Persian in Sarandib and Tamil Nadu: A study of the contributions of Sri Lanka and Tamil Nadu to Arabic, Arwi, Persian and Urdu languages, literature and education*. Madras: Imāmul ʿArūs Trust.

Thurston, Edgar. 1909. *Castes and tribes of southern India*. 7 vols. Madras: Government Press.

Tōppil Muhammatu Mīrāṉ. 2004. *Oru kaṭalōra kirāmattiṉ katai* [The story of a seaside village]. Nagercoil: Kālaccuvaṭu Patippakam. (1st ed., Tirunelveli: Jalīlā Papḷiṣiṅ Havus 1988).

Tschacher, Torsten. 2001. *Islam in Tamilnadu: Varia*. Halle (Saale): Institut für Indologie und Südasienwissenschaften der Martin-Luther-Universität Halle-Wittenberg.

Uwise, M. Mohamed [Uvais, Ma. Mukammatu]. 1976. *Muslim epics in Tamil literature*. Madras: Semmal Publications.

——. 1983. *Tamiḻ ilakkiya aṟapuc col akarāti* [Dictionary of Arabic words in Tamil literature]. Madurai: Madurai Kamaraj University.

Vaidyanathan, S. 1958. “Loan-words of Perso-Arabic origin in modern Tamil”. *Bulletin of the Deccan College Research Institute* 19:1–2.152–154.

Versteegh, Kees. 2003. “The Arabic component of the Indonesian lexicon”. *Rintisan kajian leksikologi dan leksikografi* [Preliminary studies of lexicology and lexicography], ed. Lilie Suratminto and Munawar Holil, 216–229. Jakarta: Fakultas Ilmu Pengetahuan Budaya Universitas Indonesia.

Vinson, M. Julien. 1895. “L'écriture arabe appliquée aux langues dravidiennes”. *Journal Asiatique* 9:5.153–161.

TORSTEN TSCHACHER
(University of Heidelberg)

Tamyīz

The term *tamyīz* is used by the Arab grammarians to refer to one of the functions of the accusative. Generally translated as ‘specification’ or ‘specifier’, *tamyīz*, especially in the earlier sources, has the synonyms *mumayyiz, mufassir, tafsīr, mubayyin*, and *tabyīn* (see, however, Ibn Šuqayr, *Muḥallā* 15–16, where *tafsīr* is reserved for the accusative after numbers; cf. Mubarrad, *Muqtaḍab* II, 144, 164, 173, III, 32, 91, 259, where *tabyīn* and *tamyīz* are distinguished, but not uniformly).

Although Sībawayhi (d. 180/796) discusses *tamyīz* in various places of his *Kitāb* (I, 204–211, II, 117–119, 156–182), he does not use any particular term to refer to its function. He does, however, use verbs such as *yūṣaf, ixtaṣaṣta, lam tubhim, bayyanta, yubayyin, yufassir* (II, 121, 174, 176) and the noun *tafsīr*, but not as a technical term (II, 175–178), to describe that function. *Tafsīr* becomes more of a technical term with Farrā' (d. 207/822) and 'Axfaš (d. 215/830; see Farrā', *Maʿānī* I, 226, II, 138, 166; and 'Axfaš, *Maʿānī* II, 395, 460). As a technical term, *tamyīz* is first encountered in Mubarrad's (d. 285/898) *Muqtaḍab* (see above; cf. Carter 1981:383; Owens 1990:127, 134), and has since gained supremacy over the other terms, which were, however, still recognized by the later authors as possible alternatives (e.g. Ibn ʿAqīl, *Šarḥ* 295; Suyūṭī, *Hamʿ* I, 250). As far as the types of *tamyīz* and the issues related to their syntax and regimen are concerned, they are almost exhausted by Sībawayhi, albeit in disparate parts of the *Kitāb*. The later authors did not have much to add to the basic elements surveyed by Sībawayhi, and their contribution concentrated mainly on expanding the corpus of examples or *šawāhid*, elaborating on questions related to regimen and syntax, and organizing the material in one coherent chapter usually placed next to the chapter on → *ḥāl* in the section dealing with nouns in the accusative which are considered to be ‘redundant elements’ (*faḍalāt*).

As a distinct category, *tamyīz* had to be distinguished from both → *ṣifa* and → *ḥāl*. Sībawayhi uses *waṣafa* and its derivatives to describe a general class which comprises *ṣifa, ḥāl*, and *tamyīz*, the latter only implicitly, but he also uses the term *ṣifa* more specifically to denote qualifiers (Owens 1990:65–66; Versteegh 1993:4–6). To differentiate between the functions of *tamyīz* and *ṣifa*, Sībawayhi cites the construction *hāḏā rāqūdun xallan* ‘this is a vessel of vinegar’, where *xallan* is *tamyīz*, to cite later terminology, and says that to use *xallun* as a *ṣifa* in this sentence would be *qabīḥ* ‘ugly’, hence the accusative. The later grammarians mention several distinctive features of *tamyīz* and *ḥāl*. Thus, whereas the former dispels the ambiguity of a *ḏāt* ‘entity, being, whether animate or otherwise’, the latter dispels the ambiguity of the exterior aspect (*hay'a*) of a *ḏāt* ('Astarābāḏī, *Šarḥ* I, 215–216). Furthermore, *tamyīz*, unlike *ḥāl*, may not be used as corroborative (*mu'akkid*), may occur only once in a sentence, and may only be a noun, i.e. not a sentence or prepositional phrase ('Ušmūnī, *Šarḥ* I, 226–267; Suyūṭī, *Hamʿ* I, 252–253). Semantically, *tamyīz* ordinarily implies *min*, and so the above example may be modified to become *hāḏā rāqūdun min xallin* with no apparent change of meaning. Morphologically, it is invariably an indefinite singular noun.

At the syntactic level, the grammarians distinguish between two types of *tamyīz*. The first is called *tamyīz an-nisba* or *tamyīz al-jumla* 'specification of relationship' and explains the connection between a verb and its subject or object, as in *taṣabbaba zaydun ʿaraqan* 'Zayd dripped with sweat', where the specifier, *ʿaraqan*, explains the connection or relationship between 'dripping' and 'Zayd'. This type is generally interpreted as a transformation from an original sentence, in this case *taṣabbaba ʿaraqu zaydin*. Of this type also is specification following an elative, as in *ʾanta ʾaʿlā manzilan* 'you are higher in station', which supposedly originates from *ʾanta ʿalā manziluka*. The second type is known as *tamyīz al-mufrad* 'specification of a single word' and typically follows words expressing quantity, weight, area, and the like, as in *raṭlun ʿasalan* 'a rotl-weight of honey' and *ḏirāʿun ʾarḍan* 'a cubit of land'. This type also includes specification of number, namely specification after 11 to 99, including decades (*ʿuqūd*) 20 to 90; e.g. *ʿišrūna dirhaman* 'twenty dirhams' (see Carter 1972:485–496 for the significance of this construction in Sībawayhi's *Kitāb*). Also worth noting is specification following exclamatory expressions, as in *li-l-lāhi darruhu ʿāliman* 'what a fine scholar he is!', the interrogative noun *kam*, as in *kam rajulan ʿindaka* 'how many men are there in your house?', and indefinite *kaḏā*, as in *ʿindī kaḏā dirhaman* 'I have so-and-so many *dirhams*' (for further examples, see Reckendorf 1921:94–97; Wright 1981:II, 122–128).

BIBLIOGRAPHICAL REFERENCES

Primary sources

ʾAstarābāḏī, *Šarḥ* = Raḍī d-Dīn Muḥammad ibn Ḥasan al-ʾAstarābāḏī, *Šarḥ Kāfiyat Ibn al-Ḥājib fī n-naḥw*. 2 vols. Istanbul, 1310 A.H.

ʾAxfaš, *Maʿānī* = ʾAbū l-Ḥasan Saʿīd ibn Masʿada al-ʾAxfaš al-ʾAwsaṭ, *Maʿānī l-Qurʾān*. Ed. Fāʾiz Fāris al-Ḥamad. 2nd ed. 2 vols. Kuwait, 1981.

Farrāʾ, *Maʿānī* = ʾAbū Zakariyyāʾ Yaḥyā ibn Ziyād al-Farrāʾ, *Maʿānī l-Qurʾān*. Ed. Muḥammad ʿAlī an-Najjār. 3 vols. Cairo: Dār al-Kutub al-Miṣriyya, 1955–1972.

Ibn ʿAqīl, *Šarḥ* = Bahāʾ ad-Dīn ʿAbdallāh Ibn ʿAqīl, *Šarḥ Ibn ʿAqīl ʿalā ʾAlfiyyat Ibn Mālik*. Ed. Ramzī Munīr Baʿalbakī. Beirut: Dār al-ʿIlm li-l-Malāyīn, 1992.

Ibn Šuqayr, *Muḥallā* = ʾAbū Bakr ʾAḥmad ibn al-Ḥasan Ibn Šuqayr, *al-Muḥallā, wujūh an-naṣb*. Ed. Fāʾiz Fāris. Beirut: Muʾassasat ad-Dirāsāt; Irbid: Dār al-ʾAmal, 1987.

Mubarrad, *Muqtaḍab* = ʾAbū l-ʿAbbās Muḥammad ibn Yazīd al-Mubarrad, *al-Muqtaḍab*. Ed. Muḥammad ʿAbd al-Xāliq ʿUḍayma. 4 vols. Cairo: Dār at-Taḥrīr, 1965–1968.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. Ed. ʿAbd as-Salām Muḥammad Hārūn. 5 vols. Cairo: al-Hayʾa al-Miṣriyya al-ʿĀmma, 1977.

Suyūṭī, *Hamʿ* = Jalāl ad-Dīn ʾAbū l-Faḍl ʿAbd ar-Raḥmān ibn ʾAbī Bakr as-Suyūṭī, *Hamʿ al-hawāmiʿ šarḥ Jamʿ al-jawāmiʿ fī ʿilm al-ʿarabiyya*. 2 vols. Cairo: Maṭbaʿat as-Saʿāda, 1327 A.H.

ʾUšmūnī, *Šarḥ* = ʾAbū l-Ḥasan ʿAlī ibn Muḥammad al-ʾUšmūnī, *Šarḥ al-ʾUšmūnī ʿalā ʾAlfiyyat Ibn Mālik al-musammā Manhaj as-sālik ʾilā ʾAlfiyyat Ibn Mālik*. Ed. Muḥammad Muḥyī d-Dīn ʿAbd al-Ḥamīd. 3 vols. Cairo: Dār al-Kitāb al-ʿArabī, 1955.

Secondary sources

Carter, Michael G. 1972. "'Twenty dirhams' in the *Kitāb* of Sībawaihi". *Bulletin of the School of Oriental and African Studies* 35.485–496.

——. 1981. *Arab linguistics: An introductory classical text with translation and notes*. Amsterdam: J. Benjamins.

Owens, Jonathan. 1990. *Early Arabic grammatical theory: Heterogeneity and standardization*. Amsterdam and Philadelphia: J. Benjamins.

Reckendorf, Hermann. 1921. *Arabische Syntax*. Heidelberg: C. Winter.

Versteegh, Kees. 1993. *Arabic grammar and Qurʾānic exegesis in early Islam*. Leiden: E.J. Brill.

Wright, William. 1981. *A grammar of the Arabic language*. 2 vols. 3rd ed. Rev. by W. Robertson Smith and M.J. de Goeje. Beirut: Librairie du Liban.

RAMZI BAALBAKI (American University of Beirut)

Tanāzuʿ

1. DEFINITION

Tanāzuʿ lit. 'conflict' is a syntactic phenomenon that deals with two coordinated verbs competing for assigning case to a shared constituent, as in (1).

(1) [$_1$ ḍaraba-nī] wa-[$_2$ ḍarab-tu **zayd-an**]
hit-me and-hit-I Zayd-Acc
'Zayd$_i$ hit me and I hit him$_i$'

The shared constituent (in bold) in (1) serves two distinct grammatical functions: it serves as a subject for the first verb (henceforth clause 1) and as an object for the second verb (henceforth clause 2). However, **zayd-an** belongs to clause 2, as it carries the accusative case, assigned under adjacency by the immediately preceding verb. Sentences such as in (1) have raised an intense debate among three schools of Arabic linguistic thought: the Basran and the Kufan, and that of Ibn Maḍāʾ al-Qurṭubī (d. 592/1196). The central issue for the Basran and the Kufan linguists is which verb assigns case to the shared constituent. For Ibn Maḍāʾ, who totally rejects the entire concept of case assignment (*ʾiʿmāl*), the question is which verb structurally licenses the shared constituent, in other words, which verb subcategorizes for the shared constituent. Before surveying each school's answer to this question, a quick look at some basic differences between English and Arabic with respect to *tanāzuʿ* is in order.

2. *Tanāzuʿ* between English and Arabic

English exhibits a similar phenomenon, known as the shared constituent constructions (Radford 1997:105–107), where the shared constituent, unlike Arabic, *must* serve the same grammatical function in the two coordinated clauses, as in (2).

(2) $[_1$ John is buying e_i] and $[_2$ Mary is selling a **house**$_i$]
(3) *I hit e_i and **John**$_i$ hit me

The shared constituent in (2) serves as a direct object in both clauses. The sentence in (3) is ungrammatical, as the shared constituent **John** serves two distinct grammatical functions: it functions as the object in clause (1) and the subject in clause (2). The facts in English are unproblematic, because the shared constituent triggers the deletion of its counterpart in clause 1 under identity. There exists another basic difference between English and Arabic, which can be illustrated by the contrast in (4) and (5).

(4) *Try to appeal* e_i *but try not to beg* ***for clemency***$_i$

(5) *marar-tu* e_i *wa-marra b-ī*
passed-I and-passed by-me
zayd-un$_i$
Zayd
'I passed by Zayd$_i$ and he$_i$ passed by me'

The shared constituent in (4) is a prepositional phrase which triggers the deletion of its counterpart in clause 1 under identity, as clarified by the deep structure in (6).

(6) *Try to appeal* ***for clemency*** *but try not to beg* ***for clemency***.

In contrast, the shared constituent in (5) surfaces as a noun phrase in clause 2 and triggers the deletion of a prepositional phrase in clause 1, as illustrated by the deep structure in (7).

(7) *marar-tu* ***bi-zayd-in*** *wa-marra b-ī* ***zayd-un***

The shared constituent surfaces as a prepositional phrase in clause 1 and a noun phrase in clause 2.

In sum, unlike English, shared constituents in Arabic may belong to distinct phrasal categories and can serve distinct grammatical functions.

3. Linguistic terminology: *ʾIʿmāl* vs. *taʿalluq*

ʾIʿmāl, which partially equals case assignment in Chomsky's framework (1981, 1986, 1995), is the key term used by both the Kufan and the Basran linguists to account for the syntactic behavior of the *tanāzuʿ* constructions (→ *ʿamal*). In contrast, Ibn Maḍāʾ uses *taʿalluq* lit. 'clinging to', which is a structural notion covered by Chomsky's modules of → X-bar theory and Theta theory. X-bar theory deals with a lexical head and its dependents (complements/adjuncts). Theta theory tackles the argument structure of the verb and the thematic roles carried by those arguments. *Taʿalluq* was used by Sībawayhi only to deal with predicative prepositional phrases, as in (8).

(8) *al-walad-u fī l-bayt-i*
the-boy-Nom in the-house-Gen
'The boy is in the house'

To Sībawayhi, a prepositional phrase must be subcategorized by a verb. Hence, he assumes that the deep structure of (8) has a verb which gets deleted at the surface structure level after licensing the prepositional phrase, as in (9).

(9) *al-walad-u* *staqarra* *fī*
the-boy-Nom settled in
l-bayt-i
the-house-Gen

Ibn Maḍāʾ rejects the analysis in (9) and assumes that the predicate in (8) is the prepositional phrase. He extends the structural notion of *taʿalluq* to replace the notion of ‘case assignment’, whose basic criterion for the Basran linguists is adjacency. Consider the famous example in (10) (Mubarrid, *Muqtaḍab* V, 4/37; Ibn al-ʾAnbāri, *ʾInṣāf* 64):

(10) *hāḏā* *juhr-u* *ḍabb-in*
this hole-Nom chameleon-Gen
xārib-in
wrecked-Gen
‘This is a wrecked hole of a chameleon’

The Basran linguists assign the genitive case to *xārib-in* on the grounds of adjacency, i.e., it is assigned the genitive case carried by the immediately preceding segment. In contrast, Ibn Maḍāʾ assumes that it should be assigned the nominative case because it modifies the head noun *juhr*, which carries the nominative case. Ibn Maḍāʾ’s analysis is in agreement with Chomsky’s generative approach, because the modifier should carry the case of the head noun it modifies. In (10), *ḍabbin* is a nominal complement (*muḍāf ʾilayhi*) of the head noun *juhr* ‘hole’. Because of examples like (10), Ibn Madāʾ replaces the notion of *ʾiʿmāl* by the notion of *taʿalluq* as a main criterion in syntactic analysis.

4. SĪBAWAYHI’S TREATMENT OF *TANĀZUʿ* CONSTRUCTIONS

Sībawayhi’s main diagnostic in dealing with *tanāzuʿ* constructions is case, which is assigned under adjacency. Accordingly, the shared constituent in the paradigm in (11) originates in clause 2 because it receives its case from the immediately preceding verb.

(11a) [[$_1$ ḍarab-tu] wa-[$_2$ ḍaraba-nī
hit-I and-hit-me
zayd-un]]
Zayd-Nom
‘I hit Zayd$_i$ and he$_i$ hit me’
(11b) [[$_1$ ḍaraba-nī] wa-[$_2$ ḍarab-tu
hit-me and-hit-I
zayd-an]]
Zayd-Acc
‘Zayd$_i$ hit me and I hit him$_i$’
(11c) [[$_1$ marar-tu] wa-[$_2$ marra bī
passed-I and-passed by-me
zayd-un]]
Zayd-Nom
‘I passed by Zayd$_i$ and he$_i$ passed by me’

For Sībawayhi, the shared constituent triggers the deletion of its counterpart in clause 1 if it is an object or a prepositional phrase. If it is a subject, it is pronominalized as part of the verb.

5. KUFAN TREATMENT OF *TANĀZUʿ* CONSTRUCTIONS

The Kufan linguists assume that the first verb (clause 1) assigns case to the shared constituent in clause 2 (Ibn al-ʾAnbārī, *ʾInṣāf* 61–66), as in (12).

(12a) *ʾakrama-nī*
was.generous-(to) me
wa-ʾakram-tu ***zayd-an***
and-was.generous-I Zayd-Acc
‘Zayd$_i$ was generous to me and I was generous to him$_i$’
(12b) *ʾakram-tu*
was.generous-I
wa-ʾakrama-nī ***zayd-un***
and-was.generous-(to)me Zayd-Nom
‘I was generous to Zayd$_i$ and he$_i$ was generous to me’

Their main argument is based on precedence. The first verb is stronger due to its initial position and, hence, should assign case to the shared constituent at the end of the sentence. The Kufan linguists offer no explanation as to how the initial verb in (12a), for example, with a pronominalized object *-nī*, assigns the accusative case to another noun phrase that serves as an object to another verb. In addition, they reject the concept that a pronoun may precede its coreferential noun phrase (→ cataphora).

Both the Basran and the Kufan linguists agree that the shared constituent originates in clause 2. They disagree as to how it gets its case. The former choose the verb in clause 2 on the grounds of adjacency, whereas the latter choose the verb in clause 1 due to its strength and precedence.

6. Ibn Maḍāʾ’s treatment of *tanāzuʿ*

Unlike the Basran and Kufan linguists, Ibn Maḍāʾ assumes that the shared constituent may originate in clause 2 as well as in clause 1. If it originates in clause 2, it triggers deletion. If it originates in clause 1, it gets moved (extraposed) to the end of the sentence after pronominalizing its counterpart in clause 2. The deletion and the movement analyses are illustrated by (13a) and (13b) (Ibn Maḍāʾ, *Radd* 96).

(13a) [[$_1$ ḍarab-tu e$_i$]   wa-[$_2$ ḍarab**a**-nī
hit-I   and-hit-du-me
az-zayd-ān$_i$]]
the-Zayds-du.Nom
‘I hit the two Zayds$_i$ and they$_i$ hit me’
(13b) [[$_1$ ḍarab-tu e$_i$ wa-[$_2$ ḍarab-**ā**-nī]
hit-I   and-hit-du-me
az-zayd-ayn$_i$]]
the-Zayds-du.Acc
‘I hit the two Zayds$_i$ and they$_i$ hit me’

There are two minimal differences (bolded and underlined) between (13a) and (13b): the case carried by the shared constituent (*ā/ay*) and the vocalic pronominal subject (*a/ā*) carried by the second verb. The case distinction (*ā/ay*) indicates that the shared constituent functions as the subject in clause 2 and as the object in clause 1, in (13a) and (13b) respectively. Accordingly, Ibn Maḍāʾ assumes that the shared constituent in (13a) triggers the deletion of its counterpart in clause 1. In (13b), the shared constituent originates in clause 1 and gets moved (extraposed) to the end of the sentence. (13a) and (13b) have the underlying structure of (14).

(14) [[$_1$ ḍarab-tu   **az-zayd-ayn**$_i$]
hit-I   the-Zayds-du.Acc
wa-[$_2$ ḍaraba-nī   **az-zayd-ān**$_i$]]
and-hit-me   the-Zayds-du.Nom
‘I hit the two Zayds and they hit me’

The surface structure (13a) is derived from (14) via deletion, whereby the shared constituent in clause 2 deletes its counterpart in clause 1. In contrast, the surface structure in (13b) is derived from the deep structure in (14) via two steps: pronominalization and movement (extraposition), as illustrated in (15).

(15a) [[ḍarab-tu   az-zayd-ayn$_i$]
hit-I   the-Zayds-du.Acc
wa-[ḍarab-**ā**$_i$-nī]] (pronominalization)
and-hit-du-me
(15b) [[ḍarab-tu e$_i$ wa-[ḍarab-**ā**-nī]
hit-I   and-hit-du.-me
az-zayd-ayn$_i$]] (movement)
the-Zayds-du.Acc

Ibn Maḍāʾ extends this analysis to di-transitive verbs, which, in the Arabic linguistic tradition, are divided into two categories: verbs that take two nominal complements (two objects), as in (16), and verbs that subcategorize for a direct object and an objective (predicative) complement, as in (17).

(16) *ʾaʿṭay-tu*   *zayd-an*
gave-I   Zayd-Acc
dirham-an (V + indirect object + direct object)
dirham-Acc
‘I gave Zayd a dirham’
(17) *ḏ̣anan-tu*   *zayd-an*
thought-I   Zayd-Acc
marīḍ-an (V + direct object$_i$ + objective complement$_i$)
sick-Acc
‘I thought that Zayd was sick’

As for the *give*-type verbs, the two nominal complements function as shared constituents. If they originate in clause 2, they trigger the deletion of their counterparts in clause 1, as in (18a) below. If they originate in clause 1 and get moved to the end of the sentence, they have the surface structure in (18b).

(18a) [$_1$ ʾaʿṭay-tu e$_i$ e$_x$]   wa-[$_2$ ʾaʿṭā-nī
gave-I   and-gave-me
zayd-un$_i$ **dirham-an**$_x$]
Zayd-Nom   dirham-Acc
‘I gave Zayd$_i$ a dirham and he$_i$ gave me a dirham’

(18b) $[_1$ ʾaʿṭay-tu e_i e_x wa-$[_2$ ʾaʿṯā-nī$_i$-hi$_x$]
gave-I and-(he) gave-me-it
zayd-an$_i$ **dirham-an**$_x$]
Zayd-Acc dirham-Acc

As for the *think*-type verbs in (17), there is a predicative relationship between the direct object and the objective complement, realized by the obligatory agreement between both elements in person, number, and gender. In fact, they form a small (verbless) clause (Stowell 1991), which can show up in isolation as a nominal sentence, as in (19).

(19) *zayd-un* *marīḍ-un*
Zayd-Nom sick-Nom
'Zayd is sick'

The point here is that the secondary predicate in (17) must take the direct object as its own subject, which explains the behavior of the paradigm in (20) and (21) with respect to the phenomenon of *tanāzuʿ* (Ibn Maḍāʾ, *Radd* 97).

(20) $[[_1$ ḏ̣anan-tu $[_2$ wa-ḏ̣ann-ā-**nī**
thought-I and-thought-they.du-me
šāxiṣ-an] **az-zayd-ayn**
gazing-Acc the-Zayds-du.Acc
šāxiṣ-ayn]]
gazing-du.Acc
'I thought that the two Zayds$_i$ were gazing [at me] and they$_i$ thought I was gazing [at them]'

(21) $[_1$ ḏ̣anan-tu wa-$[_2$ ḏ̣ann-ū-**nī**
thought-I and-thought-they-me
šāxiṣ-an] **az-zayd-īn**
gazing-Acc the-Zayds-pl.Acc
šāxiṣ-īn]
gazing-pl.Acc

In (20) and (21) we have movement of the two shared constituents, the direct object and secondary predicate, to the end of the sentence. Deletion is not applicable due to the fact that the secondary predicates in both clauses are not identical in agreement features. Now, contrast (20) and (21) with (22) and (23), where the secondary predicate carries the same agreement features in both clauses.

(22) $[[_1$ ḏ̣anan-tu e_i e_x] wa-$[_2$ ḏ̣anna-**nī**
thought-I and-thought-me
zayd-un$_i$ **šāxiṣ-an**$_x$]]
Zayd-Nom gazing-Acc
'I thought that Zayd was gazing [at me] and Zayd thought that I was gazing [at him]'

(23) $[[_1$ ḏ̣anan-tu e_i e_x wa-$[_2$ ḏ̣anna-**nī**$_i$**-hi**$_x$]
thought-I and-thought-me
zayd-an$_i$ **šāxiṣ-an**$_x$]]
Zayd-Acc gazing-Acc
'I thought that Zayd was gazing [at me] and Zayd thought that I was gazing [at him]'

The sentences in (22) and (23) are derived from the deep structure in (24).

(24) [[ḏ̣anan-tu zaydan šāxiṣ-an] wa-[ḏ̣anna-nī zayd-un šāxiṣ-an]]

It seems that the only diagnostic for deletion with respect to the *think*-type verbs in Arabic is the agreement features, which have to be identical in the two clauses in order to trigger deletion.

7. CONCLUSION

This entry sheds light on the controversial syntactic phenomenon of *tanāzuʿ*. In dealing with it, Sībawayhi uses the theory of case with its adjacency diagnostic. In contrast, Ibn Maḍāʾ uses the structural theory of *taʿalluq*, which is in line with Chomsky's modules of X-bar theory and Theta theory. In fact, the two analyses of Sībawayhi and Ibn Maḍāʾ are complementary in the same sense that the theories of Case, X-bar theory, and Theta theory are an inseparable part of modern linguistic theory.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Ibn al-ʾAnbārī, *ʾInṣāf* = Kamāl ad-Dīn ʾAbū l-Barakāt ʿAbd ar-Raḥmān ibn Muḥammad Ibn al-ʾAnbārī, *Kitāb al-ʾinṣāf fī masāʾil al-xilāf bayna n-naḥwiyyīna l-baṣriyyīna wa-l-kūfiyyīn*, I. Ed. Muḥammad Muhyī d-Dīn ʿAbd al-Ḥamīd. Cairo, 1945.

Ibn Maḍāʾ, *Radd* = ʾAbū l-ʿAbbās ʾAḥmad Ibn Maḍāʾ al-Qurṭubī, *Kitāb ar-radd ʿalā n-nuḥāt*. Ed. Šawqī Ḍayf. 3rd ed. Cairo: Dār al-Maʿārif, 1988. (1st ed. 1947.)

Mubarrid, *Muqtaḍab* = ʾAbū l-ʿAbbās Muḥammad ibn Yazīd al-Mubarrid, *al-Muqtaḍab*. Ed. Muḥammad ʿUḍayma. Cairo: Maṭābiʿ al-ʾAhrām at-Tijāriyya, 1386/1966.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān ibn Qanbar Sībawayhi, *al-Kitāb*, I. Ed. ʿAbd as-Salām Hārūn. Cairo: al-Hayʾa al-Miṣriyya al-ʿĀmma li-l-Kitāb, 1977.

Secondary sources
Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
——. 1986. *Barriers*. Cambridge, Mass.: MIT Press.
——. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.
Radford, Andrew. 1997. *Syntactic theory and the structure of English: A Minimalist Program*. Cambridge: Cambridge University Press.
Stowell, Tim. 1991. "Small clause restructuring". *Principles and parameters in comparative grammar*, ed. Robert Freidin, 182–218. Cambridge, Mass.: MIT Press.

Wafaa Batran Wahba
(Ain Shams University)

Tanwīn

The term *tanwīn* (called in Western terminology → 'nunation') is the *maṣdar* of the verb *nawwana* 'adding an *n*'. At first, it indicated the nasalization of the final vowel of the word, especially in the case ending of the noun (e.g. *rajulu-n*). According to Ibn Yaʿīš (d. 643/1245; *Šarḥ* IX, 29), *tanwīn* became, due to metonymy, the actual name of the final *-n*. Arabic script differentiates this *-n* from the final *n* belonging to the root of the word (e.g. *qaṭana*, *rasana*), which is why *tanwīn* is not represented graphically as a letter of the alphabet. The value of this *-n* in traditional Arabic grammar turns out to be complex and connected with various theories dealing with syntax and morphology.

The term has been used in its technical sense since the first grammatical treatise, Sībawayhi's (d. 177/793) *Kitāb*. Right at the beginning of the *Kitāb*, *tanwīn* is associated with a theory about the hierarchical organization of grammatical categories, according to which some grammatical categories are 'first' (*ʾūlā*) compared to others, e.g. the noun compared to the verb, the indefinite (*nakira*) to the definite (*maʿrifa*), the singular to the plural, the masculine to the feminine. These 'first' categories are consequently 'lighter' (*ʾaxaff*) and 'better established' (*ʾašadd tamakkun*), *tanwīn* serving as the sign (*ʿalāma*) of this, and its dropping as the sign of something they (i.e. the Arabs, the Bedouin) feel as 'heavy' (*Kitāb* I, 6.1–2).

Tanwīn is presented as the mark of a complex attribute shared by several grammatical categories. It is linked to the indefinite (*nakira*), but not exclusively. It is also linked to the syntactic categories (noun/verb), to number and gender, and, above all, to the theory of syntactic endings (*majārī ʾawāxir al-kalām*). The second chapter of the *Kitāb* explains this theory, and it is here that *tanwīn* and its value are mentioned. In this chapter, Sībawayhi examines all word endings and uses the concept of → *ʿamal* to explain the declension of both (imperfect) verbs and nouns (Ayoub 1991b). In this connection, he explains why verbs do not carry *tanwīn*, namely, because they are heavier than nouns (*Kitāb* I, 5.8–9). After each assertion about the 'lightness' and 'heaviness' of a category, he immediately states the consequences of this with regard to *ṣarf*, i.e. the presence of *tanwīn* or *kasra* as a sign of the genitive case, in other words the presence of full nominal declension. The qualifiers (*ṣifat*) *ʾabyaḍ* 'white' and *ʾaḥmar* 'red', for instance, were considered to be 'heavy' (*istaṯqalū*) by the Arabs (*Kitāb* I, 5.13), which is why they have a *fatḥa* in the genitive. The form *ʾafʿal* is lighter as a substantive (*ʾaxaff ʿalay-him*), which is why it is fully declined when indefinite (*yanṣarif fī n-nakira*). The indefinite being 'lighter' than the definite (*Kitāb* I, 5.19–20), the majority of nouns have full declension when they are indefinite. Since the singular is 'more established' than the plural (*Kitāb* I, 5.21), broken plurals do not receive full declension (*lam yaṣrifū*; e.g. *masājid*).

The properties of 'heaviness' and 'lightness' are not defined in the *Kitāb* but rather are justified by syntactic, semantic, and/or morphological arguments. Verbs are 'heavier', first because they are morphologically derived from nouns, and second because of their predicative properties (*Kitāb* I, 5.9–11): "Can't you see that the verb needs the noun, without which there wouldn't be an utterance, whereas the noun can do without the verb?". The 'qualifiers' (*ṣifāt*) "are considered 'heavy' because they resemble the resembling verb [i.e. the imperfect, which resembles agentive nouns]" (*Kitāb* I, 5.13), by the way they operate in the sentence and by their form (*bināʾ*). The feminine is heavier than the masculine, because it is derived from the latter morphologically. However, it is also 'first' semantically, which is why the word *šayʾ* 'thing', which is masculine, can be applied both to the masculine and the feminine. The definite is morphologically derived from the indefinite.

Sībawayhi's distinctions recall the markedness theory of structuralist linguistics, in which the plural, the feminine, and the definite are marked categories, marked compared to the singular, the masculine, and the indefinite. The unmarked category has a generic value in addition to the specific value it has as the counterpart of the marked category (e.g., the French masculine pronoun *ils* refers both to ‘men’ and to ‘men and women’). It therefore has a wider distribution. Sībawayhi reasons in the same way for *šayʾ* ‘thing’. This parallel is confirmed in *Kitāb* II, 22.6–7, where *ʾawwal* is synonymous with *ʾaṣl* (*al-ʾašyāʾ kullu-hā ʾaṣlu-hā t-taḏkīr fa-t-taḏkīr ʾawwal*). Now, the notion of *ʾaṣl*, at least in some of its aspects, bears some resemblance to the markedness theory (Owens 1986:226). This relationship is also found in Ibn Jinnī (d. 392/1002). In a chapter devoted to lexical semantics (*Xaṣāʾiṣ* III, 82), he extends the idea of *ṣarf*, *ʾašadd taṣarruf*, from syntax to semantics. Besides, he associates *ʾawwal* with *ʾaṣl* and *tanwīn* with *farʿ*. Nouns have a greater capability to change (*ʾašadd taṣarruf*) since they are first and thus have to be considered the *ʾaṣl*.

The category of ‘what is better established and lighter’, and consequently its marker, the *tanwīn,* is linked to a complex theory of grammatical categories, whose ultimate model is, nevertheless, the noun and the verb. Apparently, this is what the following generalization asserts: “Every [category] where the *ṣarf* has been dropped (*mā turika ṣarfu-hu*) is likened to the verb (*muḍāraʿ bi-hi l-fiʿl*) insofar as it is not as well established as other [categories], following the example of the verb, which is not as well established as the noun” (*Kitāb* I, 6.5–6).

On several occasions in the *Kitāb* (I, 2.10; I, 270; II, 1ff.; II, 52; etc.), Sībawayhi comes back to the question of the *tanwīn*, highlighting new aspects, in particular the complementary distribution of the genitive and the *tanwīn* (**kitābu-n ar-rajul*).

In later grammar, the *tanwīn* is dealt with in the chapter on the *ḥurūf al-maʿānī*. The *Kitāb*’s key ideas are repeated, but in a more rigid classification and from a more syntactic perspective, bringing together the different functions of *tanwīn* that are scattered throughout the *Kitāb*. Apart from *tanwīn at-tarannum*, which affects rhymes (whether in nouns or verbs), so as to produce a musical effect, four types of *tanwīn* are distinguished (Zajjājī, *ʾĪḍāḥ* 97–99; Versteegh 1995:168–176; Ayoub 1991a:152–155):

i. *Tanwīn al-makāna/at-tamakkun*
(e.g. *rajulu-n/rijālu-n*)

The notion of *makāna* or *tamakkun* is generally understood in connection with the theory of syntactic categories, following the *Kitāb*. But the notion of ‘heaviness and lightness’, linked with *ʾamkan* in the *Kitāb* and in Ibn Jinnī (*Lumaʿ* 94), has disappeared. According to Ibn Yaʿīš (*Šarḥ* IX, 29), the “*makāna* means that a noun is well established in its group, with no resemblance to the particle – in that case it will be undeclinable – nor to the verb – in that case it will be partially declinable [diptotic]”. The emphasis is placed on the theory of syntactic endings: Ibn Yaʿīš states that this *tanwīn* is there to differentiate what is fully declinable from what is not. For al-ʾAstarābāḏī (d. 686/1287; *Kāfiya* I, 13), *tamakkun* means that the noun is declinable (*kawn al-ʾism muʿraban*). He adds that nouns have a specific mark showing their ability to receive this ending, unlike verbs, for which this ending is only a contingent property. Thus, *tanwīn* becomes the marker of declension, which is an esssential attribute of the noun. Elsewhere, *tanwīn* is presented as a specific feature of the noun, along with the definite article, the genitive, and the ability to be a logical subject (*Kāfiya* I, 12).

ii. *Tanwīn at-tankīr* ‘*tanwīn* of indefiniteness’
(e.g. *marartu bi-ʿamrawayhi wa-ʿamrawayhin ʾāxara*)

This *tanwīn* characterizes the indefinite in some indeclinable nouns, thus confirming that in Classical theory, *tanwīn* is linked, though not restricted, to the idea of indefiniteness. It is the distinctive feature of indeclinable nouns, with the exception of diptotes (the *tanwīn* of *ʾibrāhīmi-n* in *marartu bi-ʾibrāhīma wa-ʾibrāhīmi-n ʾāxara* is analyzed as a *tanwīn at-tamakkun*). It also occurs in some interjections (→ *ism al-fiʿl*) like *ʾīhi* ‘what’s more?’ and *ṣahi* ‘be quiet’, interjections which rapidly fell into disuse due to their oral character, as shown by the controversy reported by Ibn Yaʿīš. Whereas al-ʾAsmaʿī (d. 213/828; *Šarḥ* IV, 71; cf. Ṯaʿlab [d. 291/904], *Faṣīḥ* 22.6) seems to think that *ʾīhin* is the only form employed by the ‘Arabs’ addressing someone, Ibn Jinnī (*Šarḥ* IX, 31, n.) and the Basran grammarians believe that speakers say *ʾīhi* when asking the person addressed to elaborate on a given subject. On the other hand, to say *ʾīhin* is to address someone to make them talk without any particular subject in mind. In modern terminology, in the latter enunciative situation, the reference of the conversation is indefinite, hence the use of *tanwīn*. However, Ibn Yaʕīš (*Šarḥ* IV, 71) asserts that the form without *tanwīn* (*ʾīhi*) is only justified by *qiyās*. It is rarely used, hence al-ʾAṣmaʕī's reservations.

iii. *Tanwīn al-ʕiwaḍ* '*tanwīn* of compensation'
This *tanwīn* compensates a deletion, for instance in morphophonology the deletion of the final *y* in *jawāri-n*. In syntax, this *tanwīn* 'compensates' for the deletion of the genitive (*muḍāf ʾilay-hi*), in particular in *ḥīnaʾiḏi-n*, *yawmaʾiḏi-n*, etc. (Ibn Yaʕīš, *Šarḥ* III, 29ff.; IX, 30). The deleted complement is a clause, as in Q. 99/1–4: *ʾiḏā zulzilati l-ʾarḍu zilzāla-hā wa-ʾaxrajati l-ʾarḍu ʾaṯqāla-hā wa-qāla l-ʾinsānu mā la-hā yawmaʾiḏ-i-n tuḥaddiṯu ʾaxbāra-hā*, where *yawmaʾiḏi-n* stands for *yawmaʾiḏ zulzilatu l-ʾarḍu zilzāla-hā* 'When earth is shaken with a mighty shaking and earth brings forth her burdens, and Man says "what ails her?", upon that day she shall tell her tidings'. Such sentences highlight the relationship between *tanwīn* and genitive. Note that the word with *tanwīn* takes on definite reference: *ḥīnaʾiḏ-i-n* refers to a precise moment in time, as shown by Arberry's translation.

iv. *Tanwīn al-muqābala* '*tanwīn* of correspondence' (e.g. *muslimāt-u-n* 'Muslim women')
This is the *-n* added as a suffix to feminine external plurals, corresponding to the *-n* of *muslim-ū-na* 'Muslim men'. Actually, the grammarians believe that the *-n* of *muslim-ū-na* (external masculine plural), with a graphic representation by a letter of the alphabet, and the *tanwīn* share a common value (Mubarrad [d. 285/898], *Muqtaḍab* IV, 144; ʾAstarābāḏī, *Kāfiya* I, 14, etc.). The same applies to the *-n* of the dual (Sībawayhi, *Kitāb* I, 3.18, etc.), and it is true that in certain contexts these suffixes *-n* have the same distribution.

The *tanwīn* theory has a real descriptive and explanatory value. It is generally accepted that the value of this *-n* is still an open question. Both in Orientalist theories and modern linguistics the ending *-n* is often analyzed as an indefinite article (Fleisch 1961:271; 1968:39–40) or a marker of indetermination (Blachère and Gaudefroy-Demombynes 1952:200). The difficulty lies, then, in how to account for the existence of definite nouns with *tanwīn,* such as proper nouns (*zaydu-n*), and of indefinite nouns without *tanwīn* (*ʾabyaḍu*). Kouloughli (2001), who analyzes *tanwīn* in common nouns as an indefinite article, tries to resolve these difficulties. However, his solution is still not convincing. He postulates two *tanwīn* homonyms in Arabic, the *tanwīn* of proper nouns (*zaydu-n*) and that of common nouns (*rajulu-n*). There is a 'material identity' between the two tanwīn, and not 'a functional identity' (2001:32–34). But the study gives no explanation of the functional role of the tanwīn in proper nouns; he even asserts (2001:41) that diptotic proper nouns like *ʾahmadu* do not accept *tanwīn* because they are intrinsically definite, and no explanation is given why, in this case, they do not accept the *tanwīn* of *zaydu-n*. In comparison with the Western theories, the Arabic grammarians' theory, which links the value of the *tanwīn* (*ʕalāma li-l-ʾamkan*) to several parameters, gives an elegant explanation for the presence of *tanwīn* in definite nouns (*zaydu-n*): this is because they are nouns, which happen to be 'better established' (*ʾamkan*) than verbs. Indefinite nouns in turn (*rajulu-n*) are 'better established' than definite ones. This easily explains the absence of *tanwīn* in indefinite diptotic words (*ʾabyaḍu*): they have the same form as a verb, and verbs are 'heavier' than nouns. In fact, this theory transforms the value of indefiniteness into a secondary value, appearing only in some contexts. At this point, the grammarians' explanation joins that given by Kuryłowicz, who compared the nunation (i.e. *tanwīn*) with the → mimation of East Semitic Akkadian and the suffix *–ān* of South Arabian. Hypothesizing that *tanwīn* was in Proto-Semitic a declined definite article, Kuryłowicz (1972:130–133) rejects its analysis as an indefinite article in written Arabic, which he regards as nothing but a secondary function of *tanwīn*. In his view, the main function of nunation is to mark the absence of the article *al-*. Depending on the context, this absence has two values, either neutral (or nondefinite) or negative (or indefinite): "Thus the apparently indefinite value of Ar[abic] Nunation is only a special function of the overall value 'non-definite' going back to an original value 'definite'"

(Kuryłowicz 1972:133). Kuryłowicz (1972:133, n. 21) thinks in terms of the markedness theory: the *tanwīn* is an unmarked term, and there is a material identity between the neutral and the negative term. The Arabic grammarians' theory also sheds light on a point of historical evolution. If *-n* were only an element of the system of determinants, it would be difficult to understand why it disappeared with the syntactic endings, unlike the article *al-*. The *Kitāb* enables us to understand this phenomenon, as the *tanwīn* is considered to be an integral part of the system of declension.

The strict distinction of several types of *tanwīn* is a problematic point in the classical theory. For instance, it classifies two *tanwīn* endings (that of *ʿamrawayhi* and that of *ʾibrāhīm*) as two different types, in spite of the intuitive idea that they both refer to indefiniteness. Likewise, it is dfficult to decide whether the *tanwīn* in *kull-u-n qāʾimun* is a *tanwīn ʿiwaḍ*, or a *tanwīn tamakkun*. This is a controversial issue among grammarians (Ibn Yaʿīš, *Šarḥ* IX, 31). Besides, when al-ʾAstarābāḏī (*Kāfiya* I, 12.23–24) understands *tamakkun* in terms of the theory of syntactic endings only, he is compelled to add that the *tanwīn at-tamakkun* (e.g. *rajulu-n*) also refers to indefiniteness (*Kāfiya* I, 12.23–24: *at-tanwīn fī rajul-i-n yufīdu t-tankīr ʾayḍan*), thus adding something new to the classical theory without avoiding the problem.

Nevertheless, the grammarians' theory retains a real explanatory value: for Sībawayhi and for most of the grammarians after him, the notion of *ʾamkan* is based on a theory of grammatical categories recalling basic statements of modern linguistics. In addition, by linking the final *-n* of duals and plurals with the *tanwīn*, this theory shows a great ability to generalize. Finally, a third key idea, present since Sībawayhi's *Kitāb* and best defended by al-ʾAstarābāḏī (*Kāfiya* I, 14.5), is of great interest: the idea that the common denominator of all types of *tanwīn* except *tanwīn at-tarannum* is the mark of the completeness of the noun (*kawnu-hu ʿalāma tamām al-ism*). The theoretical interest of these elements is demonstrated when *tanwīn* is analyzed within the framework of general linguistics (Ayoub 1991a; 1996:149–223). In this framework, syntactic categories are not separated by impenetrable frontiers but rather are thought of as a continuum, which enables us to understand the phenomena of verbalization or grammaticalization in languages. *Tanwīn* represents the lack of any extrinsic determination (Ayoub 1991a:176), and its analysis combines two key ideas of the grammatical tradition: (i) indefiniteness as a secondary function of *tanwīn* and (ii) *tanwīn* as the marker of the completeness of the noun.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

ʾAstarābāḏī, *Kāfiya* = Raḍī d-Dīn Muḥammad ibn al-Ḥasan al-ʾAstarābāḏī, *Šarḥ Kāfiyat Ibn al-Ḥājib fī n-naḥw*. Istanbul, 1893. (Repr., Beirut: Dār al-Kutub al-ʿIlmiyya, 1976.)

Ibn Jinnī, *Lumaʿ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Lumaʿ fī l-ʿarabiyya*. Ed. Ḥusayn Muḥammad Šaraf. Cairo: ʿĀlam al-Kutub, 1979.

——, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. Ed. Muḥammad ʿAlī an-Najjār. Cairo, 1952–1956. (Repr., Beirut: Dār al-Hudā, n.d.)

Ibn Yaʿīš, *Šarḥ* = ʾAbū l-Baqāʾ Muwaffaq ad-Dīn Yaʿīš ibn ʿAlī Ibn Yaʿīš, *Šarḥ al-Mufaṣṣal*. Beirut: ʿĀlam al-Kutub; Cairo: Maktabat al-Mutanabbī, n.d.

Mubarrad, *Muqtaḍab* = ʾAbū l-ʿAbbās Muḥammad ibn Yazīd al-Mubarrad, *Kitāb al-muqtaḍab*. Ed. Muḥammad ʿUḍayma. Cairo: Lajna ʾIḥyāʾ at-Turāṯ al-ʾIslāmī, 1968.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān ibn Qanbar Sībawayhi, *al-Kitāb*. Ed. Hartwig Dérenbourg. Paris: Imprimerie Nationale, 1881. (Repr., Hildesheim and New York: G. Olms, 1970.)

Ṯaʿlab, *Faṣīḥ* = ʾAbū l-ʿAbbās ʾAḥmad ibn Yaḥyā Ṯaʿlab, *Kitāb al-faṣīḥ*. Ed. Jacob Barth. Leipzig: J.C. Hinrich'sche Buchhandlung, 1876.

Zajjājī, *ʾĪḍāḥ* = ʾAbū l-Qāsim az-Zajjājī, *al-ʾĪḍāḥ fī ʿilal an-naḥw*. Ed. Māzin al-Mubārak. Cairo: Maktabat Dār al-ʿUrūba.

SECONDARY SOURCES

Arberry, Arthur J. 1955. *The Koran interpreted*. London: George Allen and Unwin; New York: Mac Millan.

Ayoub, Georgine. 1991a. "La nominalité du nom ou la question du *tanwīn*". *Arabica* 38.151–213.

——. 1991b. "La forme du sens: Le cas du nom et le mode du verbe". *Budapest Studies in Arabic* 3-4.37–87.

——. 1996. *Prédicats, figures, catégories: La phrase nominale en arabe littéraire*. Lille: Septentrion.

Blachère, Régis and Maurice Gaudefroy-Demombynes. 1952. *Grammaire de l'arabe classique*. Paris: Maisonneuve and Larose. 3rd ed. (Repr., 1978; 1st ed., 1937.)

Fleisch, Henri. 1968. *L'arabe classique: Esquisse d'une structure linguistique*. 2nd ed. Beirut: Dar El-Machreq. (1st ed., 1968.)

——. 1990. *Traité de philologie arabe*. I. Beirut: Imprimerie Catholique. (1st ed., 1961.)

Kuryłowicz, Jerzy. 1972. *Studies in Semitic grammar and metrics*. Wrocław: Polska Akademia Nauk.

Owens, Jonathan. 1986. *The foundations of grammar*. Amsterdam and Philadelphia: J. Benjamins.
Versteegh, Kees. 1995. *The explanation of linguistic causes: az-Zağğāğī's theory of grammar, introduction, translation, commentary*. Amsterdam and Philadelphia: J. Benjamins.

Georgine Ayoub (Institut des Langues et Civilisations Orientales)

Tanzania → East Africa

Taqdīr

The usual meaning of *taqdīr* is 'predestination, ordaining, decreeing', semantically connected with *qadr* in the sense of '(God's) decree, fate'. The word *taqdīr* may also be connected with another sense of *qadr*, 'measure, quantity', in which case it means 'measuring; estimation of value, assignment of a value to something'. In this sense it is used, for instance, for assigning the portion of war booty to which each participant in a campaign is entitled (Kofler 1933:384). This latter sense is probably the one that is behind the technical use of the term, for instance in legal theory, where *taqdīr*, apart from its usual meaning of 'estimation, evaluation', may be applied to the notion of 'legal fiction' (Brunschvig 1970–1972:44). *Taqdīr* denotes here a virtual reality and is opposed to *taḥqīq* 'realization', because it gives to something nonexisting the legal status of something existing.

In the course of the development of the Arabic grammatical tradition, *taqdīr* came to be used for the process of restitution of suppressed elements in linguistic utterances (*pace* Lane 1863–1893:VII, 2495, who derives the grammarians' use of the word from Form II of the verb *qaddara* 'to determine, decree' and asserts that in a linguistic context it signified 'to mean something to be supplied or understood'). In linguistic methodology, *taqdīr* is, therefore, the converse of → *ʾiḍmār*. The speaker 'hides' things in speech, and it is the grammarian's task to reconstruct these hidden elements in order to explain the surface structure of the sentence. The most important aim of Arabic grammar is the explanation of the case endings (*ʾiʿrāb*) in the sentence that are produced by the action (→ *ʿamal*) of a visible element in the sentence. If no such element is available, the grammarian must have recourse to an underlying structure in which these elements are made explicit (cf. Versteegh 1994).

In the technical sense of 'reconstruction of an underlying level', *taqdīr* does not belong to the oldest core of grammatical terminology. Originally, the term seems to have been used for the reconstruction of the radical structure of a word. Al-Xalīl (d. 175/791) uses it as a technical term for the assignment of a pattern to a word by means of the *f-ʿ-l* notation (*ʿAyn* I, 170). The word *qayʿūn* 'short-nosed', for instance, is said to be derived from *qaʿn* in the same way that *qayṣūm* 'southern wood [Artemisia abrotanum]' is derived from *qaṣm* (the word for derivation used here is *ištiqāq*); then, al-Xalīl adds (Talmon 1997:171):

> Many things are derived in the same way from nouns; their radicals have been obscured, but they are recognized when they are assigned to a pattern (*wa-naḥwa hāḏihi l-ʾašyāʾ uštuqqat min al-ʾasmāʾ wa-ʾumītat ʾuṣūluhā wa-lākin tuʿrafu ḏālika fī taqdīr al-fiʿl*)

This is similar to the use of *taqdīr* in the *Kitāb Sībawayhi*, where it occurs 23 times (Troupeau 1976:167; Carter 1968:296–297; Baalbaki 1979). Most instances are in a phonological context, for instance when Sībawayhi (*Kitāb* II, 21.5) states that the word *jayʾal* 'female hyena' has the *taqdīr* of *jayʿal*. The probable interpretation of this is that the original /ʾ/ in the word *jayʾal* has the status of the second radical of the word (its *ʿayn*); as a matter of fact, in one instance Sībawayhi connects *taqdīr* directly with the *f-ʿ-l* notation, when he says that the masdar *sūʾ* of the verb *sāʾa* 'to be bad' has the *taqdīr fuʿl* (*Kitāb* II, 231.16). In three cases, the context of *taqdīr* is syntactic (*Kitāb* I, 287.14, 300.14, 301.18). In *Kitāb* I, 301.18, for instance, the sentence *ʿabdullāhi niʿma r-rajula* 'Abdallah, what a man he is!' is said to have the *taqdīr*, though not the *maʿnā* 'meaning', of *ʿabdullāhi ḏahaba ʾabūhu*. Obviously, what is at stake here is the structure of the sentence, or rather the assignment of syntactic functions to the individual words in both sentences (topic, verb, agent). This syntactic use of the term *taqdīr* is new compared to the *Kitāb al-ʿayn*, but still related to its use in phonology: both in syntax and in phonology, *taqdīr* is used to clarify the structure of an utterance or a sentence. In phonology, this implies the specification of the three radicals in a word, in syntax the assignment of syntactic functions to the words in an utterance.

This is not to say that the method of *taqdīr* in the later sense of 'restitution of the underlying level' is absent in the *Kitāb*. Baalbaki (1979:8) correctly points out that "supplying or inserting parts of utterances 'missing' through elision (*iḍmār*) is a common feature of the *Kitāb*", and may in fact be demonstrated in grammarians preceding Sībawayhi, especially Yūnus ibn Ḥabīb (Baalbaki 1979:8, n. 6). This procedure probably goes back to the earliest exegetical treatises on the text of the *Qurʾān* (→ *ʾiḍmār*). The point is, however, that this procedure is not called *taqdīr* in the *Kitāb*, but rather *tamṯīl* (Ayoub 1990) or *tawahhum*, when it is the speaker who restores the missing elements (Baalbaki 1982). Baalbaki (1979) shows how the method of 'suppletive insertion', as he calls it, is linked with the central core of Arabic linguistic theory; through this method, the grammarian restores the 'harmony and hierarchy' that is lacking in the surface utterance.

After Sībawayhi, *taqdīr* for some time retained its meaning of assigning a pattern. Apparently, in Kufan grammar the term was not used: al-Farrāʾ uses *qaddara* only in the sense of God's decreeing something. ʾAbū ʿUbayda (d. 210/825) uses *taqdīr* for the assignment of morphological patterns to words (*Majāz* I, 170.4, 202.2, etc.; cf. II, 153.6 *ʿalā taqdīrihā* 'according to its pattern'). Al-Mubarrad (d. 285/898) uses *taqdīr* in the same sense (e.g. *Muqtaḍab* I, 30.6 *fa-ṣāra taqdīruhu min al-fiʿl lafʿāʾu* 'its pattern becomes *lafʿāʾu* in the notation with *f-ʿ-l*'). It is true that both ʾAbū ʿUbayda and al-Mubarrad sometimes use *taqdīr* in connection with the syntactic structure of a sentence. Al-Mubarrad (*Muqtaḍab* I, 14.1–2, 21.12), for instance, cites a sentence with *ʾan* as the paraphrase of a *maṣdar*, and elsewhere (*Muqtaḍab* I, 24.2), he cites a sentence with *allaḏī* as the paraphrase of a participial construction. Neither grammarian, however, employs *taqdīr* in its later sense of restitution of a suppressed element. For this they both use *ʾiḍmār*, which originally denoted the act of suppressing an element from the utterance by the speaker, but is applied here to the entire process of suppression, including the restitution by the grammarian.

Eventually, *taqdīr* became the central term in linguistic theory, but it was seldom discussed explicitly. Ibn Jinnī (d. 392/1002) uses the term in connection with the phonological analysis of weak verbs. He warns his readers (*Xaṣāʾiṣ* I, 256) that when grammarians analyze the verb *qāma* as /qawama/ and call this its underlying form (→ *ʾaṣl*), this does not mean that such a form was actually spoken at any time (cf. Guillaume 1981). Linguistic reconstruction involves the undoing of the action of any linguistic cause (→ *ʿilla*) in order to find the form the word would have were it not for the cause affecting it. Since these causes represent necessary constraints of Arabic linguistic structure, the reconstructed form may even turn out to be unpronounceable, for instance in the word *samāʾ* 'heaven', whose underlying form (*taqdīr*) is /sama""/, with two consecutive *ʾalif*s (*Xaṣāʾiṣ* I, 259). Ibn Jinnī explains the use of *taqdīr* as a necessary corollary of the speaker's need to be concise: the grammarian reconstructs the actual utterance by reinstalling the deleted elements (*Xaṣāʾiṣ* II, 273ff.; Méhiri 1973:368). In this reconstruction there is an element of choice or arbitrariness; in fact, Ibn Jinnī often gives his own analysis and then says *wa-ʾin šiʾta kāna taqdīruhu* 'but if you wish, its *taqdīr* is…', leaving room for alternative analyses (e.g. *Xaṣāʾiṣ* II, 362.11, 363.6).

With later grammarians, *taqdīr* simply came to mean the virtual level of speech, as opposed to the actual utterance (Carter 1981:35), as for instance in the definition of 'declension' by Ibn ʿUṣfūr (d. 670/1271): "a change in the ending of the word as a result of different operators operating on it, either overtly or virtually" (*taġayyur ʾāxir al-kalima li-xtilāf al-ʿawāmil ad-dāxila ʿalayhā lafḏ̣an ʾaw taqdīran; Šarḥ* I, 102.9–10). Here, *taqdīr* indicates everything that is not present in the actual utterance, or, as Peled (1992:95) formulates it, "The *maʿnā*-structure is recovered from the *lafẓ*-structure by a process referred to as *taqdīr*". Because of this connection with the meaning of the sentence, *taqdīr* becomes almost synonymous with → *maʿnā*. Another term that is often used in connection with *taqdīr*, and sometimes almost synonymous with it, is → *ʾaṣl*, which denotes the reconstructed underlying level.

In spite of the role of *taqdīr* in the reconstruction of the *maʿnā* level of the utterance, most grammarians handled this method without feeling obliged to provide information about the semantic aspect, which was typically taken for granted because of the inborn knowledge of the native speaker (Gully 1995:207). Unlike these grammarians, Ibn Hišām (d. 761/1359) believed that there was a much more intimate relationship between structure and semantics, and this affected his use of the term *taqdīr*. He poses several conditions for a successful reconstruction of the underlying level. In the first place, the reconstructed element must be recoverable from the context, and the suppressed/ellipted element must be known immediately (Gully 1995:215): "Good elision is when the deleted item is known immediately at the point of its reconstruction" (*ḥusn al-ḥaḏf ʾan yuʿlama ʿinda mawḍiʿ taqdīrihi; Muġnī* II, 449).

A further condition on successful *taqdīr* is that the number of ellipted elements be as small as possible. Ibn Hišām formulates this by saying that one should prefer a reconstruction that is "less in terms of reconstruction" (*ʾaqallu taqdīran*), and he motivates this by explaining that such a reconstruction "reduces the violation of the (underlying) original structure" (*taqillu muxālafat al-ʾaṣl; Muġnī* II, 615–617; Gully 1995:216).

Finally, the inserted element on the underlying level should be as much as possible in accordance with the surface level. Gully (1995: 216–217) illustrates this condition with the example *zaydan ḍarabtu-hu* 'Zayd [acc.], I hit him', in which the accusative *zaydan* needs to be explained. In this case, it is impossible to posit a change of word order (*taqdīm wa-taʾxīr*) because that would lead to **ḍarabtu-hu zaydan*, which is not acceptable because the verb *ḍarabtu* would then have two objects. Therefore, the reconstructed (*muqaddar*) sentence must be *ḍarabtu zaydan ḍarabtuhu*, where the choice of the first verb is determined by the form of the actual utterance. Only when there is a structural or semantic obstacle to this solution is it allowed to choose a different verb, e.g. in *zaydan marartu bi-hi* 'Zayd [acc.], I passed him', where the verb governs its object through the preposition *bi-*. The only solution here is to reconstruct with a different verb, e.g. *jāwaztu zaydan marartu bi-hi* 'I overtook Zayd, I passed him' (Ibn Hišām, *Muġnī* II, 448).

It is precisely this kind of reconstruction that critics of the linguistic method of *taqdīr* rejected. Ibn Maḍāʾ (d. 592/1196), in his refutation of the grammarians, believes that by reconstructing an underlying meaning, the grammarians speculate unduly about the intention of the speaker, especially when they are dealing with Qurʾānic texts (Arnaldez 1956:90). One of the examples of the unwarranted speculations of the grammarians he quotes is the one dealt with above, *ʾa-zaydan ʾakramta-hu* 'was it Zayd you honored?' (*Radd* 86). In his view, there is no compelling reason to posit an underlying verb to explain the accusative because this is the way the Arabs speak, and there is no need for any 'suppletive insertion' here. He comments (*Radd* 87.2) that such an insertion is just 'speculation and fantasy' (*taqdīr wa-taxyīl*), using the word *taqdīr* in a deprecatory sense.

The operation of the grammarian by which the underlying level of the utterance is reconstructed out of the actual utterance bears some resemblance to the method of modern transformational linguistics. An explicit comparison of the two methods is carried out by Owens (1988) and Gruntfest (1984). They both point out that Arabic grammar did not express the relationship between surface structure and underlying level in terms of a derivation or a process but rather used the underlying level as an interpretation of the surface level. Still, Gruntfest (1984) maintains that the Arabic grammarians were in fact transformationalists *avant la lettre*, but Owens is much more cautious in his conclusions. In his view (1988:196–198), there are more differences between the two frameworks. The Arabic grammarians relied on pragmatic context as much as on textual context to condition deletion. Besides, the aim of *taqdīr* was the recovery of deleted items, whereas the aim of transformations is the breakdown of complete utterances. Deletion in Arabic grammar is always seen as proceeding from a speaker, rather than obeying specific rules, and accordingly, there were no formal conditions of recoverability. It must be pointed out here that in some later versions of Arabic grammar, grammarians did attempt to bring in a certain degree of formality. In particular, Ibn Hišām's conditions on acceptable *taqdīr* (see above) go some way toward formalizing the process.

Owens' conclusion (1988:198) sums up nicely the relationship between Arabic grammar and transformational linguistics:

> In the final analysis, these are differences of emphasis more than differences of substance because the basic motivations behind the use of deletions [are] the same: the desire to maintain an overall structural coherency in the grammar and to derive the correct meaning.

The emphasis in this statement should be on the wish to maintain structural coherency (Baalbaki 1979), rather than on the derivation of 'correct meaning', because the latter does not seem to be a major motivation in transformational linguistics, and it certainly was not what the Arabic grammarians had in mind: they knew what the correct meaning was and just needed an underlying level reached through *taqdīr* in order to explain the fact that this correct meaning could also be expressed by a 'deficient' surface utterance.

Bibliographical references

Primary sources

ʾAbū ʿUbayda, *Majāz* = ʾAbū ʿUbayda Maʿmar ibn al-Muṯannā at-Taymī, *Majāz al-Qurʾān*. Ed. Muḥammad Fuʾād Sazkīn [Fuat Sezgin]. 2 vols. Cairo: Muḥammad Sāmī ʾAmīn al-Xānjī, 1954.

Ibn Hišām, *Muġnī* = Jamāl ad-Dīn ʾAbū Muḥammad ʿAbdallāh ibn Yūsuf Ibn Hišām, *Muġnī l-labīb ʿan kutub al-ʾaʿārīb*. Ed. Muḥammad Muḥyī d-Dīn ʿAbd al-Ḥamīd. N.p., n.d.

Ibn Jinnī, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ʿUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. Ed. Muḥammad ʿAlī an-Najjār. Cairo, 1952–1956. (Repr., Beirut: Dār al-Hudā, n.d.)

Ibn Maḍāʾ, *Radd* = ʾAbū l-ʿAbbās ʾAḥmad Ibn Maḍāʾ al-Qurṭubī, *Kitāb ar-radd ʿalā n-nuḥāt*. Ed. Šawqī Ḍayf. Cairo: Dār al-Fikr al-ʿArabī, 1947.

Ibn ʿUṣfūr, *Šarḥ* = ʾAbū l-Ḥasan ʿAlī ibn Muʾmin Ibn ʿUṣfūr al-ʾIšbīlī, *Šarḥ Jumal az-Zajjājī*. Ed. Ṣāḥib ʾAbū Janāḥ. 2 vols. Baghdad: Dār al-Kutub li-ṭ-Ṭibāʿa wa-n-Našr, 1980.

Mubarrad, *Muqtaḍab* = ʾAbū l-ʿAbbās Muḥammad ibn Yazīd al-Mubarrad, *al-Muqtaḍab*. Ed. Muḥammad ʿAbd al-Xāliq ʿUḍayma. 4 vols. Cairo: Dār at-Taḥrīr, 1965–1968.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. 2 vols. Bulaq, 1316 A.H. (Repr., Baghdad: Al-Muthanna Library, n.d.)

Xalīl, *ʿAyn* = ʾAbū ʿAbd ar-Raḥmān al-Xalīl ibn ʾAḥmad al-Farāhīdī, *Kitāb al-ʿayn*. Ed. Mahdī al-Maxzūmī and ʾIbrāhīm as-Samarrāʾī. 8 vols. Beirut: Muʾassasat al-ʾAʿlamī li-l-Maṭbūʿāt, 1988.

Secondary sources

Arnaldez, Roger. 1956. *Grammaire et théologie chez Ibn Ḥazm de Cordoue: Essai sur la structure et les conditions de la pensée musulmane*. Paris: J. Vrin.

Ayoub, Georgine. 1990. "De ce qui 'ne se dit pas' dans le Livre de Sībawayhi: La notion de *tamṯīl*". *Studies in the history of Arabic linguistics,* II, ed. Michael G. Carter and Kees Versteegh, 1–15. Amsterdam and Philadelphia: J. Benjamins.

Baalbaki, Ramzi. 1979. "Some aspects of harmony and hierarchy in Sībawayhi's grammatical analysis". *Zeitschrift für Arabische Linguistik* 2.7–22.

——. 1982. "*Tawahhum*: An ambiguous concept in early Arabic grammar". *Bulletin of the School of Oriental and African Studies* 45.233–244.

Brunschvig, Robert. 1970–1972. "De la fiction légale dans l'Islam médiéval". *Studia Islamica* 32.41–51.

Carter, Michael G. 1968. *A study of Sibawaihi's principles of grammatical analysis*. Ph.D. diss., University of Oxford.

——. 1981. *Arab linguistics: An introductory classical text with translation and notes*. Amsterdam: J. Benjamins.

Gruntfest, Yaakov. 1984. "Medieval Arabic grammarians: First transformationalists?". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 134. 226–236.

Guillaume, Jean-Patrick. 1981. "Le statut des représentations sous-jacentes en morphophonologie d'après Ibn Ǧinnī". *Arabica* 28.222–241.

Gully, Adrian. 1995. *Grammar and semantics in medieval Arabic: A study of Ibn-Hisham's* 'Mughni l-Labib'. Richmond: Curzon Press.

Kofler, Hans. 1933. "Handbuch des islamischen Staats- und Verwaltungsrechtes von Badr ad-Dîn Ibn Ǧamâʿah, herausgegeben, übersetzt und mit Anmerkungen versehen". *Islamica* 6.349–414.

Lane, Edward William. 1863–1893. *An Arabic-English lexicon, derived from the best and most copious Eastern sources*. 8 vols. London and Edinburgh: Williams and Norgate. (Repr., Beirut: Librairie du Liban, 1980.)

Méhiri, Abdelkader. 1973. *Les théories grammaticales d'Ibn Jinnī*. Tunis: Université de Tunis.

Owens, Jonathan. 1988. *The foundations of grammar: An introduction to medieval Arabic grammatical theory*. Amsterdam and Philadelphia: J. Benjamins.

Peled, Yishai. 1992. "Cataphora and *taqdīr* in medieval Arabic grammatical theory". *Jerusalem Studies in Arabic and Islam* 15.94–112.

Talmon, Rafael. 1997. *Arabic grammar in its formative age:* Kitāb al-ʿayn *and its attribution to Ḫalīl b. Aḥmad*. Leiden: E.J. Brill.

Troupeau, Gérard. 1976. *Lexique-index du* Kitāb *de Sībawayhi*. Paris: Klincksieck.

Versteegh, Kees. 1994. "The notion of 'underlying levels' in the Arabic grammatical tradition". *Historiographia Linguistica* 21.271–296.

Kees Versteegh (University of Nijmegen)

Tarāduf → Mutarādif

Taṣrīf → Ṣarf

Tatar

Arabic was a superstrate language for Volga Tatar (and the predecessors of the modern Tatar language) for more than one thousand years, with economic and political ties to the Islamic world established well before the Volga Bulgars' early-10th-century conversion to Islam, and the influence of Arabic on the Tatar language was significant. Up until the 19th century, Tatars shared three literary languages with the other Muslims of Russia: Classical Arabic, Classical Persian, and Chagatay. The first Arabic loanwords in Volga Turkic languages date to the early 10th century (Scherner 1977:14; Mäχmütov 1993a:5), and in the mid-19th century, a Tatar literary language arose that was based on the Kazan dialect but heavily influenced by Arabic and Persian. This literary language was written in Arabic script, which had succeeded pan-Turkic runes as the regional writing system in the 10th century. As can be seen from the approximately eight hundred pages of Arabic loanwords found in Mäχmütov a.o.'s (1993) dictionary of borrowings, by the early 20th century a significant portion of the lexicon of literary Tatar was of Arabic origin – and the majority of these words were incomprehensible to speakers of the vernacular who were not educated in Arabic and Persian, a sociolinguistic situation much like that of 19th-century Turkey.

Lexical reform begun by Tatar intellectuals in the late 19th century was undertaken with the aim of closing this gap between the literary language and the language of the people and promoted the use of native Turkic words in literary Tatar. Due to this reform movement, the Arabic-Persian element of texts by many authors, which at the turn of the century could be as high as 65 percent, decreased significantly (Mäχmütov 1993b:797). The number of Arabic loanwords used in Tatar was then drastically reduced by politically motivated Soviet-era lexical reform, when Soviet linguists replaced most of the Arabic and Persian loanwords in Tatar with Russian loanwords, such that half of the entries in today's standard Tatar-Russian dictionaries are Russian borrowings. This relexification was accompanied by two alphabet changes in quick succession: in 1927 from the Arabic script to a Latin-based alphabet and in 1938 to a Cyrillic-based alphabet. The end result is that Arabic influence in modern Tatar is significantly less than it was a century ago. In order to be comprehensible to the reading public, prerevolutionary Tatar texts are now presented in the Cyrillic Tatar alphabet and with glossaries when they are reprinted in post-Soviet newspapers and magazines (e.g. Mädriyeva 1998, where 18 percent of an excerpt from a 1908 article is glossed).

Arabic loanwords started appearing once more in written Tatar along with *perestroika* (in the mid-1980s), a symbolic gesture readily perceived by many Tatars as more than purely linguistic in nature (cf., e.g., Safiullina and Fyodorova 2000). The return of Arabic loanwords is part of purist post-Soviet linguistic reform on the part of the Tatar intellectual and political elite, where the implicit goal is the creation of a Tatar language that is maximally distinct from Russian. This lexical reform is therefore congruent with the highly contested 1999 Tatarstani legislation that decreed yet another shift in orthography, this time away from the Tatar-specific Cyrillic alphabet based on that of Russian to a new Latin-based alphabet similar to the one used for modern Turkish (Wertheim 2003).

Arabic loanwords, which are presented here in standard Turcological notation, have had a moderate effect on the phonological structure of Tatar and were integrated in a variety of ways. Table 1 below shows the relevant merger and adaptation of Arabic consonants and vowels.

Table 1. Merger and adaptation of Arabic phonemes in Tatar

Arabic	Tatar
t, ṭ	t
ḥ, x	χ
ġ, ʿ	γ
s, ṣ, ṯ	s
z, ẓ, ḍ, ḏ, ḏ̣	z
ā	a
a	ä

Arabic loanwords often violate Tatar's front/back vowel harmony, and the Tatar allophones *k/q* and *g/γ* have become phonemic due to

Arabic borrowings where *q* or *γ* is adjacent to front vowels. Tatar suffixes usually assimilate in quality to the final vowel of the loanword (e.g. *kitaplar* 'books' and not **kitaplär*); however, the *nisbe*-ending *i* in loanwords does obey vowel harmony and has the front allomorph *-i* and back allomorph *-ıy*, e.g. *ädäbi* 'literary' (< *ʾadabī*), *χosusıy* 'individual' (< *xuṣūṣī*).

Borrowed Arabic feminine nouns end in either *-a* or *-at* with occasional doublets that parallel those found in → Persian, the source for most of these nouns (Perry 1991:142), e.g. *χäräkät* 'movement' and the now archaic *χäräkä* 'short vowel diacritic' (< *ḥaraka* 'movement; vowel'); nouns without doublets include *χata* 'mistake' (< *xaṭāʾ*), *šifa* 'medicine' (< *šifāʾ*), *χökümät* 'government' (< *ḥukūma*) and *säyäχät* 'travel' (< *siyāḥa*).

Borrowed Arabic nouns are freely suffixed with Tatar inflectional and derivational morphology, e.g. *möstäqïllek* 'independence' (*möstäqïl* 'independent' < *mustaqill* + *-lek*, the Tatar abstract nominalizer) and *χalıkara* 'international' (*χalık* 'people' < *xalq* 'creation, mankind' + the Tatar postposition *-ara* 'between').

In addition, the dummy verb *itärgä* 'to do' is used to create verbs from some loanwords, e.g. *däwam* 'continuation' (< *dawām*), *däwam itärgä* 'to continue', while others are created using native derivational morphology, e.g. *riza* 'agreement' (< *riḍā*), *rizalašırga* 'to agree' (where *-la-* is the verbalizer, *-š-* the reflexive, and *-ırga* the infinitive).

The Arabic loanwords found in modern Tatar can be separated into two groups: everyday words that are encountered in regular speech, and words that are used most frequently in literary registers. Words of the first sort, which are unmarked for register, were usually not removed during the Soviet-era relexification process and are generally perceived as native. These include nouns such as *däftär* 'notebook' (< *daftar*), *iman* 'belief' (< *ʾīmān*), *kitap* 'book' (< *kitāb*), *mäktäp* 'school' (< *maktab* 'office'), and *tariχ* 'history' (< *tārīx*), and discourse-pragmatic and function words such as *älbättä* 'of course' (< *al-batta* 'definitely, positively'), *ämma* 'but, however' (< *ʾammā*), *qadär* 'as much as' (< *qadr* 'extent, quantity, amount'), and *χätta* 'even' (< *ḥattā*). By contrast, the majority of the archaic Arabic words that are in the process of being revived as part of Tatar post-Soviet lexical reform are more literary or formal words, words primarily used to describe politics, literature, culture, and religion. Their use is not yet standardized, and, more importantly, their use is not uncontroversial. The debate on the return of these archaic loanwords is played out most prominently in the pages of the daily press, which is also the primary vehicle for their reintroduction into post-Soviet Tatar. Some Arabic loanwords are found regularly in newspaper discourse, including *iǰtimaγïy* 'social' (< *ijtimāʿī*), *inqïylab* 'revolution' (< *inqilāb*), *iqtisad* 'economics' (< *iqtiṣād*), *ǰömhüriyät* 'republic' (< *jumhūriyya*), *χakimiyät* 'ruling power' (< *ḥākimiyya*), *mädäniyät* 'culture' (< *madaniyya*), and *säyäsät* 'politics' (< *siyāsa*). However, a complaint that surfaces regularly in post-Soviet Tatar discourse is that Arabic borrowings are being used overzealously, such that the language of the press is becoming increasingly alienating to average Tatar readers, or even periodically incomprehensible. Even so, the return of many Arabic loanwords to modern Tatar is a fait accompli, and part of linguistic reforms that are buttressed by political, cultural, and religious forces.

BIBLIOGRAPHICAL REFERENCES

Mädriyeva, Monǰiyä. 1998. "'Šura' kingeš digän süz" ['Šura' is a word of advice]. *Mädäni Jomγa*, September 25. 13.

Mäχmütov, M.I. 1993a. "Tözüčedän" [From the author]. *Γaräpčä-Tatarča-Rusča Alïnmalar Süzlege*, ed. M.I. Mäχmütov, K.Z. Xämzin, and G.Š. Säyfullin, 5–7. Kazan: "Iman" Näšriyatï.

——. 1993b. "Tatar ädäbi telenä kergän γarëp-farsï elementlarï" [The Arabic-Persian elements of the Tatar literary language]. *Γaräpčä-Tatarča-Rusča Alïnmalar Süzlege*, ed. M.I. Mäχmütov, K.Z. Xämzin, and G.Š. Säyfullin, 795–853. Kazan: "Iman" Näšriyatï.

——, K.Z. Xämzin, and G.Š. Säyfullin. 1993. *Γaräpčä-Tatarča-Rusča Alïnmalar Süzlege* [The Arabic-Tatar-Russian dictionary of borrowings]. Kazan: "Iman" Näšriyatï.

Perry, John. 1991. *Form and meaning in Persian vocabulary: The Arabic feminine ending*. Costa Mesa and New York: Mazda Publishers.

Safiullina, Flyora and El'vira Fyodorova. 2000. "Xäzerge tatar ädäbi telendä γaräp-farsï alïnmalarïnïŋ kire qaytu protsessï" [The process of the return of Arabic-Persian borrowings in the present-day Tatar literary language]. *Fän häm Tel* 3.16–21.

Scherner, Bernd. 1977. *Arabische und Neupersische Lehnwörter im Tschuwaschischen*. Wiesbaden: F. Steiner.

Wertheim, Suzanne. 2003. "Language ideologies and the 'purification' of post-Soviet Tatar." *Ab Imperio* 1.347–369.

SUZANNE WERTHEIM (Northwestern University)

Telugu

1. Interaction between the Arabic and Telugu speech communities

Telugu belongs to the Dravidian language family. It is spoken as the major language in Andhra Pradesh in India by more than 66 million people, and speakers of the language have spread to different parts of the world. The interaction between the Arabic and Telugu speech communities began when the Muslim king Allauddin Khilji invaded the south in 1296 C.E. Later, Giasuddin Tughlak invaded the south as well (Siddiqui 1956). Although no Muslim empire had yet been established there, during the 13th century those Muslim soldiers who did not return to Delhi settled in the south. Among them were Arabian soldiers. The Vijayanagara Kingdom was established in the south in 1336 C.E. After some years, the Bahamani Kingdom was established, also in the south. Although there was a fair amount of rivalry between the two kingdoms, the Vijayanagara kings were fond of purchasing horses from Arabia and Persia (Suravaram 1950). Thus, Persian and Arabian traders enjoyed high social prestige in the Vijayanagara Kingdom. In the reign of the Qutub Shahis, who ruled over the Telugu-speaking region during the 15th/16th century, Persian was the official language. Persian continued to be the official language during the reign of the Asafjahis, who established their dynasty in 1720 C.E. The Arabic language enjoyed a special position because the *Qurʾān* was written in this language. The kings felt that all their subjects should be well acquainted with the *Qurʾān*, which was said to contain a treasure of knowledge in many different areas of study. Thus, for sociocultural reasons the Telugu speakers came in contact with the Arabic language and borrowed several words from this language.

Arabic words are found in Telugu in the areas of administration and the judiciary, e.g. *qist* 'installment' (< Arabic *qisṭ* 'justice; share, installment'); *hissāb* 'account' (< *ḥisāb*); *arĵī* 'petition' (< *ʿarḍ* 'presentation; exhibition; submission of an application'); *tākīd* 'notice' (< *taʾkīd* 'affirmation'). There are many Arabic loanwords pertaining to day-to-day matters, e.g. *ilāj* 'medical treatment' (< *ʿilāj*); *ġarib* 'poor' (< *ġarīb* 'stranger'); *ĵid* 'adamant' (< *ḍidd* 'opposite, contrary'); *mašūr* 'well known' (< *mašhūr*). Because of the difference in structure between Telugu and Arabic, words borrowed into Telugu are adapted to the phonological and morphological structure.

2. Phonological adjustments

The voiceless velar fricative *x* of Arabic changes to an aspirated voiceless stop since Telugu does not have *x* (see Table 1). In language contact situations, the sounds which are foreign to the native structure are substituted by the phonetically closer sounds. In one word, *x* is substituted by *k* in Telugu: *xabar* > *kaburu* 'news, information'.

The voiced velar fricative *ġ* is not found in Telugu; it is substituted by a voiced stop, as in Table 2.

The uvular voiceless stop *q* changes to a voiceless velar stop *k* or an aspirated voiceless stop *kh*, as in Table 3.

Table 1. Arabic *x*/Telugu *kh*

Arabic	Telugu	
xāli	*khāli*	'empty, vacant'
xatm	*khatam*	'end, termination'
xarāb	*kharāb*	'state of ruin'

Table 2. Arabic *ġ*/Telugu *g*

Arabic	Telugu	
ġarīb 'stranger'	*garīb*	'poor'
ġalaṭ	*galat*	'mistake'
ġalīḍ 'rough'	*galīz*	'dirty'

Table 3. Arabic *q*/Telugu *k*

Arabic	Telugu	
qabḍa	*kabzā*	'capture, taking possession of'
qarār 'decision'	*karār cēyu*	'to settle, establish'
qisṭ 'justice; installment'	*kist*	'installment'
qayd 'fetter, shackle'	*khaid*	'imprisonment'

In Telugu, the voiced alveolar affricate *z* occurs before nonfront vowels, and the palatal stop *ĵ* occurs before front vowels. These two are allophones, e.g. *zuTTu* ‘hair’, *ĵila* ‘itching’. Arabic words with *z*, whether original or secondary < Persian *z* < Arabic *ḍ* or *ḏ̣*, have become *ĵ* in Telugu before front vowels (Table 4).

Before nonfront vowels, *z* in borrowed words remains as it is in Telugu because this is in conformity with the native structure (Table 5).

Arabic *j* before front vowels becomes *ĵ* in Telugu, e.g. *mustaʾjir* > Telugu *mustāĵir* ‘contractor’; *jamʿ* > *ĵemā* ‘collection’.

The loss of Arabic *ʿ* generally results in the lengthening of the preceding vowel, as in Table 6.

In consonant clusters in borrowed Arabic words, either a vowel is inserted between the two consonants, or the consonant is dropped, e.g. *masjid* > Telugu *masīdu* ‘mosque’; *mašhūr* > *mašūr* ‘popular’. Since the combination of consonants *s* and *j*, *š* and *h* does not exist in Telugu, the second consonant is dropped, which is compensated by the lengthening of the following vowel.

Liquids are not found as peaks of the syllable in the syllabic structure of Telugu. Only vowels act as peaks of the syllable. A vowel is therefore inserted between the consonants *-kr, -ql, -qm, -zn* (Table 7).

Insertion of a vowel is found optionally in words of the canonical CvCCvC, where the first member of the consonant cluster acts as coda of the first syllable and the second member as onset of the second syllable, e.g. *maḍbūt* > Telugu *mazbūtu/mazubūtu* ‘strong’; *majbūr* > *mazbūru* ~ *mazubūru* ‘compelled’.

Consonant clusters in disyllabic words with CvCCvC structure do not undergo any change, e.g. *ʿarḍ* > *arĵī* ‘application’; *qabḍa* > *kabzā* ‘occupation’; *marḍī* > *marĵī* ‘opinion’.

Table 4. Arabic *ḍ* ~ *ḏ̣*/Telugu *ĵ*

Arabic	Telugu
marḍī ‘approved, accepted’	*marĵī* ‘opinion’
māḍī ‘past’	*māĵī* ‘former’
ʿarḍ ‘presentation, application’	*arĵī* ‘application’
ḍidd ‘contrary’	*ĵiddi* ‘to contradict’

Table 5. Arabic *ḍ* ~ *ḏ̣*/Telugu *z*

Arabic	Telugu	
maḍḥaka ‘object of fun’?	*mazāk*	‘fun’
maḍbūt ‘accurate, precise’	*mazubūtu*	‘strong’
naḏ̣ar ‘look’	*nazar*	‘evil eye’
ḍāmin ‘responsible, liable’	*zāmīn*	‘to be responsible’

Table 6. Arabic *ʿ*/Telugu vowel lengthening

Arabic	Telugu	
taʿalluq ‘connection’	*tālūku*	‘to be connected to’
taʿlīm	*tālīm*	‘teaching’
taʿwīḏ ‘amulet’	*tāwīzu*	‘charm’

Table 7. Epenthetic vowels in Arabic loanwords

Arabic	Telugu	
fikr ‘thought’	*fikar*	‘fear’
naql ‘transmission’	*nakal*	‘copy’
raqm ‘number’	*rakam*	‘cash’
wazn	*wazanu*	‘weight’

3. MORPHOLOGICAL ADJUSTMENTS

Nominal forms are borrowed without any inflectional markers. Telugu has borrowed only the singular form, which is provided with a Telugu plural ending, e.g. Arabic *ḥaqq*, pl. *ḥuqūq* ‘right, claim’ > Telugu *hakku*, pl. *hakku-lu*; Arabic *kitāb*, pl. *kutub* ‘book’ > Telugu *kitāb, kitābu-lu*. Gender differentiation is made in Telugu at the pronominal level and in the verbal endings and is based upon semantic criteria. All nouns are classified into masculine and nonmasculine in the singular, and human vs. nonhuman in the plural. This difference is reflected in the verbal endings, too, where gender of the subject is marked:

rāmuDu	*vaccEE-Du*
Rama [masc.]	came [masc. sg.]

sita	*vaccindi*
Sita [nonmasc.]	came [nonmasc. sg.]

āvu	*vaccindi*
cow [nonmasc.]	came [nonmasc. sg.]
bassu	*vaccindi*
bus [nonmasc.]	came [nonmasc. sg.]
rāmuDu sita	*vaccEEru*
Rama and Sita [human pl.]	came [human pl.]
āvulu	*vaccEEyi*
cows [nonhuman pl.]	came [nonhuman pl.]
bassulu	*vaccEEyi*
buses [nonhuman pl.]	came [nonhuman pl.]

In Arabic, gender differentiation is based upon grammatical criteria, and all nouns are classified into masculine and feminine. Gender is expressed at the pronominal level and in verbal endings. Demonstrative pronouns are declined for masculine and feminine gender. Gender distinction is made in the verbal endings also. Due to the differences in the gender system, the gender of Arabic borrowings has been adapted to the Telugu gender system, e.g. *murabbā* 'jam [masc.]' > Telugu *murabba* 'a kind of sweet [nonmasc.]'. The word *masjid* 'mosque' [sg. masc.] became in Telugu *masīdu* [nonmasc.], with the plural *masīdulu* [nonhuman pl.], as in the following examples.

akkaDa	*masīdu*	*undi*
there	mosque	to be

'There is a mosque'

akkaDa	*masīdu-lu*	*unnāyi*
there	mosque-PL	to be-PL

'There are mosques'

Arabic verbs are not borrowed directly into Telugu but rather by means of Telugu conjunct verbs, which are formed by adding an auxiliary verb to a nominal form:

pani	*cēyu*
work-NOUN	to do-AUX

'to work'

Equivalents of Arabic verbs are formed in Telugu by adding the auxiliary verbs *ceyu* 'to do' and *agu* 'to happen' to nouns borrowed from Arabic, e.g.:

taraqqi	*agu*
progress	to happen

'to improve [intrans.]'

taraqqi	*cēyu*
progress	to do

'to improve [trans.]'

šarīk	*cēyu*
companion	to do

'to join'

šarīk	*agu*
companion	to happen

'to be joined'

BIBLIOGRAPHICAL REFERENCES

Siddiqui, Abdul Majeed. 1956. *History of Golconda*. Hyderabad: Literary Publications.

Suravaram, Pratap Reddy. 1950. *Andhrula Sanghika Caritra*. [In Telugu]. Hyderabad: Andhra Saraswata Parishat.

V. SWARAJYA LAKSHMI (Osmania University)

Template → Obligatory Contour Principle; Morphology

Tense

A controversy prevails in the study of Semitic languages, both ancient and modern, namely whether their verbal morphologies mark tense (e.g. past vs. non-past), or → aspect (perfective vs. imperfective), or some combination of both. Also something of an issue are → mood and modality, but they are not usually as problematic (or they are problematic in different ways) as tense and aspect.

Arabic is consistent with other Semitic languages in its nonconcatenative morphology. Verbal and nominal forms alike are typically formed by interdigitation of consonantal roots (ideally consisting of three radicals) and vocalic templates, either with or without affixes. Verbs have historically been classified as 'perfect' and 'imperfect', using a loose correlation between aspectual meaning and the morphological forms which were perceived as conveying them. These two forms are also referred to as suffix conjugation (*qatala*) and prefix conjugation (*yaqtulu*), which are the terms used in this entry. In the controversy mentioned above, the prefix conjugation is described as either ‘imperfect’ or ‘nonpast’ (or occasionally present/future), and the suffix conjugation is known as either ‘perfect’ or ‘past’.

For most Arabic grammarians, the issue was rather simple: since the verb indicates an action, and actions take place in the past, present, or future time, verbal forms by necessity indicate time/tense. In fact, the most current definition of the verb is that it is “what indicates a combination of event and time” (*al-fiʿl mā dalla ʿalā qtirān ḥadaṯ wa-zamān*), a definition to be found, for instance, in az-Zamaxšarī’s *Mufaṣṣal* (108.6). The only controversy in classical Arabic grammar centered around the question of whether there actually is such a thing as a present tense (see Zajjājī, *ʾĪḍāḥ* 86–88). According to the Kufan grammarians, the *fiʿl dāʾim*, i.e. the active participle, fulfilled this role, while the Basran grammarians followed Sībawayhi in accepting only two verbal tenses, the past and the ‘resembling’ (*muḍāriʿ*) tense, the latter being ambiguous because it could indicate both the present and the future (→ *māḍī/muḍāriʿ*). For a survey of the Arabic theories about verbal tense, see Fleisch (1979:201–206) and Versteegh (1981).

In Western reference grammars of Classical Arabic (Wright 1964:II, 18, “The *Imperfect Indicative*… does not in itself express any time”; Fleisch 1979:169–201), Modern Standard Arabic (Badawi a.o. 2004:362–371), and Arabic dialects (see Eisele 1999), the consensus seems to be that the Arabic verbal system is aspectual in nature, although Badawi a.o. (2004:362) indicate that Modern Standard Arabic “has now also a complete three tense structure replicating that of western languages (signs of which were already apparent in C[lassical] A[rabic]”. A few researchers, notably Aartun (1963), regard the tense opposition as the basic distinction in the Arabic verbal system. Others regard the distinction as irrelevant. Comrie (1976:79) states, for instance, that the basic distinction in the Arabic verbal system is neither one of tense, nor one of aspect, but incorporates both aspect and (relative) tense. It is certainly true that the two are often intertwined, and Payne emphasizes that “tense, aspect, and mode are sometimes difficult to tease apart” (1997:234).

A proposal made by Dahl (1985) regarding the crosslinguistic traditional typology of Tense(Mood)Aspect or T(M)A systems incorporates the notion of pragmatic implicature into the traditional taxonomy. Dahl (1985:11ff.) discusses ‘secondary meanings’, ‘secondary foci’, and the ‘conventionalization of implicatures’ as potentially pertaining to grammatical categories such as tense and aspect. He defines ‘implicature’ in this sense as “something that can be inferred from the use of a certain linguistic category or type of expression, although it cannot be regarded as belonging to its proper meaning”. In his analysis of the various categories available in the languages in his study, he often distinguishes between those that are overtly marked and others that are merely implicated. Such an account, if found theoretically and empirically sound, may further the resolution of the controversy regarding the Arabic verbal system.

In an experiment by Horesh (2002), six native speakers of different Arabic dialects were asked for their reactions to a number of utterances previously recorded by three native speakers of the dialect of Jaffa. They were asked to place these utterances in time (e.g. past, present, future), and in various cases to provide alternatives in their own vernacular. A second component of the interview was a fragment of Dahl’s questionnaire, in which the informants were asked to translate three similar English narratives into their vernacular.

The reactions to the sentences were fairly uniform. All suffix conjugations were consistently interpreted as denoting actions in the past. In those cases where the auxiliary verb *kān* was juxtaposed to a prefix-conjugated verb (e.g. *kān yaʿṭi* ‘he used to give’), it was interpreted as modifying an aspectual (in this case continuous) verb, again, denoting a situation or process in the past. In a sentence like (1)

(1) *ʾara* *hāda* *l-kitāb*
read.3ms this the-book
(min *ʾawwal-o* *la-ʾāxir-o)*
(from beginning-its to-end-its)
‘He read the book (from cover to cover)’

all informants agreed that the verb meaning ‘read’ was to be understood as ‘read from cover to cover’, even without the adverbial phrase explicating that, which ruled out a Slavic-type perfective interpretation of the suffix conjugation.

There was also consensus among the informants about the *yištri/uštara* contrast in sentence (2).

(2) *rāḥ* *ʿa-s-sūʾ*
went.3ms to-the-market
yištri/uštara *tuffāḥ*
buy.3ms/bought.3ms apples
'He went to the market to buy/and bought apples'

All informants agreed that the sentence with the prefix-conjugated verb contains no information as to whether the subject of the sentence has actually bought the apples. Only an explicit suffix-conjugated verb following the conjunction indicates that the purchase has indeed taken place.

The use of participial forms in (3)–(4) as denoting some kind of present perfect, i.e. an action in the past bearing consequences for the present, was controversial.

(3) *il-malik* *ṣār* *wāṣil*
the-king became.3ms arriving.Part
'The king has arrived'

(4) *miš* *min* *zamān* *šattat/mšattye*
Neg from time rained.3ms/raining.Part
'Not long ago, it (has) rained'

Several subjects rejected it altogether. In some cases, informants argued that even a simple suffix-conjugated verb (rather than a complex auxiliary + participle) would entail that the action or state described has bearing on the present.

The second part of the interview yielded somewhat more interesting results. While some speakers showed little variation across contexts, others differed quite radically when shifting from a narrative situated 'yesterday' through one pertaining to 'what just happened to me' to a narrative describing a distant past habitual. An example of the narratives is given in (5a–c); in the narratives the verbal forms to be translated were given in their base form.

(5) a. *Do you know what happened to me yesterday? I WALK in the forest. Suddenly I STEP on a snake. It BITE me in the leg. I TAKE a stone and THROW it at the snake. It DIE.*
b. *Do you know what just happened to me? I WALK in the forest. Suddenly I STEP on a snake. It BITE me in the leg. I TAKE a stone and THROW it at the snake. It DIE.*
c. *I'll tell you what happened sometimes to me when I was a child and was walking in the forest. I WALK in the forest. Suddenly I STEP on a snake. It BITE me in the leg. I TAKE a stone and THROW it at the snake. It DIE.*

While there was convergence in the use of the suffix conjugation for events in the past, there was quite some variation for some of the verbs in the narratives, especially those of movement (walking, biting, throwing, stepping), for which intermediate forms like the participle, and combinations of the auxiliary *kān* with the participle or a prefix-conjugated verb were used. With respect to the three different contexts, in the context of (5a) and (5b), the suffix conjugation was used fairly consistently by all informants, but in the context of (5c), various intermediate forms were used, including complex forms with the auxiliary *kān*.

A prefix-conjugated verb was not used for any of the verbs in the narratives, but one informant used a prefix-conjugated verb for 'you know' (which, strictly speaking, was not part of the narrative). This informant indicated a number of other stative verbs that may be used in the suffix conjugation without any reference to the past, e.g. 'to understand'. This is consistent not only with the situation in other Semitic languages (cf. Hebrew *katonti* 'I am too small; I am at a loss'), but is in fact a relic from Proto-Semitic, where presumably suffixes were productively indicative of stative verbs, as is the case in the oldest attested Semitic language, Akkadian.

On the one hand, therefore, in many instances the suffix conjugation was consistent with an action in the past. Moreover, in many cases where the informants were asked to provide an alternative form while preserving the meaning, they argued that it was impossible to do so. On the other hand, there are a number of stative verbs that may occur in the suffix conjugation without reference to the past. There are also contexts in which the suffix conjugation may be used interchangeably (for some speakers) with a participle to denote a 'perfective' aspect, a present result of a past occurrence. And finally, there is the auxiliary verb *kān* modifying various other forms (e.g. participles and prefix-conjugated verbs), for instance to place a continuous or habitual action in the past. It therefore seems plausible to conclude that the suffix conjugation has more than one meaning and is not just a past tense form. Not only that, but given the variety of uses that the prefix conjugation has, with various shades of imperfectivity (inchoativity, habituality, continuous actions, to name but a few), it makes sense to attribute the converse aspect notion, perfectivity, to the suffix conjugation.

In the preface to his edited volume on Semitic languages, Hetzron (1997:xvi) quotes Chaim Rabin, who said in a lecture, "Semitic has either aspects that express tenses or tenses that express aspects". The question remaining now is whether from a pragmatic point of view, what we have here is indeed an implicature, as proposed by Dahl for similar cases across the world's languages. Levinson (2000:261), in his chapter on grammar and implicature, argues that "the relation between syntax and pragmatics is of a fundamentally different kind than the semantics/pragmatics interface, for it is indirect". If we accept the notion that the farther away from pragmatics our facts are, the harder it is to incorporate pragmatic theory into the analysis, then the situation with respect to the Arabic verbal forms is even tougher than with the sentential anaphora phenomena examined by Levinson.

Elsewhere, Levinson (1983:77–78) follows Lyons in distinguishing between M(etalinguistic)-Tensed and (Language's)-Tensed, acknowledging that one reason for the two not being entirely compatible has to do with the latter "nearly always encod[ing] additional *aspectual* and *modal* features too". Phrases like *used to give scholarships* implicating *no longer gives scholarships* are said to be "permeated by Gricean mechanisms" (Levinson 2000:180), due to the opposition between used and unused temporal references. This may not be compatible with Sadock's (1978) critical view of the testability of implicatures for their conversational nature. Yet, it may be recalled that Dahl's arguments on TenseMoodAspect (TMA) categories was that they may be subject to the 'conventionalization of implicatures', in which case the need to calculate maxim violations, as is the case for conversational implicatures, may not be necessary.

Clearly, the results of the experiment described here represent but the tip of the iceberg, and there is need of a more extensive corpus to be examined. It is not clear whether the proper methodology exists for testing hypotheses regarding the pragmatic status of conventionalized grammatical manifestations like the ones dealt with here. Native speakers clearly have intuitions about what denotes what, but for some reason – perhaps due to simplification of grammatical explanations in school – speakers rarely go beyond labeling the forms they use with temporal terminology. Notions like 'perfective' and 'imperfective', which for linguists may be, at least for some languages, part and parcel of the verbal system (see Comrie 1976:16), are often neglected or misunderstood, and it is therefore necessary to try to devise ways to circumvent the informants' lack of expertise.

Yet, this small-scale experiment indicates that despite quite a few dialectal differences, there is a certain degree of stability in the verbal system of the Arabic dialects. This may have to do with the fact that the suffix conjugation has never in the history of Arabic been very complex. The prefix conjugation, if closely examined, will most likely turn out to be far more complex and quite more variable. This should be examined with a combination of a corpus-based quantitative analysis and a carefully constructed language-specific questionnaire.

Bibliographical references

Primary sources

Zajjājī, *ʾĪḍāḥ* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān ibn ʾIsḥāq az-Zajjājī, *al-ʾĪḍāḥ fī ʿilal an-naḥw*. Ed. Māzin al-Mubārak. Cairo: Dār al-ʿUrūba, 1959.

Zamaxšarī, *Mufaṣṣal* = ʾAbū l-Qāsim Maḥmūd ibn ʿUmar az-Zamaxšarī, *Kitāb al-mufaṣṣal fī n-naḥw*. Ed. Jens Peter Broch. Christiania: Libraria P.T. Mallingii, 1870.

Secondary sources

Aartun, Kjell. 1963. *Zur Frage altarabischer Tempora*. Oslo: Universitetsforlaget.

Badawi, Elsaid, Michael G. Carter and Adrian Gully. 2004. *Modern Written Arabic: A comprehensive grammar*. London and New York: Routledge.

Comrie, Bernard. 1976. *Aspect: An introduction to the study of verbal aspect and related problems*. Cambridge: Cambridge University Press.

Dahl, Östen. 1985. *Tense and aspect systems*. Oxford: Blackwell.

Eisele, John. 1999. *Arabic verbs in time: Tense and aspect in Cairene Arabic*. Wiesbaden: O. Harrassowitz.

Fleisch, Henri. 1979. *Traité de philologie arabe*. II. *Pronoms, morphologie verbale, particules*. Beirut: Dar El-Machreq.

Hetzron, Robert (ed.). 1997. *The Semitic languages.* New York: Routledge.
Horesh, Uri. 2002. "Tense or aspect: Is Arabic past tense an implicature?". Ms., University of Pennsylvania. Available at: www.ling.upenn.edu/~urih/Horesh_TMA_Prag.pdf.
Levinson, Stephen C. 1983. *Pragmatics.* Cambridge: Cambridge University Press.
——. 2000. *Presumptive meanings.* Cambridge, Mass.: MIT Press.
Payne, Thomas E. 1997. *Describing morphosyntax: A guide for field linguists.* Cambridge: Cambridge University Press.
Sadock, Jerold M. 1978. "On testing for conversational implicature". *Syntax and semantics.* VI. *Pragmatics*, ed. Peter Cole, 281–297. New York: Academic Press.
Versteegh, Kees. 1981. "La conception des 'temps' du verbe chez les grammairiens arabes". *Théorie Analyses* 3.47–68.
Wright, William. 1964. *A grammar of the Arabic language.* 3rd ed. rev. W. Robertson Smith and M.J. de Goeje. Repr. 2 vols. Cambridge: Cambridge University Press.

Uri Horesh
(Franklin & Marshall College)

Terminology

1. Definition

Although terminological creation is governed by the same linguistic framework as common vocabulary, it is subject to specific conditions of its own. 'Terms' can be defined as words or phrases aiming at the designation of concepts related to a particular field of knowledge or activity. New terms can be introduced either by individuals or by a limited community, in answer to particular needs and under conditions which vary from one field to another. As a result, prime occurrence of terms may – or may not – be produced through conscious activity, and new terms may – or may not – be created in an organized and methodical way. Moreover, their first appearance often leads to divergence between the models of formation that prevail in the common vocabulary and the formation of technical terminologies. Such a process cannot be without consequence for the development of common vocabulary, especially in modern times, owing to the fact that in any given language terminologies have grown at a much quicker pace than the general lexicon. Besides, it should be kept in mind that usages and traditions in creating new terms may vary from one domain of knowledge and/or activity to another.

2. Periodization of the development of the Arabic lexicon

2.1 *Periods of Arabic terminology*

Arabic terminology has developed in the overall frame of what may be described as the 'great periods' of the vocabulary of the language, which can, on a very general basis, be divided into three parts (Dichy 1998):

i. The first great period is that of the 'language of the Ancient Arabs' (*lisān al-ʿArab*), which includes the original ancient lexicon. From the viewpoint of today's lexicographers, this original lexicon appears both prior to Islam and contemporaneous to its foundation. Apart from basic notions, the pre-Islamic vocabulary, i.e. the original Arabic vocabulary, also includes a set of notions referring to the life and institutions of the Ancient Arabs, e.g. *dār* 'tent'; *šahāda* *[sense 1]* 'testimony; witnessing'; *ṭarab* *[sense 1]* 'deep joy or sadness; deep emotion [in relation to listening to poetry or singing]'; etc. In this pre-Islamic period, the great majority of terms referring to everyday life (nomadic life, parts of the Bedouin tent, camels, stars, etc.) were of Arabic or Semitic origin. But there already were some non-Semitic loanwords, most of them borrowed from Greek, Latin (often, through Greek and Syriac), or Persian (→ Greek loanwords; → Latin loanwords; → Persian loanwords). An example is the word *ṣirāṭ* 'way, road', which seems to have been directly borrowed from the Latin word *strāta* 'pavement', hence 'paved way; military road' (Rabin 1960). This word occurs 45 times in the *Qurʾān*, mostly in the noun phrase *aṣ-ṣirāṭ al-mustaqīm* 'the Straight Path' (Q. 1/6). It is nevertheless difficult to trace terminological activity so far back. After the founding of Islam, a number of words belonging to the ancient Arabic lexicon were reused in a new meaning and became part of the semantic system originating in the *Qurʾān* and the teachings of the Prophet, i.e. in the fundamental religious and/or juridical terminology of Islam, e.g. *hudā* 'right way; Guidance'; *šahāda**[sense 2]* 'ritual formula by which the Muslim believers testify of their faith in a unique God, and in the prophecy of Muḥammad'; *šahāda*$^{[sense\ 3]}$ 'testimony [as a technical term of Islamic law]'.

ii. The vocabulary of medieval Arabic civilization contains items referring to the life and institutions of cities and to Islamic civilization, from the first centuries of the Hijra to the 18th century C.E. This includes the vocabularies of religious prescription and doctrine (not only Islamic but also Christian and Jewish religious terms), Islamic law, intellectual life, literary analysis, linguistic sciences, mathematics, astronomy, medicine, philosophy, administration, chancellery, etc., e.g. *qāḍī* 'judge'; *waṭan*$^{[sense\ 1]}$ 'country of origin' (which differs from the modern meaning of the word, i.e. from *waṭan*$^{[sense\ 2]}$, below); *ʾisnād* 'chain of authority of a Tradition'. This very wide 'medieval' epoch should of course be subdivided, but given the present state of the art it is difficult to delimit narrower periods. Significant examples found in medieval dictionaries can nevertheless be analyzed. The term *manṭiq*, for instance, which originally meant 'speech', takes, in the course of time in the writings of philosophers, such as al-Fārābī (4th/10th century), the sense of 'logos' (with reference to Aristotle and Plato) and of 'logic'. The active participle *muṭrib,* which is related to the same root as *ṭarab*$^{[sense\ 1]}$ above, at one point came to mean 'singer', probably through metonymic transfer: 'he who causes *ṭarab*$^{[sense\ 1]}$'. This must have been prior to the 4th/10th century, since this meaning is attested in Ibn Fāris' (d. 385/1005) dictionary, *Maqāyīs al-luġa*.

iii. The starting point of the vocabulary of the modern age is traditionally situated at the dawn of the 19th century, i.e. at the beginning of the *Nahḍa*, the Arab 'Renaissance'. Many neologisms date from this time, e.g. *sayyāra* 'car'; *bāxira* 'steamboat; boat'; *waṭan*$^{[sense\ 2]}$ 'homeland; country' (in the contemporary nationalist meaning, as in *al-waṭan al-ʿarabī* 'the Arab homeland'); *šahāda*$^{[sense\ 4]}$ 'certificate; certificate of attestation [related to a given academic degree]'. The two words *siyāsa* 'politics' and *mujtamaʿ* 'society' (in the modern understanding) had been in use in previous periods, but did not refer then to the same concepts.

2.2 *The development of technical terminology*

The development of sciences, which started at the beginning of the Umayyad dynasty and continued during Abbasid times, was accompanied by the translation of numerous documents of the Greek heritage, often through the intermediary of Syriac. New terminologies appeared in many fields, and the Arabic language, which had until then been in use mainly in traditional knowledge (craftsmanship, traditional medicine, etc.), became in less than two centuries a language allowing discussion of high-level scientific or technical developments, just as the Ancient Greek language had been (Rabin 1960; Jacquart 1994).

The terminology of Arabic grammar, for instance, was developed on the basis of common vocabulary. Likewise, the majority of terms in mathematics and optics were created by adding a new meaning to words of the common language. A few terms were borrowed, most of which were soon replaced by a word of Arabic origin; *ʾarīṯmāṭīqā* 'arithmetic', for instance, was replaced by *ḥisāb*. Other borrowed terms remained in use, such as *mūsīqā* 'music'. On the other hand, medicine and pharmacy very often resorted to borrowed terms for substances.

The situation of the Arabic sciences in modern times, in which all new concepts are imported from Occidental languages through English or French, is not without influence on Arabic scientific and technical terminology. The creation and development of terms occurred, and continues to occur, in massive quantities on the basis of English or French terminologies.

The history of the beginning and development of Arabic scientific and technical terminologies remains to be written, despite a number of studies dealing with particular fields. The specific characteristics of each domain of scientific or technical specialty, and the various ways in which specific terminologies were developed in a language other than the one in which they had been created, are so divergent, and there are so many different factors to take into account, that one cannot deal with Arabic scientific and technical terminology as a whole, except for a few very general considerations. The historical frame can nevertheless be outlined.

The development of modern Arabic terminology began in the first period of the Arabic Renaissance (*Nahḍa*) at the turn of the 19th century, with the intrusion of European sciences and techniques in the Arab culture (Hamzaoui 1991; Ḥigāzī 1993). The three most important centers for the formation of new Arabic terms were Cairo and Beirut, together with Constantinople, although its linguistic framework was different because Ottoman Turkish borrowed most of its scientific terms from Arabic terminological usage.

Muḥammad ʿAlī's policy of modernizing Egypt led him to call on engineers and physicians from Europe, particularly from France (aš-Šayyāl 1951). Their lectures were translated into Arabic, and from the late 1830s onward, this resulted in a movement of translating scientific and technical books, accompanied by the establishment of corresponding terminologies (Sawaie 1999). Medicine, for instance, was taught in the Arabic language from the 1830s until 1887, during the English occupation, when the English language took the place of Arabic in the teaching of sciences. In Lebanon, during the same century, the sciences were taught in Arabic, even at the American University, founded in Beirut in 1866. Popular science magazines such as *al-Muqtaṭaf* (published at first in Beirut in 1876, and then in Cairo) also appeared during this period. In other countries of the Ottoman Empire, such as Syria and Iraq, the language used for official teaching was Ottoman Turkish. In the Maghreb countries, French became the language used for sciences.

After the fall of the Ottoman Empire, the sciences began to be taught in Arabic in Syria and Iraq. The Faculty of Medicine of Damascus was founded in 1919. The scientific terminology established at that time used many Arabic terms that had been previously borrowed by the Ottoman language (Monteil 1960; aš-Šihābī 1965), as the few samples in Table 1 show (Ottoman terms are written in present-day Turkish spelling).

Generally speaking, the main centers for the diffusion of Arabic terms during the second part of the 20th century appear to have been Syria and Egypt. Institutions were created in order to deal with, among other linguistic policy matters, terminological problems (→ language academies). These included the Academy of Damascus, founded in 1921 (ʿAmmār and al-Xūrī 1996) and the Academy of Cairo, founded in 1932 (Hamzaoui 1975), and, at the other end of what was to become the Arab world, the Institut d'Études et de Recherches pour l'Arabisation in Rabat, which is devoted to the dissemination of the use of Arabic in Morocco, in addition to terminological issues (Richert 1987). Other institutions provided terms, such as the University of Damascus, where the sciences continue to be taught in Arabic.

Table 1. Arabic terminology borrowed through Ottoman Turkish

English	Arabic (Egypt)	Arabic (Syria)	Ottoman Turkish
'angle of incidence'	*zāwiya(t) suqūṭ*	*zāwiya(t) wurūd*	*zaviye-i vürud*
'focus [optics]'	*buʾra*	*maḫraq/ miḫraq*	*mıhrak*
'image [optics]'	*ṣūra*	*xayāl*	*hayal*
'refractive index'	*muʿāmil*	*qarīna*	*karıne*
'coil'	*milaff*	*wašīʿa*	*veşiʾa*
'energy'	*ṭāqa*	*qudra > ṭāqa*	*kudret*
'vector [optics]'	*muttajih*	*šuʿāʿ*	*şuâʾ*
'function [mathematics]'	*dālla*	*tābiʿ*	*tabiʾ*

One of the problems for Arabic scientific and technical terminology is that of variation across the different areas of the Arab world (→ lexical variation: Modern Standard Arabic). There is no common center for the coining of new terms that could unify or standardize newly coined terms, or terms used in a particular domain (ʿAmmār and al-Xūrī 1996). In 1966, the Bureau for the Coordination of Arabization was created in Rabat under the aegis of the ALECSO (Arab League Educational, Cultural and Scientific Organization). It was assigned the task of collecting synonymous terms and organizing congresses for their unification. Various domains were concerned with the publishing of 'unified dictionaries', whose efficiency in terms of consistency or influence still has to be analyzed.

It is also to be noted that whereas scientific and technical subjects are taught in Arabic at the high school level in the state educational system of most Arabic countries, scientific teaching at the college level very often switches to English or French. An exception is Syria, where the sciences are always taught in Arabic.

3. Linguistic structure and the coinage of new terms

Except when they are borrowed from another language, newly coined terms tend to be related to the concept they express by at least one formal or semantic feature. The linguistic devices, whether of a morphological, syntactic, or semantic nature, that are used in the denomination process ensure some level of semantic motivation in new terms (Stetkevych 1970; Hamzaoui 1991; Lelubre 1992). Linguistic means resorted to in the creation of terms do not fundamentally differ from those encountered in the age-old development of the general vocabulary, although they may sometimes take the appearance of innovation. Both result from the basic strategies of lexical formation, i.e. morphological means, 'frozen' syntactic structures, semantic transfer, and borrowing (from another language).

3.1 *Morphological means*

Morphological resources for the coining of words in Arabic are of four different types:

i. 'Internal' derivation (→ *ištiqāq*). This fundamental means for the creation of new Arabic terms consists, briefly put, in the 'fitting together', 'crossing', or 'merging' of a pattern and a tri- or quadriconsonantal → root. This means is recommended by all Arabic terminological institutions as the first choice to be considered, and actually remains widely used. Examples include: *mijhar* 'microscope' (triconsonantal root *j-h-r*; pattern miR$_1$R$_2$aR$_3$ [*mifʿal*]); *inʿikās* 'reflection' (triconsonantal root *ʿ-k-s*; pattern inR$_1$iR$_2$āR$_3$ [*infiʿāl*]); *tabaʿṯur* 'scattering; dispersion' (quadriconsonantal root *b-ʿ-ṯ-r*; pattern taR$_1$aR$_2$R$_3$uR$_4$ [*tafaʿʿul*]). In all such examples, the pattern is associated with a specific grammatical meaning. The meaning of 'name of instrument' (*ism ʾāla*), for instance, which is associated with the pattern miR$_1$R$_2$aR$_3$ can be observed in *mihjar*, and the reflexive value of the pattern inR$_1$iR$_2$āR$_3$ is included in the meaning of *inʿikās* 'reflection'.

One has to take into consideration that the concepts of root and pattern, as well as that of the root-and-pattern relation, need revisiting (→ derivation). In the first place, a number of lexical entries cannot be analyzed in terms of pattern structure and do not include a root (Dichy 1989). Secondly, roots are by no means the 'origin' of lexical items and must be regarded as the 'daughters' rather than the 'mothers' of words (Dichy 1998, 2003). Finally, as a result, both root and pattern can be described as relational concepts, i.e. as sets of formal and semantic relations. All three points underlie the analysis of the examples below. The terms 'root' and 'pattern' in this article refer to the revised definition. Newly coined terms based upon morphological means are of different types:

(a) New occurrence of a 'crossing' between a root already instantiated in the language and a current pattern. Some item-to-item or pattern-to-pattern relations between lexical entries included in a given root, which are rare in the general vocabulary, are frequently resorted to in terminology. This is the case, for instance, in the derivation of a 'name of instrument' (*ism ʾāla*) from an intransitive verb or a concrete noun, e.g. in the derivation of *miṭyāf* 'spectroscope' (root *ṭ-y-f*; pattern miR$_1$R$_2$āR$_3$ [*mifʿāl*]) from *ṭayf* 'spectrum'.

(b) Frequent use of patterns other than those of the 'name of instrument' to refer to instruments, tools, or devices, such as R$_1$aR$_2$R$_2$āR$_3$(a) (*faʿʿāl(a)*) or the patterns of active participles (*ism al-fāʿil*, R$_1$āR$_2$iR$_3$(a) [*fāʿil(a)*] or muR$_1$aR$_2$R$_2$iR$_3$(a) [*mufaʿʿil(a)*]), e.g. *ṭābiʿa* 'printer' (root *ṭ-b-ʿ*; pattern R$_1$āR$_2$iR$_3$a), *muḥallil* 'analyzer' (root *ḥ-l-l*; pattern muR$_1$aR$_2$R$_2$iR$_3$). Note that this derivational process is by no means an innovation.

(c) Frequent use of patterns that were previously rare, such as R_1āR_2ūR_3 (*fāʿūl*) for the denomination of instruments, e.g. *ḥāsūb* 'computer'.

(d) Pattern innovation. There do not seem to be examples of the appearance of new patterns. Yet, attempts have been made to reorganize the use of a few existing ones. The Cairo Language Academy tried to establish a distinction between the patterns used for the 'name of instrument', with reference to specific terminological uses: miR_1R_2aR_3 (*mifʿal*) for French or English terms including the suffix *-meter*, and miR_1R_2āR_3 (*mifʿāl*) for those including *-scope*. Such an interesting proposal encounters practical difficulties: the above mentioned *mijhar*, for instance, would become *mijhār*, but this term already exists, meaning 'loudspeaker'. Reorganizing an existing terminology is a difficult endeavor.

(e) Root innovation. A number of roots were created on the basis of lexical units which themselves include a previous root, deemed 'original' in the Arabic language, e.g. *ḥawsaba* 'computerization', new root *ḥ-w-s-b*, is built on *ḥāsūb* 'computer', the original root of which is *ḥ-s-b*; *ʿawlama* 'globalization', new root *ʿ-w-l-m*, is derived from *ʿālam* 'world', original root *ʿ-l-m*. The case of *mutamarkiz* 'concentric', new root *m-r-k-z*, is significant. The term is built on *markaz* 'center', original root *r-k-z*. Traditional Arabic lexicography regards this unit as related to the root *r-k-z* (see, e.g., the dictionary of the Language Academy of Cairo, *al-Muʿjam al-wasīṭ*). Strictly speaking, such a position necessarily implies the creation of a new pattern, mutamaR_1R_2iR_3 (*mutamafʿil*), which would be the active participle (*ism al-fāʿil*) of the verb *tamarkaza* 'to take or choose as a center', new pattern tamaR_1R_2aR_3a (*tamafʿala*). Yet, neither tamaR_1R_2aR_3a (*tamafʿala*) nor mutamaR_1R_2iR_3 (*mutamafʿil*) is considered as included in the pattern system of the language. The examples given in this paragraph show that, in the eyes of most Arabic lexicographers, the derivation of a new root from an original one should be avoided at all costs. On the other hand, root innovation seems to be admitted, albeit implicitly, in the case of loanwords, some of which are (relatively) ancient, e.g. *takahrub* 'electrification' (root *k-h-r-b*, from *kahrabāʾ* 'electricity' < Persian *kah-robā* 'amber'). Contemporary examples are: *taʾaksud* 'oxidation' (root *ʾ-k-s-d* < English *oxide*, French *oxyde*); *mutalfaz* 'broadcast on television, televized' (root *t-l-f-z* < English *television*, French *télévision*). The greater part of these borrowed roots is quadriconsonantal, and they exhibit a limited degree of productivity, e.g. *takahraba* 'to be electrified; to be electrocuted'; *ʾaksada* 'to oxidize'; *talfaza* 'to broadcast on television'.

ii. Lexical use of prefixed formatives and suffixes. The Arabic language, unlike English, has only a few lexical affixes. Moreover, these are only used in nouns and adjectives, and not in verbs. Two cases are observed:

(a) Prefixed formatives. The only productive prefixed formative used in the medieval period is *lā+*, the meaning of which is that of negation, e.g. *lā+nihāya* 'infinite' (also: *al-lā+nihāya* 'the infinite'); *lā+nuqṭiyya* 'astigmatism'; *lā+lawnī* 'achromatic'. These compounds compete with syntactic constructions of equivalent meaning, using *ġayr* or *ʿadam*, e.g. *ġayr nihāya*, *ʿadam (an-)nuqṭiyya*. Attempts were made to create other prefixed formatives, but only a few authors use them. Syntactic constructions are widely preferred, although they are longer and less convenient, e.g. *qab+* (from *qabla* 'before'), in *qab+tārīxī* 'prehistorical', the corresponding syntactic-based compound of which is *mā qabla t-tārīx*, and also *qab+madārī* 'preorbital'. This process is considered by Arab lexicographers to be related to *naḥt* (see below, under iv). In some cases, prefixed formatives are borrowed, e.g. *bārā+* < English *para-*, as in *bārā+miḥwarī* 'paraxial', or *mītā+* < English *meta-*, as in *mītā+luġawī* 'metalinguistic'. Some prefixed formatives have been successful in Arabic technical terminology, as opposed to general vocabulary, where they meet with much greater resistance than, for instance, in

Modern Hebrew, e.g. *kahra+* ‘electro-’, in *kahra+miġnā-ṭīsī* ‘electromagnetic’, *kahra+ḍawʾī* ‘photoelectric’, *kahra+baṣarī* ‘photoelectric’ and at the same time ‘electro-optical’.

(b) Suffixes. Arabic suffixes are very often compositional, but some of them, such as the relative adjectival suffix *+ī* (i.e. *-iyyun*, the *yāʾ an-nisba*; → *nisba*) and the suffix *+at* (*+a(h)* in closing position), which denotes the *res generalis,* i.e. ‘the thing that…’ or ‘that which…’ (Roman 1999), are liable to be included in a ‘frozen’ lexical compound, in which stem and suffix constitute an ‘undividable’ lexical item (Dichy 2003), e.g., the 1st/7th century ‘frozen’ compound *xārij+ī* ‘belonging to the Kharijite sect’ can be contrasted with *xārij+ī* ‘outer’ (as in *al-bāb al-xārijī* ‘the outer door’), in which the relation between the stem *xārij* ‘outside’ and the suffix *+i(yy)* remains compositional; significantly, the plural form *xawārij* is related to the former ‘frozen’ compound *xārij+ī.* In the building of terminological relative adjectives and nouns, the suffix *+i(yy)* is widely used and polyvalent, e.g. *miḥwar+ī* ‘axial’ or ‘central’. Two compound variants can be observed: *+āni(yy)*, which was still used in Medieval Arabic, has not fallen out of use, e.g. *ṭūl+ānī* ‘longitudinal’, which coexists with *ṭūl+ī*, sharing the same sense; and *+awi(yy),* e.g. *nisb+awī* ‘relativistic’, as opposed to *nisb+ī* ‘relative’ (also used for ‘relativistic’). Another compound suffix including *+ī* is *+iyya(t)*, mainly used for abstract nouns or concepts, e.g. *mijhar+iyya* ‘microscopy’, *ʾišʿāʿ+iyya* ‘radioactivity’, *fusfūr+iyya* ‘phosphorescence’. Its plural form is also used to name scientific fields, such as *riyāḍ+iyyāt* ‘mathematics’, *ʾiliktrūn+iyyāt* ‘electronics’, *lisān+iyyāt* ‘linguistics’. In some cases, Arabic resorts to the plural form of the stem, e.g. *naḏ̣āʾir+ī* ‘isotopic’, built on the plural of *naḏ̣īr* (*naḏ̣āʾir*), ‘isotope’. This morphological means is neither contemporary nor recent: the term *kutub+ī* ‘dealer in manuscripts, librarian’, built on the plural form of *kitāb* ‘writing, book’, *kutub*, goes back to the first centuries of Islam.

In some productive terminological domains, particularly in chemistry, many suffixed formatives have been borrowed from English or French (although they cannot be regarded as new suffixes in Arabic), e.g. *faḥm+āt* ‘carbonate’, *kibrīt+ūr* ‘sulfide’ < French *sulfure*.

iii. Compound words encompassing two existing lexical entries are rarely used. The usually quoted example is *raʾs-u# māl* ‘capital’, written as a single word, from *raʾs-u māl-in*, the structure of which is that of a ‘construct state’ (→ *ʾiḍāfa*). Two plural forms are observed: *rasāmīl*, single word, and *ruʾūs ʾamwāl*, in two words.

iv. Composition by amalgam (or ‘portmanteau merging’) of two lexical entries. In this type of compound, two lexical units are merged, both formally and semantically, in a single ‘portmanteau word’ (“two meanings packed into one word, like a portmanteau”, from Lewis Carroll’s preface to *The hunting of the Snark* [1876]), e.g. *kahra+ṭīsī* ‘electromagnetic’, from *kahra[bāʾī]* and *[miġnā]ṭīsī*; or *zamakān* ‘space-time [continuum]’ (also: ‘time and place’ in the grammatical phrase *ism az-zamakān* ‘name of time and place’), which is built on *zamān* ‘time’ and *makān* ‘place’. This lexical device can be related to *naḥt* compounding, whose use is ancient but which is only accepted by terminological authorities ‘in case of necessity’ (→ compounds). It is often, in fact, difficult to analyze.

3.2 *Syntactic means*

Terms coined by resorting to a ‘freezing’ of syntactic structures are complex units, encompassing two words at least. All terms of this type feature a binary structure of two smaller components, a basis and an extension (Roman 1999, 2001), each of which may in turn eventually be analyzed into two smaller components, and so on. The basis-extension relation may be either coordination (*ḏahāb wa-ʾiyyāb* ‘round trip’) or, more frequently, subordination.

Subordination relations may be of several types: (i) attributive noun-and-adjective relation (*naʿt wa-manʿūt*), e.g. *ṭūl mawjī* ‘wavelength’; (ii) ‘construct state’ noun-noun relation (*ʾiḍāfa*), the second of which bears the genitive case, e.g. *ṭūl mawjat(-in)* ‘wavelength’;

(iii) prepositional phrase (*jārr wa-majrūr*) 'governed' by a word or a phrase, e.g. *istiqṭāb bi-l-inʿikās* 'polarization by reflection'; (iv) modal relation, the noun or adjective in consideration bearing the accusative case, e.g. *mustaqṭab dāʾiriyy-an* 'circularly polarized'. In addition, compound terms may feature a combination of these constructions, as in {{*intišār aḍ-ḍawʾ*} *fī* {*xuṭūṭ mustaqīma*}}, lit. '{{propagation [of] the light} in {lines straight}}', 'rectilinear propagation of light', where both basis and extension are complex components.

A great majority of terms are complex units. The proportion is much higher than, say, in English: many English terms are formed by a morphological combination of Greek and Latin, as well as English elements, including very rich sets of prefixes, suffixes, and prefixed or suffixed lexical formatives. In contrast, such combinations are much more constricted in Arabic, owing to the general structure of morpholexical relations (Dichy 2003). Arabic very often resorts to 'frozen' syntactic compounds, such as *miqyās inkisār* 'refractometer', *qābil li-l-inkisār* 'breakable', as shown above.

These constructions feature some innovations in the creation of terminological compounds (Ali 1987; Lelubre 1992, 2001), as opposed to compounds resulting from general vocabulary formation. Four of them are briefly presented here:

i. Adjective specifying another adjective, e.g. *ḥarārī nawawī* 'thermonuclear', *kahrabāʾī miġnāṭīsī* 'electromagnetic', *kahrabāʾī sākin* 'electrostatic'.

ii. Adjective preceded by a functional word expressing circumstance (*ḏ̣arf* 'adverb [of time and place]'), e.g. *taḥta ʾaḥmar* 'infrared' (*taḥta* 'under'); *fawqa ṣawtī* 'ultrasonic' (*fawqa* 'above'). This differs to some extent from more Classical Arabic constructions, such as *šibh mustaqirr* 'metastable' or *ġayr mutajānis* 'heterogeneous'. The above adjectival structures can in turn be included in another compound, which allows for their nominalization when required by the sentence structure, e.g. *mā warāʾa ṭ-ṭabīʿa* 'that which [lies] behind nature' i.e. 'metaphysical' or 'metaphysics'; *mā taḥta l-ʾaḥmar* 'infrared'; *mā qabla t-tārīx* 'prehistorical' or 'prehistory' (besides the use of the adjectival phrase *qabla tārīxī*).

iii. Asyndetical constructions between two nouns in juxtaposition, i.e. nouns that are not related to each other by a 'construct state' structure, are sometimes found, e.g. *šuʿāʿ surʿa* 'vector velocity', which gives, with the article *al-*: *aš-šuʿāʿ as-surʿa*. Such a construction coexists with the usual construct state, *šuʿāʿ as-surʿa*.

iv. Use of the hyphen or slash (→ punctuation), as in *miqyās tadāxul Fābrī – Bīrū* 'Fabry-Perot etalon'. The traditional construction *miqyās tadāxul Fābrī wa-Bīrū* is also found.

3.3 *Semantic processes*

In this type of terminological creation, one encounters no new linguistic form but rather a renewed use of forms and structures existing in either the common vocabulary or the specialized vocabulary of another field.

i. Tropic transfer of meaning is related to what Arabic rhetoric describes as → *majāz* 'tropic speech'. Semantic transfer operates according to various types of tropic processes, most of which are relatable to → metaphor, metonymy, and hypallage:

   (a) Metaphor is based on a similarity between two things or a link that can be imagined between them. Similarity is based on form or shape in the case of *ʿadas+a* 'lens' in optics (from *ʿadas* 'lentil'), and *mawj+a* 'wave' in physics (from *mawj+a* 'sea wave'). It is based on function in *ḏākir+a* 'memory' (of a computer, from *ḏākir+a* '[human] memory'). Very often, metaphors used in Arabic are borrowed from similar ones encountered in English or French, as in the previous examples. However, transposition may sometimes be problematic, when the link does not appear natural in the eyes of Arabic speakers.

   (b) Metonymy is based on a referential relation or link between two entities. This is found, for example, between the event or process described by an infinitive form (→ *maṣdar*) and its result, as in *tasjīl* 'operation of recording' (event or process), and 'recording; recorded sounds' (result). This is also the case in the 'part-for-whole' relation (the part being, classically, given for the whole), as in *ḍawʾ* meaning both 'light' and 'optics' ('the science [whole] of light [part]'). Such tropic semantic transfers may sometimes be a source of ambiguity.

(c) In hypallage, a syntactic extension related to an omitted basis is attributed to another basis (Roman 1999), as in *ittiṣālāt baṣariyya* 'optical telecommunications', which stands for 'telecommunications with optical [means]' (*ittiṣālāt [bi-wasāʾil] baṣariyya*). This process may lead to collocations that may seem strange at first sight, as in *qānūn al-jarāʾim* 'criminal law; penal code' where the English adjective 'criminal' and the second noun of the construct-state structure in Arabic, *jarāʾim* 'crimes', refer to offenses judged under this part of the law code, and not to an attribute of the law itself (cf. in French *police criminelle*). It is, on the other hand, very productive in both noun+adjective and construct-state structures, e.g. *šarīṭ al-qiyās* 'measuring tape'; *šarīṭ nāṭiq* 'sound track [of a film]; magnetic tape' – the latter being also expressed, through another hypallage, as *šarīṭ tasjīl* 'tape[-of-]recording'.

ii. Neologisms, resulting from the new use of an old term (*istinbāṭ*), are recommended by some members of the Arabic language academies, under the condition that the ancient term be out of usage. In fact, the restriction is not very efficient in most scientific fields, apart from a few successful examples, such as *qiṭār* 'train' (anciently: 'convoy of camels'), because, in many cases, the ancient and the new sense coexist, e.g. *tarqīm* 'numbering' (ancient meaning) and 'coding'.

iii. Calques (also called 'loan translations') from another language are situated between the above semantic processes and direct borrowing. In such cases, Arabic terms exhibit the same features as their English or French counterparts. In fact, most Arabic terminological creation is based on foreign terms, adapted through means provided by Arabic morphology and syntax. Most of the examples above may be considered as resulting from calque or loan translation. This is the case in most multilingual terminology and should not be considered a problem, except when new terms are regarded by native speakers as an intrusive transfer from foreign languages.

3.4 *Borrowing*

The status of borrowing (*taʿrīb*) varies according to fields and concepts. For example, borrowing is necessary in the case of most chemical elements or for physical units. It may also be resorted to in the absence of or pending the creation of a recognized Arabic equivalent. Yet, both forms may coexist for a long time; *talifūn* 'telephone', for instance, is still in use nowadays, alongside *hātif* (attested first in 1924 in this sense; originally, it meant 'unseen man whose voice is heard'), which is the only surviving denomination among other challengers. The borrowed term *kambyūtar* 'computer' may resist for a long time against its challenger *ḥāsūb*, while *mīkrūskūb* goes on coexisting with *mijhar*.

In scientific domains, borrowed terms generally consist of single words, which are more or less modified. Phonemes that do not exist in Arabic are replaced with their nearest correspondent (e.g. /b/ for /p/). Realizations of borrowed words may differ according to the way in which they are pronounced in their original language, e.g. for 'laser' *lāyzir, līzir* (English) vs. *lāzir* (French). Syllabic modification may occur, such as the addition of a prosthetic syllable to avoid initial consonant clusters, e.g. *ʾisbiktrūskūb* 'spectroscope' (alongside *sbiktrūskūb*), or the addition of a vowel to avoid three-consonant sequences, e.g. *ʾiliktirūn* 'electron', although *ʾiliktrūn* occurs more frequently. In writing, the use of long vowels before two consonant clusters, such as *ī* and *ū* in *mīkrūskūb*, helps in pronouncing words that do not have an Arabic pattern, because short vowels are generally not written. The use of long vowels for this purpose is not new; it appeared three millennia ago in other Semitic scripts, in the writing of borrowed words (cf. the *matres lectionis* of Hebrew and Aramaic, still very much in use in Modern Hebrew terminology).

Arabic may extract from the borrowed term a → root – often a quadriconsonantal one – and generate Arabic forms by means of internal derivation (→ *ištiqāq*; see the above examples of *taʾaksud* 'oxidation', etc.).

Many acronyms are borrowed, such as *DNA* (→ abbreviations). They are usually not translated but are either kept with their Latin notation or transcribed in Arabic letters. In scientific symbols and notations, we find two types of acronyms:

i. International symbols and notations (eventually including Arabic numerals) are written from left to right; this is nowadays the case in the chemical domain.
ii. Arabic symbols and notations are written from right to left, and include the use of various calligraphic styles; symbols in physics are generally transcribed in Arabic writing.

4. CONCLUSION

In medieval times, the Arabic-speaking culture efficiently coined terminology for its own needs and purposes. Nowadays, Arabic speakers can use their language to address any scientific or technical question, but specialists may still encounter some difficulties in using Arabic in their field when communication involves several countries of the Arab world.

BIBLIOGRAPHICAL REFERENCES

Ali, Abdul Sahib Mehdi. 1987. *A linguistic study of the development of scientific vocabulary in Standard Arabic*. London and Boston: Kegan Paul International.

ʿAmmār, Sām and Šahāda al-Xūrī. 1996. *at-Taʿrīb fī l-waṭan al-ʿarabī: Wāqiʿu-hu wa-mustaqbalu-hu (min ʾajl xiṭṭa ʿāmma li-t-taʿrīb)*.Tunis: ALECSO.

Dichy, Joseph. 1989. "Vers un modèle d'analyse automatique du mot graphique non-vocalisé en arabe". *Simulation de modèles linguistiques et enseignement assisté par ordinateur de l'arabe – Travaux SAMIA I*, ed. Joseph Dichy and Mohamed Hassoun, 92–158. Paris: Conseil International de la Langue Française.

——. 1998. "Mémoire des racines et mémoire des mots: Le lexique stratifié de l'arabe". *La mémoire des mots*, ed. Taieb Baccouche, André Clas, and Salah Mejri, 93–107. *Revue Tunisienne de Sciences Sociales* 117.

——. 2003. "Sens des schèmes et sens des racines en arabe: Le principe de figement lexical (PFL) et ses effets sur le lexique d'une langue sémitique". *La polysémie ou l'empire des sens*, ed. Sylvianne Rémi-Giraud and Louis Panier, 189–211. Lyon: Presses Universitaires de Lyon.

Hamzaoui, Rached [Ḥamzāwī, Muḥammad Rašād]. 1975. *L'Académie de langue arabe du Caire*. Tunis: Publications de l'Université de Tunis.

——. 1991. *al-Muʿjam al-ʿarabī: ʾIškālāt wa-muqārabāt*. Carthage: Bayt al-Ḥikma.

Ḥigāzī, Maḥmūd Fahmī. 1993. *al-ʾUsus al-luġawiyya li-ʿilm al-muṣṭalaḥ*. Cairo: Maktaba Ġurayyib.

Jacquart, Danielle (ed.). 1994. *La formation du vocabulaire scientifique et intellectuel dans le monde arabe*. Brepols, Turnhout (Belgium): CIVIMA.

Lelubre, Xavier. 1992. *La terminologie arabe contemporaine de l'optique: Faits – théories – évaluation*. Thèse de Doctorat, Université Lumière-Lyon 2.

——. 2001. "A scientific Arabic terms data base: Linguistic approach for a representation of lexical and terminological features". *ACL 39th Annual Meeting, Workshop on Arabic Language Processing; Status and Prospect*, 66–72. Toulouse.

Monteil, Vincent. 1960. *L'arabe moderne*. Paris: Klincksieck.

Rabin, Chaim. 1960. "*ʿArabiyya*". *Encyclopaedia of Islam,* I, 561–567. 2nd ed. Leiden: E.J. Brill.

Richert, Nicole. 1987. *Arabisation et technologie*. Rabat: Institut d'Études et de Recherche pour l'Arabisation.

Roman, André. 1999. *La création lexicale en arabe: Ressources et limites de la nomination dans une langue humaine naturelle*. Lyon: Presses Universitaires de Lyon.

——. 2001. *Systématique de la langue arabe*. 2 vols. Kaslik (Lebanon): Université Saint-Esprit de Kaslik.

Sawaie, Mohammed. 1999. *The crisis of Arabic terminology in the 19th century*. Damascus: Institut Français d'Etudes Arabes. [In Arabic.]

Šayyāl, Jamāl ad-Dīn aš-. 1951. *Tārīx at-tarjama wa-l-ḥaraka aṯ-ṯaqāfiyya fī ʿaṣr Muḥammad ʿAlī*. Cairo: Dār al-Fikr al-ʿArabī.

Šihābī, Muṣṭafā aš-. 1965. *al-Muṣṭalaḥāt al-ʿilmiyya fī l-luġa al-ʿarabiyya fī l-qadīm wa-l-ḥadīṯ*. 2nd ed. Damascus: Syrian Academy.

Stetkevych, Jaroslav. 1970. *The Modern Arabic literary language: Lexical and stylistic developments*. Chicago and London: University of Chicago Press.

XAVIER LELUBRE AND JOSEPH DICHY
(Université Lumière-Lyon-2)

Terms of Address

1. INTRODUCTION

The terms 'address form' and 'term of address' indicate any linguistic form used by speakers to refer to the person they are talking to (the addressee). These include pronouns, honorific 'pronoun substitutes', names, nicknames, teknonyms, titles, and other words used vocatively. Arabic varieties have overt vocative markers (the most common being the particle *yā*, which occurs before the name or term), but in many instances a term may be used vocatively without the vocative particle. Terms of address are used for a variety of communicative functions, including getting the addressee's attention; summoning, establishing, and maintaining contact with the addressee throughout a conversation; ordering and cajoling; and 'calling' the addressee something. Terms of address are often nouns with a lexical meaning that can also be used in nonvocative reference (when talking about someone rather than to someone). Although these nouns are often used as terms of address directed to their literal referent (using the noun for 'brother', for example, to address one's own brother), they are even more often extended metaphorically to nonliteral referents (using the noun for 'brother' to address one's friend, or even a stranger with whom one wants to indicate some kind of solidarity).

2. Arabic vocative grammar

In Standard Arabic, the vocative particle *yā* is normally followed by a name or an indefinite nominative noun, but without nunation (*yā ʾustāḏu* 'O Professor'). The noun may be the first term of an → *ʾiḍāfa* construction or have an attached clitic pronoun, but in that case it will be accusative instead of nominative (*yā ʾustāḏa l-ʾawlādi* 'O Professor [of the children]', *yā ʾustāḏanā* '[our] Professor!', *yā jalālata l-malik* 'Your Highness, the King'). In any case, the first noun after *yā* may not have a definite article. In contrast, the vocative particle *ʾayyuhā* must be followed by a noun with the definite article (*ʾayyuhā l-ʾustāḏu* 'O Professor'). The particle *yā* may be omitted, as in the common *sayyidātī wa-sādatī* 'Ladies and Gentlemen', while the omission of *ʾayyuhā* before a definite term is not allowed while retaining vocative force (*ʾayyuhā s-sayyidātu wa-s-sādatu* 'Ladies and Gentlemen').

Word order is always vocative particle + term of address. In compound terms of address (those that include a title and a name), the order is vocative particle + title + name (*yā ʾustāḏ Muḥammad* 'Professor Muhammad!'), except in the case of a set of borrowed Turkish terms, for which the order is vocative particle + name + title (*yā Fuʿād bēh/bak* 'Mr. Fuʿad', *yā Karīma hānim* 'Mrs. Karima', *yā ʾAḥmad bāšā* 'Mr. Ahmad'; note that these terms do not actually mean 'Mr.' or 'Mrs.', but there is no close equivalent in English; they refer to positions in the Ottoman hierarchy).

In the dialects, which in general do not mark case or have → nunation, most of the above does not apply. One additional feature in some dialects is that the vocative particle can be repeated before elements of a non-*ʾiḍāfa* multiword term of address, usually to express either exasperation or playfulness: *ya wad ya tiʾīl* 'you heavy boy', *ya Muḥammad ya ʿOsmān* 'hey, Muhammad Osman!'. Note that in Egyptian and many other dialects, the vocative particle has a short vowel rather than a long vowel as in Standard Arabic.

3. Frequency and meaning

In at least some of the Arabic dialects which have been studied (see Parkinson 1985 for Egyptian; Al-Khatib 2003, Braun 1988, and Farghal and Shakir 1994 for Jordanian; Potter 1995 for Moroccan), there are a large number of terms available for use as terms of address, and they are in fact used very frequently. Counting such things, of course, is never straightforward, and comparisons between languages are fraught with problems, but it still seems clearly to be the case that, at least in some dialects, there is a much larger number of terms in common use than in English, and speakers choose to use a term much more often than English speakers do.

Most terms have an overt 'dictionary' meaning, which may refer to a relationship (brother), a job or position (doctor, engineer, minister), a quality (stingy), and the like. However, it is clear that each dialect's system of terms of address involves a well-understood set of norms for use and nonuse so that, because of expectations of use or nonuse, a particular term in a particular situation takes on additional meanings beyond its dictionary definition. 'Expected' uses mark assumed, 'normal' relationships between speaker and addressee (often in terms of an intimacy scale, a power scale, an age scale, and a few others), and 'unexpected' uses mark some violation of the 'normal' and may indicate anger, shock, playfulness, or affection. In this system, avoiding using a term at all can be just as meaningful as using one, since hearers who are accustomed to hearing themselves addressed in a certain way will soon notice that it is not happening and assign some interpretation to that fact (Hussein 1995).

4. PRONOUNS AND PRONOUN SUBSTITUTES

All forms of Arabic have independent 'subject' pronouns and attached, clitic 'object' pronouns. Standard Arabic has a full complement of singular, dual, and plural masculine and feminine 2nd person forms (*ʾanta, ʾanti, ʾantumā, ʾantum, ʾantunna; -ka, -ki, -kumā, -kum, -kunna*), while the dialects have a reduced number, usually just three: masculine singular, feminine singular, and a single plural form (for example, Egyptian *ʾinta, ʾinti, ʾintu; -ak, -ik, -ku*). Verbs are conjugated to agree with these pronouns, so it is not necessary to include the subject pronoun when a verb is present. The pronoun system itself has no counterpart to the informal/formal 2nd person pronoun system common in European languages (as in French *tu/vous*). Everyone is addressed using the regular 2nd person forms, and the plural form is used to address a singular addressee only on relatively rare occasions, with an extremely high addressee (the king, the president). What many dialects of Arabic do have, however, is a set of what may be called 'pronoun substitutes', honorific terms which can replace either the subject or object pronoun wherever it occurs (see Parkinson 1985; Alrabaa 1987). Using Egyptian Arabic as an example, terms like *ḥaḍritak* 'Your Presence' and *siyadtak* 'Your Excellency' are used to replace *ʾinta* and its clitic counterpart. However, unlike the use of *vous* in French, this replacement does not happen every time a 2nd person form appears, and 2nd person verb forms appear both with and without it in the same conversation. It is enough to sprinkle the honorific forms every so often to have the whole speech marked as respectful. Thus, there is not the same degree of exclusivity in choice of forms. This means that a single use of the plain pronoun to a respected addressee does not have the same insulting quality that using *tu* in French would have. Other, less common, honorifics are also used, such as *faḍiltak* 'Your Nobility' and *saʿadtak* 'Your Felicity'.

Holes (1986) reports on an interesting situation in Bahrain in which the pronunciation of the clitic 2nd person pronouns was different (and thus a marker of identity) for the three main communities, the urban Sunnis, the urban Shiʿites, and the rural Shiʿites. While the two urban groups do not vary their system, the rural Shiʿites change to the urban Shiʿite system with certain types of ritualistic 2nd person reference (where Holes claims the values of the larger community are invoked), and in general when talking to members of either of the other groups. He shows how this happens variably rather than categorically, and could be explained as a means of setting the frame of reference of a conversation depending on what roles are being assumed and what is being communicated.

5. NAMES

Personal names are important markers of identity in all Arabic-speaking countries (see Parkinson 1985; → proper names). In some places with large Christian minorities, personal names are marked either as Muslim (*Muḥammad, ʿAli*), Christian (*Ḥanna*), or neutral (*Karīm*). A few names borrowed from European languages are used, particularly among Christians (*Antoine*), and names of famous world political leaders also show up (*Castro*). More conservative or religious Muslim families tend to stick to traditional Islamic names, particularly ones that resonate in Islamic history, while some of these names are felt as somewhat old-fashioned in less conservative families. It is fairly common to be named after a relative, particularly after a grandfather. Some families also enjoy giving each child a name that begins with the same letter.

The system for names beyond the first or personal name varies quite dramatically from country to country. In the Gulf it is still common to use the word *ibn* or *bin* 'son (of)' in between the personal name and the name of the father (and sometimes grandfather). In Egypt, the *ibn* is dispensed with, but one's full name consists of one's own given name, one's father's given name, and on formal occasions the grandfather's given name (*al-ism al-muṯallaṯ*). Some Egyptian families also use a 'family' name, which might be the profession or place of origin of the original ancestor (*al-ʿAṭṭār* 'the perfume salesman', *Dāġistānī* 'from Daghestan'), but use of these names is optional and sporadic, and only some members of a family may like using them. Thus, siblings in Egypt will sometimes seem to have different last names, if one chooses to use the family name and another not. Most Egyptians simply do not have one available. In countries like Lebanon, however, it has been a government requirement for several generations that families must choose a family name, so in those countries, last names are expected, and their use is closer to European last name usage.

Names may be used alone or in conjunction with a term of respect, and terms of respect may also be used without the name. Names alone imply intimacy, equality, or power (of speaker over addressee), while terms with a name imply acquaintance but not intimacy, or lack of power (of speaker over addressee), while use of a term alone implies respect without acquaintance. It has been pointed out by more than one researcher that the younger generation has begun exchanging first names much more quickly than used to be the case in traditional Arab society (see Nydell 1987). In most Arab countries, it is unusual to use a term with a last or family name; rather, almost all term + name combinations involve the first given name (*ʾustāẓ ʾAḥmad*, not *ʾustāẓ Dāġistānī*; see Potter 1995; Nydell 1987; Parkinson 1985).

Besides names, many people also have nicknames. There is a large set of nicknames (not the same for each country) that are known to 'go with' particular names (*Ḥamada* for *Muḥammad*, *ʿAdūla* for *ʿĀdil*, *Sūsu* for *Suʿād* in Egypt), but besides these, people may pick up random nicknames that are unique to themselves. Some nicknames are 'fake' teknonyms (see below), which either traditionally go with that name or are picked up randomly. Nicknames are used to mark intimacy.

Parents with children are often called *ʾabū fulān* 'father of so-and-so' and *ʾumm fulān* 'mother of so-and-so' after their oldest son, and occasionally after the oldest daughter if there is no son. Use of these terms varies. In villages and more traditional places, most people start using the teknonyms once a child is born, and the name is rarely heard thereafter, even in the family context (see Yassin 1978), while in more urban or less traditional settings the teknonym functions as a polite/nice way of addressing friends and acquaintances but does not completely replace the name or the use of title + name (→ politeness). As noted above, some teknonyms are 'fake,' not referring to a real son, and function more as nicknames.

6. KINSHIP TERMS

When addressing actual kin, people use either the first name (particularly when speaking to an addressee of the same age or younger), or a kin name, usually with a 1st person pronoun (*ya ʿammi* 'my uncle', *yabni* 'my son'). A few terms are used without the pronoun (*ya bāba, ya māma, yāba* in Egyptian). More distant relatives are often addressed with terms that bring them closer, so that an older second cousin might be addressed as 'my uncle' and the wife of one's uncle as *ya xalti* 'my aunt'. Borrowed terms like *ṭanṭ* are common in several places. For the most common classes of relatives there is sometimes a whole set of possibilities, and speaker choices often correlate with social class and level of education. Thus, when addressing parents, working-class speakers in Egypt are more likely to utter *ʾāba* or *ʾamma*, while middle-class speakers have mainly *bāba* and *māma*. Upper-class speakers sometimes add *māmi*.

These terms are not limited to addressing relatives, however. They are extended to acquaintances and strangers both in a polite way (bringing them into the family, so to speak) and in annoyance or sarcasm. Older men, even strangers, are commonly addressed as *ʿammi* 'my uncle', and older women with *xalti* 'my aunt'. The forms *ʾax* 'brother' and *ʾuxt* 'sister' are often used without a pronoun ending to mark a general solidarity, and the forms with the 1st person pronoun ending (*ʾuxti, ʾaxi, ʾaxūya*) can be used to express annoyance.

One of the more unusual aspects of the system of Arabic terms of address involves an extended use of → kinship terms. This has been called the bipolar use of the terms, or reverse-role use (see Ayoub 1962; Rieschild 1998; Yassin 1977; Parkinson 1985; Holes 1986), in which a term that is appropriate for the addressee to use to the speaker is used by the speaker instead to address the addressee. Thus, in Egypt, a father can say *ya bāba* 'dad' to his daughter or son, and in Lebanon a grandfather can call his grandson *ya židdi* 'my grandfather'. Researchers agree that these uses are affectionate, and are usually used in situations where the older relative could give a direct command and expect to be obeyed but instead wants to mitigate the force of the command and simply cajole the younger person to do what is requested on the basis of their intimate relationship.

Interestingly, in the Gulf, instead of using a 1st person pronoun, the speaker may use a 2nd person pronoun; thus, instead of a father calling his son *ʾabūya*, he calls him *ʾabūk* 'your father', using *ʾabūč* to his daughter (Yassin 1977). This kind of usage has been extended in places like Kuwait to conversations between same-age relatives (calling your sister 'your brother') and then extended, using the kin terms metaphorically, to friends and other non-kin. Thus, while most Arab countries appear to have some form of reverse-role kin term usage, the degree of use (less in urban Egypt, more in village Lebanon, even more in Kuwait) and the actual way the system works vary widely. One would venture to guess, for example, that Kuwaiti usage would confuse Egyptians hearing it for the first time.

7. Terms of respect

Common terms of respect include names of professions (*ʾustāz* 'professor', *duktōr* 'doctor', *bašmuhandis* 'chief engineer'), terms derived from army or police ranks (*šawīš* 'sergeant'), terms derived from Turkish nobility (*bēh* 'bey', *bāša* 'basha', *hānim* 'madame'), general terms or respect (*ʾafandim* 'sir', *sayyid* 'Mr.', *sitt* 'Mrs.', *madām* 'madame'), terms for working-class professions (*ʾusṭa* 'boss', *miʿallim* 'teacher/boss', *dāda* 'governess/janitress'), terms for religious leaders (*šēx* 'sheikh', *ʾabūna* '[our] father'), age-related terms (*ḥagg* 'pilgrim'), and two-word terms that involve an honorific and the name of a job (*siyadit irraʾīs* 'dominance of the president', *maʿāli l-wazīr* 'highness of the minister', *ḥaḍrit ilmušrif* 'presence of the director'). Most of these terms are used with or without names. They must be considered the core of the address form system, being used on a daily basis to establish and mark formal relationships. Some terms are most commonly restricted to those who 'deserve' them in some sense, while others are extended broadly either to any educated person or to anyone who is in a role that demands respect (a customer, for example). Thus, in Egypt, the term *duktōr* is extended only relatively rarely to addressees who are not medical doctors or those without a Ph.D., while the terms *bašmuhandis* and *ʾustāz* are used without restraint to address any respectable-looking person. This creates the interesting situation in which professors are referred to with *ʾustāz* 'professor' in the 3rd person, but addressed mainly with *duktōr*, which still retains much more prestige than *ʾustāz*. The Turkish terms were banned early in the revolution, so there are no more 'real' *bēh*s and *bāša*s, but the terms are still commonly used to address customers and bosses seriously, and to friends and others in jest.

8. Friendly and accusatory terms

Some dialects, most noticeably Egyptian, contain a large number of terms used for the functions of *mugamla* 'treating people nicely, buttering them up', *muʿākasa* 'teasing, bothering strange women on the street', *hizār* 'joking', and *lōm* 'accusing'. Friendly terms include those like *ḥabībi* 'my lover', *šāṭir* 'clever', *ʿarūsa* 'bride', *ʾamīra* 'princess', and many others, which may be used to address actual loved ones but also just about anyone on the street whom the speaker wants to treat nicely. Terms implying beauty include *gamīl* 'pretty' (often the masculine form is used to a feminine addressee), *ʾamar* 'moon', *ḥilwa* 'sweet', *bēḍa* 'white', and many others. There are many animal and food terms used in a nice or teasing way, like *ʾuṭṭa* 'kitty', *katkūta* 'chick', *sukkar* 'sugar', *šarbāt* 'sweet drink', *ʿasal* 'honey', and *ḥalāwa* 'sweetness/dessert'. Finally, there are a number of seemingly neutral terms, like *rāgil* 'man', *mara* 'woman', *gadaʿ* 'strong young man', *bani ʾādam* 'human beings', and *bint* 'girl', which are typically used without an attached name or pronoun ending to express annoyance or criticism of the addressee, without actually being abusive. A great deal of creativity in this area can be observed, with speakers creating new terms appropriate to the situation, which creates a playful undercurrent to the use of many of these kinds of terms. The literature does not indicate that speakers in other countries are quite so attached to the use of this category of terms as Egyptians are.

9. Terms of abuse

Unlike in English, where the most common way to abuse someone verbally is to call him something ('you are a ______', or 'you ______'), in Arabic the most common way is to use a term of abuse after the vocative particle *ya* (→ insults). Terms of abuse are used seriously, in anger, but also quite commonly among friends, with almost the opposite meaning (something like 'you are really a close friend'). Terms of abuse are clearly marked for social class of speaker, with 'light' terms used mainly by middle- and upper-class speakers, and 'heavy' terms limited to working-class speakers. However, upper-class young men may mimic working-class usage among themselves as a kind of joke. It is extremely rare to hear a term of abuse 'up' to someone higher on any social scale of power or age, and middle-class speakers also avoid them with peers whom they do not know well. The terms are, nevertheless, very common. Light terms of abuse include names of animals (*ḥimār* 'donkey', *kalb* 'dog'), words for negative attributes (*gabān* 'coward', *ṭifis* 'stingy', *ʾaʿma* 'blind', *magnūn* 'crazy'), or names of professions considered lowly (*mugrim* 'criminal', *ḥarāmi* 'thief', *bint ilġassāla* 'daughter of a washerwoman'). They also include the use of terms implying some harmful result to the addressee (*ya maksūr irraʾaba* 'you whose neck will be broken!'). Heavy terms imply something about the parentage of the addressee (*ibn ilkalb* 'son of a dog', *ibn ilwisxa* 'son of a prostitute'), or that something is sexually or religiously wrong with the addressee (words referring to homosexuality, prostitution, the sexual status of one's mother, or to the religious status of one's father, put in a negative light). Light terms of abuse constitute an area of creativity in the creation of new terms almost as active as the one for friendly terms; many of the heavy terms, on the other hand, are associated with a degree of → taboo and are not uttered by many women, or by anyone in polite society.

10. SUMMARY

Terms of address in Arabic-speaking communities are sensitive social markers, and speakers are very sensitive to their use, feeling offended when they do not feel they have been addressed properly. The actual terms available for use and the precise details of how they are used vary greatly from place to place in the Arab world, but all places appear to share the basic syntax of their use and the idea of their importance, their frequent use, and their extendability beyond their basic meaning. A few scholars have begun to look at modeling or predicting term usage based on social aspects of an interaction (Skousen 1989:97–100, 181–184), although this work is still in its infancy.

BIBLIOGRAPHICAL REFERENCES

Al-Khatib, Mahmoud A. 2003. "Address norms in Jordanian Arabic: A sociolinguistic perspective". *Grazer Linguistische Studien* 59.1–20.

Alrabaa, Sami. 1987. "Social usage of address forms in Egypt: A sociolinguistic study". *Zeitschrift für Arabische Linguistik* 16.123–129.

Ayoub, Millicent R. 1962. "Bi-polarity in Arabic kinship terms". *Proceedings of the Ninth International Congress of Linguists*, 1100–1106. The Hague and Paris: Mouton.

Braun, Friederike. 1988. *Terms of address: Problems of patterns and usage in various languages and cultures*. Berlin: Mouton de Gruyter. [Especially 173–195.]

Farghal, Mohammed and Abdullah Shakir. 1994. "Kin terms and titles of address and relational social honorifics in Jordanian Arabic". *Anthropological Linguistics* 36.240–253.

Holes, Clive D. 1986. "Communicative function and pronominal variation in Bahraini Arabic". *Anthropological Linguistics* 28.10–30.

Hussein, Anwar A. 1995. "The sociolinguistic patterns of native Arabic speakers: Implications for the teaching of Arabic as a foreign language". *Applied Language Learning* 6.65–87.

Nydell, Margaret K. 1987. *Understanding Arabs*. Yarmouth: Intercultural Press.

Parkinson, Dilworth B. 1985. *Constructing the social context of communication: Terms of address in Egyptian Arabic*. Berlin: Mouton de Gruyter.

Potter, Terrence M. 1995. "Moroccan terms of address in the film *Dmuʿ al-nadem (Tears of regret)*". *Georgetown Journal of Languages and Linguistics* 3.218–226.

Rieschild, Verna Robertson, 1998. "Lebanese Arabic reverse role vocatives". *Anthropological Linguistics* 40:4.617–641.

Skousen, Royal. 1989. *Analogical modeling of language*. Dordrecht: Kluwer Academic Publishers.

Yassin, M. Aziz F. 1977. "Bi-polar terms of address in Kuwaiti Arabic". *Bulletin of the School of Oriental and African Studies* 40.297–301.

——. 1978. "Personal names of address in Kuwaiti Arabic". *Anthropological Linguistics* 20.53–63.

DILWORTH B. PARKINSON
(Brigham Young University)

Textlinguistics

Textlinguistics is the discipline that concerns itself with the investigation of the regularities and features of texts. Texts, as discourse realizates and irrespective of their syntactic size, are the basic units of textlinguistics. For any stretch of language to be considered a text, thus displaying full textness or textuality, it should display both text features (TFs), i.e. internal unity (→ cohesion) and conceptual sequence (→ coherence), and user features (UFs), i.e. intentionality, acceptability, informativity, situationality, and intertextuality. Both text features and user features represent the general areas of inquiry in textlinguistics, which overall examines and normalizes aspects of structural and functional constituency of texts, including their classification into typologies, integration of sentence linguistics, stylistics, rhetoric, and text use and comprehensibility in the study of texts and textual phenomena. The constitutive and functional attributes of texts rely on how parts (elements of texts) are defined, how they relate to each other in a hierarchy, and how they are ordered so that users (producers and receivers alike) experience them in a certain configuration and effect. A further tenet in textlinguistics is that text arrangement (structure = cohesion with coherence) is a controllable variable that influences text reception, and that the same information load in a text has a particular effect depending on its structure. Of course, text structure depends heavily on user features for it to realize the communicative-functional purpose of the text.

Broadly speaking, and in both Arabic and Western linguistic traditions, textlinguistics has developed from structuralist foundations (sentence-oriented linguistic models and approaches, but also areas of discourse analysis and other theories like the Prague School) and has been integrated into an independent area of research, leaning toward stylistics and rhetoric. The historical significance of textlinguistics lies in the fact that, in both traditions, it overcame the narrow sentence-specific perspective of linguistics and thereby created a basis for the interdisciplinary study of texts.

In the Arabic linguistic tradition, the study of texts has generally been subsumed under rhetoric, which represents a systematic and comprehensive body of knowledge intended to explore all aspects of textness. In this fashion, Arabic textlinguistics has treated texts as sites of interaction relying on two fundamental principles: a social or communicative basis (text as a means of social interaction between users) and a complex interrelationship among a number of features in which the textual system evolves through specific choice-making to produce a text. A text is a product configured by the communicative need of producers to balance their own purposes and intentions with the expectations and needs of receivers. The focus in Arabic textlinguistics has been to investigate and establish norms for the arrangement of information in texts that can be described as sequences of acts their producers perform or as sequences of effects texts have on their receivers. (In terms of → speech-act theory, these correspond to the illocutionary and perlocutionary acts accomplished by texts.)

In the Arabic tradition, there are three main areas that are concerned with the study of texts: rhetoric (*balāġa*), which deals with oratory (*xaṭāba*), poetry, and the *Qur'ān*, all of which have different textual aspects but aim at convincing (argumentation) through high esthetic features; criticism (*naqd*), which focuses mainly on poetry; and exegesis (*tafsīr*), which focuses on the study of the *Qur'ān* and deals mainly with the coherence of a text that was revealed more than fourteen hundred years ago but is still coherent as a unit and as a self-regulating, balance-maintaining text involving the interaction of various elements (cf. Khatabi 1991).

For some Arabic scholars, *balāġa* subsumes textlinguistics because it examines the main issues of *faṣl* 'conjunction' and *waṣl* 'disjunction'. In his celebrated theory of *naḏm* 'construction/composition', ʿAbd al-Qāhir al-Jurjānī (d. 471/1078) viewed *balāġa* as the most difficult and demanding area of rhetoric. He stipulated it has surface/deep realizations that affect the textness of a text (cohesion and coherence). He further stipulated bases for the realization of this: a grammatical basis (text features) and a semantic and pragmatic basis (user features).

Given the diversity of texts they dealt with, Arabic linguists, particularly medieval scholars, developed approaches for the evaluation, processing, analysis, and interpretation of texts and textness. Depending on the type of texts dealt with, approaches were developed within rhetoric, literary criticism, or exegesis. Arabic linguists, according to Moutaouakil (1990), clearly distinguished between sentence-linguistics and textlinguistics because of their awareness that texts required not only formalistic grammars (focused on sentences as isolated occurrences) but also sentences as chunks of communication. In particular, poetry and the *Qurʾān* display specific characteristics that cannot be handled by sentence-linguistics. However, the beginning of such an important linguistic development is not historically clear. In this context, Versteegh writes: "Not much is known about the earliest beginnings of exegetical activities in Islam, but we do know that they all had in common a fundamental concern with the elucidation of the meaning of the text, rather than the study of its formal characteristics" (1997:13).

To provide interpretations for the Qurʾānic discourse and to argue for its inimitability (→ *ʾiʿjāz al-Qurʾān*) became, and has been, the foremost concern of early Arabic scholars. The driving force was to ascertain the textual coherence of the *Qurʾān*. For this purpose, medieval Arabic scholars distinguished texts in terms of types, a useful analytical tool in modern textlinguistics as well. According to Versteegh (1997), Muqātil (d. 150/767), for example, identified a typology of three text types in the *Qurʾān*: narrative or expository (stories about earlier generations); legal (laws and commandments); and instructive (parables and stories about paradise and hell). Typically, this is a text typology which is also current and used by a majority of modern researchers in textlinguistics and allied disciplines (e.g. Hatim 1997; Maslouh 1990; Faḍl 1996). Although methods and terminology may differ, the subject matter remains essentially the same.

Versteegh (1997) further points to a major shift in Arabic linguistics during the 11th century. The shift primarily emphasized the role of semantics and pragmatics in the investigation of linguistic occurrences. The theologian, grammarian, and literary critic ʿAbd al-Qāhir al-Jurjānī was one of the driving forces behind this shift. He formulated a coherent and comprehensive theory of rhetoric, including textlinguistics, which remains a topic of research and inquiry to date. In his two major and seminal works (*Dalāʾil al-ʾiʿjāz* and *ʾAsrār al-balāġa*), he expounded his central theme, the neglect of semantics and pragmatics by almost all of his contemporaries and predecessors. Because of his concern with the Qurʾānic text, his targets for criticism were not the linguists per se but rather the theologians who, according to him, never managed to effectively demonstrate the reasons for the inimitability of the *Qurʾān*. The central notion in his arguments was the distinction and the relationship between → *lafḏ* 'word' (pl. *ʾalfāḏ*) and → *maʿnā* 'meaning' (pl. *maʿānī*) within a theory of text he labeled *naḏm*. Because the *Qurʾān* is no ordinary text and is revered by Classical and Modern Arabic and Muslim scholars, the study of its textuality has featured in rhetorical and linguistic research since the 6th century, as in al-Jāḥiẓ (d. 255/869), for example. This interest culminated in the introduction of the fully fledged theory of *naḏm* by al-Jurjānī. With its distinct semantic and pragmatic dimensions, this theory was primarily developed to unravel the unique features of the *Qurʾān*, but the textlinguistic nature of the theory has been assessed and reassessed ever since.

Al-Jurjānī strongly believed that the inimitability of the *Qurʾān* must be based primarily on the notion of *naḏm*, which as a term was coined earlier but was generally restricted to aspects of the surface structure of language (cohesion). He raised questions that stimulated interest in the study of texts and gave impetus to semantics and pragmatics, including the sociology and psychology of language, as integral to such a study. It is in this context that al-Jurjānī redirected the whole controversy over the Qurʾānic *ʾiʿjāz* into his theory of *naḏm*, whereby structural relations and their semantico-pragmatic dimensions rather than individual words offer meanings.

Naturally, al-Jurjānī did not operate in an intellectual vacuum. His precursors and contemporaries offered significant contributions, but he remains the one who directly and forcefully linked textness to a semantic and pragmatic basis by proclaiming that words take a certain structure as servants and subordinates to meaning. Although the works of al-Jāḥiẓ, and other writers such as al-Qazwīnī, as-Sakkākī, ar-Rāzī, al-Xafājī, and al-ʿAskarī, contributed to the development of al-Jurjānī's ideas and approach to texlinguistics as a discipline and to making it more accessible and rewarding, it was al-Jurjānī's long exposure to the views of the Muʿtazilite school on the issue of the 'creation' of the *Qurʾān* that armed him with a thorough understanding of the assumptions and methods of this school and prepared him to vehemently challenge and oppose it later on in order to prove the inimitability of the *Qurʾān*. This school of thought employed, among other methods, the concept of *naḏm*, which loosely stands for 'structure' of discourse (their concept of *naḏm* was predominantly limited to the surface realizations of texts, ignoring meaning, particularly at the pragmatic level). For al-Jurjānī, the arguments of the Muʿtazilite school became unpredictable and generated more harm than good among the Muslim intelligentsia of the time. But it was because of the effectiveness of the Ashʿarite school that al-Jurjānī felt a pressing need for a more effective alternative to the Muʿtazilite linguistic and theological views concerning the *ʾiʿjāz* of the *Qurʾān*. In opposition to the Muʿtazilite doctrine, al-Jurjānī, unlike most of his predecessors and contemporaries, argued that *maʿnā* is what determines the qualities of the text, not the *lafḏ*. On this, Versteegh states: "The originality of al-Jurjānī as a rhetorician [textlinguist] is that he linked his view on meaning as the determining factor in the quality of a text to a linguistic dimension by considering it not in isolation but always as it is realized within a coherent text. Composition or cohesion (*naḏm*) is the key notion [...]" (1997:119).

In order to look at texts from the true source of excellence, i.e. meaning (semantics and pragmatics), al-Jurjānī postulated the need to go beyond the level of the individual word. Words are given meaning only in a context, and only when the context is properly ordered (*naḏm*) can we speak of texts that display full features of textness (text features and user features). For al-Jurjānī, ordering refers to the correspondence between the meanings in the 'mind of the user' and the words of the sentences. He made the relationship between meaning and words conditional on the proper application of the rules of grammar. Word order reflects the hierarchy of the elements in the mind. This process stresses the vital relationship between user and use, between text features and user features.

Although formal microelements of texts are important, the unity of texts depends more on macroelements, i.e. semantic and pragmatic relations (user features), which are mostly processed outside the internal linguistic web of the text. These relations are linked to the textual map of the text intended by the producer and expected by the receiver. Paraphrasing al-Jurjānī's language, the *kalām* (text) of a producer is activated by the 'heart' (intentions), not in order to produce lexical items that are not interrelated, but rather to establish such relations which produce stretches of language that exhibit all standards and features of textuality. Here, the linguistic structure is a reflection of the structure of meanings and ideas, which are initiated, cultivated, and fine-tuned internally based on some macrostructures that are not seen but rather gleaned from the actual words of the text. Words (*ʾalfāḏ*) exist as independent entities originally by themselves and are given collective meaning through interweaving in texts.

The implication is that neither grammatical nor lexical meanings can be thought of in and by themselves without the help of other higher meanings (user features). They are dependent on these higher meanings assigned by producers and inferred by receivers along a particular *naḏm*. Every text represents a unique unity of inputs that deal with the questions: who, to whom, when, where, why, and how. Addressing these questions implies an intrinsic relationship between lexical, grammatical, and semantic and pragmatic meanings. Thus, language is a system that mediates in a complex fashion between the universe of producers of texts and their respective receivers, based on specific purposes associated with use in context. For al-Jurjānī, a text producer conceives a *mafhūm* 'conceptual understanding' out of a combination of grammatically and semantically different lexical items. The construction or *naḏm* of this *kalām* (the text) is the realization of thought manifested linguistically, but formulated as a whole pattern psychologically.

For al-Jurjānī, it is inconceivable for thought processes to operate on the lexical meanings of words in isolation from their occurrence in particular syntactic relations. It is unreasonable to assume, for instance, that a text producer thinks of the meaning of a verb, without intending to relate it to a noun. Accordingly, text production (*naḏm* of *kalām*) is like the melting of a number of pieces of gold or silver until they become one piece. Reception of texts follows a similar process, but with the aim of undoing the melting, so to speak.

From the above discussion, it transpires that in the Arabic linguistic tradition, al-Jurjānī's textlinguistic approach constituted a major conceptual shift by giving centrality to semantics in language analysis. The producer is the one who creates the meaning (semantic structure), and the meaning determines the form. His view that the act of creating meaning is a complex mental process implies that text producers formulate structures congruous with their intentions. This congruity between the intended meaning and the mental structure is the central theme of his theory of *naḏm* and the investigation of texts. The theory centers on the principle of *taʿlīl* 'dependency' between → *lafḏ* and → *maʿnā*, providing explanatory and analytical tools for text production and reception at both macro and micro levels (Suleiman 1999). Consequently, the receiver's role is to recover the producer's utterances.

In general, Arabic linguists have examined the standards of textuality both with respect to text features and to user features. Text features (mostly cohesion and to a lesser extent coherence) have been studied most; user features, however, have not received the same attention. User features are not sufficient in and of themselves to define textual occurrences as they are not necessarily recoverable from the analysis of text features. User features are functions of the communicative interaction of the participants and form aspects of the choices made by text users in a particular context of use. Nonetheless, in an interactive, procedural approach to text, like that of al-Jurjānī's, user features link users (producers and receivers alike) with the physical text, i.e. with text features. Broadly speaking, user features include intentionality, acceptability, informativity, situationality, and intertextuality (cf. Beaugrande and Dressler 1981; Faiq 2006).

Intentionality subsumes all notions of producer intentions. This is a function of text management strategies that guide the process of choice making, which is contextualized in terms of the *naḏm* of texts and the situation of use. Acceptability is viewed as the flip side of producer intentionality. It is the perception of the receiver that the text is a coherent and cohesive communicative event. Acceptability is a rather complex issue, involving reader schemata, which al-Jurjānī relates to the psychology of texts in use.

Informativity refers to the extent to which information is new or unexpected for the text receiver. It is usually applied to content, but is processable through the internal make-up of texts (→ coherence; → cohesion). Situationality refers to factors that render a text relevant to a current or recoverable situation of occurrence. It refers to norms of mediation, monitoring, and managing strategies which text users adopt when they negotiate meanings in texts. This is subsumed by al-Jurjānī under the relationship between lexical, semantic, and pragmatic meanings and the outputs (texts) of such a relationship. Intertextuality refers to the ways in which the production and reception of a text depend on knowledge of other texts. Producers and receivers of texts use their past experience with other texts to fashion meanings.

Unlike the originality of its predecessor, modern Arabic textlinguistics has mostly preoccupied itself with adapting Western theories for the purpose of analyzing Arabic textual matter. In doing so, modern Arab scholars have bypassed their own ancestors' genius. Still, modern Arabic textlinguistics has made progress through an informed recourse to tradition, assessing, reassessing, re-editing, and rediscovering medieval Arabic texts to make them bear on the needs and the trials and tribulations of modern discourse (often competing discourses) in Arabic. Al-Musawi (2001:33) sums up this situation as follows:

> To challenge modernism, perhaps, or to face up to state manipulation of oratory, scholars have carried out extensive studies of rhetoric in its scholastic, literary, and philological dimensions, as if to counteract both fusion into mass culture and state or political polemics of expediency and urgency. Between oratory for mass communication and rhetoric for the established cultured groups, many bygone arguments regain their validity and relevance.

It is within such a context that the vitality and relevance of al-Jurjānī's views on and contributions to the development of Arabic textlinguistics lie: logic, sociology, psychology, and esthetics woven into a theory of text.

Bibliographical references

Primary sources

Jurjānī, *ʾAsrār* = ʾAbū Bakr ʿAbd al-Qāhir ibn ʿAbd ar-Raḥmān al-Jurjānī, *ʾAsrār al-balāġa*. Ed. Helmut Ritter. Istanbul: Wizārat al-Maʿārif, 1953. (2nd ed., Baghdad: Maktabat al-Muṯannā, 1989.)

——, *Dalāʾil* = ʾAbū Bakr ʿAbd al-Qāhir ibn ʿAbd ar-Raḥmān al-Jurjānī, *Dalāʾil al-ʾiʿjāz*. Ed. Yāsīn al-ʾAyyūbī. Beirut: al-Maktaba al-ʿAṣriyya, 2000.

SECONDARY SOURCES

Al-Musawi, Muhsin. 2001. “Arabic rhetoric”. *Encyclopedia of rhetoric*, ed. Thomas O. Sloane, 29–33. Oxford: Oxford University Press.

Beaugrande, Robert de and Wolfgang Dressler. 1981. *Introduction to textlinguistics*. London: Longman.

Faḍl, Ṣalāḥ. 1996. *Balāġat al-xiṭāb wa-ʿilm an-naṣṣ*. Cairo: Longman Egypt.

Hatim, Basil. 1997. *Translation theory and contrastive textlinguistics*. Exeter: University of Exeter Press.

Khatabi, Mohammed [Muḥammad Xaṭābī]. 1991. *Lisāniyyāt an-naṣṣ*. Beirut: al-Markaz aṯ-Ṯaqāfī al-ʿArabī.

Maslouh, Saad [Saʿd Maslūḥ]. 1990. *al-ʿArabiyya min naḥw ‘al-jumla’ ʾilā naḥw ‘an-naṣṣ’*. Kuwait: Kuwait University.

Moutaouakil, Ahmed. 1990. “La notion d’actes de langage dans la pensée linguistique arabe ancienne”. *Studies in the history of Arabic grammar*, II, ed. Michael G. Carter and Kees Versteegh, 229–237. Amsterdam and Philadelphia: J. Benjamins.

Suleiman, Yasir. 1999. *The Arabic grammatical tradition: A study in* taʿlīl. Edinburgh: Edinburgh University Press.

Versteegh, Kees. 1997. *Landmarks in linguistic thought*. III. *The Arabic linguistic tradition*. London: Routledge.

SAID FAIQ (American University of Sharjah)

Text Processing → Automatic Language Processing

Thai

The impact of Arabic on the Thai language is modest but raises issues of phonological, sociolinguistic, and historical interest. Arabic loanwords of two general types can be distinguished. One type consists of items in general Thai usage, mainly the result of early trade contact. The second type includes lexical items relating specifically to Islamic institutions and culture. The latter vocabulary would naturally be especially familiar to Thailand’s Muslim population, about 3 percent of the total.

In the Thai words transcribed here, colon and circumflex accent denote long vowels. Thai tones are marked by superscript numerals: 1 for low tone; 2, falling; 3, high; 4, rising; unmarked, mid. High and low back unrounded vowels are represented respectively by *u’* and *o’*. The transcription indicates modern spoken Thai, in which final /r/ and /l/ in borrowed items are assimilated to final /n/ and final /s/ to /t/, although the written language may retain etymologically conservative spelling.

Arab and Persian traders were established in Southeast Asia more than a millennium ago. The Persian term *bazaar* was known in Bali and elsewhere in the archipelago well before a cognate occurred about seven hundred years ago in the first Thai written sources. Although Islam was not to become the dominant religion of the Thai court, the king regularly appointed Muslim officials to oversee maritime trade and associated taxation. One Persian family in particular became trusted royal counselors. For several centuries, senior members of this dynastic group, the Bunnags, were admitted into the highest ranks of the Thai nobility, participated in state councils, and even had influence in determining royal succession.

Descendants of these traders and government officials have tended to be bicultural, many speaking Thai fluently and integrating into Thai society, while retaining their ancestral faith. For vocabulary referring to trade items, this group would have been pivotal in consolidating Arabic loans in common Thai usage.

The context of pluralistic language contact in which trade was conducted raises a methodological issue for diachronic linguistic analysis relevant throughout Southeast Asia. Some Arabic loans in Thai would have been introduced directly by Arabic speakers, while others would be indirect loans acquired by way of Persian, Malay, Javanese, Khmer, and other languages. Even English may be implicated, as in Arabic *safar* ‘journey’, passing through Swahili and English into Thai as *safari:* ‘safari (suit)’.

A more remote ultimate source is yet another possibility. Thus, Roman contact with Germanic in early empire times is held to account for the Latin item *sâpo*, gen. *sâpônis* ‘soap’. This should be kept in mind when considering Arabic *ṣābūn*, Thai *sabu:*^{1}, Portuguese *sabão*, Malay *sabun*, etc. Thai *krada:t*^{1} ‘paper’ invites comparison with Arabic *qirṭās* ‘chart’ as well as with Spanish and Portuguese forms. Thai *rian*^{4} ‘coin’ (< Arabic *riyāl*?), *sak*^{1}*kala:t*^{1} ‘flannel’ (but not necessarily ‘scarlet’ in color, < Persian *saqerlāt*), and *farang*^{1} ‘Westerner’ (not restricted to French or Frank, < Arabic *ʾifranj*

'Franks, Europeans') raise similar problems. Tracing exact provenance for many items such as these may not be possible and in any case is beyond the present scope.

Similarly, early Indian Ocean contact between speakers of Arabic and Sanskrit-related languages needs to be kept in mind: Arabic *nīl* 'indigo', Vedic Sanskrit *nīla* 'indigo color, dark blue' may be compared with Thai *nin* 'dark blue; precious dark stone'; Arabic *kāfūr* 'camphor', Sanskrit *karpûram* with Thai *ka:rabu:n*.

Aromatics and spices would have been traded in the early Thai bazaars. Thus Arabic *kammūn* 'cumin', *zaʿfarān* 'saffron' gave rise, directly or indirectly, to Thai *khamin*^{2} and *ya:*^{2}*-fran*^{1}. The latter item illustrates partial reanalysis, as the Thai word *ya:*^{2} refers generically to grasses and herbs. In accordance with Thai syntactic principles, *ya:*^{2} forms a number of [head + modifier] compounds, e.g. *ya:*^{2}*-kha:* 'thatch grass'. The Arabic item has been assimilated to this semantic set.

Common in Thai daily usage is *kalam*^{1}*-pli:* 'cabbage', i.e. *Brassica oleracea* (Cruciferae). This form is undoubtedly cognate with Arabic *kurumb, karumb* 'cabbage, cauliflower' and shows the change /r/ > /l/, attested in Thai more widely. The second component *pli:* in the Thai compound refers to a bulb-like outgrowth and is perhaps facilitated by the final labial in the Arabic form. In Thai, 'cauliflower' is distinguished from 'cabbage' by change of suffixal *pli:* to *do':k*^{1} 'flower': *kalam*^{1}*-do':k*^{1}.

Exotic fabrics supplied to the Thai court by Arab traders included satin: Arabic *ʾaṭlas*, Thai *at*^{1}*talat*^{1}. This item has undergone semantic specialization, as the Thai version implies interwoven threads of gold and silver. Gold embroidery is also specified for the Thai cloth *khem*^{2}*-kha:p*^{1} lit. 'dark + blue', formerly used for officials' coats; this is perhaps a folk etymology based on Arabic *kamḫā* 'damask'.

A luxury item for perfumery would have been Arabic *ʿambar* 'ambergris', Thai *amphan*. The substitution /b/ > /ph/ relates to a regular Thai devoicing with aspiration of the 15–16th centuries, indicating a date for this loan prior to contact with Western languages.

Apart from trade, another route of entry for Arabic loans into Thai was through literature, especially the Javanese Panji tales, which were the basis for the 18th-century Thai literary classic called *Inao*, along with translations of the *Arabian nights*. Thus it is through → Javanese that Arabic *ʿaraq* 'juice' appears in Thai as *a: ra*^{3} (with glottal final), implying a fermented drink (*arrack*). Arabic *xanjar* 'dagger' probably entered Thai through a similar route as *kan*^{2}*yan*^{1}. Such literary items may no longer be widely understood.

Distinct from the group of trader and court heritage mentioned above, a numerically larger Thai Muslim subgroup is centered in the southernmost Thai provinces bordering Malaysia. This area was formerly the Malay-speaking port state of Pattani, Islamic, and for several hundred years in vassal relations with the Thai kingdom prior to incorporation in the 19th century. Here, Malay is still written in an Arabic-derived script referred to as Jawi – a practice once normal throughout the Malay world. A local variety of Malay is still spoken in the Pattani area as first language, with varying degrees of Thai bilingualism.

It is mainly in this context that a second category of Arabic loans is encountered, words applying specifically to Islamic institutions and culture. Included are many vocabulary items such as *ima:m*^{1}*, kur'a:n*^{1}, etc. Forms are phonologically assimilated. For bisyllabic loans, Thai low tone is commonly assigned to the second syllable, as illustrated also in the ethnonym *a:rap*^{1} 'Arab' itself. Such items may be encountered in Thai mass media, but unlike Arabic loans of the first category, not all of those in the second would occur in standard Thai dictionaries.

Anthony Diller
(Australian Academy of the Humanities) and
Wilaiwan Khanittanan
(Thammasat University)

Thamudic

The term 'Thamudic' is used generically for a number of pre-Islamic Ancient Arabic scripts and languages epigraphically attested. This entry covers the same ground as → Old Arabic (epigraphic), but from a different perspective.

1. Terminology and taxonomy

The term 'Thamudic' covers those texts in Ancient Arabic scripts and languages which are not parceled out as Hasaitic, Dedanite/

Lihyanite, or Safaitic. On the same taxonomic level of these three terms are the designations Taymanite, Najdī, Ḥijāzī, or Hismaic 'Thamudic' (Macdonald 2000, with a slightly different terminology). From the point of view of historical linguistics, these seven languages constitute the corpus of Ancient Arabic. They need to be regarded as an ensemble and always in contrast with Old Arabic, genetically the successor of Ancient Arabic but attested in the same area at the same time. The scripts used for this family of languages derive from South Arabian script(s). Based on the script, the group is frequently called 'Ancient North Arabian'; but whereas South Arabian languages, the languages indigenous to South Arabia, are not a form of Arabic, these northern languages genetically are; hence, the term 'Ancient Arabic' is preferred. The geographical distribution of these languages is expressed by their designations. The earliest specimens can be dated to the 6th century B.C.E. (Tayma; Hayajneh 2001), the very latest to the 5th century C.E. (Khirbet es-Samrāʾ; Knauf 1998). After the 3rd century C.E., Ancient Arabic, with the definite article *han-*, disappeared quickly, and Old Arabic, with the definite article *ʾal-*, became Early (Standard) Arabic, after at least eight hundred years of coexistence with Ancient Arabic as a spoken and, very rarely, written language (Knauf, forthcoming). The other basic difference between Ancient Arabic and Old Arabic is graphic. The Ancient Arabic scripts derive from Sabaic and Minaean, i.e. from the South Semitic alphabet. The tribes and towns of North Arabia became acquainted with it after the establishment of a pan-Arabian trade network with trading colonies of South Arabians at Tayma, Dedan, and other places, in the course of the 8th century B.C.E. Old Arabic became Early Standard Arabic from the 4th century C.E. onward by adopting (or rather, by continuing) the Nabataean Aramaic script and employing the rules for the rendition of Arabic names in Aramaic writing, established between the 9th and the 6th centuries B.C.E., on the language at large.

The tribe of Ṯamūd is attested for the Ḥijāz from Sargon II (716 B.C.E.) to the *Notitia Dignitatum* (late 5th century C.E.). In the 2nd/3rd centuries C.E., the tribe had a temple at Ruwāfa with inscriptions in Nabataean Aramaic and Greek. In the *Qurʾān*, they have become a paradigmatic 'lost tribe' (Q. 7/73–79, 11/61–68, 26/141–159, 27/45–53, 51/43–45, 54/23–31, 69/4–5, 89/9, 91/11–15), supposed to have left the [Nabataean] rock architecture at Hegra/el-Ḥijr/Madāʾin Ṣāliḥ – where Salamaeans lived in the Hellenistic and Roman periods.

2. The definite article and the genetic position of Ancient Arabic

Ancient Arabic shares the following features with Old Arabic (and later stages of the language):

i. Preservation of the whole range of consonantal phonemes known from Epigraphic South Arabian (the richest yet attested in the Semitic world), with the exception of /s_3/ (originally /t^s/; Aramaic /s/, Arabic /s/).
ii. The prepositive article *han-* as in Canaanite (in opposition to the postpositive article of Epigraphic South Arabian and Aramaic).
iii. The productive broken plural, as in → South Semitic (in opposition to the other Semitic languages, in which the internal plural fossilized).
iv. On the basis of any evidence for the opposite being absent, one may assume that Ancient Arabic had the *yaqtul/yaqtulu*-imperfects of Central Semitic (in opposition to the *yaqtul/yaqattul*-imperfects of the other branches).

From this list of features, one can already gain the impression that Ancient Arabic is the genetic ancestor of Old Arabic. This impression is substantiated by the development of the definite article: (i) *han-* > *ha*C_1-; this is the form in which the Ancient Arabic article is usually attested; only in Hasaitic and Lihyanite, the *n-* is preserved in front of laryngeals and pharyngeals (*han-ʾilāt* 'the Goddess', but *hag-gamal* 'the camel'); (ii) *ha*C_1- > *ʾa*C_1- (as in the prefix of the causative stem; Voigt 1987); (iii) *ʾa*C_1 > *ʾal-* (by dissimilation of the geminate; Wensinck 1931; Tropper 2001).

The article *han-* is shared by Epigraphic → Ḥimyaritic, which has otherwise some South Arabian features and is still little understood. It might be a very archaic Ancient Arabic language, or a language of either Arabic or South

Arabian origin, which adopted forms from its neighbor.

3. PHONOLOGY

The default assumption is the preservation of the Proto-Semitic realization of the consonantal phonemes unless there is evidence for the opposite; for the reconstruction of the phonemes of Proto-Semitic and their distinctive features, and notably the sibilants, see Voigt (1979, 1981, 1998) and Sima (2001, 2004).

Ancient Arabic and Old Arabic had the sibilant s_1 /s/ and the lateral s_2 /ś/ [~ /^{t}ɬ/]; s_3 /t^{s}/ had merged with /s/. The reconstruction of Ancient and Old Arabic /s/ as /š/ by A.F.L. Beeston and others is specious, based on the Aramaic letter <š> employed for Arabic /s/ in Aramaic transcriptions of Arabic names from Imperial Aramaic to Middle Aramaic. This feature is due to Ancient West Aramaic phonology, which had s_1 /s/ (Greek *sigma*) and s_3 /ts/ (Greek *ksi*) when the Aramaic script was adopted by Greeks – and Arabs (Tropper 2000).

The evidence for word-final short vowels (and hence for case and mood/tense) is ambiguous. In Old Arabic, they sometimes do and sometimes do not appear; the proper name *Taym* is *tymw* in Nabataean (**Taymu*; the quantity of the final vowel cannot be deduced from the notation; cf. Knauf 1986), but **Tēm-* in Greek *temallatos*, **Tēm(ʾ)allāt*. The strictly consonantal orthography of Ancient Arabic (with few exceptions) is unable to indicate the presence of final short vowels; the exceptions like Safaitic *h-nbṭy* are ambiguous: *han-nabaṭī* or *han-nabaṭīyu*? On the other hand, there is *hṯmd* /haṯ-Ṯamūdī/ in Najdī ('Thamudic B'; Winnett 1937:33).

Sandhi writings are indicative of instances of lost final vowels: *lPN]bnkʿmhxṭṭ* (Ababneh 2005: no. 245) *li*-PN *bin Ka-ʿammih hax-xuṭūṭ* 'These drawings are by PN bin Kaʿammih(ī)'. Case is attested for Nabataean Old Arabic (*tymw* /Taymu/ vs. *tymʾlhy* /Taymullāhi/; in more than 95 percent of Nabataean inscriptions and graffiti, case is correctly noted), while its absence is evidenced by South Safaitic *ḏyl* /ḏīyāl/ for **ḏū ʾāl* 'the one of the tribe of TN' (Clark 1980:126, no. 3) and *hylt* for /hā ʾilāt/ 'O Goddess'. Since the case/final (short) vowel is attested for Nabataean Old Arabic, it can hardly have disappeared from Ancient Arabic completely. One might postulate the dichotomy of 'high Arabic' (with final vowels) and 'vernacular Arabic' (without final vowels) already for the earliest stages of the language. There is no evidence for → nunation except, perhaps, the habit of not writing the /n/ of *bin* 'son' in Taymanite and Najdī. **Bin Ḥaggāg* is written as *bḥgg*. Since assimilation of the *nūn* is quite unlikely in this case, *b* for *bin* was an acceptable orthography at Tayma if *ḥgg* represented *ḥaggāgin*; otherwise, there was nasalization of the /i/ vowel. In any case, *b* cannot stand for **binu*.

The distinctive features of some consonantal phonemes can partially be reconstructed from 'phunny orthografy'. From Epigraphic South Arabic the 'etymology' of the Ancient Arabic graphemes is known; sometimes it clashes with the grapheme's phonemic value. In the following list /g, ṯ/ is short for 'In Hismaic (or South Safaitic), the grapheme for /g/ is derived from the traditional sign for /ṯ/'.

In Taymanite and Dedanite/Lihyanite, we find /ṯ, s$_3$/. First, this is a good reason not to look for a phoneme /s$_3$/ in Taymanite any longer (some cases have been proposed, but not convincingly). Second, the distinctive feature [+interdental] seems to be augmented, if not replaced, by [+affricate]; /ṯ/, then, was /tṯ/ or /ts/. The latter was definitely the case in Lihyanite, where *man sāqiruhū* 'whoever wants to rob it' is written *mn ṯqrh* (Jaussen and Sauvignac 1909–1922, no. 81:6). The sound change /ns/ > /n^{t}s/ is trivial. Thus, /s$_3$/ had disappeared from Taymanite and Lihyanite (and from the rest of Ancient Arabic), and /s$_1$/ had maintained its Proto-Semitic realization (as in Arabic and Ethiopic, in opposition to Hebrew and Babylonian). If /ṯ/ was /ts/, /ḏ/ probably was /dz/.

In Ḥijāzī, we have /s$_2$, ḏ/. So /s$_2$/ still was a voiceless lateral fricative (or affricate), written by a sign derived from the grapheme for its voiced counterpart, either because /s$_2$/ had developed into something else in the Arabic or Arabian language from which the Ḥijāzīs received their peculiar script – which is unlikely, given that the lateral character of Arabic /ś/ was preserved well into the Umayyad period, or because the <n> had adopted a form hardly distinguishable from <s$_2$>.

In Hismaic/South Safaitic (or Tabuki Thamudic, ‘Thamudic E’), there are /g, ṯ/, /ṯ, ḍ/, and /ḍ, ʿ/ (King 1989). The last item indicates that /ḍ/ was still voiced, fricative, and lateral. The new sign became necessary because the traditional sign was used for /ṯ/. There seems to be only one feature that /ṯ/ and /ḍ/ have in common: [+ fricative]. So, Hismaic had /ṯ/ and /ḏ/ rather than /ts/ and /dz/; or, again, the sign change is of no significance, it simply happened because <ṯ> was needed for something else – for /g/. In the South Semitic alphabet, <ṯ> is a graphic variation of <y> (or vice versa). If the phoneme is noted as /g, y_2/, the phonetic development becomes obvious: /g/ was palatalized – the whole range of /g^y/, /j/, /ž/, /y/ is possible (Voigt 1984).

In Safaitic, /ḍ/ and /ḏ̣/ coalesced, at least sometimes (Ababneh 2005:27).

Assimilation of *n* to a following consonant is frequent and not restricted to the definite article: *ʾanā muḥibb ʾattā lafēta*/*ʾattī lafēti* (Littmann 1940:72, no. 113) ‘I am in love, you did wrong’ (or *nafēta* ‘you refused’); Safaitic *ʾfs* < /ʾanfus/ ‘persons’, *tḏ̣r* < /ntaḏ̣ar/ ‘he watched’. The diphthongs *aw* and *yi* are, as a rule, contracted, as also evidenced by Greek writings of pre-Islamic personal names, where, in accordance with (late) Aramaic phonology, /u/ is frequently rendered by <o>, and /i/ by <e>. The *-t* of the feminine singular noun in the absolute state occasionally becomes *-h*. How much of correct spelling is etymological in nature cannot even be estimated.

4. MORPHOLOGY

In verbs, radical *w* and *y* are usually written in final positions, very rarely in medial position, and in nouns usually not at all: *banay(a)* ‘he built’ vs. *ʾil(ā)* ‘to’. Again, many spellings seem to be etymological rather than phonemic or phonetic. The god/goddess Ruḍā (Esarhaddon *Ruldāyu*; Herodotus *orotalt* < /orolta:/; Ibn al-Kalbī, *ʾAṣnām: Ruḍā* and *Ruḍāʾu*) appears as *rḍw, rḍy, rḍʾ, rḍ, rḍh, ʾrṣw* in Palmyra and as *arsa-* on the Roman coins of the city of Areopolis in Moab (Knauf 1984).

The 1st person masculine singular of the perfect ends in *-t* (as opposed to *-k* in Ḥimyaritic): *wadadtu fa-/wa-katamtu* ‘I loved [+ PN] and hid’ (Littmann 1940:68; Littman prefers ‘I greeted’, which provides no good reason for hiding). In Safaitic, Forms IV, V, and VIII are visible (*ʾtrk* ‘he left behind’; *tśwq* ‘he longed for’; *tḏ̣r* ‘he watched’); the existence of Form II and III can be semantically inferred.

The dual exists at least in Lihyanite (pronominal suffix **-humay*, verb *-ā* [written *-h*], but usually replaced by the plural here), and Safaitic (*hbkrtn* /hab-bikratān/ ‘the two young she-camels’; *gmln* /gamalān/ ‘two camels’).

Winnett (1937:22) interpreted Taymanite (‘Thamudic A’) *smw* and *ʾtw* as ‘his name’ and ‘his mark’, but one might just as easily read /samaw/ ‘he is/was elated’ and /ʾataw/ ‘he came’.

The external (‘sound’) plural of the noun *-n* is rare (female: *dṣyt* /daṣyāt/ ‘antelopes’); internal (‘broken’) plurals are abundant: /gimāl/ ‘camels’; /ʾaffus/ ‘persons’; /ʾaśnāʾ/ ‘enemies’; /ʾanūq/ ‘she-camels’.

Demonstratives are *ḏā* (Lihyanite), masculine *ḏān*, feminine *ḏīn* (Thamudic; vocalization as in Maltese); in many cases, the definite article serves as a demonstrative.

5. SYNTAX, STYLISTICS, LEXICON

Texts from the Oasis towns include tombstones, especially in Hasaitic (starting with /nafs wa-qabr/ ‘stela and tomb of…’ or /wagr wa-qabr/ ‘tombstone and tomb of…’, and legal texts, claiming the fulfillment of a civic duty. At Tayma, for instance, loyalty to the community is expressed as loyalty to Ṣalm, the main deity of the city (/PN naṣar[a] li-Ṣalm bi-ḍarr TN/ ‘PN rendered assistance to Ṣalm in the war against the tribe TN’; Winnett and Reed 1970:100–102). At Dedan, landowners documented their incessant efforts to keep the irrigation system working: A *bin/bint* B *ʾaggaw* [*n-g-w*, causative]/*ʾaḏ̣all*/*ʾaḏ̣lal*/*ʾōfay*/*haḏ̣-ḏ̣ilāl li-ḏū ġābat fa-raḍā-hū* ‘A son of B has cleared/roofed/performed maintenance for the subsurface water channel for Ḏū Ġābat (main deity of Lihyan), so may He show mercy upon him’ (with variants and optional additons, Sima 1999).

Graffiti are the literary heritage of the Bedouin before the rise of classical poetry. They all share the basic function of marking someone’s presence. Safaitic and South Safaitic/Hismaic have the richest template, in its full form tripartite like the Lihyanite ‘corvée-texts’ (*superscriptio – narratio – invocatio*). The name of the author, preceded by *l-* (Ia), is compulsory (Safaitic, South Safaitic/Hismaic, Taymanite) or a demonstrative like *ḏn* (Ḥijāzī?), optionally followed by his genealogy (Ib) and/or his tribal affiliation (Ic), then by a narration commencing with *w-* (II) of his past or present doings (also in the perfect, addressing posteriority) or state of mind, and finally an invocation, starting with *f-* (III) of a deity or several of them, asking for blessings or the protection of the text (*ʿawwir/ʿawwirī ḏī yuʿawwir* ‘blind the person who effaces [this inscription]’). The standard form can be rendered as Ia [+Ib] [+Ic] [+II$_{1\ldots n}$] [+III$_{1\ldots n}$] (for a more detailed formula, see Voigt 1980). In the narrative, expressions of personal loss and longing (*nagaʿ ʿal[ā], taśawwaq ʾil[ā]* ‘he longed for’) are frequent, which will later form the *incipit* of the *qaṣīda* (Littmann 1940:114; Petráček 1973).

The elements of the Safaitic ‘memorial graffito’ had been in use for a long time, and they are individually attested in several forms:

i. Isolated names, with or without affiliation, and with or without the *l- auctoris*. In Taymanite and Najdī, *l-* co-occurs with the variants *lm* (probably **limā*) and *nm* (probably by dissimilation from the latter). In Ḥijāzī texts, the preposition may be replaced by the personal pronoun *ʾanā* or the demonstrative *ḏn*, which has a funeral connotation in at least one instance (see below; Jaussen and Sauvignac 1909–1922, no. 17).

ii. Claims of possession: *l-*PN *h-*[animal(s)/place]; when other than domestic animals are claimed, the formula is a rock artist’s signature.

iii. Isolated narratives (without *superscriptio*; those with it represent a short form of the tripartite scheme), notably for the expression of love: PN$_1$ *wdd f-*PN$_2$ (Ḥijāzī; Winnett 1937:35; Littmann 1940:64–72) is best analyzed as /N1 wādid fī N2/ ‘PN$_1$ is in love with PN$_2$’, by *wdd/wd*, or its opposite PN$_1$ *nāk[a]* PN$_2$ (also frequent in Ḥijāzī). A variant of ‘love messages’ follows the → presentative as an apposition or relative clause: *ḏn* PN$_1$ *mḥb / ʿiśq/ḏyʿśq* PN$_2$ (Littmann 1940:73–74).

iv. Prayers, notably in Najdī, starting with *hā* + DN (Nuhāy, Ruḍā, ʿAttarsamā) or *bi-*DN, with *nm* as direct-object marker. Examples: *h-rḍw w-nhy w-ʿtrsm sʿdn ʿl wddy* /hā Ruḍāw wa-Nuhāy wa-ʿAttarsamā, sāʿidnī ʿal(ā) widādī/ ‘O Ruḍā, Nuhā, and ʿAttarsamā, help me because of my love’ (Winnett and Reed 1970:80, no. 23); (1) *h-nhy sʿdn ʾlh ṯʿt* (2) *b-nhy tʿḏy nm whbnhy* (3) *bk h-srr śms* (4) *mtʿly* (5) *h-rḍw nqm whbnhy* ‘(1) O Nuhāy, help me, God of salvation (/ṯiat/?, /ṯūʿat/?) (2) By Nuhāy is the health of Wahab-Nuhāy (3) By you is the joy, rising (/mutaʿālī/) sun (5) O Ruḍā, avenge Wabah-Nuhāy’ (Winnett and Reed 1973:88(e)). At Tayma, some inscriptions starting with the usual *superscriptio* end in a confession: *mn smʿ l-ṣlm l* (/lā/) *twy* ‘who listens/is obedient to Ṣalm will not perish’. A prayer well attested in South Safaitic/Hismaic, both in isolation and integrated into the *invocatio* of the tripartite scheme, is *ḏkrt (ʾ)lt* PN ‘may Lāt commemorate PN’.

Both the graffiti and the ‘legal texts’ from the towns address a public; one cannot infer from this and the mere number of inscriptions (at present, more than 30,000) that everybody was literate, but a significant percentage of the population must have been.

The lexicon of Ancient Arabic contains some genuine Arabic words that later disappeared from the language, like *bikrat* ‘young she-camel’ (Biblical Hebrew *biḵrāh*), *ḏīrat* ‘pen, corral, enclosure’ (Biblical Hebrew *ṭīrōt*). Other words have not yet undergone later semantic developments; *naxl* still is ‘valley’, not yet ‘palm grove’ (the most common fruit tree in Arabia, commonly found in depressions where water gathers).

6. Sociolinguistics and Diglossia

The formulaic character of most of the texts indicates a kind of semiliteracy: one could write only certain types of texts, not necessarily anything that came to one’s mind. On the other hand, no scribe in the entire Ancient Near East would have thought of writing down anything that came to his mind. The maintenance of difference between the various Ancient

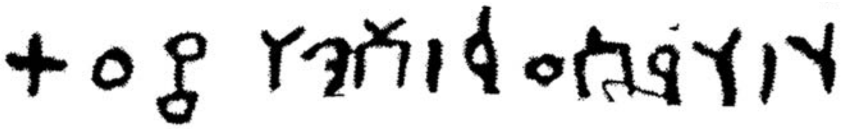


Figure 1. Pre-Hellenistic inscription from the Ḥā’il area (Najdi Thamudic), Winnett and Read 1973, No. 203e. With final vowels, the metre is *rajaz*: *hā Nuhāya sāʿidnī ʾilāha ṯūʿatī* ‘O Nuhāy, help me, God of my salvation’.

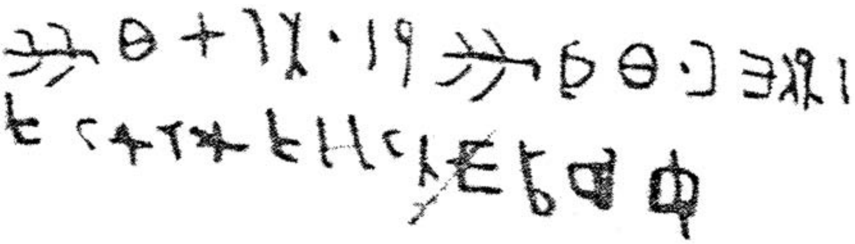


Figure 2. Hellenistic inscription (South Safaitic/Hismaic/Tabuki Thamudic), Clark 1980, No. 3: *li-Ṣāhiṭ bin Wadd ḏīyāl [< ḏī ʾāl] Nāʾilat wa-ḏakirat Lāt kalīl rāhiṭ ṣaʿiq* ‘By Ṣāhiṭ bin Wadd of the tribe Nāʾilat. May Lāt be mindful of a tired, thirstily gulping person, who is (nearly?) unconscious’.

Arabic scripts, even without phonological reasons detectable, together with the predominant ‘dropping of one’s name’ indicates the marking of territory/territoriality as the main purpose in the production of these texts. The notion is enhanced by those texts, in which *l-* does not indicate the author (who claims ownership by his mere presence) but rather the owner: *l*-PN *had-dār/hag-gimāl* ‘this place/these camels belong to PN’, sometimes accompanied by the rock drawing of the thing mentioned. In this case one might think of an artist’s signature (Ababneh 2005) but for the fact that these formulas also occur without pictures, clearly claiming the campground for the author’s next visit to the region, and the camels which one encounters in the environs.

The tribes which left these inscriptions were bilingual in several dimensions (the individuals not necessarily, and presumably not within the full range). Three kinds of Ancient Arabic diglossia are encountered:

i. Ancient Arabic and Aramaic. Several tribes left inscriptions both in Safaitic and Nabataean, e.g. Qumayr (Milik in Winnett and Reed 1970:160, no. 130).
ii. Ancient Arabic and Greek. Occasionally, a writer of Ancient Arabic could also write his name in Greek; Wahballāh ibn Ḏann’il ibn Wahballāh left a Safaitic inscription of twenty-three words (Winnett and Harding 1978, no. 1849), but also four words in Greek (Winnett a.o. 1980, no. 1860): *ouaballas tannhlou tou {u}ouaballas*.
iii. Ancient Arabic and Old Arabic. It is remarkable that Wahballāh, writing in Greek, uses the Old Arabic form of his name; in Ancient Arabic, it should be *Wahb-ha(n)-ʾIlāh (not attested at all) or Wahb(’)ilāh. The custom to switch from the Ancient Arabic article to the Old Arabic form when addressing regional or tribal ‘outsiders’ can be traced back to the 2nd century B.C.E., when a Gerrhaean, i.e. someone from al-Ḥasā, appears as *thmallatos gerraios* on Delos (Knauf, forthcoming); at home, he would have signed as *Tēm-han-ʾIlāt (Garm-han-ʾIlāt is attested).

One gets the impression that, at least among the trading communities of Arabia, (Nabataean) spoken Old Arabic had become a lingua franca well before the 4th century C.E., which would provide an elegant explanation of why it became the basis for written Early Standard Arabic. One of the latest ‘Thamudic’ inscriptions, and one of the few firmly dated, summarizes the late Nabataean tomb inscription (Jaussen and Sauvignac 1909–1922, no. 17) of

267 C.E. (Healey and Smith 1989): *ḏīn raqōš bint ʿabdmanōt* – five words; the Nabataean text has thirty-six words, of which six are Aramaic and twelve (Old or Early Standard) Arabic, another thirteen are equally intelligible in both languages, and five are proper names.

In the 3rd century C.E., the people of Hegra were more literate in Nabataean than in Ḥijāzī, and spoke Old Arabic to such an extent that they began to forget their Aramaic and were about to use the same script for Early Standard Arabic. This inscription marks the end of an epoch, in which Aramaic and Ancient Arabic coexisted as written languages with spoken Old Arabic. The Arabic diglossia of today is quite a simplification compared to the linguistic situation in Arabia before the rise of Islam.

BIBLIOGRAPHICAL REFERENCES

For a more complete survey of collections of inscriptions, → Old Arabic (epigraphic).

Ababneh, Mohammad I. 2005. *Neue safaitische Inschriften und deren bildliche Darstellungen*. Aachen: Shaker.

Clark, Vincent A. 1980. “Three Safaitic stones from Jordan”. *Annual of the Department of Antiquities of Jordan* 24.125–128.

Hayajneh, Hani. 2001. “First evidence of Nabonidus in the Ancient North Arabian inscriptions from the region of Tayma”. *Proceedings of the Seminar for Arabian Studies* 31.81–95.

Healey, John F. and G. Rex Smith. 1989. “Jaussen-Savignac 17: The earliest dated Arabic document (A.D. 267)”. *Atlal* 12.77–84.

Jaussen, Antonin and Raphaël Savignac. 1909–1922. *Mission archéologique en Arabie*. 5 vols. Paris: Ernest Leroux and Paul Geuthner. (Repr., Cairo: Institut Français d'Archéologie Orientale, 1997).

King, Geraldine. 1989. “Some inscriptions from Wadi Matakh”. *Arabian studies in honour of Mahmoud Ghul*, ed. Moawiyah M. Ibrahim, 37–55. Wiesbaden: O. Harrassowitz.

Knauf, Ernst Axel. 1984. “Arsapolis: Eine epigraphische Bemerkung”. *Studii Biblici Franciscani Liber Annuus* 34.353–356.

——. 1986. “Bemerkungen zur frühen Geschichte der arabischen Orthographie. II. Sprachen die vorklassischen Araber in Pausa **ʿAmru:?*”. *Orientalia* 55.452–453.

——. 1998. “Die altnordarabischen Inschriften”. *Fouilles de Khirbet es-Samra en Jordanie*, I, ed. Jean-Baptiste Humbert and Alain Desreumaux, 523–539. Turnhout: Brepols.

——. Forthcoming. “Arabo-Aramaic and ʿArabiyyah: From Ancient Arabic to Early Standard Arabic, 200–600 C.E.”. *The* Qurʾān *in question,* ed. Angelika Neuwirth. Beirut: Deutsches Orient-Institut.

Littmann, Enno. 1940. *Thamūd und Ṣafā: Studien zur altnordarabischen Inschriftenkunde*. Leipzig: F.A. Brockhaus.

Macdonald, Michael C.A. 2000. “Reflections on the linguistic map of pre-Islamic Arabia”. *Arabian Archaeology and Epigraphy* 11.28–79.

——. 2005. “Literacy in an oral environment”. *Writing and Ancient Near Eastern society: Papers in honour of Alan R. Millard*, ed. Piotr Bienkowski, 49–118. New York and London: T and T Clark.

Petráček, Karel. 1973. “Zur semantischen Struktur der safatenischen Inschriften”. *Archív Orientální* 41.52–57, 153–164.

Sima, Alexander. 1999. *Die lihyanischen Inschriften von al-ʿUḏayb (Saudi-Arabien).* Rahden: M. Leidorf .

——. 2001. “Der Lautwandel s^3 > s^1 und s^1 > s^3 im Sabäischen”. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 151.251–262.

——. 2004. “Der Lautwandel s^3 > s^1 im Sabäischen: Die Wiedergabe fremden Gutes”. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 154.17–34.

Tropper, Josef. 2000. “Griechisches und semitisches Alphabet: Buchstabennamen und Sibilantenentsprechungen”. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 150.317–321.

——. 2001. “Die Herausbildung des bestimmten Artikels im Semitischen”. *Journal of Semitic Studies* 46.1–31.

Voigt, Rainer M. 1979. “Die Laterale im Semitischen”. *Welt des Orients* 10.93–114.

——. 1980. “On the structure of the Safaitic inscriptions”. *Annual of the Department of Antiquities of Jordan* 24.79–93.

——. 1981. “Inkompatibilitäten und Diskrepanzen in der Sprache und das erste phonologische Inkompatibilitätsgesetz des Semitischen”. *Welt des Orients* 12.136–172.

——. 1984. “On South Safaitic”. *Annual of the Department of Antiquities of Jordan* 28.311–314.

——. 1987. “Die Personalpronomina der 3. Person im Semitischen”. *Welt des Orients* 18.49–63.

——. 1998. “Der Lautwandel s^3 > s^1 und s^1 > s^3 im Altsüdarabischen”. *Le Muséon* 111.173–186.

Wensinck, Arent Jan. 1931. “The article of determination in Arabic”. *Mededeelingen der Koninklijke Akademie van Wetenschappen (Afdeeling Letterkunde)* 71A.47–64.

Winnett, Fred V. 1937. *A study of the Lihyanite and Thamudic inscriptions*. Toronto: University of Toronto.

—— and William L. Reed. 1970. *Ancient records from North Arabia*. Toronto: University of Toronto Press.

—— and William L. Reed. 1973. “An archaeological-epigraphical survey of the Ḥāʾil area of northern Saʿudi Arabia”. *Berytus* 22.53–100.

Winnett, Fred V. and Gerald L. Harding. 1978. *Inscriptions from fifty Safaitic cairns*. Toronto: University of Toronto Press.

ERNST AXEL KNAUF
(University of Berne)

Theme/Rheme

'Theme' and 'rheme' are terms of the Prague school taken over by various types of functional grammars and by pragmatic sentence analysis. They essentially refer to the two elements that make up an ordinary sentence with a binary structure: the 'theme' expresses old, familiar information, while the 'rheme' expresses new information (which, from another perspective, can be called → 'focus'). In Anglophone linguistics (including works on the Arabic language), the equivalents are sometimes considered to be → 'topic/comment' (most often), or 'theme/→ predicate'. The terminological preferences also show the methodological preferences of various authors. In generative grammar or functional analysis, for instance, the terms 'topic' and 'comment' are sometimes used as equivalents, with interpretations specific to the respective linguistic orientations. In what follows, the terms 'theme/rheme' or 'theme/predicate' are used as a prototypical pair for a type of binary sentence analysis of Arabic with certain characteristics. Among these characteristics are the following:

i. The explicit acceptance of the distinction between thematic structure and verbal structure, as formulated, for instance, by Beeston: "Of major general importance is the contrast between a thematic sentence structure, in which the theme occupies the first position, and a verbal sentence structure in which a verb predicate comes first" (1970:63);
ii. The explicit or implicit reference to the Arabic grammatical tradition, which admits such a distinction;
iii. The strict restriction of the analysis to the boundaries of the sentence corpus provided by Classical Arabic grammars; and
iv. The borrowing of concepts from certain 20th-century European schools of linguistics (e.g. 'theme' as logical, psychological, or cognitive subject, or concepts pertaining to pragmatic sentence analysis).

Although the equivalent of the terms 'theme/rheme' in traditional Arabic grammar is often considered to be the pair *musnad ʾilayhi/musnad* (→ *ʾisnād*), the most pertinent references to thematic structure are to be found in modern works that analyze the pair *mubtadaʾ/xabar* as it appears in the work of the Arabic grammarians. The term *mubtadaʾ* (→ *ibtidāʾ*) was simply translated as 'inchoative' by Wright (1962) and others. Sometimes, the term was inadequately equated with 'subject', for instance by Reckendorf (1921) and Blachère and Gaudefroy-Demombynes (1975). In order to correct this misunderstanding, Vallaro (1993–1997) introduces the pair 'theme/rheme' as a potential equivalent of the pair *mubtadaʾ/xabar* and, implicitly, the pragmatic perspective in sentence analysis outlined by the Prague school. Although according to Vallaro the term 'theme' does not cover all the cases discussed by the Arabic grammarians, he cannot provide a more appropriate term. One element in favor of the equivalence *mubtadaʾ* – theme is the fact that the initial preverbal position is inherent to the status of theme. With respect to the nominative case (*marfūʿ*) of the theme, the Arabic grammarians spoke of an 'abstract operator', which would be precisely the initial positioning (*ibtidāʾ*). On the other hand, even the 'debatable' examples provided by the Arabic grammarians, e.g. the inversion of the two elements in the prototypical nominal sentence in locative structures, can be dealt with within the framework of syntactic-pragmatic analysis presupposed by the use of the terms 'theme/rheme' (see references to Beeston 1970 and Kouloughli 2002, below).

An argument against equating the term *mubtadaʾ* with 'subject' and for approaching the thematic structure separately from the verbal one is provided by sentences, common in Arabic, with an 'isolated natural subject'. The term is introduced by Reckendorf (1921) as an equivalent for *mubtadaʾ*; it refers to the first element in sentences of the type *zaydun jāʾa ʾabūhu* 'Zayd, his father has come' and *zaydun raʾaytu ʾabāhu* 'Zayd, I have seen his father'. The isolated natural subject, which other linguists call a 'psychological subject' or 'natural subject', perfectly overlaps with what, in later analyses, appears as theme. Starting from this concept, Bravmann (1953) suggests reconsidering the interpretation of thematic structure in Arabic. He regards the isolated natural subject in the sentences above – the only element of the statement as a whole that is known by the speaker at the beginning of the utterance – as part of an implicit interrogation, hence its pronunciation with the specific interrogative intonation. In relation to the interrogation presupposed by this 'natural subject' (*what about* ***Zayd****?*), its 'natural predicate' (i.e. a sentence added after a pause of minimum duration) represents the answer addressed by the speaker to himself. The isolated natural subject structure is, in Bravmann's view, the initial structure; a later transformation merges the two phrases into one through the penetration of the original isolated natural subject into the interior of the sentence (something like 'Zaydun is [one] whose father I saw'). In this new structure, the isolated natural subject enters into closer contact with the initial word and may eventually be governed by it syntactically. Bravmann analyzes this type of sentence within the extended frame of the development of hypotaxis (→ subordination) from parataxis (→ sentence coordination). Here, we are interested in it to the extent to which the construction with isolation of the 'natural subject' (i.e. a kind of thematic structure) is linked to the relation between protasis and apodosis: the protasis is an interrogation addressed by the speaker to himself, the answer to which is contained in apodosis. This kind of parasyntactic analysis may be of interest for the pragmatic perspective on sentence analysis.

Bravmann's exposition raises some issues encountered in various other sentence analyses that start from the bipartite structure designated by the theme/rheme, including the following:

i. The means of grammatical encoding of theme and rheme, with special reference to case and determination;
ii. The status of the elements known in Arabic as → *nawāsix* (case operators which, from a semantic perspective, largely belong to the category of modalities) in relation to the thematic structure;
iii. The theme and rheme in locative sentences;
iv. Strategies of thematization and rhematization.

These topics are also discussed by Beeston (1970, Chap. 9, "Theme and predicate"), cited in later works (see below). Beeston observes that the theme occurs in sentence-initial position if functional words are not taken into account. In principle, a theme exhibits the syntactic marker of independent status, insofar as it is morphologically capable of being marked distinctively. Some functional words assign the status of direct object to the theme, which is then marked as if it had subordinate status. Probably, such functional words originally had some degree of verbal force. Restating one of the arguments used by the Arabic grammarians to justify the ability of this and other particles to 'operate' or to 'act', Beeston (1970) shows that this is best illustrated by the functional word *ʾinna*. Beeston contrasts two types of locative structures, one of which is thematic (*al-yawmu t-tālī māta fī-hi l-malik* 'the-day the-following died on-it the-king'), whereas the other is not formally thematic (*al-yawma t-tāliya māta l-malik*). In the first example, *al-yawm****u*** is marked by a morpheme of independent status and the encapsulated structure containing the referential phrase *fi-hi* 'on it', while in the second *al-yawm****a*** is marked as having subordinate status without a resumptive pronoun (→ resumption).

With respect to the locative sentence (→ locatives) in Arabic, Kouloughli states that "Beeston is right in stressing the similarity of locative sentences with topic-comment ones but his assimilation of the two types of sentences overlooks some important syntactic and semantic proprieties which differentiate them" (2002:21). Among the arguments that Kouloughli brings in support of assigning topic status (*mubtadaʾ*) to the locator in the locative sentences is the fact that the topic must be defined, syntactically or semantically, and must denote topical information already known to the hearer, namely the case of the locator. One of the strongest arguments against considering the locator as theme in sentences of the type *fī d-dāri rajulun* 'in the house [is] a man', however, is that introducing the functional word *ʾinna* in the above-mentioned sentence produces the same modification of the case of *rajul* (from nominative to accusative), with which it would have been modified had it occurred as theme in its usual position at the beginning of the sentence. This is stressed by Carter (1981:199, n. 9.73), in whose view Beeston's hypothesis is psychologically plausible but leaves unaccounted such genuine inversions as *ʾinna fī d-dāri rajulan* 'verily, in the house is a man', where *rajulan* still has dependent form as a subject operated on by *ʾinna*. In order to refute this objection, Kouloughli puts forth an alternative hypothesis about the case-assignment (→ *ʿamal*) property of *ʾinna* and its analogues (→ *ʾinna wa-ʾaxawātuhā*). According to him, the respective particles do not have two arguments, as the Arabic grammarians thought, but just one: they assign the accusative to the first noun (or noun phrase) they encounter. In locative sentences, "the first noun that can be a candidate to accusative case-assignment is the located term" (Kouloughli 2002:19).

The theme is normally definite, as it conveys information known by the speaker and the hearer. In locative sentences and in the existential sentences that are assimilated to them, the locator is definite when fronted, while the located noun is in most cases indefinite (being a predicate of class). In fact, as Kouloughli emphasizes, compared with topic/comment predications, locative predications may be qualified as 'weak'. This structure is used as presentational, i.e. for the mere introduction of an entity into the discourse (→ presentatives).

Besides locative sentence, the sentences that are of special interest to the theme/rheme analysis are the negative and the interrogative sentences. They have been discussed mostly in relation to the strategies of thematization or → topicalization.

A related issue is whether the thematic structure is the original, initial one, as some of the Arabic grammarians suggest. According to them, the nominative on the agent of the verb is copied after the nominative of the theme (*mubtadaʾ*), which, in its turn, is the result of an 'abstract operator' (*ʿāmil maʿnawī*), representing the communicative intention of the speaker. Alternatively, this could be the result of topicalization or dislocation. Fassi Fehri (1982:40–62) argues in favor of the latter solution and supports it with the fact that the basic word order in Arabic is VSO. He argues that in Arabic, as in all real VSO languages, the verb does not agree in number with the subject. In *jāʾa l-ʾawlādu* 'has-come the-boys', for instance, the verb is in the singular, and the apparent agreement in *al-ʾawlādu jāʾū* 'the boys have come' is, in fact, the result of agreement between the topic and the subject of the verb.

To reconcile the opinions of those who claim that the thematic structures are the original ones (see Bravmann 1953) and those who oppose this opinion, a distinction can be made between thematic structures and thematized structures. The latter are the result of thematization, which, in Arabic as in other languages, is accomplished through thematizers. The best known thematizer in Arabic is *ʾammā* (the contraction of *ʾan* and *mā*), often translated by 'as for...' and followed by the theme in the nominative, whereas the rheme is often introduced by *fa-*. The structure as a whole resembles the conditional utterance (→ *jazāʾ*), with its question/answer form (see above; Bravmann 1953). Cantarino (1975:III, 196–202), who also provides many examples of the use of *ʾammā* in Modern Standard Arabic, believes that in the case of the 'isolated noun' (i.e. the theme) introduced by this particle, one can speak of an anacoluthon (its syntactic function is most often expressed by the resumptive personal pronoun): *ʾammā ṣūratu-hu fa-**hiya** ʾajmalu ṣūratin xalaqahā llāhu fī l-ʿālam* 'as for its appearance, it is the most beautiful form God ever created on this world'. The anacoluthon has been mentioned in the case of structures with an isolated natural subject in general; Fleisch (1965:169), for instance, calls such structures "phrases brisées" (broken phrases): "Ces phrases ne peuvent plus s'analyser comme des phrases régulières: on commence une phrase, on la brise et l'on reprend une autre. C'est un procédé du langage affectif dans la recherche de l'expressivité". It is not adequate to speak of a 'broken' grammatical sequence, however, or of 'affective language' to account for a type of thematic structure that is very common in Arabic. It may be assumed that thematic structures in Arabic, as in other languages, originate with affective language. In support of this, their usage in 'popular syntax', in occurrences unknown to the literary language, may be mentioned here. However, when structures at certain levels of the language – such as Modern Standard Arabic – are interpreted, it is important to specify the theoretical framework in which the analysis is being conducted. The terminological pair 'theme/rheme' sometimes interferes in analysis with the pair 'topic/comment', in which case some linguists speak of thematization as well as topicalization as a result of movement. In order to differentiate the modalities of analysis that are made explicit in terminological differences, one should take into consideration the fact that, in modern linguistics, topic/comment are seen as structural concepts, while theme/rheme are seen as semantic-pragmatic concepts.

BIBLIOGRAPHICAL REFERENCES

Beeston, A.F.L. 1970. *The Arabic language today.* London: Hutchinson University Press.

Blachère, Règis and Maurice Gaudefroy-Demombynes. 1975. *Grammaire de l'arabe classique.* 3rd ed. Paris: G.P. Maisonneuve.

Bravmann, Meir. 1953. *Studies in Arabic and general syntax.* Cairo: Institut Français d'Archéologie Orientale.

Cantarino, Vicente. 1974–1975. *Syntax of Modern Arabic prose.* 3 vols. Bloomington and London: Indiana University Press.

Carter, Michael G. 1981. *Arab linguistics: An introductory classical text with translation and notes.* Amsterdam and Philadelphia: J. Benjamins.

Fassi Fehri, Abdelkader. 1982. *Linguistique arabe: Forme et interprétation.* Rabat: Faculté des Lettres et des Sciences Humaines.

Fleisch, Henri. 1965. *L'arabe classique: Esquisse d'une structure linguistique.* Beirut: Dar El-Machreq.

Kouloughli, Djamel Eddine. 2002. "On locative sentences in Arabic". *Zeitschrift für Arabische Linguistik* 41.7–26.

Reckendorf, Hermann. 1921. *Arabische Syntax.* Heidelberg: C. Winter.

Vallaro, Michele. 1993–1997. "Fra linguistica e glottodidactica: La traduzione della terminologia grammaticale araba. A proposito del *mubtadaʾ*". *Studi Maghrebini* (Napoli) 25.385–402.

Wright, William. 1962. *A grammar of the Arabic language.* 3rd ed. 2 vols. Cambridge: Cambridge University Press.

NADIA ANGHELESCU
(University of Bucharest)

Theta Roles

Theta roles (hereafter θ-roles) are the semantic roles borne by noun phrases in a sentence that designate 'agent of action', 'patient of action', and so forth. As such, they constitute an essential part of a sentence's semantic interpretation. The concept, though not the name, is used in the beginnings of Arabic grammatical study. Saad (1982:87) observes that the medieval grammarians classified sentence elements semantically, for example as → *fiʿl* 'action', → *fāʿil* 'actor', and → *mafʿūl* 'acted upon'. Comparable designations pervade Wright's (1974) grammar of Classical Arabic and Cantarino's (1975) grammar of Modern Standard Arabic.

The study of θ-roles in generative grammar begins with the work of Fillmore (1968) and Gruber (1976, based on his 1965 Massachusetts Institute of Technology dissertation). Gruber identified roles that still form the core of thematic role inventories: Theme (T – object in motion or at a location), Location (L), Source (S), Goal (G), and Agent (A – volitional performer of an action). These relations were termed 'thematic' because of the centrality of the Theme role to the system (Gruber 2003:257). The role Patient (P – undergoer of action), not part of Gruber's taxonomy, is often conflated with Theme, and Possessor is sometimes an alternative to (human) Goals (Fassi Fehri 1993:220; Amer 1996:226–229). Fillmore's semantic roles were developed in the context of case grammar and were called 'cases'. To Gruber's roles, Fillmore added Instrument (I). Some examples may clarify the concepts.

(1) *kataba* *l-walad-u* *d-dars-a*
A P
wrote the-boy-Nom the-lesson-Acc
fī l-bayt-i
L
in-the-house-gen
'The boy wrote the lesson in the house' (Farghal 1986:7)

(2) *ʾaʿṭā* *zayd-un* *hind-an*
S (A) G
gave Zayd-Nom Hind-Acc
kitāb-an
T
book-Acc
'Zayd gave Hind a book' (Amer 1996:140)

(3) *ʾaʿṭā* *zayd-un* *kitāb-an*
S (A) T
gave Zayd-Nom book-Acc
li-hind-in
G
to-Hind-Gen
'Zayd gave a book to Hind' (Amer 1996:140)

(4) *qaṭaʿ-at* *aṭ-ṭabbāx-at-u* *as-sujūq-a*
A P
cut-3fs the-cook.f-Nom the-sausage-Acc
bi-sikkīn-in
I
with-knife-Gen
'The cook cut the sausage with a knife'

The term 'θ-role' was introduced in Chomsky (1981:2.6) as a component of θ-theory, one of the modules of → Government and Binding (GB) theory. Assuming the role inventories of Gruber and Fillmore, Chomsky designated NPs to which θ-roles were assigned as 'arguments', and those which received no θ-role (e.g. idioms and expletives) as 'nonarguments'. Both argument and nonargument NPs occupy A-positions, which are D-structure positions associated with grammatical functions (GFs). A subset of these are also θ-positions to which θ-roles are assigned according to the associated GF (Chomsky 1986:93). θ-role assignment was subject to a bi-uniqueness condition known as the θ-Criterion (Chomsky 1981:36):

(5) Each argument bears one and only one θ-role, and each θ-role is assigned to one and only one argument.

(For exceptions to both clauses of (5), see Jackendoff 1987:381–383.) θ-roles are assigned by heads (A, V, N, P) to complement NPs (internal arguments), which they head-govern. The subject NP is assigned its role 'compositionally' by the VP, due to its unique status as the external argument of the predicate (Chomsky 1981:103–105; Williams 1981:83, 1994:32).

Movement complicates this picture. NPs move in Government and Binding theory partly to receive → Case (also → case roles). However, displacement leaves θ-role assignment intact (as a consequence of the Projection Principle, the requirement that a verb's selectional properties be preserved at all levels of syntactic representation). Thus, in (6), the focused NP *kitāban* is still the Theme despite its displacement to sentence-initial position.

(6) *kitāb-an* *wajada* *muḥammad-un*
book-Acc found Muhammad-Nom
'A *book* Muhammad found' (Shlonsky 2000:329)

To account for this fact, Chomsky (1986:96) proposed that θ-roles be assigned where necessary to movement chains, consisting of a moved NP and its trace, under the Chain Condition: "[A] chain is headed by a Case-marked position and terminates in a θ-position". In (6), the chain is {*kitāban, t*}; *t* is in object position (a θ-position) after the subject NP and transmits its θ-role to the chain. (It may be that *kitāb* checks its accusative Case in the specifier of Focus Phrase [Shlonsky 2000:329–331] rather than [Spec, Agr_OP], if the Chain Condition holds in full in Arabic.)

Chomsky (1981) stimulated much research into θ-roles, raising questions that remain high on the research agenda. Central among these is what Baker (1997:73) terms the 'linking problem': how are θ-roles mapped onto syntactic arguments? Neither the Projection Principle nor the θ-Criterion answers this question. Yet, there clearly are linking regularities that must be captured; for example, crosslinguistically the Agent role maps to subjects and the Theme role to direct objects (Marantz 1984:33; Baker 1997:76). The Uniformity of Theta Assignment Hypothesis (UTAH) of Baker (1988:46) is an extremely strong proposal concerning linking regularities:

(7) Identical thematic relationships between items are represented by identical structural relationships between those items at the level of D-structure.

Challenges to the UTAH have been put forward (Beletti and Rizzi 1988; Newmeyer 2001) and contested (Pesetsky 1995; Baker 1997). In (8a) and (8b) are Arabic examples of empirical objections the UTAH faces (Fassi Fehri 1993:239; cf. Grenat 1996:149, 156–157, 201, 217–219, 223–225).

(8a) *ʾaqlaqa-nī* *intiqād-u*
annoyed-me criticizing-Nom
r-rajul-i *l-mašrūʿ-a*
the-man-Gen the-project-Acc
'The man's criticizing the project annoyed me'

(8b) *ʾaqlaqa-nī* *intiqād-u*
annoyed-me criticizing-Nom
r-rajul-i *li-l-mašrūʿ-i*
the-man-Gen to-the-project-Gen
'The man's criticizing [criticism] of the project annoyed me'

The subject in each case is an → *ʾiḍāfa* headed by a process nominal, identified in Fassi Fehri's analysis by an event affix [E]. The thematic relations within the *ʾiḍāfa*s are identical: *ar-rajuli* is the Agent and *mašrūʿ-* the Patient. By (7), then, their D-structures should be the same. However, Fassi Fehri (1993:235, 240) accounts for the differences in the Case-marking of the Patient by positing different positions for E in D-structure, along with different positions and categories for the Patient. Alternative analyses that respect the UTAH – for example regarding *mašrūʿ-* as Goal instead of Patient – seem problematic. Similar problems arise in locative alternations (Baker 1997:86–97), as in (9) from Jordanian Arabic (Fareh and Hamden 2000:73).

(9a) *mūsa sattaf al-ikyās fi*
Musa crammed the-sacks in
d-dukkān
the-store
'Musa crammed the sacks into the store'
(9b) *mūsa sattaf d-dukkān bi*
Musa crammed the-store with
al-ikyās
the-sacks
'Musa crammed the store with sacks'

Despite being near-paraphrases, their grammatical roles diverge: in both sentences, *alikyās* is Theme and *ddukkān* the Goal, but the former is direct object in (9a) and oblique in (9b), while the converse holds for the latter.

A second question concerns the locus of θ-roles in the overall theory. Chomsky (1981) conceives of the θ-Criterion syntactically, as a well-formedness condition on Logical Form. This position reappears in minimalism in classification of θ-roles as V-features that arguments must check for convergence (Hornstein 2001:37–38; for the contrary view, see Chomsky 1995c:312; Collins 1997:69–72; Gruber 2003:258). Much subsequent work has instead emphasized the relevance of θ-roles to lexical semantic representation. Jackendoff (1987:378) locates θ-roles in Conceptual Structure, an autonomous semantic component, where they are defined configurationally by their positions in representations of verb meaning (cf. Bouchard 1991). Conceptual Structure is linked to the syntax by a set of correspondence rules which allow for a direct mapping and obviate the need for any mediate level of representation (Jackendoff 1987:379). While concurring with Jackendoff's program of associating θ-roles with lexical representation, other researchers insist that lexical semantic representations must be supplemented by information about how many arguments a verb takes. Typical of such representations are a[rgument]-structure and predicate-argument structure, in (10a) and (10b), respectively.

(10a) *announce* (Agent (Goal (Theme))) (Grimshaw 1990:6)
(10b) *put*: x (y P_{loc} z) (Levin and Rappaport 1988:15, 1995:21)

The representations differ superficially in that the names of the θ-roles appear only in a-structure; Grimshaw (1990:4, 10) concurs that specific roles play no role in syntactic processes and are included merely for convenience. In both representations, then, arguments are represented by variables, whose prominence relations are indicated by the parentheses: Agent and x are the external arguments of the predicates, while Goal and Theme, y and z, are internal arguments. These prominence relations are based on a thematic hierarchy (Jackendoff 1972:43; Grimshaw 1990:9; Baker 1997:105), which ensures that the arguments will be mapped to the proper hierarchical relations in the syntax. (For arguments that such hierarchies are 'epiphenomenal', see Gruber 2003:280–282.) P_{loc} stands for locative prepositions, which mark z as a locative expression (*Jan put the book* ***on the table***).

The question of whether a level of predicate-argument structure is necessary is unavoidable in the minimalist program (→ minimalism), in which the interface levels (Phonetic Form and Logical Form) are the only ones permitted (Chomsky 1995b, 1995c, 2000:113). An affirmative answer has been argued on the evidence of Form X derived verbs. Consider the sentences in (11) and (12).

(11) *is-ta-wṭan-a zayd-un fī*
Caus-Refl-dwell-3ms Zayd-Nom in
l-qāhirat-i
the-Cairo-Gen
'Zayd settled in Cairo'
(12) *is-ta-wzar-a*
Caus-Refl-be(come).minister-3ms
zayd-un al-muhandis-a

zayd-Nom the-engineer-Acc
'Zayd appointed the engineer [masc.] as a [cabinet] minister'

The meaning of *istawṭan-* in (11) is semantically compositional: 'to cause oneself to dwell'. Thematically, the reflexive prefix *ta-* bears the Theme role (since its location is asserted by *waṭan*); it is an external argument to *waṭan* and the internal argument of an abstract predicate CAUSE (the causee). The external argument of the entire verb is *zayd* as Agent (that is, causer). Taking compositionality as the null hypothesis, the compositional reading for *istawzara* in (12) would be 'to cause someone to be minister for oneself'. As in (11), the causer of the derived verb is *zayd* as Agent, and the causee is a variable bearing the Theme role, which maps to the direct object *al-muhandis* in the syntax. Since the causee role is filled, *ta-* functions as a Goal adjunct. In both sentences, *ta-* is bound to *zayd* as its antecedent; the causee variable in (11) cannot antecede the reflexive to mean that the engineer was appointed for his own benefit. There are then two differences in the status of *ta-* in (11)–(12): an argument-adjunct binding distinction and a (consequent) difference in θ-role. But neither distinction surfaces syntactically; therefore, neither can be stated in the syntax and should be stated over predicate-argument structure representations instead (LeTourneau 1997:98–100). Whether this argument is correct or not (a question to which the UTAH is pertinent), derived verb measures seem highly relevant to the question of how parsimonious syntactic theory can be.

Applications of θ-roles to Arabic include the following: (i) adjectival passives; (ii) unaccusative and → middle verbs; (iii) → causatives; and (iv) double object constructions.

Mughazy (2001) identifies adjectival passives with passive participles in Egyptian Arabic. Whereas the active participle is usually associated with the external argument of the source verb, the passive participle is associated with one of the verb's internal arguments, typically the Patient, implying a restriction to transitives. The pair in (13) illustrates both properties.

(13a) *iš-šibbāk il-maksūr*
the-window the-broken
'the broken window'

(13b) *_iš-šibbāk il-kāsir_
the-window the-breaking
'the breaking window'

However, the verb need not be transitive, and the internal argument need not be a Patient, as in (14).

(14a) *il-kelīm il-mamši ʿalēh*
the-rug the-walked on-it
meḥtāg tanḍīf
needing cleaning
'The walked-on rug needs cleaning'

(14b) *il-ʾatbāʾ il-mettākel men-ha*
the-dishes the-eaten from-her
lāzem t-it-ġasal
must 3fs-Pass-wash
'The eaten-from dishes must be washed'

In (14a), *mamši* is the passive participle of the intransitive *mašā* 'to walk' and *ilkelīm* is a Location, while in (13b) *il-ʾaṭbāʾ* is the object of *men* and so a Source (Mughazy 2001:144–145). Mughazy (2001:150) argues that (14a) is acceptable because, while there are two participants in the action denoted by *mamši* (the implicit Agent and explicit Location of walking), the result of the action (say, dirtying the carpet) refers only to the Location participant, allowing it to surface as the head modified by the passive participle. In (14b), the act of eating culminates in its cessation; the result of the act (say, dirtying the dishes) is pragmatically relevant, rendering the Source argument *il-ʾaṭbāʾ* acceptable.

Mahmoud (1991) distinguishes 'unaccusative' from 'middle' verbs and shows that only the former occur in Arabic. Unaccusatives participate in a transitive/intransitive alternation (semantically, an inchoative/causative alternation; Pesetsky 1995:12); in (15b), the unaccusative is the intransitive Form VII *inkasara* (Mahmoud 1991:120).

(15a) *kasara zayd-un*
broke Zayd-Nom
az-zujāj-a
the-glass-Acc
'Zayd broke the glass'

(15b) *in-kasara z-zujāj-u*
Intrans-broke the-glass-Nom
'The glass broke'

zujāj- is the Patient in both sentences, but it is the direct object in (15a) and the subject in (15b). Middle verbs are exemplified by the second member of the pair in (16) (Mahmoud 1991:123).

(16a) *John translates Arabic*
(16b) *Arabic translates easily*

Arabic is the Patient (or Theme) in both sentences, but it is the direct object in (16a) and the subject in (16b). However, a detransitivized counterpart to (16b) is impossible in Arabic (Mahmoud 1991:124), as in (17b).

(17a) *tarjama* *zayd-un* *an-naṣṣ-a*
translated Zayd-Nom the-text-Acc
'Zayd translated the text'
(17b) **ya-ta-tarjamu* *n-naṣṣ-u*
3ms-Intrans-translates the-text-Nom
bi-suhūlat-in
with-ease-Gen
'The text translates easily'

Mahmoud (1991:126–129) attributes the contrast between (15b) and (17b) to the non-eventive aspect of *tarjama*, in contrast to the eventive aspect of *kasara*. (On aspect as a complement to thematic relations, see Grimshaw 1990:22–28; Baker 1997:117–120.) Inasmuch as the thematic and grammatical relations of the NPs are the same in both the unaccusative and middle sentences, this aspectual distinction offers a semantic explanation for the difference in grammaticality.

Unaccusatives and middles are of further interest for two reasons. First, they instantiate Burzio's Generalization (Chomsky 1986:139):

(18) A verb (with an object) Case-marks its object if and only if it θ-marks its subject.

In both (15b) and (16b), the Theme is subject because there is no external θ-role to be assigned; therefore, it cannot receive accusative Case from the verb and receives nominative Case instead (Perlmutter's Unaccusative Hypothesis, as presented in Baker 1997:114).

Unaccusatives and middles are also of interest because the thematic and grammatical relations of the unaccusatives display the same linking patterns in Arabic as in English, despite the fact that the directions of derivation are opposite. Intuitively, unaccusatives are semantically simpler than their transitive counterparts; inchoatives have one argument, causatives two. Yet, while English transitives can plausibly be derived from intransitives, the reverse must be true in Arabic, since the causative in (14a) is Form I and the inchoative in (14b) a derived verb.

Causatives in Arabic are of two main types, syntactic and morphological, exemplified in (19a) and (19b), respectively.

(19a) *jaʿala* *zayd-un* *hind-an*
make Zayd-Nom Hind-Acc
tarquṣu
dance
'Zayd made Hind dance' (Saad 1982:65)
(19b) *kattabat* *al-bint-u*
Cause.wrote the-girl-Nom
ṭ-ṭālib-a *d-dars-a*
the-student-Acc the-lesson-Acc
'The girl made the [male] student write the lesson'

(19a) is overtly biclausal (there being two finite verbs), so the Agent role is assigned to both *zaydun* and *hindan* by *jaʿala* and *tarquṣu*, respectively. *Zaydun* is therefore the causer and *hindan* the causee. (19b) is overtly uniclausal, but the assignment of θ-roles is similar: *al-bintu* is the causer Agent of the act of writing, *aṭ-ṭāliba* is the causee Agent of the act of writing, and *ad-darsa* is the Patient. However, there is only one verb, *kattabat*, a Form II causative, to assign the Agent roles, in violation of the θ-Criterion. The θ-Criterion is satisfied in a biclausal structure like (19a); moreover, if the Uniformity of Theta Assignment Hypothesis is correct, (19b) should be biclausal as well. Benmamoun (1991:182–183) offers evidence from Moroccan Arabic to show that this is the case (cf. Hazout 1991:196–202, 1995:392–402 for Syrian and Standard Arabic):

(20a) *ad-drari*$_i$ *gals-u* *ḥda*
the-children sat-3p beside
baʿḍeyathum$_i$
each.other
'The children$_i$ sat next to each other$_i$'

(20b) *galls-u* *ad-drari*$_i$ *hda*
made-sit-3p the-children beside
*ba*ʿ*ḍeyathum*$_i$
each.other
'They made the children$_i$ sit next to each other$_i$'

(20c) **galls-u*$_i$ *ad-drari* *hda*
*ba*ʿ*ḍeyathum*$_i$
'They$_i$ made the children sit next to each other$_i$'

In (20a), the → reciprocal anaphor *ba*ʿ*ḍeyathum* must be bound by its antecedent in the minimal clause containing them both, in order to satisfy Principle A of the binding theory. This condition is met in the single clause constituting the sentence, in which the reciprocal is bound by the subject NP *ad-drari*. (20b, c) contain the morphological causative *gallsu*. If this sentence is uniclausal, we would predict that *ba*ʿ*ḍeyathum* would be bound by the null subject of the verb, on analogy with (20a). But it is not; the reciprocal must be bound by the direct object *ad-drari*, as in (20b). This suggests that morphological causatives are underlyingly biclausal, with (19b) having the abbreviated underlying representation (21) and the intermediate representation (22).

(21) [$_{\text{AgrSP}}$ [$_{\text{AgrS}}$ [$_{\text{AgrOP}}$ [$_{\text{spec}}$ ___] $_{\text{AgrO}}$ [$_{v\text{P}}$ [$_{\text{spec}}$ al-bint] $_{v}$ [$_{\text{VP}}$ [$_{\text{spec}}$ aṭ-ṭālib] [$_{\text{V}}$ katab] [$_{\text{NP}}$ [$_{\text{N}}$ ad-dars]]]]]

(22) [$_{\text{AgrSP}}$ [$_{\text{spec}}$ al-bint$_3$] [$_{\text{AgrS}}$-kattabat$_1$] [$_{\text{AgrOP}}$ [$_{\text{spec}}$ aṭ-ṭāliba$_2$] [$_{v\text{P}}$ [$_{\text{spec}}$ t_3] [$t_{v(1)}$] [$_{\text{VP}}$ [$_{\text{spec}}$ t_2] [$t_{\text{V}1}$] [$_{\text{NP}}$ [$_{\text{N}}$ ad-darsa]]]]]

Katab- assigns the Patient θ-role to *addars* under head government. VP compositionally assigns the Agent role to *aṭ-ṭālib* in its specifier; *v*P does the same for *albint* in the specifier of *v*, a 'light' causative verb (Kitahara 1997:9). The configuration in (21) thus accounts for the dual assignment of the Agent role in accord with the θ-Criterion and the Uniformity of Theta Assignment Hypothesis. But the variant in (23) challenges the latter (LeTourneau 1998:89).

(23) *kattab-at* *al-bint-u*
Cause.wrote-3fs the-girl-Nom
ad-dars-a *li-ṭ-ṭālib-i*
the-lesson-Acc to-the-student-Gen
'The girl made the student write the lesson'

While the thematic relations among the NPs in (23) are the same as in (19b), the surface (or underlying) grammatical relations are not (Hazout 1991:202–214). The problem is much the same as for the process nominals in (8), which also exhibit an accusative-genitive alternation.

With this problem in mind, we turn finally to true double object constructions, which come in two varieties (Amer 1996:140), as in (24a) and (24b).

(24a) *ʾa*ʿ*ṭā* *zayd-un* *hind-an*
gave.3ms Zayd-Nom Hind-Acc
kitāb-an
book-Acc
'Zayd gave Hind a book'

(24b) *ʾa*ʿ*ṭā* *zayd-un* *kitāb-an*
gave.3ms Zayd-Nom book-Acc
li-hind-in
to-Hind-Gen
'Zayd gave a book to Hind'

The alternation in process nominals and causatives occurs here as well. In both sentences, *zaydun* is the Source (and Agent), *kitāban* the Theme, and *hind-* the Goal. But while both *kitāban* and *hindan* complement the verb in (24a), the latter complements the preposition *li-* in (24b). One could evade this problem for the Uniformity of Theta Assignment Hypothesis by assigning *hindan* the Possessor role in (24a) (Amer 1996:229). But this is a dubious move. Gruber (1976:55) permits NPs bearing θ-roles to be possessors, which is not permitted in his system among θ-roles themselves. Ouhalla (2000:232) argues that possession is not a θ-role but a pragmatic inference "from the semantic relation of open-ended location". Amer (1996:234) himself admits that the Possessor role overlaps with the Goal role.

A better solution would be to discard the Possessor role and account for the syntactic differences in another way under a uniform algorithm of θ-marking. Suppose, following Amer, that *hind-* in (24a) is governed by a null preposition P which assigns neither Case nor θ-role. The structure of (24a, b) will then be as in (25a, b), slightly modifying Amer (1996:229(56), 237(68)), with the Goal in [Spec VP] and the Theme as complement to V (Baker 1997:99).

(25a) [$_{v\text{P}}$ [$_{\text{spec}}$ zayd] [$_{v'}$ [$_{v}$ [$_{\text{VP}}$ [$_{\text{spec}}$ [$_{\text{PP P}}$ [$_{\text{NP}}$ hind]]] [$_{\text{V}'}$ [$_{\text{V}}$ ʾaʿṭā] [$_{\text{NP}}$ kitāb]]]]

(25b) [$_{vP}$ [$_{\text{spec}}$ zayd] [$_{v'}$ [$_{v}$ [$_{\text{VP}}$ [$_{\text{spec}}$ [$_{\text{PP P}}$ li [$_{\text{NP}}$ hind]]] [$_{\text{V}}$ ʾaʿṭā] [$_{\text{NP}}$ kitāb]]]]]

θ-marking in (25a) proceeds as follows: (i) *ʾaʿṭā* assigns the Theme role to its complement *kitāb*; (ii) the $_{v'}$ *ʾaʿṭā kitāb* assigns Goal compositionally to the PP in its specifier (implying, as currently affirmed, that [Spec, VP] is an A-position), which is transmitted to *hind-* by P; (iii) after *ʾaʿṭā* raises to *v*, $_{v'}$ assigns the Source (Agent) role(s) compositionally to *Zayd*. θ-marking in (25b) is the same, except that the Goal role is assigned by *li-* instead of by $_{v'}$ (Amer 1996:226–237). Deriving surface order would then be a matter of right-adjoining *li-hindin* to $_{v'}$, from which position it will antecedent govern its trace in [Spec, VP].

Bibliographical references

Amer, Walid Mohammad Abdelghaffar. 1996. *On double object and dative constructions*. Ph.D. diss., University of Essex.

Baker, Mark. 1988. *Incorporation: A theory of grammatical function changing*. Chicago: University of Chicago Press.

——. 1997. “Thematic roles and syntactic structure”. *Elements of grammar: Handbook in generative syntax,* ed. Liliane Haegeman, 73–137. Boston: Kluwer Academic Publishers.

Beletti, Andrea and Luigi Rizzi. 1988. “Psych verbs and θ-theory”. *Natural Language and Linguistic Theory* 6.291–352.

Benmamoun, Elabbas. 1991. “Causatives in Moroccan Arabic”. Eid and Comrie (1991:173–195).

Bouchard, Denis. 1991. “From conceptual structure to syntactic structure”. *Views on phrase structure*, ed. Katherine Leffel and Denis Bouchard, 21–35. Dordrecht: Kluwer Academic Publishers.

Cantarino, Vicente. 1975. *Syntax of Modern Arabic prose*. 3 vols. Bloomington: Indiana University Press.

Chomsky, Noam. 1981. *Lectures on government and binding: The Pisa lectures*. Dordrecht: Foris.

——. 1986. *Knowledge of language: Its nature, origin, and use*. New York: Praeger.

——. 1995a. *The minimalist program*. Cambridge, Mass.: MIT Press.

——. 1995b. “A minimalist program for linguistic theory”. Chomsky (1995:167–217).

——. 1995c. “Categories and transformations”. Chomsky (1995:219–394).

——. 2000. “Minimalist inquiries: The framework”. *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. Roger Martin, David Michaels, and Juan Uriagereka, 89–156. Cambridge, Mass.: MIT Press.

Collins, Chris. 1997. *Local economy*. Cambridge, Mass.: MIT Press.

Eid, Mushira and Bernard Comrie (eds.). 1991. *Perspectives on Arabic linguistics*, III. Amsterdam and Philadelphia: J. Benjamins.

Fareh, Shedeh and Jihad Hamdan. 2000. “Locative alternation in English and Jordanian spoken Arabic”. *Poznan Studies in Contemporary Linguistics* 36.71–93.

Farghal, Mohammad Ali. 1986. *The syntax of* Wh-*questions and related matters in Arabic*. Ph.D. diss., Indiana University.

Fassi Fehri, Abdelkader. 1993. *Issues in the structure of Arabic clauses and words*. Dordrecht: Kluwer Academic Publishers.

Fillmore, Charles. 1968. “The case for case”. *Universals of linguistic theory*, ed. Emmon Bach and Robert T. Harms, 1–88. New York: Holt, Rinehart, and Winston.

Grenat, Mohamed Hasan. 1996. *Argument structure and the Arabic* maṣdar. Ph.D. diss., University of Essex.

Grimshaw, Jane. 1990. *Argument structure*. Cambridge, Mass.: MIT Press.

Gruber, Jeffrey. 1976. *Lexical structures in syntax and semantics*. Amsterdam: North-Holland.

——. 2003. “Thematic relations in syntax”. *Handbook of contemporary syntactic theory*, ed. Mark Baltin and Chris Collins, 257–298. Malden, Mass.: Blackwell.

Hazout, Ilan. 1991. *Verbal Nouns: Theta-theoretic studies in Hebrew and Arabic*. Ph.D. diss., University of Massachusetts-Amherst.

——. 1995. “Action nominalizations and the lexicalist hypothesis”. *Natural Language and Linguistic Theory* 13.355–405.

Hornstein, Norbert. 2001. *Move! A minimalist theory of construal*. Malden, Mass.: Blackwell.

Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar*. Cambridge, Mass.: MIT Press.

——. 1987. “The status of thematic relations in linguistic theory”. *Linguistic Inquiry* 18.369–411.

Kitahara, Hisatsugu. 1997. *Elementary operations and optimal derivations*. Cambridge, Mass.: MIT Press.

Lecarme, Jacqueline, Jane Lowenstamm, and Ur Shlonsky (eds.). 2000. *Research in Afroasiatic grammar: Papers from the Third Conference on Afroasiatic Languages*. Amsterdam and Philadelphia: J. Benjamins.

LeTourneau, Mark S. 1997. “Underspecification of lexical entries for Arabic verbs”. *Perspectives on Arabic linguistic*, X, ed. Mushira Eid and Robert R. Ratcliffe, 93–121. Amsterdam and Philadelphia: J. Benjamins.

——. 1998. “Case checking by morphological causatives in Standard Arabic: A minimalist approach”. *Al-ʿArabiyya* 31.89–122.

Levin, Beth and Malka Rappaport. 1988. “What to do with θ-roles”. *Thematic relations*, ed. Wendy Wilkins, 7–36. New York: Academic Press.

—— and Malka Rappaport. 1995. *Unaccusativity: At the syntax-lexical semantics interface*. Cambridge, Mass.: MIT Press.

Mahmoud, Abdelgawad T. 1991. “A contrastive study of middle and unaccusative constructions in Arabic and English”. Eid and Comrie (1991:119–134).

Marantz, Alec. 1984. *On the nature of grammatical relations*. Cambridge, Mass.: MIT Press.

Mughazy, Mustafa. 2001. "Adjectival passives and thematic roles in Egyptian Arabic: A cognitive semantics approach". *Studies in the Linguistic Sciences* 31:2.143–153.
Newmeyer, Frederick. 2001. "Grammatical functions, thematic roles, and phrase structure: Their underlying disunity". *Objects and other subjects: Grammatical functions, functional categories, and configurationality*, ed. William D. Davis and Stanley Dubinsky, 53–75. Dordrecht: Kluwer Academic Publishers.
Ouhalla, Jama. 2000. "Possession in sentences and noun phrases". Lecarme a.o. (2000:221–242).
Pesetsky, David. 1995. *Zero syntax: Experiencers and cascades*. Cambridge, Mass.: MIT Press.
Saad, George Nehmeh. 1982. *Transitivity, causation, and passivization: A semantic-syntactic study of the verb in Classical Arabic*. Boston: Kegan Paul.
Shlonsky, Ur. 2000. "Remarks on the complementizer layer of Standard Arabic". Lecarme a.o. (2000:325–343).
Williams, Edwin. 1981. "Argument structure and morphology". *Linguistic Review* 1.81–114.
——. 1994. *Thematic structure in syntax*. Cambridge, Mass.: MIT Press.
Wright, William. 1974. *A grammar of the Arabic language*. Beirut: Librairie du Liban. (Repr. of the 3rd ed., 1896.)

Mark S. LeTourneau (Weber State University)

Topic and Comment

1. Definition of terms

The topic of an utterance/sentence is defined here as that element which expresses information that is currently in the addressee's mind, or that is more generally expected to be known by the addressee, or that is relatively predictable as a whole, or that contains significantly predictable aspects, or that is presented by the speaker/writer as if it is known or predictable. 'Comment' is the converse of topic (i.e., it is that element of the utterance/sentence which conveys information not currently in the addressee's mind, or more generally not expected to be known by the addressee, etc.).

Comment is associated intonationally with primary accent (also called main stress, or sentence stress), which normally falls toward the end of the sentence or phrase. The topic is either unaccented, if the sentence consists of one intonation unit, or is associated with a secondary accent if the sentence includes more than one intonation unit. Thus, a topic/comment structure may involve recursion. When a sentence includes two intonation units, the information unit in which the primary accent occurs can be regarded in its entirety as having primary comment status, while the information unit in which the secondary accent occurs can be regarded in its entirety as having primary topic status. Within each of these intonation units, the accented word can be regarded as being a secondary comment (or part of a secondary comment) within that intonation unit, while those unaccented elements not identified as part of the secondary comment can be regarded as part of a secondary topic.

This definition of topic and comment is based on the notion that language is fundamentally communicative, and that communication requires the conveying of an element of information that is not already apparent to the addressee. Communication also suggests a common starting point based on common knowledge between speaker/reader and addressee. This starting point provides the communicative motivation for the topic as an utterance/sentence element. It also suggests an iconic motivation (in the Peircean sense of iconic; e.g. Hervey 1982:30–31) for the typical occurrence of the comment toward the end of the sentence or phrase, and the topic at the beginning: speakers/writers begin from their own conceptual starting point (typically shared with the listeners/readers) and move on to the additional information they want to communicate in respect of this.

Topic and comment are defined here in roughly the same way as → theme/rheme in the Functional Sentence Perspective (FSP) tradition of the Prague school (e.g. Firbas 1992). This differs from a number of other approaches, such as that of Halliday (e.g. Halliday and Matthiessen 2004:64–105), which distinguish between speaker/writer-oriented aspects of information presentation (designated 'theme' and 'rheme' by Halliday) and addressee-oriented aspects of utterance division (termed 'given' and 'new' information by Halliday). The view adopted here that topic and comment are structural in nature and can involve recursion, however, owes more to Halliday than to the tradition of the Functional Sentence Perspective. The terms 'topic' and 'comment' are more typically used by American than European linguists. In some approaches (e.g. Dik 1997), the term 'topic' (and also 'comment') is used in a more restricted sense than here.

2. TOPIC/COMMENT STRUCTURE AND NOMINAL CLAUSES

The → nominal clause (*jumla ismiyya* in the Arabic linguistic tradition), a clause consisting of a simple predicand (*mubtadaʾ bihi*; → *ibtidāʾ*) and a predicate (→ *xabar*), is a good starting point for considering the operation of topic and comment in Arabic. Clauses of this type are found both in Standard Arabic (including the Classical language) and in all colloquial dialects, and the grammatically bipartite nature of this clause type interacts in a simple and transparent way with the bipartite topic/comment distinction.

The examples in Section 2 are all drawn from → Khartoum Arabic. The fundamental principles discussed here are the same for other varieties. However, Khartoum Arabic has the advantage over other varieties in expressing definiteness in all cases by *al-*; this applies to nouns/nominal phrases, adjectives/adjectival phrases, 'semiclausal' (*šibh jumla*) adverbs/adverbial phrases, verbs/verb phrases, and more complex clauses (as shown below). In most varieties of Arabic, only nouns/nominal phrases and adjectives/adjectival phrases are made definite by *al-/il-*; elsewhere, a form such as *illi-* (in most dialects) or *allaḏī* (in Standard Arabic) is used (→ subordination). There is a strong connection between definiteness and topic: a definite element typically has a degree of knownness and is therefore more likely to be chosen as topic than an indefinite one.

Various word-class combinations are possible in predicand/predicate clauses, as illustrated in (1)–(8) below.

Predicand *Predicate*
noun prepositional phrase
(1) *ar-rājil* *fi l-bēt*
'The man is in the house'

noun noun
(2) *ar-rājil* *diktōr*
'The man is a doctor'

noun adjective
(3) *ar-rājil* *ġarīb*
'The man is strange'

(4) *ar-rājil* *zaʿlān [minn-ak]*
'The man is angry [with you]'

(5) *ar-rājil* *fāhim [al-muškila]*
'The man understands [the problem]'

noun verb
(6) *ar-rājil* *ziʿil [minn-ak]*
'The man got angry [with you]'

(7) *ar-rājil* *fihim [al-muškila]*
'The man understood [the problem]'

The predicand must be either a noun or a nominal, such as an *al-* phrase (phrase beginning with *al-*), as in (8).

Predicand *Predicate*
nominal (*al-* phrase) other
(8) *al-fihim al-muškila* *fi-l-bēt/diktōr/zaʿlān [minn-ak]/ziʿil [minn-ak]*
'The one who understood the problem is in the house/is a doctor/is angry [with you]/got angry [with you]'

The → word order is predicand/predicate. However, predicate/predicand order is also possible, and is obligatory where the predicand is indefinite and the predicate is a verb, as in (9).

Predicate *Predicand*
(9) *gāmat* *nār*
'A fire broke out'

Topic/comment principles interact with basic clause structure as follows. As in natural languages generally (see Sec. 1), topics typically precede comments. Thus, the clauses beginning with *ar-rājil*..., in (1)–(8) above, all have topic/comment order: the definite known element *ar-rājil* 'the man' (topic) precedes the indefinite unknown element (comment), the indefinite element carrying clause stress. Example (9) also has standard topic/comment order. Although both elements are indefinite, *nār* 'fire' is treated as unknown, etc. (comment), because of the relative predictability of *gāmat* 'broke out'. (Typically, fires only do a limited number of things: break out, go out, spread.)

Comment/topic word orders where the comment is the predicate and the topic the predicand are also possible, as illustrated in (10)–(13).

Comment *Topic*
(10) *ġarīb* *ar-rājil da*
'That man is strange'
(11) *zaʿlān [minn-ak]* *ar-rājil da*
'That man is angry [with you]'
(12) *ziʿil [minn-ak]* *ar-rājil da*
'That man got angry [with you]'

Such clause-initial comments are stressed and typically convey a sense of emotiveness.

The interaction between predicand and predicate and between topic and comment provides insights into the word order of most subject+verb-phrase clauses, i.e. clauses with a subject predicand and a verb-phrase predicate.

SV(O)
Subject-Verb(-Object) is understood in relation to the interaction between predicand and predicate and subject and verb-phrase clauses as follows. Subjects are typically definite and, therefore, topics. Verb phrases are indefinite and, therefore, comments. Subjects thus come first in the clause and verb phrases last, since topics typically precede comments. Objects come after verbs in line with the general tendency in Arabic for the head (here the verb) to precede the modifier (here the object).

V(O)S
Verb(-Object)-Subject word order occurs either (i) where the subject is indefinite and is therefore the comment (as in *gāmat nār* 'a fire broke out'), or (ii) where the verb phrase is an initial comment (as in *fihim al-muškila ar-rājil da* 'that man understood the problem'), as is the case in Khartoum Arabic and other dialects as well. In Standard Arabic, this latter patterning is rare, perhaps because it is typical of unplanned speech (with the final topic occurring as something of an afterthought). VO word order in VOS structures and in SVO clauses (see above) reflects the general tendency in Arabic for the head to precede the modifier.

Other word orders involving V and O are typically a function of placing the object before the verb. Such preposing gives the object a comment focus (with main clause stress) or, less commonly, a topic focus (with secondary clause stress). In the examples (13)–(16) below, the preposed element is underlined.

Comment focus
(13) *jīb fūl - fasīx agṣud*
'Get beans – I mean salted fish!'
(14) *xamsa dagāyig ma tammēna*
'We hadn't been here longer than five minutes'
(15) *aḍ-ḍuhur mā ṣallēt*
'I didn't even pray the noon prayer'

In the above examples, comment preposing gives the sense of restriction to a single entity (known in traditional Arabic rhetoric as *ḥaṣr*). Thus, in (15), for example, the preposed *aḏ̣-ḏ̣uhūr* conveys the sense 'the noon prayer specifically [in contradistinction to any other prayer]'.

Topic focus

Preposed topical objects are frequently contrastive, as illustrated by (16):

(16)	*[mumkin]*	*zōl*	*gāri*
	[possible (is) that]	person [who]	read.Act. Part
	ʿarabi	*faṣīḥ*	*kalām*
	Arabic	Standard	language
	aš-šāriʿ	*ma*	*bifham*
	the-street	not	understand.3ms

'[It may be that] a person who has studied/read Standard Arabic does not understand the language of the street'

Here, the preposed object *kalām aš-šāriʿ* 'the language of the street' contrasts with the previously occurring object of the active participle *gāri* '(has) read' *ʿarabi faṣīḥ* 'Standard Arabic'. In examples (13)–(15), the object is a comment and carries primary stress (cf. Sec. 1). In the phrase *kalām aš-šāriʿ ma bifham* in (16), by contrast, the object *kalām aš-šāriʿ* carries secondary stress and is therefore a topic, while *ma bifham* carries primary accent and is therefore a comment. While *kalām aš-šāriʿ* and *ma bifham* both contain new/unknown information, *kalām aš-šāriʿ* is treated as more predictable here than *ma bifham; kalām aš-šāriʿ* can be seen as a highly predictable polar opposite of *ʿarabi faṣīḥ*, while *ma bifham* is treated as a relatively less predictable element, contrasting with *gāri*.

3. Recursion

In both Standard Arabic and the modern dialects, predicand/predicate structures occur, in which the predicate is itself further analyzable into a predicand/predicate structure, as illustrated in example (17) from Khartoum Arabic.

	Main predicand	*Main predicate*	
		Secondary predicand	*Secondary predicate*
(17)	*ar-rājil da*	*bēt-u*	*kabīr*
	that man	his house	big
	'That man's house is big'		

Here, the definite predicand *ar-rājil da* 'that man' is followed by an indefinite predicate *bēt-u kabīr* 'his house [is] big', or equally '[is] one whose house [is] big'. The definite correspondent to *bēt-u kabīr* 'his house [is] big' is *al-bēt-u kabīr* 'the one whose house [is] big'. In *al-bēt-u kabīr*, the initial *al-* relates syntactically to the entire following phrase *bēt-u kabīr*, rather than solely to *bēt* 'house' or *bēt-u* 'his house'. The element *al-* in *al-bēt-u kabīr* 'the one whose house [is] big' functions as a definite particle in Khartoum Arabic, just as it does in *al-bēt* 'the house'. English grammar, however, requires that whereas *al-* can be translated as 'the' in *al-bēt* 'the house', it has to be translated as 'the one whose' in *al-bēt-u kabīr* 'the one whose house [is] big'.

In *ar-rājil da bēt-u kabīr* 'that man's house is big [lit. 'man-that house-his big']', the indefinite predicate *bēt-u kabīr* itself consists of a secondary embedded (definite) predicand *bēt-u* 'his house' and a secondary embedded (indefinite) predicate *kabīr* '[is] big'.

Example (18) below is to be analyzed along the same lines (here the indefinite *ana ma baʿrif-u* 'I don't know him'/'[is] one whom I don't know' has the definite correspondent *al-ana ma baʿrif-u* 'the one whom I don't know').

	Main predicand	*Main predicate*	
		Secondary predicand	*Secondary predicate*
(18)	*ar-rājil da*	*ana*	*ma baʿrif-u*
	that man	I	not I know him
	'I don't know that man'		
(19)	*ar-rājil da*	*ana*	*jīt maʿā-hu*
	that man	I	came with him
	'I came with that man'		

The normal topic/comment order for clauses in which the main predicate is further analyzable into a secondary predicand/predicate structure is topic followed by comment, as in the above examples. As with simple predicand/predicate structures, however, it is possible for the comment to be preposed, as in (20).

	Main predicate		*Main predicand*
	Secondary predicand	*Secondary predicate*	
(20)	*hum*	*xawājāt*	*al-biskunu*
	they	Westerners	the-live.3p
	fi	*l-bēt*	*da*
	in	the-house	that

'They are Westerners, the people who live in that house' *or*
'The people who live in that house are Westerners'

Here, the main predicand *al-biskunu fi l-bēt da* 'the [people] who live in that house' is also the main topic. The internal structure of the main predicate, and also main comment, *hum xawājāt* 'they are Westerners' can be analyzed as follows: *hum* is a secondary predicand as well as a secondary topic, while *xawājāt* is a secondary predicate as well as a secondary comment.

4. Topic/comment structure and verbal clauses

In Standard Arabic, a → verbal clause (*jumla fiʿliyya*) is a clause which contains a main verb and either lacks an independent nominal/

pronominal subject or its subject follows the main verb. Thus, verbal clauses in Standard Arabic have VS, VSO, and VOS word orders. Clauses beginning with a verb are the norm not only in early (pre-Classical) Arabic (Dahlgren 1998:216) but also, in varying degrees, in many modern dialects of Arabic, particularly the Eastern dialects (Dahlgren 1998:189, 205). In certain dialects, for instance Khartoum Arabic, verbal clauses are, by contrast, extremely rare. As suggested in Section 2, in many modern dialects, VOS word order is better analyzed grammatically as a bipartite nominal clause (*jumla ismiyya*) consisting of the syntactic elements VO (predicate) and S (predicand) with an associated topic/comment or comment/topic structure rather than as a verbal clause of the Standard Arabic type.

In Standard Arabic, VOS structures are traditionally analyzed as verbal rather than nominal sentences, reflecting, among other things, the similarity in subject-verb agreement patterns between VSO and VOS sentences. In topic/comment terms, however, Standard Arabic VOS structures seem typically to involve a VO constituent followed by an S constituent. Consider example (21) from Holes (1995:205), describing the 1990 Iraqi invasion of Kuwait.

(21) *wa-jtāḥat hāḏā l-balad aṣ-ṣaġīr al-ʾāmin al-musālim miʾātu d-dabbābāt*

This might be translated fairly literally into English as ‘And overran this small, secure, peaceable country hundreds of tanks’. A more idiomatic translation would be ‘This small, secure, peaceable country was overrun by hundreds of tanks’. Here *hāḏā l-balad aṣ-ṣaġīr al-ʾāmin al-musālim* ‘this small, secure, peaceable country’ is clearly the topic (giving known information), or part of the topic together with the verb *(i)jtāḥat*, and is intonationally closely linked to this verb. By contrast, *miʾātu d-dabbābāt* ‘hundreds of tanks’ is a comment (giving unknown information) and is intonationally separated from what comes before.

Different writers have presented different accounts of the topic/comment structure of verbal clauses. Baker (1992:125–128) analyzes the verb in a verbal clause as the theme (topic), arguing that “a process-centred pattern is far more typical of Arabic” (1992:128) than of English, whose standard SVO order typically highlights the actor in a narrative. A weakness of Baker’s analysis is that she adopts a rather mechanistic approach to the Hallidayan notions of → theme and rheme, analyzing the element which comes first in the Arabic sentence as theme on the basis of its position in the sentence, and then assuming that it has the discourse functions ascribed in the Hallidayan model to theme, without investigating closely whether this is really the case.

Holes (1995:264–266) has argued that narrative material in Modern Standard Arabic tends to display VS(O) word order, while descriptive material tends to display SV(O) word order. These tendencies are illustrated by example (22) from *aš-Šarq al-ʾAwsaṭ* magazine (Oct. 11, 1994; reproduced in Dickins 2005:46), about the satirical Palestinian cartoonist Nāji al-ʿAli, who was murdered by an unidentified gunman in London in 1987. Relevant verbs and subjects are labeled ‘verb’ or ‘subject’.

(22) *wulida* (verb) *nāji al-ʿali* (subject) *fī qaryat aš-šajara ʿām 1936* [...]

wa-qad kānat (verb) *ḥayātu-hu tilka bayna l-xiyām* (subject) *maṣdar ʾilhāmi-hi fī rusūmi-h fīmā baʿd [...]. wa-baʿd ḥuṣūli-hi ʿalā šihādat diblōm al-mēkānīka daxala* (verb) *nāji* (subject) *ʾakādīmiyyat al-funūn fī lubnān [...]*

wa-sirr najāḥ nāji al-ʿali (subject) *yanḥaṣir fī ʿafwiyyati-hi ṣ-ṣādiqa, allatī lam yaʿhad-hā l-fann as-sāxir al-ʿarabī. wa-ʿafwiyyatu-hu* (subject) *tanṭaliq* (verb) *min xilāl rusūm-in multazima [...]*

‘Nāji al-ʿAli was born in the village of Al-Shajara in 1936 [...]

That life which he lived among the tents was the source of his inspiration in his drawings in what was to follow [...]. Following his obtaining a diploma in mechanics Nāji entered the Academy of Arts in Lebanon [...]

The secret of Nāji al-ʿAli lies in his truthful spontaneity, which Arab satirical art had not known before. His spontaneity emerges from his committed drawings [...]’

In the first two paragraphs, Nāji al-ʿAli’s life story is narrated using VS word order. By contrast, his art is described in the final paragraph using SV word order.

In the most extensive study of Arabic → word order to date, Dahlgren (1998) proposes a model for the difference between → nominal clauses and → verbal clauses in a number of dialects, as well as in early (pre-Classical) Arabic, which incorporates aspects of traditional topic/comment notions but also goes beyond them. Dahlgren suggests that in many Arabic varieties, VS(O) word order is used in narrative to convey foreground information, i.e. information which conveys the main story line (cf. also Hopper 1979), "characterized by events that come in sequence, one after the other, to give the skeleton of the narrative" (Dahlgren 1998:61). SV(O) word order, by contrast, is used to convey background or supportive material "not in sequence with the main story line: it may be concurrent or located at any other point of the time axis" (Dahlgren 1998:61).

Several writers have noted that other features may influence choice of VS(O) or SV(O) word order. In an analysis of part of a short story by the Iraqi writer Maḥmūd aḍ-Ḍāhir, Somekh (1991:32) shows how the main narrative is conveyed by SV(O) word order, while the thoughts and feelings of the hero are conveyed by VS(O) word order. Watson (1999) points out that where Arabic newspaper headlines contain a verb, the word order is almost invariably SV(O) (→ media Arabic).

It is difficult to regard the various uses of VS(O) and SV(O) word order in different varieties of Arabic identified by Holes, Dahlgren, Somekh, and Watson as purely a function of topic and comment as these notions have been defined here. Dahlgren proposes that "we may actually see the VS order in Modern Colloquial Arabic as the unmarked order in a Functional *Text* Perspective (the foreground/background distinction is relevant), and SV order as the unmarked order in a Functional *Sentence* Perspective (the theme-rheme structure is relevant)" (1998:183). One way to partially bridge the gap between Dahlgren's 'functional text perspective' and his 'functional sentence perspective', while incorporating Holes' insights, is to say that narrative sequential action is predictable or presupposed by the nature of the text type: regardless of whatever else it expresses, an initial verb in narrative involves some significantly predictable – i.e. topical – element. By contrast, in descriptive texts, the existence of descriptive elements (characters, etc.) is similarly presupposed: regardless of what else it expresses, an initial subject denoting descriptive elements involves a significantly predictable (topical) element.

An associated point of importance is that description (as a text type) is prototypically static, i.e., it does not involve a change of state and as such contrasts with the prototypically dynamic nature of narrative. Static situations are prototypically described in Arabic through verbless sentences (verbs as the primary linguistic means of indicating change of state verbs are, by contrast, the central feature of narrative). Verbless sentences are necessarily nominal. They can thus perhaps be regarded as providing the basic 'structural template' for descriptive texts: those sentences in descriptive texts which contain a verb tend to mirror the structure of sentences which do not, by employing a predicand/predicate (*jumla ismiyya* 'nominal clause') structure.

In Standard Arabic, the emphatic particle *ʾinna* is frequently used to introduce the initial nominal element in a nominal clause, and is thus by definition associated with topic/comment structures. The major discourse function of *ʾinna*, however, is not to emphasize the predicand (and topic) which it precedes and grammatically governs, but rather to emphasize the entire clause of which it is the initial element (→ *ʾinna wa-ʾaxawātuhā*; → presentatives). An outline account of different subtypes of emphasis which may be relayed by *ʾinna* is given in Dickins and Watson (1998:421–428).

5. PREPOSED ADVERBIALS

Both Standard Arabic and the dialects allow for the placing (preposing) of an adverbial element before the syntactically central parts of the clause. Some dialects, such as Khartoum Arabic, commonly accept an adverbial before a nominal clause. In Classical Arabic this kind of sentence organization may be ungrammatical, and even in Modern Standard Arabic it is rare.

Both Classical Arabic and Modern Standard Arabic allow the preposing of an adverbial before a verbal clause, the latter more commonly than the former. In Dickins and Watson (1998:337–344), it is suggested that in such clauses, the preposed adverbial is an emphatic topic, and it may serve a number of more specific contextual purposes, the most prominent of which is perhaps linkage to previous textual material, as in (23)–(26).

Spatial linkage
(23)
wa-kaṯīran mā yanbuṭu ʿinda-hum fī jabal mākūṣ, wa-min hunāka jamaʿtu-hu ʾayyām kuntu hunāk (from Ibn al-Bayṭār, *al-Jāmiʿ li-mufradāt al-ʾadwiya wa-l-ʾaġḏiya*; reproduced and analyzed in Dickins and Watson 1998:344)
‘It frequently grows in their part of the world on Mount Makus; and I collected it [from] there when I was there’

Temporal linkage
(24)
fa-lammā kāna ʾawwal al-layl ʿāda wa-qaḍā sāʿa fī ḍaḥk wa-ʿabaṯa maʿa ʾixwati-hi. wa-fī hāḏihi l-layla zaʿama li-ʾahl al-bayt ʾanna fī ʾakl aṯ-ṯūm wiqāya min al-kūlīra…(from Ṭāhā Ḥusayn, *al-ʾAyyām*; Dickins and Watson 1998:344)
‘At the start of the night he came back and spent an hour laughing and joking with his brothers. That night he told all the people of the house that eating garlic warded off cholera…’

Manner linkage
(25)
kānat ṭayyibat al-qalb, wa-bi-hāḏihi ṭ-ṭība xarrabat bayt ʾaʿazz ʾaṣdiqāʾī (from ʾAnīs Manṣūr, *Baqāyā kull šay*ʾ; Dickins and Watson 1998:344)
‘She was good natured; and through this good nature she destroyed the household of one of my dearest friends…’

Logical linkage
(26)
buʿayd wafāt ar-rasūl – ṣallā ḷḷāhu ʿalay-hi wa-sallam – badaʾ al-ʿArab futūḥa-hum allati waḍaʿat taḥta taṣarrufi-him xilāl qarnin wāḥid jamīʿ al-minṭaqa š-šāsiʿa l-mumtadda min ʾawāsiṭ ʾĀsya wa-ḥawḍ as-Sind šarqan ʾila šimāl ʾIsbāniyā ġarban, wa-bi-ḏālika ʾaqāmū dawlat al-xilāfa…(from BBC broadcast *ʿAbqariyyat al-ḥaḍāra al-ʿarabiyya*; Dickins and Watson 1998:344)
‘Shortly after the death of the Prophet Muḥammad, the Arabs began their conquests which within one century placed under their control all of the vast area stretching from central Asia and the Indus Basin in the east, to northern Spain in the west. They thus [i.e. by doing this] set up the Caliphal state…’

In each of the above cases of linkage, the linking element (*fī hāḏihi l-layla* ‘that night’, *bi-hāḏihi ṭ-ṭība* ‘through this good nature’, and *bi-ḏālika* ‘thus’) draws on information that is already known, because it has been established in the previous clause.

In terms of the topic/comment model adopted here, the initial adverbial in each of these sentences is analyzed as a main topic, while the following verbal clause is analyzed as a main comment. Within each of the main topics and the main comments, it would be possible to further analyze a secondary topic/comment structure (as already suggested in Sec. 3 for verbal sentences).

Particularly in historical or biographical narrative, Modern Standard Arabic frequently (though not as frequently as English) preposes time adverbials that may not so obviously involve an element of known information. For example, a text describing the life of Ṭāhā Ḥusayn, from *aš-Šarq al-ʾAwsaṭ* (Oct. 24, 1992; reproduced and analyzed in Dickins and Watson 1998:342–343, 469-470) preposes a number of adverbial time phrases, e.g. *wa-baʿda ʾan ʾatamma ḥifḏ̣ al-Qurʾān al-karīm* ‘[and] after he had completed the memorization of the Noble *Qurʾān*’; *wa-fī ʿām 1918* ‘[and] in the year 1918’; *wa-fī ʿām 1963* ‘[and] in the year 1963’.

In Section 4, it was suggested that initial verbs in VS(O) structures do not immediately fit the definition of topic used here, but they may be said to fit it after all, given a broader interpretation of topic to include what is presupposed by the text type. Similarly in historical or biographical narratives, preposed time adverbials are not obviously topical in that they do not relay information that is already known (etc.).

They might, however, be said to be topics in the wider sense that historical and biographical narratives presuppose forward movement through time. While the specific time information (date, etc.) given in the preposed time adverbial is not known or predictable, the passing of time itself is highly predictable in such texts.

6. WIDER PERSPECTIVES

For further information, the following publications may be referred to. Obeidat (1994) provides an account of the discoursal patterning of thematic (topic/comment) structures in Modern Standard Arabic and English. Abdul-Raof (1998) considers the relationship between subject, theme (topic), and agent in Modern Standard Arabic. Dickins a.o. (2002:116–123) consider the relationship between topic/comment (theme/rheme), grammatical mainness/subordination, and foregrounding/backgrounding (cf. Sec. 3; → grounding) in Modern Standard Arabic and English. Kammensjö (2005) uses topic/comment and related notions in analyzing discourse connectives in formal spoken Arabic.

BIBLIOGRAPHICAL REFERENCES

Abdul-Raof, H. 1998. *Subject, theme and agent in Modern Standard Arabic*. Surrey: Curzon.
Baker, Mona. 1992. *In other words: A coursebook on translation*. London and New York: Routledge.
Dahlgren, Sven-Olof. 1998. *Word order in Arabic*. Göteborg: University of Göteborg.
Dickins, James. 2005. *Thinking Arabic translation: Supplement*. www.languages.salford.ac.uk/staff/dickins/php.
——, Sandor G.J. Hervey, and Ian Higgins. 2002. *Thinking Arabic translation*. London and New York: Routledge.
Dickins, James and Janet C.E. Watson. 1998. *Standard Arabic: An advanced course*. Cambridge: Cambridge University Press.
Dik, Simon C. 1997. *The theory of Functional Grammar*. I. *The structure of the clause*. Berlin and New York: Mouton de Gruyter.
Firbas, Jan. 1992. *Functional sentence perspective in written and spoken English*. Cambridge: Cambridge University Press.
Halliday, Michael A.K. and Christian Matthiessen. 2004. *An introduction to Functional Grammar*. 3rd ed. London: Edward Arnold.
Hervey, Sandor G.J. 1982. *Semiotic perspectives*. London: George Allen and Unwin.
Holes, Clive. 1995. *Modern Arabic: Structures, functions and varieties*. London and New York: Longman.
Hopper, Paul J. 1979. "Aspect and foregrounding in discourse". *Syntax and semantics*, XII, ed. Talmy Givón, 213–241. New York: Academic Press.
Kammensjö, Heléne. 2005. *Discourse connectives in Arabic lecturing monologue*. Ph.D. diss., University of Göteborg.
Obeidat, Nawaf. 1994. *The preservation of thematic structure in English Arabic translation*. Ph.D. diss., University of Durham.
Somekh, Sasson. 1991. *Genre and language in modern Arabic literature*. Wiesbaden: O. Harrassowitz.
Tomlin, Russell S. 1987. *Coherence and grounding in discourse*. Amsterdam and Philadelphia: J. Benjamins.
Watson, Janet C.E. 1999. "The syntax of headlines and news summaries in Arabic". *Arabic grammar and linguistics*, ed. Yasir Suleiman, 161–181. London: Curzon Press.

JAMES DICKINS (University of Salford)

Topicalization

1. INTRODUCTION

'Topicalization' refers to the placing of a constituent in a place other than its default, or normal, position in order to mark it as the theme or topic of the sentence. In VSO and SVO languages, it means fronting or moving toward the front, according to the principle 'old information precedes new information'. In the Chomskyan tradition, topicalization is regarded as a transformation from an underlying deep structure, which results in a surface structure with a topicalized constituent. Many contemporary theories today do not posit any underlying level of syntactic representations (see Van Valin and Lapolla 1997:21). A passive construction, for example, which would be seen as a transformation from an active clause in Chomskyan grammar, would be regarded as a question of figure/ground organization in a cognitive grammar, such as that developed by Langacker (1987, 1991). This is illustrated in (1).

(1a) *The farmer shot the rabbit*
(1b) *The rabbit was shot by the farmer*

In (1a), *the farmer* is the figure, i.e. in the center of one's attention, and *the rabbit* is in the ground, i.e. within the scope of the viewer but outside the figure. In (1b), *the rabbit* is the figure, and *the farmer* belongs to the ground (Taylor 2002:11).

Van Valin and Lapolla (1997:8–16) refer to the Chomskyan grammar perspective, which includes Relational Grammar, as being 'syntactocentric'. In this view, language is seen as an abstract object whose structure is to be studied independently of psycholinguistic, communicative, sociocultural, and other considerations. An alternative perspective is labeled the 'communication and cognition' perspective, to which Van Valin and Lapolla relate some fifteen linguistic theories and several well-known linguists who share the same outlook without being associated with a particular school. This latter perspective views human language as a means of communication with broader cognitive processes, such as reasoning and conceptualization, which are closely linked to other cognitive systems such as perception and knowledge.

In the Prague school, with the latter perspective, topicalization belongs to what Daneš (1966) calls the third level of language. Syntactic organization takes place on the first level, with the traditional notions of subject, predicate, direct and indirect objects, and so on. The second level concerns semantic roles, e.g. agent, force, patient, and instrument. The third level describes the organization of the utterance, which is structured according to the communicative purposes of the speaker.

The representatives of the Prague school were the main pioneers in elaborating theories regarding the third level of language. Through their approach, known as the 'functional sentence perspective', they developed the concepts of → theme and rheme. In 1939, Mathesius defined the theme as "that which is known or at least obvious in a given situation and from which the speaker proceeds", and three years later as "that which is spoken about in the sentence" (quotations from Firbas 1974:13). The rheme is what the speaker says about the theme (Firbas 1974:13). Halliday (1967) elaborated the different definitions on the theme by Mathesius and proposed a division into thematic structure and information structure. He regards the thematic structure with its theme and rheme as speaker oriented, i.e., it concerns what speakers choose as the point of departure for their message and what they say about that element. Information structure, on the other hand, is listener oriented. It is divided into the *topic*, which belongs to what the speaker assumes is known information to the listener, and *focus*, the new information that the speaker imparts to the listener. The thematic structure is expressed through the ordering of words, whereas information is expressed through prosody, where the more prominent pitch falls on the focus, which is located in the larger *focus domain* (→ topic/comment).

2. Old and new information

Extensive research has been done on the concepts 'new information' and 'old information'. Chafe (1976) suggests that 'known' or 'given' were topics that have been activated in the ongoing discourse – not any known information. Sanford and Garrod (1981) speak of 'scenario' as a way of directly invoking known information that has not been activated in the discourse. They use the example of a courtroom, which has well-known participants, such as judge, lawyer, and witness, all of whom can be spoken of in definite terms without an introduction. Fillmore (1982:111) uses the term 'frame' for the same phenomenon. Prince (1981) provides a taxonomy for all types of information; she distinguishes between new, inferable, and evoked information. 'New' is partly brand new, partly (known but) unused information; 'inferable' means scenario or frame-related information; and the last category is divided into 'situational evoked entities' and 'textual evoked entities'. The former concerns topics inferred from the speech situation, such as *you* and *me*, without having been mentioned before; the latter is divided into types: current and displaced entities. 'Displaced' means already introduced in the discourse but then replaced by other topics. Givón (1990) discerns three main contextual sources for the topics, i.e. the known information from which a topic can be picked. These are the 'deictically shared context', which is the speech situation; the 'generically shared context', which is the common cultural knowledge that practically all members in a speech community have access to; and the 'textually shared context', which refers to topics already introduced in the discourse.

The work on information structure by Lambrecht (1994) is an important milestone in this field. Following Dahl (1976:38), he uses 'presupposition' for old information. A presupposition is propositional information "which the speaker assumes the hearer already knows or is ready to take for granted at the time the sentence is uttered". The pragmatic presupposition of an utterance, then, is the set of such propositions lexicogrammatically evoked in a sentence (Lambrecht 1994:52). In the sentence *I finally met the woman who moved in downstairs*, the following set of pragmatic presuppositions can be identified:

(i) The addressee can identify the female designated by the definite noun phrase.
(ii) Someone moved in downstairs from the speaker.
(iii) One would have expected to have met that individual at some point in time.

The first presupposition is evoked by the definite article *the*, the second by the relative clause *who moved in downstairs*, and the third by the adverb *finally*. To these three presuppositions about the knowledge state of the addressee, a fourth presupposition concerning the consciousness presuppositions must be added:

(iv) The addressee is aware of the referents of the pronouns *I* and *who* at the time these pronouns are uttered.

Fifth, the sentence evokes the following relevance presuppositions via the two unaccented pronouns:

(v) The proposition expressed by the sentence is construable as relevant information about the referent of *I*; the proposition expressed by the relative clause is construable as relevant information about the referent of *who*.

Lambrecht follows Strawson (1964) in regarding the following three principles as essential components of a theory of linguistic information:

I. The Presumption of Knowledge
II. The Presumption of Ignorance
III. The Principle of Relevance

The first and third principles are reflected in the five points above. The second concerns what has been labeled as 'new information' in a sentence. Lambrecht modifies this concept and calls it 'assertion'. New information and assertion belong to the focus domain, and are treated under → focus.

When it comes to topic, Lambrecht only uses the notion of 'sentence topic'; he disregards any understanding of topic on a larger level than the sentence. He accepts the notion of 'being about', but asserts also that it is insufficient; a correct conception must also include Strawson's Principle of Relevance, since one cannot add just any potential topic in the discourse and expect it to be understood. So, besides expressing 'what the sentence is about', the topic is also 'what is a matter of standing current interest or concern'. This means that although it is a concept at the sentence level, more than one sentence may be needed to determine what the topic is in the sentence. A sentence such as *The children went to school* may be an ordinary, unmarked, 'topic/comment' utterance, as in (What did the children do next?) *The children went to SCHOOL*. But it may also be an 'identificational' sentence, as in (Who went to school?) *The CHILDREN went to school*, or it may be an 'event-reporting' sentence, as in (What happened?) *The CHILDREN went to SCHOOL*. Both the identificational and the event-reporting sentence have a subject that is not a topic. The identificational sentence has the presupposed open proposition 'X went to school', but Lambrecht does not regard it as a topic, since it has no referent. The subject in this type of sentence is an 'argument focus'. In the event-reporting sentence, the pragmatic presupposition is merely that something happened, and the whole sentence belongs to the focus domain (Lambrecht 1994:119–122).

3. Topicalization and word order

In an investigation of word order in Early Arabic and modern Arabic dialects, it was found that only SVO order is possible for topicalization of the subject and only VOS order for topicalization of the (direct) object (Dahlgren 1998; for statistics on different word orders referred to below, → word order). Exceptions to this pattern were found only in → Anatolian Arabic dialects and → Uzbekistan Arabic, which have been influenced by Kurdish and Turkish, both SOV languages.

Topicalization of the subject often occurs with *ʾinna* (Dahlgren 1998:214–216). A sample of 50 instances with SV order in foregrounded narrative in the *Sīra* of Ibn ʾIsḥāq presented 39 instances out of 47 definite nouns with a preposed *ʾinna*. Two of the three instances with indefinite nouns also occurred with *ʾinna*. A study of these shows that they presented an important new topic in the discourse. The remaining instances with definite nouns were introduced by *wā*, *ʾillā ʾanna* (1), and *ʾiḏā* (1). In the *Qurʾān*, SV order is common both with *wa-* and *ʾinna*; asyndetical linkings are found after *qul* 'say!' in a few cases.

Topicalization of the object means VOS order. An example from Damascus Arabic is given in (2).

(2) *baʿd ma šreb əš-šay,*
after that drank.3ms the-tea
ʾəja šu
came.3ms what
ʾəsm-o... ʾəja ʿamāra
name-his came.3ms Amara
ʾal-l-o:
said.3ms-to-him
'After he had drunk the tea, came – what's his name... Amāra came. He said to him:

wēn əl-mənfāx? ʿṭī-ni
where the-pump give.Imper-me
l-mənfāx ḥatta
the-pump so.that
ʾənfox əl-mōtōr tabaʿ-i
pump.1s the-engine of-me
'Where is the pump? Give me the pump, so that I can pump up the engine'

ʾaxad hāda l-mənfāx
took.3ms this the-pump.engine
He took this pump,

ʿAmāra nafax əl-mōtōr...
Amāra, pumped.3ms the engine.
Amāra, he pumped up the engine...' (Grotzfeld 1965:135)

Here is a case of two topical constituents in one sentence (Lambrecht 1994:146–150). Both *əl-mənfāx* and *ʿamāra* are spoken about *and* of current interest, but in *ʾaxad hāda l-mənfāx*, *ʿAmāra*, the former receives more prominence in that it occupies the place for topicalized objects.

To activate an unused topic, both Classical Arabic and Modern Standard Arabic employ a construction with *ʾammā...fa-*. The same function can also be filled in Modern Standard Arabic by alternative constructions with *bi-n-nisbati li-* and (more casually) *ʾammā bi-n-nisbati li-* 'with respect to', *fīmā yataʿallaqu bi-*, and *fīmā yaxtaṣṣu bi-* 'concerning'.

4. LEFT-DISLOCATION

Left-dislocation has a function similar to that of *ʾammā...fa-* and its modern counterparts, but it is sometimes regarded as the topic/comment articulation par excellence within Arabic linguistics. However, in general linguistics, apart from Chomskyan grammar, it represents a rather special type of topic articulation. Left-dislocation is one type of marked topic construction, where the dislocated element is an unused topic that is picked up again in the following sentence as a suffixed pronoun (Dahlgren 1998:90). ('Left-dislocation' is, of course, an awkward term in Arabic linguistics, since what is 'left' in Latin writing becomes 'right' in Arabic; 'predislocation' would have been a better term.) In Classical Arabic, a sentence with left-dislocation is called *jumla ḏāt wajhayni* 'a sentence with two faces' (Wright 1896:II, 256).

The left-dislocated element is generally topical with specific reference, but it may also be a generic term, as in (3).

(3) *wa-li-taʿlamū ʿadad-a*
and-so.that-know.2mp number-Acc
s-sinīna wa-l-ḥisāb
the-years.Gen and-the-reckoning
'..., and that you may know the number of the years, and the reckoning;

wa-kull-a šayʾ-in
and-every-Acc thing-Gen
faṣṣalnā-hu tafṣīlā,
distinguished.1p-cl3ms distinction
and everything We have distinguished very distinctly

wa-kull-a ʾinsān-in
and-every-Acc man-Gen
ʾalzamnā-hu ṭāʾir-a-hu
fastened.1p-cl3ms bird-Acc-his
fī raqabat-i-h

in neck-Gen-his
and every man – We have fastened to him his bird of omen to his neck' (Q. 17/12–13)

At variance with the preceding example, the left-dislocated element is probably in general characterized by reduced case marking, i.e. in the nominative even if the suffixed element has another case. Example (4), having specific reference and reduced case marking, should therefore constitute a more typical example of left-dislocation.

(4) *zayd-un* *jīʾa*
Zayd-Nom came.3ms.Pass.
ʾilay-hi *bi-kitāb-in*
to-him with-book-Gen
'A letter has been brought to Zayd' (Wright 1896:II, 256)

5. Contrastive topics

Contrastive topics are accented, a syntactic trait which otherwise characterizes the focus. In the following sentences, contrastive topics are found as nouns and as pronouns in the 1st and 3rd person: *The CHILDREN went to SCHOOL, and the PARENTS went to bed*; and *I saw Mary and John yesterday. SHE says HELLO, but HE's still ANGRY with you.*

This function may be performed in Arabic by a preposed nominal clause and *ʾammā…fa-*, as in (5).

(5) *fa-hiya* *tazʿamu* *ʾanna-hu*
and-she asserts.3fs that-he
rajul-un *qaḏir-un*
man-Nom filthy-Nom
'SHE asserted that he was a filthy man

lawwaṯa *mirʾāta-hā*
stained.3ms mirror-Acc-her
n-naqiyyat-a *bi-baṣqat-in*
the-clean-Acc with-saliva-Gen
ṣafrāʾ-a
yellow-Gen
who had stained her clean mirror with saliva

mimmā *xālaṭa-hā* *min* *baqāyā*
which mixed.3ms-cl3fs with residues
t-tibġ-i *r-raxīṣ-i*
the-tobacco-Gen the-cheap-Gen
wa-ʾāṯār-i
and-traces-Gen
that he had mixed with residues of cheap tobacco and traces of

l-balġam. *ʾammā* *huwa*
the-phlegm but he
fa-yaddaʿī *ʾanna-hu*
then-claims.3ms that-he not-
lam *yabṣuq* *ʿalā* *mirʾāt-i-hā*
Neg.Past spits.3ms on mirror-Gen-cl3fs
phlegm. But HE claimed that he did not spit at her mirror,

bal *baṣaqa* *fī* *wajh-i-hi* *huwa*
but spat.3ms in face-Gen-cl3ms he
ḥīna *ṭālaʿa-hu* *fī* *l-mirʾāt.*
when inspected.3ms-cl3ms in the-mirror
but he spat at his face when he inspected it in the mirror' (Bloch 1974:57)

6. Topicality

The character of the subject or object affects the likelihood of its being topicalized. For instance, in speaking there is a tendency to give more prominence to human participants in a discourse than to animals or things. Therefore, human participants have a higher degree of topicality than nonhumans and, hence, are more likely to occur as subject and/or to undergo topicalization. It is less likely for a person to say *A bee stung me yesterday* than to say *I was stung by a bee yesterday.* Concerning the participants in a discourse, the following hierarchies were put forward in Givón (1983) and (1977), respectively; '>' means 'more topical than'.

1. more involved > less involved participants
2. speaker (1st pers.) > listener (2nd pers.) > other (3rd pers.)

Langacker (1991) speaks of four factors that have a bearing on topicality. Firstly, the semantic role of the subject defines its degree of topicality; if it is the agent and, hence, the starting point with respect to the energy flow along the action chain, its choice as subject and topic represents the default-case option. The second factor is the 'empathy hierarchy', which is a hierarchy that reflects an egocentric assessment of the various sorts of entities that populate the world. It ranks them according to their potential to attract our empathy:

speaker > hearer > human > animal > physical object > abstract entity

The speaker ranks highest and possesses the largest likelihood of becoming a topic in this hierarachy. The third factor is degree of definiteness; an indefinite subject is commonly avoided: *There is a lake in that valley? A lake is in that valley.* He presents the following hierarchy for this factor:

definite > specific indefinite > nonspecific indefinite

The fourth factor is connected with a prominent concept within cognitive linguistics, namely the figure/ground organization; it is an almost wholly subjective factor that is not inherent in a situation but a matter of construal. It stipulates that the figure, i.e. the center of one's attention, is also the topic.

A prototypical subject ranks high with respect to all four topicality factors: it is agentive, human, definite, and the figure of a person's attention.

An investigation on topicality in the *Qurʾān* gave statistical support for different degrees of topicality (Dahlgren 2001). Subjects were divided into nonrational entities (non-R), such as things, animals, and weather phenomena; rational beings (R), i.e. humans, angels, and demons; and divine beings, represented by *ar-rabb* and *Allāh*. Based on these categories, a table on VS rate for different types of definite nouns in expository discourse was presented (Table 1).

Table 1. Topicality in the *Qurʾān*

Imperfect	non-R		R		*ar-rabb*		*Allāh*	
	Total	%VS	Total	%VS	Total	%VS	Total	%VS
Qurʾān	28	85.7	35	62.9	17	35.3	179	41.3

The figures are statistically convincing, and show that SV order, i.e. topicalization of the subject, in our context increases the more rational and powerful the subject referents are. Hence, the following hierarchy was suggested in Dahlgren (2001:35).

superhuman > human > nonhuman

This hierarchy modifies an earlier one presented by Givón (1983) as human > nonhuman.

Givón (1977) suggested a so-called Topicality Hierarchy for different types of subjects:

existential > indefinite > definite > anaphoric pronoun

Here, at variance with the preceding hierarchies, the topicality increases from left to right. In the context of Arabic, this means that SV order also increases from left to right (for statistical support for this hierarchy, → word order). Note that existentials and indefinites are hardly topical at all. The former often present a new important topic, as in *kāna malikun* 'there was a king'; the latter may do so as well, but probably less often. The anaphoric pronouns refer to topics of current interest in the discourse. This is not necessarily true of the definite subjects, which explains the higher topicality of the anaphoric pronouns.

These hierarchies are also reflected in most of the instances of topicalized objects in the *Qurʾān*. In the same investigation on topicality in the *Qurʾān* (Dahlgren 2001), eleven instances were found with topicalized objects. Some of these are presented in (6).

(6) *ṯumma baddalnā makān-a*
then changed.1p place-Acc
s-sayyiʾat-i l-ḥasanat-a
the-evil-Gen the-good-Acc
'…then We gave them in the place of evil good,

ḥattā ʿafaw wa-qālū
until multiplied.3p and-said.3p
qad massa
Perfective visited.3ms
till they multiplied, and said,

ʾābāʾ-a-nā ḍ-ḍarrāʾ-u
fathers-Acc-our the-hardship-Nom
wa-s-sarrāʾ
and-the-happiness
"Hardship and happiness visited our fathers"' (Q. 7/95)

The human object here precedes the nonhuman subject. The same occurs in example (7) with the human object in the form of a demonstrative pronoun.

(7) *ʾiḏ yaqūlu l-munāfiqūna*
when said.3ms the-hypocrites
wa-llaḏīna fī
and-Rel.p in
qulūb-i-him maraḍ
hearts-Gen-their sickness
'When the hypocrites, and those in whose hearts was sickness,

ġarra hāʾulāʾi
deluded.3ms these.ones
dīn-u-hum
religion-Nom -them
said, "Their religion has deluded them"' (Q. 8/49)

However, a nonhuman topic precedes a human nontopical subject, as in example (8) from narrative discourse.

(8) *ṯumma badā la-hum min baʿd-i*
then seemed to-them from after-Gen
mā raʾawu l-ʾāyāt
what saw the-sign.pl
'Then it seemed good to them, after they had seen the signs,

la-yasjununna ḥattā ḥīn
that-imprison.3mp.Energ until while
wa-daxala
and-entered.3ms
maʿa-hu s-sijn-a fata-yān
with-him the-prison-Acc youth-du.Nom
that they should imprison him for a while. And there entered the prison with him two youths' (Q. 12/36)

In example (9), the direct object is of more current interest in the discourse than the subject, which has not been mentioned in the actual context but brought up as a (second) topic through its being within the frame of the well-known events that surround the Day of Judgment. As explained by Lambrecht (1994:146–150), whenever two topics occur together in a clause, the predication describes the relationship between them.

(9) *wa-tarā l-mujrim-īna*
and-see.2ms the-criminal-pl.Acc
yawmaʾiḏin muqarran-īna fī
that.day fettered-pl.Acc in
l-ʾaṣfād
the-fetters
'And thou shalt see the sinners that day coupled in fetters,

sarābīl-u-hum min qaṭrān
shirts-Nom-their of pitch
wa-taġšā wujūh-a-humu n-nār
and-covers.3fs faces-Acc-their the-fire
of pitch their shirts, their faces enveloped by the Fire' (Q. 14/49–50)

BIBLIOGRAPHICAL REFERENCES

Arberry, Arthur J. 1983. *The* Koran *interpreted.* Oxford: Oxford University Press.

Bloch, Ariel. 1974. *Chrestomathy of Modern Literary Arabic.* Wiesbaden: O. Harrassowitz.

Chafe, Wallace L. 1976. "Givenness, contrastiveness, definiteness, subjects, topics, and point of view". *Subject and topic*, ed. Charles N. Li and Sandra A. Thompson, 25–55. New York and London: Academic Press.

Dahl, Östen. 1976. "What is new information?" *Approaches to word order*, ed. Nils Erik Enkvist and Viljo Kohonen, 37–50. Åbo/Turku: Åbo Akademi Forskningsinstitut.

Dahlgren, Sven-Olof. 1998. *Word order in Arabic.* Göteborg: Acta Universitatis Gothoburgensis.

——. 2001. "Word order and topicality in the *Qurʾān*". *Zeitschrift für Arabische Linguistik* 39.20–35.

Daneš, František. 1966. "A three–level approach to syntax". *Travaux Linguistiques de Prague*, I, ed. Josef Vachek, 225–240. Prague: Academia.

Fillmore, Charles J. 1982. "Frame semantics". *Linguistics in the morning calm*, ed. Linguistic Society of Korea, 111–138. Seoul: Hanshin Publishing.

Firbas, Jan. 1974. "Some aspects of the Czechoslovak approach to problems of functional sentence perspective". *Papers on functional sentence perspective*, ed. František Daneš, 11–37. Prague: Academia.

Givón, Talmy. 1977. "The drift from VSO to SVO in Biblical Hebrew: The pragmatics of tense–aspect". *Mechanisms of syntactic change*, ed. Charles N. Li, 181–254. Austin and London: University of Texas Press.

——. 1983. "Topic continuity in discourse: An introduction". *Topic continuity in discourse: A quantitative cross-language study*, ed. Talmy Givón, 5–41. Amsterdam and Philadelphia: J. Benjamins.

——. 1984. *Syntax: A functional-typological introduction*, I. Amsterdam and Philadelphia: J. Benjamins.

——. 1990. *Syntax: A functional-typological introduction*, II. Amsterdam and Philadelphia: J. Benjamins.

Grotzfeld, Heinz. 1965. *Syrisch-arabische Grammatik.* Wiesbaden: O. Harrassowitz.

Halliday, Michael A.K. 1967. "Notes on transitivity and theme in English, II". *Journal of Linguistics* 3:2.199–244.

——. 1994. *An introduction to functional grammar.* 2nd ed. London, Melbourne, and Auckland: Hodder Headline Group.

Lambrecht, Knud. 1994. *Information structure and sentence form: Topic, focus and the mental representations of discourse referents.* Cambridge: Cambridge University Press.

Langacker, Ronald W. 1987. *Foundations of cognitive grammar.* I. *Theoretical prerequisites.* Stanford, Cal.: Stanford University Press.

——. 1991. *Foundations of cognitive grammar.* II. *Descriptive application.* Stanford, Cal.: Stanford University Press.

Prince, Ellen F. 1981. "Toward a taxonomy of given – new information". *Radical pragmatics,* ed. Peter Cole, 223–255. New York and London: Academic Press.

Sanford, Anthony J. and S.C. Garrod. 1981. *Understanding written language: Explorations of comprehension beyond the sentence.* Chichester: Wiley.

Strawson, Peter F. 1964. "Identifying reference and truth values". *Theoria* 30.9b-118. (Repr., *Semantics,* ed. Danny Steinberg and Leon Jacobovits, 96–118. Cambridge: Cambridge University Press, 1971.)

Taylor, John R. 2002. *Cognitive grammar.* Oxford: Oxford University Press.

Van Valin, Robert D. and Randy J. Lapolla. 1997. *Syntax: Structure, meaning and function.* Cambridge: Cambridge University Press.

Wright, William. 1896. *A grammar of the Arabic language translated from the German of Caspari and edited with numerous additions and corrections.* 3rd ed. Cambridge: Cambridge University Press.

SVEN-OLOF DAHLGREN (Göteborg University)

Toponyms

1. INTRODUCTION

Toponyms start out as meaningful nouns or combinations of nouns, and some may conserve that meaning in spite of what becomes their primarily identifying and nondescriptive function. But often they become, in analogy to → proper names, pure 'meaningless marks'. Others acquire in the course of their linguistic history a meaning that they did not have before (popular etymology). Toponyms share two contradictory qualities. They often conserve ancient linguistic stages because these names are frequently not translated from an older language into a superseding language, which leads to the emergence of etymologically interesting linguistic fossils. At the same time, toponyms often change faster and develop more rapidly than other parts of the vocabulary, probably because they are used so intensively. For both reasons, toponyms are harder to harmonize with rules of linguistic change than other parts of the vocabulary. They also lend themselves more easily to unguarded speculation.

Toponymy is understood here as the etymological study of place-names, which has an essentially diachronic approach. Toponymy disregards topographical issues, such as the history of particular settlements or the identification of archaeological remnants with particular names. Furthermore, this entry does not attempt to draw conclusions from certain names with regard to the history of settlements or to economic or social history, nor does it dwell on the considerable influence of natural geography on place-names. The vast majority of the following examples deal with the names of villages and cities, but sometimes names of rivers and mountains are also mentioned, as are names of provinces or countries.

2. FOCUS

The focus of this overview is on place-names of what is today the Arabic-speaking world. Important areas and periods are, thereby, excluded. Spain and Portugal have preserved many toponyms of Arabic origin, which document the Muslim domination of the Iberian Peninsula (→ Ibero-Romance). In southeastern Turkey, where Mesopotamian Arabic dialects are still spoken, there are numerous place-names with an Arabic past. Place-names with an Arabic etymology exist also in Senegal, Somalia, Tanzania, Ethiopia, and Iran, and mark the Muslim conquest and cultural domination of parts of Asia and Africa far beyond the borders of the Arabic language. Especially in conjunction with Arabo-Muslim personal names, place-names such as Hyderabad or Faridpur exist in India, Pakistan, and Bangladesh. Cairo (Missouri) and Medina (North Dakota) show that place-names of Arabic origin have made it as far west as the United States. Conversely, in many parts of the Arab world, there are non-Arabic toponymic adstrates and substrates, some with an old history, some modern: in the Maghreb, there are Berber place-names; in Sudan, there are Cushitic and Nilo-Saharan place-names; in Syria and Iraq, there are Kurdish, Turkish, Modern Aramaic, and possibly Circassian place-names; in Israel and the West Bank, there are Modern Hebrew place-names; on the Arabian Peninsula and in the Persian Gulf, there are South Arabian and Persian place-names. The most important historical layers of names affecting more than one area are Ottoman Turkish names (most Arabic-speaking parts of the Ottoman Empire); Greek or Latin names (practically everywhere, except for the Arabian Peninsula); Aramaic names (Lebanon, Palestine/Israel, Syria, Iraq, Egypt); Canaanite (including Phoenician and Punic) names (Lebanon, Palestine/Israel; Maghreb); Sumerian and Akkadian names (Iraq, possibly Syria). A layer of pre-Semitic forms of present-day Arabic names has often been suggested. In the case of Mesopotamia, the existence of Sumerian pre-Akkadian names has been established beyond doubt; in other cases it is likely.

Within the realm of place-names in present-day Arabic-speaking areas, research on toponymy is very uneven, even more uneven than the research done on Arabic dialects and on languages other than Arabic. For religious reasons, there has been a heightened interest in place-names occurring in the Bible and the New Testament, in their identification and their etymology. In general, easily accessible areas are far better studied than less accessible ones. The toponymies of Lebanon, Syria, Egypt, Jordan, and Palestine/Israel have, therefore, been studied more closely than those of Sudan, Saudi Arabia, Yemen, and many areas of the Maghreb and Iraq. For Yemen, Saudi Arabia, and the Gulf States, intensive studies have begun only fairly recently.

A precondition for the diachronic study of place-names and their etymologies is the correct establishment of the present-day forms of toponyms and the tracing of earlier forms of these names. The majority of place-names in Arabic countries cannot be found in written Arabic sources prior to the 19th century. Even in comparatively well-studied areas like Lebanon, Palestine/Israel, and Syria, more than 80 percent of the names of the villages do not occur in earlier sources. In the case of a handful of more important cities and places, rivers, and valleys, earlier occurrences in Arabic sources can be traced, some of which are also found in Aramaic, Phoenician, Punic, or Hebrew texts, or in Epigraphic South Arabian inscriptions. Less frequently, Greek or Latin representations of Arabic place-names are found. Sometimes, a Lebanese, Syrian, Iraqi, or Egyptian place-name is attested in Akkadian or Ancient Egyptian, or even in Sumerian sources of the 3rd millennium B.C.E. If older forms of a toponym are known, their earlier meaning can often be identified and the name may be assigned to a specific language and a specific linguistic period. In the majority of cases, however, the present-day Arabic form is the only basis, and one has to rely on typological features, such as the existence of Aramaic morphological elements in an Arabic place-name, which may suggest a pre-Arabic etymology. Only exceptionally can it be proved that a toponym looking and sounding like a normal Arabic name is in fact not Arabic at all. The Lebanese toponym *Žbayl*, in Classical Arabic sources *Jubayl*, seems to be a diminutive of purely Arabic *jabal* 'mountain'. Yet, this name is attested as *Gubla* in Sumerian texts of the 3rd millennium B.C.E. It is unknown which language was spoken on the Lebanese coast at this time, but it can hardly have been Arabic. The triconsonantal structure makes a Semitic etymology likely, and one might opt for something like early Canaanite. Without the pre-Arabic evidence, however, *Žbayl* would have to be interpreted as a fairly recent and purely Arabic name. Some place-names attested in very old non-Arabic sources, such as *Dimašq* (Damascus) or the river *ʾUrdunn* (Jordan), do not easily yield themselves to a Semitic etymology. Especially in Iraq, the pre-Arabic names of major cities and rivers defy efforts to find Akkadian or Sumerian etymologies. A pre-Semitic and pre-Sumerian origin can therefore at least not be excluded.

3. PLACE-NAME PROVINCES

Provisionally, the usual classification of the Arabic dialects into five clusters (Versteegh 1997:145) is here extrapolated to place-names. The following place-name provinces can then be distinguished (from west to east): (i) North Africa (i.e. Morocco, Algeria, Tunisia, and Libya); (ii) Egypt and Sudan; (iii) Lebanon, Syria, Palestine/Israel, and Jordan; (iv) Iraq; (v) the Arabian Peninsula, with Kuwait, the United Arab Emirates and other Gulf States, Saudi Arabia, and Yemen. These dialect clusters have been subdivided further. They are not watertight compartments; there is considerable overlapping, and modern national borders hardly ever reflect more ancient linguistic unities. In most of these areas, important supraregional linguistic differences are found between nomadic (Bedouin) and sedentary populations, the latter subdivided into rural and urban. The term 'Bedouin' includes secondarily Bedouinized dialects of what are now rural or urban areas.

4. DIGLOSSIA

Arabic place-names are influenced by → diglossia as much as other parts of the Arabic vocabulary are. Arabic toponyms exist first as parts of the local Arabic dialects and are only secondarily represented in a form closer to Standard Arabic. In their written form, they may lose all or part of their vernacular characteristics. It is theoretically desirable but practically impossible to represent each Arabic place-name in its original dialectal form, i.e. in the phonological system of its users. Therefore, the transcribed forms used in this article reflect a fairly wide spectrum between a completely standardized and very often Classicized official Arabic form, on the one hand, and a form preserving some or most dialectal features, on the other. The Moroccan Standard Arabic name *Marrākuš* (Marrakech), for instance, is pronounced *Merrākš* by its inhabitants. The increase of political state control by centralized governments and the expansion of modern media in Arab countries go along with a growing standardization of place-names. Consonant clusters in initial position may be Classicized either by a prosthetic vowel or by the insertion of a vowel after the first consonant, e.g. *Ṭarābulus* for vernacular *Ṭrāblus* < Greek *tripolis*.

5. GENERALITIES

The Arabic article is often added to earlier non-Arabic place-names. Often, the use of the article in place-names is optional. The frequent formation of a place-name of the type determined noun plus determined adjective such as *(ad-)Dār al-Bayḍāʾ* (Casablanca) usually sheds the first article and becomes *Dar al-Bayḍāʾ*, a syntagm that is very common in most Arabic dialects and already known from Classical Arabic toponyms such as *Bayt al-Muqaddas*. The types *Wādi l-ʾAḥmar* and *Wādi l-Ḥamrāʾ* are of different origins: *Wādi l-ʾAḥmar* is either a genitive construct 'wadi of the red [house etc.]', or a construction of substantive plus attributive adjective, 'red wadi'. The formation *Wādi l-Ḥamrāʾ* can only be a genitive construction 'wadi of the red [*ʿayn* 'source', or any other feminine noun]'. When a pre-Arabic place-name is taken over by speakers of Arabic, very often it is neither completely Arabicized nor taken over in its unaltered form. In such a case, the root consonants may be partly assimilated to a cognate Arabic root: *Bājisrā* < Aramaic *bē gešrā* 'place of the bridge', but /š/ is switched to /s/ because of the Arabic word *jisr* 'bridge'. In other cases, there is a shift of meaning: Hebrew *Bēṯ Leḥem* (Bethlehem) 'place of bread' becomes Arabic *Bayt Laḥm* 'place of meat' (note the absence of the Arabic article), most probably via Aramaic *bēṯ lḥam* or *bēṯ laḥmā*.

One of the most common place-name formations is the combination of a personal name with the feminine *nisba* ending. *Muḥammadiyya, Mahdiyya*, and *ʾIskandariyya* are all connected to the name of a person, who sometimes can be historically traced. A fairly common and comparable type of place-name comprises compounds with *ʾAbū* or *ʾUmm*, marking either personal names or a looser connection of a different type: *ʾUmm al-ʾAbar* (Jordan) is a place with a number of wells, *Abū Šahrayn* (Iraq) is the name of a marketplace that is inhabited only two months per year. However, the honorific title of Mecca, *ʾUmm al-Qurā* 'Mother of the villages' seems to preserve the notion of metaphorical motherhood. Place-names starting with *Nabī, Sīdi, Mawlay, Šayx, Wāli*, and *Mār* usually mark religious sites, tombs, and places of worship; *Mār* occurs only in places connected with Christians.

The diminutive form is common in place-names. In addition to the notion of smallness, it conveys an element of personal attachment. Diminutive forms corresponding to Classical Arabic *fuʿayl, fuʿayyil*, etc. are no longer productive in most dialects. In the realm of toponyms, however, they are extremely frequent. The diminutive function is lexicalized. *Tlayl* is not 'little Tall' but 'little or dear little Tall', just as a personal name like *ʿUbayd Allāh* does not mean ‘God’s little slave’ but rather ‘little ʿAbdallāh’. Frequently, the diminutive form adds a feminine ending: *ʿWayne* ‘little ʿAyn’, *Mrayži* ‘little Marž’. The diminutive form can also affect non-Arabic names: *Ṭwayri* (Lebanon) to *Ṭūrā* < Aramaic *ṭūrā* ‘rock’, *Tḥayli* (Syria) to *taḥla* <Aramaic *taḥlā* ‘watercress’, *Tlaylīn* to *Tallīn* (Lebanon) < Aramaic *tellīn* ‘hills’.

6. *Nisba* forms

The → *nisba* forms derived from place-names show many pecularities. There are double forms such as *makkī* and *makkāwī* (from *Makka*), *ḥittī* (< *ḥadtī*) from *Ḥadat* (Lebanon), *ḥaṣbānī* from *Ḥaṣbayya* (Lebanon); double *nisba* forms ending in *-ānī* or – *āwī: ṣaydāwī* and *ṣaydānī* (both from *Ṣaydā*). The *nisba* of a compound name such as *Dayr az-Zōr* (Syria) is usually the *nisba* of one component (*dayrī*); the *nisba* of *Dayr al-Qamar* (Lebanon) is *dayrānī*. However, there are numerous less regular forms, such as *talḥamī* from *Bayt Laḥm* (Palestine/Israel), *maʿarnami* from *Maʿarrat an-Nuʿmān* (Syria), and *mestīrī* from *Mnestīr* (Tunisia).

Numerous place-names in Arabic-speaking countries carry an immediately evident Arabic meaning. In most cases, the historical context of the nomenclature is lost. The Lebanese *Mʿameltayn* (Classicized: *muʿāmalatayn*) means ‘two districts’. The function of the name becomes clear only when we know that in historical sources the original form of the place-name was recorded as *jisr al-muʿāmalatayn* ‘bridge of the two districts’. This bridge connected the two important Ottoman governorships (*muamele*) Ṣaydā and Ṭrāblūs. Some of the most common pan-Arabic place-name components are *ʿayn* ‘source’, *bāb* ‘gate’, *bayt* ‘house’, *bīr* ‘well’, *burj* ‘tower’, *dār* ‘residence’, *darb* ‘path’, *xirba* ‘deserted ruins’, *maġāra* ‘cave’, *qabr* ‘grave’, *qalʿa* ‘fortress’, *qaṣr* ‘castle’, *rās* ‘summit, top’, *sūq* ‘market’, *tall* ‘hill’, *wādī* ‘valley’, *jubb* ‘cistern’. Other names with a connotation of desert formation are: *bayāḍ* ‘uncultivated land’, *ṭalaʿ* ‘deeply incised bed [of a brook]’, *ṯuġrā* ‘narrow pass’, *sahla* ‘plain’, *sayl* ‘torrent’, *šiʿb* ‘mountainous path’, *ʿirāq* ‘cave, steep cave’, *ʿarab* ‘grazing ground of a tribe’, *waʿra* ‘trackless, barely accessible area’.

7. Regional particularities

7.1 *Maghreb*

Arabic place-names coexist with Berber names and some remnants of colonial (Portuguese, Spanish, and French) domination such as *Casabranca* (Portuguese) = *Casablanca* (Spanish), which was translated into Arabic as *(ad-)Dār al-Bayḍā*ʾ (Morocco). *Colomb-Béchar* (Algeria) is a mixed French-Arabic place-name containing the name of a French colonial officer Colomb and the name of the Wādī Baššār.

The difference between dialect forms and standardized Arabic forms can be considerable. The names of the three most important cities in Algeria are *al-Jazāʾir* (French *Alger*, Algiers) < Arabic *al-jazāʾir* ‘islands’, *Wahrān* (French *Oran*) < Berber *wa-iharan* ‘place of lions’ (?) and *Qusṭanṭīna* or *Qusanṭīna* (French *Constantine*) < Latin *Constantina*. They show the three most important linguistic layers, Arabic, Berber, and Greek/Latin. There are very few names which can with certainty be derived from Punic or Numidian forms. A name of Punic origin is Tunisian *Qarṭāj*, which reflects Latin *Carthago*, which in turn is a compound, the first part of which is Phoenician *qart* ‘village’; the second part is assumed to represent a Phoenician or Punic form *hadast* ‘new’. Tunisian *Mnestīr* and *Mestīr*, Classicized *Munasṭīr* goes back to Greek *monastērion* or Latin *monasterium*; Libyan *Ṭrāblūs* (*al-Ġarb*) is like its Lebanese counterpart *Ṭrāblūs* (*aš-Šarq*), derived from Greek *Tripolis*. Berber place-names or elements of place-names are *tizi* ‘mountain pass’, *adrar* ‘mountains’, and *taurirt* ‘hill’. *Taurirt* also shows the characteristic morpheme combination of Berber feminine nouns *ta-…-t*. The oasis *Biskra* (Algeria) is a remnant of Latin *vescera*, a *limes* post. Often, there are double forms, one Berber, the other Arabic; *Malila/Melilla* (Morocco) has an Arabic form *Mlilya*, and a Berber form *Tamlilt* ‘the white one’. Berber names are also *agadir* ‘fortified enclosure’, e.g. *Agadir – Ighir* (Morocco), or *Azammur* (Morocco) < Berber *azemmur* ‘wild olive tree’; *Rabāṭ* (Morocco) also *Ribāṭ al-Fatḥ* is a ‘fortified military outpost’.

7.2 *Egypt and Sudan*

Some names can be traced back via Coptic forms to Ancient Egyptian ones: *Būṣīr* (hypercorrectly *ʾAbū Ṣīr*) is most probably Coptic **p-usir* < Hieroglyphic *pr-wsir* 'house of Osiris'; *Damanhūr* < Coptic *tmenhor* < Hieroglyphic **dmy-n-hr* 'place of Horus'; *al-Fayyūm* and *al-Bayyūm*, both < Coptic *feyom*, Hieroglyphic *pʾ-ym* 'lake, lake land'; *Šubra* < Coptic *sebro* 'hamlet'; *Qift* < Coptic *kepto, kebto* 'Egyptian, Copt'; *Atfih* < Coptic *pe-tpeh*, Hieroglyphic *pr-tp-ih.w* 'house of the cow-headed Hathor'.

Of Greek origin is *Būnumrus* (hypercorrectly *ʾAbū Numrus*) < Greek *ponmoros* [proper name], and other names ending in *-us: Banāyūs* < Greek *panayos* [proper name], *Burullus* < Greek *páralos* 'situated at the seacoast'. Other names of Greek origin do not show the Greek or Latin endings: *al-Burunbul* < Greek *parembolē* 'military camp'; *al-Fusṭāṭ* < Latin *fossatum*, Greek *fossáton* 'ditch'; *al-Qulzum* < Greek *klúsma* 'sluice'.

Frequently, *al-* in first position of an originally Latin or Greek name is reanalyzed as the /l/ of the Arabic article and is not felt to be part of the name. Greek *Alexandros* becomes *al-ʾIskandar* and can shed what has become the article to be shortened to *ʾIskandar* which gives us *(al)-ʾIskandariyya* 'Alexandria'.

Specifically Egyptian place-name elements are: *ʿizba* 'country estate, farm', *kōm* 'hill', *najʿ* 'village originally settled by Bedouin', *maḥalla* 'resting place' (= Ottoman *mahalle*?). The name *Miṣr* 'Egypt' looks like Arabic *miṣr* 'armed encampment, military settlement'; the Fatimid foundation *Miṣr al-Qāhira* 'Victorious Misr' became the name of the Egyptian capital. But the name is pre-Arabic: Hebrew *Miẓrayim* looks like a dual 'the two Miẓrs', i.e. Upper and Lower Egypt, Aramaic *meṣrīn*.

7.3 *Lebanon, Syria, Jordan, and Palestine/Israel*

Pre-Arabic names are relatively frequent in this area. Numerous place-names preserve the form of Aramaic names. Place-names ending in *-a* may represent an Aramaic masculine *status emphaticus (Dayr) Kīfā* (Lebanon) < Aramaic *kēp̄ā* 'rock'; *Kfartaʿla* (Lebanon) <Aramaic *kp̄ar taʿlā* 'village of the fox'; *Ṭurzā* (Lebanon) < Aramaic *ṭūr arzā* (with haplological elision) 'mountain of the cedar'. Place-names ending in *-tā* show the ending of the Aramaic emphatic feminine singular: *Ḥazzirtā* (Syria) < Aramaic *ḥazzūrtā* 'apple tree'; *Ḥrabtā* (Syria) < Aramaic *ḥrabtā* 'ruined place'; *ʿAyn Tantā* (Lebanon) < Aramaic *ʿayn tēntā* 'source of the fig tree'; *Biʿwirtā* (Lebanon) < Aramaic *bē ʿb̲artā* 'place of the ford'. The ending *-īn* often indicates a name with the ending of the Aramaic absolute masculine plural: *Dibbīn* (Lebanon) < Aramaic *debbīn* 'bears'; *Rišdibbīn* (Lebanon) < Aramaic *rēš debbīn* 'summit of bears'; *ʿAyn Ṭūrīn* (Syria) < Aramaic *ʿayn ṭūrīn* 'source of mountains'. The West Aramaic emphatic plural ending *-ayyā* is preserved in: *Kifrayyā* (Lebanon) < Aramaic *kap̄rayyā* 'villages'; *Rišmayyā* (Lebanon) < Aramaic *rēš mayyā* 'top of the water'; *Ṭūrzayyā* (Lebanon) < Aramaic *ṭūr arzayyā* 'mountain of cedars' (with haplological elision). A competing form of the Aramaic emphatic masculine plural is found in the ending *-āya* (in Maʿlūla: *-ōya*), as in *Qibrāya* (Syria) < Aramaic *qab̲rāya* 'graves'. The Aramaic emphatic feminine plural ending *-āṯā* can be found in *Bāʿaynāṯā* (Yāqūt, *Muʿjam* I, 472) < Aramaic *bē ʿaynāṯā* 'place of springs'. Where the Arabic dialects do not have the spirans /ṯ/, it is replaced by /t/, cf. *ʿAynāta* (Lebanon) < Aramaic *ʿaynāṯā* 'sources, springs'. Aramaic spirantic variants of /b/, /g/, /d/, /k/, /p/, /t/ are sometimes preserved: *Xirxayyā* (Lebanon) < Aramaic *karkayyā* 'villages'.

Place-names of the form *ʾifʿa/il* preserve the Aramaic masculine *status absolutus*, such as *ʾIdlib* (Palestine/Israel) < Aramaic *dlub̲* 'plane tree'; *ʾIzraʿ* (Lebanon) < Aramaic *zraʿ* 'seed'. The Aramaic masculine construct state is preserved in the numerous Lebanese village names starting with *Kfar*, such as *Kfarbadda/i* < Aramaic *kp̄ar baddā* 'village of the wine press'; *Kfartibnīn* < Aramaic *kp̄ar tibnīn* 'village of straw'; *Kfarḥāta* < Aramaic *kp̄ar ḥāṯā* 'village of the sisters'. In many cases, there is a parallel form with the more literary Arabic *kafr* instead of *kfar* as first element. In Jordan, this place-name element is pronounced *kufr*.

A frequent phenomenon is the shortening of Aramaic *bēṯ* 'house, place of' to *bā-* (mainly in Mesopotamia) and to vowelless *b-* (in Syrian and Lebanese place-names): *Bṭūrrām* (Lebanon) < Aramaic *bē ṭūr rām* 'place of the high mountain'; *Btiḥlīn* (Lebanon) < Aramaic *bē ṯaḥlīn* 'place of watercress'. This element occurs also with an Arabic second part, e.g. *Bḥamdūn* (Lebanon) < Aramaic *bē Ḥamdūn* 'place of Hamdun'.

There are some rare names showing pre-Aramaic Canaanite etymology, e.g. *Bayrūt*

(Lebanon) < Phoenician *bērōt* ‘wells’; *Rāmūṭ* (Lebanon) < Phoenician *rāmōt* ‘hills’ (with secondary velarization of the final /t/); *Qaryūt* (Israel/Palestine) < Canaanite *qaryōt* ‘villages’. In Palestine, numerous Hebrew names were preserved in Arabic toponymy, e.g. *ʿAkkā* < Hebrew *ʿAkkō; ʿAsqalān* < Hebrew *Ašqelōn; Baysān* < Hebrew *Bēṯ Šʾān; Bayt Laḥm* < Hebrew *Bēṯ Leḥem; Bīr as-Sabʿ* < Hebrew *Bʾēr Šeḇaʿ; Ġazza* < Hebrew *ʿAzza; Ḥayfā* < Talmudic Hebrew *Ḥēp̄ā; Ludd* < Hebrew *Lōd; Arīḥā* < Hebrew *Yerē/īḥō; Yāfa* < Hebrew *Yāp̄ō*.

Greek names are frequent: *Ṭrāblūs* (Lebanon) < Greek *tripolis; Nāblus* (Israel/Palestine) < Greek *neapolis, Furqlus* (Syria) < Greek *proklos, Blāṭunus* (Syria) < Greek *plátanos*. The Greek ending *-ia* is preserved in place-names such as *Sūriyā* (Syria) < Greek *Súria*. This ending *-iya* is fused with the common Arabic feminine *nisba* ending *-iyya*, very common in Arabic place-names, which gives us *Sūriyya*. *Filasṭīn* is directly derived from Greek *palaistíne*, Latin *palaestina*. The Greek form reflects a Hebrew *plešet* with an unknown etymology. *ʿAmmān* (Jordan) derives from Hebrew (*rabbat*) *ʿammōn*.

An interesting rarity is the name of a Jordanian/Iraqi border point written in Arabic <ʾl-jfwr>. This looks like an Arabic broken plural but is in reality the former pumping station no. 4 of what was the Haifa line of the Iraq Petroleum Company, in English abbreviated as ‘H 4’. This abbreviation was interpreted by Bedouin speakers as *ič-čfūr*. It makes one wonder what may lie behind other seemingly inobtrusive place-names.

7.4 *Iraq*

Iraqi place-names conserve some very old pre-Arabic toponyms. For most of the oldest place-names, which go back to the 3rd or 2nd millennium B.C.E., it is uncertain whether an Akkadian or Sumerian etymology can be assumed, or whether these place-names belong to an even older linguistic layer.

Bābil reflects Akkadian *Babilum*; the etymology *bab ilim* ‘gate of God’ may be an Akkadian popular etymology; *Niffar* is *Nippuru* in Akkadian, without known etymology; *Hīt* is Akkadian *Ita*, probably < Sumerian *id* ‘river’; *Warka* is *Uruk* in Akkadian, *Unug* in Sumerian, *ʾEreḵ* in Biblical Hebrew tradition. The names of the important rivers are all without an ascertained etymology, for instance the Tigris, Arabic *Dijla* < Akkadian *Idiglat*, Sumerian *Idigina*, Biblical Hebrew *Hiddeqel*; the Euphrates, Arabic *al-Furāt*, is *Purattum* in Akkadian.

Clearly Aramaic are names such as *al-Karx* < Aramaic *kraḵ* ‘town’; *Talkīf* < Aramaic *tel kēp̄* ‘hill of the rock’; *Barāta* < Aramaic *brāṯā* ~ *brōṯā* ‘cypress’; *ʿAbarta* < Aramaic *ʿḇartā* ‘ford’. Compounds with *bā-* ‘house, place of…’ are frequent, e.g. *Baʿqūba*, probably < Aramaic *bā Yaʿqōḇā* ‘house of Jacob’; *Bāʿaynāṯā* < Aramaic *bā ʿaynāṯā* ‘house of the wells’. Of Greek origin is *Ṭaysafūn* < *Ktesiphōn, Quṭrabbul* < *Nikatoropolis, Sulūqiyya* < *Seleukía*. *Kirkūk* is attested in Syriac sources as *Karkā ḏ-Bēṯ Selōḵ* ‘Karka of the Seleucids’.

There are numerous place-names in Iraq of Kurdish and Turkish origin.

7.5 *Arabian Peninsula*

Many of the names used are immediately comprehensible to the present-day Arab speaker; *Riyāḍ*, for instance, is a plural of *rawḍa* ‘garden’. The meaning of *Jidda*, Classical Arabic *Judda*, is obscure. In a surprisingly large number of cases, the etymology is not clear. *Makka* (Mecca), one of the few place-names mentioned in the *Qurʾān* (Q. 48/24, with the variant form *Bakka*, Q. 3/96), is etymologically obscure; so is *Yaṯrib* (Q. 33/13, also mentioned in Minaean inscriptions as *yṯrb*), which, after the Muslim takeover, was named *al-Madīna* ‘the city’ (Q. 9/101, 120; 63/8); this was later interpreted as *madīnat an-nabī* ‘the city of the Prophet’. *Ṭūr Sinīn* (Q. 95/2), *aṭ-Ṭūr*, and *Saynāʾ* (Q. 23/20) all refer to *Sinay*, the biblical Mount Sinai. An entire *sūra* (34) bears the name *Sabā*, biblical *Sba* (Sheba), also mentioned in Q. 27/22, which is a place-name with a South Arabian etymology. Specialists assert the ‘striking continuity’ of Epigraphic South Arabian, Sabaean, and Minaean place-names such as *Ḥaḍramawt* (*ḥḍrmwt*), *Mārib* (*mryb*, which may reflect *Maryab*), *Damār* (*dmr*), *Najrān* (*njrn*), and *Šibām* (*sbm, sbmm*).

8. Concluding remarks

Arabic place-names played an important stylistic role in classical Arabic poetry. The beginning verse of what is probably the best known pre-Islamic poem, *qifā nabki min ḏikrā ḥabībin* *wa-manzilī / bi-siqṭi l-liwā bayna d-Daxūli wa-Ḥawmalī* 'Let us stop, O my two friends, and weep at the memory of a loved one and a dwelling place between al-Dakhul and Hawmal' (Imruʾ al-Qays, *Muʿallaqa*, v. 1) shows two Arabic toponyms of the Arabian Peninsula. Here as elsewhere, such place-names give a spatial setting and a local color to the verse, a feature enormously popular with the Bedouin tribes. The names of camping sites, sweetwater wells, grazing grounds, mountains, desert formations, and dry river courses gave Bedouin poetry the flavor of tribal realism, just as did the mentioning of different tribes and clans. This tradition of Bedouin poetical toponymy later became a cliché in the poetry favored by the courts of the urban centers of Damascus and Baghdad. It was used by poets who described an imaginary journey on a camel they had never ridden and mentioned places in a desert they had never seen.

In politically contested areas, Arabic place-names can become a loaded issue. After the establishment of the state of Israel (1948), the Palestinian exodus, and the June war (1967) with the occupation of the West Bank and Ghaza and the subsequent annexation of the Golan Heights, the names of many Arabic villages disappeared with the villages, while others were systematically renamed with Hebrew or Hebraized names. In present-day Turkey, place-names of Arabic origin tend to be replaced by Turkish toponyms, as do names of Kurdish or Aramaic origin. The official use of Arabic place-names is frowned upon in the Iranian Arabic-speaking province of Khuzestan.

Arabic place-names created an unexpected flurry in the international press in 1985, when the Lebanese historian Kamal Salibi tried to prove that the Old Testament and Judaism came from West Arabia and not from Palestine. He based his evidence mainly on biblical place-names, which in his opinion were in reality Arabic place-names from the southern part of the Ḥijāz and the province ʿAsīr. This hypothesis never gained general acceptance.

Bibliographical references

Primary source

Yāqūt, *Muʿjam* = Yāqūt ibn ʿAbdallāh ar-Rūmī, *Muʿjam al-buldān*. Ed. Ferdinand Wüstenfeld. 6 vols. Leipzig, 1866–1870.

Secondary sources

Al-Hilou, Abdullah. 1986. *Topographische Namen des syro-palästinischen Raumes nach arabischen Geographen: Historische und etymologische Untersuchungen*. Ph.D. diss., University of Berlin.

Al-Maʿani, Sultan. 1992. *Nordjordanische Ortsnamen: Eine etymologische und semantische Untersuchung*. Hildesheim: G. Olms.

Al-Scheiba, Abdallah Hassan. 1982. *Die Ortsnamen in den altsüdarabischen Inschriften (mit dem Versuch ihrer Identifizierung und Lokalisierung)*. Ph.D. diss., University of Marburg.

ʿArrāf, Šukrī. 2004. *al-Mawāqiʿ al-juġrāfiyya fī Filasṭīn: al-ʾAsmāʾ al-ʿarabiyya wa-t-tasmiya al-ʿibriyya*. Beirut: Institute of Palestine Studies.

Bayley, Clinton. 1984. "Bedouin place-names in Sinai". *Palestine Exploration Quarterly* 116.42–57.

Björnesjö, Sophia. 1996. "Quelques réflexions sur l'apport de l'arabe dans la toponymie égyptienne". *Annis* 30.21–40.

Czapkiewicz, Andrzej. 1971. *Ancient Egyptian and Coptic elements in the toponymy of contemporary Egypt*. (= *Universitatis Iagellonica Acta Scientiarum Litterarumque*, 225, *Schedae Grammaticae*, Fasciculus 28.) Cracow: University of Cracow.

Fischer, Wolfdietrich. 1980. "Der Beitrag der Araber zur Ortsnamengebung im Vorderen Orient". *Beiträge zur Namenforschung*, N.S. 18.27–31.

Frayḥa, ʾAnīs. 1972. *Muʿjam ʾasmāʾ al-mudun wa-l-qurā al-lubnāniyya wa-tafsīr maʿānīhā (A dictionary of the names of towns and villages in Lebanon)*. Beirut: American University of Beirut.

Groom, Nigel. 1983. *A dictionary of Arabic topography and placenames: A transliterated Arabic-English dictionary with an Arabic glossary of topographical words and placenames*. Beirut: Librairie du Liban and London: Longman.

Harding, Lankester G. 1971. *Index and concordance of pre-Islamic Arabian names and inscriptions*. Toronto: University of Toronto Press.

Helck, Wolfgang and Eberhard Otto. 1975–1992. *Lexikon der Ägyptologie*. 7 vols. Wiesbaden: O. Harrassowitz.

Isserlin, Benedikt. 1989. "Arabic place name types of the Islamic period: Their derivation and distribution". *al-Masāq* 2.9–24.

Knauf, Ernst A. 1988. "The West-Arabian place name province: Its origin and significance". *Proceedings of the Seminar for Arabian Studies* 18.39–49.

Krotkoff, Georg. 1972. "Bagdader Studien. VI. Ein toponymisches Kuriosum". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 122.100.

Kuentz, Charles. 1951. "Stratification de l'onomastique égyptienne actuelle". *Actes et Mélanges du Congrès Internationale de la Toponymie et Anthroponymie*, 292ff. Louvain.

Limet, Henri. 1975. "Permanence et changement dans la toponymie de la Mésopotamie antique". *La toponymie antique: Actes du Colloque de Strasbourg, 12–14 juin 1975*, 83–115. Leiden: E.J. Brill.

Littmann, Enno. 1939. "Unsemitische Ortsnamen in Nordsyrien". *Anatolian studies presented to William Hepburn Buckler*, ed. William Moir Calder and Josef Keil, 151–160. Manchester: Manchester University Press.

Matthews, Charles D. "Non-Arabic place names in Central South Arabia". *Akten des 24. Internationalen Orientalistenkongresses in München, 28. August bis 4. September 1957*, ed. Herbert Franke, 259–262. Wiesbaden: F. Steiner.

Ramzī, Muḥammad. 1954–1960. *al-Qāmūs al-juġrāfī li-l-bilād al-miṣriyya min ʿahd qudamāʾ al-miṣriyyīna ʾilā sana 1945*. Cairo: Dār al-Kutub. (Repr., Cairo: al-Hayʾa al-Miṣriyya al-ʿĀmma li-l-Kitāb, 1993.)

Reitemeyer, Else. 1912. *Die Städtegründungen der Araber im Islam nach den arabischen Historikern*. Leipzig.

Salibi, Kamal. 1985. *The Bible came from Arabia: Radical reinterpretation of Old Testament geography*. London: J. Cape.

Sznycer, Maurice. 1977. "Recherches sur les toponymes phéniciens en méditerranée occidentale". *La toponymie antique: Actes du Colloque de Strasbourg, 12–14 juin 1975*, 163–175. Leiden: E.J. Brill.

Thilo, Ulrich. 1958. *Die Ortsnamen in der altarabischen Poesie*. Wiesbaden: O. Harrassowitz.

Wardini, Elie. 2003. *Lebanese place-names (Mount Lebanon and North Lebanon): A typology of regional variation and continuity*. Louvain: Peeters.

Wild, Stefan. 1973. *Libanesische Ortsnamen: Typologie und Deutung*. Beirut: F. Steiner.

——. 1977. "Zu aramäischen Ortsnamen in Palästina". *La toponymie antique: Actes du Colloque de Strasbourg, 12–14 juin 1975*, 65–74. Leiden: E.J. Brill.

Xammār, Qustanṭīn. 1980. *ʾAsmāʾ al-ʾamākin wa-l-mawāḍiʿ wa-l-maʿālim aṭ-ṭabīʿiyya wa-l-bašariyya wa-l-juġrāfiyya al-maʿrūfa fī Filasṭīn ḥattā l-ʿām 1948*. Beirut: al-Muʿassasa al-ʿArabiyya li-d-Dirāsāt wa-n-Našr.

Zadok, Ran. 1997. "A preliminary analysis of ancient survivals in modern Palestinian toponymy". *Mediterranean Language Review* 9.93–171.

Stefan Wild (University of Bonn)

Transcription

Transcription is usually defined as "the operation of representing the elements of a language, either sounds or signs, however they may be written originally, in any other system of either sounds or signs" (Wellisch 1975). As opposed to transliteration, the operation of representing the graphemic units of a language's writing system using another writing system, most transcription systems for Arabic attempt to represent the Arabic original more or less phonetically, including unwritten vowels, and reducing archaic orthographic forms such as عمرو *ʿAmr* to their phonetic value.

Two types of transcription of Arabic are distinguished here: ad hoc transcription, which represents Arabic in Latin script without a defined system, and scientific transcription, which attempts to define a systematic, non-ambiguous mapping from Arabic script or speech to Latin script, possibly using diacritic signs, in order to allow for the precise reconstruction of the Arabic original.

1. Ad hoc transcription

Ad hoc transcription has a wide range of everyday applications in the Arab world and in the West, wherever Arabic has to be represented in Latin script. Examples include e-mail communication, Arabic personal and brand names, geographical designations, and generally most fields where there is no need for or feasibility of systematic representation.

In general, ad hoc transcription uses idiosyncratic and sometimes inconsequential orthographic rules, often from another language, to represent Arabic without diacritic symbols, thus rendering, for example, ش as *sh* (English), *ch* (French), or *sch* (German). Several phonetic features of Arabic cannot easily be represented in this way, such as emphatic consonants or the letter ع; the former are sometimes represented by digraphic combinations with *h* (الرياض *ar-Riyadh*), the latter is usually either not represented or is indicated by an apostrophe or vowel (عيد *Eid*). Vowels are represented using the host language's inventory, leading to forms such as القاعدة *al-Qaeda* in English texts, while French or German authors prefer *al-Qaida* (the latter sometimes even *el-Kaida*). Vowel quantity is usually represented by double vowels (الجزيرة *al-Jazeera*), accented characters, especially in French-based orthographies, or not at all. Underlying forms are often chosen from dialects rather than Standard Arabic, which further affects the consonant and vowel inventories.

As a result, ad hoc transcriptions are often ambiguous and may vary considerably, depending on the host language, the underlying form, and other factors. There exist, for example, more than thirty documented variants of the name معمر القذافي, spelling the last name as *Qadhdhafi, Gaddafi, Kadafi*, etc., and making it impossible to reconstruct the original Arabic from the Latin spelling. This problem is addressed by a number of different standardization efforts for special applications such as geographic names.

A special case of ad hoc transcription is the system used in electronic communication, such as e-mail or text messages, where the communication channel is often restricted to a limited character set. Within these environments, some Arabic letters are commonly represented by numbers and apostrophes, based on graphemic similarity, as in Table 1, leading to forms such as *al7amdolellah 3alaa ssalaamah* (الحمد لله على السلامة).

Table 1. E-mail 'graphemic' transcription of Arabic

Arabic letter	Transcription
ء	2
ع	3
غ	3'
ط	6 [rare]
ح	7
خ	7'

2. Scientific transcription

There is no intrinsic necessity of transcribing Arabic for scientific purposes, as typesetting Arabic itself has been perfectly possible for centuries and the readership usually reads Arabic. In fact, in early grammars of Arabic, we often find Arabic script used without transcription, such as Erpenius' Latin-language grammar of 1613 and de Sacy's French-language grammar of 1805. Yet, Lumsden's English-language grammar of 1813 sometimes uses an ad hoc scheme for phonetic representation, and Savary's grammar of 1788 actually introduces a systematic transcription scheme based on French orthography. Looking at early dictionaries, vocalized Arabic script without transcription is found both in Golius' dictionary of 1653 and Freytag's of 1837. An exception is Mesgnien-Meninski's Arabic/Persian/Turkish dictionary of 1780, where the use of transcription was probably necessitated by the inclusion of Turkish words whose pronunciation is not always evident from the Arabic script.

The first transcription system for Oriental languages to find wide usage was probably Sir William Jones' system for Indian languages within the Royal Asiatic Society of Bengal (Jones 1788; his suggestions for Arabic were considerably less successful). In the course of the 19th century, the development of linguistics as a science emerged in a number of transcription systems, and 'universal alphabets' for general and specialized linguistic description were subjects of considerable discussion in the Orientalist community (see, e.g., Whitney 1862). For Arabic, such systems were adopted rather slowly; for quite a long time, 19th-century journals usually either printed Arabic directly or used ad hoc transcription along with journal-specific transcription schemes. The scientific community appears to have viewed this situation as increasingly unsatisfactory, and several working groups were founded in the 1890s to address the 'transcription question'. A common alphabet for transcribing Arabic, along with other languages, was presented at the Geneva Congress of Orientalists in 1896 (Socin 1895; Burgess 1897); however, transcription systems continued to be somewhat undefined well into the 20th century, even though the central features were established by this time (notably, the dotted notation for emphatic consonants and the use of macron or circumflex for long vowels). A final standardization attempt was made at the 1935 Rome Congress of Orientalists, where a system developed by the Deutsche Morgenländische Gesellschaft was presented and adopted as a congress recommendation (Brockelmann a.o. 1935).

Scientific transcription is, essentially, a hybrid system, based on phonetic transcription of Classical or Modern Standard Arabic with some elements of transliteration and morphophonemic representation. Brockelmann a.o. (1935:4) clearly attempt to classify it as transliteration but are then forced to introduce a large number of special cases and exceptions. Two basic modes of transcription exist: fully vocalized transcription, which attempts to render the original with inflectional endings, *tanwīn*, etc. (regardless of whether these are written in the text or not), and transcription in pausal form. Fully vocalized transcription is used when quoting longer passages of Classical or Qurʾānic text as well as poetry. When transcribing passages of Modern Standard Arabic prose, the *ʾiʿrāb* may be largely omitted, just as a speaker would do; proper names, isolated words, and references are usually quoted in pausal form as well.

Today, three main transcription systems are used in the scientific community: an English system, often called the Library of Congress system (abbreviated LC or ALA-LC); a French system; and the Deutsche Morgenländische Gesellschaft (DMG) system of 1935. The DMG system is the one used, with some minor modifications, in the present encyclopedia. A Cyrillic system modeled on the DMG recommendations exists as well. In addition, the *Encyclopaedia of Islam* uses its own transcription system, dating back to 19th-century conventions used in its first edition. The systems are largely similar and vary mainly with the orthography of the underlying languages.

In addition, there are several national and international standards for transcribing Arabic, such as DIN 31635, ISO 233, BS 4280, and the UN system of 1972 for geographic designations on maps. As these are either not easily available, not in general use, or very close to one of the aforementioned scientific transcription systems, they are not described here in detail.

In the following, a rough guide to transcription is given, based mainly on common usage of the Library of Congress and the Deutsche Morgenländische Gesellschaft systems. It attempts to cover most, but not all, the details. The normative works are Brockelmann a.o. (1935) and Berry (1997); a practical guide to DMG transcription is Reichmuth and Schielke (2000). For the sake of completeness, the symbol charts in Tables 2, 3, and 4 include symbols from a number of other transcription systems as well.

It is not uncommon to use the LC consonants without diacritic signs, especially in nonlinguistic publications; however, the transcription results may become increasingly ambiguous. Even with diacritics, the use of digraphs in the LC and French systems may lead to ambiguities (such as *kh* for ـخـ and ـكه) and awkward forms

Table 2. Arabic transcription systems (consonants)

Arabic	EALL	LC	DMG	French	EI	Cyrillic	Other signs
ا, ء	ʾ	ʾ	ʾ	ʾ	ʾ	ʾ	’, ʔ
ب	b	b	b	b	b	б	
ت	t	t	t	t	t	т	
ث	ṯ	th	ṯ	ṯ	th	с̱, т̱	θ
ج	j	j	ǧ	dj	d̲j̲	дж (جّ: джж)	
ح	ḥ	ḥ	ḥ	ḥ	ḥ	х̣	
خ	x	kh	ḫ	kh	k̲h̲	х̮, х̱	χ
د	d	d	d	d	d	д	
ذ	ḏ	dh	ḏ	dh	d̲h̲	з̱, д̱	ð
ر	r	r	r	r	r	р	
ز	z	z	z	z	z	з	
س	s	s	s	s	s	с	
ش	š	sh	š	ch	s̲h̲	ш	ʃ
ص	ṣ	ṣ	ṣ	ṣ	ṣ	с̣	
ض	ḍ	ḍ	ḍ	ḍ	ḍ	д̣	
ط	ṭ	ṭ	ṭ	ṭ	ṭ	т̣	
ظ	ḏ̣ (in proper names ẓ)	ẓ	ẓ	ẓ	ẓ	з̣	
ع	ʿ	ʿ	ʿ	ʿ	ʿ	ʿ	ع, ɛ, ʕ
غ	ġ	gh	ġ	gh	g̲h̲	г̇, ғ	ɣ
ف	f	f	f	f	f	ф	
ق	q	q	q	q	ḳ	к̣	
ك	k	k	k	k	k	к	
ل	l	l	l	l	l	л	
م	m	m	m	m	m	м	
ن	n	n	n	n	n	н	
ه	h	h	h	h	h	х	
و	w	w	w	w	w	в	
ي	y	y	y, j	y	y	й	

Table 3. Arabic transcription systems (vowels, diphthongs)

Arabic	LC, EALL, EI	DMG	French	Cyrillic
ا	ā	ā	â	ā
و	ū	ū	û	ӯ
ي	ī	ī	î	ӣ
ـَ	a	a	a	а
ـُ	u	u	u	у
ـِ	i	i	i	и
ـَيْ	ay	ay, ai, aj	ay	ай
ـَوْ	aw	aw, au	aw	ав
ـِيّ	iyy/īy	iyy/īy, ijj/īj	iyy/īy	ийй/ӣй
ـُوّ	uww/ūw	uww/ūw, ijj/īj	uww/ūw	увв/ӯв

Table 4. Other Arabic marks

Arabic	Transcription
ة (*ṭāʾ marbūṭa*)	In fully vocalized transcription, always rendered as *-at-* or *-āt-*; in pausal form, three cases are distinguished: *i*. *-a* or *-ah* in isolated words: مدينة *madīna, madīnah* *ii*. *-āh* or *-āt*, when preceded by *ā*: صلاة *ṣalāh, ṣalāt* *iii*. *-at* in *status constructus* in the head word of a genitive compound: مدينة النبي *madīnat an-nabī*
آ (*madda*)	Treated as *hamza* + *ʾalif*, where it represents this combination; otherwise, no special treatment
ـّ (*šadda*)	Treated as double consonant: شدّة *šadda*
ـْ (*sukūn*)	Not represented at all
ـٌ ، ـٍ ، ـًا (*tanwīn*)	Treated as vowel + *n*
ـٰ (dagger *ʾalif*) etc.	Treated as long vowel, without orthographic representation: الرحمٰن *ar-raḥmān*

such as مخ *mukhkh*. To address this problem, the signs of the International Phonetic Association (IPA) ɣ, ð, θ, χ are sometimes used for *gh, dh, th, kh*. In Russian publications, the diacritic signs are often omitted as well.

The LC recommendation of *á* for ى (*ʾalif maqṣūra*) is rarely followed.

Most systems allow for a certain amount of choice. The choice between *-a* and *-ah* for *tāʾ marbūṭa*, for example, is a matter of convention (the EALL uses *-a; -at* in genitive compounds is used only before a vowel, thus *madīnat an-nabī*, but *maktaba Bayrūt*). Many authors also use one of the variants *-īy/-ūw* vs. *-iyy/-uww* consistently, even though these two variants may express a morphological difference: قووم *qūwima* (Form III) vs. قوّم *quwwima* (Form II; this variant is probably phonetically closer to Arabic in the general case). In most of these and the optional cases outlined below, scientific communities, journals, and other groups tend to develop and follow certain common styles; as long as a given decision is followed consistently, however, the final choice is ultimately a matter of the author's preference.

The Arabic article *al-* can be transcribed in assimilated form before the *ḥurūf šamsiyya*: الشمس *aš-šams* vs. القمر *al-qamar*. EI and LC convention is to use the unassimilated form: *al-shams*, while the EALL uses the assimilated form. The article, as well as the particles *wa-, fa-, bi-, li-, ka-, ʾa-*, etc. are written separated from the word by a hyphen. For the sake of precision, enclitic pronouns can be separated, too: أبوه *ʾabū-hu* vs. *ʾabūhu*.

Hamza and *waṣla* are functionally different and thus treated differently. *Hamza* is always transliterated as ʾ except in word-initial position, where some systems omit it (أئمة *ʾaʾimma* vs. *aʾimma*). *Hamza* occurs in word-initial position if it results from a prefix, from the word's root itself, such as أخذ *ʾaxaḏa*, and in some loanwords, such as الإسكندرية *al-ʾIskandariyya*.

Waṣla is the sign ٱ for an auxiliary vowel being inserted at the beginning of a word where it would otherwise begin with a consonant cluster; typical cases are the definite article *al*-, verbal Forms VII through XV (but not IV), and some words beginning with a consonant cluster, such as ٱمرأة *imraʾa* or ٱبن *ibn*. If the preceding word ends in a vowel, this auxiliary vowel is omitted: قال ٱلرئيس *qāla r-raʾīsu*, بٱنتفاضتهم *bi-ntifāḍatihim*. Consequently, *waṣla* is not transcribed at all after a preceding vowel, otherwise, the auxiliary vowel is used to transcribe *waṣla*: ٱستفاد *istafāda* (but: أستفيد *ʾastafīdu* with morphological *hamza* from the prefix), ٱلمجمع *al-mujammaʿu*. Some systems, but not that of the EALL, indicate the presence of the *waṣla* carrier ٱ in the Arabic script by an apostrophe, even where unpronounced: *'al-mujammaʿu*, *fī 'l-mujammaʿi*, *wa-'stafāda*. Note that the auxiliary vowel varies, depending on the preceding word, such as the preposition من *min* (مِنَ ٱلمجمع *mina l-mujammaʿi*), the pronouns هم *hum* and أنتم *ʾantum* (هُمُ ٱلمؤمنون *humu l-muʾminūna*), or several words ending in *-w*, such as *ʾaw* (أو ٱلشرق *awi š-šarq*). The transcription may or may not reflect this; *hum al-muʾminūna* is common as well, and is the preferred transcription of the EALL. In DMG transcription, it is very common to add the vowel to the preceding word instead, which may be preferable for transcribing poetry: *humu l-muʾminūna*, etc., even though this may lead to awkward forms such as قد ٱنتظر *qadi ntaḏ̣ara*.

In fully vocalized transcription, some phonetic rules of Arabic concerning vowel length are usually not observed, for the sake of simplicity. Long vowels before consonant clusters are pronounced short, but transcribed long: إلى المجمع *ʾilā l-mujammaʿi* (vs. **ʾila l-mujammaʿi*). Similarly, the distinction between long and short pronunciation of enclitic pronouns (based on the preceding syllable: كتبه *katabahū* vs. رواه *rawāhu*) is usually only observed when transcribing poetry, where the original meter needs to be preserved. Rhyming vowels etc. in poetry are transcribed as long, in accordance with the pronunciation rules for → pausal forms in verse.

Inflectional endings are usually transcribed at the end of a word, even when a *waṣla* follows, and only omitted if the word is in pausal form: بسم الله الرحمن الرحيم *bi-smi llāhi r-raḥmāni r-raḥīm*.

Transcription in pausal form is used for transcribing proper names, references, single words within non-Arabic text, and quotations in Modern Standard Arabic. Here, inflectional endings are usually omitted, unless indispensable for understanding, and in adverbs, such as شكراً *šukran*, تقريباً *taqrīban*. *Waṣla* is omitted after a preceding vowel and transcribed as its auxiliary vowel otherwise: الانقلاب *al-inqilāb*, في الانقلاب *fī l-inqilāb*. The → *nisba* ending (*-iyy*) is transcribed as *-ī*, yet is treated as a consonant: الصراع العربي الإسرائيلي *aṣ-ṣirāʿ al-ʿarabī al-ʾisrāʾīlī* (and not **al-ʿarabī l-ʾisrāʾīlī*).

There are some special conventions for transcribing → proper names, where *waṣla* is usually transcribed even after vowels: مرتضى الزبيدي *Murtaḍā az-Zabīdī* (not **Murtaḍā z-Zabīdī*). Names consisting of a genitive compound are an exception to this rule: محيي الدين *Muḥyī d-Dīn* (**Muḥyī ad-Dīn* is uncommon). The words بنت *bint* and ٱبن (بن) *ibn* within a name's *nasab* chain may be shortened to *bt.* and *b.*, respectively, but this has not been done in the EALL transcription. Finally, some authors omit the article when transcribing *laqab*-style names such as البغدادي *al-Baġdādī*, leaving only a single hyphen: *-Baġdādī*, or dropping it altogether: *Baġdādī*.

Capitalization is usually observed only in proper names and for the first word within a sentence. Single Arabic terms within non-Arabic text are commonly typeset in italics. Latin punctuation signs should be given for their Arabic equivalents, with the exception of brackets, which are sometimes used as quotation marks in Arabic texts.

3. Scientific transcription in dialectology

As most dialects exhibit phonetic features absent in Classical Arabic, extended transcription systems had to be developed for dialectology. In the past, these systems served two goals: to accurately represent the dialect's phonetics and to reflect its relationship with the Classical language. A typical example is Marçais' transcription system for North African dialectology, where a system of double diacritics is used for narrow description of phonetic features such as vowel quantity and articulatory position, as described, for instance, in Singer (1984:31–36). The main problem with such a system is that the distinction between the narrow phonetic view and the broad morphophonemic or even comparative level of description can be quite hard to maintain. As a result, it has become increasingly common to use the International Phonetic Alphabet (IPA) for the narrow phonetic details, and to describe the broader and comparative levels separately using a scientific transcription system for Arabic. In the EALL, phonetic transcription within square brackets [ ] always uses IPA symbols, including the length sign [ː], whereas phonological transcription between slashes / / uses the scientific transcription.

Most transcription systems used in dialectology are extensions of the 'classical' systems outlined above, with the following modifications:

i. Substitution of LC-style digraphs by other symbols, mostly DMG-style letters or Greek letters from phonetic alphabets;
ii. Addition of symbols in analogy to existing symbols, e.g. *ž, č, ḅ* after *š, ğ, ḍ*;
iii. Addition of new symbols from other sources, such as *ś, ć* (from Polish orthography), *ä* (from German orthography), or *ə* (from phonetic alphabets);
iv. Addition of diacritic signs, such as accent characters for prosodic features;
v. Differentiation of existing symbols for the sake of precision, such as *ẓ*, which is usually split into *ẓ* and *ḏ̣* as emphatic equivalents of *z, ḏ*.

There is no formal 'standard' for dialectological transcription of Arabic so far. Most of the rules laid out above for details such as *waṣla* treatment, pausal forms, etc. are not applicable in dialectology. Instead, the transcription usually tries to reflect the structure of the language data as accurately as possible. In the EALL system, this means, for instance, that articles are not hyphenated, e.g. *iššams*, not *iš-šams*. A list of common symbols with their IPA equivalents is given in Fischer and Jastrow (1980:11–14).

BIBLIOGRAPHICAL REFERENCES

Berry, Randal K. (ed.). 1997. *ALA-LC romanization tables: Transliteration schemes for non-roman scripts.* Washington, D.C.: Library of Congress.

Brockelmann, Carl a.o. (eds.). 1935. *Die Transliteration der arabischen Schrift in ihrer Anwendung auf die Hauptliteratursprachen der islamischen Welt: Denkschrift dem 19. internationalen Orientalistenkongreß in Rom.* Leipzig: Brockhaus. (Repr., Wiesbaden: F. Steiner 1969.)

Burgess, James. 1897. "The transliteration of Oriental alphabets". *Actes du X^e^ Congrès International des Orientalistes*, I, 27–42. Leiden: E.J. Brill.

Fischer, Wolfdietrich and Otto Jastrow. 1980. *Handbuch der arabischen Dialekte.* Wiesbaden: O. Harrassowitz.

Jones, William. 1788. "A discourse on the orthography of Asiatic words". *Asiatic Researches* 1.1–46.

Reichmuth, Philipp and Samuli Schielke. 2000. *Die wissenschaftliche Umschrift des Arabischen nach den Regeln der Deutschen Morgenländischen Gesellschaft.* Bonn: Fachschaft Orientalisch-Asiatische Kulturwissenschaften.

Singer, Hans-Rudolf. 1984. *Grammatik der arabischen Mundart der Medina von Tunis.* Berlin and New York: de Gruyter.

Socin, Albert. 1895. "Referat über die Transcriptionsfrage". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 49.180–183.

Wellisch, Hans. 1975. *Transcription and transliteration: An annotated bibliography on conversion of scripts.* Silver Spring, Md.: Institute of Modern Languages.

Whitney, William D. 1862. "On Lepsius' standard alphabet". *Journal of the American Oriental Society* 7.299–332.

PHILIPP REICHMUTH (University of Halle)

Transformational Grammar

1. INTRODUCTION

There are a number of areas where Arabic, in both its Standard varieties and its modern dialects, can provide important insights into (and ultimately contribute to the search for an adequate) linguistic theory (cf. Eid 1987; Comrie 1991). Conversely, although the study of Arabic has a long tradition (cf. Bakir 1980:1ff.), and it is possible at times to find similarities between Transformational Grammar and Arabic grammarians' analyses (see below), modern linguistic theory allows for even more insightful and rigorous analyses of Arabic.

2. A SURVEY OF TRANSFORMATIONAL GENERATIVE GRAMMAR

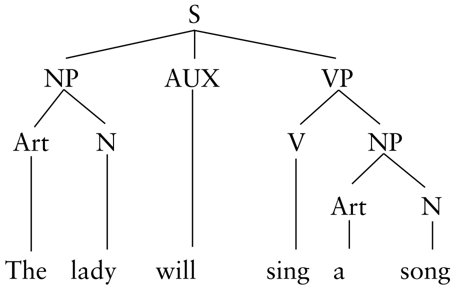
The label Transformational Grammar is used somewhat loosely in this entry to refer to models of generative grammar that precede Chomsky's *Principles and parameters* (1981), specifically versions of the Standard Theory and Extended Standard Theory models as developed in Chomsky (1957, 1965, 1977).

2.1 *Phrase Structure Grammar*

In the early Standard Theory, it was assumed that basic – as opposed to derived – tree diagrams (P-markers) like the one in (2) were generated by phrase structure rules, such as those in (1), together with lexical insertion, i.e. the selection of appropriate words from the lexicon or mental 'dictionary'.

(1) S → NP AUX VP
NP → Art N
VP → V NP

(2)



In the Standard models, grammar is assumed to include not only phrase structure rules but also a lexicon where words are listed with their grammatical categories and subcategorization information (specification of the contexts in which these words can occur). The two constitute the Base component of the grammar.

Some (pairs of) phrase structure rules, e.g. those in (3), are said to be 'recursive': they are able to introduce the same element indefinitely and hence deal with complex constructions, containing (an) embedded category/(ies), such as NP in (3a) and S in (3b).

(3a) NP → (Art) N (PP)
PP → P NP

(3b) S → NP AUX VP
VP → V (S)

This illustrates an important concept of language acquisition: creativity. In such a model, grammatical relations, such as 'Subject of' or 'Object of', are defined structurally and represented in terms of the tree configurations. In (2), the subject may be defined as the NP which is immediately dominated by S and precedes VP; and the object may be defined as the NP immediately dominated by VP.

2.2 *Transformational Grammar*

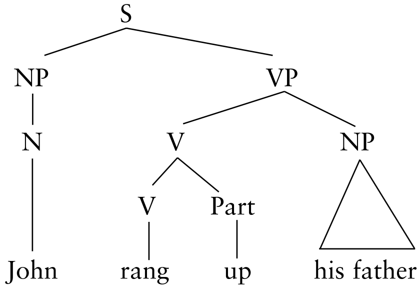
There exist certain relations, e.g. relations of modification, between parts of a sentence which are not contiguous but discontinuous, and therefore not analyzable in terms of single tree diagrams. Examples of such relations include verb/particle constructions, illustrated in (4).

(4a) *John rang up his father*

(4b) *John rang his father up*

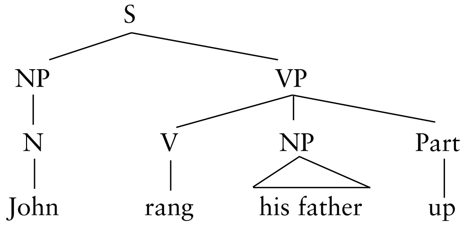
In (4a), the verb and particle form a single constituent, a phrasal verb. In (4b), the two elements are not contiguous. The theory of Transformational Grammar was developed to deal with cases like (4b), which is assumed to derive from the same tree as (4a), namely (5); this is called the 'input (base, underlying, deep) structure' for sentence (4b).

(5)



To derive an 'output (derived, surface) structure' from (5), a grammatical operation is postulated called a 'transformation' – in this case Particle-movement – which converts (5) into (6).

(6)



This analysis reflects the possibility that related elements may get separated during the process of deriving a surface structure, and, hence, local relationships sometimes result in long-distance ones.

The addition of a transformational component to the familiar base component raises a number of issues. The first issue is connected with the form of transformations: transformational rules are formalized in terms of a structural description (SD) and a structural change (SC). If a P-marker satisfies the SD of a particular transformation, this transformation will apply and change the SD according to the particular transformation. An example is (7), illustrating → WH-movement, which includes WH-questions together with other transformations believed to share certain characteristics with it, such as Relativization (→ relative clause), Clefting, and → Topicalization.

(7) WH-movement
SD: X Comp X wh X
[+Q]
1 2 3 4 5
SC: 1 4+2 3 ø 5
(SD and SC consist of either category labels or variables; + indicates that *wh* is attached to the left of 2: the two are 'sisters'; Q indicates that the sentence is a Question.)

A second issue concerns the applicability of transformations: transformations may apply repeatedly in complex sentences, and, hence, their order and domain of application must be determined. It is assumed that there are two major types of transformations: bounded transformations, including clause-bound ones, such as → Reflexive and → Passive, and transformations operating across one clause boundary, such as Equi and Subject-to-Object Raising; and unbounded transformations, operating across an indefinite number of clause boundaries, such as → WH-movement.

Bounded transformations are further classified as cyclic, and unbounded ones as postcyclic. In order to get rules to interact correctly, the principle of the transformational cycle is developed. It requires that cyclic rules (i.e. bounded transformations) should apply first to the most deeply embedded clause, then to the next most deeply embedded one, until the main clause is reached.

The third issue concerns the status of the rules. Rules are described as either obligatory or optional (see, e.g., Akmajian and Heny 1975). In fact, two versions of the Standard theory may be distinguished. In the earlier version, transformations like Passive and Question were considered optional, whereas Affix Hopping (which places affixes on the following verbs) was taken to be obligatory. In the later version, the optional transformations were reformulated as obligatory. This is made possible by having the phrase structure rules optionally introduce an element that will trigger the operation of the rule, as in (8).

(8) S → (Q) (not) NP AUX VP

The general formalism for transformations provided by such a model makes the Standard theory highly expressive in that every detail of a given construction can easily be included in one rule, with the result that the latter becomes highly stipulative as it does not explain why we must have these features and not others.

Accordingly, this general theory of transformations was gradually abandoned and replaced by a general movement rule 'move something somewhere', which is subject to certain constraints (→ minimalism).

The first attempts at establishing a general theory of conditions on transformations are Chomsky (1964) and Ross (1967). Their most important contribution was in initiating the search for general constraints on transformations, and this objective has since been the guiding principle in syntactic research. The idea is that general conditions make possible a reduction in the expressive power of rules and therefore restrict the options for grammars for the language learner (for a review, see Horrocks 1987 and Radford 1981; for a summary, see Chekili 2007).

Ross (1967), for instance, discovered that certain constructions do not allow the extraction of any type of phrase. He labeled these 'islands', and tried to account for this finding by imposing a number of constraints on transformations, thereby simplifying the statement of the transformations.

2.3 *Movement or base-generation: Topicalization and Left Dislocation*

Prior to Chomsky (1977), → Topicalization constructions such as (9) were analyzed (e.g. by Ross 1967) as involving the movement of *John* from the object position to the preverbal one:

(9) *John, I saw*

Chomsky observed that this account is not right, but that such a structure is another instance of WH-movement. For example, the movement is unbounded, as in (10).

(10) *John, Peter believes that John said I saw*

Movement is subject to island constraints, as in (11).

(11) **This man, I wonder who met*

Chomsky compared (9) with (12), and following Ross, he called the former 'Topicalization' and the latter 'Left Dislocation'.

(12) *John, I saw him*

The major difference between the two structures, Chomsky noted, following Ross, is the absence of movement in the latter – due to absence of a gap – as evidenced by the fact that it is not subject to island constraints.

(13) *This man, I wonder who met him*

Chomsky proposed the configurations in (14).

(14a) [S'' [Top John] [S' [S I saw him]]]

(14b) [S'' [Top John] [S' [S I saw e]]]

A phrase structure rule (15) generates (14a) and accounts for the absence of movement in (14a), as the latter is base generated.

(15) S'' → Top S'

Since (14a) is postulated for Left Dislocation, it can also be used for Topicalization; in other words, Chomsky argued that *John* is base generated both in Left Dislocation and Topicalization. He accounted for the difference between the two constructions by arguing that in the latter – but not the former – a null WH-item has moved. This means that Topicalization involves WH-movement (for discussion, see Lasnik and Uriagereka 1988:153ff.).

3. SELECTED TOPICS

In this section, emphasis is on syntax, rather than other aspects of structure; specifically, → word order and WH-constructions have been selected. This does not have any negative impact on the overall objective, however, which is the explanation of the relationship between Transformational Grammar and Arabic.

3.1 *Word order*

Arabic in all its varieties instantiates different word orders as illustrated in Standard Arabic (16).

(16a) *ištarā* *zayd-un* *kitāb-an* (VSO)
bought Zayd-Nom book-Acc

(16b) *zayd-un* *ištarā* *kitāb-an* (SVO)

(16c) *ištarā* *kitāb-an* *zayd-un* (VOS)

(16d) *kitāb-an* *ištarā* *zayd-un* (OVS)
'Zayd bought a book' (examples adapted from Bakir 1980:6)

This means that Arabic provides the linguist who is intent on the description of word order variation with an important database (cf. Eid 1987, and references there).

Arabic is also typologically interesting in that it seems to instantiate a VSO basic order in its standard variety, sharing certain features with other VSO languages, such as "[having] alternative basic orders among which SVO always figures" (Greenberg 1963:79), but a basic SVO order in its modern dialects (Eid 1987). These observations have led to numerous studies concerned with different aspects of word order (Eid 1987), including the determination of a 'basic' order among different orders, the criteria for selecting one order as 'basic', etc.

As far as Standard Arabic (and Classical Arabic) is concerned, these studies fall into different groups according to what they take to be the 'basic' or underlying order. For example, Arabic is taken to be underlyingly SVO (e.g. Snow 1965; Killean 1966; Lewkowicz 1967, 1971), or – in line with traditional Arabic grammarians – VSO (e.g. Russell 1977; Bakir 1980; Fassi Fehri 1981), or VOS (e.g. Anshen and Schreiber 1968; Homeidi 1987). The advantages and weaknesses of each of the above proposals are summarized in Homeidi (2003:2–8).

Prior to Transformational Grammar, 'basic' word order was determined on the basis of both pragmatic factors and syntactic distribution (cf. Bakir 1980:5–10). These two criteria have led traditional grammarians and some modern ones to take the 'basic' word order in Arabic to be VSO (Bakir 1980:8–10).

The question of basic word order took on a renewed significance with the development of Transformational Grammar with its distinction between two levels of structure, namely 'deep structure' and 'surface structure', together with a transformational component. In this view, basic word order has come to mean the order encountered at deep structure and generated by the phrase structure rules. The alternative orders are the result of transformational rules operating on these deep structures and yielding the various related structures.

Grammatical functions such as 'Subject' and 'Object' are defined configurationally at deep structure. Thus, all languages must have a VP at deep structure. As a direct consequence of this claim – and in order to maintain the configurational definition of grammatical relations – a VSO basic order is abandoned, as it does not include a VP node. This led to the postulation of different underlying orders for languages known to have a basic VSO order (cf. Bakir 1980:10ff. for discussion).

For example, Snow (1965), Killean (1966), and Lewkowicz (1971) rewrite S as NP-PredP. Snow adduces arguments from simplicity of concord and of the overall grammatical statement in favor of this order (see Bakir 1980:13–14). Anshen and Schreiber (1968) propose a VOS word order in order to maintain the configurational definition of grammatical relations.

Bakir (1980:14ff.) rejects the preceding two hypotheses and prefers to maintain an underlying VSO order for Arabic. To do this, he has to extend the definition of grammatical functions to include "the linear order of the constituents in the deep structure" (1980:14) and to address the question of a 'universal' VP constituent. Showing that the usual VP constituency tests do not apply to Arabic, he concludes that there is no VP constituent in Arabic. Bakir motivates his preference for VSO, first by arguing that the movement involved "is part of a general process of 'movement' of constituents to the beginning of the sentence" (1980:16), and second by making use of the conventional distinctions made by Transformational Grammar linguists, including a distinction between optional and obligatory transformations, the number of transformations required in the derivation of a given structure, etc. (1980:16–18).

Emonds (1979) has reacted against a flat nonconfigurational structure for VSO languages and presented arguments to the effect that VSO and SVO languages share the same underlying structure: VSO is derived from SVO by means of a process known as 'Verb Raising'.

More recently, Kuroda (1988) and Koopman and Sportiche (1988) have proposed what came to be known as the 'VP-internal subject hypothesis', according to which the subject is always generated inside VP. VSO order is derived by raising the verb to the inflectional position (cf., e.g., Koopman and Sportiche 1991; Plunkett 1993; Aoun a.o. 1994). Without going into the details of the analysis, this hypothesis is compatible with Emonds' suggestion that VSO languages have an underlying SVO structure. SVO is also assumed by Mohammad (1989), who uses agreement facts and the notion of null expletive subject (→ pro-drop) to justify his claim.

3.2 *WH-constructions*

Interest in WH-constructions, understood here to refer to WH-questions, Relative Clauses, Clefting, Topicalization, etc., goes back to the early transformational models. WH-questions, together with other transformations believed to share certain characteristics with them, were generalized as a transformation of WH-movement.

Later transformational models, such as Ross (1967) and Chomsky (1964, 1977), have generated even greater interest in 'WH-constructions', because they relied on such constructions in developing a general theory of conditions on transformations (cf. Sec. 2.2).

In order to present the relationship between Transformational Grammar and Arabic, the discussion here focuses on Left Dislocations and Topicalizations, in which Arabic is superficially similar to English; and Relativization, in which it is different as it requires a resumptive pronoun (→ resumption) where English requires a gap. Note, however, that some varieties use a mixture of the two processes (cf., e.g., Shlonsky 1992:444ff.).

The distinction between Topicalization and Left Dislocation in English (Ross 1967; Chomsky 1977) has led to a distinction in Arabic between two types of structure, which are called here, following certain linguists, → Topic/Comment (or Left Dislocations) and → Focusing (or → Topicalizations).

Within Transformational Grammar, and by analogy with the English accounts of Left Dislocation and Topicalization, the analysis of these Arabic constructions has passed through two main stages: at first, an account in terms of a transformational rule, and more recently, base-generation. Note that the latter is mainly due to Chomsky (1977) and happens to coincide with the account of the traditional Arabic grammarians, who analyze Topic/Comment structures as consisting of a *mubtadaʾ* 'inchoative' (→ *ibtidāʾ*) and a → *xabar* 'enunciative' (cf. Wright 1898:II, 251).

As the Topic/Comment (17) and Focusing (18) show, the two constructions involve the presence of an element (usually NP) in a sentence-initial position. However, whereas Topic/Comment usually requires the retention of a resumptive (or replacive) pronoun, Focusing does not. Also, while the Topic NP always gets nominative, the Focused NP retains the case assigned to its sentence-internal position. (For other differences between the two structures in Arabic not instantiated in English, see Bakir 1980:61ff.). Compare (17) and (18), taken from Al-Hroot (1994:59).

(17) *al-kitāb-u* *qaraʾa-hu* *zayd-un*
the-book-Nom read-it Zayd-Nom

(18) *al-kitāb-a* *qaraʾa* *zayd-un*
the-book-Acc read Zayd-Nom

The distinction in early Transformational Grammar of a level of deep structure (on top of the familiar level of surface structure), together with the development of the notion of transformation, has led to a transformational analysis of Topic/Comment sentences in Arabic. A number of linguists (e.g. Snow 1965; Killean 1966; Anshen and Schreiber 1968) take the Topic/Comment to be the result of a transformational rule. More specifically, a Copying rule will optionally apply and copy an NP to the left of the verb, and the original NP is changed by a Pronominalization process into a resumptive pronoun. For example, (19) is assumed to derive from (20) by undergoing first a Copying transformation, resulting in the structure (21), and then a Pronominalization transformation, yielding (19).

(19) *al-kitāb-u* *qaraʾa-hu* *zayd-un*
the-book-Nom read-it Zayd-Nom

(20) *qaraʾa* *zayd-un* *al-kitāb-a*

(21) *al-kitāb-u* *qaraʾa* *zayd-un* *al-kitāb-a*

An example of such a transformation was proposed by Anshen and Schreiber (1968) in (22).

(22) 'X-NP-Y ⇨ NP'-X-NP-Y where NP=NP'

More recent work in Transformational Grammar has rejected a transformational account of this type of structure in favor of a base-generated one: as Topic/Comment structures are base generated, they are the product of a phrase structure rule.

Lewkowicz (1971) makes use of the different treatment given to embedding within the Standard theory (the transformational version vs. the phrase structure version, which she favors) to explain how the embedding takes place. Whereas in early work an embedding transformation would replace the symbol 'Comment' by "any sentence containing an NP duplicating the topic" (Lewkowicz 1971:814), in later work, embedding is no longer introduced by a transformation but directly by the phrase structure rules. In this case, the symbol #S# is used to refer to the Comment, resulting in recursion, as in (23), from Lewkowicz (1971:812–813).

(23) *al-waladu* *ʾabū-hu*
the-boy father-his
baytu-hu *kabīrun*
house-his large
'The boy's father's house is large'

The phrase structure rules have the form in (24).

(24) S → NP PredP

$$\text{PredP} \rightarrow \left\{\begin{array}{l} \text{(copula)} + \left\{\begin{array}{l}\text{VP}\\ \text{AdjP}\end{array}\right\} \\ \\ \qquad\qquad\quad \text{NP} \end{array}\right\}$$

AdvP

#S#

Although embedding Comment sentences is a phrase structure phenomenon (the phrase structure version), a transformation – obligatory – is still required to describe the structure of the Comment, as in (25), from Lewkowicz (1971:815).

(25) # NP # X + NP + Y ## ⇨ 1 ø 3 4 5 ø
1 2 3 4 5 6
where 4 is nondistinct from 1.

The transformation will delete the 'internal boundaries' (#) just in case the structural description is satisfied (Chomsky 1965).

The principle of the transformational cycle can account for the difference in case marking between the Topic NP and the duplicate NP inside the Comment. This is because such a principle ensures that "the inner sentence (the comment) has already been subjected to the complete cycle of transformations including those for case of nouns; while the topic, which belongs only to the higher level sentence, might not yet have been subjected to the complete cycle of transformations" (Lewkowicz 1971:815, n. 11).

Later developments in Transformational Grammar, aiming at restricting the power of transformations, have maintained that pronouns are no longer to be derived by transformation from an underlying NP but must instead be base generated. This is supported by later work on Relativization in Hebrew by Chomsky (1977), where he shows that the resumptive pronoun must be base generated and interpreted in its surface position.

Bakir (1980), comparing Topic/Comment with Focusing, argues that, while Topic/Comment, like English Left Dislocation, contains a resumptive pronoun that is coreferential with the base-generated Topic NP, the focused NP, on the other hand – unlike English – undergoes movement, leaving a gap (trace) and, hence, retaining all the features of the trace (including case). Note that this proposed movement rule for Focusing, like the rule for WH-movement, is a 'chopping rule' (Ross 1967), not to be confused with the so-called copying rule that was first proposed to deal with Topic/Comment.

Relative clauses have also been the focus of much interest within a transformational framework (e.g. Snow 1965; Killean 1966; Anshen and Schreiber 1968; Lewkowicz 1971). A number of linguists within early Transformational Grammar argued that the Relative clause is derived from the Topic/Comment by means of a transformation. This is because the two structures are similar, particularly in their use of the resumptive pronoun. Compare (26) and (27) from Lewkowicz (1971:818).

(26) *al-waladu māta ʾabūhu* (Topic/Comment)

(27) *(waladun) māta ʾabūhu* (Relative)

Like the Topic/Comment, the relative clause contains an NP which duplicates the head noun modified by the relative clause. (28) is a Relative transformation (Lewkowicz 1971:819).

(28) # W+N # X+NP+Y# Z # ⇨ 1 2 ø 4 5 6 ø 8
1 2 3 4 5 6 7 8
where (a) 3 . . . 7 is a Relative,
(b) 5 is nondistinct from 2,
(c) 5 is not indefinite.

Example (27) is derived by embedding (29).

(29) *māta ʾabū l-waladi*
'The boy's father died'

As mentioned above, more recent work shows that the resumptive pronoun in relative clauses must be base generated.

Resumptive pronouns can occur not only as object of a verb, as in *al-kitāb-u qaraʾa-hu zayd-un*, above, but also as independent subject. Generally, unless used for emphasis, independent subject pronouns do not occur in the postverbal subject position.

However, both traditional Arabic grammarians and modern linguists who take (30) and (31) from Lewkowicz (1971:811) to be Dislocations (rather than simple sentences) would analyze these as containing a resumptive pronoun (see below).

(30) *al-waladu* *māta*
the-boy died

(31) *'ana* *qultu…*
I said…

Although the difference between Left Dislocation and Topicalization in English corresponds, to some extent, to the difference between Topic/Comment and Focusing, this correspondence is only partial (Bakir 1980:60ff.).

The two structures in English are sufficiently similar to motivate a similar account in terms of base-generation. In Arabic, on the other hand, the two structures display a number of differences besides the presence vs. absence of a resumptive pronoun (see above) and hence would require a separate analysis.

A single transformational account (as in Ross 1967) would have to explain why some constituents retain their case under movement while others do not. Similarly, a single base- generation account (Chomsky 1977) would be unable to account for the distributional and other peculiarities displayed by the two types of structure (Bakir 1980:61). Thus, Arabic motivates a distinct treatment for the two structures.

Furthermore, whereas Chomsky (1977) accounts for Topicalization in terms of base-generation plus WH-movement, Arabic suggests that Topicalization (Focusing) is a movement (not a WH-movement) process by which the NP moves from the right to the left of the verb. This would explain the moved NP's retaining its case. Under a WH-movement analysis, there would be a need for a 'case matching or copying process', by means of which the case of the base-generated NP is 'matched' with the case of the WH-element (for a discussion, see Bakir 1980:71ff.).

On the other hand, Arabic provides some support for Chomsky's claim that Left Dislocations are base generated. For instance, Bakir (1980:150) shows that a transformational account would have to be complicated, as it must account, for instance, for case marking and pronominalization. Similarly, structural descriptions would have to be complicated considerably in view of the requirement in Arabic that only definite NPs can occur in initial position. Compare (32) and (33), adapted from Abdul Ghany (1981:75).

(32) *al-kitāb-u qara'a-hu zayd-un*

(33) **kitāb-un qara'a-hu zayd-un*

The problem disappears under a base-generation approach and a general restriction prohibiting the initial occurrence of definite NPs (cf. Abdul Ghany 1981:75).

Finally, as pointed out above, Arabic raises the question of how to analyze the initial NP in SVO, as subject or Topic. Traditional Arabic grammarians, in particular those of the school of Basra, do not accept the idea that a subject may precede a verb, and consider the initial NP in such clauses to be a Topic, rather than a subject (cf. Bakir 1980:125ff. for a discussion).

This situation has given rise to a dual account of such structures on the part of generative linguists (see Chekili 2002 for a review of the literature): on the one hand, there are those who take the initial NP to be a 'genuine' subject with the whole structure a simple sentence (e.g. Snow 1965; Lewkowicz 1967; and, more recently, Aoun a.o. 1994), and, on the other hand, those who, following traditional grammarians, analyze it as a Topic with the whole structure being a Topic/Comment analyzed structurally as a Left Dislocation (cf., e.g., Anshen and Schreiber 1968; and, more recently, Bakir 1980; Fassi Fehri 1981; Plunkett 1993). On this account, the subject pronoun is either 'implied' (Wright 1898), or identified with a specific verbal affix (Anshen and Schreiber 1968), or is first generated and later deleted (Bakir 1980:159). Note that the process of deletion is blocked by means of a constraint similar to the one proposed by Grinder (1970) for Equi-NP deletion (cf. Bakir 1980:170–171). Eid (1977) contains a discussion of a similar constraint (an intervention constraint) on the deletion of subject pronouns in relative clauses in Egyptian Arabic. In a more recent formulation, this pronoun is a *pro* subject, i.e. the nonovert equivalent of an overt resumptive pronoun, resulting from the fact that Arabic is a null subject language (→ pro-drop).

Bibliographical references

Abdul-Ghany, Muhammad Kamal Eldine. 1981. *Government and Binding in Classical Arabic.* Ph.D. diss., University of Texas at Austin.

Akmajian, Adrian and Frank Heny. 1975. *Introduction to the principles of transformational syntax.* Cambridge, Mass.: MIT Press.

Al-Hroot, A. Khalaf. 1994. *Some aspects of deletion in Arabic.* Ph.D. diss., Mutah University.

Anshen, Frank and Peter A. Schreiber. 1968. "A focus transformation of Modern Standard Arabic". *Language* 44.792–797.

Aoun, Joseph, Elabbas Benmamoun, and Dominique Sportiche. 1994. "Agreement and conjunction in some varieties of Arabic". *Linguistic Inquiry* 25.195–220.

Bakir, Murtadha. 1980. *Aspects of clause structure in Arabic.* Bloomington: Indiana University Linguistic Club.

Chekili, Ferid. 2002. "Agreement asymmetries and the lexical/null subject parameter". *Al-ʿArabiyya* 35.87–127.

——. 2007. *Introduction to generative syntax.* Tunis: Publications de la Faculté de la Manouba.

Chomsky, Noam. 1957. *Syntactic structures.* The Hague: Mouton.

——. 1964. *Current issues in linguistic theory.* The Hague: Mouton.

——. 1965. *Aspects of the theory of syntax.* Cambridge, Mass.: MIT Press.

——. 1977. "On Wh-Movement". *Formal syntax*, ed. Peter Culicover, Thomas Wasow, and Adrian Akmajian, 71–132. New York: Academic Press.

——. 1981. *Lectures on Government and Binding.* Dordrecht: Foris.

Comrie, Bernard. 1991. "On the importance of Arabic to general linguistic theory". *Perspectives on Arabic linguistics*, III, ed. Bernard Comrie and Mushira Eid, 3–30. Amsterdam and Philadelphia: J. Benjamins.

Eid, Mushira. 1977. "Arabic relativization: Shadow deletion or pronoun drop?". *Minnesota Working Papers in Linguistics and the Philosophy of Language* 4.19–32.

——. 1987. "Arabic theoretical linguistics: The seventies and beyond". *Al-ʿArabiyya* 20:1–2.181–226.

Emonds, Joseph. 1979. "Word order in generative grammar". *Explorations in linguistics*, ed. George Bedell, Eichi Kobayashi, and Masatake Muraki, 58–88. Tokyo: Kenkyusha Press.

Fassi Fehri, Abdelkader. 1981. *Linguistique arabe: Forme et interprétation.* Rabat: Publications of the Faculty of Letters.

Greenberg, Joseph H. 1963. "Some universals of grammar with particular reference to the order of meaningful elements". *Universals of language*, ed. Joseph Greenberg, 73–113. Cambridge, Mass.: MIT Press.

Grinder, John. 1970. "Super Equi-NP deletion". *Papers from the Sixth Regional Meeting of the Chicago Linguistic Society*, ed. Mary Ann Campbell a.o., 297–317. Chicago: Chicago Linguistic Society.

Homeidi, Mohammad A. 1987. *Modality in Government and Binding.* Ph.D. diss., University of Essex.

—— 2003. "Word order in Modern Standard Arabic: A Minimalist approach". *Journal for the Arts and Humanities* (Damascus) 19:3–4.53–73.

Horrocks, Geoffrey C. 1987. *Generative grammar.* London and New York: Longman.

Killean, Carolyn. 1966. *The deep structure of the noun phrase in Modern Written Arabic.* Ph.D. diss., University of Michigan.

Koopman, Hilda and Dominique Sportiche. 1988. "Subjects". Ms., University of California Los Angeles.

—— and Dominique Sportiche. 1991. "The position of subjects". *Lingua* 85.211–258.

Kuroda, Sige-Yuki. 1988. "Whether we agree or not: A comparative syntax of English and Japanese". *Papers from the Second International Workshop on Japanese Syntax*, ed. William J. Poser, 103–143. Stanford, Calif.: Stanford University.

Lasnik, Howard and Juan Uriagereka. 1988. *A course in GB syntax: Lectures on Binding and Empty Categories.* Cambridge, Mass., and London: MIT Press.

Lewkowicz, Nancy Kennedy. 1967. *A transformational approach to the syntax of Arabic participles.* Ph.D. diss., University of Michigan.

——. 1971. "Topic-Comment and relative clause in Arabic". *Language* 47.810–825.

Mohammad, Mohammad. 1989. *The sentence structure of Arabic.* Ph.D. diss., University of Southern California.

Plunkett, Bernadette. 1993. "The position of subjects in Modern Standard Arabic". *Perspectives on Arabic linguistics*, V, ed. Mushira Eid and Clive Holes, 231–260. Amsterdam and Philadelphia: J. Benjamins.

Radford, A. 1981. *Transformational syntax: A student's guide to Chomsky's Extended Standard Theory.* Cambridge: Cambridge University Press.

Ross, John R. 1967. *Constraints on variables in syntax.* Ph.D. diss., Massachusetts Institute of Technology.

Russell, Robert. 1977. *Word order of Classical and Egyptian Arabic.* Ph.D. diss., Harvard University.

Shlonsky, Ur. 1992. "Resumptive pronouns as a last resort". *Linguistic Inquiry* 23.443–468.

Snow, James A. 1965. *A grammar of Modern Written Arabic clauses.* Ph.D. diss., University of Michigan.

Wright, William. 1898. *A grammar of the Arabic language*, II. 3rd rev. ed. W. Robertson Smith and M.J. de Goeje. Cambridge: Cambridge University Press. (Repr., 1962).

FERID CHEKILI (University of Manouba)

Transitivity

1. Basic definition of transitivity

Transitivity is defined narrowly as a state of direct dependence of a nominal phrase (i.e.

noun or noun phrase) on a verbal element (i.e. verb or verblike element). A transitive clause involves a minimum of two participants – a subject (prototypically an → agent) and an object (prototypically an affected entity), contrasting with an intransitive clause that involves a single participant subject (prototypically an agent or experiencer).

Transitivity is considered here in respect of Standard Arabic and modern Arabic dialects. In Standard Arabic, where case is expressed through vocalic endings, the agent takes the nominative case (normally *-u*), and the object takes the accusative case (normally – *a*). In modern colloquial dialects, the difference between nominative and non-nominative case is expressed in the pronouns (e.g. *ana* nominative 'I' vs. *-nī* accusative 'me'), but not in nouns.

2. TRANSITIVE ELEMENTS

In both Standard and colloquial Arabic, the following verbal elements may be transitive: verbs, active → participles, and → verbal nouns. Verbs can be taken as the most basic transitive element. The following are examples of transitivity with verbs (where no data source is mentioned, such as for Khartoum and Ṣanʿānī Arabic, data is taken from Watson and Dickins' databases):

Standard Arabic	Ṣanʿānī
raʾaynā r-rajul-a	*absarna r-rajjāl*
raʾaynā-hu	*absarnā-h*
Khartoum	
šufna r-rājil	'we saw the man'
šufnā-hu	'we saw him'

The following are examples of transitivity with active participles (problematic examples, placed in square brackets, are further discussed below):

Standard Arabic	Ṣanʿānī
huwa lābis(-un) jilbāb-an	*(hu) lābis zinnih*
[huwa lābis-u-hu]	*[(hu) lābis li-h]*
Khartoum	
(hu) lābis jallābiyya	'he's wearing a robe'
(hu) lābs-u	'he's wearing it [masc.]'

Modern Standard Arabic does not allow an object pronoun suffix following an active participle. Instead, a genitive structure has to be used: *huwa lābis-u-hu* properly means 'he is the wearer of it', and not 'he is wearing it'.

In Ṣanʿānī Arabic, participles similarly cannot take object pronoun suffixes. Instead, the preposition *li-* acts as a dummy head element for the pronoun suffix.

In Khartoum Arabic, a few adjectives other than the traditional participles may take an object pronoun suffix. Thus, with *gablān* (on the *faʿlān* pattern): *(ana) ma gablān kalām-ik* 'I don't accept what you [fem. sg.] say'; *(ana) ma gablān-u* 'I don't accept it [masc.]'.

A verbal noun can take an object only when elements corresponding to both the subject and the object of a verbal clause are found in the sentence. In this case, the subject of the verbal clause has as its correspondent a (subjective) genitive in the verbal-noun structure, while the object in the verbal clause may be retained as an object. Take the following Standard Arabic verbal clauses:

istaʿmala l-ʿummāl(-u) hāḏihi l-ʾašyāʾ(-a)	'the workers used these things'
istaʿmala-hā l-ʿummāl(-u)	'the workers used them'

These can be compared with the corresponding verbal-noun structures:

istiʿmāl(-u) l-ʿummāl(-i) hāḏihi l-ʾašyāʾ(-a)	'the workers' use of these things'
istiʿmāl(-u)-hum hāḏihi l-ʾašyāʾ(-a)	'their use of these things'

In the examples with verbal nouns, the genitive *al-ʿummāl(-i)* corresponds to the subject of the verb, *al-ʿummāl(-u)* 'the workers', in *istaʿmala l-ʿummāl(-u) hāḏihi l-ʾašyāʾ(-a)* 'the workers used these things', while *hāḏihi l-ʾašyāʾ(-a)* 'these things' and *-hā* 'them' are objects in both the verbal-clause structure and the verbal-noun structure.

In verbal-noun structures involving both subject and object correspondents, the object correspondent may be introduced by the preposition *li-*. Thus:

istiʿmāl(-u) l-ʿummāl(-i) li-hāḏihi l-ʾašyāʾ(-a)	'the workers' use of these things'
istiʿmāl(-u)-hum la-hā	'their [masc.] use of them'

In Classical Arabic, the object of a verbal noun may be introduced by *li-*, especially when the verbal noun is indefinite; likewise, the object of a participle may be introduced by *li-*, which is the usual construction when the object precedes the participle (Reckendorf 1921:249).

Subject correspondents in verbal-noun structures always appear in the genitive case. When the verbal-noun structure has an object correspondent but no subject correspondent, the object correspondent must also take genitive case: *istiʿmāl(-u) hāḏihi l-ʾašyāʾ(-i)* 'the use of these things'.

Colloquial dialects tend toward greater use of dummy elements, most commonly *li-*, in verbal-noun structures to introduce object correspondents, as in Ṣanʿānī: *istiʿmāl-hum li-hāḏa l-lašyā* 'their [masc.] use of these things'.

Khartoum Arabic allows a nonpronominal object correspondent to be expressed by an object, e.g. *istiʿmāl-um al-ḥājāt di* 'their use of these things'. However, it more commonly makes use of dummy *li-/lē-*, e.g. *istiʿmāl-um li-l-ḥājāt di* 'their use of these things'. Where the object correspondent is pronominal, this must be introduced by *li-/lē-* in Khartoum: *istiʿmāl-um lē-ha* 'their use of them'.

3. Degrees of transitivity

Taking transitivity in the narrow sense given above in Section 1, verbal elements in Arabic may be intransitive, singly transitive, doubly transitive, or, rarely, triply transitive. Verbs involving an indirect and a direct object typically occupy the semantic fields of giving, informing, forcing, and removing (cf. Diem 2002). Where verbs take two direct objects, typically either the first object is converted into the second, e.g. 'to make x into y', or the two objects are in a predicative relationship, e.g. 'to consider x to be y' (Badawi a.o. 2004: 374):

Intransitive	
māt-a	'he died' (Standard Arabic)
Singly transitive	
qatal-a l-faʾr-a	'he killed the rat' (Standard Arabic)
qtal-t-u	'I killed him' (Çukurova; Procházka 2002:106)
Doubly transitive	
addā-ni l-girūš	'he gave me the money' (Khartoum)
sug-na l-ġanam as-sūg	'we drove the goats to market' (Khartoum)
zawwaj-tu zayd-an ibnat-a ʾaxī	'I gave Zayd my brother's daughter in marriage' (Standard Arabic; Wright 1975:II, 48)
ʾaʿlam-a-nī l-ʾamr-a	'he informed me of the affair' (Wright 1975:II, 48)
Triply transitive	
aʿlama-nī zayd-an rajul-an karīm-an	'he informed me that Zayd is/was a generous man' (Standard Arabic)
sawwagnā-hum al-ġanam as-sūg	'we got them to drive the goats to market' (Khartoum)

Verbal elements in Arabic may typically take an independent subject, e.g. *(al-mara/hi) fihmat al-muškila* '(the woman/she) understood the problem' (Khartoum). A few structures exist, however, in which independent subjects are not possible. These include → impersonal passive forms in Standard Arabic: *yuttafaq-u ʿalā hāḏihi l-qaḍāyā* 'these matters are agreed on'; *ʾuʿlina ʿan wafāt-i l-ʾimām-i l-xumaynī* 'the death of Imam Khomeini was announced'.

Such forms always correspond to active verbal elements taking a prepositional complement (Sec. 6, below). The verb in impersonal forms is always in the 3rd person masculine singular.

Colloquial dialects do not typically exhibit impersonal passive forms. However, impersonal uses of the 3rd person masculine plural verb are found in many dialects, as in Khartoum *jābu lē-hu walad* 'a son was born to him' [lit: 'they bore/brought to him a son']; *bigūlu šinu li-l-kilma di bi-l-ʿarabi?* 'what do they call that word/thing in Arabic?'. Neither of these examples allows a subject noun or pronoun (such as the grammatically expected pronoun *hum* 'they').

4. Transitivity vs. adjunction

A distinction is made between transitivity in the sense of direct dependence of a nominal phrase on a verbal element (Sec. 4.1) and adjunction, an adjunct being an adverbial element, often with a prepositional head, which is dependent on the verb phrase (cf. Quirk a.o. 1985:1071–1073). Defined more broadly, transitivity can be taken to encompass not only relationships of direct dependence of a nominal phrase on a verbal element, but also at least certain kinds of adjunct (cf. Badawi a.o. 2004:372), particularly what are known as restrictive adjuncts (Quirk a.o. 1985:1075–1077).

4.1 *Nominal objects and adjuncts in Standard Arabic*

The close relationship between nominal objects and adjuncts is particularly evident in Standard Arabic, where the accusative case is the standard marker of both types of entity. This is obvious in the case of nominal objects: *raʾat al-marʾat-u r-rajul-a* 'the woman saw the man'. It is slightly less obvious in the case of adjuncts, but it becomes clear if one considers adjuncts as a syntactic subtype of the larger class of adverbials.

With very few exceptions, → adverbs (single-word adverbials) take the accusative case ending, e.g. *ṭabʿ-an* 'naturally', *sarīʿ-an* 'quickly', *ṣabāḥ-an* 'in the morning', *ġarb-an* 'westward'. Where an adverbial phrase has a nominal head, this typically also appears in the accusative, as in the case of → *ḥāl* clauses headed by a participle: *ġādara l-qāhira-t-a ʾamsi mutawajjih-an ʾilā jībūtī* 'he left Cairo yesterday, making for Djibouti' (Badawi a.o. 2004:156).

Adjuncts and nominal objects are both verb-dependent, and both are marked by the accusative. The two categories are, in the main, syntactically distinct. There are, however, occasions when nominal objects and adjuncts are syntactically nondistinct, as in Khartoum: *sakan-u l-ḥayy da* 'they settled this quarter'; *sakan-u fi l-ḥayy da* 'they settled in this quarter'. The adjunct prepositional phrase *fi l-ḥayy da* in the second example occupies the same syntactic position – that of the verbal complement – as does the object *al-ḥayy da* in the first example.

The examples *sakan-u l-ḥayy da* and *sakan-u fi l-ḥayy da* also illustrate a fairly common semantic difference between objects and adverbial complements. Objects tend to give a greater sense of exhaustion – or complete coverage – of the domain implied by the verbal element. The example *sakan-u l-ḥayy da* suggests that they settled (in) the entire quarter, while *sakan-u fi l-ḥayy da* implies more that they only settled (in) a part of it.

Cases in which objects alternate with adverbial complements are not the norm. Many verbs, described by the Arab grammarians as *ʾafʿāl mutaʿaddiya li-ġayrihā* 'verbs that pass on [to an object] through something other than themselves [viz. through a preposition]' (Wright 1975:II, 46; → *taʿaddin*), are able to take only an adverbial complement, as in Standard Arabic: *ġaḍiba ʿalā* 'to get angry with'; *šaʿara bi-* 'to know, feel'; *qadara ʿalā* 'to be able to do'. Some may take more than one adverbial complement, e.g. *samiḥa li-šaxṣ-in bi-šayʾ-in* 'to allow someone [to do] something'. And some may take a combination of adverbial complement(s) and object(s): *manaʿa šaxṣ-an min šayʾ-in* 'to prevent someone from [doing] something'.

Some verbs that take a preposition in Modern Standard Arabic took a direct object in Classical Arabic, as in the case of *ʿaliqa* 'to adhere' and *laḥiqa* 'to adhere to, overtake', both of which take the preposition *bi-* in Modern Standard Arabic (Wright 1975:II, 46).

4.2 *Objects and adjuncts in colloquial Arabic*

Modern dialects display the same general features as Standard Arabic in respect of objects and adjuncts. Perhaps more than Standard Arabic, some dialects show alternative transitivity patterns with the same verbal elements in the same basic senses. Examples from Khartoum include:

rašaḥ 'to drip'	
al-mawya rašaḥat min az-zīr	'the water dripped from the water container'
az-zīr rašaḥ mawya	'the water container dripped water'
jadaʿ	'to throw'
al-awlād jadaʿō-na bi-l-ḥijār	'the boys threw stones at us'
al-awlād jadaʿū fī-na l-ḥijār	'the boys threw stones at us' (lit. 'the boys threw us with stones')
naḍḍaf	'to clean'
naḍḍafta t-turāb	'I cleaned the dust

min aṭ-ṭarabēza	off the table'
naḍḍafta ṭ-ṭarabēza min at-turāb	'I cleaned the table of the dust'

Khartoum has also developed a generalized use of *fi* as an adverbial complement head, as an alternative to a nominal object, giving the sense of being engaged in an uncompleted activity:

bi-taktib/katabat jawābāt	'she's writing/wrote letters' [with some implication that she finished the letter-writing activity]
bi-taktib fi j-jawābāt	'she's [engaged in] letter writing'

5. Transitivity and semantic case roles

In typological terms, Arabic can be classed as an accusative language. Such languages encode the subject of an intransitive verb in the same way as the subject of a transitive verb and differently from the object of a transitive verb (Blake 1994). In this respect, Arabic is similar to most European languages. The subject of an active verb is typically an agent where action is involved, or an experiencer where subjective experience is involved.

Arabic has a large number of intransitive verbs expressing becoming a state. This may be an objective state, as in Standard Arabic *ixḍarra* 'to become green', *kabura* 'to become big/old'; or it may be a subjective state, e.g. *ḥizin* 'to become sad' (Khartoum), *ziʿil* 'to become angry' (Khartoum, Ṣanʿānī, Cairene). One class of transitive verbs expresses spatial and familial relationships: *girib* 'to be/become related to' (Khartoum), *mawwat* 'to sit beside [a dying person]' (Khartoum), *tawassaṭa* 'to be/become in the middle of' (Standard Arabic), *jāwara* 'to live close by (someone)' (Standard Arabic).

Temporality may be expressed through verbs in a way that is rather different from typical patterns in European languages. Thus, in Standard Arabic *ʾaṣbaḥa* 'to enter upon morning', *ʾamsā* 'to enter upon evening'. Through → semantic bleaching, both *ʾaṣbaḥa* and *ʾamsā* now convey the more common meaning of 'to become'. In Khartoum Arabic: *al-wāṭá maġrabat* 'evening has arrived' (lit. 'the ground has eveninged', *maġrab* being derived directly from *muġrib* '(time of) evening prayer').

A number of transitive verbs express the idea of action with incorporation of an adverbial element, e.g. in Standard Arabic *ʾaḥsana* 'to do well': *yuḥsin-u l-luġat-a l-ʿarabiyyat-a* 'he does [i.e. speaks] Arabic well'; *ʾakṯara* 'to do much': *yukṯir-u ziyārat-a-nā* 'he visits us often' (lit. 'he frequents our visiting'); in Khartoum Arabic: *tanna* 'to do a second time': *tanna l-gahwa* 'he drank a second coffee' (lit. 'he seconded the coffee'); in Ṣanʿānī: *azīdlak gahwih* 'I will give you more coffee'.

Arabic is fairly freely able to produce → causative and causative-type verbs, through the use of Form II or Form IV in Standard Arabic, and typically through the use of Form II in modern dialects. Sometimes these produce forms with a sense similar to that typically lexicalized in European languages, as in Standard Arabic, e.g. *ʾaʿlama* (Form IV) 'to inform' (normally doubly transitive), from *ʿalima* (Form I) 'to know' (normally singly transitive); *darrasa* (Form II) 'to teach' (doubly transitive), from *darasa* (Form I) 'to study' (singly transitive).

Similarly, in modern dialects: *sammaʿ* (Form II) 'to play [i.e. 'to cause to hear/listen to']' (doubly transitive), from *simiʿ* (Form I) 'to hear/listen to' (singly transitive; Khartoum); *garrā* (Form II) 'to teach' (doubly transitive), from *girī* (Form I) 'to learn' (Ṣanʿānī); *fawwat* (Form II) 'to let someone in', from *fāt* (Form I) 'to go in' (Çukurova; Procházka 2002:86).

Particularly in modern dialects, however, numerous causative and causative-type verbs have senses not typically lexicalized in European languages. Thus, in Khartoum: *garraṣ* (Form II) 'to cause to scratch [e.g. skin]' (doubly transitive), from *garaṣ* (Form I) 'to scratch' (singly transitive); *saffar* (Form II) 'to allow/cause to travel' (doubly transitive), from *sāfar* (Form III) 'to travel' (singly transitive), e.g. *nisaffir-ak* 'we'll give you money/enough money to travel' (lit. 'we will allow you to/make you travel') with the second object expressing the place to which the traveling is done omitted.

6. Transitivity and the passive

Intransitive verbs do not passivize, although one may note morphologically passive forms in Standard Arabic, such as *ġubiṭa* 'to be happy', which are not semantically simple passives of the corresponding active, in this case *ġabaṭa ʿalā* ‘to envy (someone)’. More striking are impersonal passive forms (cf. Sec. 3, above), such as *yuttafaq ʿalā hāḏihi l-qaḍāyā* ‘these matters are agreed on’, where the adverbial complement (*ʿalā hāḏihi l-qaḍāyā*) is retained in the passive, but no independent subject is possible. In the case of *ʾuġmiya ʿalā* (also *ġumiya ʿalā*) ‘to faint’, there is no corresponding active with the same sense as the impersonal passive form.

Singly transitive verbs standardly passivize, verbs expressing a relationship between two notions being an exception. Thus in Khartoum, *kaffa* ‘to be sufficient for’ does not have a passive **itkaffa*.

In the case of doubly transitive verbs, Classical Arabic allows only the indirect object to become the subject (Wright 1975:II, 52), as in: *ʿullima ʿilm-a l-hayʾat-i* ‘he was taught the science of astronomy’. Yet, according to some grammarians, in a sentence like *ʾaʿṭā zaydan dirhaman* ‘he gave Zayd a dirham’, both objects may be promoted to subjecthood, i.e., both *ʾuʿṭiya zaydun dirhaman* and *ʾuʿṭiya zaydan dirhamun* are permissible (Mubarrad, *Muqtaḍab* IV, 51, with a lengthy discussion of all possibilities). In the case of clauses with triply transitive verbs: *ʾaʿlama zayd-an ʿamr-an qādim-an* ‘he informed Zayd that ʿAmr is/was coming’, the only acceptable corresponding passive is: *ʾuʿlima zayd-un ʿamr-an qādim-an* ‘Zayd was informed that ʿAmr is/was coming’.

In many modern dialects, however, either the indirect or the direct object of the active clause may become the subject of the corresponding passive. In most dialects, when the nominal active direct object is the subject of the passive, the nominal active indirect object must be introduced by *li-*; by contrast, when the active indirect object becomes the subject of the passive, the direct object retains its object status. The examples below are from Cairene Arabic (cited from Woidich and Heinen-Nasr in Diem 2002):

samīr itfahhim id-dars	‘Samir was explained the lesson’
il-ʾawlād itʾakkilu fūl	‘the children were fed *fūl*’

but:

id-dars itfahhim li-samīr	‘the lesson was explained to Samir’
fūl itʾakkil li-l-ʾawlād	‘*fūl* was fed to the children’

While the above examples are intended to be illustrative of general tendencies, there are considerable differences in the ways in which transitivity interacts with the passive in different varieties of Arabic.

7. OBJECT-ORIENTED VS. *LI*-CONSTRUCTION VARIETIES

Some varieties of Arabic tend toward transitivity, while others tend toward intransivity, using a construction involving *li-*.

Khartoum Arabic tends markedly toward transitivity. Thus, for some Khartoum speakers, *nām* ‘to sleep’ and *gaʿad* ‘to stay; to sit’, verbs which are typically intransitive in modern Arabic dialects, are acceptable as transitive verbs: *numna l-ōḍa di* ‘we slept in that room’ (cf. the more normal *numna fi l-ōḍa di*); *gaʿadta l-bēt* ‘I stayed in the house’ (cf. the more normal *gaʿadta fi l-bēt*).

Speakers who accept objects in these cases also accept passive sentences in which the object of the active is the passive subject:

al-ōḍa di ma bi-ti-t-nām [ʿašān ḥārr]	‘that room can’t be slept in [because it is (too) hot]’
al-bēt da ma bi-t-gaʿid [ʿašān ḥārr]	‘that house can’t be stayed in [because it is (too) hot]’

Similarly, the verb *ja* ‘to come’ expresses the goal as an object in Khartoum, whereas other dialects, such as Ṣanʿānī, express the goal using a *li*-construction. Thus:

Khartoum

ja l-bēt	‘he came to the house’
al-ʿarabiyya jāt-ak	‘the car’s come for you’

Ṣanʿānī

jā li-l-bayt	‘he came to the house’
jaʾt la-k as-sayyārih	‘the car’s come for you [masc. sg.]’

In the case of doubly transitive verbs, word order and the grammatical status of the participants play a crucial role in whether both dependent participants function as objects or one is introduced by means of a *li*-construction. Many dialects allow two objects when both participants are nouns and the recipient (the indirect object) precedes:

zawwaj aḥmad bint-o	'he gave Ahmad his daughter in marriage' (Damascus; Diem 2002:52 from Cowell 1964)
warrēt is-suwwāḥ il-matḥaf	'I showed the tourists the museum' (Cairene; Woidich and Heinen-Nasr 2004:193)
dda l-bint al-kitāb	'he gave the girl the book' (Ṣanʿānī)
adda l-bitt al-kitāb	'he gave the girl the book' (Khartoum)

When the recipient does not come immediately after the verb, it must be introduced by *li-* (/*lē*):

adda l-kitāb li-l-bitt	'he gave the book to the girl' (Khartoum)
ddayt al-kitāb la-l-bint	'I gave the book to the girl' (Ṣanʿānī)
jawwaz bint-o la-ʾaḥmad	'he gave his daughter in marriage to Ahmad' (Palestinian; Diem 2002:56 from Elihai 1973)
warrēt il-matḥaf li-s-suwwāḥ	'I showed the museum to the tourists' (Cairene; Woidich and Heinen-Nasr 2004:193)

When the recipient is a pronoun, this is typically expressed as an object suffix in Khartoum and Damascus:

addā-ha l-kitāb	'he gave her the book' (Khartoum)
jawwaz-o bint-o	'he gave him his daughter in marriage' (Damascus; Diem 2002:54 from Cowell 1964)

Several dialects allow a pronominal recipient to be expressed either as an object suffix or to be introduced by *li-: warrā-hum il-matḥaf* 'he showed them the museum' or *warrā-l-hum il-matḥaf* (Cairene; Woidich and Heinen-Nasr 2004); *zawwuz-ih bint-ih* 'he gave him his daughter in marriage' or *zawwaz-l-ih bint-ih* (Eastern Libyan; Diem 2002:46 from Owens 1984). If the direct object is also a pronoun, this must then be introduced by *lē* (/*li*) in Khartoum, as in *addī-nī lē-hu* 'give me it [masc. sg.]!/give it [masc. sg.] to me!'.

Some dialects, however, allow pronominal direct and indirect objects. The direct object is suffixed directly to the indirect object in some dialects – Mekkan and Anatolian – and introduced by a reflex of *ʾiyyā-* in others – Palestinian and Damascene. Examples include: *ʿaṭāla-we* 'he gave it to her' (Anatolian; Diem 2002:60 from Jastrow 1978); *iddāha-huwwa* 'he gave it to her' (Mekkan; Diem 2002:39); *jawwazo yāha* 'he gave him her in marriage' (Damascene; Diem 2002:52 from Cowell 1964); *jawwazo iyyāha* 'he gave him her in marriage' (Palestinian; Diem 2002:56 from Elihai 1973).

Early Classical Arabic allowed free order between nominal indirect and direct objects. Thus, albeit with difference in emphasis, both of the following examples translate as 'he gave Zayd his daughter in marriage': *zawwaja zayd-an bint-a-hu* and *zawwaja bint-a-hu zayd-an*. In later Classical and Modern Standard Arabic, however, word order becomes more important, and nominal indirect and direct objects occur only when the indirect object follows the verb. Should the direct object follow the verb, the recipient must be introduced by *li-: zawwaja bint-a-hu li-zayd-in* 'he gave his daughter in marriage to Zayd'.

Classical Arabic of all eras allows a verb to take two pronoun objects as an alternative to affixation of the direct object pronoun to the particle *ʾiyyā*, as in: *isqinīhā* 'make me drink it!' (Badawi a.o. 2004:375). In Modern Standard Arabic, the pronouns may not be the same person (i.e. not both 1st, 2nd, or 3rd; Badawi a.o. 2004:375). Early Classical Arabic, however, occasionally allowed suffixation of two nonidentical 3rd person pronouns, as in: *ʾaʿṭaytuhumūhā* 'I gave it to them [masc.]' (Diem 2002:20); *ʾaʿṭāhāhu* 'he gave him to her' and *ʾaʿṭāhūhā* 'he gave her to him' (Wright 1975:II, 103), but in these cases introduction of the second (direct) object pronoun by means of the particle *ʾiyyā* was more common.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCE

Mubarrad, *Muqtaḍab* = ʾAbū l-ʿAbbās Muḥammad ibn Yazīd al-Mubarrad, *al-Muqtaḍab*. Ed. Muḥammad ʿAbd al-Xāliq ʿUḍayma. 4 vols. Cairo: Dār at-Taḥrīr, 1965–1968.

SECONDARY SOURCES

Badawi, Elsaid, Michael G. Carter, and Adrian Gully. 2004. *Modern Written Arabic: A comprehensive grammar*. London and New York: Routledge.

Blake, Barry J. 1994. *Case*. Cambridge: Cambridge University Press.

Cowell, Mark W. 1964. *A reference grammar of Syrian Arabic*. Washington, D.C.: Georgetown University Press.

Diem, Werner. 2002. *Translokative Verben im Arabischen: Eine diachronische Studie*. Wiesbaden: O. Harrassowitz.

Elihai, Yohanan. 1973. *Dictionnaire de l'arabe parlé palestinien: Français-arabe*. Paris.

Jastrow, Otto. 1978. *Die mesopotamisch-arabischen* qəltu-*Dialekte*. Wiesbaden: F. Steiner.

Owens, Jonathan. 1984. *A short reference grammar of Eastern Libyan Arabic*. Wiesbaden: O. Harrassowitz.

Procházka, Stephan. 2002. *Die arabischen Dialekte der Çukurova (Südtürkei)*. Wiesbaden: O. Harrassowitz.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London: Longman.

Reckendorf, Hermann. 1921. *Arabische Syntax*. Heidelberg: C. Winter.

Woidich, Manfred and Rabha Heinen-Nasr. 2004. Kullu Tamām: *An introduction to Egyptian Colloquial Arabic*. Cairo: American University in Cairo Press.

Wright, William. 1975. *A grammar of the Arabic language*. 3rd ed. Cambridge: Cambridge University Press.

JAMES DICKINS and JANET C. E. WATSON
(University of Salford)

Transitivity: Object

The → transitivity of a verb refers to whether or not it requires an object, where 'object' is taken here to represent a grammatical function in the sentence, in the same way → 'subject' does. Note that transitivity is not the same as → valency, even though the two terms are related. Transitivity is concerned with the relation between a verb and its object(s) only, whereas valency refers to the range of dependents that a verb may take, including its subject (see Payne 1997; Whaley 1997). While there is disagreement in the field of linguistic analysis as to whether grammatical functions such as subject and object are primitive notions, as proposed in Perlmutter's Relational Grammar, and also in Bresnan's (1982) Lexical Functional Grammar, or structurally defined, as in the Standard Theory of Chomsky (1965) and subsequent developments, a grammatical function still refers to the relation between a predicate and its dependents in sentence structure. Nothing in this entry hinges on resolving that question.

Sometimes the term → 'argument' is used to refer to nominals functioning as subjects and objects in sentences, for example to distinguish them from other optional syntactic elements like adverbials. These nonarguments are typically referred to as 'adjuncts'. Grammatical functions such as 'subject' and 'object' should therefore be distinguished from the semantic functions that arguments carry in sentence structure. These semantic functions are frequently referred to as 'thematic relations' or 'roles', and include relations such as Agent, Experiencer, Instrument, Theme, Goal, etc. This is important to note because the same thematic relation can be represented by an argument in either subject or object position. For example, a Theme argument typically occupies object position in active sentences, but occurs in subject position in passives. Consider, for example, the two sentences from Standard Arabic in (1).

(1a)	*qaraʾa*	*zayd-un*	*al-kitāb-a*
	read.3ms	Zayd-Nom	the-book-Acc
	'Zayd read the book'		
(1b)	*quriʾa*	*l-kitāb-u*	
	read.3ms.Pass	the-book-Nom	
	'The book was read'		

In both sentences, while the noun phrase *al-kitāb* represents the Theme of the event (i.e. the entity being affected by the action), it functions as object in (1a) and as subject in (1b). In Arabic traditional grammar, the passive subject is referred to as *nāʾib al-fāʿil*, which translates as 'subject by proxy' or 'substitute subject'.

This entry is organized as follows: in Section 1, the term 'transitivity' is dis-cussed, and examples of intransitive, transitive, and causative verbs in Standard Arabic are provided. In Section 2, the distinction between monotransitive and ditransitive verbs is discussed. In Section 3, some Arabic traditional grammatical notions regarding the transitivity of *believe*-type verbs and the so-called tritransitives are reviewed (→ *taʿaddin*). Section 4 illustrates with examples the other kinds of *mafāʿīl* 'objects' discussed by Arabic traditional grammarians, showing that they are all adverbial in function. In Section 5, the position of objects in Standard Arabic clause structure and its implications for current linguistic analysis is discussed. Section 6 is a summary.

1. Intransitives, transitives, and causatives

While all predicates select subjects, some predicates select objects while others do not. Verbs that require one or more objects are referred to as transitive verbs (and the clauses in which they appear are transitive clauses); those that do not require an object are intransitive (and the clauses in which they appear are intransitive clauses). For example, in Standard Arabic, the verb *nāma* 'to sleep' is intransitive, whereas the verb *wajada* 'to find' is transitive, as in (2).

(2a) *nāma* *l-walad-u*
slept.3ms the-boy-Nom
'The boy slept'
(2b) *wajada* *zayd-un* *al-kitāb-a*
found.3ms Zayd-Nom the-book-Acc
'Zayd found the book'

The templatic nature of Arabic verbal morphology allows intransitive verbs of Form I to be transitivized when used in the *ʾafʿala* (Form IV) or *faʿʿala* (Form II) template. For example, *jalasa* 'to sit down' is intransitive (3a), but *ʾajlasa* is transitive (3b). Similarly, *fariḥa* 'to be pleased' is intransitive (4a), but *farraḥa* 'to please someone' is transitive (4b).

(3a) *jalasa* *ṭ-ṭullāb-u* *fī*
sat.down.3ms the-students-Nom in
l-maktabat-i
the-library-Gen
'The students sat down in the library'
(3b) *ʾajlas-tu* *ṭ-ṭullāb-a* *fī*
seated-1s the-students-Acc in
l-maktabat-i
the-library-Gen
'I seated the students in the library'

(4a) *fariḥa* *zayd-un*
became.happy.3ms Zayd-Nom
bi-najāḥ-ī
with-success-cl1s
'Zayd was pleased by my success'
(4b) *farraḥ-tu* *zayd-an* *bi-najāḥ-ī*
pleased-1s Zayd-Acc with-success-cl1s
'I pleased Zayd by my success'

This same transitivization process can apply to transitive verbs of Form I. If these occur in the *ʾafʿala* or *faʿʿala* templates, their valency increases by one, typically rendering the verb → causative, hence appearing with two accusative-marked nominals. For example, the transitive verb *kataba* 'to write' appears with one accusative nominal (5a), whereas the causative verb *kattaba* 'to cause to write' appears with two accusative nominals (5b).

(5a) *kataba* *zayd-un* *ar-risālat-a*
wrote.3ms Zayd-Nom the-letter-Acc
'Zayd wrote the letter'
(5b) *kattab-tu* *zayd-an*
caused.to.write-1s Zayd-Acc
ar-risālat-a
the-letter-Acc
'I made Zayd write the letter'

The examples above show that nominal objects of transitive verbs are assigned accusative case in Standard Arabic and they typically follow the subject (but see below, Sec. 5, for cases where the object precedes the subject, either obligatorily or optionally).

While the object of many verbs is typically a noun phrase (NP), as illustrated by the examples in (2)–(5), a verb may also select a non-nominal object. For example, the verb *marra* 'to pass' selects a prepositional phrase (PP) as an object and is, therefore, sometimes referred to as a transitive-by-preposition verb, as in (6).

(6) *marar-tu* *bi-zayd-in*
passed-1s by-Zayd-Gen
'I passed by Zayd'

In such cases, the NP within the PP object is assigned genitive case by the preposition, and not accusative case by the verb.

Similarly, the verb *zaʿama* 'to claim' requires a clausal object introduced by the complementizer *ʾanna*, as in (7).

(7) *zaʿama* *zayd-un* *ʾanna*
claimed.3ms Zayd-Nom that
ṭ-ṭāʾirat-a *saqaṭat*
the-plane-Acc fell.down.3fs
'Zayd claimed that the airplane fell down'

In syntax, the term 'complement' is often used to refer to all these different kinds of objects that a verb may select. For example, the verb *wajada* in (2b) selects an NP complement, the verb *marra* in (6) selects a PP complement, and the verb *zaʿama* in (7) selects a clausal complement (or a Complementizer Phrase (CP), to use the technical term).

Note also that some verbs may function either transitively or intransitively, depending on whether they allow an implicit argument. For example, the verb *qaraʾa* 'to read' may be used with or without an NP object, as in (8).

(8a) *qaraʾ-tu* *l-kitāb-a*
read-1s the-book-Acc
'I read the book'
(8b) *qaḍay-tu* *ṣ-ṣabāḥ-a* *ʾaqraʾ-u*
spent-1s the-morning-Acc read.1s-Ind
'I spent the morning reading'

Similarly, some transitive-by-preposition verbs may in fact appear with an NP, rather than a PP, object. This is the case of the verb *daxala* 'to enter', as in (9).

(9a) *daxal-tu* *ʾilā* *l-bayt-i*
entered-1s to the-house-Gen
'I entered the house'
(9b) *daxal-tu* *l-bayt-a*
entered-1s the-house-Acc
'I entered the house'

Note that in the absence of the preposition, the NP *al-bayt-a* in (9b) appears in accusative, rather than genitive case (cf. (6) above).

To sum up, verbs may be either transitive or intransitive. If intransitive, they appear with only one argument, that is, the subject. If transitive, they appear with two arguments, fulfilling the two grammatical functions of subject and object. Objects can be NPs, PPs, or CPs. Arabic templatic morphology allows a verb to increase its valency by one, such that an intransitive verb becomes transitive, and an already transitive verb becomes causative when put in the *ʾafʿala* or *faʿʿala* templates. In all these cases of transitive clauses, an NP object always appears in accusative case.

2. MONOTRANSITIVE VS. DITRANSITIVE VERBS

Transitive verbs can be further classified into monotransitive and ditransitive verbs, the former selecting only one object, the latter selecting two objects. The verb *wajada* in (2b) is an example of monotransitivity, whereas the verb *ʾaʿṭā* 'to give' in (10) is an example of a ditransitive verb that selects two objects.

(10) *ʾaʿṭay-tu* *ʿaliyyan* *al-māl-a*
gave-1s Ali-Acc the-money-Acc
'I gave Ali the money'

Ditransitive structures like (10) are typically called 'double object constructions' (DOCs). In such structures, the first NP object is referred to as the 'indirect object', and the second one as the 'direct object'. Note from (10) that, unlike in several languages with double object constructions (e.g. Japanese), both direct and indirect objects of a ditransitive verb in Standard Arabic are marked accusative (rather than accusative and dative as in these other languages). Furthermore, in double object constructions, the typical word order is for the human indirect object to precede the nonhuman direct object, and the reverse order as in (11) is highly marked or rather marginal (indicated by ? at the beginning of the sentence).

(11) ?*ʾaʿṭay-tu* *l-māl-a* *ʿaliyy-an*
gave-1s the-money-Acc Ali-Acc
'I gave Ali the money'

Similarly, in passive structures, it is customary for the indirect object to become the subject of the passivized ditransitive verb, as in (12).

(12) *ʾuʿṭiya* *ʿaliyy-un* *al-māl-a*
gave.3ms.Pass Ali-Nom the-money-Acc
'Ali was given the money'

By contrast, passivization of the direct object is degraded, as in (13).

(13) *ʾuʿṭiya l-māl-u
gave.3ms.Pass the-money-Nom
ʿaliyy-an
Ali-Acc
'The money was given (to) Ali'

Note, finally, that double object constructions are often compared to the related prepositional dative construction (PDC) in (14), where the verb is followed by two complements: an NP and a PP.

(14) ʾaʿṭay-tu l-māl-a li-ʿaliyy-in
gave-1s the-money-Acc to-Ali-Gen
'I gave the money to Ali'

In prepositional dative constructions, unlike double object constructions, both complements may switch position without degrading the grammaticality of the sentence. For example, (15) is fully grammatical.

(15) ʾaʿṭay-tu li-ʿaliyy-in al-māl-a
gave-1s to-Ali-Gen the-money-Acc
'I gave the money to Ali'

Similarly, in contrast to double object constructions, passivization of the indirect object is nonproblematic in prepositional dative constructions, as in (16).

(16) ʾuʿṭiya l-māl-u
gave.3ms.Pass the-money-Nom
li-ʿaliyy-in
to-Ali-Gen
'The money was given to Ali'

In sum, transitive verbs may be either monotransitive or ditransitive, the former selecting only one object, the latter selecting two. Ditransitive verbs may either appear with two NPs, in what is called a double object construction, or with one NP and a PP, in which case the structure is referred to as the prepositional dative construction. In double object constructions, both objects appear in the accusative case, but the indirect object typically precedes the direct object in active sentences and functions as the subject in a passivized ditransitive clause. In prepositional dative constructions, by contrast, there is no restriction on word order, and it is the direct object that functions as subject in a passive clause.

3. *Ḏ̣ANNA*-TYPE VERBS AND THE SO-CALLED TRITRANSITIVE VERBS

While verbs like *to give* are truly ditransitive, given that the two objects they select represent two arguments of the verb, Arabic traditional grammarians such as Ibn ʿAqīl (d. 769/1367) and Ibn Hišām (d. 761/1360), among many others (see Ḥasan 1975), also talk about *believe*-class verbs, whose two objects are the subject and predicate of a → nominal clause, as being ditransitive verbs. An example of the usage of such verbs is given in (17).

(17) *ḏ̣anan-tu zayd-an musāfir-an*
believed-1s Zayd-Acc traveling-Acc
'I believed Zayd was traveling'

Verbs that behave like *ḏ̣anna* are traditionally assigned to three semantic classes: verbs of certainty like *ʿalima* 'to know'; verbs of probability like *ḏ̣anna* itself; and verbs of causation like *ṣayyara* 'to cause to become'. While *ḏ̣anna*-type verbs assign accusative case to both the subject and the predicate as in (17), there is a clear sense in which such verbs are not really ditransitive.

As noted earlier, a true ditransitive verb like *to give* semantically selects three arguments: Agent (the giver), Theme (the entity given), and Goal (the givee). Verbs like *to believe* in (17) select only two arguments: Experiencer (the believer) and Theme (the entity believed in). Informally put, when you believe that Zayd is traveling, you do not believe Zayd. Rather, you believe in the existence of a state of affairs (that of Zayd being a traveler). It is this state of affairs that actually represents the object of the verb *to believe*. Syntactically, such an object appears as an embedded nominal sentence in the sense of Arabic traditional grammar, or what is called a 'Small Clause' in current syntactic analysis. Verbs like *ḏ̣anna* are thus monotransitive, but their complement is a Small Clause that has both a subject and a predicate internal to it.

Believe-type verbs have always raised interesting questions for syntactic analysis because they are instances of Exceptional Case Marking

(ECM) constructions, i.e. structures in which the matrix verb assigns accusative case to an NP that it does not select as an argument. In the Arabic example in (17), the verb *ḏ̣anna* assigns accusative case to Zayd, even though Zayd is not an argument of the verb *ḏ̣anna* but rather the subject of the Small Clause and is an argument of the predicate *musāfir* within that Small Clause. That *ḏ̣anna* assigns accusative case to the subject of the Small Clause is evidenced by the fact that it becomes the subject of *ḏ̣anna* when passivized, as in (18).

(18) *ḏ̣unna* *zayd-un*
believed.3ms.Pass Zayd-Nom
musāfir-an
traveling-Acc
'It was believed that Zayd was traveling'

It is important to note here that the Exceptional Case Marking property of *ḏ̣anna*-type verbs also appears when the embedded clause is not a Small Clause, but a finite clause, as in (19).

(19) *ḏ̣anan-tu* *zayd-an* *sāfara*
believed-1s Zayd-Acc traveled.3ms
'I believed that Zayd had traveled'

Sentences such as (19) are typically referred to in the literature as 'object-raising constructions'. They have been reported in several languages; Kuno (1976) first observed them for Japanese; Davies (2005) discusses the phenomenon in Madurese; similar facts have been noted for Greek in Philippaki-Warburton (1987); and Bruening (2001) provides an analysis of the phenomenon in Passamaquoddy. For a discussion of object-raising constructions in Standard Arabic, see Soltan (2007).

Ḏ̣anna-type verbs can also select a clausal complement introduced by the complementizer *ʾanna*, as in (20).

(20) *ḏ̣anan-tu* *ʾanna* *zayd-an*
believed-1s that Zayd-Acc
sāfara
traveled.3ms
'I believed that Zayd had traveled'

Note here that the accusative case on the NP *zayd-an* in (20) is lexically assigned by the complementizer *ʾanna* and not by the matrix verb *ḏ̣anna*.

Note, finally, that some *ḏ̣anna*-type verbs may be further transitivized if they appear in the *ʾafʿala* form, thereby creating triple-accusative constructions, which have been traditionally described as constructions with tritransitive verbs, for instance by Ibn ʿAqīl and Ibn Hišām (see Ḥasan 1975). For example, the *ḏ̣anna*-type verb *ʿalima* 'to know' can be used with a Small Clause complement as noted above for *ḏ̣anna*, as in (21).

(21) *ʿalima* *ʿaliyy-un* *zayd-an*
knew.3ms Ali-Nom Zayd-Acc
musāfir-an
traveling-Acc
'Ali knew that Zayd was traveling'

If put further in the transitive *ʾafʿala* template, the original subject of *ʿalima* now becomes an NP complement in addition to the Small Clause complement, as shown by (22).

(22) *ʾaʿlam-tu* *ʿaliyy-an* *zayd-an*
informed-1s Ali-Acc Zayd-Acc
musāfir-an
traveling-Acc'
'I informed Ali that Zayd was traveling'

Structures such as (22) are obviously not tritransitive verbs, despite the presence of three accusative-marked nominals. In fact, we do not find instances of tritransitivity in human languages. Examples such as (22) are of a *ḏ̣anna*-type verb that is further transitivized by occurring in the *ʾafʿala* template. These are therefore transitive structures in which the predicate selects two complements: an NP and a Small Clause (the latter having a subject and a predicate internal to it as usual). Note that the accusatively marked *musāfir-an* in both (21) and (22) is indeed a predicate within a Small Clause structure and cannot be treated as an adverbial (or → *ḥal* in Arabic grammatical ter-minology) for two reasons. First, it modifies the NP *zayd-an*, and not the main verb of the sentence. Second, if it were an adverbial, we would expect it to be optional, contrary to fact, since (22) is ungrammatical without *musāfir-an*.

4. The term *mafʿūl* ‘object’ in traditional Arabic grammar

Traditional Arabic grammarians, such as Ibn ʿAqīl and Ibn Hišām, among many others (see Ḥasan 1975), also use the word → *mafʿūl* ‘object’ in a case-based sense: accusatively marked categories that loosely relate to the verb, though not selected by it, are typically referred to as *mafāʿīl* ‘objects’. These so-called *mafāʿīl* are not objects, however; rather, they are for the most part adverbial in function, as shown by the examples given below for each.

The first type of *mafʿūl* in the sense of traditional Arabic grammar is the cognate object (→ object, absolute), called *al-mafʿūl al-muṭlaq* ‘the absolute object’, which is usually the accusative form of the indefinite deverbal noun (known as → *maṣdar*) of the main verb, as in (23).

(23) *ġaḍib-tu* *ġaḍab-an* *šadīd-an*
got.angry-1s anger-Acc strong-Acc
‘I became very angry’

As shown by (23), the use of *al-mafʿūl al-muṭlaq* typically indicates the manner in which the event takes place, hence, its function is like that of an adverbial of manner.

A second type of *mafʿūl* in the traditional sense is *mafʿūl li-ʾajlihi* ‘object of purpose’, which is an accusatively marked nominal indicating the purpose of the event, as in (24).

(24) *waqaf-nā* *taʿḍ̱īm-an* *la-hu*
stood.up-1p reverence-Acc to-him
‘We stood up in reverence to him’

The third type is called → *mafʿūl fīhi* ‘object of time and place’, which indicates the place or time for the event denoted by the verb; it is also known as *ḍ̱arf* ‘adverbial [lit. ‘container’]’, as in (25).

(25a) *sa-nusāfir-u* *ṣabāḥ-a*
Fut-travel.1p-Ind morning-Acc
ġad-in
tomorrow-Gen
‘We will travel tomorrow morning’
(25b) *qābal-tu-hu* *ʾamām-a*
met-1s-cl3ms in.front.of-Acc
l-bayt-i
the-house-Gen
‘I met him in front of the house’

Finally, there is also the *mafʿūl maʿahu* ‘associative object’. It typically indicates an entity associated with the occurrence of the event and follows the particle *wa-*, as in (26).

(26) *mašay-tu* *wa-n-nahr-a*
walked-1s and-the-river-Acc
‘I walked by the river [lit. ‘I walked and the river’]’

As mentioned earlier, these *mafāʿīl* are treated as ‘objects’ in Arabic traditional grammar because they involve accusatively marked nominals that relate to an event denoted by the verb. It is clear from the examples, however, that none of these *mafāʿīl* represents an argument of the verb. Rather, they all indicate either the manner, purpose, time, or place of the event. Hence, they are adverbial in function.

5. Object shift in Standard Arabic

While the canonical position of the object in Standard Arabic is after the verb and the subject, there are cases where the object actually surfaces between the verb and its subject, either obligatorily or optionally. This is referred to here as ‘Object Shift’ (OS), a term commonly used to refer to object displacement in Icelandic and other Scandinavian languages (Holmberg and Platzack 1995; Thráinsson 2001). Obligatory Object Shift in Standard Arabic takes place when the object is a pronominal, in which case the only possible word order is VOS, as in (27).

(27) *qaraʾa-hu* *zayd-un*
read.3ms-cl3ms Zayd-Nom
‘Zayd read it’

Lexical NPs may optionally undergo Object Shift in Standard Arabic, whereby they surface in a position between the sentence-initial verb and the postverbal subject, as in (28b).

(28a) *qaraʾa* *zayd-un* *al-kitāb-a*
read.3ms Zayd-Nom the-book-Acc
‘Zayd read the book’
(28b) *qaraʾa* *l-kitāb-a* *zayd-un*
read.3ms the-book-Acc Zayd-Nom
‘Zayd read the book’

The interpretive effect of Object Shift in Standard Arabic is similar to that of the Icelandic type, since it is typically definite or specific DPs that undergo Object Shift. In information structure theoretic terms, a shifted object is taken to be 'given information'. This typically has the simultaneous effect of focalizing the subject NP as 'new information'.

Unlike what we see in Icelandic, however, Object Shift in Standard Arabic does apply to Prepositional Phrases, as shown by (29).

(29a) *taḥaddaṯa zayd-un maʿa ʿaliyy-in*
spoke.3ms Zayd-Nom with Ali-Gen
'Zayd spoke with Ali'
(29b) *taḥaddaṯa maʿa ʿaliyy-in zayd-un*
spoke.3ms with Ali-Gen Zayd-Nom
'Zayd spoke with Ali'

In fact, Object Shift can apply, to some degree, to CP complements of verbs as well. Consider (30).

(30a) *ʾablaġa-nī zayd-un ʾanna*
told.3ms-cl1s Zayd-Nom that
l-maṭār-a muġlaq-un
the-airport-Acc closed-Nom
'Zayd told me that the airport is closed'
(30b) *ʾablaġa-nī ʾanna l-maṭār-a*
told.3ms-cl1s that the-airport-Acc
muġlaq-un zayd-un
closed-Nom Zayd-Nom
'Zayd told me that the airport is closed'

Note that (29b) and (30b) are not cases of Subject Right Dislocation (SRD), since agreement on the verb is partial. As Ouhalla (1994) points out, in Subject Right Dislocation, agreement is always full, as shown by (31).

(31a) *taḥaddaṯ-ū maʿa ʿaliyy-in*
spoke-3mp with Ali-Gen
az-zumalāʾ-u
the-colleagues-Nom
'They spoke with Ali, the colleagues'
(31b) *ʾablaġ-ū-nī ʾanna l-maṭār-a*
told-3mp-cl1s that the-airport-Acc
muġlaq-un az-zumalāʾ-u
closed-Nom the-colleagues-Nom
'They told me that the airport is closed, the colleagues'

In (31), there is also a phonological break, just before the right-dislocated NP. Such a pause is absent in the non-Subject Right Dislocation structures in (29b) and (30b). For an elaborate discussion of Object Shift in Standard Arabic and its syntactic implications for the theory of grammar, see Soltan (2007).

6. CONCLUSION

In this entry, a descriptive characterization of transitivity in Standard Arabic has been given, with examples illustrating the difference between intransitives, monotransitives, and ditransitives, as well as causatives. The traditional classification of *ḏ̣anna*-type verbs as ditransitives has been reviewed, showing that their syntax and semantics indicate that they should be treated as monotransitive verbs selecting a Small Clause as a complement (on a par with what is found in other languages). The occurrence of multiple accusatives does not necessarily entail ditransitivity. Similarly, it has been argued that triple accusative constructions do not entail that the so-called tritransitive verbs exist. Such verbs are derived from *ḏ̣anna*-type verbs via the *ʾafʿala* template, hence selecting an NP argument in addition to a Small Clause. The broad class of objects in Arabic traditional grammar includes several types of adverbials, hence they are not really objects in the grammatical function sense used here. Finally, the phenomenon of Object Shift in Standard Arabic has been illustrated in obligatory and optional contexts.

BIBLIOGRAPHICAL REFERENCES

Bresnan, Joan (ed.). 1982. *The mental representation of grammatical relations*. Cambridge, Mass.: MIT Press.
Bruening, Benjamin. 2001. *Syntax at the edge: Cross-clausal phenomena and the syntax of Passamaquoddy*. Ph.D. diss., Massachusetts Institute of Technology.
Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Mass.: MIT Press.
Davies, William D. 2005. "Madurese prolepsis and its implications for a typology of raising". *Language* 81.645–665.
Ḥasan, ʿAbbās. 1975. *an-Naḥw al-wāfī*. Cairo: Dār al-Maʿārif.
Holmberg, Anders, and Christer Platzack. 1995. *The role of inflection in Scandinavian syntax*. New York and Oxford: Oxford University Press.
Kuno, Susumu. 1976. "Subject Raising". *Japanese generative grammar syntax and semantics*, V, ed. Masayoshi Shibatani, 17–49. New York: Academic Press.

Ouhalla, Jamal. 1994. "Verb movement and word order in Arabic". *Verb movement*, ed. David Lightfoot and Norbert Hornstein, 41–72. Cambridge: Cambridge University Press.

Payne, Thomas. 1997. *Describing morphosyntax: A guide for field linguists*. Cambridge: Cambridge University Press.

Perlmutter, David M. 1980. "Relational grammar". *Syntax and semantics: Current approaches to syntax*, XIII, ed. Edith A. Moravcsik and Jessica R. Wirth, 195–229. New York: Academic Press.

Philippaki-Warburton, Irene. 1987. "The theory of empty categories and the pro-drop parameter in Modern Greek". *Journal of Linguistics* 23.289–318.

Soltan, Usama. 2007. *On Formal Feature Licensing in Minimalism: Aspects of Standard Arabic morphosyntax*. Ph.D. diss., University of Maryland, College Park.

Thráinsson, Höskuldur. 2001. "Object shift and scrambling". *The handbook of contemporary syntactic theory*, ed. Mark Baltin and Chris Collins, 148–202. Oxford: Blackwell.

Whaley, Lindsay J. 1997. *Introduction to typology: The unity and diversity of language*. Thousand Oaks, Cal.: Sage Publications.

Usama Soltan (Middlebury College)

Translation Literature

This entry describes some of the linguistic features of the language of 8th-10th-century Greek/Arabic and Syriac/Arabic translations, links them to the history of the translation movement from Greek into Arabic, and attempts to place them in the context of contemporary linguistic change from 'Classical Arabic' to → 'Middle Arabic'.

1. The Greek/Arabic translation movement

The term 'Greek/Arabic translation movement' describes a wave of translations of Greek scientific and philosophical texts either directly into Arabic or by way of Syriac. Unsystematic translation activities probably date back to the early years of the Umayyad era (40/661–133/749 C.E.); information about these, however, is unreliable at best (Endress 1987:II, 418). The beginning of a concerted translation effort probably coincided with the advent of the Abbasid dynasty in 133/749 (Gutas 1998:2). For more than two years, the bulk of Greek medicine, science, and philosophy available to contemporary scholars was translated into Arabic.

The translation movement went through several phases, characterized by a growing refinement of translation methods and scientific and philosophical terminology (Endress 1987: III, 3–4). Chronologically, we can distinguish between several distinct groups of translators (for a historical sketch of the translation movement, see Endress 1987:II, 418–431). The division into four groups was suggested by Walzer (1970:32–35):

i. The first generations, working before the reign of the caliph al-Maʾmūn (r. 198/813–218/833), e.g. Ibn al-Muqaffaʿ (d. 139/756 or slightly later), al-Biṭrīq (d. ca. 200/815), and Theodore ʾAbū Qurrā (d. ca. 820). Most of their products are no longer extant.
ii. The circle of the philosopher ʾAbū Yūsuf Yaʿqūb ibn ʾIsḥāq al-Kindī (d. 260/870?). Among its likely members were ʾUsṭāṯ, Ibn Nāʿima al-Ḥimṣī, and possibly Yuḥannā ibn al-Biṭrīq. There are also contemporary translators, such as the Ḥarrānian Ṯābit ibn Qurra (d. 288/901), who apparently were not associated with any circle of translators. The members of the Kindī circle were close contemporaries of the third group.
iii. The group of translators working with Ḥunayn ibn ʾIsḥāq (d. 260/873 or 264/877). It included his son ʾIsḥāq ibn Ḥunayn (d. 298/910), his nephew Ḥubayš, ʿĪsā ibn Yaḥyā, and ʾAbū ʿUṯmān ad-Dimašqī (d. after 302/ 914). In addition to producing their own translations, they revised and updated numerous earlier ones.
iv. Later translators, e.g. ʾAbū Bišr Mattā ibn Yūnus (d. 328/940) and Yaḥyā ibn ʿAdī (d. 363/974). They worked after the zenith of the translation movement and used almost exclusively Syriac source texts, most of which were produced by Ḥunayn or his collaborators.

2. Describing the language of translation literature

This section and the following one rely heavily on translation analyses and terminological studies undertaken by the following authors: Afnan (1964:4, 26–27, 32, 99–100, 114–115); Alon (1985:167–168); Biesterfeldt (1973:18–

24); Daiber (1980:9, 20–23, 29–37, 62); Drossaart Lulofs and Brugman (1971:5–7, 18, 35–36); Endress (1966:34–38, 42–44, 63–72, 75–78, 85–86; 1973:154–155, 171; 1987:II, 418–419; III, 4–20; 1989:121–127, 132–134); Gätje (1967:356–360); Kruk (1979:27–28); Kunitzsch (1976:127–131); Ruland (1978:163, 196–200); Tkatsch (1928:148–150, 158–187). Other less frequently consulted sources are indicated in the text.

Many of the textual features listed below have been associated with a stratum of Arabic called 'Middle Arabic'. The use of the term is defined in Blau (1988a); the historical development of the underlying phenomenon is traced in Blau (1966–1967; 1999:1–18, 131–132, 213–228). Other commentators, treating the same phenomena without recourse to the category of 'Middle Arabic', have pointed out that translations and scientific texts display features such as 'demotic elements', which would have been regarded as grammatical mistakes or careless language from the point of view of 'Classical' Arabic usage. Zimmermann (1981: lxxvi) notes that since the translators' literary education was in Syriac, their written Arabic was much closer to the spoken Arabic used in daily intercourse.

The language of translation literature was formed by many factors. Some concern the material with which a translator worked: the quality of manuscripts and word and synonym lists; peculiarities of Greek paleography; or the character of the contents of the texts, ranging from dry and scientific to elaborate and highly literary. Some factors relate to a translator's background and could not be directly influenced: his mother tongue or dialect, which sometimes interfered with his output, his education, and his intellectual background; his experience and linguistic competence in the languages involved; and his exposure to previous translations or original Arabic texts from the same field or genre. Other factors could also play a role, for example a translator's style or his affiliation with certain schools of translators and use of their in-house methods, terminology, and phraseology. To this, we have to add a number of external factors, such as the requirements of patrons who commissioned translations (cf. Bergsträsser 1925:2–3, also nos. 5, 7, 11, 16, 38, and esp. 56) or, on a more general level, the prestige differential and variant field of use of the languages involved, which often influenced translational approaches. Some of the factors listed above made themselves again felt in the subsequent transmission of the texts, which often involved additional interventions by later translators, recipients, copyists, and redactors, as pointed out, among others, by Arnzen (1998:33).

In the majority of cases, we are unable to reconstruct a text's translation and transmission history with any precision and to distinguish between the potentially numerous layers of linguistic influences during translation, redaction, and transmission. These layers can sometimes be distinguished in a single text (cf. Ruland 1976:148–149; 1979:247). Lack of evidence also prevents us from extracting a reliable diachronic picture of the development of translation techniques, terminologies, and the like. Due to the possibility of later interventions in the texts, any information about translators and dating from secondary sources and the translations themselves has to be approached with caution (cf., e.g., Hugonnard-Roche 1990: 143).

Given the long and varied history of the translation movement, it is almost impossible to give a comprehensive description of the language of translation literature. Both the translators' methods and tools (e.g. word lists and dictionaries) and the concept of translation itself changed over time: at the beginning of organized translation activities, one encounters highly literal imitations of the word order of the source texts, but also free periphrastic adaptations. Later, translators strove for consistency in style and terminology and put more emphasis on both readability and faithfulness. Translations became more precise, smoother, and more 'Arabic' (cf. Bachmann 1965:9); older translated texts were often revised. For this reason, the following remarks can only offer an overview of some prominent linguistic phenomena found in translation literature, without reference to specific chronological phases of the translation movement. The categories selected are terminology, syntax, and phraseology/style.

3. Features of the language of translation literature

On numerous occasions, the translators had to deal with ambiguous syntactic structures, often choosing one of two or more possible interpretations. The very flexibility of Greek syntax also led to wrong reconstructions of the relation between the parts of clauses. On the other hand, syntactic ambiguities were also sometimes spelled out in the Arabic text with the help of syntactic 'doublets' that reflected both possible interpretations (Daiber 1980:31–32, 34).

The translations exhibit many syntactic traits which defy the rules of Classical Arabic. In general, complicated hypotactic Greek structures were often translated into simpler Arabic paratactic constructions, a phenomenon noted in a large number of translations. Translators frequently opted for a non-Classical Arabic word order. To put added emphasis on the subject, it was placed at the beginning, either in nominative or accusative case (preceded by *ʾinna* or *ʾanna*) and often introduced with the particle *ʾammā*. Even in unstressed positions, the subject could precede the predicate, and the object was sometimes placed before the subject. Some of these phenomena can be explained on the basis of the Greek (or Syriac) word order, which some translators slavishly reproduced, sometimes leading to semantic modifications and imprecise renderings. Another unusual syntactic structure was the sequence adverbial/prepositional clause – subject – verbal predicate. We also observe the frequent use of proleptic pronouns reflecting Syriac usage (Drossaart Lulofs and Brugman 1971:35–36; Endress 1966:34; 1987:8; Blau 1988b:215); and conjunctive particles such as *ʾinna/ʾanna*, etc. with the impersonal pronominal suffix *-hu* (Kruk 1979:27). A prominent characteristic of many translations is the use of the particle *qad* with an imperfect verbal form in assertoric statements (but see Biesterfeldt 1973:21 who maintains that this use denotes potentiality).

Subordinate clauses often diverge from the formal requirements of Classical Arabic. In conditional clauses, the verbal predicate of the apodosis frequently occurs in the indicative instead of either apocopate or perfect. At the same time, the required particle *fa-* at the beginning is often omitted (Drossaart Lulofs and Brugman 1971:36). In relative clauses, the required *ʿāʾid*, the pronoun referring back to the *nomen rectum* in the superordinated clause, was often dropped (Kruk 1979:27). Greek (real) conditional clauses were translated with Arabic conditional clauses introduced by *ʾin* or divided into an assertion (the former protasis) and a clause expressing the apodosis and bound to the protasis with certain formulas, whereas (irreal) conditional clauses were replaced with Arabic clauses with the particle *law* or similar constructions. The consecutive sense of consecutive clauses introduced in Greek by *hōste* was captured with various adverbial phrases or paraphrased (Endress 1966:76ff.).

Linguistic interference led to Arabic syntactic structures modeled on Greek and Syriac precedents. Thus, for instance, combinations of prepositions and *ʾan* and *ʾanna* fronting object clauses occur, clearly as a parallel to Syriac combinations of the particle *de-* with prepositions. Also, translators frequently employed *mā* instead of *ʾan* in conjunctional phrases, again reflecting the Syriac particle *de-*. The Syriac preposition *be-* influenced the use of *fī* as in *al-ʿilla fī ḏālika* 'the reason for this is' (Endress 1966:34). Finally, the Syriac case particle *le-* was sometimes replaced with the Arabic preposition *li-* without regard for the syntactic consequences (Tkatsch 1928:150). The formulaic expression *wa-ʾayḍan (fa-)* introducing a new section (for Greek *etí, etí dé* etc.) mirrors Syriac *ṭūb/ṭūb dēn* (cf. Georr 1948:71). Greek negative conditionals were sometimes rendered with clauses introduced by *ʾillā ʾan(na)*, occasionally elliptically shortened to *ʾillā*, also an echo of Syriac syntactic structures (Lyons 1966:xvi).

Unusual constructions abound, some of them associated with 'Middle Arabic' (cf. above). One finds inconsistencies in the use of the apocopate after *lam* or the accusative case after *laysa*, which is used like *lā* as a mere negation particle. Equally inconsistent is the occurrence of the accusative case after *sammā* 'to name' (Daiber 1980:7) or *ʾinna/ʾanna* (Daiber 1980:8), and in adverbial constructions (Drossaart Lulofs and Brugman 1971:35). Another non-Classical Arabic feature of some translations is the annexation of two (or even more) consecutive nouns in the construct state to one *nomen rectum* (Blau 1967, par. 22g). Some translators used phrases such as *bi-ḥaqqin* and *bi-ḍṭirārin*, where Standard Arabic would prescribe determination; or expressions such as *ġazā l-ʾawwal* instead of *al-ġazā l-ʾawwal*. The concord between noun and adjective attribute was apparently no longer regarded as binding (Drossart Lulofs and Brugman 1971:35–36), and, perhaps as a consequence, numerals sometimes end in *tāʾ marbūṭa* regardless of the gender of the counted noun (Blau 1988b:17; 1999:131; Kruk 1979:27). Also unusual is the use of the pronoun *kaḏā* in conjunction with *ʾan* and a conjunctive verb form. The object or psychological predicate of an *ʾammā* clause was frequently introduced with the particle *fa-* (Daiber 1980:8).

The complexities of the Greek verbal system required some ingenuity to interpret and translate because the Arabic verbal system often did not supply a straightforward equivalent. This was a frequent source of misleading translations. Paraphrase was often the only recourse, as in the case of gerundive verbal adjectives on *-téos* and *-tós*, or adhortative conjunctive verbal forms such as *legōmen* 'let us say', paraphrased with forms of *ʾarāda* 'to want', or other such expressions (Tkatsch 1928:168, 187; Endress 1966:75; 1987:III, 10–11; Ruland 1978:196). Occurrences of *estō* 'let it be' to express an assumption were rendered with the apocopate (Endress 1966:75); the Greek potentialis, constructed with optative and *án*, was paraphrased with verbal or adverbial phrases (cf. Bachmann 1965:9). To translate the Greek future tense, the translators opted for the emphatic use of the particle *qad* with a verbal predicate in the imperfect tense or the proclitic particle *li-* in conjunction with the apocopate. Also, they used phrases such as *al-battata* (Tkatsch 1928:165; Bachmann 1965:9; Endress 1966:75–76; 1987: III, 9).

As with other features of the language of translation literature, its terminology evolved over time. Our textual evidence suggests that earlier translations were produced while the development of an Arabic philosophical and scientific vocabulary was in its infancy. For many scientific and philosophical ideas introduced by the translation movement, the Arabic language lacked a preexisting technical terminology. In fact, the translations themselves often spearheaded its development. In the course of the translation movement, numerous terms were tried and discarded. The terminology for a number of fields finally matured and stabilized with the translators of Ḥunayn's generation. Most later translations and scientific and philosophical texts make use of this 'standardized' vocabulary (Rosenthal 1975:8). The continuous process of revision and alteration some translations underwent could, however, cause a certain degree of terminological contamination.

To translate terms for which there was no Arabic equivalent, translators had recourse to a number of strategies: transliteration, loan translations (calques), and the use of old or newly coined genuine Arabic terms.

Instances of transliterated terms are mostly found in older translations. Some survived and became part of the standardized vocabulary of the sciences; most, however, were subsequently replaced. The phenomenon of loanwords from the Greek was familiar from Syriac and Persian. Unsurprisingly, a number of them were taken over by the translators, e.g. *jins* 'genus' < Syriac *gensā* < Greek *génos*; or *faylasūf* 'philosopher', modeled on the Syriac term *pīlōsōpā* < Greek *philósophos*. Prior Syriac translations (as opposed to mere transliterations) of Greek technical terms also made their way into Arabic, such as *ʿilla* 'reason' < Syriac *ʿelltā*, which translated Greek *aítion*. The Persian language was the source for terms such as *jawhar* 'substance' < *gōh(a)r*, or *dūlāb* 'wheel'.

Loan translations were particularly prominent in medical terminology. They were coined either by analyzing a source term and translating its components or by isolating its supposed etymological core. The Greek group of terms *lógos, logikós*, and *logikē* is a prominent example: on the basis of etymological analogy, these were rendered as *nuṭq, nāṭiq*, and *manṭiq*, respectively. Composite terms were often paraphrased; equally as often, components of composites were suppressed, and only parts were actually translated.

Lastly, translators created new, genuinely Arabic terms, either by endowing existing words with a new meaning or by deriving new words from existing Arabic roots. This method was used from the very beginning of the translation movement and became more important over time: older transliterated terms were often replaced by genuine Arabic ones. In both scientific and philosophical translations, terms were coined by abstracting from the concrete meaning of a word. The old term *sabab*, originally denoting 'rope', 'link', or 'means; expedient', was used to translate *aítion* 'cause'. The aforementioned Persian loanword *jawhar*, originally denoting 'jewel', became the technical term for *ousía* 'essence' or 'substance'. Newly derived expressions often took the form of a verbal noun (*maṣdar*), mainly for processual terms, or, in the case of abstracts, the suffix *-iyya* was added to a concrete term. Examples for the former are *jadal* 'debate' (originally 'to tighten a rope') for *dialektikē*, and *ʾiḍāfa* 'relation' (originally 'making or declaring someone a neighbor or guest') for *prós ti*. The latter mechanism was a frequent occurrence already in earlier translations originating from the circle of the philosopher al-Kindī. Examples are *kammiyya* 'quantity' (from *kam* 'how much?') and *kayfiyya* 'quality' (from *kayfa* 'how?').

A prominent feature of the terminology of translation literature from its very beginning was the frequent use of synonymic doublets (cf. Thillet 1997). Their main function was either to reproduce the semantic range of a Greek source term with the highest possible precision or to emphasize it, especially after a negation. Early in the translation movement, their use was possibly born out of the translators' insecurity about the exact meaning of a Greek term; later, they were often consciously employed with a view to terminological accuracy. The systematic use of synonyms was not entirely unprecedented, since it was a characteristic trait of Syriac translations and a recurring phenomenon in both Syriac and Aramaic texts.

In terms of phraseology, Ancient Greek offered a varied and complex system of particles to lend emphasis and to structure sentences and arguments. One of the most challenging translation tasks was replicating both their structural component and other shades of meaning acquired in certain contexts. In Syriac, translators could utilize a range of existing particles or coin new ones – close counterparts to many Greek particles are found, e.g. *gár* > *gēr, mén* > *men, dé* < *dēn* (Tkatsch 1928:171; Endress 1966:37–38); The Arabic language proved to be more resilient to such changes (Zimmermann 1981:lxxvi–lxxvii): for the most part, translators had to work within the confines of the Arabic system of particles. Particularly in early texts, one observes that Greek particles are frequently omitted. The translators seem to have regarded their semantic impact, apart from a certain emphasis on words and clauses, as negligible. Later translators, however, while not always preserving every single particle, took pains to identify and express their exact meaning and structural role. Another strategy was to coin stock equivalents, sometimes consisting of more or less elaborate circumlocutions and to employ them schematically for every occurrence of the Greek particle in question.

Syriac-inspired phraseology is another characteristic of translation literature. Expressions such as *huwa fa-huwa* (Kruk 1979:28) or *huwa huwa*, translating *ho autós* or *tautó*, mirror the corresponding Syriac expression *hu kaḏ hu* (Tkatsch 1928:178; Drossaart Lulofs and Brugman 1971:5).

Over time, a number of translators developed a repertory of elaborate phrases and expressions modeled on the function of Greek particles, i.e. to stress elements of clauses, entire clauses, or arguments and indicate logical relations between them. As with the terminology of the translations, the phraseology at the beginning was uneven and inconsistent. Only in later texts, it became elaborate and was consistently applied throughout. Neither the influence of a particular group of translators and their style, nor Syriac source texts, can explain the full range of this phrase repertory. For some of it, one has to turn to Greek authors from Late Antiquity, who employed a similarly complex system of formulas. They are thus indicators for an unbroken tradition of scholarship that supplied translators with at least part of the phraseology observed in their products, as well as in the works of later commentators. In addition, one finds traces of personal stylistic idiosyncrasies of certain translators. Into this group belongs, for instance, Qusṭā ibn Lūqā's (d. ca. 300/912) phrase *raʾaynā ʾan yajiba ḍṭirāran ʾan nuqaddima* 'we think it [is] absolutely necessary to mention at the outset', which is found in at least two of his translations (Daiber 1980:8).

In sum, the language of translation literature reflects several aspects of linguistic change compared to Classical Arabic. On the one hand, it was part of the wider development of Arabic, displaying many elements classified as 'Middle Arabic'. In view of the religious affiliation of the almost exclusively Christian translators, the Christian Arabic dialects in particular made themselves felt. Also, it was deeply influenced by the linguistic background of the translators: their → Syriac mother tongue left its imprint on many texts, either in the form of traces of a Syriac intermediary translation or as Syriacisms involuntarily introduced by scholars whose native language was Syriac and who often came from the least Arabicized sections of the Christian community (Zimmermann 1981:lxvi).

BIBLIOGRAPHICAL REFERENCES

Afnan, Soheil. 1964. *Philosophical terminology in Arabic and Persian*. Leiden: E.J. Brill.

Alon, Ilai. 1985. "The Arabic version of Theophrastus' *Metaphysica*". *Jerusalem Studies in Arabic and Islam* 6.163–217.

Arnzen, Rüdiger. 1988. *Aristoteles* De Anima*: Eine verlorene spätantike Paraphrase in arabischer und persischer Überlieferung*. Leiden: E.J. Brill.

Bachmann, Peter. 1965. *Galens Abhandlung darüber, daß der vorzügliche Arzt Philosoph sein muß*. (= *Nachrichten der Akademie der Wissenschaften in Göttingen*, I: Phil.-Hist. Kl. Jg. 1965, Nr. 1.) Göttingen.

Bergsträsser, Gotthelf. 1925. *Ḥunayn ibn Isḥāq über die syrischen und arabischen Galen-Übersetzungen*. (= *Abhandlungen für die Kunde des Morgenlandes* 17/2.1–49.) Leipzig.

Biesterfeldt, Hans. 1973. *Galens Traktat* 'Daß die Kräfte der Seele den Mischungen des Körpers folgen' *in arabischer Übersetzung*. (= *Abhandlungen für die Kunde des Morgenlandes* 40/4.) Wiesbaden.

Blau, Joshua. 1966–1967. *A grammar of Christian Arabic, based mainly on South-Palestinian texts from the first millenium*. Louvain: Secretariat du CorpusSCO.

——. 1988a. "Classical Arabic, Middle Arabic, Middle Arabic Literary Standard, Neo-Arabic, Judaeo-Arabic and related terms". Blau (1988c:255–259).

——. 1988b. "The importance of Middle Arabic dialects for the history of Arabic". Blau (1988c:206–229).

——. 1988c. *Studies in Middle Arabic and its Judaeo-Arabic variety*. Jerusalem: Magnes Press.

——. 1999. *The emergence and linguistic background of Judaeo-Arabic: A study of the origins of Neo-Arabic and Middle Arabic*. 2nd ed. Jerusalem: Ben Zvi Institute.

Bouyges, Maurice (ed.). 1938–1952. *Averroès*: Tafsīr mā baʿd aṭ-ṭabīʿat. Beirut: Imprimerie Catholique.

Brock, Sebastian. 1984. "Aspects of translation theory in Antiquity". *Syriac perspectives on late Antiquity*, ed. Sebastian Brock, 69–87. London: Variorum Reprints.

——. 1991. "The Syriac background to Ḥunayn's translation technique". *Aram* 3.139–162.

Daiber, Hans. 1980. *Aetius Arabus: Die Vorsokratiker in arabischer Überlieferung*. Wiesbaden: F. Steiner.

Drossaart Lulofs, Hendrik Joan and Jan Brugman (eds.). 1971. *Aristotle,* Generation of animals*: The Arabic translation commonly ascribed to Yaḥyā ibn al-Biṭrīq*. Leiden: E.J. Brill.

Endress, Gerhard. 1966. *Die arabischen Übersetzungen von Aristoteles' Schrift* De Caelo. Ph.D. diss., University of Frankfurt/Main.

——. 1973. *Proclus Arabus: Zwanzig Abschnitte aus der* Institutio Theologica *in arabischer Übersetzung*. Beirut: Imprimerie Catholique.

——. 1987. "Die wissenschaftliche Literatur". *Grundriß der arabischen Philologie*, ed. Helmut Gätje, II, 400–506, III, 3–152. Wiesbaden: L. Reichert.

——. 1989. "Die griechisch-arabischen Übersetzungen und die Sprache der arabischen Wissenschaften". Endress and Schmeink (1989:103–145).

——. 1997. "The circle of al-Kindī: Early Arabic translations and the rise of Islamic philosophy". *The Ancient tradition in Christian and Islamic Hellenism: Studies in the transmission of Greek philosophy and sciences*, ed. Gerhard Endress and Remke Kruk, 43–76. Leiden: Research School CNWS.

——. and Dimitri Gutas (eds.). 1992–. *A Greek and Arabic lexicon: Materials for a dictionary of the mediaeval translations from Greek into Arabic*. Leiden: E.J. Brill.

Endress, Gerhard and Marita Schmeink (eds.). 1989. *Symposium Graeco-Arabicum*, II. Amsterdam: B.R. Grüner.

Gätje, Helmut. 1967. *Die arabische Übersetzung der Schrift des Alexander von Aphrodisias über die Farbe*. (= *Nachrichten der Akademie der Wissenschaften in Göttingen*, I: Phil.-Hist. Kl. Jg. 1967, Nr. 10.) Göttingen.

Georr, Khalil. 1948. *Les* Catégories *d'Aristote dans leurs versions syro-arabes*. Beirut: Institut Français de Damas.

Graf, Georg. 1905. *Der Sprachgebrauch der ältesten christlich-arabischen Literatur*. Leipzig: O. Harrassowitz.

Gutas, Dimitri. 1998. *Greek thought, Arabic culture*. London and New York: Routledge.

Hugonnard-Roche, Henri. 1990. "Les traductions du grec au syriaque et du syriaque à l'arabe". *Rencontres de cultures dans la philosophie médiévale: Traductions et traducteurs de l'antiquité tardive au XIV^e^ siècle*, ed. Jacqueline Hamesse and Marta Fattori, 131–147. Louvain-la-Neuve: Publications de l'Institut d'Études Médiévales.

——. 1994. "La formation de la vocabulaire de la logique en arabe". *La formation du vocabulaire scientifique et intellectuel dans le monde arabe*, ed. Danielle Jacquart, 22–38. Turnhout: Brepols.

Kruk, Remke. 1979. *The Arabic version of Aristotle's* Parts of animals*: Book XI–XIV of the* Kitāb al-ḥayawān. Amsterdam and Oxford: North-Holland.

Kunitzsch, Paul. 1976. "Zur Problematik und Interpretation der arabischen Übersetzungen antiker Texte". *Oriens* 25–26.116–132.

Lyons, Malcolm (ed.). 1966. *Kitāb tadbīr al-ʾamrāḍ al-ḥādda* li-Buqrāṭ (*Hippocrates:* Regimen in acute diseases). Cambridge: Heffer.

Mattock, John. 1989. "The early translations from Greek into Arabic: An experiment in comparative assessment". Endress and Schmeink (1989:73–102).

Rosenthal, Franz. 1975. *The Classical heritage in Islam*. London: Routledge and Kegan Paul.

Ruland, Hans-Jochen. 1976. *Die arabischen Fassungen von zwei Schriften des Alexander von Aphrodisias:* Über die Vorsehung *und* Über das liberum arbitrium. Ph.D. diss., University of Saarbrücken.

——. 1978. *Die arabische Übersetzung der Schrift des Alexander von Aphrodisias über die Sinneswahrnehmung*. (= *Nachrichten der Akademie der Wissenschaften in Göttingen*, I: Phil.-Hist. Kl. Jg. 1978, Nr. 2.) Göttingen.

——. 1979. *Zwei arabische Fassungen der Abhandlung des Alexander von Aphrodisias über die*

Universalia (Quaestio I 11a). (= *Nachrichten der Akademie der Wissenschaften in Göttingen*, I: Phil.-Hist. Kl. Jg. 1979, Nr. 10.) Göttingen.
Thillet, Pierre. 1997. "Réflexions sur les 'traductions doubles'". *Perspectives arabes et médiévales sur la tradition scientifique et philosophique grecque*, ed. Ahmad Hasnawi, Abdelali Elamrani-Jamal, and Maroun Aouad, 249–263. Louvain and Paris: Peeters.
Tkatsch, Jaroslaus. 1928. *Die arabische Übersetzung der* Poetik *des Aristoteles und die Grundlage der Kritik des griechischen Textes*. Vienna and Leipzig: Hölder-Pichler-Tempsky.
Ullmann, Manfred. 2002. *Wörterbuch zu den griechisch-arabischen Übersetzungen des 9. Jahrhunderts*. Wiesbaden: O. Harrassowitz.
Walzer, Richard. 1962. *Greek into Arabic: Essays on Islamic Philosophy*. Oxford: B. Cassirer.
——. 1970. "L'éveil de la philosophie islamique". *Revue des Études islamiques* 38.7–42, 207–242.
Zimmermann, Fritz. 1981. *Al-Farabi's commentary and short treatise on Aristotle's* De interpretatione. London and New York: Oxford University Press.

Uwe Vagelpohl
(Hampshire College)

Transliteration → Transcription

Tripoli Arabic

1. General

Tripoli, the Mediterranean port and de facto capital of the Great Socialist People's Libyan Arab Jamahiriya, is located in northwestern Libya, around 100 km from the Tunisian border. With a population of 1.7 million, it is the country's most populous city. Vernacular Arabic is the spoken mother tongue of its inhabitants. This dialect, like others in the Maghrebi group, is characterized by the prefix *n-* for the 1st person singular and the prefix *n-* with the suffix *-u* for the 1st person plural of the imperfect of the verb; this difference in 1st person verbal forms is regarded as the main distinctive isogloss between Maghrebi and Near Eastern dialects.

Tripoli is a city of historical importance in the Maghreb. A pre-Hilalian dialect, linked to the Arab conquest and the first wave of Arabization in the 7th century C.E., is attested in the literature (Rossi 1965; Mantran 1975; Féraud 2005). In the 11th century C.E., Bedouin influence came to predominate with the invasions of the Banū Hilāl and Banū Sulaym. These invasions produced a far-reaching Arabization of the eastern Maghreb. During the Ottoman period (1535–1911) and the subsequent Italian occupation (1911–1943), the spoken Arabic of Tripoli underwent further changes. Until the independence of Libya in 1951, Tripolitania was managed by the British under mandate of the United Nations (1943–1951); the British occupation, too, affected the spoken Arabic of Tripoli considerably. During the late 20th century, massive rural migration to the capital of the new Libyan state had a major impact on language. Today, the spoken Arabic of Tripoli may be considered a Bedouin dialect with certain pre-Hilalian urban characteristics. For the Jewish Arabic dialect of Tripoli, see Yoda (2005).

2. Linguistic description

2.1 *Phonology*

2.1.1 Consonants

2.1.1.1 Inventory

Tripoli Arabic has the 33 consonantal phonemes in Table 1.

The reflex of *_q_ is predominantly voiced /g/, which reveals the nomadic origin of the dialect. Examples: *gʕəd* 'he stayed', *gərgūš* 'dried meat', *nəggəl* 'he copied', *məštāg* 'nostalgic'. In the mainstream Tripoli dialect, a phoneme /q/ exists in certain words and forms minimal pairs with /g/, such as *stəgṣa* 'he went away': *stəqṣa* 'he found out'; *nəggāl* 'cheat; someone copying': *nəqqāl* 'mobile phone'; *ḥəgg* 'price': *ḥəqq* 'truth'. The /q/ is also present in certain words drawn from the religious and legal registers of written Arabic and in terms from education and new technologies, including *qāḍi* 'judge', *muṛāhəq* 'adolescent, teenager', *təḥqīq* 'questioning'. In the pre-Hilalian urban dialect of the Jews of Tripoli, the nonvoiced plosive pronunciation [q] of *_q_ remained (Yoda 2005).

For an urbanized Bedouin dialect of this type, one would expect the interdental fricatives to be preserved, but in fact they are merged with the dental stops. The unvoiced interdental fricative *_ṯ_ developed into the unvoiced dental plosive /t/, while the voiced interdental fricative *_ḏ_ merged with the voiced dental plosive /d/, e.g. * _ṯalāṯa_ > _tlāta_ 'three', *_ṯalj_ > _təlž_ 'snow', * _ḏahab_

Table 1. Inventory of consonantal phonemes in Tripoli Arabic

	bilabial	labiodental	alveolar	postalveolar	palatal	velar	uvular	pharyngeal	laryngeal
plosive	b		t, d ţ			k, g	q		
emphatic	ḅ		ṭ, ḍ						
nasal	m		n						
emphatic	ṃ								
fricatives		f, v	s, z	š ž	č, j	x, ġ		ḥ, ʿ	h
emphatic			ṣ, ẓ						
trill		r							
emphatic			ṛ						
lateral			l						
emphatic									
glides	w				y				

> *dhəb* 'gold', **ʾaxaḏ* > *xde* 'he took'. There is only one emphatic voiced plosive, /ḍ/, resulting from the merger of **ḍ* and **ḏ̣*, e.g. **marīḍ* > *mrīḍ* 'ill', **ḏ̣ull* > *ḍull* 'shade'. The development of the interdental fricatives to dental plosives is an innovation that can be observed in other pre-Hilalian sedentary dialects as well.

/ţ/ [t͡s] is a dental affricate with marginal occurrence in Italian loanwords such as *bīţa* 'pizza'.

The reflex of **j* is a voiced postalveolar fricative /ž/, as is the case in almost all Bedouin dialects, e.g. *žāwəb* 'he answered', *ṛāža* 'he waited', *ḥāžəb* 'eyebrow', *drūž* 'stairs'. In contrast to Old Arabic /j/, /ž/ is a 'sun letter' and assimilates the /l/ of the article *əl-*, e.g. *əžžnāža* 'the funeral', *əžžūʿ* 'hunger', *əžžbən* 'the cheese'; there is one frequent exception: *əlžəww* 'the environment'.

The /r/ is produced as an apico-alveolar [r], e.g. *rme* 'he threw', *ržəm* 'he stoned', *tərrās* 'man', *ṣġīr* 'small'. There is also an emphatic phoneme /ṛ/. Note the following minimal pair: *dār* [dæːr] 'he did' vs. *dāṛ* [d̴ɑːr̴] 'room'.

Some phonemes have a marginal occurrence in Italian and English loanwords: /v/ [v] is a labiodental fricative, e.g. *vīti* 'screw', *səyyəv* 'to save'; /j/ [d͡ʒ] is a palatal fricative, e.g. *jībōṭo* 'jacket', *junṭ* 'joint'; /č/ [t͡ʃ] is a palatal fricative, e.g. *kāčāvīti* 'scewdriver', *sərč* 'internet search'.

2.1.1.2 Change of sibilants

The postalveolar sibilants /š/ and /ž/ are subject to various conditioned changes. When /š/ and /ž/ are found in the same word, or in cases where one of these phonemes is found in proximity to alveolar sibilants /s/, /ṣ/, or /z/ in the same word, assimilation, and metathesis can be observed.

Postalveolar /ž/ assimilates to sibilants /s/, /z/, or /ṣ/.

i. *s – ž > s – z*
sfənž > *sfənz* 'fritter'
sərž > *sərz* 'saddle'
ii. *z – ž > z – z*
zōž > *zōz* 'two'
zəllīž > *zəllīz* 'tiled floor'
tzəwwəž > *tzəwwəz* 'he married' [old-fashioned, elderly persons]
iii. *ž – s > z – s*
žəns > *zəns* 'species, kind'
iv. *ž – ṣ > z – ṣ*
nžāṣ > *nzāṣ* 'pears'
v. *ž – z > z – z*
ʿžūz > *ʿzūz* 'old woman'
žuzzāṛ > *zuzzāṛ* 'butcher'

Assimilation also takes place when the sibilant /s/ is followed by the postalveolar /ž/; this occurs with words from the root *s – ž – l*.

səžžəl > *šəžžəl* 'he registered'
musəžžəl > *mušəžžəl* 'tape recorder'

Metathesis takes place when postalveolar /ž/ and sibilant /z/ are in the same word. When /z/ is followed by /ž/, the sequence /z-ž/ is changed into /ž-z/. This only occurs with words from the root *z-w-ž*.

z – ž > ž – z [more modern, younger persons]
tzəwwəž > *tžəwwəz* 'he married'
zāwāž > *žāwāz* 'marriage'
mətzəwwəž > *mətžəwwəz* 'married'

Similar changes of /š/ and /ž/ also occur in the dialects of the eastern Maghreb (Cantineau's 'Parlers E' and Tunisian dialects), in Ḥassāniyya, and in the Baḥariyya oasis in Egypt (Drop and Woidich 2007:13).

2.1.2 Vowels

2.1.2.1 There are five long vowel phonemes: /ā/, /ī/, /ū/, /ē/, and /ō/ (examples: *žāb* 'he brought', *žēb* 'pocket', *žīb* 'bring [imper.]!'; *dāb* 'it melted', *dīb* 'wolf'; *dāg* 'he tasted', *dōg* 'taste', *dūg* 'taste [imper.]!'; *gāl* 'he said', *gūl* 'say [imper.]!'; *ṣēf* 'summer', *ṣūf* 'wool'; *lēn* 'until', *lōn* 'color'; *fīl* 'elephant', *fūl* 'broad beans'; *līm* 'oranges', *lōm* 'blame'), and two short vowel phonemes: /u/ and /ə/ (examples: *šəxx* 'he urinated', *šuxx* 'urinate [imper.]!'; *gəṣṣ* 'he cut', *guṣṣ* 'cut [imper.]!'; *kənna* 'daughter-in-law', *kunna* 'we were'; *ḥəžṛa* 'stone', *ḥužṛa* 'bedroom'; *təˤbān* 'tired', *tuˤbān* 'snake'; *bəˤd* [baʕd] 'after', *buˤd* 'distance'). The vowels /ā/, /ī/, /ū/, /ē/, and /ō/ in final open syllables are considered phonemically long, but written here without the macron for the sake of convenience.

2.1.2.2 The phonemes /ē/ and /ō/ differ from /ā/, /ī/, and /ū/ since they represent the reduction of the diphthongs /ay/ and /aw/, respectively, e.g. *zēt* (< **zayt*) 'oil', *dēl* 'tail', *žēb* (< **jayb*) 'pocket', and *ḥōš* 'house', *yōm* (<**yawm*) 'day', *ṣōt* (< **ṣawt*) 'voice', *rōšən* (<**rawšən*) 'window'. The reduction of the diphthongs /ay/ and /aw/ to /ē/ and /ō/, respectively, is a characteristic found in other North African Bedouin dialects as well.

2.1.2.3 The *ʾimāla*

The → *ʾimāla* involves a change in timbre in the pronunciation of [a] and [aː]. It may go as far as transforming [a] to [e], or even [i]. Tripoli Arabic is characterized by first-degree *ʾimāla*, i.e., [a] is pronounced as [e] (as compared with second-degree *ʾimāla*, where [aː] may be pronounced as [iː], or even occasionally as [iːə], as in Malta).

In Tripoli Arabic, *ʾimāla* affects final *-a* (etymological **ā*), independently of the consonantal context, in independent 1st person pronouns *āna* [aːne] 'I' and *ḥne* [ħne] 'we'.

The final **ā* of the 3rd person singular masculine of the perfect of almost all IIIy verbs is pronounced [e], e.g. *že* 'he came', *mše* 'he went', and in some nouns as well, e.g. *me* 'water', *sme* 'sky', and the adverb *hne*. This is characteristic of Sulaym dialects and, more particularly, of the dialects of southern Tunisia and the eastern Sahara.

2.1.3 Morphophonology/syllabic structure

2.1.3.1 Monosyllabic words

When a suffix with an initial vowel is added to a monosyllabic word of the structure $C_1C_2əC_3$, the syllabic sequence $C_1C_2əC_3$ turns to $C_1əC_2C_3$. This mutation of the scheme is called 'ressaut'. It occurs when the suffix of the 3rd person singular feminine *-ət* or the 3rd person plural *-u* is added to a verb in the 3rd person singular masculine of the perfect: *ktəb* + *-ət* > *kətbət* 'she wrote', *fhəm* + *-u* > *fəhmu* 'they understood'. It occurs as well when a pronoun suffix of the 2nd person singular or the 3rd person singular masculine is suffixed to these verbs: *ḍṛəb* + *-ək* > *ḍəṛbək* 'he hit you', *žməl* + *-əh* > *žəmləh* 'his camel', and likewise when the feminine or singulative suffix *-a(t)* or the dual suffix *-ēn* is added to a masculine noun with the structure $C_1C_2əC_3$: *bġəl* + *-a(t)* > *bəġla(t)* 'she-mule', *nməl* + *-a(t)* > *nəmla(t)* 'one ant', *šhəṛ* + *-ēn* > *šəhṛēn* 'two months'.

2.1.3.2 Polysyllabic words

When the feminine suffix *-i* is added to the 2nd person singular and when the suffix *-u* is added to the plural of the imperfect, or when either ending is added to an imperative, the short vowel of the second syllable is elided, e.g. *túšṛub* + *-i* > *túšṛbi* 'you drink', *núṛgud* + *-u* > *núṛgdu* 'we sleep', *ə́sməˤ* + *-i* > *ə́smˤi* 'listen!', *úšṛub* + *-u* > *úšṛbu* 'drink!'.

This also affects words with the structure $CəC_1C_2C_3a(t)$. The stress on the first syllable maintains the first short vowel in an open syllable: *mə́drsa* 'school', *mə́ktba* 'library', *mə́knsa* 'broom'; *fəstka* 'one pistachio', *təlfla* 'one pimento'.

Elision also takes place when a vocalic pronoun is suffixed to a word with the structure $CəC_1C_2əC_3$, e.g. *məṭˤəm* + *-i* > *mə́ṭˤmi* 'my restaurant', *mənžəl* + *-əh* > *mə́nžləh* 'his sickle', and when a vocalic suffix is added to a feminine noun with the structure $C_1əC_2C_3a(t)$, e.g. *xədma(t)* + *-i* > *xə́dmti* 'my work', *səlfa(t)* + *-ək* > *sə́lftək* 'your sister-in-law'.

In such cases, an ultrashort epenthetic vowel may appear between the first and second consonants or between the second and third ones: /tušṛbi/ ['tʉʃʉrbɨ] 'you will drink', /əsmʿi/ ['asmăʕi] 'listen [sg. fem.]!', /məknsa/ ['mɐkĕnsa] 'broom', /xədmt-i/ ['xadə̆mti] 'my work'.

When a suffix with an initial vowel is added to the 3rd person singular feminine perfect, the final vowel of the verb ending is modified from *-ət* to *-āt*: *fəhmət* + *-ək* > *fəhmāt-ək* 'she understood you', *ḍəṛbət* + *-əh* > *ḍəṛbāt-əh* 'she hit him'. In Tripoli Arabic, the change of the final vowel of the verb ending may be attributed to the need to keep a vowel in this position; without this modification, for syllabic structure reasons, the vowel would disappear.

2.2 *Morphology*

2.2.1 Pronouns

2.2.1.1 Independent personal pronouns

In the 1st person, the pronouns are *āne* (sg.) and *ḥne* (pl.; for the final *-e*, see above, Sec. 2.1.2.3). In the 2nd person singular, there is a gender distinction in the independent personal pronoun: *ənta* (masc.) vs. *ənti* (fem.). This is not the case in certain other innovative pre-Hilalian urban dialects. The 3rd person singular pronouns are *hūwa* (masc.) and *hīya* (fem.). For the 3rd person plural, the pronoun is *humma*, as is the case in Bedouin dialects elsewhere in the Maghreb. In the 2nd and 3rd person plural, there is no gender distinction, unlike the dialects of the Fezzān (southwest Libya), Cyrenaica (East Libya), and southern Tunisia; these Bedouin dialects distinguish the feminine from the masculine in both pronouns and conjugation of the verb.

2.2.1.2 Possessive/suffixed object pronouns

In the 1st person singular, the pronominal suffix for the object is *-ni* (e.g. *šbəḥni* 'he saw me', *bāsni* 'he kissed me', *ʿžəbni* 'he pleased me, I like him'). The form *-i* is suffixed to nouns and to prepositions ending in a consonant (e.g. *ḥōši* 'my house', *mfātīḥi* 'my keys', *gəddāmi* 'in front of me'); the form *-y* is suffixed to nouns and prepositions ending in a consonant or vowel of a timbre other than [i] and [e] (e.g. *xūy* 'my brother', *būy* 'my father', *ʿšēy* 'my dinner', *mʿāy* 'with me'). When a 1st person singular pronoun is suffixed to a noun or a preposition ending in /ī/ or /ē/, the pronominal suffix is *-ya* (e.g. *kursīya* 'my chair', *īdēya* 'my hands', *ʿlēya* 'on me', *fīya* 'in me', *lēya* 'for me').

The suffix pronoun of the 3rd person singular masculine added to nouns, prepositions, and verbs ending in a vowel is *-h*. Examples: *būh* 'his father', *xūh* 'his brother', *ʿšēh* 'his dinner', *ʿlēh* 'on him', *fīh* 'in him'. After a consonant it is *-a*, as in *ḍərba u mše* 'he hit him and went away', with a pausal form *-əh*, e.g. *əlxāl fi xēṛ lēn yukbuṛ la wuld uxteh* 'the uncle is fine as long as his nephew has not grown up'. With the negational suffix *-š*, the vowel *-a* is long and stressed, e.g. *mā-ḍṛabtāš* 'he did not hit him'. As is the case for independent pronouns and the verb, there is no gender distinction in the 2nd and 3rd person plural, *-kum* and *-hum* respectively.

When the pronominal suffixes are added to the prepositions *fī-* 'in', *lē-* 'for', *ʿlē-* 'on', and *mʿā* 'with', the forms in Table 2 result.

Table 2. Prepositions with pronominal suffixes

3rd sg. masc.	*fī-h*	*lē-h*	*ʿlē-h*	*mʿā-h*
3rd sg. fem.	*fī-ha*	*lē-ha*	*ʿlē-ha*	*mʿā-ha*
3rd pl.	*fī-hum*	*lē-hum*	*ʿlē-hum*	*mʿā-hum*
2nd sg. masc.	*fī-k*	*lē-k*	*ʿlē-k*	*mʿā-k*
2nd sg. fem.	*fī-kum*	*lē-kum*	*ʿlē-kum*	*mʿā-kum*
1st sg.	*fī-ya*	*lē-ya*	*ʿlē-ya*	*mʿā-y*
1st pl.	*fī-na*	*lē-na*	*ʿlē-na*	*mʿā-na*

2.2.1.3 Demonstratives

2.2.1.3.1 Nominal determiners

The proximal deictic adjectives, linked to the speaker's space, are given in Table 3.

Table 3. Deictic adjectives: Proximal

	sg. masc.	sg. fem.	pl.
invariable form		*hā l-*	
preposed to the noun	*hādā l-*	*hādī l-*	*hādū l-* *hādūmā l-* *hādūmāyā l-*
postposed to the noun	*əl-…hāda*	*əl-…hādi*	*əl-…hādu* *əl-…hādūma* *əl-… hādūmāya*
pharyngealized form	*əl-…āhẉá*	*əl- …āhỵá*	*əl-…ā-hᵘṃá*

Examples include (invariable) *hā ṛṛāžəl ṭbīb* 'this man is a doctor', *hā lbənt uxt xāləd* 'this girl is Khaled's daughter', *hā lūlād mən* *bənġāzi* ‘these boys are from Benghazi’, *hā ssəyyāṛāt kwəyysāt* ‘those cars are good’; (sg. masc.) *hādā ləktāb mlīḥ* ‘this book is good’, *əṛṛāžəl hāda ūstādna* ‘this man is our teacher’; (sg. fem.) *hādī lbənt gnēyna* ‘this girl is cute’, *əttuffāḥa hādi bnīna* ‘this apple is tasty’; (pl.) *hādū lūlād mən ṭṛābləs* ‘these children are from Tripoli’, *hādūmā ləbnāwīt mən lībya* ‘these girls are from Libya’, *hādūmāyā ssəyyāṛāt kwəyysāt* ‘these cars are good’, *lūlād hādu mən ṭṛābləs* ‘these children are from Tripoli’, *ləbnāwīt hādūma mən lībya* ‘these girls are from Libya’, *əssəyyāṛāt hādūmāya kwəyysāt* ‘these cars are good’. The pharyngealized forms are postposed to the determined noun, e.g. *ənnəˁža āhyạ́ žābət xəṛfān* ‘this ewe has given birth to a lamb’, *lūlād u ləbnāwīt āh^{u}ṃá kīf ṭəlˁu m əlmədṛsa* ‘these boys and these girls have just left school’.

The distal deictic adjectives in Table 4 are linked to the space of the speaker and the interlocutor.

Table 4. Deictic adjectives: Distal

	sg. masc.	sg. fem.	pl.
forms preposed	*hādākā l-*	*hādīkā l-*	*hādūkā l-*
to the noun	*hādākāyā l-*	*hādīkāyā l-*	*hādūkāyā l-*
forms postposed	*əl-…*	*əl-…*	*əl-…*
	hādāka	*hādīka*	*hādūka*
to the noun	*əl-…*	*əl-…*	*əl-…*
	hādākāya	*hādīkāya*	*hādūkāya*

Examples include (sg. masc.) *hādākā ləktāb mā-ˁžəbnīš* ‘I don’t like that book’, *hādākāyā lwəld mā-yuskutš* ‘that child is not quiet’, *ləḥmāṛ hādāka mā-žāš* ‘that idiot has not come’, *əṛṛāžəl hādākāya kəddāb* ‘that man is a liar’; (sg. fem.) *hādīkā ləmdīna bˁīda bˁīda mən ṭṛābləs* ‘that city is very far from Tripoli’, *hādīkāyā lbənt šēna* ‘that girl is ugly’, *əžžumla hādīka mā-tžīš* ‘that sentence is not correct’, *əlmākla hādīkāya məš mlīḥa* ‘that food is not good’; (pl.) *hādūkā ləbnāwīt mā-səllmūš ˁlēha* ‘those girls didn’t say hello to her’, *hādūkāyā lūlād mā-yugdṛūš yəlˁbu lkūṛa hne* ‘those children cannot play football here’, *əžžūməl hādūka mā-tžūš* ‘those sentences are not correct’, *əllḥēmāt hādīkāya bāhyāt* ‘those pieces of meat are good’.

2.2.1.3.2 Demonstrative pronouns

The proximal pronouns for the space close to the speaker are given in Table 5.

Table 5. Demonstrative pronouns: Proximal

sg. masc.	sg. fem.	pl.
hāda	*hādi*	*hādu*
hādāya	*hādīya*	*hādūma*
		hādūmāya
āhwạ́	*āhyạ́*	*āh^{u}ṃá*

Examples include *hāda wəldi* ‘this one is my son’, *hādāya ṣāḥbi* ‘this one is my friend’, *hādi bənti* ‘this one is my daughter’, *hādīya ṣāḥəbti* ‘this one is my friend’; *hādu ūlādi* ‘these ones are my children’, *hādūmāya ṣḥābna* ‘these ones are our friends’, *āh^{u}ṃa lūlād u ləbnāwīt əlli dāru lḥādət* ‘these are the boys and the girls who had the accident’.

The distal pronouns are given in Table 6.

Table 6. Demonstrative pronouns: Distal

sg. masc.	sg. fem.	pl.
hādāka	*hādīka*	*hādūka*
hādākāy	*hādīkāy*	*hādūkāy*
hādākāya	*hādīkāya*	*hādūkāya*
		hādūkumma

Examples include *hādāka kəddāb kbīr* ‘this one is a big liar’, *hādākāy kān yəxdəm mˁāy f nəfs əššārīka* ‘this one was working with me in the same company’, *hādākāya mā-ˁəndi ma ngūl fīh* ‘this one, I have nothing to say about him’, *hādīka šəgga kbīra* ‘this is a big apartment’, *hādīkāy šēna* ‘this one is ugly’, *hādīkāya mā-ˁəndi ma ngūllək fīha* ‘I have nothing to say about her’; *hādūka ṭṛābəlsīya* ‘these ones are Tripolitanian’, *hādūkāy lḥēmāt bāhyāt* ‘these are good pieces of meat’, *hādūkāya mən səbha* ‘these ones are from Sebha’, *hādūkumma səyyāṛāt ṣəḥḥ* ‘these ones are very good cars’.

2.2.1.4 The relative pronoun

The most common relative pronoun is *əlli* ‘which, who’. For the animate are used *mən ~ məni ~ mənu ~ mən əlli; škūn ~ škūn əlli; škūn* + independent personal pronoun 3rd person + *əlli*. For the inanimate are used *ma* ‘which, that’; *šən, šīni, šīnu*. The relative pronoun *əlli* is invariant for both gender and number and is only used with a determined antecedent, e.g. *əllḥəm əlli žəbt-a ysədd-ək?* ‘is the meat that I bought enough for you?’. Other examples: *mā-xəlla li ma nākəl* ‘he left nothing for me to eat’, *mā-yḥəbbūš mən yəxnəb* ‘they do not like the ones who steal’, *hāda šən ṣāṛ* ‘this is what happened’.

2.2.1.5 Interrogative pronouns and adjectives

The pronouns used in Tripoli Arabic are the following: *šīn*, *šən*, *šīnu*, *šīni* ‘what?’, *škūn* and *mən* ‘who?’. For the interrogative adjective, *āma* ‘which?’ is used. The main interrogative pronouns and compound forms are *šīn*, *šīnu*, *šīni* ‘what?’, *f-šīn* ‘in what?’, *ˁlē-šīn* ‘on what, why?’, *zēy-šīn* ‘like what?’, *b-āš* ‘with what?’, *f-āš* ‘in what?’, *ˁl-āš* ‘on what, why?’, *l-āš* ‘why?’, *gədd-āš* ‘how much, which size?’, *wēn* ‘where?’, *mn-īn* ‘from where?’, *mən-wēn* ‘from where?’, *l-ēn* ‘until where?’, *škūn* ‘who?’, *lī-škūn* ‘for whom?’, *mtāˁ-škūn* ‘of whom?’, *ḥdā-škūn* ‘next to whom?’, *gədd-škūn* ‘whose size?’, *mən*, *mənu*, *məni* ‘who?’, *mˁā-mən* ‘with whom?’, *mtāˁ-mən* ‘of whom?’, *ˁlē-mən* ‘on whom?’, *ḥdā-mən* ‘next to whom?’, *gədd-mən* ‘whose size?’, *zēy-mən* ‘like who?’, *kīf* ‘how?’, *āmta* ~ *əmta* ‘when?’.

2.2.2 Adverbs

Adverbs of place include *hne*, *hnāya* ‘here’; *ġādi*, *ġādīka* ‘there’; *ġādīkāy*, *ġādīkāya* ‘over there’; *grīb* ‘near’; *bˁīd* ‘far’; *ṭūl* ‘in front, directly’; *bəṛṛa* ‘outside’; *gəddām* ‘in front’; *l-tāli* ‘to the back’; *lūṭa* ‘to the ground, downstairs’; *fəlwəsṭ* ‘in the middle, in the center’; *mən hne* ‘from here’; *mən ġādi* ‘from there’; *mən ġādīka* ‘from over there’; *mən tāli* ‘in the back’; *m-əlūṭa* ‘down, on the ground’; *ˁəl-līmīn* ‘on the right-hand side’; *ˁəl-līṣāṛ* ‘on the left-hand side’; *ˁəl-lḥāšya* ‘on the edge’; *ˁəl-lūṭa* ‘on the ground, down’.

Adverbs of time include *əlyōm* ‘today’, *əllēla* ‘tonight’, *hā-llēla* ~ *llēla hādi* ‘this night’, *əlbāṛəḥ* ‘yesterday evening’, *ġudwa* ‘tomorrow’, *mən ġədwītha* ‘the following day’, *bəˁd-ġudwa* ‘after tomorrow’, *bəˁd ġudwtēn* ‘after the day after tomorrow’, *āməs* ‘yesterday’, *wūl āməs* ‘before yesterday’, *wūl āməstēn* ‘before the day before yesterday’, *ssne* ‘this year’, *ˁāmnəwwəl* ‘last year’, *yōmləwwəl* ‘the first day’, *zmān* ‘formerly, in times past’, *kīf* ‘just’, *ˁāwəd* ‘once again’, *məṛṛāt* ‘sometimes, from time to time’, *bəˁdēn* ~ *bəˁtāli* ‘afterward, soon’, *ābādān* ‘never’, *təwwa* ‘now’, *gəbl* ‘before’, *bəkri* ‘soon, early in time’, *dīma* ‘always’, *ġāliban* ‘often’, *glīləš* ‘rarely, seldom’, *lməṛṛa hādi* ‘this time’, *fəlgəddām* ‘hereafter’ , *bəlmˁāwəd* ‘once again’, *fī-sāˁ fī-sāˁ* ‘quickly, right away’, *məṛṛa məṛṛa* ‘sometimes, from time to time’, *təwwa kīf* ‘hardly, just’, *ˁāwəd tāni* ‘once again’, *bəkri šwēya* ‘soon, a little time ago’.

Adverbs of quantity: *hākki*, *hēkke* ‘so, this way’, *hālba* ‘a lot, too much’, *šwēya* ‘not much, (too) little’, *yāsər* ‘enough’, *ḥətta* ‘even, also, either’, *glīl* ‘less, little’, *məgtūl* ‘precisely, exactly’, *swa-swa* ‘same, similarly’, *bukkull* ‘entirely, completely, not at all’.

Adverbs of manner: *hākki*, *hēkke* ‘so, this way’; *buntu* ‘deliberately’; *ḥəgga* ‘just’; *bəss* ‘only, barely’; *xālāṣ* ‘enough’; *xṣēṣən* ‘specially’; *mātālān* ‘for example’; *b šwēya* ‘gently, softly, slowly’; *b žəddīyāt* ‘seriously’; *b əlblāš* ‘for free’; *b əssēf* ‘forced’; *b əddərga* ‘secretly’; *b əlˁāni* ‘deliberately’; *zēy təwwa* ‘for example’; *bāląk* ~ *bāląk ši* ‘maybe’.

2.2.3 Nouns

2.2.3.1 Diminutives

The diminutive of triliteral nouns is formed on the patterns C_1C_2*ē*C_3, C_1C_2*ē*C_3*a*(*t*), or C_1C_2*ēyə*C_3. Note that the forms C_1C_2*ē*C_3 and C_1C_2*ē*C_3*a*(*t*) are characteristic of Bedouin dialects, as opposed to the model C_1C_2*ēyə*C_3, which is common in sedentary dialects. Examples: *kəlb* ‘dog’ and *klēb* ‘puppy’; *ṭəṛf* ‘piece’ and *ṭrēf* ‘small piece’; *ḥəlw* ‘sweet’ and *ḥlēw* ‘a little bit sweet’; *ġəṛṣ* ‘plantation’ and *ġṛēṣa* ‘small plantation’; *šəms* ‘sun’ and *šmēsa* ‘small sun’; *kərša* ‘belly, stomach’ and *krēša* ‘small belly, small stomach’; *ḥōš* ‘house’ and *ḥwēš* ‘small house’; *bənt* ‘girl’ and *bnēya* ‘small girl’; *kəbš* ‘ram’ and *kbēyəš* ‘small ram’; *žəḥš* ‘baby donkey’ and *žḥēyəš* ‘small baby donkey’. In Tripoli Arabic, the pattern C_1C_2*ēyə*C_3 is above all used to form the diminutives of nouns containing a long vowel between C_2 and C_3, e.g. *ṣġīṛ* ‘small’ and *ṣġēyəṛ* ‘very small’; *xrūf* ‘lamb’ and *xrēyəf* ‘small lamb’. This pattern is also used to form the diminutive of the diminutive of a small number of other nouns taking the model C_1C_2*ē*C_3, e.g. *klēb* ‘puppy’ and *klēyəb* ‘tiny dog’; *wlēd* ‘small boy’ and *wlēyəd* ‘wee boy’. In fact, the pattern C_1C_2*ēyə*C_3 is the diminutive pattern for nouns with a long vowel between C_2 and C_3 in Bedouin dialects.

Quadriliteral forms on the pattern C_1əC_2C_3vC_4 form their diminutives on the pattern C_1C_2*ē*C_3*ī*C_4 and C_1C_2*ē*C_3əC_4, e.g. *məftāḥ* ‘key’ and *mfātīḥ* ‘small key’; *səndūg* ‘box’ and *snēdīg* ‘small box’; *sərwāl* ‘trousers’ and *srēwīl* ‘small trousers’; *ʿəṣfūṛ* ‘bird’ and *ʿṣēfīṛ* ‘small bird’; this brings Tripoli Arabic closer to Bedouin dialects.

The following nouns have irregular diminutives: *umm* ‘mother’, diminutive *ṃṃēma*; *bu* ‘father’, diminutive *ḅḅēy; xu* ‘brother’, diminutive *uxēy; ma* ‘water’, diminutive *ṃṃēya; fumm* ‘mouth’, diminutive *f̣f̣ēm*.

2.2.3.2 Plurals

Quadriliteral nouns of the pattern $C_1vC_2C_3\bar{v}C_4$ form plurals of the pattern $C_1C_2\bar{a}C_3\bar{\imath}C_4$, e.g. *meftāḥ/mfātīḥ* ‘key’, *məṣmār/mṣāmīr* ‘nail’, *səndūg/snādīg* ‘box’, *sərwāl/srāwīl* ‘trousers’, *ʿəṣfūṛ/ʿṣāfīṛ* ‘bird’.

2.2.3.3 Colors and physical defects

The masculine singular of adjectives of color has the pattern *a*$C_1C_2vC_3$ (e.g. *əḥməṛ* ‘red’, *əzṛəg* ‘blue’, *əxḍəṛ* ‘greens’, *əkḥəl* ‘black’, *əšgər* ‘blond’). The feminine singular has the pattern $C_1vC_2C_3$*a* (e.g. *ḥəmra*, *zəṛga*, *xəḍṛa*, *kəḥla*, *səgra*). Adjectives indicating defects follow the same pattern, e.g. *əṭṛəš* (fem. *ṭəṛša*) ‘deaf’; *əkwəs* (fem. *kəwsa*) ‘having a squint [divergent strabismus]’; *əḥwəl* (fem. *ḥəwla*) ‘having a squint [converging strabismus]’. The plural has the pattern $C_1\bar{u}C_2əC_3$ (*ḥūməṛ*, *zūṛəg*, *xūḍəṛ*, *šūgəṛ*, etc.), and has the pattern $C_1\bar{u}C_2$ when C_2 = *w* (e.g. *siwəd* ‘black’, plural *sūd, əʿwəṛ* ‘one-eyed’, plural *ʿūṛ*). The diminutive is based on the pattern $C_1C_2\bar{e}C_3$*a*(*t*). Only three examples are used in Tripoli Arabic: *kḥēla* ‘a little bit black’, *zṛēga* ‘a little bit darker-skinned’, *ḥmēṛa* ‘a little bit red’.

2.2.4 Numerals

2.2.4.1 The dual

The dual is formed by adding the suffix *-ēn* to the singular noun. In Tripoli, the dual form is widely used for all sorts of nouns, e.g. *ṭrīg* ‘one road’, dual *ṭrīgēn*; *gəhwa* ‘one coffee’, dual *gəhwtēn*; *ktāb* ‘book’, dual *ktābēn*; it is even used with loanwords, e.g. *sīmāfrōwēn* ‘two traffic lights’. Nouns of units of measure and time also take the dual, as do the twin parts of the body (→ pseudodual), e.g. *wāldēn* ‘parents’; units of time: *dgīgtēn* ‘two minutes’, *yōmēn* ‘two days’, *šəhrēn* ‘two months’, *sāʿtēn* ‘two hours’, *ʿāmēn* ‘two years’, *lēltēn* ‘two nights’; units of measurement: *ždūltēn* (measure of area), *kēltēn* (measure of capacity, for grain or liquid), *drāʿēn* ‘two cubits’, *mītēn* ‘two hundred’, *ālfēn* ‘two thousand’; twin parts of the body: *wudnēn* ‘two ears’, *īdēn* ‘two hands’, *drāʿēn* ‘two arms’, *kurʿēn* ‘two legs’, *rəžlēn* ‘two feet’, *kətfēn* ‘two shoulders’. In cases where the dual is not used, the numeral *zōz* ‘two’ takes its place, followed by the noun in the plural, e.g. *zōz ḥūtāt* ‘two fish’, *zōz bībān* ‘two doors’, *zōz knāsa* ‘two brooms’.

2.2.4.2 Numerals from 3 to 10 and 11 to 19

In the numerals between 3 and 10, a short and a long form are used indiscriminately, e.g. *tlāta meṛṛāt ~ tәlt-məṛṛāt* ‘three times’, *sətta žūməl ~ sətt-žūməl* ‘six camels’, *təsʿa ktābāt ~ təsʿ-ktābāt* ‘nine books’.

The numerals between 11 and 19 form a construct state with the counted noun, which remains in the singular. The *r* of *ʿašr* reappears in the form of a liquid /l/, e.g. *ḥdāšəl wuld* ‘eleven boys’, *ṭnāšəl ʿām* ‘twelve years’, *tluṭṭāšəl bənt* ‘thirteen girls’. This /l/ is assimilated to the first consonant of the noun, just like the article *əl-*, e.g. *xəmsṭāšəž žməl* ‘fifteen camels’, *suṭṭāšən nəžma* ‘seventeen stars’.

2.2.5 Verbs

2.2.5.1 Like all Libyan dialects, Tripoli Arabic belongs to the Maghrebi family of dialects. Typically, in these dialects the 1st person prefix of the imperfect is *n-*; the plural is distinguished from the singular by the ending *-u*, e.g. *nəktəb* ‘I write’, *nəkt*^{ə}*bu* ‘we write’; *nəmši* ‘I go’, *nəmšu* ‘we go’; *ngūl* ‘I say’, *ngūlu* ‘we say’.

The conjugation of the strong verb is given in Table 7.

Table 7. Conjugation of the strong verb, Form I

	perfect *ktəb* ‘he wrote’		imperfect *yəktəb* ‘he writes’	
	sg.	pl.	sg.	pl.
3rd masc.	*ktəb*	*kətbu*	*yəktəb*	*yəktbu*
3rd fem.	*kətbət*		*təktəb*	
2nd masc.	*ktəbt*	*ktəbtu*	*təktəb*	*təktbu*
2nd fem.	*ktəbti*		*təktbi*	
1st	*ktəbt*	*ktəbna*	*nəktəb*	*nəktbu*

2.2.5.2 In verbs III*y*, there is no reconstruction of the paradigm with forms like *təlgāy* ‘you find', *yəbdāw* 'you begin' in the imperfect, and *šrāw* 'you bought', *mšāt* 'she went away' in the perfect (containing the long vowel *ā*), found in the more innovative pre-Hilalian dialects. In the imperfect, the following prefixes are found: 2nd person singular feminine *təlgi* 'you find'; in the plural, *nəšru* 'we buy', *təlgu* 'you find', and *yəbdu* 'they begin'. In the perfect, the following are found: for the 3rd person plural, *šru* 'they bought'; in the 3rd person singular feminine, Tripoli Arabic has the form C_1C_2*ət*, e.g. *mšət* 'she went away' (Table 8).

Table 8. Conjugation of the III*y* verb, Form I

	perfect		imperfect			
šra 'he bought'	sg.	pl.	sg.	pl.	sg.	pl.
3rd masc.	*šra*	*šru*	*yéšri*	*yəšru*	*yəlga*	*yəlgu*
3rd fem.	*šrət*		*təšri*		*təlga*	
2nd masc.	*šrēt*	*šrētu*	*təšri*	*təšru*	*təlga*	*təlgu*
2nd fem.	*šrēti*		*təšri*		*təlgi*	
1st	*šrēt*	*šrēna*	*nəšri*	*nəšru*	*nəlga*	*yəlgu*

2.2.5.3 No internal passives are found in Tripoli Arabic, unlike the dialects of the Fezzān. To express the simple passive in Tripoli, Form VII is used, which is obtained by prefixing *n-* to Form I of the verb, e.g. *nktəb* 'it was written', *nˁəḍḍ* 'he was bitten', *nzād* 'he was born', *ngla* 'it was fried'. Form VIII constitutes the reflexive of Form I and is often used to express an action undertaken by subjects for themselves. It is characterized by the consonant *-t-* inserted after the first radical, e.g. *rtxe* 'it went soft', *ntgəb* 'it was perforated', *ltəmmu* 'they got together', *rtāḥ* 'he relaxed'.

Combinations of Forms occur: II + VIII *btəssəm* 'he smiled', *ttəṣṣəl* 'he got in touch'; X + II *strəyyəḥ* 'he rested', *stənna* 'he waited'.

2.2.5.4 Former *hamza* verbs

The Old Arabic verbs Iʾ *ʾaxaḏa/yaʾxuḏu* 'to take' and *ʾakala/yaʾkulu* 'to eat' have equivalents in Tripoli Arabic: *xde/yāxud*, *kle/yākul*. In the perfect, they are conjugated like verbs III*y*. The imperative is formed like verbs II*w/y*: *xūd* 'take!', *kūl* 'eat!'. The active participles have the following form: *wākəl* and *wāxəd*; for the passive participles, the forms *mūkūl* and *mūxūd* are used.

2.3 *Syntax*

2.3.1 Possessive constructions

The synthetic construction is commonly used (like the dual) in Tripoli Arabic e.g. *mōlā-dukkān-ən-nəqqālāt mūš muḥḥi* 'Muhhi is not the owner of the mobile phones shop'. The analytic construction, an innovation specific to sedentary dialects, is also used. The particle used in Tripoli is *mtāˁ*, e.g. *əlmākla mtāˁ əlwuld wāṭya* 'the child's food is ready', *xde dəftar mtāˁ wuld u ktəb fīh xuṛṛāfa* 'he took a boy's notebook and wrote a story', *əlḥōš mtāˁ əlgāḍi mūš băˁīd ˁlēna* 'the judge's house is not far from us'. The use of the synthetic construction is strictly limited, for instance to naturally associated nouns: *bāb əlḥōš* 'the door of the house', *mūš ktābi lli təḥt əlkursi*, *ktāb xāləd* 'the book that is under the chair is not my book, it is Khaled's book'; to express family relations: *ˁəlt xūy* or *maṛt xūy* 'my brother's wife', *umm ṣāləḥ* 'Salah's mother', *bū mūnīr* 'Mūnīr's father'; and to refer to body parts: *ṛās wuldi* 'my son's head', *wudnēk* 'your ears'.

2.3.2 Future

In Tripoli Arabic, the preverbal marker *b(ə)-* is used to express the future of intention (only in affirmative or interrogative statements). Examples: *b-nəmšu l-bənġāzi* 'we intend to go to Benghazi', *āš bᵊtdīr ġudwa?* 'what will you do tomorrow?'. This construction with the particle *b-* can also be used to express the idea of wanting to undertake an action or as a mark of imminence, possibility, will, or finality. The preverbal marker *ḥā-* is used to express close/near future. Examples: *ḥā-nəmši l-ṭṛābləs* 'I am going to go to Tripoli'; *ḥāykəmməl xədmtəh* 'he is going to finish his work'.

2.3.3 Indefinite article

In Tripoli Arabic, no article is used to mark indeterminate noun status, and the noun appears in this case without any marker. Examples: *bīr* 'a well', *žṛāna* 'a frog', *krāsi* 'chairs'. Unlike the urban dialects of Algeria and Morocco, Tripoli Arabic does not use the indefinite construction *wāḥd əl* formed from the numeral *wāḥəd* 'one', whose use is very limited in Bedouin dialects.

3. Lexicon

The lexicon of Tripoli Arabic shows Bedouin traits. Verbs such as *dār* 'to do', *xəšš* 'to enter', *ṭləʿ* ‘to go out’, *yəbbi* ‘to want’, *šbəḥ* ‘to see’, *nšəd* ‘to ask’, *rāža* ‘to wait’, *gəʿməz* ‘to sit’, *gʿəd* ‘to stay’, *dfəʿ* ‘to pay’, *wəlla* ‘to come back’, *rkəb* ‘to go up’, *tṣubb əlmṭəṛ* ‘it’s raining’ are common. The following nouns are also used: *mṭər* ‘rain’, *dəḥya* ‘an egg’, *ḥōš* ‘a house’, *īd* ‘a hand’, *ṣubḥ* ‘a morning’, *rəžžāla* ‘men’; other words, such as *humma* ‘they’, *āməs* ‘yesterday’, and augmented forms such as *ġudwīka* ‘tomorrow’, *ġādīkāy* ‘over there’, *hādūkāy* ‘those’, which are found in other Bedouin dialects, are also employed in Tripoli Arabic.

Hālba is the word used in Tripoli Arabic for ‘a lot’; it may be derived from Hebrew *harbēh* ‘a lot’.

Tripoli Arabic contains numerous lexical borrowings from Turkish, e.g. *kāšīk* ‘spoon’, *šīša* ‘bottle’, and from Italian, e.g. *māršābēdi* (< Italian *marciapiede*) ‘pavement’, *kāčāfīti* ~ *kāčāvīti* (< Italian *cacciavite*) ‘screwdriver’, *sīmāforo* (< Italian *semaforo*) ‘traffic lights’ (→ Italian loanwords), and from English, e.g. *bōy* (< English *boy*) ‘gay’, *būfta* (< English *poofter*) ‘gay’. Note, however, that in terms of both borrowing and code-switching, Italian, the language of the 20th century’s main colonial power in Libya, has had nothing like the impact of French on spoken Arabic in the other Maghreb states.

More recently, many words dealing with new technologies have been borrowed from English, e.g. *kōmbyūtər* ‘computer’, *māws* ‘mouse’, *mōbāyl* ‘mobile phone’, *sərč* ‘internet search’, *burṭābəl* ‘laptop, portable computer’.

BIBLIOGRAPHICAL REFERENCES

Abdu, Hussein Ramadan. 1987. *Italian loanwords in colloquial Libyan Arabic as spoken in Tripoli Region*. Ph.D. diss., University of Arizona.

Caubet, Dominique. 2004. “Les parlers arabes nomades et sédentaires du Fezzān, d’après William et Philippe Marçais”. *Approaches to Arabic dialects: A collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday*, ed. Martine Haak, Rudolf de Jong, and Kees Versteegh, 67–96. Leiden: E.J. Brill.

Cesàro, Antonio. 1939. *L’arabo parlato a Tripoli*. Milan: Mondadori.

Curotti, Torquato. 1933. *Il dialetto libico: Grammatica, conversazione, dizionario*. Tripoli: Maggi.

Drop, Hanke and Manfred Woidich. 2007. *ilBaḥariyya: Grammatik und Texte*. Wiesbaden: O. Harrassowitz.

Elfitoury, Abubaker Abdalla. 1976. *A descriptive grammar of Libyan Arabic*. Ph.D. diss., Georgetown University.

Féraud, Laurent-Charles. 2005. *Annales tripolitaines*. Saint-Denis: Editions Bouchene.

Goldberg, Harvey E. 1983. “Language and culture of the Jews of Tripolitania: A preliminary view”. *Mediterranean Language Review* 1.85–102.

Griffini, Eugenio. 1913. *L’arabo parlato della Libia*. Milan: Ulrico Hoepli.

Iannotta, Elpidio. 1933. *L’arabo parlato in Cirenaica*. Benghazi.

Mantran, Robert. 1975. “La Libye des origines à 1912”. *La Libye nouvelle: Rupture et continuité*, 15–32. Paris: CNRS.

Marçais, Philippe. 1977. *Esquisse grammaticale de l’arabe maghrébin*. Paris: Adrien Maisonneuve.

——. 2001. *Parlers arabes du Fezzân: Textes, traductions et éléments de morphologie rassemblés et présentés par Dominique Caubet, Aubert Martin et Laurence Denooz*. Geneva: Librairie Droz.

Owens, Jonathan. 1983 [1987]. “Libyan Arabic dialects”. *Orbis* 23.97–117.

——. 1984. *A short reference grammar of Eastern Libyan Arabic*. Wiesbaden: O. Harrassowitz.

Panetta, Ester. 1943. *L’arabo parlato a Bengasi*. Rome: La Libreria dello Stato.

Pereira, Christophe. 2003. “Le parler arabe de Tripoli (Libye): Etat des lieux d’après les travaux de Hans Stumme, Antonio Cesàro, Eugenio Griffini…”. *Estudios de Dialectología Norteafricana y Andalusí* 7.7–57.

——. 2004. “À propos du parler arabe de Tripoli”. *AIDA 5th Conference Proceedings*, ed. Ignacio Ferrando and Juan Sánchez Sandoval, 431–443. Cadiz: Universidad de Cádiz.

——. 2006. “Eléments de description du parler arabe de Tripoli”. *L’arabe dialectal: Enquêtes, descriptions, interprétations. Actes d’AIDA 6*, ed. Salah Mejri, 371–385. Tunis: Centre d’Etudes et de Recherches Economiques et Sociales.

——. 2007. “Urbanization and dialect change: The Arabic dialect of Tripoli (Libya)”. *Arabic in the city: Studies in dialect contact and language variation,* ed. Catherine Miller, Enam Alwer, Dominique Caubet, and Janet Watson, 2008. London: Routledge.

——. 2008. *Le parler arabe de Tripoli (Libye): Phonologie, morphosyntaxe et catégories grammaticales*. Ph.D. diss., Institut National des Langues et Civilisations Orientales, Paris.

Rossi, Ettore. 1968. *Storia di Tripoli e della Tripolitania dalla conquista araba al 1911*. Ed. Maria Nallino. Rome: Istituto per l’Oriente.

Stumme, Hans. 1894. *Tripolitanisch-tunesische Beduinenlieder*. Leipzig: J.C. Hinrich’sche Buchhandlung.

——. 1898. *Märchen und Gedichte aus der Stadt Tripolis in Nordafrika*. Leipzig: Hinrich’sche Buchhandlung.

Trombetti, Alfredo. 1912. *Manuale dell’arabo parlato a Tripoli*. Bologna: L. Beltrami.

Türkmen, Erkan. 1988. “Turkish words in the Libyan dialect of Arabic”. *Erdem* 4.227–243.

Yoda, Sumikazu. 2005. *The Arabic dialect of the Jews of Tripoli: Grammar, text and glossary*. Wiesbaden: O. Harrassowitz.

CHRISTOPHE PEREIRA (LACNAD-CRÉAM-INALCO)

Triptosis → Diptosis

Triradicalism → Biradicalism; Lexicon: Matrix and Etymon Model

Truncation

1. DEFINITION

The term 'truncation' can be understood generally as referring to a process of word shortening which is phonologically predictable. Morphological processes of clipping and acronymy also belong to this category. The former process derives a shorter form from a simple lexeme, while the latter derives forms consisting of the first letters of a compound lexeme (cf. Kreidler 2000:956ff.). In the context of Arabic linguistics, the term 'truncation' makes sense as an umbrella term for various types of morphophonological shortening, notably in the imperative and the jussive of the verb, but also in the nominal realm (→ abbreviations; → compounds). Truncation phenomena have also caught the attention of theorists working in the framework of prosodic morphology (cf. McCarthy and Prince 1990:22; McCarthy and Prince 1998:289ff.; → morphology), one important point being the 'minimal word' in the context of hypocoristic forms, notably the → diminutive (typically in an endearing meaning).

2. TRUNCATION IN CLASSICAL ARABIC AND MODERN STANDARD ARABIC

Except for the apocopate/jussive and the imperative, truncation is not a prominent feature in either Classical or Modern Standard Arabic. When occurring, word-final shortening can affect both vowels and, in rarer circumstances, consonants. The Arabic term *jazm* – a purely morphological notion – denotes both the 'apocopate' (Greek 'cut off'), as surfacing in the negated past, *lam yafʿal* 'he didn't do', and the 'jussive' *yafʿal* 'let him do'. While the apocopate can be associated diachronically with the Akkadian preterite *iprus* 'he cut off' and the Hebrew consecutive *wayyiqṭol* 'and he killed', the jussive has a precedent in the Akkadian precative *liprus* 'let him cut off', and vetitive *ayy-iprus* 'let him not cut off'. It may be noted that the jussive has a *yv*C_1C_2*v*C_3 structure also in those branches of Semitic (Akkadian and Ethio-Semitic) whose imperfect conjugation has a *yv*C_1*v*C_2C_2*v*C_3 structure. In the case of Chaha, a *yä*C_1*ə*C_2c_3 structure is found, depending on the position of C_2 and C_3 on the sonority scale (cf. Leslau 1964; McCarthy and Prince 1995:330f.). Paradigmatically, the imperative (2nd person) stands in a suppletive position vis-à-vis the jussive (1st and 3rd person), which is also used for the negation of the imperative: *ifʿal* 'do!' vs. *lā tafʿal* 'don't do!'. For easy reference, Table 1 gives an overview of the basic forms of the imperfect (3rd pers. masc. sg.), jussive (3rd pers. masc. sg.) and the imperative (2nd pers. masc. sg.) of sound verbs in contrast to weak verbs that are shortened in the prefix conjugations and in the imperative, both word-externally and word-internally.

Table 1. Imperfect, jussive, and imperative forms

	imperfect	apocopate/ jussive	imperative
sound	*yafʿalu*	*yafʿal*	*ifʿal*
I*w/y*	*yaqifu/* *yaysiru*	*yaqif/* *yaysir*	*qif/isir*
II*w/y*	*yaqūlu/* *yasīru*	*yaqul/* *yasir*	*qul/sir*
III*w/y*	*yadʿū/* *yarmī*	*yadʿu/* *yarmi*	*udʿu/irmi*
$C_2 = C_3$	*yaruddu*	*yardud ~* *yarudda/i/u*	*urdud ~* *rudda/i/u*

Some verbs, notably 'doubly weak' verbs, undergo even farther reaching truncation, for instance in the case of *raʾā* 'to see', whose basic imperfect, apocopate/jussive, and imperfect forms are *yarā*, *yara*, and *rah*, respectively. As in the case of the imperative *fih* of the doubly weak verb *wafā* 'to fulfil', the final *h* in *rah* 'see!' (the so-called *hāʾ as-sakt*) seems to have the function of upholding a 'minimal word constraint' (CvC structure) for graphically independent Arabic words (but see Versteegh 2004 on the status of 'one-letter words' in Arabic grammar).

Word-final vowel shortening in → pause (*waqf*), as practiced in the pronunciation of Classical Arabic (see, for instance, Birkeland 1940) and as being the norm in Modern Standard Arabic, is also a case in point. Sībawayhi devotes 15 chapters to this topic in the *Kitāb* (Chaps. 490–504). While a word-final short vowel drops completely, a word-final long vowel is shortened, e.g. *mā* > *ma*. While → nunation (*tanwīn*) is elided altogether in prose (except in the accusative), the case-marking vowel in an indefinite case ending may be replaced by the respective long vowel in poetry, e.g. *marartu bi-zayd* (prose) vs. *marartu bi zaydī* (poetry) 'I passed Zayd' (for details, see Wright 1967:II, 368–373; for word-final orthography and phonotactics in general, see Diem 1981). Occasionally, long final vowels (notably *ī*) in the *Qurʾān* are shortened in context as well. Examples that occur passim include *rabb-i* 'my Lord', *fa-ʿbudū-ni* 'worship [pl.] me!' and *al-wādi* 'the valley'.

Vollers (1906:156–159) mentions a number of Qurʾānic *qirāʾāt* in which the feminine ending (*tāʾ marbūṭa*) is clipped before a suffix pronoun, e.g. *ʿidda-hū* for *ʿiddata-hū* 'his preparation [acc.]'.

In some instances, the absence of nunation has a specific grammatical function, as in the vocative *yā zaydu* 'oh, Zayd', or in the absolute negation *lā ʾilāha* '[there is] no God' (cf. e.g., Birkeland 1940:43).

Due to the tendency to integrate nouns with more than four radicals into quadriradical patterns, the last radical of a noun may be clipped both in the formation of the plural and in a process of linear derivation (typically by means of the → *nisba*): *ʾandalīb* > *ʾanādil* 'nightingale', *ḥaḍramawt* > *ḥaḍramī* 'one from Ḥaḍramawt'.

The consonant *n* is occasionally elided word-finally, notably in poetry, due to *ḍarūrat aš-šiʿr* → 'poetic license' lit. 'necessity [to uphold meter]'. Examples include dual formations like *xuṭṭatā* 'two alternatives' (*xuṭṭatāni*), shortened tribal names such as *ba-l-ʿanbari* for *banū l-ʿanbari*, and the clipped apocopate *lam yaku* 'he wasn't'. In rare instances, whole (groups of) syllables can be truncated, e.g. *mar* for *marḥaban* 'welcome' (cf. Wright 1967:II, 379–382).

Clipping processes are in no way restricted to the coda of a word. The process of haplological syllable elision can occur word-initially, word-medially, and word-finally. An example of the first kind is the reading of *taqattala* for *tataqattala*, as applied by the Kufan *Qurʾān* readers (see, for instance, Brockelmann 1908:261). Some instances of haplology have found their way into the Cairene *textus receptus* of the *Qurʾān*, e.g. *lā takallamu nafsun* 'nobody talks' (Q. 11/105). Poetic licenses such as *yasṭīʿu* for *yastaṭīʿu* 'he is able' are instances of word-medial truncation (cf. Wright 1967:II, 380). An example of word-final truncation is the variant *yaqtulū-nī* for *yaqtulūna-nī* 'they are killing me' (for an analysis of such reductions in terms of preference laws for syllable structure, see Vennemann 1988; for an analysis in terms of Optimality Theory, see Edzard 2000).

In the nominal realm, clipping occurs in a more systematic way in the context of blending, a process to which the Arab grammarians referred as *naḥt* (lit. 'sculpture'; → compounds). Clipping in such formations typically applies to the end of the first constituent and the beginning of the second. Grünert (1893) provides an extensive overview (cf. also for Classical Arabic: Stetkevych 1970:49ff.; and for Modern Standard Arabic: Monteil 1960:131–142; Versteegh 2001:181ff.; Badawi a.o. 2004:751–762; Holes 2004:312ff.; this phenomenon is pervasive in Modern Hebrew, see, e.g., Kutscher 1982:221). Clipped prefixes in neologisms such as *bay-ṯaqafī* 'intercultural' and *qab-tārīxī* 'prehistorical' have not always caught on in common language use. Acronymity (cf., e.g., Blau 1981:174) can also be subsumed under the category of 'clipping' (→ abbreviations).

In Classical Arabic, mainly in verse, the truncation of a final consonant in proper names in the vocative is known as *tarxīm* 'softening' (see Wright 1967:I, 174; II, 88). The process applies to CvvCvC nouns (e.g. *māzin* > *māzi*); CvCvvC nouns (e.g. *suʿād* > *suʿā*); CvCCvC nouns (e.g., *jaʿfar* > *jaʿfa*); and CvCCvvC nouns (e.g., *ʿuṯmān* > *ʿuṯma*—here the last vowel is also shortened; but not to CvCC or CvCvC nouns. In formal morpho-phonology, this circumstance has been interpreted as striking evidence for the 'minimal word', i.e., a structure that cannot be further phonologically diminished (cf. McCarthy and Prince 1990:22).

3. Truncation in the Arabic dialects

In the Arabic dialects, the opposition between the imperfect and the apocopate/jussive no longer exists, because the inflectional mood endings of the imperfect are reduced. The resulting short imperfect forms have a modal aspect; nonmodal aspects are expressed by a number of preformative elements. In general, the imperative forms in the dialects are not markedly different from those in Modern Standard Arabic.

The abbreviation (viz. truncation) of words in the spoken language is also attested in the Arabic colloquials. Arbitrarily selected examples include *kamsayta* (< *kayfa ʾamsayta*) ‘good evening [lit. ‘how did you get in the evening?’]’ in Ṣanʿānī Arabic and *bi-šak* (< *bi-šakl*) in varieties of Sudanese Arabic (personal information by Janet Watson and James Dickins, respectively).

An example of an abbreviated proper noun is *ʿAbbūd* for *ʿAbdallāh*. An example of a productive hypocoristic diminutive pattern is C_1*a/i/ə*C_2C_2*o*, as in *Widdo* for *Widād* and *Fatto* for *Fatḥallāh* (cf. Wild 1982:155). In this latter case, the third radical is truncated (in cases where $C_2 = C_3$). Abbreviated forms of foreign terms (sometimes themselves of Arabic origin) likewise occur in spoken Arabic. An example is *Alex* (*ʾāliks*) as a short form of *Alexandria* (*al-ʾiskandariyya*) in Arabic-spoken dialogues among students of the American University in Cairo.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCE

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān ibn Qanbar Sībawayhi, *al-Kitāb*. Ed. ʿAbd as-Salām Muḥammad H̱ārūn. 5 vols. Cairo: al-Hayʾa al-Miṣriyya al-ʿĀmma li-l-Kitāb, 1966–1977.

SECONDARY SOURCES

Badawi, El-Said, Michael G. Carter, and Adrian Gully. 2004. *Modern Written Arabic: A comprehensive grammar*. London: Routledge.

Birkeland, Harris. 1940. *Altarabische Pausalformen*. Oslo: Norske Videnskabs-Akademi.

Blau, Joshua. 1981. *The renaissance of Modern Hebrew and Modern Standard Arabic: Parallels and differences in the revival of two Semitic languages*. Berkeley and Los Angeles: University of California Press.

Booij, Geert, Christian Lehmann, and Joachim Mugdan (eds). 2000. *Morphologie: Ein internationales Handbuch zur Flexion und Wortbildung*. Berlin and New York: Walter de Gruyter.

Brockelmann, Carl. 1908. *Grundriß der vergleichenden Grammatik der semitischen Sprachen*. I. *Laut- und Formenlehre*. Berlin: Reuther und Reichard. (Repr., Hildesheim: G. Olms, 1999.)

Diem, Werner. 1981. “Untersuchungen zur frühen Geschichte der arabischen Orthographie. III. Endungen und Endschreibungen”. *Orientalia* 50:4.322–383.

Edzard, Lutz. 2000. “Sībawayhi’s observations on assimilatory processes and re-syllabification in the light of Optimality Theory”. *Journal of Arabic and Islamic Studies* 3.48–65.

Grünert, Max. 1893. “Die Mischwörter im Arabischen (Quellenbeiträge zur Kenntniss der ‘*naʿt*’)”. *Actes du Huitième Congrès International des Orientalistes, tenu en 1889 à Stockholm et à Christiania [Oslo]*. II, 133–172. Leiden: E.J. Brill.

Holes, Clive. 2004. *Modern Arabic: Structures, functions, and varieties*. 2nd rev. ed. Washington, D.C.: Georgetown University Press.

Kreidler, Charles. 2000. “Clipping and acronymy”. Booij a.o. (2000:956–963).

Kutscher, Eduard. 1982. *A history of the Hebrew language*. Jerusalem: Magnes Press.

Leslau, Wolf. 1964. “The jussive in Chaha”. *Language* 40.53–57.

McCarthy, John and Alan Prince. 1990. “Prosodic morphology and templatic morphology”. *Perspectives on Arabic linguistics*, II. ed. Mushira Eid and John McCarthy, 1–54. Amsterdam and Philadelphia: J. Benjamins.

——, and Alan Prince. 1995. “Prosodic morphology”. *The handbook of phonological theory*, ed. John Goldsmith, 318–366. Oxford: Blackwell.

——, and Alan Prince. 1998. “Prosodic morphology”. *The handbook of morphology*, ed. Andrew Spencer and Arnold M. Zwicky, 283–305. Oxford: Blackwell.

Monteil, Vincent. 1960. *L’arabe moderne*. Paris: C. Klincksieck.

Stetkevych, Jaroslav. 1970. *The modern Arabic literary language: Lexical and stylistic developments*. Chicago: University of Chicago Press.

Vennemann, Theo. 1988. *Preference laws for syllable structure and the explanation of sound change*. Berlin: Mouton de Gruyter.

Versteegh, Kees. 2001. *The Arabic language*. 2nd ed. Edinburgh: Edinburgh University Press.

——. 2004. “One letter does not make a word: The doubly weak imperatives in Arabic”. *Words, texts and concepts cruising the Mediterranean Sea: Studies on the sources, contents and influences of Islamic civilization and Arabic philosophy and science dedicated to Gerhard Endress on his sixty-fifth birthday*, ed. Rüdiger Arnzen and Jörn Thielmann, 131–141. Leuven: Peeters.

Vollers, Karl. 1906. *Volkssprache und Schriftsprache im alten Arabien*. Strasbourg: K.J. Trübner. (Repr., Amsterdam: APA – Oriental Press, 1981.)

Wild, Stefan. 1982. “Arabische Eigennamen”. *Grundriß der arabischen Philologie*. I. *Sprachwissenschaft*, ed. Wolfdietrich Fischer, 154–164. Wiesbaden: L. Reichert.

Wright, William. 1967. *A grammar of the Arabic language*. Trans. from German. 3rd ed., rev. by W. Robertson Smith and M.J. de Goeje. 2 vols. Cambridge: Cambridge University Press.

LUTZ EDZARD (University of Oslo)

Ṯuluṯ

Although the *ṯuluṯ* 'one-third' script appears in some early classical texts, not much is known about this ancient script, except that one of its main features, in its smaller version, was the roundness of its letters. Ibn an-Nadīm (d. 380/990), for example, reports that the large *ṯuluṯ* (*aṯ-ṯuluṯ al-kabīr aṯ-ṯaqīl*) was 'invented' by Quṭba (d. 154/771) as one of the four leading scripts, the others being *jalīl, ṭūmār al-kabīr*, and *niṣf aṯ-ṯaqīl*. The smaller (*xafīf*) version apparently developed from a small and round script called *al-mudawwar aṣ-ṣaġīr*, which was used for writing in registers (*dafātir*) and recording traditions (*ḥadīṯ*) and poetry (*šiʿr*; Ibn an-Nadīm, *Fihrist* 17; Abbott 1939:31).

Ṯuluṯ script became established as the principal chancery script, with pronounced curvilinear features, as a result of the reform of handwriting by Ibn Muqla (d. 328/940) and later Ibn al-Bawwāb (d. 413/1022), during the 4th/10th and early 5th/11th centuries. A good early example of the use of *ṯuluṯ* can be found in a multiscript composition presenting a collection of poems by Salāma ibn Jandal executed before 456/1064. Even though its attribution to Ibn al-Bawwāb is disputed, there appears to be no doubt that the manuscript was produced in the middle of the 5th/11th century (Rice 1955:22; Blair 2006:169).

By the 7th/13th century, *ṯuluṯ* was practiced from Egypt to Iran. Indeed, it became established as one of the 'Six Pens' in the Eastern tradition (Iraq and Iran), as well as the main curvilinear chancery script in Mamluk Syria and Egypt.

Just like → *muḥaqqaq, ṯuluṯ* in the Arabic Mamluk tradition is one of the two fundamental scripts (*ʾaṣl*) from which others are derived (*farʿ*). This same tradition reports that *ṯuluṯ* is called thus from the fact that only one-third of its letter forms are rectilinear, or because it was written with a pen whose nib cut was eight horsehairs wide, i.e. one-third of the ancient *ṭūmār* script, or that it was written on the size of paper known as *ṯulṯ* 'one-third'. Like all major scripts, it had its larger (*ṯaqīl, kabīr*) and smaller (*xafīf*) varieties (Gacek 1987:127; 1989:147). Furthermore, al-Qalqašandī (*Ṣubḥ* III, 58) reports that it was written with a pen that was cut obliquely (*muḥarraf*) because it used hairlines (*tašʿīrāt*) to link letters and words, which can only be executed with such a pen.

Being the largest script in the curvilinear (*layyin, muqawwar, muraṭṭab*) family, *ṯuluṯ* script was serifed, with a pronounced right-sloping head-serif (*tarwīs*). Its *ʾalif*, unlike the *ʾalif* of *muḥaqqaq*, was slightly bent and had a left-turned (*muḥarraf*) tail/foot. Some sources describe the shape of this *ʾalif* as "a man looking at his feet", and being either seven or nine rhombic dots in height. Indeed, the foot-serif on the *ʾalif* often joins to the following letter. *Ṯuluṯ*, moreover, leaves all the counters ('eyes') of its letters open (*maftūḥ, munawwar*), as opposed to closed (*maṭmūs, muʿallaq*). Just as in *muḥaqqaq*, its most visible feature and the one that distinguishes the two scripts, lies in the shape of the descenders (sublinear strokes). Here most of the descenders, which fall quite far below the baseline, on such letters as *jīm, rāʾ, sīn, ṣād, ʿayn, mīm, nūn*, curve upward (*muqawwar*), while the tails of some of them are joined to the next letter by means of hairlines (*tašʿīrāt*; Gacek 1989). Indeed, the use of hairlines is characteristic of the whole of the *ṯuluṯ* family of scripts. In *ṯuluṯ*, in contrast to *muḥaqqaq*, we find many assimilated/contracted (*mudġam, muʿallaq*) letterforms. Thus, for instance, the final *hāʾ* often has the form of *hāʾ muḥdawdiba*, that is, a *hāʾ* 'bowed upward', as opposed to a *hāʾ* with an open counter (*hāʾ mardūfa*, *muḥaqqaqa*, ـه), which is typical of *muḥaqqaq* script. Also, again in contrast to *muḥaqqaq, ṯuluṯ* appears not to have favored the use of the 's'-shaped *kāf* (*al-kāf al-mabsūṭa*), in its initial and medial positions. Instead, we find proportionately more of the other type, the cross-barred *kāf*, ك (*al-kāf al-maškūla*).

In the chancery, *ṯuluṯ* was used for important documents, such as edicts, whereas in codices it was used mostly for book titles and chapter headings, especially in manuscripts of the *Qurʾān* (Gacek 1989).

Ṯuluṯ remained the ideal style for epigraphy and was used on virtually any material and anywhere. Verses were written on pillows and curtains, goblets and flasks, garments and headgear, belts and kerchiefs, golden and silver vessels, as well as on porcelain (Schimmel 1984:22, 25–26; Soucek 1979:14).

For monumental inscriptions, a large (*jalīl*) *ṯuluṯ* form was used (→ epigraphy; see Fig. 1).

A number of disciples of Yāqūt al-Mustaʿṣimī (d. 698/1298) excelled both in Qurʾānic calligraphy and in the layout of huge architectural inscriptions. Thus, for instance, one of the outstanding examples of the use of *ṯuluṯ* in Iranian architecture involved the calligrapher Ḥaydar al-Mašhadī, one of the six famous pupils of Yāqūt. Probably the best known example of his work is a stucco *miḥrāb* in the mosque in Isfahan, dated 710/1310. The inscriptions, carved on three levels, are in various sizes of *ṯuluṯ* on a floral arabesque ground (Blair 2006:253–255).

Among the Ottoman architectural inscriptions in *ṯuluṯ* one should mention Bayazid's mosque in Istanbul, inscribed by Ḥamd Allāh al-ʾAmāsī, as well as the Suleymaniye (Istanbul) and Selimiye (Edirne) mosques designed by ʾAḥmad Qaraḥisārī and Ḥasan Jalabī (Celebi) in the middle of the 10th/16th century. And in 11th/17th century India, we have the superb inscriptions on the Taj Mahal (outside of Agra) executed by ʿAbd al-Ḥaqq, known as ʾAmānat Xān Šīrāzī (Blair 2006:479, 496–497, 546–548).

Ottoman calligraphers used large (*celi*) *ṯuluṯ* not only for architecture but also for individual compositions or panels (*levha*); they were later mounted, framed, and hung on walls. In the Maghreb, a type of *ṯuluṯ* known as *mašriqī* script (→ *maġribī*), was also used for titles and chapter headings, as well as in large inscriptions. Other uses of *ṯuluṯ* include calligraphic specimens (*qiṭʿa*), alphabetic exercises (*mufradāt*), practice sheets, pilgrimage (*ḥajj*) certificates, and zoomorphic calligraphy (e.g. the famous calligraphic lion drawn in Iran in the 10th/16th century, or the stork containing the *basmala* penned by the Ottoman calligrapher ʾIsmāʿīl Zuhdī (d. 1213/1798; Blair 2006:215–216, 449–451, 506–507).

A smaller version of *ṯuluṯ* (*xafīf aṯ-ṯuluṯ*) was known as *tawqīʿ* 'signature'. It is also referred to in its plural forms, *at-tawāqīʿ* or *at-tawqīʿāt*. The *tawqīʿ* is characterized by an even more liberal use of hairlines (*tašʿīrāt*) to create ligatures between letters and words. The larger version of *tawqīʿ* was known also as *at-tawāqīʿ aṯ-ṯuluṯiyya* or *at-tawāqīʿ al-kibār* and its smaller version as *at-tawāqīʿ ar-riqāʿiyya*, showing clearly the close relationship between these scripts (Gacek 1989). The term *at-tawqīʿ al-muṭlaq* for a regular highly ligatured script is used by al-Qalqašandī. This script was probably what Ibn Xalaf, the 5th/11th-century author of *Mawādd al-bayān*, had in mind when he made a distinction between *muḥaqqaq* and *muṭlaq* (Soucek 1979:14; Stern 1964:105).

According to Mamluk authors, the *tawqīʿ* script was to be seriffed or mostly seriffed, even though in practice we find this script also serifless (sans serif). From this script, it is believed, the Persian *taʿlīq* developed (→ *nastaʿlīq*). A highly ligatured version of the *tawqīʿ* was *musalsal* 'chained; chainlike', with its *lām ʾalif* looking like links in a chain. It was difficult to read, though, and Jalāl ad-Dīn Rūmī (d. 672/1273), complains in one of his poems, "You wrote to me in *musalsal*, that means you do not want me to read it!" (Schimmel and Rivolta 1992:15). Like its larger version, *ṯuluṯ, tawqīʿ* was rarely used for full texts. An excellent example, however, of the text fully penned in *tawqīʿ*, with *ṯuluṯ* for chapter headings, is appropriately a copy of the manual on the art of letter writing, *Tashīl as-sabīl ʾilā taʿallum at-tarsīl*, by ʾAbū ʿAbdallāh Muḥammad al-Ḥumaydī (d. 488/1095), penned in the 7th/13th century and published in facsimile in Frankfurt am Main in 1985.

Ṯuluṯ or *tawqīʿ* written in gold ink was known as *ḏahab* (*qalam aḏ-ḏahab* 'the script of gold'), except that it had no hairlines due to the fact that its letters were outlined (*muzammak*) in another color. The earliest example of the 'golden script' in the form of *tawqīʿ* (probably not *ṯuluṯ*, on account of its medium size) can be found in the Chester Beatty Qurʾān of Ibn al-Bawwāb, dated 391/1000 or 1001, which he used for *sūra* headings, *juzʾ* statements, an omission (inscribed in a *tabula ansata*), and tables (Rice 1955).

The smaller version of *tawqīʿ* script, in turn, was *riqāʿ*, properly *qalam ar-riqāʿ*, not to be confused with the Turkish *riqʿa* (*rikʾa*; → *ruqʿa*). Ibn an-Nadīm mentions that it was derived from *xafīf aṯ-ṯuluṯ al-kabīr* and used for edicts (*tawqīʿāt*) and the like (Ibn an-Nadīm, *Fihrist* 17). It was finer, more ligatured, and more curvilinear than *tawqīʿ*. Although many surviving examples of *riqāʿ* have head-serifs, according to some major sources the use of head-serifs in this script was optional. Other features may include the *ʾalif* slightly inclined to the right (*al-mumāla ʾilā jihat al-yamīn*; Gacek 1989:146; 2003). In the Persian context, *riqāʿ* and *tawqīʿ* were difficult to distinguish, and for

Qāḍī ʾAḥmad, the author of *Gulistān-i hunar*, an early-11th/17th-century Persian treatise on calligraphers and painters, the terms are virtually synonymous (Soucek 1979:12, 14).

The *tawqīʿ/riqāʿ* scripts in Ottoman Turkey came to be known as *xaṭṭ al-ʾijāza,* used almost exclusively for the writing of diplomas (*ʾijāzāt*), hence its appellation, and in colophons (Schimmel 1984:15). This script often has a long head-serif (*zulf*) which wraps around the shaft of such letters as *ʾalif* and *lām* (Gacek 2003). Another feature of the *ʾijāza* script is the highly intertwined (*musalsal*) final *hāʾ*.

The *ṯuluṯ* script in the Mamluk period also appears to have had a variant known as *qalam al-ʾašʿār* (also *aš-šiʿr, al-mušaʿʿar, al-muʾannaq*). Although some authors, such as aṭ-Ṭayyibī (Ṭībī), viewed it rather as an offshoot of *muḥaqqaq* and *nasx*, or a hybrid of *muḥaqqaq* and *ṯuluṯ*, others viewed it as *ṯuluṯ* related. The script takes it name from its association with writing/copying poetry and not so much perhaps from its use of hairlines as previously thought (Gacek 1989:145).

A good example of this script is Baybars Jāšnagīr's Qurʾān (preserved in the British Library), executed by the celebrated calligrapher Ibn al-Waḥīd (d. 711/1311 or 1312). Aṣ-Ṣafadī (d. 764/1363) refers to it as *qalam al-ʾašʿār* and qualifies it as *ṯuluṯ kabīr*, whereas Ibn ʾIyās (d. ca. 930/1524) identifies this script in his work *Badāʾiʿ az-zuhūr* as *qalam aš-šiʿr* (James 1988:38; Blair 2006:345–349).

Ṯuluṯ and its variants survived in books printed by lithography, especially in Iran, as display scripts, and are still practiced today, alongside *nasx*, by many modern calligraphers.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Ibn an-Nadīm, *Fihrist* = Muḥammad ibn ʾIsḥāq ibn an-Nadīm al-Warrāq al-Baġdādī, *Kitāb al-fihrist*. Ed. Yūsuf ʿAlī Ṭawīl. Beirut: Dār al-Kutub al-ʿIlmiyya, 1996.

Qalqašandī, *Ṣubḥ* = ʾAḥmad ibn ʿAlī al-Qalqašandī, *Ṣubḥ al-ʾaʿšā fī ṣināʿat al-ʾinšāʾ*. 14 vols. Cairo: al-Muʾassasa al-Miṣriyya al-ʿĀmma, 1963–1965. (First publ., aṭ-Ṭabʿa al-ʾAmīriyya, 1913–1918.)

Secondary sources

Abbott, Nabia. 1939. *The rise of the North Arabic script and its Ḳurʾānic development, with a full description of the Ḳurʾān manuscripts in the Oriental Institute*. Chicago: University of Chicago Press.

Blair, Sheila S. 2006. *Islamic calligraphy*. Edinburgh: Edinburgh University Press.

Gacek, Adam. 1987. "Al-Nuwayrī's classification of Arabic scripts". *Manuscripts of the Middle East* 2.126–130.

——. 1989. "Arabic scripts and their characteristics as seen through the eyes of Mamluk authors". *Manuscripts of the Middle East* 4.144–149.

——. 2003. "The head-serif (*tarwīs*) and the typology of Arabic scripts: Preliminary observations". *Manuscripta Orientalia* 9:3.27–33.

James, David. 1988. *Qurʾāns of the Mamlūks*. London: Thames and Hudson.

Rice, David Talbot. 1955. *The unique Ibn al-Bawwāb manuscript in the Chester Beatty Library, Dublin*. Dublin: Emery Walker.

Schimmel, Annemarie. 1984. *Calligraphy and Islamic culture*. New York: New York University Press.

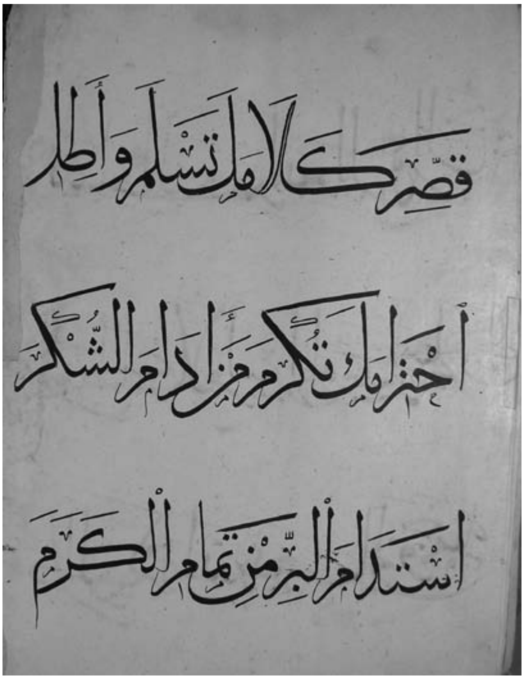
—— and Barbara Rivolta. 1992. "Islamic calligraphy". *Metropolitan Museum of Art Bulletin* [published independently].

Soucek, Priscilla P. 1979. "The arts of calligraphy". *The arts of the book in Central Asia, 14th–16th centuries*, ed. Basil Gray, 7–33. Paris and London: Serinda.

Stern, Samuel Miklos. 1964. *Fāṭimid decrees: Original documents from the Fāṭimid chancery*. London: Faber and Faber.

ADAM GACEK (McGill University)

Figure 1. *Jalīl aṯ-ṯuluṯ* by Muḥammad aṭ-Ṭayyibī (Ṭībī). ca. 908/1505 (Istanbul). TSM K.882. f. 24b)



Tunis Arabic

1. GENERAL

The Arabic spoken in Tunis is by nature something of a contact variety, given the vast population influx from all over → Tunisia during the 20th century. What is described here is the Muslim variety spoken by the majority of those who now live in Tunis, a koineized variety, which Singer (1984:28) describes as the "allgemeine stadttunisische Koiné" (general koine of the town of Tunis), rather than the somewhat similar traditional dialect of the long-established dwellers of the medina (now mainly living in the *quartiers nords),* often referred to as *baldi*, a variety whose very existence is in jeopardy. The latter is what Singer (1984) describes; it includes distinctive forms used only by women.

Tunis Arabic functions as a de facto prestige variety (Gibson 2002), toward which some other varieties and speakers in Tunisia are shifting and which is widely used in the oral and visual media (especially theater and film) and for written slogans on billboards. Apart from these domains of widespread use, it can be found in written form in the popular weekly *al-Sarih*, as well as in cartoons in other papers, and has its own translation of *The little prince* (St. Exupéry 1997), along with collections of proverbs such as Balegh (1993), but beyond informal letters between friends and advertising slogans, it is not generally accepted in written form by the majority of the population.

The dialect is, in North African terms, solidly urban, and has much in common with the dialects of Kairouan, Bizerte, Nabeul, Sousse, and, to a lesser extent, Sfax, while being highly distinct, especially in morphology, from the Bedouin dialects spoken a few kilometers away. Singer (1984) and Marçais (1950) outline the commonalities and dialect types of Tunisian Arabic, while Gibson (1998) looks at the changes occurring in Tunisian Arabic today. The first reliable documentation on Tunis Arabic itself is found in Stumme (1893, 1896), while the most thorough and extensive description is Singer (1984), stretching to more than seven hundred pages. Other useful sources on the dialect are Cohen (1970), a comparison of Muslim and Jewish varieties; Talmoudi (1986); and Jabeur (1987). There are some published pedagogical materials (e.g. Jourdan 1956; Quitot 2002), while the distinctive Jewish variety, now less likely to be heard in Tunis than on the Boulevard de Belleville in Paris, or in Israel, is thoroughly documented by Cohen (1964, 1975).

2. LINGUISTIC DESCRIPTION

2.1 *Phonology*

2.1.1 Consonants

The established consonantal phonemes of Tunis Arabic are given in Table 1.

The bracketed phonemes are ones to which not all linguists would necessarily attribute phonemic status, but the case for their being separate phonemes is convincing. There are two sources for these bracketed consonants: the pharyngealized phonemes are internal developments, while the other three are due to substantial borrowing from French, and Modern Standard Arabic in the case of /ʾ/. Minimal pairs are not always easy to find for these tokens, but there are nonetheless many examples showing that these marginal forms are not conditioned variants, e.g. *bāb* [bɛːb] 'door'/*ḅāḅa* [ƀɑːƀɑ] '[my] father', alongside a minimal pair *gāz* [gɛːz] 'petrol'/*gāẓ* [gɑːƶ] 'gas'.

Table 1. Consonants of Tunis Arabic

(p)				t	ṭ		k	q		(ʾ)
b	(ḅ)			d			g			
m	(ṃ)			n	(ṇ)					
f		ṯ		s	ṣ	š		x	ḥ	h
(v)		ḏ	ḏ̣	z	(ẓ)	ž		ġ	ʿ	
				l	ḷ					
w				r	ṛ	y				

It is to be noted that the realization of the vowels in these words is dramatically different (see below, Sec. 2.1.2). Phonetic pharyngealization on the consonants themselves is relatively weak, the main realization being on adjacent vowels, and is being lost among some speakers; an example is in *sbāḥ* 'morning', where there is no vowel to carry any pharyngealization on the first consonant. Singer notices this phenomenon in *ṛāžil/ržāl* 'man/men', but attributes it to lack of consistency of pharyngealization in the stem. There are other words, such as *ṇāṇa* 'old lady', whose form, while not having any minimal or near-minimal pairs, cannot be accounted for by conditioned variation, and which lead us to posit an (admittedly rare) phoneme /ṇ/. Minimal pairs for the more commonly admitted phonemes /ṛ/ and /ḷ/ can be given, as in *žra* [ʒrɛ] 'he ran'/*žṛa* [ʒɍɑ] 'it happened', and *walla* [wɛlːɛ] 'or'/*waḷḷa* [wɑɫːɑ] 'by God!'.

Singer (1984:37–60) gives a full list of oppositions for each phoneme. Tunis Arabic has substantial borrowing from French, and many words and expressions used by those who do not speak French maintain /p/ and /v/, e.g. *pisīn* 'swimming pool', *mgarrap* 'suffering from influenza' (< French *grippe*), *ynarvisni* 'he annoys me', *gāriv* 'on strike' (< French *grève*).

/ʾ/ tends to occur in the learned register, in loans from Modern Standard Arabic, often in *maṣdar* forms at the onset of the word but also in other words like *bīʾa* 'environment' and sometimes *yisʾal* 'he asks', although the original Tunis form here is *yishal*; more highly educated speakers especially avoid this form.

Where the pronunciation of consonants is concerned, Tunis Arabic is relatively conservative, with interdentals retained (except in *tlāṯa* 'three' [but *ṯāliṯ* 'third'], and the variation *ṯamma ~ famma* 'there is/are') alongside /q/, as is the case with almost all Tunisian varieties in the first case, and urban dialects in the second. A substantial number of words do have /g/ as a reflex of **q*, such as *bagra* 'cow'; an extensive list is given by Baccouche (1972). There are minimal pairs, such as *daqdaq* 'to knock' and *dagdag* 'to grind'. **ḍ* has merged with **ḏ̣* into /ḏ̣/, and **j* gives /ž/, although this has assimilated to /z/ (evidently via /ž/) where another /z/ was already present in the word, giving us such words as *zazzār* 'butcher', *zūz* 'two', and the root *z-w-z*, e.g. *zāz* 'he passed by'.

2.1.2 Vowels

If we accept pharyngealization as being a property of consonants, then Tunis Arabic possesses three vowel qualities /i, a, u/, while length is also contrastive for all three, giving us six vowel phonemes, as in Modern Standard Arabic. The length distinction is suspended word-finally, with the final vowel being realized long in accent-bearing words of one syllable, as in *ža* [ʒɛː] 'he came', otherwise short, although underlying length is realized once a suffix is added to the word: *tˁašša* [ˈtʕɑʃːa] 'he dined', *ma-tˁaššāš* [matʕɑˈʃːaːʃ] 'he did not dine'. Note the concomitant shift of stress. The dialect maintains a robust distinction between all short vowels, with examples such as *qimt* 'I resided' and *qumt* 'I rose' showing this distinction to be maintained.

In non-*baldi* varieties, Old Arabic diphthongs are not preserved, while the feminine *baldi* variety has diphthongs in words such as *layl* 'evening', *lawn* 'color', which are otherwise *līl*, *lūn*. These traditional forms for women are now somewhat stigmatized among the younger generation in Tunis, despite the fact that they represent a variety that used to have prestige; one interviewee in Gibson (1998) admitted to having been made fun of at school for using these forms, and desisted. Many others stated that their mothers or grandmothers spoke this way, but they did not. There is no evidence for intermediate forms such as [leːl], [loːn] as found in Sfax or phonemically in places such as Monastir.

In nonpharyngealized environments, Tunis Arabic shows a strong fronting and closing of /ā/, which, especially among younger speakers, can reach as far as [eː]. The short vowel /a/ also shows this raising, but to a lesser extent, and this is not apparent word-finally.

2.1.3 Syllable

Tunis Arabic, like many other North African varieties, has a very different syllable structure from Modern Standard Arabic. CC in the onset is common, the nucleus may contain a short or long vowel, and the coda may have maximally three consonants. Word-internal syllables are generally heavy in that they either have a long vowel in the nucleus or consonant in the coda. Nonfinal Cv syllables where the vowel is short (i.e. light syllables) are very rare, and are generally loans from Modern Standard Arabic; generally short vowels in this position have been lost, resulting in the many initial CC clusters. We can see the results of this development in *žwāb* 'letter', which exists alongside the loan *žawāb* 'reply', in much of the verb morphology (see below), and in the vowel shifting and elision which occurs in possessed nouns, e.g.

dbiš	'stuff'	*ṛāžil*	'man'
dibši	'my stuff'	*ṛāžli*	'my husband'
dbišha	'her stuff'	*ṛāžilha*	'her husband'

This regular morphological change prevents nonfinal syllables from being light throughout, but it is not exceptionless; many words given in this article do not conform to the 'rule'. However, we assume that this aversion to light open syllables also accounts for the development of such forms as *ṛāžil* 'man', *lūġa* 'language', and *ṯīqa* 'trust'.

A coda in CCC occurs mainly with the suffixation of /š/ to verbs as part of the negative circumfix, as in *ma-šuft-š* 'I did not see', where the suffix is sometimes preceded by an epenthetic vowel. Phonetic epenthesis also occurs at a word boundary where the onset of the second word is CC and the preceding word ends in at least one consonant, e.g. *ṛāžil qwi* ['ɫɑːʒil ə'qwiː] 'a strong man'. This is one of the few cases where a Berber substratum may be usefully invoked.

2.1.4 Word stress

Stress is predictable, occurring on the final syllable if this is heavy, with the rime being at least v̄C or vCC, otherwise on the penult, e.g.

žā́bit	'she brought'
ma-žābítš	'she did not bring'
žābítha	'she brought it/her'
žābithā́lu	'she brought it to him'
ma-žābithūlhā́š	'she did not bring him to her'

This is helpful in defining the phonological word, which here, by two tests, includes the indirect object pronominal suffix.

2.2 *Morphophonology*

In common with other Arabic varieties, Tunis Arabic has maintained the distinction between construct and absolute state. This is expressed morphologically in regular singular feminine nouns and duals of body parts and the word *wāldīn* 'parents'.

The feminine morpheme represented in Modern Standard Arabic by the *tāʾ marbūṭa* behaves thus:

kišra	'skin'	*karhaba*	'car'
kišrti	'my skin'	*karhabti*	'my car'
kišrit uxti	'my sister's skin'	*karhabit škūn*	'whose car?'

Note the presence and absence of the /i/ vowel in the construct state, which depends on whether a vowel follows the /t/ within that word. The word *karhaba*, derived by metathesis from *kahrabāʾ*, a word without *tāʾ marbūṭa* in Modern Standard Arabic, nevertheless expresses the construct state with /t/, by analogy with other feminine words ending in /a/. For the same reason, we find more instances of *maˁnitha* 'its meaning; that is to say' than the more conservative *maˁnāha*. 'Woman', which in absolute state is *mra*, has a special construct state *mart*.

Duals, where present, also have a special construct state: *ˁīnīn* 'two eyes'; *ˁīnīya* [ʕiːneːja] 'my two eyes'; Singer (1984:175) notes the lowering of the high vowel in such environments.

The construct state is not used when two nouns are contiguous through → *tamyīz*, as in *dabbūza ḥlīb* 'a bottle of milk'. However, we find the construct state with these two words for a different meaning, as in *dabbūzit ḥlīb* 'a milk bottle'.

An alternative to the construct state is the uninflected particle *mtāˁ*, as in *karhaba mtāˁi* 'my car'. The length of the vowel varies, depending on whether there is a vowel-initial suffix following, and 'hers' and 'theirs' come out as *mtaḥḥa* and *mtaḥḥum*.

2.3 *Morphology*

2.3.1 Pronouns

Tunis Arabic is distinctive among Arabic dialects in having no gender marking whatsoever in the 2nd person, a characteristic it shares with other urban Tunisian dialects and Maltese. This leaves us with only seven personal pronouns (Table 2).

Table 2. Pronouns in Tunis Arabic

	Personal pronoun	Direct suffix object suffix	Indirect object suffix	Preposition with suffix	Possessive
3rd sg. m.	*hūwa*	*-u, -h**	*-lu*	*fīh*	*-u, -h**
3rd sg. f.	*hīya*	*-ha*	*-ilha*	*fīha*	*-ha*
3rd pl.	*hūma*	*-hum*	*-ilhum*	*fīhum*	*-hum*
2nd sg.	*inti*	*-ik, -k**	*-lik*	*fīk*	*-ik, -k**
2nd pl.	*intūma*	*-kum*	*-ilkum*	*fīkum*	*-kum*
1st sg.	*āna*	*-ni*	*-li*	*fīya*	*-i, -ya**
1st pl.	*aḥna*	*-na*	*-ilna/-inna*	*fīna*	*-na*

* = after vowels

The only form not noted here is the 3rd person singular masculine direct object suffix when followed by an indirect object in the same word, which is *-hū-*, as in *žābithūlha* 'she brought him to her', which also indicates the inclusion of pronominal indirect objects into the verb. Indirect object suffixes do not start with /i/ when they follow a vowel. Other Arabs quickly find the use of *inti* with all interlocutors somewhat distinctive.

2.3.2 Demonstratives

The usual demonstrative for 'this' is *ha* ~ *haḏ* for both genders, as in *haḏ ṛṛāžil* 'this man'. When contrast is indicated, then *haḏāya* ~ *haḏa* (masc.), *haḏīya* ~ *hāḏi* (fem.), or *haḏūma* (pl.) may be added after the noun, as in (*ha*) *lktāb haḏāya* 'this book'. For more distant objects *hak* functions in the same way as *ha* ~ *haḏ* before the noun, and *haḏāka* (masc.), *hāḏīka* (fem.), and *haḏūkum* (pl.) after it.

2.3.3 Function words

The relative pronoun, used for definite nouns only, has the form *illi*, and *waqtilli* ~ *waqt* for temporal clauses. Many interrogative pronouns and adverbs are formed with the element *-āš*:

āš ~ *šnūwa* ~ *šnīya*	'what?'
šnūma	'what [pl.]?'
škūn	'who?'
āna	'which?'
wīn ~ *fīn*	'where?'
lwīn	'whither?'
mnīn	'whence?'
ʿlāš ~ *lwāš*	'why?'
kīfāš	'how?'
waqtāš	'when?'
qaddāš	'how many/much?'

Āš is not used for emphasis, and is used mainly within a verbal sentence (but *āš qawlik* 'what is your opinion?'). It is also the form used when the interrogative is governed by a preposition, e.g. *fāš, bāš* 'in what, with what?'. The other two forms were evidently originally distinguished for gender, but most speakers make no such distinction.

The final four words can be pronounced with the final /š/ replaced by /h/, especially when pronounced in isolation. Interrogatives come first in the clause, except in echo questions, or where they are to be given special emphasis.

The definite article is found in various forms. Before sun letters it is represented by /i/ and a doubling of the initial consonant, e.g. *iššams* 'the sun'. /ž/ normally functions as a sun letter, e.g. *iždīd* 'the new one', but this is not consistent with learned vocabulary, e.g. *ilžarīda* ~ *ižžarīda* 'the newspaper'. Before moon letters, the article is *il-* where the onset has one consonant *ilgamra* 'the moon', and *li-* where the word starts with two: *liktāb* 'the book'.

Conjunctions include *u* (*w* before vowels) 'and', *walla* 'or', *amma* and *lākin* 'but', (*ʿla*) *xāṭr* 'because', *ki* ~ *kīf* 'as', and *yāxxi* 'so, and then' (especially in narrative). The latter also functions as an interrogative marker (see below).

The existential is *famma* ~ *ṯamma*, the first form being the more common. It takes the verbal negative circumfix, and uses the verb *k-w-n* to carry aspect and tense marking.

Prepositions include *fi* 'in, at', *b* 'with [instrumental]; for', *l* 'to', *maʿ* 'with', *min* 'from', and *ʿla* 'on'. The latter two have special forms when found before the definite article, *m*, *ʿa*, as in *milḥanūt* 'from the shop', and *ʿaṭṭāwla* 'on the table'. When *maʿ* has a pronominal complement, it takes the form *mʿā* as in *mʿāya/mʿāh* 'with me/him', and *b* and *l* have allomorphs *bī* and *lī* in the same environment.

2.3.4 Adjectives

Adjectives (including participles) are marked for gender in the singular but not in the plural, e.g. *ṛāžil kbīr/mra kbīra/ržāl kbār/nsa kbār* 'an old man/an old woman/old men/old women'. The regular ending for the plural is *-īn*, as in *xaybīn* 'bad [pl.]'. Inanimate objects can take either a feminine singular adjective or a plural one (with the congruent verb agreement). There is much variation on this, however, *nās* 'people' generally takes singular feminine.

2.3.5 Adverbials

Common adverbials include *tawwa* 'now', *waqtha* 'then', *minqbil* 'beforehand', *mbaʿd* 'afterward', *lbāriḥ* 'yesterday', *wutt lbāriḥ* 'the day before yesterday', *lyūm* 'today', *ġudwa* 'tomorrow', *baʿd ġudwa* 'the day after tomorrow', *hūni* ~ *hna* 'here', *ġādi* ~ *ġadīka* '[over] there', *hakka* 'in this way'.

2.3.6 Numerals

When counting objects, *wāḥid* (masc.), *waḥda* (fem.) 'one', unlike other numbers, is placed after the noun. It is also the only numeral to show any gender agreement. 'Two' is *zūz* when in apposition to a noun, but when counting sequentially, *ṯnīn* is used, also in compound numerals: *ṯnīnuʿašrīn* 'twenty-two'. From three to twenty we have *tlāṯa, arbʿa, xamsa, sitta, sabʿa, ṯmānya, tisʿa, ʿašra, ḥaddāš, ṯṇāš, ṯluṭṭāš, arbaṭṭāš, xumsṭāš, sṭṭāš, sabʿaṭṭāš, ṯimnṭāš, tsaʿṭāš, ʿašrīn*. The numbers 11–20 take a final *-n* when in apposition with a noun, e.g. *ḥaddāšn ktāb* 'eleven books'. *Mīya* 'one hundred' behaves in a similar way, becoming *mīyit* in the same environment. The dual form is used with counting units, as well as some body parts, e.g. *kaʿbtīn yuġūrt* 'two yogurts'. To say 'two or three hours', a combination of dual and figures is used: *sāʿtīn tlāṯa*.

Table 3. Verbal Forms in Tunis Arabic

Form	Perfect	Imperfect	Participle
I	*ktib* 'to write'	*yiktib*	*kātib, maktūb*
	miss 'to touch'	*ymiss*	*mās, mamsūs*
	bna 'to build'	*yibni*	*bāni, mibni*
	qām 'to stand up'	*yqūm*	*qāyim, maqyūm*
	kla 'to eat'	*yākul*	*mākil, mūkūl*
II	*ʿarris* 'to marry'	*yʿarris*	*mʿarris*
	šammam 'to cause to smell'	*yšammam*	*mšammam*
	qayyid 'to record'	*yqayyid*	*mqayyid*
	šahha 'to provoke appetite'	*yšahhi*	*mšahhi*
	wakkil 'to feed'	*ywakkil*	*mwakkil*
III	*šārik* 'to share'	*yšārik*	*mšārik*
	nāda 'to call'	*ynādi*	*mnādi*
V	*tkallim* 'to speak'	*yitkallim*	*mitkallim*
	tʿadda 'to pass'	*yitʿadda*	*mitʿaddi*
VI	*tqābil* 'to meet [reciprocal]'	*yitqābil*	*mitqābil*
	tlāqa 'to find [reciprocal]'	*yitlāqa*	*mitlāqi*
VIII	*rtāḥ* 'to rest'	*yirtāḥ*	*mirtāḥ*
	stwa 'to become ripe'	*yistwi*	*mistwi*
X	*staḥšim* 'to behave modestly'	*yistaḥšim*	*mistaḥšim*
	stḥaqq 'to deserve'	*yistḥaqq*	*mistḥaqq*
	stanna 'to wait'	*yistanna*	*mistanni*
XI	*kbār* 'to grow old'	*yikbār*	*mukbār*
	rqāq 'to become thin'	*yirqāq*	*murqāq*
	byāḍ 'to become white'	*yibyāḍ*	*mubyāḍ*

2.3.7 Verb

2.3.7.1 Verbal Forms

Table 3 lists the common verbal Forms.

Note that weak-final verbs have only one pattern in Forms II-XI. Only in Form I is there variation on the vowel between different aspects. Form VIII is no longer productive. Participles for Forms II–XI can have either active or passive meaning, but in Form I there are two distinct participles. Form II often has a causative meaning; III can be reciprocal, along with VI; V is often reflexive; X coming or beginning to do something; and XI being or becoming, from an adjectival root, like the meaning of Form IX in Modern Standard Arabic. As with any Arabic variety, the semantics do not always match the morphology.

Roots with four radicals have two forms, a base form, e.g. *farkis/yfarkis*, participle *mfarkis* 'to look for', while the derived form has an initial /t/ with a passive or reflexive meaning, e.g. *narvis* 'to annoy', *tnarvis* 'to be annoyed', *mitnarvis* 'annoyed'. There are also many such verbs with a reduplicative root, e.g. *daqdaq* 'to knock'.

The passive in Tunis Arabic is formed with the prefixation of *t-* to the stem of the verb, e.g. *tšrab/yitšrab* 'he was/is drunk'. The vowel of both aspects here is copied from the active perfective form, e.g. *tqāl/yitqāl* 'it was/is said', *tšra/yitšra* 'it was/is bought'. 'To eat' has the forms *tākil/yitākil* ~ *yittākil*. It will be noted that the passive of Form II is identical to Form V, and many Form V verbs are in fact passive in meaning. See Maalej (n.d.) for further discussion of the use of the passive in Tunisian Arabic.

2.3.7.2 Aspect

Tables 4 and 5 give the forms of the verb in the imperfect and the perfect.

The plural forms of the triconsonantal verb are usually pronounced with some vowel in the cluster, most commonly [yiktibu], but this vowel is most probably epenthetic, not taking an accent. Forms such as [ykitbu] can also be heard, where the vowel, taking an accent, is phonemic. Forms II, III, V, VI, and X have vowel elision in the plural, as in *yfakkir* 'he thinks', *yfakkru* 'they think'. Note the retention of final weak root vowels in the conjugation, as in other urban Tunisian dialects. There are no mood prefixes in Tunis Arabic, hence there is no indicative/subjunctive distinction.

Where a 3rd person singular feminine perfect ending in *-it* is followed by a vowel-initial suffix, the /t/ doubles, avoiding an accented light syllable, e.g. *šāfittu* 'she saw him'. Forms

Table 4. Conjugation of the imperfect in Tunis Arabic

	I	II gem.	II*w/y*	III*y*	III*y*
3rd sg. masc.	*yiktib*	*yḥabb*	*ykūn*	*yimši*	*yinsa*
3rd sg. fem.	*tiktib*	*tḥabb*	*tkūn*	*timši*	*tinsa*
3rd pl.	*yiktbu*	*yḥabbu*	*ykūnu*	*yimšīw*	*yinsāw*
2nd sg.	*tiktib*	*tḥabb*	*tkūn*	*timši*	*tinsa*
2nd pl.	*tiktbu*	*tḥabbu*	*tkūnu*	*timšīw*	*tinsāw*
1st sg.	*niktib*	*nḥabb*	*nkūn*	*nimši*	*ninsa*
1st pl.	*niktbu*	*nḥabbu*	*nkūnu*	*nimšīw*	*ninsāw*

Table 5. Conjugation of the perfect in Tunis Arabic

	I	II gem.	II*w/y*	III*y*	III*y*
3rd sg. masc.	*ktib*	*ḥabb*	*kān*	*mša*	*nsa*
3rd sg. fem.	*kitbit*	*ḥabbit*	*kānit*	*mšāt*	*nsāt*
3rd pl.	*kitbu*	*ḥabbu*	*kānu*	*mšāw*	*nsāw*
2nd sg.	*ktibt*	*ḥabbīt*	*kunt*	*mšīt*	*nsīt*
2nd pl.	*ktibtu*	*ḥabbītu*	*kuntu*	*mšītu*	*nsītu*
1st sg.	*ktibt*	*ḥabbīt*	*kunt*	*mšīt*	*nsīt*
1st pl.	*ktibna*	*ḥabbīna*	*kunna*	*mšīna*	*nsīna*

II and V have vowel elision, as in *fakkir, fakkru*. The leveling of forms between the different classes of weak verbs stands out, as does the identity of the forms for 1st and 2nd person singular, and the movement of the initial vowel in regular verbs due to the preference for avoiding light syllables. Two verbs vary in the stem used for the different aspects: ‘to take’ and ‘to eat’ have *āxiḏ ~ āxu, ākil* as stems in the imperfect, with imperatives *xūḏ ~ xu, kūl* and *xḏa, kla* as perfect. The verb *ža/yži* ‘to come’ has the imperative forms *īža, īžāw*. For these aspects, the negative is formed with the circumfix *ma -…-š* as in *ma-ḥabbīt-š* ‘I did not want’.

Future tense is marked by preposing the particle *bāš ~ biš*. Negation is performed by placing the negative copula before this marker rather than on the verb itself, e.g. *manīš ~ muš biš niržaʿ* ‘I am not going to return’.

The imperfect past is formed as in Standard Arabic, with the perfect of *k-w-n* and the imperfect of the main verb, as in *kān yuxzzurli* ‘he was looking at me’. The future particle could be added between the two verbs to give the sense ‘he was going to look at me’: *kān biš yuxzzurli*.

One of the most distinctive things about the tense and aspect system of Tunis Arabic is the form of the progressive. As in many Arabic varieties, this is expressed in verbs of motion and state by a present participle, e.g. *wīn māši* ‘where are you going?’; *mrawwaḥ* ‘going home’. For other verbs, there are two ways of expressing this aspect, often combined. The first is the use of the present participle *qāʿid*, stripped of its normal meaning ‘sitting’, inflected for number but not normally for gender, e.g. *qāʿid nixdim* ‘I’m working’. It is relatively rare to find *qāʿid* without the other main marker of progressive, which can, however, be used only with transitive verbs. This is the use of *fi* before the direct object in a way that appears prepositional, e.g. *nistanna fīk* ‘I’m waiting for you’, *qāʿdīn yibnīw fi dār* ‘they’re building a house’.

Interestingly, whenever the progressive is used, whatever the marker, negation is marked by use of the negative copula, and the main verb does not take the usual negative circumfix, e.g. *manīš qāʿid nḏ̣ayyaʿ fi waqti* ‘I’m not wasting my time’, *māni naʿmil fi ḥatta šāy* ‘I’m not doing *anything*’.

The negative copula is not the only nonverbal form to have verbal negation. The existential *famma ~ ṯamma* takes *ma-…-š* as do *fibāl-* ‘to know/be aware’ and *ʿand-* ‘to have’ (but not when used prepositionally), e.g. *ma-fibālū-š* ‘he doesn’t know’, *ma-ʿandi ḥatta fṛank* ‘I don’t have a penny’. The latter two present additional evidence of shifting in a verbal direction, with some speakers using forms such as *ma-kunt-š fibāli* and *kunt ʿandi*, where the verb is conjugated in the 1st person, apparently in agreement with the ‘main verb’, instead of the more usual *ma-kān-š fibāli* ‘I didn’t know’ and *kān ʿandi* ‘I had’.

Mood markers include *ṛā-*, *mā-*, and *hā-*, to which a pronominal suffix must be attached. All have some sort of emphatic meaning, different for each one. The marker *ṛā-* is used to draw attention to the hearer that something is evident, e.g. *ṛāni mazrūb* ‘you see, I’m busy’. This marker is not used with the frequency that is heard in Algiers, having a more restricted function. Like the other markers, it can precede any sort of complement. The marker *hā-* draws attention to the presence of the referent, e.g. *hāni naḥki mʿāk* ‘here I am, talking to you’ (likely to be said after the speaker has been blamed for not coming), or *hāh ža* ‘see, he has come’.

2.3.7.3 Verbal negation

Negation of the verb is marked, as in other North African dialects, by the circumfix *ma-…-š*, e.g. *ma-nifham-š* ‘I do not understand’. Where another negative particle is present, the suffix should normally be omitted. The relevant particles are *šāy* ‘nothing’, *ḥadd* ‘no one’, *ʿumr-il-ik* etc. ‘I/you etc. never’, *waḷḷāhi* ‘by God’, and *ḥatta*, which must be followed by another noun and gives the meaning of ‘not a’. For example:

ma-nifham (ḥatta) šāy	‘I don’t understand anything [a thing]’
ma-ža (ḥatta) ḥadd	‘no one [at all] came’
ma-ʿumru yaqḏ̣i min ʿandi	‘he never shops at my place’
waḷḷāhi ma-šuftu	‘I swear I didn’t see him!’
ma-mša lḥatta buqʿa	‘he didn’t go anywhere’

There is a negative copula of the following forms: *manīš* 'I am not', *makš* ~ *mākš* 'you are not', *mahūš* 'he is not', *mahīš* ~ *mahyāš* 'she is not', *manāš* 'we are not', *makumš* 'you [pl.] are not', *mahumš* 'they are not'. It is also used extensively in negation of certain tenses and aspects. On occasion, the unconjugated form *muš* is used instead of these forms. The regular copula is filled either by zero in the imperfect, or by the verbs *k-w-n* 'to be', *ž-y* 'to come', or *b-d-a* 'to begin'.

2.4 *Syntax*

2.4.1 Word order

The question of what is the basic order in a → pro-drop language such as Arabic is a difficult one to answer. A better way of approaching this question is to find out in which circumstances Subject-Verb or Verb-Subject order predominates.

New topics, where definite (because of accessibility or because they function as a contrastive topic), come before the verb, resulting in a Subject-Verb order, which is very common in conversation, e.g. *lixtilāf ykūn bsīṭ* 'the difference is simple', *iṛṛāžil ma-yixtaṛš martu būh wummu yixtāṛūhālu* 'the man does not choose his wife, his parents choose her for him'. In each of these cases, newly activated or contrastive topics precede the main verbs.

However, when a topic has already been established, it is not normally represented lexically but sometimes can be, by way of a reminder, e.g. *u kīf yitkallmu zāda ma-yḥabbūš ržāl kbār* 'and also as they speak, old men do not like…'. In this case the topic is indefinite, which may further encourage Verb–Subject order here, but such an order can also occur with definite topics.

In fact, it seems that it is rather the new or contrastive participant which precedes the verb, rather than the grammatical subject, e.g. *la āna baʿṯitni ummi xāṭr ummi u bāba ils sont divorcés* 'no, I was sent by my mother [lit. my mother sent me]…because my parents are divorced…'; *bu ummi ʿumri ma-ḥkīt mʿāh ma-nʿarfūš* 'as for my maternal grandfather, I never got to talk with him: I never knew him'.

Interrogative sentences can be formed by three different mechanisms, the first by using an interrogative pronoun or adverb, listed above. Yes/no questions can be marked by an initial *yāxxi*, usually in its own intonation group. Alternatively, the expression *walla la* 'or not' can be added at the end of the sentence, e.g. *mšītlu walla la* 'did you go to [see] him or not?'. A second option is to use a rising intonation pattern. Finally, interrogation can be expressed by the suffix *-ši*, sometimes shortened to *-š* attached to the verb or to a verb-like element, e.g. *žīt-ši waḥdik* 'did you come alone?', *ʿandik-ši ṣarf* 'do you have any change?'.

2.4.2 Conditional sentences

Conditional clauses are introduced with (*īḏa*) *kān*, as in *kān tḥabb ḥāža uxṛa, qulli bark* 'if you want something else, just tell me'. The use of the perfect within the clause is also common. Counterfactuals are introduced by *lu* with an imperfect, while the main clause starts with *ṛā-*, e.g. *lu klītha ṛāk lqītha bnīna* 'if you had eaten it, you would have found it to be delicious!'.

3. LEXICON

The main source of vocabulary in Tunis Arabic is from the pool of Arabic itself, but there are many other sources as well (→ Tunisia). Berber loans are evident in words such as *šlāġim* 'moustache', *fakrūn* 'tortoise', and *ʿallūš* 'sheep', among others. This last word, along with *qaṭṭūs* 'cat', may ultimately have its origin in Latin, via Berber. Turkish loans include common words such as *bālak* 'perhaps', *barša* 'very, a lot', and *gawri* 'European'. From Spanish we have *ṣabbāṭ* 'shoe', alongside many words used in the manufacture of the *šāšiya*, a brimless red felt cap that used to be worn by Tunisian men. Italian immigration was a major factor in the 19th and early 20th centuries, and there are many words from Italian, such as *famīlya* 'family', *dākūrdu* 'okay', and *trīnu* 'train', now largely being replaced by the French form *trã*.

It is, however, French which has the most influence in the vocabulary of modern Tunis Arabic, with some words still showing evidence of French phonology, such as the form of the word 'train' above (but the uvular pronunciation of French /r/ is not used in borrowed words). Words often used today, without this being explained by code-switching, include *dossier* 'file', with vowels that cannot be put into the system described above, and with the accent on the final syllable; such words have not been integrated into Tunisian phonology, unlike older loans such as *zufri* 'laborer' (< *les ouvriers*). The longevity of some of these loans is not certain: a loan from Italian *avūkātu* 'lawyer' is often neglected in preference for the Arabic form *muḥāmi*. There is also a lot of code-switching, evidenced by either a fixed expression, a clause, or sentence being used in French: obviously this is common only among highly educated speakers, and is not found among those who know no French.

BIBLIOGRAPHICAL REFERENCES

Baccouche, Taieb. 1972. "Le phonème 'g' dans les parlers arabes citadins de Tunisie". *Revue Tunisienne de Sciences Sociales* 9:30/31.103–137.

Balegh, Hédi. 1993. *Proverbes tunisiens:* Himl al-jamaah rish *tous pour un et un pour tous*. Tunis: Editions La Presse.

Cohen, David. 1964. *Le parler arabe des juifs de Tunis: Textes et documents linguistiques et ethnographiques*. The Hague and Paris: Mouton.

——. 1970. "Les deux parlers arabes de Tunis". David Cohen, *Études de linguistique sémitique et arabe*, 150–171. The Hague and Paris: Mouton.

——. 1975. *Le parler arabe des juifs de Tunis*. II. *Etude linguistique*. The Hague and Paris: Mouton.

Gibson, Michael L. 1998. *Dialect contact in Tunisian Arabic: Sociolinguistic and structural aspects*. Ph.D. diss., University of Reading.

——. 2002. "Dialect levelling in Tunisian Arabic: Towards a new spoken standard". *Language contact and language conflict in Arabic: Variations on a sociolinguistic theme*, ed. Aleya Rouchdy, 24–40. London and New York: RoutledgeCurzon.

Jabeur, Mohamed. 1987. *A sociolinguistic study in Rades: Tunisia*. Ph.D. diss., University of Reading.

Jourdan, J. 1956. *Cours pratique d'arabe dialectal tunisien: Grammaire et vocabulaire, exercices, versions et thèmes*. Tunis: Abela.

Maalej, Zouhair. N.d. *Passives in Modern Standard and Tunisian Arabic*. [Available in cached form via Google search *Maalej passive Tunisian*.]

Marçais, William. 1950. "Les parlers arabes". *Initiation à la Tunisie,* ed. André Basset a.o., 195–219. Paris: Adrien-Maissonneuve.

Quitot, Michel. 2002. *Parlons l'arabe tunisien: Langue et culture*. Paris: L' Harmattan.

St. Exupéry, Antoine de. 1997. *Le petit Prince*. Trans. Hédi Balegh. Tunis: Maison Tunisienne de l'Edition.

Singer, Hans-Rudolf. 1984. *Grammatik der arabischen Mundart der Medina von Tunis*. Berlin: de Gruyter.

Stumme, Hans. 1893. *Tunisische Märchen und Gedichte: Eine Sammlung prosaischer und poetischer Stücke im arabischen Dialekt der Stadt Tunis nebst Einleitung und Übersetzung*. Leipzig: J.C. Hinrichs'sche Buchhandlung.

——. 1896. *Grammatik des tunisischen Arabisch, nebst Glossar*. Leipzig: J.C. Hinrichs'sche Buchhandlung.

Talmoudi, Fathi. 1986. *A morphosemantic study of Romance verbs in the Arabic dialects of Tunis, Susa, and Sfax*. Göteborg: Acta Universitatis Gothoburgensis.

MAIK GIBSON (Nairobi)

Tunisia

1. THE ARABIC LANGUAGE

Tunisia's Constitution, promulgated after independence in 1956, establishes Arabic as the official language of Tunisia. However, what is meant by Arabic is not the spoken language but rather written Arabic. In this entry, the first questions to be dealt with are where this language comes from, and when and under what circumstances it spread in Tunisia to the extent of becoming the official language.

Arabic is a Semitic language which has been spoken for at least three millennia by the tribes of the Arabian Peninsula, notably those in the northern regions. Although it is more related to the → South Semitic languages (high rate of gutturals, affixed verbs, internal plurals, etc.), it also displays common features with the → Northwest Semitic languages (nasalized suffixation of the masculine plural, internal passive, diminutive, etc.). The median geographical position could explain the median structural character, which probably makes it the nearest language to ancient Semitic. Its evolution, from a spoken language in Antiquity and the beginning of the Middle Ages to written and codified Classical Arabic with the advent of Islam at the beginning of the 7th century C.E., then to a modern language since the Renaissance (*Nahḍa*) of the 19th century, has endowed it with vitality, coupled with an exceptional longevity (Cohen 1993; Baccouche 2001b). This evolution has allowed Arabic to change from a local interdialectal koine into a universal language of culture with great influence all over the medieval world (Baccouche 2001a).

The most ancient language known on Tunisian territory was Libyc, from which a number of Berber dialects seem to be derived. Its written form was certainly very restricted. Among the ancient traces (a few hundred inscriptions), some are bilingual (Libyco-Punic or Libyco-Latin). However, the ancient names of the country (Africa) and its population (the Berbers), come from the Romans. Bilingualism was the keynote of the entry of the region into history three millennia ago, with the foundation of the Carthaginian Empire (814–146 B.C.E.) by Phoenicians coming from Tyre (Lebanon). Their language, a Semitic one, even became the official language in the Numidian kingdoms and continued to be used for several centuries after the destruction of Carthage by the Romans. Latin then became the official language of the region, which became multilingual. In fact, to the three languages mentioned above, Greek was added as a fourth, especially after the Byzantine reconquest (533–698 C.E.) of the region. These four languages constitute the linguistic substratum to Arabic, having left behind traces in toponymy, anthroponymy, and the lexicon.

Berber is the most important language, being the original native language before it was gradually replaced by Arabic. According to the *Ethnologue*, Nafusi Berber or Chilha is still spoken by approximately 26,000 people on the island of Jerba, in isolated villages in southern Tunisia, and in the villages of Tamezret and Taoujjout. Spoken Tunisian Arabic still contains words from Berber, such as *suwa* 'countryside' and *fakrūn* 'tortoise'.

The Punic lexical heritage is not always easy to detect because of its relatedness with Arabic, except when it is attested before the Arabic conquest. The Greco-Roman heritage, on the other hand, is easier to identify, for instance in the names of the months from the Julian calendar, e.g. *ayinnār* 'January' < *ianuarius*, and in the loanwords *fallūs* 'chick' < *pullus, qattūs* 'cat' < *cattus, barkūs* 'ram' < *porcus* 'pork' (with a shift in meaning), *furka* 'fork' < *furca*. Words like *ḥarqūs* 'beauty substance made basically of burned copper', *ʿatrūs* 'billy goat', and *maʿdanūs* 'parsley' are derived from Greek. These few hundreds of lexical items inherited from the linguistic substratum represent a small proportion of the lexicon, compared to subsequent borrowings (Baccouche and Skik 1976; Février 1976; Mandouze 1976; Moscati 1976).

On the eve of the Arab conquest, Africa was trilingual and even quadrilingual, if we take into account the strong preference of the learned elite for Greek. The first Arab incursions started in the middle of the 7th century, but due to fierce resistance, it took half a century of wars before Carthage was finally captured, in 698 C.E. The introduction of Arabic took place slowly but gradually, with the help of the Islamization of the country at the expense of the Christian Roman civilization. At first, Latin was used in the Muslim administration, as is attested by the coinage of that time. The Arab army corps included Persians and Arabs from several tribes, including the Tamīm, from whom the Aghlabid dynasty in Kairouan descended.

For about three centuries, Arabic coexisted with Latin (probably in the form of a Romance variety), alongside Berber, in several regions, including the southwest. From the 11th century, Arabic/Berber bilingualism prevailed for about half a millennium, with the balance swinging more and more in favor of Arabic. This was the result of the invasion of new tribes of Arab Bedouin, sent by the Fatimids of Cairo to punish their Berber vassals, the Zirids, to whom they had delegated power when leaving their Tunisian capital, Mahdia (east center of Tunisia), in order to conquer Egypt (W. Marçais 1961a; Ph. Marçais 1975).

The historian Ibn Xaldūn (14th century) describes this phenomenon, focusing on the differences between this newer Arabic and the Old Arabic as described by the first grammarians in the 8th century (*Muqaddima* 554–562). The Arabization of ʾIfrīqiyā, as Tunisia was then called, created a new situation that proved irreversible. Arabic/Berber bilingualism lasted only for a few centuries. New events and new human contacts changed the linguistic situation in the region.

2. Contacts with Turkish and the Romance languages

Toward the end of the Middle Ages, Tunisia found itself at the center of events, conflicts, and exchanges that brought about extremely varied linguistic contacts, growing more intense from one century to the next. Among the most important were the Norman incursions (1148–1160), the crusades of St. Louis (1270), the Reconquista in Spain, the attacks of the Aragonese (1284–1325) and of the Franco-Genovese (1390), the exodus of the Moriscoes, the fall of Constantinople to the Turks in 1453 and Granada to the Spanish in 1492, the incorporation of Tunisia into the Ottoman Empire, and lastly, centuries later, the establishment of the French Protectorate in 1881. All these conflicts and exchanges left their traces on the languages of Tunisia and the region (Abdelwahab 1970; Epalza 1973; Tlili 1974).

2.1 *Arabic in contact with Turkish*

The Muslim empire of the Ottoman Turks waged a long battle against Austro-Spanish forces to achieve dominance in Tunisia in 1574 C.E. Tunisia then became an Ottoman province. Although it does not belong to the same language family as Arabic, Turkish was deeply impregnated with Arabic borrowings and used the Arabic script. The use of Turkish in the administration was more and more limited to relations with the Sublime Porte in Istanbul.

The army of Turkish-speaking janissaries in Tunisia, coming from everywhere in the empire, were extremely diverse ethnically. Their descendants, the Kouloughlis (born from mixed and bilingual marriages), ended up by integrating into the population of the country and Arabizing with time.

The dynasty of the Beys, who came to power in Tunis in 1705, kept only formal links of fealty with the 'Porte' until the French Protectorate in 1881. In fact, Turkish prevailed only in the administration and in the army. Thus, Arabic/Turkish bilingualism remained limited to the elite and the administrative and military staff; the majority of the population communicated in dialectal Arabic. The Ottoman authorities never tried to impose Turkish since they were a Muslim power which respected Arabic.

The influence of the Turkish language is still perceptible at the lexical level, on the one hand in administrative and military terminology, and on the other hand in musical, clothing, and culinary vocabulary. Borrowings from Turkish are found in both literary and dialectal Arabic, e.g. *brik* 'fried crepe', *šawš* 'usher', *fistān* 'dress', *kāhia* 'assistant, associate', *kešk* 'tobacco shop', *qišla* 'barracks', *tersāna* 'arsenal', *tārzi* 'tailor', *yaġurt* 'yogurt'.

Some of these loanwords have become patronymics, such as *Šawš, Kāhia*, along with *Bay* 'bey', *Bāša* 'pasha', *Šelbi* 'graceful', *Zarqūn* 'minium, red lead' (Baccouche 1994; Sebag 1989).

2.2 *Arabic in contact with Romance languages: The lingua franca*

Alongside Arabic and Turkish, some Romance languages were quite widely spoken in some towns, particularly in the ports, in the form of a mixture of French, Italian, and Spanish, together with Arabic and Turkish borrowings. It is this mixture, called → lingua franca, which made communication with Europeans possible, as well as with all those around the Mediterranean, such as sailors, merchants, travelers, corsairs, and migrants (Cifoletti 1989). Accounts of travels, memoirs, and glossaries of that time show the intensity of such contacts, which encouraged the borrowing of words from one language to another in the Mediterranean region. The Tunisians of the coastal towns in particular regularly mixed with Sicilians, Venetians, Genovese, Livornese, Maltese, Provençals, Spanish, etc., along with Turks and citizens of all races from the Ottoman Empire.

It is not always easy to sort out the various linguistic influences that resulted. Tunisian Arabic has integrated hundreds of loanwords going back to this period, e.g. *babūr* 'steamboat', *bastūn* 'piston', *bīntu* 'twenty [francs]', *blaj* 'bolt', *dmān* 'rudder', *dūru* 'five [francs]', *fāška* 'flask', *fišṭa* 'festival', *jurnāṭa* 'daily wages', *jūrni* 'day', *karrīṭa* 'cart', *karrūsa* 'coach', *līra* 'lira', *markānti* 'merchant; rich', *qardāš* 'card', *qarnīṭ* 'octopus', *qmajja* 'shirt', *sbāṭa* 'spatula', *silūn* 'silo', *šqaf* 'skiff', *šqāla* 'a big terracotta plate', etc.

Some loanwords probably entered Tunisian Arabic through the Moriscoes who were driven from Spain. The extraordinary ethnolinguistic diversity of the populations whose paths crossed in Tunisia at that time is also reflected by the great variety of foreign patronymics, which are still current: *Anglīz* 'English', *Arnaūt* (from Arnont in Albania), *Burtugīz* 'Portuguese', *Busnāq* (Bosnian), *Dingizli* (from Karadanize in the Black Sea), *Fransīs* 'French', *Gallāti* (from Galata in Constantinople), *Jinwīz* 'Genoese', *Maltīz* 'Maltese', *Nabulitān* 'Neapolitan', *Rodesli* (from Rhodes), *Sbanyūl* 'Spanish', *Zmerli* (from Smyrna in Turkey), *Blānku* (< Spanish *blanco* 'white'), *Brīma* (< Italian *prima*), *Balasīra* (< Italian *bella sera*), *Senyūr* (< Italian *signore*, Spanish *señor*), *Krīstu* (< *Cristo* 'Christ'), *Lūngu* (< Italian *longo*), *Nīgru* (< Spanish *negro* 'black'), etc.

Even toponyms still show some traces of this mixing of races, e.g. Cap Blanc, Cap Serrat, Porto-Farina, Tunnara, Zambra, Zambretta (Baccouche 1994; Institut Supérieur des Langues 2000).

2.3 *Arabic and Spanish*

The temporary occupation by the Spanish (1535–1574) of some sites on the Tunisian coast (Tunis, Mahdia, and Jerba) had no tangible impact compared to that of migratory movements from Spain to Tunisia following the Reconquista (from 12th to 16th century) and the general deportation of tens of thousands of Muslims and Jews from Spain at the beginning of the 17th century.

Most of these 'Moriscoes' settled in the northeast and in Tunis, giving Spanish names to whole districts: *Murkād* < *mercado, Bīga* < *vega, Bārdu* < *Prado*. Whole towns were planned for them (no fewer than twenty). Patronymics are still reminiscent of their origin: *Blānko, Bonatīru, Burīga, Katalān, Mālqi, Merrīšku, Mnāra, Mnakbi, Nīgru, Qabādu, Sānšu*, etc.

Most of the Moriscoes spoke Aragonese, Castilian, or Catalan. Having already lost their Arabic, some continued using their Spanish dialect until the 18th century. The Moriscoes introduced their science, trades, crafts, and agricultural techniques, as well as their lexicon; a notable example was their manufacture of the *šāšiya* 'red cap'. They also introduced the *hindi* 'prickly pear', the variety of figs called *gūti* 'Gothic', and several varieties of citrus fruit (Epalza 1973; Baccouche 1994).

3. Arabic in contact with Italian and French

The above-mentioned contacts were spread over many centuries. Those with Turkish and Spanish stabilized, whereas contacts with Italian and then French developed in a dramatic way.

3.1 *The impact of Italian*

In the 19th century, the Italian colony became the most important one through the growth of immigration, notably of Sicilians. To the seven thousand Italians resident in 1870, about eleven hundred Livornese Jews were added, most of whom settled in a district in Tunis which is still called Grana. In spite of the establishment of the French Protectorate (1881), they were still a large majority among the Europeans half a century later, in all sectors of the economy except public service. The Italian colony had its own newspaper, theater, and hospital, nine schools in Tunis, four banks, a Chamber of Commerce, and control over four of the five cinemas. The Dante Alighieri Society taught Italian.

This dynamic colony was close to the Tunisian people. Hence, hundreds of Italian words entered Tunisian Arabic. Among Italian loanwords which are still current are *bagno* > *bānu* 'bath', *batteria* > *batrīyya* 'battery', *birra* > *bīrra* 'beer', *blusa* > *blūza* 'blouse', *buono* > *būnu* 'good', *cambiale* > *kimbyāl* 'bill of exchange', *cantina* > *kantīna* 'basement, cellar', *casseruola* > *kasarūna* 'casserole', *cioccolata* > *šuklāṭa* 'chocolate', *crema* > *krīma* 'cream', *cucina* > *kujīna* 'kitchen', *dozzina* > *tuzzīna* 'dozen', *fabbrica* > *fabrīka* 'factory', *falso* > *fālsu* 'false', *farina* > *farīna* 'flour', *fattura* > *fatūra* 'invoice', *maccheroni* > *makarūna* 'macaroni', *macchina* > *makīna* 'machine', *marca* > *mārka* 'brand', *molla* > *mūlla* 'spring, coil', *pacco* > *bāku* 'package', *prova* > *brūfa* 'test, rehearsal', *ricotta* > *rigūta* 'ricotta cheese', *sala* > *sāla* 'hall', *sigaro* > *sigāru* 'cigar', *tonnellata* > *ṭurnāṭa* 'ton', *veste* > *fīsta* 'vest', etc. (Baccouche 1994).

At this stage, it may be said that the borrowings from Italian and the above-mentioned languages into Tunisian Arabic constitute a closed and integrated set. The similarities of Italian borrowings to Spanish ones certainly facilitated their integration.

3.2 *The impact of French*

Through the process of colonization, French soon became the official language of administration and public education. After 1881, apart from traditional scholars, all Tunisians who went to school became bilingual. Borrowings from French entered Arabic by the hundreds, whether literary or dialectal, and this is true of all fields of activity. The status of French, a foreign language that was at first imposed and suffered as an official language, did not change with the end of the protectorate in 1956.

Because of its importance in asserting identity, Arabic was used as a means of resistance to the occupation, and the Arabization of public life constituted a recurrent political claim. The new state of independent Tunisia, however, maintained the status of French as a foreign language to fulfill a strategic choice of maintaining links with the French-speaking world (*Francophonie*). Arabization then became the watchword for the Opposition, first the Arab nationalists, then the Islamic fundamentalists. However, the policy of Arabic-

French bilingualism in its broad outlines has been maintained up to now (Ganiage 1959; Grandguillaume 1983).

4. THE PRESENT-DAY LINGUISTIC SITUATION

4.1 *Bilingualism at the present time*

Arabic-French bilingualism is still perceived today as an impediment to Arabization, which some advocate for political and religious reasons. The political aspects of bilingualism are considered an alienating factor by them, while others see in it an avenue to modernity and development.

This bilingualism is nowadays practiced in the following manner. Arabic may be the only language used at the level of basic education, but for secondary and higher education French remains the language in the so-called exact sciences, in medicine, and in some fields of law and the social sciences (Garmadi 1968; Grandguillaume 1983). This is reflected in the choice of the language for scientific publications (theses, dissertations, articles in specialized reviews). However, Arabic, which began to be used in academic research in the 1970s, now accounts for more than a third of the publications in law and the political sciences and for two-thirds in human and social sciences, that is to say about half the papers in the journals of all these disciplines together (Baccouche 1998).

The progress of Arabization at the university level is echoed in the administration, where it is encouraged and even recommended by the authorities. But in the economic sector, Arabic is still used in a limited way. It is rare, for example, to have checks made out in Arabic, although they are bilingual (Baccouche 2001a). While appearing to be complementary in some cases, French and Arabic are rather in a situation of dynamic competition, with French fairly stable in some fields and in decline in others (Institut Supérieur des Langues 2000). This evolution is also noticeable in daily spoken use, in which social bilingualism is still current. The most notable features of this bilingualism are:

i. The widespread use of French among those who have carried on their studies in French or those who belong to well-off and educated urban circles.
ii. The use on the part of many bilinguals of a hybrid dialect lacking homogeneity, a dialect which employs code-mixing of Arabic statements containing many borrowings from French or French expressions, or even French statements containing Arabic expressions (Jamoussi 1984; Daoud 2001). However, the form observed more and more among the younger generations who master neither French nor Arabic is a spoken dialect in which the use of French words does not necessarily suggest that their direct Arabic counterparts are lacking (lexical or incidental gaps). These loanwords often undergo Arabic inflection. Verbs, in particular, are submitted to the same processes of derivation and conjugation (the amalgamation of an Arabic morphological pattern and a French lexical root, e.g. *patiner* > *ypatīni* (*p-t-n*) 'to skate', *doublage* > *dablaj* (*d-b-l-j*) 'lining', etc.
iii. The practice of integrating French loanwords is a sign of decreasing mastery of French, in spite of its predominance at the university. This development does not go hand in hand with a better command of Standard Arabic, which, in turn, is very much influenced by the Tunisian dialect.

4.2 *Diglossia at the present time*

Tunisian dialectal Arabic presents phonological as well as lexical regional variants which, though important, do not usually hinder mutual comprehension. The *Tunisian Linguistic Atlas*, soon to be published, will map this phenomenon (Baccouche and Mejri 2000). Some of the varieties have practically disappeared, for example the Maltese variety, or the Arabic of Tunisian Jews, who also had an Arabic literature written in Hebrew characters (Cohen 1964–1975).

Like Classical Arabic, Tunisian Standard Arabic makes use of 28 consonantal phonemes and 6 vocalic ones. But some phonemes have lost their ancient features (notably /q/, /ṭ/, /j/, /ḍ/). /ḍ/ is no longer distinguishable from /ḏ̣/. /g/ has been introduced because of its occurrence in Bedouin dialects and loanwords.

In morphology, the main differences are found on the level of syllabic and vocalic structures of nominal as well as the verbal patterns. The syllabic system of dialectal Arabic has become more flexible and consequently more permeable to borrowings.

In syntax, spoken Arabic no longer has a declensional system, due to the loss of the inflectional endings.

Tunisian dialectal Arabic shares with Modern Standard Arabic much of the lexicon, but it differs from it due to its heritage and its various contacts. Standard and spoken Arabic in Tunisia constitute two related but different systems. They tend, however, to complete each other, to get reciprocally enriched thanks to literacy and to the media, and to function as two levels of the same language. This interrelationship favors the emergence of intermediary or mixed registers, which could be perceived as a continuum.

Because it is not systematically codified and does not have a well-established written tradition, Tunisian dialectal Arabic remains essentially an oral language (Marçais 1961b; Baccouche 1994, 1998).

5. Conclusion

Of the ancient languages mentioned above, what remains today in Tunisia are just more or less clearly identifiable traces in diverse forms. Only French competes today with Arabic, notably at the level of higher education. In spite of its privileged status as a foreign language, English does not represent any threat to French, whose importance is decreasing only to the benefit of Arabic.

Modern Standard Arabic usage in Tunisia is evolving under the influence of both French and Tunisian spoken Arabic, and also of some varieties of Arabic conveyed by the media from other Arab states, in particular the Middle East, where the impact of English is more obvious.

Arabization is steadily progressing in spite of the absence of any planning, but the affiliation with the French-speaking world (*Francophonie*) is no longer explicitly assumed nor politically claimed. However, the future of Arabization at the university will probably govern the future of the present bilingualism.

Finally, the growing complementarity and permeability between Modern Standard Arabic and the spoken dialect is a common feature in most Arab countries, interacting mutually, thanks to the media and to social mobility. This process may lead to the birth of a new modern Arabic koine which is dynamic and open to foreign languages, and which integrates Tunisia as well as the other Arab countries into a larger linguistic and cultural area that may justifiably be called the Arab world.

Bibliographical references

Primary source

Ibn Xaldūn, *Muqaddima* = Walī d-Dīn ʿAbd ar-Raḥmān ibn Muḥammad Ibn Xaldūn, *al-Muqaddima*. Cairo, n.d.

Secondary sources

Abdelwaheb, Hassan Hosni. 1970. "Coup d'oeil général sur les apports ethniques étrangers en Tunisie". *Cahiers de Tunisie* 69–70.151–169.

Baccouche, Taïeb. 1994. *L'emprunt en arabe moderne*. Carthage: Beit el Hikma.

—— 1998. "L'arabe dans le monde arabe". *L'information grammaticale*, special issue *Tunisie*, 49–54. Paris: Société pour l'Information Grammaticale.

—— 2001a. "L'arabe, d'une koiné dialectale, à une langue de culture". *Mémoires de la Société Linguistique de Paris* 9.87–93.

—— 2001b. "Tunisie: Langues". *Encyclopédie de l'Islam* X, 657–660. 2nd ed. Leiden: E.J. Brill.

—— and Salah Mejri. 2000. "Langage et altérité: L'expérience de l'*Atlas linguistique de Tunisie*". *Revue Tunisienne des Sciences Sociales*, special issue 120.

—— and Hichem Skik. 1976. "Aperçu sur l'histoire des contacts linguistiques en Tunisie". *Actes du 2ème Congrès International d'Etude des Cultures de la Méditerranée Occidentale*, ed. *Association Internationale d'Etude des Civilisations Méditerranéennes*, 157–195. Algiers: Sned.

Cifoletti, Guido. 1989. *La lingua franca-mediterranea*. Padua: Unipress.

Cohen, David. 1964–1975. *Le parler arabe des juifs de Tunis*. I. *Textes et documents*. 2. *Etude descriptive et historique*. The Hague: Mouton.

—— 1993. "Arabe: Langue arabe". *Encyclopaedia Universalis* II, 707–714. Paris: Encyclopaedia Universalis France.

Daoud, Mohamed. 2001. "The language situation in Tunisia". *Current Issues in Language Planning* 2.11–52.

Epalza, Miguel. 1973. "Moriscos et Andalous en Tunisie au 17ème siècle". *Etudes sur les Moriscos andalous en Tunisie*, ed. Miguel Epalza and Ramon Petit, 150–186. Madrid: Dirección General de Relaciones Culturales.

Février, Paul Albert. 1976. "Colonisation romaine et forme artistique dans les provinces de la Méditerranée occidentale". *Actes du 2ème Congrès International d'Etude des Cultures de la Méditerranée Occidentale*, ed. *Association Internationale d'Etude des Civilisations Méditerranéennes*, 49–102. Algiers: Sned.

Ganiage, Jean. 1959. *Les origines du protectorat français en Tunisie (1861–1881)*. Paris: Presses Universitaires de France.

Garmadi, Salah. 1968. "La situation linguistique actuelle en Tunisie". *Revue Tunisienne des Sciences Sociales* 13.13–24.

Grandguillaume, Gilbert. 1983. *Arabisation et politique linguistique au Maghreb*. Paris: Maisonneuve et Larousse.
Institut Supérieur des Langues. 2000. *Les langues en Tunisie: Etat des lieux et perspectives*. Tunis.
Jamoussi, Souad. 1984. *A case study in multilingual behavior: Code switching among educated speakers in Tunisia*. Ph.D. diss., University of Tunis.
Mandouze, André. 1976. "*L'Afrique chrétienne: Aventure originale ou avatar de la romanisation?*". *Actes du 2ème Congrès International d'Etude des Cultures de la Méditerranée Occidentale*, ed. *Association Internationale d'Etude des Civilisations Méditerranéennes*, 103–117. Algiers: Sned.
Marçais, Philippe. 1975. "Les dialectes occidentaux: ʿArabiyya". *Encyclopédie de l'Islam* I, 597–601. 2nd ed. Leiden: E.J. Brill.
Marçais, William. 1961a. "Comment l'Afrique du Nord a été arabisée?". *Articles et conférences*, 171–192. Paris: A. Maisonneuve.
—— 1961b. "La diglossie arabe". *Articles et conférences*, 83–97. Paris: A. Maisonneuve.
Moscati, Sabatino. 1976. "L'expansion phénico-punique dans la méditerranée occidentale". *Actes du 2ème Congrès International d'Etude des Cultures de la Méditerranée Occidentale*, ed. *Association Internationale d'Etude des Civilisations Méditerranéennes*, 9–33. Algiers: Sned.
Sebag, Paul. 1989. *Tunis au 17ème s.: Une cité barbaresque au temps de la course*. Paris: L'Harmattan.
Tlili, Béchir. 1974. *Les rapports culturels et idéologiques entre l'Orient et l'Occident en Tunisie au 19ème s. (1830–1880)*. Ph.D. diss., University of Tunis.

Taïeb Baccouche (University of Tunis)

Turkey

In Turkey, as in all Muslim countries, Arabic is a language of religion. However, it is also a minority language traditionally spoken by communities along the eastern section of the Mediterranean coast and along the southern border shared with Syria and Iraq. The presence of Arabic in Turkey is reasonably secure as the population is almost entirely Muslim (98.8%; Courbage and Fargues 1997:115), but Turkey is officially a secular state, and the implicit association of Arabic with conservative Islam colors the status of Classical Arabic in Turkey. As a minority language in Turkey, the status of Arabic is largely determined by attitudes and policies relating to all minority languages, → Kurdish being a more visible example. The Arabic-speaking community is also notably diverse in terms of its speakers' dialect, religion, and socioeconomic status. This is one reason why Arabic's status as a minority language is not offset by its status as a religious language.

1. Background

The Turkish Republic emerged from the collapse of the → Ottoman Empire and its subsequent division by the European Allies at the close of the First World War. The Ottoman Empire was notably tolerant of other languages existing alongside the administrative language, Ottoman Turkish, although those languages were not used for official purposes. By contrast, the Turkish Republic, founded in 1923, has pursued an active policy of homogenization through the construction of a single ethnic and linguistic identity.

The principles on which the republic was founded have, from the beginning, had an impact on the status of Arabic, both as a minority language and as a language of religion. There are even parallels in the treatment of Arabic borrowings into Ottoman Turkish. The founding principles (as in the 1931 party program) were republicanism (that there be no monarchy); secularism (that religion be removed from public life and, ironically, that the state control religious institutions); nationalism (that a national identity be created); populism (that the interests of the whole nation come before those of any group or class); statism (that the state be preeminent in the economic field); and revolutionism/reformism (that the nation modernize; Zürcher 1993:189–190).

For the Turkish language (Doğançay-Aktuna 1995:226–227, drawing on König 1987:261), these principles correspondingly translated into:

i. Turkish being established as the new official language;
ii. Arabic and Persian borrowings (grammatical and lexical) being expunged and the Arabo-Persian script being abandoned;
iii. The Turkish language being used to unify the heterogeneous population, along with 'language authentification and purification efforts' (including ii, with replacement borrowings taken from other Turkic languages);
iv. Erosion of the divide between the language of the Ottoman literary elite and that of the masses;
v. Establishment of the Turkish Language Society (Türk Dil Kurumu, TDK) to imple-

ment much of the language policy relating to Turkish; and

vi. Development of the new terms needed by a Westernizing nation.

The principle of secularism led to the banning of Arabic and Persian in primary and secondary schools (Doğançay-Aktuna 1998:26), while the principles of nationalism and populism led to the promotion of a monoethnic Turkish identity at the expense of the republic's diverse minorities.

2. ARABIC AS THE LANGUAGE OF ISLAM

The fortunes of Arabic as the language of religion have been largely determined by the State's treatment of Islam and Islamists, most notably in the sphere of religious education. Atatürk's reforms of the education system were a key component in the secularization of the country (Winter 1984:186). Locally run *medrese*s were closed down and replaced by state-controlled *imam-hatip* schools for the training of prayer leaders and preachers. The curriculum of these schools focused on secular subjects; religious subjects were secondary, and Arabic was not even taught (Winter 1984:188). Consequently, the schools did not flourish and were closed in 1931 (Winter 1984:188). The following year the directorate of religious affairs required that the *ezan* 'call to prayer' be recited in Turkish rather than Arabic (Lewis 1999:46). The Arabic *ezan* was restored in the 1950s (Progler 1998, par. 4). The *imam-hatip* schools were reopened in 1951 with a more appropriate format, including Arabic lessons, and became very successful (Winter 1984:188). In fact, the semiautonomous schools became so successful that in 1997 compulsory education in the secular school system was extended to eight years to curb their influence on young Turkish citizens. The *imam-hatip* schools are the most widespread source of Arabic education. In the academic year 2001/2002, approximately 73,700 students were enrolled in these schools (including distance education; Republic of Turkey, Ministry of National Education 2002). However, the level of achievement attained can be quite limited, sometimes not exceeding the ability to read out prayers or follow along, with comprehension not being seen as crucial. Information on tertiary-level students studying Arabic for religious or other reasons is difficult to find (as Uzunoğlu 2003, notes).

In other domains, Arabic is visible in Turkey. Prayers in Arabic script may decorate a home or a vehicle, but occupants need not necessarily understand them. In some cases, prayers are transliterated into Roman script, particularly *Bismillah/irrahmanirrahim* 'in the name of God/ the most gracious the most merciful'. Similarly, during a religious festival such as Ramadan, the month of fasting, a religious television channel may broadcast the ninety-nine names of God in Roman script. As comprehension is generally limited, sermons in the mosque are given in Turkish, while standard prayers are in Arabic. To most Turkish citizens, Arabic and Ottoman Turkish are indistinguishable. Calligraphy from both languages may be copied, particularly to decorate souvenirs for the tourist market.

3. ARABIC AS A MINORITY LANGUAGE

The Turkish language was a key component in unifying the heterogenous population of the republic. This was facilitated by the removal of the two most notable non-Muslim minorities of the Ottoman Empire, the Greeks and the Armenians, by the end of 1923 (Zürcher 1993:189–190; Pope and Pope 1997:350).

For those linguistic minorities that remained, the cultural climate in the republic was inhospitable. Just as Arabic and Persian were removed from the curriculum, Kurdish schools were also closed down (Blau and Suleiman 1996:158). The treatment of linguistic minorities in the Ottoman Empire and the Turkish Republic can be gauged by comparing the Treaty of Sèvres (1920, between the Ottoman Empire and the European Allies) and the Treaty of Lausanne (1923, between the Turkish Republic and the Allies). The Treaty of Sèvres assured "racial, religious or linguistic minorities... [of] the right to use their own language" as well as the right to set up institutions including schools (Art. 147). However, with the founding of the republic, the Treaty of Sèvres was superseded by the Treaty of Lausanne (1923). Although the provisions regarding language were taken largely verbatim from the earlier treaty, the above rights were provided only for non-Muslim minorities. Thus, the majority of Arabic speakers, being officially Muslim (see below), were not afforded protection of their linguistic rights, whereas the small Christian and Jewish minority of Arabic speakers ostensibly were. Greek, Armenian, and Judeo-Spanish schools in Turkey remain under these provisions, but there are no comparable Arabic schools.

In 1934, Turkish surnames became compulsory. While this reform was primarily designed to Europeanize the Turks, it also served to 'Turkify' the linguistic minorities. It was often done unsympathetically, as some families were assigned Turkish surnames with no regard for existing Arabic surnames (Smith/Kocamahhul 2004). According to anecdote, Arabic was banned in Hatay by the local judiciary. It was not unheard of for authorities, such as teachers, to assault someone who disregarded the ban, even if the individual was unable to speak another language.

Official intolerance of minority languages peaked following the 1980 coup. In 1982, the military government revised the Constitution and, in its irrevocable general principles, enshrined the notion of the monolingual Turkish state (Art. 3: "The Turkish State, with its territory and nation, is an indivisible entity. Its language is Turkish"). Furthermore, the new Constitution prohibited languages other than Turkish being taught as a mother tongue (Art. 42.9). Thus, Arabic could be taught as a foreign language but not as a mother tongue, not even to those for whom it was the latter.

This prohibition was reinforced a year later when legislation deemed Turkish to be the mother tongue of all Turkish citizens (Rumpf 1989:82). The use of languages other than Turkish was banned in a range of situations, including private conversations on a public street (Rumpf 1989:82). The law was repealed in 1991.

More recently, the outlook for minority languages has improved as Turkey tries to meet the 'Copenhagen criteria' for entry into the European Union. Numerous reforms have been introduced, particularly ones aimed at permitting broadcasting and education in Kurdish. As the climate for Kurdish language maintenance improves, smaller linguistic groups are likely to be more inclined to attempt language maintenance efforts for themselves (Smith/Kocamahhul 2001:47).

In Turkey, more regard is given to religious divisions than to ethnolinguistic ones, and this is also reflected in official statistics. Thus, it is difficult to know the number of Arabic speakers. The last census to include information on language use was in 1965 (Metz 1995: Linguistic and ethnic groups, par. 1), at which time mother tongue speakers of Arabic numbered 1.2 percent of the population (approximately 365,000; Dewdney 1971:88). Arabic was the third-ranked mother tongue after Turkish (90.1%) and Kurdish (7.1%; Dewdney 1971:88). The greatest concentration of speakers was in the province of Hatay, where, according to Dewdney (1971:88), they numbered 36 percent of the population. Dündar (1999), on the other hand, using the same census results, calculates Arabic speakers as only 29 percent of Hatay's Arabic speakers. However, Dündar lists Arabic as a Muslim minority language, which may account for the discrepancy (see below). The rest were in the then provinces of Adana, Urfa, Mardin, and Siirt (Dewdney 1971:89 for a map). Dündar (1999) calculates the total number of Arabic speakers in the 1965 census (mother tongue and second language) as 1.7 percent of the republic's population (mother tongue 365,340; second language 169,724).

More recent figures can only be estimates. Grimes estimates 400,000 speakers of Northern Mesopotamian Arabic in the Mardin and Siirt provinces (predominantly Kurdish areas) and 500,000 speakers of North Levantine Arabic (Grimes 1996:792, 776). Procházka (2003) additionally estimates 150,000 speakers of Bedouin dialect in the Urfa region.

Yağmur (2001:414) amalgamates Andrews' (1989) figures for Arabs as an ethnic group (569,058) as a way of assessing the number of Arabic speakers. However, it is inadvisable to equate the two: the term 'Arab' can be used pejoratively in Turkey (Arıkan a.o., n.d.), even by one Arabic speaker to another. Thus, some Arabic speakers will deny being ethnically Arab. This, in turn, throws some doubt on reported language use because, as with Kurdish (Hassanpour 1992:135), reports of language use may vary according to the degree of political pressure at the time of survey. Additionally, speakers of languages other than Turkish may claim Turkish as their first language and their other language as their second language (Andrews 1989:53). Interviewees who claim Arabic as a second language (approximately 360,000 in the 1965 census) may be Sunni Arab (as Andrews 1989:148 assumes); alternatively, they may include Turks who have studied Classical Arabic or Arabic as a foreign language. In the southeast, this may also include Kurds (Dündar 1999) and Süryanis (Neo-Aramaic-speaking Christians), who speak local Arabic dialects in addition to their own minority languages (Dündar 1999:216 reports 4,283 Kurdish/Arabic speakers).

Importantly, the population of Arabic speakers in Turkey is heterogeneous. There are three broad dialect groups: North Mesopotamian (*Qəltu*), North Levantine (Syro-Lebanese; Grimes 1996), and Bedouin (Andrews 1989:211; Arnold 1995; Procházka 1999:116). Furthermore, the speakers come from four different religious groups: Sunni Muslim, Alevi (Alawite) Muslim, Christian (Greek Orthodox and Catholic), and Jewish. One should note that some Alevi reject the label 'Muslim'; Arabic-speaking Alevi are also known in the literature as Nusairis. In the İçel, Adana, and Hatay provinces, the majority of the Arabic-speaking population are Alevi (Andrews 1989:214, 209; Procházka 1999:115), while in the southeast, Sunni Muslims are in the majority (Andrews 1989:214, 209). The three broad dialect divisions also have finer divisions according to geographical location and religion (Arnold 2002; Smith/Kocamahhul 2004).

Interestingly, Arabic is the religious language not only for the two Muslim communities but also for Arab Orthodox Christians. Andrews (1989:211) also mentions a Uniate Catholic liturgy in Arabic for Hatay and elsewhere, but this is no longer evident in the two main cities (Antioch and İşkenderun). As in Turkey generally, endogamy is the norm. Intermarriage with members of different ethnicity but the same religion is considered preferable to intermarriage with Arabic speakers of a different religion. Notwithstanding this, intermarriage between Christians and Alevis is reasonably common in Hatay.

There is also a wide range in socioeconomic background. Across Turkey, many Arabic speakers live a peasant lifestyle in rural villages, but other occupations are also associated with certain groups. In Hatay's urban centers, for example, members of the Christian community are strongly represented in the gold trade, while the Jewish community is associated with clothing retail (Arnold 2002). In Cilicia, while most are farmers, Arabic speakers in towns may also be unskilled workers, craftsmen (Procházka 1999:117), fishermen, or clerks (Andrews 1989:154). Others are seasonal workers from Hatay (Andrews 1989:152). With increasing urbanization, still others are university graduates.

Contact between Arabic speakers from different provinces is limited. The sedentary and Bedouin dialects are not mutually intelligible, even in areas such as Hatay where members of the two dialect groups come into contact. Arnold (2002) notes that Bedouin may have command of a sedentary dialect, although the reverse is unheard of. If an Arabic speaker cannot infer that a stranger speaks Arabic, Turkish is used by default (Procházka 1999:120). Because Arabic speakers may deny being Arab, even when asked explicitly, this further limits potential contact with Arabic speakers outside one's immediate area.

Almost all Arabic speakers are illiterate in Arabic, unless they have learned Modern Standard Arabic or Classical Arabic for religious purposes. Usually, this only applies to men in leadership positions. Some learn Modern Standard Arabic in order to study in an Arabic-speaking country. This is mandatory for Arab Orthodox men training to become priests, as such training is unavailable in Turkey.

4. Arabic and language contact in Turkey

Research on language contact between Arabic and other languages in Turkey is scant with the exception of the effect of Arabic in the → Ottoman Empire. Yet, the Turkish Language Reform (literally the 'language revolution', *dil devrimi*) significantly reduced the influence of Arabic on Turkish without eliminating it entirely.

There is little research on language contact in Turkey between Arabic and languages other than Turkish. Arnold (1995:7; 2000:348, 351) alludes to contact with Aramaic. In Kurdish-speaking areas, Arabic comes third in status after Turkish and Kurdish (Lahdo 2003), and there are some instances where Kurdish words have been borrowed into Arabic dialects (Procházka 2003). Similarly, one dialect may influence another. In Hatay, for example, the dialects of smaller groups (for example the Bedouin, village, and Jewish dialects) show evidence of contact with the larger dialects (notably the city Alevi dialect) in the pronunciation of certain phonemes (e.g. /q/ pronounced as postvelar /k/ by villagers coming to the city; Arnold 2002).

The influence of Turkish on Arabic (→ Turkish loanwords) is evident in contact phenomena such as borrowing (Lahdo 2003; Procházka 2002), code-switching (Arnold 2002; Procházka 2002; Smith/Kocamahhul 2004), and language shift (Lahdo 2003; Procházka 1999; Smith/Kocamahhul 2004).

Borrowing from Turkish has occurred on the phonological, lexical, and syntactic levels. Phonemes such as /j/ (Smith/Kocamahhul 2004), /č/, /g/, and /p/ (Nevo 1999:67; Procházka 2002) were borrowed from Turkish. Similarly, word-final consonants are devoiced (Lahdo 2003; Procházka 2002). Lexical borrowing of Turkish words into Arabic has continued from Ottoman times; 'recent' borrowings relate to technology, administration (Arnold 2000:359), everyday items (food, clothes, etc.), and medicine (Procházka 2002). Additionally, some Turkish borrowings from Arabic were borrowed back into Arabic dialects with a meaning modified by the Turkish meaning (Procházka 2002; Lahdo 2003). For example, *misafir* in Turkish means 'guest', and this is the meaning of *msāfir* in Cilician Arabic rather than 'traveler' as in Standard Arabic (Procházka 2000). Calquing (word-by-word translation) also occurs and may include lexical borrowing, for example *telefon açmak* 'telephone open [infinitive]' = 'to telephone' > *fataḥ telefōn* (Procházka 2002). There is also evidence of syntactic influence in word order and in the use of pronouns for emphasis (Procházka 2002).

Code-switching between Turkish and Arabic also occurs. Both individual words and more complex words and phrases from one language may be used in a discourse largely in the other language (Arnold 2002; Procházka 2002; Smith/Kocamahhul 2004). For example: *bi-l-ʿağīn ma-ḥiṭṭ-illu* **pak maya**, *ma-ḥiṭṭ-illu* **eritiyorum** *bi-l-qadaḥ*... "*into the dough, I put* **Pak yeast** *into it, I put it into it*, **I dissolve** *it in the glass*..." (Procházka 2002; Arabic in italic, Turkish in bold). Conversely, Arabic words and phrases may also occur in a discourse that is largely in Turkish (Smith/Kocamahhul 2004), for example: **benim bir. aa bir adam vardı.** *ismo*. H— Y—. **boyu iki metre. M—'tan daha böyle uzun. bir ayakları var şu kadar 'I had a aa- a man** [i.e. a friend] . *his name*. **H—Y—. his size two metres. taller than M—like this. he has feet this big'.**

5. Language shift

While Arabic coexisted with Ottoman Turkish for centuries, the Republican policies outlined above have contributed to a decline in the amount of Arabic spoken in certain domains and a proportional decline in the numbers speaking it (Lahdo 2003; Procházka 2002; Smith/Kocamahhul 2004). However, not only an explicit policy on minority languages has led to Arabic's decline. The political structure and administration of Turkey is strongly centralized, hence, there is always a Turkish presence in Arabic-speaking areas. Additionally, particularly in Hatay, the low cost of living has led to an influx of monolingual Turkish students to the local university, while local students leave Hatay to study at more prestigious universities. Furthermore, despite its proximity to Syria, Arabic is not available as a foreign language at the local university, therefore, students leave Hatay to study Arabic language and literature elsewhere. Emigration for educational purposes is common to all Arabic-speaking areas. Similarly, for young men, compulsory military service necessitates the use of Turkish. Also, many internal refugees migrated to the main cities due to the conflict in southeastern Turkey. This furthers the general urbanization of the Arabic-speaking population, thus depleting the population in Arabic-speaking areas.

State policies (and their implementation) not only affect the transmission of Arabic to the younger generations; they also reflect and reinforce negative attitudes about minorities, and Arabs in particular. In the 1940s, a popular schoolboys' rhyme exhorted the speakers of Arabic, Persian, and French to return to their respective lands (Lewis 1999:24–25).

Due to such pressures against minority languages, younger generations of the Arabic-speaking communities, mirroring a more general trend, are increasingly using Turkish as their mother tongue (Lahdo 2003; Procházka

2002; Smith/Kocamahhul 2004). There is also evidence of 'leveling', that is shift from one dialect to another, supposedly more prestigious, variety (Arnold 2000:365). The extent of the shift from Arabic to Turkish is so great, however, that the effects of leveling are minor by comparison. In other cases, communities have moved out of Turkey altogether, to Syria (Lahdo 2000:61) and Israel (Nevo 1999:66).

Broadly, the language shift can be seen in the domains in which Arabic is not used or is no longer used. Arabic has always played a minimal role in education, government, and the media. Now, it is used less in the domains of business and friendship, and Turkish is even being used in the family and neighborhood domains (Procházka 1999:120). In the Orthodox Christian community in Antioch, Turkish is even encroaching into the religious domain: separate services in Turkish are held for the congregation's youth.

It is clear to the external observer that language shift is underway. This fact is also generally evident to community members. When members of the Arabic-speaking community in Antioch were explicitly asked whether Arabic was declining, the vast majority (84%) said it was (Smith/Kocamahhul 2004). When asked which language they used with members of their family, only 5 percent claimed to use Arabic with all their relatives irrespective of the generation (Smith/Kocamahhul 2004). When asked why Arabic is declining, however, the lack of Arabic in schools was the only external factor widely blamed for the decline in Arabic (50% of respondents). In addition to this, many respondents claimed that older community members do not teach the language (47%), while others said younger community members do not want to learn it (29%; Smith/Kocamahhul 2004). In fact, these three causes are interconnected. Parents report being told by teachers to speak Turkish with their children, as speaking Arabic is assumed to adversely affect acquisition of Turkish. Likewise, children discover that Arabic is not acceptable at school and then refuse to speak it at home. While it is theoretically possible for children to learn Arabic in an *imam-hatip* school, this is only appropriate for Sunni students, as the theology taught is incompatible with Alevi, Christian, or Jewish belief.

Language shift is, therefore, a crucial issue for Arabic in Turkey: although research has documented some of the dialects, there is a grave risk that many of the Arabic dialects in Turkey will disappear without such documentation.

Bibliographical references

Andrews, Peter Alford. 1989. *Ethnic groups in the Republic of Turkey*. Wiesbaden: L. Reichert.

Arnold, Werner. 1995. "Arabian dialects in the Turkish province of Hatay". *Proceedings of the 2nd International Conference of l'Association Internationale pour la Dialectologie Arabe held at Trinity Hall in the University of Cambridge 10–14 September 1995*, ed. Joe Cremona, Clive Holes, and Geoffrey Khan, 1–10. Cambridge: University of Cambridge.

——. 2000. "The Arabic dialects in the Turkish province of Hatay and the Aramaic dialects in the Syrian mountains of Qalamūn: Two minority languages compared". *Arabic as a minority language*, ed. Jonathan Owens, 347–370. Berlin: Mouton de Gruyter.

——. 2002. "Code switching and code mixing in the Arabic dialects of Antioch". *Aspects of the dialects of Arabic today: Proceedings of the Fourth Conference of the International Arabic Dialectology Association (AIDA), Marrakesh April 1–4.2000*, ed. Abderrahim Youssi, Fouzia Benjelloun, Mohamed Dahbi, and Zakia Iraqui-Sinaceur, 163–168. Rabat: Amapatril.

Arıkan, Necmettin, Gülderen Yenal, and Gülsevin Taşpınar. N.d. *The golden dictionary: Altın sözlük, ingilizce/Türkçe, Türkçe/ingilizce* [English/Turkish, Turkish/English]. Istanbul: Altın Kitaplar Yayınevi.

Blau, Joyce and Yasir Suleiman. 1996. "Language and ethnic identity in Kurdistan: An historical overview". *Language and identity in the Middle East and North Africa*, ed. Yasir Suleiman, 153–164. London: Curzon Press.

Courbage, Youssef and Philippe Fargues. 1997. *Christians and Jews under Islam*. London: I.B. Tauris.

Dewdney, John C. 1971. *Turkey*. London: Chatto and Windus.

Doğançay-Aktuna, Seran. 1995. "An evaluation of the Turkish Language Reform after 60 years". *Language Problems and Language Planning* 19.221–49.

——. 1998. "The spread of English in Turkey and its current sociolinguistic profile". *Journal of Multilingual and Multicultural Development* 19.24–39.

Dündar, F. 1999. *Türkiye nüfus sayımlarında azınlıklar* [Minorities in Turkish censuses]. Istanbul: Doz Yayıncılık.

Grimes, Barbara F. (ed.) 1996. *Ethnologue: Languages of the world*. Dallas, Tex.: Summer Institute of Linguistics.

Hassanpour, Amir. 1992. *Nationalism and language in Kurdistan, 1918–1985*. San Francisco: Mellen Research University Press.

König, Wolf Dietrich. 1987. "On some sociolinguistic aspects of language reform in Turkey". *Studies on* *modern Turkish: Proceedings of the Third Conference on Turkish Linguistics*, ed. Ludo Verhoeven, 259–270. Tilburg: Tilburg University Press.

Lahdo, Ablahad. 2000. "An Arabic dialect in Qamashli, II". *Orientalia Suecana* 49.61–69.

——. 2003. *The Arabic dialect of Tillo in the region of Siirt: (South-Eastern Turkey).* Ph.D. diss., Uppsala University.

Lewis, Geoffrey. 1999. *The Turkish language reform: A catastrophic success.* Oxford: Oxford University Press.

Metz, Helen Chapin. 1995. *Turkey: A country study.* Washington, D.C.: Federal Research Division, Library of Congress [http://rs6.loc.gov/frd/cs/trtoc.html (accessed 9 January 2002, last updated January 1995)].

Nevo, Moshe. 1999. "Notes on the Judaeo-Arabic dialect of Siverek". *Zeitschrift für Arabische Linguistik* 36.66–84.

Pope, Nicole and Hugh Pope. 1997. *Turkey unveiled: Atatürk and after.* London: J. Murray.

Procházka, Stephan. 1999. "Language contact to language death: The example of the Arabic spoken in Cilicia (southern Turkey)". *Orientalia Suecana* 48.115–25.

——. 2002. "Contact phenomena, code-copying, and code-switching in the Arabic dialects of Adana and Mersin (southern Turkey)". *Aspects of the dialects of Arabic today: Proceedings of the Fourth Conference of the International Arabic Dialectology Association (AIDA), Marrakesh April 1–4.2000*, ed. Abderrahim Youssi, Fouzia Benjelloun, Mohamed Dahbi, and Zakia Iraqui-Sinaceur, 133–139. Rabat: Amapatril.

——. 2003. "The Bedouin Arabic dialects of Urfa". *Association Internationale de Dialectologie Arabe 5th Conference Proceedings, Cádiz, September 2002*, ed. Ignacio Ferrando and Juan José Sánchez Sandoval, 75–88. Cadiz: Universidad de Cádiz.

Progler, Yusuf. 1998. "In Muslim Turkey, battle looms on the education front". *Muslimedia* September 16–30. [www.muslimedia.com/archives/special98/eduturk.htm, accessed 28 January 2004.]

Republic of Turkey. 1982. *Türkiye cumhuriyeti anayasası* [Constitution of the Turkish Republic] 1982–1995. Istanbul: Kocaoluk Yayın Evi. [In English at www.mfa.gov.tr/grupc/ca/cag/I142.htm, last accessed 19 February 2004.]

——. Ministry of Education. 2002. www.meb.gov.tr/stats/apk2002ing/apage130–140.htm; www.meb.gov.tr/stats/apk2002ing/apage166–172.ht (accessed 9 November 2004).

Rumpf, Christian. 1989. "Das Sprachenverbot in der Türkei unter besonderer Berücksichtigung ihrer völkerrechtlichen Verpflichtungen". *Orient* 30.413–27. [Also: "The Turkish law prohibiting languages other than Turkish", *Documentation of the International Conference on Human Rights in Kurdistan, 14–16 April 1989*, 68–89. Bremen: Hochschule.]

Smith/Kocamahhul, Joan. 2001. "In the shadow of Kurdish: The silence of other ethnolinguistic minorities in Turkey". *Middle East Report* 31.45–47.

——. 2004. *Language choice, code-switching and language shift in Antakya, Turkey.* Ph.D. diss., University of Canterbury.

Uzunoğlu, M. Vecih. 2003. "Türkiye'de Arap dili alanında yapılan yüksek lisans ve doktora tezleri" [M.A. and Ph.D. theses about the Arabic language in Turkey]. *Nüsha Şarkiyat Araştırmaları Dergisi* III, 8. [Also available as http://vecih.tripod.com/Turkiyedekitezler.htm, accessed 9 November 2004.]

Winter, Michael. 1984. "The modernization of education in Kemalist Turkey". *Atatürk and the modernization of Turkey*, ed. Jacob M. Landau, 183–194. Boulder, Colo.: Westview Press.

Yağmur, Kutlay. 2001. "Languages in Turkey". *The other languages of Europe*, ed. Durk Gorter, 407–428. Clevedon: Multilingual Matters.

Zürcher, Eric J. 1993. *Turkey: A modern history.* London: I.B. Tauris.

JOAN SMITH KOCAMAHHUL
(University of Canterbury)

Turkish

1. INTRODUCTION

Since the beginning of the Islamization of the Turks more than one millennium ago, Arabic loanwords have formed an important layer in the lexicon of the Turkish language. Only a certain part of the Arabic layer in the Turkish lexicon is the result of direct language contact; more important is the Arabic component which entered the Turkish lexical stock through Persian. For many centuries, nearly every Arabic lexeme, with the exception of finite verbs, could be used in a Turkish text as a learned or poetic expression. In the last decades of the → Ottoman Empire, a vast number of pseudo-Arabic words were created by Turkish scientists to substitute for European terminology. Together with Arabic lexical elements, several features of Arabic morphology and – to a lesser extent – syntax made their entrance into the Turkish language. Even the elimination of thousands of Arabic loanwords during the Turkish Language Reform in the course of the 20th century left traces, such as loan-translations and imitations, in the Turkish lexical stock.

Because Arabic loanwords are easily discernible in the Turkish lexicon, they have been the topic of several serious studies for more than one hundred years (Tietze 1990:113–114). Bittner (1900) is still impressive and full of interesting prelinguistic thoughts, while some other contributions are repetitive to a certain degree. Many new insights and data result from the outstanding research of Andreas Tietze (1914–2004). Few writers from the Turkish Republic have made any contribution to this field since, in the context of the predominant official linguistic nationalism, Arabic loans have been regarded as suspicious elements of the past. Despite interesting results, this field of research is dominated by traditional etymological and philological methods and still lacks significant contributions with a linguistic approach.

Turkish differs typologically from Arabic in certain aspects. While Arabic is a root-inflecting language with gender, dual forms, and dozens of irregular plural patterns, Turkish is agglutinative, has no gender, and only one regular pattern of plural formation without dual. Whereas Arabic has three qualities of vowel phonemes and a length distinction, Turkish possesses eight short vowel phonemes and exhibits vowel harmony. The system of Arabic consonants displays several features (emphatics and pharyngeals) that have no counterpart in the Turkish system. Other Turkish typological features, like a left-branching clause and sentence structure (in contrast to the right-branching structure of Arabic), are less important in this context.

2. OTTOMAN TURKISH

Since Turkic-speaking tribes were Islamized before the conquest of Anatolia in the late 11th century, it does not come as a surprise that Arabic loanwords can also be found in early Turkic literary sources from Central Asia. The first Turkish texts from Anatolia clearly predate the Ottoman principality which emerged at the beginning of the 14th century in western Anatolia (Kerslake 1998:179). While Persian was the language of the Seljuk court in Konya, Turkish was established more and more as an official language in the small independent Anatolian principalities. In this context of Persian cultural dominance, a great many Arabic loanwords were introduced together with Persian loans into the still unsettled variants of Anatolian Turkish (Prokosch 1983:342). Ottoman Turkish in its classical form emerged in the late 15th century. Since that time, thousands of learned Arabic expressions were established in the Turkish language, through, for example, the *madrasa* system of higher education. While the rate of Persian loans remains on the same level, we see an increase of Arabic words, especially in written formal discourse (Römer 1981:101; Kirchner 1996:153–154). The dominance of Arabo-Persian terminology, the increasing replacement of Turkish words by stylish words from the more prestigious Islamic languages, especially Arabic, and the adoption of several aspects of Arabic nominal inflection, together with some Persian techniques of nominal syntax, led to a highly artificial jargon in written (and, to a certain degree, in spoken) discourse of the educated classes. In the course of the Europeanization of Ottoman society in the last decades of the 19th century, the resulting diglossic situation was criticized by intellectuals. Yet, even in the last years of the Ottoman Empire, Turkish equivalents for new European (scientific) concepts were formed by Turkish scholars on the core of the Arabic lexical fund (Tietze 1958:257). Through Ottoman Turkish, Arabic loans entered several languages of the Balkans, e.g. Albanian and Bulgarian (→ Slavonic languages). Despite the fact that the Arabic heartlands were under Ottoman rule for centuries, only a few → Turkish loanwords made their way to Arabic (Procházka 2005).

The principle of historical graphic representation of loanwords from the prestigious Islamic languages in Ottoman texts, some peculiarities of the Arabic script, and the fact that vocalization was rarely used complicate research on the phonological adaptation of Arabic loanwords in different layers of language history. Important sources for a phonetic reconstruction are the so-called transcription texts, documents in Latin script mostly intended for the practical use of foreigners in the Ottoman Empire, which may give us an impression of the real pronunciation in Ottoman times (cf. Hazai 1973). Arabic *muʿāmala* 'treatment', for example, is transcribed in various texts as *muamele, mamele, müamele, muāmele*, etc. (Stachowski 1975–1986:II, 116). Over the centuries, a codified system of pronunciation was developed for learned Arabic loanwords. In this system, still in use in the modern language, the Turkish distinction of front and back vowels was employed to indicate several phonemic distinctions of Arabic consonants (Tietze 1992:353).

Since a great deal of the Arabic loans in Ottoman Turkish were pan-Islamic internationalisms and learned expressions, restricted to formal written discourse, it is not surprising that semantic change occurs less often than in other cases of language contact. Widespread semantic change is attested in early Arabic loanwords via Persian intermediation. Arabic *ixtiyār* 'choice, selection' has gained in Ottoman Turkish the more frequent meaning of 'old man, chief', while the primary meaning is restricted to some contexts. Arabic loanwords have entered basic concepts of the language, like pronouns and designations of colors. Ottoman Turkish *beyāż* 'white' < Arabic *bayāḍ* covers all concrete meanings, while the original Turkish lexeme *ak* 'white' was more and more reduced to abstract meanings and place-names. In the 19th century, calques of new European concepts were created on the core of Arabic material (Röhrborn 2003:26–29): *taxtelbahir* 'submarine' (Lewis 1999:109; cf. Standard Arabic *ġawwāṣa* 'id.').

The root structure of Arabic loans and the semantic interrelation between derivations from one root was to a certain degree discernible for educated Turks not only because of their adherence to the rules of Arabic orthography but also because of their sheer number, e.g. *suʾāl* 'question', *mesʾele* 'problem', *mesʾūl* 'responsible' (less evident in modern Turkish: *sual*, *mesele*, *mesul*). The same goes for patterns having a certain function, e.g. *faʿʿāl*: *jellād* 'hangman', *qassāb* 'butcher', *ḥammāl* 'carrier' (this relationship is also less obvious in the modern language: *cellât, kasap, hamal*; Tietze 2002:127). The distinction between noun and adjective, not deeply rooted in the structure of Turkish, was strengthened by the suffixlike Arabic *nisba* in hundreds of loans: *dīnī* 'religious', *medenī* 'civilized'. Arabic gender marking was partly discernible for speakers of Ottoman Turkish (*muʿallim* 'teacher' vs. *muʿallime* 'woman teacher'). In *iżāfet* constructions following the Persian model, even the Arabic rules of gender agreement were adhered to more strictly than in some Persian texts: *devlet-i ʿālīye-yi ʿosmānīye* 'the sublime Ottoman government' (Bittner 1900:46–49). Some Arabic feminine loanwords in Ottoman Turkish show the ending *-at/-et* (early loans via Persian, dissyllabic words, like *jennet* 'paradise', several abstract nouns), while others have *-a/-e* (*qalʿe* 'castle'). In the refined styles of written Ottoman Turkish Arabic, broken plurals are preferred to singular forms, with the Turkish suffix *-lar/-ler*: *vüzerāʾ* instead of *vezīrler* 'ministers'. Sometimes even Persian loans are pluralized analogous to Arabic plural patterns: *šehir* < Persian 'city', pl. *šuhūr*. The use of the regular plural of Arabic masculine nouns is restricted to the designation of public groups: *ḥavāriyūn* 'apostles'; regular feminine plurals are widespread, sometimes in singular meaning: *edebīyāt* 'literature'. Double marking with Arabic and Turkish plurals is a phenomenon that dates back to Ottoman times: *tüjjār* (Arabic pl. of *tājir*) 'merchant' > *tüjjārlar* 'merchants' (Battersby 1966:110). Dual nouns or noun phrases are for the most part lexicalized, showing agreement even in number: *ḥaremeyn-i muḥteremeyn* 'the sacred territories of Mekka and Medine' (Buğday 1999:72). Some (compound) Arabic morphemes, like *-iyya*, were extensively used in Ottoman word formation on Arabic roots as quasi-suffixes: *muvaffaqıyet* 'success' < *muvaffaq* 'successful' (Toparlı 1985:170). Arabic case endings are sometimes preserved, for instance in coined expressions (*müšārün ʾileyh* 'the aforementioned') and in adverbial forms (*nisbeten* 'relatively'). Verbal concepts are derived from Arabic nominal forms using the typical agglutinative techniques of Turkish (*šübhe* 'suspicion' > *šübhelenmek* 'to suspect'), more often in phraseological constructions with auxiliary verbs: *muqābele* 'confronting' > *muqābele et-* 'to confront' (cf. Persian *muqābala kardan*). Finite verbs are very seldom borrowed from Arabic; a considerable number of them can be found in nominalized constructions with the verbal negation *lā* (Bittner 1900:73): *lā-yenkesir* 'unbreakable'.

The Arabic influence on Ottoman Turkish syntax is much more limited than the Persian influence. Even in an Ottoman Turkish interlinear translation of the *Qurʾān*, only a few syntactic structures show Arabic interference (Kirchner 2005). Several postpositions of Arabic origin were integrated into the Turkish system and are used in several constructions involving adverbial clauses. *Evvel* 'before' (< *ʾawwal*) and others are still used together with Turkish converbs in order to express a temporal relationship: *gelmedin evvel*... 'before (s)he came...'. The conjunction *ve* (< *wa*) and others were firmly established on the phrasal and clausal level of spoken and written Ottoman Turkish (Ivanov 1973:51–52). In elaborate style, Arabic verbal nouns behaved like verbs in taking complements as if an auxiliary had been present: *milleti* [acc.] *iḥyā ičün* 'in order to revive the state' (Kerslake 1998:201).

The fact that Persian and Arabic culture, especially literature with all its genres, was the overall model for the Ottoman Turks facilitated the acceptance of loanwords from these languages. Ottoman poetry was based on Arabic and Persian meters that did not fit into Turkish syllable structure and the system of eight short vowels. Thus, many Arabic loanwords were used in Ottoman poetry to fill gaps. When, for example, the sequence short syllable/long syllable (. –) was needed, *nebīẕ* 'wine' fit better into the scheme than the original Turkish word *süci*. In the artistic *ʾinšāʾ* prose, Arabic and Persian words, expressions, and patterns were used wherever possible regardless of whether there was a Turkish expression for the concept. The concept 'to walk (on foot)', which can be expressed in Turkish as *yaya(n) yürümek*, would be rendered in such a text as *rājilen mešy etmek* (Bittner 1900:27). It is not surprising that titles of books and headings of chapters were preferably entirely Arabic or Persian in such a context.

In the classical period of written Ottoman Turkish, Arabic loanwords outnumbered the native elements in the language by far. However, most of the Arabic loans were nominal expressions integrated into an (open) subsystem of loanwords from prestigious Oriental languages dominated by Persian rules for the formation of noun phrases. Even though for a long time loanwords from Arabic and Persian formed more than 80 percent of the vocabulary, Ottoman Turkish never was a 'mixed language'. The basic grammatical relations and the verbal core remained Turkish. This was the point of departure for far-reaching changes and reforms in the course of the 20th century.

3. Modern Turkish

The development of Ottoman Turkish educated and written styles based on a huge number of Arabic and Persian elements did not exert much influence on the language spoken by the uneducated masses in rural districts. The resulting diglossia was not felt to be a problem until the second half of the 19th century, when Turkish intellectuals and officials, influenced by Western concepts of equality, unity of the nation, and participation of the people in the political process, began to understand that the linguistic situation in the empire could form an obstacle to reforms in society. As a result of the founding of an 'ethnically homogeneous' and secular Turkish Republic after the dissolution of the multiethnic Islamic Ottoman Empire, linguistic nationalism began to dominate the discussion of the language issue. Radical purists, who wanted to substitute all foreign, especially Arabic and Persian, elements in the language, had the upper hand, while 'simplifiers', who mainly wanted to reduce unnecessary grammatical elements (e.g. Arabic broken plurals), exercised less influence. The process of purification of the Turkish language was speeded up after the transition to Latin script in 1928. As a result of this, the percentage of Arabic loanwords in educated written discourse was reduced heavily. After establishing a multiparty system, language policy led to severe disputes in Turkish society. Especially in the 1970s, a diglossic situation in written styles began to develop anew as social democrats and leftists tried to avoid loanwords from Oriental languages wherever possible while the conservatives insisted on the use of traditional expressions. After the coup d'état of 1980, the military government and most of the elected successors supported a more conservative attitude toward language reform (Brendemoen 1990; Lewis 1999).

Arabic loanwords still form an inevitable lexical stock in all Turkish styles, especially in the spheres of law, religion, education, etc. Despite all efforts in purification, words like *kitap* 'book', *kütüphane* (Arabic pl. *kutub* + Persian *xāna* 'house') 'library', *mektup* 'letter', all based on the Arabic root *k-t-b*, are without Turkish equivalents. Some Arabic loanwords are still in use but semantically restricted: Ottoman Turkish *ʿamelīyāt* 'the deeds of a man, practical performances' (Redhouse 1890:1321) means 'surgical operation' in modern Turkish (New Redhouse 1996:56 *ameliyat*). Many 'pure Turkish' neologisms were inspired by the semantics of the respective Arabic terms: *devre* 'period, cycle' < *dawra* was replaced by *dönem*, a derivation from the Turkish verb *dön-* 'to turn around' (cf. *dāra* 'to turn around'; Röhrborn 2003:119). Sometimes, even Arabic word fields were transferred into Turkish: *soru* 'question' (from *sor-* 'to ask') instead of *suʾāl* 'question' (both with s-anlaut followed by a labial vowel!;

cf. *saʾala* ‘to ask’), *mesʾele* ‘problem’ ⇒ *sorun*, *mesʾūlīyet* ‘responsibility’ ⇒ *sorumluluk*.

The new Turkish Latin script, which was successfully introduced in 1928, was phonetic as well as phonemic in character. As a rule, Arabic loanwords were written in accordance with their educated pronunciation (Steuerwald 1964 is the best introduction to related problems). The complete loss of Arabic and Persian orthographic traditions and distinctions, which had been preserved in the Ottoman Turkish writing system, was welcomed by the reformers as an act of liberation. The relation between the phonetic structure of Arabic loanwords in Turkish and their original pronunciation is too complex to be treated in this framework in detail (brief introduction in Lewis 1975:1–24; Deny 1921:79–88). Generally, Arabic long vowels are preserved with respect to their quality, but only partially with respect to their quantity (Prokosch 1981:91). In open, stressed syllables, Arabic length is more likely to be preserved (Prokosch 1983:345): *cevap* ‘answer’ (< *jawāb*); *ceva:bı* ‘his answer’. Secondary length in Arabic loanwords is a result of the loss of consonants, especially ʾ: *te:sir* ‘effect’ (< *taʾṯīr*). Arabic short vowels split up according to certain features of the surrounding consonants that trigger front or back vowels (Prokosch 1981), so that with adjacent ʿ, *ḍ*, *ġ*, *x*, [*ḥ*], *q*, *ġ*, *ṭ*, *ḏ̣*, the following changes take place: *a* > *a*, *e*; *i* > *ı*, *i*; *u* > *u*, *ü* (*o*, *ö*); some examples: *kutup* ‘pole’ (< *quṭb*); *kütüphane* ‘library’ (< *kutub* ‘books’); *saf* ‘row’ (< *ṣaff*); *sefer* ‘voyage’ (< *safar*). As a surrounding consonant, ʿ causes further deviations (Bouvat 1903:318), for example: *i* > *a*: *aşk* ‘love’ (< *ʿišq*). Turkish does not possess diphthongs, thus, Arabic diphthongs in loanwords are analyzed as a sequence of vowel + consonant: *tevhid* ‘unification’ (< *tawḥīd*). Not unlike the situation in Persian, many phonemic oppositions in the system of Arabic consonants are lost when Arabic words are borrowed into Turkish. Several distinction are lost without traces (*ṯ* vs. *s* > *s*), while others are recoverable through the front or back quality of the surrounding vowels (*q* vs. *k* > *k*). Some marginal Turkish phonemes, like *f* and *h*, were reinforced as a result of Arabic loanwords. Geminated Arabic consonants are not preserved in auslaut position and tend to be simplified between vowels: *kere* ‘time’ (< *karra*; Toparlı 1985:165). Vowel harmony as an overall principle of Turkish phonetics generally does not operate on Arabic loanwords, *müzakere* ‘negotiation’ (< *muḏākara*) having front and back vowels. Exceptions are a result of popular pronunciation: *mümkün* ‘possible’ (< *mumkin*). Some Arabic loanwords with a back vowel in the last syllable take front-vowel suffixes, but the underlying rules and tendencies have not yet been sufficiently described (Prokosch 1996:38–39): *sıhhat* ‘health’ ⇒ *sıhhati* ‘his health’ (< *ṣiḥḥa*). Neither Arabic nor Turkish allows consonant clusters in anlaut position, but in auslaut position the situation is quite different. Here, most of the Arabic clusters are broken up by high epenthetic vowels in the process of adaptation: *keşif* ‘discovery’ (< *kašf*), but *keşfi* ‘his discovery’.

Even if the Arabic *nisba* ending is of minor importance in modern Turkish, it is still made visible as *-î* in order to make a distinction between *dinî* ‘religious’ and *dini* (*din* + possessive suffix) ‘his religion’. The explicit distinction between nouns and adjectives once imported from Arabic and Persian is now reproduced with the neologistic suffix *–(s)al/-(s)el*: *dinsel* ‘religious’. Gender marking has lost importance in Arabic loans: *müdür* (< *mudīr*) ‘director’ is nowadays also used for women, replacing the feminine form *müdire*. Surviving broken plurals are now lexicalized: *ahbap* < *ʾaḥbāb* ‘friend’ (Arabic pl. of *ḥabīb*). Many other Arabic morphological elements are still identifiable etymologically, but they no longer have any function: *ender* ‘very rare’ is the elative of *nadir* ‘rare’. Since there are fewer than five lexemes of this type left in living Turkish, this relationship cannot form a paradigm, not even in a nonproductive sense.

The few Turkish syntactic structures influenced by Arabic patterns were not in the focus of linguistic purists. An exception was the Arabic coordinative conjunction *ve*, whose use was suppressed in Turkish schools. Some writers tried to avoid its use completely, employing several Turkish syntactic devices instead, for example converbs (Johanson 1992:265–269). In some cases, one observes a functional differentiation between synonymous elements of Arabic (*lâzım* ‘necessary’) and Turkish origin (*gerek* ‘id.’) in specific modal constructions (Ersen 2001:173).

Direct Arabic Turkish language contact is still an important factor in southeastern

Anatolia (→ Turkey). Turkish generally is the dominant language and exercises considerable influence on the Arabic dialects of the region. But there are also many lexical elements borrowed from Arabic dialects into Turkish dialects (Eren 1960). Unlike Arabic elements in the standard language, there are a large variety of forms. Arabic *šatl* 'seedling plant', for example, is found throughout Anatolia as *şetil*, *şitil*, *sitil*, *çetil*, *çitil* (Tietze 1958:285; Eren 1999:390). Dialectal *kada* 'fate' (< *qaḍāʾ*) is supposed to be a Syrian Arabic loanword, whereas *kaza* 'id.' in the standard language represents Ottoman Turkish literary tradition (Tietze 1958:312). Several popular Turkish forms of Arabic loanwords are now used in the standard language: *mutfak* 'kitchen' (< *maṭbax*).

Arabic loanwords still form an important part of the Turkish lexicon, but they no longer constitute the lexical core of the educated language as part of an independent and open subsystem. Mainly as a consequence of alphabet reform, the internal lexical structure of borrowed Arabic lexemes is concealed. It depends on the further development in Turkish society and politics whether Arabic loanwords will take their historical place between other foreign components in the lexicon, such as French, Persian, or Greek elements, or whether they will recover some of the lost domains.

Bibliographical references

Battersby, Harold K. 1966. "Arabic and Persian elements in Ottoman Turkish (Osmanlıca)". *Reşit Rahmeti Arat için*, 94–141. Ankara: Türk Kültürünü Araştırma Enstitüsü.

Bittner, Maximilian. 1900. *Der Einfluss des Arabischen und Persischen auf das Türkische*. Vienna: Carl Gerold's Sohn.

Bouvat, Lucien. 1903. "Les emprunts arabes et persans en turc osmanli". *Keleti Szemle* 4.316–334.

Brendemoen, Bernt. 1990. "The Turkish language reform and language policy in Turkey". *Handbuch der türkischen Sprachwissenschaft,* I, ed. György Hazai, 454–493. Budapest: Akadémiai Kiadó.

Buğday, Korkut. 1999. *Osmanisch: Einführung in die Grundlagen der Literatursprache*. Wiesbaden: O. Harrassowitz.

Deny, Jean. 1921. *Grammaire de la langue turque (dialecte osmanli)*. Paris: Presses Universitaires de France.

Eren, Hasan. 1960. "Anadolu ağızlarında Rumca, İslâvca ve Arapça kelimeler" [Greek, Slavic, and Arab lexemes in Anatolian dialects]. *Türk Dili Araştırmaları Yıllığı-Belleten*, 295–371.

——. 1999. *Türk dilinin etimolojik sözlüğü* [Etymological dictionary of the Turkish language]. Ankara: Bizim Büro.

Ersen, Margarete. 2001. *Türkische Grammatik*. Ismaning: Hueber.

Hazai, György. 1973. *Das Osmanisch-Türkische im 17. Jahrhundert: Untersuchungen an den Transkriptionstexten von Jakab Nagy de Harsány*. Budapest: Akadémiai Kiadó.

Ivanov, Sergej Nikolaevič. 1973. *Arabizmy v tureckom jazyke* [Arabic loanwords in Turkish]. Leningrad: Izdatel'stvo Leningradskogo Universiteta.

Johanson, Lars. 1992. *Strukturelle Faktoren in türkischen Sprachkontakten*. Stuttgart: F. Steiner.

Kerslake, Celia. 1998. "Ottoman Turkish". *The Turkic languages*, ed. Lars Johanson and Éva Ágnes Csató, 179–202. London: Routledge.

Kirchner, Mark. 1996. "Zwei osmanische Bearbeitungen des persischen *Qâbûsnâma* als Quelle zur türkischen Sprachgeschichte". *Symbolae Turcologicae: Studies in honor of Lars Johanson…*, ed. Árpád Berta, Bernt Brendemoen, and Claus Schönig, 143–155. Istanbul: Swedish Research Institute.

——. 2005. "Adverbial clauses in an Old Ottoman Turkish interlinear version of the *Koran*". *Linguistic convergence and areal diffusion*, ed. Éva Ágnes Csato, Bo Isaksson, and Carina Cahani, 305–315. London: RoutledgeCurzon.

Lewis, Geoffrey. 1975. *Turkish grammar*. Oxford: Oxford University Press.

——. 1999. *The Turkish language reform: A catastrophic success*. Oxford: Oxford University Press.

New Redhouse. 1996. *Turkish-English dictionary*. 16th ed. Istanbul: Redhouse.

Procházka, Stephan. 2005. "The Turkish contribution to the Arabic lexicon". *Linguistic convergence and areal diffusion*, ed. Éva Ágnes Csato, Bo Isaksson, and Carina Cahani, 191–202. London: RoutledgeCurzon.

Prokosch, Erich. 1981. "Der Vokalismus der arabischen Fremdwörter im Osmanisch-Türkischen und im Azärī". *Österreichische Osthefte* 23: 1.91–98.

——. 1983. "Vorstudien zum arabischen Wortgut im Osmanisch-Türkischen: Grundsätzliches und Konsequenzen für die Osmanistik". *Journal of Turkish Studies* 7.337–355.

——. 1996. "Zur Konsonantenvokalharmonie im Osmanisch-Türkischen und im heutigen Türkischen: Der Vokalismus der Sproßvokale, der persischen Izafet und der türkischen Suffixe an Fremdwörtern". *Grazer linguistische Studien* 45.33–43.

Redhouse, Sir James W. 1890. *A Turkish and English lexicon*. Constantinople: The American Mission.

Röhrborn, Klaus. 2003. *Interlinguale Angleichung der Lexik: Aspekte der Europäisierung des türkeitürkischen* Wortschatzes. Göttingen: Vandenhoeck and Ruprecht.

Römer, Claudia. 1981. "Der Einfluß der Übersetzungen aus dem Persischen auf die Entwicklung des Osmanischen im 14. und 15. Jahrhundert". *Wiener Zeitschrift für die Kunde des Morgenlandes* 73.89–114.

Stachowski, Stanisław. 1975–1986. *Studien über die arabischen Lehnwörter im Osmanisch-Türkischen*. 4 vols. Wrocław: Polska Akademia Nauk.

Steuerwald, Karl. 1964. *Untersuchungen zur türkischen Sprache der Gegenwart.* II. *Zur Orthographie und Lautung des Türkischen.* Berlin: Langenscheidt.

Tietze, Andreas. 1958. "Direkte arabische Entlehnungen im anatolischen Türkisch". *Mélanges Jean Deny*, ed. János Eckmann, Agâh Sırrı Levend, and Mecdut Mansuroğlu, 255–333. Ankara: Türk Dil Kurumu. (Repr. in Tietze 1999.)

——. 1990. "Die fremden Elemente im Osmanisch-Türkeitürkischen". *Handbuch der türkischen Sprachwissenschaft,* I, ed. György Hazai, 104–118. Budapest: Akadémiai Kiadó.

——. 1992. "Überlegungen über die lautliche Form der arabischen und persischen Lehnwörter im anatolischen Türkisch". *Wiener Zeitschrift für die Kunde des Morgenlandes* 82.349–358.

——. 1999. *Wörterbuch der griechischen, slavischen, arabischen und persischen Lehnwörter im Anatolischen Türkisch.* Istanbul: Simurg.

——. 2002. *Tarihi ve etimolojik Türkiye Türkçesi Lugatı: Sprachgeschichtliches und etymologisches Wörterbuch des Türkei-Türkischen*, A.-E. Istanbul and Vienna: Simurg.

Toparlı, Recep. 1985. "Türkçe'de galatlar" [Barbarisms in Turkish]. *Türk Dünyası Araştırmaları*, Şubat [February] 1985.

MARK KIRCHNER (Giessen)

Turkish Loanwords

1. HISTORY OF CONTACTS

Close contacts between the Arabs and speakers of Turkic languages go back to the first half of the 9th century, when the Abbasid caliphs began recruiting Turks from Central Asia as Praetorian guards. Although some of these mercenaries – for instance ʾAḥmad ibn Ṭūlūn and his successors – were even de facto rulers of Egypt (868–905), their language left hardly any traces in Arabic. The same is true of such later Turkish dynasties as the Ikhshidids and Seljuks. However, during the rule of the Mamluks in general (13th–16th centuries), and of the Baḥrī Mamluks (whose sultans were Kipchak Turks) in particular, at least the Arabic of Egypt was significantly influenced by a Turkic language. Nevertheless, the vast majority of Turkish loans in both written and colloquial Arabic date from the time of the → Ottoman Empire, which for about four hundred years dominated a large part of the Arab world. The influence of the Turkish language during that time even reached regions not under direct control of the Ottomans, such as Oman and Morocco. The disintegration of the Ottoman Empire after the First World War brought to an abrupt end Arabic/Turkish language contacts in most of the former provinces of the empire and resulted in a relatively rapid decrease in the use of Turkish words in all registers of Arabic. The remaining small number of Turkish-speaking minorities in Syria, and the so-called Turkmens of Iraq, have played only a limited regional role for the further transmission of Turkish loans into Arabic. The only exception to this general tendency are the Arabic minorities within the boundaries of present-day → Turkey. Because of the strong impact of Turkish in the educational system and the media, the dialects of these Arabic speakers are still influenced by Turkish not only in vocabulary but also to some degree in morphology and syntax (for the situation in Cilician Arabic, see Procházka 2002:184–203).

In the following sections, the region is always specified for dialectal words; examples taken from Modern Written Arabic are not explicitly indicated as such. Although Ottoman was the origin of nearly all Turkish loans into Arabic, the Turkish etyma – unless otherwise indicated – are cited in modern Turkish orthography and phonological form to enable the reader to find them in Turkish dictionaries. The term 'Turkish loan' is used in the sense that Ottoman Turkish was the transmitter of the words in question, whether these words were originally Turkish or borrowed by Turkish from another language.

2. WAYS AND PERIODS OF BORROWING

The unavailability of historical dictionaries of Arabic, together with our limited knowledge of Kipchak Turkish, makes it sometimes difficult to distinguish Mamluk from (especially earlier) Ottoman loans. The Turkic language of the Mamluk upper class influenced personal names and official titles, as in the names Aq Sunqur (< *aq sonqur* 'white falcon') and Tīmūr Buġā (< *timur boġa* 'iron bull'). However, the extent to which other words entered Arabic during the Mamluk period, and which of them have survived, has not yet been investigated. A pre-Mamluk origin is in fact claimed for *dabbūs* < *topuz* 'mace; globular knob' (see Ateş 1966:30) and for *xāzūq* < **xazuq* (modern: *kazık*) 'pale, stake; pile' (the Egyptian *ʾazaʾ*

< *qazaq* < *qazıq* 'telegraph poles' is clearly a later borrowing reflecting the Ottoman form of the same word). Likely Mamluk loans are *buqša* < *boqča* (modern *bohça*) 'bundle', and *xāšūga/xāšū'a* < **xašuq* (modern *kaşık*) 'spoon', the latter used in Iraqi and Syrian dialects (see Procházka 2004).

On the basis of the semantics and the phonological shape of the loanwords, it can be assumed that the bulk of Ottoman-Turkish loans found their way into Arabic during the 18th and 19th centuries. The three most important routes of transmission of Turkish loans were through official written Ottoman as the administrative language of the empire; through the standardized Ottoman Turkish spoken by the military class, government officials, and probably traders; and through vernacular Turkish in those areas with a denser Turkish population (for direct loans from Turkish dialects, see Procházka 2004).

Except for the technical terms of Ottoman civil and military bureaucracy, most Turkish words found their way into written Arabic and (for obvious reasons) especially into Arabic dialects, due to direct contacts, that is, through oral rather than literary borrowing. In many instances, the spelling of the Turkish loans in Arabic do not precisely follow the orthography of Ottoman, and clear cases of spelling pronunciations are rare (cf. Procházka 2004).

Because a remarkably large proportion of the Ottoman lexicon consisted of foreign elements, Arabic also picked up a number of non-Turkish lexemes from Ottoman, including words not only from Persian, Greek, and Slavonic but also from French and Italian. Western words entered Arabic through Ottoman because, until the 19th century, Turkish was the main language for the transmission of Western ideas and culture into the Arab world (for the special case of reborrowed Arabic words, see Sec. 6, below). Although it is not always possible to separate direct loans from these other languages into Arabic from those which entered Arabic indirectly via Turkish, in many cases a Turkish transmission can be detected because of phonological (e.g. restricted use of geminated consonants in colloquial Turkish) or semantic adaptations these words underwent in Turkish.

Of Persian origin are, for instance, *xumbara* 'small jar' > Ottoman *kumbara* 'bombshell' > *qunbula* 'bomb, grenade'; *čādor-šab* > *çarşaf* > *šaršaf* 'bedsheet'. Ultimately Greek is *sēmadoúra* > *şamandıra* > *šamandūra* 'buoy'; Italian (see Behnstedt 1996, esp. 64–65) is represented by *timone* > *dümen* > *dūmān* 'rudder, helm' and *caciocavallo* > *kaşkaval* > *qašqawān* 'a kind of cheese'. French in origin is *capote* 'soldier's coat' > *kaput* 'soldier's coat; condom' > Egypt *kabbūt* 'condom'. From (South) Slavonic is the toponym *an-Nimsā* 'Austria' < *Nemçe* 'German speakers, Austrians' < *niemce* 'German'.

3. Phonology of Turkish loans

In most cases, Ottoman Turkish words in Arabic have undergone phonological changes in order to substitute sounds not known to Arabic and to adapt the loans to common morphological patterns. Although some sound changes appear quite regularly, on the whole there are no phonetic rules for them since there are so many exceptions and no conformity in sound changes among the different dialects. Therefore, the following sketch should not be taken to be a complete description of all possible sound shifts. Turkish consonants, of which only six, *č, g, p, v, ñ, ž* (the latter two play no role for the loans), are unknown to Standard Arabic, were generally much less affected by changes than the nine vowel phonemes.

In both written Arabic and most Arabic dialects, Turkish *ç* [tʃ] usually has become *š*, e.g. *çanta* > *šanṭa* 'suitcase', *çuval* > *šuwāl* 'sack'; Tunis *bıçkı* > *bəšqi* 'mincing knife'. Sometimes, however, and especially in final position, it is *j*, e.g. *saç* > *ṣāj* 'thin sheet iron'. The sound *č* is maintained in the Iraqi dialects, which have a phoneme *č* (< *k*), and in the Anatolian and North Syrian dialects, which, doubtless because both regions had the earliest and most direct contacts with the Ottomans, adopted this new phoneme under the influence of Turkish (and Kurdish). Examples are: Iraq *suç* > *ṣūč* 'blame, fault', *alçak* > *alčaġ* 'rotten, low-down'; Anatolia *bekçi* > *bakči* 'guard'; and Syria *çay* > *čāy* 'tea', *çöl* > *čōl* 'steppe' (for the occurrence of *č* in Syria, see Behnstedt 1997, maps 18–30).

The sound *g* is generally reflected as *ġ* if followed by a back vowel (e.g. *damga* > *damġa* 'stamp, hallmark'), but as *k* if followed by a front vowel (e.g. *sergi* > *sarkī* 'bill of exchange'). The *j* in *jumruk* < *gümrük* 'customs' (ultimately

< Latin *commercium*) is probably a spelling pronunciation from the Egyptian use of the letter *jīm* for *g*. In dialects which already possess the phoneme *g* (e.g. Egypt, Yemen, and Bedouin dialects), the Turkish *g* remains unchanged. The same is true for the Anatolian and numerous Syrian dialects, in which the large number of Turkish loans have led to the adoption of a marginal phoneme *g*, e.g. Anatolia *gəmləke* < *gömlek* 'shirt'.

The voiceless *p* nearly always becomes *b*, e.g. *paşa* > *bāšā* 'pasha'; Iraq *tepsi* > *tabsi* 'tray'; Syria *top* > *ṭōb* 'cannon'; Tunisia *sepet* > *sbät* 'basket'. The labiodental *v* is either reflected as *w* (e.g. *çavdar* > *jawdār* 'rye') or, especially in initial position, as *b* (e.g. *vapur* > *bābūr* 'steamship').

The Turkish *ğ* (in Ottoman, pronounced *ġ* in the vicinity of back consonants) normally appears as *ġ*, e.g. *bağa* > *bāġa* 'celluloid', but sometimes as *k*, e.g. in *ığdış* (modern *iğdiş*) > *kadīš* 'cart horse, nag' (probably via dialectal *gdīš*). Ottoman *q* is sometimes reflected as *k*, as in *bakraç* > *bakraš* 'kettle'. The shift *q* > *x*, however, might be indicative of an older loan (see above, *xāzūq*).

Quite frequent, but with regional variations, is the velarization of *d, t, z, s* in the vicinity of back vowels, e.g. Egypt *oda* > *ʾōḍa* 'room'; written Arabic *tava* > *ṭawwāya* 'frying pan', *boza* > *būẓa* 'a beerlike beverage', and *sağ* > *ṣāġ* 'right, proper'. Metathesis of consonants is also sometimes found, e.g. *zincir* > *jinzīr* 'chain'; Syria *başlamak* > *ballaš* 'to begin', and *çapkın* 'good-for-nothing' > *šaqban* 'to twaddle'.

In the dialects, Turkish consonants are often affected by the same sound shifts as the corresponding Old Arabic consonants. Thus, Turkish *c* [dʒ] appears as *g* in Egypt, but as *ž* in parts of Syria and the Maghreb. Ottoman *q* is usually reflected as ʾ in those urban dialects where Old Arabic *q* has become a glottal stop, e.g. *qazma* > Cairo *ʾazma* 'pickax', Jerusalem *qışla* > *ʾišle* 'barrack'.

The treatment of the nine different Ottoman Turkish vowels is not homogeneous, neither in written Arabic nor in the dialects. The following very rough rules have countless exceptions: *a, ä > a, ā; e, i, ı > i, ī; o, ö, u, ü > u, ū*. Many vowel changes, though, are clearly the result of the tendency to reshape Turkish words in Arabic patterns. This is often true for nouns and always the case for verbs. Examples are *başlık* > *bašnūqa* 'kerchief', *çizme* > *jazma* 'boots', *fişek* > *fašak* 'cartridges'. The dialects usually also treat the vowels of loanwords according to their own phonetic rules: unstressed short vowels (especially *i, u, ə*) in open syllables are often elided, e.g. Syria *konak* > *qnāq* 'halting place', Tunisia *börek* > *brīk* 'a kind of fried pastry'.

Initial Turkish *a* is sometimes preceded by ʿ, e.g. *araba* > *ʿaraba* 'carriage, wagon'. Several loanwords differ from the vocalization of modern Turkish since they reflect older Ottoman, e.g. *duğrı* (modern *doğru*) > *duġri* 'strait' (in, e.g., Palestine, Egypt), and *vergü* (modern *vergi*) > Palestine *wērko* 'real estate tax'.

4. MORPHOLOGY OF TURKISH LOANS

Many Turkish loans, nouns as well as verbs, have been integrated into Arabic by adapting them to Arabic patterns. Borrowed nouns are usually masculine unless they end in *-a* and are therefore regarded as feminine, e.g. *çevirme* > *šāwirma* 'charcoal-broiled mutton'; *tencere* > *ṭanjara* 'casserole'. The construct form of these feminine nouns is regularly used, e.g. *çorba* > *šurba* 'soup', *šurbat ʿadas* 'lentil soup'; Syria *oda* > *ʾōḍa* 'room', *ʾōḍet səfra* 'dining room'. Adjectives, however, are often invariable for gender, e.g. *sade* > *sāda* 'simple, plain'; Syria *çürük* > *čərok* 'rotten'; Iraq *yasak* > *yaṣaġ* 'illegal'; but Syria *zengin* > *zangīn*, fem. *zangīne* 'rich'. As the number of borrowed adjectives is relatively small, elative forms are rare and seem to exist only in the dialects, e.g. Palestine *ašlab* 'prettier' < *šalabi* < *çelebi*.

Collective nouns such as *fişek* > *fašak* 'cartridges' or Syria *bürüncük* > *brənjok* 'fine crêpe' form a *nomen unitatis* according to the usual rules, i.e. *fašaka, brənjke*. The most frequent form of pluralization is the suffix *-āt* (or *-w/yāt*), e.g. *kıskaç* > *quṣāj, quṣājāt* 'pliers', *paşa* > *bāšā, bāšāwāt* 'pasha'. Nouns whose singular is in an Arabic pattern, however, often exhibit internal plural forms, e.g. *balta* > *balṭa, bulaṭ* 'ax'; *kemer* > *kamar, ʾakmār* 'belt'; *parmak* > *barmaq, barāmiq* 'spike'; *tabur* > *ṭābūr, ṭawābīr* 'queue'.

Verbs are usually regarded as a more intensive integration of foreign words into a language than nouns. In both written and dialectal Arabic, all verbs of Turkish origin have been altered to Arabic patterns for the sake of inflection. Many of these verbs were not directly borrowed from Turkish verbs but derived rather from Turkish nouns already integrated into Arabic. At least in written Arabic, the number of verbs derived from Turkish is not large; examples are *baṣama* 'to print' < *baṣma* < *basma* 'print', *jamraka* 'to take toll' < *jumruk* < *gümrük* 'customs', *farraša* 'to brush' < *furša* < *fırça* 'brush', *dawzana* 'to tune' < *dūzān* < *düzen* 'tune'. In most dialects, the number of verbs going back to Turkish is much larger. For Syria, Halasi-Kun (1969:29, 82–84) reports 75 verbs, e.g. Syria *yasak* 'forbidden' > *yassaq* 'to forbid'; Iraq *çizmek* > *čazz* 'to mark over'; Egypt *kılavuz* 'screw-tap' > *ʾalwaẓ* 'to screw down'.

Turkish derivational suffixes have usually been borrowed as a unit with the root word and thus are found in all layers of Arabic. In some dialects (Iraq, Syria, Egypt), a few very frequent Turkish suffixes have become productive and are used to a limited extent with Arabic words in combinations unknown to Turkish. The best survey on this topic is Masliyah (1996), who covers the Iraqi dialects; for written Arabic, see the lists in Gülensoy (1975:129–133). By far the most common suffix is *-ci*, which is used for professions and (in Arabic, almost exclusively negative) characterizations. Halasi-Kun (1969:68–70) provides a list of about one hundred words for the Syrian dialects; for Egypt, see the explanations, including a long list, in Prokosch (1983a:70–73). Examples are *boyacı* > *būyaji* 'house painter, shoeshine', *hurdacı* > *xurdaji* 'dealer in miscellaneous smallwares'; Iraq *bāysikilči* 'bicycle seller', *ʿaragči* 'drunkard [i.e. addicted to arrack]'; Syria *batakçı* > *baṭaqji* 'swindler, gangster'; Egypt *makwagi* 'laundryman'. In addition to the relational suffix *-li* (e.g. Syria *Mardilli* 'a person from Mardin') and the suffix *-siz*, added to nouns to form adjectives meaning 'without' (e.g. Iraq *šarafsizz* 'without honor'), there also appears *-lik*, e.g. *tozluk* > *ṭūzluq* 'gaiters'. In Iraq, the latter suffix is, in a pleonastic combination with the Arabic ending *-iyya*, used to form abstract nouns, e.g. *zmāl* 'donkey' > *zmālləġiyya* 'stupidity'.

Elements of Ottoman compound nouns, either of Turkish or Persian origin, have also been borrowed into Arabic, e.g. *-xāne* 'house', *-dār* 'carrier of'. The noun *baş* 'head, chief of' is particularly frequent in newly created combinations in Arabic, for instance Syria *bāš-argəlži* 'senior waiter responsible for the hookah'; Algeria *bāš-kəḏḏāb* 'big liar'.

5. Lexical importance and semantic domains

For obvious reasons, one hundred years ago the number of Turkish loanwords in both written and spoken Arabic was considerably larger than it is today. The decline in direct contact, the Arabization of the official language, and, last but not least, the negative perception of the Ottoman era in today's Arab world have resulted in the rapid decrease of Turkish loanwords. Although no systematic research has been done, studies such as Barbot (1961), Prokosch (1983a), and Reinkowski (1998) suggest a drastic decline of the Turkish influence on spoken Arabic. Approximately half of the words quoted in the dictionaries of the Arabic dialects in question are no longer in active use or have even become unintelligible. A similar situation can be assumed for written Arabic (for Turkish loans in written Arabic in general, see Ateş 1965; Mutawallī 1991; Zahidi 1977). Thus, it is very difficult, if not impossible, to estimate how many Turkish loanwords still exist in Arabic. For contemporary written Arabic, the number certainly does not exceed 250. In the dialects, the somewhat dated figures are 3,000 for Syria (Halasi-Kun 1969:20), 1,150 for Egypt (Prokosch 1983a), and about 250 in Iraq (Reinkowski 1995).

Particularly symptomatic of the quantitative and semantic decay of Turkish loanwords in Arabic is the history of the old Ottoman-Turkish titles since the breakdown of the Ottoman Empire. Many of the Ottoman titles for military ranks, such as *ʾōmbāšī* < *onbaşı* 'corporal', *yūzbāšī* < *yüzbaşı* 'captain', and *mīr ālāy* < *miralay* 'brigadier general', were officially used in some Arab armies until the 1950s. However, today these words appear almost exclusively in historical novels or films. Some of the old Ottoman titles are now applied to menial jobs or positions, or used with ironic or even pejorative connotations. For instance, in Tunisia *šāwuš* < *çavuş*, formerly 'sergeant', now means 'office boy, gatekeeper'. In Egypt, *balṭagi* < *baltacı* 'pioneer' is now used in the sense of 'gangster, rowdy, bouncer'. In written Arabic, *šalabī* < *çelebi* 'gentleman, prince' sometimes appears in the sense of 'dandy, fop'.

In spite of the aforementioned examples, and the universal tendency to a lowering of social status in titles, several Turkish words are still used in many Arab countries as polite forms of address (→ terms of address; for Jordan, see Prokosch 1989). In Egypt, titles such as *bāšmuhandis* < *başmühendis* ‘(chief) engineer’ or *hānim* < *hanım* ‘lady’ are widespread and are regarded as very polite (Rosenbaum 1998:100).

Several other Turkish loanwords are now used infrequently simply because they denote tools, dresses, or fixtures which have become obsolete (see examples below). Most Turkish loans in Arabic fall into the following domains: administration and government, army and war, crafts and tools, house and household, dress, and food and dishes. The influence of Turkish on Arabic in these particular categories is obviously the consequence of the presence of the Ottoman bureaucracy and army in the Arab world in particular, and of the influence of centuries-long relations on everyday life in general. Moreover, many new things, such as fashions in dress or improved tools for craftsmen, reached the Arabs via Istanbul, for centuries the cultural center of the Islamic world. A few specific examples of Turkish loans into Arabic in each of the above-listed cultural categories follow:

Administration and government: *dönüm* > *dūnum* ‘a square measure’, *damga* > *damġa* ‘stamp’ (and *damaġa* ‘to stamp’), *gümrük* > *jumruk* ‘customs’, *zindan* > *zinzāna* ‘prison cell’.

Army and war: *binbaşı* > *bikbāšī* (with spelling pronunciation!) ‘lieutenant colonel’, *kol* > *qōl* ‘army corps’, *tabur* > *ṭābūr* ‘battalion, queue’, *lağım* > *laġam* ‘mine’, *tabanca* > *ṭabanja* ‘pistol’.

Crafts and tools: *takım* > *ṭaqm* ‘set (of tools), service’, *çengel* > *šankal* ‘hook’, *sinara* > *ṣinnāra* ‘fish hook’, *kılavuz* ‘screw-tap’ > *qalāwūz* ‘screw’, *yay* > *yāy* ‘spiral spring’.

House and household: *çeşme* ‘fountain’ > *šašma* ‘toilet’, *edephane* > *ʾadabxāna* ‘water closet’, *köşk* > *kušk* ‘kiosk’, *soba* ‘stove’ > *ṣōba* ‘stove, hothouse’ (e.g. in *taʾṯīr ṣōbī* ‘greenhouse effect’); Syria, Egypt *oda* > *ʾōḍa* ‘room’; Iraq *çekmece* > *čakmača* ‘drawer’.

Household vessels: *bakraç* > *bakraj* ‘kettle’, *kazan* > *qazān* ‘large boiler’, *leğen* > *lakan* ‘basin’, *tawa* > *ṭawwāya* ‘frying pan’, *teneke* > *tanaka* ‘tin can’; Tunis *cezve* > *zazwa* ‘coffeepot’.

Dress: *çintiyan* > *šintiyān* ‘loose trousers’, *çizme* > *jazma* ‘boots’, *kayış* > *qāyiš* ‘belt, girth’, *kundura* > *kundura* ‘(Western-style) shoe’, *yaka* > *yāqa* ‘collar’; Syria *şapka* > *šabqa* ‘hat’.

Food (including fruits) and dishes: colloquial *burgul* (Standard Turkish *bulgur*) > *burġul* ‘cracked wheat’; *dondurma* > *dandurma* ‘ice cream’, *kavurma* > *qāwirma* ‘fried meat’, *sucuk* > *sujuq* ‘sausage’, *meze* > *māza/mazza* ‘hors d’oeuvres’; fruits: Ottoman *yusuf efendi* > *yūsuf afandi, yūsufī* ‘tangerines’, *hıyar* > *xiyār* ‘cucumber’; *yemiş* > *yāmīš* ‘dried fruit’.

Not yet investigated systematically are the formally Arabic words coined by the Ottomans and then borrowed back into Arabic (see the preliminary study by Prokosch 1999). Many of these words are abstract nouns denoting ideas and concepts imported to the Middle East from Europe during the 18th and 19th centuries (see Lewis 1996). Among them are *jumhūriyya* ‘republic’, *baladiyya* ‘municipality’, *qawmiyya* ‘nationalism’, *madaniyya* ‘civilization’. Frequently found are semantic extensions of Arabic words under Turkish influence, the result of the new meanings the Ottomans gave to already existing Arabic words. Examples are (the Turkish forms in parentheses) *fiʾa* (*fiyat*) ‘price’, *sajjāda* (*seccade*) ‘prayer rug’, *fāʾiḍ*, also *fāyiẓ* (*faiz*) ‘interest [on money]’, *kīs* (*kese*) ‘Turkish towel’ (and likely > *mukayyis* ‘masseur’); Syria *xəṭyār* (*ihtiyar*) ‘old’.

6. PHRASEOLOGICAL AND SYNTACTIC INFLUENCES

Especially in the dialects, there are a large number of calques. But because of the lack of detailed studies on phraseology in both Arabic and Turkish, it is often impossible to decide whether these loan translations actually went from Ottoman into Arabic or vice versa. The following examples are, however, definitely of Ottoman origin: Syria *aslan sütü* ‘arrack’ (lit. ‘milk of the lion’) > *ḥalīb əsbāʿ*, *kazık yemek* ‘to be cheated' (lit. 'to eat a fraud') > *akal xāzūq*; Palestine *şöyle böyle* 'so and so' > *šēle bēle*; Iraq *ne var ne yok* 'what's new' (lit. 'what is and what is not') > *šaku māku*.

Apparently, there is no Turkish influence on the syntactic level, except in those Arabic dialects spoken within the boundaries of today's Turkey (for examples from the dialects of Cilicia, see Procházka 2002:199–202).

7. SUMMARY

Given the long-term and very intensive contacts between Turks and Arabs, there are surprisingly few traces of Turkish in (written) Arabic, especially in contrast with the Turkish influence on the Balkan languages. The two main reasons for this difference in impact are, first, the relatively small number of ethnic Turks who actually lived in the Arab provinces and, second, the fact that Arabic was a much more highly developed and, especially because of its status in Islam, prestigious language than the languages of the Balkans. Today, in spite of a relatively stable core of Turkish borrowings (many of them neither phonologically nor semantically recognized by Arabic speakers as being of Turkish origin), Turkish loanwords in Arabic continue to undergo quantitative decrease, semiological diminution, and marginalization.

BIBLIOGRAPHICAL REFERENCES

Ateş, Ahmed. 1965. "Arapça yazı dilinde Türkçe kelimeler üzerine bir deneme II: XI. yüzyıldan bugüne kadar" [About Turkish words in written Arabic. II. From the 11th century until today]. *Türk Kültürü Araştırmaları* 1–2.5–25.

——. 1966. "Arapça yazı dilinde Türkçe kelimeler (X. yüzyıla kadar)" [Turkish words in written Arabic until the 10th century]. *Reşid Rahmeti Arat için*, 26–31. Ankara: Türk Kültürünü Araştırma Enstitüsü.

Aytaç, Bedrettin. 1994. *Arap lehçelerindeki Türkçe kelimeler* [Turkish words in Arabic dialects]. Istanbul: Türk Dünyası Araştırmaları Vakfı.

Barbot, Michel. 1961. "Emprunts et phonologie dans les dialectes citadins syro-libanais". *Arabica* 8.174–188.

Behnstedt, Peter. 1996. "Romanisches Lehngut im Syrisch-Arabischen". *Romania Arabica,* ed. Jens Lüdtke, 63–72. Tübingen: G. Narr.

——. 1997. *Sprachatlas von Syrien.* Wiesbaden: O. Harrassowitz.

Ben Cheneb, Mohammed. 1922. *Mots turcs et persans conservés dans le parler algérien.* Algiers: Université d'Alger.

Gülensoy, Tuncer. 1975. "Yaşayan Farsça ve Arapçadaki Türkçe kelimeler üzerinde notlar" [Remarks about Turkish words in Persian and Arabic] . *İstanbul Üniversitesi Edebiyat Fakültesi Türk Dili ve Edebiyatı Dergisi* 21.127–157.

Halasi-Kun, Tibor. 1969–1982. "The Ottoman elements in the Syrian dialects." *Archivum Ottomanicum* 1.14–91; 5.17–95; 7.117–267.

Lewis, Bernard. 1996. "The Ottoman legacy to contemporary political Arabic". *Imperial legacy: The Ottoman imprint on the Balkans and the Middle East,* ed. L. Carl Brown, 203–213. New York: Columbia University Press.

Masliyah, Sadok. 1996. "Four Turkish suffixes in Iraqi Arabic: *-li, lik, siz* and *-çi*". *Journal of Semitic Studies* 41.291–300.

Mutawallī, ʾAḥmad Fuʾād. 1991. *al-ʾAlfāḏ̣ at-turkiyya fī l-lahajāt al-ʿarabiyya wa-fī luġat al-kitāba.* Cairo.

Parzymies, Anna. 1984. "Appellatifs turcs dans le dialecte arabe de Tunis". *Rocznik Orientalistyczny* 43.123–129.

Procházka, Stephan. 2002. *Die arabischen Dialekte der Çukurova.* Wiesbaden: O. Harrassowitz.

——. 2004. "The Turkish contribution to the Arabic lexicon". *Linguistic convergence and areal diffusion: Case studies from Iranian, Semitic and Turkic*, ed. Eva Csató Johanson a.o., 191–204. London: RoutledgeCurzon.

Prokosch, Erich. 1983a. *Osmanisches Wortgut im Ägyptisch-Arabischen.* Berlin: K. Schwarz.

——. 1983b. *Osmanisches Wortgut im Sudan-Arabischen.* Berlin: K. Schwarz.

——. 1989. "Zum osmanischen Wortgut im arabischen Dialekt von ʿAmmān". *Materialia Turcica* 15.14–23.

——. 1999. "Vorstudien zum Einfluß des Osmanisch-Türkischen auf die moderne arabische Schriftsprache". *Grazer Linguistische Studien* 52.101–120.

Reinkowski, Maurus. 1995. "Phonologische Veränderungen des türkischen Wortgutes im Bagdadisch-Arabischen, mitsamt einer Wortliste". *Folia Orientalia* 31.89–115.

——. 1998. "Türkische Lehnwörter im Bagdadisch-Arabischen: Morphologische Adaption an die arabische Schemabildung und Bedeutungsveränderung". *Turkologie heute: Tradition und Perspektive*, ed. Nurettin Demir and Erika Taube, 239–253. Wiesbaden: O. Harrassowitz.

Rosenbaum, Gabriel. 1998. "Address forms in Egyptian literature". *The Arabist: Budapest Studies in Arabic* 19–10.99–108.

Zahidi, Abbas. 1977. *Muasir Äräb dilindä Turk mänşäli sözlär* [Turkish words in contemporary Arabic]. Baku: Elm.

STEPHAN PROCHÁZKA (University of Vienna)

Turku Arabic → Pidgin Arabic: Bongor Arabic

Turkic Languages → Kazakh; Tatar; Turkish; Uzbek

U

Uganda → East Africa

Unaccusative → Middle Verbs

Urdu/Hindustani

1. HISTORICAL BACKGROUND AND DEFINITIONS

The linguistic label 'Hindustani' refers both to what is now known as Hindi and to what is known as Urdu, or, rather, to their recent common linguistic history. After 1858, this label, frequently used by colonial British authorities, came to refer to the North Indian common language that was to replace Persian for lower administration purposes, in addition to its already established use in the British colonial army. Indian Persian itself was an administrative and courtly language that had borrowed intensely from Arabic, often after Persian patterns, though in an Indian cultural context; as far as lexicology is concerned, it is referred to here as 'Indo-Persian'.

For a long time. Urdu and Hindustani were more or less regarded as identical, although the term *ūrdū* was often restricted to the language used in literary and poetic circles. Up to the middle of the 20th century, Hindustani was, in its official use, heavily loaded with loanwords from Persian and Arabic, while modern Hindi was growing apart under the influence of Sanskritization, itself a nationalistic reaction to colonialism. On the other hand, Urdu, since the end of 19th century, has gained in prestige and has become established in urban and popular use. Combining North Indian or Hindi syntax and basic vocabulary, it continued as official Hindustani, freely admitting loanwords and phrases from Persian. While Hindi was being written in Devanagari (derived from Sanskrit), Urdu retained the Arabo-Persian alphabet, in the → *nastaʿlīq* style, conserving all Arabic and Persian orthography and forging new combinations of letters for the transliteration of the Indian retroflex and aspirated consonants (Table 1), as well as diacritical signs (Table 2) and combinations of glides (Table 3) for the transliteration of the Indian vowels.

Some of the Arabic consonants became homophonous in this alphabet, such as *ḥāʾ* and *ḥāʾ*, both being pronounced as *h*, and *zāy*, *ḏ̣āʾ*, *ḍād*, and *ḏāl*, all being pronounced *z*. The pronunciation of *xāʾ*, *ġayn*, and *qāf* followed the Arabic convention (although *qāf* is also pronounced as *k*; for *ź* a letter was borrowed from the Persian alphabet (→ Arabic alphabet for other languages). Hindustani, especially for military purposes, had in addition a tradition of being written in Latin alphabet, although this would be discontinued after the colonial era.

Later on, Urdu became one of the regional and constitutional languages of the Indian Federation, as well as the national language of the newly created Muslim state of → Pakistan, also a federation, with its own official languages at the state (or province) level.

As far as literature is concerned, Arabic prosody and poetical genres, through Persian, and Indo-Persian conventions and images, were at the roots of the ever-growing literary tradition of Urdu. Indeed, Urdu poetry, in its early times (14th century C.E.), migrated with the language itself from the Delhi Court to the southern Dakhani courts, and in the mid-18th century, when it returned to the north at the end of the Mughal Empire, it was stronger than ever, with such lyrical genres as *masnavī, qasīdā, marsiyā*, and *ġazal*, all of them inherited from the Arabo-Persian tradition. Indeed, from the mid-19th century onward, when the great poet Ghalib (which was the *nom de plume* or *taxallus* under which he wrote) published his Urdu *dīwān* consisting mainly of *ġazal*s, *ġazal* became the favorite of Urdu-speaking audiences; since then, it has gained in popularity thanks to its use in the Indian cinema.

The continuous Persianization of Hindustani and Urdu (including Arabic elements), in a changing political frame extending roughly from the Delhi Sultanate through Mughal India and British colonization, and up to the modern independent South Asian countries, is illustrated in Figure 1. This figure shows the linguistic divide that took place under two diverging sociolinguistic trends still at work today, Persianization and Sanskritization, without trying to solve questions of linguistic classifications and chronology. The Urdu/Hindi labels (replacing colonial Hindustani) are practical sociolinguistic reconstructions from the end of the 19th century (schema adapted from Désoulières 1995, which was inspired by Southworth 1971:5).

Figure 1. Schema of Urdu-hindi historical links

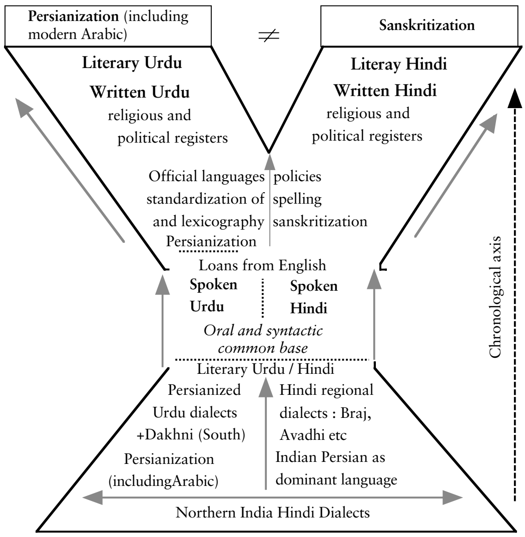

Persianization (including modern Arabic)
≠
Sanskritization
Literary Urdu
Written Urdu
religious and political registers
Literay Hindi
Written Hindi
religious and political registers
Official languages standardization of and lexicography
Persianization
policies spelling sanskritization
Loans from English
Spoken Urdu
Spoken Hindi
Oral and syntactic common base
Literary Urdu / Hindi
Persianized Urdu dialects +Dakhni (South)
Persianization (includingArabic)
Hindi regional dialects : Braj, Avadhi etc
Indian Persian as dominant language
Northern India Hindi Dialects
Chronological axis

Table 1. Urdu consonants in Arabic script

Urdu letter	name of the letter	transcription
ا	*alif*	–
ب	*be*	*b*
بھ	*bhe*	*bh*
پ	*pe*	*p*
پھ	*phe*	*ph*
ت	*te*	*t*
تھ	*the*	*th*
ٹ	*ṭe*	*ṭ*
ٹھ	*ṭhe*	*ṭh*
ث	*se*	*s*
ج	*jīm*	*j*
چ	*ce*	*c*
چھ	*che*	*ch*
ح	*he*	*h*
خ	*xe*	*xe*
دھ	*dāl*	*d*
ڈھ	*ḍāl*	*ḍ*
ذ	*zāl*	*z*
رھ	*re*	*r*
ڑ	*ṛe*	*ṛ*
ڑھ	*ṛhe*	*ṛh*
ز	*ze*	*z*
ژ	*źe*	*ź*
س	*sīn*	*s*
ش	*śīn*	*ś*
ص	*soad*	*s*
ض	*zoad*	*z*
ط	*ṭā*	t
ظ	*zoe*	*z*
ع	*ʿain*	ʿ(or corresponding Urdu vowel)
غ	*ġain*	*ġ*
ف	*fe*	*f*
ق	*qāf*	*k*
ک	*kāf*	*k*
کھ	*khāf*	*kh*
گ	*gāf*	*g*
گھ	*ghāf*	*gh*
ل	*lām*	*l*
م	*mīm*	*m*
ن	*nūn*	*n*
و	*vau*	*v*, *w*
ھ	*he*	*h*
ی	*ye*	*y*

Table 2. Diacritical vocalic signs

name	sign	transcription
zabar	ـَ	*a*
zer	ـِ	*i*
peś	ـُ	*u*
hamza	ـٔ	dieresis, not transliterated
hamza	آ	apostrophe (’)

Table 3. Urdu vowels

	long vowels	
final *ā* ا	medial *ā* ا	initial *ā* آ
final *ī* ی	medial *ī* ـِیـ	initial *ī* ای
final *ū* وُ	medial *ū* وُ	initial *ū* وُ
	middle vowels	
final *e* ے	medial *e* ـیـ	initial *e* ایـ
final *o* و	medial *o* ـو	initial *o* او
	short vowels	
initial *a* اَ		
initial *i* ا		
initial *u* اُ		
	short diphthongs	
initial *au* اَو		
initial *ai* اَے		
	nasal vowels	
long *ā* آں		
long *ī* ایں		
long *ū* وُں		
ẽ ایں		
o وں		
final *ā* آں	medial *ā* آ	initial *ā* آ

2. INTEGRATION AND ADAPTATION OF ARABIC LOANWORDS

Except when stated otherwise, the following observations are valid both for former Hindustani and for present-day Urdu. As a rule, in Urdu, the great majority of Arabic loanwords came through Indo-Persian, which explains their phonetic changes as well as their semantic adaptation. The corollary is that many Arabic words are found with Persian affixes (the opposite being less frequent), and there are many Arabo-Persian compounds (mostly nouns).

Although some lexicological fields related to religion (law and justice, traditional medicine, historiography, linguistics, etc.) rely distinctly more than others on Arabic loanwords, it would be unrealistic to equate Arabic (indirect) linguistic influence with Islam. In fact, loanwords from Arabic may be present in any lexicological field, from the earliest times onward, because of the prevalence of Arab navigation and, later on, the universal presence of Indo-Persian. Indo-Persian was once the administrative and court language of the Mughal Empire, not to mention the earlier Delhi Turkish and Pathan Sultanates or the southern Golkunda Kingdom and others, which also used Persian language and culture.

Even the vocabulary of Hindu states was influenced by the Arabo-Persian fashion: for administrative purposes, Persian (and Classical Urdu) forged Arabic plurals for the very words *hindū* ‘a Hindu person’ and *mandir* ‘a Hindu temple’, which may have *hinūd* and *manādir* as plural forms. Indeed, *manādir* may have been modeled after *manāzil*, plural of Arabic *manzil* ‘place of residence’ (present-day meaning in Urdu ‘stage, goal, destination; floor [of a building]’). Properly speaking, these are not Arabic loanwords but instead are interesting calques of the Arabic plural.

Because Indo-Persian was, for nearly a century and a half, the official language of the East India Company, English influence should not be neglected, even in apparently ‘pure’ Arabic loanwords. Some Arabic loanwords may turn out to be either semantic calques of English or compound neologisms dating from the colonial period and consisting of an Arabic and an Indian word.

2.1 *Phonetic adaptation*

Coming through the Indo-Persian crucible, Arabic loanwords lost all those phonemes that were foreign to that phonemic inventory. Thus, for instance, the glottal or posterior aspiration /ḥ/ was lost, while the anterior one (like the English /h/ phoneme) was retained. Similarly, these loanwords lost all interdental phonemes, which were replaced by /z/ sounds (unlike English). Another characteristic phonetic adaptation is the systematic replacement of the pharyngeal /ʿ/ in Arabic words by the nearest vowel. Those Arabic consonants that were common in the Indo-Persian stock remained in use in the pronunciation of modern standard Urdu, while they were foreign to the original North Indian stock and therefore were discarded by modern popular Hindi. This is the case with the uvular stop /q/ (Indo-Persian and Urdu name *qāf*), the voiced fricative /ġ/ (*ġayn*), the unvoiced labiodental /f/ (*fe*), the voiced dental /z/ (*ze*), and the palatal fricative /x/ (*xe*). Most of these consonants are regarded as phonetic hallmarks of proper Urdu diction, /q/ being in a weaker position. However, Arabic loanwords keep their Arabic spelling but lose diacritics that signal Arabic declension; some Arabic loanwords in Urdu are written with additional signs meant to help the Urdu-speaking reader when the original Arabic spelling may conflict with the Urdu graphic conventions. Modern Urdu phonology does not readily accept short vowels in word-final position, so Indo-Persian final *a* is often written as long *ā*.

2.2 *Morphological adaptation*

An interesting case of integration of Arabic morphology in modern Urdu, as well as in Classical Hindustani, is the use of Arabic nominal plurals. In Urdu, such plurals, e.g. *kitāb/kutub*, are not mandatory. Their use denotes a rather high literary level of language. Speakers using them wish to show that they have received some education. But an Arabic noun used with its genuine Arabic plural does not vary in the oblique or indirect case. In other words, the Arabic plural of nouns remains impervious to Urdu declension, although the Urdu gender is not affected.

In some cases, like clichés or calques from Latin plurals used in English, Arabic plurals are a must. As early as the end of 19th century, for instance, the English academic plural (borrowed from Latin) *alumni*, was calqued by the plural of an Arabic loanword *tulabā* 'students', while the singular *tālib e ilm* 'a student [masc.]' (later form *tālib ilam* without Persian *izafet* and adapted to Urdu phonology) shows a different process of word creation. The corresponding feminine noun ('a female student') was created later on; it was adopted with a more literary Arabic pattern, *ṭāliba*, and became *tālibā* to fit popular Urdu phonology, which has a constraint against short *a* in word endings. Because the literary plural of *tālibā, tālibāt*, is mandatory, no Indian plural can be used with this word.

In other cases, the Arabic plural may help in creating neologisms. A well-known instance of such a creation, which imitates Modern Arabic word creation, is the noun *xabar* 'a piece of news', whose plural *axbār* came to mean 'a newspaper'. Another (older) case is the Indo-Persian *nawwāb* 'Nabab' or, in the Mughal Empire, 'provincial governor; prince', from the singular noun *nāʾib*, itself historically abbreviated from *nāʾib sultān* 'vice-sultan'. Modern Urdu now has *nawāb*, from Hindustani *nawwāb*, a singular noun with the same meaning but also used to refer to the ruling elite of former princely states of colonial India, and their descendants. As in many other cases, popular characters in the cinema helped in imposing the popular form. An interesting lexical derivation by suffix is the feminine noun *nawābī*, for people belonging to the rich aristocracy, but with a somewhat derogatory secondary meaning. Note that the singular *nāʾib*, a masculine noun in Urdu, meaning 'substitute', is less common as a noun than as a prefix, like the English *vice-* in *viceroy*.

Another instance of adapting an academic Arabic plural to create a neologism is the literal plural of *hādisā*, in Urdu a masculine noun meaning 'a traffic accident', which gave *hādisāt*, a collective masculine noun and a semantic calque for the English term *emergencies*, referring to a hospital context.

2.3 *Syntactic adaptation*

Apart from some function words and a few prefixes, Arabic loanwords in Urdu are confined to the nominal field (nouns, adjectives, and adverbs, adverbs being close to adjectives and nouns, even when borrowed from Arabic). In very rare cases, an Urdu verb is derived from an Arabic root (see below).

Grammatical gender is an essential feature in Urdu for classifying nouns, depending on phonetic and semantic criteria (unlike Persian, which has suffixes); hence, all borrowed nouns are assigned a grammatical gender. Except in obvious cases, when a female semantic feature is present (e.g. *šāirā* 'poetess'), the gender of the borrowed Arabic noun is determined by its phonetic structure and ending, following a

North Indian and Hindi pattern. Some learned nouns adopted the gender of equivalent Sanskrit learned words. Urdu lexicographers issued rules for this purpose, more or less along these lines, but there are dubious cases. The Urdu gender of a noun borrowed from Persian may also influence the gender of an Arabic synonym loanword, e.g. Persian *sipāh* 'army' is a feminine noun in Urdu, and the more frequent Arabic synonym *fauj* 'crowd; troop; battalion' was also classified as feminine.

Urdu has borrowed from Persian some function words that are frequently used in modern speech. The most obvious of these words is *bād* 'after' < Arabic *baʿd* (with deletion of /ʿ/). It has become a common word in some extremely frequent compounds, e.g. the postposition *ke bād* 'after that [lit. "that after"]', where *ke* is an Indian postposition. It is also used in other compounds, such as *bād men* 'afterward'.

Another basic function word is the conjunction *lekin* 'but'. What is remarkable in this case is its incorporation in the Urdu system. Compared to other conjunctions that are close in meaning, such as the Indian *par* 'but; though', also used as a locative postposition meaning 'on', from which the meaning 'thereupon' is derived, or the Persian loanword *magar*, preferably used in a negative context, *lekin* has a more general meaning devoid of locative or negative nuance. It is comparable to English *but* and may not even signal a literary level of language, despite its Arabic origin. Phonetically, it was also easy to adapt, none of its components being foreign to the Indian or Persian phonetic system.

An interesting case of a function word integrated into Urdu syntax is the compound conjunction *hattā ke* 'insofar as' < Arabic *ḥattā* 'until' + Persian *ke* 'that', which in Urdu governs the optative or subjunctive mood. An interesting parallel is Spanish *hasta que*, Portuguese *até que*, both compounds of Arabic *ḥattā* with Romance *que*.

Some very frequent adverbs or adverbial phrases were borrowed in modern Urdu, which may also be regarded as function words, e.g. *bilkul* 'completely, totally' (< *bi-l-kull*) or *fauran* 'immediately' (< *fawran*). Adverbs formed from an Arabic noun in accusative case with the ending *-an* are a typical feature of modern Urdu. Such adverbs are freely produced by native speakers, although their use is more frequent for educated speakers.

A remarkable feature of borrowing Arabic words into Indo-Persian is that entire lexical series may be borrowed starting from a key word and an Arabic root that motivates the series, thus borrowing several derived forms according to Arabic morphology, although their Indo-Persian grammatical status and meaning may differ from the Arabic pattern. Such serial borrowings, or motivated lexical series, may occur in Urdu (after Indo-Persian), with yet another set of meanings and grammatical status. For example, although the root *k-t-b* is no longer present in Urdu as such, the words *kitāb* 'book', *kātib* 'copyist, transcriber', *kitābat* 'the act of copying before proofreading', and others, still exist. Nowadays, *kitāb* and *kitābat* are both feminine nouns, which they are not in Persian. The locative noun *maktab* lit. 'elementary school' is now mostly used in a figurative sense, for instance in the usual phrase *maktab e xeāl* 'school of thought', with a typical Persian *izafet* link between the two Arabic loanwords; this compound is classified as a masculine noun, according to the Urdu gender of its last component *xeāl*. The use of the Arabic plural form of *kitāb, kutub*, is considered pedantic, except in the official terminology for 'library', *kutub xānā* (with *xānā* as a Persian locative suffix). Again, *kutub xānā* is treated as a masculine noun in Urdu because of its ending.

An interesting case of Urdu lexical derivation from an Arabic loanword is the series derived from the word *badal*, which in Urdu means 'change'. First of all, the root-lexeme *badal* occurs not as an independent (masculine) noun but instead within a compound transitive verb, *badal karnā* 'to change' lit. 'to make change'. This type of derivation noun + *karnā* is a frequent pattern for verbal derivation in Urdu and indeed quite a common one with Arabic loanwords (after → Persian, it seems, but not exclusively; see Table 4).

A parallel case is the use of Arabic nouns with the Urdu verb *honā* 'to be' to form intransitive verbs (Table 5); intransitive verbs may also be formed with the Urdu verb *-ānā* 'to come' (Table 6).

In some cases the transitive and intransitive verbs form a pair, e.g. *śurū honā* 'to begin [intrans.]'/*kā śurū karnā* 'to begin [trans.]'; *xatam honā* 'to finish [intrans.]'/*xatam karnā* 'to finish [trans.]'.

Table 4. Frequent transitive compound verbs: Arabic noun/adjective + *karnā* 'to do'

Urdu compound	gloss	Arabic noun/adjective
madad karnā	'to help'	*madad* 'help, aid'
intazār karnā	'to wait'	*intiḏ̣ār* 'waiting, expectation'
śāmil karnā	'to include'	*šāmil* 'comprehensive'
tawajjū karnā	'to pay attention'	*tawajjuh* 'attention'
itirāf karnā	'to confess'	*iʿtirāf* 'confession'
qabūl karnā	'to accept' (formal)	*qabūl* 'acceptance'
jama karnā	'to add'	*jamʿ* 'addition'
kisī se muhabbat karnā	'to love somebody'	*maḥabba* 'love'
taslīm karnā	'to acknowledge'	*taslīm* 'surrender; greeting; concession; recognition'
tamir karnā	'to build [a monument]'	*taʿmīr* 'construction'
kā śurū karnā	'to begin something'	*šurūʿ* 'beginning'
kā faislā karnā	'to make a decision to'	*fayṣal* 'decisive criterion'
manā karnā	'to forbid'	*manʿ* 'prevention'
bahas karnā	'to discuss'	*baḥṯ* 'search; discussion'
tabsarā karnā	'to evaluate [a situation]'	*tabṣira* 'instruction, information'
śayā karnā	'to publish, edit'	*šāʾiʿ* 'widespread, well-known'
istimāl karnā	'to use, utilize'	*istiʿmāl* 'use, employment'
hisāb karnā	'to do the accounts'	*ḥisāb* 'account'
xatam karnā	'to finish'	*xatm* 'conclusion, closure'

Table 5. Frequent intransitive compounds: Arabic noun/adjective + *honā* 'to be'

Urdu compound	gloss	Arabic noun
śurū honā	'to begin [intrans.]'	*šurūʿ* 'beginning'
xatam honā	'to finish [intrans.]'	*xatm* 'conclusion, closure'
ehsās honā	'to have a feeling, to feel'	*ʾiḥsās* 'feeling'
kī zarūrat honā	'to be in need of, to need'	*ḍarūra* 'need, necessity'

Table 6. Frequent intransitive compounds: Arabic noun/adjective + *ānā* 'to come'

Urdu compound	gloss	Arabic noun/adjective
nazar ānā	'to appear [to somebody]'	*naḏ̣ar* 'eyesight, vision'
xeāl ānā	'to have an idea'	*xiyāl* 'imagination'
xauf ānā	'to be frightened'	*xawf* 'fear'

Actually, *badal* is one of the rare instances of a simple Urdu verb (not a compound), borrowed from an Arabic root. There is an intransitive verb *badalnā* 'to change', *nā* being the infinitive suffix, and a secondary derivation *badlānā* for the corresponding transitive verb, meaning 'to transform; to transfer'. According to a common Hindustani pattern for deriving verbs from nouns, and transitive verbs from intransitive ones, such derived verbs may be called 'causative'. There is even a double causative verb (again a typical Hindi-Urdu lexical derivation), with the verb *badalwānā* 'to have somebody make a change'. Another common instance of deriving a simple verb from an Arabic root-lexeme is the transitive verb *dafnānā* 'to bury [a dead body]', derived from the word *dafan* 'burial'. But for this word, there is no serial lexical derivation, as in the case of *badal*. Indeed, the serial lexical derivation from *badal* is quite a rich one; up to eighteen different entries may be counted in the monolingual reference dictionary *Firoz ul Lughat* (Firoz ud Din ca. 1995). Among these entries, the more frequent ones are the verbal phrase *badal lenā* 'to exchange' lit. 'to take in exchange', *lenā* being a semi-auxiliary verb systematically used to form verbal phrases with a lexical feature signaling an action that benefits the subject; *badlā* (masc.) 'exchange, vengeance', *badlāʾī* (fem.) 'commission in money exchange', *tabdīl* 'changed; transferred', an invariable adjective (just like most adjectives derived from Arabic or Persian in Urdu); *tabdīl karnā* 'to transfer', a transitive compound verb; and, again from *tabdīl, tabdīlī* (fem.) 'transformation, transfer'. The phrases *adle kā badlā* 'tit for tat' and *adal badal* 'mutatis mutandis', 'a good substitute', both masculine compound nouns, are interesting Urdu popular derivations with typical echo words.

3. Semantic fields of loanwords and semantic adaptation

From the point of view of semantics, Arabic loanwords are found in almost any semantic field in Urdu, mostly through Indo-Persian influence, which pervaded the entire Indian society. In learned and technical terminology, the proportion of Arabic loanwords is higher, only rivaled by English neologisms.

The field of religious terminology – and not only Qurʾānic and Islamic terms, as exemplified by Hindustani and Urdu Christian texts produced by missionaries – is pervaded by Arabic lexemes. Such very common lexemes as *masjid* 'mosque' or *mullā* 'priest' are indeed familiar terms. Some religious terms are recent loans or words that were reintroduced under changing political and religious circumstances, for instance the legal term *zakāt* for the traditional Islamic tax in Pakistan. In the field of religious terminology, too, Persianized Arabic loanwords are frequent, for instance the word *musulmān* 'Muslim [noun and adjective]', perhaps derived from Arabic *muslim* + Persian generic or plural suffix *-ān; īdgāh* 'open mosque for Muslim festivals', derived from Arabic *ʿīd* + Persian suffix *gāh; qurān xā* [written *qurʾān xān*] 'person specialized in the reading of the *Qurʾān*', derived from Arabic *qurʾān* + Persian suffix *xān*. At the level of everyday speech, Persian terms are preferred to Arabic ones, e.g. *xudā* 'God', *rozā* 'ritual daily fast', *namāz* 'ritual prayer' (the equivalents *allāh, saum*, and *salāt* are used only in a learned context); *saum o salāt* 'fast and prayers', used to denote the most common religious obligations for Muslims, are formed with the Persian conjunction *o* 'and', but this is a fixed expression.

Present-day Urdu onomastics and toponymy (i.e. the Indo-Persian toponymy of Urdu-speaking areas) could reveal a wealth of Arabic terminology – mostly related to the religious field and to religious titles – but it is sometimes obscured by modern toponyms. Both onomastics and toponymy are at the fringe of Urdu lexicology and sociohistorical heritage, and are beyond the scope of the present entry.

Semantic need is a crucial factor in adopting Arabic loanwords. They found their way into Hindustani and Urdu often by filling gaps in semantic fields created by changes in the referent or by the emergence of new referents, for instance through Arab navigation, which introduced new terms that coexisted with old Indian words, or sometimes even replaced them. Thus, the modern term for 'navy' is *bahariā*, along with the still-current Classical Indian word *nāo* 'ship'. The Urdu masculine noun *jahāz*, borrowed from Arabic, must have been introduced as a substitute of *nāo* for a new kind of ship, but then it adapted to the semantic shift that took place (after the English pattern) from navigation to aviation. Modern *havāʾī jahāz* was created to fit the English noun 'aircraft'. Given the importance of aviation in the subcontinent as well as in an emigration context, the Urdu masculine noun *havāʾī jahāz* is commonly abbreviated into *jahāz* to fit the popular English term 'plane'.

Generally speaking, the adjective *havāʾī* 'aerial' (< Arabic *hawāʾ* 'air') followed the English prefixed noun 'air', thus 'airport' gave the masculine Urdu noun *havāʾī aḍḍā*, a rare instance of a compound noun formed with an Arabic prefixed adjective and a North Indian masculine noun with a typical retroflex consonant *ḍ*.

The common word for 'ocean' is still *samandar*, the Indian word derived from Sanskrit, while in the lexical field of geography, as in Persian navigation documents, a given ocean is called *bahar* (< Arabic *baḥr* 'sea'). This brings about a clear lexical opposition between the popular concept and the learned one, the latter appearing in textbooks.

Arabic loanwords are very frequent in some lexical fields, mostly through Indo-Persian influence, which pervaded the whole Indian society. In learned and technical terminology, the proportion of Arabic loanwords increases even more, only to be challenged by English neologisms. One very basic and restricted lexical field linked with daily life experience is that of ‘house’. Such a field may be regarded as a more or less receptive or open semantic structure: the concept of ‘house’, though quite universal, may vary greatly according to social and historical context or referent. Therefore, loanwords or neologisms are welcome for filling the gaps. Indeed, in modern Urdu, the basic lexeme *ghar* ‘house; home’ (a North Indian word with aspirated /gh/) is now opposed to the Arabic loanword *makān* ‘place’, a locative masculine noun that in Urdu indicates a ‘house’ or ‘cottage’. This lexeme is used for an isolated house or building, rather than a ‘home’ or *ghar*, which may be a unit in an apartment building as well as a single house. From the administrative point of view, the equivalent of the English ‘residence’ is the Persian noun *rihāiś* with an *-iš* ending that classifies it as a feminine noun in Urdu. If we go for the ‘rich’ and ‘historical’ distinctive features of the ‘house’ lexical field, we find *mahal* ‘palace’ (< Arabic *maḥall* ‘place, site, location’, a locative noun that is masculine in Urdu). Within this field, *mahal* is frequently opposed to the ‘poor’ and ‘rural’ features of *jhõpṛī* ‘tumbledown or dilapidated house’ (with nasal *o* and retroflex *ṛ*).

Closely related to *mahal* is the lexeme *qilā*, which combines two semes, ‘fort’ and ‘palace’, as well as the distinctive feature ‘historical’. Also related to *mahal* is the term *havelī*, belonging to the same root (although the derivation is not motivated by the Arabic root). In Urdu, *havelī*, a feminine noun, carries the meanings of ‘mansion or noble house’.

A large building, a block of apartments, or a monument may be referred to by the word *imārat* (< Arabic *ʿimāra* ‘building’), in which the initial /ʿ/ has been deleted; this word combines at least two semes or distinctive lexical features: ‘large building’ and ‘monument’, to which a ‘historical’ seme may be added. *Imārat* is classified as a feminine noun in Urdu because of its ending *-at*; the literary (Arabic-styled) plural *imārāt* is not a rare occurrence and competes with the Indian plural *imārāten.*

The usual term for ‘floor [of a building]’ would be *manzil*, and a common phrase such as ‘first floor’ would be *pahlī manzil. Manzil* is a frequent word in Urdu, mentioned above as an Arabic loanword with the broader meanings of ‘stage, goal, destination’, but classified as a feminine noun in Urdu (a gender perhaps strengthened by the English ‘story’ if we overlook the complicated set of rules laid down by Urdu lexicographers and grammarians for Arabic loanwords). The Urdu lexeme *manzil* is certainly a very old South Asian borrowing from the Arabo-Persian stock (it has retained in modern Urdu the Qurʾānic meaning), but when referring to this particular lexical field with a ‘story’ seme as a distinctive feature, given the relatively recent referent in urban architecture, it may be rated as a neologism. Indeed, the common English phrase, ‘two-story house’ is calqued by the Urdu equivalent *do manzilā ghar* (with a Persian derivation, attested by Platts, from 1884).

The entire perfectly integrated lexical series that is motivated by the seme ‘house’ counts nine terms, none of them belonging to a literary level of language, including six Arabic loanwords (through Indo-Persian), plus one Persian loanword, which is quite a remarkable proportion.

The lexical field of ‘politics’ provides an even more striking example of significant Arabic neologisms in modern Urdu. Here, we find the same borrowing pattern as signaled before (following the Indo-Persian model, but also a persisting English model because of the colonial and postcolonial heritage). For historical and political reasons, Arabic loanwords like *vazīr* ‘minister’ or *sultān* ‘sultan’ are present as in so many other languages, but the word *hukūmat* (< Arabic *ḥukūma*) was connected with the meaning of ‘authority’ before it became an equivalent of the English ‘government’. After the 1947 Partition, *hukūmat* became the official designation of the Pakistani government, while the former (Indo-Persian) *sarkār*, which was the official designation of the British colonial government, was kept by the Indian Federation. The adjective *sarkārī* ‘governmental’ is still in use in both countries, but Pakistani official terminology also has *hukūmatī*, and, at the popular level, the lexeme *sarkār* (a feminine word in Urdu, like *hukūmat*) is still used, sometimes with the meaning of 'boss'. The political British heritage is such that the English loanword 'party' (i.e. 'political party'), borrowed as *pārṭī* (a feminine noun with a retroflex /ṭ/), has never been effectively replaced by the rarely used Arabic loanword *hizb* (< Arabic *ḥizb*, a masculine noun in Urdu), usually found in the phrases *hizb e ixtilāf* 'opposition, opposition party' and *hizb e iqtidār* 'party in power, government party', compound masculine nouns built with Arabic lexemes and Persian *izafet e*.

No hasty conclusions should be drawn on the basis of these samples of Urdu lexicology permeated by Arabic loanwords. It would be easy to find other similar instances where Arabic loanwords are dominant, but one should remember that the Indo-Persian channel is the key factor, and in common modern speech, actual occurrences of relatively recent Arabic loanwords may still be challenged by their English (or Anglo-Indian equivalent). In monolingual dictionaries, for instance, one may still find *floor* instead of *manzil*; in a connected field, the use of the English term *chambers* (in the plural, in its commercial context) is still frequently preferred to its Arabic equivalent *daftar*. Nevertheless, another conclusion is that Arabic neologisms in Urdu, whether they come through Indo-Persian or directly through Modern Arabic, as a calque from English, or because of a political and religious referent, are part of the very nature of the Urdu lexical system, inherited from Hindustani. Because of its constant evolution, because of easier communications with the Arab and Muslim world, and because Persian influence has not disappeared completely, it is difficult to give precise statistics, but one may safely assume that Arabic loanwords and Arabic neologisms will definitely remain an intrinsic part of the Urdu language, as will the Arabo-Persian literary models and the *nastaʿlīq* script.

BIBLIOGRAPHICAL REFERENCES

Désoulières, Alain. 1995. "Le rôle de l'ourdou dans le sentiment d'identité nationale pakistanaise, de 1947 à 1971 et de 1972 à nos jours". *Langues et pouvoir: Bicentenaire de l'INALCO (1795–1995)*, ed. Salem Chaker, 227–257. Aix-en-Provence: Edisud. [Deals with the consequences of government policies on Urdu lexicology and lexicography.]

——. 2003. "Le projet linguistique libéral et colonial de Fallon". *La langue libérée*, ed. Anne-Marie Laurian, 29–50. Bern: P. Lang. [A case of colonial lexicography, an attempt to reform Hindustani by suppressing loanwords.]

——. (with Najma Khan). 2005. *Petit lexique ourdou français*. Paris: INALCO. [A basic modern Urdu vocabulary in the form of a working paper and manual in a provisional edition, containing 3,000 Urdu lexemes and signaling Arabic and Persian loanwords.]

Firoz ud Din, Al Haj Maulavi. Ca. 1995. *Firoz ul lughat*. Lahore: Firozsanz. (New ed., 2005.) [Monolingual and etymological modern Urdu dictionary, with more than 150,000 words.]

Khaweshgi, Muhammad Abdullah Khan. 1937. *Farhang e Amra*. Islamabad. (2nd ed., Islamabad, 1989.) [Dictionary of Arabic and Persian loanwords in Urdu.]

Platts, John T. 1884. *A dictionary of Urdu, Classical Hindi, and English*. Oxford: Oxford University Press.

Qureshi, Bashir Ahmad. Ca. 1980. *Kitabistan's 20th-century standard dictionary Urdu into English*. 2nd ed. Lahore and Delhi: Kitabistan Publishing Company. [50,000 words.]

Southworth, Franklin C. 1971. *Hindi-Urdu reference manual*. Tucson, Ariz.: University of Arizona Press.

Alain Désoulières
(INALCO Urdu/South Asia and CERLOM)

Uyghur

1. Arabic-Uyghur relations

Uyghur (usually called Modern Uyghur, to distinguish it from the Old Uyghur language which was once spoken in roughly the same area) belongs to the Chaghatay (Southeast) subgroup of the Turkic language family. It is one of the closest relatives of the dialect that provided the basis for the Chaghatay lingua franca and literary language. It is the literary language used by the speakers of a diverse group of dialects in eastern Turkistan. Most of the estimated ten million native speakers reside in the Xinjiang Uyghur Autonomous Region in the People's Republic of China and in the Ferghana Valley west of it.

Although some of the Arabic words used in eastern Turkistan may have been adopted more than a thousand years ago, none of them can be demonstrated to have been borrowed directly from native speakers of Arabic.

The Arabic loanwords in Uyghur can be divided roughly into two groups. The first group entered through the mediation of Persian and/or via neighboring Turkic languages. These are mostly common everyday words known by all, often appearing in a form close to → Persian. Many phonetic and semantic deviations from Arabic are actually of Persian rather than Uyghur origin.

The second group appears to have entered through the literary language, as well as through religious schools and tractates. These words mostly belong to elevated language and are not necessarily known by all speakers. They have remained closer to Arabic in meaning and pronunciation.

2. Arabic Uyghur script

The present version of the Uyghur Arabic script (in use since 1983) is based on phonetic principles (see Table 1).

Table 1. Uyghur Arabic characters

The 'usual' letters: *a, b, t, j, x, d, r, z, ž, s, š, ʿ, ġ, f, q, k, l, m, n*

Persian letters: *g, ŋ, p, č*

Abolished characters from the Arabic alphabet: *ʾ, ṯ, ḥ, ḏ, ṣ, ḍ, ṭ, ḏ̣*

Extra letters are the following:

و	o	
ۆ	ö	
ۇ	u	
ۈ	ü	
ۋ	w	
ى	i	(undotted in all positions)
ې	e (é)	(always with vertical dots)
ي	y	(always with horizontal dots)
ە	ä	(never ھ)
ھ	h	(never ە)
ئ	[']	(not considered a consonant in Uyghur; all initial vowels are preceded by the '*hamza* carrier' rather than by *ʾalif*)

Examples

ئەرەپ	*äräp* 'Arab'	مۇسۇلمان	*musulman* 'Muslim'
ئاي	*ay* 'month'	ۋەتەن	*wätän* 'homeland'
ھەرپ	*härp* 'letter'	ئەنئەنە	*änʾänä* 'tradition'
تېلېفون	*telefon* 'phone'	دىن	*din* 'religion'
ئاپتوۋۇز	*aptowuz* 'bus'	ئۆيلەردە	*öylärdä* 'in the houses'

All eight Uyghur vowels have their own grapheme, and all vowels are written. Diacritic signs are used to distinguish *e, y*, and *i* (the latter always written undotted). Plain *wāw* stands for *o*, whereas *w, ö, u*, and *ü* are each marked with a different diacritic. *Hāʾ* was split into two separate graphemes for *h* and *ä*. Initial vowels are preceded by a '*hamza* carrier' in writing. The number of consonant letters has been reduced because many distinctions that are relevant for Arabic are not pronounced in Uyghur (see below). This spelling system is also applied to Arabic loanwords.

3. Phonological adaptation of the vowels

In view of the divergent phoneme inventories of Arabic and Uyghur and the syllable structures they permit, it is not surprising that many Arabic words were adapted in the borrowing process. Uyghur vowel harmony only permits certain vowel patterns in a single word, while Uyghur syllable structure typically alternates vowels and consonants. In the following, it has to be kept in mind that many phonetic changes in Uyghur are in fact of Persian origin.

The long Arabic vowels *ā, ī, ū* usually retained their quality as *a, i, u*, e.g. *xal* (< *xāl*) 'birthmark', *din* (< *dīn*) 'religion', *nur* (< *nūr*) 'light'. In many words, these vowels have remained long in Uyghur. The distinction between long and short vowels cannot be indicated in Uyghur Arabic spelling, but they may be indicated in scholarly publications (see Hashim and Mikhri 1986).

The short Arabic vowels display a wider variety of counterparts in Uyghur. Short *a* mostly appears as *a* or *ä*, as in *mäšq* (< *mašq*) 'exercise', *haraq* (< *ʿaraq*) 'brandy'; short *i* mostly appears as *i* or *e*, as in *išq* (< *ʿišq*) 'love', *elan* (< *ʾiʿlān*) 'announcement'; short *u* may appear as *o, ö, u, ü*, as in *örp* (< *ʿurf*) 'tradition', *nopus* (< *nufūs*) 'population'. The choice of vowel is not entirely predictable; it mostly depends on the phonetic environment, which can palatalize or velarize vowels.

Short *a* in the first syllable undergoes so-called umlauting before *ī*, as in *bexil* (< *baxīl*) 'stingy', *heqiq* (< *ʿaqīq*) 'agate', *peqir* (< *faqīr*) 'poor', *xemir* (< *xamīr*) 'dough'. Long *ā* is not affected in this way, e.g. *tarix* (< *tārīx*) 'history', *hami* (< *ḥāmī*) 'protector'.

Secondary vowel rounding can be caused by the following vowel, as in *qobul* (< *qabūl*) 'acceptance', *zörüriy* (< *ḍarūrī*) 'requisite', or by adjacent consonants, as in *süpät* (< *ṣifat*) 'quality', *muaš* (< *maʿāš*) 'wages'.

The typically Uyghur phenomenon of 'vowel raising' in middle syllables has also affected the short *a* in Arabic words, e.g. *mäsilä* (< *masʾala*) 'matter', *ölima* (< *ʿulamāʾ*) 'theologians'. In some words, the raised middle vowel is subsequently elided (often in the vicinity of *r*), e.g. *härkät* (< *ḥarakat*) 'movement', *harpa* ~ *ärpä* (< *ʿarafa*) 'eve', *puqra* (< *fuqarāʾ* 'poor [pl.]') 'crowd'.

Consonant sequences not allowed in Uyghur are broken up by inserting a high vowel *i*, *u*, or *ü*, as in *äsir* (< *ʿaṣr*) 'epoch', *kupur* (< *kufr*) 'unbelief', *hököm* (< *ḥukm*) 'command'.

Diphthongs and other vowel sequences, normally not allowed in Turkic, arose in words that dropped an intervocalic ʾ or ʿ, e.g. *mäišät* (< *maʿīšat*) 'livelihood', *zäip* (< *ḍaʿīf*) 'weak', *saät* (< *sāʿat*) 'hour'.

4. PHONOLOGICAL ADAPTATION OF THE CONSONANTS

Many Arabic consonants were left more or less unchanged, or at least they systematically correspond to the same Uyghur consonants. This applies to *d, t, h, x, ġ, q, s, š, j, z, k, l, r, n, m*.

The 'emphatic' consonants generally merged with 'nonemphatic' counterparts, although traces of this opposition survive in the treatment of the adjacent vowels. Usually these consonants developed as in Persian.

H and *ḥ* merged into Uyghur *h*, e.g. *himmät* (< *himmat*) 'care', *hisap* (< *ḥisāb*) 'counting'. Final *-h/-ḥ* was generally dropped from pronunciation and spelling, as in *alla* (< *allāh*) 'God', *nika* (< *nikāḥ*) 'marriage', *täswi* (< *tasbīḥ*) 'prayer beads', although there is now a tendency to restore the *h* in spelling, by means of the Uyghur letter *h*. Preconsonantal *h/ḥ* is often spelled *h* but pronounced *x*, as in *mähsulat* (< *maḥṣūlāt*) 'products', *rähmät* (< *raḥmat* 'mercy') 'thanks'.

The 'alien' consonant ʿ also became *h* in a handful of words, such as *hamma* (< *ʿamma*) 'aunt', *hasa* (< *ʿaṣā*) 'cane', *häsäl* (< *ʿasal*) 'honey', but merged with vocalic onset in most cases. In the middle of words, it was either lost, e.g. *äpi* (< *ʾafʿā*) 'species of snake', *mämur* (< *maʿmūr*) 'official', *möjizä* (< *muʿjiza*) 'miracle', *wädä* (< *waʿda*) 'promise', or represented by vocalic onset, e.g. *qitʾä* (< *qiṭʿa* 'piece') 'continent', *sünʾi* (< *ṣunʿī*) 'artificial'. Final ʿ is normally lost, as in *näp* (< *nafʿ*) 'profit', *mawzu* (< *mawḍūʿ*) 'subject', *täläy* (< *ṭāliʿ*) 'fortune'.

Initial ʾ is not considered a consonant in Uyghur. Medial ʾ is usually lost, as in *mömin* (< *muʾmin*) 'believer', but before *i* it may change into a *y* glide, as in *mulayim* (< *mulāʾīm*) 'mild'.

Arabic *x* generally remained *x*, although alternations *x* ~ *q* do occur, as in *zix* ~ *ziq* (< *sīx*) 'skewer, spit'.

S, *ṣ*, and *ṯ* merged into *s*, e.g. *säpär* (< *safar* 'journey') 'occurrence', *sada* (< *ṣadā*) 'voice', *miras* (< *mīrāṯ*) 'inheritance'. Similarly, *z*, *ḏ̣* (*ẓ*), *ḍ*, and *ḏ* merged into *z*, e.g. *zukam* (< *zukām*) 'cold [disease]', *näzär* (< *naḏ̣ar*) 'view', *ziyapät* (< *ḍiyāfat*) 'feast', *zäki* (< *ḏakiyy*) 'clever'. These mergers all reflect the situation in Persian.

T and *ṭ* merged, e.g. *töhpä* (< *tuḥfa*) 'contribution', *talaq* (< *ṭalāq*) 'divorce'.

Arabic *f* is consistently represented by *p*, e.g. *pikir* (< *fikr*) 'thought', *peil* (< *fiʿl*) 'verb', *sähipä* (< *ṣaḥīfa*) 'page'.

Voiced plosive *b* became voiceless in final position, which is reflected in spelling, as in *kitap* (< *kitāb*) 'book', although 'archaizing' spellings with *-b* seem to become more popular. Devoicing in nonfinal positions is rare, as in *ispat* (< *ʾiṯbāt*) 'proof'. Intervocalic *b* has often become *w*, e.g. *nöwät* (< *nawbat*) 'turn', *xäwär* (< *xabar*) 'message'. Note that final *-b* can become intervocalic due to suffixation, as in *säwäp* (< *sabab*) 'reason' ⇒ *säwiwi* 'his/her reason'.

Arabic *aw* was often contracted into *o* or *ö*, e.g. *dölät* (< *dawlat*) 'state', *ronaq* (< *rawnaq*) 'splendor'.

Voiced plosive *d* generally became voiceless in final position, e.g. *heyt* (< *ʿīd*) 'religious holiday', but rarely elsewhere in the word, as in *mästär* (< *maṣdar*) 'gerund'.

Word-final consonant clusters not tolerated in Uyghur were broken up by epenthetic vowels, e.g. *äqil* (< *ʿaql*) 'mind', *hösün* (< *ḥusn*) 'beauty', *isim* (< *ism*) 'name', *jinis* (< *jins*) 'gender', *qulup* (< *qufl*) 'lock'. The epenthetic vowel, which is always *i/u/ü* due to Uyghur phonological constraints, is dropped when such words receive suffixes beginning with a vowel, such as the possessive suffixes, e.g. *äqli*, *hösni*, *ismi* 'his/her mind, beauty, name'. Sonorant + obstruent clusters are acceptable in Turkic and not necessarily simplified, e.g. *xälq* (< *xalq*) 'people', *härp* (< *ḥarf*) 'letter [character]'.

Geminates are usually preserved between vowels, as in *minnät* (< *minnat*) 'obligation', *ippät* (< *ʿiffat*) 'chastity'; in final position, they are simplified, as in *šäk* (< *šakk*) 'doubt', *sir* (< *sirr*) 'secret', *zit* (< *ḍidd*) 'opposite', but, like consonant clusters, they may reappear when suffixes are added, e.g. *häq* (< *ḥaqq*) 'right' ⇒ *häqq-i* 'his right', *räb* (< *rabb*) 'lord' ⇒ *ya räbb-im* 'o my Lord'.

In some words, Uyghur features unexpected geminates not found in Arabic, as in *qimmät* (< *qīmat*) 'expensive', *zinnät* (< *zīnat*) 'ornament', *addiy* (< *ʿādī*) 'ordinary'.

Some words with more eroded variants include *ištiha* ~ *ištäy* (< *ištihāʾ*) 'appetite', *märtiwä* ~ *märtä* (< *martaba* 'degree, rank') 'times', *išarät* ~ *šärät* (< *ʾišārat*) 'sign', *waqit* ~ *wax* (< *waqt*) 'time', *mähäl* ~ *mäl* (< *maḥall*) 'place', *mäpä* (< *maḥāfa*) 'cart'. There are many other deviating forms, not all of which are reflected in writing.

5. INCORPORATION AND INFLECTION OF ARABIC WORDS

Nouns, adjectives, adverbs, and conjunctions were simply borrowed as lexical items without morphological problems. In case of morphological variants within Arabic, Uyghur usually features one variant only, as in Persian. An example of this is the treatment of *tāʾ marbūṭa*, which in Uyghur is either present in all inflected forms, as in *pursät* (< *furṣat*) 'opportunity', *niyät* (< *niyyat*) 'intention', or absent throughout, as in *hikayä* (< *ḥikāyat*) 'story', *hädiyä* (< *hadiyyat*) 'gift' (for the distribution of these variants in Turkic languages and the relationship with the distribution of the feminine ending in Persian, see Perry 1991:178–180, 183–188).

As in other Turkic languages, Arabic verbs could not be borrowed directly, due to the different verbal systems. Instead, Uyghur creates new verbs by combining an Arabic noun or verbal noun with an Uyghur auxiliary verb like *qil-* 'to do' or *bol-* 'to become', e.g. *wapat bol-* (< *wafāt* 'death') 'to die', *täšäkkür eyt-* (< *tašakkur* 'giving thanks') 'to thank', *pärq qil-* (< *farq* 'distinction') 'to distinguish', *jawap bär-* (*jawāb* 'answer [noun]') 'to answer'. Alternatively, a verbalizer suffix is used, as in *šuġullan-* (< *šuġl* 'occupation') 'to occupy oneself with', *rawajlandur-* (< *rawāj* 'circulation, spreading') 'to develop', *zamaniwilaštur-* (< *zamān* 'time') 'to modernize'. From these Uyghur stems all verb forms can be derived.

Etymological hybrids abound, e.g. Persian/Arabic/Turkic *na-häq-liq* 'injustice', with prefixed Persian negative and Uyghur suffix.

Nominal inflection, such as case endings and possessive suffixes, may involve phonetic alterations of the stem, depending on the original vowel and consonant, as in *kitap* ⇒ *kitaw-im* 'my book' (with original *b* and long *ā*), as opposed to *täräp* ⇒ *tärip-im* 'my side' (with original *f* and short *a*). As mentioned above, Arabic long *ā* resists the Uyghur umlauting rules, e.g. *hal-i* 'his situation' as compared to native *baš* 'head' ⇒ *beš-i* 'his head'.

Most Arabic words follow the normal vowel harmony rules when suffixes are added. Monosyllables with the vowel *i* (or *e*) receive back-vocalic suffixes, e.g. plurals *pil-lar* 'elephants', *heyt-lar* 'holidays', and verbalizers such as *zit-laš-* (< *ḍidd* 'opposite') 'to oppose'.

Uyghur created a number of additional suffix variants to accommodate Arabic words, e.g. the dative variants *-ġä, -qä*, as in *xälq-qä* 'to the people', *qälb-i-ġä* 'to his heart' (*ä* would not co-occur with *q* or *ġ* in native words). There are also newly created possessive forms to add to stems ending in a long vowel, usually due to loss of a final consonant, e.g. *ġiza-yiŋ/ġiza-yi/ġiza-yimiz* (< *ġiḏāʾ*) 'your/her/our food', cf. the short vowel forms *harwa-ŋ/harwi-si/harwi-miz* (< *ʿaraba*) 'your/her/our cart'.

As demonstrated above, consonants and consonant sequences which are unpronounceable to Uyghurs and simplified or changed in final position, may resurface when suffixes are attached.

Although many Arabic morphological processes are reflected in Uyghur, none of these are actively used by the Uyghurs to make new forms, either on borrowed stems or on native stems.

Arabic plurals were apparently borrowed as separate lexical entities and are often used with a slight semantic difference, as in the case of *xäwär* 'message' and *axbarat* 'news', *xälq* 'people' and *xalayiq* 'the masses'. Most Arabic words now receive Uyghur plurals.

Only a few Arabic feminine forms are still in use, e.g. *muällimä* (< *muʿallima*) 'female teacher’, *mäbudä* (< *maʿbūda*) ‘goddess’, *räpiqä* (< *rafīqa*) ‘wife’. Adjectives can be gender specific in some formulas, like *märhum atam* ‘my late father’, *märhumä anam* ‘my late mother’ (< *marḥūm*).

Even Arabic endings that occur frequently in Uyghur, such as *-i* ~ *-iy* ~ *-(i)wi* in *omumi* ‘general’, *ġärbiy* ‘western’, *dunyawi* ‘worldly’, and the adverbial ending *-än* in *mäsilän* ‘for example’, *häqiqätän* ‘indeed’, etc., cannot be attached to native words.

Many formulas and phrases were adopted as a whole, e.g. *ästaġpurulla* (< *ʾastaġfiru llāh*) ‘I ask God for forgiveness’, *hašqalla* (< *ʿišq allāh*) ‘thank you’, *wahakaza* (< *wa-hākaḏā*) ‘et cetera’. Most compounds are just felt to be single lexemes, e.g. *elipbä* (< *ʾalīf bāʾ*) ‘alphabet’, *äksilhärkät* (< *ʿaks* + *ḥarakat*) ‘countermovement’, *hayat-mamat küriši* (< *ḥayāt* + *mamāt*) ‘life-and-death struggle’, *häbbulmülük* (< *ḥabb al-mulk*) [plant name], *bäynälmiläl* (< *bayn al-milal*) ‘international’, *bäytulla* (< *bayt allāh*) ‘house of God [i.e. place of worship]’, *darilpunun* (< *dār al-funūn*) ‘academy’. Many instances are in fact Persian coinages based on Arabic stems: *äpuomumi* (< Persian *afv-e umumī* < Arabic *ʿafw* ‘pardon’ + *ʿumūmī* ‘general’) ‘amnesty, general pardon’, *nuqtiinäzär* (< Persian *nuqṭa-ye naẓar* < Arabic *nuqṭa* ‘point’ + *naḏar* ‘view’) ‘point of view’, etc.

A popular feature of the Uyghur lexicon are doublets of (nearly) synonymous words, especially rhyming ones, or derivates from the same root, e.g. *örp-adät* (< *ʿurf* + *ʿādat*) ‘customs and traditions’, *zoq-šoq* (< *ḏawq* + *šawq*) ‘enjoyment’, *ġälwä-ġoġa* (< *ġalaba* + *ġawġa*) ‘tumult’, *ġerip-ġurwät* (< *ġarīb* + *ġurbat*) ‘alone and deserted’, many of which are Persian creations. Arabic/Uyghur hybrids include *ilim-bilim* (< *ʿilm*) ‘knowledge’, *soal-soraq* (< *suʾāl*) ‘questions’, *histuyġu* (*ḥiss*) ‘feeling’, and *uruq-äwlat* (< *ʾawlād*) ‘offspring’.

6. SEMANTIC AND FUNCTIONAL FIELDS OF BORROWING

Nearly all categories of the lexicon are represented. First names are predominantly of Arabic and Persian origin. However, the core lexicon, such as numerals, basic adjectives, body parts, family members, names of familiar animals, clothes, and household utensils, is not affected as much. Only a few names for basic body parts were borrowed, such as *mäydä* (< *miʿda*) ‘chest; stomach’ and *hälqum* (< *ḥulqūm*) ‘throat’. Among animal names we find exotic species, such as *timsax* (< *timsāḥ*) ‘crocodile’, *maymun* (< *maymūn*) ‘monkey’, *pil* (< *fīl*) ‘elephant’.

Common religious terms may be more eroded, such as *mäzin* (< *muʾaḏḏin*) ‘muezzin’, *mečit* (< *masjid*) ‘mosque’. Less common terms, and elevated vocabulary not known to all speakers, include *hazazul* (< *ʿazāzīl*) ‘Azazil [epithet of the devil]’, *mäqbärä* (< *maqbara*) ‘cemetery’, and *munäjjim* (< *munajjim*) ‘astronomer’. Some words may have been adopted for euphemistic reasons, e.g. *hajätxana* (< *ḥājat* ‘need’ + Persian *xāna*) ‘toilet’, and *tanasil äza*, based on *ʾaʿḍāʾ tanāsuliyya* ‘genitals’.

Uyghur conjunctions are usually of Arabic origin: *päqät* (< *faqaṭ* ‘only’), *ämma* (< *ʾammā*), *lekin* (< *lākin*), all meaning ‘but, however’; *wä* (< *wa*) ‘and’; and many modal adverbs: *hätta* (< *ḥattā*) ‘even’, *ġayät* (< *ġāyat* ‘utmost limit, goal’) ‘very’. Interestingly, some Arabic words have become postpositions in Uyghur, e.g. *häqqidä* (derived from *ḥaqq* ‘right [noun]’) ‘about’, *täripidin* (from *ṭaraf* ‘edge, side’) ‘by [with passive verbs]’.

A number of ultimately non-Arabic loanwords, mostly *Wanderworte* from Greek, have been borrowed in their Arabicized shape, such as *bälġäm* (< *balġam*) ‘phlegm’, *qanun* (< *qānūn*) ‘law’, *iqlim* (< *ʾiqlīm*) ‘climate’, *toz* (< *ṭāwūs*) ‘peacock’, *panus* (< *fānūs*) ‘lantern’.

As expected, the meanings of the borrowed lexemes may be different from or narrower than the Arabic source words; for instance, Uyghur *ali* (< *ʿālī*) means only ‘high’ in an abstract sense, *ayal* means ‘wife’ (< *ʿiyāl* ‘family members’), *xät* means ‘letter [epistle]’ (< *xaṭṭ* ‘line, etc.’), *mäqät* means ‘anus’ (< *maqʿad* ‘sitting’), *iplas* means ‘dirty’ (< *ʾiflās* ‘bankrupt’), and *usul* means ‘dance’ (< *ʾuṣūl* ‘principles, basic rules’).

BIBLIOGRAPHICAL REFERENCES

Hahn, Reinhard F. 1991. *Spoken Uyghur*. Seattle and London: University of Washington Press.

——. 1998. “Uyghur”. *The Turkic Languages*, ed. Lars Johanson and Éva Csátó, 379–396. London and New York: Routledge.

Hashim and Mikhri. 1986. “Weiwuer konyu lide zhang duan yuanyin. [Long and short vowels in spoken Uyghur]”. *Minzu Yuwen* 3.39–43.

Jarring, Gunnar. 1964. *An Eastern Turki-English dialect dictionary*. Lund: CWK Gleerup.

Nadžip, Èmir N. 1968. *Ujgursko-russkij slovar'* [Uyghur-Russian dictionary]. Moscow: Izdatel'stvo "Sovetskaja Ènciklopedija".

Perry, John R. 1991. *Form and meaning in Persian vocabulary: The Arabic feminine ending.* Costa Mesa, Cal.: Mazda Publishers.

Pritsak, Omeljan. 1959. "Das Neuuigurische". *Philologiae Turcicae Fundamenta*, ed. Jean Deny a.o., 525–563. Wiesbaden: F. Steiner.

Hans Nugteren (Dordrecht)

Uzbek

1. Arabic-Uzbek relations

Uzbek (Ozbek, Özbek) is a Turkic language spoken by some 23 million people, most of whom live in Uzbekistan, while the remainder are distributed over neighboring countries, the largest groups being found in Afghanistan and Tajikistan.

Arabic influence on modern Uzbek is restricted to loanwords; the entirely different structures of Arabic and Uzbek seem to have prevented any enduring grammatical influence. The lexical influence is pervasive, especially in the field of abstract notions, as exemplified by the following sentence: *Insåniyat***niṇ** *madaniyat taraqqiyåt tarixi***da** ***ozbek*** *milla***ti** *ajåyib tuhfa***lar** ***qošgan*** 'In the history of the cultural development of humanity, the Uzbek people have made amazing contributions'. Apart from the ethnonym and the final verb, only the endings are of Turkic origin (native elements in bold).

Most Arabic loanwords in Uzbek seem to have entered through → Persian, as well as literary Chaghatay and other written languages. This is not surprising in view of the fact that many speakers of Uzbek are bilingual in varieties of Persian (→ Tajik and Dari). Borrowing directly from Central Asian Arabs cannot be demonstrated.

2. Writing

Linguistically, Uzbek has developed gradually from Chaghatay, which was written in Arabic script and did not consciously change the spelling of its Arabic loanwords. At the beginning of the Soviet period, the Arabic notation of Uzbek was abandoned in favor of a Latin-based orthography, which in its turn was replaced by the Cyrillic alphabet in 1940. After the collapse of the Soviet Union, it was decided that a new Latin alphabet would be introduced. The latter, which is steadily replacing the Cyrillic alphabet, although this is still widely used, is mostly a transliteration of its Cyrillic predecessor. After the replacement of the Arabic alphabet, orthographic distinctions that were not reflected in pronunciation were largely abolished. For instance, the new alphabets do not contain equivalents for Arabic letters such as *ḥ, ṣ, ṭ, ḏ̣* (*ẓ*). Instead, these were merged with their 'nonemphatic' counterparts. In other regards, modern spelling is conservative, in that it meticulously preserves many instances of unpronounced ʿ and ʾ, as well as foreign consonant clusters, which are simplified or broken up in spoken Uzbek.

3. Phonological adaptation of the vowels

Instead of the canonical Turkic vowel system with eight vowels, standard Uzbek has a simplified system with six vowels, transliterated here as *a, å, e, i, o, u* [*å* and *o* are now written *o* and *o'* respectively]. The rounded front vowels *ö* and *ü* can be found in other varieties of Uzbek. This is essentially the Persian (Tajik) vowel system. The six Arabic vowel phonemes, especially the short ones, do not correspond straightforwardly with the Uzbek vowels.

The long Arabic vowels *ā, ī, ū* are usually represented in Uzbek by *å, i, u*, e.g. *hål* (< *ḥāl*) 'state', *fil* (< *fīl*) 'elephant', *ruh* (< *rūḥ*) 'soul'. Long *ā* may also appear as *a*, as in *tarix* (< *tārīx*) 'history', and many variant forms are found, e.g. *åxir* ~ [colloquial] *axir* (< *ʾāxir*) 'end', *ġawġå* ~ [colloquial] *ġåwġa* (< *ġawġāʾ*) 'hubbub'. Long *ī* is occasionally represented by *iy* in Uzbek, as in *siymå* (< *sīmāʾ*) 'appearance', *qåriyb* (< *qarīb*) 'approximately'. When shortened, it may also appear as *e*, as in *merås* (< *mīrāṯ*) 'inheritance'.

The short vowel *a* typically becomes *a*, e.g. *dars* (< *dars*) 'lesson', *harf* (< *ḥarf*) 'letter', but *e* (perhaps due to the following long *ī*) appears in *metin* (< *matīn*) 'sturdy', *yetim* (< *yatīm*) 'orphan'.

Short *i* normally becomes *i*, as in *irq* (< *ʿirq*) 'race', *jism* (< *jism*) 'body', but *e* often appears before final ʿ (which itself was dropped), or before *h*/*ḥ*, e.g. *mawze* (< *mawḍiʿ*) 'place', *zåye*

(< *ḍāyiʿ*) ‘in vain’, *zehn* (< *ḏihn*) ‘intellect’, *sehr* (< *siḥr*) ‘magic’. However, in several dialects, including that of Tashkent, these words are pronounced *i* instead of the literary *e*.

Short *u* typically becomes *u*, as in *urf* (< *ʿurf*) ‘tradition’, *huquq* (< *ḥuqūq* ‘rights’) ‘justice’. Uzbek *o* is the only vowel that is rare in Arabic words, e.g. *tofån* (< *ṭūfān*) ‘hurricane’, *momin* (< *muʾmin*) ‘believer’. Irregular unrounding of *u* is common in nonfirst syllables, e.g. *tasådifan* (< *taṣādufan*) ‘coincidentally’, *tawåze* (< *tawāḍuʿ*) ‘politeness’.

Further unexpected vowel changes are found in the words *hibs* (< *ḥabs*) ‘prison’, *xislat* (< *xaṣlat*) ‘characteristic’, *zilzila* (< *zalzala*) ‘earthquake’, *mazax* ~ *mazah* (< *muzāḥ*) ‘joke’, *halqum* (< *ḥulqūm*) ‘throat’, *xartum* (< *xurṭūm*) ‘elephant’s trunk’. Most of these deviations are not restricted to Uzbek. Some longer words feature the elision of the middle vowel, e.g. *hadya* (< *hadiyyat*) ‘gift’.

The loss of intervocalic ʾ or ʿ has given rise to vowel sequences that are normally not allowed in Turkic, as in *inšååt (< ʾinšāʾāt)* ‘construction’, *mudåfaa* (< *mudāfaʿa*) ‘defense’, *tabiiy (< ṭabīʿī*) ‘natural’, *iåna* (< *ʾiʿānat*) ‘alms’, *såat* (< *sāʿat*) ‘hour’.

4. Phonological adaptation of the consonants

Many Arabic consonants were adopted by Uzbek more or less unchanged: *d, t, h, x, ġ, q, s, š, j, z, k, l, r, n, m*. The remaining consonants, alien to Uzbek, usually merged with existing consonants, in a fashion similar to Persian.

H and *ḥ* merged into Uzbek *h*, e.g. *haykal* (< *haykal*) ‘statue’, *hiyla* (< *ḥīla*) ‘ruse’, *qabih* (< *qabīḥ*) ‘bad’, *ablah* (< *ʾablah* ‘foolish’) ‘bad’, *tasbeh* (< *tasbīḥ*) ‘prayer beads’. Final *-h/-ḥ* was occasionally dropped, as in the name *ållå* (< *allāh*) ‘God’ and formulas including it.

Arabic *x* generally remained *x*, as in *xizmat* (< *xidmat*) ‘service’. In the spoken language, *x* and *q* are not always distinguished, as in *nusxa* ~ *nusqa* (< *nusxa*) ‘sample’ and, in reverse, *nuqta* ~ *nuxta* (< *nuqṭa*) ‘point’. In the Tashkent dialect, *h* tends to merge with *ḥ* before another consonant, in words like *rahmat* (< *raḥmat* ‘mercy’) ‘thanks’ and *mehnat* (*mihnat*) ‘work’.

T and *ṭ* merged into *t*, e.g. *tuhmat* (< *tuhmat*) ‘false accusation’, *talab* (< *ṭalab*) ‘demand’.

S, *ṣ*, and *ṯ* merged into *s*, e.g. *sayåhat* (< *siyāḥat*) ‘journey’, *sinf* (< *ṣinf*) ‘class’, *marsiya* (< *marṯiya*) ‘elegy’. Similarly, *z, ḏ, ḏ̣* (*ẓ*), and *ḍ* merged into *z*, e.g. *zahmat* (< *zaḥmat*) ‘effort’, *zavq* (< *ḏawq*) ‘pleasure’, *zulm* (< *ḏ̣ulm*) ‘tyranny’, *håwuz* (< *ḥawḍ*) ‘water basin’. Some words with *ḍ* display variant forms, e.g. *ġazab* ~ *ġadab* (< *ġaḍab*) ‘anger’.

Initial ʿ is represented by *h* in a handful of words, such as *aså* ~ *hassa* (< *ʿaṣā*) ’walking cane’, *haqiq* (< *ʿaqīq*) ‘agate’, but it is generally merged with vocalic onset (ʾ), as in *arafa* (< *ʿarafa*) ‘eve’, *asal* (< *ʿasal*) ‘honey’. Variant forms include *iyd* ~ *hayit* (< *ʿīd*) ‘religious holiday’. Intervocalic ʿ is often dropped from spelling and/or pronunciation, especially between vowels, as in *zaif* (< *ḍaʿīf*) ‘weak’. If at all pronounced, it is realized as vocalic onset. In word-final position, ʿ is dropped from spelling as well as pronunciation, as in *mawzu* (< *mawḍūʿ*) ‘subject’. In preconsonantal position, it is represented by the grapheme <’> (the hard sign ь in Cyrillic), whereas in pronunciation it is usually realized as lengthening of the preceding vowel, as in the spellings *malum* (< *maʿlum*) ‘known’, *ba’zi* (< *baʿḍ*) ‘some’, *še’r* (< *šiʿr*) ‘poem’, *fe’l* (< *fiʿl*) ‘deed’, *ta’m* ~ *tam* (< *ṭaʿm*) ‘taste’, which are in fact pronounced as [maːlum, baːzi, ʃeːr, feːl, taːm]. The spoken language may also display *ġ, x, h*, as in *na’l* ~ [colloquial] *naġal* (< *naʿl*) ‘hobnail’, *ne’mat* ~ [colloquial] *niġmat* (< *niʿmat*) ‘good deed’, *tåle* ~ *tålix* ~ *tålei* (< *ṭāliʿ*) ‘good luck’, *matå* ~ [colloquial] *matåh* (< *matāʿ*) ‘fabric’. Postconsonantal ʿ is usually pronounced as vocalic onset, as in *qit’a* (< *qiṭʿa* ‘piece’) ‘continent’.

Final ʾ is neither written nor pronounced, as in *sahrå* (< *ṣaḥrāʾ*) ‘steppe’. Postconsonantal ʾ is usually preserved in spelling but not always pronounced, as in *hay’at* (< *hayʾat*) ‘staff’, *juz’iy* (< *juzʾī*) ‘minor’. Between vowels, it is commonly dropped from both spelling and pronunciation, but occasionally it is replaced by *y*, as in *mulåyim* (< *mulāʾīm*) ‘mild’. In some words medial ʾ is dropped from the orthography, e.g. *imån* ~ *iymån* (< *ʾīmān*) ‘belief’.

Arabic *f* is normally represented by Uzbek *f*, but occasionally by *p*, as in *supa* (< *ṣuffa*) ‘pedestal’, *supra* ~ *sufra* (< *sufrat*) ‘floor covering’, *patir* (< *faṭīr*) ‘kind of unleavened bread’. In spoken varieties of Uzbek, the pronunciation *p* is quite common.

Arabic *b* is pronounced *p* in final position but remains written as *b*, as in *sabab* (< *sabab*) 'reason'. Intervocalic *b* occasionally became *w* in the spoken language, e.g. *arawa* (< *ʿaraba*) 'cart', *tabaqa* ~ [colloquial] *tawaqa* (< *ṭabaqa*) 'layer', *tabaq* ~ [colloquial] *tåwåq* (< *ṭabaq, ṭabāq*) 'dish', *xabar* ~ [colloquial] *xawar* (< *xabar*) 'news', *tawba* ~ *tåwwa* (< *tawba*) 'repentance'.

Unlike elsewhere in Turkic, the epenthetic vowels used to break up non-Turkic consonant groups are rarely reflected in the orthography, so that spellings like the following are common: *izn* (< *ʾiḏn*) 'permission', *ufq* ~ *ufuq* (< *ʾufq*) 'horizon', *matn* (< *matn*) 'text', *sidq* (< *ṣidq*) 'sincerity', *hajw* (< *hajw*) 'humor'. Spoken Uzbek does require epenthetic vowels in these sequences, in the form of a reduced *i* or *u*, depending on the preceding vowel, as in *ism* > *isĭm* (< *ism*) 'name', *šakl* > *šakĭl* (< *šakl*) 'form', but *hukm* > *hukŭm* (< *ḥukm*) 'judgment'. These nonwritten vowels disappear again if a suffix is attached that starts with a vowel, as in *ism-im* 'my name'. This has a parallel in the behavior of some native stems, e.g. *burun* 'nose' ⇒ *burn-im* 'my nose'.

Geminates are usually preserved between vowels, as in *tabassum* (< *tabassum*) 'smile [noun]', *zarra* (< *ḏarra*) 'particle', but may be simplified, as in *sihat* (< *ṣiḥḥat*) 'health'. In final position, they are simplified, as in *had* (< *ḥadd*) 'frontier', *his* (< *ḥiss*) 'feeling', but they usually resurface when suffixes are added, e.g. *hadd-i* 'its border', *hiss-i* 'her/his feeling'.

Some Uzbek words feature geminates not found in Arabic, e.g. *zakiy* ~ *zakki* (< *ḏakiyy*) 'intelligent', *qimmät* (< *qīmat*) 'expensive', *udda* (< *ʿuhdat*) 'task'.

Cases of metathesis (also found in Tajik) include *kift* (< *katf* ~ *kitf*) 'shoulder' and *qulf* (< *qufl*) 'lock'.

Words with divergent colloquial variants include *la'nat* ~ *na'lat* (< *laʿnat*) 'curse', *laqab* ~ *laqam* (< *laqab*) 'nickname', *izn* ~ *izm* (< *ʾiḏn*) 'permission'. For further examples, see the treatment of ʿ.

5. Incorporation and inflection of Arabic words

As most suffixes have only one variant, due to the loss of vowel harmony in Uzbek, nominal inflection of Arabic nouns (such as case endings and possessive suffixes) is quite straightforward. Furthermore, final consonants of borrowed stems do not undergo alternations, unlike in native words, compare *tåwåq* (< *ṭabāq*) 'dish' ⇒ *tåwåqim* 'my dish' with native *qulåq* 'ear' ⇒ *qulåġim* 'my ear'.

Normally, all Arabic nouns can be provided with Uzbek suffixes, such as plural, possessive, and case endings. Arabic stems ending in a consonant follow the same pattern as native stems, e.g. (*niyat*) *niyat-im, niyat-iŋ, niyat-i, niyat-imiz* 'my, your, her/his, our intention'. The same applies to stems in a short vowel, e.g. (*xāṭira*) *xåtira-m, xåtira-ŋ, xåtira-si, xåtira-miz* 'my, your, her/his, our memory'. There is some hesitation as to the possessive forms of words originally ending in ʿ. In spite of the disappearance of the final consonant itself, these words are inflected as if ending in a consonant. *Mawzu* (< *mawḍūʿ*) 'subject' should result in the following possessive forms: *mawzu-im* 'my subject', *mawzu-i* 'her/his subject', *mawzu-imiz* 'our subject', but *mawzu-m, mawzu-si, mawzu-miz* can now also be found, as well as *mawzuyim, mawzuyi, mawzuyimiz*, with added *y*-glide. Likewise, it is possible to find *jåme-si* 'its mosque' alongside traditional *jåme-i*. Similar variants can be found in words with final ʾ. Although the 3rd person suffix seems to behave as in native stems, e.g. (*xaṭāʾ* >) *xatå-si* 'her/his mistake', (*ʾaʿḍāʾ* >) *a'zå-si* 'its member', rather than **xatå-i, *a'zå-i*, there seems to be less certainty concerning the other possessives, leading to variants like *xatåim* ~ *xatåm* 'my mistake', *a'zåimiz* ~ *a'zåmiz* 'our members'.

Etymological hybrids abound, most commonly Arabic-Persian-Turkic compounds, e.g. *tarix-šunås-liq* '(study of) history', *zarar-kunanda-čilik* 'sabotage'.

Like all Turkic languages, Uzbek is unable to adopt Arabic verbs directly. There are two methods to form Uzbek verbs based on Arabic stems. First, an Arabic noun can be combined with a native verb such as *bol-* 'to become', *et-* or *qil-* 'to do', *ber-* 'to give', as in *rad et-/qil-* (< *radd* 'refusal') 'to refuse', *tark et-/qil-* (< *tark* 'abandoning') 'to abandon', *bahs et-* (< *baḥṯ* 'discussion') 'to discuss', *tarjima qil-* (< *tarjama*) 'to translate', *ruxsat ber-* (< *ruxṣat* 'permit') 'to allow', *mahrum bol-* (< *maḥrūm* 'bereaved') 'to lose', *wåqe bol-* (< *wāqiʿ* 'occurring') 'to occur, take place'. Alternatively, a verbalizer suffix is used, as in *håzirla-* (< *ḥāḍir* 'present, ready') 'to prepare', *åxirlaš-* (< *ʾāxir* 'last') 'to come to an end', *xawfsira-* (< *xawf* 'danger') 'to be scared'.

All inflected verb forms can be derived from the Uzbek verb stems thus created.

Although numerous derivations from the same Arabic root may occur in Uzbek, and a large part of Arabic nominal morphology is therefore reflected, not a single process of Arabic-type derivation has become productive.

Arabic plurals were apparently borrowed as separate lexical entities, and may have developed a different meaning. The plurals may lose their plural connotation, as in the case of *waqt* 'time' as opposed to *åwqat* 'food' (< *ʾawqāt* 'times [pl.]'); sometimes there is no difference in meaning, as in *tilsim(åt)* (< *ṭilsim(āt)*) 'mystery'. Double plurals like *ahwålåt* (< *ʾaḥwāl* + feminine plural ending *-āt*) 'conditions', *rusumåt* (< *rusūm* + *-āt*) 'customs' were also borrowed as such. Normally, plurals are formed with the native plural ending *-lar*, e.g. *waqtlar* 'times', even if the borrowed stem is itself an Arabic plural, e.g. *ahwållar* 'circumstances', *atråflar* 'surroundings', *haywånåtlar* 'animals'.

Some Arabic feminine forms are used to denote female persons, e.g. *ålima* (< *ʿālima*) 'learned woman', *ilåha* (< *ʾilāha*) 'goddess', *raqqåsa* (< *raqqāṣa*) 'female dancer', *såhiba* (< *ṣāḥiba*) 'lady, mistress', *zaifa* (< *ḍaʿīfa* 'weak') 'wife'. The feminine forms of some adjectives are still used, e.g. *mahbuba* 'loved', as opposed to masculine *mahbub* (< *maḥbūb*). However, feminine Arabic nouns normally take the 'masculine' version of the adjective, unlike in literary Chaghatay.

Even the most frequently occurring Arabic endings, such as the adjectival ending *-iy* ~ *-wiy* ~ *-yi* in *ilmiy* 'scientific', *wåqeiy* 'real', *måliyawiy* 'financial', *sahråyi* 'steppe-', and the adverbial ending *-an* in *aslan* 'originally', *šaxsan* 'personally', are not used to make new Uzbek derivations.

Many formulas and phrases were adopted as a whole, e.g. *barakalla* (< *bāraka llāhu*) 'bravo', *fawqulådda* (< *fawqa l-ʿāda*) 'extraordinary'. Compound words, whether of genuine Arabic origin or Persian coinages from Arabic elements, are often perceived as single lexemes, e.g. *aks-sadå* (< *ʿaks* + *ṣadā* 'counter' + 'voice') 'echo', *ġayriinsåniy* (< *ġayri* + *ʾinsānī*) 'inhumane', *låqayd* (< *lā* + *qayd*) 'indifferent', *zullisånayn* (< *ḏu l-lisānayn*) 'bilingual', *såhib-jamål* (< *ṣāḥib* + *jamāl*) 'beautiful'.

As in other Turkic languages, doublets of rhyming or related words are very popular, e.g. *ins-jins* 'evil spirits', *fikr-zikr* 'thoughts', *rasm-rusum* 'customs', *sir-asrår* 'secrets', *šak-šubha* 'doubt', *tawba-tazarru* 'remorse', *uzr-ma'zur* 'apologies'.

6. SEMANTIC AND FUNCTIONAL FIELDS OF BORROWING

Arabic loanwords are represented in nearly all categories of the lexicon. Arabic first names continue to be very popular. However, much of the basic Turkic lexicon was retained, and in most lexical fields the influence of Persian has been stronger than that of Arabic. Contrary to expectation, even common religious terms used by laypeople, for instance the words for 'prophet', 'hell', 'sin', 'fasting', and 'prayer' are often of Persian rather than Arabic origin.

Some lexical fields where Arabic predominates are words related to writing and literature, e.g. *harf* (< *ḥarf*) 'letter', *qalam* (< *qalam*) 'pen'; tribal and societal terms, e.g. *qabila* (< *qabīla*) 'tribe', *rais* (< *raʾīs*) 'chief', *dallål* (< *dallāl*) 'marriage broker'; legal and political terms, e.g. *šåhid* (< *šāhid*) 'witness', *talab* (< *ṭalab*) 'demand', *wakil* (< *wakīl*) 'representative', *wazifa* (< *waḏ̣īfa*) 'duty'.

Elevated language contains more Arabic phrases, which were adopted as a whole, e.g. *abadulåbåd* (< *ʾabadu l-ʾābād*) 'eternity', *alhål* (< *al-ḥāl*) 'presently', *filhål* (< *fī l-ḥāl*) 'immediately', *låyamut* (< *lā yamūt*) 'immortal', *nafsilamr* (< *nafsu l-ʾamr*) 'as a matter of fact'. Not all speakers are familiar with these.

Many common Uzbek adverbs, conjunctions, interjections, and exclamations are of Arabic origin. Adverbs include *ba'zan* (< *baʿḍan*) 'sometimes', *hattå* (< *ḥattā*) 'even', *nihåyatda* (< *nihāyat* 'end' + Uzbek locative case) 'very'. Conjunctions include *ammå* (< *ʾammā*) 'however', *faqat* 'but' (< *faqaṭ* 'only'). Interjections and exclamations include *ajabå* (< *ʿajabā*) 'wow!', *ya'ni* (< *yaʿnī*) 'that is to say', *ilåhim* ~ *ilåyim* ~ *ilåyå* (< *ʾilāhī*) 'my God! [exclamation]'. Postpositions based on Arabic nouns and adverbs include *åid (< ʿāʾid)* 'regarding', *dåir* (< *dāʾir*) 'concerning', *binåan* (< *bināʾan*) 'based on', *e'tibåran* (< *iʿtibāran*) 'starting from'.

Words that underwent a semantic shift include the following: *kasal* 'ill' (< *kasal* 'lazy'), *maišat* 'the good life, pleasure' (< *maʿīšat* 'livelihood'), *naġma* 'harmonica' (< *naġma* 'tune'), *naql* 'legend, tale' (< *naql* 'carrying'), *saġir* 'orphan' (< *ṣaġīr* 'small').

In some cases, an Arabic word form has been supplanted by its Russian counterpart, as in *musiqa* 'music', *qahwa* 'coffee', which were replaced by *muzika* and *kåfe*. In other cases, Arabic words were replaced by international words, e.g. *firqa* 'party' and *sayyåra* 'planet' by *partiya* and *planeta*.

BIBLIOGRAPHICAL REFERENCES

Boeschoten, Hendrik. "Uzbek". 1998. *The Turkic languages*, ed. Lars Johanson and Éva Csátó, 357–378. London and New York: Routledge.
Borovkov, Aleksandr K. (ed.). 1959. *Uzbeksko-russkij slovar'* [Uzbek-Russian dictionary]. Moscow: Gosudarstvennoe Izdatel'stvo Inostrannyx i Nacional'nyx Slovarej.
Kononov, Andrej N. 1960. *Grammatika sovremennogo uzbekskogo jazyka* [Grammar of modern Uzbek]. Moscow and Leningrad: Izdatel'stvo Akademii Nauk SSSR.
Rahmatullaev, Šavkat. 2003. *Uzbek tilining etimologik luġati*. II. *Arab sozlari va ular bilan håsilalar* [Etymological dictionary of Uzbek. II. Arabic words and their derivations]. Tashkent: Universitet.
Wurm, Stefan. 1959. "Das Özbekische". *Philologiae Turcicae Fundamenta*, ed. Jean Deny a.o., I, 489–524. Wiesbaden: F. Steiner.

HANS NUGTEREN (Dordrecht)

Uzbekistan Arabic

1. GENERAL

1.1 *History: Origin of the Arab minority in Uzbekistan*

The Central Asian Arabs settled in the area of the Khanat Bukhara and the northern plains of Afghanistan-Turkistan. There are different theories about their origin. Barfield (1981:3–4) considers two possibilities. Either their origin goes back to the Islamization of Uzbekistan at the beginning of the 8th century (see also Karmysheva 1964:272; Akiner 1983:366), or Tamerlane settled the Arabs in Central Asia in 1401 after his conquests in the west, deporting Arabs from Damascus and Aleppo. Regarding the first explanation, Arabs invaded Transoxania from a staging point in Khorasan (eastern Iran), and several Arab tribes are reported to have moved to Bukhara after these conquests. Spuler (1960:530) considers a third possibility. In his opinion, the Central Asian Arabs could have originated in Iraq and northern Afghanistan, moving to Central Asia in the 16th century. Akiner (1983:366) and Karmysheva (1964:272) support this thesis. Unfortunately, they suggest no motive for this emigration, nor do they state its extent.

The Tamerlane theory is the least reliable. The legend of Tamerlane seems to be an Arab motif for explaining their origin. Spuler's theory is unsubstantiated by other sources. The most reasonable explanation, therefore, is the first, that the current population ultimately traces its origins to the late 7th and early 8th centuries. This accords, moreover, with the linguistic evidence, which shows Central Asian Arabic to share traits with dialects far removed from it (see, e.g., Owens 1998:72; Jastrow 1997, 1998). Such linguistic affinities are explicable in terms of an early split of a dialect group, with shared traits being preserved at the fringes of the Arabic-speaking world (*Sprachinseln*).

During the 19th century, the Arabs of Central Asia were still living as nomads and subsisting on sheep breeding. They constituted a community of about thirty thousand people in the area of Bukhara. After the constitution of the Soviet Union, an effort was made to force them to settle, and as a reaction to this, many of them emigrated to Afghanistan. There are villages in northern Afghanistan (Bactria) where a nearly identical dialect is spoken (Ingham 1994; Kieffer 2000). The clan system of the Arab minority in the Soviet Union split during the forced settlement, and in the course of this process, their consciousness of identity was weakened. At the beginning of their settlement, they lived separately from the Uzbek and Tajik population in special quarters or villages. Today, they are mixed with Uzbeks and Tajiks, and many of them declare affiliation with either of these groups. Only in distinct rural areas have they been able to preserve their characteristic attributes and their original way of living. In these areas, they marry exclusively within the Arab community (Akiner 1983:366–367). Because of migration and integration into the Uzbek community, official censuses have seen the number of Arabs drop from 27,977 to 2,007 between 1926 and 1956 (Burykina and Izmajlova 1930:1; Karmysheva 1964:271). The regression has gone so far that Nowak (1995) does not mention an Arab minority in his report about the ethnopolitical situation in Uzbekistan.

1.2 *Dialects*

There are two main dialects of Uzbekistan Arabic:

i. The Bukhara dialect (approximately four hundred speakers in 1938) in the villages of Djogari, Tchardari, Shahan-Bek (Gijduvon district), and Arabkhona (Vabkend district)
ii. The Qašqa-Darya dialect (approximately one thousand speakers in 1938) in the villages of Qamashi and Djeinau (Qašqa-Darya district; Schippers and Versteegh 1987:136)

Dialects in three of these villages – Djogari, Arabkhona, and Djeinau – have become well known through publications of Russian and Georgian scholars.

The Central Asian Arabic speakers are bilingual or even trilingual, also speaking → Tajik and/or → Uzbek (Tsereteli 1970b:169). The speakers of the two main dialects speak to each other in Tajik and/or Uzbek. According to Tsereteli (1970b:168), this is because there are considerable differences between the dialects. However, it is also the case that Tajik and Uzbek are lingua francas and, therefore, strangers of any linguistic group use these languages as a matter of course.

2. Linguistic description

2.1 *Phonology*

The examples in this entry are based on the consonantal and vowel inventories of the cited authors, who give different realizations of sounds, often without clear identification of the dialect.

2.1.1 Consonants

The inventory of consonants in Table 1 is based on different dialects of Uzbekistan Arabic, often not clearly differentiated by the authors. Zimmermann, Dereli, and Fischer describe the consonants of the dialect of Djogari. Tsereteli (1939), Fischer (1961), Axvlediani (1985), Dereli (1997), and Zimmermann (2002) agree on the consonants in Table 1.

Only Dereli (1997) does not list /w/. Other authors list additional consonants:

/ƀ/	Tsereteli (1939); an implosive bilabial [ɓ], limited to Djogari
/ḏ/	Tsereteli (1939)
/ž/	Tsereteli (1939), Zimmermann (2002)
/ʾ/	Tsereteli (1939), Zimmermann (2002); limited to the dialect of Djogari
/ṣ/	Tsereteli (1939), Fischer (1961)
/ṭ/	Tsereteli (1939), Fischer (1961), Axvlediani (1985)
/ḍ (ẓ)/	Fischer (1961)
/v/	Dereli (1997), Zimmermann (2002)
/ŋ/	Dereli (1997)

Old Arabic ***q* is represented by /q/ and /g/ (Fischer 1961:237): *gidir* ‘kettle’, *qalb* ‘heart’ (Djogari), *galib* ‘heart’ (Djeinau). In Djeinau, there are also cases of **ġ* > /q/: *qēm* ‘clouds’ < **ġaym*. The phonemes /p/ and /č/ only exist in borrowings.

As far as the interdentals are concerned, there is confusion about emphasis and the merger of several interdentals. The old interdental spirants often have double correspondences: **ḏ* > /d/, /z/; **ṯ* > /s/; **ḏ̣* > /ḍ/, /ẓ/, /d/ but mostly /z/; **ẓ* > /z/.

The following words have the plosives /d/, /ḍ/ in partially fixed and partially varying form:

Table 1. Inventory of consonants of Uzbekistan Arabic

	bilabial	labiodental	dental	alveolar	palatal	velar	uvular	pharyngeal	glottal
plosive	p, b		t, d			k, g	q		
fricative	w	f	s, z	š	y	x, ġ		ḥ, ʿ	h
affricate				č, j					
liquid			l	r					
nasal	m		n						

Djeinau *xadā* ~ *xazā* (< **ʾaxaḏa* 'to take'); Djogari *darab* ~ *zarab* 'to beat', *duk*, fem. *dīki* 'that one', *orḍ*, *rōḍ* 'earth, soil'; Arabkhona *hād*, fem. *hādi* 'this one', *dahab* 'to go away', *ḥaddar* 'to protect'.

In many cases, however, /z/, /ẓ/ are lexically fixed: *zabaḥ* 'to slaughter', *abyaz* 'white', *ixzarra* 'green, spread', *ʿazim* 'bone'. The emphatic pronunciation has mostly been given up in the case of /ḏ̣/ (> /z/; Fischer 1961:238). According to Tsereteli (1956:xxiff.), historical emphasis has also been lost for /ṣ/, /ḍ/, /ṭ/. According to Zimmermann (2002:75), the variants /ṣ/ of /s/ and /ṭ/ of /t/ are conditioned by back vowels /u/, /a/: *naṭṭā* 'he jumped', *naṭṭu* 'they jumped', but *nattit* 'she jumped', *nattī* 'you [sg. fem.] jumped'. Other consonantal variants, not summarized here, are attested by one author or the other.

According to Fischer (1961:238), the old interdental spirants have partially double correspondences: /ṯ/ > /s/; /ḏ/ > /d/, /z/; and /ḏ̣/ > /ḍ/, /ẓ (z)/. In his opinion, this sound shift originated when the population which emigrated to Central Asia spoke two different dialects prior to the emigration. In one dialect, /ṯ/, /ḏ/, /ḏ̣/ became /t/ (no evidence), /d/, /ḍ/ (as in urban dialects). In the other dialect, /ṯ/, /ḏ/, /ḏ̣/ were preserved (as in Bedouin dialects), but after the emigration to Central Asia, the consonants /ṯ/, /ḏ/, /ḏ̣/ shifted to /s/, /z/, /ẓ (z)/ under Tajik influence. This would mean that Tajik has borrowed words with /ṯ/ as well as words with /ḏ/ and /ḏ̣/ from Arabic, and afterward adapted them to its inventory. Finally, Central Asian Arabic reborrowed the same words from Tajik. However, there is no explanation why this should be true only for the interdentals and not, for example, for /ʿ/ > /ʾ/ (Cowan 1967:136).

Finally, there is no obvious explanation for the sound shift in Central Asian Arabic. The general drift of Semitic languages seems to be continued in */ṯ/, */ḏ/, */ḏ̣/ > /s/, /z/, /ẓ/, the interdentals shifting to spirants (Cowan 1967:136).

Apart from these explanations, there are other possibilities. The Arabs may have borrowed Persian/Tajik words with /s/, /z/, /ẓ/ and also started to shift the interdentals /ṯ/, /ḏ/, /ḏ̣/ in words of Arabic origin toward /s/, /z/, /ẓ/. Alternatively, the speakers may have grown up as bilinguals (Tajik/Arabic) and taken over /s/, /z/, /ẓ/ instead of /ṯ/, /ḏ/, /ḏ̣/.

2.1.2 Vowels

Uzbekistan Arabic has the following vowels:

short	i		u	long	ī	ū
		(ə)			ē	ō
		a			ā	

See Axvlediani (1985:10–11) for different realization of the vowels. In unstressed syllables, /i/ and /u/ may develop to /ə/: *mudxul* ~ *mədxul* 'he enters' (Fischer 1961:236).

The vowel /ō/ is a variant of /ā/. The dialects show different values of /ā ī ō/, e.g. Djogari *kōyin* 'to be [perfect participle]', *ḥisōb* 'bill', as against Arabkhona *kāyin, ḥisāb.*

Within a paradigm, /ā ī ō/ sometimes alternate, e.g. *warō* 'behind': *warōha, warāi* 'behind her, behind me'; and the negation *ma: mōhi* 'she is not' (Djogari, Arabkhona). The vowel /ō/ remains qualitatively [o] after shortening (Fischer 1961:235), e.g. *šodd* < *šādd* (*šaddā* 'to bind'). In loanwords from Tajik, where /ā/ is always pronounced /ō/, /ō/ is always taken over, e.g. *libōs* (< Tajik) as against *ilbās* 'clothes' (Fischer 1961:234–235).

The pronunciation of /ū/ corresponds to Tajik /yy/ [ʏː] (Fischer 1961:235; Tsereteli 1956: xiv–xv).

/i/ and /u/ in stressed and closed syllables vary, based on dialect, e.g. *jifir* (Djogari) as against *jufur* (Djeinau) 'well'; *miḥiṭṭ* (Djogari) as against *miḥuṭṭ* (Arabkhona) 'he puts down'.

/i/, /u/ are deleted in unstressed, post-stress open syllables, e.g. *simiʿ* 'he hears', but *simʿu* 'he hears him'; *xuluṣ* 'he was rescued', *xulṣit* 'she was rescued' (Fischer 1961:236).

Tajik has influenced the vowel system of Central Asian Arabic. Classical Arabic /ā/ is sometimes represented by a low, front open /ā/ [æː] in the two main dialects of Uzbekistan Arabic. According to Tsereteli (1970b:168), it comes close to the short Tajik vowel /a/. Mostly, /ā/ shifts to the mid-high back rounded vowel /ō/. This applies both to borrowings and Arabic words (Tsereteli 1970b:168–169). In some cases, the same word is pronounced differently in the different dialects: *mōt* 'he died', *qōl* 'he said' (Bukhara), but *rās* 'head'; and *māt* 'he died', *gāl* 'he said', *rās* 'head' (Qašqa-Darya). Tsereteli (1970b:169) also reports that one and the same speaker may use both forms.

The realization of Classical Arabic /ā/ as /ō/ is based on Tajik influence. Middle Persian /ā/

is represented by /ō/ in Tajik, which is identical to the Central Asian Arabic vowel /ō/ (Tsereteli 1970b:169).

There is a tendency to shorten long vowels in closed syllables and at the end of words. If the shortening of the long vowel precedes the shifting /ā/ to /ō/, /ā/ is preserved as /a/; in all other cases it shifts to /ō/. *ʾImāla* at the end of the word also prevents /ā/ from shifting to /ō/ (Tsereteli 1970b:169).

/ō/ < *aw shifted to [ʏː]. This is a rounded front vowel, located between /ū/ and /ō/, with a tendency to be closer to /ū/. In Central Asian Arabic, we therefore have **zawj* > [zʏːdʒ] 'husband'. In closed syllables, [ʏː] is often shortened to [ʏ]: [zʏdʒki] 'your [sg. fem.] husband' (Tsereteli 1970b:170). Hence, this development can be summarized as follows: /ā/ > /ō~o/; /aw/ > /ō/ > [ʏː] ~ [ʏ].

2.1.3 Diphthongs

Zimmermann (2002:73) lists the following diphthongs, all cognate with Old Arabic diphthongs.

[au] ~ [aw]	[sawb] 'clothes'
[ɑu] ~ [ɑw]	[dɑuvɔ] 'medicine'
[ay]	[zayf] 'guest'
[ɑy]	[lɑyl] 'night'
[ey]	[beyt] 'house'

As reflexes of Old Arabic /ay/ and /aw/, Tsereteli (1956:xxiii–xiv) lists the falling diphthongs /ai̯/ and /au̯/, but also the pure vowels /ē/, /ō/. Furthermore, /ai̯/ may occur as /i/, as in *zin* < /zai̯n/ 'nice', and /au̯/ often shifts to /u/ in unstressed syllables, as in *suda* < /sau̯da/ 'black [fem.]' (Tsereteli 1939:262).

2.2 *Morphology*

2.2.1 Pronouns

2.2.1.1 Independent personal pronouns

	near	distant
3rd sg. masc.	*hat, hāz* (young)	*dūk* ~ *duk*
3rd sg. fem.	*hai* (young)	*dīki*
3rd pl. masc.	*halō* (adult)	*dukalō*
3rd pl. fem.	*halān* (adult)	*dikalān*
2nd sg. masc.	*hint*	
2nd sg. fem.	*hinti*	
2nd pl. masc.	*hintu*	
2nd pl. fem.	*hintin*	
1st sg.	*anā*	
1st pl.	*nəḥna*	

The pronouns of the 3rd person singular masculine and the 3rd person singular feminine are used with reference to younger persons only; the 3rd person plural masculine and the 3rd person plural feminine pronouns may refer to an older, respected person.

The 3rd person forms are identical with the demonstratives (see Sec. 2.2.1.3). The different forms of personal pronouns are also sensitive to the distance of the person spoken about (Zimmermann 2002:75).

2.2.1.2 Possessive/object suffixes

The following possessive/object suffixes occur in Bukhara.

	singular	plural
3rd masc.	-*ū*(*h*) after consonants -*hu* after vowels	-*um* after consonants, -*hum* after vowels
3rd fem.	-*ā*(*h*) after consonants -*ha*, -*hā* after vowels	-*in* after consonants, -*hin* after vowels
2nd masc.	-*ak* after consonants -*k* after vowels	-*kum*
2nd fem.	-*kī*	-*kin*
1st sg.	-*ī*, with verbs -*nī*	-*nā*

The /h/ in brackets in the 3rd person singular indicates that the object suffixes of the 3rd person are sometimes followed by a slight /h/.

The pronominal suffixes are used as possessive pronouns, as object suffixes on verbs, and they mark a subject as well as an object in combination with participles (see Sec. 2.2.3.5).

2.2.1.3 Demonstratives

sg. masc., fem.; pl.	*hā* 'this one, he', 'this one, she'
sg. masc.	*hād, hat, hāz*
sg. fem.	*hai* (Djogari), *hādī* (Arabkhona)
pl. masc.	*halō, halāu*, or *hadlāu* (Arabkhona)

pl. fem.	*halān, hadlān* (Arabkhona)
sg. masc.	*dūk, duk* 'that one, he'; *zōk, zōka* (Djeinau)
sg. fem.	*dikī, dukī* (Arabkhona); *zīk* (Djeinau)
pl. masc.	*dukalā, dukalāu* (Arabkhona)
pl. fem.	*dikalān, dikilān, dukalān* (Arabkhona)

The same forms are used for the 3rd person of the personal pronouns (Sec. 2.2.1.1).

2.2.1.4 Interrogatives

Fischer 1961:259	Axvlediani 1985:91
ēš, īš (Arabkhona), *yēš* (Djeinau) 'what?' < **ʾayy-šayʾ*	*iš* 'what?'
bišān ēš 'why?'	
ēšt-in, yēšin (Djeinau) 'whatever' < **ʾayy-šīt-in*	
išayt 'which, which one < **ēš-aḥid*	
išṭūr, ištūr 'how' < **ṭūr* < **ṭawr* 'status'	*išṭur, ušṭur* 'how?'
ēṣāb, īṣāb; īṣāb (Arabkhona) 'where?' < **ṣawb* 'direction'	*eysāb, esāb, iysāb, isab*
hen 'where?'	
minṣāb, min īṣāb (Arabkhona) 'from where?'	*min-eysāb* 'from where?'
iščailī, išzayil (Arabkhona) 'which?' < **ēš-zayy-lī*	
matā, īmit (Djeinau) 'when?' < **ʾayy-matā*	
kam 'how many?'	
lī, šinī (Arabkhona) 'why?' < **li-ʾayy; šī-in-ʾayy*	*li, šini* 'why?'
mīn 'who?'	

In Arabkhona, *ēš* may be used with a suffix, e.g. *īš-ak xaditu* 'what did she take away from you?'. Likewise, *hēn-, hīn-* may be used with a suffix to express questions like 'where [is]?'.

2.2.1.5 Reflexive pronoun

The reflexive pronoun is *rūḥ* or *eb-rūḥ*, as in *warō ḥajara zammit rūḥa* 'she hid behind the stone', *hint brūḥak min iṣāb* 'where are you yourself from?' (Vinnikov 1962:97).

2.2.2 Noun

Diminutives can be formed with the suffix *-eika*, as in *binteika* 'little daughter', which derives from Persian *-ak* (Fischer 1961:244), or with the Uzbek suffix *-čá*, as in *gidirčá* 'little pot' (Dereli 1997:17).

2.2.2.1 Definiteness and indefiniteness of the noun

As in Uzbek and Tajik, there is no definite article in the Bukhara dialect, the noun itself being interpreted as definite: *xēṭ qaṣṣu wē sakkīn* 'they cut the rope with the knife' (Fischer 1961:263; Dereli 1997:18). Traces of the Old Arabic article **al-* are preserved only in some lexical items, e.g. *balbeyt* 'door' (< **bāb al-bayt*), *roxer* 'another' (< **al-ʾāxar*; Vinnikov 1962:19). The Qašqa-Darya dialect, on the other hand, displays a definite article *il-*, e.g. *rōs il-iḥmōr* 'the donkey's head' (Chicovani 2005:131).

The indefinite article in both dialects is *fad* ~ *fat* ~ *fadaḥid* (< **fard*), e.g. *fad baqara kun ʿendu* 'he had a cow' (Vinnikov 1962:187), *fat domōna tuffāḥ* 'a lapful of apples' (Dereli 1997:81).

2.2.2.2 Construct state

Feminine nouns ending in *-a* form a construct state with *-at* ~ *-it: jalla* 'dung' – *jallet dabba* 'piece of horse dung' (Vinnikov 1962:51); *mayna* 'head' – *maynitki* 'your head' (Dereli 1997:68). The vowel may be elided: *arīza* 'money' – *arīztu* 'his money' (Vinnikov 1962:51); *dabba* 'horse' – *dab(b)tu* 'his horse' (Dereli 1997:72).

Nouns ending in *-ā* or *-ī* exhibit construct forms with *-t* as well, even if they are masculine: *akā* 'elder brother' – *akātu* 'his elder brother' (Vinnikov 1962:16); *bībī* 'grandmother' – *bībīti* 'my grandmother' (Vinnikov 1962:28); *asā* – *asōti* 'my stick' (Dereli 1997:69).

2.2.2.3 Dual and plural formation

According to Fischer (1961:243), the dual suffix *-ēn* is commonly used, e.g. *darbēn* 'two ways', *dūdtēn* 'two worms'. → Pseudoduals exist as well, but show up only in connection with possessive suffixes; thus, *ʿeyn* 'eye' has a plural *ʿuyūn* or *ʿeyniyāt*, but with suffixes *ʿeynēha* 'her eyes', *ʿeynēki* 'your [fem.] eyes' (Vinnikov 1962:149).

The plural suffix *-īn* added to the singular is confined to masculine human beings (Fischer 1961:243), e.g. *ṣayyōdīn* 'fishermen', *ʿaskarīn* 'soldiers', and often replaces broken plurals, as in *wazīrīn* 'viziers', *zūjīn* 'husbands'. For all other nouns, the suffix *-āt* may be used, again often instead of a broken plural, e.g. *dukkōnāt* 'shops', *sanduqāt* 'chests', *kitobāt* 'books' (Vinnikov 1962:180), or alternating with a broken plural, e.g. *darbāt* ~ *durūb* 'ways, streets' (Vinnikov 1962:83), *qabrāt* ~ *qubūr* 'tombs' (Vinnikov 1962:162), *ruḥāt* ~ *arwōḥ* 'souls' (Vinnikov 1962:97). Other nouns display broken plurals only, e.g. *kilāb* 'dogs' (Vinnikov 1962:184), *gidūr* 'kettles' (Vinnikov 1962:163).

2.2.2.4 Adjectives

Adjectives follow the substantive, which often receives the linker *-in* (see 2.3.3), and agree with it in number and gender, e.g. *jabalt-in ʿalīya* 'high mountain' (Vinnikov 1962:145), *šiyāt-in zīnāt* 'beautiful things' (Vinnikov 1962:106). Common adjectives such as *ṭawīl* 'long', *qaṣīr* 'short' form broken plurals, *ṭuwōl* (Vinnikov 1962:133), *quṣōr* (Vinnikov 1962:170).

The elative exists only in set phrases, such as *afzal* 'very good!', and is not productive, according to Fischer (1961:245). The comparative is formed in the Turkish way with the preposition *min*, e.g. *hint zaġīr min ruḥna* 'you are younger than we' (Vinnikov 1962:208), whereas the superlative may be expressed by a possessive suffix, e.g. *kabīru* 'the oldest of them'. An alternative construction uses the Uzbek particle *en* 'most', e.g. *en zörin kabīrin* 'the strongest and biggest of them' (Dereli 1997:86), or, in Qašqa-Darya, *čüdo*, e.g. *čüdo kbīr* 'biggest' (Chicovani 2005:130).

2.2.2.5 The cardinal numbers

Fischer (1961:245) gives a series of Arabic cardinal numbers: *fataḥid* ~ *wāḥid* ~ *ḥad* 'one', *isnēn* 'two', *salās* 'three', *orbaʿ* 'four', *xams* 'five', *sitt* 'six', *sabiʿ* 'seven', *samānia* 'eight', *tisiʿ* 'nine', *ʿašir* 'ten', *ʿašrīn* 'twenty'. Dereli (1997:86) found that in Djogari from seven onward the Tajik numbers are used: *orbaʿ, xams, sitt, haft, hašt, ʿašara*. This is confirmed by the list given by Chicovani (2005:130) for Qašqa-Darya, where the Tajik numbers start with 'five': *xamsa* or *panč* 'five', *sitta* or *šiš* 'six'.

Compound numbers either follow the Turkish order, e.g. *ʿašir wāḥid* 'eleven', *ʿašir isnēn* 'twelve', or the Tajik order, e.g. *ʿašr-u fataḥid* 'eleven', *ʿašr-u isnēn* 'twelve', *salās ʿašrāt* 'thirty', *orbaʿ ʿašrāt* 'forty' (*čil*; Dereli 1997:86), *xams ʿašrāt* 'fifty', *miʿa* ~ *sad* 'hundred'.

According to Fischer (1961:246), from two through ten and after twenty, the noun follows in the plural, whereas from 'eleven' onward (except for 'twenty'), the singular noun follows, e.g. *ʿašara salās hōiṭ* 'thirteen houses' (Chicovani 2005:130).

2.2.3 Verb

2.2.3.1 Inflection of imperfect and perfect

2.2.3.1.1 Imperfect

The imperfect is neutral as to present and future tense: *ualad-i m-iji, qāḷet, ana ilḗ-k a-zumm-ik* 'my son will come, she says, I [subj.] shall hide thee'. Verbal forms with particle *m/mi-* are very often used for representing the future: *naḥna i pāšā eš n-ən-qūl* 'what shall we tell the pasha?'.

The imperfect verb may be marked by the prefix *m-*, which appears to be cognate with the *b-* prefix of other Arabic dialects. In both the Uzbekistan and the Afghanistan (Ingham 1994:111–112) dialects of Central Asian Arabic, *m-* does not co-occur with the subjunctive prefix *ta-*: *mu-ġdi* 'he goes' (indicative), *daʿu ta-yilʿab* 'let him play' (subjunctive). In other contexts, however, Zimmermann (2002:82) could not establish consistent semantic differences in the use of verb forms with or without the particle *mi-*. Based on the information of native speakers, the individual speaker can decide which form to use, without change of meaning, although forms with particle *m/mi-* are used more frequently.

The particle *m-* must form a syllable adjacent to the verb stem. It will thus syllabify with the 3rd person singular masculine, e.g. *m-il-bis* (*mil.bis*), and with the 2nd person of weak verbs, *m-it-qul* (*mit.qul*). It cannot occur with the 2nd person of other Form I verbs like *t-ilbis*; in this case, *m-* does not syllabify with a stem-adjacent syllable since the nonstem *t-* will always prevent stem-adjacent syllabification, e.g. **mit.il.bis* or **im.til.bis*, etc.

Unless otherwise noted, the following paradigms are based on Zimmermann (2002:78–81).

Table 2. Imperfect of sound triliteral verbs

‘to put on’

	singular	plural
3rd masc.	*m-ílbis*	*m-ilbis-ū́n*
3rd fem.	*t-ílbis*	*m-ilbis-ín*
2nd masc.	*t-ílbis*	*t-ilbis-ū́n*
2nd fem.	*t-ilbis-ī́n*	*t-ilbis-ín*
1st	*m-a-lbís*	*n-ílbis*

Table 3. Imperfect of verbs II*w*/*y*

‘to say’

	singular	plural
3rd masc.	*m-iqū́l*	*m-iqūl-ū́n*
3rd fem.	*m-itqū́l*	*m-iqūl-ín*
2nd masc.	*m-itqū́l*	*m-itqūl-ū́n*
2nd fem.	*m-itqūl-ī́n*	*m-itqūl-ín*
1st	*m-a-qū́l*	*n-in-qū́l*

For verbs containing the particle *m*/*mi-* with all persons, Axvlediani (1985:66) notes that in the 1st person plural *n*/*ni-* arises via assimilation, *mi-n* > *ni-n*.

Fischer (1961:249ff.) and Tsereteli (1970a: 293) have the same pattern of conjugation for verbs II*w*/*y*. Here and in other paradigms, Axvlediani (1985:71) shortens a long /ū/ to /u/ in 2nd singular feminine, 2nd plural masculine, 2nd plural feminine, 3rd plural masculine and 3rd plural feminine, e.g. *itqulīn* ‘you [sg. fem.] say’, *itqulūn* ‘you [pl. masc.] say’.

Here and in other paradigms listed below, Axvlediani (1985:71) notes the ending *-īn* in the 2nd person plural feminine and the 3rd person plural feminine.

Table 4. Imperfect of verbs III*y*

‘to go’

	singular	plural
3rd masc.	*m-ímši*	*m-imš-ū́n*
3rd fem.	*t-ímši*	*m-imš-ín*
2nd masc.	*t-ímši*	*t-imš-ū́n*
2nd fem.	*t-imš-ī́n*	*t-imš-ín*
1st	*m-á-mši*	*ní-mši*

Final vowel length, it may be noted, is not consistent from author to author.

Table 5. Imperfect of geminated verbs

‘to put’

	singular	plural
3rd masc.	*m-əḥítt*	*m-əḥitt-ū́n*
3rd fem.	*m-ətḥítt*	*m-əḥítt-in*
2nd masc.	*m-ətḥítt*	*m-ətḥitt-ū́n*
2nd fem.	*m-ətḥitt-ī́n*	*m-ətḥitt-ín*
1st	*m-a-ḥítt*	*n-ən-ḥítt*

Table 6. Quadriliteral verbs

‘to take off’ (Axvlediani 1985:69)

	singular	plural
3rd masc.	*m-ijilmid*	*m-ijilmid-ū́n*
3rd fem.	*m-itjilmid*	*m-ijilmid-īn*
2nd masc.	*m-itjilmid*	*m-itjilmid-ū́n*
2nd fem.	*m-itjilmid-ī́n*	*m-itjilmid-īn*
1st	*m-a-jilmid*	*n-en-jilmid*

2.2.3.1.2 Perfect

Unless noted otherwise, the following paradigms are based on Zimmermann (2002:82–84).

Table 7. Perfect of sound triliteral verbs

zárab ‘to beat’

	singular	plural
3rd masc.	*zárab*	*zárab-ū*
3rd fem.	*zarab-ít*	*zarab-ín*
2nd masc.	*zaráb-t*	*zaráb-tū*
2nd fem.	*zaráb-ti*	*zaráb-tin*
1st	*zaráb-t*	*zaráb-nā*

Fischer (1961:247–248) lists the ending *-ī* in the 2nd person singular feminine. Axvlediani (1985:64) gives *-at* rather than *-it* as the suffix vowel in the 3rd person singular feminine.

Table 8. Perfect of verbs II*w*/*y*

‘to say’

	singular	plural
3rd masc.	*qṓl*	*qṓl-ū*
3rd fem.	*qṓl-it*	*qṓl-in*
2nd masc.	*qúl-t*	*qúl-tī*
2nd fem.	*qúl-ti*	*qúl-tin*
1st	*qúl-t*	*qúl-nā*

Shifting from /ā/ to /ō/ in the 3rd person singular and plural of some words is characteristic for verbs II*w*/*y*, e.g. *qōl* 'to say' (Vinnikov 1962:175–176, 186). Original /ā/ is preserved in other words, e.g. *šāl* 'to lift' (Vinnikov 1962:124–125).

Table 9. Perfect of verbs III*y*

'to go'	singular	plural
3rd masc.	*mašā́*	*maš-ū́*
3rd fem.	*maš-ít*	*maš-ín*
2nd masc.	*mašḗ-t*	*mašḗ-tū*
2nd fem.	*mašē-tí*	*mašē-tín*
1st	*mašḗ-t*	*mašḗ-na*

Axvlediani (1985:72) notes long /ē/ instead of short /i/ in the 3rd person singular feminine; in the 2nd person plural masculine and feminine, he shortens /ē/ to /e/.

Table 10. Perfect of geminated verbs

'to put'	singular	plural
3rd masc.	*ḥaṭṭā́*	*ḥaṭṭ-ū́*
3rd fem.	*ḥatt-ít*	*ḥatt-ín*
2nd masc.	*ḥattḗ-t*	*ḥattḗ-tū*
2nd fem.	*ḥattḗ-ti*	*ḥattḗ-tin*
1st	*ḥattḗ-t*	*ḥattḗ-na*

About the alternation of /t/ and /ṭ/, see Section 2.1.1. Axvlediani (1985:69) specifies long /ē/ instead of short /i/ in the 3rd person singular feminine. There are also examples of verbs being described containing short /i/ in all persons. Fischer (1961:247) states that the perfect verb is used only in historical narratives, while the perfect participle takes over the function of the perfect. Against this point of view, Zimmermann (2002:85) observes that the perfect is used for marking recent events, while the perfect participle marks occurrences that happened in the more distant past.

Table 11. Perfect of quadriliteral verbs

'to take off' (Axvlediani 1985:69)	singular	plural
3rd masc.	*jalmad*	*jalmad-ū*
3rd fem.	*jalmad-et*	*jalmad-īn*
2nd masc.	*jalmad-t*	*jalmad-tū*
2nd fem.	*jalmad-ti*	*jalmad-tīn*
1st	*jalmad-t*	*jalmad-nā*

2.2.3.2 Imperative

The imperative is constructed with verb forms of the 2nd person imperfect (Table 12). The prefix and the /n/ of the ending *-īn* (2nd pers. sg. fem.) and *-ūn* (2nd pers. pl. masc.) are lost (Fischer 1961:253; Axvlediani 1985:75–76). A double consonant at the beginning of the word is avoided by the insertion of a prosthetic vowel.

2.2.3.3 Irregular verbs

The irregular verb *sawā* 'to do' has in the perfect *sawā, sawit, sawēt*, etc.; in the imperfect, /t/ is assimilated to /s/: *misū, missī, missīn, misūn, misin, missūn, missin, ninsū*; perfect participle *sōwī*, fem. *sōwā*; imperative *sū*, fem. *sūwī*.

2.2.3.4 Derived verb Forms

Form II consists in the gemination of the second root-consonant: $C_1aC_2C_2aC_3$; it has causative and intensive meaning and is used for denominal verbs, e.g. *šarrab* 'to give to drink', *qassaḥ* 'to scatter, disperse', *pakkaz* 'to clean' (< Tajik). The meaning of Form II often coincides with that of the base Form.

There are only a few Form III verbs, C*ā*C*a*C, usually without a corresponding Form I, e.g. *ʿāyan* 'to see'.

Table 12. Imperative

	sg. masc.	sg. fem.	pl. masc.	pl. fem.
Form I	*udxul* 'enter!'	*udxul-ī*	*udxul-ū*	*udxul-in*
II*w*/*y*	*qōl* 'say!'	*qōlī*	*qōlū*	*qōlin*
III*y*	*imš* 'go!'	*imšī*	*imšū*	*imšin*
quadriliteral	*jilmid* 'take off!'	*jilmidī*	*jilmidū*	*jilmidin*
Form II	*wudd* 'bring!'	*wuddī*	*wuddū*	*wuddin*

Form V follows the scheme *it*C*a*C$_1$C$_1$*a*C. It is the reflexive/intransitive of Form II. Sometimes, there is a relation between Forms V and I, if Form II of a verb does not exist or if the meaning of the Forms I and II of this verb is identical, e.g. *itḥarrak* 'to move [intrans.]'. Weak verbs (I*y*, II*w*/*y*, III*w*) in Form V are formed by analogy with sound triliteral verbs, e.g. *ityabbas* 'to dry (up)'; *iṭṭayyab* 'to be decorated'; *itfawwar, itfōwar* 'to boil [intrans.]'; *itḥallaw* 'to become sweet'.

Form VI has the scheme *it*C*ā*C*a*C (sometimes /ā/ > /ō/). There are only a few verbs in this Form, e.g. *itʿārak* 'to argue', *itbōxal* 'to envy; to be greedy'.

Form VII is the reflexive or passive of Form I. It follows the model *in*C*a*C*a*C, e.g. *inḥasad* 'to be harvested', *inwakal* 'to be eaten'. Verbs II*w*/*y* follow the scheme *in*Cv̄C, e.g. *inqōl* 'to be said'; verbs II*y* have the scheme *in*C*a*C*ā*, e.g. *inḥašā* 'to be eaten'.

Form VIII is derived from sound triliteral verbs and geminated verbs. This Form almost completely coincides with Form VII and often with Form V, both in function and meaning. It is constructed as *i*C*ta*C*a*C and *i*C*ta*CC*ā*, e.g. *iftaraq* 'to disperse [intrans.])'; *iltammā* 'to assemble, to meet'.

Form IX is used for the expression of changing color. It follows the scheme *i*CC*a*CC*ā*; with a long last vowel, e.g. *iḥmarrā* 'to blush; to become red'.

Form X is rare, e.g. *istaġōz* 'to be angry'.

Derived Forms are used relatively rarely. At the same time, innovations have developed under the influence of coterritorial languages. One common mechanism is to use a noun with an auxiliary verb, especially *sawā* 'to do' (comparable to the use of the dummy verb *kardan* in Arabic loanwords in → Persian), as in *milḥ sawā* 'to salt' (Axvlediani 1985:57–59; Kiefer 2000:187).

2.2.3.5 The perfect participle

What is historically the active participle has a perfective value. Uniquely among varieties of Arabic, this form has developed a set of subject suffixes. The 1st and 2nd person are formed with the linking particle *-in* (the etymological origin is unclear; see Retsö 1988; → participle) with the object suffix that functions as subject or agent, as in the following examples:

1st sg.	*qaʿd-in-ī* 'I have sat down'
	lōq-in-ī 'I have found'
2nd sg. masc.	*qaʿd-in-ak* 'you have sat down'
	lōq-in-ak 'you have found'
2nd sg. fem.	*qaʿd-in-kī* 'you have sat'
	lōq-in-kī 'you have found'
1st pl.	*qaʿd-in-ā* 'we have sat down'
	lōq-in-ā 'we have found'
2nd pl. masc.	*qaʿd-in-kum* 'you have sat down'
	lōq-in-kum 'you have found'
2nd pl. fem.	*qaʿd-in-kin* 'you have sat down'
	lōq-in-kin 'you have found'

According to Fischer (1961:254), *-n-* is geminated before a V-initial suffix, *qaʿd-inn-ī* 'I have sat down', *qaʿd-inn-ak* 'you have sat down', etc. In the 3rd person, forms are used which are identical to familiar participial paradigms. However, from a syntactic perspective the suffixes function as subject markers, 3rd person singular masculine *fāʿil*, 3rd person plural masculine *faʿl-īn*, 3rd person singular feminine *faʿl-a*, 3rd person plural feminine *faʿl-āt*. Long /ā/ is shortened to /a/ before two consonants in the 3rd person singular feminine, the 3rd person plural masculine and the 3rd person plural feminine (cf. the 3rd pers. sg. fem. *faʿla* < *fāʿila*, etc., above); /ā/ has partly shifted to /ō/, e.g. *lōqinī* < **lāqinī* 'I have found'.

According to Zimmermann (2002:91–100), in the perfect participle of the 1st and 2nd person both subject and object are marked by object suffixes (see Sec. 2.2.1.2), whereas in the 3rd person, subject markers are indicated by what, historically, are 'normal' number/gender participial suffixes, -Ø, *-a, -īn, -āt*. In the following paradigms, the 2nd person singular masculine *hint zorb-in-ak* 'you [sg. masc.] have beaten' (with subject suffix *-k*) and the 3rd person masculine *hāz/hāt zōrib* 'he has beaten' are used with object suffixes.

	singular	plural
3rd masc.	*zorb-in-ak-āh*	
3rd fem.	*zorb-in-ak-āhā*	*zorb-in-ak-āhin*
1st	*zorb-in-ak-ānī*	*zorb-in-ak-ānā*

	singular	plural
3rd masc.	*zōrib-ū*	*zōrib-um* ~ *zōrib-un*
3rd fem.	*zōrib-āh*	*zōrib-in*

2nd masc.	*zōrib-ak*	*zōrib-kum*
2nd fem.	*zōrib-kī*	*zōrib-kin*
1st	*zōrib-nī*	*zōrib-nā*

Fischer (1961:255) specifies that the linking particle *-in* is used in the perfect participle of the 3rd person masculine, e.g. *zorb-in-nī* 'he has beaten me'.

The perfect participle of the 3rd person singular feminine is *hai zorb-a*. However, this form is not used with object suffixes. Instead of this, the simple past tense is used. Fischer (1961:255; not confirmed by Zimmermann's informants) describes an alternative of the 3rd person singular feminine perfect participle with object suffixes (see also Seeger 2002:635), in which the linking particle *-in* is inserted between participle and object suffixes, the feminine ending singular *-a* becoming *-it* before the object suffix, e.g. *zorb-it-in-nī* 'she has beaten me'.

When both subject and object are indicated by suffixes, there is a strong tendency for the subject suffix to end in a long vowel. This entails, in conspiratorial fashion, that subject suffixes in CvC drop the final –C, and those in v(v)C- add /ā/ after the consonants. In the latter case, it is unclear whether this /ā/ can be identified etymologically with the Classical Arabic forms ending in *-ā*, e.g. *hintu zorb-in-kum* 'you [pl. masc.] have beaten' vs. *hintu zorb-in-kū-yā* ~ *zorb-in-kū-hā* 'you [pl. masc.] have beaten her'; *hintin zorb-in-kī-nā* 'you [pl. fem.] have beaten us' (note that the latter form is identical with the 2nd pers. sg. fem.); *halō zorb-īn-ā-kum* (cf. the paradigm above for the 2nd pers. sg. masc.). Alternative forms without the C + *ā* form are also attested (dialectally/ideolectally), e.g. *halō zorb-īn-kum*. The basis of this variation is unclear.

2.3 *Syntax*

2.3.1 Word order in Central Asian Arabic

Word order in Central Asian Arabic is remarkable in having dominant SOV order, one of a number of observations which leads Owens (2001:352) to speak of Central Asian Arabic as a mixed language. Versteegh (1984–1986) focuses on the word order of Uzbekistan Arabic. Evaluating texts published by Vinnikov (1956), he found four sentence orders, more than 98 percent of them with verb-final order (number of occurrences in parentheses): OVS (2), SVO (1), OSV (29), SOV (109). SOV seems to be the unmarked sequence.

OSV order is used with WH-questions, e.g. *bōy ilā-y iš kon-qōyil* 'what did the bey say to me?'. If the interrogative is the subject, it too moves immediately before the verb, creating an OSV sequence, e.g. *hamalān mīn jab-en* 'who brought them?'. A topic position before the subject is used for purposes of stress, cohesion, and focus. An example of cohesion is provided by the sentence *xaṭīb ṣayyōt jab-u* 'the fisherman brought the mullah'. In a neutral context, this sentence would mean 'the mullah brought the fisherman'. But in the preceding sentence, the emir commanded the fisherman to bring him the mullah. The cohesion between the sentences is established by the word *xaṭīb*, which is topicalized, so that the final meaning of the sentence is 'and this mullah was brought by the fisherman'.

Central Asian Arabic is V-final, so that other verbal complements are usually preverbal as well, for instance with an indirect object *ʿō samaka anā m-a-ṣōr-mi* 'shall I then remain a fish?', or with a predicate *zaġīr ḥajara fī id-u xadā-ha* 'the young man took the stone in his hand'.

The coterritorial languages (Tajik) and (Uzbek) have been very important in the development of the SOV order of Uzbekistan Arabic. Alternative constructions have been introduced in the language, and the frequency of their use is influenced by them (Versteegh 1984–86:452).

2.3.2 Direct object marker

Jastrow (2005:136) interprets the preposition *i* 'in, to' (< *il*; Fischer 1961:263; Vinnikov 1962:18) as a definite object marker: *i-xaṭīb jābtu* 'she brought the mollah', and compares it to Persian and Turkish markers of specificity, and to Iraqi constructions of the type *čābata lxaṭīb* 'she brought the mollah'. Without the referential suffix on the verb, it seems to be used for a dative object, as in *i-hamay intu* 'give him to her!' (Dereli 1997:79). *Il-* may be used in the same way, e.g. *il-mīta dafanūa* 'they buried the dead one' (Vinnikov 1962:17).

2.3.3 Linker *-in*

The linker *-in* is suffixed to nouns in determined contexts. It appears in nominal phrases and in front of attributes, e.g. *bayt-in kabīr* 'a big house'. If the linker *-in* is suffixed to feminine nouns, it follows the feminine suffix *–t*, e.g. *baqarat-in kabīr-a* ‘a big cow’. Mostly, the part of a sentence connected by *-in* appears as attribute (adjective, numeral, or relative clause), e.g. *ḏ̣iyāt-in ġāli ġāli* ‘very expensive goods’, *mū-hin aḥmar* ‘the golden water’ (Fischer 1961:244–245).

The basic sequence in the NP is N + Adj + Numeral + Dem. The definite article is not functional. Kieffer (2000) observes that it may occur in fixed contexts, for instance always after the preposition *fi*, but that it has no semantic function.

2.3.4 Genitive constructions

Genitive constructions may be formed as in Arabic, with the modifier following the head of the noun phrase, e.g. *rās adami* ‘the head of the man’, but under Uzbek influence, the modifier may as well precede the head, which in this case carries a referential pronominal suffix, e.g. *adami rāsu* ‘the man’s head’ (Dereli 1997:18), *waladak farasu* ‘your boy’s horse’ (Chicovani 2005:131). The linker *-in* may serve the same purpose, a fact which reminds of the Tajik *iẓafet* construction, e.g. *waxt-in ṣalō* ‘prayer time’ (Dereli 1997:18; Fischer 1962:244).

2.3.5 Present tense (durative)

The participle *nāyim*, of the verb *nām/minām* ‘to lie (down)’, with the imperfect of a verb serves as a present tense, comparable to *ʾāʿid/gāʿid* in other Arabic dialects, e.g. *nōkul nāyminni* ‘we are now eating’ (Chicovani 2005:131), *iyṣāb toġdi nayminak* ‘where are you going to?’ (Vinnikov 1962:226), *šims tiqʿad nayma* ‘the sun is setting’ (Dereli 1997:25). In a similar meaning, the participles *qāʿid* and *wōquf* may be used (Fischer 1961:256), as well as *ġōdi*, e.g. *miyoġdi ġōdi* ‘he was going on’ (Dereli 1997:85).

2.3.6 Gerundial constructions

As stated by Jastrow (2005:136), under the impact of the Uzbek language the verbal noun is used like a gerund, e.g. *iš sovahānu ma yiʿrif* ‘he did not know what to do’ (Dereli 1997:25).

2.3.7 Questions

The Uzbek particle *-mi* may be used to mark yes/no questions (Jastrow 2005:136), e.g. *hāt bačča baččīt hayta mi yō hayta mi* ‘is this child from her or from her?’ (Dereli 1997:78).

2.3.8 Relative sentences

In general, relative sentences follow asyndetically, e.g. *li hama ādami ʿal mū ġadāk qazabu* ‘he seized the man who went to the water‘; rarely, Tajik *-ki* may be used as a relative marker, as in *tapsi ki jābun giddamkum* ‘the bowl they brought to you’ (Fischer 1961:241).

An alternative construction is with the pronoun *il-*, in which case the relative sentence precedes the noun, e.g. *rabīʿ il mizraʿūna ḥunṭa* ‘the wheat which they sow in spring’ (Vinnikov 1962:17). If the relative pronoun is independent, it is prefixed to the verb and the object introduces the relative sentence, e.g. *ʿeyš ilmíṭbax* ‘the one who cooks food [i.e. the cook]’ (Vinnikov 1962:17).

Bibliographical references

Akiner, Shirin. 1983. *Islamic peoples of the Soviet Union.* London: Kegan Paul.

Axvlediani, Vladimir. 1985. *Buxarskij arabskij dialekt* [The Arabic dialect of Bukhara]. Tbilisi: Mezniereba.

Barfield, Thomas. 1981. *The Central Asian Arabs of Afghanistan.* Austin: University of Texas Press.

Burykina, N. and M. Izmajlova. 1930. “Nekotorye dannye po jazyku arabov kišlaka Džugary buxarskogo okruga i kišlaka Džejnau kaška-dar’inskogo okruga Uzbekskoj SSR” [Some facts about the Arabic language in Djogari village, district of Bukhara, and Djeinau village, district of Qašqa-Darya, in Uzbekistan]. *Zapiski Kollegii Vostokovedov* 5.527–549.

Chicovani, Guram. 2000. “The Qashqa-Darya Arabic dialect of Central Asia (nominal structures)”. *Proceedings of the Third International Conference of AIDA, held in Malta, 29 March–2 April 1998*, ed. Manwel Mifsud, 189–194. Malta: Salesian Press.

——. 2005. “Linguistic contacts in Central Asia”. Csató a.o. (2005:127–132).

Cowan, William. 1967. “Sound change in Central Asian Arabic”. *Der Islam* 43.134–138.

Csató, Éva Ágnes, Bo Isaksson, and Carina Jahani (eds.). 2005. *Linguistic convergence and areal diffusion.* London and New York: RoutledgeCurzon.

Dereli, Belgin. 1997. *Het Uzbekistaans Arabisch in Djogari* [Uzbekistan Arabic in Djogari]. MA thesis, University of Nijmegen.

Fischer, Wolfdietrich. 1961. “Die Sprache der arabischen Sprachinsel in Uzbekistan”. *Der Islam* 36.232–263.

Ingham, Bruce. 1994. “The effect of language contact on the Arabic dialect of Afghanistan”. *Actas del Congreso Internacional sobre Interferencias Lingüísticas Arabo-Romances y paralelos Extra-Iberos*, ed. by Jordi Aguadé, Federico Corriente, and Marina Marugán, 105–118. Saragossa: Navarro y Navarro.

Jastrow, Otto. 1997. “Wie arabisch ist Uzbekistan-Arabisch?”. *Built on solid rock: Studies in honour of Professor Ebbe Egede Knudsen on the occasion* *of his 65th birthday, April 11th, 1997*, ed. Elie Wardini, 141–152. Oslo: Novus.

——. 1998. "Zur Position des Uzbekistan-Arabischen". *Annäherung an das Fremde. XXVI. Deutscher Orientalistentag vom 25. bis 29.9.1995 in Leipzig*, ed. Holger Preissler and Heidi Stein, 173–184. Stuttgart: F. Steiner.

——. 2005. "Uzbekistan Arabic: A language created by Semitic-Iranian-Turkic linguistic convergence". Csató a.o. (2005:123–141).

Karmysheva, Balkyz Kh. 1964. "The Arabs and Jews of Central Asia". *Central Asian Review* 12.271–274.

Kieffer, Charles. 2000. "The Arabic speech of Bactria (Afghanistan)". *Arabic as a minority language*, ed. Jonathan Owens, 181–198. Berlin: Mouton de Gruyter.

Nowak, Jürgen. 1995. "Die ethnopolitische Situation Usbekistans". *Berichte des Instituts für Ostwissenschaftliche und Internationale Studien Nr. 30*. Cologne: Bundesinstitut für Ostwissenschaftliche und Internationale Studien.

Owens, Jonathan. 1998. *Neighborhood and ancestry: Variation in the spoken Arabic of Maiduguri (Nigeria)*. Amsterdam and Philadelphia: J. Benjamins.

——. 2001. "Arabic creoles: The orphan of all orphans". *Anthropological Linguistics* 43.348–78.

Ratcliffe, Robert R. 2005. "Bukhara Arabic: A metatypized dialect of Arabic in Central Asia". Csató a.o. (2005:142–160).

Retsö, Jan. 1988. "Pronominal suffixes with *-n(n)*- in Arabic dialects and other Semitic languages". *Zeitschrift für Arabische Linguistik* 18.77–94.

Schippers, Arie and Kees Versteegh. 1987. *Het Arabisch: Norm en realiteit*. Muiderberg: D. Coutinho.

Seeger, Ulrich. 2002. "Zwei Texte im Dialekt der Araber von Chorasan". *Sprich doch mit deinen Knechten Aramäisch, wir verstehen es! 60 Beiträge zur Semitistik. Festschrift für Otto Jastrow zum 60. Geburtstag*, ed. Werner Arnold and Hartmut Bobzin, 629–645. Wiesbaden: O. Harrassowitz.

Spuler, Berthold. 1960."Appendix: Arabs in Central Asia at the present day". *Encyclopaedia of Islam* I, 530–531. 1st ed. Leiden: E.J. Brill.

Tsereteli, George. 1939. "Materialy dlja izučenija arabskix dialektov Srednej Azii" [Materials for the study of Central Asian Arabic dialects]. *Zapiski Instituta Vostokovedov Akademii Nauk SSSR* 7.254–283.

——. 1956. *Arabskie dialekty Srednej Azii* [The Arabic dialects of Central Asia]. Tbilisi: Izdatel'stvo Akademii Nauk Gruzinskoj SSR.

——. 1970a. "The verbal particle *m/mi* in Bukhara Arabic". *Folia Orientalia* 12.291–295.

——. 1970b. "The influence of the Tajik language on the vocalism of Central Asian Arabic dialects". *Bulletin of the School of Oriental and African Studies* 33.167–170.

Versteegh, Kees. 1984–1986. "Word order in Uzbekistan Arabic and universal grammar". *Orientalia Suecana* 33–35.443–453.

Vinnikov, Isaak Natanovič. 1956. *Fol'klor buxarskix Arabov* [The folklore of the Arabs of Bukhara]. *Acta Orientalia Academiae Scientiarum Hungaricae* 6.181–206.

——. 1962. *Slovar' dialekta buxarskix arabov* [A dictionary of the Bukharian Arabic dialect]. Moscow: Akademii Nauk SSSR.

Zimmermann, Gerit. 2002. *Das Arabische von Buchara zwischen alten Quellen und neuen Forschungsergebnissen*. MA thesis, University of Bayreuth.

Gerit Zimmermann (Bayreuth)

V

Valency

Valency is a verb-centered phenomenon, consisting in cataloguing the different nominal structures that participate in achieving verb complementation (Spencer 1991:190). Under the influence of the formal paradigm in linguistics, valency is, thus, reduced to a syntactic dimension. In the functional/cognitive paradigm, however, there is a grammatical valency relation when "two or more symbolic structures combine to form a more elaborate expression" (Langacker 1987:277). The result of valency is a "composite structure". Criticizing the formal view of valency, Langacker agrees that "a relational predication bearing a valence relation to a nominal participant is no doubt prototypical, but it is nonetheless only a special case in the full spectrum of possibilities" (1987:284).

To study valency in Arabic, frame semantics may profitably be used as a tool to explain it. Frame semantics is associated with the work of Fillmore (1970, 1982); its central idea is that word meaning is organized around a conceptual frame, including elements that capture the various semantic participants or roles related to and required by the frame in question in the experience of the language users. Fillmore argues that "what a speaker of language knows about the individual 'words' of his language and the conditions that determine their appropriate use is perhaps the most accessible aspect of linguistic knowledge" (1970:120). We arrive at this linguistic knowledge when we acquire the lexicon of our mother tongue; it relates to how words are pronounced, how they are built, what syntactic structures they can enter into, what sense and reference they have, and what pragmatic factors constrain their use in the world.

The Arabic verb system is organized around two major classes known as the triliteral (*ṯulāṯī*) pattern, consisting of a three-consonant tier as represented, for instance, by *ḍ-r-b* 'to beat', and the quadriliteral (*rubāʿī*) pattern, consisting of a four-consonant tier as represented, for instance, by *z-x-r-f* 'to adorn'. Whatever pattern is used, the number of consonants in an Arabic verb cannot exceed six. Basically, the triliteral pattern is the more productive since it may yield three other subpatterns by adding one, two, or three consonants to the root tier. However, no classification according to these two major classes of verbs is given here. The → transitivity system in Arabic includes two types of verbs, *lāzim* or *ġayr mutaʿaddin* lit. 'not crossing over to an object', corresponding to the intransitive pattern, and *mutaʿaddin* lit. 'crossing over to an object', corresponding to the transitive pattern in English (→ *taʿaddin*). When the verb is *lāzim* 'intransitive', it requires one noun – a subject. When it is *mutaʿaddin* 'transitive', it requires up to four nouns, one of which must be the subject noun, while the others function as direct object(s). For the sake of simplification, verbs requiring one noun are referred to here as 'one-place verbs', those requiring two nouns are called 'two-place verbs', and so on.

1. One-place verbs

One-place verbs are intransitive, not requiring for the completion of their meaning more than the nominal subject (→ *fāʿil* lit. 'doer'), as in *ḏahaba zaydun* 'Zayd left'. In frame semantic terms, the frame for these verbs presupposes only one nominal participant, whereby, for instance, the process of GOING only requires a GOER in our experience in the world we live in. The meanings of such verbs have to do with the body in itself, its creation, movement, and shape (Ibn as-Sarrāj, *ʾUṣūl* I, 169–170). These meanings have been specified as disposition and temper, e.g. *ḥasuna* 'to get better'; shape, e.g. *qaṣura* 'to get/become short'; cleanliness, e.g. *naḏufa* 'to become clean'; impurity, e.g. *wasixa* 'to become filthy'; temporary psychological state, e.g. *mariḍa* 'to become sick'; color, e.g. *iḥmarra* 'to turn red'; defect, e.g. *ʿawira* 'to become a one-eyed person'; etc. (al-ʾAnṭākī n.d.:167).

2. TWO-PLACE VERBS

Two-place verbs correspond to Quirk a.o.'s (1972) monotransitive verbs, which, apart from their nominal subject, require for the completion of their meaning a direct object, as in *ḍaraba zaydun ʿamran* 'Zayd hit ʿAmr'. The two nominals are distinguished through case in Arabic, with the subjective/nominative receiving the vowel *-u(n)*, while the objective/accusative nominal takes the vowel *-a(n)*. The logic of this two-place frame is that, for instance, the process of HITTING presupposes a HITTER and a person that is HIT (Ibn Yaʿīš, *Šarḥ* VI, 72). In cognitive semantic terms, "[HIT] designates a process in which a trajector moves through space until it makes forceful contact with a landmark" (Langacker 1987:317). HIT is "conceptually dependent because it presupposes, as an inherent part of its own internal structure, the two things participating in the correspondences" (Langacker 2002:170). Thus, the trajector associated with *-un* case marking denotes agency, while the landmark is associated with *-an* case marking, denoting affectedness. HIT organizes the scene in space between two participants in this process, where the trajector is able to use physical force and the landmark as capable of receiving the force directed against him. However, dependence is a matter of degree, as *ʿamran* seems to be more linked in a relation of dependency to *ḍaraba* than to *zaydun* (Langacker 1987:298).

An interesting distinction has been made between those two-place verbs that are *ʿilājī* 'requiring a tool' and those that are *ġayr ʿilājī* 'not requiring one' (Ibn Yaʿīš, *Šarḥ* VI, 72), the former requiring an instrument for the completion of the process and the latter not requiring one, as in *ḍarabtu ʿamran* 'I hit Amr' and *fahimtu l-ḥadīṯa* 'I understood the story', respectively. The meanings of two-place verbs have to do with every movement of the body that meets another, the senses of the body, and the body's psychology (Ibn as-Sarrāj, *ʾUṣūl* I, 170). Verbs of sensory perception require an obligatory object related to the sense in question. For instance, seeing requires something that is seen, tasting requires something that is tasted, etc.

It is worth noting that two-place verbs may see their valency reduced to that of one-place verbs, as in *fulān yabnī wa-yahdimu* 'X builds and destroys' (Sakkākī, *Miftāḥ* 228). The meaning is grasped as 'X performs the action of building up and tearing down', which gives an impression of exaggeration. This phenomenon of valency reduction has been explained as affording "reduction of expression and maximization of meaning" (Sakkākī, *Miftāḥ* 228–229). The reduction of expression has to do with deleting the direct object, whose deletion gives an impression of generality and exaggeration. This is a case of detransitivization, giving the verbal process a wider applicability than the speaker applies it to in normal circumstances.

3. THREE-PLACE VERBS

Three-place verbs are the second type of transitive verb, but they require three nominal constructions. Roughly, these correspond in Quirk a.o.'s (1972) terminology to ditransitive verbs. Three-place verbs can be exemplified by the paradigm case *ʾaʿṭā zaydun ʿamran nuqūdan* 'Zayd gave ʿAmr money'. As part of their frame, such verbs require a nominal in the nominative case represented by the ending *-un* (as in *zaydun*) and acting as subject agent, and two nominals in the accusative marked by the ending *-an* as in *ʿamran* and *nuqūdan*. It should be noted that the first accusative nominal is prototypically human in Arabic, while the second may not be. But the two nominals may be human in Arabic, as in *ʾahdā zaydun ʿamran jāriyatan* 'Zayd offered ʿAmr an odalisque'. However, when the verb is used metaphorically, it requires two nonhuman complements, as in *ʾahdā zaydun al-maʿlūmata kulla htimāmihi* 'Zayd gave the information all his attention'. The philosophy of this frame is that GIVING presupposes a GIVER, a GIVEE, and something GIVEN. It is, thus, specified in the frame semantics of this Arabic verb that the thing GIVEN can be human or nonhuman, depending on whether GIVING is used nonmetaphorically or metaphorically.

Arabic provides a great deal of flexibility in the movement of nouns, due to case markings. For instance, it is possible to have the following shift of positions between clause elements, without occasioning major differences in meaning: *ʾaʿṭaytu zaydan dirhaman* 'I gave Zayd a dirham', *ʾaʿṭaytu dirhaman zaydan* 'I gave a dirham to Zayd', and *zaydan ʾaʿṭaytu dirhaman* 'Zayd I gave a dirham' (Ibn Yaʿīš, *Šarḥ* 73). Ibn Yaʿīš distinguishes two classes of three-place verbs: *muʾaṯṯira* 'affecting' and *ġayr muʾaṯṯira* 'nonaffecting'. Affecting verbs can be exemplified through *ʾaʿṭaytu zaydan dirhaman* 'I gave Zayd a dirham', where the first object (*zaydan*) affects the second object by receiving it. This can be explained and paraphrased as: 'I gave Zayd a dirham, and Zayd received it'. The impact can, thus, be felt in the fact that the first object is actually the agent to the second object. Nonaffecting verbs, however, include verbs of belief, certainty, and knowledge, such as *ḏanna* 'to believe', *ʿalima* 'to know', *zaʿama* 'to claim', respectively. With such verbs, the first object entertains a subject-predicate relation with the second as in *ḏanantu zaydan munṭaliqan* 'I believed Zayd to be leaving', where *zaydan munṭaliqan* can be read as *zaydun munṭaliqun* 'Zayd is leaving' (Ibn as-Sarrāj, *ʾUṣūl* I, 171). The fact that nonaffecting examples allow less movement in the structure of the sentence shows that the two objects are closely linked in a subject-predicate unit. Another factor that confirms this classification is deletion of one of the objects. While affecting verbs allow the deletion of one of the objects as in *ʾaʿṭaytu zaydan* 'I gave Zayd', *ʾaʿṭaytu dirhaman* 'I gave a dirham', the nonaffecting verbs allow no such deletion, as in **ḏanantu zaydan* 'I believed Zayd to be' and **ḏanantu munṭaliqan* 'I believed to be leaving'.

4. FOUR-PLACE VERBS

The third type of transitive verb requires four nominal structures, as in *ʾanbaʾtu zaydan xālidan muqīman* 'I informed Zayd that Khalid is staying' (Ibn Yaʿīš, *Šarḥ* VI, 67). The case marker *-an* in *zaydan xālidan muqīman* attests to the existence of three nouns in the accusative. Because this class of transitivity uses verbs of belief, certainty, and knowledge, it does not affect the relation between the various accusative nouns, disallowing flexibility in noun movement and deletion between the second and third objects, as in **ʾanbaʾtu zaydan muqīman xālidan* or **ʾanbaʾtu zaydan xālidan*. Like three-place verbs that disallow movement and deletion of their accusative nouns, four-place verbs include accusative nouns that are closely linked in a subject-predicate unity.

5. VALENCY REDUCTION AND EXPANSION

Valency reduction is usually associated with *al-mabnī li-l-majhūl* 'the apophonic passive' and *al-muṭāwaʿa* 'compliance pattern'. The apophonic passive and the compliance pattern turn the accusative noun into a subjective one, detransitivizing the verb and obscuring the logical subject. However, the reasons for the disappearance of the logical subject from surface structure are different. While the absence of the logical subject from the apophonic passive is due to ignorance of who the agent is, fear for mentioning the agent's name, or fear that something might happen to someone by mentioning their name (an-Nādirī 1995:503), that of the compliance pattern has to do with causation, i.e., this pattern gives the impression that the compliant noun is the logical agent of the process involved. Valency expansion is, however, associated with *al-mušāraka* 'coparticipation pattern', which transitivizes an essentially intransitive verb, giving the impression that one of the participants (the direct object) is affected by the verbal process (→ middle verbs).

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Ibn as-Sarrāj, *ʾUṣūl* = ʾAbū Bakr ibn as-Sarī Ibn as-Sarrāj, *al-ʾUṣūl fī n-naḥw*. Ed. ʿAbd al-Ḥusayn

al-Fatlī. 3 vols. Beirut: Muʾassasat ar-Risāla, 1985.
Ibn Yaʿīš, *Šarḥ* = Muwaffaq ad-Dīn Yaʿīš Ibn Yaʿīš, *Šarḥ al-Mufaṣṣal*. 10 vols. Beirut: ʿĀlam al-Kutub, n.d.
Sakkākī, *Miftāḥ* = ʾAbū Yaʿqūb Yūsuf ibn ʾAbī Bakr as-Sakkākī, *Miftāḥ al-ʿulūm*. Ed. Naʿīm Zarzūr. 2nd ed. Beirut: Dār al-Kutub al-ʿIlmiyya, 1987.

Secondary sources
ʾAnṭākī, Muḥammad al-. N.d. *al-Muḥīṭ fī ʾaṣwāt al-ʿarabiyya wa-naḥwihā wa-ṣarfihā*. Beirut: Dār aš-Šarq al-ʿArabī.
Fillmore, Charles J. 1970. "The grammar of *hitting* and *breaking*." *Readings in English transformational grammar*, ed. Roderick A. Jacobs and Peter S. Rosenbaum, 120–133. London and Cambridge, Mass.: Ginn and Co.
——. 1982. "Frame semantics." *Linguistics in the morning calm*, ed. Linguistic Society of Korea, 111–137. Seoul: Hanshin Publishing Company.
Langacker, Ronald W. 1987. *Foundations of cognitive grammar: Theoretical prerequisites*, I. Stanford, Cal.: Stanford University Press.
——. 2002. *Concept, image, and symbol: The cognitive basis of grammar*. Berlin and New York: Mouton de Gruyter.
Nādirī, Muḥammad ʾAsʿad an-. 1995. *Naḥw al-luġa al-ʿarabiyya*. Beirut: al-Maktaba al-ʿAṣriyya.
Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1972. *A grammar of contemporary English*. London: Longman.
Spencer, Andrew. 1991. *Morphological theory: An introduction to word structure in generative grammar*. Oxford: Blackwell.

Zouhair Maalej (University of Manouba)

Variation

1. Introduction

The study of linguistic variation is based on two maxims:

i. Variation is an inherent characteristic of every living human language. This means that in every language there is more than one way of saying the same thing, and no individual speaks in exactly the same manner all the time and in all situations. In evolutionary terms, the very fact that variation in human languages has always existed implies that it is somehow conducive to human life. We can at least demonstrate that the ability to vary one's speech in accordance with the situation is an asset to one's ability to maximize gain and minimize loss in social interaction. Therefore, variation in language can be thought of as functional in itself. (For an insightful discussion of these points, see Chambers 2003.). Human languages are structured so as to make variation in structure and form possible. At the phonological level, citing an example from Arabic, the following seven pronunciations for the word /qāl/ 'to say' are possible: [qaːl], [gaːl], [ʔaːl], [kaːl], [ꝁaːl], [gɑːl], and [ʔɑːl]. These different pronunciations have the same referential meaning, but their social meanings can be very different.

ii. Variation in language is not random but structured. Native speakers have an internalized knowledge of the possible variations in their own language, although the lay speaker is normally unaware of the formal linguistic constraints on the occurrence of each form of variation. Variation in language is structured by two types of constraints:

a. Internal linguistic constraints
These are the constraints that operate within the linguistic system itself and which permit or prohibit the occurrence of certain forms under specific conditions. For example, in many Arabic dialects there is variation in the realization of /k/ such that it is pronounced either [k] or [tʃ], but this variation is constrained by the phonological environment, the variant [tʃ] occurring after front vowels only. Another example from Arabic concerns the realization of the feminine ending /ah/. In the urban Levantine dialects, the feminine ending is a variable which has two variants: [e] is the default variant which occurs everywhere except after velarized and pharyngeal consonants, and the open variant [a] is used after velarized and pharyngeal sounds. In the same region, some rural dialects have the low open variant [a] as the default variant, except after coronal sounds, in which case the ending is raised to [ɛ] (Al-Wer 2002a). At the morphophonemic level, an example can be cited from some urban Palestinian dialects as spoken in Amman. In these dialects, verbs in the imperfect of 3rd persons masculine are prefixed either by [bi]

or [bji] (i.e. with or without [j]). This variation is constrained by the structure of the syllable: [bi] occurs in open syllables, e.g. [biʔuːl] 'he says', and [bji] is used in closed syllables, e.g. [bjisʔal] 'he asks'. Because linguistic constraints are embedded in the system, they are relatively hard to break. For example, in dialect contact situations, if linguistic accommodation occurs, the process can be prohibited or at least delayed in cases where adoption of a target variant is linguistically constrained by the recipient dialect (for elaboration and examples, see Trudgill 1986). This does not mean, however, that linguistic constraints on variation cannot be broken; indeed, they can. Normally, if language change is in progress, the attested variation in the use of the old form and the new form tends to be heavily constrained at the beginning of the change; the constraints ease off as the change progresses to completion. A good illustration of this process comes from the tensing of /æ/ in American English (part of the Northern Cities chain shift), explained in detail in Labov (1994), which was first reported during the 1950s. In Arabic, a number of contemporary linguistic changes can also demonstrate the process. For example, the traditionally Najdī feature of *kaskasa* (→ *kaškaša/kaskasa*) is reported by Al-Essa (forthcoming) to be disappearing from the speech of the younger Najdī speakers who grew up in the Ḥijāzī city of Jeddah. Importantly, according to Al-Essa's data, the disappearance of the feature follows a consistent route: it has become almost obsolete when [k] occurs in the stem of a word, but it progresses at a slower rate when [k] occurs in the suffix /ik/. So, the original rule which affricated /k/ in front vowel environments has been broken, and this development is applied almost categorically to stem /k/, but partially in the case of /k/ in the suffix. It is possible to predict on the basis of the statistics provided by Al-Essa that this is a change in progress which is likely to go to completion; when it is completed, all of the original constraints will have been broken.

b. External constraints

Linguistic variation can be equally structured by social, stylistic, and geographical and spatial factors. Geographical barriers, such as rivers and mountain series, can have a significant effect on linguistic variation, as demonstrated for Arabic by Behnstedt and Woidich (2005) and de Jong (ms.). Social factors include speaker variables, such as age, sex, social class, and ethnicity. Stylistic factors concern the context of the speech situation. An alternative model of style shifting considers the addressees or the audience to be the governing factor. Research in this area views style shifting as an active responsive phenomenon, motivated by sociopsychological factors. The framework of analysis is called Accommodation Theory, which was developed by Howard Giles and his associates, and operates on a number of principles (→ speech accommodation). One such principle is that speakers converge their speech to the speech of those they like and want to be liked by, or diverge their speech from the speech of those they dislike and want to dissociate themselves from. A study by Allen Bell (1984) on style shifting by New Zealand radio broadcasters shows that style of speech shifts in accordance with the attributes or the perceived attributes of the audience. The external constraints on variation can be thought of as dimensions in human societies according to which language can vary. Arabic varies according to all of these factors, and a few additional ones, as explained in some detail in the sections to follow.

2. The role of Standard Arabic

The findings from a number of variationist studies, especially over the past two decades, provide observations in relation to the role of the standard variety in linguistic variation and change in Arabic-speaking communities. The earliest studies in the field were based on the assumption that standard linguistic features were the target forms. This assumption may have been influenced by a number of factors. Firstly, the analytical framework in use in quantitative sociolinguistics at the time was based on a hierarchical stratification of social groups, and the social meanings assigned to linguistic features were mainly based on the abstract notions of prestige and stigmatization, which tally to a large extent with the relative social status of the users of these features. Thus, the highest and most powerful social group are the most consistent users of the standard (prestigious) features, since these features are characteristic of their native dialects, and the lowest, least powerful social group are the most consistent users of the (stigmatized) vernacular features, while showing a tendency to adopt standard features for upward social mobility (since these are the features used by the social groups at the top). The notion of prestige was used more or less ready-made and on the basis of stereotypes, but its exact relevance to linguistic variation has been poorly understood. In Western societies, where the vast majority of early sociolinguistic research was conducted, the lack of a scientifically analyzed interpretive tool created a problem at the level of interpretation only, but the linguistic analyses yielded plausible correlations between the independent and dependent variables. In studies on variation in Arabic, the appeal to notions of 'prestige' and 'stigmatization', with the former connoting the use of Standard Arabic features and the latter connoting the use of the bulk of non-Standard features, has been considerably more problematic. Importantly, it has yielded paradoxical results in both analyses and interpretation. Linguistic change in spoken Arabic was approached as a case of 'standardization', by which it was meant that the target features would be the Standard Arabic features. The most seriously paradoxical result was that the most upwardly mobile speakers and the highest educated groups were in most cases the leaders of change away from the Standard feature, and the least mobile speakers and the speakers with least exposure to the Standard features used them most consistently. To illustrate, consider the case of the interdental variable (ṯ) in Jordan. It has two variants: the variant [θ] is characteristic of the traditional Jordanian dialects and at the same time is a Standard feature; the variant [t] is characteristic of the new city dialects. The results from five studies in Jordan (Abdel-Jawad 1981; Abdel-Jawad and Awwad 1989; Al-Khatib 1988; Al-Wer 1991; Al-Tamimi 2001) showed that the oldest and least educated speakers used [θ] most consistently (often categorically), and the youngest and the most educated speakers used [t] most consistently. Clearly, the Standard Arabic feature is not the target in this change. It would be equally inaccurate to interpret this result as an abandonment of a Standard Arabic feature (although it might seem so; see below). The data from a milieu of Arabic-speaking communities, reporting on structural features (phonology, morphology, and morphosyntax), which show the same pattern, are overwhelming. Given the abundance and variety of the available data, and given the fact that Standard Arabic is held in high esteem, possibly by the vast majority of Arabic speakers, interpretation of these results clearly requires a reconsideration of the basic assumptions. One such reconsideration would disentangle the question of esteem and psychological claim, which Standard Arabic undoubtedly enjoys, from the Standard variety's cognitive and practical involvement in processes of variation and change in spoken Arabic. The data available would suggest that the status of linguistic features in relation to the Standard variety is simply irrelevant. Native speakers of Arabic who have a functional knowledge of the Standard variety do, of course, resort to the use of Standard Arabic lexical items and Standard constructions in formal contexts. This is partly due to the appropriateness of the Standard variety in such domains, and partly because learned items are often only available in Standard Arabic. This usage, however, is a device in a stylistic league of its own and does not influence stylistic shifts in the vernacular (see the important article by Ibrahim 1986; and Al-Wer 1997).

3. Dimensions of variation in Arabic

The most commonly used speaker variables in classic variation studies of Arabic are age, sex (gender), and level of education. A few studies additionally included socioeconomic class and sect. There are only a few variation studies in Arabic which have incorporated a post-Labovian methodology and analytical tools. A study in point is Jabeur (1987) in Tunis, which is probably the first and only study to apply the social network approach as developed by Lesley Milroy. Two more studies are worth mentioning because they too have ventured into new ground, utilizing state-of-the-art interpretive models such as those developed by Penelope Eckert (1989), namely Hachimi

(2005) on Casablanca, and Ismail (2008) on Damascus.

The variables chosen for discussion in this entry are the ones for which empirical data are available, and which provide the possibility for making as many generalizations as possible.

3.1 *Age*

In sociolinguistic research, age is used to give the investigation a depth in time. The assumption is that the linguistic behavior of different age groups represents different stages of evolution in the language itself. Therefore, differences in linguistic usage between different age groups are taken as indicators of linguistic change. Age is measured in units (years), or it is represented in terms of life stages (childhood, adolescence, adulthood). For many sociolinguists, categorizing speakers in terms of life stages is preferred, and is seen to be consistent with a more profound analysis, especially since it is unclear why differences in age per se should correlate with differences in linguistic usage; what age differences reflect are probably differences in life experiences (for elaboration, see Milroy and Gordon 2003, Chap. 1). In all human societies, life experiences of individuals, as well as expectations of how they should behave, are different at different stages of their lives. At the same time, life experiences and society's expectations at different life stages vary across cultures. In Arab societies in general, the divisions of life stages are similar to those found in Western societies, but only at the conceptual level. In practice, youngsters, generally speaking, socialize separately from their families and form their own friendship networks at a later age than in Western societies. In sociolinguistic research, it is important to incorporate a sophisticated understanding of the social meanings of age divisions, so that it can be shown that the sampling techniques and the analytical tools are motivated in some way, for instance by the community's social practices.

The use of age differences as a substitute for time in studying language change is a practical method since it provides instantaneous results. However, it is not without problems. The main problem associated with the assumption that every linguistic difference between the speech of the young and that of the old is an instance of language change, is due to the phenomenon of age grading. This refers to the possibility that speakers may change their linguistic behavior as they grow older, and hence the linguistic differences observed between young and old at a particular point in time may not be permanent, but rather temporary alterations that are abandoned later on in life and are not transmitted to the next generation. Age-graded variations tend to recur in successive generations. It is because of age grading that sociolinguists warn that even though generational linguistic differences are symptoms of change, they are not hard and fast evidence of it. A simple demonstration of a linguistic usage that in many languages, including Arabic, tends to be age-graded is the use of trendy words and expressions and the use of obscenities, which are normally abandoned in adulthood (→ youth speech). In some cases, existing words may acquire new meanings; a good example from English is the use of the adjectives *awesome* to mean 'excellent' and *wicked* to mean 'good', which are more likely to be used by the younger generation than by the older generation in various English-speaking communities. A similar transformation in meaning, but definitely unrelated to the English case, occurs in many Arabic dialects, where the word *rahīb* 'awesome' is also used by the younger generation to mean 'marvelous' or 'excellent'. Words and expressions that are age-specific may also take the form of creations using existing linguistic material, which come about as a result of shared experience or shared life style among a group of people of similar ages. Another example of similar usages is the use of the discourse marker *ʾənno* 'that', which is originally a complementizer but has been shown by Germanos (2006) to be used in Beirut in a variety of ways, mainly as a discourse marker. For example, compare the usage of *ʾənno* as a complementizer in *walākin hay mēdde mafrūḍa yaʿne yuftaraḍ ʾənno yaˁrif leġġto* 'but this subject is obligatory; it is therefore assumed that he should know his language', with its usage as a discourse marker in *baʾa betlāḥze šeʿro ʾamrār bikūn hužūme ktīr ʾamrār bikūn ġazale ktīr ʾənno ḥasab kīf bikūn waḍˁo huwwe* 'so, you notice that his poetry is sometimes very aggressive and sometimes very amorous *ʾənno* it depends on how he feels and on his state of mind' (Germanos 2006, examples 7 and 26).

All of the studies available on variation in Arabic have shown that age is a significant variable. As would be expected, the pattern found shows that the younger generation use the incoming and new forms more often than the older generations. A careful analysis of age-related variation often finds links between the linguistic behavior of different age cohorts and social change in the community at large. Studies of variation in Arabic in real-time are as yet rare, but we may expect to see an increase in the number of projects in this area by the next generation of linguists interested in variation in Arabic. One such attempt is by Muna Al-Qouz (forthcoming) from Bahrain, whose research focuses on the speech of Manāma youth between the ages of seven and sixteen, taking as a starting point the findings by Clive Holes in the late 1970s (published 1987). Al-Qouz provides data in real time of the progression of the changes reported by Holes.

3.2 *Gender*

Students of sociolinguistics will be aware that the term → 'gender' has been more widely used than the term 'sex' in sociolinguistics especially since the 1980s. The change is not simply terminological but also reflects increasing sophistication in methods of analysis and interpretation. Chambers sums up the distinction between the two terms succinctly as recognizing "biological and sociocultural differences" (2003:117). Sex then is a biological category (being born female or male), which is determined naturally, prenatally, and over which the individual has no control. Gender, on the other hand, is determined by sociocultural factors. Sex is involved in gender in the sense that the social construction of one's self is always going to be influenced by one's sex, albeit to varying degrees in different societies. The extent to which the roles of the male and female members of society are prescribed will have a direct influence on the freedom of individuals to determine their gender identity.

As a sociological term, gender has been translated (Arabicized) as *an-nawˁ il-ijtimāˁī*, or sometimes simply called *jandar*, and gender studies are referred to as *al-jandariyya*. Of all the speaker variables included in sociolinguistic research, gender is the least transparent and the most difficult to explain. In the cases of age and social class, most people will have some intuition as to why language should vary along these dimensions. In the case of language differentiation between male and female speakers, it is not readily clear why such differentiation should exist, given that in most societies the norm is for a man and a woman to live together, and, as pointed out by Trudgill, "in most societies men and women communicate freely with one another, and there appear to be few social barriers likely to influence the density of communication between the sexes" (1995:63).

One of the most robust findings in sociolinguistic research on gender differentiation is that female speakers on average produce or maintain linguistic features that are closer to the standard or prestigious features more often than male speakers. This generalization is valid so long as we bear in mind that it is based on statistics arrived at through averaged data. An important reformulation by James and Lesley Milroy (J. Milroy a.o. 1994) of the sociolinguistic findings of gender-related variation looks at male and female preferences not in terms of standard vs. nonstandard, but in terms of localized vs. supralocal features. Localized linguistic features are features that tend to be characteristic of a particular dialect, or dialects of a particular region, and are generally not found in other dialects. Supralocal features are features that can be found, or have spread, in a larger geographical area and, therefore, are not peculiar to a particular dialect or particular group of dialects. The restatement in terms of localized vs. supralocal features stretches the generalization in relation to female vs. male linguistic behavior to include patterns of variation in languages where there is a formal standard and de facto standards, such as is the case in Arabic. Because of the controversy that has surrounded the issue of gender differentiation patterns in Arabic communities as compared with other languages, which sometimes has led to misconstrued conclusions in the literature, the issue is treated here in some detail (see also Ibrahim 1986; Haeri 1987; Al-Wer 1997). Early sociolinguistic studies of Arabic treated variation and change as an approximation to Standard Arabic; crucial to the present discussion is that the variants investigated were defined in terms of Standard vs. non-Standard Arabic. For example, in the case of the variable (ṯ) mentioned earlier, the variant [θ] is described as the Standard variant and [t] as the non-Standard variant. These definitions fall short of recognizing the relevant social meanings of these variants. In Jordan, for instance, the variant [θ] signifies an outdated lifestyle, and the variant [t] symbolizes a modern lifestyle. Study results which showed that women use [θ] less often than men led to the conclusion that unlike female speakers in the West, Arab women use the Standard prestigious linguistic features less often than men. The correct interpretation is that the male speakers are more consistent in the use of the traditional feature [θ] and the female speakers are more innovative. In a sense, the false interpretations deceived attempts to explain gender differentiation in language in general. Labov (1994), for instance, treated the reported pattern (falsely interpreted) of male vs. female speech in Arabic as an anomaly. If a generalization is needed in this regard, the true situation is that male and female patterns in Western and Arab communities are identical in that men tend to use the localized features more often than women, and Arab women prefer supralocal features (which are often non-Standard features). Keeping with the same variable to demonstrate this point, in the Levant, the variant [θ] can easily be shown to be a localized feature, characteristic as it is of the nonurban traditional dialects in the region, whereas the variant [t] is supralocal, characterizing as it does all of the metropolitan dialects (Jerusalem, Damascus, Beirut, and now Amman).

A further important generalization to emerge from sociolinguistic research is that where language change is in progress, in the majority of cases women are ahead of men in using the new features, i.e., women lead linguistic change. This has been found to be true in all of the studies conducted in North America, Britain, Latin America, Hong Kong, Korea, and the Middle East. In a small minority of cases of change in progress, men were found to be in the lead, and in these cases the changes involved were relatively minor. In the data available from Arabic, women are sometimes ahead of men by a whole generation. For example, in Al-Wer (1991), which investigated the use of four phonological variables in three towns in Jordan, the men were found not to participate in the variation between the local and nonlocal features at all; hence, they were excluded from the analysis. In a small-scale restudy of the town of Sult in 1997, the young men were beginning to participate in the variation interdental-stop, and affricate-fricative realizations of /j/. In the dialect of Amman (see Al-Wer 2007), all cases of divergence from the traditional features were found to be led by the young women. A very interesting pattern is reported in Al-Essa's (forthcoming) work on the patterns of accommodation by Najdī speakers to Ḥijāzī features in the city of Jeddah. Overall, she found that among the oldest generation, the women were considerably more conservative with respect to the Najdī traditional features; this finding is explained with reference to the older women's lack of contact with and access to the target features as a result of restrictions in the range of socialization. However, as social restrictions were relaxed in succeeding generations, the women suddenly outscored the men in adopting the new features. This finding is significant because it shows the interplay between two types of factors, social and biological, in the following way. The social conditions provide a plausible explanation in the case of differences among the male and female speakers in the older generation. But note that changes in the social conditions alone cannot explain the finding that the younger women outscore (rather than just equalize) the younger men. Findings of this order from sociolinguistic research over three decades, which indicate some sort of verbal superiority on the part of the female sex, have prompted scholars to give due consideration to the possibility that some of the observed linguistic differences between men and women may be sex-based. For example, it has been suggested that the female sex has a neuropsychological advantage over the male sex, i.e. that women have an innate linguistic advantage (for a review of evidence in this area, see Chambers 2003, Chap. 3).

3.3 *Education*

Level of education as a speaker variable can be found in the early large-scale studies as well as in the later studies, for instance Schmidt (1974), Abdel-Jawad (1981), Jabeur (1987), Al-Muhannadi (1991), Al-Wer (1991), Al-Khatani (1992), and Al-Shehri (1993). In some cases, education simply refers to whether the speakers are literate or not, and in other cases it is quantified to include lower, medium, or higher levels of education. A common pattern was found in all studies, namely that the higher the educational level of speakers, the more innovative and the less conservative they are in their linguistic behavior. In other words, the educated speakers were found to be the leaders of language change and change away from the Standard Arabic variety, if one wants to define variation in these terms. Hence, what 'education' does not do is to promote the language associated with education. This observation has prompted Al-Wer (2002b) to revisit this variable in order to understand the exact denotations of education as a speaker variable. She suggests that although correlations can be established between the speakers' level of education and their linguistic choices, the correlations in themselves cannot ensure that education is not a proxy variable, which acts on behalf of other less obvious independent variables. The suggestion is that level of education is an indicator of the nature and extent of the speakers' social contacts (i.e. their social network). It just so happens that generally in the Arab world, access to education, especially at the higher level, often involves leaving one's hometown, changes in familial links, expansion in social contacts, interaction with speakers of other dialects, exposure to different social values, shifting of one's loyalties and attachments to various social groups, changes in priorities and ambitions, etc. All of these, and others of a similar nature, are important factors in shaping the individuals' linguistic behavior. It follows then that the reason why the classification of speakers according to level of education has thus far provided researchers with fairly accurate results, especially in terms of locating the social groups who initiate and/or diffuse new features, is that in the Arabic-speaking communities, particularly those with a recent history of urbanization, education is the major channel through which members of the community have opportunities of contact with the speakers of the target features. It is very likely that when this developmental phase is completed and education ceases to be the only prerequisite for social mobility and contact, it will no longer show straightforward correlations with linguistic usage.

3.4 *Ethnicity, sect, and religion*

Ethnicity, sect, and religion often coincide in Arab societies in the sense that a group of a different ethnic origin can also be of a different sect and belong to a different religious group in relation to the majority. Differences in Arabic dialects along these lines have been well documented by early European and Arab linguists, as well as by contemporary scholars (→ communal dialects). One clear example of variation according to the speakers' sect is the case of Bahrain and its two main denominations, the Arab Sunni group and the Baḥārna Shiʿi group. Perhaps the most useful points to discuss in this domain are the social and political forces that contribute to the maintenance of linguistic divisions along these lines, and the repercussions on divisions along ethnic, sectarian and religious lines of the modern and the currently more inclusive – whether by compulsion or by will – urban Arab societies. It is reasonable to suggest that the persistence of linguistic differences in one community between groups of various heritages are due in the first place to social barriers that would prevent individuals from fully interacting with members of the outside groups. The existence of social barriers for a lengthy period of time leads to linguistic features that may have been regional features in the first place, becoming ethnic/sectarian/religious group markers (see the discussion in Edwards 1985). A well-documented case which illustrates this point is that of the linguistic differences between Christian and Muslim Baghdadis, as reported in Blanc (1964). Behnstedt and Woidich (2005:43–46) provide further elaboration and linguistic maps of the communal dialects in Iraq (Mesopotamia), Djerba (Tunisia), and Bahrain.

Naturally, the linguistic features which symbolize the boundaries of the group can over time also become markers of identity, and their maintenance can become associated with loyalty to one's own group. Not all Arab societies in which different ethnicities or different religious groups have coexisted for a considerable period of time show linguistic differences along these lines. For instance, no such differences can be detected in Jordanian society between the Christian and Muslim Jordanians, and this undoubtedly is related to the fact that, except for cross-marriage, no social or physical barriers have existed which would prevent the members of either group from fully interacting with members of the other group. But this situation can alter since sociolinguistic situations are not static. For instance, in the Jordanian case, the prevailing marriage customs can indirectly lead to linguistic-religious group associations. The country's population has grown from one million during the early 1960s to well over six million forty years later. In the 1960s, Christian Jordanians formed roughly 15 percent of the population. The raw number of Christians has fallen significantly due to emigration, and the proportion to the total population has now fallen to probably half of this figure. At the same time, the vast majority of the immigrants are Muslims. Cases of cross-marriage between Christians and Muslims have always been extremely rare, and they continue to be prohibited by both groups (apparently without hard feelings on either side). This situation has meant that over the years the Muslim population's mix through marriage with nonlocals has increased dramatically, while the choices of partners for the Christians have remained largely locally based (i.e. between local Jordanian partners). The result is that through marriage the mix with speakers of non-Jordanian dialects has risen markedly in the case of the Muslim population. By a stretch of imagination, a few decades down the line, the traditional features may become part of the native linguistic repertoire for the Christians only (or at least more consistently so); they may hence also become religious markers. One may conclude that, especially in contemporary Arab societies, sociolinguistic situations are increasingly mutable.

3.5 *Social class*

Social class refers to the hierarchical stratification of society's members. In modern societies, the basis of the stratification into different classes is usually economic ability, which, in turn, translates into and determines social power. Therefore, social class is generally used to mean socioeconomic class (see Milroy and Gordon 2003, Chap. 2). In sociology as in sociolinguistics, there are no absolute criteria that can be used for class stratification. Wealth, income, occupation, father's occupation, educational level, and housing type are among the criteria that have been used to determine social class. What is clear, however, is that the social class system is a fluid system, in the sense that it is generally possible for individuals to move up and down the scale.

Similarities in socioeconomic status between groups of people are often paralleled by similarities in lifestyle, attitudes, ambitions, and values, and lead to differentiation between different classes along these lines. It is also largely true that people socialize more frequently with others in their own socioeconomic class. This observation leads us to explain the existence of social dialects or sociolects, i.e. dialects that are differentiated according to social class, on the basis of frequency of contact.

Social class is not a universal sociolinguistic variable, for although it can be argued that differences in wealth exist in every society, these differences do not always translate into sociolinguistic correlations. In many Arab societies, linguistic variation cuts across socioeconomic differences and does not correlate with them. The effect of the type of social structure is key to understanding why this difference exists, which can be demonstrated by comparing a hierarchical social structure, such as the class system, with an egalitarian tribal system. Not all tribal systems are egalitarian, but in those which are, individuals can differ in wealth but not in social status as members of the clan. This leads to a situation where there are no social barriers between individuals of varying wealth and, perhaps therefore, no linguistic differences according to economic status.

One can expect sociolinguistic correlations to alter in accordance with changes in the social structure of a community. A good illustration comes from Middle Eastern societies that have experienced increasing degrees of urbanization and population mobility (especially from countryside to city). A series of recent sociolinguistic investigations point to the importance of distinguishing two types of cities: old, well-established cities and new cities. The first type includes cities with a long-established history of urban life in social characteristics and in economic activities (see Miller a.o. 2007). Cairo and Damascus are good examples of this type. In Cairo, Haeri (1997) shows that social class is an important variable which clearly correlates with linguistic variation, blurring differences among the speakers in terms of their ethnic or regional origins. The second type of cities are those which have had a recent history of urban life, and can be exemplified by the cities of the Gulf region (Dubai, Kuwait, etc.) and Amman. In Amman, for instance, no straightforward correlation can as yet be established between linguistic usage and socioeconomic class; tribal divisions and ethnic or regional origins continue to play the major role in structuring variation (→ Jordanian Arabic: Amman). This situation is expected to change as reliance on an urban-based economy will ultimately lead to homogenization and the blurring of ethnic, sectarian, and other differences.

4. VARIATION, SPACE, AND POLITICAL BORDERS

The political borders which divide the Arab world of today were fixed at various times following the collapse of the Ottoman Empire. As elsewhere in the world, political borders do not suddenly alter the linguistic maps of regions, but gradually they create different realities, which divide one family of dialects or force different groups of dialects together into a new political entity. The extent to which dialects that suddenly find themselves border dialects of different political entities diverge over time is very much influenced by the frequency of contact that the speakers on both sides are able to maintain. Situations of this sort can be found all over the Arab world, and a treatment of each one of these cases would be complex and certainly rewarding. In this section, the aim is to highlight the importance to the study of variation in Arabic of sociopolitical changes by drawing on one example of such a situation, the region of Ḥōrān, across which runs the political border which has divided Jordan from Syria since 1921. The plains of Ḥōrān stretch some three hundred kilometers from the southern outskirts of Damascus to the Jordanian district of Muʾab. It is a rich agrarian society and includes no fewer than eight hundred villages. The dialects of the region belong to one southern Levantine group (see Behnstedt 1997). After 1921, adjacent villages were split between the two countries. Along the border, many tribes carry the same family names. The dialect spoken by the older generations on one side of the border is indistinguishable from the dialect spoken by the same generations on the other side, but the younger generations diverge markedly from their forefathers' dialects. The reason for this divergence is that while the Ḥōrānis on the Syrian side look north to the dialect of Damascus as their standard, their relatives on the Jordanian side look south to the dialect of Amman as a standard. Contact and familial links were not severed by the political division, even at points in history when the central governments of Syria and Jordan were not on the best of terms. It is rather the new political reality in the region which has shaped the linguistic developments in Ḥōrān. It can be further argued that the political separation of the southern part of Greater Syria is altogether responsible for the development of a separate local standard in Jordan; what might have been expected had the separation not occurred are linguistic adjustments which would have converged the Jordanian dialects toward the regional standard (e.g. the dialect of Damascus), i.e. a development similar to what has happened in the Syrian part of Ḥōrān.

BIBLIOGRAPHICAL REFERENCES

Abdel-Jawad, Hassan Rashid E. 1981. *Lexical and phonological variation in spoken Arabic in Amman.* Ph.D. diss., University of Pennsylvania.

——. and Mohammad Awwad. 1989. "Reflexes of Classical Arabic interdentals: A study of historical sociolinguistics". *Beiträge aus Forschung und Anwendung* 122.259–282.

Al-Ahdal, H. 1989. *A sociostylistic description of speech in Mekkah.* Ph.D. diss., University of Reading.

Al-Essa, Aziza. Forthcoming. *Najdi speakers in Hijaz: A sociolinguistic investigation of dialect contact in Jeddah.* Ph.D. diss., University of Essex.

Al-Khatani, Ali. 1992. *The impact of social change on linguistic behaviour: Phonological variation in spoken Arabic, Asir, Saudi Arabia.* Ph.D. diss., University of Essex.

Al-Khatib, Mahmoud. 1988. *Sociolinguistic change in an expanding urban context: A case study of Irbid city, Jordan.* Ph.D. diss., University of Durham.

Al-Muhannadi, Muneera. 1991. *A sociolinguistic study of women's speech in Qatar.* Ph.D. diss., University of Essex, Colchester.

Al-Qouz, Muna, Forthcoming. *Dialect contact, acquisition and change among Manama youth.* Ph.D. diss., University of Essex.

Al-Shehri, Abdullah. 1993. *Urbanisation and linguistic variation and change: A sociolinguistic study of the impact of urbanisation on the linguistic behaviour of urbanised rural immigrants in Hijaz, Saudi Arabia.* Ph.D. diss., University of Essex, Colchester.

Al-Tamimi, Fida. 2001. *Phonetic and phonological variation in the speech of rural migrants in a Jordanian city.* Ph.D. diss., University of Leeds.

Al-Wer, Enam. 1991. *Phonological variation in the speech of women from three urban areas in Jordan.* Ph.D. diss., University of Essex, Colchester.

——. 1997. "Arabic between reality and ideology". *International Journal of Applied Linguistics* 7:2.251–265.

——. 1999. “Gender and sociolinguistic variation”. *Language and gender*, ed. Jennifer Coates, 64–75. Oxford: Blackwell.

——. 2002a. “Jordanian and Palestinian dialects in contact: Vowel raising in Amman”. *Language change: The interplay of internal, external and extra-linguistic factors*, ed. Mari Jones and Edith Esch, 63–79. Berlin: Mouton de Gruyter.

——. 2002b. “Education as a speaker variable”. *Language contact and language conflict in Arabic: Variations on a sociolinguistic theme*, ed. Aleya Rouchdy, 41–53. London: Routledge Curzon.

——. 2007. “The formation of the dialect of Amman: From chaos to order”. Miller a.o. (2007:55–76).

Behnstedt, Peter. 1997. *Sprachatlas von Syrien*. Wiesbaden: O. Harrassowitz.

——. and Manfred Woidich. 2005. *Arabische Dialektgeographie*. Leiden: E.J. Brill.

Bell, Allen. 1984. “Language style as audience design”. *Language in Society* 13:2.145–204.

Blanc, Haim. 1964. *Communal dialects in Baghdad*. Cambridge, Mass.: Harvard University Press.

Chambers, Jack. 2003. *Sociolinguistic theory*. Oxford: Blackwell.

Eckert, Penelope. 1989. *Jocks and burnouts: Social categories and identity in the high school*. New York: Teachers College Press.

Edwards, John. 1985. *Language, society and identity*. Oxford: Blackwell.

Germanos, Marie A. 2006. “From complementizer to discourse marker: The functions of *ʾenno* in Lebanese Arabic”. Ms.

Hachimi, Atiqa. 2005. *Dialect levelling, maintenance and urban identity in Morocco*. Ph.D. diss., University of Hawaii.

Haeri, Niloofar. 1987. “Male/female differences in speech: An alternative interpretation”. *Proceedings of the Fifteenth Annual Conference on New Ways of Analyzing Variation*, ed. Keith M. Denning, Sharon Inkelas, Faye C. McNair-Knox, and John Rickford, 173–182. Palo Alto, Cal.: Department of Linguistics, Stanford University.

——. 1997. *The sociolinguistic market of Cairo: Gender, class, and education*. London: Kegan Paul International.

Holes, Clive. 1987. *Language variation and change in a modernising Arab state*. London: Kegan Paul International.

——. 1995. “Community, dialect, and urbanization in the Arabic-speaking Middle East”. *Bulletin of the School of Oriental and African Studies* 58.270–287.

Ibrahim, Muhammad Hassan. 1986. “Standard and prestige language: A problem in Arabic sociolinguistics”. *Anthropological Linguistics* 28:1.115–126.

Ismail, Hanadi. 2008. *Suburbia and the inner city: Patterns of linguistic variation and change in Damascus*. Ph.D. diss., University of Essex.

Jabeur, Muhammad. 1987. *A sociolinguistic study in Tunisia: Rades*. Ph.D. diss., University of Reading.

Jong, Rudolf de. Ms. “A grammar of the Bedouin dialects of the middle and southern Sinai desert”. Unpublished Ms.

Labov, William. 1994. *Principles of linguistic change: Internal factors*. Cambridge, Mass.: Blackwell.

Miller, Catherine. 2007. “Urban Arabic vernacular: Development and change”. Miller a.o. (2007:1–32).

——, Enam Al-Wer, Dominique Caubet, and Janet Watson (eds.). 2007. *Arabic in the city*. New York: Routledge Curzon.

Milroy, James, Lesley Milroy, Sue Hartley, and David Walshaw. 1994. “Glottal stops and Tynside glottalization: Competing patterns of variation and change in British English”. *Language Variation and Change* 6:3.327–359.

Milroy, Lesley and Matthew Gordon. 2003. *Sociolinguistics: Method and interpretation*. Oxford: Blackwell.

Schmidt, Richard. 1974. *Sociolinguistic variation in spoken Egyptian Arabic: A re-examination of the concept of diglossia*. Ph.D. diss., Georgetown University.

Trudgill, Peter. 1974. *The social differentiation of English in Norwich*. Cambridge: Cambridge University Press.

——. 1986. *Dialects in contact*. Oxford: Blackwell.

——. 1995. *Sociolinguistics: An introduction to language and society*. Harmondsworth: Penguin.

ENAM AL-WER (University of Essex)

Velarization

Velarization is a secondary articulation that refers to the raising of the tongue body toward the back of the soft palate. Classical Arabic has four velarized phonemes, /ṭ/, /ḍ/, /ṣ/, and /ḏ̣/, all of which are coronal consonants with a primary articulation in the dental-alveolar region of the vocal tract. Sībawayhi uses the term *muṭbaq* ‘covered’ in his *Kitāb* to describe the velarized consonants of Arabic (→ *ʾiṭbāq*). These consonants are often referred to as pharyngealized. In terms of articulation, a pharyngealized consonant is one made with the root of the tongue drawn toward the back wall of the pharynx. Ladefoged and Maddieson (1996:365) note that Arabic dialects vary as to whether they have velarization or pharyngealization. There is very little acoustic distinction between the two; both are characterized by the lowering of the second formant. The term ‘emphatic’ is also used to refer to the velarized consonants. This is a translation of the Arabic term *mufaxxama*, which is traditionally understood to include the uvular consonants as well. Both uvular and velarized consonants are made with the tongue body raised toward the back of the soft palate; with the uvular consonants, the tongue body raising is the primary articulation, whereas with the velarized consonants it is a secondary articulation that accompanies the primary contact in the dental-alveolar region. The English term ‘emphatic’ is typically used only for the velarized consonants.

A phonological analysis of velarization in an Arabic dialect must answer the following two questions: what are the velarized phonemes of the dialect, and to what extent does velarization affect neighboring sounds? With respect to the first question, while almost all modern dialects have maintained the velarized consonants, there are some differences from Classical Arabic. For example, colloquial Cairene Arabic has /ẓ/ rather than Classical Arabic /ḏ̣/, while the northern Palestinian dialect described in Herzallah (1990) lacks /ḍ/. Furthermore, most dialects are reported to have a variety of other velarized consonants that are sometimes referred to as ‘secondary emphatics’. These include the liquids [r̴] and [ɫ] as well as the labials [b̴] and [m̴]. However, the phonemic status of the secondary emphatics is controversial in that they are often marginal, not occurring in many words or only occurring in the environment of a low vowel. As an example, although many dialects are reported to have a velarized [r̴], its phonological behavior is usually distinct from the ‘primary emphatics’ /ṭ/, /ḍ/, /ṣ/, and /ḏ̣/. This is because velarized [r̴] usually only occurs in the context of a low vowel and alternates with its nonvelarized counterpart when the low vowel is not present. An illustration of this comes from Cairene Arabic, where the /r/ in the word [kubɑːr̴] ‘big [pl.]’ is pronounced as velarized when adjacent to a low vowel, but pronounced as nonvelarized in [kabiːr] ‘big [sg.]’ when the low vowel is not adjacent. Primary emphatics do not alternate in such a manner; they are pronounced as velarized regardless of the vocalic context, as exemplified by the Cairene Arabic pair [naʃiːt̴] ‘active’/[ʔanʃɑt̴] ‘more active’.

The second question that a phonological analysis of velarization in an Arabic dialect must address is the extension of velarization beyond the single phoneme. In most dialects, when there is a velarized phoneme in the word, neighboring sounds are also pronounced as velarized. In some dialects it is only the vowel immediately adjacent to the velarized consonant that is pronounced as velarized (e.g. certain dialects of Saudi Arabia), whereas in other dialects it may be the entire word that is velarized (e.g. Cairene Arabic). Thus, depending on the dialect, the common Arabic word /ṣabāḥ/ ‘morning’ will either be pronounced with just the first syllable velarized or with the entire word velarized. The effect of velarization is particularly strong with the low vowel, where it is generally identified as allophonic variation between a low back vowel (velarized) and a low front vowel (nonvelarized). Moreover, detailed investigations of velarization in specific dialects (e.g. Ghazeli 1977; Heath 1987; Herzallah 1990; Younes 1993; Davis 1995; Shahin 1997; Watson 1999; and Zawaydeh 1999) reveal that some dialects display an asymmetry between velarization that is regressive and velarization that is progressive. For example, in the northern Palestinian dialect discussed in Herzallah (1990), velarization extends (or spreads) regressively to the beginning of the word, but the progressive spreading of velarization is limited. Thus, when a word has a velarized consonant at the end, the entire word is velarized, as exemplified by /xayyāṭ/ ‘tailor’, which is pronounced entirely velarized. But when the velarized consonant is at the beginning, only part of the word would usually be velarized, as exemplified by /ṣabāḥ/, in which only the first syllable is velarized. Another manifestation of the asymmetry between the regressive and progressive spreading of velarization is the observation that in some dialects certain sounds (typically those made with a high tongue position) block the progressive spread of velarization. In the rural Palestinian dialect discussed by Younes (1993), the consonants /š/, /y/, and /w/ block progressive spreading, as exemplified by the word /ṣiyām/ ‘fasting’, in which only the initial /ṣ/ is pronounced velarized; these consonants do not prevent regressive spreading, as illustrated by the word /xayyāṭ/ ‘tailor’, in which the entire word is pronounced as velarized. A third asymmetry between regressive and progressive spreading of velarization, documented by Zawaydeh (1999) for Ammani Jordanian Arabic, is that regressive spreading is categorical, whereas progressive spreading is gradient. For example, the word /tasallaṭ/ ‘he overruled’ is pronounced with the first two syllables just as velarized as the third syllable, whereas in the word /ʿaḍalātak/ ‘your muscles’, the third syllable shows less velarization than the second, and the fourth syllable shows even less velarization than the third. The asymmetries manifested between regressive and progressive assimilation reflect that regressive velarization is stronger than progressive velarization, and this may have a basis in articulation, as suggested by Watson (1999).

A largely unexplored topic with respect to velarization is its sociolinguistics, namely to what extent, if any, does velarization in a single dialect vary with respect to gender, style, social class, etc. An investigation undertaken by Royal (1985) documented acoustically stronger velarization patterns for males than for females.

Bibliographical references

Davis, Stuart. 1995. "Emphasis spread in Arabic and grounded phonology". *Linguistic Inquiry* 26.465–498.

Eid, Mushira and Clive Holes (eds.). 1993. *Perspectives on Arabic linguistics*, V. Amsterdam and Philadelphia: J. Benjamins.

Ghazeli, Salem. 1977. *Back consonants and backing coarticulation in Arabic.* Ph.D. diss., University of Texas at Austin.

Heath, Jeffrey. 1987. *Ablaut and ambiguity: Phonology of a Moroccan Arabic dialect.* Albany: SUNY Press.

Herzallah, Rukayyah. 1990. *Aspects of Palestinian Arabic phonology: A nonlinear approach.* Ph.D. diss., Cornell University.

Ladefoged, Peter and Ian Maddieson. 1996. *The sounds of the world's languages.* Oxford: Blackwell.

Royal, Anne Marie. 1985. *Male/female pharyngealization patterns in Cairo Arabic: A sociolinguistic study of two neighborhoods.* Ph.D. diss., University of Texas at Austin.

Shahin, Kimary. 1997. *Postvelar harmony: An examination of its bases and crosslinguistic variation.* Ph.D. diss., University of British Columbia.

Watson, Janet. 1999. "The directionality of emphasis spread in Arabic". *Linguistic Inquiry* 30.289–300.

Younes, Munther. 1993. "Emphasis spread in two Arabic dialects". Eid and Holes (1993:119–145).

Zawaydeh, Bushra Adnan. 1999. *The phonetics and phonology of gutturals in Arabic.* Ph.D. diss., Indiana University.

Stuart Davis (Indiana University)

Verb

1. Introduction

The Arabic word for verb is → *fiʕl*. This term has undergone a double metalinguistic reinterpretation and generalization: 'action' > 'expression designating an action' > '(action or nonaction) verb'. In the Arabic grammatical tradition, while syntactically (*naḥw*) representing merely the second element of the tripartite division of the parts of speech (after the noun and before the particle), the verb occupies a central place in the morphology (→ *ṣarf* or *taṣrīf*). This entry deals first with the morphology of the verb, according to the main subdivision inflection/derivation. Clearly, the opposition inflection/derivation is not a rigid one. In fact, inflection can use the same processes as derivation, i.e. the adding of affixes to a base, just like derivation may be performed through 'internal inflection', i.e. → apophony and vowel lengthening, or consonant gemination. Inflection and derivation do not cover the entire field of verbal morphology; for this reason, the last part of this entry is devoted to those issues that are related neither to the former nor to the latter, but rather to the intersection between morphology and phonology, i.e. morphophonology. The scope of the present entry is limited to Classical Arabic, with a few references to colloquial Arabic. For a detailed description of the verb in the latter, see the entries devoted to the different dialects.

2. Inflection

All the inflected forms of a verb together constitute its conjugation. The Arabic tradition distinguishes three subgroups, namely → *māḍī* 'past', *muḍāriʕ* 'the one resembling [the *nomen agentis*]', and *ʾamr* 'order'. Arabists follow this tradition, but they rename the first two 'perfect' and 'imperfect'. The passage from a directly or indirectly temporal terminology to an explicitly aspectual one has resulted in a huge amount of literature on the nature of the correlation between the two forms, → tense and/or → aspect. This issue is dealt with here. Many scholars of Arabic prefer to avoid it by reintroducing the terminology used by scholars of Semitic languages, i.e. 'suffixed form' and 'prefixed form'.

2.1 *Suffixed form*

The suffixed form (Table 1) adds personal, gender, and number markers to the verb stem, which coincides with the pausal form of the 3rd person masculine singular, following the markedness vs. unmarkedness principle. The 1st and 2nd persons are marked by *-t-*, in opposition to the unmarked 3rd person form. There is nonetheless a discrepancy, in the 1st person, between singular and plural ('we' is not a plural of 'I'), and the gender distinction involves not only the 3rd person but the 2nd person as well. The feminine is represented by the marked form (*-t*) in opposition to the unmarked masculine, at least in the singular; in the plural, there is a feminine (*-na*) and a masculine form (*-ū*), which reappears in the 2nd person when followed by an affix (*qataltum* > *qataltum-ū-hu*). The only dual marker is *-ā*, which is added to the singular form in the 3rd person but to the plural form in the 2nd person.

2.2 *Prefixed forms*

The so-called prefixed form (Table 2) adds both prefixes and suffixes to the verb stem. Prefixes mark mainly the person; in the 1st person, though, the person marker *ʾ-/n-* is at the same time a number marker; in the 3rd person feminine singular, the person marker *t-* is at the same time a feminine marker, contrasting with the masculine marker *y-*; the 2nd person masculine singular and 3rd person feminine singular are identical. Elsewhere, *t-* and *y-* are markers of the 2nd and 3rd person, respectively. Suffixes are divided into two series. The first series marks gender and/or number: *-ī* (fem. sg.), *-na* (fem. pl.), *-ū* (masc. pl.), and *-ā* (du.). The second series marks the → 'mood', according to the marked/unmarked principle. This series is itself divided into two subseries: the vowels *-u* and *-a*, which are added to the stem, mark the 'indicative' (*marfūʿ*) and the 'subjunctive' (*manṣūb*) as opposed to ø, marker of the apocopate (*majzūm*). The consonant *-n*, which is added to gender and number markers, vocalized *a* in the feminine singular and masculine plural and *i* in the dual, marks the indicative, while its absence marks the subjunctive and the apocopate. The second series of suffixes is therefore interpreted in terms of moods by Western scholars of Arabic. Yet, the term 'apocopate', directly translated from Arabic, shows their uneasiness vis-à-vis this form. It seems difficult to reconcile its use as a jussive and a conditional with the use of *lam yaqtul* as a negation of *qatala* in assertive sentences. This dilemma is explained diachronically as a legacy of the Akkadian perfect form (Kuryłowicz 1949). In the Arabic tradition, the series of 'modal' endings is analyzed as so-called → *ʾiʿrāb* endings (declensional inflection). This view has the merit of recognizing that at a syntactic level the indicative is a free form, as opposed to the subjunctive and the apocopate, which are bound forms.

Table 1. Inflection of the suffixed verb

	singular	plural	dual
1st	*qatal-t-u*	*qatal-nā*	
2nd masc.	*qatal-t-a*	*qatal-tum*	*qatal-tum-ā*
2nd fem.	*qatal-t-i*	*qatal-tun-na*	*qatal-tum-ā*
3rd masc.	*qatal-(a)*	*qatal-ū*	*qatal-ā*
3rd fem.	*qatal-at*	*qatal-na*	*qatal-ā*

Table 2. Inflection of the prefixed verb

	indicative	subjunctive	apocopate
1st sg.	*ʾa-qtul-u*	*ʾa-qtul-a*	*ʾa-qtul*
2nd masc. sg.	*ta-qtul-u*	*ta-qtul-a*	*ta-qtul*
2nd fem. sg.	*ta-qtul-ī-na*	*ta-qtul-ī*	
3rd masc. sg.	*ya-qtul-u*	*ya-qtul-a*	*ya-qtul*
3rd fem. sg.	*ta-qtul-u*	*ta-qtul-a*	*ta-qtul*
1st pl.	*na-qtul-u*	*na-qtul-a*	*na-qtul*
2nd masc. pl.	*ta-qtul-ū-na*	*ta-qtul-ū*	
2nd fem. pl.	*ta-qtul-na*		
3rd masc. pl.	*ya-qtul-ū-na*	*ya-qtul-ū*	
3rd fem. pl.	*ya-qtul-na*		
2nd du.	*ta-qtul-ā-ni*	*ta-qtul-ā*	
3rd masc. du.	*ya-qtul-ā-ni*	*ya-qtul-ā*	
3rd fem. du.	*ta-qtul-ā-ni*	*ta-qtul-ā*	

2.3 *Imperative*

The imperative (Table 3) coincides with the stem of the imperfect, to which gender and/or number markers are added. The imperative is used only in the 2nd person of the active voice and in the affirmative form. As for the other persons of the active voice and the negative form, periphrastic structures are used, namely the injunctive *li-yaqtul* and the prohibitive *lā yaqtul*. The imperative appears, therefore, in the 2nd person, as a variant of the jussive. In school grammars, the imperative is usually derived from the 2nd person of the apocopate by removing the prefixed person marker *t-*. Once the prefix is removed, if the imperative

Table 3. Inflection of the imperative

masc. sg.	*(u)-qtul*
fem. sg.	*(u)-qtul-ī*
masc. pl.	*(u)-qtul-ū*
fem. pl.	*(u)-qtul-nā*
du.	*(u)-qtul-ā*

starts with CC, a prosthetic *hamza* is added. This *hamza* is called *hamzat al-waṣl*, because it disappears in context (*uqtul/wa-qtul*). Its vowel quality is consistent with that of the stem (*u/u*, but *i/i* or *i/a*).

2.4 *Energicus*

Western scholars of Arabic add to these paradigms another imperfect, called → 'energicus', which can be 'heavy' or 'light'. This actually consists in a suffix *-an* or *-anna* added to the imperfect stem. This suffix can be expanded with gender and number markers and even with the indicative marker *-n*(*a*), but in the latter case, it occurs only with *-ī* and *-ū* markers and with the suffix *-an* (for more details on the formation of the energicus and the alterations related to it, see Ġalāyīnī, *Jāmiᶜ* I, 90–99). Therefore, this suffix provides some 'energ(et)ic' forms, either in the imperative or in the imperfect (mainly jussive), but sometimes in the indicative, too. It never occurs in a context where the subjunctive is likely to be found, although many Arabists tend to link it to the subjunctive because of the vowel *-a* of this suffix. Its Arabic name (*taʾkīd* 'corroboration') stresses its modal value: *udxul* 'come in!'/*udxul-an* 'do come in!/*udxul-anna* 'just come in!'; it belongs almost exclusively to pre-Classical Arabic.

3. DERIVATION

From the derivational point of view, Arabic grammarians as well as Western scholars of Arabic divide verbs first of all into triliteral and quadriliteral ones, depending on whether the → 'root' consists of three or four consonants. Many Arabists believe that the base Form and the 'derived' verb classes of the triliteral and quadriliteral verb are derived from or at least modeled on the root; they even go so far as to include the noun forms related to them. Arabic grammarians deal with this in a completely different way, namely through a double horizontal and vertical derivation: horizontal derivation (→ *ištiqāq*), consisting in deriving from the *maṣdar* (*nomen actionis*) all the inflected forms (> perfect > imperfect > imperative), and from the imperfect, most of the noun forms related to the verb; and vertical derivation (*ziyāda*), consisting in deriving from the base Form (*mujarrad*) all the 'augmented' forms (*mazīd fīhi*). These two approaches show two different conceptions of → derivation and word formation in Arabic: the root/pattern principle of Western scholars of Arabic and the 'from word to word' principle of the Arabic grammarians.

3.1 *The triliteral verb*

3.1.1 The base Form

The base Form can have three possible vocalizations: *faᶜala* (action verbs), *faᶜula* (stative verbs), and *faᶜila* (action or stative verbs). The fourth possible vocalization, *fuᶜila*, is generally viewed as secondary, representing a → 'passive' of the action verb. Nevertheless, this peculiarity prompts us to consider these vocalizations as originally marking not verb classes but rather, by apophony, the different → diatheses of a verb. One can find many vestiges of this in the lexicon of Classical Arabic as well as in modern dialects (e.g. Egyptian). Arabic grammarians were well aware of this, since they provided the example of *ḥazantu-hu fa-ḥazina*. This correlation designates *ḥazina* as a stative-resultative verb in relation to the active form with causative meaning *ḥazana-hu* 'to make someone sad; to sadden someone'. It is therefore better to consider this as being originally the → 'middle voice' ('to sadden' = 'to become sad') of an action verb, reinterpreted as a stative verb ('to be sad'). The above-mentioned correlation allows us to understand what happened: the internal inflection disappeared as the systematic marker of the diathesis, except for the passive, because it was challenged by derivation.

3.1.2 Derived verb classes

In addition to the base Form, numbered I, Western scholars of Arabic list 14 other 'derived' Forms of the triliteral verb, numbered II to XV, 9 of which are considered 'usual' (II–X) and 5 'rare' (XI–XV). Apart from the lack of numbering and some difference in the order, the above-mentioned list comes straight from a small handbook called *Kitāb al-ᶜIzzī* by az-Zanjānī (d. 654/1256–1257), translated in Rome (1610). This handbook is based on the increasing number of *ḥarf*s in the perfect, reinterpreted as number of additional letters in relation to the radicals (Åkesson 2001:113). In the great Arabic grammatical treatises, though, the number of triliteral verb forms is higher (between 20 and 25), because it includes the triliteral verb that is linked (*mulḥaq*; see below) to the quadrilateral. Together with Form I, Forms II to

Table 4. The derived Forms

	perfect	imperfect	imperative	participle	*nomen actionis*
II active	*faʿʿala*	*yufaʿʿil-*	*faʿʿil*	*mufaʿʿil*	*tafʿīl*
passive	*faʿʿala*	*yufaʿʿil-*		*mufaʿʿal*	
III active	*fāʿala*	*yufāʿil-*	*fāʿil*	*mufāʿil*	*fiʿāl* or
passive	*fūʿila*	*yufāʿal-*		*mufāʿal*	*mufāʿala*
IV active	*ʾafʿala*	*yufʿil-*	*ʾafʿil*	*mufʿil*	*ʾifʿāl*
passive	*ʾufʿila*	*yufʿal-*		*mufʿal*	
V active	*tafaʿʿala*	*yatafaʿʿal-*	*tafaʿʿal*	*mutafaʿʿil*	*tafaʿʿul*
passive	*tufuʿʿila*	*yutafaʿʿal-*		*mutafaʿʿal*	
VI active	*tafāʿala*	*yatafāʿal-*	*tafāʿal*	*mutafāʿil*	*tafāʿul*
passive	*tufūʿila*	*yutafāʿal-*		*mutafāʿal*	
VII active	*infaʿala*	*yanfaʿil-*	*infaʿil*	*munfaʿil*	*infiʿāl*
passive	*unfuʿila*	*yunfaʿal-*		*munfaʿal*	
VIII active	*iftaʿala*	*yaftaʿil-*	*iftaʿil*	*muftaʿil*	*iftiʿāl*
passive	*uftuʿila*	*yuftaʿal-*		*muftaʿal*	
IX active	*ifʿalla*	*yafʿal[i]l-*	*ifʿalil*		*ifʿilāl*
X active	*istafʿala*	*yastafʿil-*	*istafʿil*	*mustafʿil*	*istifʿāl*
passive	*ustufʿila*	*yustafʿal-*		*mustafʿal*	

X constitute the well-known chart of the 'ten Forms' of the triliteral verb used by Western scholars of Arabic. Form I having already been dealt with, Table 4 only shows Forms II to X. In order to give an idea of the verbal morphology, the active and passive diatheses of the inflected forms are presented, as well as the main noun forms related to the verb.

This list can easily be organized into a system. From Form I, a first series of Forms is derived: II by doubling of the second radical; III by lengthening of the first radical vowel, and IV by prefixation of ʾ-. From the base Form and the first series of Forms derives a second series of Forms, with the affix *t*: from I, VIII (the *t* is infixed); from II and III, V and VI (the *t* is prefixed); from IV, X (the *t* is infixed and the form keeps the ancient causative prefix *s*-; the relation IV/X, morphologically opaque, is nonetheless syntactically as well as semantically evident). From Form I also derives VII by prefixation of *n*-. This Form is not concerned by the opposition –*t*/+*t*, although it is close to VIII, which replaces it when VII cannot be formed for phonological reasons. Finally, Form IX falls out of the system (see below). Table 5 (Larcher 2003) summarizes the situation.

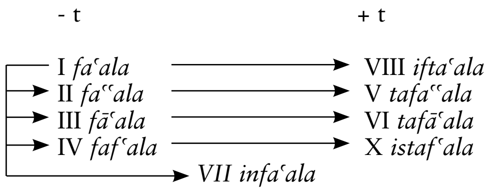
3.1.3 Syntax and semantics

3.1.3.1 Forms without *t*

Form II has, in comparison to Form I, an intensive (or rather, iterative) or a causative value, depending on whether Forms I and II have the same or a different construction. Examples of an intensive value are *ṭāfa*/*ṭawwafa* 'to make one/several circumambulation(s)' (both verbs are intransitive); *qasama-hā* 'to divide something [into two]'/*qassama-hā* 'to share something' (both verbs are transitive). Examples of a causative value are *nazala* 'to go down'/*nazzala-hu/hā* 'to bring someone or something down' (I is intransitive, II becomes transitive); *ʿalima-hā* 'to know something'/*ʿallama-hu -hā* 'to make someone aware of something' (I is simply transitive, while II becomes bitransitive). This is what is called *taʿdiya* 'transitivation' in Arabic grammatical terminology (→ *taʿaddin*).

Form III has, in comparison to Form I, primarily an insistence value (*mubālaġa*). This is the only value when Forms I and III have the same object (e.g. *saʾala-hu* 'to question' and *sāʾala-hu* 'to interrogate'). If Form I is intransitive, Form III becomes transitive and the insistence focuses on the object (e.g. *ḥakama* 'to judge' and *ḥākama-hu* 'to try someone'). Insistence implies repetition, but within continuity. This is what distinguishes it from Form II, which has a repetitive meaning as well, but with discontinuity, e.g. II *ḍaʿʿafa* and III *ḍāʿafa*. Both verbs

Table 5. The relationship between the derived Forms



mean 'to double' in grammatical terminology, but *taḍʿīf* designates the meaningful doubling of the second radical, typical of Form II, whereas *muḍāʿaf* designates the 'doubled' verb with a meaningless doubling, resulting from the identity of the second and third radical. It is, therefore, the idea of repetition within continuity which explains the meaning of implicit reciprocity; indeed, what both these values have in common is the fact that an action is being done in addition to another (e.g. *ṭāraqa an-naʿl* 'to resole a shoe by means of a hammer', but *ṭirāq (an-niʿāl)* 'alternate hammering [of the hooves of a camel]').

Form IV is the causative of Form I, with the same syntactic features as Form II. When Forms IV and II both occur, there is always a difference of meaning between the two (Leemhuis 1977), e.g. *ʾaʿlama/ʿallama-hu -hā* 'to inform someone about something'/'to teach someone something' (the former conveys the general meaning of Form I, the latter, though linked to Form I, is reread through *ʿilm/ʿulūm* 'science(s)'); *ʾaskana/sakkana-hu* 'to lodge someone; to calm someone down' (the former refers to *sakana sakanan* 'to dwell', the latter to *sakana sukūnan* 'to be calm'). In Classical Arabic, Form IV, as in the German *lassen machen*, is neutral between 'to make someone do' and 'to let someone do' (e.g. *ʾaškā-hu* 'to cause someone to complain' or 'to let him complain', and therefore 'to accept, to stop his complaint'). This Form can be implicitly reflexive (e.g. *ʾaslama* 'to become a Muslim' lit. 'to surrender [oneself] to [God]' (*ʾaslama ʾamra-hu ʾilā llāhi*).

3.1.3.2 Forms with t

The affix *t* is the only marker of the reflexive. Classical Arabic knows all types of reflexivity, direct and/or indirect on the syntactic level, passive as well as middle voice (in this latter case, with an agentive or nonagentive subject) on the semantic level.

The reflexivity is direct if the *t-* affixed verb is intransitive or simply transitive, whereas the verb without *t-* is simply or doubly transitive, e.g. IV *ʾaʿadda-hu* 'to equip someone'/X *istaʿadda* 'to equip oneself'; II *ʿallama-hu -hā*/V *taʿallama-hā* 'to become learned, to learn'. This is the equivalent of what the Arabic grammarians call *muṭāwaʿa*, symmetrical to the *taʿdiya*, and not to be confused with reflexivity. In fact, *muṭāwaʿa* excludes indirect reflexivity on the syntactic level but includes, on the semantic level, the resultative correlation which can be marked by Form VII in relation to I (see below) or by Form I in relation to the causative (e.g. *ʿalima-hā* in relation to *ʿallama-hu-hā*).

Reflexivity is indirect if the *t-* affixed verb is simply or doubly transitive, like the verb without *t-*, e.g. IV *ʾaxraja-hu/hā* 'to get someone or something out'/X *istaxraja-hu/hā* 'to get someone or something out for oneself', hence 'to extract something' and 'to ask someone to go out'; IV *ʾaktaba-hu/hā* 'to make someone write something'/X *istaktaba-hu/hā* 'to have someone write something for oneself'.

Reflexivity has a passive meaning if the subject of the *t*-affixed verb is the object of the verb without *t-* and is not the one who 'benefits' from the action designated by this verb, e.g. II *kassara-hā* 'to break something into pieces'/V *takassara* 'to become broken into pieces'.

Otherwise, reflexivity has a middle voice meaning, because the subject of the *t-* affixed verb may be the agent of the action designated by the verb without *t-*, e.g. I *ʾaxaḏa-hā* 'to take something'/VIII *ittaxaḏa-hā* 'to take something for oneself', or may not, as in I *nafaʿa-hu* '(something) benefits someone'/VIII *intafaʿa min-hā* '(someone) benefits from something'.

Syntax often 'forces' derivation: for example, *istaxbara ʿanhā minhu* 'to be informed about something by someone' is a middle reflexive verb with a nonagentive subject, but *istaxbara-hu ʿanhā* becomes a true reflexive-causative with the meaning 'to have someone inform us about something'.

Finally, it is worth noting the existence of doublets such as *basama/ibtasama* 'to smile [intrans.]' or *jahada/ijtahada fī-hā* 'to make an effort toward something' (both verbs have the same construction). This *t-* cannot be explained syntactically. Synchronically, one can analyze them as pure → 'middle verbs'; diachronically, Zaborski (2004) recognizes in these forms a remnant of the Akkadian perfect *iptarasa*.

3.1.3.3 Form VII

Finally, to oversimplify, Form VII is the 'resultative' of Form I (which can be assimilated to a direct reflexive with a passive meaning), e.g. *qasama-hā* 'to divide something'/*inqasama ʾilayhā* 'to be divided into'). This Form shows the tendency to substitute derivation by affixes with internal inflection as a diathesis marker. Actually, Form VII is used in many dialects as a passive of the base form, even when in Classical

Arabic it cannot be formed due to phonological or semantic reasons.

3.1.3.4 Stative vs. action verbs

The same values can be found, *mutatis mutandis*, with the derived forms of stative verbs. For a stative verb 'to be A', where A is an adjective (e.g. *ḥasuna* 'to be good'), one can have II, with the causative meaning 'to make A' (*ḥassana-hu*), and III 'to be A toward someone' (*ḥāsana-hu*); IV is causative as well, but with an explicit or implicit object, which is the *maṣdar* of another verb: thus, *ʾaḥsana-(hā)* means both 'to do something well' and 'to act well'. In the Classical Arabic lexicon, Form IV often has an estimative meaning: 'to find A' (e.g. *ʾakbara-hu* 'to find great (*kabīr*)'). Among the most frequent *t*-forms are the Forms V and X; the former is a direct reflexive with a passive and/or middle meaning, 'to be made A' or 'to prove to be A', depending on whether the object of Form II is someone or something (e.g. *taḥassana*). The latter is a direct and/or indirect reflexive, with a middle meaning, e.g. *istakbara* 'to prove', but *istakbara-hu* 'to find someone or something *kabīr*'.

3.1.3.5 Vocalization/construction of the base Form and derived Forms

A derived Form of the verb can be syntactically or semantically linked to the passive of an action verb and, in the Classical Arabic lexicon, to the different diatheses (see above) and constructions of the same verb. Among these are the type *naqaṣa d-dirhamu wa-naqaṣtu-hu* 'the dirham has decreased and I decreased the dirham', or the type *faraqa-(hā)* 'to part something' or 'to part'. Examples include *ʾawjada-hu* 'to create something' = 'to bring something into existence', the only object of which represents the subject of the passive *wujida* 'to exist'; *kārama-hu* 'to compete with someone in generosity', which refers to *karama-hu* 'to surpass someone in generosity', but not to *karuma* 'to be generous'; *fāraqa-hu* 'to break up with someone', which refers to the intransitive *faraqa*. In this framework, one can easily link diachronically and even synchronically the intensive and causative meanings of Form II; *ḥazzana-hu* and *naqqaṣa-hu* would then be intensive forms in relation to *ḥazana-hu* and *naqaṣa-hu*, but causative forms in relation to *ḥazina* and *naqaṣa; farraqa-hu* can be interpreted both as the intensive of *faraqa-hu* and the causative of *faraqa*.

3.1.3.6 Form IX and 'rare' Forms

The so-called rare Forms appear as derived Forms of stative verbs, remaining at the same time stative verbs. Conversely, in the group of the so-called usual Forms, the derived Forms of stative verbs become action verbs. In this respect, Form IX should be added to these Forms, though labeled as 'usual'. They are directly or indirectly derived from these stative verbs, via the verbal adjective (*ṣifa mušabbaha*) that is connected to them, and whose form seems to determine the derived Form of the verb. When this adjective exists (or remains) alone, one can talk about formally conditioned denominative verbs. This is the case of Form IX *ifʿalla*, and its variant XI *ifʿālla*, well-attested in Maghrebi dialects (Fleisch 1979:319). Presently, these are exclusively linked to a color adjective *ʾafʿal* (e.g. *iswadda* 'to blacken') or to a defect adjective (e.g. *iʿwarra* 'to be or become one-eyed'), but *ʾaswad* is the *ṣifa mushabbaha* of the verb *sawida* 'to be black', of which only the *maṣdar sawād* 'blackness' is used. It is also the case of Form XV *ifʿanlā*, e.g. *iḥbanṭā* 'to be swollen', immediately linkable to *ḥabanṭā* 'swollen, big-bellied', although this latter can be linked to *ḥabiṭa* 'to have a swollen belly'. As for the Forms XII *ifʿawʿala*, XIII *ifʿawwala*, and XIV *ifʿanlala*, they too can be linked to a verb base both directly and indirectly (XIII may be a variant of XII, and XIV another variant of IX), e.g. *iḥdawdaba* 'to be bumpy [road]', cf. *ḥadiba* 'to be hunchbacked'; *ixrawwaṭa* 'to be long and fine [beard]', cf. the expression *maxrūṭ al-liḥya* 'having a beard in which is length without breadth' (the verb is linked to the passive of an action verb which functions as a stative verb); *iqʿansasa* 'to have a hump in front', cf. *qaʿisa* 'id.'. Form IX and the rare Forms constitute a formally and semantically better structured subgroup than is often asserted.

3.1.3.7 Denominative verbs

All verb classes, including Form I in the Classical Arabic lexicon, can be denominative, some of them more than others. The main features of this formation, generally viewed as secondary, may be listed as follows: (i) Morphologically, the nominal base is represented by its root, at least a triplet of consonants extracted from the base. (ii) Semantically, the denominative takes one of the meanings of the deverbative of the same form by adapting it. Thus, for Form II, the most productive denominative form, one passes from 'to cause to perform the action designated by the Form I verb' to 'to do what one does with the object referred to by N[oun]'; for Form III one passes from implicit reciprocity to symmetrical relation; etc. (iii) Denominative formations are in a contrastive way indicated by morphology (e.g. *istajwaba* 'to interview' vs. *istajāba* 'to fulfill'), syntax (e.g. *sallama ʿalayhi* 'to greet' vs. *sallama-hu/hā* 'to preserve'), complementary distribution of the forms (e.g. *ʿawwada-hu* 'to give someone a habit' vs. *ʾaʿāda-hā* 'to repeat something [lit. 'to get something back']').

3.2 *Quadriliteral verb*

Quadriliteral verbs have one base Form (I *faʿlala*) and three augmented Forms: a usual one (II *tafaʿlala*) and two rare ones (III *ifʿanlala* – encoded like the triliteral Form XIV – and IV *ifʿalalla*). The Forms II, III, and IV at first sight appear to be to the quadriliteral verb what the Forms V, VII, and IX are to the triliteral one. However, as is the case with the triliteral verb, one must distinguish between deverbative and denominative formations. While Form II is mainly a deverbative with *t-* of Form I (whatever the origin of I), III appears less often linked to a verbal than to a nominal base, and IV, as a stative verb, is linked to a nominal base, if not completely isolated (e.g. *iḍmaḥalla* 'to disappear'). Morphologically speaking, one may distinguish between three types. The first one is widespread in the lexicon of Classical Arabic, as well as in the dialects. It is the triliteral that Arab grammarians used to 'relate' etymologically to the quadriliteral, rightly or wrongly. This 'relation' is due either to the → reduplication of one of the radicals, or to the insertion of some consonants, mainly *w* and *y* (generating, among others, the types *fawʿala* and *fayʿala*). The second type results from the reduplication of a syllable: it underlies not only onomatopoetic formations but also formations linked to geminate triliteral verbs (e.g. *xaṣxaṣa* 'to privatize'). The third type is formed by derivation from a base with more than three consonants, whatever the nature of this base (word, phrase, or clause) or the nature of this additional radical (e.g. *ta-markaza* 'to concentrate', *salṭana* 'sultanate', *rasmala* 'to capitalize', *basmala* 'to say *bismillāh*', etc.).

3.3 *Nominal forms related to the verb*

Arab grammarians list eight nominal forms related to the verb. One of them is always present for all verbs: the *maṣdar* (→ verbal noun). Its form is almost unique for the derived Forms of the triliteral and quadriliteral verb, but it has multiple forms for the base Form of the triliteral verb. These forms are first and foremost related to the base verb's different vocalizations and constructions, but they also allow for qualification and quantification of the action (the so-called *maṣdar* or *ism al-marra* and the *maṣdar* or *ism an-nawʿ*), e.g. *jalsa* 'a single sitting' and *jilsa* 'a manner of sitting', in relation to *julūs* 'the act of sitting'. Two other nominal forms, active and passive → participles, are never missing for action verbs, either base or derived Forms. The remaining nominal forms concern only the triliteral base verb, either the action one (intensive forms of the active participle, nouns of time/place and instrument) or the stative one (*ṣifa mušabbaha*, of variable form, and → elative, bearing the form *ʾafʿal*). Arab grammarians used to consider these as deverbal forms, except for the *maṣdar*, which they regarded as a 'source' of derivation, for semantic reasons. This is what they actually are, synchronically speaking, including the *maṣdar*: they are all in a more or less tight formal relation with the verb (sometimes vocalized), mainly in its imperfect but sometimes in the perfect, too (this is the case of some *masdar*s, and, in the base Form, of the active participle and of the assimilated adjective). The relation with the perfect is not surprising: diachronically, the ancestors of the *nomen agentis* and assimilated adjectives are the basis of the perfect. Even synchronically, one can sometimes note a coincidence between the perfect stem and either the assimilated adjective, e.g. *fariḥa* 'to be happy (*fariḥ*)', or a form of the *nomen agentis*, e.g. *ḥakam* 'arbiter', *ḥakama* 'to judge'.

4. MORPHOPHONOLOGY

The last section of this entry deals with the issues related either to the presence of a *wāw* or a *yāʾ* among the radicals of the triliteral verb, or to the similarity of the second and third radical. Following the first criterion, Arabic grammarians divided the verb into two categories: *ṣaḥīḥ* 'strong' and *muʿtall* 'weak'. Postclassical handbooks of Arabic divided each category into three subcategories: the first one into *sālim* 'completely strong', *mahmūz* 'hamzated', and *muḍāʿaf* 'geminate'; and the second one into *miṯāl* 'assimilated', *ʾajwaf* 'hollow', and *nāqiṣ* 'defective', according to the position of the *wāw* or the *yāʾ* in the first, second, or third position. Some doubly weak verbs are called *lafīf*. Western Arabic grammar has entirely adopted these categories and subcategories, although the trend has been to oppose, under the name of strong and → weak verbs, the completely strong (*sālim*) verb to the rest, and to overlook hamzated and assimilated verbs, while concentrating on geminate, hollow, and → defective verbs. Although according to the Arabic grammarians the geminate verb belongs to the 'strong' verbs and the 'hollow' and 'defective' verbs belong to the 'weak' verbs, they share the deformation of their stems in comparison to the completely strong verb. This deformation results either in the appearance, next to the 'expected form', of a new stem (through → resyllabication for geminate verbs, complete disappearance of the *w* or *y* for defective verbs), or in the disappearance of the expected stem and, thus, the appearance of two new stems for the hollow verb. Much ink has been poured on this subject, first of all by Arabic grammarians, then by Western scholars of Arabic, and finally by linguists, the members of each category considering themselves more 'scientific' than their predecessors. Actually, they all do the same thing: in terms of generative linguistics, they derive concrete forms from underlying abstract structures (e.g. **radada* > *radda, *qawama* > *qāma, *daʿawa* > *daʿā*, etc.) and sometimes even more abstract structures (e.g. McCarthy 1979). The only things that change are the terminology, the number, the order, and the justification of the operations.

5. Conclusion

Thanks to the simplicity of its inflection and the complexity of its derivation and morphophonology, the Arabic verb continues to fascinate Western scholars of Arabic. The small number of inflected forms results in the high number of their types of use and consequently in some difficulties in the organization of its system according to one or several principles. The apparent complexity of derivational and morphophonological processes is counterbalanced by their remarkable regularity. No doubt much is still to be discovered, not only through internal research but also through the contribution of Arabic dialects as well as other Semitic languages, without neglecting the legacy of the Arabic grammatical tradition.

Bibliographical references

This bibliography is limited to some general references containing broad bibliographies, except for some reference works dealing with specific points.

Primary sources

Ġalāyinī, *Jāmiʿ* = Muṣṭafā al-Ġalāyinī, *Jāmiʿ ad-durūs al-ʿarabiyya*. 3 vols. 10th ed. Beirut and Sidon: al-Maktaba al-ʿAṣriyya li-ṭ-Ṭibāʿa wa-n-Našr, 1386/1966.
Zanjānī, *ʿIzzī* = ʿIzzī ad-Dīn ʾAbū l-Faḍāʾil ʿAbd al-Wahhāb ibn ʾIbrāhīm az-Zanjānī, *at-Taṣrīf al-ʿIzzī*. Rome: Jo. Bapt. Raymundus, 1610.

Secondary sources

Åkesson, Joyce. 2001. *Arabic morphology and phonology based on the* Marāḥ al-ʾarwāḥ *by Aḥmad b. ʿAlī b. Masʿūd presented with an introduction, Arabic edition, English translation and commentary*. Leiden: E.J. Brill.
Cuvalay-Haak, Martine. 1997. *The verb in literary and colloquial Arabic*. (= *Functional Grammar Series*, XIX.) Berlin and New York: Mouton de Gruyter.
Fischer, Wolfdietrich and Otto Jastrow. 1980. *Handbuch der arabischen Dialekte*. Wiesbaden: O. Harrassowitz.
Fleisch, Henri. 1979. *Traité de philologie arabe*. II. *Pronoms, morphologie verbale, particules*. Beirut: Imprimerie Catholique.
Kuryłowicz, Jerzy. 1949. "Le système verbal du sémitique". *Bulletin de la Société Linguistique de Paris* 45:1.47–56.
Larcher, Pierre. 2003. *Le système verbal de l'arabe classique*. Aix-en-Provence: Presses Universitaires de Provence.
Leemhuis, Fred. 1977. *The D and H stems of Koranic Arabic: A comparative study of the function and meaning of the* faʿʿala *and* ʾafʿala *forms in Koranic usage*. Leiden: E.J. Brill.
McCarthy, John Joseph. 1979. *Formal problems in Semitic phonology and morphology*. Ph.D. diss., Massachusetts Institute of Technology. (Publ., Bloomington: Indiana University Linguistics Club, 1982; republ., New York: Garland, 1985.)
Retsö, Jan. 1989. *Diathesis in the Semitic languages: A comparative morphological study*. Leiden: E.J. Brill.
Zaborski, Andrzej. 2004. "Traces of *iptaras* in Arabic". *Egyptian and Semito-Hamitic: Afro-Asiatic studies in memoriam W. Vycichl*, ed. Gábor Takács, 160–171. Leiden: E.J. Brill.

Pierre Larcher (Université de Provence)

Verb Phrase

1. INTRODUCTION

The concept of the verb phrase (VP) is central to contemporary theoretical approaches to Arabic, and, indeed, to modern syntactic theory in general, with its status as a theoretical construct being controversial. The controversy revolves around what is being claimed by saying that a language 'has a VP'. The weak claim is simply that in at least some data types, a discrete constituent consisting of a verb stem and its dependents can be identified. In the case of Arabic, as in many languages, establishing this claim is as simple as finding coordination structures in which an auxiliary verb conjoins two verbs combined with their dependents. The stronger meaning of 'has a VP' is that every expression containing a verb includes a VP category, regardless of whether a discrete VP constituent can be identified in the pronounced form of the expression.

Section 2 considers arguments for the weak version of the claim that Arabic has a VP, and it is shown that Arabic clearly does. Section 3, on the other hand, considers more abstract conceptions of the VP, and arguments that have been made that Arabic has a VP in the strong sense.

2. DIRECT EVIDENCE FOR A VERB PHRASE

A naive definition of 'verb phrase' is a discrete word group consisting of a verb stem and any of its dependents, such as arguments, complements, or certain kinds of adverbial modifiers, such as instrumental, locative, temporal, and aspectual modifiers. The verb is the 'head' of the phrase because it determines the category or type of phrase and because it provides the semantic 'nucleus' of the phrase.

Traditionally, VP constituents are identified on the basis of constituency tests (see Wells 1947; Harris 1951; Zwicky 1978), including coordination, ellipsis, and pro-form substitution. Of these, → sentence coordination is perhaps the most widely used. Coordination involves two word groups being paired as one using a conjunction or disjunction morpheme. Examples of this sort abound in both Standard Arabic (1) and the dialects (2)–(4).

(1) *mā ʾamkana-nī ʾan*
not Perf-be.possible.3ms-cl1s that
ʾaʿmal-a šayʾ-an walā
do.1s-Subj thing-Acc or
ʾaqṭaʿ-a ʾamr-an
finish.1s-Subj matter-Acc
'It was impossible for me to do anything or finish any matter' (Standard Arabic: Wright 1898:II, 327)

(2) *ʿam-b-yrūḥu yiḥku*
Progr-Ind-go.3p speak.3p
w-yitfalsafu w-yithabalu
and-philosophize.3p and-be.stupid.3p
'They're going to talk and philosophize and be stupid' (Levantine Arabic: Maamouri a.o. 2005:321.59)

(3) *rāyiḥ yitfaṣṣax*
going.ms undress.3ms
w-yilbas d-dišdāša
and-put.on.3ms the-dishdasha
'[He's] going to undress and put on his dishdasha' (Gulf Arabic: Qafisheh 1977:181)

(4) *ṛ-ṛažel ma-ka-ixalli-ni la*
the-man not-Progr-let.3ms-cl1s neither
nexrož wala nedrež
go.out.1s nor step.out.1s
'The man won't let me go out at all' (Moroccan Arabic: Harrell 2004:235)

Predicate ellipsis is frequently used as a test for VP-constituents, on the assumption that ellipsis of this kind targets a VP-category. Ellipsis following a tense auxiliary of the sort found in English generally does not seem to occur in Arabic, although Moroccan Arabic appears to be an exception, as in (5), based on Kortobi's (2002) data.

(5) *yasin kan ka-ilʿəb l-kura*
Yasin was.3ms Progr-play.3ms the-ball
w-yusre kan ḥətta huwa
and-Yousre was.3ms even he
'Yasin was playing football and Yousre was too' (Moroccan Arabic: Kortobi 2002:225)

However, other kinds of example can be found in Standard Arabic and the dialects in which an ellipsis occurs in parallel with a verb phrase in a preceding clause, as in (6)–(8).

(6) *ḥāwala* *ʾan* *yaktub-a*
Perf.try.3ms that write.3ms-Subj
ʾanimay *wa-yuxrij-a-hu*
Animé and-direct.3ms-Subj-cl3ms
fī *nafs-i* *l-waqt-i* *lākinna-hu*
in self-Gen the-time-Gen but-cl3ms
lam *yaqdar*
not.past be.able.Juss.3ms
'He tried to write Animé and direct it at the same time, but he wasn't able to' (Modern Standard Arabic: Internet source)

(7) *ʿāyiza* *ʾaʿmal* *kida* *bass* *mā*
want.fs do.1s that but not
b-aʾdar-š
Ind-be.able.1s-Neg
'I want to do that but I can't' (Egyptian Arabic: Internet source)

(8) *kān* *b-yəʾdar* *ykūn*
was.3ms Ind-be.able.3ms be.3ms
ʾawwal *wāḥəd* *b-ṣ-ṣaff*
first one in-the-class
law *rād*
if wanted.3ms
'He could be the first in the class if he wanted' (Syrian Arabic: Stowasser and Ani 1964:34)

Substitution tests involve placement of a pronominal form in a subordinated position that would otherwise be occupied by a verb phrase, and which refers back to a verb-phrase meaning salient in the discourse. Such examples are also easily found in Arabic. For example, in (9), the → demonstrative pronoun *ḏālika* 'that' occurs as the object of *ʿamila* 'to make, do' and refers back to an activity described in a preceding sentence:

(9) *yajibu* *ʾan* *ʾastaʿidd-a*
Ind.must.3ms that prepare.1s-Subj
wa-ʾatamarran-a *min* *jadīd-in*
and-practice.1s-Subj from new-Gen
liʾanna-nī *lam* *ʾaʿmal*
because-cl1s not.Past do.1s.Juss
ḏālika *xilāla* *sanatayn*
that since year-fd.Gen
'I must prepare and practice again because I have not done that in two years' (Modern Standard Arabic: Internet source)

Likewise, (10) shows an example from Syrian Arabic in which the demonstrative *hēk* refers back to action performed by an interlocutor.

(10) *lēš* *sāwēt* *hēk*
why did.2s so
'Why did you [have to] do that?' (Syrian Arabic: Stowasser and Ani 1964:237)

However, the presence of ellipsis and substitution does not present clear, theory-neutral evidence for VP-constituents, given that there are theories of anaphoric reference and ellipsis that do not assume a phrase-structure grammar specification, and for these theories, ellipsis and substitution do not provide evidence for a VP-labeled constituent (see Dalrymple a.o. 1991; Miller 1992).

To summarize, coordination data clearly show that constituents consisting of a verb stem and its dependents occur in both Standard and dialectal Arabic. This shows that Arabic 'has a VP' in the weak sense. The question is therefore whether there is evidence that Arabic 'has a VP' in the strong sense.

3. THE ABSTRACT VERB PHRASE

To see whether Arabic 'has a VP' in the strong sense requires consideration of sentence types other than those mentioned above and, in particular, clauses in VS word order. Much of the modern structuralist literature on Arabic includes analyses of Arabic clauses in VSO word order (→ verbal clause). Given the order of constituents in VS clauses, there is no direct evidence for the presence of a VP constituent properly contained within the root clause constituent. This has led Bakir (1980), Al-Haq (1992), Beller (2006), and others to assume a flat exocentric structure like the one in (11) for VSO clauses:

(11) [$_S$ ʾakala [$_{NP}$ zaydun] [$_{NP}$ tuffāḥatan]]

Researchers working in the Principles and Parameters (Chomsky 1981; Rizzi 1990) and Minimalist Program (Chomsky 1995, 2000, 2001, 2004) traditions have argued that VSO clauses include an abstract VP, which may be only partially pronounced, or unpronounced altogether.

The seminal work on Arabic VP theory is Mohammad (1989), later revised and expanded as Mohammad (2000). Mohammad argues for the presence of an abstract VP in VSO clauses on the basis of evidence indicating hierarchical relationships between the arguments of the verb and, in particular, asymmetries between subject and object arguments in terms of their extraction potential and scope possibilities. The evidence is summarized in (12) and (13) (see → nominal clause for more discussion).

(12) The subject argument is ordered higher than the object in terms of its scope (→ binding) potential.

(13) The object can be extracted more freely than the subject, due to it being a sister to (and hence governed by) the verb.

The implication is that the subject argument in an Arabic VS clause is ordered above the object argument at some level of representation.

A fundamental assumption in Principles and Parameters and the Minimalist Program is that phrase structure encodes hierarchical relations between arguments in addition to dominance and precedence relations between pronounced constituents. In particular, the c-command relation holding between a node A and its sister node B or B's daughter nodes is argued to be the fundamental relation in the definition of syntactic binding constraints or scopal relationships between operators (see, e.g., Heim and Kratzer 1998), and in argument hierarchies (see, e.g., Baker 1988; Grimshaw 1990; Williams 1994). It follows from these assumptions that asymmetries in scope or binding potential between two constituents A and B indicate an asymmetric c-command relation between them, which in turn indicates that they are ordered in terms of a hierarchical structure.

Based on these assumptions, Mohammad concludes from his observations that a VSO clause must include a partially unpronounced VP constituent in which the hierarchical ordering between the subject and the object is established. For example, (11) has the structure in (14), assuming the widely adopted 'VP-internal subject hypothesis' (e.g. Koopman and Sportiche 1991).

(14) [$_{\text{IP}}$ ʾakala$_{\text{i}}$ [$_{\text{VP}}$ [$_{\text{NP}}$ zaydun] [$_{\text{V'}}$ t$_{\text{i}}$ [$_{\text{NP}}$ tuffāḥatan]]]]

Mohammad's analysis has been widely assumed and elaborated on in the extensive literature on Arabic in the Principles and Parameters and Minimalist Program traditions (e.g. Demirdache 1991; Benmamoun 1992, 1997, 2000; Bahloul and Harbert 1993; Eid 1991, 1993; Fassi-Fehri 1993; Diesing and Jelinek 1995; Khalaily 1997; Shlonsky 1997; Harbert and Bahloul 2002; Jelinek 2002; Kortobi 2002; Kremers 2003; Souag 2006).

However, Mohammad's conclusion rests crucially on the assumption that argument and scope hierarchies are encoded in phrase-structural representations (see Williams 1994, Manning 1996, and Sag a.o. 2003 for alternatives). It follows that Mohammad's arguments for an abstract VP are theory-internal.

Gapping in (15)–(17) and right-node raising in (18)–(19) in different varieties of Arabic provide evidence from which support can be adduced for an abstract VP constituent.

(15) *yasin* *kla* *t-təffaḥ*
Yasin ate.3ms the-apple
w-yusre *el-banan*
and-Yousre the-banana
'Yasin ate the apple and Yousre the banana' (Moroccan Arabic: Kortobi 2002: 220)

(16) *suʿād* *ʿand-ha* *karhba*
Souad at-cl3fs car
w-karīm *bisklet*
and-Karim bicycle
'Souad has a car and Karim a bicycle' (Tunisian Arabic: Halila 1992:237)

(17) *aḥmad* *rāḥ* *ʿa-ʿammān*
Ahmad went.3ms to-Amman
w-hišām *ʿa-ʾirbid*
and-Hisham to-Irbid
'Ahmad went to Amman, and Hisham to Irbid' (Jordanian Arabic: elicited)

(18) *xallī* *yilḥag*
let.3ms-cl3ms chase.3ms
yrudd-in-na *aj-jmāl*
return.3ms-to-cl1p the-camels
'Let him chase and bring us back the cam-

els' (North Israel Bedouin Arabic: Rosenhouse 1984:124)

(19) *badu* *ənˁšu* *w-yəḥyu*
began.3p stimulate.3p and-revive.3p
l-ʾadab *l-ˁarabi* *l-ʾadīm*
the-culture.ms the-Arabic.ms the-old.ms

'They began to stimulate and revive the old Arabic culture' (Syrian Arabic: Cowell 1964:392)

The difficulty that gapping and right-node raising pose to a headed phrase-structure grammar (such as the → X-bar theory schema) is the lack of rules like those in (20) and (21), which would admit (15)–(17) and (18)–(19), respectively.

(20) XP → YP ZP

(21) XP → X X YP

Rule (20) is an exocentric rule, as it does not include a head on the right-hand side, while (21) provides two heads. Both violate the assumptions of X-bar theory, and its analogues in more recent Minimalist theory (Chomsky 1995, 2000, 2001, 2004). Because the grammar does not include rules like these, gapping and right-node raising are referred to as 'nonconstituent' coordination, and challenge the assumption that coordination operates on constituents.

Kortobi (2002) analyzes gapping in Moroccan Arabic as resulting from 'across-the-board' (or ATB) movement, a kind of movement operation in which a verbal head that has two identical instances (*modulo* → agreement marking) in two conjuncts can raise as one instance to a higher position in the first clause, 'stranding' the arguments of the second conjunct (22).

(22) $[_{IP}$ kla$_i$ $[_{VP}$ yasin $[_{V'}$ t$_i$ t-təffaḥ]]
w-$[_{VP}$ yusre $[_{V'}$ t$_i$ l-banan]]].

Similarly, right-node raising can be treated as rightward ATB movement of an object argument (23).

(23) $[_{IP}$ ənˁšu$_i$ $[_{VP}$ pro $[_{V'}$ t$_i$ t-təffaḥ]]
w-$[_{VP}$ yusre $[_{V'}$ t$_i$ l-banan]]].

Across-the-board movement out of an abstract VP is also used by Aoun a.o. (1994) in their analysis of first-conjunct → agreement in Lebanese and Moroccan Arabic. They note that, in a clause with a postverbal conjoined subject, the verb can either agree in the plural with both conjuncts, or with the first conjunct only (24).

(24) *kēn/kēno* *karīm*
was.3ms/was.3mp Karim
w-marwān *ˁam-yilˁabo*
and-Marwan Progr-play.3mp
'Karim and Marwan were playing' (Lebanese Arabic: Aoun a.o. 1994:208)

Aoun a.o. (1994) analyze the first-conjunct agreement pattern as involving both gapping and right-node raising: the verb undergoes ATB movement to the left, while the complements or clausal remnant undergoes rightward ATB movement (25).

(25) $[_{IP}$ kēn$_i$ $[_{AuxP}$ $[_{AuxP}$ karīm t$_i$ t$_{VP}$]
w-$[_{AuxP}$ marwān t$_i$ t$_{VP}$]]
$[_{VP}$ ˁam-yilˁabo]]

Although it is not clear how their analysis would account for the plural marking on *ˁam-yilˁabo*, they are able, nevertheless, to account for agreement variation on the initial verb by supposing an abstract VP (for discussion and a different analysis, see Bahloul and Harbert 1993; Harbert and Bahloul 2002).

Eid (1993) uses an abstract VP (generalized to all lexical categories; see also Jelinek 2002 and references) to account for patterns of negation marking in Egyptian Arabic (for more on negation in Arabic, see Ouhalla 1990, 1997, 2002; Benmamoun 1992, 1997, 2000; Al-Tamari 2001; Hoyt 2007). As in many dialects of Arabic, negation is marked by some permutation of the clitics *ma-* and *-š*. These bracket words marked with agreement morphology expressing person, or occur in the free allomorph *miš* 'not' otherwise (26)–(30).

(26) *ma-katab-š*
**miš katab*
'He did not write'

(27) *ma-b-yiktib-š*
miš b-yiktib
'He is not writing'

(28) *miš ḥa-yiktib*
**ma-ḥa-yiktib-š*
'He will not write'

(29) *nadya miš farḥāna*
**nadya ma-farḥāna-š*
'Nadya is not happy'

(30) *ma-ʿand-ī-š migalla*
**miš ʿand-ī-š migalla*
'I don't have a journal'

Eid claims that cliticization of *ma-* and *-š* to a person-marked stem indicates that the stem has raised out of the VP (or other category) to the head of IP, where negation is realized (31)–(32).

(31) [$_{\text{IP}}$ pro$_i$ [$_{\text{I}}$ ma-b-yiktib$_j$-š] [$_{\text{VP}}$ t$_i$ t$_j$]]

(32) [$_{\text{IP}}$ nadya$_i$ [$_{\text{I}}$ miš] [$_{\text{AP}}$ t$_{\text{i}}$ farḥāna]

The optionality in negation marking for the *b*-marked imperfect stems (see also Jelinek 1982) thus indicates optionality in verb raising (33).

(33) [$_{\text{IP}}$ pro$_i$ [$_{\text{I}}$ miš] [$_{\text{VP}}$ t$_i$ b-yiktib]]

The abstract VP therefore allows Eid to capture the correspondence between inflectional properties and negation form (although her analysis unexpectedly implies that the future-marked imperfect stem in (28) cannot raise to tense).

In sum, several prominent studies have been able to account for a wide variety of data by supposing that an abstract VP occurs even in Arabic clauses in which no such constituent can be identified on the basis of a word group.

4. The extended verb phrase

Since Larson (1988), researchers in the Principles and Parameters and the Minimalist Program traditions have argued that what was traditionally referred to as a VP actually has a more articulated structure containing one or more functional categeries dominating the VP proper, which is simply the combination of a verbal head with its complement. Proposed categories include Agr(eement)-O(bject) (AgrO-P; e.g. Fassi-Fehri 1993; Plunkett 1993, 1996; Shlonsky 1997); Chomsky's (1995) 'little-*v*' (*v*P; e.g. Kremers 2003), Aspect (AspP; Diesing and Jelinek 1995; Benmamoun 2000; Kortobi 2002), or Pred(ication) (PredP; Bowers 1993; Harbert and Bahloul 2002). This approach is sometimes called the 'split VP' hypothesis.

The motivations for the different labels vary, but many proposals follow Chomsky (1995) in using 'little-*v*' and assume it to be the head that both licenses an external → theta-role and assigns accusative case to an object argument (some proposals also attribute aspectual meaning to it). Given these assumptions, it has been argued that the presence of an external argument (whether expressed or implicit) and accusative case marking (i.e. licensing a direct object argument) indicate the presence of a VP-complex. This is particularly the case in the analysis of the Arabic → *maṣdar* (Fassi-Fehri 1993; Kremers 2003; → noun phrase).

Additional support for the split-VP hypothesis in Arabic can be adduced from examples like the following, in which a proclitic, such as the indicative *bi-* used widely in the Levant and Egypt (34) (Blau 1960; Cowell 1964; Mitchell and El-Hassan 1994; Eisele 1999; Brustad 2000), the Urban Levantine progressive clitic *ʿam-* (35) (Cowell 1964; Mitchell and El-Hassan 1994; Brustad 2000), or its analog *ka-* used in Moroccan and some western Algerian dialects (36) (cf. Harrell 2004; Benmamoun 2000; Brustad 2000; Kortobi 2002) can take scope over two conjoined verb phrases.

(34) *yaʿni* *n-nās* *bi-taʾkul*
mean.3ms the-people Ind-eat.3fs
wa-tišrab *min* *il-ḥāgāt*
and-drink.3fs from the-things
di
these.fs
'In other words, people eat and drink from these things' (Egyptian Arabic: Internet source)

(35) *hallaʾ* *ṣār-l-na* *sanīn*
now became.3ms-to-cl1p years
ʿam-niḥki *wi-nʾūl* *innu*
Progr-talk.1p and-say.1p that
ḥāliyyan *mudīran* *māl*
actually director Poss
žurž wasūf
George Wasūf
'Now, for years we having been talking and saying that he is actually George Wasūf's director' (Syrian Arabic: Internet source)

(36) *u-fin-emma* *tfekkeṛ* *l-baṛaka*
and-when think.3ms the-blessing
dyal-ha, *ka-ixerrež-ha*
Poss-cl3fs Progr-remove.3ms-cl3fs
w-ibus-ha *w-yaʕṭi*
and-kiss.3ms-cl3fs and-give.3ms
le-ḥbabu *u-l-uladu*
to-friends-cl3ms and-children-cl3ms
ibusu-ha *mʕa-h*
kiss.3mp-cl3fs with-cl3ms
'…and when he thinks of it, he takes it out and kisses it and gives [it] to his friends and children to kiss with him' (Moroccan Arabic: Harrell 2004:242)

Given the assumption that scope relationships correspond to phrase-structural hierarchy and that conjunction joins two similar constituents, these examples imply an analysis in which the aspectual particles head a category that takes a VP complement, perhaps with the morphological attachment of the particle being part of a postsyntactic phonological readjustment process.

BIBLIOGRAPHICAL REFERENCES

Al-Haq, Fawwaz al-Abed. 1992. "Functional or anaphoric control in Jordanian Arabic?". *Language Sciences* 1–2.1–28.

Al-Tamari, Emad Ahmad. 2001. *Sentential negation in English and Arabic: A Minimalist approach.* Ph.D. diss., University of Kansas.

Aoun, Joseph, Elabbas Benmamoun, and Dominique Sportiche. 1994. "Agreement, word order, and conjunction in some varieties of Arabic". *Linguistic Inquiry* 25.195–220.

Bahloul, Maher and Wayne Harbert. 1993. "Agreement asymmetries in Arabic". *Proceedings of WCCFL 11*, ed. David Lightfoot and Norbert Hornsten, 15–31. Stanford, Cal.: CSLI.

Baker, Mark. 1988. *Incorporation.* Chicago: University of Chicago Press.

Bakir, Murtada. 1980. *Aspects of clause structure in Arabic.* Bloomington: Indiana University Linguistics Club.

Beller, Charley. 2006. "The role of argument structure in Arabic impoverished agreement constructions". *University of Rochester Working Papers in the Language Sciences* 3:1. 2–25.

Benmamoun, Elabbas. 1992. *Functional and inflectional morphology: Problems of projection, representation, and derivation.* Ph.D. diss., University of Southern California.

——. 1997. "Licensing of negative polarity items in Moroccan Arabic". *Natural Language in Linguistic Theory* 15.263–287.

——. 2000. *The feature structure of functional categories: A comparative study of Arabic dialects.* New York: Oxford University Press.

Blau, Joshua. 1960. *Syntax des palästinensischen Bauerndialektes Bir Zeits.* Walldorf (Hessen): Verlag für Orientkunde.

Bowers, John. 1993. "The syntax of predication". *Linguistic Inquiry* 24:4.591–656.

Brustad, Kirsten. 2000. *The syntax of spoken Arabic: A comprehensive study of Moroccan, Egyptian, Syrian, and Kuwaiti dialect.* Washington, D.C.: Georgetown University Press.

Chomsky, Noam. 1981. *Lectures on government and binding.* Dordrecht: Foris.

——. 1995. *The Minimalist program.* Cambridge, Mass.: MIT Press.

——. 2000. "Minimalist inquiries: The framework". *Step by step: Essays on Minimalist Syntax in honor of Howard Lasnik*, ed. Roger Martin, David Michaels, and Juan Uriagereka, 89–155. Cambridge: MIT Press.

——. 2001. "Derivation by phase". *Ken Hale: A life in language*, ed. Michael Kenstowicz, 1–52. Cambridge, Mass.: MIT Press.

——. 2004. "Beyond explanatory adequacy". *The cartography of syntactic structures.* III. *Structures and beyond*, ed. Adriana Belletti, 104–131. Oxford: Oxford University Press.

Cowell, Mark. 1964. *A reference grammar of Syrian Arabic.* Washington, D.C.: Georgetown University Press.

Dalrymple, Mary, Stuart M. Shieber, and Fernando C.N. Pereira. 1991. "Ellipsis and higher-order unification". *Linguistics and Philosophy* 14.399–454.

Demirdache, Hamida. 1991. *Resumptive chains in restrictive relatives, appositives and dislocation structures.* Ph.D. diss., Massachusetts Institute of Technology.

Diesing, Molly and Eloise Jelinek. 1995. "Distributing arguments". *Natural Language and Linguistic Theory* 3.123–176.

Eid, Mushira. 1991. "Verbless sentences in Arabic and Hebrew". *Perspectives on Arabic linguistics*, III, ed. Bernard Comrie and Mushira Eid, 31–61. Amsterdam and Philadelphia: J. Benjamins.

——. 1993. "Negation and predicate heads in Arabic". *Principles and predication: The analysis of natural language*, ed. Mushira Eid and Gregory Iverson, 135–152. Amsterdam and Philadelphia: J. Benjamins.

Eisele, John C. 1999. *Arabic verbs in time: Tense and aspect in Cairene Arabic.* Wiesbaden: O. Harrasowitz.

Fassi-Fehri, Abdelkader. 1993. *Issues in the structure of Arabic clauses and words.* Dordrecht: Kluwer Academic Publishers.

Grimshaw, Jane. 1990. *Argument structure.* Cambridge, Mass.: MIT Press.

Hale, Kenneth and Samuel J. Keyser. 1993. "On argument structure and the lexical expression of syntactic relations". *The view from Building 20*, ed. Kenneth Hale and Samuel J. Keyser, 53–109. Cambridge, Mass.: MIT Press.

Halila, Hafedh. 1992. *Subject specificity effects in Tunisian Arabic.* Ph.D. diss., University of Southern California.

Harbert, Wayne and Maher Bahloul. 2002. "Postverbal subjects in Arabic and the theory of agreement". Ouhalla and Shlonsky (2000: 45–70).

Harrell, Richard S. 2004. *A short reference grammar of Moroccan Arabic*. Washington, D.C.: Georgetown University Press.

Harris, Zellig S. 1951. *Structural linguistics*. Chicago: University of Chicago Press.

Heim, Irene and Angelika Kratzer. 1998. *Semantics in generative grammar*. London: Blackwell.

Hoyt, Frederick M. 2007. "An Arabic Wackernagel clitic? The morphosyntax of negation in Palestinian Arabic". *Perspectives on Arabic linguistics*, XX, ed. Mustafa Mughazy, 105–131. Amsterdam and Philadelphia: J. Benjamins.

Jelinek, Eloise. 1982. *On defining categories: AUX and PREDICATE in Egyptian Colloquial Arabic*. Ph.D. diss., University of Arizona.

——. 2002. "Agreement, clitics, and focus in Egyptian Arabic". Ouhalla and Shlonsky (2002:71–105).

Khalaily, Samir. 1997. *One syntax for all categories: Merging nominal atoms in multiple adjunction categories*. Ph.D. diss., Leiden University.

Koopman, Hilda and Dominique Sportiche. 1991. "The position of subjects". *Lingua* 85.211–258.

Kortobi, Ibtissam. 2002. "Gapping and VP-deletion in Moroccan Arabic". Ouhalla and Shlonsky (2002:217–240).

Kratzer, Angelika. 1996. "Severing the external argument from its verb". *Phrase structure and the lexicon*, ed. Johan Rooryck and Laurie Zaring, 109–137. Dordrecht: Kluwer Academic Publishers.

Kremers, Joost M. 2003. *The Arabic noun phrase: A minimalist approach*. Ph.D. diss., University of Nijmegen.

Larson, Richard. 1988. "On the double object construction". *Linguistic Inquiry* 19:3.335–391.

Maamouri, Mohamed, Tim Buckwalter, and Hubert Jin. 2005. *Levantine Arabic QT training data set 4 (speech + transcripts) LDC2005S14*. CD-ROM. Philadelphia: Linguistic Data Consortium.

Manning, Christopher D. 1996. "Argument structure as a locus for binding theory". *Online Proceedings of the First LFG Conference*. www-csli.stanford.edu/publications/LFG/lfg1.html.

Miller, Philip H. 1992. *Clitics and constituents in phrase-structure grammar*. Ph.D. diss., Free University of Brussels.

Mitchell, Terence F. and Shahir A. El-Hassan. 1994. *Modality, mood, and aspect in spoken Arabic, with special reference to Egypt and the Levant*. London and New York: Kegan Paul International.

Mohammad, Mohammad. 1989. *The sentence structure of Arabic*. Ph.D. diss., University of Southern California.

——. 2000. *Word order, agreement, and pronominalization in Standard and Palestinian Arabic*. Amsterdam and Philadelphia: J. Benjamins.

Ouhalla, Jamal. 1990. "Sentential negation, relativized minimality, and the aspectual status of auxiliaries". *Linguistic Review* 7.183–231.

——. 1994. "Verb movement and word order in Arabic". *Verb movement*, ed. David Lightfoot and Norbert Hornstein, 41–72. Cambridge: Cambridge University Press.

——. 1997. "Remarks on focus in Standard Arabic". *Perspectives on Arabic linguistics*, X, ed. Mushira Eid and Robert Ratcliffe, 31–61. Amsterdam and Philadelphia: J. Benjamins.

——. 2002. "Negative sentences in Arabic". Ouhalla and Shlonsky (2002: 299–320).

——. and Ur Shlonsky (eds.). 2002. *Themes in Arabic and Hebrew syntax*. Dordrecht: Kluwer Academic Publishers.

Plunkett, Bernadette. 1993. "The position of subjects in Modern Standard Arabic". *Perspectives on Arabic linguistics*, V, ed. Mushira Eid and Clive Holes, 231–260. Amsterdam and Philadelphia: J. Benjamins.

——. 1996. "Binding in Arabic and the Lexical Clause". *Perspectives on Arabic linguistics*, VIII, ed. Mushira Eid, 86–97. Amsterdam and Philadelphia: J. Benjamins.

Qafisheh, Hamdi A. 1977. *A short reference grammar of Gulf Arabic*. Tucson: University of Arizona Press.

Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge, Mass.: MIT Press.

Rosenhouse, Judith. 1984. *The Bedouin Arabic dialects: General problems and a close analysis of North Israel Bedouin dialects*. Wiesbaden: O. Harrassowitz.

Sag, Ivan A., Tom Wasow, and Emily Bender. 2003. *Syntactic theory: A formal introduction*. Stanford, Calif.: CSLI Publications.

Schmidt, Hans and Paul Kahle. 1918. *Volkserzählungen aus Palästina*, I. Göttingen: Vandenhoek und Ruprecht.

——. 1930. *Volkserzählungen aus Palästina*, II. Göttingen: Vandenhoek und Ruprecht.

Shlonsky, Ur. 1997. *Clause structure and word order in Hebrew and Arabic: An essay in comparative Semitic syntax*. New York: Oxford University Press.

Souag, Mostafa Lameen. 2006. *Explorations of the syntactic cartography of Algerian Arabic*. M.A. thesis, University of London.

Stowasser, Karl and Moukhtar Ani. 1964. *A dictionary of Syrian Arabic: English-Arabic*. Washington, D.C.: Georgetown University Press.

Wells, Rulon S. 1947. "Immediate constituents". *Language* 23.81–117.

Williams, Edwin. 1994. *Thematic structure*. Cambridge, Mass.: MIT Press.

Wright, William. 1898. *A grammar of the Arabic language*. 3rd rev. ed. W. Robertson Smith and M.J. de Goeje. 2 vols. Cambridge: Cambridge University Press.

Zwicky, Arnold M. 1978. "Arguing for constituents". *Proceedings of the 14th Regional Meeing of the Chicago Linguistic Society*, 503–512. University of Chicago: Chicago Linguistics Society.

Frederick M. Hoyt
(University of Texas at Austin)

Verbal Clause

1. Introduction

The term 'verbal clause' (*jumla fiʿliyya*) is taken from traditional Arabic grammatical theory, and is used in contrast to → nominal clause (*jumla ismiyya*). While the status of both terms in contemporary Western linguistic theory is unclear, the verbal clause seems to have elicited less theoretical interest than nominal clauses have, except with relation to → agreement phenomena. This entry presents a comparison of two different ways in which the term 'verbal clause' has been used, then examines its role in the discussion of word order in Arabic and how it is represented in different theoretical frameworks.

In what follows, SV word order is referred to as 'agent-initial' (or A-initial) word order, so as to avoid taking sides in the theoretical debate over the structure of nominal clauses and, in particular, over the issue of whether the S is actually a grammatical subject or a grammatical topic. Accordingly, the S in SV word order is referred to herein as the 'initial agent', and VS word order is referred to as 'verb-initial' (or V-initial) word order, so as to include within the category → 'pro-drop' clauses with implicit or anaphoric subjects.

2. What is a verbal clause?

'Verbal clause' is used with at least two senses. Some sources use it to describe clauses in V-initial word order (Wright 1898:II, 251; Cantarino 1974–1975:I, 41; Bakir 1980:125; Plunkett 1993, 1996; Badawi a.o. 2004:344). Elsewhere, 'verbal clause' describes clauses that have a conjugated verb stem as → predicate, referred to here as 'verbally headed' (or V-headed) clauses (Abu-Haidar 1979:123; Rosenhouse 1984:127; Cowell 1964:407; Qafisheh 1977:203; Belyayeva 1997:52; Fassi-Fehri 1993:47, 87; Shahin 2000:39; Al-Tamari 2001:9–10, 35).

The V-initial sense of 'verbal clause' describes a clause the nucleus of which begins with a verb or, in some cases, with a participle or an adjective, all of which show some form of → agreement marking. The verb either precedes an independent NP expressing its agent, or expresses its agent by means of agreement morphology.

In contrast, the nucleus of a nominal clause begins with a nominal expression (*mubtadaʾ* 'that which is begun with', referred to here as 'initial NP'; → *ibtidāʾ*) of which the remainder of the clause (→ *xabar* 'news, report, comment') is predicated. Also included in the contrast is the locative clause (*jumla ḏ̣arfiyya* 'locative clause'; → *mafʿūl fīhi*), the nucleus of which begins with a locative prepositional phrase. The V-initial sense focuses on the order in which the elements of a predication are presented, and therefore may imply a theory of syntax concerned with 'information structure', namely, how syntactic structure encodes transactional or pragmatic meaning.

According to the 'V-headed' sense, a verbal clause is headed by a verb stem, regardless of its position relative to its agent argument. In contrast, a nominal clause is a copular clause (→ copula), the predicate of which is a → noun phrase, → adjective phrase, or prepositional phrase. The contrast here is not between different orders of presentation but rather between different kinds of predicate, and may imply a theory more concerned with thematic or descriptive meaning.

The V-initial usage treats (1)–(2) as verbal clauses and (3)–(4) as nominal clauses, while the V-headed usage treats (1)–(3) as verbal clauses and (4) as a nominal clause.

(1) *ʿabar-na* *ṭ-ṭarīq-a*
cross.Perf.1p the-street-Acc
l-ʿarīḍ-at-a
the-wide-fs-Acc
'We crossed the wide street' (Badawi a.o. 2004:344)

(2) *lam* *yatawāfar*
Neg.Past to.be.available.Imperf.3ms
l-ī *hāḏ-āni*
to-cl1s these.md.Nom
š-šarṭ-āni
the-conditions.md.Nom
'These two conditions were not available to me' (Badawi a.o. 2004:347)

(3) *hāḏāni* *š-šarṭ-āni* *lam*
these.md the-conditions.md Neg.Past
yatawafar-ā
to.be.available.Imperf.3md

l-ī
to-cl1s
‘These two conditions were not available to me’ (Badawi a.o. 2004:347)

(4) *al-masʾal-at-u* *basīṭ-at-un*
the-question-fs-Nom simple-fs-Nom
‘The question is simple’ (Badawi a.o. 2004:307)

From the perspective of contemporary Western linguistics, the V-initial usage captures certain generalizations more perspicuously than does the V-headed usage, and draws finer distinctions between clause types. The first generalization is that initial NPs in clauses headed by any kind of predicate, whether verbal or otherwise, are interpreted as being ‘specific’ (→ specificity). Treating both copular clauses and A-initial clauses as kinds of nominal clause captures this restriction directly, while conflating clauses in A-initial word order with those in V-initial word order does not.

The second generalization is the celebrated agreement asymmetry between A-initial and V-initial clauses. In A-initial word order, the verb is inflected as matching the agent in gender and number, while in V-initial word order with an independent agent NP, the verb is inflected as matching the agent in gender only. This generalization is easily captured by treating A-initial and V-initial clauses as different types, whereas treating them as belonging to the same type requires auxiliary statements to capture the agreement asymmetry (→ nominal clause).

According to some theoretical analyses, even present-tense copular clauses (→ copula) include a verb phrase headed by an abstract or ‘null’ copula (Bakir 1980; Fassi-Fehri 1993; Eisele 1999; see Benmamoun 2000 for criticism), as in (5)–(6).

(5) *al-masʾal-at-u* *basīṭ-at-un*
the-question-fs-Nom simple-fs-Nom
‘The question is simple’

(6) [$_{\text{IP}}$ *al- masʾalat-u* [$_{\text{I}}$, BE$_{\text{J}}$ [$_{\text{VP}}$ t_j
[$_{\text{V'}}$ *basīṭat-un*]]]]

This kind of analysis treats all Arabic clauses as verbal clauses, including those that do not contain a pronounced verb stem, with the consequence that the term ‘verbal clause’ becomes synonymous with the term ‘clause’ itself and makes no useful distinction between clause types. Accordingly, using ‘verbal clause’ for describing clauses in V-initial word order allows well-known generalizations about Arabic word order to be captured directly and is therefore more useful.

3. THE VERBAL CLAUSE AND ARABIC WORD ORDER

V-initial word order arises frequently in discussion of ‘basic’, ‘default’, or ‘neutral’ → word order order in different varieties of Arabic. Where Standard Arabic is described as having a basic word order, it is almost always claimed to be V-initial word order (Fassi-Fehri 1993; Mohammad 2000; Badawi a.o. 2004:345), although in some contexts, such as newspaper headlines, A-initial word order predominates (7) (Dahlgren 1998; Badawi a.o. 2004:349).

(7) *ʿiṣāb-at-tun* *musallaḥ-at-un*
band-fs-Nom armed-fs-Nom
taʾsiru *waḥd-at-a*
capture.Imperf.3fs unit-fs-Acc
sawārīx-a *rūsiyy-at-an*
rocket.mp-Gen Russian-fs-Acc
‘Armed band captures Russian rocket unit’ (Badawi a.o. 2004:349)

V-initial word order is also reported to be the basic order in dialects including Moroccan (Harrell 2004:160), Jordanian (Al-Haq 1992:2), varieties of Palestinian (Shlonsky 1997:7–8; Shahin 2000:38–39), Najdi Bedouin (Ingham 1994:38), and Syrian Arabic (Grotzfeld 1965: 98–99). In these sources, A-initial word order is described as → topicalization of the agent argument (Ingham 1994:38; Harrell 2004:160; Shahin 2000:38–39), or as expressing a greater degree of ‘emphasis’ on the agent argument (Grotzfeld 1965:98–99; Abu-Haidar 1979).

For example, Harrell (2004:160) reports that in Moroccan Arabic, the subject usually follows the predicate in a simple sentence:

(8) *žaw* *ḍ-ḍyaf*
come.Perf.3mp the-guests
‘The guests have come’

With regard to agent-initial word order, Harrell reports that “in many cases a noun which seems to be the subject of a following verb is separated from the verb by a slight pause and is to be interpreted as a prestated topic rather than directly as the subject":

(9) *le-mġaṛba* *ka-yaklu*
the-Moroccans Progr.-eat.Imperf.3p
tlata *wella* *ṛebʿa* *de-l-meṛṛat*
three or four of-the-times
fe-n-nhaṛ
in-the-day
'The Moroccans, they eat three or four times a day'

A-initial word order is reported to be basic in dialects including Moroccan (Kortobi 1997:218; Benmamoun 2000; in contrast to Harrell 2004) and Libyan Arabic (Owens 1984:212–215).

El-Yasin (1985) presents an interesting argument in favor of treating A-initial order as basic in Northern Jordanian Arabic (spoken in the Irbid region of northwestern Jordan). El-Yasin reports that both V-headed and copular clauses have a progressive limit on the number of initial NPs with which they can begin, with two being the largest felicitous number, three being deviant, and four entirely unacceptable. For example, the NP headed by *ṭullāb* 'students' in (10) is a single initial-NP. The paraphrases in (11)–(13) break up this complex NP into longer sequences of simpler initial NPs, with a corresponding reduction in acceptability.

(10) [*ṭullāb* *mʿallim* *midrast* *il-giryih*]
students teacher school the-village
mlāḥ *rāḥū* *š-šām*
good.mp went.3p Damascus
'The village schoolteacher's students are good/went to Damascus
(11) [*mʿallim midrast il-giryah*] [*ṭullāb-u*] *mlāḥ/rāḥu š-šām*
(12) ?[*midrist il-giryah*] [*mʿallim-ha*] [*ṭullāb-u*] *mlāḥ/rāḥu š-šām*
(13) *[*il-giryih*] [*midrasit-ha*] [*mʿallim-ha*] [*ṭullāb-u*] *mlāḥ/rāḥu š-šām*

Assuming that the NPs headed by *ṭullābu* 'his students' in the copular versions of (10)–(13) are subjects (Jelinek 1981; Fassi-Fehri 1988; Eid 1991; Plunkett 1993; Mohammad 1989), El-Yasin concludes from the fact that the two sentence types seem to tolerate the same number of initial NPs that *ṭullābu* in the V-headed versions of (10)–(13) must be subjects as well, and that Northern Jordanian Arabic has A-initial word order as its basic word order. El-Yasin's argument rests on the assumption that the initial NPs in (10)–(13) are grammatical subjects. However, this assumption is not innocent because, according to the V-initial usage of 'verbal clause', (10)–(13) are all nominal clauses. Therefore, El-Yasin's observation has no theory-independent implications about the word order in Northern Jordanian Arabic.

Some sources, rather than asserting that one word order or the other is basic, report variation between the two (see for Nablus Palestinian Arabic, Belyayeva 1997:52–53; and for Jordanian Arabic, Al-Haq 1992:2). The variation is often reported as being conditioned by various grammatical or pragmatic factors. The first of these is a generalization that has been made for several varieties of colloquial Arabic that V-initial word order is more common with verbs in the perfect stem, while A-initial order is more common with verbs in the imperfect (see → word order for discussion). This has been noted for Jordanian, Egyptian, Moroccan, Syrian, and Saudi Arabian Arabic (Al-Tamari 2001), Moroccan Arabic (Benmamoun 2000:62–63; 2003), Najdi Bedouin Arabic of Saudi Arabia (Ingham 1994:38–39), and Syrian Arabic (Cowell 1964:409; Ambros 1977:79–81).

For example, Al-Tamari (2001:19–21) claims that in Jordanian Arabic, agent-initial order (14) is preferred with a verb in the imperfect stem, whereas V-initial order is preferred with a verb in the perfect stem (15), although the opposite word orders are acceptable with each:

(14) *le-wlād* *b-yilʿabu* *fi*
the-boys Progr-play.Imperf.3mp in
s-sāḥa
the-field
'The boys are playing in the field'
(15) *ʾakal* *māhir* *iṭ-ṭabīx*
ate.Perf.3ms Maher the-food
'Maher ate the food'

Benmamoun (2003:759) suggests a historical explanation for the preference of V-initial word order with perfect verb stems. He notes that such preferences also arise in so-called God-wishes (after Ferguson 1983), in which a verb in the perfect stem precedes the subject, whereas in paraphrases of these expressions verbs in the imperfect stems follow the subject, as in examples (16)–(19) from Moroccan Arabic.

(16)	*jāzā-ka*		*llāh*
	reward.Perf.3ms-cl.2ms		God
	‘May God reward you!’		
(17)	*llāh*	*yjāzī-k*	
	God	reward.Imperf.3ms-cl2ms	
	‘May God reward you!’		
(18)	*llāh*	*ybārik*	*fī-k*
	God	bless.Imperf.3ms	in-cl2ms
	‘May God bless you!’		
(19)	*bārak*	*llāhu*	*fī-k*
	bless.Perf.3ms	God	in-cl2ms
	‘May God bless you!’		

Noting that formulaic expressions often retain properties of an earlier stage of a language, Benmamoun points to the parallel between V-initial word order and the word order in the → construct state. He suggests that this asymmetry may be a residue of a stage of pre-Classical Arabic, at which the perfective stems were verbal nouns that combined with their subjects in a construct state.

Another widely noted generalization is that A-initial word order disambiguates subjects and objects, when the agreement form and meaning of the verb fail to do so. For example, Souag (2006:23–37) reports that in the Dellys dialect of Algerian Arabic, the preference for SV in clauses containing independent NP objects in examples like (20) is “overwhelming”.

(20)	*u-ṭŭlba*	*yətʿəllmu*	*fi-ha*
	and-students	learn.Imperf.3p	in-cl3fs
	l-qŭrʾan		
	the-Qurʾān		
	‘And students learn the Qurʾān in it’		

In contrast, if the verb hosts an encliticized object pronoun, or if it is intransitive, V-initial word order is preferred, as in (21).

(21)	*yəqraw*	*fi-ha*	*ṭ-ṭəlba*
	read.Imperf.3p	in-cl.3fs	the-students
	‘The students study in it’		

Souag (2006:23) claims that this alternation is not a matter of transitivity because the unmarked position for the subject of both intransitive verbs and transitive verbs hosting object clitics is following the verb.

Blau (1960:207–208), analyzing data from Bir Zeit rural Palestinian Arabic (Schmidt and Kahle 1918, 1930), notes that the grammatical ambiguity between the subject and object is resolved through the use of agreement marking on the verb (which indicates the agent), particularly through the use of ‘linear-progressive word order’ placing the agent before the object. He notes that, perhaps for this reason, A-initial word order “becomes ever more frequent”, although V-initial word order is still quite ordinary. Very similar generalizations are reported by Abu-Haidar (1979:126) for the Baskinta variety of Lebanese Arabic, and by Al-Tamari (2001) for Jordanian Arabic.

An even more important influence on word order is the discourse genre in which a sentence is uttered, and the kind of information that the sentence contributes to the discourse. Numerous sources report that indefinite or referentially nonspecific subject NPs tend to follow the verb, including Damascene Syrian Arabic (Cowell 1964:407–411; Ambros 1977:79–81), the Nablus dialect of Palestinian Arabic (Belyayeva 1997), and Gulf Arabic (in particularly Bahraini; Qafisheh 1977:202–203). This correlates with a strong tendency in Arabic for new information to be expressed later in the sentence (Moutaouakil 1989), and for indefinite NPs to express new information, whether introduced as a new discourse topic or as background information (Brustad 2000).

A very widely noted generalization is that V-initial word order expresses particular kinds of ‘information structure’, meaning that it is used to express particular instructions about how the descriptive information expressed in a clause is to be incorporated into a discourse context. V-initial word order, particularly in the case of → ‘pro-drop’ clauses, is used to introduce an event into a discourse, and is most frequently used to describe sequences of events in narrative genres, such as storytelling or news reporting (Khan 1988; Moutaouakil 1989; Fakhri 1995; Dahlgren 1998; Brustad 2000).

This very broad overview shows that, although both A-initial and V-initial word orders are asserted as being basic in various dialects, both are attested, and the preference for one or the other is affected by a variety of grammatical and pragmatic factors. In general, V-initial word order, and hence verbal clauses, retains a number of functions both grammatically and pragmatically. These include:

(i) expressing existential or presentative predication
(ii) expressing new-information focus on the subject or on the sentence as a whole
(iii) introducing an eventuality into a discourse

The fact that V-initial word orders are used for multiple functions suggests that verbal clauses as a sentence type are still alive and well in the Arabic dialects, however different theoretical frameworks and methodologies may analyze them. As Brustad (2000:328) notes, "[V-initial] typology can be shown to be prominent in spoken Arabic narratives, and remains a basic word order of the language".

4. THE VERBAL CLAUSE IN CONTEMPORARY SYNTACTIC THEORY

Verbal clauses have uncertain status in contemporary Western syntactic theory. Some analyses make use of nonheaded or 'flat' structure, derived from context-free rewrite rules, such as in (22), which contrast with a rule for nominal clauses such as in (23) (subscripted *subj* and *obj* indicate non-case-marked NPs, and subscripted *case* indicating any case marking).

(22) S → V (NP$_{\text{subj}}$) (NP$_{\text{obj}}$)
S → V (NP$_{\text{case}}$)*
(23) S → NP$_{\text{nom}}$ V NP$_{\text{obj}}$

This approach is found in earlier stages of transformational grammar (Bakir 1980; Jelinek 1981), in Lexical Functional Grammar (Al-Haq 1992), and in Head-Driven Phrase Structure Grammar (Chahristan 1991; Moosally 1994; Borsley 1995; Beller 2006). Each of these approaches allows verbal clauses to be identified with a particular rule schema, and therefore to be represented directly as a theoretical concept, rather than merely as a descriptive label.

In the Government and Binding and the Minimalist traditions (Chomsky 1981, 1986, 1995, 2000, 2001), the verbal clause has no independent theoretical status (Mohammad 1989, 2000; Eid 1991, 1993; Bahloul and Harbert 1993; Aoun a.o. 1994; Benmamoun 1992, 2000; Fassi-Fehri 1993; Ouhalla 1994; Diesing and Jelinek 1995; Shlonsky 1997; Harbert and Bahloul 2002, among others). A major reason for this is that these frameworks make extensive use of 'null' or unpronounced abstract structure, and as such, 'word order' as it is traditionally known does not correspond directly to constituent order (Kayne 1994; Chomsky 1995:334; Zubizarreta 1998; Kremers 2003).

In these frameworks, it is almost universally assumed that the underlying structure of a verbal clause like (24) is a structure like (25), from which V-initial word order is derived by head-movement of the verb to the head of an inflectional projection (26) (→ verb phrase). A-initial order is then derived by further movement of the subject NP into the IP projection (27).

(24) *yuḥibbu* *ʿamr-un* *ʿalāʾ-an*
love.Imperf.3ms Amr-Nom Alaa-Acc
'Amr loves Alaa'
(25) [$_{\text{IP}}$ I [$_{\text{VP}}$ *ʿamr-un* [$_{\text{V'}}$ *yuḥibbu ʿalāʾ-an*]]]
(26) [$_{\text{IP}}$ [$_{\text{I}}$ I *yuḥibbu*$_{\text{i}}$] [$_{\text{VP}}$ *ʿamr-un* [$_{\text{V'}}$ t$_{\text{i}}$ *ʿalāʾ-an*]]]
(27) [$_{\text{IP}}$ *ʿamr-un*j [$_{\text{I'}}$ [$_{\text{I}}$ I *yuḥibbu*$_{\text{i}}$] [$_{\text{VP}}$ t$_{\text{j}}$ [$_{\text{V'}}$ t$_{\text{i}}$ *ʿalāʾ-an*]]]

How these movement operations are motivated varies from proposal to proposal, but they share the implication that a verbal clause is one in which raising of the verb to I is motivated, while raising of the subject is not. A verbal clause, then, describes the pronounced word orders resulting from derivations of this kind.

A variation on this approach reduces all verbal clauses to A-initial order by supposing that in V-initial order, an abstract or 'null' expletive pronoun is present in the IP projection, and that the agreement marking on the verb is 'split' between the expletive pronoun and the subject NP (Mohammad 1989, 2000), as in (28)–(29).

(28) *yaltamisu* *l-luġawiyy-ūna*
seek.Imperf.3ms the-linguists-mp.Nom
tafsīr-an
explanation-Acc
'The linguists are seeking an explanation'
(Badawi a.o. 2004:352)
(29) [$_{\text{IP}}$ pro$_{\text{3P}}$ [$_{\text{I'}}$ *yaltamisu*$_{\text{j}}$
[$_{\text{VP}}$ *l-luġawiyyūna*$_{3M}$ [$_{\text{V'}}$ t$_{\text{j}}$ *tafsīr-an*]]]]

According to this approach, verbal clauses do not exist as a particular kind of syntactic structure but rather are simply an artifact of how the structure is pronounced or 'spelled out' (see also Aoun and Benmamoun 1998; Benmamoun 2000, 2003; Benmamoun and Lorimor 2006 for a more recent variation on this approach).

BIBLIOGRAPHICAL REFERENCES

Abu-Haidar, Farida. 1979. *A study of the spoken Arabic of Baskinta*. Leiden: E.J. Brill.

Al-Haq, Fawwaz al-Abed. 1992. "Functional or anaphoric control in Jordanian Arabic?". *Language Sciences* 1–2.1–28.

Al-Tamari, Emad Ahmad. 2001. *Sentential negation in English and Arabic: A minimalist approach*. Ph.D. diss., University of Kansas.

Aoun, Joseph, Elabbas Benmamoun, and Dominique Sportiche. 1994. "Agreement, word order, and conjunction in some varieties of Arabic". *Linguistic Inquiry* 25.195–220.

Badawi, Elsaid, Michael G. Carter, and Adrian Gully. 2004. *Modern Written Arabic: A comprehensive grammar*. London: Routledge.

Bahloul, Maher and Wayne Harbert. 1993. "Agreement asymmetries in Arabic". *Proceedings of the Eleventh West Coast Conference on Formal Linguistics (WCCFL 11)*, ed. David Lightfoot and Norbert Hornsten, 15–31. Stanford, Cal.: Center for the Study of Language and Information.

Bakir, Murtada. 1980. *Aspects of clause structure in Arabic*. Bloomington: Indiana University Linguistics Club.

Beller, Charley. 2006. "The role of argument structure in Arabic impoverished agreement constructions". *University of Rochester Working Papers in the Language Sciences* 3:1.2–25.

Benmamoun, Elabbas. 1992. *Inflectional and functional morphology: Problems of projection, representation, and derivation*. Ph.D. diss., University of Southern California.

——. 2000. *The feature structure of functional categories: A comparative study of Arabic dialects*. New York: Oxford University Press.

——. 2003. "Agreement parallelism between sentences and noun phrases: A historical sketch." *Lingua* 113.747–764.

——. and Heidi Lorimor. 2006. "Featureless expressions: When morphophonological markers are absent". *Linguistic Inquiry* 37.1–23.

Blau, Joshua. 1960. *Syntax des palästinensischen Bauerndialektes Bir Zeits*. Walldorf (Hessen): Verlag für Orientkunde.

Bolotin, Naomi. 1995. "Arabic and parametric VSO agreement". Eid and Parkinson (1995:7–27).

Borsley, Robert D. 1995. "On some similarities and differences between Welsh and Syrian Arabic". *Linguistics* 33.99–122.

Cantarino, Vicente. 1974–1975. *Syntax of modern Arabic prose*. 3 vols. Bloomington: University of Indiana Press.

Chahristan, Kamil F. 1991. *Aspects of phrases and clauses in Syrian within the framework of Head-driven Phrase Structure Grammar*. Ph.D. diss., University of Wales.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

——. 1986. *Knowledge of language*. New York: Praeger.

——. 1995. *The Minimalist program*. Cambridge, Mass.: MIT Press.

——. 2000. "Minimalist inquiries: The framework". *Step by step: Essays on Minimalist syntax in honor of Howard Lasnik*, ed. Roger Martin, David Michaels, and Juan Uriagereka, 89–155. Cambridge, Mass.: MIT Press.

——. 2001. "Derivation by phase". *Ken Hale: A life in language*, ed. Michael Kenstowicz, 1–52. Cambridge, Mass.: MIT Press.

Cowell, Mark. 1964. *A reference grammar of Syrian Arabic*. Washington, D.C.: Georgetown University Press.

Diesing, Molly and Eloise Jelinek. 1995. "Distributing arguments". *Natural Language and Linguistic Theory* 3.123–76.

Eid, Mushira. 1991. "Verbless sentences in Arabic and Hebrew". *Perspectives on Arabic linguistics*, III, ed. Bernard Comrie and Mushira Eid, 31–61. Amsterdam and Philadelphia: J. Benjamins.

——. 1993. "Negation and predicate heads in Arabic". *Principles and predication: The Analysis of natural language*, ed. Mushira Eid and Gregory Iverson, 135–152. Amsterdam and Philadelphia: J. Benjamins.

——, and Dilworth Parkinson (eds). 1995. *Perspectives on Arabic linguistics*, VII. Amsterdam and Philadelphia: J. Benjamins.

Eisele, John C. 1992. "The category of AUX and Cairene Arabic auxiliaries". *Perspectives on Arabic linguistics*, IV, ed. Ellen Broselow, Mushira Eid, and John McCarthy, 143–165. Amsterdam and Philadelphia: J. Benjamins.

——. 1999. *Arabic verbs in time: Tense and aspect in Cairene Arabic*. Wiesbaden: O. Harrassowitz.

El-Yasin, Mohammed. 1985. "Basic word order in Classical Arabic and Jordanian Arabic". *Lingua* 65.107–122.

Fakhri, Ahmad. 1995. "Topic continuity in Arabic narrative discourse". Eid and Parkinson (1995:141–155).

Fassi-Fehri, Abdelkader. 1988. "Agreement in Arabic, binding, and coherence". *Agreement in natural language: Approaches, theories, description*, ed. Michael Barlow and Charles Ferguson, 107–158. Stanford, Cal.: Center for the Study of Language and Information.

——. 1993. *Issues in the structure of Arabic clauses and words*. Dordrecht: Kluwer Academic Publishers.

Grotzfeld, Heinz. 1965. *Syrisch-Arabisch grammatik*. Wiesbaden: O. Harrassowitz.

Harbert, Wayne and Maher Bahloul. 2002. "Postverbal subjects in Arabic and the theory of agreement". Ouhalla and Shlonsky (2002:45–70).

Harrell, Richard S. 2004. *A short reference grammar of Moroccan Arabic*. Washington, D.C.: Georgetown University Press. (1st ed., 1962.)

Ingham, Bruce. 1994. *Najdi Arabic: Central Arabian*. Amsterdam and Philadelphia: J. Benjamins.

Jelinek, Eloise. 1981. *On defining categories: AUX and PREDICATE in Egyptian Colloquial Arabic*. Ph.D. diss., University of Arizona.

——. 2002. "Agreement, clitics, and focus in Egyptian Arabic". Ouhalla and Shlonsky (2002:71–105).

Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, Mass.: MIT Press.

Kortobi, Ibtissam. 2002. "Gapping and VP-deletion in Moroccan Arabic". Ouhalla and Shlonsky (2002:217–240).

Kremers, Joost. 2003. *The Arabic noun phrase: A Minimalist approach*. Ph.D. diss., University of Nijmegen.

Mohammad, Mohammad. 1989. *The sentence structure of Arabic*. Ph.D. diss., University of Southern California.

——. 2000. *Word order, agreement, and pronominalization in Standard and Palestinian Arabic*. Amsterdam and Philadelphia: J. Benjamins.

Moosally, Michelle Jamila. 1994. *Resumptive pronouns in Modern Standard Arabic: A Head-driven Phrase Structure Grammar Account*. M.A. thesis, University of Texas at Austin.

Moutaouakil, Ahmad. 1989. *Pragmatic functions in a functional grammar of Arabic*. Dordrecht: Foris.

Ouhalla, Jamal. 1990. "Sentential negation, relativized minimality, and the aspectual status of auxiliaries". *Linguistic Review* 7.183–231.

——. 1994. "Verb movement and word order in Arabic". *Verb movement*, ed. David Lightfoot and Norbert Hornstein, 41–72. Cambridge: Cambridge University Press.

——. 1997. "Remarks on focus in Standard Arabic". *Perspectives on Arabic linguistics*, X, ed. Mushira Eid and Robert Ratcliffe, 31–61. Amsterdam and Philadelphia: J. Benjamins.

——. 2002. "The logical form of negative sentences". Ouhalla and Shlonsky (2002:299–320).

——. and Ur Shlonsky (eds.). 2002. *Themes in Arabic and Hebrew syntax*. Dordrecht: Kluwer Academic Publishers.

Owens, Jonathan. 1984. *A short reference grammar of Eastern Libyan Arabic*. Wiesbaden: O. Harrassowitz.

Plunkett, Bernadette. 1993. "The position of subjects in Modern Standard Arabic". *Perspectives on Arabic linguistics*, V, ed. Mushira Eid and Clive Holes, 231–260. Amsterdam and Philadelphia: J. Benjamins.

——. 1996. "Binding in Arabic and the lexical clause". *Perspectives on Arabic linguistics*, VIII, ed. Mushira Eid, 86–97. *Amsterdam and Philadelphia: J. Benjamins.*

Qafisheh, Hamdi A. 1977. *A short reference grammar of Gulf Arabic*. Tucson: University of Arizona Press.

Rosenhouse, Judith. 1984. *The Bedouin Arabic dialects: General problems and a close analysis of North Israel Bedouin dialects*. Wiesbaden: O. Harrassowitz.

Shahin, Kimary N. 2000. *Rural Palestinian Arabic (Abu Shusha dialect)*. Munich: Lincom.

Schmidt, Hans and Paul Kahle. 1918. *Volkserzählungen aus Palästina*, I. Göttingen: Vandenhoek und Ruprecht.

——. and Paul Kahle. 1930. *Volkserzählungen aus Palästina*, II. Göttingen: Vandenhoek und Ruprecht.

Shlonsky, Ur. 1997. *Clause structure and word order in Hebrew and Arabic: An essay in comparative Semitic syntax*. New York: Oxford University Press.

Souag, Mostafa Lameen. 2006. *Explorations of the syntactic cartography of Algerian Arabic*. M.A. thesis, University of London.

Wright, William. 1898. *A grammar of the Arabic language*, ed. W. Robertson Smith and M.J. de Goeje. 3rd ed. 2 vols. Cambridge: Cambridge University Press.

Zubizarreta, Maria-Luisa. 1998. *Prosody, focus, and word order*. Cambridge, Mass.: MIT Press.

Frederick M. Hoyt
(University of Texas at Austin)

Verbal Noun

1. Introduction

A verbal noun (→ *maṣdar*) is a fixed nominal form associated with the derived forms or patterns of the verb (*ʾawzān*). Following the Arabic tradition since Sībawayhi, who calls it also *ḥadaṯān* and *ḥadaṯ* (cf. Levin 1998:917, n. 3; Ditters 1985), Šarṭūnī (1958) defines the *maṣdar* as "what denotes a situation or event without time". Verbal nouns are a basic structure in Classical Arabic, Literary Arabic (or Modern Standard Arabic or Educated Standard Arabic), and Colloquial Arabic. Differences exist among their forms and use. The lack of consensus about verbal nouns or their description shows how multifaceted they are.

2. Morphology of the verbal noun in Modern Standard Arabic

In Modern Standard Arabic, Form I uses numerous nominal patterns as verbal nouns (e.g. *ʿilm* 'knowledge', *ṭalab* 'request', *dirāsa* 'learning', *sariqa* 'theft', *duxūl* 'entrance', *suʾāl* 'question', *maʿād* 'return', *majrā* 'flow'), while Form II (e.g. *taʿlīm* 'teaching', *tarbiya* 'education'), Form III (e.g. *firāq* 'parting (with)', *muʿālaja* 'looking after, caring for'), Form IV (*ʾikrām* 'respecting', *ʾišāra* 'pointing at'), and Form X (*istiʿmāl* 'making use of, using' and *istiqāma* 'straightening out') may each have two patterns. The second pattern in Form II, III, IV, and X is usually limited to verbs II*w/y* (→ weak verbs). In other cases, the form with the feminine ending may reflect a simple noun contrasting with the verbal noun (e.g. *tadkīr* 'reminding'/*tadkira* 'ticket'; cf. Larcher 2003:33). The verbal nouns of Form V (*taġallub* 'overcoming'), VI (*takātub* 'correspondence'), VII (*infijār* 'explosion'), VIII (*ištirāk* 'participation'), and IX (*iḥmirār* 'reddening') have one form each.

In addition, Classical and Modern Standard Arabic employ the form of the passive participle singular in certain derived forms in the function of verbal nouns, e.g. *munṭalaq* 'starting point', *murtafaᶜ* 'rising [landscape]', *mustaqbal* 'time of future meeting', etc. Some of these modern words are also used in the higher registers of Colloquial Arabic.

Sībawayhi mentions other patterns used as verbal nouns by analogy to several other noun patterns (e.g. *takrār* for *takrīr* 'repetition', *ʾiqām* for *ʾiqāma* 'organizing', etc.). See also Wright's (1962, vol. I) detailed list of verbal nouns in Classical Arabic.

Modern grammars of Arabic usually pay little attention to verbal nouns. Verbal nouns are mentioned in chapters dealing with the morphology of nouns or verbs, sometimes listed with the derived forms, and with only short notes about their features. Thus, for instance, Šarṭūnī (1958) mentions certain syntactic functions and features of the verbal nouns, based on Classical Arabic.

Modern Standard Arabic verbal nouns are inherited from Classical Arabic and do not show morphological innovations. With respect to Colloquial Arabic, Brockelmann (1908) refers to verbal nouns as nouns that appear with various affixes, and he adds explanations about their meanings, which are related to the basic meanings of the respective derived form of the verb.

Morphologically, a verbal noun is a singular form, but in some cases, mainly in Modern Standard Arabic, it may receive a plural suffix and reflect plural activities, either as a real noun or as multiple occurrences of the verb activity, e.g. *istiᶜlāmāt* 'information'; *muxābarāt* 'intelligence [military or police]'. Thus, semantic processes affect the lexical meaning and syntactic function of the word form.

Cantarino (1975:401–406), in his syntactic discussion of the verbal nouns, follows the Arab grammarians' approach, although his examples are drawn from Modern Literary Arabic texts. There are, however, semantic innovations yielding modern meanings for verbal nouns and enabling their use in new contexts. Such innovations distinguish these verbal nouns from those of Classical and Colloquial Arabic.

3. Verbal nouns between Colloquial Arabic and Modern Standard Arabic

With respect to verbal nouns, a number of common features are shared between Modern Standard Arabic and Colloquial Arabic dialects (cf. Rosenhouse 2004):

i. Action/result ambiguity of the verbal noun, e.g. *xalq* 'creation; creatures'.
ii. Active/passive ambiguity, e.g. *taʾlīf, muʾallaf* 'composing, writing a literary work' and 'a composed work'.
iii. Suppletion, due to discrepancy between verbal Forms and verbal nouns. Although this feature is common to Modern Standard Arabic and Colloquial Arabic, certain lexical details differ: *ʾixrāj* 'taking out; producing' is not used in Colloquial Arabic for 'taking out', but it is used for 'producing', a modern notion borrowed from Modern Standard Arabic.
iv. The number of verbal nouns of Form I is by far larger than verbal nouns in the other Forms in all the studied texts. This is in line with the high frequency of verbs in Form I compared to other Forms as a basic characteristic of Arabic.

Differences between Colloquial Arabic and Modern Standard Arabic exist on all levels: phonological, morphological, syntactic, semantic, and lexical. Some differences between verbal nouns in Colloquial Arabic and Modern Standard Arabic are due to differences in the phonological systems, for instance in the feminine ending (Modern Standard Arabic *daᶜwā*, but in some varieties of Colloquial Arabic *daᶜwe* 'invitation') or the realization of phonemes.

Morphological differences exist for certain dialects only (see below). Other differences are found in the use of verbal nouns in relation to a different verb pattern. This kind of suppletion was already mentioned by Sībawayhi. Thus, for instance, Modern Standard Arabic *tarāsul* is in Colloquial Arabic *mrāsale* 'correspondence'.

A syntactic difference is, for instance, the governance of an object by the verbal noun in a noun phrase. Since in Colloquial Arabic such synthetic structures are avoided, they remain characteristic of Modern Standard Arabic and Literary Arabic.

Lexical differences also exist, e.g. in Modern Standard Arabic *taḥadduṯ* ‘talking’ vs. *ḥaki, ḥadīṯ* in Colloquial Arabic.

The differences between verbal nouns in Standard Arabic and Colloquial Arabic may be summarized as follows (cf. Rosenhouse 2004):

i. There is limited frequency of verbal nouns of active/passive/reflexive Forms (V, VI, VII) in Colloquial Arabic compared to Modern Standard Arabic, where especially Form VI verbal nouns develop novel semantic meanings: *tafattuḥ* ‘opening [e.g. a blossom]’, *tamāluk* ‘(self-)control’, *infitāḥ* ‘opening [e.g. a door by itself]’, *iftitāḥ* ‘opening [e.g. an inauguration]’.

ii. Verbal nouns used as *mafʿūl muṭlaq* ‘absolute object’, as in *nāma nawman ʿamīqan* ‘he slept a deep sleep’, are frequent in Classical and Modern Standard Arabic (cf. e.g. Wright 1958: II; Levin 1998; Talmon 1996; Larcher 2003; → object, absolute), but not in Colloquial Arabic. According to Talmon (1996), there are a total of 64 tokens in the *Qurʾān*, from 52 different verbs. Only 3 verbs do not take the stem of their corresponding verbs, whereas 61 do (Form I: 34; Form II: 26; Form IV: 3; Form X: 1). The total number of verbal nouns of Forms above III is 4 (= 6.25% of the 64 tokens). This *mafʿūl muṭlaq* seems to be used in Colloquial Arabic mainly to express exaggeration in narratives.

iii. Deverbal or denominal verbal nouns expressing procedures in Modern Standard Arabic may be formed freely from modern nouns and verbs. Such forms do not exist in Colloquial Arabic, but may be borrowed from Modern Standard Arabic into Colloquial Arabic when needed, e.g. ‘flexurization’: Literary Arabic: *jaʿlu(hu) marinan/tamrīn/tamarrun/talyīn; iḥmirār* ‘becoming red, blushing’, *taʾaqlum/ʾaqlama* ‘climatization’.

iv. Suppletion in Colloquial Arabic, due to nonexistence of verbal nouns, e.g. *ʾinno yṣīr* + adjective ‘that he/it may become...’ instead of *jaʿlu(hu)* + adjective.

v. Suppletion in case of causativity. For example, for Literary Arabic *ʾalbasa*/Colloquial Arabic *labbas* ‘to dress someone’, the verbal noun in Literary Arabic is *ʾilbās*, but for Colloquial Arabic it is probably *talbīs* (if at all, due to iii above).

vi. Suppletion for reciprocity in certain Form III verbs such as Literary Arabic *baraka*/Colloquial Arabic *barake* ‘blessing’, rather than the expected **birāk/?mubāraka* in this Form. Form VI verbal nouns are used in Classical Arabic and Modern Standard Arabic, e.g. *taḍārub* ‘mutual hitting’, but relatively rarely in Colloquial Arabic. Reciprocity in Colloquial Arabic is usually expressed analytically, using the correlative particle, e.g. *ḍarabu* (*baʿḍhom*) *baʿḍ* ‘they hit each other’. The verbal nouns of some Form III verbs can be used in Colloquial Arabic.

vii. The preference for inflected verb forms in Colloquial Arabic. Nominal forms are relatively abundant in Modern Standard Arabic, probably influenced by Western languages and styles, such as English and French journalese. Colloquial Arabic prefers inflected verb forms to verbal nouns or other nominal forms. Therefore, Modern Standard Arabic uses verbal nouns more often than Colloquial Arabic. This has been observed in the context of suppletion, especially with intransitive and denominal (or de-adjectival) verbs.

viii. Pronunciation of verbal nouns (morphophonology). In contrast with Classical Arabic and Modern Standard Arabic, vowels of verbal nouns in Colloquial Arabic are often deleted according to local Colloquial Arabic phonological rules: Literary Arabic *bišāra*/Colloquial Arabic *bšāra* ‘delivering good news’; Literary Arabic *muwāfaqa*/Colloquial Arabic *mwāfaqa* ‘agreement’; Literary Arabic *ḥisāb*/Colloquial Arabic *ḥsāb* ‘calculation’. In certain phonetic environments and lexical items, such vowels remain also in Colloquial Arabic, e.g. *muḏāhara* ‘picketing’, *firāq* ‘parting with’, *taʿallum* ‘learning’.

ix. Avoidance of verbal nouns due to phonetic considerations. Consonant clusters involving homorganic adjacent consonants often lead to general avoidance of verbal nouns (and some other word structures with the same clusters). Thus, for instance, for Colloquial Arabic *daššar* ‘to leave, neglect’, the verbal noun would be Colloquial Arabic *tadšīr*; but actually this form does not occur (but compare Literary Arabic and Colloquial Arabic *tadšīn* ‘inauguration’). In certain cases, this phonetic issue leads to a different verbal noun form (by suppletion).

4. INNOVATIONS IN COLLOQUIAL ARABIC

Morphological and semantic innovations are found in the Arabic dialects, especially in North African dialects and peripheral dialects such as Malta. Little information about verbal nouns is usually given in dialect studies. Often, but not always, lists of verbal noun forms are presented with verbal derived forms. On the other hand, some works cite only the morphologically special forms, which appear to be compared implicitly with the ‘traditional’ and well-known verbal noun patterns of Classical and Modern Standard Arabic (e.g. Behnstedt 1987).

New forms of verbal nouns occur mainly in the western region of Colloquial Arabic. The following forms are noteworthy. Egyptian Arabic: *ʾuʿād* ‘sitting’ for Form I, *tarbeyya* ‘education, upbringing’, *tekrār* ‘repetition’ for Form II, *minadya* ‘calling’ for Form III; Algeria: *tʿanīd* ‘rivalry’ as verbal noun of Form III; Morocco: *metgārsa* ‘wrestling’, *tmesxīr* ‘making oneself a laughing stock’ for Form VI; Maltese: *tkebbir* ‘self-importance, enlarging’ for Form V, *tqabid* ‘fighting’ for Form VI, *stenbih* ‘awakening’ for Form X; Ḥassāniyya of Mauritania: *texmām* ‘thinking’ for Form II, *tġarbīl* ‘sieving [e.g. flour]’.

5. Distribution and frequency

Literary Arabic texts have a much larger rate of verbal nouns (tokens and percentage) than Colloquial Arabic texts do (see Rosenhouse 2004; Tables 1 and 2). This may be related to the preference of nominal structures in Modern Literary Arabic compared to their avoidance in Colloquial Arabic. Regarding verbal noun use, no sociolinguistic differences have been detected so far between sedentary and Bedouin Colloquial Arabic dialects.

Classical Arabic, Modern Standard Arabic, and Colloquial Arabic vary as to the frequency level of verbal nouns of verbs above Form I. Further study is required to elucidate this picture. Verbal nouns of Modern Standard Arabic verbs above Form II are more frequent than verbal nouns of such measure in Colloquial Arabic.

Texts of Palestinian sedentary speakers from two periods in the beginning (Schmidt and Kahle 1918) and the end of the 20th century have been examined in order to answer the question of whether verbal nouns in Eastern Colloquial Arabic dialects developed in this period. Few differences between these time periods have been revealed; apparently, such processes take a longer time.

A large difference exists between the higher frequency of Modern Literary Arabic verbal nouns and the lower frequency of verbal nouns in the Classical Arabic texts studied. One of the factors leading to this finding could be the use of verbal nouns in contexts that sometimes blur the semantic boundaries between verbal nouns and ‘regular’ nouns.

6. Conclusions

The similarities and differences in use between verbal nouns in Literary (Classical and Modern Standard Arabic) and Colloquial Arabic can be summarized as follows. A simple analysis of dictionary lists reveals differences between Literary Arabic and Colloquial Arabic on the morphological, semantic, syntactico-semantic, and phonological levels (Rosenhouse 2002). Additional analysis of texts from the various Arabic periods and levels (Rosenhouse 2004) reveals clear differences between Literary Arabic and Colloquial Arabic in the frequency and distribution of verbal nouns: verbal nouns are more frequent in Classical Literary Arabic and Modern Literary Arabic than in Colloquial Arabic. Differences in use of verbal nouns were found also between different Modern Standard Arabic and Classical Arabic texts. Interdialectal differences of verbal nouns use are relatively minor, compared to the difference between Literary Arabic and Colloquial Arabic.

Purely syntactic structural differences between Literary Arabic and Colloquial Arabic concerning verbal nouns have not been found. Governed subject and object complements of a verbal noun in a noun phrase are, however, apparently lost in Colloquial Arabic, unlike Classical and Modern Standard Arabic.

It appears that the existence or generation of a verbal noun depends on semantic criteria of the inflected verb or on morphological criteria such as the Form (Measure) of the inflected verb. Thus, not every verb automatically has a verbal noun in Colloquial Arabic. This fact contributes to the general difference between Colloquial Arabic and Literary Arabic; the latter restricts the generation of verbal nouns much less than Colloquial Arabic.

The distribution rules in both Modern Standard Arabic and Colloquial Arabic depend mainly on the semantic category of each individual verb (root). These semantic categories relate to those active in Arabic in general and include transitivity/intransitivity, causativity, passivity, reciprocation, ‘verbs of the heart’, verbs of motion, etc. Some of these categories are usually but not only expressed in morphological verb measures.

Phonetic differences between Literary Arabic and the various Colloquial Arabic dialects also

Table 1. Verbal nouns in Literary Arabic texts

text	number of pages; number of words	number of verbal nouns	verbal nouns/total number of words	number of verbal nouns of non-Form I verbs	non-Form I verbal nouns/total number of verbal nouns (%)	verbal nouns above Form III/ non-Form I verbal nouns (rate and %)
*Qurʾān**	(see Talmon 1996)	64		30	46.8%	4/30 (13.3%)
Ibn al-Muqaffaʿ, *Kalīla wa-Dimna*	30; 3,555	187	5.3%	43	23%	20/43 (46.5%)
Maḥfūẓ, *Zaʿbalāwi*	12; 2,515	121	4.8%	43	35.5%	32/43 (74.4%)
Ḥuḍayr, *ʾUrjūḥa*	11; 2,042	32	1.6%	13	40%	10/13 (76.9%)
Takarlī, *Tannūr*	6; 1,588	27	1.7%	16	59.25%	10/16 (62.5%)

* Verbal nouns noted here are only of the *mafʿūl muṭlaq* type as described in Talmon (1996).

Table 2. Verbal nouns in various Colloquial Arabic texts

text	number of pages; number of words	number of verbal nouns	number of verbal nouns/ total number of words (%)	verbal nouns of non-Form I verbs	non-Form I verbal nouns/ total number of verbal nouns (%)	verbal nouns above Form III/non-Form I verbal nouns (rate and %)
Alexandria (Behnstedt and Woidich 1987)	12; 3,001	13	0.43%	0	–	0/0 (0)
Sinai Bedouin (Stewart 1988)	14; 2,909	24	0.8%	2	8.3%	0/2 (0)
Negev Bedouin (Henkin 1985)	13; 3,322	10	0.3%	0	–	0/0 (0)
Galilee Bedouin (Rosenhouse 1984)	16; 2,970	21	0.7%	2	9.5%	0/2 (0)
Palestinian (Meron and Kabha 1993)	48; 3,102	11	0.35%	2	18.6%	2/2 (100%)

Table 2 (*cont.*)

text	number of pages; number of words	number of verbal nouns	number of verbal nouns/ total number of words (%)	verbal nouns of non-Form I verbs	non-Form I verbal nouns/ total number of verbal nouns (%)	verbal nouns above Form III/non-Form I verbal nouns (rate and %)
Jordanian (Palva 1969, 1970)	12; 3,785	28	0.7%	6	21.4%	4/6 (66.%)
Mardin (*qəltu*) (Jastrow 1981)	10; 2,055	9	0.44%	0	–	0/0 (0)
Damascus (Grotzfeld 1965)	8; 3,078	17	0.55%	1	5.8%	0/1 (0)
Aleppine (Sabuni 1980)	17; 2,825	26	0.9%	1	3.8%	0/1 (0)
Baghdadi Judeo-Arabic (Mansour 1977)	28; 2,253	12	0.53%	3	25%	2/4 (50%)

affect verbal nouns. These differences contribute to the overall distinction between Literary Arabic and Colloquial Arabic. The pronunciation of the verbal nouns suggests lexical interaction between Literary Arabic and Colloquial Arabic verbal nouns. Often (in forms that can be phonetically modified), when the verbal noun is pronounced in the same manner in both Literary Arabic and Colloquial Arabic, it is an 'authentic' Literary Arabic form borrowed into Colloquial Arabic. If the pronunciation of a verbal noun in Colloquial Arabic differs from that of its Literary Arabic equivalent, however, it is usually an 'originally' Colloquial Arabic form. This point is related to the use of Modern Standard Arabic in speech, which has been studied as part of the features of the modern 'middle language' or Standard → Educated Arabic (e.g. Blanc 1960; Meiseles 1980), but is beyond the scope of the present description.

Many of the issues raised here remain open, awaiting further study. To refine this description, more texts and text types should be studied, in all the registers of the Arabic language, including the Eastern and Western Colloquial Arabic dialects, Arabian Bedouin dialects, and peripheral dialects such as those in the Nilo-Saharan area. The use of verbal nouns as an artistic device in the *mafʿūl muṭlaq* pattern in Modern Standard Arabic and Colloquial Arabic is also worthy of future study.

BIBLIOGRAPHICAL REFERENCES

Behnstedt, Peter. 1987. *Die Dialekte der Gegend von Saʿdah (Nord Jemen)*. Wiesbaden: O. Harrassowitz.

——. and Manfred Woidich. 1987. *Die ägyptisch-arabischen Dialekte*. III. *Texte*. I. *Delta Dialekte*. Wiesbaden: L. Reichert.

Blanc, Haim. 1960. "Style variations in spoken Arabic". *Contributions to Arabic linguistics*, ed. Charles A. Ferguson, 81–158. Cambridge, Mass.: Harvard University.

Brockelmann, Carl. 1908. *Grundriss der vergleichenden Grammatik der semitischen Sprachen*, I. Berlin: Reuther und Reichard. (Repr., Hildesheim: G. Olms.)

Cantarino, Vicente. 1975. *Syntax of Modern Arabic prose*. II. *The expanded sentence*. Bloomington, Ind. and London: Indiana University Press.

Ditters, Everhard. 1985. "The structure of the *maṣdar* noun phrase according to Abū Ḥayyān al-Andalusī". *Zeitschrift für Arabische Linguistik* 15.67–79.

Fischer, Wolfdietrich. 1972. *Grammatik des klassischen Arabisch*. Wiesbaden: O. Harrassowitz.

——. 1991. "Classical Arabic". *The Semitic Languages*, ed. Robert Hetzron, 187–219. London: Routledge.
Grotzfeld, Heinz. 1965. *Syrisch-arabische Grammatik*. Wiesbaden: O. Harrassowitz.
Henkin, Ron. 1985. *The verbal system in Negev Bedouin dialects: Tense and aspect*. Ph.D. diss., Tel Aviv University.
Ḥuḍayr, M. 1972. "al-ʾUrjūḥa (The seesaw)" [from *The Black Kingdom*]. *Selected Modern Arabic short stories*, ed. Ami Elad-Bouskila, 87–97. Jerusalem: Academon.
Ibn al-Muqaffaʿ. N.d. *Kalīla wa-Dimna*. Beirut: Dār al-Kutub al-ʿIlmiyya.
Jastrow, Otto. 1981. *Die mesopotamisch-arabischen* Qəltu *Dialekte*. II. *Volkskundliche Texte in elf Dialekten*. Wiesbaden: F. Steiner.
Larcher, Pierre. 2003. *Le système verbale de l'arabe classique*. Aix-en-Provence: Université de Provence.
Levin, Aryeh. 1998. *Arabic linguistic thought and dialectology*. Jerusalem: Magnes Press.
Maḥfūẓ, Nagīb. 1963. "Zaʿbalāwi" [from *God's World*]. *Selected Modern Arabic short stories*, ed. Ami Elad-Bouskila, 19–30. Jerusalem: Academon.
Mansour, Jacob. 1977. *The Judaeo-Arabic dialect of Baghdad*. II. *Texts*. Haifa: University of Haifa.
Meiseles, Gustav. 1980. "Educated spoken Arabic and the Arabic language continuum". *Archivum Linguisticum* 11.118–143.
Meron, Yoram and Riad Kabha. 1993. *Tales of the wadi: Selected folktales from Wadi Arah*. Givat Havivah: Institute of Arabic Studies.
Palva, Heikki. 1969. "Balgāwi Arabic. II. Texts in the dialect of the *yigūl* group". *Studia Orientalia* 40:2.3–15.
——. 1970. "Balgāwi Arabic. III. Texts from Ṣāfūṭ". *Studia Orientalia* 43:1.3–26.
——. 1992. *Artistic Colloquial Arabic*. Helsinki: Finnish Oriental Society.
Rosenhouse, Judith. 1984. *The Bedouin Arabic dialects: General problems and a close analysis of North Israel Bedouin dialects*. Wiesbaden: O. Harrassowitz.
——. 1990. "Tendencies to nominalization in Modern Literary Arabic as compared with Classical Arabic". *Zeitschrift für Arabische Linguistik* 22.23–43.
——. 1991–1992. "The occurrence of the passive in some Bedouin dialects". *Journal of Afroasiatic Languages* 3.9–21.
——. 2002. "Some features of verbal nouns in Colloquial and Literary Arabic (compared to Hebrew)". *Helkat Lashon* 33–34.127–147. (In Hebrew.)
——. 2004. "On verbal nouns in Colloquial and Literary Arabic". *Approaches to Arabic dialects: A collection of articles presented to Manfred Woidich on the occasion of his 60th birthday*, ed. Martine Haak, Rudolf de Jong, and Kees Versteegh, 275–289. Leiden: E.J. Brill.
Sabuni, Abdulghafur. 1980. *Laut- und Formenlehre des arabischen Dialekts von Aleppo*. Frankfurt a. Main: P. Lang.
Šarṭūnī, Rašīd aš-. 1958. *Mabādiʾ al-ʿarabiyya fī ṣ-ṣarf wa-n-naḥw*. Beirut: al-Maṭbaʿa al-Kāṯūlīkiyya.
Schmidt, Hans and Paul Kahle. 1918. *Volkserzählungen aus Palästina gesammelt bei den Bauern von Bir-Zet*, I. Göttingen: Vandenhoeck und Ruprecht.
Stewart, Frank H. (with the assistance of Haim Blanc and S. Ḥsēn). 1988. *Texts in Sinai Bedouin law: The texts in Arabic; glossary (Part 2)*. Wiesbaden: O. Harrassowitz.
Takarli, Fūʾād at-. 1960. "at-Tannūr (The oven)". *Selected modern Arabic short stories*, ed. Ami Elad-Bouskila, 81–86. Jerusalem: Academon.
Talmon, Rafael. 1996. "The syntactic category of *mafʿūl muṭlaq*: A study in Qurʾānic syntax". *Arabic grammar and linguistics*, ed. Yasir Suleiman, 107–129. Richmond: Curzon Press.
Wright, William. 1962. *A grammar of the Arabic language, translated from the German of Caspari and edited with numerous additions and corrections*. 3rd ed. rev. by W. Robertson Smith and M.J. de Goeje. 2 vols. Cambridge: Cambridge University Press.

Judith Rosenhouse
(Swantech-Sound Waves Analysis and Technologies, Ltd.)

Vernacular → Colloquial

Voice (Morphology) → Diathesis

Voice (Phonetics)

1. Introduction

This entry stays close to the most common definition of 'voice', according to which 'voice' corresponds to the component of speech that is mainly attributed to the vibration of the vocal folds. The phoneticians distinguish, in fact, between several modes of vocal fold vibration (or voice quality). For some of them, 'voice' is the nonmarked mode considered 'neutral' or 'modal' (see below, Sec. 3.1). The component of speech assigned to the vocal fold vibration is just one example of possible differences between speech sounds that are produced mainly by modifying the states of the vocal folds. These modifications can express several functions: linguistic (the phonetic contrasts between the segments, the prosodic patterns), paralinguistic (the speaker's emotional state), sociolinguistic (to signal regional or social linguistic differences), or extralinguistic (the speaker's identification: age, sex, etc.).

2. Aerodynamic and articulatory considerations

Every sound of speech possesses at least one source and one front resonating cavity which modify its spectral properties (Kent a.o. 1996). The source of a vowel is a periodic sound produced by the vibrations of the vocal folds and modulated by the vocal tract, which enhances the amplitude of some of its harmonics and weakens others. Depending on the specific shape of its supralaryngeal cavity, each vowel has a specific quality that depends on the natural frequency resonance of this cavity. For a whispered (low) vowel, the source is turbulence produced in the larynx and concentrated at the level of the natural frequency resonance of its supralaryngeal cavity, which is not very different from that of its voiced counterpart. The source, during the fricatives and the frication phase of the aspirated stops release, is a turbulence generated through a very narrow laryngeal or supralaryngeal constriction and modulated by the front cavity. A sound can also combine periodic laryngeal and nonperiodic supralaryngeal sources (voiced fricative). The difference between a voiced and a whispered vowel, between a voiced and a voiceless consonant, is a difference of phonation type (laryngeal adjustments) and not of segmental sound quality (supralaryngeal adjustments). In fact, there exist, in languages in general and in Arabic in particular, several phonation types that can be simple or compound. Special attention is given here to those that may have a linguistic function.

During speech, the vocal folds constitute the first and main obstacle that airflow, expelled by the lungs, encounters. The vocal folds are composed of an anterior (or ligamental) part and a posterior (or cartilaginous) part. The space between the two vocal folds is commonly called the glottis. The glottis can, therefore, remain fully open (vocal folds abducted), or closed (vocal folds adducted) completely or at its anterior part only. Above the vocal folds are the ventricular bands, or the false vocal folds, that can also participate in the modification of the phonation type (Catford 1964, 1968, 1977). The tips of the arytenoids, the aryepiglottic folds, and the base of the epiglottis form the aryepiglottic sphincter that can also be involved in the regulation of the phonation types (Catford 1964, 1968, 1977; and mainly Esling 1996). The ventricular folds and the aryepiglottic sphincter constitute the supraglottic level.

3. Examples of simple and compound phonation types

3.1 *Simple phonation types*

Breath: This is a turbulent airflow produced at the glottis when the vocal folds are fully separated, the glottal opening is smaller than the supralaryngeal one, and the transglottic airflow is sufficiently large. It is typically observed during the release of aspirated stops, and the voiceless laryngeal fricative [h]. Classical Arabic [t k] are sometimes considered to have been aspirated stops (Roman 1982; Ghazeli 1977). This analysis cannot be adopted for these consonants in the modern Arabic dialects because even their release, as in Classical Arabic, takes more time. Indeed, during the voiceless nonaspirated stops, the glottis is fully open, but its size is smaller than during aspirated ones (for review: Dixit 1989). Since the voiceless unaspirated stops are produced with a significant increase of the intraoral pressure, the transglottic airflow is not sufficient to generate turbulence (breath or aspiration) at the glottis during their closure, and the frication noise produced during their release in fact is initiated at the supraglaryngeal level. Fibroscopic and transillumination observations show that the maximal opening of the glottis during [s], [t k] (long release), and [q] (short release) of Moroccan Arabic is very similar (Zeroual 2000; Zeroual a.o. 2006). Based on these physiological arguments, among others, one may say that all Moroccan Arabic voiceless stops [t k q] are nonaspirated. The laryngeal fricative is voiceless in Classical Arabic and many modern Arabic dialects (Cantineau 1960), while in Moroccan Arabic, its production combines two phonation types (see below, Sec. 3.2 *Whispery voice*).

Whisper: This is a turbulent airflow produced with different glottis shapes depending on its intensity (Weitzman a.o. 1976). For example, if the ligamental and cartilaginous parts of the vocal folds are moderately and widely open respectively, a whisper or turbulent airflow, more intense than breath, is initiated at the cartilaginous glottis and the supraglottic level. This phonation type induces an adduction of the ventricular folds (Catford 1964, 1968, 1977) and an anterior-posterior compression of the aryepiglottic sphincter (Weitzman a.o. 1976; Gao and Esling 2003) that appears as a closeness mainly between the arytenoids and the base of the epiglottis. Usually, whisper does not have a distinctive function in languages, being generally used, including Arabic, when speakers want to communicate in secret with their interlocutor (paralinguistic function). The Arabic voiceless pharyngeal consonant has some features of the whisper: auditively, a very similar frication noise, and a more significant compression of the aryepiglottic sphincter (Zeroual 1999; Zeroual a.o. 2004a, b).

Modal voice: During voiced sounds, the vocal folds are moderately in contact over their entire length, and form an obstacle for air compressed by the lungs. The subglottal pressure increases to attain a higher value, which separates the vocal folds and allows the air to pass through the glottis. The elastic structure of the muscles and ligaments of the vocal folds, and the abrupt fall in subglottal pressure during the opening phase of the glottis, free the vocal folds to return to their initial closed position. This movement is accelerated by another aerodynamic phenomenon. Indeed, when the airflow passes quickly through the glottis, the pressure between the vocal folds decreases in a very significant manner and 'suction' takes place between the two vocal folds (Bernoulli effect). The subglottal pressure increases again, separates the vocal folds, and the cycle restarts (aerodynamic-myoelastic theory). Then, a periodic laryngeal sound (modal voice) is produced.

This phonation type is responsible for the phonetic contrast between the voiced and voiceless sounds present in almost all languages. With respect to this opposition, the situation in Classical Arabic is complex. Indeed, if we put aside [ṭ q ʔ], voiced vs. voiceless correspond to the opposition → *majhūra* vs. *mahmūsa* in the Arabic grammarian's analysis since Sībawayhi (for review: Blanc 1967; Roman 1982; Al-Nassir 1993). [ṭ q ʔ] are considered *majhūra* although they are voiceless in the majority of the modern Arabic dialects. The phonetic categorization of Classical Arabic [ṭ q] is very controversial, whereas [ʔ] is considered by the majority of the authors to be voiceless. Here, the hypothesis of, among others, Ghazeli (1977) is adopted, according to which [ṭ q ʔ] are voiceless, but classified with the *majhūra* because the duration of their release (or positive VOT) is very short.

In speech, it is principally the voiced segments that carry the melodic patterns, i.e. the variations of pitch that express some discursive functions (affirmation, question, focus, etc). In general, these variations are expressed with the help of a mode of vocal fold vibration that is considered neutral because, in principle, it does not add any special paralinguistic functions. This phonation type is called 'modal voice' by Laver (1980), and 'voice' by Catford (1964, 1968, 1977) and Ladefoged (1971).

Creak: This is a phonation type that is possible but rarely used in speech (Ladefoged 1971), except if it is combined with voicing (see Sec. 3.2, *Creaky voice*). During this mode, the posterior part of the glottis is tightly closed; only the anterior parts of the vocal folds can vibrate periodically, but with a very low frequency. Catford (1968) underlines the fact that "its auditory effect resembles a stick being dragged along a railing".

3.2 *Compound phonation types*

There are, in fact, several types of vocal cord vibration, generally characterized according to differences between their articulatory, acoustic, aerodynamic, and auditory properties and those of modal voice. These phonation types are achieved while combining, among others, breath, whisper, and creak with voicing. As in the phonetic literature, these are called 'compound modes': breathy voice, whispery voice, and creaky voice.

Creaky voice: Ladefoged and Maddieson (1996) identify two different laryngeal mechanisms associated with this phonation type. Indeed, if the posterior part of the glottis is firmly closed, the anterior parts of the vocal folds will vibrate in a nonperiodic manner. However, if the closing of the posterior part is not too strong, the ligamental and cartilaginous parts of the vocal folds will vibrate, but out of phase, and then a nonperiodic sound is initiated. During creaky voice, even the ventricular folds are adducted

(Hollien a.o. 1966, cited by Laver 1980); their bases are in contact with the superior margins of the vocal folds. The vibrating parts of the vocal folds are thicker and their mass more significant than in the modal voice. This particular posture explains why the frequency of vocal fold vibration (Fo) in creaky voice is low. Murry and Brown (1971, cited by Laver 1980) also attribute this to the low value of the subglottal pressure during creaky voice. This phonation type can also induce a compression of the aryepiglottic sphincter (Gauffin 1977). Creaky voice is used, in many languages, as a primary or secondary acoustic correlate of certain phonological contrasts. In modern Arabic dialects, the voiced pharyngeal consonant is systematically pronounced with a creaky voice (Mitchell 1993). This is predictable because this consonant is produced with a very strong compression of the aryepiglottic sphincter, the ventricular folds being also adducted (Zeroual 1999; Zeroual a.o. 2004a, b; Laufer and Condax 1979).

Breathy voice: According to Laver (1980), during this mode the vocal folds are relaxed and brought closer together along their entire length; they vibrate, but without meeting one another. The result is a periodic laryngeal sound combined with a friction noise. The periodic laryngeal sound is perceptively more prominent. The distinction between this phonation type and whispery voice is subtle, as is clear from the example given in connection with the phonetic correlates of the latter mode.

Whispery voice: This phonation type, called 'murmur' by Ladefoged (1971), combines voicing with a friction noise that is more intense and more prominent than in breathy voice (Laver 1980). Catford (1977) states that during whispery voice, the anterior part of the glottis, where voicing is initiated, is generally closed, while its posterior part stays open and generates a whisper. He adds that a whispery voice can also be achieved if the anterior parts of the vocal folds are relaxed and are brought closer together than during breathy voice, and the cartilaginous part of glottis is closed. The shape of the glottis during Moroccan Arabic [ɦ] (Zeroual 1999; Zeroual a.o. 2004a, b) corresponds to the first type of whispery voice. In spite of this, it seems more accurate to say that [ɦ] is produced with breathy voice. Indeed, whisper and whispery voice require a relatively higher tension of the vocal fold muscles (Laver 1980), while Moroccan Arabic [ɦ] is associated with a low degree of tension through the laryngeal (and supralaryngeal) cavity. Moreover, whisper (and probably whispery voice) is accompanied by a supraglottic constriction, unlike [ɦ]. Note that the majority of authors use breathy voice for all phonation types that combine breath and voicing without controlling all their phonetic correlates.

Bibliographical references

Al-Nassir, A.A. 1993. *Sibawayh the phonologist*. London and New York: Keegan Paul International.

Blanc, Haim. 1967. "The 'sonorous' vs. 'muffled' distinction in old Arabic phonology". *To honor Roman Jakobson: Essays on the occasion of his seventieth birthday*, 295–308. The Hague and Paris: Mouton.

Cantineau, Jean. 1960. *Cours de phonétique arabe*. Paris: C. Klincksieck.

Catford, John. 1964. "Phonation types: The classification of some laryngeal components of speech production". *In honour of Daniel Jones*, ed. David Abercrombie a.o., 26–37. London: Longmans, 1964.

——. 1968. "The articulatory possibilities of man". *Manual of phonetics*, ed. Bertil Malmberg, 309–333. Amsterdam: North-Holland.

——. 1977. *Fundamental problems in phonetics*. Edinburgh: Edinburgh University Press.

Dixit, R. Prakash. 1989. "Glottal gestures in Hindi plosives". *Journal of Phonetics* 17.213–237.

Esling, John H. 1996. "Pharyngeal consonants and the aryepiglottic sphincter". *Journal of the International Phonetic Association* 26.65–88.

Gao, Man and John H. Esling. 2003. "Articulatory features of tones in whispered Chinese". *Proceedings of the XVth International Congress of Phonetic Sciences*, 2629–2632. Barcelona.

Gauffin, Jan. 1977. "Mechanisms of larynx tube constriction". *Phonetica* 34.307–309.

Ghazeli, Salem. 1977. "On the controversy of the Arabic 'qaf'". *Texas Linguistic Forum* 6.40–50.

Kent, Raymon D., James Dembowski, and Norman J. Lass. 1996. "The acoustic characteristics of American English". *Principles of experimental phonetics*, ed. Norman J. Lass, 185–225. St. Louis, Mo.: Mosby-Year Book.

Ladefoged, Peter. 1971. *Preliminaries to linguistic phonetics*. Chicago and London: University of Chicago Press.

——. and Ian Maddieson. 1996. *The sounds of the world's languages*. Cambridge, Mass. and Oxford: Blackwell.

Laufer, Asher and Iovanna D. Condax, 1979. "The epiglottis as an articulator". *UCLA Working Papers in Phonetics* 45.60–83.

Laver, John. 1980. *The phonetic description of voice quality*. Cambridge: Cambridge University Press.

Mitchell, T.F. 1993. *Pronouncing Arabic*. Oxford: Clarendon Press.

Moscati, Sabatino (ed.). 1980. *An introduction to the comparative grammar of the Semitic languages: Phonology and morphology*. 3rd ed. Wiesbaden: O. Harrassowitz.

Roman, André. 1982. *Etude de la phonologie et de la morphologie de la koine arabe*. 2 vols. Aix-en-Provence: Université de Provence.

Weitzman, Rayman S., Masayaki Sawashima, Hajime Hirose, and Tatsujiro Ushijima. 1976. "Devoiced and whispered vowels in Japanese". *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics, University of Tokyo* 10.61–79.

Zeroual, Chakir. 1999. "A fiberscopic and acoustic study of 'guttural' and emphatic consonants of Moroccan Arabic". *Proceedings of the XIVth International Congress of Phonetic Sciences*, 997–1000. San Francisco.

——. 2000. *Propos controversés sur la phonétique et la phonologie de l'arabe marocain*. Thèse de Doctorat, Université de Paris 8.

——, Lise Crevier-Buchman, and John H. Esling. 2004a. "Ajustements laryngaux durant les consonnes gutturales et l'emphatique [S] simples et géminées de l'arabe marocain". *XXVèmes Journées d'Etude sur la Parole, 19–22 avril, Fès, Maroc*, 469–472. Fes.

——, John H. Esling, and Lise Crevier-Buchman. 2004b. "Ajustements supralaryngaux durant les consonnes gutturales et l'emphatique [S] simples et géminées de l'arabe marocain". *XXVèmes Journées d'Etude sur la Parole, 19–22 avril, Fès*, Maroc, 473–476. Fes.

——, Susanne Fuchs, Phil Hoole, and John H. Esling. 2006. "Kinematic study of Moroccan Arabic simple and geminate obstruents: Evidence from transillumination". *Proceedings of the 7th International Seminar on Speech Production*, 287–294. Ubatuba-SP-Brazil.

Chakir Zeroual (Université Sidi Mohamed Ben-Abdellah / CNRS / Sorbonne–Nouvelle)

Vowel → Phonetics; Phonology

Vowel Backing

Ordinarily, vowels occur in coarticulation with consonants in the context of the syllable. Coarticulation refers to the overlapping gestures that occur during the pronunciation of any sequence of speech sounds. During speech, the movements of the articulators for the production of successive phonetic segments overlap in time and interact with one another. As a consequence, the vocal tract configuration at any point in time is influenced by more than one segment. Thus, the acoustic information in coarticulated vowels is fused and carried in parallel with the consonantal information (Strange a.o. 1976, 1983; Farnetani 2000). As for *backing coarticulation* in Arabic, the phonetic phenomenon of emphasis has been intensively investigated, but in most cases, it has been addressed from the consonantal point of view (Marçais 1948; Abramson and Ferguson 1962; Ali and Daniloff 1972; Giannini and Pettorino 1982; Al-Ani 1970; Delattre 1971; Wood 1979; Al-Bamerni 1983; Elgendy 2001; Ghali 1989; McCarthy 1994; El-Halees 1985; Herzallah 1990; Davis 1993, 1995; Younes 1994; Sakai a.o. 1995; Yeou 1996; Newman 2002).

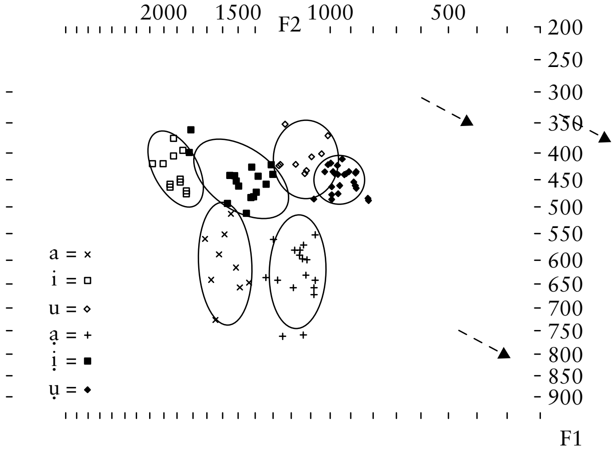
Indeed, it is well known that, in addition to oral consonants, speakers of Arabic distinguish between a number of sounds which exhibit an anterior primary articulation coupled with a secondary articulation resulting from the pulling of the back of the tongue toward the back wall of the oropharynx at the level of the second cervical vertebra. Although this articulatory specificity is traditionally referred to as 'emphasis', it appears, both acoustically and physiologically, that the area involved in the production of these posterior sounds is pharyngeal. Therefore, it seems that these sounds should more accurately be referred to as 'pharyngealized sounds'. In Modern Standard Arabic, the only segments that have this secondary articulation are [tˤ], [dˤ], [sˤ], and [ðˤ], but obviously their number is not the same in all Arabic dialects since in fact most vernaculars possess a larger number than these four, Cairo Arabic being reported to be one of the most productive in terms of functional load (Tomiche 1964; Wahba 1994).

A recent review of the research on pharyngeal articulation in Arabic shows that most physiological studies tackling this issue oriented their approach to obtain sagittal sections of the oral cavity and show the shape of the tongue during the production of pharyngealized sounds as produced by speakers of various Arabic dialects (Elgendy 2001:32). As far as the effect of these sounds on adjacent vowels is concerned, the notion of emphasis spread and allophonic variation is mentioned as far back as the Middle Ages by Arabic philologists. They had observed that emphasis spreads over a domain that is larger than one single segment and, consequently, that vowel quality could differ considerably as a function of consonantal environment (→ *tafxīm*). As for the coarticulation dynamics involved in the production and/or perception of vocalic segments in the context of these sounds, one should bear in mind that, in the case of Arabic, vowel backing and lowering are *complementary phenomena* that are directly linked with back consonant coarticulation. As a matter of fact, all vowels are backed when accompanying pharyngealized segments because emphatic consonants induce a tongue backing gesture over a distance of one to three neighboring segments. Articulatorily, backed vowels are characterized by a constriction in the pharyngeal cavity caused by Retracted Tongue Root. As for vowel lowering, it results from the lowering of the body of the tongue which is simultaneously correlated with the retraction movement of the whole body of the tongue, especially its back. Most of the works that attempted to describe pharyngeal coarticulation were phonological studies: Harris (1942); Harrell (1957); Lehn (1963); Cohen (1969); Ali and Daniloff (1972); Broselow (1976); Baddredine (1977), and Kouloughli (1978). Although the importance of these works in the field of Arabic linguistics is highly significant, they are rather impressionistic and based on limited linguistic resources. For example, Baddredine's work on emphasis in the dialect of Kairouan (Tunisia) superficially addresses the effect of pharyngealization on adjacent vowels using a corpus of only two words. Although Ali and Daniloff's study on backing coarticulation in Iraqi Arabic was based on a larger corpus, it appears that many variables, such as the nature of the emphatic consonants and the duration and quality of the vowels under study, were not controlled. Next to these, a remarkable articulatory study dealing with emphatic sound assimilation was carried out by Ghazali (1977), whose work on back consonants and backing coarticulation in some dialects of Arabic (i.e. Tunisian, Algerian, Libyan, Egyptian, Jordanian, and Iraqi) should be regarded as the first valuable experimental investigation on this topic. The cinefluorographic data were acquired together with oral and nasal airflow, spectrographic analysis being made on the corresponding audio signal. From then on, the articulatory and acoustic proprieties of pharyngealized vowels have been investigated in a number of experimental studies combining spectrographic data with X-ray pictures (Bonnot 1977; Giannini and Pettorino 1982; Ghazali 1983; Wood 1982); electromyographic investigation (Kuriyagawa a.o. 1988); or tracings of lip rounding and protrusion: Adem (1983), so as to establish an accurate acoustic-articulatory conversion for the phenomena of backing.

The acoustic effect of backing coarticulation can be observed with spectrographic analysis, any acoustic event defined as a phonetic segment showing the influence of its neighboring phones. Thus, descriptive accounts of coarticulation became possible only after acoustical methods of speech analysis became available and widespread, i.e. since the 1970s. Since Obrecht's (1968) pioneering study on the acoustic correlates of emphasis spread in Lebanese Arabic, several studies have shown that the spectral properties of vowels in pharyngealized environment are highly different from those of their nonpharyngealized equivalent (Al-Ani 1970; Al-Ani and El-Dalee 1984; Ghazali 1983; Norlin 1987; Younes 1993; Zawaydeh 1997; Barkat-Defradas a.o. 2003). Since there is a reasonably good correlation between the articulatory description of a vowel and its position in an acoustic F_1/F_2 diagram – where F_1 is related to the open/close vertical axis corresponding to the openness of the jaw (i.e., F_1 is low for high – or close – vowels and high for low – or open – vowel) and F_2 on the horizontal axis corresponds approximately to front/back position of the tongue in the mouth (i.e. high F_2 values for front or anterior vowels and low F_2 for back or posterior vowels), it is thus possible to observe that pharyngealized vowels commonly display an increase of F_1 (i.e. an opening of the jaw) and a lowering of F_2 (i.e. a retraction of the tongue), resulting in a more compact spectrum (i.e., F_1 and F_2 are closer; see Fig. 1).

The backing of the tongue can be observed on F_2 transitions since they reflect the distance the tongue should cover from the place of articulation of the pharyngealized consonant to the target of the adjacent vowel, the degree of lowering of F_2 values depending on the position of the vowel in the articulatory space. Then, the influence of emphasis seems to be related to the features [high]~[low] and [front]~[back] of the adjacent vowel, so that high vowels such as [i] and [u] seem to be less affected by pharyngealization than low vowels like [a], and front vowel like [i] more than back vowels [u] and [a]. Besides, it has been repeatedly shown that short vowels exhibit a more drastic move-

Figure 1. Formant chart for $F_1 - F_2$ for plain and emphatic [i a u] in Egyptian Arabic with arrows showing change in formant values from plain to emphatic environment (adapted from Norlin 1987:37, where emphasized vowels are transcribed with a subscript dot)



ment of backing and lowering than their long counterparts. One possible explanation for this may be that the distinctive feature for emphasis is primarily a consonantal property. During the articulation of a long vowel, the coarticulatory influence of a pharyngealized consonant on an adjacent vowel has time to decrease. In the case of short vowels, this cannot happen because, during the short span of articulation, the coarticulatory effect of pharyngealization remains through the vowel, which therefore cannot reach its target value (Norlin 1987).

Table 1. Emphaticized allophones of vocalic segments in several Arabic dialects

Vocalic segments	IPA symbols	Mean Formant Frequencies (in Hz)	Dialect	Authors	Year
/i/	[ɨ]	F_1 : 1000 F_2 : 1100–1300	Iraqi	Al-Ani	1970
/u/	[ʊ]	F_1 : 285–290 F_2 : 1000			
/a/	[ɑ]	F_1 : 600–675 F_2 : 1050–1250			
/a/	[ɑ]	F_1 : 733 F_2 : 1143	Saudi	Bonnot	1977
/i/	[ɪ]	F_1 : 500	Tunisian	Ghazali	1983
/iː/	[iː]	F_2 : 1150–1500 F_1 : – F_2 : 2000			
/u/	[ʊ]	F_1 : –			
/uː/	[uː]	F_2 : 900 F_1 : 400–550 F_2 : 930–1125			
/a/	[ɑ]	F_1 : 600–800 F_2 : 1150–1450			
/i/	[ɪ]	F_1 : 433–495 F_2 : 1631–1667	Egyptian	Al-Ani and El-Dalee	1984
/u/	[ʊ]	F_1 : 310–371 F_2 : 771–867			

Table 1 (*cont.*)

Vocalic segments	IPA symbols	Mean Formant Frequencies (in Hz)	Dialect	Authors	Year
/a/	[ɑ]	F_1 : 720–736			
		F_2 : 1029–1081			
/i/	[i]	F_1 : 450	Egyptian	Norlin	1987
/iː/	[iː]	F_2 : 1485			
		F_1 : 325			
		F_2 : 2220			
/u/	[u]	F_1 : 450			
/uː/	[uː]	F_2 : 955			
		F_1 : 330			
		F_2 : 765–795			
/a/	[ɑ]	F_1 : 630			
/aː/	[ɑː]	F_2 : 1165			
		F_1 : 620			
		F_2 : 1085			
/eː/	[eː]	F_1 : 410			
		F_2 : 2100			
/oː/	[oː]	F_1 : 400			
		F_2 : 845			
/a/	[ɑ]	F_1 : –	Rural Palestinian	Younes	1993
		F_2 : 1400–1600			
/a/	[ɑ]	F_1 : –	Jordanian	Zawaydeh	1997
		F_2 : 1100–1142			
/i/	[i]	F_1 : 437	Moroccan	Barkat-Defradas a.o.	2003
/iː/	[iː]	F_2 : 1005			
		F_1 : 252			
		F_2 : 1331			
/u/	[u]	F_1 : 373			
/uː/	[uː]	F_2 : 710			
		F_1 : 271			
		F_2 : 584			
/a/	[ɑ]	F_1 : 442			
/aː/	[ɑː]	F_2 : 781			
		F_1 : 487			
		F_2 : 865			
/eː/	[eː]	F_1 : 350			
		F_2 : 1166			

Although available data on formant frequencies from different Arabic dialects are not immediately comparable because of the widely differing sets of test material – which are collected for different purposes – it is possible to make rough comparisons and get some idea of dialectal variations. On the basis of some of the studies mentioned before, Table 1 presents the emphaticized allophones for the vocalic segments attested in several Arabic dialects; each vocalic allophone is characterized, when available, on the acoustic level.

BIBLIOGRAPHICAL REFERENCES

Abramson, A. and Charles A. Ferguson. 1962. X-ray sound motion picture film in Arabic. Produced by Columbia University, Columbia-Presbyterian Medical Center, and Haskins Laboratories.

Adem, S.A. 1983. "Are the emphatic consonants of

Egyptian Arabic rounded? An experimental investigation of rounding and lip-protrusion". *Reports from Uppsala University, Department of Linguistics* 11.

Al-Ani, Salman H. 1970. *Arabic phonology: An acoustic and physiological investigation*. The Hague and Paris: Mouton.

——. and M.S. El-Dalee. 1984. "*Tafkhim* in Arabic: The acoustic and psychological parameters". *Proceedings of the Tenth International Congress of Phonetic Sciences, Utrecht, The Netherlands*, 386–389. Dordrecht: Foris.

Al-Bamerni, A. 1983. *Oral, velic and laryngeal coarticulation across languages*. Ph.D. diss., University of Oxford.

Ali, Latif H. and Raymond G. Daniloff. 1972. "A contrastive cinefluorographic investigation of the articulation of emphatic vs. non-emphatic cognate consonants". *Studia Linguistica* 26:2.81–105.

Baddredine, Belhassen. 1977. *Le parler de Kairouan-ville: Etude phonologique et phonétique*. Ph.D. diss., Université de Paris X.

Barkat-Defradas, Melissa, Jalal Eddin Al-Tamimi, and Thami Benkirane. 2003. "Phonetic variation in production and perception of speech: A comparative study of two Arabic dialects". *Proceedings of the Fifteenth International Congress of Phonetic Sciences, Barcelona*, 857–860. Barcelona.

Bonnot, Jean-François. 1977. "Recherche expérimentale sur la nature des consonnes emphatiques de l'arabe classique". *Travaux de l'Institut Phonétique de Strasbourg* 9.47–88.

Broselow, Ellen. 1976. *The phonology of Egyptian Arabic*. Ph.D. diss., University of Massachusetts.

Cohen, David. 1969. "Sur le statut phonologique de l'emphase en arabe". *Word* 25.59–69.

Davis, Stuart. 1993. "Arabic pharyngealization and phonological features". *Perspectives on Arabic linguistics*, V, ed. Mushira Eid and Clive Holes, 149–162. Amsterdam: J. Benjamins.

——. 1995. "Emphasis spread in Arabic and grounded phonology". *Linguistic Inquiry* 26:3.465–498.

Delattre, Pierre. 1971. "Pharyngeal features in the consonants of Arabic, German, French and American English". *Phonetica* 23.129–155.

Elgendy, Ahmed M. 2001. *Aspects of pharyngeal coarticulation*. Ph.D. diss., Universiteit van Amsterdam.

El-Halees, Y. 1985. "The role of F_1 in the place of articulation distinction in Arabic". *Journal of Phonetics* 13.287–98.

Farnetani, Ena. 2000. "Coarticulation and Connected Speech Processes". *The handbook of phonetic sciences*, ed. William Hardcastle and John Laver, 371–404. Oxford and Cambridge, Mass.: Blackwell.

Ghali, M. 1989. "Pharyngeal articulation". *Topics in linguistics*, ed. M.M. Ghali, A.H. Hussien, M. Shafi, and A. Wagialla. Jeddah: Dar Al-Shourouq.

Ghazali, Salem. 1977. *Back consonants and backing coarticulation in Arabic*. Ph.D. diss., University of Austin, Texas.

——. 1983. "La coarticulation de l'emphase en arabe". *Arabica* 28.251–277.

Giannini, A. and P. Pettorino. 1982. "The emphatic consonants in Arabic". *Speech Laboratory Report*, IV. Naples: Istituto Universitario Orientale.

Harrell, Richard Slade. 1957. "The phonology of Colloquial Egyptian Arabic". *American Council for Learned Societies*. New-York.

Harris, Zellig. 1942. "The phonemes of Moroccan Arabic". *Journal of the American Oriental Society* 62.309–318.

Herzallah, Rukayyah. 1990. *Aspects of Palestinian Arabic phonology: A non-linear approach*. Ph.D. diss., Cornell University.

Kouloughli, Jamel Eddine. 1978. *Contribution à la phonologie générative de l'arabe: Le système verbal du parler du Sra*. Thèse de troisième cycle, Université de Paris-7.

Kuriyagawa, F., M. Sawashima, S. Nimi, and H. Hirose. 1988. "Electromyographic study of emphatic consonants in Standard Jordanian Arabic". *Folia Phoniatrica* 40.117–122.

Lehn, William. 1963. "Emphasis in Cairo Arabic". *Language* 39.29–39.

Marçais, Pierre. 1948. "L'articulation de l'emphase dans un parler arabe maghrébin". *Annales de l'Institut d'Etudes Orientales* (Algiers) 7.5–28.

McCarthy, John. 1994. "The phonology of Semitic pharyngeals". *Phonological structure and phonetic form*, ed. Patricia Keating, 191–233. Cambridge: Cambridge University Press.

Newman, Daniel. 2002. "The phonetic status of Arabic within the world's languages: The uniqueness of the *luʁat al-dˤaad*". *Antwerp Papers in Linguistics* 100.63–75.

Norlin, Kjell. 1987. *A phonetic study of emphasis and vowels in Egyptian Arabic*. Lund: Department of Linguistics, Lund University.

Obrecht, Dean H. 1968. *Effects of the second formant on the perception of velarization consonants in Arabic*. The Hague and Paris: Mouton.

Sakai, S., F. Tanokuchi, and Y. Fujiwara. 1995. "Laryngeal fricatives in Arabic". *Folia Phoniatrica et Logopedia* 47:2.101.

Strange, Winifred, James Jenkins, and Thomas L. Johnson. 1983. "Dynamic specification of coarticulated vowels". *Journal of the Acoustical Society of America* 74:3.695–705.

——, Robert Verbrugge, Donald Shankweiler, and Thomas Edman. 1976. "Consonant environment specifies vowel identity". *Journal of the Acoustical Society of America* 60:1.213–224.

Tomiche, Nada. 1964. *Le parler arabe du Caire*. The Hague and Paris: Mouton.

Wahba, Kassem M. 1994. "An acoustic study of vowel variation in colloquial Egyptian Arabic". Paper presented at the Fifth Annual Symposium of the Arabic Linguistic Society, University of Michigan, Ann Arbor.

Wood, Sidney. 1979. "A radiographic analysis of constriction location for vowels". *Journal of Phonetics* 7.25–43.

——. 1982. *X-ray and model studies of vowel articulation*. Lund: Department of Linguistics, Lund University.

Yeou, Mohamed. 1996. *Une étude expérimentale des consonnes postérieures et pharyngalisées de l'arabe standard*. Ph.D. diss., Sorbonne Nouvelle, Paris 3.

Younes, Munther. 1993. "Emphasis spread in two Arabic dialects". *Perspectives on Arabic linguistics*, V, ed. Mushira Eid and Clive Holes, 119–145. Amsterdam: J. Benjamins.

——. 1994. “On emphasis and /r/ in Arabic”. *Perspectives on Arabic linguistics*, VI, ed. Mushira Eid, Vincente Cantarino, and Keith Walters, 215–235. Amsterdam: J. Benjamins.

Zawaydeh, Bushra Adnane. 1997. “Gradient uvularization spread in Ammani-Jordanian Arabic”. *Perspectives on Arabic linguistics*, XI, ed. Elabbas Benmamoun, Mushira Eid, and Niloofar Haeri, 117–141. Amsterdam: J. Benjamins.

Melissa Barkat (University of Lyon-2)

Vowel Elision

Elision refers to the deletion of sound segments of differing natures (i.e. consonants, vowels, syllables). The factors responsible for this may be related to various phonetic, phonological, and prosodic constraints in the languages of the world. In French, for example, elision occurs so as to prevent vowel-to-vowel contact where vowels are attested at the initial position of words, as in the following examples:

le + *ami* → *l’ami* [lami] ‘friend’
si + *il* → *s’il* [sil] ‘if he’

Such a situation is impossible in Standard Arabic, where words always start with a consonant. If, nonetheless, two vowels do come into contact, less drastic solutions than vowel deletion are available, as illustrated in the following examples from Moroccan Arabic:

1(a) /bali + a/ → [bæljɑ] ‘old’
‘old + feminine singular morpheme’
1(b) /bldi + a/ → [bəldijɑ] ‘traditional’
‘traditional + feminine singular morpheme’
1(c) /ndu + a/ → [nduwɑ] ‘enemy’
‘enemy + feminine singular morpheme’

The possible solutions to avoid vowel-to-vowel contact are either to transform the first vowel [i] into the approximant [j], as in (1a), or to insert an approximant of the same quality as the first vowel, as in (1b) and (1c). The process of vowel elision in Arabic has often been dealt with in comparative studies from a diachronic point of view in an attempt to establish links between Classical Arabic and its vernacular forms. From this perspective, it is easy to understand the typical vocalic erosion attested in the vocalic systems of Western Arabic dialects. Indeed, Marçais (1977:24) underlined the fact that “les mots d’arabe maghrébin, lorsqu’on les rapporte aux prototypes anciens dont ils procèdent, présentent une très importante diminution du matériel vocalique” (Western Arabic words are characterized by an important reduction of vocalic material when compared to their ancient equivalents) and that “la diminution de la substance vocalique s’accroît d’Est en Ouest, les parlers marocains étant ceux où elle apparaît le plus réduite” (this diminution is more and more noticeable from east to west, Moroccan dialects attesting the most important vocalic reduction). This development and its effects on syllabic structure constitute a major typological feature distinguishing Western Arabic dialects from their Eastern counterparts (Barkat 2001).

Without lapsing into diachronic debates, the following items help in understanding the importance of short vowel elision:

	Standard Arabic	Translation		Moroccan Arabic
2a	*litāmuhā*	‘her veil’	2b	*ltæmha*
3a	*nadā*	‘dawn’	3b	*nda*
4a	*kitāb*	‘book’	4b	*ktæb*
5a	*katabtu*	‘I wrote’	5b	*ktəbt*
6a	*musāfir*	‘traveling’	6b	*msæfər*
7a	*yatīm*	‘orphan’	7b	*ytim*
8a	*safarjala*	‘quince’	8b	*sfržl*
9a	*ibnu ʾādam*	‘Adam’s son’	9b	*bnædəm*

On the one hand, short vowel deletion shatters both the opposition of quantity and the rhythmic organization that is directly linked to vocalic duration. On the other hand, this gives rise to the emergence of complex consonantal clusters where phonotactic rules are remarkably flexible. Besides, comparing basic word forms with their realization in pausal position in Standard Arabic leads us to see short vowel elision as an ancient, if limited, feature. Indeed, specialists are already familiar with the pausal rule: “Arabs do not end with a short vowel” (*al-ʿArab lā taqifu ʿalā mutaḥarrik*). Indeed, final short vowels are elided in pausal position (→ pausal forms): /## ʔal + mudiiru/ → [ʔalmudiːr] ‘the director’

This process also occurs in some Arabic vernaculars in prepausal context. In Moroccan Arabic, for example, the phonetic realization of the final vowel is optional in a large number of polysyllabic words:

10(a) [ntin(ɑ)] ‘you’
10(b) [ljum(ɑ)] ‘today’
10(c) [təmmæk(i)] ‘there’
10(d) [mitr(u)] ‘meter’
10(e) [tomobil(ɑ)] ‘car’

This list of examples is not exhaustive, as revealed by the treatment of the discontinuous morpheme of negation /ma…ši/:

(11) [ma klæʃ(i)] ‘he did not eat’

This vowel apocope can be explained by a prosodic factor, word → stress. As a matter of fact, in Western Arabic, where topicalization is preferentially expressed by stressing the penultimate syllable of the word, final vowels are ipso facto unstressed and subject to → vowel reduction (Benkirane 1998). The same type of predetermined vocalic elision due to stress contrast has been attested in several Arabic dialects, such as those spoken in the Levant (Brame 1973; Bohas 1988). In Palestinian Arabic, for example, the high vowel [i] is elided in unstressed open syllables, as in 13(a) and 13(b):

12(a) ['katab] ‘he wrote’
12(b) ['katabu] ‘they wrote’
12(c) [ka'tabna] ‘we wrote’
13(a) ['fihim] ‘he understood’
13(b) ['fihmu] ‘they understood’
13(c) [f'himna] ‘we understood’

Kenstowicz’s (1981:21) analysis of vowel elision in Levantine Arabic reveals that unstressed vowel elision is not applicable to the low vowel [a].

The influence of stress on vowel elision versus preservation occurs at another level whenever two accentuated words come into contact:

14(a) /msa + k/ → ['msæk] ‘your evening’ [evening + pronoun 2nd person, singular]
14(b) /msa# + l-xir/ → [msl'xer] ‘good evening to you’ [evening + definite article – good]
15(a) /tlata/ → ['tlætɑ] ‘three’
15(b) /tlata # snin/ → [tlt'snin] ‘three years old’

These examples from Moroccan Arabic show that vowel elision can be due to stress subordination. Indeed, stressed forms such as ['msa] and ['tlætɑ] lose their vowels when they are followed by a stressed word, their reduced form corresponding to the consonantal frame, i.e. [ms] and/or [tlt]. In other words, one could say that in a nominal syntagm of two words, a hierarchical rule requires that the rightmost word carry the stress. In conclusion, it should be noted that the tendency to interstress subordination, as well as the phenomena of vowel reduction and elision in unstressed syllables, seems to constitute important evidence for the description of rhythmic structure in Arabic dialects (Ghazali a.o. 2001; Hamdi a.o. 2004).

BIBLIOGRAPHICAL REFERENCES

Barkat, Melissa. 2001. “Vers l’identification automatique des parlers arabes”. *Langue et Linguistique* 7.47–76.

Benkirane, Thami. 1998. “Intonation in Western Arabic (Morocco)”. *Intonation systems*, ed. Daniel Hirst and Albert Di Cristo, 345–359. Cambridge: Cambridge University Press.

Bohas, Georges. 1988. “Accentuation et effacement de schwa dans le parler de Damas”. *Revue Québécoise de Linguistique* 16:1.39–60.

Brame, Michael K. 1973. “On stress assignment in two Arabic dialects”. *A Festschrift for Morris Halle*, eds. Stephen R. Anderson and Paul Kiparsky, 14–25. New York: Holt, Rinehart and Winston.

Ghazali, Salem, Rym Hamdi, and Melissa Barkat-Defradas. 2001. “Speech rhythm variation in Arabic dialects”. *Proceedings of the First International Conference on Speech Prosody, Aix-en-Provence (France), April 2001*, 331–334. www.isca-speech.org/archive/sp2002.

Hamdi, Rym, Melissa Barkat-Defradas, Emmanuel Ferragne, and François Pellegrino. 2004. “Speech timing and rhythmic structure in Arabic dialects: A comparison of two approaches”. *Proceedings of ICSLP Interspeech, Jeju (Korea) 4–8 October 2004*. CD-ROM.

Kenstowicz, Michael. 1981. “A note on cyclic stress in Levantine Arabic”. *Theoretical issues in the grammar of Semitic languages* (= *MIT Working Papers in Linguistics*, 3.), ed. Hagit Borer and

Joseph Aoun, 21–31. Cambridge, Mass.: Department of Linguistics and Philosophy, Massachusetts Institute of Technology.
Marçais, Philippe. 1977. *Esquisse grammaticale de l'arabe maghrébin*. Paris: Librairie d'Amérique et d'Orient and Adrien Maisonneuve.

Thami Benkirane (University of Fes)

Vowel Fronting → Vowel Raising

Vowel Harmony

Vowel harmony is a long-distance phonological process. It is defined as the phonetic influence of one vowel on another. For example, the phonetic quality of being a high vowel (i.e. /i/ or /u/) may influence or 'spread to' a local non-high vowel. If vowel height is harmonized in a particular language, then we expect a non-high vowel to surface as a high vowel. The phonetic quality that spreads and the direction it spreads is language-specific. Some languages may exhibit only leftward spreading, while other languages may exhibit only rightward spreading, while still others may exhibit both.

In Arabic, it appears that the Eastern dialects of the pre-Islamic period exhibited some form of vowel harmony (Versteegh 1997). The examples in (1) illustrate the differences in pronunciation between the Eastern and Western dialects. The phonetic quality of /i/ is harmonized in the Eastern dialects but not the Western dialects.

(1) Eastern	Western	
biʿīr	*baʿīr*	'camel'
minhim	*minhum*	'from them'

The bulk of the literature on vowel harmony in Arabic, however, analyzes the Palestinian dialect. Palestinian Arabic exhibits 'rounding harmony'. This means that the phonetic quality of being a round vowel (i.e. /u/) is harmonized. The vowel system of Palestinian Arabic is controversial between a three- and a five-vowel system. The traditional generative view adopts the five-vowel system, illustrated in (2a) (Younes 1993), whereas others have adopted the three-vowel system, as in (2b) (Shahin, forthcoming), in accordance with the standard view of other dialects of Arabic. Because of the nature of rounding harmony in Palestinian Arabic, simply adopting the three vowel system (2b) is sufficient.

(2a) *ī ē* *ū ō*
ā
(2b) *ī* *ū*
ā

In Palestinian Arabic, if the root vowel has the phonetic quality of being round, then the quality of roundness spreads to other high vowels in the prefix (Kenstowicz 1981; Abu-Salim 1987; Yoshida 1993; Monahan 2002). This is illustrated in (3). Suffixes in Palestinian Arabic are immune to rounding harmony.

(3a)	/yi-drus/	[yúdrus]	'he studies'
(3b)	/ti-ktub/	[túktub]	'she writes'
(3c)	/ti-ktub-i/	[túktubi]	'you [sg. fem.] write'

In (3), the root vowels are round, while the prefix vowels are high unrounded vowels. Therefore, the phonetic quality of roundness spreads to the prefix vowel in a regressive, or right-to-left, manner. Notice that the final vowel in (3c) is a constituent of the suffix and is, therefore, not subject to harmony.

Vowel harmony, however, is not limited to the verbal domain in Palestinian Arabic. It is also observable in the nominal domain, as in (4).

(4a)	/furin/	[fúrun]	'oven'
(4b)	/ʿuris/	[ʕúrus]	'wedding'

An interesting fact of rounding harmony in Palestinian Arabic is that vowels left of the stressed syllable are not vulnerable to harmony, as in (5a) vs. (5b).

(5a)	*bitšū́f*	'she sees'
(5b)	*túktub*	'she writes'

The next examples of vowel harmony are from dialects of northern Yemen, as presented by Behnstedt (1985). Some lexical forms in the dialects of Yemen are differentiated between *i*-type and *a*-type verbs. This is illustrated in the harmony exhibited in the endings of the examples in (6), all of which are in the 3rd person singular feminine.

(6) *a*-type ‘to write’	*i*-type ‘to drink’
a. *katabat*	*širibit*
b. *katabah*	*širibih*

The examples in (6a) were recorded from the mideastern regions of the North Yemen dialect, whereas the examples in (6b) were recorded in the southwestern regions of the North Yemen dialect. Notice the difference in the endings between the *a*-type and *i*-type verbs. For example, in (6a), the ending for the *a*-type is *-at*, whereas the ending for the *i*-type is *-it*.

The final examples of vowel harmony discussed here are from the Bedouin dialects of the Northern Sinai littoral (de Jong 2000). The particular dialects presented are Balawiy Arabic, the dialect of Biliy; Rmēliy Arabic, the dialect of Rmēlāt; and Swērkiy Arabic, the dialect of Sawārkah. The initial vowels of the morphological templates of the regular imperfect form in these dialects differ depending upon the root vowel. The morphological templates of the regular imperfect are either $yaC_{1}C_{2}aC_{3}$, $yuC_{1}C_{2}uC_{3}$, or $yiC_{1}C_{2}iC_{3}$. The vowel of the imperfect prefix harmonizes with the root vowel. This is illustrated in (7).

(7) imperfect	‘to drink’ (*a*-type)	
	sg.	pl.
3rd masc.	*yášṛaḅ*	*yášṛaḅaw*
3rd fem.	*tášṛaḅ*	*yášṛaḅan*
2nd masc.	*tášṛaḅ*	*tášṛaḅaw*
2nd fem.	*tášṛaḅay*	*tášṛaḅan*
1st c.	*ášṛaḅ*	*nášṛaḅ*

imperfect	‘to sit’ (*u*-type)	
	sg.	pl.
3rd masc.	*yúgʿud*	*yúguʿduw*
3rd fem.	*túgʿud*	*yúguʿdin*
2nd masc.	*túgʿud*	*túguʿduw*
2nd fem.	*túguʿdiy*	*túguʿdin*
1st c.	*ágʿud*	*núgʿud*

imperfect	‘to grab, take hold of’ (*i*-type)	
	sg.	pl.
3rd masc.	*yímsik*	*yímiskuw*
3rd fem.	*tímsik*	*yímiskin*
2nd masc.	*tímsik*	*tímiskuw*
2nd fem.	*tímiskiy*	*tímiskin*
1st c.	*ámsik*	*nímsik*

Closer inspection of the data in (6) reveals that not only are the prefix vowels harmonized but so are the ending vowels (de Jong 2000). For example, compare the endings of *yášṛaḅaw* ‘drink [3rd pers. pl. masc.]’ and *yúguʿduw* ‘sit [3rd pers. pl. masc.]’. The verb ‘to drink’ is *a*-type, whereas the verb ‘to sit’ is *u*-type. Because the endings are also harmonized, the ending of ‘drink [3rd pers. pl. masc.]’ is *-aw*, while the ending of ‘sit [3rd pers. pl. masc.]’ is *-uw*.

Vowel harmony is not an uncommon process in various dialects of Arabic, and the examples are not limited to those presented here. The Palestinian facts, which are considerably well documented, exhibit both left and rightward spreading of the phonetic quality round. The vowel harmony demonstrated in the Yemen and Bedouin dialects appears to be the influence of all the phonetic qualities of the root vowel, and the spreading seems to be bidirectional. Although a productive phonological process, vowel harmony in Arabic does not appear to bear on semantic interpretation.

BIBLIOGRAPHICAL REFERENCES

Abu-Salim, Issam. 1987. “Vowel harmony in Palestinian Arabic: A metrical perspective”. *Journal of Linguistics* 23.1–24.

Behnstedt, Peter. 1985. *Die nordjemenitschen Dialekte*. I. *Atlas*. Wiesbaden: L. Reichert.

Jong, Rudolph E. de. 2000. *A grammar of the Bedouin dialects of the Northern Sinai littoral: Bridging the linguistic gap between the eastern and western Arab world*. Leiden: E.J. Brill.

Kenstowicz, Michael. 1981. “Vowel harmony in Palestinian Arabic: A suprasegmental analysis”. *Linguistics* 19.449–465.

Monahan, Philip. 2002. “Harmony and metrical structure in Palestinian Arabic”. *Proceedings of the 2002 Texas Linguistics Society*, ed. Augustine Agwuele and Sang-Hoon Park, 182–198. Austin, Tex.: Texas Linguistic Forum.

Shahin, Kimary. Forthcoming. “Vowel innovation in Arabic: Inductive grounding and pattern symmetry”. *Research in Afroasiatic grammar*, III, ed. Jacqueline Lecarme, Jean Lowenstamm, and Ur Shlonsky. Amsterdam and Philadelphia: J. Benjamins. [Also available from Rutgers Optimality Archive: http://roa.rutgers.edu, ROA573-0203.]

Versteegh, Kees. 1997. *The Arabic language*. New York: Columbia University Press.

Yoshida, Shohei. 1993. “Licensing of empty nuclei: The case of Palestinian vowel harmony”. *Linguistic Review* 10.127–159.

Younes, Munther. 1993. “Emphasis spread in two Arabic dialects”. *Perspectives on Arabic linguistics*, V, ed. Clive Holes and Mushira Eid, 119–145. Amsterdam and Philadelphia: J. Benjamins.

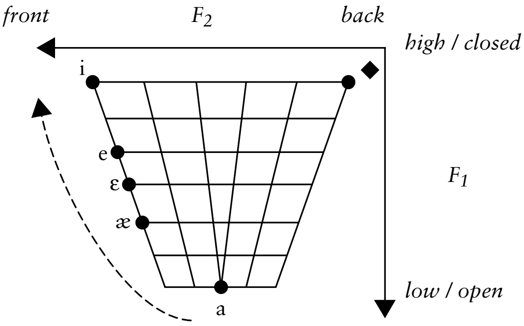
PHILIP J. MONAHAN
(University of California at San Diego)

Vowel Lowering → Vowel Backing

Vowel Raising

In spoken Arabic, vocalic segments are subject to a wide range of qualitative variations, most of them being determined by the consonantal environment (→ vowel backing). One of these changes concerns the raising and fronting of the vowel [a] in medial position (i.e. inside the word) and/or in final position (i.e. at the end of feminine nouns and adjectives that are morphologically marked by the ending [a] in Arabic). This process – referred to by the Arab grammarians with the term → *ʾimāla* – is often connected with the phenomenon of *umlaut* and defined as "the inclination of [a] due to the umlauting influence of [i]" (Kaye 1997:198). But in the case of Arabic, this definition is not entirely satisfactory. Indeed, *ʾimāla* should be regarded as an unconditioned process that applies to all [a] inside a word – unless they appear in the environment of any blocking segment (i.e. pharyngealized consonants) – as well as final [a] if such a phenomenon is attested in the dialect in question. In phonetic terms, *ʾimāla* corresponds to a raised and fronted realization of the open central vowel [a], which could be pronounced [æ] (raised low vowel), [ɛ] (lower mid vowel), [e] (higher mid vowel), or even [i] (high vowel). According to this principle, words such as [ʔiba:d] 'slaves' could be pronounced [ʔibæ:d], [ʔibɛ:d], [ʔibe:d], or [ʔibi:d] and [nsa] 'women' as [nsæ] or [nse], depending on the speaker's dialectal background. Variations in the degree of raising and fronting enable us to distinguish between 'strong' vs. 'medium' *ʾimāla* (Fig. 1). Basically, internal *ʾimāla* – which is widely developed over all the Arabic domain – is of the medium type (i.e. [a] > [æ] [ɛ] or [e]), final *ʾimāla* being strongest where attested (i.e. [a] > [e] or [i]).

Figure 1. Raising and fronting of the central vowel [a]; the arrow shows the *ʾimāla* movement and its extent in the vocalic space



On the acoustic point of view, *ʾimāla* corresponds to a lowering of F_1 and a raising of F_2 (where F_1 is related to the open/closed vertical axis, corresponding to the openness of the jaw and F_2 on the horizontal axis, corresponding to front/back position of the tongue in the mouth; Benkirane 1982). The values observed for F_1 and F_2 in various dialects of Arabic confirm the idea that – in the same way as in the pre-Islamic Arabic dialects (Vollers 1906) – modern vernaculars are variable with regard to this feature. In a comparative acoustic study dealing with final *ʾimāla* in Algerian, Moroccan, Syrian, and Lebanese vernaculars, Barkat a.o. (1997) showed the acoustic variations of the vowel [a] at the end of feminine words, adjectives, and perfect verbs. It appears that final *ʾimāla* is not attested in Western dialects (i.e. Moroccan and Algerian dialects), in which the values of F_1 and F_2 correspond to a central realization of [a] (i.e. mean F_1 = 657 Hz; mean F_2 = 1326 Hz in Algerian and mean F_1 = 634 Hz; mean F_2 = 1173 Hz in Moroccan), whereas it is present in Syrian (i.e. mean F_1 = 707 Hz; mean F_2 = 1568 Hz) and Lebanese dialects (i.e. mean F_1 = 340 Hz; mean F_2 = 2025 Hz), where it produces a distinctive higher pronunciation of the same vowel (Fig. 2).

The spectrograms below show the spectral characteristics of final [a] in the word [malika] 'queen' in the different dialects under study. The degree of *ʾimāla*, moving gradually from none, in Western dialects, to medium in the Syrian dialect spoken in the city of Abu-Kamal (located at the border of Iraq), to strong in the Lebanese dialects spoken in Beirut as well as in the Syrian dialect of Homs.

Following Cantineau (1960), Eastern dialects appear to be typically raising dialects. Indeed, strong *ʾimāla* – though it is regarded in the literature as relatively rare – is reported to be attested in the oasis of Sukhne east of Palmyra (e.g. [ʃi:reb] 'moustache') and in the dialect spoken in Homs (e.g. [mæliki] 'queen'). As for medium *ʾimāla* – which is, on the contrary, widely attested – it covers the major part of the oriental Arabic domain, like, for example, in the Syrian dialects of the Ḥōrān (e.g. [bense] 'he

Figure 2. Distribution of [a] in word-final position in F_1/F_2 acoustic space. Eastern dialects (in white) attest an anterior realization compared to Western dialects (in black).

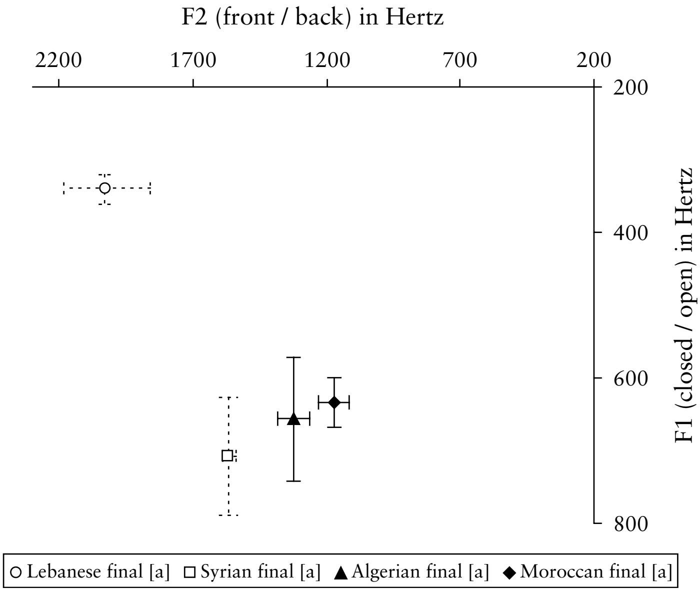
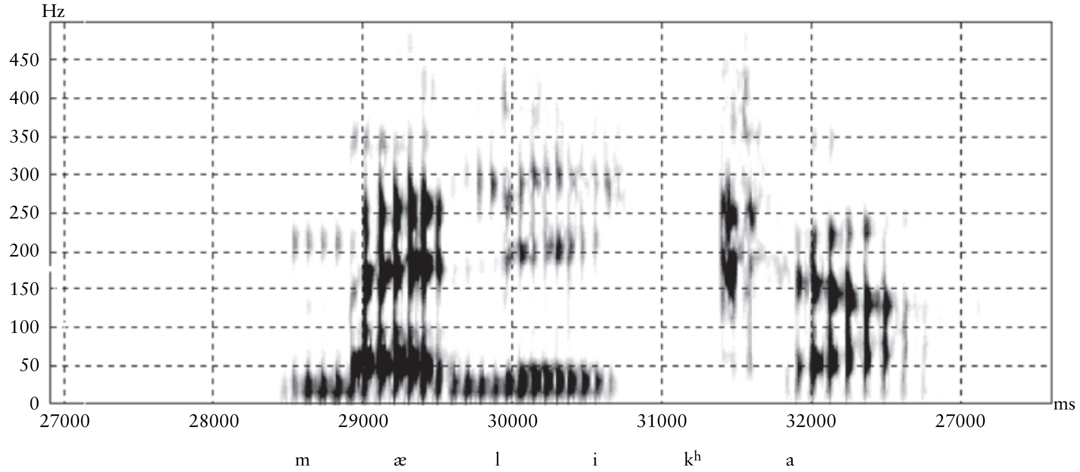
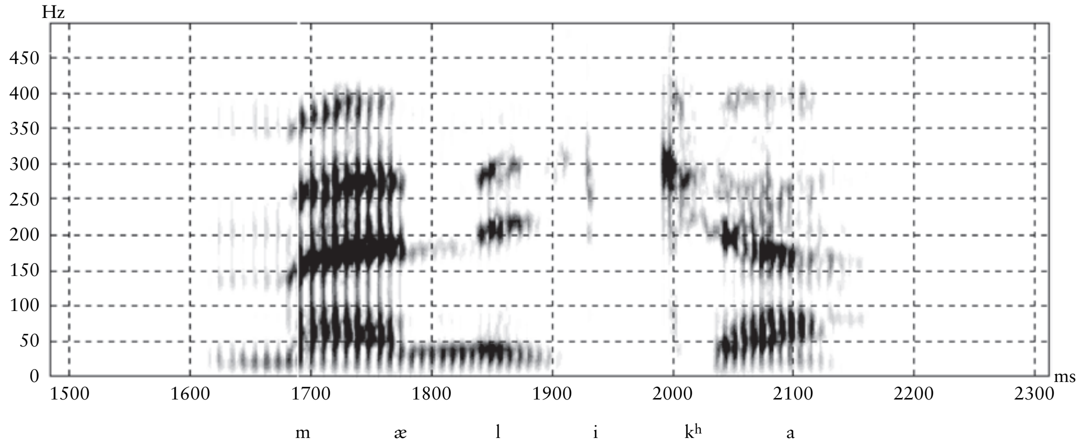


Figure 3. Pronunciation of *malika* in Arabic dialects

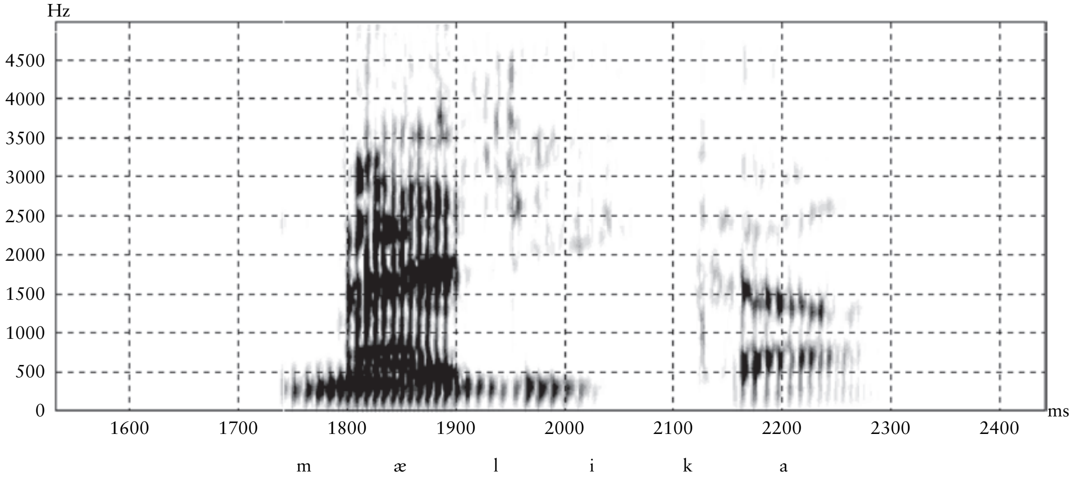
(3a) Moroccan (spoken in Casablanca)



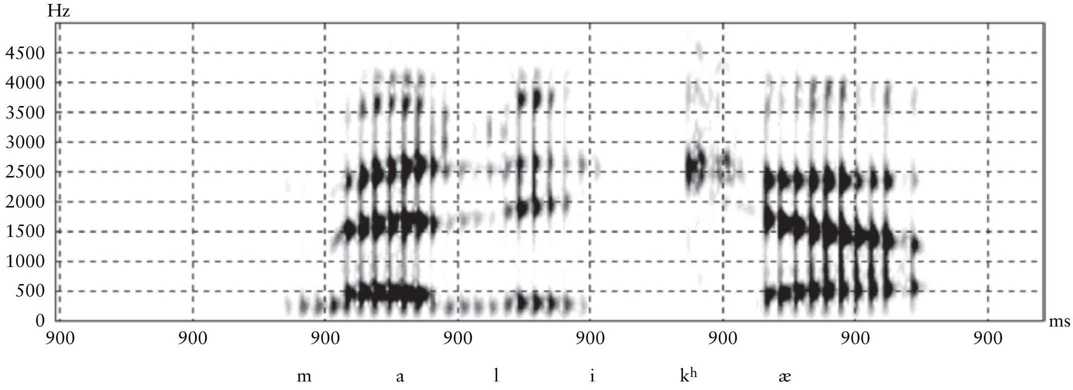
(3b) Algerian (spoken in Batna)



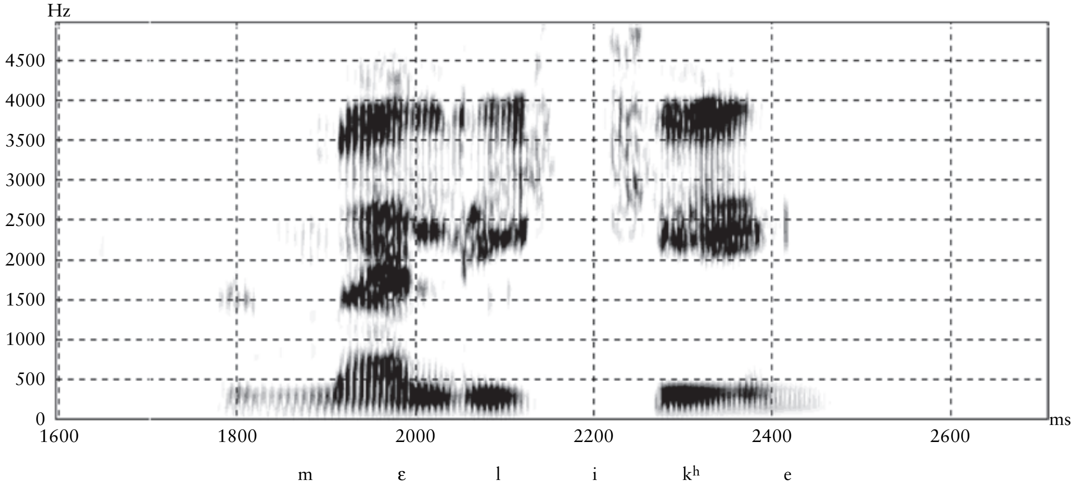
(3c) Algerian (spoken in Wahran)



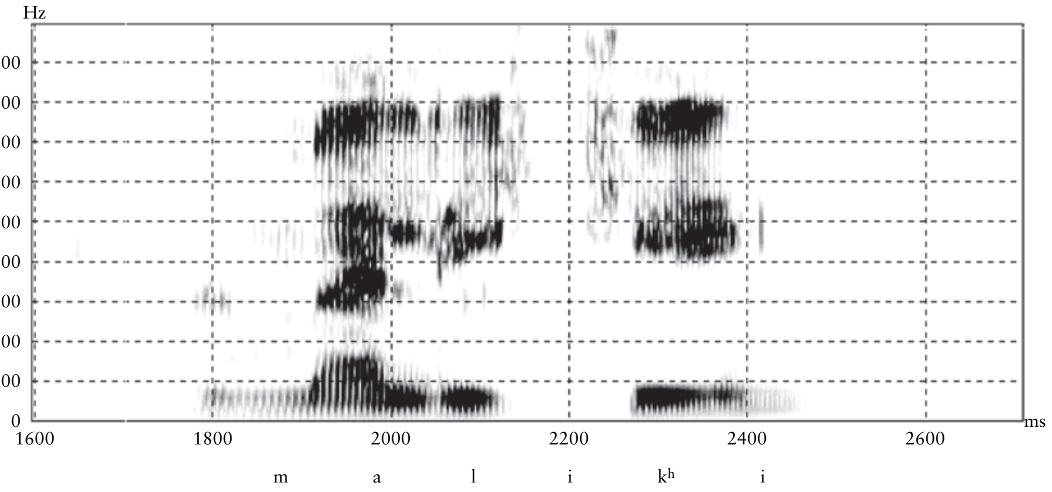
(3d) Syrian (spoken in Abu-Kamal)



(3e) Syrian (spoken in Homs)



(3f) Lebanese (spoken in Beirut)



will forget her'; Bergsträßer 1915) and Damascus (e.g. [samake] 'fish'; Angoujard 1981) as well as in Beirut (e.g. [mɛli:ke] 'queen') (Naim-Sanbar 1986). Now, although Maghrebi dialects are said to be nonraising, medium *ʾimāla* does exist in the region of Annaba in northeast Algeria and in the south of the territory. It is also the case in some northern coastal Tunisian villages, such as Bizerte (e.g. [meʔ] 'water'), as well as in the Bedouin dialects of southern Tunisia (in Gabes and Neffzaoua, e.g. [mʃeʔ] 'he went'; Marçais 1977) and in Korba (Walter 1989). Nevertheless, the fact that *ʾimāla* is attested in some dialect or other should not lead one to consider it as a set phenomenon. Indeed, in a study dealing with vowel raising in the emerging dialect of Amman, Al-Wer (1998) notes it can be subjected to various sociolinguistic factors, for instance competition between urban and rural vernaculars. In this point, and at the level of the Levant region, the nonraising dialects are generally nonurban, whereas socially dominant dialects are raising. This factor generally furthers the emergence of *ʾimāla* in the former vernaculars. Note that the contrary is true in Egypt, where it is the nonraising dialect of Cairo which is dominant (Woidich 1994). The parameters of speakers' age and gender also play an important role in the evolution of *ʾimāla*. Kaye (1997) notes that *ʾimāla* is especially characteristic of women's speech. In her study, Al-Wer (1998) shows that young females are generally leading in the leveling out of regional features related to traditional Jordanian Arabic and/or to Palestinian Arabic, so as to adhere to the rules attested in prestigious Levantine dialects (Damascus, Beirut, etc.). Conversely, in the small town of Korba (in Tunisia), young male speakers are more active in the disappearance of *ʾimāla*, which is regarded there as a local nonprestigious feature (Walters 1991). Last but not least, Versteegh (1997) has proposed the hypothesis of a phonemic contrast between central [a] and sounds that are generally held to be its raised allophones. He points out such contrastive pairs as [tˤe:leb] 'striving' vs. [tˤa:leb] 'student' in North Syrian Arabic (dialect spoken in Aleppo) This situation may be explained on the one hand by the general rule of [a]-raising/fronting in *all* consonantal contexts in North Syrian, and on the other hand by the synchronic rules governing the pronunciation of Classical Arabic loanwords. The same type of argument is advanced for some contrasting pairs found in Palestinian Arabic (Ibrahim 1982).

Bibliographical references

Al-Wer, Enam. 1998. "Raising of /a/ and related vocalic movements in the emerging dialect of Amman". *Proceedings of the Third International Conference on Arabic Dialectology, March 29 – April 2, Malta*, ed. Manwel Mifsud, 77–82. Malta: Salesian Press.

Angoujard, Jean-Pierre. 1981. "Marqueur du féminin et système vocalique dans l'arabe de Damas". *Arabica* 28.345–357.

Barkat, Melissa, Jean-Marie Hombert, and Catherine Taine-Cheikh. 1997. "Détermination d'indices acoustiques pour l'identification des parlers arabes:

Etats d'[a]mes vocaliques". *Actes des Journées d'Etudes Linguistiques: La voyelle dans tous ses états, Université de Nantes*, 141–146. Nantes: Université de Nantes.

Benkirane, Thami. 1982. "Durée, prosodie et syllabation en arabe marocain". *Travaux de l'Institut de Phonétique d'Aix* 8.49–83.

Bergsträßer, Gotthelf. 1915. "Sprachatlas von Syrien und Palästina". *Zeitschrift des Deutschen Palästina-Vereins* 38.169–222 (and maps XX–LXII).

Cantineau, Jean. 1960. *Cours de phonétique arabe*. Paris: Klincksieck.

Ibrahim, Muhammad. 1982. "On the contrast between [æ] and [a] in Modern Arabic". *Proceedings of the 13th International Congress of Linguistics*, 624–628. Tokyo: Tokyo Press.

Kaye, Alan S. (ed.). 1997. *Phonologies of Asia and Africa*. Winona Lake, Ind.: Eisenbraun.

Marçais, Philippe. 1977. *Esquisse d'arabe maghrébin*. Paris: Maisoneuve.

Naim-Sanbar, Samia. 1986. *Le parler arabe de Râs-Beyrouth (Liban)*. Paris: Geuthner.

Versteegh, Kees. 1997. *The Arabic language*. Edinburgh: Edinburgh University Press.

Vollers, Karl. 1906. *Volkssprache und Schriftsprache im alten Arabien*. Strasburg: K.J. Trübner. (Repr., 1981.)

Walters, Keith. 1989. *Social change and linguistic variation in Korba, a small Tunisian town*. Ph.D. diss., University of Texas at Austin.

——. 1991. "Women, men and linguistic variation in the Arab world". *Perspectives on Arabic linguistics*, III, ed. Bernard Comrie and Mushira Eid, 199–229. Amsterdam: J. Benjamins.

Woidich, Manfred. 1994. "Cairo Arabic and the Egyptian dialects". *Proceedings of the First International Conference on Arabic Dialectology*, ed. Martine Vanhove and Dominique Caubet, 493–507. Paris: Institut National des Langues et Cultures Orientales.

Melissa Barkat (University of Lyon-2)

Vowel Reduction

Basically, the phenomenon of vowel reduction results from the shortening of a linguistic unit. This shortening can be due to sociolinguistic factors, phonetic evolution, apocope, phonotactic constraints, temporal adjustments dictated by prosodic events such as → sandhi (i.e. the fusion of sounds across word boundaries, or alteration of sounds due to neighboring sounds), accentual contrast, increase in the speech rate, etc.

Consider the following examples from Moroccan Arabic:

(1) /## l + magana# djal# xa+ i ## /
the + watch # of # brother + pronoun 1st person singular
'my brother's watch'
(1a) [lmagæna djæl xɑj]
(1b) [lmagæna di xɑj]
(1c) [lmagæna dxɑj]

(2) /## ras # mal+ u # kbir ## /
head + money # pronoun, 3rd person singular # big
'He is wealthy'
(2a) [rɑs mælu kbir]
(2b) [rɑs mælu kbir]

Although (1a), (1b), and (1c) exhibit remarkable formal differences, they are still semantically equivalent. In fact, the functional particle connecting the two substantives can be encoded either in a full form, [djæl], or in reduced forms, [di] (a variant attested in eastern Morocco) or [d]. Similarly in (2a) and (2b), the open vowel [ɑ] of [rɑs] 'head' is reduced to schwa [ə].

Examples (1a), (1b), and (1c) show a steady diminution in the number of sound segments: segments have been purely and simply deleted. For vocalic segments, it is important to notice that reduction concerns the length of the vowel and not the vowel itself. In other words, there is a limit on the process of reduction, making vowel reduction a different process than → vowel elision. Thus, examples (1a) and (1b) illustrate the deletion of more than one segment, while examples (2a) and (2b), which have the same number of segments, illustrate vowel reduction. Traditional studies in Classical Arabic report that final vowels are not always deleted in pausal context. Indeed, according to Cantineau (1960:115), Arabic grammarians already observed "diverses réductions de ces voyelles à la quantité ultra-brève qu'ils nomment *rawm* et *ɔišmām*, l'articulation de la voyelle se bornant à une simple 'intention', à une 'odeur' fugitive" (several occurrences of reduction of these overshort vowels that they call *rawm* and/or *ʾišmām*, of which the articulation is nothing else than an 'articulatory intention', a fleeting 'smell'). Vocalic quantity does not function the same way in Modern Standard and in Moroccan Arabic. In the latter, vowels occurring in open syllables at the end of words are realized and perceived as being short: "The stable vowels are relatively long except at the end of words where they are short" (Harrell 1962:10). In fact, in prepausal context there is a significant, supraliminal temporal contrast between vowels occurring in CvC sequences compared with Cv sequences (where C corresponds to 'Consonant' and v to 'Vowel'). Other things being equal, the vowel in CvC is realized with a longer duration than in Cv sequences (Benkirane 1998; Abu-Mansour 1992). This result seems to contradict the Closed Syllable Vowel Shortening principle described by Maddieson (1985:206). Vocalic reduction is often referred to in phonological descriptions. For example, in the following rule vv → v/_CC, Benhallam (1980:4) states that a long vowel (vv) is reduced when it is followed by two consonants. This rule can be applied to several verbal paradigms in Arabic:

(3a) [tafiːqu] 'you are waking up' (imperfect, 2nd person singular, masculine)
(3b) [tafiqna] 'you are waking up' (imperfect, 2nd person plural, feminine)
(4a) [taxaːfu] 'you are scared' (imperfect, 2nd person singular, masculine)
(4b) [taxafna] 'you are scared' (imperfect, 2nd person plural, feminine)
(5a) [tasuːqu] 'you are driving' (imperfect, 2nd person singular, masculine)
(5b) [tasuqna] 'you are driving' (imperfect, 2nd person plural, feminine)

In these examples, the three long vowels of Standard Arabic [iː], [uː], and [aː] are respectively shortened to [i], [u], and [a] when they occur in closed syllables. In other words, this vocalic reduction constitutes a temporal adjustment that can be explained by the existence of a syllable weight constraint: the vocalic shortening prevents the emergence of internally superheavy syllables. Vocalic reduction also occurs for the same reason at word boundaries in Modern Standard Arabic:

(6a) /fii # l + maktabati/
'in # the + library'
(6b) [filmaktaba]

However, this rule of vocalic reduction in the context of a closed syllable is subject to many exceptions (see Brockelmann 1961 for further details). On the phonetic and acoustic level, vocalic reduction modifies the prosodic proprieties of the segments, i.e. duration, intensity, and fundamental frequency (Fo). The spectral characteristics of the reduced vowel show clear centralization close to that of the neutral vowel schwa, i.e. [ə] (Lindblom 1963). As for the relation existing between vocalic reduction and destressing, this topic is often referred to in the acoustic literature. It is generally observed that vowel quality is more marked in stressed syllables (where vowels are long and plain) than in unstressed syllables (where vowels are centralized and reduced). It is also often observed that stressed composed nouns are subject to subordination rules that give rise to destressing as well as syllabic nuclei reduction. Such a process can be observed in (2b).

BIBLIOGRAPHICAL REFERENCES

Abu-Mansour, Mahasen Hasan. 1992. "Closed syllable shortening and morphological levels". *Perspectives on Arabic linguistics*, IV, ed. Ellen Broselow, Mushira Eid, and John McCarthy, 47–75. Amsterdam and Philadelphia: J. Benjamins.
Benhallam, Abderrafi. 1980. *Syllable structure and rule types in Arabic*. Ph.D. diss., University of Florida.
Benkirane, Thami. 1998. "Intonation in Western Arabic (Morocco)". *Intonation systems*, ed. Daniel Hirst and Albert Di Cristo, 345–359. Cambridge: Cambridge University Press.
Brockelmann, Carl. 1961. *Grundriss der vergleichenden Grammatik der semitischen Sprachen*, I. Repr., Hildesheim: G. Olms.
Cantineau, Jean. 1960. *Cours de phonétique arabe*. Paris: Klincksieck.
Harrell, Richard S. 1962. *A short reference grammar of Moroccan Arabic*. Washington, D.C.: Georgetown University Press.
Lindblom, Bjorn. 1963. "Spectrographic study of vowel reduction". *Journal of the Acoustical Society of America* 35.1773–1781.
Maddieson, Ian. 1985. "Phonetic cues to syllabification". *Phonetic linguistics: Essays in honor of Peter Ladefoged*, ed. Victoria A. Fromkin, 203–221. New York: Academic Press.

THAMI BENKIRANE (University of Fes)

W

Waḍᶜ al-Luġa

The phrase *waḍᶜ al-luġa*, which may be translated as 'the founding of language', represents a concept that is central to classical Muslim scholarly thinking about language. Language in that thinking was entirely invented. That is to say, it owed its existence to a process of deliberate assignment of patterned vocal utterances – or components of such utterances – to meanings, of *ʾalfāḏ̣* (→ *lafḏ̣*) to *maᶜānī* (→ *maᶜnā*). The relationship between the utterances and the meanings was not a natural or intrinsic relationship. In principle, an utterance could have any meaning. That an utterance had a particular meaning was due entirely to its being assigned to that meaning. The meaning of an utterance had to be learned; it could not automatically be known from the utterance.

The question of whether language arose from *thésis* or *phúsis*, so famously debated among the Greeks, never assumed major importance in Islam. Greek philosophers were not concerned with the question of who had established language, but they wished to find out to what extent words reflect reality (→ sound symbolism). Muslim thinkers were hardly interested in this aspect of the debate. They almost universally agreed that language had its origin in *thésis* in the sense of 'imposition' (this term being more or less equivalent to *waḍᶜ*). Much more debated was the question of agency: who assigned the *ʾalfāḏ̣* to their *maᶜānī* – God or human society? For most Muslims, the point of departure was the Qurʾānic verse 2/31, "He taught Adam all the names" (*ᶜallama ʾĀdama l-ʾasmāʾa kullahā*). Orthodox grammarians like Ibn Fāris (d. 395/1005) believed that this was enough evidence that God had created language (*Ṣāḥibī* 5–11). According to most Muᶜtazilite thinkers, however, language was a human invention. One of the most comprehensive accounts of the origin of speech is that by Ibn Jinnī (d. 392/1002). He states (*Xaṣāʾiṣ* I, 40–47) that the question is whether language was created by God (*waḥy, tawqīf*), or instituted by human convention (*iṣṭilāḥ, tawāḍuᶜ*). After presenting arguments for both positions, he declares himself unable to decide the matter (Méhiri 1973:91–118). For a survey of opinions on the origin of speech, see Loucel (1963–1964) and Versteegh (1997:101–114).

The question was not intensely debated by later grammarians, since how one answered the question had no bearing on the actual character of language as an invention (Weiss 1974). As a thing invented, language, if it was to remain constant over time, had to be preserved. This required an effective transmission process, implemented partly through consistency in language usage across generations (*tawātur*; Weiss 1984), but also in large part through the conscious efforts of linguistic scholars – in particular, lexicographers – to secure normative usage in written documents.

The idea of the founding of language appears quite early in Muslim scholarly writing. However, it did not become the subject matter of a separate linguistic discipline until the science of the founding of language (*ᶜilm al-waḍᶜ*) arose in the 16th century (Weiss 1966, 1987). The stimulus for this late Muslim science was a short treatise by al-ʾĪjī (d. 756/1355) entitled *ar-Risāla al-waḍᶜiyya*, which in the century after its writing spawned a number of commentaries (van Ess 1966). Commentaries eventually led to manuals that summed up the science and could be memorized by students. Replete with a basic text, a plethora of commentaries and glosses, and manuals for students, the science of the founding of language was able to enter the curriculum of Muslim higher education, where it remained into the beginning of the modern age.

The emergence of this science is probably to be explained by the sense that an idea so crucial to the entire body of linguistic sciences as well as to sciences dependent on those sciences, such as *fiqh, ʾuṣūl al-fiqh* (Makdisi 1984), and *tafsīr,* had to be thoroughly explained in order to be considered viable. The method of explanation developed in the science of the founding of language entailed, first of all, a dissection of language into all its meaning-bearing components and then a classification of the various ways in which these components were assigned by the founder(s) of language to their meanings. Meaning-bearing components included not only words but also forms of words, as distinct from the substrates in which the forms adhered. Each could have its separate meaning. Beyond words, there were also prefixes, suffixes, inflections, and so on. Also included were syntactic and other structures made up of words. The founder of language was in fact something of an architect, the one who set up all structural features of language. For example, by assigning a vocable to a meaning made up of a 'happening' (*ḥadaṯ*) considered in respect of its attribution to something lying outside the meaning, the founder of language created a verb. The verb was thus understood to be a type of word that was dependent for the intactness of its meaning on something outside that meaning. It was a word that had to be linked to a subject (*ḏāt*) to make full sense. The subject was designated by means of a noun. Unlike the verb, the noun is intact without need for an outside element to complete its meaning. For that reason, a noun can have a purely naming function. A verb does more than name; it demands linkage with another word. If one wants merely to name an action, one does not use a verb but rather a → *maṣdar*, or verbal noun. This difference between verb and noun is among the many features of language that the founder designed in the process of instituting language (Weiss 1976).

The science of the founding of language, as developed in the commentaries and glosses on al-ʾĪjī's treatise, eventually set up two different classifications of the founding event, the assignment of a vocable or element of a vocable to its meaning. One such classification distinguished a generic assignment from an individualized assignment (Weiss 1966:114–117). A generic assignment (*waḍᶜ nawᶜī*) consisted of the assignment of an element common to a group of vocables to a particular meaning, whereas an individualized assignment (*waḍᶜ šaxṣī*) consisted of the assigning of an element unique to a particular vocable or of an entire but indivisible vocable to a particular meaning. Thus, the form common to *kātib*, *ḍārib, ḥāfiẓ, ᶜālim, šākir, kāfir*, etc. – the *fāᶜil* form – is assigned to its meaning by way of a generic assignment, whereas the consonantal root *ḍ-r-b*, being a unique consonantal root, is assigned to its meaning by way of a indi-vidualized assignment. Nouns such as *ʾasad, zayd*, etc., which are not divisible into formal and material elements, are also assigned to their meanings by way of an individualized assignment.

The second classification reflects jurisprudential interests. It is concerned with the founding of general terms and specific terms (*ᶜāmm* and *xāṣṣ*). The assignment of a vocable to a species to form a common noun such as *man*, or *bird*, or *mountain* is considered a general assignment (*al-waḍᶜ al-ᶜāmm*), while the assignment of a vocable to a particular individual within a species to form a proper name, such as *zayd*, is considered a 'specific assignment' (*al-waḍᶜ al-xāṣṣ*). Two categories of words, however, proved problematic: pronouns and particles. A pronoun such as 'he' (*huwa*) could be said in a certain sense to have been assigned to the general notion of a single male person; and yet, in actual usage, it was not this notion that one was trying to get across; rather, the word was used to refer to a particular male person as opposed to the idea of a single male person. Particles such as 'and' (*wa-*) likewise could be said in a certain sense to have been assigned to the general notion of 'and-ness'; and yet, in actual usage, they were clearly meant to convey not the general notion, but a specific context-dependent relationship of conjunction between two things. With both kinds of words there seemed to be both general and specific elements. Therefore, al-ʾĪjī proposed the rubric *al-waḍᶜ al-ᶜāmm li-mawḍūᶜ lahu xāṣṣ* as a category distinct from the other two types that entailed the categories *ᶜāmm* and *xāṣṣ*, which he modified by placing them under the extended rubrics *al-waḍᶜ al-ᶜāmm li-mawḍūᶜ lahu ᶜāmm* and *al-waḍᶜ al-xāṣṣ li-mawḍūᶜ lahu xāṣṣ*.

Combined with the division into *ᶜāmm* and *xāṣṣ*, the following three categories were distinguished in individual assignment (*waḍᶜ šaxṣī*; Weiss 1966:128):

i. *al-waḍᶜ aš-šaxṣī al-ᶜāmm li-mawḍūᶜ lahu ᶜāmm*. Linguistic elements established in this way include 'solid [i.e. nonderived concrete] nouns' (*ism jāmid*), *maṣdar*s, the 'matter' (i.e. the radicals forming the root) of derived nouns, the 'matter' of verbs. These elements have been universally established for a universal meaning, for instance the word 'man', whose idea is identical with its meaning (Weiss 1966:110).
ii. *al-waḍᶜ aš-šaxṣī al-ᶜāmm li-mawḍūᶜ lahu xāṣṣ*: Linguistic elements established in this way include personal pronouns, demonstrative pronouns, relative pronouns, and particles, which have been established as universal means to refer to a particular meaning (Weiss 1966:95ff.; for the particles, see Weiss 1966:110).
iii. *al-waḍᶜ aš-šaxṣī al-xāṣṣ li-mawḍūᶜ lahu xāṣṣ*: Linguistic elements established in this way include proper names and generic names; for such expressions, concept and that for which the expression is established are identical, i.e., the idea itself becomes the meaning (Weiss 1966:109).

The same categories were distinguished in generic assignment (*waḍᶜ nawᶜī*). Although there was some disagreement concerning the assignment of radicals and forms, most authors agreed that the three radicals *ḍ-r-b* had been assigned individually to all words containing these radicals, whereas the pattern (form) of these words had been assigned generically (Weiss 1966:126–127). To the categories of generic assignment also belonged constructions like the vocative phrase or the verbal sentence. There was considerable discussion among writers on *waḍᶜ al-luġa*, for instance about the vocative phrase *yā fulān*: some writers maintained that it had been established for the universal idea of requiring someone's presence, whereas others asserted that it had been established for particular instances of requiring the presence of individual persons by means of a universal idea (Weiss 1966:129–130). In other words, they debated the question of whether this construction had been established for a meaning of its own, or for a meaning it acquired when expressed in words. Likewise, concerning the → *ʾisnād* construction of a subject (*musnad ʾilayhi*) and a predicate (*musnad*), writers were divided about the question of whether construction had been established for the general idea of a 'complete relation' (*nisba tāmma*), or for particular ideas, such as Zayd's hitting in *zayd ḍārib*, which is subsumed under this general idea.

As can be seen from the many manuals written for students wishing to study the science of the founding of language, the business of that science, as it developed in later Islamic learning, was to determine the correct placement of every last component of language into the category to which it properly belonged. By doing so it was able to provide a complete and satisfactory account of how language – every jot and title of it – came to be invented. Who invented language was not of crucial importance. That had to remain something of a mystery. What *was* of crucial importance was the question of how sense could be made of its invention.

The principal commentaries on al-ʾĪjī's *ar-Risāla al-waḍᶜiyya* are those of ᶜAlī ibn Muḥammad al-Jurjānī (d. 816/1413), ᶜAlāʾ ad-Dīn ibn Muḥammad al-Qušjī (d. 879/1474), ʾAbū l-Qāsim as-Samarqandī (fl. 888/1483), and ᶜIṣām ad-Dīn al-ʾIsfarāʾinī (d. 944/1573). Manuscripts of these works may be found in the Princeton manuscript collection (Mach 1977). Manuals and summaries are listed in the *Fihris al-kutub al-mawjūda bi-l-Maktaba al-ʾAzhariyya* (*Fihris* 1946–1952), vol. IV.

BIBLIOGRAPHICAL REFERENCES

PRIMARY SOURCES

Ibn Fāris, *Ṣāḥibī* = ʾAbū l-Ḥusayn ʾAḥmad Ibn Fāris, *aṣ-Ṣāḥibī fī fiqh al-luġa*. Ed. Moustafa Chouémi [Muṣṭafā Šuʾaymī]. Beirut: Badran, 1964.
Ibn Jinnī, *Xaṣāʾiṣ* = ʾAbū l-Fatḥ ᶜUṯmān Ibn Jinnī, *al-Xaṣāʾiṣ*. Ed. Muḥammad ᶜAlī an-Najjār. 3 vols. Cairo, 1952–1956. (Repr., Beirut: Dār al-Hudā, n.d.)
ʾĪjī, *Risāla* = ᶜAḍud ad-Dīn ᶜAbd ar-Raḥmān ibn ʾAḥmad al-ʾĪjī, *ar-Risāla al-waḍᶜiyya*. Ed. in *Majmūᶜ al-mutūn*, 33–35. Cairo: al-Maṭbaᶜa al-Jāmiᶜiyya, n.d.

SECONDARY SOURCES

Ess, Josef van. 1966. *Die Erkenntnislehre des ᶜAḍudaddīn al-Īcī*. Wiesbaden: F. Steiner.
Fihris al-kutub al-mawjūda bi-l-Maktaba al-ʾAzhariyya. 1946–1952. Cairo: Maṭbaᶜat al-ʾAzhar.
Loucel, Henri. 1963–1964. "L'origine du langage d'après les grammairiens arabes". *Arabica* 10.188–208, 253–281; 11.57–72, 151–187.

Mach, Rudolph. 1977. *Catalog of Arabic manuscripts (Yahuda section) in the Garret Collection, Princeton University Library*. Princeton, N.J.: Princeton University Press.

Makdisi, George. 1984. “The juridical theology of Shâfiʿî: Origins and significance of the *uṣûl al-fiqh*”. *Studia Islamica* 59.5–47.

Méhiri, Abdelkader. 1973. *Les théories grammaticales d'Ibn Jinnî*. Tunis: Université de Tunis.

Versteegh, Kees. 1997. *The Arabic linguistic tradition*. London and New York: Routledge.

Weiss, Bernard. 1966. *Language in orthodox Muslim thought: A study of* waḍʿ al-lughah *and its development*. Ph.D. diss., Princeton University.

——. 1974. “The medieval discussions of the origin of language”. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 125.33–41.

——. 1976. “A theory of the parts of speech in Arabic: A study of *ʿilm al-waḍʿ*”. *Arabica* 32.23–36.

——. 1984. “Language and tradition in medieval Islam: The question of *al-Ṭarīq ilā maʿrifat al-lugha*”. *Der Islam* 61.91–99.

——. 1987. “*ʿIlm al-waḍʿ*: An introductory account of a later Muslim philological science”. *Arabica* 34.339–356.

BERNARD WEISS (University of Utah)

Wādī Ḥaḍramawt Arabic

1. GENERAL

Ḥaḍramī Arabic is an Arabic dialect spoken by the people living in Ḥaḍramawt, now a governorate in the Republic of Yemen. It is also spoken by many Yemeni emigrants, who migrated from Ḥaḍramawt to the Gulf States, particularly Saudi Arabia, and to East Africa (Kenya, Somalia, and Tanzania) and Southeast Asia (Indonesia, Malaysia, and Singapore).

This entry deals mainly with Ḥaḍramī Arabic in the Wādī region, but occasional reference is also made to Ḥaḍramī Arabic in the coastal region, especially when discussing aspects of phonotactics, as this constitutes a major phonological difference between the two varieties.

2. LINGUISTIC DESCRIPTION

2.1 *Phonology*

2.1.1 Consonants

The consonants of Ḥaḍramī Arabic are listed in Table 1, as pronounced in an intervocalic environment of a front open unrounded vocoid for plain consonants, and a back unrounded vocoid for emphatic or emphatic-like consonants, e.g. [ata] and [ɑtɑ] for /t/ and /ṭ/, respectively.

The dialect is also characterized by its **g*-yodization, i.e. changing **j* to /y/ [j]. In educated speech, **j* is realized as a voiced palatal plosive or affricate in some lexical items, which are marked [+religious] or [+educated] (see **q* below). In the Wādī Ḥaḍramī Arabic, where [ɟ] exists, [ʤ] does not, and either of them alternates with /y/, thus [ɟ]*umʿah* ~ [ʤ]*umʿah* ~ [j]*umʿah* for ‘Friday’.

A distinction is made between /t/ – /ṯ/ and /d/ – ḏ/, but /ḍ/ and /ḏ̣/ are both pronounced as /ḏ̣/ [ð̴]. Coastal Ḥaḍramī Arabic merged all these pairs into the stops /t/, /d/, and /ḍ/ [d̴], respectively.

The **q* reflex is pronounced as a voiced velar /g/ in all lexical items throughout the dialect.

2.1.2 Vowels

2.1.2.1 Monophthongs

There are eight monophthongs, three of which are short and the remaining five long. Phonologically, their front quality in a nonemphatic environment is striking for native speakers of Eastern Arabian dialects in words with an *ā* [æː], as in *zamān* ‘time’, *klāb* ‘dogs’, *blād* ‘country’, which are pronounced in some Eastern Arabian dialects with a centralized or back quality ranging from [aː] to [ɑː].

The vowels and their allophones have the following IPA values:

i	[ɪ] as in *binit* ‘girl’, [ɨ] as in *ṭibb* ‘medicine’, [e] as in *biʿ* ‘sell!’
a	[æ̆] as in *fann* ‘art’, [ă] as in *kabb* ‘he poured’, [ɑ̆] as in *gaḏ̣ḏ̣* ‘it made a click’
u	[ə] as in *baḥur* ‘sea’, [o] as in *xubiz*, [ʊ] as in *bunn* ‘coffee beans’, [ʉ] as in *muyy* ‘sip of water’
ī	[iː] as in *sīf* ‘seashore’, [iːə] as in *bīḏ̣* ‘white [pl. masc.]’
ē	[eː] as in *sēf* ‘sword’, [ɛːə] as in *bēḏ̣* ‘eggs’
ā	[æː] as in *bāb* ‘door’ , [ɑː] as *dār* ‘house’
ō	[oː] as in *yōm* ‘day’, [ɔː] as in *ṣōṭ* ‘voice; sound’, [ɔːə] as in *dōr* ‘garden wall’
ū	[uː] as in *ʿūd* ‘lute; twig’

2.1.2.2 Diphthongs

Diphthongs are combinations of any monophthong with /y/ or /w/. They occur in word-final positions, as in *ḥayy* ‘alive, live; quarter’. In the /j~y/ style or sociolinguistic alternation, diphthongs with a /y/-element also occur in word-medial positions, e.g. *majnūn* ~ *maynūn*. The diphthongs and their allophones have the following IPA values:

Table 1. Consonants of Ḥaḍramī Arabic

	labial	labiod.	dental	alveo.	postalv.	palatal	velar	phar.	glot.
plosive	b			t, d		j([ɟ])*	k, g		
emphatic				ṭ, (ḍ)**					
affricate					j([ʤ])*				
nasal	m			n					
fricative		f	ṯ, ḏ	s, z	š		x, ġ	ḥ, ʿ	ʾ
emphatic			(ḏ̣)**	ṣ					
trill				r					
lateral				l					
emphatic				ḷ					
approximant	w					y			

* [ɟ] / [ʤ] alternate (i.e. either [ɟ] or [ʤ] represents the *jīm* reflex)
** *ḏ̣/ḍ* merge (i.e., either *ḏ̣* or *ḍ* represents the merged reflexes of *ḏ̣āʿ* and *ḍād*)

ay	[ǣɪ] as in *ḥayy* ‘alive’, [ʌ̆ɪ] as in *ṭayy* ‘folding’
aw	[ăʊ] as in *jaw* ‘air’
uy	[ɵ̆ɪ] as in *muyy* ‘sip of water’
ūy	[uːɪ] as in *xūy* ‘my brother’
ēw	[eːʊ] as in *grēw* ‘Qurayw [place name]’

2.2 *Phonotactics*

Ḥaḍramī Arabic can be divided into two main categories in terms of consonant clustering. In initial positions, there is a difference between the Wādī and the coastal varieties. Coastal Ḥaḍramī elides all short vowels in open prestress syllables and thus displays initial clusters, as in *bṣal* ‘onions’, *brīd* ‘mail [noun]’, and *kda* [kdeʔ] ‘so’, whereas Wādī Ḥaḍramī preserves /a/ in this case, as in *baṣal*, *barīd*, and *kaḏa*.

Coastal and Wādī Ḥaḍramī do not have final clusters. Thus, *bint* ‘girl’ and *sabt* ‘Saturday’ are realized in both varieties as *binit* and *sabit* respectively. Before pause, the 1st person singular and 2nd person singular masculine of the perfect show free variation, as in *ruḥtih* ~ *ruḥit* ‘I went; you went’, *mtāh ruḥtih* ~ *ruḥit* ‘when did you [sg. masc.] leave?’. The citation form is *ruḥit*. In context before a vowel, the inserted vowel is deleted: *ruḥt issūg* ‘I went to the market’.

When liaisoned to another word, final geminates are realized as single consonants, e.g. *šilli* ‘take! [sg. fem.]’, but *šal abūh* ‘he took his father’; *haddana* ‘he hit me’, but *had ʾamīn* ‘he beat Amin’. Final geminates can also change phonetic features due to regressive assimilation, e.g. *ʿidd ṣābiʿak* > *ʿiṣ ṣābiʿak* ‘count [sg. masc.] your fingers!’; *rudd ṯōbi* > *rùṯ ṯōbi* ‘give [sg. masc.] back my dress!’.

2.3 *Simplification*

Some words undergo simplification processes conditioned by certain morphological forms, for instance when linked to certain pronouns. Thus, /f/ in *šuf* ‘look!’ is elided in *šu naḥna jīna* ‘look, we have arrived!’ (instead of **šūfu* or **šfu*); *šīš farḥānih* ‘look, you [sg. fem.] are happy!’ (instead of **šūfīš*).

Other grammatical words with segment elision include *la* < *ila* < **ʾiḏā* ‘when [conj.]’; *ba* < *bġa* ‘want’ as in *ma ba* ‘I don’t want’; *baʿdēn* [bʕə̃] *xaraj* ‘then, he left’; *kuḏ* [kɵɪə] *ya (ja) walla ʿāduh* ‘has he come yet?’.

Lexical items may occur as ‘strong’ or ‘weak’ forms, retaining the same meaning. The weak form occurs as a result of syllable reduction in certain rhythmic patterns or tempi. Simplification here is contextualized, i.e., it occurs because of such contexts. Thus, in the following examples, pretonic syllables are reduced as the speaker pronounces them in anticipation of the tonic syllable which is the main focus of the utterance: *sabab* > [saβab] > [saːβ] > [səβ-ː] ‘reason’, e.g. *w ssabab innuh māhu dā̀rī bši* ‘…and the reason is that he doesn’t know’. Simplification takes place according to a scale of different tempi ranging from low to high speed. Another example is *w baʿdēn* [wəbaʕdɛːn] > [wəβəʕə̃n] > [ʷβ̞əʕə̃] > [ʷəʕə̃] ‘and then’.

2.4 *Morphology*

As in most Yemeni dialects, there is gender distinction in the 3rd and 2nd persons of pronouns and verbs, as well as in the 1st person singular of independent and direct object pronouns.

2.4.1 Pronouns

2.4.1.1 Independent pronouns

3rd sg. masc.	*hu*	2nd sg. masc.	*intih*	1st sg. masc.	*ana*
3nd sg. fem.	*hi*	2nd sg. fem.	*inti*	1st sg. fem.	*ani*
3rd pl. masc.	*hum*	2nd pl. masc.	*intu*	1st pl. com.	*naḥna*
3rd pl. fem.	*hin*	2nd pl. fem.	*intēn*		

2.4.1.2 Direct object pronouns

	postconsonantal	*ḍarab-* 'he hit'	postvocalic	*ja-* 'he came'
3rd sg. masc.	*-uh*	*ḍarabuh*	*-h*	*jāh*
3rd sg. fem.	*-ha*	*ḍarabha*	*-ha*	*jāha*
3rd pl. masc.	*-hum*	*ḍarabhum*	*-hum*	*jāhum*
3rd pl. fem.	*-hin*	*ḍarabhin*	*-hin*	*jāhin*
2nd sg. masc.	*-ak*	*ḍarabak*	*-k*	*jāk*
2nd sg. fem.	*-iš*	*ḍarabiš*	*-š*	*jāš*
2nd pl. masc.	*-kum*	*ḍarabkum*	*-kum*	*jākum*
2nd pl. fem.	*-kin*	*ḍarabkin*	*-kin*	*jākin*
1st sg. masc.	*-na*	*ḍarabna*	*-na*	*jāna*
1st sg. fem.	*-ni*	*ḍarabni*	*-ni*	*jāni*
1st pl. com.	*-naḥna*	*ḍarabnaḥna*	*-naḥna*	*jānaḥna*

2.4.1.3 Possessive pronouns

	postconsonantal	*kitāb* 'book'	postvocalic	*ʿaṣa* 'stick'
3rd sg. masc.	*-i*	*ktābi*	*-y*	*ʿaṣāy*
3rd sg. fem.	*-ha*	*ktābha*	*-ha*	*ʿaṣāha*
3rd pl. masc.	*-hum*	*ktābhum*	*-hum*	*ʿaṣāhum*
3rd pl. fem.	*-hin*	*ktābhin*	*-hin*	*ʿaṣāhin*
2nd sg. masc.	*-ak*	*ktābak*	*-k*	*ʿaṣāk*
2nd sg. fem.	*-iš*	*ktābiš*	*-š/-iš*	*ʿaṣāš*
3rd sg. masc.	*-uh*	*ktābuh*	*-h/-uh*	*ʿaṣāh*
1st sg. com.	*-i*	*ktābi*	*-y/-ih/-i*	*ʿaṣāy* ~ *ʿaṣāyih*
1st pl. com.	*-na*	*ktābna*	*-na*	*ʿaṣāna*

2.4.1.4 Demonstrative pronouns

The basic form from which demonstrative pronouns are derived is *ḏa* (masc.), *ḏi* (fem.), *ḏēl(a)* ~ *ḏōl(a)* (pl.) 'this'. With affixes like *hā-*, *hō*, *hū*, *-k*, *ka-*, these forms refer to distance (*ḏāk/ḏīk/ḏēlāk; hāḏa ~ hūḏa/hāḏi ~ hīḏi/hāḏēla ~ hāḏōla; hāḏāk/hāḏīk/hāḏēlāk*), to direction (*kaḏa, kaḏāk*), or to manner (*hōkaḏa*), when used as deictic terms or, with varying degrees of modification, to satisfy the rhythmic patterns in conveying (focal) emphasis in discourse.

Demonstratives can be annexed to a phrase to convey manner, reason, or spatial and temporal senses: *hākaḏa/hōkaḏa* 'thus, this way; like this'; *hūḏa* 'this; therefore, for this reason'.

2.4.1.5 Relative pronoun

The relative pronoun is *li-* or, less commonly, *illi: fyēn ilktāb li garētuh?* 'where is the book that you read?'.

2.4.1.6 Interrogative pronouns and adverbs

The basic interrogative pronouns are the following:

āh	'what?'	*min*	'who?'
yāt	'which one?	*lēh*	'why?'
fēn ~ *fyēn*	'where?'	*mtāh*	'when?'
kamm	'how many [fem.]?'	*kammih*	'how many [masc.]?'
kēh	'how?'		

There are other forms of these pronouns with affixes *hu-* or *kā-*: *hwāh* (*hu* + *āh*), e.g. *hwāh hūḏa* 'what is this?'; *hūmin*, e.g. *hūmin hūḏa* 'who is this?'; *kākēh*, e.g. *kākēh naggi (nalgi) buh* 'how can we deal with it?'

The pronouns *ēš* 'what?' and *lēš* 'why?', which are common in other Yemeni dialects such as ʿAdanī and Ṣanʿānī, are now heard quite commonly in the speech of Ḥaḍramī Arabic speakers, especially when in contact with such dialects.

2.4.2 Adverbs

The most common adverbs of place, time, and manner are:

Place			
hna	'here'	*hnāk*	'there'
kaḏa	'this direction'	*kaḏāk*	'that direction'
Time			
ḏalḥīn ~ *ḏaḥḥīn*	'now'	*ilyōm*	'today'
ġudwah	'tomorrow'	*ilgāblih*	'tomorrow afternoon/ evening'
šāʿatha	'the evening after tomorrow'	*amis*	'yesterday'
ilbāriḥ	'yesterday night'	*bāriḥtlūla*	'the day before yesterday'
Manner			
rayyiḍ	'good, okay'	*sawa*	'okay'
jamm	'very much'	*b-surʿah*	'quickly'

Such forms can be intensified or exaggerated by adding the suffixation *-īnih, jamm* 'very, much' > *jammīnih* 'great many, plenty of'.

2.4.3 Particles

2.4.3.1 Definite article

The definite article is *il-*, with assimilation of *l-* to the preceding consonant if this consonant is [+coronal], i.e. dental, alveolar, or alveo-palatal (except *j* and *y*), e.g. with nonassimilated *l*, *ilmanārah* 'the minaret', and with assimilated *l ittūm* 'the garlic'. The vowel *i* in the definite article *il-* is elided when the word begins with a vowel, e.g. *āṯār* 'archaeological sites' > *lāṯār* 'the archaeological sites'.

2.4.3.2 Prepositions

The following are the main prepositions in Ḥaḍramī Arabic; most of them have the same meaning as in Standard Arabic.

b	'with'	*fi*	'in'
ʿala	'on'	*min*	'from'
bēn	'between'	*janib*	'near, beside'
giddām	'in front of'	*taḥit*	'under'
l	'to; for'	*wara*	'behind'
kama	'like'	*ʿind*	'at, by, near'
mʿa	'by, with'		

2.4.3.3 Conjunctions

The following are the main conjunctions:

ila ~ *la*	'if; when; up to'	*walla*	'or'
bass	'but'	*ya… walla…*	'either… or…'
lākann ~ *kann*	'but'	*w*	'and'
baʿdēn	'then; next'	*lākuḏ* ~ *lakuḏ*	'when'
ištgūl ~ *štgūl* ~ *tgūl*	'as if'	*lmād (lmēd)*, *layil (lajil)*	'in order to; for'
yōm	'because; since'	*kama*	'also'
lammān	'up to'		

2.4.4 Nominal morphology

2.4.4.1 Gender

With few exceptions, feminine gender has an *-ah* or *-ih* ending: *dīrih* 'household; family', *maktabih* 'library, bookshop', *maḥḏarah* 'room', *manārah* 'minaret'.

2.4.4.2 Derivation

Derivational patterns are mainly from verbs and, in a few cases, from nouns. Forms derived from verbs are, for example: $C_1aC_2iC_3$/$C_1aC_2uC_3$ *kasir* 'breaking' < *kasar* 'he broke', *sarib* 'insertion' < *sarab* 'he inserted '; $C_1aC_2C_3$ with doubled final radical *gazz* '(feeling) disgust for something' < *gazz* 'he felt disgust for something', *lazz* 'joining [two things together firmly]' < *lazz* 'he joined'.

Forms derived from other nouns include *manārah* ‘minaret’ < *nūr* ‘light’, *madajjih* ‘hens’ cage’ < *djāj* ‘hens, chickens’.

Simple nonderivational substantives include loanwords such as *naʿʿ* ‘sweet, bonbon [a children’s word]’, *šāh* ‘goat’, *mōtar* ‘car’.

2.4.4.3 Substantive patterns

Common substantive patterns include:

Cv̄C	*dār* ‘house’, *jār* ‘neighbor’, *nōm* ‘sleep’
Cv̄CC	*ḥāll* ‘tenant’
CvCC	*yadd* ‘hand’, *kaff* ‘palm of the hand’, *ḥadd* ‘border; edge’
CCvC	*ḥmah* ‘stamina’, *ʿrus* ‘marriage’, *ʿnib* ‘grape’
CCv̄C	*brāṣ* ‘(brass) bearing’, *blāk* ‘device that makes ignition spark in an engine’
CvCvC	*jabal* ‘mountain’
Cv̄CvC	*ḥāris* ‘guard’, *sāyig* ‘driver’, *zāriʿ* ‘farmer’
CvCv̄C	*ṭabīb* ‘physician’, *ṣadīg* ‘friend’
Cv̄Cv̄C	*fārān* ‘butcher’s knife; chopper’, *ṣārūn* ‘sarong’
CvCC*a*	*fuxṭah* ‘fertilization of palm trees’, *rabših* ‘chaos, turmoil’
$C_1vC_2C_1vC_2$	*jiljil* ‘sesame’, *ḏubḏub* ‘tiny premature green dates (buds)’
$CvC_2C_2\bar{v}C$	*sakkīn* ‘knife’, *xabbāz* ‘baker’, *sabbāḥ* ‘swimmer’, *dabbāġ* ‘tanner’
$\mathit{m}vC_1C_2\bar{v}C_3$	*miswāk* ‘cleaning stick for teeth’, *mizḥāh* ‘ax-shaped spade’
$C_1vC_2C_3\bar{v}C_4$	*mandīl* ‘handkerchief’, *ziʿṭār* ‘a variety of lizard’
$C_1vC_2C_3vC_4\bar{v}C_5$	*ʿankabūt* ‘spider’, *zaʿfarān* ‘saffron’, *zinjabīl* ‘ginger’

2.4.4.4 Compounds

There are a number of substantives made of *bā* ‘father (of)’ and another substantive, e.g. *bā jʿālih* ‘dung beetle’ (< *jʿālih* ‘wage’), *bā farārah* ‘flying beetle that infects palm trees’ (< *f-r-r* ‘to fly’), *bā ṣwāb* ‘drizzles, tiny drops of rain’ (< *ṣwāb* ‘lice egg’), *bā magdaḥ* ‘[ladle-shaped] young frog, tadpole’ (< *magdaḥ* ‘ladle’), *bā ʿōdih/ʿawdah* ‘a ‘Yemeni immigrant visiting his home for some time and going back abroad’ (< *ʿawdah* ‘reentry visa of the host country’).

2.4.4.5 Adjectives

The most common patterns of adjectives are the following; their feminine is formed by the suffix *-ih* or *-ah*.

CvCv̄C	*jadīd, jadīdih* ‘new’
Cv̄CvC	*yābis, yābsih* ‘dry’
CvCCv̄C	*bahlūl, bahlūlih* ‘naive, stupid’
CvCCv̄C	*kaḏḏāb, kaḏḏābih* ‘liar’
$C_1vC_2C_3$*ān*	*sakrān, sakrānih* ‘drunk’

For color adjectives, stress is on the last syllable in the masculine but not in the feminine: vCCvC, CvCCv (fem.) *axḍár, xáḍra* ‘green’.

The following adjectives are derived from Form II or from quadriliterals: *m*CvCCvC *mnignig, mnignigah ~ mnignigih* ‘nagger, grumbler’.

Relative adjectives are formed by affixing *-i* to proper names, e.g. *ʿadni* [masc.], *ʿadniyih* [fem.], *ʿadāniyih* [pl.] ‘native of Aden’ < *ʿadan* ‘Aden’.

2.4.4.6 Elatives

The elative is formed with the pattern *a*CC*a*C and is followed by *min* as a comparative, e.g. *aḥmad akbar min sālim* ‘Ahmad is bigger than Salim’. The superlative is followed by the noun without *min*, e.g. *waladak aḥsan wāḥad f iṣṣaff* ‘your [masc.] son is the best in the class’. The exclamatory pattern *ma-ʾafʿal* is represented in *maḥsan*, e.g. *maḥsanuh min makān* ‘what a beautiful place!’.

2.4.4.7 Number

The sound plural is made by adding *-īn* to the masculine and *-āt* to the feminine nouns or adjectives, e.g. *fallāḥ* [sg. masc.], *fallāḥīn* [pl. masc.], *fallāḥih* [sg. fem.], *fallāḥāt* [pl. fem.] ‘farmer’.

Relative adjectives ending in *-i* for the singular masculine and *-iyih* for the feminine take the sound plural *-yīn* for the masculine and *-yāt* for the feminine, e.g. *maṣri* [sg. masc.], *maṣriyīn* [pl. masc.], *maṣriyih* [sg. fem.], *maṣriyāt* [pl. fem.] ‘Egyptian’. Some relative adjectives have a different form in the masculine plural, e.g. *badwi*, plural masculine *badu*, plural feminine *badwiyāt* ‘Bedouin’.

2.4.4.8 Broken plural

Some of the most common broken plural patterns are the following:

	plural	singular
CCvC	*sgal* ‘children’	*suglih* (masc./ fem.)
CCv̄C	*xlāf* ‘windows’	*xalfih*
CCv̄CvC	*ṭbābih* ‘doctors, physicians’	*ṭabīb*
CvCv̄C	*sawēḥ* ‘squares [in towns]’	*sāḥib*
Cv̄Cv̄C	*kīzān* ‘jugs’	*kūz*
CvCv̄Cv	*ʿaḏāra* ‘adolescent girls’	*ʿaḏra*
CvCCv̄C	*ḥillān* ‘dwellers’	*ḥāll* masc./ *ḥāllih* fem.
CvCv̄Cv̄C	*ḥašākīl* ‘miscellaneous small articles’	*ḥaškūl*
CCv̄Cv̄C	*bhālīl* ‘stupid [pl.]’	*bahlūl*
CvCC*a*C	*gaʿṭah* ‘Quʿayṭīs’	*gʿēṭi* (masc.)
CvCCv̄CvC	*baggārah* ‘plowmen’	*baggār*
Cv̄Cv̄C	*ḏ̣āfīr* ‘fingernails’	*ḏ̣āfūr*

2.4.5 Cardinal numbers

masc.	fem.	
wāḥad	*waḥdih*	‘one’
iṯnēn	*ṯintēn*	‘two’
ṯalāṯih	*ṯalāṯ*	‘three’
arbʿah	*arbaʿ*	‘four’
xamsih	*xamis*	‘five’
sittih	*sitt*	‘six’
sabʿih	*sabiʿ*	‘seven’
ṯamāniyh	*ṯamān*	‘eight’
tisʿih	*tisiʿ*	‘nine’
ʿašrah	*ʿašir*	‘ten’

Examples: *ṯalāṯih rjāl* ‘three men’, *ṯalāṯ ḥarīm* ‘three women’, *xamsih dyār* ‘five houses’.

Numerals from 11 to 19 do not have gender distinction. They can be used as cardinal or ordinal (see below), as in *ja tartībuh f iṣṣaff ḥadaʿšar* ‘he became number 11 in the class’, *ja tartībuh f iṣṣaff ilḥadaʿšar* ‘he is the 11th in the class’.

2.4.6 Verbal morphology

The verbs show the most elaborate morphology, marking number, gender, person, tense, and aspect, and transitivity type. Verbs denote two aspects, perfective (nonprogressive) and imperfective (progressive).

2.4.6.1 Form I

Form I displays a transitive *a*-type *katab/yúktub* ‘to write’ (Table 2), and an intransitive *u*-type *kbur/yúkbur* ‘to grow’ (Table 3).

2.4.6.2 The derived Forms II – X

The derived Forms, with the exception of V and VI, display a morphological distribution of /a/, /i/ in the ultima with /a/ in the perfect and /i/ in the imperfect. Forms II, VII, VIII, and X show /i/ in the penultimate of the imperfect as well (Table 4).

Quadriliteral roots follow Form II and Form V as to vowels: *bahḏal/ybihḏil* ‘to deform something’, active participle *mbihḏil*, passive participle *mbahḏal; tgaḥzal/yitgaḥzal* ‘to show off’, participle *mitgiḥzil*.

2.4.6.3 Intensive verbs

A particularity of Ḥaḍramī Arabic are the intensive verbs, which are formed by adding the infix |aw| -*ō*- to Form I, e.g. *katab* ‘to write’, intensive |kawtab| *kōtab* ‘to write a lot’. These give the sense of intensity or exaggeration. Consider the following examples (ʿAqīl 1981:139): *katab/yiktub* ‘to write’, *kōtab/ykōtib* ‘to write a lot’; *ṣalag/yiṣlug* ‘to shout’, *ṣōlag/yṣōlig* ‘to shout repeatedly’; *ṣabaṭ/yiṣbuṭ* ‘to hit, to beat’, *ṣōbaṭ/yṣōbiṭ* ‘to beat repeatedly’.

Another form of intensive can be made by metathesis of /w/ (from |aw|) with the following consonant: |kawsar| > |kaswar|. This may lead to some lexical expansion of meaning, e.g. *kōsar fi illiʿib* ‘he played rough’, but *kaswar/ykiswir min iḏ̣ḏ̣aḥkāt* ‘he made a series of giggles or laughs’.

2.4.6.4 Weak verbs

2.4.6.4.1 Geminate verbs

l-m-m ‘to keep, protect, gather’

Person	Perfect	Imperfect
3rd sg. masc.	*lamm*	*ylumm*
3rd sg. fem.	*lammat*	*tlumm*
3rd pl. masc.	*lammaw*	*ylummūn*
3rd pl. fem.	*lammēn*	*ylummēn*
2nd sg. masc.	*lammēt*	*tlumm*
2nd sg. fem.	*lammēti*	*tlummēn*
2nd pl.masc.	*lammētu*	*tlummūn*
2nd pl. fem.	*lammētḗn*	*tlummḗn*
1st sg. com.	*lammēt*	*lumm*
1st pl. com.	*lammēna*	*nlumm*

Table 2. Conjugation of *k-t-b* ‘to write’

	Perfect	Imperfect	Imperative
3rd sg. masc.	*katab*	*yúktub*	–
3rd sg. fem.	*katbat*	*túktub*	–
3rd pl. masc.	*katbaw*	*yuktubū́n*	–
3rd pl. fem.	*katabin*	*yuktúbin*	–
2nd sg. masc.	*katabit* ~ *-tih*	*túktub*	*ktub*
2nd sg. fem.	*katabti*	*tuktubī́n*	*kutbí*
2nd pl. masc.	*katabtu*	*tuktubū́n*	*kutbú*
2nd pl. fem.	*katabtḗn*	*tuktúbin*	*ktúbin**
1st sg. com.	*katabit* ~ *-tih*	*úktub*	–
1st pl. com.	*katabna*	*núktub*	–

**kutbán* in al-Qaṭn and some other Bedouinized accents.

Table 3. Conjugation of *k-b-r* ‘to grow’

	Perfect	Imperfect	Imperative
3rd sg. masc.	*kbur*	*yúkbur*	–
3rd sg. fem.	*kúbrit*	*túkbur*	–
3rd pl. masc.	*kúbru*	*yukburū́n*	–
3rd pl. fem.	*kbúrin*	*yukbúrin*	–
2nd sg. masc.	*kbúrit*	*túkbur*	*kbur*
2nd sg. fem.	*kbúrti*	*tukburī́n*	*kubrí*
2nd pl. masc.	*kbúrtu*	*tukburū́n*	*kubrú*
2nd pl. fem.	*kburtḗn*	*tukbúrin*	*kbúrin*
1st sg. com.	*kbúrit*	*úkbur*	–
1st pl. com.	*kbúrna*	*núkbur*	–

Table 4. Derived Forms

	Perfect	Imperfect	Imperative	Participles	
II	*šarrab*	*yiširrib*	*širrib*	*mširrib* *mšarrab*	‘to shave’
III	*ʿālag*	*yʿālig*	*ʿālig*	*mʿālig*	‘to reproach’
V	*txabbar*	*yitxabbar*	*txabbar*	*mitxibbir*	‘to ask about someone’
VI	*tʿālag*	*yitʿālag*	*tʿālag*	*mitʿālig*	‘to have a dispute with someone’
VII	*ngabaṣ*	*yingibiṣ*	*ngubuṣ*	*mingbūṣ*	‘to be stung’
VIII	*(ə)štabaḥ*	*yištibiḥ*	*(ə)štbiḥ*	*mištibiḥ*	‘to boast; to stretch’
X	*(ə)staḥmal*	*yistiḥmil*	*(ə)stiḥmil*	*mistiḥmil*	‘to endure’

Passive participle

malmūm ‘kept intact, kept together’

2.4.6.4.2 Verbs Iʾ

ʾ-k-l ‘to eat’

Person	Perfect	Imperfect	Imperative
3rd sg. masc.	*ʾakál~ kall ~ kalá*	*yōkul ~ yōkil*	–
3rd sg. fem.	*kalát*	*tōkil*	–
3rd pl. masc.	*kaláw*	*yōklū́n*	–
3rd pl. fem.	*kálin*	*yōkúlin*	–
2nd sg. masc.	*kalēt(-ih)*	*tōkul ~ tōkil*	*kull*
2nd sg. fem.	*kalēti*	*tōklīn*	*klí* [kʷli] ~ [kli]
2nd pl. masc.	*kalḗtu*	*tōklū́n*	*klū* [kʷlu]
2nd pl. fem.	*kalētḗn*	*tōkúlin*	*klēn* [kleːn]
1st sg. com.	*kalēt (-ih) ~ kálit ~ káltih*	*ʾōkul ~ ʾōkil*	–
1st pl. com.	*kalḗna*	*nōkul*	–

Passive participle

mākūl ‘eaten’

2.4.6.4.3 Verbs I*w*

w-z-n ‘to weigh’

Person	Perfect	Imperfect	Imperative
3rd sg. masc.	*wazán*	*yōzin*	–
3rd sg. fem.	*wáznat*	*tṓzin*	–
3rd pl. masc.	*wáznaw*	*yōznū́n*	–
3rd pl. fem.	*wazánin*	*yōzínin*	–
2nd sg. masc.	*wazánit~ wazántih*	*tṓzin*	*wzín* [ŭ'zɪn]
2nd sg. fem.	*wazánti*	*tōznī́n*	*wizní*
2nd pl. masc.	*wázantu*	*tōznū́n*	*ʾṓznu ~ wiznú*
2nd pl. fem.	*wazantēn*	*tōzínin*	*ʾōzínin*
1st sg. com.	*wazánit ~ wazántih*	*ʾōzin*	–
1st pl. com.	*wazánna*	*nōzin*	–

Passive participle

mōzūn ‘balanced, adjusted’

2.4.6.4.4 Verbs II*w/y*

g-w-m ‘to stand’, *n-w-m* ‘to sleep’, *s-y-b* ‘to leave’

Person	Perfect	Imperfect	Imperative
3rd sg. masc.	*gām, nām, sāb*	*ygūm, ynām, ysīb*	–
3rd sg. fem.	*gāmat, nāmat, sābat*	*tgūm, tnām, tsīb*	–
3rd pl. masc.	*gāmaw, nāmaw, sābaw*	*ygūmūn, ynāmūn, ysībūn*	–
3rd pl. fem.	*gumin, nimin, sibin*	*ygumin, ynimin, ysibin*	–
2nd sg. masc.	*gumit ~ gumtih, nimit ~ nimtih sibit ~ sibtih*	*tgūm, tnām, tsīb*	*gum, nim, sib*

2nd sg. fem.	*gumti, nimti, sibti*	*tgūmīn, tnāmīn, tsībīn*	*gūmi, nāmi, sībi*
2nd pl. masc.	*gūmu, nāmu, sību*	*ygūmūn, ynāmūn, ysībūn*	*gūmu, nāmu, sību*
2nd pl. fem.	*gumin, nimin, sibin*	*ygumin, ynimin, ysibin*	*gumin, nimin, sibin*
1st sg. com.	*gumit ~ gumtih, nimit ~ nimtih sibit ~ sibtih*	*gūm, nām, sīb*	–
1st pl. com.	*gumna, nimna, sibna*	*ngūm, nnām, nsīb* ([n̩guːm] etc.)	–

Passive participle

msayyab ‘left on the loose; unprotected’

2.4.6.4.5 *Verbs* III*y*

r-m-y ‘to throw’, *n-s-y* ‘to forget’

Person	Perfect	Imperfect	Imperative
3rd sg. masc.	*rama, nsi*	*yirmi, yinsa*	–
3rd sg. fem.	*ramat, nsit ~ nisyit*	*tirmi, tinsa*	–
3rd pl. masc.	*ramaw, nsu* [nsu]	*yirmūn, yinsūn*	–
3rd pl. fem.	*ramēn, nsēn*	*yirmēn, yinsēn*	–
2nd sg. masc.	*ramēt ~ ramētih, nsīt ~ nsītih*	*tirmi, tinsa*	*ʾirim ~ rmi, ʾinis ~ nsa*
2nd sg. fem.	*ramēti, nsīti*	*tirmīn*	*rmi*
2nd pl. masc.	*ramētu, nsītu*	*tirmūn, tinsūn*	*rmu, nsu*
2nd pl. fem.	*ramētēn, nsītēn*	*tirmēn, tinsēn*	*rmēn, nsēn*
1st sg. com.	*ramēt ~ ramētih*	*ʾirmi*	–
1st pl. com.	*ramēna, nsīna*	*nirmi, ninsa*	–

Passive participle

mansi ‘forgotten’

2.5 *Syntax*

The syntax of Ḥaḍramī Arabic, like the phonology and morphology, is not as diverse as that of other Peninsular Arabic dialects.

2.5.1 Sentence types

2.5.1.1 Nominal sentences

The subject can be a noun or a pronoun, and the predicate can be a noun, an adjective, an adverb of place, or a prepositional phrase. Examples: [adjectival predicate] *iljābiyih malānih* ‘the swimming pool is full [of water]’; [noun predicate] *ʿali mdirris* ‘Ali is a teacher’; [adverbial predicate of place] *ilklān dāxal* ‘the bridegroom is inside’; [prepositional predicate] *ilʿummāl min ilgarin* ‘the workers are from al-Qarn’.

Nominal sentences with another sentence as the predicate are manifested by a silent stress separating the subject and the predicate, as in *ḥāmid, ʾabūh ʿammi* ‘Hamid’s father is my paternal uncle’.

When the predicate is a question word, it may precede the subject: *kam gīmituh?* ‘how much does it cost?’

2.5.1.2 Pseudoverbal sentences

Some closed-system items like the prepositions *ʿind, maʿ/mʿa*, and *l-* can be suffixed with nouns or pronouns to form verb-like constructions with the general meaning of ‘to have; to own’ (Qafisheh 1977:200 on Gulf Arabic; Qafisheh 1992:231 on Ṣanʿānī Arabic): *mʿi ktub* ‘I have books’ (→ pseudoverb).

2.5.1.3 Other nominal sentences with special particles

Some special or verblike particles occur in both nominal and verbal sentences. These include *kuḏ, ʿād, mʿād,* and *yirjaʿ* (see below, Sec. 2.5.9).

2.5.2 Verbal sentences

The following word orders of the constituents (S)ubject, (V)erb, (O)bject, (C)omplement, and

(A)dverbial occur in the verbal sentence in Ḥaḍramī Arabic.

V	*ba(a) yiḏ̣har* 'he will appear', *rijˁū* 'they [masc.] came back'
VO	*šufnāh* 'we saw him'
	jibtīhin 'you [sg. fem.] brought them [fem.]'
SV	*issgal kubrū* 'the children grew up'
	aḥmad wṣil 'Ahmed arrived'
	ilˁummāl aḏ̣rabaw 'the workers struck'
SVO	*iššēbih kala ˁaṣīd* 'the old man ate porridge'
	wad-dxūi ištara sēkal nār 'my brother's son bought a motorcycle'
VOA	*gabaḏ̣ īduh b-guwwah* 'he seized his hand firmly'
VOO	*ˁaṭanaḥna galīl tamur* 'he gave us some dates'
SVOO	*ˁali ˁaṭa xūh sāˁih* 'Ali gave his brother a watch'
	iššēbih warraṯ bittuh nuṣṣ iḏḏabur 'the old man bequeathed his daughter half of the land'
VA	*rāḥ kaḏāk* 'he went over there, in that direction'
SVOA	*ḥāmid yḥibb bittuh jamm* 'Hamid loves his daughter very much'
VOOA	*ˁaṭāh ilˁaddi kāš* [kɑːʃ] 'he gave him the money in cash'
VOCA	*yaggīna (yalgīna) ṣāḥibuh kulma šāfana* 'he makes me his friend whenever he sees me'
	yšūfūnuh zaˁīmhum ḥīn yuxṭub 'they see in him their leader when he delivers a speech'
SVOCA	*iṭṭamaˁ xallāh insān bala iḥsās* 'greediness made him hard-hearted'
	innūrah raddat ḏi lmaḥḏ̣arah bārdih jamm 'the lime (coating) made this room cooler'

2.5.3 Complex sentences

A nominal clause can be followed by a verbal clause. Both of them can be independent: *āh min ruzz w laḥim, yxalli lwāḥad yōkil ṣābiˁuh* 'what a [nice] meal [of meat and rice] that will make one eat his fingers [with it when eating]'.

A dependent clause can be followed by an independent clause and a dependent clause: *ṣbur ˁalay galīl // w baˁṭīh daris // ma bayinsāh* 'wait [sg. masc.] for me a little and I'll give him a lesson he'll never forget'.

2.5.4 Coordination

The most common coordination particles are *w* 'and', *kann*, *(w-)lākann* 'but, nevertheless, though', *walla* 'or', *ma* (Standard Arabic *ʾammā*) 'as for…' (*kann/lākann* can also be used as question particles at the beginning of a sentence to convert it into a question). Two verbal sentences are conjoined by *wi: sālim yrūḥ libtidāʾi w ṣāliḥ yrūḥ iṯṯanawi* 'Salim goes to primary (school) and Salih goes to secondary (school)'. Two nominals are conjoined with *wi*: *ilmaṭār garīb w ittakāsi wājdih* 'the airport is nearby and the taxis are many'. A verbal conjoin may be followed by a nominal or verbal conjoin introduced by *wi-*; this is the case with *ḥāl* sentences: *ilmsāfrīn wiṣlū w ilˁaša hamiyyih* 'the guests have arrived while the supper is not ready yet'; *kullīn ja w rāḥ w sˁādih ˁa lmirḏ̣āḥ* 'everybody comes and goes while Saˁadah is still on her grinder [grinding date stones]', i.e. 'everybody is standing still, no improvement'; *ilklān gāˁid w innās yizfnūn* 'the bridegroom is sitting while people are dancing'. Adversative *kann: illasˁār tiṭlaˁ, kann ilmrattabāt makānha* 'prices are soaring, but salaries are the same'.

The particle *yān* (**ˁay ˁanna*) introduces explanations: *āl awwal nšāṭ, yān ilwāḥad minhum bayōkul ṭali kāmil* 'the old generation are fit people, one of them can eat a whole lamb'. It can precede the construction for emphasis, as in *(waḷḷah) yānak tikḏib, māḥad smiˁ ši xabar* 'by God's name [i.e. definitely] you are lying. No one has heard any news'.

Some of the auxiliaries are originally independent verbs like *gām* 'to stand', *gˁad* 'to sit', *ṣabbar* 'to start', *gafaz/ṭamar* 'to jump', *rjiˁ* 'to return', *sagaṭ* 'to fall down', etc. They can occur as auxiliaries with an imperfect, e.g. *la gām yaġanni šaˁˁagaw* 'when he sings/starts singing, they tear off their clothes [due to excessive enthusiasm and admiration]'.

2.5.5 Negation

Simple sentences are negated by adding one of the negative particles *ma, māhu, la, lˁād, mˁād* to the affirmative sentence. Such particles, like many others, manifest variant pronunciation when contracted with personal pronouns, e.g. *mā kull ˁūd aglīd* 'not every stick is a (wooden) key'.

Compound negative sentences: *la gatlna wa la ṣalla ˁalay* 'he neither killed me nor prayed for me [i.e. 'he left me in suspense; he didn't give a clear answer']'.

2.5.6 Adverbial clauses

Adverbs as modifiers of verbs: *gṭaˁ samāḥ māhu šeīz* ‘cut [the piece of cloth, paper, etc.] straight, not crooked!’ In the following example, an adverb modifies an adjective: *ma ṣōṭuh* ʌ *ˁeīf lāšag* ‘as for his voice, it is too bad’ (ʌ refers to silent stress or + juncture).

Adverbial clauses with *la* for time: *la xlug, sammeīnāh* ‘when he is born, we will name him’ (a proverb).

Adverbial clause with *štgūl* for comparison: *šuf izzbeīr ja aġbar (š)tgūl ḥadd xadar buh sōm* ‘look at al-Zubair when he came, covered with dust, as if someone had drilled a sand barrier with him [i.e. used him like a digging tool in the sand]’.

Adverbial clause with *lmād* for purpose: *yxilṭūn iṭṭīn izzibir b ittibil, lmād ilmadar yigaˁ gawi* ‘they mix zibir mud with (chopped) straw, so that bricks become firm’. The pan-Arabic conjunction *ˁalašān* is also used, especially by literate people: *ḏākir sawa ˁalašān tinjaḥ* ‘study [sg. masc.] hard in order to succeed!’.

Adverbial clause with *lmād* for cause with *yōm* or *yōmak: šufha ba tikrahak, yōmak tgūl kaḏa* ‘she will hate you for saying such things’, *xarjaw farḥānīn yōm ilwāḥad gabaḏ̣ mšāharatuh* ‘they came out happily, as everybody received their salary’.

Adverbial clauses with *kama* or *miṯilma* involve comparison or equation: *ˁaṭa kull wāḥad minhum kama/miṯilma li ˁaṭāh l ilnafar min ˁyāluh* ‘he gave every one of them the same as what he had given to each of his sons’.

2.5.7 Interrogative sentences

The following interrogative particles are used: *kamm* ‘how many’, with its derivatives free variant *kammih* and *kammīnih* (intensified form) followed by *min*. Other particles are *āh*, *hwāh* ‘what?’, *fyēn* ‘where?’, *kēh, kākēh* ‘how?’, *kann, lēh*, ‘why?’, and *tara* ‘but why?’.

kamm/kammih/kammīnih ‘how many’: *kamm* and *kammih* are used alternatively before a consonant: *kamm iddūlār ḏaḥḥīn?* ‘how much is the dollar these days?’, *kamm ryālāt mˁak?* ‘how many rials do you have?’; *kammih* cannot be used before a vowel: *kammih minnikum* ‘how many of you [pl. masc.]?’, but not **kammih iddūlar ḏaḥḥīn?*.

The question particle *kammih* has the intensive form *kammīnih*, e.g. *kammīnih minnihum* ‘how many of them?’.

The particle *kann* means ‘why?’, e.g. *kann innās ġurmit* ‘why are people becoming mad?’; *āh/hwāh* means ‘what?’, e.g. *āh isim jiddak?* ‘what is the name of your grandfather?’

Sometimes, quasi-interrogative independent clauses are used (to suspend the attention) for emphasis, e.g. *warāh gāl luh āh ísmak? taw ˁaṭāh ilˁaddi* ‘did he ask him what his name was? (No), he gave him the money immediately’.

2.5.8 Passive voice

In Ḥaḍramī Arabic, the apophonic passive is restricted to a few verbs and situations in which it is not important to define the agent, as in *jābōh min ilwādi baˁd ilˁša. iḏ̣ḏ̣āhir innuh gbuṣ* ‘they brought him from the valley after *ˁiša* time. It seems that he was bitten/stung’.

The most common form of passivization in Ḥaḍramī Arabic is mediopassive, which is realized by adding prefixes like *t-*, *i+(t)*, and *in-* to different verbal stems (Forms).

Besides being used to express the active voice (i.e. causative and factitive meanings), Form II can also render mediopassive, as in the following examples: *ilmandīl ballaḥ* ‘[the color of] the handkerchief faded’; *ilmaṭīrih xaḏ̣ḏ̣rat* ‘the patch [of field] became green’.

2.5.9 Particles

2.5.9.1 The particle *kuḏ*

The particle *kuḏ* ~ *guḏ* means ‘yet/already’, denoting completion; although the sentence is nominal, it implies a perfective aspect. It can be added to a pronoun and function as a subject in a nominal (or verbal sentence). *kuḏ* ~ *guḏ* is, together with another pair of variants, *gad* ~ *gid*, the cognate of Standard Arabic *qad*, which in Ḥaḍramī Arabic developed semantically as follows: *kuḏ/guḏ* ‘yet, already, almost, nearly’; *gad/gid* ‘maybe, perhaps’.

2.5.9.2 The particle *ˁād*

The ‘imperfective’ counterpart of *kuḏ* is the particle *ˁād* ‘yet; still; anymore; already’. It is common also in other Yemeni dialects (see ˁAqīl 1981:140 on Ḥaḍramī Arabic; Qafisheh 1992:228, and Watson 1993:53 on Ṣanˁānī Arabic). It indicates the continued presence of something, but it can also indicate the presence of something (i.e. ‘to be, exist’), or that something has just happened (i.e. ‘just’; Watson, 1993:5). Thus, *kuḏ* and *ʿād* are particles used to indicate time in nominal sentences. When *ʿād* is used in verbal sentences with negation, it must be used with *mā*, but, being merged into one particle, the realization is *mʿād* rather than **mā ʿād*: *ʿād galīl ma fi lxazzān* 'still there is some water in the tank'.

2.5.9.3 Preverbal particles

Bā + imperfect denotes a future event or action, e.g. *ilmsāfrīn bā yṣalūn illeīlih* 'the travelers will arrive tonight'.

kōd + imperfect 'hardly': *kōd ḏa lgmāš ykaffi lbanṭlōn* 'this piece of cloth would hardly be enough for (making) a pair of trousers'; *habbin* is the same as *kōd* except that it is commonly used with a pronominal suffix.

2.5.9.4 The verb *yirjaʿ*

Although the tense of the verb can change from {yafʿal} to {faʿal}, the other elements associated with the verb and other contextual elements determine the aspect. In the following pair of sentences, although the tense of the verb *yirjaʿ* changes, the aspect remains the same:

irrajjāl li ḥaṣṣalnāh yirjaʿ axu ʿali 'the man whom we found is Ali's brother'
irrajjāl li ḥaṣṣalnāh rjiʿ axu ʿali 'the man whom we found turned out to be Ali's brother'.

2.5.9.5 Conditionals

Conditional particles in Ḥaḍramī Arabic are *iḏa, (i)la, inn*, and *la* (< *law*), e.g. *inn batismaʿ kalami, bathaṣṣil li bġeītuh* 'if you obey me, you will get what you want [sg. masc.]'. Conditionals can be real or unreal. An example of a real conditional is as follows: *inn ba taʿṭīna ʿadiyyi mʿād ba tšūfana* 'if you give me my money, you will not see me (again)'. The conjunctions *iḏa/(i)la* introduce a real condition: *šariṭ iḏa ~ la xallaṣtuh baʿṭīk šagāk kāmil* 'I will give your full wage on condition that you finish it'; or an unreal condition: *iḏa ~ la šargat iššamiš sarḥaw* 'when the sun rises, they go [to work]'.

2.5.9.6 Other auxiliaries in verbal sentences

Some of the auxiliaries are originally independent verbs like *gām* 'to stand', *gʿad* 'to sit', *sabbar* 'to start', *gafaz/ṭamar* 'to jump', *rjiʿ* 'to return', *sagaṭ* 'to fall down', etc. They can occur as auxiliaries with the imperfect, e.g. *gʿad yaḍḥak* 'he went on laughing'.

3. LEXICON

Native speakers of Ḥaḍramī Arabic see their language as being very near to Classical Arabic. This view may reflect the fact that they find many Classical Arabic words in their dialect which do not exist in the vernaculars of other Arabs. In fact, apart from recent innovations in which Ḥaḍramī Arabic borrowed some words from other Arabic dialects and foreign languages, Ḥaḍramawt (and indeed Yemen in general) has preserved its language over the centuries due to its geographical isolation. There is an Old South Arabian element in the dialect (Bâfaqîh 1987:10.) Arabists like Serjeant have repeatedly emphasized the fact that, unlike urban centers in the Arab world, this isolation was a major factor in preserving the type of life in Ḥaḍramawt, which resembles the time of the rise of Islam (Lewcock 1986:37).

Due to a long history of immigration of Ḥaḍramīs to many parts of the world, Ḥaḍramī Arabic had and still has a number of loanwords taken from different sources including English, Indonesian/Malay, Swahili, and the languages of the Indian subcontinent.

Words related to food, like *ṣambal* 'fried (vegetable) cooking', in combinations like *ṣambal ḥanīd* 'dried ground fish fried with onions', *ṣambal blēšān* 'a sambal of shrimp paste (*blacan*)', *ṣambal hāb* 'dry sambal', etc., can be heard, not only among families with Southeast Asian links in big towns of the Wādī but also among others, including families with no history of immigration (Al-Saqqaf 2006).

BIBLIOGRAPHICAL REFERENCES

Al-Saqqaf, Abdullah Hassan. 1999. *A descriptive linguistic study of the spoken Arabic of Wādī Ḥaḍramawt*. Ph.D. diss., University of Exeter.

——. 2005. "Ḥaḍramī Arabic verbal morphology". *Matériaux Arabes et Sudarabiques (MAS-Gellas) 2002–2005*, NS 10.91–128.

——. 2006. "The linguistics of loanwords in Ḥaḍramī Arabic". *International Journal of Bilingual Education and Bilingualism* 9.75–93.

ʿAqīl, ʿAli. 1981. "Namūḏaj min al-lahja al-yamaniyya fī Wādī Ḥaḍramawt". *Journal of the Gulf and Arabian Peninsula Studies* 7-8.131–144.

Bâfaqîh, Mohammad. 1987. "Quelques racines communes au sudarabique épiraphique et à l'arabe" (in Arabic: Min al-juḏūr al-muštaraka bayna ʿarabiyyat an-nuqūš wa-l-ʿarabiyya al-fuṣḥā). *Sayhadic*, ed. Christian Robin and Mohammad Bâfaqîh, 1–12. Paris: Geuthner.

Doe, Brian and Robert Bertram Serjeant. 1975. "A fortified tower house in Wādī Jirdān". *Bulletin of* *the School of Oriental and African Studies* 38.276–295.

Johnstone, Thomas Muir. 1967. *Eastern Arabian dialect studies*. London: Oxford University Press.

Lewcock, Ronald. 1986. *Wādī Ḥaḍramawt and the walled city of Shibām*. Paris: UNESCO.

Qafisheh, Hamdi A. 1977. *A short reference grammar of Gulf Arabic*. Tucson, Ariz.: University of Arizona Press.

——. 1992. *Yemeni Arabic reference grammar*. Kensington, Md.: Dunwoody Press.

Quʿayṭī, Sayf ibn Ḥusayn al-. N.d. *Xaṣāʾiṣ al-luġa al-ḥaḍramiyya*. Hyderabad. (Ms., Osmania University, India.)

Watson, Janet C.E. 1993. *A syntax of Ṣanʿānī Arabic*. Wiesbaden: O. Harrassowitz.

Abdullah Hassan Al-Saqqaf
(Sultan Qaboos University)

Waqf → Pausal Forms

Waṣf → Ṣifa

Waṣl → Ṣila

Weak Verbs

1. Introduction

Apart from functional words such as *ḥattā* 'until', *li-* 'to', *wa-* 'and', all nominal and verbal forms in Arabic are made up of three or four essential elements, called radicals. Any phoneme, except the vowel *a*, can be the radical of a → root. To give an example, in Modern Standard Arabic (Wehr 1985), the following roots can be formed from the strong consonants *q* and *l* and the so-called weak radicals *u̯* (i.e. the consonantal realization of the vowel *u*), *i̯* (i.e. the consonantal realization of the vowel *i*), and ʾ (glottal stop): *q-u̯-l* 'to say', *q-i̯-l* 'to have a nap after dinner', *q-l-l* 'to be little/few', *q-l-u̯/i̯* 'to fry, bake', *u̯-q-l* (Form V) 'to climb', *q-l-q* 'to sway', *q-l-q-l* 'to move, shake', *l-q-i̯* 'to occur, meet', *l-i̯-q* 'to be appropriate, suit', *q-u̯-i̯* 'to be strong, powerful', *u̯-q-i̯* 'to protect, preserve', *q-i̯-ʾ* 'to vomit', *l-u̯-i̯* 'to turn, bend', *u̯-l-i̯* 'to be near, to follow', *ʾ-u̯-i̯* 'to seek refuge', *ʾ-u̯-l* 'to return', *ʾ-l-i̯* 'to omit', *u̯-ʾ-l* 'to take refuge', *l-ʾ-l-ʾ* 'to shine, glitter', *l-ʾ-i̯* 'slowness', and possibly others. Nouns of probably foreign origin, like *qāqulla* 'cardamom', are disregarded here. These examples show clearly that the weak elements, *u̯*, *i̯*, and ʾ, cannot simply be regarded as prop elements that were added later, as some scholars have argued.

Some root types are more common than others: type 1–2–2 (geminated verbs), for instance, is common, type 1–2–1 (e.g. shortened from type *1–2–1–2) is rare, and type 1–1–2 is even rarer. Because Arabic has a constraint against having two identical consonants – except when reduplicated or lengthened – in the same root, the original character of the verb type 1–2–2 has been disputed. To some extent, this is justified because in roots of this nature it is not the case that two consonants come together, as in verb type 1–2–2 (< *1–2–1–2), but a lengthened consonant appears in two positions. If one were to posit a rule **q-l* > *q-l-l*, one would equally have to posit **q-l* > *q-i̯-l*, *q-u̯-l*, *q-l-u̯*, etc., which would lead to great difficulties in the reconstruction of Semitic roots. It seems preferable to assume for the geminated roots a structure *1–2:, with a lengthened second consonant (indicated here with :). Lengthening the second radical leads to the structure (*1–2: >) 1–2–2, as in the reduplicated form *(1–2): > 1–2–1–2.

Concerning consonants, with the exception of the semivowels, only a small number of special processes are encountered, such as → assimilation to emphatic consonants (e.g. Form VIII *(i)ḍtaraba* > *(i)ḍṭaraba* 'he loitered') or to voiced consonants (e.g. Form VIII *(i)ztahara* > *(i)zdahara* 'he blossomed').

In contrast, the semivowels in the position of first, second, or third radical may show extensive assimilations and morphological adaptations. The general rule is that the semivowels that are assumed to be original (S: *u̯*, *i̯*) are elided between two short vowels (VSV) and between a short and a long vowel (VSVV). Yet, the sound sequences *i*S*a* (> *iya*), *u*S*a*, VS*aa* are exceptional in that they are generally preserved. On contact between two vowels, the vowel *a* proves the strongest because it can never be assimilated to *u* or *i*. Similarly, on contact between *i*/*i* und *u*/*u*, *i* is found to be stronger (**ui*, **iu* > *ii*). Hence, a vowel hierarchy (*a* > *i* > *u*) may be observed. In one case, the position of the stress has to be taken into consideration as the distinguishing factor between *aSV > *aa* and V, respectively. Furthermore, morpheme juncture (noted as =) plays a part.

For ease of derivation, the weak verbal forms are always shown in their triradical form. A few deviations from the transcription system are used here. For clarity's sake, a distinction is made, depending on the position in the syllable, between *u* and *u̯*, and *i* and *i̯*. Yet, this is not important for the formulation of the sound laws. The semivowels that do not assimilate in consonantal position are transcribed *w* and *y*, as, for instance, the personal suffix in *yaʾkulu* 'he eats'. With long vowels, a distinction is made between contractional length (denoted with double vowels *aa, ii, uu* for *ā, ī, ū*, as in *qaala* 'he said') and morphological length (as in *qa:ʾil* 'saying'). About the problems of weak verbs in general, see Voigt (1988).

2. Verbs Iʾ

i. With hamzated verbs, attention must be drawn to the orthographic rules governing the choice of the → *hamza* bearer, i.e. the written character (*w, y* without dots, and *ʾalif*) above which or under which the *hamza*, if necessary, is to be placed.

   (a) A *hamza* (glottal stop) does not need a bearer after a long or closed syllable, as in *tala:ʾamū* 'they reconciled with each other', *yasʾalu* 'he asks'.
   (b) Word-initially and adjacent to the *a*-vowel, *ʾalif* is the bearer, as in *ʾakala* 'he ate', *ʾukila* 'it was eaten', *yaʾṯuru* 'he tells [a story]'.
   (c) In the vicinity of the vowel *i*, the *hamza* bearer is *y*, as in *xaṭiʾa* 'he sinned', *suʾila* 'he was asked'.
   (d) In the vicinity of the vowel *u*, the *hamza* bearer is *w* if no *i* appears, as in *baʾusa* 'he was brave', *yuʾminu* 'he believes'.
   (e) If *hamza* comes to stand at the end of the core morpheme (here marked by =), the preceding vowel determines which bearer must be chosen. Thus, in *ʾaqraʾ=u* 'he reads', it is *ʾalif*, and in *yuqriʾ=u* 'he lets read', it is *y* that is the *hamza* bearer.

   These rules point to a preference hierarchy in choosing vowels as *hamza* bearers, with *a* being the weakest and *i* the strongest: *a* < *u* < *i*. These rules reflect the pronunciation of certain dialects in the Ḥijāz in which *hamza* was lost and replaced by a glide. In these dialects, *xaṭiʾa* and *baʾusa* were pronounced **xaṭiya* and **bawusa*, and it was to these forms that *hamza* was later added.

ii. Dissimilation rules ʾVʾ > ʾVV, e.g. **ʾaʾkulu* > *ʾaakulu* 'I eat', **ʾuʾḏina* > *ʾuuḏina* 'he was given the news'.
iii. In the imperative of the verbs *ʾaxaḏa* 'he took', *ʾamara* 'he ordered', and *ʾakala* 'he ate', the initial syllable is elided: *xuḏ* (< **ʾxuḏ*, imperfect *yaʾxuḏu*), *kul* (imperfect *yaʾkulu*), *mur* (imperfect *yaʾmuru*), but **iʾsar* > *ʾiisar, wa-ʾsar* '(and) bind!'.
iv. With some verbs in Form VIII, *hamza* assimilates to the *t*, characteristic of the formation of this stem: **(i)ʾtaxaḏa* > *(i)ttaxaḏa* 'he took for himself'.
v. Word-internally, *hamza* can sometimes be elided. Thus, the imperative of *saʾala* 'he asked' is *sal* (< **(i)sʾal*). As for *raʾaa* 'he saw', see below, Section 8.

3. Verbs Iu̯

i. In the basic stem, the verbs Iu̯ lose their first radical in the imperfect, imperative, and infinitive, if the imperfect stem is characterized by the vowel *i*. From *walada* 'he begot', *wariṯa* 'he inherited', *waḍaʿa* 'he laid', the following are formed:

   (a) imperfect: *yalidu* (< **yau̯lidu*), *yariṯu*, *yaḍaʿu* (< **yau̯ḍiʿu* with *i* > *a* due to the laryngeal)
   (b) imperative: *lid* (< **ulid*), *riṯ, ḍaʿ*
   (c) infinitive: *lida* (< **uVlidat*^{un}), *riṯa*, *ḍaʿa*

ii. Verbs with *a* as characteristic vowel of the imperfect pattern as strong forms, e.g. *wabiqa/yawbaqu* 'he perished', imperative *iibaq* (< **(i)u̯baq*).
iii. The sound sequence **iu̯* becomes *ii*, as in the infinitive of Form IV: *ʾiira:d* (< **ʾiu̯ra:d*) 'to lead to water' (*u̯-r-d*), and in the noun *miiṯa:q* (< **miu̯ṯa:q*) 'agreement, contract' (*u̯-ṯ-q*).
iv. In Form VIII, *u̯* is assimilated to the *t* characteristic of the formation of this Form: *(i)ttafaqa* (< **(i)u̯tafaqa*) 'he came to an agreement' (*u̯-f-q*).

Form IV *yuu̯qifu* = *yuuqifu* ‘he lets stand’ does not constitute a sound change in the proper sense.

The rules mentioned so far are of a quite divergent nature. The dissimilation rule *ʾVʾ > ʾVV and the assimilation rule **iu̯* > *ii* are of a purely phonetic nature, for instance. Also phonetic in origin is the elision of initial *u̯* in the imperative of the verbs *u̯*-2–3 with imperfect vowel *i*: **u̯lid* > *lid* and in the imperative of some verbs Iʾ. If this is correct, the imperative must go back to an early period when a prefixed prop vowel developed, e.g. **ktub* > *(u)ktub, wa-ktub* ‘(and) write!’.

Partial phonetic motivation can be seen in the assimilation of ʾ, *u̯*, and *i̯* as first radical to the *t* characteristic of this stem (**ʾt*, **u̯t*, **i̯t* > *tt*). The lengthening of the *t* compensates for the impending elision of the first root radical through lengthening of the dental.

Differently again, a morphological process can be seen in the elision of initial *u̯* in the imperfect (*yalidu*) and the infinitive (*lida*). In the imperfect, it is the morphological proximity to the imperative (*lid*) that causes the shortening.

4. Verbs IIu̯

Some forms of verbs II*u̯* pattern like strong verbs, for instance the imperfect of the basic stem *yaquumu* = *yaqu̯umu* ‘he gets up, stands up’, which corresponds to the imperfect of the strong verb (*yaktubu* ‘he writes’). For most other forms, sound assimilations and reductions must be posited.

i. Perfect active I

The most remarkable morpheme variation occurs in the perfect active: *qaama/qumta* (imperfect *yaquumu*) ‘to get up, stand up’, *naama/nimta* (imperfect *yanaamu*) ‘to sleep’. These forms can only be explained if one posits as original forms **qau̯uma* and **nau̯ima*. In addition, stress must be introduced as a distinguishing factor.

qáu̯uma* > **qáuma* (elision of *u̯*) > *qaama* (áu* > *aa*) ‘he got up’
qau̯úmta* > **qaúmta* (elision of *u̯*) > *qumta* (aú* > *u*) ‘you [masc. sg.] got up’
náu̯ima* > **náima* (elision of *u̯*) > *naama* (ái* > *aa*) ‘he slept’
nau̯ímta* > **naímta* (elision of *u̯*) > *nimta* (aí* > *i*) ‘you [masc. sg.] slept’

It is obvious that in the sound sequence *au̯*V the first vowel is predominant (**au̯*V > *aa*), except when the second vowel (V) is stressed (**au̯*V́ > V́). Only here, i.e. in the perfect of verbs with weak second radical, is the vowel alternation *aa* (in open syllable) ~ *u* (in closed syllable) encountered. Otherwise, a long vowel is always shortened in closed syllable, e.g. *ʾaqaama* ‘he made rise’ – *ʾaqamta* ‘you [masc. sg.] made rise’.

ii. Perfect passive I

The form *qiila* ‘it was said’, which corresponds to the form *fuʿila* of the strong verb, can be explained by the rule **u(u̯)i* > *ii*: **quu̯ila* > **quila* (elision of *u̯*) > *qiila* (*ui* > *ii*).

In a closed syllable, the vowel is shortened: **qiilta* > *qilta*. In this case, the different position of the stress plays no further part (**qúu̯ila* > *qiila*, **quu̯ílta* > *qilta*).

iii. Imperfect active I

Proceeding from the form *yaquumu* ‘he gets up’ and *yanaamu* ‘he sleeps’, a long vowel undergoes shortening in a closed syllable, as in the jussive and imperative: *yaqum/qum, yanam/nam*.

iv. Imperfect passive I

The form *yuqaalu* ‘it is said’ can be derived according to the rule *§*u̯a* > *aa* (the sign § marks the beginning of a new syllable) from **yuqu̯alu* (like *yuktabu*). In a closed syllable, this long vowel is shortened: *yuqal*.

v. Active and passive participle I

Active participle **qa:u̯il* > *qa:ʾil* ‘saying’. The rule **a:u̯i* > *a:ʾi* only applies to Form I. For the imperfect and active participle of Form III, see below. Passive participle **maqu̯u:l* > *maquul* ‘said’.

vi. Perfect and imperfect in the derived Forms

The following sound laws must be taken into account here:

- **au̯a* > *aa*: Perfect active VIII: **(i)qtau̯ada* > *(i)qtaada* ‘he led, he let himself be led’

- *_§u̯a_ > _aa_: Perfect active IV: *_ʾaqu̯ama_ > _ʾaqaama_ ‘he remained’
- *_au̯i_ > _aa_: Imperfect active VIII: *_yaqtau̯idu_ > _yaqtaadu_ ‘he leads, he lets himself be led’
- *_u̯i_ > _ii_: Imperfect active IV: *_yuqu̯imu_ > _yuqiimu_ ‘he remains’, imperative active IV: *_ʾaqu̯imuu_ > _ʾaqiimuu_ ‘remain [pl.]!’
- *_uu̯i_ > *_ui_ > _ii_: Perfect passive VIII: *_(u)ṣtuu̯ida_ > _(u)ṣṭiida_ ‘he was hunted down’ (with *_ṣt_ > _ṣṭ_)
- *_iu̯a:_ > _iya:_: Infinitive VIII: *_(i)qtiu̯a:d_ > _(i)qtiya:d_ ‘to let oneself be led’

In the derived stems, i.e. in the imperfect and active participle of Form III, the sound sequence _a:u̯i_ is preserved. This is due to the fact that the verbs in Form III and also in Form II pattern like strong verbs, which suggests the adoption of the spelling _aw:i_ and _a:wi_, respectively, e.g. II _yuṣaw:iru_ ‘he paints’ and III _yuja:wiru_ ‘he is (somebody’s) neighbor’. Furthermore, the sound sequence _au̯:a_ (> _aw:a_) and _a:u̯a_ (> _a:wa_) remains generally preserved, e.g. II _ṣaw:ara_ ‘he painted’ and _ja:wara_ ‘he was (somebody’s) neighbor’, respectively. The formations of the Forms II and III with _w_ as second radical were only developed at a later stage in the language and thus cannot be reconstructed as Proto-Semitic.

These are the rules that apply to verbs IIu̯:

- *Vu̯V > VV
- *_au̯_V > _aa_
- *_au̯_V > V
- *_§u̯a_ > _aa_
- *V_iV_i (in closed syllable) > V_i
- *_(u)u̯i_ > _ii_
- *_iu̯a:_ > _iya:_
- *_a:u̯i_ > _a:ʾi_ (in Form I only)

5. VERBS IIIu̯

Verbs IIIu̯ have three basic types:

- perfect _ġazaa_ (< *_ġazau̯a_), imperfect _yaġzuu_ ‘to make a raid’
- perfect _raḍiya_ (< *_raḍiu̯a_), imperfect _yarḍaa_ ‘to be content, wish’
- perfect _saruwa_ (< *_saruu̯a_), imperfect _yasruu_ ‘to be noble’

The following sound changes may be observed:

i. Basic type _faʿala_

Perfect: 3rd person masculine singular *_ġazau̯a_ > _ġazaa_; 3rd person masculine plural *_ġazau̯uu_ > _ġazauu_ (elision of _u̯_) > _ġazaw_ (elision of _u_). All the other forms are regular, e.g. 2nd person masculine singular _ġazau̯ta_ (= _ġazauta_ = _ġazawta_), 3rd person masculine singular *_ġazau̯at_ > _ġazaat_ > _ġazat_ (with vowel shortening). The spelling of the form _ġazaa_ (< *_ɣazaː_) with _ʾalif_ as opposed to the spelling of _ramaa_ (< *_ramai̯a_) ‘he threw’ with _y_ indicates that the phonetic realization of the two final vowels must have been different at the time their orthography was established, possibly [ɣɑzɑː] with an open back vowel, versus [ramæː] with an open front vowel.

Imperfect: Although a number of regular forms can be found here (e.g. subjunctive 3rd pers. masc. sg. _yaġzuu̯a_), most forms are shortened. The indicative can easily be derived with the rules stated so far.

- 3rd person masculine singular *_yaġzuu̯u_ > _yaġzuu_
- 2nd person feminine singular *_taġzuu̯iina_ > *_taġzuiina_ (elision of _u̯_) > _taġziina_ (predominance of the long vowel of the ending)
- 3rd person masculine plural *_yaġzuu̯uuna_ > *_yaġzuuuna_ (elision of _u̯_) > _yaġzuuna_

The jussive, though, proves more difficult. It cannot have been different from the indicative forms with no ending since the jussive *_yaġzuu̯_ (like _yaktub_) must have been identical with the indicative _yaġzuu_ (< *_yaġzuu̯u_ like _yaktubu_). The indicative is usually the longer form compared to the jussive. From a historical perspective, one arrives at the formula: jussive _yaktub_ plus endings _-u_ (3rd pers. masc./fem. sg., 1st pers. pl.)/_-na_ (2nd pers. fem. sg., 3rd pers. masc. pl., 2nd pers. masc. pl.)/_-ni_ (du.)/-Ø (3rd. pers. fem. pl., 2nd. pers. fem. pl.) ⇒ indicative. In contrast to this, speakers seem to have perceived the jussive as the outcome of a shortened form, which leads to the formula: indicative minus endings _-u/-na/-ni/-Ø_ ⇒ jussive, hence the synchronous connection in Table 1.

Table 1. Indicative and jussive endings of the strong verb

	indicative	minus ending	⇒	jussive
sg. 3rd m.	_yaktubu_	_-u_	⇒	_yaktub_
2nd f.	_taktubiina_	_-na_	⇒	_taktubii_
du. 3rd m.	_yaktubaani_	_-ni_	⇒	_yaktubaa_
pl. 3rd m.	_yaktubuuna_	_-na_	⇒	_yaktubuu_
3rd f.	_yaktubna_	-Ø	⇒	_yaktubna_

Verbs with _u̯_ as last radical follow the same pattern, as in Table 2.

Table 2. Indicative and jussive of verbs with _u̯_ as last radical

	indicative	minus ending	⇒	jussive
sg. 3rd m.	_yaġzuu_	_-u_	⇒	_yaġzu_
2nd f.	_taġziina_	_-na_	⇒	_taġzii_
du. 3rd m.	_yaġzuwaani_	_-ni_	⇒	_yaġzuwaa_
pl. 3rd m.	_yaġzuuna_	_-na_	⇒	_yaġzuu_
3rd f.	_yaġzuuna_	-Ø	⇒	_yaġzuuna_

It follows from this that some forms have word-finally a short _-u_, which cannot be seen as the outcome of a regular sound change from *_-uu̯_. The same applies to the imperative _(u)ġzu_, which corresponds to strong _(u)ktub_.

Active participle nominative, genitive *_ġa:ziu̯=un/in_ > _ġa:zin_, accusative *_ġa:ziu̯=an_ > _ġa:ziyan_, with article nominative, genitive *_(a)l-ġa:ziu̯=u/i_ > _(a)l-ġa:zii_, accusative *_(a)l-ġa:ziu̯=a_ > _(a)l-ġa:ziya_ ‘raider, warrior’. Also important is the rule *_iu̯=u/i_ > _ii_, with shortening to _i_ in closed syllable (> _in_).

ii. Other verbal Forms

All other verbal Forms pattern on the verbs with _i̯_ as the last radical. This is clearly demonstrated by the perfect forms 3rd person masculine singular, which are written with final <y>, and by the forms with _ay_ (like Form IV _ʾaʿṭaytu_ ‘I gave’ from the root _ʿ-ṭ-u̯_), which cannot be traced back to *_au̯_ relying on the regular sound changes.

As for the infinitive in _a:ʾ_, the juncture (=) that follows must be taken into consideration, e.g. IV *_ʾirḍa:u̯=_ > *_ʾirḍa:i̯=_ (through analogy) > _ʾirḍa:ʾ_ = ‘satisfaction’. The rule: *_i̯/a:_ ___= > ʾ applies.

iii. Type _faʿila_

The finite verbal forms no longer show any contrast between verbs with _u̯_ or with _i̯_ as their last radical. In the perfect *_raḍiu̯a_ > _raḍiya_ and the imperfect (_yarḍaa_ ‘he is content’), the forms look as if they were derived from a verb III_i̯_.

The diphthongs _au_ and _ai_, which occur in the imperfect of the basic stem, can only be derived when consideration is given to the length of the vocalic ending and their weight:

*_yarḍau̯uuna_ > *_yarḍauuna_ (elision of _u̯_) > _yarḍauna_ (= _yarḍawna_) (_auu_ > _au_) ‘they are content’
*_tarḍau̯iina_ > *_tarḍaiina_ (elision of _u̯_) > _tarḍaina_ (= _tarḍayna_) (_aii_ > _ai_) ‘you [fem. sg.] are content’

This leads to the rule *_au̯_VV > _a_V, or put more generally, *_a_SVV > _a_V.

Only the infinitive _riḍwa:n_ still shows the original third radical; see also the infinitive _riḍan_, with article _(a)r-riḍaa_, which was previously written with final _ʾalif_ (< *_riḍau̯=u/i/an_) but later with _y_ (_ʾalif maqṣūra_; < *_riḍai̯=u/i/an_).

iv. Perfect passive I

The form _qiila_ ‘it was said’, which corresponds to the form _fuʿila_ of the strong verb, is explained by the rule *_u(u̯)i_ > _ii_: *_quu̯ila_ > *_quila_ (elision of _u̯_) > _qiila_ (_ui_ > _ii_).

v. Type _faʿula_

With the perfect of this type, it should be borne in mind that it is not a deep sound change, like _saruwa_ (< *_saruu̯a_), 2nd person masculine singular _saruuta_ (= _saruu̯ta_ = _saruwta_). For the imperfect _yasruu_, one may compare the imperfect _yaġzuu_ (see Sec. 10.iii).

vi. Rooted in the realm of morphology is the rule that the infinitive of Form IV is not *_ʾiqu̯a:m_ but _ʾiqa:ma_ (with feminine ending compensating for the lost _u̯_), likewise in Form X.

To recapitulate the rules which apply to the verbs type III_u̯_, in addition to the rules that apply to verbs of type II_u̯_:

- *_au̯_VV > *_a_VV > _a_V
- *_iu̯a_ > _iya_

- **ui* > *ii*
- **a:u̯=* > *a:ʾ=*

6. VERBS Ii̯

The verbs Ii̯ are formed almost totally regularly, e.g. *yasira*, imperfect *yaysiru* ‘to be easy’.

In formations with *u* before *i̯* as the first radical, the *u*-vowel prevails (**ui̯* > *uu*), cf. the imperfect of Form IV **yui̯siru* > *yuusiru* ‘he is rich’. This is possibly nothing more than an assimilation to the verbs Iu̯, which also show *uu* in Form IV. It is also possible that the vowel -*u* of the personal prefixes (ʾ-, *t*-, *y*-, *n*-) in Form IV and in the passive of Form I carries morphologically more weight than the first root radical that follows it. The imperfect **yui̯siru* (according to the rule **ui̯* > *ii*) could not result in **yiisiru*, in order to avoid confusion in the verbal stem derivation. In this case, the coherence of the verbal stem derivations would prove stronger than the coherence of root derivations. However, if one remembers that very few verbs Ii̯ exist, the assumption of an assimilation to the more common verbs Iu̯ appears less adventurous.

In Form VIII, assimilation of *i̯* to *t* can occur, as in *(i)ttasara* < **(i)i̯tasara* ‘he played at some hazard’. This is an assimilation to the more common verbal type Iu̯, with assimilation of the semivowel to the glottal stop.

7. VERBS IIi̯

i. Perfect active I

The morpheme variation in the perfect active *saara*/*sirta* (imperfect *yasiiru*) ‘to travel’ and *ḫaaba*/*ḫibta* (imperfect *yaḫaabu*) ‘to fear’ is striking. These forms can only be explained if one posits as basic forms **sai̯ira* and **ḫai̯iba*. Furthermore, stress has to be introduced as a distinguishing factor.

sái̯ira* > **sáira* (elision of *i̯*) > *saara* (ái* > *aa*) ‘he traveled’
sai̯írta* > **saírta* (elision of *i̯*) > *sirta* (aí* > *i*) ‘you [masc. sg.] traveled’
ḫái̯iba* > **ḫáiba* (elision of *i̯*) > *ḫaaba* (ái* > *aa*) ‘he was frightened’
ḫai̯íbta* > **ḫaíbta* (elision of *i̯*) > *ḫibta* (aí* > *í*) ‘you [masc. sg.] were frightened’

Compare this with the corresponding forms of the verbs IIu̯ in Table 3.

Table 3. Perfect active of verbs IIu̯ and verbs IIi̯

IIu̯		IIi̯	
**qáu̯um=a*	> *qaama*	**sái̯ir=a*	> *saar=a*
**qau̯úm=ta*	> *qum=ta*	**sai̯ír=ta*	> *sir=ta*
**náu̯im=a*	> *naam=a*	**ḫái̯ib=a*	> *ḫaab=a*
**nau̯ím=ta*	> *nim=ta*	**ḫai̯íbta*	> *ḫib=ta*

Rules applicable to both verbs IIu̯ and IIi̯ in Form I can be formulated as follows: **á*SV > *aa*, **a*SV́ > V́ (S = semivowel, V = *u*, *i*). In the sound sequence *á*SV, it is the first vowel that prevails (> *aa*), except when the second vowel is stressed (*a*SV́ > V́).

ii. Passive participle of the basic Form: **mabi̯u:ʿ* = **mabiuuʿ* > **mabiuʿ* (elision of *u*) > *mabiiʿ* (**iu* > *ii*) ‘sold’

iii. In the derived Forms, the following sound changes must be observed:

- *Vi̯V > VV
- **ai̯*V > *aa*: perfect active VIII **(i)btai̯aʿa* > *(i)btaaʿa* ‘he bought’, imperfect active VIII: **yabtai̯iʿu* > **yabtaaʿu* ‘he buys’
- *§*i̯a* > *aa*: perfect active IV *ʾ*abi̯ana* > *ʾabaana* ‘he was, made clear’
- V_iV_i (in closed syllable) > V_i
- **(u)i̯i* > **ui* > *ii*: perfect passive VIII **(u)btui̯iʿa* > *(u)btiiʿa* ‘it was bought’, imperfect active IV **yubi̯inu* > *yubiinu* ‘he is, makes clear’
- **a:i̯i* > *a:ʾi* (only in Form I)

Since the relevant rules concerning verbs IIu̯ agree with the ones above, a more general formulation is possible:

- *VSV > VV
- **á*SV > *aa*
- **a*SV́ > V́
- *§S*a* > *aa*
- *V_iV_i (in closed syllable) > V_i
- **(u)*S*i* > *ii*
- **i*S*a:* > *ii̯a:*
- **a:*S*i* > *a:ʾi* (only in Form I)

8. Verbs IIIi̯

The above-mentioned rules concerning verbs IIIu̯ are also valid when u̯ is replaced by i̯:

- **ai̯a* > *aa*: perfect **ramai̯a* > *ramaa* 'he threw', written with final <y>, thus hinting at a pronunciation as a long open front vowel [ramæː], which contrasts with the long open back vowel in *ġazaa* [ɣɑzɑː] (< **ġazau̯a*) written with final *ʾalif*. The long vowel is shortened in closed syllable: **ramai̯at* > **ramaat* > *ramat* 'she threw'.
- **aĭu* > *aa*: imperfect passive **yurmai̯u* > *yurmaa* (elision of *i̯*) 'he is thrown'.
- **ai̯uu* > *au*: perfect **ramai̯uu* > **ramauu* (elision of *i̯*) > *ramau* (*auu* > *au*) 'they threw'.
- **ai̯ii* > *ai*: **tubkai̯iina* > **tubkaiina* (elision of *i̯*) > *tubkaina* (*aii* > *ai*) 'you [fem. sg.] are wept over)'.
- **ii̯u* > *ii*: imperfect I **yarmii̯u* > **yarmiu* (elision of *i̯*) > *yarmii* (*iu* > *ii*) 'he throws'; imperfect V **yanbaġii̯u* > **yanbaġiu* (elision of *i̯*) > *yanbaġii* 'it is necessary, it is proper'.
- **ii̯uu* > *uu*: imperfect I **yarmii̯uuna* > **yarmiuuna* (elision of *i̯*) > *yarmuuna* (*iuu* > *uu*) 'they throw'. The long suffix vowel *uu* prevails.
- **ii̯ii* > *ii*: imperfect **tarmii̯iina* > **tarmiiina* (elision of *i̯*) > *tarmiina* 'you [fem. sg.] throw'.
- **ii̯=un/in* > *in*, **ii̯=an* > *iyan*, **ii̯=u/i* > *ii*, **ii̯=a* > *iya*: active participle nominative, genitive *ra:min*, accusative *ra:miyan*, with article nominative, genitive *(a)r-ra:mii*, accusative *(a)r-ra:miya* 'throwing'.

Slightly unusual is the formation of the passive participle **mahdu:i̯=* = **mahdu:y=* > *mahdi:y=* (i.e. *mahdiyy=*) 'rightly guided', where no vowel assimilation across the juncture occurs (as in the case of *ra:min*). The solution lies with the long vowel *u:*, which requires the strong treatment of the following radical; **u:y* has to become *i:y* / *iyy*.

The rules **ai̯a*, **ai̯u* > *aa* can be generalized to **ai̯*V > *aa* (and **a*SV > *aa*), e.g. **maʿnai̯=u/i/a* > *maʿnaa*, shortened with nunation to *maʿnan*.

The different treatment of **ai̯u* and **ai̯i* (> *aa*) in contrast to **ai̯uu* (> *au*) and **ai̯ii* (> *ai*) is important. The long vowels **uu* and **ii* are so strong that they cannot be assimilated to **a* – unlike the short vowel **u* that is assimilated to **a* (**ai̯u* > *aa*). In the case of **ii̯uu* > *uu*, the weight of the ending must account for the prevalence of the long *uu*.

The jussive of Form I and all derived Forms has a short final vowel (e.g. I *yarmi*, V *yatašaf:a*). The above-mentioned rule applies: indicative minus certain endings ⇒ jussive, e.g. imperfect (indicative) I *yarmii* (< **yarmii̯u* 'he throws'), V *yatašaf:aa* (< **yatašaf:ai̯u* 'he achieved healing') minus vowel ⇒ jussive *yarmi*, *yatašaf:a*.

9. Doubly weak verbs

Doubly weak verbs are verbs that belong simultaneously to two weak verbal classes and share the features of both, as long as there is no conflict between the two. Apart from hamzated verbs, there is only one type containing the weak elements u̯ and i̯ as its first and third radical (Iu̯/IIIi̯), respectively, e.g.

- *waqaa* (< **waqai̯a*), imperfect *yaqii* (< **yau̯qii̯u*) 'to guard, preserve' with imperative *qi* (strengthened to *qih*)
- *waʾaa* (< **waʾai̯a*), imperfect *yaʾii* (< **yau̯ʾii̯u*) 'to promise' with imperative *i* (strengthened to *ih*)

If the *hamza* is regarded as a further weak element, it is possible to establish many verbal types having, in addition to the semivowel, an additional *hamza* in the first, second, or third position. Only in rare cases can special formations be observed.

- Iʾ and IIu̯: *ʾaaba*, imperfect *yaʾuubu* 'to return', imperative *ub*, have regular formation.
- IIʾ and IIIi̯: *raʾaa*, imperfect *yaraa* (< **yarʾai̯u*) 'to see', imperative *ra* (strengthened to *rah*) with loss of *hamza* in some forms.

10. Verbs with strong *w* and *y* as the second radical

Verbs with a weak third radical must treat u̯/i̯ as second radical as strong (II*w*/*y*), e.g.

- *šawaa* (< **šawai̯a*), imperfect *yašwii* 'to broil' with imperative *(i)šwi*
- *qawiya*, imperfect *yaqwaa* (< **yaqwai̯u*) 'to be strong' with imperative *(i)qwa*
- *ḥayiya*, *ḥayya* (< **ḥayiu̯a*), imperfect *yaḥyaa* (irregularly word-finally written with

ʾ_alif_) ‘to live’ with imperative _(i)ḥya_; this verb may undergo shortening in Form X: _(i)staḥaa_ (< _(i)staḥyaa_), imperfect _yastaḥii_ (< _yastaḥyii_)

- ʾ_awaa_, imperfect _yaʾwii_ ‘to betake oneself to’

There are strong verbs with a consonantal _w_ or _y_ as their middle radical, like _sawida/yaswadu_ ‘to be black’, ʿ_awira/yaʿwaru_ ‘to be one-eyed’, which do not differ from regular strong verbs (like ʿ_ajiba/yaʿjabu_ ‘to be astonished’). Here, _w_ is strong because the corresponding adjectives ʾ_aswadu_ ‘black’ and ʾ_aʿwaru_ ‘one-eyed’ also display strong formation.

In the Forms II and III, the semivowels show up in their consonantal forms _w_ and _y_. This can also influence the strong formation in the other Forms, such as VIII _(i)zdawaja/yazdawiju_ ‘to be in pairs’ (_zd_ < *_zt_), formed from the denominal Form II (_zaw:aja/yuzaw:iju_ ‘to pair’). Denominal stems tend to form strong roots with _w_ and _y_, e.g. III ʿ_a:yana_ ‘to view’, denominally derived from ʿ_ayn_ ‘eye’, and III _qa:wala_ ‘to confer, dispute’, derived from _qawl_ ‘word, speech’.

11. Summary of rules

In the formation of weak verbal types, different kinds of rule interact: purely phonetic rules, morphophonological rules, and morphological rules.

i. _Phonetic rules_

Phonetic rules are those that are phonetically motivated:

- The dissimilation rule ʾVʾ > ʾVV with hamzated verbs
- The weakness (i.e. loss) of _hamza_ (ʾ) occurring in the formation of some (frequently used) roots
- The elision of _u̯_ as the first radical of a root in the imperative, e.g. _lid_ (< *_u̯lid_)
- *VSV(V) > VV(V), most common sound law (intervocalic semivowel is elided), exceptions are: _i_S_a_ (> _iya_, e.g. _raḍiya_), _u_S_a_ (e.g. _saruwa_), VS_aa_ (e.g. _yarḍayaani, yaġzuwaani, raḍiyaa_; _i_S_aa_ > _iyaa_); in some cases, this triggers a further sound development (like *_quu̯ila_ > *_quila_ > _qiila_ ‘it was said’)
- *_a_SV > _aa_ (it is important that V is short and unstressed), including *_a_S=V > _aa_ (this contraction occurs also across the juncture =)
- *_a_SV́ > V́ (in the perfect of verbs IIS)
- *V_iV_i (in closed syllable) > V_i
- *_ui_ (_ui̯_ and _u̯i_) > _ii, iu_ > _ii_ (*_mabi̯u:_ ʿ = *_mabi̯uu_ʿ > *_mabiu_ʿ > _mabii_ʿ ‘sold’)
- A series of three identical vowels is shortened by one element, if the syllable structure allows this: *_maqu̯u:l_ > _maquul_ ‘utterance’, but _maġzu:w_ = _maġzuww_ ‘assaulted’

ii. _Morphophonological rules_

Morphophonological rules apply when a phonetic rule needs additional morphophonological input. To this group belong the assimilation of ʾ, _u̯_, and _i̯_ as first radical to the dental of Form VIII (> _tt_). *§S_a_ > _aa_ – this rule does not apply to adjectives IIS of the pattern ʾ_afʿalu_, like ʾ_aʿwaru_ ‘one-eyed’, ʾ_axyaru_ ‘better’.

Two rules have been mentioned involving the sound sequence *_a:_S. In the active participle of Form I, the rule *_a:_S_i_ > _a:ʾi_ must be observed, and in some infinitives (e.g. of Form IV: ʾ_irḍa:_ʾ=), the rule *_a:_S= > _a:_ʾ=. Both rules can be collapsed into the rule that *_a:_S becomes _a:_ʾ with the exception of Forms II and III.

iii. _Morphologically determined rules_

- Loss of _u̯_ as first radical in the imperfect (_yalidu_) and infinitive (_lida_) of the verbs of the _i_-class. Solely the loss of _u̯_ in the imperative (_lid_) could be given a phonetic explanation.
- The final short vowel in the jussive and the imperative of the verbs III_S_ (e.g. _yarmi, yaġzu, yarḍa_) must be seen as the result of a shortening process applied to the longer forms of the indicative (here < _yarmii, yaġzuu, yarḍaa_).
- The roots III_u̯_ become III_i̯_ in the derived Forms through the phenomenon of analogy. The _i_-vowel in the imperfect and participle active of some stems has triggered the convergence with verbs III_i̯_, e.g. Form IV *_yuftiu̯u_ > _yuftii_ ‘he delivers a legal opinion’.
- The verbs I_i̯_ change in Form IV into verbs I_u̯_ (*_yui̯bisu_ > _yuubisu_). It is possible that here, too, the heavier weight of the prefix vowel vis-à-vis the core morpheme plays a role (see the next item).
- For morphological reasons, the personal endings of the perfect and imperfect that

have a long vowel, and the endings of the masculine plural carry special weight. The endings in question are

- perfect 3rd person masculine plural *=uu*
- imperfect 2nd person feminine singular *=ii(na)*
- imperfect 2nd/3rd person masculine plural *=uu(na)*
- nominal ending plural nominative *=uu(na)*, genitive, accusative *=ii(na)*

In front of all these endings there is VS=, i.e. the semivowel as the third radical preceded by any kind of vowel. Altogether, these are the combinations that can occur: *a*S= (**ġazau̯=*, **yuġzau̯=*, *ramai̯=*, *yurmai̯=*), *u*S (**saruu̯=*, **yaġzuu̯=*) and *i*S= (**rumii̯=*, **yarmii̯=*, **raḍiu̯=*, **yurḍiu̯=*). An example for the nominal ending is nominative **qa:ḍii̯=uuna* > *qa:ḍuuna*, genitive, accusative **qa:ḍii̯=iina* > *qa:ḍiina* ‘decisive [pl.]’.

VS coalesces with the long vowel that follows in this way:

- If no *a*-vowel precedes the semivowel, then the vowel of the ending prevails: **i*S*ii* > *ii*, **i*S*uu* > *uu*, **u*S*ii* > *ii*, **u*S*uu* >*uu*.
- If *a* precedes the semivowel, the diphthongs *ai* (= *ay*) or *au* (= *aw*) arise respectively: **a*S*ii* > *ai*, **a*S*uu* > *au*.

This may be seen as a tendency to give paradigmatic coherence (i.e. distinctiveness of endings) preference over root coherence. This leads to the loss of distinction between III*i̯* and III*u̯*. Perhaps it is no coincidence that the endings of the dual (perfect 3rd pers. masc. *-aa*; imperfect 2nd/3rd pers. masc. *-aa(ni)*) are impervious to assimilation (VS*aa* remains intact, apart from the sound change **iu* > *ii*). However, it is possible to see in this case a wide-ranging merger of III*S* into III*i̯* caused by the vicinity of the vowel *i* and caused by analogy in the derived Forms (with only the transitive verbs III*u̯* preserving their original third radical in Form I).

12. Geminated verbs (II=III)

Geminated verbs have identical second and third radicals. From a diachronic point of view, the original root structure cannot have been 1–2–2 but rather 1–2:. For clarity’s sake, it seems nevertheless preferable to use the notation 1–2–2.

The conjugation of the geminated verbs links up with the conjugation of the strong verb, if the following rules are observed. These rules apply to the vowel between the two identical consonants, if a vowel follows the last consonant (X2V2V).

i. If X is a vowel, then the (short) vowel between the second and third radical is elided, (*V2V2V > V22V), **farara* > *farra* ‘he fled’, **masisa* > *massa* ‘he touched’, but *farartu* ‘I fled’.
ii. If X is a consonant, then the vowel between the second and third radical goes in front of the second radical (*C2V2V > CV22V), **yafriru* > *yafirru* ‘he flees’, but *yafrir*, imperative masculine plural **(i)friruu* > *firruu* ‘flee!’, but feminine *(i)frirna*.
iii. Occasionally, rule (ii) applies even if the third radical is not followed by a vowel: imperative masculine singular *firra/i* (with prop-vowel *-a/i*) instead of and beside *(i)frir*, jussive *yafirra/i* instead of and beside *yafrir*.
iv. Rule (i) also applies to the active participle of the basic stem, even though this leads to a syllable structure which is normally not acceptable, e.g. **ʿa:mimun* > *ʿa:mmun* ‘general, common’. Also, cf. perfect III *fa:rra* instead of and beside *fa:rara* ‘he confided a secret’.

If the last radical is followed by a consonant, it conjugates according to the strong pattern, e.g. *farartu* ‘I fled’.

13. Nominal forms

Finally, a number of nominal forms must be mentioned that are characterized by roots with weak radicals since they have sound sequences that do not occur in verbs:

- **uu̯u:* > *uʾu:*, **ṯuu̯u:bun* > *ṯuʾu:bun* pl. of *ṯaubun* ‘garment’
- In nominal forms connected to Form III, the semivowel is treated as strong, e.g. *jiwa:run* ‘neighborhood’ (cf. *ja:wara/yuja:wiru* ‘to be a neighbor’)

Other forms can be subsumed under the stated rules without any problems even though the exact rule formulation may be slightly different, e.g. *ṭau̯i̯ > ṭayy, infinitive of ṭawaa (< *ṭawai̯a)/ yaṭwii 'to fold' (general rule *ui > ii).

BIBLIOGRAPHICAL REFERENCES

Brockelmann, Carl. 1962. *Arabische Grammatik*. 15th ed. Leipzig: VEB Verlag Enzyklopädie.

Caspari, C.P.: See Wright (1964).

Fleisch, Henri. 1961, 1979. *Traité de philologie arabe*, I, II. Beirut: Imprimerie Catholique.

Voigt, Rainer. 1988. *Die infirmen Verbaltypen und das Biradikalismus-Problem*. Stuttgart: F. Steiner.

Wehr, Hans. 1985. *Arabisches Wörterbuch für die Schriftsprache der Gegenwart*. 5th ed. Wiesbaden: O. Harrassowitz. [English transl. J.M. Cowan, *A dictionary of Modern Written Arabic*. Ithaca, N.Y.: Cornell University Press, 1976.]

Wright, William. 1964. *A grammar of the Arabic language*. Trans. from the German of Caspari..., I. 3rd ed. Cambridge: Cambridge University Press.

RAINER VOIGT (Free University of Berlin)

West Sudanic Arabic

1. ARABIC OF CHAD, CAMEROON, AND NIGERIA

The Arabic of → Chad, → Cameroon, and → Nigeria forms a broadly homogeneous dialect region characterized by a number of features either unique to Arabic dialects or found only rarely outside of the region (see Sec. 7). Within this homogeneity, at least two clear subdialects are discernible. In the southern fringe of the area, beginning in eastern Nigeria near the Cameroonian border and stretching through Cameroon and on to an as-yet-unresearched border in Chad is what Owens has termed 'Bagirmi Arabic', after the historical political area where this dialect is found. It is the only dialect in the Sudanic region, for instance, with syllable-final disyllabic stress Cv'CvC, as in *ka'tab* 'he wrote'. The other subdialect includes some of the urban varieties of Chad, as described especially for Ndjamena (Worbe 1962; Pommerol 1990, 1999a,b) and for Abbéché (Roth 1979). These are characterized by certain simplification tendencies, as noted in Section 8.

The Arabs came to the region as early as the late 14th century – there is a letter written in 1391 by the Mai of Kanem, the region adjoining Lake Chad to the east, to the Mamluk sultan Barqūq of Egypt – enjoining him to restrain the marauding, pillaging Arab nomads. These Arabs came from Upper Egypt in the wake of the Mamluk conquest of Nubia, a conquest which itself was part of a Mamluk policy to rid Egypt of unruly Bedouin. The nomadic Arabs

Map 1. Approximate area of Baggara Belt, Western Sudanic Arabic

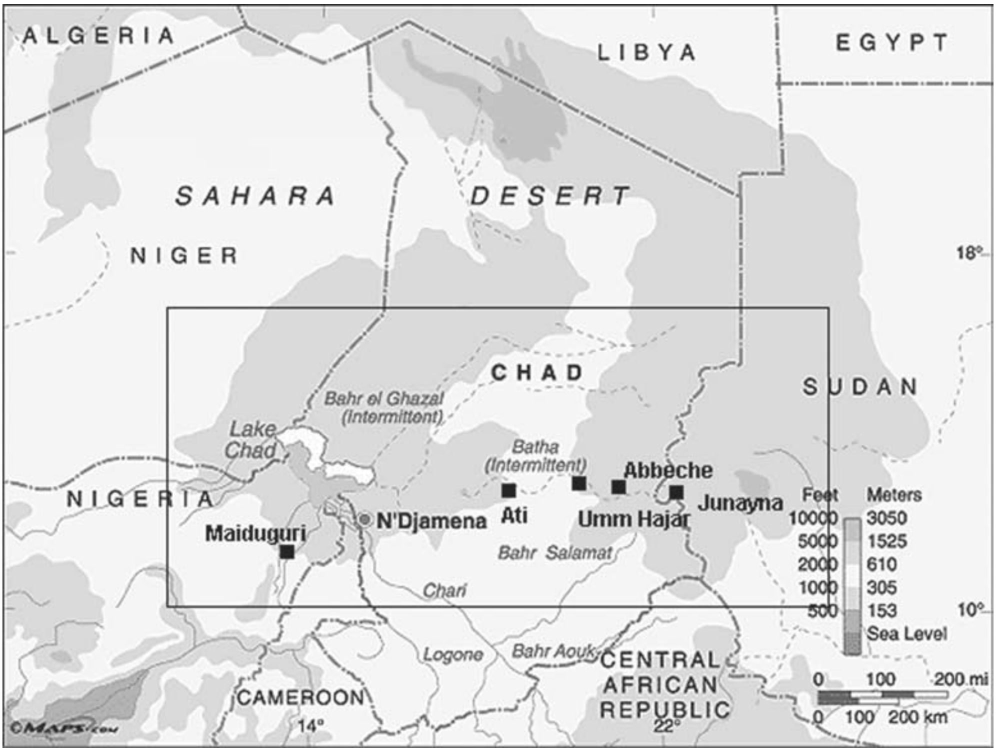

ALGERIA
LIBYA
EGYPT
SAHARA
DESERT
NIGER
18°
CHAD
SUDAN
Bahr el Ghazal (Intermittent)
Lake Chad
Batha (Intermittent)
Abbeche
NIGERIA
Maiduguri
N'Djamena
Ati
Umm Hajar
Junayna
Bahr Salamat
Chari
Feet
Meters
10000
3050
5000
1525
2000
610
1000
305
500
153
Sea Level
10°
Logone
Bahr Aouk
CENTRAL AFRICAN REPUBLIC
CAMEROON
14°
22°
0 100 200 mi
0 100 200 km
©MAPS.com

Table 1. Consonants in West Sudanic Arabic

	labial	dental	prepalatal	palatal	postpalatal	laryngeal
plain	b	d	j	g		
emphatic		ḍ				
implosive		ɗ̣				
plain	p	t	c	k	q	ʾ
plain	f	s	š	x		h
emphatic		ṣ				
plain		l, r				
emphatic		ḷ, ṛ				
plain	m	n	ny [ɲ]	ng [ŋ]		
emphatic	ṃ					
plain	w		y			

Table 2. Correspondences between Old Arabic and West Sudanic Arabic

Old Arabic	West Sudanic Arabic	
**ġ*	/q/, /x/	*qanna*/*xanna* 'to sing'
**ḥ*	/h/	*hilim* 'to dream' (merges with *h* as in *šahar* 'month')
*ʿ	/ʔ/	*baʾarif* 'I know'
**ṭ*	/ɗ̣/	*ṛabaɗ̣* 'he tied' (/ɗ̣/ = implosive + emphatic)

spread throughout the region, and today the western Sudanic dialect ('West Sudanic Arabic', as Owens terms it) is broadly coterminous with what the anthropologist Ulrich Braukämper terms the 'Baggara belt' (see Map 1). This is named after the *baggāra* cattle-herding culture of Kordofan and Darfur, although in Nigeria, Cameroon, and western Chad the term *baggāra* as a self-designation is not used. It has its eastern border well into the Sudan and its western in Nigeria. Its economic foundation was, and to a degree still is, based on nomadic cattle herding. The greater part of native Arab speakers in the region, though, are either farmers or urban dwellers. Geographically, it is still expanding. In Nigeria, for instance, nomadic Arabic groups have moved south, out of Bornu, and are now found in neighboring Adamawa State.

Pommerol (1997) estimates that there are 600,000 (10% of the population) native Arab speakers in Chad, while in Nigeria there are up to another 500,000. No estimates are available for Cameroon, although the number is probably not more than 100,000. In addition, Arabic is the main lingua franca of Chad.

The oldest attested linguistic evidence for West Sudanic Arabic is found in Koelle's (1854) word list (Owens 1993b:148ff.). Detailed works began appearing during colonial times, and work has continued to appear up to the present, so that despite many gaps, particularly in the dialectology and sociolinguistics of Cameroon and Chad, the area is fairly well described.

In the following, the main structural features of West Sudanic Arabic are summarized, mainly on the basis of the Arabic of Nigeria. Important dialectal or sociolectal variants are noted, as well as cultural aspects of the variety.

2. PHONOLOGY

2.1 *Consonants*

The inventory of the consonants is given in Table 1.

In general, there is a one-to-one correspondence between Old Arabic etymology and West Sudanic Arabic, allowing for familiar adjustments. /g/, for instance, corresponds to Old Arabic /q/. The correspondences in Table 2, which are specific for West Sudanic Arabic, should be noted, however.

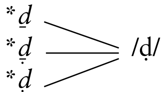
In addition, the following prominent mergers may be noted.

$$\begin{tikzcd}
{*d} \arrow[dr, dash] & \\
& /d/ \\
{*\underline{d}} \arrow[ur, dash] &
\end{tikzcd}$$

Table 3. Vowels of West Sudanic Arabic

short		*long*	
i	u	ī	ū
e	o	ē	ō
	a	ā	

Examples: *dalla* 'to descend, get down' (< Old Arabic *dallā* 'to cast down'), *axad* 'to take' (< *ʾaxaḏa*)



Examples: *ḍallal* 'to tame, domesticate' (< Old Arabic *ḏallala* 'to debase, humble'), *ḍahar* 'back' (< Old Arabic *ẓahr*), *muruḍ* 'he got sick' (< Old Arabic *mariḍa*)

The emphatic sounds listed are phonemic, representing phonemic splits, e.g. *aṃṃ* 'mother' vs. *ʿamm* or *ʿemm* 'father's uncle'; *kaṛṛa* 'to drag' vs. *karra* 'to tear'.

In addition, secondary emphaticization occurs, as in *gaṇaṣ* 'he hunted' (< *ganaṣ*). /c/ [tʃ] occurs in loanwords, frequently ideophones (*cu* ideophone for 'very red'), irregularly from /š/, e.g. *cakka* (cf. Old Arabic *šakka* 'to transfix, pierce') in words of unknown etymology, *calla* 'to hack off', and as a palatalization of /t/ in the Bagirmi dialect, *cāl* 'come'. /p/, /ny/, /ng/ are introduced via loanwords, *nyangūr* 'towel' < Mimi; Pommerol 1999b:1018), *pōlīs* 'police'.

Emphatic sounds are not reported for Abbéché by Roth-Laly, nor for Ndjamena by Pommerol. In rural Chad and Cameroon, however, they occur as listed above. For the /q/ ~ /x/ alternation, /x/ occurs especially in eastern Chad and Ndjamena, otherwise /q/. Eastern Chadian dialects around Amm Timan retain /ḥ/, /ʿ/, *aniḥna* 'we'. /j/ is phonetically affricate [ʤ] or palatal [ɟ]; these variants have not yet been systematically localized, although each variant has been noted in both Nigerian and Chadian Arabic. Other variants are noted in Section 8 below.

2.2 *Vowels*

The vowels of West Sudanic Arabic are given in Table 3.

Long vowels correspond in general to Old Arabic long vowels or diphthongs, e.g. *bāt* 'to spend the night', *bēt* 'house', *bīt* 'spend the night!', *gūl* 'say!', *gōl* 'expression'. Short /a/ corresponds generally to Old Arabic /a/. In Nigerian Arabic, there is no contrast between high /i/ and /u/. Nonetheless, the distribution of the two is not lexically predictable and will vary from village to village (see Owens 1998:39–42; e.g. *bitimm* ~ *butumm* 'he finishes'), and so can be given phonemic status. For Ndjamena, Pommerol (1999a:15) gives the minimal pair *jurr* 'pull!' vs. *jirr* 'fermented millet' (though *jīr* in Nigerian Arabic). The short mid vowels generally are introduced in loanwords, *cokol* 'spoon' < Hausa (Nigerian Arabic), although *-e* word-finally may appear as a conditioned variant of *-a* (*kabīr-e* 'big [fem.]').

In contrast to many dialects, low [a] and [aː] are stable low central vowels, lacking prominent front and back variants (e.g. in nonemphatic/emphatic contexts). [a] may be raised to [e] in the context of /l, r/, however, e.g. [leben] 'milk'.

The high vowels generally are in a harmonic relation. This is particularly evident in the preformative vowel of the imperfect verb (see Sec. 4.1), but, subject to factors such as speed of speech, it applies elsewhere as well, e.g. *b-u-rudd* 'he returns' vs. *b-i-timm* 'he finishes', *b-u-ktub* 'he writes' vs. *b-i-nzil* 'he descends'. The unmarked value is /i/, which will appear when no other conditioning context is present, e.g. *b-i-šarab* 'he drinks'.

2.3 *Syllable structure*

The most prominent facet of syllable structure are four different rules of epenthesis. In the following, the epenthetic vowel is underlined.

i. Guttural epenthesis (→ *gaḥawa*-syndrome), /h, ʾ, q, x/ + C → C_{gut} a C, e.g. *xaɗ̣ar* 'he left', *baxa̱ɗar* 'I leave', *ḥamra* 'red [fem.]', *aḥa̱mar* 'red [masc.]'
ii. Sonorant epenthesis (→ *bukaṛa*-syndrome), C + /r, l, n/ → Ci/uC_{son}, e.g. *tiji̱ri* 'she runs, *bimi̱la* 'he fills', *bugu̱nuṣ* 'he hunts'
iii. CCC epenthesis. A sequence of three consonants is broken up in one of two ways:
    a. Word-finally: C_1C_2-v-C_3 where $C_{1/2}$ are either stem consonants or stem-final C + suffix (verbal person marker, feminine

nominal suffix), C_3 a suffix, e.g. *limm* 'gather!', *limm-ha* 'gather it [fem.]!' > *limm-a̱-ha; darb* 'road', *darb-kan* > *darb-i̱-kan* 'your [pl.fem.] road'; *kammal*(*t*) 'I finished', *kammal-t-hum* > *kammalt-u̱-hum* 'I finished them [masc.]'. In this case, the epenthetic vowel is harmonic with the suffix: /a/ before *-ha*, /u/ before /u/, otherwise /i/.

b. Word-medially the breakup occurs between C_1 and C_2. Medial sequences of three consonants occur via vowel elision. Vowels are elided in verb and noun stems under various conditions, some of them lexically driven (i.e. lexically idiosyncratic; cf. Owens 1998:26), when a vowel-initial suffix leads to a final open syllable, e.g. *bu-rgud-an* > *bu-rgd-an, kibd-it-a* 'his liver' > *kibdt-a*. Such CCC sequences are then broken up by an epenthetic vowel between the first two consonants, e.g. *buru̱gdan*, *kibi̱dta* [kiˈbitːa].

Other phonological processes include devoicing of voiced stops and fricatives word-finally or before a voiceless C, e.g. *rizz* > *riss* 'rice', *kibidta* > *kibitta*, as in (b). Nasals + C are assimilated to a following consonant, e.g. *saŋgā* 'koranic school', pl. *sanāgi*.

2.4 *Stress*

Generally, stress is placed on the first heavy sequence (-vCC, -v̄C) from the end of the word, otherwise on the first syllable, except that primary stress cannot be more than three syllables from the end (in the following, determining heavy sequence underlined), e.g. *ˈkatab* 'he wrote', *kataˈbōha* 'they [masc.] wrote it [fem.]', *kataˈbanha* 'they [fem.] wrote it [fem.]', *biˈdanguru* 'they bend down'.

In the case of sequences of four syllables or more, the last three of which are open, the penultimate bears primary stress, e.g. *kataˈbata* 'she wrote it', *kammaˈlata* 'she finished it'.

Epenthetic vowels may be stressed, e.g. *ˈtatin* 'you [sg. masc.] fold', vs. *taˈtinha* 'you [sg. masc.] fold it [fem.]'.

3. MORPHOLOGY

3.1 *Nominals*

Nouns can be inflected for the following categories: number (singular, dual, plural, singulative), gender (masculine, feminine), definiteness (definite, indefinite, nonspecific linkage), possessive (possessive pronoun). In all respects, these categories are those familiar in all varieties of Arabic, except for the restricted nature of the dual and the linker particle.

3.1.1 Morphological realizations

Plural is marked either by broken plurals or sound, suffix plurals. The sound plurals are masculine *-īn*, e.g. *muslim-īn* 'Muslims [masc.]', feminine *-āt*, e.g. *muslim-āt* 'Muslims [fem.]'. Looking through all the dialects, about twelve frequently used broken plurals are attested in the region, e.g. *buyūt* 'houses', *kibār* 'big [pl.]', *hamīr* 'donkeys', *kubšān* 'rams', *karāde* 'pagans' (sg. *kirdi*). Two which are common to the region are attested only rarely outside it. The first, $C_1aC_2aC_3a$, is common among singulars with C_2 or C_3 = *w*, *y*, or in certain nouns with geminate $C_{2/3}$, e.g. *fayala* 'elephants' (sg. *fīl*), *jarawa* 'puppies' (sg. *juru*), *ṣaṃaṃa* 'names' (sg. *ṣuṃṃ*), and the other occasionally with CvCC*e* nouns, e.g. *rikabb* 'knees' (sg. *rikbe*), *guraʾʾ* 'gourds' (sg. *girʾe*). In addition, idiosyncratic plurals occur, e.g. *abbahāt* 'fathers' (sg. *abu*). The dual is largely restricted to extents of time (*dōrt-ēn* 'two weeks', *šahar-ēn* 'two months'), even it can be successfully elicited for virtually any count noun. The singulative is marked by the feminine suffix *-a, -e* (or *-ā, -āya*) and is lexically restricted: *qanam-a* 'one sheep' (*qanaˈmā, qanaˈmāya*), *šadar-a* 'one tree'.

Gender is masculine or feminine. Feminine nouns are usually marked by *-a, -e* or *-ā, -āy*, although some unmarked nouns are feminine, such as *baɗun* 'stomach'. For the suffix form, the choice of *-a, -e* is phonologically conditioned, *-e* after a high vowel, *-a* after a low (examples above). The choice between *-a, -e* or *-ā, -āy* is lexically determined on an irregular basis (e.g. the singulative of *baṣal* 'onions' is *baṣal-a* or *baṣal-āya* according to speaker or dialect). The choice between *-ā* or *-āy* is dialectal or sociolectal (*baṣal-ā* also occurs). In all cases, the feminine suffix has an *ʾiḍāfa* variant with -v*t*: *rikb-it-ha* 'her knee', *qanam-āy-it-ha* 'her sheep'.

Definiteness is marked by *al-*, which assimilates to a dental or alveolar consonant. Indefiniteness is unmarked. A contextually determined nonspecific linker suffix *-an* may be added to an indefinite noun modified by an adjective, number, or relative clause, *rājil* *abu watīr-an hamra* ‘a man with a red car’. The indefinite form is used in first mention, the definite in previous mention or for generic statements.

The possessive pronouns are suffixed to nouns and are identical to object pronouns, except for the 1st person singular. The 1st person singular suffix is stressed. Forms with initial *-h* often drop the *h* phonetically. They are presented in Table 4 suffixed to the noun *buyūt* ‘houses’. Note that plural feminine is a morphological category.

Table 4. Possessive suffixes in West Sudanic Arabic

	singular	plural
3rd masc.	*buyūt-a*	*buyūt-hum*
fem.	*buyūt-ha*	*buyūt-hin*
2nd masc.	*buyūt-ak*	*buyūt-ku*
fem.	*buyūt-ki*	*buyūt-kan*
1st	*buyūˈt-i*	*buyūt-na*

3.1.2 Modifying categories

Adjectives modify nouns and agree with the noun in number, gender, and definiteness. They have elative forms in many of the areas, e.g. *akˈbar* ‘bigger’ or *ˈakbar*. In some Chadian dialects, comparison is with simple adjective + *min*, *kabīr min* ‘bigger than’.

Demonstratives follow the noun and also agree with it in number and gender, e.g. *albuyūt alkibār dēl* ‘these big [pl.] houses’, *watīr kabīr-e* ‘a big car’. The complete demonstrative paradigm is given in Table 5, with near and distal forms. C-final demonstratives may be followed by a final *-a*, according to conditions still to be defined.

In Chad, *dilāk* is frequently used as a common plural distal.

Numerals follow the noun. Count and non-count forms are identical, *wāhid* (fem. *wāhde*), *tinēn, talāta, arbaʾ, xamsa, sitte, sabʾa, ṭaṃāṇe, tisʾe, ʾašara*. The numerals 11–19 are formed on the pattern ‘10 (and) unit’, *ʾašara* (*haw*) *wāhid* ‘11’, *ʾašara* (*haw*) *tinēn* etc., or they are formed from old compounds, *ihdāšar* etc. The numeral *wāhid* also functions as an article (both specific and nonspecific), ‘a, a certain’; the plural forms *wād-īn, wād-āt* are used for ‘some, certain ones, some…others’.

Table 5. Demonstratives in West Sudanic Arabic

	near		distal	
	singular	plural	singular	plural
masc.	*da*	*dōl*	*ḍāk*	*ḍōlak*
fem.	*di*	*dēl*	*ḍīke*	*ḍēlak*

3.2 Verbs

3.2.1 Inflection

The inflectional elements of the perfect, imperfect, and imperative can be illustrated with the verb *katab* (Table 6).

The perfect inflections are quite uniform throughout the West Sudanic Arabic area. The 1st and 2nd person singular masculine has the form *kaˈtab-t-* before object suffixes, e.g. *katab-t-a* ‘he wrote it’, before nominal objects beginning with the definite article, e.g. *katab-t almaktūb* ‘I wrote the letter’ (cf. *kaˈtab maktūb-a* ‘I wrote his letter’), and after conjugations in *-ē*, e.g. *badē-t* ‘I began’ (see below). The 3rd person singular feminine has allomorphs *-it* and *-t*, before vowel-initial suffixes in II*w/y* and III*y* verbs, respectively, e.g. *jāb-t-a* ‘she brought it’. In the imperfect, many dialects have a high vowel *i/u* as preformative vowel, e.g. *t-u-ktub* ‘you/she writes’. Many Chadian dialects (for instance some in Ndjamena) lack *b-* altogether, e.g. *yaktub* ‘he writes’, and in those dialects with *b-*, it functions as an indicative, or as a noncontrol marker: if the action is out of the control of the actor, *b-* does not occur, as in the contrast *gūl leya ij-i/gūl leya b-ij-u* ‘tell him to come!/tell him that they are coming!’. In Chad, the 1st person singular is *n-*, 1st plural *n-…-u*, e.g. *n-uktub, n-uktub-u*. The imperfect v-initial suffixes are subject to harmony, mid-variants occurring after low stem vowels, *ta-šarb-e, ta-šarb-o* ‘you [sg. fem.] drink/you [pl. masc.] drink’. The plural feminine suffix doubles before a v-initial suffix, *katab-ann-a* ‘they [fem.] wrote it [masc.]’.

High and low vowel conjugations are maintained throughout the region, e.g. *libis* ‘he dressed’, *b-ilbas* ‘he dresses’.

Among the basic verbs, II*w/y* and III*y* occur, each with conjugations dependent on the vowel quality. Examples before -C and v-initial suffixes are given in Table 7.

III*y* verbs drop a final *-i* in the 2nd person singular masculine, so that in this conjugation

Table 6. Conjugation of the verb *katab* 'to write'

	perfect singular	plural	imperfect singular	plural	imperative	
3rd masc.	*'katab*	*katab-o*	*b-u-ktub*	*b-u-ktub-u*		
fem.	*katab-at*	*katab-an*	*t-a-ktub*	*b-u-ktub-an*		
2nd masc.	*ka'tab*	*katab-tu*	*t-a-ktub*	*t-a-ktub-u*	*a-ktub*	*a-ktub-u*
fem.	*ka'tab-ti*	*katab-tan*	*t-a-ktub-i*	*t-a-ktub-an*	*a-ktub-i*	*a-ktub-an*
1st	*ka'tab*	*katab-na*	*b-a-ktub*	*n-a-ktub*		

Table 7. Verbs II*w/y* and III*y*

	perf.	imperf.	perf.	imperf.	perf.	imperf.
II*w/y*		'to say'		'to carry'		'to fear'
1st sg.	*gul*	*b-a-gūl*	*šil*	*b-a-šīl*	*xif*	*b-a-xāf*
3rd sg. fem.	*gāl-at*	*t-a-gūl*	*šāl-at*	*t-a-šīl*	*xāf-at*	*t-a-xāf*
III*y*		'to buy'		'to forget'		
	perf.	imperf.	perf.	imperf.		
1st sg.	*šar-ēt*	*b-a-širi*	*nisī-t*	*b-a-nsa*		
3rd sg. fem.	*šar-at*	*t-a-širi*	*nis-at* ~ *nisiy-at*	*t-a-nsa*		

a contrast exists between 2nd person singular masculine vs. 3rd person singular feminine, e.g. *tamiš* 'you [masc.] go' vs. *tamši* 'she goes'. This applies to derived verbs as well, e.g. *taxall* 'you allow' vs. *taxalli* 'she allows'. Geminated verbs have 3rd person singular masculine forms in final *-a*, e.g. *tamm-ēt* 'I finished', *tamma* 'he finished', *ba-timm* 'I finish'.

I*w/y* verbs are rare. Most shift to Form II in the imperfect, *warad/bawarrid* 'to water cattle', although there are a few so idiosyncratic that they must be listed separately, e.g. *'wagaf/b-a-gīf* 'he stopped/I stop', *wagēt/b-a-ga* 'I fell/I fall'. Other verbs that can be noted are *jīt/b-a-ji* 'I came, I come', *ga''ad, b-a-gōd* 'I stood, I stand', *a'kal, b-ākul* 'I ate, I eat' (3rd pers. sg. masc. *bi-yākul* 'he eats') and *sawwa*, which on a variational basis may lose its *ww* altogether or have it replaced by *y*: *b-a-sawwi* ~ *b-a-seyy*, *saww-ēt* ~ *sayy-ēt* 'I do, I did'.

3.2.2 Derivation

Productive derived classes are Form II, which has the same form as the class of quadriliterals, reduplicated verbs, Form III, Forms V and VI with stative prefix *al-*, and VII. Form II is most often causative or denominative, V and VI stative, VI reciprocal, and VII usually passive. Reduplicates are distributive or frequentive. Examples: *tammam/bitammim* 'to finish something', *baddal* 'to exchange', *maskan/bimaskin* 'to make poor', *tamtam/bitamtim* 'to finish many', *bādal/bi-bādil* 'to interact', *albaddal/b-ilbaddal* 'to be exchanged', *almaskan/b-ilmaskan* 'to become poor', *albādal/b-ilbādal* 'to exchange with each other', *anbadal/b-inbadil* 'to be exchanged'. Pommerol (1999b:176) also gives Form VIII, *n-ijtahid* 'I exert' and X, *n-istaxrab* 'I am surprised', although Roth-Laly (1969) notes that Form VIII is rare and she does not list Form X; Tourneux and Zeltner (1986) do not list either one.

Derived participles are the active, as *kātib* 'having written, *tāmm* 'having finished', derived with *mi-*, *mi-tammim* 'having finished', *mi-baddil* 'having been exchanged', etc. The passive is maintained formally only in basic verbs, *maktūb* 'written', *mišri* 'bought'. In derived verbs, it falls together with the active, *mitammim* 'having been finished'.

Three frequent verbal nouns of basic verbs end in -īn, *katb-īn* 'writing', or *-ān, gadam-ān* 'becoming old', or have the pattern C*a*CC, *daxil* 'entering'. In Form II, there is an ablaut form with prefix *ti*-, e.g. *tibiddil* 'exchanging', *timiskin* 'becoming poor'.

3.2.3 Object suffix

The object suffixes are as for possessed nouns (see Sec. 3.1.1 above), except that the 1st person singular is *-ni* (in Chad, stressed *-ˈni*), e.g. *šāf-ni* 'he saw me'. In many Chadian varieties, an object suffix added to the 3rd person singular masculine perfect verb induces the insertion of *-ā-*, e.g. *kammal* 'he finished', *kammal-ā-hum* 'he finished them [masc.]'. An indirect object is marked by the preposition *le*, which does not cliticize to the verb, e.g. *šarēt lē-hum* 'I bought for them'.

3.3 *Pronouns and other word classes*

3.3.1 Pronouns

The paradigm for independent pronouns is given in Table 8.

The reflexive pronoun is *ḍumma* (always possessed) in Nigerian Arabic; *nafis* is found throughout the region.

The relative pronoun is *al-*, identical in form to the definite article, *hu da al-gabul xadamna beya sei* 'it's the one which we worked with before, right?'. The genitive marker is *hana*, *hīl, hinē*, or *hintāt* (as alternative to *ʾiḍāfa*), e.g. *alxidime hīl alhakūma* 'the government's work'. The *ʾiḍāfa* genitive juxtaposes possessed and possessor, as in *qanamāyit jāˈri* 'my neighbor's sheep'.

3.3.2 Other word classes

Question words: *miˈne, yaˈtu, yaˈti* 'who?'; *šuˈnu* 'what?'; *wēn, yēn* 'where?'; *mata, mitēn, waqtēš* 'when?'; *maˈla* 'why?'; *kam* 'how many?'; *yēˈnu, yēˈni* 'which?'; *kēf, kikkēf, kefkef* 'how?'. A yes-no question marker is *-wa* or *-a* cliticized to the end of the questioned sentence, *mašēt-a* 'did you go?'.

Table 8. Independent pronouns in West Sudanic Arabic

		singular	plural
3rd	masc.	*hu*	*hum*
	fem.	*hi*	*hinna*
2nd	masc.	*inta*	*intu*
	fem.	*inti*	*intan*
1st		*ana*	*anīna*

Common adverbs are *hinēn, hine* 'here'; *hinăk, qādi* 'there'; *amis* 'yesterday'; *alōm* 'today'; *ambākir* 'tomorrow'; *bukura* 'day after tomorrow'; *hassa, duggut* 'now'; *tawa* 'before'; *bilhēn* 'very'. Prepositions include *le* 'for'; *be* 'by means of'; *giddām* 'in front'; *wara* 'behind, after'; *ʾind* 'at'; *fōg* 'at, on' (largely replacing *fi* found in other dialects); *maʾa, miya* 'with'; and *ambēn* 'between'. Subordinating conjunctions include *kan* 'if'; *waqit, sāʾit* 'when'; *awān, lamma, damma(n), namma(n)* 'when'; *min* 'since'; *kadar, hadar, xadar le* 'before'; *ašān, mišān, lenjān* 'because'. Note that there is a good deal of lexical variability in these forms.

The negation marker is *ma* placed before the verb, e.g. *ma ligīta* 'I didn't get it', a negative imperative with *yā*, e.g. *yā taj gede* 'don't come again!'. There is a large class of ideophones used in various functions (see Owens and Hassan 2004); an example is *kurok*, in (3) below.

A presentative is formed by the construction demonstrative + *le-* pronoun (+ noun), as in (1).

(1) *ḍōlaka* *le-hum* *r-rijāl*
those to-them Def-men
'There are the men!'

4. BASIC SYNTAX

The usual sentence order is SVO or SP(redicate) for nominal sentences, as in (2)–(4).

(2) *kūre* *al-banāt* *bi-gīf-an* *fi*
formerly Art-girls Ind-stand-3fp in
l-mada *b-ilʾab-an*
Art-field Ind-dance-3fp
'In former days the girls would stand in the field and dance'

(3) *aš-šam* *kurok* *kan* *wag-at*
Art-sun Ideophone if fell-3fs
dugo *buxarm-an*
then come.out-3fp
'When the sun's set 'kurok', then they [the cattle] go out'.

(4) *al-bagar-a* *marīḍ-a*
Art-cow-fs sick-fs
'The cow is sick'

For presentatives with *ʾind* 'have', *fi* 'there is', *māfi* 'there is not', as predicted, the sequence is usually Predicate-Subject, this sequence being determined by pragmatic factors; subjects here are overwhelmingly indefinite, and indefinite nominals in general (e.g. including objects) are in the majority of cases indefinite, as in (6).

(6) *'fi* *nādim* *b-idōr*
exist person Ind-want.3ms
bi-kallim-ak
Ind-talk3ms-you.2ms
'There is a man who wants to talk to you'

5. DIALECTS

As mentioned in the introduction, the entire West Sudanic Arabic is relatively homogeneous and characterized by very specific isoglosses setting it off from all or most other Arabic dialects. These include the following: **ḏ* > *ḍ*, e.g. *ḍahab* 'gold'; the particular reflex of the 1st and 2nd person singular masculine suffix in the imperfect paradigm (see above); and *al-* as Form V and VII derived verb marker (see Sec. 4.2).

There is one major dialect boundary in the West Sudanic Arabic area, running roughly along an east-west line, beginning in the area of Kala, Banki in the west, extending through Cameroon and into Chad (see Map 1). The eastern boundary is unknown. This dialect region lies in the sociohistorical region known as Bagirmi, and Owens therefore has termed it the Bagirmi dialect. The other dialect, which covers a wider area, is left unnamed. The major differences between the Bagirmi and non-Bagirmi dialects are set out in Table 9. An excellent description of the Bagirmi dialect is Tourneux and Zeltner (1986).

In Table 9, (iii) is a special case of (ii), although in the Bagirmi dialect this leads to basic verbs with an Old Arabic pharyngeal as initial consonant moving into the derived Form III (e.g. *ba-tērib* 'I cultivate'). Note that the lengthened vowel is historically an epenthetic vowel of the *gahawa* class (see Sec. 2.3 above), which indicates that the ancestral population had the original pharyngeal and the conditioned shift to /e/ when they moved into the region, the merger of the pharyngeals with the glottals being a subsequent development. Not all of the features occur in concert, although the dialectal situation is well documented only for Nigeria (Owens 1998: Chap. 5). In Nigeria, (i)-(iii) and (vii) occur in the triangle between Kala-Gulumba-Banki, (iv) and (vi) are restricted to the Banki area, and (v) occurs in the Gulumba-Banki region, not in Kala, but also extends westward outside of the Bagirmi dialect region toward Dikwa.

Salient dialect differences coinciding with national boundaries are rare, although two do tend to differentiate Nigeria from Chad (with Cameroon underresearched). As noted above, Chad marks the 1st person singular imperfect with *n-*, *n-...-u*, and the 1st person singular object suffix on verbs is stressed. In Nigeria, only the 1st person singular possessor suffix is stressed. Another parameter of difference, as yet not explored systematically on a comparative basis, occurs in the vocabulary, where Chadian Arabic is more strongly influenced by French and perhaps by Sudanese Arabic, and Nigerian Arabic by English, Hausa, and Kanuri (Owens 2000b).

Throughout the region, other prominent differences are found, some of which can be men-

Table 9. Differences between Bagirmi and non-Bagirmi dialects

	Non-Bagirmi	Bagirmi	Non-Bagirmi	Bagirmi
i.	'CvCv(C)	Cv'Cv(C)	*'bana*	*ba'na* 'he built'
ii.	**ḥ, ʿa > h, ʾa*	*h, ʾe*	*'laham/'baʾar*	*le'hem/be'ʾer* 'meat/manure'
iii.	**-ḥ, ʿa*	*-hē, ʾē* in imperfect	*bahalim*	*bahēlim* 'I dream'
iv.	**ṯ > t*	*s*	*taḷāta*	*saḷāsa* 'three'
v.	*-hin*	*-han*	*bēt-hin*	*bēt-han* 'their [pl. fem.] house'
vi.	*-a*	*-e* after front vowel	*bēt-a*	*bēt-e* 'his house'
vii.	AP + suffix	AP + *in* + Suffix	*kātib-ha*	*kātb-in-ha* 'has written it'

tioned here. Urban Chadian Arabic (Ndjamena, as described by Pommerol 1999; Abbéché, as described by Roth 1979) appears to have undergone a number of simplificatory processes: emphatic consonants are lost, the plural feminine may be lost as a morphological category, and, in Abbéché, the two phonologically conditioned suffixes of the imperfect are collapsed in *-i* (sg. fem.), *-u* (pl.).

There is a range of features not found at all in Nigeria or Cameroon but attested in various places in Chad. As noted in Owens (1993b:134–135), however, these features do not meld into coherent dialect areas, at least given our current dialect knowledge of Chad. They include the following, all of which are attested in the triangle formed by Amm Timan, Abbéché, and Atia, although they are not necessarily restricted to this region, nor do all varieties in this triangle have all of the features.

- *ġ* > *x* (see Sec. 2.1)
- *ḥ*, ʿ maintained (see Sec. 2.1)
- *nt* > *tt* in the 2nd person pronoun *itta* 'you [sg. masc.]
- *-āy* singulative very common (see Sec. 3.1)
- intrusive *-ā-* in perfect verb (see Sec. 4.3)
- short low vowel raising: C*a*-Cv → C*i*/*u*-Cv; combined with epenthetic rule (Sec. 2.3 iii) above, this produces chains as follows: *taktub* 'you [sg. masc.] write', *taktub-u* 'you [pl. masc.] write' (via deletion, epenthesis) > *ta'kudbu* > (raising of preformative vowel in open syllable) *tu'kudbu*.

6. SOCIOLINGUISTIC AND OTHER STUDIES

Pommerol (1997) summarizes the national language debate in Chad. The countrywide consensus appears to be that Chadian Arabic is well suited to be a national language, but not Standard Arabic. Pommerol notes that only 10 percent of the Chadian population are native Arabic speakers but almost 40 percent of the total population speak some form of Arabic.

Owens (1995, 1998, 1999) presents a detailed variationist account of Nigerian Arabic in Maiduguri, based on the examination of twelve linguistic variables. He shows that the spoken Arabic of the city is not developing in the direction of a normalized variety or a koine, arguing instead that the ancestral founding principle, coupled with neighborhood contact effects, is the major variable explaining the observed variation. He suggests that the status of the language as a minority one inhibits the spread of wider norms. Given this, and against the background of urban immigration, dialectal heterogeneity is currently increasing in the city.

Recently, Owens (2005b) has investigated the multilingual repertoire of Maiduguri Arabs, describing a hierarchicalized multilingualism, whereby Nigerian Arabic and Hausa serve to structure bilingual discourse, with English and Standard Arabic providing a lexical reserve. In Owens (2002, 2005a), the same multilingual corpus is analyzed for the occurrence of Ø morphological forms, described in terms of degree of iconicity and related to psycholinguistic processing issues. Multilingualism is also implicitly an issue in Owens (1996a), where he describes an extreme degree of semantic convergence of Nigerian Arabic toward what he terms Lake Chad area idioms. The expression *rās albēt* 'roof [lit. 'head of the house']', in most Arabic dialects expressed by a single lexeme (e.g. Cairene Arabic *saṭḥ, sa'f*), is explained as a calque on Kanuri *kəla fato-be* 'head house-of', but also a concept expressed in the same form in Kotoko, Glavda, Fulfulde, Bura, and other languages of the area.

There are a number of lexica of dialects in the region (Lethem 1920; Roth-Laly 1969; Kaye 1982). However, special mention needs to be made of Pommerol (1999b), a 1,640-page tour de force for Chadian (particularly Ndjamena) Arabic.

7. POPULAR CULTURE

The Nigerian Arabs have a number of traditional dances and songs, most of which are also found in Ndjamena. These are mainly performed live, with little in the way of recorded media. Dances accompanied by singing include *korokoro, bazaga, cardi* (dialectal variants), and *bāra, bandīl* and *am-dagura fanne* have singers accompanied by drummers; *šilašila* is a flute, usually performed by three musicians; and *bandala* and *hūdu* have a singer accompanied by percussion gourds and cowries. The *qalab alxēl* or *kāro* are praise songs sung by women at horse-riding tournaments. The music may sometimes be performed by standing groups, who are paid to perform at various social functions such as circumcisions, naming ceremonies, weddings, and religious holidays.

BIBLIOGRAPHICAL REFERENCES

Alzubair, Ibrahim Hussain. 1999. *The use of Shuwa Arabic in Maiduguri.* B.A. thesis, University of Maiduguri.

Bassi, Ruth J. 1998. *Language shift and management among the Lebanese families in Kano.* B.A. thesis, University of Maiduguri.

Braukämper, Ulrich. 1993. "Notes on the origin of Baggara Arab culture with special reference to the Shuwa". *Sprache und Geschichte in Afrika* 14.13–46.

Carbou, Henri. 1913. *Méthode pratique pour l'étude de l'arabe parlé au Ouaday et à l'est du Tchad.* Paris: Geuthner.

Décobert, Christian. 1985. *Phonologie arabe du Tchad.* Paris: Geuthner.

Goni, Fati Abba. 2003. *The structure of the noun phrase in Nigerian Arabic.* B.A. thesis, University of Maiduguri.

Hagège, Claude. 1973. *Profil d'un parler arabe du Tchad.* Paris: Geuthner.

Hamid, Ali Tahir. 1997. *Reduplication processes in Nigerian Arabic.* B.A. thesis, University of Maiduguri.

Hassan, Jidda. 1995. *Phonological variation among middle-aged Shuwa Arabs: A network analysis.* M.A. thesis, University of Maiduguri.

——. 2004. *An analysis of codeswitching among Nigerian Arabs.* Ph.D. diss., University of Maiduguri.

Howard, C.G. 1921. *Shuwa Arabic stories, with an introduction and vocabulary.* Oxford: Oxford University Press.

Ilyas, Mohammad Umar. 1996. *Deletion and epenthesis rule in Nigerian Arabic.* B.A. thesis, University of Maiduguri.

Jidda, Ali Dawud. 2003. *The morphology of nominals in Nigerian Arabic.* B.A. thesis, University of Maiduguri.

Kaye, Alan. 1975. *Chadian and Sudanese Arabic in the light of comparative Arabic dialectology.* The Hague and Paris: Mouton.

——. 1982. *A dictionary of Nigerian Arabic.* Malibu: Undena.

Koelle, Sigismund W. 1854. *Polyglotta africana.* London: Church Missionary Society.

Lethem, G. 1920. *Colloquial Arabic: Shuwa dialect of Borno, Nigeria and of the Lake Chad region.* London: Crown Agent for the Colonies.

Mohammad, Abba Mohammad. 1998. *Assimilation processes in Nigerian Arabic.* B.A. thesis, University of Maiduguri.

Musa, Jibril Musa. 1997. *Ideophones in Nigerian Arabic.* B.A. thesis, University of Maiduguri.

Owens, Jonathan. 1985. "Arabic dialects of Chad and Nigeria". *Zeitschrift für Arabische Linguistik* 14.45–61.

——. 1993a. *A grammar of Nigerian Arabic.* Wiesbaden: O. Harrassowitz.

——. 1993b. "Nigerian Arabic in comparative perspective". *Sprache und Geschichte in Afrika* 14.85–176.

——. 1995. "Minority languages and urban norms: A case study". *Linguistics* 33.305–358.

——. 1996a. "Grammatisierung, Semantisierung und Sprachkontakt: Arabisch im Tschad-See-Gebiet". *Sprachtypologie und Universalienforschung* 49.79–85.

——. 1996b. "Idiomatic structure and the theory of genetic relationship". *Diachronica* 13.283–318.

——. 1997. "Nigerian Arabic (Shuwa Arabic)". *Borno Museum Society Newsletter.*

——. 1998. *Neighborhood and ancestry: Variation in the spoken Arabic of Maiduguri (Nigeria).* Amsterdam and Philadelphia: J. Benjamins.

——. 1999. "Uniformity and discontinuity: Toward a characterization of speech communities". *Linguistics* 37.663–698.

——. (ed.) 2000a. *Arabic as a minority language.* Berlin: de Gruyter.

——. 2000b. "Loanwords in Nigerian Arabic: A quantitative approach". Owens (2000a:259–346).

——. 2002. "Processing the world piece by piece: Iconicity, lexical insertion and possessives in Nigerian Arabic codeswitching". *Language Variation and Change* 14.173–209.

——. 2005a. "Ø forms and lexical insertion in codeswitching". *Bilingualism: Language and Cognition* 8.23–38.

——. 2005b. "Hierarchicalized matrices: Codeswitching among urban Nigerian Arabs". *Linguistics* 43.957–994.

——. and Jidda Hassan. 2000. "Making a fish of a friend. Waris: the secret language of Arabic Koran school students in Borno". Owens (2000a:221–258).

——. and Jidda Hassan. 2004. "Remarks on ideophones in Nigerian Arabic". *Approaches to Arabic dialects: A collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday*, ed. Martine Haak, Rudolf de Jong, and Kees Versteegh, 207–220. Leiden: E.J. Brill.

Pommerol, Julien de. 1990. Da Hayyin*: Méthode d'arabe véhiculaire parlé au Tchad.* 4 vols. Ndjamena: CEFOD.

——. 1997. *L'arabe tchadien: Emergence d'une langue véhiculaire.* Paris: Karthala.

——. 1999a. *Dictionnaire arabe tchadien – français.* Paris: Karthala.

——. 1999b. *Grammaire pratique de l'arabe tchadien.* Paris: Karthala.

Roth [-Laly], Arlette. 1969. *Lexique des parlers arabes tschado-soudanais.* Paris: Editions CNRS.

——. 1971. "Esquisse de la phonologie du parler arabe d'Abbéché". *Comptes rendus du GLECS* 16.33–79.

——. 1979. *Esquisse grammaticale du parler arabe d'Abbéché.* Paris: Geuthner.

Tourneux, Henri and Jean-Claude Zeltner. 1986. *L'arabe dans le bassin du Tchad.* Paris: Karthala.

Umar, Mohammad Abba. 2002. *The morphology of verbs in Nigerian Arabic: A case study of the Salamat dialect.* M.A. thesis, University of Maiduguri.

Worbe, André. 1962. *Etude de l'arabe parlé au Tchad.* (mimeo).

Yunus, Jibril Yunus. 1998. *Language use among the Arabic speaking families in Maiduguri: A case study of Yemenis, Lebanese and Fezzanis.* B.A. thesis, University of Maiduguri.

Zeltner, Jean-Claude and Maurice Fournier. 1971. "Notice pour suivre un enregistrement en arabe

Salamt de la region du Lac Tchad". Fort Lamy (mimeo).

JONATHAN OWENS (University of Bayreuth and University of Maryland) and JIDDA HASSAN (Maiduguri)

WH-Movement

WH-movement plays a key role in the syntax of long-distance dependencies. Typically, it is involved in the derivation of *wh*-interrogatives, but it is also involved in the formation of other constructions, such as topicalized constructions (→ topicalization) and → relative clauses. This type of movement affects the maximal projections of many syntactic categories and phrasal constituents; noun phrases, prepositional phrases, adjective phrases, and adverb phrases are its primary targets. In a seminal paper on the topic, Chomsky (1977) outlines the properties that characterize WH-movement as follows: (i) the COMP position is involved; (ii) it leaves a gap; (iii) it is unbounded; and (iv) it is constrained by Subjacency. These characteristics have become the diagnostics according to which a given type of displacement can be identified as WH-movement.

1. WH-MOVEMENT IN *WH*-INTERROGATIVES

An identifying characteristic of *wh*-interrogatives in Modern Standard Arabic is that the *wh*-constituent must appear in sentence-initial position if the sentence is simplex, as illustrated in (1).

(1a) *man* *ḍaraba* *ʿaliyy-un*
who hit.3ms Ali-Nom
'Who did Ali hit?'
(1b) *ʾayna* *taraka* *salīm-un*
where left.3ms Salim-Nom
miḥfaḏ̣at-a-hu
wallet-Acc-his
'Where did Salim leave his wallet?'

WH-movement in Modern Standard Arabic applies by preposing a *wh*-phrase (2a, b) or a constituent containing the *wh*-phrase (2c) to the sentence-initial position, generally leaving a gap in the base/original position within the sentence.

(2a) *ʾayna* *zaʿamta* *ʾanna* *ʿaliyy-an*
where claimed.2ms that Ali-Acc
qābala *fāṭimat-a*
met.3ms Fatima-Acc
'Where did you claim Ali met Fatima?'
(2b) *man* *qatalta?*
who killed.2ms
'Who did you kill?' (Classical Arabic, Wright 1862:223)
(2c) *maʿa* *man* *takallamta*
with who spoke.2ms
ʿan *al-qaḍiyyat-i*
about the-affair-Gen
l-muʿaqqadat-i *hāḍihi*
the-complicated-Gen this.fs
'With whom did you speak about this complicated affair?'

In indirect questions, the *wh*-constituent must appear in clause-initial position, as illustrated in (3).

(3) *nasītu* *ʾayna* *qābala*
forgot.1s where met.3ms
ʿaliyy-un *fāṭimat-a*
Ali-Nom Fatima-Acc
'I forgot where Ali met Fatima'

It can be argued that the landing site of the *wh*-phrase or the constituent containing it is the specifier of the complementizer phrase, [Spec CP], the highest functional projection dominating a clause. Thus, in WH-questions, the fronted *wh*-phrase (or constituent containing it) is in complementary distribution with the complementizers *ʾinna* and *ʾanna* (4)–(5).

(4a) *ʾinna* *l-fatāt-a* *kānat* *fī*
Comp the-girl-Acc was.3fs in
l-bayt-i
the-house-Gen
'The girl was in the house'
(4b) (**ʾinna*) *man* (**ʾinna*) *kāna*
(*Comp) who (*Comp) was.3ms
fī *l-bayt-i*
in the-house-Gen
'Who was in the house?'

(5a) *taḏakkartuʾanna*
remembered.1s that
l-mubār-āt-a *kānat*
the-match-fs-Acc was.3fs
ḥāsim-at-an
decisive-fs-Acc
'I remembered that the match was decisive'

(5b) *taḏakkartu* (**ʾanna*) *kayfa*
remembered.1s (*that) how
(**ʾanna*) *kānat* *l-mubār-āt-u*
(*that) was.3fs the-match-fs-Nom
'I remembered how the match was'

This complementarity between the *wh*-phrase and the complementizer has been attributed to the Doubly Filled COMP filter in syntax (van Riemsdijk and Williams 1986), which seems to operate in Modern Standard Arabic, as it does in many other languages.

Modern Standard Arabic *wh*-interrogatives have an additional requirement that the *wh*-constituent must also be immediately followed by the verb, a verb-second requirement, as shown in (6) (Bakir 1980).

(6a) *matā* *ištarā* *ˁaliyy-un*
when bought.3ms Ali-Nom
al-kitāb-a
the-book-Acc
'When did Ali buy the book?'
(6b) **?matā* *ˁaliyy-un* *ištara*
when Ali-Nom bought.3ms
l-kitāb-a
the-book-Acc
'When did Ali buy the book?'

Sentences with multiple *wh*-words exhibit a parallel requirement. That is, only one *wh*-constituent is preposed (7), and it must immediately precede the verb (8).

(7a) *man* *ḏahaba* *maˁa* *man*
who left.3ms with who
'Who left with whom?'
(7b) **man* *maˁa* *man* *ḏahaba*
who with who left.3ms
'Who left with whom?'

(8a) *man* *ˁarrafat* *hind-un*
who introduced.3fs Hind-Nom
ˁalā *man*
on who
'Who did Hind introduce to whom?
(8b) **man* *hind-un* *ˁarrafat*
who Hind-Nom introduced.3fs
ˁalā *man*
on who
'Who did Hind introduce to whom?'

The contrast between (6a) and (6b), on one hand, and (8a) and (8b), on the other, can be taken as a further indication that in Arabic, *wh*-phrases that are displaced occupy a position in the CP domain. This is so under the assumption that, in verb-second constructions, the verb marks the left edge of the Inflection Phrase (IP), immediately dominated by CP.

The adjacency requirement holding between *wh*-phrases and the verb in Modern Standard Arabic is relaxed in various Arabic dialects. For example, both Egyptian Arabic and Lebanese Arabic allow the subject to intervene between the *wh*-constituent and the verb (9).

(9a) *maˁa* *mīn* *mona* *xaragit*
with who Mona went.3fs
'With whom did Mona leave?' (Egyptian Arabic)
(9b) *ʾayya* *walad* *zeina* *šēfit*
which child Zeina saw.3fs
'Which child did Zeina see?' (Lebanese Arabic)

Even more interesting is the fact that various Arabic dialects diverge from Modern Standard Arabic with respect to the fronting of the *wh*-constituent. In Lebanese Arabic, for instance, WH-movement alternates with the in-situ strategy, whereby the *wh*-constituent remains in the base/original position inside the sentence, as illustrated by the acceptability of (10a), as well as (10b) from Lebanese Arabic.

(10a) *ˁa-mīn* *tˁarraf* *ˁali*
on-who met.3sm Ali
'Who did Ali meet?'
(10b) *tˁarraf* *ˁali* *ˁa* *mīn*
met.3ms Ali on who
'Who did Ali meet?'

Although (10b), like (10a), is interpreted as a direct question seeking information about the identity of the person that Ali met, the constituent containing the *wh*-word in (10b) appears inside the sentence, instead of on its left periphery.

In other dialects, like Egyptian Arabic, *wh*-in-situ is the default strategy for forming *wh*-interrogatives and, in some contexts, WH-movement is strictly prohibited (Wahba 1984). In Egyptian Arabic, nominal *wh*-words, like *ʾeh* 'what?', as well as non-nominal *wh*-constituents like *maʿa mīn* 'with whom?', can occur in situ, as illustrated in (11).

(11a) *mona nisyit tiktib ʾeh*
Mona forgot.3fs write.3fs what
'What did Mona forget to write?'
(11b) *mona ʾiftakarit ʾinn bāba*
Mona thought.3fs that father
xarag maʿa mīn
left.3ms with who
'With whom did Mona think that Father left?'

However, while *maʿa mīn* 'with whom?' can undergo WH-movement in simplex sentences (12a), it cannot do so in complex sentences (12b) (Wahba 1984).

(12a) *maʿa mīn mona xaragit*
with who Mona went.3fs
'With whom did Mona leave?'
(12b) **maʿa mīn bāba ʾiftakar*
with who father thought.3ms
ʾinn mona xaragit
that Mona left.3fs
'With whom did Father think that Mona left?'

Nominal *wh*-words like *ʾeh* cannot undergo WH-movement, as seen in (13).

(13a) **ʾeh mona ʾaret*
what Mona read.3fs
'What did Mona read?'
(13b) **ʾanhi walad mona šafit*
which boy Mona saw.3fs
'Which boy did Mona see?'

Wh-in-situ constructions raise a question regarding the necessity of (overt) displacement of *wh*-constituents as a defining characteristic for WH-movement (Huang 1982).

2. Gap as variable

Another important diagnostic of WH-movement is that it leaves a gap in the variable position. While it is true that *wh*-interrogatives in Modern Standard Arabic can always involve a gap in the base/original position, some *wh*-interrogatives involve a resumptive pronoun in that position (→ resumption), which shares relevant agreement features with the fronted *wh*-constituent (14). Wright (1862) refers to this pronoun as *ʿāʾid* or *rājiʿ*.

(14a) *man bi-yadi-hi malakūt-u*
who in-hand-his kingdom-Nom
kull-i šayʾ-in
every-Gen thing-Gen
'In whose hand is the kingdom of everything?' (Wright 1862:224)
(14b) *ʾayy-u rajul-in qālat*
which-Nom man-Gen said.3fs
fāṭimat-un ʾanna ʿaliyy-an
Fatima-Nom that Ali-Acc
raʾā(-hu)
saw.3ms-cl3ms
'Which man did Fatima say that Ali saw?'

Thus, within the class of *wh*-words, a distinction can be made between those that can be related to a resumptive pronoun, namely *man* 'who?' and *ʾayy NP* 'which NP?', and those that cannot (see Table 1). The illustrative examples in (15) contrast in that respect with those in (14).

(15a) *māḏā ʾakalat(*-hu)*
what ate.3fs(-*cl3ms)
fāṭimat-un
Fatima-Nom
'What did Fatima eat?'
(15b) *ʾayna ḏahabat(-*hu) mona*
where went.3fs(-*cl3ms) Mona
'Where did Mona go?'

Table 1. *Wh*-words in Modern Standard Arabic

Wh-words that can be linked to resumptive pronouns	*Wh*-words that cannot be linked to resumptive pronouns
	mā(ḏā) 'what?'
man 'who?'	*matā* 'when?'
	ʾayna 'where?'
	limā(ḏā) 'why?'
ʾayy(-u) NP 'which NP?'	*kayfa* 'how?'

In more recent studies of *wh*-in-situ in Arabic (Aoun and Choueiri 1999), it was observed that, in addition to being the only *wh*-words that can be related to a resumptive pronoun, *man* 'who?' and *ʾayy(-u) NP* 'which NP?' share the fact that both can appear in partitive constructions (16) and can be d-linked (17)–(18).

(16a) *man* *min-kum* *ḥafiḏ̣a*
who of-you.mp memorized.3ms
dawra-hu
role-his
'Which one of you has memorized his lines?'
(16b) *ʾayya* *lāʿib-in* *min-kum*
which player-Gen of-you.mp
raʾat *nadia*
saw.3fs Nadia
'Which player among you did Nadia see?'

(17) A: *hunā* *ʿiyādat-u* *ṭ-ṭabīb*
here clinic-Nom the-doctor
ʿomar *wa* *ṭ-ṭabīb* *ʿali*
Omar and the-doctor Ali
'This is Doctor Omar and Doctor Ali's clinic'
man *turīdu* *ʾan* *tarā*
who want.2ms that see.2ms
'Who do you want to see?'

(18) A: *hunā* *ʿiyādat-u* *ṭ-ṭabīb*
here clinic-Nom the-doctor
ʿomar *wa* *ṭ-ṭabīb* *ʿali*
Omar and the-doctor Ali
'This is Doctor Omar and Doctor Ali's clinic'
ʾayya *wāḥid-in* *turīdu*
which one-Gen want.2ms
ʾan *tarā*
that see.2ms
'Which one do you want to see?'

'D-linking' is a term coined to describe a crucial difference between *which*-phrases and *what* (Kuroda 1968; Pesetsky 1987): a *which*-phrase, but not *what*, is used under the assumption that both speaker and hearer have a set in mind, and that the answer to the *which*-phrase will be drawn from that set.

3. UNBOUNDED DEPENDENCIES AND SUBJACENCY

In *wh*-interrogatives involving gaps, the relation between the WH-moved constituent and its corresponding gap is unbounded. Thus, a *wh*-word can be extracted from an embedded tensed clause to the matrix complementizer position, as in (19).

(19a) *matā* *qulta* *l-ī* *ʾanna*
when said.2ms to-me that
muḥammad-an *raʾā* *ʿaliyy-an*
Muhammad-Acc saw.3ms Ali-Acc
'When did you say to me that Muhammad saw Ali?'
(19b) *māḏā* *zaʿama* *ʿaliyy-un* *ʾanna*
what claimed.3ms Ali-Nom that
muḥammad-an *ištara*
Muhammad-Nom bought.3ms
l-yawm-a
the-day-Acc
'What did Ali claim that Muhammad bought today?'

However, the unbounded relation between the WH-moved constituent and its corresponding gap is constrained by 'subjacency'. As illustrated in (20), gaps created by WH-movement are prohibited inside islands, such as the WH-island or the complex NP island (Ross 1967).

(20a) *_māḏā_ *saʾala* *ʿaliyy-un*
what asked.3ms Ali-Nom
man *ištara* *l-yawm-a*
who bought.3ms the-day-Acc
'What did Ali ask who bought today?'
(20b) *_māḏā_ *zaʿama* *r-rajul-u*
what claimed.3ms the-man-Nom
llaḏī *ištara* *ʾanna*
Rel bought.3ms that
muḥammad-an *qābala* *ʿaliyy-an*
Muhammad-Acc met.3ms Ali-Acc
'What did the man who bought claim that Muhammad met Ali?'

In contrast with the relation between the WH-moved constituent and its corresponding gap, the relation between a resumptive pronoun and its antecedent is not constrained by subjacency (21).

(21) *ʾayyu* *kitāb-in* *saʾala* *ʿaliyy-un*
which book-Gen. asked.3sm Ali-Nom
man *ištarā-hu* *l-yawm-a*
who bought.3sm-it the-day-Acc
'Which book did Ali ask who bought it today?'

The pronominal element found in (21) agrees in person, number, and gender with the sentence-initial *wh*-constituent. In addition, the relation between the resumptive pronoun and the sentence-initial *wh*-constituent is unbounded. Unlike gaps, however, resumptive pronouns can occur inside islands.

Based on the fact that *wh*-interrogatives in Arabic involve movement of a *wh*-constituent to the complementizer phrase, leaving a gap, and that the relation between this gap and its corresponding *wh*-constituent is unbounded but subject to subjacency, those constructions therefore display the typical characteristics of WH-movement. Resumptive *wh*-interrogatives involve unbounded dependencies, but they are not constrained by subjacency. Therefore, it is not clear whether they involve WH-movement.

4. Other WH-movement constructions

Other constructions in Arabic which display typical characteristics of WH-movement include relative clauses and topicalized constructions (22).

(22a) *al-māl-u* *llaḏī*
the-money-Nom Rel
taštahī *ʾanfus-u-nā*
desire.3fs souls-Nom-cl1p
'the money that our souls desire'
(22b) *al-kitāb-a/kitāb-a-n*
the-book-Acc/book-Acc-Indef
wajada *muḥammad-un*
found.3ms Muhammad-Nom
'the book/a book, Muhammad found'

In what has traditionally been termed 'topicalized constructions' (22b) (→ topicalization), a fronted constituent appears in clause-initial position, leaving a gap in its base/original position in the sentence. Bakir (1980) and Shlonsky (1996) have aptly termed these constructions 'focalization constructions' (→ focus) because of the interpretation of the preposed constituent as a contrastively focused element.

Relative constructions in Arabic, especially indefinite relatives, also involve unbounded dependencies with resumptive pronouns (→ resumption; → relative clause), as in (23).

(23a) *al-māl-u* *llaḏī*
the-money-Nom Rel
taštahī-hi *ʾanfus-u-nā*
desire.3fs-cl3ms souls-Nom-cl1p
'the money that our souls desire'
(23b) *māl-u-n* *taštahī-hi*
money-Nom desire.3fs-cl3ms
ʾanfus-u-nā
souls-Nom-cl1p
'money our souls desire'

The possible relation between the resumptive strategy and the gap strategy for forming unbounded dependencies and the question regarding the necessity of (overt) displacement of *wh*-constituents raised by *wh*-in-situ constructions are issues that have kept the topic of WH-movement at the forefront of the theoretical debate in → syntax.

Bibliographical references

Aoun, Joseph and Lina Choueiri. 1999. "Modes of interrogation". *Perspectives on Arabic linguistics*, XII, ed. Elabbas Benmamoun, 7–26. Amsterdam and Philadelphia: J. Benjamins.

Bakir, Murtadha. 1980. *Aspects of clause structure in Arabic*. Ph.D. diss., Indiana University.

Chomsky, Noam. 1977. "On Wh-movement". *Formal syntax*, ed. Adrian Akmajian, Peter W. Culicover, and Thomas Wasow, 71–132, New York: Academic Press.

Huang, C.-T. James. 1982. "Move-wh in a language without wh-movement". *Linguistic Review* 1.369–416.

Kuroda, Sige-Yuki. 1968. "English relativization and certain related problems". *Language* 44.244–266.

Pesetsky, David. 1987. "*Wh*-in-situ: Movement and unselective binding". *The representation of (in)definiteness*, ed. Eric Reuland and Alice ter Meulen, 98–129. Cambridge, Mass.: MIT Press.

Riemsdijk, Henk van and Edwin Williams. 1986. *Introduction to the theory of grammar*. Cambridge, Mass.: MIT Press.

Ross, John Robert. 1967. *Constraints of variables in syntax*. Ph.D. diss., Massachusetts Institute of Technology.

Shlonsky, Ur. 1996. "Remarks on the complementizer layer of Standard Arabic". *Research in Afroasiatic grammar*, ed. Jacqueline Lecarme, Jean Lowenstamm, and Ur Shlonsky, 325–344. Amsterdam and Philadelphia: J. Benjamins.

Wahba, Wafaa. 1984. Wh-*constructions in Egyptian Arabic*. Ph.D. diss., University of Illinois at Urbana-Champaign.

Wright, William. 1862. *A grammar of the Arabic language*, I. 3rd ed. Edinburgh: Williams and Norgate.

Lina Choueiri (American University of Beirut)

Wolof

1. INTRODUCTION

Wolof is the major lingua franca of → Senegal. The country has six national languages, Wolof, Mandinka, Joola, Seereer, Soninke, and Pulaar, as well as French, the official language. The population of the country is about 11 million people today. The population consists of about 92 percent Muslim, 6 percent Animist, and 2 percent Christian.

The high percentage of Muslims is due to the early Islamization of the country. According to Gellar (1995), northern Senegal's long involvement in the trans-Saharan trade exposed it to Islamic influences from North Africa. From northern Senegal (which corresponds to the present regions of Saint-Louis and Matam) spread the Almoravid movement, which swept through Morocco and Spain during the last third of the 11th century. Over the years, northern Senegal became a training ground for Muslim clerics and missionaries operating throughout modern Senegal and West Africa (Gellar 1995:3). By the 14th century, Islamic schools were established in Senegal, and most Senegalese Muslims were already able to use Arabic script to write their own languages (Diop 1989).

This early Islamization of the country accounts for the strong linguistic influence of Classical Arabic found in Wolof today. Yet, the Arabic language has never been used as a medium of communication in the daily life of the Senegalese people. Its use was and still is primarily restricted to religious spheres (Ngom 2004).

Arabic loanwords in Wolof fall into the following categories: (i) loans with no phonological adaptations and those with phonological adaptations (which have been integrated into the Wolof linguistic system); (ii) loans with semantic changes and truncations; (iii) loans with morphological adaptations (hybrid loans consisting of Arabic and Wolof morphemes); and (iv) loans with reanalysis of the Arabic definite article (whereby the Arabic article and other lexical items merge to form one new lexical unit). The following examples illustrate these types of Arabic loans found in Wolof.

2. PHONOLOGICAL ADAPTATION OF ARABIC LOANS

Some loans are adapted to Wolof phonological rules (which may or may not be identical to the Arabic forms), whereas others retain their Arabic forms for some speakers (especially those who have some knowledge of Arabic). Thus, Arabic loans which do not violate the Wolof linguistic system are used by all speakers. In contrast, adapted loans are often used by ordinary people, and the 'learned loans' (which have retained their Arabic traits) are used by people with some knowledge of Arabic. The following examples illustrate Arabic loans that did not need to be adapted to the Wolof linguistic system because of the similarity between the two languages: *xādim* 'servant', *salām* 'peace', *rasūl* 'prophet', *mūsa* (Arabic *mūsā*) 'Moses', *xalīfa* 'caliph', *miskīn* 'poor', *dunya* (Arabic *dunyā*) 'world', *sunna* 'customs of the Prophet', *islām* (Arabic *ʾislām*) 'Islam'. These Arabic loans show no phonological changes. This is due to the fact that the phonemes and the phonotactic properties of these words are identical to those in Wolof, a language that favors CV(C) syllable structures (Diouf 2001).

In contrast, loans which have retained their Arabic forms and contain phonemes that do not exist in the Wolof phonological system are typically used by a minority of Wolof speakers who are educated in Standard Arabic. Among the majority of Wolof speakers who have not been exposed to Standard Arabic, such Arabic loans are strongly influenced by Wolof and undergo significant linguistic adaptations to conform to Wolof linguistic constraints. The following examples illustrate the adaptation processes of such Arabic loans in Wolof: *alxurān* < *al-qurʾān* 'the *Qurʾān*', *waxtu* < *waqtu* 'time', *malāka* 'angel' < *malāʾika* [pl.], *ad(d)una* < *ad-dunyā* 'the world', *aj(j)ana* < *až-žanna* 'the heaven', *aj(j)uma* < *až-žumʿa* 'Friday', *yalla* 'God' < *yā ʾallāh* ~ *yaḷḷāh* 'O God!', *lëwël* 'first cup of tea' < *al-ʾawwal* 'the first', *nām* < *naʿam* 'yes', *aliw* < *ʿalī* 'Ali [proper name]'.

In the first two examples, although the Arabic voiceless uvular consonant *q* is considered to be part of the Wolof consonantal system in most Wolof grammars today, its frequent substitution with the Wolof voiceless velar fricative *x* in naturalized Arabic loans suggests that it may have been introduced in Wolof through the extensive borrowing of Arabic loans in the language, and that it was not initially part of the Wolof consonantal system. The deletion of the glottal stop, the vowel deletions, and other consonantal deletions are all triggered by the Wolof linguistic system. Since Wolof does not have the pharyngeal ʿ, this consonant is usually deleted in Arabic loanwords in Wolof. Similarly, Arabic geminates resulting from assimilation, such as *al-dunyā* 'the world' > *ad-dunyā*, *al-janna* ~ *al-žanna* 'the heaven' > *až-žanna*, and *al-jumʿa* ~ *až-žumʿa* 'Friday' > *až-žumʿa* in Arabic are typically maintained or simplified in Wolof. The postalveolar consonant *ž* is also replaced by its closest Wolof counterpart, *j*.

3. LOANS WITH SEMANTIC CHANGES AND TRUNCATIONS

In addition to these phonological adaptations, semantic shifts and truncations are also common in Arabic loans found in Wolof. Semantic changes are exemplified by *tubāb* 'white person' < *ṭabīb* 'doctor'; *bilāl* 'muezzin; [male proper name]' < *bilāl* '[male proper name]'; *alkāti* 'police officer' < *al-qādī* 'judge'; *nūrayni* '[first name of a Senegalese religious leader]' < *nūrayni* 'two lights'. These words with semantic changes are fully adapted loans. The meanings of the examples are extended to cover other semantic domains in Wolof. These words are historical remnants of past influences and refer to Arabic items, professions, or cultural or religious concepts that came along with Islam in Senegal. For instance, the Arabic word *bilāl* refers to one of the most prominent black persons in the history of Islam, whose specialty was to call people for prayer. Today, the meaning of the word has been generalized in Wolof and serves both as a first name and to indicate the religious function of calling people to prayer in Wolof.

Examples of truncations are *xadi* '[female proper name]' < *xadīja; mamadu* '[male proper name]' < *muḥammad; bis* '[interjection]' < *bismillāh* 'in the name of God'; *sallāw* 'if God please' < *ʾinšāllāh*. All deleted segments in these loans contain phonological elements that do not exist in Wolof. This shows that these truncations are also motivated by the Wolof linguistic system and are designed to enable Arabic loans to function effectively in the language.

4. MORPHOLOGICAL ADAPTATION

In addition to these Arabic loans, Arabic loans with morphological adaptations (which consist of Arabic and Wolof morphemes) are also common in Wolof. The examples in Table 1 illustrate these types of morphological adaptions and show instances in which Arabic words were integrated into Wolof to such an extent that they are combined morphologically with Wolof morphemes.

Table 1. Morphological integration of Arabic loanwords

1. *arāmul* 'it is not forbidden'	Arabic *harām* 'forbidden' + Wolof *-ul* [3rd person negative]
2. *gennad(d)una* 'to die'	Wolof *genn* 'to go out, exit' + Arabic *ad-dunyā* 'the world'
3. *bāylād* 'name of religious leader in the Mourid brotherhood'	Wolof *bāy* 'father' + Arabic *al-ʾaḥad* 'the one'
4. *saraxe* 'to give charity to someone'	*sarax* < Arabic *ṣadaqa* + Wolof *-e* [ablative]
5. *sēxal* 'to give someone the title of sheikh'	Arabic *šayx* 'sheikh' + Wolof *-al* [causative]
6. *saraxāt* 'to give charity again'	*sarax* < Arabic *ṣadaqa* 'charity' + Wolof *-āt* [iterative]
7. *siyārewāt* 'to pay a visit to someone again'	Arabic *ziyāra* 'visit' + Wolof *-āt* [iterative]
8. *saraxendo* 'to give charity together'	*sarax* < Arabic *ṣadaqa* 'charity' + Wolof *-ando* 'together'
9. *tafsīrkat* 'Islamic preacher'	Arabic *tafsīr* 'Qurʾānic exegesis' + Wolof *-kat* [agent]

These examples illustrate the types of adaptation processes involved in hybrid loans in Wolof. Similar to the adapted loans mentioned above, these examples also show that Arabic structures and segments that do not exist in Wolof are usually replaced by their closest Wolof counterparts. The deletion of the glottal consonants *h*, *ḥ*, ʾ in examples 1 and 3, the replacement of the Arabic *š* and *z* by *s* (in examples 5 and 7), and *q* by *x* (in examples

4, 6, and 8) show the extent to which Arabic loans are transformed to meet the phonological requirements of Wolof. This is consistent with example 7, in which *w* is introduced between the Wolof iterative morpheme *-āt* and the last vowel of the word [e] to avoid the formation of vowel cluster [eā], which is unacceptable in Wolof. Furthermore, the consonants *ṣ* and *d* in examples 4 and 6 are replaced by the Wolof consonants *s* and *r*, respectively.

5. Cases of reanalysis of the Arabic definite article *al-* 'the'

Loans which involve the reanalysis of the Arabic article *al-* comprise a major part of Arabic loans found in Wolof. These loans are usually formed by merging several elements (generally the Arabic definite article *al-* and other independent lexical units) to form one new lexical entry in Wolof, as exemplified in Table 2.

Table 2. Reanalysis of definite article in Arabic loanwords

Wolof	Arabic
alāxira 'judgment day'	*al-ʾāxira* 'Judgment Day'
ilimān 'imam'	*al-ʾimām* 'the imam'
asaka 'charity'	*az-zakāt* 'charity'
ajana 'heaven'	*al-janna* 'heaven'
alkāti 'police officer'	*al-qāḍī* 'the judge'
alxamis 'Thursday'	*al-xamīs* 'Thursday'
asamān 'sky'	*as-samāʾ* 'the sky'
aduna 'world'	*ad-dunyā* 'the world'
lëwël 'first cup of tea'	*al-ʾawwal* 'the first'
yalla 'God'	*yā ʾallāh* ~ *yaḷḷāh* 'O, God!'

With the exception of the last example *yalla* 'God', which is also pronounced as such in Arabic, the examples above are among the most common instances of reanalysis of the Arabic definite article found in Wolof. All these examples share one property. While the Arabic words consist of two elements (the article *al-* 'the' followed by a noun, as in examples 1 through 9, and the vocative *yā* followed by a noun, as in example 10), the reanalyzed forms in Wolof only consist of one lexical element. Thus, the Arabic article *al-* and the vocative *yā* have merged to form a new lexical unit in Wolof. Wolof speakers regard these reanalyzed units as single lexical elements and consequently use them with Wolof determiners, as in *asamān-bi* 'the sky', *alkāti-bi* 'the police officer', *ilimān-bi* 'the imam', etc.

6. Conclusion

Most Arabic loans in Wolof today have been in the language for so long that they have become part of Wolof speakers' lexical competence, and most Wolof speakers are unaware of the Arabic origins of these words. Based upon these Arabic loanwords in Wolof, Senegal can be regarded as a culturally creolized society (Swigart 1994), deriving one part from Wolof culture and another from Islamic culture. The considerable number of Arabic loans found in Wolof today is symptomatic of the linguistic consequences of this cultural mixture.

Bibliographical references

Diop, Amadou H. 1989. *Language contact, language planning and language policy in Senegal.* Ph.D. diss., Pennsylvania State University.

Diouf, Jean L. 2001. *Grammaire du wolof contemporain.* Tokyo: Institute for the Study of Languages and Cultures.

Gellar, Sheldon. 1995. *Senegal: An African nation between Islam and the West.* Boulder, Colo.: Westview Press.

Ngom, Fallou. 2004. "The social status of Arabic, French and English in the Senegalese speech community". *Language Variation and Change* 15.349–366.

Swigart, Leigh. 1994. "Cultural creolization and language use in post-colonial Africa: The case of Senegal". *Africa* 64:2.177–189.

Fallou Ngom (Boston University)

Word Order

1. Introduction

Since the 1960s, the concept of word order has received widespread interest within linguistics. Through a seminal article by Joseph H. Greenberg (1966) on universals of grammar, a new branch opened up within general linguistics, that of language typology and language universals. Greenberg presented a wide range of different word orderings in about thirty languages, but it was the order of the verb (V), the subject (S), and the object (O) that received the most attention in studies on the classification of the world's languages. Later surveys showed that the two most common variants were SOV and SVO, with SOV around 45–50 percent and SVO 40–45 percent. VSO occupies the third place, with about 10 percent of the world's languages; VOS scored 2–3 percent, OVS 0–1 percent, and OSV 0.0–0.2 percent (Dahlgren 1998:97).

Having noticed the overwhelmingly high SO order in the world's languages, Dik (1981:21) suggested the following language-independent schema:

P2, P1 (V) S (V) O (V), P3

Here, S and O stand for the unmarked positions of Subject and Object, the Vs for possible positions of verbs, and the Ps for 'special positions' used for certain purposes. The commas stand for breaks in the intonation. P2 is the 'left dislocation' position (such as *Bill* in *Bill, I saw him yesterday*). P1 is the slot for certain marked constructions like question words, topic, or focus. P3 is the 'right dislocation', which is an explanatory afterthought, as in *I like him, Bill.*

Greenberg also studied the correlations between basic word order and other parameters, and found that other orders could often be predicted from the basic word order. For instance, in a language with VSO, one can predict that it will also have prepositions (Pr) and not postpositions (Po); the head in a genitive construction will precede the genitive (NG); and the adjective will follow the noun (NA). However, this pattern is also valid for SVO, which led some linguists to drop the notion of 'subject' and speak of two main word order types (Comrie 1989:95):

VO, Pr, NG, NA
OV, Po, GN, AN

The first type is characteristic of Arabic in nearly all of its different types, diachronically and synchronically.

In classifications based on the relative positions of the subject, object, and verb, Classical Arabic has commonly been classified as VSO among Arabists, whereas the modern spoken variants have often been referred to as SVO varieties. Mohammad (2000:2–3) notes that traditional grammars seem to accept all six orders as possible and grammatical, but he regards SOV and OSV as marginal and questions their validity. According to Mohammad (2000:83), VSO is the "discourse neutral word order", but SVO is the underlying word order. Majdi (1990) and Anshen and Schreiber (1968) propose that Arabic is a VOS language. These three investigations were all written within the generative model, based on the deductive tradition, with a multilevel syntactic structure (underlying and surface), and the relationship between the two levels accounts for the variation in word order. In the following discussion, the treatment of word order is based on the inductive tradition, within an empirical, discourse-pragmatic perspective, characteristic of the Prague school's 'Functional Sentence Perspective', and on the 'Functional Grammar' developed by Halliday (1994) and Dik (1981), among others, who share a similar approach but do not necessarily adhere to one particular linguistic school.

2. Word order in Modern Standard and Early Arabic and in the Eastern dialects

2.1 *Modern Standard Arabic: Newspaper language*

In Modern Standard Arabic, Parkinson (1981: 28) found a fairly high percentage of SVO in newspaper language, as shown in Table 1.

Table 1. SVO in newspaper language

Type	% SVO
Headlines	92
Political writing	48
Editorials	39
Short stories	39
Linguistics dissertations	34
Magazines	30
Scholarly journals, political science	27
News articles	8

Parkinson (1981:25) notes that even VOS and OVS turned out to be possible word orders and in fact occurred quite frequently.

2.2 *Early Arabic and Eastern dialects*

A recent larger investigation has confirmed the VSO character of Classical Arabic and of most of the Eastern dialects that are represented with sufficiently large text corpora (Dahlgren 1998). The dialects in question range from Cairo to Mesopotamia and include those of

Central Arabia. All Bedouin variants over the area were classified as VSO, as well as most other village and town dialects. Exceptions were found in Egypt, where both VSO and SVO were found in different sources, whereas the Anatolian dialects clearly exhibited a main SVO pattern, due to contact influence from Kurdish and Turkish, both of which are SOV languages. In → Uzbekistan Arabic, contact influence has affected the language to the extent that it has become a SOV language (Versteegh 1984–1986:446–447, 452). The predominant SVO order in → Anatolian Arabic shows a smaller degree of contact influence than the SOV order in Uzbekistan Arabic. In the former, the subject has been preposed to the verb, but the next step, with the object preposed to the verb, is only common in Uzbekistan, due to the heavy contact influence on Uzbekistan Arabic and its isolated existence.

The mixed pattern in Egypt may be due to both diachronic and geographical factors. The Sharqiyya texts from the northeastern Delta displayed VSO, at variance with modern Cairene texts, suggesting that Cairo itself and other parts of Egypt belong to a dialect area where SVO is dominant (Dahlgren 1998:169, 174–176). The occurrence of different forms for the passive supports the view of different dialect areas (Retsö 1983:79–80), especially the use of the *n*-prefix vs. the *t*-infix, the former being more common in the Syro-Palestinian area, whereas the latter dominates in Egypt. Cairo experienced a large influx of rural migrants in the 20th century, which may explain the VSO character in the texts of Spitta-Bey (1880) and the dominant SVO order of the much later ones in Hassan (1971, 1981) and Nakano (1982).

3. Occurrence of independent subjects

Before examining the factors that affect word order in Arabic, it is important to examine the distribution of basic sentences with independent verb, subject, and object. These are relatively rare in both spoken and written Arabic because the subject is typically marked as an affix on the verb. It often lacks an independent form in the sentence, making sentences of this type irrelevant to the determination of basic word order, based as it is on the relative order of subject, verb, and object. Independent subjects in Arabic are most common when a subject shift takes place in a discourse/text (Dahlgren 1998:130–134, 170). Approximately only every third or fourth sentence contains an independent subject in written and spoken Arabic. Objects occur even more rarely. In Dahlgren's (1998) study, only 301 objects were found in narrative discourse in a sample of 2,657 instances with independent subjects in the Early Arabic of Ibn ʾIsḥāq's *Sīrat an-nabī*, which indicates that approximately only 3 percent of all sentences in this type of discourse include a verb and independent subjects and objects (Dahlgren 1998:207, 249–250).

4. Discourse type

Recent research has shown that discourse type is an important factor in grammatical analysis. A number of major and minor types of discourse have been identified, many of which have characteristics that influence word order. Longacre (1983:1–5, 39) makes a main division into 'dialogue' and 'monologue', where the former is characterized by uncontrolled interaction, often with unexpected turns and twists, and the latter by complete control of the discourse by the speaker or writer. Monologue can be further divided according to the parameters of 'contingent succession' and 'agent orientation'. For narrative discourse, both parameters are set to + (positive), whereas procedural discourse (how to do it, how it was done, how it takes place) has contingent succession but lacks agent orientation. Behavioral discourse is a broad category that includes eulogies and political speeches of candidates; it lacks contingent succession but has agent orientation. Expository discourse ranges from the familiar essay to the scientific article and is assigned – (negative) for both parameters.

Another important distinction is that between written and spoken language. Written language is characterized by the fact that more time is spent on both producing and processing, which significantly reduces the need for redundant but supporting elements. Spoken language contains many incomplete sentences, often simple sequences of phrases; it typically contains little subordination; passives occur to a lesser extent than in written language (Brown and Yule 1983).

Only narrative and dialogue are treated in the present entry, since other discourse types have not been sufficiently investigated. Expository discourse tends to contain a high proportion of S(ubject)V(erb)COMP(lement), as Holes (1995:205) observes.

4.1 *Narrative: Foreground and background*

Several quite intricate language phenomena hide behind the simple classification of most varieties of Arabic as VSO. VSO is regarded as the basic word order, which means that it is the default, or normal, word order in verbal sentences. A sentence is 'default' if it is the most common variant and exhibits least structural complexity and/or least cognitive complexity. The basic sentence has been defined as the foregrounded, main, declarative, affirmative, active, and continuative clause (Dahlgren 1998:93). For Arabic, the last criterion was excluded in Dahlgren's investigation because continuative clauses do not contain independent subjects, as explained below.

One of the great advances within → text linguistics in recent years is the notion of foreground and background (→ grounding), elaborated mainly by Labov and Waletzky (1967) and Hopper (1979). Foreground gives the main story line in a narrative; it presents events in sequence or consecutively, i.e. events that follow each other sequentially. Backgrounded events and situations are those which deviate from this story line, being anterior to it, simultaneous, or farther in the future. Hopper regards this distinction as a universal; it may be displayed through morphology as the alternation between the *ki-* and *ka-* morphemes on Swahili verbs (verbs preceded by *ka-* are used consecutively, i.e. in foreground, and verbs preceded by *ki-* expand the story with suppletive, backgrounded sentences in this language); through verbal forms like the *passé simple* for foreground and the *imparfait* in written French; or through word order, of which Hopper found manifestations in older types of English (VS and OV clauses were connected with foreground and the SV type with background).

Table 2 presents the word order in the *Sīra* of Ibn ʾIsḥāq, as representative of Early Arabic; the narrative discourse is divided into foreground and background (Dahlgren 1998:249–250). Left-dislocated elements are disregarded as subjects, since the dislocated element occurs in the following sentence, most commonly as a suffixed or separate pronoun (→ topicalization). In a sentence like *Khalid, I* saw *him yesterday*, *Khalid* is the left-dislocated element, and the word order in the sentence is accordingly SVO. In Arabic, however, only SV would have been noted, since the pronoun would be suffixed, and suffixed pronouns were not counted as

Table 2. Word order in Early Arabic

i. Narrative discourse, foreground, definite subjects

	VS	VSO		VOS		OVS		Tot.	%	SV	SVO		SOV		OSV		Tot.	%
		Od	Oi	Od	Oi	Od	Oi				Od	Oi	Od	Oi	Od	Oi		
Perfect																		
Def. n.	1,799	142	30					1,978	98	26	13	1					40	2
Pers. pr.	1																1	
Dem. pr.	15	6	1	1				23	100									
Imperfect																		
Def. n.	20	1	1					22	92	2							2	8
Pers. pr.	0							0	0	3							3	100

ii. Narrative discourse, background, definite subjects

	VS	VSO		VOS		OVS		Tot.	%	SV	SVO		SOV		OSV		Tot.	%
		Od	Oi	Od	Oi	Od	Oi				Od	Oi	Od	Oi	Od	Oi		
Perfect																		
Def. n.	275	37	7	7				326	985	2			7	2				
Dem. pr.	4							4	1000				0	0				
Imperfect																		
Def. n.	24	1						25	6113	3			16	39				
Pers. pr.	0							0	08				8	100				
Dem. pr.	1							1	0				0					

objects. (In the tables, Od indicates 'definite object' and Oi 'indefinite object'.)

There is no great difference in word order between foreground and background. The distinction is achieved mainly through the conjunctions *fa-* and *wa-* in Early Arabic, the former designating foreground, the latter background (Dahlgren 1998:207–209).

The very low occurrence of pronouns in narrative discourse is explained by the absence of the subject requirement explained above (→ pro-drop). The more extensive presence of pronouns in dialogue is a reflection of their wider use in live speech (see Sec. 4.2).

Modern spoken Arabic lacks morphological devices like *fa-* and *wa-* for marking foreground and background. A presentation of the material from the modern dialects in Syria, Lebanon, Jordan, and Palestine (Eastern Mediterranean Modern Dialects), Bedouin dialects, and Mesopotamian and Egyptian dialects, according to the parameters VS/SV order and foreground/background, showed that word order was relevant for the latter distinction, with parts of Egypt and the whole of Anatolia as notable exceptions. The figures for Eastern Mediterranean Modern Arabic are presented in Table 3 as representative of those dialects that have remained unaffected by any other language in terms of word order. Note that Anatolian Arabic represents a different type (Dahlgren 1998: 230, 232). Objects are unaccounted for here.

The figures present VS as the unmarked order in foregrounded narrative discourse in the Eastern Mediterranean dialect, while in backgrounded discourse, SV is the unmarked order. SV is also the unmarked order in both foregrounded and backgrounded discourse in Anatolia. Deviations from the general pattern are due to different pragmatic articulations, such as focusing, topicalization, and peak (Dahlgren 1998:183–186).

4.2 *Dialogue*

Dialogue differs from narrative discourse through its lack of the foreground/background distinction, in addition to other differences due

Table 3. Narrative discourse: Foreground and background in modern dialects

i. Eastern Mediterranean Modern Arabic

	FOREGROUND				BACKGROUND			
	VS	%	SV	%	VS	%	SV	%
Perfect								
Def. n.	912	87	131	13	61	42	84	58
Pers. pr.	121	63	70	37	9	12	65	88
Dem. pr.	243	66	125	34	7	10	65	90
hā + def. n.	85	79	22	21	8	53	7	47
Imperfect								
Def. n.	89	85	16	15	30	27	82	73
Pers. pr.	9	36	16	64	8	14	51	86
Dem. pr.	14	56	11	44	7	9	68	91
hā + def. n.	9	69	4	31	8	28	21	72

ii. Anatolia

	FOREGROUND				BACKGROUND			
	VS	%	SV	%	VS	%	SV	%
Perfect								
Def. n.	73	28	186	72	16	19	68	81
Pers. pr.	12	13	83	87	1	2	41	98
Dem. pr.	11	17	54	83	0	0	24	100
hā + def. n.	1		0		0		1	
Imperfect								
Def. n.	4	9	40	91	3	7	41	93
Pers. pr.	0	0	11	100	1	4	24	96
Dem. pr.	4	36	7	64	1	6	16	94
hā + def. n.	1		0		0		1	

to text production. In a real dialogue, there is a lower degree of planning the discourse, as opposed to narrative, and a greater need for redundant material to support the presented information, unlike written narrative, which can be read slowly and repeatedly so as not to lose any information. One should note that dialogues from Early Arabic are artificial, i.e. construed, and written, whereas in the dialect material, they are authentic and spoken. Table 4 shows word order in dialogue in Early Arabic and two modern dialects. In Early Arabic, the independent objects are included, but they have been omitted in the case of the modern dialect in order to give a better overview. The orders of objects do not differ to a large degree in Early Arabic, compared to modern Arabic dialects. The material exhibits noticeably more VOS forms in the modern dialects, and even some scattered instances of OVS, SOV, and OSV (Dahlgren 1998:246). The difference between definite object (Od) and indefinite object (Oi) in Table 4 corresponds fairly well with the distinction between given and new information.

There is a significantly larger use of personal pronouns in dialogue in Early Arabic as compared to its narrative discourse. This is most likely a reflection of spoken Arabic, because it appears in both old and modern times. Tables 3 and 4 show that in Eastern Mediterranean Modern Arabic, the number of personal pronouns is 22 percent of the number of definite nouns in narrative, but 133 percent in dialogue.

Unlike narrative discourse, there is not a larger difference in the VS-SV word order in this discourse type (i.e. dialogue) between Early Arabic and many modern dialects. This leads to the conclusion that no significant change has occurred diachronically in this respect.

Dialogue exhibits a larger degree of SV with imperfect verb forms, and is generally characterized by a higher SV rate than narrative discourse. This probably stems from the rather different character of dialogues; its volatile

Table 4. Dialogue in Early Arabic and modern dialects

i. Early Arabic, definite subjects

	VS	VSO		VOS		OVS		Tot.	%	SV	SVO		SOV		OSV		Tot.	%
		Od	Oi	Od	Oi	Od	Oi				Od	Oi	Od	Oi	Od	Oi		
Perfect																		
Def. n.	130	15	2					147	74	33	17	2					52	26
Pers. pr.	0							0		3	3						6	
Dem. pr.	2							2		4	1						5	
Imperfect																		
Def. n.	18	3	1					22	49	15	7	1					23	51
Pers. pr.	1							1	3	24	8	2					34	97
Dem. pr.	0							0		6							6	

ii. Eastern Mediterranean Modern Arabic

	Perfect				Imperfect			
	VS	%	SV	%	VS	%	SV	%
Def. n.	77	54	66	46	33	39	52	61
Pers. pr.	11	9	113	91	20	11	159	89
Dem. pr.	4	27	11	73	3	23	10	77
hā + def. n.	4		4		6	60	4	40

iii. Anatolia

	Perfect				Imperfect			
	VS	%	SV	%	VS	%	SV	%
Def. n.	7	9	72	91	7	16	37	84
Pers. pr.	4	6	66	94	4	4	108	96
Dem. pr.	2	15	11	85	0	0	10	100
hā + def. n.	0		3					

nature, with frequent switches of direction and a larger need for redundancy, makes it more subject to marked structures such as topicalization and focusing. The fairly ample use of SV syntax in dialogue is probably also due to the common occurrence of sentences similar to beginning sentences in narrative discourse, i.e. sentences that are found at the very beginning of a story (Dahlgren 1998:187, 191). However, since the VS rate dominates in clauses with definite nouns and perfect verbs in Eastern Mediterranean, it is considered to be the unmarked order, while other orders are the result of marked structures.

5. OV SYNTAX IN ARABIC

When an independent subject is absent, the tolerance for OV order is larger. A focused object, for example, appears in sentence-initial position in (1) from the *Qurʾān* (Q. 17/20).

(1) *kull-an* *numiddu* *hāʾulāʾi* *wa-hāʾulāʾi*
all-Acc succour.1p these and-these
min *ʿaṭāʾ-i* *rabb-i-ka*
from gift-Gen lord-Gen-your
'Each We succour, these and those, from thy Lord's gift'

In this example, *kullan* is an indefinite object and is accordingly regarded as a focused constituent, which appears much less frequently than the topicalized constituents.

6. WORD ORDER FACTORS

6.1 *Foreground and background*

Having established the importance of studying each discourse type separately, it is now important to consider more specific factors that have a bearing on word order. One such factor is the foreground/background distinction (Sec. 4.1; see discussion of Longacre's discourse types and narrative discourse in Sec. 4), which is a crucial factor in modern spoken Arabic of the Eastern Mediterranean type.

6.2 *Topicalization*

Topicalization constitutes another important factor that influences word order. SV order in foreground and dialogue is analyzed as → topicalization of the subject. This slot, i.e. the initial position to be occupied by a verb, subject, object, or adverbial element, may also be used for a marked → focus construction, indicated by indefinite subjects in preverbal position, as shown in Table 7 (dealing with the topicality hierarchy). Objects may undergo topicalization, as Table 5 shows.

Table 2 shows that VSO was the most common word order in Early Arabic sentences that included independent subjects and objects. The somewhat astonishing figures in Table 5 present both VOS and SVO as far more common than VSO in Eastern Mediterranean Modern Arabic. So, quite paradoxically, one may

Table 5. Objects with definite subjects in narrative discourse

i. Eastern Mediterranean Modern Arabic

	VSO		VOS		OVS		SVO		SOV		OSV	
	Od	Oi	Od	Oi	Od	Oi	Od	Oi	Od	Oi	Od	Oi
Perfect												
Foreground	18	6	52	4	0	0	58	12	0	0	0	0
Background	2	0	0	0	0	0	0	0	0	0	0	0

ii. Anatolia

	VSO		VOS		OVS		SVO		SOV		OSV	
	Od	Oi	Od	Oi	Od	Oi	Od	Oi	Od	Oi	Od	Oi
Foreground	2	0	2	1	0	1	34	18	0	0	0	0
Background	2	0	0	0	0	0	5	11	0	1	0	1

say that this dialect is of unmarked VS order in narrative discourse (according to Table 3), although instances where all three constituents are present do not indicate unmarked VSO. This leads to the conclusion that SVO is the word order for topicalization or focusing of the subject, and VOS the word order for topicalization or focusing of the object (→ topicalization). The absence of OV order in Eastern Mediterranean Modern Arabic in Table 5 indicates that this dialect type is of rigid VO order, just as Early Arabic, as is evident from Table 2.

The Anatolian dialects indicate that SVO is the unmarked word order, which is in accordance with the previous conclusion that these dialects have unmarked SV order. There are also scattered instances of all other orders—an obvious consequence of the contact influence from the SOV languages, referred to above.

Statistics from the analysis of dialogue discourse data give a similar picture, as shown in Table 6.

The figures for the Eastern Mediterranean dialects are similar to those obtained from the analysis of narrative discourse in Table 5; the supposedly unmarked VS order is not manifested when all three constituents appear together. The SVO tendency in dialogue is, however, stronger than it is in narrative discourse.

6.3 *Focusing*

Table 6 also indicates, through the instances with indefinite objects, that VOS may constitute a marked → focus construction, as illustrated by the following example from the *Qurʾān* (Q. 20/67).

(2) *fa-ʾawjasa* *fī* *nafs-i-hi*
and-conceived.3ms in soul-Gen-his
ḫīfat-an *mūsā*
fear-Acc Moses
'And Moses conceived a fear within him'

To some extent, the statistics presented here concerning Early Arabic and Eastern Mediterranean support the schema suggested by Moutaouakil (1989:60) for Classical Arabic:

Theme, CO {Foc/Top} V S O X, Tail

'Theme' and 'Tail' correspond to Dik's P2 and P3 slots (see Sec. 1). The slot for the focus or topic here corresponds to Dik's P1. CO stands for clause operators, such as *ʾinna*, *ʾinnamā*, *qad*, *laqad*, and X for other constituents to which no syntactic function has been assigned, such as adverbs and prepositional phrases. Yet, as we have seen, the preverbal slot for pragmatic functions serves only for subjects, so for pragmatic objects a modification is required:

Theme, CO {Foc_{subj}/Top_{subj}} V {Foc_{obj}/Top_{obj}} S O X, Tail

6.4 *Topicality*

Related to topicalization is the concept of topicality. Different degrees of topicality indicate a subject's propensity for being topicalized; higher topicality means higher SV rate in Dahlgren's analysis of the tables on the Topicality Hierarchy below (Dahlgren 1998:172–173, 234). Table 7 presents statistical support for this conclusion.

Table 6. Objects with definite subjects in dialogue

i. Eastern Mediterranean Modern Arabic

	VSO		VOS		OVS		SVO		SOV		OSV	
	Od	Oi	Od	Oi	Od	Oi	Od	Oi	Od	Oi	Od	Oi
Perfect	0	2	7	0	0	0	26	14	0	1	0	0
Imperfect	2	0	2	1	0	0	30	17	0	0	0	0

ii. Anatolia

	VSO		VOS		OVS		SVO		SOV		OSV	
	Od	Oi	Od	Oi	Od	Oi	Od	Oi	Od	Oi	Od	Oi
Perfect	1	0	1	0	0	0	10	6	0	1	0	0
Imperfect	0	0	0	0	0	0	5	7	0	0	0	0

Table 7. The Topicality Hierarchy in narrative discourse

i. Eastern Mediterranean Modern Arabic

	Existential		Indefinite		Definite		Anaphoric pronoun			
							Pers. pr.		Dem. pr.	
	Tot.	%VS	Tot.	%VS	Tot.	%VS	Tot.	%VS	Tot.	%VS
Perfect										
Foreground			122	96	1,043	87	192	64	368	66
Background	34	100	25	60	145	42	74	12	72	10
Imperfect										
Foreground			14	93	105	85	25	36	25	56
Background	4	100	14	43	112	27	51	14	75	9

ii. Anatolia

	Existential		Indefinite		Definite		Anaphoric pronoun			
							Pers. pr.		Dem. pr.	
	Tot.	%VS	Tot.	%VS	Tot.	%VS	Tot.	%VS	Tot.	%VS
Perfect										
Foreground			68	71	259	28	95	13	65	17
Background	48	65	25	48	84	16	42	2	24	0
Imperfect										
Foreground			7	71	44	9	11	0	11	36
Background	6	50	5	20	44	7	25	4	17	6

'Tot.' refers to the total number of instances in the actual sample. Existential expressions, such as Classical Arabic *kāna rajulun* 'there was a man', exhibit 100 percent VS order in Classical Arabic and in all other dialects studied that have not been exposed to contact influence from an S(O)V language. In Anatolian Arabic, 48 existential expressions with the perfect verb forms show 65 percent VS order. The overall tendency is toward a larger degree of SV order as one moves to the right in the table, which also means columns with an increasing degree of topicality; demonstrative and personal pronouns belong to one single category, the anaphoric pronouns.

Table 7 shows a slight tendency toward a more dominant SV order in clauses with imperfect verbs, a tendency that was prevalent in other dialect areas as well. This tendency prevails even in dialogue in dialects with little contact influence on word order, as shown in Table 8.

6.5 *Animacy*

The concept of 'animacy' is also connected with topicality: human beings have more topicality than inanimate beings, and superhuman beings have higher topicality than human beings; hence, they receive SV syntax to a larger degree (→ topicalization).

6.6 *Aspect*

Tables 7 and 8 also show that the aspect of the verb has a bearing on word order. SV syntax is more common with imperfect verbs than perfect verbs. This is shown even more clearly by verbs that denote punctual aspect. In a sample in the Eastern Mediterranean Modern Arabic type dialects with 164 verbs expressing punctual aspect, with definite nouns in the perfect and foreground in narrative discourse, 99 percent VS order was found (Dahlgren 1998:237).

Table 8. The Topicality Hierarchy in dialogue

i. Eastern Mediterranean Modern Arabic

	Existential		Indefinite		Definite		Anaphoric pronoun			
							Pers. pr.		Dem. pr.	
	Tot.	%VS	Tot.	%VS	Tot.	%VS	Tot.	%VS	Tot.	%VS
Perfect	3	100	40	92	143	54	124	9	15	27
Imperfect	0		11	73	85	39	179	11	13	23

ii. Anatolia

	Existential		Indefinite		Definite		Anaphoric pronoun			
							Pers. pr.		Dem. pr.	
	Tot.	%VS	Tot.	%VS	Tot.	%VS	Tot.	%VS	Tot.	%VS
Perfect	2	50	26	77	79	9	70	6	13	15
Imperfect	0		5	80	44	39	112	4	10	0

6.7 *Rhythm*

According to Holes (1995:206), there is a rhythmic principle in Arabic that, apart from the verb, causes the 'heavier' sentence constituents to follow the lighter ones. This is because there is a rhythmic break in Arabic that falls at or slightly before the halfway point in the sentence at a major constituent boundary. Sometimes, this principle conflicts with the principle of 'definite first', where the latter usually prevails (Holes 1995:206).

7. Word order in subordinate clauses

Word order in subordinate clauses has not yet been treated in a larger investigation. *Ḥāl*-clauses generally have SVO order, whereas temporal clauses may exhibit VOS order when the object is more topical than the subject, as in (3).

(3) *fa-lammā jāʾa ʾāl-a lūṭ-in*
and-when came.3ms folk-Acc Lot-Gen
al-mursal-ūna, qāla
the-envoy-mp.Nom said.3ms
'So, when the envoys came to the folk of Lot, he said,...' (Q. 15/61–62)

8. Adjuncts

Adjuncts are prepositional phrases or adverbs that belong to the periphery of a clause, outside the core, i.e. the verb with its arguments (Van Valin and Lapolla 1997:26). The X-slot in Moutaouakil's schema given above (Sec. 6.3) is their unmarked position, but they may sometimes be fronted for highlighting or scene setting, as in (4) and (5).

(4) *ʾān-un li-balīġ ʾafandī ʾan*
time-Nom for-Baligh Afandi that
yuʾmin-a bi-naṣīḥat-i
believe.3ms-Subj on-advice-Gen
ṣadīq-i-hi l-mutaṭabbib-i
friend-Gen-his the-medical-Gen
fa-li-yartaḥil-a ʿalā ʾajal
and-in.order.to-go3ms-Subj on haste
'It was high time for Baligh Afandi to believe the advice of his medically inclined friend, and he had better go quickly' (Bloch 1974:24)

(5) *fī ṣabāḥ-i l-yawm-i t-tālī*
in morning-Gen the-day-Gen the-next
taʾālat tahālīl-u l-faraḥ-i
be.raised.3fs shouts-Nom the-joy-Gen
bayna l-junūd
among the-soldiers
'The next morning shouts of joy were raised among the soldiers' (Bloch 1974:51)

The adjuncts discussed here occupy Dik's P1 position. They are likely to appear in this position as alternatives to topics or focus elements, making the P1 position a slot for one single pragmatic activation. The modified schema by Moutaouakil, presented in Section 6.3, should therefore be rewritten once more as:

Theme, CO {Foc_{subj}/$\text{Top}_{\text{subj/}}$Adjunct} V {$\text{Foc}_{\text{obj}}$/$\text{Top}_{\text{obj}}$} S O X, Tail

There are also prepositional phrases that occur in a fixed position in the sentence. These belong to the core, such as the prepositional objects, as in *raġiba fī šayʾin* ‘he wished for something’, and the prepositional attributes, as in (6).

(6)	*bayna*	*l-muslim-īna*	*fi*
	between	the-Muslim-mp.Gen	in
	l-ʿirāq-i	*wa-l-muslim-īna*	
	the-Iraq-Gen	and-the-Muslim-mp.Gen	
	fī	*jazīrat-i*	*l-ʿarab*
	in	Island-Gen	the-Arabs

‘between the Muslims in Iraq and the Muslims in the Arabian Peninsula’ (Cantarino 1974–1975:II, 81)

Adjuncts that modify the whole sentence, and not just the core or a core argument, are more likely to appear in the beginning of a sentence (cf. Van Valin and Lapolla 1997:166–167), as shown in (7).

(7)	*ḥaqīqat-a*	*lastu*	*ʾafhamu*
	truth-Acc	not.be.1s	understand.1s
	kaṯīr-an		
	much-Acc		

‘Actually, I do not understand much’ (Cantarino 1974–1975:II, 248)

The adverb *kaṯīr-an* modifies the core verb and, hence, naturally follows it.

Holes notes that for Modern Standard Arabic there “seems to be no preferred order for any additional adverbial complements of location or time which follow VSCOMP or VCOMPS”. He adds that “the principle that the longer, ‘heavier’ elements are normally further to the right than the lighter also applies, and that the positioning of an adverbial element at the end is an indication that it may be taken up as the theme of the succeeding text” (Holes 1995:208).

BIBLIOGRAPHICAL REFERENCES

Anshen, Frank and Peter Schreiber. 1968. “A focus transformation of Modern Standard Arabic”. *Language* 44.792–797.

Bloch, Ariel. 1974. *A chrestomathy of Modern Literary Arabic.* Wiesbaden: O. Harrassowitz.

Brown, Gillian and George Yule. 1983. *Discourse analysis.* Cambridge: Cambridge University Press.

Cantarino, Vicente. 1974–1975. *The syntax of modern Arabic prose.* 3 vols. Bloomington: Indiana University Press.

Comrie, Bernard. 1989. *Language universals and linguistic typology.* 2nd ed. Oxford: Basil Blackwell.

Dahlgren, Sven-Olof. 1998. *Word order in Arabic.* Göteborg: Acta Universitatis Gothoburgensis.

Dik, Simon C. 1981. *Functional grammar.* Dordrecht and Cinnaminson: Foris Publications.

Greenberg, Joseph H. 1966. “Some universals of grammar with particular reference to the order of meaningful elements”. *Universals of language*, ed. Joseph H. Greenberg, 73–113. 2nd ed. Cambridge, Mass.: MIT Press.

Halliday, Michael A.K. 1994. *An introduction to functional grammar.* 2nd ed. London, Melbourne, Auckland: Edward Arnold.

Hassan, Motie Ibrahim. 1971. *In-nās wil-malik, ḥikāyāt šaʿbʿiyya bil-lahga l-maṣriyya, mansūxa bil-ḥurūf il-lātīniyya, ṭarīqit in-nasxə wi-mrāgʿit-ha wil-muqaddima.* Copenhagen: Akademisk Forlag.

——. 1981. *ʾAġrāb fi-balad-hum, qiṣaṣ maktūba bil-lahga l-qāhiriyya ʿan il-mugtamaʿ il-maṣri fi-ʾawāʾil il-xamsīnāt, bil-ḥurūf il-ʿarabiyya wil-lātīniyya.* Copenhagen: Akademisk Forlag.

Holes, Clive. 1995. *Modern Arabic: Structures, functions and varieties.* London and New York: Longman.

Hopper, Paul J. 1979. “Aspect and foregrounding in discourse”. *Syntax and semantics*, XII, ed. Talmy Givón, 213–241. New York: Academic Press.

Labov, William and Joshua Waletzky. 1967. “Narrative analysis: Oral versions of personal experience”. *Essays of the verbal and visual arts: Proceedings of the 1966 Annual Spring Meeting of the American Ethnological Society*, ed. June Helm, 12–44. Seattle and London: University of Washington Press.

Longacre, Robert E. 1983. *The grammar of discourse.* London and New York: Plenum Press.

Majdi, Basim. 1990. “Word order and proper government in Classical Arabic”. *Perspectives on Arabic linguistics*, I, ed. Mushira Eid, 127–154. Amsterdam and Philadelphia: J. Benjamins.

Mohammad, Mohammad A. 2000. *Word order, agreement and pronominalization in Standard and Palestinian Arabic.* Amsterdam and Philadelphia: J. Benjamins.

Moutaouakil, Ahmed. 1989 *Pragmatic functions in a functional grammar of Arabic.* Dordrecht: Foris.

Nakano, Aki’o. 1982. *Folktales of Lower Egypt.* I. *Texts in Egyptian Arabic*, 1. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.

Parkinson, Dilworth B. 1981. “VSO to SVO in Modern Standard Arabic: A study in diglossia syntax”. *al-ʿArabiyya* 14.24–37.

Retsö, Jan. 1983. *The finite passive voice in modern Arabic dialects.* Göteborg: Acta Universitatis Gothoburgensis.

Spitta-Bey, Wilhelm. 1880. *Grammatik des arabischen Vulgärdialects von Aegypten.* Leipzig: J.C. Hinrich’sche Buchhandlung.

Van Valin, Robert D. and Randy J. Lapolla. 1997. *Syntax: Structure, meaning and function.* Cambridge: Cambridge University Press.
Versteegh, Kees. 1984–1986. "Word order in Uzbekistan Arabic and Universal Grammar". *Orientalia Suecana* 33–35.443–453.

SVEN-OLOF DAHLGREN
(Göteborg University)

World Wide Web → Internet

X

Xabar

1. THE GRAMMARIANS' USAGE OF THE TERM *XABAR*

The term *xabar*, like many other medieval grammatical terms, developed out of the literal meaning of the word, which is 'message', 'story', etc. The verb *ʾaxbara* means 'to relate something (new) about something'. Right from the early stages of the medieval Arab grammatical tradition, the term *xabar* and its derivatives have been used in the sense of 'predicate', cutting across the two sentence types *jumla ismiyya* and *jumla fiʿliyya*. One often finds the grammarians using *xabar* in the sense of 'predicate', whether to a *mubtadaʾ* (→ *ibtidāʾ*) or to a → *fāʿil*. The underlying principle is that the *xabar*'s function is to predicate of (*ḥaddaṯa, ʾaxbara*) the subject (convey the 'new' information), irrespective of its type. When dealing with sentences such as *fī-hā zaydun qāʾiman* 'in it [fem.] Zayd is standing', Sībawayhi (*Kitāb* I, 222) starts off his discussion using the term *xabar* with reference to *qāʾiman*. Subsequently, he makes the point that *fī-hā* implements the function of *xabar* to *zayd*, analyzing *qāʾiman* as *ḥālun mustaqirrun fī-hā* 'the circumstances of his staying in it'. In later grammarians' writings, the → *ḥāl* in sentences such as *jāʾa zaydun rākiban* 'Zayd came riding' is often described as a secondary *xabar* (*xabar ṯānin*), as opposed to the primary *xabar* - *jāʾa* (see, e.g., Ibn Yaʿīš, *Šarḥ* II, 62).

In many cases, however, the term *xabar* and its derivatives are not clearly differentiated, in the sense that one can hardly determine whether *ʾaxbara*, for instance, is intended to mean 'predicating of'/'functioning as *xabar* to', or rather as 'relating something about something/someone'. Furthermore, *ʾaxbara* and *ʾixbār* are also used in the sense of 'to xabarize' and 'xabarization', respectively; that is, moving a constituent into a *xabar* position, thus assigning it the function of *xabar*. For instance, the *mubtadaʾ zaydun* in *zaydun qāʾimun* 'Zayd is standing' may be moved into the second position in the sentence; in the resulting sentence, *al-qāʾimu zaydun* 'the standing one is Zayd', *zaydun* is the *xabar* (and *al-qāʾimu* is the *mubtadaʾ*). *ʾIxbār* may similarly be implemented through relativization with *allaḏī*: *allaḏī qāma zaydun* 'the one who stood up is Zayd'. For discussion, see Goldenberg (1988:67–69).

Finally, *xabar* may be used to refer to a declarative sentence, as opposed, for instance, to *istifhām*, denoting an interrogative sentence (see, e.g., Jurjānī, *Dalāʾil* 108).

However, there can be no doubt that the most common usage of the term *xabar* in medieval Arabic grammatical literature is in the sense of predicate to a *mubtadaʾ*. Literally, there is no symmetry between the two terms *mubtadaʾ* and *xabar*; one might have expected the counterpart for the latter to be *muxbar ʿan-hu*, which was, indeed, used by later grammarians, though normally in a pragmatic rather than syntactic sense.

2. TYPES OF *XABAR: MUFRAD* VS. *JUMLA*

In his *Bāb al-ibtidā'*, Sībawayhi (*Kitāb* I, 239) distinguishes between three types of *xabar*: a nominal, identical in reference to the *mubtada'* (*mā huwa huwa*), an adverbial of place, and an adverbial of time. Later grammarians draw a distinction between a single phrased *xabar* (*mufrad*) and a clausal *xabar* (→ *jumla*), emphasizing, however, that the basic form (→ *'aṣl*) is the former rather than the latter (see, e.g., Ibn Yaʿīš, *Šarḥ* I, 87–88).

By way of exemplifying the basic form of the *xabar*, the grammarians typically use an active participle such as *qā'imun* 'standing' or *munṭaliqun* 'leaving'. It is always stated that, in cases of disagreement in definiteness between the two predicative constituents, the definite one should, as a rule, come first and be assigned the function of *mubtada'*; the indefinite one should follow and be analyzed as *xabar*. This is consistent with the principle that the *mubtada'* conveys the 'known', whereas the *xabar* is designed to carry the new information conveyed by the sentence. The identity in reference between the *mubtada'* and the *xabar* is reflected in the fact that they normally agree in number and gender, as well as in case: the two share the *rafʿ* case. The grammarians differ with regard to the *ʿāmil* (→ *ʿamal*) assigning *rafʿ* to the *mubtada'* and the *xabar*. But many seem to follow Sībawayhi's theory (*Kitāb* I, 239) that the assigner of *rafʿ* to the *mubtada'* is the → *ibtidā'* and that the *mubtada'*, in its turn, assigns *rafʿ* to the *xabar*, since the latter has the same status (*manzila*) as the former.

The theory of referential identity between *mubtada'* and *xabar* runs into difficulties in two cases: first, when the *xabar* is realized as a clause (*jumla*), and second, when the *xabar* is a locative such as *xalfa-ka* 'behind you' or *fī d-dār* 'in the house'. In the former case, the *xabar* must contain a resumptive pronoun (→ *ʿā'id*) referring to the *mubtada'*, so as to signal the predicatival relationship between the two constituents. The *ʿā'id* is obligatory regardless of whether the *xabar* is a *jumla fiʿliyya* or a *jumla ismiyya*, as is illustrated by the following sentences, respectively: *zaydun qāma 'abū-hu* 'Zayd, his father stood up' and *zaydun 'abū-hu qā'imun* 'Zayd, his father is standing'. As for cases such as *zaydun xalfa-ka* 'Zayd is behind you' or *zaydun fī d-dār* 'Zayd is in the house', the grammarians normally posit an underlying (*muḍmar*; → *'iḍmār*) verb of existence, such as *istaqarra* 'to settle', preceding the locative: *zaydun [istaqarra] fī d-dār*. They claim that the *ʿā'id* in these cases is implicit in the verb. The latter is normally described as the assigner of *naṣb* to the locative.

Regarding the locative occupying the *xabar* position, the grammarians indicate that, while an adverbial of place may be used freely, an adverbial of time is restricted to cases where the *mubtada'* is a verbal noun signaling an event. Thus, *zaydun xalfa-ka* 'Zayd is behind you' and *al-qitālu xalfa-ka* 'the battle is behind you' are both perfectly grammatical sentences. By contrast, a time adverbial such as *ġadan* 'tomorrow' may implement the function of *xabar* to *al-qitāl* (*al-qitālu ġadan* 'the battle is tomorrow'), but not to *zayd*. This apparently universal rule is explained by the grammarians, as could be expected, in terms of communicative value (*fā'ida*). A sentence such as **zaydun al-yawma* 'Zayd is today' lacks communicative value, since the qualification 'today' necessarily applies to all objects (and people) existing at the moment of utterance; 'today' cannot count as a piece of information specific to Zayd, and the sentence is therefore excluded. By contrast, *al-qitālu l-yawma* 'the battle is today' is acceptable, since the battle could potentially take place at any time; by stating the specific time of the battle, the adverbial *al-yawma* in this case provides the addressee with new information, so the sentence is communicatively well formed.

3. THE *XABAR*'S POSITION IN THE SENTENCE

Stipulating that the *xabar* must, in principle, follow the *mubtada'*, the grammarians were well aware of sentence constructions where the reverse is the case. Their discussion normally concentrates upon the following structures:

(1) *rajulun ʿabdu llāhi*
'A man is ʿAbdallāh'
(2) *qā'imun zaydun*
'Standing is Zayd'

(3) *fī d-dāri zaydun*
'In the house is Zayd'
(4) *fī d-dāri rajulun*
'In the house there is a man'

The first of the four cases was treated quite straightforwardly as an inverted version of a *mubtadaʾ* + *xabar* sentence: *rajulun* was analyzed as a preposed (*muqaddam*) *xabar* and *ʿabdu llāhi* as a postposed (*muʾaxxar*) *mubtadaʾ*. The same analysis was also applied to sentences such as *ʾabū-hu qāʾimun zaydun* 'his father is standing, Zayd'; they were described as cases displaying a preposed clausal *xabar*. The other three cases gave rise to an extensive debate among the grammarians. Sentences like (2) were outrightly rejected by the Kufans on the ground that the pronoun implicit in *qāʾimun*, referring forward to *zayd*, constitutes an inadmissible case of → cataphora (*al-ʾiḍmār qabla ḏ-ḏikr*). Some grammarians proposed for (2) and (3) the analysis of (1): a preposed *xabar* followed by a postposed *mubtadaʾ*. Under this analysis, the order *xabar-mubtadaʾ* was presented as a secondary (*farʿ*) formal (*lafḏī*) construction resulting from *taqdīm wa-taʾxīr*, whereas in the basic (*ʾaṣl*) underlying (*maʿnā*) structure, the *xabar* follows the *mubtadaʾ* and not the other way around. This assumption was intended as a solution to two problems attached to *qāʾimun zaydun*: it was a response to the Kufan objection to a cataphoric pronoun, and it dealt effectively with the claim that in *qāʾimun zaydun* the operator (*ʿāmil*) follows the operated-upon constituent (*maʿmūl*), in violation of a major → *ʿamal* principle.

However, other grammarians advocated a different analysis altogether. For them, the main problem was that in (2) the *xabar* position is occupied by a participle, a form recognized by all as akin to the verb. In their view, a participle in presubject position invites an analogy to a verbal sentence displaying a *fiʿl* + *fāʿil* construction. Yet, the vast majority of grammarians were reluctant to analyze *qāʾimun zaydun* as a verbal sentence, because this would be inconsistent with the nominal status of the participle (see, e.g., Sībawayhi, *Kitāb* I, 239). This gave rise to two alternative solutions. In one solution, the verbal analysis of (2) was accepted but restricted to cases where the *xabar* is 'supported' (*yaʿtamidu ʿalā*) by some preceding constituent, such as a noun implementing the function of *mubtadaʾ* or head to a relative clause, or, otherwise, by the interrogative particle *ʾa-* (see, e.g., Sībawayhi, *Kitāb* I, 239; Ibn as-Sarrāj, *ʾUṣūl* I, 60). As an alternative solution, some later grammarians analyzed *qāʾimun* in *qāʾimun zaydun* as *mubtadaʾ*, whereas *zaydun* was presented as *xabar sadda masadd al-fāʿil* (see, e.g., Ibn ʿAqīl, *Šarḥ* I, 189). This, indeed, appears to be a conflation of both the nominal and the verbal analyses.

Similarly, the majority of grammarians analyzed *fī d-dāri zaydun* (3) as an inverted version of *zaydun fī d-dāri*, with a preposed *xabar* and a postposed *mubtadaʾ*. Others, positing the underlying verb *istaqarra* (see above), modeled their analysis on that of a verbal sentence. Still others argued that in such cases it is the locative *fī d-dār* that assigns the *rafʿ* case to *zaydun*. This argument implies, however, that cases such as (3) represent a third type of sentence: *jumla ḏarfiyya*. This was indeed the conclusion of the proponents of this approach, expounded and firmly advocated by Ibn Hišām al-ʾAnṣārī (d. 761/1360; *Muġnī* 492).

The fourth case, *fī d-dāri rajulun*, differs from the third in that the *mubtadaʾ rajulun* is indefinite. This makes the *xabar-mubtadaʾ* order in such cases obligatory. In the grammarians' words, this is a case of an obligatory inversion (*wujūb at-taqdīm*). They all agree that **rajulun fī d-dāri* is unacceptable as an independent sentence (as opposed to a noun phrase with *fī d-dāri* functioning as an attribute (→ *ṣifa*) to *rajulun* (for discussion, see, e.g., Ibn Yaʿīš, *Šarḥ* I, 86). Moreover, some of the later grammarians, going into the pragmatic aspects of *fī d-dāri rajulun*, point out that in this case it is *rajulun* that predicates of *ad-dār* and not the other way around. Ibn ʾAbī r-Rabīʿ (d. 688/1289; *Basīṭ* I, 588) argues that the sentence in question is a shortened version of the underlying (*ʾaṣl*) sentence *ad-dāru maʿmūratun bi-rajulin* 'the house is inhabited by a man'. The constituent *ad-dār*, he maintains, is placed sentence-initially because it is the 'real' topic (*muxbar ʿan-hu*). Some grammarians state clearly that *rajulun* occupies the *xabar* rather than the *mubtadaʾ* position in the sentence (see, e.g., Ibn Yaʿīš, *Šarḥ* I, 86–87). Another case where the *xabar* obligatorily precedes the *mubtadaʾ* is that of interrogative sentences such as *ʾayna zaydun* 'where is Zayd?'. Here, the argument is that interrogative words have a 'right' to occur sentence-initially (*ḥaqq aṣ-ṣadāra*). For a discussion of this and other cases where the *xabar* must obligatorily precede the *mubtadaʾ*, see, e.g., Ibn ʾAbī r-Rabīʿ (*Basīṭ* I, 586–588).

4. The *xabar* in *ʾinna* and *kāna* sentences

The particle *ʾinna* and the auxiliary verb *kāna*, with their respective 'sisters' (→ *ʾinna wa-ʾaxawātuhā*; → *kāna wa-ʾaxawātuhā*), are described by the medieval grammarians as operators (*ʿawāmil*) entering upon (*yadxulna ʿalā*) a *mubtadaʾ-xabar* construction. Later grammarians introduced the terms *xabar ʾinna* and *xabar kāna*, reflecting the idea that the *xabar* in these cases is assigned the *rafʿ* case and the *naṣb* case by *ʾinna* and *kāna*, respectively. A question frequently raised by the grammarians is whether the position of *xabar ʾinna/kāna* is available for any constituent admissible as *xabar* after a *mubtadaʾ* that is not preceded by *ʾinna/kāna*. Zajjājī's (d. 337/949; *Jumal* 42, 53) answer is in the positive. But his position is sharply criticized by his commentator al-Baṭalyūsī (d. 521/1127; *Ḥulal* 165–166, 180–182) as an overgeneralization. First, al-Baṭalyūsī remarks that in both *ʾinna* and *kāna* sentences, a nonaffirmative *xabar* is excluded: while *zaydun hal laqīta-hu* 'Zayd, did you meet him?' is perfectly acceptable, *kāna/ʾinna* sentences do not admit a *xabar* that is a question, a command, a prohibition, an inducement (*taḥḍīḍ*), or a plea (*duʿāʾ*). Next, he points out that while in *kāna* sentences the *xabar* may be a past tense verb, this is disallowed in sentences introduced by *laysa, ṣāra*, and any auxiliary with *mā* preceding the verb; in other cases, such as *kāna, ʾaṣbaḥa*, and *ʾamsā*, it is controversial (unless the past tense main verb is preceded by *qad*). For further discussion, see Ibn ʿUṣfūr (*Šarḥ* I, 380–382) and Ibn ʾAbī r-Rabīʿ (*Basīṭ* II, 681–683), defending Zajjājī's position regarding *xabar kāna*.

Finally, in all sentences containing a *ḍamīr aš-šaʾn*, the latter is analyzed by the medieval grammarians as *mubtadaʾ* and the clause following it as a clausal *xabar*. They explain that the *ḍamīr aš-šaʾn* is a 'vague' (*mubham*) pronoun referring to the whole 'matter' described by the following *xabar* (→ *ḍamīr*). The latter serves as an exponent (*tafsīr*) of the *mubtadaʾ*. Thus, for instance, in *ʾinnahu ʾamatu llāhi ḏāhibatun* 'the female slave of God is going', the clause *ʾamatu llāhi ḏāhibatun* is analyzed as an exegetical *xabar* to the 'vague' pronoun in *ʾinnahu*. For a detailed discussion, see, e.g., Ibn Yaʿīš (*Šarḥ* III, 114–118).

Bibliographical references

Primary sources

Baṭalyūsī, *Ḥulal* = ʾAbū Muḥammad ʿAbdallāh ibn Muḥammad ibn as-Sayyid al-Baṭalyūsī, *Kitāb al-ḥulal fī ʾiṣlāḥ al-xalal min Kitāb al-jumal*. Ed. Saʿīd ʿAbd al-Karīm Saʿʿūdī. Baghdad: Dār ar-Rašīd li-n-Našr, 1980.

Ibn ʾAbī r-Rabīʿ, *Basīṭ* = ʿUbaydallāh ibn ʾAḥmad ibn ʿUbaydallāh Ibn ʾAbī r-Rabīʿ, *al-Basīṭ fī šarḥ Jumal az-Zajjājī*. Ed. ʿAyyād ibn ʿĪd aṯ-Ṯabītī. Beirut: Dār al-Ġarb al-ʾIslāmī, 1986.

Ibn ʿAqīl, *Šarḥ* = Bahāʾ ad-Dīn ʿAbdallāh Ibn ʿAqīl, *Šarḥ Ibn ʿAqīl ʿalā ʾAlfiyyat Ibn Mālik*. Ed. Muḥammad Muḥyī d-Dīn ʿAbd al-Ḥamīd. N.p.: Dār Ṣaʿb, n.d.

Ibn Hišām, *Muġnī* = Jamāl ad-Dīn ʾAbū Muḥammad ʿAbdallāh ibn Yūsuf Ibn Hišām al-ʾAnṣārī, *Muġnī l-labīb ʿan kutub al-ʾaʿārīb*. Ed. Māzin al-Mubārak and Muḥammad ʿAlī Ḥamdallāh. Beirut: Dār al-Fikr, 1985.

Ibn as-Sarrāj, *ʾUṣūl* = ʾAbū Bakr Muḥammad ibn Sahl Ibn as-Sarrāj, *al-ʾUṣūl fī n-naḥw*. Ed. ʿAbd al-Ḥusayn al-Fatlī. Beirut: Muʾassasat ar-Risāla, 1987.

Ibn ʿUṣfūr, *Šarḥ* = ʾAbū l-Ḥasan ʿAlī ibn Muʾmin ibn Muḥammad Ibn ʿUṣfūr, *Šarḥ Jumal az-Zajjājī*. Ed. Ṣāḥib ʾAbū Janāḥ. Mosul: ʾIḥyāʾ at-Turāṯ al-ʾIslāmī, 1980–1982.

Ibn Yaʿīš, *Šarḥ* = Muwaffaq ad-Dīn Yaʿīš ibn ʿAlī Ibn Yaʿīš, *Šarḥ al-Mufaṣṣal*. Cairo: Maktabat al-Mutanabbī, n.d.

Jurjānī, *Dalāʾil* = ʾAbū Bakr ʿAbd al-Qāhir ibn ʿAbd ar-Raḥmān al-Jurjānī, *Dalāʾil al-ʾiʿjāz fī ʿilm al-maʿānī*. Ed. Muḥammad Rašīd Riḍā. Beirut: Dār al-Kutub al-ʿIlmiyya, 1988.

Sībawayhi, *Kitāb* = ʾAbū Bišr ʿAmr ibn ʿUṯmān Sībawayhi, *al-Kitāb*. Ed. Hartwig Derenbourg. Repr., Hildesheim and New York: G. Olms, 1970.

Zajjājī, *Jumal* = ʾAbū l-Qāsim ʿAbd ar-Raḥmān ibn ʾIsḥāq az-Zajjājī, *Kitāb al-jumal fī n-naḥw*. Ed. ʿAlī Tawfīq al-Ḥamad. Beirut: Muʾassasat ar-Risāla and Dār al-ʾAmal, 1988.

Secondary source

Goldenberg, Gideon. 1988. "Subject and predicate in Arab grammatical tradition". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 138:1.39–73.

Yishai Peled (Tel-Aviv University)

X-Bar Syntax

1. Introduction: The X-bar schema

Early generative grammar, from the late 1950s to the late 1960s, factored grammatical descriptions into two main components: the *base* component, consisting of (the lexicon and) phrase-structure rules, and the → *transformational* component. X-bar syntax is the designation for a theory of phrase structure that originated in the context of an intramural debate within the generativist camp in the late 1960s concerning the proper way to characterize relationships among expressions like (1a,b,c).

(1) a. *John proved the theorem*
b. *John's proving of the theorem*
c. *John's proof of the theorem*

(1b) is a gerund (*-ing*) nominal, and (1c) a derived nominal, a noun derived from a verb. At issue was whether both (1b) and (1c) should be derived transformationally from the deep structure corresponding to sentence (1a). Chomsky (1970) argued that while gerund nominals like (1b) derived by transformations from deep structure sentences, derived nominals like (1c) were formed in the lexicon by ordinary processes of derivational morphology. The ground for this conclusion was that derived and gerund nominals differ markedly in three areas: morphological productivity, semantic regularity, and internal syntactic structure. Gerund nominals are highly productive and are semantically compositional; structurally, they are clauses. Derived nominals, by contrast, are relatively unproductive and semantically idiosyncratic, and structurally they are noun phrases (Chomsky 1970:187–200). Chomsky dubbed his position the 'lexicalist' hypothesis, in contrast to the 'transformationalist' approach.

In distinguishing between the sources of the two types of nominals, the lexicalist hypothesis left open the question of how to capture the similarities between them and sentence (1a) that motivated the transformationalist analysis. Toward this end, Chomsky proposed two innovations to phrase-structure theory. The first was to analyze syntactic categories as bundles of syntactic features, exploiting an idea first proposed in Chomsky (1965:Chap. 2). The second built on the observation that the three lexical categories noun (N), verb (V), and adjective (A) could all take complements. In (1a), the verb *proved* takes the complement *the theorem*; in (1c), the noun *proof* takes the complement *(of) the theorem*; and in *afraid of theorems, afraid* takes the complement *(of) theorems*. To capture this cross-categorial generalization, Chomsky (1965:210) proposed the generalized phrase-structure schema in (2), in which X is a category variable ranging over N, V, and A.

(2) X' → X…

(Here and throughout, the standard practice of using prime notation instead of bars will be followed, while retaining the name 'X-bar syntax'.) (2) says that a sequence of an N, V, or A and its complement(s), indicated by "…", forms a single-bar constituent with X as its head. An X-bar can in turn combine with categories that precede the head, referred to as its specifiers, to form a noun phrase (NP), verb phrase (VP), or adjective phrase (AP) by a second schema for X-double bar, as in (3).

(3) X'' → [Spec, X'], X'

Thus, N'' = NP, V'' = VP, and A'' = AP. The categories that function as specifiers depend on the category of the head: determiners like *the* for N, adverbs like *hardly* for V, and degree words such as *very* for A.

The base rule for generating a sentence (S) was (4).

(4) *S* → N'' V'' (= S → NP VP)

At this stage, the S rule stood outside the X-bar schema. While phrases as defined by (2) and (3) were endocentric (the phrase being of the same category type as its head), sentences were still exocentric (having no head and being of a different category from either of its immediate constituents). In addition to endocentricity, phrases were subject to several other conditions (Emonds 1976:12–20; Webelhuth 1995:23; Koizumi 1995:137):

i. peripherality, the requirement that the head X be initial (or, later, final) within X';
ii. maximality, the requirement that specifiers and complements be X''s;
iii. binarity, the requirement that an X'' or X' node dominate no more than two other nodes;
iv. asymmetricity, the requirement that a phrase be projected from only one of its daughters.

2. Revision and application of the X-bar schema

Jackendoff (1977) revised and extended the X-bar schema of Chomsky (1970) in two main ways. First, he built (1977:31) on Chomsky's suggestion about using syntactic features to define four universal categories, N, V, A, and P (preposition), in terms of binary (+ or –) values for the categorial features [N, V], as in (5).

(5) a. N: [+N, –V]
b. V: [–N, +V]
c. A: [+N, +V]
d. P: [–N, –V]

This analysis was retained in later work without modification (Chomsky 1981:48, 1986b:2). A proposal by Adnan (1993:186) to extend [+N] to nonlexical categories such as D (determiner) has not gained acceptance. Second, Jackendoff (1977:36) proposed that the maximal number of bar (prime) levels for all four categories in (5) was three, designating this as the Uniform Three-Level Hypothesis.

Al-Aboudi (1985) employs both these innovations in his study of the → construct state (CS) in (Modern) Standard Arabic. His tree diagrams for construct states rely on feature matrices for syntactic categories (1985:136–137), and he uses the Uniform Three-Level Hypothesis to show that the definite article *al-* and the nominal suffix *-n* (→ *tanwīn*) in Modern Standard Arabic are not in complementary distribution as markers of (in)definiteness and, consequently, project to different bar levels (1985:138–148). His evidence comes from quantified construct states with numbers, as in (6), and with *kull* 'all' or 'every', as in (7).

(6) *ištaray-tu* *l-ʾarbaʿīna* *kitāb-in*
bought-1s the-forty book-Gen
'I bought the forty books'

(7a) *ištar-ā* *sālim* *kull-a* *baqarat-in*
bought-3ms Salim all-Acc cow-Gen
'Salim bought every cow'
(7b) *ištar-ā* *sālim* *kull-a*
bought-3ms Salim all-Acc
l-baqarat-i
the-cow-Gen
'Salim bought the whole cow'

In (6), definite *al-* and indefinite *-n* occur in the same phrase, which is possible because *al-* is generated in [Spec, N''] and *-n* is Det''', with the entire NP projection being N''' (cf. Al-Aboudi 1985:144). In (7), the difference in interpretation between 'every' in (a) and 'all' in (b) is not simply a function of definiteness, which indicates that the determiners are [Spec, N''] and Det''', respectively.

Unlike decomposition of syntactic categories into feature bundles, the Three-Level Hypothesis did not gain wide support. On the one hand, most analysts followed Chomsky in taking the upper limit on phrasal projections to be two. On the other hand, with the advent of functional categories (see below), certain theories left the number of bar levels indeterminate, either because lexical categories were argued to be able to iterate specifiers indefinitely at the X' level (Fukui 1995:11), or because bar level was removed from the set of primitives for X-bar theory (Speas 1990:39–42). With the advent of bare phrase structure, and its elimination of bar levels from syntactic representation (Chomsky 1995c:397–400; 1995d:243–246), the issue has become moot (→ minimalism).

One aspect of the Three-Level Hypothesis did stimulate fruitful research into the outstanding problem of the exceptional status of S. At the time, current work sought to minimize the exocentricity of S by adding X-bar schematic variants to the inventory of S-expansion rules in order to accommodate complement clauses and topics (Chomsky 1977:91):

(8) a. S'' → TOP S'
b. S' → COMP S

Studies of both root and embedded → topicalization in Egyptian Arabic made use of such phrase-structure rules (Farghaly 1981:120–124; Farghal 1986:82–112). Jackendoff (1977:41) proposed that S should be integrated into the X-bar schema as V'''. With the general acceptance of a two-bar system, some alternative was needed. In discussing the S-expansion rule, Chomsky (1981:51–52) speculated that S might properly be analyzed as a projection of the head I[inflection] and S' as a projection of the head C[omplementizer]. Chomsky (1986a) took this step for I, asserting S = I'' = IP (Inflection Phrase) (1986a:160–161). Chomsky (1986b:3) completed the assimilation, taking C to head the maximal projection C'' = CP (Complementizer Phrase). IP and CP were designated as functional categories to distinguish them from lexical categories such as NP and VP. (For discussion of the complementizer *ʾanna* in Standard Arabic, see Shlonsky 2000).

Once IP and CP were recognized, they proved unstable, and a plethora of new functional projections emerged (see Webelhuth 1995:76 for an enumeration). An early and extremely influential catalyst for this proliferation was Pollock (1989), who elegantly explained a complicated set of differences in the order of finite and nonfinite verbs relative to negation and adverbs in French and English by decomposing IP into TP, T[ense] Phrase, and AgrP, Agreement Phrase, and adding NegP (Negative Phrase) between them (Pollock 1989:366–384, 397). Chomsky (1995b:147) then divided AgrP into subject (S) and object (O) agreement projections: Agr$_S$P and Agr$_O$P, the former dominated by TP and the latter by NegP. The articulated clause structure that resulted is given in (9).

(9) [TP [T [AGR$_S$P [AGR$_S$ [NEGP [NEG [AGR$_O$P [AGR$_O$ [VP…]]]]]

Subsequent work has further decomposed Agr$_O$P and VP and countenanced layered specifiers (Koizumi 1995:Chaps. 2, 5, 6).

The recognition of IP also stimulated intensive inquiry into the structure of NP, as researchers noticed that the parallelism between noun phrases and sentences exemplified in (1) no longer held under the IP analysis. In the X-bar theoretic analysis of derived nominals, the subject of the deverbal noun was in [Spec, NP] and its complement under N' with the N head, so that all the arguments of the head were contained within its maximal projection (Ritter 1988:911), as in (10).

(10) [NP [NP the army's] [N' [N destruction [NP the city]]]

Under the IP analysis, by contrast, the subject of the clause *the barbarians destroyed the city* does not appear in [Spec, VP], the position which should correspond to [Spec, NP] in (10). Conversely, there is no functional projection in the NP corresponding to IP in the clause (cf. Bernstein 2003:536). To rectify the latter problem, Abney (1987) and Stowell (1989) proposed that NPs be reanalyzed as DPs (Determiner Phrases) headed by a (sometimes phonetically null) D head taking an NP complement. (10) would then be assigned the structure in (11) (Ritter 1988:913).

(11) [DP [the army] [D' [D -s] [NP [N' [N destruction] [DP [D the] [NP city]]]]]]

This solution to the latter problem exacerbated the former one, for now neither the subject of the DP nor the clause is in the maximal projection of the lexical head (*destruction* or *destroy*, respectively). To correct this, it was proposed that the subject in each case originates in the specifier of a lexical head and raises to the specifier of the functional head that dominates it. Thus, in (11), *the army* originates in [Spec, NP] and raises to [Spec DP], while in the analogous clause it originates in [Spec, VP] and raises to [Spec, IP] (Sportiche 1988:425). This widely accepted thesis, that all the arguments of a lexical head are contained within its maximal projection, is known as the Lexical Clause Hypothesis (Plunkett 1996) or the 'VP-Internal Subject Hypothesis' (Webelhuth 1995:60–64).

3. Functional categories in Arabic

Of the various functional projections proposed in the literature, the following have had the greatest impact on the study of Arabic: TP, AgrP, NegP, DP. In addition, Mood Phrase (MoodP) and Topic Phrase (TopP) have been proposed. Debates in this area have mainly turned on three issues: (i) the syntactic motivation for functional projections; (ii) their hierarchical order; and (iii) their proper structural characterization.

The existence of TP and NegP in Arabic is widely accepted (cf. M. Bahloul 1996; R. Bahloul 1996). Arguments for their existence are based on their syntactic and semantic saliency and, for TP, evidence of its presence in the absence of a lexical host. The discontinuous bound morpheme *ma-...š* 'not' is the syntactic head Neg in Moroccan Arabic, despite its affixal status, because it creates scopal ambiguities, as in (12a), and licenses negative → polarity items, as in (12b) (Benmamoun 1992:10–11).

(12a) *ma-qra-š* *bezzāf*
not-read.3ms-Neg many
d-le-ktuba
of-the-books
'He did not read many of the books' (negation > quantifier)
'Many are the books that he did not read' (quantifier > negation)
(12b) *ma-qra* *ḥatta* *ktāb*
not-read.3ms any book
'He did not read any book'

Similarly, Tense heads a syntactic projection because it licenses temporal adverbs even in the absence of an overt present tense auxiliary (which does not exist in Arabic) with the participle *ġādi* 'going' (Benmamoun1992:12–15), as in (13).

(13a) *ʿomar* *ġādi* *lmadrasa* *daba*
Omar going to-school now
'Omar is going to school now'
(13b) *ʿomar* *kān* *ġādi* *lmadrasa*
Omar was going to-school
lbāriḥ
the-yesterday
'Omar was going to school yesterday'

Benmamoun (2000:14) proposes that copular sentences have only the feature [+D] (determiner), while those with *kān* have the feature complex [+V, +D].

As for the relative order of TP and NegP, Benmamoun asserts, in a much-cited argument, that Neg is situated between T and V because it blocks an imperfect verb form from adjoining to T under the Minimality Condition on movement (Rizzi 1990:7); instead, Neg itself raises to amalgamate with T to yield a tensed negative, with present, past, or future time reference (Benmamoun 1992:156, Ouhalla and Shlonsky 2002:10; → negation), as in (14a,b,c).

(14a) *lā* *ya-ktub-u*
not 3ms-write-Ind
'He does not write'
(14b) *lam* *ya-ktub-ø*
not 3ms-write-Juss
'He did not write'
(14c) *lan* *ya-ktub-a*
not 3ms-write-Subj
'He will not write'

(14) agrees with (9) in generating NegP below TP, but says nothing about the position of AgrP. In evaluating the six possible orders of Neg, Agr, and T, Benmamoun (1992:22–30) concludes that only the order T-Neg-Agr respects all grammatical constraints, in accord with Pollock's original proposal. Ouhalla (1991:57) concurs but also presents comparative evidence that the order T-Agr is not universal, noting that the converse order occurs in Italian, as in (15b).

(15a) *sa-ya-štar-ī* *zayd* *dār*
Fut-3ms-buy-Ind Zayd house
'Zayd will buy a house'
(15b) *legge-va-no*
read-impf-3p
'They read'

In Arabic, Agr = *ya-* appears inside Tense = *sa-* in Arabic, but in Italian, Tense = *-va-* occurs inside Agr = *-no*. Ouhalla presents similar data for variation in the order of Neg relative to T and Agr. On the strength of these observations, he proposes (1991:67–69) that the order of functional projections is parametrized. This assumption is explicit in some current work (e.g. Al-Harbi 2000), although the issue seldom arises.

The contrary view that the hierarchy of functional categories is fixed by Universal Grammar (UG) is defended by Shlonsky (1997:3–4) on the grounds of descriptive parsimony and learnability. Contrary to Pollock and Chomsky, Shlonsky supposes Agr$_S$P to be higher than TP. He argues (1997:98) that the agreement features on the imperfect verbs in (14) are not Agr$_S$ features but rather AgrPart[iciple] features, on the basis of several types of evidence. Sentences in Modern Standard Arabic with the finite auxiliary *kāna* '[he] was' (cf. [13b]) exhibit subject agreement features (1997:97, 99), as does the imperfect main verb.

(16) *kān-a* *l-walad* *ya-lʿab-u*
be.Perf-3ms the-boy 3ms-play-Ind
'The boy was playing'

Minimally, such sentences require postulating two Agr projections. The lower Agr is AgrPartP because the imperfect distributes with participles and nonverbal predicates in co-occurring with *laysa* 'is not' (Shlonsky 1997:105–106), as in (17).

(17a) *mōna lays-at fī l-bayt*
Mona neg-3fs in the-house
'Mona is not in the house'
(17b) *mōna lays-at ḍārib-a zayd*
Mona neg-3fs hitting-3fs Zayd
'Mona is not hitting Zayd'
(17c) *mōna lays-at ta-ḍrib-u*
Mona neg-3fs 3fs-hit.Imperf-Ind
zayd
Zayd
'Mona does not hit Zayd'

The upper Agr is Agr_SP because it may host an initial subject with negated verbs in the perfect but not the imperfect (Shlonsky 1997:107, with second example modified):

(18a) **zaynab mā ta-ḍrib-u*
Zaynab-Nom not 3fs-hit.Imperf.-Ind
zayd
Zayd
'Zaynab does not hit Zayd'
(18b) *zaynab mā ḍarab-at zayd*
Zaynab-Nom not hit.Perf-3fs Zayd
'Zaynab did not hit Zayd'

Ouhalla and Shlonsky's analyses assume that Agr projects syntactically in Arabic. This has, in fact, been a controversy on which consensus has not yet been achieved. Benmamoun (1993:66–67) argues that agreement differs from tense and negation in that it is semantically inert and that its syntactic function, mediating between tense and the subject to check the nominative Case of the latter, is unrealized in three contexts in Standard Arabic: (i) negated sentences like those in (14); (ii) equational sentences like (19a); and (iii) DPs with posthead quantifiers that include a clitic agreeing with the head of NP like (19b).

(19a) *aṭ-ṭālib-a fī l-bayt*
the-student-fs in the-house
'The [female] student is in the house'
(19b) *ra'ay-tu l-ʾawlād-a*
saw-1s the-children-Acc
kull-a-hum
all-Acc-them
'I saw all the children'

In (19a), nominative Case is assigned to the subject *aṭ-ṭāliba* in the absence of Agr; in (19b), accusative Case is checked by the verb *ra'aytu*, while → agreement in Case and person-number holds between *al-ʾawlād-a* and the floated quantifier *kullahum*. In addition, LeTourneau (1998:99–106) argues that Agr_O must be absent from the lower VP of double-accusative causative constructions like *hiya kattabat al-ʾawlāda d-darsa* 'she made the boys write the lesson' under Minimalist assumptions about clause structure, economy, and feature checking.

Arguments in favor of the existence of AgrP assume or seek to complete the parallelism discussed in connection with (10)-(11) above between TPs and DPs, specifically, → construct state (CS) DPs (→ case theory). The standard analysis of a construct state DP such as *bayt-u l-mudarris-i hunā* 'the [male] teacher's house is here' is given in (20) (Fassi Fehri 1993:215–216; Benmamoun 1999:624; 2003:748–751; Mohammad 1999:33).

(20a) [DP [D] [NP [DP al-mudarris-]
[N bayt-]]]
(20b) [DP [D [N bayt-u_1] [NP [DP l-mudarris-i]
[N t_1]]]

The NP complement of D is headed by the possessee *bayt-* and contains in its specifier the DP possessor *al-mudarris-*. The surface word order is derived by N-D raising. N checks its Case against some higher inflectional head, and the DP in [Spec NP] has its genitive Case by the N head.

Mohammad (1999) argues that construct state DPs parallel sentences in exhibiting both partial and full agreement. In sentences, VS order exhibits partial agreement between verb and subject, excluding number, while SV order exhibits full agreement (Mohammad 1990:95–96; Bolotin 1995:7–8; Benmamoun 1998:112; LeTourneau 2003:85–86; among others). In construct state DPs, the first and second terms exhibit partial definiteness agreement, since the first term lacks (in)definiteness morphology but agrees in definiteness with the second (28). They also exhibit full agreement in φ-features – gender, number, and Case (→ Case theory) – between N and POSS (*tabaʿ* or *ʾabu*) in possessive constructions like those in (21), from Palestinian Arabic, which Mohammad (1999:38, 40) takes to be construct states rather than relative clauses, since at least (21a) cannot include the relative marker *(ʾe)lli.*

(21a) *lektāb (*elli) tabaʿ*
the-book which belong-to.ms
elbent
the-girl
'the book belonging to the girl'

(21b) *elwalad abu lgamīṣ leḥmar*
the-boy father the-shirt the-red
'the boy with the red shirt'

Mohammad (1999:40) proposes that both types of construct state contain at least one Agr projection between D and NP, as in (22).

(22) [DP [D [AGRP AGR [NP [DP lektāb] [N tabaʿ] [DP elbent]]]]]

This node accounts for agreement in φ-features between *lektāb* and *tabaʿ*. Mohammad (1999:39) associates the Agr projection in (22) with what Ritter (1991) calls Num[ber]P[hrase]. Ritter (1991:55) analyzes quantified construct states in Modern Hebrew as consisting of a Q (quantifier) such as *kol* 'all', a Num head which takes no NP complement. LeTourneau (1995:40–41) adopts this analysis for Modern Standard Arabic and proposes that expressions like *kullu l-mudarrisīna* 'all the teachers' are generated by Q-D raising, parallel to N-D raising in nonquantified construct states (cf. Benmamoun 1999).

Two other functional projections have figured in recent work: TopP and MoodP. Al-Shorafat (1999) seeks to overcome defects of earlier analyses of → topicalization, such as those of Bakir (1980), Farghaly (1981), and Abdul-Ghany (1981), with a Minimalist analysis in which TopP is left-peripheral to CP (cf. Rizzi 1997) and topics raise from internal to TP to [Spec, TopP] to check their formal features. Plunkett (1993a,b) argues from the modal distinction among imperfect verb forms, illustrated in (14), that MoodP projects between NegP and Agr$_{O}$P. Ayed (2003:30–37) offers very interesting arguments that *ʾan*, which is conventionally classified as a subjunctive complementizer (Aoun 1985:56–57), actually heads the category Modal Phrase, on the basis of its selectional similarities to the future negative particle *lan* and its complementary distribution with that particle.

4. CLAUSAL MODIFIERS IN ARABIC

X-bar theory in the 1970s assumed that specifiers are left-peripheral phrasal modifiers to a head X and that, in addition to complements, posthead phrasal modifiers called 'adjuncts' occur right-peripherally, such as *hard* in *John hit the ball hard.* Two innovations came to call into question the distinction between specifiers and adjuncts. One was to factor out linear order from the X-bar schema, limiting its function to specifying hierarchical relations among constituents and leaving → word order to be determined by other modules of grammar, such as → Case theory or θ-theory (Travis 1989; → theta roles). The other was Kayne's (1994:6) antisymmetry proposal, the core of which is the Linear Correspondence Axiom (LCA), represented in (23).

(23) d(A) is a linear ordering of T

A is the set of nonterminal symbols in a grammar, d(A) the set of terminal symbols (words) dominated by the members of A in a syntactic tree diagram, and T the set of terminal symbols. Assuming asymmetric c-command, (23) states that if a nonterminal node X asymmetrically c-commands another nonterminal Y, then the word(s) dominated by X will linearly precede the one(s) dominated by Y. In addition to deriving a number of axioms of X-bar theory as theorems of the Linear Correspondence Axiom, Kayne (1994:17) deduced that specifiers and adjuncts are the same, since the latter cannot be generated to the right of a head under the Linear Correspondence Axiom. (For a relativized version of antisymmetry, see Moro 2000.)

Fassi Fehri (1998:32–33) appeals to the Linear Correspondence Axiom for his conclusion that both prenominal and postnominal adjectives in Arabic are generated in the specifiers of functional heads generated to the left of their heads instead of on right branches. The Linear Correspondence Axiom also pertains to the analysis of two types of clausal postmodifiers: restrictive → relative clauses and adverbial (circumstantial) clauses (→ *ḥāl*).

In the Modern Standard Arabic relative clause in (24), *al-fatātu* is a relativized direct object.

(24) *al-(fatāt-u [llatī ʾu-ḥibb-u-hā])*
the-girl-Nom that 1s-love-Ind-her
'the girl that I love (her)'

Restrictive relatives have been subject to two phrase-structure analyses. On the adjunct analysis, the bracketed string (the relative clause proper) is adjoined to the DP *al-fatātu*. On the complement analysis, the material in parentheses is a CP complement to the (overt or covert) D head of the entire DP (Choueiri 2002:193). Under the Linear Correspondence Axiom, the complement analysis is the only option, since right adjuncts are excluded. Choueiri designates this as the 'External Determiner' (ED) hypothesis.

The External Determiner hypothesis is supported by several strands of evidence. First, in contrast to a nonrestrictive relative, a restrictive relative is interpreted in the scope of its determiner. This is evident from how their entailments diverge, as in (25).

(25a) *Mary knows few boys who enjoy knitting*
(25b) *Mary knows few boys, who enjoy knitting*
(25c) *Mary knows few boys*

Only the nonrestrictive (25b) entails (25c); the restrictive (25a) does not (Choueiri 2002:199). Second, the subject of an existential *there* sentence cannot be definite, yet it can appear in a definite relative clause, as in (26), suggesting that the definite determiner is not a constituent of the postverbal subject of the existential (Aoun and Li 2003:103).

(26a) **There were the men in the garden*
(26b) *The men*$_1$ *there were* t_1 *in the garden were all diplomats*

For discussion of contrary evidence associating D with the raised NP, see Ouhalla (2004).

If the D *al-* in (24) originates outside the relative clause, the question arises as to the status of *fatātu*. Choueiri (2002:Chap. 2) and Aoun and Li (2003:127–128) provide evidence, based on reconstruction, that the nominal head of a definite relative in Lebanese Arabic undergoes raising out of the relative clause proper to [Spec, NumP]. An N head cannot raise to a specifier by itself, since specifiers host only maximal projections. Nor can it be an NP, since [Spec, CP] hosts DP arguments, as shown by Lebanese Arabic **l-piano lli blˁab-o* 'the piano that I play it', which is ungrammatical (on the relevant reading) because *piano* is nonreferential; cf. *I play piano* (Choueiri 2002:226–227). Therefore, the phrase that moves to [Spec, CP] is a DP with a null determiner, which incorporates into the definite D heading the relative clause (Choueiri 2002:227; Aoun and Li 2003:129).

Adverbial clauses, known as → *ḥāl* 'condition' clauses in traditional Arabic grammar (Wright 1974:II, 330), are exemplified by (27).

(27) *waṣal-a l-wazīr (wa-hwa)*
arrived-3ms the-minister and-he
ya-ḥmil-u kitāb
3ms-carry-Ind book
'The minister arrived [while] carrying a book'

In Classical Arabic, the circumstantial clause follows the main clause; in Standard Arabic, it now occasionally precedes (Holes 2004:266, 270). The same holds true in colloquial Arabic, in which initial circumstantial clauses serve as topics to frame the main clause, as in (28) from Egyptian Arabic (Brustad 2000:339–340).

(28) *wa humma mašy-in nisi*
and they going-3p forgot.3ms
šanṭit-u maˁā-ha
bag-his with-her
'When they were leaving, he forgot his briefcase with her'

Circumstantials have been little studied by generative linguists, Mohammad (1991) being an exception, so their phrase structure is poorly understood. The Linear Correspondence Axiom would seem to require that the base order be the one in (28), with raising of the main clause to the specifier of some functional projection to generate (27), subject, perhaps, to lexical semantic properties of the main verb (Abboud 1986).

Bibliographical references

Abboud, Peter. 1986. “The *ḥāl* construction and the main verb in the sentence”. *The Fergusonian impact: In honor of Charles A. Ferguson on the occasion of his 65th birthday*. I. *From phonology to society*, ed. Joshua A. Fishman a.o., 191–196. New York: Mouton de Gruyter.

Abdul-Ghany, Mohammad Kamal Eldeen. 1981. *Government binding in Classical Arabic*. Ph.D. diss., University of Texas at Austin.

Abney, Stephen. 1987. *The English noun phrase in its sentential aspect*. Ph.D. diss., MIT.

Adnan, George. 1993. “Functional heads and Case assignment in NPs in Standard Arabic”. *Dilbilim araştırmaları* 175–188. Ankara: Hitit Yayınevi.

Al-Aboudi, Fahmi Basil Shukri. 1985. *The syntax of* “idaafah”*: Annexation in Arabic*. Ph.D. diss., Indiana University.

Al-Harbi, Abdullah. 2000. “The NEG elements in clause structure”. *Syntaxis* 3.45–73.

Al-Shorafat, Mohammad O. 1999. “Topics in Arabic: A Minimalist approach”. *al-ʿArabiyya* 32.1–21.

Aoun, Joseph. 1985. *The grammar of anaphora*. Cambridge, Mass.: MIT Press.

—— and Audrey Li. 2003. *Essays on representation and derivation in grammar: The diversity of* wh-*constructions*. Cambridge, Mass.: MIT Press.

Ayed, Hela Ben. 2003. *Mood and functional projections*. Ph.D. diss., McGill University.

Bahloul, Maher. 1996. “Extending the NegP Hypothesis”. Eid (1996:31–46).

Bahloul, Raja Ballek. 1996. “Negation in French and Tunisian Arabic”. Eid (1996:67–83).

Bakir, Murtadha J. 1980. *Aspects of clause structure in Arabic: A study of word order in Literary Arabic*. Bloomington, Ind.: Indiana University Linguistic Club.

Baltin, Mark R. and Anthony S. Kroch (eds.). 1989. *Alternative conceptions of phrase structure*. Chicago: University of Chicago Press.

Benmamoun, Elabbas. 1992. *Inflectional and functional morphology: Problems of projection, representation, and derivation*. Ph.D. diss., University of Southern California.

—— 1993. “The status of agreement and the agreement projection in Arabic”. *Studies in the Linguistic Sciences* 23.61–71.

—— 1998. “Spec-head agreement and overt Case in Arabic”. *Specifiers: Minimalist approaches*, ed. David Adger, Susan Pintzuk, and Bernadette Plunkett, 110–125. Oxford: Oxford University Press.

—— 1999. “The syntax of quantifiers and quantifier float”. *Linguistic Inquiry* 30.623–642.

—— 2000. *The feature structure of functional categories: A comparative study of Arabic dialects*. Oxford: Oxford University Press.

—— 2003. “Agreement parallelism between sentences and noun phrases”. *Lingua* 113.747–764.

Bernstein, Judy. 2003. “The DP hypothesis: Identifying clausal properties in the nominal domain”. *Contemporary syntactic theory*, ed. Mark Baltin and Chris Collins, 536–561. New York: Blackwell.

Bolotin, Naomi. 1995. “Arabic and parametric VSO agreement”. Eid (1995:7–27).

Brustad, Kristen E. 2000. *Spoken Arabic: A comprehensive study of Moroccan, Egyptian, Syrian, and Kuwaiti dialects*. Washington, D.C.: Georgetown University Press.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Mass.: MIT Press.

—— 1970. “Remarks on nominalization”. *Readings in English transformational grammar*, ed. Roderick Jacobs and Peter Rosenbaum, 184–221. Waltham, Mass.: Ginn.

—— 1977. “On *wh*-movement”. *Formal syntax*, ed. Peter Culicover, Thomas Wasow, and Adrian Akmajian, 77–132. New York: Academic Press.

—— 1981. *Lectures on government and binding*. Dordrecht: Foris.

—— 1986a. *Knowledge of language: Its nature, origin, and use*. New York: Praeger.

—— 1986b. *Barriers*. Cambridge, Mass.: MIT Press.

—— 1995a. *The Minimalist Program*. Cambridge, Mass.: MIT Press.

—— 1995b. “Economy of derivation and representation”. Chomsky (1995:129–166).

—— 1995c. “Bare phrase structure”. Webelhuth (1995:383–439).

—— 1995d. “Categories and transformations”. Chomsky (1995a:218–394).

Choueiri, Lina. 2002. *Issues in the syntax of resumption: Restrictive relatives in Lebanese Arabic*. Ph.D. diss., University of Southern California.

Eid, Mushira (ed.). 1995. *Perspectives on Arabic linguistics*, VII. Amsterdam and Philadelphia: J. Benjamins.

—— (ed.). 1996. *Perspectives on Arabic linguistics*, VIII. Amsterdam and Philadelphia: J. Benjamins.

Emonds, Joseph. 1976. *A transformational approach to English syntax: Root, structure-preserving, and local transformations*. New York: Academic Press.

Farghal, Mohammad Ali. 1986. *The syntax of* wh-*questions and related matters in Arabic*. Ph.D. diss., Indiana University.

Farghaly, Ali Ahmed Sabry. 1981. *Topics in the syntax of Egyptian Arabic*. Ph.D. diss., University of Texas at Austin.

Fassi Fehri, Abdelkader. 1993. *Issues in the structure of Arabic clauses and words*. Dordrecht: Kluwer Academic Publishers.

—— 1998. “Layers in the distribution of Arabic adverbs and adjectives and their licensing”. *Perspectives on Arabic linguistics*, XI, ed. Elabbas Benmamoun, Mushira Eid, and Niloofar Haeri, 9–46. Amsterdam and Philadelphia: J. Benjamins.

Fukui, Naoki. 1995. *Theory of projection in syntax*. Stanford, Cal.: Center for the Study of Language and Information.

Holes, Clive. 2004. *Modern Arabic: Structures, functions, varieties*. 2nd rev. ed. Washington, D.C.: Georgetown University Press.

Jackendoff, Ray. 1977. *X' syntax: A study of phrase structure*. Cambridge, Mass.: MIT Press.

Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, Mass.: MIT Press.

Koizumi, Masatoshi. 1995. *Phrase structure in Minimalist syntax*. Ph.D. diss., Massachusetts Institute of Technology.

LeTourneau, Mark S. 1995. “Internal and External Agreement in quantified construct states”. Eid (1995:29–57).

—— 1998. "Case-checking by morphological causatives in Standard Arabic: A Minimalist approach". *al-ʿArabiyya* 31.89–122.
—— 2003. "Interpretability, feature strength, and impoverished agreement in Standard Arabic". *Perspectives on Arabic linguistics*, XV, ed. Dilworth B. Parkinson and Samira Farwaneh, 85–131. Amsterdam and Philadelphia: J. Benjamins.
Mohammad, Mohammad A. 1990. "The problem of subject-verb agreement in Arabic: Towards a solution". *Perspectives on Arabic linguistics*, I, ed. Mushira Eid, 95–125. Amsterdam and Philadelphia: J. Benjamins.
—— 1991. "The *ḥāl* construction in Palestinian Arabic". Paper presented at the Fifth Annual Arabic Linguistics Symposium, 1–2 March, 1991, University of Michigan, Ann Arbor.
—— 1999. "Checking and licensing inside DP in Palestinian Arabic". *Perspectives on Arabic linguistics*, XII, ed. Elabbas Benmamoun, 27–44. Amsterdam and Philadelphia: J. Benjamins.
Moro, Andrea. 2000. *Dynamic antisymmetry*. Cambridge, Mass.: MIT Press.
Ouhalla, Jamal. 1991. *Functional categories and parametric variation*. New York: Routledge.
—— 2004. "Semitic relatives". *Linguistic Inquiry* 35.288–300.
—— and Ur Shlonsky. 2002. Introduction to *Themes in Arabic and Hebrew syntax,* ed. Jamal Ouhalla and Ur Shlonsky, 1–43. Dordrecht: Kluwer Academic Publishers.
Plunkett, Bernadette. 1993a. *Subjects and specifier positions*. Ph.D. diss., University of Massachusetts, Amherst.
—— 1993b. "The position of subjects in Modern Standard Arabic". *Perspectives on Arabic linguistics*, V, ed. Mushira Eid, 231–260. Amsterdam and Philadelphia: J. Benjamins.
—— 1996. "Binding in Arabic and the lexical clause". Eid (1996:85–97).
Pollock, Jean-Yves. 1989. "Verb movement, universal grammar, and the structure of IP". *Linguistic Inquiry* 20.365–424.
Ritter, Elizabeth. 1988. "A head-movement approach to construct-state noun phrases". *Linguistics* 26.909–929.
—— 1991. "Two functional categories in noun phrases: Evidence from Modern Hebrew". *Perspectives on phrase structure: Heads and licensing*, ed. Susan J. Rothstein, 37–62. New York Academic Press.
Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge, Mass.: MIT Press.
—— 1997. "The fine structure of the left periphery". *Elements of grammar: Handbook in generative syntax*, ed. Liliane Haegeman, 281–337. Dordrecht: Kluwer Academic Publishers.
Shlonsky, Ur. 1997. *Clause structure and word order in Hebrew and Arabic: An essay in comparative Semitic syntax*. Oxford: Oxford University Press.
—— 2000. "Remarks on the complementizer layer of Standard Arabic". *Research in Afroasiatic grammar: Papers from the Third Conference on Afroasiatic Languages,* ed. Jacqueline Lecarme, Jean Lowenstamm, and Ur Shlonsky, 325–343. Amsterdam and Philadelphia: J. Benjamins.
Speas, Margaret. 1990. *Phrase structure in natural language*. Dordrecht: Kluwer Academic Publishers.
Sportiche, Dominique. 1988. "A theory of floating quantifiers and its corollaries for constituent structure". *Linguistic Inquiry* 19.425–449.
Stowell, Tim. 1989. "Subjects, specifiers, and X-bar theory". Baltin and Kroch (1989:232–262).
Travis, Lisa. 1989. "Parameters of phrase structure". Baltin and Kroch (1989:263–279).
Webelhuth, Gert (ed.). 1995. *Government and binding theory and the Minimalist Program*. New York: Blackwell.
Wright, William. 1974. *A grammar of the Arabic language*. 3rd ed. Repr. Beirut: Libraire du Liban.

Mark S. LeTourneau (Weber State University)

Xafḍ → ʾIʿrāb

Y

Yemen

Arabic dialects in Yemen are spoken as a mother tongue in most parts of the country, except in the eastern province of Mahra and on the island of Soqotra, where Modern South Arabian languages are the native languages of the inhabitants. The Yemeni Arabic dialects are characterized by a great diversity and by a number of unique traits not found elsewhere in the Arabic-speaking community. Although not the best known Arabic varieties, some of them have been studied since the end of the 19th century. For the center of the country (the former People's Democratic Republic of Yemen), the first descriptions are due to the Swedish scholar Carlo de Landberg (1901, 1905, 1909, 1913, 1920–1943) for the sultanates of Ḥaḍramawt, Faḏlī, High and Low ʿAwlaqī, ʿAwāḏil, and the tribal Confederacy of Daṯīnah (today, all but Ḥaḍramawt are situated in Abyan Province). The first dictionary (English/Arabic, in Arabic script) was published by Stace (1893) for the dialect of Aden. A few decades later, during the British domination, Emerson and Ghanem (1943a,b) wrote a grammar with exercises about the same variety. For the western part of the country (the former Yemen Arab Republic, previously an independent imamate) during the 1930s, Arabic dialects were studied by the Italian scholar Rossi (1938, 1939) in Ṣanʿāʾ and a few other places, and, later on, by the Yemeni scholars Nāmī (1946), al-ʾAkwaʿ (1968), and Šaraf ad-Dīn (1970). Landberg's publications on Ḥaḍramawt and Daṯīnah Arabic are still unequaled today and have hardly been updated. Since then, only short studies have been carried out, and no study of a similar scale has been achieved in these areas. The opening of Yemen since the 1970s has allowed linguistic fieldwork to be carried out again and made it possible to supplement the previous studies. The dialects of the western part of the country were the first to be investigated on a larger scale by Western researchers such as Diem (1973), Diem and Radenberg (1994), Jastrow (1984), Behnstedt (1985, 1987, 1992, 1993, 1996, 2006), Bettini (1985), Prochazka (1987), and Werbeck (2001). Their studies, in particular the dialect atlas of Behnstedt and his monographs, followed later on by those of Al-Selwi (1987) on the medieval Yemeni lexicon, Watson (1993, 1996) on → Ṣanʿāni Arabic, and the publications of a few others, such as Naïm-Sanbar (1994) and Simeone-Senelle (1996a,b), have enriched and updated our knowledge of this geographical zone. As for the center, it is only in the late 1980s and early 1990s that studies about dialects not previously described were undertaken by a Yemeni researcher Habtoor (1989–1990) for the valley of Ġayl Ḥabbān in Šabwa Province, and by Vanhove for the mountainous areas of Yāfiʿ and Ḏ̣ālaʿ (Vanhove 1993, 1995a,b, 1996a, 1997). Fodor's (1970) short monograph of the Lahej dialect, the only available publication based on the speech of migrants just settled in Egypt, must also be mentioned. Another important source of documentation on Yemeni Arabic are the studies undertaken by Goitein (1960, 1970) and Piamenta (1990–1991) of the speech of Yemeni Jewish migrants in Israel and the manuscripts, written in Hebrew script, they brought with them. For the purpose of this presentation, data from unpublished fieldwork by Vanhove for Aden, Lahej, Abyan, Mukeyras, and Ḥaḍramawt areas are also used.

Today, our better knowledge of the Yemeni Arabic varieties, although still fragmented, shows that the dialectal limits do not strictly correspond to the political boundaries of the 20th century.

From the available data, it is possible to draw a typological classification (see map). It is still provisional because of the lack of data in many places, and some of the geographical limits are still vague, particularly in the central part of the country where no extensive study has been achieved yet.

For the former Yemen Arab Republic in the west, Behnstedt (1985) proposed, after Diem (1973) with some slight modifications, a division into eleven main types (with subdivisions), all but one named after geographical zones.

I. Tihāmah dialects
II. *-k* dialects (these include the 'Southern Mountain Range' type of Diem)
III. Ḥugariyyah dialects
IVa. Jabal Wašḥah
IVb. alMaḥābišah (IVa and IVb are transition zones sharing features with both zones I and Va)
Va. Northern High Plateau
Vb. Arḥab
VI. Ṣanʿāʾ and central mountains (→ Ṣanʿānī Arabic)
VII. Southern High Plateau
VIII. Mārib and alBayḏ̣āʾ dialects (former Southeast dialects)
IX. Northern High Plateau 2 (perhaps a transition zone between Va, Vb, and XI)
X. alJawf dialects (former Northeast dialects)
XI. Ṣaʿdah

Since the reunification of Yemen, some of the geographical designations have become inappropriate. Because those of zones VIII and X could be misleading, they have been relabeled here. In a recent work, Behnstedt groups together the Jawf and Mārib dialects (but apparently not alBayḏ̣āʾ) and classifies them as "Bedouin-type dialects intermediate between those of the highlands, North Arabian dialects and those of the south-east of the Arabian peninsula (Ḥaḍramawt, Ḏ̣ufār, ʿUmān)" (2001:23). So, it seems this classification could be reduced to ten types.

Furthermore, it is now known that a few of the above-mentioned groups extend further east. In particular, the *-k* perfect dialectal zone (zone II) goes on into Ḏ̣ālaʿ and Yāfiʿ, two areas located in Ḏ̣ālaʿ Province, the northern part of Lahej Province, and the west of Abyan. The same type of dialectal enclave as in the west is found in the city of Ḏ̣ālaʿ itself, i.e. dialects with a final *-tu* in the 1st person singular perfect, a feature considered typical of the Ḥugarīyyah (zone III) but also found in the city of Taʿizz, for instance. Diem supposed that zones IIa and III may "have originally been one connected group, from which the Ḥugarīyeh dialects split off in losing – perhaps owing to a greater accessibility of the region – the common *k*-perfect and adopting instead the *t*-perfect" (1973:149).

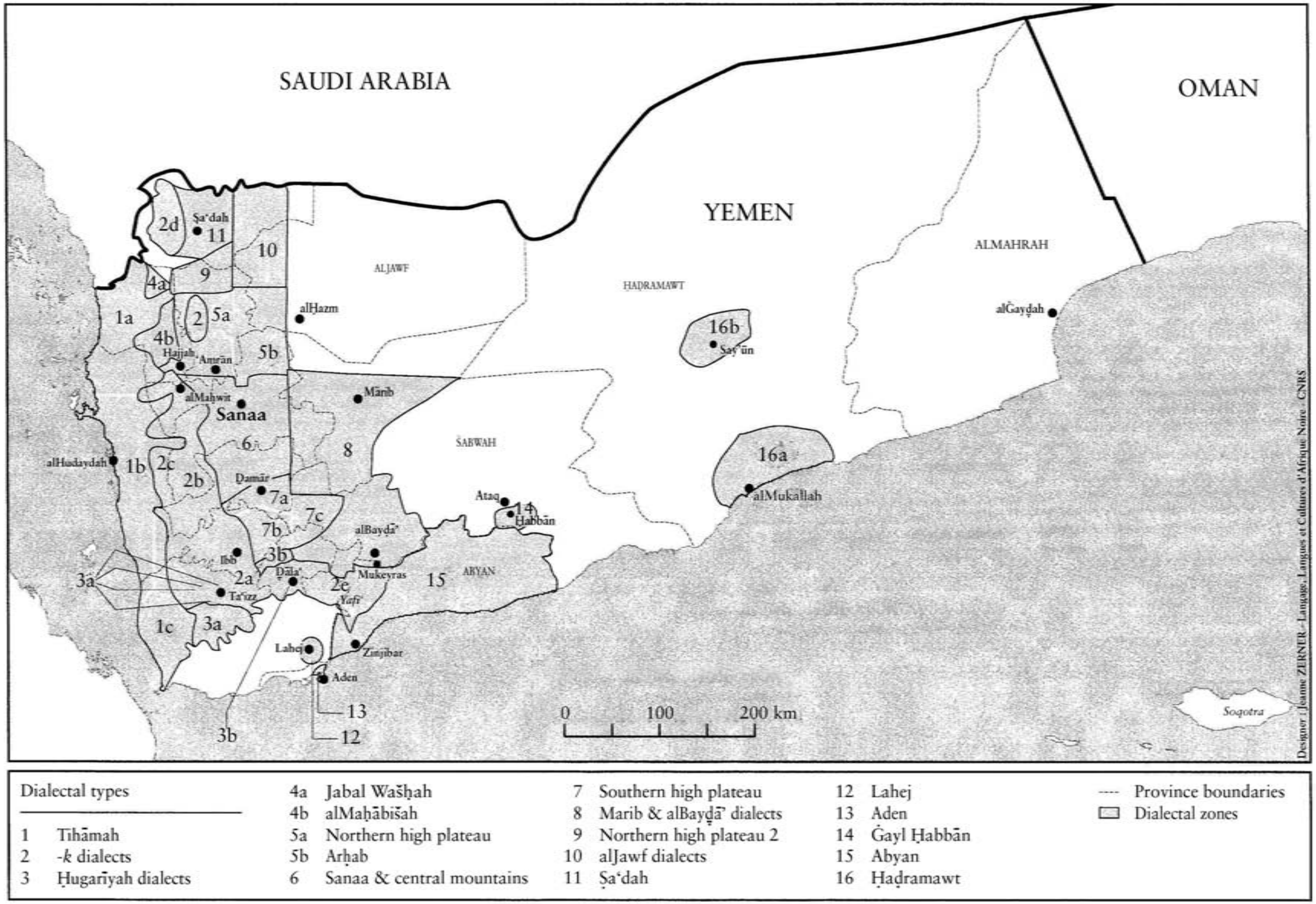
The Tihāmah dialectal zone extends further to the south into Lahej Province, but as there are only data available for alṢabbēha (= Ṯōr alBāha), a village situated to the west of Lahej City, it is not mentioned on the map.

The dialect of Mukeyras, situated on the high plateau to the north of Abyan Province, near alBayḏ̣āʾ City, shows a number of similarities with zone VIII, e.g. the article with *m-* (*am-* in Mukeyras, *im-* in alBayḏ̣āʾ) or the change /*q/ > /g/. But to group them together is still debatable because of the scarce data, the possibility of contact phenomena, and some significant differences, such as the change from pharyngeal /ʿ/ to laryngeal /ʾ/.

In addition to these eleven (or ten) types, it is possible to recognize at least five more main dialectal zones, with subgroupings, exclusively located in the former People's Democratic Republic. These are:

XII. Lahej
XIII. Aden
XIV. Ġayl Ḥabbān
XV. Abyan (which includes the Daṯīnah dialects described by Landberg)
XVI. Ḥaḍramawt

Ḥaḍrami dialects share some features with other varieties spoken in the Arabian Peninsula, in particular the palatalization of /*g/ into /y/ (a feature which cuts across the dialect boundaries;


SAUDI ARABIA
OMAN
YEMEN
ALJAWF
ḤADRAMAWT
ALMAHRAH
ŠABWAH
ABYAN
Ṣa'dah
alḤazm
Ḥajjah
'Amrān
alMaḥwit
Sanaa
Mārib
alḤudaydah
Ḏamār
Ibb
Ta'izz
Ḍāla'
alBayḍā'
Mukeyras
Yafi'
Ataq
Ḥabbān
Laḥej
Zinjibar
Aden
Say'ūn
alMukallah
alĠayḍah
Soqotra
1a
1b
1c
2
2a
2b
2c
2d
2e
3a
3b
4a
4b
5a
5b
6
7a
7b
7c
8
9
10
11
12
13
14
15
16a
16b
0
100
200 km
Dialectal types
1 Tihāmah
2 -k dialects
3 Ḥugarīyah dialects
4a Jabal Wašḥah
4b alMaḥābišah
5a Northern high plateau
5b Arḥab
6 Sanaa & central mountains
7 Southern high plateau
8 Marib & alBayḍā' dialects
9 Northern high plateau 2
10 alJawf dialects
11 Ṣa'dah
12 Lahej
13 Aden
14 Ġayl Ḥabbān
15 Abyan
16 Ḥaḍramawt
---- Province boundaries
▨ Dialectal zones

Designer : Jeanne ZERNER - Langage, Langues et Cultures d'Afrique Noire - CNRS

Map 1: The Arabic dialects of Yemen

see Johnstone 1965), but it seems there are enough differences to set them apart as a dialectal entity (→ Wādi Ḥaḍramawt Arabic).

It must be pointed out that the extreme diversity of Yemeni dialects prevents, in many instances, mutual understanding. For example, it is not possible for other Yemeni speakers, even the closest neighbors, to understand uneducated Tihāmi, Ṣaʿdi, or Yāfiʿi speakers, nor are the three dialects mutually understandable. If dialectal continuity and transition zones do exist in some parts of Yemen, this country could be best characterized as a zone of discontinuities.

The above dialectal classification is founded on a number of phonetic, phonological, morphological, and, to a lesser extent, lexical isoglosses, some of which have already been mentioned. A few others can be added. Diem (1973) sets apart Ṣanʿāʾ and the central mountains from the Northern High Plateau zone, in spite of many common features, because of, among other things, the perfect flexion in the 2nd and 3rd persons plural (*sirtayn*/*sārayn* vs. *sirtin*(*na*)/*sirna* 'you [pl.]/they went'), and from the Southern High Plateau zone because of the relative pronoun, which is *ḏi* (instead of *allaḏi* in Ṣanʿāʾ).

The coastal plain of the Tihāmah is characterized by the remarkable ending *-an* in the 3rd person singular feminine, the definite article *am-*, the indefinite article *-un*, the negative adverb *dawʾ* ~ *daʾ* ~ *duwwayy* 'no' (in the south) close to the Sabaic and → Ḥimyaritic forms (a little known language which Rabin considers to be "basically an Arabic dialect of the Yemenite type, but with some archaic features, and with a great deal of South-Arabian loanwords" [1951:2]), etc.

As for central Yemen, Tables 1 and 2 show the main distinctive criteria taken into consideration.

Comments on Table 1

1. The various degrees of affrication and palatalization for /*g/, so common in western Yemen, are rarer in the center, with the noticeable exceptions of Ḥaḍramawt, Ġayl Ḥabbān, and Mukeyras. It seems that the southwest of the country, from the Tihāmah coast up to Abyan, forms a sort of enclave in the Yemen where the reconstructed Semitic velar stop /*g/ was preserved.

2. Unlike in west Yemen (Behnstedt 1985:41), [q], [ɣ], and [ʁ] are not in complementary distribution in the center. All three allophones (only the first two ones for /ġ/) are free variants, often for both /*q/ and /*ġ/.

3. /ġ/ has an additional allophone which it does not share with /q/, i.e. the laryngeal stop [ʔ], or, in Yāfiʿ and Ḏālaʿ, a velarized [ʔ̴], whose velar trait spreads to all the consonants of a word (e.g. [ʔ̴an̴am̴] for *ġanam* 'goat', [m̴aʔ̴r̴eb̴] for *maġreb* 'west' (this 'emphatic' characteristic of the laryngeal stop is not described by Landberg for Abyan or Fodor for Lahej). It is possible, although not certain, that the voiced velar fricative /ġ/ has undergone a change to

Table 1. Main consonantal reflexes
(The data between brackets belong either to a restricted number of subgroups or to lost variants; slashes indicate phonetic variants.)

Proto-Arabic	Yāfiʿ (II)	Ḏālaʿ (II)	Ḏālaʿ City (III)	Mukeyras (?)	Lahej (XII)	Aden (XIII)	Ḥabbān (XIV)	Abyan (XV)	Ḥaḍramawt (XVI)
*g	g	g	g	dy	g	g	j	g	y (j, ž)
*q	q/ɣ/ʁ	q/ɣ	q	g	q	q	g	q/ɣ (ʕ)	g
*ġ	ɣ/q/ʔ̴ (ʕ)	ɣ/ʔ̴	ġ	ġ	ɣ/q/ʔ	ġ	ġ	ɣ/q/ʔ (ʕ)	ġ
*ʕ	ʕ (ʔ)	ʕ	ʕ	ʔ (ʕ)	ʕ	ʕ	ʕ	ʕ/ʔ	ʕ
*lʒ̣	ḏ̣ (ḍ)	ḏ̣	ḏ̣	ḏ̣	ḏ̣	ḍ	ḷ	ḷ	ḏ̣
*ḏ̣	ḏ̣	ḏ̣	ḏ̣	ḏ̣	ḍ	ḍ	ḷ	ḷ (ḏ̣)	ḏ̣
interdentals	+	+	+	+	+	–	+	+ (f)	+ (f)
*k > š	+	–	+	+	+	–	+	+	+
emphasis	+	+	+	+	+	–	+	+/–	+

a laryngeal stop, retaining and spreading the velar, i.e. emphatic, trait. This may have been thwarted by sociolinguistic factors as well as phonetic ones (e.g. the fricative allophone [ɣ̴] of /q/). Today, [ʔ̵] is stigmatized. Only the elderly use it frequently, while it is retained in only a few words by the younger generations.

4. The [ʕ] variant for /q/ in Abyan, mentioned by a Yemeni colleague (p.c.) for a coastal village, needs to be checked. Whatever the real status of this variant, phonetically it would not be surprising: Landberg (1901:271) reports, for Ḥaḍrami and Classical Arabic, an alternation between the pharyngeal and the uvular in the verb meaning ‘to be able’, *ṭāʿ* and *ṭāq* respectively, and an alternation [q] ~ [ʕ] is attested in classical Aramaic script (possibly through an emphatic velar fricative allophone [ɣ̴] of /q/ (Cohen 1988:85), for the reflex of the Semitic emphatic fricative lateral phoneme */lʒ̣/, corresponding to Arabic *ḍād* (not to **q*).

5. As for the coalescence of /ġ/ with /ʿ/ in Abyan, although it was often noted by Landberg a century ago, it seems now to be quickly disappearing, at least in the ten places surveyed by Vanhove in 1989 and 1991. A trace of the phenomenon is still found in cases of hypercorrection, such as [ɣazaf] for [ʕazaf] ‘wickerwork’. But in Jabal Yazīdī, a village in the Yāfiʿ area, the merging of /ġ/ and /ʿ/ into /ʿ/ is still alive, although receding. Another change is still heard in the speech of Bedouin women of the hamlet of Ṯuwa in Abyan: the merger of /ġ/ and /ʿ/ into the laryngeal stop /ʾ/. Landberg already pointed to the decline of the pronunciation of /ʿ/ as a laryngeal stop /ʾ/. It is even more true today.

6. Regarding *ḍād* and *ḏ̣āʾ*, the dialects of Abyan and Ġayl Ḥabbān are the only ones to have kept the lateral trait of the Semitic emphatic fricative lateral phoneme */lʒ̣/ (Arabic → *ḍād*), which has become a velarized /ḷ/ (*ʿaḷam* ‘bones’). In some cities and villages of the Abyan zone, such as Mudia and Lawdar, it changed into a velarized /ṛ/ (*ʿaṛam* ‘bones’). Furthermore, Yemen is also the sole country where remnants of a distinction between the two Semitic phonemes */lʒ̣/ and */ṯ̣/ are found. Jastrow (1980:106) interpreted Landberg’s indication about Daṯīnah as a distinction between /ḷ/ and /ḏ̣/, but it seems to have disappeared today. [ð̴̣] is only a combinatory variant of /ḷ/, before the resonants /l/, /r/, and /n/: [ð̴̣ofər] ‘nail’. For Lahej dialect, Fodor’s description shows a (nonsystematic) distinction between an emphatic voiced interdental /ḏ̣/ and a dental /ḍ/ (*baʿḏ̣* ‘some’, *ḍalf* ‘claw’), and Behnstedt (1987:5) also mentions it for some subgroups in the Ṣaʿdah zone, with different realizations, respectively /ḏ̣/ and /ṯ̣/. It is not possible to find in the literature a single minimal pair to prove their phonological status, but /ḏ̣/ is a rare phoneme in Arabic anyway. Yemeni Arabic is also exceptional among Arabic varieties, for the affricated retroflex articulation [tʃ] of *ḍād*, as mentioned by Behnstedt (1987:5, 136) for some villages in the Ṣaʿdah zone. Until then, retroflexion was thought to be only a Cushitic reflex of the old emphatic consonant.

7. The merger of the interdentals with the corresponding dentals is rare in Yemen, and is only mentioned as a regular change for alḤudaydah and Aden. Still, some dialects have preserved the interdentals, fusing /ḏ/ and /d/ into /d/, but only in a few words. Such is the case for Ġayl Ḥabbān and Ḏ̣ālaʿ, where the relative pronoun is *di* and not *ḏi*. Other subdialects, in Abyan and Ḥaḍramawt, have a fricative free variant [f] for the unvoiced interdental /ṯ/, which seems to be limited today to a few words, such as [dafiːnah] for [daθiːnah] or [falaːf] for [θalaːθ] ‘three’.

8. The change /k/ > /š/ concerns only the 2nd person singular feminine pronominal suffix, and, for part of the -*k* dialects, also the flexion of the perfect 2nd person singular feminine (→ *kaškaša/kaskasa*). It is noteworthy that in central Yemen, only Aden and Ḏ̣ālaʿ -*k* dialects have not undergone this quite common form of → palatalization, absent only in the western part of the country, in the Tihāmah, Ḥugariyyah, and the southern -*k* dialects (IIa) – and in four villages of the Va zone (Behnstedt 1985:82). Therefore, Ḏ̣ālaʿ -*k* dialects could be grouped with the subtype of southern -*k* dialects, while the area of Yāfiʿ most probably constitutes an eastern subgroup (IIe). Aden thus represents the southernmost point of this morphophonetic retention.

9. Although not mentioned by previous authors, the loss of the emphatic trait is also a characteristic of the old Adeni dialect. But it is rapidly receding and is only heard today in the speech of old residents of Crater, the

city center, mainly old women. In Yāfiʿ, as in Abyan, it is also a regressive feature, as already noticed by Landberg for Daṯīnah, but emphasis can still be very light or even nonexistent, in the speech of old women.

Comments on Table 2

1. The distinction between a masculine and a feminine form in the 1st person singular pronoun, *ana* vs. *ani*, quite common in the west (Behnstedt 1985:71), is also found in central Yemen, in a zone which continues the western one. It is only marginal in Ḥaḍramawt.

2. The 1st person plural pronoun in central Yemen presents a unique feature, i.e. the change of the initial nasal /n/ to the resonants /l/ or /r/: *laḥna*, in part of Abyan, as already noted by Landberg, or *raḥna* in Yāfiʿ and part

Table 2. Morphological features
(The data between brackets belong to a restricted number of subgroups).

	Yāfiʿ (II)	Ḏ̣ālaʿ (II)	Ḏ̣ālaʿ City (III)	Mukeyras (?)	Lahej (XII)	Aden (XIII)	Ḥabbān (XIV)	Abyan (XV)	Ḥaḍramawt (XVI)
1st sg. masc., fem. pronoun	*ana*, *ani*	*ʾana*, *ʾani*	*ana*, *ani*	*ana*	*ʾana*, *ʾani*	*ana*	*ʾana*	*ana*	*ana*, *ani*
1st pl. pronoun	*raḥna*	(*raḥna*)	*naḥna*	*naḥna*	*naḥna*	*naḥna*	*naḥna*	*laḥna* (*ḥna*)	*naḥna*
1st pl. imperf.	*n-*	*n-*	*n-*	?	*n-*	*n-*	*l-*	*l- (n-)*	*n-*
III*w/y* imperf.		*yēbək*							
2nd pl. fem. imperf. 3rd pl.masc. imperf.	*-īn -ūn*	*-ĩ: ũ*	*-ī -ū*	? ?	? *-ū*	*-i -u*	*-īn -ūn*	*-ī ~ -īn* *-ū ~* *-ūn*	*-in -ūn*
1st, 2nd sg. and pl. perf.	*-k*	*-k*	*-tu*	*-t*	*-t*	*-t*	*-t*	*-t*	*-t*
3rd sg. fem. perf.	*-ah*	*-ah*	*-ah*	-v*t*	*-at*	*-at*	-v*t*	-v*t*	-v*t*
II gem. 3rd sg. masc. perf.		*-iy*	*-a*						
gender difference in plural of verbs	+	+	+	+	−	−	−	+	+
progressive particle	(*bi-*)	*ya- ~* *ta- ~* *qa-*	*la-*	?	Ø	Ø	Ø	Ø	Ø
article	*al-*	*al-*	*al-*	*am-*	*al-*	*al-*	*al-*	*am- ~* *um-*	*al-*

of Ḏālaᶜ; *raḥna* is not found in the other *-k* dialects of western Yemen.

3. The change of resonants has also affected the /n/ of the 1st person plural prefix of the imperfect, which is changed to /l/ in part of Abyan dialects (*lənṣa* 'we forget') and in Ġayl Ḥabbān (*lilᶜab* 'we play'), a zone which has not undergone the change in the independent pronoun.

4. Yemen has both the vocalic and the nasal endings in the 2nd and 3rd person plural imperfect. In the west, the nasal ending is almost entirely restricted to north and center Tihāmah. In the center, it is found in Yāfiᶜ, Ġayl Ḥabbān, part of Abyan, and Ḥaḍramawt. It is not certain that the geographical variants of Ḏālaᶜ, [ĩː] and [ũː], are reflexes of a nasal consonant; it could be a pausal nasalization, quite common in the southern *-k* dialects (with the addition of [ŋ]; Behnstedt 1985:57).

5. The dialect of Gihaf, in Ḏālaᶜ, has another remarkable peculiarity in the form of the imperfect of verbs III*y*. It is restructured on the pattern of verbs I*w* (with a different long vowel): *yēbək* 'he weeps'. Unique in Arabic, this analogical formation is known in Semitic, namely in Hebrew (Gesenius 1910:211).

6. The *-k* perfect is typical of the dialects of the high mountain range of Yemen. They share this feature with → Ḥimyaritic, which was spoken in the same mountainous area, and with the South Semitic languages (South Arabian stocks, Modern South Arabian languages, Afro-Semitic). Substratal influence is thus possible.

7. In central Yemen, the extension of the *-ah* ending in the 3rd person singular feminine perfect, analogous to the nominal feminine ending, directly continues the zone of southern *-k* dialects and Ḥugariyyah dialects of western Yemen.

8. The vocalic ending in the 3rd person singular masculine perfect for geminated verbs, typical of Chadic and Sudanese Arabic, is also found sporadically in Yemen, mainly in Northern Tihāmah. Behnstedt (1985:136) notes either /-a/, /-ī/ or /-ē/. In central Yemen, both Ḏālaᶜ City and Ḏālaᶜ share this feature, respectively with /-a/ and /-iy/: *ḥabba* 'he loved', *šelliy* 'he took'.

9. The nondistinction of gender in the plural verbal forms is rather rare in Yemen. In the west, it is mainly restricted to the Tihāmah; in the center, three dialectal zones are concerned: Lahej, Aden, and Ġayl Ḥabbān.

10. Central Yemen seems to use few particles – five only – to express the progressive, as against some ten in the west, and they are limited to Ḏālaᶜ City, Ḏālaᶜ, and Yāfiᶜ. But further research is needed. Two of the particles, *la-* and *bi-*, are also used in the west, but *bi-* is not frequent in most Yāfiᶜi dialects, while it has become an obligatory marker of the imperfect (no longer of the progressive) in the subvariety of Jabal Yazīdī. Particles *ya-*, *ta-*, and *qa-*, particular to central Yemen, are geographical variants in the Ḏālaᶜ zone.

11. The definite article with *m-* is used in a few areas in both western and central Yemen, and nowhere does it assimilate to any initial consonant of the following noun: *am-ṣanīni* 'the little finger' (Abyan). It is now receding. Landberg (1909:286) for Upper ᶜAwlaqī and Marxah in Abyan, and Behnstedt (1987) for the area of Ṣaᶜdah, also mention an article *en-* or *an-*, which corresponds to the old Ḥimyaritic article.

12. Another morphological feature, not mentioned in Table 2, is worth referring to, i.e. the existence of an internal passive for verbs of Form I. It is more or less productive, depending on the dialect, but it is still quite common all over Yemen.

As for the lexicon, Yemeni dialects present many peculiarities, which can be found in the dictionaries mentioned in the bibliography. Some of them are traces of → Ḥimyaritic, or of Old South Arabian languages, as already mentioned for *dawʾ* 'no'. For others, the form and/or the meaning are unknown in other Arabic varieties, e.g. *tqambas* 'to sit down' (Ṣaᶜdah, Yāfiᶜ), *ḥusan* 'cat' (Yāfiᶜ). There are also interesting cases of grammaticalization toward prepositions and adverbs: (Yāfiᶜ) *dōr* 'next to' (< *d-w-r*), *rās* 'on' (< *r-ʾ-s*), *tī* 'like, as' (deictic), *ʾublah* 'together' (cf. Classical Arabic *ʾabala* 'to be numerous [camels]').

Syntactic studies, apart from Watson (1993) for the Ṣanᶜāʾ dialect, are still rare and fragmentary. Studies about negation, conditional and concessive sentences, discourse particles, and focalization in the Yāfiᶜi dialects can be found in Vanhove (1996a,b, 2000, 2002a) and for negation and verbs of movement in the Tihāmah dialects in Simeone-Senelle (1996a,b).

Landberg's studies on Ḥaḍramawt and Daṯīnah also contain a number of interesting, although unorganized, observations in various syntactic domains.

As elsewhere in the Arab world, Yemeni dialects are subjected to a process of → leveling, for various social reasons. So far, no systematic study of the linguistic outcome of social changes has been undertaken, so, in addition to the previous scattered remarks, the following general comment is mainly based on Vanhove's fieldwork in the central part of the country (Vanhove 1997, 2002b). Among the phonetic and morphological characteristics discussed above, only three features seem to resist even when speakers migrate to different places: the alternation [q] ~ [ɣ] for both /q/ and /ġ/ in Yāfiʿ and Abyan, the 1st person singular feminine pronoun *ani*, and the presence or absence of gender opposition in the plural verb. Even the widespread development /k/ > /š/ (→ *kaškaša*/*kaskasa*) or the fusion of the interdentals with the corresponding dentals are slowly receding among educated people. On the whole, men are more likely to change their speech toward more koineized forms than women, but young educated women are also very sensitive, as is usually the case, to more modern or fashionable forms. In Yemen, the problem is to know how to characterize the koine (or koines). It seems that the capital city Ṣanʿāʾ is not a center of attraction, because of its many peculiarities, which set it apart from any version of Modern Standard Arabic. Furthermore, the influence of Egyptian Arabic is more limited than it was once thought. In the center of the country, province capitals, such as Aden or Mukalla, tend to be prestigious models, but their sociolinguistic situation is quite complex and little known. In spite of the leveling process, Yemen is still a very conservative and fascinating dialectal area, where much remains to be done and discovered.

BIBLIOGRAPHICAL REFERENCES

ʾAkwaʿ, I. al-. 1968. *al-ʾAmṯāl al-yamaniyya maʿa muqāranatihā bi-naḏāʾirihā min al-ʾamṯāl al-fuṣḥā wa l-ʾamṯāl al-ʿāmmiyya fī l-bilād al-ʿarabiyya*, I. Cairo.

Al-Selwi, Ibrahim. 1987. *Jemenitische Wörter in den Werken von al-Ḥamdānī und Našwān und ihre Parallelen in den semitischen Sprachen*. Berlin: D. Reimer.

Behnstedt, Peter. 1985. *Die nordjemenitischen Dialekte*. I. *Atlas*. Wiesbaden: L. Reichert.

——. 1987. *Die Dialekte der Gegend von Ṣaʿdah (Nord-Jemen)*. Wiesbaden: O. Harrassowitz.

——. 1992. *Die nordjemenitischen Dialekte*. II. *Glossar Alif – Dāl*. Wiesbaden: L. Reichert.

——. 1993. *Glossar der jemenitischen Dialektwörter in Eduard Glasers Tagebüchern (II, III, VI, VII, VIII, X)*. Vienna: Österreichische Akademie der Wissenschaften.

——. 1996. *Die nordjemenitischen Dialekte*. II. *Glossar Ḏāl – Ġayn*. Wiesbaden: L. Reichert.

——. 2001. "Notes on the Arabic dialects of eastern North-Yemen (ilǦawf, Ṣirwāḥ, Mārib, Banī Ḏabyān)". *Linguistic and cultural studies on Arabic and Hebrew: Essays presented to Professor Moshe Piamenta for his eightieth birthday*, ed. Judith Rosenhouse and Ami Elad-Bouskila, 23–40. Wiesbaden: O. Harrassowitz.

——. 2006. *Die nordjemenitischen Dialekte*. III. *Glossar Fāʾ – Yāʾ*. Wiesbaden: L. Reichert.

Bettini, Lidia. 1985. "Note sull'arabo parlato al Barradūn (Yemen del Nord)". *Studi Yemeniti, Quaderni di Semitistica* 14, ed. Pelio Fronzaroli, 117–159. Florence: Istituto di Linguistica e di Lingue Orientali, Università di Firenze.

Cohen, David. 1988. "L'araméen". *Les langues dans le monde ancien et moderne: Langues chamito-sémitiques*, ed. Jean Perrot and David Cohen, 84–104. Paris: CNRS.

Diem, Werner. 1973. *Skizzen jemenitischer Dialekte*. Beirut: F. Steiner.

——. and Hans-Peter Radenberg. 1994. *Dictionary: The Arabic material of S.D. Goitein's* A Mediterranean society. Wiesbaden: O. Harrassowitz.

Emerson, L.H.S. and S.M.A. Ghanem. 1943a. *Aden Arabic exercises*. Aden: Al-Maaref Press.

——, and S.M.A. Ghanem. 1943b. *Aden Arabic grammar*. Aden: British Council, Al-Maaref Press.

Fodor, I. 1970. "Two tales in the dialect of Lahej". *Mitteilungen des Instituts für Orientforschung* 16:1.33–74.

Gesenius, Wilhelm. 1910. *Hebrew Grammar*. Ed. and enlarged by Emil Kautzsch. 2nd English ed. trans. and rev. in accordance with the 28th German ed. (1909) by A.E. Crowley. Oxford: Clarendon Press.

Goitein, Shlomo Dov. 1960. "The language of al-Gades: The main characteristics of an Arabic dialect spoken in Lower Yemen". *Le Muséon*. 351–394.

——. 1970. *Jemenica: Sprichwörter und Redensarten aus Zentral-Jemen mit zahlreichen Sach- und Worterläuterungen*. Leiden: E.J. Brill. (1st publ., 1934.)

Haḅtoor, Mohamed. 1989–1990. *L'arabe parlé à Ġayl Ḥabbān: Phonologie et morphologie*. Thèse de doctorat, Université Paris III.

Jastrow, Otto. 1980. "Die Dialekte der Arabischen Halbinsel". *Handbuch der arabischen Dialekte*, ed. Wolfdietrich Fischer and Otto Jastrow, 103–129. Wiesbaden: O. Harrassowitz.

——. 1984. "Zur Phonologie und Phonetik des Ṣanʿānischen". *Entwicklungsprozesse in der Arabischen Republik Jemen* (= *Jemen Studien*, I.), ed. Horst Kopp and Günther Schweizer, 289–304. Wiesbaden: L. Reichert.

Johnstone, Thomas M. 1965. "The sound change *j* > *y* in the Arabic dialects of peninsular Arabia".

Bulletin of the School of Oriental and Asian Studies 28.233–241.

Landberg, Carlo de. 1901. *Etudes sur les dialectes de l'Arabie méridionale.* I. *Ḥaḍramaût.* Leiden: E.J. Brill.

——. 1905. *Etudes sur les dialectes de l'Arabie méridionale.* II. *Daṯînah.* 1. *Textes et traduction.* Leiden: E.J. Brill.

——. 1909. *Etude sur les dialectes de l'Arabie méridionale.* II. *Daṯînah.* 2. Leiden: E.J. Brill.

——. 1913. *Etudes sur les dialectes de l'Arabie méridionale.* II. *Daṯînah.* 3. *Commentaire des textes poétiques, articles détachés et indices.* Leiden: E.J. Brill.

——. 1920–1943. *Glossaire daṯînois.* Leiden: E.J. Brill.

Naïm-Sanbar, Samia. 1994. "Contribution à l'étude de l'accent yéménite: Le parler des femmes de l'ancienne génération". *Zeitschrift für Arabische Linguistik* 27.65–89.

Nāmī, Yaḥyā Xalīl. 1946. "Min al-lahajāṯ al-yamaniyya al-ḥadīṯa". *Majalla Kulliyat al-ʾĀdāb* 8.69–84.

Piamenta, Moshe. 1990–1991. *Dictionary of post-Classical Yemeni Arabic.* 2 vols. Leiden: E.J. Brill.

Prochazka, Theodore Jr. 1987. "Remarks on the spoken Arabic of Zabīd". *Zeitschrift für Arabische Linguistik* 17.58–68.

Rabin, Chaim. 1951. *Ancient West-Arabian.* London: Taylor's Foreign Press.

Rossi, Ettore. 1938. "Appunti di dialettologia del Yemen". *Rivista degli Studi Orientali.* 17.230–265.

——. 1939. *L'arabo parlato a Ṣanʿâʾ: Grammatica, testi, lessico.* Rome: Istituto per l'Oriente.

Šaraf ad-Dīn, ʾAḥmad Ḥusayn. 1970. *Lahajāt al-Yaman qadīman wa-ḥadīṯan.* Cairo.

Simeone-Senelle, Marie-Claude. 1996a. "Le fonctionnement de certains verbes de mouvement dans des dialectes de la Tihama du Yémen". *Proceedings of the 2nd International Conference of l'Association Internationale pour la Dialectologie Arabe held at Trinity Hall in the University of Cambridge, 10–14 September*, ed. Joe Cremona, Clive Holes, and Geoffrey Khan, 227–235. Cambridge: University of Cambridge.

——. 1996b. "The negation in some Arabic dialects of the Tihama of the Yemen". *Perspectives on Arabic linguistics*, IX, ed. Mushira Eid and Dilworth Parkinson, 207–221. Amsterdam and Philadelphia: J. Benjamins.

Stace. E.V. 1893. *An English-Arabic vocabulary for the use of students of the Colloquial.* London: B. Quaritch.

Vanhove, Martine. 1993. "Note sur le dialecte *qəltu* de Dhalaʿ (province de Lahej, Yémen)". *Matériaux Arabes et Sudarabiques* N.S. 5.165–190.

——. 1995a. "Notes on the Arabic dialectal area of Yāfiʿ (Yemen)". *Proceedings of the Seminar for Arabian Studies* 25.141–152.

——. 1995b. "À propos du verbe dans les dialectes arabes de Yāfiʿ (Yémen)". *Dialectologia arabica: A collection of articles in honour of the sixtieth birthday of Professor Heikki Palva* (= *Studia Orientalia* 75), 257–269. Helsinki: Finnish Oriental Society.

——. 1996a. "The negation *maašii* in a Yaafiʿi dialect (Yemen)". *Perspectives on Arabic linguistics*, IX, ed. Mushira Eid and Dilworth Parkinson, 195–206. Amsterdam and Philadelphia: J. Benjamins.

——. 1996b. "Les particules *qad* et *raʿ* dans un dialecte arabe de Yāfiʿ (Yémen)". *Proceedings of the 2nd International Conference of l'Association Internationale pour la Dialectologie Arabe held at Trinity Hall in the University of Cambridge, 10–14 September*, ed. Joe Cremona, Clive Holes, and Geoffrey Khan, 243–252. Cambridge: University of Cambridge.

——. 1997. "Diversité dialectale et nivellement linguistique: Le cas de l'arabe au Yémen (sud et est)". *Proceedings of the 16th International Congress of Linguists, Paris 20–25 July 1997*, ed. Bernard Caron, paper no. 115. Oxford: Pergamon.

——. 2000. "Coordination, subordination, deixis et focalisation en arabe de Yāfiʿ (Yémen)". *Oriente Moderno* 19 (LXXX), N.S. 1.209–223.

——. 2002a. "Conditionnelles et concessives en arabe de Yāfiʿ (Yémen)". *"Sprich doch mit deinen Knechten aramäisch, wir verstehen es!": 60 Beiträge zur Semitistik. Festschrift für Otto Jastrow zum 60. Geburtstag*, ed. Werner Arnold and Hartmut Bobzin, 755–775. Wiesbaden: O. Harrassowitz.

——. 2002b. "Processus de koïnisation dans le dialecte arabe de Yāfiʿ (Yémen)". *Aspects of the dialects of Arabic today: Proceedings of the 4th Conference of the International Arabic Dialectology Association (AÏDA) Marrakesh Apr. 1–4. 2000,* ed. Abderrahman Youssi, Fouzia Benjelloun, Mohamed Dahbi, and Zakia Iraqi-Sinaceur, 151–159. Rabat: Amapatril.

Watson, Janet. 1993. *A syntax of Ṣanʿānī Arabic.* Wiesbaden: O. Harrassowitz.

——. 1996. *Ṣbaḥtū! A course in Ṣanʿānī Arabic.* Wiesbaden: O. Harrassowitz.

Werbeck, Wolfgang. 2001. *Laut- und Formenlehre des nordjemenitisch-arabischen Dialekts von Manāḫa.* Münster: Rhema-Verlag.

MARTINE VANHOVE
(LLACAN – INALCO, CNRS)

Yoruba

1. YORUBA AND ISLAM

The term 'Yoruba' identifies a people, as well as the language they speak. The Yoruba, with an estimated population of about 24,000,000 (Fagborun 1994:13), live mainly in → Nigeria's southwestern states and in parts of the West African republics of Benin and Togo. They are made up of approximately 24 dialect groups, all of which are bound to one another by common traditions of origin (Johnson 1921:3–25).

The use of the term 'Yoruba' to refer to both the language and the people gained currency in the early 1840s when Bishop Ajayi Crowther, the first African bishop of the Church Missionary Society, started reducing the Yoruba language to writing mainly to facilitate the translation of the Bible and other Christian literature into Yoruba (Ajayi 1970). The Standard or Koine Yoruba that emerged from this exercise was initially based on the dialects and accents of the Oyo and Egba subgroups (Fagborun 1994:16, 17). Over time, this form of Yoruba, which functioned as Literary Yoruba, the language of the elite, acquired elements not only from many of the other dialects but also from the English language, from which Christian literature was initially translated into Yoruba.

The earliest use of the term ‘Yoruba’ can be traced to Arab writers such as ʾAḥmad Bābā (d. 1627) in his *Miʿrāj as-suʿūd* and Muhammad Bello (d. 1837) in his *ʾInfāq al-maysūr* (Hunwick 1964:568–593; Hunwick and Harrak 2000:392ff.; Bello 1964:48).

One of several widely believed Yoruba traditions locates their origin as Iraq (Bello 1964), while yet another says the Yoruba have been so influenced by Ancient Egyptian language and culture that many of their present words have cognates in Hebrew or the Ancient Egyptian language (Oduyoye 1971:8, 17–28, 37). The proponents of a pure Arab stock for the Yoruba people hold that Arabic is the parent language to Yoruba, and that most contemporary Yoruba vocabulary is traceable to an Arabic origin (Al-Ilori 1972:28–31; Mazhar 1976:1–3; Salloum 1976:7–25).

There is evidence from Arabic sources (Bello 1964:48; Hunwick and Harrak 2000:39–70) and from the accounts of early European travelers and explorers in West Africa (Gbadamosi 1978: Chap. 3), as well as from the morphological features of Arabic loanwords in Yoruba to suggest that Yoruba contact with Islam dates farther back than the 19th century (Reichmuth 1988:272–291). This contact inevitably implies regular contact with Arabic on account of the daily reading and recitation of portions of the *Qurʾān* in daily worship.

2. Arabic loanwords in Yoruba

The daily use of the Arabic language and the accompanying practice of establishing Arabic schools to teach new converts facilitated the inroad of a considerable number of Arabic-derived words and expressions relating to Islamic worship and other subjects into the Yoruba language, with or without morphophonemic modifications to harmonize them with the Yoruba sound system. Some scholars have attempted to trace virtually all Yoruba words to an Arabic root (Mazhar 1976; Salloum 1976; Al-Ilori 1972:34–38), whereas others base their identification of Arabic loanwords on a linguistic analysis (Abdul 1976:37–42; Ogunbiyi 1984:147–177; Reichmuth 1988:272–291; Malik 1995:425–439).

The following data on Arabic loanwords in Yoruba were obtained from previous publications on Arabic loanwords in Yoruba, such as Oduyoye (1971), Al- Ilori (1972), Mazhar (1976), Abdul (1976), Gbadamosi (1978), Ogunbiyi (1984), Malik (1995), as well as from dictionaries of the Yoruba language, especially Crowther (1852) and Abraham (1962).

As a consequence of the absence of certain Arabic sounds from the Yoruba system, they are generally substituted with phonologically proximate sounds in Yoruba, such as /s/ replacing /ṯ/, /z/, /ḏ/, /s/, and /z/ (Ogunbiyi 1980:109–127). In Yoruba orthography, *s* and the vowels *e, o* are written with a dot underneath to represent the Yoruba consonants /š/ and the vowels /ɛ/ and /ɔ/, respectively. Yoruba does not usually allow consonant clusters or a consonant that is not followed by a vowel. Consonant clusters in Arabic loanwords are therefore usually broken up by the insertion of a vowel.

The most easily identifiable loanwords are those in which the basic triliteral consonant phonemes of the Arabic root are retained either in their original Arabic form or in their harmonized form to conform to the Yoruba sound system, e.g. *àlùbáríkà* from the root *b-r-k*, and *bàlágà* from the root *b-l-ġ*.

In the majority of the loanwords, only two of the root consonants remain, either because one of the three original Arabic consonants is replaced by a vowel, as in *àkíìkà* (root *ḥ-q-q*) or because geminate consonants are simplified, as in *àláàjì* (root *ḥ-j-j*), *àlàmísì* (root *x-m-s*). In a number of cases, the last consonant in the root is the one that is dropped, as in *àlùbọ́sà* (root *b-s-l*).

A number of loans have only one consonant or trace of it remaining from the original root, e.g. *èrè, orí*. Some scholars believe that these words represent the earliest stages of the assimilation of Arabic-related loanwords into Yoruba – probably going back to the pre-Islamic migrations of the presumed ancestors of the Yoruba from the Middle East (Al-Ilori 1972:39). At the present stage of scholarship on the subject of Arabic loanwords in Yoruba, it would appear that only those words that have retained three or two consonant phonemes from the Arabic could be categorically said to be of Arabic origin.

3. SEMANTIC FIELDS COVERED BY THE LOANWORDS

Arabic loanwords feature in the speech and writing by Yoruba of all religious and educational backgrounds, in live radio and television broadcasts, in mass-circulation Yoruba newspapers, and even in the Yoruba Bible in current usage (Malik 1995:425–435; Ogunbiyi 2001). In a quasi-technical publication by the Nigerian Educational Research and Development Council entitled *Quadrilingual glossary of legislative terms* (NERDC:1991), many Arabic loanwords, especially for tools, implements, and literacy, such as *tánganran, àlìkámọ̀, ìmọ̀*, and *kálàmù*, occur quite frequently. Even general words such as *àlááfíà, àníyàn*, and *làákàyè* occur frequently.

On account of the relationship between Islam and Arabic, there is a higher frequency of Arabic phrases and loanwords, especially from the *Qurʾān* and other Islamic literature, in the speech of those with some level of Arabic/Islamic education, when addressing a preponderantly Muslim audience.

The speech of ordinary Yoruba speakers in highly Islamized Yoruba communities such as Ilorin and Iwo is usually replete with a higher percentage of loanwords and direct borrowings from Arabic. The incorporation of many Arabic words and phrases into the lyrics of forms of popular Yoruba music is also an important factor in the progressive popularization and assimilation of such words and phrases into the common Yoruba language.

Several semantic fields may be distinguished:

i. Personal names

As is to be expected, Islamic names from Arabic roots constitute the most common form of loanwords. They are usually phonologically harmonized with the Yoruba sound system. Abdullah is usually shortened to *Aábúdù*; Ahmad becomes *Amódú*; Khadija, Ghaffar, Rahman become *Adíjá, Kàfárù*, and *Ràmónì*, respectively.

ii. Religion

Words relating to religious concepts and functionaries form the largest proportion of the Arabic loanwords, due to the symbiotic relationship between Islam and Arabic: *àdúrà, àlùwàlá, àlùbáríkà, àlùkíámọ̀, àráámù, kádàrá, làdání, lèmọ́nmù*, and *yìgì* are illustrative of this category.

iii. Time

Accurate timing is very important for major Islamic religious observances, hence the loanwords *wákàtí* for ‘time’, *àlàmísì* etc. for days of the week, and *àsùbáà* etc. for the prescribed times of the five daily Muslim prayers.

iv. Ethics, morals

Words relating to desirable ethical qualities that must be cultivated and undesirable traits that must be shunned are included in the loans, e.g. *àkíìkà, àlùfànṣá, àlùmọ̀kọ́rọ́yí, mọ̀nọ́fòkí, sàdáńkátá, sùúrù, gááfárà*.

v. Trade, tools, and implements

Muslim traders also propagated the language of trade in their contacts with the Yoruba. It is believed that basic domestic tools such as *abẹ́rẹ́* and *fìtílà* date back to the earliest contact between Yoruba and Arabic. Many of the tools and implements ended up retaining their original names in adapted forms, e.g. *àlìkámò, àlùbọ́sà, fàdákà, gàásì, kafitáàni, tàjíà, tánganran*.

vi. Literacy and education

Islam’s emphasis on education – learning to read and write – is reflected in Arabic loanwords like *ìmọ̀, kálàmù, hàntú, làbárè, tàdáà, tákàdá*.

It is obvious from the list of Arabic loanwords in Table 1 that no compartment of life is untouched by the loanwords. The word *àlááfíà* is uttered several times a day by every Yoruba speaker, regardless of age, sex, social status, or religion, in the exchange of greetings and pleasantries. A survey of the list reveals a wide variety of overlapping topics, which cannot be fully classified into categories. These include words commonly used in various interpersonal interactions, such as *àlámọ̀rí, àǹfààní, àṣírí*; words relating to defects and parts of the human body, such as *akúra, arọ, bàlágà, fùrọ̀, orúkún, orí*; words relating to the immediate environment and the cosmos, such as *àrá* ‘thunder’, *dúníyàn, sọ́nmọ́*; and social and kinship relationships, such as *baálé, ìyáálé*.

Table 1. Selected Arabic loanwords in Yoruba

Yoruba	Arabic	gloss
abẹ́rẹ́	*ʾibra*	‘needle’
àdúrà	*ad-duʿāʾ*	‘prayer, supplication’
àdéùn	*ʿahd(un)*	‘agreement, promise’
àkîìkà	*ḥaqīqa*	‘truth; correctness’
akúra	*ʿaqr*	‘impotent, sterile’; Arabic ‘sterility’
àlááfíà	*al-ʿāfiya*	‘good health; peace’
àláájì	*al-ḥājjī*	‘pilgrim [male]’
àláásàrì	*al-ʿaṣr*	‘afternoon prayer’
àlàmísì	*al-xamīs*	‘Thursday’
àlámọ̀rí	*al-ʾamr*	‘matter, affair’
àléébù	*al-ʿayb*	‘fault; defect’
àléèfà	*xilāfa*	‘(successor’s) seat, caliph’; Arabic ‘caliphate’
àlíámúdù	*al-ḥamd*	‘praise’
àlìkámọ̀	*al-qamḥ*	‘wheat’
àlùbáríkà	*al-baraka*	‘blessing’
àlùbọ́sà	*al-baṣal*	‘onion’
àlùfaǹṣá	*al-faḥšāʾ*	‘obscene, careless [talk, speech]’; Arabic ‘obscene deed’
àlùkíámọ̀	*al-qiyāma*	‘resurrection’
àlùkùránì	*al-Qurʾān*	‘the *Qurʾān*’
àlùmọ́kọ̀rọ́yí	*al-makr*	‘craftiness’
àlùmọ́ọ́nì	*al-ʾamāna*	‘wealth, material possession’; Arabic ‘safety’
àlùmúńtù	*al-mawt*	‘death’
àlùwàlá	*al-wuḍūʿ*	‘ablution’
àǹfààní	*an-naf ʿ/* *ʾanfaʿanī*	‘usefulness, utility’, ‘it benefited me’
àníyaǹ (*àníyọ̀n*)	*an-niyya*	‘intention, concern’
àráámù/ hàráámù	*al-ḥarām*	‘forbidden, unlawful cheating’
arọ	*ʾaʿraj*	‘lame’
àṣírí	*as-sirr*	‘secret, private matter’
aásìkì (*arísìkí*)	*ar-rizq*	‘blessing of God, fortune provision’
àsùbáà	*aṣ-ṣubḥ*	‘dawn prayer’
baálé	*baʿl*	‘husband; patriarch; owner’
bàlágà	*balaġa*	‘to reach [puberty]’
báríkà	*baraka*	‘blessing, congratulation’
dábírà	*dabbara*	‘to perform; wonderful feat’; Arabic ‘to design; manage’
dúńbú	*ḏabaḥa*	‘slaughter’
dúníyaǹ	*dunyā*	‘world’
èrè	*ar-ribḥ*	‘profit, interest’
fàdákà	*fiḍḍa*	‘silver’

Table 1 (*cont.*)

Yoruba	Arabic	gloss
fìtílà	*fatīla*	‘wick lamp’
fìtínátù	*fitna*	‘plague; enchantment; pestering worry’
fùrọ̀	*farj*	‘anus’; Arabic ‘female genitals’
gááfárà	*ġafara*	‘to excuse, pardon, forgive’
gàásì	*kaʾs*	‘wine glass’
hàntú	*xaṭṭu*	‘[Arabic] script’
ìbaada	*ʿibāda*	‘worship, devotion; anything at which one works with seriousness’
ìgbà	*ḥiqba*	‘time, period’
ìmẹ́lẹ́	*al-malal*	‘idleness’
ìmọ̀	*ʿilm*	‘knowledge’
irun	*ar-rakʿa*	‘Muslim canonical prayers’
iwájú	*wajh*	‘front part’
ìyáálé	*ʿiyāl*	‘(senior) wife’; Arabic ‘dependents [pl.]’
jímọ́-ọ̀n	*jumʿa*	‘Friday’
jọ̀nmọ́ọ̀n	*jamāʿa*	‘congregation, assembly’
kádàrá	*qadar*	‘fate, destiny’
kafitáànì	*qafṭān*	‘long-sleeved outer garment’
káfúrà	*kāfūr*	‘camphor’
kálàmù	*qalam*	‘pen’
kèfèrí	*kāfir*	‘unbeliever, heathen’
làbárè	*al-xabar*	‘news, information’
làdánì	*al-ʾaḏān/* *al-muʾaḏḏin*	‘call/caller to Muslim prayer, muezzin’
làákàyè	*al-ʿaql*	‘common sense’
lèmọ́mù	*al-ʾimām*	‘the imam’
màlékà	*malāʾika* [pl.]	‘angel(s)’
màkàrúrù	*mankūr* or *makrūh*	‘dishonest act’
mógíríbù	*maġrib*	‘prayer at sunset, west’
mọ̀ndàla	*madḥ Allāh*	‘may Allah be praised!’; Arabic ‘praise of God’
mọ̀nọ́fìkí	*munāfiq*	‘hypocrisy’
múmínì	*muʾmin*	‘believer’
músííbà	*muṣība*	‘calamity, misfortune’
nọ̀fílà	*nāfila*	‘supererogatory prayer’
nọ̀síà/nàsíà	*naṣīḥa*	‘Muslim sermon, advice’; Arabic ‘advice’
ọdún	*ʿīd*	‘celebration
ọlọ́ṣà	*al-liṣṣ*	‘thief’
orúkún	*ar-rukba*	‘knee’
oṣù	*aš-šahr*	‘month’
orí	*raʾs*	‘head’
rìbá	*ribā*	‘bribery’
sáà	*sāʿa*	‘time, period, hour’
sààrì	*sahūr*	‘predawn meal by Muslims who are fasting’
sábàbí	*sabab*	‘cause, reason’
sàdáńkátà	*ṣadaqta*	‘to praise someone for being truthful, honest’; Arabic ‘you are right’
sàmáánì	*zamān*	‘epoch, time, duration’
sàráà	*ṣadaqa*	‘charity, alms’

Table 1 (*cont.*)

Yoruba	Arabic	gloss
(da) ṣẹ̀ríà	*šariʿa*	‘to impose deserved punishment’
ṣẹ̀ríà	*šarīʿa*	‘sharia, Islamic legal system’
sọ́nmọ̀	*samāʾ*	‘sky’
sìná	*zinā*	‘adultery, illicit sexual relation’
súnà, súnnọ̀n	*sunna*	‘to conform to (Muslim) law, proper, tradition’
tàjíà	*tāj*	‘skull cap’
tàdáà	*dawā*	‘inkwell’
tákàdá	*qirṭās*	‘paper’
tánganran	*tanjara*	‘crockery, ceramics, porcelain’
tànmọ́ọ̀n	*tamannī* or *ṭamʿ*	‘wish, thought, desires’
tásà	*ṯāsa*	‘flat plate’
tira	*ṭīra*	‘amulet’
tááka	*ṭāqa*	‘to show off power, compete with force’; Arabic ‘power, force’
túbá	*tawba*	‘to repent’; Arabic ‘repentance’
wàlááhì	*wa-llāh*	‘I swear by God’; Arabic ‘by God’
wákà	*waqʿa*	‘music’; Arabic ‘rhythm’
wákàtí	*waqt*	‘hour, time’
wáàsí	*waʿḏ̣*	‘sermon’
wàásù	*waʿaḏ̣a*	‘to preach a sermon’
wòlímà	*walīma*	‘marriage, banquet’
yìgì	*ʿiqd*	‘nuptial engagement’

BIBLIOGRAPHICAL REFERENCES

Abdul, M.A. 1976. “Arabic loan words in Yoruba”. *Yoruba: Journal of the Yoruba Studies Association of Nigeria* 2.37–52.

Abraham, R.C. 1962. *Dictionary of Modern Yoruba*. London: Hodder Stoughton.

Ajayi, J.F.A. 1970. “Bishop Ajayi Crowther: An assessment”. *ODU: Journal of West African Studies* (Ile-Ife) N.S. 4.3–17.

Al-Ilori, Adam A. 1972. *ʾAṣl qabāʾ il Yuruba*. Agege, Lagos.

Bello, Muhammad. 1964. *ʾInfāq al-maysūr*. Cairo.

Crowther, Samuel Ajayi. 1852. *A grammar of the Yoruba language*. London: Seeleys.

Fagborun, J. Gbenga. 1994. *The Yoruba koine: Its history and linguistic innovations*. Munich: Lincom Europa.

Gbadamosi, T.G.O. 1978. *The growth of Islam among the Yoruba, 1841–1908*. London: Longman.

Hunwick, John O. 1964. “A new source for the biography of Ahmad Baba al-Tinbukti (1556–1627)”. *Bulletin of the School of Oriental and African Studies* 27.568–593.

——, and Fatima Harrak. 2000. Miʿrāj as-Suʿūd: *Ahmad Baba’s replies on slavery*. Rabat: Institute of African Studies.

Johnson, Rev. Samuel. 1921. *The history of the Yorubas*, ed. O. Johnson. Lagos: Church Missionary Society.

Malik, Sayed H.A. 1995. “The impact of Arabic on the linguistic and cultural life of the Yoruba people”. *Language in Nigeria: Essays in honour of Ayo Bamgbose*, ed. Kolawole Owolabi, 425–439. Ibadan: Group Publishers.

Mazhar, M.A. 1976. *Yoruba traced to Arabic*. Lagos: Ahmadiyya Muslim Mission.

Nigerian Educational Research and Development Council (NERDC). 1991. *Quadrilingual glossary of legislative terms: English/Hausa/Igbo/Yoruba*. Lagos: NERDC.

Oduyoye, Modupe. 1971. *The vocabulary of Yoruba religious discourse*. Ibadan: Daystar Press.

Ogunbiyi, Isaac A. 1980. “Phonological problems faced by Yoruba-speaking learners of the Arabic language” [in Arabic]. *Journal of the Nigerian Association of Teachers of Arabic and Islamic Studies* 11:1.109–127.

——. 1984. “Arabic loan words in Yoruba language in the light of Arab-Yoruba relations from prehistoric times”. *Arab Journal of Language Studies* (Khartoum) 3.147–177.

——. 2001. “Arabic-Yoruba translations of the *Qurʾān*: A socio-linguistic perspective”. *Journal of Qurʾānic Studies* 3.21–45.

Reichmuth, Stefan. 1988. “Songhay-Lehnworter im Yoruba und ihr historischer Kontext”. *Sprache und Geschichte in Afrika* 9.269–299.

Salloum, D. 1976. “al-ʾAlfāḏ al-mustaʿāra min al-ʿArabiyya fī luġat al-Yurub̄a”. *Majalla Kulliyyat al-ʾAdab* 20.

Isaac A. Ogunbiyi (Lagos State University)

Youth Speech

Youth speech is no longer a fringe phenomenon in the Arab world. Rather, it is widespread and has gained momentum by acquiring a normative status through forming its own linguistic and sociolinguistic rules; this is evident in the ongoing documentation and classification of its lexicon, and the public attention paid to this language form. This entry provides examples from youth speech in Egypt (specifically from Cairo) and the Gulf States.

1. Youth language in Egypt

In Egypt, the interest in this variety of speech has culminated in a two-volume dictionary first published in 1999. The second edition (2006) contains more lexemes, expressions, and sentences than the first, leading one to believe that this speech variety will continue to develop and expand.

The dictionary is entitled *The cool dictionary* (*Riwiš ṭaḥn*), where *ṭaḥn* means 'extremely'. *Riwiš*, however, is more problematic. Ibn Manẓūr (d. 711/1311) mentions the etymology of the word in his famous dictionary *Lisān al-ʿArab*, under the two entries *r-š-n* and *r-y-š* (*Lisān* XIII, 180–181; VI, 308–310). The active participle *rāšin* means 'a hungry person, one who is seeking food'. It also means 'a person who takes the opportunity, when he sees people eating, to start begging for food; a parasite, a moocher". *Rāšin* also means 'somebody who does not conform to proper manners and social customs'. Another form is *ʾaryaš*, an adjective which means 'a wealthy person'. There is a noun *riyāš* or *rīš* 'outer surface', from the verb *rāša/yarīšu* 'to be wealthy' (wealth being here *rīš*, which means money and/or other belongings). The origin of this expression is that bird feathers (another meaning of *rīš*), an expensive commodity, were used by the rich as bedding. Ibn Manẓūr quotes Ibn as-Sikkīt (d. 243/857) as saying that *riyāš* is furniture or belongings in general, whereas *rīš* describes belongings or colors.

The etymology of *riwiš* reflects two main concepts closely related to young people today: rebellion and an appreciation for superficial signs of wealth. Rebellion is a property of youth speech, a language which breaks the rules of *fuṣḥā* and *ʿāmmiyya* alike, although it uses the same sounds, and is characterized by the use of coinage and metaphor. The connection between *riwiš* and superficial signs of wealth lies in the fact that a 'cool' person (*riwiš*) is viewed among his peers as someone with a great sense of style. The word 'extremely' (*ṭaḥn*) is used in this context to emphasize the underlying meaning of wealth and coolness.

Upon reviewing the literature on this topic, the following conclusions can be drawn. There are many groups whose argot could be the source for youth speech, such as thieves, convicts, drug addicts, coffee shop waiters, bartenders, barbers, electricians, car mechanics, and, most notably, students and college-educated people (→ jargon). People use → slang for various purposes; thieves use it to avoid the law, convicts use it to conceal existing and potential secrets, and drug addicts use it even among members of the same group in order to avoid being discovered. On the other hand, one of the reasons for young people to use slang, in addition to all the aforementioned reasons, may be in order to vent and forget their frustration with life, boredom, poverty, sense of loss, and despair. These reasons are too intertwined and complicated for a researcher to claim that one is the main motivation behind such language. The pressures of life have caused young people to lose the ability to create a meaningful dialogue with authority and society, and traditional language does not help them change their reality. The boredom, tediousness, and meaninglessness of their lives have driven them to invent their own internal code of communication; society has ignored their ambitions, dreams, and desires, and they face that by ignoring social norms as a way of self-assertion.

To prove this, one only has to consult *The cool dictionary*, whose compiler, Yassir Ḥimāya (the name could be a pseudonym), states in the introduction:

> We [the writer and his generation of cool people] are not frivolous or ignorant, and we have not given up on our causes. Adults, how can you miss that? You were just like us when you were our age. Young people's message to you is only this: 'Understand us!'

There is a sense of complete lack of trust between the youth and the older generations, and a sense of unique cohesion among the youth. One can find evidence for this in the increasing number of secret marriages, another form of common-law marriage (*zawāj ʿurfi*). However, whereas people married under common law were always afraid of being exposed, secret marriages among cool, young people are the result of intergroup trust and lack of faith in the current rules of family and society.

Youth speech is common in universities and other institutions, streets, markets, and coffee shops, the last of which are identified by Web sites as the place to find all fads and trends. As mentioned above, young people feel pressured, and rebellious against society, and therefore abandon their everyday language and adopt a new language of their own, which involves creating their own fields and places, as if they have severed their ties with society, irreversibly alienating themselves from it. Moreover, young people view their language as merely one more aspect of a much-needed change, equal in this respect to changes in the media, theater, women's rights, education, the job market, law, and fashion. This change in language is performed through a distinctly different perspective.

Historically, youth speech, which is now exceedingly common, was documented systematically and academically in 1916 in a pioneering study in the field of folklore by Muḥammad Luṭfī Gumʿa. His essay, entitled *Studies in folklore*, comprises a dictionary of language misuses and workers' slang among common people in Egypt. These include in addition to youth speech misuses and slang used by pickpockets, thieves, con artists, drug addicts, drug dealers, thugs, beggars, jewelers and goldsmiths, upholsterers, belly dancers and their male assistants, who were usually gay, actors, musicians, tent makers, carpenters, construction workers, public bath workers, wheat merchants, cooks, cage makers, gamblers, coffee shop waiters, shoemakers, and monkey trainers.

The essay contains nearly 550 words, expressions, and sentences, some of which can still be observed in speech, and 360 of which were found in the *Riwiš ṭaḥn* dictionary. The researcher classifies them under a number of categories. They are mostly used to describe temporary or permanent states and conditions of young people, especially girls. Intoxication and immorality are also depicted in this language, as well as figurative language, perhaps to hide despair and frustration and sarcasm, and to exaggerate one's experience in order to impress peers. All this usually takes place at parties, where young people engage in social taboos. Tables 1 (one-word lexemes) and 2 (imperatives) list some examples of this language.

Table 1. One-word lexemes (nouns/adjectives/verbs) in youth speech

mitanšin	'unconscious; sad; angry'
tanšana	'complete intoxication following drug use'
miʾantax	'an unemployed graduate'
ʾadrugi	'playboy; business savvy'
fatrīna	'a girl who wears eyeglasses'
buʾla	'business' (as in 'this is my own business')
tirill	'rash, crazy, unpredictable'
tirilla	'a severely overweight girl' (< English *trailer*)
xumfišāri/ xunfišāri	'ambiguous; complicated; a stupid person'
dabbūr	'playboy'
zallūma	'a curious person'
ʿaṣfūra	'a squealer'
mifayyaṣ	'exhausted; useless'
laḥlūḥ	'an Egyptian pound'
ʾastīk	'one thousand Egyptian pounds'
ʾarnab	'a million Egyptian pounds'
ʾasitūk	*ʾastīka* (< French *élastique*?) 'a girl who draws attention to her beauty'
ḥilwa kek	'as sweet/beautiful as cake'
šankūti	'an intrusive/curious/parasitic person'
zanbūʿa/ zanbūʾa	'a person with a long nose; a curious person'
ibtašam	'to torture to death'
ʾaʿallimu	'I'll permanently scar (him)!'
ʾukša/muzza	'a very beautiful girl'
sakamūra	'girlfriend'
sakamūtis	'a nice, well-bred girl'
sīka	'a young army recruit who serves an officer's family off-duty'
xalbūṣ	'playboy'
gilyāṭ	'impolite, ungentlemanlike'
baks	'gullible'
zaġrūf	'a kind of drugs'
ṭabliyya	'a person with a large head'
bastafa	'sarcastic insults'
istrubyā	'a person not to be taken seriously' (< French *estropié* 'crippled'?)
ʾitim	'unfunny, unliked'
zumbagi	'a person who maliciously spreads gossip to hurt people'

Table 2. Imperative sentences/commands

uhruš	'give me money!'
insāni w xud ʿinwāni	'don't call me again, it's over!'
idʿak ilfanūs	'danger, get out of here!'
inzil min ʿala widni	'leave me alone, you've given me a headache!'

Table 1 (*cont.*)

uʾʿudli f ḥitta našfa	‘you’re too young and inexperienced to talk to me!’
ʾaṣṣar wi limm ilmitkassar	‘change the subject, don’t argue with me anymore!’
inzil min ʿal masraḥ	‘stop arguing, you’ve had your chance to talk, good-bye!’
kabbar ilgumguma w rayyaḥ iddāl	‘try to understand, don’t give yourself a headach over it!’
dablig ilmasāʾil	‘help me tone down the tension!’
rawwaʾ ilʾanāni	‘don’t be angry!’
ixlaʿ ʾird	‘be quiet and change the subject!’ (said when an unwanted person approaches)

Complete sentences (declarative or predicative) are sometimes used in youth language. Although such sentences might give the impression of reporting information, they are, in fact, a reflection of the speaker’s experience. Examples include the following: *iddafš illi fašlak gaḥmizu w ḥazzibu* ‘benefit from him as much as you can!’; *sakkaḥ maragīḥ ʿaddayyaʾ* ‘life’s tough’; *iššaxramūn ṭāḥ f ittaralullī* ‘there’s no use trying, it’s all in vain’; *ilʿafša ʾafašit* ‘it’s difficult’; *fi rāsu bargal* ‘he is a good planner’ (usually about plots and conspiracies); *markit ʾiftaḥ ilkabbūt w iʾfilu* ‘like a dishonest auto mechanic’; *ana hāriš malḥitu* ‘I know everything about him, he can’t get away from me’; *miš f il furma* ‘not in the mood’; *dayra ʿala ḥall šaʿraha* ‘she is an irresponsible, immoral woman’; *ilʿamaliyya f innamliyya* ‘don’t worry, everything is fine’; *maʿāk ʿal-xaṭṭ* ‘I agree with everything you say’; *ʾē ʾāxir iftikasātak?* ‘what will you come up with next?’ (sarcastic); *balāš nixabbat f ilḥilal* ‘no need to expose each other/air out our dirty laundry’; *ẓarrafni tiʿrafni/ʾabbigni tagidni* (said when asking for a bribe); *ʾalbu ʾalb xaṣṣāya* ‘he is very kind and pure’; *iḥna-lli dahanna lhawa dūkū (būya)/iḥna-lli xaramna ttaʿrīfa/iḥna-lli bahyaẓna lfahāyiṣ* ‘we know what we’re doing’; *gab gāz* (said of someone who has reached a severe degree of exhaustion); *gālli garab fi wdāni* ‘I am sick of listening to your nonsense’; *ḥaṭṭ ʾīdu ʿal-balf* ‘he has figured out the truth’; *f ilhazīʿ w ilhazīʿ* ‘in the throes of passion’; *f ilʾawanṭa* ‘a liar, a know-it-all’; *da f ilhušūm* ‘he is disconnected from reality’.

Coffee shop orders that are characteristic of youth speech include: *wāḥid mustašfa* ‘one aniseed drink’; *wāḥid wilāda* ‘one fenugreek’, *wāḥid ṣuwān* ‘one black Turkish coffee’. Youth speech contains a number of English expressions, e.g. *ist-morning* ‘food/drinks/money that would help one recover from a hangover’; *another stone* ‘more coal for the hookah’; *enter* ‘watch out [because another participant in the conversation cannot keep a secret]!’ (the same as the computer command). Perhaps as an expression of absurdity, a frequent phenomenon in Egyptian youth speech is the use of meaningless sentences, e.g. *ḥuṭṭ šwayyit mayya našfa maʿ šwayyit mayya mablūla* ‘put some dry water with some wet water!’.

2. Youth speech in the Gulf

The situation of youth speech in the Gulf States is similar to that in Egypt. A United Arab Emirates daily newspaper, *The Gulf*, published a report that took up almost an entire page on 25 December 2007 (issue no. 10446). According to the report, written by Azza Sanad, youth speech has become widespread in the Gulf States, particularly in Bahrain, Saudi Arabia, Kuwait, and the United Arab Emirates. The report maintains that this variety has recently evolved among young people in the Gulf, either to rebel against adults or as a playful attempt to invent something new and individualistic. This language does not conform to traditional linguistic rules and standards, as it either consists of existing words and expressions whose meaning has been twisted or altered, or new words and expressions, puns, and borrowed words. It has, therefore, acquired its own style, which is distinctive from the language used by older generations. Many of its words are borrowed from English, mostly related to computers and the Internet.

The following expressions have been noted in Bahrain: *ʿṭeh mir-rixīṣ* ‘lie to him!’; *ʾazrufhā* ‘steal it!’, *fattīt* ‘a great soccer player’, *ʾarūḥ malḥ* ‘I melt’, *dabj* ‘beating’, *ʾizbidha* ‘cut to the chase!’, *condition* (< English) ‘a very beautiful girl’, *tibṭi ʿaḏ̣m* ‘there’s no hope’, *deṣ šeṣ* ‘subtly’, *fīk laffa* ‘you’re crazy’, *mitgarwiš* ‘worried, running away from something’, *mhawjis* ‘not paying attention’.

The following expressions have been noted in Jeddah (Saudi Arabia): *ʾabū kalb* ‘broken [e.g. cell phone or pen]’, *ʾiddīh ʾuškul* ‘ignore him!’, *inbaraš* (said to someone who makes irrelevant comments), *balf* ‘rash, irrational’, *tašbīk* ‘crush’, *taġrīz* ‘falling in love’, *ʾiʿṭīlu janṭ* ‘ignore him!’, *dāfūr* ‘clever; hardworking’, *daʿʿis* ‘come on, move along!’, *zayy ilkartūt* ‘like an idiot’, *šakūš* ‘an unattractive girl’.

The following expressions have been noted in Kuwait: *ana rāyiḥ alʿab maʿa ruby* ‘I’m going to the gym’, *ʾinta bitʾallid ruby* ‘you’re being secretive, hiding something’, *mikayyiš ilʾufaf* ‘having a large amount of money’, *ʾukša* ‘a beautiful girl’, *ḥitta šampū* (< English) ‘it smells nice’, *ʾinta taksi* (< English) ‘you’re a playboy’, *miliskubitš* ‘stingy’, *xufšari* ‘air-headed’.

Some young people in the United Arab Emirates believe that their sense of pride in their language prevents them from using such forms of speech, although some still use it playfully. Others consider it among the rights of young people in any country to coin their own language that reflects their own problems, and that this is the generation of trends, common-law marriage, and unemployment. There is also a third opinion, claiming that this language is an extension of an older variety, coined by past generations of young people in the United Arab Emirates, which has by now become a fixed part of the language. To this older variety belong expressions like *xabīla* ‘unintelligent’, *xarīša* ‘an unwise young man’, *suk daʿna* ‘sit down!’, *giṭʿa* ‘a beautiful girl’. New expressions include *zabūṭ innagʿa* ‘someone who is slightly built’ (< *zabūṭ* ‘type of shellfish’, and *nagʿa* ‘shore’), *jimbāzi* ‘lying, conniving’, *ʿaflangī* ‘manipulative, disreputable’, *minfihra* ‘impolite’, *yitzankaḥ* ‘he is walking slowly and indifferently’, *imzahlag* ‘straightened up after being in a state of chaos’.

3. Conclusion

Questions for consideration are where this language is headed, and whether it is expanding or receding. Presumably, it will expand as long as its motivating factors continue to exist, especially because it gives young people a feeling of indifference and irresponsibility, and therein lies its ultimate attraction.

In Egypt, although youth speech has spread in Cairo and other major Egyptian cities, there is no evidence of its spread in rural areas in Upper or Lower Egypt. The only exception would be some features that have crept in by way of satellite channels through television shows. This language has also been observed in modern Egyptian popular songs, which have established a wide fan base among young people and are constantly played in the media, making youth speech less controversial.

In the Gulf States, youth speech remains controversial. On the one hand, those who support it claim that it is more realistic, economic, comfortable, and richer than adult speech, which is more complex and prone to abstraction and description. The lack of complex, rigorous rules in this language also simplifies communication, which makes young people in the Gulf use it as a flexible communication tool for all topics. On the other hand, those who oppose it consider it a temporary phenomenon that will not last due to its being limited to a single group of users, and even a single ‘clique’. This belief may stress the pride people in the United Arab Emirates take in their language and local dialects.

Bibliographical references

Primary source

Ibn Manẓūr, *Lisān* = ʾAbū l-Faḍl Jamāl ad-Dīn Mukarram ibn Mukarram Ibn Manẓūr, *Lisān al-ʿArab*. 20 vols. Cairo: Dār al-Maʿārif, n.d.

Secondary sources

Allam, Jehan. 2000. “Youngsters and new vocabulary”. *al-Lugha, langue, language: Focus on Egyptian Arabic*, ed. Gerda Mansour and Madiha Doss, 111–135. Cairo: Cairo Linguistics Group, Arab Research Center.

Gumʿa, Muḥammad Luṭfī. 1916. “Studies in folklore”. *Library of Public Studies* 34.

Ḥimāya, Yassir. 2006. *Qāmūs riwiš ṭaḥn*. Cairo: Al-Hussam Publishing and Distribution. (1st ed., 1999.)

Peterson, Jennifer Leigh. 2002. “Contemporary Cairene youth terminology: Linguistic deviation or social art?”. *Aspects of the dialects of Arabic today*, ed. Abderrahim Youssi Fouzia Benjelloun, Mohamed Dahbi, and Zakia Iraqi-Sinaceur, 422–429. Rabat: Amapatril.

Ahmad Taher Hassanein
(Misr University for Science and Terminology).
Translation: Injie Hejazi

Zāʾid → Ṣila

Encyclopedia of Arabic Language and Linguistics

Volume V
Index

GENERAL EDITOR
Kees Versteegh
(University of Nijmegen)

ASSOCIATE EDITORS
Mushira Eid
(University of Utah)
Alaa Elgibali
(University of Maryland)
Manfred Woidich
(University of Amsterdam)
Andrzej Zaborski
(University of Cracow)

ADVISORY BOARD
Ramzi Baalbaki *(American University of Beirut)*
Elsaid Badawi *(American University of Cairo)*
Dominique Caubet *(INALCO, France)*
Clive Holes *(University of Oxford)*
Manfred Kropp *(Orient-Institut Der Deutschen Morgenländischen Gesellschaft)*
Jérôme Lentin *(Institut National des Langues et Civilisations Orientales)*
John McCarthy *(University of Massachusetts)*
Jamal Ouhalla *(University College Dublin)*
Jan Retsö *(Göteborg University)*
Sabah Safi *(King Abdulaziz University)*

ENCYCLOPEDIA OF ARABIC LANGUAGE AND LINGUISTICS

VOLUME V
Index

General Editor
Kees Versteegh

Associate Editors
Mushira Eid
Alaa Elgibali
Manfred Woidich
Andrzej Zaborski

BRILL

LEIDEN • BOSTON
2009

Cover Illustration: Handwritten manuscript page by Naguib Mahfouz (1911–2006).

Despite our efforts we have not been able to trace all rights holders to some copyrighted material. The publisher welcomes communications from copyrights holders, so that the appropriate acknowledgements can be made in future editions, and to settle other permission matters.

© Copyright 2009 by Koninklijke Brill NV, Leiden, The Netherlands
Koninklijke Brill NV incorporates the imprints Brill, Hotei Publishing, IDC Publishers,
Martinus Nijhoff Publishers and VSP

All rights reserved. No part of this publication may be re-produced, translated, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without prior written permission from the publisher.

Authorization to photocopy items for internal or personal use is granted by Brill provided that the appropriate fees are paid directly to The Copyright Clearance Center, 222 Rosewood Drive, Suite 910,
Danvers MA 01923, USA.
Fees are subject to change.

PRINTED IN THE NETHERLANDS

ISBN: 978 90 04 14973 1 (Set)
ISBN: 978 90 04 17484 9 (Index)

This book is printed on acid-free paper.

Cover design: BEELDVORM, Pijnacker

LEMMA LIST

VOLUME ONE

A – Ed

Articles *Page numbers*

A

Abbreviations (Adam Gacek) 1–5
Acehnese (Edwin Wieringa) 5–12
Adjective Phrase (Alain Kihm) 12–16
Adjectives (Wolfdietrich Fischer) 16–21
Adverbs (Janet C.E. Watson) 21–25
Affrication (Chakir Zeroual) 26–28
Afghanistan Arabic (Bruce Ingham) 28–35
Afro-Asiatic Languages (Andrzej Zaborski) 35–40
Agent (Hussein Abdul-Raof) 40–43
Agreement (Maher Bahloul) 43–48
ʿĀʾid (Yishai Peled) 49–50
Aktionsart (Johannes Reese) 50–53
Algeria (Jacques Grand'Henry) 53–58
Algiers Arabic (Aziza Boucherit) 58–66
ʿAmal (Valeriy Rybalkin) 67–74
Analogy (Robert Ratcliffe) 74–82
Analytic Genitive (Kerstin Eksell) 82–85
Anaptyxis (Petr Zemánek) 85–86
Anatolian Arabic (Otto Jastrow) 87–96
Andalus (Otto Zwartjes) 96–101
Andalusi Arabic (Federico Corriente) 101–111
Antiochia Arabic (Werner Arnold) 111–119
Apophony (Sabrina Bendjaballah) 119–123
Apposition (Mohssen Esseesy) 123–126
ʿArab (Jan Retsö) 126–133
Arabic Alphabet for Other Languages (Alan S. Kaye) 133–147
Arabic Alphabet: Origin (Beatrice Gruendler) 148–165
Arabic Studies in Europe (Alastair Hamilton) 166–172
ʿArabiyya (Yasir Suleiman) 173–178

Aramaic/Syriac Loanwords (Jan Retsö) 178–182
Argument (John C. Eisele) 182–187
Article, Definite (Andrzej Zaborski) 187–188
Article, Indefinite (Lutz Edzard) 188–191
ʾAṣl (Ramzi Baalbaki) 191–195
Aspect (John C. Eisele) 195–201
Asseverative Particle (David Testen) 201–204
Assimilation (Petr Zemánek) 204–206
Automatic Language Processing (Noureddine Chenfour) 206–216
Auxiliary (Martine Haak) 216–221

B

Baghdad Arabic (Farida Abu-Haidar) 222–231
Baghdad Arabic Jewish (Jacob Mansour) 231–241
Bahraini Arabic (Clive Holes) 241–255
Bambara (Dinie Bouwman) 255–257
Bangladesh (Ali Riaz) 257–259
Bedouin Arabic (Judith Rosenhouse) 259–269
Bedouinization (Farida Abu-Haidar) 269–274
Beirut Arabic (Samia Naïm) 274–286
Bengali (James M. Wilce) 286–289
Berber (Jordi Aguadé and Peter Behnstedt) 289–293
Berber Loanwords (Abderrahman El Aissati) 293–299
Bʿēri Arabic (Manfred Woidich) 299–308
Bināʾ (Ramzi Baalbaki) 308–310
Binding (Joseph Aoun) 310–313
Biradicalism (Andrzej Zaborski) 313–316
Braille (Muhammad Al-Sharkawi) 316–320
Bukaṛa-syndrome (Rudolf de Jong) 320–322

C

Cairo Arabic (Manfred Woidich) 323-333
Cameroon Arabic (George Echu and Mohamadou Aminou) 333–339
Caretaker Talk (Khawla Aljenaie) 339–343
Case Roles (Hussein Abdul-Raof) 343–347
Case Theory (Mark S. LeTourneau) 347–353
Cataphora (Yishai Peled) 353–355
Causative (Adrian Măcelaru) 355–360
Chad Arabic (Patrice Jullien de Pommerol) 360-368
Child Bilingualism (Ghada Khattab) 368–375
Child Language (Mary-Ann Walter) 375–378
China (Leila Chérif-Chebbi) 378–383
Christian Middle Arabic (Jacques Grand'Henry) 383–387
Cilician Arabic (Stephan Procházka) 388–397
Classical Arabic (Wolfdietrich Fischer) 397–405
Classicism (Heikki Palva) 405–411
Clitic (Joseph Aoun) 411–414
Code-switching (Gunvor Mejdell) 414–421
Cognitive Linguistics (Sami Boudelaa) 421–427
Coherence (Said Faiq) 427–430
Cohesion (Esam N. Khalil) 430–433
Collective (Ignacio Ferrando) 433–434
Collocation (Dalal Mahmoud El-Gemei) 434–439

Colloquial (Soha Abboud-Haggar) 439–442
Communal Dialects (Keith Walters) 442–448
Comoros (Ahmed Chanfi) 448–451
Compounds (Abdul Sahib Mehdi Ali) 451–455
Computational Linguistics (Everhard Ditters) 455–465
Concessive Clause (Michael Langer) 465–467
Conjunctions (Michael Waltisberg) 467–470
Connectives (Heléne Kammensjö) 470–477
Construct State (Elabbas Benmamoun) 477–482
Contrastive Grammar (Utz Maas) 482–489
Convergence (Lutz Edzard) 489–495
Coptic (Tonio Sebastian Richter) 495–501
Coptic Loanwords (Peter Behnstedt) 501–505
Copula (Maher Bahloul) 506–511
Corpus Linguistics (Everhard Ditters and Jan Hoogland) 511–518
Creole Arabic (Jonathan Owens) 518–527
Culture and Language (Niloofar Haeri) 527–536
Cypriot Maronite Arabic (Alexander Borg) 536–543

D

Ḍād (Kees Versteegh) 544–545
Damascus Arabic (Jérôme Lentin) 546–555
Ḍamīr (Yishai Peled) 555–559
Declension (Lutz Edzard) 559–564
Defective Verbs (Rainer Voigt) 564–565
Deixis (Mihai Dat) 565–569
Demonstratives (Ángeles Vicente) 569–573
Derivation (Pierre Larcher) 573–579
Determiners (Lina Choueiri) 579–583
Dialect Geography (Peter Behnstedt) 583–593
Dialect Koine (Catherine Miller) 593–597
Dialect Literature (Humphrey Davies) 597–604
Dialects: Classification (Heikki Palva) 604–613
Dialects: Genesis (Soha Abboud-Haggar) 613–622
Diathesis (Jan Retsö) 622–626
Ḍidd (Lidia Bettini) 626–629
Diglossia (Naima Boussofra-Omar) 629–637
Diminutive (Tamar Zewi) 637–640
Diphthongs (Tamás Iványi) 640–643
Diptosis (André Roman) 643–647
Discourse Analysis (Ahmed Fakhri) 647–653
Dissimilation (Petr Zemánek) 653–654
Djibouti/Eritrea (Marie-Claude Simeone-Senelle) 654–656
Dysphemism (Mohammed Farghal) 656–659

E

East Africa (Abdul Aziz Y. Lodhi) 660–666
Educated Arabic (Karin C. Ryding) 666–671

VOLUME TWO

Eg – Lan

Articles *Page numbers*

E

Egypt (David Wilmsen and Manfred Woidich) 1–12
Elative (Alain Girod) 12–16
Elision (Michael G. Carter) 16–18
Ellipsis (Mustafa Mughazy) 18–21
Enclisis (David Testen) 21–22
Energicus (Tamar Zewi) 22–25
English (Alan S. Kaye) 25–29
English Loanwords (Ahmad Atawneh) 29–35
Epenthesis (Samira Farwaneh) 35–39
Epigraphy (Islamic) (Sheila S. Blair) 40–50
Ethiopia (Andreas Wetter) 51–56
Ethiopic Loanwords (Stefan Weninger) 56–57
Ethnicity and Language (Eirlys E. Davies and Abdelâli Bentahila) 58–65
Etymology (Václav Blažek) 65–69
Euphemism (Mohammed Farghal) 69–72
Europe (Jan Jaap de Ruiter) 72–79
Exclamation (Daniela Rodica Firanescu) 79–81

F

Fāʿil (Hassan Salam Bazzi Hamzé) 82–84
Faṣīḥ (Georgine Ayoub) 84–90
Fiʿl (Hassan Salam Bazzi Hamzé) 90–96
First Language Acquisition (Fatima Badry Zalami) 96–104
First Language Teaching (Kassem M. Wahba) 104–113
Focus (Sven-Olof Dahlgren) 113–116
Foreigner Talk (Muhammad al-Sharkawi) 116–122
French Loanwords (Carole Paradis and Darlene LaCharité) 122–131
Fronting (Wafaa ʿAbd al-Fahim Batran Wahba) 131–136
Frozen Expression (Elizabeth M. Bergman) 136–137
Fulfulde (Rolf Theil) 137–142
Functional Grammar (Ahmed Moutaouakil) 143–150

G

Gahawa-Syndrome (Rudolf de Jong) 151–153
Gemination (Samuel Rosenthall) 153–155
Gender (Atiqa Hachimi) 155–164
Glide (Abdellah Chekayri) 164–169
Government (Mark S. LeTourneau) 169–175
Grammatical Tradition: Approach (Jean-Patrick Guillaume) 175–182
Grammatical Tradition: History (Michael G. Carter) 182–191
Grammaticalization (Mohssen Esseesy) 191–198
Greek Loanwords (Dimitri Gutas) 198–202
Greetings (Nagwa Elzeiny) 202–207
Grounding (Esam N. Khalil) 207–210
Gulf States (Clive Holes) 210–216
Gypsy Arabic (Yaron Matras) 216–222

H

Ḥaḏf (Jacques Grand’Henry) 223
Ḥāl (Monique Bernards) 224–228
Hamza (Andries W. Coetzee) 228–232
Ḥaraka (Kees Versteegh) 232–236
Ḥarf (Samvel Karabekyan and Marat Yavrumyan) 236–239
Ḥassāniyya Arabic (Catherine Taine-Cheikh) 240–250
Hausa (Al-Amin Abu-Manga) 250–256
Ḥimyaritic (Christian Julien Robin) 256–261
History of Arabic (Ignacio Ferrando) 261–268
Horn of Africa (Marie-Claude Simeone-Senelle) 268–275
Hypercorrection (Benjamin Hary) 275–279

I

ʾIbdāl (Jaakko Hämeen-Anttila) 280–281
Ibero-Romance (Reinhard Kiesler) 281–286
Ibero-Romance Loanwords (Federico Corriente) 287–290
Ibtidāʾ (Choukri Iraqi Houssaini) 290–293
ʾIḍāfa (Karin C. Ryding and Kees Versteegh) 294–298
ʾIdġām (Janusz Danecki) 298–300
ʾIḍmār (Michael C. Carter and Kees Versteegh) 300–302
ʾIʿjāz (Sophia Vasalou) 302–307
ʾIlġāʾ (Kees Versteegh) 307–308
ʿIlla (Kees Versteegh) 308–311
ʾImāla (Aryeh Levin) 311–315
Impersonal Verb (Alain Girod) 315–318
Implicational Scale (Maik Gibson) 318–325
India (Tahera Qutbuddin) 325–331
Indirect Speech (Madiha Doss) 331–333
Indonesia (Karel Steenbrink) 333–339
Indonesian/Malay (Stuart Campbell) 340–345
Inflection (Djamel Eddine Kouloughli) 345–354
ʾInna wa-ʾaxawātuhā (Kees Versteegh) 354–358
ʾInšāʾ (Pierre Larcher) 358–361
Insults (Evelyne Larguèche) 361–365
Interface Linguistics (Peter Hallman) 365–372
Interference (David Wilmsen) 373–378
Interjection (Herbert Eisenstein) 378–380
Internet (Waheed Samy) 380–387
Interrogative Pronoun (Kees Versteegh) 387–389
Interrogative Sentences (Moha Ennaji) 389–395
Intonation (Dana Chahal) 395–401
ʾIʿrāb (Kinga Dévényi) 401–406
Iran (Éva M. Jeremiás) 406–414
Iraq (Otto O. Jastrow) 414–424
Ism (Monique Bernards) 424–429
Ism al-fāʿil (Michael G. Carter) 429–432
Ism al-fiʿl (Aryeh Levin) 432–434
ʾIsnād (Kees Versteegh) 434–437
Israel (Muhammad Hasan Amara) 437–441
Istiʿāra (Udo Simon) 441–447
Ištiqāq (Abdellah Chekayri) 447–451

Istiṯnāʾ (Monique Bernards) 451–454
Italian (Guido Cifoletti) 454–456
Italian Loanwords (Guido Cifoletti) 456–459
ʾIṭbāq (Muhammad Hasan Bakalla) 459–461
Ivrit (Aharon Geva-Kleinberger) 461–464
Ivrit Loanwords (Muhammad Hasan Amara) 464–467

J

Jargon (Elizabeth M. Bergman) 468–472
Javanese (Rochayah Machali) 472–477
Jazāʾ (Kinga Dévényi) 477–481
Jerusalem Arabic (Judith Rosenhouse) 481–493
Jīm (Andrzej Zaborski) 494–496
Jinās (Lidia Bettini) 496–498
Jordan (Mohammed Sawaie) 498–505
Jordanian Arabic (Amman) (Enam Al-Wer) 505–517
Juba Arabic (Catherine Miller) 517–525
Judaeo-Arabic (Geoffrey Khan) 526–536
Jumla (Tamás Iványi) 536–540

K

Kalām (Tamás Iványi) 541–545
Kalima (Aryeh Levin) 545–548
Kāna wa-ʾaxawātuhā (Aryeh Levin) 548–552
Kanuri (Sergio Baldi) 552–555
Kaškaša and Kaskasa (Munira Ali Al-Azraqi) 555–557
Kazakh (Rob Ermers) 557–559
Khartoum Arabic (James Dickins) 559–571
Khuzestan Arabic (Bruce Ingham) 571–578
Kināya (Joseph Dichy) 578–583
Kinship Terms (Mohammed Farghal) 583–588
Ki-Nubi (Ineke Wellens) 588–597
Kufic (Sheila S. Blair) 597–603
Kurdish (Michael Chyet) 604–608
Kuwaiti Arabic (Clive Holes) 608–620

L

Labiovelarization (Dominique Caubet) 621–623
Lafḏ̣ (Djamel Eddine Kouloughli) 623–628
Laḥn (Georgine Ayoub) 628–634
Language Academies (Mohammed Sawaie) 634–642
Language and Gender (Fatima Sadiqi) 642–650
Language Attitudes (Keith Walters) 650–664
Language Contact (Sarah G. Thomason) 664–674
Language Impairment (Mohammad T. Alhawary) 674–681
Language Loss (Abderrahman El Aissati) 681–687
Language Pathology (Sabah M.Z. Safi) 687–694
Language Policies and Language Planning (Kassim Shaaban) 694–707
Language Shift: Amazigh (Yamina El Kirat El Allame) 707–716

VOLUME THREE

Lat – Pu

Articles *Page numbers*

L

Latin America (Ernesto Capello) 1–6
Latin Loanwords (Irfan Shahîd) 6–8
Leveling (Reem Bassiouney) 8–13
Lexical Variation: Modern Standard Arabic (Zeinab Ibrahim) 13–21
Lexicography: Bilingual Dictionaries (Jan Hoogland) 21–30
Lexicography: Classical Arabic (Tilman Seidensticker) 30–37
Lexicography: Monolingual Dictionaries (Ahmed Taher Hassanein) 37–45
Lexicon: Matrix and Etymon Model (Georges Bohas and Mihai Dat) 45–52
Libya (Christophe Pereira) 52–58
Lingua Franca (Catherine Miller) 58–64
Linguistics and Arabic (Bernard Comrie) 64–71
Lisān (Ahmad Taher Hassanein) 71–74
Literacy (Mohamed Maamouri) 74–80
Locatives (Mushira Eid) 80–88
Luġa (Tamás Iványi) 88–95

M

Māḍī and Muḍāriʿ (Michael G. Carter) 96–100
Mafʿūl (Zeinab Ahmed Taha) 100–106
Mafʿūl fīhi (Kees Versteegh) 106–110
Maġribī (Adam Gacek) 110–113
Maḥmūl (Miklós Maróth) 113–116
Majāz (Udo Simon) 116–123
Majhūra/Mahmūsa (Janusz Danecki) 123–125
Malagasy (Narivelo Rajaonarimanana) 125–128
Malayalam (R.E. Asher and M.N. Karassery) 128–135
Mali (Dinie Bouwman) 135–141
Malta (Joseph M. Brincat) 141–145
Maltese (Manwel Mifsud) 146–159
Maʿnā (Djamel Eddine Kouloughli) 159–164
Maṣdar (Everhard Ditters) 164–169
Mauritania (Catherine Taine-Cheikh) 169–176
Mawḍūʿ (Miklós Maróth) 176–178
Meccan Arabic (Mahasen Hasan Abu-Mansour) 179–187
Mechanisms of Linguistic Change (Reinhard Stempel) 187–192
Media (Ragia M. Effat) 192–199
Media Arabic (Ragia M. Effat and Kees Versteegh) 199–204
Metathesis (Elizabeth Hume) 204–207
Meter (Dmitry Frolov) 207–215
Middle Arabic (Jérôme Lentin) 215–224
Middle Verbs (Zouhair Maalej) 224–227
Mimation (Michael P. Streck) 227–228
Minimalism (Joost Kremers) 228–233
Modal Verbs (Daniela Rodica Firanescu) 233–238
Modern Standard Arabic (Ernest N. McCarus) 238–262

Mood (Standard Arabic) (Robert Fradkin) 269–273
Moroccan Arabic (Dominique Caubet) 273–287
Morocco (Jorge Aguadé) 287–297
Morphology (John J. McCarthy) 297–307
Muḫaqqaq (Adam Gacek) 307–311
Multilingualism (Abdelali Bentahila and Eirlys E. Davies) 311–320
Muštarak (Lidia Bettini) 320–323
Mutarādif (Lidia Bettini) 323–325

N

Najdi Arabic (Bruce Ingham) 326–334
Nasalization (Muhammad Hasan Bakalla) 334–336
Nastaʿlīq (Adam Gacek) 336–339
Nasx (Adam Gacek) 339–343
Nationalism and Language (Elie Kallas) 343–353
Nawāsix (Nadia Anghelescu) 353–354
Negation (Jamal Ouhalla) 355–360
Negev Arabic (Roni Henkin) 360–369
Neo-Aramaic (Werner Arnold) 370–373
Nigeria (ʿAbdulFattah Olayiwola) 373–377
Nisba (Jean N. Druel and René-Vincent du Grandlaunay) 377–381
Nominal Clauses (Frederick Hoyt) 381–388
Nominalization (Mustafa Mughazy) 388–394
North America (R. Kirk Belnap) 394–400
Northwest Arabian Arabic (Heikki Palva) 400–408
Northwest Semitic Languages (Rebecca Hasselbach and John Huehnergard) 408–422
Noun (Lutz Edzard) 422–428
Noun Phrase (Frederick Hoyt) 428–434
Nubian (Angelika Jakobi) 435–439
Number (Robert R. Ratcliffe) 439–447
Numerals (Catherine Taine-Cheikh) 447–453
Nunation (Franz-Christoph Muth) 453–454

O

Object, Absolute (Moheiddin Ali Homeidi) 455–461
Obligatory Contour Principle (Samuel Rosenthall) 461–464
Old Arabic (Epigraphic) (Michael C.A. Macdonald) 464–477
Omani Arabic (Clive Holes) 478–491
Optimality Theory (Lutz Edzard) 491–499
Orality (Elizabeth M. Bergman) 499–501
Ottoman Empire (Gottfried Hagen) 501–505

P

Pakistan (Tariq Rahman) 506–512
Palaeography (Petra M. Sijpesteijn) 513–524
Palatalization (Chakir Zeroual) 524–526
Palestinian Arabic (Kimary N. Shahin) 526–538
Paronomasia (Avihai Shivtiel) 538–541
Participle (Jonathan Owens and Marat Yavrumyan) 541–546
Parts of Speech (Michael G. Carter) 546–552
Passive (Vít Bubeník) 552–558
Passive (Syntax) (Amira Agameya) 558–564
Pausal Forms (Robert D. Hoberman) 564–570

Performatives (Mustafa Mughazy) 570–573
Persian (John R. Perry) 573–580
Persian Loanwords (Asya Asbaghi) 580–584
Personal Pronoun (Arabic Dialects) (Ángeles Vicente) 584–588
Personal Pronoun (Standard Arabic) (David L. Appleyard) 588–593
Phonetics (Salman H. Al-Ani) 593–603
Phonological Merger (Enam Al-Wer) 603–605
Phonological Split (Enam Al-Wer) 605–607
Phonology (Ellen Broselow) 607–615
Phonology: Metrical (Bruno Paoli) 615–623
Phonotactics (Stefan A. Frisch) 624–628
Phraseology (Avihai Shivtiel) 628–634
Pidgin Arabic: Bongor Arabic (Xavier Luffin) 634–639
Pidginization (Mauro Tosco) 639–644
Poetic Koine (Kees Versteegh) 644–647
Poetic License (Geert Jan van Gelder) 647–652
Polarity (Frederick M. Hoyt) 652–658
Politeness (Avihai Shivtiel) 658–663
Political Discourse and Language (Nathalie Mazraani) 663–671
Possession (Samia Naïm) 671–676
Pragmatics (Mustafa Mughazy) 676–683
Predicate (Frederick M. Hoyt) 683–689
Pre-Islamic Arabic (Muhammad Al-Sharkawi) 689–699
Prepositions (Stephan Procházka) 699–703
Presentatives (Geoffrey Khan) 703–705
Pro-Drop (Mushira Eid) 705–713
Pronominalization (Ahmed-Sokarno Abdel-Hafiz) 713–717
Proper Names (Franz-Christoph Muth) 717–724
Prosody (Thami Benkirane) 724–728
Prothetic Vowel (Munther A. Younes) 728–730
Proverb (Avihai Shivtiel) 730–736
Pseudodual (Kerstin Eksell) 736–739
Pseudoverb (Bernard Comrie) 739–740
Punctuation (Siamak Rezaei) 740–742

VOLUME FOUR

Q – Y

Articles *Page numbers*

Q

Qāf (Lutz Edzard) 1–3
Qirāʾāt (Mustafa Shah) 4–11
Qiyās (Miklós Maróth) 11–14
Quantifier (Peter Hallman) 14–20
Qurʾān (Muhammad Abdel Haleem) 21–31

R

Rajaz (Jaakko Hämeen-Anttila) 32–37
Reanalysis (Mohssen Esseesy) 37–43
Reciprocals (Joost Kremers) 43–45
Redundancy (Ali Farghaly) 45–50
Reduplication (Dina El Zarka) 50–53
Reflexives (Joost Kremers) 53–57
Register (Reem Bassiouney) 57–60
Relative Clause (Mustafa Mughazy) 60–70
Relative Pronoun (Arabic Dialects) (Ángeles Vicente) 70–72
Religion and Language (Stefan Reichmuth) 72–80
Resumption (Joseph Aoun) 80–85
Rhyme (Soheir G. Mahfouz) 86–92
Root (Petr Zemánek) 93–98
Ruqʿa (Adam Gacek) 98–100

S

Sabab (Michael G. Carter) 101–102
Sajʿ (Gert Borg) 103–106
Ṣanʿānī Arabic (Janet Watson) 106–115
Sandhi (Robert D. Hoberman) 115–118
Ṣarf (Joyce Åkesson) 118–122
Saudi Arabia (Bruce Ingham) 123–130
Script and Art (Adam Gacek) 130–138
Second Language Acquisition (Mohammad T. Alhawary) 138–146
Second Language Teaching (Helle Lykke Nielsen) 146–156
Secret Languages (Abderrahim Youssi) 156–160
Semantic Bleaching (Mohssen Esseesy) 160–164
Semantic Extension (Mohssen Esseesy) 164–169
Semitic Languages (Rainer Voigt) 170–179
Senegal (Fiona McLaughlin) 179–185
Sentence (Moha Ennaji) 185–191
Sentence Coordination (Heidi Lorimor) 191–195
Serial Verbs (Kees Versteegh) 195–199
Šiʿr (Kamal Abu Deeb) 199–214
Sicily (Alex Metcalfe) 215–219
Ṣifa (Kees Versteegh) 219–222
Sign Languages (Bernadet Hendriks and Ulrike Zeshan) 222–235
Ṣila (Kees Versteegh) 235–237
Sinai Arabic (Rudolf de Jong) 237–251

Slang (John Eisele) 251–259
Slavonic Languages (Václav Blažek) 259–267
Slips of the Tongue (Sabah M.Z. Safi) 267–272
Somali (Andrzej Zaborski) 272–275
Somalia (Mohamed Haji Mukhtar) 275–278
Songhay (Sergio Baldi) 279–282
Sound Symbolism (Kees Versteegh) 282–290
South Africa (Iris Hoedemaekers and Kees Versteegh) 290–295
South Arabian Loanwords (Martin R. Zammitt) 295–297
South Arabian, Modern (Antoine Lonnet) 297–300
South Semitic Languages (Anna Gr. Belova) 300–315
Specificity (Frederick M. Hoyt) 315–320
Speech Accommodation (Sonia Shiri) 320–328
Speech Acts (Daniela Rodica Firanescu) 328–334
Speech Errors (Stefan A. Frisch) 334–338
Stem (Adamantios I. Gafos) 338–344
Stress (René Kager) 344–353
Subject (Amira Agameya) 354–359
Subordination (Mark S. LeTourneau) 360–367
Substrate (Wouter Kusters) 367–375
Sudan (Al-Amin Abu-Manga) 375–381
Swahili (Sergio Baldi) 381–387
Syllable Structure (Maher Jesry) 387–389
Syncope (David Teeple) 389–391
Syntax (Elabbas Benmamoun) 391–402
Syria (Peter Behnstedt) 402–409

T

Taʿaddin (Zeinab Taha) 410–416
Taboo (Avihai Shivtiel) 416–421
Tafxīm (Muhammad Hasan Bakalla) 421–424
Tajik (John R. Perry) 424–425
Tajwīd (Kristina Nelson) 425–428
Taʿlīq (Valeriy Rybalkin) 428–430
Taltala (Jacques Grand'Henry) 430–431
Tamakkun (Janusz Danecki) 431–433
Tamil (Torsten Tschacher) 433–436
Tamyīz (Ramzi Baalbaki) 436–437
Tanāzuʿ (Wafaa Batran Wahba) 437–442
Tanwīn (Georgine Ayoub) 442–446
Taqdīr (Kees Versteegh) 446–449
Tatar (Suzanne Wertheim) 450–451
Telugu (V. Swarajya Lakshmi) 452–454
Tense (Uri Horesh) 454–458
Terminology (Xavier Lelubre and Joseph Dichy) 458–466
Terms of Address (Dilworth B. Parkinson) 466–471
Textlinguistics (Said Faiq) 471–476
Thai (Anthony Diller and Wilaiwan Khanittanan) 476–477
Thamudic (Ernst Axel Knauf) 477–483
Theme/Rheme (Nadia Anghelescu) 484–487
Theta Roles (Mark S. LeTourneau) 487–494
Topic and Comment (James Dickins) 494–501
Topicalization (Sven-Olov Dahlgren) 501–508

Toponyms (Stefan Wild) 508–515
Transcription (Philipp Reichmuth) 515–520
Transformational Grammar (Ferid Chekili) 520–528
Transitivity (James Dickins and Janet C.E. Watson) 528–535
Transitivity: Object (Usama Soltan) 535–542
Translation Literature (Uwe Vagelpohl) 542–548
Tripoli Arabic (Christophe Pereira) 548–556
Truncation (Lutz Edzard) 557–559
Ṯuluṯ (Adam Gacek) 560–562
Tunis Arabic (Maik Gibson) 563–571
Tunisia (Taïeb Baccouche) 571–577
Turkey (Joan Smith Kocamahhul) 577–583
Turkish (Mark Kirchner) 583–589
Turkish Loanwords (Stephan Procházka) 589–594

U

Urdu/Hindustani (Alain Désoulières) 595–603
Uyghur (Hans Nugteren) 603–608
Uzbek (Hans Nugteren) 608–612
Uzbekistan Arabic (Gerit Zimmermann) 612–623

V

Valency (Zouhair Maalej) 624–627
Variation (Enam Al-Wer) 627–636
Velarization (Stuart Davis) 636–638
Verb (Pierre Larcher) 638–645
Verb Phrase (Frederick M. Hoyt) 646–652
Verbal Clause (Frederick M. Hoyt) 653–659
Verbal Noun (Judith Rosenhouse) 659–665
Voice (Phonetics) (Chakir Zeroual) 665–669
Vowel Backing (Melissa Barkat) 669–674
Vowel Elision (Thami Benkirane) 674–676
Vowel Harmony (Philip J. Monahan) 676–677
Vowel Raising (Melissa Barkat) 678–682
Vowel Reduction (Thami Benkirane) 682–683

W

Waḍʿ al-Luġa (Bernard Weiss) 684–687
Wādī Ḥaḍramawt Arabic (Abdullah Hassan Al-Saqqaf) 687–699
Weak Verbs (Rainer Voigt) 699–708
West Sudanic Arabic (Jonathan Owens and Jidda Hassan) 708–718
WH-Movement (Lina Choueiri) 718–722
Wolof (Fallou Ngom) 723–725
Word Order (Sven-Olov Dahlgren) 725–736

X

Xabar (Yishai Peled) 737–740
X-Bar Syntax (Mark S. LeTourneau) 741–749

Y

Yemen (Martine Vanhove) 750–758
Yoruba (Isaac A. Ogunbiyi) 758–763
Youth Speech (Ahmad Taher Hassanein) 764–767

Index

A

ʾa- 2.22, 2.115, 2.133–135, 2.143, 2.389
/a/, elision of 4.404
/a/, raising of → *ʾimāla*
Aamer, Mahmoud 2.404
Aarts, Rian 2.75
Aartun, Kjell 1.196, 3.99, 4.455
ʿAbābda 2.2
Ababneh, Mohammad I. 4.480, 482
ʾAbādir ʿUmar ar-Riḍā 2.53
A-bar position 1.350
Abas, Husen 2.341
ʿAbbād ibn Sulaymān 4.286–287
abbāla 1.363
ʿAbbās, ʾIḥsān 2.305
Abbasid bookhand 3.110, 3.112
Abbasid Dynasty 1.402, 4.542, 4.589
Abbasid scripts 1.151
Abbasid Style, Early 2.597
Abbasid Style, New 3.110–112, 4.130–132
Abbasid Style, Old 2.597
Abbate, Giliberto 3.142
Abbéché 1.363–365, 1.367, 3.61
Abbéché Arabic 3.672, 3.674, 4.708, 4.716
Abbot Samson 1.96
Abbott, Nabia 2.373, 3.112, 3.307, 3.339, 3.514, 3.517–519, 4.560
Abboud, Peter F. 1.606, 2.424, 3.429, 4.168, 4.318, 4.350, 4.747
Abboud-Haggar, Soha 1.439–442, 1.613–622
abbreviation 1.1–5, 1.453, 3.43, 3.742
Abcarius, John 3.26
ʿAbd al-ʿĀl, ʿAbd al-Munʿim Sayyid 3.440
ʿAbd al-ʿAzīz, Muḥammad Ḥasan 1.435, 3.14
ʿAbd al-ʿAzīz ibn ʾAmīr Hāšim 2.52
ʿAbd al-Ġaniy, ʾAyman 4.61
ʿAbd al-Ḥamīd al-Kātib 4.74
ʿAbd al-Ḥaqq 4.561
ʿAbd al-Jabbār, al-Qāḍī 1.428, 2.304, 2.306, 2.579–580, 3.162–163
ʿAbd al-Karīm 1.361
ʿAbd al-Karīm, Ṣubḥī ʿAbd al-Ḥamīd Muḥammad 3.91, 3.94
Abd al-Kuri 4.297, 4.301
Abd al-Majīd, Sultan 4.98
ʿAbd al-Malik, Caliph 1.151, 1.401, 2.40, 2.199, 2.264, 2.599, 3.31, 3.516–517, 4.130
ʿAbd al-Malik ibn Manṣūr, Chamberlain 2.44
ʿAbd al-Malik ibn Marwān, Caliph 3.346
Abd al-Masih, George Mitri 3.41
ʿAbd al-Munʿim, Muḥammad 1.601
ʿAbd al-Munʿim, Ṣafāʿ 1.599
ʿAbd al-Muṭṭalib 1.361
ʿAbd al-Qays 2.212
ʿAbd al-Qays Arabic 3.94
ʿAbd an-Nāṣir, Jamāl 4.76
ʿAbd ar-Raḥīm 3.338
ʿAbd ar-Raḥmān al-Kātib al-Mālikī 3.309
ʿAbd ar-Raḥmān II, Caliph 2.46
ʿAbd ar-Rāziq Sālim 4.98
ʿAbd aṣ-Ṣabūr, Ṣalāḥ 4.214, 4.92
ʿAbd at-Tawwāb, Ramaḍān 1.614, 1.616, 1.617
ʿAbdalī, ʿAḥmad Faḍl al- 3.348
Abdalla, Fauzia 2.680, 2.689

Abdallah, A.G. 3.727
Abdallah, Ahmed 1.450
ʿAbdallāh Ibn Masʿūd 4.5
ʿAbdallāh ibn Muqla 3.339
Abdel Haleem, Muhammad 4.21–31
Abdelfattah, Nabil M.S. 3.193, 3.201–202
Abdel-Hafiz, Ahmed-Sokarno 3.713–717
Abdel-Jawad, Hassan Rashid E. 1.596, 2.501–503, 2.506, 3.12, 3.665, 3.681, 3.403, 4.268, 4.336–337, 4.629, 4.632
Abdelkarim, Abdourahman 1.336
Abd-El-Kuri → Abd al-Kuri
Abdel-Masih, Ernest T. 1.630, 1.668
Abdel-Nour, Jabbour 1.603
Abdelwahab, Hassan Hosni 4.572
Abdo, Daud A. 2.366, 4.350
ʿAbdū, ʾIbrāhīm 3.193
ʿAbdū, Muḥammad 3.193
Abdul, M.A. 4.759
Abdul Aleem 2.303
Abdul Aziz, Mohammad Hassan 2.331
Abdulaziz, Mohamed H. 1.631, 1.668, 3.135, 3.137
Abdul-Fattah, Hussein S. 3.681
Abdul-Ghany, Muhammad Kamal El-Deen [Eldine] 2.171, 2.173, 4.360–361, 4.366, 4.527, 4.746
Abdülhamid, Sultan 2.335, 2.498
Abdul-Karim, Kamal Wadih 2.396, 4.348–349
Abdullah, Abdul-Samad 4.182
Abdullah, Emir 2.498–500, 2.640
Abdul-Raof, Hussein 1.40–43, 1.339–347, 4.501
Abdurahiman, Muhammad 3.131
Abela, Gian Francesco 3.141
Abercrombie, David 3.725
Abha Arabic 4.127
ʾabhama 2.85
Abi Shaqra, Shawqi 4.213
ʿAbīda Arabic 3.606
ʾabjad 1.148–150, 1.153, 3.111
ʾabjada 3.21
ʾabjadiyya 2.471
ʾablaġu 2.582
ablaut 1.37, 1.119, 1.522, 1.525, 1.574, 1.624, 2.346, 2.350, 2.592, 3.297, 3.302
Abney, Steve 4.15, 4.743
ʾAbnūdī, ʿAbd ar-Raḥmān al- 1.599, 1.601
Abou, Selim 3.349
Abou Abdallah, Albert 2.456
Abou El Aazm, Abdel Ghani 3.27
Abou Sall, Ibrahima 3.314
Abou-Deia 1.362
Aboudi, Fahmi B.S. 3.432
Abraha Wende, Tekle 1.654–655
Abraham, Roy Clive 2.251, 4.759
Abramson, A. 4.669
ʿabrānī 2.471
ʿabriya 2.471
absolute accusative 3.455
absolute state suffix *-n* 3.568
abstract noun 1.18, 1.187, 2.424
abstraction 4.160
abstraction mechanism 3.684
ʾabū 2.578
Abu Abbas, Khaled Hassan 2.501
ʾAbū ʿAlī, Vizier 3.339
ʾAbū ʿAmr ibn al-ʿAlāʾ 2.442, 2.628, 3.90, 3.92, 4.5, 4.7, 4.121
ʾAbū Bakr, Caliph 2.47, 4.4
Abu Bakr Effendi 4.291
ʾAbū Bakr Ibn al-Fāḍil 4.293
ʾAbū Bakr ibn Muḥsin 2.328
ʾAbū Buṯayna 1.601
ʾAbū d-Dardāʾ 4.5
Abu Deeb, Kamal 1.425, 2.304–305, 4.199–214, 4.213
Abu Dhabi 3.78
Abu Dhabi Arabic 3.702
ʾAbū Ḏuʿayb al-Huḏalī 4.200
ʾAbū Dulaf al-Xazrajī 4.255–256
ʾAbū Farrāj 1.601
ʾAbū Ḥanīfa 1.529
ʾAbū Hāšim 3.163
ʾAbū Ḥayyān al-ʾAndalusī 1.641, 2.186, 2.226, 2.235, 2.291, 2.295, 2.354–355, 2.357, 2.433, 2.437, 2.453, 2.550, 3.72, 3.92, 3.94
ʾAbū Hurayra 3.14, 3.732
ʾAbū Jahl 4.79

Abu Jirj Arabic 2.5
ʾAbū Juday 2.598
ʾAbū l-ʿAlāʾ al-Maʿarrī 2.629, 2.637, 3.7, 4.89–90, 4.214
ʾAbū l-ʿAmayṯal 3.34, 3.321
ʾAbū l-ʾAswad ad-Duʾalī 1.152, 1.400, 2.86, 2.177, 2.232, 2.401, 2.425, 2.632
ʾAbū l-ʾAxzar 2.85
ʾAbū l-Faraj 2.598
ʾAbū l-Faraj, Muḥammad ʾAḥmad 1.435
ʾAbū l-Fatḥ ʿUmar ibn al-Malik al Muẓaffar Yūsuf ibn ʿAlī ibn Rasūl 2.258
ʾAbū l-Hindī 4.205, 4.208
Abu l-Xaṣīb Arabic 1.272
ʾAbū Mādī, ʾIliyyā 4.91
Abu Manga, Al-Amin 1.525, 2.250–256, 2.274, 4.375–381
ʾAbū Mūsā, Muḥammad M. 2.580
ʾAbū n-Najm 2.84, 4.34, 4.36
ʾAbū Nuwās 3.540, 4.32, 4.88, 4.90, 4.205, 4.209–210
ʾAbū Qurra, Theodore 4.542
Abu Rabia, Salim 2.97, 2.464, 2.660, 3.79
ʾAbū Šādī, ʾAḥmad Zakī 4.91
ʾAbū Šāma 3.93
ʾAbū s-Suʿūd, ʿAbbās 3.440
ʾAbū Tammām 2.446, 3.539, 3.717, 4.90, 4.104, 4.202, 4.205–206, 4.208–209, 4.213
Abu Tīj Arabic 2.5
ʾAbū ṭ-Ṭayyib al-Luġawī 1.313, 1.627, 2.86, 2.20–281, 2.401, 3.34–35
ʾAbū ʿUbayd al-Qāsim ibn Sallām al-Harawī 1.627, 3.33–36, 3.89, 3.321
ʾAbū ʿUbayda, Maʿmar ibn al-Muṯannā 2.301, 2.579, 3.36, 3.117, 3.689, 4.9, 4.23, 4.447
ʾAbū ʿUṯmān ad-Dimašqī 4.542
ʾAbū Xamīra 2.598
ʾAbū Zayd 4.33–34
ʾAbū Zuʿayb al-ʿAbšamī 4.287
Abu-Absi, Samir 2.696–699, 2.701–702
Abubakar, Ali 3.374
Abubakre, Razaq Deremi 3.376
Abu-Chacra, Faruk 4.148
Abudacnus, Josephus 1.169, 1.222–231, 1.269–274, 1.444, 1.610
Abu-Haidar, Farida 2.415, 4.653–654, 4.656
Abu-Kamal Arabic 4.678, 4.680
Abul-Fadl, Fahmi 1.584, 3.402, 3.404, 4.238
Abu-Lughod, Ibrahim 1.528, 3.200
Abu-Mansour, Mahasen Hasan 2.36, 2.38, 3.179–187, 3.494, 4.683
Abu-Melhim, Abdel-Rahman Husni 4.323, 4.325–326
Abu-Risha, Zuleikha 2.643
Abu-Salim, Issam M. 2.37–38, 4.268, 4.336–337, 4.347, 4.676
Abyan 4.750–751
Abyan Arabic 4.751, 4.753–756
Abyssinia 1.654, 4.275
Abyssinian 2.258, 4.295
Académie Française 2.634, 2.636–637
academies, language 2.189, 2.376, 2.634–642
Academy, Arabic Language 1.629
Academy, Bangla 1.288
Academy, Berber 3.350
Academy, Cairo 1.453–454
Academy, Hebrew language 2.462
Academy, Kurdish 3.351
accent 2.113–114
accent, bitonal 2.400
accent, nuclear 2.396, 2.398
accent, phrase 2.396
accent, pitch 2.395–399
accent, prenuclear 2.396
accent, qualitative 4.103
accent, tonic 3.635
acceptability 1.429–432, 1.453, 4.472, 4.475
accident 2.432
accommodation → speech accommodation
Accommodation Theory 4.628
accomplishment 1.52, 1.197–200, 3.253
accusation 2.363
accusative 1.347–350, 1.352, 1.400, 1.559–560, 1.562–563, 1.580, 2.146–147, 2.626, 3.101, 3.392

accusative, absolute 3.455
accusative, cognate 1.436, 3.102, 3.256, 3.555
accusative, indefinite 1.408
accusative, specificational 3.684
Aceh 1.5–6, 1.8–9, 2.334–335, 2.338
Acehnese 1.5–12, 1.486
Achaemenid Empire 1.178
Acheron 3.350
achievement 1.52, 1.197–198, 3.253
Achit, Abdelmadjid 1.458
Acholi 4.386
Ackema, Peter 3.231–232
acoustic analysis 1.458
acquisition → first language acquisition, second language acquisition
acquisition, natural order of 4.140
acquisition, partial 2.711
acquisition, phoneme 1.375–376
acquisition, phonological 2.98
acquisition, universals of 4.370
acquisition of auxiliaries 1.377
acquisition of bilingual lexicon 1.371
acquisition of broken plurals 2.101
acquisition of causative 1.377, 2.102
acquisition of consonants 2.98
acquisition of derivation 2.102
acquisition of diphthongs 2.98
acquisition of inflection 2.101
acquisition of interrogatives 2.100
acquisition of lexicon 2.99
acquisition of middle voice 2.102
acquisition of morpheme order 4.140
acquisition of negation 2.100, 4.140–141
acquisition of passive 1.377
acquisition of past tense in English 2.101
acquisition of patterns 1.377
acquisition of plural 2.101
acquisition of questions 2.100
acquisition of reciprocals 2.102
acquisition of spatial terms 1.377
acquisition of temporal terms 1.377
acquisition of vowels 2.98
Acre 2.439
acrolect 2.266
acronym 1.1, 5, 1.175, 1.452, 3.426, 4.466, 4.557–558
act 3.252–253
act, evaluative 2.80
actant 1.622–623
actant, first 1.624–626
actant, second 1.625–626
actant, third 1.625
action, change-of-state 3.253
action, non-sequenced 2.208–209
action error 4.271
action nominalization 3.388
activation, spreading 4.270
active 1.625
activity 1.197–198, 1.200, 3.253
activity verb 1.52
actualization 4.40–41
ad hoc formation 4.34–35
ʾadā 2.237, 2.306, 2.362, 3.549, 4.26
ʾadab 1.403, 2.86, 2.446, 2.632, 3.658, 4.104
ʾaḍāfa 2.294
Adal 2.52
Adam 2.182
ʾĀdam, ʾAḥmad ibn 2.53
Adamawa State 2.137, 2.139, 4.709
Adamovic, Milan 3.501
Adams, William Y. 3.435
ʾaḏān 1.528
Adana 1.388, 4.579–580
Adana Arabic 1.391, 1.394–395, 2.313
adaptation 2.31–32, 2.35, 2.128–129, 2.131, 2.200
adaptation, morphological 2.139–140, 2.285
adaptation, phonological 2.284, 2.473
adaptation, semantic 2.474
adaptation of loanwords 2.123–124, 2.266
ʾadāt at-taḥḍīḍ 3.573
ʾaḍdād → *ḍidd*
Addis Ababa Accord 2.700
additive 1.431, 1.472–473
address, bipolar terms of 2.616, 3.486
address, terms of 1.550, 2.586, 3.658, 4.466–471
Adem, S.A. 4.670
Aden 1.260, 1.654, 2.273–274
Aden Arabic 1.23, 1.260, 1.501, 2.388, 4.690, 4.750–751, 4.754, 4.756–757
ʾadġama 2.298

Adger, David 4.392
Adiabene 1.126
ʿAdiyy ibn Zayd al-ʿIbādī 1.150, 4.202
adjacency 1.653, 4.719
adjacency of sounds 3.206
adjacency pair 2.202, 2.204
adjacent autosegment 4.95
adjective 1.12–21, 1.68, 1.477–481, 1.561, 1.567, 1.572, 1.582, 1.626, 2.171, 2.431–432, 3.244, 3.422, 3.424, 3.443, 4.219–220, 4.746
adjective, agentive 1.346
adjective, attributive 4.221
adjective, epicene 1.14
adjective, experiencer 1.346
adjective, intensive 1.16
adjective, postnominal 3.431
adjective, predicative 112–13
adjective, relational 1.80
adjective, relative 1.14, 1.17, 1.452, 3.153, 3.220, 4.691
adjective, stative 1.346
adjective, verbal 1.197, 3.424
adjective agreement 1.266
adjective in Swahili 1.664
adjective noun 1.13–14
adjective of color and defect 1.14, 1.17, 1.20, 1.107, 1.251, 1.304, 1.390, 1.393, 1.550–551, 1.576, 1.607, 2.14, 2.244, 2.485, 2.616, 3.279, 3.304, 3.532, 4.307, 4.310, 4.554, 4.691
adjective of possibility 3.556
adjective ordering 3.232
adjective phrase 1.12–16, 1.20, 3.431–432
Adji, Haoua 1.334
adjunct 1.481, 1.486, 1.644, 3.107–108, 4.392, 4.535, 4.746–747
adjunct, conjunctive 1.475
adjunct, modal 1.475
adjunct, restrictive 4.531
Adjunct Island Constraint 4.80
adjunction 2.133, 4.530–531
Adler, Jakob George Christian 1.151, 2.597
administration 1.401–402, 1.404
admiration 2.80
ʿAdnān 1.617, 3.345
Adnan, George 4.742
adnominal complement 1.312
Adonis [ʾAdūnīs] [ʿAlī ʾAḥmad Saʿīd] 4.77, 4.90, 4.204, 4.211, 4.213–214
adoption 2.32
adposition → apposition
Adrar 3.169–171
adstratal influence 1.29, 1.34, 1.54, 1.102, 1.490, 1.611, 2.262, 3.189, 4.368–369, 4.372, 4.374
adstratal influence of African languages 1.366
adstratal influence of Berber 1.263, 1.571
adstratal influence of French in English 4.374
adstratal influence of Greek 1.537, 1.539
adstratal influence of Hebrew 2.666
adstratal influence of Kurdish in Anatolian Arabic 4.727
adstratal influence of Romance 2.288
adstratal influence of Romance in Maltese 3.159
adstratal influence of Tajik in Uzbekistan Arabic 4.621–622
adstratal influence of Turkish 1.396
adstratal influence of Turkish in Anatolian Arabic 4.727
adstratal influence of Uzbek in Uzbekistan Arabic 4.616–617, 4.621–622, 4.727
Adulis 1.654
ʾAdūnīs → Adonis
Advanced Tongue Root 2.520
adverb 1.21–25, 1.240, 1.248, 1.279, 1.303, 1.327, 1.392, 1.404, 1.408, 1.476, 1.549, 1.562, 1.567
adverb, Bantu-Arabic 1.664
adverb, demonstrative 1.91, 1.567, 1.569–573
adverb, grade 1.331
adverb, interrogative 1.92, 1.116, 1.279, 1.303, 1.327, 1.491, 2.243, 2.390, 3.278
adverb, locative 1.569
adverb, manner 1.22
adverb of direction 1.22
adverbial 1.21, 1.195, 1.344, 1.371, 1.432, 2.209, 2.356, 4.531, 4.539–540

adverbial, fronted 1.476
adverbial, preposed 4.499–501
adverbial, pre-topical 1.475
adverbial, sentential 1.476, 3.392
adverbial, speech-act 1.476
adverbial clause 1.217, 2.525
adverbial complement 3.97
adverbial of place 3.101, 3.103, 3.107, 3.224, 4.738
adverbial of time 1.199, 3.101, 3.103, 3.107, 3.166, 4.738
adverbialization 2.192
adversative 1.202, 1.431, 1.465, 1.472
advertising 1.600, 2.30, 3.540
Adygey 3.350
Adyghe 2.506
Aescoly, Aaron Zeev 4.95
Afade 1.334
Afāḏ̣le Arabic 4.407
ʾafʿāl → fiʿl
ʾafʿāl al-qulūb → fiʿl al-qalb
Afambo 2.272–273
Afar 1.35, 1.38, 1.451, 1.655, 2.51, 2.53–55, 2.269–274, 3.62
Afar-Saho 1.38–39
affectedness 4.625
affection 2.79
affective suffix 2.80
afficient 2.348
affix 1.209, 1.411, 1.573–574, 2.346–347, 2.353, 2.449, 3.297
Affix Grammar 1.458
Affix Hopping 4.522
affix, derivational 1.574
affix, inflectional 1.574
affixal morpheme 3.297
affixation 1.576, 2.192
affixation, agreement 2.366
affixation, clitic 2.366
affricate 1.26–27, 1.60–61, 1.260, 1.299, 1.571, 2.494–495, 2.507, 3.189, 3.598
affrication 1.26–29, 1.53, 1.103, 1.242, 1.260, 1.291, 1.295, 1.585, 1.606, 1.612, 2.212, 2.416, 2.610, 3.275, 3.290, 3.293, 3.361, 3.403–404, 3.479, 3.587, 3.606, 4.128–129
ʾAfġānī, Jamāl ad-Dīn al- 3.193, 3.346
ʾAfġānī, Saʿīd al- 3.344, 3.379
Afghanistan 2.61, 2.219, 3.312, 4.612
Afghanistan Arabic 1.28–35, 1.537, 1.583, 3.84, 4.617
ʿAfīfī, Fawzī Sālim 4.98
ʿĀfir, al-Maskīnī al- 1.603
ʾafiyās 4.105
Af-Jiddu 4.275
Aflaq, Michel 3.345, 4.76
Af-Maay 4.275, 4.277
Af-Mahaa 4.275
Afnan, Soheil M. 3.113, 3.176, 4.542
Afrasian languages 1.35
Africa, sub-Saharan 2.588
Afrikaans 2.669, 4.290–295
Afro-Asiatic languages 1.35–40, 1.315, 1.334, 1.563, 1.643, 2.67, 2.250, 2.621, 3.52, 3.64, 3.67, 3.69, 3.204, 3.420, 3.591, 3.593, 4.96
ʾafṣaḥ 2.86–88, 3.93
ʾafṣaḥa 2.84, 2.543
Afshar Turks 2.572
afterthought 2.148, 4.726
ag 2.679
Agadir 2.715
Agameya, Amira 3.558–564, 4.354–359
Agatharchides 1.131
ʿAgāylah 4.238
ʿAgāylah Arabic 3.401
age as a variable 4.630
age grading 4.630
agency 1.649, 3.225, 4.412, 4.625
agent 1.40–43, 1.67, 1.344–347, 1.436, 1.556, 1.563, 1.622–623, 2.95, 2.145–146, 2.187, 2.291, 2.315, 2.348, 2.357, 2.429–431, 2.435–437, 2.537, 2.539, 2.550, 3.100–105, 3.202, 3.225–226, 4.354, 4.487, 4.489, 4.490, 4.492, 4.501, 4.529, 4.535, 4.538
agent → *fāʿil*
agent, implicit 3.226
agent, initial 4.653
agent, logical 4.626
agent, pronominalized 4.357
agent construction 1.623
agent extension 1.624–625

agent marker 3.99
agent noun 2.430, 2.432, 3.98
agent orientation 4.727
agent phrase 1.624, 3.560–562
agent pronoun 3.98–99
agenthood 1.41
agentive phrase 3.555–556
agentive verb 1.41–42
agentivity 1.41–42, 2.367
Agesthialingom, S. 4.433
AGFL 1.460
agglutination 1.538, 2.221, 2.346
agglutination of article 4.384
Agheysisi, Rebecca 2.652
Aghlabid Dynasty 4.215, 4.218, 4.572
Agiùs, Dionisius A. 1.668, 1.670, 2.158, 3.142, 3.144, 3.146, 4.151, 4.217–4.218
Agiùs De Soldanis, Giovan Pietro Francesco 3.141, 3.146
ʾAġlab al-ʿIjlī, al- 4.33
agnomen 2.578
Agostini, Francesco 4.272–274
agrammatism 2.691
agraphia 2.688
agree 3.229
agreement 1.12–16, 1.18–20, 1.43–48, 1.110, 1.185–186, 1.230, 268, 1.308, 1.348, 1.396, 1.434, 1.479, 1.486, 1.554, 1.582, 2.99, 2.157, 2.272, 2.429–431, 2.491, 2.597, 2.631, 2.643–645, 3.67, 3.69, 3.84–85, 3.221, 3.229–232, 3.258, 3.285, 3.430, 3.439, 3.446, 3.537, 3.588, 3.709, 3.714, 4.187, 4.392, 4.653, 4.745
agreement, acquisition of 1.377, 4.142–145
agreement, conjunct 4.397
agreement, default 4.355
agreement, deflected 1.14–15, 2.162–163, 2.378, 2.645, 4.356–357
agreement, dual 4.356
agreement, equivocal 2.162
agreement, first conjunct 2.174, 4.193–194, 4.649
agreement, full 3.382, 4.318–319
agreement, gender 2.161, 3.231
agreement, mixed 4.194
agreement, natural 1.616
agreement, neutralized 4.318–319
agreement, number 1.491, 2.102, 3.231–232, 4.396
agreement, partial 4.397, 4.745
agreement, 'phonological' 3.231
agreement, reduced 4.357
agreement, reverse 1.378
agreement, rich 3.706
agreement, specifier/head 2.174, 3.231
agreement, split 3.231
agreement, strict 2.162
agreement, subject 3.298
agreement, subject/verb 1.184, 1.349, 1.649, 3.231, 4.193, 4.355, 4.357, 4.396
agreement, syntactic 3.232
agreement, verbal 1.622–623
agreement asymmetry 1.45–46, 1.182
agreement clitic 4.362
agreement error 2.689
agreement feature 1.182–183
agreement in code-switching 1.417
agreement in Northwest Semitic 3.419
agreement marker 1.186, 1.411, 1.488, 2.149
agreement marking 4.653, 3.383
agreement mismatch 2.680
agreement of the dual 3.737–738
agreement relation 1.349
agreement under government 3.231
Agri 2.607
Agrigento 3.142
AgrP → phrase, Agreement
Aguadé, Jordi 1.289–293, 1.295, 1.595–596, 3.287–297, 3.280–281, 4.71
Aguilar 1.99
-ah 2.80
ʾaḥad 4.13
Ahaywāt Arabic 3.401, 3.403, 4.238
ʾahdāb 4.133
ʾahl al-qurra 1.129
ʾAhl aš-Šimāl 3.401
Ahl Giblī 3.401
ʾAhl Ḥabät 3.172
ʾAḥmad, ʿAbd as-Samīʿ Muḥammad 3.42

Ahmad, Azimuddin 2.257, 2.260
ʾAḥmad, Jamīl 2.328
Ahmad, Kusay 2.229
ʾAḥmad, Muḥammad ʿAbd al-Qādir 1.626
Ahmad, Syed Barakat 3.343
Ahmad, Zubaid 3.508
ʾAḥmad ibn Mīr Munši al-Ḥusaynī al-Qummī, Qāḍī 3.337
ʾAḥmadī 2.609
Aḥmadī Gīwī 2.410
Ahmadiyyah 2.337
ʾAḥmar, ʿAlī ibn al-Ḥasan al- 4.119
Ahmedabad 2.329
Ahmed-i Khani 2.604
ahong 1.380
ʾAhrām, al- 3.200–201
Ahrens, Margaret 1.341
ʾahú 1.571
ʾAḥwaṣ, al- 4.202
Ahwaz 2.571–572, 2.576, 3.370
ʿāʾid 1.49, 1.106, 2.132, 3.245, 3.261, 4.102, 4.544, 4.720, 4.738
Ain Temouchent 1.56
ʿĀʾiša 3.91
Aisenstadt, Esther 1.437–438
Aissati, Abderrahman El 1.293–299
ʿajʿaja 1.614, 2.494
ʿajam 1.99, 1.131, 2.53–55, 2.60, 2.84–85, 2.88, 3.345–546
ʿajamī 1.131, 1.135–136, 1.194, 2.85, 4.183
Ajar Fazzan 2.552
ʿAjārma Arabic 1.626, 3.405
Ajayi, J.F. Ade 3.374
ʿAjjāj, al- 3.650, 4.287
Ajloun Arabic 2.506
ʿAjmān Arabic 1.590, 3.326
ʾajnabī 1.49, 2.357
ʾĀjurrūmiyya 1.168
ʿajuz 4.87, 4.89
ʾajwaf 4.121, 4.645
Akalay, Lotfi 3.294
ʾakalūnī l-barāġīṯ 1.616, 2.430
Akbar 2.327
Åkesson, Joyce 1.574, 1.641, 2.180, 2.448, 4.118–122, 4.640
Akhenaten, Pharao 3.409
Akhetaten 3.409
Akhisari, Hasan Kafi 3.503
Akiner, Shirin 2.558, 4.612
Akkadian 1.38–39, 85, 1.122, 1.126, 1.132, 1.178, 1.203, 1.489, 1.560, 1.563–564, 1.624, 2.66–67, 2.212, 2.387, 3.58, 3.227, 3.409–410, 3.413, 3.417–418, 3.422, 3.424, 3.453, 3.494, 3.553, 3.592–593, 3.738, 4.1, 4.96, 4.104, 4.170–171, 4.174, 4.176–178, 4.296, 4.302, 4.313, 4.444, 4.456, 4.509
Akkadian, El-Amarna 2.22–23, 2.67, 3.41
Akkadian, Kamid El-Loz 2.23
Akkadian, Old 3.424
Akkadian, perfect in 4.639, 4.642
Akkadian, precative in 4.557
Akkadian, preterite in 4.557
Akkadian, Taanach 2.23
Akkadian, vetitive in 4.557
Akl, Saïd 1.603, 3349
Akmajian, Adrian 4.522
Akrawi, Matta 2.104
Aksum 2.57
ʾAkṯam, ibn Sayfī al- 3.71
ʾakṯar 1.192
Aktionsart 1.50–53, 1.195–197, 4.196
ʾAkwaʿ, Muḥammad ʿAlī...al-Ḥiwalī al- 2.258–259, 4.750
ʾĀl Bu Šāmis 2.213
Al Busaid 1.448
Al Gindi, Anwar 3.193
ʾĀl Murrah 4.125, 3.326
ʾĀl Murrah Arabic 3.326–327, 4.127, 4.124
ʾĀl Ṣubāḥ 2.572
ʾĀl Wahība 2.213
ʾĀl Wahība Arabic 3.478–479, 3.483, 3.489
ʾĀl Yāsīn, Muḥammad ʿAbd al-Qādir 1.627
ʾĀl Yāsīn, Muḥammad Ḥasan 1.626
ʾĀl Zubayr Arabic 1.272
Alaattin 1.115
Al-Aboudi, Fahmi Basil Shukri 4.742
Alafin Obalokun 3.375
Alagoa, E.J. 3.374
Al-Ajmi, Hassan 3.28
Al-Akeel, Abdulrahman I. 2.99
Al-Ali, Monira Dawod 4.126–127

Al-Ali, Salih 4.99
ʿ*alam* 2.404, 2.677, 3.717, 3.719
Alam, Muzaffar 2.325
Alam, Yahya Mir 2.448
ʿ*alāma* 1.556, 2.82, 2.405, 4.442, 4.445
ʿ*alāma li-l-ʾamkan* 4.444
ʿ*alāmat al-ʾiʿrāb* 1.343
ʿ*alāmat al-istifhām* 3.741
ʿ*alāmat al-istišhād* 3.741
ʿ*alāmat al-muḍmar* 4.221
ʿ*alāmat al-waqf* 3.740
ʿ*alāmat at-taʿajjub* 3.741
ʿ*alāmat at-taḍbīb* 1.4
ʿ*alāmat at-tarqīm* 3.740
Alami, Yusef Amin el- 3.295
Al-Amri, Abd al-Salam 4.107
Al-Ani, Salman H. 2.229–231, 2.365, 2.396–397, 2.642, 3.593–603, 3.610–611, 3.726, 4.139–140, 4.344, 4.387–388, 4.669, 4.671
Alaoui, Ahmed 2.292
Alatas, Ali 2.337
Al-Attas, Syed Muhammad Naquib 1.8, 2.340, 2.344
ʿAlawī, al-Muẓaffar al- 3.649, 3.651
ʿAlawī, Sayyid ʾAbū Bakr al- 3.508
Alawi Arabic 1.391, 1.395–396
ʿAlāwīn Arabic 3.401
Alawis 1.111, 1.388, 2.59
Alawne 1.362
ʿAlāylī, Sheikh ʿAbdallah al- 3.345
Al-Azraqi, Munira Ali 2.555–557
Al-Bamerni, A. 4.669
Albanian 1.133, 4.371
Albanian in Greece 2.684
Al-Batal, Mahmoud 1.431–432, 1.471–473, 1.648
Al-Bawab, Marwan 2.165, 2.448
Al-Bazi, Mohammed 4.62, 4.64
Albū Kmāl Arabic 4.407
Al-Buainain, Haifa 4.141
Alcalá, Pedro de 1.99–100, 1.168–169, 1.171, 2.289, 3.23
alchemy 1.498, 1.500
Alcock, Antony 2.439
Alderete, John 3.494, 627
Aldrete, Bernardo de 1.96–97
Alegat Arabic 3.401
Alem, Shadia 4.204
Al-Enazi, Mohamed Husein 1.372
Aleppo 1.172, 2.217–218, 2.220, 4.402, 4.408
Aleppo Arabic 1.445, 1.537, 1.608, 1.625, 2.313, 2.494, 2.534, 2.608, 3.450, 3.607, 4.64, 4.407–408, 4.664, 4.681
Aleppo Arabic, Jewish 1.111
Al-Essa, Aziza 3.606, 4.628, 4.632
alethic 3.234–235
Alevis 4.580
Alexander the Great 3.6, 3.645
Alexandretta Arabic 2.313
Alexandria 1.589, 2.1, 2.470, 4.157
Alexandria Arabic 1.439, 1.441, 1.589, 1.604, 2.1–3, 2.11, 2.495, 4.663
alexia 2.690
Alexopoulou, Theodora 3.386–387, 3.684, 3.687
ʾ*Alf layla wa-layla* 1.598, 3.219
ʾ*alġā* 2.308
Algeria 1.23, 1.26, 1.53–58, 1.534, 2.363, 2.704, 2.707–710, 2.712–715, 3.53, 3.60, 3.316–317
Algeria, Arabic in 1.260, 1.502, 1.590, 1.606, 1.642, 3.672
Algerian Arabic 1.53, 1.56–57, 1.83–84, 1.418, 1.492, 1.502, 1.584, 2.74, 2.374, 2.494, 62.57, 2.692, 2.697, 3.545–546, 4.67, 4.306, 4.662, 4.678
Algerian Arabic, Jewish 1.57–58
Algiers 1.23, 1.53–55, 1.58, 2.715
Algiers Arabic 1.58–66, 1.440, 1.443–444, 1.596, 1.610, 2.532, 2.534, 3.587, 4.569
Algiers Arabic, Jewish 1.59
algorithm, finite-state 1.514
algorithm, learning 1.515
Al-Haj, Majed 2.105, 2.441
Al-Haq, Fawwaz Al-Abed 2.656, 2.661, 4.647, 4.654–655, 4.657
Al-Harbi, Awwad Ahmad al-Ahmadi 1.10
Alharbi, Lutfi 2.397, 3.727
Alhawary, Mohammad T. 2.674–681, 4.138–146
Al-Hazmi, A. 4.127
Al-Hroot, A. Khalaf 4.525
Al-Muhtaseb, Husni 1.458

Ali, Abdul Sahib Mehdi 1.451–455, 3.426, 4.167, 4.464
ʿAlī, Jawād 2.640
Ali, Latif H. 4.669–670
ʿAlī, Muḥammad → Muḥammad ʿAlī
ʿAlī, Muṣṭafā 3.502
ʿAlī, Nabīl 1.460
Ali Awad, Fahade 1.654–655
ʿAlī ibn ʾAbī Ṭālib, Caliph 1.362, 2.47, 2.425, 2.598, 3.14, 3.344.5, 4.134
ʿAlī ibn Maʿṣūm, Sayyid 2.328
ʿAlī ibn Šaḏān 2.601
ʿAlī ibn Ṭalḥa ar-Rāzī al-Baqašlan 3.309
Ali well Manu 3.173
alienable 1.83, 1.110, 3.671
alienable possession 2.594, 3.537
ʾalif 1.151–153, 1.403, 2.164, 2.233–234, 2.311, 3.220, 4.133
ʾalif, accusative 1.383, 1.386, 1.400
ʾalif, dagger 4.518
ʾalif, otiose 1.152
ʾalif, prosthetic 1.385, 1.638
ʾalif al-ʾimāla 2.233
ʾalif al-mumāla ʾilā jihat al-yamīn, al- 4.561
ʾalif at-taʾsīs 4.89
ʾalif fāṣila 1.384
ʾalif lām 1.487, 2.425–426
ʾalif mamdūda 1.384, 2.156, 3.426
ʾalif maqṣūra 1.384, 1.401, 1.640, 2.156, 2.527, 3.242, 3.426, 4.518
Aligarh 2.326, 2.329
alignment 3.492
Alignment Constraint 2.36
Alilat 1.126–127
Al-Ilori, Adam Abdallah 3.373, 3.375, 4.759–760
ʿālim 1.451
Ali-Sabieh 1.655
Al-Issa, Ahmad 3.681
ʿAlj, aṭ-Ṭayyib al- 1.603
Al-Jaġbūb 3.56
aljamiado 1.100
Al-Jazeera 1.533, 2.382, 3.19, 3.198, 3.203, 4.226
Aljenaie, Khawla 1.333–343, 2.102
Al-Jubouri, Adnan 1.648
Al-Kahtany, Abdallah Hady 2.657–658
ʾalkan 2.677
Al-Kasimi, Ali M. 2.332, 3.28
Al-Kharrat, Edouar → Kharrat, Edouard
Al-Khatani, Ali 4.632
Al-Khatib, Ahmed Sh. 3.26
Al-Khatib, Hayat 1.372
Al-Khatib, Mahmoud A. 2.500–501, 2.506, 3.681, 4.467, 4.629
Al-Khuli, Muhammad A. 1.457
Al-Kurdi, Wasim 2.111
allaḏī 1.309, 2.479
Allam, Jehan 4.252–254
ʿAllāmī, ʾAbū l-Faḍl 2.598
Allan, Keith 1.657, 1.659, 2.69
Allāt 1.127
Allauddin Khilji 4.452
allegory 2.441, 2.445, 3.121
allegro form 2.591
Allen, John W.T. 1.660–661, 1.663
Allen, Roger 4.150
Allen, W. Sidney 4.345
Alliance Israélite Universelle 2.64
alliteration 2.496, 3.538
allograph 1.150
allomorph 4.40, 4.115
allomorphy 1.78, 3.440
allophonic variation 3.605
allusiveness 2.305
Al-Madi, Muneeb 2.498
Almasude, Amar 2.63
Almeida, Monica 3.2
Almohad Dynasty 1.97–98, 1.100, 1.589, 2.62–63, 2.532
Almoravid Dynasty [al-Murābiṭūn] 1.97–98, 1.100, 2.62–63, 3.169–170, 4.180, 4.723
Al-Mozainy, Hamza Qublan 3.205, 4.344, 4.351, 4.389
Al-Muhannadi, Muneera 4.632
Al-Musawi, Muhsin 4.475
Al-Muslat, Zaid Abdullah 4.224
Al-Najjar, Balkees 4.40, 4.162
Al-Nassir, Abdulmunim Abdalamir 1.544, 1.641, 2.228–229, 2.232–233, 2.237–238, 2.298–300, 3.73, 3.525, 3.602, 3.681, 4.2, 4.423, 4.667
AlNeami, Ahmed 1.460
Alon, Ilai 4.542
Alosh, Mahdi 1.669
Alparslan, Ali 3.338

alphabet 1.207
alphabet, Arabic 1.133–165, 3.31, 4.130
alphabet, finger-spelling 4.225
alphabet, order of 1.148, 150, 1.487
alphabet, South Semitic 4.480
alphabet, universal 4.516
alphabetical order 3.21
Alphonse VI, King 1.97
ALPNET 1.460
Al-Qahtani, Saad 2.702
ʿAlqama 3.646
Al-Quoz, Muna 4.631
Al-Rawi, Sabah 1.438
Al-Saqqaf, Abdullah Hassan 4.687–699
Al-Sasi, Omar 3.179, 4.126
Alsayed, Adnan 2.173–174
Al-Selwi, Ibrahim 2.257, 4.750
Al-Shatti 1.339–340, 1.342
Al-Shehri, Abdullah 4.632
Al-Shorafat, Mohammad O. 4.746
Al-Sughayer, Khalil Ibrahim 2.501
Al-Sulaiti, Latifa 1.513
ʾalṯaġ 2.676–677
Al-Tahir Mikky, Ahmad 2.107, 2.109
Altaic languages 2.67, 3.204
Al-Tajir, Mahdi A. 1.271, 1.446
Al-Tamari, Emad Ahmad 4.398, 4.649, 4.653, 4.655–656
Al-Tamimi, Fida 4.629
Altay 2.557
Al-Tayyan, Mohamad Hassan 2.448
Altena, Nelleke 2.97, 3.315
alternative 1.431, 1.465, 1.472
Al-Thebyan, Qusai 4.288
Altheim, Franz 4.402
although 1.620
Altinözü 1.114–115
Altoma, Salih J. 2.104–106, 2.111
Alur 2.597
ʾAlūsī, Maḥmūd Šukrī al- 2.637
ʾAlūsī, Sālim al- 2.639
Alvarus of Cordoba 1.96
alveolar ridge 3.599
ʿAlwa 1.361
Al-Wer, Enam 1.596, 2.482, 2.501, 2.505–517, 3.11–12, 3.603–607, 4.627–636, 4.681
ʾam 1.281, 1.468, 2.490, 3.237
ʾa-mā 1.305
Ama, Morris 2.518
Amadasi Guzzo, Maria Giulia 3.409
Amado, Jorge 3.4
Amajur 2.600
ʿamal 1.67–74, 1.347, 1.558, 2.44, 2.92, 2.169, 2.181, 2.184, 2.187, 2.290, 2.307, 2.355, 2.403, 2.542, 2.625–626, 3.101, 3.108, 3.354, 4.221, 4.432, 4.438, 4.442, 4.446, 4.486, 4.739
amalgamation, V-N 1.352
Amara, Muhammad Hasan 1.419, 2.104–105, 2.437–441, 2.464–467, 2.572, 2.661, 2.699
Amari, Michele 4.217
Amarna, El- 3.409
ʾAmāsī, Ḥamd Allāh al- 3.341, 4.131
ʿAmāyra, ʾIsmāʿīl ʾAḥmad 4.26
Amayreh, Ismail 2.479
Amayreh, Mousa M. 1.376, 2.97–99, 2.679, 2.688
Amazigh 2.62–63, 2.74, 2.659, 2.704, 2.707–716, 3.312, 3.317–319, 3.350
Amazigh in education 2.715
Amazigh in Morocco 2.707
Amazigh literature 2.709
Amazighity 3.350
ambiguity 1.456, 1.627–628, 2.368, 3.678, 4.39, 4.200, 4.208
ambivalence 1.627–628
Ambouli 1.655
Ambros, Arne A. 1.562, 2.30, 3.700–701, 4.656
Amenophis IV, Pharao 3.409
Amer, Attia 4.13
Amer, Walid Mohammad Abdelghaffar 4.487, 4.492–493
American Army Language School 4.148
American Council for Teaching Foreign Languages (ACTFL) 4.150
American University 1.175
Amerindian languages 4.371
Amhara 2.67
Amharic 1.203, 1.493, 1.638–639, 2.53–55, 3.494, 3.592, 4.301–302, 4.305–306, 4.309–310, 4.312
ʾĀmidī, Sayf ad-Dīn ʾAbū l-Ḥasan ibn ʾAbī ʿAlī al- 1.428, 2.359, 3.324, 4.209, 4.213
Amiji, Hatim M. 3.135, 3.137, 3.139
ʿāmil 1.67, 1.309, 1.471, 1.556, 2.169–170, 2.237, 2.290, 2.294, 2.403,

2.405, 2.430, 2.478, 2.536, 2.538, 2.542, 2.549–550, 2.625–626, 3.178, 3.234, 3.353, 4.447, 4.738–740
ʿāmil lafḏī 2.291
ʿāmil maʿnawī 2.290–291, 4.486
ʿamila 2.44
ʾAmīn, ʿAlī 3.193
ʾAmīn, Muṣṭafā 2.578, 2.581, 3.193
Aminou, Mohamadou 1.333–339
ʾAmīr, ʿAbdallāh ʾAḥmad 3.348
ʾAmīr Talāʾī ibn Ruzziq 2.601
Amiritic 3.227
ʾamkan 4.445
ʾamkaniyya 4.432
ʿāmm 4.685–686
Amm Timan Arabic 1.362, 4.710, 4.716
ʾammā 1.476, 2.224, 3.35, 3.200, 4.486, 4.504–505, 4.544–545
ʾammā…fa- 1.183, 1.472, 2.148, 3.260, 3.589
ʿammāl 1.331, 1.608
Amman 2.498–500, 2.502, 2.505–517, 3.11
Amman Arabic 1.439, 1.596, 1.650, 2.117, 2.499, 2.505–517, 3.11, 3.403, 3.726, 4.325, 4.632, 4.634–635, 4.637, 4.681
ʿAmmāniyyīn 2.507
ʿAmmār, ʿAbbās 1.589
ʿAmmār, Sām 4.460
Ammar, Wafaa 2.677–678
ʿāmmat al-ʿArab 2.630
ʿāmmiyya 1.405, 1.409, 1.419, 1.483, 1.527, 1.629, 1.667, 2.649, 4.252
Ammonite 3.409
Ammonius 3.547
Amorite 3.227, 3.410, 4.170
Amour, Salmin 1.662
amplitude 2.396
ʾamr 1.310, 2.359, 2.478, 4.332, 4.638
ʿAmr ibn al-ʿĀṣ 1.361, 3.53
ʾamṣār → *miṣr*
ʿĀmūda Arabic 4.407
amulet 2.45, 4.181
ʾan 1.71, 1.109, 1.196, 1.502, 2.91, 2.171–173, 2.195, 2.301, 2.355, 2.427, 2.627, 3.105, 3.221, 3.272, 4.60, 4.359–360, 4.363, 4.432, 4.544–545, 4.746
ʾan al-maṣdariyya 4.360–361, 4.363
ʾan muxaffafa 2.171, 4.361, 4.363
ʾan with indicative 3.691, 3.695
ʿĀna 4.407
ʿĀna Arabic 2.415
anacoluthon 4.486
Anakara 3.127
analogical explanation 4.14
analogical extension 2.194
analogical reasoning 4.11, 4.14
analogy 1.74–82, 1.175, 1.193, 1.245, 1.452–453, 1.590, 1.616, 1.644, 2.159, 2.165, 2.184–185, 2.193, 2.85, 2.305, 2.310, 2.425, 2.448, 3.118, 3.191, 3.446, 4.41–42, 4.119
analogy, pattern-based 1.79–80
analogy, proportional 1.74, 1.76, 1.79
analogy, semantic 2.160
analysis, morphological 1.209
analytic 1.356–357, 1.359, 1.486, 1.499, 1.552, 1.562, 1.614, 1.624, 2.161, 2.197, 2.265, 2.489, 2.684, 3.182, 3.184, 3.280, 3.671, 4.311
analytic construction 2.195, 2.522, 2.525, 2.594
analytic feature 3.59
analytic form 4.306
analytic genitive 1.64, 1.82–85, 1.109, 1.246, 1.351, 1.491, 1.605–607, 2.197, 2.213, 1.263, 1.266, 2.570, 2.613, 3.62–63, 3.365, 3.368, 3.390, 3.404, 3.406, 3.432, 3.482, 3.587, 3.701, 4.247
analytic passive 3.553
analytic structure 1.606, 2.118, 2.120–121, 3.152, 3.641
analyzer, morphological 1.459–460
ʿanʿana 4.125
anaphora 1.49, 1.310–312, 1.354–355, 1.377, 1.431, 1.493, 1.514, 1.568–569, 2.83, 2.161, 3.589, 3.713, 3.716, 4.392
anaphora, adverbial 1.569
anaphora, nominal 1.569
anaphora, verbal 1.569
anaphoric expression 2.19
anaphoric island 1.481
anaptyctic vowel 1.234, 1.261, 1.263, 1.664, 2.152, 2.482, 2.484, 2.572–573, 2.577, 4.243

anaptyxis 1.85–86, 1.90, 1.113, 1.225, 1.235, 1.262, 1.320–321, 1.390, 1.400, 1.608, 2.35, 2.417–418, 3.266, 3.327–328, 3.363, 3.403, 3.611, 3.692, 3.695, 4.243
ʾAnas, Muḥammad 1.316
anʾAsər Arabic 2.2
Anatolia 1.87, 1.181, 2.414, 4.584, 4.588
Anatolian Arabic 1.84, 1.87–96, 1.537–538, 1.584, 1.607, 1.612, 2.313, 2.420, 2.423, 4.407, 4.503, 4.534, 4.590–591, 4.727, 4.729, 4.732–734
ʿAnaze 3.326, 4.125, 4.403
ʿAnaze Arabic 1.260, 1.606, 1.610–611, 2.152, 3.326, 3.705
ʿAnaze confederation 3.402
ʾAnbār 1.149, 3.693
ʿAnbarīn-qalam 3.338
ʾAnbāṭ 1.617
ancestry 2.59
ʾAndalus, al- 1.96–102, 1.545, 1.620, 2.60, 2.62, 2.64, 2.287–288, 2.671
ʾandalusī 3.112
Andalusian Arabic 1.57, 1.84, 1.96, 1.99, 1.101–111, 1.168, 1.170, 1.608, 1.611, 2.158, 2.282, 2.284–287, 2.289–290, 2.388, 2.360, 2.630, 3.84, 3.237, 3.738, 4.71, 4.217
Andersen, Roger W. 2.683
Anderson, John M. 1.343–344
Anderson, Stephen R. 3.302, 4.339
Andhra Pradesh 4.452
Andrews, Avery 1.41
Andrews, Peter Alford 4.580
androcentricity 2.642–643, 2.645
Andronov, Michail S. 4.433–434
Anfūši 2.1
anggeĩon 3.106
Anghelescu, Nadia 1.567, 3.234–235, 3.353–354, 4.484–487
Angoujard, Jean-Pierre 2.165, 2.448, 3.612–614, 3.616, 3.620–623, 3.726, 4.51, 4.350, 4.681
Ani, Moukhtar 3.655, 4.647
animacy 1.41, 2.424, 4.733
animacy hierarchy 2.594
animal sounds 4.286
animals, calls to 2.379
animals, names of 1.504, 2.27, 2.289, 4.50, 4.285
animate 1.34, 1.41–42, 1.344, 1.433, 1.436, 1.645
animateness 2.99
ʾAnīs, ʾIbrāhīm 1.614–615, 1.617, 1.640, 2.555, 3.14–15, 3.72, 3.93–94, 3.689–690, 3.697, 4.87–88
ʿAnizah → ʿAnaze
ʿAnizah Arabic → ʿAnaze Arabic
Anjouan [Ndzuwani] 1.450
Anjra Arabic 3.290–293
ʾanna 1.203, 1.350, 1.467–469, 1.473, 1.476, 2.53, 2.172–174, 2.331–332, 2.354–355, 3.221, 3.591, 3.687, 3.691, 4.60, 4.358, 4.360, 4.362–364, 4.366–367, 4.430, 4.544, 4.743
ʾanna muxaffafa 1.69
Annaba 1.54
Annaba Arabic 4.681
An-Nakel al-Arabi 1.460
annexation → annexion; construct state
annexion 1.18, 1.20, 1.367, 1.560, 2.15–16, 2.294–298
annexion → construct state
annexion, binomial/polynomial 2.296
annexion, extended 2.296
annexion, improper 2.225, 4.102
annexion, joint 2.296
ʾAnnī, Muḥammad al- 2.53
annotation, multi-level 1.515
anomia 2.690
Ansaba 1.655
ʾanṣār 1.128–130
ʾAnṣārī, ʿAbd ar-Raḥmān T. al- 3.467, 3.471
ʾAnṣārī, ʾAbu Zayd al- 3.34, 3.689
Anshen, Frank 4.524–527, 4.726
answering system 1.455
Anṭākī, al- 3.224, 4.625
Antakya 1.111–112, 1.114–116
Antambahoaka 3.127
antanaclasis 3.538
ʿAntara 3.211
antecedent 1.311, 1.431, 1.556, 2.18–21, 83, 2.171, 2.226, 2.452, 4.71, 4.392–393

antecedent, quantificational 4.81, 4.83–84
antecedent government 2.171–172, 2.174
Antemoro 3.125, 3.127
Antep 2.220
anterior 1.524
anteriority 1.220, 3.99, 3.643
anthem, national 1.660
anthropology, linguistic 2.651, 2.654
anthropomorphism 2.580, 3.122
Anti-Atlas 2.62
anticipation 1.230
anticipatory suffix 1.230, 1.284
Anti-Lebanon 3.372
Antioch 1.111, 1.388, 2.220, 4.580
Antiochia Arabic 1.111–119, 2.313, 4.404
antiphrasis 1.628
antisymmetry 3.230, 3.232–233, 4.746
antithesis 4.211
antonomasia 3.121
antonymy 2.71, 3.38, 4.335
Antoun, Richard T. 4.74, 4.77
Antwerp Polyglot Bible 1.169
Anwar, Mohamed Sami 3.81, 3.83–84, 4.335
Anwarī, Ḥ. 2.410
aorist, Semitic 4.178
Aoun, Joseph 1.310–313, 1.350, 1.411–414, 1.457, 2.170, 2.173–174, 2.369–370, 3.384, 3.432, 4.64, 4.67–68, 4.80–85, 4.193–194, 4.363, 4.365, 4.394, 4.397, 4.400, 4.524, 4.527, 4.649, 4.657–658, 4.721, 4.746–747
aphaeresis 2.457
aphaeretic verb form 1.186
aphasia 1.425, 2.675, 2.681, 2.688, 2.690, 2.692, 3.305, 4.96, 4.271
aphasia, agrammatical 2.689
aphasia, Broca's 2.690
aphasia, childhood 2.689
aphasia, conduction 2.690
aphasia, congenital 2.689
aphasia, decoding 2.690
Aphasia, Developmental 2.689
aphasia, efferent 2.690
aphasia, encoding 2.690
aphasia, expressive 2.690
aphasia, fluent 2.690
aphasia, non-fluent 2.690
aphasia, sensory 2.690
aphasia, verbal 2.690
aphonia 2.675, 2.688
aphorism 4.214
Aphrodito Archive 2.23
apocopate 2.16, 2.92, 3.269, 4.557, 4.639
apocope 2.17, 2.36, 2.67, 4.682
apocope, vowel 4.675
apodosis 1.468, 2.477–479, 3.261
apology 3.661
Apophonic Path 1.121–122
apophony 1.37, 1.119–123, 1.574, 1.624, 2.168, 2.346, 3.191, 3.243, 3.250, 3.297, 3.302, 4.638
A-position 4.395, 4.488
Appel, René 2.97, 3.315
Appleyard, David L. 3.588–593
applicative 3.181
apposition 1.13, 1.123–126, 1.473, 1.559, 1.645, 3.101, 4.221
apposition, close 1.125
apposition, conjugated 3.66
apraxia, verbal 2.690
Apuleius 4.215
ʿAqaba 2.500, 3.401, 3.404
ʿAqaba Arabic 2.498
ʿāqil 2.15
ʿAqīl, ʿAlī 4.692, 4.697
ʿAql, Saʿīd → Akl, Saïd 1.603
ʾaqlām → qalam
ʾaqlām al-ʾuṣūl, al- 3.340, 4.131
ʾaqlam as-sitta, al- 1.152, 2.598, 3.341, 4.131
Aqmar, mosque of al- 2.41, 2.43, 2.600
ʿAqqād, ʿAbbās Maḥmūd al- 1.472, 1.648, 4.87, 4.89–90
ʿAqra Arabic 2.415–416, 2.419–423, 2.534, 3.448
ʾaqṣā l-ḥalq 4.426
ʾaqṣā l-lisān 4.1
ʾaqsām al-kalām 2.425, 4.329
Aquilina, Joseph 3.142–144, 3.146–147, 3.156, 3.159
ʾaqyasu 2.543, 4.10
ʿarab al-ʿāriba 1.129

Arab Americans 2.60–61
Arab ana bongor 3.634
Arab emigration 3.1–2
Arab League 1.131, 1.448–450, 1.663, 2.61, 3.174, 3.344, 3.351–352
Arab nation 3.344
Arab Revolt 3.402
Arab unity 2.656
Arab world 2.654
ʿaraba 1.131, 2.85, 2.401
arabe choa → Arabic, Shuwa
arabe médian 1.58, 3.664
arabes 1.129
ʿarabī 1.131, 3.345–346
Arabia Petraea 3.400
Arabian, Ancient North → North Arabian, Ancient
Arabian, Ancient South → South Arabian, Ancient
Arabian, East → Arabic, East Arabian
Arabian Nights → *ʾAlf layla wa-layla*
Arabian Peninsula 1.23, 1.617, 2.555–556
Arabic 2.74, 2.374, 2.494, 62.57, 2.692, 2.697, 3.545–546, 4.67, 4.306, 4.662, 4.678
Arabic, Abbéché 3.672, 3.674, 4.708, 4.710, 4.716
Arabic, ʿAbd al-Qays 3.94
Arabic, Abha 4.127
Arabic, ʿAbīda 3.606
Arabic, Abu Dhabi 3.702
Arabic, Abu Jirj 2.5
Arabic, Abu l-Xaṣīb 1.272
Arabic, Abu Tij 2.5
Arabic, Abu-Kamal → Arabic, Albū Kmāl
Arabic, Abyan 4.751, 4.753–756
Arabic, Adana 1.391, 1.394–395, 2.313
Arabic, Aden 1.23, 1.260, 1.501, 2.388, 4.690, 4.750–751, 4.754, 4.756–757
Arabic, Afāḏle 4.407
Arabic, Afghanistan 1.537, 1.583, 3.84, 4.617
Arabic, ʿAgāylah 3.401
Arabic, Aḥaywāt 3.401, 3.403, 4.238
Arabic, ʿAjārma 1.626, 3.405
Arabic, Ajloun 2.506
Arabic, ʿAjmān 1.590, 3.326
Arabic, ʾĀl Murrah 3.326–327, 4.127, 4.124
Arabic, ʾĀl Wāhiba 3.478–479, 3.483, 3.489
Arabic, ʾĀl Zubayr 1.272
Arabic, al-ʿArīš → Arabic, El-Arish
Arabic, Alawi 1.391, 1.395–396
Arabic, ʿAlāwīn 3.401
Arabic, alBayḏāʾ 4.751
Arabic, Albū Kmāl 4.407, 4.678, 4.680
Arabic, Alegat → Arabic, ʿLēgāt
Arabic, Aleppo 1.445, 1.537, 1.608, 1.625, 2.313, 2.494, 2.534, 2.608, 3.450, 3.607, 4.64, 4.407–408, 4.664, 4.681
Arabic, Alexandretta 2.313
Arabic, Alexandria 1.439, 1.441, 1.589, 1.604, 2.1–3, 2.11, 2.495, 4.663
Arabic, Algerian 1.418, 1.492, 1.502, 1.584
Arabic, Algiers 1.440, 1.443–444, 1.596, 1.610, 2.532, 2.534, 3.587, 4.569
Arabic, Altinözü 1.117
Arabic, Amm Timan 1.362 4.710, 4.716
Arabic, ʿĀmūda 4.407
Arabic, ʿĀna 2.415
Arabic, anʾĀṣər 2.2
Arabic, Anatolian 1.537–538, 1.584, 1.607, 1.612, 2.313, 2.420, 2.423, 4.407, 4.503, 4.534, 4.590–591, 4.727, 4.729, 4.732–734
Arabic, ʿAnaze 1.606, 1.610–611, 2.152, 3.326, 3.705
Arabic, Ancient 4.478–479, 4.481
Arabic, Ancient South Palestinian 1.386
Arabic, Andalusian 1.57, 1.84, 1.96, 1.99, 1.101–111, 1.168, 1.170, 1.608, 1.611, 2.158, 2.282, 2.284–287, 2.289–290, 2.388, 2.360, 2.630, 3.84, 3.237, 3.738, 4.71, 4.217
Arabic, ʿAnizah → Arabic, ʿAnaze
Arabic, Anjra 3.290–293
Arabic, Annaba 4.681
Arabic, Anti-Lebanon 2.494
Arabic, Antiochia 2.313, 4.404
Arabic, ʿAqaba 2.498
Arabic, ʿAqra 2.415–416, 2.419–423, 2.534, 3.448

Arabic, Arabian Peninsula 1.260, 1.273, 1.444, 1.584, 1.625, 2.152
Arabic, Arabkhona 4.613–616
Arabic, Arak 4.406
Arabic, Arbīl 2.415, 2.417–419, 2.421, 2.532, 2.534
Arabic, ʾAsad 2.494, 3.14, 3.91–92
Arabic, ʾAsd 2.555
Arabic, Asir 1.504, 1.590, 2.556
Arabic, Aswan 2.2, 2.5
Arabic, Asyūṭ 1.586, 2.2, 2.5, 3.405
Arabic, Atia 4.716
Arabic, Awlād Saʿīd 4.240
Arabic, ʿAwlaqī 4.756
Arabic, Axārsah 3.401, 3.403, 4.238
Arabic, ʿAyāydah 3.401, 4.238
Arabic, ʿAzāzmih 3.400
Arabic, ʾAzd 3.692, 3.697, 4.376
Arabic, ʾAzd-Šanūʾa 3.89, 3.94
Arabic, ʾAzd-ʿUmān 3.94
Arabic, Azəx 4.407
Arabic, Badārah 4.238
Arabic, Badāri 2.5
Arabic, Baggāra 3.62
Arabic, Baghdadi 1.271–272, 1.442, 1.444–445, 2.8, 2.414–424, 2.532, 2.572, 2.577, 3.9, 3.84, 3.237, 3.525, 3.569, 3.666–668, 3.672, 3.674, 3.703, 4.62, 4.117, 4.194, 4.304, 4.325
Arabic, Baghdadi Christian 1.273, 1.537, 1.607, 1.610, 2.276, 2.313, 2.532, 4.633
Arabic, Baghdadi Jewish 1.351, 1.537, 1.607, 1.610, 1.625, 2.414, 2.276, 2.313, 2.532–534, 4.664
Arabic, Baghdadi Muslim 1.271, 1.612, 1.625–626, 1.642, 2.276, 2.611, 2.617, 3.495, 3.585, 4.128, 4.633
Arabic, Bagirmi 4.708, 4.715
Arabic, Baḫariyya 1.270, 1.586, 1.610, 2.6, 4.550
Arabic, Bāhila 3.94
Arabic, Bahrāʾ 3.94
Arabic, Bahraini 2.494, 2.609, 2.615–619, 3.478, 3.729, 4.64, 4.656
Arabic, Bahraini Shiʿite 4.127–128
Arabic, Bakr ibn Wāʾil 2.555, 3.94
Arabic, Bal Ḥāriṯ ibn Kaʿb 1.641, 3.89
Arabic, Bal Qarn 4.127
Arabic, Balāṭ 1.585, 2.2
Arabic, Balawiy 4.677
Arabic, Balġāwi 2.506
Arabic, Balkh 4.311
Arabic, Balqāʾ 2.506
Arabic, Balṭīm 2.2–3
Arabic, Balyana 1.572
Arabic, Bani ʿAṭiye 3.401–402, 3.404, 3.407, 4.240
Arabic, Bani Bišr 4.127
Arabic, Bani Hamida 2.503
Arabic, Bani Hasan → Arabic, Beni-Hasan
Arabic, Banī Ṣaxar 1.406
Arabic, Bani Šihr 4.125, 4.127
Arabic, Bani Swayf → Arabic, Beni Swef
Arabic, Bani Ṯōr 4.126
Arabic, Baniy Wāṣil 4.240
Arabic, Barīs 2.2, 4.196
Arabic, Bašandi 2.2, 4.286
Arabic, Baskinta 1.270, 4.656
Arabic, Basra 2.312, 2.415–416, 2.421, 2.631, 3.525
Arabic, Bāṭina 3.479, 3.481, 3.484–485, 3.489
Arabic, Batman 2.608
Arabic, Bawīṭi 2.2
Arabic, Bdūl 3.402, 3.404–405
Arabic, Bedouin 2.587, 2.632, 3.205, 3.265, 3.360, 3.585, 3.609, 3.621, 4.2–3, 4.124
Arabic, *begūl* 3.190
Arabic, Bəḥzāni 1.541, 1.571, 1.607
Arabic, Beirut 1.269, 1.295, 1.323, 1.538, 1.607–608, 2.231, 2.314, 2.509, 3.11, 3.674–675, 4.349, 4.630, 4.632, 4.678, 4.681
Arabic, Beirut Sunni 1.283
Arabic, Benghazi 2.532, 3.55–56
Arabic, Beni Swef 1.322, 1.608, 1.611, 2.2, 2.5, 2.494
Arabic, Beni-Hasan 2.173, 2.503, 3.622, 3.707, 4.349, 4.351
Arabic, Bʿēri 1.271, 1.299–308, 2.2, 5, 2.151–153
Arabic, Biba 2.5
Arabic, Biḥēra 1.589, 2.1–2
Arabic, Bilī 3.400–401, 3.405
Arabic, Bimbashi 2.517–518
Arabic, Bir Biʿēri 2.2

Arabic, Bir Zeit 1.625, 3.430, 3.684, 3.686, 4.317, 4.656
Arabic, Bišmizzīn 1.540, 3.607, 4.70
Arabic, Biyyāḍiyyah 3.401, 3.403, 4.238
Arabic, Bizerte 4.563, 4.681
Arabic, Blida 1.609
Arabic, Bongor 1.364, 3.61, 3.634–639
Arabic, Bornu 3.634
Arabic, Brak 2.622
Arabic, Branes 1.295
Arabic, Btiġrīn 1.270
Arabic, Buḥayra → Arabic, Biḥēra
Arabic, Bukayriyyah 4.127
Arabic, Bukhara 4.311, 4.613–616
Arabic, Bulāg 2.2
Arabic, Buraymi 3.481
Arabic, Burullus 2.2–3
Arabic, Cairene 1.270, 1.275, 1.300, 1.304, 1.306, 1.312, 1.406–410, 1.439, 1.441, 1.445, 1.492, 1.570, 1.589, 1.595–596, 1.604, 1.608, 1.625, 2.1–3, 2.5, 2.8, 2.117–118, 2.22, 2.120, 2.231, 2.354, 2.395–396, 2.404, 2.495, 2.533, 2.657–658, 3.9–11, 3.305, 3.449, 3.451, 3.426–427, 3.430, 3.542, 3.545–546, 3.569, 3.584, 3.586, 3.592, 3.607, 3.609–611, 3.613–614, 3.616–618, 3.620–621, 3.665–668, 3.673–675, 3.704, 3.737, 4.70–71, 4.85–86, 4.117, 4.168, 4.194, 4.255, 4.304, 4.326, 4.344, 4.346, 4.389–390, 4.533–534, 4.634, 4.637, 4.669, 4.681, 4.716, 4.727
Arabic, Cairene Jewish 2.532–533
Arabic, Casablanca 1.596, 1.609, 2.621, 3.59, 3.273–274, 3.276–278, 3.285, 3.288–289, 3.291–293, 4.630, 4.679
Arabic, Central African 1.537
Arabic, Central Arabian 1.570, 1.624, 2.213, 3.693
Arabic, Central Asian 2.290, 3.429, 3.702, 4.70, 4.612–613, 4.617, 4.621
Arabic, Central Delta 1.323, 1.586
Arabic, Central Iraqi 1.269, 1.272
Arabic, Central Najdi 1.272, 1.606
Arabic, Central Palestine 1.608
Arabic, Chadian 1.360–368, 1.521, 2.388, 3.634–636, 3.638, 4.715–716
Arabic, chairs of 1.166–167, 1.169, 1.171
Arabic, Chaouen 3.288–292
Arabic, Cherchell 1.270, 1.609, 3.452
Arabic, Christian 1.270, 2.23, 2.277
Arabic, Christian Middle 1.383–387, 2.266, 3.218
Arabic, Cilician 1.388–397, 2.160, 3.674, 3.702, 4.404, 4.589, 4.594
Arabic, Classical 1.397–411, 1.490, 1.527–534, 1.578, 1.581–582, 1.592, 1.594, 1.614–620, 1.630–631, 1.633–634, 1.636, 1.669–670, 2.263, 2.267, 2.388, 2.404, 2.695, 2.708, 3.12, 3.94, 3.217, 3.318, 3.465
Arabic, Collo 1.609
Arabic, Colloquial 1.405–408, 1.410, 1.419, 1.439–442, 1.527, 1.597–598, 1.604, 1.617, 1.631, 2.708, 3.24
Arabic, Constantine 1.270, 1.596, 1.609, 2.534
Arabic, Crater 4.754
Arabic, creole 1.205, 1.518–527, 4.376
Arabic, criticism of 2.336
Arabic, Çukurova 1.469, 1.502, 1.540, 4.532
Arabic, Cypriot (Maronite) 1.365, 1.536–543, 1.607, 1.612, 2.160, 2.231, 3.84, 3.672, 3.674–675
Arabic, Cyrenaican 2.313–314, 3.55–56, 3.611, 4.551
Arabic, Cyrenaican Jebel 2.151–152
Arabic, Dabba 3.94, 3.693
Arabic, Ḍafīr 3.326
Arabic, Dakhla 1.300, 1.586, 2.2, 2.6, 2.8
Arabic, Ḍālaʿ 4.750–751, 4.753–754, 4.756
Arabic, Damascene 1.298, 1.323, 1.395, 1.407, 1.493, 1.537, 1.540, 1.546–555, 1.608, 1.625, 2.22, 2.30, 2.162–163, 2.231, 2.278, 2.502, 2.509, 2.658, 3.11, 3.265, 3.373, 3.448–449, 3.564–565, 3.569, 3.612, 3.617–622, 3.702, 3.736–737, 4.194, 4.304, 4.344, 4.349, 4.404, 4.407–408, 4.504, 4.630, 4.632, 4.656, 4.664, 4.681
Arabic, Damascus Plain 1.607

Arabic, Damietta 1.323
Arabic, Daqahliyya 2.2
Arabic, Daragözü 1.571, 2.160, 3.84, 3.672, 3.674
Arabic, Darj 3.55–56
Arabic, Daṯīna 1.544, 2.388, 3.588, 3.705, 4.750–751, 4.754–755, 4.757
Arabic, Dawāġrah 3.401, 3.404–407, 4.238
Arabic, Daxla → Arabic, Dakhla
Arabic, Dbūr 4.238
Arabic, Debdou 3.289, 3.291, 4.70
Arabic, Dellys 1.270, 1.609, 4.656
Arabic, Delta 1.608, 2.2, 4, 2.11
Arabic, Dēr ʿAṭiye 4.406
Arabic, Dēr izZōr 2.414, 4.407–408
Arabic, Dhofar 4.304, 4.751
Arabic, Dirbasīye 4.407
Arabic, Diyarbakır 1.538, 1.607, 2.420, 4.70
Arabic, Djeinau 4.613–614, 4.616
Arabic, Djerba 4.633
Arabic, Djidjelli 1.609, 3.674, 3.737, 4.70, 4.371
Arabic, Djogari 4.613–615, 4.617
Arabic, Dōsiri 2.609
Arabic, Drayj 4.406
Arabic, Druze 1.275–276, 1.442, 1.445, 2.314
Arabic, Druze Golan 2.314
Arabic, Druze Horan 2.314
Arabic, Dubai 4.62
Arabic, Ḍufār → Arabic, Dhofar
Arabic, Duhūs 2.2
Arabic, Ḍullām 3.400, 4.238
Arabic, Dūš 2.2
Arabic, Early Christian 3.592
Arabic, Early Christian Middle 1.383
Arabic, Early Standard 4.478
Arabic, East Arabian 1.28, 1.490, 1.618, 2.160, 2.617, 3.94, 3.691, 3.645, 4.9–10, 4.124, 4.127, 4.286, 4.302, 4.304–305, 4.676
Arabic, East Egyptian 1.266
Arabic, East Libyan 3.53, 3.56, 4.534
Arabic, East Moroccan 1.650
Arabic, East Sudanese 4.376
Arabic, East Syrian 1.589
Arabic, East Yemeni 4.304
Arabic, Eastern 1.259–269, 1.536, 1.538, 1.540–541, 1.605, 1.614, 2.8, 2.263–264, 3.697
Arabic, Eastern Bedouin 1.572
Arabic, Eastern Delta 1.589, 2.389
Arabic, Eastern Desert 2.2
Arabic, Eastern Mediterranean 4.729–734
Arabic, Edfu 1.306
Arabic, Educated 1.630, 1.666, 2.377, 2.696, 3.14, 3.18–19
Arabic, Educated Colloquial 1.409–410
Arabic, Educated Spoken 1.630–632, 1.634, 1.666–668, 2.661, 4.325–326
Arabic, Egyptian 1.270, 1.323, 1.340–342, 1.351, 1.384, 1.417, 1.419, 1.483, 1.495, 1.501–504, 1.538, 1.554, 1.562, 1.570–572, 1.584, 1.595, 1.604, 1.606–609, 1.611, 1.632, 1.636, 1.657–659, 2.20, 2.32, 2.36–38, 2.71, 2.74, 2.96–100, 2.136, 2.160–161, 2.194, 2.199, 2.203, 2.344, 2.375, 2.388, 2.390, 2.455–459, 2.517, 2.560, 2.584, 2.587, 2.597, 2.649, 2.656, 2.658, 2.678, 3.12, 3.65–68, 3.70, 3.81–85, 3.87, 3.97, 3.190, 3.198, 3.200–201, 3.236, 3.240, 3.264–268, 3.386, 3.389–393, 3.398, 3.456, 3.545, 3.560–563, 3.572–573, 3.586–587, 3.609, 3.612, 3.669, 3.674, 3.676, 3.680, 3.684, 3.687, 3.705–706, 3.708–711, 4.45, 4.50–51, 4.62–64, 4.66, 4.71, 4.195–198, 4.284–285, 4.317, 4.354, 4.372, 4.399, 4.468, 4.470, 4.490, 4.649, 4.662, 4.671–672, 4.719, 4.742, 4.747, 4.757
Arabic, Egyptian Jewish 2.276, 2.278–279
Arabic, Egyptian Nile Valley 2.151
Arabic, Egyptian Oases 1.586, 1.590
Arabic, Egyptian Radio 2.278
Arabic, El Jadida 3.288
Arabic, El-Arish 4.240
Arabic, El-Karak 2.506
Arabic, Empty Quarter 3.479, 3.484, 3.489–490
Arabic, Esna 2.5
Arabic, Essaouira 3.288
Arabic, Farafra 1.586, 1.610–611. 2.2, 2.6, 2.8

Arabic, Fašn 2.2
Arabic, Fawaxriyyah 4.240
Arabic, Fayfa 4.123
Arabic, Fayyum 1.320–321, 1.608, 1.611, 1.640, 2.2, 2.5, 2.494
Arabic, Fes 1.444, 1.571, 1.596, 2.532–534, 2.621, 3.273–274, 3.276, 3.285, 3.288–290, 3.292–293, 3.585
Arabic, Fezzan 1.265–266, 2.622, 3.55–56, 3.587, 4.551, 4.555
Arabic, Formal Spoken 1.668
Arabic, Gabes 3.451, 4.681
Arabic, Galilee 1.265, 1.608, 2.501, 4.663
Arabic, Ġāmid 4.126
Arabic, Ġanī 3.94
Arabic, Garabulli 3.55
Arabic, Garāršа 3.401, 3.403, 4.240
Arabic, Ġarbiyya 2.2
Arabic, Gaṣir 2.2
Arabic, Ġaṭafān 3.94
Arabic, Ġaṭīf 3.89
Arabic, Ġayl Ḥabbān 4.750–751, 4.756
Arabic, Gaza 1.608, 2.32, 3.527, 3.531
Arabic, Gḥaṭān 3.326
Arabic, Gihaf 4.756
Arabic, *gilit* 1.610–611, 4.3, 4.325
Arabic, Gina 2.2
Arabic, Giza 2.2, 2.5
Arabic, Gorda 2.622
Arabic, Goulimine 3.288
Arabic, Granada 1.103, 1.107, 3.23, 3.585, 4.431
Arabic, Greater Syria 1.444
Arabic, Gulf 1.270, 1.272, 1.595, 1.606, 1.611, 1.624, 2.36, 2.38, 2.96, 2.161, 2.210–213, 2.374–375, 2.388, 2.390, 2.555, 2.587, 3.9, 3.12, 3.19, 3.63, 3.68, 3.264, 3.451–452, 3.478, 3.545, 3.585, 3.587, 3.592, 3.609, 4.124, 4.317, 4.646, 4.656, 4.695
Arabic, Gulf Pidgin 3.63
Arabic, Gurna 2.5
Arabic, Gūṣ 2.5
Arabic, Gypsy 2.216–222, 3.527
Arabic, Ḥaḍramawt 1.260, 2.270, 2.274, 3.449, 4.687–699, 4.750–751, 4.753, 4.754–757
Arabic, Ḥaḍrami → Arabic, Ḥaḍramawt
Arabic, Haifa 1.608
Arabic, Ḥalāyib 2.2
Arabic, Hama 1.608, 4.408
Arabic, Ḥamādah 4.240
Arabic, Hamar 4.376
Arabic, Hamdān 3.94
Arabic, Hanādwah 4.240
Arabic, Ḥanīfa 3.94
Arabic, Ḥarb 4.126–127
Arabic, Ḥāriṯ 3.94
Arabic, Ḥasa 4.127–128
Arabic, Ḥasake 4.407
Arabic, Ḥassāniyya 1.570, 1.572, 1.625, 2.240–250, 2.388, 2.622, 3.53, 3.60–61, 3.135, 3.288, 3.290, 3.314, 3.352, 3.450–451, 3.585, 3.587, 3.664, 4.70, 4.313, 4.376, 4.550, 4.662
Arabic, Hatay 1.584
Arabic, Hawāzin 2.555, 3.94
Arabic, Ḥāyil 4.126
Arabic, Hebron 1.608, 2.30, 3.527
Arabic, Ḥijāzī 1.295, 1.399–400, 1.490, 1.606, 1.641, 2.88, 2.199, 2.263–264, 2.279, 2.312, 2.355, 2.387, 2.499, 2.680, 2.689, 2.691, 3.14, 3.90, 3.93, 3.179, 3.404, 3.497, 3.609, 3.621–622, 3.645, 3.690–691, 3.693–698, 4.9–10, 4.62, 4.70, 4.124–127, 4.351, 4.389, 4.431
Arabic, Ḥijāzī Jewish 2.259
Arabic, Hilāli 1.609, 3.288
Arabic, Ḥimṣ → Arabic, Homs
Arabic, Ḥimyar 3.94
Arabic, Hispano- → Arabic, Andalusian
Arabic, history of 2.261–268
Arabic, Hīt 2.415, 2.533–534
Arabic, Hodeida → Arabic, Ḥudaydah
Arabic, Homs 1.608, 4.678, 4.680
Arabic, Ḥōrān 1.546, 1.583, 1.605, 1.607–608, 2.501, 2.506, 2.622, 4.407, 4.635, 4.678
Arabic, Ḥudaydah 4.754
Arabic, Huḏayl 1.627, 3.90, 3.92, 3.94, 3.692, 3.695, 3.697, 4.376
Arabic, Hufūf 4.126
Arabic, Ḥugariyyah 4.751, 4.754, 4.756
Arabic, Ḥuwayṭāt → Arabic, Ḥwēṭāt
Arabic, Huwwērīn 4.406
Arabic, Ḥwēṭāt 3.401–403, 3.407, 4.238, 4.240

Arabic, ʿIbrī 3.481
Arabic, Īgli 3.290–291
Arabic, Īgni → Arabic, Īgli
Arabic, informal written 2.214
Arabic, Inner Oman 1.642
Arabic, Inter- 1.668
Arabic, intermediate 1.630–633
Arabic, Iraqi 1.417, 1.432, 1.443–444, 1.493, 1.540, 1.639–641, 2.38, 2.96, 2.174, 2.231, 2.312, 2.388, 2.390, 3.68, 3.70, 3.429, 3.595, 3.597, 3.611, 3.613–614, 3.654, 3.702, 3.729, 4.45, 4.85–86, 4.303, 4.311, 4.318, 4.590, 4.592, 4.670–671
Arabic, Irbid 2.498–499, 2.503, 2.506, 4.655
Arabic, Ismint 2.2
Arabic, ʿIzbit al-Basīlī 1.271
Arabic, Jabal Axḍar 3.479, 3.482–485, 3.489
Arabic, Jabal idDrūz 3.592, 4.407
Arabic, Jabal Wašḥah 4.751
Arabic, Jabal Yazīdi 4.754, 4.756
Arabic, Jabali 2.314
Arabic, Jaffa 1.608, 2.502
Arabic, Jarājrah 4.238
Arabic, Jawf 4.751
Arabic, Jbāla 1.439, 1.590, 1.609, 3.288–290, 3.292–293, 4.70
Arabic, Jbāliyyah 3.401, 3.403, 4.240
Arabic, Jeddah 1.271, 4.126, 4.632
Arabic, Jenin 2.321
Arabic, Jerusalem 1.323, 1.444–445, 1.608, 2.32, 2.221, 2.231, 2.481–493, 2.502, 2.509, 2.532, 2.534, 3.8–9, 3.11, 3.658, 3.704, 4.632
Arabic, Jerusalem Jewish 2.491
Arabic, Jewish 1.54, 1.323, 1.443, 1.445–447, 2.23, 2.266, 2.277, 2.313, 2.470, 2.494, 2.526–536, 2.659, 2.661, 3.218, 3.220, 3.451, 3.472, 4.58, 4.70, 4.157, 4.216, 4.217
Arabic, Jiblah 1.571
Arabic, Jidda → Arabic, Jeddah
Arabic, Jordanian 1.656–659, 2.71–72, 2.97–99, 2.120, 2.230–231, 2.395, 2.397–398, 2.498–517, 2.585–587, 3.11, 3.263–268, 3.527, 3.604, 3.611–612, 3.678–679, 3.686, 4.407, 4.489, 4.629, 4.632, 4.648, 4.654–656, 4.664, 4.672, 4.681
Arabic, Jordanian Bedouin 4.325
Arabic, Juba 1.489, 1.520, 1.525, 2.517–525, 3.10, 3.62, 3.635, 3.639, 3.642–643, 4.198, 4.376
Arabic, Judaeo- 1.443, 2.23, 2.266, 2.277, 2.313, 2.470, 2.494, 2.526–536, 2.659, 2.661, 3.218, 3.220, 3.472, 3.451, 4.70, 4.216, 4.217
Arabic, Judaeo- → Arabic, Jewish
Arabic, Juhayna 3.94
Arabic, Kafr išŠēx 2.2
Arabic, Kairouan 1.609, 4.563, 4.670
Arabic, Kalb 3.94
Arabic, Karaduvar 1.395
Arabic, Karak 2.498–499, 2.503, 2.506–507
Arabic, Kayışlı 1.392
Arabic, Kazanlı 1.392
Arabic, Kfar ʿAbīda 1.540, 3.451
Arabic, Khabura 3.478
Arabic, Khalil 3.527
Arabic, Khanaqin 2.415
Arabic, Kharga 1.586, 2.2, 2.5–6, 2.8, 2.389, 4.196
Arabic, Khartoum 1.525, 2.271, 2.559–571, 3.10, 3.62, 3.643, 4.376, 4.495–497, 4.499, 4.529–535
Arabic, Khatuniya 2.414, 4.407
Arabic, Khurasan 1.583, 3.544
Arabic, Khuzestan 1.584, 2.414, 2.571–578, 3.84
Arabic, Kināna 3.94
Arabic, Ki-Nubi 1.357, 1.520, 3.62, 3.635, 3.642–643
Arabic, Kirkuk 2.415, 2.417–418, 2.421, 2.532
Arabic, Korba 4.681
Arabic, Kordofan 4.376
Arabic, Kormakiti 3.450
Arabic, Kozluk 4.70
Arabic, Kozluk-Sason 1.607
Arabic, Kufa 2.312
Arabic, Kufra 3.55–56
Arabic, Kurdistan 2.419, 2.423, 2.532
Arabic, Kuwaiti 1.340, 1.342, 1.657–659, 2.97, 2.102, 397, 2.608–620, 3.609, 3.679, 3.740, 4.65, 4.163, 4.303

Arabic, Lahej 4.750–751, 4.753–754, 4.756
Arabic, Lake Chad 1.611
Arabic, Latakia 1.608
Arabic, Late Christian Middle 1.383, 1.386
Arabic, Lawdar 4.754
Arabic, Lebanese 1.274–275, 1.281, 1.311–312, 1.411–412, 1.432, 1.473, 1.538, 1.540–541, 1.554, 1.568, 1.580, 1.603, 1.642, 2.4, 2.161, 2.231, 2.314, 2.367, 2.369–371, 2.395–400, 2.506, 3.240, 3.264, 3.266–268, 3.349, 3.359, 3.384, 3.387, 3.413, 3.429, 3.446, 3.569, 3.612, 3.726, 4.64, 4.67, 4.80, 4.82–84, 4.194, 4.348, 4.397, 4.649, 4.670, 4.678, 4.719, 4.747
Arabic, ʿLēgāt 3.401, 4.240
Arabic, Levantine 1.312, 1.412, 1.668, 2.4–5, 2.36–38, 2.71, 2.74, 2.136, 2.173, 2.204, 2.395, 2.510, 2.584, 2.658, 3.84, 3.359, 3.495, 3.614, 3.688, 3.707, 4.318, 4.323, 4.646, 4.675
Arabic, Libyan 1.270, 1.642, 2.6, 2.374, 2.390, 2.658, 3.54–55, 3.587, 4.655
Arabic, Libyan oases 2.8
Arabic, L-iḪmēra 4.406
Arabic, literary 1.602
Arabic, Lmnabha 4.286
Arabic, Lower Egyptian 1.608, 1.642, 3.702
Arabic, Lower Iraq 1.443
Arabic, Luxor 2.2, 2.5
Arabic, Maʿān 2.498
Arabic, Mʿaḏ̣ḏ̣amīye 4.406–407
Arabic, Maḏḥij 3.94
Arabic, Maghrebine 1.270, 1.291, 1.296, 1.492, 1.585–586, 1.590, 1.611, 2.7, 2.74, 2.160, 2.455, 2.621, 3.702, 4.281, 4.323
Arabic, Maḥābišah 4.751
Arabic, Maḥāyil 4.125, 4.127
Arabic, Mahdiya 1.609
Arabic, Maiduguri 4.716
Arabic, Malālḥah 4.238
Arabic, Mālik ibn Kināna 3.89
Arabic, Malkiye 4.407
Arabic, Maltese 1.416, 1.609, 3.65, 3.67, 3.69–70, 3.141–159, 3.237, 3.298, 3.312, 3.352, 3.545, 3.569, 3.739–740, 4.565, 4.662
Arabic, Manāma 4.325, 4.631
Arabic, Mandisa 2.2
Arabic, Mangalla 2.517
Arabic, Manzala 2.2
Arabic, Maʿqili 1.609
Arabic, Marāzīg 1.264
Arabic, Mardin 1.607, 2.608, 3.84, 3.237, 3.448, 4.70, 4.664
Arabic, Maʿrib 4.751
Arabic, Maridi 3.59, 642
Arabic, Marrakech → Arabic, Marrakesh
Arabic, Marrakesh 1.609, 2.621–622, 3.274–276, 3.288–289, 3.291–292
Arabic, Marxah 4.756
Arabic, Masāʿīd 3.401, 4.238
Arabic, Mauritanian 3.702, 4.159
Arabic, Meccan 1.300, 1.400–401, 1.501, 1.611, 1.615, 2.355, 3.90, 3.179–187, 3.335, 3.404, 3.612, 4.126, 4.534
Arabic, Medea 1.609
Arabic, Media 1.474, 1.667, 2.296, 3.17, 3.19, 3.198–204
Arabic, Meknes 1.609, 2.621, 3.85, 3.291
Arabic, Mersin 1.390, 1.392, 1.394–395, 2.313
Arabic, Mesopotamian 1.571, 1.607, 1.606
Arabic, Mḥallamīye 3.237
Arabic, Mhamid 3.288
Arabic, Mḥardi 4.403–404
Arabic, Mhīn Fruqlus 4.406
Arabic, Midan 1.547
Arabic, Middle 1.383, 1.403, 1.410, 1.496, 1.593, 1.597, 1.619, 1.631, 1.642, 1.668, 2.265–266, 2.312–313, 2.332, 2.388, 2.529, 3.466, 3.704, 4.196, 4.543–544, 4.546
Arabic, Middle Egyptian 1.320–321, 2.6
Arabic, Mila 1.56
Arabic, Miliana 1.609
Arabic, Minufiyya 2.2
Arabic, Minya 1.322, 2.2, 5
Arabic, Misurata 3.54–56
Arabic, Modern Standard (MSA) 1.399, 1.406–410, 1.423, 1.434, 1.436, 1.441, 1.446, 1.456, 1.469, 1.472–473, 1.476,

1.490, 1.493, 1.528, 1.532, 1.578, 1.580–582, 1.595, 1.624, 1.630, 1.633–634, 1.636, 1.647, 1.650, 1.652, 1.655–656, 1.667, 1.669–670, 2.267, 3.10–12, 3.14, 3.18–19, 3.76
Arabic, Mogador 1.609
Arabic, Monastir 4.564
Arabic, Moroccan 1.294–298, 1.351, 1.407–408, 1.418–419, 1.479, 1.481, 1.483, 1.486, 1.488, 1.502, 1.512, 1.568, 1.571–572, 1.585, 1.604, 2.21, 2.71–72, 2.75–76, 2.96–97, 2.102, 2.116, 2.156, 2.159, 2.161, 2.174, 2.194, 2.205, 2.231, 2.374, 2.388, 2.390–391, 2.393, 2.395–396, 2.621, 2.658, 2.686, 3.48, 3.68–70, 3.84, 3.237, 3.273, 3.287, 3.305, 3.358–359, 3.382, 3.429, 3.446, 3.554, 3.584–587, 3.592, 3.611, 3.655–656, 3.684, 3.705, 3.707, 3.724–726, 3.740, 4.45, 4.51–52, 4.67, 4.71, 4.117, 4.141, 4.160, 4.167, 4.185–191, 4.286, 4.306, 4.317–318, 4.339, 4.351, 4.393, 4.397–399, 4.491, 4.646, 4.648–649, 4.651, 4.654–4.656, 4.662, 4.666, 4.668, 4.672, 4.674–675, 4.678, 4.682
Arabic, Moroccan Jewish 1.444–445, 2.533, 3.288, 3.292–293
Arabic, Mosul 1.417, 1.537, 1.610, 2.312, 2.415–417, 2.419–420, 2.423, 2.534, 3.448, 3.674, 4.372
Arabic, Mount Hermon 4.407
Arabic, Mount Lebanon 1.269–271
Arabic, Muḍaybi 3.481, 3.489
Arabic, Mudia 4.754
Arabic, Mukalla 4.757
Arabic, Mukeyras 4.751, 4.753
Arabic, Murād 3.94
Arabic, Murayna 3.94
Arabic, Mürseloğlu 1.392
Arabic, Mušiyya 2.2
Arabic, Mūṭ 2.2
Arabic, Mzāb 1.264
Arabic, Mzēnah 2.152, 3.401, 4.240
Arabic, Nabataean 2.198, 3.467, 3.696–697
Arabic, Nabataean Old 4.479
Arabic, Nabeul 4.563
Arabic, Nabk 4.406
Arabic, Nablus 2.32, 3.527, 4.655
Arabic, Naj Hammadi 2.2, 2.5
Arabic, Najdi 1.399–400, 1.606–607, 1.625, 1.627, 2.151–152, 2.161, 2.353, 2.404, 2.506, 2.609, 3.12, 3.94, 3.360, 3.403, 3.429, 3.478, 3.585, 3.587, 3.606, 4.2, 4.123–124, 4.126–127, 4.129, 4.238, 4.241, 4.318, 4.628, 4.632, 4.654–655
Arabic, Najrān 3.326, 3.606, 4.124, 4.127
Arabic, Namir 3.94
Arabic, N’djamena 4.708, 4.710, 4.712, 4.716
Arabic, Nedroma 1.609
Arabic, Nefzaoua 4.681
Arabic, Negev 1.260, 1.266, 3.400–401, 3.403–406, 3.585, 4.238, 4.240, 4.350
Arabic, Negev Bedouin 3.360, 3.531, 3.536–537, 3.569, 3.609, 3.613–614, 4.351, 4.663
Arabic, Negev Fallāḥi 3.360
Arabic, Nʿēmāt 3.402, 3.404–405
Arabic, Neo- 1.613–614, 1.616–617, 1.619, 2.265–266
Arabic, New Fes 1.609
Arabic, newspaper 3.200, 3.202–203
Arabic, Nigerian 1.502, 2.471, 3.542, 3.544–546, 4.118, 4.710, 4.714
Arabic, Nile Delta 1.270, 1.503, 2.2, 2.7, 2.494
Arabic, Nile Valley 1.586, 1.605, 1.608, 1.625, 2.4–5, 2.7–8
Arabic, North African 1.606–607, 2.36, 2.251, 2.351
Arabic, North Algerian 1.642
Arabic, North Arabian 1.259, 1.584, 1.642, 2.494, 3.403–405, 3.407, 4.403, 4.407, 4.751
Arabic, North Israel 1571, 4.649
Arabic, North Israel Bedouin 4.649
Arabic, North Jordanian 4.655
Arabic, North Lebanese 1.607
Arabic, North Middle Egyptian 1.586
Arabic, North Palestine 1.442, 1.445

Arabic, North Sinai 1.570–571, 1.605, 2.509, 4.677
Arabic, North Syria 1.538, 1.612, 4.590, 4.677
Arabic, North Tunisian 1.625
Arabic, North Yemeni 1.584, 1.605, 4.676–677, 4.751, 4.753
Arabic, Northern Najdi 1.584, 1.606
Arabic, Northern Nile valley 1.585
Arabic, Northwest Arabian 1.606, 3.94, 3.360, 3.400, 3.646, 3.695, 4.238, 4.240, 4.241
Arabic, Northwest Moroccan 1.625
Arabic, Nubi → Arabic, Ki-Nubi
Arabic, Nubian 2.494, 3.60
Arabic, Nusaybin 2.534
Arabic, Nusayri 1.642
Arabic, Oases 2.7
Arabic, Old 1.399–400, 1.614, 2.265, 2.267, 3.415, 3.417, 3.464, 4.478–479, 4.482
Arabic, Old Fes 1.609
Arabic, Old Rabat 3.273
Arabic, Omani 1.663, 2.211, 2.619, 3.84, 3.478–491, 3.544, 3.658, 4.124, 4.298–300, 4.383
Arabic, Omdurman 4.376
Arabic, Oran 1.443, 1.596, 1.610, 2.532, 4.680
Arabic, Orania 1.609
Arabic, Oujda 3.288–289, 3.292
Arabic, Palestinian 1.312, 1.351, 1.375, 1.377, 1.481, 1.538, 1.540–541, 1.608, 2.100–101, 2.230–231, 2.365–366, 2.370, 2.461, 2.463, 2.466, 2.499, 2.502, 2.506–507, 2.509, 2.587, 2.691, 3.11, 3.82–84, 3.86, 3.263–268, 3.360, 3.403, 3.430, 3.433, 3.526–538, 3.569, 3.609, 3.656, 3.685, 3.687, 3.704–705, 4.117, 4.194, 4.317–318, 4.346–349, 4.361, 4.393, 4.407, 4.534, 4.627, 4.637, 4.654, 4.662–663, 4.672, 4.675–676, 4.681, 4.746
Arabic, Palmyra 1.538, 2.314, 2.494, 4.406–408
Arabic, Pan- 1.668
Arabic, peninsular 1.606
Arabic, peripheral 3.640
Arabic, pidgin 1.364, 1.619, 3.634–639, 4.123
Arabic, Port Said 1.604
Arabic, post-Classical 2.528, 3.216
Arabic, pre-ʿAnaze 1.606
Arabic, pre-Classical 1.491
Arabic, pre-Hilali 1.609–610, 3.53–54, 3.273–274, 3.288
Arabic, pre-Islamic 1.399–401, 1.491, 1.562, 1.593–594, 1.613–614, 1.616–618, 1.653, 2.88–89, 2.198, 2.263, 2.355, 3.14, 3.90, 3.465, 3.644, 3.689–699
Arabic, Proto- 1.313, 1.490, 1.494, 2.263, 2.495, 3.411
Arabic, Proto Neo- 1.618
Arabic, Proto-Colloquial 2.495
Arabic, Qahabah 4.127
Arabic, Qalamūn 1.538, 1.546, 1.584, 1.608, 4.403, 4.406, 4.408
Arabic, Qalhāt 3.488
Arabic, Qāmišli 2.389, 2.534, 4.407–408
Arabic, Qarītēn 1.538, 2.314, 4.406
Arabic, Qāsim 3.326, 3.606
Arabic, Qašqa-Darya 4.613–614, 4.616–617
Arabic, Qaṭar 2.609
Arabic, Qaṭīf 4.124, 4.128
Arabic, Qaṭn 4.693
Arabic, Qayrawān → Arabic Kairouan
Arabic, Qays 1.641, 4.376
Arabic, *qəltu* 1.396, 1.536–538, 1.584, 1.607, 1.610, 1.612, 2.313, 4.3, 4.70, 4.325
Arabic, *qētel* 1.538
Arabic, Quḍāʿa 1.614, 2.427
Arabic, Qurʾānic 1.615, 1.617, 2.647–648, 2.264, 2.705, 2.705, 2.708, 3.10, 3.645, 3.689, 4.21–31
Arabic, Qurayš 1.400, 1.594, 1.614–617, 1.614, 1.617, 2.263–264, 3.14, 3.89, 3.91–93, 4.377
Arabic, Rabat 1.596, 1.609, 2.621, 3.9, 3.273, 3.288–289, 3.292, 3.274
Arabic, Rabīʿa 1.614, 2.555
Arabic, Ramallah 1.584
Arabic, Rās ilʿAyn 4.407
Arabic, Rās Ṣadr 4.238
Arabic, Rašīd 2.2–3

Arabic, Rašidīyya → Arabic, Rosetta
Arabic, Rašiyda 3.62
Arabic, Rāziḥī 4.106
Arabic, reform of 2.635
Arabic, revival of 1.531
Arabic, Rijāl Alma 4.127
Arabic, Ristāq 2.388, 3.481
Arabic, Riyadh 4.62
Arabic, Rmēlāt 3.401, 4.238, 4.677
Arabic, Rosetta 2.2–3
Arabic, Rubatab 4.376
Arabic, Rufayda 4.127
Arabic, Rustāq → Arabic, Ristāq
Arabic, Rwālah 4.125–126
Arabic, Ṣabbēha 4.751
Arabic, Saʿd 3.94
Arabic, Saʿd ibn Bakr 4.376
Arabic, Saʿd Ibn Zayd Manāt 4.376
Arabic, Ṣʿadah 4.751, 4.754, 4.756
Arabic, Sadr City 1.272
Arabic, Safed 1.608
Arabic, Sahara Bedouin 1.642, 3.672
Arabic, Sāḥil Maryūṭ 2.314
Arabic, Saida 1.608
Arabic, Saʿīdi → Arabic, Upper Egyptian
Arabic, Ṣalāla 3.479
Arabic, Salé 1.609, 3.273–274
Arabic, Ṣalṭ 1.595, 2.498–499, 2.506, 2.514, 4.632
Arabic, Samānah 3.401
Arabic, Samandağ 1.117
Arabic, Samaritan 3.218
Arabic, Samāwa 2.416
Arabic, Šammari 1.260, 1.606, 3.326, 4.126
Arabic, Ṣanʿāni 1.445, 1.596, 2.39, 2.231, 3.84, 3.450–452, 3.611, 3.613, 3.620, 3.674–675, 3.702, 4.62, 4.106–115, 4.117, 4.529–535, 4.559, 4.690, 4.695, 4.697, 4.750–751, 4.753, 4.756–757
Arabic, Sarāt 3.606
Arabic, Sarāt ʿAbīdah 4.127
Arabic, Šarqiyya 1.584, 1.589, 2.2, 3.401–402, 3.404, 3.406, 4.238, 4.727
Arabic, Saudi Arabian 1.657–658, 2.37–38, 2.97, 2.99, 2.658, 3.237, 3.695, 4.655, 4.671
Arabic, Sawālha 3.401, 4.240
Arabic, Sawārkah 3.401, 4.238
Arabic, Šāwi 1.546, 4.403, 4.407–408
Arabic, Sebha 3.55
Arabic, Sefrou 2.534, 3.288, 3.292, 4.70
Arabic, Sekik 3.607
Arabic, Sendor 2.415, 2.534
Arabic, Sfax 1.609, 1.642, 4.563–564
Arabic, Šḥīm 2.314
Arabic, Shuwa 1.333–334, 1.338, 3.373, 4.284, 4.376
Arabic, Sicilian 1.416, 2.158, 2.630, 3.146, 4.215
Arabic, Sidi Kasem 3.288
Arabic, Sidon 1.269
Arabic, Siirt 1.538, 2.608, 3.585, 4.70
Arabic, simplified 1.632
Arabic, Sinai 1.260, 1.266, 1.320, 2.2, 2.4, 2.388, 3.401, 3.405–406, 4.237–251, 4.663
Arabic, Širgāṭ 2.419
Arabic, Skikda 1.609
Arabic, Skūra 3.288, 3.290, 3.293, 4.71
Arabic, Sohāg 1.306, 2.2
Arabic, Ṣōran 4.406
Arabic, Sorman 3.55
Arabic, Soukhne 1.538, 1.585, 1.607–608, 2.314, 2.494, 4.406–408, 4.678
Arabic, Sousse 1.609, 4.563
Arabic, South Algerian 1.642
Arabic, South Arabian 1.492, 1.502
Arabic, South Iraqi 1.272, 1.584
Arabic, South Jordanian 4.240
Arabic, South Kurdistan 2.533
Arabic, South Lebanese 1.608
Arabic, South Mesopotamian 1.570
Arabic, South Middle Egyptian 1.586
Arabic, South Tunisian 1.642, 3.738
Arabic, South Turkey 1.537
Arabic, South Yemeni 4.751, 4.753
Arabic, Southeast Turkey 1.541
Arabic, Southern Levantine 2.506
Arabic, Southwest Arabian 1.259–260
Arabic, spoken → Arabic, Colloquial
Arabic, Spoken Literary 1.511
Arabic, spread of 2.664, 3.59–60
Arabic, Standard 1.632, 1.636
Arabic, Standard Egyptian 1.585

Arabic, Standard Spoken 1.668
Arabic, study of 1.166–172, 4.147
Arabic, Sudanese 1.300, 1.363, 1.525, 1.570, 2.2, 2.5, 2.38, 2.51, 2.160, 2.255, 2.517–519, 2.590, 3.10, 3.435, 3.437, 3.635, 3.643, 4.376–377, 4.380, 4.559, 4.715, 4.756
Arabic, Sudayr 4.125–127
Arabic, Suhāg → Arabic, Sohāg
Arabic, Šukriyya 2.388, 3.62, 3.404, 3.544, 4.376
Arabic, Sulaymi 1.609, 4.550
Arabic, Sult → Arabic, Ṣalṭ
Arabic, superiority of 2.178, 2.181, 2.406–407, 2.630, 2.654, 3.117
Arabic, Susa 1.571, 1.642, 4.62
Arabic, Suxne → Arabic, Soukhne
Arabic, Swērkiy 4.677
Arabic, Syrian 1.281, 1.340–342, 1.408, 1.492, 1.538, 1.540, 1.571, 1.604, 1.657, 2.38, 2.96, 2.156, 2.388, 2.390, 2.397, 2.499, 2.658, 3.67–68, 3.70, 3.237, 3.264, 3.266–268, 3.446, 3.456, 3.553, 3.556–557, 3.568, 3.611–612, 3.614, 3.656, 3.684–685, 3.729, 3.739–740, 4.45, 4.52, 4.67, 4.116, 4.163, 4.195–196, 4.309, 4.312, 4.317–318, 4.339, 4.341, 4.491, 4.647, 4.649, 4.654–655, 4.678
Arabic, Syro-Lebanese 1.546, 3.585, 3.587
Arabic, Syro-Mesopotamian 1.625
Arabic, Syro-Mesopotamian Bedouin 1.607, 4.70
Arabic, Syro-Palestinian 1.388, 1.642, 3.237, 4.285–286
Arabic, Tafilalt 3.290–291
Arabic, Taġlib 3.94
Arabic, Taʿizz 4.751
Arabic, Takrouna 3.448
Arabic, Tamīm 1.399, 1.490, 1.614, 1.616, 1.641, 2.263, 2.427, 2.494, 2.555, 3.14, 3.89, 93.2–94, 3.186, 3.497, 3.645, 3.698, 4.10
Arabic, Tamīx 3.94
Arabic, Tangier 1.439, 1.609, 3.274, 3.288, 3.291–293, 3.672
Arabic, Tannūmah 4.127
Arabic, Tarābīn 3.365–366, 4.238
Arabic, Tarsus 1.390–392, 1.394–395
Arabic, Tata 3.288
Arabic, Ṭawāra 3.401, 3.403
Arabic, Tayāha 4.238
Arabic, Ṭayybit IlImām 4.406–407
Arabic, Ṭayyiʾ 1.310, 1.399, 1.628, 2.494, 3.89, 3.94, 3.692, 3.694–695, 3.697, 4.62, 4.126, 4.376
Arabic, Taza 1.609, 3.288, 3.290–291
Arabic, teaching of → teaching of Arabic
Arabic, Tenes 1.609
Arabic, Tetouan 1.596, 1.609, 3.273, 3.288, 3.290–291, 3.293
Arabic, Thōr 4.127
Arabic, Tiberias 1.608
Arabic, Tigris 1.607
Arabic, Tihāma 1.589, 1.655, 2.259, 2.270–273, 3.179, 3.454, 4.240, 4.751, 4.753–754, 4.756
Arabic, Tikrit 2.415, 2.417, 2.421
Arabic, Tlemcen 1.444, 1.609, 2.532, 2.534, 3.452, 4.70
Arabic, Tlemcen Jewish 2.533
Arabic, Tobrouk 3.55
Arabic, Ṯōr alBāha 4.751
Arabic, Tourkou → Arabic Turku
Arabic, Trans-Jordanian 2.499
Arabic, Tripoli 2.532, 2.534, 3.53, 3.55, 3.607, 3.666–668, 4.548–556
Arabic, Tripoli Jewish 4.548
Arabic, Tripolitanian 3.55–56
Arabic, Tunis 1.260, 1.444, 1.446–447, 1.596, 1.609, 1.642, 2.534, 3.12, 3.85, 3.607, 3.737, 3.739, 4.563–571, 4.629
Arabic, Tunis Jewish 1.609, 2.277, 3.449, 4.563
Arabic, Tunisian 1.423–424, 1.457, 1.632, 1.635–636, 2.74, 2.161, 2.318, 2.320–322, 2.374, 2.388, 2.455–458, 2.653, 2.659, 3.11–12, 3.84, 3.226, 3.349, 3.352, 3.585, 3.607, 3.729, 4.95, 4.318, 4.323–324, 4.326, 4.648, 4.671
Arabic, Tunisian Jewish 4.575
Arabic, Turku 1.364, 3.61, 3.634
Arabic, Tuwa 4.754
Arabic, Ṭuwāra → Arabic, Ṭawāra
Arabic, Tuz Khurmatu 2.415

Arabic, Tyre 1.269
Arabic, Ulād Brāhīm 3.704
Arabic, Upper Egyptian 1.270, 1.299, 1.306–308, 1.492, 1.590, 2.2, 2.4–5, 2.11, 2.160, 2.314, 2.494, 3.404, 3.613–614, 3.702, 4.349
Arabic, Upper Iraqi 1.444
Arabic, ʿUqayl 3.14, 3.94
Arabic, Urban Cultivated 1.668
Arabic, Urfa 2.608
Arabic, Uzbekistan 1.537, 1.618, 2.414, 2.420, 2.670, 3.84, 3.237, 3.449, 3.544, 3.588, 3.672, 4.198, 4.503, 4.612–613, 4.727
Arabic, Valencia 4.431
Arabic, vehicular 1.364, 1.366, 3.60–62
Arabic, vernacular 1.527–528
Arabic, Wādi Ḥaḍramawt 3.478, 4.687–699
Arabic, Wādi Naṭrūn 2.2
Arabic, Wahrān → Arabic, Oran
Arabic, Wasṭa 2.5
Arabic, West Arabian 1.490, 1.618, 3.94, 3.693–694, 3.696, 3.698
Arabic, West Libyan 3.54–56
Arabic, West Sudanic 1.525–526, 1.618, 3.544, 3.634, 4.376, 4.708–718
Arabic, West Yemeni 2.271, 4.304
Arabic, Western 1.259–269, 1.605, 1.614, 2.8, 2.263, 4.674–675
Arabic, Western Delta 1.589, 1.605
Arabic, Western Desert 1.320, 1.322, 2.2, 2.5
Arabic, Xabura → Arabic, Khabura
Arabic, Xaṯʿam 3.94
Arabic, Xatuniyye → Arabic, Khatuniya
Arabic, Yāfiʿ 4.750–751, 4.753–757
Arabic, Yemen Plateau 4.349
Arabic, Yemeni 1.314, 1.504, 2.257–258, 2.260, 2.263, 2.270, 2.388, 2.390, 2.462, 2.556, 3.63, 3.237, 3.413, 3.416, 3.637, 3.691, 3.694, 3.726, 3.729, 4.71, 4.296, 4.298, 4.303–305, 4.309–310, 4.312, 4.314, 4.373, 4.750–758
Arabic, Yemeni Jewish 2.534, 4.750
Arabic, Zabīd 3.675
Arabic, Zarzis 3.226
Arabic, Zawia 3.55
Arabic, Zʿīr 3.288–290, 3.292
Arabic, Zrigat 3.290
Arabic as a foreign language 3.398–399
Arabic as language of science 2.705
Arabic as lingua franca 4.375
Arabic as official language 3.316
Arabic as trade language 1.655
Arabic broadcasting 3.195
Arabic college 3.129
Arabic Dependency Treebank 1.460
Arabic in Algeria 1.260, 1.502, 1.590, 1.606, 1.642, 3.672
Arabic in Australia 3.313
Arabic in Bahrain 4.62, 1.269, 1.271, 1.439, 1.446, 1.501, 1.596, 1.610, 2.323, 2.609, 2.615, 3.12, 4.70, 4.323, 4.325, 4.631, 4.633
Arabic in Cameroon 4.708
Arabic in Chad 1.260, 3.315, 3.448, 4.198, 4.310, 4.708, 4.712, 4.714, 4.756
Arabic in Coptic script 1.495
Arabic in Egypt 1.444, 1.605, 2.272, 3.15–19, 3.672, 3.675
Arabic in Europe 2.665, 3.313
Arabic in India 4.595–603
Arabic in Indonesia 2.337–338
Arabic in Iran 2.407
Arabic in Iraq 2.555, 3.237, 3.525
Arabic in Israel 1.584, 2.437–441
Arabic in Jordan 1.260, 2.31, 2.494, 2.555, 3.525, 3.613, 3.726
Arabic in Kuwait 3.525
Arabic in Latin America 2.665, 3.1–6
Arabic in Lebanon 1.432, 1.445, 1.492, 1.502, 1.571, 1.584, 1.607, 1.642, 2.31, 2.278, 2.374, 2.494, 3.15–19, 3.237, 3.613, 3.672, 3.674, 3.702–703
Arabic in Libya 1.260, 1.571, 1.642, 2.456, 3.702
Arabic in Mali 3.135–141
Arabic in Malta 2.313–314, 3.672, 3.675, 3.737, 4.550
Arabic in Mauritania 1.260, 1.624, 1.642, 3.175, 3.702
Arabic in Morocco 1.260, 1.440, 1.584, 1.592, 1.626, 1.642, 3.15–19, 3.452, 3.672, 3.674, 3.702, 3.737, 4.62

Arabic in Nigeria 1.260, 3.449, 4.310, 4.708, 4.758–763
Arabic in Oman 2.272, 3.89, 3.544, 3.672, 3.674, 4.304, 4.751
Arabic in Pakistan 2.672, 3.506–512, 4.595–603
Arabic in Palestine 1.571, 1.583, 1.605, 1.607, 2.4, 2.31–32, 2.494, 3.8, 3.702, 3.726
Arabic in Saudi Arabia 2.205
Arabic in Senegal 4.179–185
Arabic in Sicily 4.215–219
Arabic in Somalia 4.272, 275–278
Arabic in Sudan 1.260, 1.611, 2.272, 2.454, 2.456, 2.494, 3.429, 3.525, 3.674, 4.310, 4.375–381
Arabic in Syria 1.407, 1.502, 1.571, 1.583–584, 1.592, 1.605, 1.607–608, 2.31, 2.204, 2.374, 2.532, 2.555, 3.588
Arabic in Tanzania 2.672
Arabic in the Emirates 2.609
Arabic in the Ottoman Empire 3.503–504
Arabic in the United States 2.684
Arabic in Tunisia 1.440, 1.571, 1.584, 1.592, 3.672, 4.550
Arabic in Turkey 4.577–583
Arabic in Uzbekistan 4.608–612
Arabic in Yemen 1.260, 1.502, 1.590, 1.606, 1.624–625, 2.204, 2.272, 2.494, 3.429, 3.585, 3.672–673, 3.693, 3.695–697, 3.702, 4.2
Arabic language processing 1.215
Arabic Language Resources 1.214
Arabic on the internet 3.316
Arabic portals 2.381
Arabic script 1.7, 1.371, 1.379
Arabic Teaching Institute for Foreigners 4.149
Arabicization 1.102, 1.619–620, 2.58, 2.406, 2.526, 2.634, 2.638, 2.641, 2.696, 3.142, 4.375, 4.408
arabico-malgache 3.126
Arabism 2.282–283, 2.285–286, 2.666
Arabistan 2.407, 2.572
ʿarabiyya 1.131–132, 1.153, 1.173–181, 1.397–400, 1.402, 1.405, 1.490, 1.527, 1.532, 1.534, 1.615, 1.617–618, 1.624–626, 2.60, 2.86–87, 2.89, 2.236, 2.264, 2.401, 2.403, 2.541, 2.624, 2.629, 2.631–632, 3.7, 3.90, 3.93, 3.344, 3.465, 3.499, 3.646, 4.7
Arabization 1.53, 1.55–56, 1.58, 1.97–98, 1.178, 1.269, 1.289, 1.403, 1.451, 1.495, 1.534, 1.589, 1.594, 1.611, 1.613, 2.199, 2.282, 2.375, 2.656, 2.659, 2.672, 2.694–695, 2.697–699, 2.701–703, 2.705, 2.707, 3.13, 3.25, 3.52–53, 3.57, 3.62, 3.169, 3.171, 3.175, 3.274, 3.296, 3.314–316, 3.347, 3.349–350, 3.435, 3.641, 4.460, 4.574
Arabization in Algeria 1.56
Arabization in Morocco 3.287
Arabization of education 2.697, 2.698 2.700
Arabized pronunciation 4.382–383
Arabkhona 4.613
Arabophonia 3.501–502, 3.504
Arabs 1.126–133, 1.178, 1.241, 1.271, 1.445, 1.490, 2.1, 2.60, 2.87–88, 2.212, 2.262–263, 2.499, 2.541, 2.543–544, 2.571–572, 2.653, 2.656, 3.343, 3.352, 3.464, 4.170
Arabs, Central Asian 4.612
Arabs, Indonesian 2.337
ARABTALK 1.210
ʿarabūn 1.132
ʿaraḍ 1.617
Arada 1.362
Aragon 1.98
Aragonese 4.572, 4.574
a-raising → ʾimāla
Arak Arabic 4.406
Arakkal 3.129
Aram Naharayim 1.178
Aramaeans 1.541, 1.617, 1.619, 4.170
Aramaic 1.84, 1.88, 1.95, 1.110, 1.126, 1.153, 1.166, 1.170, 1.178–179, 1.181, 1.401, 1.469, 1.493, 1.501, 1.536–538, 1.540–541, 1.554, 1.564, 1.638, 2.59, 2.66, 2.68, 2.84, 2.198–199, 2.211–212, 2.218–219, 2.262–263, 2.265, 2.286, 2.373, 2.388, 2.526–528, 2.531, 2.533, 2.560, 2.605, 2.607, 3.58–59, 3.218, 3.408–409,

3.411, 3.413–419, 3.465, 3.467–468, 3.553, 3.574, 3.584, 3.641, 4.170–171, 4.175, 4.296, 4.302–303, 4.313, 4.372–373, 4.403, 4.478–479, 4.482–483, 4.580
Aramaic, Ancient West 4.479
Aramaic, Biblical 1.85, 2.67, 3.411
Aramaic, Christian Palestinian 3.409
Aramaic, Galilean 1.541
Aramaic, Imperial 1.148, 2.23, 3.409–410, 4.479
Aramaic, Jewish 1.541, 2.67, 3.466
Aramaic, Maʿlūla 2.67
Aramaic, Middle 3.370, 3.409
Aramaic, Modern 4.303
Aramaic, Nabataean 3.410, 3.465–468, 3.697, 4.402, 4.479, 4.482–483
Aramaic, Neo- 2.67, 3.409, 4.176, 4.402
Aramaic, Neo West- 4.402–403
Aramaic, Northwest 4.406
Aramaic, Old 2.23, 2.67, 3.370, 3.372, 3.409, 3.411, 4.1
Aramaic, Palmyrene 3.409–410, 3.467
Aramaic, Proto- 3.412
Aramaic, Samaritan 3.409
Aramaic, Syro- 4.6
Aramaic, Western 3.372
Aramaic, Western Late 3.409
Aramaic inscriptions 1.401
Aramaic script 1.150, 153
Aramaic/Syriac 1.126, 3.6
arap 1.130
AraParse 1.460
ʿaraqat 3.308
ʾaraṯṯ 2.677
Arbāya 1.127, 1.129
Arberry, Arthur J. 4.429, 4.444
Arbīl 2.604, 2.608
Arbīl Arabic 2.415, 2.417–419, 2.421, 2.532, 2.534
arbiter of correctness 2.87
arbitrariness of the linguistic sign 3.188
archaism 3.189
Archangeli, Diana 2.38, 2.365
ʾArdabīlī, al- 2.403–404
area, dialect 1.604–605
area, isolated 1.590
area, language 4.371
area, lateral 1.590
area, linguistic 1.541, 1.590
area, transitional 1.604–605, 1.610
area studies 4.149
areal feature 1.37, 1.89
areal influence 4.372
areal norm 1.590
Arends, Jacques 3.634
Arent, Eugene R. 3.681
Arent, Russell 1.650
Areopolis 4.480
ʾArġawānī, Saläma al- 3.348
Argentina 2.671, 3.1–3, 3.5
Argobba 2.53–55, 2.67, 4.301, 4.305, 4.312
argot 2.468, 2.470, 4.251–252, 4.255–256, 4.258, 4.764
argot, Jewish 2.533–534
argument 1.182, 4.535
argument, external 1.182–184, 1.186, 3.685
argument, quasi-external 1.185
argument focus 2.113–115
argument hierarchy 4.648
argument structure 3.391, 4.489
argument text 1.651
argumentative writing 1.648
Arḥab Arabic 4.751
Ari 1.35
ʿārib 1.132
ʿāriba 1.130, 3.345
ʿāriḍ 1.309, 3.113
Arikan, Necmettin 4.579
ʿArīš, al- → El-Arish
aristocracy 2.8
Aristotle 2.66, 2.90–91, 2.185–186, 2.200–201, 2.426, 2.435, 2.442–443, 3.106, 3.109, 3.113–115, 3.176–178, 3.547, 3.549, 4.220, 4.235, 4.333, 4.459
Armant 1.299
Armbruster, Charles Hubert 3.437–438
Armenia 2.607
Armenian 1.222, 1.444, 1.527, 2.1, 2.59–60, 2.217, 2.500, 2.607, 2.704, 3.313, 3.349, 3.351, 4.262, 4.402, 4.578–579
Armenian, language loss in 2.500
Arnaldez, Roger 2.310, 4.448

Arnold, Gordon 1.515
Arnold, Werner 1.111–119, 1.584, 2.313, 3.370–373, 4.403, 4.580–582
Arnzen, Rudiger 4.543
Aronoff, Mark 2.684, 3.299, 4.339
Arpahan 1.115–116
arsis 3.209
ʾArslān, Šakīb 3.346
ʾArsūzī, Zakī al- 1.173, 3.346, 4.76
árthron 3.547
Arthur, Bradford 2.117, 2.121
article 1.488, 4.280
article, Arabic, in Swahili 4.384
article, assimilation of 1.187–188, 2.455, 4.96, 4.115
article, Coptic 1.498, 1.503
article, definite 1.50, 1.69, 1.75, 1.92, 1.94, 1.106, 1.109, 1.187–188, 1.190, 1.204, 1.226, 1.229, 1.239, 1.244, 1.248, 1.256, 1.262, 1.266, 1.279, 1.281–291, 1.304, 1.327, 1.340, 1.386, 1.392, 1.397, 1.407, 1.419, 1.473, 1.481, 1.491, 1.497, 1.525, 1.548–549, 1.560, 1.569, 1.579–583, 2.140, 2.161, 2.243, 2.247, 2.251, 2.272, 2.285, 2.343, 2.353, 2.419, 2.422, 2.483, 2.486, 2.511, 2.516, 2.527–528, 2.569, 2.612, 2.614, 2.669, 3.182, 3.192, 3.278, 3.293, 3.330, 3.365, 3.403, 3.415, 3.428–429, 3.453, 3.466, 3.471, 3.481, 3.484, 3.531, 3.579, 3.609, 3.729, 4.15–16, 4.22, 4.26, 4.116, 4.280–281, 4.312, 4.444, 4.478, 4.566, 4.616, 4.690, 4.742
article, definite *al-* 1.581–582, 2.547–548, 2.554
article, definite *am-* 1.491, 2.259, 3.692–693, 4.376, 4.751, 4.756
article, definite *-an* 2.260
article, definite *h-* 2.263
article, definite *han-* 4.478
article, definite *hn-* 1.491, 2.257, 2.259
article, definite *im-* 4.127
article, definite, in Northwest Semitic 3.415
article, definition of 3.244
article, deletion of definite 2.77
article, fossilized 1.30
article, French 1.418
article, Himyaritic 4.756
article, indefinite 1.31, 1.92, 1.106, 1.188–191, 1.205, 1.239, 1.248, 1.266, 1.283, 1.418, 1.497, 1.605, 2.17, 2.129, 2.243, 2.248, 2.575, 2.592, 2.614, 3.55, 3.278, 3.291, 3.429, 3.450, 3.484, 4.318, 4.444, 4.555, 4.616
article, indefinite *fad-* 2.419, 2.422, 2.575
article, indefinite *fadd-* 1.189, 1.239
article, indefinite *fat-* 1.189
article, indefinite *ši-* 1.189
article, indefinite *wāḥid* 1.189, 1.248, 1.418, 2.614
article, *l-* 1.491, 2.263
article, *n-* 1.491
article, postpositive 4.478
article, prepositive 4.478
article, reanalysis of 2.554, 2.669, 4.723, 4.725
articulation, compensatory 2.676
articulation, pharyngeal 4.669
articulation, place of 2.299
articulation, point of 1.653
articulation, secondary 2.460, 3.599, 3.610, 4.637, 4.669
articulation feature 3.626
articulator 3.594
articulatory-perceptual system 3.228
ʿarūḍ 2.238, 3.207, 3.209–210, 3.212, 3.214, 3.574, 4.201, 4.206
ʿarūḍī 4.104
Aruri, Naseer H. 2.499, 2.503, 4.77
Arusha Declaration 1.661
ʿArūżī, Niẓāmī 2.410
aryepiglottic folds 2.229, 4.666
aryepiglottic sphincter 4.666–668
arytenoids 4.666–667
ʿasā 1.557–558, 1.564, 3.235
ʾAʿšā, al- 2.496, 3.213, 4.202
ʿaṣabāt 2.363
ʿaṣabiyya 1.130
ʾAsad 2.88, 2.312, 3.698, 4.430
ʾAsad Arabic 2.494, 3.14, 3.91–92
Asafjahis 4.452
Asakir 3.340
Asala 1.362
Asanbayev, B. 2.557
ʿašara al-mubaššara, al- 2.47

ʾAšʿarī, al- 2.580, 3.122, 4.5
Asbaghi, Asya 3.579–584
Ascalon 2.46
ʾAsd Arabic 2.555
Asfour, John Mikhail 4.91
Asfour, Mohammed 3.28
Ashʿarites 2.303–306, 3.119, 3.164, 3.170, 4.474
Asher, R.E. 2.31, 3.128–135
Ashkelon 3.409
Ashkenazi Jews 2.461
Ashtiany, Julia 3.200–203
Ashton, E.O. 4.382, 4.385–386
Ashuri 2.414
ʿĀṣim 3.92, 4.5
Asim, Muterjim Ahmed 3.504
ʿAṣīr 4.125, 4.127
ʿAṣīr Arabic 1.504, 1.590, 2.556
ʿAskarī, ʾAbū Hilāl al- 1.428, 2.70, 2.304, 2.443, 2.496, 2.579, 2.582, 2.632, 3.117, 3.160, 3.162, 3.172, 3.323–324, 3.647, 3.649, 4.103, 4.473
ʾaṣl 1.50, 1.191–195, 1.309–310, 1.573–574, 2.179–180, 2.186, 2.309–310, 2.402–403, 2.447, 2.452, 2.543, 2.546, 2.624–625, 3.90, 3.119, 3.208, 3.308, 3.423, 4.13, 4.119, 4.287, 4.422, 4.443, 4.446–447, 4.560, 4.738
ʾaṣl al-waḍʿ 1.194
ʾaṣl aṯ-ṯanāyā 1.194
ʾaṣlī 2.239
ʾaṣlī, ḥarf 2.238, 2.447
ʾaṣliyya, istiʿāra 2.445
ʾasmāʾ → *ism*
ʾasmāʾ al-ḥusnā, al- 2.46, 3.719
ʾasmāʾ aš-širāf, al- 2.46
ʾasmāʾ as-sitta, al- 1.310, 2.234
Asmaa Dynasty 4.276
ʾAṣmaʿī, ʿAbd al-Malik ibn Qurayb al- 1.402, 1.627–628, 2.89, 2.496, 2.628, 3.36, 3.323, 4.443–444
ʾAsmar, Jirjis ʿĪsā al- 2.402
ʾasnada 2.435–436
aspect 1.50–51, 1.110, 1.195–201, 1.216–217, 1.508, 1.510, 1.523, 1.553, 1.568, 1.606, 2.91, 3.67, 3.99, 3.186, 3.247
aspect, continuous 1.95, 1.253, 1.611, 2.516, 2.524, 2.569, 3.536, 3.643, 4.457
aspect, durative 1.165, 1.282, 2.490, 3.221, 3.266, 4.622
aspect, eventive 4.491
aspect, experiential perfect 1.248
aspect, habitual 1.305, 1.330, 1.508, 2.514, 2.516, 2.524, 2.595
aspect, imperfective 1.195, 3.68, 3.99, 3.247, 4.454, 4.457
aspect, lexical 1.199
aspect, non-eventive 4.491
aspect, perfect 1.331, 3.186
aspect, perfective 2.171, 2.423, 4.456–457, 4.620, 4.697
aspect, progressive 3.186–187, 3.247, 3.266, 3.573, 4.569
aspect, punctual 4.733
aspect marker 4.114, 4.162
aspect marking 1.485
aspect marking in sign language 4.228–229
aspectual value, inner 1.219
aspectualizer 3.285
aspiration 1.26–27, 1.90, 1.135, 2.7, 2.397, 2.560, 3.597, 4.666
aspiration, final 1.262
ʿAsqalānī, Ibn Ḥajar al- → Ibn Ḥajar al-ʿAsqālanī
ʿaṣr, fuṣḥā al- 1.405
Assab 2.271
ʿAssāl, Fatḥiyya al- 1.599
assertion 2.144, 2.360, 4.503
assertive 3.248, 4.329
assessment 1.202
asseverative particle 1.201–204, 2.22
assignment, generic 4.685–686
assignment, individualized 4.685
assimilated word 2.448
assimilation 1.60–61, 1.75, 1.104, 1.187–188, 1.204–206, 1.225, 1.233, 1.244, 1.246, 1.262, 1.277, 1.301, 1.314–315, 1.325, 1.445, 1.492, 1.497, 1.548, 2.37, 2.242, 2.270, 2.280, 2.298–300, 2.311, 2.483, 2.494, 2.591, 2.612, 3.15, 3.204, 3.335, 3.481, 3.529, 4.109, 4.121, 4.244, 4.324, 4.549

assimilation, full 4.427
assimilation, mutual 1.206
assimilation, partial 1.204–206, 2.280, 4.427
assimilation, progressive 1.204–206, 2.299–300
assimilation, reciprocal 1.204
assimilation, regressive 1.204–205, 2.299–300, 4.688
assimilation, regressive vowel 2.591
assimilation, total 1.204–205
assimilation, voice 3.181
assimilation of minority groups 2.713
Association culturelle islamique du Cameroun 1.336
assonance 3.669
Assyrian 1.126–127, 2.59, 2.61, 2.64, 2.67, 2.84, 2.695, 3.350, 3.466, 4.147, 4.174, 4.402
Assyrian, Modern 2.61
Assyro-Babylonian 3.738
ʾAstarābāḏī, Raḍī d-Dīn al- 1.192, 2.81, 2.83, 2.91–94, 2.165, 2.181, 2.189, 2.291, 2.299, 2.359–360, 2.447, 2.450, 3.224–225, 3.548, 4.119, 4.430, 4.436, 4.443–445
Asterix 4.284
astronyms 2.25–26, 2.282, 2.455
ʿAšūr, Nuʿmān 1.601
ʾAswad ibn Yaʿfur, al- 4.200
Aswan 2.1, 3.435
Aswan Arabic 2.2, 2.5
aswan arti 2.1
Aswan High Dam 3.60
Asymmetric Rhythmic Theories 3.616, 3.620
asymmetricity 4.742
asymmetry 1.183, 1.418, 1.446
asymmetry, agreement 1.474.396–397, 4.654
asymmetry, subject/object 2.172
asymmetry in code-switching 1.419
asymmetry of codes 1.416
asyndetic 1.473, 1.488, 1.491, 1.552, 2.331, 2.333, 2.491, 4.191, 4.197
asyndetic construction 1.295, 1.397, 1.468–469, 2.224–225, 4.195, 4.198
Asyut 1.299, 1.589, 2.2
Asyut Arabic 1.586, 2.2, 2.5, 3.405
Atam, Aladji kouotto Malam 1.337
ʾĀṯārī, al- 3.309, 4.132, 4.134
Atatürk 1.135, 4.578
Atawneh, Ahmad 1.372, 2.29–35, 3.318
Atbara 2.559
atelic verb 1.52
Ateş, Ahmed 4.592
ʿaṭf 1.124, 4.410
ʿaṭf bayān 1.125, 4.221
ʿaṭf nasaq 4.221
Athanasius of Qūṣ 3.72
Athenaeus 2.66
Ati 1.362, 1.364, 2.91
Atia Arabic 4.716
Aṭiyya 1.361
Atkinson, Max 3.669
Atlas 1.293, 2.62
ʾaṭnāb 4.133
Atoui, H. 1.449
Atsiz, [Huseyn Nihal] 3.502
ʿAṭṭār, ʾAḥmad ʿAbd al-Ġafūr 3.14
ʿAṭṭār, Busti 3.37
ʿAṭṭār, Dāwūd al- 1.598
ʾaṯṯara 4.412, 4.415
attentional view 1.423
attention-maintaining 1.249
Attia, A. 1.632
attitude 1.415, 1.430, 2.31, 2.144, 2.504, 2.650
attitude, gender-based 2.657
attitude, negative 1.657–659
attitude, speaker's 1.631
attitude specification 2.148
attribute 1.16, 1.18, 1.20, 1.70, 4.219–220
attribute, negative 1.658
attributive 1.19
attributive adjective 1.13
attributive construction 1.645
attrition 2.194
attrition, first language 1.370
attrition, lexical 2.682–683
Audebert, Claude-France 2.303
auditory signal 3.205
Auer, Peter 2.320, 3.664
augment 1.120, 2.346–352, 2.354, 2.450
augment, nasal 2.347
augmentation 1.120

augmentative 1.638, 2.352
Aulia, Sheikh Nizam ad-Din 3.508
Aurangzeb Alamgir 2.328, 3.338, 3.508
Austin, John Langshaw 1.657, 2.80, 2.359–360, 3.570, 3.679, 4.328–329, 4.331–332
Australia, Arabic in 3.313
Austronesian languages 1.7, 2.472, 3.204, 4.372
automata 1.457
automatic language processing 1.206–216
automatic text processing 1.214
automaton, finite-state 1.209
autonomy of cognitive systems 2.691
autosegmental association 3.462
autosegmental morphology 4.93
autosegmental phonology 1.121
autosegmental representation 3.628
autosegmental theory 1.205
autosegmental tier 1.653, 4.95–96
Autosegmental-Metrical (AM) model 2.395–398
auxiliaries, acquisition of 1.377
auxiliary 1.51, 1.56, 1.64, 1.216–221, 1.229, 1.281–282, 1.396, 1.432, 1.488, 1.524, 1.553–554, 2.193–194, 2.221, 2.271, 2.273, 3.68, 3.187, 3.236–237, 3.285, 3.386, 3.654, 4.40, 4.162–163, 4.167, 4.456
auxiliary, narrative 3.368
auxiliary, tense 4.646
auxiliary verb 2.77, 4.306
Avishur, Yitzhak 2.527, 2.534
Avoid a Pronoun Principle 3.711
Avoid Clash 3.617
Avoid Lapse 3.617
avoidance strategy 2.162
a-Vorschlag 1.322
ʾaw 1.467–468
Aw Jamaʾ, ʿUmar ʿĪsā 4.277
ʾa-wa 2.144
ʿAwaḍ, Lūwīs [Louis] 1.599, 3.349
Awad, Maher 3.537, 4.362
ʿAwāḏil 4.750
ʿawālim 4.157
ʿAwāzim 3.326
Awdaġust 3.170
Awdal 4.276
Awdal-Harar Corridor 4.275
Awestian 3.581
Awfāt 2.52
Awjila 3.56
ʾAwlād ʿAlī 1.299, 1.528, 2.2
ʾAwlād Saʿīd Arabic 4.240
ʿAwlaqī 4.750
ʿAwlaqī Arabic 4.756
ʿAwra, ʾIbrāhīm al- 3.222
Awrangzeb → Aurangzeb Alamgir
ʾawṣala 4.235, 4.414
ʿAwwād, Bahāʾ 1.599
ʿAwwād, Louis → ʿAwaḍ, Lūwīs
Awwad, Mohammad Amin 3.386, 4.629
ʿAwwād, Mūrīs (Maurice) 1.603
ʾawwal 2.402, 4.443
ʾaxaḏa 2.193
ʾaxaff 3.99
Axārsah Arabic 3.401, 3.403, 4.238
ʾaxawāt ʾinna 2.354
ʾaxawāt kāna 2.549–551
ʾaxbara 2.435, 4.737
ʾAxfaš al-ʾAwsaṭ, ʾAbū l-Ḥasan Saʿīd ibn Masʿada al- 2.227, 2.294, 2.426, 2.551, 2.631, 3.104, 3.208, 3.210, 3.214, 3.548, 4.9, 4.12, 4.89, 4.119, 4.235, 4.329, 4.436
ʾāxir al-kalima 4.447
Axtarī 3.502
Axumite Kingdom 2.51–52, 2.56
Axvlediani, Vladimir Grigor’evič 2.238, 4.613–614, 4.616, 4.618–620
Ayaji Crowther, Bishop 4.758
Ayalon, Ami 4.76
Ayalon, David 3.23, 3.380
Ayari, Salah 3.78
ʾāyāt 1.8
ʿAyāydah Arabic 3.401, 4.238
ʿayb 3.173, 4.416
ʿAydarūs, Šarīf ʿAydarūs ibn Šarī ʿAlī 4.277
Ayed, Hela Ben 4.363
Aymara 4.374
ʾayna 1.309, 2.427
aynalī 4.134
Ayoub, Georgine 1.457, 2.83–90, 2.301, 2.434, 2.628–634, 3.384, 3.428, 3.432, 3.453, 4.395–396, 4.442–447

Ayoub, Millicent R. 4.469
Ayoubi, Kenneth K. 3.397
Ayuthia 2.334
ʾayy 2.80, 3.358, 3.653
ʾayyām al-ʿArab 1.402, 3.465
ʾayyun 2.387–388
Azamgarh 2.329
ʿAzāyza 2.2
ʿAzāzmih 3.360, 3.401
ʿAzāzmih Arabic 3.400
ʾAzd 3.692
ʾAzd Arabic 3.692, 3.697, 4.376
ʾAzd Sarāt 4.430
ʾAzd ʿUmān 2.212, 3.692
ʾAzd ʿUmān Arabic 3.94
ʾAzdī, al- 3.379–380
ʾAzd-Šanūʾa Arabic 3.89, 3.94
Azeez, Waheed O. 3.374
Azer 3.170
Azerbaijan 3.370
Azerbaijani 2.607, 3.350
Azeri 3.313, 3.370
Āzəx 1.87–89, 1.91–93, 1.95
Āzəx Arabic 4.407
Azhar University 1.382, 1.448, 1.451, 2.337, 4.21, 4.76
ʾAzharī, ʾAbū Manṣūr Muḥammad ibn ʾAḥmad al- 3.32–33
ʾAzharī, aš-Šayx Xālid al- 2.189
ʾAzharī, ʾIsmāʿīl ibn Muḥammad Ḥanīf al- 4.291
Azid, al- 1.361–362
Aziz, Y. 3.713, 3.715
Azoury, Negib 1.130
Azra, Azyumardi 2.335
Azraq Oasis 2.498, 2.502
ʿAzzām, ʿAbdallāh 4.78
Azzam, Rima 3.78
ʿAzzawi, Diya al- 4.204
Azzopardi-Alexander, Marie 3.144, 3.147, 3.157, 3.739–740
Azzoune, Hamid 1.458

B

Bā ʿAlwī ibn ʿAbdallāh al-ʿAllāma aṭ-Ṭāhir 2.335
Baalbaki, Munir 3.26–28
Baalbaki, Ramzi 1.191–194, 1.308–310, 2.80, 2.177–178, 2.180–181, 2.404, 4.7, 4.119–120, 4.284, 4.295, 4.414, 4.436–437, 4.446–447, 4.449
Baalbaki, Rohi 3.26–27
Baars, Bernard 4.270, 4.335
Baath 2.61
Baâziz 2.78
bāb 1.193
Bab al-Mandeb 1.655
Baba, Ahmad 3.136, 4.759
Baba, Mamoudou 1.333
Babatorun 1.114–116
Babault, Sophie 2.703
babbling 2.97, 4.50
Babel 3.345, 3.352
Babikian, Salem 4.76
Babtī, ʿAzīza 3.558
baby talk 1.339–343, 4.50, 4.59, 4.285–286, 4.321
baby talk → caretaker talk
Babylonian 2.212, 4.174, 4.479
Babylonian, Old 3.414, 3.738
Baccouche, Taieb 4.564, 4.571–577
Bachmann, Peter 4.543, 4.545
back consonant 4.670
background 1.485, 2.207–209, 4.499, 4.501, 4.728–729, 4.731, 4.733
background information 4.656
backgrounding 1.648
backing, vowel → vowel backing
backward talking 4.159
Bacri, N. 3.725
Bactria 3.518, 4.612
baʿḍ 1.284, 3.202, 4.43–45
Badʿ, al- 3.402
badāʿat al-lisān 2.362
badal 1.4, 1.123–126, 1.552, 1.559, 2.280–281, 2.546, 3.101, 3.650, 4.121, 4.428
badal al-ḥaraka 2.234
Badārah Arabic 4.238
Badāri Arabic 2.5
Badaun 2.326
badawī 1.273, 1.605, 2.212
Badawi, Elsaid Muhammad [as-Saʿīd Badawī] 1.405, 1.409, 1.419, 1.440, 1.472, 1.630–6321.666–670, 2.104, 2.157, 2.159, 2.162, 2.224–225, 2.296, 2.315, 2.317, 2.355, 2.390, 2.424, 2.452–453, 2.477–478, 3.14, 3.17, 3.27, 3.101, 3.168, 3.200–202, 3.239, 3.272, 3.423–424, 3.426–427, 3.456, 3.559, 3.561, 3.589, 3.683, 3.686, 3.700, 3.728–729, 3.740–741, 4.192–193, 4.284, 4.354, 4.357, 4.360–361, 4.419, 4.455, 4.530–531, 4.534, 4.558, 4.653–654, 4.657
Badawī, Muḥammad Muṣṭafā 1.600–601, 2.104
Baddredine, Belhassen 4.670
Badecker, William 2.691, 4.194
Badeir, Salih 2.465
Bader, Youssef 1.372
Badger, George P. 3.143
badīʿ 1.151, 2.305, 2.327, 2.446, 2.496, 3.121, 4.205, 4.210, 4.213
bādiya 1.128, 1.130
Badr ad-Dīn Ibn Mālik 4.119
Badran, Margot 2.645
Badri, Kamal 4.379

Baduel, Pierre-Robert 3.169
badw 1.130, 2.212
badwi → *badawī*
Baerman, Matthew 1.644
Bafaqih, Mohammad 4.698
Bafour 3.170
baġā 1.220
Baġdādī, ʿAbd al-Qādir ibn ʿUmar al- 3.647
Baġdādī, ʿAbdallāh al- 4.132
Baġdadi, Nadya 2.703
baggāra 1.363
Baggāra Arabic 3.62
Baggāra Belt 4.709
Baghdad 1.26, 1.86, 1.222, 1.271–272, 1.403, 1.443–445, 2.220, 2.406, 2.414, 2.532
Baghdad, Christians of 1.222
Baghdad, Jews of 1.222
Baghdad, school of 2.185
Baghdadi Arabic 1.222–231, 1.271–272, 1.442, 1.444–445, 2.8, 2.414–424, 2.532, 2.572, 2.577, 3.9, 3.84, 3.237, 3.525, 3.569, 3.666–668, 3.672, 3.674, 3.703, 4.62, 4.117, 4.194, 4.304, 4.325
Baghdadi Arabic, Christian 1.222–232, 1.273, 1.537, 1.607, 1.610, 2.276, 2.313, 2.532, 4.633
Baghdadi Arabic, Jewish 1.23–24, 1.231–241, 1.625, 2.414, 4.664
Baghdadi Arabic, Muslim 1.23, 1.25, 1.79, 1.232, 1.271, 1.595, 1.612, 1.625–626, 1.642, 2.276, 2.611, 2.617, 3.495, 3.585, 4.128, 4.633
Bāgirmi 1.363, 4.715
Bāgirmi Arabic 4.708, 4.715
Bahadur, Nawab Aziz Jang 2.325
Bahamani Kingdom 4.452
Bahar, M.T. 2.410
Baḥariyya, il- 2.2
Baḥariyya Arabic 1.270, 1.586, 1.610, 2.6, 4.550
Baḥārnah 1.241, 1.271, 1.446, 1.610, 2.212–213, 3.478, 4.324, 4.633
Bahasa Indonesia 1.135, 1.404
Bahasa Indonesia → Indonesian
Bahasa Melayu → Malay
Bahat, Yaaqov 2.534
Bahdinan 2.607
Bahdînanî 2.604
Bāhila Arabic 3.94
Bahlawān 2.217, 2.220
Bahloul, Maher 1.43–48, 1.506–511, 2.174, 3.81, 3.231, 4.193–194, 4.365, 4.395–398, 4.648–650, 4.657, 4.743
Bahloul, Raja Ballek 4.743
Bahnasāwī, Ḥusām al-Bahiyy ʿAlī al- 3.38
baḥr 3.210
Bahr al-Ghazal 1.333, 1.362, 2.517, 3.61
Bahrāʾ Arabic 3.94
Bahrain 1.241, 1.273, 1.446, 2.105, 2.107, 2.212–215, 2.696, 4.468, 4.766
Bahrain, Arabic in 4.62, 1.269, 1.271, 1.439, 1.446, 1.501, 1.596, 1.610, 2.323, 2.609, 2.615, 3.12, 4.70, 4.323, 4.325, 4.631, 4.633
Bahrain, youth speech in 4.766–767
Bahraini Arabic 1.241–255, 1.260, 2.494, 2.609, 2.615–619, 3.478, 3.729, 4.64, 4.656
Bahraini Arabic, Shiʿite 4.127–128
Bahrick, Harry P. 2.683
baʿīd, maʿnā 3.538
Bailey, Carolyn Madden 4.140
Bailey, Charles-James N. 3.642
Bailey, Clinton 3.360–361, 3.400–401, 4.238
Bakalla, Muhammad Hasan 1.615, 2.96, 2.459–461, 2.468, 3.179–180, 3.334–336, 4.1, 4.157, 4.421–424
Baker, Carl L. 3.653
Baker, Mark 3.706, 4.488–489, 4.491, 4.648
Baker, Mona 1.436–437, 1.519, 3.713, 4.498
Bakhtin, Mikhail M. 1.636
Bakir, Murtada J. 1.507, 3.382–384, 3.387, 3.683, 4.317, 4.361, 4.395, 4.520, 4.524–525, 4.527, 4.647, 4.653–654, 4.657, 4.719, 4.722, 4.746
Bakker, Dik 3.66
Bakr 4.304
Bakr ibn Wāʾil 2.212, 2.223, 2.555, 3.94
Bakrī, ʾAbū ʿUbayd al- 1.256, 1.361, 1.518, 3.59, 3.373, 3.642, 4.179, 4.279

Bakrī, Muḥammad Tawfīq al- 2.634, 2.672
Bakry, Hasbullah 2.337
Baku 4.425
bal 1.468, 3.356–357, 4.35
Bal Ḥāriṯ ibn Kaʿb 1.641, 3.89
Bal Qarn 4.127
Balāḏurī, al- 3.7
balāġa 1.428–429, 1.471, 2.88, 2.304–305, 2.579, 4.472
balāġī 2.358
balāġiyyūn 2.441
balah 3.500
balamo-group 2.220
Balaqaev, M. 2.557
Balāṭ Arabic 1.585, 2.2
Balawiy Arabic 4.677
Balban, Ghias Uddin 3.508
Balboni, Luigi 2.457
Baldauf, Ingeborg 2.558
Baldi, Sergio 1.256–257, 1.660, 2.251, 2.253, 2.552–555, 4.279–282, 4.381–387
Baldissera, Eros 3.23
Baleares 1.98
Balegh, Hedi 3.349, 4.563
Balga 2.506
Balgāwī Arabic 2.506
Bālī 3.691, 4.476
Balig, Al-Hadi al- 1.603
Balkan linguistic area 4.371
Balkh Arabic 4.311
Ball, Martin J. 3.336
Balochi 1.133, 2.211, 3.350, 3.506
Balochis 3.478
Balqāʾ 3.405
Balqāʾ Arabic 2.506
Balqar 4.263
Balṭīm 2.4
Balṭīm Arabic 2.2–3
Baltin, Mark 2.18
Balyāna Arabic 1.572
Bamako 3.137–138
Bamba Mbacke, Shaykh Amadu 4.180–182, 4.184
Bambara 1.255–257, 2.669
Bamford, John 2.675
Banadir 4.276
Banadir Coast 4.275
Banda Aceh 1.7
Bandanījī, al- 3.32
Bangala 2.525
Bangilani, Kiyahi Haji Misbakh ibn Zain Mustafa 2.339
Bangla 3.574
Bangla Academy 1.288
Bangla Musalmani 1.287–288
Bangladesh 1.257–259, 1.288, 2.211, 2.326, 3.511, 4.508
Bani ʿAṭīye 3.401
Bani ʿAṭīye Arabic 3.401–402, 3.407, 4.240
Bani Bišr Arabic 4.127
Bani Ḥamīda Arabic 2.503
Bani Ḥasan → Beni-Hasan
Bani Kaʿb 2.572
Bani Lām 2.571
Bani Maṭar 4.305
Bani Sāsān 2.469–470, 4.255
Banī Ṣaxar Arabic 1.406
Bani Sēd 1.362
Bani Šihr Arabic 4.125, 4.127
Bani Swayf → Beni Swef
Bani Ṭōr Arabic 4.126
Bani Ṭuruf 2.571
Bani ʿUqayl 1.260
Bani Wāʾil 1.362
Bani Wāṣil Arabic 4.240
Bani Yāsīn, Raslan 2.501, 3.402–405
Banki 4.715
Bankipore 2.329
Banna 1.35
Bannā, Ḥasan al- 4.77
bannāʾī 2.600
Banskira, Degif Petros 3.305
Banten 2.335
Bantu class-prefix system 4.385
Bantu languages 1.357, 1.448, 1.663–665, 2.589, 2.668, 2.670–671, 3.126, 4.381–382, 4.385
Bantu plural 2.670
Banū Dabbāb 3.53
Banū Ġaṭīf 3.89
Banū Ḥarb 3.693
Banū Ḥassān 2.240, 3.53, 3.170–171
Banū Hilāl 1.53, 1.59, 1.260, 1.289, 1.361–362, 1.595, 1.609, 2.240,

3.52–53, 3.142, 3.170, 3.288, 3.691, 4.125, 4.548
Banū ʿIjl 4.36
Banū Maʿqil 1.53, 3.170, 3.288
Banū Nahd 3.14
Banū Qays 3.454
Banū Sayf 1.361
Banū Sulaym 1.53, 1.260, 1.289, 1.609, 2.240, 3.52–53, 3.170, 3.288, 3.691, 4.548
Banū Tamīm 2.223, 3.454
Banū Zīd 1.655
Banū Zugba 3.53
Banyās 1.388
Banyo 1.337
Bāqillānī, ʾAbū Bakr al- 2.303–306
Barakat 3.56
Barakat, Halim 2.697
Barakat, Salim 4.213
Baranov, Charlampij Karpovič 3.23
Baranquilla 3.2
Bar-Asher, Meir 2.533–534
Bar-Asher, Shalom 3.274
Barāwī, Sheikh ʿUways al- 4.277
Barbacoan languages 4.374
Barbar 4.377
Barbar, Hector Azar 3.396
Barbarchi 3.375
Barbot, Michel 2.471, 4.157, 4.592
Bardera 4.276
Barfield, Thomas 4.612
bargaining 1.650
Bargery, G.P. 2.251
Bargoens 2.217
Bar-Hillel, James 1.565
Barhum, Musa 2.111
Bari 1.519, 2.519–520, 2.525, 2.597, 3.643, 4.375
Bariou, Jacques 3.173
Barīs Arabic 2.2, 4.196
Barkat-Defradas, Melissa 3.724–726, 4.669–674, 4.678–682, 4.672, 4.674
bar-Lev, Zev 4.288
Barlow, Michael 1.458
barma 1.363
Barqa Salam 1.361
Barqūq, Sultan 4.708
Barrada, F. 3.194
barrier 2.172
barrier relic 1.589
Barriers Model 3.86
Barrs, Andrew 2.370
Bartalī, Muḥammad ibn ʾAbī Bakr aṣ-Ṣiddīq al- 3.172
Barth, Fredrik 2.58
Barth, Jacob 1.577–578, 1.638, 2.23–24, 3.417, 3.446, 4.176
Barth, Johann 3.423
Barthélemy, Adrien 2.313, 4.403, 4.408
Barthélemy, Abbé Jean-Jacques 3.141
Barth-Ginsberg Law 3.417
Barth's Law 1.491, 4.431
Bartoli, Matteo 1.590
Barton, Ellen 2.19
Barūdī, Maḥmūd Sāmī al- 4.87, 4.90
Barwise, Jon 4.15
Barzilai, Gad 2.439
Bāšā, Ḥasan al- 3.718
Basaa 1.333
Bašandi Arabic 2.2, 4.286
base 1.574, 3.423, 4.50
base, morphological 1.577
base, nominal 1.575, 1.578
base, richness of 3.491
base, semantic 1.574
base component 4.521–522, 4.741
base generation 4.525
base node 1.47–48
base structure 4.521
Basheer, Vaikom Muhammad 3.131
Bashkir 2.557
basic variety 1.485
basilectal 1.525
basīṭ 3.211, 3.214
Baskinta Arabic 1.270, 4.656
basmala 2.43
Basque 1.36
Basque Country 2.439
Basra 1.618, 2.406, 2.572
Basra Arabic 2.312, 2.415–416, 2.421, 2.631, 3.525
Basran school 1.67–74, 1.310, 2.185, 2.291, 2.298, 2.355, 2.427, 2.453, 3.91, 3.301, 3.104–105, 3.108, 3.165, 3.424, 3.548, 3.550, 3.649, 4.8–9, 4.12, 4.120, 4.235–236, 4.438–439, 4.443, 4.455, 4.527
Baššār ibn Burd 1.402, 4.204–205

Bassiouney, Reem 1.419, 3.8–13, 4.57–60
Bastani, Tanus Jorge 3.4
Bataʿ Dam 2.260
Baṭāʾiḥī, Maʾmūn al- 2.41
Bataineh, Rula F. 3.681
Baṭalyūsī, ʾAbū ʿAbdallāh Ibn Sīd al- 2.291–292, 2.356, 3.33, 3.177, 4.132, 4.740
Batavia 2.335
Bat-El, Outi 3.205, 3.305–206, 3.494
Bates, Daniel G. 2.60
Bates, Elizabeth 2.686, 2.689–690
Bateson, Mary Catherine 2.224, 3.499–500, 3.728–729
Baʿth Party 4.76
Batha 1.360, 1.362
Baṭḥari 3.350, 4.297–298, 4.301
Batibo, Herman M. 4.382
bāṭin 2.360
Bāṭina 3.478
Bāṭina Arabic 3.479, 3.481, 3.484–485, 3.489
Bāṭina Coast 2.211
Batman Arabic 2.608
Battenburg, John 2.710
Battersby, Harold K. 4.585
Battisti, Carlo 2.456
Baudelaire, Charles 4.213
Bauer, Gertrud 2.235, 3.72
Bauer, Hans 3.409–410
Bauer, Leonhard 2.481, 2.502, 3.403
Bauer, Thomas 3.119
Bausani, Alessandro 1.628, 2.340, 4.433–435
Bawīṭi Arabic 2.2
Bāwiyah 2.571
Bawwāb, al- 4.94
Baxʿa 1.178
bayān 1.4, 2.85, 2.305, 3.136, 3.160, 4.209, 4.221
Bayani, Manijeh 3.337
bayān-i ḥaraka 2.408
Bayḏ̣āʾ, al- 4.751
Bayḏ̣āʾ Arabic 4.751
Bayḍāni 1.655
Bayḍāwī, ʿAbdallāh ibn ʿUmar al- 4.430
Bayḍūn, Aḥmad 2.703
Baye Fall 4.180
bayt 3.208, 3.210
Bayt al-ḥikma 1.403
Bayt Dafaʿ 2.260
bayyana 4.436
Bayyātī, ʿAbd al-Wahhāb al- 4.77
Bayyūmī, H. 2.167
Bdūl Arabic 3.402, 3.404–405
Beaudouin, Mondry 1.536
Beaugrande, Robert de 1.430, 4.475
Beaujard, Philippe 3.127
Bechar 1.56
Beck, Edmund 4.8
Beck, Maria-Luise 4.144–145
Beckman, Mary 2.395–396
Bedarfsentlehnungen 1.410
Bedawi 3.350
bədd 1.553
Bedeutungslaute 1.573
Bedja → Beja
Bedouin 1.53–57, 1.59, 1.84, 1.111, 1.127–128, 1.130, 1.190, 1.217, 1.222, 1.241, 1.259, 1.299–300, 1.324, 1.402, 1.440, 1.530, 1.589, 1.592, 1.605, 1.614–1.615, 2.1, 2.3, 2.87, 2.176, 2.501, 2.628, 3.171, 3.360, 3.402
Bedouin, North African 2.3
Bedouin, Sinai 4.237–238
Bedouin Arabic 1.259–274, 1.406, 1.443, 1.544, 1.569–572, 1.589, 2.587, 2.632, 3.205, 3.265, 3.360, 3.585, 3.609, 3.621–622, 4.2–3, 4.124
Bedouin Arabic, Jordanian 4.325
Bedouin Arabic, Negev 3.360, 3.531, 3.536–537, 3.569, 3.609, 3.613–614, 4.351, 4.663
Bedouin Arabic, North Israel 4.649
Bedouin Arabic, Sahara 1.642, 3.672
Bedouin Arabic, Syro-Mesopotamian 1.607, 4.70
Bedouin features 3.10
Bedouin influence 2.4–5
Bedouin informants 1.400, 1.402, 2.176, 2.184, 2.401, 2.403, 2.542–543, 2.628, 3.90
Bedouin lexicon 1.308, 4.380
Bedouin migration 3.94
Bedouin narrative style 1.266
Bedouin settlement 1.595

Bedouinization 1.269–274, 1.605, 1.610, 1.612, 2.414
Bedwell, William 1.166, 1.169, 1.172
Beebe, Leslie M. 4.320–321
Beene, Wayne 2.415
Beersheba 3.400–401
Beesley, Kenneth R. 4.94
Beeston, Alfred Felix Landon 1.491, 2.224, 2.256, 2.258, 2.260, 2.296–297, 2.542, 3.227, 3.446, 3.467, 3.470, 4.104, 4.192, 4.295–296, 4.307, 4.310, 4.479, 4.484–485
Beg, Muhammad Abdul Jabbar 2.341, 2.344
begadkefat consonants 1.536, 1.538, 1.540, 3.372
begadkefat shift 1.179–181
Beghelli, Filippo 2.369
begūl Arabic 3.190
behavior, nonverbal 2.205–206
Behaviorism 4.139
Behnstedt, Peter 1.190, 1.289–293, 1.299, 1.322, 1.501–505, 1.571–572, 1.583–593, 1.595, 1.604–605, 1.607–612, 2.2, 2.4, 2.6–7, 2.151, 2.260, 2.270–273, 2.388–389, 2.456, 2.494, 3.287, 3.289, 3.291, 3.402, 3.404–405, 3.429, 3.584, 3.606, 4.70–71, 4.107, 4.238, 4.303–305, 4.402–409, 4.590, 4.628, 4.633, 4.635, 4.661, 4.663, 4.676, 4.750–751, 4.753–756
Bəḥzāni Arabic 1.541, 1.571, 1.607
Beijing 1.382
Beirut 1.175, 1.274, 1.603, 2.220, 2.505–506
Beirut Arabic 1.269, 1.274–286, 1.295, 1.323, 1.538, 1.607–608, 2.231, 2.314, 2.509, 3.11, 3.674–675, 4.349, 4.630, 4.632, 4.678, 4.681
Beirut Arabic, Sunni 1.283
Beit-el-Ras 1.661
Beja 1.35, 1.38–39, 1.656, 2.1–2, 2.560, 3.420, 3.592, 4.375–377, 4.379
Bejāya 1.54, 2.712, 2.715
Beland, Renee 2.691–692, 4.96
Belazi, Hedi M. 3.318
Beletti, Andrea 4.488
Belgium 2.439
Belgium, Arabic in 2.73
Belkacemi, Manchester 4.67
Belkaid, Yamina 2.396
Bell, Allan 4.322, 4.628
Bell, Herman 3.435
Bellamy, James A. 3.469
Beller, Charley 4.647, 4.657
Bellermann, J.J. 3.141
Bello, Muhammad 4.759
Belnap, Kirk R. 1.630, 1.636, 2.162, 3.394–400, 3.430, 3.446, 3.500, 4.318, 4.356
Belot, Jean-Baptiste 3.26
Belova, Anna Gr. 2.257–260, 4.300–315
Belyayeva, Dina 3.81, 3.86–87, 4.317, 4.653, 4.655
Ben Ahmed, Mohamad 1.458
Ben Ali, Damir 1.663
Ben Busta, ez-Zubir 3.295
Ben Cheneb Mohammed 2.537, 4.423
Ben Jelloun, Tahar 2.703, 3.294
Benabbou, Mostafa 1.584, 3.287, 3.289
Benali, Abdelkader 2.78
Benali, Ismail 3.725
Ben-Amotz, Dan 2.462
Bencheikh, Djamaleddin 2.497
Bencheneb, Rachid 2.470
Bender, M. Lionel 3.435
Bendjaballah, Sabrina 1.119–123
benediction 3.139
benefactive 1.40, 1.43, 1.344, 1.346
Bengal 1.286, 1.288
Bengali 1.257–259, 1.286–289, 2.333, 2.668, 2.674, 3.313, 3.511, 3.574
Benghazi Arabic 3.55–56
Bengio, Ofra 3.664, 3.669
Benhallam, Abderrafi 3.725–726, 4.683
Beni Shangul 2.51
Beni Swef Arabic 1.322, 1.608, 1.611, 2.2, 2.5, 2.494
Beni-Hasan Arabic 2.503, 3.622, 3.707, 4.349, 4.351
Benin 4.758
Benjelloun, Said 2.74, 3.274
Benkirane, Thami 2.116, 2.395–396, 3.724–728, 4.674–676, 4.678, 4.682–683

Benmamoun, Elabbas 1.46–47, 1.349–350, 1.352, 1.411, 1.477–482, 1.507–508, 1.510, 1.582, 2.174, 2.369–370, 2.645, 3.68–70, 3.84, 3.232–233, 3.356–359, 3.382, 3.384, 3.429, 3.432, 3.446, 3.653, 3.655–656, 3.684, 3.707, 3.709, 4.15, 4.82, 4.194, 4.339, 4.365, 4.391–402, 4.648–650, 4.654–656, 4.658, 4.745
Benrabah, Mohamed 2.695, 2.697
Ben-Rafael, Eliezer 2.465, 2.661
Bentahila, Abdelali 1.372, 1.418, 2.653, 2.659–660, 2.695, 3.294, 3.311–320
Bentin, Shlomo 2.692
Bentolila, Yaakov 2.661
Benveniste, Emile 1.565–566, 2.82, 4.328
Ben-Yehuda, Eliezer 2.462–463, 2.666
Benzakour, Fouzia 2.122
Ben-Zeev, Y. 2.481
Berber 1.35–39, 1.53–54, 1.56, 1.58, 1.98, 1.102, 1.106, 1.110, 1.133, 1.189, 1.260, 1.289–299, 1.370, 1.482–483, 1.487, 1.501, 1.527, 1.534, 1.619, 1.652, 2.1, 2.26, 2.59–60, 2.62–64, 2.66–68, 2.73, 2.160, 2.251–252, 2.265, 2.285, 2.374, 2.470, 2.526, 2.621, 2.648, 2.659, 2.673, 2.695, 2.697–698, 3.52–53, 3.56, 3.142, 3.170–171, 3.189, 3.273, 3.287–288, 3.294, 3.312, 3.314, 3.317, 3.346, 3.350–351, 3.446, 3.491, 3.493, 3.592, 3.642, 4.190, 4.279, 4.371, 4.565, 4.572
Berber → Tamazight, Amazigh
Berber, Agadir 2.714
Berber, Chaouia 1.290–291, 2.714
Berber, Chefchaouen 1.294
Berber, Chleuh → Berber, Tachelhit
Berber, determiner in 1.296
Berber, feminine marker in 1.291, 1.296
Berber, Figuig 1.289–290, 1.292–293
Berber, Imdlawn Tashilhiyt 3.493
Berber, Kabylian 1.38, 1.56, 1.122, 1.289, 1.291, 1.370, 2.711–712, 2.714
Berber, Nador 2.714
Berber, Nafusi → Berber, Nefusa
Berber, Nefusa 1.290, 4.572
Berber, recognition of 1.534
Berber, Rif 1.290–291, 1.293, 1.571, 1.590, 2.711–712, 2.714, 3.273, 3.287
Berber, Shawiya → Berber, Chaouia
Berber, Siwa 1.290, 2.1, 2.5, 3.313
Berber, Souss 2.711–712, 2.714
Berber, Tachelhit 1.289–294, 2.712, 3.274, 3.287, 3.494
Berber, Tafilalt 1.298
Berber, Tamazight 1.290–293, 2.673, 3.274, 3.287
Berber, Tarifit → Berber, Rif
Berber, Tashelhit → Berber, Tachelhit
Berber, teaching of 1.534, 3.294
Berber, Tuareg 1.35, 1.289, 2.709
Berber, Zenaga 3.169–171
Berber dialects 1.584
Berber in education 2.704
Berber in Europe 2.712
Berber in school 3.350–351
Berber in Tunisia 4.572
Berber influence in Arabic 3.279
Berber Manifesto 2.63
Berber nationhood 2.62
Berbers, Lemtuna 3.169
Berchem, Max van 2.42, 2.45
Berent, Iris 3.305
Berg, L.W.C. van den 2.337
Berg, Thomas 3.628, 4.268, 4.337
Berger, Anne-Emmanuelle 1.534
Bergman, Elizabeth M. 2.136–137, 2.468–472, 2.560, 3.499–501
Bergstrasser, Gotthelf 1.583, 1.605, 1.608, 2.506, 3.402, 4.543, 4.681
Bergter, Annette 2.354–355
Bʿērī Arabic 1.299–308, 2.2, 2.5, 2.151–153
Beriberi 3.373
Berjaoui, Nasser 2.471, 3.500, 4.159
Berkley, Deborah Milam 3.627
Berman, Ruth A. 1.485, 2.102, 3.713
Bernards, Monique 2.224–228, 2.424–429, 2.451–454, 3.550, 3.690, 4.8
Bernhardt, Barbara H. 2.676
Bernier, Francis 3.508
Bernoulli Effect 4.667
Bernstein, Basil 4.58
Bernstein, Judy 4.743
Berque, Jacques 2.705

Berry, Randal K. 4.517
Berthier, Anne 4.132
Bertolini, Umberta 2.689
Bertoncini, Elena 4.382
Beška, Emanual 1.458
Bettini, Lidia 1.626–629, 2.496–498, 3.320–325, 4.750
Beys 4.573
beytol 3.171
Beziehungslaute 1.573
Bhat, D.N. 3.524
Bhatia, Tej K. 2.684
Bhutto, Z.A. 3.511
bi- 1.72, 1.217, 1.266, 1.330, 1.596, 1.607–608, 1.611–612, 2.22, 2.316, 2.377, 2.568–569, 2.595, 4.531
Bialystok, Ellen 4.141
Biba Arabic 2.5
Biber, D. 1.483, 2.384
Bible, Antwerp Polyglot 1.169
Bible, Arabic translation of 2.527
Bible, Hebrew 2.526
Bible, study of 1.166–167
Bible, Swahili translation of 1.663
Bible, translation of 1.287, 1.603
Bible, Yoruba 4.760
Bibliander 1.170
Bickerton, Derek 2.321
biconsonantal → biradical
biculturalism 2.441
Bīḏ̣ān 2.240, 3.60
bidialectalism 1.440, 1.444, 3.527
Bidwell, R. 3.314
Biʿēṛāt, il- 1.299
Bierwisch, Manfred 4.269
Biesterfeldt, Hans 4.542, 4.544
biḫārī 4.131–132
Biḥēra Arabic 1.589, 2.1–2
bi-imperfect → imperfect, *bi-*
bi-lā kayfa 2.580, 3.122
bilabial 3.597
bilabial, implosive 4.613
Bilād aš-Šām 2.505, 3.6
Bilād Tukrūr 3.373
Bilge, Kilisli Rifat 3.501, 3.503
Bilgrāmī, Ġulām ʿAlī Āzād 2.328
bilingual 1.369, 1.371–372, 1.417, 1.419, 1.482–484, 2.77, 2.97, 2.131
bilingual, Arabic 1.370
bilingual, balanced 1.368, 2.652
bilingual, compact 1.368
bilingual, coordinated 1.368
bilingual, dominant 1.368
bilingual, primary 1.368
bilingual, secondary 1.368
bilingual, subordinate 1.368
bilingual community 1.415
bilingual development 2.76
bilingual education 1.370
bilingualism 1.56, 1.97–99, 1.100, 1.368, 1.370, 1.414, 1.440, 1.629, 2.31, 2.441, 2.671, 2.703, 2.710, 3.294, 3.314, 3.318, 4.371
bilingualism, Algerian Arabic/French 2.692
bilingualism, Amazigh/Arabic 2.710–711, 2.715
bilingualism, Arabic/Aramaic 1.178
bilingualism, Arabic/Berber 2.708, 2.710, 4.572
bilingualism, Arabic/Dutch 1.418–419
bilingualism, Arabic/English 1.376, 2.661, 2.685
bilingualism, Arabic/French 1.56, 1.534, 2.660–661, 2.696, 3.175, 3.349, 4.574–575
bilingualism, Arabic/French Lebanese 2.691
bilingualism, Arabic/Greek 4.215–218
bilingualism, Arabic/Hebrew 2.438, 3.313
bilingualism, Arabic/Ivrit 2.464, 2.466
bilingualism, Arabic/Latin 4.572
bilingualism, Arabic/Malay 2.340
bilingualism, Arabic/Nubian 3.435
bilingualism, Arabic/Persian 2.571
bilingualism, Arabic/Portuguese 2.665, 2.671, 3.2–5
bilingualism, Arabic/Romance 2.288
bilingualism, Arabic/Spanish 2.665, 2.671, 3.2–3
bilingualism, Arabic/Swedish 2.679
bilingualism, Arabic/Tajik 4.614
bilingualism, Arabic/Turkish 4.573
bilingualism, Arabic/Uzbek 4.613
bilingualism, Aramaic/Arabic 1.180

bilingualism, Berber/Arabic 3.59–60, 3.317, 4.369
bilingualism, child 1.368–376, 2.97
bilingualism, Domari/Arabic 2.221
bilingualism, early 1.368
bilingualism, French/Chadian Arabic 1.365
bilingualism, French/English 2.684
bilingualism, Greek/Arabic 3.516
bilingualism, Hausa/Arabic 2.255, 2.671
bilingualism, Hungarian/English 2.684
bilingualism, Kurdish/Arabic 4.580
bilingualism, late 1.368
bilingualism, Latin/Arabic 4.216
bilingualism, Lebanese Arabic/English 1.371
bilingualism, Malay/Thai 4.477
bilingualism, Maltese/English 3.143–144
bilingualism, Maltese/Italian 3.143
bilingualism, Moroccan Arabic/Dutch 1.371, 2.685–686
bilingualism, Nigerian Arabic/Hausa 4.716
bilingualism, Palestinian Arabic/English 1.371
bilingualism, Persian/Arabic 3.574
bilingualism, Romance/Arabic 2.287, 4.216
bilingualism, Sindhi/Arabic 2.325, 2.671
bilingualism, South Arabian/Arabic 2.671, 3.60
bilingualism, Uzbek/Tajik 4.608
bilingualism in al-Andalus 2.671, 3.59
bilingualism in Canada 2.652
bilinguals 1.420, 1.634
biliteral → biradical
Biliy Arabic 3.400–401, 3.405, 4.238, 4.677
Biltine 1.360
Bimbashi Arabic 2.517–518
bimoraic 3.612–613
bimorphemic question word 2.388–389
Bin Laden → ʾUsāma bin Lādin
bināʾ 1.191, 1.308–310, 2.401–404, 2.479, 2.542, 3.429, 4.432, 4.442
bināʾ li-l-majhūl 1.309
binarity 4.742
binary feature 3.229
binder 4.393
Binder, Friedrich 2.501
Binder, Leonard 4.77
binding 1.310–313, 1.413, 2.369–370, 3.383–384, 4.81
Binding Principle 1.312, 377
Binding Theory 1.310–311, 2.170
Bingol 2.607
binomial 3.631–632
Binxet 2.607
binya 2.448, 4.22
binyanim 3.225, 3.626
biological gender 2.159
Biondi Assali, Estela 3.3
bioprogram 4.373
bipolarity 1.628
bipolarity in address forms 1.550
bipolarity in kinship terms 4.469
Biqāʾ Valley 1.126, 1.178
Bir Biʿērī Arabic 2.2
Bir Zeit Arabic 1.625, 3.430, 3.684, 3.686, 4.317, 4.656
biradical 1.80, 1.491, 3.610
biradical noun 1.80, 1.105
biradical root 1.81, 4.51
biradical word 1.81
biradicalism 1.312–316, 2.168, 2.450, 3.45–46, 4.93, 4.96, 4.254–255, 4.288
biradicalism, Semitic 4.174
bird names 2.27, 4.285
bird sounds 4.283, 4.285
Bird, Charles S. 2.202
Birenbaum, Menucha 2.661
Birgid (Birgit) 1.38, 3.435
Birkāwī, Muḥammad ʾAmīn al- 1.380
Birkeland, Harris 1.616, 1.641, 2.404, 3.454, 3.564, 3.726, 4.558
Birnbaum, Eleazar 3.501, 3.503
Bīrūnī, ʾAbū r-Rayḥān al- 3.579
biʾsa 1.73–74, 1.564, 2.80, 3.326, 4.123
Bišāri 2.1
Bishai, Wilson B. 1.630, 1.668, 2.373
Bishop, Brian 1.630, 1.636, 3.500
Bishop, Dorothy V.M. 2.676, 2.689
Bišmizzīn Arabic 1.540, 3.607, 4.70
Bišr ibn ʾAbī Xāzim 3.213
Bišr, Kamāl Muḥammad 4.89
Biṭār, Ṣalāḥ ad-Dīn al- 3.345
Bitjaa Kody, Zachee Denis 1.333
bitonal accent 2.400

bitransitive 1.556, 4.641
Biṭrīq, al- 4.542
Bittner, Maximilian 4.583, 4.585–586
bitūᶜ il-malāhi 4.255
bi-uniqueness condition 4.488
Bivar, A.D.H. 3.112
Biyyāḏiyyah Arabic 3.401, 3.403, 4.238
Bizerte Arabic 4.563, 4.681
Björnesjö, Sophia 2.1
Blacas Ewer 2.45
Blachère, Régis 2.14–15, 2.84, 2.86–88, 2.92, 2.224, 2.315, 2.317, 2.452, 3.689, 4.444, 4.484
Black Bedouin 3.360
Black September 4.325
Blair, Philip 2.560
Blair, Sheila S. 2.40–47, 2.597–602, 3.111, 3.308–309, 3.337–338, 3.340–341, 4.131, 4.133–134, 4.138, 4.560–561
Blake, Barry J. 4.532
Blakemore, Diane 1.471
Blanc, Haim 1.189, 1.222–223, 1.234–235, 1.406, 1.442–447, 1.492–493, 1.536, 1.538, 1.595, 1.607, 1.610, 1.612, 1.617, 1.630–632, 1.668, 2.8, 2.151, 2.276, 2.313, 2.414–417, 2.462, 2.481, 2.495, 2.509, 2.532, 2.534, 2.546, 2.571, 2.696, 3.8–9, 3.125, 3.237, 3.361–367, 3.369, 3.400, 3.403–406, 3.569, 3.585, 3.609, 3.613, 3.736–738, 4.1–3, 4.127, 4.238, 4.242, 4.323, 4.325, 4.350, 4.633, 4.664, 4.667
Blangoua 1.334
BLARK 1.215
blasphemy 4.416
Blau, Joshua 1.383–384, 1.404, 1.476, 1.493–494, 1.496, 1.562, 1.593, 1.613, 1.615–616, 1.618–619, 1.642, 2.22–23, 2.87, 2.226, 2.265–266, 2.275–279, 2.313, 2.463, 2.527, 2.529–530, 2.534, 2.629–630, 2.632, 3.216–218, 3.221–222, 3.420, 3.426, 3.466, 3.568, 3.640, 3.645, 3.654–655, 3.671, 3.673–674, 3.689, 3.696, 3.704, 3.738, 4.58–59, 4.316–317, 4.543–545, 4.558, 4.650, 4.656
Blau, Joyce 4.578
Blažek, Václav 2.65–69, 4.259–267
bleaching, semantic 4.160–164, 4.197, 4.532
blend 3.426, 3.494, 3.629
blending 1.175, 1.451, 3.187, 4.268, 4.270, 4.336, 4.379, 4.558
Blevins, Juliette 3.204–206
Blida 1.54
Blida Arabic 1.609
blind 1.316
Bloch, Alfred 3.650
Bloch, Ariel A. 1.491, 1.556, 2.278, 2.354, 3.417, 3.448, 3.704–705, 4.52
Bloch, Oscar 3.51
blocking of sandhi 2.397
Blohm, Dieter 1.51, 3.543
Bloom, Jonathan M. 2.45
Bloomfield, Leonard 1.125
Blue Nile 1.361
Blum-Kulka, Shoshana 3.713
Boas, Franz 1.484–485
bodily defect 2523, 3.279
body, parts of the 1.249, 1.281, 1.304, 1.327–328, 1.341, 1.392, 1.504, 2.160, 2.193, 2.222, 2.272, 2.594, 2.615, 3.152, 3.279, 3.672, 3.736–738, 4.554, 4.565
Bogo 1.334
Bohas, Georges 1.315, 1.562, 1.641, 2.83, 2.165, 2.168, 2.176–181, 2.233, 2.290, 2.299, 2.434, 2.448–449, 2.536, 2.538, 2.544, 2.562, 2.624, 3.45–52, 3.208, 3.210, 3.214, 3.423, 3.445, 3.617, 3.619, 4.14, 4.97 4.119–120, 4.288, 4.329–330, 4.332, 4.431, 4.675
Bohras, Daudi 2.328–329
boko 1.135
Bolinger, Dwight L. 2.113, 2.395
Bolotin, Naomi 3.382, 4.745
Bolozky, Shmuel 3.427
Bombay 2211, 2.329
Bombo 1.520, 1.522
Bonebakker, Seeger A. 3.538–539
Bongo Lendu 1.519
Bongor 3.634
Bongor Arabic 1.364, 3.61, 3.634–639
Bonnot, Jean-François 4.670–671
Bonte, Pierre 2.364, 3.172

Boogert, Nico van den 3.112, 3.519
bookhand 3.519
bookhand, Abbasid → Abbasid bookhand
Boomer, Donald 4.337
Booth, Marilyn 1.601–603
Bopp, Franz 4.37
Bordreuil, Pierre 3.469
Borer, Hagit 3.388, 3.429, 3.432
Boretzky, Norbert 3.643
Borg, Albert J. 3.141, 3.147, 3.157–159, 3.370, 3.739–740
Borg, Alexander 1.536–543, 1.607, 1.612, 2.231, 3.142, 3.144, 3.146, 3.361, 3.569, 3.672, 3.674
Borg, Gert 4.106
Bori Peninsula 2.268
Borisov, V.M. 3.23
Borkou 1.360
Borno → Bornu
Borno Empire → Bornu Empire 3.376
Bornou → Bornu 1.362
Bornouans 1.334
Bornu 1.362, 2.471, 3.373–375, 4.709
Bornu Arabic 3.634
Bornu Empire 3.376
Borromeo, Federico 1.170
Borrow, George 2.219
borrowing 1.95, 1.175, 1.294–295, 1.402, 1.407, 1.416, 1.418, 1.495, 1.497, 1.501, 2.29–31, 2.33, 2.35, 2.123, 2.354, 3.188, 4.368, 4.371, 4.461, 4.465
borrowing → loanwords
borrowing, lexical 2.666
borrowing, nonce 1.416, 1.496
borrowing of discourse markers 2.668
borrowing of function words 4.599
borrowing of numerals 2.668
borrowing of pronouns 2.668
borrowing of verbs 1.180, 2.668, 2.670
borrowings, Bedouin 1.269–270
borrowings, layering of 2.340, 2.343, 2.553
borrowings from Standard Arabic 1.224
Borsley, Robert D. 4.657
Bos, Petra 1.371, 2.76
Boša 2.217
Bosha, Ibrahim 1.660–663
Bosworth, Clifford Edmund 2.469, 3.717–721, 4.159, 4.255–257
Bot, Kees de 2.681–683
Bouazza, Hafid 2.78
Boubakeur, Si Hamza 4.426
Bouchard, Denis 4.489
Boucherit, Aziza 1.58–66, 1.596, 1.610, 3.587
Bouchon, G. 2.325
Boudelaa, Sami 1.421–427, 3.305
Boudlal, Abdelaziz 3.305, 4.351
Bouissef Rekab, Mohamed 3.295
Boujedra, Rashid 2.703
Boukous, Ahmed 2.696, 2.715, 3.60
Boula de Mareuil, Philippe 3.725
Boullata, Issa J. 2.306
boulomaic 3.234–235
Boumans, Louis 1.415, 1.418–419, 1.484, 2.76–77, 2.642, 3.318, 3.669
Boumédienne, President Houari 2.697
boundary loss 4.38
boundary marker 1.474
boundary shift 4.38
boundary strength 2.397
boundary tone 2.396
boundedness 1.50, 3.616, 3.619
Bounding Theory 2.170
Bourguiba, President Habib 3.349
Bourhis, Richard Y. 2.704, 4.321
Boussofara-Omar, Naima 1.372, 1.629–637
Bouteflika, President Abdelaziz 3.317
Bouvat, Lucien 4.587
Bouwman, Dinie 1.255–257, 3.135–141
Bowerman, Melissa 2.102
Bowers, John 3.685, 4.650
Bowersock, G.W. 2.373
Boyd, Douglas A. 3.194, 3.197, 3.199
Boyd, Sally 2.710
Brābīč 2.240
Brac de la Perrière, Eloise 3.338
Bracketed Grid Theory 3.616, 3.618
Bradac, James J. 2.652
Braga, Teofilo 3.4
Brahim, Ahmed 2.316–317, 3.226
Brahimi, Belynda 3.725
Brahimi, Fadila 3.314, 3.317
braille 1.316–320

Braille, Louis 1.316–320
Brak Arabic 2.622
Brakel, Lode Frank 2.334
Brakna 3.171
Brame, Michael K. 2.37, 2.165, 2.168, 2.366, 2.448, 3.240, 3.299, 3.302, 3.495, 3.726, 4.116, 4.344, 4.347–348, 4.350, 4.675
Bramon, Dolors 1.99
Branes Arabic 1.295
Brau, Hans Hermann 3.718
Braukämper, Ulrich 2.52, 4.709
Braun, Friederike 4.467
Bravmann, Meier Max 1.641–642, 2.13–14, 2.238, 2.477, 3.125, 3.671, 4.2, 4.40, 4.42, 4.484–486
Brazil 3.1–5
breath 4.666
Brelje, Herman W. 4.222
Brendemoen, Bernt 4.586
Brenner, Louis 3.136–139
Bresnan, Joan 4.535
Breton 2.712
BrightSpeech 1.210
Brincat, Joseph M. 3.141–145
broadcasting 1.213–214, 1.512, 2.29, 3.199
broadcasting → radio
broadcasting, Arabic 3.195, 3.376
broadcasting, Arabic in China 1.382
broadcasting, Arabic in Djibouti 1.654–655
broadcasting, Arabic in the United States 3.397
Broca's Area 2.690
Brockelmann, Carl 1.566, 1.574, 1.617, 1.637–640, 1.643, 2.156, 2.328, 3.7, 3.46, 3.227, 3.420, 3.423, 3.426, 3.645, 4.2, 4.41, 4.51, 4.70, 4.175, 4.196, 4.304, 4.349, 4.516–517, 4.558, 4.660, 4.683
Brockett, Adrian 2.57, 3.478, 3.484
Broe, Michael 3.610
Broeder, Peter 2.75
broken cursive 3.111, 4.130–132
broken plural 1.14, 1.54, 1.78, 1.80, 1.119, 1.643, 2.157, 2.250, 2.375, 2.410, 3.191, 3.243, 3.203–204, 3.329, 3.430, 3.440–447, 4.94, 4.174–175, 4.309, 4.478
broken plurals, acquisition of 2.101
Bronsted, K. 2.676
Broselow, Ellen 2.36, 2.38, 3.427, 3.494–495, 3.529, 3.607–615, 4.85, 4.344–345, 4.388, 4.390, 4.670
Brosh, Hezi 2.661
Brown, Gillian 4.727
Brown, H. Douglas 4.141
Brown, Penelope 1.657, 2.70, 3.658
Brown, Roger 4.140
Brown, W.S., Jr. 4.668
Browne, Wayles 3.302
Browning, Marguerite 4.67
browser 1.455
Bruce, Gosta 2.395, 3.724
Bruening, Benjamin 4.539
Brugman, Jan 4.543–544, 4.546
Brugnatelli, Vermondo 3.450
Bruhn de Garavito, Joyce 4.144–145
Bruinessen, Martin van 2.338
Bruni, Francesco 2.456
Brunner, Edmund 3.195
Brunot, Louis 1.295, 2.534, 3.287
Brunschvig, Robert 2.52, 4.446
Brustad, Kristen E. 1.221, 1.484, 2.194–195, 2.197, 3.68, 3.387, 3.398, 3.429–430, 3.432, 3.446, 3.554, 3.557, 3.739–740, 4.52, 4.62, 4.64–65, 4.162, 4.167, 4.316–318, 4.650, 4.656–657, 4.747
Bruyn, Adrienne 2.389
Btiġrīn Arabic 1.270
Bu Štā, az-Zubayr bin 1.603
Bubeník, Vít 3.552–558, 3.645
Bubiyan 2.608
Bucaram, Assad 3.2
Buckingham, Hugh 4.271
Buckley, Eugene 3.627
Buckwalter, Tim 1.460
buʿdā 1.567
Buddhism 2.474
Buduma 2.554
Buenos Aires 3.2–3
Buffalo 3.395
Buğday, Korkut 4.585
Bugia 2.455

Buginese 2.334
Bugūm 3.326
Buḥayra Arabic → Biḥēra Arabic
Bühler, Karl 4.328
Buḥturī, al- 2.305, 4.88, 4.202, 4.209
Buitelaar, Marjo 2.647
bukara-syndrome 1.320–322, 1.586, 1.610, 2.5, 2.7, 4.244
Bukayriyyah Arabic 4.127
Bukhara 1.29, 3.574, 4.424–425, 4.612
Bukhara Arabic 4.311, 4.613–616
Bukhara Arabs 1.28
Bukhara, Khanat 4.612
Bukharis 4.123
Bulāg Arabic 2.2
Būlāq 1.175
Bulgarian 4.371
Bulgarian, Arabic loanwords in 4.259–267
Bullāṭa, Kamāl 4.204
Būlus, Jawād 3.348
Bung 1.333
Bunnags 4.476
Bura 4.716
Buraymi Arabic 3.481
Burchardt, Bishop 3.142
Burckhardt, John Lewis 3.326, 3.402
Bureng, George Vincent 2.518
Burg Miġīzil 2.4
Burger, S. 2.98
Burgess, James 4.516
Burkina Faso 1.255
Burridge, Kate 1.657, 1.659, 2.69
Burriel, Andres Marcos 1.96
burst 1.26
Burt, Marina 4.140
Burton, John 4.6, 4.23
Burton, Richard F. 4.275
Burton-Page, J. 2.325
Burton-Roberts, Noel 1.125
Burullus 2.4
Burullus Arabic 2.2–3
Burundi 1.661–662, 4.381
Burykina, N. 4.612
Burzio's Generalization 4.491
Busa 4.281
Būsayrī, Murtaḍā 3.375
Busīrī, ʾAbū ʿAbdallāh Muḥammad ibn Ḥasan al- 1.8, 2.52, 2.335
Busse, Heribert 3.720
Bussmann, Hadumod 2.378
Bustān as-Salāṭīn 1.6, 1.8
Bustānī, ʿAbdallāh al- 3.39
Bustānī, Buṭrus al- 3.39, 3.42, 3.348, 3.539
Butana 4.380
Butcher, Andrew 2.229
Butros, Albert 2.456
Butterworth, Brian 4.269
Buxārī, Niẓām ad-Dīn al- 1.4, 2.335, 3.717, 4.138
Bybee, Joan L. 2.101, 2.193–194, 4.160–161
Byblos 3.409
Byram, Michael 4.148, 4.150, 4.153
Byrd, Roy J. 4.165
Byzantine Empire 2.198, 2.499
Byzantines 3.52, 4.215
Bzoch, Kenneth R. 2.676

C

Cachia, Pierre J.E. 2.224, 2.378, 2.452, 3.121
Cadora, Frederic J. 1.269, 2.230–231
caesura 4.89
Caetani, Leone 3.379, 3.717, 3.720
Čagatay → Chagatay
Cairene Arabic 1.24–25, 1.183–186, 1.189, 1.270, 1.275, 1.300, 1.304, 1.306, 1.312, 1.323–333, 1.406–410, 1.439, 1.441, 1.445, 1.492, 1.570, 1.589, 1.595–596, 1.604, 1.608, 1.625, 2.1–3, 2.5, 2.8, 2.117–118, 2.22, 2.120, 2.231, 2.354, 2.395–396, 2.404, 2.495, 2.533, 2.657–658, 3.9–11, 3.305, 3.449, 3.451, 3.426–427, 3.430, 3.542, 3.545–546, 3.569, 3.584, 3.586, 3.592, 3.607, 3.609–611, 3.613–614, 3.616–618, 3.620–621, 3.665–668, 3.673–675, 3.704, 3.737, 4.70–71, 4.85–86, 4.117, 4.168, 4.194, 4.255, 4.304, 4.326, 4.344, 4.346, 4.389–390, 4.533–534, 4.634, 4.637, 4.669, 4.681, 4.716, 4.727
Cairo 1.23, 1.83, 1.324, 1.589, 2.1, 2.30, 2.518, 4.157
Cairo Academy 1.453–454
Cairo-Damietta Corridor 1.589
Caitucoli, Claude 2.703
Calcutta 1.172, 1.287, 2.329
calembour 3.538
Calicut 3.128
Callegari, Laura 4.273–274
calligraphy 1.150, 1.152, 2.40, 2.383, 2.601, 3.308, 3.514, 3.520, 4.130–138
calligraphy, Chinese Muslim 1.379
calligraphy, modern 4.138
calligraphy, Mughal 4.133
calligraphy, Ottoman 4.134
calligraphy, zoomorphic 4.134
Callow, Kathleen 2.207
Caló 2.219
calque 1.175, 1.335, 2.10, 2.29, 2.201, 2.282, 2.285, 2.375, 2.377, 2.412, 2.607, 2.638, 2.641, 2.666, 2.684, 3.202, 3.222, 3.241, 3.579, 3.630, 4.167, 4.465, 4.545, 4.581, 4.585, 4.593
calque, stylistic 3.201
calque, Turkish in Arabic 4.593
calquing 1.297
calumny 2.363
Calvet, Louis-Jean 1.257
Camaioni, Lugia 1.341
Cambridge 1.171
Cameroon 1.333–334, 1.337–338, 2.30, 2.33, 2.137, 2.250, 3.61, 3.634, 4.709–710, 4.716
Cameroon, Arabic in 4.708
Cameroon Arabic 1.333–339
Cameroon Pidgin English 1.333
Campbell, Lyle 2.191, 2.197, 4.37–39, 4.41
Campbell, Stuart 2.340–345
Canaanite 1.561, 1.616, 3.408, 3.411–415, 3.418–419, 4.170–171, 4.171, 4.313, 4.478, 4.509
Canaanite, Amarna 3.416
Canaanite, Old 3.227
Canada 2.439
Canada, Arabic in 3.395, 3.398
candidates (in Optimality Theory) 3.492
canned phrase 1.648

Cannon, Garland 2.26
canonical syllable shape 3.440
Cant 2.217, 2.468, 4.251–252, 4.255
Cantarino, Vicente 1.472, 1.486, 2.80, 2.171, 2.173, 2.224–225, 2.295, 2.452–453, 3.164, 3.167–168, 3.272, 3.455, 3.553, 3.555–556, 4.163, 4.191–193, 4.317, 4.360, 4.486–487, 4.653, 4.660, 4.735
Cantineau, Jean 1.27, 1.260, 1.269, 1.544, 1.573, 1.578, 1.583–584, 1.605, 1.607–608, 1.640–642, 2.151, 2.280, 2.494–495, 2.501, 2.621–622, 3.46, 3.124, 3.149, 3.405, 3.410, 3.445, 3.525, 3.586, 3.604, 3.607, 3.609, 3.622, 3.726, 4.1, 4.95 4.339, 4.402–403, 4.406–407, 4.666, 4.678, 4.682
Canton 1.378
Canut, Cécile 2.77
Cape Colony 4.290–291
Cape Malays 4.290
Cape Town 2.335
Capello, Ernesto 3.1–6
capitalization 3.742
Caracausi, Girolamo 2.455, 4.216, 4.218
Caramazza, Alfonso 2.691
Carbou, Henri 1.361
cardinal number 2.102, 3.447–451, 4.18
cardinality 4.15, 4.18
caregiver talk → caretaker talk
caretaker talk 1.339–343
Carey, William 1.287
Cargile, Aaron Castelan 2.652
Carib 3.204
Carlsen, Martin 4.155
Carlson, Greg 4.67
Carnie, Andrew 4.392
Carpenter, Susan 2.51
Carre, Olivier 4.76
Carroll, Lewis 4.463
Carter, Michael G. 1.472, 1.620, 1.643, 2.16–18, 2.182–191, 2.224–226, 2.236–237, 2.292, 2.300–302, 2.307, 2.315, 2.317, 2.355, 2.390, 2.401, 2.403, 2.424–426, 2.429–432, 2.452–453, 2.477–478, 2.536, 2.539, 2.541–543, 2.625, 3.17, 3.21, 3.96–101, 3.107–109, 3.353, 3.453, 3.497, 3.546–552, 3.565, 4.2, 4.6, 4.9, 4.13, 4.101–102, 4.119, 4.219, 4.221, 4.436–437, 4.446–447, 4.485
Carthage 3.52, 3.409, 4.571–572
cartoon 1.487
CASA → Center for Arabic Studies Abroad
Casablanca 2.470, 2.715
Casablanca Arabic 1.596, 1.609, 2.621, 3.59, 3.273–274, 3.276–278, 3.285, 3.288–289, 3.291–293, 4.630, 4.679
Casagrande, Carla 2.362
Casati, Gaetano 1.519
case 1.343, 1.347–351, 2.170, 2.353
case, absolutive 1.563
case, abstract 1.347
case, active 1.644
case, default 1.349
case, functional 2.146
case, inherent 1.348
case, morphological 1.347
case, predicative 1.563, 644
case, structural 1.348, 2.146
case, subject 1.563
case, surface 2.146–147
case, underlying 2.146–147
case assigner 1.349–350
case assignment 1.348, 1.508, 2.170, 2.625, 3.354, 4.392
case-checking 1.348, 1.352
case ending 1.400–401, 1.562, 1.615–616, 1.632, 1.644, 2.36, 2.353, 2.401–402, 2.404, 2.625
case endings, loss of 1.493, 1.616, 2.629, 2.684, 3.568, 3.729
Case Filter 1.348
case frame 1.42, 1.344, 1.346
Case Grammar 1.41, 1.343–344, 1.457, 4.487
case inflection 1.153
case leveling 1.562
case marker 1.40, 2.353, 2.546–547, 2.625
case marking 1.350, 1.622–623, 2.146, 2.626, 3.429, 4.89
Case Marking, Exceptional 4.538–539
case relation 1.343–344

case role 1.41–42, 1.343–347, 4.488
case role, coreferential 1.345
case role, covert 1.345
case role, implicit 1.345
case system 1.562–563
case theory 1.347–353, 2.170, 4.746
Caselli, Maria C. 2.689
Caskell, Werner 2.198, 3.472, 3.719
Caspari, Karl Paul 1.169, 2.358–359
Cassels, David A. 1.660, 1.664
Cassola, Arnold 3.142
caste system 3.129
Castell, Edmund 1.167
Castilian 1.169, 2.287–289, 4.574
Castilian, Old 2.290
Castro, President Fidel 1.97
casus obliquus 1.559
casus pendens 1.476
casus rectus 1.559
catachresis 3.538
Catalan 2.282–286, 2.288–289, 3.286, 4.574
Catalonia 2.439
cataphora 1.353–355, 1.493, 1.556, 1.568, 2.149, 2.302, 2.357, 3.715–716, 4.439, 4.739
categorical change 4.38, 4.40
Categorical Grammar 1.457
Categorical Unification-based Grammar 1.457
category of speech 4.329
category switch 2.684
category, empty 1.348
category, functional 4.743
category, grammatical 1.343
category, lexical 4.743
catenation 3.268
catenatives 3.268
Catford, John 4.666–668
Catherine, Monastery of Saint 1.383
Caton, Steven 1.528, 2.202–204, 2.206, 3.500
Caubet, Dominique 1.189, 1.270, 1.273, 1.417–418, 1.420, 1.569, 1.571–572, 2.77–78, 2.388, 2.621–623, 2.642, 3.53, 3.55, 3.273–287, 3.318, 3.584–587, 3.705, 3.736–737, 4.71
Caucasian languages 1.133, 2.229
causality, linguistic 2.187
causation 1.359, 1.649, 3.225–226
Causation Hypothesis 3.226
causative 1.93, 1.252, 1.265, 1.328–329, 1.355–360, 1.551, 1.623, 1.626, 2.348, 2.367, 2.486, 2.564, 2.617, 3.184, 3.187, 3.251, 3.301, 3.563, 4.51, 4.177, 4.307, 4.313, 4.490–491, 4.532, 4.536, 4.568, 4.642–643, 4.697
causative prefix 1.399
causative, acquisition of 1.377, 2.102
causative, analytical 1.356–357, 359–360
causative, lexical 1.356–357
causative, morphological 1.356–357, 1.359–360
causative, passivized 3.553
causative-factitive 2.245
causativity 1.356, 1.358, 4.661
causativization 2.367
cause 2.309–310
cause, grammatical 2.179
cause, grammatical → ʿ*illa*
cavity, supralaryngeal 4.666
Caxaro, Pietro 3.142, 3.147, 3.151
c-command 1.311–312, 1.413, 2.134–135, 2.170, 2.174, 4.81, 4.393, 4.397, 4.648, 4.746
Cejador y Frauca, Julio 1.98
Celebes 2.335
Čelebi → Jalabi
Celtic languages 1.45, 3232, 4.368, 4.373
censorship 3.193
Center for Arabic Studies Abroad (CASA) 4.149
Central African Republic 1.519, 2.250, 3.634
Central Asia 1.205
Central Asian Arabic 1.189
Central Indian languages 2.217
Central Semitic languages 3.420, 4.313
Ceret'eli, Giorgi V. 1.28
Çermik 1.87
Cerulli, Enrico 2.52, 4.275–276
Çetenli 1.114
Ceuta 3.287
Chad 1.334, 1.360, 1.362, 1.518, 1.521, 2.137, 2.250, 3.60–61, 3.314–315,

3.634, 3.643, 4.376, 4.709–710, 4.715–716
Chad, Arabic in 1.260, 3.315, 3.448, 4.198, 4.310, 4.708, 4.712, 4.714, 4.756
Chad Arabic 1.37, 1.83, 1.190, 1.360–368, 1.521, 2.388, 3.52, 3.634–636, 3.638, 4.715–716
Chad Arabic Pidgin 1.37
Chadic languages 1.35–36, 1.38–39, 2.68, 2.250, 3.446
Chadic languages, East 1.36
Chafe, Wallace L. 1.41, 1.344, 1.474, 3.665, 4.502
Chafik, Mohamed 1.294–298
Chagatay 3.502, 3.574, 4.450, 4.603, 4.608
Chaha 2.67, 3.305, 3.494, 4.301, 4.557
Chahal, Dana 2.395–401, 3.726–727
Chahine, Jerome 3.717
Chahristan, Kamil F. 4.657
Chaib, M. 1.632
Chaika, Elaine 2.202, 2.205
Chaiken, Shelly 2.650–651, 2.653
chain 2.191
Chain Condition 4.488
Chaker, Salem 1.289
Chaldaean 2.414, 4.147, 4.402
Chaldaean Catholics 2.61
Chaldaean Neo-Aramaic 3.350
Cham 1.8
Chambers, Jack 4.627, 4.631–632
Chambers, John K. 1.613, 3.605
Champa 2.333, 2.340
Champa Muslims 2.334
Chan, Yuet-Hung 4.144–145
chancellery 3.110, 3.516, 3.518, 4.131
chancellery, Umayyad 4.74
chancellery hands 3.519
chancellery script 1.151–152, 4.130
chancery → chancellery
Chandra, Bharat 1.287
Chanfi, Ahmed 1.448–451
Chang’an 1.378–379
change of state 1.200–201
change, analogic 1.74–82
change, analogical 3.605
change, contact-induced 2.320–321
change, language 3.187–192
change, morphological 3.190–192
change, phonological 3.191
change, semantic 4.160–164
change, universals of 2.319
change, word order 4.372
change from above 3.605
change from below 3.605
channel 2.191
Chaouen Arabic 3.288–292
Chari 1.360
Chari division 1.334
Chari River 1.362
Chari-Baguirmi 1.360–362, 1.364
Chari-Logome 3.61
Charles, Marie-Claire 4.279
Charnay, Jean-Paul 1.628
chat forums, Arabic 2.382–383
chat room 3.500
chat session 3.274
chatbot 1.461
Chatterji, Suniti Kumar 1.287
Chaub 2.572
Chechenian 4.402
Chechens 2.498, 2.500, 3.350, 4.123
checking 2.170
Chehadeh, George 2.703
Chejne, Anwar G. 3.689
Chekayri, Abdellah 2.164–169, 2.447–451
Chekili, Ferid 1.457, 4.520–528
Chemor, Emilio Chuayffet 3.396
Chen, Da-Sheng 2.340
Chenfour, Noureddine 1.206–216
Cheraman Masjid 3.129
Cheraman Perumal 3.129
Cherchell 1.23, 1.25, 1.54
Cherchell Arabic 1.270, 1.609, 3.452
Cherif-Chebbi, Leila 1.378–383
Cherkaoui-Iqbal, Ahmed [ʾAḥmad aš-Šarqāwī ʾIqbāl] 4.158
Cheshire, Jenny 2.275
Chetrit, Joseph 1.446–447, 2.471, 2.531, 2.534, 4.157
Chicago 3.397
Chicovani, Guram 4.616–617, 4.622
Chiha, Michel 3.348–349
child bilingualism 1.368–375
child language 1.369, 1.375–378, 2.96, 4.50, 4.285

child-directed speech 1.339
child syntax 1.377
children, bilingual 1.370
children's rhyme 4.59
Chile 3.1
Chilha 4.572
China 1.378–383, 2.557
China, People's Republic of 1.382
Chinese 1.379–382, 2.9, 2.148, 2.333, 2.337, 2.341, 2.345, 3.706
Chinese, Mandarin 2.400, 4.195, 4.268
Chiu, Chi-Yue 2.653
Chomsky, Noam 1.151, 1.310, 1.342, 1.347–348, 1.350, 1.422, 1.457, 1.511, 2.37, 2.96, 2.131–132, 2.135, 2.170–171, 2.174, 2.370, 2.644, 2.687, 2.691, 3.86, 3.204, 3.228–230, 3.233, 3.302, 3.388, 3.525, 3.706, 3.711, 4.46, 4.49, 4.54, 4.67, 4.80–81, 4.342, 4.391, 4.396, 4.438–439, 4.441, 4.488–489, 4.491, 4.501–502, 4.504, 4.521–527, 4.535, 4.647, 4.649–650, 4.657, 4.718, 4.741–743
Choua Arabs → Shuwa Arabs
Choueiri, Lina 1.312, 1.457, 1.579–583, 2.370, 3.359, 4.64, 4.67–68, 4.82–84, 4.400, 4.718–722, 4.747
Choukri, Khalid 1.513
Chraibi, Driss 3.294
Christian Arabic 1.269–270, 2.23, 2.277, 3.592
Christian Baghdadi Arabic 1.94
Christian Middle Arabic 1.383–387, 2.266, 3.218
Christian mission 1.661
Christian terminology 1.181, 1.503
Christianization 1.654, 3.142
Christians 2.59, 2.61, 2.203–204, 3.218, 3.370
Christians, Baghdadi 4.633
Christians, Coptic 1.500
Christians, Iraqi 2.414–415
Christians, Jordanian 4.633–634
Christians, Maronite → Maronites
Christians, Sicilian 4.215
Christians, Syrian 3.129
Christie, W. 2.501
Christmann, Jakob 1.169
Chtatou, Mohamed 1.295
Chugtai, Shaheen 2.695, 2.705
church, Latin 4.216
Church Slavonic, Old 4.263
Churchyard, Henry 3.189
Chuvash 4.263
Chyet, Michael 2.604–608
Cifoletti, Guido 2.454–459, 4.573
Cigal 1.334
Cilicia 1.388, 4.404, 4.580
Cilician Arabic 1.87, 1.388–397, 2.160, 3.674, 3.702, 4.404, 4.589, 4.594
cinema 1.492, 1.546, 2.30, 2.32
Circassian 4.402, 4.508
Circassians 2.59–60, 2.498–499, 2.502, 2.506, 3.313
circumfix 2.194, 3.232, 3.297–298, 3.386, 4.565
circumfix, negative 1.185–186, 3.357–358, 3.406, 3.531, 4.569, 4.649–650
circumlocution 2.71, 4.141
circumstant 1.21–22, 1.567
circumstantial 3.108, 4.747
circumstantial clause 1.50
circumstantial predicative 1.16, 1.18–20
Cisse, Seydou 3.137–138
Cissoko, Sekene Mody 3.136
Cizre 1.87
Clahsen, Harald 2.689
Clapperton, Hugh 1.360
Clarity, Beverly E. 2.415
Clark, Eve V. 3.80
Clark, Vincent A. 4.479, 4.482
Clarke, Peter B. 2.552, 3.135, 4.279
class, morphological 1.208
class, social 4.634
class distinction 2.8
class prefix, Bantu 2.670
Classic Tree Theory 3.617
Classical Arabic 1.397–411, 1.490, 1.527–534, 1.578, 1.581–582, 1.592, 1.594, 1.614–620, 1.630–631, 1.633–634, 1.636, 1.669–670, 2.263, 2.267, 2.388, 2.404, 2.695, 2.708, 3.12, 3.94, 3.217, 3.318, 3.465
classicism 1.405–411, 2.163, 2.289, 3.220, 4.323
classicization 1.405–411, 1.419, 1.632, 3.9, 4.323

classification of Northwest Semitic 3.420
classificatory construction 4.229
classifier 4.231
Claudi, Ulrike 2.191, 2.193, 2.196
Claudot-Hawad, Hélène 3.172
clausal complement 4.60
clause 1.430, 2.143, 2.148, 2.207, 2.536, 2.538, 2.544
clause, adjective 1.481, 1.569
clause, adverbial 1.354, 1.468, 1.475–476, 4.747
clause, adversative 1.468
clause, asyndetic 1.387
clause, causal 1.387
clause, circumstantial 1.468–469, 1.473, 1.554, 1.649, 2.224, 2.491, 3.261, 4.195, 4.360, 4.746
clause, complement 1.468, 550, 4.188–189, 4.360, 4.362, 4.364–366, 4.742
clause, concessive 1.465–468
clause, conditional 1.72, 1.308, 1.396, 1.465, 1.468, 2.491, 3.293, 4.60, 4.544
clause, conjunctive 1.465
clause, connective 1.475
clause, continuative 4.728
clause, copular 3.683–684, 3.687, 4.653–654
clause, dependent 1.488
clause, embedded 4.65, 4.67, 4.721
clause, equational 3.683
clause, final 1.467–468
clause, finite 1.348, 488
clause, fronting object 4.544
clause, full 4.395–396
clause, *ḥāl* 1.473, 4.360, 4.531
clause, interrogative 2.521
clause, locative 1.567
clause, matrix 1.469, 4.80–82, 4.362
clause, nominal 1.506–507, 4.495–497, 4.499, 4.538, 4.653
clause, non-finite 4.361
clause, object 1.354
clause, pro-drop 4.656
clause, purpose 4.60
clause, relative 1.387, 1.466–467, 1.473, 1.479, 1.481, 1.552, 1.554, 2.146, 2.248, 2.489, 2.594, 3.261, 3.285, 3.389, 3.536, 4.60–70, 4.84, 4.187, 4.235, 4.317–318, 4.360, 4.524, 4.526, 4.746–747
clause, root 1.347, 1.506, 2.171
clause, small 1.350, 1.507–508, 4.395–396, 4.441, 4.538–539, 4.541
clause, subject 1.473
clause, subordinate 1.664, 2.92, 2.171, 2.538, 2.544, 3.272, 4.360
clause, temporal 1.387, 1.468
clause, verbal 1.506, 2.149, 4.497, 4.499, 4.647, 4.653–659
clause, verbally headed 4.653–654
clause, verb-initial 3.685
clause, verbless 1.350, 4.441
clause chaining 3.537
clause type in sign language 4.233
Clayman, Steven 2.382
cleft focus 2.115
cleft lip 3.336
cleft palate 2.678, 2.688, 3.336
cleft palate impairment 2.676
clefted question 4.62
clefting 4.522, 4.524
Clements, Georges N. 3.524–525
Clemetson, Lynette 2.60
Clenardus, Nicolaus 1.168
Cleveland, Ray L. 2.501, 2.506
cline 2.191
clipping 4.557–558
clitic 1.245–246, 1.312, 1.350, 1.411–414, 2.21, 2.353, 2.369, 3.236, 3.297, 3.429, 4.187, 4.192
clitic, pronominal 1.411
clitic-doubling 1.412, 1.414
clitic left dislocation 4.400
clitic pronoun 1.479
cliticization 1.412, 2.192, 3.587, 4.650
clock time 4.112
close apposition 1.125
Closed Syllable Vowel Shortening 3.724–725, 4.683
closure 1.26
cluster 1.86, 1.90
cluster, consonant 1.342, 3.362, 3.480, 3.528, 3.530, 3.601, 4.688
cluster, word-final 1.271, 1.300, 3.611
cluster, word-initial 1.261, 1.300, 4.688
cluster simplification 2.678
Clyne, Michael 2.684, 3.313

coalescence 2.346, 2.349, 2.353
coarticulation 2.460, 2.621, 3.596, 3.724, 4.669–671
coarticulation, backing 4.669–670
coarticulation, pharyngeal 4.670
Cobban, Helen 2.704
Cochin 3.128
coda 2.36, 3.611–612, 3.614, 3.729, 4.387–389
coda, complex 3.611–612
code, discreteness of 1.420
code-mixing 1.259, 1.369, 1.406, 1.408, 1.414, 1.419, 1.484, 1.669, 2.374, 2.377, 3.200
code-mixing, Arabic/French 4.575
codependency 2.322
code-switching 1.176, 1.369–370, 1.372, 1.406–407, 1.414–421, 1.440, 1.484, 1.495, 1.499, 1.585, 1.633–635, 1.650, 2.77, 2.274, 2.374, 2.652, 2.661, 2.703, 3.60, 3.318, 3.664–665, 3.669–670, 3.677, 4.368
code-switching, alternational 1.414–415
code-switching, Arabic/Dutch 3.318
code-switching, Arabic/English 1.417, 2.685, 3.318
code-switching, Arabic/French 1.417–418, 3.318, 4.408
code-switching, Arabic/Hebrew 3.318
code-switching, Arabic/Ivrit 2.465
code-switching, Arabic/Nubian 1.416
code-switching, Arabic/Spanish 3.318
code-switching, Arabic/Turkish 4.581
code-switching, cross-dialectal 4.326
code-switching, diglossic 1.419
code-switching, Dutch/Moroccan Arabic 2.76
code-switching, Dutch/Turkish 1.418
code-switching, English/Spanish 1.415, 1.417
code-switching, French/Moroccan Arabic 2.77
code-switching, French/Tunis Arabic 4.570
code-switching, *fuṣḥā*/dialect 3.76
code-switching, insertional 1.414–415
code-switching, inter-sentential 1.414
code-switching, intra-sentential 1.414, 1.416
code-switching, metaphorical 1.415
code-switching, Moroccan Arabic/Dutch 1.418–419, 2.77
code-switching, Tunisian Arabic/French 4.324, 4.326
code-switching, Welsh/English 4.321
codification of the *Qur'ān* 4.4
coding 1.513
coding algorithm 1.212
Coetsem, Frans van 2.341
Coetzee, Andries W. 2.228–232
Coetzee, Anna 4.291
cognate accusative 3.103
cognition 1.421–422
cognitive linguistics 1.421–427
cognitive orientation 1.485
cognitive psychology 1.425–426
cognitive system 3.228
Cognitivism 2.687
Cohen, Andrew 4.141
Cohen, David 1.444–445, 1.575–576, 1.593–594, 1.609, 1.617, 1.626, 1.628, 2.240, 2.242, 2.534, 2.622, 2.642, 3.45, 3.274, 3.451, 3.585, 3.640, 4.70, 4.563, 4.571, 4.575, 4.670, 4.754
Cohen, Marcel 1.35, 1.61, 2.271, 2.534, 3.46
coherence 1.427–430, 1.474, 1.568, 1.651, 2.145, 4.472, 4.475
coherency test 1.586
cohesion 1.427–433, 1.472–474, 1.648, 1.651, 3.666, 3.713, 4.472, 4.475
cohesion, collocational 1.431
cohesion, lexical 1.431
cohesion, sequential 1.432
cohesive device 3.713, 715
cohortative 2.23, 3.416
cohortative, Hebrew 4.178
co-indexing 3.684, 4.65
coins 2.41–42, 2.45, 2.341, 2.599–600
Cole, Peter 1.457
Coleridge, Samuel Taylor 4.91
Colin, Georges 1.102, 1.577, 3.274, 3.287–288, 3.293
collation 1.4
collective 1.249, 1.433, 1.522, 2.156–157, 2.160, 2.522, 2.645, 3.243–244, 3.426, 3.430, 3.439, 3.451, 4.94, 4.356

collective noun 1.237
College of Fort William 1.172
Collins, Chris 4.489
Collo 1.53–54
collocation 1.434–439, 2.136, 3.28, 3.540, 3.629, 3.668, 4.465
collocation, bound 1.437
collocation, English 1.437–438
collocation, macaronic 3.579
collocation, open 1.437
collocation, restricted 1.437
colloquial Arabic 1.405–408, 1.410, 1.419, 1.439–442, 1.527, 1.597–598, 1.604, 1.617, 1.631, 2.708, 3.24
colloquial drama, Egyptian 1.600
colloquial literature, Egyptian 1.598–602
colloquial poetry, Upper Egyptian 1.601
colloquial variety 1.483, 1.487–488, 1.528, 1.617, 1.620, 1.629, 1.631–633, 1.635, 1.667, 1.669, 4.251–252
colloquial, educated 1.405
colloquial, Egyptian 1.599
Colloquial, Standard Egyptian 2.1
colloquialism 1.406, 2.376–378, 3.39
colloquialization 1.419
Colombia 3.2, 3.5
colon 3.741
colonialism 1.530, 2.696, 3.294
colonialism, British 2.694, 2.699
colonialism, French 1.417, 2.62, 2.122, 2.660, 2.694, 2.699
colonialism, French in West Africa 3.137
colonization 1.448, 3.312, 4.574
colonization, British 3.376
colonization, European 2.654
colonization, French 3.174, 3.314
colophon 1.3
color term 2.99
colors and defects → adjective of color and defect; verb of color and defect
Combe, Etienne 2.42
comics 4.284–285
comma 3.741
command 2.331, 2.397
commedia dell'arte 4.255
comment 1.485, 2.113–115, 2.291, 2.308, 2.354–357, 2.434, 2.437, 2.537–539, 3.231, 3.683, 4.484–487, 4.494–501, 4.525–526
comment focus 4.496
commissive 3.248, 3.570, 4.329
communal dialect 1.87, 1.102, 1.271, 1.442–448, 1.596, 1.610, 2.8, 2.491, 2.532, 4.325
communal differences 1.246
communication 1.423, 1.427, 1.429, 1.461, 1.566, 4.494
Communication Accommodation Theory 4.320–328
communication, cross-dialectal 4.323, 4.325
communication, interethnic 2.714
communication, nonverbal 3.676
communicative act 2.205
communicative competence 3.398
communicative intention 3.160, 4.333
communicative strategy 2.196
communicative value 2.536, 4.738
communities, Jewish 1.445
communities, religio-ethnic 1.443
community code 1.419
community, religious 1.442
Comorian 1.448, 1.450, 1.662, 3.126, 3.314
Comorian → Shikomoro
Comoros 1.448–451, 1.661–662, 3.314, 3.351
comparative 1.18, 1.465, 2.13, 2.15, 2.195, 2.221, 2.244
comparative construction 1.502
comparative, feminine 1.594
comparison 2.15, 2442, 2523
compatibility, phonotactic 3.624
compensatory lengthening 3.695
competence 1.371
competence, bilingual 1.369, 2.77
competence, communicative 3.667, 4.150
competence, native 4.145
competence, second language 4.144
complement 1.217, 1.351, 1.354, 1.486, 1.643–644, 3.230, 3.684, 4.392, 4.488, 4.537
complement, adnominal 1.412
complement, adverbial 4.531–532, 4.735

complement, circumstantial 1.567
complement, clausal 1.350, 4.364, 4.539
complement, clause 1.626
complement, CP 4.747
complement, determinative 1.367
complement, infinitival 1.488
complement, nominal 1.622
complement, prepositional 1.412–413, 1.625, 3.559–560
complement, Small Clause 4.539
complement clause 1.350, 2.525
complement clause, finite 1.350
complement clause, interrogative 2.393
complementation, verb 4.624
complementizer 1.350, 1.417, 1.525, 2.132–133, 2.171–173, 2.332–333, 3.268, 3.272, 3.392, 4.60, 4.62, 4.64, 4.66, 4.188–189, 4.193, 4.198, 4.358, 4.360–364, 4.392, 4.630, 4.743
complementizer *ʾanna* 4.537, 4.539
complementizer, definite 4.64
complementizer, relative 4.63, 4.65
complementizer, subjunctive 4.746
complete functional complex 3.86
completeness 1.220
complex predicate 1.51
complexity 1.484
complexity, structural 2.118
Complex-NP Constraint 3.384, 4.80
compliance 3.225–226
compliance pattern 4.626
component, morphological 1.208
compositional 2.368
compositionality 4.490
Compositionality Principle 2.368
compound 1.451–455, 1.494, 1.522, 1.663, 2.267, 2.411, 3.203, 3.241, 3.425–426, 3.629–630, 4.462–463, 4.558, 4.691
compound verb 1.286
compound, hybrid 1.240
compound, phrasal 2.411–412
compound, verbal 4.599–601, 4.610
comprehensibility, mutual 4.322–323, 4.326
comprehension 2.99
computational 1.455–465
computational system 3.228–229
Comrie, Bernard 1.183–186, 3.64–71, 3.81, 3.85, 3.158, 3.675, 3.739–740, 4.63, 4.455, 4.457, 4.520, 4.726
conative 2.348
concatenation 1.209, 1.211, 1.213, 2.37
concatenation, morpheme 2.36, 3.611
concatenative morphology 3.149
concatenative synthesis 1.210
Conceptual Structure 4.489
conceptual-intentional system 3.228
conceptualization 1.422
concessive 1.465–467, 1.549
conclusive 1.431
concomitance 2.271
concord 1.12–15, 1.57, 1.230, 1.268, 1.386, 1.479–480, 2.410, 2.491, 3.285
concord, loss of 1.493
concord, negative 3.358–359, 3.655–656
concordance in the *Qurʾān* 4.26
concordance program 1.455
concordancer 1.515
Condax, Iovanna D. 4.668
condensation 2.193
condition 1.268, 2.478
condition on pronominal binding 4.81
conditional 1.202, 1.476, 1.509, 1.553, 2.195, 3.369, 3.643, 4.42
conditional clause 1.72, 1.308, 1.396, 1.465, 1.468, 2.491, 3.293, 4.69, 4.544
conditional particle 2.195
conditional phrase 1.220
conditional sentence 1.332, 3.186
configuration, syntactic 1.347–348
configurationality, discourse 3.387
Confucius 1.382
Congo Republic 1.519
congratulations 3.661
conjoined subject 1.47
conjugation 1.559–560, 2.91, 2.512, 2.522, 2.592, 3.96
conjunct 4.221, 4.397
conjunction 1.48, 1.240, 1.249, 1.280, 1.303, 1.328, 1.392, 1.404, 1.429–432, 1.467, 1.470–476, 1.499, 1.524, 1.550, 1.566, 3.183, 3.202, 4.191–192, 4.472
conjunction, complex 1.467
conjunction, conditional 2.615

conjunction, coordinating 2.243
conjunction, double 1.466
conjunction, final 3.220
conjunction, simple 1.467
conjunction, subordinating 2.243
Conjunction Phrase 1.47–48, 2.174
connecting vowel 2.22, 2.24
connection 1.432
Connectionism 2.687, 4.270
connective 1.431–432, 1.470–477, 1.648, 2.360, 4.190–192
connective, discourse 4.501
connective, hybrid 1.473
connective participle 3.544
connective slot 1.474
connective vowel 2.23
connectives in English 1.470
connectivity 1.430–431
connector 4.190–191
Connell, Dan 1.655
Connor, Ulla 1.648
Connor, W. 2.58
connotation 1.434, 657
connotation, negative 1.657
consecutive 1.465
consecutive, Hebrew 4.557
consensus 1.194
consistency 1.77, 2.165
consonant 2.237–238
consonant, emphatic 3.610
consonant, geminate 3.613
consonant, guttural 3.610
consonant, oral 4.669
consonant, palatal 3.524
consonant, pharyngealized 4.671
consonant, radical 1.575
consonant, 'stranded' 3.611
consonant, velarized 2.235
consonant cluster 1.90, 1.113, 1.225, 1.235, 1.244, 1.262, 1.277–278, 1.300, 1.325–326, 1.390, 1.548, 1.608, 2.35–37, 2.230, 2.242, 2.374–375, 2.417–418, 2.482, 2.560, 2.572, 2.577, 2.611, 3.276, 3.597
consonant cluster, final 2.263
consonant cluster, initial 1.295
consonant co-occurrence 3.463
consonant frequency 3.624
consonant harmony 1.341
consonant lengthening 3.566
consonantal skeleton 2.346
consonantism 1.296
consonants, acquisition of 2.98
consonants, Semitic 4.171–173
conspiracy, phonotactic 4.342
Constance, Queen 3.142
Constantine 1.53–55, 1.66
Constantinople 4.460
constative 2.360, 4.328, 4.332
constituency 1.457, 3.621–623
constituency test 4.646
constituent, adverbial 1.354
Constituent, Extra Clausal 2.148–149
constituent, fronted 4.722
constituent, topical 1.354
constituent boundary 4.102, 4.115, 4.734
constituent focus marker 2.115
constituent-command → c-command
constraint 4.627–628
constraint, binding 4.648
constraint, bottom-up 4.84
constraint, cognitive 1.484–485
constraint, combinatory 4.389
Constraint, Continuous Column 3.617
constraint, co-occurrence 4.95
Constraint, Coordinate Structure 3.386–387, 3.687
constraint, equivalence 1.416, 1.418
constraint, external 4.628
constraint, focus assignment 2.145
Constraint, Free Morpheme 1.415–416
constraint, island 4.68, 4.721–722
constraint, metrical 2.129, 4.389
constraint, morpheme structure 4.342
constraint, morpho-syntactic 1.634
Constraint, Nuclear Harmony 3.493
Constraint, Onset 3.493, 4.85
constraint, phonotactic 2.152, 4.389
constraint, processing 2.100
Constraint, Stop Lenition 1.540
constraint, structural 1.372, 1.484, 1.633, 1.635
Constraint, Superiority 4.84
constraint, syllable structure 2.35
constraint, syllable weight 4.683
constraint, syntactic category 4.268
constraint, top-down 4.84
Constraint, WH-Island 3.384, 4.81

constraint hierarchy 3.492
constraint on code-switching 1.415–418
constraint on transformations 4.522
constraint on word combinations 1.436
constraint ranking 2.38, 3.491–492, 3.497
constraint system 2.37
Constraint Tableau 3.492–496
constraint violation 2.38
constriction 3.596
construal 1.422
construct 3.67–69
construct, multi-noun 2.296
construct state 1.13, 1.82, 1.246, 1.263, 1.266, 1.283, 1.302, 1.326–327, 1.348, 1.350–351, 1.378, 1.386, 1.393, 1.395, 1.477–482, 1.552, 1.581–583, 1.644–645, 2.156, 2.248, 2.294–298, 2.484, 2.489, 2.613, 3.83, 3.152, 3.232–233, 3.276, 3.284, 3.303, 3.368, 3.389–390, 3.429, 3.432–433, 3.440, 3.535, 3.672, 4.15–17, 4.19, 4.63, 4.176, 4.244, 4.311, 4.400, 4.565, 4.742, 4.745–746
construct state, lexical 1.352
construct state, quantified 1.351–352
construct state in Northwest Semitic 3.415
construct state nominal 1.14
construction 1.428, 4.197
construction, causative 4.745
construction, classifier 4.231
construction, focalization 4.722
construction, focus 4.731–732
construction, interrogative 2.144
construction, partitive 4.721
construction, possessive 4.746
construction, *tanwīn-ʾalif* 2.226
construction, *tanwīn-naṣb* 2.226
construction, topicalized 4.722
constructions, asyndetic 3.221
contact, dialect → dialect contact
contact, language → language contact
containment hypothesis 4.160
contamination 1.314, 2.285, 4.270
content words, overuse of 2.675
context 1.429, 2.581, 4.166–167, 4.193
context, deictically shared 4.502
context, generically shared 4.502
context, situational 1.428
context, supranational speech 3.11
context, textually shared 4.502
contextual factor 1.647
contextual feature 1.650
contextual form 3.240
contextual indication 3.119
contextualization 1.484, 3.543
contingent succession 4.727
Contini, Riccardo 3.60, 3.467
Contini-Morava, Ellen 4.381–383
continuant 1.27
continuation 1.220, 1.331
continuative 1.475, 1.553
continuity 1.431
continuity, thematic 1.648
continuous 1.95, 1.253, 1.611, 3.536, 3.643, 4.457
continuous action 1.268
Continuous Column Constraint 3.620
continuous determiner 4.19
continuous past 2.516, 2.524
continuous present 2.569
continuum 1.633–635, 2.319–320, 2.324
continuum, dialectal 1.607
contour, falling 2.397, 2.400
contour, falling-rising 2.396
contour, level 2.397, 2.400
contour, pitch 2.395
contour, plateau 2.399–400
contour, rising 2.397, 2.400
contour segment 1.27
contradictory meaning 1.626–628
contrafactivity 3.263
contrariness 3.121
contrarium 3.121
contrast 2.321, 4.399, 4.505, 4.570
contrast in the *Qurʾān* 4.29
contrastive analysis 4.139
control theory 2.170
convention, human 4.684
conventionalization 4.455, 4.457
conventionalization of meaning 4.165–166
converb 1.51, 4.306, 4.587
convergence 1.419, 1.46, 1.489–495, 2.222, 2.262, 3.9, 3.60, 4.321, 4.324, 4.326–327, 4.368, 4.371
convergence, lexical 1.494

convergence, speech 4.321, 4.324
convergence, stereotypical 4.322, 4.324
conversation 2.207
conversation analysis 2.384, 2.202
conversational exchange 2.202
conversational implicature 1.657, 2.80, 3.677, 3.679, 3.681
converseness 2.585
conversion 2.59
conversion, speech-to-text 1.455
conversion, text-to-speech 1.455
converts 2.667, 3.59
co-occurrence 1.435–436
co-occurrence, consonant 3.624–626
co-occurrence class 3.626–627, 4.95–96
co-occurrence constraints 4.95
co-occurrence restriction 3.464, 4.1
Cook, James 4.416
Cook, Michael 4.4
Cook, Walter A. 1.41, 1.344
cookbook 4.58
Cooper, Robert L. 2.51
Cooper, Robin 4.15
Coordinate Structure Constraint 3.386–387, 3.687
coordination 1.69, 1.280, 1.417, 1.468, 1.472–474, 1.476, 1.648, 1.651, 2.144, 4.189–190
coordination, asyndetic 4.191–192
coordination, clausal 4.194, 4.397
coordination, nonconstituent 4.649
coordination, sentence 4.191–195, 4.393, 4.485, 4.646
coordination, syndetic 4.191–192
coordination of verbs 4.437
coordinator 2.148, 3.183, 4.190
coparticipation 3.224–225
Coptic 1.35, 1.37, 1.110, 1.387, 1.495–503, 2.1, 2.9, 2.59, 2.68, 2.235, 2.373, 2.560, 3.535, 3.592, 3.642, 4.1, 4.73, 4.370, 4.372, 4.511
Coptic, Bohairic 1.495–497, 3.72
Coptic, Fayyumic 1.495–496
Coptic, infinitive in 1.498–499
Coptic, Sahidic 1.495–497, 2.67, 3.72
Coptic, stative in 1.499
Coptic, verb in 1.499
Coptic church 2.1
Coptic in Arabic script 1.495
Coptic scientific terminology 1.499
Copts 1.323, 403, 527, 619
copula 1.19, 1.90–91, 1.217, 1.506–511, 1.607, 1.612, 2.100, 2.195, 2.273, 2.419, 2.431–432, 2.549, 3.81, 3.84–86, 3.114, 3.150, 3.236, 3.393, 3.585, 3.590, 3.683, 3.708–709, 3.740, 4.653–654
copula, negative 3.709, 4.569–570
copula, null 4.654
copula, omission of 1.340
copula, pronominal 3.710, 4.400
copular construction 3.708
copy 4.80–81
copying, constituent 2.19
Corbett, Greville G. 2.157, 3.69, 4.194
Corder, Stephen Pit 4.139–140
Cordoba 1.98, 2.44
core argument 1.182, 4.735
core sentence 2.538
core syntax 3.230–232
coreference hierarchy 1.124
coreferential 1.344, 1.622
coreferentiality 1.311–312, 4.53–54, 4.56
Corm, Charles 3.349
Corominas, Joan 2.287, 4.75
coronal 1.204
coronal obstruent 1.654
coronal sonorant 1.653–654
corpus 1.214–215, 1.460, 1.511–518, 2.187
corpus, lexical 2.702
corpus, multilingual 1.513
corpus, Spanish/Arabic 1.513
corpus, speech 1.455, 1.512–513
corpus, spoken Arabic 1.668
corpus, text 1.455, 1.512
corpus linguistics 1.438, 1.457, 1.460, 1.511–518, 2.384
corpus of grammar 2.264
corpus used by grammarians 2.87, 2.188, 2.265, 2.401–402
correctness of speech 2.176
correlation coefficient 2.322
correspondence, structural 1.485
correspondence between sound and meaning 2.178
Corriente, Federico C. 1.100–111, 1.400, 1.614–618, 2.262, 2.282–290,

2.629, 3.23, 3.585, 3.604, 3.696, 4.9, 4.71, 4.296
corroboration 1.436
corroborative 1.125, 4.436
corruption of speech 2.184, 3.59
corvee-texts 4.481
Corvetto, Ines 2.397
Coseriu, Eugenio 1.614
Côte française des Somalis 1.654
Couba 1.655
Coulmas, Florian 2.331
Council of Florence 1.166
counterfactual 1.331–332, 1.524, 1.553, 3.369, 3.537
counterfactual conditional 3.186
counterintuitive 1.202
counting unit 4.567
Couper-Kuhlen, Elizabeth 2.208
Coupland, Nikolas 4.321–322
Courbage, Youssef 4.577
court proceedings 2.214
courtesy expression 2.136
covariation pattern 2.321–322
coverb → converb
Cowan, David 4.358
Cowan, J. Milton 2.30, 3.26, 3.624
Cowan, William G. 2.495, 3.448, 4.614
Cowell, Mark W. 1.199, 2.96, 2.388, 2.390, 3.263, 3.429, 3.553, 3.556, 3.565, 3.568–569, 3.611–612, 3.614, 3.655, 3.684–685, 3.729, 3.737, 3.739–740, 4.62, 4.318, 4.339, 4.341, 4.348, 4.534, 4.649–650, 4.653, 4.655–656
Cowie, Anthony P. 1.437–438
CP → phrase, complementizer
CP domain 4.392
Crago, Martha 2.689
Craig, Colette 2.193
Crain, Stephen 1.342
Crass, Joachim 3.438
Crater 4.754
Crater Arabic 4.754
Crawford, David 2.63
creak 4.667
creaky voice 2.560, 4.299
creation of language 4.684
creation of language → imposition of speech; invention of language
creation of the *Qurʾān* → *Qurʾān*, creation of the
creativity 4.521
Cremona, Gherardo da → Gerard of Cremona
Cremona, Joseph 3.142
creole 1.37, 1.102, 1.595, 1.665, 2.320, 2.517, 2.588, 2.672, 3.10, 3.12, 3.639, 3.643, 4.368
creole, Atlantic 1.526
creole, Jamaican English 2.319–320
Creole Arabic 1.205, 1.518–527, 4.376
Creole Arabic → Ki-Nubi
Creole English, Krio 4.198
creolistics 2.319, 2.321
creolization 1.489, 1.519–520, 1.526, 1.594–595, 1.619, 2.255, 2.265, 2.588–589, 3.635, 3.640, 4.218
creolization, cultural 4.725
cristianos algarabiados 1.98
criteria for correct speech 2.86, 2.184
Critical Discourse Analysis 2.384
critical period 4.144–145
crossing 1.573–577
cross-marriage 4.634
Crowther, Samuel Ajayi 4.759
cryptography 1.150, 4.159
cryptolalic formation 2.217–218
Crystal, David 2.136, 2.690, 3.713, 3.728, 4.57
CS → construct state
C-tokens 1.669
cue, syntactic 2.686
Cuenca 2.44
Çukurova 1.111, 1.388
Çukurova Arabic 1.469, 1.502, 1.540, 4.532
Cultural Revolution 1.381–382
culture and language 1.527–536
culture, high 1.527
Culy, Cristopher 4.58
Cuny, Albert 2.156
Curl, Traci S. 4.44
curriculum, language teaching 2.107, 2.109, 2.111
curse 1.106, 2.136, 2.378, 2.463, 4.419–420
cursing 2.362, 2.364, 3.681
cursive 1.149

cursive, broken 2.601, 3.307
cursive writing 3.514
Cusa 4.217
Cushitic, retroflex consonants in 4.754
Cushitic languages 1.35–39, 2.53, 2.68, 2.268, 2.273, 3.52, 3.62, 3.438, 3.446, 3.493, 4.272, 4.371, 4.508, 4.754
Cutchi 1.665
Cutler, Anne 3.724, 4.269
Cuvalay-Haak, Martine 2.143, 3.237
CV-skeleton 3.301
cycle, transformational 4.522, 4.526
cyclical analysis 3.496
cyclical rule 3.619, 4.522
Cypriot (Maronite) Arabic 1.83–84, 1.189, 1.365, 1.536–543, 1.607, 1.612, 2.160, 2.231, 3.84, 3.672, 3.674–675
Cyprus 1.205, 536
Cyrenaica 3.52–53, 3.56
Cyrenaican Arabic 2.313–314, 3.55–56, 3.611, 4.551
Cyrene 3.52
Czech 3.555, 3.726
Czech, Arabic loanwords in 4.259–267

D

/ḍ/ 1.10, 1.103, 1.397, 2.473, 4.434
/ḍ/, lateral 1.102
D’Andrea, Daniel 2.681
daanu leer 4.182
daara 4.180, 4.183
dabba 1.4
Dabba Arabic 3.94, 3.693
ḍabṭ 3.308
ḍād 1.544–545, 2.252, 3.124, 3.189, 3.601, 3.604, 4.171, 4.273, 4.302–303, 4.434, 4.754
ḍād ḍaʿīfa 1.544
ḍād, luġat aḍ- 1.544
Dadan 3.467, 4.478, 4.480
Dadanitic 3.414, 3.418, 3.465, 3.468, 4.477, 4.479
Ḏ̣afīr 3.326, 4.125
Ḏ̣afīr Arabic 3.326
daftar 4.560
Dagana 1.362
Dagbani 4.281
Dagorn, Rene 3.719
ḏahāb 2.16
Ḏahabī, Šams ad-Dīn Muḥammad ibn ʾAḥmad aḏ- 2.362, 3.717
Dahalik 1.655, 2.268–270, 2.272
Dahbi, Mohamed 3.725
Dahdah, Antoine el- 2.224, 2.424–425, 2.452
Daher, Jamil 1.630, 2.30
Daher, Nazih 2.642
ḏ̣āhir 2.82–83
dahir berbère 3.294, 3.314
Ḏ̣āhir, Maḥmūd aḏ̣- 4.499
Dahl, Osten 4.163, 4.455, 4.457, 4.503
Dahl, Otto Chr. 3.125
Dahlak Archipelago 2.268, 2.271
Dahlak Kebir 1.655, 2.268–270, 2.272–273
Dahlgren, Sven-Olof 2.113–116, 4.654, 4.656, 4.498–499, 4.501–508, 4.725–736
Dahomey 4.279
Dai, John Xiang-ling 4.195
Daiber, Hans 2.200, 4.543–546
Daibul 2.325
dāʾim 2.431, 3.97, 3.543, 3.549, 4.455
Dakar 4.185
Dakhani 4.596
Dakhla Arabic 1.300, 1.586, 2.2, 2.8
Ḏ̣ālaʿ 4.751
Ḏ̣ālaʿ Arabic 4.750–751, 4.753–754, 4.756
dalāʾil an-nubuwwa 2.304
dalāla 2.403, 3.38, 4.163
dalāla ʿaqliyya 3.119
dalāla waḍʿiyya 3.119
dalālat al-iltizām 3.120
dalālat al-muṭābaqa 3.120
dalālat at-taḍammun 3.120
Dālī, Muḥammad ad- 1.382
dalil 2.302, 2.306, 3.122
Dallamale 2.55
Dalman, Gustav 3.410, 4.285
Dalrymple, Mary 4.647
Damāmīnī, ad- 2.189
Ḏamār 2.258
Damascus 1.25, 1.187, 1.546, 2.30, 2.199, 2.471, 2.505–506, 4.408, 4.534, 4.634–635
Damascus Arabic 1.15, 1.24, 1.298, 1.323, 1.395, 1.407, 1.493, 1.537,

1.540, 1.546–555, 1.608, 1.625, 2.22, 2.30, 2.162–163, 2.231, 2.278, 2.502, 2.509, 2.658, 3.11, 3.265, 3.373, 3.448–449, 3.564–565, 3.569, 3.612, 3.617–622, 3.702, 3.736–737, 4.194, 4.304, 4.344, 4.349, 4.404, 4.407–408, 4.504, 4.630, 4.632, 4.656, 4.664, 4.681
Damietta 1.324, 1.586, 2.4
Damietta Arabic 1.323
ḍamīr 1.49, 1.488, 1.555–559, 1.568, 2.300–301, 2.425, 3.422
ḍamīr ʿāʾid 2.132
ḍamīr al-fāʿil 1.556
ḍamīr al-faṣl 1.558–559, 3.385, 3.589–590
ḍamīr al-ġāʾib 2.317
ḍamīr al-qiṣṣa 2.316
ḍamīr aš-šaʾn 1.558, 2.316–317, 2.548, 2.552, 4.740
ḍamīr at-taʾkīd 1.559
ḍamīr fiʿl 2.301
ḍamīr munfaṣil 1.566
ḍamīr mustatir 4.101
ḍamīr muttaṣil 1.566–567
ḍamm 1.152, 1.309, 2.92, 2.232–234, 2.402, 3.163
ḍamma 1.560, 2.232, 2.234
Damurdašī, ʾAḥmad ad- 3.222
Dana 2.53
Danault, S. 2.691–692
Danecki, Janusz 2.298–300, 3.23, 3.123–125, 4.431–433
Daneš, František 4.502
Danfodio, Sheikh Utman 3.375
Dangaleat 1.36
Dānī, ʾAbū ʿAmr ʿUṯman ibn Saʿīd ad- 4.4
Daniel, Robert W. 3.467
Daniloff, Raymond G. 2.688, 4.669–670
Dann, Uriel 2.498
ḏ̣anna 1.558, 2.308, 2.355, 3.101, 3.234, 3.354
ḏ̣anna wa-ʾaxawātu-hā 1.71
Daoud, Mohamed 2.694, 2.696, 2.698, 2.703, 2.705, 2.660, 4.575
Daqahliyya 2.4
Daqahliyya Arabic 2.2
Daqqāq, ʿUmar 3.40
Dār al-ʾalsun 1.175
Dār as-Sūdān 1.361
Dār Bagli 1.362
Dar Salem 1.362
Daragözü 1.83–84, 1.89–90, 1.92, 1.94, 1.206
Daragözü Arabic 1.571, 2.160, 3.84, 3.672, 3.674
Daraw 2.1
ḍarb 3.210
Dardic languages 1.29, 2.220
Dar es Salaam 4.382
Dar es Salaam, University of 1.661
ḏ̣arf 1.310, 1.567, 2.186, 2.225, 2.294, 2.07, 2.357, 2.426, 3.97, 3.101, 3.103, 3.106–107, 3.109, 3.548, 4.12, 4.220, 4.464, 4.540
ḏ̣arf al-makān 3.108
ḏ̣arf az-zamān 3.108
ḏ̣arfiyya 2.186
Darfouf, Noreddine 3.48
Darfur 1.361, 1.519, 3.435, 4.375–376, 4.709
Dari 1.28, 2.406, 3.574, 3.577, 4.608
dārija 1.483, 1.487, 1.600, 2.648, 4.252
darj 3.340
Darj Arabic 3.55–56
Darrow, James 4.67–68
dars 3.129
darsa 1.665
Dars-i Niẓāmī 2.327
ḍarūra 2.404, 2.544, 3.94, 3.647–649, 3.651, 4.32, 4.200, 4.207–208, 4.210, 4.212
ḍarūrat aš-šiʿr 2.543–544, 4.558
Darwīš, ʿAbdallāh 3.40
Darwish, Kareem 1.460, 1.655
Darwish, Mahmud 4.213
Das, Rahul Peter 4.434
dash 3.741
ḏāt 3.178, 3.323, 4.436, 4.685
Dat, Mihai 1.565–569, 4.288
data collection 1.512
data routing 1.458
database 1.455, 1.458–459
database, linguistic 1.208
database, morphological 1.207–208
database, speech 1.210, 1.460

Daṯīnah Arabic 1.544, 2.388, 3.588, 3.705, 4.750–751, 4.754–755, 4.757
Daṯīnah 4.750
dative 1.412, 1.553
dative, ethical 1.308
dative construction 1.298
dative shift 1.349
dative suffix, deictic 3.368
Daudi Bohras 2.328–329
Dauer, Rebecca M. 3.725
Davids, Achmat 4.290–291
Davidson, Basil 4.279
Davies, Eirlys E. 1.372, 1.418, 2.58–65, 2.659, 2.695, 3.311–320
Davies, Humphrey 1.603
Davies, William D. 4.539
Davis, Stuart 2.365, 3.305, 3.529, 4.636–638, 4.669
Davis, Susan Schaefer 2.646
daʿwa 1.450–451, 2.51, 2.407, 4.78
Dawāġrah Arabic 3.401, 3.404–407, 4.238
Dawāsir 3.326
Dawlat 3.338
Dāwūd, Muftī 2.53
Dawwāra, Fuʾād 1.601
Dawwe 2.53
Dāxila → Dakhla 2.6
Daxla → Dakhla
Daybajat 2.259
Ḍayf, Šawqī 1.176, 1.614–615, 1.617, 1.629, 2.641, 3.105
Daylamī, ʿAbd ar-Rašīd 3.338
days of the week 4.111, 4.183
Dbūr Arabic 4.238
De Blois, F.C. 3.6
De Camp, David 2.319–320
De Houwer, Annick 1.368
De Jong, Kenneth 2.395–398, 3.726
De Simone, Adalgisa 4.218
de-accenting 2.396
Deaf 4.222, 4.224–225
Deaf clubs 4.225–226
deaf education 4.224
deaf villages 4.223
deaffrication 3.290, 4.172–173
deafness 2.677, 4.222–223
Dearborn 3.396–397
Debba 3.435
Debdou Arabic 3.289, 3.291, 4.70
Deboo, Jeffrey 2.260
DeBose, Charles E. 2.117
decategorialization 3.236
declaration 3.248, 3.570
declaration, performative 3.571
declarative 1.190, 4.329
declarative utterance 2.396
declension 1.559–564, 1.614–615, 1.643–646, 1.668, 2.179–180, 2.263, 2.401–406
declension, Akkadian 4.174–175
declension, loss of 1.616, 1.666, 3.697
declension, Northwest Semitic 3.414
declension, reduction of 3.696
declension, Semitic 4.174–175
declinability 1.308
declinable 1.309, 1.559
Decobert, Christian 1.367
decoding, pseudo-word 3.79
decoding skills 3.78
decolonization 1.661
decoupé work 4.133
decreolization 1.525, 1.595, 1.619, 2.163, 2.265, 2.517–518, 3.10, 3.12, 3.639–640
Dedan → Dadan
Dedanitic → Dadanitic
de-emphaticization 1.315, 2482
de-emphatization → de-emphaticization
deep structure 1.343–345, 2.300, 4.488–489, 4.521, 4.524–525, 4.741
defective 1.80
defective verb 1.209, 4.37
defective word 2.448
Defense Language Institute 4.148, 4.150
definite 1.190, 1.522, 1.560
definite article → article, definite
Definite Clause Grammar 1457, 1.514
Definite First Principle 4.734
definiteness 1.46–47, 1.187, 1.229, 1.297, 1.478, 1.480–481, 1.487, 1.552, 1.579–580, 1.582–583, 2.146, 2.296, 2.424, 2.592, 2.594, 3.86–87, 3.284, 3.389, 3.428, 3.548, 4.311–312, 4.495
definiteness marker 1.488, 1.491
definition, lexical 3.35
definition, word 3.42
degemination 2.591

Degen, Rainer 3.409
deglottalization 4.173
degrammaticalization 2.197
degree adverb 1.22
deictic 1.199, 1.471, 1.474, 3.82
deictic expression 1.565
deictic *hā-* 2.612
deictic *la-* 2.614
deictic value 1.571
Deir Ezzor → Dēr izZōr
Deir Ezzor Arabic → Dēr izZōr Arabic
deixis 1.565–569, 3.192, 3.677, 3.703
deixis, distant 3.645
deixis, far 1.569–570
deixis, middle 1.571–572
deixis, near 1.569–570
deixis, personal 1.566
deixis, proximate 3.645
deixis, spatial 1.566–567
deixis, temporal 1.567
deixis, textual 1.568
deixis, third 1.303, 1.392
delabialization 3.494
delateralization 4.173
Delattre, Pierre 4.368, 4.669
deletion 1.49–50, 1.245, 1.429, 1.507, 1.522, 2.18, 2.21, 2.35, 2.127–129, 2.167, 3.204, 3.617, 4.448
deletion of definite article 2.77
deletion, Equi-NP 4.527
deletion, glide 2.233, 2.366
Deletion, Low Vowel 3.621–622
deletion, segment 2.128
deletion, short vowel 4.674
deletion, vowel 1.522
Delhi 2.325–326, 4.596
Dell, François 3.493
Dell, Gary S. 4.269–270, 4.335
Dellys 1.54
Dellys Arabic 1.270, 1.609, 4.656
Delos 4.482
Delta 2.1
demarking 2.149
Deme, Abdoulaye 4.183
Demirdache, Hamida 3.382, 4.648
Demirel, Hamide 3.502
demonstrative 1.23–24, 1.105, 1.115, 1.239, 1.247, 1.279, 1.327, 1.391, 1.430–431, 1.522, 1.549, 1.565–566, 1.569–573, 1.579–582, 2.146, 2.161, 2.243, 2.511, 2.561, 2.590–591, 2.614, 3.182, 3.277, 3.364, 3.483, 4.245
demonstrative adverbial 1.22–23
demonstrative as article 3.634–635
demonstrative base 1.566
demonstrative phrase 1.22
demonstrative prefix 4.316
demonstrative pronoun 1.49, 1.263, 1.610, 1.612
demonstrative, attributive 1.570–572
demonstrative, proclitic 1.582
demonstrative, pronominal 1.572, 1.582
demotic elements 4.543
Dempwolff, Otto 2.66
Dendi 4.279–280
dendrogram 1.586
Denham, Dixon 1.360
Deniz, Fuat 2.61
denominative 1.252, 1.551–552, 1.575–577, 2.245, 2.620
denotative of state 2.224
dento-alveolar 3.597–599
Deny, Jean 4.587
Deobandh 2.326
deontic 3.234–235
dependency 1.344, 1.457–458, 1.514, 3.383, 4.625
dependency, unbounded 4.721–722
Dependency Grammar 1.457
dependency relation 4.392
depharyngealization 1.278, 1.302, 3.275, 3.695
depidginization 3.638, 3.642
depletion, semantic 4.160–161
Deprez, Viviane 3.655
Dēr ʿAṭīye Arabic 4.406
Dēr izZōr 1.270, 4.407
Dēr izZōr Arabic 2.414, 4.407–408
Dereli, Belgin 4.613, 4.616–617, 4.622
Derendiger, R. 3.61, 3.634
derivation 1.194, 1.313, 1.452–453, 1.573–579, 2.103, 2.165, 2.345, 2.447–448, 3.229–230, 3.297, 3.423, 4.93–95, 4.119, 4.461, 4.638, 4.640, 4.642
derivation, acquisition of 2.102
derivation, crashing of 3.228

derivation, formal 1.576–577
derivation, pattern 1.578
derived form 2.93, 4.306, 4.640
derived stem 4.306–307
derived structure 4.521
Derman, M. Ugur 3.338, 4.99, 4.138
Déroche, François 1.151, 2.597, 2.599–601, 3.110–112, 3.513–514, 3.517–518, 4.132
derogative 4.51
Dersim 2.607
Derwing, Tracey M. 2.117
Desai, Ziyaud-Din 2.326, 2.329
Descartes, René 1.449
description, structural 4.526
descriptor, morphological 1.208
desemanticization 2.194, 4.40, 4.160–161, 4.196
desibilation 4.173
desiderative 1.331, 3.333
Désoulières, Alain 4.595–603
despirantization 1.490
Desreumaux, Alain 3.469
Dessoulavy, Charles Louis 3.142
Destaing, Edmond 3.287
destressing 3.617, 4683
determination 3.244
determination system, Coptic 1.497
determiner 1.352, 1.479, 1.499, 1.552, 1.569, 1.579–583, 2.67, 3.232, 3.428, 4.16, 4.747
determiner, discontinuous 4.19
determiner, null 4.747
determiner, possessive 1.569
determiner feature 1.349
determiner head 1.351
Determiner Phrase 1.46, 1.351, 1.413–414, 3.429
detransitivization 4.625–626
Deuchar, Margaret 1.368–369
Devanagari 2.334, 4.595
development, parallel 1.492–493
developmental stage 2.100
Developmental Studies 4.141
Dévényi, Kinga 2.301, 2.401–406, 2.477–481, 2.542
deverbal 1.577
deverbal noun 4.306
deviation 2.309
device, anaphoric 1.431
device, cohesive 1.430–432
device, face-saving 1.650
device, lexical 1.430–431
devoicing 1.60–61, 1.90, 1.278, 1.300, 1.321, 1.325, 1.376, 1.548, 2.5, 2.590, 2.678–679, 3.147–148, 4.110, 4.711
Dewdney, John C. 4.579
Dez, Jacques 3.126
Dhaka 1.286
Dhaka, University of 1.258–259
Dhofar 1.260, 3.478, 3.694, 4.297, 4.373
Dhofar Arabic 4.304, 4.751
Dhuhab 1.655
Di Cristo, Albert 3.724
Dia Kossoi 4.279
Diab, Hassan 2.701
Diab, Rula 2.660
diachronic 1.604
diachrony 1.575
diacritic 1.150–152, 1.207, 3.77–78
diacritic(al) dot 2.183, 3.515–516
diacritic(al) mark 2.598, 2.630, 2.701
Diakite, Drissa 3.138
Diakonoff, Igor M. 3.420, 4.309
dialect 1.439, 1.442, 1.589, 4.251–252
dialect, communal 1.442–448, 1.492, 1.589, 2.212, 4.633
dialect, mixed 1589, 1.595
dialect, peripheral 3.672
dialect, rural 1.595, 1.605, 1.607–608, 2.8
dialect, sectarian → dialect, communal
dialect, standard 1.604
dialect, urban 1.595–596, 1.605, 1.618, 2.1, 2.8, 2.117, 2.161, 2.330–331, 3.55
dialect, written 1.323
dialect atlas 1.583–584, 1.590, 1.605
dialect atlas, Moroccan 3.287
dialect border 1.585
dialect classification 1.586, 1.604–613
dialect contact 1.585, 1.590, 1.593–595, 1.611, 1.617, 2.3, 3.406, 4.374, 4.628
dialect continuum 1.618
dialect corridor 2.4
dialect feature, primary 1.585

dialect feature, secondary 1.585
dialect features in Judaeo-Arabic 2.530
dialect geography 1.583–593, 604–605
dialect geography, generative 1.590
dialect koine 1.490–491, 1.593–597
dialect leveling 3.639
dialect literature 1.597–604
dialect literature in Egypt 1.598–602
dialect literature in Lebanon 1.602–603
dialect literature in Morocco 1.603
dialect literature in Tunisia 1.603
dialect literature in Upper Egyptian 1.599
dialect mixing 1.595
dialect network 1.445
dialect ranking 2.658
dialect rating 2.658–659
dialecticism 2.631, 3.217
dialectology 1.583
dialectology, Arabic 1.583
dialectology, Dutch 1.585
dialectology, German 1.585
dialectology, Japanese 1.585
dialectology, perceptive 1.585
dialectology, Romansh 1.585
dialectology, United States 1.585
dialectometry 1.585–586
dialects in the *Qurʾān* 3.14
dialects, emergence of 1.593, 1.615–622, 2.265–266, 3.12, 3.63
dialogic couplet 2.202
dialogue 4.727, 4.729–732, 4.734, 4.640–643
Diamare 1.334
Diamare Division 1.334
diana 1.660, 1.665
diaspora, Arab 2.665, 3.59
diathesis 1.486, 1.622–626
Diaz de Vivar, Rodrigo 1.98
Dib, Pierre 1.536
Dichtersprache 3.645
Dichy, Joseph 1.460, 1.483, 2.578–583, 4.458–466
Dickens, Charles 3.715
Dickins, James 2.559–571, 4.494–501, 4.528–535, 4.559
dictionary 1.487, 1.574, 2.184, 2.637, 2.702
dictionary, active 3.25
dictionary, Arabic 1.170–171
dictionary, Arabic/Arabic 3.37–45
dictionary, Arabic/Berber 1.294
dictionary, Arabic/Coptic 3.22
dictionary, Arabic/Dutch 1.438, 2.74, 2.702, 3.27
dictionary, Arabic/English 1.437, 3.22, 3.26–28, 3.30
dictionary, Arabic/French 3.27
dictionary, Arabic/German 3.26–28, 3.30
dictionary, Arabic/Hebrew 3.23
dictionary, Arabic/Indonesian 2.336
dictionary, Arabic/Italian 3.23
dictionary, Arabic/Latin 1.169–170, 3.22, 3.30
dictionary, Arabic/Persian 1.171, 3.22, 3.576
dictionary, Arabic/Russian 3.23
dictionary, Arabic/Spanish 3.23
dictionary, Arabic/Swedish 3.23
dictionary, Arabic/Turkish 1.170–171, 3.502
dictionary, bidirectional 3.26–27
dictionary, bilingual 1.214, 1.437, 3.21–30
dictionary, Cairo Arabic/Danish 3.23
dictionary, Chinese/Arabic 1.382
dictionary, Coptic 1.502–503
dictionary, Coptic/Arabic 1.496
dictionary, Dutch/Arabic 2.74, 3.27
dictionary, Egyptian Arabic/English 3.27
dictionary, electronic 1.455
dictionary, English/Arabic 1.438, 3.26
dictionary, etymological 3.43, 3.51
dictionary, French/Arabic 3.26
dictionary, German/Arabic 3.26–27
dictionary, Hausa 2.254
dictionary, historical 2.638, 3.23
dictionary, Italian/Arabic 3.23
dictionary, Juba Arabic 2.518
dictionary, Latin/Arabic 1.169–170, 3.23
dictionary, monolingual 1.435, 3.21, 3.37–45
dictionary, multilingual 1.460
dictionary, online 1.214
dictionary, onomasiological 3.31, 3.33
dictionary, passive 3.25
dictionary, Portuguese/Arabic 3.4
dictionary, Qurʾānic 3.34
dictionary, Russian/Arabic 3.23

dictionary, semasiological 3.31
dictionary, Spanish/Arabic 1169, 3.23
dictionary, Spanish/Latin 1.169
dictionary, specialized 3.24–25, 3.34
dictionary, Swahili 1.664
dictionary, Swedish/Arabic 3.23
dictionary, Syrian Arabic 1.546
dictionary, thematic 2.702, 3.33
dictionary of terminology 2.641
didactic poem 2.185
ḍidd 1.626–628, 3.34, 3.50, 4.167
Diekema, Anne R. 1.460
Diem, Werner 1.419, 1.489, 1.492, 1.494, 1.502, 1.562, 1.595, 1.615–617, 1.619–620, 1.638, 1.642, 2.160, 2.262–263, 2.265, 2.426, 2.534, 3.200, 3.216, 3.227, 3.293, 3.419–420, 3.422, 3.453, 3.518, 3.567–568, 3.584, 3.671, 3.696–697, 3.721, 4.102, 4.220, 4.304, 4.310, 4.349, 4.369–370, 4.372–374, 4.403, 4.530, 4.533–534, 4.558, 4.750–751, 4.753
Diephuis, Rob 2.75
Diesing, Molly 3.684, 4.648, 4.650, 4.657
Dietrich, Albert 2.200, 3.7, 3.720
differential → *différentiel*
differentiation, language 1.369
différentiel 1.114, 1.325, 1.548, 1.584, 3.406, 3.622, 4.243
diffusion 1.594, 1.619, 2.220
diffusion, lexical 3.604–605
diffusion, linguistic 1.586, 1.589
diffusion model 1.586
digital signal processing 1.211
digitization 1.512–513
diglossia 1.173, 1.323, 1.370–372, 1.397, 1.419, 1.439, 1.446, 1.483, 1.533, 1.537, 1.593, 1.599, 1.629–637, 1.647, 1.667, 2.10, 2.29, 2.111, 2.214, 2.265, 2.414, 2.501, 2.630, 2.642, 2.654–655, 2.657, 2.662, 2.695, 2.708, 3.10, 3.13–14, 3.16, 3.24, 3.70, 3.74, 3.76, 3.78–79, 3.216–217, 3.219, 3.239, 3.274, 3.293, 3.317, 3.348, 3.640, 3.645, 3.664, 3.726, 4.58, 4.252, 4.510
diglossia and bilingualism 1.102
diglossia in Saudi Arabia 4.129
diglossia in Syria 4.408
diglossia model 4.150
diglossic variable 2.657
digraph 4.517
Dihlavī, Amīr Xusraw 3.338
dihqān 2.406
Dihxudā, ʿA.A. 2.410–411
DIINAR-MBC 1.460, 1.512, 1.516
Dijk, Teun A. van 1.470–471, 2.207, 2.384, 3.665
Dijken, Lianne van 4.231
Dīk al-Jinn 4.104
Dik, Simon C. 1.457, 2.143, 2.145–146, 4.189, 4.495, 4.726, 4.732, 4.734
Dikhil 1.655
ḏikr 1.49, 2.329, 2.537, 4.739
ḏikr aš-šayʾ 3.121
Dikwa 4.715
dilāla → *dalāla*
Diller, Anthony 4.476–477
Dillmann, August 4.306
Dimašqī, al-Kātib ad- 3.339, 4.132
diminutive 1.54–55, 1.64, 1.66, 1.76–79, 1.107, 1.116, 1.192, 1.251, 1.264, 1.280, 1.297, 1.304, 1.328, 1.331, 1.341, 1.491, 1.550, 1.576, 1.637–640, 2.156, 2.244, 2.350–351, 2.424, 2.616, 2.631, 3.153, 3.171, 3.183, 3.279, 3.304, 3.366, 3.423–424, 3.427, 3.486, 3.532, 4.52, 4.95, 4.113, 4.553, 4.557, 4.559, 4.616
diminutive, adverbial 1.639
diminutive, in place names 4.510–511
diminutive, verbal 1.639, 2.244
Dimlî 2.604
ḏimmī 1.444, 1.537, 3.346
Dīnawarī, ʾAbū Ḥanīfa ad- 3.32
Dindān 3.646
Dinka 1.519, 4.375
Dinnsen, Daniel 2.688
Dinshaway 3.193
Diodorus of Sicily 1.127
Dionysius Halicarnassensis 2.66
Diop, Amadou H. 4.723
Diouf, Jean L. 4.723
diphthong 1.89, 1.103–104, 1.113, 1.224, 1.235 1.242, 1.261, 1.390,

1.539, 1.547, 1.610, 1.616, 1.640–643, 1.646, 2.4–5, 2.7, 2.164, 2.242, 2.482, 2.611, 3.276, 3.289, 3.362, 3.405, 3.599
diphthong, secondary 1.641
diphthongization 1.640, 1.642, 2.313, 2.349, 2.351, 4.109, 4.406
diphthongization, secondary 1.642
diphthongs, acquisition of 2.98
diptosis 1.17, 1.38, 1.188, 1.398, 1.559–561, 1.580, 1.643–647, 2.353, 2.403, 3.241, 3.303, 3.414, 3.426, 3.453, 3.721, 4.443–444
Dirbaşiye Arabic 4.407
direct object 1.350
direct object, introduction of 1.553
direct object, marking of 1.395, 1.540, 1.612
direct speech 1.649
direct speech in the *Qurʾān* 4.30
directional syllabification 4.85
directionality 2.36, 2.39, 3.616, 3.619
directive 3.248, 3.570, 4.329
dirge 4.104
disambiguating device 1.558
disambiguation 1.430, 1.458, 2.114
discourse 1.427–431, 1.472–473, 1.568, 2.19, 2.143, 2.207
discourse, behavioral 4.727
discourse, business 1.649
discourse, expository 4.727
discourse, indirect 3.638
discourse, narrative 2.208, 4.727, 4.729–733
discourse, news 2.208
discourse, persuasive 3.500
discourse, religious 1.650
discourse, spoken 1.650, 1.652
discourse, universe of 3.87
discourse analysis 1.647–653, 2.202, 4.142
discourse connective 1.471
discourse execution 2.148
discourse marker 1.471, 1.474, 2.668, 4.192
discourse organization 1.472, 1.647, 2.148
discourse pattern 1.651
discourse strategy 1.415, 1.429
discourse topic 4.319
discourse type 1.475, 4.727, 4.731
discussion forum 2.382
Dīse, ad- 3.402
disfluency 2.685
disjunction 1.429, 1.471, 4.190–191, 4.472
dislocation 4.486
disorder, grammatical 2.679
disorder, language → language disorder
disorder, phonological 2.676–677, 2.688
displacement 2.369–370
Displacement Theory 4.80
dissimilation 1.104, 1.245, 1.262, 1.313–315, 1.539, 1.653–654, 2.483, 2.494, 2.558, 2.612, 3.148, 3.481, 3.494, 3.592, 4.33
dissimilation, progressive 1.653
dissimilation, regressive 1.653
distance of tension 4.212
distance, structural 1.483
distich 3.173
distributional method 2.627
distributor 4.43–45
di-syllable 1.211, 1.213
ditransitive 1.67, 3.224, 4.537–538
Ditters, Everhard 1.455–465, 1.511–518, 3.164–169, 3.428, 4.155, 4.659
divergence 1.489, 1.494, 2.192, 2.194, 2.262
divergence of speech 4.321
dīwān 3.516
dīwānī 4.98, 4.132
Dixit, R. Prakash 4.666
Diyāb, Maḥmūd 1.601
Diyarbakır 1.87–89, 1.91–93, 1.95
Diyarbakır Arabic 1.538, 1.607, 2.420, 4.70
Djajadiningrat, Hoesein 1.6–7, 1.9–10
Djebar, Assia 1.529
Djebel Nefoussa 3.56
Djegere 1.334
Djeinau 4.613
Djeinau Arabic 4.613–614, 4.616
Djenné 3.135
Djerba Arabic 4.633
Djibouti 1.451, 1.654–656, 2.105, 2.268, 2.274, 3.312, 3.314, 3.351, 4.275

Djibouti Town 1.634–635, 1.655
Djidjelli 1.24, 1.53–54, 1.84, 1.189
Djidjelli Arabic 1.609, 3.674, 3.737, 4.70, 4.371
Djogari 4.613
Djogari Arabic 4.613–615, 4.617
D-linking 4.721
Dobhāsī 1.287–288
DO-construction 1.197, 1.419, 2.221
document indexing 1.207
document processing 1.459
Dodd, Erica Cruikshank 2.45
Dogancay-Aktuna, Seran 4.577–578
doggerel 1.340
Dokainish, Mona 2.695
Dolgopolsky, Aron 2.67
Dols, Michael W. 1.589
Dom 2.217
domain 1.483
domain, language 3.317
domain, source 1.424–425
domain, target 1.424–425
domain of use 1.633, 1.636
domain overlap 2.710
Domari 2.217–221, 3.313, 3.350
Domari, Jerusalem 2.221
Dombay, Franz von 3.287
Dome of the Rock 1.151, 2.40, 2.599, 4.130, 4.134
dominant code 1.419
Dominant Language Hypothesis 1.416, 1.420
domination, social 4.368
Dongola 1.361
Dongolawi 3.438, 4.375
Dongolese → Dongolawi
Doniach, Nakdimon 3.26
Donitsa-Schmidt, Smadar 2.464
Donner, Fred McGraw 1.595
Donohue, John J. 3.346
Dorian, Nancy 2.684
Doron, Edit 2.367, 3.386–387, 3.684, 3.687–688, 4.316, 4.395, 4.400
Dōsiri Arabic 2.609
Doss, Madiha 1.570, 1.572, 2.2, 2.331–333, 2.377
dots 1.400
double construct 2.489
double entendre 2.497, 3.538
double object construction 4.490, 4.492, 4.537–538
double object, pronominal 1.237
doublet 1.233–234, 1.314–315, 1.366, 1.546, 2.254, 2.280, 2.342, 2.388, 3.577
doublet, synonymic 4.546
doubling 1.413–414
doubling, final 4.51
doubling, initial 4.51
doubling rule 1.474
Doubly Filled COMP filter 4.719
Doughty, Charles Montagu 3.402
Doutte, Edmond 2.647
DO-verb 1.34, 2.77
Dowty, David 1.197
Dozy, Reinhart Pieter Anne 1.171, 3.6–7, 3.23, 3.30
DP → phrase, Determiner
Dra Wadi 3.170
dragomanes 1.98
Drake-Brockman, Ralph E. 4.275
Dravidian languages 2.67, 3.204, 4.372, 4.433, 4.452
Dravidian languages, South 3.128
Drayj Arabic 4.406
Dreisilbengesetz → stress rule, three-syllable
Dressler, Wolfgang 1.430, 3.207, 4.475
Drewes, Abraham Johannes 2.52, 2.55
Drewes, G.W.J. 2.334, 2.340, 2.344
drift 2.267
drift, general 3.641
drift, internal 1.595
drift, linguistic 1.616
drift, natural 1.594, 1.616
Driss, Rhaid 2.124
Drossaart Lulofs, Hendrik Joan 4.543–544, 4.546
Drozdík, Ladislav 1.668, 2.156, 2.158, 4.64
Druel, Jean N. 3.377–381
Druze 2.59, 2.502, 2.658, 4.407
Druze Arabic 1.275–276, 1.442, 1.445, 2.314, 3.592, 4.407
Druze Arabs 2.464
Dry, Helen Aristar 2.207
Dryer, Matthew S. 3.64–66, 3.68, 3.705
D-stem 4.307

D-structure → deep structure
ḏū 1.310, 1.399
Ḏū Ġābat 4.480
ḏū l-ḥāl 2.224–225
Ḏū Raydān 2.256
Ḏū r-Rumma 1.402, 4.200, 4.202
Du Val, Jean-Baptiste 3.23
duʿāʾ 1.381, 2.46, 2.329, 4.740
dual 1.54, 1.107, 1.266, 1.386, 1.550, 1.554, 1.559–560, 1.562, 1.570–571, 1.582, 1.616, 1.646, 1.668–669, 2.119–120, 2.205, 2.353, 2.484, 3.67, 3.154, 3.183, 3.243, 3.280, 3.426, 3.439–440, 3.446, 3.450, 3.532, 3.736, 3.738, 4.304, 4.554, 4.565, 4.567, 4.616, 4.711
dual, loss of 1.490, 1.493, 1.594, 3.67
dual ending 2.402, 2.547, 3.691, 3.693, 3.696
dual in Northwest Semitic 3.415
dual marker 2.67
Duala 1.333
Duʾali, ad- → ʾAbū l-ʾAswad ad-Duʾalī
duality of word and meaning 1.428
Dubai Arabic 4.62
dū-bayt 4.91
Ducrot, Oswald 1.575, 2.468, 3.233
Ducroz, Jean-Marie 4.279
Ḍufār → Dhofar
Ḍufār Arabic → Dhofar Arabic
Duhūs Arabic 2.2
Dulay, Heidi 4.140
Ḍullām 3.401
Ḍullām Arabic 3.400, 4.238
Ḍum 2.217
Dūmah 1.126–127, 1.129
Duman 2.217, 2.220
Dūmat al-Jandal 4.123
Dumestre, Gérard 3.138
dummy element 1.411, 1.418, 2.218, 2.317, 4.530
dummy subject 2.315
dummy verb 2.412, 2.606, 2.668
Dumyāṭ → Damietta
Dun Karm Psaila 3.143
dūna 1.469
Dundar, F. 4.579–580
Dunn, Carla 2.688
dūr 1.129
Durand, Olivier 3.287, 3.289, 3.587
duration 1.211
durative 1.265, 1.282, 2.490, 3.221, 3.266, 4.622
durativity 1.331, 3.186
Durayhimi 1.655
durhīmi 1.655
Durie, Mark 1.9
Durkheim, Emile 4.73
Durrani, Atash 3.511
Durūʿ 2.213
Dūš Arabic 2.2
Dussaud, René 3.469
Dūst Muḥammad 2.598
Duszak, Anna 1.649
Dutch 1.7, 2.77, 2.333, 2.395, 2.400, 2.685, 3.231, 4.53, 4.262–263, 285, 4.370
Dutch, plural marker in 2.686
Dutch East Indies Company 4.290
Dutch in Australia 2.684
Dweik, Bader S. 2.500, 2.652, 2.657, 3.395
Dyerma 4.279
dynamic verb 1.42
dysarthria 2.690
dysfluency 2.675
dysgraphia 2.689
dyslexia 2.688
Dyslexia, Deep 2.692
Dyslexia, Developmental 2.689
Dyson, Alice T. 1.376, 2.97–99, 2.679, 2.688
dysphasia 2.675
dysphemism 1.656–659, 4.416
dysphemism, euphemistic 1.659
dyspraxia, verbal 2.675
Dzhalil, Dzhalile 2.607–608
Dzhalil, Ordikhane 2.607–608

E

Eagly, Alice H. 2.650–651, 2.653
East Africa 1.520, 1.619, 1.660–666, 2.211
East African coast 3.126
East Bank 2.498, 2.502
East India Company 2.326
East Morocco 1.27
East Semitic 4.170–171
East Sudan 1.23
Eastern Desert 2.1
Eblaic 2.66–67
Ebrahim, Mogamat Hoosain 4.291
echo question 1.266, 2.393, 4.566
echo word 1.246, 4.50
Echols, John M. 2.341
Echu, George 1.333–339, 2.30, 2.33
Eckert, Penelope 4.629
école de village 1.336
école régionale 1.336
Ecole spéciale des langues orientales 1.171
economy 3.228–229
economy, linguistic 3.189
Ecuador 3.1–2, 3.3, 3.5
Edfu Arabic 1.306
edge tone 2.395–398
editing of manuscripts 2.699
Edomite 3.409
Educated Arabic 1.630, 1.666, 2.377, 2.696, 3.14, 3.18–19
Educated Colloquial Arabic 1.409–410
Educated Spoken Arabic 1.630–632, 1.634, 1.666–668, 2.661, 4.325–326
education 1.441, 1.533, 2.8, 2.104–105, 2.254, 3.11–12, 3.74
education, foreign-language 2.695
education, Islamic 1.256, 3.136
education, language of 1.404
education, multilingual 1.370–371, 2.704
education, primary 1.654
education, religious 1.258
educational policy 2.110
educational system 3.75
Edwards, John 2.58, 4.633
Edwards, Malcolm 1.372, 1.458
Edzard, Lutz 1.188–191, 1.489–495, 1.559–564, 3.422–428, 3.453, 3.478–499, 4.1–3, 4.557–559
Effat, Ragia M. 3.192–204
Egba 4.759
Eggs, Ekkehard 2.442, 3.121
Eguilaz y Yanguas, Leopoldo 1.97
Egypt 1.531, 1.533, 1.597, 2.1–12, 2.105–107, 2.109, 2.111, 3.316
Egypt, Arabic in 1.444, 1.605, 2.272, 3.15–19, 3.672, 3.675
Egypt, oases of 1.611
Egypt, youth speech in 4.764–766
Egyptian, Old 1.35, 1.37–39, 1.563, 2.560, 3.592, 4.104, 4.511, 4.759
Egyptian Arabic 1.24, 1.37, 1.75, 1.80–81, 1.83, 1.86, 1.183, 1.197, 1.199, 1.206, 1.217–221, 1.260, 1.270, 1.323, 1.340–342, 1.351, 1.384, 1.417, 1.419, 1.483, 1.495, 1.501–504, 1.538, 1.554, 1.562, 1.570–572, 1.584, 1.595, 1.604, 1.606–609, 1.611, 1.632, 1.636, 1.657–659, 2.20, 2.32, 2.36–38, 2.67–68, 2.71, 2.74, 2.96–100, 2.136, 2.160–161, 2.194, 2.199, 2.203, 2.344, 2.375, 2.388, 2.390, 2.455–459, 2.517, 2.560, 2.584,

2.587, 2.597, 2.649, 2.656, 2.658, 2.678, 3.12, 3.65–68, 3.70, 3.81–85, 3.87, 3.97, 3.190, 3.198, 3.200–201, 3.236, 3.240, 3.264–268, 3.386, 3.389–393, 3.398, 3.456, 3.545, 3.560–563, 3.572–573, 3.586–587, 3.609, 3.612, 3.669, 3.674, 3.676, 3.680, 3.684, 3.687, 3.705–706, 3.707–711, 4.45, 4.50–51, 4.62–64, 4.66, 4.71, 4.195–198, 4.284–285, 4.317, 4.354, 4.372, 4.399, 4.468, 4.470, 4.490, 4.649, 4.662, 4.671–672, 4.719, 4.742, 4.747, 4.757
Egyptianism 1.407
Egyptianization 3.349
Eḥklo’ 4.297
Ehret, Christopher 1.315, 2.68, 3.46, 3.435, 4.97
Eickelman, Dale F. 2.60, 2.64
Eid, Mushira 1.312, 1.417, 1.419–420, 1.484, 1.634, 1.669, 2.642, 3.80–88, 3.318, 3.358, 3.385–386, 3.585, 3.664, 3.705–713, 4.49, 4.395, 4.400, 4.520, 4.523, 4.527, 4.648–650, 4.655, 4.657
Eilers, Wilhelm 2.67
Eisele, John C. 1.182–187, 1.195–201, 1.220, 1.630, 2.654, 3.97–98, 3.385, 3.545, 3.683–684, 4.251–259, 4.455, 4.650, 4.654
Eisenstein, Herbert 2.378–380
ejective 2.269, 3.63, 3.190, 4.299
ejective stop 2.53
Ejieh, Michael U.C. 2.703
Eksell, Kerstin 1.82–85, 3.237, 3.432, 3.672–674, 3.736–739
El Aissati, Abderrahman 2.75–76, 2.681–687
El Bayad 1.56
El Cid 1.98
El Fathi, Abderrahman 3.295
El Hassan, Shahir A. 1.633
El Jadida Arabic 3.288
El Kirat, Yamina 2.707–716
El Maleh, Edmond Amran 2.64
El Medlaoui, Mohamed 3.60
El Zarka, Dina 4.50–53
El-Ali, Nasser 2.501, 2.655–656, 2.658
Elamrani-Jamal, Abdelali 3.113, 3.177
El-Araby, Salah 3.676
El-Arish 4.237
El-Arish Arabic 4.240
elative 1.14, 1.17–18, 1.76, 1.108, 1.251, 1.304, 1.395, 1.550, 1.565, 1.643, 2.12–16, 2.246, 2.411, 2.424, 2.485, 2.489, 2.594, 2.616, 2.631, 3.171, 3.244, 3.280, 3.431, 3.443–444, 3.486, 3.532, 3.536, 3.631, 4.95, 4.617, 4.644, 4.691
elative, feminine 1.490
El-Ayoubi, Hashem 3.426
Elazig 2.607
Elbert, Mary 2.688
Elbinck, van 1.8
El-Dalee, M.S. 4.670–671
El-Dash, Linda 2.652, 2.656, 2.660
e-learning 4.155
electronic technology 3.500
element, redundant 4.436
elementary prosodic unit 3.208
elementary school 1.655
elements, Demotic 4.543
elementum 1.148
Elephantine Island 2.1
El-Gemei, Dalal Mahmoud 1.434–439
Elgendy, Ahmed M. → El-Gindi Ahmad A.
Elgibali, Alaa 1.409, 1.668, 2.672, 3.59, 3.642
El-Gindi, Ahmed A. 3.604, 3.690, 3.697, 4.669
Elhakim, Gehan H. 3.202
El-Halees, Y. 4.669
El-Hassan, Shahir A. 1.435, 1.572, 1.630–632, 1.668, 2.11, 2.399, 2.501, 3.262–269, 4.62, 4.325, 4.650
El-Hijr 4.478
Elias, Antun Elias 3.26–27
Elihai, Yohanan 4.534
El-Imam, Youssef 2.397
elision 1.226, 1.235–236, 1.243, 1.245, 1.262, 1.321–322, 1.383–385, 1.391, 1.548, 2.5, 2.16–18, 2.20–21, 2.223, 2.300–301, 2.366, 2.529, 2.591, 3.117, 3.148, 3.530, 4.193, 4.243, 4.447, 4.688
elision, of *a* 4.404
elision, segment 4.688
elision, vowel 1.300, 2.612, 3.568, 4.674–675
elision in pausal position 4.674

El-Karak Arabic 2.506
Elkhafaifi, Hussein M. 2.638, 2.699, 2.705, 3.20
Ellaty, Youhanna 3.355
Elliott, H.M. 3.506
ellipsis 1.430, 1.474, 1.651, 2.18–21, 2.300–301, 2.366, 3.34, 3.122, 4.193–194, 4.448, 4.646–647
ellipsis mark 3.741
ellipsis of verb phrase 2.21
Ellis, Andrew 2.689, 4.271
Ellis, Rod 2.117, 2.121
Elman, Jeffrey 2.689
Elmdari, Fouad 2.621
Elmedlaoui, Mohamed 3.493, 3.624, 3.626, 4.286
Elmgren, Sven Gabriel 3.403
eloquence 1.428, 2.304–306, 2.441, 2.579, 3.160
ELRA 1.215
El-Sayed, Ali 3.658, 3.680
El-Shiyab, Said 2.19
ELSNET 1.215
Elusa 3.467
Elwell-Sutton, Lawrence P. 3.576
Elyaacoubi, Mohammad 3.287, 3.291–292
El-Yasin, Mohammed K. 2.501, 4.655
Elzeiny, Nagwa 2.202–207
e-mail 1.455, 1.600, 3.274, 3.317, 4.516
e-mail, transcription in 4.516
Emam, Ossama 1.210
Embarki, Mohamed 1.371
embedded language 1.416, 2.77
embedding 1.417, 2.118
eme.sal 4.33
Emeneau, Murray B. 4.372
Emerson, L.H.S. 4.750
Emery, Peter G. 1.437, 3.658, 3.681
Emesa 1.178
emigration 2.60, 3.294, 3.313, 4.91
Emin Pasha 1.519–520, 2.588
Emirates, Arabic in the 2.609
Emonds, Joseph E. 2.370, 3.388, 4.524, 4.741
emotional modal feature 2.144
Empeirikoi 4.11
emphasis 1.205–206, 1.231, 1.260, 1.276, 1.278, 1.302, 1.440–441, 1.569, 2.241, 2.365, 2.459, 2.482, 2.503, 3.65, 3.124, 3.239, 3.275, 3.711, 4.171, 4.421–424, 4.669–671
emphasis, loss of 1.525, 2.269, 2.461, 2.558, 2.560, 3.62, 3.146, 3.643, 4.754
emphasis, secondary 1.324, 2.3–4, 4.710
emphasis blocking 2.365
emphasis spread 1.105, 1.260, 1.278, 1.302, 1.325, 1.548, 2.365, 2.482, 2.529, 2.612, 3.65, 3.181, 3.276, 3.363, 3.481, 3.527, 3.529, 3.600, 3.610, 4.118, 4.637
emphasizer 1.202–203, 1.555–556, 1.559, 4.221, 4.236
emphatic 1.342, 1.376, 1.539, 1.544, 1.547, 2.97, 2.269, 2.327, 2.560, 3.180, 3.190, 4.172, 4.637
emphatic, marginal 3.600
emphatic, primary 4.637
emphatic, secondary 4.637
emphatic consonant 2.529, 3.597, 3.599–601, 3.608, 4.670
emphatic state 3.220
emphatics, disappearance of 1.29
emphatization 2.483
Empson, William 4.200
Empty Category Principle 2.171–174
empty pronoun 3.709
Empty Quarter 3.693, 4.123
empty subject position 3.709
ʿĒn Avdat, inscription of 1.149
enaciados 1.98
enantiosemantics 4.167
Enç, Mürvet 2.371, 4.316
enclave → island, linguistic
enclave, linguistic 1.590
enclave, relic 1.590
enclaves, innovation in 1.590
enclisis 2.21–22
enclitic 2.21, 3.298, 4.116
enclitic suffixing 4.250
End Rules 3.617
endearment, terms of 1.227
Endegen 4.301
endocentricity 4.741
endogamy 1.388, 1.446, 3.2, 4.580
endophoric 1.568
Endress, Gerhard 2.199–201, 3.177, 3.426, 3.453, 3.719–721, 4.13, 4.542–546
energeticus → energicus

energicus 1.38–39, 1.202, 1.310, 2.22–25, 2.259, 3.250, 3.270–273, 3.298, 3.565, 3.573, 4.121, 4.178, 4.341, 4.640
energicus in Northwest Semitic 3.416
Engaddi Papyri 1.149
Engelhardt, Miriam 3.393
England, Liz 2.110
English 1.24, 1.52, 1.176, 1.197, 1.200–201, 1.222, 1.231, 1.233, 1.241, 1.260, 1.274, 1.278, 1.333, 1.346, 1.353–354, 1.370, 1.407, 1.448, 1.471, 1.473, 1.475–476, 1.484–485, 1.494, 1.507–508, 1.514, 1.529, 1.537, 1.574, 1.625, 1.630, 1.649–652, 1.654, 1.660–662, 1.665, 2.8, 2.25–29, 2.101, 2.114, 2.216, 2.286, 2.326, 2.332–333, 2.338, 2.345, 2.374, 2.393, 2.395–396, 2.399–400, 2.430, 2.482, 2.491, 2.525, 2.560, 2.609, 2.699, 3.11, 3.62, 3.69, 3.81, 3.84, 3.86, 3.99, 3.129, 3.144, 3.147, 3.224–225, 3.228, 3.232, 3.316, 3.560, 3.706, 3.725, 4.53, 4.56–57, 4.197, 4.438
English, African American Vernacular 2.275
English, American 1.436, 2.276, 2.398–399, 3.18, 4.628
English, British 3.18
English, British Black 2.275
English, Cameroon Pidgin 1.333
English, connectives in 1.470
English, Jamaican Creole 2.319–320
English, Krio Creole 4.198
English, left dislocation in 4.525–527
English, Middle 3.606, 3.640
English, Old 3.206
English, Pidgin 1.619
English, shared constituents in 4.438
English, South African 2.276
English, Standard 2.320
English, syllable structure in 4.388–389
English, teaching of 2.338
English, topicalization in 4.525–527
English, word order in 4.394
English, plural in 3.154
English in Egypt 2.699
English in India 4.369
English in Jordan 2.504
English in Palestine 2.699
English in Saudi Arabia 4.123
English in Sudan 2.699
English in the Gulf States 2.700
enjambment 3.499, 3.646
Ennaji, Moha 1.456, 1.488, 2.389–395, 2.642, 2.659, 3.287–296, 4.185–191
Ennedi 1.360
Ennemore 4.301
enrichment of data 1.513
Entwistle, William J. 1.99
Epalza, Mikel de 1.98, 4.572, 4.574
EPEC 1.211
epenthesis 1.85, 1.243–245, 1.539, 1.547, 2.35–39, 2.128, 2.612, 3.181, 3.412, 3.480, 3.494, 3.528–530, 3.611, 3.615, 3.622, 4.85–86, 4.108, 4.110, 4.116–118, 4.344–345, 4.348, 4.351, 4.565, 4.710–711, 4.716
epenthetic schwa 1.244
epenthetic vowel 1.254, 2.242, 2.253, 2.353, 2.374, 3.299
epicene adjective 1.14
Epicureans 4.11
Epie, Ahmed Ibrahim Nzube 1.336
epiglottis 3.594, 4.666–667
epigraphic script 1.151
Epigraphic South Arabian → South Arabian, Epigraphic
epigraphy 2.40–50
epigraphy, monumental 2.601
epigraphy, Old Arabic 3.464–477
Epiphanius of Eleutheropolis 3.467
epistemic 2.144, 3.234–235
epistemic modality 1.248
epithet 3.632, 3.733
Epstein, Eliahu 2.498, 2.503
equational sentence 1.197, 1.349, 2.432, 3.81, 3.590–591, 4.745
Equatoria Province 2.588
Equatoria Region 2.517–518
Equi 4.522
Equipotentiality, Theory of 2.690
equivalence 2.319
equivalence constraint 1.416, 1.418
Eratosthenes 1.127
Eren, Hasan 4.588
ergative 1.38, 1.563, 1.644

Ergin, Osman 3.503–504
Eritrea 1.654–656, 2.268, 2.270–274, 2.454, 3.60, 3.62, 3.312
Eritrean People's Liberation Front 1.655
Erman, Adolf 2.68
Ermers, Robert 2.235, 2.404–405, 2.557–559, 3.22, 3.72
erosion, phonological 4.161
Erpenius, Thomas 1.8, 1.169, 1.170, 1.172, 4.516
Errihani, Mohammed 2.659, 3.294
error analysis 1.485, 4.139–140
error, aphasic 4.271
error, between-word 4.337
error, competence 4.140
error, gesture 4.335
error, grammatical 4.335–336
error, learners' 4.139
error, lexical 4.268
error, morpheme 4.336
error, performance 4.139–140
error, phonological 4.335–336
error, retrieval 4.270
error, root-and-pattern 4.337
error, segmental 4.268–270, 4.335–337
error, semantic 2.692, 4.335
error, stranding 4.269
error, substitution 2.680, 4.270
error, within-word 4.337
Ersen, Margarette 4.587
Ervin-Tripp, Susan M. 2.202
Erwin, Wallace M. 2.96, 2.388, 2.390, 2.415, 3.495, 3.614, 3.654, 3.729, 4.318
Erythraic 1.35
Esarhaddon 4.480
ʾəṣāṣā 2.232
Esling, John H. 4.666–667
Esna Arabic 2.5
Espinosa, Aina 3.524
Esposito, John L. 3.346
Ess, Josef van 3.551, 4.684
Essaouira Arabic 3.288
Essawi, Hala 2.332
essay writing 1.432
Esseesy, Mohssen 1.123–126, 2.191–198, 4.37–43, 4.160–169
Estefi 1.379
estimative 3.251
ethical dative 1.308
Ethiopia 1.37, 2.51–56, 2.250, 3.60, 3.62, 3.312, 4.508
Ethiopian languages 1.639, 2.51, 2.53–55, 2.262, 2.560, 3.420, 3.446, 3.641
Ethiopian Semitic → Ethio-Semitic
Ethiopic 1.638–639, 2.199, 3.189, 3.417, 4.170–171, 4.175–177, 4.371, 4.373, 4.479
Ethiopic, Classical 2.56, 4.174, 4.177
Ethiopic, Old 4.175, 4.301
Ethio-Semitic 1.37, 1.489–490, 1.624, 3.494, 3.553, 3.592, 4.1, 4.170, 4.301–307, 4.309–310, 4.312–314
ethnicity 1.443, 2.58–65, 4.633
ethnography of speaking 2.202
ethnography, linguistic 2.653
ethnolect 2.217
etumología 2.66
étumon 1.575, 2.66, 3.46–52, 4.97, 4.288
etymological fallacy 2.68
etymology 1.194, 1.575, 2.65–69, 2.447, 4.287
etymology, external 2.66
etymology, internal 2.66
Eubank, Lynn 4.143
eulogy 1.3, 3.661
euphemism 1.656–657, 1.659, 2.69–72, 2.342, 3.448, 3.632, 3.662, 3.733, 4.165–167, 4.416–417
euphemizer 2.72
Euphrates 2.571
Europe 2.72–79
Europe, Arabic in 2.665, 3.313
European Language Resources Association 1.512
European languages 1.531
European, Standard Average 1.493
evaluation 1.202, 3.235
evaluative 3.234
evaluative act 2.80
event 1.195, 1.197, 1.200–201, 2.114
event, countersequential 2.209
event, sequenced 2.208
event affix 1.352, 4.488
event nominal 3.388–393
event time 1.199–200

eventiveness 2.367
eventuality 4.657
Everett, Daniel 2.668
Evers, Vincent 3.525
evidential 1.396
evolution, parallel 1.594
Ewald, Georg 3.207
Ewe 4.282
Ewondo 1.333
exactness of copying relation 3.492
exact-probability test 2.322
exaptation 3.577
exception 2.451–452
Exceptional Case Marking 4.54, 4.56, 4.538–539
Exceptional Case-Marked Contexts 1.412
exchange 2.202
exchange, lexical 4.269–270
exchange of roots 4.336
exclamation 1.106, 1.192, 1.249, 1.279–280, 1.548, 2.79–81, 2.94, 2.144, 2.195, 2.360, 2.397, 2.615, 2.631, 3.278, 3.485, 3.532
exclamation mark 3.741
exegesis 2.327
exegesis, Jewish 2.579
exegesis, mystical 4.75
exegesis of the *Qur'ān* 2.578, 2.580, 3.34, 3.36, 3.122, 3.551, 4.73, 4.235, 4.329
Exhaustivity Condition 3.619
exhortation 3.536
exhortative 1.246, 3.248
existence 1.248
existential 1.249, 1.280, 1.392, 1.524, 1.549, 1.554, 1.564, 2.615, 3.80–84, 3.86–87, 3.186, 3.202, 3.278, 3.365, 3.485, 4.112, 4.318, 4.566, 4.733–734
existential clause 1.285
existential construction 3.675
existential copula 2.273
existential marker 1.106, 3.81–82, 3.87
existential sentence 2.249, 3.187, 4.358, 4.486, 4.747
exocentricity 4.741–742
exophoric 1.568
expansion 1.342, 1.492, 3.120
expansion, valency 4.626
experience 4.529
experiencer 1.40, 344–346, 4.535
experiential perfect 1.248
experiential view 1.423
expletive 1.45. 3.84–85
expletive 'there' 3.81
expletive element 2.627
expletive hypothesis 4.363–364
expletive pro 2.174
expletive pronoun 1.45
explicative 1432, 4.221
expository 4.473
expository text 1.474, 1.649, 1.651
expression 3.160
expression, circumstantial 2.209
expression, concealed 2.578
expression, explicit 2.578, 2.581
expression, figurative 3.117
expression, formulaic 2.202
expression, indexical 1.565
expression, indirect 2.578
expression, metaphorical 3.32
expression, referring 3.710
expression, ritualistic 2.585
expression, spatio-temporal 2.209
expression, tropic 2.579
Expression Rules 2.146
expressive 3.570, 4.329
expressive act 4.331
expressive renewal 2.351
expressiveness 2.79
Extended Projection Principle 4.365
Extended Standard Theory 1.457, 4.521
extension, metaphorical 2.218
extension, semantic 4.164–169, 4.593
External Determiner Hypothesis 4.747
external influence 1.594
Extra, Guus 3.315
extraction 1.413, 2.171–172, 2.393, 3.686, 4.363–364, 4.648, 4.721
extraction, subject 2.173
extra-linguistic method 1.590
extra-linguistic variable 1.440
extrametricality 3.616–617, 3.619–621, 4.345–348, 4.351–352
extrasyllabic 3.613
Eža 2.67, 4.301, 4.04–305
Ezha → Eža
Ezzat, Ali 1.435

F

/f/ 4.302
fa- 1.71, 1.467–468, 1.473–476, 1.566, 1.648, 2.22, 2.208, 2.478–479, 2.542, 3.200, 3.225, 3.646, 3.654, 4.192, 4.194, 4.486, 4.544, 4.729
fa-, resumptive 1.472
faʿal 2.424
faʿala 3.224–225
faʿāli 2.379, 2.432–433, 3.35
Faber, Alice 1.616, 3.420, 4.313
fable 3.734
face affront 1.657
face saving 1.657
facial expression in sign language 4.229
factitive 1.553, 1.623, 1.626, 2.348, 4.51, 4.307, 4.697
factitivity 1.329
Fadak 1.129
Fadicca 2.1, 3.60
Fāḍil, Yūsuf [Yusef Fadel] 1.603, 3.295
fading 2.192
Faḍl al-Mawlā 2.588
Faḍl, Ṣalāḥ 4.473
faḍla 2.225, 2.291, 2.430, 3.100, 3.102–103, 3.105, 4.413, 4.436
Fadli 4.750
faʾfāʾ 2.676–677
Fafunwa, Aliyu Babs 3.374
Fagborun, J. Gbenga 4.758–759
Fahahil 2.609
Faḥḥām, Šākir al- 2.635
fāḥiš 2.362
Fahmi, Mahir Hasan 2.634
faḥš 2.362
fāʾida 1.2, 1.4, 2.186, 2.536–538, 2.542, 4.413, 4.738
fāʿil 1.40–41, 1.182, 1.436, 1.488, 1.556, 1.624, 2.82–84, 2.92–95, 2.315–316, 2.356, 2.427, 2.429–431, 2.434, 2.436, 2.537–539, 2.550–551, 3.100, 3.224, 3.424, 3.683, 3.685, 4.354, 4.410–412, 4.487, 4.625, 4.737, 4.739
Failaka 2.212, 2.608
Failed Functional Features Hypothesis 4.144
Faiq, Said 1.427–430, 4.471–476
faithfulness 3.492–493, 3.496–497
Faithfulness Condition 3.619
Fakhri, Ahmed 1.484–485, 1.647–653, 4.141, 4.656
fakk al-ʾidġām 4.10
faʿlān 1.197
fallacy of perfection 3.492
fallāḥī 1.605
fallāḥīn 2.2, 2.498, 2.501
Fallāḥiyyah 2.572
falling contour 2.397, 2.400
falling melody 2.397
family names 2.71, 2.586
family tree model 1.490
Fanselow, Frank S. 4.433
Fansuri, Hamzah 2.334
fa-qad 1.474
Faqʿasī, ʾAbū Ḥayyān al- 4.34
farʿ 1.193, 2.310, 2.408, 2.539, 4.13, 4.422, 4.443, 4.560, 4.739
far deixis 1.106, 115
Far North Province 1.334
Fārābī, ʾAbū Naṣr al- 2.85, 2.90, 2.235, 2.425, 2.435, 3.32–34, 3.73, 3.113–115, 3.176–177, 3.547, 4.219, 4.235, 4.459

Farāfra 1.589
Farāfra Arabic 1.586, 1.610–611, 2.2, 2.6, 2.8
Farah, Rola 2.101
Faraj 2.44
Faraj, ʾAlfirīd 1.601
Faraj, Faraj as-Sayyid 1.601
Faraj, Samīḥ 1.599
faraka 4.133
Farazdaq, al- 2.94, 2.543, 3.650, 4.200, 4.202, 4.209
Farber, Bernard 2.583
farḍ 1.189
farḍ ʿayn 2.337
Fareh, Shehdeh 2.209, 4.489
Farghal, Mohammed Ali 1.656–659, 2.69–72, 2.169–172, 2.501, 2.583–588, 3.678, 3.681, 4.365, 4.467, 4.487, 4.742
Farghali, Ellen 3.676
Farghaly, Ali Ahmed 1.511, 4.45–50, 4.742, 4.746
Fargues, Philippe 2.59, 4.577
Farḥāt, Germanus [Jarmānūs] 1.174, 2.189, 2.497
Farḥat, Ilyās Ḥabīb 3.4
Faridpur 4.508
fārisī 3.582
Fārisī, ʾAbū ʿAlī al-Ḥasan ibn ʾAḥmad al- 1.173, 2.82, 2.188, 2.235, 2.406, 2.538, 3.103, 3.321, 3.649, 4.9, 4.119–200
Fārisī, ʾAbū l-Ḥusayn al- 3.122
Farnetani, Ena 4.669
Farooqi, F.U. 2.326
Farrāʾ, al- 1.49, 1.70, 1.73–74, 1.402, 1.618, 1.641, 2.79, 2.178, 2.227, 2.232, 2.294, 2.298, 2.301, 2.316, 2.402, 2.426, 2.436, 2.478, 2.536, 3.14, 3.72, 3.91, 3.101, 3.107, 3.109, 3.116, 3.123, 3.549, 3.689, 4.8, 4.119, 4.219–220, 4.235, 4.431, 4.436. 4.447
Farrugia, G. 3.154
Farruxī 2.410
Farsi 1.222, 1.231, 1.259, 1.403, 2.40, 2.406–409, 2.411–412, 2.673, 3.313, 3.574, 3.576, 4.424
Faršīdvard, Khosrow 2.410–411, 3.578
Fārūqī, Zubayr ʾAḥmad al- 2.326
Faruqi, Harith Suleiman 3.26
Farwaneh, Samira 2.35–39, 3.305, 3.494, 3.611, 4.85
fasād al-luġa 1.615
faṣāḥa 1.429, 1.636, 2.84–88, 2.304–305, 4.202, 4.209
fāsī 3.112
faṣīḥ 1.131, 1.402, 1.530, 2.84–90, 2.257, 2.629, 2.632, 3.15, 3.92, 4.22
fāṣila 3.208–209, 3.214, 3.741, 4.103
fāṣila manqūṣa 3.741
faṣl 1.429, 1.471, 1.558–559, 2.292, 4.472
Fašn Arabic 2.2
Fasold, Ralph W. 1.619, 2.58–59, 2.275, 2.319, 2.321, 2.652, 2.707, 3.19
fassara 4.436
Fassberg, Steven E. 2.24
Fassi Fehri, Abdelkader 1.45–46, 1.350, 1.358, 1.457, 1.480–481, 1.486, 1.488, 1.507, 1.510, 1.580, 2.173, 2.195, 2.367, 2.369, 3.168, 3.225, 3.232, 3.356, 3.382, 3.384–385, 3.388, 3.393, 3.428–429, 3.431–433, 3.707, 3.709–710, 4.15, 4.42, 4.365–366, 4.394–396, 4.398, 4.400, 4.486–489, 4.524, 4.527, 4.648, 4.650, 4.653–655, 4.657, 4.745–746
faṣuḥa 2.85
Fatani, Afnan H. 4.288
fatḥ 1.152, 1.309, 2.92, 2.232–234, 2.402, 4.10
fatḥ al-ʿuqad 3.341
fatḥa 1.560, 2.232, 2.311
fatḥa maṣūba bi-kasra 2.233
Fātiḥa, Sūrat al- 1.529
Fāṭima 2.47
Fāṭima bint ʿAbd aṣ-Ṣamad 4.275
Fatimi, S. Qudratullah 2.340
Fatimid Dynasty 4.218, 4.572
fatla 4.211
Fatlī, ʿAbd al-Ḥusayn M. al- 1.614–615
Fatwa Academy 1.450–451
Fauconnier, Giles 3.653
faʿula 4.412
Faulkner, Raymond O. 2.68
Fawaxriyyah Arabic 4.240
Faxruddin Dynasty 4.276
Fāxūrī, Ḥannā 3.71
Fayḍī 2.327

Fayfa Arabic 4.123
Fayrūzābādī, al- → Fīrūzābādī al-
Fayyāḍ, Sulaymān 4.61
Fayyum Arabic 1.320–321, 1.608, 1.611, 1.640, 2.2, 2.5, 2.494
Fayyūmī, al- 3.32, 3.42
Faẓāʾilī, Habib Allah 4.138
Fazāṛa 1.299
Fazlıoğlu, İhsan 3.501–502
Fazlıoğlu, Şukran 3.503
feature 3.228–229
feature, localized 4.631
feature, Φ- 2174, 4.745–746
feature, supralocal 4.631
feature, syntactic 4.741–742
feature checking 4.745
feature frequency method 1.586
feature geometry 3.627, 3.463–464
Feature Implication Hypothesis 2.321
feedback, orosensory 2.365
feeding role 2.152
Feghali, Habaka 4.62, 4.124
Feghali, Michel T. 1.275, 1.541, 1.642, 2.156, 3.429, 3.674, 4.403
Feldman, Laurie Beth 3.305
felicity conditions 4.328
Fellag 2.78
Fellata 4.375
Fellbaum, Christiane 4.45, 4.169
feminine 1.15, 1.434, 1.646, 2.155–164
feminine, unmarked 2.158–160
feminine ending 1.263, 1.302, 1.304, 1.326, 1.328, 1.392, 1.490, 1.497, 1.538, 1.547, 1.550, 1.638, 2.156–160, 2.312, 2.343, 2.420, 2.509, 2.557, 2.606, 2.645, 4.565, 4.627
feminine ending, pausal form of 3.568
feminine ending in Northwest Semitic 3.415
feminine ending in Persian 3.575, 3.577–578
feminine ending in Tajik 4.425
feminine ending in Tamil 4.434
feminine ending in Tatar 4.451
feminine ending in Turkish 4.585
feminine ending in Uyghur 4.606
feminine ending in Uzbek 4.611
feminine marker 1.400, 1.479, 1.556, 2.101, 2.250, 2.583, 3.154, 3.183, 3.278, 3.303, 3.430, 4.64, 4.244
feminine marker, Berber 2.252
feminine marker, Hausa 2.253
feminine marker, pausal form of 3.565–566
feminine marker, raising of 4.243
feminine marker, Romance 3.154
feminine plural 1.393
feminine reading 2.314
feminine suffix 1.78, 1.646, 3.220
feminism 2.642
feminist literature 1.599
feminist movement 2.648
Fenton, Paul B. 2.526, 2.579
Feraud, Laurent-Charles 4.548
Ferber, Andrew 2.202
Ferdowsi → Firdawsi
Fergani, Nader 2.705
Ferghana Valley 4.603
Ferguson, Charles A. 1.173, 1.340–342, 1.419, 1.439–440, 1.442, 1.447, 1.490–491, 1.533, 1.593–595, 1.615–616, 1.618, 1.629–631, 1.634–635, 1.667, 1.669, 2.10, 2.29, 2.51, 2.104, 2.117, 2.162–163, 2.203–204, 2.267, 2.501, 2.642, 2.654–655, 2.658, 2.695, 2.708, 3.10, 3.62, 3.70, 3.238, 3.346, 3.348, 3.444, 3.446, 3.451, 3.640, 3.647, 3.658, 3.664, 3.670, 3.726, 4.50, 4.57–59, 4.150, 4.399, 4.423, 4.655, 4.669
Ferrando, Ignacio 1.433–434, 2.61, 2.261–268, 3.237, 3.439
Fes 1.24, 3.135, 3.294, 4.156
Fes Arabic 1.444, 1.571, 1.596, 1.609, 2.532–534, 2.621, 3.273–274, 3.276, 3.285, 3.288–290, 3.292–293, 3.585
Fes Arabic, Jewish 1.84
Fes lithographs 3.112
Fəsken 1.91
Fevrier, Paul Albert 4.572
Fezzan 1.273, 1.361–362, 2.552, 3.53, 3.56, 4.376
Fezzan Arabic 1.265–266, 2.622, 3.55–56, 3.587, 4.551, 4.555

Fezzanis 1.362

fī 2.273, 3.109, 3.675, 4.544

Fibonacci, Leonardo 2.455

Fierro, Maribel 3.379

Figali, Hanna al-Xuri al- 1.603

figura etymologica 2.496

figurative expression 2.70, 2.443

figurative language 2.305, 3.651

figurative sense 1.437

figure of speech 3.629

figure/ground 1.423

figure/ground organization 4.501, 4.506

fiʿl 1.309, 2.82, 2.90–96, 2.184, 2.225, 2.227, 2.424, 2.434, 2.537–539, 2.542, 2.545, 3.241, 3.424, 3.546, 3.550, 3.683, 3.685, 4.354, 4.414, 4.487, 4.638, 4.739

fiʿl → verb

fiʿl al-ʿibāra 2.551

fiʿl al-ʿilla 1.576

fiʿl al-ʾinšāʾ 2.193

fiʿl al-lawn 1.576

fiʿl al-madḥ wa-ḏ-ḏamm 1.73, 2.94

fiʿl al-mūqaraba 1.70, 2.94, 3.105, 3.234, 3.354

fiʿl al-mutakallim 2.360

fiʿl al-muṭāwaʿa 4.414

fiʿl al-qalb 1.71, 1.203, 2.93, 2.186, 3.166, 4.56, 4.429

fiʿl al-qasam 4.414

fiʿl ar-rajāʾ 3.105

fiʿl aš-šakk wa-l-yaqīn 3.354

fiʿl aš-šurūʿ 2.94, 2.193

fiʿl aš-šuʿūr 3.105

fiʿl at-taʿajjub 1.564–565, 2.80, 2.94, 3.550

fiʿl at-tafḍīl 1.69, 2.13, 3.105

fiʿl at-taḥwīl 3.235

fiʿl dāʾim 2.431, 3.97, 3.543, 3.549, 4.455

fiʿl fāʿil 4.415

fiʿl fī l-lafḏ̣ 2.551

fiʿl mafʿūl 4.415

fiʿl muḍāriʿ 2.431, 2.625, 3.96, 3.98

fiʿl mutaʿaddin li-ġayrihi 2.94, 4.531

fiʿl mutaʿaddin li-ġayrihi

fiʿl nāqiṣ 1.70, 1.217, 2.549–551

fiʿl tāmm 2.549–550

Filipino 3.313

fiʿliyya 2.186, 2.434

filler 1.457

Fillmore, Charles J. 1.41, 1.343–344, 1.346, 1.457, 4.487, 4.502, 4.624

film 2.11

film, Indian 2.211

finality 1.306

Finck, Franz N. 2.217

Findlow, Sally 2.694, 2.705

finger spelling 4.225

finite state 1.457

finite state automaton 1.208, 1.456

Finnish 2.684, 4.53, 4.194

fiqh 1.336, 1.338, 1.380, 1.427, 1.451, 2.360, 2.543, 4.11, 4.685

Firanescu, Daniela Rodica 2.79–81, 2.360, 3.233–238, 4.196, 4.198, 4.283, 4.328–334

Firbas, Jan 4.494, 4.502

Firdawsi 2.406, 3.575, 3.720,

Firoz ud Din, Al Haj Maulavi 4.601

First Cataract 2.1

first language acquisition 1.339, 1.342, 1.368–376, 1.482, 2.96–104, 3.724, 4.37, 4.140

first language attrition 1.370

first language teaching 2.104–113

Firth, J.R. 1.435

Firth, Raymond 2.202, 2.204–205

Fīrūzābādī, Muḥammad ibn Yaʿqūb al- 1.3–4, 2.581, 2.597, 3.21–22, 3.33, 3.40, 3.42, 3.50, 3.172, 3.321, 3.504, 3.693, 4.103

Fischer, August 2.638, 3.23, 3.43, 3.453, 3.674

Fischer, J.B. 3.547

Fischer, Wolfdietrich 1.16–21, 1.383, 1.397–405, 1.562, 1.571–572, 1.609, 1.614–615, 1.618–619, 1.638–639, 1.642, 2.195, 2.230–231, 2.236–237, 2.296, 2.388, 2.452, 3.190–191, 3.216, 3.237, 3.404, 3.410, 3.422, 3.424, 3.429, 3.444, 3.453–454, 3.553, 3.556, 3.568, 3.593, 3.609, 3.700, 3.704, 3.718–721, 4.2, 4.40, 4.112, 4.197–198, 4.404, 4.520, 4.613–614, 4.616–618, 4.621–622

Fishman, Joshua A. 1.636, 2.58, 2.652, 3.343

f-ʿ-l notation 4.446

f-languages 4.314
Flashner, Vanessa 2.208
Fleisch, Henri 1.560, 1.566–567, 1.573, 1.580–581, 1.584, 1.614–615, 1.638–639, 1.641, 1.643–644, 1.653, 2.22, 2.80, 2.91, 2.94, 2.156, 2.159, 2.165, 2.223, 2.236–237, 2.295, 2.346, 2.348, 2.350, 2.387, 2.426, 2.428, 2.494–495, 3.46, 3.192, 3.269, 3.300, 3.453–454, 3.457, 3.564–565, 3.567, 3.591, 3.645, 3.689, 3.696, 3.700, 3.721, 4.1–2, 4.12–13, 4.93–94, 4.97, 4.339, 4.444, 4.455, 4.486, 4.643
Fleischer, Cornell 3.502
Fleischer, Heinrich Leberecht 1.594, 2.358, 2.360, 2.545, 3.454
Flemming, Barbara 3.502–503
Flemming, Edward 2.365, 3.524
flexible syntax 3.231
Florence 2.454
flower names 3.581
Fo accentual peak 2.397
Fo contours 2.395
Fo peak 2.395, 2.397
Fo valleys 2.395
focality 2.145
focalization 2.115, 4.722
focalizer 3.591
focus 1.485, 2.113–116, 2.131, 2.143, 2.145, 2.147, 2.321, 3.260, 3.383, 3.387, 3.672, 4.358, 4.400, 4.484, 4.496, 4.502–503
focus, argument 2.113
focus, contrastive 2.143, 2.145, 2.147, 4.399
focus, new 2.145
focus, predicate 2.113
focus, replacing 2.145
focus, requestive 2.143
focus, restricting 2.145
focus, selecting 2.145
focus, sentence 2.113
focus domain 2.113, 2.116, 4.502–503
focus element 4.734
focus marker 2.115
Focus Phrase 4.488
focusing 4.525–527, 4.729, 4.731–732
Fodor, I. 4.750, 4.753–754
Fodor, Janet Dean 4.316
Fodor, Jerry 2.691
Foley, J. 2.165
Foley, William A. 4.197
folk etymology 1.575, 2.68, 2.285
folk linguistics 1.173
folk literature 2.531
folk poetry 1.323
folk story 1.597
Folmer, M.L. 3.410
Fontinoy, Charles 3.736
foot 3.613, 3.620
foot, binary 3.616
foot, degenerate 4.346, 4.348
foot, iambic 3.613–614, 3.620, 4.389
foot, left-dominant 3.613, 3.616
foot, metrical 2.238, 3.616
foot, poetic 4.89, 4.92
foot, right-dominant 3.613, 3.616
foot, trochaic 3.613, 3.620, 4.390
foot assignment rule 3.617
foot dominance 3.616
Force 4.360
foreground 1.485, 2.207–209, 4.499, 4.501, 4.728–729, 4.731, 4.733
foregrounding 1.422–423
foreign accent 3.725
foreign language, fluency in 2.9
foreign language learning 1.432
foreign phonemes 1.260
foreign word 2.71
foreigner talk 2.116–122, 3.63, 3.640
form, derived 1.237–239
form, fundamental 4.341–342
Form II 4.51, 4.307, 4.532
Form III 1.576, 4307
Form IV 1.93, 1.108, 1.611, 1.626, 2.277–278, 2.348–349, 2.421, 2.486, 2.565, 2.576, 2.617, 2.631, 3.184, 3.221, 3.251, 3.367, 3.403, 3.488, 3.532, 4.34, 4.45, 4.113, 4.241, 4.307, 4.532, 4.536, 4.642
Form VI 4.57
Form VII 4.57
Form IX 1.576, 4.307, 4.341
formal speech 2.647
formal style 2.323
formality 1.650, 2.661, 3.664
Formangleichung 1.577

formant frequency 4.672
formant frequency in deaf speech 2.678
formative, prefixed 4.462
formula 1.452, 1.473, 1.476
formula, emergent 1.474
formula, frozen 1.266
formula, prior-text 1.474
formula, religious 2.647, 2.649, 4.75
formulaic element 3.646
formulaic expression 1.268, 3.570, 4.141
formulaic pattern 1.648
formulaic schema 3.646
formulaic technique 3.499–500
formulaicity 3.646
Forrest, Karen 2.688
Forster, Kenneth I. 1.425
Fort Lamy 1.363
fortis 3.123, 3.125, 4.2
fossilization 1.485, 1.554, 1.565, 1.615
fossilized form 4.44
Fotokol 1.334
Foumban 1.337
founder of language → creation of language 2.177
Fournier, Jean-Marie 2.177
Fox, Andrew 4.288
Fox, Barbara A. 2.208
Fox, Joshua 1.638, 4.176
fracture 1.642
Fradkin, Robert 3.269–273
Fraenkel, Siegmund 3.6–7, 3.719
Fragner, Bert 3.501
Frajzyngier, Zygmunt 4.44
frame 1.422–423, 4.502, 4.624
framed slot 1.476
framing 1.475
France 2.73
France, Arabic in 2.73–75
Francophonie 4.574
Frank, Richard M. 4.329
Frayḥa, ʾAnīs 1.176, 2.701, 3.732
Frederick II, King 4.216
Free Morpheme Constraint 1.415–416
free state 1.350–351, 1.479
Freed, Barbara F. 2.681
Freeman, Andrew 1.460
Freeze, Ray 3.80, 3.83–86
Frege, Gottlob 2.368
Freller, Thomas 3.146
French 1.55–56, 1.58–60, 1.176–177, 1.222, 1.231, 1.233, 1.260, 1.274, 1.278, 1.290, 1.292, 1.333, 1.335, 1.353, 1.364–365, 1.370, 1.407, 1.411, 1.448–449, 1.451, 1.494, 1.529, 1.534, 1.554, 1.625, 1.630, 1.650–652, 1.654, 2.8, 2.64, 2.114, 2.122–123, 2.286, 2.345, 2.373, 2.392, 2.400, 2.482, 2.504, 2.560, 2.651, 2.659–660, 2.699, 3.16–17, 3.51, 3.60, 3.137, 3.175, 3.312–314, 3.706, 3.724–726, 4.53, 4.285, 4.728
French, Acadian 3.524
French, attitude toward 2.651
French, infinitive in 1.418
French, Jersey 3.204
French, Le Hâvre 3.204
French, Louisiana 2.684
French, teaching of 3.295
French, word order in 4.394
French in Algeria 2.697
French in Chad 3.315
French in Lebanon 2.699, 3.314
French in Mauritania 3.174
French in Morocco 2.698, 3.294, 3.296
French in North Africa 2.697
French in Senegal 4.723
French in Syria 2.699
French in Tunisia 2.698, 4.574–575
frenulum 2.677
frequency 2.449–450
frequency, sound 4.88
frequency distinction 2.677–678
frequency of roots 4.94
frequency of use 2.86–87, 3.43, 3.90
frequency program 1.455
frequentative 1.551
Freud, Sigmund 4.267
Freytag, Georg Wilhelm 1.171, 3.22, 3.207, 3.649, 4.516
frication 1.26
fricative 1.27, 3.596–598
fricative, laryngeal 4.666
fricative, voiced 4.666
fricativization 4.172

Friedmann, Naʾama 2.691
Friedrich, Johannes 3.409
Friggieri, Albert 3.146
Frisch, Stefan A. 3.624–628, 3.463, 3.610, 4.334–338
Frolov, Dmitrij V. 2.236, 2.238, 3.207–215
Fromkin, Victoria 4.268–270, 4.335, 4.337
fronted adverbial 1.476
fronting 1.413–414, 2.5, 2.131–136, 2.145, 2.147, 2.311, 2.688, 3.327, 4.722, 4.734
fronting, topic 3.333–334
fronting, vowel 4.678
fronting of adverbial clause 1.476
fronting of constituent 3.382
fronting of *wh*-constituent 4.719
Frost, Ram 2.692, 3.305
frozen expression 2.136–137
frozen formula 1.266
frozen structure 2.136
Frye, Richard 3.501
Fsadni, Michael 3.142
Fu, Jingqi 3.388
Fuad Library 1.382
fuʿāl 4.168
Fück, Johann 1.400, 1.562, 1.593, 1.614–616, 1.618–620, 2.86, 2.406, 2.628–629, 2.642, 3.59, 3.689
Fujimura, Osamu 3.525
Fukui, Naoki 3.685, 4.742
Ful 1.36
Fula → Fulfulde
Fulani → Fulfulde 1.133, 2.137, 3.373, 4.375
Fulayfil, Muḥammad ʾIbrāhīm 1.382
Fulbe 2.137
Fulfulde 1.133, 2.137, 1.256, 1.333–336, 1.363, 1.545, 2.137–142, 2.252, 2.668–669, 2.673, 3.59, 3.169–171, 3.314, 3.373, 3.375, 4.179–180, 4.182–183, 4.279, 4.281–282, 4.375, 4.716, 4.723
Full Interpretation 3.228–230
full stop 3.741
Full Transfer/Full Access Hypothesis 4.143–144
Fumo (Sultan) Laiti Nabhani 4.381
function word 2.676
function word, omission of 1.340
function words, borrowing of 4.599
function words, underuse of 2.675
functional category 4.742
functional feature 1.510
Functional Grammar 1.457, 2.143–150, 4.726
Functional Grammar, Lexical 1.457, 4.535, 4.657
functional head 1.350–351
functional load 3.603–604, 3.696
Functional Sentence Perspective 4.494, 4.502, 4.726
functional yield 1.616
fundamentalism 2.189
fuqahāʾ 2.178, 2.362
Fur 2.560, 4.375–376
Furley, Oliver 1.662
furūʿ → *farʿ*
fuṣaḥāʾ al-ʿArab 2.632
fuṣḥā 1.173–174, 1.405–406, 1.408–409, 1.419, 1.483–484, 1.527, 1.532, 1.599, 1.603, 1.629–630, 1.632–636, 1.667, 2.84, 89, 2.107, 2.110–111, 2.630, 2.636, 2.641, 2.649, 2.651–657, 2.659, 2.661–662, 2.695, 2.697, 3.18, 3.76, 3.79, 3.344, 3.347–349, 3.465, 4.9, 4.129, 4.252
fuṣḥā, attitude toward 2.651
fuṣḥā, cultivated 2.701
fuṣḥā, learning of 3.76
fuṣḥā al-ʿaṣr 1.533, 3.14
fuṣḥā al-ḥadīṯa, al- 1.533
fusion 1.451, 1.539, 2.591, 4.38
fusion, morpheme 4.37
Fusṭāṭ, al- 1.618
Futayyiḥ, ʾAḥmad al- 2.635–637
future 1.199–200, 1.217, 1.306, 1.330–331, 1.377, 1.383, 1.508, 1.568, 2.91, 2.272, 2.490, 3.186, 3.265, 3.536, 4.545, 4.569
future intent 1.396, 2.248, 3.285, 3.332, 3.536
future marker 1.186, 2.192, 2.423, 2.514, 2.516, 3.237

future marker *bġa* 1.220
future particle 4.569
future particle *ġadī* 1.568
future particle *ḥa-* 1.148, 1.217, 1.219–220, 1.306, 1.330, 1.408, 1.553, 1.572, 2.272, 2.514, 2.516, 2.568–569, 3.265, 3.406
future particle *raḥ* 1.65, 1.229, 1.281, 1.306, 1.553, 1.568, 1.612, 2.423, 2.490, 2.514, 2.516, 2.577
future particle *sa-* 1.217–218, 1.568, 1.607, 3.97
future, imminent 3.332
Fuuta Jaloo 2.137
Fuuta Toro 4.179–180, 4.182–183
Fuẓūlī 3.502
Fyodorova, El'vira 4.450

G

/g/ 1.324
Gaballa, Hassan 2.169, 3.558, 4.357
Gabelentz, von der 4.161
Gabès 3.53
Gabès Arabic 3.451, 4.681
Gabrieli, Giuseppe 3.379, 3.717, 3.720
Gabučan, Gracija Mikajelovič 2.238
Gacek, Adam 1.1–5, 3.110–113, 3.307–311, 3.336–343, 4.98–100, 4.130–138, 4.560–562
Gadalla, Hassan 3.393
Gaddafi, Colonel Muammar 3.664–666
Gaddo 2.53
Gadoua, Abdulhamid H. 2.230
Gaelic 2.712
Gaelic, East Sutherland 2.684
Gafat 2.67, 4.301, 4.304–305, 4.312
Gafos, Adamantios 3.305, 3.611, 3.626, 4.338–344
gahawa syndrome 1.114, 1.117, 1.243, 1.245, 1.261–262, 1.264–265, 1.272, 1.301, 1.612, 2.2, 2.5, 2.151–153, 2.573, 2.611–612, 3.205, 3.363, 3.367, 3.403, 3.405, 3.478, 3.481, 3.532, 3.609, 4.126, 4.241, 4.244, 4.389, 4.710, 4.715
ġāʾib 1.555, 1.566–568, 1.612, 2.82
Gairdner, William Henry Temple 2.459, 2.494, 3.125, 3.602
Gajar 2.217–219, 2.469
Gajda, Iwona 2.256
Gal, Susan 2.662
Galadanci, Said 3.374–376
Galal, Mohammad 4.60, 4.63, 4.68
galālīf 2.215
Galalo 2.270, 2.272
qalam al-ḥilya 4.134
ġalaṭ 1.124, 2.629, 3.648
Ġalāyinī, Muṣṭafā al- 1.573–574, 3.392, 4.640
Galen 3.72
Galen, Pseudo- 4.134
ġalīḏ̣ 3.340
Galilee Arabic 1.263, 1.265, 1.608, 2.501, 4.663
Galland, Antoine 1.167
Galletti, Mirella 3.351
Galley, Micheline 3.380
Gallop, Annabel Teh 2.344
Galloway, Douglas 3.478
Galmés de Fuentes, Álvaro 1.99–100, 2.286
Gambia 1.255
Ġāmid Arabic 4.126
Gammon 2.217
Gàmmu 4.181
Gandour, Jack 4.268
Ġanī Arabic 3.94
Ganiage, Jean 4.575
Gansu 1.381
Gao 4.279
Gao, Man 4.667
gap 1.590, 2.18–19, 2.369, 4.212, 4.525–526, 4.718, 4.720, 4.722
gap, lexical 1.419, 2.29–30, 2.661
gap, terminological 1.453
gap strategy 4.392
gapping 4.194, 4.397, 4.648–649
Garabulli Arabic 3.55
ġaraḍ 1.428, 4.333

Garārša Arabic 3.401, 3.403, 4.240
Ġarb Aswān 2.1
Garba, Mohamadou 1.336
Garbell, Irene 1.620, 2.276, 2.508, 4.2
Ġarbi, Ḥasan aṣ-Ṣaġīr al- 1.361
Garbini, Giovanni 4.372
Ġarbiyya Arabic 2.2
García Yebra, Valentín 1.617
Garde, Paul 3.725–726
Gardet, Louis 1.628
Gardner, Howard 2.691
Gardner-Chloros, Penelope 1.372
ġarīb 4.203
ġarīb al-ḥadīṯ 3.31, 3.34
ġarīb al-Qurʾān 3.31, 3.34
Gārim, ʿAlī al- 4.87
Garmadi, Salah 1.632
Garoua 1.336–337
Garrett, Andrew 3.204–206
Garrett, Merrill 2.691, 4.268–271
Garrod, S.C. 4.502
Gascon 1.590
Gash-Barka 1.655
Gasim, Awn al-Sharif 3.435
Gasir Arabic 2.2
Gašmū ha-ʿarḇī 1.126
Gass, Susan 2.117
Ġassānī, Ḥassān al- 3.53
Ġatafān Arabic 3.94
Gätje, Helmut 4.543
Gaudefroy-Demombynes, Maurice 2.92, 2.224, 2.45, 4.484, 4.444
Gauffin, Jan 4.668
gav 3.173
Ġawalma 1.362
Gawhar Sad, mosque of 2.601
gaws 2.468, 2.470, 4.157
ġāya 3.109
Ġayl Ḥabbān Arabic 4.750–751, 4.756
Gayo 1.6
ġayr 2.297, 4.41
ġayr al-xabariy(ya) 2.358
ġayr ʿāqil 2.15
ġayr ʿilājī 4.625
ġayr muʾaṯṯir 4.626
ġayr mulāqin 4.412
ġayr munṣarif 1.559, 1.643, 2.403, 2.424, 2.426, 3.453, 4.432
ġayr muṣawwit 2.235
ġayr mutaʿaddin 4.410, 4.624
ġayr mutamakkin 1.309, 1.643, 2.426, 4.431–432
ġayr mutaṣarrif 3.109
ġayr šaxṣ 3.548
ġayr ṭalab 4.332
ġayr ṭalabī 2.360
Gaza 3.401
Gaza Arabic 1.608, 2.32, 3.527, 3.531
Gaza Strip 2.465, 3.360, 3.526
ġazal 3.173, 4.202, 4.596
Ġazālī, ʾAbū Ḥāmid al- 2.304, 2.339, 2.362, 3.42, 3.176–177, 3.321, 3.324, 3.502, 3.578–579, 4.219
Gazdar, Gerald 1.458
Ġazzālī, al- → Ġazālī, al-
GB → Government and Binding
Gbadamosi, T.G.O. 3.374, 4.759
Gdalī, Yaḥyā ibn ʾIbrāhīm al- 3.169
Geblo 4.297
Gee, John 2.162
Geers, F.W. 4.96
Geertz, Clifford 2.63
Geʿez 1.38–39, 1.85, 1.122, 1.203, 1.493, 2.51–52, 2.54–55, 2.66–67, 2.223, 3.414, 3.418, 4.1, 4.301–306, 4.309–310, 4.312–313
Gehin, Paul 3.515
gejmi 4.157
Gelb, Ignace Jay 3.227
Gelder, Geert Jan H. van 2.304–305, 3.646–652
Gelderen, Elly van 4.395–396
gelet → *gilit* Arabic
Gellar, Sheldon 4.723
Gellner, Ernest 2.62–63, 2.647
geminate 1.238, 1.243, 1.277, 1.341, 1.521, 2.153, 2.553, 2.591, 3.530, 3.608, 4.50, 4.688
geminate reduction 2.685, 4.241
geminate verb 1.81, 1.85, 1.109, 1.118, 3.220
gemination 1.207, 1.496–497, 1.525, 1.574, 2.37, 2.153–155, 2.270, 2.298–299, 2.346, 2.349–351, 2.353, 2.367, 3.181, 3.423, 3.635, 4.51, 4.70, 4.279, 4.310, 4.638
gemination, medial 2.154–155

gender 1.43, 1.46–48, 1.386, 1.441, 1.529, 1.643, 2.155–164, 2.424, 2.642–650, 3.66, 3.242, 3.426, 3.429–430, 4.631
gender, feminine 2.643, 4.355
gender, natural 2.156
gender agreement 2.162–163, 4.567
gender crossing 3.450
gender distinction 1.278, 1.326, 1.391, 1.491, 1.570–571, 1.606–608, 1.611–612, 2.213, 2.418, 2.484, 2.510, 2.513, 2.555, 2.583, 2.642, 3.54–55, 3.149, 3.157, 3.181, 3.291, 3.403, 3.530, 4.110, 4.241, 4.304, 4.551, 4.565, 4.631, 4.639
gender distinction, absence of 2.522, 3.634–636
gender distinction, loss of 1.590, 1.596, 2.684
gender distinction in pronouns 3.585
gender distinction in verb 1.270, 1.272, 1.302
gender in Coptic 1.498
gender in Indo-European 2.156
gender marking 2.99, 2.318, 2.320–321
gender of loanwords 4.591, 4.598–599
gender polarity 3.69, 4.274
gender replacement 2.160
gender studies 4.631
gender system in Telugu 4.454
genealogical tree 1.36, 4.371, 4.373
General Phrasal Accent 2.116
General Phrasal Accent Principle 2.113
general term 4.685
generality of meaning 4.22
generalization 2.118, 4.145, 4.162, 4.166
generalization in the *Qurʾān* 4.29
generalization of plural markers 3.63
Generalization, Kayne's 1.412
generalization, semantic 4.160
Generalized Phrase Structure Grammar 1.458
Generative Grammar 1.312, 1.347, 1.422, 1.478, 1.488, 1.507, 3.204, 3.228, 3.232–233, 4.521, 4.741
generator, morphological 1.460
generic 1.248, 1.250, 2.644, 3.247
generic nominal 3.393
generic noun 1.187
generic reference 4.317
Genesee, Fred 1.368, 4.322
genetic relationship 1.294–295, 1.298
genetic relationship of the Semitic languages 4.313
genitive 1.347, 1.351–352, 1.479, 1.559–560, 1.563, 1.580, 1.643, 2.146–147
genitive, analytic 3.634, 3.672–673, 4.247, 4.555
genitive, epexegetic 3.674
genitive, synthetic 3.672–673
genitive construct 2.294
genitive construction 1.292, 1.540, 4.622
genitive exponent 1.82–83, 1.226, 1.491, 1.493, 1.562, 1.611, 2.119–120, 2.161, 2.489, 3.587, 3.671
genitive marker 1.248, 1.279, 1.283, 1.303, 1.327, 1.367, 1.392, 1.478, 1.549, 1.552, 1.607, 1.612, 2.248, 2.420, 2.511, 2.521, 2.575, 2.592, 2.594, 2.614, 3.183, 3.278, 3.291, 3.365, 3.484, 3.637, 4.112, 4.555, 4.714, 4.311
genitive particle 1.263, 3.182, 3.220, 4.565
genitive particle *bitāʿ* 1.596, 1.608, 2.570, 3.390
genitive particle *ḏil(a)* 1.92
genitive particle *dyal* 1.64, 3.278, 3.291
genitive particle *ḥagg* 1.248, 1.611, 2.570, 2.614, 3.182, 3.673
genitive particle *hana* 1.367
genitive particle *ihnīn* 1.303, 1.608
genitive particle *lē* 1.92
genitive particle *līt* 2.420
genitive particle *māl* 1.226, 1.240, 1.248, 1.351, 1.417, 1.607, 2.420, 2.614, 3.672
genitive particle *mtāʿ* 1.64
genitive particle *ntāʿ* 3.291
genitive particle *šel* 1.351
genitive particle *šīt* 1.552

genitive particle *šuġl* 1.608, 3.365
genitive particle *tabaᶜ* 1.279, 1.283, 1.351, 1.549, 1.552, 2.489, 2.511
genitive relation 1.477
Genizah 1.598, 2.529–530
Genoa 2.454
genre, discourse 1.648
genre analysis 1.649
gentilics 2.285
Geoghegan, B. 2.683
geography, dialect 1.583
geography, lexical 1.590
geomantic 3.126
Geonim 2.527
George, Alain 3.514
Georgin, Pierre 1.61–62
Georr, Khalil 3.113, 3.176
Gerard of Cremona 1.166, 2.455
Gerard, Albert S. 4.182
German 1.119, 1.122, 1.484–485, 2.345, 2.699, 3.725
German, Swiss 2.684
German, syllable structure in 4.388–389
German in Morocco 3.295
German in Texas 2.684
Germania 2.217
Germanic languages 1.475, 1.501, 3.228, 3.232–233
Germanos, Marie A. 4.630
Germany 2.72
Germany, Arabic in 2.73, 2.75
gerund 1.521, 1.523, 2.590, 4.306, 4.622
gerund nominal 4.741
gerundival construction 3.554
gerundive 3.556, 4.306
gerundive nominal 1.477, 1.479
Gesem 1.126
Gesenius, Wilhelm 3.46, 3.141, 3.409, 3.420, 4.756
Gessi 1.519, 1.521
gesture 2.202, 2.505, 2.579, 3.676
Geva-Kleinberger, Aharon 2.461–464
Geyer, Rudolf 1.615, 3.689, 3.718
Gezer 3.409
Gezi, Kal 2.105
Gezira 4.376–377
Ghabban, ᶜAlī ibn ᵓIbrāhīm 3.515, 3.517
Ghadames 3.56
Ghaith, Ghazi M. 2.694, 2.699, 2.703, 3.315
Ghali, M. 4.669
Ghali, Wagdy Rizk 3.23–25
Ghalib 4.596
Ghana 2.250, 3.170
Ghanem, Choucri 3.348
Ghanem, S.M.A. 4.750
Gharb 3.147
Ghassanid Dynasty 2.499
Ghat 3.56
Gḥaṭān 3.326
Gḥaṭān Arabic 3.326
Ghazali, Nazim il- 1.528
Ghazeli, Salem 3.48, 3.610, 4.637, 4.666–667, 4.670–671, 4.675
Ghaznavid Dynasty 1.403, 3.508
Ghedira, Ameur 1.449
Gherdaya 2.715
Ghobrial, Naguib 3.680
Ghoneim, Wafiq 3.471
Ghul, Mahmud Ali 1.614, 1.617
Ghul, Omar al- 3.467
Ghuri, Muhammad 2.325–326
Giannikadou, Anastasia 3.653, 3.655
Giannini, A. 4.669–670
Giasuddin Tughlak 4.452
Gibaliya 1.655
Gibb, Hamilton A.R. 3.732, 4.105
Gibran, Gibran Khalil → Jibran (Jubran), Jibran Khalil
Gibson, Michael [Maik] 2.161, 2.318–325, 3.11–13, 4.563–571
Gide, André 4.76
Gieling, Saskia 2.407
Giggei, Antonio [Giggeius, Antonius] 1.170, 3.30
Gihaf Arabic 4.756
Ğihēna 2.2
Gijduvon 4.613
Gil, David 3.64–66, 3.68
Gil, Moshe 2.259
Giles, Howard 4.320, 4.628
gilit Arabic 1.86 1.93, 1.122, 1.232, 1.610–611, 1.260, 2.414–424, 2.533, 2.571–572, 4.3, 4.325
Gilliéron, Jules 1.590
Gilliot, Claude 3.36, 4.8

Gimira 3.494
Gina Arabic 2.2
Gincburg, David 3.208
Gindibu 4.170
ġinnāwas 1.528
Girgas, Viktor 4.429
Girod, Alain 2.12–16, 2.315–318, 3.200
giṣīdih 3.500
given/new information 1.430
Givenness Hierarchy 3.86–87
Givón, Talmy 1.42, 1.649, 2.192, 2.594, 3.711, 4.40, 4.161, 4.316, 4.502, 4.505–506
Giza Arabic 2.2, 2.5
Glavda 4.716
Glaß, Dagmar 4.76
Gleßgen, Martin-Dietrich 2.455
glide 1.235, 1.640, 1.642, 2.164–169, 2.228, 2.233–235, 2.349, 2.448, 3.241
glide deletion 3.240
gliding 2.688
globalization 1.176, 1.420, 2.655, 2.660, 2.695–696, 2.700, 2.705, 4.147, 4.154
glottal 3.599, 4.715
glottal catch 4.242
glottal co-occlusion 4.3
glottal replacement 2.678
glottal stop 1.81, 1.204, 1.260, 1.295, 1.300, 1.323, 1.376, 1.384, 1.397, 1.399–400, 1.407, 1.490, 1.492, 1.546, 1.608–610, 1.632, 1.653, 2.98, 2.228–232, 2.263, 2.279, 2.343, 2.350, 2.408, 2.416, 2.473, 2.482, 2.503, 2.508, 2.610, 3.125, 3.146, 3.594, 3.599, 4.125
glottal stop, loss of 1.594
glottal stop, pharyngealized 1.276, 2.231
glottal stop, prothetic 2.5
glottalization 1.27, 1.103, 1.299, 2.3, 2.460, 4.1, 4.109, 4.172, 4.422
glottalization, pausal 2.5, 3.363–364
glottis 2.229, 3.599, 4.666–668
glottochronology 4.312–313
Glubb, John Bagot 2.498, 3.199, 3.402
gnä 3.173
gnome 3.730
Goa 2.335
goal 1.344–345, 2.145, 3.230, 4.487, 4.489–490, 4.492–493, 4.535, 4.538
God, attributes of 2.432, 4.219
God, names of 2.600, 4.74
God, qualities of 2.432
God, speech of 2.306, 4.6, 4.73, 4.105
God wish 4.399, 4.419, 4.655–656
Goddard, Cliff 4.166
Goebl, Hans 1.585–586
Goedemans, Rob 3.66
Goffman, Ervin 2.202
Goggot 4.301
Goitein, Shlomo Dov 1.589, 1.598, 2.534, 4.106, 4.108, 4.750
Golan 4.402
Gölbaşı 1.115–116
Golconda 2.328
Gold, Toni 2.675
Goldberg, Harvey E. 2.471
Goldenberg, Gideon 1.573, 1.575, 1.638, 2.435, 2.539, 2.544, 4.339, 4.737
Goldsmith, John 3.301, 3.461, 3.617
Goldstein, Louis 4.335
Goldziher, Ignaz 3.88–89, 3.91, 3.719–720
Golius, Jacob 1.167, 1.170, 1.172, 3.22, 3.30, 4.516
Golkunda Kingdom 4.597
Gontor 2.338
Goodglass, Harry 2.690
Goodman, Morris 3.641, 4.368
Goody, Jack 2.583, 3.135
Goossens, Jan 1.585, 1.590
Gopnik, Myrna 2.689
Goran 1.360
Gorani 2.604
Gorda Arabic 2.622
Gordon, C.H. 1.643
Gordon, Matthew 4.288–289, 4.350, 4.630, 4.634
Gordon, Raymond G. Jr. 3.312
Gordon, Yehuda Leib 1.494
Gorea, Maria 3.470
Gorter, Durk 3.315
ġoṣ 4.157
ġoṣ Hedawa 4.157
ġoṣ lkaf 4.157

Goulfey 1.334
Goulimine 2.240
Goulimine Arabic 3.288
governance 1.67, 2.92–93, 2.290–292, 3.108, 4.221
governance of the object 3.543
government 1.347–349, 2.169–175, 2.181, 2.187, 2.307–308, 3.178, 3.231, 4.429
Government and Binding Theory 1.347, 1.457, 2.170, 3.229–230, 3.706, 3.713, 4.488, 4.657
government relation 2.538
government, cancellation of 4.428–430
government, head 1.350
Government-Agreement analysis 1.46–47
Government-Agreement Constraint 1.47
governor 1.67, 1.349–350, 2.92, 2.169, 2.171, 2.292, 2.294–295, 2.356, 2.538
governor, abstract 2.290
governor, subordinate 1.344
Gozo 3.143, 3.147
grade repetition 3.75
gradience of similarity 3.627
Gradual Differentiation Theory 1.368
Graf, Dafna 3.494
Graf, David F. 3.468
graffiti, Bedouin 4.480
graffiti, memorial 4.481
graffiti, Nabataean 3.467
Gragg, Gene 4.302, 4.305, 4.307
gram 2.191, 2.193–194
grammacization 2.191
grammar 1.530, 2.335
Grammar, Categorical 1.457
Grammar, Categorical Unification-based 1.457
grammar, Classical Arabic 1.402
grammar, cognitive 4.501
grammar, constraint 1.514
grammar, context-free 1.456, 1.514
grammar, contrastive 1.482–489
Grammar, Definite Clause 1457, 1.514
Grammar, Dependency 1.457
grammar, feature augmented 1.456
Grammar, Functional 1.457, 2.143–150, 4.726
Grammar, Generalized Phrase Structure 1.458
Grammar, Generative 1.312, 1.347, 1.422, 1.478, 1.488, 1.507, 2.143, 3.204, 3.338, 3.232–233, 4.521, 4.741
Grammar, Lexical Functional 1.457, 4.535, 4.657
Grammar, Lexical Tree Adjoining 1.457
grammar, natural language 1.347
grammar, origin of 2.177
grammar, pedagogical 2.178, 2.184, 2.188
Grammar, Phrase Structure 1.514
Grammar, Relational 1.457, 4.502, 4.535
grammar, school 1.575
grammar, simplification of 1.176
grammar, speculative 1.486
Grammar, Systemic 1.457
grammar, teaching of 1.534, 2.107, 2.110, 2.188, 2.336
Grammar, Transformational 1.457, 1.562, 2.356, 4.448, 4.520–528, 4.657, 4.741
Grammar, Tree Adjoining 1.514
Grammar, Universal 1.342, 1.484–485, 1.488, 2.96, 2.102, 2.170, 4.143–144, 4.391,
Grammar, Word 1.457
grammar instruction 4.21
grammarians, Baghdadian → Baghdad, school of
grammarians, Basran → Basran school
grammarians, Kufan → Kufan school
grammar-translation method 4.148
grammatical function splitting 1.350
grammatical tradition, Arabic 1.488, 2.175–191, 2.265
grammatical tradition, Coptic 2235, 3.72
grammatical tradition, Greek 2.309, 2.425, 3.107, 3.547–548
grammatical tradition, Hebrew 1.313
grammatical tradition, Indian 3.123–124
grammatical tradition, Persian 2.235, 2.411
grammatical tradition, Syriac 2.232, 2.294, 2.309, 2.402, 2.411, 3.107, 3.547

grammatical tradition, Turkic 2235, 2.405, 3.72
grammaticality 2.117, 2.121, 2.193
grammaticality, cline of 2.195
grammaticality, degree of 2.192
grammaticalization 1.23–24, 1.203, 1.220, 1.328, 1.395, 1.476, 1.540, 1.554, 1.658, 2.144, 2.191–198, 2.215, 3.187, 3.191, 3.236, 4.37, 4.39–40, 4.42, 4.160–164, 4.166, 4.196–198, 4.250, 4.756
grammaticalization cycle 2.192
grammaticalization, layers of 2.195
grammaticization 1.485, 4.161
grammatization 2.191
Grammont, Maurice 3.205
Granada 1.168, 174
Granada Arabic 1.103, 1.107, 3.23, 3.585, 4.431
Grand Màggal 4.181
Grand'Henry, Jacques 1.53–58, 1.387, 1.609, 2.223, 4.70, 4.430–431
Grande, Bentsion M. 4.429
Grandguillaume, Gilbert 2.697, 3.347, 3.349, 4.575
Grandlaunay, René-Vincent du 3.377–381
grantha letters 4.434
grapheme 2.237–238
Grapow, Hermann 2.68
Gratzl, Emil 3.718
Greek 1.110, 1.126–127, 1.132, 1.170, 1.175, 1.401, 1.404, 1.494, 1.498–499, 1.527, 1.537, 1.593, 2.1, 2.68, 2.201, 2.263, 2.265–266, 2.286, 2.289, 2.469, 2.560, 3.6–7, 3.59, 3.72, 3.141, 3.218, 3.313, 3.422, 3.426, 3.435, 3.465, 3.470, 3.553, 3.641, 4.73, 4.194, 4.371, 4.539, 4.572, 4.578–579
Greek, Ancient 1.539, 1.565, 2.345
Greek, Cypriot 1.537, 1.539, 1.612
Greek, Modern 1.539, 2.201
Greek, object-raising in 4.539
Greek, Sicilian 4.215, 4.218
Greek, verb in 4.545
Greek in Asia Minor 2.684
Greek in Syria 4.402
Greek philosophy 2.176
Greek writings, translation of 1.403
Green, Georgia M. 2.19, 3.676–678
Greenberg, Joseph H. 1.484, 2.250–251, 2.253, 2.319, 3.435, 3.445–446, 3.462, 3.610, 3.625, 4.51, 4.94–95, 4.161, 4.279, 4.375, 4.523, 4.725
Greenfield, Jonas C. 3.472
greetings 2.148, 2.202–207, 3.540, 3.659, 3.661, 3.681
greetings, Biblical 2.204
greetings, Islamic 2.203, 2.206
greetings, Malagasy 3.126
greetings, Wolof 4.183
Grenat, Mohamed Hasan 4.488
Grendel, Marjon 2.682–683
Grice, H. Paul 2.19, 2.70, 3.677
Grice's Cooperative Principle 2.19, 3.677, 3.681
Gricean maxims 1.429
grid 3.616
Grid Theory 3.616
Griffin, David 2.287
Grigore, George 3.237, 4.70
Grimes, Barbara F. 3.506, 4.580
Grimes, Joseph E. 2.208, 4.317
Grimm, Jacob 1.119
Grimme, Hubert 3.469
Grimshaw, Allen D. 2.202
Grimshaw, Jane 3.388, 3.393, 4.489, 4.491, 4.648
Grinder, John 4.527
griot 3.173, 4.180
gris-gris 3.139–140
Grohmann, Adolf 1.149, 3.469–470, 3.513–516, 3.519
Groome, Francis H. 2.220
Grosjean, François 1.369
Grotzfeld, Heinz 2.278, 3.190, 3.737–738, 4.196, 4.654, 4.664
ground/figure 1.423
grounded phenomenon 2.365
grounding 1.422, 1.473, 1.648–649, 2.207–210, 4.728
grounding values 2.207
group identity 2.222
group solidarity 4.251
group stereotypes 4.324
groupness 2.58
Gruber, Jeffrey 4.487, 4.489, 4.492

Gruendler, Beatrice 1.148–155, 2.263, 3.517–518
Grunebaum, Gustave E. von 2.304–305
Grünert, Max 4.558
Gruntfest, Yaakov 4.448
gṭāʿ 3.173
Guanche 1.35, 2.67
Guayaquil 3.2
Guban 2.572
ġubār 1.151–152, 3.309, 4.134
gudayən 3.170
Gudykunst, William B. 4.322
Guellala 2.709
Guera 1.360, 362–364, 366
Guerssel, Mohand [Mohamed] 1.121–122, 2.37, 2.168, 3.302
Gueunier, Noel J. 3.126
Guignard, Michel 3.173
Guillaume, Jean-Patrick 1.641, 2.91, 2.165, 2.168–182, 2.233, 2.291, 2.299, 2.309–310, 2.425, 2.448, 2.536, 2.538, 2.544, 2.562, 3.454, 4.119, 4.447
Guimier, Claude 4.160
Guinea 1.255, 4.179
Guinea Bissau 4.179
Guion, Susan G. 3.524
Guirao, Jose M. 1.513
Guiraud, Pierre 3.47
Gujarati 1.662, 1.665
Gujeratis 2.333
Guled, Said Utman 4.277
Gülensoy, Tuncer 4.592
Gulf Arabic 1.270, 1.272, 1.595, 1.606, 1.611, 1.624, 2.36, 2.38, 2.96, 2.161, 2.210–213, 2.374–375, 2.388, 2.390, 2.555, 2.587, 3.9, 3.12, 3.19, 3.63, 3.68, 3.264, 3.451–452, 3.478, 3.545, 3.585, 3.587, 3.592, 3.609, 4.124, 4.317, 4.646, 4.656, 4.695
Gulf Arabic Pidgin → Pidgin Gulf Arabic
Gulf Co-operation Council 2.215
Gulf of Aqaba 3.401
Gulf States 2.111, 2.210–216
Gulf States, youth speech in 4.766–767
Gully, Adrian J. 1.471–472, 2.17, 2.224–225, 2.315, 2.317, 2.355, 2.388, 2.390, 2.424, 2.452–453, 2.477–478, 2.539, 3.17, 3.101, 3.107, 3.323, 4.448
Gulumba 4.715
gulzār 4.134
Gumʿa, Muḥammad Luṭfī 4.765
Gumbad-i Qabus 2.408
Gumperz, John J. 1.415, 1.417, 1.442, 1.484–485, 3.664, 4.58
Gundel, Jeannette K. 3.86
ġunna 3.335, 4.427–428
Gura 4.301
Gurage 2.53–54, 2.66–67, 4.312
Gurage, Central 4.301
Gurage, East 4.301, 4.304–305, 4.310, 4.312
Gurage, West 4.301
Gurage-Selte 2.51
Gurani 3.351
Gurna Arabic 2.5
Gurnat Marʿi 1.299
ġūṣ 2.468, 2.470, 4.157
Ġūṣ Arabic 2.5
ġuṣ trisa 4.157
ġusl 1.381
guslar 3.646
Gussenhoven, Carlos 2.395
Ġūṭa 1.546
Gutas, Dimitri 2.198–202
ġuṯm 2.259
Gutt, Ernst-August 4.305–307
Guttman, Louis 2.319
Guttman Scale 2.319
guttural 2.38, 2.229–230, 2.365, 3.205, 3.328–330
Guyard, Stanislas 3.208
ġuyūb 2.303
Gwadur 2.211
Gyeto 2.67, 4.301
Gypsies 2.469, 2.576
Gypsy Arabic 2.216–222, 3.527
Gypsy language 4.255

H

hāʾ as-sakt 3.565
hāʾ marbūṭa 3.92
hāʾ mardūfa 4.560
hāʾ muḥaqqaqa 4.560
hāʾ muḥdawdiba 4.560
hāʾ musalsala 4.562
Haak, Martine 1.216–221
ḥabasa 3.603
Habash, Nizar 1.460
ḥabbaḏā 2.80
Habermas, Jürgen 4.73
Ḥabīb, Mūnā 2.698
habitual 1.38, 1.95, 1.195, 1.198–199, 1.220, 1.229, 1.253, 1.305, 1.330, 1.508, 2.514, 2.516, 2.524, 2.595, 3.247, 3.266, 3.536, 3.573
habitual, past 2.516
habituality 1.331, 4.457
Habré, President Hissène 1.365
Habtoor, Mohamed 4.750
Hachimi, Atiqa 2.155–164, 4.629
ḥaḍar 2.212, 2.571
ḥaḍarī 1.605, 1.617, 2.212
ḥadaṯ 2.180, 2.549, 3.455, 4.659, 4.685
ḥadaṯān 4.659
ḥadd 1.192, 2.544
Haddad, Ayed 1.668, 1.670
Haddad, Bassam 1.458
Haddad, E. Nasrallah 2.481
Ḥaddād, Fuʾād al- 1.601
Haddad, Ghassan F. 2.36, 4.64, 4.348
Haddad, Qasim 4.213
Haddad, Thuraya 1.668, 1.670
Haḏḏāl, Xalaf al- 4.129
ḥaddaṯa 2.435, 4.737
ḥaḏf 2.16–17, 2.165, 2.223, 2.290, 2.300–302, 2.544, 3.91, 3.649–650, 4.27, 4.121, 4.448
ḥāḍir 2.91, 3.97, 3.548
Ḥadīṯ 1.2, 1.380, 1.427, 1.472, 1.512, 1.558, 2.46, 2.109, 2.183, 2.203, 2.327–328, 2.335, 2.337, 2.436–437, 2.552, 3.122, 3.136, 3.629
Ḥadīṯi, Xadīja al- 2.426
Hadj-Salah, Abdurrahman 1.457, 3.73
Ḥaḍramawt 1.260, 1.448, 1.451, 1.545, 1.662, 2.56, 2.273, 2.337, 2.340, 3.694, 4.295, 4.687–699, 4.750
Ḥaḍramawt Arabic 1.260, 2.270, 2.274, 3.449, 4.687–699, 4.750–751, 4.753, 4.754–757
Ḥaḍrami Arabic → Ḥaḍramawt Arabic
Ḥaḍramī, ʿAbdallāh al- 4.209
Ḥaḍramīs 2.336–337
Hadramitic 2.212, 4.295, 4.301
Haegeman, Liliane 3.655, 4.392
Haeri, Niloofar 1.527–536, 1.631, 1.633, 1.636, 2.8, 2.11, 2.642, 2.653, 2.655–657, 2.694, 2.700–701, 2.705, 3.607, 4.631, 4.634
ḥafala 4.34–35
Hafez Osman 3.341, 3.343
Hafez, Sabry 1.528
Ḥāfiẓ al-Qurʾān 3.341
Ḥāfiẓ, aṭ-Ṭāhir 1.438
Ḥāfiẓ, Ḥusayn 1.382
Hagar 1.129
Hagège, Claude 3.61
Hagen, Admiral van der 3.128
Hagen, Gottfried 3.501–505
Hagenbucher, F. 1.333

Haggan, Madeline 1.339, 1.341–342
hagiography, martyr 4.78
Hago, Abdel Moneim Nasir 2.518
Haichour, El Houcine 1.649
Haifa 2.439, 2.464, 2.505–506
Haifa Arabic 1.608
Hāʾil 4.482
Haim, Soleiman 2.410–411
Haim, Sylvia 3.345–346
Haiman, John 4.161
Haiting 1.381
Ḥāja bint Miqdām 4.275
Ḥajarī, Nāṣir ad-Dīn ʾAḥmad ibn Qāsim al- 1.169, 3.220, 3.222
Haji Abdulmalik ibn Karim ibn Amrullah 2.337
Hajič, Jan 1.458, 1.516
ḥajj 1.448
Ḥājj, Kamāl Yūsuf al- 3.349
Ḥājj, ʾUnsi al- 4.213
Hajj Malik Sy, al- → Malik Sy, al-Hajj
Hajj Umar Tal, al- 4.179
Hajjaj, Ali 2.32
Ḥajjāj ibn Yūsuf, al- 1.152, 3.120, 3.517, 3.641, 4.204
Ḥājji Xalīfa 3.501–503
Haj-Salah, Abdurrahman → Hadj-Salah, Abdurrahman
Ḥakam II, Caliph al- 2.46
Håkansson, Gisela 2.97, 2.117, 2.679
Ḥaketiyya 4.156
Ḥakīm, Tawfīq al- 1.409, 1.600–601, 3.556
Hakmi 1.655
Hakuta, Kenji 2.681
ḥāl 1.19, 1.50, 1.68, 1.563, 2.91, 2.115, 2.132, 2.143, 2.185, 2.224–228, 2.356, 2.389, 3.97, 3.101, 3.103, 3.105, 3.108, 3.114, 3.160, 3.200, 3.261, 3.286, 3.542, 3.548, 3.685–687, 3.703, 4.221, 4.360, 4.412, 4.414, 4.436, 4.539, 4.737, 4.746–747
ḥāl al-fiʿl 3.96
ḥāl sentence 1.308, 1.332, 1.396, 1.465–466, 3.186, 3.537, 3.543
ḥāl, ṣāḥib al- 2.132, 2.224
ḥāla 2.226
Ḥalab → Aleppo
Ḥalab is-Sūdān 2.217
Ḥalabī, Šihāb ad-Dīn Maḥmūd ibn Sulaymān al- 2.497, 4.167
Halasi-Kun, Tibor 4.592
Ḥalawānī, Muḥammad Xayr → Ḥilwānī, Muḥammad Xayr
Ḥalāyib Arabic 2.2
Halenqa 1.656
Halhed 2.673
Halhed, Nathaniel B. 1.287
Halila, Hafedh 4.318, 4.398, 4.648
Ḥallāj, al- 4.211
Hallāniyyāt 3.478
Halle, Morris 2.38, 2.396, 3.204, 3.302, 3.525, 3.615–616, 3.618–622, 4.342, 4.344–346, 4.348–349
Hallel, Michael 2.464
Hallett, Jennifer S. 2.652
Halliday, Michael A.K. 1.430, 1.457, 1.472, 1.474–475, 2.113, 3.664, 3.713, 3.715, 4.57, 4.494, 4.498, 4.502, 4.726
Hallman, Peter 2.365–372, 4.14–20
Halper, B. 3.472
ḥalq 2.228–229
Ham, Anthony 3.56
Hama Arabic 1.608, 4.408
Ḥamad, Muntaṣir al- 3.467
Ḥamādah Arabic 4.240
Hamadan 3.693
Hamaḏānī, ʿAbd ar-Raḥmān ibn ʿĪsā al- 3.33, 3.693
Hamaḏānī, Badīʿ az-Zamān al- 1.403, 1.512, 2.469
Hamar Arabic 4.376
Ḥamāt 1.361
Ḥamd Allāh al-ʾAmāsī 4.561
Hamdān 4.295
Hamdān Arabic 3.94
Hamdānī, al-Ḥasan al- 2.257–260, 4.295, 4.303, 4.312
Hamden, Jihad 4.489
Hamdi, Rym 3.725, 4.675
Hämeen-Anttila, Jaakko 2.159, 2.280–281, 4.32–37
Hamer 1.35
Ḥāmid ibn al-Faqīh Ṣiddīq al-Hararī al-Ḥimyarī 2.52
Hamid, Abdel Halim M. 2.38, 2.560
Hamīde 1.362

Ḥamīdīye, il 4.402
Ḥāmidun (ibn), Muxtār 3.172
Hamilton, Alastair 1.166–172
Hamitic languages 1.35, 2.250
Hamito-Semitic languages 1.35, 1.315, 2.250
Hamito-Semitic languages → Afro-Asiatic languages
Hamka 2.337
Ḥammādīye 1.362
Hammarberg, Björn 2.677
Hammond, Michael 3.304, 3.618, 4.339
Hammoumi, Abdeljebbar 3.726
Ḥammūd, Ṣalāḥ ad-Dīn 2.694, 2.698, 2.703
hams 3.123–124
hamz 2.263–264
hamza 1.148, 1.152–153, 1.204, 1.273, 1.490, 1.617, 2.228–232, 2.279, 2.408–409, 2.427, 3.77, 3.125, 3.200, 3.220, 3.594, 3.599, 3.691–697, 4.6, 4.121, 4.125, 4.518–519
hamza bayna bayna 2.229
hamza bearer 4.700
Ḥamza ibn Ḥabīb 4.5
hamza musahhala 4.427
hamza, initial 1.616
hamza, intervocalic 1.640
hamza, loss of 1.564, 1.614, 1.616
hamza, prothetic 4.640
hamza, talyīn al- 3.91
Hamzaoui, Rached [Hamzāwī, Muḥammad Rašād] 2.636–637, 4.460–461
*hamzat al-qaṭ*ʿ 2.230, 3.111, 4.432
hamzat al-waṣl 1.400, 2.230, 2.327, 3.111, 3.650, 3.729, 4.432, 4.640
Hamzé, Hassan Salam Bazzi 2.82–84, 2.90–96
Han Muratpaşa 1.114
Hanādwah Arabic 4.240
Hanafis 1.380, 2.337, 2.609
Hanaway, William L. 3.337–338
Hanbalis 1.380
Ḥanīfa Arabic 3.94
Hankamer, Jorge 3.705
Hanks, Hilary 2.703
Hanna. Nelly 1.602
Hanna, Nabil Sobhi 2.217–218, 2.220
Ḥannāš, Muḥammad al- 2.705
Hannum, Hurst 2.439
Hanssen, Federico [Friedrich] 1.97
Hansson, Kristina 2.679
Hantýrka 2.217
haplology 2.299, 3.592, 4.558
Haq, Anwar ul 2.170
Haq, Muhammad Enamul 1.287
ḥaqīqa 2.579–581, 3.116–117, 3.119–122
ḥaqīqat al-kalām 4.333
ḥaqīqī 3.121
Ḥaql Kitāb 2.258
ḥaqq aṣ-ṣadāra 4.740
ḥaraka 2.183, 2.232–238, 2.309, 2.402, 2.404, 2.545–546, 3.208, 4.428
Harar 2.51–52, 4.276
Harari 1.133, 2.53–55, 2.66–67, 4.301, 4.303, 4.305, 4.312
Haravī, Muḥammad Šafīʿ 3.338
Harawī, Mīr Muḥammad Zāhid al- 2.327
Ḥarb 3.326, 4.125
Ḥarb Arabic 4.126–127
Harbert, Wayne 1.46, 1.48, 2.174, 3.231, 3.382, 4.193–194, 4.365, 4.396–397, 4.648–650, 4.657
Ḥarbī, ʾAbū ʾIsḥāq ʾIbrāhīm ibn ʾIsḥāq al- 3.34
Ḥarbīye 1.113
Ḥardallū 3.348
Harding, Gerald Lankester 3.227, 4.482
Hardt, Daniel 2.18
ḥarf 1.309, 1.471, 1.557–558, 2.90, 2.184, 2.232, 2.234–239, 2.402, 2.404, 2.424, 2.429, 2.435, 2.545–546, 3.14, 3.92–93, 3.208–209, 3.241, 3.546, 3.549, 3.602, 4.4, 4.6, 4.431
ḥarf al-badal 2.546
ḥarf al-ibtidāʾ 2.290
ḥarf al-ʾiḍāfa 1.72, 2.294, 3.103, 3.699
ḥarf al-ʿilla 2.309–310
ḥarf al-istiṯnāʾ 1.72, 2.452
ḥarf al-jarr 1.72, 1.310, 2.237, 2.295, 2.425–426, 3.699
ḥarf al-jazāʾ 2.477, 2.479
ḥarf al-līn wa-l-madd 2.233, 2.238, 2.309
ḥarf al-muḍāraʿa 1.574
ḥarf al-xafḍ 1.72, 3.699

ḥarf ʾaṣl 4.422
ḥarf ʾaṣlī 2.238–239, 2.447
ḥarf aš-šarṭ 2.479
ḥarf aṣ-ṣifa 4.220, 4.236
ḥarf aš-širāk 2.237
ḥarf at-tawkid 2.115
ḥarf farʿ 4.422
ḥarf ḥalqī 2.327
ḥarf istifhām 2.237
ḥarf istiʿlāʾ 4.423
ḥarf jarr 1.310, 2.237
ḥarf maʿnā 4.443
ḥarf munfatiḥ 2.459
ḥarf muqaṭṭaʿ 1.2
ḥarf mušabbah 1.69
ḥarf mustaʿlin 2.312, 4.2
ḥarf mutaḥarrik 2.238
ḥarf muʿtall 2.233
ḥarf muṭbaq 2.459
ḥarf nāṭiq 2.235
ḥarf nidāʾ 2.237
ḥarf qamarī 1.204
ḥarf qasam 2.237
ḥarf sākin 2.238
ḥarf šamsī 1.204, 4.518
ḥarf ṣawtī 2.235
ḥarf zāʾid 2.238–239, 2.546–547
Ḥarīrī, ʾAbū Muḥammad al-Qāsim ibn ʿAlī al- 2.85–86, 2.88, 2.189, 2.446, 2.469, 2.629–633, 3.508, 4.105, 4.157
Ḥāriṯ Arabic 3.94
Ḥāriṯ ibn Jabala, al- 3.470
harmony, iambic 4.389–390
Harms, Robert 3.617
Harning, Kerstin Eksell → Eksell, Kerstin
Haron, Muhammad 4.291
Haroon, Mohammed 2.328
Harrak, Fatima 4.759
Ḥarrān, inscription of 1.149
Harrell, Richard Slade 2.96, 2.645, 2.678, 2.686, 3.274, 3.279, 3.284, 3.287, 3.293, 3.429, 3.611, 3.613, 3.655–656, 4.62, 4.317, 4.344, 4.351, 4.646, 4.650–651, 4.654–655, 4.670, 4.682
Harris, Alice C. 2.191, 4.37–39, 4.41
Harris, Rachael 3.678
Harris, Zellig S. 2.542, 4.646, 4.670
Harrison, Christopher 3.137
Ḥarsūsi 2.66–67, 2.211, 3.350, 3.478, 4.297–302, 4.304, 4.309–310, 4.312
Hart, David M. 2.62
Hart, Johan ’t 2.397
Hartmann, Martin 3.214
Hartmann, Regina 1.457
Hartmann, R.R.K. 2.378
Hārūn ar-Rašīd, Caliph 3.690
Hārūn, ʿAbd as-Salām 2.80
Hary, Benjamin 1.447, 1.630, 1.667–668, 2.275–279, 2.495, 2.529, 2.534, 3.667
Ḥasa, al- 4.124, 4.126–127, 4.482
Ḥasa Arabic 4.127–128
Ḥaṣaitic 2.198, 2.263, 3.411, 4.477–478
Ḥasake Arabic 4.407
ḥasan 1.192, 2.184, 2.542, 3.432
Ḥasan ibn ʿArabšāh 2.46
Hasan II, King 1.534, 2.63, 2.714, 3.350
Ḥasan, ʿAbbās 1.40, 2.296, 2.424–425, 3.17, 3.392, 4.61, 4.411, 4.432, 4.538–540
Hasan, Ruqaiya 1.430, 1.472, 1.474, 3.713, 3.715
Ḥasan, Yūsuf Faḍl 3.435
Ḥasanayn, ʾAḥmad Ṭāhir → Hassanein, Ahmed Taher
Ḥasanī, Mīr ʿImād al- 3.337–338
Ḥasawūna 1.361–362
Ḥasb Allāh, ʿAlī 2.578
Hashabeiky, Forogh 2.408
Hashim 4.604
Hāšimī, ʾAḥmad al- 4.30
ḥāšiya 1.2–4
Hasjim, Kiyahi Haji Abdul Wahid 2.338
Hasköy 1.92
Haspelmath, Martin 2.197, 3.64–66, 3.68
ḥaṣr 4.27, 4.496
ḫāṣṣa 1.610, 2.306
Ḥassan, ʿAbbās → Ḥasan, ʿAbbās
Ḥassān, ʾAḥmad 2.337
Hassan, Jidda 2.471, 4.714
Hassan, Kamel M. 3.201
Hassan, Motie Ibrahim 4.727
Hassan, S. al- 3.654
Ḥassān, Tammām 1.432, 1.436, 2.80, 4.22, 4.283

Hassanayn, Khalid S.A. 2.205–206, 3.681
Hassanein, Ahmad Taher 1.472, 3.37–45, 3.71–74, 4.763–767
Ḥassān-i Hind 2.328
Ḥassāniyya 1.24, 1.265, 1.357, 1.570, 1.572, 1.625, 2.240–250, 2.388, 2.622, 3.53, 3.60–61, 3.135, 3.169, 3.171, 3.173–175, 3.288, 3.290, 3.314, 3.352, 3.450–451, 3.585, 3.587, 3.664, 4.70, 4.313, 4.376, 4.550, 4.662
Hassanpour, Amir 2.659, 3.351, 4.579
Hasselbach, Rebecca 3.408–422
Hasting 3.718
ḥasuna 2.426
ḥašw 2.307, 3.210, 4.235–236
Hatari 4.402
Hatay 1.87, 1.112, 3.315, 4.579–581
Hatay Arabic 1.584
Hatch, Evelyn 2.118
Hatim, Basil 2.209, 2.331–332, 4.473
Ḥātimī, Muḥammad ibn al-Ḥasan al- 3.117
Hatra 1.126–127, 129
Hatran 3.409
ḥattā 1.71, 1.465, 1.467–469, 1.557, 1.612, 2.193, 4.39
Haugen, Einar 2.31
Hausa 1.35, 1.38–39, 1.133, 1.135–136, 1.333, 1.404, 1.487, 1.545, 2.137, 2.139, 2.250–256, 2.553–554, 2.668, 2.670, 3.59, 3.350, 3.373, 3.375, 4.279–281, 4.384, 4.710, 4.715–716
Hausa, Sudanese 2.255
Hausa, West African 2.255
Hausa Language Board 2.254
Hausaland 2.251, 3.375
Häusler, Martina 1.599
Have, Paul ten 2.384
hawāʾ 2.228
Hawar 3.351
Hawar Islands 2.215
ḥawāšī 3.309, 4.134
Hawāwīnī, Najīb 4.98
Ḥawāzin 4.304, 4.430
Ḥawāzin Arabic 2.555, 3.94
Hawi, Khalil 4.213
Hawkins, John A. 1.484
Hawkins, Roger 4.144–145
Ḥawrān → Horan
hayʾa 4.436
Hayajneh, Hani 4.478
Hayani, Ibrahim 3.395
Haydar al-Mašhadī 4.561
Haydar, Y. 2.395–397, 3.727
Ḥaydara, ʾAbū l-Ḥasan ʿAlī ibn Sulaymān al- 3.124–125
Hayes, Bruce 2.365, 3.613–616, 3.620–621, 3.627–628, 4.344–345, 4.347–348, 4.350, 4.389
Hayes, Donald 1.341
Ḥāyil Arabic 4.126
Haykal, Muḥammad Ḥusayn 3.349, 3.555
ḥayṯu 1.309–310, 1.387, 1.467, 1.566, 3.200
Haywood, John Alfred 2.86, 2.230, 3.21–22, 3.36, 3.558–559, 4.148, 4.358
ḥayyiz 2.228
Hazai, György 4.584
hazaj 3.210–211, 3.214
Hazaras 1.29
Hazout, Ilan 3.388, 3.393, 3.432, 4.400, 4.491–492
head 1.457, 1.481, 2.146, 2.397, 3.229–230, 4.392, 4.488, 4.646
head, tense 4.392
head, verbal 4.392
Head-Driven Phrase Structure Grammar 4.657
headedness 3.619
head-government 2.171–172, 2.174
headlines 3.203, 3.540, 4.726
headlines, French 1.651
headlines, newspaper 4.499, 4.654
head-movement 4.657
Healey, John F. 3.471, 4.483
hearer-new status 3.87
hearing, loss of 2.675, 2.688
hearing impairment 2.677–678, 2.688
Heath, Jeffrey 1.27, 1.296, 1.406–407, 1.418, 1.445, 1.571, 2.122–123, 2.231, 2.240, 2.249, 2.374, 2.468, 2.533–534, 3.274, 3.286–288, 3.290–293, 3.305, 3.664, 4.52, 4.70, 4.288–289, 4.339, 4.637

heaviness 2.178, 2.350, 2.425, 3.45, 3.108, 4.432, 442–443
ḥəḇāṣā 2.232
Hebbo, Ahmed Irhayem 1.181, 2.57, 2.199
Hebrew 1.39, 1.100, 1.110, 1.126–127, 1.132, 1.166–168, 1.170–171, 1.233, 1.237, 1.240, 1.350–351, 1.376–377, 1.404, 1.445–446, 1.467–468, 1.479, 1.490, 1.493–494, 1.508, 1.541, 1.564, 1.639, 2.23–24, 2.64, 2.66–68, 2.181, 2.223, 2.257, 2.262, 2.354, 2.373, 2.438–440, 2.464–467, 2.469–471, 2.482, 2.504, 2.526–527, 2.529, 2.531, 2.533, 2.659–660, 2.684, 2.689, 2.691, 2.699, 3.46, 3.48, 3.64, 3.191–192, 3.218, 3.227, 3.233, 3.312–313, 3.315, 3.360, 3.398, 3.408, 3.410–411, 3.413–417, 3.419–420, 3.429, 3.431–432, 3.553, 3.592, 3.641, 3.705, 3.713, 3.731, 4.1, 4.35, 4.53, 4.72–73, 4.96, 4.104, 4.147, 4.157, 4.172–173, 4.175–177, 4.271, 4.302–303, 4.313, 4.395, 4.456, 4.479, 4.526, 4.756, 4.759
Hebrew, acquisition of 1.376
Hebrew, attitude toward 2.660–661
Hebrew, Biblical 1.196, 1.203, 1.313, 1.494, 1.540, 1.638, 2.23, 2.66, 3.408, 3.412, 3.414, 3.418, 3.629, 3.738, 4.481
Hebrew, clipping 4.558
Hebrew, consecutive in 4.557
Hebrew, construct state in 4.746
Hebrew, Massoretic 2.491
Hebrew, Mishnaic 1.85, 1.493, 2.66, 3.408, 3.414
Hebrew, Modern 1.476, 1.493, 2.461–462, 2.464, 3.205–206, 3.232, 3.493–494, 4.463, 4.558
Hebrew, Modern → Ivrit
Hebrew, plural in 1.237
Hebrew, Rabbinic 2.528
Hebrew, relativization in 4.527
Hebrew, speech errors in 4.271
Hebrew, suffix plural in 1.237
Hebrew, teaching of 3.398
Hebrew, Tiberian 3.621–622
Hebrew grammatical tradition 1.313
Hebrew jargon 2.470
Hebrew verbal Forms 2.531
Hebron 2.30, 2.33, 2.46, 2.506
Hebron Arabic 1.608, 2.30, 3.527
Hebron hills 3.360
Hedberg, Nancy 3.86
Heddawa 4.157
Heggie, Lorie 4.395
Hegra 3.466, 3.471, 4.478, 4.483
Hegra → Ḥijr, al-
Heijer, Johannes den 2.201
Heim, Irene 3.685, 4.15, 4.43, 4.648
Heine, Bernd 1.521–522, 2.191, 2.193, 2.196, 2.589, 3.642, 3.674, 4.38–39, 4.160–163
Heinen-Nasr, Rabha 4.152, 4.533–534
Heinrichs, Wolfhart P. 2.200, 2.442, 2.496–497, 2.578–579, 3.117, 3.122
Heliel, Mohamed Hilmi 1.437–438
Heligoland Treaty 1.660
Heliogabalus 4.402
hellēnismós 2.401
Hellenistic culture 1.593, 3.52
Hellenistic education 2.425
Hellenization 2.199
Heller, Monica 1.415
Hēmāt 1.363
hemistich 4.87, 89, 4.9192
Hemmādiye 1.362
Hendriks, Bernadet 4.222–235
Henein, Nessim Henry 1.503
Hengeveld, Kees 2.143
Henkin, Ron 3.360–369, 3.400, 3.405–406, 3.536, 4.663
Heny, Frank 4.522
Henzl, Vera 2.117
Herat 2.601, 4.133
Herbolich, James B. 2.658
Herburger, Elena 3.655
Herdan, Gustav 4.94
heritage, Arabo-Islamic 2.637–640, 2.655
Heritage, John 2.382
heritage speakers 3.399
Hermann of Carinthia 1.166
hermeneutics, legal 3.551
Herodotus 1.126, 1.132, 3.466, 4.480
Herrero Muñoz-Cobo, Bárbara 3.274
Hersi, Ali Abdirahman 4.276
Hertevin 1.540
Hervey, Sandor G.J. 4.494

Herzallah, Rukayyah 2.38, 4.637, 4.669
Herzog, Martin 2.319, 2.321
Herzog's Principle 3.603
Hess, Jean-Jacques 3.718
heterogeneity, structured 2.318
heteronym 1.592
Hetzron, Robert 2.262, 3.417, 3.420, 4.18, 4.170, 4.302, 4.305–307, 4.457
Heusinger, Klaus von 4.316
Hewramî 2.604
Heycock, Caroline 3.386–387, 3.684–685, 3.687–688, 4.316, 4.395, 4.400
Heyworth-Dunne, James 2.109
hiatus 1.235
hiatus avoidance 4.86
Hidāyat Allāh 3.338
Hidden Markov Models 1.213–214
hierarchy, animacy 2.594, 3.158–159
hierarchy, similarity 3.627
hierarchy, sonority 3.620, 3.622
hierarchy, subject/object assignment 2.146
hierarchy, topicality 4.734
hierarchy of categories 2.180
hierarchy of rhymes 3.620
Ḥigāzī, Maḥmūd Fahmī 4.460
Higginbotham, James 2.368, 4.15, 4.81
High Dam 2.1
High variety 1.419, 1.439, 1.633–634, 1.667, 2.10–11, 2.265, 2.654, 2.695, 2.708, 3.70
Highest Subject Constraint 1.312
ḥijāʾ 2.237, 3.111, 3.173
Ḥijāb, Sayyid 1.601
Ḥijāz 1.133, 1.260, 1.618, 2.499, 3.14, 3.360, 3.401–402, 3.645, 3.691–694, 4.92, 4.123, 4.240
ḥijāzī 1.70, 1.151, 1.153, 1.260, 3.517
Ḥijāzī Arabic 1.259, 1.295, 1.399–400, 1.490, 1.606, 1.641, 2.88, 2.199, 2.263–264, 2.279, 2.312, 2.355, 2.387, 2.499, 2.680, 2.689, 2.691, 3.14, 3.90, 3.93, 3.179, 3.404, 3.497, 3.645, 3.690–691, 3.693–698, 4.9–10, 4.62, 4.70, 4.124–127, 4.351, 4.389, 4.431
ḥijāzī script 1151, 3.517–518
Ḥijāzī, ʿAbd al-Muʿṭī 4.92
Ḥijīz 4.123
Ḥijr, al- 3.471
Ḥijr, al- → Hegra
ḥikāya 2.537, 4.23, 4.283–284
Hikaya Iskandar 2.334
Hikaya Muhammad Hanafiyya 2.334
hikayat 1.6, 8–9
Hikayat Prang Sabi 1.6
ḥikma 2.178, 2.188, 3.731
ḥikmat al-ʿArab 1.173
Hilāl, Bani 1.260
hilālī 1.55
Hilali, Shaikh Ghulam Maqsud 1.287
Hilāli Arabic 1.53, 1.596, 1.609, 3.288
Hilé Alifa 1.334
Hillelson, S. 2.560
Ḥillī, Ṣafī d-Dīn al- 2.313, 2.469, 3.732, 3.654, 4.255, 4.257
ḥilm 4.205–206
Ḥilmī, Mūsā 2.450
Ḥilwānī, Muḥammad Xayr al- 3.226, 3.393
ḥilya 2.14
Ḥimāya, Yassir 4.764
Ḥimṣ → Homs
Ḥimyar 2.56, 2.256, 2.258, 3.693
Ḥimyar Arabic 3.94
Ḥimyarī, Našwān al- 3.32, 3.141–142
Ḥimyaritic 1.127, 1.491, 2.256–261, 2.631, 3.15, 3.92, 3.693–694, 4.123, 4.295–296, 4.301–302, 4.305–306, 4.312–313, 4.478, 4.480, 4.753, 4.756
ḥīna 1.467
Ḥinawr 2.326
Hind 4.104–105
Hindāwī, Ḥasan 4.119
Hindi 1.241, 1.662, 2.210, 2.215, 2.219, 2.673
Hindi, Abu Zafar al- 3.504, 3.508
Hindi, Trinidad 2.684
Hinds, Martin 3.27, 3.518, 4.284, 4.419
Hindu 1.286–288, 3.313
Hinduism 2.340, 2.474, 3.129
Hindustani 2.668, 4.595–603
Hinnebusch, Thomas 1.660, 1.663, 4.381–382
Hinskens, Frans 4.370
Hinton, Leanne 4.282

Ḥīra, al- 1.150, 1.400, 3.210
Hirsh-Pasek, Katryn 1.339–340
Hirst, Daniel 3.724
ḥisāb 1.338
Hišām I, Caliph ʾAbū l-Walīd 1.97, 2.44
Hiskett, Mervyn 3.135, 3.376, 4.279
Ḥisma 3.401–402
Hismaic 3.468, 4.479–482
Hispano-Arabic → Andalusian Arabic
hispano-godos 1.100
Hispano-Romans 1.98
historical narrative 3.248
Hīt Arabic 2.415, 2.533–534
Hītī, ʿAbdallāh ibn ʿAlī al- 4.132
Hitti, Philip K. 1.536–537
Hittite 2.67
Hoberman, Robert D. 3.305, 3.564–570, 4.115–118, 4.339
Ḥobyōt 2.66, 3.63, 4.297–298, 4.301
Hock, Hans Heinrich 3.205–206, 3.606, 4.368–369, 4.372–373
Hockett, Charles F. 2.546
Hodeida Arabic → Ḥudaydah Arabic
Hoedemaekers, Iris 4.290–295
Hoenigswald, Henri M. 3.189
Hoffiz, Benjamin 4.62
Hoffman, Paul R. 2.688
Hofheinz, Albrecht 4.75
Hogariya → Ḥujariyya
Holes, Clive D. 1.241–255, 1.271, 1.406–407, 1.419, 1.440, 1.444–447, 1.532, 1.595, 1.611, 1.651, 2.96, 2.160, 2.194, 2.210–216, 2.261, 2.267, 2.320, 2.323, 2.388, 2.390, 2.608–620, 2.642, 3.9–12, 3.14, 3.70, 3.200–203, 3.239, 3.424, 3.432, 3.446, 3.567, 3.585, 3.609, 3.664–665, 3.668–669, 3.689, 3.713, 3.715, 3.728–729, 4.62, 4.162, 4.168, 4.192–194, 4.286, 4.317, 4.322–326, 4.360, 4.468–469, 4.498, 4.558, 4.631, 4.728, 4.734– 735, 4.747
Hollien 4.668
hollow words 2.448
Holm, John 3.639, 4.195, 4.198
Holmberg, Anders 4.540
holy water 4.181
home language 1.370
Home Language Instruction 2.73, 2.75
Homeidi, Moheiddin Ali 1.348–349, 3.455–461, 4.524
Homer 3.499
Homeric poems 3.645–646
homograph 1.150, 1.152
homonymic clash 1.590, 2.4
homonymy 1.626, 2.496, 2.624, 3.34, 3.45, 3.49, 3.320–323, 3.538–539, 4.164, 4.168
homorganic consonants 3.462–464, 4.95
homosexuals, jargon of 2.470
Homs 1.178, 4.402
Homs Arabic 1.608, 4.678, 4.680
honor 2.363–364, 2.587, 4.417, 4.419–420
honorific 1.286, 2.587, 3.720, 4.468, 4.470
Hoogland, Jan 1.438, 1.511–518, 3.21–30, 3.274
Hooker, M.B. 2.341
Hookham Frère, John 3.143
Hooper, Joan Bybee 3.491
Hopkins, Simon 1.614, 1.617, 1.640, 2.23, 2.194, 2.226, 2.313, 2.534, 3.472, 3.568, 3.641
Hoppenbrouwers, Cor 1.586
Hopper, Paul J. 2.191–196, 2.208–209, 4.37–39, 4.41–42, 4.161, 4.499, 4.728
Hōr al-Ḥammār 2.572
Hōr al-Ḥuwayzah 2.571–572
Horan 1.126, 2.506, 3.466, 3.716, 4.402, 4.407
Horan Arabic 1.546, 1.583, 1.605, 1.607–608, 2.501, 2.506, 2.622, 4.407, 4.635, 4.678
Horesh, Uri 4.454–458
Horn, Laurence 3.653, 3.680
Horn, Paul 2.408
Horn of Africa 1.654, 2.51, 2.268–275, 3.62, 4.275–276
Hornstein, Norbert 2.369–370, 3.685, 4.400, 4.489
Horovitz, Josef 3.718–719
Horrocks, Geoffrey C. 4.522
hortative 3.272
Hotopf, W.H.N. 4.271
Houdas, Octave 3.110–112
Houn 3.53, 3.56

Hourani, Albert H. 2.60, 2.498
Hourani, George Fadlo 4.275
House, A. 2.229
Houssaini, Choukri Iraqi 2.290–293
Hout, Roeland van 2.681
Houtman, Frederick de 1.8
Howell, Mortimer Sloper 2.225–226, 2.424–426, 2.452, 3.453–454, 3.720, 4.119
Hoyland, Robert G. 2.45, 3.470, 3.517
Hoyt, Frederick M. 2.370, 3.381–388, 3.428–434, 3.652–658, 3.676–689, 4.315–320, 4.400, 4.645–652, 4.653–659
Hu Dengzhou 1.380
Huai-sheng, mosque of 1.379
Huang, C.-T. James 4.720
Ḥubayš 4.542
ḥubsa 2.676
Ḥudaydah Arabic 4.754
Huḏayl 1.627, 3.90, 3.92, 3.94, 3.691–695, 3.697, 4.376, 4.430, 4.663
Hudson, Alan 1.631, 4.57–58
Hudson, Grover 3.305, 4.302, 4.305–307
Hudson, Richard Anthony 1.457, 1.613, 1.619, 2.642, 4.189
Huehnergard, John 3.408–422, 3.424, 4.172
Hufūf 1.129
Hufūf Arabic 4.126
Hug 3.410
Ḥujariyyah 4.751
Ḥujariyyah Arabic 1.654–655, 2.274, 4.751, 4.754, 4.756
hugiēs 2.309
Hugonnard-Roche, Henri 4.543
ḥujja 2.85, 2.304, 4.9
ḥukla 2.676
ḥukm 1.193, 2.310, 3.177, 4.12–13
ḥukm šarʿī 2.360
ḥukm taklīfī 2.360
Hulsen, Madeleine 2.681, 2.683
Hulst, Harry van der 3.66
Humai Jilme 2.552
human language technology 1.215
Ḥumaydī, ʾAbū ʿAbdallāh Muḥammad al- 4.132–133, 4.561
Humayun 3.508
Humbert, Geneviève 2.425
Humboldt, Wilhelm von 4.328
Hume, Elizabeth 3.204–207, 3.524–525
humor 2.661
hunāka 3.202
Ḥunayn ibn ʾIsḥāq 4.542, 4.545
Hungarian in the United States 2.683
Hunnemeyer, Friderike 2.191, 2.193, 2.196
hunting poem 4.32
Hunwick, John O. 3.136, 4.759
Hunza Valley 2.217
hupokeímenon 3.176
hupóthesis 3.176
Ḥurr, Ṭannūs al- 1.603
ḥurūf → *ḥarf*
ḥurūf as-sitta, al- 2.223
Husain, Tafuzzul 3.508
Ḥuṣām ad-Dīn, K. Zaki 2.583
Ḥusayn, Šāh Sulṭān 3.341
Ḥusayn, Ṭāhā 1.601, 1.615, 1.617, 2.264, 2.701, 3.349, 3.555–556, 4.500
Ḥusaynī, ʾAḥmad al- 4.98
Ḥusaynī, Mīr ʿAlī 3.337
ḥusnā, al-ʾasmāʾ al- 3.719
Ḥuṣrī, Sāṭiʿ al 1.131, 1.175, 1.453, 2.60, 3.311, 3.345–346
Hussein, Ahmed 2.53–54
Hussein, Anwar A. 4.467
Hussein, Lutfi 3.654, 4.196–197
Hussein, Riyad Fayez 2.501, 2.655–656, 2.658, 2.661, 3.318
Hutchison, John P. 2.553
Ḥuwayṭāt → Ḥwēṭāt
Ḥuwayzah 2.572
huwiyya 3.115
Huwwerin Arabic 4.406
Ḥwēṭāt 3.360, 3.405
Ḥwēṭāt Arabic 3.401–403, 3.407, 4.238, 4.240
hybrid 1.495, 1.633
hybrid act 4.331
hybrid form 1.409–410, 3.12, 3.154, 3.667, 4.218, 4.320, 4.326–327, 4.425, 4.606, 4.610
hybrid loan 4.723
hybridization 1.409
Hyderabad 2.329, 4.508
Hyderabad-Deccan 2.329
hypallage 4.464–465

hyperadaptation 2.276
hyperbole 3.631–632, 3.733
hypercharacterization 2.158–159
hypercoristic 2.17
hypercorrection 1.60, 1.103, 1.106, 1.153, 1.384–386, 1.407, 1.410, 2.255, 2.266, 2.274–279, 2.530, 2.605, 2.631, 3.200, 3.217
hyperforeign 2.277
hyphen 4.464
hypocoristic 1.243, 1.251, 1.328, 1.341, 4.557, 4.559
hypocorrection 2.276–279, 2.530
hyponym 2.585
hypotaxis → subordination
hypothetical 1.554

I

iambic 3.613
iambic pattern 2.350
Iambic/Trochaic Law 3.620
Ibadites 2.552, 3.170
ʿibāra 2.539, 2.551
ʾibdāl 1.313, 2.280–281, 2.299, 3.35, 3.91, 3.649–650, 4.121
ʾibdāl luġawī 2.280
ʾibdāl naḥwī 2.280
Iberian Peninsula, toponyms in 4.508
Iberians 1.619
Ibero-Romance languages 1.545, 2.281–290
ʾibhām 3.539
Ibn ʿAbbās, ʿAbdallāh 3.71, 4.158
Ibn ʿAbd al-Wahhāb, Muḥammad 1.380
Ibn ʿAbd as-Salām 4.27
Ibn ʿAbd Rabbihi 3.208
Ibn ʾAbī Dāʾūd 4.132
Ibn ʾAbī ʾIsḥāq, ʿAbdallāh 2.227
Ibn ʾAbī l-ʾIṣbaʿ 4.103–104
Ibn ʾAbī r-Rabīʿ 2.433, 4.739–740
Ibn ʾAbī Ṭāhir, ʾAḥmad 4.104
Ibn ad-Dahhān 1.627
Ibn al-ʾAjurrūm, ʾAbū ʿAbdallāh Muḥammad ibn Dāʾūd as-Sanhājī 2.189, 2.336, 2.403, 2.539, 2.544
Ibn al-ʾAnbārī, ʾAbū Bakr 2.82, 2.86, 2.165, 2.176, 2.187–188, 2.227, 2.234, 2.291, 2.310, 2.355, 2.402, 2.405, 2.427, 2.433, 2.448, 2.453, 2.539, 2.551, 2.625, 2.630, 2.642–643, 4.13
Ibn al-ʾAnbārī, ʾAbū l-Barakāt 1.626–628, 1.194, 1.310, 3.34, 3.101, 3.104–105, 3.107, 3.109, 3.165, 3.424, 3.542, 3.548, 4.33–34, 4.120, 4.167, 4.220, 4.439
Ibn al-ʾAṯīr, Diyāʾ ad-Dīn 3120, 4.103
Ibn al-ʾAṯīr, ʿIzz ad-Dīn 3.380
Ibn al-ʾAṯīr, Majd ad-Dīn 3.33–34
Ibn al-Basīs 4.132
Ibn al-Bawwāb 1.152, 2.598, 2.601, 3.308–309, 3.340–341, 4.131–133, 4.560–561
Ibn al-Bayṭār 1.100
Ibn al-Furāt 2.186
Ibn al-Ḥājib, ʾAbū ʿAmr ʿUṯmān 1.71, 1.169, 2.188–189, 2.359, 4.119
Ibn al-Jawzī 3.34–35, 4.62, 4.65–66
Ibn al-Jazarī 1.3, 2.293, 3.92, 4.7
Ibn al-Kalbī, Hišām ibn Muḥammad 4.480
Ibn al-Mināwī, Zayn ad-Dīn ʿAbd ar-Raʾūf 2.578
Ibn al-Muʾaddib 2.93
Ibn al-Muqaffaʿ 1.402, 2.331, 2.435, 2.579, 3.109, 3.547, 3.574, 4.62, 4.66, 4.220, 4.416, 4.542, 4.663
Ibn al-Muʿtazz 2.443, 2.496–497, 2.579, 3.117, 4.210–211
Ibn al-Qayyim 4.78
Ibn al-Waḥīd 4.132, 4.562
Ibn al-Warrāq 2.356–357, 3.105
Ibn al-Xabbāz 2.91
Ibn ʿĀmir 4.5
Ibn an-Nadīm 1.151–152, 2.87, 2.598, 3.307, 3.339, 3.514, 3.517, 4.132, 4.295, 4.560–561

Ibn ʿAqīl 2.189, 2.311, 2.479, 3.224, 3.377–378, 4.61, 4.436, 4.538–540, 4.739
Ibn ʿArabī 4.204, 4.211
Ibn ʿArabšāh 3.540
Ibn ar-Rāwandī 2.303
Ibn ʾAsad 3.308, 3.340
Ibn aṣ-Ṣāʾiġ 3.308, 4.132
Ibn aš-Šajarī 3.34, 3.101–105, 3.208, 3.214, 3.226, 3.321, 3.324, 3.345, 3.455, 3.542, 4.165
Ibn as-Sarrāj, ʾAbū Bakr ibn as-Sarī 1.193, 1.398, 1.556–557, 1.559, 2.83, 2.90, 2.94, 2.176, 2.180, 2.184, 2.186, 2.294–295, 2.308–311, 2.316, 2.426–427, 2.433, 2.435–436, 2.447, 2.545, 2.549, 2.551–552, 4.9, 4.102, 4.235–236, 4.412–415, 4.625–626, 4.739
Ibn as-Sikkīt 1.313, 1.627–628, 2.84, 2.280–281, 2.629, 3.35, 3.323, 4.764
Ibn aṭ-Ṭayyib 3.176
Ibn ʿAwf, ʿAbd ar-Raḥmān 4.79
Ibn az-Zamlakānī → Zamlakānī, ʿAbd al-Waḥīd ibn ʿAbd al-Karīm Ibn az-
Ibn Bābašāḏ 2.188
Ibn Bābawayhi 3.346
Ibn Barrī 2.194, 3.33, 4.163
Ibn Bārūn 2.66
Ibn Baṭṭūṭa 2.326, 3.128, 4.276
Ibn Bišr, ʿUṯmān 4.129
Ibn Buklāriš 1.100
Ibn Dāniyāl 2.469, 4.255
Ibn Ḍayfallāh 4.75
Ibn Durayd 2.443, 3.21, 3.32, 3.118, 3.322, 3.346, 3.693, 3.717
Ibn Durustawayhi 1.627, 2.316, 3.321, 4.132
Ibn ʿEzra → Moše ibn ʿEzra
Ibn Fāris, ʾAbū l-Ḥusayn ʾAḥmad 1.194, 1.203, 1.435–436, 1.452–453, 1.627, 1.641, 2.16–17, 2.80, 2.87–88, 2.187, 2.405, 2.426, 2.555, 2.578–579, 2.630, 3.31, 3.35, 3.90–91, 3.93, 3.162, 3.320–321, 3.323–324, 3.346, 3.549, 3.648, 3.651, 3.690, 4.329–330, 4.459, 4.684
Ibn Ḥajar al-ʿAsqalānī 2.52, 3.92, 3.380
Ibn Ḥanbal 3.122
Ibn Ḥawqal 2.325, 3.506
Ibn Ḥayyān 2.298
Ibn Ḥazm 1.98–99, 3.72
Ibn Ḥijja 2.496–497
Ibn Hišām al-ʾAnṣārī 1.49–50, 1.202–203, 1.471, 2.17, 2.82, 2.94, 2.189, 2.193, 2.225, 2.293, 2.316, 2.355, 2.357, 2.359, 2.388, 2.539, 2.543–544, 3.107, 3.539, 3.691, 4.26, 4.38–39, 4.61, 4.275, 4.430, 4.448, 4.538–540, 4.739
Ibn Ḥumayra 2.598
Ibn ʾIsḥāq, ʾAbū Bakr Muḥammad 1.180, 2.199, 3.690, 4.504, 4.727–728
Ibn ʾIyās 1.598, 4.562
Ibn Jinnī, ʾAbū l-Fatḥ ʿUṯmān 1.173, 1.192, 1.273, 1.544, 1.575, 1.614, 2.82, 2.84–85, 2.88, 2.91, 2.94, 2.96, 2.165, 2.176–178, 2.184, 2.188, 2.226, 2.233–236, 2.299, 2.309–311, 2.403–404, 2.406, 2.428, 2.447, 2.450, 2.453, 2.538, 2.544, 2.546–547, 2.551–552, 2.555, 2.626, 2.630, 3.15, 3.45, 3.72, 3.89, 3.92, 3.98, 3.104, 3.109, 3.115, 3.124, 3.320, 3.324, 3.335, 3.345, 3.649, 3.651, 3.717, 3.729, 4.1, 4.9–10, 4.119–121, 4.200, 4.254–255, 4.287, 4.421–422, 4.443, 4.447, 4.684
Ibn Jubayr 2.315, 4.215
Ibn Kamāl Bāšā 2.629, 3.502
Ibn Kaṯīr, ʿImād ad-Dīn, ʾIsmāʿīl ibn ʿUmar 4.5
Ibn Kaysān, ʾAbū l-Ḥasan Muḥammad ibn ʾAḥmad 2.184, 2.427, 4.220
Ibn Kemal Paşa → Ibn Kamāl Bāšā
Ibn Maḍāʾ al-Qurṭubī 1.176, 2.83, 2.181, 2.187, 2.189, 2.310, 4.438–439, 4.441, 4.448
Ibn Makkī 2.629, 4.217
Ibn Mālik, Jamāl ad-Dīn, ʾAbu ʿAbdallāh Muḥammad ibn ʿAbdallāh 1.71, 1.73, 1.191, 1.471, 2.52, 2.82, 2.189, 2.291, 2.355, 3.102, 3.120, 3.178, 4.119

Ibn Manẓūr 1.577, 2.13, 2.84, 2.192, 2.194, 2.280–281, 2.361, 2.428, 2.536, 2.541, 2.555, 2.578, 2.628, 3.21–22, 3.33, 3.40, 3.50, 3.72, 3.89, 3.101, 3.457, 4.33–34, 4.38–40, 4.118–119, 4.163, 4.165, 4.764
Ibn Masʿūd, ʿAbdallāh 2.227, 4.5, 4.8
Ibn Masʿūd, ʾAḥmad ibn ʿAlī 4.119
Ibn Masʿūd, ʿUbaydallāh 1.380, 1.574
Ibn Mayyāda 1.402
Ibn Muḥayṣin 4.7
Ibn Mujāhid 4.5–7, 4.9
Ibn Muqla 1.152, 2.598, 3.307–308, 3.339–340, 4.131–132, 4.560
Ibn Nāʿima al-Ḥimṣī 4.542
Ibn Qalāqis 4.216
Ibn Qurayš, Yehudah 2.66
Ibn Qutayba, ʾAbū Muḥammad ʿAbdallāh ibn Muslim 1.428, 2.87, 2.157, 2.304, 2.443, 2.579, 2.581–582, 2.629–631, 3.34, 3.92, 3.117, 4.4, 4.6, 4.132, 4.202
Ibn Quzmān 1.57
Ibn Rašīq 1.428, 2.497, 3.117, 3.209, 3.647–649, 3.651
Ibn Rušd 3.113, 3.122
Ibn Sābir 2.433, 4.283
Ibn Saḥnūn 2.184
Ibn Sallām 2.304, 4.7
Ibn Šanabūḏ 2.598
Ibn Sīdah, ʿAlī ibn ʾIsmāʿīl 1.99, 1.627, 2.84–85, 2.281, 2.362, 3.32–33, 3.35, 3.321, 4.283, 4.285
Ibn Sīnā 2.200, 2.234–235, 2.237, 3.113–115, 3.176, 3.178, 3.502, 3.579, 3.602–603, 4.32
Ibn Sūdūn, ʾAbū l-Ḥasan ʿAlī 1.598
Ibn Šuqayr, ʾAbū Bakr ʾAḥmad ibn al-Ḥasan 4.436
Ibn Ṭabāṭabā 3.117
Ibn Ṭaymiyya 1.380, 4.78
Ibn Ṭulūn, ʾAḥmad 4.589
Ibn Tulun, mosque of 2.42
Ibn ʾUmm Šaybān 2.598
Ibn ʿUṣfūr al-ʾIšbīlī 1.192–193, 1.555, 2.91, 2.351, 2.357, 2.537–538, 2.623–624, 3.226, 3.648, 3.651, 4.119–121, 4.200, 4.447, 4.740
Ibn Wahb 2.194
Ibn Xalaf an-Nayramānī 3.307, 4.132, 4.561
Ibn Xālawayhi, ʾAbū ʿAbdallāh al-Ḥusayn ibn ʾAḥmad 2.82, 2.88, 2.350, 4.9
Ibn Xaldūn 1.130, 1.259, 1.273, 2.178, 2.184, 3.110, 3.112, 3.170, 3.373, 3.375, 3.510, 4.3, 4.105, 4.572
Ibn Xallikān 2.597, 3.720
Ibn Xaššāb 2.433
Ibn Xaṭīb ad-Dahša 3.717
Ibn Yaʿīš, Muwaffaq ad-Dīn Yaʿīš 1.49, 1.555–559, 2.79, 2.91, 2.93, 2.165, 2.168, 2.226, 2.237, 2.291, 2.311, 2.351, 2.427, 2.433, 2.448–450, 2.453, 2.478, 2.539, 2.545–550, 2.552, 2.555, 3.224, 4.119–121, 4.235–236, 4.442–445, 4.625–626, 4.737–740
Ibn Yāsīn 3.170
Ibn Zurʿa 3.114, 3.177–178
Ibra Fall, Shaykh 4.180
ʾIbrāhīm, ʿAbd al-ʿAlīm 2.110
ʾIbrāhīm, ʾAbū l-Faḍl 2.628
Ibrahim, Hafiz 4.88, 4.90, 4.213
ʾIbrāhīm, Muḥammad Ḥassān 1.595, 1.631, 1.667–668, 2.109, 2.156–159, 2.642, 2.657, 3.12, 3.76, 4.51, 4.629, 4.631, 4.681
Ibrahim, Zeinab 2.654–655, 2.657–658, 3.13–21
ʿibrī 1.132
Ibriszimow, Dymitr 2.68
Ibro, Salim Alio 4.275
ʾibtāʿ 1.303
ibtidāʾ 1.182, 1.202, 1.488, 1.557, 2.132, 2.184, 2.290, 2.357, 2.537, 3.177–178, 4.317, 4.426, 4.484, 4.738
Icel 4.580
Icelandic 4.540
Icelandic, object displacement in 4.540–541
i-coloring environment 2.417
iconicity 1.357, 4.50, 4.223, 4.226–227

iconicity, semantic 1.74, 1.76
Iconicity Principle 2.113–114, 2.116
ictus 3.209
ʾiḏ 1.467–469, 1.476, 3.200
ʾiḏā 1.110, 1.466–468, 1.476, 2.477–478, 3.703, 3.705, 4.60
ʾiḍāfa 1.55, 1.124, 1.386, 1.560, 1.563, 2.197, 2.294–298, 2.327, 2.403, 2.409, 2.426–427, 2.436–437, 3.68, 3.244, 3.258, 3.377, 3.440, 3.568, 3.630, 3.650, 3.672, 4.27, 4.430, 4.488–489
ʾiḍāfa ġayr ḥaqīqiyya 2.297, 3.431, 3.684, 4.102
ʾiḍāfa ġayr maḥḍa 2.294
ʾiḍāfa maḥḍa 2.295
ʾiḍāfa taqdīriyya 2.411
ʾīḍāḥ 2.579
ʾiḏan 1.467
iddiʿāʾ 2.442–444
iddiġām → *ʾidġām*
identification 4.321
identification, automatic 1.207
identity 2.655, 3.11, 4.321
identity, African 2.518
identity, Algerian 3.350
identity, Amazigh 2.710, 2.713–715
identity, Arab 1.531, 2.60
identity, Assyrian 2.61
identity, Berber 2.63, 3.350
identity, collective 2.64
identity, ethnic 1.415
identity, group 2.217, 2222
identity, Islamic 1.258
identity, Jewish 2.64
identity, Moroccan 2.64
identity, Muslim 1.528
identity, national 1.370, 3.20
identity, regional 2.215
identity, religious 2.59
identity, social 2.202
identity marker 1.525
identity test 1.586
ideologization 2.58
ideophone 1.367, 4.284, 4.714
ʾidġām 1.492, 2.233, 2.280, 2.298–300, 2.311, 3.92, 3.335, 4.6, 4.10, 4.23, 4.121, 4.427
ʾidġām ʾaṣġar 2.299
ʾidġām kabīr 2.299
ʾiḏ̣hār 2.302, 3.335
Idi Amin 2.588
idiom 1.658, 3.28, 3.540, 3.572, 3.629–630, 3.633, 3.655, 3.731
idiom, allusive 3.632
idiom, aphetic 3.630
Idiom Principle 1.658
idiomatic dual 3.630
idiomaticity 3.628
idiomaticization 4.38
ʾiḍmār 2.16, 2.82, 2.290, 2.300–302, 4.446–447, 4.738
ʾiḍmār qabla ḏ-ḏikr 4.739
ʾiḍrāb 1.124
ʾIdrīs, ʾAḥmad 2.328
Idris, Helene Fatima 3.435
ʾIdrīs, Muḥammad 1.603
ʾIdrīs, Nisma 1.599
ʾIdrīs, Yūsuf 1.599, 1.601, 3.633
ʾIdrīsī, Muḥammad Ibn aš-Šarīf al- 4.216
Idriss, Moḥammed 2.159–160
Idriss, Souheil 3.28
Idrissi, Ali al- 1.97, 1.361, 1.425, 2.367, 2.691–692, 3.304, 4.96, 4.358
iḍṭirār 3.90
Ifrīqiya 4.215, 4.218, 4.572
ʾifṣāḥ 2.579
iftaʿala 3.226
iftixār 4.329
Igbo 3.373
Igli [Igni] Arabic 3.290–291
ʾiġrāʾ 2.432–433
ʾiḥāla 4.24
ʾīhām 3.538
iḥtijāj 4.9, 4.22
iḥtimāl 3.322
ʾiʿjām 1.151
ʾiʿjāz al-Qurʾān 1.174, 1.427–428, 2.17, 2.264, 2.302–307, 2.443, 3.117, 3.162–163, 3.346, 4.21, 4.105, 4.203, 4.473–474
ʾijāza 3.116, 3.136, 4.134, 4.562
ʾĪjī, ʿAḍud ad-Dīn ʿAbd ar-Raḥmān ibn ʾAḥmad al- 4.684–685
ʿIjlī, ʾAbū n-Najm al- 3.650
ʾijmāʿ 1.194, 2.88, 2.305, 4.13

ʿIjmān 3.326, 4.125
ijtihād 2.185
ijtizāʾ 1.616
Ikhshidid Dynasty 4.589
Ikhwan Movement 1.380–381
ʿilājī 4.625
ʿilal → *ʿilla*
Ilappais 2.325
ʾilġāʾ 1.558, 2.307–308, 2.356, 4.410, 4.428
ʿIlī, Šayx ʿAbdī 4.277
Ilkhan 1.403
ʿilla 1.71, 1.73, 2.164, 1.176, 1.193–194, 1.467, 1.557, 2.178–179, 2.186, 2.306, 2.308–311, 2.452–453, 2.629, 3.101, 3.104, 3.210, 4.7, 4.12–13, 4.122, 4.447, 4.545
ʾillā 2.452–453
ʿilla jadaliyya 2.179
ʿilla naḏariyya wa-jadaliyya 2.187, 2.310
ʿilla qiyāsiyya 2.179, 2.187, 2.310
ʿilla taʿlīmiyya 2.179, 2.187, 2.310
ʿillat al-ʿilla 2.309–310
illiteracy 1.533, 2.75, 2.498, 2.704, 3.74–80
illiteracy, female 2.648, 3.75
illocution, conventional 4.332
illocution, denial 2.144
illocution, derived 2.144
illocution, disapproval 2.144
illocution, nonconventional 4.332
illocution, nonliteral 2.144
illocutionary 2.360, 4.328, 4.472
illocutionary act 2.80–81, 3.679–680
illocutionary commitment 4.328
illocutionary derivation 4.331
illocutionary force 1.428, 1.657, 3.268, 3.570, 3.680–681, 4.328
ʿilm al-ʾaṣwāt 2.183
ʿilm al-ʿawāmil 2.403
ʿilm al-badīʿ 3.538
ʿilm al-balāġa 2.358, 3.538, 3.650
ʿilm al-bayān 2.306, 2.441, 3.116, 3.120, 3.538
ʿilm al-kalām 3.122
ʿilm al-lisān 3.73
ʿilm al-luġa 2.183, 3.73
ʿilm al-maʿānī 2.306, 2.358, 2.360, 3.118, 3.120, 3.161, 3.538, 3.650, 4.330
ʿilm al-waḍʿ 3.35, 4.684
ʿilm an-naḥw 2.184
ʿilm aṣ-ṣarf 2.183, 4.119, 4.207
Ilorin 4.760
iltibās 2.624, 4.102
iltifāt 2.333, 4.23–25, 4.29–30
iltiqāʾ as-sākinayn 3.208, 3.728, 4.116
iltizām 3.120
Iltutmish 2.326
ʾīmāʾ 2.579
ʿimād 1.557
image, coherence of the 4.213
ʾiʿmāl 2.169, 4.438
ʾimāla 1.89, 1.103, 1.113–114, 1.234, 1.243, 1.245, 1.262–263, 1.278, 1.300, 1.324, 1.390, 1.497, 1.537–538, 1.547, 1.571, 1.610, 1.612, 1.614, 1.640, 2.233, 2.242, 2.263, 2.311–315, 2.417, 2.483, 2.533–534, 2.612, 3.55, 3.91, 3.148–149, 3.220, 3.276, 3.327, 3.362–363, 3.481, 3.527, 3.530, 3.585, 3.595, 3.607, 3.610, 3.693–694, 4.6, 4.10, 4.109, 4.218, 4.404, 4.406–407, 4.421, 4.427, 4.550, 4.615, 4.678, 4.681
ʾimāla, final 2.312–314
ʾimāla, internal 4.678
ʾimāla, medial 2.311–314
ʾimāla, pausal 2.3–4
ʾimām 4.4
ʾImām, ʾAḥmad ʿAlī al- 3.94
Imaziġen 2.709
Imberg, Gosta 1.660, 1.663
imitation 4.283–284
ʾimmā 1.469
immigrants, Arab in Senegal 4.179
immigrants, Arab in Sudan 4.375
immigrant communities 1.370
immigrant workers 1.418
immigration 1.482
immigration, Italian in Tunisia 4.217
immigration, urban 4.716
immigration in Saudi Arabia 4.123
imminence 1.268
impairment, processing 2.689

Impairment, Specific Language 4.145
imparfait 4.728
imperative 1.217–218, 1.239, 1.265, 1.282, 1.310, 1.340–342, 1.498–499, 1.509, 1.523, 1.538, 1.568, 1.574, 2.91, 2.359, 3.249, 3.263, 3.266–267, 3.269, 3.272, 3.299, 4.639
imperative, narrative 1.110, 2.490, 3.368, 3.536
imperative, negative 1.249, 2.596, 3.248
imperative statement 4.331
imperfect 1.38, 1.44, 1.196–197, 1.217–218, 1.220, 1.309, 1.568, 1.574, 1.624, 2.426, 2.625, 3.67, 3.96
imperfect, *aktib/niktib* 4.247
imperfect, *ʿam-* 1.553, 3.265–266
imperfect, *ʿammāl* 1.331, 1.608
imperfect, apocopated 3.269
imperfect, *b(i)-* 1.197–199, 1.219–220, 1.327, 1.331, 1.408–409, 1.538, 1.553, 1.607, 2.271, 2.516, 2.618, 3.264–265, 3.404, 3.481, 3.533, 3.573, 4.241, 4.249, 4.650, 4.712
imperfect, *bayn-* 1.607
imperfect, *dā-* 1.229, 1.239, 1.610, 1.612
imperfect, dependent 3.269
imperfect, *gāʿid* 2.568–569, 2.618
imperfect, *gi-* 2.595
imperfect, *ḥa-* → future particle *ḥa-*
imperfect, *kū-* 2.423
imperfect, *niktib/niktibu* 1.323, 1.589, 1.605, 1.610, 2.3–5, 2.7, 2.533–534, 2.631, 3.53, 3.99, 3.146, 3.220, 3.291, 4.554
imperfect, *qā-* 1.239, 1.610, 2.423
imperfect, Semitic 4.177
imperfect prefix 1.399, 1.533
imperfective aspect 1.195, 3.68, 3.99, 3.247, 4.454, 4.457
impersonal 1.553
impersonal passive → passive, impersonal
impersonal structure 3.558, 3.562
impersonal verb → verb, impersonal
impersonalization 2.317
implication 2.578, 2.581
implication, bilateral 2.319
implication, relationship 2.319
implication, statistical 2.322
implication, unilateral 2.319
implicational analysis 2.320–321
implicational relationship 2.320–322
implicational scale 2.318–325, 3.606
implicational statement 2.318
implicature 1.428, 4.455, 4.457
Implicature Model 4.163
implicature, conversational 2.70, 4.331
implosive 1.366
importation 2.31
importation of phonemes 2.127
imposition 4.370–371
imposition of speech 2.178, 4.684–685
imposition of speech → creation of language; invention of language
imprecative 2.71, 2.586–587
Imruʾ al-Qays 1.128, 2.305, 2.581, 3.211, 3.469, 3.499, 3.646–647, 3.651, 4.52, 4.201, 4.210, 4.212, 4.514
imtināʿ 4.331
ʾin 1.202, 1.468, 1.509, 2.355, 2.477–480, 2.542, 3.272, 4.60, 4.544
in (kān) 1.110
ʾin al-muxaffafa 1.202, 2.355
ʾin an-nāfiya 1.202
in situ 2.390, 2.596, 4.719–722
inalienable 1.83, 1.110, 1.283, 1.396, 1.522, 1.540, 2.523, 3.152, 3.183, 3.432, 3.671–674
inalienable possession 1.327, 2.570, 2.594
ʿInānī, Muḥammad Muḥammad 4.285
inanimate 1.34, 1.41–43, 1.220, 1.344, 1.346, 1.359, 1.433, 1.436, 3.446
ʿināya 2.627
inception 1.219–220
inceptive 1.199
inchoative 1.198–199, 1.282, 1.524, 1.553, 2.245, 2.272, 2.524, 3.221, 3.234–236, 3.252–253, 4.196, 4.490–491
inchoativity 4.457
inclusion 4.370
inclusiveness 3.492

incorporation, numeral 4.228
incorporation, pronoun 1.486
Increasing Complexity Principle 2.147
ʿinda 3.675
indeclinability 1.308
indeclinable 1.559, 4.443
indefinite 1.190, 1.522, 1.560, 1.644–645
indefinite article → article, indefinite 1.31, 1.92, 1.106, 2.17
indefinite marker 1.491, 1.606, 1.611, 2.419, 2.422
indefiniteness 1.94, 1.188–190, 1.226, 1.283, 1.478, 1.480–481, 1.579–580, 1.582–583, 2.132, 2.146, 2.247, 2.296, 2.353, 2.424, 2.594, 3.55, 3.82, 3.87, 3.186, 3.284, 3.429, 3.453, 3.535, 3.548, 4.311–312, 4.444
indefiniteness, sentence-initial 3.87
indeterminacy 3.205
indeterminacy in the signal 3.206
indetermination 3.227
indetermination, marker of 4.444
indexicality 2.654
India 2.211, 2.325–331, 2.439
India, Arabic in 4.595–603
Indian Ocean 1.660, 2.333
Indian Ocean trade 1.545
Indians 2.212, 2.609
indicative 1.562, 2.23, 2.92, 3.98, 3.263–264, 3.269, 3.272, 3.298
indirect object 3.182, 3.277
indirect object suffix 1.247, 1.302–303, 1.326, 1.391, 2.613, 3.276, 3.364, 3.483
indirect question 2.393
indirect speech 1.367, 1.649, 2.331–333
indirect speech, free 2.331–332
individuation 2.353, 3.428, 3.430, 4.316
Indo-European languages 1.573–574, 2.67, 2.403, 3.141, 3.192, 3.204, 3.700, 4.96–97
Indonesia 1.529, 2.333–339, 2.472
Indonesia, Arabic in 2.337–338
Indonesian [Bahasa Indonesia] 1.545, 2.26, 2.333–345, 2.668–670
Indonesian, Betawi 2.668
Indo-Persian 4.595, 4.597, 4.599, 4.601–603
infaʿala 3.226
infinitive 1.488, 1.521, 1.523, 3.166
infinitive, Coptic 1.498–499
infinitive, French 1.418
infirād 3.322
infix 2.349, 2.352, 3.297
infixation 1.576
INFL 3.84, 3.86
INFL → inflection
inflection 1.348, 1.350, 1.486, 1.574, 2.171, 2.174, 2.345–354, 3.297, 4.93, 4.638–639, 4.642, 4.743
inflection, acquisition of 2.101
inflection, external 2.346–347, 2.349, 2.352–353
inflection, internal 1.37, 1.574, 2.346–347, 2.349, 2.353, 4.638
inflection, loss of 3.61
inflection, verbal 3.708–709
informants 2.87
information 1.427, 1.568, 2.186, 2.207–208
information, displaced 4.502
information, frame-related 4.502
information, given 4.494, 4.497, 4.502
information, known 4.738
information, new 2.113, 2.435, 4.494, 4.497, 4.502–503, 4.657, 4.738
information, old 2.113, 4.502–503
information, propositional 4.503
information and communication technology 2.695–696
information retrieval 1.207, 1.209, 1.214, 1.455, 1.458–460, 1.514–515, 2.380
information storage 2.380
information structure 1.486, 4.503
informativity 4.472, 4.475
Ingela 1.655
Ingham, Bruce 1.28–35, 1.272, 1.570, 1.584, 1.590, 1.606, 1.611, 1.617, 2.38, 2.152, 2.213, 2.404, 2.407–408, 2.571–578, 2.609,

3.179–180, 3.326–334, 3.403, 3.429, 3.478, 3.585, 3.587–588, 3.606, 4.123–130, 4.318, 4.612, 4.617, 4.654–655
Ingram, David 2.676, 2.688
ingressive 1.491
in-group 1.415, 4.255, 4.321
in-group vocabulary 2.217
inimitability of the *Qurʾān* → *ʾiʿjāz al-Qurʾān*
initial position 1.474
initialism 1.1, 1.5
injective 1.366
injunctive 1.281–282, 3.191, 4.639
ʾinna 1.69, 1.202–203, 1.349, 1.468–469, 1.471, 1.473, 1.476, 1.558, 1.563, 1.649, 2.144, 2.147, 2.171, 2.179–180, 2.187, 2.237, 2.290–291, 2.316, 2.333, 2.354–358, 3.101, 3.105, 3.161, 3.221, 3.234, 3.256, 3.260, 3.268, 3.353–354, 3.589, 3.591, 3.691, 3.705, 4.358–362, 4.430, 4.485–486, 4.499, 4.504, 4.544, 4.732, 4.740
ʾinna muxaffafa 1.69
ʾinna wa-ʾaxawātuhā 1.69
ʾinna with accusative predicate 3.691
ʾinna with nominative subject 3.695
ʾinnamā 1.471, 2.115, 4.732
innateness 2.96, 2.689, 3.228
Inneqor 4.301
innovation 1.492–493, 2.7, 3.189, 3.202–203
innovation in enclaves 1.590
Inor 4.304–305
input structure 4.521
inqaṭaʿa 3.96
inqirāʾiyya 3.74
ʾinšāʾ 2.81, 2.327, 2.358–361, 4.331–333, 4.586
ʾinšāʾ ʾīqāʿī 2.358–359
ʾinšāʾ kitābī 2.358
ʾinšāʾ ṭalabī 4.332
inscription, foundation 2.43
inscription, Greek 2.212
inscription of al-Aqmar Mosque 2.600
inscription of al-ʿUla 1.129, 1.149, 3.467–469, 3.471–472
inscription of ʿĒn Avdat 1.149, 3.467–469
inscription of Friday Mosque Isfahan 2.601
inscription of Ḥarrān 1.149, 3.467, 3.470
inscription of ʿIgl bn Hfʿm 3.467
inscription of Jabal ʾUsays 3.467, 3.470
inscription of Namāra 3.467, 3.469
inscription of Qaryat al-Fāw 3.467
inscription of the Dome of the Rock 1.401, 2.599, 3.515, 3.517, 4.130, 4.134
inscription of ʾUmm al-Jimāl 1.149, 3.470
inscription of Zabad 1.149, 3.470, 3.476
inscriptions 2.40, 2.42–43
inscriptions, Alhambra 3.112
inscriptions, Arabic in India 2.329
inscriptions, Aramaic 1.401
inscriptions, bilingual 4.482
inscriptions, Ḥimyaritic 2.256–257, 2.260
inscriptions, historical 2.44
inscriptions, Jordanian 2.263
inscriptions, Kufic 2.599–600
inscriptions, Libyco-Latin 4.571
inscriptions, Libyco-Punic 4.571
inscriptions, monumental 4.560
inscriptions, Nabataean 2.198, 3.696, 4.479
inscriptions, North-Arabian 2.263
inscriptions, Ottoman architectural 4.561
inscriptions, Palmyrene 1.401
inscriptions, pre-Islamic 2.260
inscriptions, Qurʾānic 2.45
inscriptions, Sabaean 2.256
inscriptions, Syrian 2.263
inscriptions, Taj Mahal 4.561
inscriptions of Jabal Ramm 3.469
insertion 1.418, 2.127, 3.617
insertion, lexical 4.521
insertion, suppletive 2.302
insijām 1.427
Institut Bourguiba des Langues Vivantes 4.149
Institut pédagogique polyvalent Bella Assan 1.337

Institute of Kiswahili and Foreign Languages 1.661
Institute of Kiswahili Research 1.661
instrument, noun of 1.328, 3.279, 3.153, 3.303, 3.424, 4.95, 4.461–462, 4.487, 4.535
instrumental 1.344, 1.346
insult 2.136, 2.361–365, 2.586–587, 2.622, 4.419
integration 4.370
integration of loanwords 2.288, 2.342, 2.465, 4.724
intensification 1.308, 4.51
intensification strategy 2.206
intensifier 1.306
intensity 1.211, 1.329, 1.331, 3.186
intensive 2.156–157, 4.177, 4.643, 4.692
intensive adjective 1.16
intensive aorist 1.38
intent 1.284–285
intent, declaration of 2.578
intention 1.330–331, 1.607, 2.301, 3.160, 4.329
intention, communicative 2.209, 2.623
intention, speaker's 2.186
intentional 1.553
intentional meaning 4.329
intentionality 1.359, 1.429–430, 4.472, 4.475
interaction management 2.148
interaction, face-to-face 4.322
interaction, formal 4.58
interaction, informal 4.58
interdental 1.29, 1.53–54, 1.59–61, 1.88, 1.102, 1.112, 1.223, 1.232, 1.242, 1.260, 1.270–271, 1.275, 1.291, 1.300, 1.323–324, 1.365–366, 1.384, 1.390, 1.405, 1.407, 1.409–410, 1.440, 1.490, 1.493, 1.501, 1.546, 1.571, 1.585, 1.590, 1.607, 1.609–611, 2.2, 2.31, 2.213, 2.415, 2.473, 2.482, 2.507, 2.572, 2.610, 3.55, 3.146, 3.180, 3.189–190, 3.220, 3.275, 3.289–290, 3.361, 3.478–479, 3.598, 3.604, 3.608–609, 3.666, 3.668, 4.124, 4.126, 4.128, 4.240–241, 4.302, 4.370, 4.372, 4.548–549, 4.564, 4.613–614, 4.629, 4.631–632, 4.754, 4.757
interdentals, Semitic 4.172
interdialectal contact 3.8
interdialectal contact → speech accommodation
interdialectal conversation 1.406–407
interdigitation 1.573
interface, morphology-syntax-semantics 2.366
interface, phonology-morphology 2.366
interface, syntax-semantics 2.368–371
interface linguistics 2.365
interference 1.416, 1.440, 1.485, 2.373–378, 4.368, 4.370
interference → substratal influence; adstratal influence
interference, Arabic in Ottoman Turkish 4.585
interference, first language 4.139–140
interference, Greek in Arabic translations 4.544
interference, linguistic 2.666
interference, Persian in Ottoman Turkish 4.585
interference, structural 2.669
interference, Syriac in Arabic 4.546
interference, Tajik in Uzbekistan Arabic 4.614–615
interference, Turkish in Arabic 4.581, 4.593–594
interference, vernacular 2.529
intergroup encounter 4.321–322
interjection 1.73, 1.367, 1.566, 2.272, 2.378–380, 2.432, 2.585, 3.97, 3.103, 3.257, 4.287
interjection, demonstrative 1.63
interlanguage 1.482, 1.485, 4.139–141, 4.143–144, 4.369
intermarriage 2.59, 2.711
internal cause of change 3.189
internal drift 1.594
internal inflection → inflection, internal
internal passive → passive, internal

internal plural → plural, internal
International Language Institute 4.149
International Phonetic Alphabet 4.520
Internet 1.207, 1.455, 1.512, 1.602, 2.380–387, 2.701, 3.193, 3.199, 3.274, 3.317, 4.320, 4.322
Internet, Arabic on the 3.316
interpretation 1.429
interrogation 2.101, 2.143
interrogation, implicit 4.485
interrogative 1.230, 1.240, 1.392, 1.502, 1.509, 2.171, 2.195, 2.397, 2.521, 2.561–562, 3.293, 3.365, 4.247
interrogative, indirect 2.393
interrogative, phrasal 2.388
interrogative, postposition of 1.502
interrogative, *wh-* 4.719–720, 4.722
interrogative adverbial 1.22, 1.24–25
interrogative determiner 2.392
interrogative marker 4.566
interrogative pronoun 1.279, 1.303
interrogative statement 4.331
interrogative utterance 4.332
interrogatives in Egyptian Arabic 4.372
interrogatives, acquisition of 2.100
intertextuality 4.472, 4.475
interval 1.197–199, 1.201
intihāʾ 2.191
intimacy 2.587
intonation 1.445, 2.116, 2.391, 2.395–401, 2.490, 2.532, 2.574, 3.240, 3.569, 3.726–727
intonation, interrogative 1.249
intonational contour 2.395
intonational phrase 2.396
intransitive 1.328, 1.348, 1.356–357, 4.410–416, 4.530, 4.536, 4.624, 4.641–642
intransitive verb 1.221, 1.281, 1.622–623, 1.625, 2.93–95, 3.224, 4.532
intransitivity 3.225
intransitivization 1.623–624
introduction, topic 1.432
introflective 4.93
introflexion 3.192
intrusive *-in* 3.544
invective 2.362
invention of language 2.178, 4.684–686
invention of language → imposition of speech; creation of language
inversion 2.224, 3.83, 4.739
inversion strategy 3.82
inversion, free subject 3.706
invocation, divine 2.364
Inzidenzschema 1.468
iotacism 2.201
IP → phrase, Inflection
IPA → International Phonetic Alphabet
ʾīqāʿ 2.359
ʾīqāʿī 2.358
ʾiqḥām 2.296
ʾiqlāb 3.335, 4.427
iqtiṣār 2.17
ʾiqwāʾ 4.88, 4.207–208
ʾiʿrāb 1.67, 1.308–310, 1.471, 1.559–560, 1.562, 1.601, 1.614–616, 1.666, 1.668, 2.85, 2.179–180, 2.232, 2.263, 2.293, 2.308, 2.353, 2.401–406, 2.479, 2.536–539, 2.542–543, 3.90, 3.92, 3.429, 3.551, 3.645, 3.650–651, 3.696, 4.22–23, 4.93, 4.432, 4.446, 4.639
Irafayle 2.268–269
Iram 3.345
Iran 1.403, 1.529, 2.219, 2.406–414, 4.508
Iran, Arabic in 2.407
Iranian languages 2.219, 2.604
Iranian Revolution 4.77–78
Iran-Iraq war 2.407
Iraq 1.26, 1.222, 2.105–107, 2.109, 2.424
Iraq, Arabic in 2.555, 3.237, 3.525
ʿirāqī 1.80, 1.83–84, 1.93–94, 1.151, 1.189, 3.307, 4.131
Iraqi Arabic 1.87, 1.92, 1.95, 1.190, 1.222, 1.417, 1.432, 1.443–444, 1.493, 1.540, 1.639–641, 2.38, 2.96, 2.174, 2.231, 2.312, 2.388, 2.390, 3.68, 3.70, 3.429, 3.595, 3.597, 3.611, 3.613–614, 3.654, 3.702, 3.729, 4.45, 4.85–86, 4.303, 4.311, 4.318, 4.590, 4.592, 4.670–671

Iraqui-Sinaceur, Zakia 1.609
Irbid 2.499
Irbid Arabic 2.498–499, 2.503, 2.506, 4.655
ʿirḍ 2.363, 3.658
ʾirdāf 2.579–581
Iregat 1.362
Irish 3.706, 4.392
irony 3.121
ʿirq 3.345
irrealis 1.554, 2.477, 2.524, 3.269, 3.537, 3.643
ʾirsāl 3.91
Irshied, Omar Musleh 2.501, 3.610, 4.349, 4.351
Irvine, Judith 2.651, 2.662
ʿIryānī, Muṭṭahar ʿAlī al- 2.258, 4.107
ʿĪsā ibn ʿUmar aṯ-Ṯaqafī 2.86, 2.227, 3.543, 4.7
ʿĪsā ibn Yaḥyā 4.542
Isa, Ali 2.469
Isaksson, Bo 3.585, 4.407
ʾišāra 1.565, 2.579, 2.581, 3.160
ʾišbāʿ 1.616, 2.234, 3.124
ʾIṣbahānī, ʾAbū l-Faraj al- → ʾIṣfahānī, ʾAbū l-Faraj al-
ʾIšbīlī, ʾAbū l-Xayr al- 1.3
ʾIsbīr, Muḥammad Saʿīd 3.539
ʾIṣfahānī, ʾAbū l-Faraj al- 1.616, 2.433–434, 2.550, 3.216
ʾIṣfahānī, ar-Rāġib al- 3.34
ʾIṣfahānī, Bābā Šāh 3.337
ʾIṣfahānī, Ḥabīb 2.407, 2.601
ʾIṣfahānī, Muḥammad al- 3.540
ʾIsfarāʾinī, ʿIṣām ad-Dīn al- 4.686
ʾIsḥāq ibn Ḥunayn 4.542
ʾIsḥāq ibn ʾIbrāhīm 3.117
ʾIsḥāq, Muḥammad Qamar 2.326, 2.328
ʾIsḥāqī, ʿAbd al-Bāqī al- 1.598
Ishihara, Tadayoshi 3.287
Isidorus of Seville 1.98
ʾiskān 1.614–616, 2.16, 2.165
Iskenderun 1.111, 1.114–116, 4.580
Isla 3.147
Islam 2.664
Islam, language of 1.530
Islam, Shiʿite 2.406
Islam, spread of 2.667
Islamic College 1.661
Islamic conquests 1.174
Islamic Cultural Association of Cameroon 1.336
Islamic education 3.138–139
Islamic fundamentalism 2.698, 2.703–704
Islamic law 2.578
Islamic law → *fiqh*
Islamic law, terminology of 2.341
Islamic learning 1.450, 2.51–53, 2.325, 3.136, 3.170, 3.374–376
Islamic learning in Indonesia 2.336
Islamic World League 1.450–451
Islamicization 2.58
Islamists 1.531
Islamiya college 3.129
Islamization 1.55–56, 1.98–99, 1.286, 1.654, 2.340, 3.169–170, 3.641
Islamization of nationalism 4.77
Islamization of Pakistan 3.511
Islamization of Senegal 4.723
Islamization of Somalia 4.272
Islamization of South India 4.433
Islamization of Uzbekistan 4.612
island 2.369–370, 4.522
island constraint 4.82, 4.523
island violation 4.81
Island-sensitivity 3.384
ism 1.185, 1.309, 1.488, 1.565, 2.90, 2.184, 2.424–429, 2.542, 2.545, 3.241, 3.323–324, 3.422, 3.424, 3.546, 3.717, 3.719, 4.220–221
ism → noun
ism aḏ̣-ḏ̣arf 3.424
ism al-ʿadad 3.422
ism al-ʾāla 2.424, 3.424, 4.461
ism al-ʿalam 3.717, 3.719
ism al-fāʿil 1.68, 2.297, 2.424, 2.429–432, 2.550–551, 2.624, 3.96–98, 3.424, 3.459, 3.542, 3.548–549, 4.220
ism al-fiʿl 1.73, 2.80, 2.379, 2.432–434, 3.103, 3.679, 4.282–283, 4.414, 4.443

ism al-ḥadaṯ 3.102
ism al-ḥadaṯān 3.455
ism al-ʾišāra 1.566, 1.568, 3.422
ism al-jamʿ 1.433
ism al-jins 1.433
ism al-kaṯra 1.577, 3.425
ism al-mafʿūl 1.68, 2.424, 2.550–551, 2.624, 3.424, 3.459, 3.542, 3.548
ism al-marra 3.167, 3.424, 4.644
ism al-maṣdar 1.68, 3.165
ism al-muṯallaṯ, al- 4.468
ism al-waḥda 1.433, 3.424, 3.439
ism al-wiʿāʾ 3.424
ism an-nawʿ 3.424, 4.644
ism at-tafḍīl 2.13
ism az-zamān 3.424
ism fiʿl 3.101
ism ʾiḍāfa 2.294
ism ʾinna 1.69, 2.355, 2.357
ism jāmid 1.573, 2.445, 4.686
ism kāna 1.70, 2.431, 2.549, 2.551–552, 3.354, 3.687
ism li-l-fiʿl 2.432
ism makān 2.424
ism mawṣūl 1.488, 1.568, 3.422
ism mubham 1.73, 4.221
ism muḍāf 1.69
ism muḍmar 1.557
ism muḥaqqar 1.637
ism muṣaġġar 1.637
ism ṣawt 3.176
ism tāmm 1.69
ism ṯulāṯī 3.721
ism, definition of 2.424, 2.427
ism, tamām al- 4.445
ʾIsmāʿīl 3.345
Ismail, Hanadi 4.630
ʾišmām 2.233, 3.566, 4.427, 4.682
Ismint Arabic 2.2
ismiyya 2.434
ʾIšmūnī, Aḥmad al- 4.291
Isna Arabic 2.2
ʾisnād 1.2, 2.291, 2.294, 2.316, 2.427, 2.424–437, 2.538–539, 3.136, 3.683, 4.354, 4.484, 4.686
ʾisnādī, tarkīb 2.404
isogloss 1.583–586, 1.604–605
isogloss, distinctive 1.585
isogloss, formal 1.585
isogloss, non-distinctive 1.585
isogloss, structural 1.585
isogloss bundle 1.604, 3.406
isogloss clustering 3.10
isomorphism 3.612
Israel 1.131, 1.176, 1.232, 2.60, 2.64, 2.97, 2.106, 2.437–441, 2.531, 2.660, 3.312, 3.315
Israel, Arabic in 1.584, 2.437–441
Israelis 2.438
ʾIsrāʾīlī, ʾIsḥāq ibn Sulaymān al- 3.111
ʾišrāk 2.292
Issaq 4.277
Isserlin, Benedict S.J. 3.142, 3.147
istaʾnafa 2.293
Istanbul 1.168, 1.171–172, 4.425
ištaqqa 4.446
istaṯqala 4.442
istiʿāra 1.424–425, 2.305, 2.441–447, 2.578, 3.116–121, 3.651, 4.165, 4.207, 4.210
istiʿāra ʾaṣliyya 2.444–445
istiʿāra bi-l-kināya 2.445
istiʿāra dāxiliyya 2.445
istiʿāra ġayr dāxiliyya 2.445
istiʿāra ġayr mufīda 2.443–444, 3.121
istiʿāra ʿinādiyya 2.445
istiʿāra makniyya 2.445
istiʿāra mufīda 2.443
istiʿāra mujarrada 2.445
istiʿāra muraššaḥa 2.445
istiʿāra muṭlaqa 2.445
istiʿāra ṭabaʿiyya 2.444–445
istiʿāra taḥakkumiyya 2.446
istiʿāra taḥqīqiyya 2.445, 3.120
istiʿāra tamlīḥiyya 2.446
istiʿāra taṣrīḥiyya 2.445
istiʿāra taxyīliyya 2.444–445, 3.120
istiʿāra wifāqiyya 2.445
istiʿāra xāṣṣiyya 2.445
istidlāl 2.304
istidlāl bi-l-kaṯra 2.449
istidlāl bi-l-miṯāl 2.449
istifhām 2.184, 2.237, 2.292, 2.388, 2.478, 2.543, 3.549, 4.333, 4.737
istifhām, ʿalāmat al- 3.741

ištiġāl 1.50, 2.149
istiġāṯa 4.329
istiḥsān 4.13
istiˁlāʾ 4.423
iṣṭilāḥ 4.684
iṣṭilāḥī 3.629
ištimal 1.124, 2.86, 2.359, 4.332
istiʾnāf 2.293
istinbāṭ 4.465
ištiqāq 1.453, 1.573–575, 2.165, 2.447–451, 2.496, 2.638, 2.641, 3.15, 3.45, 3.114, 3.539, 4.119–120, 4.446, 4.461, 4.465, 4.640
ištiqāq ʾakbar 1.575
ištiqāq ʾaṣġar 2.447
ištiqāq kabīr 1.575, 4.254, 4.287
ištiqāq qiyāsī 1.452
ištiqāq ṣaġīr 1.573, 1.575
ištirāk 3.320, 3.322, 3.538
istiṣḥāb al-ḥāl 1.194, 4.30
istiṯnāʾ 1.72, 2.451–454, 3.101, 3.549, 4.221
istiṯnāʾ mufarraġ 2.452, 3.104
istiṯnāʾ munqaṭiˁ 2.452
istiṯnāʾ muttaṣil 2.452
istixdām 3.539
istixfāf 2.17, 2.233
Išuruniyya 4.157
Italian 1.340, 1.537, 1.554, 1.562, 2.1, 2.26, 2.285, 2.375, 2.454–459, 2.470, 2.482, 2.491, 2.504, 2.560, 2.659, 2.668–669, 2.699, 3.141, 3.143–144, 3.228, 3.312–313, 3.706, 4.53, 4.216, 4.574
Italian, North 1.45
Italian, Tuscan 3.146
Italian in Egypt 2.457
Italian in Libya 4.556
Italian in Tunisia 4.574
ʾitbāˁ 3.631–632
ʾiṭbāq 1.544, 2.459–461, 3.124, 3.600, 3.608, 4.2, 4.211, 4.421, 4.636
ʾiṯbāt 2.444, 3.118, 3.121
iteration 4.51
iterative 1.201, 1.253, 1.282, 2.245
iterativity 3.616
iˁtibār 3.324
iʾtilāf 4.213
iˁtilāl 2.308
iˁtimād 3.124–125, 4.739
itinerants 2.216
Itô, Junko 2.36, 3.611, 4.85, 4.374
ittaṣala 4.235
ittibāˁ 1.435–436
ittifāq 3.176, 3.320
ittiḥād al-maḥall 3.324
ittisāˁ 2.17, 3.116, 3.125, 3.651
Ivanov, Sergej Nikolaevič 4.585
Iványi, Tamás 1.640–643, 2.536–545, 3.88–95, 3.648
Ivrit 2.461–467, 2.670, 3.409, 3.426, 4.72
Ivrit → Hebrew, Modern
ˁiwaḍ 3.453, 4.444–445
Iwo 4.760
ʾixbār 2.358–360, 2.436, 4.332, 4.737
ʾixfāʾ 2.299–300, 4.426–427
ixtazala 2.16
ixtilāf 3.320
ixtilāṣ 3.92
ixtiṣār 2.17
ixtiṣāṣ 2.295, 3.118
ixtiyār 4.4–7
ʾIxwān Movement → Ikhwan Movement
Iyebi-Mandjek, Olivier 1.334
Iyesiye 1.362
ʾiyyā- 1.263, 1.557, 1.559, 3.591, 4.534–535
iza 1.554
izafet 4.585, 4.598, 4.603, 4.622
ʿIzbit al-Basīlī Arabic 1.271
izdiwāj(iyyat) al-luġa 1.173, 1.533, 1.630, 3.349
İzgi, Cevat 3.503
Izmajlova, M. 4.612
ʿIzzat Afandi, Muṣṭafā → Izzet Efendi
Izzet Efendi 3.338

J

/j/ 1.75, 1.271, 1.608, 2.494–496, 3.10
jāʾa bi- 4.38
Jaʾaliyyin 2.559
Jaʾāme 1.362
Jabal Axḍar Arabic 3.479, 3.482–485, 3.489
Jabal Ramm 3.469
Jabal Ramm, inscription of 1.149
Jabal Šammar 3.326, 4.126
Jabal Says 3.470
Jabal Ṭayy 4.126
Jabal Ṭubayq 3.401
Jabal ʾUmm Jadāyid 3.469
Jabal ʾUsays, inscription of 1.149
Jabal Wašḥah Arabic 4.751
Jabal Yazidi Arabic 4.754, 4.756
Jabali Arabic 2.314
Jabartī, ʿAbd ar-Raḥmān ibn Ḥasan al- 2.332
Jabeur, Mohamed 3.607, 4.563, 4.629, 4.632
Jābir ibn Ḥayyān 2.234, 4.235, 4.287
Jābirī, Majdī al- 1.600
Jackendoff, Ray 2.113, 4.488–489, 4.742
Jacobi, Renate 3.646
Jacobites 4.402
Jacquart, Danielle 4.459
Jaeggli, Osvaldo 3.706
Jaʿfar ibn Ṣābir 2.425
Jaʿfar, Muṣṭafā 4.138
Jaʿfar, Ṭalḥa 2.54
Jaffa 2.464, 2.506
Jaffa Arabic 1.608, 2.502
Jafr, al- 3.401
Jāhān, Shah 3.338
Jahāngīr, Shah 3.338
Jahani, Carina 2.409
jāhiliyya 4.76, 4.78
Jāḥiẓ, al- 1.403, 1.428, 1.434, 2.96, 2.303–304, 2.442, 2.469, 2.579, 2.630, 2.633, 2.676–677, 3.71–72, 3.117, 3.160–161, 3.690, 4.64, 4.74, 4.202, 4.473
jahr 3.124
Jahriyya 1.380
Jahrū, ʾAsmahān al- 2.258
Jaimoukha, Amjad 2.498, 2.500
jāʾiz 2.184, 2.299
Jakarta 2.335
Jakobi, Angelika 3.435–439
Jakobson, Roman 2.98, 2.319, 2.560, 2.687, 3.600, 4.199, 4.328, 4.422
Jalabī [Celebi], Ḥasan 4.561
Jalāl, Muḥammad ʿUṯmān 1.600
jalīl 1.151, 3.308
jaʿliyya 3.225
jamʿ 1.309, 1.576, 3.426, 3.439
jamʿ al-jamʿ 2.352
jamʿ al-kaṯra 3.439
jamʿ al-qilla 3.439
jamʿ maksūr 3.439
jamʿ mukassar 3.439
jamʿ sālim 3.439
Jamāʿa al-ʾIslāmiyya, al- 4.77–78
Jamaican Creole English 2.319–320
Jamālī, Badr al- 2.46
James, David Lewis 3.112, 3.308–309, 3.340, 3.519
jamīʿ 2.445, 2.537, 4.17
Jāmī, ʿAbd ar-Raḥmān al- 1.380

jāmid 1.573, 2.94, 4.686
jāmid, ism 2.445
Jamīl ibn Muʿammar 4.202
Jamīlī, Xuḍayr ʿA. al- 1.619
Jamoussi, Souad 4.575
Janaba 2.213
Janjero 1.35
janna 1.174
Jansen, Johannes J.G. 4.78
Jansen, Willy 4.432
Janssens, Gerard 1.618, 3.726, 4.349
Japanese 2.9, 3.232, 3.706, 3.725, 4.537, 4.539
Japanese, object-raising in 4.539
jarāʾid 4.217
Jarājrah Arabic 4.238
Jarema, Gonia 1.425
jargon 2.8, 2.468–472, 3.59, 3.639, 3.642, 4.252, 4.255, 4.380
jargon, Cairene 2.469
jargon, criminal 2.470
jargon, Egyptian 2.469–470
jargon, Hebrew 2.470
jargon, Islamic 2.471
jargon, Jewish 2.471
jargon, North African 2.470
jargon, secret 3.127
jargon of boatbuilders 2.215
jargon of homosexuals 2.470
jargon of students 4.380
Jarīm, ʿAlī al- 2.578, 2.581
Jarmī, ʾAbū ʿUmar al- 1.193, 3.690, 4.119
jarr 1.67, 1.309, 1.347, 1.555–556, 1.559–560, 2.232–233, 2.294, 2.402, 2.425, 2.427, 3.699
Jarrāḥi 2.572
Jarrār, Samīr A. 2.104–105, 2.110
Jarrār, Walid Sadeq 2.426–427
jars 4.2
Jašar-Nasteva, Olivera 4.259
Jaspaert, Koen 2.681
Jaṣṣāṣ, ʾAbū Bakr ʾAḥmad ibn ʿAlī al- 3.122
Jastrow, Otto 1.87–96, 1.270, 1.388, 1.444–445, 1.536, 1.540, 1.562, 1.571, 1.584, 1.607, 1.609–610, 1.612, 1.615, 1.642, 2.230–231, 2.272–273, 2.313, 2.388, 2.414–424, 2.533–534, 3.190, 3.237, 3.568, 3.585, 3.609, 4.2, 4.109, 4.302–303, 4.318, 4.404, 4.407, 4.520, 4.612, 4.621–622, 4.664, 4.750, 4.754
Jat 2.217
Jatlāwī, al-Hādī al- 2.70
Jaussen, Antonin 4.479, 4.481–482
Java 2.340, 2.472
javanais 2.468, 4.156
Javanese 2.334, 2.336, 2.339, 2.472–477, 2.668, 2.673, 4.476–477
jawāb 1.72, 1.202, 2.477, 2.479
jawāb al-jazāʾ 2.478
jawāb ʾin 2.478
jawāb law 2.478
Jawād, Muṣṭafā 1.453
Jawālīqī, ʾAbū Manṣūr Mawhūb ibn ʾAḥmad al- 1.627, 3.36
jawāzim → *jāzim*
Jawbarī, ʿAbd ar-Raḥīm ibn ʿUmar al- 2.469
Jawf Arabic 4.751
Jawf, al- 4.123
jawhar 1.574, 2.186
Jawhar ibn Ḥayḍar 2.53
Jawharī, ʾAbū Naṣr ʾIsmāʿīl ibn Ḥammād al- 1.4, 3.21, 3.33, 3.42, 3.44, 3.214, 3.502
Jawi 1.135, 4.477
Jawnpūrī, Mullā Maḥmūd 2.327
Jaww, al- 3.401
Jawwās ibn Huraym 3.647
Jawzjānī, al- 3.379
Jayakar, Atmaram 2.211, 3.478
Jaybī, al-Fāḍil al- 1.603
jayyid 1.192
Jayyusi, Salma Khadra 4.92
jazāʾ 1.49, 1.72, 2.477–481, 2.542, 3.261
Jazāʾirī, Ṭāhir al- 2.637
Jazayery, Mohammad Ali 2.411–412, 3.579
Jazeera, al- → Al-Jazeera
jāzim 1.71
Jazīra, al- → Al-Jazeera
Jazirah 3.326, 4.125
Jazīrat al-Qamar 3.125

jazm 1.67, 1.71–72, 1.152, 1.309, 2.16, 2.233, 2.402, 2.404, 2.478–479, 4.132, 4.557
Jazūlī 2.52
Jbāla Arabic 1.439, 1.590, 1.609, 3.288–290, 3.292–293, 4.70
Jbāliyyah 3.401
Jbāliyyah Arabic 3.401, 3.403, 4.240
Jebala 3.288
Jebayr 1.363
Jebel South 1.655
Jeddah 3.179, 4.628, 4.767
Jeddah, youth speech in 4.767
Jeddah Arabic 1.271, 4.126, 4.632
Jefferson, Gail 2.202
Jeffery, Arthur 2.57, 2.198–199, 2.373, 2.666, 3.7, 4.295
Jelinek, M. Eloise 1.220, 3.385–386, 3.684, 3.709, 4.395, 4.399–400, 4.648–650, 4.655, 4.657
Jenābī 3.502–503
jenggot 2.336, 2.339
Jenin Arabic 2.321
Jenkins, Major 2.518
Jenssen, Herbjøn 2.360
Jephson, Arthur Jeremy Mounteney 1.519
Jerash 2.498–499, 2.502
Jerba 2.709, 4.572, 4.574
Jeremiás, Éva M. 2.406–414, 3.114
Jernudd, Björn 1.668
Jerome, Saint 1.167
Jerusalem 1.23, 2.439, 2.481, 2.505, 3.409
Jerusalem Arabic 1.323, 1.444–445, 1.608, 2.32, 2.221, 2.231, 2.481–493, 2.502, 2.509, 2.532, 2.534, 3.8–9, 3.11, 3.658, 3.704, 4.632
Jerusalem Arabic, Jewish 2.491
Jespersen, Otto 1.616, 1.619, 2.158, 2.331–332, 3.653, 3.655
Jesry, Maher 4.387–389
Jesuit College de Louis-le-Grand 1.171
Jesus 3.730
Jewish Arabic 1.54, 1.323, 1.443, 1.445–447, 2.23, 2.266, 2.277, 2.313, 2.470, 2.494, 2.526–536, 2.659, 2.661, 3.218, 3.220, 3.451, 3.472, 4.58, 4.70, 4.157, 4.216, 4.217
Jewish Arabic, Baghdadi 1.351, 1.537, 1.607, 1.610, 1.625, 2.414, 2.276, 2.313, 2.532–534, 4.664
Jewish Arabic, Cairene 2.532–533
Jewish Arabic, Egyptian 2.276, 2.278–279
Jewish Arabic, Ḥijāzī 2.259
Jewish Arabic, Jerusalem 2.491
Jewish Arabic, Moroccan 1.444–445, 2.533, 3.288, 3.292–293
Jewish Arabic, Tlemcen 2.533
Jewish Arabic, Tripoli 4.548
Jewish Arabic, Tunis 1.609, 2.277, 3.449, 4.563
Jewish Arabic, Tunisian 4.575
Jewish Arabic, Yemeni 2.534, 4.750
Jewish communities 2.531
Jews 1.99, 1.444–448, 1.527, 1.589, 2.60–61, 2.526, 3.218, 3.370
Jews, Egyptian 2.534
Jews, Iraqi 2.414–415, 2.526, 2.532
Jews, Karaite 2.471
Jews, Libyan 2.471
Jews, Moroccan 2.63–64, 2.470
Jews, Portuguese 2.527
Jews, Rabbanite 2.533
Jews, Sephardic 2.64, 2.437–438, 2.461
Jews, Spanish 2.527, 2.532
Jews, Tunisian 2.532
Jews in Malta 3.142
Jews in Morocco 3.294, 3.314
Jews in Sicily 4.216
Jews of Bukhara 4.424
Jews of Cairo 1.323
Jews of Cochin 3.129
Jews of Jerusalem 2.481
Jews of Samarkand 4.424
Jezreel, Plain of 3.402
Jibbāli 1.493, 2.66–67, 2.211, 3.478, 3.593, 4.297–303, 4.312
Jiblah 1.187, 2.273
Jibran (Jubran), Jibran Khalil 2.703, 3.395, 4.91, 4.211

Jidda Arabic → Jeddah Arabic
Jiha, Michel 4.70
jihād 4.78
Jihadist Movements 3.136, 3.170
Jilālī, Ḥulām al- 3.38, 3.43
Jilānī, Sheikh ʿAbd al-Qādir al- 4.277
Jilyānī al-ʾAndalusī, al- 4.204
jīm 1.26–27, 2.494–496, 3.609
Jimma 2.51, 2.53
Jin, Zhou 4.268
jinās 2.496–498, 3.538–539, 4.211
jinās al-ištiqāq 2.496, 3.539
jinās al-qalb 3.539
jinās al-xaṭṭ 2.497
jinās at-tarkīb 3.539
jinās lafḏī 3.539
jinās maʿnawī 3.539
jinās maqlūb 2.497
jinās muḍāriʿ 2.497
jinās muḥarraf 2.497
jinās nāqiṣ 2.497
jinās tāmm 2.496–497
jinās taṣḥīfī 4.157
Jindi, Ahmad Alam ad-Din al- → Jundi, Ahmad Alam ad-Din
jingtang yu 1.380
jins 2.234, 4.545
Jīwan, Mulla 2.327
jiwār 1.72, 2.403–404
jizya 3.346
JMODEL 1.208–209
Job, Book of 1.167
Johansen, Baber 2.360
Johanson, Lars 4.587
John of Segovia 1.168
John of Seville 1.166
Johns, Jeremy 4.217
Johnson, C. Douglas 4.347
Johnson, Frederick 1.663
Johnson, Mark 1.423, 1.425, 2.689, 4.165
Johnson, Martin Elmer 1.595
Johnson, Rev. Samuel 4.758
Johnstone [Koch], Barbara 1.473–474, 1.648, 3.500, 3.669, 3.678, 4.49
Johnstone, Penelope 2.677
Johnstone, Thomas Muir 1.446, 1.584–585, 1.611, 1.639, 2.151–152, 2.211, 2.609–610, 3.622, 4.125, 4.297, 4.302–306, 4.403, 4.753
jokes 2.653, 2.658
Joly, André 2.82
Jomier, Jacques 3.584, 3.586–587
Jones, Russell 2.334, 2.341–342, 2.476
Jones, William 1.167, 3.207, 4.516
Jong, Rudolf E. de 1.270, 1.320–322, 1.570–571, 1.605, 2.2, 2.151–153, 2.509, 3.360, 3.401, 3.403–406, 4.237–251, 4.628, 4.677
Jonge, Huub de 2.337
Joola 4.723
Joos 1.632
Jordan 1.26, 1.260, 2.105, 2.498–505, 2.658, 3.316
Jordan, Arabic in 1.260, 2.31, 2.494, 2.555, 3.525, 3.613, 3.726
Jordan valley 3.402
Jordanian Arabic 1.83, 1.217, 1.656–659, 2.71–72, 2.97–99, 2.120, 2.230–231, 2.395, 2.397–398, 2.498–517, 2.585–587, 3.11, 3.263–268, 3.527, 3.604, 3.611–612, 3.678–679, 3.686, 4.407, 4.489, 4.629, 4.632, 4.648, 4.654–656, 4.664, 4.672, 4.681
Joseph, Brian 4.371
Joshi, Aravind 1.457
Joufra 3.56
Joüon, Paul 1.576–577, 3.409
Jourdan, J. 4.563
journalism 2.636, 2.655, 3.19, 3.39, 3.193
Journalism, Folkloristic 3.200
journals, Arabic 1.259
Joyce, James 4.36
Jozami, Gladys 3.3
Juan [Hispalense] 1.96
Juba 2.517, 3.10
Juba Arabic 1.37, 1.489, 1.520, 1.525, 2.517–525, 3.10, 3.62, 3.635, 3.639, 3.642–643, 4.198, 4.376
Jubb ʿAdīn 1.178
Jubbāʾī, ʾAbū Hāšim al- 3.162–163
Jubran, Jubran Khalil → Jibran, Jibran Khalil

Juburī, ʿAbdallāh al- 2.638–639
Judaean Desert 3.360
Judaeo-Arabic → Jewish Arabic
Judaeo-Aramaic 1.240
Judaeo-Spanish 4.579
Jufra 3.53
Juhayna 1.361–362
Juhayna Arabic 3.94
Juhaynī, ʿAbdullāh al- 1.361
Julien de Pommerol, Patrice 1.360–368, 3.61, 3.634, 4.708–710, 4.713, 4.716
Jumʿa 4.104–105
Jumʿa, ʾIbrāhīm 3.349
Jumaḥī, ʾAbū Dihbil al- 3.89
Jumbe, Alhajj Aboud 1.661–662
jumla 1.49, 1.558, 2.186, 2.292, 2.358, 2.536–540, 2.544, 4.738
jumla ḏ̣arfiyya 1.49, 1.567, 4.318, 4.653, 4.739
jumla ḏāt wajhayni 4.504
jumla fiʿliyya 1.182, 2.430, 3.114, 3.177, 4.497, 4.653, 4.737–738
jumla ḥāliyya 2.224
jumla ibtidāʾiyya 2.292–293
jumla ʾinšāʾiyya 2.358, 4.29
jumla ismiyya 1.182, 1.202, 2.430, 3.114, 3.177, 3.354, 3.381, 3.683, 3.687, 4.56, 4.495, 4.499, 4.653, 4.737–738
jumla ʾixbāriyya 2.358
jumla kubrā 2.539
jumla marwiyya 2.331
jumla ṣuġrā 2.539
jumla, tamyīz al- 4.437
jumlat al-istiṯnāʾ 2.451
Jun, Jongho 2.365
Junaydī, Bilāl 3.539
Junaydī, Muḥammad 4.429
junction 1.430, 1.432
junction, morphophonological 4.236
juncture, clausal 4.197
juncture, core 4.197
juncture, morpheme 4.700
juncture, nuclear 4.197
Jundī, ʾAḥmad ʿAlam ad-Dīn al- 1.641, 2.255–256, 3.89, 3.91–92, 3.94
Jundī, ʿAlī al- 2.496–497
Jungraithmayr, Hermann 2.68
Junker, Wilhelm 1.519
Junus, Mahmad 2.336
Junusi, Zainuddin Labai el- 2.337
Jurafsky, Daniel 1.456, 511
Jurjānī, ʿAbd al-ʿAzīz al-Qāḍī al- 1.428, 2.582, 3.118, 4.209
Jurjānī, ʾAbū Bakr ʿAbd al-Qāhir ibn ʿAbd ar-Raḥmān al- 1.67, 1.424–425, 1.428–429, 1.471, 2.17, 2.79–80, 2.83, 2.95, 2.131–135, 2.181, 2.188, 2.290–291, 2.295, 2.303, 2.336, 2.441–444, 2.550, 2.580, 3.105, 3.116–118, 3.161, 3.163, 3.539, 4.21, 4.166, 4.203, 4.207, 4.213, 4.472–473, 4.737
Jurjānī, ʿAlī ibn ʿAbd al-ʿAzīz al- 2.443
Jurjānī, ʿAlī ibn Muḥammad al- 4.686
jussive 1.203, 1.219, 2.23, 2.92, 3.248–249, 3.269, 3.272, 3.299, 3.355, 4.557, 4.639
jussive, Hebrew 4.178
Justi, Ferdinand 3.717
Justice, David 4.167
Juwaynī, ʾAbū l-Maʿālī ʿAbd al-Malik ibn ʿAbdallāh al- 2.304
juxtaposition 1.473, 3.672
juzʾ 3.208–209
juzʾī 3.178
Juzur al-Qamar 1.448

K

ka- 1.297, 1.557–558, 1.609, 2.22, 3.281, 3.285, 3.292, 3.702
Kaado 4.281
ka-ʾanna 2.354–355, 2.357
Kaapmaleiers 4.290
ka-ʾayyin 1.73
Kaʿb 2.571
Kaba, Lansine 3.136, 3.138, 3.140
Kabha, Riad 4.663
Kaʿbīye 1.92
Kabrah, Rawiah S. 3.179–180
Kabyli 1.290
Kabylia 1.53–54
Kabylian Berber 1.38, 1.56, 1.122, 1.289, 1.370
Kabylians 2.704
Kachru, B. Braj 2.30–31
kāda 1.73, 3.105, 3.354
Kadmon, Nirit 3.653
Kaempfert, Manfred 4.72, 4.76
kāf al-mabsūṭa, al- 4.560
Kafa 1.35
Kafawī, al- 2.358–359
kaff 2.16
Kafiri 1.29
Kafr išŠēx Arabic 2.2
Kafrāwī, ʾAsʿad ʿAbd al-Ġanī as-Sayyid al- 3.322
Kager, René 3.493, 3.496, 4.344–353
Kahane, Henry 2.456
Kahane, Renée 2.456
kāhin 4.104, 4.203
kahir 4.105
Kahle, Paul 1.598, 1.615, 3.431, 3.654–655, 3.684, 3.686, 3.689, 4.317, 4.656, 4.662
Kähler, Hans 4.291, 4.293–294
Kahn, David 4.159
Kahn, Margaret 4.289
kāʾin 3.96
Kairouan 1.361, 2.532, 3.53, 4.572
Kairouan Arabic 1.609, 4.563, 4.670
Kakwa 1.519
Kala 4.715
kalām 1.380, 1.429, 2.83–85, 2.87, 2.184, 2.186, 2.304, 2.358, 2.401–404, 2.536–539, 2.541–545, 2.623, 2.627, 3.73, 3.91, 3.115–116, 3.550, 3.648, 4.11, 4.13, 4.332, 4.474
kalām al-ʿArab 2.85–86, 2.176–179, 2.181, 2.263, 2.265, 2.543, 2.628
kalām Allāh 2.183, 2.541, 4.6
kalām an-nafs 2.306
kalām as-sabab 4.252
kalām ʾawwal 2.478
kalām nafsī 2.306
kalām, ʾaqsām al- 4.329
kalamo 1.544
kalamon'Antesitesy 3.127
Kalb 1.545
Kalb Arabic 3.94
Kalbī, Muḥammad ibn as-Sāʾib al- 1.309, 2.401, 3.89–90, 3.116, 4.235
Kaleši, Hasan 3.502
Kalfou 1.337

kalim 2.624
kalim wujūdiyya 3.115
kalima 2.424, 2.544–548, 3.547, 3.550
Kalimantan 3.125
kallā 3.550
Kallas, Elie 1.603, 3.343–353
Kalus, Ludvik 2.42
kam 1.73, 3.90
kam al-xabariyya 4.430
kamā 3.200
Kamal, Ribhi 1.626
Kamal al-Din, Adeeb 4.204
Kamberli 1.114, 1.117
Kamel, Boshra 3.681
Kamel, Mustafa A. 3.680
Kamel, Salwa 1.425
Kamel, Sheikh Saleh 3.197
kāmil 3.211–212, 3.214
Kamil, Murad 3.437
Kāmil, Muṣṭafā 3.193
Kammensjö, Heléne 1.470–477, 4.192, 4.501
Kampala 1.520
Kampffmeyer, Georg 3.287
kān 1.64, 1.281, 1.553–554, 2.271, 2.273, 2.595, 3.81, 3.85
kāna 1.216–220, 1.432, 1.486, 1.488, 1.506, 1.558, 1.563, 2.93, 2.180, 2.194, 2.290–291, 2.355–356, 2.431, 2.537, 2.548–552, 2.578, 2.582, 3.68, 3.99, 3.101, 3.105, 3.115, 3.234–236, 3.260, 3.353–354, 3.386, 3.687, 3.691, 3.695, 4.163, 4.395, 4.412, 4.414, 4.740
kāna an-nāqiṣa 2.548–551
kāna at-tāmma 2.548–549, 2.551, 4.163
kāna az-zāʾida 2.548, 2.551
kāna waʾ-axawātuhā 1.70, 2.548–552
kāna with nominative predicate 3.691, 3.695
kāna yafʿalu 1.218
Kanafani-Zahar, Aida 2.363–364
Kane, Cheikh Hamidou 4.184
Kanem 1.360–363, 2.552, 4.708
Kanem Bornu Empire 3.373
Kanembou 1.366
Kankan Musa 3.135
Kannur 3.129
Kano 1.337, 2.250, 3.375–376
Kanun, Abdallah 3.201
Kanuri 1.133, 1.333, 2.137, 2.251–252, 2.552–555, 3.314, 4.279, 4.282, 4.284, 4.715–716
Kanzi 2.1
Kanzleisprache 1.490
Kaolack 4.182
Kapeliuk, Olga 1.490
Kaplan, Edith 2.690
Kaplan, Mary 2.646, 2.649
Kaplan, Robert 1.648
Kapliwatzky, Yohanan 2.481
Kaptein, Nico 2.337
Karabacek, Josef von 1.151
Karabekyan, Samvel 2.236–239
Karachay-Balkhar 2.557
Karači 2.220
Karaduvar Arabic 1.395
Karahan, Abdülkadir 3.502
Karahisari, Ahmad → Qarahḥisārī, Ahmad
Karaites 2.528, 2.533
Karak 2.499
Karak Arabic 2.498–499, 2.503, 2.506–507
Karakalpak 2.557
Karassery, M.N. 3.128–135
Karbala 2.407
Karibʾil Watar 1.127
Karimov, Akbar U. 4.305
Karmī, Ḥassān Saʿīd al- 3.41
Karmiloff-Smith, Annette 2.689
Karmysheva, Balkyz Kh. 4.612
Karnataka 2.325
Karouri, Abdel Muneim al- 1.457
Kars 2.607
Karshuni 1.495–497, 1.603, 3.218
Kartvelian languages 2.67
Karume, Sheikh Abeid Aman 1.661
Kārūn 2.571–572
Karxī, Mollā ʿAbbūd al- 3.348
Kasab 4.402
Kāsānī, ʾAbū Bakr ibn Masʿūd al- 2.359
Kāšġarī, al- 3.22
Kashan 2.46
Kashmiri 1.133–134, 2.220
Kasimin, Amran 2.341

kaskasa/kaškaša 1.26, 1.571, 1.614, 2.213, 2.555–557, 3.93, 3.478, 3.525, 3.587, 3.606, 4.36, 4.126–127, 4.304, 4.628, 4.754, 4.757
Kašmīrī, Muḥammad Ḥusayn 3.338
kasr 1.152, 1.309, 2.232–234, 2.311, 2.402
kasra 1.560
Kassala 4.376
Kassian 4.76
Kassim Mohamed, Souad 1.655
Kästner, Hartmut 2.229–230, 3.594, 4.1
katēgoroúmenon 4.220
kātib 1.398
kaṯrat al-istiʿmāl 2.86
Katz, Yehuda 3.360, 3.400
Kaufman, Dorit 2.684
Kaufman, Elana 2.439
Kaufman, Terrence 1.294, 1.501, 1.504, 2.29, 2.666, 2.668, 4.368, 4.370–373
Kaun, Abigail 2.365
Kaur, Kuldip 2.326
Kavalcık 1.115, 1.117
Kawākibī, ʿAbd ar-Raḥmān al- 3.346
Kawar 1.361, 2.552
Kawāwila 2.576
kay 1.467
Kay, Martin 1.457, 1.511
Kay, Paul 4.58
Kaye, Alan S. 1.133–136, 1.615–616, 1.631, 2.25–29, 2.230, 2.495, 2.518, 2.560, 2.655, 3.141–142, 3.146, 3.427, 3.451, 3.634, 3.639–641, 3.643, 3.726, 4.2, 4.302–303, 4.305–306, 4.678, 4.681, 4.716
kayfa 1.72, 2.427
Kayne, Richard 1.412, 3.230, 3.232–233, 4.746
Kayne's Generalization 1.412
Kayyal, Ziad N. 1.668, 1.670
Kazakh 2.557–559, 4.263
Kazakhstan 2.557
Kazan 4.425, 4.450
Kazanlı Arabic 1.392
Kazimirski, Alfred de Biberstein 3.23, 4.96–97
Kazziha, Walid 2.498
k-dialecs → *k*-perfect
Kearns, Kate 1.40–42
Kedourie, Elie 2.59
Keegan, John M. 4.351
Keeling, Captain 3.128
Keenan, Edward 4.15
Kehayia, Eva 1.425
Kehr, Georg Jacob 2.42
Kendon, Adam 2.202
Kənfāwī, ʿAbd əṣ-Ṣamad əl- 3.295
Kenner, Charmian 1.371
Kenny, Keith Dallas 2.684–685
Kenstowicz, Michael 2.173, 2.396, 3.207, 3.495–496, 3.610, 3.613–614, 3.619, 3.621–622, 3.706–707, 3.709, 3.726, 4.64, 4.344–345, 4.347–351, 4.390, 4.400, 4.675–676
Kent, Raymon D. 4.666
Kenya 1.518, 1.520, 1.661–663, 2.588–589, 3.642, 4.386
Kenya Coastal Strip Protectorate 1.660
Kenya Muslim Association 1.662
Kenzi 3.437–438
Kenzi-Dongolawi 2.668, 3.313, 3.435
Kepel, Gilles 4.77–78
Kerak 2.506
Kerala 2.325, 3.128–129, 4.123
Kerawān 1.361
Kermanšahı 3.351
Kermanshah 2.604
Kerslake, Celia 4.584, 4.586
Keskinci 1.116
Kettani, M. Ali 1.661–662
Kfār ʿAbīda Arabic 1.540, 3.451
Khabur 4.402
Khabura 1.24
Khabura Arabic 3.478
Khairallah, Shereen 2.45
Khal, Yusuf al- 4.214
Khalafallah, Abdelghany A. 2.2, 4.349
Khalaily, Samir 3.654, 4.648
Khalil Arabic 3.527
Khalil, Aziz 3.248, 3.573
Khalil, Esam N. 1.430–433, 1.648, 2.207–210
Khalil, Muhamed al- 3.500
Khamis, Cornelia 1.520–522

Khan, Geoffrey 1.151, 2.415, 2.471, 2.526–536, 3.370, 3.419, 3.518, 3.703–705, 4.317, 4.319, 4.656
Khan, Gulfishan 3.508
Khan, Sher 3.508
Khānaqīn 2.533
Khānaqīn Arabic 2.415
Kharga Arabic 1.258, 2.2, 2.5–6, 2.8, 2.389, 4.196
Khargird 2.600
Kharijites 3.170
Kharma, Nayef 2.32
Kharraki, Abdennour 1.650
Kharrat, Edouard 2.332, 4.288
Kharrat, Muhammad 2.397, 3.725
Khartoum 1.23–25, 1.525, 2.518, 2.559–571, 4.376, 4.381
Khartoum Arabic 1.525, 2.271, 2.559–571, 3.10, 3.62, 3.643, 4.376, 4.495–497, 4.499, 4.529–535
Khashm el-Girba 3.435
Khatabi, Mohammed [Muḥammad Xaṭābī] 4.472
Khater, Mahmoud Rushdi [Maḥmūd Rušdī Xāṭir] 2.105–107, 2.109
Khattab, Ghada 1.368–375
Khatuniya Arabic 2.414
Khawlāni 1.528
Khaybar → Xaybar
Khazendar, Walid 4.213
Khirbet es-Samra 4.478
Khmer 4.476
Khogali, Hišām 1.436, 1.438
Khoisan 4.291
Khokha 2.270–271, 2.273
Khomeini, Ayatollah 2.407, 3.370
Khōr Mūsa 2.572
Khōr Rōrī 2.212
Khorasan → Khurasan
Khorramshahr 2.572
Khouri, Rašīd Salīm al- 3.4
Khouzam, Joseph 3.584, 3.586–587
Khurasan 1.28, 2.406, 2.408
Khurasan Arabic 1.583, 3.544
Khuzestan 1.260, 2.407, 2.571–578, 3.370
Khuzestan Arabic 1.584, 2.414, 2.571–578, 3.84
Kibera 1.520
kiḏb 2.362
Kieffer, Charles 2.61, 4.612, 4.620, 4.622
Kienast, Burkhart 1.644, 3.453
Kiesler, Reinhard 2.281–286
Kihm, Alain 1.12–16
kilā 4.17–18
Kilani-Schoch, Marianne 3.207
Kildān 4.402
Kılıçtutan 1.111
Killean, Carolyn G. 1.572, 4.524–526
Killion, Tom 1.654–655
Kināna 3.94, 3.692–693
kināya 1.555, 2.305, 2.445, 2. 578–583, 3.116–121, 4.6, 4.157, 4.165
Kinberg, Naphtali 1.476, 1.545, 2.178, 2.227, 2.294, 2.298, 2.301, 2.436, 2.477–478, 2.536, 3.107, 3.109, 3.549, 4.219–220, 4.235, 4.429, 4.431
Kinda 1.129, 3.467, 4.295
kindergarten 3.79
Kinderib 1.88–93, 1.95
Kindī, ʾAbū Yūsuf Yaʿqūb ibn ʾIsḥāq al- 3.336, 4.542, 4.546
King, D.A. 2.201
King, Geraldine 4.480
kinship 3.674
kinship terminology 2.583–588
kinship terms 1.83, 1.116, 1.227, 1.283, 1.292, 1.327, 1.540, 1.550, 2.222, 2.272, 2.289, 2.594, 3.50, 3.366, 4.469
kinship terms, affectionate 2.587
kinship terms, Bengali 1.286
kinship terms, distant 2.587
kinship terms, Malayalam 3.134
kinship terms, reversed 3.366
Ki-Nubi 1.37, 1.86, 1.357, 1.520, 2.518, 2.525, 2.588–597, 3.62, 3.635, 3.642–643
Ki-Nubi, Arua 1.524
Ki-Nubi, Bombo 1.522, 1.524
Ki-Nubi, East African 1.520–521
Ki-Nubi, Gulu 1.524
Ki-Nubi, Kenyan 1.521, 1.524, 2.589

Ki-Nubi, Ugandan 1.521, 1.524, 2.589
Kiparsky, Paul 2.36–38, 2.396, 3.494–496, 3.604, 4.344–345, 4.348
Kipchak 2.557, 4.263–264, 4.589
Kipp, Sandra 3.313
Kiraz, George Anton 1.457, 1.484, 4.94, 4.388
Kirchner, Mark 4.583–589
Kirchner, Robert 2.365
Kirghiz 2.557, 4.263
Kirkuk 2.533, 2.604, 2.608
Kirkuk Arabic 2.415, 2.417–418, 2.421, 2.532
kirš 3.112
Kisāʾī, ʾAbū l-Ḥasan ʿAli ibn Ḥamza al- 1.70, 1.402, 2.79, 2.355, 2.629, 2.631, 3.35, 3.690, 4.5, 4.267
Kisseberth, Charles 4.342
Kister, Menachem J. 3.518
kitāb 1.131
kitab kuning 2.338
kitāba 1.338
Kitahara, Hisatsugu 2.369, 4.492
Kittredge, Richard 4.58
Kjamilev, Said H. 4.306
klām əl-Bīḏ̣ān 2.240
Klat, Hector 3.349
Klatt, Dennis H. 2.229, 2.365
Kleiber, Georges 1.568
Klein, Ewan 1.458
Klemz 4.322
Klich, Ignacio 3.2
Klima, Edward 3.653
Klingenheben, August 2.139
Kluge, Eva-Maria 2.224–225
Knauf, Ernst A. 3.472, 4.171, 4.477–483
Kniffka, Hannes 1.484
Knudtzon, J.A. 3.409
Knutsson, Bengt 1.384
Koalib 4.375
Kodungallur 3.129
Koelle, Sigismund W. 4.709
Köfler, Hans 1.626–627, 3.93, 4.36, 4.446
Kogan, Leonid E. 2.67, 4.295, 4.302–303, 4.305–307, 4.310
koine 1.364, 1.490–491, 1.593–597, 1.617, 1.654, 1.669, 3.59, 3.188, 4.716, 4.757
koine, Arabic 1.594, 1.667
koine, Cairene 1.570
koine, Chadian Arabic 3.61
koine, commercial 1.617
koine, dialect 1.490–491, 1.593–597
koine, Greek 1.593
koine, Gulf 2.609
koine, inter-Arabic 1.668
koine, interdialectal 4.571
koine, international 1.668
koine, Jordanian 3.405
koine, literary 1.593, 2.529, 3.59
koine, local 2.609
koine, military 1.593–594, 1.618, 2.267, 3.640, 3.647
koine, Moroccan 3.273, 3.274
koine, national dialect 1.596
koine, Palestinian 3.360
koine, poetic 1.153, 1.400, 1.593–594, 1.614–618, 2.264, 2.353, 2.373, 3.59, 3.465, 3.499, 4.9, 4.201
koine, poetico-Qurʾānic 1.615, 1.617, 3.644–645
koine, post-Islamic 1.618
koine, regional 2.214, 2.531
koine, Sara 1.363
koine, spoken 1.618
koine, Syrian 1.546
koine, urban 1.59, 593–594, 1.596, 2.508
koineization 1.259, 1.406–407, 1.445, 1.483, 1.489, 1.572, 1.593–596, 1.632, 1.667, 1.669, 3.9–11, 3.360, 3.403–404, 3.639, 3.645
koineized colloquial 4.323
Koizumi, Masatoshi 4.741, 4.743
Kokan, Muhammad Yousuf 2.328
Kōken, Yusuf M. 3.506
Kom Ombo 2.1, 3.435
König, Ekkehard 2.193, 2.196, 4.161
König, Wolf Dietrich 4.577
Kontzi, Reinhold 1.585, 3.141–142
Konya 4.584

Koopman, Hilda 3.685, 4.524, 4.648
Kopf, Lothar 2.86, 3.36, 4.295
Koplewitz, Immanuel 2.438, 2.465
Koppe, Reiner 2.407, 3.575, 3.579
Korandje 4.279
Korba Arabic 4.681
Kordofan 1.519, 3.435, 4.375, 4.380, 4.709
Kordofan Arabic 4.376
Korean 2.400
Kormakiti 1.536
Kormakiti Arabic 3.450
Korotajev, Andrej 4.302, 4.305
Kortobi, Ibtissam 2.21, 4.646, 4.648–650, 4.655
Korucam 1.536
Kōsa 1.87, 1.92, 1.95
Koschmieder, Erwin 3.98
Kossmann, Maarten 1.289, 2.68
Kosti 2.559
kōteubah 1.8
Kotoko Sultanates 1.334
Kotokos 1.334
Kotzé, Ernst F. 4.294
Koufra 3.56
Kouloughli [family name] 4.573
Kouloughli, Djamel Eddine 1.580, 2.165, 2.168, 2.178–179, 2.181, 2.345–354, 2.448, 2.536, 2.538, 2.544, 2.562, 2.623–628, 3.159–164, 3.453, 3.617, 3.726, 4.444, 4.670, 4.484–486
Kousseri 1.334, 1.336–337
Kouwenberg, N.J.C. 3.418
Kovalam 2.329
Koya, Mohammad 2.325
Koyunoğlu 1.114
Kozhikode → Calicut
Kozlowska, J. 3.23
Kozluk 1.87–88
Kozluk Arabic 4.70
Kozluk-Sason Arabic 1.607
Kozluk-Sason-Muş 1.94–95
k-perfect 1.607, 1.314, 2.259, 3.694, 4.177, 4.305, 4.314, 4.751, 4.754, 4.756
Kraemer, Jorg 3.36
Kraemer, Roberta 2.661
Krahl, Günther 3.27, 3.453, 4.151
Kramers, Johannes Hendrik 3.720
Krapf, J. Ludwig 1.663
Krapyak 2.336
Krashen, Stephen 4.140, 4.145
Kratzer, Angelika 3.685, 4.648
Kraus, Paul 2.234, 4.235, 4.287
Krauss, Robert M. 2.653
Kreemer, J. 1.9–10
Kreidler, Charles 4.557
Kreindler, Isabelle T. 2.557
Kremer, Alfred von 2.218, 2.220
Kremers, Joost Merijn 3.228–233, 3.428, 3.432, 4.43–45, 4.53–57, 4.648, 4.650, 4.657
Kress, Günther 2.384
Kretzmer, David 2.438
Krier, Fernande 3.142
Krifka, Manfred 3.653
Krio Creole English 4.198
Krishnamurti, Bh. 3.128
Kristensen, Kjeld 2.320
Kroon, Sjaak 2.681
Kropfitsch, Lorenz 1.490, 1.492, 3.17, 3.27
Kropp, Manfred 2.51–52, 2.57, 3.469, 3.471
Kroskrity, Paul V. 2.651
Krotkoff, Georg 2.629, 3.27, 3.125
Kruk, Remke 4.543–546
Krumm, Bernard 1.660, 1.663
Kubrawiyya 1.380
Kufa 1.618, 2.597
Kufa Arabic 2.312
Kufan school 1.49, 1.67–74, 1.309, 1.559, 2.185, 2.232, 2.291, 2.298, 2.301, 2.355, 2.427, 2.431, 2.453, 3.91, 3.97–98, 3.101, 3.105, 3.107–108, 3.116, 3.165, 3.424, 3.543, 3.548–550, 4.8–9, 4.12, 4.120, 4.219–220, 4.235, 4.431, 4.438–439, 4.447, 4.455, 4.558, 4.739
kufesque 2.600
kūfī 1.151, 2.597
Kufic, angular 2.601
Kufic, bordered 2.600
Kufic, broken 2.601, 3.519
Kufic, cursive 3.110

Kufic, eastern 2.601, 3.110
Kufic, Floriated 2.600
Kufic, New Style 2.601, 3.110
Kufic, Persian 3.110
Kufic, pseudo- 2.600
Kufic, Qarmathian 2.601
Kufic, semi- 3.110
Kufic, square 2.600–601
Kufic, western 2.601
Kufic script 2.597–604, 3.517–518, 4.130
Kufra Arabic 3.55–56
Kühnel, H. 3.7
Kukiya 4.279
Kulk, Friso 3.727
kull 4.17, 4.20
kullī 3.178
Kumari, T.C. 3.128
Kumayt ibn Zayd al-ʾAsadī, al- 3.89
Kumba 1.336
Kumyk 2.557
Kumzārī 2.211, 3.350, 3.478
Kunitzsch, Paul 3.7, 4.543
Kuno, Susumu 4.539
Kunta 2.240
Kunuzi 2.1
kunya 1.50, 2.87, 2.578–579, 3.717–719
Kurāʿ an-Naml 3.32, 3.34, 3.321, 4.89
Kurānī, ʾIbrāhīm al- 2.335
kurasa 1.450
Kurbat [Qurbāṭ] 2.717, 2.220
Kurd ʿAlī, Muḥammad 2.635–638, 2.640
kurdī 3.351
Kurdī, Muḥammad al- 4.98
Kurdish 1.24, 1.88–89, 1.92, 1.95, 1.133–135, 1.176, 1.222, 1.469, 1.482, 2.219–221, 2.414, 2.604–608, 2.639, 2.659, 2.670, 2.695, 2.704, 3.312–313, 3.317–319, 3.350–351, 3.370, 3.583, 4.402, 4.503, 4.508, 4.513, 4.577, 4.590, 4.727
Kurdish, Arbil 2.605
Kurdish, Kurmanji 2.604–608, 3.351
Kurdish, Serdeşt 2.605
Kurdish, Şino 2.605
Kurdish, Sorani 2.604–608, 3.351
Kurdish in Turkey 4.578–579
Kurdistan 2.60, 2.414, 2.604
Kurdistan, Arabic in 2.419, 2.423, 2.532
Kurds 1.527, 2.59, 2.61, 3.351, 4.580
Kuria Muria Islands 3.478
Kuriyagawa, F. 4.670
Kurmanji 2.604–608, 3.351
Kuroda, Sige-Yuki 4.524, 4.721
Kurpershoek, P. Marcel 3.327, 3.500, 3.646, 4.129
kursī 2.228
Kuryłowicz, Jerzy 1.119–122, 1.188, 1.565, 2.191, 3.190, 3.446, 4.339, 4.341, 4.444–445, 4.639
Kushitic languages → Cushitic languages
Kusters, Wouter 3.491, 4.367–375
Kūt 2.571
kutab 1.660
Kutscher, Eduard Yechezkel 2.276, 4.558
kuttāb 1.533, 2.10, 2.630, 3.375, 4.131
kuttur 2.220
Kutubī, Ibn Šākir al- 3.540
Kutubiyya, mosque of 2.43–44
Kuwait 1.26, 1.339, 1.451, 2.105, 2.213, 2.215, 2.572, 2.608–620, 3.63, 3.313, 4.123, 4.470
Kuwait, Arabic in 3.525
Kuwait, youth speech in 4.767
Kuwait City 2.609
Kuwaiti Arabic 1.186, 1.221, 1.340, 1.342, 1.657–659, 2.97, 2.102, 397, 2.608–620, 3.609, 3.679, 3.740, 4.65, 4.163, 4.303
Kuwaitis 2.609
Kuzeytepe 1.114–115
Kwayriš 1.23
Kypchak → Kipchak
Kyrghyz → Kirghiz

L

L 3.239
L1 → first language
L2 → second language
lā 3.202, 3.248, 3.257, 3.272, 3.355–356, 3.359, 4.398–399, 4.544
la- 1.70–71, 1.92, 1.201–204, 1.454, 1.564, 2.22, 2.115, 2.308, 2.355, 2.477, 2.491, 2.577, 3.98, 4.429
lā li-nafy al-jins 2.186
lā li-n-nafy 1.70
lā of absolute negation 4.41
laʿalla 2.354–355, 3.591
Labai, Zainuddin 2.337
Labaxi, al- 2.260
Labbais 2.325
labial, velarized 1.342
labialization 1.302, 1.390, 2.123, 2.460, 2.621, 3.290, 3.293, 3.525, 4.107, 4.109, 4.302
Labīd ibn Rabīʿa 2.628, 3.646
labiodental 3.598
labiovelar 3.599, 4.302
labiovelarization 2.621–623, 3.276
Labov, William 1.439–440, 1.632, 2.33, 2.207, 2.275, 2.319–321, 2.506, 3.603–605, 3.655, 4.628, 4.632, 4.728
labs 4.25
Lac 1.360
Laccadive Islands 3.128–129
Lacerenza, Giancarlo 3.469
LaCharité, Darlene 2.122–131
Ladd, Robert D. 2.395, 2.399, 3.724
Ladefoged, Peter 2.229, 2.460, 3.123, 3.125, 3.600, 73.25, 4.636, 4.667–668
Ladino 1.98, 1.100, 2.491
ladun 1.310, 1.467
Ladusaw, William 3.653, 3.655
lafāf 2.677
lafḏ̣ 1.428–429, 2.88, 2.304, 2.424, 2.443, 2.445, 2.541, 2.579, 2.623–628, 3.45, 3.116, 3.159–164, 4.202–203, 4.287, 4.447, 4.473–475, 4.684
lafḏ̣a 2.548, 2.623, 4.103
lafḏ̣ī 1.67, 2.625, 3.234, 3.539, 4.739
Laferrière, Pierre 1.503
lafīf 4.121, 4.645
lafīf mafrūq 2.448
lafīf maqrūn 2.448
Lagarde, Michel 1.256–257
laġw 2.307–308, 3.91, 4.235
Lagwan 1.334
Laḥana 2.628
Lahchiri, Mohamed 3.295
Lahdo, Ablahad 4.580–581
Lahej 4.751
Lahej Arabic 4.750–751, 4.753–754, 4.756
Laḥḥām, Durayd 1.546
lahinna- 1.203
lāḥiq 3.539
Lahiri, Aditi 3.525
Lahiri, Utpal 3.653
lahja 3.88, 4.252
lahja ʿāmmiyya dārija 1.173
lahja ḥirfiyya 2.468
Lahlou, Moncef 1.417, 3.726–727
Lahmeur, Karim 1.457
laḥn 1.174, 1.398, 1.618, 2.178, 2.281, 2.628–634, 3.35, 3.92, 3.346, 4.336

laḥn al-ʿāmma 1.544, 2.629, 3.35, 4.217
Laitin, David D. 3.316
lajlaj 2.676
Laka, Itziar 3.655
Lake Albert 2.588
Lake Chad 1.333, 1.360–363, 1.605, 2.552, 3.373, 4.708
Lake Chad Arabic 1.611
Lake Fitri 1.333, 1.362
Lake Nasser 2.1, 3.435
Lakhmid Dynasty 3.580
lākin 1.468–469, 2.355
lākinna 1.202, 1.648, 2.354–355, 2.357, 2.360, 3.591
Lakoff, George 1.41, 1.423, 1.425, 4.165
Laks, Bernard 3.618, 622
Lakshadweep 3.128–129
Lakshmi, V. Swarajya 4.452–454
Lalami, Laïla 3.382, 3.384
lam 1.196, 1.408, 2.263, 2.279, 2.427, 3.97, 3.221, 3.248, 3.355, 3.359, 4.398–399, 4.544
lām al-ʾamr 1.203, 2.327
lām al-fāriqa 2.355
lām al-ibtidāʾ 1.557
lām al-ʾiḍāfa 2.294
lām ʾalif 1.317, 3.308, 3.337, 3.341, 4.133, 4.561
lām ʾalif muḥaqqaqa 4.133
lām ʾalif muxaffafa 4.133
lām ʾalif warrāqiyya 3.110, 3.112, 3.341, 4.133
lām al-qasam 4.429
lām at-taʾkid 1.202–203, 1.559
lām at-tawkīd 1.202
lām jawāb al-qasam 1.202
lām muʾḏina 1.202
lām muwaṭṭiʾa 1.202
lām muzaḥliqa 1.203
lām zāʾida 1.202
Lambert, Richard D. 2.681
Lambert, Wallace E. 2.652, 2.660–661
Lamberti, Marcello 4.275
Lambrecht, Knud 2.113–116, 4.316, 4.503–504, 4.507
lament 2585, 4.104
lamḥa 2.579
lammā 1.467–468
Lamu archipelago 1.662
lan 1.196, 2.362, 2.364, 3.355, 3.359, 4.38, 4.361, 4.363, 4.398–399, 4.746
Landau, Jacob 2.438
Landberg, Carlo de 1.544, 1.598, 2.258, 3.478, 4.302–303, 4.306, 4.310, 4.750, 4.753–757
Landelinus, P. 2.334
landing site 4.718
Landman, Fred 3.653
Lane, Edward William 1.171, 2.308, 2.433, 2.451, 3.22–23, 3.30, 4.103, 4.118–119, 4.446
Lang, Jürgen 1.590
Langacker, Ronald W. 4.37–39, 4.501, 4.505, 4.624–625
Langen, August 4.76
Langendoen, D. Terence 2.365, 4.346
Langer, Michael 1.465–467
language 1.442, 1.613
language, accusative 4.532
language, administrative 4.73, 4.577
language, agglutinative 2.346, 2.353
language, colonial 1.529, 1.531, 1.630, 2.694, 2.703
language, community 1.418
language, corrupted 1.483
language, corruption of 1.615, 1.619, 2.628
language, creation of 2.178, 4.684–685
language, diplomatic 2.456
language, dominant 1.447, 2.659
language, ethnic 2.659
language, figurative 3.122
language, foreign 2.696, 2.703
language, Homeric 3.646
language, indigenous 2.695
language, inflected 2.346
language, institution of 2.178, 2.426
language, literacy 1.597
language, liturgical 2.1
language, majority 1.370, 1.372
language, maritime 2.456
language, matrix 1.635

language, middle 1.632–633
language, minority 1.370, 1.416, 1.444, 2.659, 2.696, 2.704, 2.713
language, mixed 2.217, 3.58, 3.60, 3.640, 4.621
language, moraic-timed 3.725
language, mystical 4.76
language, national 1.483, 1.529, 1.534, 1.655, 2.703
language, national standard 1.596
language, native 2.702–703, 4.139
language, natural 1.457
language, official 2.438–440, 2.464, 2.635, 2.647, 2.695, 2.699, 3.31
language, oral 2.648, 3.76, 3.79
language, origin of 4.684
language, play 2.468, 2.471
language, political 4.77–79
language, public 2.642, 2.645–646
language, recipient 4.371
language, sacred 2.656, 4.72
language, scientific 1.475
language, secret 2.217, 2.468, 4.156–160, 4.252
language, sign → sign language
language, source 3.27, 4.370
language, spoken 1.615–616
language, standard 2.655, 3.12
language, street 2.463
language, stress-accent 2.396, 2.400
language, stress-timed 3.725
language, syllable-timed 3.725
language, target 3.27, 4.139, 4.141
language, third 1.597, 1.632, 1.634
language, third level of 4.502
language, trade 3.60
language, transmission of 3.188
language, vehicular 3.58
language, verb-initial 4.392
language, written 1.483, 1.529, 1.652
language academies 1.453, 1.532, 2.266–267, 2.634–642, 2.699, 2.701, 2.703, 3.13–14, 3.18, 3.241, 3.347, 4.460
language academies, union of 2.640–641
language academy, Amman 2.640
language academy, Arabic 2.30
language academy, Baghdad 2.638–640
language academy, Cairo 2.637–638
language academy, Damascus 2.635–637
language academy, Kurdish 2.639
language academy, Rabat 2.635
language academy, Syriac 2.639
language academy, Tunis 2.635
language acquisition 1.375, 1.432, 2.689
language acquisition → first language acquisition; second language acquisition
language and culture 1.527–536
language and gender 2.642–650
language area 4.371–372
language as a social phenomenon 1.439
language attitude 2.650–664, 2.682, 2.710
language attrition 2.681–687, 4.320
language awareness 2.714
language broadcasts 2.712
language change 1.444, 1.532, 2.262, 2.275, 3.10, 3.13, 3.36, 4.372, 4.628
language change, external 1.298
language change, internal 1.298
language choice 1.271
language comprehension 1.421
language contact 1.440, 1.489, 1.495, 1.502, 1.584, 1.607, 1.635, 2.29, 2.276, 2.341, 2.373, 2.664–674, 2.684, 2.707, 2.713, 3.189, 3.406, 4.368, 4.370–371, 4.373
language contact situation 2.664
language contact, Arabic/Turkish 4.581, 4.587, 4.589–594
language death 2.715, 4.320–322
language delay 2.675, 4.145
language development 1.340, 1.342, 1.368, 1.370, 1.372, 2.676, 2.679
language disorder 2.674–675, 2.687
language disorder, acquired 2.688
language disorder, developmental 2.688
language distance 2.682
language drift 2.262
language education 2.705
language engineering 1.458
language environment 2.682
language games 2.692, 3.305
language identification 1.455, 1.459

language ideology 2.651, 2.654–655, 4.325
language impairment 1.372, 2.97, 2.674–681, 2.687–694
language instruction 1.651
language learning 1.484–485
language learning, computer-assisted 4.155
language learning, imperfect 2.588
language loss 1.370, 2.73, 2.76, 2.681–687, 2.711, 2.713, 2.715, 4.320
language maintenance 2.710, 2.713–714, 3.396
language mixing 1.593
language of administration 3.346
language of education 2.700
language of orality 3.76
language of the chancery 2.599
language of the press → Media Arabic
language of the Quʾrān → Quʾrānic Arabic
language pathology 1.425, 2.674, 2.687–694
language pedagogy 4.139
language performance 1.457
language planning 1.441, 2.672, 2.694–707
language policy 1.55–56, 1.289, 1.370, 1.482, 2.694–707, 3.20, 3.74
language policy in Turkey 4.578, 4.581
language practice 2.653
language processing 1.455–456, 1.460, 2.689
language processing, Arabic 1.215
language production 1.342, 1.369, 1.421, 2.691
language proficiency 2.73–75, 2.682, 4.150–151
language purification in Turkey 4.577
Language Question in Malta 3.143
language reform 2.672, 2.701
language reform, Persian 2.673
language reform, Turkish 2.666, 2.673, 4.580, 4.586
language schools 2.8
language shift 1.445, 1.536, 2.59, 2.255, 2.499, 2.659, 2.707–716, 4.368, 4.370–371, 4.581
language shift in the United States 3.396
language shift, Amazigh 2.707–716
language shift, Arabic in Turkey 4.582
language spread 2.659
language standardization 2.264
language target 3.25
language teaching 1.482
language technology 1.455
language transfer 1.651
language typology 2.319, 4.725
language universals 4.396, 4.725
language use 1.647
language variation → variation, language
language variety 1.613
language vitality 2.73
Language Weaver 1.460
Lanza, Elizabeth 1.368
Lapesa, Rafael 1.99
Lapolla, Randy J. 4.501–502, 4.734–735
Lappish 3.206
laqab 3.717, 3.720, 4.221
laqad 2.115, 4.732
Laqbo 4.276
Larache 4.156
Larcher, Pierre 1.359, 1.573–579, 1.593–594, 1.613, 2.81, 2.357–361, 3.52, 3.56, 3.236, 3.423, 4.329, 4.331–332, 4.638–645, 4.660–661
Larguèche, Evelyne 2.361–365
Larkin, Margaret 2.303, 2.305–306, 2.578, 2.580
Larnaca 1.536
Larsen-Freeman, Diane 2.117, 4.140
Larson, Richard 4.650
Larusa Syntax Parser 1.460
laryngeal 1.206, 1.376, 1.538, 2.229, 3.609, 4.668
laryngeal adjustment 4.666
laryngeal stop 4.754
laryngealization 4.299
larynx 2.365, 3.594, 4.666
Lāšīn, ʿAbd al-Fattāḥ 2.132
Lasnik, Howard 3.230, 4.523
Lass, Roger 4.389
Lašuniyya 2.471, 4.157
Lāt 4.481–482
lāta 1.71, 564

lateral 1.544–545, 2.252, 3.124, 3.189, 3.599, 3.604, 4.107, 4.171, 4.273, 4.302–303, 4.434, 4.479
lateral, fricative 4.754
laterality 2.460
lateralized 1.544, 3.601
Latin 1.97–99, 1.110, 1.126, 1.148, 1.166, 1.187, 1.296–297, 1.332, 1.404, 1.514, 1.562, 1.590, 2.27–28, 2.200, 2.287, 2.289–290, 2.345, 2.457, 2.655, 2.669, 2.705, 3.6–8, 3.51–52, 3.141–142, 3.553–557, 3.726, 4.217, 4.293, 4.572
Latin America 3.1–6
Latin America, Arabic in 2.665, 3.1–6
Latin, Classical 1.614
Latin, Late 1.296
Latin, Low 2.287–288
Latin, Medieval 2.456
Latin, Middle 2.282, 284
Latin, Neo- 1.614
Latin, Pig 2.468
Latin, Sicilian 4.218
Latin, stress in 4.347
Latin, Vulgar 1.562
Latin/Arabic dictionary 3.23
latinier 1.98
Lattakia 1.388
Latvian 4.372
Laufer, Asher 4.668
Lausanne, Treaty of 4.578
Lavandera, Beatriz R. 3.667
Laver, John 3.597, 3.601, 4.337, 4.387, 4.667–668
law 1.110, 1.465, 1.467–469, 1.509, 1.554, 2.477–479, 3.272, 4.60, 4.544
Law of Coordination of Likes 4.193
Law, Vivien 3.31, 3.124
law. Islamic → Islamic law
Lawdar Arabic 4.754
lawḥa 4.134
lawlā 1.557–558
lawm 2.362
Lawrence, T.E. 3.402
Lawson[-Sako], Sarah 2.660–661, 4.323–324
lax articulation 1.278
Laxmī, Ibn Hišām al- 2.628–631
Layard, A.H. 2.572
Layene Order 4.180
layer 2.143
layer, illocutionary 2.143
layer, locality 2.145
layer, modal 2.143, 2.146
layer, modality 2.144
layer, quality 2.145–146
layer, quantity 2.145
layers of borrowing 2.340, 2.343, 2.553
laysa 1.70, 1.191, 1.217, 1.220, 1.386, 1.564, 2.195, 2.550, 3.221, 3.246, 3.355–356, 3.359, 3.739, 4.38, 4.544, 4.745
layta 1.70, 2.354–355, 3.235, 3.591
Layton, Scott C. 3.227
layyin 2.234, 4.560
Lazard, Gilbert 2.406–410, 2.412, 3.574, 3.579, 4.425
lāzim 1.67, 2.581, 3.113, 3.120, 3.224, 4.410, 4.624
LDC-approach 1.516
Lea, Cyril A.E. 3.435
Leander, Pontus 3.409–410
learnability 1.483
learner's error 4.139
learner's variety 1.482
learning, Islamic 1.448
learning algorithm 1.459
learning difficulty 1.482
learning of Arabic 2.648
learning of Arabic in Mali 3.135–141
learning of Arabic in Senegal 4.180
learning strategy 4.140
Lebanese 1.23–24, 1.187
Lebanese Arabic 1.180, 1.274–275, 1.281, 1.311–312, 1.411–412, 1.432, 1.473, 1.538, 1.540–541, 1.554, 1.568, 1.580, 1.603, 1.642, 2.4, 2.161, 2.231, 2.314, 2.367, 2.369–371, 2.395–400, 2.506, 3.240, 3.264, 3.266–268, 3.349, 3.359, 3.384, 3.387, 3.413, 3.429, 3.446, 3.569, 3.612, 3.726, 4.64, 4.67, 4.80, 4.82–84, 4.194, 4.348, 4.397, 4.649, 4.670, 4.678, 4.719, 4.747
Lebanese in North America 3.395

Lebanon 1.370, 1.531, 1.597, 1.602, 2.105, 2.363, 2.699, 3.315, 3.347–348, 4.470
Lebanon, Arabic in 1.432, 1.445, 1.492, 1.502, 1.571, 1.584, 1.607, 1.642, 2.31, 2.278, 2.374, 2.494, 3.15–19, 3.237, 3.613, 3.672, 3.674, 3.702–703
Lebeaux, David 3.388, 3.393
Leben, William 3.461
Lebock, Anne 2.19
Lecerf, Jean 1.603
Leclerc, Jacques 3.312, 3.314
Lecomte, Gérard 1.449, 2.184, 2.578
Leech, Geoffrey N. 1.513, 2.70
Lees, Robert B. 3.388
Leeuwen, Theo van 2.384
Leezenberg, Michiel 1.425
left dislocation 1.414, 2.115, 3.382–384, 4.82, 4.504–505, 4.523, 4.525–527, 4.728
left peripheral 1.350
left-dominant 3.613
legal theory 2.185, 2.187, 2.310
ʿLēgāt (ʿAlēgāt) Arabic 4.240
Leguest, Charles 3.50
Lehiste, Ilse 3.724
Lehmann, Christian 2.191–197, 4.160–161
Lehn, William 4.670
Leibniz, Gottfried W. 2.66
Leiden 1.169–171
Lelubre, Xavier 4.458–466
lemmatization 1.514
lemmatizer 1.515
Lendu 2.588
length distinction 4.564
lengthening 1.383
lengthening, compensatory 1.204
lengthening, phrase-final 2.397
lenis 3.123, 3.125, 4.2
Lentin, Jérôme 1.546–555, 1.596, 3.215–224, 3.237
Leonard, Laurence B. 2.676, 2.689
Leopold, Werner F. 1.368
Lerchundi, José de 3.287
Leslau, Wolf 1.639, 2.52, 2.55, 2.57, 4.1, 4.297, 4.303, 4.307, 4.371, 4.557
Lessau, Donald 2.192–193
Lesser, Jeffrey 3.2
Lethem, G.J. 3.634, 4.716
LeTourneau, Mark S. 1.343–353, 2.169–175, 4.360–367, 4.487–494, 4.741–749
letter 2.238, 2.546
letter form, categories of 4.133
letter frequency 1.150
letter symbolism 1.7
letters, mystic meaning of 4.181
letters, numerical value of 1.150, 4.181
Levantine Arabic 1.223, 1.312, 1.412, 1.668, 2.4–5, 2.36–38, 2.71, 2.74, 2.136, 2.173, 2.204, 2.395, 2.510, 2.584, 2.658, 3.84, 3.359, 3.495, 3.614, 3.688, 3.707, 4.318, 4.323, 4.646, 4.675
level 1.409, 1.419, 1.439–440
level, educational 1.441
level, linguistic 1.633
level, personal 2.143
level, representational 2.143, 2.145
level, structural 2.143
level, stylistic 1.408
level contour 2.397, 2.400
level of informality 4.251
level ordering 2.366
level tone 2.395
level tune 2.397
leveling 1.56, 1.107, 1.259, 1.269, 1.406, 1.593–594, 1.596, 1.632, 1.669, 2.7, 3.8–13, 3.59, 4.323, 4.582, 4.757
leveling, analogical 2.351
leveling, phonological 4.126
levels of politeness in Javanese 2.472
levels of speech 1.632, 1.667
Levelt, Willem 4.269–271
Levenshtein Distance 1.586
Levi, Judith N. 3.388
Levin, Aryeh 1.310, 1.556, 2.94, 2.176, 2.227, 2.237, 2.311–315, 2.432–436, 2.481, 2.545–552, 3.103, 3.105, 3.455, 3.683, 4.1, 4.8, 4.282, 4.414, 4.659
Levin, Beth 4.489
Levine, Baruch A. 3.466

Levinson, Stephen C. 1.485, 1.568, 1.657, 2.19, 2.70, 2.384, 2.587, 3.570, 3.658, 3.677, 3.713, 4.457
Lévi-Provençal, Evariste 1.99, 3287
Levites 1.127
Levtzion, Nehemia 3.135–136
Levy, Mary Mansnerus 3.299, 3.303–304, 3.430, 3.440–441
Levy, Simon 3.286, 3.288, 3.293, 4.157
Lewcock, Ronald 4.698
Lewis, Bernard 4.593
Lewis, Geoffrey 3.504, 4.578, 4.581, 4.585–587
Lewis, Ioan Myrddin 4.277
Lewis, Jonathan E. 2.61
Lewkowicz, Nancy Kennedy 4.42, 4.524–527
lexeme 1.573, 4.339
lexical access 2.692
lexical aspect 1.195–197, 1.199
lexical category 4.742
lexical choice 1.657
Lexical Clause Hypothesis 4.743
lexical compression 1.656
lexical decision task 2.692
Lexical Diffusion Model 3.604–605
lexical error 4.268
Lexical Functional Grammar 1.457, 4.535, 4.657
lexical gap 1.419, 2.29–30, 2661
lexical stress 2.395–396
Lexical Tree Adjoining Grammar 1.457
lexical variation 1.56, 1.95, 3.13021, 3.203
Lexicalist Hypothesis 3.388, 3.393, 4.741
lexicalization 1.313–314, 1.658
lexicography 1.530, 1.574, 2.630, 3.30–38, 3.93, 3.162
lexicography → dictionaries
lexicography, Arabic 1.435, 3.21–22
lexicography, Classical Arabic 1.402
lexicography, Greek 3.31
lexicon 3.228
lexicon, Bedouin 3.55, 4.555–556
lexicon, disguised 2.217
lexicon, mental 2.683, 4.269
lexicon, organization of the 1.654
lexicon, Semitic 3.48, 3.50
lexicon, Swahili 1.665
lexicon acquisition, bilingual 1.371
Lexicon Pentaglotton 1.166
lexicostatistics 4.313
lexifier 4.369, 4.373
Lhomond, Charles François 1.449
li- 1.71, 1.467–468, 1.471, 1.491, 1.625, 2.22, 2.80, 2.294–295, 2.301, 2.531, 3.107, 3.201, 3.248, 3.272, 3.674, 3.703, 4.529–530, 4.533–534, 4.544–545
Li, Charles N. 4.317
Li, Yen-Hui Audrey 3.384, 4.64, 4.67–68, 4.400, 4.747
li-ʾanna 2.354–355
Liberman, Mark 3.615–616
Liblī, ʾAḥmad ibn Yūsuf al- 2.93
libraries 3.172
libraries, Indian 2.329
libraries in Jakarta 2.335
Libya 1.589, 2.105, 2.250, 2.255, 2.454, 3.52–58, 3.316, 4.548
Libya, Arabic in 1.260, 1.571, 1.642, 2.456, 3.702
Libyan Arabic 1.270, 1.642, 2.6, 2.374, 2.390, 2.658, 3.52, 3.54–55, 3.587, 4.655
Libyan Arabic → Tripolis Arabic
Libyan Arabic, Eastern 1.269
Libyans 1.293
Libyc 4.571
Libyco-Berber 1.35, 1.644, 3.52
licensing 2.171, 2.174
licensing, formal 2.174
Lieberson, Stanley 2.59
life stage 4.630
ligature 3.77, 3.514, 4.133
ligature, *lām ʾalif* 3.110
Lightfoot, David W. 3.685, 3.706, 4.39
lightness 2.93, 2.178, 2.353, 2.425, 3.45, 3.108, 4.432, 442–443
lightness scale 2.349
Liguria 4.215
L-iḤmēra Arabic 4.406
Liḥyān 3.468
Liḥyanitic 1.187, 2.198, 3.468, 2.263, 3.227, 4.478–479, 4.481

Limaamu Laye, Seydina 4.180
Limanköyü 1.114
Limassol 1.536
Limsīḥ, ʾAḥmad 1.603
Lindblom, Björn 4.683
Lindholm, Kathryn 1.368
Lindström, Per 4.15
Line Conflation 3.619
Linear Correspondence Axiom 4.746–747
linear order 4.524
Linebarger, Marcia C. 3.653
Lingala 4.386
lingo 4.251
Lings, Martin 3.308, 3.337, 3.340
lingua franca 1.222, 1.255, 1.338, 1.519, 1.521, 1.525, 1.554, 1.593, 1.619, 1.654, 1.667, 1.669, 2.216, 2.250–251, 2.268–270, 2.272, 2.274, 2.456, 2.517–518, 2.588, 2.647, 2.671–672, 2.708, 3.58–64, 3.135, 3.142, 3.286, 4.290, 4.375, 4.573
lingua franca, African 2.252
lingua franca, Afrikaans as 4.291
lingua franca, Aramaic 3.409
lingua franca, Melkite 3.218
lingua franca, Nabataean 4.482
linguicide 1.534
linguistic correctness 2.85, 2.88, 2.275
Linguistic Data Consortium 1.512–513
linguistic insecurity 3.76
linguistic island 1.589
linguistic island → enclave
linguistics, applied 1.647
linguistics, areal 1.586
linguistics, comparative 1.574, 2.66, 3.188
linguistics, computational 1.455–465, 1.511, 1.515
linguistics, computer 1.484
linguistics, contact 2.286
linguistics, descriptive 1.457
Linguistics, Generative → Generative Grammar
linguistics, historical 1.575
Linguistics, Transformational → Transformational Grammar
Linjān 2.600
linkage, logical 4.500
linkage, manner 4.500
linkage, spatial 4.500
linkage, temporal 4.500
linking 1.472
linking element 3.544, 3.587–588, 4.500
linking problem 4.488
Linux/Unix 2.380
Linxia 1.381
lip reading 2.675, 4.223
lip rounding 4.670
Lipiński, Edward 1.544, 1.563, 1.638–639, 1.644, 3.49, 3.227, 3.423, 3.453, 3.717, 3.719, 3.721, 3.738, 4.1–2
Lippi-Green, Rosina 2.655
liquid 3.625
lisān 2.85, 2.359, 2.544, 3.71–74, 3.88, 3.90, 3.93, 3.689
Lisramic 1.35
lissa 1.198
literacy 1.602, 2.648, 3.12, 3.57, 3.74–80, 4.324
literacy, adult 3.76
literacy, bilingual 1.371
literacy statistics 3.78
literary criticism 2.304
literature, Arabic in Indonesia 2.334–336
literature, Arabic in Kerala 3.129
literature, Arabic in Senegal 4.181–182
literature, *mahjar* 3.4
literature, Mamluk 4.132
literature, popular 2.214
literature, teaching of 2.107–109
literature, Zoroastrian 3.574
Lithuanian 4.372
litotes 3.632
little arrow method 1.586
little-v 4.650
Littmann, Enno 2.218, 3.470, 3.472, 3.718, 4.75, 4.480–481
Liu Zhi 1.381
Liverani, Mario 3.414
Livingstone, Alasdair 3.466
Llobregat, Enrique 1.98
Lloyd, Geoffrey E.R. 4.11
Llull, Ramon 1.171, 2286
Lmnabha Arabic 4.286
loan affix 2.285

loan blend 2.285, 2.474
loan meaning 2.286
loan mismatch 2.474
loan name 2.285
loan phrase 2.342
loan transfer 2.474
loan translation 2.10, 2.255, 2.285, 2.375, 2.462, 2.465, 2.684, 3.151, 3.241, 4.41, 4.167–168, 4.465, 4.545
loan usage 2.286
loan, hybrid 4.723
loans, typology of 2.286
loanverb 1.297, 1.417–418, 1.496, 1.498, 1.664, 2.55, 2.214, 2.253, 2.255, 2.290, 2.342, 2.667–668, 2.670, 3.156–158, 3.438, 4.274, 4.385, 4.434, 4.585, 4.591–592, 4.599–601, 4.606, 4.610–611, 4.724
loanverb, Maltese 3.156–158
loanword 1.85, 1.88–89, 1.95, 1.233, 1.235, 1.242, 1.280, 1.403, 1.416, 1.501, 1.575, 3.188, 3.201, 3.426–427, 4.95, 4.281, 4.311
loanword, adaptation of 2.665
loanword, adoption of 2.665–666
loanwords, Arabic 1.133, 1.135, 1.545
loanwords, Arabic in Acehnese 1.5–12
loanwords, Arabic in Afrikaans 4.293–295
loanwords, Arabic in Albanian 4.584
loanwords, Arabic in Amharic 2.54–55
loanwords, Arabic in Argobba 2.54–55
loanwords, Arabic in Bambara 1.256–257
loanwords, Arabic in Bengali 1.259, 1.286–289, 2.674
loanwords, Arabic in Berber 1.289
loanwords, Arabic in Bulgarian 4.259–267, 4.584
loanwords, Arabic in Coptic 1.495–496, 1.498–501
loanwords, Arabic in Czech 4.259–267
loanwords, Arabic in Domari 2.221
loanwords, Arabic in English 2.25, 2.31, 2.33
loanwords, Arabic in Ethiopian 2.54–55
loanwords, Arabic in French 2.26
loanwords, Arabic in Fulfulde 1.545, 2.137–142
loanwords, Arabic in Geʿez 2.51–52
loanwords, Arabic in Hausa 1.545, 2.250, 2.667
loanwords, Arabic in Ibero-Romance 1.545, 2.281–286
loanwords, Arabic in Indonesian/ Malay 1.545, 2.334, 2.340–345
loanwords, Arabic in Italian 2.454
loanwords, Arabic in Italo-Romance 4.216
loanwords, Arabic in Ivrit 2.461
loanwords, Arabic in Javanese 2.472–477
loanwords, Arabic in Kanuri 2.552
loanwords, Arabic in Kazakh 2.557–559
loanwords, Arabic in Kurdish 2.604–608
loanwords, Arabic in Malagasy 1.544, 3.126
loanwords, Arabic in Malay 2.340–345, 4.293, 4.434
loanwords, Arabic in Malayalam 3.128
loanwords, Arabic in Mande 1.256
loanwords, Arabic in Modern South Arabian 4.297–300
loanwords, Arabic in Nabataean 3.466
loanwords, Arabic in Neo-Aramaic 2.671, 3.370–373
loanwords, Arabic in Nubian 2.671, 3.435
loanwords, Arabic in Persian 2.406–413, 2.666–667, 3.573–580, 3.582, 4.424–425
loanwords, Arabic in Polish 4.259–267
loanwords, Arabic in Romance 1.103–104
loanwords, Arabic in Russian 4.259–267
loanwords, Arabic in Serbo-Croatian 4.259–267
loanwords, Arabic in Slavonic 4.259–267
loanwords, Arabic in Somali 4.272–275
loanwords, Arabic in Songhay 4.279
loanwords, Arabic in South Arabian 2.668

loanwords, Arabic in Spanish 3.189
loanwords, Arabic in Swahili 1.660–666, 2.668, 2.670–671, 4.381–387
loanwords, Arabic in Tajik 4.424
loanwords, Arabic in Tamil 4.433–436
loanwords, Arabic in Tatar 4.450
loanwords, Arabic in Telugu 2.669, 4.452–454
loanwords, Arabic in Turkish 2.673, 3.503, 4.583
loanwords, Arabic in Urdu 3.511, 4.434, 4.595–603
loanwords, Arabic in Uyghur 4.603–608
loanwords, Arabic in Uzbek 4.608–612
loanwords, Arabic in Wolof 2.670, 4.182, 4.723–725
loanwords, Arabic in Yoruba 3.373, 4.758–763
loanwords, Arabo-Persian in Turkish 4.577
loanwords, Aramaic 1.88, 1.178, 1.197, 1.401, 1.403, 1.541
loanwords, Aramaic in Arabic 1.384, 2.265, 2.373
loanwords, Aramaic in Damascus Arabic 1.554
loanwords, Aramaic in Judaeo-Arabic 2.531, 2.533, 3.218
loanwords, Aramaic in Syrian Arabic 4.403
loanwords, Baghdad Arabic 1.231
loanwords, Bantu in Nubi 1.522
loanwords, Beja in Sudanese Arabic 4.377
loanwords, Berber 1.102
loanwords, Berber in Arabic 1.293–299, 2.373
loanwords, Berber in Ḥassāniyya 2.244, 2.249
loanwords, Berber in Moroccan Arabic 3.286, 3.293
loanwords, Berber in Tunis Arabic 4.570
loanwords, Berber in Tunisian Arabic 4.572
loanwords, Chinese in Indonesian 2.341
loanwords, Coptic in Arabic 1.332, 1.495, 1.501–505, 2.373
loanwords, cultivated 2.455
loanwords, Egyptian in Damascus Arabic 1.554
loanwords, English in Arabic 1.240–241, 1.332, 1.592, 2.29–35
loanwords, English in Arabic 2.29–35
loanwords, English in Arabic Afrikaans 4.293
loanwords, English in Egyptian Arabic 2.9, 2.374
loanwords, English in Gulf Arabic 2.214
loanwords, English in Hausa 2.254
loanwords, English in Jordanian Arabic 2.501
loanwords, English in Libyan Arabic 3.56
loanwords, English in Malayalam 3.129
loanwords, English in Maltese 3.152–154
loanwords, English in Meccan Arabic 3.187
loanwords, English in Palestinian Arabic 2.465
loanwords, English in Sudanese Arabic 4.377
loanwords, English in Swahili 1.663
loanwords, English in Syrian Arabic 2.30
loanwords, English in Tripoli Arabic 4.556
loanwords, English in Urdu 4.602
loanwords, Ethiopic in Arabic 2.56–57, 4.295–296
loanwords, European 1.445
loanwords, foreign 2.637, 2.696
loanwords, French in Arabic 1.60–61, 1.233, 1.240, 1.332, 1.418
loanwords, French in Bambara 1.257
loanwords, French in Bongor Arabic 3.638
loanwords, French in Damascus Arabic 1.547, 1.554
loanwords, French in Egyptian Arabic 2.9, 2.374
loanwords, French in Hausa 2.254

loanwords, French in Levantine Arabic 2.374
loanwords, French in Meccan Arabic 3.187
loanwords, French in Modern Standard Arabic 3.17
loanwords, French in Moroccan Arabic 2.122–131, 2.374, 3.286
loanwords, French in Shuwa Arabic 1.335
loanwords, French in Syrian Arabic 2.30
loanwords, French in Tunisian Arabic 4.564, 4.570, 4.574
loanwords, from Standard Arabic 2.565
loanwords, function words as 4.599
loanwords, Gallo-Romance in Sicilian Arabic 4.218
loanwords, Greco-Roman in Tunisian Arabic 4.572
loanwords, Greek in Arabic 1.332, 1.403, 1.499, 2.198–202, 2.265, 2.373, 4.1, 4.545
loanwords, Greek in Coptic 1.498
loanwords, Greek in Hebrew 4.1
loanwords, Greek in Persian 4.545
loanwords, Greek in Sicilian Arabic 4.218
loanwords, Greek in Syriac 4.545
loanwords, Hebrew in Arabic 1.240, 2.666
loanwords, Hebrew in Cairene argot 4.255
loanwords, Hebrew in Jerusalem Arabic 2.482
loanwords, Hebrew in Judaeo-Arabic 2.266, 2.531, 2.533, 3.218
loanwords, Ḥimyaritic in Arabic 2.259
loanwords, Hindi/Urdu 1.241
loanwords, Ibero-Romance in Andalusian Arabic 2.287–290
loanwords, Indian in Gulf Arabic 2.211
loanwords, Italian in Arabic 2.456–459, 2.669
loanwords, Italian in Cairene argot 4.255
loanwords, Italian in Damascus Arabic 1.554
loanwords, Italian in Libyan Arabic 2.374, 3.56
loanwords, Italian in Maltese 3.148
loanwords, Italian in Meccan Arabic 3.187
loanwords, Italian in Syrian Arabic 2.30
loanwords, Italian in Tripoli Arabic 4.549, 556
loanwords, Italian in Tunis Arabic 4.570
loanwords, Italian in Tunisian Arabic 4.574
loanwords, Ivrit in Arabic 2.464–467
loanwords, Ivrit in Palestinian Arabic 2.670
loanwords, Judaeo-Aramaic 1.240
loanwords, Kurdish 1.88–89, 1.95, 1.469
loanwords, Kurdish in Arabic 2.608
loanwords, Kurdish in Domari 2.219
loanwords, Latin in Arabic 2.199, 3.6–8
loanwords, Latin in English 2.669–670
loanwords, Lebanese in Damascus Arabic 1.554
loanwords, lingua franca in Tunisian Arabic 4.573
loanwords, Malay in Arabic Afrikaans 4.293
loanwords, Modern Standard Arabic 1.233
loanwords, nativization of 2.668–669, 2.671
loanwords, Neo-Aramaic in Arabic 3.218
loanwords, Nubian in Sudanese Arabic 2.667, 3.435, 4.377
loanwords, Ottoman 1.332
loanwords, Persian 1.233, 1.240–241, 1.275, 1.403, 1.452, 1.617
loanwords, Persian in Arabic 2.265, 2.373, 2.406, 3.580–584
loanwords, Persian in Arabic slang 4.257
loanwords, Persian in Damascus Arabic 1.554
loanwords, Persian in Gulf Arabic 2.210, 2.374
loanwords, Persian in Iraqi Arabic 2.415–417, 2.423

loanwords, Persian in Malay 2.341
loanwords, Persian in Meccan Arabic 3.187
loanwords, Persian in Ottoman Turkish 2.670
loanwords, Persian in *rajaz* 4.33
loanwords, Persian in Sudanese Arabic 4.377
loanwords, Persian in Syrian Arabic 2.30
loanwords, Persian in Turkish 4.584
loanwords, Persian in Urdu 4.595–603
loanwords, Perso-Arabic in Bengali 1.286–289
loanwords, Portuguese 1.241
loanwords, Portuguese in Brazilian Arabic 3.4–5
loanwords, Romance 1.99–100, 1.102, 1.110
loanwords, Romance in Maltese 3.149, 3.152
loanwords, Russian in Tajik 4.425
loanwords, Russian in Tatar 2.673, 4.450
loanwords, Russian in Uzbek 4.612
loanwords, Russian in Yupik 2.671
loanwords, serial 4.599
loanwords, South Arabian in Arabic 4.295–297
loanwords, South-Arabian in Ḥimyaritic 4.753
loanwords, Spanish in Moroccan Arabic 3.286, 3.290
loanwords, Spanish in Syrian Arabic 2.30
loanwords, Standard Arabic 1.324, 1.332, 1.366, 1.405, 1.407–410, 1.546–548, 1.554, 2.323, 4.565
loanwords, Sudanese Arabic in Hausa 2.255
loanwords, Swahili in Nubi 1.523
loanwords, Syriac 1.178
loanwords, Tajik in Uzbekistan Arabic 4.614
loanwords, Turkish 1.60, 1.88–89, 1.95, 1.118, 1.233, 1.240–241, 1.275, 1.332, 1.389, 1.392–393, 1.395, 1.397
loanwords, Turkish in Anatolian Arabic 4.591
loanwords, Turkish in Arabic 2.373, 2.667, 4.581, 4.584, 4.589–594
loanwords, Turkish in Damascus Arabic 1.546, 1.554
loanwords, Turkish in Iraqi Arabic 2.415–417, 2.423
loanwords, Turkish in Jerusalem Arabic 2.491
loanwords, Turkish in Jordanian Arabic 2.500
loanwords, Turkish in Libyan Arabic 3.56
loanwords, Turkish in Meccan Arabic 3.187
loanwords, Turkish in Sudanese Arabic 4.377
loanwords, Turkish in Syrian Arabic 2.30, 2.668, 4.592
loanwords, Turkish in Tripoli Arabic 4.556
loanwords, Turkish in Tunis Arabic 4.570
loanwords, Turkish in Tunisian Arabic 4.573
loanwords, verbal → loanverb
loanwords from Classical Arabic 2.242, 2.565, 3.607
loanwords in Arabic 4.458
loanwords in Ottoman Turkish 2.666
loanwords in the *Qurʾān* 2.86, 2.666, 3.7, 3.36, 4.295
loanwords in Turkish 2.666
local adverb 1.22
Local Impairment hypothesis 4.144
locality conditions 2.170
localization 4.229
localization, temporal 1.568
localization in sign language 4.229
localizationist theory 2.690
localizer 1.567
location 1.184, 4.487, 4.490
locational 3.80, 3.83
locative 1.344, 2.146, 2.273, 3.68, 3.80–88, 3.107, 3.109, 4.318, 4.485, 4.738
locative clause 1.285

locative marker 3.83–84
locative paradigm, universal 3.80
locutionary 4.328
locutionary act 3.679
Lodhi, Abdul Aziz Y. 1.666–671
Löffler, Heinrich 1.585
Loftus, Elizabeth F. 2.681
Loftus, Geoffrey A. 2.681
Loftus, William Kenneth 2.572
logic 2.176, 2.183, 2.185–187, 2.310, 2.427, 3.115
logic, Arabic 2.435
logic, Greek 2.234–235, 2.425–426, 2.434, 2.479, 2.544, 3.113, 3.115, 3.161, 4.235
logical form 2.368, 3.228, 3.230, 3.232, 4.489
logograph 1.1–2
Logone Birni 1.334, 1.360
logophoric 4.54–55
lógos 2.66, 4.545
Löhr, Max 2.481
Loi Corvetto, Ines 3.727
Lom 2.217
Lomavren 2.217
Lombardy 4.215
London 2.518
London Polyglot 1.167
Long, James 1.287
Long, Michael H. 2.117, 4.140, 4.145
Longacre, Robert E. 4.727, 4.731
Longobardi, Emiddia 1.341
Longobardi, Giuseppe 3.232
Lonnet, Antoine 4.297–300
Lord, Albert B. 3.499, 3.645–646, 4.201
Lorimer, D.L.R. 2.217, 2.572
Lorimer, J.G. 2.572
Lorimor, Heidi 4.191–195, 4.396–397, 4.658
Los Angeles 3.397
loss of grammar 2.690
loss-and-gain model 4.162–163
Lotuko 1.519
Loubignac, Victorien 3.287, 3.290
Loucel, Henri 2.178, 4.684
Low Variety 1.419, 1.439, 1.447, 1.633–634, 1.667, 2.10, 2.265, 2.654, 2.695, 2.708, 3.70
Low Vowel Deletion 3.621–622
Löwenstamm, Jean 1.121–122, 2.168, 3.302
lowering, vowel 4.671
LSPERT 1.210
Lucera 4.216
Lucknow 2.326, 2.329
Ludolphus de Suchem 3.142
Luffin, Xavier 3.634–639, 3.642
luġa 1.399, 2.89, 2.355, 2.544, 2.628–629, 3.15, 3.72–73, 3.88–95, 3.689–690, 4.252
luġa ʾakalūnī l-barāġīṯ 1.616, 1.556, 2.430, 3.691
luġa bahlawāniyya 2.220
luġa gajariyya 2.220
luġa ḥalabiyya 2.220
luġa sirriyya 2.468
luġa wusṭā 1.668, 2.377
luġa xāṣṣa 2.468
Luganda 2.597
Lugard, Captain 2.588
luġat al-muṯaqqafīn 1.631, 1.668
luġat iṣṣād 4.157
luġat issīm 4.157
luġat issīn 4.157
luġawī 3.119–121
Luġda 2.427
luġz 4.157
Lujamī, ʾAdīb al- 3.41
Luke, Saint 3.141
Lüling, Gerhard 4.6
lullaby 4.59
Lumsden, Michael 4.316, 4.516
lunar letters → moon letters
Luo 1.333, 2.597
Lur 2.588
Lūṭ ibn Sām ibn Nūḥ 3.345
luṯga 2.677
Luxenberg, Christoph 1.180, 3.719, 4.6
Luxor 1.299, 2.2, 2.217
Luxor Arabic 2.2, 2.5
luzūm 4.103
LVCSR 1.213
Lydda 2.464
lying 2.362
Lykke-Nielsen, Helle 2.679
Lyon 3.262, 4.457
Lyons, John 1.474, 3.263, 3.265, 3.570

M

M’barek, Mohamed Nait 1.418
mā 1.70, 1.248, 1.268, 1.408, 1.467, 1.476, 2.133–135, 2.194–195, 2.387–388, 2.479, 2.490, 2.521, 2.570, 2.577, 2.615, 3.221, 3.356–357, 3.359, 3.691, 4.26, 4.40, 4.61, 4.64, 4.38, 4.164, 4.399
Ma Fuchu 1.380–381
mā ḥijāziyya 3.691
mā huwa huwa 4.738
Ma Jian [Muḥammad Mākīn aṣ-Ṣīnī] 1.382
Ma Lianyuan [Muḥammad Nūr al-Ḥaqq ibn Luqmān] 1.380–381
Maʾanyan 3.125
maʿa 3.675
Maʿadd 3.345
Maalej, Zouhair 1.423–424, 3.224–227, 4.568, 4.624–627
Maalouf, Amin 2.61, 2.703
Maamouri, Mohammed 1.516, 1.632, 2.111, 2.694, 2.701–702, 2.705, 3.74–80, 4.646
Maʿān 2.499
Maʿān Arabic 2.498
Maʿānī l-kalām 1.428, 4.330
Maʿānī l-Qurʾān 3.92, 4.8–9
maʿānī n-naḥw 2.305, 3.161
maʿānī, ʿilm al- 1.428–429
Maʿarrī, al- → ʾAbū l-ʿAlāʾ al-Maʿarrī
Maas, Utz 1.221, 1.482–489, 2.74, 3.60, 3.725, 4.52
Maasina 2.137
Maʿāza 2.2
Maba 4.376
mabnī 1.191, 1.309–310, 1.559, 2.424, 3.429, 4.23, 4.119, 4.432
mabnī ʿalā l-mubtadaʾ 4.221
mabnī ʿalayhi 2.435
mabnī li-l-fāʿil 2.93
mabnī li-l-mafʿūl 2.93, 2.625
mabnī li-l-majhūl 3.224–225
mabsūṭ 3.112, 3.308, 4.133
Mac 2.380
Macalister, R.A.S. 2.220–221
Macdonald, Michael C.A. 3.411, 3.413–415, 3.418, 3.420, 3.464–475, 4.478
MacEachern, Margaret 3.627
Macedonian 3.726, 4.371
Macedonian, Arabic loanwords in 4.259–267
Măcelaru, Adrian 1.355–360
Mach, Rudolph 4.686
Machali, Rochayah 2.472–477
Machek, Václav 4.259
machine translation 1.207, 1.214, 1.455, 1.458–460
machine translation system 1.515
MacKay, Donald 4.335
MacKenzie, D.N. 2.605, 2.607
Mackenzie, John Lachlan 2.143
Macler, Frédéric 3.469
MacMichael, Harold 3.435
macro-discrimant 1.585
Macro-Fluency Model 2.684
Macuch, Rudolf 3.410
MacWhinney, Brian 2.686, 2.690
maḍā 3.96–97
Madaba 2.506

Madagascar 1.448, 1.544, 1.663, 3.125–128
Madāʾin 3.574
Madāʾin Ṣāliḥ 3.471, 4.478
madanī 1.605
Madanī, al- 4.78
Madariaga 1.99
madarsa 1.665
madd 1.152, 4.426–428
madda 1.1, 1.574, 4.518
Mʿaḏḏamīye Arabic 4.406–407
Maddieson, Ian 2.229, 3.65, 3.123, 3.125, 3.600, 3.643, 3.725, 4.1, 4.636, 4.667, 4.683
Madenli 1.115
madḥ ən-näbi 3.173
Maḏḥā 2.256
maḏhab al-kalāmī, al- 4.210–211
Madhabic 4.301
Maḏḥij 3.94, 3.467
māḍī 1.309, 1.568, 2.91, 2.358, 3.96–100, 3.270, 3.542, 3.548, 3.551, 4.455, 4.638
madīd 3.211, 3.214
madīḥ 3.173
Madina, Maan Z. 3.26
Madinah → Medina
Madīnat az-Zahrāʾ 2.44
Madjid, Nurcholis 2.338
Madkūr, ʾIbrāhīm 2.634, 2.638
Madouni, Jihane 1.56
Madras 2.329
madrasa 1.258, 1.449–451, 1.660, 1.662, 2.40, 2.43, 2.188, 2.325–327, 2.329, 2.338, 3.129, 3.137–139, 3.347, 3.508, 4.291, 4.584
madrasa, Persian 2.327
madrasa, Shiʿite 2.327
Madriyeva, Monjiya 4.450
Madurese 4.539
mafhūm 2.426, 2.581, 4.474
Mafraq 2.500
mafrūq 3.173, 3.539
mafrūq, watid 3.209–210
maftūḥ 3.307, 4.133, 4.560
mafʿūl 1.40, 2.93–94, 2.356, 2.427, 2.551, 3.100–106, 3.424, 3.558, 4.237, 4.318, 4.410–411, 4.413–414, 4.487, 4.540
mafʿūl ʿalayhi 3.101
mafʿūl bihi 1.182, 2.429, 3.100, 3.102–105, 3.108, 3.224, 3.455, 3.460
mafʿūl fī l-maʿnā 3.103
mafʿūl fīhi 2.426, 3.83, 3.97, 3.101, 3.104, 3.106–110, 3.455, 3.460, 3.548, 4.414, 4.540, 4.653
mafʿūl ġayr ṣarīḥ 3.560
mafʿūl ḥaqīqī 3.102
mafʿūl lahu 3.108
mafʿūl li-ʾajlihi 3.101, 3.455, 3.460, 4.540
mafʿūl maʿahu 3.101, 3.108, 3.455, 3.460, 4.540
mafʿūl minhu 3.101
mafʿūl muṭlaq 1.436, 3.101, 3.103, 3.108, 3.455, 3.460, 4.52, 4.414, 4.540, 4.661, 4.664
mafʿūliyya 3.102, 3.104
Maga 1.334
Mağaracık 1.114
magazines 1.598, 1.601, 1.603
Maghrebine Arabic 1.75, 1.189, 1.269–270, 1.291, 1.296, 1.492, 1.585–586, 1.590, 1.611, 2.7, 2.74, 2.160, 2.240, 2.455, 2.621, 3.702, 4.281, 4.323, 4.370
Maghut, Muhammad al- 4.213
Maġīlī, Muḥammad ibn ʿAbd al-Karīm al- 3.374–375
Magri, Emmanuel 3.141
maġribī 1.152, 1.605, 3.110–113, 4.132–133, 4.138, 4.218
Maġribī, Ibn Saʿīd al- 4.215
Mahabad 2.604, 2.608
Mahābišah Arabic 4.751
maʿhad 1.449–450
Mahāʾimī, ʿAlāʾ ad-Dīn 2.328
maḥājir 4.133
maḥall 2.293, 2.316, 3.107
Mahāmīd 1.362
Mahamūd, ʾIbrāhīm Ḥāšī 4.277
Maharashtra 2.325
Maḥāyil Arabic 4.125, 4.127
maḥḍ 2.294–295
maḥḍ, ism 3.323

Mahdi 1.519, 1.532
Mahdi, Muhsin 3.219
Mahdi, Qasim R. 2.415
Mahdia 4.572, 4.574
Mahdist revolt 2.588
Mahdiya Arabic 1.609
maḥḏūf 2.82, 2.301, 3.307
Maher, Julianne 2.684
Mahfouz, Naguib [Nagīb Maḥfūẓ] 1.601, 3.236, 3.554, 3.633, 3.715–716, 4.663
Mahfouz, Soheir G. 4.86–92
Maḥfūẓ, Nagīb → Mahfouz, Naguib
mahgrāyē 1.129
mahjar 3.4, 4.91
maḥkī 2.537–538
maḥkūm bihi 3.121
maḥkūm lahu 3.121
Mahmoud, Abdelgawad T. 4.490–491
Mahmoud, Youssef 1.630–631, 1.668
Mahmud Kemal İnal 4.98
Mahmud, Ushari A. [ʿUšārī Maḥmūd] 1.521, 1.525, 2.517–518, 3.10, 3.642
Mahmudiyya Canal 1.589
maḥmūl 2.434–435, 3.113–116, 3.177
mahmūs 1.27, 3.123–125, 2.233, 3.602–603, 4.2, 4.667
mahmūz 4.121, 4.645
Mahra 4.750
Mahri → Mehri
Mahriye 1.362
maḥsūs 2.424
maḥw al-ʾummiyya 3.74
Mai of Kanem 4.708
Maiduguri 1.337
Maiduguri Arabic 4.716
māʾil 1.151
Maila, Joseph 2.59
Maimonides 2.529, 3.218, 4.58
Maʿīn 4.295
maintenance, language → language maintenance
maintenance, tense 1.432
maintenance, topic 1.651
maintenance of referent 1.649
maintenance of topic 1.651
Mairi, Lies 1.61
Maizel, S.S. 4.96–97
majalla majmaʿ al-luġa → language academies
majāz 2.305, 2.441, 2.443, 2.579–581, 2.638, 2.641, 3.116–123, 3.651, 4.9, 4.165, 4.210, 4.330, 4.464
majāz ʿaqlī 2.445, 3.117, 3.120–121
majāz at-tarkīb 3.121
majāz at-tašbīh 3.121
majāz fī l-ʾiṯbāt 3.121
majāz ḥukmī 3.121
majāz ʾisnādī 3.121
majāz luġawī 2.445, 3.117, 3.120
majāz murakkab 2.445, 3.120
majāz mursal 3.120–121, 4.165
majāzī 4.203
Majdi, Basim 2.171–172, 4.363–364, 4.366, 4.726
majhūl 1.309, 1.624, 2.93–94, 2.317, 3.224–225, 4.424
majhūr 3.602–603
majhūra 1.27, 3.123–125, 4.2, 4.667
majlis 3.30
majmaʿ al-luġa → language academies
majmūʿ 3.173
Majnūn Layla 4.202
majrā 2.232, 2.302, 2.402, 2.624, 4.13, 4.442
majrūr 2.316, 3.429
Majunga 3.126
Majus, J.H. 3.141
majzūm 1.71–72, 1.196, 2.16, 2.92, 4.639
Makāhīl, Šayx Muḥammad ʿAbdī 4.277
makān 3.108, 3.424,
makāna 4.443
Makary 1.334, 1.337
Makassarese 2.335
Makassarī, Muḥammad Yūsuf al- 2.335
Makdisi, George 3.508, 4.685
Makha 1.655
Makhlouf, Nabila 1.651
maktab 2.325–327, 3.508
maktaba 3.347
Malabar Coast 2.325–326, 3.128, 3.506
Malacca 1.8, 2.340

Malagasy 1.133, 1.544, 1.663, 3.125–128
Malāʾika, Nāzik al- 4.91–92
Malāʾika, Nisār 2.415
Malālḥah Arabic 4.238
Malallah, Seham 2.660
Malawi 1.661–662
Malay 1.8–10, 1.133, 1.135, 1.404, 1.545, 2.26, 2.333–336, 2.338, 2.340–345, 2.668–670, 3.59, 4.290–291, 4.293, 4.434, 4.476–477
Malay Muslims 4.290
Malaya 2.340
Malayalam 1.133, 2.669, 3.128–135
Malayo-Polynesian languages 3.125
Malaysia 2.344
malediction 2.362, 2.378
malḥūn 2.538, 2.541, 3.219, 3.295
Mali 1.255, 2.137, 2.240–241, 3.135–141, 3.312, 4.279
Mali Empire 2.137
Mali, Arabic in 3.135–141
Mālik 2.60
Malik al-ʾAšraf, al- 2.258
Mālik ibn ʾAnas 2.401, 3.111, 3.122
Mālik ibn Dīnār 3.129
Mālik ibn Kināna 3.89
Malik Sy, al-Hajj 4.180, 4182
Malik, Jamal 3.511
Malik, Sayed H.A. 4.759–760
Malikites 2.43, 3.136, 3.170
Malindi 1.662
Malinowski, Bronislaw 2.202
Māliṭī, al- 3.142
Malka, Elie 2.534
Malkīye Arabic 4.407
mallam 3.375
mallum 1.336
Malouf, Louis → Maʾlūf, al-ʾAb Lūwīs al-Yasūʿī
Malta 1.205–206, 1.365, 2.660, 2.673, 3.141–145, 4.215
Malta, Arabic in 2.313–314, 3.672, 3.675, 3.737, 4.550
Maltam 1.334
Maltese 1.79, 1.84, 1.105, 1.183–186, 1.416, 1.487, 1.489, 1.537, 1.539, 1.609, 2.231, 2.290, 2.388, 3.65, 3.67, 3.69–70, 3.141–159, 3.189, 3.204, 3.237, 3.298, 3.312, 3.352, 3.545, 3.569, 3.739–740, 4.215, 4.217–218, 4.304, 4.565, 4.662
Maltese in Canada 3.398
Maltese, purism in 2.674
Maʾlūf, al-ʾAb Lūwīs al-Yasūʿī 2.336, 3.39
Maʾlūf, Fawzī al- 3.4
Maʾlūf, Michel 3.4
maʿlūl 2.308
Maʿlūla 1.178, 1.181, 3.370, 4.402
maʿlūm 2.14, 2.93
malzūm 3.120
mamdūd 2.234
Mamluk Dynasty 1.174, 4.589
Mamluk Dynasty, Baḥri 4.589
Mamluk literature 4.132
mamnūʿ min aṣ-ṣarf 1.309, 4.209, 4.432
maʿmūl 1.67, 2.169–170, 2.536, 4.739
Maʾmūn, Caliph al- 1.403, 4.542
Mamvu 1.519
man 1.309, 2.143, 2.387–388, 2.479, 4.39, 4.61, 4.64
maʿnā 1.428, 2.88, 2.290, 2.295, 2.304, 2.403–404, 2.426–427, 2.443, 2.579, 2.623, 3.38, 3.45, 3.116, 3.119, 3.159–164, 3.538, 3.549, 4.202–203, 4.205, 4.236–237, 4.287, 4.329, 4.447–448, 4.473, 4.475, 4.684, 4.739
maʿnā baʿīd 3.538
maʿnā qarīb 3.538
maʿnā l-kalām 2.538
maʿnā nafsī 1.429
Manāma 1.241–242
Manāma Arabic 4.325, 4.631
maʿnawī 1.67, 3.234
Manāwī, Dāʾūd al- 1.598
Manāxa 1.23
manbaha 1.192
mand 3.264, 3.266
Mandaean → Mandaic
Mandaic 1.178–179, 2.67, 3.370–373, 3.409–410, 4.302

Mandaic, Modern 3.370, 3.373
Mande 1.36, 1.255–256, 4.183
Manding 3.59
Mandinka 4.179, 4.183, 4.282, 4.723
Mandisa Arabic 2.2
Mandouze, Andre 4.572
mandūm 4.200
Mandūr, Muḥammad 1.409, 3.556
Manfalūṭī, Muṣṭafā Luṭfī al- 2.337, 2.637
Manfredi, Stefano 2.518
Manga 1.362
Mangalla Arabic 2.517
Mangbetu 1.519
Mangkunagera IV 2.336
manhiyy 4.414
manhiyy ʿanhu 4.414
manḥūt 1.452–453
māniʿ aṣ-ṣarf 1.643
Mann, Traugott 3.717
Mannāʿ, Hāšim Ṣāliḥ 2.264
Mannan, Abdul 1.287
Manning, Christopher D. 1.456, 1.511, 4.648
Mansoor, Menahem 1.668
Mansour, Jacob 1.231–241, 2.415, 2.534, 4.664
Mansouri, Fethi 1.485, 4.143
manṣūb 1.71, 1.196, 2.92, 2.225, 2.356, 3.101, 3.269, 3.429, 4.131, 4.639
Mansur, Paulo Jorge 3.2
Manṣūra 2.4, 2.325, 3.506
Manṣūrī, ʿAlī ibn Sulaymān al- 1.545
manṭaqa al-wusṭā 1.655
manṭiq 3.123, 3.136, 4.459, 4.545
manṭiqiyyūn 3.177
Mantran, Robert 3.53, 4.548
manṯūr 3.340, 4.200
Manūnī, Muḥammad al- 3.112
manuscripts, Arabic 2.636
manuscripts, Berber 3.171
manuscripts, editing of 2.637–640
manuscripts, West African 3.136
manuscripts in Mauritania 3.172
manʿūt 1.49, 3.422
Manzala Arabic 2.2
Manzala, il 2.4
manzil 2.598, 4.158
manzila 2.82, 2.184, 2.356, 2.537, 4.12, 4.738
Manzini, Maria Rita 2.367
Mao, Chairman 1.35, 1.382
Maore 1.662
map, lexical 1.590
map, perceptive 1.586
Mappiḷappaṭṭu 3.131
Mappilas 2.325, 3.129
maqāl al-qawl 2.331
maqām 3.104, 4.331–332
maqāma 2.446, 4.105, 4.210
maqāmat al-kalām 4.332
Maqdisī, al- 4.132
Maqdoom, Sheikh Zainuddin 3.129
Maʿqil 1.609, 2.240, 3.53, 3.170
Maʿqilī Arabic 1.55, 1.609, 2.598, 2.600
maqlūb 2.497, 3.539
Maqqarī, al- 1.97
Maqrīẓī, al- 2.41, 3.6–7, 3.380
maqṣadiyya 1.429
maqṭaʿ 4.103
maqʿūd 4.2
maʿqūs 4.134
marabout 3.139–140, 3.174, 4.180–181, 4.183
Marantz, Alec 3.301, 4.488
Maras 2.220
Marāšī, al- 2.257
Marazīg Arabic 1.264
Marçais, Philippe 1.55, 1.189, 1.571, 1.605–606, 1.642, 2.622, 3.53, 3.55, 3.587, 3.655, 3.672, 3.737, 4.70, 4.371, 4.519, 4.572, 4.674, 4.681
Marçais, Pierre 4.669
Marçais, William 1.173, 1.642, 2.534, 3.287, 3.645, 4.70, 4.563, 4.572, 4.576
Mardin 1.23–25, 1.87–95, 2.607, 4.579
Mardin Arabic 1.607, 2.608, 3.84, 3.237, 3.448, 4.70, 4.664
Mardyckers 4.290

marfūʿ 1.71, 1.196, 2.92, 2.356, 3.429, 4.639
marfūw 3.539
marginal element 4.387
Margoliouth, David Samuel 3.718–721
Marʾi, Abd Al-Rahman 2.104–105, 2.438, 2.464
Mariana, Juan de 1.96
Maʿrib Arabic 4.751
Maʿrib Dam 4.125
Marīdī Arabic 3.59, 642
maʿrifa 3.115, 3.453, 4.442
Marin-Guzman, Roberto 3.395
Marius, Sala 4.156
Marka 4.275
marked/unmarked 1.194
markedness 2.276–277, 2.279, 3.11, 3.191, 3.492–493, 4.638
Markedness Theory 4.442–443, 4.445
markedness, dialectal 1.408
markedness, loss of 2.149
markedness, stylistic 1.408
marker, accusative 3.674
marker, anterior 2.595
marker, aspect 2.209, 4.114, 4.162
marker, continuous 2.524
marker, definite 3.637
marker, definiteness 1.488, 1.491
marker, direct object 3.703, 4.621
marker, discourse 1.471, 1.474, 2.209, 2.668, 4.192
marker, dual 2.67
marker, emphasis 3.589
marker, end-of-paragraph 3.741
marker, existential 2.597
marker, focus 2.215, 2.592, 2.595, 3.357
marker, future 1.186, 2.192, 2.423, 2.514, 2.516, 2.595, 3.237, 3.406, 4.555
marker, genitive → genitive marker
marker, identity 1.525
marker, indefinite 1.491, 1.606, 1.611, 2.419, 2.422, 3.440
marker, locative 3.83–84
marker, mode 2.577
marker, mood 1.473, 1.668, 4.659
marker, negation 3.355, 3.638, 4.714
marker, object 3.591
marker, plural 3.636
marker, present tense 2.423, 3.292
marker, preverbal 3.643
marker, progressive 2.595
marker, question 3.585–586, 4.714
marker, relative 4.622, 4.746
marker, Tense-Mood-Aspect 3.635
marker, theme 2.149
marker, time-change 3.665
marker, topicalization 4.60
marker, vocative 2.585, 4.466
marker of indetermination 4.444
marker of specificity 4.621
market 1.650, 1.662
marking, negation 4.649
Markov models 1.456
Marková, Ivana 4.322
Marley, Dawn 2.122–123, 2.659–660, 2.694–698, 2.703–704
Maronite College 1.171
Maronites 1.171, 1.175, 1.536, 1.602, 2.59–61, 3.2, 3.396
Maróth, Miklós 3.113–116, 3.176–178, 4.11–14
Maroua 1.336–337
Marracci, Lodovico 1.172
Marrakech → Marrakesh
Marrakesh 2.43–44, 2.470, 4.157
Marrakesh Arabic 1.609, 2.621–622, 3.274–276, 3.288–289, 3.291–292
marshland dialects 2.571–572
marsiyā 4.596
Marslen-Wilson, William D. 1.425, 3.305
martaba 1.567
Martin, James H. 1.456, 1.511
Martin, Richard C. 2.303
Martinet, André 2.495, 3.603, 4.1
Martin-Granel, Nicolas 3.172
*mar*t*iya* 4.104
Marxah Arabic 4.756
Marzubānī, ʾAbū ʿUbaydallāh Muḥammad ibn ʿImrān al- 3.648
Masa 3.634
masadd 4.739
maṣāḥif 3.340, 4.131, 4.133
maṣāḥif → *maṣḥaf*
Masāʿīd Arabic 3.401, 4.238

masʾala zunbūriyya 1193, 2.632
Masalit 4.375–376
Masana 3.634
Mascara 1.54, 1.56
Mascitelli, Danieli 3.469, 3.472
masculine 2.155–164
masculine, precedence of 2.643
maṣdar 1.14, 1.68, 1.352, 1.436, 1.573–574, 1.577, 2.90, 2.92, 2.94–95, 2.165, 2.297, 2.351–352, 2.424, 2.430, 2.433, 2.448, 3.18, 3.101, 3.108, 3.164–169, 3.191, 3.255, 3.423–425, 3.428, 3.432, 3.455, 3.457–459, 3.497–498, 3.548, 4.52, 4.540, 4.644, 4.650, 4.659–665, 4.685–686
maṣdar mīmī 1.574, 1.577, 4.40
maṣdariyya 4.26
Mašhadī, Sultān ʿAlī 3.337
maṣḥaf 3.307–309
Mashlahi 1.655
Mashūr, al- 2.598
Masira 1.131, 3.478
Maskan 4.301
Maslam 1.334
Mas-Latrie, M.L. de 1.536
Masliyah, Sadok 1.639, 3.680, 4.592
Maslouh, Saad [Saʿd Maslūḥ] 4.473
masnavī 4.596
mašq 1.151–152, 2.598, 3.307
maṣri 2.11
mašriq 2.659
mašriqī 1.605, 3.112, 4.561
maṣrūf 4.432
mass media 2.518
Massakory 1.362
Massawa 1.654, 1.656, 2.268
Massenbach, Gertrud von 3.435
Massenya 1.362
Massialas, Byron G. 2.104–105, 2110
Massignon, Louis 1.445, 4.75
Masʿūd, Jibrān 3.41
Masʿūdī, ʾAbū l-Ḥasanʿ Alī ibn al-Ḥusayn al- 3.7
matā 1.509, 4.39
maṯal 2.442, 2.444, 3.89, 3.174, 3.629, 3.731
Matam 4.723
Maṭāʿna 2.2
Maṭar, ʿAbd al-ʿAzīz A. 2.609
Maṭarī, Muḥammad al- 4.107
Maṭariyya, il- 2.4
matched-guise technique 2.652–653, 2.655–656, 2.658, 3.318, 4.324
matching analysis of relative clause 4.67
Mathesius, Vilém 4.502
Matheson, Virginia 2.341
Matisoff, James A. 2.191
maṭlaʿ 4.103–104
Maṭlabī, Ġālib Fāḍil al- 3.94
maṭlūb 2.360
Maṭlūb, ʾAḥmad 2.70
maṭmūs 4.133, 4.560
matn 1.3, 3.309
matn al-luġa 3.43
matoki 2.1, 3.60
Matras, Yaron 2.216–222
matrix 1.409, 1.419, 2.77, 3.46–52, 4.97, 4.288
Matrix and Etymon Model 3.45–52, 4.97, 4.288
matrix clause 1.311, 1.468
Matrix Frame Theory 1.484
matrix language 1.416–417, 2.661
Matrix Language Frame Model 1.416–418, 1.634–635
matrix verb 1.412
matrix, implicational 2.319, 2.321, 2.324
Matsuura, Kochiro 2.704
Mattā ibn Yūnus, ʾAbū Bišr 2.186, 3.177, 4.542
Matthews, Peter H. 1.125, 3.80, 3.224, 4.339
Matthews, Roderic D. 2.104
Matthiessen, Christian 4.494
Maʿtūq, ʾAḥmad Muḥammad al- 3.41–42, 3.44
maʾṯūr, *qawl* 3.629
Mauretania Tingitana 3.287
Mauritania 2.137, 2.240, 2.363, 3.60, 3.169–176, 3.314
Mauritania, Arabic in 1.260, 1.624, 1.642, 3.175, 3.702
Mauritanian Arabic 1.52, 3.702, 4.159

Mauss, Marcel 2.202
Mavroudi, Maria 3.472
mawālī → *mawlā*
mawḍiʿ 2.184, 2.308–309, 2.357, 2.479, 2.538, 3.107, 3.124, 4.410, 4.448
mawḍūʿ 2.434, 3.113, 3.119, 3.176–178, 3.324, 4.332
Mawdūdī, ʾAbū l-ʿAlāʾ 3.346
Mawhoub, Mourad 3.727
mawlā 1.128, 1.402, 3.59
mawlid 2.53
mawṣilī 2.45
Mawṣilī, ʾIlyās al- 3.1
mawṣūf 2.435, 3.113, 3.178, 3.422, 4.219, 4.221
mawṣūl 1.488, 2.537, 4.221, 4.235–236
maxims, cooperative 3.681
Maxims, Grice's Cooperative 3.678
maxim of manner 2.70
maxim of quality 2.70
maxim of quantity 2.19
maxim of relevance 2.19
maximal projection 2.171
maximality 4.742
maxims of conversation 2.70
Maxmutov, M.I. 4.450
maxraj 2.233, 2.237, 3.73, 3.125, 4.426
Maxwell, Edith M. 2.688
Maxzūmī Dynasty 2.52, 4.276
Maxzūmī, Mahdī al- 1.40, 4.12
May, Darlene R. 3.611, 4.387–388
May, Robert 2.134, 2.368, 4.15
Mayan 3.204
Maydan, mosque of 2.46
Maydānī, ʾAbū l-Faḍl ʾAḥmad ibn Muḥammad al- 3.731
Mayer, Carl 4.267–268, 4.335
Mayer, Mercer 2.685
Mayo-Danay 1.334
Mayo-Kebbi 1.360, 1.364
Mayo-Kebbi area 3.634
Mayo-Sava 1.334
Mayotte 1.662
Mayyāḥ, ʿAlī M. al- 1.617
Mazār Sharīf 1.28
Maze, Camille 3.204
Mazhar, M.A. 4.759
mazīd 2.93
mazīd fīhi 4.640
Māzinī, ʾAbū ʿUṯmān Bakr ibn Muḥammad al- 1.192–193, 2.79, 2.234, 2.355, 2.453, 3.323, 3.553, 4.119–120
Maznī, ʾIbrāhīm al- 4.90
Mazouna 1.54
Mazraani, Nathalie 1.406, 1.409, 1.419, 3.663–671
Mbale 1.661–662
Mbiti, John S. 3.135
McArthur, Tom 3.193
McCall, Daniel F. 1.663
McCarthy, John Joseph 1.121, 1.457, 1.653, 2.37–38, 2.153–155, 2.229–230, 2.365, 2.367, 2.396, 2.692, 3.66, 3.204, 3.225, 3.297–307, 3.423, 3.430, 3.440, 3.445, 3.461–463, 3.492–493, 3.496, 3.617–618, 3.624–627, 3.726, 4.51, 4.93–94, 4.117, 4.269, 4.339, 4.345–346, 4.348–349, 4.388–389, 4.557–558, 4.645, 4.669
McCarthy, Richard J. 3.672, 4.62
McCarus, Ernest N. 3.203, 3.238–262, 3.573, 4.147, 4.168
McCawley, James 4.63
McEnery, Tony 1.512
McIntyre, Joseph 2.251
McLaughlin, Barry 1.368
McLaughlin, Fiona 4.179–185
McOmber, Michael L. 4.339
McWhorter, John 1.526
meaning 1.428–429, 2.304–305, 2.624, 3.159–164
meaning, conventional 2.445, 3.120
meaning, direct 2.581
meaning, indirect 2.581
meaning, induced 2.581
meaning, intended 3.120
meaning, intentional 2.80
meaning, literal 1.428, 3.119, 3.122
meaning, primordial 3.49
meaning, semantic-grammatical 2.624
meaning, tropic 2.582, 3.122
meaning, vulgar 1.657

MECAS → Middle East Centre for Arab Studies
Mecca 1.448, 1.451, 3.88, 3.179, 3.691, 3.697, 4.377
Meccan Arabic 1.300, 1.400–401, 1.501, 1.611, 1.615, 2.355, 3.90, 3.179–187, 3.335, 3.404, 3.612, 4.126, 4.534
Meccan script 1.151
mechanism of change 2.193, 4.42
mechanism of grammaticalization 2.193
Medea 1.54
Medea Arabic 1.609
medersa 3.174
media, Arabic 3.397
media, Chadian 1.364
media, news 2.381, 3.199, 3.201
Media Arabic 1.474, 1.629, 1.667, 2.9, 2.212, 2.214, 2.296, 3.10–14, 3.17, 3.19, 3.192–204
media language 1.419, 1.474
medial 3.596
Medial Gemination Rule 2.154
medicine, Greek 4.542
Medina 1.128, 1.451, 3.179, 3.691, 3.694, 4.127
Medina, Barak 2.438
mediopassive 1.39, 3.553, 3.557, 4.697
Mediterranean Sea 3.60
Mednine 2.709
medrese 3.503, 4.578
Megiser, H. 3.141, 3.146
Mehall, David John 1.668
Mehdi, Saad A. 1.457
Méhiri, Abdelkader 2.233–234, 2.299, 2.309–310, 2.538, 2.544, 4.287, 4.447, 4.684
Mehlem, Ulrich 2.74–75
Mehler, Jacques 3.725
Meḥmed, Yazıjıoğlı 3.502
Mehmet Izzet Efendi 4.98
Mehren, A.F.M. von 3.116, 3.120
Mehri 1.39, 1.203, 1.493, 2.66–67, 2.211, 3.63, 3.350, 3.478, 3.593, 3.693, 4.175, 4.297–305, 4.307, 4.310, 4.312
Meier, Fritz 2.408
Meiering, Gregor 2.700
Meillet, Antoine 1.614, 2.191
Meinhof, Carl 2.250
Meisel, Jürgen M. 1.368
Meiseles, Gustav 1.410, 1.629–633, 1.668, 3.567, 4.325, 4.664
Mejdell, Gunvor 1.406, 1.414–421, 4.71
Mejri, Salah 1.584, 4.575
Meknès 3.294
Meknès Arabic 1.609, 2.621, 3.85, 3.291
mekteb 3.503
Melbourne 3.313
melḥoun → *malḥūn*
Melkites 1.403
melody 2.397, 2.628
melody, falling 2.397
memorization 1.336, 2.110, 2.327, 3.170, 4.184
memorization of the *Qurʾān* 2.110, 2.338, 3.57, 3.136
Menassa, Laila 1.503
Mende 3.461
Menem, President Carlos 3.3
Menéndez Pidal, Ramón 1.98
Menovščikov, G.A. 2.671
mental lexicon 1.425
mental map 1.585
mental process 1.421, 1.425
mental representation 1.369, 1.425
mental representation of root 4.95–96
Merah Silau 2.334
Merchant, Jason 4.193
Mercier, Henry 3.287
Meredith, David 3.472
Merge 3.229
merger 1.46, 1.103, 1.384–385, 1.481, 1.490–491, 1.493, 1.609, 2.7, 2.242, 2.688–689, 3.63, 3.147, 3.232, 4.371
merger of /ḍ/ and /ḏ̣/ 1.544
merger of /i/ and /u/ 4.404
merger, consonant 1.539
merger, phonological 1.590, 3.190, 3.603–605
merger, sibilant 2.533–534
merging 3.228
merging, portmanteau 4.463

Meringer, Rudolf 4.267–268, 4.335
merismus 3.633
Mernissi, Fatima 2.642, 2.645, 2.647
Meroitic 1.36
Meron, Yoram 4.663
Mersin 1.388
Mersin Arabic 1.390, 1.392, 1.394–395, 2.313
Merx, Adalbert 2.401, 3.106, 3.547
Mešaʿ 3.409
Mesgnien-Meninski, François de 4.516
Mesopotamia 2.212
Mesopotamian Arabic 1.23, 1.87, 1.189, 1.222.1.236, 1.571, 1.607, 1.606
Messaoudi, Leila 1.596, 2.642, 3.236
Messelink, J. 2.683
Mester, R. Armin 2.36, 3.462, 4.374
metadiscourse 1.649
metalanguage 1.487, 3.550
metalinguistic comment 2.653
metaphor 1.422–423, 1.425, 1.429, 1.649, 1.658, 2.193, 2.196, 2.305, 2.441–446, 2.578, 2.638, 2.641, 3.35, 3.116–122, 3.321, 3.629, 3.632, 3.651, 3.733, 4.165–166, 4.210, 4.213, 4.464
metaphor, conceptual 1.424
metaphor, ontological 1.424
metaphor, orientational 1.423
metaphor, structural 1.423–424
metathesis 1.104, 1.192, 1.225, 1.245, 1.256, 1.262, 1.314, 1.377, 1.391, 1.425–426, 1.548, 1.668, 2.103, 2.140, 2.242, 2.280, 2.366, 2.483, 2.612, 2.678–679, 3.15, 3.45, 3.49, 3.148, 3.204–207, 3.289, 3.304, 3.361, 3.481, 3.530, 3.566, 4.96–97, 4.117, 4.244, 4.286, 4.379, 4.549, 4.591
Metcalf, Barbara Daly 2.326
Metcalfe, Alex 4.215–219
meter 1.601, 3.173, 3.207–215, 3.568, 3.647, 4.86, 4.92, 4.103, 4.201, 4.206, 4.212
meter, *rubāʿī* 4.206–207
meter in *rajaz* 4.34
meter Persian 4.586
Methodikoi 4.11
metonymy 1.425, 2.193, 2.196, 2.305, 2.443, 2.578, 3.117, 3.121, 3.187, 3.321, 3.537, 3.632, 3.733, 4.165–166, 4.335, 4.464
metrical constituency 3.621
metrical parsing 3.492
metrical theory 3.615–616
metrics 2.184
Metz, Helen Chapin 4.579
Meulen, Daniel van der 2.341
Mexico 3.1, 3.3, 3.5
Mexico, Arabic in 3.395–396
Mey, Jacob L. 3.677
Meyer-Bahlburg, Hilke 2.251
Mez, A. 1.577
Mezzine, Larbi 1.298
mġannyīn 3.173
Mḥallamī 1.87–88, 1.91–92, 1.95
Mḥallamīye Arabic 3.237
Mhamid Arabic 3.288
Mḥardi Arabic 4.403–404
Mhīn Fruqlus Arabic 4.406
micrography 4.134
microvariation 3.233
Mīdān Arabic 1.547
Middle Arabic 1.57, 1.84, 1.383, 1.403, 1.410, 1.496, 1.544, 1.593–594, 1.597, 1.601, 1.613, 1.619, 1.631, 1.642, 1.668, 2.265–266, 2.312–313, 2.332, 2.388, 2.529, 3.215–224, 3.466, 3.704, 4.196, 4.543–544, 4.546
Middle Arabic, Christian 1.383, 1.386
Middle Arabic, Jewish → Judaeo-Arabic
Middle East Centre for Arab Studies (MECAS) 4.148
Middle Egypt 2.2
Middle Persian → Pahlavi
middle verb 2.566, 3.221, 3.224–227, 4.307
middle voice 3.446, 3.553, 4.642
middle voice, acquisition of 2.102
midground 1.649, 2.209
Midob 3.435

Midrashim 2.527
Mifsud, Manwel 3.144, 3.146–159, 3.545
Migama 1.36
Migdadi, Fathi H. 2.501, 3.681
migrants 2.73, 2.75–76, 2.97, 2.682–684
migrants, Arab 2.336
migrants, foreign 3.63
migration 1.483, 1.593, 1.595, 1.609, 1.611, 1.617, 2.72, 2.76, 2.212–213, 2.250, 2.337, 2.499, 2.506–507, 2.712
migration, Arab to Central Asia 4.612
migration, Islamic 3.125
migration, Jewish 2.532
migration from Yemen to Egypt 4.750
migration from Yemen to Israel 4.750
migration in Saudi Arabia 4.125
migration of speakers 1.586, 1.589
migration to Cairo 4.727
miḥrāb 2.45
Mikhri 4.604
Miklosich, Franz von 1.622, 1.625
Mila 1.53–54, 1.56
Mila Arabic 1.56
Mīlād, Xālid 2.359
Milani, Celestina 2.456
Milelli, Jean-Pierre 4.78
Miles, George 3.7
Miles, Tim R. 2.689
Miliana 1.54
Miliana Arabic 1.609
Milik, Joseph T. 3.472
Militarev, Alexander 2.67, 4.295, 4.302–303, 4.310, 4.313
Millar, Fergus G.B. 3.467
Miller, Ann 1.491, 1.593–594
Miller, Catherine 1.269, 1.271, 1.521, 1.525, 1.572, 1.593–597, 2.1, 2.274, 2.517–525, 2.642, 3.58–64, 3.643, 4.198, 4.634
Miller, Philip H. 4.647
Miller, R.E. 2.325
Millet-Gérard, Dominique 1.620
Milroy, James 2.655, 3.605, 4.631
Milroy, Lesley 2.655, 4.629–631, 4.634
mimation 1.491, 1.560, 3.191, 3.227–228, 3.414, 3.453
mimetic 2.619, 4.286
Mimi 4.710
mīmī, maṣdar 1.574, 1.577, 4.40
mimophony 3.47–48
Mimouni, Zohra 1.425, 2.691–692
min 1.72, 1.625, 2.15, 2.144, 2.147, 2.195, 2.295
Minaean 4.295, 4.301, 4.313, 4.478
Minangkabau 2.337
minbar 2.40, 2.43–44, 2.46
Mines, M. 2.325
Ming Dynasty 1.379
Mingshale 1.380
minimal pair 3.604, 3.607
minimal projection 1.351
Minimal Tree Hypothesis 4.143
Minimalist Program 1.348–349, 2.131, 3.228–233, 3.684, 4.489, 4.522, 4.647–650, 4.657, 4.742
minimality 1.47, 3.229, 4.392, 4.399
Minimality Condition 2.171, 2.174, 4.363
minimality effects 2.171
MINMWMAX algorithm 1.586
Minnis, Denise D. 1.372
minority 1.444, 2.439, 2.441
minority, Arab 2.72
minority, linguistic 2.217
minority group 2.58, 2.61, 2.63–64, 2.204
minority language 2.414, 2.437, 2.684, 2.686, 3.314, 4.577
minority language, Arabic in Turkey 4.578
minqār 3.341
Minsky, Marvin Lee 1.458
Minufiyya Arabic 2.2
Minya Arabic 1322, 2.2, 2.5
Mīr Sayyid Aḥmad 2.598
Mir, Mustansir 2.304–305
Miramar 1.171
Mirkin 2.462
mirqam 4.132

Mischlich, Adam 2.251
Misfalawiyyah 4.159
Mishkurov, Edward N. 4.306
Mishnah 2.527
Mishnaic Hebrew 1.85, 1.493, 2.66, 3.408, 3.414
Misirīye humur 1.362
Misirīye zurug 1.362
Miskīn, Muḥammad 1.603
mismatch, phonological 2.31
Misnāwī, ʾIdrīs al- 1.603
miş-perfect in Turkish 1.396
miṣr 1.129
miṣrāʿ 3.208
misreading 3.78
Miṣrī, ʿAbd ar-Rāziq ʿAwaḍ al- 4.98
Miṣrī, Ḥusayn Šafīq al- 1.599
Miṣrī, Muḥammad 3.511
Missing Surface Inflection Hypothesis 4.144
missionary school 2.8, 2.699
misspelling 3.78
mistake 3.200
Misurata 3.53
Misurata Arabic 3.54–56
miṯāl 2.90, 2.449, 4.13, 4.121, 4.645
Mitchell, Terence F. 1.502, 1.630–632, 1.667–668, 2.96, 2.151, 2.228, 2.501, 3.18, 3.263–264, 3.495, 3.525, 3.544, 3.594–595, 3.609, 3.611, 3.613, 3.654, 3.705, 4.62, 4.99–100, 4.148–149, 4.317, 4.325, 4.344, 4.346, 4.351, 4.650, 4.668
miṯl 2.302
Mitwalli, M. 1.589
mixed form 1.633
mixed variety 1.669
Mixe-Zoque 3.204
mizbar 4.132
Mlaḥsō 1.541, 3.370
Mleiha (Mulayḥa) 3.472
mnemonic phrase 4.121
mnemotechnic 4.105
Moab 3.467, 4.480
Moabite 3.409, 3.414
moagaritai 1.129
Moatassime, Ahmed 3.347
mobile phone 3.199, 3.274
mobility, population 4.634
mobility, upward 4.629
mobilization 2.58
Mobius, Bernd 4.388
modal 1.509, 3.260
modal adverb 1.22
modal auxiliary 3.268
modal expression 1.219
modal verb 1.221, 1.266, 3.233–238, 4.365
modality 1.510, 2.23, 3.234
modality, emotional 2.144
modality, reinforcement 2.144
modality, volitional 2.144
Modarressi, Hossein 3.117, 3.120
mode marker 2.577
Modern Standard Arabic (MSA) 1.399, 1.406–410, 1.423, 1.434, 1.436, 1.441, 1.446, 1.456, 1.469, 1.472–473, 1.476, 1.490, 1.493, 1.528, 1.532, 1.578, 1.580–582, 1.595, 1.624, 1.630, 1.633–634, 1.636, 1.647, 1.650, 1.652, 1.655–656, 1.667, 1.669–670, 2.267, 3.10–12, 3.14, 3.18–19, 3.76, 3.202, 3.238–262
modernity 1.531
modernization 1.532, 1.636, 2.655, 2.700
modernization of Arabic 1.531, 2.698, 2.701
modibbo 3.375
modification, phonological 2.374
modifier 1.457, 2.290, 3.431, 4.221, 4.317, 4.392
modifier, adjectival 1.474
modifier, appositive 2.370
modifier, left-peripheral phrasal 4.746
modifier, noun 1.579
modifier, specified 3.688
modularity of cognitive systems 2.691
Mogadishu 4.275–276
Mogador Arabic 1.609
Mohamed VI, King 2.63, 3.317
Mohamed, Ibrahim 2.100, 3.357
Mohamed, L. 3.653
Mohamed, Satifa 1.337

Mohammad, Mohammad A. 1.45, 1.312, 2.133, 2.171–172, 2.369, 3.81–84, 3.232, 3.382–383, 3.385, 3.429, 3.432–433, 3.560, 3.683, 3.685–687, 3.709–710, 3.713, 4.15, 4.316–317, 4.359, 4.362–365, 4.393–396, 4.400, 4.524, 4.648, 4.654–655, 4.657, 4.726, 4.745–747
Mohammed Koya, C.H. 3.129
Mohéli [Mwali] 1.450
Moïnfar, Mohammad Djafar 2.407, 3.575
Mokhlis, Hassan 2.165, 2.448–449, 4.120
Molan, Peter D. 2.630–631
Molière 1.600
Mombasa 1.662
Mommersteeg, Geert 3.136–139
Monahan, Philip J. 4.676–677
Monastir Arabic 4.564
Mongolia 2.557
Mongolian 1.29, 1.378
Mongols 1.174, 1.379, 1.403, 1.589, 2.414, 2.527
Mon-Khmer 1.8
monoconsonantal 4.93
monogenesis 3.640, 3.647
monolingual 1.369, 1.371–372, 1.416, 1.483
monolingualism 1.629, 3.315
monolingualism, Amazigh 2.710
monolingualism in Nubia 3.425
monolingualism in Turkey 4.579
monologue 1.409, 2.207, 4.727
monologue, formal 1.406
monomoraic 3.612–613
monophthong 3.405
monophthongization 1.103, 1.235, 1.261, 1.300, 1.383, 1.390, 1.410, 1.492, 1.547, 1.608, 1.640–642, 2.416–417, 2.509, 3.54, 3.362, 4.406
monorhyme 4.87, 4.89
monotransitive 3.224, 4.537–538
Monroe, James 3.500
Montague, Richard 1.457, 4.15
Monteil, Vincent 3.168, 3.201, 3.426, 4.460, 4.558
months, names of the 1.3, 2.258, 2.289
mood 1.216–217, 1.398, 1.508, 1.510, 2.92, 3.98, 3.233, 3.248, 3.262–273, 3.298
mood ending 1.386, 1.614
mood marker 1.475, 1.668, 4.569
mood of speech 4.332
moodibo 1.336
MoodP → phrase, Mood
moon letter 1.187, 1.204
Moors 2.240, 3.174
Moosa, Ebrahim 4.294
Moosally, Michele Jamila 4.657
Moplah 1.133
Mora 1.334, 1.337, 2.154, 2.238, 3.612
Morag, Shelomo 2.534
Moravcsik, Edith 4.51
Moreh, Shmuel 1.598
Morin, Didier 2.269–270
Moriscos 1.96, 3.294, 4.572, 4.574
Moritz, Bernard 3.515
Moro 2.525
Moro, Andrea 4.746
Moroccan Arabic 1.27, 1.51, 1.55, 1.77–81, 1.84–186, 1.220–221, 1.294–298, 1.351, 1.407–408, 1.418–419, 1.479, 1.481, 1.483, 1.486, 1.488, 1.502, 1.512, 1.568, 1.571–572, 1.585, 1.604, 2.21, 2.71–72, 2.75–76, 2.96–97, 2.102, 2.116, 2.156, 2.159, 2.161, 2.174, 2.194, 2.205, 2.231, 2.374, 2.388, 2.390–391, 2.393, 2.395–396, 2.621, 2.658, 2.686, 3.48, 3.68–70, 3.84, 3.237, 3.273, 3.287, 3.305, 3.358–359, 3.382, 3.429, 3.446, 3.554, 3.584–587, 3.592, 3.611, 3.655–656, 3.684, 3.705, 3.707, 3.724–726, 3.740, 4.45, 4.51–52, 4.67, 4.71, 4.117, 4.141, 4.159–160, 4.167, 4.185–191, 4.286, 4.306, 4.317–318, 4.339, 4.351, 4.393, 4.397–399, 4.491, 4.646, 4.648–649, 4.651, 4.654–4.656, 4.662, 4.666, 4.668, 4.672, 4.674–675, 4.678, 4.682
Morocco 1.23–24, 1.293, 1.483, 1.533–534, 1.597, 2.63–64, 2.122,

2.240, 2.471, 2.648, 2.698, 2.704, 2.707–710, 2.712–715, 3.53, 3.60, 3.287–297, 3.316–317
Morocco, Arabic in 1.260, 1.440, 1.584, 1.592, 1.626, 1.642, 3.15–19, 3.452, 3.672, 3.674, 3.702, 3.737, 4.62
Morocco, literacy in 3.78
Moron 1.99
moros latinados 1.98
morpheme 1.573, 2.237, 2.546–548
morpheme, content 1.416
morpheme, function 1.416
morpheme, grammatical 1.416, 1.573
morpheme, interphrasal 4.142
morpheme, lexical 1.416, 1.573
morpheme boundary 4.96
morpheme loss 2.192
morpheme order, acquisition of 4.140
morphological generator 1.460
morphology 1.575, 3.297–307
morphology, autosegmental 4.93
morphology, camouflage 2.218
morphology, derivational 3.297, 3.300
morphology, inflectional 3.297, 3.300
morphology, nonconcatenative → nonconcatenative morphology
morphology, prosodic 3.618
morphology, root-and-pattern 3.301, 3.305
morphology, root-based 1.576, 3.149
morphology, simultaneous 4.228–229
morphology, stem-based 3.149, 3.156, 4.338–344
morphology, templatic 3.301
morphology, two-level 1.457
morphophonology 2.180
Morsi, Ranya 1.376
Mörth, Karlheinz 3.448
Moru-Madi 1.519
Morzouk 3.56
Moscati, Sabatino 1.638–639, 1.643, 2.156, 3.227, 3.440, 3.446, 3.525, 4.572
Moscoso, Francisco 3.290, 4.70
Moše ibn ʿEzra 1.99
Mosel, Ulrike 2.294, 2.355, 2.479, 2.541, 3.100–101, 3.103–108, 3.164–165, 4.102, 4.220–221, 4.236, 4.414, 4.431
Moser, Hugo 4.72
mosque 2.40
mosque education 1.380–381
mosque of al-Aqmar 2.41, 2.43, 2.600
mosque of Cheraman 3.129
mosque of Gawhar Sad 2.601
mosque of Huai-sheng 1.379
mosque of Ibn Tulun 2.42
mosque of Kutubiyya 2.43–44
mosque of Maydān 2.46
mosque of Qingjing 1.379
mosque of Shengyu 1.379
Mostaganem 1.54
Mostari, Hind A. 2.694, 2.697–698
Mostowski, Andrzej 4.15
Mosul 1.178, 1.181, 2.45, 2.414, 3.370
Mosul Arabic 1.417, 1.537, 1.610, 2.312, 2.415–417, 2.419–420, 2.423, 2.534, 3.448, 3.674, 4.372
mother tongue 2.645
motherese → caretaker talk 1.339–340
motion verb 1.52, 1.268
motivation 1.485
motivation, language learning 2.651
motivation, learning 1.483
motivation, social 1.635
Motley, Michael 4.270, 4.335
Mouchaweh, Lina 4.395
Mouchaweh, Lubana 1.508
Mouhssine, Ouafae 2.660
Moujib, Ilham Dupont 4.95
Mount Hermon Arabic 4.407
Mount Lebanon 2.60
Mount Lebanon Arabic 1.269–271
Mouradian, Jacques 4.182
Moussa, Ali Helmy 1.458, 1.512
Moussaoui, Abderrahmane 2.363
Moustafa, Ahmad 4.138
Moutaouakil, Ahmed 1.457, 2.114, 2.143–150, 3.356–357, 4.329–331, 4.398, 4.472, 4.656, 4.732, 4.734
movement 3.617, 4.82, 4.392, 4.488, 4.524, 4.657
movement, across-the-board 4.649

movement blocking 2.370
movement chains 2.370
movement site 2.171
Mowrer, D.E. 2.98
Mozaʿ 1.655
Mozambique 1.661–662
Mozambique Channel 1.448
Mozarabic 1.98, 2.287
Mozarabs 1.96–100, 1.167, 1.169, 1.620, 2.61, 3.4
Mozayni, Hamza Qublan al- 3.621–622
Mpade 1.334
Mrayati, Mohamad 2.395–397, 2.448, 2.450, 3.727
MSA → Modern Standard Arabic
msahafu 1.451
Mser 1.334
Muʾab 4.635
Muʾaddib, al- 4.119
muʾakkid 4.436
muʾallaf 2.538
muʿallaq 3.307, 4.560
Muʿallaqāt 4.200, 4.202, 4.210
muʿāmalāt 2.360
muʿammā 4.133, 4.157
muʾannaṯ 2.159, 3.426, 3.651
muʾannaṯ ḥaqīqī 2.157
muʾannaṯ lafḏ̣ī 2.157
muʾannaṯ majāzī 2.157
muʿāqaba 2.280
muʿaqqaf 3.308, 4.133
muʿāraḍa 1.4, 2.303
muʿarrab 1.403
muʿarraq 4.133
muʾaṯṯir 4.415, 4.626
Muʿāwiya, Caliph 2.599
muʿawwar 4.133
muʾaxxar 4.739
muʿayyan 3.549
muʿāzalat at-tarkib 4.209
mubālaġa 3.101, 3.119
Mubārak ibn Muṭlub 2.572
Mubarak, President Hosni 3.197
Mubarakpurī, Qazi Athar 3.508
Mubarrad, ʾAbū l-ʿAbbās Muḥammad ibn Yazīd al- 1.191, 1.193, 1.627, 2.178, 2.185, 2.193, 2.226, 2.291–292, 2.295, 2.311, 2.356, 2.426, 2.433, 2.453, 2.536–538, 2.544–545, 2.547, 2.551–552, 2.579, 3.34, 3.100, 3.102, 3.104–105, 3.107–108, 3.124, 3.321, 3.323, 3.550, 3.731, 4.9, 4.411–412, 4.414–415, 4.119, 4.282, 4.439, 4.413, 4.436, 4.444, 4.447, 4.533
Mubarrid, al- → Mubarrad, al-
mubayyin 4.436
mubdal minhu 1.123
mubham 1.73, 1.558, 2.316, 3.109, 4.221, 4.740
Mubi 1.36
mubīn 1.174, 3.346
mubtadaʾ 1.41, 1.49–50, 1.68–69, 1.182, 1.436, 1.557–558, 2.132, 2.290–292, 2.354, 2.425, 2.430, 2.434–436, 2.537–539, 2.542, 2.550, 2.552, 2.562, 2.627, 3.115, 3.177–178, 3.231, 3.234–235, 3.353–354, 3.381, 3.557, 3.683, 4.56, 4.317, 4.354, 4.429, 4.484–486, 4.525, 4.653, 4.737–740
mubtadaʾ bihi 4.495
mubtadaʾ muʾaxxar 2.132, 4.354
muḏ 1.72, 1.557
muḍāʿaf 4.121, 4.642, 4.645
muḍāf 1.557, 2.294, 2.296, 2.425, 2.437
muḍāf ʾilayhi 2.294–295, 2.425, 2.437, 3.429, 4.439, 4.444
muḏakkar 3.426, 3.651
Muḍar 4.304
muḍāraʿa 2.92, 2.239, 2.403, 2.425, 2.625, 3.98
muḍāriʿ 1.196, 1.309, 1.568, 2.91, 2.426, 2.431, 2.436, 2.497, 2.542, 3.96–100, 3.213–214, 3.270, 3.539, 3.547–548, 3.551, 3.651, 4.455, 4.638
mudawwar aṣ-ṣaġīr, al- 4.560
Muḍaybi Arabic 3.481, 3.489
muḏayyal 3.539
mudġam 2.300, 3.307, 4.560
mudġam ʾilayhi 2.300
muḏ̣har 2.301–302, 3.307
Mudia Arabic 4.754

muḍmar 2.301–302, 3.422, 4.221, 4.738
mufāḍala 2.14
mufassir 4.436
mufaxxam 2.235, 3.600, 4.421–423, 4.636
mufīd 2.538–539, 2.541, 3.121
mufrad 1.558, 2.426, 2.536–537, 2.539, 3.214, 3.439, 3.547, 3.603
Mughal Empire 4.596–597
Mughals 1.186–287, 2.325–326
Mughazy, Mustafa 2.18–21, 3.83, 3.87, 3.388–394, 3.570–573, 3.676–683, 4.60–70, 4.490
Muġrabī, Ibn Saʿīd al- 1.361
muḥaddaṯ ʿanhu 2.436–437
muḥājāt 4.157
muhājirūn 1.128–129
muḥāl 2.184
Muhalhil 3.211
Muhallabī, al- 4.279
Muḥammad, Fāyiz Yūsuf 3.39
Muḥammad, Hāšim al-Xaṭṭāṭ 4.98
Muḥammad, Manẓūr 3.343
Muḥammad, Muḥammad Saʿd 3.38, 3.42
Muḥammad, Xālid S. 3.679
Muḥammad ʿAbd ar-Raḥmān 4.98
Muḥammad ʿAlī, Khedive 1.174–175, 1.494, 1.519, 2.456, 2.699, 3.344, 3.349, 4.460
Muḥammad ibn Ḥasan ibn ʿAlī 1.3
Muḥammad ibn Mānī 3.374–375
Muḥammad ibn Qāsim 3.506
Muḥammad ʿIzzat Afandī → Mehmet Izzet Efendi 4.98
Muḥammad Muḥsin 2.598
Muḥammad wəll al-Muxtār wəll Billaʿmaš 3.172
Muhammadan Rose 2.46
Muḥammarah 2.572
muḥaqqaq 1.152, 2.598, 3.307–311, 3.340–341, 3.519, 4.131, 4.133, 4.560–562
muḥaqqiqūna, al- 1.627
muḥarraf 3.539
Muḥarraq 1.241, 2.497
muḥaṣṣal 2.427, 3.549
Muhawi, Ibrahim 2.658
Muḥaysin 2.571–572
Muḥaysin, Muḥammad Sālim 4.426
muḥdaṯ 3.214, 3.648
muḥdaṯūn 2.446
Muher 2.67, 4.301, 4.304–305
Mühlhäusler, Peter 2.117
muhmal 1.152, 3.178, 3.215
Muḥsin, Ḥikmat 1.546
muʿjam 2.629
mujānas 2.496
mujānasa 2.496
mujarrad 2.93, 3.178, 4.640
mujawhar 3.112
Mujāwir, Muḥammad Ṣalāḥ 2.104
mujāzā 2.478
muʿjiza 2.302
mujmal 2.536
mujtaṯṯ 3.213
Mukalla Arabic 4.757
Mukamatu Ārip Miyāṉ 4.433
mukassar 3.426
Mukeyras Arabic 4.751, 4.753
Mukhtar, Mohamed Haji 4.275–278
mulāʾama 1.436
mulābasa 3.118
mulaffaq 3.539
mulāḥaḏ̣a 3.118–119
mulāqin 4.412–413
mulġā 2.308, 4.410
mulḥaq 4.640
Müller, David Heinrich 2.257–258, 4.303
Müller, Walter W. 2.198, 2.258–259, 2.263, 3.227, 3.414, 3.467–468
Müller-Kessler, C. 3.410
Multan 2.325, 3.506
multidimensional scaling 1.586
Multi-Dimensional Scaling Plot 4.238
multiglossia 2.529
multilingualism 1.519, 1.629, 1.662, 2.439, 2.671, 2.695, 2.703, 2.710, 3.311–320, 3.641, 4.716
multilingualism in Djibouti/Eritrea 2.268
multilingualism in Ethiopia 2.51
multilingualism in South Arabia 4.297

multilingualism of Moroccans in the Netherlands 2.74
mumāṯil 3.539
mumayyiz 4.436
Mumbai → Bombay
mumtaniʿ 2.299, 4.25
Mumtāz Afandī 4.98
Mumtāz Bik 4.98
munādā 1.73
munaḏ̣ara 2.92
munājāt 2.329
Munajjid, al- 3.321
Munawar, Muhammad 3.508
munawwar 4.133, 4.560
munḏu 1.72, 1.309, 1.467, 2.538
Münejjimbaşı 3.502
munfaṣil 1.555, 1.557
munfatiḥ, ḥarf 2.459
Munjid 2.336
Munjid, Ṣalāḥ ad-Dīn al- 3.14
Munn, Alan 4.397
munqaḍin 3.97
munqaṭiʿ 2.293, 3.96
munṣarif 1.559, 1.643, 2.403–404, 2.424, 2.426, 3.453, 4.119, 4.432
munsariḥ 3.210, 3.212, 3.214
Munšī, Muḥammad Jamāl ad Dīn ibn Badr ad-Dīn al- 1.627
muntaḏ̣ar 3.97, 3.548
muntaṣab 4.133
Munthe, Ludvig 1.663, 3.126
munxafiḍ 4.421
Munzel, Kurt 1.83, 1.502
muqābal 1.628
muqābala 1.4, 3.453, 4.444
muqaddam 4.739
muqaddar 1.49, 4.448
Muqaddasī, al- 1.99, 1.620, 2.325
muqāraba 1.70, 3.105, 3.234
Muqātil ibn Sulaymān 2.301, 3.89, 4.235, 4.473
muqaṭṭaʿ 2.426
muqawwar 4.133, 4.560
muqawwas 4.133
muqayyad 2.404
muqīm 3.97, 3.548
muqtaḍā l-ḥāl 4.332
muqtaḍab 3.210, 3.213
muʿrab 1.309–310, 1.559, 2.92, 2.538, 3.429, 4.119, 4.443
Murābiṭūn, al- → Almoravids
murād 2.80, 2.186, 3.693, 4.331
Murad Arabic 3.94
murakkab 2.538, 2.541, 3.214, 3.539, 3.548, 3.603
murakkab mazjī 1.454
Muraoka, Takamitsu 3.410
muraqqaʿ 4.134
muraqqaʿāt 3.337
muraqqaq 2.235, 4.423
murattab 1.152, 4.133, 4.560
Murayna Arabic 3.94
Muraz, Gaston 3.61, 3.634, 3.642
Murcia 1.99
Murgida, Jacqueline 3.203
muʿrib 1.132
Murid Sufis 4.180
Murīdiyya 4.180–183
murmur 4.668
Murray, G.W. 2.2
Murry, T. 4.668
Mürseloğlu Arabic 1.392
Mursy, Ahmad A. 3.681
Murtonen, Aimo 3.440
Muş 1.87–88
Mūsā ar-Rāhib 2.599
Musa, Monsur 1.288
Mūsā, Salāma 1.602, 3.349
Mūsā, Sulaymān 2.498–499
mušabbah 2.444–445, 3.120, 4.220–221
mušabbah bihi 2.444–445, 3.120
mušaddad 3.336
muṣāḥaba 1.435
muṣaḥḥaf 3.539
muṣaḥḥaḥ 3.426
Musalmani Bangla 1.287–288
musalsal 1.151, 4.561–562
musalsalāt 4.408
musammā 2.428, 3.323–324
Musandam Peninsula 2.211, 3.478
muṣāraḥa 2.578
mušāraka 3.224, 4.626
Mušarrafa, Muṣṭafā 1.599
muṣarraḥ 2.445
musattar 4.87
muṣawwat 2.238

muṣawwit 2.234–235
Musaylima ibn Ḥabīb 2.303
mušbaʿ 2.234
Muscat 2.213, 2.215, 3.478
Musgu 2.554
Musgums 1.334
muṣḥaf 1.492, 4.4
Musil, Alois 3.402
Mušiyya Arabic 2.2
muškil al-ʾiʿrāb 4.200
Muslim 3.717, 4.79
Muslim Brotherhood 4.77
Muslim ibn al-Walīd 2.446, 4.205
Muslim names 1.286
Muslims, Chinese 1.379–382, 2.334
Muslims, Indian 2.325, 2.340, 2.344
Muslims, Indonesian 2.333
Muslims, Malay 4.290
Muslims, Tatar 2.557
Muslims, Thai 4.477
Muslims, Vietnamese 2.334
Muslims in India 3.131
Muslims in Kerala 3.129, 3.131–133
Muslims in Malabar 3.129
Muslims in Mali 3.137
musnad 1.429, 2.435–437, 3.112, 3.683, 3.685, 4.219, 4.354, 4.484, 4.686
musnad ʾilayhi 1.429, 2.425, 2.435–437, 3.683, 3.685, 4.219, 4.354, 4.484, 4.686
mustaʿār 2.442, 2.444–445
mustaʿār lahu 1.424–425, 2.441–442, 2.445, 3.120
mustaʿār minhu 1.424–425, 2.441–442, 2.444–445, 3.120
Muṣṭafā, ʾIbrāhīm 2.189
mustafʿil 4.421
mustaġnin 2.537
muṣṭalaḥ 3.122
mustaʿlin 2.312, 4.421
mustalʿin, ḥarf 4.2
muštamal 3.119, 3.215
Mustanṣir bi-llāh, Caliph al- 1.361
Mustapha, Abdel Rahman 2.560
Mustapha, Abubakar 3.374–375
mustaqbal 3.97, 3.542, 3.548
mustaqbil 2.91
mustaqill 1.49, 2.538
mustaqīm 2.184
muštaqq 2.432, 3.539
muštaqqāt 1.574
muštarak 1.626, 2.496, 3.49, 3.162, 3.320–323
mustaʿriba 1.130, 3.345
mustatar 2.301, 2.316
mustatir 2.82, 2.430–431, 3.99
mustaṯnā 1.73, 2.451, 3.104
mustaṯnā minhu 2.451
mustawfī 3.539
Mūt Arabic 2.2
mutaʿaddin 1.67, 2.316, 3.102–103, 3.116, 3.224, 4.410, 4.412–415, 4.624
mutaʿaddin, ġayr 1.67
mutaʿarriba 1.130, 3.345
muṭābaqa 3.120, 4.26, 4.332
muṭābiqa(t)-i ṣifat wa mawṣūf 2.410
mutadārik 3.210, 3.214
mutaḥarrik 2.238, 2.299, 3.208
Muṭahhar, ʿAbd al-Rahmān 4.107
mutakallim 1.555, 1.566, 2.537, 4.332
mutakallimūn 2.178, 4.12
muʿtall 2.93, 2.164, 2.309, 2.448, 3.91, 4.644
mutamakkin 1.309, 1.643, 2.425–426, 4.431–432
mutamakkin ʾamkan 4.432
mutamakkin ġayr ʾamkan 4.432
Mutanabbī, al- 2.637, 4.90, 4.200, 4.202, 4.214
mutanawwirīn, ʿāmmiyyat al- 1.405
muṯannā 3.439, 4.134
mutaqaddimūn 2.446
mutaqārib 3.213–214
muṯaqqafin, ʿāmmiyyat al- 1.405
mutarādif 3.162, 3.320–321, 3.323–325
muṭarraf 3.539
Muṭarrizī, al- 1.380, 2.189
mutašābih 3.122, 3.539
mutaṣarrif 2.14, 3.109

mutaṣayyad 4.24

muṭāwaʿa 3.224, 4.411, 4.414, 4.642

Mutawallī, ʾAḥmad Fuʾād 2.30, 4.592

mutawaqqaʿ 3.97, 3.548

mutawāṭiʾ 3.321

mutawātir 4.9

Muṭayr 3.326, 4.125

Muʿtazilites 2.303–304, 2.306, 2.310, 2.580, 3.102, 3.119, 3.122, 3.162–164, 4.13, 4.473–474

muṭbaq 4.421, 4.636

muṭbaq, ḥarf 2.459

Muth, Franz-Christoph 3.453–454, 3.717–724

muṭlaq 2.404, 2.445, 3.101, 3.307, 3.539, 4.561

muṭlaq, mafʿūl 1.436, 3.101, 3.103, 3.108, 3.455, 3.460, 4.52, 4.414, 4.540, 4.661, 4.664

Mutrah 2.213, 3.478

Muṭrān, Xalīl 4.90

Muttaqī, ʿAlī al- 2.328

muṭṭarid 1.641, 2.177, 3.92, 4.22

muttaṣil 1.555–558

muttazin 4.104

muwāḍaʿa 3.114, 3.122

muwāfiq 3.323

muwalladūn 1.97, 1.615, 1.620, 4.205

muwarrā ʿanhu 3.538

muwarrā bihi 3.538

muwaššaḥa 1.100, 1.102, 4.87, 4.92

muwāṭin 3.344

muwāṭiniyya 3.344

Muwayliḥ 3.401

muxaffaf 1.69

muxāṭab 1.555, 1.566–567, 2.302, 2.537, 4.332

muxbar ʿanhu 2.435, 3.177, 4.219, 4.737, 4.739

muxbar bihi 2.435, 4.219

muxfā 3.307

muxtalis 3.307

muxtaṣṣ 3.109

Muysken, Pieter 1.415, 1.420, 2.389

Muẓaffar Dynasty 4.276

muzammak 4.561

muzdawaj 3.539, 4.87

muzdawij 4.32

Mwali → Mohéli

Myasa 1.363

Myers-Scotton, Carol 1.415–417, 1.419, 1.484, 1.634–635, 2.654, 3.11, 4.57

Mzab 3.170

Mzab Arabic 1.264

Mzēnah Arabic 2.152, 3.401, 4.240

N

/n/ 2.300
/n/, pronunciation of 2.235
naʿal 2.364
nāba ʿan 2.316
nabāṭ 1.178, 617
Nabataea 1.127–128
Nabataean Arabic 2.198, 3.467, 3.696–697, 4.479
Nabataean Aramaic 1.126–128, 1.150, 1.153, 1.178, 1.617, 3.339, 3.409–410, 3.465–468, 3.697, 4.402, 4.479, 4.482–483
Nabataean inscriptions 1.401
Nabataean Kingdom 3.698
Nabataean script 1.148–149, 1.151, 1.401, 2.263, 3.472
Nabataeans 2.198
Nabaṭī 1.102, 1.617–618
Nabaṭī poetry 1.269, 1.273, 3.500
nabaz 3.717
Nabeul Arabic 4.563
Nabhan, Neuza Neif 3.4–5
Nābiġa aḏ-Ḏubyānī, an- 3.646, 4.88
Nabk Arabic 4.406
Nablus 2.220, 2.506
Nablus Arabic 2.32, 3.527, 4.655
naḏ̣āʾir 2.280
nadba 3.573
Nader, Laura 2.654, 2.658
Nadīm, ʿAbdallāh an- 1.598–599, 1.601–602, 3.193, 3.348
nādir 2.88, 2.92, 3.34
Nādirī, Muḥammad ʾAsad an- 3.224–225, 4.626
naḏ̣m 1.428–429, 1.471, 2.305, 3.117, 3.162–164, 4.103, 4.473–475
Nador 2.715
Naḍr ibn Šumayl, an- 3.33
Nadvi, Habibul Haq 2.326
Nadvi, S.A. Zafar 2.329
Nadwī, Syed Sulaiman 3.506
nafaḏa/ʾanfaḏa 4.414
nafas 3.124
Nāfiʿ 4.5
Nāfiʿ, ʾAḥmad 3.200, 3.202–203, 3.740–741
nafs 1.428, 4.54–57
nafs, kalām an- 2.306
nafy 2.184, 3.97, 3.549
Naḥal Ḥever 1.149–150
nahḍa 1.404, 1.531, 1.602, 2.266, 3.216, 3.344, 4.459–460, 4.571
Nahdlatul Ulama 2.336
Naḥḥās, ʾAbū Jaʿfar ʾAḥmad ibn Muḥammad an- 4.9–10, 4.132
Nahmad, Hayim M. 2.230, 3.558–559, 4.148, 4.358
naḥt 1.1, 1.451–455, 2.638, 2.641, 3.425–426, 3.630, 4.462, 4.558
naḥt xaṭṭī 4.138
naḥw 1.336, 1.338, 1.381, 2.84, 2.86, 2.187, 2.401–402, 2.625, 2.627, 3.136, 3.160, 4.119
naḥw al-jumal 2.290, 2.292
nahy 2.359, 4.332
nāʾib (ʿan) al-fāʿil 2.82, 2.315, 2.429, 3.104, 3.107, 3.558, 4.354, 4.535
Naïm-Sanbar, Samia 1.274–286, 3.671–676, 4.107–108, 4.681, 4.750
Nainar, S. Muhammad Husayn 4.433

Naʿīr 2.498
Nairobi 1.520, 1.662–663, 2.588
Naitias 2.325
Naït-Zerrad, Kamal 2.68
Naj Hammadi Arabic 2.2, 2.5
Najaf 2.407
Najd 1.241, 3.14, 3.691, 3.697
Najdi Arabic 1.220, 1.399–400, 1.606–607, 1.625, 1.627, 2.151–152, 2.161, 2.353, 2.404, 2.506, 2.609, 3.12, 3.94, 3.326–334, 3.360, 3.403, 3.429, 3.478, 3.585, 3.587, 3.606, 4.2, 4.123–124, 4.126–127, 4.129, 4.238, 4.241, 4.318, 4.628, 4.632, 4.654–655
Nāji al-ʿAli 4.498
Najjār, ʿAlī an- 1.598
Najjār, Ḥāzim 1.649
Najjār, ʾIbrāhīm 3.25
Najjār, Muḥammad an- 1.601
Najrān 3.326, 4.123, 4.125
Najrān Arabic 3.326, 3.606, 4.124, 4.127
Nakano, Aki'o 4.727
Nakao, T. 3.206
nakira 1.558, 2.225, 3.115, 3.453, 4.442
Nalborczyk, Agatha Skowron 2.29
Nalout 3.56
Namāra, inscription of 1.128, 1.150, 2.263, 3.469
Namboodiripad, E.M.S. 3.128
names → nicknames
names → proper names
names, Algerian 3.718
names, animal 3.720
names, apotropaic 3.719
names, Bedouin 3.718
names, Biblical 3.718–719, 3.721
names, Christian 4.468
names, compound 1.401
names, Druze 3.718
names, Egyptian 3.718
names, family 3.721, 4.468
names, female 3.718, 3.721
names, flower 3.720
names, foreign 1.645
names, Hebrew 3.717
names, Iranian 3.717
names, jewelry 3.720
names, last 4.469
names, modern Hebrew 3.717
names, Muslim 1.286
names, Qurʾānic 1.381
names, theophoric 3.719
names, totem 3.719
names, tribal 3.719
names, Western 3.721
names of animals 3.718
names of birds 4.285
names of God 2.600, 3.719, 4.74
names of plants 3.718
names of professions 1.296, 3.279, 3.303
names of stars 2.25–26, 2.282, 2.455
names of the days of the week 3.126
names of the months 2.342, 3.126
names of the months, Libyan 3.57
names of the signs of the Zodiac 3.126
names of tribes 3.445
Nāmī, Yaḥyā Xalīl 4.750
Namir, Fāris 3.18
Namir Arabic 3.94
Naml, Kurāʿ an- → Kurāʿ an-Naml
Napoleon 1.589, 2.8, 2.42, 2.374, 2.634, 2.699, 3.344
naqd 4.472
nāqiṣ 2.93, 2.234, 2.497, 2.548–551, 3.539, 4.23, 4.121, 4.645
naql 2.442–443, 2.638, 3.116–117, 3.121, 3.566, 4.13
Naqšbandiyya 1.380–381
naqṭ 1.151
narration 2.209
narrative 1.485–486, 4.317, 4.473, 4.499, 4.507, 4.656–657, 4.727–728, 4.730
narrative, periphrastic 1.308
narrative background 3.369
narrative imperative 1.266
narrative pattern 3.368
narrative skill 2.76
narrative structure 1.649
narrative style 1.308, 1.331, 1.552
narrative style, Bedouin 1.266

narrative text 1.649
narrative verbal noun 1.308
nasab 3.717, 3.719
Nasafī, ʿUmar an- 1.9, 1.380
nasal 3.335–336, 3.598
nasal cavity 2.677, 3.594
nasal compound 2.589
nasal emission 2.678
nasality 3.334–336, 3.598
nasalization 1.205, 2.271, 2.676, 3.334–336, 3.454, 3.598, 3.694, 4.218, 4.427
nasaxa 3.353
naṣb 1.67, 1.71, 1.309, 1.347, 1.556, 1.559–560, 1.562–563, 2.92, 2.187, 2.232–233, 2.402, 2.404, 2.430, 2.538, 3.92, 3.651
nasīb 3.173
nāṣib 1.71
Nāṣif, Ḥafnī 2.634
Nasir, Gamal Abd an- → Nasser, President Gamal Abdel
Nāṣir, ʿAbd al-Munʿim ʿAbd al-ʾAmīr al- → Al-Nassir, Abdalmunim Abdalamir
Nāṣir, Muḥammad 1.599
Nāṣir-i Xusraw 2.409
Nasirism 4.76
Nāṣiriyyah 2.571–572
nāsix 2.290, 2.308, 2.355, 3.234, 3.353–354, 4.485
nāsix al-ibtidāʾ 1.69
nāsixa 3.353
nasq 4.221
Naṣr ibn ʿĀsim 1.152
Naṣr, ʿAṭiyya Qābil 4.422
Naṣr, Muḥammad Makkī 4.426
Naṣṣ, ʾIḥsān an- 2.636
Naššār, ʿAlī Sāmī an- 4.13
Naššār, Ḥusayn 3.14–15, 3.36, 3.42, 3.689–690
Naššār, Nāṣif 3.344
Nasser, President Gamal Abdel 1.331, 1.406, 1.419, 1.531, 3.193–194, 3.196, 3.199, 3.201, 3.344–345, 3.348, 3.664, 3.666, 3.668–669
nastaʿlīq 1.287, 3.336–339, 3.341, 3.508, 3.519, 4.98, 4.132, 4.134, 4.138, 4.561
Našwān al-Ḥimyarī 2.257, 2.260, 3.693, 4.303
nasx 1.151–152, 2.601, 3.112, 3.308–309, 3.336–337, 3.339–343, 3.519, 4.131–134, 4.138, 4.562
nasx al-faḍḍāḥ, an- 3.341
nasx al-waḍḍāḥ, an- 3.341
nasxī 3.309, 3.339–340, 3.519
nasx-i ġubār 4.134
nasx-i taʿlīq 3.336–337
nasxī, Kufic 3.307
naʿt 1.124, 2.329, 3.422, 4.219–220
naʿt sababī 1.50, 4.101–102, 4.221
nation 2.59
nation, Arab 1.130–131
Nation, Paul I.S. 2.683
National Swahili Council 1.662
nationalism 2.58, 2.107–108, 2.267, 2.654, 2.662, 2.697, 3.343–353, 3.664
nationalism, Arab 1.130, 1.602, 2.656, 2.659
nationalism, Egyptian 3.349
nationalism, Islamic 1.531
nationalism, Islamization of 4.77
nationalism, Lebanese 3.348–349
nationalism, pan-Arab 1.530–531
nationalism, Syrian 3.348
nationalism, Tunisian 3.349
nationalism and language 4.76
nationality 2.59, 2.73
nationhood 2.59
nationhood, Berber 2.62
nāṭiq 4.545
nāṭiq, ḥarf 2.235
Native American languages 2.667
native speaker 2.117
nativism 2.687
nativization 1.519, 1.521, 2.31, 2.588
nativization of loanwords 2.668–669, 2.671
naṯr 3.648
naṯr mursal 4.103
Natto, Ibrahim 3.180
natural language 1.460
natural language processing 1.207, 1.511

natural order of acquisition 4.140
naturalness, phonological 3.497
Naumkin, Vitalij 4.304
Navarrese 2.285
Navāyat 2.325
nawājiḏ 4.133
Nawala 1.362
Nawar 2.217, 2.219
nawāṣib → *nāṣib*
nawāsix → *nāsix*
Nawawī, Muḥyī d-Dīn Yaḥyā ibn Šaraf ad-Dīn an- 2.335
Nawaybe 1.362
Naxl 3.401
Nayrīzī, ʾAḥmad 3.341
Nāzik al-Malāʾika → Malāʾika, Nāzik al-
Naẓīm, Maḥmūd Ramzī 1.601
Nazzal, Ayman 1.650
Naẓẓām, ʾAbū ʾIsḥāq an- 2.303–304
Nazzi, Thierry 3.724
N'djamena 1.334, 1.363, 3.61–62
N'djamena Arabic 4.708, 4.710, 4.712, 4.716
Ndzuwani → Anjouan
Nebes, Norbert 1.218, 2.355, 3.99, 3.420
Nebrija, Antonio de 1.168–169
Nedroma 1.53–54, 56
Nedroma Arabic 1.609
Neeleman, Ad 3.231–232
Nefzaoua Arabic 4.681
negation 1.64, 1.92, 1.183, 1.185, 1.217–218, 1.230, 1.239, 1.268, 1.279–280, 1.285, 1.303, 1.327, 1.377–378, 1.392, 1.396, 1.549, 1.607–608, 2.101, 2.116, 2.133–134, 2.195, 2.249, 2.490, 2.512, 2.570, 2.596, 3.69, 3.84, 3.152, 3.183, 3.221, 3.232, 3.257–258, 3.273, 3.355–360, 3.366, 3.386, 3.485, 3.681, 3.739, 4.247, 4.398–399, 4.650, 4.696, 4.745
negation, absolute 1.563, 2.308, 4.112
negation, absolute → *lā li-nafy al-jins*
negation, acquisition of 2.100, 4.140–141
negation, constituent of 3.356, 3.391
negation, double 2.596
negation, English 4.140
negation, metalinguistic 3.680–681
negation, nominal 1.64, 1.249, 1.266, 1.280, 1.285, 3.84, 3.485
negation, sentential 3.356, 4.398–399
negation, verbal 3.740
negation complex 3.357–358
negation emphasizer 2.194
negation in sign language 4.233
negation marker 1.106, 1.524, 2.521
negation marking 4.650
negation *miš* 2.194–195
negation morpheme 1.377
negation of possessive 3.675
negation particle 1.227, 1.564, 2.195, 3.70
negation suffix 2.194
negative attitude 1.656
negative circumfix 1.185–186, 3.357–358, 3.406, 3.531, 4.569, 4.649–650
negative polarity 3.359
negative polarity licensing 4.400
negative-polarity items 3.653
negative-restrictive 2.145
negator 2.135, 2.577
negator, absolute 3.202
Negev 1.83, 1.260
Negev Arabic 1.260, 1.266, 3.360–369, 3.400–401, 3.403–406, 3.585, 4.238, 4.240, 4.350
Negev Bedouin Arabic 1.265
Negev, Avraham 3.360, 3.400, 3.469
Nehme, Laila 3.471
Nejmi, H. 3.726
Nelson, Gayle L. 3.681
Nelson, Kristina 2.662, 3.567, 4.425–428
NEMLAR 1.215
Neo-Arabic 1.613–614, 1.616–617, 1.619, 2.265–266
Neo-Aramaic 1.222, 1.490, 1.492, 1.540, 2.61, 2.414, 2.607, 2.671, 3.370–373

Neo-Aramaic, Ahwaz 3.373
Neo-Aramaic, Eastern 1.493, 3.370–372
Neo-Aramaic, Hertevin 1.540, 3.371
Neo-Aramaic, Jewish 3.372
Neo-Aramaic, Maʿlūla 3.370–372, 3.409
Neo-Aramaic, Mandaic 1.178–179, 2.67, 3.370–373, 3.409–410, 4.302
Neo-Aramaic, Northeastern 1.189
Neo-Aramaic, Salamas 3.371
Neo-Aramaic, Turoyo 1.540, 3.370–371, 3.409
Neo-Aramaic, Urmi 1.540
Neo-Aramaic, Western 3.370–372
Neo-Aramaic, Zakho 3.372
Neogrammarians 1.615, 3.604–605
neologism 1.247, 1.403, 2.30, 2.214, 3.28, 3.378, 3.426, 4.459, 4.465
neologisms in *rajaz* 4.34, 4.36
neologisms, Persian 3.579
Neophyte College 1.171
Neo-Syriac 2.61
Nerbonne, John 1.586
nesih 3.341
Nespor, Marina 2.397
Nessana ostraca 1.149
nestalik (*nastaʿlīq*) 3.338
Nestorians 2.61, 4.402
Netherlands, Arabic in the 2.73–74, 2.97
Netherlands, the 2.72
Nettelbladt, Ulrika 2.679
neutral lexis 1.658
neutralization 1.61, 1.206, 3.606
neutralization of vowels 1.276
Neuwirth, Angelika 4.103
Neville, Edward 4.275
Nevo, Moshe 2.534, 4.581
New Caledonia 2.472
New Nubia 2.1
New York City 3.395, 3.397
Newbold, Douglas 3.435
Newbold, F.R.S. 2.217–219
Newby, Gordon Darnell 2.534
Newman, Daniel 4.669
Newman, Paul 2.251
Newmeyer, Frederick 4.488
news 1.213–214, 1.648, 2.10, 2.208, 2.214, 2.332
news → broadcasting
news media 2.381, 3.199, 3.201
newspaper language 4.726
newspaper language → Media Arabic
newspaper style 1.624
newspapers 1.175, 1.214, 1.446, 1.531–532, 1.534, 1.598, 1.601, 1.603, 1.649, 1.651, 1.6552.466, 2.503, 3.193, 3.200–201, 3.318
newspapers → press
newspapers, Arabic 2.381–382
newspapers, Arabic in Nigeria 3.376
newspapers, Arabic in the United States 3.396–397
newspapers, Bengali 1.288
newspapers, Moroccan 3.295
newspapers, online 2.381
Newton, Brian 1.536
Ngajenan, Mohamad 2.341
Ngala 3.373
Ngaoundéré 1.336–337
Ngaoundéré, University of 1.338
Ngazidja 1.450
Ngom, Fallou 2.30, 4.723–725
Nhial, Abdon J. 2.518
Niasse, Al-Hajj Ibrahima 4.182
Niasse, Mohammadou 4.182
Niasse, Roqaya 4.182
Nicaise, A. 3.725
Nichols, Johanna 4.372
Nicholson, Reynold A. 3.731
nicknames 1.341, 1.646, 3.183, 3.305, 3.720, 4.469
Nicolai, Robert 4.279
Nicosia 1.536
nidāʾ 2.80, 2.425, 3.549
Niebuhr, Carsten 2.42
Niehoff-Panagiotidis, Jannis 3.6
Nielsen, Helle Lykke 2.679, 4.142–143, 4.146–156
Nietzsche, Friedrich 4.76
Nieuwkerk, Karin van 2.470
Niffārī, Muḥammad ibn ʿAbd al-Jabbār an- 2.601, 4.204, 4.211

Niger 1.334, 2.250, 3.312, 4.279
Niger-Congo languages 2.137, 4.182
Nigeria 1.334, 1.337, 1.529, 2.137, 2.250, 2.471, 3.312–314, 3.373–377, 3.634, 4.279, 4.376, 4.709, 4.715–716, 4.758
Nigeria, Arabic in 1.260, 3.374, 3.449, 4.310, 4.708
Nigerian Arabic 1.37, 1.502, 2.471, 3.542, 3.544–546, 4.118, 4.710, 4.714
Nigm, Fuad → Nijm, ʾAḥmad Fuʾād
Nijm, ʾAḥmad Fuʾād 1.599, 1.601, 3.348
Nikkhou, Mahtab 1.513
Nile Valley 1.23
Nilo-Saharan languages 1.35, 1.37, 2.553, 3.56, 4.279, 4.376, 3.435, 4.508
Nilotic languages 2.597
Nilsen, Don Lee Fred 1.343
niʿma 1.73–74, 1.564, 2.80, 4.283
Niʿma, Fuʾād 3.379
Nīsābūrī, Šāh Maḥmūd an- 3.337
Nisan, Mordechai 2.59
nisba 1.14, 1.17, 1.38, 1.192, 1.386, 1.404, 1.452, 1.490, 1.638, 1.640, 2.45, 2.285, 2.294, 2.347, 2.353, 2.424, 2.631, 3.153, 3.241, 3.279, 3.377–381, 3.424, 3.426–427, 3.430, 3.445, 3.532, 3.576, 3.578, 3.717, 3.720, 4.93, 4.463, 4.510
nisba, tamyīz an- 4.437
nisba in place names 4.511, 4.513
nisba in Turkish 4.585, 4.587
nisba tāmma 4.686
Nishio, Tetsuo 2.2, 3.401, 3.614, 4.240
nisyān 1.124
niyāḥa 4.104
niyya 2.301
Niẓām ad-Din, Mulla 2.327, 3.510
Niẓām al-Mulk 2.600
Niẓāmī 3.337–338
Nkrumah Teachers College 1.661
Nobiin 2.59, 2.668, 3.435, 3.437–438, 4.375
Noja, Sergio 3.469, 3.472
Nokou 1.363
Nolan, Francis 2.229
Nöldeke, Theodor 1.149, 1.400, 1.489, 1.492, 1.576, 1.615, 2.56–57, 2.264, 2.350, 3.7, 3.93, 3.410, 3.414, 3.420, 3.645, 3.689, 3.717, 3.719, 3.721, 4.24, 4.285
Noll, Volker 2.285
nomadic dialects 3.609
nomadic tribes 1.269, 1.363, 1.492
nomadism 1.492, 3.402, 4.124
nomads, commercial 2.216
nomen abundantiae 1.577
nomen actionis 1.573
nomen agentis 2.350
nomen instrumenti 1.574
nomen instrumenti → noun of instrument
nomen loci 1.577
nomen locis et temporis 1.574
nomen unitatis 2.156–157
nomina segolata 1.85
nominal, derived 3.388, 4.743
nominal, process 3.168
nominal, result 3.168
nominal clause 3.381–388, 4.317
nominal derivational site 1.13–14
nominal negation 1.64, 1.249, 1.266, 1.280, 1.285, 3.84, 3.485
nominal phrase → noun phrase
nominal sentence 1.182, 1.348–349, 3.202, 4.354
nominal sentence → nominal clause
nominalization 1.352, 2.352, 3.203, 3.388–394
nominalizing morpheme, Hausa 2.253
nominative 1.347–350, 1.559–560, 1.562–563, 1.580, 1.643, 2.146–148, 2.403, 2.626
nómos 3.176
non-anaphoric 1.565
non-Arabs 3.349–351
nonce borrowing 1.416, 1.496
nonce word 2.101
nonconcatenative morphology 1.376, 1.457, 1.484, 1.573, 3.423, 4.93, 4.85, 4.269, 4.336, 4.454

non-configurational 4.524
non-differential → *non-différentiel*
non-différentiel 1.114, 1.391, 1.584, 3.148, 3.406
non-echo questions 2.393
non-factual 1.331
non-inchoative verb 1.198
nonlinearity 3.48
nonmanual signal 4.232–233
non-Muslims 3.346
Nonnos 1.127
non-person 1.566–567
nonpunctual 3.643
non-telic 1.198
nonverb 3.740
Norlin, Kjell 2.395–396, 3.727, 4.670–672
norm 1.191
norm, linguistic 1.667
Norman, Donald 4.271
Normans 3.141–142, 4.215, 4.218
Norris, H.T. 3.169, 3.173
North Africa 1.260
North African Arabic → Maghrebine Arabic
North America 3.394–400
North America, Arabic in 2.665, 3.313, 3.394–400
North America, Islam in 3.394
North Arabian 2.263, 3.227
North Arabian → Thamudic
North Arabian Arabic 1.259, 4.751
North Arabian, Ancient 2.263, 3.227, 3.410, 3.413–415, 3.417, 3.419, 3.465–466, 3.468, 3.471, 4.478
North Arabic, Early 2.198
North Caucasian languages 3.65
North Israel Arabic 1.265
North Israel Bedouin Arabic 1.269, 271
North Yemen 1.528
Northern Cities chain shift 4.628
Northwest Arabian Arabic 1.606, 3.94, 3.360, 3.400, 3.646, 3.695, 4.238, 4.240, 4.241
Northwest Semitic languages 1.188, 1.203, 2.262, 3.408–422, 4.170–171, 4.301–302, 4.313
Nortier, Jacomine 1.418, 1.484, 2.76–77, 3.318
Norwegian 4.53
Nostratic languages 1.36, 2.67
Nouaouri Izrelli, Nadi Hamdi 2.75–76
noun 1.16, 1.18, 2.90, 2.92, 2.171, 2.180, 2.184, 2.424–429, 3.178, 3.422–428, 3.546
noun → *ism*
noun, abstract 2.156
noun, canonical 3.303
noun, collective 1.237, 1.249, 3.243, 4.356
noun, common 1.645
noun, count 1.551, 2.146
noun, definition of 2.424, 2.427
noun, denominal 3.423–424
noun, derivative 3.423
noun, deverbal 3.232, 3.423–424, 4.540
noun, event 3.165–166, 3.168
noun, generic 1.579
noun, inanimate 2.156–157, 2.159
noun, indefinite 1.109, 1.580
noun, instance 2.246
noun, mass 2.146
noun, noncanonical 3.303
noun, primitive 3.423
noun, psych 3.709
noun, quadriradical 1.580, 2.352
noun, relational 2.584, 3.445
noun, topic 4.362
noun, unit 1.237, 1.433–434, 1.551, 2.156–157
noun, verbal → verbal noun
noun, weak 1.561
noun class 2.140, 2.157
noun of instrument 1.328, 3.153, 3.279, 3.303, 3.424, 4.52, 4.95, 4.175, 4.461–462, 4.487, 4.535
noun of place 3.153, 3.424, 4.95, 4.175
noun of profession 1.578, 3.445, 4.52, 4.95
noun of species 1.433
noun of time 3.424
noun phrase 1.13, 1.19, 1.283, 1.331, 1.343, 1.347, 1.354, 1.480,

1.507, 1.581–583, 3.186, 3.258, 3.428–434
noun phrase, genitive 1.479
noun phrase, left-peripheral 4.364, 4.366
noun phrase, *maṣdar* 3.167
noun phrase, quantificational 3.687
novelists, female 1.529
Nowak, Jürgen 4.612
NP → noun phrase
NP raising 3.709
n-passive 1.606
NP-trace 1.348
Nuʿayma, Mixāʾīl 4.91
Nuba Mountains 4.375
Nubi → Ki-Nubi
Nubia 1.361, 2.559, 4.708
Nubian 1.36, 1.519, 1.527, 2.1, 2.219, 2.560, 2.671–672, 2.695, 3.60, 3.313, 3.435–439, 4.375, 4.377
Nubian → Dongolawi
Nubian → Kenzi-Dongolawi
Nubian → Nobiin
Nubian, Darfur 3.435
Nubian, Kordofan 3.435
Nubian, Nile 3.435
Nubian, Old 3.435
Nubian Arabic 2.494, 3.60
nuclear accent 2.396, 2.398
nuclear element 4.387
Nuclear Harmony Constraint 3.493
nucleus 2.143, 2.146, 3.601–602, 3.611, 3.729, 4.387–389, 4.646
nudba 1.73, 2.80, 4.329
Nuer 1.519, 4.375
Nufud 3.401, 3.692
Nugteren, Hans 4.603–612
Nuhāy 4.481–482
null copula 1.349, 1.507
null pronominal 4.395
null pronoun 1.45–46
null restriction 4.17
null subject 2.174, 3.706–711, 4.524, 4.527
null-subject language 2.173
Numayrī, ʾAbū Ḥayyā an- 3.650
number 1.14, 1.43, 1.47, 1.434, 2.157, 2.347, 2.424, 3.303, 3.426, 3.430, 3.439–447, 4.183
Number Phrase 1.351
numeral 1.108, 1.190, 1.228, 1.237, 1.251, 1.264, 1.280, 1.292, 1.304, 1.328, 1.386, 1.393, 1.477, 1.550, 1.643, 2.221, 2.244, 2.253, 2.420, 2.485, 2.512, 2.523, 2.593, 2.616, 3.69, 3.154, 3.183, 3.200, 3.280, 3.447–453, 3.652, 4.15, 4.18–19
numeral, cardinal 2.102, 3.447–451, 4.18
numeral, ordinal 3.431, 3.452
numeral phrase 1.266
numerals, Arabic in Swahili 4.385
numerals, Bantu 2.671
numerals, borrowing of 2.668
numerals, compound 1.309
numerals, Modern South Arabian 4.300
numerals, Swahili 1.662
numerals in Tajik 4.617
numerator 2.146
numerical value of letters 3.111
Numidian 4.511
Numidian Kingdom 4.571–572
nūn an-niswa 2.327
nūn at-tawkīd 2.22, 2.327, 3.573
nūn paragogicum 1.493
nūn sākina 2.239, 3.335
nunation 1.188, 1.190, 1.397, 1.400, 1.491, 1.559–560, 1.562, 1.579–580, 1.615, 1.668, 2.296, 2.424–425, 3.227, 3.240, 3.244, 3.333, 3.414, 3.453–454, 3.569, 4.357, 4.442–446, 4.479, 4.558
nunation in Thamudic 4.479
nunation, absence of 4.558
Nuñez, Hernán 1.168
Nupe 4.282
nuqṣān 3.119, 3.649
nuqṭa 2.232, 3.741
Nūr ad-Dīn ar-Ranīrī → Ranīrī, Nūr ad-Dīn ar-
Nūr ad-Dīn, Emir 2.601
Nūr al-Ḥaqq ibn Luqmān, Muḥammad → Ma Lianyuan
Nuristani 1.29

Nurse, Derek 1.660, 1.663, 1.665, 2.671, 4.381–382
Nuṣaybīn 1.87
Nuṣaybīn Arabic 2.534
Nuṣayri Arabic 1.642
Nuṣayri-Alawis 1.388
Nuṣayris 1.111, 4.580
Nuṣṣ, ʿAbd al-ʿAzīz an- 1.599
nusxa 1.2, 4
nuṭq 4.545
Nuwayrī, Šihāb ad-Dīn ʾAḥmad ibn ʿAbd al-Wahhāb an- 3.339, 4.132, 4.134
Nuwīrī, Muḥammad an- 2.580
Nwēbiʿ 4.238
Nwyia, Paul 4.75
Nydell, Margaret K. 4.469
Nyimang 4.375

O

O’Connor, Joseph 1.515
O’Connor, Michael 3.409, 3.466
oath 1.106, 1.202, 2.23, 2.80, 3.257, 4.75, 4.236
Oba 3.374
Obdeijn, Herman 2.73
Obeidat, Hussein A. 3.713, 3.715
Obeidat, Nawaf 4.501
object 1.41–42, 1.67–68, 1.182, 1.344–348, 1.428, 1.622–623, 1.625, 2.146, 2.149, 3.100, 4.529, 4.535–542
object → *mafʿūl*
object, absolute 3.101, 3.108, 3.166, 3.455–461, 4.52, 4.540, 4.661
object, associative 4.540
object, clausal 4.537
object, dative 1.625, 4.621
object, definite 4.729
object, direct 1.356, 1.563, 2.550, 3.102–103, 3.224, 3.226, 4.413
object, double 3.591
object, focused 4.731
object, focusing of 4.732
object, fronted 4.544
object, indefinite 4.729, 4.732
object, indirect 4.533–534
object, locative 3.101
object, pragmatic 4.732
object, prepositional 1.347, 3.554, 3.701, 4.735
object displacement 4.540
object marker 1.562, 3.158–159
object marking 1.34, 1.95, 2.422–423
object of purpose 4.540
object of time and place 4.540
object pronoun suffix 1.30, 3.329
object raising 2.369, 3.103, 3.226, 4.539
object shift 4.540–541
object suffix 1.279, 1.302, 1.326, 1.391, 3.277
object suffix, anticipatory 2.531
objective conjugation 2.349
obligatorification 2.195
Obligatory Contour Principle 1.204, 1.653, 3.461–464, 3.626, 4.95–96
oblique 1.562
Obock 1.655
Obodas, King 3.469
Obrecht, Dean H. 2.231, 2.461, 4.670
obscenity 4.630
obsolete word 3.632
obstruent 3.596, 3.625
obviation requirement 4.84
Ochs, Elinor 2.654, 3.665
OCP → Obligatory Contour Principle
OCP rule 4.94–96
OCP violation 4.96
OCP-Place 3.626–627
OCR → Optical Character Recognition
Odden, David 3.461
Odisho, Edward 4.2
Oduyoye, Modupe 4.759
Offer 2.144
official language → language, official 2.635
offset 3.596
Ogawa, Yoshiki 3.388
Ogunbiyi, Isaac A. 3.374, 4.758–763
Ohala, John 2.365, 4.336

Okesanya, Ayo 3.196
Ola, Louis Biajo 1.525, 3.642
Olayiwola, Abdul Fattah 3.373–377
Old Arabic 1.399–400, 1.614, 2.265, 2.267, 3.415, 3.417, 3.464, 4.478–479, 4.482
Old Iraqi School 2.226, 2.436, 3.107
Old Testament 1.126–127, 1.232
Olmo-Lete, G. del 4.93
Olshtain, Elite 2.661
Olson, Mike 4.197
Oman 2.211–213, 2.215, 3.478
Oman, Arabic in 2.272, 3.89, 3.544, 3.672, 3.674, 4.304, 4.751
Omani Arabic 1.241, 1.243–246, 1.250, 1.252, 1.259–260, 1.663, 2.211, 2.619, 3.84, 3.478–491, 3.544, 3.658, 4.124, 4.298–300, 4.383
Omani Arabs 4.381–382
Omanis, Zanzibari 2.211
Omar, Ahmed Mokhtar 3.321
Omar, Margaret 1.341, 2.96, 2.98–102
Omara, Salma 3.678
Omdurman 2.559, 4.381
Omdurman Arabic 4.376
omission 4.22, 4.26
omission in the *Qurʾān* 4.27
Omotic languages 1.35
one-word stage 2.100
Ong, Walter J. 3.499
online newspapers 2.381
ónoma 3.547
onomastics 3.717–724
onomatopoeia 1.331, 1.367, 1.395, 1.503, 2.288, 3.185, 3.537, 4.93, 4.96, 4.212, 4.226, 4.282–290, 4.421
onset 3.596, 3.601, 3.611, 3.729, 4.387, 4.389
Onset Constraint 3.493, 4.85
onset, complex 3.612
Onsori 3.575
Ontario 3.398
opacity 1.121, 2.37–38, 2.365
opacity, structural 4.39
opaqueness 1.325
Open Principle 1.658
operation 2.187
operator 2.135, 2.143, 2.169, 3.234
operator, abstract 4.484
operator, clause 4.732
operator, illocutionary 2.144
operator, interrogative 2.143
operator, overt 2.430
Oppenheim, Max Freiherr von 3.401, 4.402
optative 1.281, 1.406, 2.524, 3.248, 3.333
Optical Character Recognition 1.455
optimality 4.39
Optimality Theory 2.36–38, 3.179, 3.463, 3.491–499, 3.626–628
Optional Infinitive Stage 1.377
optionality 3.231
oral literature 3.172, 3.174, 3.731, 4.129
oral literature, Amazigh 2.715
oral performance 1.597
oral poetry 3.646, 4.90
oral proficiency interview 4.150
oral tradition 3.645–646
oral transmission 4.105
orality 1.648, 3.499–501, 3.645, 4.201
orality, secondary 3.500
Oran 1.54–56, 4.156
Oran Arabic 1.443, 1.596, 1.610, 2.532, 4.680
Orania Arabic 1.609
order, *ʾabjad* 3.31
order, alphabetical 3.27, 3.32
order, canonical 2.369
order, linear 4.746
order, permutative 3.32
order, phonetic 3.31
order, rhyming 3.32, 3.44
order, Xalīlian 3.31–32
ordinals 3.431, 3.452
Orientalische Akademie 1.171
Orientel 1.513
origin of speech 2.178, 4.684–685
Orioles, Vincenzo 2.456
Orma 4.381

Oromo 1.35, 2.52, 2.54–55, 3.494
Oromo Region 2.51
oropharynx 4.669
Ortaylı, İlber 3.503–504
orthographic sign 1.152
orthography 1.153, 1.209, 1.371, 1.487, 1.532, 2.326, 2.408, 2.692, 3.79, 3.111, 3.220
orthography of pausal forms 3.566
orthography, Aramaic 1.401
orthography, Judaeo-Arabic 2.527–529, 2.531
orthography, Kurdish 2.605
orthography, Qurʾānic 1.401
orthography, standardization of 2.636
Ortner, Sherry 1.636
Os, Charles van 2.683
Osama bin Laden → ʾUsāma bin Lādin
Oseni, Zakariyau I. 3.374
Osgood Scale 2.652
Osman Ali, Mohammed 1.655–656
Osman, Mariam 3.710
Osmanli 1.133–134
Osmanov, M.N. 3.575
Osrhoene 1.126–128
Oster, AnneMarie 2.677
Ötençay 1.115
Othman, Emam 1.460
otitis media 2.675
Ottenheimer, Harriet 1.663
Ottenheimer, Martin 1.663
Otting, Janna 2.689
Ottoman Empire 1.175, 1.404, 1.602, 2.505, 2.654, 3.501–505, 4.572–573, 4.577–578, 4.580, 4.583–584, 4.589
Ottoman Empire, Arabic in the 3.503–504
Ottoman Turkish 1.95, 1.133–135, 4.584–586
Ottomans 1.174, 1.589
Ouaddai 1.360–361, 1.364
Ouali, Hamid 2.673
Ouaouicha, Driss 1.651
Oubari 3.56
Ouhalla, Jamal 2.100, 2.115, 2.133, 2.369, 3.232–233, 3.355–360, 3.382, 3.653, 3.655, 3.684, 4.15, 4.64, 4.84, 4.354, 4.362–363, 4.398–400, 4.492, 4.541, 4.649, 4.657, 4.745, 4.747
Oujda 2.470
Oujda Arabic 3.288–289, 3.292
Oulad Tayma 3.287
Ould al-Barra, Yahya 2.362
Ould Bah, Mohamd El-Moktar 3.172–173
Ould Cheikh, Abdel Wedoud 2.362, 3.170, 3.174–175
Ould Daddah, Mokhtar 3.174
Ould Ebnou, Moussa 3.173–174
Ould Mohamed Baba, Ahmad-Salem 3.173, 4.70
Ould Zenagui, Cheikh El-Bou 3.173
Oum-Hadjer 1.362
ousía 4.220
out-group 1.415, 4.321
output structure 4.521
overgeneralization 2.102, 2.162, 4.139
overgeneralization error 2.101
overt governor 1.69, 1.71
Oyo 4.759
Owens, Jonathan 1.194, 1.269–270, 1.458, 1.490, 1.518–527, 1.562, 1.611, 1.613–614, 1.618, 2.176, 2.227, 2.232, 2.236–238, 2.290–291, 2.294–295, 2.308, 2.356, 2.390, 2.402, 2.425–427, 2.436, 2.471, 2.501, 2.517–518, 2.539, 2.588–589, 2.597, 2.642, 2.672, 2.695, 3.53, 3.55–56, 3.61, 3.101–102, 3.104, 3.107–109, 3.314, 3.317, 3.402–405, 3.423, 3.429, 3.541–546, 3.566, 3.568, 3.634, 3.639–643, 3.689, 4.9, 4.220–221, 4.330, 4.351, 4.410, 4.429, 4.431, 4.436, 4.443, 4.448–449, 4.612, 4.621, 4.655, 4.708–718
Oxford 1.171–172
Oxford, Rebecca 4.141
Oxford Text Archive 1.512
oxymoron 3.733

P

/p/ 2.415, 3.65, 3.189–190
padding device 3.201
Padgett, Jaye 2.36, 3.462–463, 3.625, 3.627
Padilla, Amado M. 1.368
Padova 1.171
Pagliuca, William 2.193–194
Pahlavi 1.401–402, 2.199–200, 2.286, 2.406, 2.673, 3.574, 3.576–577, 3.580–583, 4.424, 4.614
Pajas, Petr 1.458, 1.516
Pakistan 1.257, 1.288, 1.529, 2.211, 2.326, 2.673, 3.506–512
Pakistan, Arabic in 2.672, 3.506–512, 4.595–603
Pakistanis 2.215
palatalization 1.25, 1.260, 1.324, 1.390, 1.539, 1.608, 2.3, 2.7, 2.212, 2.235, 2.270, 2.482, 2.495, 2.555, 2.590, 3.192, 3.478–479, 3.524–526, 3.606, 4.2, 4.110, 4.173, 4.303, 4.373, 4.710, 4.751, 4.753–754
palatalization, Neo-Aramaic 3.372
palato-alveolar 3.598
paleography 1.151, 3.513–524
paleography, Qur'ānic 2.597
Palermo 2.454, 3.142, 4.215–216
Palestine 1.589, 2.105, 2.437, 2.439, 2.464, 2.498, 2.500–502, 2.660, 3.526
Palestine, Arabic in 1.571, 1.583, 1.605, 1.607, 2.4, 2.31–32, 2.494, 3.8, 3.702, 3.726
Palestinian Arabic 1.23, 1.312, 1.351, 1.375, 1.377, 1.481, 1.538, 1.540–541, 1.546, 1.608, 2.100–101, 2.230–231, 2.365–366, 2.370, 2.461, 2.463, 2.466, 2.499, 2.502, 2.506–507, 2.509, 2.587, 2.691, 3.11, 3.82–84, 3.86, 3.263–268, 3.360, 3.403, 3.430, 3.433, 3.526–538, 3.569, 3.609, 3.656, 3.685, 3.687, 3.704–705, 4.117, 4.194, 4.317–318, 4.346–349, 4.361, 4.393, 4.407, 4.534, 4.627, 4.637, 4.654, 4.662–663, 4.672, 4.675–676, 4.681, 4.746
Palestinian poetry 4.77
Palestinians 2.105–106, 2.437–438, 2.464–466, 2.498–499, 2.502, 2.506, 2.510, 2.658, 3.11
Palestinians, Christian 2.59
Palfreyman, David 3.500
Pali 2.286
Palmer, H.R. 2.250
Palmyra 1.23, 2.198, 4.402–403, 4.480
Palmyrene Arabic 1.538, 2.314, 2.494, 4.406–408
Palmyrene Aramaic 3.409–410, 3.467
Palmyrene inscriptions 1.401
Palva, Heikki 1.269, 1.405–411, 1.502, 1.593, 1.595–596, 1.604–613, 2.501, 2.506, 3.12, 3.360, 3.368, 3.400–408, 4.240–241, 4.664
Pamir 1.29
Pamplona 2.44
pan-Arabism 1.173, 1.531, 3.194, 3.344–345, 3.347, 3.352
Pancasila 2.337
panegyric 2.446

Pāṇini 1.40, 3.497
Pankhurst, Alula 2.53–54
pan-nationalism 2.654
Pantelleria 3.141, 4.215–216
Paoli, Bruno 3.208, 3.210, 3.615–623
paper 3.519
Papuan languages 4.372
papyri 1.148–149, 1.151–152, 2.23, 2.199, 2.256, 3.216, 3.516–517, 3.519
papyri, Aphrodito 2.23
papyri, Arabic 3.515
papyri, Egyptian 3.518
papyri, Greek 3.467
papyri, Greek/Arabic 3.516
Pará 3.1
paradigm 2.192
paradigmaticity 2.196
PARADIS 1.210, 1.213
Paradis, Carole 2.122–131
paragraph 1.473
paralinguistic strategy 4.326
parallel evolution 1.594
parallelism 1.647–648, 1.651, 2.146, 3.492, 3.500, 3.669
parameter 2.319, 3.228
parameter, null subject 3.706
parameter, pro-drop 3.706–707
paraphasia 2.690
paraphasia, semantic 2.691
paraphrase 3.500, 4.545
parataxis 1.648, 4.485, 4.544
parataxis → coordination
paratone 1.473
parentese → caretaker talk 1.339
parenthesis 3.742
parenthesis in the *Qurʾān* 4.28
parenthetical comment 1.406
Paris Polyglot 1.167
Parisi, Domenico 2.689
Park, Tschang-Zin 1.368
Parker, Simon 3.409
Parkinson, Dilworth B. 1.532, 1.630–631, 1.633–635, 2.8, 2.11, 2.657, 2.695, 3.14, 3.201–202, 3.708, 3.711–712, 4.466–471, 4.726
paroimía 3.730
paronomasia 2.496, 3.456, 3.538–541, 3.669, 3.733, 4.52, 4.211
paronomastic construction 1.188
Parry, Milman 3.499, 3.645–646, 4.201
parse tree 1.459
parser 1.460, 1.515
parser, morphological 1.207–209, 1.214
parser, syntactic 1.459
Parsi 2.407
Pārsī-i Darī 2.406
parsing 1.459, 1.514
parsing, automatic 1.514
participant observation 2.653
participative 2.348
participle 1.13, 1.16, 1.20, 1.32, 1.39, 1.52, 1.68, 1.229, 1.265, 1.306, 1.330, 1.551, 1.553, 1.574, 1.576, 2.424, 2.431, 2.449, 2.486, 2.514, 3.185–186, 3.254, 3.259, 3.282, 3.422, 3.424, 3.533, 3.541–546, 4.456
participle, active 1.65, 1.196, 1.198–199, 1.201, 1.218–219, 1.246, 1.302, 1.331, 1.394, 1.396, 1.578, 2.248, 2.490, 2.613, 3.97, 3.101, 3.156, 3.368, 3.536, 3.541, 3.543–546, 4.52
participle, active, in Northwest Semitic 3.417
participle, ‘developmental’ 3.253
participle, passive 1.68, 1.197, 1.201, 1.394, 1.625, 2.367, 3.101, 3.157, 3.541, 3.544, 3.553–554, 3.556
participle, passive, as verbal noun 4.660
participle, perfect 4.620–621
participle verbalization 4.306
particle 2.90, 2.180, 2.184, 2.236–237, 2.427, 2.429, 3.546
particle → *ḥarf*
particle, aspectual 4.622
particle, asseverative 1.201–204, 2.22, 2.477
particle, conditional 2.479
particle, connective 3.587
particle, definite 2.522–523, 2.569–570

particle, determinative 2.521
particle, emphatic 4.358, 4.499
particle, exceptive 3.550
particle, exclamatory 4.420
particle, existential 2.522, 2.575, 3.369, 4.356
particle, genitive → genitive particle
particle, Greek 4.546
particle, interrogative 1.230, 2.491, 2.511, 3.278, 4.27
particle, linking 4.617, 4.620–622
particle, modal 3.235
particle, narrative 3.364–365
particle, negation 1.248, 2.614, 3.183, 3.278, 3.634, 3.694, 4.746
particle, oath 3700, 4.429
particle, performative 3.570, 3.573
particle, phrasal 3.630
particle, preverbal 4.698
particle, proclitic 4.545
particle, question 2.134–135, 3.278, 3.485, 4.622, 4.697
particle, subordinating 1.669
particle, topicalizing 3.161
particle, verbal 1.216
particle, vocative 1.249, 1.304, 2.592, 3.257, 3.278, 3.366, 3.485, 3.532, 4.467, 4.471
Partition 1.288
partitive 1.84, 1.190, 2.248, 2.297
parts of speech 1.513, 2.425, 3.546–552
parts of speech → *ʾaqsām al-kalām*
parts of the body 2.160, 2.193, 2.272, 2.594, 2.615, 3.152, 3.279, 3.672, 3.736–738, 4.554, 4.565
Pasai 2.334
Paşaköy 1.115–116
Pasch, Thelma 1.525
Pascual, Jose A. 4.75
Pashto 1.28, 34, 133–134, 3.350, 3.506, 3.510, 3.574
Passamaquoddy 4.539
passé simple 4.728
passive 1.67, 1.297, 1.408, 1.410, 1.493, 1.522–523, 2.93–94, 2.99, 2.347–348, 2.449, 2.523, 2.566, 2.568, 2.576, 2.590, 2.593, 3.77–78, 3.171, 3.202, 3.251, 3.281, 3.298, 3.301–302, 3.533, 3.552–564, 4.306, 4.522, 4.533, 4.568, 4.640, 4.727
passive, acquisition of 1.377
passive, adjectival 4.490
passive, agentless 3.555
passive, analytic 3.553
passive, apophonic 2.271, 3.191, 3.224–225, 3.552, 4.626, 4.697
passive, impersonal 1.625, 2.317, 3.221, 3.563, 3.572, 4.357, 4.365, 4.530, 4.533
passive, internal 1.39, 1.108, 1.253, 1.398, 1.606, 1.611, 1.616, 1.624, 2.213, 2.244, 2.421, 2.486, 2.530, 2.576, 2.618, 3.184, 3.326, 3.332, 3.478, 3.488, 4.114, 4.124, 4.176, 4.306, 4.555
passive, periphrastic 2.195, 3.200
passive, reflexive 3.181
passive, stative 2.593, 2.596
passive, ‘subjectless’ 3.710
passive construction 1.623–626
passive in Ki-Nubi 1.522
passive participle → participle, passive
passive transformation 1.654
passive verb 1.75
passive voice 1.346
passivization 1.252, 2.618, 3.225, 3.560, 3.563, 4.538
past continuous 1.229
past tense 1.199–200, 1.331, 1.377, 1.349, 1.508–509, 1.568, 2.91–92, 3.186
past tense, acquisition of English 2.101
Patai, Rafael 2.498–499
Patani 2.334
Pate 4.381
paternity 2.58–59
páthos 2.309, 3.114
pathway 2.191–192
patient 2.625, 3.100, 3.102–103, 4.487, 4.489–491
patient, semantic 3.103
Patkanoff, K.P. 2.217, 2.220
Patna 2.329

Patriarchi, Christina 3.435
patriarchy 2.642, 2.645–646
patrimony 2.58–59
patronymic 4.573
patronymics, Tunisian 4.573
Pattani 4.477
pattern 1.573–577, 3.240, 3.299, 3.301, 4.93, 4.96, 4.446
pattern drill 1.484
pattern morpheme 2.346
patterns, acquisition of 1.377
paucity 1.283
Paul, Hermann 1.575
Paulston, Christina Bratt 2.58
Paulussen, Hans 2.702
pausa → pause
pausal allomorph 3.585
pausal context 4.682
pausal form 1.262, 1.301, 1.391, 1.400, 1.548, 1.616, 2.5, 2.353, 2.404, 2.418, 3.208, 3.240, 3.564–570, 3.645, 4.115
pausal form in transcription 4.519
pausal forms, orthography of 3.566
pause 1.190, 2.16, 2.255, 3.240, 4.558
pausing phenomena 2.397
Payne, Thomas E. 4.197, 4.535, 4.455
peak 4.387, 4.729
Peake, Frederick G. 2.498, 3.402
Pearson, Barbara Zurer 1.368
pedagogy 1.651, 2.110
Pedro de Alcalá → Alcalá, Pedro de
pegon 2.334
Pehlevi → Pahlavi
pejorative 1.637
Peled, Yishai 1.49–50, 1.353–355, 1.555–559, 2.290, 2.477–479, 4.447, 4.737–740
Pellat, Charles 1.618, 2.406, 2.578, 2.629–630
Pellegrini, Giovan Battista 2.454–456, 4.216
Pencheon, Tom 2.710
Penn Arabic Treebank 1.460
Pentateuch, translation of 2.23, 2.528
Perdue, Clive 1.485
Pereira, Christophe 1.457, 3.52–58, 4.548–556
Perestroika 4.450
Pérez Lázaro, José 2.630
perfect 1.44, 1.196, 1.198–201, 1.217–218, 1.220–221, 1.309, 1.568, 1.574, 1.624, 3.67, 3.96
perfect, experiential 1.248, 3.333
perfect, *k*- → *k*-perfect
perfect, resultative 2.248
perfect aspect 1.331, 3.186
Perfect Grid 3.617
perfective aspect 1.95, 1.195, 1.197, 1.199, 2.171, 2.423, 2.524, 2.596, 3.247, 4.454, 4.456–457, 4.620, 4.697
perfectivity 4.457
Perfetti, Charles A. 3.77
performance, bilingual 1.369
performance, linguistic 4.46
performative 1.284, 2.359–360, 3.248, 3.263, 3.570–573, 4.328, 4.331–332
performative, commissive 3.571
performative, directive 3.571–572
performative, explicit 4.328
performative, expressive 3.571
performative, nonexplicit 4.328
performative, representative 3.571, 3.573
performative declaration 3.571–573
performative of order 2.359
performative verb 4.329
peripatetic communities 2.217, 2.219
peripheral 1.416
peripheral constituent 1.41
peripheral dialect 3.672
peripherality 4.742
periphery 4.734
periphery, left 4.746
periphery, right 4.746
periphrasis 1.491, 2.578
periphrastic construction 2.684
periphrastic genitive 1.82
periphrastic narrative 3.285, 3.536
Periphrastic Verbal Construction 1.38

Perkell, Joseph 2.365
Perkins, Revere D. 2.193–194, 2.321
Perlmutter, David M. 3.705, 3.710, 4.535
Perlmutter's Unaccusative Hypothesis 4.491
perlocutionary 2.360, 4.328, 4.472
perlocutionary effect 4.333
perlocutionary force 1.657
permansive 4.176
permutation 2.280, 3.21
permutative 4.221
Perry, John R. 2.344, 2.410–411, 3.573–580, 4.424–425, 4.451, 4.606
Persatuan Islam 2.337
Persian 1.24, 1.29, 1.33–34, 1.110, 1.126, 1.133–135, 1.167, 1.171, 1.173–174, 1.181, 1.189, 1.233, 1.240–242, 1.260, 1.275, 1.379–381, 1.402–404, 1.487, 1.494, 1.529, 1.554, 1.590, 1.617, 1.619, 1.665, 2.9, 2.26, 2.30, 2.60, 2.199–200, 2.210–212, 2.219, 2.221, 2.235, 2.265, 2.286, 2.325–326, 2.329, 2.334, 2.340, 2.344, 2.373, 2.469, 2.504, 2.560, 2.571, 2.605–607, 2.609–610, 2.628, 2.670, 3.6, 3.21, 3.59, 3.72, 3.131, 3.133–134, 3.346, 3.370, 3.501–502, 3.504, 3.508, 3.510, 3.573–584, 3.597, 4.377, 4.434–435, 4.450, 4.452, 4.476, 4.599, 4.601, 4.608, 4.616
Persian, Classical 2.407, 2.409–410
Persian, Middle → Pahlavi
Persian, Modern → Farsi
Persian, Old 3.478, 3.583
Persian, plural marker in 2.409–410, 3.578
Persian in India 4.595
Persian writings, translation of 1.402
Persianism 2.631
Persianization 4.596
persistence 2.196
personal pronoun 1.302, 2.82, 2.161, 2.484, 2.510, 2.521, 2.561, 2.574, 2.591, 2.613, 3.149–150, 3.182, 3.584–593, 3.713–714, 3.716, 4.110
perspective 1.422–423
perspective, writer's 2.208–209
Persson, Andrew 2.271, 2.560
Persson, Janet 2.271, 2.560
Persson, Maria 4.362
pesantren 2.336, 2.338
pesantren salafiyah 2.338
Pesetsky, David 4.488, 4.490, 4.721
pəṯāḥā 2.232
Peters, Johannes R.T.M. 2.541
Petersen, Jennifer 1.416
Peterson, Jennifer Leigh 2.8
Petherick, John 1.519
Petra 1.126, 2.198, 3.402, 3.404, 3.467
Petráček, Karel 1.86, 1.644, 4.94, 4.481
Petté 1.334
Pettorino, P. 4.669–670
Peul → Fulfulde
PF interface 1.481
pharyngeal 1.341–342, 1.376, 1.378, 1.390, 1.539, 2.55, 2.229, 2.685, 3.48, 3.65, 3.180, 3.599, 3.608–609, 4.667–669, 4.715
pharyngeal cavity 4.670
pharyngealization 1.103, 1.276, 1.278, 1.302, 1.304, 1.490, 1.525, 1.609, 2.365, 2.460, 2.560, 3.65, 3.181, 3.275–276, 3.290, 3.525, 3.527, 3.599, 4.107, 4.118, 4.171, 4.289, 4.422, 4.427, 4.564, 4.636, 4.670–671
pharyngealized 4.669–670, 4.678
pharyngeals, loss of 1.102
pharynx 2.229, 2.365, 3.48, 3.594, 3.599, 3.610
phasal aspect 1.221
phatic communion 2.202
Φ-*features* 1.583, 2.174
Philadelphia 2.506
Philipinos 4.123
Philippaki-Warburton, Irene 4.539
Philippi's Law 1.103
Philippus Arabs 4.402

philosophy, Greek 3.119, 3.176
Phoenician 1.293, 1.603, 2.23, 2.59, 2.66–67, 3.52, 3.409–410, 3.412, 3.414–415, 3.418, 3.731, 4.303, 4.313, 4.511, 4.571
Phoenician past of Lebanon 3.348–349
Phoenician script 1.148, 1.150, 3.409, 3.411
phonaestheme 4.254–255
phonation 3.124
phonation type 4.666–667
phōnē ásēmos 3.547
phonēen 2.235
phoneme 2.237–238, 3.607
phoneme, marked 2.685
phoneme, overdifferentiation of 2.684
phoneme, underdifferentiation of 2.684
phoneme substitution 2.685
phonemicization 3.605
phonemization, secondary 1.391
phonetic cue 3.206
Phonetic Form 4.489
phonetics 3.593–603
Phonological Form 1.480, 2.368, 2.450, 3.228, 3.232
phonology 3.607–615
phonology, autosegmental 3.66, 3.301, 3.626
phonology, generative 3.491
phonology, lexical 2.37, 3.496
phonology, metrical 3.615–623
phonology, moraic 2.38
phonology, nonlinear 3.615
phonology, postlexical 4.115
phonotactics 2.483, 3.624–628
phrasal compound 2.412
phrase, adjective 1.507, 4.653
phrase, adverb 1.507
phrase, adverbial 4.531
phrase, agreement 3.684, 4.743
phrase, annexive 2.146
phrase, broken 4.486
phrase, circumstantial 3.703–704
phrase, circumstantial → *ḥāl*
phrase, complement 1.635
phrase, complementizer 4.362, 4.537, 4.718, 4.722, 4.743
phrase, conjunctive 1.473
phrase, determiner 1.480, 4.743, 4.745
phrase, epithet 4.83
phrase, infinitive 1.664
phrase, inflection 4.719, 4.743
phrase, inflectional 1.510
phrase, intermediate 2.396–397
phrase, intonational 2.397–398
phrase, locative 4.220
phrase, modal 1.510, 4.746
phrase, mood 4.743
phrase, negative 4.743
phrase, noun → noun phrase
phrase, possessive noun 1.582
phrase, prepositional → prepositional phrase
phrase, quantifier 3.259
phrase, tense 3.684
phrase, topic 4.743, 4.746
phrase, verb → verb phrase
phrase, verbal noun 3.259
phrase structure 4.747
Phrase Structure Grammar 4.521
phrase structure rules 1.457, 4.521
phrase structure rules → rules, phrase structure
phraseology 3.201, 3.628–634
phúsis 3.176, 4.684
Piamenta, Moshe 1.444–445, 2.258, 2.260, 2.272–273, 2.462, 2.481, 2.491, 2.534, 3.361, 3.658, 4.75, 107, 4.750
Pianel, Georges 2.470–471
picture-naming task 2.682
pidgin 1.37, 1.85, 1.518–521, 1.526, 1.595, 1.665, 2.214–215, 2.517, 2.588, 2.672, 3.12, 3.58–59, 3.61, 3.127, 3.634–644
pidgin Arabic 1.364, 1.619, 3.634–639, 4.123
Pidgin Arabic, Chad 3.61
Pidgin Gulf Arabic 2.211, 3.63, 3.643
Pidgin Sudanese Arabic 3.62
pidgin, expanded 3.639, 3.642
Pidgin, Persian Gulf 1.518
pidginization 1.364, 1.485, 1.489, 1.518–519, 1.593–595, 1.619,

2.119, 2.255, 2.265, 2.267, 3.58, 3.62, 3.478, 3.634–635, 3.639–644, 4.218
pidginization, 'first' 3.643
Piedmont 4.215
pied-piping 2.20, 2.390, 4.66–67
Pienemann, Manfred 1.485, 2.679, 4.142
Pierrehumbert, Janet B. 2.395, 3.206, 3.463, 3.610, 3.624–625
Pierret, Roger 2.249
Pig Latin 2.217, 4.156, 4.159
Pike, Kenneth L. 3.725
pilgrimage 3.179
pilgrimage → *ḥajj*
Pinker, Steven 3.74, 3.78, 4.46–47
Piplāi, Bipradās 1.287
Pīr-i Bakrān 2.600
Pisa 2.454–455
Pisowicz, Andrzej 2.406, 2.409
pitch 1.211, 1.213, 2.395–396, 2.400, 2.590, 2.593, 4.667
pitch, high 1.340
pitch, tonal 1.367
pitch accent 2.395, 2.397–399
pitch accent, final 2.396
pitch contour 2.395
pivot 4.318–319
place names 4.371, 4.508–515
place names, Akkadian 4.513
place names, Arabic in Sicily 4.215
place names, Aramaic 4.512–513
place names, Berber 1.296, 4.511, 4.508
place names, Canaanite 4.513
place names, Coptic 4.512
place names, Greek 4.512–513
place names, Hebrew 4.513–514
place names, Latin 4.512
place names, Minaean 4.513
place names, Phoenician 4.513
place names, Sabaean 4.513
place names, Sicilian 4.215
place names, Spanish in Tunisia 4.574
place names, Sumerian 4.513
place names, Turkish 4.514
place names, Urdu 4.601
place names in Iberian Peninsula 4.508
place names in poetry 4.513–514
place names in Tunisia 4.573
place names in Turkey 4.508
place of articulation 3.597, 3.625
Placement Rules 2.147
Plant, Geoff 2.677
Plantin, Christophe 1.169
plants, names of 1.504, 2.26–27, 2.283, 2.289
Plascov, Avi 2.499
plateau contour 2.399–400
plateau tune 2.397–398
Plato 2.65, 3.176, 4.459
Platts, John T. 4.602
Platzack, Christer 4.540
pleonasm 4.35
Plonka, Arkadiusz 3.349
Plunkett, Bernadette 2.173, 3.382, 3.384–385, 3.685, 4.527, 4.524, 4.650, 4.653, 655, 4.743, 4.746
Plunkett, Kim 2.689
plural 1.433–434, 1.562, 1.576, 3.439
plural, acquisition of 2.101
plural, Bantu 2.670
plural, broken 1.237, 1.560–561, 1.576, 1.664, 2.352, 3.154, 3.189, 3.191, 3.243, 3.302–304, 3.329, 3.430, 3.440–447, 4.94, 4.174–175, 4.309
plural, deictic 1.550
plural, English 3.154
plural, external 1.250, 1.560, 4.309–311
plural, Hebrew suffix 1.237
plural, iambic 3.304–305
plural, internal 1.38, 1.116, 1.250, 1.560, 1.653, 2.352, 2.616, 4.174, 4.309, 4.311
plural, internal 2.616
plural, masculine 1.646
plural, paucal 3.430
plural, Persian 2409, 3.578
plural, quadriradical 1.646, 2.354
plural, quadrisyllabic 1.646
plural, sound 1.560, 3.430, 3.439–440
plural, sound masculine 1.644

plural ending 2.402
plural formation 1.119, 2.76, 2.684–686, 4.309
plural marker 2.410, 2.592
plural marker, Dutch 2.686
plural marker, Persian 2.410
plural of abundance 1.230, 1.434
plural of large number 2.352
plural of multiplicity 3.439
plural of paucity 1.434, 3.439, 3.450
plural of plenty 3.430
plural of plural 2.352
plural of small number 2.352
pluralis majestatis 4.25
plurality 1.305, 1.329, 1.358, 1.433, 1.523, 3.450
plurality, phase-internal 4.51
plurality, verbal 4.50
plurals, 'productive' 3.445
plurals, broken, in Northwest Semitic 3.414
P-marker 4.521–522
Pococke, Edward 1.167, 1.172
Podor 4.179
poem, prose 4.91
Poerbatjaraka 2.336
poetic jargon 3.646
poetic koine → koine, poetic
poetic language 1.625
poetic lexicon 3.645
poetic license 1.192, 2.177, 2.184, 2.404, 2.541, 2.543–544, 3.90, 3.94, 3.116, 3.453, 3.645, 3.647–652, 3.94, 4.32, 4.200, 4.558
poetic license → *ḍarūrat aš-šiʿr*
poetics, Arabic 2.443
poetry 1.399, 1.402, 1.456, 1.512, 1.528, 1.546, 1.597–598, 1.602, 1.615–616, 1.651, 2.176–177, 2.364, 2.446, 2.541, 2.543–544, 2.546, 3.4, 3.118, 3.173, 4.191, 4.199–214
poetry → *šiʿr*
poetry, Abbasid 2.497
poetry, Andalusian 4.59, 4.92
poetry, Bedouin 1.273, 3.646
poetry, Classical Arabic 3.499
poetry, colloquial 1.601
poetry, dialectal 1.241
poetry, Greek 3.499, 4.210
poetry, Hebrew 4.59
poetry, Jāhiliyya 1.617
poetry, love 4.202, 4.205
poetry, Mauritania 3.172
poetry, Modernist 4.202, 4.204, 4.211–213
poetry, Moorish 3.173
poetry, Nabaṭī 1.269, 1.273, 3.500
poetry, Najdī 3.500
poetry, oral 3.646, 4.90
poetry, Ottoman 4.586
poetry, Palestinian 4.77
poetry, Persian 1.404
poetry, popular 4.75
poetry, pre-Abbasid 3.35
poetry, pre-Islamic 2.183, 2.198–199, 2.264, 3.10, 3.30, 3.210, 3.500, 3.644–646, 3.648, 3.689
poetry, recitation of 3.173
poetry, Romance 4.59
poetry, Serbo-Croatian 3.499
poetry, Sufi 4.75, 4.204, 4.211, 4.214
poetry, Swahili 1.665, 4.381
poetry, vernacular 2.414, 2.609, 3.348
poetry, wine 4.202, 4.205, 4.208–209
poetry, Yemeni 3.500
poets, female 1.529
pointing 3.111
poíon 4.220
polarity 1.121, 1.216, 1.328, 2.145, 3.652–658, 4.18
polarity, gender 1.251, 1.491, 3.450, 3.652
polarity, negative 1.220, 1.481, 1.508, 3.358, 3.387, 3.652–658, 3.681
Polarization, Law of 1.121
polarization, stylistic 1.408
polemics 2.303
Polish 3.726
Polish, Arabic loanwords in 4.259–267

politeness 1.657, 2.69–70, 2.202, 2.205, 2.587, 3.658–664, 3.679, 3.719, 4.419
Politeness Principle 2.70
political discourse 3.663–671
Pollak, Isidor 3.176
Pollock, Jean-Yves 1.348, 4.394, 4.398, 4.743
Polomé, Edgar C. 1.660, 1.663–665, 4.382–383, 4.385
polygenesis 3.640
polygrammaticalization 2.193, 4.39–40
polymorphemic 1.453
polyptoton 3.538, 3.633, 3.733
polysemy 1.626, 2.496, 3.45, 3.49, 3.320–323, 3.538–539, 4.164–168
polysemy, antithetical 3.45, 3.49–50
Ponnani 3.129
Ponty, William 1.338
Poonawala, Ismail K. 2.325, 2.328
Pope, Nicole 4.578
Poplack, Shana 1.415–416, 418
popular literature 1.57
population mobility 4.322, 4.324, 4.326
Port Said Arabic 1.604
Port Sudan 4.376
portal, Arabic 2.381
Porten, Bezalel 3.410
portmanteau 1.1, 1.452, 3.629
portmanteau merging 4.463
Portuguese 1.241, 1.665, 2.26, 2.282–289, 2.333, 2.609, 2.671, 3.1, 4.476
Porxomovskij, Viktor 4.304
position, clause-initial 4.718, 4.722
position, COMP 4.718
position, left-dislocated 4.395
position, postverbal 4.395
position, preverbal 4.394–395, 4.731
position, sentence-initial 4.718, 4.722
position, specifier 4.396
possessed 1.351
possession 1.84, 1.110, 1.183–184, 1.226, 1.230, 1.240, 1.248, 1.279, 1.283, 1.285, 1.344, 1.396, 1.563, 2.248, 2.295, 2.489, 2.594, 2.615, 3.80–81, 3.587, 3.671–676, 3.701, 4.15–16
possession, inalienable 1.248
possession, transitive 3.674
possession exponent 3.671
possession exponent → genitive exponent
possessive 1.302, 1.326, 1.391, 1.417, 2.99, 2.161, 2.273, 2.297, 3.80–83, 3.86, 3.277, 3.365, 3.432
possessive clause 1.285
possessive construction 1.350, 2.594, 3.67–69, 3.85, 4.555
possessive construction, Turkic 2.405, 3.67
possessive marker 1.540
possessive particle 2.161
possessive pronoun 1.185, 1.187, 2.521, 2.591, 3.182, 4.689, 4.712
possessive suffix 1.247, 2.561, 2.613
possessor 1.351, 3.232, 3.432, 4.487, 4.492
possibility 1.248
Postal, Paul 4.54
postclitic 2.21–22
post-creole continuum 1.525
Postel, Guillaume 1.167–169, 1.171
postmodifier, clausal 4.746
postposition 1.34
potential 1.553
Pott, August F. 2.220, 4.50–51
Potter, Terrence M. 4.467, 4.469
Poulisse, Nanda 4.271
Pouplier, Marianne 4.335
Pouss 1.334
power distribution 4.325
power, male 2.642
PP → prepositional phrase
PP-fronting 1.413
praedicatum 3.113
pragmatic activation 4.734
pragmatic axiom 2.19
Pragmatic Highlighting Principle 2.147
pragmaticalization 1.476

pragmaticization 2.192
pragmatics 1.429, 3.676–683
Prague 1.460
Prague Arabic Dependency Treebank 1.458
Prague School 1.458, 1.515–516, 4.472, 4.484, 4.494, 4.502, 4.726
praise song 4.716
Prasse, Karl-Gottfried 2.68
Prätor, Sabine 3.504
Pratt, Marie-Louise 1.636
prayatna 3.124
prayer 2.136
prayer, Friday 2.649
prayer, language of 1.529
prayer, public 2.647
precative, Akkadian 4.557
predicand 2.562, 2.570, 4.495, 4.497, 4.499
predicate 1.16, 1.18–20, 1.71, 1.182, 1.184, 1.349–350, 1.413, 1.428, 1.436, 1.485–486, 1.507, 2.186, 2.431, 2.434, 2.537, 2.539, 2.550–551, 3.81, 3.113–115, 3.381, 3.385, 3.683–689, 4.354, 4.484, 4.495, 4.497, 4.499, 4.737
predicate → *xabar*
predicate, change of state 1.198
predicate, clausal 1.203, 1.354
predicate, collective 3.69
predicate, complex 1.488
predicate, derived 3.688
predicate, developmental 1.199
predicate, embedded 2.132
predicate, inceptive 1.199
predicate, interval 1.198
predicate, locative 1.184, 3.80
predicate, matrix 4.364
predicate, momentaneous 1.198
predicate, natural 4.485
predicate, nominal 3.113–114, 3.177
predicate, non-stative 1.198
predicate, non-verbal 1.183
predicate, one-place 3.685
predicate, phrasal 1.203
predicate, punctual 1.199
predicate, secondary 2.132
predicate, small-clause 2.132
predicate, telic 1.198
predicate, transformed 2.227
predicate, two-place 2.584
predicate, verbal 1.622, 3.113–115, 3.177
predicate focus 2.113–114
predicate raising 2.133
predicate/argument structure 4.489
Predicate-internal Subject Hypothesis 3.685
predication 1.12, 1.182, 1.184, 1.186, 2.145–146, 2.291, 2.431, 2.435, 3.683, 3.685, 3.687
predication, clausal 3.386
predication, incomplete 1.220
predication, possessive 1.183, 3.675
predication, primary 3.685, 3.688
predication, relational 4.624
predication, secondary 3.685
predication, topic/comment 4.486
predication phrase 3.685
predicative 1.16, 19
predicative adjective 1.12–13
predicatoid 1.285, 3.675
predicator 3.684
predictability 4.46
predictability, semantic 4.47
predictability, syntactic 4.47
predictive 3.247
prefix 1.208, 1.454, 1.493, 2.21, 2.350, 2.547, 3.203, 3.297, 4.462
prefix conjugation 1.36, 1.491, 2.23–24, 3.191, 4.454–457
prefix conjugation in Northwest Semitic 3.416
prefix conjugation, Semitic 4.177
prefix, modal 1.491
prefix, verbal 1.33, 1.196
prefixation 1.576
prefixed form 4.638–639
pre-Hilalian Arabic 1.53, 1.58–59, 1.609–610, 2.532, 3.53–54, 3.273–274, 3.288

pre-Islamic Arabic 1.399–401, 1.491, 1.562, 1.593–594, 1.613–614, 1.616–618, 1.653, 2.88–89, 2.198, 2.263, 2.355, 3.14, 3.90, 3.465, 3.644, 3.689–699
prelinguistic stage 2.97
prenasalization 1.256
preposing 2.393
Preposing, NP 2.172
preposition 1.21, 1.72, 1.183, 1.185–186, 1.240, 1.249, 1.303, 1.327, 1.348, 1.386, 1.392, 1.524, 1.549, 1.567, 1.585, 1.641, 1.644, 2.95, 2.99, 2.146, 2.161, 2.171, 2.193, 2.237, 2.243, 2.274, 2.512, 2.522, 2.592, 3.183, 3.221, 3.669–703
preposition, compound 3.221, 3.700
preposition, enclitic 2.391–392
preposition, nominal 2.424
preposition, primary 3.699, 3.701
preposition, secondary 3.699–701
prepositional 3.700
prepositional dative construction 4.538
prepositional idiom 3.201
prepositional phrase 1.21, 1.24, 1.331, 1.413, 1.432, 3.431, 4.354, 4.436, 4.438–439, 4.531, 4.541, 4.653, 4.734–735
prepositions, nominal origin of 3.700
prescriptive rule 2.178
present, habitual 1.396
present, immediate 1.396
present, narrative 1.553
present, Semitic 4.178
present continuous 1.229, 2.516
present tense 1.39, 1.195, 1.200–201, 1.331, 1.349, 1.508–509, 1.566.1.568, 2.91–92, 3.114, 3.186
present tense marker 2.423
present tense of address 1.284
presentative 1.116, 1.246–247, 1.279, 1.303, 1.327, 1.387, 1.392, 1.549, 2.243, 2.354, 2.511, 2.614, 3.277, 3.364–365, 3.368, 3.483, 3.530, 3.703–705, 4.111, 4.246, 4.486, 4.657, 4.714
press 1.175–176, 1.654, 2.11, 2.267
press → newspapers
press, Arab-American 3.397
prestige 1.440–441, 1.667, 2.29, 2.31, 2.35, 2.275–276, 2.278, 2.654, 3.10–12, 4.3, 4.321, 4.628–629
prestige, covert 3.12
prestige, overt 2.655, 3.12
prestige, social 2.532
prestige norm 2.33
prestige pronunciation 2.495
prestige variety 2.657
Preston, Dennis R. 1.585, 2.384
presupposition 2.113, 4.503
presupposition, pragmatic 4.503
presuppositional background 4.58
pretense 3.184
preterite, Akkadian 4.178
preverb 1.295, 2.271, 4.306
preverbal marker 1.218
preverbal particle 3.221
prime 1.425
priming 1.425
primitive, morphological 1.208
Prince, Alan S. 2.153–155, 2.367, 3.299, 3.303–305, 3.423, 3.430, 3.440, 3.445, 3.491–493, 3.496, 3.614–618, 3.621, 4.269, 4.339, 4.345–346, 4.389, 4.557–558
Prince, Ellen F. 4.316, 4.502
Principle, Projection 4.488
Principle, Root-and-Pattern .76, 1.376, 3.425, 3.440, 4.461, 4.640
Principle, Unaccentable Element 2.113
Principle of Compositionality 2.368
Principle of Definite First 4.734
Principle of Increasing Complexity 2.147
Principle of Pragmatic Highlighting 2.147
Principle of Relevance 4.503
Principle of Separation and Non-identity 2.226, 2.453, 3.107
Principle of Unidirectionality 2.192, 4.39
Principles and Parameters Theory 3.228, 3.684, 3.706, 4.391–394,

4.396–399, 4.521, 4.647–648, 4.650
printing 1.603
private space 2.645
pro 1.348, 1.350, 3.706, 3.709–710
pro, expletive 3.710
probe 3.230
procedure, category 4.142
procedure, phrasal 4.142
process 1.197
process nominal 1.352
processability model 2.679
Processability Theory 4.142–144
processing, auditory 2.689
processive 1.199
Procházka, Stephan 1.388–397, 1.585, 2.158, 2.160, 2.313, 3.699–703, 4.50, 4.285–286, 4.532, 4.579–581, 4.589–594
Prochazka, Theodore 1.446, 2.152, 3.606, 4.127–128, 4.305, 4.750
proclitic 2.21–22, 3.298
pro-constituent 1.354
pro-drop 1.486, 3.228, 3.386, 3.561, 3.705–714, 4.46, 4.49, 4.356–357, 4.395, 4.524, 4.527, 4.653, 4.656, 4.729
pro-drop language 2.321
production error 2.676, 2.691
profanity 4.416
professionals 1.328
professions, names of 1.296, 3.279, 3.303
proficiency 2.117
proficiency, language 4.150–151
pro-form 3.84, 3.86
pro-form, existential 3.83
pro-form, forward-referring 1.353
Progovac, Liljiane 3.653, 3.655
progressive 1.42, 1.197, 1.199, 1.220, 1.281, 1.367, 1.523, 1.553, 3.186–187, 3.247, 3.266, 3.573, 4.569
prohibition 2.23, 2.491, 3.267
prohibitive 4.639
prohibitive statement 4.331
Projection Principle 4.488
projection, functional 4.392, 4.398, 4.743
projection, inflectional 4.657
projection, IP 4.657
projection, lexical 4.392
projection, maximal 1.481, 4.743, 4.747
projection, minimal maximal 2.170
projection, null 4.398
projection, syntactic 4.398
projection, Tense 4.396, 4.398–399
projection, verbal 4.399
projection, VP 4.395
Prokosch, Erich 2.517–518, 2.670, 2.672, 4.584, 4.587, 4.592–593
proleptic suffix 1.395
prominence 2.114, 2.396, 2.400, 3.704, 3.725–726
prominence hierarchy 2.395, 2.400
Prominence view 1.423
pronominal 1.310–311, 1.350, 2.432, 3.713
pronominal suffix 2.22
pronominal suffixes, Northwest Semitic 3.413
pronominal system 4.304
pronominalization 1.354, 1.432, 2.208, 2.300–301, 3.710, 3.713–717, 4.525
pronoun 1.311–312, 1.377, 1.555–559, 2.300, 3.589
pronoun, accusative 1.581
pronoun, adnominal 1.312
pronoun, anaphoric 1.353–354, 4.101
pronoun, anticipatory 1.353
pronoun, bound 3.590–591
pronoun, cataphoric 1.353–355, 1.552–553, 2.149
pronoun, concealed 4.101
pronoun, copular 3.385–386, 3.547, 3.740
pronoun, demonstrative 1.226, 1.384, 1.559–560, 1.569–573, 1.608, 1.611, 2.484, 2.574, 3.151, 3.220, 3.246, 3.329, 3.645, 3.694, 4.110, 4.689

pronoun, direct object 1.236, 3.182, 3.588, 4.689
pronoun, dummy 2.316
pronoun, expletive 3.560, 4.657
pronoun, full 1.485
pronoun, implicit 2.83, 2.94, 2.552, 4.357
pronoun, incorporated 4.395
pronoun, indefinite 1.384
pronoun, independent 1.247, 2.120–121, 3.531
pronoun, indirect object 1.236
pronoun, interrogative 1.30, 1.106, 1.116, 1.226, 1.247, 1.327, 1.384, 1.392, 1.549, 1.566, 2.143, 2.243, 2.387–392, 2.511, 2.530, 2.574, 2.614, 3.151, 3.182, 3.484, 3.531, 4.111, 4.553, 4.566, 4.689
pronoun, negative 3.386, 3.586
pronoun, non-anaphoric 1.558
pronoun, Northwest Semitic 3.413
pronoun, null 3.710, 4.400, 4.657
pronoun, personal 1.30, 1.105, 1.114, 1.226, 1.247, 1.302, 1.326, 1.391, 1.548, 1.555, 1.569, 2.82, 2.161, 2.484, 2.510, 2.521, 2.561, 2.574, 2.591, 2.613, 3.149–150, 3.182, 3.584–593, 3.713–714, 3.716, 4.110
pronoun, possessive 1.185, 1.187, 1.230, 1.283, 1.579–581, 2.521, 3.182, 4.110, 4.689, 4.712
pronoun, proleptic 1.540, 4.544
pronoun, reciprocal 1.523
pronoun, reflexive 3.220, 4.67, 4.616
pronoun, relative 1.13, 1.91, 1.106, 1.116, 1.226, 1.239, 1.247, 1.263, 1.273, 1.392, 1.399, 1.491, 1.549, 1.559–560, 1.566, 1.606, 1.639, 1.668, 2.161, 2.195–196, 2.243, 2.271, 2.484, 2.530, 2.537, 2.614, 3.151, 3.192, 3.245, 3.261, 3.277, 3.365, 3.403, 3.484, 3.531, 3.636, 3.694–695, 4.62–63, 4.70–72, 4.111, 4.187–188, 4.221, 4.236, 4.246, 4.392, 4.552, 4.622, 4.689
pronoun, resumptive 1.49, 1.279, 1.303, 1.312, 1.350, 1.493, 1.552, 2.132, 2.148, 2.171–172, 2.296, 2.595, 3.220, 3.383, 3.385, 3.560, 3.589, 4.60, 4.65, 4.68, 4.71, 4.82, 4.186, 4.366, 4.485, 4.525–527, 4.720–722, 4.738
pronoun, subject 1.236, 1.350
pronoun deletion 3.705
pronoun doubling 1.417–418
pronoun incorporation 1.46
pronoun shift 1.342
pronoun substitute 4.468
proper names 1.561, 1.580, 1.582, 1.643–646, 2.17, 2.586, 3.426, 3.717–724, 4.201, 4.468, 4.508, 4.686
proper names, female 1.643
proper names, Kurdish 2.604–605
proper names, Malay 2.341
proper names, Malayalam 3.132–133
proper names, *nisba* 3.379
proper names, Sicilian 4.217
proper names, transcription of 4.519
proper names, Turkish 4.579
proper names, Urdu 4.601
proper names, Uyghur 4.607
proper names, Yoruba 4.760
proper names in India 2.329
proper names in Indonesia 2.334
proper names in Malaysia 2.344
proper names in Senegal 4.183
proper names in Tunisia 4.573
property, morphological 1.208
property assignment 1.12
prophethood 2.303–304
proportioned writing 4.131
proposition 1.343–344, 1.430–431, 1.471, 2.113, 2.186, 2.207, 2.209, 3.176–177
propositional content 1.485, 3.679, 4.328
prose 1.528, 1.598
prosodeme 2.237
prosodic circumscription 3.304
prosodic feature 1.205, 1.369

prosodic hierarchy 2.395
prosodic peak 2.113
prosodic template 2.366–367
prosodic tier 3.225
prosodic unit 1.480–481
prosody 1.515, 2.395, 3.647, 3.665, 3.724–728, 4.502
prosody, Persian 3.574
prosody generation 1.210–211
prospective 1.220
prosthesis → prothesis
prosthetic vowel → prothetic vowel
protasis 1.468, 2.477–479, 3.261
prothesis 1.85, 3.728–730
prothetic vowel 1.86, 1.235, 1.243, 1.261, 1.277, 2.14, 2.611
Proto-Afro-Asiatic 1.313–314, 2.67, 3.446
Proto-Arabic 1.313, 1.490, 1.494, 2.263, 2.495, 3.411
Proto-Hamito-Semitic 1.314
protolanguage 1.490, 2.66, 3.188, 4.373
Proto-Northwest Semitic 3.410
Proto-Semitic 1.313–314, 1.358, 1.467, 1.544, 1.563, 1.640–641, 1.643–644, 2.23, 2.67, 3.188, 3.227, 3.411, 3.446, 3.568, 3.591–593, 4.97, 4.170–173, 4.175–176, 4.178, 4.302
prototype 2.180
Provençal, Old 2.286
proverb 1.286, 1.554, 2.136, 2.177, 2.444, 2.607–608, 3.35, 3.174, 3.179, 3.540, 3.629, 3.730–736
proverbium 3.730
Provincia Arabia 1.126, 1.131
proximate intent 1.246, 1.253
Prunet, Jean-François 1.425, 2.123, 2.692, 3.305, 4.96, 4.271, 4.286
pseudoclassicism 3.220
pseudoclassicization 1.407
pseudocleft 1.197, 2.145
pseudocomplementation 4.197
pseudoconstruct 3.431–432, 3.684
pseudocorrection 2.266, 2.275–276, 2.278–279, 3.217
pseudodual 1.107, 1.280, 1.304, 1.328, 1.393, 1.550, 1.616, 2.244, 2.512, 3.67, 3.154, 3.276, 3.279, 3.366, 3.446, 3.736–739, 4.554, 4.616
pseudoimpersonal construction 1.554
pseudoliterary feature 2.530
pseudoloan 3.577
pseudolongitudinal research 2.683
pseudonym 3.718
pseudopreposition 4.93
pseudoverb 1.183, 1.185–186, 1.268, 1.553, 1.564, 2.197, 2.249, 2.491, 3.68, 3.675, 3.739–740, 4.39, 4.695
psycholinguistics 2.96
ptōsis 1.559
public space 2.645
Publilius Syrus 3.730
publishing 1.532
Puech, Emile 3.472
Pulaar → Fulfulde
Pular 2.137
Pullum, Geoffrey K. 1.458, 4.193, 4.197
Pumbeditha 2.526
pun 2.496, 3.538
punctual 1.199
punctuation 1.152, 1.207, 1.211, 2.332–333, 3.240, 3.740–742, 4.192
punctuation, Arabic Braille 1.318
Punic 2.66–67, 3.52, 3.141, 3.409, 3.738, 4.511
Punic, Neo- 3.409, 4.215
Punic Theory 3.141
Punjabi 2.210, 3.506, 4.1
Punt 4.275
Pure Grid Theory 3.617–618, 3.622–623
purism 2.10, 2.632, 2.674
purism in Turkey 4.586
purpose 1.429
purpose, communicative 4.502
purpose, speaker's 2.80
push-chain mechanism 4.1–2
Pustejovsky, James 4.166
Puthi 1.288

Q

/q/ 1.54, 1.59–61, 1.87, 1.112, 1.223, 1.232, 1.260, 1.270, 1.272, 1.275–276, 1.323, 1.365–366, 1.376, 1.390, 1.407–410, 1.440–441, 1.546, 1.584–586, 1.589, 1.605, 1.607–612, 2.3–4, 2.7, 2.99, 2.212, 2.241, 2.263, 2.270, 2.275–276, 2.322–324, 2.409, 2.414, 2.416, 2.481, 2.494, 2.502, 2.507–508, 2.533, 2.610, 2.656–657, 3.10, 3.12, 3.124, 3.147, 3.290, 3.402, 3.497, 3.527, 3.607, 3.666, 3.668, 4.1–3, 4.408, 4.548, 4.564, 4.581, 4.613, 4.687, 4.709
qabīh 2.184, 2.542
qabīḥ al-kalām 2.361
Qaboos, Sultan 2.211
qad 1.217–219, 1.248, 1.468, 1.509, 1.648, 2.208–209, 3.248, 3.257, 3.326, 3.548, 4.124, 4.544–545, 4.732, 4.740
qaddara 4.447
qad̲f 2.361–363
Qāḍī ʾAḥmad 2.598, 4.562
Qāḍī, Muḥammad Yūnus al- 1.601
Qāḍī, Wadād al- 4.74
Qadīm 1.380
Qādirī, Murād al- 1.603
Qādiriyya 1.380, 2.53, 4.180, 4.277
Qādisiyya, al- 1.618
qaḍiyya ḥamliyya 3.176
qadr 2.362, 4.446
Qadru 1.127
qāf 4.1–3
qāf ġayr maʿqūda 4.2
qāf maʿqūda 4.2
Qafisheh, Hamdi A. 2.36, 3.655, 4.646, 4.653, 4.656, 4.695, 4.697
qāfiya 1.150, 3.210, 4.86–87, 4.105
qāfiya muqayyada 2.404
qāfiya muṭlaqa 2.404
Qahabah Arabic 4.127
Qaḥṭān 1.617, 3.345, 3.467
Qaḥṭanīye, il 4.402
Qāʾid, Yūsuf al- 1.599
Qāʿida, al- 4.78
qāla ʿalā 3.114
qāla as complementizer 4.198
qāla fī 3.114
qalam 2.40, 3.516, 4.132
qalam → ʾaqlām
qalam ad̲-d̲ahab 4.561
qalam al-ʾašʿār 4.562
qalam al-baṭāʾiq 4.134
qalam al-ḥalba 4.134
qalam al-ḥawāšī 3.340
qalam al-janāḥ 4.134
qalam al-maṣāḥif 3.309
qalam al-maṣḥif 3.308
qalam al-matn 3.340
qalam ar-riqāʿ 4.561
qalam aš-šiʿr 4.562
Qalamūn 4.402–403, 4.406
Qalamūn Arabic 1.538, 1.546, 1.584, 1.608, 4.403, 4.406, 4.408
Qalamūn Mountains 3.370, 3.372
Qalāʾūn, Sultan 2.41
qalb 1.71, 1.192, 1.203, 2.165, 2.280
Qalfāṭ 2.185

Qalhāt Arabic 3.488
Qālī, ʾAbū ʿAlī ʾIsmāʿīl ibn Qāsim al- 2.280
qalīl 1.192
Qaliyasqarûlï, Qabidolda 2.558
qalqala 4.426–427
Qalqašandī, Šihāb ad-Dīn ʾAḥmad ibn ʿAlī al- 1.152, 3.307–308, 3.340, 3.513–514, 3.718, 4.103, 4.132–134, 4.134, 4.159, 4.560–561
qāma 2.195–196
qāma bi- 3.200
Qamar ad-Dīn 2.326
Qamashi 4.613
Qāmišlī 4.402
Qāmišlī Arabic 2.389, 2.534, 4.407–408
Qandīl, Bayyūmī 1.600
Qāniya 2.256–257, 2.259
Qānṣawh al-Ġawrī 2.598
qarāba 4.221
Qarahḥisārī, Ahmad 2.601, 3.311, 4.561
qarawī 1.605
Qāriʾ, ʿAlī ibn Sulṭān Muḥammad al- 4.426
qarīb, maʿnā 3.538
qarīḍ 4.103–104
qarīna 2.80, 2.581, 3.119, 3.122, 4.103, 4.332
qarīna ḥāliyya 2.442
qarīna maqāliyya 2.442
qarīnat al-ḥāl 4.331
Qaritēn, il- 4.403, 4.408
Qaritēn Arabic 1.538, 2.314, 4.406
qarmaṭa 1.151
Qarqar 1.127
Qarṭajannī, Ḥāzim ibn Muḥammad al- 3.649, 4.333
Qarṭmīn 1.95
Qaryat al-Fāw 3.471
qāsa 4.12
qasam 1.202, 2.237, 3.549, 4.414, 4.429
qaṣd 1.429, 1.558, 2.80, 2.539, 2.541, 4.333
qaṣīd 2.629, 3.211
qaṣīda 2.329, 3.500, 3.646, 4.481, 4.596
qaṣīda sāsāniyya 4.255
qaṣīdat an-naṯr 4.91
Qāsim 3.326
Qāsim, ʿAwn aš-Šarīf 2.560, 4.375–377
Qāsim Arabic 3.606, 3.326
Qašqa-Dārya Arabic 4.613–614, 4.616–617
qaṭ 1.248, 2.292, 4.432
Qaṭabān 2.56, 2.257, 4.295
Qaṭabānian 1.127, 2.66, 3.593, 4.295, 4.301, 4.313
qatala 1.39
Qaṭar 1.176, 2.213, 2.215, 2.696, 4.123, 4.306
Qaṭar Arabic 2.609
Qaṭīf 4.124
Qaṭīf Arabic 4.124, 4.128
Qaṭn Arabic 4.693
qaṭṭ 4.133
qawl 2.539, 2.541, 2.579, 3.731
qawl maʾṯūr 3.629
qawl sāʾir 3.629
qawm 3.344
qawmiyya 3.344
qawsāni 3.742
Qayrawān → Kairouan
qayrawānī 3.112
Qayrawānī, ʾAbū Zayd al- 3.375
Qays 1.641, 2.88, 2.312, 4.376, 4.430
Qaysarānī, ʾAbū l-Faḍl Muḥammad ibn Ṭāhir al- 3.380
Qayyūm, Muḥammad ʿAbdul 1.287
Qazwīnī, Jalāl ad-Dīn al- 2.79, 2.81, 2.86, 2.358, 2.441, 2.445, 2.581, 3.120–121, 4.332, 4.473
Qazzāz, Muḥammad ibn Jaʿfar al- 3.648, 3.651
Qedar 1.127–128
qəltu Arabic 1.86–87, 1.90, 1.93, 1.181, 1.189, 1.222, 1.232, 1.234, 1.270, 1.396, 1.536–538, 1.584, 1.607, 1.610, 1.612, 2.313, 2.414–424, 2.532–533, 4.3, 4.70, 4.325
qētel Arabic 1.538
qibla ʾīwān 2.43
Qidru 1.127

Qilāʿī, Jibrāʾīl ibn al- 1.603
Qingjing mosque 1.379
Qipchaq → Kipchak
qirāʾa 2.281, 3.74
qirāʾa ʾimlāʾ 1.338
qirāʾāt 1.2, 1.399–400, 1.427, 1.492, 2.177, 2.301, 3.14, 3.89, 3.92, 4.4–11, 4.426, 4.558
qis 1.183
qiṣṣa 1.558
qiṭʿa 4.134
qiyās 1.191–194, 2.86–87, 2.176, 2.184, 2.200, 2.310, 2.360, 2.425, 2.448, 2.543, 2.629, 3.427, 4.11–14, 4.119, 4.444
qiyās šarʿī 4.13
qiyāsī 1.67, 2.177
Qōm Išqawh 1.151
quadriconsonantal → quadriradical
quadriliteral → quadriradical
quadriradical 1.55, 1.255, 1.307, 1.331, 1.452, 3.301, 4.50–51, 4.93, 4.244, 4.411, 4.558, 4.465, 4.568, 4.624, 4.640, 4.692
quadriradical noun → noun, quadriradical
quadriradical root 2.346, 3.462
quadriradical verb → verb, quadriradical
qualifier, circumstantial 2.224
qualities of God 2.432
quantificational phrase 4.82
quantificational relation 4.15, 4.19
quantificational relationship 4.16
quantifier 1.125, 1.188, 1.283, 1.351–352, 1.413, 1.477, 1.552, 1.579, 2.20, 2.146, 2.368, 2.489, 2.594, 3.221, 3.285, 3.428, 3.535, 4.14–20, 4.67, 4.81–82, 4.235
quantifier, nominal 4.15–16, 4.18–19
quantifier, numeric 4.18
quantifier, phrasal 4.15, 4.19
quantifier, proportional 4.18
quantifier, universal 4.64
quantifier floating 4.400
quantifier raising 2.368–369
quantitative adjective 1.18
quantity opposition 3.289
quantity sensitivity 3.616–617, 3.619, 4.344, 4.348
Quantity-Sensitivity Rule 3.618, 3.622
Quanzhou 1.378–379
quasi-auxiliary 1.183–184, 1.186
quasi-verb → pseudoverb
quatrain 4.90
Quay, Suzanne 1.368–369
Quḍāʿa 1.614, 2.427, 4.430
Qudāma ibn Jaʿfar 1.428, 2.304, 2.496, 2.579–581, 3.117
Quebec 3.398
Quechua 4.373–374
question 1.249, 1.303, 1.327, 1.342, 2.23, 2.144, 2.331
question, clefted 4.62
question, echo 1.266, 2.393, 4.566
question, indirect 2.393
question, phatic 2.205
question, rhetorical 1.266, 3.118, 3.365
question, yes/no 1.249, 2.115, 2.133–134, 2.388–389, 2.393–394, 2.397, 2.399
question answering system 1.458, 1.460
question mark 3.741
question particle 1.245, 1.280, 2.388–390
question word 1.524, 2.592, 2.596
question word, bimorphemic 2.388–389
questionnaire-based studies 2.651, 2.653
questions, acquisition of 2.100
Quhandizī, al- 2.189
Quilon 3.128
quinquiconsonantal → quinquiradical
quinquiliteral → quinquiradical
quinquiradical root 4.95
Quintilian 2.442
Quintin, Jean 3.141
Quiring-Zoche, Rosemarie 3.503
Quirk, Randolph 3.224, 4.531, 4.625
Quitout, Michel 3.274, 4.563

Qulammas al-Kinānī, al- 4.104
Qulunqul 4.276
Qumayr 4.482
Qumrān 3.409
Qŭnanbayŭlï, Ibrahim [Abay] 2.558
Qūnyawī, Muḥammad al- 3.341
quotation 2.537–538
quotation mark 2.332, 3.741
qurā 1.129
Qurʾān 1.151, 1.166–168, 1.170, 1.173, 1.179–181, 1.313, 1.336, 1.379, 1.381–382, 1.397–403, 1.427–428, 1.434–435, 1.448, 1.450, 1.456, 1.461, 1.472, 1.484, 1.487, 1.490, 1.512, 1.528–530, 1.532, 1.562, 1.564, 1.614–615, 1.617, 2.17, 2.25, 2.45–46, 2.70, 2.83, 2.85–86, 2.104, 2.108–109, 2.111, 2.176–177, 2.182–183, 2.187, 2.198–199, 2.263, 2.302–307, 2.331, 2.334, 2.373, 2.654–655, 3.14, 3.30–31, 3.71, 3.117, 3.139, 3.465, 3.629
Qurʾān, Amajur 2.600–601
Qurʾān, Baybars Jasnagir's 4.562
Qurʾān, Blue 2.598, 2.600
Qurʾān, Chester Beatty 4.131, 4.561
Qurʾān, codification of 4.4
Qurʾān, contrast in the 4.29
Qurʾān, copying of the 4.131
Qurʾān, creation of the 4.473
Qurʾān, dialects in the 3.14
Qurʾān, direct speech in the 4.30
Qurʾān, eternity of the 2.306
Qurʾān, exegesis of 2.232, 2.578, 2.580, 3.34, 3.36, 3.122, 3.551, 4.73, 4.235, 4.329
Qurʾān, figurative language in the 3.122
Qurʾān, generalization in the 4.29
Qurʾān, in Latin script 1.487
Qurʾān, inimitability of → *ʾiʿjāz al-Qurʾān*
Qurʾān, interpretation of the 1.427
Qurʾān, language of the → Qurʾānic Arabic
Qurʾān, loanwords in the 2.86, 2.666, 3.7, 3.36, 4.295
Qurʾān, *maġribī* 3.111
Qurʾān, manuscripts of the 2.597, 2.599, 2.600–601
Qurʾān, meaning of the 1.627
Qurʾān, memorization of the 2.110, 2.338, 3.57, 3.136
Qurʾān, metaphorical interpretation of the 3.117–118
Qurʾān, omission in the 4.27
Qurʾān oral transmission of the 4.105
Qurʾān, parenthesis in the 4.28
Qurʾān, Qarmaṭian 4.134
Qurʾān, readers of the 4.4, 4.8
Qurʾān, recitation of the 2.178, 2.293, 2.311, 2.327, 2.383, 3.194, 3.207, 3.503, 3.567, 3.726, 4.4, 4.181, 4.184, 4.346, 4.425–428, 4.759
Qurʾān, repetition in the 4.24
Qurʾān, rhyme in the 3.567, 4.30, 4.105
Qurʾān, study of the 3.690
Qurʾān, text of the 4.4
Qurʾān, translation of the 1.166, 1.170, 1.172, 1.529–530, 2.335, 3.501, 4.585
Qurʾān, variant readings of the 3.92–93 4.4–11
Qurʾān, word order in the 4.27
Qurʾān on the Internet 2.383
Qurʾān readers, Kufan 4.558
Qurʾānic Arabic 1.489, 1.491–492, 1.530, 1.532, 1.534, 1.594, 1.613, 1.615, 1.617, 2.264, 2.647–648, 2.705, 2.708, 3.10, 3.645, 3.648, 3.689, 4.21–31
Qurʾānic commentary → exegesis of the *Qurʾān*
Qurʾānic orthography 1.401
Qurʾānic paleography 2.597
Qurʾānic school 1.335, 2.137, 3.135, 3.113, 3.137–139, 4.184, 4.276
Qurayš 1.28, 1.128, 1.131, 1.361, 1.400, 1.594, 1.614–617, 1.614, 1.617, 2.88, 2.263–264, 3.14, 3.89, 3.91–93, 3.345, 4.201, 4.275, 4.377
qurba 1.567

Qurbāṭ → Kurbat **2**.217
qurrāʾ **4**.8
Qurra ibn Šarīk **1**.151, **1**.401, **3**.517
qurrāʾ, ʾahl al- **1**.129
Qušāšī, al- **2**.335
Qušjī, ʿAlāʾ ad-Dīn ibn Muḥammad al- **4**.686
Qusṭā ibn Lūqā **4**.546
Quṯām ibn ʿAbbās **2**.46
Quṭb, Sayyid **4**.77
Quṭba **4**.560
Quṭbī, al- **4**.78, **4**.277
Qutbuddin, Tahera **2**.325–331
Quṭrub, ʾAbū ʿAlī Muḥammad ibn al-Mustanīr **1**.626–628, **2**.403, **3**.34, **4**.329
Qutub Shahis **4**.452
quwwa **4**.287, **4**.414

R

/r/ 2.416
/r/, velarized 1.103
ra- 1.64, 572
Rabat 1.23–24, 2.470–471, 2.715
Rabat Arabic 1.596, 1.609, 2.621, 3.9, 3.273, 3.288–289, 3.292, 3.274
Rabbath Ammon 2.506
Rabeh 1.521, 3.61, 3.634
Rabīʿa 1.614, 2.555, 3.698, 4.304, 4.430
Rabin, Chaim 1.490, 1.562, 1.593, 1.613–618, 1.641, 2.85–86, 2.88, 2.257–259, 2.263, 2.355, 2.373, 2.387, 2.403, 2.539, 2.629, 3.14, 3.89, 3.93–94, 3.404, 3.410–411, 3.413, 3.416, 3.419, 3.454, 3.465, 3.497, 3.565, 3.645, 3.689–690, 3.692–694, 4.2, 4.36, 4.70, 4.123, 4.125–126, 4.295, 4.430, 4.457–459, 4.753
Rabinowitz, Isaac 3.466
rābiṭ(a) 1.49–50, 3.381–385, 3.547, 4.102
rabṭ 1.432, 4.22, 4.24
rabṭ, ʾadawāt ar- 1.471–472
Raby, Julian 3.341
racine 1.574, 1.600
racism 3.314
Radenberg, Hans-Peter 4.750
Radford, Andrew 2.131, 4.438, 4.522
radīʾ 1.192, 2.88, 4.12
radical 1.191, 1.313, 3.240
radical conjugation 1.209
radical consonant 1.313
radio 1.364, 1.528, 1.546, 1.654, 2.11, 2.503, 2.517–518, 2.658, 3.11, 3.194–195, 3.295
radio → broadcasting
radio, Kurdish 3.351
radio, New Zealand 4.628
radio Arabic 2.261, 2.278
radio Arabic → Media Arabic
Radio Omdurman 2.518
radix 1.574
Radkan 2.600, 2.256
Radtmann, Bartholomaeus 1.169
rafʿ 1.49, 1.67, 1.71, 1.309, 1.347, 1.556–560, 1.562–563, 2.92, 2.187, 2.232–234, 2.292, 2.402–404, 2.429–430, 2.537, 3.92, 3.651, 4.357
Rafea, Ahmed 1.460
Raffouli, Faraj 3.672, 4.62
Rafʾī, ʾAmīn ar- 3.193
rāfiʿa 2.225
Rāfiʿī, ar- 3.42
Raghib, Yusuf 3.516
Rahayta 2.272
Rahman, Parez Islam Syed Mustafizur 3.338
Rahman, Tariq 3.506–512
Rahman, Yusuf 2.306
Raimondi, Giovanni Battista 1.168
Rainey, Anson F. 2.22, 3.409
raising 2.311
raising analysis of relative clause 4.67
raising of noun phrase 4.365

raising, NP 4.747
raising, object 2.369, 3.103, 3.226, 4.539
raising, predicate 2.133
raising, Q-D 4.746
raising, quantifier 2.368–369
raising, right-node 4.648–649
Raising, Subject-to-Object 4.522
raising, verb 4.394, 4.399, 4.524, 4.650, 4.657
raising, vowel → vowel raising
rajāʾ 3.105
Rajaonarimanana, Narivelo 3.125–128
rajaz 2.281, 3.210, 3.212, 3.214, 3.219, 3.648, 4.32–37, 4.202, 4.287–288
rājiʿ 1.49, 4.102, 4.720
Rājiḥī, ʿAbdallāh ar- 3.455
Rājiḥī, ʿAbduh ar- 3.93
Rājiḥī, Šaraf ad-Dīn ʿAlī ar- 3.91
rājiz 4.32
ramal 3.210, 3.212, 3.214
Ramallah Arabic 1.584
Ramla 2.464
Ramlī, ʾAḥmad ibn ʿAlī ar- 2.541
Rammuny, Raji M. 2.395–397, 3.727, 4.140
Rampur 2.329
Ramsay, I.T. 4.72
Ramus, Franck 3.725
Ramzī, ʾIbrāhīm 1.600
Ranīrī, Nūr ad-Dīn ar- 1.8, 2.334–335
Rao, Aparna 2.217
Raphelengius, Franciscus 1.169–171
Rapoport, Tova 1.508
Rappaport, Malka 3.621, 4.489
Rās ilʿAyn Arabic 4.407
Rās Rummān 1.242
Rās Ṣadr Arabic 4.238
Rās Samra 3.408
Rasāʾil ʾIxwān aṣ-ṣafāʾ 4.220
Rašāyda 1.654, 1.656, 2.2, 3.326
Rašāyda Arabic 3.62
Rashidi, Linda 2.645
Rašīd, ir- 2.4
Rašīd Arabic 2.2–3
Rāšidiyya → Rosetta
rasm 3.515, 4.4
Rasmussen, Stig T. 3.23
Rassam, Amal 2.60
raṭan 2.259
raṭāna 2.468, 4.251–252
Ratcliffe, Robert R. 1.74–82, 1.576, 2.367, 3.304–305, 3.414, 3.420, 3.423, 3.439–447, 4.52, 4.94, 4.309–310, 4.339
rate of speech 3.725
Ratzaby, Yehudah 2.534
Ravid, Dorit 2.101
Rāwa 4.407
rāwī 1.4, 1.616, 2.542–543, 3.89, 3.645, 4.103
Rāwī, Nājiḥ ar- 2.635, 2.641
rawiyy 4.87–89
rawm 3.566, 4.427, 4.682
Raydān Palace 2.256
Raydanitic 2.256
rayḥān 1.152, 3.308–309, 3.340–341, 3.519, 4.131, 4.133–134
rayḥānī 3.309
Rayḥānī, ʿAlī ibn ʿUbayda ar- 3.309, 3.556
Raz, Shlomo 2.271, 4.305–307
Rāzī, ʾAbū Bakr Muḥammad ibn Zakariyyā [Rhazes], ar- 2.677
Rāzī, ʿAlī ibn Šaḏān ar- 2.601
Rāzī, Farīda 3.575
Rāzī, Faxr ad-Dīn ar- 2.444, 2.580, 2.677, 3.45, 3.119–120, 3.122, 3.322, 3.324, 4.8, 4.473
Rāziḥī Arabic 4.106
readers of the *Qurʾān* 4.4, 4.8
reading 1.153, 1.371, 2.108, 2.110–111, 2.692, 2.701, 3.74
reading, learning of 3.76, 3.78
reading, teaching of 2.106, 2.108
reading accuracy 3.79
reading difficulty 3.79

reading error 3.78, 4.271
reading fluency 3.79
realis 1.554, 3.269
reanalysis 1.80–81, 1.183–184, 1.256, 1.291, 1.304, 1.314, 1.491, 1.503, 1.590, 2.67–68, 2.140, 2.193–194, 2.287, 2.375, 3.578, 3.624, 4.37–43
reanalysis of definite article 2.554, 2.669, 4.723, 4.725
Rebhan, Helga 4.76
reborrowing 2.254
rebracketing 4.38
Rebstock, Ulrich 3.172
Recanati, François 4.328
Recasens, Daniel 3.524
Rechabites 1.127
recipient 2.145
reciprocal 1.284, 1.310, 1.329, 1.549, 2.245, 2.565, 2.568, 4.43–45, 4.57, 4.307, 4.397, 4.413, 4.492, 4.568
reciprocals, acquisition of 2.102
reciprocator 4.43–45
reciprocity 1.284, 2.524, 3.184, 3.202, 3.226, 4.43–45, 4.661
recitation 1.528, 1.533, 2.110
recitation of the *Qurʾān* 1.492, 2.183, 2.329, 2.494, 2.546, 2.649, 3.375, 3.503, 3.576, 3.726, 4.4, 4.181, 4.346, 4.425–428, 4.759
Reckendorf, Hermann 1.196, 1.472, 1.639, 2.224, 2.354–355, 2.389, 2.452, 2.539, 3.202, 3.453, 3.456, 3.556, 3.701, 3.705, 3.719, 3.721, 4.57, 4.196, 4.437, 4.484, 4.530
Reconquista 2.61, 2.64, 2.282, 4.572
reconstruction 2.369–370, 4.81–82
reconstruction of Proto-Semitic 4.479
reconstruction, syntactic 3.385
recursion 4.494, 4.497, 4.525
Red Sea 2.51–52, 2.268
Red Sea Coast 3.62
Redfern, Walter 3.538
Redhouse, Sir James W. 4.586
re-diphthongization 1.410
Redjala, Mbarek 3.141
Redlinger, Wendy E. 1.368
reduction 1.492–493, 1.539, 1.594, 1.611, 2.119–120, 2.195, 4.161
reduction, geminate 2.685, 4.241
reduction, morphological 2.520
reduction, phonological 2.194, 3.59
reduction, semantic 4.160
reduction, syllabic nuclei 4.683
reduction, syllable 4.688
reduction, valency 4.625–626
reduction, vowel 4.674–675, 4.682–683
reduction of consonant clusters 2.611
reduction of geminates 4.241
reduction of grammar 4.376
reduction of phonetic inventory 3.61, 3.63
reduction of short vowels 4.302, 4.307
redundancy 1.562, 2.19, 2.118, 2.226, 2.274, 3.696, 4.45–50, 4.193, 4.731
redundant element 4.235
redundant expression 3.201
reduplicant 4.50
reduplication 1.331, 1.341–342, 1.358, 1.395, 1.523, 1.639, 2.270, 2.349, 2.379, 2.474, 2.566–567, 2.619, 3.297, 3.626, 3.638, 4.50–53, 4.94–95, 4.286, 4.310–311, 4.644
reduplication, full 4.50–51
reduplication, morphological 4.50
reduplication, partial 1.313, 4.50
reduplication, phonological 4.50
reduplication, pronoun 4.52
reduplication, syndetic 4.52
reduplication, syntactic 4.50
reduplication, verbal 2.524
Reed, William L. 4.480, 4.482

Reese, Johannes 1.50–53
reference 1.12, 1.430, 1.474
reference time 1.199–200
reference, pronominal 1.431
referent 1.568
referent, maintenance of 1.649
referent shift 1.649
referential enunciation 4.332
referential expression 1.310–312, 1.486
referentiality 4.316
reflexive 1.284, 1.297, 1.310, 1.329, 1.377, 1.523, 2.245, 2.348–349, 2.525, 2.565, 2.568, 3.184, 3.553, 4.43, 4.53–57, 4.307, 4.522, 4.642
reflexive, logophoric 4.54–55
reflexive, simplex 4.53
reflexive complex 4.53–54
reflexive construction 1.622
reflexive verb 1.75
reflexive-passive 2.348, 3.251
reflexivity 4.642
reflexivization 1.284
reflexivizer 4.54
reform of Arabic 2.635
register 1.406, 1.409, 1.437, 1.483, 2.8, 2.10–11, 2.117, 4.57–60, 4.194
register, Bedouin 4.240
register, colloquial 1.616
register, conventionalization of 3.670
register, familiar 1.483
register, formal 1.483, 1.666
register, informal 1.483
register, sacred 2.662
register, third 1.632
register marker 4.57
Regnier, Jean Pierre 3.195
regression, dialect in 1.605
regression, false 2.276–277
regularity 2.177
Reh, Mechthild 4.38, 4.161
Reich, Peter 4.335
Reichenbach, Hans 1.196, 1.199–200
Reichmuth, Philipp 4.515–520
Reichmuth, Stefan 1.190, 2.388, 3.404, 3.437, 3.544, 4.72–80, 4.376, 4.759
Reig, Daniel 1.628, 3.26–27
Reim 4.305
Reimwortbildung 1.315
Reinaud, Joseph Touissant 2.42
reinforcement 1.342
Reinhardt, Carl 1.663, 3.478, 3.481, 3.489, 3.544
Reinhart, Tanya 2.170, 2.208, 4.53–55
Reinkowski, Maurus 4.592
Reiske, Johann Jakob 1.167
Rejwan, Nissim 4.76
Rejzek, Jiří 4.259
relation, specifier/head 1.352
relational adjective 1.80
Relational Grammar 1.457, 4.502, 4.535
relationship, sound and meaning 4.286–287
relative, cleft 4.66
relative, definite 4.68
relative, indefinite 4.68, 4.722
relative, nonrestrictive 4.64, 4.747
relative, restrictive 4.747
relative adjective 1.14, 1.17, 1.452, 3.153, 3.220, 4.691
relative clause 1.13, 1.20, 1.49–50, 1.124, 1.266, 1.311–312, 1.395, 1.417, 2.595, 3.202, 3.431
relative clause, matching analysis of 4.67
relative clause in Northwest Semitic 3.419
relative identity value 1.586
relative marker 1.488, 1.524, 4.60–61, 4.63, 4.65, 4.67
relative marker, Tajik 4.622
relative pronoun 1.13, 1.91, 1.106, 1.116, 1.226, 1.239, 1.263, 1.273, 1.279, 1.303, 1.392, 1.399, 1.491, 1.549, 1.559–560, 1.566, 1.606, 1.639, 1.668, 2.161, 2.195–186, 2.243, 2.271, 2.484, 2.530, 2.537, 2.614, 3.151, 3.192, 3.245, 3.261, 3.277, 3.365, 3.403, 3.484, 3.531, 3.636, 3.694–695, 4.62–63, 4.70–72, 4.111, 4.187–188, 4.221, 4.236, 4.246, 4.392, 4.552, 4.622, 4.689

relativization 2.118, 4.522, 4.525–526
relativizer 4.61, 4.64–66
release 1.26
Relevance Principle 4.503
relic area 1.537
relic area → enclave
religion 1.650
religion and language 4.72–80
religious community 1.442–448
religious expressions 1.650
remodeling 2.71
Renan, Ernst 2.86
Rendille 1.35
Rendsburg, Gary 1.493
Renfrew, Colin 3.141
renominalization 2.208
Rentz, G. 3.402
reordering 3.204
repetition 1.246, 1.342, 1.431, 1.471, 1.474, 1.640, 1.647–648, 1.651, 2.205, 3.500, 3.669, 4.49–50, 4.52, 4.103
repetition, lexical 1.431–432, 2.208, 3.713
repetition, root 4.52
repetition in the *Qurʾān* 4.24
repetition of consonants 4.94
reported speech 3.665
representation, autosegmental 3.626
representative 3.570
request 3.659, 3.680, 4.331–332
resegmentation 4.38
resemblance 2.92
resolution, anaphoric 1.514
resonant 3.596–597
resonating cavity 4.666
response, affective 2.650
response, behavioral 2.650
response, cognitive 2.650
response, evaluative 2.650
restitution, false 2.276–277
restriction 4.27
restriction, co-occurrence 1.436
restriction, definite 4.19
restriction, minimality 4.392
restriction, selectional 1.436
restriction of quantifier 4.15, 4.17–19
restructuring 1.593–595, 2.120, 3.61–62
restructuring, linguistic 3.58
result nominal 3.393
resultative 1.199, 1.201, 1.331, 1.396, 1.553
resumption 1.312, 1.413, 3.383, 4.67, 4.80–85, 4.400
resumption, direct object 4.67
resumptive pronoun 1.49, 1.312, 1.350, 1.493, 1.552, 2.132, 2.148, 2.171–272, 2.296, 2.595, 3.220, 3.383, 3.385, 3.560, 3.589, 4.60, 4.65, 4.68, 4.71, 4.82, 4.186, 4.366, 4.485, 4.525–527, 4.720–722, 4.738
resyllabication 1.28, 1.246, 2.152, 2.230, 2.575, 2.611, 2.685, 3.405, 3.480, 3.488, 3.611, 4.85–86, 4.116, 4.124, 4.126, 4.243, 4.351, 4.390–391
resyllabication, guttural 3.403
resyllabication, Najdī 4.244
resyllabification → resyllabication
retention 1.80
Retracted Tongue Root 4.670
retraction of the tongue root 2.365
retrieval 2.681, 4.270
retroflex 1.134–135, 1.521, 2.269, 4.273, 4.434
retroflexion 4.754
Retsö, Jan 1.126–133, 1.178–182, 1.490, 1.614–617, 1.622–626, 2.24, 3.192, 3.424, 3.544, 3.553, 3.556–557, 3.564–565, 3.588, 4.62, 4.64, 4.70, 4.304, 4.402, 4.620, 4.727
Reuland, Eric 4.53–55
Reuschel, Wolfgang 1.484, 3.543, 4.151
revelation 1.173, 4.21–31
Revell, E.J. 2.232
reversal of kinship terms 2.587
reverse-role use 4.469, 4.470
Reves, Thea 2.464
Revised Extended Standard Theory 1.457
revival of Amazigh 2.713
revival of Arabic 1.531

R-expression 1.310–312
Rey Bouba 1.337
Reyhanlı 1.111, 1.113–114, 1.116–118
Reynolds, Dwight Fletcher 3.500
Reza Shah 2.572
Rezaei, Siamak 3.739–742
rhapsode 3.645–646
Rhazes → Rāzī, ʾAbū Bakr Muḥammad ibn Zakariyyā ar-
rhēma 3.547
rhematization 4.485
rheme 1.429, 1.485, 2.569, 4.484–487, 4.494, 4.498, 4.501–502
rheme marker 1.476
rhetoric 2.88, 2.358–359, 2.445, 2.627, 4.472
rhetoric, Arabic 2.443, 3.116–117
rhetoric, contrastive 1.648
rhetoric, Indian 2.579
rhetorical question 1.266, 3.118, 3.365
Rhodokanakis, Nikolaus 3.478, 4.302–304, 4.306
rhyme 3.210, 3.566–568, 3.601, 3.611–612, 3.647, 4.86–92, 4.103, 4.208, 4.387
rhyme, free 4.88
rhyme, restricted 4.88
rhyme in *rajaz* 4.32, 4.34
rhyme in the *Qurʾān* 3.567, 4.30, 4.105
rhymed prose 2.264, 2.305, 3.32, 4.103
rhyming letter 1.150
rhyming order 3.21
rhythm 3.620, 4.734
rhythmic principle 4.734
rhythmic stress shift 3.617
Riad, M. 1.503
Riaz, Ali 1.257–259
Ribera, Julián 1.99
Rice, David Talbot 2.680, 3.309, 3.340, 4.131, 4.560–561
Rice, Frank A. 2.481, 4.59
Rice, Mabel 2.680, 2689
Richard, Francis 3.336–338, 3.341
Richert, Nicole 4.460
Richter, Tonio Sebastian 1.495–501
Riḍā, ʾAḥmad 3.40
Riḍā, ʿAlī 1.382
Riḍā, Muḥammad 3.346
ridda 1.129
Riddell, Peter 2.335
riddle 1.57, 3.174, 4.158
ridf 4.88
Riebeeck, Jan van 4.290
Riemsdijk, Henk van 4.719
Rieschild, Verna Robertson 4.286, 4.469
Rif 2.62
Rifaat, Khaled 2.677
Riffaterre, Michael 4.199
right-dislocation 2.148, 4.541, 4.726
right-dominant 3.613
Rights and Obligations 1.415
Riguet, Maurice 2.660
Rihani, Amin 2.703, 3.395
Rīḥānī, Najīb ar- 1.600
Rijāl Alma Arabic 4.127
Rijkhoff, Jan 2.143
rika 4.98
rikʿa kırması 4.98
Rimbaud, Arthur 4.212
Ringel, Heinrich 3.718
Río del Oro 2.240, 3.169
Rippin, Andrew 2.373, 3.7, 3.36, 3.690, 4.295
riqāʿ / rigʿa 1.152, 3.309, 4.98, 4.131, 4.134, 4.561
Rīrāš, Sayx ʾAḥmad ʿAbdallāh 4.277
rising contour 2.397, 2.400
rising melody 2.397
Ristāq Arabic 2.388, 3.481
Ritter, Elizabeth 1.479, 3.232, 3.429, 4.15, 4.743, 4.746
ritual 2.202, 2.206, 2.646, 4.73
Rivolta, Barbara 3.337, 4.561
riwāya 2.543
riwāya bi-l-lafḏ̣ 2.331
riwāya bi-l-maʿnā 2.331
Riyadh 4.123
Riyadh Arabic 4.62

Rizēgāt 1.362
Rizzi, Luigi 2.171, 2.173–174, 3.706, 4.362, 4.366, 4.488, 4.647, 4.746
Rizzitano, Umberto 2.629
Rmēlāt Arabic 3.401, 4.238, 4.677
Roach, Peter 4.387
Robert of Ketton 1.166
Roberts, Allen F. 4.181
Roberts, Mary Nooter 4.181
Robertson Smith, William 3.719
Robin, Christian Julien 1.617, 2.256–261, 2.263, 3.467, 3.469–472, 3.516–517, 4.295
Robins, R.H. 1.343
Robinson, Charles Henry 2.251
Robinson, David 4.180
Robinson, Francis 3.510
Robinson, Jane J. 1.457
Rocco, Benedetto 4.216
Rockey, Denyse 2.677
Rodgers, J. 2.262
Roger II, King 3.142
Röhrborn, Klaus 4.585
Röll, Walter 2.456
Röllig, Wolfgang 3.409
Řom 2.217, 4.255
Romaine, Suzanne 1.416, 1.419, 2.321, 2.684, 2.704
Roman, André 1.544–545, 1.641, 1.643–647, 2.83, 2.228–229, 3.564, 4.463, 4.465, 4.666–667
Romance languages 1.96–99, 1.100, 1.102, 1.562, 2.173, 2.256, 3.232–233, 3.524
řomanes 2.217
Romani 2.73, 2.217, 2.219–221, 2.469–470
řomani čhib 2.217
Romanian 4.371
Romans 1.293
Romansch 2.684
Romany 4.255
Rome 1.171
Römer, Claudia 4.584
Ronjat, Jules 1.368
Ronkel, Philippus Samuel van 2.334
root 1.13, 1.15, 1.76, 1.191, 1.207, 1.313, 1.376–377, 1.425, 1.573–577, 2.103, 2.346–347, 2.447–450, 2.692, 3.21, 3.31, 3.33, 3.40, 3.45–46, 3.66, 3.240, 3.301, 3.423, 3.425, 3.610, 3.627, 4.93–98, 4.336–337, 4.338–344, 4.461
root, assimilated 2.165
root, biradical 1.638, 1.640, 2.155, 2.347, 2.349, 3.49, 3.462, 3.610, 4.93, 4.95, 4.97
root, consonantal 1.574–575
root, defective 2.167–168
root, discontinuous 1.376, 3.225
root, doubly weak 2.164, 2.448
root, geminated 1.313–314, 4.93
root, hollow 2.166–167
root, homorganic 3.610
root, lexical 3.626
root, meaning of 1.576
root, mental representation of 4.95–96
root, quadriradical 2.448, 2.485, 3.627, 4.93, 4.95
root, quinquiradical 4.95
root, reduplicated 4.94
root, sound 2.165
root, triradical 1.644–645, 2.347, 2.349, 2.352, 2.375, 2.448, 3.46, 3.49, 4.93–94
root, verbal 3.624
root, weak 1.314, 2.164–165, 2.168, 2.448
root augment 1.315
root consonant 2.346
root determinative 1.315, 4.96–97
root evidence 1.378
root extension 4.97
root innovation 4.462
root morpheme 2.346, 3.297
root radical 2.449
root segmentation 4.95
root system, Semitic 4.173–174
root tier 3.225

root variation 1.82
root vowel 1.313, 2.167
root-and-pattern morphology 3.445
Root-and-Pattern Principle 1.76, 1.376, 3.425, 3.440, 4.461, 4.640
root-echo response 2.203
Rose, Sharon 3.305, 3.494, 3.628
Rosen, Lawrence 2.62–64
Rosenbaum, Gabriel M. 2.378, 2.534, 4.593
Rosenhouse, Judith 1.259–269, 1.271, 1.571, 1.616, 2.230, 2.481–493, 3.141–142, 3.146, 3.360, 3.403, 3.427, 3.584, 3.586, 3.717, 3.721, 3.727, 4.2, 4.192, 4.195, 4.302–303, 4.305–306, 4.649, 4.653, 4.659–665
Rosenthal, Franz 3.7, 3.112, 3.307, 3.309, 3.410–411, 3.689, 4.132–133, 4.545
Rosenthall, Samuel 2.153–155, 3.461–464
Rosetta Arabic 2.2–3
Ross, Jeffrey A. 2.59, 2.64
Ross, John Robert 3.384, 3.386, 3.687, 4.81, 4.522–527, 4.721
Rossi, Ettore 2.456, 4.106, 4.108, 4.548, 4.750
Rossi, Mario 3.724
Rössler, Otto 4.170
rote memorization 3.499
Roth [-Laly], Arlette 1.190, 1.365, 1.536, 2.363, 2.388, 3.61, 4.198, 4.708, 4.710, 4.713, 4.716
Rothstein, Susan 3.685
Rothweiler, Monika 2.689
Rottland, Franz 3.435
Rotwelsch 2.217, 4.255
Rouaud, Alain 1.654
Rouchdy, Aleya 1.416–417, 2.1, 2.642, 3.59, 3.313, 3.435, 3.664
rounding harmony 4.676
Rousseau, Jean 1.574
routine 2.202
routinization 2.196
Roux, Arsène 2.470–471
Rowson, Everett 2.468, 2.470, 4.255–256, 4.258
Roxburgh, David J. 3.341
Royal Institute for Amazigh Culture 2.63
Royal, Anne Marie 2.8, 4.638
Rozwadowska, Bozena 3.388
Rubᶜ al-Xāli → Empty Quarter
Ruʾba ibn al-ᶜAjjāj 4.32–36, 4200
rubāᶜī 4.624
rubāᶜī meter 4.206–207
Rubatab Arabic 4.376
rubba 1.72, 1.558, 3.700, 4.161
Rubenson, Samuel 2.1
Rubin, Aaron 3.415
Rubino, Carl 4.50
Rubinstein, Amnon 2.438
Ruch, R. 4.271
Ruḍā 1.127, 4.480–481
Rufayda Arabic 4.127
Ruiter, Jan Jaap de 2.72–79
rujḥān 1.71, 3.322
rukn 2.496–497
Ruland, Hans-Jochen 4.543, 4.545
rule 1.208–209
rule, chopping 4.526
rule, copying 4.525–526
rule, cyclic 2.38, 3.619, 4.522
rule, exocentric 4.649
rule, movement 4.522
rule, obligatory 4.522
rule, optional 4.522
rule, phrase structure 4.524, 4.526, 4.741–742
rule, prescriptive 2.178
rule, S 4.741
rule, S-expansion 4.742–743
rule, subcategorization 1.418, 1.635
rule, transformational 4.522, 4.524–525
rule change 4.38
rule extension 1.74
rule ordering 2.38
rule ordering, extrinsic 3.496
rule system 1.422, 1.457
rūmī 1.99, 3.341
Rūmī, Jalāl ad-Dīn 4.561

Rūmiyya 3.7
Rummānī, ʿAlī ibn ʿĪsā ar- 2.188, 2.303–304, 2.443–444, 3.323, 4.119
Rumpf, Christian 4.579
Rundgren, Frithiof 3.31, 3.450
Rundī, ar- 2.316
runes, Turkic 4.450
run-on lines 4.92
ruqʿa 3.519, 4.98–100, 4.132
ruqʿa, broken 4.98
ruqʿa qirmasī 4.98
rural area 1.595
rural dialect 1.595, 1.605, 1.607–608, 2.8
Rušdī, Rašād 1.601
Russell, Robert 4.524
Russian 1.45, 1.51–52, 1.196, 2.66, 2.345, 2.440, 2.699, 3.524, 3.621–622, 3.725, 4.372, 4.395, 4.612
Russian, Arabic loanwords in 4.259–267
Rustāq → Ristāq
rutba 4.22
rutta 2.676
Ruwwāfa 4.478
Rūx, Drīs ar- 1.603
ruxṣa 3.647, 3.649, 3.651
Růžička, Karel F. 1.660
Rwālah Arabic 1.260, 4.125–126, 4.403
Rwanda 1.661–662
Rybalkin, Valeriy 1.67–74, 4.428–430
Ryckmans, Jacques 3.471, 4.296
Ryding, Karin C. 1.666–671, 2.294–298, 2.424, 3.100–101, 4.192, 4.360–361

S

S rule 4.741
sa- 2.22, 2.192, 2.547
saʿa 3.104, 4.411
Saad, Elias N. 3.135–136
Saad, George Nehmed 1.357, 3.555, 4.487
Saʿāda, ʾAnṭūn 3.348
Saadawi, Nawal 2.642, 2.645, 2.647
Saadeddin, Mohamed Akram 1.432, 1.648
Saʿadī, as- 3.373
Saadi-Mokrane, Djamila 1.534
Saʾādne 1.362
Saadya Gaon 1.170, 2.23, 2.66, 2.528, 2.534, 3.218
saafara 4.181
ṣaʿālīk → *ṣuʿluk* 4.208
saʿat al-kalām 3.116
Ṣaʿayda 2.2
Sabaʾ 2.56, 2.256, 3.467, 4.295
sabab 1.49, 2.306, 2.309, 2.357, 3.208–210, 3.213–214, 4.101, 4.236, 4.545
sababī 4.101 -102, 4.221
Sabaean 1.127, 2.56, 2.66–67, 2.256–258, 3.227, 3.465, 3.467, 3.471–472, 3.592, 4.175, 4.295, 4.301, 4.303, 4.309–310, 4.312, 4.478, 4.753
Sabaeo-Ḥimyaritic 2.256
Sabaeo-Raydanitic 2.256
Sabaic → Sabaean
Sabaki 4.381
Sabʿān, L.K. as- 2.609
Šaʿbān, Sultan 3.308
sabb 2.361–363
šabbaha 4.12–13
Ṣabbān, aṣ- 2.189
Ṣabbēha Arabic 4.751
Sabines, Jaime 3.396
Sabuni, Abdulghafur 4.664
Šabwa 4.750
Sachdev, Itesh 2.660–661, 4.323–324
Sacks, Harvey 2.202
Sacleux, Charles 1.663
sacralization 4.73
sacred language 1.174
Saʿd Arabic 3.94
Saʿd ibn Bakr 3.694
Saʿd ibn Bakr Arabic 4.376
Saʿd Ibn Zayd Manāt 4.376
Ṣaʿdah Arabic 4.751, 4.754, 4.756
ṣadara 4.740
Sadat, President Anwar as- 3.194–196, 3.200, 3.540, 4.77
šadd 1.152, 1.192, 2.177, 3.90, 3.92, 3.94, 4.9, 4.22
šadda 1.2, 1.134, 3.77, 4.518
Saddam Hussein, President 1.657, 3.370, 3.664, 3.666, 3.670
Säddūm wəll Ndyartu 3.173
Šādigān 2.572
Ṣādiqī, ʿA.A. 2.407, 2.409
Sadiqi, Fatima 1.456, 2.642–650, 2.659, 4.190
šaḍiyya 3.341
Sadock, Jerold M. 1.457, 4457
ṣadr 4.87, 4.89
Sadr City 1.272
Sadr City Arabic 1.272

Ṣadrī Afšār, Gh.Ḥ. 2.410–411
Saeed, Aziz 1.650
Ṣafadī, aṣ- 2.497, 3.380, 4.562
Ṣafadī, Yāsīn Ḥamīd 3.337, 340, 4.134
Safady, Jorge Salim 3.4
Ṣafaitic 2.198, 2.263, 3.227, 3.411, 3.465, 3.468, 3.471, 4.478–482
Ṣafaitic, South 4.480–482
šafatāni 4.426
Safavid Dynasty 2.572, 3.519
Safed Arabic 1.608
Saffar, Abdul-Emir al- 1.457
Safī d-Dīn al-Ḥillī, → Ḥillī, Ṣafī d-Dīn al-
Šafī, ʿMuḥammad 3.341
Safid Buland 2.40
Šāfiʿī law 2.337
Šāfiʿī, Jalāl ad-Dīn Muḥammad ibn ʾAḥmad aš- 1.380, 1.448, 1.451, 2.60, 3.36, 3.122, 3.346, 4.13, 4.295
Safir, Kenneth 3.706
Safi-Stagni, Sabah 2.687–694, 3.305, 4.267–272, 4.336
Safiullina, Flyora 4.450
ṣafra 4.157–158
Ṣafwān, Muṣṭafā 1.600
Sag, Ivan A. 1.458, 1.2.18, 4.316, 4.648
Ṣaġānī, ʾAbū l-Faḍāʾil Ḥasan ibn Muḥammad aṣ- 1.627, 2.328, 3.34–35
Sagey, Elizabeth 1.27, 3.525
Ṣaġīr, Muḥammad aṣ- 3.222
Saguer, A. 2.347
Sagya el-Ḥamra 2.240, 3.169
Saharan languages 3.438
Sahel region 1.360, 1.362–363
Sahhaf, as-, Minister 1.657
ṣāḥib al-ḥāl 2.132, 2.224
šāhid 1.627, 2.177, 2.188, 3.23, 3.38, 3.43, 3.323, 3.648, 3.690, 4.32, 4.35, 4.436
ṣaḥīḥ 2.88, 2.93, 2.309, 2.448, 3.91, 4.121, 4.644
Sāḥil Maryūṭ Arabic 2.314
Šāhīn, Ṣalāḥ 1.601
Saho 1.35, 2.268–272, 3.62
Sahraoui, Cheb 2.78
Šahrastānī, Tāj ad-Dīn Muḥammad ibn ʿAbd al-Karīm aš- 2.303
Saʿīd, ʿAlī ʾAḥmad → Adonis
Said, Edward 1.533
Said, Majed F. 2.481
Saʿīd, Naffūsa Zakariyyā 3.348
Saʿīd, Nūrī as- 3.348
Saʿīd, Ṭalāl ʿUṯmān as- 1.273
Saida Arabic 1.608
Saidat, Ahmad Mahmoud 2.501
Ṣaʿīdī Arabic → Arabic, Upper Egyptian
Saidi, Redouan 2.75
Saiegh-Haddad, Elinor 3.79
Saifawa Dynasty 2.552
Ṣāʾiġ, Fāʾiz aṣ- 3.18
Saint-Exupéry, Antoine de 1.603
Saint-Louis 3.174, 4.723
šāʿir 3.645
ṣāʾit 2.238
sajʿ 1.8, 2.264, 2.305, 3.32, 3.648, 4.73, 4.103–106, 4.203
sajʿa 4.103–104
šajāʿat al-ʿarabiyya 3.651, 4.24
Sakai, S. 4.669
Sakarna, Ahmad Khalaf 2.501
Sakhr 1.214, 1.460, 1.513
Sakhr products 1.210
sākin 2.238, 3.208
Šākir, ʿAbdullāh 4.107
Sakkākī, ʾAbū Yaʿqūb Yūsuf ibn ʾAbī Bakr as- 2.79, 2.81, 2.359, 2.441, 2.444, 2.447, 2.580–581, 3.119–120, 3.214, 4.287, 4.330–331, 4.473, 4.625
Sakkal, Mamoun 2.601–602
Sakkout, Hamdi 3.518
Sakthivel, S. 4.433
Sala 3.287
Šalabī, ʿAbd al-Fattāḥ ʾIsmāʿīl 4.8
Saladin [Ṣalāḥ ad-Din], Emir 2.601
Salafis 1.381
Salafiyya 1.380, 3.137–138
Ṣalāla 2.212, 3.478
Ṣalāla Arabic 3.479
Salama ibn Jandal 3.309, 3.340, 4.560
Salamaeans 4.478
Salamanca 1.168
Salamas 3.370
Salamat 1.360–363

Salameh, Eva-Kristina 2.97, 2.678–679

ṣalāt 2.45

salb ad-dalāla 4.163

Salé 3.294

Salé Arabic 1.609, 3.273–274

Salegh-Haddad, Elinor 2.111

Salem, Babikian 4.77

salfih 3.500

Saliba, George 2.258

Salibi, Kamal 4.514

saliency 2.117–118

saliency, perceptual 3.206

Ṣāliḥ, aṭ-Ṭayyib 2.559, 2.598, 3.308–309

Ṣāliḥ, Maḥmūd H. 1.457, 3.678, 3.681

Ṣāliḥiyya 4.277

sālim 4.645

Sālim, ʿAlī 1.601

Sālim, Ḥamdī as-Sayyid 4.276

Salim, Major 1.519

Salishan 3.204

Sallam, A.M. 1.668, 3.430

Sallo, Ibrahim-Khidhir 1.417

Salloum, D. → Sallūm, Dāwūd

Sallūm, Dāwūd 3.91, 4.304, 4.759

Ṣalm 4.480–481

Salmanasar III, King 1.126

Salmāniye 1.362

Salmiyya 2.609

Salmons, Joe 4.373

Salonen, Erkki 2.418

Šalqānī, ʿAbd al-Ḥamīd aš- 3.89

Ṣalṭ Arabic 1.595, 2.498–499

Ṣalṭ, aṣ- 2.499

Salzburg School 1.586

Šām, aš- 2.481

samāʿ 1.72, 1.148, 1.194, 2.87

samāʿī 1.67

Samānah 4.238

Samānah Arabic 3.401

Samandağ 1.112, 1.117–118

Samandağ Arabic 1.117

Samʿānī, ʾAbū Saʿd as- 3.380, 717

Samanid Dynasty 1.174, 1.403, 2.406, 3.574

Samarin, William J. 4.72

Samaritan Arabic 3.218

Samarqand 2.46

Samarqandī, ʾAbū l-Qāsim as- 4.686

Samarra-Fallujah Line 2.414

Sāmarrāʾī, Fāḍil Ṣāliḥ as- 2.80

Samarrāʾī, ʾIbrāhīm as- 2.263, 3.718

Samārum 2.212

Šamatrānī, Šams ad-Din as- 2.335

Samāwa Arabic 2.416

Samawʾal ibn ʿAdī, as- 2.526, 3.732

Samb, Amar 4.182

saṃdhi 4.115

Samḥa 4.301

Samhan, Helen 3.313

ṣāmit 2.238

Šammar 3.326, 4.125, 4.403

Sammarco, A. 2.457

Šammari Arabic 1.260, 1.606, 3.326, 4.126

Sampson, John 2.219

Samuels, M.L. 4.288

Samy, Doaa 1.513

Samy, Waheed 2.380–387

šaʾn 1.558, 2.552

San Eulogio 1.96

Ṣanʿāʾ 3.693, 4.757

ṣanʿā 2.44

ṣanʿā, ʾahl aṣ- 2.305

Sanad, Azza 4.766

Sanandaj 2.604

Ṣanʿānī Arabic 1.23, 1.25, 1.445, 1.596, 2.39, 2.231, 3.84, 3.450–452, 3.611, 3.613, 3.620, 3.674–675, 3.702, 4.62, 4.106–115, 4.117, 4.529–535, 4.559, 4.690, 4.695, 4.697, 4.750–751, 4.753, 4.756–757

Sančes [Sánchez], Alcaén A. 2.238, 3.208–209

Sánchez-Albornoz, Claudio 1.98

sandhi 1.187, 1.237, 1.262, 1.322, 2.299, 2.397, 4.115–118, 4.682

sandhi, blocking of 2.397

sandhi, external 4.115

sandhi, internal 4.115–116

sandhi writings 4.479

Sandler, Wendy 4.223
Sandoval, Antonio Moreno 1.513
Sanford, Anthony J. 4.502
Sango 1.363, 4.386
sanjak 1.8
Sankoff, David 1.418
Sanni, Amidu 3.646
Ṣannūʿ, Yaʿqūb 1.176, 1.598, 1.600, 1.602, 3.193
šanšana 2.556, 4.304
Sanskrit 1.8, 1.286, 2.286, 2.326, 2.337–338, 3.129, 3.133, 3.422, 3.553, 3.621–622, 4.262, 4.434, 4.477
Sanskrit, Vedic 4.477
Sanskritization 1.287–288, 2.674, 4.596
Santiago de Chile 3.2
Santon Sidi Haddi 4.157
São Paulo 3.2, 3.4
Sapir, Edward 1.484, 1.616, 2.67
saqṭa 2.629
Sara 1.360, 1.363
Sara, Solomon I. 2.228
Sara Mountains 3.401, 3.404
Saracens 1.128
šaraf 3.658
Šaraf ad-Dīn 4.750
Sarapis 1.131
Sarāt 3.693
Sarāt ʿAbīdah Arabic 4.127
Sarāt Arabic 3.606
Sarauw, Christian 1.614, 2.312, 3.93
Sardinia 4.215
ṣarf 1.119, 1.381, 1.643, 2.402, 2.404, 2.426, 2.448, 2.623, 3.96, 3.107, 3.136, 4.118–122, 4.207, 4.432, 4.442–443
ṣarf, mamnūʿ min aṣ- 1.309
ṣarfa 2.303–304
Sargon II, King 4.478
šarḥ 1.3–4, 1.232, 2.527, 2.531
šarʿī 3.119, 3.210, 3.212, 3.214
šarīʿa 1.6, 1.403, 1.665, 2.338, 2.646, 2.648, 3.135, 3.352
Sarig, Lea 1.474
Sarmaan 4.276
Šarqiyya 1.589, 2.4
Šarqiyya Arabic 1.584, 1.589, 2.2, 3.401–402, 3.404, 3.406, 4.238, 4.727
Sarrāj, Muḥammad as- 3.392
šarṭ 1.49, 1.72, 2.477–479
šarṭa 3.741
Šarṭūnī, Rašīd aš- 2.13–15, 2.316, 4.659–660
Šarṭūnī, Saʿīd al-Xūrī aš- 3.39–40, 3.72
Sarwāni, ʿAbbās Khan 3.508
Sason 1.87
Sassanid Empire 1.378, 2.406, 4.424
Sasse, Hans-Jürgen 1.563, 3.237
Šaṭaṭ, Yūnis 2.705
satellite 2.29, 2.143–144, 3.194, 3.198
satellite channel 4.154
satellite television 3.317
satellite-framed 1.485
Šāṭibī, ʾAbū ʾIsḥāq ʾIbrāhīm ibn Mūsā aš- 1.3, 4.28, 4.21
šatm 2.361–363, 3.173
šaṭr 4.87
Šatt al-ʿArab 2.571
Satterthwait, Arnold 3.179
Satti, Nasir 3.435
Satzinger, Helmut 1.496
Saudi Arabia 1.258, 1.450–451, 2.105, 2.107, 2.250, 2.255, 2.268, 2.696, 3.179, 3.313, 4.123–130
Saudi Arabia, Arabic in 2.205
Saudi Arabian Arabic 1.657–658, 2.37–38, 2.97, 2.99, 2.658, 3.237, 4.655, 4.671
Saudi Arabian Bedouin Arabic 1.266
Saunders, Elaine 2.675
Saussure, Ferdinand de 1.573, 1.575, 3.188
Sauvaget, Jean 2.42
Sauvignac, Raphael 4.479, 4.481–482
Savary, Claude 4.516
ṣawāb 2.629

sawād al-ʿIrāq 1.617
šawāhid → *šāhid*
Sawaie, Mohammed 1.596, 1.668, 2.498–505, 2.634–642, 2.656, 2.658, 4.460
Ṣawālḥa Arabic 3.401, 4.240
Sawārkah Arabic 3.401, 4.238
sawfa 1.217, 1.509, 1.568, 2.192, 3.97, 3.548
Šāwi 4.403
Šāwi Arabic 1.546, 4.403, 4.407–408
šāwiyya 1.333
Šawkānī, Muḥammad ibn ʿAlī aš- 3.380
Šawqi, ʾAḥmad 4.87, 4.90, 4.106, 4.213
ṣawt 2.237, 2.426, 2.433, 2.541, 3.123, 3.603, 4.283–284
ṣawt murakkab 1.640
ṣawtī, ḥarf 2.235
Sawyer, John F.A. 4.72
Saxāwī, as- 3.380
Saxon genitive 3.69, 3.232
Ṣaxr, Bani 1.260
šaxṣ 2.427, 3.547
šaxṣi 3.178
šayʾ 1.191, 2.388, 4.443
Sayah, Mansour 2.696
Šaybānī, ʾAbū ʿAmr aš- 2.85, 2.359, 3.21, 3.31–32, 3.689
Ṣaydāwī, aṣ- 4.132
Sayf 1.361
Sayf ad-Dawla, ʿAyda 1.600
Sayf ad-Dīn, ʿAbd ʿAlī 2.328
Sayf ad-Dīn, Ṭāhir 2.327–328
Sayf ibn Ḏī Yazan 1.361
Ṣayhadic 1.491, 2.262, 4.301
Ṣāyiġ, Mājid al- 3.14
Ṣāyiġ, Samīr al- 4.214
Ṣāyiġ, Tawfīq 4.214
Sayrambaev, T. 2.557
šayʾun 2.194
Sayyāb, Badr Šākir as- 4.90–91
Šayyāl, Jamāl ad-Dīn aš- 4.26, 4.460
Sayyid, ʾAḥmad Luṭfī as- 1.602, 2.634, 3.193, 3.349
Sbai, Noufissa 3.294
Šbanāt 3.170
SC → structural change
scalability 2.321–322
Scaliger, Joseph Justus 1.167, 1.169–170
Scarbrough, Earl of 4.149
scenario 4.502
Sceptics 4.11
Schaade, Arthur 2.233, 3.123
Schabert, Peter 3.65, 3.142
Schacht, Joseph 3.7
Schack, Adolf Friedrich von 1.98
Schaeffer, Jean-Marie 1.575, 3.233
Schall, Anton 2.199, 2.201, 3.426
Scheer, Tobias 2.165, 2.167–168, 2.448
Schegloff, Emanuel 2.202
scheme 1.573
Schenkel, W. 1.503–504
Scherer, G.A.C. 2.683
Scherner, Bernd 4.450
Schiffrin, Deborah 1.471, 1.474, 2.384
Schimmel, Annemarie 2.325, 2.328, 3.337–338, 3.379, 3.718, 3.721, 4.98, 4.131, 4.133–134, 4.138, 4.560, 4.562
Schindler, Valentin 1.166
Schippers, Arie 2.66, 4.59
Schirmunski, Viktor 1.585
Schlauch 1.586
Schleifer, Abdallah S. 3.193, 3.197–198
Schmidt, Annette 2.684
Schmidt, Hans 3.431, 3.654–655, 3.684, 3.686, 4.317, 4.656, 4.662
Schmidt, Karl Horst 3.188
Schmidt, Richard 1.633, 1.668, 4.632
Schneider, Irene 3.6
Schöck, Cornelia 2.479
Schoeler, Gregor 4.4
scholasticism 2.188
Scholz, Arno 2.455
school 1.97, 1.175, 1.177, 1.258, 2.338
school, Apollo 4.91
school, Basran → Basran school
school, bilingual 1.440
school, Christian Mission 1.662
school, Diwan 4.90–91
school, foreign 1.532

school, Franco-Arabic 3.137
school, French 1.448, 1.450, 3.139
school, French/Berber 3.294
school, *imam-hatip* 4.578, 4.582
school, Jewish 2.64
school, Kufan → Kufan school
school, language 1.382
school, missionary 2.8, 2.699
school, mosque 2.109
school, Muslim 1.382
School, Old Iraqi → Old Iraqi school
school, preparatory 2.107
school, primary 1.655, 1.660–662, 2.73, 2.106–107, 3.78
school, Qurʾānic 1.335, 1.448–451, 1.487.1.660–662, 1.665, 2.51, 2.137, 3.135, 3.137–139, 3.113, 4.184, 4.276
school, religious 1.528, 1.533
school, secondary 1.654, 1.660–661, 2.109
school, Sino-Arabic 1.382
school curriculum 2.105, 3.347
school for the blind 1.316
school for the deaf 4.223–224
school for translators 1.171
school grammar 1.488
school of interpreters 1.171
school performance 3.75
school reform 2.111
school system, Arab 2.105
schools, Arab in Latin America 3.3
schools, Dutch 2.337
schools, Hebrew, in Morocco 3.295
schools, Malaysian 2.340
schools, Moorish 3.174
schools, Qurʾānic, in Senegal 4.183–184
schools in Cameroon 1.336–337
schools in Chad 1.365
schools in China 1.381
Schramm, Gene M. 4.339, 4.341
Schregle, Götz 3.27–28
Schreiber, Giselher 3.179, 3.404, 4.126
Schreiber, Peter A. 4.524–527, 4.726
Schriftsprache 1.615
Schröder, Christoph 2.74
Schub, Michael B. 4.66
Schulthess, Friedrich 2.379
Schultz, David E. 1.668, 2.661, 3.567
Schulz, Eckehard 3.323–324
Schumann, John 4.141
Schumans, J. 2.683
Schütze, Hinrich 1.456, 1.511, 4.166
schwa 2.484, 4.682–683
Schwartz, Bonnie D. 4.143
Schwarz, Klaus 3.502
Schwarzschild, Roger 4.316
Schweinfurth, Georg 1.519
science, transmission of 1.500
Scirha, Lydia 2.660
scope 2.116, 2.134–135, 2.144–146, 2.368–369, 4.316
scope, narrow 2.135
scope, wide 2.135
scope of quantifier 4.15
scopeless 4.316
scrambling 3.231
scribes 4.73, 4.481
scribes, Sicilian 4.218
script, Abbasid bookhand 3.110, 3.112
script, ambiguity of 1.153–154
script, Ancient Arabic 4.478, 4.482
script, Ancient South Arabian 1.148
script, *ʾandalusī* 3.110, 3.112, 3.519
script, Arabic 1.7, 1.133–165, 1.371, 1.379, 1.381, 1.446, 1.487, 1.641, 1.661–663, 2.40, 2.74, 2.263, 2.329, 2.334, 2.338, 2.340, 2.374, 2.533, 3.59, 3.77–78, 3.135–137, 3.139, 3.513–524, 4.5, 130–138
script, Arabic for African languages 2.673
script, Arabic for Afrikaans 4.291–295
script, Arabic for Amharic 2.53
script, Arabic for Berber 3.171
script, Arabic for Fulfulde 2.137, 4.179, 4.183
script, Arabic for Kazakh 2.557

script, Arabic for Malagasy 3.126
script, Arabic for Malay 4.291
script, Arabic for Mandinka 4.179
script, Arabic for other languages 2.672
script, Arabic for Persian 2.408, 2.673, 3.574
script, Arabic for Pulaar 4.183
script, Arabic for Somali 4.272, 4.277
script, Arabic for Swahili 4.381
script, Arabic for Tajik 4.424
script, Arabic for Tamil 4.433–434
script, Arabic for Tatar 4.450
script, Arabic for Turkish 2.673, 4.291, 4.584
script, Arabic for Urdu 2.673, 3.508, 4.595–598
script, Arabic for Uyghur 4.604
script, Arabic for Uzbek 4.608
script, Arabic for Wolof 4.179, 4.183
script, Arabic in Senegal 4.723
script, Arabic-Malayalam 3.129–130
script, Aramaic 1.148, 1.150, 1.153, 1.401, 2.673, 4.754
script, Berber 2.63, 3.294, 3.317, 3.350
script, *bihārī* 3.519
script, bilingual 1.371
script, broken cursive 3.519
script, broken Kufic 3.519
script, Canaanite 1.148
script, chancellery 4.98, 4.130, 4.560
script, Chinese 1.379, 1.381
script, Chinese in Arabic 1.381
script, Coptic 1.384, 1.495, 2.235
script, Coptic for Arabic 1.495
script, cuneiform 3.408
script, cursive 3.519
script, Cyrillic 1.133, 2.558
script, Cyrillic for Tajik 4.424
script, Cyrillic for Tatar 2.673, 4.450
script, Cyrillic for Uzbek 4.608
script, Dadanitic 1.148, 3.467, 3.470, 3.472
script, Dadanitic for Arabic 3.467
script, Dadano-Arabic 3.470
script, Dedanic → script, Dadanitic
script, Devanagari 2.334, 4.595
script, Early Abbasid 3.517, 4.130
script, Eastern Kufic 3.519
script, epigraphic 1.151
script, Epigraphic South Arabian 1.148
script, Estefi 1
script, Ethiopic 1.148
script, *fāsī* 3.112
script, grantha letters 4.434
Script, Greek 1.383, 3.472
script, Greek for Arabic 3.465, 3.472, 3.568
script, Hasaean → script, Ḥasaitic
script, Ḥasaitic 1.148
script, Hatran 1.148
script, Hausa 1.135
script, Hebrew 1.232, 1.446, 1.598, 2.266, 2.526–528, 2.531, 3.472
script, Hebrew for Arabic 3.218
script, Hebrew for Arabic Moroccan 3.274
script, *ḥijāzī* 1.151, 3.517–518
script, Hismaic for Arabic 3.468
script, Jawi 1.135, 2.334, 2.341, 2.343–344
script, Kashmiri 1.134
script, Kufic 1.151, 2.597–604, 3.517–518, 4.130, 4.134
script, Kurdish 1.134
script, Latin 1.133, 1.487, 1.600, 2.74, 2.518, 3.57, 3.60
script, Latin for Arabic 3.4, 3.317
script, Latin for Egyptian Arabic 3.349
script, Latin for Juba Arabic 2.518
script, Latin for Lebanese Arabic 3.349
script, Latin for Maltese 3.142
script, Latin for Moroccan Arabic 3.274
script, Latin for Somali 4.272
script, Latin for Tatar 4.450
script, Latin for Turkish 4.586–587
script, Latin for Uzbek 4.608
script, Liḥyanic 1.148
script, *mabsūṭ* 3.112

script, *maġribī* 3.110–113, 3.519, 4.130
script, Malay 1.135
script, *mašriqī* 3.112
script, Meccan 1.151
script, *mujawhar* 3.112
script, *musnad* 3.112
script, Nabataean 1.148–149, 1.151, 1.401, 2.263, 3.472
script, Nabataean Aramaic 4.478
script, Nabataean for Arabic 3.465, 3.468
script, Nabataeo-Arabic 3.471
script, *nastaʕlīq* for Urdu 3.338
script, *nasxī* 3.110, 3.519
script, New Abbasid Style 3.110, 3.112
script, New Style 4.134
script, North Arabian 1.148
script, North Arabian for Arabic 3.465
script, Ottoman 4.132
script, Ottoman Turkish 1.135
script, Palaeo-Hebrew for Arabic 3.218
script, Palmyrenian 1.148
script, Pashto 1.134
script, Persian 1.134
script, Persian Kufic 3.519
script, Phoenician 1.148, 1.150, 3.409, 3.411
script, proportioned 3.111, 4.98
script, Punic 3.141
script, *qayrawānī* 3.112
script, Qurʾānic 4.132
script, Roman 1.661
script, Roman for Lebanese Arabic 3.349
script, Roman for Swahili 4.381
script, round 2.600
script, Sabaean 3.467, 3.471
script, Sabaean for Arabic 3.465, 3.467
script, Safaeo-Arabic 3.471
script, Ṣafaitic 1.148
script, Ṣafaitic for Arabic 3.468
script, Sanskrit 2.334
script, *sertā* 3.110
script, Sindhi 1.135
script, *ṣīnī* 3.519
script, *sorabe* 3.126
script, South Arabian 2.258, 2.263
script, *sūdānī* 3.112, 3.519
script, Swahili 1.136
script, Swahili-Arabic 1.663
script, Syriac 1.148–151, 1.603, 3.110
script, Syriac for Arabic 3.218
script, Tamil 4.433–434
script, Thamudic 1.148
script, Thamudic for Arabic 3.469
script, Tifinagh 1.133, 1.294, 2.63, 2.673, 2.709, 3.171, 3.294, 3.317, 3.350–351
script, Turkic runes 4.450
script, Ugaritic 3.411
script, Urdu 1.135
script, Uyghur 1.135
script, *zimāmī* 3.112
script for Kurdish 3.351
scriptio defectiva 2.599, 4.4
scriptio plena 1.384, 3.220
scripture hall education 1.380
SD → structural description
seals 2.42, 2.45
search engine 1.455, 2.381
Searle, John R. 2.80, 2.202, 2.581, 3.570, 3.679, 4.328–329, 4.331
Seattle 3.397
Sebag, Paul 1.446, 4.573
Sebba, Mark 4.195
Sebha 3.53
Sebha Arabic 3.55
second language acquisition 1.368–375, 1.482–489, 1.595, 1.647, 1.651, 2.96, 4.138–146
second language learning 1.432, 1.594, 2.209, 3.641, 4.142, 4.144, 4.370
second language learning, massive 4.370
second language teaching 1.482, 4.138, 4.146–156
secondary focus 4.455
secondary meaning 4.455
secret language 4.156–160
sectarian dialect → communal dialect

secularization 1.531, 1.636, 2.338, 4.76, 4.578
secularization of grammar 2.701
sedentarization 1.259, 1.605, 1.611, 2.2–3, 2.414, 2.609, 3.402
sedentary dialect 1.259, 1.443–444, 1.569, 1.571, 1.589, 1.594–595, 1.605–609, 1.612, 2.532
sedentary speech 1.269
Seeger, Ulrich 1.583–584, 3.527, 4.621
Seereer 4.723
Seetzen, Ulrich Jasper 2.220
Sefrioui, Ahmed 3.294
Sefrou Arabic 2.534, 3.288, 3.292, 4.70
segment of speech 2.236
segmental error 4.268–269
segmental harmony 3.492
segmental matrix 3.615
segmentation 1.214
Segou 3.137
segregation 1.443
Séguy, Jean 1.586
Sehel 2.1
Seidensticker, Tilman 3.30–37
Seignobos, Christian 1.334
šekaste → *šikasta*
Sekik Arabic 3.607
sekolah 2.338
šəkr 3.173
Seleucia-Ctesiphon Madāʾin
self-awareness 2.58
Seliger, Herbert W. 2.683–684
Selim Bey 2.588
Selinker, Larry 4.139
Seljuks 2.406, 4.584, 4.589
Selkirk, Elizabeth 2.36, 2.395–396, 3.494, 3.611–612, 3.618, 3.726, 4.85
Sellheim, Rudolf 3.721
Sells, Peter 4.81
Selms, Adrianus van 4.291
Selte 2.53–55, 2.67
Selu, Jorge Hayeg 3.396
Semaan, Khalil I. 2.234, 3.90
semantic bleaching 1.23, 2.192, 2.194, 4.160–164, 4.197
semantic depletion 4.160
semantic devaluation 2.254
semantic domain 2.33
semantic engendering 4.331
semantic extension 1.175, 2.254, 2.466, 2.683, 4.164–169
semantic gain 4.163
semantic generalization 4.160, 4.166
semantic intensification 2.254
semantic interference 2.274
semantic link 4.101–102
semantic load 2.624
semantic mismatch 2.18
semantic networks 1.456, 1.458
semantic opposition 4.191
semantic reduction 4.160
semantic relationship 3.51
semantic role 1.622, 4.624
semantic shift 2.469
semantic shrinking 2.254
semantic specialization 4.167
semantic weakening 4.160
semantics 1.427–429, 1.458, 1.575, 2.627
semantics, conceptual 3.43
semantics, frame 4.624–625
semi-auxiliary 2.524, 3.236
semicolon 3.741
semi-idiom 3.630
semi-lingualism 2.703
semi-literacy 4.481
semi-speaker 2.684
semi-syllable 3.494
Semitic, case endings in 2.403
Semitic, Central 4.170–171
Semitic, East 1.188, 1.203, 2.262, 3.408–422, 4.170–171
Semitic, Ethio 1.37, 1.489–490, 1.624, 3.494, 3.553, 3.592, 4.1.4.170, 4.301–307, 4.309–310, 4.312–314
Semitic, Ethiopian → Semitic, Ethio-
Semitic, Middle 4.170–171
Semitic, Neo- 4.170
Semitic, Northwest 4.170–171, 4.301
Semitic, Proto- 1.313–314, 1.358, 1.467, 1.544, 1.563, 1.640–641, 1.643–644, 2.23, 2.67, 3.188, 3.227, 3.411, 3.446, 3.568, 3.591–593, 4.97, 4.170–173, 4.175–176, 4.178, 4.302

Semitic, South 1.38–39, 1.616, 2.262, 3.189, 3.420, 4.170–171, 4.300–315, 4.478, 4.756
Semitic, Southwest 2.262, 3.191, 4.171
Semitic, West 1.358, 1.616, 3.191, 4.171, 4.176, 4.178, 4.313
Semitic languages 1.35–36, 1.85, 1.166, 1.195, 1.203, 1.313, 1.490, 1.492, 1.494, 1.544, 1.560, 1.564, 1.573, 1.575, 1.583, 1.615–616, 1.623, 1.637–639, 1.643, 2.158, 2.223, 2.262, 2.462, 2.692, 3.48, 3.52, 3.64, 3.89, 3.188–189, 3.641, 3.738, 4.170–179, 4.371
Semitic languages, classification of 2.262
semivowel 1.205, 1.653, 3.599, 4.699–700
Sen, Dinesh Chandra 1.288
Sendor Arabic 2.415, 2.534
Senegal 1.255, 1.529, 2.137, 4.179–185, 4.508, 4.723–725
Senegal, Arabic in 4.179–185
Senegal River 2.240, 3.169, 3.174
Senfft, Joachim 3.717–718
Senglea 3.147
Sennar 2.559
sentence 1.474, 2.143, 2.148, 2.186, 2.358, 2.536, 2.538, 2.542, 4.185–191
sentence → *jumla*
sentence, assertive 4.639
sentence, complex 4.185, 4.187
sentence, compound 1.355
sentence, conditional 1.110, 1.554, 2.478, 2.516, 3.261, 3.272, 3.286, 3.537, 4.570
sentence, copular 1.507–508, 1.510, 3.68, 3.81
sentence, equational 1.197, 1.349, 2.432, 3.81, 3.590–591, 4.745
sentence, event-reporting 2.114, 4.503
sentence, exceptive 2.308, 2.451–452
sentence, existential 2.249, 2.597, 3.187, 3.260, 3.286, 3.369, 3.537, 3.638, 4.318, 4.358, 4.486, 4.747
sentence, garden-path 3.77–78
sentence, *ḥāl* 1.308, 1.332, 1.396, 1.465–466, 3.186, 3.537, 3.543
sentence, here-and-now 1.342
sentence, identificational 2.114, 4.503
sentence, interrogative 1.554, 1.609, 2.389–395, 2.478, 3.118, 4.330, 4.333, 4.570, 4.737, 4.740
sentence, locative 4.318, 4.486
sentence, multiclausal 4.360
sentence, negated 4.745
sentence, nominal 1.182, 1.346, 1.348–349, 1.398, 1.486, 1.509, 2.290, 2.430, 2.435, 3.81, 3.202, 4.354
sentence, nominal → nominal clause
sentence, passive 4.354
sentence, possessive 3.638
sentence, presentational 2.114
sentence, presentative 3.590
sentence, pseudoverbal 4.695
sentence, relative 2.118, 4.622
sentence, relative → relative clause
sentence, simple 4.185, 4.187
sentence, small 2.539
sentence, subject-initial 4.187
sentence, subjectless 1.623–625, 4.356–357
sentence, thetic 2.114
sentence, topic/comment 2.114, 4.354
sentence, ungrammatical 1.342
sentence, verbal 1.398, 1.486, 1.488, 1.623, 2.291, 2.430, 2.435, 3.102, 3.260, 4.354
sentence, verbless 2.562, 4.395–396, 3.81, 3.259, 4.499
sentence accent 2.116
sentence coordination 4.191–195, 4.646
sentence focus 2.113, 2.114–115
sentence linking 1.647–648
sentence processing 2.686
sentence-completion tasks 2.660
sentence-initial position 4.485, 4.488, 4.525, 4.731, 4.740

sententia 3.730
sentential connective 3.654
Separate Development Theory 1.368
Separation and Non-identity Principle 2.226, 2.453, 3.107
sequencer 3.365
sequential 1.432
sequential progression 1.651
Serbian, Arabic loanwords in 4.259–267
Serbo-Croatian 1.133, 4.371
Serbo-Croatian, Arabic loanwords in 4.259–267
Šeredāt 1.363
Serer 4.282
Serhedan 2.607
serial 1.528
serial verb 1.51, 3.187
serialization 1.295, 4.195–199
serialization of consonants 4.96–97
serif 3.308–309, 3.337, 3.341
Serjeant, Robert Bertram 3.731, 4.698, 4.159
sermon 2.10, 2.214, 2.649, 3.665, 4.73–74
sermon, Friday 4.73
sermons, language of 3.396
Serracino Inglott 3.153
Şeşen, Ramazan 3.503
set expression 2.136
settlement history 1.610
Seuren, Pieter 4.197
Seville 1.100
Sèvres, Treaty of 4.578
Sezgin, Fuat 3.36, 3.717
Sfax 1.23
Sfax Arabic 1.609, 1.642, 4.563–564
Sgroi, Salvatore 4.218
Shaaban, Kassim [Qāsim Šaʿbān] A. 2.694–707, 3.315
Shaalan, Khaled 1.460
Shaban, M.A. 1.618
Shabaneh, Osama 3.446
Shadily, Hassan 2.341
shadow theater 1.546, 2.469, 4.255
Shah, Mustafa 4.4–11
Shahabuddin, Syed 2.325
Shahan-Bek 4.613
Shah-i Zinda 2.46
Shahid, Irfan 3.6–8, 3.697, 4.402
Shahidullah, Muhammad 1.259
Shahin, Kimary N. 1.371, 2.229–231, 2.678, 3.526–538, 4.637, 4.653–654, 4.676
Shahpur 3.370
Shaiqi, Old 3.435
Shakespeare 1.600, 4.211
Shakir, Abdullah 2.501, 2.587, 4.467
Shalaby, Ahmad Q. 3.510
Shams-i Faxrī 2.407
shared constituent construction 4.438–439
Sharkawi, Muhammad al- 1.316–320, 2.116–122, 3.689–699
Ṩhäro᾽ 4.297
Shattuck-Hufnagel, Stefanie 4.268, 4.270, 4.335
Shawqi, Ahmad → Šawqī, ʾAḥmad
Shboul, Ahmad M.H. 3.674, 3.701
She Bashir Umar 2.55
Sheba → Sabaʾ
Sheba, Queen of 1.127
Shechem 3.409
Shehri 3.350
Sheikh, Hanan al- 1.529
Sheikholeslami, Cynthia 1.651
Shellabear, W.G. 2.340
Shelley, Percy Bysshe 4.91
Shemlan 4.148–149
Shengyu, mosque of 1.379
Shibarghān 1.28
shibboleth 4.3, 4.404
shift 4.324
shift → language shift
shift, /a/ > /i/ 3.361
shift, grammatical 4.24–25
shift, referent 1.649
shift, semantic 4.724
shift, thematic 1.648
shift, topic 1.432, 1.651
shifter 1.565, 1.567
Shihab, Alwi Abdurrahman 2.337
Shihab, Muhammad Quraish 2.337

Shiʿite speech 1.271–272, 1.445–446, 1.610, 1.596, 2.213
Shiʿites 1.271–272, 1.445–446, 2.41, 2.47, 2.583, 2.609, 4.633
Shiʿites, Ṭayyibi Musṭalawi Ismaʿili 2.328
Shiʿites, Twelver 2.328–329
Shiʿites, Zaydite 2.303
Shiʿites in Bahrain 4.633
Shikomoro 1.448, 1.662–663, 3.126
Shikomoro → Comorian
Shilha 4.1
Shilluk 1.519
Shim Arabic 2.314
Shimron, Joseph 1.573, 3.305
Shinar, Pessah 3.23
Shingazija 1.662
Shinnie, Peter L. 3.435
Shiraz 2.45, 2.572
S'hiri, Sonia 2.659, 4.320–328
Shivtiel, Avihai 3.538–541, 3.629–634, 3.658–663, 3.730–736, 4.286, 4.416–421
Shlonsky, Ur 1.312, 1.352, 1.481, 2.172, 3.232–233, 3.356, 3.386, 3.429, 3.432, 4.15, 4.360, 4.362–363, 4.366, 4.395, 4.400, 4.488, 4.525, 4.648, 4.650, 4.654, 4.657, 4.722, 4.743, 4.745
Shoa 2.52
Shoa, Sultanate of 2.52
Shohamy, Elana 2.438, 2.464, 3.313, 3.315
Shokeid, Moshe 2.64
Shona 4.386
Shonke 2.53
shop signs 3.316–318
Shorrab, Ghazi A. 3.318
shortening of final vowel 4.33
Shua Arabs → Shuwa Arabs
Shuayb, Tayka 4.433
Shunnaq, Abdullah 4.288
Shuwa 3.373
Shuwa Arabic 1.333–334, 1.338, 1.373, 4.284, 4.376
Shuwa Arabs 1.333–334, 1.362
shwa → schwa
Siagh, Zohra 1.56
Siam 2.334
Sibari, Mohamed 3.295
Sībawayhi, ʾAbū Bišr ʿAmr ibn ʿUṯmān 1.26–27, 1.71, 1.124, 1.173, 1.191–193, 1.309, 1.402, 1.433, 1.492, 1.506, 1.544–545, 1.555–557, 1.559, 1.617, 1.640–641, 2.16, 2.80, 2.82, 2.86–87, 2.90–94, 2.96, 2.165, 2.176, 2.178, 2.180, 2.184–186, 2.188–189, 2.223, 2.225–229, 2.232–233, 2.237–238, 2.280, 2.292, 2.295, 2.298–302, 2.307, 2.309, 2.311–312, 2.351, 2.354–356, 2.402–404, 2.406, 2.425–427, 2.432–436, 2.453, 2.459–460, 2.478–479, 2.495, 2.536, 2.541–543, 2.544–551, 2.555–556, 2.578, 2.624–625, 2.627, 2.629, 2.631, 2.676, 3.37, 3.72–73, 3.90–92, 3.96–97, 3.101–102, 3.104–109, 3.116, 3.123–125, 3.164–166, 3.209, 3.224, 3.320–321, 3.335, 3.377–378, 3.422, 3.455, 3.497, 3.542–543, 3.546–550, 3.565, 3.568, 3.601–603, 3.609, 3.648–649, 3.690, 3.704, 3.728–729, 4.1–2, 4.7–9, 4.10, 4.12, 4.21, 4.35, 4.101–102, 4.119, 4.121, 4.219–221, 4.235, 4.282, 4.284, 4.303, 4.329, 4.354, 4.410–416, 4.422, 4.430–432, 4.436–439, 4.441–443, 4.445–446, 4.495, 4.558, 4.636, 4.659–660, 4.667, 4.737–739
sibilant 2.257, 3.190–275, 4.172, 4.244, 4.549
sibilants, Semitic 4.172
Sicilian 2.455, 3.141–142, 3.144, 3.149, 3.151, 3.286, 4.216–218
Sicilian Arabic 1.416, 2.158, 2.630, 3.146, 4.215
Sicilian dialects 2.454
Sicily 2.26, 2.454, 2.671, 3.141–142, 4.215–219
Sicily, Arabic in 4.215–219
Siculi 4.215
Sīd al-Muxtār, aš-Šayx 3.172
Sidarus, Adel 3.22
Siddiqui, Abdul Majeed 4.452
side-head 1.3

Šidērāt 1.362
Sidi Bel Abbes 1.56
Sidi Brahim (ibn), Muhammadun 3.173
Sidi Kasem Arabic 3.288
Sīdi Muḥammad al-Kuntī, aš-Šayx 3.172
Sīdi Muḥammad wəll Ḥabät 3.172
Sidon 3.409
Sidon Arabic 1.269
Šidyāq, Fāris aš- 2.189
Siedle, Christine 1.586
Siegel, Jeff 1.593
Sieny, Mahmoud Esmail [Maḥmūd ʾIsmāʿīl Ṣīnī] 3.179, 3.404, 4.62, 4.124
Sierra Leone 1.255
Siewierska, Anna 3.66, 3.554
ṣifa 1.49–50, 2.82, 2.183, 2.426, 2.431, 2.435, 2.453, 3.101, 3.107–108, 3.113, 3.178, 3.261, 3.324, 3.422, 3.548, 4.219–222, 4.436, 4.442, 4.739
ṣifa, ism 3.323
ṣifa mušabbaha 1.50, 1.68, 2.432, 3.101, 4.220–221, 4.643–644
ṣifāt al-ḥurūf 4.426
ṣifāt Allāh 3.324, 4.219
Sifēra 1.362
ṣīġa 1.574, 2.13, 2.448
ṣīġat al-mafʿūl 3.558
ṣīġat at-taʿajjub 2.80
siglum 1.1
sign 1.573, 2.428
sign, arbitrary nature of 3.45
sign, homemade 4.223
sign, linguistic 3.188
sign, nonarbitrary nature of 3.50
sign language 4.222–235
Sign Language, Brazilian 4.223
Sign Language, Egyptian 4.223
Sign language, Gulf 4.223
Sign Language, Iraqi 4.223
Sign Language, Israeli 4.223
Sign Language, Jordanian 4.226
Sign Language, Levantine Arabic 4.223, 4.226–230, 4.233–234
Sign Language, Turkish 4.223
Sign Language, unified Arabic 4.226
sign language, village-based 4.223, 4.226
Sign Language, Yemeni 4.223
sign space 4.230–231
signal processing 1.209
signification, absence of 4.213
signified 3.160
signifier 2.623, 3.160
signing 2.675
signing, deaf 4.225
signing, hearing 4.225
Šihāb, Mufīd 1.629
Šihābī, Muṣṭafā as- 4.460
Šiḥḥī 2.211
Šiḥūḥ 2.211, 3.478
Siirt 1.87–95, 4.579
Siirt Arabic 1.538, 2.608, 3.585, 4.70
Sijistānī, ʾAbū Ḥātim Sahl ibn Muḥammad as- 1.173, 1.627–628
Sijpesteijn, Petra M. 3.513–524
šikasta 1.134, 3338
šikasta-amīz 3.338
šikasta-nastaʿlīq 3.337
ṣila 1.2, 1.49–50, 1.152, 2.478, 2.537, 4.221, 4.235–237
ṣila fī l-kalām 4.235
silence 2.579
Siloni, Tal 1.480, 3.388, 3.393, 3.431–432
Silte 4.301
Silva, Clare M. 3.206
Silverman, Daniel 2.365
Silvestre de Sacy, Antoine-Isaac 1.169, 1.449, 1.573–574, 2.358–359, 4.516
sīm 2.468, 4.252, 4.255
sīm → *luġat issīm*
sīm al-xawanīn 4.256
sīm il-ʿawālim 4.255
sīm il-fannānīn 4.255–256
sīm il-ḥarāmiyya 4.256
sīm il-kawanin 4.256
sīm il-miʿarriṣīn 4.256
sīm il-mumassilīn 4.255
sīm il-xarṭiyya 4.255
sīm in-naššālīn 4.256
sīm is-sāġa 4.255–256

sīm is-sakka 4.256
sīm it-tuggār 4.255
sīm xān il-xalīli 4.255
Sima, Alexander 4.479
Simeone-Senelle, Marie-Claude 1.654–656, 2.268–275, 3.62, 4.303–307, 4.750, 4.756
similarity avoidance 3.628
simile 1.425, 2.441–446, 3.117–118, 3.120–121, 3.632, 3.733
similitudo 3.121
Simon, Udo Gerald 2.360, 2.441–447, 3.116–123, 4.330
Simonet, Francisco Javier 1.99, 2.287
simplicity, linguistic 4.39
simplification 1.526, 1.596, 1.616, 1.633, 2.120–121, 2.158, 3.11, 3.640, 3.642–643, 4.162, 4.688
simplification, cluster 1.341, 2.35
simplification, phonological 2.679
simplification of Arabic 1.531
simplification of grammar 2.189, 2.701
simplified Arabic 1.632
Simpson, J.M.Y. 4.72
simultaneity 1.432, 1.466
sīn 2.468, 4.255
ṣināʿa 2.304
ṣināʿa lafḏiyya 2.623, 2.625, 2.627, 3.160
sinād at-tawjīh 4.88–89
Sinai 1.260, 1.383, 1.589, 3.360, 3.400, 4.237–251
Sinai Arabic 1.260, 1.266, 1.320, 2.2, 2.4, 2.388, 3.401, 3.405–406, 4.237–251, 4.663
Sinaouen 3.56
sincerity conditions 4.328
Sinclair, John 1.658
Sind 2.325, 3.506
Sindhi 1.133, 1.135, 1.665, 2.325, 3.506, 3.510
Singer, Amy 3.502
Singer, Hans-Rudolf 1.609, 2.388, 3.287, 3.737–738, 4.520, 4.563–465
Šingīṭi, ʾAḥmad ibn al-ʾAmīn as- 3.170, 3.172
Singkīlī, ʿAbd ar-Rāʾūf as- 2.335
singular 1.433–434, 3.439
singulative 2.156–157, 3.439, 3.445
Ṣīnī, Muḥammad Mākīn aṣ- → Ma Jian
Sini, Muhammad Nur al-Haqq ibn Luqman as- 1.381
Sinneh 2.604
Šinqīṭ 3.169
sins of language 2.362
Sinte 2.73
Sipāhīzāde 3.502–503
šiʿr 2.543, 3.173, 3.645, 3.648, 4.199–214
šiʿr ḥurr 4.91
šiʿr mursal 4.91
sīra 1.338
Sirach 3.730
Sīrāfī, ʾAbū Saʿīd al-Ḥasan ibn ʿAbdallāh as- 2.82, 2.86, 2.186, 2.311–312, 2.427, 2.436, 2.632, 3.92, 3.101, 3.104, 3.116, 3.177, 3.648–650
Siraiki 3.506
Sirāj Šīrāzī 3.337
širāk 2.237
Šīrānī, Ḥāfiẓ Maḥmūd 3.508
Sīrat Banī Hilāl 3.500
ṣirāṭ mustaqīm 2.187
Šīrāzī, aš- 2.45
Šīrāzī, Rūzbihān 3.341
Šīrāzī, Viṣāl-i 3.341
Širbīn 2.4
Širbīnī, Yūsuf aš- 1.598, 2.189, 2.308, 3.101, 3.108, 3.453, 3.550, 4.255
Širgāṭ Arabic 2.419
Širjān 2.256–257, 2.260
šīš qalam 2.598
Šittiye 1.362
situationality 4.472, 4.475
Siverek 1.87
Siwa 2.1, 2.5
Siwi 1.290, 2.1, 2.5, 3.313
Six Pens 2.598, 2.601, 4.560
ṣiyaġ al-ʿuqūd wa-l-fusūx 2.360
siyāq 4.22, 4.28

Sïzdïqova, Rèbiġa 2.557
Skalička, Vladimir 3.192
Skalmowski, W. 2.407
Skik, Hichem 4.572
Skikda 1.54
Skikda Arabic 1.609
Skok, Petar 4.259
Skousen, Royal 4.471
Skūra 3.287
Skūra Arabic 3.288, 3.290, 3.293, 4.71
Skutnabb-Kangas, Tove 2.704
slang 2.217, 4.156–157, 4.251–259, 4.764
slang, teenage 4.253–254
slave trade 1.519, 2.57, 2.60
slaves 2.588
Slavonic languages 1.50, 1.196–197, 2.684, 3.524, 4.259–267
Slavs 4.259
slip, action 4.271
slip of the ear 4.271
slip of the eye 4.271
slip of the hand 4.271
slip of the tongue 4.267–272, 4.334
Slobin, Dan Isaac 1.485, 2.97, 2.100–102
slot 1.457
Slovak, Arabic loanwords in 4.259–267
Slovenian 4.194
Slovenian, Arabic loanwords in 4.259–267
sluicing 2.19–20
Small Clause Analysis 1.507–508
Smart, Jack R. 2.215, 3.478, 3.643
Smeaton, B. Hunter 2.214
Smith, G. Rex 3.471, 4.483
Smith, Ian 2.518
Smith, Norval 2.389
Smith, W. Robertson 2.583
Smith Kocamahhul, Joan 4.577–583
Smolensky, Paul 3.491, 3.493, 3.496
Smooha, Sami 2.438, 2.441
Smrž, Otakar 1.458, 516
Šnaidauf, Jan 1.458
Snouck Hurgronje, Christiaan 1.6–8, 3.179
Snoussi, Mohamed Larbi 1.447
Snow, James A. 4.524–527
soap opera 2.609, 3.19, 4.326
Sōba 1.361
Sobelman, H. 3.429
social dialect 2.498
social distance 1.650
social hierarchy 2.658
social interaction 1.647, 1.649–650, 2.205, 2.587
social marking 2.320
social network approach 4.629
social status 2.205
socialization 1.527
sociative construction 1.358
Socin, Albert 3.718, 4.516
sociolect 2.217
sociolinguistics, variationist 1.445–446, 3.604–605
Socrates 2.46
Soddo 4.301, 4.304–305, 4.310
Soden, Wolfram von 1.638, 3.738, 4.1
software applications 2.380
Sogdiana 3.517
Sohāg Arabic 1.306, 2.2
Sokna 3.53, 3.56
Sokoto 1.337, 2.250
Sokoto Caliphate 3.375–376
solar letters → sun letters
solidarity 2.205–206, 2.587, 2.661, 4.467, 4.469
solidarity, group 4.251
Solomonic dynasty 2.51
Soltan, Usama 4.396–397, 4.535–542
Somali 1.35, 1.451, 1.655, 2.51, 2.54–55, 3.314, 3.316, 4.1, 4.272–275, 4.381
Somalia 1.451, 2.105, 3.314–316, 3.351, 4.272, 4.275–278, 4.508
Somalia, Arabic in 4.272, 275–278
Somekh, Sasson 3.539, 4.77, 4.499
Somel, Selçuk Akşin 3.503–504
Somers, H.L. 1.343
Somerset 2.498
Soncino, Gerson 1.170
Songhay 1.36, 1.256, 2.668, 3.59, 4.279–282

Songhay Empire 4.279
Soninke 3.169–171, 3.314, 4.723
Sonni Ali Beer 4.279
sonority 3.621, 4.2
sonority, scale of 2.310
sonority hierarchy 1.86
sonority scale 3.620–621
sonority sequencing 2.36
soothsaying 4.104–105, 4.203
Soqoṭra 4.297–298, 4.301, 4.373, 4.750
Soqoṭri 1.203, 2.66–67, 3.63, 3.350, 4.297, 4.299–305, 4.307, 4.309–311
sorabe 3.126
Ṣōran Arabic 4.406
Sorani 2.604–608, 3.351
Soravia, Giulio 4.272, 4.274
šoreš 1.574
Sorman Arabic 3.55
Souag, Mostafa Lameen 4.648, 4.656
Soucek, Priscilla P. 3.308–309, 3.337, 4.560–562
Soudan, Frédérique 2.42
Soueif, Ahdaf 1.529, 2.703
Soukhne 4.403, 4.406
Soukhne Arabic 1.538, 1.585, 1.607–608, 2.314, 2.494, 4.406–408, 4.678
sound change 1.74, 1.77–78, 2.33, 3.204, 3.604–606
sound correspondence 1.36, 2.66
sound effect 4.287
sound law 3.190
sound symbolism 4.50, 4.93, 4.226, 4.254, 4.282–290, 4.684
sounds, animal 4.282–284, 4.286
sounds of birds 4.283, 4.285
source 1.344–345, 1.622–623, 4.487, 4.490, 4.492–493
source language 1.437, 1.482, 2.209
source text 1.432
Sourdel, Dominique 3.6
Sourdel-Thomine, Janine 3.513, 3.519
Souriau, Christiane 3.56
Sous 3.287
Sous Valley 1.293
Souss 2.62
Sousse Arabic 1.609, 4.563
South Africa 2.439, 4.290–295
South Arabia 1.128
South Arabian 1.126, 1.493, 1.545, 1.617, 1.639, 2.263, 2.265, 3.446, 3.453, 3.693–694, 4.127, 4.444, 4.478
South Arabian, Ancient 2.23, 2.257, 2.262, 3.420, 3.593, 4.175, 4.295–297, 4.301, 4.698, 4.756
South Arabian, Epigraphic 1.468, 2.262, 3.189, 4.170–171, 4.173, 4.295–297, 4.301–302, 4.305, 4.307, 4.309–314, 4.478–479
South Arabian, Modern 1.38–39, 1.188, 1.544, 1.639, 2.66, 2.211–212, 2.262, 2.668, 3.63, 3.420, 3.478, 3.593, 3.738, 4.123, 4.170–172, 4.175, 4.177, 4.297–301, 4.303–307, 4.309–310, 4.313–314, 4.750, 4.756
South Arabian, Old → South Arabian, Ancient 1.259–260
South Arabian script 1.149
South Semitic languages 1.38–39, 1.616, 2.262, 3.189, 3.420, 4.170–171, 4.300–315, 4.478, 4.756
Southwest Arabian Arabic 1.259–260
Southwest Semitic languages 2.262, 3.191, 4.171
Southworth, Franklin C. 4.596
Sowayan, Saad 1.273, 3.327, 3.500
space, notion of 2.645
space, private 2.646
space, public 2.646
space, signing 4.230
space dichotomy 2.645–646
spacing of letters 2.599
Spain 2.72
Spain, Arabic in 2.73
Spanish 1.122, 1.260, 1.448, 1.497, 1.545, 2.26, 2.75–76, 2.282–286, 2.319, 2.470, 2.491, 2.504, 2.532, 2.659, 2.671, 3.1, 3.3, 3.228, 3.312, 3.314, 3.579, 3.706, 3.725, 4.53, 4.75, 4.157, 4.570, 4.599

Spanish, Old 2.283
Spanish, teaching of 3.295
Spanish in Morocco 3.294–295
Spanish Sahara 2.240
Spanish/Arabic dictionary 3.23
Sparse Morphology Hypothesis 2.676
spatial marking in sign language 4.229
spatial terms, acquisition of 1.377
spatio-temporal indexing 1.649
Spaulding, J.L. 3.435
speaker comments 1.476
speaking, ethnography of 1.415
Spear, Russell 3.129
Spear, Thomas 4.381–382
Speas, Margaret 3.685, 4.742
Spec-Head agreement 1.47–48
Spec-Head relation 1.46, 4.397
specialization 2.195
species 1.433
Specific Language Impairment 2.675–676, 2.678–680, 2.689, 4.145
specific term 4.685
specification 2.225, 3.672–673
specificity 1.283, 1.304, 1.331, 1.565, 1.579, 2.370, 3.83, 3.186, 3.284, 3.382, 3.385, 3.428–430, 3.535, 3.687, 4.52, 4.315–320, 4.654
specificity, descriptive 4.316, 4.319
specificity, epistemic 4.316–319
specificity, partitive 4.316
specificity, quantificational 4.316–317, 4.319
specificity, scopal 4.316
specifier 1.348–350, 1.479, 3.230, 4.394, 4.396, 4.488, 4.746
specifier/head relation 1.348
specifying complement 3.451
spectral property 4.666
speech 2.542
speech, adult/child 1.339–341
speech, affective 2.79
speech, Alawi 1.388
speech, allegro 1.320
speech, aphasic 4.96
speech, autonomous 4.58
speech, Bedouin 2.506, 4.8
speech, Christian 1.271, 1.273, 1.323, 1.388, 1.443–448, 1.492, 1.537, 1.547, 1.595, 1.610
speech, classroom 1.419
speech, Coptic 1.323
speech, corpus of spoken 1.214, 1.460
speech, correctness of 2.176
speech, corruption of 2.177, 2.184, 3.59
speech, deaf 2.677–678
speech, delayed 2.678
speech, direct 1.649, 2.331–332
speech, *fallāḥī* 2.506–508
speech, female 1.441, 1.446, 1.529, 1.547, 1.572, 1.596, 1.609, 1.642, 1.650, 2.587, 4.70, 4.289, 4.631, 4.754–755
speech, figurative 1.425
speech, formal 2.377
speech, Hindu 1.286–287
speech, indirect 1.367, 1.649, 2.331–333
speech, interdialectal 1.669
speech, intergroup 4.321–322
speech, Jewish 1.323, 1.443–448, 1.492, 1.595, 1.609–610, 2.531
speech, lento 1.322, 1.325
speech, levels of 1.405, 1.409
speech, *madanī* 2.506–508, 2.514
speech, male 1.658, 4.631–632
speech, Muslim 1.271–272, 1.286–287, 1.323, 1.443–448
speech, non-autonomous 4.58
speech, of God 2.306, 4.6, 4.73, 4.105
speech, oral 4.75
speech, origin of 2.178, 4.684–685
speech, political 1.406, 3.663, 3.670
speech, production of 4.334–335
speech, reported 2.331, 3.670
speech, reporting 2.331
speech, represented 2.332
speech, rural 2.658, 3.11, 4.322
speech, Shiᶜite 1.271–272, 1.445–446, 1.596, 1.610, 2.213
speech, simplified 1.342

speech, Sunni 1.271, 1.273, 1.278, 1.283, 1.388, 1.445–446, 1.596, 1.610
speech, tropical 3.116
speech, urban 1.390, 1.417, 1.441, 1.443, 1.446, 2.318, 2.322, 2.658, 4.322
speech, women's 1.243, 1.247, 1.272, 1.278, 2.8, 2.33–34, 4.380, 4.563–564, 4.681
speech, youth 1.446, 1.596, 2.8, 4.251–252, 4.254, 4.630, 4.763–767
speech accommodation 1.271, 1.419, 1.444, 1.446, 2.659, 3.9, 3.12, 4.320–328, 4.628
Speech Accommodation Theory 4.320–328
speech accommodation, convergent 4.321
speech accommodation, divergent 4.321
speech act 3.572, 3.677, 4.328–334
speech act, constative 4.331
speech act, conventional 2.81
speech act, illocutionary 4.328–333
speech act, indirect 2.80–81, 2.581, 3.681
speech act, locutionary 4.328–329, 331–332
speech act, performative 4.331
speech act, perlocutionary 4.328, 332
Speech Act Theory 4.330–331, 4.333, 4.472
speech acts, expressive 2.202
speech coding algorithm 1.211
speech community 1.439–441, 1.633, 1.635, 1.652, 2.651, 2.662, 4.58
speech data, Arabic 1.511
speech defect 2.676
speech deficit 2.676
speech disorder 1.375
speech disorder → language disorder
speech error 2.629–633, 2.636, 2.691–692, 4.334–338, 4.267–272
speech errors in Chinese 4.268
speech errors in English 4.268, 4.335–337
speech errors in German 4.268, 4.337
speech errors in Thai 4.268
speech generation 1.455–456
speech pathology 2.688
speech processing 1.209, 1.456, 1.515
speech processing → automatic language processing
speech production 3.594, 4.269–270
speech recognition 1.207, 1.213, 1.455–456
speech recognizer 1.214
speech strategy 1.369
speech style 4.321
speech synthesis 1.210, 1.212, 1.460
speech therapy 1.372, 4.223
speech time 1.199–200
speech-to-speech translation 1.461
speech-to-text conversion 1.458
speech-to-text converter 1.515
Speiser, E.A. 2.13, 2.156
spelling, Chadian Arabic 1.364–365
spelling, Classical Arabic 1.400–401
spelling checker 1.455
Spell-out 1.481, 3.228
Spence, Nicole 3.204
Spencer, Andrew 3.298, 4.624
Spey, Rutger 1.169
spirantization 1.179, 1.205, 1.295, 1.314, 2.270, 3.293
spirantization, Neo-Aramaic 3.372
spiritus asper 2.201
Spiro Bey, Socrates 2.457
Spitaler, Anton 2.57, 3.689, 3.720
Spitta[-Bey], Wilhelm 1.602, 4.727
split → phonological split
split, allophonic 3.606
split, phonological 3.190, 3.605–607
split, vocalic 3.607
Split VP Hypothesis 4.650
Split-Complementizer Hypothesis 4.362
splitting 2.194
Spoer, Hans H. 2.481

Spolsky, Bernard 2.110–111, 2.438, 2.464, 2.661, 3.313, 3.315
spontaneity 3.666
Spooner, Brian 3.337–338
Spooner, William 4.267
Sportiche, Dominique 2.170, 3.685, 4.394, 4.397, 4.524, 4.648, 4.743
Sprachbund 4.170
Sprachinsel 4.612
Sprachinsel → language island
Sprachlandschaft 1.541
spread, emphasis 1.105, 1.260, 1.278, 1.302, 1.325, 1.548, 2.365, 2.482, 2.529, 2.612, 3.65, 3.181, 3.276, 3.363, 3.481, 3.527, 3.529, 3.600, 3.610, 4.118, 4.637, 4.669–670
Sproat, Richard 3.525
Sprouse, Rex A. 4.143
Spuler, Bertold 3.7, 3.720, 4.612
Sranan 4.195
Sreedhara Menon, A. 3.129
Sri Lanka 2.335, 3.506
Sri Lankans 4.123
Stace, E.V. 4.750
Stachowski, Stanisław 4.584
Stager, Sheila 2.688
Stampe, Dennis W. 3.570, 3.679
standard 4.322
standard, interregional 1.631, 1.668
standard, neutral 1.446
standard, regional 3.9
standard language 1.630, 2.104, 2.111
Standard Theory 1.457, 4.521–522, 4.525, 4.535
standardization 1.364, 1.397–402, 1.618, 2.176, 2.640, 2.647, 3.9, 3.11, 3.37, 3.347, 4.629
standardization, language 2.700
standardization of dialect 2.656
standardized variety 2.657
Stanley, Henry Morton 1.519–520
Stanley, Tim 3.110, 3.112, 3.308, 3.337
stanza 4.92
Starcky, Jean 1.149
Stark, Werner 4.72
Starostin, Sergei 4.313
stars, names of 2.25–26, 2.282, 2.455
Stassen, Leon 3.85, 3.674
state 1.195, 1.197, 1.219
state, non-sequenced 2.209
state of affairs 1.197
statement, veracity of 1.202
stative 1.197, 1.200–201, 1.358, 1.624, 2.348, 3.252, 4.176
stative in Coptic 1.499
stative passive 2.593, 2.596
stative verb 1.12, 1.42, 1.52, 1.108, 2.595, 2.618, 3.247, 3.252, 3.545, 3.643, 4.456
status 1.201
Stavi, Jonathan 4.15
Steenbrink, Karel 2.333–339
Steiger, Arnald 1.99
Steiner, Richard C. 1.545, 4.172
stem 1.574, 1.578, 4.338–344
stem, imperfect 1.574
stem, perfect 1.574
stem template 2.155
stem vowel 2.367, 3.250
Stemberger, Joseph P. 2.676, 4.335–337
stemmer 1.515
stemming 1.459, 1.514
Stempel, Reinhard 3.187–192, 3.227, 3.453
Stepanov, Arthur 3.388
Steppat, Fritz 4.150
stereotype, group 4.322
stereotyped language 2.675
Steriade, Donca 2.365, 3.205, 3.628
Stern, Samuel Miklos 3.307, 3.309, 4.561
Stetie, Salah 2.703
Stetkevych, Jaroslav 1.532, 1.633, 2.104, 3.425, 4.461, 4.558
Steuerwald, Karl 4.587
Stevens, Kenneth 2.229, 2.365
Stevens, Paul B. 2.660, 3.681
Stewart, Ahmad 4.103
Stewart, Devin J. 3.681, 4.104, 4.106
Stewart, Frank Henderson 3.361, 3.401, 3.403, 3.405, 4.238, 4.663

Stewart, Tony K. 1.286
Stiehl, Ruth 4.402
stigma 2.711
stigmatization 2.320, 2.712, 2.714, 3.11, 4.628–629
stigmatized 2.650, 2.652, 2.658, 2.661, 2.708
Stillman, Norman A. 2.64, 2.534, 3.287, 3.292
stimulus, poverty of 1.342
Stiver, Dan R. 4.72
Stoa 3.121
Stocqueler, J.H. 2.572
Stoel-Gammon, Carol 2.688
Stoetzer, Willem F.G.J. 3.208, 4.87
Stolbova, Ol'ga V. 2.68
stop, aspirated 4.666
stop, unaspirated 4.666
stop, voiceless 4.666
stop consonant 3.206, 3.596–597
Stop Lenition Constraint 1.540
Stork, F.C. 2.378
Storrs, Ronald 1.536
story-telling 3.665
Stowasser, Karl 3.655, 4.647
Stowasser-Freyer, Barbara 3.719–720
Stowell, Timothy 2.369, 3.685, 4.441, 4.743
Strabo 2.66
stranding 2.20, 4.649
stranding of preposition 2.390, 3.560
Strange, Winifred 4.669
strategy, avoidance 2.119, 2.587
strategy, camouflage 2.218
strategy, communication 2.207, 4.141
strategy, discourse-organizing 3.665
strategy, intensification 2.204
strategy, learning 4.141, 4.150
strategy, problem-solving 2.196
strategy, public-speaking 3.669
strategy, teaching-learning 4.140
strategy in foreigner talk 2.117
Strawson, Peter F. 4.503
Streck, Bernhard 2.218–220
Streck, Michael P. 3.227–228, 3.410
street signs 2.661, 3.316
stress 1.30, 1.63, 1.89, 1.104, 1.113, 1.211, 1.225, 1.234–235, 1.244, 1.261, 1.278, 1.301, 1.323, 1.390, 1.521, 1.540, 1.548, 1.616, 2.4–5, 2.7, 2.37, 2.117–118, 2.120, 2.242, 2.350, 2.353, 2.366, 2.417, 2.483, 2.509, 2.520, 2.561, 2.574, 2.590, 2.611–612, 3.66, 3.180, 3.240, 3.494, 3.612–615, 3.725, 4.108, 4.344–353, 4.675
stress, antepenultimate 4.347–349
stress, contrastive 4.354–355
stress, enclitic 4.348
stress, final 1.265
stress, fixed 3.726
stress, free 3.725
stress, grammatical 4.379
stress, iambic 4.351–352, 4.389
stress, lexical 2.395–396
stress, main 4.494, 4.496
stress, opaque 4117, 4.344, 4.348
stress, penultimate 4.349, 4.675
stress, primary 4.497
stress, secondary 4.346, 4.494, 4.496–497
stress, sentence 4.494
stress, tonic 3.725
stress, trochaic 4.390
stress, word 1.325
stress assignment 3.617, 3.620–622
stress fluctuation 4.108
stress group 1.540
stress in Cairene Arabic 4.345–346
stress in Classical Arabic 4.349–350
stress in Damascene Arabic 4.348
stress in Italian 2.458
stress in Lebanese Arabic 4.348–349
stress in Moroccan Arabic 4.351–352
stress in Negev Arabic 4.351
stress in Palestinian Arabic 4.346–349
stress in Sudanese Arabic 4.378
stress in Western Arabic 4.350–351
stress marking 1.515
stress rule, Eastern Arabic 4.347–348
stress rule, three-syllable 4.350

stress shift 1.522, 1.524, 2.366, 2.592–593, 3.621–622, 3.725, 4.564
stress subordination 4.675
Strickland, Gerald 3.141
Stroomer, Harry 2.211
structural change 4.522
structural description 4.522
structural difference 1.483
structure, abstract 4.657
structure, composite 4.624
structure, exocentric 4.647
structure, flat 4.657
structure, information 4.653, 4.656
structure, non-configurational 4.524
structure, non-headed 4.657
Stumme, Hans 3.146, 4.563
stuttering 2.675–677, 2.688
style, narrative 1.553
style, ornate 1.651
style books 3.200
style shifting 2.33, 3.665, 4.628
style switching 1.439
style variation 4.323
style variety 4.323
Suaieh, Saadun 4.63
Subayʿ 3.326
subcategorization rule 1.418, 1.635
Ṣubḥ 2.46
subiectum 3.177
subjacency 4.718, 4.721
subject 1.40–46, 1.49, 1.182–184, 1.186, 1.345–350, 1.354, 1.356, 1.412–413, 1.428, 1.474, 1.479, 1.485–486, 1.488, 1.558, 1.563, 1.622–624, 1.626, 2.146, 2.149, 2.186, 2.316, 2.348, 2.369, 2.430, 2.434–435, 2.550, 3.113–115, 3.177, 3.224–226, 3.231, 3.381, 3.385, 3.683, 4.354–359, 4.484, 4.535
subject, anaphoric 4.653
subject, broad 3.386–387
subject, clausal 4.359
subject, conjoined 3.69, 4.397
subject, delayed 4.358
subject, dummy 1.623, 1.625
subject, embedded 4.54, 4.56
subject, expletive 3.562, 4.359, 4.394, 4.396
subject, focalized 1.563
subject, focusing 4.732
subject, grammatical 4.653
subject, impersonal 3.563, 4.359
subject, inanimate 1.220
subject, indefinite 2.132–133, 2.369, 3.82, 3.201, 4.354, 4.356, 4.366, 4.400, 4.656
subject, isolated natural 4.484
subject, logical 1.217, 4.626
subject, matrix 4.365
subject, natural 4.484–485
subject, nonspecific 3.82
subject, null 4.363, 4.492
subject, null expletive 4.365
subject, overt 1.217
subject, passive 4.535
subject, postponed 2.149
subject, postverbal 4.318
subject, pronominal 3.710
subject, prototypical 2.146
subject, psychological 4.484
subject agreement 1.183
subject assignment 3.554
subject movement 4.394
subject NP 1.348
subject position 4.395
subject relativization 2.174
Subject Right Dislocation 4.541
subject shift 4.727
subject switch 3.711
Subject-Auxiliary-Inversion 2.390
subjecthood 1.185, 2.371, 3.81–82
subjective mood of speech 2.81
subjectless construction 1.625
subjunctive 1.38, 1.71, 1.468, 1.524, 1.562, 2.22–23, 2.92, 2.490, 3.98, 3.191, 3.248–249, 3.263, 3.267–269, 3.272–273, 3.298, 3.355, 4.360–363, 4.639–640
Sublet, Jacqueline 3.378–379, 3.718, 3.720
subordinate clause 4.734
subordinating particle 1.669

subordination 1.280, 1.404, 1.417, 1.468, 1.472–474, 1.476, 1.648, 2.301, 3.268, 4.60, 4.191, 4.194, 4.360–367, 4.485, 4.501, 4.727
subordinative, Akkadian 4.178
subordinator 4.113
subsegmental property 3.206
substance 2.432
substantive noun 1.13, 1.16, 1 .18
substantivization 1.18
substitute, pronoun 4.468
substitution 1.430–431, 2.31, 2.691, 4.647
substitution, lexical 4.269–270
substitution, semantic 4.336
substitution, sound 3.14, 4.335
substitution, word 2.690
substitution error 2.688–689
substitution test 4.647
substratal influence 1.84, 1.102, 1.104, 1.241, 1.619, 2.160, 2.262, 2.373, 3.189, 3.640
substratal influence, Aramaic 1.469, 1.536, 1.538–541, 1.612, 2.665, 3.159, 3.584, 3.674, 3.703
substratal influence, Basque in Spanish 4.369
substratal influence, Berber 1.294–295, 1.609, 2.621–622, 2.665, 3.279–280, 3.290, 3.293
substratal influence, Berber in Arabic 4.369
substratal influence, Berber in Tunis Arabic 4.565
substratal influence, Coptic 1.332, 1.501–505, 2.665
substratal influence, Coptic in Egyptian Arabic 4.370
substratal influence, Dravidian in Indian languages 4.369, 372
substratal influence, Hebrew 2.665
substratal influence, Late Latin 1.296
substratal influence, Mehri 3.63
substratal influence, Nubian 3.60
substratal influence, Romance 1.106–107, 2.287
substratal influence, Syriac 1.284
substratal influence in Maltese 3.141
substratal influence in Persia 2.665
substratal influence of Arabic in Dutch 4.370
substratal influence of Aramaic in Palestinian 4.373
substratal influence of Aramaic in Syrian Arabic 4.372, 4.403
substratal influence of Barbacoan languages in Quechua 4.374
substratal influence of Bari in Juba Arabic 3.643
substratal influence of Berber in Djidjelli Arabic 1.189, 4371
substratal influence of Celtic in Germanic language 4.373
substratal influence of Celtic in Romance language 4.373
substratal influence of Coptic in Egyptian Arabic 4.372
substratal influence of Cushitic in Ethiopic 4.371
substratal influence of South Arabian in Yemeni Arabic 4.756, 4.310, 4.313
substratal influence of Southsemitic in Yemeni Arabic 4.373
substratal influence of Spanish in Maghrebi Judaeo-Arabic 2.532
substratal influence of Uralic in Baltic languages 4.372
substrate 1.489, 1.526, 2.265, 3.176, 3.642–643, 4.367–375
substrate, Berber 2.240, 3.674
Sudan 1.24, 1.83, 1.206, 1.361, 1.518–519, 1.656, 2.1, 2.59, 2.105, 2.107, 2.137, 2.250, 2.255, 2.268, 2.517–518, 2.588, 2.700, 3.314, 3.351, 3.642–643, 4.375–381
Sudan, Arabic in 1.260, 1.611, 2.272, 2.454, 2.456, 2.494, 3.429, 3.525, 3.674, 4.310, 4.375–381
Sudan, Southern 1.519–521, 3.60
Sudanese Arabic 1.24–25, 1.300, 1.363, 1.525, 1.570, 2.2, 2.5, 2.38, 2.51, 2.160, 2.255, 2.517–519, 2.590, 3.10, 3.435, 3.437, 3.635, 3.643, 4.376–377, 4.380, 4.559, 4.715, 4.756
sūdānī 3.112
Sudanic languages 1.519

Sudanic languages, Eastern 3.435
Sudayr 3.326
Sudayr Arabic 4.125–127
šudūd 4.22
Sueif, Ahdaf → Soueif, Ahdaf
Suez Canal 2.337, 4.237
suffix 1.208, 1.555, 1.557, 1.576, 2.21, 3.297
suffix, anticipatory 1.230, 1.284
suffix, indirect object 1.115
suffix, linker 4.711
suffix, object 1.105, 1.226, 4.620–621, 4.714
suffix, possessive 1.105, 1.115
suffix, pronominal 3.586
suffix, subject 4.620–621
suffix, verbalizer 4.606
suffix, verbalizer in Uzbek 4.610
suffix conjugation 4.454–456
suffix conjugation, Semitic 4.178
suffix conjugation in Northwest Semitic 3.416
suffixation, double 1.263
suffixed form 4.638
suffixes, Romance in Maltese 3.153–154, 3.157
suffixes, Turkish 4.592
Sufi, G.M.D. 3.508
Sufi, Sheikh 4.277
Sufi order 1.380–381, 3.170
Sufi poetry 4.75
Sufi saints 4.181
Sufis 2.328, 3.122, 4.181, 4.204
Suhāg Arabic → Sohāg Arabic
Suharto, President 2.337
Suhaylī, ʾAbū l-Qāsim ʿAbd ar-Raḥmān ibn ʿAbdallāh as- 2.90, 2.181, 2.316
Suhrawardī, ʾAḥmad as- 3.308, 4.204
ṣuḥufiyyūn 1.148
Suhūl 3.326
Sukhnah 2.498
Šukrī, ʿAbd ar-Raḥmān 4.90–91
Šukrī, Al-Xūrī 1.603
Šukriyya 1.24, 1.190
Šukriyya Arabic 2.388, 3.62, 3.404, 3.544, 4.376
sukūn 1.152, 1.207, 1.310, 2.92, 2.238, 2.402, 3.91, 3.208, 4.518
sukūt 2.541
Sulaimania 2.604
Sulawesi → Celebes
Sulaym, Banū → Banū Sulaym
Sulaymān 1.362
Sulaymī Arabic 1.55, 1.609, 4.550
Suleiman, Michael W. 2.60, 3.311
Suleiman, Saleh M. 2.501
Suleiman, Yasir 1.173–178, 1.193, 1.530, 2.60, 2.105, 2.177, 2.181, 2.187, 2.310, 2.425, 2.654–656, 2.658, 2.661, 2.642, 3.313, 3.346, 3.349, 3.664, 3.669, 4.12, 4.323, 4.325, 4.475, 4.578
Süleyman I, Sultan 2.40, 2.601, 3.503
Ṣūlī, ʾAbū Bakr Muḥammad ibn Yaḥyā aṣ- 3.307, 4.132
Ṣulṭ Arabic → Ṣalṭ Arabic
Sultan Iskandar Muda 1.5, 1.8
Sulṭānī, ʿAlī ibn Ḥasan as- 3.337
Sulu 1.133
ṣuʿluk 4.208
Suluk Gatoloco 2.336
Suluk Wujil 2.336
sulus 3.341
Šūmān, Masʿūd 1.600
Sumatra 2.340
Sumerian 1.560, 4.33, 4.509
summons 2.202
sun letter 1.187, 1.204, 1.262, 1.301, 1.325, 1.548–549, 2.243, 2.483
sundesmós 3.547, 4.235
Sung Dynasty 1.378–379
sunna 2.187, 3.651
Sunni 1.271, 1.275, 1.445–446
superheavy 3.614
Superiority Constraint 4.84
superiority of Arabic 2.88
superlative 1.18, 2.13, 2.15, 2.523
superstratal influence 3.189, 3.643
superstratal influence, Romance 2.288
superstrate language 3.639, 3.643, 4.368–369, 4.371, 4.373
suppletion 1.522, 1.553, 2.592, 4.660–661

suppletive insertion 4.447–448
suppression 2.300–301
supralaryngeal adjustment 4.666
suprasegmental 3.724
suprasegmental property 3.206
Sūq al-ʾAhwāz 2.571
Šuqayr, Naʿīm 4.240
suqūṭ 2.16
Ṣūr Arabic 3.479, 3.481, 3.489
Sura 2.526
ṣūra ʿaqliyya 2.444
Surat 2.326, 2.329
Suravaram, Pratap Reddy 4.452
surface order 3.232
surface structure 2.208, 4.521, 4.524–525
Surinam 2.472
Surūr, Najīb 1.601
Suryanis 4.580
Sūsa Arabic 1.571, 1.642, 4.62
Sutcliffe, Edmund F. 3.159
Šuʿūbiyya 1.626, 2.303, 3.30, 3.34, 3.346, 4.295
šuʿūr 3.105
Suwāriqiyya 1.129
Suwaydi, ʿIzz ad-Dīn ʾAbū ʾIsḥāq ibn Muḥammad ibn Tarxān as- 1.536
Suwaylam, Ahmad 4.92
Suxne → Soukhne
Suyūṭī, Jalāl al-Dīn as- 1.194, 1.380, 1.567, 1.614, 1.626–627, 2.84–89, 2.92, 2.156, 2.158, 2.176, 2.189, 2.291, 2.298, 2.304, 2.316, 2.401, 2.425, 2.433, 2.496–497, 2.555, 2.628, 2.630, 3.7, 3.33, 3.90, 3.93, 3.101, 3.321–324, 3.380, 3.648–649, 3.651, 4.21, 4.283, 4.295, 4.436–437
Swadesh, Morris 4.313
Swahili 1.133–134, 1.136, 1.448, 1.545, 1.660–666, 2.26, 2.211, 2.525, 2.597, 2.669–670, 2.673, 3.59, 3.126–127, 3.635, 3.640, 3.643, 4.281, 4.381–387, 4.728
Swahili in Oman 3.478
Swahilization 1.663
Swain, Michael 1.368
Swales, John 1.649
Swan 2.67
Swan, Toril 1.476
Ṣwayliḥ 2.498
swear word 4.418–421
Sweden 2.72
Sweden, Arabic in 2.73, 2.97
Sweetser, Eva 4.162–163
Swērkiy Arabic 4.677
switch, single-word 2.374
switching 1.415
switching, diglossic 1.634–635
switching, diglossic → code-switching
switching, register 1.634
Switzerland 2.439
Šxauri 4.301
Sykes-Picot Agreement 2.699
syllabary, Sumero-Akkadian 4.1
syllabic cluster 3.209
syllabic economy 1.63
syllabic harmony 3.492
syllabic weight 3.726
syllabication 1.225, 2.35–36, 3.493, 3.530, 3.611, 3.617, 3.729, 4.344
syllabication, directional 4.85
syllabification → syllabication
syllable 3.208, 3.547, 3.610, 4.387
syllable, bimoraic 2.154
syllable, canonical 1.342
syllable, closed 4.388, 4.683
syllable, degenerate 2.36, 4.389
syllable, heavy 3.613, 4.388, 4.390
syllable, lexically stressed 2.396, 2.399–400
syllable, light 4.388, 4.390
syllable, monomoraic 2.154
syllable, open 4.388, 4.390
syllable, pitch-accented 2.396
syllable, superheavy 2.396, 3.612–613, 3.616–621, 4.108, 4.344, 4.388
syllable, ‘vowelless’ 3.611
syllable complexity 3.725
syllable juncture 3.612
syllable margin 1.487
syllable onset 3.728
syllable structure 2.32, 2.36, 2.350, 3.240, 3.301, 3.328–329, 3.601, 3.615–616, 4.116–117, 4.387–389, 4.564
syllable type 1.261, 3.611

syllable weight 3.612–613
syllepsis 3.538
syllogism 4.14
symbolism 4.212
symbolism, semantic 4.287
synchronic 1.604
synchrony 1.575
syncope 2.17, 2.36, 2.123, 3.494–495, 3.530, 3.729, 4.116, 4.344–345, 4.348, 4.351, 4.389–391
syncretism 1.644
syndetic 4.191
synecdoche 3.121, 3.632, 3.733
synesthesis 4.282, 4.286
synonym 1.435, 1.656, 2.69, 2.442, 3.35, 3.38
synonym → *mutarādif*
synonym, rhyming 4.607, 4.611
synonymy 1.435, 2.624, 3.45, 3.161–162, 3.320–321, 3.323–324
syntactic function 2.625
syntactic reconstruction 3.385
syntactic transfer 4.38
syntax 4.391–402
synthetic 1.400, 1.486, 1.562, 1.614, 1.616, 3.671
synthetic construction 2.247
synthetic genitive 1.82–85, 2.570
Syria 1.531, 2.500, 2.699, 3.315, 3.347, 4.402–409
Syria, Arabic in 1.407, 1.502, 1.571, 1.583–584, 1.592, 1.605, 1.607–608, 2.31, 2.204, 2.374, 2.532, 2.555, 3.588
Syriac 1.132, 1.166–167, 1.171, 1.180–181, 1.284, 1.493, 1.540–541, 1.602, 2.66, 2.182, 2.199–201, 2.211, 2.266, 2.639, 2.658, 3.218, 3.313, 3.351, 3.409–412, 3.414–415, 3.418–420, 3.426, 3.592, 3.696, 4.1, 4.73 4.168, 4.175, 4.177, 4.303, 4.402, 4.459, 4.542–544
Syriac, West 2.67
Syriac script 1.148–151, 1.603, 3.110
Syriacism 4.546
Syrian Arab Republic 1.546
Syrian Arabic 1.187, 1.221, 1.281, 1.340–342, 1.408, 1.492, 1.538, 1.540, 1.571, 1.604, 1.657, 2.38, 2.96, 2.156, 2.388, 2.390, 2.397, 2.499, 2.658, 3.67–68, 3.70, 3.237, 3.264, 3.266–268, 3.446, 3.456, 3.553, 3.556–557, 3.568, 3.611–612, 3.614, 3.656, 3.684–685, 3.729, 3.739–740, 4.45, 4.52, 4.67, 4.116, 4.163, 4.195–196, 4.309, 4.312–313, 4.317–318, 4.339, 4.341, 4.491, 4.647, 4.649, 4.654–655, 4.678
Syrian Desert 1.260
Syrian Plates of Kottayam 3.129
Syrian Protestant College 1.175
Syro-Lebanese Arabic 1.546, 3.1–2, 3.4, 3.585, 3.587
Syro-Mesopotamian Arabic 1.260
Syro-Palestinian Arabic 1.84, 1.111, 1.388, 1.642, 3.237, 4.285–286
Syrt region 3.53
system 1.369
system, one-or-two-systems debate 1.369
Systemic Grammar 1.457
Szabolsci, Anna 3.653

T

tāʾ 1.297, 1.609, 2.521, 2.594, 3.152, 3.281, 3.285, 3.292
tāʾ at-taʾnīṯ 2.156, 2.544
tāʾ at-taṯniya 2.239
tāʾ maftūḥa 2.327
tāʾ marbūṭa 1.78, 1.134, 1.136, 1.152, 1.384, 1.497, 1.561, 1.612, 1.639–640, 1.643, 2.156, 2.327, 2.342–344, 2.484, 2.527, 2.583, 2.606, 3.67, 3.92, 3.111, 3.220, 3.240, 3.242, 3.426, 3.430, 3.439, 3.445–446, 3.566, 3.590, 4.518
tāʾ ṭawīla 1.134, 1.384
Taʾabbaṭa Šarran 4.202
taʿaddā 3.102, 4.410
taʿaddin 1.356, 2.94, 3.101, 3.108, 3.224, 4.236, 4.410–416, 4.531, 4.624, 4.641
taʿajjub 2.79–80, 3.101, 3.550
taʿajjub, ʿalāmat at- 3.741
Ṯaʿālibī, ʾAbū Manṣūr ʿAbd al-Malik ibn Muḥammad aṯ- 1.173, 1.435, 2.443, 2.555, 2.582, 3.34, 3.117, 3.731, 4.283–284, 4.286
taʿalluq 4.438–439, 4.441
Ṭaʿʿān, Hāšim aṭ- 1.614, 1.616
taʿarrin 2.291
ṭabʿ 2.176
tabalbul al-luġāt 3.345
Ṭabarī, ʾAbū Jaʿfar Muḥammad ibn Jarīr aṭ- 2.334, 3.6–7, 3.36, 3.93, 3.574, 4.295
Tabbaa, Yasser 4.131
Tabet, Jacques 3.348
tābiʿ 1.124, 3.101, 3.201, 4.221
Ṯābit ibn Qurra 4.542
tabʿiyya 4.221
Tablig Jamayat 1.286
taboo 1.657, 1.659, 2.363–364, 2.646, 3.662, 4.166, 4.416–421, 4.471, 4.765
taboo, speaking 2.714
taboo expression 2.71
Tabriz 2.220
Tabrīzī, Mīr ʿAlī 3.337
Tabrīzī, Mīrzā Jaʿfar 3.337
tabsīṭ 1.176, 1.531
Tābūk 3.401–402, 3.404
tabyīn 4.436
Tachelhit → Berber, Tachelhit
taḍāfur al-qarāʾin 4.22
taḍāmm 4.22, 4.27
taḍammun 3.120
tadāwul 1.429
tadāxul al-luġāt 3.321, 4.167
taḍʿīf 1.358, 3.566, 4.642
taʿḏ̣īm 1.558, 4.329
taʿdiya 1.356, 4.410, 4.641–642
Tadjoura 1.654–655
taḏkīr 3.92, 651, 4.443
taḍmīn 1.628, 3.646
Taeschner, Traute 1.368
Taez → Taʿizz
tafaʿʿala 3.226
tafāʿul 1.628
tafḍīl 1.69, 2.13, 2.79, 3.105
tafʿila 3.208–209, 4.89–92, 4.103
Tafilalt 3.53
Tafilalt Arabic 3.290–291
tafjīr al-luġa 4.211
TAFL → Teaching of Arabic as a Foreign Language

tafrīd 3.430, 4.316
tafsīr 1.336, 1.427, 1.512, 1.558, 2.383, 3.136, 4.436, 4.472, 4.685, 4.740
Taftazānī, Saʿd ad-Dīn at- 1.380
tafxīm 1.205, 1.558, 2.233, 2.235, 2.503, 2.529, 3.91, 3.149, 3.600, 4.10, 4.218, 4.421–424, 4.427, 4.669
tafxīm, ʾalif at- 3.90
tag 1.342, 2.148, 2.615
tag, discourse 1.514
Tagant 3.171
Tägdāwəst 3.170
tagger 1.460
tagger, parts-of-speech 1.514–515
tagging 1.455, 1.459, 1.513, 1.515
Taġlib 2.121, 3.94
Tagore, Rabindranath 1.288
taġyīr 2.403, 3.649, 4.121
taġyīr al-ʾiʿrāb 3.650
Ṭāhā, ʿAlī Maḥmūd 4.90
Taha, Zeinab Ahmed 3.100–106, 4.410–416, 4.235–236
tahaddud 4.329
taḥḍīḍ 3.573, 4.331, 4.740
taḥdīṯ 1.531
taḥliya 4.221
taḥmīd 2.327
taḥqīq 3.92, 3.119, 3.307, 4.446
taḥqīq al-hamza 4.6, 4.10
taḥqīr 1.637
taḥrīf 2.630
Ṭahṭāwī, Rifāʿa Rāfiʿ aṭ- 1.494, 3.193, 3.349
taḥwīl 3.92, 3.235
Ṭāʾif 3.179
Theme and Tail constituents 2.148–149, 2.397, 4.732
Taine-Cheikh, Catherine 1.419, 1.585, 2.240–250, 2.363, 3.60, 3.169–176, 3.447–453, 3.664, 3.667
Taʿizz 1.655, 3694
Taʿizz Arabic 4.751
Tāj al-ʿarūs 3.21, 3.30–33, 4.2, 4.35, 4.103
Taj Mahal 4.561
Tāj as-salāṭīn 1.8
tajāḏaba 2.580
tajānus 2.496
tajawwuz 3.118
Tajfel, Henri 2.713
Tajik 1.29, 1.34, 1.590, 2.408, 2.670, 3.578, 4.424–425, 4.608, 4.612–614, 4.621
Tajik, numerals in 4.617
Tajik suffix in Uzbekistan Arabic 4.616
tajmiyt 2.470
tajnīs → *jinās*
tajwīd 1.338, 1.528, 3.207, 3.335, 3.567, 3.726, 4.422, 4.425–428
Takarlī 4.663
taʾkīd 1.202, 1.555, 1.559, 4.221, 4.640
takrār 1.471
Takrouna Arabic 3.448
ṭalʿa 3.173
Ṯaʿlab, ʾAbū l-ʿAbbās ʾAḥmad ibn Yaḥyā 1.627, 2.79, 2.86–89, 2.234, 2.359–360, 2.443, 2.496, 2.629–631, 3.35, 3.107, 3.117, 3.323–324, 4.330–332, 4.443
ṭalabī 2.358, 2.360
talāqin 4.413
talaṭṭuf 2.70
talāʾum 1.429
Talay, Shabo 4.407
talāzum 1.435
Taleb Ibrahimi, Khaoula 2.660
ṭālib al-luġa 2.630
talik 3.338
taʿlīl 1.194, 2.177, 4.12–13, 4.475
taʿlīq 3.336–338, 4.21–22, 4.98, 4.131, 4.428–430, 4.561
Tāliqānī, ʿAbd al-Majīd 3.338
taʿlıq-qirması 3.338
Talmon, Rafael 1.309, 1.562, 1.583–584, 2.176–178, 2.226–227, 2.234, 2.294, 2.307, 2.402, 2.435–436, 2.453, 2.536, 2.539, 2.542, 2.544, 2.632, 3.21, 3.106–107, 3.109, 3.123, 3.176, 3.315, 3.361, 3.456, 3.547, 4.220, 4.235, 4.284, 4.431, 4.446, 4.661
Talmoudi, Fathi 1.571, 3.729, 4.62, 4.563
Talmud 1.131, 2.527, 3.409

Talmy, Leonard 1.485
Talo, Ewa S. 4.271
taltala 1.399, 1.490–491, 1.493, 1.594, 1.614, 2.263, 3.14, 3.691–692, 4.430–431
talṭīf 2.70
talyīn al-hamza 3.91
tamakkana 4.431
tamakkun 3.453, 4.431–433, 4.445
tamām al-ism 4.445
Tamari, Tal 3.137–139
Tamasheq 4.280
tamāsuk 1.427
Tamazight 1.534, 2.73, 2.698, 2.704, 2.709, 2.712, 3.274, 3.350
Tamerlane 1.28, 4.612
Tamezret 4.572
Tamil 2.333, 2.669, 3.725, 4.433–436
Tamil, Kanniyakumari 4.433, 4.435
Tamil Nadu 2.325
Tamīm 1.399, 1.490, 1.614, 1.616, 1.641, 2.88, 2.263, 2.312, 2.427, 2.494, 2.555, 3.14, 3.89, 93.2–94, 3.186, 3.497, 3.645, 3.691, 3.697–698, 4.10, 4.22, 4.304, 4.430, 4.572
Tamīmi Arabic 1.71
Tamīx 3.94
tamkīn 3.453
tāmm 1.624, 2.14, 2.548–551
ṯamma 1.566, 3.200
tamṣīr al-luġa 3.349
tamtām 2.676–677
tamṯīl 2.301, 2.305, 2.403, 2.443–444, 2.446, 3.118, 4.447
tamṯīliyya 4.157
Ṯamūd 4.478
Ṯamūdic → Thamudic
tamyīz 2.15, 2.186, 2.225, 3.101, 3.103–105, 3.241, 3.432, 3.451–452, 3.548, 3.684, 4.436–437, 4.565
tamyīz al-jumla 4.437
tamyīz al-mufrad 4.437
tamyīz an-nisba 4.437
ṯanāʾ 3.173
Tana River 4.381
tanāzuʿ 4.193, 4.437–442
tanbīh 2.237
tandīm 4.331
Tandjile 1.360
Tang Dynasty 1.378
Tanganyika 1.660, 1.663
Tangier 2.64, 3.287, 3.294, 4.156
Tangier Arabic 1.439, 1.609, 3.274, 3.288, 3.291–293, 3.672
taʾnīṯ 2.156, 3.92, 3.651
tankīr 1.188, 3.453, 4.443, 4.445
Tannen, Deborah 3.665
Tannūmah Arabic 4.127
Tanūxī, ʿIzz ad-Dīn at- 2.637, 3.216
tanwīn 1.68, 1.73, 1.109, 1.134, 1.152, 1.188–189, 1.207, 1.246, 1.248, 1.266, 1.273, 1.309, 1.559, 1.562, 1.580, 1.606, 1.615, 1.644–646, 2.20, 2.24, 2.87, 2.226, 2.239, 2.278, 2.353, 2.379, 2.403–404, 2.411, 2.424–427, 2.453, 2.546–547, 3.98, 3.191, 3.220, 3.270, 3.335, 3.403, 3.429, 3.440, 3.453, 3.482, 3.548, 3.551, 3.590, 3.648, 3.693, 3.696, 4.15–16, 4.124, 4.318, 4.357, 4.442–446, 4.558, 4.742
tanwīn al-ʿiwaḍ 4.444–445
tanwīn al-makāna 4.443
tanwīn al-muqābala 4.444
tanwīn at-tamakkun 4.443, 4.445
tanwīn at-tankīr 4.443
tanwīn at-tarannum 4.443
tanwīn fatḥa 2.327
tanwīn-naṣb construction 2.453
Tanzania 1.661–663, 3.642, 4.386, 4.508
Tanzania, Arabic in 2.672
tanzīl 1.173
Tanzimat 3.504
Taoujjout 4.572
Ṯaqafī, ʿĪsā ibn ʿUmar aṯ- → ʿĪsā ibn ʿUmar aṯ-Ṯaqafī
Ṯaqafī, Muḥammad ibn al-Qāsim aṯ- 2.325
taqdīm 2.131, 3.649–650

taqdīm wa-taʾxīr 4.739
taqdīm, wujūb at- 4.739
taqdīr 2.290, 2.301–302, 2.402, 2.404, 2.537–538, 2.626, 4.13, 4.446–449
taqdīrī 1.561
Ṯaqīf 3.691
ṯaqīl 2.233, 3.551, 4.431–432
Taqla, Bušārā 3.193
Taqla, Salīm 3.193
taqlīd 2.305
taqwīʿ 3.341
Taṛābīn 3.360, 3.401
Taṛābīn Arabic 3.365–366, 4.238
tarāduf 3.323
Ṭarafa ibn al-ʿAbd 3.211–212, 3.646, 4.201, 4.208–209
tarannum 3.453, 4.443, 4.445
ṭardiyya 4.32, 4.35
target 1.425, 2.192–193
target language 1.437, 1.459, 1.482, 1.485, 3.639, 4.139, 4.370
target text 1.432
targum 3.409
taʿrīb 2.638, 2.641, 2.696, 4.465
taʿrīḍ 2.578–579, 2.581–582
taʿrīf 3.428
taʿrīf, ʾadāt at- 1.488
Tarifit → Berber, Rif
ṭarīqa al-ḥamdiyya, at- 3.341
tarjama 4.157
tarkīb ismī 3.428
tarkīb ʾisnādī 2.404
Tarone, Elaine 4.141
tarqīm 3.740
tarṣīʿ 4.103–104
Tarsus 1.388
Tarsus Arabic 1.390–392, 1.394–395
tartīb 1.575
tartīl 4.426
tarwīs 1.152, 3.308, 3.341, 3.650, 4.134, 4.560
tarxīm 1.638, 2.17, 4.558
Tāsāftī 3.222
taṣarruf 2.179, 2.356, 2.426, 3.658, 4.432, 4.443
taṣawwur 1.429
tasbīb an-naʿt 3.432
tašbīh 2.329, 2.425, 2.441, 3.116, 3.119, 3.122, 3.321
tašbīhī 2.444
tašdīd 1.2, 1.152, 2.299
taṣġīr 1.576, 1.637, 2.424
taṣḥīf 2.630
tashīl al-hamz 4.10
tašʿīra 4.560–561
Tasiwit → Berber, Siwa
tasjīʿ 4.103–104
taškīl 1.487
taskīn 1.615–616, 2.16
Ṭāšköprüzāde 3.502–503
tasliya 2.46
taṣrīʿ 4.87, 4.92
taṣrīʿ jāmiʿ ʾabyātihi 4.87
taṣrīf 1.119, 1.122, 1.192, 1.573, 2.180, 2.298, 2.402, 2.426, 2.623–624, 3.423, 4.119, 4.431
taṣrīf al-ʾafʿāl 2.404
taṣrīḥ 2.578–579, 2.581, 3.120
Tata Arabic 3.288
Tataouine 2.709
Tatar 2.557, 2.670, 2.673, 4.263–264, 4.450–451
Tatar, Crimean 4.263
Tatar, Kazan 4.450
Tatar, Volga 4.450
tatimma 4.221
taʾṯīr 3.114
taṯniya 2.184
Tauzin, Aline 3.173
Tavla 1.116
tawāḍuʿ 4.684
tawahhum 4.447
tawallud 4.330–331
Ṭawāra Arabic 3.401, 3.403
tawārud 1.436, 1.438
tawassuʿ 3.116, 3.120
tawassuʿāt 3.651
tawātur 4.13, 4.684
Taweel, Ahmad ibn Ahmad Muhammad Abdalla [Ṭawīl, ʾAḥmad ibn ʾAḥmad Muḥammad ʿAbdallāh aṭ-] 3.336
Tawfīq, Muḥammad 1.601
tawḥīd 1.336, 1.338, 3.136

Tawḥīdī, ʾAbū Ḥayyān at- 1.403, 3.112, 3.161, 3.177, 3.307, 3.309, 4.64, 4.66, 4.132
tawhīm 3.539
ṭawīl 1.429, 2.579, 3.119, 3.211, 3.214
tawjīh 3.539
tawkīd 1.124, 1.202, 1.435–436, 1.569, 2.22, 2.82, 2.115, 3.573, 4.562
tawqīʿ 1.152, 3.336, 3.519, 4.131, 4.561
tawqīʿ al-muṭlaq, at- 4.561
tawqīf 1.3, 4.684
tawriya 2.497, 2.578, 3.538–539, 4.157
tawriya mubayyana 3.538
tawriya muhayyaʾa 3.538
tawriya mujarrada 3.538
tawriya muqtarina 3.538
tawriya muraššaḥa 3.538
Tawwazī, at- 1.627
taxalluṣ 3.717, 3.720
taxāṭub 3.122
taxfīf 2.17
taʾxīr 3.649–650
taxnīn 3.336
taxrīj 1.628
taxṣīṣ 1.190, 3.428, 4.316
Tayāha Arabic 4.238
Tayash, Fahad A. 3.397
taʿyīn 1.190
Taylor, John R. 1.343–344, 4.166, 4.502
Tayma 4.478–479, 4.481
Tayman 3.411
Taymanite 4.478–481
Taymūr, Maḥmūd 1.409, 1.601
Taymūr, Muḥammad 1.600, 1.602
taysīr 1.176
ṭayyāyē 1.128
Ṭayybit ilImām Arabic 4.406–407
Ṭayyiʾ 1.187, 1.310, 1.399, 1.628, 2.494, 3.89, 3.94, 3.692, 3.694–695, 3.697, 4.62, 4.126, 4.376, 4.430
Ṭayyib Ṣāliḥ aṭ- → Ṣāliḥ, aṭ-Ṭayyib
Ṭayyibī, aṭ- 3.121, 4.132, 4.562
Taza Arabic 1.609, 3.288, 3.290–291
Tchardari 4.613
Tchédé 1.334
TD-PSOLA 1.210, 1.212
teacher training 2.110
Teaching Arabic as a Foreign Language (TAFL) 4.154
teaching, mosque 1.381
teaching Arabic writing 1.651
teaching English writing 1.651
teaching materials 1.670, 2.74, 4.148, 4.152–153, 155
teaching of Amazigh 2.712
teaching of Arabic 1.176, 1.258, 1.336–338, 1.364–365, 1.370, 1.449–451, 1.472, 1.484, 1.654, 1.660–661, 1.667, 1.669, 2.75, 2.104–113, 2.184, 2.214, 2.337–338, 2.438, 2.672, 2.698, 2.700
teaching of Arabic in Europe 2.73, 3.315
teaching of Arabic in India 2.326–327
teaching of Arabic in Israel 2.440, 3.315
teaching of Arabic in Jordan 2.503
teaching of Arabic in Mali 3.137–138
teaching of Arabic in Malta 3.143–144
teaching of Arabic in Nigeria 3.374, 3.376
teaching of Arabic in North America 3.395
teaching of Arabic in Pakistan 3.508, 3.511
teaching of Arabic in Senegal 4.185
teaching of Arabic in Somalia 4.276
teaching of Arabic in the Ottoman empire 3.504
teaching of Arabic in the United States 3.398
teaching of Arabic in Turkey 4.578, 4.582
teaching of Berber 2.63
teaching of English 2.8–9, 2.31, 2.216, 2.696, 3.316
teaching of foreign languages 2.696, 2.703, 3.313, 3.315

teaching of French 2.8–9, 2.698
teaching of French in Jordan 2.504
teaching of French in Morocco 2.122
teaching of German 2.9
teaching of grammar 2.701
teaching of Hebrew 3.398
teaching of Italian 2.9
teaching of Ivrit 2.464
teaching of Moroccan Arabic 2.74
teaching of oral communication 2.107
teaching of reading 2.107
teaching of reading and writing 3.75
teaching of Russian 2.9
teaching of second language → second language teaching
teaching of Spanish 2.9
teaching of writing 2.107
təbrāᶜ 3.173
Teda 1.36, 1.360
Teeple, David 4.85–86, 4.389–391
téere 4.181
Teeuw, Andries 2.337, 2.341
Teixera, Pedro 2.572
Teixidor, Javier 3.472
Tekna 2.240
teknonym 2.586, 2.588, 4.469
Tekrur 4.179
Telegdi, Zsigmund 2.407, 2.411
telegraphese 4.47–49
telephone conversation 1.214
television 1.273, 1.275, 1.408, 1.528, 1.546, 1.655, 2.32, 2.214, 2.503, 2.609, 2.648, 2.658, 3.11, 3.19, 3.194, 3.196–198, 3.295, 3.318, 4.408
television, multichannel 3.199
television, satellite 3.197, 4.322
television serial 4.377
telic 1.198, 1.200
telic verb 1.52
telicity 1.52
Tell al-Masxūṭa 3.466
Telugu 2.669, 4.452–454
Telugu, administrative terminology in 4.452
Telugu, gender system in 4.454
template 1.573, 2.447–450, 3.225, 3.301, 3.610, 4.93–95
template satisfaction 3.492
temporal 1.472, 3.234
temporal adverb 1.21–23
temporal terms, acquisition of 1.377
temporality 2.76
Temporary Impairment Hypothesis 4.145
Ten Great Books 1.380
Tenes 1.54
Tenes Arabic 1.609
tenor 2.442
tense 1.110, 1.195–196, 1.216–217, 1.348–350, 1.508, 1.510, 1.523, 1.568, 1.606, 2.90–92, 2.174, 3.67, 3.98–99, 3.247–248, 4.306, 4.392, 4.395, 4.398–399, 4.454–458, 4.745
tense, Greek 3.97
tense auxiliary 4.646
tense maintenance 1.432
tense marker 1.485, 1.567, 2.568, 2.577
tense mood 1.220
Tense Projection 4.396, 4.398–399
tense shift 1.432
Tense-Mood-Aspect system 1.486, 2.523–524, 3.67, 4.455, 4.457
term, general 4.685
term, specific 4.685
terminative 3.414
terminology 1.402, 2.634, 2.637, 3.18, 4.458–466
terminology, administrative 2.199, 2.636, 4.590, 4.593
terminology, administrative, in Telugu 4.452
terminology, agricultural 2.284, 2.289, 2.667, 4.312
terminology, alchemical 2.27–28
terminology, anatomical 2.26
terminology, astronomical 2.26, 2.667
terminology, bicycle 2.637
terminology, botanical 2.26, 2.200
terminology, building 4.312
terminology, car 2.34

terminology, Christian 1.181, 1.503
terminology, coining of 2.635–640, 4.168
terminology, commercial 2.458
terminology, computer 2.9, 2.30, 2.34, 2.375, 2.667
terminology, Coptic scientific 1.499
terminology, creation of 1.175
terminology, culinary 2.29
terminology, dictionaries of 2.641
terminology, education 2.341–342
terminology, grammatical 1.168, 2.186
terminology, Indo-European 2.539
terminology, irrigation 1.504
terminology, Islamic 2.29, 4.78
terminology, Islamic jurisprudence 2.341
terminology, Jewish 1.240
terminology, kinship 2.583–588
terminology, legal 3.26, 3.32
terminology, mathematical 2.28, 2.667
terminology, medical 4.460, 4.545
terminology, military 3.6, 4.593
terminology, musical 2.28
terminology, mystical 4.76
terminology, pharmacological 2.200
terminology, philosophical 2.200, 4.545
terminology, religious 1.181, 1.257, 1.372, 1.443, 1.541, 2.140, 2.253, 2.342, 2.667, 4.168, 4.601
terminology, rural 1.503–504, 1.541
terminology, scientific 1.403–404, 2.254, 4.459–460, 4.545
terminology, technical 1.176, 1.332, 1.437, 1.453, 1.531, 1.665, 2.9, 2.30, 2.374–376, 2.468, 2.700, 2.702, 3.25–26, 3.203, 3.241, 3.426
terminology, trade 2.253, 2.455, 2.667
terms of abuse 4.470–471
terms of address 1.251, 2.8, 2.205, 2.622, 3.719, 4.466–471
terms of address, Ottoman 4.593
terms of address, Turkish 3.660–661, 4.467, 4.470
terms of endearment 4.470
terms of respect 4.470
Territoire français des Afars et des Issas 1.654
territory 2.60
Tersis, Nicole 4.279
Tesnière, Lucien 1.41, 1.343, 2.317
Testen, David 1.201–204, 2.21–23, 2.66, 3.191, 3.419
Tetouan 2.64, 3.294, 4.156
Tetouan Arabic 1.596, 1.609, 3.273, 3.288, 3.290–291, 3.293
Teubner, J.K. 3.7
text 1.430–432, 1.472, 1.568, 1.648, 1.651, 2.143, 2.148, 2.207
text, expository 1.474
text categorization 1.459
text corpus 1.214
text database 1.460
Text Encoding Initiative 1.513
text feature 4.472, 4.474–475
text linguistics 1.430, 1.470
text organization 2.207
text processing 1.455, 1.459, 1.515
text production 4.474
text reception 4.474
text type 1.432, 4.473
text typology 4.473
textbook 1.448–449, 1.651, 1.655, 1.670, 2.110–111, 2.189, 2.336–338, 2.704
textiles 2.40
textlinguistics 4.471–476
text-to-speech conversion 1.460
text-to-speech converter 1.515
text-to-speech synthesis 1.207
text-to-speech synthesizer 1.210
text-to-speech system 1.211–212
textual theme 1.475
textuality 1.430, 4.472
textus receptus 4.4
Thackston, Wheeler M. 2.598, 4.148
Thai 4.268, 4.476–477
Thai, Arabic loanwords in 4.476

Thai Muslims 4.477
Thakerar, Jitendra N. 4.322
Tham Seong Chee 2.341, 2.344
Thamudic 2.198, 2.263, 4.477–483
Thamudic, Hijazi 4.478–479, 4.481, 4.483
Thamudic, Hismaic 4.478
Thamudic, Najdi 4.478–479, 4.481–482
Thamudic, Tabuki 4.480, 4.482
Thamudic, Taymanite 4.479–481
thanks 3.660
that-trace violation 3.707
thäydīn 3.173
tḫāži 3.174
theater 1.175, 1.600–601, 2.77
theater, colloquial 1.600
theater play 1.323
Theil, Rolf 2.137–142
Thelander, Mats 2.320
Thelin, Nils 2.208
Thelwall, Robin 1.525, 3.62, 3.435
thematic continuity 1.648
thematic development 1.427
thematic hierarchy 4.489
thematic pro 2.174
thematic relation 4.535
thematic role 1.40–41, 4.535
thematic shift 1.648
thematic slot 1.474–475
thematic structure 4.484–486, 4.501–502
thematic unity 3.646
thematic vowel 1.119, 1.121
thematization 1.623, 2.249, 4.485–486
thematized structure 4.486
thematizer 4.486
theme 1.347, 1.429, 1.474, 1.485, 2.148–149, 2.569, 3.81, 4.484–487, 4.489–494, 4.498, 4.501–502, 4.535, 4.538, 4.732
theme, textual 1.475
theme, topical 1.475
theme marker 2.149
Theophanes 4.259
theophoric names 3.696
thésis 3.176, 3.209, 4.684
Θ criterion 4.488–489, 4.491–492
Θ marking 4.491, 4.493
Θ position 4.488
Θ role 4.487–494
Θ Theory [Theta Theory] 2.170, 4.438, 4.441, 4.488, 4.746
Thillet, Pierre 4.546
third deixis 1.106
third distance 1.23
Thirteen Classics 1.380
Thomas, Bertram 2.211, 3.478, 4.124
Thomas, Jenny 1.657
Thomas, Saint 3.129
Thomason, Sarah Grey 1.294, 1.501, 1.504, 2.29, 2.664–674, 3.59, 3.642, 4.368, 4.370–373
Thompson, Billy Bussell 1.98
Thompson, Sandra A. 2.208, 4.317
Thōr Arabic 4.127
Thorbecke, Heinrich 2.629
Thorburn, D.H. 2.518
Thráinsson, Höskuldur 4.540
Thurston, Edgar 4.435
Tiaret 1.56
Tiberias Arabic 1.608
Tibesti 1.360, 1.362
Ṭībī, aṭ- → Ṭayyibī, aṭ-
Tibrīzī, al-Xaṭīb at- 2.538
tier, autosegmental 3.626
tier, skeletal 3.626
tier conflation 4.93–94, 4.96
Tietz, Renate 2.477, 2.479
Tietze, Andreas 2.456, 4.583–584, 4.588
Tifinagh 1.133, 1.294, 2.63, 2.673, 2.709, 3.171, 3.294, 3.317, 3.350–351
Tiglath Pileser III, King 1.126, 3.466
Tigray → Tigriñya
Tigré 1.639, 2.66–67, 2.269, 2.271, 3.592, 4.175, 4.301, 4.304–306, 4.309, 4.312
Tigriñya 1.654–655, 2.53–54, 2.66–67, 2.269, 3.62, 3.627, 4.301–302, 4.304–306, 4.309–310, 4.312

Tigris Arabic 1.91, 1607
Tīh, at- 3.400, 4.237
Tihāma 1.654–655, 2.274, 3.691, 3.695, 3.697, 4.125, 4.127, 4.751
Tihāma Arabic 1.589, 1.655, 2.259, 2.270–273, 3.179, 3.454, 4.240, 4.751, 4.753–754, 4.756
Tijāniyya 4.180–182
Tikrīt Arabic 2.415, 2.417, 2.421
Tikrītī, Bahjat K. at- 1.617
tilde 1.1
Till, Walter C. 1.502
Tilmatine, Mohand 2.73
Timberlake, Alan 4.37, 4.39–40
Timbuktī, at- 3.373
Timbuktu 2.240, 3.135–136, 3.138, 3.394
time 2.91–92
time reference 1.195–196, 1.553, 4.48
Tindouf 2.240
Tintin 4.284
Tio 2.273
tip of the tongue 2.683
Tipografia Medicea 1.168
ṯiqal 1.191, 2.425, 4.431
ṭirāz 2.41, 2.600
Tirmidh 2.600
Tirol 2.439
Tirosh-Becker, Ofra 2.534
Tišīt 3.170
title 1.125, 3.658–659, 3.720, 4.467, 4.468–470
title, Ottoman 4.592
Tivaouane 4.182
Tiyāha 3.360
Tiyaha Confederation 3.401
Tizi Ouzou 2.712, 2.715
Tkatsch, Jaroslaus 4.543–546
Tlemcen 1.24, 1.53–54, 1.56
Tlemcen Arabic 1.444, 1.609, 2.532, 2.534, 3.452, 4.70
Tlili, Bechir 4.572
TMA → Tense Mood Aspect
tmaziġt → Tamazight
Tobi, Y. 2.534
Toblig Jamat 1.286
Tobrouk Arabic 3.55
Todorov, Tzvetan 2.468
Togo 2.250, 4.758
tokenizer 1.515
ṭolba 4.158–159
Toledo 1.96–97, 1.100, 1.166–167, 1.169, 1.171, 4.217
Tombalbaye, President François 1.363
Tombouctou → Timbuktu
tombstone 2.46
Tomiche, Nada 4.669
tone 2.139, 2.253, 2.590, 2.593, 2.671, 3.461
tone, edge 2.395–398
tone, high 2.560
tone, level 2.397
tone, lexical 1.521, 2.400
tone language 3.437
tone of voice, abnormal 2.675
Tongan 4.416
tongue click 4.233
tongues, confusion of 3.345
tonic accent 1.367
Tonk 2.329
Topaloğlu, Ahmed 3.501
Toparlı, Recep 4.587
topic 1.45, 1.50, 1.69–71, 1.73–74, 1.183–184, 1.350, 1.354, 1.431, 1.436, 1.485–486, 1.488, 1.507, 1.558, 1.651, 2.113–115, 2.132, 2.145–147, 2.149, 2.291–292, 2.308, 2.321, 2.354–357, 2.434, 2.436–437, 2. 537–539, 3.105, 3.231, 3.381, 3.383, 3.386–387, 3.557, 3.560, 3.683, 4.319, 4.354, 4.366, 4.484–487, 4.494–502, 4.504, 4.527, 4.570, 4.734
topic → *mubtada'*
topic, contrastive 4.505
topic, emphatic 4.500
topic, grammatical 4.653
topic, sentence 4.503
topic accessibility 3.711
topic continuity 1.649
topic focus 4.496
topic introduction 1.432
topic maintenance 1.651
topic management 2.384
topic marker 1.649

topic position 4.186–187
topic reanalysis 1.185
topic shift 1.432, 1.651
topic shifting 1.432
Topic/Comment 4.42, 4.525–527
topical expression 2.113
topical structure 1.651
topicality 2.145, 4.319, 4.366, 4.505–506, 4.732–733
Topicality Hierarchy 4.506, 4.731–733
topicality, degree of 3.86
topicalization 1.183–184, 1.186, 1.349, 1.476, 1.553, 1.623, 2.114, 2.131, 2.145, 3.85, 3.383, 3.557, 3.589, 4.354–355, 4.395, 4.400, 4.486, 4.501–508, 4.522–525, 4.527, 4.654, 4.675, 4.722, 4.728–729, 4.731–733, 4.742, 4.746
topicalization marker 4.60
topicalizer 1.473
topichood 4.317
topic-prominent language 3.557
toponyms → place names
Tōppil Muhammatu Mīra̱n 4.433
Ṯōr alBāha Arabic 4.751
Torobe 3.170
Torodɓe 4.184
torsimanys 1.98
Tosco, Mauro 1.521, 2.518, 2.523, 3.61, 3.634, 3.639–644
Toshabim 2.64, 3.312
Touareg → Tuareg
Touba 4.180, 4.182
Toubou 3.56
Toulmin, Stephen 1.651
Touma, Pierre 3.23
Touratier, Christian 2.83
Touré, Saada 3.137–138
Touré, Saliou Almahady 1.256–257
tourism 2.9
Tourkou Arabic → Arabic, Turku
Tourneux, Henri 4.713, 4.715
Toygarlı 1.117
TP → phrase, Topic
trace 2.171–173, 4.80–81, 4.364, 4.395, 4.488, 4.493, 4.526
trace, clitic 1.350
Ṭṛād, Mīšāl (Michel) 1.603
trade 1.588–590, 2.325, 2.333, 2.552, 2.588, 2.671, 3.58, 3.76, 3.135, 3.137
trade, incense 4.295
trade, Indian Ocean 4.298, 4.381, 4.434–435, 4.477
trade, Levant/Mecca 3.698
trade, Mecca/Yemen 3.698
trade, Mediterranean 2.454, 4.573
trade, Southeast Asia 4.476
trade, Southeast India 4.433
trade, trans-Saharan 3.374, 4.723
trade jargon 2.215, 3.62
trade language 1.256, 2.51, 2.219, 2.268, 2.274, 3.642
trade language, Arabic as 1.655
trade routes 2.667, 3.698
trade routes, Arabian peninsula 3.697–698
trade settlement 1.617
trading camp 1.519
transcription 1.214, 1.514–515, 1.600, 4.515–520
transcription, ad hoc 4.515
transcription, Greek 1.642
transcription, Latin 1.603
transcription, scientific 4.515–516, 4.519
transcription in email 4.516
transcription of speech 1.210, 1.213
transducer 1.456
transfer 1.416, 1.459, 1.651, 4.41, 4.370–371
transfer, first language 4.143–145
transfer, semantic 4.461, 4.464–465
transference 2.442–444, 3.116, 3.121
transformation 4.521–522
transformation, Affix Hopping 4.522
transformation, bounded 4.522
transformation, cyclic 4.522
transformation, Equi 4.522
transformation, noncyclic 4.522

transformation, Passive 4.522
transformation, Pronominalization 4.525
transformation, Question 4.522
transformation, Reflexive 4.522
transformation, Relative 4.526
transformation, unbounded 4.522
transformational component 4.524, 4.741
Transformational Grammar 1.457, 1.562, 2.356, 4.448, 4.520–528, 4.657, 4.741
Transformational Linguistics → Transformational Grammar
transformational rule 4.522
transhumance 4.403
transient 1.26
transitional area 2.3, 2.5
transitive 1.328–329, 1.348, 1.356–357, 4.410–416, 4.536, 4.624, 4.641–642
transitive, complex 3.224
transitive, doubly 4.530, 4.533
transitive, singly 4.530
transitive, triply 4.530
transitivity 1.356, 2.94, 3.101–102, 3.108, 3.224, 3.251, 4.236, 4.410–416, 4.528–535, 4.624, 4.626
transitivization 3.224, 4.536, 4.626
Trans-Jordan 2.498–506
Trans-Jordanian Arabic 2.499
translation 1.432, 1.436–438, 1.531–532, 2.209, 3.201–202
translation from Arabic into Amharic 2.54
translation from Arabic into Geʿez 2.51
translation from Arabic into Latin 1.166
translation from Arabic into Malay 2.334
translation from Arabic into Oromo 2.54
translation from Arabic into Turkish 3.503
translation from Arabic into Urdu 3.510
translation from Brazilian into Arabic 3.4
translation from English into Arabic 2.332
translation from Greek into Arabic 1.166, 2.185, 2.200, 4.542–548
translation from Persian into Arabic 2.200
translation from Persian into Malay 2.334
translation from Syriac into Arabic 4.542
translation literature 4.542–548
translation movement 4.460
translation of Arabic grammar in Europe 2.189
translation of foreign works 2.699
translation of scientific texts 1.175
translation of the Bible 1.175, 1.383, 2.334, 4.759
translation of the Pentateuch 2.23, 2.528
translation, etymological 2.201
translation, interlinear 2.336
translation, morpheme-by-morpheme 2.666
translators 2.200–201, 2.209, 4.543
translators, Christian 4.546
transliteration 1.514, 2.375, 4.515
transliteration of foreign words in Arabic 4.465
transliteration, Greek of Arabic 4.218
transliteration, Greek of Arabic names 4.480
trans-locative 1.199
Transoxania 1.403, 4.612
transparency 2.37
Trāra 1.54
Trarza 3.171
Trask, R.L. 3.728
Traugott, Elizabeth Closs 2.191–196, 4.37–39, 4.41–42, 4.161, 4.163
Travancore 3.128
travelers 3.60

Travis, Lisa 4.746
tree, derivation 1.514
tree, metrical 3.616
tree diagram 4.521
Tree Pruning Hypothesis 2.691
tree structure 3.228, 230
t-reflexive 1.606
Treiman, Rebecca 1.339–340
Trengganu Stone 2.344
tribal affiliation 2.583
tribes, names of 1.17
triconsonantal → triradical
trilingualism 2.696
trilingualism, Arabic/Uzbek/Tajik 4.613
triliteral → triradical
Trimingham, J. Spencer 2.39, 2.560
triple-accusative construction 4.539
Tripoli 1.362, 1.589, 1.610, 2.552, 3.53, 4.548
Tripoli Arabic 2.532, 2.534, 3.53, 3.55, 3.607, 3.666–668, 4.548–556
Tripolitania 3.52–53,
Tripolitania Arabic 3.55–56
triptosis 1.398, 1.559–561, 1.580, 1.643
triradical 1.191, 3.624, 4.548
triradical root 3.66
triradicalism 1.80, 1.313–315, 3.45, 4.96
tritransitive 1.67, 4.539, 4.541
trochaic 3.613
trochaic pattern 2.350
trochee, bimoraic 4.347, 4.349
trochee, even 4.348
trochee, moraic 4.345, 4.348, 4.390
trochee, uneven 4.347–348
trope 2.444–445, 2.579, 3.116–118, 3.121–122
tropic meaning 2.580–582
Tropper, Josef 3.227, 3.408, 3.416–419, 4.478–479
Troupeau, Gérard 1.574, 2.232, 2.294, 2.298, 2.308, 2.436, 2.536, 2.542, 2.625, 3.72, 3.90, 4.119, 4.219, 4.235, 4.414, 4.431, 4.446
Trubetzkoy, Nikolaj S. 2.238
Trucial Oman 4.124
Trudgill, Peter 1.585, 1.613, 2.33, 2.507, 3.12, 4.628, 4.631
trujamanes 1.98
Trumpp, Ernest 2.477
truncation 1.1, 2.17, 2.165, 3.297, 3.299, 3.650, 4.202, 4.557–559, 4.724
truncation, vowel 4.37
truth condition 3.570, 4.15, 4.18
truth value 1.423, 3.652, 4.15
Truzzi, Oswaldo 3.2
Tschacher, Torsten 4.433–436
Tsereteli, George 4.613–615, 4.618
Tsereteli, Konstantin G. 4.303
Tsiapera, Maria 1.536
Tuareg 1.38, 1.133, 3.56, 3.592
Tubu 1.36, 2.554
Tucker, Archibald 4.381–383
Tucker, G. Richard 2.652, 2.656, 2.660
Tucuman 3.3
Ṭufayl, S. Muḥammad 4.426
ṭuġrā 3.308, 4.134
Tukrūr, Bilād 3.373
ṯulāṯī 2.14, 4.624
Ṭulba (Ibn at-), M. 3.172
ṯuluṯ 1.152, 3.112, 3.307–308, 3.340–341, 3.519, 4.131, 4.138, 4.560–562
ṯuluṯ kabīr 4.562
ṯuluṯ maġribī 3.112
ṯuluṯ rayḥān 1.151
ṭumār 1.152
ṯumma 1.467, 1.475
ṭumṭumāniyya 3.693
Tunceli 2.607
tune 2.395, 2.397–400
Tunis 1.171, 4.574
Tunis Arabic 1.260, 1.444, 1.446–447, 1.596, 1.609, 1.642, 2.534, 3.12, 3.85, 3.607, 3.737, 3.739, 4.563–571, 4.629
Tunis Arabic, Jewish 1.189
Tūnisī, Bayram at- 1.599, 1.601, 1.603, 4.105
Tunisia 1.24, 1.56, 1.597, 2.105, 2.363, 2.531, 2.698, 2.707–710,

2.713–715, 3.11, 3.53, 3.316, 3.347, 4.563, 4.571–577
Tunisia, Arabic in 1.440, 1.571, 1.584, 1.592, 3.672, 4.550
Tunisian Arabic 1.23, 1.51, 1.187, 1.423–424, 1.457, 1.632, 1.635–636, 2.74, 2.161, 2.318, 2.320–322, 2.374, 2.388, 2.455–458, 2.653, 2.659, 3.11–12, 3.84, 3.226, 3.349, 3.352, 3.585, 3.607, 3.729, 4.95, 4.318, 4.323–324, 4.326, 4.648, 4.671
Tunjur 1.361–362
Tupuri 3.634
Ṭūr ʿAbdīn 3.370, 4.402
Ṭūr, aṭ- 4.237
turāṯ, fuṣḥā at- 1.405
turcos 3.1, 3.4
túrgu 3.634
Turkey 1.176, 1.388, 1.529, 3.312, 4.577–583
Turkey, Arabic in 4.577–583
Turkey, language policy in 4.578, 4.581
Turkey, toponyms in 4.508
Turkic languages 1.33–34, 1.133, 1.259, 2.221, 2.235, 2.405, 2.557, 3.72
Turkish 1.24, 1.60, 1.88–89, 1.95, 1.167, 1.170–171, 1.174–175, 1.189, 1.222, 1.231, 1.233, 1.240–241, 1.260, 1.275, 1.328, 1.388, 1.390, 1.396, 1.404, 1.529, 1.546, 1.554, 1.590, 2.1, 2.9, 2.26, 2.40, 2.200–201, 2.219, 2.221, 2.469–470, 2.491, 2.504, 2.605–607, 2.635, 2.669, 2.673, 3.59, 3.67, 3.350–351, 3.370, 3.555, 3.578–579, 3.640, 4.377, 4.503, 4.583–589, 4.727
Turkish, Anatolian 4.584
Turkish, Arabic loanwords in 4.259–267
Turkish, Diyarbakır 2.605
Turkish, evidential in 1.396
Turkish, Kipchak 4.589
Turkish, *miş*-perfect in 1.396
Turkish, Ottoman 2.560, 2.636, 2.670, 4.259–460, 4.584–586
Turkish, possessive construction in 2.405, 3.67
Turkish, Van 2.605
Turkish, West 3.502
Turkish as language of prayer 4.578
Turkish as official language 3.504
Turkish in Syria 4.402
Turkish in the Middle East 2.699
Turkish terms of address 3.660–661, 4.467, 4.470
Turkish/Arabic dictionary 3.22
Turkmens 2.59, 3.313, 3.351, 4.589
Turkoman 2.414
Turks, Kipchak 2.557, 4.263–264, 4589
Turku 1.521, 3.634, 3.643
Turku Arabic 1.364, 3.61, 3.634
Ṭurōyō 1.88, 1.540, 3.350, 3.370, 4.402
Ṭurōyōs 4.402
ṭurra 1.4
Ṭūsī, Nāṣir ad-Dīn aṭ- 3.502
Ṯuwa Arabic 4.754
Ṯuwāra Arabic 3.401
Tuz Khurmatu Arabic 2.415
twarab 1.449
Tweissi, Adel 2.117–120
Twelver Shiʿites 2.328–329
two-phase verb 1.51–52
two-word stage 2.100
Tyane, Elie 3.349
typology 1.483–484
Tyre 2.600, 3.409, 4.571
Tyre Arabic 1.269

U

ʿUbāda, ʿAbd al-Fattāḥ 2.634–635
ʿUbāda, Muḥammad ʾIbrāhīm 4.22
ʾUbayy ibn Kaʿb 4.5
Üçgüllük 1.115
Üçırmak 1.116
u-coloring environment 2.417
Udovitch, Avraham 3.7
ʿuḏrī 4.202
ʿudūl ʿan al-ʾaṣl 4.30
Uganda 1.518–520, 1.661–662, 2.588–589, 3.642, 4.381
Uganda Muslim University 1.662
Uganda Supreme Islamic Council 1.662
Ugaritic 1.38, 1.132, 1.203, 1.537, 1.643, 2.23, 2.66–67, 3.227, 3.408, 3.410–419, 4.104, 4.171, 4.302–303, 4.313
ʾuḥjiya 4.157
ʾujma 2.85, 2.258, 3.345
ʿUkbarī, ʾAbū l-Baqāʾ ʿAbdallāh ibn al-Ḥusayn al- 2.291, 4.119
Ukrainian, Arabic loanwords in 4.259–267
ʿUla, al- inscription of 1.129, 1.149, 3.467–469, 3.471–472
Ulād Brāhīm Arabic 1.58, 3704
ʿulamāʾ 2.53, 2.630
Ullendorf, Edward 4.296
Ullmann, Manfred 4.33–36, 4.287–288, 1.564, 3.7, 3.30, 3.219, 3.648
Ullmann, Stephen 4.282
ʿulqa 4.102
Ultan, Russell 3.206
ultra-short vowel 1.62
ʿulūj 1.617
Ūlyāytū, Sultan 3.308
ʿUmar ibn ʾAbī Rabīʿa 4.202
ʿUmar ibn Šabba 2.630
ʿUmar II, Caliph 1.361
ʿUmar Tal, al-Ḥājj 4.182
ʿUmar, ʾAḥmad Muxtār 2.638, 3.43
ʿUmar, Caliph 2.47, 2.628, 4.5
ʿUmari, al- 4.11
Umayyad Dynasty 1.402
ʿumda 2.291, 2.430, 3.100
umlaut 1.300, 4.678
Umlautung 1.301, 2.152–153
ʾumm 2.578
ʾUmm al-Jimāl, inscription of 1.149
Umm Kulthum 1.528
Umm ul-Quṭṭayn 2.502
ʾumma 1.629, 2.647, 3.344, 4.76, 4.78
ʾummiyya 3.74
ʾummiyyīn, ʿāmmiyyat al- 1.405
ʿUmrān 4.127
Unaccentable Element Principle 2.113
unaccusative 1.358, 2.367, 3.563, 4.490–491
Unaccusative Hypothesis 3.224
underlying form 1.422
underlying level 4.446–448
underlying representation 2.626, 3.492

underlying structure 2.626
understatement 2.71
unergative 1.358
ungrammaticalness 4.200–201, 4.207–208
Unicode 1.515, 2.380
unidirectionality 2.197
Unidirectionality Principle 2.192, 4.39
unification 1.457–458
unification, feature 1.459
Uniform Three-Level Hypothesis 4.742
Uniformity of Theta Assignment Hypothesis 4.488, 4.491–492
uniformity, morphological 3.706
United Arab Emirates 2.212–213, 2.215, 2.696, 3.63
United Kingdom 2.73
United States, Arabic in the 2.684
Universal Grammar 1.342, 1.484–485, 1.488, 2.96, 2.102, 2.170, 4.143–144, 4.391
universal idea 4.686
universality 3.492
universals 1.484, 2.319, 2.321
universals of acquisition 4.370
universals of linguistic change 2.319
unmarkedness 4.638
Unterbeck, Barbara 2.157, 2.163
Upper Egypt 1.24, 1.589, 2.1
Upper Nile 2.517
Upper Volta 4.279
upward mobility 4.629
ʿuqad 3.160
ʿUqayl 3.14, 3.94
ʿUqba ibn Nāfiʿ 3.53
Uralic languages 2.67, 4.372
ʿurbān 1.129, 1.132
urban area 3.60
urban center 1.595
urban culture 1.594
urban dialect 1.271, 1.407, 1.595–596, 1.604–605, 1.607–610, 1.618, 2.1, 2.8, 2.117, 2.161, 2.230–231, 3.55
urbanization 1.363, 1.593, 1.595, 3.12, 3.58, 4.77, 4.322, 4.633–634
ʿurbī 1.273
Urdu 1.133–135, 1.241–242, 1.257–258, 1.487, 1.662, 2.26, 2.210, 2.215, 2.610, 2.668, 2.673, 3.59, 3.131, 3.313, 3.504, 3.506, 3.510, 3.574, 3.578, 4.434, 4.595–603
Urdu, Christian 4.601
Urfa 1.87, 1.388, 4.579
Urfa Arabic 2.608
ʿurfī 3.119
Uriagereka, Juan 4.523
ʾurjūza 3.212, 4.32–33, 4.36
Urmi 1.540
Urmia 2.607
ʿurūba 3.344, 4.76
ʿUrwa ibn al-Ward 4.208
ʾUsāma bin Lādin 2.71, 4.78–79
user feature 4.472, 4.474–475
ʾUšmūnī, al- 2.189, 4.436
Ussishkin, Adam 2.367, 3.305, 3.494
ʾUsṭāṯ 4.542
ʾuṣūl → *ʾaṣl*
ʾuṣūl al-fiqh 2.176, 2.187, 2.359, 3.122, 4.332, 4.685
ʾuṣūl an-naḥw 1.194, 2.176, 2.187, 2.625
ʾuṣūliyyūn 2.360, 3.321–322, 3.324, 4.332
Utas, Bo 2.407, 2.412, 3.575
ʿUtaybah 3.326
ʿUṯmān ibn ʿAffān, Caliph 2.47, 2.264, 4.4–5
ʿUṯmān ibn ʿAlī 3.341
ʿUṯmān, Ṣabrī 1.599
ʿUṯmānic codex 4.5–7, 4.9–10
Utrecht 2.76–77
utterance 1.474, 2.358, 2.536, 2.542, 2.623
utterance, mean length of 1.340
utterance starter 1.474
uvula 2.365, 3.599
uvular 2.229, 2.231, 2.270, 3.65, 3.598, 3.608–609, 4.1, 4.636, 4.754
uvularization 3.527, 4.173
Uwise, M. Mohamed [Uvais, Ma. Mukammatu] 4.433
Uyghur 1.135, 2.26, 4.603–608

Uyghur, vowel raising in 4.605
ʿUzayzī, Rukus ibn Zaʾid 2.502–503
Uzbek 1.28–29, 1.34, 1.590, 2.26, 2.557, 4.424–425, 4.608–613, 4.616–617, 4.621–622
Uzbek, question particle in 4.622
Uzbek, Tashkent 4.609
Uzbek suffix in Uzbekistan Arabic 4.616–617
Uzbekistan 1.84, 2.557, 4.608–623
Uzbekistan Arabic 1.537, 1.618, 2.414, 2.420, 2.670, 3.84, 3.237, 3.449, 3.544, 3.588, 3.672, 4.198, 4.503, 4.612–613, 4.727
Uzbekistan, Arabic in 4.608–612
Uzunçarşılı, İsmail Hakkı 3.502–505
Uzunoğlu, M. Vecih 4.578

V

Vabkend 4.613
Vagelpohl, Uwe 2.200–201, 4.542–548
Vago, Robert M. 2.683–684
Vaidyanathan, S. 4.433
Vainikki, Anne 4.143
Vaissière, Jacqueline 3.724
valence → valency
Valencia 1.98, 1.102
Valencia Arabic 1.108, 3.286, 4.431
valency 1.182, 1.343, 1.356–357, 3.225, 4.535, 4.624–627
valency, deep structure 1.346
valency expansion 4.626
valency reduction 3.225–226, 4.625–626
Valkhoff, Marius F. 4.291
Vallaro, Michele 4.484
Vallduvi, Enric 3.653, 3.655
valorization 1.483
Van Cleve, J.V. 4.222
Van Coetsem, Frans 4.370
Van Ess, John 3.429
Van Mol, Mark 2.702
Van Rensburg, M.C.J. 4.294
Van Valin, Robert D. 4.501–502, 4.734–735
Vandals 3.52
Vanderveken, Daniel 2.80, 4.328, 4.331
Vanhove, Martine 1.609, 2.270, 3.142, 3.144, 3.237, 4.750–758
Vankuli 3.502
variability 1.439, 2.318, 2.320–322
variability, intraparadigmatic 2.196
variable 1.440
variable, linguistic 1.440–441
variable, proxy 4.633
variable, situational 3.664
variable, speaker 4.628–629
variant readings of the *Qurʾān* → *Qurʾān*, variant readings of
variation 4.627–636
variation, age-related 4.630–631
variation, class-related 4.634–635
variation, cross-linguistic 1.312
variation, diastratic 1.572
variation, education-related 4.632–633
variation, ethnicity-related 4.633–634
variation, free 2.97, 2.318
variation, gender-related 4.631–632
variation, grammatical 3.15
variation, language 1.440, 2.58, 2.628, 2.655, 3.188
variation, lexical 1.56, 1.95, 3.13–21, 3.203
variation, parametric 3.706
variation, register 1.483–484, 634
variation, religion-related 4.633–634
variation, sex-related 4.632
variation, sociostylistic 1.419
variation, speech 1.271
variation, style 1.631
variation, word order 2.114
variety, basic 1.485
variety, colloquial 1.616
variety, discrete 1.635
variety, High 1.636
variety, intermediate 1.634

variety, Low 1.636
variety, matrix 1.634
variety, middle 1.634
Varisco, Daniel M. 2.258–259
varition, micro- 3.233
Varley, Rosemary 2.690
Varma, Marthanda 3.128
Varonis, Marlos 2.117
Varvaro, Alberto 3.142, 4.218
Vasalou, Sophia 2.302–307
Vasco da Gama 3.128
Vasmer, Max 4.259
Vassalli, Michele Antonio 3.141–143, 3.146, 3.153
Vatikiotis, Panayiotis J. 2.498
Vecchio, Silvana 2.362
vehicular language 1.363
vehicular language, Arabic as 1.364, 1.366, 3.60–62
Veinstein, Gilles 3.503
velar 3.598, 4.1
velarization 1.60–61, 1.103, 1.233, 1.390, 1.441, 1.544, 1.548, 2.152, 2.235, 2.459–461, 2.482, 2.503, 2.560, 2.612, 2.621–622, 3.65, 3.190, 3.275, 3.363, 3.481, 3.525, 3.599, 3.608, 3.610, 4.171, 4.173, 4.218, 4.421–424, 4.422, 4.591, 4.636–638
velarization, secondary 1.390, 3.361, 3.363
velarization of demonstratives 3.364
velarization spread 1.245, 4.637
velarized 1.544, 1.547
velarized /r/ 1.103
velic closure 3.597
Vella, abate Giuseppe 3.144
Velten, Carl 1.663
velum 3.598
Vendler, Zeno 1.197, 1.199, 2.367
Venetian 2.457, 3.286
Venezuela 3.2
Venice 1.171, 2.454–455
Vennemann, Theo 3.491–492, 4.558
ventive, Akkadian 4.178
Venzlaff, Helga 4.283
verb 1.183, 1.344, 1.428, 2.90, 2.171, 2.180, 2.184, 2.437, 2.537, 3.546, 4.638–645
verb → *fiʿl*
verb, ‘carrier’ 2.221
verb, ‘inceptive’ 3.253
verb, ‘qualitative’ 3.252
verb, act 3.252
verb, action 1.344–346, 4.643
verb, activity 1.52, 3.352–353
verb, affecting 4.626
verb, agentive 1.345
verb, asseverative 4.361
verb, atelic 1.52
verb, auxiliary 1.331, 1.419, 1.557–558, 1.625, 2.21, 2.77, 2.121, 2.146, 3.221, 3.260, 3.536, 4.306, 4.394, 4.620, 4.740
verb, believe-class 4.538
verb, benefactive 1.344, 346
verb, bi-transitive 1.557
verb, causative 3.563
verb, change-of-state 1.253, 3.225
verb, cognition 1.281, 1.346
verb, cognitive 1.358, 1.556, 1.558
verb, compound 2.606
verb, Coptic 1.499
verb, copular 4.395–396
verb, core 4.735
verb, declarative 1.359
verb, defective 1.209, 1.564–565, 1.616, 2.195, 3.246, 3.258, 3.739, 4.37, 4.645
verb, denominative 1.358, 3.532, 4.643
verb, derived 2.450
verb, ditransitive 3.70, 3.559, 3.591, 4.625
verb, double object 1.284
verb, doubly weak 4.705
verb, dummy 1.248, 2.557, 2.615, 3.576, 4.293, 4.451, 4.599–601, 4.606, 4.610–611, 4.620
verb, duration 3.236
verb, durative 3.543
verb, dynamic 1.42, 1.345, 3.226

verb, effective 3.563
verb, elative 3.105
verb, embedded 1.419
verb, epistemic 4.361
verb, estimative 1.359
verb, exceptional case-marking 3.687
verb, existential 3.114–115, 3.237
verb, experiential 1.344, 1.346
verb, finite 1.350, 1.474
verb, four-place 4.626
verb, full lexical 1.220–221
verb, future 3.97
verb, geminated 1.306, 1.330, 1.385, 1.394, 1.490, 1.498, 1.669, 2.247, 2.514, 2.563–564, 2.576, 2.618, 3.185, 3.282, 3.331, 3.462, 3.480, 3.534, 3.541, 3.691, 4.339–343, 4.619, 4.645, 4.699, 4.707, 4.713
verb, hamzated 4.700
verb, hollow 4.645
verb, III*w* 4.702–705
verb, II*w* 4.701–702, 4.704
verb, imperfect 2.431, 4.733
verb, impersonal 2.315–318, 3.252
verb, inchoative 1.198–199, 3.235–236, 3.252–253
verb, ingestive 1.358
verb, intensive 4.692
verb, intransitive 1.221, 1.281, 1.622–623, 1.625, 2.93–95, 3.224, 4.532
verb, I*w* 4.700–704
verb, labile 1.357
verb, locative 1.344
verb, manipulative 4.362
verb, matrix 4.362
verb, meteorological 1.625
verb, middle 1.281, 1.551, 3.563, 4.490–491, 4.626, 4.642
verb, modal 1.221, 1.266, 3.233–238, 4.365
verb, monotransitive 4.625
verb, negative 1.658
verb, nonaffecting 4.626
verb, non-dynamic 1.219
verb, non-inchoative 3.545
verb, non-real 4.412, 4.414–415
verb, non-stative 2.523–524, 2.595, 3.643
verb, one-place 4.624
verb, onomatopoetic 1.331
verb, passive 2.197, 3.103
verb, perfect 4.733
verb, performative 2.80, 3.570, 3.572, 3.679, 3.681, 4.329, 4.332
verb, phrasal 3.253, 3.630, 4.521
verb, postural 2.195–196
verb, prepositional 3.630
verb, pseudo- 1.185–186, 1.564, 3.739–740, 4.39, 4.695
verb, punctual 3.543
verb, quadriradical 1.109, 1.229, 1.245, 1.395, 1.552, 2.245, 2.566, 2.568, 2.612, 2.619, 3.185, 3.251, 3.491, 3.534, 4.113, 4.286, 4.619, 4.644
verb, quasi- 3.739
verb, real 4.412, 4.414–415
verb, reflexive 1.281
verb, sensory perception 4.625
verb, serial 1.221, 1.491, 3.268, 3.536, 4.195–199
verb, spatial agreement 4.230–231
verb, stative 1.12, 1.42, 1.52, 1.108, 1.219, 1.626, 2.523–524, 2.595, 2.618, 3.247, 3.252, 3.545, 3.643, 4.456, 4.569, 4.643
verb, subjectless 1.626
verb, telic 1.52
verb, three-place 4.625
verb, transitive 1.281, 1.625, 2.93–94, 2.551, 3.102–103, 3.224–225
verb, triradical 1.573–574
verb, tropative 1.359
verb, two-actant 1.623–624
verb, two-phase 1.51–52
verb, two-place 4.625
verb, unaccusative 2.370, 4.490
verb, unipersonal 2.315
verb, weak 1.32, 1.94, 1.109, 1.118, 1.122, 1.209, 1.254, 1.306, 1.330, 1.385, 1.551, 1.564, 1.641, 1.669, 2.180, 2.247, 2.349, 2.563,

2.576, 3.12, 3.282, 3.299, 4.37, 4.41, 4.95, 4.645, 4.699–708
verb, weather 2.315
verb classes, derived 1.576
verb inflection, acquisition of 2.102
verb modifier 1.606–607
verb movement 1.378, 1.479, 1.567, 4.394, 4.399
verb nominalization 2.593
verb of beginning 2.196
verb of belief 4.626
verb of blame and praise 1.564, 2.80, 2.94
verb of causation 4.538
verb of certainty 4.538, 4.626
verb of cognition 4.198, 4.413
verb of color and defect 2.366, 2.566, 4.51, 4.411, 4.643
verb of desire 1.346
verb of emotion 1.346
verb of exclamation 4.412
verb of existence 4.738
verb of incomplete predication 1.220
verb of knowledge 4.626
verb of motion 1.221, 1.265, 1.5541.281, 1.332, 1.345, 3.156, 3.236–237, 3.253, 3.368, 3.545, 4.162–163, 4.198, 4.569
verb of motion, in sign language 4.230
verb of perception 1.19, 1.281, 1.332, 1.346
verb of physical defect 4.411, 4.643
verb of probability 4.538
verb of saying 2.331–332
verb of surprise 1.564, 1.639
verb of transfer 1.346
verb one-actant 1.622–623, 1.625
verb phrase 1.197, 1.217, 1.284, 1.331, 1.348, 1.507, 3.285, 3.432, 4.393, 4.646–652
verb phrase, extended 3.684
verb preposition idiom 4.38
verb raising 2.133, 4.365
verb string 1.669
verbal adjective 1.17–20
verbal aggression 2.363
verbal clause 3.381, 3.432–433
verbal complex 1.217–221
verbal compound, Persian 2.412
verbal event 1.195
verbal Forms 3.299
verbal noun 1.15, 1.265, 1.436, 1.551, 2.246, 2.424, 2.430, 2.433, 2.486, 2.618, 3.18, 3.101, 3.157, 3.164–169, 3.709, 3.255, 3.282, 3.388, 3.432–433, 3.533, 4.529, 4.644, 4.659–665
verbal noun, governance of 4.660
verbal noun, object of 4.529–530
verbal sentence 1.182, 1.348
verbal stems, Hebrew 4.177
verbal stems, Semitic 4.176–177
verbalization of the Sacred 4.74
verbalizing morpheme, Hausa 2.253
verbatim rendition 2.331
verb-framed 1.485
verbhood 1.185
verbless clause 1.349
verbs in Romance 3.157
verbs, borrowing of 2.667
verb-second requirement 4.719
verbum 3.547
verbum sentiendi 1.203
Vergnaud, Jean-Roger 2.396, 3.615–616, 3.618–622, 4.344–346
Verhaar, John W.M. 2.343, 2.669
Verhoeven, Ludo Th. 2.75
Verkaik, Paula 2.683
Verkuylen, H. 2.683
vernacular 1.440, 2.11, 2.696
vernacular, Arabic 1.529
Vernoit, Stephen 4.131
Verona, Jacobus de 1.536
verse, blank 4.91
verse, free 4.91–92
Versteegh, Kees 1.10, 1.85, 1.194, 1.295, 1.309, 1.401, 1.489, 1.494, 1.502, 1.530, 1.544–545, 1.562, 1.575, 1.593–595, 1.614–615, 1.617, 1.619, 1.634, 1.641, 2.80, 2.83, 2.86, 2.88, 2.91, 2.157, 2.161, 2.163, 2.176, 2.179, 2.183, 2.232–237, 2.262, 2.265, 2.267,

2.290, 2.293–298, 2.307–311, 2.340, 2.343–344, 2.354–358, 2.373, 2.387–389, 2.401–403, 2.425–428, 2.434–437, 2.448, 2.479, 2.536–538, 2.544, 2.629–631, 2.642, 2.658–659, 2.666, 2.668, 2.670, 2.672, 2.705, 3.9–12, 3.14, 3.45, 3.59, 3.72, 3.88–90, 3.96, 3.102, 3.106–110, 3.113–114, 3.116, 3.127, 3.160, 3.176–179, 3.237, 3.239, 3.348, 3.377, 3.426, 3.453–454, 3.465, 3.547, 3.549, 3.607, 3.613, 3.639–647, 3.673–674, 3.689, 3.714, 4.1–3, 4.7–9, 4.12–13, 4.41, 4.119–120, 4.147, 4.168, 4.199, 4.219–222, 4.235–237, 4.290–295, 4.329–330, 4.350, 4.369, 4.371–372, 4.374, 4.411, 4.434, 4.436, 4.443, 4.446–449, 4.455, 4.473–474, 4.557–558, 4.621, 4.676, 4.681, 4.684, 4.727
vetitive, Akkadian 4.557
V-features 4.489
Vicente, Angeles 1.569–573, 3.280, 3.287, 3.289, 3.292–293, 3.584–588, 4.70–72
Viciana, Martin de 1.97
vicinitas 3.121
vid 1.196
videocassette 3.197
Vienna 1.172
Viereck, Wolfgang 1.586, 1.589
Vihman, Marilyn M. 2.98
Vijayanagara Kingdom 4.452
Vijver, Ruben van de 3.726
Vikokotoni 1.660
Vincent, Nigel 3.643
Vinnikov, Isaak Natanovič 4.616–617, 4.619, 4.621–622
Vinson, M. Julien 4.433–434
violability 3.492
Violet, Bruno 3.472
Viré, F. 3.6
Vitry, Jacobo de 1.99
vocabulary → lexicon
vocal fold vibration 4.668
vocal folds 3.594, 4.665–668
vocal tract 3.594, 3.596–597, 4.666, 4.669
vocalic melody 1.119, 1.121, 2.346
vocalic representation 3.77
vocalization 1.207, 3.77–79, 3.111
vocalization sign 2.529
vocative 1.106, 1.251, 1.297, 1.309, 1.475, 1.550, 1.563, 2.80, 2.378, 2.585, 2.616, 3.101, 3.486, 4.686
vocative particle 1.227, 1.550, 2.585
vocative statement 4.331
vocomimetic 3.47–48
Vogel, Irene 2.397
voice 1.205–206, 1.216, 3.251
voice, breathy 4.668
voice, creaky 4.667–668
voice, modal 4.667
voice, phonetic 4.665–669
voice, whispery 4.668
voice contrast 2.689
voice contrast, collapse of 2.689
voice disorder 2.675
voice loss 2.675
voice onset time 1.376, 2.689
voiced 3.123, 3.125, 3.594, 4.669
voiceless 3.123, 3.125, 3.594, 4.665–669
voicelessness 3.124
voicing 2.591, 3.124
voicing, intervocalic 4.110
voicing, neutralization of 1.539
Voigt, Rainer M. 1.545–565, 2.262, 3.418, 3.420, 3.593, 4.97, 4.170–179, 4.481, 4.699–708
Vokalumsprung 2.418
Volga Bulgars 4.450
volition 2.595, 4.250, 4.361
volitional 1.220, 1.282, 2.144, 3.253, 3.265
volitive 1.607, 1.611, 3.416
volitive mood 2.22
Volkssprache 1.492, 1.615
Vollers, Karl 1.401, 1.492, 1.560, 1.614, 2.401, 2.459, 3.93, 3.689, 4.558, 4.678
Voltaire 3.43
Volterra, Virginia 1.368
Volubilis 1.296, 3.287

Von Fintel, Kai 3.653
von Wartburg, Walter 3.51
Voorhoeve, P. 1.8
vowel 1.207, 1.235, 1.242, 2.232, 2.237–238, 3.594, 4.669–683
vowel → *ḥaraka*
vowel, anaptyctic 1.234, 1.261, 1.263, 1.664, 2.152, 2.482, 2.484, 2.572–573, 2.577, 4.243
vowel, anterior 4.670
vowel, back 2.235, 3.595
vowel, central 3.595
vowel, epenthetic 2.612, 4.550
vowel, final 4.675, 4.682
vowel, final short 1.383–386
vowel, front 2.235, 3.595
vowel, 'helping' 3.614
vowel, high 4.676, 4.678
vowel, long 2.164, 2.233–234, 2.546, 3.239, 4.683
vowel, loss of short 1.605
vowel, low 4.678
vowel, mid 4.678
vowel, neutralization of 1.261
vowel, pausal 3.454
vowel, performative 1.491
vowel, pharyngealized 4.670
vowel, posterior 4.670
vowel, prothetic 2.612, 3.728–730
vowel, rounded 2.235, 3.595, 4.676
vowel, short 4.670, 4.674
vowel, ultra-short 1.62
vowel, unrounded 2.235, 3.595
vowel, whispered low 4.666
vowel alternation 1.538, 2.346–347, 2.350, 2.353
vowel backing 2.230, 3.610, 4.669–674, 4.678
vowel deletion 2.366, 2.591, 3.181, 3.204, 3.617, 3.621, 3.729
vowel dots 2.599
vowel elision 1.301, 2.483, 2.591, 2.612, 3.266, 3.289, 3.482, 3.611–612, 3.622
vowel fronting 2.311, 3.530, 3.595, 4.678
vowel harmony 1.86, 1.103, 1.135, 1.206, 1.265, 1.286, 1.341, 1.522, 1.526, 1.547, 2.38, 2.223, 2.263, 2.311, 2.520, 2.558, 2.589, 2.670, 3.405, 3.590, 3.615, 3.691–692, 3.697, 4.450, 4.587, 4.604, 4.676
vowel harmony, loss of 4.610
vowel insertion 2.151, 2.484, 3.204, 3.609, 3.611, 3.614
vowel length 1.605, 2.343, 2.346, 2.590, 2.684, 3.643, 3.725
vowel length, loss of 1.539, 2.589
vowel length, neutralization of 3.276
vowel lengthening 1.245, 1.538, 1.548, 1.616, 2.234–235, 2.242, 2.350–351, 2.367, 2.484, 2.612, 3.266, 3.482, 3.530, 3.565, 3.569, 3.724, 4.117, 4.638
vowel lowering 2.38, 2.230, 3.610, 4.565, 4.670
vowel markers 2.183–184
vowel quality 3.596
vowel quantity 3.596, 3.618
vowel raising 2.151–152, 2.311, 2.418, 3.530, 3.595, 4.678–682
vowel raising in Uyghur 4.605
vowel reduction 1.179–180, 1.490, 3.293, 3.566, 3.725, 4.682–683
vowel rounding 2.38
vowel shift 1.538
vowel shortening 1.245, 1.326, 1.608, 1.616, 1.638, 2.233, 2.612, 3.181, 3.240, 3.299, 3.482, 3.530, 3.566, 3.612–613, 3.691–692, 4.116–117, 4.558
vowel sign 1.152, 1.400, 3.453, 3.515
vowel syncope 3.240
vowel system 3.609
vowels, acquisition of 2.98
vowels, Semitic 4.173
VP → verb phrase
VP constituent 4.392–394, 4.396, 4.524, 4.646, 4.648
VP Internal Subject Hypothesis 1.348, 4.364, 4.524, 4.648, 4.743
VP-constituent 4.647
Vrolijk, Arnoud 1.598
Vycichl, Werner 1.496, 1.500, 2.68, 2.217–218, 2.220

W

/w/ 2.164–168
wa- 1.71, 1.467–469, 1.471–472, 1.473–475, 1.499, 1.557, 1.648, 2.22, 2.80, 2.132, 2.224, 2.547, 3.104, 3.161, 3.200, 3.202, 3.646, 3.741, 4.190, 4.192, 4.194, 4.585, 4.685, 4.729
waḍʿ 1.194, 2.80, 2.359, 2.538, 2.541, 2.581, 3.119, 3.176, 4.286, 4.684–687
waḍʿ al-ʿāmm, al- 4.685–686
waḍʿ al-luġa 2.178, 3.35, 3.88
waḍʿ al-xāṣṣ, al- 4.685–686
waḍʿ ḥaqīqī 2.581
waḍʿ majāzī 2.581
waḍʿ nawʿī 4.685–686
waḍʿ šaxṣī 4.685–686
Wadān 3.170, 3.172
Waddai-Darfour 1.363
Waddan 3.53, 3.56
wāḍiʿ 3.121
Wāḍiʿ ad-Dāma 3.401
wāḍiʿ al-luġa 2.179
Wadi al-Sīr 2.498
Wādi ʿAraba 3.400
Wādī Ḍuraʾ 2.260
Wādi Fāṭima 4.127
Wādi Ḥaḍramawt 2.212
Wādi Ḥaḍramawt Arabic 3.478, 4.687–699
Wādi Naṭrūn Arabic 2.2
Wādi Ramm 3.402–404
Wādi Sirḥān 3.401
Wādiy aṭ-Ṭūr 4.240
Wādiy Fēṛān 4.237
Wadud, Amina 2.647
Waer, Mazen al- 1.457
Wafd 3.193
wāfir 3.211, 3.214
Wagner, Daniel A. 1.533, 3.78
Wagner, Ewald 2.52–53, 4.305–307
Wahba, Kassem M. 1.441, 2.8, 2.104–113, 4.669
Wahba, Magdi 3.26
Wahba, Saʿd ad-Dīn 1.601
Wahba, Wafaa Batran 2.131–136, 2.174, 4.437–442, 4.719–720
Wahhabis 1.380, 3.402
Wahhabism 1.287, 1.381
wāḥid 1.189, 1.248, 1.418, 2.614
Wahid, Abdurrahman 2.336–337
Wahiga 1.655
wahm 2.629, 3.176
Wahrān → Oran
Wahrmund, Adolf 3.26
waḥšī 2.88, 4.203
waḥy 4.684
waʿīd 4.329
Waidyar, Moyinkutty 3.131
wa-ʾin 1.465
wajh 1.192, 2.402, 2.404, 3.92
wajh al-kalām 2.629
wajh al-qawl 2.629
wājib 2.299
Wālašmaʿ Dynasty 2.52
Walāta 3.170
Waletzky, Joshua 2.207, 4.728
Walī Allāh, Šāh → Walīullāh, Shāh
Walīd ibn al-Muġīra, al- 2.304

Walīd ibn Yazīd, Caliph al- 4.202, 4.205
Walīullāh, Shāh 2.328, 3.508
Walker, Rachel 3.494
Wallace, Erwin 3.429
Wallace, Stephen 2.207
Wallin, George Augustus 3.402–403
Wallo 2.51, 2.53
Walter, Mary-Ann 1.375–378
Walters, Keith 1.442–448, 1.629–634, 1.636, 2.650–664, 4.323–324, 4.681
Walther, Wiebke 3.718, 3.720–721
Waltisberg, Michael 1.467–470
Waltke, Bruce K. 3.409
Walzer, Richard 4.542
Wandala 1.333
Wang, William 3.604
Wansbrough, John 4.6
waqaʿa 3.96, 3.106, 3.108, 3.122
wa-qad 1.474, 2.224
waqf 1.2, 1.309, 1.614, 2.16, 2.183, 2.233, 2.293, 2.402, 3.564, 3.567, 4.236, 4.426, 4.558
waqf, ʿalāmat al- 3.740
Waqialla, Osman 4.138
Warbah 2.608
waris 2.471
Warner, Levinus 1.167
warrāq 3.307
warrāqī 3.307–308, 4.131, 4.133
Warrāqī, Nārīmān Nāʾlī al- 1.472
Warren, David 1.457
Warschauer, Mark 2.695–696, 3.317
wa-rubba 3.648
waṣafa 4.436
waṣala 3.103, 4.235, 4.412, 4.415
waṣf 2.184, 2.453, 3.422, 4.221
waṣf jāmiʿ 2.580–581
wāṣil 3.103, 4.236–237, 4.412–415
wāṣila 4.235, 4.415
Wāsiṭ 2.572
waṣl 1.152, 1.429, 1.432, 1.471, 2.292, 4.235–236, 4.432, 4.472
waṣla 1.2, 4.518–519
Wasow, Thomas 2.18
Wasserstein, David 1.97
Wasta Arabic 2.5
Watanabe, Akira 3.655
watid 3.173, 3.208–210, 3.212–214, 3.304, 4.101
Watson, Janet C.E. 1.21–25, 2.39, 2.229, 2.231, 2.390, 2.556, 2.562, 3.201, 3.203, 3.305, 3.494–495, 3.608, 3.610–612, 3.620, 3.728–729, 4.62, 4.96, 4.106–115, 4.117, 4.346, 4.559, 4.499–500, 4.528–535, 4.637–638, 4.697, 4.750, 4.756
Watson, Richard L. 1.521, 1.525, 2.518, 2.520, 3.642
Wau 1.519
wave model 1.594
wāw 1.152, 2.233–234, 2.546
wāw al-ḥāl 1.649, 2.132
wāw al-maʿiyya 3.701
wāw an-nadba 3.573
wāw conversive 1.196
wāw rubba 3.646, 4.35
-wayhi 1.310
wazn 2.223, 2.447–448, 3.299, 3.445, 4.86, 4.207, 4.659
weak 2.93
weak consonant 1.204, 1.313, 2.308
Weak Mora Stress Correction 3.617
weak roots 1.653, 4.93
Weak Transfer hypothesis 4.143
weak verb → verb, weak
weakening 2.192
Webelhuth, Gert 4.741, 4.743
Weber, Doris 2.157
Weber, Max 2.59
Webi Shebelle River 4.381
websites, Arabic 1.207
Webster, Roger 3.478
Wedeen, Lisa 1.636
Weeda, Donald 3.299

Weerman, Fred 3.231
Wehr, Hans 1.562, 2.13, 2.224, 2.347, 2.349–351, 2.427, 2.451, 3.26, 3.72, 3.300–301, 3.425, 3.624, 3.689, 3.702, 4.165, 4.699
Weijnen, Anton 1.585–586
Weil, Gotthold 3.208–209, 3.214, 4.12–13
Weil, Gustav 1.626–628
Weinhold, Norbert 1.586, 1.590
Weinreich, Uriel 1.368, 1.618–619, 2.318–319, 2.321, 2.684, 4.160
Weipert, Reinhard 3.36, 3.50, 3.718
Weir, S. 4.106
Weismer, Gary 2.688
Weiss, Bernard G. 2.178, 2.304, 2.306, 2.425, 3.88, 3.321, 3.324, 4.287, 4.684–687
Weiss, Josef 2.236–237
Weitzman, Rayman S. 4.666–467
Welden, Ann 3.726, 4.346, 4.390
Wellens, Ineke 1.521–524, 2.588–597, 3.642, 4.198, 4.370, 4.373
Welles, Elizabeth B. 4.154
well-formedness 1.431, 3.492
Wellisch, Hans 4.515
Wells, John 3.606
Wells, Rulon S. 4.646
Welmers, W.E. 2.269
Welsh 4.194, 4.321
Weltens, Bert 2.681–683
Weninger, Stefan 2.56–57, 3.31, 3.719, 3.721
Wensinck, Arent Jan 2.46, 3.718–719, 4.478
Werbeck, Wolfgang 4.750
Werner, Jürgen 3.402–405
Werner, Roland 3.437–438
Wernicke's Area 2.690
Wernicke-Lichteim Model 2.690
Wernitz, Corinna 2.77
Wertheim, Suzanne 4.450–451
West Africa 2.250–251, 3.135–136
West African languages 1.526, 3.461
West Arabian, Ancient 1.187
West Arabian Arabic 1.490, 1.618, 3.94, 3.693–694, 3.696, 3.698
West Bank 2.465, 2.498–499, 2.502, 3.526
West Nile 2.588
West Semitic languages 1.358, 1.616, 3.191, 4.171, 4.176, 4.178, 4.313
West Sudanic Arabic 1.525–526, 1.618, 3.544, 3.634, 4.376, 4.708–718
Westerlund, David 1.660
Western Algerian Arabic 1.264
Western Arabic 1.37, 1.58, 1.259–269, 1.605, 1.614, 2.8, 2.263, 4.674–675
Western Bedouin Arabic 1.260
Wetter, Andreas 2.51–56
Wettinger, Godfrey 3.142, 3.153
Wetzstein, Janos Gottfried 2.151
Wexler, Paul 2.252, 2.680
WH- element 1.30, 2.19–20, 4.80–81
WH- question 1.350, 2.389, 2.393–394, 2.397, 2.399, 2.574, 3.359, 4.522, 4.524
Whaley, Lindsay J. 4.535
WH-complementizer 2.394
WH-construction 4.524
WH-extraction 2.173, 3.707
WH-interrogative 4.718
WH-island 3.383
WH-Island Constraint 3.384, 4.81
whisper 4.666
White Nile 1.361
White, Lydia 4.144–145
Whiteley, Wilfred H. 1.660
Whitney, William D. 4.516
WH-movement 1.413–414, 1.508, 2.390, 2.393, 3.329, 3.706, 3.709, 4.522–523, 4.526–527, 4.718–722
WH-object 2.172
WH-operator 2.172

Whorf, Benjamin 1.493, 2.642
WH-phrase 4.82
WH-trace 1.348
Wieringa, Edwin 1.5–12
Wierzbicka, Anna 1.42, 3.678
Wiet, Gaston 2.42
Wijst, Per van der 2.683
Wilce, James M. 1.286–289
Wilcox, William 1.602
Wild, Stefan 3.31, 3.36, 3.426, 3.717–721, 4.77, 4.508–515
Willett, Thomas 4.160, 4.163
Williams, Edwin 3.684–685, 4.488, 4.648, 4.719
Williams, M.P. 1.651
Wilmsen, David 1.419, 1.668, 2.1–12, 2.373–378
Wilson, Andrew 1.512
Wilson, Jessica 2.33
Wilson, John 3.681
Winder, R. Bayly 4.19
Windfuhr, Gernot L. 2.220, 2.409
Windows 2.380
Wink, André 2.325
Winkelhane, Gerd 3.502
Winkler, Hans Alexander 1.503, 2.2, 2.218
Winnebago 3.622
Winnett, Fred V. 4.479–481
Winter, Michael 2.464, 4.578
Wirba, Malam Innua 1.337
Wise, Hilary 1.457, 3.388, 3.392, 4.62, 4.64, 4.67
wish 1.249, 1.268, 1.284–285, 1.303
Wissman, Herman von 2.341
Wiswall, Abdul-Qadir 3.63
Witkam, Jan Just 3.340
Wode, Henning 4.140
Woidich, Manfred 1.221, 1.270–271, 1.299–308, 1.320, 1.323–333, 1.503–504, 1.570, 1.572, 1.584, 1.586, 1.590, 1.595, 1.602, 1.604–605, 1.609–611, 2.151, 2.153, 2.388–390, 2.494–495, 3.11, 3.236–237, 3.401–402, 3.404–405, 3.426–427, 3.451, 3.595, 3.653, 3.655, 3.705, 4.70–71, 4.152, 4.196–197, 4.238, 4.285, 4.533–534, 4.550, 4.628, 4.633, 4.681
Wolane 2.67, 4.301
Wolfe, Ronald G. 2.310
Wolfram, Walt 4.321
Wolfson, Harry Austryn 3.113
Wollamo 1.35
Wolof 2.670, 2.673, 3.60, 3.169–170, 3.314, 4.179–182, 4.184, 4.723–725
Wolof, iterative in 4.725
wolofal 4.183
women, position of 2.643
women's speech 1.243, 1.247, 1.272, 2.8, 2.33–34, 4.380, 4.563–564, 4.681
Wood, Sidney 4.669–670
Woodhead, Daniel R. 2.415
Woolard, Kathryn A. 1.629
Worbe, André 4.708
word 2.545, 2.623–624
word, biconsonantal 4.97
word, minimal 4.346, 4.557
word, monosyllabic 4.388
word, obsolete 3.632
word, one-letter 4.557
word, prosodic 2.396
word, pseudo-Arabic 4.583
word accent 1.608
word boundary 2.118, 2.120, 3.611
word classification 3.109
word decoding 3.79
word derivation 4.93
word edge 2.397
word formation 1.452, 2.345, 2.470, 3.297, 4.461
word formation, secondary 1.577
Word Grammar 1.457
word juncture 3.612
word order 1.34, 1.45–47, 1.110, 1.183–185, 1.266, 1.284, 1.331, 1.353–354, 1.396, 1.440, 1.479, 1.486, 1.491, 1.493, 1.553, 1.616, 1.622, 1.635, 1.649, 2.76, 2.100, 2.115, 2.149, 2.221, 2.249, 2.390, 2.430, 2.491, 2.569, 2.596, 3.62, 3.68, 3.81, 3.186, 3.201–202, 3.222, 3.231–232, 3.260, 3.285, 3.536,

3.685, 4.186, 4.312, 4.364–365, 4.379, 4.393, 4.496, 4.498–499, 4.503, 4.506, 4.523, 4.570, 4.657, 4.725–736, 4.746
word order, agent-initial 4.653–657
word order, basic 4.486, 4.654, 4.657, 4.728
word order, inverted 3.83
word order, narrative 1.473
word order, verb-first 3.685
word order, verb-initial 4.653–657
word order in the *Qurʾān* 4.27
word order in Uzbekistan Arabic 4.621
word recognition 1.456, 3.79
word recognition experiment 2.682
word recognizer 1.214
word retrieval 1.341
word size, minimal 3.614
word-finding difficulty 2.690
word-formation 1.575–578
wordplay 3.174, 3.540
word-string, phonological 4.288
Wordsworth, William 4.91
word-tree dominance 3.616
Worrell, William H. 1.496
Worsely, Allan 2.560
Wouden, Ton van der 3.653, 3.655
Wright, Elaine 3.337
Wright, Roger 1.100
Wright, William A. 1.169, 1.471–472, 1.638–639, 2.13–15, 2.23–24, 2.156, 2.169, 2.196, 2.224–226, 2.228, 2.230, 2.294–295, 2.315–316, 2.358–359, 2.367, 2.378, 2.387, 2.389, 2.426, 2.433, 2.452, 2.550, 2.645, 3.166–167, 3.246, 3.255, 3.304, 3.422–424, 3.426, 3.440, 3.450, 3.453, 3.455–456, 3.489, 3.542, 3.558, 3.561, 3.563–564, 3.613, 3.649, 3.699, 3.705, 3.717, 3.719–721, 3.729, 4.18–19, 4.45, 4.54, 4.56, 4.61–62, 4.65, 4.116, 4.118–119, 4.316–317, 4.335, 4.341, 4.349, 4.355, 4.357–359, 4.361, 4.437, 4.455, 4.484, 4.487, 4.504, 4.525, 4.527, 4.531, 4.533–534, 4.558, 4.646, 4.653, 4.660–661, 4.720, 4.747
writing 1.483, 1.487, 1.647, 2.40, 2.111, 3.170
writing, Akkadian 4.1
writing, consonantal 1.487
writing error 4.271
writing of Arabic 2.74
wujūd 2.580
Wulād Abu Digin 1.362
Wulād Abu Īse 1.362
Wulād Abu Xidēr 1.362
Wulād Alwān 1.362
Wulād Atīye 1.361–362
Wulād Bilāl 1.362
Wulād Eli 1.362
Wulād Emir 1.362
Wulād Hēmat 1.361–362
Wulād Himād 1.362
Wulād Humrān 1.362
Wulād Isa 1.362
Wulād Maharēb 1.362
Wulād Rašīd 1.361–363
Wulād Sālem 1.362
Wulād Salmān 1.362
Wulād Sarrār 1.362
Wulād Slimān 1.362–363
Wulād Tālib 1.362
Wulād Xānem 1.362
Wulfeck, Beverly 2.690
wuṣla 4.102
wusṭā 1.567
Wzzino, Dun Franzisk 3.146

X

xabar 1.49–50, 1.202, 1.428, 1.436, 1.488, 1.557–558, 2.132, 2.227, 2.290, 2.354, 2.358–360, 2.432, 2.434–435, 2.437, 2.537–539, 2.551, 2.627, 3.113, 3.115, 3.177, 3.231, 3.234–235, 3.353, 3.381, 3.683, 3.687, 4.330, 4.332–4.333, 4.354, 4.484, 4.495, 4.525, 4.653, 4.737–740
xabar al-maʿrifa 2.227
xabar ʾinna 1.69, 2.355, 4.740
xabar kāna 1.70, 1.557, 2.431, 2.549, 2.551, 3.105, 3.354, 4.740
xabar li-l-maʿrūf 2.227
xabar manṣūb 2.227, 4.738
xabar sadda masadd al-fāʿil 4.739
xabar ṯānin 4.737
xabariy(ya) 2.358
xabīṯ 1.192
Xafājī, Ibn Sinān al- 1.173, 1.428, 4.473
xafḍ 1.67, 1.559–560, 2.232, 2.402
xafīf 2.233, 3.213–214, 3.308, 3.340, 3.551, 4.431–4.432, 4.442
Xalaf al-ʾAḥmar 2.178, 2.290, 2.628, 4.204–207, 4.220, 4.283, 4.431, 4.446
xalīfa 2.425, 2.433, 4.283
Xalīl ibn ʾAḥmad al-Farāhīdī, al- 1.67, 1.152, 1.400, 1.402, 1.453, 2.85, 2.184, 2.227–228, 2.232, 2.238, 2.295, 2.299, 2.309, 2.402, 2.436, 3.21, 3.31–32, 3.36–38, 3.73, 89, 3.107, 3.123, 3.207–210, 3.213–215, 3.602, 3.648, 4.12, 4.89, 4.201, 4.206
xalq al-Qurʾān → *Qurʾān*, creation of the
xalṭ 3.118
Xamdān 1.378
Xānfū 1.378
xarajāt 1.100
xaras 2.677
xārij 2.360
xarja 1.100
Xarrāt, ʾAdwār al- → Kharrat, Edouard
xašana-group 2.220
xāṣṣ 4.685–4.686
Xāṣṣ, al- 4.104
xaṭaʾ 2.629, 2.632, 3.648
xaṭaba 4.472
Xaṭābī, Muḥammad → Khatabi, Mohammed
Xaṯʿam 3.693
Xaṯʿam Arabic 3.94
Xaṭīb, ʿAdnān 2.636, 2.640
Xaṭīb, Muḥibb ad-Dīn al- 2.259
xātim 4.181
Xāṭir, Maḥmūd Rušdī → Khater, Mahmoud Rushdi
xaṭṭ 3.160, 3.307, 4.130–4.131
xaṭṭ al-ʾijāza 4.562
xaṭṭ ʾanasī 1.316
xaṭṭ fārisī 3.336
Xaṭṭābī, ʾAbū Sulaymān Ḥamd Muḥammad al- 2.303, 3.324
xaṭṭ-i nāxun 4.134
xaṭṭ-i šikasta 3.337
xaṭṭ-i tarassul 3.338

Xaṭuniyye Arabic 4.407
Xawalme 1.362
xawāṣṣ 2.630
Xaybar 1.129, 4.125
Xayqānī, ʾAḥmad ibn ʾAbī l-Qāsim al- 2.600
Xayrī, Badīʿ 1.600–601
Xazʿal, Šayx 2.572
xazala 2.16
Xazrajī, ʾAbū Dulaf al- 2.469
X-bar schema 2.174
X-bar Syntax 4.741–749
X-bar theory 3.229–230, 4.438, 4.441, 4.649
Xereibah, al- 3.468
Xhosa 4.386
Xi'an 1.378–379
xiao'erjin 1.381
xiaojing 1.381
xiffa 1.122, 2.425, 3.108, 4.431
xilāf 3.107
Xinjiang 1.382, 4.603
Xirbat al-Muxayyit 3.472
xiṭāb 2.183, 2.360
Xošḥālābād 1.28
Xrakovskij, V.S. 3.557
X-slot 4.734
Xuḍarī, al- 3.102
Xuḍayr, Muḥammad ʾAḥmad 3.101–102, 3.104–105
Xūlī, Muḥammad ʿAlī al- 4.423
Xūri, al- 4.460
Xuri, Rašīd Sālim al- → Khouri, Rasid Salim
Xušanī 1.98
Xušmān 2.2
xuṭba 2.42, 2.629, 4.74, 4.78
xuṭūṭ → *xaṭṭ*
xuṭūṭ al-kuttāb 4.132
xuṭūṭ al-maṣāḥif 4.132
xuṭūṭ al-warrāqīn 4.132
Xūz 2.571
Xuzistān 2.572
Xuzistan → Khuzestan
Xwārizmī [Xwārazmī], ʾAbū ʿAbdallāh Muḥammad ibn ʾAḥmad al- 2.28, 2.188, 2.234, 2.427, 2.455, 3.6–7, 3.337, 4.219, 4.260

Y

/y/ 2.164–168
yā 1.152, 2.80, 2.233–234, 2.311, 2.378, 2.546, 3.101, 3.659, 4.686
yāʾ al-ʾiḍāfa 2.294, 4.6,
yābis 1.152, 3.308
Yacine, Kateb 2.703
Yacoub, Adil Y. 3.203
Yadin, Yigael 3.466
Yāfiʿ Arabic 4.750–751, 4.753–757
Yagi, Viviane A. 2.560
Yağmur, Kutlay 4.579
yāhūdiyya 2.259
Yaḥyā ibn ʿAdī 4.542
Yalin 2.462
Yaltkaya, Şerefettin 3.501, 3.503
Yamāni, ʾAḥmad al- 3.508
Yangzhou 1.378–379
Yaounde 1.336–337
yaqīn 1.71
Yaʿqūb 3.337
Yāqūt al-Mustaʿṣimī al-Hamawī 1.152, 3.308–309, 3.340–341, 4.131, 4.133, 4.561
Yarur, Juan 3.2
Yasarī, Muḥammad Asʿad al- 3.338
Yaseen, Mustafa 1.458
Yassin, M. Aziz F. 2.586, 2.616, 4.469
Yathrib 1.128
Yavrumyan, Marat 2.236–239
Yavuz, Kemal 3.501
Yawis 4.123
Yaycıoğlu, Ali 4.98
Yayladağ 1.111, 1.114, 1.117
Yaylıca 1.114
Yazbak, ʾAnṭūn 1.600–601
Yazīdī, ʾAbū ʿAbdallāh al- 3.34
Yāzijī, ʾIbrāhīm al- 1.175, 1.435, 2.629, 3.346
Yaziji, Jorge Suleiman 3.4
Yāzijī, Nāṣif al- 2.189, 2.401, 3.71
Yazıȷıoğlı, Meḥmed → Meḥmed Yazıȷıoğlı
Yefren 3.56
Yefsah, Abdelkader 3.347
Yem 1.35
Yemen 1.190, 1.260, 1.361, 1.451, 1.491, 1.655, 2.105, 2.212, 2.257–258, 2.268, 2.531, 3.63, 4.687, 4.750–758
Yemen, Arabic in 1.260, 1.502, 1.590, 1.606, 1.624–625, 2.204, 2.272, 2.494, 3.429, 3.585, 3.672–673, 3.693, 3.695–697, 3.702, 4.2
Yemen, reunification of 4.751
Yemeni Arabic 1.24, 1.103–104, 1.129–130, 1.187, 1.314, 1.504, 1.617, 1.654, 1.654, 2.257–258, 2.260, 2.263, 2.270, 2.388, 2.390, 2.462, 2.556, 3.63, 3.237, 3.413, 3.416, 3.637, 3.691, 3.694, 3.726, 3.729, 4.71, 4.296, 4.298, 4.303–305, 4.309–310, 4.312, 4.314, 4.373, 4.750–758
Yeou, Mohamed 3.727, 4.669
yes/no question 2.115, 2.133–134, 2.388–389, 2.393–394, 2.397, 2.399
Yesari Esad Efendi 3.338
Yezidis 2.414–415, 2.605
Yīfāt 2.52–53
Yip, Moira 3.462, 4.339
Yobe 3.373

Yoda, Sumikazu 2.534, 4.548
yod-occlusivization 1.539
Yokwe, Eluzai M. 2.518
Yola 1.337
Yoruba 1.545, 3.373–375, 3.725, 4.195, 4.282, 4.758–763
Yoruba, Literary 4.759
Yoruba Bible 4.760
Yorubaland 3.374–375
Yoshida, Shohei 3.207, 4.676
Younes, Munther A. 1.484, 2.230, 3.532, 3.728–730, 4.347, 4.637, 4.669–4.670, 4.672, 4.676
Young Turks 3.344, 3.504
Young-Scholten, Martha 4.143
Youssef, Amr 1.210
Youssef, F. 3.194
Youssef, Zafer 3.542
Youssi, Abderrahim 2.136, 2.470, 3.237, 3.274, 3.287, 4.156–160
youth speech 4.763–767
Yrttiaho, Kaarlo 3.402, 3.404–405
Yuan Dynasty 1.378
Yucatán 3.395
Yuḥannā ibn al-Biṭrīq 4.542
Yule, George 4.727
Yunnan 1.379–380
Yūnus ibn Ḥabīb 3.543, 3.689, 4.447
Yupik, Siberian 2.671
Yūsuf, ʿAlī 3.193
Yūsuf, Ḥājjī 4.277
Yusuf, Harun-Al-Rashid 1.458
Yūsuf, Ḥusnī 1.599
Yusuf, Ibrahim A.G. 2.202
Yusuf, S.M. 2.325, 2.671

Z

Zā 4.279
Zā Kosoï 4.279
Zabad, inscription of 1.149
Zabada 1.362
zabān-i ištiqāqī 2.411
zabān-i tarkībī 2.411
Zabīd 1.655, 2.335
Zabīd Arabic 3.675
Zabīdī, Murtaḍā az- 2.427, 3.21–22, 3.30, 3.40
Zaborski, Andrzej 1.35–40, 1.187–188, 1.313–316, 1.584, 2.23, 2.161, 2.494–496, 3.192, 3.420, 3.585, 3.592, 4.70, 4.96, 4.272–275, 4.314, 4.642
Zacharski, Ron 3.86
Ẓafār 2.256
Zafrani, Haim 2.64
Zaghawa 1.360, 4.376
Zagros Mountains 2.571
Zahid, A. 3.726
Zahidi, Abbas 4.592
Ẓāhirites 2.83, 2.310
Zaiback, Abdelnour 1.668, 1.670
zāʾid 1.191, 2.238–239, 2.307, 2.316, 2.449, 2.546–548, 2.551, 2.627, 4.235
zāʾida 1.72, 2.316
Zaire 1.662, 4.381
Zajączkowski, Ananiasz 3.501
zajal 1.102, 1.601, 1.603, 3.219, 4.217
Zajjāj, ʾAbū ʾIsḥāq ʾIbrāhīm ibn as-Sarī az- 2.295, 2.302, 2.426, 4.9
Zajjājī, ʾAbū l-Qāsim ʿAbd ar-Raḥmān ibn ʾIsḥāq az- 1.124, 1.203, 2.90, 2.92–94, 2.176, 2.179, 2.184, 2.187, 2.234, 2.280, 2.291, 2.295, 2.310, 2.355–356, 2.401–404, 2.425–427, 2.432, 2.436, 2.453, 2.537–538, 2.551–552, 2.625, 2.630, 3.72, 3.102–103, 3.105, 3.108, 3.176, 3.543, 3.548, 3.550, 4.219, 4.221, 4.283, 4.443, 4.455, 4.740
Zakaria, A.A. 3.726
Zakariya, Mohamed 4.138
Zakariya, Mohamed U. 4.98
Zakariyyā (Saʿīd), Naffūsa 1.601–602
Zaki, Abdel-Malek 1.570
Zalami, Fatima Badry 2.96–104
zamān 2.180, 2.424, 2.426–427, 2.549, 3.96–97, 3.108, 3.176
zamān muʿayyan 3.549
zamān muḥaṣṣal 3.549
Zaman, Muhammad Qasim 2.326
Zamaxšarī, ʾAbū l-Qāsim Maḥmūd ibn ʿUmar az- 1.3, 1.72–73, 1.471, 1.558, 1.567, 2.79, 2.93–95, 2.181, 2.226, 2.306, 2.311, 2.354–355, 2.379, 2.387, 2.403–404, 2.427, 2.433–434, 2.478–479, 2.538, 2.549–551, 2.555, 2.580, 2.625–626, 3.22, 3.32, 3.108–109, 3.214, 3.455, 3.542, 3.728–729, 4.2, 4.36, 4.119, 4.121, 4.282–4.283, 4.285, 4.330, 4.455
Zamlakānī, ʿAbd al-Waḥīd ibn ʿAbd al-Karīm Ibn az- 2.445
Zammit, Martin R. 3.144, 4.295–297, 4.313
zandaqa 2.303
Zande 4.375

Zanjānī, ʾAbū l-Qāsim Saʿd ibn ʿAlī az- 1.168, 1.545, 4.119, 4.640
Zanten, Ellen van 3.66
Zanuttini, Rafaella 3.655
Zanzibar 1.448, 1.451, 1.660–662, 4.381–4.382
Zanzibar Town 1.660
Zanzibar University 1.661
Zappa, Francesco 3.139
Zaragoza 1.98
zärg 3.174
Zaria 1.337
Zarīn-qalam 3.338
Zarkašī, az- 2.304, 2.359, 4.23, 4.333
Zarqa, al- 2.498
Zarrinkūb, Abd al-Ḥusain 2.406
Zarzis Arabic 3.226
Zavadovskij, Jurij N. 2.241, 4.304, 4.306
Zawawi, Sharifa 1.660, 1.663
Zawaydeh, Bushra Adnan 2.229–231, 2.395–398, 3.305, 3.494, 3.726, 4.637, 4.670, 4.672
Zāwī, az- 3.42
Zawia Arabic 3.55
Zawila 2.552
zāwiya 4.182
Zaxue 1.381
Zayd ibn ʿAlī 3.34
Zayd ibn Ṯābit 4.4
Zaydite Shiʿites 2.303
Zaylaʿī, ʿAbd ar-Raḥmān az- 4.277
Zaylaʿī, ʿAbdallāh ibn Yūsuf az- 4.277
Zaylaʿī, ʿUṯmān ibn ʿAlī az- 4.277
Zaytūn 1.378
Zaza 2.604, 3.351
Zebed 3.470
Zeenni, Umar ez- → Ziʿinni, ʿUmar az-
Zeila 1.654
Zeine, Zeine N. 3.344
Zeltner, Jean-Claude 1.361, 4.713, 4.715
Zemánek, Petr 1.85–86, 1.204–206, 1.315, 1.458, 1.516, 1.653–654, 4.93–98
Zenaga 2.240–241, 2.249, 3.169–171
Zeraoui, Zidane 3.395
zero 3.447
zero copula 1.13
zero morpheme 2.82
Zeroual, Chakir 1.26–28, 3.524–526, 4.665–669
Zeroual, Liemin 3.350
Zeshan, Ulrike 4.222–235
zeugma 3.538, 3.733
Zewi, Tamar 1.637–640, 2.22–25, 3.416
Zheng He 3.128
Zia ul Haq, President 3.511
Ziadeh, Farhat J. 4.19
Ziftāwī, az- 4.132
ziḥāfa 3.210
Ziʿinni, ʿUmar az- [Zeenni, Umar ez] 1.603, 3.348
zimāmī 3.112
Zimmermann, Fritz W. 3.547, 4.220, 4.543, 4.546–4.547
Zimmermann, Gerit 3.544, 4.612–623
zināʾ 2.362, 4.104
Zīrid Dynasty 4.572
Ziyād ibn ʾAbīhi 1.152
ziyāda 3.119, 3.649, 4.121, 4.640
ziyādāt 2.238
Zizi, Khadija 1.651
Zniber, Mohamed 3.237
Zoroastrian literature 3.574
Zoroastrians 3.581
Zouaghi, Anis 1.458
Zrigat Arabic 3.290
Zrigui, Mounir 1.458
Zubayd 3.693
Zubaydī, ʿAbd al-Karīm az- 4.295
Zubaydī, ʾAbū Bakr Muḥammad ibn al-Ḥasan az- 2.86, 2.88, 2.401, 2.628–630, 3.690
Zubayr, Muḥammad az- 3.718
Zubizarreta, María Luisa 3.388, 4.657
Ẓufār → Dhofar

Zughoul, Muhammad Raji [Muḥammad Rājī Zuġūl] 2.501, 2.696
Zuhayr ibn ʾAbi Sulma 2.389, 3.211, 4.200
Zuhdī, ʾIsmāʿīl 4.561
Zûla 1.654
zulf 4.562
Zulf-i arūs 4.134
Zürcher, Eric J. 4.577, 4.578
Zuyūd 1.362
Zwara 3.56
Zwartjes, Otto 1.96–101, 1.620
Zwarts, Frans 3.653
Zway 2.67, 4.301
zwaya 3.170, 3.174
Zwettler, Henri 3.689
Zwettler, Michael 1.615–617, 2.176, 2.373, 3.468–469, 3.499, 3.645–646, 3.689, 4.9
Zwicky, Arnold M. 3.298, 4.193, 4.646